



HAL
open science

Data informativity for the prediction error identification of MIMO systems: identification of a MEMS gyroscope

Kévin Colin

► **To cite this version:**

Kévin Colin. Data informativity for the prediction error identification of MIMO systems: identification of a MEMS gyroscope. Other. Université de Lyon, 2020. English. NNT : 2020LYSEC018 . tel-03114994

HAL Id: tel-03114994

<https://theses.hal.science/tel-03114994>

Submitted on 19 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ÉCOLE
CENTRALE LYON

N° d'ordre NNT : 2020LYSEC18

Thèse de Doctorat de L'UNIVERSITÉ DE LYON
opérée au sein de l'ÉCOLE CENTRALE DE LYON

Ecole Doctorale N° 160 Electronique, Electrotechnique et Automatique (EEA)

Spécialité de doctorat : Automatique

préparée au Laboratoire Ampère, UMR CNRS 5005

**Data Informativity for the Prediction Error
Identification of MIMO Systems.
Identification of a MEMS gyroscope.**

par Kévin COLIN

Soutenue publiquement le 17/09/2020 devant le jury composé de :

Présidente	Mme Marion GILSON	Professeure des Universités, CRAN
Rapporteur	M. Johan SCHOUKENS	Professeur, Vrije Universiteit Brussel
Rapporteur	M. Guillaume MERCÈRE	Maître de Conférences (HDR), LIAS
Examineur	M. Håkan HJALMARSSON	Professeur, KTH Royal Institute of Technology
Directeur de thèse	M. Xavier BOMBOIS	Directeur de Recherche CNRS, Ampère
Co-Directeur de thèse	M. Laurent BAKO	Maître de Conférences (HDR), Ampère

Remerciements

Pendant ces trois ans de thèse, beaucoup de personnes ont contribué sur les plans personnel et professionnel. Je tiens ici à les nommer et les remercier.

Pour commencer, je veux remercier Bpifrance pour le (grand) soutien financier dans le cadre du projet PSPC Next4MEMS sans quoi cette thèse n'aurait pas eu lieu. J'ai pu aussi participer à plusieurs conférences, séminaires et autre grâce à ce financement et je ne peux qu'exprimer ma gratitude à Bpifrance pour ces opportunités.

Je tiens ensuite à remercier les membres de mon jury de thèse pour leur investissement dans l'évaluation de cette thèse et pour m'avoir décerné le grade de Docteur de l'Université de Lyon. Plus particulièrement, je remercie les rapporteurs Johan SCHOUKENS et Guillaume MERCÈRE qui ont évalué ce présent mémoire de thèse dans des conditions assez difficiles en raison de la pandémie du COVID-19 et je salue ici leur courage. En plus de Johan SCHOUKENS et Guillaume MERCÈRE, je remercie également la présidente de mon jury Marion GILSON et mon examinateur Håkan HJALMARSSON pour leurs questions pertinentes lors de ma soutenance de thèse.

Pour mon encadrement, je tiens à remercier chaleureusement Xavier BOMBOIS, Laurent BAKO et Anton KORNIENKO. Plus particulièrement, je tiens à remercier Xavier pour m'avoir fait confiance en me laissant libre dans mes travaux de recherche et pour m'avoir guidé dans mon avenir professionnel. Par exemple, je le remercie de m'avoir mis en contact avec Håkan Hjalmarsson pour l'offre de post-doctorat à KTH et pour m'avoir donné la possibilité de valoriser mes travaux de recherche devant la communauté européenne sur l'identification des systèmes (ERNSI) et ceci deux fois ! Je remercie Anton pour m'avoir encadré sur la partie pratique pour l'identification MEMS et m'avoir fait goûter des alcools ukrainiens pendant ses pots. Jamais je n'aurais pensé dire un jour que j'étais ivre un mardi soir au travail en face de mes supérieurs hiérarchiques ! J'exprime aussi ma gratitude envers Laurent pour avoir accepté de rejoindre l'encadrement de cette thèse en cours de route et pour m'avoir donné des pistes et des points de vue très intéressants pendant ma thèse. Je remercie mes encadrants d'avoir pu continuer à assurer ma supervision (correction du mémoire entre autre) sans problème lors du confinement en raison de la pandémie du COVID-19. Je remercie aussi Christophe LE BLANC pour m'avoir assisté au début de la thèse sur les aspects techniques de la carte électronique embarquant le MEMS afin que je puisse faire l'identification du gyroscope MEMS.

L'envie de poursuivre en thèse je la dois à Gérard SCORLETTI et Anton KORNIENKO qui m'ont d'abord convaincu de faire un Master recherche à la fin de mon stage de deuxième année au laboratoire Ampère puis de continuer en thèse à la fin de mon stage de fin d'études. Je les en remercie pour cette suggestion. C'est aussi grâce à leur encadrement et leur accueil ainsi que celui de Xavier pendant mon stage de fin d'études.

Je remercie aussi conjointement Anton, Gérard et Xavier pour accepté ma candidature pour cette thèse.

Pendant ce voyage dans la recherche, j'ai eu le pur bonheur de faire partie de la grande et heureuse famille Next4MEMS composée de 3 autres doctorants (Fabricio SAGGIN, Federico MORELLI et Jorge AYALA) et de 3 stagiaires (Eva MARKIEWICZ, Cécile PERNIN et Anthony EL HAJJ). Les 3 années en leur compagnie ont été un véritable bonheur et je ne suis pas prêt d'oublier les moments que j'ai partagés avec eux. Je remercie donc Fabrício pour avoir su apprécier la qualité du modèle que je lui ai fourni pour la synthèse de correcteur du gyroscope MEMS. Je n'oublierai pas non plus cette nuit d'hôtel à Grenoble dans une chambre appelée "Nuit Câline", ces superbes caïpirinhas et cette fameuse blague avec Hugo. Je remercie Federico, mon homologue en identification des systèmes (mais en plus grand), qui m'a appris à faire une pâte à pizza de qualité (je n'ai toujours pas réussi à la faire mais je ne perds pas espoir), qui m'a assisté pendant ma thèse et qui a tenté de défendre avec moi le caractère scientifique de l'identification des systèmes face aux critiques d'autres collègues. En parlant de ce sujet, je remercie donc Jorge pour sa "sagesse" qu'il a exprimé tout au long de ma thèse sur mes travaux de recherche. J'aimerais citer ce grand sage : "L'identification est une science exacte avec des incertitudes !". Ses paroles résonnent encore dans mon cœur. Bien que j'admets tout de même que j'ai forcé Jorge à dire cela, j'ai bien vu dans ses yeux¹ l'émerveillement que constituait une telle révélation pour lui ! Je remercie Maman Provence a.k.a. Thatcher a.k.a. Eva a.k.a. Madame "Tu peux mettre de l'eau à bouillir pour moi aussi s'il te plaît?" pour m'avoir donné le courage de faire le Silverstar et le Blue Fire d'EuropaPark² mais aussi pour la bonne humeur qu'elle a apportée à l'équipe et le partage de thé que nous avons pratiqué pendant ma thèse ! C'est en cela que j'admets la supériorité alsacienne face à la Lorraine. Je remercie Anthony et Cécile pour avoir participé aussi à nos moments détente Next4MEMS, rajoutant encore plus de bonne humeur dans l'équipe. Je me souviendrai toujours de la soirée où nous avons tous dansés la Macarena, YMCA (d'ailleurs Federico, en parlant de ça, ...), etc. Si mes encadrants lisent cette partie, je précise que c'était bien évidemment pour travailler sur le gyroscope MEMS et rien d'autre. Je remercie bien sûr la femme de Fabrício, Priscilla, pour m'avoir fait prendre conscience (de façon calme et reposée) qu'utiliser un sachet plastique pour faire des glaçons, c'est mal pour l'environnement. Je remercie aussi la copine de Jorge, Anaïs, pour le soutien et la bonne humeur toujours présente ! Comme dit précédemment, je pars avec plein de souvenirs. Car être Next4MEMS, c'est foirer un escape game organisé par l'école, mais le foirer ensemble ! C'est faire des playlists Deezer lors du confinement, c'est faire des séances "débat politique" le vendredi après midi. Mais c'est aussi trafiquer le PC des autres pour faire des bêtises comme mettre un second bureau avec la musique It's Raining Men en boucle pendant 10 heures (et non It's Raining in Maine comme il l'a été écrit dans la barre de recherche YouTube). J'espère que cette liste non exhaustive s'allongera dans l'avenir !

¹Lors d'une des rares occasions où ses yeux étaient ouverts bien sûr.

²La version officielle parle plutôt d'un forçage pour monter dans ces attractions à sensations horribles.

Je remercie aussi les autres doctorants que j'ai côtoyés pendant ces 3 ans de thèse + 2 ans auparavant en tant que stagiaire. Sans trop rentrer dans les détails, je remercie donc Arthur, seul doctorant avec qui je pouvais discuter sans lever la tête, Alexandre avec qui j'ai partagé le "bocal à stagiaires" du laboratoire en 3A, Ouassim dont j'attends toujours l'invitation au mariage, Hussein seul doctorant que j'ai cotoyé qui a eu le courage de tenter Ma thèse en 180 secondes (formidable talent d'acteur au passage !), François dont sa tasse M&Ms l'attend toujours dans l'armoire du bureau Next4MEMS et Peng avec qui j'ai collaboré depuis ma 2A. Je remercie aussi (en espérant que j'ai oublié personne) Yanis, Grégory, Qui-lin, Fatma, Ayyoub, Abdelghani, Jérémy, Quentin, Marion et Benoît. Je remercie les doctorants de l'ED EEA en général pour m'avoir élu représentant des doctorants afin que je puisse agir pour le bien commun des doctorants³. Je sais que certains doctorants s'attendaient à ce que je lutte pour avoir des hamacs dans les bureaux mais je n'ai pas eu le courage de faire la demande auprès du directeur de l'école doctorale a.k.a. Gérard. Pour me faire pardonner, voici une photo d'un hamac:



Figure 1: Une photo d'un hamac. Source: <https://www.tropical-hamac.com/>

Je remercie aussi les autres représentants des doctorants de l'école doctorale EEA : Nelly, Charly, Suzanne et Carlos avec qui j'ai pu participer à des projets intéressants pour réunir les doctorants de l'école doctorale.

Je me dois aussi de remercier les permanents du laboratoire Eric BLANCO, Julien HUILLERY, Eric VAGNON, Edith BERGEROUX (qui a du gérer ma procrastination assez coriace, notamment sur les remboursements d'abonnement TCL), Giacomo CASADEI, Laure FRANQUEVILLE, Marie-Christine HAVGOUDOUKIAN, François BURET, Arnaud BRÉARD, Christian VOLLAIRE (qui m'a certainement traumatisé à vie avec la blague du scotch et du hamster) mais aussi Richard et Alice (qui a eu le courage chaque jour de voir mon bureau dans un état approximatif). Je remercie Emmanuel BOUTLEUX et Catherine MUSY pour m'avoir donné l'opportunité de participer à l'enseignement du cours d'Automatique Linéaire pendant 3 ans. Même si au début il était bizarre pour

³Mon honnêteté dirait plutôt que c'était pour pouvoir valider 10h de formations obligatoires !

moi d'enseigner à des élèves pas plus âgés que moi, cette expérience a été pour moi un vrai plaisir confortant mon envie de continuer dans le monde académique.

Enfin, je remercie aussi ma famille pour leur soutien : ma mère Karine, ma sœur Mélissa, mon frère Jérémy et mon beau-père Serge. Je remercie aussi mes colocataires Théo et Anne et mes proches amis Samuel et Diogo avec qui le confinement s'est bien passé avec nos sessions "brainstorming" (i.e., apéros) Skype du samedi soir. Enfin, après avoir gardé le meilleur pour la fin, je remercie mon copain Aurélien pour m'avoir fortement soutenu pendant les deux dernières années de ma thèse et pour avoir apporté de la bonne humeur dans les moments où j'en avais besoin !

Abstract of the thesis

UNIVERSITÉ DE LYON

Ecole Centrale de Lyon - Laboratoire Ampère

Data Informativity for the Prediction Error Identification of MIMO Systems. Identification of a MEMS Gyroscope.

by Kévin Colin

Mathematical models have a crucial place in every engineering field. Models can be used for several purposes such as the design of a controller, the prediction, the health monitoring of a system, etc. Often, the performances of the controller, the accuracy of the prediction and the ability to detect faults occurring in a system strongly depend on the quality of the model. System Identification is the scientific field consisting in the modeling of a system with experimental data collected on this system. Therefore, it gives a more realistic modeling of the system since it can allow to identify some dynamics that are not modeled by first principles. This type of modeling relies on three ingredients: the model structure (set of candidate models), the experimental data and the identification criterion. These three ingredients must be carefully chosen in order to guarantee a good quality for the model.

In order to obtain an accurate identified model, it is essential to guarantee the consistency of the estimate (identified model). This property directly depends on the three aforementioned ingredients. It guarantees the convergence of the identified model to the most accurate candidate model available in the model set when the number of data increases endlessly. When considering the classical Prediction Error estimator (with a least-square criterion), it has been shown that the consistency is achieved when two properties are verified: the identifiability of the model structure and the informativity of the data with respect to the model structure. For the identifiability, a lot of works have been done to derive linear parametrized model structures having strong identifiability properties such as, e.g., Box-Jenkins, Output-Error, Finite Impulse Response, ARX and ARMAX model structures.

For the data informativity, a lot of works have been done for the identification of the single-input single-output systems within the Prediction Error framework. Necessary and sufficient conditions have been derived to verify the data informativity with respect to the classical single-input single-output linear model structures (Box-Jenkins, Output-Error, Finite Impulse Response, ARX and ARMAX model structures) for the open-loop and the direct closed-loop identification. However, for multiple-inputs multiple-outputs systems, the only results for the data informativity available in the literature are only

sufficient and can be very restrictive with respect to the experiment design limitations of the to-be-identified system. Therefore, in this dissertation, we consider the study of data informativity for linear multiple-inputs and multiple-outputs systems in open-loop and closed-loop and we develop less conservative conditions for the verification of this property within the Prediction Error framework.

The second limitation of the data informativity studies available in the literature is the lack of results for the identification of nonlinear systems which can be accurately identified within the Prediction Error framework. This limitation is moreover very restricting in practice since a majority of real-life systems have nonlinear behaviors. Consequently, in this dissertation, we address the problem of the data informativity for a particular type of nonlinear model structures with an input monomial nonlinearity. This particular analysis is motivated in this dissertation by the problem of the identification of a real-life system with an input monomial nonlinearity: the inertial gyroscope used for the measurement of angular rate (rotation speed).

This type of gyroscopes presents a lot of advantageous properties such as its cheap manufacturing price, its low energy consumption and its small size. However, it suffers from several inaccuracies due to manufacturing imperfections and its instrumentation. This problem can be tackled by putting the sensor in closed-loop with a controller. In order to derive an optimal controller, we need an accurate model of the inertial gyroscope and system identification advanced techniques such as Prediction Error can solve this problem. However, in the literature, there is no complete model structure and no identification method for the modeling of the main dynamics in an inertial gyroscope. Therefore, the last problem addressed in this thesis is the development of a complete model structure and an identification approach to accurately estimate the main dynamics of the inertial gyroscope.

Key-words: Prediction Error identification; consistency analysis; MIMO systems; data informativity; Hammerstein system identification; MEMS gyroscope

Résumé de la thèse

UNIVERSITÉ DE LYON

Ecole Centrale de Lyon - Laboratoire Ampère

Informativité des Données pour l'Identification par Erreur de Prédiction des Systèmes MIMO. Identification d'un Gyroscope MEMS.

de Kévin Colin

Les modèles mathématiques ont une place très importante dans beaucoup de domaines de l'ingénierie. Ces modèles peuvent être utilisés dans plusieurs applications de l'automatique comme, par exemple, la synthèse de correcteurs, le diagnostic des systèmes, la prédiction, etc. Souvent, les performances du contrôle, la qualité de la prédiction et de la capacité de détection des pannes sur un système dépendent directement de la qualité du modèle mathématique à décrire les dynamiques de ce système. L'identification des systèmes est le domaine scientifique qui consiste à déterminer un modèle mathématique des dynamiques d'un système en utilisant des données expérimentales récoltées sur ce système. Cette approche donne une modélisation plus réaliste du système puisqu'elle permet, par exemple, de modéliser des dynamiques qui ne peuvent pas être décrites par des équations de la physique. L'identification des systèmes repose sur trois ingrédients: la structure de modèle (l'ensemble des modèles candidats), les données récoltées et le critère d'identification. Le choix de ces trois ingrédients est important puisqu'il a un impact sur la qualité du modèle identifié.

Pour avoir une estimée précise des dynamiques du système identifié, il est important de garantir la consistance de l'estimée et cette propriété dépend fortement du choix de ces trois ingrédients. La consistance est la propriété qui assure la convergence de l'estimée vers le modèle le plus précis disponible dans la structure de modèle choisie lorsque le nombre de données tend vers l'infini. Lorsque l'estimateur considéré est celui de la méthode par erreur de prédiction (Prediction Error) avec un critère des moindres carrées, la consistance est garantie lorsque deux propriétés sont vérifiées: l'identifiabilité de la structure de modèle et l'informativité des données par rapport à la structure de modèle. L'identifiabilité a été largement étudiée et des structures de modèle linéaires paramétrées avec des propriétés d'identifiabilité fortes ont été développées: c'est le cas des structures Box-Jenkins, Output-Error, Réponse Impulsionnelle Finie, ARX and ARMAX.

Pour l'informativité des données, plusieurs études ont été menées pour les structures de modèle linéaire une-entrée une-sortie avec la méthode Prediction Error. Des conditions nécessaires et suffisantes ont été développées pour vérifier l'informativité des

données pour les structures de modèle classiques (Box-Jenkins, Output-Error, Réponse Impulsionnelle Finie, ARX and ARMAX) pour l'identification en boucle ouverte et l'identification directe en boucle fermée. Cependant, pour l'identification des systèmes plusieurs-entrées plusieurs-sorties, il y a peu de résultats pour vérifier l'informativité des données. Les seules conditions disponibles dans la littérature scientifique sur ce sujet sont des conditions qui ne sont que suffisantes et parfois très restrictives par rapport aux limitations du choix de l'excitation pour l'identification. Ainsi, la première motivation de cette dissertation est d'étudier l'informativité des données pour les systèmes linéaires plusieurs-entrées plusieurs-sorties et de développer des conditions de vérification moins restrictives dans les cas boucle ouverte et boucle fermée avec la méthode par erreur de prédiction.

La deuxième limitation des études de l'informativité des données avec la méthode Prediction Error est qu'elles se limitent au cas linéaire alors que la plupart des systèmes réels présentent des dynamiques non-linéaires. La méthode Prediction Error peut être utilisée pour identifier efficacement certains types de systèmes non-linéaires. Ainsi, le deuxième problème abordé dans cette dissertation est l'informativité des données pour une structure de modèle non-linéaire avec une non-linéarité entrée de type puissance. L'étude de ce type de structure est motivée par l'identification d'un système réel qui présente ce type de non-linéarité: le gyroscope inertiel pour la mesure de vitesses de rotation.

Ce type de gyroscope a de nombreux avantages comme son faible prix, sa faible consommation d'énergie et sa petite taille. Cependant, il n'est pas très précis à cause des problèmes de fabrication imparfaite et de son instrumentation. Ainsi, ce capteur est souvent mis en boucle fermée avec un correcteur optimal pour améliorer la mesure de la vitesse de rotation. Afin d'obtenir un correcteur optimal, nous avons besoin d'un modèle précis du gyroscope et l'identification des systèmes peut résoudre ce problème. Cependant, dans la littérature, il n'y a pas de travaux qui proposent une modélisation complète des phénomènes les plus importants du gyroscope. Ainsi le dernier problème étudié dans cette thèse est le développement d'une structure de modèle complète et d'une technique d'identification pour les principales dynamiques du gyroscope inertiel dans le but de synthétiser un correcteur optimal.

Mots-clefs: identification par erreur de prédiction; analyse de consistance; informativité des données; identification des systèmes Hammerstein; gyroscope MEMS

Contents

Contents	xi
List of Figures	xv
List of Tables	xviii
1 Introduction	1
1.1 System Identification	1
1.2 Consistency of the identification	4
1.3 State of the art on data informativity and first problem formulations . .	6
1.4 From LTI systems to nonlinear systems identifiable within the Prediction Error framework	8
1.5 Real-life example of this thesis: MEMS gyroscope	9
1.6 Contributions and outline of the thesis	16
2 Prediction Error Identification	21
2.1 Introduction of Chapter 2	21
2.2 Discrete-time quasi-stationary signals	22
2.3 Prediction Error identification of MIMO systems in open-loop	23
2.4 Consistency and Informativity in closed-loop	26
2.5 MIMO model structure considered in this dissertation	28
2.6 Previous results on MIMO data informativity	30
2.7 Persistency of a vector of quasi-stationary signals	32
2.8 Conclusion of Chapter 2 and reformulation of Research Question 1 . . .	33
3 Data Informativity for MIMO Systems in Open-Loop	35
3.1 Introduction of Chapter 3	36
3.2 Simplification of the study of data informativity for MIMO systems . . .	36
3.3 Data informativity for MISO structures in open-loop	38
3.4 Multisine excitation yielding informativity	49
3.5 Filtered white noise excitation yielding informativity	52

3.6	Checking informativity without input regressor	54
3.7	A synthetic verification approach of the data informativity for the open-loop identification of MIMO systems	55
3.8	Numerical example	56
3.9	Conclusion of Chapter 3	61
4	Data Informativity for MIMO Systems in Closed-Loop	63
4.1	Introduction of Chapter 4	64
4.2	Simplification of the study	64
4.3	Rewriting of the left-hand side of (4.3)	66
4.4	Rewriting of $\bar{E} [\mathbf{Y}_{u,i}(z)\mathbf{r}(t) ^2] = 0$ into (4.15)	74
4.5	Interpretation of Theorem 4.4	75
4.6	A synthetic approach for the verification of the data informativity for the direct closed-loop identification of MIMO systems	77
4.7	Numerical example	78
4.8	Conclusion of Chapter 4	84
5	Data informativity for a particular parallel Hammerstein system	85
5.1	Introduction of Chapter 5	85
5.2	System of the study, Prediction Error framework and data informativity	86
5.3	White Gaussian noise excitation	89
5.4	Multisine excitation	89
5.5	Synthesis of the results	92
5.6	A numerical example	92
5.7	Conclusion of Chapter 5	96
6	MEMS gyroscope and modeling objectives	99
6.1	Introduction of Chapter 6	99
6.2	Description of the MEMS gyroscope architecture	101
6.3	General principle of angular rate deduction	101
6.4	Main dynamics in the MEMS gyroscope	103
6.5	Controller and principle of angular rate deduction	112
6.6	Conclusion of Chapter 6	115
7	Identification for the control of the MEMS gyroscope	117
7.1	Introduction of Chapter 7	118
7.2	Identification procedure and model structure	119
7.3	Determination of an initial model	122
7.4	MISO identification and data informativity	133
7.5	Block "Parasitic Compensation $\mathbf{K}_{comp}(z)$ "	137
7.6	Block "Tracking Controller $\mathbf{K}(z)$ "	142
7.7	Conclusion	150
8	Conclusion	153
8.1	Summary of the dissertation	153

8.2 Leads for future works	155
Bibliography	159
Appendices	167
A Appendix for Chapter 2	169
A.1 Degeneration of the power spectrum matrix of multisine vectors	169
B Appendix for Chapter 3	173
B.1 Proof of Theorem 3.1	173
B.2 Proof of Lemma 2.1	175
B.3 Proof of Theorem 3.2	175
B.4 Proof of Theorem 3.3	176
B.5 Data informativity for MIMO ARMAX model structures	178
B.6 Proof of the persistency of $\varphi^{(sin)}$	184
B.7 Proof of Lemma 3.5	185
B.8 Lemma 3.5 for the ARMAX model structure	186
C Appendix for Chapter 4	189
C.1 Proof of Theorem 4.1	189
C.2 Proof of Lemma 2.2	190
C.3 Proof of Theorem 4.2	191
C.4 Proof of Lemma 4.1	191
C.5 Proof of Theorem 4.3	193
C.6 Orders of the polynomials $\Upsilon_{y,ii}(z)$ and $\Upsilon_{u,ik}(z)$	193
D Appendix for Chapter 5	195
D.1 Proof of Lemma 5.1	195
D.2 Proof of Theorem 5.1	197
E Appendix for Chapter 6	201
E.1 Inertial effects for objects in rotation	201
F Appendix for Chapter 7	203
F.1 Comparison between SISO-by-SISO approach and MISO identification	204
F.2 Hankel-singular-value-based reduction	208
F.3 Initialization method for the grey-box modeling	213
F.4 Bode of the tracking controller $\mathbf{K}(z)$	216
G Long résumé : Introduction de la thèse	221
G.1 Identification des systèmes	221
G.2 Consistance d'un estimateur pour l'identification des systèmes	223
G.3 État de l'art sur l'informativité des données et premières questions de recherche	224

G.4	Et pour les systèmes non-linéaires identifiables avec la méthode Prediction Error ?	226
G.5	Étude pratique de la thèse : gyroscope MEMS	227
G.6	Organisation du long résumé des contributions	232
H	Long résumé : L'identification par erreur de prédiction	233
H.1	Les signaux quasi-stationnaires à temps discret	233
H.2	La Méthode Prediction Error pour l'identification des systèmes MIMO en boucle ouverte	234
H.3	Identification en boucle fermée et consistance	237
H.4	Structures de modèles considérées dans cette thèse	238
H.5	Résultats sur l'informativité des données dans le cas MIMO présents dans la littérature	240
H.6	Persistance d'excitation de vecteurs de signaux quasi-stationnaires	242
H.7	Conclusion de l'Annexe H	243
I	Long résumé : Résumé des Chapitres 3, 4 et 5	245
I.1	Résumé du Chapitre 3 (Question de Recherche 1)	246
I.2	Résumé du Chapitre 4 (Question de Recherche 2)	256
I.3	Résumé du Chapitre 5 (Question de Recherche 3)	266
J	Long résumé : Gyroscope MEMS et objectifs de modélisation	273
J.1	Présentation de l'architecture du gyroscope MEMS	274
J.2	Principe général de mesure de la vitesse angulaire Ω	274
J.3	Dynamiques principales du gyroscope MEMS	275
J.4	Correcteur et mesure de la vitesse angulaire	279
J.5	Conclusion sur le résumé du Chapitre 6	281
K	Long résumé : Résumé du Chapitre 7	283
K.1	Méthode d'identification	284
K.2	Limitations de l'expérimentation et informations a priori	286
K.3	Initialisation : approche SISO-par-SISO	288
K.4	Identifications MISO et résultats	290
K.5	Réduction d'ordre du modèle par identification boîte grise	293
K.6	Conclusion sur le résumé des résultats du Chapitre 7	296
L	Long résumé : Conclusion de la thèse	297
L.1	Conclusion sur les travaux réalisés pendant la thèse	297
L.2	Ouvertures pour des recherches futures	299

List of Figures

1	Une photo d'un hamac. Source: https://www.tropical-hamac.com/	v
1.1	The three ways to derive a mathematical model and the place of system identification.	3
1.2	Block scheme of the open-loop system (top) and of the closed-loop system (bottom).	4
1.3	A MEMS gyroscope GYPRO® designed by Tronics (Source: [EL]).	10
1.4	Description of the MEMS gyroscope (MIMO system) with two inputs u_1 and u_2 and two outputs y_1 and y_2 into four subsystems with one input and one output.	16
2.1	Scheme of the closed-loop.	27
5.1	Scheme of the true system \mathcal{S} considered in Chapter 5.	86
6.1	Simplified scheme of the MEMS gyroscope.	102
6.2	Block scheme of the mechanical dynamics of the MEMS gyroscope.	103
6.3	Simplified scheme of a comb-drive (Source: [KBS ⁺ 17]).	104
6.4	Simplified scheme of the MEMS gyroscope, instrumented with capacitive comb-drives.	105
6.5	Block scheme of the instrumented mechanical dynamics of the MEMS.	107
6.6	Illustration of one part of the parasitic electrical coupling.	108
6.7	Block scheme of the main dynamics of the MEMS gyroscope.	110
6.8	Block scheme of the MEMS gyroscope dynamics described by the three transfer function matrices $\mathbf{G}_0(z)$, $\mathbf{Q}_0(z)$ and $\mathbf{H}_0(z)$	111
6.9	General scheme of the MEMS gyroscope in closed-loop with a controller.	112
6.10	A detailed scheme block of the two-degree-of-freedom controller.	113
6.11	Comparison between the RF transmission and the Coriolis effect in the MEMS gyroscope.	116
7.1	Bode diagrams of the mechanical coupling from the SISO-by-SISO approach.	131

7.2	Bode diagrams of the parasitic capacitive coupling from the SISO-by-SISO approach.	132
7.3	Estimated PSD of v_x (left) and v_y (right) from the SISO-by-SISO approach.	133
7.4	Illustration of the sets of frequency for the multisine excitation.	134
7.5	Scheme of the experiment for the verification of the parasitic coupling compensation.	138
7.6	FFT magnitudes of the drive output before compensation \tilde{x} (blue), after compensation \tilde{x}_c (orange) and the simulated drive output \tilde{x}_{sim} from $\hat{\mathbf{G}}(z)$ (yellow).	140
7.7	Zoom of Figure 7.6 around the resonance frequencies.	140
7.8	FFT magnitudes of the sense output before compensation \tilde{y} (blue), after compensation \tilde{y}_c (orange) and the simulated sense output \tilde{y}_{sim} from $\hat{\mathbf{G}}(z)$ (yellow).	141
7.9	Zoom of Figure 7.8 around the resonance frequencies.	141
7.10	Comparison of Bode diagram magnitudes between the identified black-box model (blue solid line) and the identified grey-box model (orange dotted line).	146
7.11	Comparison of Bode diagram phases between the identified black-box model (blue solid line) and the identified grey-box model (orange dotted line).	147
7.12	Block scheme for the comparison between the measured output \tilde{x} in closed-loop and the simulated one \tilde{x}_{mod} . The term \mathbf{G}_{gb} denotes the identified grey-box model.	148
7.13	Comparison of the drive measured output \tilde{x} (blue) and the drive reference signal x_{ref} (yellow).	149
7.14	Zoom of Figure 7.13 around $t_c = 6$ s.	150
E.1	An object M moving into a rotating frame of reference.	202
F.1	Bode diagram magnitudes of the identified \mathbf{G} from the SISO-by-SISO approach (solid blue line) and the identified \mathbf{G} from the MISO identifications (orange dashed line).	204
F.2	Bode diagram phases of the identified \mathbf{G} from the SISO-by-SISO approach (solid blue line) and the identified \mathbf{G} from the MISO identifications (orange dashed line).	205
F.3	Bode diagram magnitudes of the identified \mathbf{Q} from the SISO-by-SISO approach (solid blue line) and the identified \mathbf{Q} from the MISO identifications (orange dashed line).	206
F.4	Bode diagram phases of the identified \mathbf{Q} from the SISO-by-SISO approach (solid blue line) and the identified \mathbf{Q} from the MISO identifications (orange dashed line).	207
F.5	Comparison of the DSPs of the measurement noises v_x (left) and v_y (right) between the SISO-by-SISO approach (blue solid line) from the MISO identifications (orange dotted line).	208
F.6	Bode diagram magnitudes of the black-box model (blue solid line) and of the reduced-order model based on Hankel singular values of order 8 (orange dotted line).	209

F.7	Bode diagram phases of the black-box model (blue solid line) and of the reduced-order model based on Hankel singular values of order 8 (orange dotted line).	210
F.8	Bode diagram magnitudes of the identified black-box model (blue solid line) and of the reduced-order model based on Hankel singular values of order 6 (orange dotted line).	211
F.9	Comparison of Bode diagram phases of the identified black-box model (blue solid line) and of the reduced-order model based on Hankel singular values of order 6 (orange dotted line).	212
F.10	Bode diagram magnitudes of the MIMO tracking controller $\mathbf{K}(z)$	217
F.11	Bode diagram phases of the MIMO tracking controller $\mathbf{K}(z)$	218

List of Tables

3.1	Mean of the identified parameter vector for the first channel with $N = 1000$ for different cases over 1000 experiments.	60
3.2	Mean of the identified parameter vector for the second channel with $N = 1000$ for different cases over 1000 experiments.	61
4.1	Mean of the identified parameter vector for the first channel with $N = 500$ for different cases over 1000 experiments.	83
4.2	Mean of the identified parameter vector for the second channel with $N = 500$ for different cases over 1000 experiments.	83
5.1	Mean of the identified parameter vector for the first channel with $N = 1000$ for different cases over 1000 experiments.	95
5.2	Mean of the identified parameter vector for the second channel with $N = 1000$ for different cases over 1000 experiments.	95
7.1	Obtained degrees of the polynomials for the mechanical coupling functions.	128
7.2	Obtained degrees of the polynomials for the parasitic coupling transfer functions.	128
7.3	Obtained degrees of the polynomials for the noise transfer functions.	128
7.4	Best Fit for the mechanical coupling from the SISO-by-SISO approach.	130
7.5	Best Fit for the parasitic coupling from the SISO-by-SISO approach.	130
7.6	Best Fit for the models obtained from the MISO approach with $N = 50000$	137

Acronyms

ADC	Analog-to-Digital Converter
BJ	Box-Jenkins
DAC	Digital-to-Analog Converter
DC	Direct Current
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
LPV	Linear Parameter Varying
LTI	Linear Time Invariant
MEMS	Micro-ElectroMechanical Structure
MIMO	Multiple-Inputs Multiple-Outputs
MISO	Multiple-Inputs Single-Output
OE	Output-Error
ONERA	Office National d'Études et de Recherches Aérospatiales
PE	Persistently Exciting
PI	Proportional-Integral
PID	Proportional-Integral-Derivative
PSD	Power Spectrum Density
RBS	Random Binary Sequence
RF	Radio-Frequency / Radio-Fréquence
SISO	Single-Input Single-Output
SNR	Signal-to-Noise Ratio
ZOH	Zero-Order Hold

Notations

Sets

- \mathbb{N} Set of nonnegative integers
 \mathbb{N}^* Set of nonnegative integers excluding 0
 \mathbb{Z} Set of integers
 \mathbb{R} Set of real scalars

For two nonnegative integers n and p such that $p \geq n$, we denote by $\llbracket n, p \rrbracket$ the set of consecutive integers between n and p , i.e., the set $\{n, n + 1, \dots, p - 1, p\}$.

Matrices, vectors, scalars

The matrices are denoted with uppercase letters in bold. The column vectors are denoted with lowercase letters in bold. Finally, scalars are denoted either with uppercase or lowercase letters but not in bold.

For a matrix \mathbf{A} , we denote by $\mathbf{A}_{:k}$ its k^{th} column and by $\mathbf{A}_{i:}$ its i^{th} row. Finally, we denote by A_{ik} the (i, k) -entry of the matrix \mathbf{A} . A positive semidefinite (resp. definite) matrix \mathbf{A} is denoted by $\mathbf{A} \succeq 0$ (resp. $\mathbf{A} \succ 0$). The identity matrix of dimension $n \times n$ is denoted by \mathbf{I}_n . The notation $\mathbf{0}_{m \times n}$ denotes the matrix of dimension $m \times n$ whose entries are all equal to 0.

The notation $\text{diag}(a_1, \dots, a_n)$ refers to the $n \times n$ diagonal matrix whose elements in its diagonal are the scalars a_1, \dots, a_n . For given matrices $\mathbf{A}_1, \dots, \mathbf{A}_n$, the notation $\text{bdiag}(\mathbf{A}_1, \dots, \mathbf{A}_n)$ denotes the following block diagonal matrix

$$\begin{pmatrix} \mathbf{A}_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{A}_n \end{pmatrix}$$

For a column vector \mathbf{x} , we denote by x_k its k^{th} entry. The notation $\|\mathbf{x}\|$ refers to the Euclidean norm, i.e., $\|\mathbf{x}\| = \sqrt{x_1^2 + \dots + x_n^2}$.

Polynomials

The degree of a polynomial $P(X)$ is denoted by $\deg(P(X))$. We will use the following notation for a polynomial: $P(X) = p_0 + p_1X + p_2X^2 + \dots + p_nX^n$. When the subscript index becomes heavy, we will adopt the following notation $P(X) = p^{(0)} + p^{(1)}X + p^{(2)}X^2 + \dots + p^{(n)}X^n$.

When $X = z^{-1}$ with z the discrete-time forward shift operator, we say that ρ is the delay of $P(z^{-1})$ when the first nonzero coefficient is linked to $z^{-\rho}$, i.e., $P(z^{-1}) = p_\rho z^{-\rho} + p_{\rho+1}z^{-(\rho+1)} + p_{\rho+2}z^{-(\rho+2)} + \dots$.

Operators

E	Expectation operator
\otimes	Matrix Kronecker product
\wedge	Vector product
det	Determinant operator
tr	Trace operator
z	Forward shift operator (same notation for the \mathcal{Z} transform variable)
!!	For odd integers p , $p!! = p \times (p-2) \times (p-4) \times \dots \times 3 \times 1$

Variables

ω	Angular frequency (in rad/s)
Ω	Angular rate around a given axis that the MEMS gyroscope has to measure
\mathcal{S}	True system to be identified
\mathbf{u}	Input vector of the true system \mathcal{S}
\mathbf{y}	Output vector of the true system \mathcal{S}
\mathbf{e}	White noise vector of the true system \mathcal{S}
\mathbf{G}_0	Discrete-time transfer function from \mathbf{u} to \mathbf{y}
\mathbf{H}_0	Discrete-time transfer function from \mathbf{e} to \mathbf{y}
\mathbf{K}	Controller
p	Laplace variable
t_c	Continuous time variable
t	Integer indexing discrete-time signals

Introduction

Contents

1.1	System Identification	1
1.2	Consistency of the identification	4
1.3	State of the art on data informativity and first problem formulations	6
1.4	From LTI systems to nonlinear systems identifiable within the Prediction Error framework	8
1.5	Real-life example of this thesis: MEMS gyroscope	9
1.5.1	Presentation of the challenges in the MEMS gyroscope . . .	9
1.5.2	Goal and brief presentation of the Next4MEMS project . .	11
1.5.3	State of the art in MEMS gyroscope identification and problem formulation	13
1.6	Contributions and outline of the thesis	16

1.1 System Identification

Mathematical models of the dynamics of real-life systems play an important role in a variety of science and engineering fields such as chemistry, mechatronics, electromagnetism, biology, etc. These models are needed for different types of purposes such as controller design, prediction, health-monitoring, reverse-engineering, etc:

- controller design: in a lot of engineering applications, the system must be designed such that its behavior satisfies the desired one. For this purpose, a controller is synthesized in order to design the optimal actuation that will guarantee the desired behavior. A mathematical model of the system has a crucial part in this controller design, as illustrated in [NDG⁺04] for a catalytic distillation process.

- prediction: when the forthcoming values of the output of a system need to be predicted under some given inputs and operating conditions, a model can be used for this purpose. An example is the forecast of weather [Vas01].
- health monitoring: it is important to detect the faults occurring in a system. A model of the nominal behavior of the system is then used to detect any deviation of the observed data from the nominal (desired) ones [Din08].
- reverse-engineering: a model can also be used in order to deduce some behaviors that are invisible for the user at first glance. It is the case, e.g., in biology where the feedbacks between all the components of a cell are investigated [Kha08, TA05].

The obtained performances of the controller, the prediction, etc are often dependent on the quality of the model, i.e., on the ability of the model on explaining the observed dynamics of the system. It is therefore important to guarantee an accurate model for the system.

The ways to obtain a model can be divided into three mainstream categories. In the first one, the model is completely determined from first principles (e.g., Newton's laws for the mechanics, Maxwell's laws for the electromagnetism, etc). It corresponds to the white-box modeling category. In the second one, the model is only derived from experimental data collected on the system. This corresponds to the black-box modeling. The last category is a case between the white and black-box models and is naturally called grey-box modeling. It corresponds to the determination of all the (or some) parameters in the first principle equations by the use of experimental data collected on the system. An illustration of these three ways to derive a mathematical model is given in Figure 1.1.

System identification is the scientific field consisting in determining a mathematical model with experimental data collected on the system to be modeled. It regroups the grey-box and black-box modeling approaches (see Figure 1.1). Among all the identification methods that have been derived in the literature, we can cite, e.g., Prediction Error [Lju99], frequency domain identification [PS12], subspace identification [VODM96], etc.

To perform an identification, the user needs three main ingredients: a model structure, a data set and an identification criterion. A model structure corresponds to a set of candidate models. Several characteristics of this model structure must be chosen by the user. For instance, the set of candidate models can be grey-box/black-box, linear/nonlinear, parametrized/non-parametrized¹, etc. The second main ingredient that is the very essence of system identification is the collected data set on the true system. Finally, the last ingredient is the identification criterion that will be used to

¹When is it parametrized, the user has to choose, e.g., the order of the parametrized transfer functions.

do a ranking of all the candidate models in the model structure and to compute the optimal model among them (the one with the highest ranking) that explained at best the observed data with respect to the identification criterion.

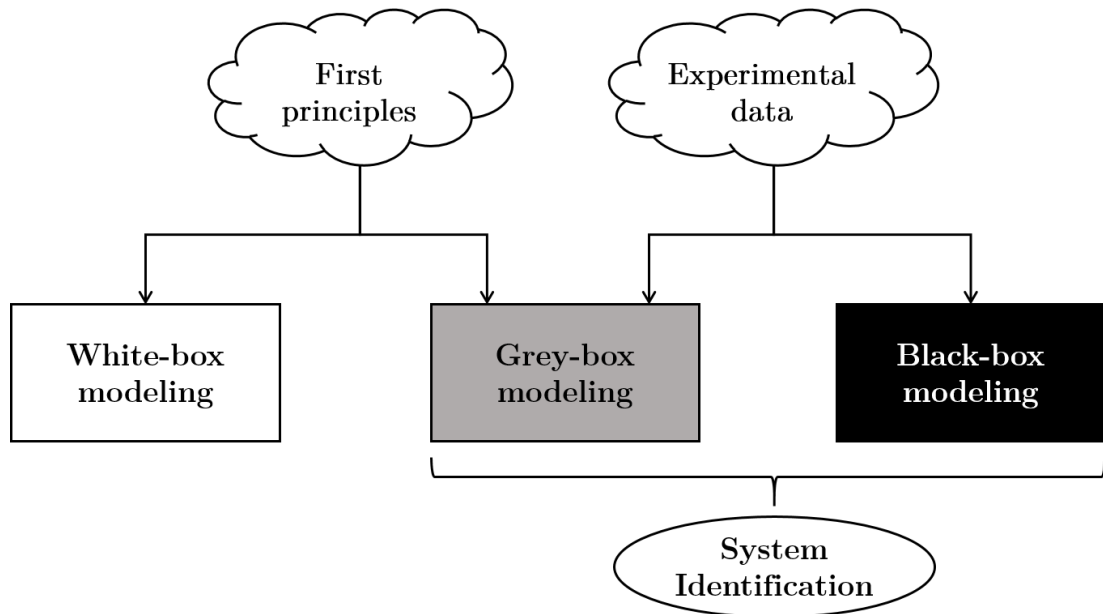


Figure 1.1: The three ways to derive a mathematical model and the place of system identification.

The identification of a system can be done in the open-loop or closed-loop configuration (see Figure 1.2). When the system is put under feedback control with a stabilizing controller, we often want to identify a model of the system without disconnecting the controller. In this case, an external excitation can be added to the control effort for identification purposes as illustrated in Figure 1.2.

In order to guarantee an accurate model, it is important that the chosen model structure can capture all the dynamics of the to-be-identified system. It is the so-called full-order property of the model structure. This means that there is one model in the model set that explains exactly all the dynamics of the system and which is called the true model. Then, the goal of the modeling procedure is to estimate this true model. When the model structure is parametrized, the parameter vector links to the true model is called the true parameter vector and we search for an estimate of this true parameter vector.

However, due to the presence of some unknown disturbances such as noise in the data set used for the identification, it is never possible to estimate exactly this true model (or the true parameter vector for parametrized model structures) from a real identification experiment since the number of data and the excitation power are limited in practice.

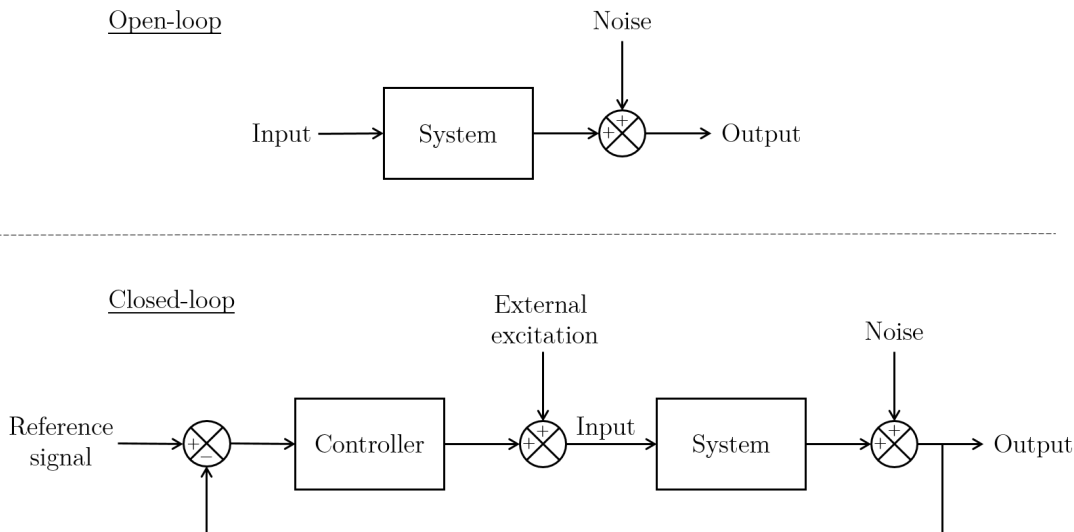


Figure 1.2: Block scheme of the open-loop system (top) and of the closed-loop system (bottom).

Consequently, in all identification experiments, we look for an accurate estimate of this true model, i.e., an estimate close to the true model. For this purpose, it is essential to guarantee the consistency of the estimate. We deal with the notion of consistency in the next paragraph.

1.2 Consistency of the identification

The consistency is the crucial property guaranteeing the fact that the identified estimate (or the identified parameter vector for parametrized model structures) converges to the true model (or the true parameter vector for full-order parametrized model structures) when the number of data increases endlessly. If this property is not ensured, the identified model will be an inaccurate estimate of the true model. The consistency depends on the chosen model structure, the collected data set and the identification criterion.

In this dissertation, we will focus on the classical black-box Prediction Error identification with a least-square identification criterion. The model structures considered in this method are parametrized but the parameters do not have necessarily physical meanings (black-box). Based on a chosen parametrized model structure and some data collected on the system to-be-identified, we construct the one-step ahead predictor and the identification criterion which is a least-square criterion on the difference between the measured output and the predictor. This difference is called the prediction error. We then compute an estimate of the true parameter vector by minimizing this least-square criterion. The resulting estimate is desired to be consistent with respect to the true parameter vector.

The consistency property for this classical method has been widely studied in the 1970's, like, e.g., in [Lju76a, Lju76b]. A lot of works have been derived on the direct closed-loop identification of systems with different assumptions on the system and the external excitation [SGL75, SLG76, AG82, GLS77]. However, these conditions are only sufficient for the consistency.

Later, the work of Ljung in [Lju99] proposed new conditions for the consistency by connecting it to the uniqueness of the minimizer of the asymptotic identification criterion² (which is the true parameter vector) or, equivalently, by linking it to the uniqueness of the parameter vector that guarantees the whiteness of the prediction error. For this purpose, two concepts have been introduced: the identifiability of the model structure and the informativity of the data with respect to the chosen model structure.

The identifiability of a model structure parametrized by a given parameter vector is guaranteed when two different parameter vectors do not yield the same model. The identifiability depends on the parametrization of the considered model structure. A lot of works have been done to develop parametrized model structures which have interesting identifiability properties. It is, for instance, the case of the classical single-input single-output (SISO) model structures which are Finite Impulse Response (FIR), ARX, ARMAX, Output-Error (OE) and Box-Jenkins (BJ). These model structures have been also extended for multiple-inputs multiple-outputs (MIMO) systems.

The data informativity ensures the fact that two different models in the considered model structure do not yield the same prediction error. Combined with the identifiability of the model structure, the informativity guarantees that the true parameter vector is the only parameter vector that makes the prediction error white or, equivalently, the true parameter vector is the unique minimum of the power of the prediction error (i.e., the asymptotic least-square criterion). The data informativity property directly depends on the model structure complexity and on the quantity of information about the true system contained in the collected data set. More precisely, the data informativity is guaranteed when there is enough information about the true system in the collected data set in order to guarantee the distinctiveness of the predictor (or prediction error) for all the models in the model structure. If the data are not informative with respect to the chosen model structure, then the consistency will not be guaranteed and so the parameter vector computed from an identification experiment will be an inaccurate estimate of the true parameter vector, even if the number of data is large. On the contrary, if the data are informative, then any level of accuracy can be achieved for the identified model, provided that the number of data can be increased. The data informativity property is thus a necessary property to be guaranteed in order to derive an accurate model. A state of the art on the data informativity for Prediction Error identification is given in the next section.

²For the Prediction Error method with a least-square criterion, the asymptotic identification criterion is the power of the prediction error.

1.3 State of the art on data informativity and first problem formulations

The data informativity has been largely studied for the classical SISO linear time invariant (LTI) model structures (FIR, ARX, ARMAX, OE, BJ) in both the open-loop identification and the direct closed-loop identification. These results are based on a notion of signal richness for the excitation signal that is used during the identification: an excitation signal is sufficiently rich of an order ν if and only if its input power spectrum has a nonzero amplitude in ν different frequencies in the interval $]-\pi, \pi]$ (see e.g. [Lju99, SS89]).

Based on this notion of signal richness, necessary and sufficient conditions have been derived for the data informativity in both the open-loop and closed-loop case in [GBM08, GBBM09]. In particular, these papers derive the minimal order of signal richness that the excitation signal must have to ensure data informativity and this is done for all classical model structures. As shown in [GBM08, GBBM09], this minimal order depends on the type and the complexity of the considered model structure. In the closed-loop case, this minimal order is also related to the complexity of the controller present in the loop during the identification. In [GBM08, GBBM09], it is also shown that, if the controller is sufficiently complex, the data can be informative even if the excitation signal is equal to zero (costless identification).

As already mentioned, due to its link with the consistency property, the data informativity is necessary to guarantee that the identified model can be made as accurate as desired by an appropriate choice of the experiment duration, i.e., the number of data. This link between accuracy and data informativity has led to research connecting the data informativity to the positive definiteness of the covariance matrix of the identified parameter vector [BBG12, GBBM09]. When the model structure is full-order, this covariance matrix is indeed a measure of model accuracy.

While the data informativity seems to be a mature research area in the SISO case for LTI systems, this cannot be said for the MIMO case. However, the study of the data informativity for the MIMO case is of interest since the systems get more and more complex with an increase in the number of inputs and outputs. In the restricted literature on data informativity for MIMO systems, attention has been mostly given to the costless identification case in closed-loop, i.e., the problem of verifying whether the controller is sufficiently complex to ensure that informative data are obtained when the external excitation is zero. In [BGM10], a sufficient condition is provided for this situation considering general MIMO model structures³. Restricting attention to multiple-inputs single-output (MISO) ARMAX model structures (and at the cost of some model structure and controller simplifications), the paper [YZ18] presents a necessary and sufficient condition for the costless identification situation.

³Another condition which is only necessary is also provided in [BGM10].

The costless identification case is a very particular situation. In the vast majority of the closed-loop identification experiments, a (nonzero) external excitation will be required to yield informative data and, in open-loop identification, data informativity can of course only be obtained with a nonzero excitation signal. Perhaps, the only result that allows to check whether a given external excitation yields informative data is formulated in [BGM10]⁴. This result is as follows: if the excitation signal has a strictly positive definite power spectrum matrix at (almost) all frequencies, then this excitation signal yields informative data and this holds for both open-loop and closed-loop identifications. This condition is of course only sufficient and is moreover very restrictive. As an example, a multisine excitation will never respect this condition while such excitation may well be the only type of excitations that is actually possible due to experiment design restrictions. Consequently, we will study the data informativity for the open-loop and direct closed-loop identification of LTI MIMO systems, i.e., we will consider the two following research questions in this dissertation:

Research Question 1

Given a MIMO LTI system in *open-loop*, excited by a given input excitation and identified within a given MIMO linear model structure, how can we verify that the data collected on the system are informative for the open-loop identification of the system within the Prediction Error framework?

Research Question 2

Given a MIMO LTI system in *closed-loop* with a given stabilizing controller, excited by a given external excitation and identified within a given MIMO linear model structure, how can we verify that the data collected on the system are informative for the identification of the system with the direct closed-loop approach within the Prediction Error framework?

To address Research Questions 1 and 2, we will first show that the data informativity with respect to (w.r.t.) MIMO generalizations of the classical SISO model structures (ARX, FIR, etc) can be treated by analyzing the data informativity channel-by-channel in both the open-loop and the closed-loop cases.

⁴In [BGM10], direct closed-loop identification is considered, but the result can be extended to open-loop identification which is a special case where the controller is equal to zero.

For the open-loop identification case (Research Question 1), an analysis of the data informativity channel-by-channel boils down to the analysis of data informativity for Multiple-Input Single-Output (MISO) systems. We will see that a sufficient condition for data informativity with respect to a MISO model structure⁵ is the linear independence of the elements of a vector of signals that we will call *regressor vector*. This regressor vector $\phi_{\mathbf{u}}$ contains delayed versions of the elements of the input vector. There is, of course, an analogy here with the SISO case since a scalar input signal is sufficiently rich of a given order ν if the elements of such a regressor vector of dimension ν are linearly independent [GBM08].

We then show that the linear independence of the elements of $\phi_{\mathbf{u}}$ is equivalent to the fact that a given matrix is full row rank when the input \mathbf{u} of the system is either a multisine or a filtered white $\mathbf{u}(t) = \mathbf{M}(z)\mathbf{v}(t)$ where $\mathbf{M}(z)$ is a matrix of stable transfer functions and $\mathbf{v}(t)$ a vector of independent white noises. This matrix will be a function of the input parametrization, i.e., the amplitudes, the phase shifts and the frequencies of the different sinusoids in the multisine case and the coefficients in the matrix $\mathbf{M}(z)$ in the filtered white noise case. We then analyze which conditions are necessary to make this matrix full row rank. We therefore derive, for the multisine case, conditions on the number of sinusoids that are present in the multisine input vector and, for the filtered white noise case, conditions on the complexity of the matrix $\mathbf{M}(z)$ and on the number of white noises in $\mathbf{v}(t)$.

For the direct closed-loop case (Research Question 2), we will follow a similar procedure. We will see that the data informativity will be guaranteed in the MIMO closed-loop case if, for each channel/output, a certain matrix is full row rank. This matrix depends on the model structure complexity, on the controller coefficients and on the external excitation parametrization (i.e., amplitudes, phase shifts and frequencies for multisine and filter coefficients for filtered white noise). We will also observe that this matrix clearly separates the contributions of the controller and of the external excitation to the informativity of the collected input-output data.

1.4 From LTI systems to nonlinear systems identifiable within the Prediction Error framework

It is to be noted that the literature and Research Questions 1 and 2 restrict the study of data informativity to LTI systems while most of real-life systems have some nonlinear behaviors. Fortunately, Prediction Error can also be used as an efficient tool to identify certain classes of nonlinear systems. It is, e.g., the case of block-oriented systems with static nonlinearities such as Hammerstein/Wiener systems [GB10, ST17]. Up to our knowledge, the problem of data informativity has never been studied in the literature for nonlinear systems which can be accurately identified with the Prediction Error framework. For instance, in this dissertation, we will consider the following discrete-

⁵When the model structure is ARX or FIR, the condition is also necessary.

time particular Hammerstein system given by

$$y(t) = G_{0,1}(z)u(t) + G_{0,2}(z)u^n(t) + H_0(z)e(t) \quad (1.1)$$

where u is the input, y is the output, $n \in \mathbb{N}^*$, e is a white noise and z the forward shift operator. The terms $G_{0,1}(z)$, $G_{0,2}(z)$ and $H_0(z)$ are three unknown discrete-time transfer functions that we want to estimate. Then, by considering a full-order nonlinear model structure and an excitation u , we have to ensure the informativity of the collected data on the to-be-identified system in order to guarantee an accurate estimate. Therefore, we will consider the following research question

Research Question 3

Given the following Hammerstein system in open-loop

$$y(t) = G_{0,1}(z)u(t) + G_{0,2}(z)u^n(t) + H_0(z)e(t) \quad n \in \mathbb{N}^*$$

a nonlinear model structure and an excitation u , how can we verify that the data collected on this nonlinear system are informative for its identification within the Prediction Error framework?

To address Research Question 3, we will show that we can recast this type of system as an equivalent MISO LTI one with one output y and two inputs given by $u_1 = u$ and $u_2 = u^n$. The idea is then to use the data informativity results for MISO systems that we have derived to address Research Question 1. We will address this study for two types of excitations u : white Gaussian noise and multisine.

Even though (1.1) is a rather particular structure, it is, e.g., representative of the dynamics of a micro-electromechanical structure (MEMS) gyroscope, presented in the next section.

1.5 Real-life example of this thesis: MEMS gyroscope

1.5.1 Presentation of the challenges in the MEMS gyroscope

Besides the theoretical part described above, this Ph.D. thesis will also consider an applied problem, i.e., the identification of a MEMS gyroscope [Apo16, Kem11]. It is an inertial sensor used to measure an angular rate (or rotation speed) around a given axis and which is based on the Coriolis effect. A MEMS gyroscope (GYPRO[®]) is depicted in Figure 1.3. Among the types of gyroscopes that exist such as the optical and atomic gyroscopes [FQ12], the MEMS gyroscope presents some advantageous characteristics like its microscopic size, its low price and its low energy consumption.

The MEMS gyroscope is made up of two microscopic masses vibrating in orthogonal directions. Each mass has a resonant behavior with a given resonance frequency⁶. When the angular rate is nonzero, both mass vibrations are coupled with the Coriolis effect which depends linearly on the angular rate. Therefore, the measurement principle of the MEMS gyroscope is to estimate the Coriolis effect in order to deduce the angular rate from this estimation. To estimate the Coriolis effect, we need to actuate and measure both masses vibrations. Each mass is excited with an electrostatic force generated from excitation voltages, one for each mass. Both excitation voltages are the inputs of the MEMS gyroscope. There are two outputs for the MEMS gyroscope which are also voltages, each voltage being an image of the displacement of one mass. Therefore, the MEMS gyroscope is a MIMO system.

Due to manufacturing imperfections, there is a mechanical cross-coupling⁷ between both masses and if this effect is not considered in the estimation of the angular rate, the estimate will be inaccurate. This coupling can be modeled with cross-coupling stiffness and damping coefficients. The cross-coupling stiffness coefficients illustrate the anisotropy effect and the cross-coupling damping coefficients are due to the nonproportional damping effect [Sau08]. However, the mechanical dynamics are not the only dynamics that are involved in the MEMS gyroscope. Indeed, because of the capacitive instrumentation used in the MEMS to actuate and measure both mass motions, additional dynamics can be observed such as the parasitic capacitive coupling and noise measurement [AS05, LWZ⁺17]. Moreover, a square input nonlinearity is also observed due to the actuation instrumentation. If these additional dynamics are ignored during the angular rate estimation, the quality of the estimate will be seriously hampered.

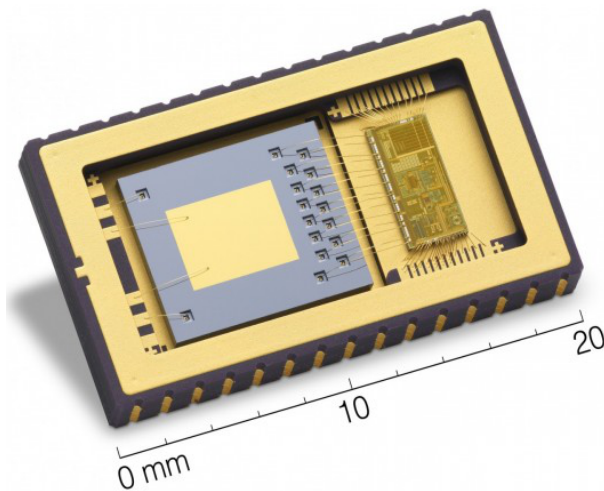


Figure 1.3: A MEMS gyroscope GYPRO® designed by Tronics (Source: [EL]).

⁶There are then two different resonance frequencies, one for each mass.

⁷This means that, even if there is no angular rate, a motion of the first mass leads to a motion of the second mass.

Consequently, the micro-sensor is often put in closed-loop with a MIMO controller in order to tackle these additional dynamics during the angular rate estimation. The control objectives are chosen such that the angular rate can be deduced in three steps. First, the displacement of one of both masses (called the drive mass) must track a sinusoidal reference signal. Moreover, the frequency of the sinusoidal reference signal must match the resonance frequency of the drive mass in order to decrease the power of the control effort on the drive mass⁸. Secondly, for the second mass (called the sense mass), the controller must guarantee that its motion caused by the Coriolis force is canceled out. Since the Coriolis force can be seen as a disturbance that generates a motion of the sense mass, the control effort on the second mass ensuring the motion compensation⁹ is an image of this Coriolis force. Therefore, the control effort applied to the sense mass serves as an estimation the Coriolis force. Finally, the angular rate is deduced from the estimation of the Coriolis force by exploiting the fact that the Coriolis force is proportional to the angular rate.

A lot of works have been dedicated to the design of linear and nonlinear controllers such as PI, PID and amplitude gain controllers, etc for the MEMS gyroscope [Sau08, EMK12, PTKH09, MV99]. More advanced control techniques such as adaptive sliding mode controllers and observer-based controllers have also been considered [FB09b, FB09a, Fei10, ZDLG08]. These works provide good accuracy performances for the measurement of the angular rate.

1.5.2 Goal and brief presentation of the Next4MEMS project

It might be possible to get better performances with other controller design approaches in order to compete better, e.g., with the optical gyroscopes in the market. This is the main objective in the project funding this thesis, called the Next4MEMS project and funded by Bpifrance¹⁰. In other words, the goal of this project is to develop MEMS gyroscopes with better performances than the current ones in the market.

The Next4MEMS project is a consortium of three private companies (Tronics/TDK, Asygn and iXblue) and two public laboratories (ONERA¹¹ and Ampère). The company Tronics/TDK manufactures MEMS gyroscopes such as the MEMS gyroscope GYPRO® depicted in Figure 1.3. The company Asygn designs the electronic card used for the controller implementation and the treatment of the angular rate measurement. The company iXblue manufactures inertial units for aerospace and (sub)marine applications. The task of the laboratory ONERA (specialized in aerospace studies) in this project is the study of the use of MEMS gyroscopes in some aerospace applications as well as the improvement of finite element model of mechanical parts of MEMS gy-

⁸The control effort of the drive mass is the excitation voltage used to generate the electrostatic force applied to this mass.

⁹The control effort of the sense mass is the excitation voltage used to generate the electrostatic force applied to this mass.

¹⁰French public investment bank.

¹¹Office National d'Études et de Recherches Aéropatiales (French aerospace laboratory).

rosopes. Finally, the goal of the laboratory Ampère in the Next4MEMS project is the development of a method for the design of optimal controllers for the MEMS gyroscope. Four Ph.D. students of the laboratory Ampère are participating in this project:

- Fabrício SAGGIN (supervised by Xavier BOMBOIS, Gérard SCORLETTI and Anton KORNIENKO). He works on the controller design. The approach that he has followed throughout his Ph.D. thesis is the \mathcal{H}_∞ design method. It allows to synthesize an optimal control law obtained from an optimization procedure and whose goal is to stabilize the system with guaranteed performances. This design method is quite powerful for MIMO systems with cross-coupling transfer functions such as the MEMS gyroscope.
- Kévin COLIN (supervised by Xavier BOMBOIS, Laurent BAKO and Anton KORNIENKO). The design of the controller is based on a model of the dynamics of the MEMS gyroscope. As aforementioned in Section 1.1, the performances of the controller are often dependent on the quality of the model and so it is important to compute an accurate model in order to synthesize an optimal controller with the \mathcal{H}_∞ design method. System identification can then be used to derive this accurate model and this will be the practical problem studied in this dissertation.
- Jorge AYALA (supervised by Gérard SCORLETTI and Anton KORNIENKO). He works on the robustness analysis part of the project which links the modeling and control design parts. As aforementioned, the controller is designed based on the identified model of the MEMS gyroscope. However, due to the uncertainties on the identified model and some possible neglected dynamics during the identification procedure, the performances of the controller might be deteriorated when implemented on the real system. Therefore, his goal is to develop some methods with reasonable computational cost which verify that the performances of the controller are still valid for all the models contained in a set of models described by the uncertainties and the neglected dynamics and for which the true (unknown) model belongs to.
- Federico MORELLI (supervised by Xavier BOMBOIS, Laurent BAKO and Anton KORNIENKO). For the angular rate measurement, the displacement of one of both masses (the drive mass) of the MEMS gyroscope must track a sinusoidal reference signal. Moreover, the frequency of the sinusoidal reference signal must match the resonance frequency of the drive mass, in order to decrease the control effort power. However, the mechanical properties of both masses in the MEMS gyroscope are dependent on the temperature [Sau08] and so do the resonance frequencies. Therefore, the frequency of the sinusoidal reference signal and/or the controller parameters must be adapted w.r.t. the resonance frequency of the drive mass. Two approaches have been considered for this purpose in the Ph.D. thesis of Federico MORELLI: extremum seeking [ZO11] and recursive-identification-based approach [LS83].

Since this dissertation focuses on the identification part of the project, we give a state of the art on the identification of MEMS gyroscopes and we derive the fourth and final research question of this dissertation in the next paragraph.

1.5.3 State of the art in MEMS gyroscope identification and problem formulation

There are some works in the literature that treat the identification of a MEMS gyroscope. The early work on system identification of MEMS gyroscopes can be found in [MGH01, CHM03]. In [MGH01], an open-loop identification method is developed to obtain a MIMO ARX model with chirp excitations. However, because of the slow dynamics of the MEMS, a very long experiment is needed in order to derive an accurate model. Therefore, in [CHM03], the sensor is put in closed-loop to fasten the dynamics and a closed-loop identification method is derived. Although this method provides a good model of the MEMS, the modeling works provided in [MGH01, CHM03] cannot be used in our case since the architecture of the MEMS in [CHM03, CHM03] considers angular vibrations of the masses while the one considered in this dissertation relies on linear vibrations of both masses.

For the type of MEMS gyroscope considered in the Next4MEMS project, most of the papers in the literature only cover the identification of the mechanical dynamics of both masses. In [PS02], three methods are provided to determine the cross-coupling stiffness coefficients. In the first one, a static analysis is made to determine these transfer functions. In the second method, a Lissajous-ellipsoid-based approach combined with a principal component analysis is provided to estimate the cross-damping coefficients. However, both algorithms rely on the assumption that there is no damping which is a restrictive assumption with respect to the real dynamics. Therefore, in the third approach, the sensor is put in closed-loop in order to guarantee sustainable oscillations (compensation of the damping effect). While these methods can provide an accurate estimate of the anisoelectricity effect, it is not enough to determine completely the cross-transfer functions since the nonproportional damping effect is not considered. The work in [PS04] provides a result in that matter since a frequency domain method is provided to identify accurately these transfer functions by estimating simultaneously both the anisoelectricity and nonproportional damping effects. The same purpose is followed in [FY11, FY12] but with an online-observer-based method in the open-loop and closed-loop configuration.

As already mentioned, all these aforementioned works do not consider the modeling of the parasitic coupling. The only work in that matter is the one in [AS05] where the authors propose a frequency-based approach which relies on the linearization of the square input nonlinearity for sinusoidal excitations in order to identify a part of the dynamics of the mechanical and parasite capacitive effects. This approach shows good results but the model might lose its validity at other frequencies due to the linearization. The model must indeed be valid for a wider frequency range for the design of the optimal controller. Moreover, the work in [AS05] does not consider possible cross-coupling effects

of the parasite capacitive coupling, which is the case of the MEMS gyroscope considered in this dissertation.

Finally, it is to be noted that there is no estimation of the noise measurement in the literature while it is important to have this estimation for the design of the optimal controller with the \mathcal{H}_∞ design approach. Therefore, for the design of the controller, we must choose a complete model structure and develop a methodology based on advanced identification techniques in order to accurately estimate all the dynamics of the MEMS gyroscope. This is formulated in the next research question:

Research Question 4

For the design of an optimal controller, how can we choose a complete model structure than the ones proposed in the literature for the identification of the three main dynamics of the MEMS gyroscope which are the mechanical dynamics, the parasite capacitive coupling and the measurement noise? Based on this complete model structure, how can we develop a methodology for the identification of the aforementioned dynamics without linearizing the square input nonlinearity?

To answer the first part of Research Question 4, we will develop the complete model structure of the MEMS gyroscope by using the first principles. We will see that this model structure is nonlinear and corresponds to the class of nonlinear systems that can be accurately identifiable within the Prediction Error framework. Therefore, we will consider this identification method to address the second part of Research Question 4.

As already mentioned, the MEMS gyroscope has two inputs and two outputs. Therefore, it can be seen as the collection of four subsystems with one input and one output as illustrated in Figure 1.4. Since this system is a MIMO one, we could consider a MIMO identification by exciting both inputs and measure both outputs in the same time. However, if we excite both inputs (u_1 and u_2 in Figure 1.4), some dynamics can be hidden by others in the measured outputs (y_1 and y_2 in Figure 1.4). It is, e.g., the case of the mechanical cross-coupling dynamics that can be hidden by the direct ones. Therefore, in order to avoid that some dynamics hide others, we will split the identification problem within the nonlinear complete model structure into the identification of four less complex nonlinear systems which are the subsystems illustrated in Figure 1.4. The idea is to excite only one input (e.g., the input u_1 in Figure 1.4) and set the other to 0 (e.g., the input u_2 in Figure 1.4). Then, in this case, we have to identify the two corresponding subsystems (Subsystems 1 and 3 in Figure 1.4), one for each output (y_1 and y_2 in Figure 1.4). We then excite u_2 by setting u_1 to 0 and perform the identification of the last two subsystems (Subsystems 2 and 4 in Figure 1.4).

We will furthermore see that the dynamics of these four subsystems are of the form given in (1.1). Hence, we will use the same trick proposed to address Research Question 3 by reformulating the identification of these subsystems into MISO LTI subsystems. Then, in order to guarantee an accurate model for the subsystems, we have to provide excitations such that the data are informative for the identification of the four subsystems. This data informativity analysis boils down to the data informativity study addressed by Research Question 3 and so we will exploit the developed results for the identification of the MEMS gyroscope. In the case where the data are informative, we can reach any level of uncertainties on the identified parameters by increasing the number of data. However, for some (unlucky) excitation choices that yield informative data, we might need to consider very long experiment (very large number of data) to ensure small uncertainties. In practice, increasing largely the number of data is not suitable. In order to get small uncertainties for a reasonable number of data, it is also important to design an excitation that increases the signal-to-noise ratio (SNR).

However, even if the data are informative and the SNR is increased for the identification of the four subsystems, the MISO LTI identification of the four subsystems are quite complex nonlinear optimization problems and so initial estimates of the transfer functions are needed to initialize properly these algorithms. A method will be proposed for this purpose which consists in isolating each dynamic and performing SISO identifications for each LTI transfer function involved in the nonlinear model structure, which are less complex nonlinear optimization problems.

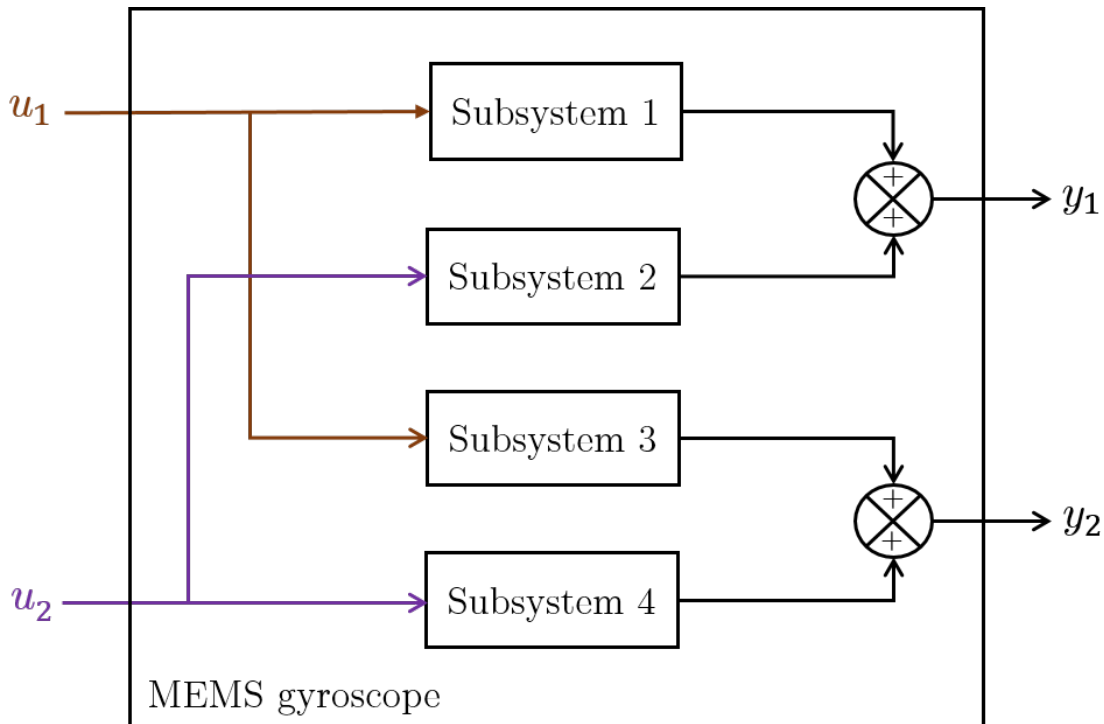


Figure 1.4: Description of the MEMS gyroscope (MIMO system) with two inputs u_1 and u_2 and two outputs y_1 and y_2 into four subsystems with one input and one output.

1.6 Contributions and outline of the thesis

The dissertation is divided into two parts, one theoretical part and another practical one. In the first (theoretical) part of the dissertation, we deal with Research Questions 1, 2 and 3.

- First, we introduce background information on Prediction Error identification in Chapter 2. The concept of data informativity is introduced for the open-loop and direct closed-loop identification.
- In Chapter 3, we study the data informativity with respect to natural MIMO generalizations of the classical SISO model structures FIR, ARX, ARMAX, OE and BJ. The approach that is considered in this chapter is inspired from the work in [GBM08] where the data informativity is connected to the persistency of a particular signal vector called regressor. We provide an extension of the input regressor persistency for the considered MIMO model structures. The results given in this chapter are based on the following conference and journal papers:

Kévin Colin, Xavier Bombois, Laurent Bako and Federico Morelli. Informativity: how to get just sufficiently rich for the Identification of MISO FIR Systems with Multisine Excitation? In *2019 18th European Control Conference (ECC)*, pages 351– 356, Napoli, Italy, June 2019. [doi:10.23919/ECC.2019.8795997](https://doi.org/10.23919/ECC.2019.8795997)

Kévin Colin, Xavier Bombois, Laurent Bako and Federico Morelli. Data informativity for the identification of MISO FIR system with filtered white noise excitation. In *2019 58th Conference on Decision and Control (CDC)*, pp. 1734-1739, Nice, France, December 2019. [doi:10.1109/CDC40024.2019.9029300](https://doi.org/10.1109/CDC40024.2019.9029300)

Kévin Colin, Xavier Bombois, Laurent Bako and Federico Morelli, Data informativity for the open-loop identification of MIMO systems in the prediction error framework, *Automatica*, Volume 117, 2020. [doi:10.1016/j.automatica.2020.109000](https://doi.org/10.1016/j.automatica.2020.109000)

- In Chapter 4, we give attention to the informativity of the direct closed-loop identification of MIMO systems. We derive conditions to verify the informativity for a given MIMO LTI controller and external excitation. Some comparisons with previous results [BGM10, YZ18] are provided. The material given in this chapter can be found in the following journal paper

Kévin Colin, Xavier Bombois, Laurent Bako and Federico Morelli. Closed-loop Identification of MIMO Systems in the Prediction Error Framework: Data informativity Analysis. *Automatica*, Volume 121, 2020. [doi:10.1016/j.automatica.2020.109171](https://doi.org/10.1016/j.automatica.2020.109171)

- In Chapter 5, we consider the first part of Research Question 3 by deriving results to verify the data informativity for a particular class of parallel Hammerstein systems given by $y(t) = G_{0,1}(z)u(t) + G_{0,2}(z)u^n(t) + v(t)$ for a given nonzero integer n and a given excitation u . The content of Chapter 5 is built from the following conference paper:

Kévin Colin, Xavier Bombois, Laurent Bako and Federico Morelli. Data informativity for the Identification of particular Parallel Hammerstein Systems. Accepted for presentation at *IFAC 2020*, 2020. [hal-02392113](https://hal.archives-ouvertes.fr/hal-02392113)

In the second (practical) part of the dissertation, we deal with Research Question 4. The real-life example of the MEMS gyroscope is then considered in Chapters 6 and 7.

- In Chapter 6, an introduction of the MEMS gyroscope is given. We build a complete model structure for the three main dynamics of the MEMS gyroscope (first part of Research Question 4). The control objectives are also given.

- In Chapter 7, we develop a method for the modeling of the main dynamics of the MEMS gyroscope (second part of Research Question 4). Among all the challenges tackled in this chapter, we will need to perform a data informativity verification. The results of the performances of the controller designed based on this model (performed by Fabrício SAGGIN) are also given in this chapter. A part of the results given in this chapter has been published in the following conference paper:

Kévin Colin, Fabrício Saggin, Christophe Le Blanc, Xavier Bombois, Anton Kornienko and Gérard Scorletti. Identification-Based Approach for Electrical Coupling Compensation in a MEMS Gyroscope. In *2019 IEEE International Symposium on Inertial Sensors and Systems (INERTIAL)*, pp. 1-4, Naples (Florida), USA, April 2019. [doi:10.1109/ISISS.2019.8739573](https://doi.org/10.1109/ISISS.2019.8739573)

In the next chapter, we introduce background information on Prediction Error identification and we will define the data informativity property that will be studied in Chapters 3, 4 and 5.

Theoretical work on the data informativity

Prediction Error Identification

Contents

2.1	Introduction of Chapter 2	21
2.2	Discrete-time quasi-stationary signals	22
2.3	Prediction Error identification of MIMO systems in open-loop	23
2.4	Consistency and Informativity in closed-loop	26
2.5	MIMO model structure considered in this dissertation	28
2.6	Previous results on MIMO data informativity	30
2.7	Persistency of a vector of quasi-stationary signals	32
2.8	Conclusion of Chapter 2 and reformulation of Research Question 1	33

2.1 Introduction of Chapter 2

In Chapters 2, 3 and 4, we will focus on addressing Research Questions 1 and 2 given in Chapter 1, i.e., the development of conditions to verify the data informativity for both the open-loop and direct closed-loop identification of MIMO LTI systems within the Prediction Error framework. In order to address this problem, we first need to introduce some background notions on Prediction Error identification.

As explained in Chapter 1, each identification method such as Prediction Error is based on three ingredients: a model structure, a set of experimental data and an identification criterion. Therefore, in this chapter, we present each ingredient that we will consider for Research Questions 1 and 2:

- **Model structure:** The model structures that will be considered in this dissertation are natural generalizations of the classical SISO LTI discrete-time model

structures that are commonly-used in Prediction Error: FIR, ARX, ARMAX, OE and BJ.

- **Data Set:** In the Prediction Error method, the two types of signals that are the most used are the following ones: multisine or filtered white noise. Consequently, we will consider both aforementioned types of excitations in this dissertation. By applying one of these types of excitation to the to-be-identified system, the data set that will be considered is made up of input/output data collected on this system which can be put in the open-loop or the closed-loop configuration.
- **Identification criterion:** the identification criterion that will be considered in this dissertation will be the least-square criterion on the prediction error, i.e., a criterion minimizing the power of the prediction error.

This chapter is constructed as follows. First, we give some notions on the discrete-time quasi-stationary signals in Section 2.2. Then in Section 2.3, we present the Prediction Error framework and more precisely the least-square Prediction Error estimator whose consistency must be ensured in order to guarantee an accurate estimate of the system. For this purpose, the notion of data informativity is first introduced for the open-loop identification. In Section 2.4, this definition is extended for the direct closed-loop identification. In Section 2.5, the MIMO model structures that will be considered for the data informativity analysis are presented. As aforementioned, we will consider natural generalizations of the classical SISO model structures that are commonly-used in Prediction Error: FIR, ARX, ARMAX, OE and BJ. The results on data informativity in the literature are provided in Section 2.6. Finally, the notion of persistency of quasi-stationary signal vectors is presented in Section 2.7, since it will be needed for the derivation of the results to address Research Questions 1 and 2.

2.2 Discrete-time quasi-stationary signals

In this dissertation, we will consider discrete-time scalar signals and signal vectors. This choice is motivated by the fact that the signals are sampled in real-life applications. We will assume that the signals are uniformly sampled at a constant sampling time T_s . Consequently, a discrete-time signal vector \mathbf{x} can be seen as a collection of values $\{\mathbf{x}(tT_s)\}_{t \in \mathbb{N}^*}$. However, in this dissertation, we will remove the information of the sampling time by normalizing the time, i.e., \mathbf{x} can be seen as a collection of values $\{\mathbf{x}(t)\}_{t \in \mathbb{N}^*}$.

We will use the class of quasi-stationary signals in this dissertation. To define this signal class, let us denote by \bar{E} the following operator on a discrete-time stochastic or deterministic signal vector \mathbf{x}

$$\bar{E}[\mathbf{x}(t)] = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{t=1}^N E[\mathbf{x}(t)] \quad (2.1)$$

assuming that the limit exists and the quantity is well-defined. The operator E is the expectation operator and the operator \bar{E} is often called full expectation operator in the literature. Based on \bar{E} , we can define the set of quasi-stationary signals:

Definition 2.1. *A discrete-time vector \mathbf{x} of n complex- (or real-) valued signals x_1, \dots, x_n is said to be quasi-stationary if the following conditions are met*

- (a) $\bar{E}[\mathbf{x}(t)]$ is well-defined and is equal to a constant.
- (b) the following matrix is well-defined and is a function that depends only on $\tau \in \mathbb{Z}$

$$\bar{E}[\mathbf{x}(t)\mathbf{x}^T(t-\tau)] = \begin{pmatrix} \bar{E}[x_1(t)x_1(t-\tau)] & \bar{E}[x_1(t)x_2(t-\tau)] & \cdots & \bar{E}[x_1(t)x_n(t-\tau)] \\ \bar{E}[x_2(t)x_1(t-\tau)] & \bar{E}[x_2(t)x_2(t-\tau)] & \cdots & \bar{E}[x_2(t)x_n(t-\tau)] \\ \vdots & \vdots & \ddots & \vdots \\ \bar{E}[x_n(t)x_1(t-\tau)] & \bar{E}[x_n(t)x_2(t-\tau)] & \cdots & \bar{E}[x_n(t)x_n(t-\tau)] \end{pmatrix}$$

The set of quasi-stationary signals contains deterministic periodic signals (e.g., sum of sinusoids or multisine) and stationary stochastic signals (e.g., filtered white noise). In this dissertation, we will consider both aforementioned types of quasi-stationary signals.

2.3 Prediction Error identification of MIMO systems in open-loop

Consider a MIMO system \mathcal{S} with a quasi-stationary input vector $\mathbf{u} \in \mathbb{R}^{n_u}$ and an output vector $\mathbf{y} \in \mathbb{R}^{n_y}$, described by

$$\mathcal{S}: \quad \mathbf{y}(t) = \mathbf{G}_0(z)\mathbf{u}(t) + \mathbf{H}_0(z)\mathbf{e}(t) \quad (2.2)$$

where $\mathbf{G}_0(z)$ is a stable matrix of (rational) transfer functions of dimension $n_y \times n_u$, $\mathbf{H}_0(z)$ is a stable, inversely-stable and monic matrix¹ of (rational) transfer functions of dimension $n_y \times n_y$ and \mathbf{e} is a vector consisting of n_y white noise signals such that $\bar{E}[\mathbf{e}(t)\mathbf{e}^T(t)] = \mathbf{\Sigma}_0 \succ 0$. We will also make the following additional assumption on $\mathbf{H}_0(z)$:

Assumption 2.1. *The transfer function matrix $\mathbf{H}_0(z)$ will be assumed to be diagonal, i.e.,*

$$\mathbf{H}_0(z) = \text{diag}(H_{0,1}(z), \dots, H_{0,n_y}(z))$$

where the notation *diag* has been introduced in the section *Notations* (see page *xx*).

¹i.e. \mathbf{H}_0 and \mathbf{H}_0^{-1} are stable and $\mathbf{H}_0(z = \infty) = \mathbf{I}_{n_y}$

Due to Assumption 2.1, it is interesting to note that the MIMO system (2.2) can also be seen as the collection of n_y MISO systems:

$$\mathcal{S}_i : \quad y_i(t) = \mathbf{G}_{0,i}(z)\mathbf{u}(t) + H_{0,i}(z)e_i(t) \quad (i = 1, \dots, n_y) \quad (2.3)$$

where $\mathbf{G}_{0,i}(z)$ is the i^{th} row of \mathbf{G}_0 . The different channels y_i can however be connected by the eventual correlation between the white noises in \mathbf{e} (if $\mathbf{\Sigma}_0$ is not diagonal).

When the MIMO system \mathcal{S} in (2.2) is operated in open-loop, the excitation signal that will be used for identification purposes will be the input vector \mathbf{u} which is assumed to be independent of \mathbf{e} . Moreover, we will consider two different types of quasi-stationary excitation \mathbf{u} .

In the first type, each entry u_k of \mathbf{u} ($k = 1, \dots, n_u$) is a multisine consisting of sinusoids at s different frequencies $\omega_l \in]0, \pi[$ ($l = 1, \dots, s$), i.e.,

$$u_k(t) = \sum_{l=1}^s \lambda_{kl} \cos(\omega_l t + \psi_{kl}) \quad k = 1, \dots, n_u \quad (2.4)$$

where λ_{kl} and ψ_{kl} are respectively the amplitude and phase shift of the sinusoid at the frequencies ω_l . Note that λ_{kl} can be zero for some value(s) of k ($k = 1, \dots, n_u$), but, for each $l = 1, \dots, s$, there exists (at least) a value of k for which $\lambda_{kl} \neq 0$. For further reference, we will denote by s_k ($k = 1, \dots, n_u$) the number of sinusoids for which the amplitude λ_{kl} in the expression (2.4) for u_k is nonzero ($s_k \leq s \forall k$).

In the second type of input \mathbf{u} , \mathbf{u} is generated as $\mathbf{u}(t) = \mathbf{M}(z)\mathbf{v}(t)$ via a stable transfer function matrix $\mathbf{M}(z) = (M_{kq}(z))_{(k,q) \in \llbracket 1, n_u \rrbracket \times \llbracket 1, f \rrbracket}$ and a vector $\mathbf{v} = (v_1, \dots, v_f)^T$ containing f independent white noise signals v_q ($q = 1, \dots, f$). In other words, each entry u_k of \mathbf{u} is given by:

$$u_k(t) = \sum_{q=1}^f M_{kq}(z)v_q(t) \quad k = 1, \dots, n_u. \quad (2.5)$$

Note that f is not necessarily equal to n_u and that $M_{kq}(z)$ can be identically zero for some value(s) of k ($k = 1, \dots, n_u$), but, for each $q = 1, \dots, f$, there exists (at least) a value of k for which $M_{kq}(z) \neq 0$.

To identify \mathcal{S} , we consider a full-order parametrized model structure

$$\mathcal{M} = \{(\mathbf{G}(z, \boldsymbol{\theta}), \mathbf{H}(z, \boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \mathcal{D}_\theta\}$$

where $\boldsymbol{\theta} \in \mathcal{D}_\theta \subset \mathbb{R}^n$ is a parameter vector. This model structure is full-order if $\exists \boldsymbol{\theta}_0 \in \mathcal{D}_\theta$ (called the true parameter vector) such that $(\mathbf{G}(z, \boldsymbol{\theta}_0), \mathbf{H}(z, \boldsymbol{\theta}_0)) = (\mathbf{G}_0(z), \mathbf{H}_0(z))$. The set \mathcal{D}_θ restricts the parameter vector $\boldsymbol{\theta}$ to those values for which $\mathbf{G}(z, \boldsymbol{\theta})$ is stable and

$\mathbf{H}(z, \boldsymbol{\theta})$ is stable, inversely stable and monic. The matrix $\mathbf{H}(z, \boldsymbol{\theta})$ is moreover diagonal because of Assumption 2.1.

We will further suppose that this model structure is globally identifiable at the true parameter vector $\boldsymbol{\theta}_0$, i.e., for all $\boldsymbol{\theta} \in \mathcal{D}_\theta$,

$$(\mathbf{G}(z, \boldsymbol{\theta}), \mathbf{H}(z, \boldsymbol{\theta})) = (\mathbf{G}(z, \boldsymbol{\theta}_0), \mathbf{H}(z, \boldsymbol{\theta}_0)) \implies \boldsymbol{\theta} = \boldsymbol{\theta}_0$$

We can define, based on \mathcal{M} , the one-step ahead predictor $\hat{\mathbf{y}}(t, \boldsymbol{\theta})$ for the output $\mathbf{y}(t)$ at time t :

$$\hat{\mathbf{y}}(t, \boldsymbol{\theta}) = \mathbf{W}_y(z, \boldsymbol{\theta})\mathbf{y}(t) + \mathbf{W}_u(z, \boldsymbol{\theta})\mathbf{u}(t) = \mathbf{W}(z, \boldsymbol{\theta})\mathbf{x}(t) \quad (2.6)$$

where

$$\mathbf{W}_y(z, \boldsymbol{\theta}) = \mathbf{I}_{n_y} - \mathbf{H}^{-1}(z, \boldsymbol{\theta}) \quad (2.7)$$

$$\mathbf{W}_u(z, \boldsymbol{\theta}) = \mathbf{H}^{-1}(z, \boldsymbol{\theta})\mathbf{G}(z, \boldsymbol{\theta}) \quad (2.8)$$

$$\mathbf{W}(z, \boldsymbol{\theta}) = (\mathbf{W}_y(z, \boldsymbol{\theta}), \mathbf{W}_u(z, \boldsymbol{\theta}))$$

$$\mathbf{x}(t) = \begin{pmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{pmatrix}$$

From a set of data \mathcal{Z}^N collected on the true system \mathcal{S} and defined by

$$\mathcal{Z}^N = \left\{ \mathbf{x}(t) = (\mathbf{y}^T(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, N \right\}$$

we compute the optimal parameter vector $\hat{\boldsymbol{\theta}}_N$ that provides the minimum value of the following least-square criterion $V_N(\boldsymbol{\theta})$ on the prediction error $\boldsymbol{\epsilon}(t, \boldsymbol{\theta}) = \mathbf{y}(t) - \hat{\mathbf{y}}(t, \boldsymbol{\theta})$:

$$\hat{\boldsymbol{\theta}}_N = \arg \min_{\boldsymbol{\theta} \in \mathcal{D}_\theta} V_N(\boldsymbol{\theta}, \mathcal{Z}^N) \quad (2.9)$$

$$V_N(\boldsymbol{\theta}, \mathcal{Z}^N) = \frac{1}{N} \sum_{t=1}^N \boldsymbol{\epsilon}^T(t, \boldsymbol{\theta}) \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\epsilon}(t, \boldsymbol{\theta})$$

where $\boldsymbol{\Sigma}_0$ is supposed known for simplicity (it can however be estimated together with $\hat{\boldsymbol{\theta}}_N$ (see, e.g., [Lju99, Chapter 15])). In most cases, the optimization problem in (2.9) is a nonlinear one which can raise some local minima issues. Several optimization algorithms can be considered such as gradient descent algorithms, Gauss-Newton algorithm, trust region reflective algorithm, etc. More details can be found in [Lju99, Chapter 10].

We want $\hat{\boldsymbol{\theta}}_N$ to be a consistent estimate for $\boldsymbol{\theta}_0$, i.e., $\hat{\boldsymbol{\theta}}_N$ must converge to $\boldsymbol{\theta}_0$ with probability 1 when $N \rightarrow +\infty$. This property is indeed crucial since, with the consistency property, we ensure that the modeling error $\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0$ can be made arbitrarily small by choosing N sufficiently high.

In [Lju99, Chapter 8], it has been proven that the consistency is guaranteed if the true parameter vector $\boldsymbol{\theta}_0$ is the unique minimizer of the asymptotic criterion given by $\bar{E} [\boldsymbol{\epsilon}^T(t, \boldsymbol{\theta}) \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\epsilon}(t, \boldsymbol{\theta})]$ or, equivalently, $\boldsymbol{\theta}_0$ is the unique parameter vector $\boldsymbol{\theta}$ for which $\boldsymbol{\epsilon}(t, \boldsymbol{\theta}) = \boldsymbol{e}(t)$.

To guarantee this property, the input vector \boldsymbol{u} used to excite \mathcal{S} must yield informative data w.r.t. the model structure \mathcal{M} .

Definition 2.2 (Data Informativity in open-loop [BGM10, Lju99]). *Consider the framework defined above. Suppose that a quasi-stationary input vector $\boldsymbol{u}(t)$ is applied to the true system \mathcal{S} from $t = 1$ till $t = +\infty$ yielding the infinite data set \mathcal{Z}^∞ :*

$$\mathcal{Z}^\infty = \left\{ \boldsymbol{x}(t) = (\boldsymbol{y}^T(t), \boldsymbol{u}^T(t))^T \mid t = 1, \dots, +\infty \right\} \quad (2.10)$$

Consider a model structure \mathcal{M} yielding the predictor $\hat{\boldsymbol{y}}(t, \boldsymbol{\theta}) = \mathbf{W}(z, \boldsymbol{\theta})\boldsymbol{x}(t)$. Let us also define the set

$$\boldsymbol{\Delta}_{\mathbf{W}} = \left\{ \Delta \mathbf{W}(z) = \mathbf{W}(z, \boldsymbol{\theta}') - \mathbf{W}(z, \boldsymbol{\theta}'') \mid \boldsymbol{\theta}' \text{ and } \boldsymbol{\theta}'' \text{ in } \mathcal{D}_\theta \right\} \quad (2.11)$$

The data set \mathcal{Z}^∞ is said to be informative w.r.t. the model structure \mathcal{M} when, for all $\Delta \mathbf{W}(z) \in \boldsymbol{\Delta}_{\mathbf{W}}$, we have

$$\bar{E} [|\Delta \mathbf{W}(z)\boldsymbol{x}(t)|^2] = 0 \implies \Delta \mathbf{W}(z) \equiv \mathbf{0}_{n_y \times (n_y + n_u)} \quad (2.12)$$

where $\Delta \mathbf{W}(z) \equiv \mathbf{0}_{n_y \times (n_y + n_u)}$ means that $\Delta \mathbf{W}(e^{j\omega}) = \mathbf{0}_{n_y \times (n_y + n_u)}$ for almost all ω and where \bar{E} is the operator defined in (2.1).

From Definition 2.2, if the input vector \boldsymbol{u} yields informative data, we have that $\boldsymbol{\epsilon}(t, \boldsymbol{\theta}) = \boldsymbol{e}(t)$ if and only if $(\mathbf{G}(z, \boldsymbol{\theta}), \mathbf{H}(z, \boldsymbol{\theta})) = (\mathbf{G}(z, \boldsymbol{\theta}_0), \mathbf{H}(z, \boldsymbol{\theta}_0))$. Since we suppose that the model structure is globally identifiable at $\boldsymbol{\theta}_0$, the latter property implies that $\boldsymbol{\theta}_0$ is indeed the unique parameter vector $\boldsymbol{\theta}$ for which $\boldsymbol{\epsilon}(t, \boldsymbol{\theta}) = \boldsymbol{e}(t)$. In other words, the combination of the identifiability and data informativity conditions implies the consistency of $\hat{\boldsymbol{\theta}}_N$.

In Chapter 3, we will therefore derive conditions to ensure the data informativity for input vectors \boldsymbol{u} of the types (2.4)-(2.5).

2.4 Consistency and Informativity in closed-loop

In the previous section, we considered open-loop identification and data informativity/consistency in this context. We will also consider the case where an estimate of the true parameter vector $\boldsymbol{\theta}_0$ will be derived using direct closed-loop identification. For this purpose, we suppose that the MIMO system \mathcal{S} introduced in the previous section is operated in closed-loop with a known stabilizing MIMO LTI controller $\mathbf{K}(z)$ (see

Figure 2.1). In this closed-loop configuration, the excitation signal that will be used for identification purposes is the signal vector \mathbf{r} such that \mathbf{u} is given by

$$\mathbf{u}(t) = -\mathbf{K}(z)\mathbf{y}(t) + \mathbf{r}(t) \quad (2.13)$$

The signal vectors \mathbf{r} and \mathbf{e} are assumed to be independent and we will suppose the same type of excitation as in the open-loop case, i.e., multisine and filtered white noise excitation. We just repeat the equations for further references:

$$r_k(t) = \sum_{l=1}^s \lambda_{kl} \cos(\omega_l t + \psi_{kl}) \quad k = 1, \dots, n_u. \quad (2.14)$$

$$r_k(t) = \sum_{q=1}^f M_{kq}(z)v_q(t) \quad k = 1, \dots, n_u. \quad (2.15)$$

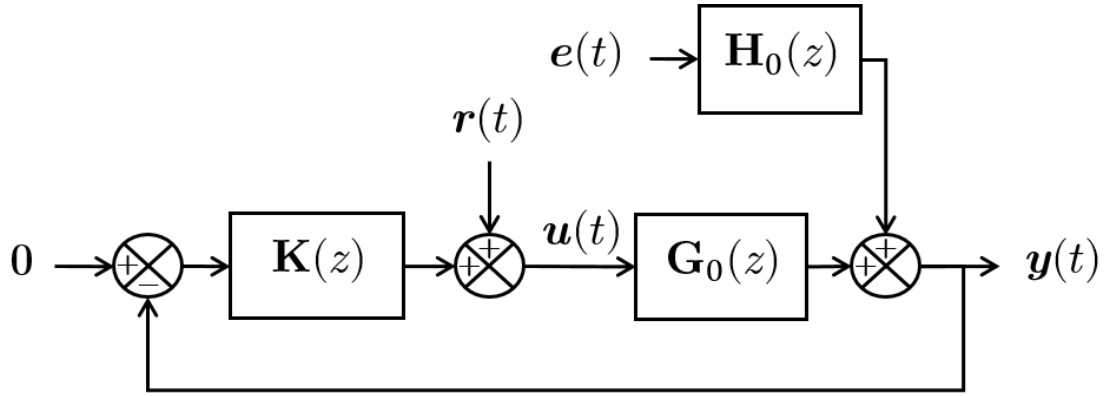


Figure 2.1: Scheme of the closed-loop.

If the excitation vector \mathbf{r} is applied to the closed-loop system represented in Figure 2.1 from $t = 1$ till $t = N$, we can collect the input-output data set

$$\mathcal{Z}_{cl}^N = \left\{ \mathbf{x}(t) = (\mathbf{y}^T(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, N \right\}$$

and identify an estimate $\hat{\boldsymbol{\theta}}_N$ of $\boldsymbol{\theta}_0$ using the same prediction error criterion (2.9) as in open-loop and the same full-order model structure \mathcal{M} . If we add the assumption that all the entries of $\mathbf{G}_0(z)\mathbf{K}(z)$ contain at least a delay, the consistency of this closed-loop estimate $\hat{\boldsymbol{\theta}}_N$ is also guaranteed under the condition that \mathcal{M} is globally identifiable at $\boldsymbol{\theta}_0$ and that the excitation vector \mathbf{r} yields informative data. The definition of data informativity in closed-loop is almost equivalent to the one in open-loop. The only difference lies in the way the infinite data set is collected.

Definition 2.3 (Data Informativity in closed-loop [BGM10, Lju99]). *Consider that, in the framework defined above, an excitation vector \mathbf{r} is applied to the closed-loop system*

represented in Figure 2.1 from $t = 1$ till $t = +\infty$ and suppose that the corresponding infinite input-output data set \mathcal{Z}_{cl}^∞ is collected:

$$\mathcal{Z}_{cl}^\infty = \left\{ \mathbf{x}(t) = (\mathbf{y}^T(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, +\infty \right\} \quad (2.16)$$

Consider a model structure \mathcal{M} yielding the predictor $\hat{\mathbf{y}}(t, \boldsymbol{\theta}) = \mathbf{W}(z, \boldsymbol{\theta})\mathbf{x}(t)$. The data set \mathcal{Z}^∞ is said to be informative w.r.t. the model structure \mathcal{M} if (2.12) holds for all $\Delta \mathbf{W}(z) \in \Delta \mathbf{W}$ ($\Delta \mathbf{W}$ and $\Delta \mathbf{W}(z)$ are defined as in Definition 2.2).

2.5 MIMO model structure considered in this dissertation

We will derive conditions for data informativity for the identification of the true system \mathcal{S} in open-loop (Chapter 3) and in closed-loop (Chapter 4). These conditions will depend on the parametrization of the model structure \mathcal{M} used for the identification. The model structures \mathcal{M} that will be considered in this manuscript are natural generalizations of the classical model structures used in SISO identification, i.e., the ARX, FIR, ARMAX, BJ and OE model structures. These extensions are described in this section. For this purpose, let us decompose the parameter vector $\boldsymbol{\theta}$ as follows

$$\boldsymbol{\theta} = \begin{pmatrix} \tilde{\boldsymbol{\theta}} \\ \boldsymbol{\eta} \end{pmatrix}$$

where $\tilde{\boldsymbol{\theta}}$ is a vector made up by the parameters uniquely found in $\mathbf{G}(z, \boldsymbol{\theta})$ and $\boldsymbol{\eta}$ contains the rest of the parameters.

ARMAX: The ARMAX model structures are described by

$$\begin{cases} \mathbf{G}(z, \boldsymbol{\theta}) &= \mathbf{A}^{-1}(z, \boldsymbol{\eta})\mathbf{B}(z, \tilde{\boldsymbol{\theta}}) \\ \mathbf{H}(z, \boldsymbol{\theta}) &= \mathbf{A}^{-1}(z, \boldsymbol{\eta})\mathbf{C}(z, \boldsymbol{\eta}) \end{cases} \quad (2.17)$$

where $\mathbf{A}(z, \boldsymbol{\eta})$ and $\mathbf{C}(z, \boldsymbol{\eta})$ are monic and *diagonal*² polynomial matrices of dimension $n_y \times n_y$ and $\mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ is a polynomial matrix of dimension $n_y \times n_u$. The ARX and FIR model structures are special cases of the ARMAX ones with $\mathbf{C}(z, \boldsymbol{\eta}) = \mathbf{I}_{n_y}$ in the ARX case and with $\mathbf{C}(z, \boldsymbol{\eta}) = \mathbf{A}(z, \boldsymbol{\eta}) = \mathbf{I}_{n_y}$ in the FIR case.

The matrix $\mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ is made up of $n_y n_u$ independently parametrized polynomials $B_{ik}(z, \tilde{\boldsymbol{\theta}}) = B_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) = z^{-\rho_{ik}} \hat{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ ($i = 1, \dots, n_y, k = 1, \dots, n_u$) where the delays ρ_{ik} can be all different and $\hat{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ is a polynomial with a zero-delay parametrized by a parameter vector $\tilde{\boldsymbol{\theta}}_{ik}$. The matrix $\mathbf{H}(z, \boldsymbol{\eta})$ is composed of n_y transfer functions $H_i(z, \boldsymbol{\eta})$ which are independently parametrized with a parameter vector $\boldsymbol{\eta}_i$, i.e., $H_i(z, \boldsymbol{\eta}) = H_i(z, \boldsymbol{\eta}_i) = C_i(z, \boldsymbol{\eta}_i)/A_i(z, \boldsymbol{\eta}_i)$ where $C_i(z, \boldsymbol{\eta}_i)$ and $A_i(z, \boldsymbol{\eta}_i)$ are monic polynomials. The coefficients of the polynomial \hat{B}_{ik} are the parameters in $\tilde{\boldsymbol{\theta}}_{ik}$ while the ones

²Because the matrix $\mathbf{H}(z, \boldsymbol{\theta})$ is considered to be diagonal.

in the polynomials $A_i(z, \boldsymbol{\eta}_i)$ and $C_i(z, \boldsymbol{\eta}_i)$ are the ones in $\boldsymbol{\eta}_i$. Consequently, we have that:

$$\begin{aligned}\mathring{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) &= \tilde{\theta}_{ik,1} + \sum_{m=1}^{\deg(\mathring{B}_{ik})} \tilde{\theta}_{ik,(m+1)} z^{-m} \\ C_i(z, \boldsymbol{\eta}_i) &= 1 + \sum_{m=1}^{\deg(C_i)} \eta_{i,m} z^{-m} \\ A_i(z, \boldsymbol{\eta}_i) &= 1 + \sum_{m=1}^{\deg(A_i)} \eta_{i,(m+\deg(C_i))} z^{-m}\end{aligned}$$

where $\tilde{\theta}_{ik,m}$ denotes the m^{th} entry of $\tilde{\boldsymbol{\theta}}_{ik}$ and $\eta_{i,m}$ the m^{th} entry of $\boldsymbol{\eta}_i$. The number of parameters to be identified in \mathring{B}_{ik} , C_i and A_i is thus equal to $\deg(\mathring{B}_{ik}) + 1$, $\deg(C_i)$ and $\deg(A_i)$ respectively.

BJ: In the BJ model structures, the matrices $\mathbf{G}(z, \boldsymbol{\theta})$ and $\mathbf{H}(z, \boldsymbol{\theta})$ are independently parametrized. They are given by

$$\begin{cases} \mathbf{G}(z, \boldsymbol{\theta}) &= \mathbf{G}(z, \tilde{\boldsymbol{\theta}}) \\ \mathbf{H}(z, \boldsymbol{\theta}) &= \mathbf{H}(z, \boldsymbol{\eta}) \end{cases} \quad (2.18)$$

The OE model structure is a particular case of the BJ one where $\mathbf{H}(z, \boldsymbol{\eta}) = \mathbf{I}_{n_y}$.

The matrix $\mathbf{G}(z, \tilde{\boldsymbol{\theta}})$ is made up of $n_y n_u$ independently parametrized transfer functions $G_{ik}(z, \tilde{\boldsymbol{\theta}}) = G_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) = z^{-\rho_{ik}} \mathring{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) / F_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ where $\mathring{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ and $F_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ are polynomials parametrized by a parameter vector $\tilde{\boldsymbol{\theta}}_{ik}$. Moreover, $F_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ is monic and $\mathring{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ has a zero-delay. The diagonal matrix $H(z, \boldsymbol{\eta})$ is composed of n_y independently parametrized transfer functions $H_i(z, \boldsymbol{\eta}_i)$ with $H_i(z, \boldsymbol{\eta}_i) = C_i(z, \boldsymbol{\eta}_i) / D_i(z, \boldsymbol{\eta}_i)$ where $C_i(z, \boldsymbol{\eta}_i)$ and $D_i(z, \boldsymbol{\eta}_i)$ are monic polynomials parametrized by a parameter vector $\boldsymbol{\eta}_i$. The coefficients of the polynomials \mathring{B}_{ik} and F_{ik} are the parameters in $\tilde{\boldsymbol{\theta}}_{ik}$ while the coefficients of the polynomials $C_i(z, \boldsymbol{\eta}_i)$ and $D_i(z, \boldsymbol{\eta}_i)$ are the ones in $\boldsymbol{\eta}_i$. Consequently, we have that:

$$\begin{aligned}\mathring{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) &= \tilde{\theta}_{ik,1} + \sum_{m=1}^{\deg(\mathring{B}_{ik})} \tilde{\theta}_{ik,(m+1)} z^{-m} \\ F_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) &= 1 + \sum_{m=1}^{\deg(F_{ik})} \tilde{\theta}_{ik,(m+\deg(\mathring{B}_{ik})+1)} z^{-m} \\ C_i(z, \boldsymbol{\eta}_i) &= 1 + \sum_{m=1}^{\deg(C_i)} \eta_{i,m} z^{-m} \\ D_i(z, \boldsymbol{\eta}_i) &= 1 + \sum_{m=1}^{\deg(D_i)} \eta_{i,(m+\deg(C_i))} z^{-m}\end{aligned}$$

where $\tilde{\theta}_{ik,m}$ denotes the m^{th} entry of $\tilde{\theta}_{ik}$ and $\eta_{i,m}$ the m^{th} entry of $\boldsymbol{\eta}_i$. The number of parameters to be identified in \tilde{B}_{ik} , F_{ik} , D_i and C_i is thus equal to $\deg(\tilde{B}_{ik}) + 1$, $\deg(F_{ik})$, $\deg(D_i)$ and $\deg(C_i)$ respectively.

For all these model structures, the parameter vector $\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{\tilde{n}}$ is the concatenation of all $\tilde{\theta}_{ik}$ ($i = 1, \dots, n_y$, $k = 1, \dots, n_u$), i.e., $\tilde{\boldsymbol{\theta}} = (\tilde{\theta}_{11}^T, \tilde{\theta}_{12}^T, \dots, \tilde{\theta}_{1n_u}^T, \dots, \dots, \tilde{\theta}_{n_y 1}^T, \tilde{\theta}_{n_y 2}^T, \dots, \tilde{\theta}_{n_y n_u}^T)^T$. The parameter vector $\boldsymbol{\eta} \in \mathbb{R}^{n_\eta}$ is also the concatenation of all $\boldsymbol{\eta}_i$ ($i = 1, \dots, n_y$), i.e., $\boldsymbol{\eta} = (\boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T, \dots, \boldsymbol{\eta}_{n_y}^T)^T$.

The MIMO FIR/ARX model structures are globally identifiable at any $\boldsymbol{\theta}_0 \in \mathcal{D}_\theta$. For the other MIMO model structures, they are globally identifiable at the true parameter vector $\boldsymbol{\theta}_0$ if $\boldsymbol{\theta}_0$ does not lead to a pole/zero cancellation. For MIMO OE model structures, the proof of this result can be found in [SS89, Complement C6.1].

Index notations: In the sequel, we will use the index $i \in \llbracket 1, n_y \rrbracket$ to specify the entries y_i of \mathbf{y} while $k \in \llbracket 1, n_u \rrbracket$ will be used to specify the entries u_k of \mathbf{u} .

2.6 Previous results on MIMO data informativity

As mentioned in the introduction (Chapter 1), there are very few results on MIMO data informativity in the literature. Among these existing results, we have the following lemmas (one for the open-loop situation and one for the closed-loop situation)

Lemma 2.1. *Consider Definition 2.2 and the infinite data set \mathcal{Z}^∞ corresponding to an open-loop experiment with excitation vector \mathbf{u} . If the power spectrum density $\boldsymbol{\Phi}_{\mathbf{u}}(\omega)$ of \mathbf{u} is strictly positive definite at almost all frequencies ω in the interval $] -\pi, \pi]$, then the data set \mathcal{Z}^∞ is informative with respect to any model structure, irrespectively of its type and of its complexity.*

Lemma 2.2 ([BGM10]). *Consider Definition 2.3 and the infinite data set \mathcal{Z}_{cl}^∞ corresponding to an open-loop experiment with excitation vector \mathbf{r} . If the power spectrum density $\boldsymbol{\Phi}_{\mathbf{r}}(\omega)$ of \mathbf{r} is strictly positive definite at almost all frequencies ω in the interval $] -\pi, \pi]$, then the data set \mathcal{Z}_{cl}^∞ is informative with respect to any model structure, irrespectively of its type and of its complexity.*

The proof of Lemma 2.2 is given in [BGM10] and Lemma 2.1 is a special case of Lemma 2.2 when $\mathbf{K} = \mathbf{0}_{n_u \times n_y}$ and $\mathbf{r} = \mathbf{u}$. A formal proof of these lemmas will also be deduced in the next chapters.

Based on the above lemmas, excitation vectors \mathbf{u} (resp. \mathbf{r}) such that $\boldsymbol{\Phi}_{\mathbf{u}}(\omega) \succ 0$ for almost all frequencies ω in the interval $] -\pi, \pi]$ (resp. $\boldsymbol{\Phi}_{\mathbf{r}}(\omega) \succ 0$ for almost all frequencies ω in the interval $] -\pi, \pi]$) always yield informative data and thus consistent estimates $\hat{\boldsymbol{\theta}}_N$.

Within the class of excitations we consider, this property is satisfied for the filtered white noise excitation such that $f \geq n_u$ and $\mathbf{M}(e^{j\omega})$ is full row rank for almost all frequencies ω in the interval $] -\pi, \pi]$. This is summarized in the following theorem restricting attention to the open-loop case (the closed-loop case is equivalent)

Theorem 2.1. *Consider Definition 2.2 and the infinite data set \mathcal{Z}^∞ given in (2.10) corresponding to an open-loop experiment with an excitation vector \mathbf{u} . If \mathbf{u} is of the type (2.5) with $f \geq n_u$ and $\text{rank}(\mathbf{M}(e^{j\omega})) = n_u$ for almost all frequencies ω in the interval $] -\pi, \pi]$, then the data set \mathcal{Z}^∞ is informative with respect to any model structure, irrespectively of its type and of its complexity.*

Proof. When \mathbf{u} is of the type (2.5), its power spectrum $\Phi_{\mathbf{u}}(\omega)$ is given by $\Phi_{\mathbf{u}}(\omega) = \mathbf{M}(e^{j\omega})\Sigma_v\mathbf{M}^*(e^{j\omega})$ where $\Sigma_v = E[\mathbf{v}(t)\mathbf{v}^T(t)]$ is a positive definite diagonal matrix. Under the condition on $\mathbf{M}(e^{j\omega})$ in the statement of the theorem, we observe that $\Phi_{\mathbf{u}}(\omega)$ is thus positive definite at almost all frequencies ω in the interval $] -\pi, \pi]$. As shown in Lemma 2.1, an input vector \mathbf{u} having this property will yield informative data with respect to any model structure, irrespectively of its type and of its complexity. This completes the proof. \square

A special case of the above theorem is the case where $f = n_u$ and where $\mathbf{M}(z)$ is a diagonal matrix since, in this case, $\mathbf{M}(e^{j\omega})$ will always be full row rank at all frequencies ω in the interval $] -\pi, \pi]$. This particular case for the excitation is quite classical in practice: it indeed corresponds to the case where all channels are excited with independent filtered white noises. However, in some cases, experiment design limitations may impede such classical choice and we may be faced with filtered white noise excitation with $f < n_u$. In this case, $\det(\Phi_{\mathbf{u}}(\omega)) = 0$ at all frequencies ω in the interval $] -\pi, \pi]$ and we cannot use Lemma 2.1 to prove the data informativity. However, we will show in the sequel that many of these particular excitations may lead to data informativity.

The multisine excitation case is also a case where Lemmas 2.1 and 2.2 cannot be of any help. Indeed, all multisine excitations of the type (2.4) in open-loop and (2.14) in closed-loop have a spectrum whose determinant is zero at all frequencies (proof in Appendix A.1). We will nevertheless also show in the sequel that many multisine excitations can yield informative data.

In [BGM10], there exists another result that pertains to data informativity in closed-loop when the excitation is zero. Indeed, as also mentioned in the introduction, consistency can sometimes be ensured by the sole excitation due to the disturbance $\mathbf{H}_0(z)\mathbf{e}(t)$ fed back by the controller $\mathbf{K}(z)$. This result is recalled in the following theorem

Theorem 2.2. *Let the true system \mathcal{S} in (2.2) be identified in closed-loop with a stabilizing control law $\mathbf{u}(t) = -\mathbf{K}(z)\mathbf{y}(t)$ without external excitation ($\mathbf{r} = \mathbf{0}_{n_u \times 1}$), using a model structure \mathcal{M} . Consider*

- a left-coprime row reduced factorization³ for $\mathbf{K}(z)$ of the form

$$\mathbf{K}(z) = \mathbf{V}_c^{-1}(z)\mathbf{N}_c(z)$$

where $\mathbf{N}_c(z)$ and $\mathbf{V}_c(z)$ are coprime and of dimension $n_u \times n_y$ and $n_u \times n_u$ respectively, $\mathbf{V}_c(z)$ being row-reduced. Denote c_{min} the smallest row degree of the matrix $(-\mathbf{N}(z), \mathbf{V}(z))$, which is also called the minimum observability index.

- a left-coprime row reduced factorization for a given $(\Delta\mathbf{W}_y(z), -\Delta\mathbf{W}_u(z))$ of the form

$$(\Delta\mathbf{W}_y(z), -\Delta\mathbf{W}_u(z)) = \mathbf{O}^{-1}(z)(\mathbf{Y}(z), \mathbf{Z}(z))$$

where $\mathbf{Y}(z)$, $\mathbf{Z}(z)$ and $\mathbf{O}(z)$ are polynomial matrices of dimension $n_y \times n_y$, $n_y \times n_u$ and $n_y \times n_y$ respectively, $\mathbf{O}(z)$ being row reduced. Let us denote k_{max} the maximal row degree of $(\mathbf{Y}(z), \mathbf{Z}(z))$ among all $(\Delta\mathbf{W}_y(z), \Delta\mathbf{W}_u(z)) \in \Delta\mathbf{W}$.

Then, the data set \mathcal{Z}_{cl}^∞ in (2.16) is informative with respect to \mathcal{M} with $\mathbf{r} = \mathbf{0}_{n_u \times 1}$ (costless identification) if $c_{min} > k_{max}$.

A similar result pertaining to only the ARMAX model structure is given in [YZ18]. In Chapter 4, we will propose an alternative condition which is based on another approach than the one in [BGM10]. This approach presented in Chapter 4 will have the advantage that it can be extended to the classical case where the excitation vector \mathbf{r} is nonzero. We will also show via an example that our alternative condition for the case $\mathbf{r} = \mathbf{0}_{n_u \times 1}$ can be less conservative than the one in Theorem 2.2.

2.7 Persistency of a vector of quasi-stationary signals

An important notion that will be needed to derive data informativity conditions in the MIMO case is the notion of persistency of a vector of signals. The persistency of a (complex- or real-valued) vector of signals is defined as follows:

Definition 2.4 (Persistency). *A quasi-stationary real- or complex-valued vector ϕ of signals is persistently exciting (PE) if and only if $\bar{E}[\phi(t)\phi^*(t)] \succ 0$.*

We have also the following useful result regarding real-valued vector of signals.

Lemma 2.3. *Consider a vector ϕ containing p real-valued quasi-stationary signals ϕ_h ($h = 1, \dots, p$). Then, the three following propositions are equivalent:*

- ϕ is PE.
- $\bar{E}[\phi(t)\phi^T(t)] \succ 0$.

³See [Kai80] for technical details on left-coprime row reduced factorization.

(c) $\forall \alpha \in \mathbb{R}^p$, $\bar{E}[(\alpha^T \phi(t))^2] = 0$ if and only if $\alpha = \mathbf{0}_{p \times 1}$.

Proof. (a) \Leftrightarrow (b): this follows from a straightforward application of Definition 2.4 for a real-valued vector of signals.

(b) \Leftrightarrow (c): $\bar{E}[\phi(t)\phi^T(t)] \succ 0$ is equivalent to the fact that, for any $\alpha \in \mathbb{R}^p$,

$$\alpha^T \bar{E}[\phi(t)\phi^T(t)]\alpha = 0 \Leftrightarrow \alpha = \mathbf{0}_{p \times 1}$$

which is in turn equivalent to condition (c). □

Remark 2.1. *The condition (c) of Lemma 2.3 means that the elements ϕ_h of ϕ are linearly independent signals.*

From Definition 2.4 and Lemma 2.3, we have also the property that, for any $m \in \mathbb{N}$, $\phi(t - m)$ is PE if and only if $\phi(t)$ is PE and the property that, if ϕ is PE, then all subvectors of ϕ are also PE.

Finally, let us consider a vector signals ϕ_u made up of elements of a quasi-stationary scalar signal u

$$\phi_u(t) = \begin{pmatrix} u(t - \rho) \\ u(t - \rho - 1) \\ \vdots \\ u(t - \mu) \end{pmatrix} \quad (2.19)$$

for arbitrary values of $\rho \in \mathbb{N}$ and $\mu \in \mathbb{N}$ such that $\mu \geq \rho$. Note that the dimension of ϕ_u is $\nu = \mu - \rho + 1$. Then, we have the following result relating the persistency of ϕ_u and the power spectrum of u :

Lemma 2.4 ([GBM08, Lju99]). *Consider the vector ϕ_u defined in (2.19) with a quasi-stationary signal u . The vector ϕ_u is PE if and only if the power spectrum $\Phi_u(\omega)$ of u is nonzero in at least $\nu = \mu - \rho + 1$ frequencies in the interval $]-\pi, \pi]$.*

2.8 Conclusion of Chapter 2 and reformulation of Research Question 1

In this chapter, we have presented background information on the Prediction Error method for the open-loop and the direct closed-loop identification of MIMO systems. We have introduced the definition of data informativity for both configurations. The model structures and the excitations that will be considered in this dissertation have also been given. Based on these notions, we can now give a reformulation of Research Question 1 for the open-loop case and Research Question 2 for the closed-loop case.

Reformulation of **Research Question 1**

Given one of the MIMO model structures defined in Section 2.5 and an excitation \mathbf{u} of the types (2.4)-(2.5) applied to the system \mathcal{S} in *open-loop*, how can we verify that the data set \mathcal{Z}^∞ defined in (2.10) is informative w.r.t. this model structure (see Definition 2.2)?

Reformulation of **Research Question 2**

Given one of the MIMO model structures defined in Section 2.5, an external excitation \mathbf{r} of the types (2.14)-(2.15) applied to the system \mathcal{S} operated in *closed-loop* with a stabilizing controller $\mathbf{K}(z)$ (see Figure 2.1), how can we verify that the data set \mathcal{Z}_{cl}^∞ defined in (2.16) is informative w.r.t. this model structure (see Definition 2.3)?

Data Informativity for MIMO Systems in Open-Loop

Contents

3.1	Introduction of Chapter 3	36
3.2	Simplification of the study of data informativity for MIMO systems	36
3.2.1	Simplification of the study	36
3.3	Data informativity for MISO structures in open-loop	38
3.3.1	Introduction and notations for a MISO model structure	38
3.3.2	From data informativity to input regressor persistency	39
3.3.3	Necessary and sufficient condition for the persistency of the input regressor	42
3.3.4	Necessary and sufficient condition for ϕ_u to be PE	48
3.4	Multisine excitation yielding informativity	49
3.5	Filtered white noise excitation yielding informativity	52
3.6	Checking informativity without input regressor	54
3.7	A synthetic verification approach of the data informativity for the open-loop identification of MIMO systems	55
3.8	Numerical example	56
3.8.1	True system, model structure and input regressor	56
3.8.2	Identification with multisine	57
3.8.3	Identification with filtered white noise	58
3.8.4	Consideration of a second channel	59
3.8.5	Verification of the results with Monte-Carlo simulations	60
3.9	Conclusion of Chapter 3	61

3.1 Introduction of Chapter 3

The purpose of this chapter is to address Research Question 1 (see Section 2.8) which corresponds to the study of data informativity in open-loop for the MIMO model structures defined in Section 2.5 and the input vector \mathbf{u} of the type (2.4)-(2.5).

For this purpose, in Section 3.2.1, we will see that, for the MIMO model structures defined in Section 2.5, the study of data informativity w.r.t. MIMO model structures can be done channel-by-channel. In other words, to check the data informativity for the MIMO system (2.2) with n_y outputs, we can successively check the data informativity for the n_y MISO systems (2.3). Consequently, we can restrict attention to data informativity with respect to a MISO model structure. In Section 3.3.2, we will see that a sufficient condition for data informativity with respect to a MISO model structure is that a given vector of signals $\phi_{\mathbf{u}}$ is persistently exciting (PE). This vector of signals $\phi_{\mathbf{u}}$ will be made up of elements of the quasi-stationary scalar signals u_k ($k = 1, \dots, n_u$) in \mathbf{u} . This will be done in Section 3.3. This condition is necessary and sufficient for the data informativity w.r.t. MISO ARX/FIR model structures while it is only sufficient for the data informativity w.r.t. MISO ARMAX/BJ/OE model structures.

Then, necessary and sufficient conditions are developed to guarantee the persistency of $\phi_{\mathbf{u}}$ for input vector \mathbf{u} of the type (2.4)-(2.5). We will see that its persistency is equivalent to the fact that a given matrix is full row rank. This matrix will be a function of the input parametrization, i.e., the amplitudes, phase shifts and frequencies of the different sinusoids in the multisine case (see (2.4)) and the coefficients in the matrix $\mathbf{M}(z)$ in the filtered white noise case (see (2.5)). We then analyze which conditions are necessary to make this matrix full row rank. We therefore derive, for the multisine case, conditions on the number of sinusoids that are present in the multisine input vector and, for the filtered white noise case, conditions on the complexity of the matrix $\mathbf{M}(z)$ and on the number of white noises in \mathbf{v} .

Since the persistency of $\phi_{\mathbf{u}}$ is only a sufficient condition for the BJ, OE and ARMAX case, we consider in Section 3.6 alternative data informativity conditions. After having synthesized the results in Section 3.7, a numerical example is provided in Section 3.8. Finally, we give a conclusion in Section 3.9.

3.2 Simplification of the study of data informativity for MIMO systems

3.2.1 Simplification of the study

In this chapter, we want to derive conditions on input vectors \mathbf{u} of the type (2.4) or (2.5) to verify that an open-loop identification experiment with these input vectors within one of the model structures described in Section 2.5 will yield a consistent estimate $\hat{\boldsymbol{\theta}}_N$ (see (2.9)). In other words, we want to find conditions on these input vectors for

the data set \mathcal{Z}^∞ defined in Definition 2.2 to be informative with respect to the chosen model structure \mathcal{M} . For this purpose, we first derive simpler conditions to verify the informativity of \mathcal{Z}^∞ . These simpler conditions are subsequently given in Theorems 3.1 and 3.2.

Note that, in the core of the text, we will for the sake of clarity only consider the ARX, FIR, BJ and OE model structures. The results pertaining to the ARMAX model structures will be presented in an appendix (see later).

Theorem 3.1. *Consider Definition 2.2 and consider the notations of Section 2.5 for the ARX, FIR, BJ and OE model structures \mathcal{M} . Define the set*

$$\Delta_{\mathbf{X}} = \{\Delta\mathbf{X}(z) = \mathbf{X}(z, \boldsymbol{\theta}') - \mathbf{X}(z, \boldsymbol{\theta}'') \mid \boldsymbol{\theta}' \text{ and } \boldsymbol{\theta}'' \text{ in } \mathcal{D}_\theta\} \quad (3.1)$$

where $\mathbf{X}(z, \boldsymbol{\theta})$ is defined by

- $\mathbf{X}(z, \boldsymbol{\theta}) = \mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ for ARX/FIR model structure.
- $\mathbf{X}(z, \boldsymbol{\theta}) = \mathbf{G}(z, \tilde{\boldsymbol{\theta}})$ for BJ/OE model structures.

Then, the data set \mathcal{Z}^∞ introduced in Definition 2.2 is informative w.r.t. the model structure \mathcal{M} , if and only if, for all $\Delta\mathbf{X}(z) \in \Delta_{\mathbf{X}}$, the following property holds

$$\bar{E}[|\Delta\mathbf{X}(z)\mathbf{u}(t)|^2] = 0 \implies \Delta\mathbf{X}(z) \equiv \mathbf{0}_{n_y \times n_u} \quad (3.2)$$

Proof. See Appendix B.1 for the proof. □

Remark 3.1. *Using Theorem 3.1, we are now able to give a formal proof of Lemma 2.1 in Chapter 2. This proof is given in Appendix B.2.*

Theorem 3.2. *Consider Definition 2.2 and Theorem 3.1 and let us define, based on $\Delta_{\mathbf{X}}$ given in (3.1), the sets*

$$\Delta_{\mathbf{X},i} = \{\Delta\mathbf{X}_{i,:}(z) \mid \Delta\mathbf{X}_{i,:}(z) \text{ is the } i\text{th row of } \Delta\mathbf{X}(z) \in \Delta_{\mathbf{X}}\}$$

for all $i = 1, \dots, n_y$. Then, the data set \mathcal{Z}^∞ (see (2.10)) is informative w.r.t. \mathcal{M} if and only if, for all $i = 1, \dots, n_y$, the following property holds for all $\Delta\mathbf{X}_{i,:}(z) \in \Delta_{\mathbf{X},i}$

$$\bar{E}[|\Delta\mathbf{X}_{i,:}(z)\mathbf{u}(t)|^2] = 0 \implies \Delta\mathbf{X}_{i,:}(z) \equiv \mathbf{0}_{1 \times n_u} \quad (3.3)$$

Proof. See Appendix B.3 for the proof. □

Theorems 3.1 and 3.2 are important since they allow one to check whether an input vector \mathbf{u} yields informative data for the identification of a MIMO system \mathcal{S} (see (2.2)) by

successively checking that this input vector yields informative data for the identification of the n_y MISO systems \mathcal{S}_i in (2.3) (see Section 2.1). To see this, let us define, based on \mathcal{Z}^∞ in (2.10), the infinite data sets

$$\mathcal{Z}_i^\infty = \{\mathbf{x}_i(t) = (y_i(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, \infty\}$$

with y_i as given in (2.3). Then, by using Theorem 3.1, we observe that verifying (3.3) for all $\Delta \mathbf{X}_i(z) \in \mathbf{\Delta}_{\mathbf{X}, \mathbf{i}}$ is equivalent to verifying that \mathcal{Z}_i^∞ is informative with respect to the MISO model structure \mathcal{M}_i

$$\mathcal{M}_i = \{(\mathbf{G}_i(z, \boldsymbol{\theta}_i), H_i(z, \boldsymbol{\theta}_i)) \mid \boldsymbol{\theta}_i \in \mathcal{D}_{\boldsymbol{\theta}, i}\}$$

where \mathbf{G}_i is the i^{th} row of \mathbf{G} and $\mathcal{D}_{\boldsymbol{\theta}, i}$ is the set that restricts $\boldsymbol{\theta}_i$ to those values such that $\mathbf{G}_i(z, \boldsymbol{\theta}_i)$ is stable and $H_i(z, \boldsymbol{\theta}_i)$ is stable and inversely stable.

To sum up, in order to verify that \mathcal{Z}^∞ is informative w.r.t \mathcal{M} , we just have to verify that, for all $i = 1, \dots, n_y$, \mathcal{Z}_i^∞ is informative with respect to \mathcal{M}_i .

3.3 Data informativity for MISO structures in open-loop

3.3.1 Introduction and notations for a MISO model structure

Consequently, in this chapter, we will from now onwards only consider MISO model structures of the type \mathcal{M}_i . Let us redefine some notations pertaining to the row vectors $\mathbf{X}_i(z)$ corresponding to the MISO model structure \mathcal{M}_i . For a MISO model structure \mathcal{M}_i as defined at the end of the previous section and that can be derived from the MIMO structures of Section 2.5, $\mathbf{X}_i(z, \tilde{\boldsymbol{\theta}}_i)$ is given by:

- $\mathbf{X}_i(z, \tilde{\boldsymbol{\theta}}_i) = \mathbf{B}_i(z, \tilde{\boldsymbol{\theta}}_i)$ for ARX/FIR model structures.
- $\mathbf{X}_i(z, \tilde{\boldsymbol{\theta}}_i) = \mathbf{G}_i(z, \tilde{\boldsymbol{\theta}}_i)$ for BJ/OE model structures.

For simplicity, we will drop the indexes i and i : from now onwards.

In a MISO ARX/FIR model structure, we have thus $\mathbf{X}(z, \tilde{\boldsymbol{\theta}}) = \mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ where $\mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ is a row vector consisting of n_u independently parametrized polynomials $B_k(z, \tilde{\boldsymbol{\theta}}_k)$ ($k = 1, \dots, n_u$). In the BJ/OE case, $\mathbf{X}(z, \tilde{\boldsymbol{\theta}}) = \mathbf{G}(z, \tilde{\boldsymbol{\theta}})$ where $\mathbf{G}(z, \tilde{\boldsymbol{\theta}})$ is a row vector consisting of n_u independently parametrized transfer functions $G_k(z, \tilde{\boldsymbol{\theta}}_k)$. In both cases, the parameter vector $\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{\tilde{n}}$ is the concatenation of all $\tilde{\boldsymbol{\theta}}_k$ ($k = 1, \dots, n_u$), i.e., $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\theta}}_1^T, \tilde{\boldsymbol{\theta}}_2^T, \dots, \tilde{\boldsymbol{\theta}}_{n_u}^T)^T$ and where \tilde{n} is the dimension of $\tilde{\boldsymbol{\theta}}$. Let us also introduce the following notation for each entry of these vectors:

$$\begin{aligned} B_k(z, \tilde{\boldsymbol{\theta}}_k) &= z^{-\rho_k} \mathring{B}_k(z, \tilde{\boldsymbol{\theta}}_k) \\ G_k(z, \tilde{\boldsymbol{\theta}}_k) &= z^{-\rho_k} \frac{\mathring{B}_k(z, \tilde{\boldsymbol{\theta}}_k)}{F_k(z, \tilde{\boldsymbol{\theta}}_k)} \end{aligned} \quad (3.4)$$

where the delays ρ_k can be different for each $k \in \llbracket 1, n_u \rrbracket$ and where both \mathring{B}_k and F_k are polynomials (F_k is moreover a monic polynomial and \mathring{B}_k has a zero-delay). In the sequel, we will denote the degree of these polynomials by $\deg(\mathring{B}_k)$ and $\deg(F_k)$, respectively. The coefficients of the polynomials \mathring{B}_k and F_k are the parameters in $\tilde{\theta}_k$. Consequently, we have that:

$$\begin{aligned}\mathring{B}_k(z, \tilde{\theta}_k) &= \tilde{\theta}_{k,1} + \sum_{m=1}^{\deg(\mathring{B}_k)} \tilde{\theta}_{k,(m+1)} z^{-m} \\ F_k(z, \tilde{\theta}_k) &= 1 + \sum_{m=1}^{\deg(F_k)} \tilde{\theta}_{k,(m+\deg(\mathring{B}_k)+1)} z^{-m}\end{aligned}$$

where $\tilde{\theta}_{k,m}$ denotes the m^{th} entry of $\tilde{\theta}_k$. The number of parameters to be identified in \mathring{B}_k and F_k is thus equal to $\deg(\mathring{B}_k) + 1$ and $\deg(F_k)$, respectively.

In Section 2.5, we have restricted the value of $\theta \in \mathbb{R}^n$ to the ones in \mathcal{D}_θ . Since only $\tilde{\theta} \in \mathbb{R}^{\tilde{n}}$ is relevant for data informativity, we introduce the set $\mathcal{D}_{\tilde{\theta}}$ as $\mathcal{D}_{\tilde{\theta}} = \{\tilde{\theta} \in \mathbb{R}^{\tilde{n}} \mid \mathbf{B}(z, \tilde{\theta}) \text{ is stable}\}$ for ARX/FIR model structures and as $\mathcal{D}_{\tilde{\theta}} = \{\tilde{\theta} \in \mathbb{R}^{\tilde{n}} \mid \mathbf{G}(z, \tilde{\theta}) \text{ is stable}\}$ for BJ/OE model structures. Note that $\mathcal{D}_{\tilde{\theta}}$ covers the whole parameter space $\mathbb{R}^{\tilde{n}}$ in the ARX/FIR case since $\mathbf{B}(z, \tilde{\theta})$ is always stable.

3.3.2 From data informativity to input regressor persistency

Using the notion of persistency defined in Section 2.7, we show in the next theorem that the data informativity in the MISO case can be guaranteed by the persistency of a particular vector of signals depending on the input vector \mathbf{u} (see Definition 2.4). This vector will be denoted $\phi_{\mathbf{u}}$ and will be called *regressor*.

Theorem 3.3 (MISO informativity and regressor persistency). *Consider the data set \mathcal{Z}^∞ obtained by applying an input vector \mathbf{u} on a MISO system (see Definition 2.2). Consider a full-order model structure \mathcal{M} for this MISO system and the notations introduced in Section 3.3.1. Then, the data set \mathcal{Z}^∞ is informative with respect to \mathcal{M} if the regressor $\phi_{\mathbf{u}}$ defined below is PE:*

$$\phi_{\mathbf{u}}(t) = \begin{pmatrix} \phi_{u_1}(t) \\ \vdots \\ \phi_{u_{n_u}}(t) \end{pmatrix} \quad \text{with} \quad \phi_{u_k}(t) = \begin{pmatrix} u_k(t - \rho_k) \\ \vdots \\ u_k(t - \mu_k) \end{pmatrix} \quad (3.5)$$

The dimension ν of the regressor $\phi_{\mathbf{u}}$ in (3.5) is given by $\nu = n_u + \sum_{k=1}^{n_u} (\mu_k - \rho_k)$ (the dimension of ϕ_{u_k} being equal to $\mu_k - \rho_k + 1$ ($k = 1, \dots, n_u$)) and the scalars μ_k ($k = 1, \dots, n_u$) in (3.5) are respectively given by:

- $\mu_k = \deg(B_k)$ for ARX and FIR model structures.
- $\mu_k = \deg(B_k) + \deg(F_k) + 2 \sum_{m=1, m \neq k}^{n_u} \deg(F_m)$ for BJ and OE model structures.

For the ARX and FIR model structures, the persistency of $\phi_{\mathbf{u}}$ is not only a sufficient condition for data informativity but also a necessary condition.

Proof. See Appendix B.4 for the proof. □

If we apply Theorem 3.3 to the SISO case ($n_u = 1$), we observe that the condition for data informativity is that the regressor $\phi_{\mathbf{u}} = \phi_{u_1}$ of dimension $\nu = \mu_1 - \rho_1 + 1$ is PE. Using Lemma 2.4 (see Section 2.7), this condition is equivalent to the fact that the power spectrum $\Phi_{\mathbf{u}}(\omega)$ of $u = u_1$ is nonzero in at least $\nu = \mu_1 - \rho_1 + 1$ frequencies in $]-\pi, \pi]$. By computing μ_1 when $n_u = 1$, we then retrieve the results for data informativity with respect to model structures in the SISO case [GBM08].

In the SISO case, the persistency of ϕ_{u_1} is a necessary and sufficient condition for data informativity for all model structures. However, in the MISO case, the persistency of the regressor $\phi_{\mathbf{u}}$ defined in (3.5) is only a necessary and sufficient condition for data informativity w.r.t. ARX/FIR model structures while it is just a sufficient condition in the BJ/OE case. In Section 3.8, we will show in an example that this sufficient condition is nevertheless not overly conservative.

Remark 3.2. *A result similar to the one of Theorem 3.3 can also be derived for MISO ARMAX model structures. This result is given in Appendix B.5. The reason to have limited Theorem 3.3 to BJ, OE, ARX, FIR structures is the sake of clarity. The results for ARMAX are indeed a bit more cumbersome to present.*

In the next sections, we will derive a necessary and sufficient condition to verify that a given input vector \mathbf{u} of the type (2.4) and (2.5) yields a regressor $\phi_{\mathbf{u}}$ that is PE. In the ARX/FIR case, this condition will then be a necessary and sufficient condition for data informativity while it will be only sufficient in the BJ/OE case. Before deriving this necessary and sufficient condition for $\phi_{\mathbf{u}}$ to be PE, let us formulate the following lemma that will give us a first idea on the properties that \mathbf{u} must satisfy for this purpose.

Lemma 3.1. *Consider a quasi-stationary input vector \mathbf{u} of dimension n_u and the corresponding regressor $\phi_{\mathbf{u}}$ defined in (3.5). If $\phi_{\mathbf{u}}$ is PE, then we have necessarily the following two properties:*

- (i) each ϕ_{u_k} ($k = 1, \dots, n_u$) is PE
- (ii) for any $m \in \mathbb{N}$, the vector $\phi_{\mathcal{U}_m}$ consisting of the elements of the set $\mathcal{U}_m = \{u_k(t) \mid u_k(t - m) \text{ is present in } \phi_{\mathbf{u}}\}$ is PE. In other words, the inputs u_k that appear with the same delay in $\phi_{\mathbf{u}}$ must be linearly independent.

Moreover, for any $k = 1, \dots, n_u$, the vector ϕ_{u_k} of dimension $\mu_k - \rho_k + 1$ is PE if and only if the power spectrum $\Phi_{u_k}(\omega)$ of u_k ($k = 1, \dots, n_u$) is nonzero in at least $\mu_k - \rho_k + 1$ frequencies in $]-\pi, \pi]$.

Proof. The first point directly follows from condition (c) of Lemma 2.3 and the expression (3.5) for $\phi_{\mathbf{u}}$. Indeed, the linear independence condition can only hold for $\phi_{\mathbf{u}}$ if it also holds for each ϕ_{u_k} ($k = 1, \dots, n_u$). The second point follows from a similar reasoning and from the fact that a quasi-stationary vector ϕ is PE if and only if $\phi(t - m)$ is PE $\forall m \in \mathbb{N}$. Finally, the last part of the lemma is a straightforward application of Lemma 2.4 in Section 2.7. \square

Let us first discuss the necessary condition (i). This condition will be always respected in the filtered white noise case (see (2.5)) since $\Phi_{u_k}(\omega)$ will be nonzero at (almost) all frequencies ω (u_k is indeed generated by at least one filtered white noise). In the multisine case (see (2.4)), the condition (i) entails that, for each $k = 1, \dots, n_u$, the number s_k of sinusoids that are effectively present in u_k (i.e., for which $\lambda_{kl} \neq 0$) satisfies the following constraint:

$$s_k \geq \frac{\mu_k - \rho_k + 1}{2} \quad (3.6)$$

Note that this condition is independent of the choice of the amplitudes and phase shifts in (2.4).

Let us now consider the property (ii). In the majority of the model structures, there will be a value of the delay m for which \mathcal{U}_m will contain all entries of \mathbf{u} (e.g., this happens when the delays ρ_k are the same for all $k \in \llbracket 1, n_u \rrbracket$). When this is the case, property (ii) requires that the signals u_k ($k = 1, \dots, n_u$) are linearly independent (the vector of signals \mathbf{u} is PE). In the multisine case, we will see in the sequel that \mathbf{u} is PE if and only if a certain matrix consisting of the amplitudes λ_{kl} and the phase shifts ψ_{kl} is full row rank. In the filtered white noise case, this entails that the rows of $\mathbf{M}(z)$ are linearly independent. Even though choosing an input vector \mathbf{u} that is PE is a good practice, this property is only strictly necessary for $\phi_{\mathbf{u}}$ to be PE for the model structures having the property described in the beginning of this paragraph, as shown in the following example that also illustrates the construction of $\phi_{\mathbf{u}}$.

Example 3.1. Consider a MISO OE model structure with $n_u = 3$ and for which the predictor $\hat{y}(t, \boldsymbol{\theta}) = \mathbf{G}(z, \boldsymbol{\theta})\mathbf{u}(t)$ is given by:

$$\hat{y}(t, \boldsymbol{\theta}) = \underbrace{\frac{\tilde{\theta}_{1,1}}{1 + \tilde{\theta}_{1,2}z^{-1}}}_{G_1(z, \boldsymbol{\theta})} u_1(t) + \underbrace{(\tilde{\theta}_{2,1}z^{-4} + \tilde{\theta}_{2,2}z^{-5})}_{G_2(z, \boldsymbol{\theta})} u_2(t) + \underbrace{\frac{\tilde{\theta}_{3,1}z^{-10}}{1 + \tilde{\theta}_{3,2}z^{-1}}}_{G_3(z, \boldsymbol{\theta})} u_3(t)$$

with $\boldsymbol{\theta}^T = \tilde{\boldsymbol{\theta}}^T = (\tilde{\theta}_{1,1} \ \tilde{\theta}_{1,2} \ \tilde{\theta}_{2,1} \ \tilde{\theta}_{2,2} \ \tilde{\theta}_{3,1} \ \tilde{\theta}_{3,2})$. We observe that, in this model structure, the delays are given by $(\rho_1, \rho_2, \rho_3) = (0, 4, 10)$ and that the transfer functions

$G_k(z, \boldsymbol{\theta})$ ($k = 1, 2, 3$) in $\mathbf{G}(z, \boldsymbol{\theta})$ have the form in (3.4) with

$$\begin{aligned} \mathring{B}_1(z, \boldsymbol{\theta}) &= \tilde{\theta}_{1,1} & \mathring{B}_2(z, \boldsymbol{\theta}) &= \tilde{\theta}_{2,1} + \tilde{\theta}_{2,2}z^{-1} & \mathring{B}_3(z, \boldsymbol{\theta}) &= \tilde{\theta}_{3,1} \\ F_1(z, \boldsymbol{\theta}) &= 1 + \tilde{\theta}_{1,2}z^{-1} & F_2(z, \boldsymbol{\theta}) &= 1 & F_3(z, \boldsymbol{\theta}) &= 1 + \tilde{\theta}_{3,2}z^{-1} \end{aligned}$$

Let us deduce the regression order μ_k defined in Theorem 3.3:

- $\mu_1 = \deg(B_1) + \deg(F_1) + 2\deg(F_2) + 2\deg(F_3) = 3$
- $\mu_2 = \deg(B_2) + \deg(F_2) + 2\deg(F_1) + 2\deg(F_3) = 9$
- $\mu_3 = \deg(B_3) + \deg(F_3) + 2\deg(F_1) + 2\deg(F_2) = 13$

and let us form the input regressor $\boldsymbol{\phi}_u$ defined in (3.5), i.e., $\boldsymbol{\phi}_u(t) = (\phi_{u_1}(t)^T, \phi_{u_2}(t)^T, \phi_{u_3}(t)^T)^T$ with

$$\begin{aligned} \phi_{u_1}^T(t) &= (u_1(t) \quad u_1(t-1) \quad u_1(t-2) \quad u_1(t-3)) \\ \phi_{u_2}^T(t) &= (u_2(t-4) \quad u_2(t-5) \quad u_2(t-6) \quad u_2(t-7) \quad u_2(t-8) \quad u_2(t-9)) \\ \phi_{u_3}^T(t) &= (u_3(t-10) \quad u_3(t-11) \quad u_3(t-12) \quad u_3(t-13)) \end{aligned}$$

We can see that, for all $m \in \mathbb{N}$, each set \mathcal{U}_m defined in Lemma 3.1 contains at most one signal. This means that the condition (ii) of Lemma 3.1 simply says that the inputs u_k must not be identically equal to 0. It is therefore possible to get the persistency of $\boldsymbol{\phi}_u$ even if there is a linear dependence between u_1 , u_2 and u_3 . In particular, we can choose $u_1 = u_2 = u_3 = v$ where v is a white noise signal. In this case, $\boldsymbol{\phi}_u$ is indeed PE since it consists of different elements of a white noise sequence. This choice of \mathbf{u} thus guarantees data informativity (Theorem 3.3).

It is important to stress that Lemma 3.1 only gives necessary conditions for $\boldsymbol{\phi}_u$ to be PE. In the next section, we will therefore continue our analysis by deriving a necessary and sufficient condition for the input regressor to be PE. This result will entail (among other aspects) that, to guarantee the persistency (which implies the data informativity), we also require an additional condition on the total number s of sinusoids in the multisine case and on the complexity of the filter matrix $\mathbf{M}(z)$ in the filtered white noise case.

3.3.3 Necessary and sufficient condition for the persistency of the input regressor

Approach for the development of a necessary and sufficient condition

For input signals u_k of types (2.4)-(2.5), the necessary and sufficient condition will be derived by rewriting $\boldsymbol{\phi}_u$ (see (3.5)) into the form $\boldsymbol{\phi}_u(t) = \mathbf{T}\boldsymbol{\varphi}(t)$ where \mathbf{T} is a time-independent and deterministic (possibly complex) matrix whose entries are functions

of the input parametrization and $\boldsymbol{\varphi}$ is a vector of (possibly complex-valued) quasi-stationary signals that is always PE (see Definition 2.4). This decomposition is important since we have then the following result:

Lemma 3.2. *Consider a vector of quasi-stationary signals $\boldsymbol{\phi}$ given by $\boldsymbol{\phi}(t) = \mathbf{T}\boldsymbol{\varphi}(t)$ where \mathbf{T} is a time-independent, deterministic and complex matrix and $\boldsymbol{\varphi}$ is a vector of complex-valued quasi-stationary signals that has the property of being PE. Then, $\boldsymbol{\phi}$ is PE if and only if \mathbf{T} is full row rank. \square*

Proof. Due to Definition 2.4, the lemma will be proven if we show that $\bar{E}[\boldsymbol{\phi}(t)\boldsymbol{\phi}^*(t)] = \mathbf{T}\bar{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^*(t)]\mathbf{T}^* \succ 0$. Since $\boldsymbol{\varphi}$ is PE, we have $\bar{E}[\boldsymbol{\varphi}(t)\boldsymbol{\varphi}^*(t)] \succ 0$. Consequently, a full row rank \mathbf{T} is indeed a necessary and sufficient condition for $\bar{E}[\boldsymbol{\phi}(t)\boldsymbol{\phi}^*(t)] \succ 0$. \square

In the next two paragraphs, we will show that we can rewrite $\boldsymbol{\phi}_{\mathbf{u}}$ in this way for multisine and filtered white noise excitation. We will denote by $\mathbf{T}^{(sin)}$ and $\boldsymbol{\varphi}^{(sin)}$ the corresponding \mathbf{T} and $\boldsymbol{\varphi}$ when \mathbf{u} is a multisine. When it is a filtered white noise, we will use the notations $\mathbf{T}^{(fwn)}$ and $\boldsymbol{\varphi}^{(fwn)}$.

Rewriting of the input regressor for multisine excitation

In this paragraph, \mathbf{u} is a multisine signal vector where each of its entry u_k is given by (2.4). By using Euler's formula, we have that, for all $m_k \in \llbracket \rho_k, \mu_k \rrbracket$,

$$\begin{aligned} u_k(t - m_k) &= \sum_{l=1}^s \lambda_{kl} \cos(\omega_l(t - m_k) + \psi_{kl}) \\ &= \frac{1}{2} \sum_{l=1}^s \left(\bar{\lambda}_{kl} e^{-jm_k\omega_l} e^{j\omega_l t} + \bar{\lambda}_{kl}^* e^{jm_k\omega_l} e^{-\omega_l t} \right) \end{aligned}$$

where j is the complex number such that $j^2 = -1$ and $\bar{\lambda}_{kl} = \lambda_{kl} e^{j\psi_{kl}}$ is a phasor. Consequently, for all $k = 1, \dots, n_u$, the regressor $\boldsymbol{\phi}_{u_k}$ defined in (3.5) can be rewritten as

$$\begin{aligned} \boldsymbol{\phi}_{u_k}(t) &= \begin{pmatrix} u_k(t - \rho_k) \\ u_k(t - \rho_k - 1) \\ \vdots \\ u_k(t - \mu_k) \end{pmatrix} = \frac{1}{2} \sum_{l=1}^s \begin{pmatrix} \bar{\lambda}_{kl} e^{-j\rho_k\omega_l} e^{j\omega_l t} + \bar{\lambda}_{kl}^* e^{j\rho_k\omega_l} e^{-\omega_l t} \\ \bar{\lambda}_{kl} e^{-j(\rho_k+1)\omega_l} e^{j\omega_l t} + \bar{\lambda}_{kl}^* e^{j(\rho_k+1)\omega_l} e^{-\omega_l t} \\ \vdots \\ \bar{\lambda}_{kl} e^{-j\mu_k\omega_l} e^{j\omega_l t} + \bar{\lambda}_{kl}^* e^{j\mu_k\omega_l} e^{-\omega_l t} \end{pmatrix} \\ &= \mathbf{T}_k^{(sin)} \boldsymbol{\varphi}^{(sin)}(t) \end{aligned}$$

where the matrix $\mathbf{T}_k^{(sin)}$ of dimension $(\mu_k - \rho_k + 1) \times 2s$ and the signal vector $\boldsymbol{\varphi}^{(sin)}$ of dimension $2s$ are successively defined by

$$\begin{aligned}
 \mathbf{T}_k^{(sin)} &= \begin{pmatrix} \bar{\lambda}_{k1}e^{-j\rho_k\omega_1} & \bar{\lambda}_{k1}^*e^{j\rho_k\omega_1} & \dots & \bar{\lambda}_{ks}e^{-j\rho_k\omega_s} & \bar{\lambda}_{ks}^*e^{j\rho_k\omega_s} \\ \bar{\lambda}_{k1}e^{-j(\rho_k+1)\omega_1} & \bar{\lambda}_{k1}^*e^{j(\rho_k+1)\omega_1} & \dots & \bar{\lambda}_{ks}e^{-j(\rho_k+1)\omega_s} & \bar{\lambda}_{ks}^*e^{j(\rho_k+1)\omega_s} \\ \vdots & \vdots & \dots & \vdots & \vdots \\ \bar{\lambda}_{k1}e^{-j\mu_k\omega_1} & \bar{\lambda}_{k1}^*e^{j\mu_k\omega_1} & \dots & \bar{\lambda}_{ks}e^{-j\mu_k\omega_s} & \bar{\lambda}_{ks}^*e^{j\mu_k\omega_s} \end{pmatrix} \\
 \boldsymbol{\varphi}^{(sin)}(t) &= \frac{1}{2} \begin{pmatrix} e^{j\omega_1 t} \\ e^{-j\omega_1 t} \\ \vdots \\ e^{j\omega_s t} \\ e^{-j\omega_s t} \end{pmatrix} \tag{3.7}
 \end{aligned}$$

Consequently, the input regressor $\boldsymbol{\phi}_u$ can be rewritten as follows

$$\begin{aligned}
 \boldsymbol{\phi}_u(t) &= \mathbf{T}^{(sin)} \boldsymbol{\varphi}^{(sin)}(t) \tag{3.8} \\
 \mathbf{T}^{(sin)} &= \begin{pmatrix} \mathbf{T}_1^{(sin)} \\ \mathbf{T}_2^{(sin)} \\ \vdots \\ \mathbf{T}_{n_u}^{(sin)} \end{pmatrix}
 \end{aligned}$$

where the matrix $\mathbf{T}^{(sin)}$ is of dimension $\nu \times 2s$ with $\nu = n_u + \sum_{k=1}^{n_u} (\mu_k - \rho_k)$ (see Theorem 3.3). An illustration of the construction of $\mathbf{T}^{(sin)}$ is given in the following example.

Example 3.2. Consider an input regressor $\boldsymbol{\phi}_u$ with $n_u = 2$, $(\rho_1, \rho_2) = (1, 0)$, $(\mu_1, \mu_2) = (3, 2)$. Consider the following multisine excitation with $s = 3$ frequencies

$$\mathbf{u}(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = \begin{pmatrix} \cos(\omega_1 t + 0.2) + 3 \cos(\omega_3 t + 0.7) \\ 2 \cos(\omega_2 t - 0.5) - 0.1 \cos(\omega_3 t) \end{pmatrix}$$

where $\omega_1 = 0.1$, $\omega_2 = 0.5$ and $\omega_3 = 2$. The phasors $\bar{\lambda}_{kl}$ ($k = 1, \dots, n_u$, $l = 1, \dots, s$) are given by

$$\begin{aligned}
 \bar{\lambda}_{11} &= e^{j0.2} & \bar{\lambda}_{12} &= 0 & \bar{\lambda}_{13} &= 3e^{j0.7} \\
 \bar{\lambda}_{21} &= 0 & \bar{\lambda}_{22} &= 2e^{-j0.5} & \bar{\lambda}_{23} &= -0.1
 \end{aligned}$$

First, $\boldsymbol{\varphi}^{(sin)}$ is given by $\boldsymbol{\varphi}^{(sin)}(t) = (e^{j\omega_1 t}, e^{-j\omega_1 t}, e^{j\omega_2 t}, e^{-j\omega_2 t}, e^{j\omega_3 t}, e^{-j\omega_3 t})^T$. Therefore, the matrices $\mathbf{T}_k^{(sin)}$ with $k = 1, 2$ are given by

$$\begin{aligned}
 \mathbf{T}_1^{(sin)} &= \begin{pmatrix} e^{j0.1} & e^{-j0.1} & 0 & 0 & 3e^{-j1.3} & 3e^{j1.3} \\ 1 & 1 & 0 & 0 & 3e^{-j3.3} & 3e^{j3.3} \\ e^{-j0.1} & e^{j0.1} & 0 & 0 & 3e^{-j5.3} & 3e^{j5.3} \end{pmatrix} \\
 \mathbf{T}_2^{(sin)} &= \begin{pmatrix} 0 & 0 & 2e^{-j} & 2e^j & -0.1e^{-j2} & -0.1e^{j2} \\ 0 & 0 & 2e^{-j1.5} & 2e^{j1.5} & -0.1e^{-j4} & -0.1e^{j4} \\ 0 & 0 & 2e^{-j2} & 2e^{j2} & -0.1e^{-j6} & -0.1e^{j6} \end{pmatrix}
 \end{aligned}$$

Consequently, the matrix $\mathbf{T}^{(sin)}$ is given by

$$\mathbf{T}^{(sin)} = \begin{pmatrix} \mathbf{T}_1^{(sin)} \\ \mathbf{T}_2^{(sin)} \end{pmatrix} = \begin{pmatrix} e^{j0.1} & e^{-j0.1} & 0 & 0 & 3e^{-j1.3} & 3e^{j1.3} \\ 1 & 1 & 0 & 0 & 3e^{-j3.3} & 3e^{j3.3} \\ e^{-j0.1} & e^{j0.1} & 0 & 0 & 3e^{-j5.3} & 3e^{j5.3} \\ 0 & 0 & 2e^{-j} & 2e^j & -0.1e^{-j2} & -0.1e^{j2} \\ 0 & 0 & 2e^{-j1.5} & 2e^{j1.5} & -0.1e^{-j4} & -0.1e^{j4} \\ 0 & 0 & 2e^{-j2} & 2e^{j2} & -0.1e^{-j6} & -0.1e^{j6} \end{pmatrix}$$

Remark 3.3. For the sequel, it is important to note that the matrix $\mathbf{T}^{(sin)}$ in (3.8) will have the property that none of its columns is identically equal to 0 since the set of frequencies $\{\omega_1, \dots, \omega_s\}$ is defined in such a way that, for each frequency ω_l ($l = 1, \dots, s$), there will be always at least one u_k such that $\lambda_{kl} \neq 0$ (see below (2.4)).

Rewriting of the input regressor for filtered white noise excitation

In this paragraph, we consider input vectors \mathbf{u} generated as $\mathbf{u}(t) = \mathbf{M}(z)\mathbf{v}(t)$ with a stable matrix $\mathbf{M}(z)$ of transfer functions and with a vector \mathbf{v} consisting of f independent white noises, i.e., each $u_k(t)$ is given by (2.5). To get the desired form for $\phi_{\mathbf{u}}$, we will rewrite $\mathbf{M}(z)$ as $\mathbf{M}(z) = \mathbf{L}(z)/w(z)$ where $\mathbf{L}(z)$ is a $n_u \times f$ matrix of polynomials (matrix of FIR filters) and $w(z)$ the least common multiple of the denominators of $\mathbf{M}(z)$. Each entry $L_{kq}(z)$ of $\mathbf{L}(z)$ ($k = 1, \dots, n_u$, $q = 1, \dots, f$) will always be of the form

$$L_{kq}(z) = \sum_{h=d_{kq}}^{n_{kq}} l_{kq}^{(h)} z^{-h} \quad (3.9)$$

with $n_{kq} = \deg(L_{kq}(z))$, d_{kq} the delay of $L_{kq}(z)$ and $l_{kq}^{(h)}$ the coefficients of the polynomials $L_{kq}(z)$. For all $m_k \in \llbracket \rho_k, \mu_k \rrbracket$, we can thus express $u_k(t - m_k)$ as

$$\begin{aligned} u_k(t - m_k) &= w^{-1}(z) \sum_{q=1}^f L_{kq}(z) v_q(t - m_k) \\ &= w^{-1}(z) \left(\sum_{h=d_{k1}}^{n_{k1}} l_{k1}^{(h)} v_1(t - h - m_k) + \dots + \sum_{h=d_{kf}}^{n_{kf}} l_{kf}^{(h)} v_f(t - h - m_k) \right) \end{aligned} \quad (3.10)$$

Let us consider the above expansions for all entries in $\phi_{\mathbf{u}}$ and let us define a vector $\varphi^{(fwn)}$ containing all elements $v_q(t - h - m_k)$ of the white noise sequences v_q ($q = 1, \dots, f$) that are present in these expansions. For a given q , $\varphi^{(fwn)}$ will thus contain all elements of $v_q(t)$ from $v_q(t - \alpha_q)$ till $v_q(t - \beta_q)$ with

$$\begin{cases} \alpha_q &= \min_{k \in \llbracket 1, n_u \rrbracket} d_{kq} + \rho_k \\ \beta_q &= \max_{k \in \llbracket 1, n_u \rrbracket} n_{kq} + \mu_k \end{cases} \quad (3.11)$$

The vector $\varphi^{(fwn)}$ of dimension $\xi = f + \sum_{q=1}^f (\beta_q - \alpha_q)$ is thus given by:

$$\varphi^{(fwn)}(t) = \begin{pmatrix} \mathbf{v}_1(t) \\ \vdots \\ \mathbf{v}_f(t) \end{pmatrix} \quad \text{with} \quad \mathbf{v}_q(t) = \begin{pmatrix} v_q(t - \alpha_q) \\ v_q(t - \alpha_q - 1) \\ \vdots \\ v_q(t - \beta_q) \end{pmatrix} \quad (3.12)$$

Using $\varphi^{(fwn)}$ and the expression for $u_k(t - m_k)$ in (3.10), we have that:

$$\phi_{u_k}(t) = \underbrace{\left(\mathbf{T}_{k1}^{(fwn)} \quad \dots \quad \mathbf{T}_{kf}^{(fwn)} \right)}_{\mathbf{T}_k^{(fwn)}} w^{-1}(z) \varphi^{(fwn)}(t)$$

where $\mathbf{T}_{kq}^{(fwn)}$ ($q = 1, \dots, f$) is a matrix of dimension $(\mu_k - \rho_k + 1) \times (\beta_q - \alpha_q + 1)$ defined as:

$$\mathbf{T}_{kq}^{(fwn)} = \left(\mathbf{0}_{(\mu_k - \rho_k + 1) \times (d_{kq} + \rho_k - \alpha_q)} \quad \tilde{\mathbf{T}}_{kq}^{(fwn)} \quad \mathbf{0}_{(\mu_k - \rho_k + 1) \times (\beta_q - n_{kq} - \mu_k)} \right)$$

$$\tilde{\mathbf{T}}_{kq}^{(fwn)} = \begin{pmatrix} l_{kq}^{(d_{kq})} & \dots & \dots & \dots & l_{kq}^{(n_{kq})} & \dots & \mathbf{0} \\ & \ddots & \dots & \dots & \dots & \ddots & \\ \mathbf{0} & & l_{kq}^{(d_{kq})} & \dots & \dots & \dots & l_{kq}^{(n_{kq})} \end{pmatrix}$$

We can therefore also write that:

$$\phi_{\mathbf{u}}(t) = \mathbf{T}^{(fwn)} w^{-1}(z) \varphi^{(fwn)}(t) \quad (3.13)$$

where $\mathbf{T}^{(fwn)}$ is a matrix of dimension $\nu \times \xi$ given by

$$\mathbf{T}^{(fwn)} = \begin{pmatrix} \mathbf{T}_1^{(fwn)} \\ \vdots \\ \mathbf{T}_{n_u}^{(fwn)} \end{pmatrix}$$

To illustrate the construction of this matrix $\mathbf{T}^{(fwn)}$, let us consider the following example.

Example 3.3. Consider for the input regressor $n_u = 3$, $(\rho_1, \rho_2, \rho_3) = (1, 1, 3)$, $(\mu_1, \mu_2, \mu_3) = (1, 3, 4)$, $f = 2$ and

$$\mathbf{M}(z) = \begin{pmatrix} \frac{0.2 - z^{-2}}{1 - 0.2z^{-1}} & \frac{z^{-1} - 0.8z^{-2} + 3z^{-3}}{1 - 0.2z^{-1}} \\ \frac{0.5z^{-2}}{1 - 0.2z^{-1}} & -z^{-1} \\ 0 & 0.4z^{-1} + 5z^{-2} \end{pmatrix}$$

First, we have to rewrite $\mathbf{M}(z)$ as $\mathbf{M}(z) = \mathbf{L}(z)/w(z)$ where $\mathbf{L}(z)$ is a $n_u \times f$ matrix of polynomials (matrix of FIR filters) and $w(z)$ the least common multiple of the denominators of $\mathbf{M}(z)$. In our case, $w(z) = 1 - 0.2z^{-1}$ and so $\mathbf{L}(z)$ is given by

$$\mathbf{L}(z) = \begin{pmatrix} 0.2 - z^{-2} & z^{-1} - 0.8z^{-2} + 3z^{-3} \\ 0.5z^{-2} & -z^{-1} + 0.2z^{-2} \\ 0 & 0.4z^{-1} + 4.92z^{-2} - z^{-3} \end{pmatrix}$$

The input regressor $\phi_{\mathbf{u}}$ is thus consisting of $u_1(t-1)$, $u_2(t-1)$, $u_2(t-2)$, $u_2(t-3)$, $u_3(t-3)$ and $u_3(t-4)$ (i.e., $\nu = 6$). To determine $\varphi^{(fwn)}(t)$, we observe that $\alpha_1 = 1$, $\alpha_2 = 2$, $\beta_1 = 5$ and $\beta_2 = 7$ (see (3.11)). Consequently, $\varphi^{(fwn)}(t) = (\mathbf{v}_1^T(t) \quad \mathbf{v}_2^T(t))^T$, where

$$\mathbf{v}_1(t) = \begin{pmatrix} v_1(t-1) \\ v_1(t-2) \\ v_1(t-3) \\ v_1(t-4) \\ v_1(t-5) \end{pmatrix} \quad \text{and} \quad \mathbf{v}_2(t) = \begin{pmatrix} v_2(t-2) \\ v_2(t-3) \\ v_2(t-4) \\ v_2(t-5) \\ v_2(t-6) \\ v_2(t-7) \end{pmatrix}$$

We can now construct successively the matrices $\tilde{\mathbf{T}}_{kq}^{(fwn)}$ ($k = 1, \dots, n_u$, $q = 1, \dots, f$ with $n_u = 3$ and $f = 2$) and the matrix $\mathbf{T}^{(fwn)}$:

$$\begin{aligned} \tilde{\mathbf{T}}_{11}^{(fwn)} &= \begin{pmatrix} 0.2 & 0 & -1 \end{pmatrix} & \tilde{\mathbf{T}}_{12}^{(fwn)} &= \begin{pmatrix} 1 & -0.8 & 3 \end{pmatrix} \\ \tilde{\mathbf{T}}_{21}^{(fwn)} &= \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.5 \end{pmatrix} & \tilde{\mathbf{T}}_{22}^{(fwn)} &= \begin{pmatrix} -1 & 0.2 & 0 & 0 \\ 0 & -1 & 0.2 & 0 \\ 0 & 0 & -1 & 0.2 \end{pmatrix} \\ \tilde{\mathbf{T}}_{31}^{(fwn)} &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} & \tilde{\mathbf{T}}_{32}^{(fwn)} &= \begin{pmatrix} 0.4 & 4.92 & -1 & 0 \\ 0 & 0.4 & 4.92 & -1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \mathbf{T}^{(fwn)} &= \left(\begin{array}{c|c} \mathbf{T}_{11}^{(fwn)} & \mathbf{T}_{12}^{(fwn)} \\ \mathbf{T}_{21}^{(fwn)} & \mathbf{T}_{22}^{(fwn)} \\ \mathbf{T}_{31}^{(fwn)} & \mathbf{T}_{32}^{(fwn)} \end{array} \right) = \left(\begin{array}{ccc|ccc} \mathbf{0}_{1 \times 0} & \tilde{\mathbf{T}}_{11}^{(fwn)} & \mathbf{0}_{1 \times 2} & \mathbf{0}_{1 \times 0} & \tilde{\mathbf{T}}_{12}^{(fwn)} & \mathbf{0}_{1 \times 3} \\ \mathbf{0}_{3 \times 2} & \tilde{\mathbf{T}}_{21}^{(fwn)} & \mathbf{0}_{3 \times 0} & \mathbf{0}_{1 \times 0} & \tilde{\mathbf{T}}_{22}^{(fwn)} & \mathbf{0}_{3 \times 2} \\ \mathbf{0}_{2 \times 2} & \tilde{\mathbf{T}}_{31}^{(fwn)} & \mathbf{0}_{2 \times 1} & \mathbf{0}_{1 \times 2} & \tilde{\mathbf{T}}_{32}^{(fwn)} & \mathbf{0}_{2 \times 0} \end{array} \right) \\ &= \left(\begin{array}{ccccc|ccccc} 0.2 & 0 & -1 & \underline{0} & \underline{0} & 1 & -0.8 & 3 & \underline{0} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & 0.5 & 0 & 0 & -1 & 2 & 0 & 0 & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & 0 & 0.5 & 0 & 0 & -1 & 2 & 0 & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & 0 & 0 & 0.5 & 0 & 0 & -1 & 2 & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & 0 & 0 & \underline{0} & \underline{0} & \underline{0} & 0.4 & 4.92 & -1 & 0 \\ \underline{0} & \underline{0} & 0 & 0 & \underline{0} & \underline{0} & \underline{0} & 0 & 0.4 & 4.92 & -1 \end{array} \right) \end{aligned}$$

where the underlined zeros $\underline{0}$ correspond to the zeros of $\mathbf{0}_{(\mu_k - \rho_k + 1) \times (\beta_q - n_{kq} - \mu_k)}$ and $\mathbf{0}_{(\mu_k - \rho_k + 1) \times (d_{kq} + \rho_k - \alpha_q)}$ for all $\mathbf{T}_{kq}^{(fwn)}$.

Remark 3.4. *If all the coefficients $l_{kq}^{(h)}$ are nonzero in the expression of the filters $L_{kq}(z)$ in (3.9) ($k = 1, \dots, n_u$, $q = 1, \dots, f$), all the columns in $\mathbf{T}^{(fwn)}$ will be nonzero (as it was the case for $\mathbf{T}^{(sin)}$, see Remark 3.3). However, if some of these coefficients appear to be zero, this can no longer be the case. As an example, the second column of $\mathbf{T}^{(fwn)}$ in Example 3.3 is identically zero. This happens because $L_{11}(z)$ misses the term in z^{-1} and, consequently, the white noise element $v_1(t-2)$ will never appear in the expansions describing $u_k(t-m_k)$ given in (3.10).*

3.3.4 Necessary and sufficient condition for ϕ_u to be PE

We can now combine Lemma 3.2 with the matrices $\mathbf{T}^{(sin)}$ and $\mathbf{T}^{(fwn)}$ introduced in the previous subsections to deduce our main theorem giving a necessary and sufficient condition for ϕ_u to be PE:

Theorem 3.4. *Consider the regressor ϕ_u of dimension ν (see Theorem 3.3) and the expressions (3.8) and (3.13) for this regressor corresponding, respectively, to the multisine case (see (2.4)) and the filtered white noise case (see (2.5)). Then, ϕ_u is PE if and only if*

- $\mathbf{T}^{(sin)}$ is full row rank (i.e. $\text{rank}(\mathbf{T}^{(sin)}) = \nu$) in the multisine case
- $\mathbf{T}^{(fwn)}$ is full row rank (i.e. $\text{rank}(\mathbf{T}^{(fwn)}) = \nu$) in the filtered white noise case.

Proof. Let us first consider the multisine case. We first observe that $\varphi^{(sin)}$ in (3.7) is (always) PE since $\bar{E}[\varphi^{(sin)}(t)\varphi^{(sin)*}(t)] = \frac{1}{4}\mathbf{I}_{2s} \succ 0$ (see Appendix B.6). Consequently, the result follows from a direct application of Lemma 3.2. Let us now consider the filtered white noise case. We observe that $\varphi^{(fwn)}$ in (3.12) is (always) PE since it contains elements of independent white noise sequences. By noticing that $w(z)$ is a stable and inversely stable filter, we have that $w^{-1}(z)\varphi^{(fwn)}(t)$ is PE. Indeed, for such $w(z)$, it is straightforward to show using condition (c) of Lemma 2.3 that $\varphi^{(fwn)}(t)$ is PE if and only if $w^{-1}(z)\varphi^{(fwn)}(t)$ is PE. Consequently, an application of Lemma 3.2 yields the equivalence between a full row rank matrix $\mathbf{T}^{(fwn)}$ and the fact that $\phi_u(t)$ is PE. \square

Given the expressions (2.4) and (2.5) for \mathbf{u} , it is easy to construct, respectively, $\mathbf{T}^{(sin)}$ and $\mathbf{T}^{(fwn)}$ and check whether the corresponding input signal will yield a persistently exciting ϕ_u and thus informative data for the considered model structure \mathcal{M} . Indeed, these matrices are a function of the model structures parameters ρ_k and μ_k and of the input vector parametrization. This input vector parametrization is characterized by the amplitudes λ_{kl} , the phase shifts ψ_{kl} and the frequencies ω_l (for the multisine case) and by the matrix $\mathbf{M}(z)$ (filtered white noise case).

We can, e.g., apply the above theorem on the situations described in Examples 3.2 and 3.3. In Example 3.2, the matrix $\mathbf{T}^{(sin)}$ is full row rank, i.e., $\text{rank}(\mathbf{T}^{(sin)}) = \nu = 6$.

Consequently, the multisine input vector $\phi_{\mathbf{u}}$ will therefore yield informative data for any model structure corresponding to the regressor $\phi_{\mathbf{u}}$ used in Example 3.2. In Example 3.3, the matrix $\mathbf{T}^{(fwn)}$ is also full row rank, i.e., $\text{rank}(\mathbf{T}^{(fwn)}) = \nu = 6$. Consequently, the filtered white noise input vector $\phi_{\mathbf{u}}$ will therefore yield informative data for any model structure corresponding to the regressor $\phi_{\mathbf{u}}$ used in Example 3.3. It is interesting to note that this is the case in both examples even though the power spectrum matrix $\Phi_{\mathbf{u}}$ is such that $\det(\Phi_{\mathbf{u}}(\omega)) = 0$ for all frequencies ω in $]-\pi, \pi[$.

Theorem 3.4 gives us thus a tool to verify whether a given input vector will yield informative data for a given model structure. However, it does not give much hints on the properties that the input vector must satisfy in order to yield full row rank matrices $\mathbf{T}^{(sin)}$ and $\mathbf{T}^{(fwn)}$. Such hints will be given in the next sections.

Remark 3.5. *The necessary and sufficient condition for the persistency of $\phi_{\mathbf{u}}$ is the right invertibility of a matrix depending on the model structure complexity and the input parametrization. Although this condition seems easy to verify, the computation of the rank of a matrix, e.g., with MATLAB can be inaccurate.*

3.4 Multisine excitation yielding informativity

In Section 3.3.2, it was discussed that choosing an input vector \mathbf{u} with linearly independent elements (\mathbf{u} is PE) is a good practice and is furthermore a necessary condition to yield a persistently exciting $\phi_{\mathbf{u}}$ for the majority of model structures. We have given in that section a necessary and sufficient condition for \mathbf{u} to be PE in the filtered white noise case. We can now do the same for the multisine case. If \mathbf{u} is described as in (2.4), \mathbf{u} is PE if and only if the following phasor matrix \mathcal{L} is full row rank:

$$\mathcal{L} = \begin{pmatrix} \bar{\lambda}_{11} & \bar{\lambda}_{11}^* & \cdots & \bar{\lambda}_{1s} & \bar{\lambda}_{1s}^* \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \bar{\lambda}_{n_u 1} & \bar{\lambda}_{n_u 1}^* & \cdots & \bar{\lambda}_{n_u s} & \bar{\lambda}_{n_u s}^* \end{pmatrix} \quad (3.14)$$

Let us indeed observe that $\mathbf{u}(t) = \mathcal{L}\varphi^{(sin)}(t)$ with $\varphi^{(sin)}$ as defined in (3.7). Consequently, the above result directly follows from Lemma 3.2 and the fact that $\varphi^{(sin)}$ is PE (see the proof of Theorem 3.4).

To help us verifying that a multisine input vector will yield a persistently exciting $\phi_{\mathbf{u}}$ (and thus data informativity), we have already discussed a number of aspects in Section 3.3.2. In particular, we cannot have a persistently exciting $\phi_{\mathbf{u}}$ if the number s_k of nonzero sinusoids in each u_k does not respect (3.6). This condition can also be explained by the fact that the matrix $\mathbf{T}^{(sin)}$ introduced in Section 3.3.3 can only be full row rank if all the matrices $\mathbf{T}_k^{(sin)}$ are full row rank. Since $\mathbf{T}_k^{(sin)}$ has $\mu_k - \rho_k + 1$ rows and contains $2s_k$ nonzero columns, it is necessary that (3.6) holds for $\mathbf{T}_k^{(sin)}$ to be full row rank. Using a similar reasoning on the full matrix $\mathbf{T}^{(sin)}$, we can now derive the following additional condition on the total number s of sinusoids in the input vector \mathbf{u} .

Lemma 3.3. *Consider a regression vector $\phi_{\mathbf{u}}$ of dimension ν (see Theorem 3.3). Suppose that the corresponding input signal \mathbf{u} of dimension n_u is a multisine containing s*

different frequencies (see (2.4)). Then, for $\phi_{\mathbf{u}}$ to be PE, it is necessary to have that

$$s \geq \frac{\nu}{2} \quad (3.15)$$

Proof. The number of columns in $\mathbf{T}^{(sin)}$ is $2s$ and they are all nonzero (see Remark 3.3) while its number of rows is given by ν . Consequently, it is clear that (3.15) is a necessary condition for $\mathbf{T}^{(sin)}$ to be full row rank. \square

Recall that $\nu = n_u + \sum_{k=1}^{n_u} (\mu_k - \rho_k)$ is larger than $\mu_k - \rho_k + 1$ and that $s \geq s_k \forall k$. It is therefore clear that this additional condition is in no way implied by (3.6). Actually, when $s_k = s$ for each k , it is (3.15) that implies (3.6).

As shown below Theorem 3.3, in the SISO case, the data informativity also boils down to the persistency of a regressor. Using Lemma 2.4 in Section 2.7 for a scalar input multisine, we see that data informativity in the SISO case will be achieved if and only if the scalar multisine input contains a number of frequencies/sinusoids that is larger than the half of the dimension of the input regressor corresponding to the considered model structure. We have an interesting analogy in the MISO case. For $\phi_{\mathbf{u}}$ to be PE, the number s of frequencies in the multisine input vector must be larger than the half of the dimension of the input regressor $\phi_{\mathbf{u}}$ and the number s_k of frequencies in each scalar input u_k of this input vector must also be larger than the half of the dimension of the corresponding part ϕ_{u_k} of $\phi_{\mathbf{u}}$. The main difference is however that the combination of these conditions are not sufficient to ensure that $\phi_{\mathbf{u}}$ is PE for input vector \mathbf{u} (even if we add condition (ii) of Lemma 3.1). Indeed, some phasor and frequency choices can yield a rank deficient matrix $\mathbf{T}^{(sin)}$ or, equivalently, a regressor $\phi_{\mathbf{u}}$ that is not PE. An example is a phasor choice that would yield $u_1(t) = -3 u_2(t - 2)$ while $u_1(t)$ and $u_2(t - 2)$ both lie in $\phi_{\mathbf{u}}(t)$ (see Example 3.4).

However, the above phenomenon can generally be avoided if we choose the phasor value in some random manner. Consider thus the case where, due to, e.g., experiment design limitations, we need to choose the input vector as (2.4) with a given set of s frequencies and with $s_k = s$ for all $k = 1, \dots, n_u$. This defines a class of multisine input vector for which a number of phasors have to be determined. Let us suppose that the necessary conditions (3.6) and (3.15) are respected for these s and s_k and that we can determine the phasors freely. Then, a good practice is to choose these phasors in a random and independent manner and to subsequently check whether the corresponding $\mathbf{T}^{(sin)}$ is indeed full row rank. If it is not the case, a new realization is performed, etc.

Example 3.4. Consider an input regressor $\phi_{\mathbf{u}}$ with $n_u = 2$ and with $(\rho_1, \rho_2) = (0, 1)$ and $(\mu_1, \mu_2) = (1, 2)$. Therefore, the regression vector is given by

$$\phi_{\mathbf{u}}(t) = \begin{pmatrix} u_1(t) \\ u_1(t-1) \\ u_2(t-1) \\ u_2(t-2) \end{pmatrix}$$

which has dimension $\nu = 4$. Consider that the input vector \mathbf{u} is generated with $s = 2$ frequencies $\omega_1 = 0.1$ and $\omega_2 = 0.2$ and that $s_1 = s_2 = 2$. This defines the following class of input vectors:

$$\begin{aligned} u_1(t) &= \lambda_{11} \cos(0.1t + \psi_{11}) + \lambda_{12} \cos(0.2t + \psi_{12}) \\ u_2(t) &= \lambda_{21} \cos(0.1t + \psi_{21}) + \lambda_{22} \cos(0.2t + \psi_{22}) \end{aligned}$$

This input vector class respects both (3.15) and (3.6).

Consider first the following choice for the four phasors $\bar{\lambda}_{kl} = \lambda_{kl}e^{j\psi_{kl}}$ ($k = 1, 2$, $l = 1, 2$): $\bar{\lambda}_{11} = 0.7e^{j0.2}$, $\bar{\lambda}_{12} = -1.8$, $\bar{\lambda}_{21} = -2.1e^{j0.6}$ and $\bar{\lambda}_{22} = 5.4e^{-j0.4}$. This phasor choice respects condition (ii) of Lemma 3.1 since u_1 and u_2 are linearly independent. We indeed observe that the phasor matrix \mathcal{L} in (3.14) has here the following expression and is full row rank:

$$\mathcal{L} = \begin{pmatrix} 0.72e^{j0.2} & 0.72e^{-j0.2} & -1.81 & -1.81 \\ -2.16e^{j0.4} & -2.16e^{-j0.4} & 5.43e^{j0.4} & 5.43e^{-j0.4} \end{pmatrix}$$

Even though all three necessary conditions are respected, the input regressor $\phi_{\mathbf{u}}$ is nevertheless not PE since we can observe that $u_1(t) = -3u_2(t-2) \forall t$ and that both $u_1(t)$ and $u_2(t-2)$ lie in $\phi_{\mathbf{u}}$. This can be also verified by constructing the corresponding matrix $\mathbf{T}^{(sin)}$:

$$\mathbf{T}^{(sin)} = \begin{pmatrix} \mathbf{T}_1^{(sin)} \\ \mathbf{T}_2^{(sin)} \end{pmatrix} = \begin{pmatrix} 0.72e^{j0.2} & 0.72e^{-j0.2} & -1.81 & -1.81 \\ 0.72e^{j0.1} & 0.72e^{-j0.1} & -1.81e^{j0.2} & -1.81e^{-j0.2} \\ -2.16e^{j0.3} & -2.16e^{-j0.3} & 5.43e^{j0.2} & 5.43e^{-j0.2} \\ -2.16e^{j0.2} & -2.16e^{-j0.2} & 5.43 & 5.43 \end{pmatrix}$$

We observe that the fourth row is obtained by multiplying the first one by -3 , i.e., $\text{rank}(\mathbf{T}^{sin}) < \nu = 4$. Consequently, $\phi_{\mathbf{u}}$ is not PE with this phasor choice. This is however a consequence of a particularly unlucky choice of phasors. To show this, let us, e.g., just modify the phasor $\bar{\lambda}_{11}$ to $\bar{\lambda}_{11} = e^{j0.2}$ and the matrix $\mathbf{T}^{(sin)}$ then becomes full row rank. In fact, $\mathbf{T}^{(sin)}$ will be full row rank for the majority of the phasor choices. To show this, let us generate randomly the amplitudes and the phase shifts: we have generated 1000 realizations for the four amplitudes λ_{kl} and the four phase shifts ψ_{kl} ($k = 1, 2$, $l = 1, 2$) using independent continuous uniform distributions, the amplitudes varying between 1 and 4 and the phase shifts between 0 and 2π . For these 1000 realizations, the rank of the corresponding matrix $\mathbf{T}^{(sin)}$ was indeed equal to $\nu = 4$. Consequently, the 1000 realizations of the input vector \mathbf{u} would yield informative data for any model structure corresponding to the regressor $\phi_{\mathbf{u}}$ used in this example.

3.5 Filtered white noise excitation yielding informativity

As opposed to the multisine case, condition (i) of Lemma 3.1 is always respected in the filtered white noise case. However, the condition on the relation between the number of rows and of the number of nonzero columns in a full row rank matrix will, like in the multisine case, lead to a necessary condition for $\phi_{\mathbf{u}}$ to be PE. This condition will impose a constraint on the complexity of $\mathbf{M}(z)$. Let us recall that, for the filtered white noise case, there can be some zero-columns among the ξ columns of $\mathbf{T}^{(fwn)}$ (see Remark 3.4).

Lemma 3.4. *Consider a regression vector $\phi_{\mathbf{u}}$ of dimension ν (see Theorem 3.3) and that corresponds to an input vector \mathbf{u} as in (2.5). Suppose that the matrix $\mathbf{T}^{(fwn)}$ of dimension $\nu \times \xi$ corresponding to this input vector (see Section 3.3.3) has χ nonzero columns ($\chi \leq \xi$). Then, for $\phi_{\mathbf{u}}$ to be PE, it is necessary to have that*

$$\chi \geq \nu$$

Proof. The number of nonzero columns in $\mathbf{T}^{(fwn)}$ is χ while its number of rows is given by ν . Consequently, it is necessary that $\chi \geq \nu$ for $\mathbf{T}^{(fwn)}$ to be full row rank. \square

As shown in Section 3.3.3, the larger the complexity of the filters in $\mathbf{M}(z)$ and the dimension f of the vector \mathbf{v} are, the larger the number χ of nonzero columns in $\mathbf{T}^{(fwn)}$ is (which is, in the vast majority of the cases, equal to ξ , i.e., the dimension of $\varphi^{(fwn)}$). In Example 3.3, this necessary condition is respected ($\chi = 10 > \nu = 6$) which is logical since we already observed that $\mathbf{T}^{(fwn)}$ is full row rank. It is however important to stress that $\chi \geq \nu$ is only a necessary condition for $\phi_{\mathbf{u}}$ to be PE (even if we add condition (ii) of Lemma 3.1). In other words, even though the complexity of $\mathbf{M}(z)$ coupled with the dimension of \mathbf{v} is sufficient to ensure $\chi \geq \nu$, there are particular values of the filter matrix $\mathbf{M}(z)$ that yield a rank deficient $\mathbf{T}^{(fwn)}$. Similarly as for the multisine case, this happens for very specific choices of the corresponding filter coefficients and we can avoid this phenomenon by generating these filter coefficients in a random manner (and checking subsequently whether the corresponding $\mathbf{T}^{(fwn)}$ is full row rank).

Example 3.5. *Consider an input regressor $\phi_{\mathbf{u}}$ with $n_u = 3$ and with $(\rho_1, \rho_2, \rho_3) = (0, 1, 1)$ and $(\mu_1, \mu_2, \mu_3) = (1, 2, 2)$. The input regressor $\phi_{\mathbf{u}}$ is thus given by*

$$\phi_{\mathbf{u}}(t) = \begin{pmatrix} u_1(t) \\ u_1(t-1) \\ u_2(t-1) \\ u_2(t-2) \\ u_3(t-1) \\ u_3(t-2) \end{pmatrix}$$

and is of dimension $\nu = 6$. Consider the following input vector class described as follows:

$$u(t) = \begin{pmatrix} l_{11}^{(1)} z^{-1} + l_{11}^{(2)} z^{-2} + l_{11}^{(3)} z^{-3} + l_{11}^{(4)} z^{-4} + l_{11}^{(5)} z^{-5} \\ l_{21}^{(0)} + l_{21}^{(1)} z^{-1} + l_{21}^{(2)} z^{-2} + l_{21}^{(3)} z^{-3} \\ l_{31}^{(0)} + l_{31}^{(1)} z^{-1} + l_{31}^{(2)} z^{-2} + l_{31}^{(3)} z^{-3} \end{pmatrix} v_1(t)$$

where v_1 is a white noise and where all $l_{kq}^{(h)}$ are nonzero. By following the procedure of the rewriting of $\phi_{\mathbf{u}}$ in Section 3.3.3, the vector $\varphi^{(fwn)}$ corresponding to this matrix $\mathbf{M}(z) = \mathbf{L}(z)$ is a vector of dimension $\xi = 6$ containing $v_1(t-1)$, $v_1(t-2)$, $v_1(t-3)$, $v_1(t-4)$, $v_1(t-5)$ and $v_1(t-6)$. Since the expansion (3.9) of the two polynomial filters in $\mathbf{M}(z)$ do not miss any coefficients, we will have $\chi = \xi = 6$ and, consequently, the necessary condition of Lemma 3.4 is respected. Let us choose the following values for $l_{kq}^{(h)}$

$$\mathbf{M}(z) = \mathbf{L}(z) = \begin{pmatrix} -z^{-1} + 2z^{-2} - 0.4z^{-3} - 5.7z^{-4} - 1.4z^{-5} \\ 1 - z^{-1} + 3.1z^{-2} + 0.9z^{-3} \\ 1 + 2.7z^{-1} - 4.8z^{-2} - 1.4z^{-3} \end{pmatrix}$$

We observe that the rows of $\mathbf{M}(z)$ are linearly independent and so \mathbf{u} is PE (condition (ii) of Lemma 3.1). However, we also observe that $u_1(t) = -u_2(t-1) + u_3(t-2) \forall t$ and that $u_1(t)$, $u_2(t-1)$ and $u_3(t-2)$ lie in $\phi_{\mathbf{u}}$. Consequently, the input regressor $\phi_{\mathbf{u}}$ is not PE and the corresponding matrix $\mathbf{T}^{(fwn)}$ cannot be full row rank. Let us verify this by calculating $\mathbf{T}^{(fwn)}$. In our case, $f = 1$, $n_u = 3$, $\alpha_1 = 1$, $\beta_1 = 6$. The matrices $\tilde{\mathbf{T}}_{11}^{(fwn)}$, $\tilde{\mathbf{T}}_{21}^{(fwn)}$ and $\tilde{\mathbf{T}}_{31}^{(fwn)}$ are given by

$$\begin{aligned} \tilde{\mathbf{T}}_{11}^{(fwn)} &= \begin{pmatrix} -1 & 2 & -0.4 & -5.7 & -1.4 & 0 \\ 0 & -1 & 2 & -0.4 & -5.7 & -1.4 \end{pmatrix} \\ \tilde{\mathbf{T}}_{21}^{(fwn)} &= \begin{pmatrix} 1 & -1 & 3.1 & 0.9 & 0 \\ 0 & 1 & -1 & 3.1 & 0.9 \end{pmatrix} \\ \tilde{\mathbf{T}}_{31}^{(fwn)} &= \begin{pmatrix} 1 & 2.7 & -4.8 & -1.4 & 0 \\ 0 & 1 & 2.7 & -4.8 & -1.4 \end{pmatrix} \end{aligned}$$

Therefore, the matrix $\mathbf{T}^{(fwn)}$ is given by

$$\mathbf{T}^{(fwn)} = \begin{pmatrix} \mathbf{0}_{2 \times 0} & \tilde{\mathbf{T}}_{11}^{(fwn)} & \mathbf{0}_{2 \times 0} \\ \mathbf{0}_{2 \times 0} & \tilde{\mathbf{T}}_{21}^{(fwn)} & \mathbf{0}_{2 \times 1} \\ \mathbf{0}_{2 \times 0} & \tilde{\mathbf{T}}_{31}^{(fwn)} & \mathbf{0}_{2 \times 1} \end{pmatrix} = \begin{pmatrix} -1 & 2 & -0.4 & -5.7 & -1.4 & 0 \\ 0 & -1 & 2 & -0.4 & -5.7 & -1.4 \\ 1 & -1 & 3.1 & 0.9 & 0 & \underline{0} \\ 0 & 1 & -1 & 3.1 & 0.9 & \underline{0} \\ 1 & 2.7 & -4.8 & -1.4 & 0 & \underline{0} \\ 0 & 1 & 2.7 & -4.8 & -1.4 & \underline{0} \end{pmatrix}$$

where the underlined zeros $\underline{0}$ correspond to the zeros of $\mathbf{0}_{(\mu_k - \rho_k + 1) \times (\beta_q - n_{kq} - \mu_k)}$ and $\mathbf{0}_{(\mu_k - \rho_k + 1) \times (d_{kq} + \rho_k - \alpha_q)}$ for all $\mathbf{T}_{kq}^{(fwn)}$. We have indeed that the last row is equal to the sum of the third row and the first one and so this matrix cannot be full row rank.

However, if we just modify $l_{21}^{(0)}$ to $l_{21}^{(0)} = 3$, the corresponding matrix $\mathbf{T}^{(fwn)}$ becomes full row rank.

The phenomenon observed above is thus a consequence of an unlucky choice of the filter coefficients. To confirm this, let us, like in Example 3.5, generate randomly the 13 coefficients $l_{kq}^{(h)}$ using independent zero-mean uniform distributions where each l_{kq} can vary between -1 and 1 . For 1000 of such realizations, the rank of the corresponding matrix $\mathbf{T}^{(fwn)}$ was indeed equal to $\nu = 6$. Consequently, the 1000 realizations of the input vector \mathbf{u} would yield informative data for any model structure corresponding to the regressor $\phi_{\mathbf{u}}$ used in this example.

3.6 Checking informativity without input regressor

In some particular cases of the input excitation \mathbf{u} , we do not need to consider the input regressor of Theorem 3.3 to check whether this input vector yields informative data.

For the filtered white noise case, such a situation is the one described in Theorem 2.1 (Section 2.6).

For the multisine case, it is, e.g., the case for the multisines described in the following lemma.

Lemma 3.5. *Consider the data set \mathcal{Z}^∞ in (2.10) obtained by applying a multisine input vector \mathbf{u} (see (2.4)) on a MISO system with n_u inputs. Consider a full-order model structure \mathcal{M} for this MISO system and the notations introduced in Section 3.3.1. Then, the data set \mathcal{Z}^∞ is informative with respect to \mathcal{M} if, for each $k = 1, \dots, n_u$,*

(i) *the scalar input u_k contains a number s_k of nonzero sinusoids that satisfies:*

- $s_k \geq \frac{\deg(B_k) - \rho_k + 1}{2}$ when \mathcal{M} is ARX or FIR.
- $s_k \geq \frac{\deg(B_k) - \rho_k + \deg(F_k) + 1}{2}$ when \mathcal{M} is BJ or OE.

(ii) *the s_k frequencies of the sinusoids in u_k are different from the s_m frequencies of the sinusoids in all u_m ($m \neq k$), i.e., the signals u_k are mutually uncorrelated.*

Moreover, if (i) and (ii) are respected, \mathcal{Z}^∞ is informative whatever the values of the s_k (nonzero) frequencies, the s_k (nonzero) amplitudes and the s_k phase shifts defining u_k ($k = 1, \dots, n_u$).

Proof. See Appendix B.7 for the proof. □

Remark 3.6. *A similar result for MISO ARMAX model structure is given in Appendix B.8.*

For the FIR/ARX case, the conditions (i) in Lemma 3.5 correspond to the necessary conditions in (3.6). In the BJ/OE case, the regressor $\phi_{\mathbf{u}}$ (see (3.5)) corresponding to the multisine input vectors defined in Lemma 3.5 will not be necessarily PE since the necessary conditions for this property (i.e. (3.6) and (3.15)) will not be always fulfilled. This is however not a contradiction since a persistently exciting $\phi_{\mathbf{u}}$ is only a sufficient condition for data informativity in the MISO BJ/OE case.

Remark 3.7. *Note that the situations in Theorem 3.1 and Lemma 3.5 are certainly not the only situations for which we can derive the data informativity property without using Theorem 3.3. Another example of such a situation will, e.g., be given in Chapter 5 (see Lemma 5.1).*

3.7 A synthetic verification approach of the data informativity for the open-loop identification of MIMO systems

In the previous sections, we have extensively analyzed the data informativity for the open-loop identification of MISO systems. In this section, we give a procedure that the user can follow to verify the data informativity in open-loop given an input vector \mathbf{u} and a MIMO model structure \mathcal{M} .

First, if the input vector is of the type (2.5) with $f \geq n_u$ and with $\text{rank}(\mathbf{M}(e^{j\omega})) = n_u$ for almost all frequencies ω in $] - \pi, \pi]$, then we have data informativity by virtue of Lemma 2.1 in Section 2.6 ($\Phi_{\mathbf{u}}(\omega) \succ 0$ for almost all ω in $] - \pi, \pi]$).

If it is not the case, as said below Theorem 3.2, we must check, for each $i = 1, \dots, n_y$, the informativity of \mathcal{Z}_i^∞ with respect to the n_y MISO model structures \mathcal{M}_i that can be derived from \mathcal{M} . For each $i = 1, \dots, n_y$, this can be done using the results presented in Sections 3.3.2, 3.4, 3.5, 3.6.

In the next two paragraphs, we give a procedure for the multisine case and another one is for the filtered white noise case when we cannot conclude with Lemma 2.1.

For multisine inputs \mathbf{u} of the type (2.4), if \mathbf{u} satisfies the conditions in Lemma 3.5, then \mathcal{Z}_i^∞ is informative. If it is not the case, we construct the matrix $\mathbf{T}^{(sin)}$ corresponding to the MISO model structure \mathcal{M}_i and we check whether it is full row rank. Note that $\mathbf{T}^{(sin)}$ can only be full row rank if (3.6) and (3.15) are satisfied. If $\mathbf{T}^{(sin)}$ is full row rank, then \mathcal{Z}_i^∞ is informative with respect to \mathcal{M}_i .

For filtered white noise excitations \mathbf{u} of the type (2.5), we construct $\mathbf{T}^{(fwn)}$ corre-

sponding the MISO model structure \mathcal{M}_i and we check whether this matrix is full row rank. Note that this can only be the case if the condition in Lemma 3.4 is satisfied. If $\mathbf{T}^{(fun)}$ is full row rank, then \mathcal{Z}_i^∞ is informative with respect to \mathcal{M}_i .

Of course, it can happen that, for one i in these procedures, the matrix $\mathbf{T}^{(fun)}$ (or $\mathbf{T}^{(sin)}$ in the multisine case) is not full row rank. Then, if \mathcal{M} is ARX or FIR, we know that \mathcal{Z}_i^∞ is not informative with respect to \mathcal{M}_i and thus that \mathcal{Z}^∞ is not informative with respect to \mathcal{M} . If \mathcal{M} is BJ or OE, then we cannot conclude neither that the data are informative nor that the data are not informative (since Theorem 3.3 only gives a sufficient condition for these model structures).

3.8 Numerical example

In this section, we are going to illustrate the results developed in this chapter on a numerical example. We will, e.g., illustrate that the result of Theorem 3.3 is not overly conservative for the data informativity w.r.t. BJ/OE model structures.

3.8.1 True system, model structure and input regressor

Consider the following MISO OE system \mathcal{S} with $n_u = 2$ inputs given by

$$y(t) = \underbrace{\begin{pmatrix} 7.11z^{-1} & 6.45 \\ 1 - 0.51z^{-1} & 1 + 0.39z^{-1} \end{pmatrix}}_{\mathbf{G}_0(z)} \underbrace{\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}}_{\mathbf{u}(t)} + e(t)$$

where the covariance σ_0 of the zero-mean white noise e is here equal to $\sigma_0 = 0.0001$.

We will identify \mathcal{S} within a full-order model structure \mathcal{M} , where all the transfer functions in $\mathbf{G}(z, \tilde{\boldsymbol{\theta}}) = (G_1(z, \tilde{\boldsymbol{\theta}}_1), G_2(z, \tilde{\boldsymbol{\theta}}_2))$ are independently parametrized with delays given by $(\rho_1, \rho_2) = (1, 0)$. They can therefore be put into the form given in (3.4) with

$$\begin{aligned} \mathring{B}_1(z, \tilde{\boldsymbol{\theta}}_1) &= \tilde{\theta}_{1,1} & F_1(z, \tilde{\boldsymbol{\theta}}_1) &= 1 + \tilde{\theta}_{1,2}z^{-1} \\ \mathring{B}_2(z, \tilde{\boldsymbol{\theta}}_2) &= \tilde{\theta}_{2,1} & F_2(z, \tilde{\boldsymbol{\theta}}_2) &= 1 + \tilde{\theta}_{2,2}z^{-1} \end{aligned}$$

We have $\boldsymbol{\theta}^T = \tilde{\boldsymbol{\theta}}^T = (\tilde{\boldsymbol{\theta}}_1^T \ \tilde{\boldsymbol{\theta}}_2^T) = (\tilde{\theta}_{1,1} \ \tilde{\theta}_{1,2} \ \tilde{\theta}_{2,1} \ \tilde{\theta}_{2,2})^T$. Let us now determine the regressor $\boldsymbol{\phi}_u$ corresponding to this MISO OE model structure \mathcal{M} . For this purpose, let us calculate the scalars μ_k (see Theorem 3.3)

- $\mu_1 = \deg(B_1) + \deg(F_1) + 2\deg(F_2) = 4$
- $\mu_2 = \deg(B_2) + \deg(F_2) + 2\deg(F_1) = 3$

Consequently, the input regressor is given by $\phi_{\mathbf{u}} = (\phi_{u_1}^T, \phi_{u_2}^T)^T$ with

$$\begin{aligned}\phi_{u_1}^T(t) &= (u_1(t-1) \quad u_1(t-2) \quad u_1(t-3) \quad u_1(t-4)) \\ \phi_{u_2}^T(t) &= (u_2(t) \quad u_2(t-1) \quad u_2(t-2) \quad u_2(t-3))\end{aligned}$$

The dimension ν of $\phi_{\mathbf{u}}$ is here equal to $\nu = 8$.

3.8.2 Identification with multisine

In this paragraph, we consider that \mathbf{u} is a multisine given by (2.4).

As shown in Theorem 3.3, the persistency of $\phi_{\mathbf{u}}$ is only a sufficient condition for data informativity w.r.t. OE model structures. However, we will show that this sufficient condition is not too restrictive. This will be done as follows. We will present an example where $\phi_{\mathbf{u}}$ is not PE and where the data set \mathcal{Z}^∞ is not informative. Moreover, we will also show that if the dimension of $\phi_{\mathbf{u}}$ is reduced by removing $u_1(t-4)$ from this regressor, this reduced $\phi_{\mathbf{u}}$ would be PE. In this example, the regressor $\phi_{\mathbf{u}}$ in (3.5) is in fact the only one that is not PE.

Case 1: We choose u_1 as a multisine containing $s = s_1 = 4$ frequencies $\omega_1 = 0.001\pi$, $\omega_2 = 0.01\pi$, $\omega_3 = 0.1\pi$ and $\omega_4 = 0.5\pi$ and with phasors $\bar{\lambda}_{1l} = 1$ ($l = 1, \dots, 3$). The signal u_2 is chosen in such a way that

$$(G_2(z, \tilde{\theta}') - G_2(z, \tilde{\theta}_0))u_2(t) = -(G_1(z, \tilde{\theta}') - G_1(z, \tilde{\theta}_0))u_1(t)$$

for $\tilde{\theta}' = (3, 0.4, 10, 0.2)^T$ different from $\tilde{\theta}_0 = (7.11, -0.51, 6.45, 0.39)^T$. This yields:

$$\mathcal{L}^T = \begin{pmatrix} \bar{\lambda}_{11} & \bar{\lambda}_{21} \\ \bar{\lambda}_{12} & \bar{\lambda}_{22} \\ \bar{\lambda}_{13} & \bar{\lambda}_{23} \\ \bar{\lambda}_{14} & \bar{\lambda}_{24} \end{pmatrix} = \begin{pmatrix} 1 & 3.3488e^{j3.1344} \\ 1 & 3.3451e^{j3.0696} \\ 1 & 3.0257e^{j2.4511} \\ 1 & 1.2333e^{j0.7279} \end{pmatrix}$$

This phasor and frequency choice respects both necessary conditions (3.6) and (3.15). Moreover, the phasor matrix \mathcal{L} is full row rank.

It is however clear that this input vector choice does not yield informative data (see Theorem 3.1) and this fact can be detected via Theorem 3.3. Indeed, the regressor $\phi_{\mathbf{u}}$ defined in (3.5) is not PE since the matrix $\mathbf{T}^{(sin)}$ corresponding to this $\phi_{\mathbf{u}}$ and the above phasor and frequency choice has a rank equal to 7 while $\nu = 8$. This regressor is furthermore the one with the smallest number of signals that can detect the fact that the data are not informative w.r.t. \mathcal{M} . Indeed, if we consider all sub-vectors of $\phi_{\mathbf{u}}$, we obtain signal vectors that are PE. For instance, the subvector of $\phi_{\mathbf{u}}$ of dimension 7 obtained by removing $u_1(t-4)$ from $\phi_{\mathbf{u}}$ is PE. It is important to stress that we can generate infinite number of pathological cases as the previous one by considering all $\theta' \in \mathcal{D}_\theta$.

Case 2: Consider now a multisine input vector \mathbf{u} with the same $s = 4$ frequencies as in Case 1, but with the following choice of phasors:

$$\begin{pmatrix} \bar{\lambda}_{11} & \bar{\lambda}_{21} \\ \bar{\lambda}_{12} & \bar{\lambda}_{22} \\ \bar{\lambda}_{13} & \bar{\lambda}_{23} \\ \bar{\lambda}_{14} & \bar{\lambda}_{24} \end{pmatrix} = \begin{pmatrix} -1.72e^{j1.54} & 4.85e^{-j2.37} \\ -0.573 & 2.46e^{-j0.123} \\ 4.39e^{-j0.02} & 7.63e^{j3.104} \\ 1.46e^{j0.14} & 3.19j \end{pmatrix}$$

For this choice of phasors, the matrices $\mathbf{T}^{(sin)}$ is full row rank, i.e., $\text{rank}(\mathbf{T}^{(sin)}) = 8$. This multisine input vector will thus yield informative data with respect to \mathcal{M} . However, in order to guarantee that $\mathbf{T}^{(sin)}$ has more columns than rows, we need at least $s = 4$ frequencies to guarantee the persistency of $\phi_{\mathbf{u}}$. We will see in the next case that we can reduce it to $s = 3$ with the result in Lemma 3.5.

Case 3: Consider now a multisine input vector with $s = 3$ sinusoid frequencies $\omega_1 = 0.001\pi$, $\omega_2 = 0.01\pi$, $\omega_3 = 0.1\pi$ and with the following phasor choice:

$$\begin{pmatrix} \bar{\lambda}_{11} & \bar{\lambda}_{21} \\ \bar{\lambda}_{12} & \bar{\lambda}_{22} \\ \bar{\lambda}_{13} & \bar{\lambda}_{23} \end{pmatrix} = \begin{pmatrix} 1.83e^{-j1.75} & 0 \\ -2.36e^{j0.19} & 0 \\ 0 & 4.02e^{j2.51} \end{pmatrix}$$

The scalar inputs u_1 and u_2 do not share any common frequencies (i.e., they are mutually uncorrelated) and we have that $s_1 = 2 \geq (\deg(B_1) - \rho_1 + \deg(F_1) + 1)/2$ and $s_2 = 1 \geq (\deg(B_2) - \rho_2 + \deg(F_2) + 1)/2$. Consequently, we are in the case of the class of multisine input vector proposed in Lemma 3.5. This class yields informative data independently of the choice of (nonzero) phasors and (nonzero) frequencies.

3.8.3 Identification with filtered white noise

In this paragraph, we consider that \mathbf{u} is a filtered white noise given by (2.5).

Case 4: Consider an input vector generated with $f = 1$ white noise $v_1(t)$, i.e.,

$$\mathbf{u}(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix} = \underbrace{\begin{pmatrix} M_{11}(z) \\ M_{21}(z) \end{pmatrix}}_{\mathbf{M}(z)} v_1(t)$$

where

$$M_{11}(z) = \frac{-4z^{-1} + 9.5z^{-2} - 6.1z^{-3}}{1 - 1.1z^{-1} - 0.22z^{-2} + 0.41z^{-3} - 0.05z^{-4}}$$

$$M_{21}(z) = 5z^{-1}$$

We rewrite $\mathbf{M}(z)$ as follows $\mathbf{M}(z) = \mathbf{L}(z)/w(z)$ where $w(z) = 1 - 1.1z^{-1} - 0.22z^{-2} + 0.41z^{-3} - 0.05z^{-4}$ and

$$\mathbf{L}(z) = \begin{pmatrix} -4z^{-1} + 9.5z^{-2} - 6.1z^{-3} \\ 5z^{-1} - 5.5z^{-2} - 1.1z^{-3} + 2.05z^{-4} - 0.25z^{-5} \end{pmatrix}$$

Let us calculate $\xi = \chi$ (since all coefficients $l_{kq}^{(h)}$ are nonzero) to verify the necessary conditions in Lemma 3.4. For this purpose, let us compute α_1 and β_1 defined in (3.11): $\alpha_1 = 1$ and $\beta_1 = 8$ and so $\chi = \sum_{q=1}^r (\beta_q - \alpha_q) + q = \beta_1 - \alpha_1 + 1 = 8$.

Consequently, we have that $\chi \geq \nu$, i.e., the condition of Lemma 3.4 is respected. We verify the rank of the corresponding matrix $\mathbf{T}^{(fwn)}$ and we obtain $\text{rank}(\mathbf{T}^{(fwn)}) = 8 = \nu$. Consequently, this input vector will yield informative data.

Case 5: The choice to use $f = 1$ white noise to generate \mathbf{u} in Case 4 forced us to use entries with high order for $\mathbf{M}(z)$ in order to get the informativity. Let us now consider an input vector \mathbf{u} with $f = n_u = 2$ and where $\mathbf{M}(z) = \mathbf{L}(z)/w(z)$ is characterized by $w(z) = 1 - 0.5z^{-1}$ and

$$\mathbf{L}(z) = \begin{pmatrix} -0.25z^{-1} + 7.81z^{-2} & -3.45 + 2.1z^{-1} \\ 0.73z^{-1} & 0.6z^{-2} + 0.59z^{-3} \end{pmatrix}$$

We are here in the situation of Lemma 2.1 and we can easily verify that the frequency response $\mathbf{M}(e^{j\omega})$ of $\mathbf{M}(z)$ is full row rank for all $\omega \in] - \pi, \pi]$. Therefore, from Theorem 2.1, the data set \mathcal{Z}^∞ is informative w.r.t. \mathcal{M} .

3.8.4 Consideration of a second channel

Let us suppose now that the system \mathcal{S} has two channels and is given by

$$\underbrace{\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix}}_{\mathbf{y}(t)} = \underbrace{\begin{pmatrix} \frac{7.11z^{-1}}{1 - 0.51z^{-1}} & \frac{6.45}{1 + 0.39z^{-1}} \\ 1 - 1.84z^{-1} + 2.54z^{-2} & -3.91z^{-1} + 5.02z^{-2} \end{pmatrix}}_{\mathbf{G}_0(z)} \underbrace{\begin{pmatrix} u_1(t) \\ u_2(t) \end{pmatrix}}_{\mathbf{u}(t)} + \underbrace{\begin{pmatrix} e_1(t) \\ e_2(t) \end{pmatrix}}_{\mathbf{e}(t)}$$

where the covariance matrix Σ_0 of \mathbf{e} is here equal to $\Sigma_0 = 0.0001\mathbf{I}_2$. This MIMO system is composed of two MISO ones where the first one is the MISO OE system studied in Section 3.8.1 and the second one is a MISO FIR system.

We will identify the MIMO system within a full-order MIMO OE model structure \mathcal{M}_{mimo} where all the transfer functions in $\mathbf{G}(z, \boldsymbol{\theta})$ are independently parametrized with delays given by $\rho_{11} = 1$, $\rho_{12} = 0$, $\rho_{21} = 0$ and $\rho_{22} = 1$. The transfer functions of the first row of $\mathbf{G}(z, \boldsymbol{\theta})$ are parametrized as given in Section 3.8.1, except that now we add the index 1 to specify that it is linked to the first channel. As a consequence, \hat{B}_1 , \hat{B}_2 , F_1 and F_2 become \hat{B}_{11} , \hat{B}_{12} , F_{11} and F_{12} respectively, $\tilde{\boldsymbol{\theta}}_1$ and $\tilde{\boldsymbol{\theta}}_2$ become $\tilde{\boldsymbol{\theta}}_{11}$ and $\tilde{\boldsymbol{\theta}}_{12}$ respectively and $\tilde{\boldsymbol{\theta}}$ becomes $\tilde{\boldsymbol{\theta}}_1$. For the second channel, the entries of $\mathbf{G}_2(z, \boldsymbol{\theta})$ are parametrized as given in Section 2.5

$$\begin{aligned} \hat{B}_{21}(z, \tilde{\boldsymbol{\theta}}_{21}) &= \tilde{\theta}_{21,1} + \tilde{\theta}_{21,2}z^{-1} + \tilde{\theta}_{21,3}z^{-2} & F_{21}(z, \tilde{\boldsymbol{\theta}}_{21}) &= 1 \\ \hat{B}_{22}(z, \tilde{\boldsymbol{\theta}}_{22}) &= \tilde{\theta}_{22,1} + \tilde{\theta}_{22,2}z^{-1} & F_{22}(z, \tilde{\boldsymbol{\theta}}_{22}) &= 1 \end{aligned}$$

We have $\tilde{\boldsymbol{\theta}}_2^T = (\tilde{\boldsymbol{\theta}}_{21}^T \ \tilde{\boldsymbol{\theta}}_{22}^T) = (\tilde{\theta}_{21,1} \ \tilde{\theta}_{21,2} \ \tilde{\theta}_{21,3} \ \tilde{\theta}_{22,1} \ \tilde{\theta}_{21,2})^T$. The input regressor for the second channel is then given by $\boldsymbol{\phi}_u = (\boldsymbol{\phi}_{u_1}^T, \boldsymbol{\phi}_{u_2}^T)^T$ with

$$\begin{aligned}\boldsymbol{\phi}_{u_1}^T(t) &= (u_1(t) \ u_1(t-1) \ u_1(t-2)) \\ \boldsymbol{\phi}_{u_2}^T(t) &= (u_2(t-1) \ u_2(t-2))\end{aligned}$$

The dimension ν of the regressor $\boldsymbol{\phi}_u$ corresponding to the second channel is here equal to $\nu = 5$. By following the same approach as for the first channel, we can prove that, for the second channel, the matrix $\mathbf{T}^{(sin)}$ for Cases 2 and 3 and the matrix $\mathbf{T}^{(fwn)}$ for Case 4 are full row rank. Therefore, for Cases 2, 3 and 4, the matrices $\mathbf{T}^{(sin)}$ and $\mathbf{T}^{(fwn)}$ are full row rank for both channels: the data are informative with respect to the MIMO model structure \mathcal{M}_{mimo} .

For Case 5, since the input power spectrum matrix $\Phi_u(\omega)$ has been proven to be positive definite at all frequencies ω in $]-\pi, \pi]$, then the informativity is always guaranteed for any MIMO model structures (see Theorem 2.1) and thus also for \mathcal{M}_{mimo} .

3.8.5 Verification of the results with Monte-Carlo simulations

In order to confirm that the input choice in Cases 2, 3, 4 and 5 yield informative data, we have applied each of these input vectors to the MIMO true system in Section 3.8.4 in 1000 identification experiments (with different realizations of the white noise \mathbf{e}) and we have identified the estimate $\hat{\boldsymbol{\theta}}_N$ (see (2.9)) for each experiment. We initialize each identification computation with a random parameter vector $\boldsymbol{\theta}_{init}$ that lies in the set $\{\boldsymbol{\theta}_{init} \in \mathcal{D}_\theta \mid \|\boldsymbol{\theta}_{init} - \boldsymbol{\theta}_0\|^2 < 0.01\}$. For each input vector, we have computed the mean of these 1000 estimates. The results are given in Tables 3.1 and 3.2 and we have in each case observed that this mean is almost equal to the true parameter vector $\boldsymbol{\theta}_0$, suggesting the consistency.

Table 3.1: Mean of the identified parameter vector for the first channel with $N = 1000$ for different cases over 1000 experiments.

	$\tilde{\theta}_{11,1}$	$\tilde{\theta}_{11,2}$	$\tilde{\theta}_{12,1}$	$\tilde{\theta}_{12,2}$
$\boldsymbol{\theta}_0$	7.11	-0.51	6.45	0.39
Case 2	7.1094	-0.5100	6.4489	0.3897
Case 3	7.1035	-0.5105	6.4479	0.3896
Case 4	7.1100	-0.5098	6.4999	0.3900
Case 5	7.1100	-0.5100	6.4500	0.3900

Table 3.2: Mean of the identified parameter vector for the second channel with $N = 1000$ for different cases over 1000 experiments.

	$\tilde{\theta}_{21,1}$	$\tilde{\theta}_{21,2}$	$\tilde{\theta}_{21,3}$	$\tilde{\theta}_{22,1}$	$\tilde{\theta}_{22,2}$
θ_0	1	-1.84	2.54	-3.91	5.02
Case 2	0.9926	-1.8362	2.5409	-3.9104	5.0204
Case 3	0.9985	-1.8242	2.5325	-3.9054	5.0158
Case 4	0.9998	-1.8402	2.5397	-3.9104	5.0197
Case 5	0.9999	-1.8400	2.5400	-3.9098	5.0196

3.9 Conclusion of Chapter 3

The purpose of Chapter 3 was to address Research Question 1 by developing conditions to verify the data informativity for the open-loop identification of MIMO model structures and this for multisine and filtered white noise excitations \mathbf{u} . The main result developed in this chapter is the verification of the data informativity by checking the fact that a certain matrix is full row rank for each channel. This matrix depends on the model structure complexity and on the excitation parametrizations. This new condition is clearly less conservative than the only condition available in the literature and given in Lemma 2.1. We also derived results that allow to verify the data informativity without checking the regressor persistency since this condition is conservative for BJ/OE model structures. Based on the results derived in this chapter, we will see in the next one that we can address Research Question 2.

Data Informativity for MIMO Systems in Closed-Loop

Contents

4.1	Introduction of Chapter 4	64
4.2	Simplification of the study	64
4.3	Rewriting of the left-hand side of (4.3)	66
4.3.1	Considered approach	66
4.3.2	Factorization of (4.3) into a polynomial matrix form	66
4.3.3	Main result for data informativity	68
4.3.4	Rewriting of $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))\Xi_i(z) \equiv \mathbf{0}$ into (4.14)	71
4.4	Rewriting of $\bar{E} [\ \Upsilon_{u,i}(z)\mathbf{r}(t)\ ^2] = 0$ into (4.15)	74
4.5	Interpretation of Theorem 4.4	75
4.6	A synthetic approach for the verification of the data informativity for the direct closed-loop identification of MIMO systems	77
4.7	Numerical example	78
4.7.1	True system to be identified and model structure	78
4.7.2	Factorization of the model structure with Lemma 4.1	79
4.7.3	Costless identification	80
4.7.4	Identification with external excitation ($\mathbf{r} \neq \mathbf{0}_{n_u \times 1}$)	81
4.7.5	Verification of the results with Monte-Carlo simulations	83
4.8	Conclusion of Chapter 4	84

4.1 Introduction of Chapter 4

In this chapter, we address Research Question 2 in Section 2.8 by deriving conditions to verify the data informativity of MIMO model structures defined in Section 2.5 for a given MIMO controller $\mathbf{K}(z)$ and a given external excitation \mathbf{r} of the type (2.14)-(2.15).

In Section 4.2, we will show that, like in the open-loop case, we can verify that an excitation vector \mathbf{r} will yield informative data with respect to a MIMO model structure in a channel-by-channel manner. This channel-by-channel verification will involve, for each channel, the verification of the rank of a matrix. The matrix for each channel will be made up of two parts: the first part will depend on the controller and the model structure and the second part will depend on the parametrization of the excitation vector \mathbf{r} and on the model structure. We will see that this second part will be constructed in the same way as the matrices $\mathbf{T}^{(sin)}$ and $\mathbf{T}^{(fwn)}$ in Chapter 3.

A synthesis of the results is given in Section 4.6. To illustrate the results developed in this chapter, a numerical MIMO example is provided in Section 4.7. Finally, we conclude in Section 4.8.

4.2 Simplification of the study

As mentioned above, we will now consider that the MIMO system \mathcal{S} in (2.2) with n_u inputs and n_y outputs is in the closed-loop configuration with a stabilizing controller $\mathbf{K}(z)$ as illustrated in Figure 2.1 (see Section 2.4) and we will derive conditions to verify the data informativity for the excitation vectors \mathbf{r} given in (2.14) and (2.15) (see Definition 2.3). As in the open-loop case, these conditions will depend on the model structure complexity and the richness of the excitation signal \mathbf{r} . We will also see that, in addition, these conditions will depend on the controller complexity (such as was observed in the SISO case [GBM08]).

To derive these conditions let us, as was done in Theorem 3.1 for the open-loop case, specify more precisely Definition 2.3.

Theorem 4.1. *Consider Definition 2.3 and, in particular, the set \mathcal{Z}_{cl}^∞ obtained via the application of the excitation signal \mathbf{r} to the closed-loop system depicted in Figure 2.1. Consider also one of the model structures \mathcal{M} defined in Section 2.5. Recall that \mathbf{r} and \mathbf{e} are assumed to be independent. For each $\Delta\mathbf{W}(z) \in \mathbf{\Delta}_W$ (see (2.11) in Definition 2.3), we define similarly $\Delta\mathbf{W}_y(z)$ and $\Delta\mathbf{W}_u(z)$ (see (2.7) and (2.8) respectively). Then, the data set \mathcal{Z}_{cl}^∞ in (2.10) is informative w.r.t. the model structure \mathcal{M} if and only if, for all $\Delta\mathbf{W}(z) = (\Delta\mathbf{W}_y(z), \Delta\mathbf{W}_u(z)) \in \mathbf{\Delta}_W$, the following property holds*

$$\left\{ \begin{array}{l} \Delta\mathbf{W}_y(z) - \Delta\mathbf{W}_u(z)\mathbf{K}(z) \equiv \mathbf{0}_{n_y \times n_y} \\ \bar{E} [\|\Delta\mathbf{W}_u(z)\mathbf{r}(t)\|^2] = 0 \end{array} \right\} \implies (\Delta\mathbf{W}_y, \Delta\mathbf{W}_u) \equiv (\mathbf{0}_{n_y \times n_y}, \mathbf{0}_{n_y \times n_u}) \quad (4.1)$$

Proof. See Appendix C.1 for the proof. \square

Remark 4.1. *The proof of Lemma 2.2 (that can also be found in [BGM10]) is in fact a straightforward consequence of Theorem 4.1 and is given in Appendix C.2.*

Thanks to the parametrization introduced in Section 2.5, we will show that we can simplify the result of Theorem 4.1 in the next theorem.

Theorem 4.2. *Consider Definition 2.3, Theorem 4.1 and one of the model structures \mathcal{M} defined in Section 2.5. Based on the set $\Delta_{\mathbf{W}}$ (see (2.11) in Definition 2.3), define the sets*

$$\Delta_{\mathbf{W},i} = \{\Delta \mathbf{W}_{i,:}(z) \mid \Delta \mathbf{W}_{i,:}(z) \text{ is the } i^{\text{th}} \text{ row of } \Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}}\} \quad (4.2)$$

for all $i = 1, \dots, n_y$. For each $\Delta \mathbf{W}_{i,:}(z) \in \Delta_{\mathbf{W},i}$, we define similarly $\Delta \mathbf{W}_{u,i}(z)$ and $\Delta \mathbf{W}_{y,i}(z)$ which are the i^{th} row of $\Delta \mathbf{W}_u(z)$ and $\Delta \mathbf{W}_y(z)$ respectively. Then, the data set \mathcal{Z}_{cl}^∞ is informative w.r.t. the model structure \mathcal{M} if and only if, for all $i = 1, \dots, n_y$, the following property holds for all $\Delta \mathbf{W}_{i,:}(z) = (\Delta \mathbf{W}_{y,i}(z), \Delta \mathbf{W}_{u,i}(z)) \in \Delta_{\mathbf{W},i}$:

$$\begin{cases} \Delta \mathbf{W}_{y,i}(z) - \Delta \mathbf{W}_{u,i}(z) \mathbf{K}(z) & \equiv \mathbf{0}_{1 \times n_y} \\ \bar{E} [|\Delta \mathbf{W}_{u,i}(z) \mathbf{r}(t)|^2] & = 0 \end{cases} \implies (\Delta \mathbf{W}_{y,i}, \Delta \mathbf{W}_{u,i}) \equiv (\mathbf{0}_{1 \times n_y}, \mathbf{0}_{1 \times n_u}) \quad (4.3)$$

Proof. See Appendix C.3 for the proof. \square

Theorem 4.2 allows to simplify the data informativity verification: it can be done channel-by-channel¹ (or output-by-output). Hence, in the sequel, we will consider an arbitrary i and restrict attention to (4.3) for that particular i .

To simplify the presentation in the sequel, we will often use shorthand notations for each transfer of the form $M(z)$, $M(z, \theta')$ and $M(z, \theta'')$ by dropping the argument and denote them by M , M' and M'' respectively.

¹However, we do not have the same simplification as in the open-loop case (Chapter 3). In the open-loop case, we can simplify the data informativity study w.r.t. a MIMO model structure into the analysis of the data informativity w.r.t. a MISO model structure. Here, note that the first equation of the left-hand side of (4.3) still contains the MIMO controller $\mathbf{K}(z)$.

4.3 Rewriting of the left-hand side of (4.3)

4.3.1 Considered approach

In its actual form, (4.3) is function of the rational transfer function matrix $(\Delta \mathbf{W}_{y,i}(z), \Delta \mathbf{W}_{u,i}(z)) \in \Delta \mathbf{W}_i$ where $\Delta \mathbf{W}_i$ is defined in (4.2). We need to transform (4.3) into a polynomial matrix form for the development of the conditions on the data informativity. Let us for this purpose first observe that the first equation of the left-hand side of (4.3) can be rewritten as follows

$$(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z)) \begin{pmatrix} \mathbf{I}_{n_y} \\ \mathbf{K}(z) \end{pmatrix} \equiv \mathbf{0}_{1 \times n_y} \quad (4.4)$$

By considering a right-factorization of the controller $\mathbf{K}(z)$ and a left-factorization of the term $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$, we will prove that we can transform (4.4) into a polynomial matrix form. Moreover, we will see that this left-factorization of $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ will allow us to transform the second equation of the left-hand side of (4.3) into a polynomial matrix form.

4.3.2 Factorization of (4.3) into a polynomial matrix form

Let us first deduce a left-factorization for the rational matrix $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ which should be valid for all $\Delta \mathbf{W}_i(z) = (\Delta \mathbf{W}_{y,i}(z), \Delta \mathbf{W}_{u,i}(z)) \in \Delta \mathbf{W}_i$ with $\Delta \mathbf{W}_i$ as defined in (4.2). The one that we will choose is obtained by putting all entries of $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ on the same denominator. It is given in the next lemma.

Lemma 4.1. *Consider the model structures defined in Section 2.5. For these model structures, the rational block-matrix $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ can be left-factorized into the form $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z)) = Q_i^{-1}(z)(\Upsilon_{y,i}(z), \Upsilon_{u,i}(z))$ where the row polynomial vector $\Upsilon_{y,i}(z)$ of dimension n_y has all its entries equal to 0 except possibly the i^{th} one denoted $\Upsilon_{y,ii}(z)$. The row polynomial vector $\Upsilon_{u,i}(z)$ is of dimension n_u . The scalar polynomials $Q_i(z)$ and $\Upsilon_{y,ii}(z)$ and the row polynomial vector $\Upsilon_{u,i}(z)$ are given by*

- for FIR: $Q_i = 1$, $\Upsilon_{y,ii} = 0$, $\Upsilon_{u,i} = \mathbf{B}_i'' - \mathbf{B}_i'$.
- for ARX: $Q_i = 1$, $\Upsilon_{y,ii} = A_i'' - A_i'$, $\Upsilon_{u,i} = \mathbf{B}_i'' - \mathbf{B}_i'$.
- for ARMAX: $Q_i = C_i'' C_i'$, $\Upsilon_{y,ii} = C_i' A_i'' - C_i'' A_i'$, $\Upsilon_{u,i} = C_i' \mathbf{B}_i'' - C_i'' \mathbf{B}_i'$.
- for OE: $Q_i = \prod_{k=1}^{n_u} F_{ik}' F_{ik}''$, $\Upsilon_{y,ii} = 0$ and $\Upsilon_{u,i} = (\Upsilon_{u,ik})_{k \in \llbracket 1, n_u \rrbracket}$ with

$$\Upsilon_{u,ik} = (B_{ik}'' F_{ik}' - B_{ik}' F_{ik}'') \prod_{l=1, l \neq k}^{n_u} F_{il}' F_{il}''$$

- for BJ: $Q_i = C_i' C_i'' \prod_{k=1}^{n_u} F_{ik}' F_{ik}''$, $\Upsilon_{y,ii} = D_i'' C_i' \prod_{k=1}^{n_u} F_{ik}' F_{ik}'' - D_i' C_i'' \prod_{k=1}^{n_u} F_{ik}' F_{ik}''$ and $\Upsilon_{u,i} = (\Upsilon_{u,ik})_{k \in \llbracket 1, n_u \rrbracket}$ with

$$\Upsilon_{u,ik} = D_i'' C_i' B_{ik}'' F_{ik}' \prod_{l=1, l \neq k}^{n_u} F_{il}' F_{il}'' - D_i' C_i'' B_{ik}' F_{ik}'' \prod_{l=1, l \neq k}^{n_u} F_{il}' F_{il}''$$

Proof. See Appendix C.4 for the proof. \square

For the controller $\mathbf{K}(z)$, we will consider the right-factorization consisting in bringing all entries of $\mathbf{K}(z)$ on the same denominator $J(z)$ ($J(z)$ is a polynomial). The common denominator $J(z)$ that we will choose² is the least common multiple of the denominators of the entries of $\mathbf{K}(z)$. The controller $\mathbf{K}(z)$ can be rewritten as

$$\mathbf{K}(z) = \mathbf{N}(z)\mathbf{V}^{-1}(z) \quad (4.5)$$

where $\mathbf{N}(z)$ is a FIR matrix of dimension $n_u \times n_y$ and $\mathbf{V}(z) = \text{diag}(\underbrace{J(z), \dots, J(z)}_{n_y \text{ times}})$.

Based on the factorization of $\mathbf{K}(z)$ in (4.5) and the one in Lemma 4.1, we can transform (4.3) into a polynomial matrix form. For this purpose, based on the set $\Delta_{\mathbf{W},\mathbf{i}}$ defined in (4.2) ($i = 1, \dots, n_y$), let us define the set $\mathbf{Y}_{\mathbf{i}}$ made up of all polynomial matrices $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))$ obtained by considering the left-factorization of $(\Delta\mathbf{W}_{y,i}(z), -\Delta\mathbf{W}_{u,i}(z))$ given in Lemma 4.1, i.e.

$$\begin{aligned} \mathbf{Y}_{\mathbf{i}} = \{ & (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \mid (\Delta\mathbf{W}_{y,i}(z), -\Delta\mathbf{W}_{u,i}(z)) = Q_i^{-1}(z)(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \\ & \text{with } \Delta\mathbf{W}_{i}(z) = (\Delta\mathbf{W}_{y,i}(z), \Delta\mathbf{W}_{u,i}(z)) \in \Delta_{\mathbf{W},\mathbf{i}} \text{ and } Q_i(z) \text{ as defined} \\ & \text{in Lemma 4.1} \} \end{aligned} \quad (4.6)$$

Theorem 4.3. Consider Definition 2.3, Theorem 4.2 and one of the model structures \mathcal{M} defined in Section 2.5. Consider the right-factorization of $\mathbf{K}(z)$ given in (4.5). For all $i = 1, \dots, n_y$, denote

$$\Xi_i(z) = \begin{pmatrix} \mathbf{V}_{i}(z) \\ \mathbf{N}(z) \end{pmatrix} \quad (4.7)$$

where $\mathbf{V}_{i}(z)$ is the i^{th} row of $\mathbf{V}(z)$.

For each $\Delta\mathbf{W}_{i}(z) = (\Delta\mathbf{W}_{y,i}(z), \Delta\mathbf{W}_{u,i}(z)) \in \Delta_{\mathbf{W},\mathbf{i}}$, consider the left-factorization of $(\Delta\mathbf{W}_{y,i}(z), -\Delta\mathbf{W}_{u,i}(z)) = Q_i^{-1}(z)(\Upsilon_{y,i}(z), \Upsilon_{u,i}(z))$ given in Lemma 4.1 and denote $\Upsilon_{y,ii}$ the i^{th} element of the row polynomial vector $\Upsilon_{y,i}$. Then, the data set $\mathcal{Z}_{cl}^{\infty}$ in (2.10) is informative w.r.t. the model structure \mathcal{M} if and only if, for all $i = 1, \dots, n_y$,

$$\begin{cases} (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))\Xi_i(z) & \equiv \mathbf{0} \\ \bar{E} [||\Upsilon_{u,i}(z)r(t)||^2] & = 0 \end{cases} \implies (\Upsilon_{y,ii}, \Upsilon_{u,i}) \equiv (0, \mathbf{0}_{1 \times n_u}) \quad (4.8)$$

for all $(\Upsilon_{y,ii}, \Upsilon_{u,i}) \in \mathbf{Y}_{\mathbf{i}}$ where $\mathbf{Y}_{\mathbf{i}}$ is defined in (4.6).

Proof. See Appendix C.5 for the proof. \square

²As we did in Chapter 3 for the matrix $\mathbf{M}(z)$ when considering filtered white noise input $\mathbf{u}(t) = \mathbf{M}(z)\mathbf{v}(t)$.

Remark 4.2. *The factorization of $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ and $\mathbf{K}(z)$ are of course not unique. The chosen factorizations are thus only one of the possible choices. However, it is important to note that this particular choice does not introduce any conservatism since the data informativity condition given in Theorem 4.2 and the one in Theorem 4.3 are equivalent.*

4.3.3 Main result for data informativity

In this section, we derive the main result of this chapter. This result will allow us to check data informativity by verifying for each channel $i = 1, \dots, n_y$ whether a given matrix is full row rank. As we will see in the sequel, this matrix will depend on the complexity of the model structure, on the controller coefficients and on the external excitation parametrization (amplitudes, phase shifts, frequencies for multisine excitation and filter coefficients for filtered white noise excitation).

For this purpose, a first step is to give a formal expression of the polynomial $\Upsilon_{y,ii}$ and the row vector of polynomials $\Upsilon_{u,i} = (\Upsilon_{u,i1}, \dots, \Upsilon_{u,in_u})$ in \mathbf{Y}_i where \mathbf{Y}_i is defined in (4.6). Using Lemma 4.1 and the notations introduced in Section 2.5, we can determine the scalars $\mu_{y,i}, \mu_{u,ik}$ ($i = 1, \dots, n_y, k = 1, \dots, n_u$) such that all $\Upsilon_{y,ii}$ and the entries of all $\Upsilon_{u,i}$ in \mathbf{Y}_i can be expressed as follows:

$$\Upsilon_{y,ii}(z) = \tilde{\boldsymbol{\delta}}_{y,i}^T \mathbf{Z}_{y,i}(z) \quad (4.9)$$

$$\Upsilon_{u,ik}(z) = \tilde{\boldsymbol{\delta}}_{u,ik}^T \mathbf{Z}_{u,ik}(z) \quad (4.10)$$

where $\mathbf{Z}_{y,i}(z) = (z^{-1}, \dots, z^{-\mu_{y,i}})^T$, $\mathbf{Z}_{u,ik}(z) = (z^{-\rho_{ik}}, \dots, z^{-\mu_{u,ik}})^T$ and $\tilde{\boldsymbol{\delta}}_{y,i}, \tilde{\boldsymbol{\delta}}_{u,ik}$ are vectors of coefficients. The values of $\mu_{y,i}, \mu_{u,ik}$ ($i = 1, \dots, n_y, k = 1, \dots, n_u$) as a function of the model structure type and the model structure complexity are given in Appendix C.6.

Using (4.9) and (4.10) and defining $\tilde{\boldsymbol{\delta}}_{u,i} = (\tilde{\boldsymbol{\delta}}_{u,i1}^T, \dots, \tilde{\boldsymbol{\delta}}_{u,in_u}^T)^T$, the vectors of polynomials $(\Upsilon_{y,ii}, \Upsilon_{u,i})$ in the set \mathbf{Y}_i (see (4.6)) can be rewritten as follows:

$$(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) = \tilde{\boldsymbol{\delta}}_i^T \text{bdiag}(\mathbf{Z}_{y,i}(z), \mathbf{Z}_{u,i}(z)) \quad (4.11)$$

where $\tilde{\boldsymbol{\delta}}_i = (\tilde{\boldsymbol{\delta}}_{y,i}^T, \tilde{\boldsymbol{\delta}}_{u,i}^T)^T$ and

$$\mathbf{Z}_{u,i}(z) = \text{bdiag}(\mathbf{Z}_{u,i1}(z), \dots, \mathbf{Z}_{u,in_u}(z)) \quad (4.12)$$

We will denote by $\zeta_i, \zeta_{u,i}$ and $\zeta_{y,i}$ the dimension of $\tilde{\boldsymbol{\delta}}_i, \tilde{\boldsymbol{\delta}}_{u,i}$ and $\tilde{\boldsymbol{\delta}}_{y,i}$. These dimensions can be easily deduced based on $\mu_{y,i}, \mu_{u,ik}$ ($i = 1, \dots, n_y, k = 1, \dots, n_u$) given in Appendix C.6. Let us also introduce the set

$$\mathbf{D}_i = \{ \tilde{\boldsymbol{\delta}}_i \mid (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) = \tilde{\boldsymbol{\delta}}_i^T \text{bdiag}(\mathbf{Z}_{y,i}(z), \mathbf{Z}_{u,i}(z)) \in \mathbf{Y}_i \} \quad (4.13)$$

Note that, except in the FIR case where $\mathbf{D}_i = \mathbb{R}^{\zeta_i}$, \mathbf{D}_i is a subset of \mathbb{R}^{ζ_i} . Note that such a set \mathbf{D}_i was also considered in the open-loop case (see Appendix B.4).

In Sections 4.3.4 and 4.4, we will describe a procedure to construct matrices $\mathbf{R}^{(i)}$, $\mathbf{S}^{(i)}$ and $\mathbf{T}^{(i)}$ such that the left-hand side of (4.8) for a given $(\Upsilon_{y,ii}, \Upsilon_{u,i}) \in \mathbf{Y}_i$ (with \mathbf{Y}_i as defined in (4.6)) can be expressed as a function of the parameter vector $\tilde{\delta}_i \in \mathbf{D}_i$ defining $(\Upsilon_{y,ii}, \Upsilon_{u,i})$ (see (4.11)). In particular, with the matrices $\mathbf{R}^{(i)}$ and $\mathbf{S}^{(i)}$ obtained with the procedure presented in Section 4.3.4, the first equation of the left-hand side of (4.8) is equivalent to (4.14):

$$\left\{ \begin{array}{l} \underbrace{\begin{pmatrix} \tilde{\delta}_{y,i}^T & \tilde{\delta}_{u,i}^T \end{pmatrix}}_{\tilde{\delta}_i^T} \begin{pmatrix} \mathbf{R}^{(i)} \\ \mathbf{S}^{(i)} \end{pmatrix} = \mathbf{0} \quad (\text{BJ/ARX/ARMAX case}) \\ \tilde{\delta}_{u,i}^T \mathbf{S}^{(i)} = \mathbf{0} \quad (\text{OE/FIR case}) \end{array} \right. \quad (4.14)$$

The difference in the above equation follows from the fact that $\Upsilon_{y,ii} \equiv 0$ for OE/FIR model structures (see Lemma 4.1). With the matrix $\mathbf{T}^{(i)}$ obtained with the procedure that will be described in Section 4.4, the second equation of the left-hand side of (4.8) is equivalent to (4.15):

$$\tilde{\delta}_{u,i}^T \mathbf{T}^{(i)} = \mathbf{0} \quad (4.15)$$

As will be shown later, the dimension and the elements of the matrices $\mathbf{R}^{(i)}$, $\mathbf{S}^{(i)}$ and $\mathbf{T}^{(i)}$ will be function of the complexity of the model structure and will depend on the controller coefficients and the external excitation parametrization.

Theorem 4.4. *Consider Definition 2.3 and, in particular, the set Z_{cl}^∞ obtained via the application of the external excitation signal \mathbf{r} to the closed-loop system of Figure 2.1. Consider Theorem 4.3 and one of the model structures \mathcal{M} defined in Section 2.5. Consider finally the notations introduced in Section 4.3.3 and, in particular, the equations (4.14) and (4.15) that are respectively equivalent to the first and second equations of the left-hand side of (4.8). Then, the data set Z_{cl}^∞ is informative with respect to \mathcal{M} if, for each $i = 1, \dots, n_y$, at least one of the matrices $\mathbf{P}^{(a,i)}$, $\mathbf{P}^{(b,i)}$ and $\mathbf{P}^{(c,i)}$ defined below are full row rank.*

- (a) The matrix $\mathbf{P}^{(a,i)}$ is equal to $\begin{pmatrix} \mathbf{R}^{(i)} \\ \mathbf{S}^{(i)} \end{pmatrix}$ for the BJ/ARX/ARMAX case and to $\mathbf{S}^{(i)}$ for the OE/FIR case.
- (b) The matrix $\mathbf{P}^{(b,i)}$ is equal to $\mathbf{T}^{(i)}$.
- (c) The matrix $\mathbf{P}^{(c,i)}$ is equal to $\begin{pmatrix} \mathbf{R}^{(i)} & \mathbf{0} \\ \mathbf{S}^{(i)} & \mathbf{T}^{(i)} \end{pmatrix}$ for the BJ/ARX/ARMAX case and to $\begin{pmatrix} \mathbf{S}^{(i)} & \mathbf{T}^{(i)} \end{pmatrix}$ for the OE/FIR case.

Moreover, when $\mathbf{r} = \mathbf{0}_{n_u \times 1}$, the data set Z_{cl}^∞ is informative with respect to \mathcal{M} if, for each $i = 1, \dots, n_y$, the matrix $\mathbf{P}^{(a,i)}$ defined above is full row rank.

Proof. We will prove the theorem in the BJ/ARX/ARMAX case. The proof for the OE/FIR case can be derived using the same argumentation. Let us first note that $\tilde{\delta}_i = (\tilde{\delta}_{y,i}^T, \tilde{\delta}_{u,i}^T)^T$ in (4.14)-(4.15) is constrained to lie in \mathbf{D}_i with \mathbf{D}_i as defined in (4.13). When $\tilde{\delta}_i \in \mathbf{D}_i$, having a full row rank $\mathbf{P}^{(a,i)}$ is a sufficient condition³ for $\tilde{\delta}_i = \mathbf{0}$ to be the unique solution of (4.14). Note also that $\tilde{\delta}_i = \mathbf{0}$ is equivalent to $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \equiv (0, \mathbf{0})$ (see (4.11)). Consequently, using the equivalence recalled in the statement of the theorem, when $\mathbf{P}^{(a,i)}$ is full row rank, we have also that $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))\Xi_i(z) \equiv \mathbf{0} \implies (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \equiv (0, \mathbf{0})$ for all $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \in \mathbf{Y}_i$ with \mathbf{Y}_i as defined in (4.6). It is clear by Theorem 4.3 that the latter implies that \mathcal{Z}_{cl}^∞ is informative (in particular when $r = \mathbf{0}_{n_u \times 1}$).

Note now that (4.14) and (4.15) can be combined into $\tilde{\delta}_i^T \mathbf{P}^{(c,i)} = \mathbf{0}$. If $\mathbf{P}^{(c,i)}$ is full row rank, then $\tilde{\delta}_i^T \mathbf{P}^{(c,i)} = \mathbf{0}$ implies $\tilde{\delta}_i = \mathbf{0}$. Consequently, using the equivalence recalled in the statement of the theorem, when $\mathbf{P}^{(c,i)}$ is full row rank, we have also that (4.8) holds for all $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \in \mathbf{Y}_i$. It is clear by Theorem 4.3 that the latter implies that \mathcal{Z}_{cl}^∞ is informative.

Finally, if $\mathbf{P}^{(b,i)}$ is full row rank, the equation (4.15) implies $\tilde{\delta}_{u,i} = \mathbf{0}$. Consequently, using the equivalence recalled in the statement of the theorem, when $\mathbf{P}^{(b,i)}$ is full row rank, we have also that: $\bar{E} [|\Upsilon_{u,i}(z)r(t)|^2] = 0 \implies \Upsilon_{u,i}(z) \equiv \mathbf{0}$ for all $\Upsilon_{u,i}(z) \in \{\Upsilon_{u,i}(z) \mid (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \in \mathbf{Y}_i\}$. Since $\Upsilon_{u,i}(z) \equiv \mathbf{0}$, the first equation of the left-hand side of (4.8) reduces to $\Upsilon_{y,ii}(z)J(z) \equiv 0$ where $J(z)$ is a given polynomial defined before in (4.5). The latter always implies that $\Upsilon_{y,ii}(z) \equiv 0$. Consequently, by virtue of Theorem 4.3, we have also data informativity when $\mathbf{P}^{(b,i)}$ is full row rank. \square

Remark 4.3. *As already mentioned in the proof Theorem 4.4, we cannot easily conclude that the rank verification condition of the matrices $\mathbf{P}^{(c,i)}$ is necessary for almost all model structures, except for the FIR model structures since we have that $\mathbf{D}_i = \mathbb{R}^{\zeta_i}$. Indeed, from the proof of Theorem 4.4, we will have the data informativity if and only if, for all $i = 1, \dots, n_y$, the following property holds*

$$\delta_i^T \mathbf{P}^{(c,i)} = \mathbf{0} \implies \delta_i = \mathbf{0}$$

Therefore, a necessary and sufficient condition for the data informativity is that, for all $i = 1, \dots, n_y$, the intersection of left null space of $\mathbf{P}^{(c,i)}$ and the set \mathbf{D}_i is equal to $\{\mathbf{0}\}$. Hence, a geometrical study of \mathbf{D}_i is necessary and this study is challenging for all the model structures considered in this dissertation (except FIR ones).

Let us now show how we can rewrite the left-hand side of (4.8) into (4.14)-(4.15).

³If \mathbf{D}_i would be equal to \mathbb{R}^{ζ_i} , this would be a necessary and sufficient condition.

4.3.4 Rewriting of $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))\Xi_i(z) \equiv \mathbf{0}$ into (4.14)

We will consider the BJ/ARX/ARMAX case since it is the more general. Using (4.11) and (4.7), we have that:

$$(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))\Xi_i(z) = (\tilde{\delta}_{y,i}^T, \tilde{\delta}_{u,i}^T)\mathbf{U}^{(i)}(z) \quad (4.16)$$

with $\mathbf{U}^{(i)}(z)$ a polynomial matrix of dimension $\zeta_i \times n_y$:

$$\mathbf{U}^{(i)}(z) = \begin{pmatrix} \mathbf{Z}_{y,i}(z)\mathbf{V}_i(z) \\ \mathbf{Z}_{u,i}(z)\mathbf{N}(z) \end{pmatrix}$$

where $\mathbf{Z}_{y,i}(z)$ and $\mathbf{Z}_{u,i}(z)$ are defined in (4.9) and (4.12) respectively and where ζ_i is the dimension of $\tilde{\delta}_i = (\tilde{\delta}_{y,i}^T, \tilde{\delta}_{u,i}^T)^T$.

Each entry $U_{ml}^{(i)}(z)$ of $\mathbf{U}^{(i)}(z)$ ($m = 1, \dots, \zeta_i$, $l = 1, \dots, n_y$) is a polynomial in z^{-1} and can therefore be rewritten as $\alpha_{i,ml}^T \mathbf{Z}_{U^{(i)}}(z)$ with

$$\mathbf{Z}_{U^{(i)}}(z) = (z^{-\gamma_{i,min}}, z^{-\gamma_{i,min}-1}, \dots, z^{-\gamma_{i,max}})^T$$

where $\alpha_{i,ml}$ is a vector containing the coefficients⁴ of the polynomial $U_{ml}^{(i)}(z)$ and where $z^{-\gamma_{i,min}}$ and $z^{-\gamma_{i,max}}$ are the monomials in z^{-1} with the smallest and largest degrees respectively among all entries $U_{ml}^{(i)}$ of $\mathbf{U}^{(i)}$. This yields to the following expression for $\mathbf{U}^{(i)}(z)$:

$$\mathbf{U}^{(i)}(z) = \begin{pmatrix} \mathbf{Z}_{y,i}(z)\mathbf{V}_i(z) \\ \mathbf{Z}_{u,i}(z)\mathbf{N}(z) \end{pmatrix} = \begin{pmatrix} \mathbf{R}^{(i)} \\ \mathbf{S}^{(i)} \end{pmatrix} (\mathbf{I}_{n_y} \otimes \mathbf{Z}_{U^{(i)}}(z)) \quad (4.17)$$

where \otimes represents the Kronecker product and the matrix $(\mathbf{R}^{(i)T}, \mathbf{S}^{(i)T})^T$ is such that its m^{th} row ($m = 1, \dots, \zeta_i$) is given by $(\alpha_{i,m1}^T, \dots, \alpha_{i,mn_y}^T)$.

Using now (4.16) and (4.17), we see that $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))\Xi_i(z) \equiv \mathbf{0}$ is equivalent to (4.14). The matrix $(\mathbf{R}^{(i)T}, \mathbf{S}^{(i)T})^T$ has ζ_i rows and a number of columns equal to $n_y(\gamma_{i,max} - \gamma_{i,min} + 1)$. This number of columns therefore depends both on the complexity of the controller and on the complexity of the model structure (via $\gamma_{i,min}$, $\gamma_{i,max}$, $\mu_{u,ik}$, $\mu_{y,i}$). The more complex the model structure and the controller are, the larger this number of columns is. While the number of columns of the matrix $(\mathbf{R}^{(i)T}, \mathbf{S}^{(i)T})^T$ depends on both the model structure complexity and the controller complexity, its number of rows ζ_i only depends on the model structure complexity.

Example 4.1. Consider the following MISO ARX model structure \mathcal{M} with $n_u = 2$ inputs and $n_y = 1$ output:

$$\mathbf{G}(z, \boldsymbol{\theta}) = \begin{pmatrix} \frac{\tilde{\theta}_{11,1}z^{-1}}{1 + \eta_{1,1}z^{-1}} & \frac{\tilde{\theta}_{12,1}z^{-1} + \tilde{\theta}_{12,2}z^{-2}}{1 + \eta_{1,1}z^{-1}} \end{pmatrix}$$

$$H(z, \boldsymbol{\theta}) = \frac{1}{1 + \eta_{1,1}z^{-1}}$$

⁴These coefficients correspond to the ones in $\mathbf{N}(z)$ and $J(z)$ of the right-factorization of the controller $\mathbf{K}(z)$ in (4.5).

where $\boldsymbol{\theta} = (\tilde{\theta}_{11,1}, \tilde{\theta}_{12,1}, \tilde{\theta}_{12,2}, \eta_{1,1})^T \in \mathcal{D}_\theta$. In this case, we have the left-factorization of $(\Delta W_{y,1}(z), -\Delta \mathbf{W}_{u,1}(z))$ in Lemma 4.1 of the form $(\Delta W_{y,1}(z), -\Delta \mathbf{W}_{u,1}(z)) = Q_1^{-1}(z)(\Upsilon_{y,1}(z), \Upsilon_{u,1}(z))$ with $Q_1(z) = 1$. From Appendix C.6, the polynomial $\Upsilon_{y,1}(z) = \Upsilon_{y,11}(z)$ has an order $\mu_{y,1}$ equal to 1 and is given by

$$\Upsilon_{y,11}(z) = \tilde{\delta}_{y,1}^T Z_{y,1}(z)$$

with $\tilde{\delta}_{y,1} = \eta_{1,1}'' - \eta_{1,1}'$ and $Z_{y,1}(z) = z^{-1}$.

From Appendix C.6, the polynomial entries $\Upsilon_{u,11}(z)$ and $\Upsilon_{u,12}(z)$ of $\Upsilon_{u,1}(z)$ have a degree of $\mu_{u,11} = 1$ and $\mu_{u,12} = 2$ respectively and are successively given by

$$\begin{aligned} \Upsilon_{u,11}(z) &= \tilde{\delta}_{u,11}^T Z_{u,11}(z) \\ \Upsilon_{u,12}(z) &= \tilde{\boldsymbol{\delta}}_{u,12}^T \mathbf{Z}_{u,12}(z) \end{aligned}$$

where

- $\tilde{\delta}_{u,11} = \tilde{\theta}_{11,1}'' - \tilde{\theta}_{11,1}'$ and $Z_{u,11}(z) = z^{-1}$.
- $\tilde{\boldsymbol{\delta}}_{u,12} = (\tilde{\theta}_{12,1}'' - \tilde{\theta}_{12,1}', \tilde{\theta}_{12,2}'' - \tilde{\theta}_{12,2}')^T$ and $\mathbf{Z}_{u,12}(z) = (z^{-1}, z^{-2})^T$.

Then, we have that

$$(\Upsilon_{y,11}(z), \Upsilon_{u,1}(z)) = \underbrace{(\tilde{\delta}_{y,1}^T, \tilde{\boldsymbol{\delta}}_{u,1}^T)}_{\tilde{\boldsymbol{\delta}}_1^T} \text{bdiag}(Z_{y,1}(z), \mathbf{Z}_{u,1}(z))$$

where $\mathbf{Z}_{u,1}(z) = \text{bdiag}(Z_{u,11}(z), Z_{u,12}(z))$ and $\tilde{\boldsymbol{\delta}}_{u,1} = (\delta_{u,11}, \tilde{\boldsymbol{\delta}}_{u,12}^T)^T$.

The true system \mathcal{S} is put under feedback control with the following stabilizing controller $\mathbf{K}(z)$:

$$\mathbf{K}(z) = \underbrace{\begin{pmatrix} 0.4 + 0.29z^{-1} - 0.07z^{-2} \\ 0.04 + 0.01z^{-1} - 0.3z^{-2} \end{pmatrix}}_{\mathbf{N}(z)} \underbrace{\frac{1}{1 - 0.62z^{-1} + 0.07z^{-2}}}_{\mathbf{V}^{-1}(z)}$$

In this case, the least common multiple of the denominators of $\mathbf{K}(z)$ is directly $J(z) = 1 - 0.62z^{-1} + 0.07z^{-2}$. Consequently, the matrix $\boldsymbol{\Xi}_1(z)$ in (4.7) (see Theorem 4.3) is equal to $\boldsymbol{\Xi}_1(z) = (J(z), \mathbf{N}^T(z))^T$. Therefore, $(\Upsilon_{y,11}(z), \Upsilon_{u,1}(z))\boldsymbol{\Xi}_1(z)$ can be rewritten as

$$(\tilde{\delta}_{y,1}^T, \tilde{\boldsymbol{\delta}}_{u,1}^T) \underbrace{\begin{pmatrix} Z_{y,1}(z)J(z) \\ \mathbf{Z}_{u,1}(z)\mathbf{N}(z) \end{pmatrix}}_{\mathbf{U}^{(1)}(z)} = (\tilde{\delta}_{y,1}^T, \tilde{\boldsymbol{\delta}}_{u,1}^T) \underbrace{\begin{pmatrix} z^{-1} - 0.62z^{-2} + 0.07z^{-3} \\ 0.4z^{-1} + 0.29z^{-2} - 0.07z^{-3} \\ 0.04z^{-1} + 0.01z^{-2} - 0.3z^{-3} \\ 0.04z^{-2} + 0.01z^{-3} - 0.3z^{-4} \end{pmatrix}}_{\mathbf{U}^{(1)}(z)}$$

The smallest and largest value of the monomials $z^{-\gamma}$ among the entries of $\mathbf{U}^{(1)}(z)$ are $\gamma_{1,min} = 1$ and $\gamma_{1,max} = 4$ respectively. Therefore, the latter can be rewritten as follows

$$(\tilde{\delta}_{y,1}^T, \tilde{\delta}_{u,1}^T) \begin{pmatrix} \mathbf{R}^{(1)} \\ \mathbf{S}^{(1)} \end{pmatrix} \begin{pmatrix} z^{-\gamma_{1,min}} \\ \vdots \\ z^{-\gamma_{1,max}} \end{pmatrix} = (\tilde{\delta}_{y,1}^T, \tilde{\delta}_{u,1}^T) \begin{pmatrix} 1 & -0.62 & 0.07 & 0 \\ 0.4 & 0.29 & -0.07 & 0 \\ 0.04 & 0.01 & -0.3 & 0 \\ 0 & 0.04 & 0.01 & -0.3 \end{pmatrix} \begin{pmatrix} z^{-1} \\ z^{-2} \\ z^{-3} \\ z^{-4} \end{pmatrix}$$

The entries in the matrices $\mathbf{R}^{(1)}$ and $\mathbf{S}^{(1)}$ are the polynomial coefficients of the numerators in $\mathbf{N}(z)$ and the least common multiple denominator $J(z)$ of the controller $\mathbf{K}(z)$. The number of rows of $\mathbf{R}^{(1)}$ and $\mathbf{S}^{(1)}$ are respectively equal to the dimension of $\tilde{\delta}_{y,1}$ and $\tilde{\delta}_{u,1}$, which are 1 and 3 respectively.

The matrix $\mathbf{P}^{(a,1)} = (\mathbf{R}^{(1)T}, \mathbf{S}^{(1)T})^T$ has a rank of 4: it is full row rank. Therefore, from Theorem 4.4, the data are informative with respect to \mathcal{M} in the costless framework ($\mathbf{r} = \mathbf{0}_{2 \times 1}$). The paper [BGM10] proposes another (sufficient) condition to verify whether $r = \mathbf{0}_{n_u \times 1}$ can yield informative data and given in Theorem 2.2 (Section 2.6). In this example, the condition in Theorem 2.2 is not satisfied since $c_{min} = 2$ and $k_{max} = 2$. Consequently, the sufficient condition of Theorem 4.4 is in this example less restrictive than the one in [BGM10].

Remark 4.4. For MISO ARMAX structures and a SIMO controller, [YZ18] gives a necessary and sufficient condition for data informativity when $r = \mathbf{0}_{n_u \times 1}$. Note however that some restrictions are imposed on the considered controller $\mathbf{K}(z)$ and the ARMAX model structures in [YZ18] to be able to derive this necessary and sufficient condition. In order to compare these results with ours, we have derived the matrix $\mathbf{P}^{(a,1)}$ for the MISO ARMAX examples in [YZ18]. In the three cases where the necessary and sufficient condition for data informativity is not respected, we logically also observe that $\mathbf{P}^{(a,1)}$ is rank deficient. In the last case (case 2 of example 2), the matrix $\mathbf{P}^{(a,1)}$ is also rank deficient while the necessary and sufficient condition of [YZ18] shows that the data are informative in this case. This illustrates that Theorem 4.4 comes with a dose of conservatism.

4.4 Rewriting of $\bar{E} [||\Upsilon_{u,i}(z)\mathbf{r}(t)||^2] = 0$ into (4.15)

This equation has already been studied for the open-loop case⁵ in Chapter 3. The idea for the rewriting of $\bar{E} [||\Upsilon_{u,i}(z)\mathbf{r}(t)||^2] = 0$ is based on the introduction of the regressor concept. For this purpose, let us observe that the n_u entries $\Upsilon_{u,ik}$ of the polynomial vector $\Upsilon_{u,i}$ are the FIR filters given in (4.10) with the orders $\mu_{u,ik}$ given in Appendix C.6. Hence, we have that

$$\begin{aligned}\Upsilon_{u,i}(z)\mathbf{r}(t) &= \sum_{k=1}^{n_u} \Upsilon_{u,ik}(z)r_k(t) \\ &= \sum_{k=1}^{n_u} \tilde{\delta}_{u,ik}^T \phi_{r_k,i}(t) \\ &= \underbrace{\left(\tilde{\delta}_{u,i1}^T \quad \tilde{\delta}_{u,i2}^T \quad \cdots \quad \tilde{\delta}_{u,in_u}^T \right)}_{\tilde{\delta}_{u,i}^T} \phi_{\mathbf{r},i}(t)\end{aligned}$$

where

$$\phi_{\mathbf{r},i}(t) = \begin{pmatrix} \phi_{r_1,i}(t) \\ \phi_{r_2,i}(t) \\ \vdots \\ \phi_{r_{n_u},i}(t) \end{pmatrix} \quad \text{with} \quad \phi_{r_k,i}(t) = \begin{pmatrix} r_k(t - \rho_{ik}) \\ r_k(t - \rho_{ik} - 1) \\ \vdots \\ r_k(t - \mu_{u,ik}) \end{pmatrix}$$

We have that $\bar{E} [||\Upsilon_{u,i}(z)\mathbf{r}(t)||^2] = 0 \Leftrightarrow \tilde{\delta}_{u,i}^T \bar{E} [\phi_{\mathbf{r},i}(t)\phi_{\mathbf{r},i}^*(t)] \tilde{\delta}_{u,i} = 0$. By following the same approach as in Chapter 3, we can show that, when⁶ \mathbf{r} is given either by (2.14) or (2.15), the regressor vector $\phi_{\mathbf{r},i}$ can be rewritten as $\phi_{\mathbf{r},i}(t) = \mathbf{T}^{(i)}\varphi_i(t)$ where φ is a complex- or real-valued signal vector of dimension d such that $\bar{E} [\varphi_i(t)\varphi_i^*(t)] \succ 0$ and $\mathbf{T}^{(i)}$ a complex or real-valued time independent deterministic matrix of dimension $\zeta_{u,i} \times d$ where $\zeta_{u,i}$ is the dimension of $\tilde{\delta}_{u,i}$. Consequently, we have that $\bar{E} [||\Upsilon_{u,i}(z)\mathbf{r}(t)||^2] = 0 \Leftrightarrow \tilde{\delta}_{u,i}^T \mathbf{T}^{(i)} \bar{E} [\varphi_i(t)\varphi_i^*(t)] \mathbf{T}^{(i)*} \tilde{\delta}_{u,i} = 0 \Leftrightarrow \tilde{\delta}_{u,i}^T \mathbf{T}^{(i)} = \mathbf{0}$. Therefore, $\bar{E} [||\Upsilon_{u,i}(z)\mathbf{r}(t)||^2] = 0$ is rewritten as

$$\tilde{\delta}_{u,i}^T \mathbf{T}^{(i)} = \mathbf{0}$$

As shown in Chapter 3, when the external signal \mathbf{r} is a multisine given by (2.14), we have that $\mathbf{T}^{(i)} = \mathbf{T}^{(sin,i)}$ where $\mathbf{T}^{(sin,i)}$ is given by (3.8) if we replace μ_k ($k = 1, \dots, n_u$) in this expression by $\mu_{u,ik}$ ($k = 1, \dots, n_u$). When \mathbf{r} is a filtered white noise given by (2.15), we have that $\mathbf{T}^{(i)} = \mathbf{T}^{(fwn,i)}$ where $\mathbf{T}^{(fwn,i)}$ is given by (3.13) if we replace μ_k ($k = 1, \dots, n_u$) in this expression by $\mu_{u,ik}$ ($k = 1, \dots, n_u$).

⁵In Chapter 3, it was $\bar{E} [||\Delta\mathbf{X}(z)\mathbf{u}(t)||^2] = 0$.

⁶In Chapter 3, it is not \mathbf{r} but \mathbf{u} .

4.5 Interpretation of Theorem 4.4

Theorem 4.4 shows that we can verify whether a given external excitation \mathbf{r} yields informative data by checking, for each channel, if one of the three matrices $\mathbf{P}^{(a,i)}$, $\mathbf{P}^{(b,i)}$ or $\mathbf{P}^{(c,i)}$ is full row rank.

In particular, if $\mathbf{P}^{(a,i)}$ is full row rank for each channel i , then the data set \mathcal{Z}_{cl}^∞ will yield informative data even if the external excitation \mathbf{r} is equal to zero.

To have that $\mathbf{P}^{(a,i)}$ is full row rank, the number of its columns should be larger than its number of rows ζ_i . For this necessary condition to hold, the controller $\mathbf{K}(z)$ must have sufficient complexity (see the discussion above Example 4.1).

Even if $\mathbf{P}^{(a,i)}$ is not full row rank, we can of course obtain informative data by adding a nonzero excitation \mathbf{r} of the type (2.14) or (2.15). In this case, the data informativity can be checked by verifying whether $\mathbf{P}^{(c,i)}$ is full row rank. We observe that $\mathbf{P}^{(a,i)}$ and $\mathbf{P}^{(c,i)}$ have the same number ζ_i of rows, but the number of columns of $\mathbf{P}^{(c,i)}$ is larger than the one of $\mathbf{P}^{(a,i)}$ (due to the matrix $\mathbf{T}^{(i)}$ linked to the external excitation). Consequently, even if $\mathbf{P}^{(a,i)}$ is not full row rank, the addition of extra columns can compensate the rank deficiency of $\mathbf{P}^{(a,i)}$ and make $\mathbf{P}^{(c,i)}$ full rank. In particular, a multisine excitation of the type (2.14) with s different sinusoids adds $2s$ columns to the ones already present in $\mathbf{P}^{(a,i)}$. For a filtered white noise excitation of the type (2.15), we have, as shown in Section 3.5, the following relation: the larger the number f of white noises in \mathbf{v} and the larger the complexity of the filters in $\mathbf{M}(z)$, the larger the number of added columns in $\mathbf{P}^{(c,i)}$ with respect to $\mathbf{P}^{(a,i)}$.

The above consideration allows to derive a condition for the data informativity when $\mathbf{P}^{(a,i)}$ is not full row rank for each channel i . The excitation richness (e.g., the number of sinusoids in the multisine case) must be sufficient to guarantee that, for each channel i , the number of columns in $\mathbf{T}^{(i)}$ is larger than the difference between the number ζ_i of rows of $\mathbf{P}^{(a,i)}$ and the rank of $\mathbf{P}^{(a,i)}$. If that is the case, it should be still verified that $\mathbf{P}^{(c,i)}$ is full row rank for each i since rank deficiency can occur due to some unlucky combinations of controller coefficients and excitation parametrizations (amplitudes and phase shifts for the multisine case and coefficients of the filters in $\mathbf{M}(z)$ for the filtered white noise case). A similar phenomenon was also observed in the open-loop case in Chapter 3.

In the open-loop case, we have shown that our MIMO data informativity condition (Theorem 3.3) exactly boils down to the SISO data informativity condition of [GBM08] when we use $n_u = n_y = 1$ in Theorem 3.3. This was possible due to the fact that input regressor persistency was at the roots of both the SISO and MIMO case results. In the closed-loop case, this comparison is much more complicated since the matricial rank verification of Theorem 4.4 is difficult to relate to the conditions on the degrees of the controller and the model structure polynomial in the SISO case. However, we

can still do some comparisons between the MIMO and SISO case. To illustrate this, we are going to apply the methodology of this chapter to verify the data informativity for the identification of a SISO system in closed-loop with two SISO controllers. We will see that the condition in Theorem 4.4 can be conservative for SISO ARMAX model structures.

Example 4.2. Consider a SISO true system \mathcal{S} given by

$$\mathcal{S}: \quad y(t) = \underbrace{\frac{-0.23z^{-1}}{1+0.87z^{-1}}}_{G_0(z)} u(t) + \underbrace{\frac{1-0.3z^{-1}}{1+0.87z^{-1}}}_{H_0(z)} e(t)$$

To identify \mathcal{S} , we consider the following full-order ARMAX model structure

$$\mathcal{M} = \{(G(z, \boldsymbol{\theta}), H(z, \boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \mathcal{D}_\theta\}$$

with

$$\begin{aligned} G(z, \boldsymbol{\theta}) &= \frac{B(z, \tilde{\boldsymbol{\theta}})}{A(z, \boldsymbol{\eta})} = \frac{\tilde{\theta}_1 z^{-1}}{1 + \eta_2 z^{-1}} \\ H(z, \boldsymbol{\theta}) &= \frac{C(z, \boldsymbol{\eta})}{A(z, \boldsymbol{\eta})} = \frac{1 + \eta_1 z^{-1}}{1 + \eta_2 z^{-1}} \end{aligned}$$

where $\boldsymbol{\theta} = (\tilde{\boldsymbol{\theta}}^T, \boldsymbol{\eta}^T)^T$ with $\tilde{\boldsymbol{\theta}} = \tilde{\theta}_1$ and $\boldsymbol{\eta} = (\eta_1, \eta_2)^T$.

The system \mathcal{S} is put under feedback control with the following stabilizing controller $K(z)$

$$K(z) = \frac{0.4147 + 0.5038z^{-1} + 0.1507z^{-2}}{1 + 1.568z^{-1} + 0.6154z^{-2}}$$

First, by virtue of the necessary and sufficient condition in [GBM08], we prove that the costless experiment ($r = 0$) yields informative data w.r.t. \mathcal{M} (i.e., the controller is enough complex w.r.t. \mathcal{M}). Let us verify this result with Theorem 4.4 by constructing the matrix $\mathbf{P}^{(a,1)}$ and verify that it is full row rank.

The polynomial $\Upsilon_{y,11}(z)$ has an order of $\mu_{y,1} = 2$ and the polynomial $\Upsilon_{u,11}(z)$ has an order of $\mu_{u,11} = 2$ (see Appendix C.6). Then, by following the procedure in Section 4.3.4, the matrix $\mathbf{P}^{(a,1)}$ is given by

$$\mathbf{P}^{(a,1)} = \begin{pmatrix} \mathbf{R}^{(1)} \\ \mathbf{S}^{(1)} \end{pmatrix} = \begin{pmatrix} 0.4147 & 0.5038 & 0.1507 & 0 \\ 0 & 0.4147 & 0.5038 & 0.1507 \\ 1 & 1.568 & 0.6154 & 0 \\ 0 & 1 & 1.568 & 0.6154 \end{pmatrix}$$

It has a rank of 4: it is therefore full row rank. Consequently, from Theorem 4.4, we indeed verify that the data are informative.

4.6. A synthetic approach for the verification of the data informativity for the direct closed-loop identification of MIMO systems

Let us now consider another SISO controller for this system:

$$K(z) = \frac{1.384 + 0.5329z^{-1}}{1 - 0.7536z^{-1}}$$

This controller still stabilizes the system \mathcal{S} . From the necessary and sufficient condition in [GBM08], we prove that the costless experiment with this new controller yields informative data w.r.t. \mathcal{M} (i.e., the controller is enough complex w.r.t. \mathcal{M}). In this case, the matrix $\mathbf{P}^{(a,1)}$ is given by

$$\mathbf{P}^{(a,1)} = \begin{pmatrix} \mathbf{R}^{(1)} \\ \mathbf{S}^{(1)} \end{pmatrix} = \begin{pmatrix} 1.384 & 0.5329 & 0 \\ 0 & 1.384 & 0.5329 \\ 1 & -0.7536 & 0 \\ 0 & 1 & -0.7536 \end{pmatrix}$$

It has only 3 columns for 4 rows: it cannot be full row rank. Consequently, we cannot conclude on the informativity in this case, illustrating the conservatism of this condition in the SISO ARMAX case.

4.6 A synthetic approach for the verification of the data informativity for the direct closed-loop identification of MIMO systems

In this chapter, we have extensively analyzed the data informativity for the direct closed-loop identification of MIMO systems. In this section, we give a synthesis of the results by giving a procedure that the user can follow to verify this data informativity given a controller $\mathbf{K}(z)$, a nonzero external excitation \mathbf{r} and a MIMO model structure \mathcal{M} .

First, if the external excitation vector \mathbf{r} is a filtered white noise (see (2.15)) with $f \geq n_u$ and with $\text{rank}(\mathbf{M}(e^{j\omega})) = n_u$ for almost all ω in $] -\pi, \pi[$, then we have data informativity from Lemma 2.2 in Section 2.6.

If it is not the case, we have to use the results developed in this chapter. First, for all channels $i = 1, \dots, n_y$, we construct and compute the rank of the matrices $\mathbf{P}^{(a,i)}$. If they are all full row rank, we conclude that the controller $\mathbf{K}(z)$ alone guarantees the data informativity and that the data are informative also in the presence of an external excitation. However, if the matrix $\mathbf{P}^{(a,i)}$ is not full row rank for some channels, we construct the corresponding matrices $\mathbf{P}^{(c,i)}$ and we verify that they are all full row rank. If it is the case, then the data set \mathcal{Z}_{cl}^∞ is informative with respect to \mathcal{M} . If not, we cannot conclude for the informativity except for FIR model structures for which the rank condition is a necessary and sufficient condition for data informativity.

Remark 4.5. As already explained in this chapter, there are two conditions in the literature for the data informativity w.r.t. MIMO model structures in the costless framework [BGM10, YZ18]. These results can also be used to verify the data informativity

even if $\mathbf{r} \neq \mathbf{0}_{n_u \times 1}$. Indeed, if the data are informative with $\mathbf{r} = \mathbf{0}_{n_u \times 1}$, then they will be informative when $\mathbf{r} \neq \mathbf{0}_{n_u \times 1}$. Therefore, in the case where there is at least one channel for which both matrices $\mathbf{P}^{(a,i)}$ and $\mathbf{P}^{(c,i)}$ are not full row rank, we can still try to verify the data informativity with Theorem 2.2. For MISO ARMAX model structures, there is a necessary and sufficient condition for the data informativity in [YZ18], provided that the controller $\mathbf{K}(z)$ and the model structure satisfy the assumptions considered in [YZ18].

4.7 Numerical example

In this section, we will illustrate the result developed on data informativity for the direct closed-loop identification. Another example with a MISO BJ model structure can be found in [CBBM20a].

4.7.1 True system to be identified and model structure

Consider the following MISO OE system \mathcal{S} with $n_u = 3$ inputs and $n_y = 2$ outputs given by

$$\mathbf{y}(t) = \underbrace{\begin{pmatrix} \frac{-0.1z^{-1}}{1+0.1z^{-1}} & -0.8z^{-1} - 0.7z^{-2} & 0.5z^{-1} \\ z^{-1} - 0.5z^{-2} & -1.5z^{-1} + z^{-2} & 4.2z^{-1} - 3z^{-2} + 2z^{-3} \end{pmatrix}}_{=\mathbf{G}_0(z)} \mathbf{u}(t) + \mathbf{e}(t)$$

where $\mathbf{u}(t) = (u_1(t), u_2(t), u_3(t))^T$, $\mathbf{y}(t) = (y_1(t), y_2(t))^T$ and $\mathbf{e}(t) = (e_1(t), e_2(t))^T$. We will consider that the covariance matrix $\bar{E}[\mathbf{e}(t)\mathbf{e}^T(t)]$ of \mathbf{e} is equal to $0.05\mathbf{I}_2$.

We will identify \mathcal{S} within a full-order model structure \mathcal{M} (see Section 2.5) with all delays ρ_{ik} equal to 1 and

$$\begin{aligned} B_{11}(z, \tilde{\boldsymbol{\theta}}_1) &= \tilde{\theta}_{11,1}z^{-1} & F_{11}(z, \tilde{\boldsymbol{\theta}}_1) &= 1 + \tilde{\theta}_{11,2}z^{-1} \\ B_{12}(z, \tilde{\boldsymbol{\theta}}_1) &= \tilde{\theta}_{12,1}z^{-1} + \tilde{\theta}_{12,2}z^{-2} & F_{12}(z, \tilde{\boldsymbol{\theta}}_1) &= 1 \\ B_{13}(z, \tilde{\boldsymbol{\theta}}_1) &= \tilde{\theta}_{13,1}z^{-1} & F_{13}(z, \tilde{\boldsymbol{\theta}}_1) &= 1 \\ B_{21}(z, \tilde{\boldsymbol{\theta}}_2) &= \tilde{\theta}_{21,1}z^{-1} + \tilde{\theta}_{21,2}z^{-2} & F_{21}(z, \tilde{\boldsymbol{\theta}}_2) &= 1 \\ B_{22}(z, \tilde{\boldsymbol{\theta}}_2) &= \tilde{\theta}_{22,1}z^{-1} + \tilde{\theta}_{22,2}z^{-2} & F_{22}(z, \tilde{\boldsymbol{\theta}}_2) &= 1 \\ B_{23}(z, \tilde{\boldsymbol{\theta}}_2) &= \tilde{\theta}_{23,1}z^{-1} + \tilde{\theta}_{23,2}z^{-2} + \tilde{\theta}_{23,3}z^{-3} & F_{23}(z, \tilde{\boldsymbol{\theta}}_2) &= 1 \end{aligned}$$

with

$$\boldsymbol{\theta} = (\tilde{\boldsymbol{\theta}}_1^T, \tilde{\boldsymbol{\theta}}_2^T)^T = (\tilde{\theta}_{11,1}, \tilde{\theta}_{11,2}, \tilde{\theta}_{12,1}, \tilde{\theta}_{12,2}, \tilde{\theta}_{13,1}, \tilde{\theta}_{21,1}, \tilde{\theta}_{21,2}, \tilde{\theta}_{22,1}, \tilde{\theta}_{22,2}, \tilde{\theta}_{23,1}, \tilde{\theta}_{23,2}, \tilde{\theta}_{23,3})^T$$

This MIMO model structure can be seen as two independent MISO model structures where the first channel is an OE model structure and the second channel is a FIR one.

The system \mathcal{S} is put under feedback control with a stabilizing controller $\mathbf{K}(z)$ given by

$$\mathbf{K}(z) = \begin{pmatrix} K_{11}(z) & K_{12}(z) \\ K_{21}(z) & K_{22}(z) \\ K_{31}(z) & K_{32}(z) \end{pmatrix}$$

where

$$\begin{aligned} K_{11}(z) &= 0.4 \\ K_{12}(z) &= 0.6 \\ K_{21}(z) &= \frac{-0.2}{1 - 0.9z^{-1}} \\ K_{22}(z) &= 0.3 \\ K_{31}(z) &= \frac{0.32 - 0.888z^{-1} - 0.04z^{-2} + 0.0036z^{-3}}{1 - 0.9z^{-1} - 0.01z^{-2} + 0.009z^{-3}} \\ &= \frac{0.32 - 0.888z^{-1} - 0.04z^{-2} + 0.0036z^{-3}}{(1 - 0.9z^{-1})(1 - 0.1z^{-1})(1 + 0.1z^{-1})} \\ K_{32}(z) &= \frac{1.44 + 0.612z^{-1} - 0.0048z^{-2} - 0.0054z^{-3}}{1 - 0.01z^{-2}} \\ &= \frac{1.44 + 0.612z^{-1} - 0.0048z^{-2} - 0.0054z^{-3}}{(1 - 0.1z^{-1})(1 + 0.1z^{-1})} \end{aligned}$$

4.7.2 Factorization of the model structure with Lemma 4.1

For data informativity, we have two channels to analyze in our example. First, we have to do the left-factorization of $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ given in Lemma 4.1 of the form

$$(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z)) = Q_i^{-1}(z)(\Upsilon_{y,i}(z), \Upsilon_{u,i}(z))$$

for each channel ($i = 1, 2$).

- **Channel 1:** Since the first channel of \mathcal{M} is a OE, from Lemma 4.1, $\Delta \mathbf{W}_{y,1}(z) = \Upsilon_{y,1}(z) = \mathbf{0}_{1 \times 2}$ and $\Upsilon_{u,1}(z) = (\Upsilon_{u,11}(z), \Upsilon_{u,12}(z), \Upsilon_{u,13}(z))$ is given by

$$\begin{aligned} \Upsilon_{u,11}(z) &= F'_{12}F''_{12}F'_{13}F''_{13}(B''_{11}F'_{11} - B'_{11}F''_{11}) = \tilde{\delta}_{u,11}^T \mathbf{Z}_{u,11}(z) \\ \Upsilon_{u,12}(z) &= F'_{11}F''_{11}F'_{13}F''_{13}(B''_{12}F'_{12} - B'_{12}F''_{12}) = \tilde{\delta}_{u,12}^T \mathbf{Z}_{u,12}(z) \\ \Upsilon_{u,13}(z) &= F'_{11}F''_{11}F'_{12}F''_{12}(B''_{13}F'_{13} - B'_{13}F''_{13}) = \tilde{\delta}_{u,13}^T \mathbf{Z}_{u,13}(z) \end{aligned}$$

which have degrees of $\mu_{u,11} = 2$, $\mu_{u,12} = 4$ and $\mu_{u,13} = 3$ respectively (see Appendix C.6).

- **Channel 2:** Since the first channel of \mathcal{M} is a FIR, from Lemma 4.1, $\Delta \mathbf{W}_{y,2}(z) = \Upsilon_{y,2}(z) = \mathbf{0}_{1 \times 2}$ and $\Upsilon_{u,2}(z) = (\Upsilon_{u,21}(z), \Upsilon_{u,22}(z), \Upsilon_{u,23}(z))$ is given by

$$\begin{aligned}\Upsilon_{u,21}(z) &= B_{21}'' - B_{21}' = \tilde{\delta}_{u,21}^T \mathbf{Z}_{u,21}(z) \\ \Upsilon_{u,22}(z) &= B_{22}'' - B_{22}' = \tilde{\delta}_{u,22}^T \mathbf{Z}_{u,22}(z) \\ \Upsilon_{u,23}(z) &= B_{23}'' - B_{23}' = \tilde{\delta}_{u,23}^T \mathbf{Z}_{u,23}(z)\end{aligned}$$

which have degrees of $\mu_{u,21} = 2$, $\mu_{u,22} = 2$ and $\mu_{u,23} = 3$ respectively (see Appendix C.6).

4.7.3 Costless identification

We want first to verify if the controller $\mathbf{K}(z)$ is enough complex to guarantee the informativity with respect to \mathcal{M} for a costless identification ($\mathbf{r} = \mathbf{0}_{3 \times 1}$). For this purpose, we should construct and verify the rank of the matrices $\mathbf{P}^{(a,1)}$ and $\mathbf{P}^{(a,2)}$ of Theorem 4.4. However, the controller has been chosen such that, for the first channel, we have that

$$\underbrace{(\mathbf{G}_{0,1}(z) - \mathbf{G}_{1:(z, \theta_1'')})}_{\Delta \mathbf{W}_u(z)} \mathbf{K}(z) \equiv \mathbf{0}_{1 \times 2}$$

with $\tilde{\theta}_1'' = (-0.3, -0.5, -0.8, -0.1, 0)^T$. From Theorem 4.2, a zero excitation (i.e, $\mathbf{r} = \mathbf{0}_{3 \times 1}$) will therefore not yield informative data with respect to \mathcal{M} . Let us verify that the condition in Theorem 4.4 is not in contradiction with this conclusion. First, we have to factorize the controller $\mathbf{K}(z) = \mathbf{N}(z)\mathbf{V}^{-1}(z)$ as in (4.5) where $\mathbf{V}(z) = J(z)\mathbf{I}_2$ with $J(z)$ the least common factor of the denominators of $\mathbf{K}(z)$. In this case, we observe that the least common factor is $J(z) = (1 - 0.9z^{-1})(1 - 0.1z^{-1})(1 + 0.1z^{-1}) = 1 - 0.9z^{-1} - 0.01z^{-2} + 0.009z^{-3}$ and so $\mathbf{N}(z)$ is given by

$$\mathbf{N}(z) = \begin{pmatrix} N_{11}(z) & N_{12}(z) \\ N_{21}(z) & N_{22}(z) \\ N_{31}(z) & N_{32}(z) \end{pmatrix}$$

with

$$\begin{aligned}N_{11}(z) &= 0.4 - 0.36z^{-1} - 0.004z^{-2} + 0.0036z^{-3} \\ N_{12}(z) &= 0.6 - 0.54z^{-1} - 0.006z^{-2} + 0.0054z^{-3} \\ N_{21}(z) &= -0.2 + 0.002z^{-2} \\ N_{22}(z) &= 0.3 - 0.27z^{-1} - 0.003z^{-2} + 0.0027z^{-3} \\ N_{31}(z) &= 0.32 - 0.888z^{-1} - 0.04z^{-2} + 0.0036z^{-3} \\ N_{32}(z) &= 1.44 - 0.684z^{-1} - 0.5556z^{-2} - 0.00108z^{-3} + 0.00486z^{-4}\end{aligned}$$

Channel 1: By following the steps given in Section 4.3.4 on the first channel, we can deduce the matrix $\mathbf{P}^{(a,1)} = \mathbf{S}^{(1)}$ where

$$\mathbf{S}^{(1)} = \begin{pmatrix} \mathbf{S}_{11}^{(1)} & \mathbf{S}_{12}^{(1)} \\ \mathbf{S}_{21}^{(1)} & \mathbf{S}_{22}^{(1)} \\ \mathbf{S}_{31}^{(1)} & \mathbf{S}_{32}^{(1)} \end{pmatrix}$$

with

$$\begin{aligned}
\mathbf{S}_{11}^{(1)} &= \begin{pmatrix} 0.4 & -0.36 & -0.004 & 0.0036 & 0 & 0 & 0 \\ 0 & 0.4 & -0.36 & -0.004 & 0.0036 & 0 & 0 \end{pmatrix} \\
\mathbf{S}_{12}^{(1)} &= \begin{pmatrix} 0.6 & -0.54 & -0.006 & 0.0054 & 0 & 0 & 0 \\ 0 & 0.6 & -0.54 & -0.006 & 0.0054 & 0 & 0 \end{pmatrix} \\
\mathbf{S}_{21}^{(1)} &= \begin{pmatrix} -0.2 & 0 & 0.002 & 0 & 0 & 0 & 0 \\ 0 & -0.2 & 0 & 0.002 & 0 & 0 & 0 \\ 0 & 0 & -0.2 & 0 & 0.002 & 0 & 0 \\ 0 & 0 & 0 & -0.2 & 0 & 0.002 & 0 \end{pmatrix} \\
\mathbf{S}_{22}^{(1)} &= \begin{pmatrix} 0.3 & -0.27 & -0.003 & 0.0027 & 0 & 0 & 0 \\ 0 & 0.3 & -0.27 & -0.003 & 0.0027 & 0 & 0 \\ 0 & 0 & 0.3 & -0.27 & -0.003 & 0.0027 & 0 \\ 0 & 0 & 0 & 0.3 & -0.27 & -0.003 & 0.0027 \end{pmatrix} \\
\mathbf{S}_{31}^{(1)} &= \begin{pmatrix} 0.32 & -0.888 & -0.04 & 0.0036 & 0 & 0 & 0 \\ 0 & 0.32 & -0.888 & -0.04 & 0.0036 & 0 & 0 \\ 0 & 0 & 0.32 & -0.888 & -0.04 & 0.0036 & 0 \end{pmatrix} \\
\mathbf{S}_{32}^{(1)} &= \begin{pmatrix} 1.44 & -0.684 & -0.5556 & -0.00108 & 0.00486 & 0 & 0 \\ 0 & 1.44 & -0.684 & -0.5556 & -0.00108 & 0.00486 & 0 \\ 0 & 0 & 1.44 & -0.684 & -0.5556 & -0.00108 & 0.00486 \end{pmatrix}
\end{aligned}$$

First, we can see that there are more columns (14 columns) than rows (9 rows) for the matrix $\mathbf{P}^{(a,1)}$. However, by computing the rank of the matrix $\mathbf{P}^{(a,1)} = \mathbf{S}^{(1)}$, we obtain 8 which is less than its number of rows: $\mathbf{P}^{(a,1)}$ is not full row rank. This is normal since we have constructed the controller such that the data set \mathcal{Z}_d^∞ is not informative with respect to \mathcal{M} in the costless identification. This particular choice of the controller $\mathbf{K}(z)$ shows that having a matrix $\mathbf{P}^{(a,1)}$ with more columns than rows does not necessarily mean that this matrix is full row rank. A similar phenomenon was observed for the open-loop case in Chapter 3.

Channel 2: By following the steps given in Section 4.3.4 on the second channel, we can deduce the matrix $\mathbf{P}^{(a,2)} = \mathbf{S}^{(2)}$ and we observe that $\mathbf{P}^{(a,2)}$ has 13 columns for 7 rows. Moreover, by computing the rank of $\mathbf{P}^{(a,2)}$, we observe that it is full row rank (rank equal to 7).

As the costless identification cannot yield informative data, we need to add an external excitation \mathbf{r} to guarantee the informativity. Several choices of external excitations are presented in the next paragraph.

4.7.4 Identification with external excitation ($\mathbf{r} \neq \mathbf{0}_{n_u \times 1}$)

For each external excitation \mathbf{r} that is considered in this paragraph, we will have to check if either $\mathbf{P}^{(b,1)}$ or $\mathbf{P}^{(c,1)}$ is full row rank since $\mathbf{P}^{(a,1)}$ is not. For the second channel, recall that we have shown in the last paragraph that $\mathbf{P}^{(a,2)}$ is full row rank

and so the verification of the rank of $\mathbf{P}^{(b,2)}$ or $\mathbf{P}^{(c,2)}$ will not be necessary. Therefore, we will focus on the first channel from now onwards, i.e., focus on studying the rank of $\mathbf{P}^{(b,1)}$ or $\mathbf{P}^{(c,1)}$ for each external excitation considered.

Case 1: Let us consider a multisine excitation. The first way to obtain the informativity is to make $\mathbf{P}^{(b,1)}$ full row rank. For this purpose, we need at least five sinusoids in \mathbf{r} since $\zeta_{u,1} = 9$. Let us therefore choose a signal $\mathbf{r} = (r_1, r_2, r_3)^T$ with five sinusoids at $\omega_1 = 0.0001\text{rad/s}$, $\omega_2 = 0.001\text{rad/s}$, $\omega_3 = 0.01\text{rad/s}$, $\omega_4 = 0.1\text{rad/s}$ and $\omega_5 = 1\text{rad/s}$ with the following phasors $\bar{\lambda}_{kl}$ ($k = 1, 2, 3$, $l = 1, \dots, 5$)

$$\begin{pmatrix} \bar{\lambda}_{11} & \bar{\lambda}_{12} & \bar{\lambda}_{13} & \bar{\lambda}_{14} & \bar{\lambda}_{15} \\ \bar{\lambda}_{21} & \bar{\lambda}_{22} & \bar{\lambda}_{23} & \bar{\lambda}_{24} & \bar{\lambda}_{25} \\ \bar{\lambda}_{31} & \bar{\lambda}_{32} & \bar{\lambda}_{33} & \bar{\lambda}_{34} & \bar{\lambda}_{35} \end{pmatrix} = \begin{pmatrix} 0.2 & 4e^{-j0.3} & -1 & 3e^j & -2 \\ 4.6e^{-j0.4} & 0 & 0.6e^{j0.64} & 8e^{-j0.6} & 0 \\ 0 & -11.1 & 3.2e^{j0.7} & -0.6e^{-j0.87} & 2e^{j0.1} \end{pmatrix}$$

We construct the matrix $\mathbf{P}^{(b,1)} = \mathbf{T}^{(\text{sin},1)}$ as explained in Section 4.4. It is full row rank. Therefore, by virtue of Theorem 4.4, we conclude the the data set \mathcal{Z}_{cl}^∞ will be informative.

Case 2: To make $\mathbf{P}^{(b,1)}$ full row rank, we needed at least five sinusoids in \mathbf{r} since $\zeta_{u,1} = 9$. However, we can use less sinusoids and still get informativity. For this purpose, we have to focus on the matrix $\mathbf{P}^{(c,1)}$. To make $\mathbf{P}^{(c,1)}$ full row rank, we have to choose the number of sinusoids s such that $2s \geq \zeta_1 - \text{rank}(\mathbf{P}^{(a,1)}) = 1$. Therefore, one sinusoid may be sufficient. Let us verify this property by choosing a signal $\mathbf{r} = (r_1, r_2, r_3)^T$ with one sinusoid at $\omega_1 = 0.3\text{rad/s}$ and with the following phasors $\bar{\lambda}_{kl}$ ($k = 1, 2, 3$, $l = 1$)

$$\bar{\lambda}_{11} = 5e^{j0.3} \quad \bar{\lambda}_{21} = 5e^{j0.3} \quad \bar{\lambda}_{31} = 5e^{j0.3}$$

i.e., $r_1(t) = r_2(t) = r_3(t) = 5 \cos(\omega_1 t + 0.3) \forall t$.

The rank of the corresponding matrix $\mathbf{P}^{(c,1)}$ is equal to 9: it is full row rank. From Theorem 4.4, the data set \mathcal{Z}_{cl}^∞ generated with this excitation \mathbf{r} is thus informative with respect to \mathcal{M} .

Case 3: Let us consider the same sinusoid frequency as in Case 2 but with the following phasors $\bar{\lambda}_{kl}$ ($k = 1, 2, 3$, $l = 1$):

$$\bar{\lambda}_{11} = 5e^{j0.3} \quad \bar{\lambda}_{21} = 0 \quad \bar{\lambda}_{31} = 0$$

i.e., only one entry of \mathbf{r} is excited and the other ones are set to 0. The matrix $\mathbf{P}^{(c,1)}$ has a rank of 9: it is full row rank. From Theorem 4.4, the data set \mathcal{Z}_{cl}^∞ is informative with respect to \mathcal{M} . In this case, we illustrate that we can generate informative data by only exciting one signal in \mathbf{r} (as was observed in [MKBG08]) or, as in the previous case, by putting $r_1 = r_2 = r_3$.

Case 4: It is also possible to have the data informativity without checking the rank of $\mathbf{P}^{(b,1)}$ nor $\mathbf{P}^{(c,1)}$ by using Lemma 2.2. For this purpose, let us consider the following

filtered white noise excitation

$$\mathbf{r}(t) = \underbrace{\begin{pmatrix} \frac{0.12z^{-1}}{1-0.41z^{-1}} & \frac{2.3+5z^{-1}}{1-0.4z^{-1}+0.04z^{-2}} & 0 \\ 1.2z^{-1}-0.5z^{-2} & \frac{7.9}{1+0.5z^{-1}} & 0 \\ 0 & 0 & \frac{-1}{1-0.3z^{-1}} \end{pmatrix}}_{\mathbf{M}(z)} \underbrace{\begin{pmatrix} v_1(t) \\ v_2(t) \\ v_3(t) \end{pmatrix}}_{\mathbf{v}(t)}$$

where $\bar{E}[\mathbf{v}(t)\mathbf{v}^T(t)] = \mathbf{I}_3$. We have that $\Phi_{\mathbf{r}}(\omega) \succ 0$ for almost all frequencies ω . Consequently, from Lemma 2.2, the data set $\mathcal{Z}_{cl}^{\infty}$ is informative with respect to \mathcal{M} .

4.7.5 Verification of the results with Monte-Carlo simulations

In order to confirm that the input choice in Cases 1, 2, 3 and 4 yield informative data, we have applied each of these input vectors to the true system in 1000 identification experiments (with different realizations of the white noise e) and we have identified the estimate $\hat{\theta}_N$ (see (2.9)) for each experiment. We initialize each identification computation with a random parameter vector θ_{init} that lies in the set $\{\theta_{init} \in \mathcal{D}_{\theta} \mid \|\theta_{init} - \theta_0\|^2 < 0.01\}$. For each input vector, we have computed the mean of these 1000 estimates and we have in each case observed that this mean is close to the true parameter vector θ_0 , suggesting the consistency.

Table 4.1: Mean of the identified parameter vector for the first channel with $N = 500$ for different cases over 1000 experiments.

	$\tilde{\theta}_{11,1}$	$\tilde{\theta}_{11,2}$	$\tilde{\theta}_{12,1}$	$\tilde{\theta}_{12,2}$	$\tilde{\theta}_{13,1}$
θ_0	-0.1	0.1	-0.8	-0.7	0.5
Case 1	-0.1004	0.1039	-0.7998	-0.7003	0.5001
Case 2	-0.0990	0.0964	-0.8013	-0.6986	0.4996
Case 3	-0.0964	0.1233	-0.8018	-0.6980	0.4987
Case 4	-0.1000	0.0999	-0.8000	-0.7000	0.5000

Table 4.2: Mean of the identified parameter vector for the second channel with $N = 500$ for different cases over 1000 experiments.

	$\tilde{\theta}_{21,1}$	$\tilde{\theta}_{21,2}$	$\tilde{\theta}_{22,1}$	$\tilde{\theta}_{22,2}$	$\tilde{\theta}_{23,1}$	$\tilde{\theta}_{23,2}$	$\tilde{\theta}_{23,3}$
θ_0	1	-0.5	-1.5	1	4.2	-3	2
Case 1	0.9747	-0.4787	-1.5121	0.9658	4.1924	-2.9920	2.0040
Case 2	0.9861	-0.4910	-1.4891	0.9795	4.1954	-2.9962	1.9987
Case 3	0.9794	-0.4838	-1.4581	0.9653	4.1763	-2.9936	1.9973
Case 4	0.9995	-0.5000	-1.4996	1.0000	4.1999	-2.9999	1.9999

4.8 Conclusion of Chapter 4

To address Research Question 2, we developed in this chapter a condition to verify whether a given excitation vector \mathbf{r} of the type (2.14)-(2.15) yields informative data when the system is operated in closed-loop with a given LTI MIMO controller $\mathbf{K}(z)$. This condition is similar to the one developed in Chapter 3 since it consists on checking if a matrix is full row rank for each channel. This matrix depends on the model structure complexity, on the controller coefficients and on the external excitation parametrization. In the literature, only the costless case has been considered for the derivation of results for the data informativity of MIMO systems. The only result that exists in the literature for nonzero external excitations is the positive definiteness of the power spectrum of \mathbf{r} which is too restrictive. The main contribution of our condition is that we can verify the data informativity for external excitations \mathbf{r} for which the power spectrum of \mathbf{r} degenerates at all frequencies. This condition also clearly separates the contribution of the controller and of the external excitation to the data informativity.

After having extensively studied the linear case in Chapters 3 and 4, we consider the nonlinear case in the next chapter by studying the data informativity for the identification of a particular class of parallel Hammerstein systems.

Data informativity for a particular parallel Hammerstein system

Contents

5.1	Introduction of Chapter 5	85
5.2	System of the study, Prediction Error framework and data informativity	86
5.3	White Gaussian noise excitation	89
5.4	Multisine excitation	89
	5.4.1 Method for the computation of m_2	90
	5.4.2 Result without computation of m_2	91
5.5	Synthesis of the results	92
5.6	A numerical example	92
	5.6.1 True system	92
	5.6.2 Multisine excitation	93
	5.6.3 White Gaussian excitation	95
	5.6.4 Verification of the results with Monte-Carlo simulations	95
5.7	Conclusion of Chapter 5	96

5.1 Introduction of Chapter 5

In Chapters 3 and 4, we have studied the data informativity for MIMO LTI systems in both the open-loop and closed-loop operation. In this chapter, we focus on addressing Research Question 3 by studying the data informativity for a particular class of nonlinear systems. We thus consider the nonlinear class of systems (1.1). This class of systems is a particular class of parallel Hammerstein systems [SPR11, ST17]. Up to

our knowledge, there is no result in the literature that treats the data informativity for the identification of such nonlinear systems within the Prediction Error framework.

For the considered class of Hammerstein systems, we derive results in order to verify whether a given set of data is informative or not. For this purpose, in Section 5.2, we show that we can recast this type of system as an equivalent MISO LTI system with one output and two inputs, i.e., $u_1 = u$ and $u_2 = u^n$. The idea is then to use data informativity results for MISO systems that we have derived in Chapter 3 to tackle Research Question 3. We will address this study for two types of excitations: white Gaussian noise (Section 5.3) and multisine (Section 5.4). The results are summarized in Section 5.5 and illustrated in Section 5.6 in one numerical example. Finally, a conclusion is provided in Section 5.7.

5.2 System of the study, Prediction Error framework and data informativity

Consider the parallel Hammerstein system \mathcal{S} with one input u and one output y described by

$$\mathcal{S}: \quad y(t) = G_{0,1}(z)u(t) + G_{0,2}(z)u^n(t) + H_0(z)e(t) \quad (5.1)$$

where $G_{0,1}(z)$ and $G_{0,2}(z)$ are stable transfer functions, $H_0(z)$ is a stable, inversely stable and monic transfer function, e a white noise and n an integer that will be assumed to be known and such that $n \geq 2$. We will suppose that u and e are independent (i.e., the identification experiment is performed in open-loop). Therefore, u^n and e are also independent. The Figure 5.1 illustrates the system \mathcal{S} considered in (5.1).

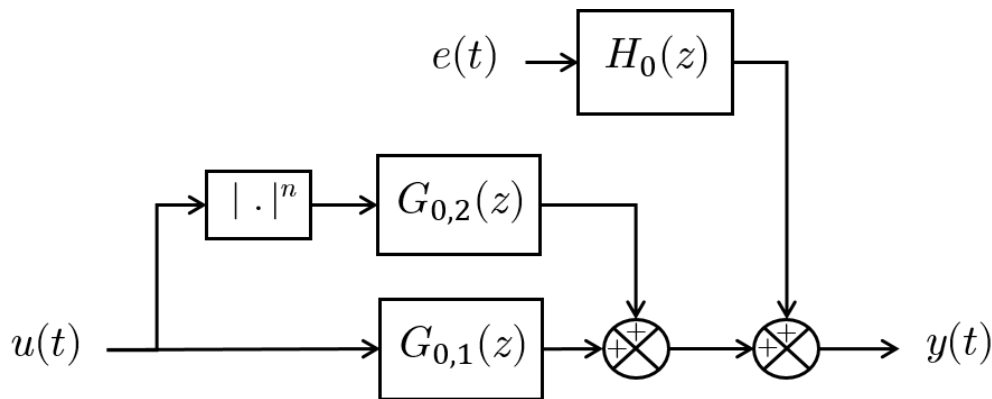


Figure 5.1: Scheme of the true system \mathcal{S} considered in Chapter 5.

As already mentioned in the introduction, we want to identify a model of \mathcal{S} by using the PE identification framework. For this purpose, one could consider a parametrized nonlinear model structure $\mathcal{M}_{nl} = \{(f(u(t), \boldsymbol{\theta}), H(z, \boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \mathcal{D}_\theta\}$ with $f(u(t), \boldsymbol{\theta})$ a nonlinear function in $u(t)$ defined by $f(u(t), \boldsymbol{\theta}) = G_1(z, \boldsymbol{\theta})u(t) + G_2(z, \boldsymbol{\theta})u^n(t)$ where

$G_1(z, \boldsymbol{\theta})$, $G_2(z, \boldsymbol{\theta})$ and $H(z, \boldsymbol{\theta})$ are parametrized rational transfer functions¹ and $\boldsymbol{\theta}$ the parameter vector. The set $\mathcal{D}_\theta \subset \mathbb{R}^m$ restricts $\boldsymbol{\theta}$ to those values for which $G_1(z, \boldsymbol{\theta})$ and $G_2(z, \boldsymbol{\theta})$ are stable and $H(z, \boldsymbol{\theta})$ is stable and inversely stable.

Instead of choosing this model structure, one can notice that the system \mathcal{S} in (5.1) is equivalent to the following MISO system with the input vector $\mathbf{u} = (u, u^n)^T$:

$$\mathcal{S}: \quad y(t) = \mathbf{G}_0(z)\mathbf{u}(t) + H_0(z)e(t)$$

where $\mathbf{G}_0(z) = (G_{0,1}(z), G_{0,2}(z))$.

Therefore, it can be identified within the following MISO (linear) model structure

$$\mathcal{M} = \{(\mathbf{G}(z, \boldsymbol{\theta}), H(z, \boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \mathcal{D}_\theta\} \quad (5.2)$$

with $\mathbf{G}(z, \boldsymbol{\theta}) = (G_1(z, \boldsymbol{\theta}), G_2(z, \boldsymbol{\theta}))$ where $G_1(z, \boldsymbol{\theta})$, $G_2(z, \boldsymbol{\theta})$ and $H(z, \boldsymbol{\theta})$ are the transfer functions defined in \mathcal{M}_{nl} . We will consider that the model structure in (5.2) is full-order and globally identifiable at $\boldsymbol{\theta}_0$. Finally, we will restrict attention to MISO model structures having the ARX, FIR, BJ, OE form to keep the presentation simple (see Section 3.3.1 for the notations of these particular model structures²).

Let us slightly adapt Definition 2.2 to formally define data informativity for the problem considered in this chapter.

Definition 5.1. *Consider the framework defined above and suppose that we apply a quasi-stationary input $u(t)$ to the true system \mathcal{S} in (5.1) from $t = 1$ till $t = +\infty$ yielding the infinite data set \mathcal{Z}^∞ :*

$$\mathcal{Z}^\infty = \left\{ \mathbf{x}(t) = (y(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, +\infty \right\} \quad (5.3)$$

where $\mathbf{u} = (u, u^n)^T$. Consider a MISO model structure \mathcal{M} with $n_u = 2$ yielding the predictor $\hat{y}(t, \boldsymbol{\theta}) = \mathbf{W}(z, \boldsymbol{\theta})\mathbf{x}(t)$. Let us also define the set

$$\Delta_{\mathbf{W}} = \{ \Delta \mathbf{W}(z) = \mathbf{W}(z, \boldsymbol{\theta}') - \mathbf{W}(z, \boldsymbol{\theta}'') \mid \boldsymbol{\theta}' \text{ and } \boldsymbol{\theta}'' \text{ in } \mathcal{D}_\theta \}$$

The data set \mathcal{Z}^∞ is said to be informative w.r.t. the model structure \mathcal{M} when, for all $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}}$, we have

$$\bar{E} [\|\Delta \mathbf{W}(z)\mathbf{x}(t)\|^2] = 0 \implies \Delta \mathbf{W}(z) \equiv \mathbf{0}_{1 \times 3}$$

where $\Delta \mathbf{W}(z) \equiv \mathbf{0}_{1 \times 3}$ means that $\Delta \mathbf{W}(e^{j\omega}) = \mathbf{0}_{1 \times 3}$ for almost all ω and where \bar{E} is the operator defined in (2.1).

¹ $H(z, \boldsymbol{\theta})$ is moreover assumed monic, i.e., $H(z = \infty, \boldsymbol{\theta}) = 1$.

²Note that we have here that $n_u = 2$.

Since we want a consistent estimate of $(G_{0,1}(z), G_{0,2}(z), H_0(z))$, the objective of this chapter is to develop results in order to verify if a given experiment with an excitation u will yield informative data \mathcal{Z}^∞ (see (5.3)).

For this purpose, we will use the results in the literature and the ones that we have already developed for MISO systems in Chapter 3. Indeed these results are valid for arbitrary input vectors \mathbf{u} and thus also for input vectors $\mathbf{u} = (u, u^n)^T$. These results can be found in Lemma 2.1 (positive definiteness of the input power spectrum matrix), Theorem 3.1 (informativity for open-loop MIMO model structures) and Theorem 3.3 (informativity and persistency of the input regressor). In the next lemma, we give an additional result for data informativity w.r.t. MISO systems in open-loop that will be useful in the sequel.

Lemma 5.1. *Consider a MISO linear system with two inputs ($n_u = 2$) and that can be described with a full order MISO model structure as defined in Section 3.3.1. Assume that the input vector $\mathbf{u} = (u_1, u_2)^T$ is a multisine applied to the true system \mathcal{S} in (2.2). Denote by Φ_{u_1} (resp. Φ_{u_2}) the power spectral density (PSD) of u_1 (resp. u_2). Suppose that Φ_{u_1} (resp. Φ_{u_2}) is nonzero at m_1 (resp. m_2) frequencies in $] -\pi, \pi]$. Finally, suppose that Φ_{u_1} and Φ_{u_2} are both nonzero at $m_{1,2}$ frequencies in this interval ($m_{1,2}$ can be equal to zero). Then, the data set \mathcal{Z}^∞ is informative w.r.t. \mathcal{M} if $m_1 \geq \varpi_1 - \rho_1 + 1$ and $m_2 - m_{1,2} \geq \varpi_2 - \rho_2 + 1$ where ρ_1 and ρ_2 are the delays of $G_1(z, \boldsymbol{\theta})$ and $G_2(z, \boldsymbol{\theta})$ respectively. The scalars ϖ_k ($k = 1, 2$) are given by*

- $\varpi_k = \deg(B_k)$ when \mathcal{M} is ARX/FIR.
- $\varpi_k = \deg(B_k) + \deg(F_k)$ when \mathcal{M} is OE/BJ.

Proof. See Appendix D.1 for the proof. □

As already mentioned, Lemma 5.1 and the results in Chapter 3 pertain to an arbitrary input vector \mathbf{u} with $n_u = 2$. They can therefore also be used in the case where the input vector $\mathbf{u} = (u_1, u_2)^T$ is of the form $u_1 = u$ and $u_2 = u^n$. This fact will be used in the sequel to derive data informativity results for the Hammerstein system \mathcal{S} in (5.1). We will consider two types of excitation: white Gaussian noise and multisine. In the next section, we deal with the white Gaussian excitation.

5.3 White Gaussian noise excitation

As shown in Appendix D.2, if the input signal u of (5.1) is chosen as a zero-mean Gaussian white noise, the power spectrum matrix of $\mathbf{u} = (u, u^n)^T$ is strictly positive definite at (almost) all frequencies (whatever the value of $n \geq 2$). Consequently, we have the following result.

Theorem 5.1. *Consider the system \mathcal{S} in (5.1) that can be described by the full-order model structure \mathcal{M} given in (5.2) when $\mathbf{u} = (u, u^n)^T$. Consider also Definition 5.1. Assume that we apply a zero-mean white Gaussian noise u to the system \mathcal{S} in (5.1). Then, the data set \mathcal{Z}^∞ with $\mathbf{u} = (u, u^n)$ is informative, irrespectively of the model structure complexity.*

Proof. See Appendix D.2 for the proof. □

This type of stochastic excitation is interesting since it can allow to identify any system of the type (5.1). Let us now consider the case where the input u of the system (5.1) is a multisine.

5.4 Multisine excitation

In this section, u is a sum of s_1 sinusoids given by

$$u(t) = \sum_{l=1}^{s_1} \lambda_l \cos(\omega_l t + \psi_l) \quad (5.4)$$

where λ_l and ψ_l are respectively the amplitude and the phase shift of the sinusoid at the *nonzero* frequency ω_l belonging to the normalized frequency interval $]0, \pi[$. Since u is a multisine, u^n is also a multisine, made up of more sinusoids than u . In Chapter 3, we gave a condition to verify if a given multisine input vector \mathbf{u} will yield informative data w.r.t. \mathcal{M} by verifying the rank of a matrix depending on the model structure complexity and on the amplitudes, phase shifts and frequencies of the sinusoids in \mathbf{u} (Theorem 3.4). To apply it to our problem, we have to compute the amplitudes, phase shifts and frequencies of u^n which can become computationally expensive when n and s_1 increase. Fortunately, using Lemma 5.1, we can derive a sufficient condition for data informativity that only requires the knowledge of the number of frequencies present in u^n (and not its full expression).

Theorem 5.2. *Consider that the system \mathcal{S} in (5.1) is excited with the multisine (5.4) where $\omega_l \in]0, \pi[$ ($l = 1, \dots, s_1$). Define m_2 and $m_{1,2}$ as in Lemma 5.1 for $\Phi_{u_1} = \Phi_u$ and $\Phi_{u_2} = \Phi_{u^n}$. Then, the data set \mathcal{Z}^∞ in (5.3) is informative w.r.t. \mathcal{M} (see (5.2)) if $2s_1 \geq \varpi_1 - \rho_1 + 1$ and if $m_2 - m_{1,2} \geq \varpi_2 - \rho_2 + 1$ where ϖ_1 and ϖ_2 are defined in Lemma 5.1.*

Proof. Straightforward consequence of Lemma 5.1. □

As already mentioned, to use the (sufficient) data informativity condition of Theorem 5.2, we only need to know how many frequencies are present in u^n and to compare them with the ones in u (see (5.4)).

For this purpose, one could compute the Fast Fourier Transform (FFT) of the sequence $\{u^n(t) \mid t = 1, \dots, N'\}$ for a sufficiently large value of N' . Another procedure to this end will be derived in the next subsection. In this procedure, we will formally suppose that $\lambda_l > 0$ and $\psi_l = 0$ ($l = 1, \dots, s_1$) in (5.4). See Remark 5.1 in the sequel for further details.

5.4.1 Method for the computation of m_2

Let us thus consider (5.4) with $\lambda_l > 0$, $\psi_l = 0$ and $\omega_l \in]0, \pi[$ ($l = 1, \dots, s_1$) and let us observe that $u^n(t) = u^{n-1}(t)u(t)$, $u^{n-1}(t) = u^{n-2}(t)u(t)$, \dots , $u^2(t) = u(t)u(t)$. Therefore, to get the sinusoids of u^n , we can first determine the ones of u^2 from u , then the ones of u^3 from u^2 , \dots , the ones of u^{n-1} from u^{n-2} and finally the ones of u^n from u^{n-1} .

Let us now study how to obtain the frequencies of u^p from the ones of u^{p-1} ($p = 2, \dots, n$). For this purpose, we know that u^{p-1} is a multisine and so $u^p(t) = u^{p-1}(t)u(t)$ can be written as the sum of the products of each sinusoid in u by each sinusoid in u^{p-1} . By using the fact that $\cos(\alpha)\cos(\beta) = 1/2(\cos(\alpha + \beta) + \cos(\alpha - \beta)) \forall (\alpha, \beta) \in \mathbb{R}^2$, all the products in this sum can be written as the sum of two sinusoids.

By doing this, we see that the set of frequencies in u^p can be determined by adding and subtracting ω_l ($l = 1, \dots, s_1$) to the frequencies of each sinusoid present in u^{p-1} . Of course, in the obtained set, we have to remove the duplicates (e.g., the frequency $-\omega'$ is equivalent to the frequency ω' and the frequency $\omega' + 2k\pi$ with $k \in \mathbb{Z}$ is equivalent to ω'). We then obtain a set of frequencies in the interval $[0, \pi]$. Let us denote by s_2 the number of frequencies in this set. Then,

- $m_2 = 2s_2$ if the set contains neither the frequency 0 nor π .
- $m_2 = 2s_2 - 1$ if the set contains either 0 or π .
- $m_2 = 2s_2 - 2$ if the set contains both 0 and π .

The above procedure supposes that no terms in the sum of products can cancel out. This is the reason why we suppose that $\lambda_l > 0$ and $\psi_l = 0$. Consequently, all the terms in the summation will be characterized by a positive amplitude and a zero phase shift and no cancellations can then occur (see Remark 5.1). Let us illustrate the above procedure on an example.

Example 5.1. Consider $n = 3$, $s_1 = 2$ and

$$u(t) = \cos(\omega_1 t) + \cos(\omega_2 t)$$

with $\omega_1 = 0.2$ and $\omega_2 = 0.3$. To compute s_2 , we need the frequencies of the sinusoids in u^3 . For this purpose, we need first to determine the ones in u^2 from u .

► For each frequency in u (i.e., the frequencies 0.2 and 0.3), we add and subtract $\omega_1 = 0.2$ and $\omega_2 = 0.3$. With ω_1 we obtain the terms 0.4, 0, 0.5, 0.1 and with ω_2 we have 0.5, -0.1 , 0.6, 0. By removing the duplicates, we obtain the following frequencies in u^2 : 0, 0.1, 0.4, 0.5 and 0.6.

► For each frequency in u^2 (i.e., the frequencies 0, 0.1, 0.4, 0.5 and 0.6), we add and subtract $\omega_1 = 0.2$ and $\omega_2 = 0.3$. With ω_1 , we obtain the terms 0.2, -0.2 , 0.3, -0.1 , 0.6, 0.2, 0.7, 0.3, 0.8 and 0.4 and with ω_2 , we have 0.3, -0.3 , 0.4, -0.2 , 0.7, 0.1, 0.8, 0.2, 0.9 and 0.3. By removing the duplicates, we obtain the 8 following frequencies in u^3 : 0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8 and 0.9, all belonging to $]0, \pi[$. We see that $s_2 = 8$ and that $m_2 = 16$. Since the frequencies ± 0.2 and ± 0.3 are also in Φ_u , we have that $m_{1,2} = 4$ and thus that $m_2 - m_{1,2} = 12$. \square

Remark 5.1. As already mentioned, the above procedure is not guaranteed to yield the right m_2 if the phase shifts are not all equal to zero. To show this let us consider the following $u(t) = 1/2 \cos(0.4t) + \cos(0.6t) + \cos(0.5t + \pi/2)$ and let us apply the above procedure for $n = 2$: we obtain the set of frequencies 0, 0.1, 0.2, 0.8, 0.9, 1, 1.1, and 1.2. However, the frequency 1 does not appear in u^2 due to the cancellations of the terms at this frequency. Note, however, that such a situation is rare in practice and we can therefore carefully use the procedure in this section when the phase shifts are nonzero. However, a verification using the FFT can be always useful. \square

5.4.2 Result without computation of m_2

In some cases, we even do not have to compute m_2 to check the data informativity.

Lemma 5.2. Consider that the excitation of (5.1) (with $n \geq 2$) is given by $u(t) = \sum_{l=1}^{s_1} \lambda_l \cos(\omega_l t)$ with $\lambda_l > 0$ and $\omega_l \in]0, \pi[$ ($l = 1, \dots, s_1$) and such that the s_1 frequencies $n\omega_l$ are

- (i) all different between them (modulo 2π).
- (ii) all different from π and 0 (modulo 2π).
- (iii) all different from each ω_h ($h = 1, \dots, s_1$) (modulo 2π).

Then, the data set \mathcal{Z}^∞ is informative with respect to \mathcal{M} if

$$s_1 \geq \frac{1}{2} \max(\varpi_1 - \rho_1 + 1, \varpi_2 - \rho_2 + 1)$$

where ϖ_1 and ϖ_2 are defined in Lemma 5.1.

Proof. Following the procedure in Section 5.4.1, we see that $n\omega_1, \dots, n\omega_{s_1}$ are frequencies of the multisine u^n . Since these s_1 positive frequencies are all different (modulo 2π) and different from π (modulo 2π), this implies that the PSD of u^n will be nonzero at at least $2s_1$ frequencies in $] -\pi, \pi[$, i.e., $m_2 \geq 2s_1$. Due to the fact that the s_1 frequencies $n\omega_l$ are (modulo 2π) different from the frequencies ω_h ($h = 1, \dots, s_1$) in u , we have also that $m_2 - m_{1,2} \geq 2s_1$. The result then follows from Theorem 5.2. \square

Remark 5.2. *The idea of Lemma 5.2 is to give a lower bound for $m_2 - m_{1,2}$ which is $2s_1$ when the three conditions (i), (ii) and (iii) are guaranteed. However, this bound is conservative. In the example of Section 5.4.1 where $n = 3$ and $s_1 = 2$, we have seen that $m_2 - m_{1,2} = 12$ while the lower bound is equal to $2s_1 = 4$.* \square

5.5 Synthesis of the results

Let us now summarize the different results of this chapter by giving a general approach to verify the data informativity w.r.t. \mathcal{M} . If u is a white Gaussian noise, then the data set \mathcal{Z}^∞ is informative, irrespectively of the model structure complexity.

Let us now focus on the multisine case. First, we verify that $2s_1 \geq \varpi_1 - \rho_1 + 1$. If it is not the case, the data set \mathcal{Z}^∞ is certainly not informative with respect to \mathcal{M} . If $2s_1 \geq \varpi_1 - \rho_1 + 1$, we can check whether the (conservative) condition of Lemma 5.2 is satisfied. If it is not the case, we compute m_2 and $m_{1,2}$ (using the FFT approach or the procedure of Section 5.4.1) and we verify the (less conservative) condition of Theorem 5.2. If this condition is still not validated, we need to compute the full expression of u^n (with all amplitudes and phase shifts) and use the results of Chapter 3 on this expression and the one of u .

5.6 A numerical example

In this section, we provide a numerical example to illustrate the developed results in this chapter.

5.6.1 True system

Consider the following parallel Hammerstein system \mathcal{S} given by

$$\mathcal{S} : \quad y(t) = G_{0,1}(z)u(t) + G_{0,2}(z)u^3(t) + \underbrace{\frac{1}{1 + 0.5z^{-1} - 0.24z^{-2}}}_{H_0(z)} e(t)$$

where e is a white Gaussian noise of variance 0.001 and with

$$G_{0,1}(z) = \frac{0.1 + 2z^{-1} - 0.5z^{-2} + 0.14z^{-3}}{1 + 0.5z^{-1} - 0.24z^{-2}}$$

$$G_{0,2}(z) = \frac{-0.5 + 0.27z^{-1} + 0.6z^{-2} - 1.5z^{-3} + 0.24z^{-4} + 0.5z^{-6}}{1 + 0.5z^{-1} - 0.24z^{-2}}$$

This system will be identified within the following MISO ARX model structure $\mathcal{M} = \{(\mathbf{G}(z, \boldsymbol{\theta}), H(z, \boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \mathcal{D}_\theta\}$ where $\mathbf{G}(z, \boldsymbol{\theta}) = (G_1(z, \boldsymbol{\theta}), G_2(z, \boldsymbol{\theta}))$ with the delays ρ_k ($k = 1, 2$) given by $\rho_1 = 0$ and $\rho_2 = 0$ and

$$G_1(z, \boldsymbol{\theta}) = z^{-\rho_1} \frac{\mathring{B}_1(z, \tilde{\boldsymbol{\theta}}_1)}{A(z, \boldsymbol{\eta})}$$

$$= \frac{\tilde{\theta}_{1,1} + \tilde{\theta}_{1,2}z^{-1} + \tilde{\theta}_{1,3}z^{-2} + \tilde{\theta}_{1,4}z^{-3}}{1 + \eta_1z^{-1} + \eta_2z^{-2}}$$

$$G_2(z, \boldsymbol{\theta}) = z^{-\rho_2} \frac{\mathring{B}_2(z, \tilde{\boldsymbol{\theta}}_2)}{A(z, \boldsymbol{\eta})}$$

$$= \frac{\tilde{\theta}_{2,1} + \tilde{\theta}_{2,2}z^{-1} + \tilde{\theta}_{2,3}z^{-2} + \tilde{\theta}_{2,4}z^{-3} + \tilde{\theta}_{2,5}z^{-4} + \tilde{\theta}_{2,6}z^{-5} + \tilde{\theta}_{2,7}z^{-6}}{1 + \eta_1z^{-1} + \eta_2z^{-2}}$$

$$H(z, \boldsymbol{\theta}) = \frac{1}{A(z, \boldsymbol{\eta})} = \frac{1}{1 + \eta_1z^{-1} + \eta_2z^{-2}}$$

where $\boldsymbol{\theta} = (\tilde{\boldsymbol{\theta}}_1^T, \tilde{\boldsymbol{\theta}}_2^T, \boldsymbol{\eta}^T)^T$ with $\tilde{\boldsymbol{\theta}}_1 = (\tilde{\theta}_{1,1}, \tilde{\theta}_{1,2}, \tilde{\theta}_{1,3}, \tilde{\theta}_{1,4})^T$, $\tilde{\boldsymbol{\theta}}_2 = (\tilde{\theta}_{2,1}, \tilde{\theta}_{2,2}, \tilde{\theta}_{2,3}, \tilde{\theta}_{2,4}, \tilde{\theta}_{2,5}, \tilde{\theta}_{2,6}, \tilde{\theta}_{2,7})^T$ and $\boldsymbol{\eta} = (\eta_1, \eta_2)^T$. In this case, the scalars ϖ_k ($k = 1, 2$) of Lemma 5.1 are given by $\varpi_1 = 3$ and $\varpi_2 = 6$. The model structure \mathcal{M} is full-order and globally identifiable at all $\boldsymbol{\theta} \in \mathcal{D}_\theta$. Indeed, there exists a unique true parameter vector $\boldsymbol{\theta}_0$ given by

$$\boldsymbol{\theta}_0 = (0.1, 2, -0.5, 0.14, -0.5, 0.27, 0.6, -1.5, 0.24, 0, 0.5, 0.5, -0.24)^T$$

5.6.2 Multisine excitation

Case 1: We consider the following multisine excitation for u with $s_1 = 2$

$$u(t) = \cos(\omega_1 t) + \cos(\omega_2 t)$$

where $\omega_1 = 0.1$ and $\omega_2 = 0.3$. The PSD of u is then nonzero at $m_1 = 2s_1 = 4$ frequencies, which guarantees the necessary condition $m_1 \geq \varpi_1 - \rho_1 + 1$ with $\varpi_1 = 3$. It is clear that, in this case, we cannot apply Lemma 5.2 since $\omega_2 = n\omega_1 = 3\omega_1$ and so the condition (iii) is not guaranteed.

Let us then verify if we can conclude with Lemma 5.1. Since all amplitudes are positive and the phase shifts are equal to 0, we can apply the procedure in Section 5.4.1 to compute m_2 and $m_{1,2}$. We obtain the following sinusoid frequencies in u^3 : 0.1, 0.3, 0.5, 0.7 and 0.9 which implies that $m_2 = 10$ and $m_{1,2} = 4$. Then, we observe that the inequality $m_2 - m_{1,2} \geq \varpi_2 - \rho_2 + 1$ is not guaranteed since $m_2 - m_{1,2} = 6$

and $\varpi_2 - \rho_2 + 1 = 7$. Therefore, we cannot conclude on the data informativity with Lemma 5.2.

The remaining option for the data informativity verification is Theorem 3.3 (regressor persistency). The regressor $\phi_{\mathbf{u}}$ is given by $\phi_{\mathbf{u}} = (\phi_{\mathbf{u}}^T, \phi_{\mathbf{u}^3}^T)^T$ with

$$\begin{aligned}\phi_{\mathbf{u}}^T(t) &= (u(t), u(t-1), u(t-2), u(t-3)) \\ \phi_{\mathbf{u}^3}^T(t) &= (u^3(t), u^3(t-1), u^3(t-2), u^3(t-3), u^3(t-4), u^3(t-5), u^3(t-6))\end{aligned}$$

and so its dimension is equal to $\nu = 11$. However, from the computation of the frequencies in u^3 given in the previous paragraph, we observe that the total number s of frequencies in $\mathbf{u} = (u, u^3)^T$ is equal to 5 (and which are 0.1, 0.3, 0.5, 0.7 and 0.9). Consequently, $\phi_{\mathbf{u}}$ cannot be PE since the condition (3.15) of Lemma 3.3 is not guaranteed. By recalling that the persistency of $\phi_{\mathbf{u}}$ is a necessary and sufficient condition for data informativity in the ARX case (Theorem 3.3), we conclude that the data set \mathcal{Z}^∞ with this input choice is not informative w.r.t. \mathcal{M} .

Since the condition (3.15) is not respected in Case 1, we can add another frequency in u in order to yield data informativity. However, we will see in the next case that a multisine with $s_1 = 2$ frequencies can also yield data informativity.

Case 2: We consider the same multisine excitation for u as before with $s_1 = 2$

$$u(t) = \cos(\omega_1 t) + \cos(\omega_2 t)$$

but with the new frequencies $\omega_1 = 0.4$ and $\omega_2 = 0.8$. In this case, the three conditions (i), (ii) and (iii) of Lemma 5.2 are respected. However, we have that $\frac{1}{2} \max(\varpi_1 - \rho_1 + 1, \varpi_2 - \rho_2 + 1) = 3$ but $s_1 = 2$ and so the inequality $s_1 \geq \frac{1}{2} \max(\varpi_1 - \rho_1 + 1, \varpi_2 - \rho_2 + 1)$ is not respected. Therefore, we cannot conclude on the informativity with Lemma 5.2.

Let us then verify if we can conclude with Lemma 5.1. Since all amplitudes are positive and the phase shifts are equal to 0, we can apply the procedure in Section 5.4.1 to compute m_2 and $m_{1,2}$. We obtain the following sinusoid frequencies in u^3 : 0.4, 0.8, 1.2, 1.6 and 2 and 2.4 which implies that $m_2 = 12$ and $m_{1,2} = 4$. Then, we observe that the inequality $m_2 - m_{1,2} \geq \varpi_2 - \rho_2 + 1$ is indeed respected since $m_2 - m_{1,2} = 8$ and $\varpi_2 - \rho_2 + 1 = 7$. Therefore, from Lemma 5.1, the data set \mathcal{Z}^∞ is informative w.r.t. \mathcal{M} .

Case 3: Finally, we consider a multisine excitation with $s_1 = 3$

$$u(t) = \cos(\omega_1 t) + \cos(\omega_2 t) + 2 \cos(\omega_3 t)$$

with the frequencies $\omega_1 = 0.4$, $\omega_2 = 0.8$ and $\omega_3 = 0.01$. In this case, the three conditions (i), (ii) and (iii) of Lemma 5.2 are respected. Moreover, the inequality $s_1 \geq \frac{1}{2} \max(\varpi_1 - \rho_1 + 1, \varpi_2 - \rho_2 + 1)$ is guaranteed. Therefore, from Lemma 5.2, the data set \mathcal{Z}^∞ is informative w.r.t. \mathcal{M} .

5.6.3 White Gaussian excitation

Case 4: let us consider now consider a white Gaussian noise excitation with a variance equal to 1. From Theorem 5.1, the data are always informative when u is a white Gaussian noise.

5.6.4 Verification of the results with Monte-Carlo simulations

In order to confirm that the input choice in Cases 2, 3 and 4 yield informative data, we have applied each of these input vectors to the true system in 1000 identification experiments (with different realizations of the white noise e) and we have identified the estimate $\hat{\theta}_N$ (see (2.9)) for each experiment. We initialize each identification computation with a random parameter vector θ_{init} that lies in the set $\{\theta_{init} \in \mathcal{D}_\theta \mid \|\theta_{init} - \theta_0\|^2 < 0.01\}$. For each input vector, we have computed the mean of these 1000 estimates. The results are given in Tables 5.1 and 5.2 and we have in each case observed that this mean is almost equal to the true parameter vector θ_0 , suggesting the consistency.

Table 5.1: Mean of the identified parameter vector for the first channel with $N = 1000$ for different cases over 1000 experiments.

	$\tilde{\theta}_{1,1}$	$\tilde{\theta}_{1,2}$	$\tilde{\theta}_{1,3}$	$\tilde{\theta}_{1,4}$
θ_0	0.1	2	-0.5	0.14
Case 2	0.1014	1.9970	-0.4971	0.1389
Case 3	0.1005	1.9987	-0.4988	0.1394
Case 4	0.1000	2.0000	-0.5000	0.1400

Table 5.2: Mean of the identified parameter vector for the second channel with $N = 1000$ for different cases over 1000 experiments.

	$\tilde{\theta}_{2,1}$	$\tilde{\theta}_{2,2}$	$\tilde{\theta}_{2,3}$	$\tilde{\theta}_{2,4}$	$\tilde{\theta}_{2,5}$	$\tilde{\theta}_{2,6}$	$\tilde{\theta}_{2,7}$
θ_0	-0.5	0.27	0.6	-1.5	0.24	0	0.5
Case 2	-0.5001	0.2701	0.5999	-1.5000	0.2401	0.0001	0.5000
Case 3	-0.5000	0.2700	0.6000	-1.5000	0.2400	0	0.5000
Case 4	-0.5000	0.2700	0.6000	-1.5000	0.2400	0	0.5000

5.7 Conclusion of Chapter 5

The goal of this chapter was to address Research Question 3 given in Section 1.4 by studying the data informativity for a particular class of parallel Hammerstein systems given in (5.1) within the Prediction Error framework. The idea that had been followed in this chapter was to recast this nonlinear system into a MISO LTI one with two inputs $u_1 = u$ and $u_2 = u^n$. This allowed us to use previous results developed in Chapter 3 such as Theorem 3.1 in order to address Research Question 3. Then, we analyzed the data informativity for two types of excitations: white Gaussian noise and multisine. The white Gaussian excitation is an efficient choice for the excitation since it always guarantees informative data for this type of parallel Hammerstein systems. When one goes for a multisine excitation, then simple verification can be performed with Lemmas 5.1 and 5.2 that had been derived in order to simplify the data informativity verification.

Based on this study, we will see later that some results developed in this chapter will be useful to address the second and practical part of the thesis: the identification of a MEMS gyroscope. The presentation of this real-life system is given in the next chapter.

Practical work on the identification of a MEMS gyroscope

MEMS gyroscope and modeling objectives

Contents

6.1	Introduction of Chapter 6	99
6.2	Description of the MEMS gyroscope architecture	101
6.3	General principle of angular rate deduction	101
6.4	Main dynamics in the MEMS gyroscope	103
6.4.1	Dynamics of both masses	103
6.4.2	Actuators and sensors: comb-drives	104
6.4.3	Additional dynamics of the MEMS gyroscope	107
6.5	Controller and principle of angular rate deduction	112
6.5.1	Closed-loop scheme for angular rate deduction	112
6.5.2	Angular rate deduction	113
6.5.3	Modeling approach considered for the control of the MEMS gyroscope	114
6.6	Conclusion of Chapter 6	115

6.1 Introduction of Chapter 6

In the second (practical) part of this dissertation, we consider the problem treated in Research Question 4, i.e., the identification of a MEMS gyroscope for its control. In this chapter, we present the MEMS gyroscope, its control and modeling objectives. The objective of this chapter is to develop a global model structure for the identification of the dynamics of the MEMS gyroscope (first part of Research Question 4).

In many applications such as avionics and aerospace engineering, it is important to measure the motion properties of a flying object during its flight. One of these properties is the angular rate (or rotation speed) around a given axis. The sensor used for this purpose is called a gyroscope. There are different types of gyroscopes, each of them based on a different physics principle. For instance, we can cite the atomic gyroscopes based on the spinning of electrons [FQ12, KKD11], the optical gyroscopes exploiting the Sagnac effect [DTCA14, RA81] and the MEMS gyroscopes which use the Coriolis effect [Kem11, Apo16, Sau08]. The last type of gyroscopes is a cheap, low-energy consuming and small solution for the measurement of angular rates. However, it is less accurate than its optical and atomic counterparts, suffering from important problems such as manufacturing imperfections. Moreover, the capacitive instrumentation that is often used for this application adds parasitic dynamics that can deteriorate the quality of the angular rate estimate if not considered in the process of the angular rate measurement. A large amount of studies focus on its manufacturing process in order to attenuate these parasitic effects [RASB05, BHY⁺02]. Another research line that is considered is the development of methods for the design of optimal controllers in order to increase the accuracy of the angular rate estimation [Sau08, EMK12, PTKH09, FB09b].

The Next4MEMS project funding this thesis intends to develop optimal MEMS gyroscopes. The goal of the laboratory Ampère participating in this project is the development of an optimal method for the design of a controller. The approach for the controller design that is considered in the thesis of Fabrício SAGGIN (another Ph.D. student in the Next4MEMS project) is based on the \mathcal{H}_∞ method. However, for the controller design, it is important to design an accurate model of the dynamics of the MEMS gyroscopes. As mentioned in Section 1.5, the literature does not propose a global model structure and an identification method for the three main dynamics of the MEMS gyroscope which are the mechanical coupling, the parasitic capacitive coupling and the noise measurement.

For this purpose, we will construct a global model structure in this chapter, based on the first principles. In Sections 6.2 and 6.3, we give a brief presentation of the MEMS gyroscope and the principle of angular rate deduction. Then, the instrumentation and the additional different dynamics that come with it are presented in Section 6.4. In Section 6.5, we give the control objectives in order to estimate accurately the angular rate. After this presentation of the controller architecture, we present the modeling objectives for the controller design in Section 6.5.3. Finally, in Section 6.6, we give a conclusion of this chapter.

Notation: we will denote by t_c the continuous time while t will still correspond to the integer indexing the values of a discrete-time signals $x(t)$ (see Section 2.2). For a signal x , the notations \dot{x} and \ddot{x} denote its first and second derivative w.r.t. the time.

6.2 Description of the MEMS gyroscope architecture

The MEMS (Micro Electro-Mechanical Structure) gyroscope is an inertial sensor used to measure an angular rate (or rotation speed) of an object around a given axis. Here, we will consider a rotation motion around the z -axis. This micro-sensor consists of a structure fixed to the object whose angular rate is measured. Attached to this structure with micro silicon-based beams, a first mass m_x oscillates along the x -axis. A second mass m_y , oscillating along the y -axis, is attached to the mass m_x with micro silicon-based beams too. These micro-beams are mechanically acting like springs and dampers due to viscous effects. They are designed such that both masses form lightly damped oscillators and we will denote by $\omega_{0,x}$ (resp. $\omega_{0,y}$) the resonance frequency of m_x (resp. m_y). The angular rate around the z -axis is denoted Ω and we are interested in measuring it. The Figure 6.1 depicts a simplified scheme of the MEMS gyroscope.

Each mass can be actuated with external forces: $\vec{f}_x(t_c)$ for the mass m_x and $\vec{f}_y(t_c)$ for the mass m_y . The instrumentation to generate both forces f_x and f_y and to measure both mass displacements (x_{mec} for the mass m_x and y_{mec} for the mass m_y) will be explained later. In the next section, we give a *high level* explanation of the principle of angular rate deduction in a MEMS gyroscope. A more technical presentation of this principle will be explained later in this chapter. To avoid any possible confusion in the sequel, the term *measurement* will be used for the measurement of the mass displacements and the term *deduction* will be used for the angular rate measurement.

6.3 General principle of angular rate deduction

In this paragraph, we present briefly the way to obtain the angular rate Ω . A more technical presentation will be given later in Section 6.5. When $\Omega \neq 0$, both masses are subjected to an inertial effect that is called the Coriolis effect. This effect is explained in Appendix E.1. The Coriolis force on the mass m_x is $f_{cor,x} = 2m_y\Omega\dot{y}_{mec}$ and the one on the mass m_y is $f_{cor,y} = -2m_x\Omega\dot{x}_{mec}$. It is important to note that the Coriolis force is proportional to the angular rate Ω . This fact will be exploited later for its estimation. The idea of the deduction of the angular rate Ω is based on the estimation of the Coriolis effect on the mass m_y , i.e., the term $f_{cor,y} = -2m_x\Omega\dot{x}_{mec}$. This term explains the role of both masses:

- Mass m_x : for the term $-2m_x\Omega\dot{x}_{mec}$ to exist, it is necessary to have a nonzero velocity \dot{x}_{mec} . Therefore, the role of the mass m_x is the generation of this term through its motion. This mass is also called the drive mass.
- Mass m_y : we estimate the Coriolis term $-2m_x\Omega\dot{x}_{mec}$ thanks to the displacement of the mass m_y . This mass is also called the sense mass.

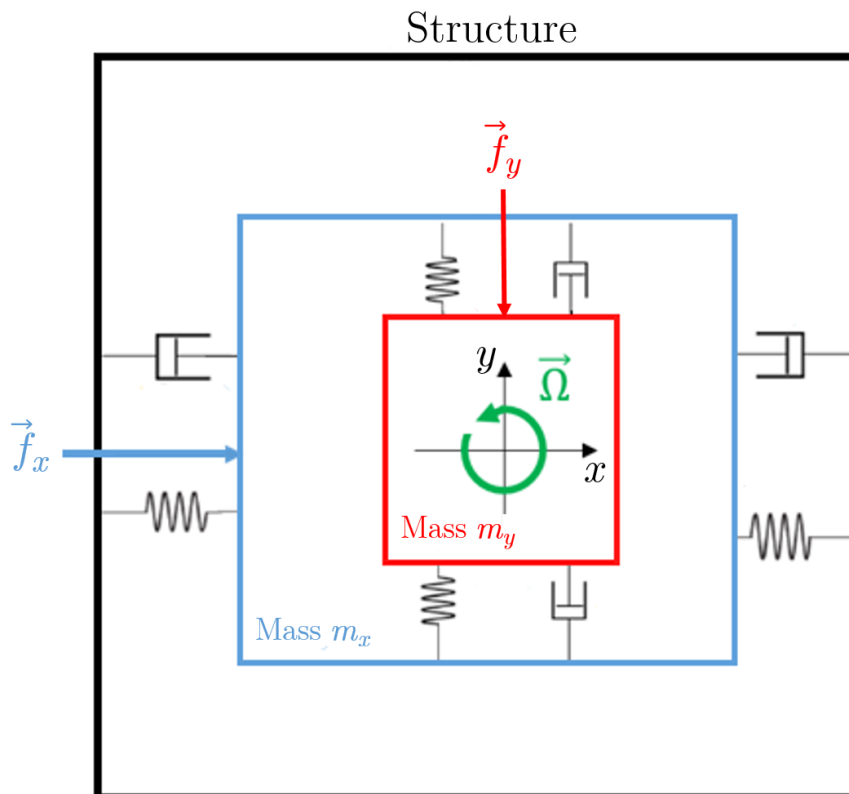


Figure 6.1: Simplified scheme of the MEMS gyroscope.

To be more specific, the general idea for the angular rate deduction is based on three general steps. First, the mass m_x is excited in order to have a nonzero velocity \dot{x}_{mec} . Secondly, the Coriolis term $f_{cor,y} = -2m_x\Omega\dot{x}_{mec}$ is estimated from the motion of the sense mass. Finally, the third step consists in deducing Ω from the estimate of $f_{cor,y} = -2m_x\Omega\dot{x}_{mec}$. We will give later a more technical explanation on how these three steps are realized in practice (see Section 6.5). In order to answer the first part of Research Question 4, we search for a more global model structure for the dynamics of the MEMS gyroscope than those that have been proposed in the literature. For this purpose, in the next section, we present the main dynamics of the MEMS gyroscope that will be modeled in this dissertation and we will obtain this general model structure.

6.4 Main dynamics in the MEMS gyroscope

In this section, we present the main dynamics of the MEMS gyroscope.

6.4.1 Dynamics of both masses

When $\Omega \neq 0$, the Newton's second law on each mass gives¹

$$\begin{aligned} m_x \ddot{x}_{mec}(t_c) + d_{xx} \dot{x}_{mec}(t_c) + k_{xx} x_{mec}(t_c) + d_{xy} \dot{y}_{mec}(t_c) + k_{xy} y_{mec}(t_c) \\ = f_x(t_c) + 2m_y \Omega(t_c) \dot{y}_{mec}(t_c) \end{aligned} \quad (6.1)$$

$$\begin{aligned} m_y \ddot{y}_{mec}(t_c) + d_{yy} \dot{y}_{mec}(t_c) + k_{yy} y_{mec}(t_c) + d_{yx} \dot{x}_{mec}(t_c) + k_{yx} x_{mec}(t_c) \\ = f_y(t_c) - 2m_x \Omega(t_c) \dot{x}_{mec}(t_c) \end{aligned} \quad (6.2)$$

where k_{xx} , k_{xy} , k_{yx} , k_{yy} are stiffness coefficients and d_{xx} , d_{xy} , d_{yx} , d_{yy} are damping coefficients. The cross-coupling coefficients k_{xy} , k_{yx} , d_{xy} and d_{yx} are due to manufacturing imperfections which cause anisoelectricity (k_{xy} and k_{yx}) and nonproportional damping (d_{xy} and d_{yx}) [Sau08]. The Coriolis effect is also cross-coupling both modes when $\Omega \neq 0$. A scheme block of these mechanical dynamics are given in Figure 6.2.

When $\Omega = 0$, the dynamics of both masses can be seen as a MIMO system where the input vector is $(f_x(t_c), f_y(t_c))^T$ and the output vector is $(x_{mec}(t_c), y_{mec}(t_c))^T$.

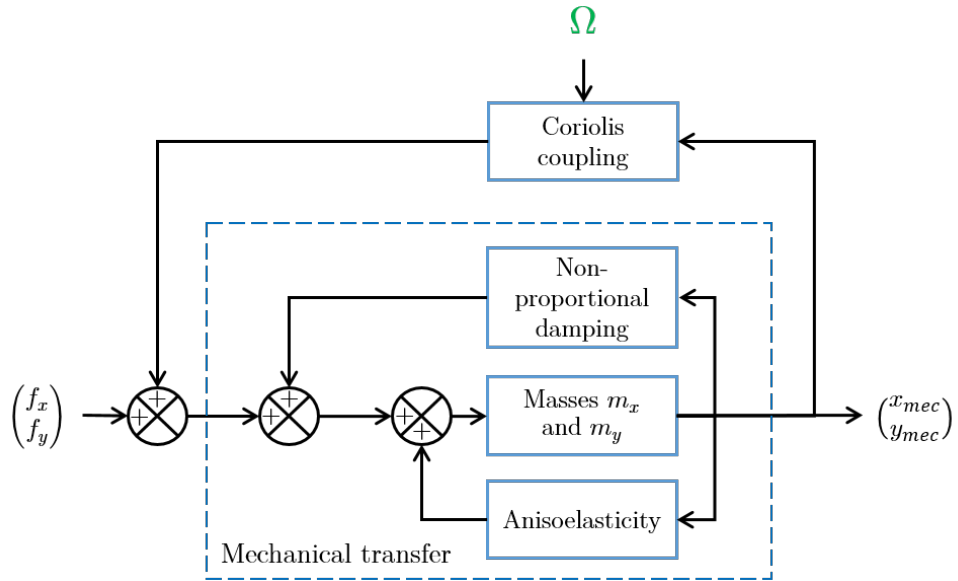


Figure 6.2: Block scheme of the mechanical dynamics of the MEMS gyroscope.

¹Since the displacements of the masses have a microscopic magnitude order, we have assumed that we can neglect the centrifugal and Euler's forces (see Appendix E.1 for more details). The effect of gravity is also neglected due to the scale effect.

However, in practice, the external forces f_x and f_y are not the real inputs of the MEMS as well as x_{mec} and y_{mec} are not the real measured outputs. Indeed, we have to consider the instrumentation of both masses. The one considered in the gyroscope of this study is described in the next paragraph.

6.4.2 Actuators and sensors: comb-drives

The mass motions are actuated and measured with an electrostatic instrumentation² using comb-drives. Comb-drives are microelectromechanical actuators and sensors which are made up of two electrically conductive combs (finger structures) that are interdigitated, i.e., the teeth of one comb is opposite to the holes of the other comb. One comb is fixed (static) to a structure and the other one is moving. A comb-drive is illustrated in Figure 6.3. The main advantageous property of this architecture is that the generated capacitance between both combs is proportional to the distance between them (distance d in Figure 6.3). However, this property is only valid for small distances. Therefore, comb-drives are often used in micro- and nano-scale applications such as the MEMS gyroscope.

In the MEMS gyroscope, there are four comb-drives in total, two per mass. For each mass, one comb-drive is used for the force actuation and the other one is used for the displacement measurement. For the drive mode, for each comb-drive, the static comb is fixed on the structure and the moving one is fixed on the drive mass. For the sense mode, for each comb-drive, the static comb is fixed on the drive mass and the moving one is fixed on the sense mass. The comb-drives are placed as illustrated in the Figure 6.4.

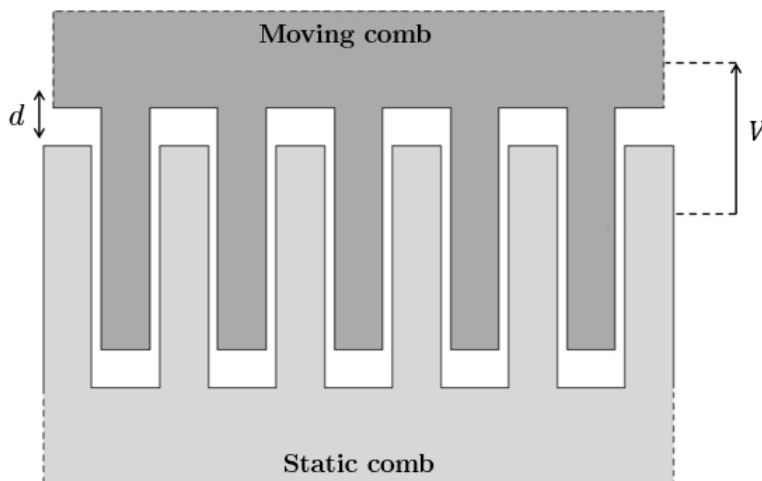


Figure 6.3: Simplified scheme of a comb-drive (Source: [KBS⁺17]).

²Note that there exist other ways to actuate and measure the modes displacements. In [YAN98] they consider piezo-electrical instrumentation for their MEMS gyroscope.

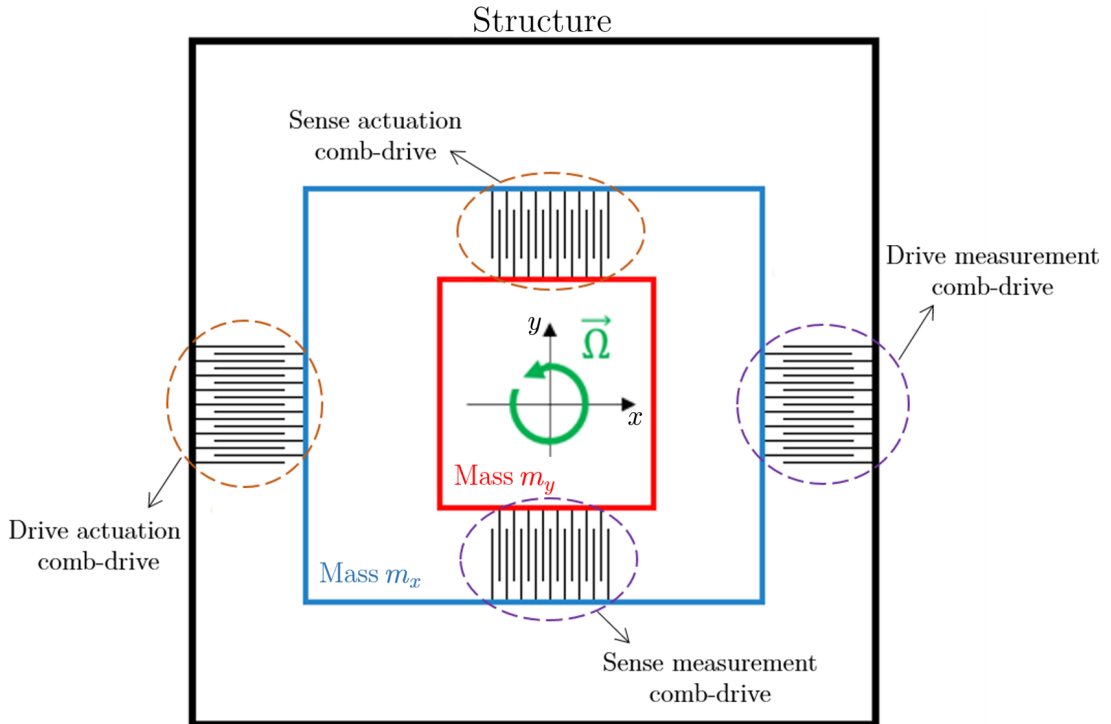


Figure 6.4: Simplified scheme of the MEMS gyroscope, instrumented with capacitive comb-drives.

For the actuation, we excite the MEMS gyroscope with a discrete-time input $u_x(t)$ for the drive mode (resp. $u_y(t)$ for the sense mode) that is converted into a voltage $V_x(t_c)$ (resp. $V_y(t_c)$) through a digital-to-analog converter (DAC) which is a zero-order hold (ZOH). The voltage V_x (resp. V_y) is applied between the two combs of the drive (resp. sense) actuation comb-drive (see Figure 6.3 with the voltage V). This results in electrostatic forces f_x and f_y proportional to the square of the corresponding voltage, i.e.,

$$f_x(t_c) = \alpha_x V_x^2(t_c) \quad (6.3)$$

$$f_y(t_c) = \alpha_y V_y^2(t_c) \quad (6.4)$$

where α_x and α_y are constant³ depending on the electrostatic and geometric properties of the comb fingers. We can observe that there is a square input nonlinearity.

The discrete-time signals u_x and u_y constitute the inputs of the MEMS gyroscope yielding the excitation voltages V_x and V_y through the ZOH. Consequently, the voltages

³As long as the displacements of the masses are kept small, these constant do not depend on the time and on the displacements of the masses. This advantage justifies that this type of actuator is often used in the micro-scale.

$V_x(t_c)$ and $V_y(t_c)$ are given by

$$\begin{aligned} V_x(t_c) &= u_x(t) \quad \text{for } tT_s \leq t_c < (t+1)T_s \\ V_y(t_c) &= u_y(t) \quad \text{for } tT_s \leq t_c < (t+1)T_s \end{aligned}$$

where T_s is the sampling time. The sampling frequency $\omega_s = 2\pi/T_s$ is equal to $2\pi \cdot 62500$ rad/s in the gyroscope considered in this study.

With the ZOH, V_x and V_y have staircase form and the square nonlinearity is static. Therefore, both equations (6.3)-(6.4) can be rewritten as follows

$$\begin{aligned} f_x(t_c) &= \alpha_x u_x^2(t) \quad \text{for } tT_s \leq t_c < (t+1)T_s \\ f_y(t_c) &= \alpha_y u_y^2(t) \quad \text{for } tT_s \leq t_c < (t+1)T_s \end{aligned}$$

For the measurement of both displacements $x_{mec}(t_c)$ and $y_{mec}(t_c)$, we use the measurement comb-drives illustrated in Figure 6.4 and we exploit the fact that the capacitance between two combs is proportional to the distance between these combs (distance d in Figure 6.3). If we can measure this capacitance, then the distance d will be known up to a multiplicative constant. This is the idea which is used in the MEMS gyroscope to measure both displacements $x_{mec}(t_c)$ and $y_{mec}(t_c)$.

For this purpose, each measurement comb-drive of the MEMS (see Figure 6.4) is put into an electronic circuit that proportionally converts the capacitance into a voltage. This voltage is also proportional to the capacitance and so proportional to the displacement of the corresponding mass. Then, these voltages are amplified, filtered through a first order anti-aliasing filter of cut-off frequency $2\pi \cdot 15000$ rad/s and converted in the discrete-time domain with an analog-to-digital converter (ADC). Consequently, the outputs of the ADC are two discrete-time signals which will be denoted $\tilde{x}_{mec}(t)$ and $\tilde{y}_{mec}(t)$ and which are assumed to be proportional to $x_{mec}(t_c = tT_s)$ and $y_{mec}(t_c = tT_s)$ respectively, i.e.,

$$\tilde{x}_{mec}(t) = \beta_x x_{mec}(t) \tag{6.5}$$

$$\tilde{y}_{mec}(t) = \beta_y y_{mec}(t) \tag{6.6}$$

where β_x and β_y are constant depending on the electrostatic and geometric properties of the comb-drives and on the properties of the electronic converter circuit and amplifiers. In Figure 6.5, we give a scheme block of the instrumented mechanical dynamics of the MEMS.

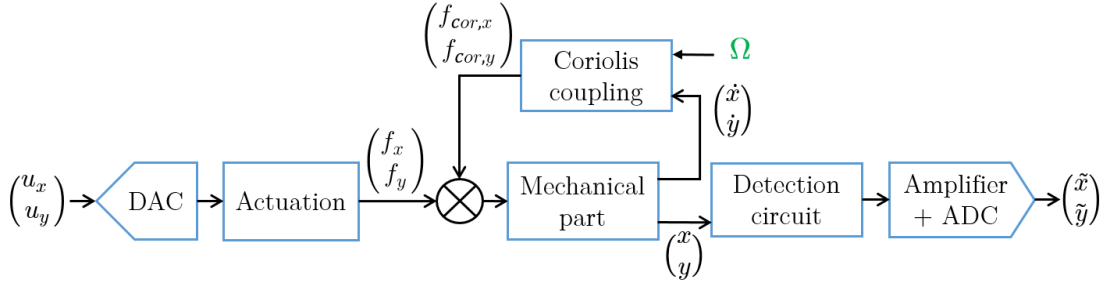


Figure 6.5: Block scheme of the instrumented mechanical dynamics of the MEMS.

Therefore, the outputs of the mechanical dynamics are the discrete-time signals $\tilde{x}_{mec}(t)$ and $\tilde{y}_{mec}(t)$. Since we have seen that the inputs of the mechanical dynamics are u_x^2 and u_y^2 , the mechanical dynamics can be represented by a 2×2 discrete-time LTI matrix $\mathbf{G}_0(z)$ of transfer functions when $\Omega = 0$, i.e.,

$$\begin{pmatrix} \tilde{x}_{mec}(t) \\ \tilde{y}_{mec}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} G_{0,xx}(z) & G_{0,xy}(z) \\ G_{0,yx}(z) & G_{0,yy}(z) \end{pmatrix}}_{\mathbf{G}_0(z)} \begin{pmatrix} u_x^2(t) \\ u_y^2(t) \end{pmatrix}$$

The transfer functions in $\mathbf{G}_0(z)$ depend on the parameters in (6.1) and (6.2) and on the parameters α_x , α_y , β_x et β_y . Each transfer function $G_{0,xx}(z)$ and $G_{0,yy}(z)$ has a lightly damped resonant behavior at the resonance frequency $\omega_{0,x}$ for $G_{0,xx}(z)$ and at the resonance frequency $\omega_{0,y}$ for $G_{0,yy}(z)$. The transfer functions $G_{0,xy}(z)$ and $G_{0,yx}(z)$ have two lightly damped resonances at both resonance frequencies $\omega_{0,x}$ and $\omega_{0,y}$.

However, the use of the instrumentation creates some additional dynamics and so the real measured outputs are not equal to $\tilde{x}_{mec}(t)$ and $\tilde{y}_{mec}(t)$. These additional dynamics are presented in the next paragraph.

6.4.3 Additional dynamics of the MEMS gyroscope

In this paragraph, we will present some additional dynamics involved in the MEMS gyroscope due to the use of the capacitive instrumentation. We will see that the measured outputs of the MEMS are two discrete-time signals \tilde{x} and \tilde{y} which are each the sum of three signals:

$$\begin{aligned} \tilde{x} &= \tilde{x}_{mec} + \tilde{x}_{par} + v_x \\ \tilde{y} &= \tilde{y}_{mec} + \tilde{y}_{par} + v_y \end{aligned}$$

where \tilde{x}_{mec} and \tilde{y}_{mec} are the mechanical outputs given in (6.5) and in (6.6) respectively. In the next two paragraphs, we explain what are behind the signals \tilde{x}_{par} , v_x , \tilde{y}_{par} and v_y .

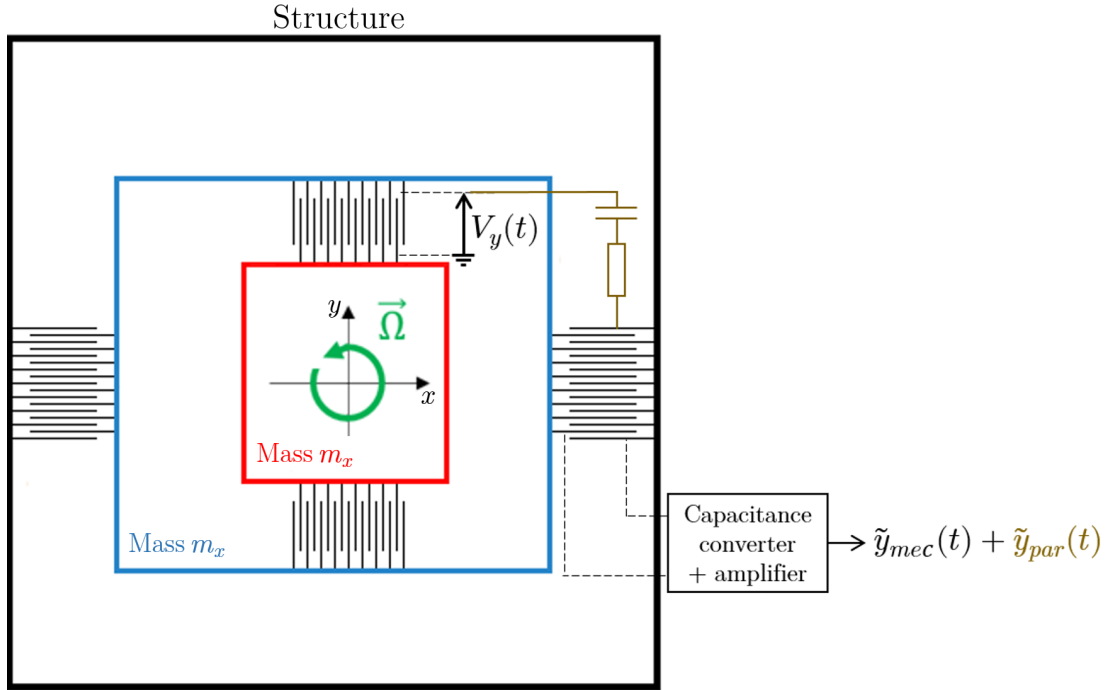


Figure 6.6: Illustration of one part of the parasitic electrical coupling.

Second dynamic: parasite capacitive coupling

The use of comb-drives for both actuators and sensors create some interactions between the generated electrical fields. A parasitic resistor-capacitive bond is made between the actuation and measurement comb-drives [AS05]. This results in a feed-through of the excitation voltages V_x and V_y to the measurement circuits, parasiting the mechanical discrete-time outputs \tilde{x}_{mec} and \tilde{y}_{mec} . Figure 6.6 illustrates one of these couplings.

In the literature, this parasitic effect is modeled as a disturbance vector $(\tilde{x}_{par}, \tilde{y}_{par})^T$ added to the signal vector $(\tilde{x}_{mec}, \tilde{y}_{mec})^T$. Moreover, in the literature, the disturbance vector $(\tilde{x}_{par}, \tilde{y}_{par})^T$ can be modeled as the output vector of a highpass linear filter where the discrete-time input vector is $(u_x, u_y)^T$.

In a systemic point of view, the dynamics of the parasitic coupling corresponds to the following discrete-time expression

$$\begin{pmatrix} \tilde{x}_{par}(t) \\ \tilde{y}_{par}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} Q_{0,xx}(z) & Q_{0,xy}(z) \\ Q_{0,yx}(z) & Q_{0,yy}(z) \end{pmatrix}}_{\mathbf{Q}_0(z)} \begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix}$$

where $\mathbf{Q}_0(z)$ is a 2×2 matrix of transfer functions.

Third dynamic: measurement noises

The third type of dynamics is the measurement noise. In this dissertation, we will assume that this effect can be modeled as stochastic disturbances v_x and v_y added to $\tilde{x}_{mec} + \tilde{x}_{par}$ and $\tilde{y}_{mec} + \tilde{y}_{par}$ respectively such that the measured outputs \tilde{x} and \tilde{y} are given by, when $\Omega = 0$,

$$\begin{pmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} Q_{0,xx}(z) & Q_{0,xy}(z) \\ Q_{0,yx}(z) & Q_{0,yy}(z) \end{pmatrix}}_{\mathbf{Q}_0(z)} \begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix} + \underbrace{\begin{pmatrix} G_{0,xx}(z) & G_{0,xy}(z) \\ G_{0,yx}(z) & G_{0,yy}(z) \end{pmatrix}}_{\mathbf{G}_0(z)} \begin{pmatrix} u_x^2(t) \\ u_y^2(t) \end{pmatrix} + \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} \quad (6.7)$$

This is illustrated in the block scheme in Figure 6.7. The output voltages are noisy due to two major phenomena : flicker noise and mechanical-thermal noise [Sau08]

- Flicker noise: this noise comes from the presence of active electronic components such as amplifiers. The power spectrum of this noise is proportional to $1/\omega$ in the low frequency range. After a particular frequency, this noise is overshadowed by the other noise sources.
- Mechanical-thermal noise: as the gyroscope is a dissipative system, it is affected by thermal noise. From the equipartition theorem of energy and the Nyquist relation, the power spectral density of the noises on the drive and sense mode outputs ($i = \{x, y\}$) are given by $\Phi_{therm}(\omega) = 4k_b T d_{ii}$ with k_b the Boltzmann constant, T the absolute temperature and d_{ii} the damping of the mode represented by the direction $i = \{x, y\}$. The power spectrum of this noise source is then flat.

For the controller design, it is also important to estimate the noise spectrum of $(v_x, v_y)^T$. For this purpose, we will assume that there exist two transfer functions $H_{0,x}(z)$ and $H_{0,y}(z)$ such that

$$\begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} = \underbrace{\begin{pmatrix} H_{0,x}(z) & 0 \\ 0 & H_{0,y}(z) \end{pmatrix}}_{\mathbf{H}_0(z)} \begin{pmatrix} e_x(t) \\ e_y(t) \end{pmatrix} \quad (6.8)$$

where $(e_x, e_y)^T$ is a white noise vector. To estimate the power spectrum of $(v_x, v_y)^T$, we will need to identify $H_{0,x}(z)$ and $H_{0,y}(z)$ as well as to estimate the variances of e_x and e_y .

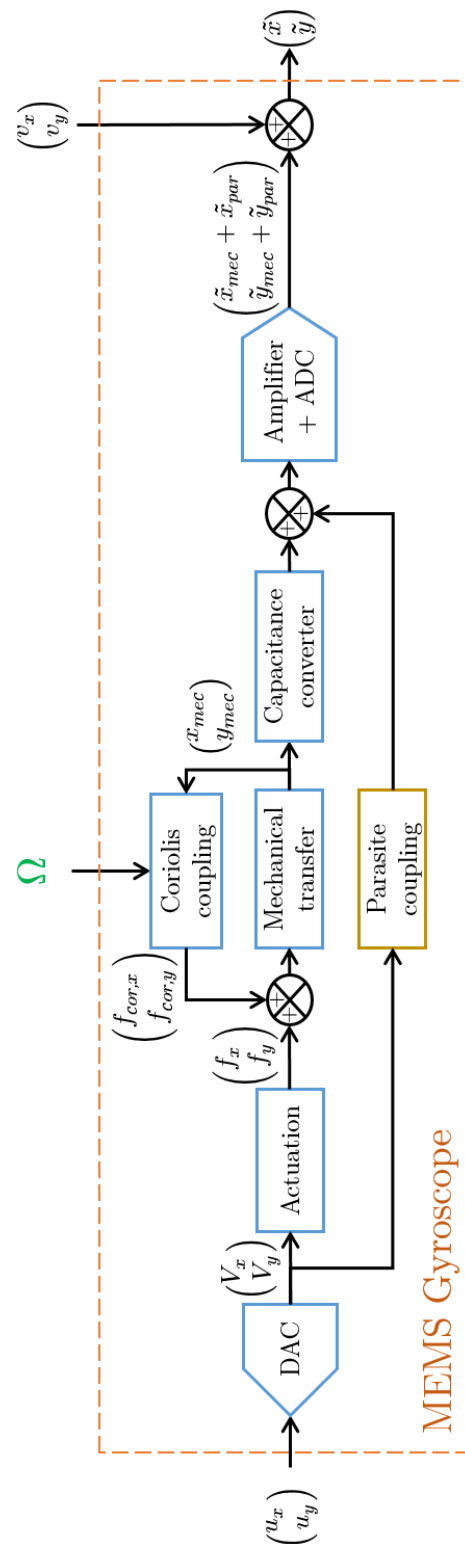


Figure 6.7: Block scheme of the main dynamics of the MEMS gyroscope.

Other nonlinear phenomena can be involved in the MEMS gyroscope such as the temperature dependency, the Spring Softening effect, etc. However, we will assume that we can limit these effects⁴. For the temperature dependency, the temperature around the MEMS gyroscope will be kept constant and equal to 20°C by putting it in a thermal chamber.

The model given in (6.7) combined with (6.8) is more complete than the ones available in the literature. A scheme block is provided in Figure 6.8 to illustrate this complete model structure when $\Omega = 0$. In the next paragraph, we give the technical details on how to obtain the angular rate Ω . We will also see how essential it is to estimate all the transfer functions in (6.7) in order to guarantee an accurate estimate of Ω .

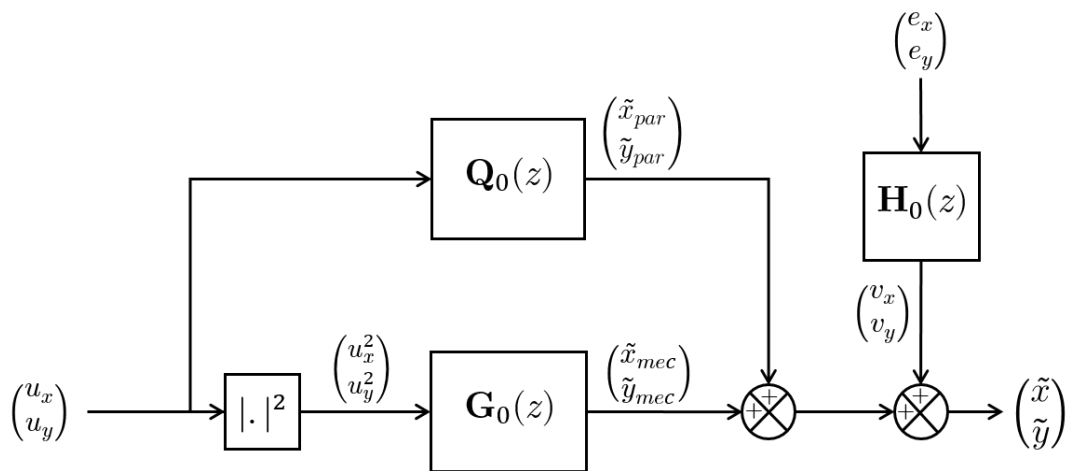


Figure 6.8: Block scheme of the MEMS gyroscope dynamics described by the three transfer function matrices $\mathbf{G}_0(z)$, $\mathbf{Q}_0(z)$ and $\mathbf{H}_0(z)$.

⁴See [Sau08] for the technical details on how to limit, e.g., the effects of the Spring softening.

6.5 Controller and principle of angular rate deduction

6.5.1 Closed-loop scheme for angular rate deduction

For the angular rate deduction, the sensor is often put in closed-loop with a controller as depicted in Figure 6.9.

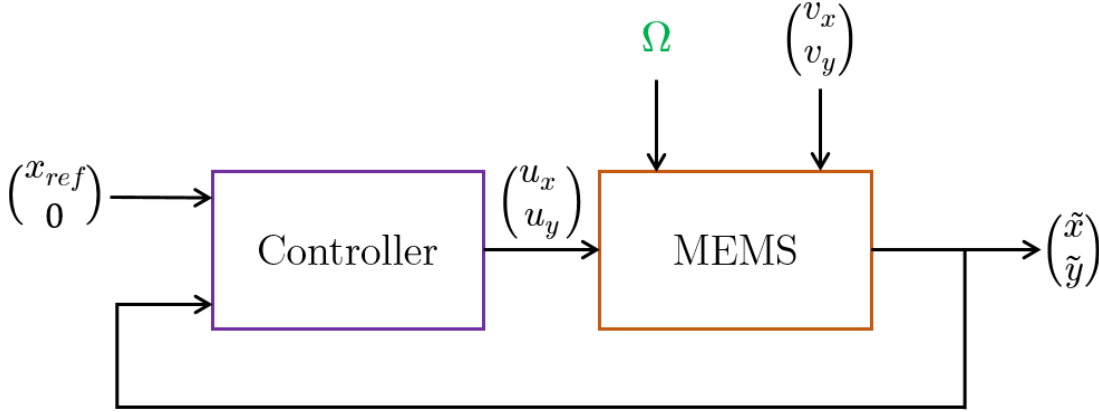


Figure 6.9: General scheme of the MEMS gyroscope in closed-loop with a controller.

The goal of the controller is to get the output vector $(\tilde{x}_{mec}, \tilde{y}_{mec})^T$ to track a reference vector equal to $(x_{ref}, 0)^T$ where x_{ref} is a sinusoidal signal at a given frequency ω_{ref} . Since $\tilde{x}_{mec} = \beta_x x_{mec}$ and $\tilde{y}_{mec} = \beta_y y_{mec}$, then we will have $x_{mec} = x_{ref}/\beta_x$ and $y_{mec} = 0$. We will see in the next paragraph how these objectives can allow us to obtain the angular rate Ω .

In order to guarantee these objectives, the controller architecture that has been chosen by Fabrício SAGGIN is given in Figure 6.10. This controller has two degrees of freedom which are the two blocks "Tracking Controller $\mathbf{K}(z)$ " and "Parasitic Compensation $\mathbf{K}_{comp}(z)$ ". Both blocks will be independently synthesized.

The purpose of the block "Parasitic Compensation $\mathbf{K}_{comp}(z)$ " is to suppress the parasitic signals \tilde{x}_{par} and \tilde{y}_{par} from the measured outputs \tilde{x} and \tilde{y} . The compensated signals are denoted \tilde{x}_c for the drive output and \tilde{y}_c for the sense output. By denoting $\mathbf{K}_{comp}(z)$ the 2×2 transfer function matrix used for this compensation, we have that

$$\begin{pmatrix} \tilde{x}_c(t) \\ \tilde{y}_c(t) \end{pmatrix} = \begin{pmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{pmatrix} + \mathbf{K}_{comp}(z) \begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix} \quad (6.9)$$

From (6.7), the optimal expression for $\mathbf{K}_{comp}(z)$ in order to guarantee $\tilde{x}_c(t) = \tilde{x}_{mec} + v_x$ and $\tilde{y}_c(t) = \tilde{y}_{mec} + v_y$ is $\mathbf{K}_{comp}(z) = -\mathbf{Q}_0(z)$. Therefore, if we have an accurate estimate $\hat{\mathbf{Q}}(z)$ of $\mathbf{Q}_0(z)$, the parasite capacitive coupling will be well compensated.

The tracking controller $\mathbf{K}(z)$ must ensure an appropriate rejection of the noises v_x and v_y and the tracking of the reference $(x_{ref}, 0)^T$. With the tracking controller, the control efforts u_x and u_y are given by

$$\begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix} = \mathbf{K}(z) \begin{pmatrix} x_{ref}(t) - \tilde{x}_c(t) \\ -\tilde{y}_c(t) \end{pmatrix}$$

In order to increase the amplitude of the vibration of the drive mass m_x with a minimal control effort u_x , the frequency ω_{ref} of the sinusoidal reference signal x_{ref} should be equal to the resonance frequency $\omega_{0,x}$.

If $\mathbf{K}(z)$ and $\mathbf{K}_{comp}(z)$ attain their respective objectives, we will have the following property: $\tilde{x}_{mec} \approx x_{ref}$ and $\tilde{y}_{mec} \approx 0$. In the next paragraph, we explain how to obtain the angular rate Ω when the MEMS gyroscope is operated as described above.

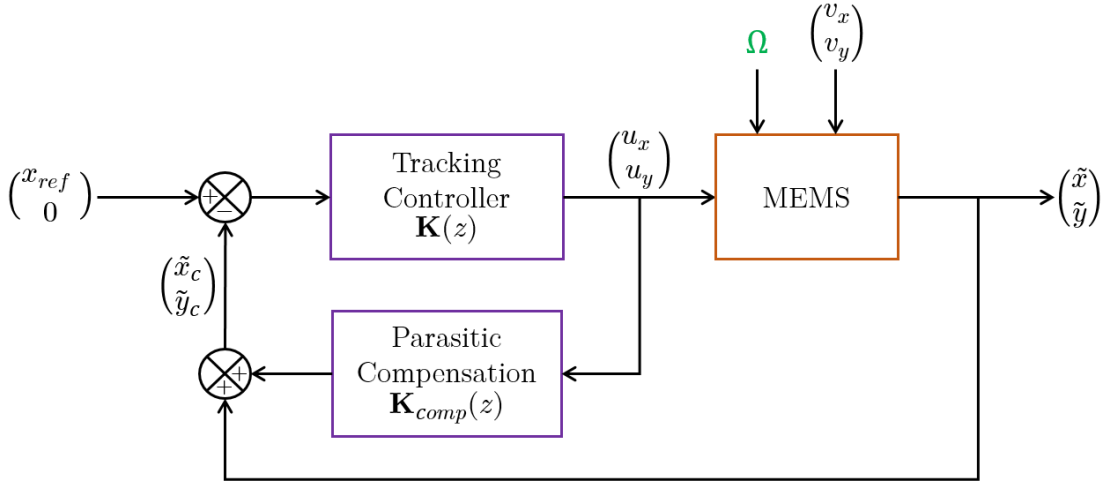


Figure 6.10: A detailed scheme block of the two-degree-of-freedom controller.

6.5.2 Angular rate deduction

We will suppose that the objectives of $\mathbf{K}(z)$ and $\mathbf{K}_{comp}(z)$ are attained and thus that $\tilde{x}_{mec} = x_{ref}$ and $\tilde{y}_{mec} = 0$. Since the mechanical dynamics are the only remaining dynamics, let us consider equation (6.1) and (6.2). In these equations, x_{mec} will have a sinusoidal shape which is proportional to x_{ref} and y_{mec} will be equal to zero. Consequently, we must have that the force f_y applied by the actuator is given by:

$$f_y = 2m_x\Omega\dot{x}_{mec} + d_{yx}\dot{x}_{mec} + k_{yx}x_{mec} \quad (6.10)$$

Since $f_y = \alpha_y u_y^2$ and $\tilde{x}_{mec} = \beta_x x_{mec} = x_{ref}$, (6.10) becomes

$$u_y^2 = \underbrace{(a\Omega + b)}_{\hat{\Omega}} \dot{x}_{ref} + \frac{k_{yx}}{\alpha_y \beta_x} x_{ref}$$

where $a = 2m_x/(\alpha_y\beta_x)$ and $b = d_{yx}/(\alpha_y\beta_x)$.

Due to the sinusoidal shape of x_{ref} , if we demodulate⁵ u_y^2 with \dot{x}_{ref} , we obtain $\hat{\Omega}$. Then, if we know precisely the values of d_{yx} , m_x , α_y and β_x , we can obtain Ω . We will see later in Chapter 7 that we can estimate them up to a multiplicative constant. Note that it is also possible to estimate directly both constants a and b by doing a calibration of the sensor with different values of known angular rate Ω .

It is to be noted that the principle of the angular rate deduction in the MEMS gyroscope is similar to the principle of radio-frequency (RF) transmission. Indeed, in RF transmission, a signal $q(t_c)$ that we want to transmit is first modulated by a sinusoidal carrier signal $q_{car}(t_c)$ and then sent through a channel. Then, this modulated signal $q(t_c)q_{car}(t_c)$ is received by an antenna and demodulated to retrieve the initial signal $q(t_c)$. For the MEMS gyroscope, we have that $q(t_c) = \hat{\Omega}(t_c)$ and $q_{car}(t_c) = \dot{x}_{ref}(t_c)$. The transmission is due to the Coriolis effect and the antenna is the sense mass. The Figure 6.11 illustrates this analogy.

6.5.3 Modeling approach considered for the control of the MEMS gyroscope

For the modeling, we will consider that $\Omega = 0$. To design optimally both regulator blocks "Parasitic Controller $\mathbf{K}_{comp}(z)$ " and "Tracking Controller $\mathbf{K}(z)$ " in Figure 6.10 in order to guarantee $\tilde{x}_{mec} = x_{ref}$ and $\tilde{y}_{mec} = 0$, we need to model/estimate accurately the following phenomena: (i) the mechanical coupling, (ii) the parasitic electrical coupling and (iii) the PSDs of the noises v_x and v_y within the structure given in (6.7).

In Chapter 1, we have explained that there are three categories of modeling methods: white-box, grey-box and black-box modeling (see Figure 1.1). We cannot perform a white-box modeling since some of the mechanical parameters (e.g., d_{xy} and d_{yx}) are unknown. Therefore, we should perform a system identification but we have to decide between a grey-box and a black-box approach. In this chapter, we have mainly presented the dynamics linked to the gyroscope and the instrumentation by assuming that the dynamics of the detection circuit (implemented on the electronic card) can be modeled as a constant. However, this might not be the case and we should be able to model these extra dynamics for which we do not have an idea on how to model them with physics principles. Therefore, we should consider a black-box approach⁶ in order to estimate all the transfer functions in $\mathbf{G}_0(z)$, $\mathbf{Q}_0(z)$ and in $\mathbf{H}_0(z)$. For the identification

⁵In [Sau08], the demodulation technique which is used is the synchronous demodulation. It is the same technique considered in the MEMS of this study. See [Sau08] for more details on the synchronous demodulation.

⁶By black-box approach, we mean that we will determine the complexity of the transfer functions in $\mathbf{G}(z)$, $\mathbf{Q}(z)$ and $\mathbf{H}(z)$ such that they satisfy the full-order assumption, i.e., such that all the dynamics observed in the data are modeled. But the modeling approach is not entirely black-box since the input of $\mathbf{G}(z)$ and $\mathbf{Q}(z)$ has been determined based on first principles.

of these transfer functions, we are going to include some priors from the physics laws and from the documentation given by the MEMS gyroscope manufacturer:

- the dynamics of $G_{0,xx}$ and $G_{0,yy}$ are characterized by a resonance at the resonance frequency $\omega_{0,x}$ for $G_{0,xx}$ and $\omega_{0,y}$ for $G_{0,yy}$. The cross-transfer functions $G_{0,xy}$ and $G_{0,yx}$ have two resonances each, at both the resonance frequencies $\omega_{0,x}$ and $\omega_{0,y}$.
- at 20°C, $\omega_{0,x} \approx 2\pi \times 11836.8$ rad/s and $\omega_{0,y} \approx 2\pi \times 11843.4$ rad/s.
- at 20°C, the quality factor of the drive resonance has an order of magnitude of 10^5 and the one of the sense resonance has an order of magnitude of 10^4 .
- the transfer functions in \mathbf{Q}_0 are highpass filters.

6.6 Conclusion of Chapter 6

In this chapter, we have given a brief description of the MEMS gyroscope and a picture of the control challenges that exist regarding this sensor in order to increase its measurement quality of angular rates. In order to derive an optimal controller, it is important to have a model of the main dynamics of the MEMS gyroscope. There are three main dynamics in the MEMS gyroscope: mechanical coupling, parasite capacitive coupling and noise measurement but the literature does not propose a complete model structure to model them. Therefore, we have developed a more complete structure that can represent these three dynamics through the matrices $\mathbf{G}_0(z)$, $\mathbf{Q}_0(z)$ and $\mathbf{H}_0(z)$ (see Figure 6.8), answering the first part of Research Question 4.

We have then presented the chosen control strategy for the angular rate deduction. It consists in the design of two blocks (degrees of freedom): one block is dedicated to the parasite capacitive coupling compensation and the other block is used for the tracking objectives for the angular rate deduction and the noise measurement reduction. In the next chapter, we will focus on developing an identification approach for the modeling of the transfer functions in $\mathbf{G}_0(z)$, $\mathbf{Q}_0(z)$ and $\mathbf{H}_0(z)$.

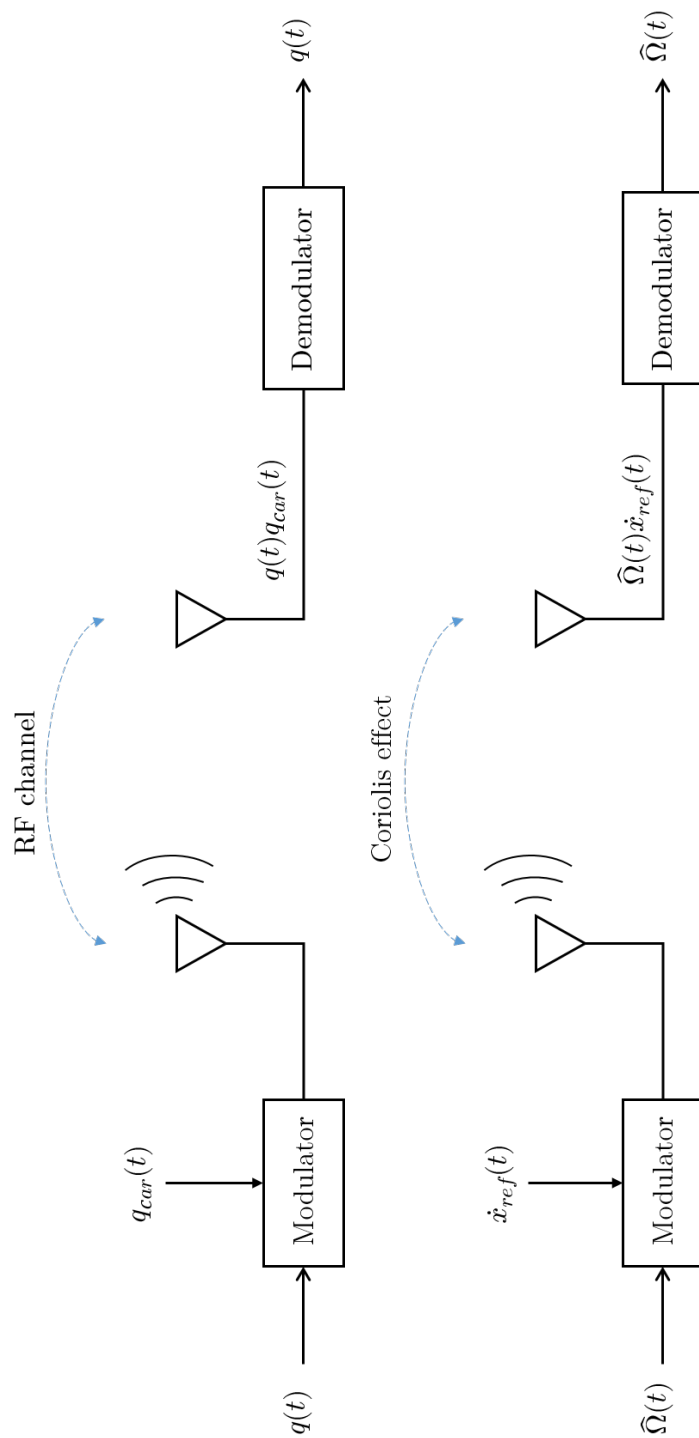


Figure 6.11: Comparison between the RF transmission and the Coriolis effect in the MEMS gyroscope.

Identification for the control of the MEMS gyroscope

Contents

7.1	Introduction of Chapter 7	118
7.2	Identification procedure and model structure	119
7.2.1	Simplification	119
7.2.2	Model structures	120
7.2.3	Data sets	121
7.2.4	Identified parameter vectors	121
7.3	Determination of an initial model	122
7.3.1	Limitation on the experiment design	122
7.3.2	SISO-by-SISO approach	122
7.3.3	Results of the initial identifications and interpretation	128
7.4	MISO identification and data informativity	133
7.4.1	Input choice for the MISO identification	134
7.4.2	Study of the data informativity	135
7.4.3	Identification results and comparison with initial model	136
7.5	Block "Parasitic Compensation $\mathbf{K}_{comp}(z)$ "	137
7.5.1	Strategy: model-based compensation	137
7.5.2	Results of the model-based compensation	137
7.6	Block "Tracking Controller $\mathbf{K}(z)$ "	142
7.6.1	Issue for the design: high complexity of $\mathbf{G}(z, \hat{\boldsymbol{\theta}}_N)$	142
7.6.2	Grey-box-based model reduction	143
7.6.3	Controller design and validation	148
7.7	Conclusion	150

7.1 Introduction of Chapter 7

In this chapter, we focus on addressing the second part of Research Question 4 formulated in Section 1.4. We will propose an identification method based on Prediction Error for the three main dynamics of the MEMS gyroscope: the mechanical coupling, the parasitic capacitive coupling and the noise measurement. More precisely, we want to derive an accurate estimate of the entries of the matrices $\mathbf{G}_0(z)$, $\mathbf{Q}_0(z)$ and $\mathbf{H}_0(z)$ involved in the complete model structure developed in Chapter 6 (see (6.7)).

As explained in Chapter 6, the MEMS has two outputs \tilde{x} and \tilde{y} and two inputs u_x and u_y . Therefore, one option should be to identify this system by exciting both inputs and measure both outputs at the same time. However, by doing some first experiments, we have observed that some dynamics might hide the other ones when both inputs are excited at the same time. For instance both parasite capacitive transfer functions $Q_{0,xx}$ and $Q_{0,xy}$ can completely hide the mechanical ones $G_{0,xx}$ and $G_{0,xy}$ when u_x and u_y are excited simultaneously.

Therefore, in Section 7.2, we will propose a simplification of the identification to circumvent this effect. For this purpose, a MIMO system with two inputs and two outputs can be seen as the collection of four subsystems with one input and one output, as explained in Section 1.5. By only exciting one input, we can identify two subsystems whose input is the one that is excited. Then, the converse is done by exciting the other input and by identifying the two other subsystems. In our case, we first excite the input u_x and set u_y to 0 and we identify the two subsystems whose input is u_x , i.e., we will identify the transfer functions $G_{0,xx}$, $G_{0,yx}$, $Q_{0,xx}$, $Q_{0,yx}$ and H_x . Then, we set u_x to 0, we excite u_y and we identify the remaining subsystems. These four subsystems are of the form given in (5.1) with $n = 2$ and so we can reformulate these four subsystems as MISO systems as we did in Chapter 5. We then must design u_x and u_y such that the MISO identifications are accurate.

One property that must be guaranteed for this purpose is the data informativity. Therefore, we can think at first glance that we can choose any excitation and verify that the data will be informative for the four MISO identifications in order to yield accurate estimates. However, this is true *if we can increase the number of data*. For some (unlucky) input choices, this solution will not be suitable in practice since we would need to do very long experiments and the computational cost of the identification algorithm would increase a lot. Therefore, we should design properly the inputs u_x and u_y by increasing the signal-to-noise ratio (SNR) in order to get small variances on the computed parameter vectors for a reasonable number of data (provided that the data informativity is ensured with these input designs for the four MISO identifications). In Section 7.4.1, we explain how to do this design by using the priors given in Section 6.5.3. A data informativity analysis will be done to verify that this design is appropriate.

However, even if the input is well chosen to guarantee small variances, the algorithms used for these MISO identifications are complex nonlinear optimizations and some local

minimum problems can be met when computing the estimates. Therefore, in Section 7.3, a method will be provided to determine initial estimates for all the transfer functions in (6.7) by isolating the dynamics of each transfer function in (6.7) and performing SISO identifications. After that, a choice is proposed for the input u_x (and then for u_y) in order to accurately identify the transfer functions $G_{0,xx}$, $G_{0,yx}$, $Q_{0,xx}$, $Q_{0,yx}$ and H_x .

After that, the performances of the controller that is synthesized by Fabrício SAGGIN based on the identified model will be assessed. Finally, we give a conclusion in Section 7.7.

7.2 Identification procedure and model structure

7.2.1 Simplification

As explained in Chapter 6, the main dynamics of the MEMS gyroscope with $\Omega = 0$ (illustrated in Figure 6.8) can be modeled as follows

$$\begin{aligned} \begin{pmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{pmatrix} &= \begin{pmatrix} Q_{0,xx}(z) & Q_{0,xy}(z) \\ Q_{0,yx}(z) & Q_{0,yy}(z) \end{pmatrix} \begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix} + \begin{pmatrix} G_{0,xx}(z) & G_{0,xy}(z) \\ G_{0,yx}(z) & G_{0,yy}(z) \end{pmatrix} \begin{pmatrix} u_x^2(t) \\ u_y^2(t) \end{pmatrix} + \mathbf{v}(t) \\ \mathbf{v}(t) &= \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} = \begin{pmatrix} H_{0,x}(z) & 0 \\ 0 & H_{0,y}(z) \end{pmatrix} \begin{pmatrix} e_x(t) \\ e_y(t) \end{pmatrix} \end{aligned} \quad (7.1)$$

We have, in this case, a MIMO parallel Hammerstein system with a square (static) nonlinearity, with two inputs u_x and u_y and two outputs \tilde{x} and \tilde{y} . However, as will be seen later, some transfer functions such as $G_{0,xx}(z)$ and $G_{0,yy}(z)$ can be dominant with respect to the other transfer functions $G_{0,xy}(z)$ and $G_{0,yx}(z)$ and so inaccurate models can be obtained for $G_{0,xy}(z)$ and $G_{0,yx}(z)$ if we perform a MIMO identification by exciting the MEMS gyroscope with nonzero inputs u_x and u_y at the same time. Therefore, we will simplify the identification as follows:

- we only excite u_x and set u_y to 0. Then, (7.1) becomes

$$\begin{pmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{pmatrix} = \begin{pmatrix} Q_{0,xx}(z) \\ Q_{0,yx}(z) \end{pmatrix} u_x(t) + \begin{pmatrix} G_{0,xx}(z) \\ G_{0,yx}(z) \end{pmatrix} u_x^2(t) + \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} \quad (7.2)$$

- we only excite u_y and set u_x to 0. Then, (7.1) becomes

$$\begin{pmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{pmatrix} = \begin{pmatrix} Q_{0,xy}(z) \\ Q_{0,yy}(z) \end{pmatrix} u_y(t) + \begin{pmatrix} G_{0,xy}(z) \\ G_{0,yy}(z) \end{pmatrix} u_y^2(t) + \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} \quad (7.3)$$

Let us then observe that (7.2) is made up of two particular parallel Hammerstein systems \mathcal{S}_{xx} and \mathcal{S}_{yx} given by

$$\mathcal{S}_{xx} : \quad \tilde{x}(t) = Q_{0,xx}(z)u_x(t) + G_{0,xx}(z)u_x^2(t) + H_{0,x}(z)e_x(t) \quad (7.4)$$

$$\mathcal{S}_{yx} : \quad \tilde{y}(t) = Q_{0,yx}(z)u_x(t) + G_{0,yx}(z)u_x^2(t) + H_{0,y}(z)e_y(t) \quad (7.5)$$

Therefore, when we only consider the input u_x , \mathcal{S}_{xx} and \mathcal{S}_{yx} have to be identified.

The same holds for (7.3) with the following Hammerstein systems \mathcal{S}_{xy} and \mathcal{S}_{yy} given by

$$\begin{aligned}\mathcal{S}_{xy} : \quad & \tilde{x}(t) = Q_{0,xy}(z)u_y(t) + G_{0,xy}(z)u_y^2(t) + H_{0,x}(z)e_x(t) \\ \mathcal{S}_{yy} : \quad & \tilde{y}(t) = Q_{0,yy}(z)u_y(t) + G_{0,yy}(z)u_y^2(t) + H_{0,y}(z)e_y(t)\end{aligned}\quad (7.6)$$

The four systems \mathcal{S}_{xx} , \mathcal{S}_{xy} , \mathcal{S}_{yx} and \mathcal{S}_{yy} correspond to the four subsystems aforementioned in Sections 1.5 and 7.1 and which constitute the main dynamics of the MEMS gyroscope. Let us observe that these four subsystems have the particular parallel Hammerstein structure studied in Chapter 5 with $n = 2$ (see (5.1)).

As done in Chapter 5, we can rewrite (7.4) and (7.5) successively as follows

$$\begin{aligned}\mathcal{S}_{xx} : \quad & \tilde{x}(t) = (Q_{0,xx}(z) \quad G_{0,xx}(z)) \mathbf{u}_x(t) + H_{0,x}(z)e_x(t) \\ \mathcal{S}_{yx} : \quad & \tilde{y}(t) = (Q_{0,yx}(z) \quad G_{0,yx}(z)) \mathbf{u}_x(t) + H_{0,y}(z)e_y(t)\end{aligned}$$

where $\mathbf{u}_x = (u_x, u_x^2)^T$. The same rewriting can be done for (7.5) and (7.6) with $\mathbf{u}_y = (u_y, u_y^2)^T$.

7.2.2 Model structures

In order to identify the transfer functions in \mathcal{S}_{xx} and \mathcal{S}_{yx} , we will consider two linear MISO BJ model structures

$$\begin{aligned}\mathcal{M}_{xx} &= \left\{ (Q_{xx}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}}), G_{xx}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}}), H_x(z, \boldsymbol{\eta}_x)) \mid \boldsymbol{\theta}_{xx} = (\tilde{\boldsymbol{\theta}}_{Q_{xx}}^T, \tilde{\boldsymbol{\theta}}_{G_{xx}}^T, \boldsymbol{\eta}_x^T)^T \in \mathcal{D}_{\boldsymbol{\theta}_{xx}} \right\} \\ \mathcal{M}_{yx} &= \left\{ (Q_{yx}(z, \tilde{\boldsymbol{\theta}}_{Q_{yx}}), G_{yx}(z, \tilde{\boldsymbol{\theta}}_{G_{yx}}), H_y(z, \boldsymbol{\eta}_y)) \mid \boldsymbol{\theta}_{yx} = (\tilde{\boldsymbol{\theta}}_{Q_{yx}}^T, \tilde{\boldsymbol{\theta}}_{G_{yx}}^T, \boldsymbol{\eta}_y^T)^T \in \mathcal{D}_{\boldsymbol{\theta}_{yx}} \right\}\end{aligned}$$

where $\mathcal{D}_{\boldsymbol{\theta}_{xx}}$ and $\mathcal{D}_{\boldsymbol{\theta}_{yx}}$ are defined similarly as in the previous chapters.

Since \mathcal{M}_{xx} is a BJ model structure, the transfer functions $Q_{xx}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}})$, $G_{xx}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}})$ and $H_x(z, \boldsymbol{\eta}_x)$ are given by

$$\begin{aligned}Q_{xx}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}}) &= z^{-\rho_{Q_{xx}}} \frac{\dot{B}_{Q_{xx}}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}})}{F_{Q_{xx}}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}})} \\ G_{xx}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}}) &= z^{-\rho_{G_{xx}}} \frac{\dot{B}_{G_{xx}}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}})}{F_{G_{xx}}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}})} \\ H_x(z, \boldsymbol{\eta}_x) &= \frac{C_x(z, \boldsymbol{\eta}_x)}{D_x(z, \boldsymbol{\eta}_x)}\end{aligned}$$

where $\mathring{B}_{Q_{xx}}$ and $\mathring{B}_{G_{xx}}$ are polynomials with a zero-delay and where $F_{Q_{xx}}$, $F_{G_{xx}}$, C_x and D_x are monic polynomials. The integers $\rho_{Q_{xx}}$ and $\rho_{G_{xx}}$ are respectively the delays of Q_{xx} and G_{xx} . We define similarly the polynomials $\mathring{B}_{Q_{yx}}$, $\mathring{B}_{G_{yx}}$, $F_{Q_{yx}}$, $F_{G_{yx}}$, C_y and D_y for the transfer functions Q_{yx} , G_{yx} and H_y in the model structure \mathcal{M}_{yx} .

We also define similarly the model structures \mathcal{M}_{xy} and \mathcal{M}_{yy} with all the notations introduced in this paragraph.

7.2.3 Data sets

By applying a quasi-stationary u_x to the MEMS gyroscope and by setting u_y to 0 from $t = 1$ till $t = N$, we collect simultaneously the following data sets

$$\begin{aligned}\mathcal{Z}_{xx}^N &= \{(\tilde{x}(t), \mathbf{u}_x^T(t))^T \mid t = 1, \dots, N\} \\ \mathcal{Z}_{yx}^N &= \{(\tilde{y}(t), \mathbf{u}_x^T(t))^T \mid t = 1, \dots, N\}\end{aligned}$$

We also define similarly the data sets \mathcal{Z}_{xy}^N and \mathcal{Z}_{yy}^N by applying a quasi-stationary u_y to the MEMS gyroscope and by setting u_x to 0 from $t = 1$ till $t = N$.

Based on the data sets \mathcal{Z}_{xx}^N and \mathcal{Z}_{yx}^N , we define, for each $(Q_{xx}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}}), G_{xx}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}}), H_x(z, \boldsymbol{\eta}_x)) \in \mathcal{M}_{xx}$ and $(Q_{yx}(z, \tilde{\boldsymbol{\theta}}_{Q_{yx}}), G_{yx}(z, \tilde{\boldsymbol{\theta}}_{G_{yx}}), H_y(z, \boldsymbol{\eta}_y)) \in \mathcal{M}_{yx}$, the one step-ahead predictors

$$\begin{aligned}\hat{x}(t, \boldsymbol{\theta}_{xx}) &= H_x^{-1}(z, \boldsymbol{\eta}_x)(Q_{xx}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}}), G_{xx}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}})\mathbf{u}_x(t) + (1 - H_x^{-1}(z, \boldsymbol{\eta}_x))\tilde{x}(t) \\ \hat{y}(t, \boldsymbol{\theta}_{yx}) &= H_y^{-1}(z, \boldsymbol{\eta}_y)(Q_{yx}(z, \tilde{\boldsymbol{\theta}}_{Q_{yx}}), G_{yx}(z, \tilde{\boldsymbol{\theta}}_{G_{yx}})\mathbf{u}_x(t) + (1 - H_y^{-1}(z, \boldsymbol{\eta}_y))\tilde{y}(t)\end{aligned}$$

7.2.4 Identified parameter vectors

The models are then obtained by computing the optimal parameter vectors $\hat{\boldsymbol{\theta}}_{N,xx}$ and $\hat{\boldsymbol{\theta}}_{N,yx}$ given by

$$\hat{\boldsymbol{\theta}}_{N,xx} = \arg \min_{\boldsymbol{\theta}_{xx} \in \mathcal{D}_{\boldsymbol{\theta}_{xx}}} \frac{1}{N} \sum_{t=1}^N (\tilde{x}(t) - \hat{x}(t, \boldsymbol{\theta}_{xx}))^2 \quad (7.7)$$

$$\hat{\boldsymbol{\theta}}_{N,yx} = \arg \min_{\boldsymbol{\theta}_{yx} \in \mathcal{D}_{\boldsymbol{\theta}_{yx}}} \frac{1}{N} \sum_{t=1}^N (\tilde{y}(t) - \hat{y}(t, \boldsymbol{\theta}_{yx}))^2 \quad (7.8)$$

We define similarly the optimal parameter vectors $\hat{\boldsymbol{\theta}}_{N,xy}$ and $\hat{\boldsymbol{\theta}}_{N,yy}$ based on \mathcal{Z}_{xy}^N and \mathcal{Z}_{yy}^N and on the model structures \mathcal{M}_{xy} and \mathcal{M}_{yy} . In order to guarantee an accurate model of the MEMS gyroscope, we wonder how to design the excitation u_x to guarantee accurate estimates $\hat{\boldsymbol{\theta}}_{N,xx}$ and $\hat{\boldsymbol{\theta}}_{N,yx}$. The same problem holds for \mathcal{M}_{xy} and \mathcal{M}_{yy} with the design of u_y .

Since $\hat{\boldsymbol{\theta}}_{N,xx}$, $\hat{\boldsymbol{\theta}}_{N,yx}$, $\hat{\boldsymbol{\theta}}_{N,xy}$ and $\hat{\boldsymbol{\theta}}_{N,yy}$ are solutions of complex non-linear optimization problems, we will first find relatively accurate models of each unknown transfer functions

in \mathcal{S}_{xx} , \mathcal{S}_{yx} , \mathcal{S}_{xy} and \mathcal{S}_{yy} using a SISO-by-SISO approach. In this approach, we will make use of the fact that the input non-linearity is known.

7.3 Determination of an initial model

Before presenting this SISO-by-SISO approach, we should take into account the input limitations of the electronic instrumentation. They are presented in the next paragraph.

7.3.1 Limitation on the experiment design

For the MEMS gyroscope considered in this study, the input voltage u_x must be non-negative, cannot exceed 2V and must have a constant¹ DC voltage u_{DC} equal to 1V, i.e.,

$$u_x(t) = u_{DC} + \bar{u}_x(t) \quad (7.9)$$

where $-1V \leq \bar{u}_x(t) \leq 1V$. The same holds for u_y . Therefore, we should focus on the signals \bar{u}_x and \bar{u}_y for the experiment design choice.

As mentioned in the introduction of this chapter (Section 7.1), in order to obtain identified models with a small variance, it is important to maximize the signal-to-noise ratio (SNR). Therefore, we should design \bar{u}_x such that it has a maximal power and it guarantees $-1V \leq \bar{u}_x(t) \leq 1V$. However, maximizing the input power w.r.t. the input limitations is not enough for the input signal to be appropriate for the identification. Indeed, the choice of \bar{u}_x must also optimize the input power spectrum of u_x and u_x^2 for both outputs \tilde{x} and \tilde{y} . Recalling Section 6.5.3, the transfer function $G_{0,yx}$ should be excited at the resonances frequencies $\omega_{0,x}$ and $\omega_{0,y}$ while $G_{0,xx}$ should be excited at $\omega_{0,x}$. Therefore, to perform an identification of $G_{0,xx}$ and $G_{0,yx}$ using an unique experiment, the power spectrum of the input of both transfer functions (which is u_x^2) should be non zero at frequencies near $\omega_{0,x}$ and $\omega_{0,y}$. For the transfer functions $Q_{0,xx}$ and $Q_{0,yx}$, the dynamics are distributed on a wider frequency range than $G_{0,xx}$ and $G_{0,yx}$. Therefore, the input of both transfer functions $Q_{0,xx}$ and $Q_{0,yx}$ (which is u_x) should excite them at a wide frequency range, preferably in the high frequencies since both transfer functions are highpass filters. We will have to verify if this input choice yields the data informativity for the four MISO identifications.

Remark 7.1. *The reasoning given in the previous paragraph can be repeated for the input u_y and the transfer functions $G_{0,xy}$, $G_{0,yy}$, $Q_{0,xy}$ and $Q_{0,yy}$.*

7.3.2 SISO-by-SISO approach

In this section, we present the SISO-by-SISO approach to determine initial estimates of all the unknown transfer functions in the four subsystems \mathcal{S}_{xx} , \mathcal{S}_{yx} , \mathcal{S}_{xy} and \mathcal{S}_{yy} . The general idea of the considered method is to isolate and identify the dynamics of each

¹The DC voltage needs to be kept constant to avoid the Spring Softening effect which makes both resonances to vary when this constant is changed [Sau08].

transfer function separately with SISO identifications (hence the name SISO-by-SISO approach). First, we can separate the noise dynamics from the mechanical coupling and parasitic capacitive coupling as explained in the next paragraph.

In the next paragraphs, we will present the SISO-by-SISO approach for the estimation of all the transfer functions in \mathcal{S}_{xx} and \mathcal{S}_{yx} . The approach is similar for \mathcal{S}_{xy} and \mathcal{S}_{yy} . Therefore, the input u_y will be set to 0 in the presentation of the method.

Initial estimation of $H_{0,x}$ and $H_{0,y}$

To estimate $H_{0,x}$ and $H_{0,y}$, we also set u_x to 0 (i.e., the MEMS gyroscope is not excited). In this case, (7.4) and (7.5) become successively

$$\begin{aligned}\mathcal{S}_{xx} : \quad \tilde{x}(t) &= H_{0,x}(z)e_x(t) \\ \mathcal{S}_{yx} : \quad \tilde{y}(t) &= H_{0,y}(z)e_x(t)\end{aligned}$$

Let us consider the following model structures

$$\begin{aligned}\mathcal{M}_{H_x} &= \{H_x(z, \boldsymbol{\eta}_x) \mid \boldsymbol{\eta}_x \in \mathcal{D}_{\eta_x}\} \\ \mathcal{M}_{H_y} &= \{H_y(z, \boldsymbol{\eta}_y) \mid \boldsymbol{\eta}_y \in \mathcal{D}_{\eta_y}\}\end{aligned}$$

where $H_x(z, \boldsymbol{\eta}_x)$ and $H_y(z, \boldsymbol{\eta}_y)$ are defined in the model structures \mathcal{M}_{H_x} and \mathcal{M}_{H_y} in Section 7.2.2.

We collect $\tilde{x}(t)$ and $\tilde{y}(t)$ for $t = 1$ till $t = N = 10000$ for $u_x = u_y = 0$ and we use the following criteria to identify a model in \mathcal{M}_{H_x} and \mathcal{M}_{H_y} :

$$\begin{aligned}\hat{\boldsymbol{\eta}}_{x,init} &= \arg \min_{\boldsymbol{\eta}_x \in \mathcal{D}_{\eta_x}} \frac{1}{N} \sum_{t=1}^N \epsilon_{H_x}(t, \boldsymbol{\eta}_x)^2 \\ \hat{\boldsymbol{\eta}}_{y,init} &= \arg \min_{\boldsymbol{\eta}_y \in \mathcal{D}_{\eta_y}} \frac{1}{N} \sum_{t=1}^N \epsilon_{H_y}(t, \boldsymbol{\eta}_y)^2\end{aligned}$$

where $\epsilon_{H_x}(t, \boldsymbol{\eta}_x) = H^{-1}(z, \boldsymbol{\eta}_x)\tilde{x}(t)$ and $\epsilon_{H_y}(t, \boldsymbol{\eta}_y) = H^{-1}(z, \boldsymbol{\eta}_y)\tilde{y}(t)$. We will see later in Section 7.3.3 how the degrees of the polynomials C_{H_x} and D_{H_x} in H_x and the polynomials C_{H_y} and D_{H_y} in H_y are chosen.

Initial estimation of $G_{0,xx}$ and $G_{0,yx}$

While the dynamics of the noise are quite easy to isolate for the identification of H_x and H_y , the same cannot be said for the transfer functions $Q_{0,xx}$, $G_{0,xx}$, $Q_{0,yx}$ and $G_{0,yx}$. Indeed, both inputs u_x and u_x^2 of the aforementioned transfer functions are linked by the square nonlinearity. In this paragraph and the next one, we propose an experiment design that allows the separation of both dynamics.

Let us first start with the mechanical transfer functions $G_{0,xx}$ and $G_{0,yx}$ whose input is u_x^2 . Recall that, when u_y is set to 0, the dynamics of the MEMS gyroscope are given by (7.4) and (7.5). The main idea is to separate the effect of $Q_{0,xx}$ from $G_{0,xx}$ and the effect of $Q_{0,yx}$ from $G_{0,yx}$ in the frequency domain by designing a quasi-stationary input u_x such that it exists a bandpass filter $T(z)$ guaranteeing $T(z)u_x \approx 0$ and $T(z)u_x^2 \neq 0$. Consequently, with this particular excitation u_x applied to the MEMS (with $u_y = 0$) and by filtering the outputs \tilde{x} and \tilde{y} with $T(z)$, (7.4) and (7.5) successively become

$$T(z)\tilde{x}(t) = G_{0,xx}(z)T(z)u_x^2(t) + T(z)H_{0,x}(z)e_x(t) \quad (7.10)$$

$$T(z)\tilde{y}(t) = G_{0,yx}(z)T(z)u_x^2(t) + T(z)H_{0,y}(z)e_x(t) \quad (7.11)$$

We should now decide which model structure to use for the identification of both transfer functions $G_{0,xx}(z)$ and $G_{0,yx}(z)$. If we consider a SISO BJ model structure for, e.g., (7.4) then we will identify $T(z)H_{0,x}(z)$ for the noise transfer function. The same holds for (7.4) and $T(z)H_{0,y}(z)$. But, since some initial estimates have been obtained for $H_{0,x}(z)$ and $H_{0,y}(z)$, the identification of the noise transfer functions $T(z)H_{0,x}(z)$ and $T(z)H_{0,y}(z)$ are not necessary to be performed. From the results of [Lju99], we know that we can get consistent estimates of $G_{0,xx}(z)$ and $G_{0,yx}(z)$ by identifying these transfer functions with OE model structures. Therefore, we do not have to identify the noise transfer functions which will also reduce the complexity of the identification algorithm. Therefore, we will consider the following SISO OE model structures for the identification of $G_{0,xx}(z)$ and $G_{0,yx}(z)$

$$\begin{aligned} \mathcal{M}_{G_{xx}} &= \{G_{xx}(z, \tilde{\theta}_{G_{xx}}) \mid \tilde{\theta}_{G_{xx}} \in \mathcal{D}_{\theta_{G_{xx}}}\} \\ \mathcal{M}_{G_{yx}} &= \{G_{yx}(z, \tilde{\theta}_{G_{yx}}) \mid \tilde{\theta}_{G_{yx}} \in \mathcal{D}_{\theta_{G_{yx}}}\} \end{aligned}$$

where $G_{xx}(z, \tilde{\theta}_{G_{xx}})$ and $G_{yx}(z, \tilde{\theta}_{G_{yx}})$ are defined in the model structures \mathcal{M}_{xx} and \mathcal{M}_{yx} in Section 7.2.2.

In this case, the data sets that will be considered for the identification of $G_{0,xx}(z)$ and $G_{0,yx}(z)$ are the data collected by applying this particular excitation u_x to the MEMS gyroscope (with $u_y = 0$) and by filtering the outputs $\tilde{x}(t)$ and $\tilde{y}(t)$ and the input u_x^2 with $T(z)$, i.e., the following data sets

$$\begin{aligned} \mathcal{Z}_{G_{xx}}^N &= \{(T(z)\tilde{x}(t), T(z)u_x^2(t))^T \mid t = 1, \dots, N\} \\ \mathcal{Z}_{G_{yx}}^N &= \{(T(z)\tilde{y}(t), T(z)u_x^2(t))^T \mid t = 1, \dots, N\} \end{aligned}$$

Based on these data sets, initial estimates $\hat{\theta}_{G_{xx},init}$ and $\hat{\theta}_{G_{yx},init}$ can be computed as

defined below

$$\hat{\boldsymbol{\theta}}_{G_{xx},init} = \arg \min_{\tilde{\boldsymbol{\theta}}_{G_{xx}} \in \mathcal{D}_{\theta_{G_{xx}}}} \frac{1}{N} \sum_{t=1}^N \epsilon_{G_{xx}}(t, \tilde{\boldsymbol{\theta}}_{G_{xx}})^2$$

$$\hat{\boldsymbol{\theta}}_{G_{yx},init} = \arg \min_{\tilde{\boldsymbol{\theta}}_{G_{yx}} \in \mathcal{D}_{\theta_{G_{yx}}}} \frac{1}{N} \sum_{t=1}^N \epsilon_{G_{yx}}(t, \tilde{\boldsymbol{\theta}}_{G_{yx}})^2$$

where

$$\epsilon_{G_{xx}}(t, \tilde{\boldsymbol{\theta}}_{G_{xx}}) = T(z)\tilde{x}(t) - G_{yx}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}})T(z)u_x^2(t)$$

$$\epsilon_{G_{yx}}(t, \tilde{\boldsymbol{\theta}}_{G_{yx}}) = T(z)\tilde{y}(t) - G_{yx}(z, \tilde{\boldsymbol{\theta}}_{G_{yx}})T(z)u_x^2(t)$$

In order to perform this identification, we have to determine (i) the bandpass filter $T(z)$ and (ii) this particular external excitation u_x .

For the filter $T(z)$, recall that the main dynamics of the transfer function $G_{0,xx}$ are mainly focused around $\omega_{0,x}$ and that the ones of transfer function $G_{0,yx}$ are focused around $\omega_{0,x}$ and $\omega_{0,y}$ (priors given in Section 6.5.3). Therefore, the frequency band \mathcal{I} of $T(z)$ should include the resonance frequencies $\omega_{0,x} \approx 2\pi \times 11836.8$ rad/s and $\omega_{0,y} \approx 2\pi \times 11843.4$ rad/s. We will consider the following interval $\mathcal{I} = 2\pi[11750, 11950]$ rad/s.

Now, we should design u_x such that the PSD of u_x^2 should be non zero at some frequencies belonging to $2\pi[-11950, -11750] \cup 2\pi[11750, 11950]$ rad/s while the one of u_x should be zero for all frequencies in $2\pi[-11950, -11750] \cup 2\pi[11750, 11950]$ rad/s. For this purpose, let us choose the signal u_x as in (7.9) with \bar{u}_x a multisine whose frequencies are located around $\omega_{0,x}/2$ and $\omega_{0,y}/2$. More precisely, we have chosen some excitation frequencies belonging to this interval $2\pi[5875, 5975]$ rad/s, which corresponds to the half of the frequencies belonging to \mathcal{I} . If \bar{u}_x is chosen in this way, the signal exciting the mechanical coupling will be:

$$u_x^2(t) = (1 + \bar{u}_x(t))^2 = 1 + 2\bar{u}_x(t) + \bar{u}_x^2(t) \quad (7.12)$$

Consequently, we observe that the mechanical coupling will be excited at many frequencies among which the frequencies in the interval \mathcal{I} (but also the frequency 0 and the frequencies in the interval $2\pi[5875, 5975]$ rad/s, etc). As opposed to this, the parasite electrical coupling will not be excited at the frequencies in the interval \mathcal{I} (it will only be excited at the frequency 0 and the frequencies in the interval $2\pi[5875, 5975]$ rad/s).

Then, if the considered filter $T(z)$ strongly filters out any frequency that does not belong to \mathcal{I} (and more especially the ones belonging to $2\pi[5875, 5975]$), then both (7.10) and (7.11) hold. For this purpose, $T(z)$ will be chosen as a Butterworth filter of order 6.

Let us give the details of the excitation and of the identification. We choose the following set of excitation frequencies belonging to $2\pi[5875, 5975]$ for u_x , all regularly distributed as follows

- Set 1 (around $\omega_{0,x}/2$): from $2\pi \times 5918.415$ till $2\pi \times 5918.49$ with a frequency bin of $2\pi \times 0.0005$. In total, there are 151 frequencies.
- Set 2 (between $\omega_{0,x}/2$ and $\omega_{0,y}/2$): from $2\pi \times 5918.5$ till $2\pi \times 5921$ with a frequency bin of $2\pi \times 0.5$. In total, there are 6 frequencies.
- Set 3 (around $\omega_{0,y}/2$): from $2\pi \times 5921.5$ till $2\pi \times 5922.2$ with a frequency bin of $2\pi \times 0.005$. In total, there are 141 frequencies.

The Sets 1 and 3 are denser in frequency content than the Set 2 in order to describe well the resonances at $\omega_{0,x}$ and $\omega_{0,y}$. Moreover, with this large number of frequencies, we are sure that the data informativity will be guaranteed for the SISO identifications for high degrees of the polynomials $B_{G_{xx}}$, $F_{G_{xx}}$, $B_{G_{yx}}$ and $F_{G_{yx}}$ involved in $\mathcal{M}_{G_{xx}}$ and $\mathcal{M}_{G_{yx}}$.

However, as explained in Section 7.1, ensuring the data informativity is not enough in practice to guarantee an accurate model. We have also to choose the corresponding amplitude and phase shift of the sinusoids in order to maximize the power of \bar{u}_x . In order to simplify the choice, we will consider the same amplitude for all the sinusoids. For the phase shifts, we should optimize this choice in order to decrease the crest factor of \bar{u}_x . This would allow to increase the amplitude of each sinusoid, which will subsequently maximize the power of \bar{u}_x and so the SNR. Therefore, we can expect estimates $\hat{\theta}_{G_{xx},init}$ and $\hat{\theta}_{G_{yx},init}$ with a small variance.

To decrease the crest factor, we consider 1000 realizations of the phase shifts with independent uniform distributions between 0 and 2π for each of them and we search for the maximal absolute value among the time samples $u_x(t)$ from $t = 1$ till $t = N = 10000$. We then consider the phase shift realization that provides the least maximal absolute value. After that, we normalize this optimized multisine with its maximal absolute value such that the obtained normalized signal \bar{u}_x is between -1 and $1V$. By adding the DC level u_{DC} to \bar{u}_x , we obtain an input u_x for the identification of $G_{0,xx}$ and $G_{0,yx}$.

The discussion on how to choose properly the degrees of the polynomials $B_{G_{xx}}$, $F_{G_{xx}}$, $B_{G_{yx}}$ and $F_{G_{yx}}$ will be given later in Section 7.3.3.

Initial estimation of $Q_{0,xx}$ and $Q_{0,yx}$

For the parasitic electrical coupling identification (i.e., the identification of the transfer functions $Q_{0,xx}$ and $Q_{0,yx}$), we cannot use the technique considered for the identification of $G_{0,xx}$ and $G_{0,yx}$ since the mechanical coupling is excited at the same frequencies as the

parasitic coupling does (as explained below (7.12)). However, we have initial estimate of $G_{0,xx}$ and $G_{0,yx}$ ($G_{xx}(z, \hat{\theta}_{G_{xx},init})$ and $G_{yx}(z, \hat{\theta}_{G_{yx},init})$) and we will use them to isolate the parasitic capacitive coupling.

In order to isolate the parasitic electrical effect from the mechanical one, we remove the latter from the measured outputs \tilde{x} and \tilde{y} by using both previously identified models $G_{xx}(z, \hat{\theta}_{G_{xx},init})$ and $G_{yx}(z, \hat{\theta}_{G_{yx},init})$, i.e., we will consider the outputs $\varrho_{xx}(t) = \tilde{x}(t) - G_{xx}(z, \hat{\theta}_{G_{xx},init})u_x^2(t)$ and $\varrho_{yx}(t) = \tilde{y}(t) - G_{yx}(z, \hat{\theta}_{G_{yx},init})u_x^2(t)$ for the identification of $Q_{0,xx}(z)$ and $Q_{0,yx}(z)$ respectively.

For the input choice, we can use any excitation signal u_x given this isolation. A random binary sequence (RBS) is then a good idea since we want a model over the whole frequency range. We then consider for \bar{u}_x a RBS whose amplitude is 1 such that the obtained input u_x has two possible values : 0V and 2V. Note that this excitation is the one that will provide the maximal power w.r.t. the input limitation given in Section 7.3.1. Moreover, this type of excitation yields informative data w.r.t. any SISO model structure, irrespectively of its complexity.

We apply this excitation to the MEMS gyroscope (with $u_y = 0$) from $t = 1$ till $t = N = 10000$ and we collect the following data sets for the identification

$$\begin{aligned}\mathcal{Z}_{Q_{xx}}^N &= \{(\varrho_{xx}(t), u_x(t))^T \mid t = 1, \dots, N\} \\ \mathcal{Z}_{Q_{yx}}^N &= \{(\varrho_{yx}(t), u_x(t))^T \mid t = 1, \dots, N\}\end{aligned}$$

For the same reason that we gave for the SISO identification of $G_{0,xx}(z)$ and $G_{0,yx}(z)$, we will also consider the following SISO OE model structures for the SISO identification of $Q_{0,xx}(z)$ and $Q_{0,yx}(z)$

$$\begin{aligned}\mathcal{M}_{Q_{xx}} &= \{Q_{xx}(z, \tilde{\theta}_{Q_{xx}}) \mid \tilde{\theta}_{Q_{xx}} \in \mathcal{D}_{\theta_{Q_{xx}}}\} \\ \mathcal{M}_{Q_{yx}} &= \{Q_{yx}(z, \tilde{\theta}_{Q_{yx}}) \mid \tilde{\theta}_{Q_{yx}} \in \mathcal{D}_{\theta_{Q_{yx}}}\}\end{aligned}$$

where $Q_{xx}(z, \tilde{\theta}_{Q_{xx}})$ and $Q_{yx}(z, \tilde{\theta}_{Q_{yx}})$ are defined in the model structures \mathcal{M}_{xx} and \mathcal{M}_{yx} in Section 7.2.2.

Therefore, initial estimates $\hat{\theta}_{Q_{xx}}$ and $\hat{\theta}_{Q_{yx}}$ can be computed with the following least-square optimizations as defined below

$$\begin{aligned}\hat{\theta}_{Q_{xx},init} &= \arg \min_{\tilde{\theta}_{Q_{xx}} \in \mathcal{D}_{\theta_{Q_{xx}}}} \frac{1}{N} \sum_{t=1}^N \epsilon_{Q_{xx}}(t, \tilde{\theta}_{Q_{xx}})^2 \\ \hat{\theta}_{Q_{yx},init} &= \arg \min_{\tilde{\theta}_{Q_{yx}} \in \mathcal{D}_{\theta_{Q_{yx}}}} \frac{1}{N} \sum_{t=1}^N \epsilon_{Q_{yx}}(t, \tilde{\theta}_{Q_{yx}})^2\end{aligned}$$

where $\epsilon_{Q_{xx}}$ and $\epsilon_{Q_{yx}}$ are given by $\epsilon_{Q_{xx}}(t, \tilde{\theta}_{Q_{xx}}) = \varrho_{xx}(t) - G_{xx}(z, \tilde{\theta}_{Q_{xx}})u_x(t)$ and $\epsilon_{Q_{yx}}(t, \tilde{\theta}_{Q_{yx}}) = \varrho_{yx}(t) - Q_{yx}(z, \tilde{\theta}_{Q_{yx}})u_x(t)$.

The discussion on the choice of the degrees of $B_{Q_{xx}}$, $F_{Q_{xx}}$, $B_{Q_{yx}}$ and $F_{Q_{yx}}$ will be later given in Section 7.3.3.

Remark 7.2. *It is clear that the accuracy of the initial estimates $Q_{xx}(z, \hat{\theta}_{Q_{xx},init})$ and $Q_{yx}(z, \hat{\theta}_{Q_{yx},init})$ will depend on the accuracy of the estimates $G_{xx}(z, \hat{\theta}_{G_{xx},init})$ and $G_{yx}(z, \hat{\theta}_{G_{yx},init})$ since these models are used to remove the outputs of the mechanical dynamics in the measured outputs.*

Remark 7.3. *The input choices for u_y and the procedure for initial estimation is the same for the transfer functions in both model structures \mathcal{M}_{xy} and \mathcal{M}_{yy} .*

7.3.3 Results of the initial identifications and interpretation

In the previous section, we have not discussed how the degrees of the different polynomials involved in the definition of the different model structures are chosen. In a nutshell and as usually done in system identification, we choose these degrees equal to the smallest values which guarantee that the prediction error evaluated at the identified parameter vector is close to a white noise [Lju99]. The obtained degrees are given in Tables 7.1, 7.2 and 7.3. The delays of all the transfer functions in \mathbf{Q} and in \mathbf{G} are equal to 1.

Table 7.1: Obtained degrees of the polynomials for the mechanical coupling functions.

$B_{G_{xx}}$	$B_{G_{xy}}$	$B_{G_{yx}}$	$B_{G_{yy}}$	$F_{G_{xx}}$	$F_{G_{xy}}$	$F_{G_{yx}}$	$F_{G_{yy}}$
2	2	2	2	2	4	4	2

Table 7.2: Obtained degrees of the polynomials for the parasitic coupling transfer functions.

$B_{Q_{xx}}$	$B_{Q_{xy}}$	$B_{Q_{yx}}$	$B_{Q_{yy}}$	$F_{Q_{xx}}$	$F_{Q_{xy}}$	$F_{Q_{yx}}$	$F_{Q_{yy}}$
3	3	3	3	3	3	3	3

Table 7.3: Obtained degrees of the polynomials for the noise transfer functions.

C_{H_x}	C_{H_y}	D_{H_x}	D_{H_y}
6	5	6	5

We have also evaluated the ability of the different identified models to explain another set of $N = 10000$ data (having the same spectrum as the data used for the identification). For the transfer function G_{xx} , we generate a new set of data $\tilde{x}(t)$ and $u_x^2(t)$

($t = 1, \dots, N$) using the procedure presented in Section 7.3.2 and we compute the best fit of the identified model $G_{xx}(z, \hat{\theta}_{G_{xx},init})$ as follows:

$$\text{Best Fit for } G_{xx} = 100\% \left(1 - \frac{\sqrt{\sum_{t=1}^N (\tilde{x}(t) - G_{xx}(z, \hat{\theta}_{G_{xx},init}) u_x^2(t))^2}}{\sqrt{\sum_{t=1}^N (\tilde{x}(t) - \bar{x})^2}} \right)$$

where \bar{x} is the mean value of the set $\{\tilde{x}(t) \mid t = 1, \dots, N\}$. We define similarly the Best Fits of G_{xy} , G_{yx} and G_{yy} .

As another example, for Q_{xx} , we generate a new set of data $q_{xx}(t)$ and $u_x(t)$ ($t = 1, \dots, N$) using the procedure presented in Section 7.3.2 and we compute the best fit of the identified model $Q_{xx}(z, \hat{\theta}_{Q_{xx},init})$:

$$\text{Best Fit for } Q_{xx} = 100\% \left(1 - \frac{\sqrt{\sum_{t=1}^N (q_{xx}(t) - Q_{xx}(z, \hat{\theta}_{Q_{xx},init}) u_x(t))^2}}{\sqrt{\sum_{t=1}^N (\tilde{x}(t) - \bar{x})^2}} \right)$$

We define similarly the Best Fits of Q_{xy} , Q_{yx} and Q_{yy} .

All the computed Best Fits are given in Tables 7.4 (for the transfer functions in \mathbf{G}) and 7.5 (for the transfer functions in \mathbf{Q}). We observe high Best Fits for the identified models $G_{yy}(z, \hat{\theta}_{G_{yy},init})$, $Q_{xx}(z, \hat{\theta}_{Q_{xx},init})$ and $Q_{yx}(z, \hat{\theta}_{Q_{yx},init})$ which might allow us to conclude that these estimates are reliable. For the other models, the Best Fits are low and we might conclude that they are not reliable. For $Q_{yx}(z, \hat{\theta}_{Q_{yx},init})$, the Best Fit is close to 0%: this estimate is clearly unreliable.

The Bode diagrams of all identified models are given in Figures 7.1, 7.2 and 7.3. Let us discuss these Bode diagrams, starting with the ones corresponding to the parasitic capacitive coupling (See Figure 7.2). As expected, the identified models for Q_{xx} , Q_{xy} and Q_{yy} present all an highpass effect. We also observe that the gain of these identified models are important and this is particularly the case in the frequency range of the resonance frequencies of the gyroscope (around $2\pi \times 11836.8$ rad/s). It seems that there is almost no capacitive coupling from the drive mode to the sense mode since $Q_{yx}(z, \hat{\theta}_{Q_{yx},init}) \approx 0$ but this estimate is unreliable due to the low Best Fit obtained. Finally, it seems that there is a lowpass effect in high frequencies. This may be due to extra dynamics coming from the electronic card and so our conjecture given in Section 6.5.3 seems to be true here. Therefore, it seems that choosing a black-box approach was a good initiative after all.

Secondly, let us interpret the results of the identified mechanical coupling transfer functions (see Figure 7.1). First, we can observe a resonance behavior in all the transfer functions. For G_{xy} and G_{yx} , we can observe two resonances why for the direct terms G_{xx} and G_{yy} only one resonance is visible. The quality factors of the resonances of the identified transfer functions G_{xx} and G_{yy} have the same magnitude order as the ones

given by the gyroscope manufacturer. For the direct transfer functions G_{xx} and G_{yy} , an unstable zero has been identified but it does not correspond to a physical phenomenon. It might come from the electronic card dynamics. However, it is important to note that all the validity of the models in \mathbf{G} can only be guaranteed in the frequency range $2\pi \times [11750, 11950]$ rad/s due to the data-prefiltering used in the SISO-by-SISO approach. The validity of the models in \mathbf{Q} is also questionable since the models in \mathbf{G} were used for the determination of the parasitic capacitive coupling transfer functions (Remark 7.2).

Finally, for the noise measurement (see Figure 7.3), we see a slope -20dB/decade in the low frequency range, which is the effect of the flicker noise. Then, a constant part is observed in H_x and which is due to the mechanical-thermal noise. It is however not observed in H_y . A small peak is observed in high frequencies in H_x and a slope of -40dB/dec is present in high frequencies for both PSD. Further works should be done on determining the physical interpretation of this peak and the slope of -40dB/dec in high frequencies. For the latter, the extra dynamics coming from the electronic card might be the reason.

It is clear that we cannot use these models for the design of the controller since we have some validity issues with the transfer functions in \mathbf{G} and \mathbf{Q} . In order to circumvent this validity issue, we are going to perform the MISO identifications explained in Section 7.2 without doing some data-prefiltering. As we will see in the next section, the models obtained using this MISO approach will be very close to the ones obtained using the SISO-by-SISO approach. However, since the input excitation used in the MISO identification will be more broadband (than it was the case, e.g, for the identification of the transfer functions in \mathbf{G}), we will have more confidence in the obtained models.

Table 7.4: Best Fit for the mechanical coupling from the SISO-by-SISO approach.

	G_{xx}	G_{xy}	G_{yx}	G_{yy}
Best Fit	55.69%	57.51%	75.18%	97.47%

Table 7.5: Best Fit for the parasitic coupling from the SISO-by-SISO approach.

	Q_{xx}	Q_{xy}	Q_{yx}	Q_{yy}
Best Fit	90.1%	93.59%	-0.28%	59.01%

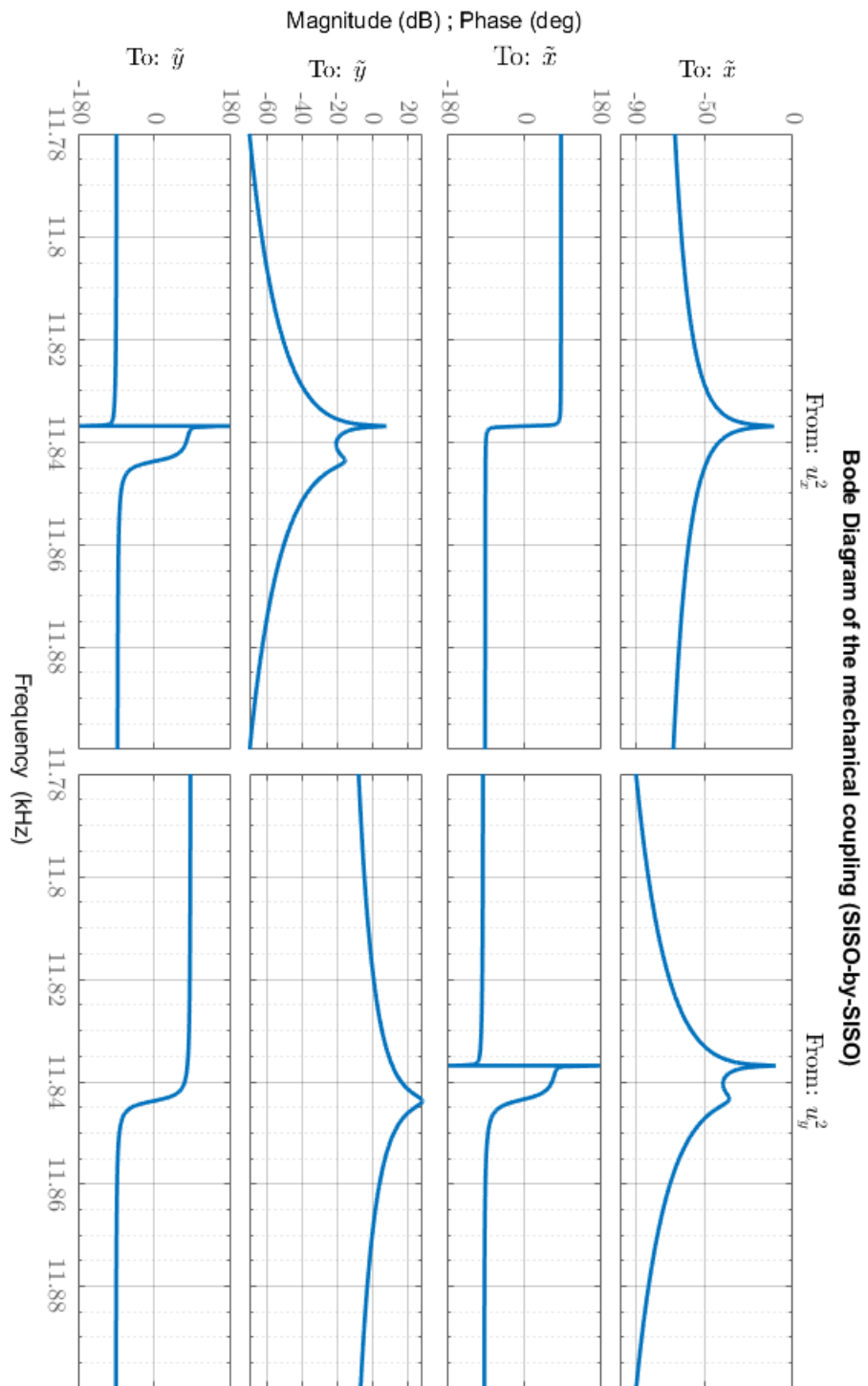


Figure 7.1: Bode diagrams of the mechanical coupling from the SISO-by-SISO approach.

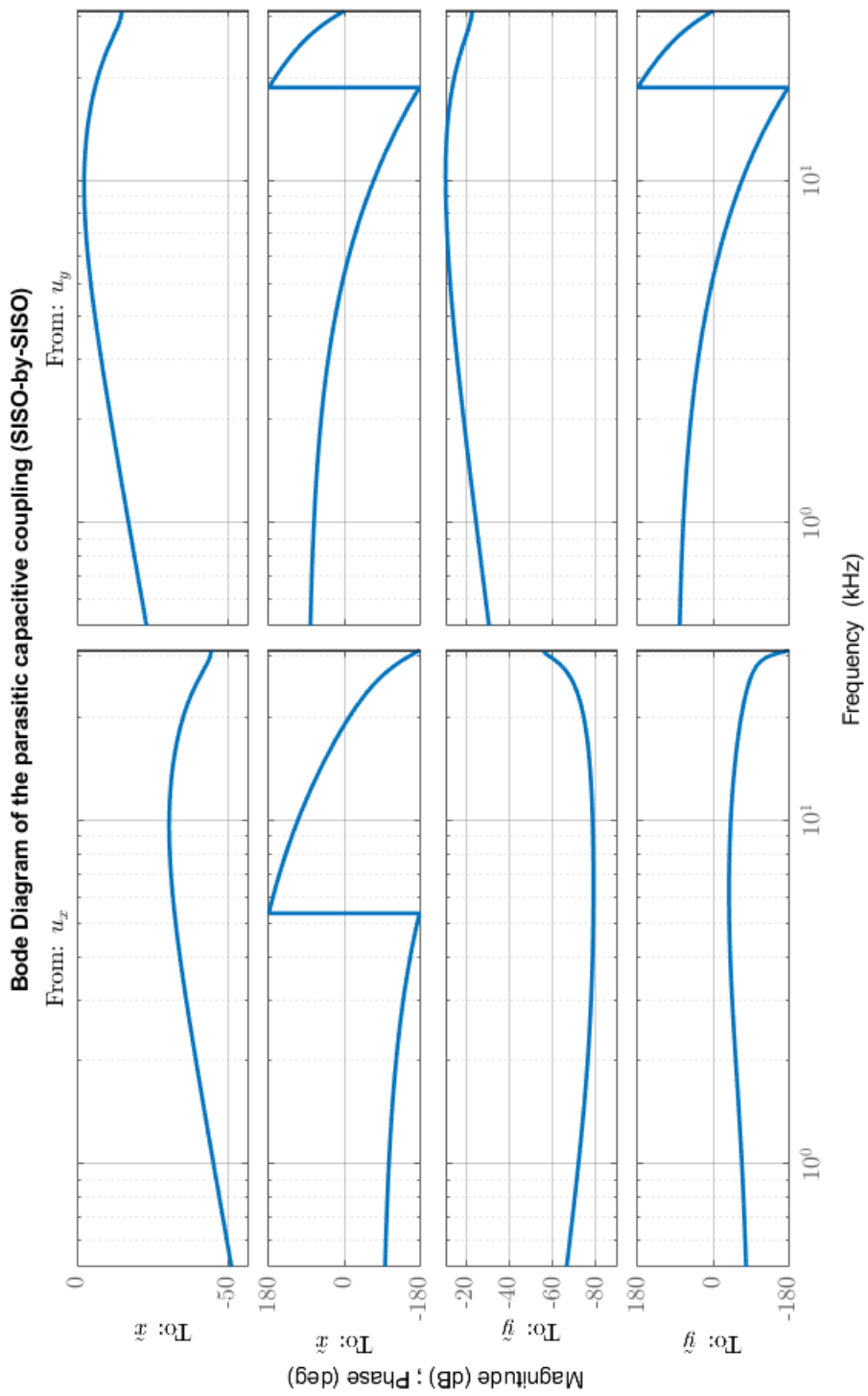


Figure 7.2: Bode diagrams of the parasitic capacitive coupling from the SISO-by-SISO approach.

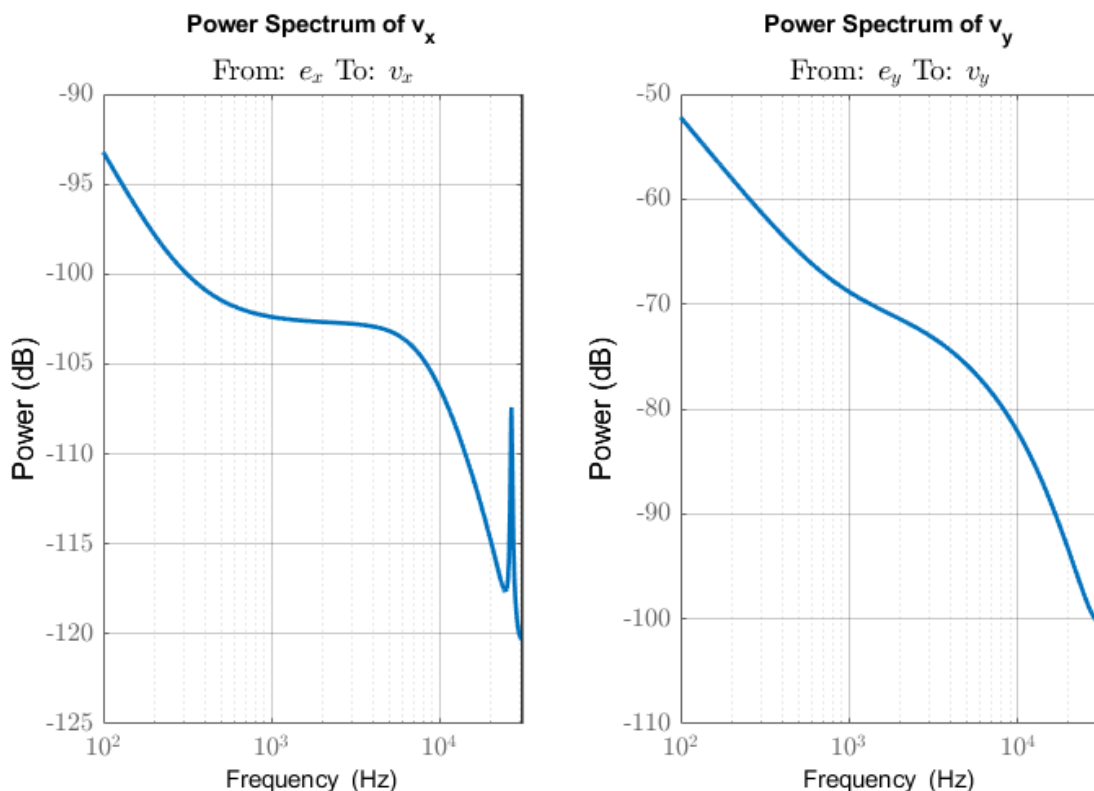


Figure 7.3: Estimated PSD of v_x (left) and v_y (right) from the SISO-by-SISO approach.

7.4 MISO identification and data informativity

Using the identification criteria (7.7) and (7.8), we can derive models for the subsystems (7.4) and (7.5) corresponding to the MEMS gyroscope (7.1) when $u_y = 0$. For this purpose, we need to perform an experiment by applying a signal u_x to (7.1) while keeping $u_y = 0$. The choice of the excitation signal u_x will be described in the sequel. The optimization procedure to solve the criteria (7.7) and (7.8) will be initialized with the models obtained in the previous section. Consequently, the model structures that will be used in the criteria (7.7) and (7.8) must be the same as the model structures used during the initial identification (see Tables 7.1, 7.2 and 7.3).

Note that we will only describe the MISO identification of the subsystems (7.4) and (7.5) with $u_x \neq 0$ and $u_y = 0$ (see (7.7) and (7.8)). A similar approach can be taken for the identification of \mathcal{S}_{xy} and \mathcal{S}_{yy} , i.e., when $u_x = 0$ and $u_y \neq 0$.

7.4.1 Input choice for the MISO identification

First, it is important that the spectrum of u_x^2 is non zero in the direct neighbourhood of $\omega_{0,x}$ and $\omega_{0,y}$ in order to excite well the mechanical dynamics. Therefore, as in Section 7.3, we will choose for the multisine excitation the same excitation frequencies around $\omega_{0,x}/2$ and $\omega_{0,y}/2$ as in Section 7.3.2 (Sets 1, 2 and 3) in order to excite only the mechanical effects around the resonances frequencies $\omega_{0,x}$ and $\omega_{0,y}$ thanks to the square nonlinearity. Secondly, for the parasitic coupling, it is also important to excite in low and high frequencies since the parasitic coupling has also dynamics in these frequency ranges. For this purpose, we will add sinusoidal frequencies to Sets 1, 2 and 3 in low and high frequencies. Therefore, the final multisine excitation is made up of 250 frequencies distributed over the interval $[0, \omega_s/2]$. They are regrouped in sets and distributed regularly in each set as follows

- Set 1 (around $\omega_{0,x}/2$): from $2\pi \times 5918.415$ till $2\pi \times 5918.49$ with a frequency bin of $2\pi \times 0.0005$. In total, there are 151 frequencies.
- Set 2 (between $\omega_{0,x}/2$ and $\omega_{0,y}/2$): from $2\pi \times 5918.5$ till $2\pi \times 5921$ with a frequency bin of $2\pi \times 0.5$. In total, there are 6 frequencies.
- Set 3 (around $\omega_{0,y}/2$): from $2\pi \times 5921.5$ till $2\pi \times 5922.2$ with a frequency bin of $2\pi \times 0.005$. In total, there are 141 frequencies.
- Set 4 (low frequency range): from $2\pi \times 500$ till $2\pi \times 5000$ with a frequency bin of $2\pi \times 100$. In total, there are 46 frequencies.
- Set 5 (after $\omega_{0,y}/2$ and before $\omega_{0,x}$): from $2\pi \times 6000$ till $2\pi \times 11000$ with a frequency bin of $2\pi \times 1000$. In total, there are 11 frequencies.
- Set 6 (after $\omega_{0,y}$ until $\omega_s/2$): from $2\pi \times 12000$ till $2\pi \times 31000$ with a frequency bin of $2\pi \times 1000$. In total, there are 20 frequencies.

It is also illustrated in Figure 7.4.

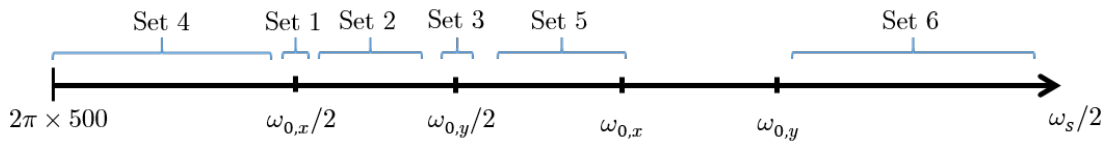


Figure 7.4: Illustration of the sets of frequency for the multisine excitation.

This multisine contains sines over the whole frequency range, in particular in the direct neighbourhood of $\omega_{0,x}/2$ and $\omega_{0,y}/2$ near *but not* in the direct neighbourhood of $\omega_{0,x}$ and $\omega_{0,y}$. According to (7.12), this means that both u_x and u_x^2 will be broadband, but only u_x^2 will have contributions in the direct neighbourhood of $\omega_{0,x}$ and $\omega_{0,y}$. This

is important to guarantee an accurate identification of both \mathbf{Q} and \mathbf{G} . Observe indeed that the gain of \mathbf{Q} is maximal (and sometimes higher than the one of \mathbf{G}) in the direct neighbourhood of $\omega_{0,x}$ and $\omega_{0,y}$, which is the frequency range where the "only" dynamics of \mathbf{G} are present.

For the amplitude choice, we decide to choose the same amplitude for all sinusoids. For the phase shifts, we optimize this choice in order to decrease the crest factor of the multisine signal as we did in Section 7.3.2. However, instead of $N = 10000$ as it was the case in Section 7.3.2, we will consider $N = 62500$ which will be the number of data that will be used for the MISO identification. Although this input choice allows to increase the maximal power in order to get small variances, it must yield informative data so that it will be the case. The study of the data informativity is done in the next paragraph.

7.4.2 Study of the data informativity

It is interesting to note that the results of Chapter 5 allow us to verify that the chosen excitation u_x will yield informative data for the identification criteria (7.7) and (7.8).

First, in u_x , we have $s_1 = 250$ nonzero frequencies and one constant level. Consequently, the PSD of u_x is nonzero at $m_1 = 2s_1 + 1 = 501$ frequencies.

We cannot use Lemma 5.2 with $n = 2$ to conclude on the data informativity since we have some frequencies and their double in the set of frequencies, violating the condition (iii) of Lemma 5.2. For instance, both frequencies $2\pi \times 1000$ and $2\pi \times 2000$ are part of the set of excitation frequencies. Let us then try to use Lemma 5.1 by computing m_2 and $m_{1,2}$. For this computation, we could use the procedure given in Section 5.4.1. However, even though all the amplitudes are taken positive, the phase shifts are not equal to 0. Therefore, we compute the FFT of u_x^2 with a large number N' of data ($N' = 10^7$) in order to give a lower bound for m_2 . We also compute the FFT of u_x in order to determine a lower bound for $m_{1,2}$. From the observation of the FFT, We conclude that $m_2 - m_{1,2} \geq 100$. Therefore, with this excitation, we have that

$$\begin{aligned} \deg(B_{Q_{xx}}) + \deg(F_{Q_{xx}}) &\leq 501 = m_1 \\ \deg(B_{G_{xx}}) + \deg(F_{G_{xx}}) &\leq 100 \leq m_2 - m_{1,2} \end{aligned}$$

and that

$$\begin{aligned} \deg(B_{Q_{yx}}) + \deg(F_{Q_{yx}}) &\leq 501 = m_1 \\ \deg(B_{G_{yx}}) + \deg(F_{G_{yx}}) &\leq 100 \leq m_2 - m_{1,2} \end{aligned}$$

where these degrees are the ones determined from the SISO-by-SISO approach and which are given in Tables 7.1 and 7.2.

Consequently, the chosen excitation u_x will yield informative data for the identification criteria (7.7) and (7.8).

7.4.3 Identification results and comparison with initial model

Using an experiment with $u_y = 0$ and the signal u_x presented in Section 7.4.1 from $t = 1$ till $t = N = 62500$, we determine the solutions of the identification criteria (7.7) and (7.8), initializing these optimizations problems at the parameter vectors obtained in Section 7.3. This identification experiment leads to models for the transfer functions in \mathcal{S}_{xx} and \mathcal{S}_{yx} (see 7.4 and 7.5). Similarly, using an experiment with $u_x = 0$ and a signal u_y having the same expression as the signal u_x described in Section 7.4.1, we can determine models for the transfer functions in \mathcal{S}_{xy} and \mathcal{S}_{yy} .

Note that, even if the MISO identification criteria are initialized using the models obtained in Section 7.3, we could observe some local minimum problems since identified models of $G_{0,xy}$ and $G_{0,yx}$ gave us results that were absurd (resonance frequencies that were far away from $\omega_{0,x}$ and $\omega_{0,y}$). We have to further constrain each parameter of $\theta_{G_{xy}}$ and $\theta_{G_{yx}}$ to range from -5% to $+5\%$ of the initial (identified) value. This means that, during the optimization, the parameter vectors $\theta_{G_{xy}}$ and $\theta_{G_{yx}}$ are constrained to lie in a hyperrectangle whose center is $\hat{\theta}_{G_{xy},init}$ for $\theta_{G_{xy}}$ and $\hat{\theta}_{G_{yx},init}$ for $\theta_{G_{yx}}$ with a maximal deviation of 5% for each parameter. This can be implemented with the `System Identification Toolbox` of MATLAB[®].

The Bode diagrams of the obtained models compared with the ones from the SISO-by-SISO approach are given in Figures F.1 and F.2 in Appendix F.1. As already mentioned, we observed that the models identified using the MISO approach are very close to the initial models obtained using the SISO-by-SISO approach. For G_{xy} and G_{yx} , this is surely explained by the constrained optimization.

We have performed two validation experiments, using the same excitation signals as the experiments used for the identification. This allows to compute the Best Fit of the identified models for the broadband excitation signal defined in Section 7.4.1. The first validation experiment ($u_y = 0$ and u_x as in Section 7.4.1) yields validation data u_x , u_x^2 , \tilde{x} and \tilde{y} for $t = 1$ till $t = N = 50000$. Using these data, we can define:

$$\text{Best Fit}_{xx} = 100\% \left(1 - \frac{\sqrt{\sum_{t=1}^N (\tilde{x}(t) - Q_{xx}(z, \hat{\theta}_{N,xx})u_x(t) - G_{xx}(z, \hat{\theta}_{N,xx})u_x^2(t))^2}}{\sqrt{\sum_{t=1}^N (\tilde{x}(t) - \bar{x})^2}} \right)$$

$$\text{Best Fit}_{yx} = 100\% \left(1 - \frac{\sqrt{\sum_{t=1}^N (\tilde{y}(t) - Q_{yx}(z, \hat{\theta}_{N,yx})u_x(t) - G_{yx}(z, \hat{\theta}_{N,yx})u_x^2(t))^2}}{\sqrt{\sum_{t=1}^N (\tilde{y}(t) - \bar{y})^2}} \right)$$

where \bar{x} is the mean value of the set $\{\tilde{x}(t) \mid t = 1, \dots, N\}$ and \bar{y} is the mean value of the set $\{\tilde{y}(t) \mid t = 1, \dots, N\}$ with $N = 50000$. The second validation experiment ($u_x = 0$ and u_y with the spectrum given in Section 7.4.1) yields similar Best Fit_{xy} and Best Fit_{yy}.

These four Best Fits are given in Table 7.6 where we observe that they are all extremely high except Best Fit $_{xy}$ which is relatively smaller. This low Best Fit can be explained by the fact that Q_{yx} seems to be equal to 0 and that only the dynamics of G_{yx} are visible in the data. Then, the fact that there is no data-prefiltering may explain why Best Fit $_{xy}$ is lower than the Best Fit obtained for the initial identification of G_{yx} since the contribution of the noise is higher on the identified parameter without data-prefiltering. We will nevertheless deem the models identified using this MISO approach validated and use them for the controller design. In other words, we will use the identified model of (7.1) to design the blocks "Parasitic Compensation $\mathbf{K}_{comp}(z)$ " and "Tracking Controller $\mathbf{K}(z)$ " (see Figure 6.10). In the next section, we present the strategy for the block "Parasitic Compensation $\mathbf{K}_{comp}(z)$ ".

Table 7.6: Best Fit for the models obtained from the MISO approach with $N = 50000$.

Best Fit $_{xx}$	Best Fit $_{yx}$	Best Fit $_{xy}$	Best Fit $_{yy}$
96.68%	90.53%	59.33%	92.9%

7.5 Block "Parasitic Compensation $\mathbf{K}_{comp}(z)$ "

7.5.1 Strategy: model-based compensation

In this section, we focus on the block "Parasitic Compensation $\mathbf{K}_{comp}(z)$ " in Figure 6.10. The main purpose of this block is to compensate the parasitic capacitive coupling in order to guarantee that both signals \tilde{x}_c and \tilde{y}_c defined in (6.9) are almost equal to $\tilde{x}_{mec} + v_x$ and $\tilde{y}_{mec} + v_y$ respectively.

As explained in Section 6.5.3, the block "Parasitic Compensation $\mathbf{K}_{comp}(z)$ " is chosen as $-\hat{\mathbf{Q}}(z)$ where $\hat{\mathbf{Q}}(z)$ is an estimate of $\mathbf{Q}_0(z)$. In our case, it is the transfer function matrix made up of model of the identified parasitic capacitive coupling transfers $Q_{xx}(z, \hat{\boldsymbol{\theta}}_{N,xx})$, $Q_{xy}(z, \hat{\boldsymbol{\theta}}_{N,xy})$, $Q_{yx}(z, \hat{\boldsymbol{\theta}}_{N,yx})$ and $Q_{yy}(z, \hat{\boldsymbol{\theta}}_{N,yy})$ obtained from the MISO identifications in Section 7.4 and defined as below

$$\hat{\mathbf{Q}}(z) = \begin{pmatrix} Q_{xx}(z, \hat{\boldsymbol{\theta}}_{N,xx}) & Q_{xy}(z, \hat{\boldsymbol{\theta}}_{N,xy}) \\ Q_{yx}(z, \hat{\boldsymbol{\theta}}_{N,yx}) & Q_{yy}(z, \hat{\boldsymbol{\theta}}_{N,yy}) \end{pmatrix} \quad (7.13)$$

7.5.2 Results of the model-based compensation

In this paragraph, we are going to verify if the model identified in Section 7.4 is sufficiently accurate to ensure an efficient compensation of the capacitive coupling. For this purpose, we consider the setup represented in Figure 7.5 where the gyroscope is operated in open loop and where only the block "Parasitic Compensation $\mathbf{K}_{comp}(z)$ " $-\hat{\mathbf{Q}}(z)$ in (7.13) is implemented (no block "Tracking Controller \mathbf{K} "). In this configuration, we simultaneously apply an excitation on u_x and u_y . These excitation signals are chosen (see (7.9)) as $u_x(t) = \bar{u}_x(t) + 1$ and $u_y(t) = \bar{u}_y(t) + 1$ with \bar{u}_x and \bar{u}_y are independent RBS of amplitude equal to 1V. We then collect the corresponding output signals before the compensation (i.e, \tilde{x} , \tilde{y}) and after the compensation (i.e, \tilde{x}_c and \tilde{y}_c).

If the compensation works well, we should expect the FFTs of the compensated outputs \tilde{x}_c and \tilde{y}_c to be almost equal to the FFT of the mechanical outputs \tilde{x}_{mec} and \tilde{y}_{mec} in the neighbourhood of the resonance frequencies. Indeed, in this frequency range, the effect of the noise is negligible due to the high quality factors of both resonances. For this purpose, as shown in Figure 7.5, we also simulate the outputs \tilde{x}_{sim} and \tilde{y}_{sim} from the identified model of the mechanical transfer functions as given below

$$\begin{pmatrix} \tilde{x}_{sim}(t) \\ \tilde{y}_{sim}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} G_{xx}(z, \hat{\theta}_{N,xx}) & G_{xy}(z, \hat{\theta}_{N,xy}) \\ G_{xy}(z, \hat{\theta}_{N,xy}) & G_{yy}(z, \hat{\theta}_{N,yy}) \end{pmatrix}}_{\hat{\mathbf{G}}(z)} \begin{pmatrix} u_x^2(t) \\ u_y^2(t) \end{pmatrix} \quad (7.14)$$

where u_x and u_y are the same RBS inputs which were used to excite the gyroscope. We will then compare the FFT of these simulated mechanical outputs with the ones of the compensated outputs \tilde{x}_c and \tilde{y}_c . If both FFTs match well in the neighbourhood of the resonance frequencies, this will allow us to conclude that the compensation works well in this frequency range and that the model of the mechanical coupling is quite accurate in this frequency range.

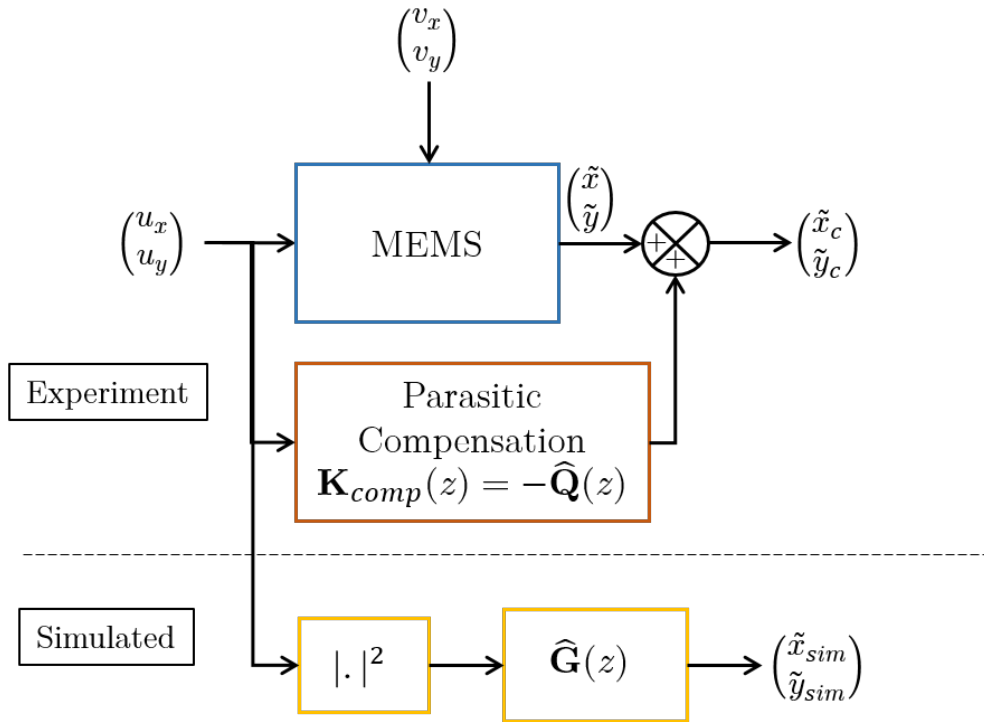


Figure 7.5: Scheme of the experiment for the verification of the parasitic coupling compensation.

The FFTs of the drive output before compensation (\tilde{x} , blue curve) and after compensation (\tilde{x}_c , orange curve) are depicted in Figure 7.6. We also add the FFT of the signal (\tilde{x}_{sim} , yellow curve) simulated with the identified model of the mechanical coupling. A zoom around the resonance frequency is given in Figure 7.7. By looking at the blue curve in Figure 7.6, we observe that both resonances are completely hidden by the parasitic capacitive coupling. The digital compensation allows an important reduction of this effect and the resonance peaks at $\omega_{0,x}$ and $\omega_{0,y}$ are now visible as illustrates the orange curve in Figure 7.6. Around the resonance frequencies $\omega_{0,x}$ and $\omega_{0,y}$, we observe that the FFTs of the compensated output \tilde{x}_c (orange curve) and of the simulated one \tilde{x}_{sim} (yellow curve) are close, except in the direct neighbourhood of $\omega_{0,y}$. This is not surprising since we have already observed that the model of the subsystem from u_y to \tilde{x} is not very accurate. However, it seems that the compensation performs quite well near $\omega_{0,x}$ since both FFTs match well in this frequency range.

For the sense mode output, Figure 7.8 gives the FFTs of the sense mode output before compensation (\tilde{y} , blue curve) and after compensation (\tilde{y}_c , orange curve). We also add the FFT of the signal \tilde{y}_{sim} simulated with the identified model of the mechanical coupling in Section 7.4 (\tilde{y}_{sim} , yellow curve). A zoom is also provided around the resonance frequencies in Figure 7.9. In this case, the resonance peak is visible without the compensation (blue curve) but with high gains before and after the resonances. The compensation (orange curve) allows to reduce the gain before and after the resonances. Moreover, the compensated output \tilde{y}_c corresponds to the one simulated from the model (yellow curve) around the resonance frequencies.

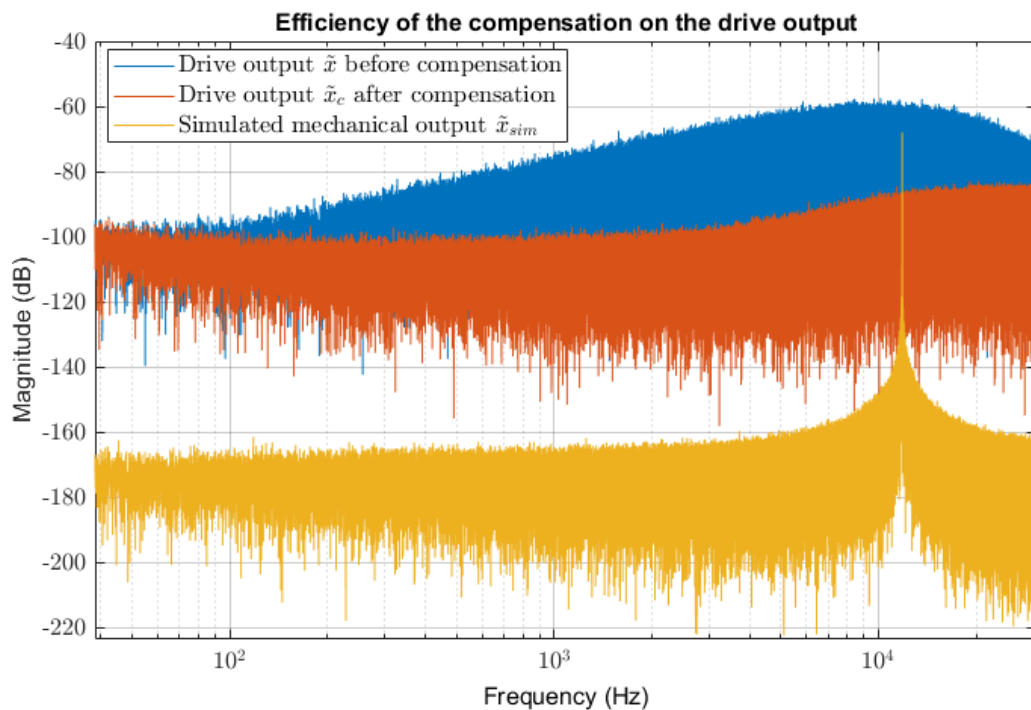


Figure 7.6: FFT magnitudes of the drive output before compensation \tilde{x} (blue), after compensation \tilde{x}_c (orange) and the simulated drive output \tilde{x}_{sim} from $\hat{\mathbf{G}}(z)$ (yellow).

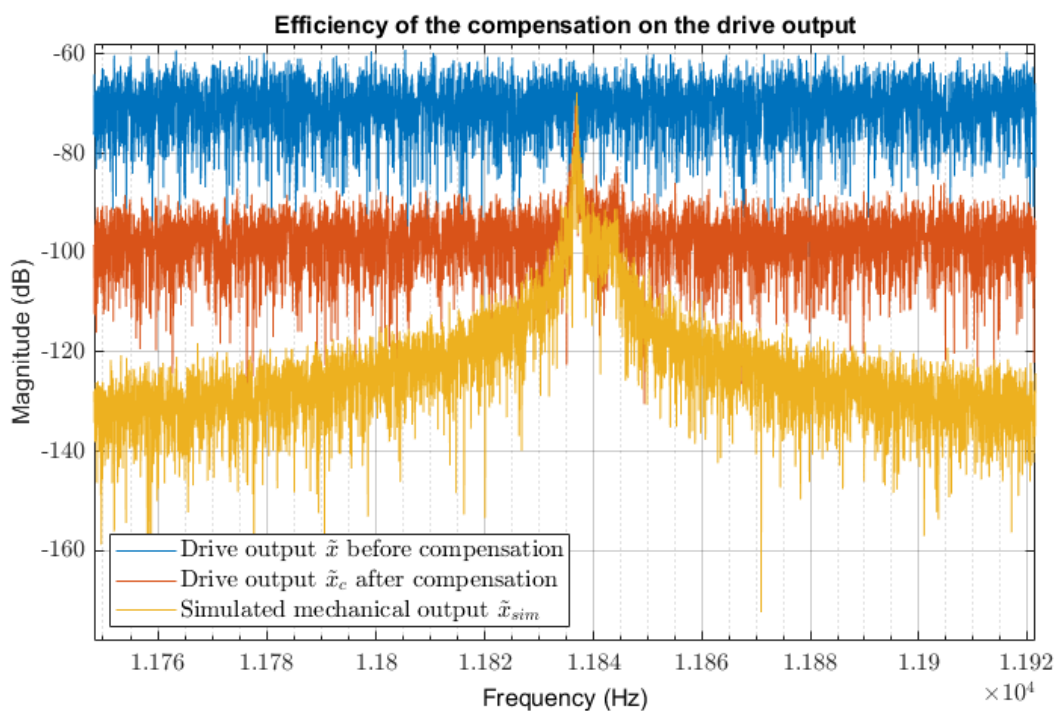


Figure 7.7: Zoom of Figure 7.6 around the resonance frequencies.

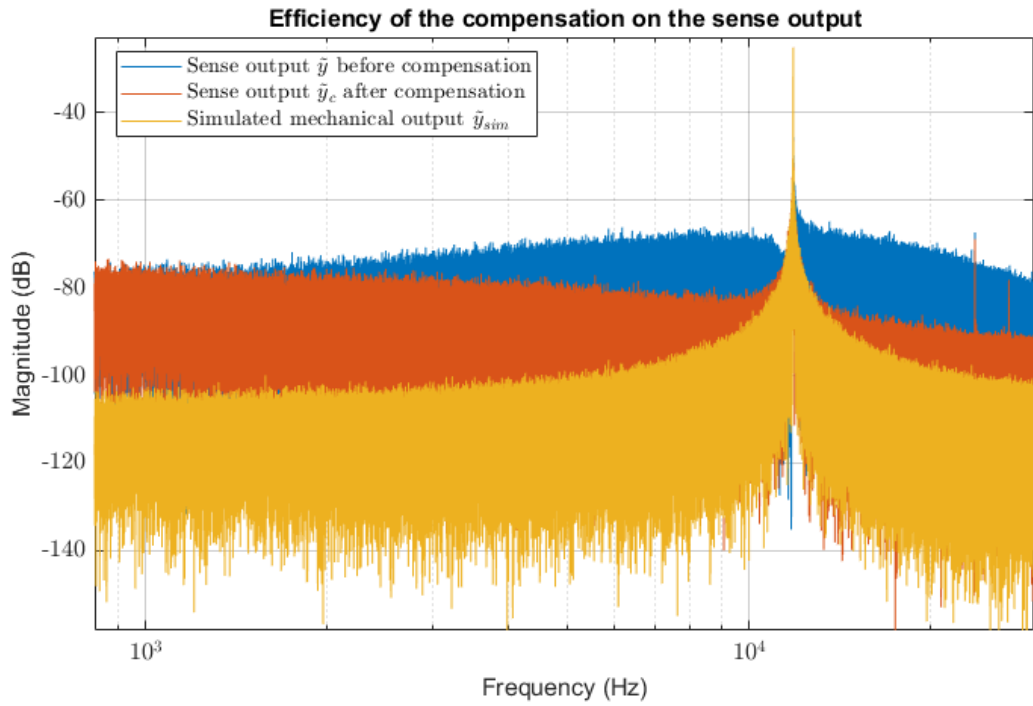


Figure 7.8: FFT magnitudes of the sense output before compensation \tilde{y} (blue), after compensation \tilde{y}_c (orange) and the simulated sense output \tilde{y}_{sim} from $\hat{\mathbf{G}}(z)$ (yellow).

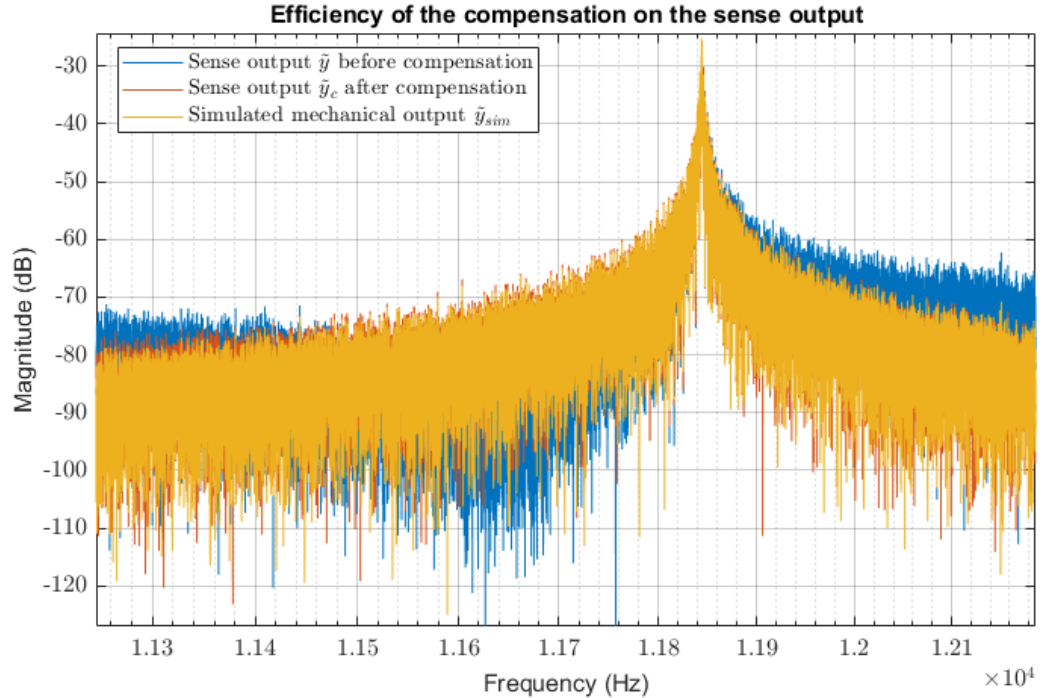


Figure 7.9: Zoom of Figure 7.8 around the resonance frequencies.

Therefore, we can fairly consider that the parasitic electrical coupling is well compensated and so the signal vector that is sent to the block "Tracking Controller $\mathbf{K}(z)$ " in Figure 6.10 depends only on the noise and the mechanical dynamics of the MEMS gyroscope. Consequently, we can design the latter by only considering the identified models of the noise $H_x(z, \hat{\boldsymbol{\eta}}_{N,x})$ and $H_y(z, \hat{\boldsymbol{\eta}}_{N,y})$ and the mechanical transfer function matrix $\hat{\mathbf{G}}(z)$. This block design is presented in the next section.

7.6 Block "Tracking Controller $\mathbf{K}(z)$ "

7.6.1 Issue for the design: high complexity of $\mathbf{G}(z, \hat{\boldsymbol{\theta}}_N)$

In this section, we will focus on the block "Tracking Controller $\mathbf{K}(z)$ ". This block is synthesized with the \mathcal{H}_∞ method and it will be based on the identified models $H_x(z, \hat{\boldsymbol{\eta}}_{N,x})$ and $H_y(z, \hat{\boldsymbol{\eta}}_{N,y})$ of the noise and the mechanical coupling function matrix $\hat{\mathbf{G}}(z)$ defined in (7.14). However, a \mathcal{H}_∞ controller has an order which is equal to the sum of the model orders and the weighting function orders used in the design. Unfortunately, if we compute the order of the MIMO model $\hat{\mathbf{G}}(z)$ given in (7.14), we obtain 12 which is quite high. Therefore, the controller will be too complex to be implemented in the electronic card of the MEMS gyroscope. Moreover, it is quite surprising that a MIMO transfer function matrix that describes two resonances has an order of 12. Even by using the MATLAB function `minreal`, we could not reduce this order.

We should subsequently find an approach to reduce the model complexity of $\hat{\mathbf{G}}(z)$. The first method that has been tried is the Hankel-singular-value-based reduction by using the MATLAB function `balred`. The results are given in Appendix F.2 and they show that this method is not successful.

One reason that might explain why the Hankel singular values based reduction did not perform well is linked to our black-box identification modeling approach. Indeed, we have identified the transfer functions in $\hat{\mathbf{G}}(z)$ (see (7.14)) one by one *with independently parametrized transfer functions*. This modeling way prevents the physical links between all the identified transfer functions. A grey-box modeling approach may help us to circumvent this problem. This is presented in the next paragraph.

Remark 7.4. *The latter conclusion might let us think that the black-box modeling approach that we have considered so far seems to be inappropriate. However, it is important to recall that it allowed us to model extra dynamics from the electronic card that could not be modeled with a grey-box model. Moreover, we will see that the black-box model will be useful to prevent local minimum issues that we can meet with the nonlinear optimization algorithm used in the grey-box identification. Therefore, the black-box approach is still relevant for the modeling of the mechanical dynamics involved in the MEMS.*

7.6.2 Grey-box-based model reduction

We will determine a more compact representation of \mathbf{G}_0 using the model description presented in Chapter 6. By assuming that the parasitic capacitive coupling is perfectly compensated, the noises v_x and v_y are equal to 0 and that $\Omega = 0$, the main dynamics can be represented by the following equation system

$$\begin{cases} m_x \ddot{x}(t_c) + d_{xx} \dot{x}(t_c) + k_{xx} x(t_c) + d_{xy} \dot{y}(t_c) + k_{xy} y(t_c) &= \alpha_x u_x^2(t_c) \\ m_y \ddot{y}(t_c) + d_{yy} \dot{y}(t_c) + k_{yy} y(t_c) + d_{yx} \dot{x}(t_c) + k_{yx} x(t_c) &= \alpha_y u_y^2(t_c) \end{cases} \quad (7.15)$$

We should also add both measurement equations

$$\begin{cases} \tilde{x}(t_c) &= \beta_x x(t_c) \\ \tilde{y}(t_c) &= \beta_y y(t_c) \end{cases} \quad (7.16)$$

This model structure will nevertheless not be sufficient. Indeed, as observed in Figure F.3 for the parasite capacitive coupling, we observe extra dynamics with the slope of -20dB/decade after the resonance frequencies. We also observe an unexplained slope in the noise DSPs in Figure 7.3 in the same frequency range. This effect, coming from the electronic card, seems to be modeled by a lowpass filter. Therefore, we will replace the measurement equations in (7.16) by:

$$\begin{cases} \frac{1}{\lambda_x} \dot{\tilde{x}}(t_c) + \tilde{x}(t_c) &= \beta_x x(t_c) \\ \frac{1}{\lambda_y} \dot{\tilde{y}}(t_c) + \tilde{y}(t_c) &= \beta_y y(t_c) \end{cases} \quad (7.17)$$

where λ_x and λ_y are the cut-off frequencies frequencies of this lowpass filter effect. This effect may come from the anti-aliasing filters which have a cutoff frequency of $2\pi \times 15000$ rad/s. The identified values of λ_x and λ_y that we will obtain from this grey-box modeling might help us to conclude on this conjecture.

By considering the following state vector

$$\mathbf{w}(t_c) = \begin{pmatrix} x(t_c) \\ \dot{x}(t_c) \\ y(t_c) \\ \dot{y}(t_c) \\ \tilde{x}(t_c) \\ \tilde{y}(t_c) \end{pmatrix}$$

the equation system (7.15) combined with (7.17) can be put into the following state-space representation

$$\begin{cases} \dot{\mathbf{w}}(t_c) &= \mathbf{A}_w \mathbf{w}(t_c) + \mathbf{B}_w \begin{pmatrix} u_x^2(t_c) \\ u_y^2(t_c) \end{pmatrix} \\ \begin{pmatrix} \tilde{x}(t_c) \\ \tilde{y}(t_c) \end{pmatrix} &= \mathbf{C}_w \mathbf{w}(t_c) \end{cases} \quad (7.18)$$

where

$$\mathbf{A}_w = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{k_{xx}}{m_x} & -\frac{d_{xx}}{m_x} & -\frac{k_{xy}}{m_x} & -\frac{d_{xy}}{m_x} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -\frac{k_{yx}}{m_y} & -\frac{d_{yx}}{m_y} & -\frac{k_{yy}}{m_y} & -\frac{d_{yy}}{m_y} & 0 & 0 \\ \lambda_x \beta_x & 0 & 0 & 0 & -\lambda_x & 0 \\ 0 & 0 & \lambda_y \beta_y & 0 & 0 & -\lambda_y \end{pmatrix}$$

$$\mathbf{B}_w = \begin{pmatrix} 0 & 0 \\ \frac{\alpha_x}{m_x} & 0 \\ 0 & 0 \\ 0 & \frac{\alpha_y}{m_y} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{C}_w = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

We want to identify the parameters in the matrices \mathbf{A}_w and \mathbf{B}_w by using the MATLAB function `greyest`. It consists in the computation of the optimal physical parameters that minimizes a least-square criterion on the error between the measured output and the output of the state space model. However, in the actual form of the state-space representation, there is a problem of global identifiability due to the divisions by the masses m_x and m_y . We can tackle this problem by replacing all the corresponding parameters by the following normalized parameters

$$\begin{aligned} \bar{k}_{xx} &= k_{xx}/m_x & \bar{d}_{xx} &= d_{xx}/m_x \\ \bar{k}_{xy} &= k_{xy}/m_x & \bar{d}_{xy} &= d_{xy}/m_x \\ \bar{k}_{yy} &= k_{yy}/m_y & \bar{d}_{yy} &= d_{yy}/m_y \\ \bar{k}_{yx} &= k_{yx}/m_y & \bar{d}_{yx} &= d_{yx}/m_y \\ \bar{\alpha}_x &= \alpha_x/m_x & \bar{\alpha}_y &= \alpha_y/m_y \end{aligned}$$

In order to perform the grey-box identification, we have to choose the excitations u_x and u_y and the corresponding collected data. The first possibility is to use some experimental data collected on the true system \mathcal{S} by applying some excitation u_x and u_y . However, we just want to reduce the complexity of the model $\mathbf{G}(z, \tilde{\theta}_N)$ which is already a quite accurate estimate of the real dynamics. Therefore, we can directly simulate the data with this transfer function. This has the advantage to consider

noise-free data, allowing better grey-box modeling performances. Moreover, we are not anymore constrained by the input limitations for u_x and u_y (see Section 7.4.1). For the excitation, the input signals u_x and u_y will be independent white Gaussian noises.

Even if there is no noise, we can meet some local minimum problems since the grey-box identification is also based on a nonlinear optimization algorithm. Therefore, we should find a good initialization for this algorithm. The idea for the initialization that we will consider is based on the identified transfer function $\hat{\mathbf{G}}$. The method for the initialization is explained in Appendix F.3.

The Figures 7.10 and 7.11 give the Bode diagram magnitudes and phases respectively of the identified grey-box model of order 6 (that has been converted into a discrete-time model with the ZOH assumption) and the identified black-box model of order 12. Both Bode diagrams are really close for all transfer functions, suggesting that the grey-box model reduction performs well. However, the identified parameters λ_x and λ_y are close to $2\pi \times 7000$ rad/s and not to $2\pi \times 15000$ rad/s. Therefore, the effect of the electronic seems to not be explained by the anti-aliasing filter. Further works should be done to determine the nature of this effect.

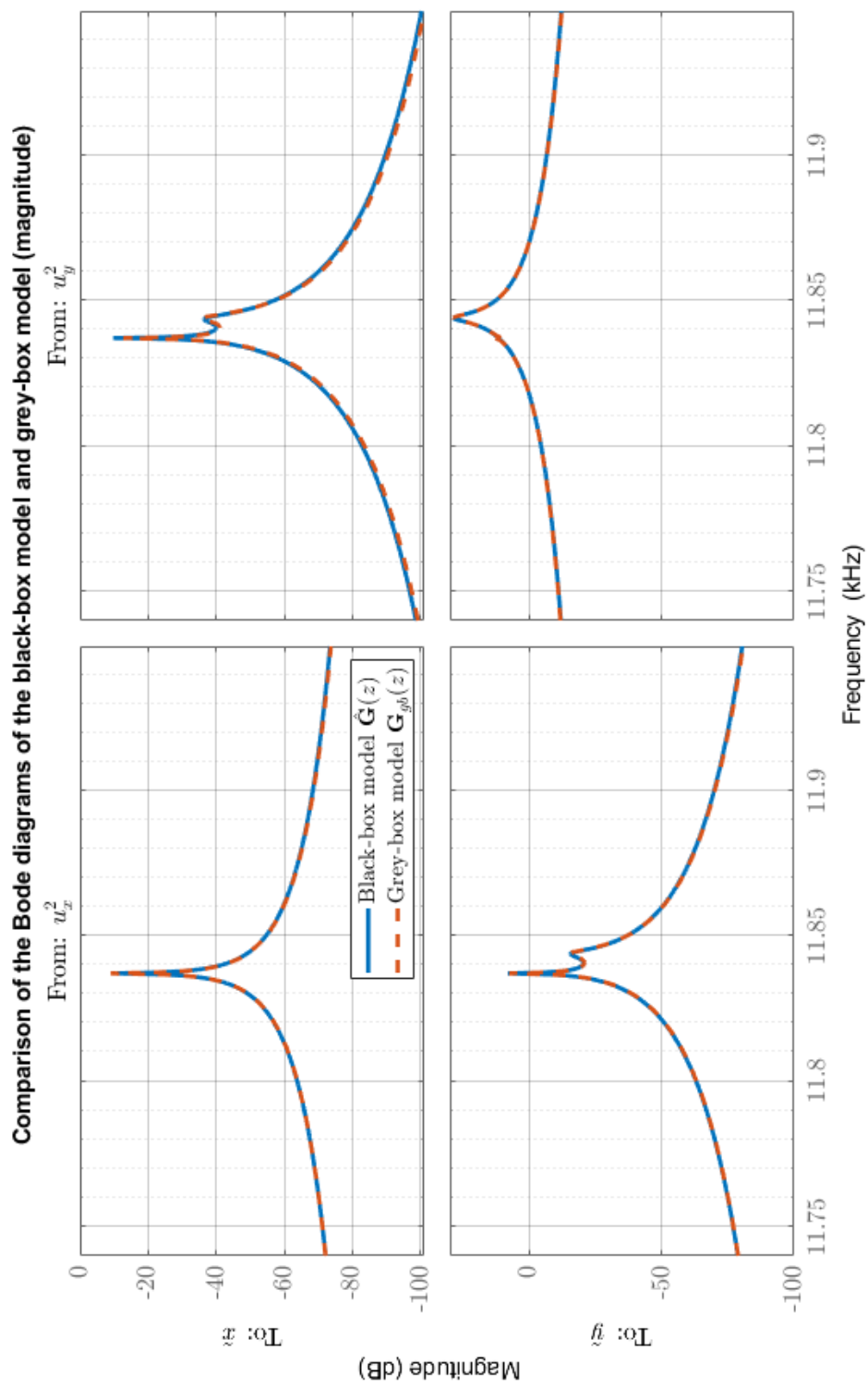


Figure 7.10: Comparison of Bode diagram magnitudes between the identified black-box model (blue solid line) and the identified grey-box model (orange dotted line).

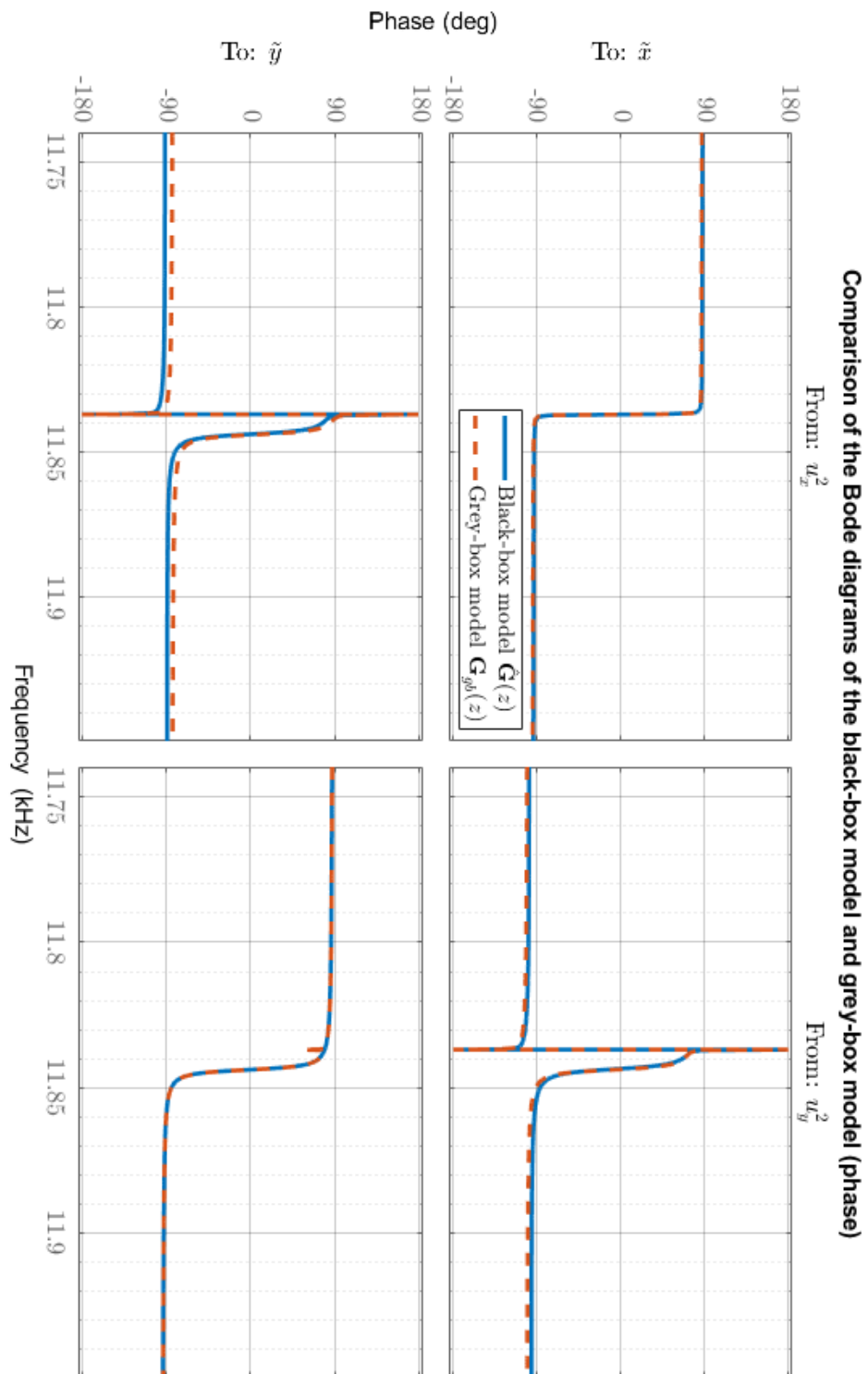


Figure 7.11: Comparison of Bode diagram phases between the identified black-box model (blue solid line) and the identified grey-box model (orange dotted line).

7.6.3 Controller design and validation

Based on the grey-box model obtained in the previous section, a \mathcal{H}_∞ controller of order 18 is designed in order to guarantee the control objectives given in Section 6.5. The controller is designed by Fabrício SAGGIN, another Ph.D. student participating in the Next4MEMS project funding this thesis. Its Bode diagram is given in Appendix F.4. In the next paragraphs, we are going to evaluate the performances of the MIMO controller and the modeling performances on one closed-loop experiment. For the reference signals, recall that the sense mode reference signal is equal to 0 and the one of the drive mode is a sinusoidal one. In order to avoid any control effort saturation, the drive reference signal x_{ref} is first modulated by a ramp signal during 2s until it reaches the reference amplitude of 0.05V.

Modeling performances

We are going to verify once more the quality of the computed model with a closed-loop experiment this time. For this purpose, as illustrated in Figure 7.12, we will compare the measured drive output \tilde{x} and the drive output \tilde{x}_{mod} simulated by the reduced-order model obtained in Section 7.6.2, put in closed-loop with the same controller $\mathbf{K}(z)$ and the same sinusoidal reference signal x_{ref} .

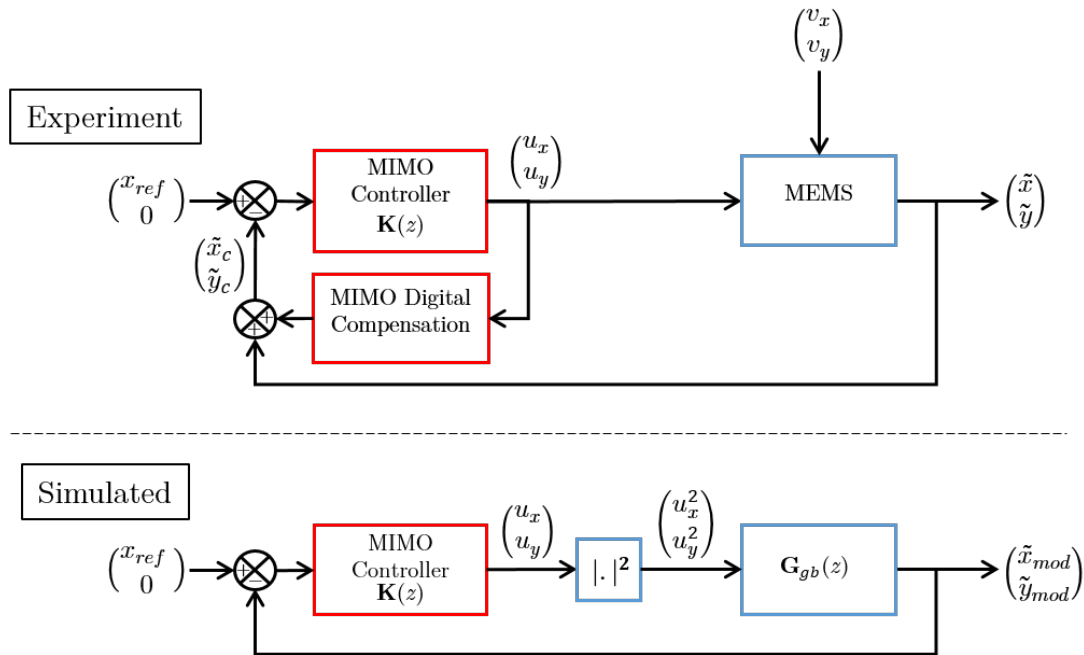


Figure 7.12: Block scheme for the comparison between the measured output \tilde{x} in closed-loop and the simulated one \tilde{x}_{mod} . The term \mathbf{G}_{gb} denotes the identified grey-box model.

To give a quantitative measure of the quality of the model for the drive channel, we compute the Best Fit which is defined in this case by

$$\text{Best Fit}_{cl} = 100\% \left(1 - \frac{\sqrt{\sum_{t=N_1}^{N_2} (\tilde{x}_c(t) - \tilde{x}_{mod}(t))^2}}{\sqrt{\sum_{t=N_1}^{N_2} (\tilde{x}_c(t) - \bar{x}_c)^2}} \right)$$

where \tilde{x}_c is the measured drive output after compensation of mean value \bar{x}_c over the $N_2 - N_1 + 1$ collected data and $\tilde{x}_{mod}(t)$ the drive mode output simulated from the grey-box model obtained in Section 7.6.2. For the computation of the Best Fit with the experimental data, we consider $N_1 = 6 \times 62500$ and $N_2 = 9 \times 62500$ (i.e., 3s of experiment from $t_c = 6$ s till $t_c = 9$ s). We obtain a Best Fit equal to 96.19% suggesting great modeling performances for the drive channel.

Controller performances

In this section, we verify that the controller guarantees the tracking objectives given in Section 6.5. In Figure 7.13, we compare in the time-domain the reference signal x_{ref} and the measured drive output \tilde{x}_c . The Figure 7.14 gives a zoom of Figure around $t_c = 6$ s. The drive output seems to track very well the reference signal x_{ref} .

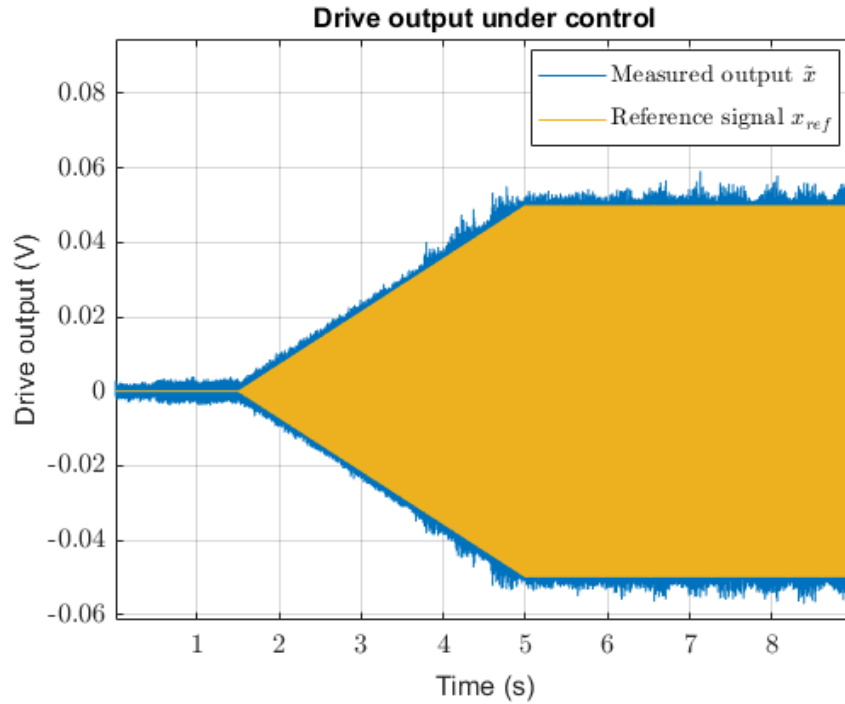
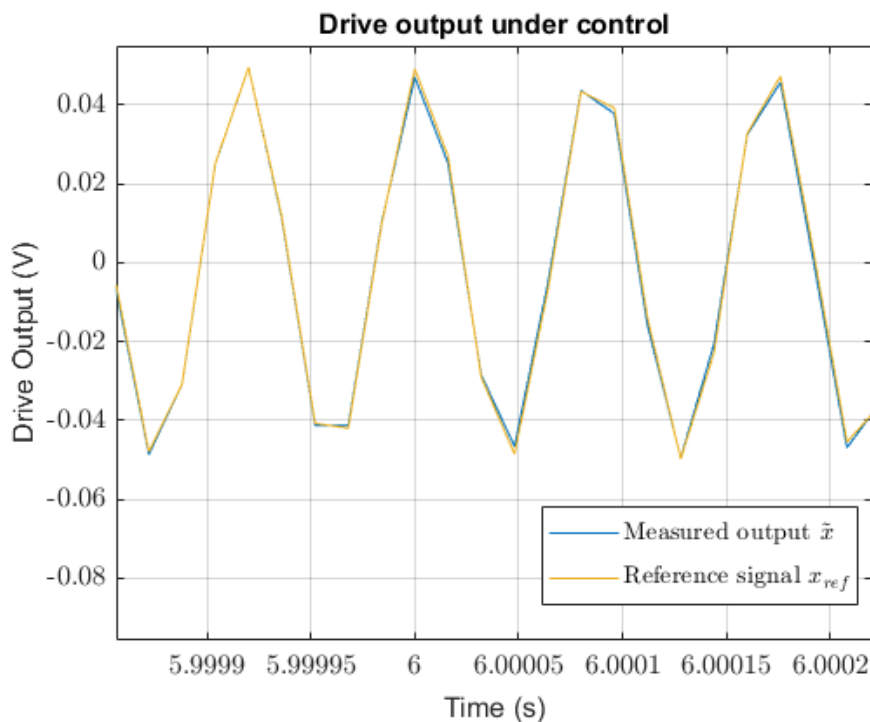


Figure 7.13: Comparison of the drive measured output \tilde{x} (blue) and the drive reference signal x_{ref} (yellow).

Figure 7.14: Zoom of Figure 7.13 around $t_c = 6$ s.

7.7 Conclusion

In this chapter, we have developed a method to identify the main dynamics of a MEMS gyroscope for the design of a controller with four separate MISO identifications, i.e., we have answered the second part of Research Question 4. However, an initialization was needed in order to avoid local minimum issue during these MISO identifications and we used some priors for this purpose. Then, a signal has been designed in order to increase the SNR and to provide the data informativity simultaneously so that the estimates are accurate. Almost all identified transfer functions are accurately identified and so we could design a controller based on these models.

As explained in Chapter 6, the controller has two degrees of freedom. For the block "Parasitic Compensation $\mathbf{K}_{comp}(z)$ " in Figure 6.10, we have seen that the parasitic capacitive coupling is well compensated by using the identified model $\hat{\mathbf{Q}}(z)$. For the block "Tracking Controller $\mathbf{K}(z)$ ", we had to reduce the complexity of the model $\hat{\mathbf{G}}(z)$ in order to avoid high order controller that cannot be implemented. For this purpose, we have considered a grey-box identification but where the data are simulated from the identified model $\hat{\mathbf{G}}(z)$. This allowed an important reduction of the model order with the same modeling performances and a controller has been designed with the \mathcal{H}_∞ design approach based on this reduced-order model. We have observed that the tracking objectives are well achieved for the drive mode.

It seems that Research Question 4 has been answered for the drive control objectives given in Chapter 6. Of course, there are still some experiments to do such as angular rate experiments in order to verify the performances of the designed controller for the accuracy of the angular rate deduction and for the sense mode control objectives.

In the next chapter, we give the conclusion of this dissertation and some leads for future works are provided.

Conclusion

Contents

8.1	Summary of the dissertation	153
8.2	Leads for future works	155
8.2.1	Restrictions of the considered MIMO model structures . . .	155
8.2.2	Conservatism of the developed conditions and possible inaccuracies for the rank computation	155
8.2.3	Data informativity for nonlinear systems identifiable within the Prediction Error framework	156
8.2.4	From the data informativity analysis to the least-costly experiment design	156
8.2.5	MEMS gyroscope: uncertainties, angular rate analysis and temperature-dependency	157

In this chapter, we give some concluding remarks on the research work presented in this dissertation. After that, some ideas for future works are suggested regarding the data informativity analysis and the identification of the MEMS gyroscope.

8.1 Summary of the dissertation

In this dissertation, we have first studied the data informativity w.r.t. MIMO linear model structures which are natural generalizations of classical SISO model structures: FIR, ARX, ARMAX, OE and BJ. We have considered two cases which are the open-loop identification (Research Question 1) and direct closed-loop identification (Research Question 2). In both cases, the types of excitation that have been considered are multi-sine and filtered white noise, which are two commonly-used types of excitations within the Prediction Error framework. This study was motivated by the lack of not-overly

conservative conditions for the verification of the data informativity for the identification of MIMO systems. We have seen in Chapters 3 and 4 that the data informativity can be verified by checking, for each channel, the right invertibility of a matrix depending on the model structure complexity, on the controller (for closed-loop experiments) and on the excitation parametrization. Even if this new condition comes with a dose of conservatism for most of the considered MIMO model structures, it is still less conservative than the ones developed in the literature, except¹ for the closed-loop costless identification of MISO ARMAX systems [YZ18]. Therefore, this new condition for the verification of the data informativity answers Research Questions 1 and 2.

However, the previous study only restricts attention to linear systems while most of real-life systems have some nonlinear behaviors that can be accurately identified within the Prediction Error framework. Therefore, in Chapter 5, we have studied the data informativity for a particular class of nonlinear systems given in (5.1) (Research Question 3). This particular study was motivated by the real-life system studied in this thesis: the MEMS gyroscope. Its main dynamics can be modeled within such structure. In order to answer Research Question 3, we have seen that the system (5.1) can be recast as a MISO linear system with 2 inputs: $u_1 = u$ and $u_2 = u^n$. Therefore, the developed results in Chapter 3 can be used to verify the data informativity for the identification of such nonlinear system. However, in the case where n is large and the excitation u is a multisine, the main result developed in Chapter 3 (persistency of a particular regressor) might need a high computational cost to verify the data informativity since we have to calculate all the amplitudes, the phase shifts and the frequencies of the multisine u^n . Therefore, additional results requiring less computation (but with more conservatism) than the ones derived in Chapter 3 have been developed in Chapter 5 to verify the data informativity in this case.

Finally, in the practical part of this dissertation, we consider the problem of the identification of the MEMS gyroscope. We have seen in Chapter 6 that it is important to design an optimal controller for the MEMS gyroscope in order to increase the accuracy of the angular rate estimation. Therefore, an accurate model of the main dynamics of the MEMS gyroscope is needed. However, the literature does not propose a complete model structure and an identification method to identify the three main dynamics of the MEMS gyroscope which are the mechanical coupling, the parasitic capacitive coupling and the measurement noise (Research Question 4). Therefore, in Chapter 6, we have developed a complete parametric model structure for the identification of these three phenomena (first part of Research Question 4). In Chapter 7, an identification method has been developed in order to estimate all the transfer functions in this model structure (second part of Research Question 4). Based on the identified model of the MEMS, a controller with two degrees of freedom has been designed by Fabrício SAGGIN, another Ph.D. student working in the Next4MEMS project. We could observe great performances for

¹However, the result derived in [YZ18] is valid for restrictive class of MISO ARMAX model structures and for controllers with particular coprimeness assumptions. The condition developed in this dissertation do not need to do these assumptions/restrictions.

the tracking objectives and the rejection of the parasite capacitive coupling. Therefore, for the control objectives considered in the thesis, our identification methodology seems to answer Research Question 4. However, we still have to analyze the quality of the angular rate estimation with this controller.

8.2 Leads for future works

In this section, we develop some leads for future works.

8.2.1 Restrictions of the considered MIMO model structures

In Chapters 3 and 4, we have considered MIMO systems for which the noise transfer function matrix $\mathbf{H}_0(z)$ has been assumed to be diagonal. This is the first limitation of the study presented in this dissertation. Fortunately, for the open-loop case and for FIR, ARX, OE and BJ model structures, we can still link the data informativity to the persistency of the same regressor vectors ϕ_u given in Theorem 3.3 even if $\mathbf{H}_0(z)$ is not assumed to be diagonal. This has been proven in [CBBM20b]. However, for ARMAX model structures, this diagonal assumption is necessary to derive the result in Appendix B.5.

We have also considered that all the numerators and denominators of the entries of $\mathbf{G}(z, \theta)$ and $\mathbf{H}(z, \theta)$ do not share common parameters (except for the parameters in η for the ARX and ARMAX model structures). However, for some real-life systems, the dynamics can be, e.g., modeled within a model structure for which the transfer functions have some common parameters. It is the case of the MEMS gyroscope since, by converting the grey-box state-space model into a transfer function matrix form (see Appendix F.3), we can observe that the transfer functions have some common parameters. Identifying a system within a full-order model structure with a common parametrization reduces the number of parameters which subsequently decreases the variance on the identified parameters. Moreover, a common parametrization often implies physical links between the transfer functions of the system which are also important to be identified. Therefore, studying the data informativity with common parameters is an interesting but also a challenging subject to work on in the future.

8.2.2 Conservatism of the developed conditions and possible inaccuracies for the rank computation

As explained in Chapters 3 and 4, we can at the moment conclude that the rank verification criterion developed in this dissertation is a sufficient condition for ARMAX, OE and BJ in open-loop and for ARX, ARMAX, OE and BJ in closed-loop. For instance, in the closed-loop case, we cannot easily conclude that the rank verification of $\mathbf{P}^{(c,i)}$ given in Theorem 4.4 is necessary since the sets \mathbf{D}_i in (4.13) do not cover the whole space \mathbb{R}^{ζ_i} except for FIR model structures for which we have $\mathbf{D}_i = \mathbb{R}^{\zeta_i}$. But we can still deduce a necessary and sufficient condition. As mentioned in Remark 4.3, in the closed-loop case with a nonzero external excitation ($\mathbf{r} \neq \mathbf{0}_{n_u \times 1}$), a necessary and

sufficient condition for the data informativity is that, for each channel $i = 1, \dots, n_y$, the intersection of the left null space of $\mathbf{P}^{(c,i)}$ and the set \mathbf{D}_i is equal to $\mathbf{0}$. However, this requires a geometrical study of the sets \mathbf{D}_i which will be done in future works. A similar reasoning applies for the open-loop case where the corresponding sets \mathbf{D}_i are discussed in the proof of Theorem 3.3.

The second main issue is linked to the fact that the new developed condition in this dissertation is based on a rank verification of a certain matrix that depends on the excitation parametrization, the controller (for closed-loop experiments) and the model structure complexity. However, as mentioned in Remark 3.5, a rank computation can be inaccurate. Further works should be done in order to tackle this issue.

8.2.3 Data informativity for nonlinear systems identifiable within the Prediction Error framework

In Chapter 5, we have considered a very particular type of parallel Hammerstein model structure given in (5.1). This choice for such restriction was motivated by the identification of the MEMS gyroscope whose main dynamics can be modeled within such model structures with $n = 2$. However, this particular nonlinear model structure is quite limiting for other real-life applications. We can then wonder how to generalize the data informativity study for more complex parallel Hammerstein model structures with different types of input static nonlinearities such as polynomial ones, dead-zone, saturation, etc. We can also consider other types of block-oriented parallel model structures such as Wiener, Hammerstein-Wiener, Wiener Hammerstein, etc [GB10, ST17].

8.2.4 From the data informativity analysis to the least-costly experiment design

As mentioned in Chapter 1, the data informativity is a crucial property in order to get an accurate model, *provided* that the number of data N can be increased. Indeed, the data informativity (combined with the global identifiability at $\boldsymbol{\theta}_0$) only guarantees the convergence of the identified parameter vector $\hat{\boldsymbol{\theta}}_N$ to the true parameter vector $\boldsymbol{\theta}_0$. However, for a fixed N , all excitations that yield informative data w.r.t. the model structure will not provide the same uncertainties on the identified parameters (i.e., the same covariance matrix which is a measure of the uncertainties). For some (unlucky) choices of excitations that yield informative data, we might need to largely increase the number of data N to guarantee a small level of uncertainties, which is not ideal in practice. That is why, in Chapter 7, we have designed the excitations for the identification of the MEMS gyroscope such that the data are informative and the signal-to-noise ration (SNR) is increased simultaneously in order to get small variances on the identified parameters for a reasonable number of data N .

Therefore, the next step after the data informativity study is naturally the choice of the optimal excitation that guarantees desired uncertainty levels: it is, e.g., the case of the least-costly experiment design framework where the optimal input corre-

sponds to the one with the minimal power satisfying some desired uncertainty constraints [BSG⁺06, BBHS08, JH05].

The study provided in this dissertation can help for the experiment design choice in order to do this optimization. Indeed, the data informativity provides the positive definiteness of the covariance matrix, which is a necessary condition to do the optimal input design. For instance, for MISO ARX model structure with multisine excitation in open-loop, we have seen that there is a necessary number of sinusoids that must be guaranteed for the data informativity. In other words, for the optimal design of the multisine excitation, it is necessary to satisfy this richness condition, otherwise the optimization will never give a solution for which the covariance matrix is positive definite. Future works should be done on how to connect the developed results of data informativity for the identification of MIMO systems with the least-costly experiment design framework.

8.2.5 MEMS gyroscope: uncertainties, angular rate analysis and temperature-dependency

For the MEMS gyroscope considered in this dissertation, a MIMO controller has been designed based on the identified model. The drive mode tracking objectives, presented in Chapter 6, have been evaluated and verified. However, in these control objectives, there is no specification on the angular rate estimation. In future works, it will be important to verify the quality of the angular rate deduction based on this MIMO controller. This type of experiments will also be useful to verify the sense mode tracking performances.

In Chapter 7, we have derived a model for the three main dynamics of the MEMS gyroscope. However, a model alone is not enough for the derivation of a controller. It is also important to take into account the uncertainties on the identified parameters. Since we have designed the multisine excitation such that the SNR is increased, we could expect small uncertainties. Fortunately, with Prediction Error, we can estimate the uncertainties² for a given probability α to verify this fact [BSG⁺06]. We then obtain a set of models (uncertain set) for which we know that the true model belongs to with a probability of α . A first study on the direct transfer functions G_{xx} , E_{xx} , G_{yy} and E_{yy} and with $\alpha = 0.99$ have shown that these uncertainties are indeed relatively small. However, further works should be dedicated to verify that the uncertainties on all the identified transfer functions are small enough to guarantee that the whole performances of the controller (drive mode *and* sense mode control objectives) are still valid for all models in the uncertainty set and so for the true (unknown) system. This study is the very essence of the scientific field called robustness analysis. In the Next4MEMS project, this part is studied by Jorge AYALA, one of the four Ph.D. students working in the Next4MEMS project.

As aforementioned, the MIMO controller has been designed based on the identifica-

²If the consistency of the estimator is provided.

tion of the three main phenomena involved in the MEMS gyroscope: the mechanical transfer, the parasitic capacitive coupling and the noise measurement. However, there are other dynamics that can deteriorate the performances of this MIMO controller. For instance, the dynamics of the mechanical transfer strongly depend on the temperature. More precisely, both resonances $\omega_{0,x}$ and $\omega_{0,y}$ vary when the temperature changes due to stiffness variations of the silicon-based micro-beams used to attach both masses of the MEMS (see Chapter 6). In Chapter 7, we have considered a fixed temperature (20°C) for the MEMS gyroscope and so fixed resonances $\omega_{0,x}$ and $\omega_{0,y}$. Therefore, this temperature-dependency did not raise any issue during the identification procedure presented in Chapter 7. However, the aim of the Next4MEMS project is to use these sensors for aerospace applications for which the temperature varies a lot. Moreover, as mentioned in Chapter 6, it is important that the frequency ω_{ref} of the reference signal x_{ref} be equal to $\omega_{0,x}$. To verify that the computed MIMO LTI controller can handle this temperature-dependency issue, a robustness analysis has been performed by Jorge AYALA for different values of $\omega_{0,x}$ and $\omega_{0,y}$ (i.e., for different temperatures). He has shown that the performances (such as the stability of the loop) are deteriorated with the MIMO LTI controller designed at the temperature 20°C (the one presented in Chapter 7) for the whole temperature range $[-40, 80]^{\circ}\text{C}$, which is in the set of specifications of the MEMS gyroscope considered in this project.

To tackle this problem, the possibility that is followed by Fabrício SAGGIN is to design a MIMO Linear-Parameter-Varying (LPV) controller whose varying parameter is $\omega_{0,x}$ and not the temperature since there is no thermometer in the considered electronic card of the MEMS gyroscope. For this purpose, we need two ingredients: a LPV model of the mechanical transfer for the synthesis of the MIMO LPV controller and a method to estimate the resonance frequency $\omega_{0,x}$ in order to adapt the controller and the frequency ω_{ref} of the reference signal x_{ref} (which must match $\omega_{0,x}$). The latter is studied by Federico MORELLI, another Ph.D. student participating in the Next4MEMS project. Two approaches are considered in his work: extremum seeking [ZO11] for the adaptation of the reference signal and recursive identification [LS83] for the online identification of $\omega_{0,x}$ in order to adapt ω_{ref} .

For the derivation of a LPV model for the mechanical transfer, one can consider the local approach [Töt10, GBH⁺18, LGPW14, MREL18] in order to use the developed method in Chapter 7. The idea of this approach is in two steps. First, we compute different grey-box reduced-order models by following the approach in Chapter 7 (i.e., first a black-box identification and secondly a grey-box reduction of the black-box model) for different temperatures³ and, for each model, we identify also the corresponding resonance frequency $\omega_{0,x}$. Secondly, we determine the function that describes the evolution of each parameter of the grey-box model with respect to the resonance frequency $\omega_{0,x}$ by doing, e.g., a polynomial interpolation for each parameter. This possibility will be studied in future works.

³For this purpose, the MEMS gyroscope is put in a thermal chamber in order to fix the temperature at the desired one.

Bibliography

- [AG82] Brian D.O. Anderson and Michel Gevers. Identifiability of linear stochastic systems operating under linear feedback. *Automatica*, 18(2):195–213, 1982. [doi:10.1016/0005-1098\(82\)90108-X](https://doi.org/10.1016/0005-1098(82)90108-X).
- [Apo16] Vladislav Apostolyuk. *Coriolis Vibratory Gyroscopes: Theory and Design*. Springer, 2016. [doi:10.1007/978-3-319-22198-4](https://doi.org/10.1007/978-3-319-22198-4).
- [AS05] Cenk Acar and Andrei M Shkel. An approach for Increasing Drive-Mode Bandwidth of MEMS Vibratory Gyroscopes. *Journal of Microelectromechanical Systems*, 14(3):520–528, 2005. [doi:10.1109/JMEMS.2005.844801](https://doi.org/10.1109/JMEMS.2005.844801).
- [BBG12] Alexandre S. Bazanella, Xavier Bombois, and Michel Gevers. Necessary and sufficient conditions for uniqueness of the minimum in prediction error identification. *Automatica*, 48(8):1621–1630, August 2012. [doi:10.1016/j.automatica.2012.06.018](https://doi.org/10.1016/j.automatica.2012.06.018).
- [BBHS08] Märta Barenthin, Xavier Bombois, Hakan Hjalmarsson, and Gérard Scroletti. Identification for control of multivariable systems: Controller validation and experiment design via LMIs. *Automatica*, 44(12):3070–3078, December 2008. [doi:10.1016/j.automatica.2008.05.022](https://doi.org/10.1016/j.automatica.2008.05.022).
- [BGM10] Alexandre Sanfelice Bazanella, Michel Gevers, and Ljubiša Mišković. Closed-Loop Identification of MIMO Systems: A New Look at Identifiability and Experiment Design. *European Journal of Control*, 16(3):228–239, 2010. [doi:10.3166/ejc.16.228-239](https://doi.org/10.3166/ejc.16.228-239).
- [BHY⁺02] Sam Y. Bae, Ken J. Hayworth, Karl Y. Yee, Kirill Shcheglov, and Dean V. Wiberg. High-performance MEMS microgyroscope. In Bernard Courtois, Jean Michel Karam, Karen W. Markus, Bernd Michel, Tamal Mukherjee, and James A. Walker, editors, *Design, Test, Integration, and Packaging of MEMS/MOEMS 2002*, volume 4755, pages 316 – 324. International Society for Optics and Photonics, SPIE, 2002. [doi:10.1117/12.462826](https://doi.org/10.1117/12.462826).

- [BSG⁺06] Xavier Bombois, Gérard Scorletti, Michel Gevers, Paul M.J. Van den Hof, and Roland Hildebrand. Least costly identification experiment for control. *Automatica*, 42(10):1651–1662, October 2006. doi:[10.1016/j.automatica.2006.05.016](https://doi.org/10.1016/j.automatica.2006.05.016).
- [CBBM20a] Kévin Colin, Xavier Bombois, Laurent Bako, and Federico Morelli. Closed-loop identification of mimo systems in the prediction error framework: Data informativity analysis. *Automatica*, 121:109171, 2020. URL: <http://www.sciencedirect.com/science/article/pii/S0005109820303691>, doi:[10.1016/j.automatica.2020.109171](https://doi.org/10.1016/j.automatica.2020.109171).
- [CBBM20b] Kévin Colin, Xavier Bombois, Laurent Bako, and Federico Morelli. Data informativity for the open-loop identification of MIMO systems in the prediction error framework. *Automatica*, 117, 2020. doi:[10.1016/j.automatica.2020.109000](https://doi.org/10.1016/j.automatica.2020.109000).
- [CHM03] Yen-Cheng Chen, Jason Hui, and Robert T. M'Closkey. Closed-loop identification of a micro-sensor. In *42nd IEEE International Conference on Decision and Control (CDC)*, pages 2632–2637, 2003. doi:[10.1109/CDC.2003.1273020](https://doi.org/10.1109/CDC.2003.1273020).
- [Din08] Steven X. Ding. *Model-based fault diagnosis techniques: design schemes, algorithms, and tools*. Springer Science & Business Media, 2008.
- [DTCA14] Francesco Dell'Olio, Teresa Tatoli, Caterina Ciminelli, and Mario N. Armenise. Recent advances in miniaturized optical gyroscopes. *Journal of the European optical society-Rapid publications*, 2014. doi:[10.2971/jeos.2014.14013](https://doi.org/10.2971/jeos.2014.14013).
- [EL] Eletronic-Lab. URL: <https://www.electronics-lab.com/gypro-high-performance-mems-gyroscopes>.
- [EMK12] Markus Egretzberger, Florian Mair, and Andreas Kugi. Model-based control concepts for vibratory MEMS gyroscopes. *Mechatronics*, 22(3):241–250, 2012. doi:[10.1016/j.mechatronics.2011.06.003](https://doi.org/10.1016/j.mechatronics.2011.06.003).
- [FB09a] Juntao Fei and Celal Batur. A novel adaptive sliding mode control with application to MEMS gyroscope. *ISA transactions*, 48(1):73–78, 2009. doi:[10.1016/j.isatra.2008.10.008](https://doi.org/10.1016/j.isatra.2008.10.008).
- [FB09b] Juntao Fei and Celal Batur. Robust adaptive control for a MEMS vibratory gyroscope. *The International Journal of Advanced Manufacturing Technology*, 42(3-4):293–300, 2009. doi:[10.1007/s00170-008-1591-5](https://doi.org/10.1007/s00170-008-1591-5).
- [Fei10] Juntao Fei. Robust adaptive vibration tracking control for a micro-electro-mechanical systems vibratory gyroscope with bound estimation. *IET Control Theory & Applications*, 4(6):1019–1026, 2010. doi:[10.1049/iet-cta.2008.0199](https://doi.org/10.1049/iet-cta.2008.0199).

- [FQ12] JianCheng Fang and Jie Qin. Advances in Atomic Gyroscopes: A View from Inertial Navigation Applications. *Sensors*, 12(5):6331–6346, 2012. doi:[10.3390/s120506331](https://doi.org/10.3390/s120506331).
- [FY11] Juntao Fei and Yuzheng Yang. System identification of MEMS vibratory gyroscope sensor. *Mathematical Problems in Engineering*, 2011. doi:[10.1155/2011/829432](https://doi.org/10.1155/2011/829432).
- [FY12] Juntao Fei and Yuzheng Yang. Comparative study of system identification approaches for adaptive tracking of MEMS gyroscope. *International Journal of Robotics and Automation*, 27(3):322, 2012. doi:[10.2316/Journal.206.2012.3.206-3693](https://doi.org/10.2316/Journal.206.2012.3.206-3693).
- [GB10] Fouad Giri and Er-Wei Bai. *Block-oriented Nonlinear System Identification*, volume 1. Springer, 2010.
- [GBBM09] Michel Gevers, Alexandre S. Bazanella, Xavier Bombois, and Ljubiša Miškovic. Identification and the Information Matrix: How to Get Just Sufficiently Rich? *IEEE Transactions on Automatic Control*, 54(12):2828–2840, December 2009. doi:[10.1109/TAC.2009.2034199](https://doi.org/10.1109/TAC.2009.2034199).
- [GBH⁺18] Debarghya Ghosh, Xavier Bombois, Julien Huillery, Gérard Scorletti, and Guillaume Mercère. Optimal identification experiment design for lpv systems using the local approach. *Automatica*, 87:258–266, 2018. doi:[10.1016/j.automatica.2017.10.013](https://doi.org/10.1016/j.automatica.2017.10.013).
- [GBM08] Michel Gevers, Alexandre S. Bazanella, and Ljubiša Miškovic. Informative data: How to get just sufficiently rich? In *2008 47th IEEE Conference on Decision and Control (CDC)*, pages 3425–3430. IEEE, December 2008. doi:[10.1109/CDC.2008.4738735](https://doi.org/10.1109/CDC.2008.4738735).
- [GLS77] Ivar Gustavsson, Lennart Ljung, and Torsten Söderström. Identification of processes in closed loop—identifiability and accuracy aspects. *Automatica*, 13(1):59–75, 1977. doi:[10.1016/0005-1098\(77\)90009-7](https://doi.org/10.1016/0005-1098(77)90009-7).
- [JH05] Henrik Jansson and Håkan Hjalmarsson. OPTIMAL EXPERIMENT DESIGN IN CLOSED LOOP. *IFAC Proceedings Volumes*, 38(1):488–493, 2005. doi:[10.3182/20050703-6-CZ-1902.00082](https://doi.org/10.3182/20050703-6-CZ-1902.00082).
- [Kai80] Thomas Kailath. *Linear Systems*, volume 156. Prentice-Hall Englewood Cliffs, NJ, 1980.
- [KBS⁺17] Anton Kornienko, Xavier Bombois, Gérard Scorletti, Thomas Dehaeze, and Roman Bohoslavets. Rapport d’Étude Bibliographique et d’Étude Déterminant les Fonctionnalités Devant Être Intégrées à la Plate-forme pour la Mise en Oeuvre des Algorithmes de Conception, 2017.
- [Kem11] Volker Kempe. *Inertial MEMS: principles and practice*. Cambridge University Press, 2011. doi:[10.1017/CB09780511933899](https://doi.org/10.1017/CB09780511933899).

- [Kha08] Mustafa Khammash. Reverse engineering: the architecture of biological networks. *Biotechniques*, 44(3):323–329, 2008. doi:[10.2144/000112772](https://doi.org/10.2144/000112772).
- [KKD11] John Kitching, Svenja Knappe, and Elizabeth A. Donley. Atomic sensors – A review. *IEEE Sensors Journal*, 11(9):1749–1758, 2011. doi:[10.1109/JSEN.2011.2157679](https://doi.org/10.1109/JSEN.2011.2157679).
- [LGPW14] Qin Liu, Joseph Gross, Sven Pfeiffer, and Herbert Werner. A local approach for the LPV identification of an actuated beam using piezoelectric actuators and sensors. *Mechatronics*, 24(4):289–297, 2014. doi:[10.1016/j.mechatronics.2013.12.012](https://doi.org/10.1016/j.mechatronics.2013.12.012).
- [Lju76a] Lennart Ljung. On consistency and identifiability. In *Stochastic Systems: Modeling, Identification and Optimization, I*, pages 169–190. Springer, 1976. doi:<https://doi.org/10.1007/BFb0120772>.
- [Lju76b] Lennart Ljung. On the consistency of prediction error identification methods. In Raman K. Mehra and Dimitri G. Lainiotis, editors, *System Identification Advances and Case Studies*, volume 126 of *Mathematics in Science and Engineering*, pages 121 – 164. Elsevier, 1976. URL: <http://www.sciencedirect.com/science/article/pii/S0076539208608711>, doi:[10.1016/S0076-5392\(08\)60871-1](https://doi.org/10.1016/S0076-5392(08)60871-1).
- [Lju99] Lennart Ljung. *System identification: Theory for the user*. Prentice Hall information and system sciences series. Prentice Hall PTR, Upper Saddle River (NJ), second edition edition, 1999.
- [LS83] Lennart Ljung and Torsten Söderström. *Theory and practice of recursive identification*. MIT press, 1983.
- [LWZ⁺17] Mingxiao J. Li, Ynjiun P. Wang, QC Zhao, Chunhua He, Liang T. Lin, Jianfeng F. Yang, and G.Z. Yan. Study on the influence induced by electrical coupling among interconnection lines in MEMS gyroscopes. In *2017 19th International Conference on Solid-State Sensors, Actuators and Microsystems (TRANSDUCERS)*, pages 1088–1091. IEEE, 2017. doi:[10.1109/TRANSDUCERS.2017.7994241](https://doi.org/10.1109/TRANSDUCERS.2017.7994241).
- [MGH01] Robert T. M’Closkey, Steve Gibson, and Jason Hui. System identification of a MEMS gyroscope. *Journal of Dynamic Systems, Measurement, and Control*, 123(2):201–210, 2001. doi:[10.1115/1.1369360](https://doi.org/10.1115/1.1369360).
- [MKBG08] Ljubisa Mišković, Alireza Karimi, Dominique Bonvin, and Michel Gevers. Closed-loop Identification of Multivariable Systems: With or Without Excitation of All References? *Automatica*, 44(8):2048–2056, 2008. doi:[10.1016/j.automatica.2007.11.016](https://doi.org/10.1016/j.automatica.2007.11.016).
- [MREL18] Koffi M.D. Motchon, Lala H. Rajaoarisoa, Lucien Etienne, and Stephane Lecoeuche. On experiment design for local approach identification of LPV

- systems. *IFAC-PapersOnLine*, 51(15):221–226, 2018. doi:10.1016/j.ifacol.2018.09.138.
- [MV99] Robert T. M’closkey and Alex Vakakis. Analysis of a microsensor automatic gain control loop. In *Proceedings of the 1999 American Control Conference (Cat. No. 99CH36251)*, volume 5, pages 3307–3311. IEEE, 1999. doi:10.1109/ACC.1999.782377.
- [NDG⁺04] Christoph Noeres, Kai Dadhe, Ralf Gesthuisen, Sebastian Engell, and Andrzej Górak. Model-based design, control and optimisation of catalytic distillation processes. *Chemical Engineering and Processing: Process Intensification*, 43(3):421–434, 2004. doi:10.1016/j.cep.2003.05.001.
- [PP02] Athanasios Papoulis and Unnikrishna Pillai. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Europe, 4th edition, 2002. URL: https://www.academia.edu/35834044/Probability_Random_Variables_and_Stochastic_Processes_Fourth_Edition_Papoulis.
- [PS02] Chris C. Painter and Andrei M. Shkel. Identification of anisoelectricity for electrostatic trimming of rate-integrating gyroscopes. In *Smart Structures and Materials 2002: Smart Electronics, MEMS, and Nanotechnology*, volume 4700, pages 157–168. International Society for Optics and Photonics, 2002. doi:10.1117/12.475027.
- [PS04] Abhilash S. Phani and Ashwin A. Seshia. Identification of anisoelectricity and nonproportional damping in MEMS gyroscopes. In *NSTI-Nanotech*, volume 2, pages 343–346, 2004. URL: https://www.researchgate.net/publication/255606825_Identification_of_Anisoelectricity_and_Nonproportional_Damping_in_MEMS_Gyroscopes.
- [PS12] Rik Pintelon and Johan Schoukens. *System identification: A Frequency Domain Approach*. John Wiley & Sons, 2012. doi:10.1002/9781118287422.
- [PTKH09] Sungsu Park, Chin-Woo Tan, Haedong Kim, and Sung Kyung Hong. Oscillation Control Algorithms for Resonant Sensors with Applications to Vibratory Gyroscopes. *Sensors*, 9(8):5952–5967, 2009. doi:10.3390/s90805952.
- [RA81] J.J. Roland and Govind P. Agrawal. Optical gyroscopes. *Optics & Laser Technology*, 13(5):239–244, 1981. doi:10.1016/0030-3992(81)90049-9.
- [RASB05] Tony Rogers, Nick Aitken, Kevin Stribley, and James Boyd. Improvements in mems gyroscope production as a result of using in situ, aligned, current-limited anodic bonding. *Sensors and Actuators A: Physical*, 123:106–110, 2005. doi:10.1016/j.sna.2005.03.009.

- [Sau08] Mikko Saukoski. *System and circuit design for a capacitive MEMS gyroscope*. Ph.D. dissertation, Helsinki University of Technology, Faculty of Electronics, Communications and Automation, Department of Micro and Nanosciences, 2008. URL: <https://aaltodoc.aalto.fi/handle/123456789/4413>.
- [SGL75] Torsten Söderström, Ivar Gustavsson, and Lennart Ljung. Identifiability conditions for linear systems operating in closed loop. *International Journal of control*, 21(2):243–255, 1975. doi:10.1080/00207177508921984.
- [SLG76] Torsten Söderström, Lennart Ljung, and Ivar Gustavsson. Identifiability conditions for linear multivariable systems operating under feedback. *IEEE Transactions on Automatic Control*, 21(6):837–840, 1976. doi:10.1109/TAC.1976.1101376.
- [SP05] Sigurd Skogestad and Ian Postlethwaite. *Multivariable Feedback Control: Analysis and Design*. John Wiley & Sons, 2005. URL: https://www.researchgate.net/publication/229058308_Multivariable_Feedback_Control_Analysis_and_Design.
- [SPR11] Maarten Schoukens, Rik Pintelon, and Yves Rolain. Parametric identification of parallel Hammerstein systems. *IEEE Transactions on Instrumentation and Measurement*, 60(12):3931–3938, 2011. doi:10.1109/TIM.2011.2138370.
- [SS89] Torsten Soderstrom and Petre Stoica. *System Identification*. Prentice Hall, 1989.
- [ST17] Maarten Schoukens and Koen Tiels. Identification of block-oriented nonlinear systems starting from linear approximations: A survey. *Automatica*, 85:272–292, 2017. doi:10.1016/j.automatica.2017.06.044.
- [TA05] Claire J. Tomlin and Jeffrey D. Axelrod. Understanding biology by reverse engineering the control. *Proceedings of the National Academy of Sciences*, 102(12):4219–4220, 2005. doi:10.1073/pnas.0500276102.
- [Tót10] Roland Tóth. *Modeling and identification of linear parameter-varying systems*, volume 403. Springer, 2010.
- [Vas01] Dimitris Vassiliadis. System identification, modeling, and prediction for space weather environments. *Plasma Science, IEEE Transactions on*, 28:1944 – 1955, 01 2001. doi:10.1109/27.902223.
- [VODM96] Peter Van Overschee and Bart L.R. De Moor. *Subspace identification for linear systems: Theory—Implementation—Applications*. Springer Science & Business Media, 1996. doi:10.1007/978-1-4613-0465-4.

- [YAN98] Navid Yazdi, Farrokh Ayazi, and Khalil Najafi. Micromachined inertial sensors. *Proceedings of the IEEE*, 86(8):1640–1659, 1998. doi:[10.1109/5.704269](https://doi.org/10.1109/5.704269).
- [YZ18] Wengang Yan and Yucai Zhu. Informative conditions for identification of miso armax model in closed-loop systems. *IFAC-PapersOnLine*, 51(15):455–460, 2018. 18th IFAC Symposium on System Identification SYSID 2018. doi:<https://doi.org/10.1016/j.ifacol.2018.09.187>.
- [ZDLG08] Qing Zheng, Lili Dong, Dae H. Lee, and Zhiqiang Gao. Active disturbance rejection control for mems gyroscopes. In *2008 American control conference*, pages 4425–4430. IEEE, 2008. doi:[10.1109/ACC.2008.4587191](https://doi.org/10.1109/ACC.2008.4587191).
- [ZO11] Chunlei Zhang and Raúl Ordóñez. *Extremum-seeking control and applications: a numerical optimization-based approach*. Springer Science & Business Media, 2011.

Appendices

Appendix for Chapter 2

Contents

A.1	Degeneration of the power spectrum matrix of multisine vectors . . .	169
-----	--	-----

A.1 Degeneration of the power spectrum matrix of multisine vectors

We are going to prove that the power spectrum matrix $\Phi_{\mathbf{u}}(\omega)$ of the input vector \mathbf{u} with $n_u \geq 2$ has a zero-determinant at all frequencies ω when the entries u_k of \mathbf{u} are multisine as given in (2.4). Since (2.4) is a general formulation to describe any multisine vector, we will prove that $\Phi_{\mathbf{u}}(\omega)$ degenerates at all frequencies ω for any multisine vector of dimension greater than or equal to 2.

First, the power spectrum matrix $\Phi_{\mathbf{u}}(\omega)$ is given by

$$\Phi_{\mathbf{u}}(\omega) = \begin{pmatrix} \Phi_{u_1}(\omega) & \Phi_{u_1 u_2}(\omega) & \cdots & \Phi_{u_1 u_{n_u}}(\omega) \\ \Phi_{u_2 u_1}(\omega) & \Phi_{u_2}(\omega) & \cdots & \Phi_{u_2 u_{n_u}}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{u_{n_u} u_1}(\omega) & \Phi_{u_{n_u} u_2}(\omega) & \cdots & \Phi_{u_{n_u}}(\omega) \end{pmatrix}$$

where $\Phi_{u_k u_{k'}}$ is the cross-power spectrum between u_k and $u_{k'}$ ($k = 1, \dots, n_u, k' = 1, \dots, n_u$) and it is denoted by Φ_{u_k} when $k = k'$. In order to prove that $\det(\Phi_{\mathbf{u}}(\omega)) = 0$ for all ω when \mathbf{u} is of the type (2.4), we are going to prove that the following matrix has a zero-determinant at all ω

$$\mathcal{A}(\omega) = \begin{pmatrix} \Phi_{u_1}(\omega) & \Phi_{u_1 u_2}(\omega) \\ \Phi_{u_2 u_1}(\omega) & \Phi_{u_2}(\omega) \end{pmatrix} \quad (\text{A.1})$$

Indeed, if we prove that $\mathcal{A}(\omega)$ has a zero-determinant at all frequencies ω , then $\Phi_{\mathbf{u}}(\omega)$ will have necessarily a zero-determinant at all frequencies ω since $\mathcal{A}(\omega)$ is a principal submatrix of $\Phi_{\mathbf{u}}(\omega)$.

To calculate $\Phi_{u_1}(\omega)$, $\Phi_{u_1 u_2}(\omega) = \Phi_{u_2 u_1}(-\omega)$ and $\Phi_{u_2}(\omega)$, we have to determine the Fourier transform of the correlation functions $R_{u_1}(\tau)$, $R_{u_1 u_2}(\tau)$ and $R_{u_2}(\tau)$ respectively which are defined by

$$\begin{aligned} R_{u_1}(\tau) &= \bar{E}[u_1(t)u_1(t-\tau)] \\ R_{u_1 u_2}(\tau) &= \bar{E}[u_1(t)u_2(t-\tau)] \\ R_{u_2}(\tau) &= \bar{E}[u_2(t)u_2(t-\tau)] \end{aligned}$$

For the multisine considered in (2.4), these correlations functions are equal to

$$\begin{aligned} R_{u_1}(\tau) &= \sum_{l=1}^s \frac{\lambda_{1l}^2}{2} \cos(\omega_l \tau) \\ R_{u_1 u_2}(\tau) &= \sum_{l=1}^s \frac{\lambda_{1l} \lambda_{2l}}{2} \cos(\omega_l \tau + \psi_{1l} - \psi_{2l}) \\ R_{u_2}(\tau) &= \sum_{l=1}^s \frac{\lambda_{2l}^2}{2} \cos(\omega_l \tau) \end{aligned}$$

Consequently, by taking the Fourier transform of the correlation functions, the matrix $\mathcal{A}(\omega)$ in (A.1) is given by

$$\begin{aligned} \mathcal{A}(\omega) &= \pi \sum_{l=1}^s \underbrace{\begin{pmatrix} \frac{\lambda_{1l}^2}{2} & \frac{\lambda_{1l} \lambda_{2l}}{2} e^{j(\psi_{1l} - \psi_{2l})} \\ \frac{\lambda_{1l} \lambda_{2l}}{2} e^{-j(\psi_{1l} - \psi_{2l})} & \frac{\lambda_{2l}^2}{2} \end{pmatrix}}_{\mathcal{P}_l} \delta(\omega - \omega_l) \\ &+ \pi \sum_{l=1}^s \underbrace{\begin{pmatrix} \frac{\lambda_{1l}^2}{2} & \frac{\lambda_{1l} \lambda_{2l}}{2} e^{-j(\psi_{1l} - \psi_{2l})} \\ \frac{\lambda_{1l} \lambda_{2l}}{2} e^{j(\psi_{1l} - \psi_{2l})} & \frac{\lambda_{2l}^2}{2} \end{pmatrix}}_{\mathcal{Q}_l} \delta(\omega + \omega_l) \end{aligned}$$

where the notation δ here refers to the Kronecker distribution.

First, for all the frequencies ω that are different from $\pm\omega_l$, $\mathcal{A}(\omega) = \mathbf{0}_{4 \times 4}$ and so its determinant is equal 0. When $\omega = \omega_l$, $\mathcal{A}(\omega_l) = \mathcal{P}_l$ which has a zero-determinant. The same holds when $\omega = -\omega_l$ with $\mathcal{A}(-\omega_l) = \mathcal{Q}_l$. Consequently, the matrix $\mathcal{A}(\omega)$ has a zero-determinant at all frequencies ω . However, as mentioned at the beginning of the proof, if $\mathcal{A}(\omega)$ has a zero-determinant at all frequencies ω , then it is also the case for

$\Phi_{\mathbf{u}}(\omega)$. Consequently, the power spectrum matrix $\Phi_{\mathbf{u}}(\omega)$ has a zero-determinant at all ω for any multisine vector with a dimension greater than or equal to 2, which concludes the proof. \square

Appendix for Chapter 3

Contents

B.1	Proof of Theorem 3.1	173
B.2	Proof of Lemma 2.1	175
B.3	Proof of Theorem 3.2	175
B.4	Proof of Theorem 3.3	176
B.5	Data informativity for MIMO ARMAX model structures	178
	B.5.1 Simplification of the study	178
	B.5.2 Notation for MISO ARMAX model structures	180
	B.5.3 Result	180
	B.5.4 Proof of Theorem B.3	180
B.6	Proof of the persistency of $\varphi^{(sin)}$	184
B.7	Proof of Lemma 3.5	185
B.8	Lemma 3.5 for the ARMAX model structure	186

B.1 Proof of Theorem 3.1

To prove Theorem 3.1, we will need the following lemma

Lemma B.1. *Consider a quasi-stationary input vector \mathbf{u} of the type (2.4)-(2.5). Consider also a nonzero stable and inversely stable matrix $\mathbf{V}(z)$ of transfer functions and another stable matrix $\mathbf{W}(z)$ of transfer functions such that $\mathbf{V}(z)\mathbf{W}(z)$ and $\mathbf{V}(z)\mathbf{W}(z)\mathbf{u}(t)$ are valid operations yielding a matrix of transfer functions and a vector of signals respectively. Then, the following equivalences hold*

$$\begin{aligned}
 (a) \quad & \bar{E} [\|\mathbf{V}(z)\mathbf{W}(z)\mathbf{u}(t)\|^2] = 0 \iff \bar{E} [\|\mathbf{W}(z)\mathbf{u}(t)\|^2] = 0 \\
 (b) \quad & \mathbf{V}(z)\mathbf{W}(z) \equiv \mathbf{0} \iff \mathbf{W}(z) \equiv \mathbf{0}
 \end{aligned}$$

Proof. First, since \mathbf{u} is of the type (2.4)-(2.5) and $\mathbf{V}(z)$ and $\mathbf{W}(z)$ are stable filter matrices, $\mathbf{V}(z)\mathbf{W}(z)\mathbf{u}(t)$ and $\mathbf{W}(z)\mathbf{u}(t)$ are quasi-stationary signals. Then, $\bar{E} [|\mathbf{W}(z)\mathbf{u}(t)|^2]$ and $\bar{E} [|\mathbf{V}(z)\mathbf{W}(z)\mathbf{u}(t)|^2]$ are well defined.

Let us begin with the first equivalence. Since $\mathbf{V}(z)$ is inversely stable, we have that the equation $\bar{E} [|\mathbf{V}(z)\mathbf{W}(z)\mathbf{u}(t)|^2] = 0$ leads to $\bar{E} [|\mathbf{V}^{-1}(z)\mathbf{V}(z)\mathbf{W}(z)\mathbf{u}(t)|^2] = \bar{E} [|\mathbf{W}(z)\mathbf{u}(t)|^2] = 0$. Now, since $\mathbf{V}(z)$ is stable, $E [|\mathbf{W}(z)\mathbf{u}(t)|^2] = 0$ implies that $\bar{E} [|\mathbf{V}(z)\mathbf{W}(z)\mathbf{u}(t)|^2] = 0$. This completes the proof of the first equivalence. We can follow a similar reasoning for the second equivalence, i.e., $\mathbf{V}(z)\mathbf{W}(z) \equiv \mathbf{0} \implies \mathbf{V}^{-1}(z)\mathbf{V}(z)\mathbf{W}(z) = \mathbf{W}(z) \equiv \mathbf{0} \implies \mathbf{V}(z)\mathbf{W}(z) \equiv \mathbf{0}$. \square

Let us now prove Theorem 3.1. By using (2.2) and (2.6), $\Delta\mathbf{W}(z)\mathbf{x}(t)$ in (2.12) can be rewritten as follows

$$\Delta\mathbf{W}(z)\mathbf{x}(t) = (\Delta\mathbf{W}_u(z) + \Delta\mathbf{W}_y(z)\mathbf{G}_0(z))\mathbf{u}(t) + \Delta\mathbf{W}_y(z)\mathbf{H}_0(z)\mathbf{e}(t)$$

where $\Delta\mathbf{W}_u(z)$ and $\Delta\mathbf{W}_y(z)$ are defined similarly as $\Delta\mathbf{W}(z)$.

Since \mathbf{u} and \mathbf{e} are assumed independent, $\bar{E} [|\Delta\mathbf{W}(z)\mathbf{x}(t)|^2] = 0$ is equivalent to

$$\begin{aligned} \bar{E} [|\Delta\mathbf{W}_u(z) + \Delta\mathbf{W}_y(z)\mathbf{G}_0(z)|^2] &= 0 \\ \bar{E} [|\Delta\mathbf{W}_y(z)\mathbf{H}_0(z)\mathbf{e}(t)|^2] &= 0 \end{aligned} \quad (\text{B.1})$$

Observe that, due to the whiteness of the vector \mathbf{e} and the invertibility of $\mathbf{H}_0(z)$, the spectrum $\Phi_{\tilde{\mathbf{v}}}(\omega)$ of $\tilde{\mathbf{v}}(t) = \mathbf{H}_0(z)\mathbf{e}(t)$ has the property that $\Phi_{\tilde{\mathbf{v}}}(\omega) \succ 0$ at all frequencies ω . By using Parseval's theorem, (B.1) is equivalent to

$$\text{tr} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \Delta\mathbf{W}_y(e^{j\omega}) \Phi_{\tilde{\mathbf{v}}}(\omega) \Delta\mathbf{W}_y^*(e^{j\omega}) d\omega \right) = 0$$

Since $\Phi_{\tilde{\mathbf{v}}}(\omega) \succ 0$ at all frequencies ω , the latter implies $\Delta\mathbf{W}_y(z) \equiv \mathbf{0}_{n_y \times n_y}$. Consequently, we can also say that $\bar{E} [|\Delta\mathbf{W}(z)\mathbf{x}(t)|^2] = 0$ is equivalent to the combination of the following two conditions:

$$\begin{aligned} \bar{E} [|\Delta\mathbf{W}_u(z)\mathbf{u}(t)|^2] &= 0 \\ \Delta\mathbf{W}_y &\equiv \mathbf{0}_{n_y \times n_y} \end{aligned} \quad (\text{B.2})$$

Due to the expression of $\mathbf{W}_y(z)$ in (2.7), (B.2) yields $\mathbf{H}(z, \boldsymbol{\theta}') = \mathbf{H}(z, \boldsymbol{\theta}'')$. Let us now deduce an expression for $\Delta\mathbf{W}_u$. In the BJ case, the fact that $\mathbf{H}(z, \boldsymbol{\theta}') = \mathbf{H}(z, \boldsymbol{\theta}'')$ implies that $\Delta\mathbf{W}_u$ can be rewritten as $\Delta\mathbf{W}_u(z) = \mathbf{H}^{-1}(z, \boldsymbol{\eta}')(\mathbf{G}(z, \tilde{\boldsymbol{\theta}}') - \mathbf{G}(z, \tilde{\boldsymbol{\theta}}''))$. In the ARX/FIR cases, $\Delta\mathbf{W}_u(z) = \mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - \mathbf{B}(z, \tilde{\boldsymbol{\theta}}'')$ and, in the OE case, $\Delta\mathbf{W}_u(z)$ is given by $\Delta\mathbf{W}_u(z) = \mathbf{G}(z, \tilde{\boldsymbol{\theta}}') - \mathbf{G}(z, \tilde{\boldsymbol{\theta}}'')$. Thus, in general, $\Delta\mathbf{W}_u(z)$ can be rewritten as $\Delta\mathbf{W}_u(z) = \mathbf{Y}^{-1}(z, \boldsymbol{\eta}')\Delta\mathbf{X}(z)$ with $\Delta\mathbf{X}(z)$ as defined in the statement of the theorem and $\mathbf{Y}(z, \boldsymbol{\eta}') = \mathbf{I}_{n_y}$ for ARX/OE/FIR model structures and $\mathbf{Y}(z, \boldsymbol{\eta}') = \mathbf{H}(z, \boldsymbol{\eta}')$ for BJ model structures.

Since we restrict attention to $\boldsymbol{\theta}' = (\tilde{\boldsymbol{\theta}}'^T, \boldsymbol{\eta}'^T)^T \in \mathcal{D}_\theta$, $\mathbf{Y}(z, \boldsymbol{\eta}')$ is a (nonzero) stable and inversely-stable matrix of transfer functions in both situations. Consequently, by using Lemma B.1, $\bar{E} [|\Delta \mathbf{W}_u(z) \mathbf{u}(t)|^2] = 0$ is equivalent to $\bar{E} [|\Delta \mathbf{X}(z) \mathbf{u}(t)|^2] = 0$.

From the reasoning above, for all $\Delta \mathbf{W}(z) \in \boldsymbol{\Delta}_\mathbf{W}$ (with $\boldsymbol{\Delta}_\mathbf{W}$ as given in (2.11)), the left-hand side of (2.12) is equivalent to $\bar{E} [|\Delta \mathbf{X}(z) \mathbf{u}(t)|^2] = 0$. From the reasoning above, we can also conclude that, for all $\Delta \mathbf{W}(z) \in \boldsymbol{\Delta}_\mathbf{W}$, the left-hand side of (2.12) always implies $\Delta \mathbf{W}_y(z) \equiv \mathbf{0}_{n_y \times n_y}$. Consequently, the right-hand side of (2.12) can be restricted to $\Delta \mathbf{W}_u(z) \equiv \mathbf{0}_{n_y \times n_u}$ or, equivalently, via Lemma B.1, to $\Delta \mathbf{X}(z) \equiv \mathbf{0}_{n_y \times n_u}$. In other words, the condition (3.2) in Theorem 3.1 (for all $\Delta \mathbf{X}(z) \in \boldsymbol{\Delta}_\mathbf{X}$) is equivalent to the condition (2.12) (for all $\Delta \mathbf{W}(z) \in \boldsymbol{\Delta}_\mathbf{W}$). This completes thus the proof. \square

B.2 Proof of Lemma 2.1

From Parseval's theorem, the term $\bar{E} [|\Delta \mathbf{X}(z) \mathbf{u}(t)|^2]$ in (3.2) is equal to

$$\text{tr} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \Delta \mathbf{X}(e^{j\omega}) \boldsymbol{\Phi}_u(\omega) \Delta \mathbf{X}^*(e^{j\omega}) d\omega \right)$$

Consequently, if $\boldsymbol{\Phi}_u(\omega) \succ 0$ at almost all frequencies ω , it is clear that (3.2) will hold for all $\Delta \mathbf{X}(z) \in \boldsymbol{\Delta}_\mathbf{X}$, irrespectively of the model structure complexity. \square

B.3 Proof of Theorem 3.2

From Theorem 3.1, the data are informative if the property (3.2) holds for all $\Delta \mathbf{X}(z) \in \boldsymbol{\Delta}_\mathbf{X}$ with $\boldsymbol{\Delta}_\mathbf{X}$ as defined in (3.1). We are going to prove that the latter is equivalent to the fact that (3.3) holds for all $\Delta \mathbf{X}_{i:}(z) \in \boldsymbol{\Delta}_{\mathbf{X},i}$ for all $i = 1, \dots, n_y$ with $\boldsymbol{\Delta}_{\mathbf{X},i}$ as defined in the theorem statement.

Let us first observe that $\Delta \mathbf{X}(z) \equiv \mathbf{0}_{n_y \times n_u}$ is equivalent to $\Delta \mathbf{X}_{i:}(z) \equiv \mathbf{0}_{1 \times n_u}$ ($i = 1, \dots, n_y$). Secondly, let us prove that

$$\bar{E} [|\Delta \mathbf{X}(z) \mathbf{u}(t)|^2] = 0 \Leftrightarrow \bar{E} [|\Delta \mathbf{X}_{i:}(z) \mathbf{u}(t)|^2] = 0 \quad (i = 1, \dots, n_y)$$

For this purpose, let us rewrite the term $\bar{E} [|\Delta \mathbf{X}(z) \mathbf{u}(t)|^2]$ as follows:

$$\bar{E} [|\Delta \mathbf{X}(z) \mathbf{u}(t)|^2] = \sum_{i=1}^{n_y} \bar{E} [|\Delta \mathbf{X}_{i:}(z) \mathbf{u}(t)|^2]$$

with $\Delta \mathbf{X}_{i:}(z)$ as defined in the theorem. Since the terms $\bar{E} [|\Delta \mathbf{X}_{i:}(z) \mathbf{u}(t)|^2]$ are non-negative ($i = 1, \dots, n_y$), the equation $\bar{E} [|\Delta \mathbf{X}(z) \mathbf{u}(t)|^2] = 0$ is equivalent to the n_y equations $\bar{E} [|\Delta \mathbf{X}_{i:}(z) \mathbf{u}(t)|^2] = 0$ ($i = 1, \dots, n_y$).

Therefore, we have shown that the condition (3.2) is equivalent to

$$\bar{E} [|\Delta \mathbf{X}_{i:}(z) \mathbf{u}(t)|^2] = 0 \quad (i = 1, \dots, n_y) \implies \Delta \mathbf{X}_{i:}(z) \equiv \mathbf{0}_{1 \times n_u} \quad (i = 1, \dots, n_y)$$

Noting that there are no common parameters in $\Delta\mathbf{X}_{i:}(z)$ and $\Delta\mathbf{X}_{i':}(z)$ ($i' \neq i$), the fact that (3.2) holds for all $\Delta\mathbf{X}(z) \in \mathbf{\Delta}_\mathbf{X}$ is therefore equivalent to the fact that (3.3) holds for all $\Delta\mathbf{X}_{i:} \in \mathbf{\Delta}_{\mathbf{X},i}$ and for all $i = 1, \dots, n_y$. This completes thus the proof. \square

B.4 Proof of Theorem 3.3

First note that Theorem 3.1 also applies to the case $n_y = 1$ and let us distinguish the ARX/FIR case and the BJ/OE case in the sequel.

Let us first consider ARX/FIR model structures and let us first observe that, using the notations of the model structures in the MISO case (see Section 3.3.1) and the fact that $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\theta}}_1^T, \tilde{\boldsymbol{\theta}}_2^T, \dots, \tilde{\boldsymbol{\theta}}_{n_u}^T)^T$, we have that $B_k(z, \tilde{\boldsymbol{\theta}}_k)u_k(t) = z^{-\rho_k} \mathring{B}_k(z, \tilde{\boldsymbol{\theta}}_k)u_k(t) = \boldsymbol{\phi}_{u_k}(t)^T \tilde{\boldsymbol{\theta}}_k$ and $\mathbf{B}(z, \tilde{\boldsymbol{\theta}})\mathbf{u}(t) = \boldsymbol{\phi}_u^T(t)\tilde{\boldsymbol{\theta}}$ with $\boldsymbol{\phi}_{u_k}$ and $\boldsymbol{\phi}_u$ as defined in (3.5). Let us now rewrite the left-hand side of condition (3.2), i.e., $\bar{E} [|\Delta\mathbf{X}(z)\mathbf{u}(t)|^2] = 0$. For ARX/FIR model structures, we have that $\Delta\mathbf{X}(z) = \Delta\mathbf{B}(z)$ where $\Delta\mathbf{B}(z)$ is a row vector of n_u transfer functions whose entries $\Delta B_k(z)$ are polynomials defined as $\Delta B_k(z) = z^{-\rho_k} (\mathring{B}_k(z, \tilde{\boldsymbol{\theta}}'_k) - \mathring{B}_k(z, \tilde{\boldsymbol{\theta}}''_k))$. We have thus that

$$\begin{aligned} \Delta B_k(z)u_k(t) &= \boldsymbol{\phi}_{u_k}(t)^T (\tilde{\boldsymbol{\theta}}'_k - \tilde{\boldsymbol{\theta}}''_k) \\ \Delta\mathbf{B}(z)\mathbf{u}(t) &= \boldsymbol{\phi}_u(t)^T (\tilde{\boldsymbol{\theta}}' - \tilde{\boldsymbol{\theta}}'') \end{aligned}$$

Consequently, $\bar{E} [|\Delta\mathbf{B}(z)\mathbf{u}(t)|^2] = 0$ is equivalent to

$$(\tilde{\boldsymbol{\theta}}' - \tilde{\boldsymbol{\theta}}'')^T \bar{E} [\boldsymbol{\phi}_u(t)\boldsymbol{\phi}_u^T(t)] (\tilde{\boldsymbol{\theta}}' - \tilde{\boldsymbol{\theta}}'') = 0 \quad (\text{B.3})$$

Let us now observe that the right-hand side of (3.2) is here equivalent to $\tilde{\boldsymbol{\theta}}' = \tilde{\boldsymbol{\theta}}''$. Recall also that $\mathcal{D}_{\tilde{\boldsymbol{\theta}}} = \mathbb{R}^{\tilde{n}}$ and that, here, the dimension \tilde{n} of $\tilde{\boldsymbol{\theta}}$ is equal to the dimension ν of $\boldsymbol{\phi}_u$. Using the above reasoning and Theorem 3.1, we will thus have data informativity if and only if, for any $\tilde{\boldsymbol{\theta}}'$ and $\tilde{\boldsymbol{\theta}}''$ in $\mathbb{R}^{\tilde{n}}$, (B.3) implies $\tilde{\boldsymbol{\theta}}' = \tilde{\boldsymbol{\theta}}''$. This latter condition is equivalent to $\bar{E} [\boldsymbol{\phi}_u(t)\boldsymbol{\phi}_u^T(t)] \succ 0$ or that $\boldsymbol{\phi}_u$ is PE (Definition 2.4). This completes thus the proof for the ARX/FIR model structures.

Let us now consider BJ/OE model structures and let us here also rewrite the left-hand side of condition (3.2), i.e., $\bar{E} [|\Delta\mathbf{G}(z)\mathbf{u}(t)|^2] = 0$. For this purpose, we can observe that the n_u entries $\Delta G_k(z)$ of the row vector $\Delta\mathbf{G}(z) = (\Delta G_k(z))_{k \in [1, n_u]}$ are given by

$$\Delta G_k(z) = z^{-\rho_k} \left(\frac{\mathring{B}_k(z, \tilde{\boldsymbol{\theta}}'_k)}{F_k(z, \tilde{\boldsymbol{\theta}}'_k)} - \frac{\mathring{B}_k(z, \tilde{\boldsymbol{\theta}}''_k)}{F_k(z, \tilde{\boldsymbol{\theta}}''_k)} \right)$$

Using obvious short-hand notations, the latter equation is rewritten as:

$$\Delta G_k(z) = z^{-\rho_k} \left(\frac{\mathring{B}'_k}{F'_k} - \frac{\mathring{B}''_k}{F''_k} \right)$$

Let us now put all entries of $\Delta \mathbf{G}(z)$ on the same denominator which will be denoted by $U(z)$, i.e., $U(z) = \prod_{k=1}^{n_u} F'_k F''_k$. Consequently, we have that $\Delta \mathbf{G}(z) = \frac{\mathbf{V}(z)}{U(z)}$ where $\mathbf{V}(z) = (V_k)_{k \in \llbracket 1, n_u \rrbracket}$ is a vector of n_u polynomials $V_k(z)$ having the following expression:

$$V_k(z) = z^{-\rho_k} (\mathring{B}'_k F''_k - \mathring{B}''_k F'_k) \prod_{m=1, m \neq k}^{n_u} F'_m F''_m$$

Each $V_k(z)$ is thus a polynomial of delay ρ_k and of degree μ_k (with μ_k as defined in the statement of the theorem) and can thus be rewritten as follows:

$$V_k(z) = \sum_{h=\rho_k}^{\mu_k} \delta_k^{(h)} z^{-h}$$

where the coefficients $\delta_k^{(h)}$ are known functions of $\tilde{\theta}'$ and $\tilde{\theta}''$. Combining the above elements successively yields:

$$\begin{aligned} \Delta \mathbf{G}(z) \mathbf{u}(t) &= \frac{1}{U(z)} \left(\sum_{k=1}^{n_u} \sum_{h=\rho_k}^{\mu_k} \delta_k^{(h)} u_k(t-h) \right) \\ \Leftrightarrow \Delta \mathbf{G}(z) \mathbf{u}(t) &= \frac{1}{U(z)} \underbrace{(\phi_{\mathbf{u}}(t)^T \boldsymbol{\delta})}_{=\boldsymbol{\varpi}(t)} \end{aligned}$$

with $\boldsymbol{\delta}$ the vector of dimension ν consisting of the concatenation of all $\delta_k^{(h)}$ and $\phi_{\mathbf{u}}$ as given in the statement of the theorem.

Since $\tilde{\theta}'$ and $\tilde{\theta}''$ are in $\mathcal{D}_{\tilde{\theta}} = \{\tilde{\theta} \mid \mathbf{G}(z, \tilde{\theta}) \text{ is stable}\}$, $U(z)$ is a stable and inversely stable filter. Consequently, via Lemma B.1 in Appendix B.1, the left-hand side of (3.2) (i.e. $\bar{E} [|\Delta \mathbf{G}(z) \mathbf{u}(t)|^2] = 0$) is equivalent to

$$\bar{E} [\boldsymbol{\varpi}^2(t)] = \boldsymbol{\delta}^T \bar{E} [\phi_{\mathbf{u}}(t) \phi_{\mathbf{u}}^T(t)] \boldsymbol{\delta} = 0 \quad (\text{B.4})$$

When $\phi_{\mathbf{u}}$ is PE, i.e., when $\bar{E} [\phi_{\mathbf{u}}(t) \phi_{\mathbf{u}}^T(t)] \succ 0$ (Lemma 2.3), we have that, for any $\boldsymbol{\delta}$ generated as above with $\tilde{\theta}'$ and $\tilde{\theta}'' \in \mathcal{D}_{\tilde{\theta}}$, (B.4) implies $\boldsymbol{\delta} = \mathbf{0}$ which in turn implies that $\mathbf{V}(z) \equiv \mathbf{0}_{1 \times n_u}$ and thus that $\Delta \mathbf{G}(z) \equiv \mathbf{0}_{1 \times n_u}$ (i.e. the right-hand side of (3.2)). The sufficient condition for the BJ/OE case in this theorem is then a consequence of Theorem 3.1. Note that this condition is not necessary in the BJ/OE case since the vectors $\boldsymbol{\delta}$ that are generated by all $\tilde{\theta}'$ and $\tilde{\theta}'' \in \mathcal{D}_{\tilde{\theta}}$ will not cover the whole vectorial space \mathbb{R}^{ν} . Observe indeed that the dimension ν of $\boldsymbol{\delta}$ is (much) larger than the dimension of $\tilde{\theta}'$ and $\tilde{\theta}''$. \square

B.5 Data informativity for MIMO ARMAX model structures

B.5.1 Simplification of the study

In this section, we derive the results for the data informativity w.r.t. MIMO ARMAX model structures (see Section 2.5). In the next two theorems, we will see that we can simplify the study of the data informativity w.r.t. MIMO ARMAX model structures into the analysis of the data informativity w.r.t. MISO ARMAX model structures as we did for the other model structures in Theorems 3.1 and 3.2.

Theorem B.1. *Consider Definition 2.2 and consider the notations of Section 2.5 for MIMO ARMAX model structures. Based on the set $\Delta_{\mathbf{W}}$ (see (2.11) in Definition 2.3), define the set*

$$\Delta_{\mathbf{W}_0} = \{\Delta \mathbf{W}(z) = (\Delta \mathbf{W}_y(z), \Delta \mathbf{W}_u(z)) \in \Delta_{\mathbf{W}} \mid \Delta \mathbf{W}_y(z) \equiv \mathbf{0}_{n_y \times n_y}\} \quad (\text{B.5})$$

Then, the data set \mathcal{Z}^∞ is informative w.r.t. the MIMO ARMAX model structure \mathcal{M} if and only if the following property holds for all $(\Delta \mathbf{W}_y(z), \Delta \mathbf{W}_u(z)) \in \Delta_{\mathbf{W}_0}$:

$$\bar{E} [|\Delta \mathbf{W}_u(z) \mathbf{u}(t)|^2] = 0 \implies \Delta \mathbf{W}_u(z) \equiv \mathbf{0}_{n_y \times n_u} \quad (\text{B.6})$$

Proof. From Definition 2.2, the data are informative if the property (2.12) holds for all $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}}$ with $\Delta_{\mathbf{W}}$ as defined in (2.11). We are going to prove that the latter is equivalent to the fact that (B.6) holds for all $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}_0}$ with $\Delta_{\mathbf{W}_0}$ as defined in the statement of the theorem.

First, we have shown in the proof of Theorem 3.1, that the left-hand side of (2.12) is equivalent to

$$\begin{aligned} \bar{E} [|\Delta \mathbf{W}_u(z) \mathbf{u}(t)|^2] &= 0 \\ \Delta \mathbf{W}_y(z) &\equiv \mathbf{0}_{n_y \times n_y} \end{aligned}$$

Consequently, the property (2.12) holds for all $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}}$ if and only if the following property holds for all $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}}$

$$\begin{aligned} \bar{E} [|\Delta \mathbf{W}_u(z) \mathbf{u}(t)|^2] &= 0 \\ \Delta \mathbf{W}_y(z) &\equiv \mathbf{0}_{n_y \times n_y} \end{aligned} \implies (\Delta \mathbf{W}_y(z), \Delta \mathbf{W}_u(z)) \equiv (\mathbf{0}_{n_y \times n_y}, \mathbf{0}_{n_y \times n_u})$$

The latter is equivalent to the fact that (B.6) holds for all $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}_0}$, which concludes the proof. \square

Theorem B.2. *Consider Definition 2.2 and Theorem B.1 and consider the notations of Section 2.5 for MIMO ARMAX model structures. Based on the set $\Delta_{\mathbf{W}_0}$ (see (B.5) in Theorem B.1), define, for all $i = 1, \dots, n_y$, the sets*

$$\Delta_{\mathbf{W}_{0,i}} = \{\Delta \mathbf{W}_{i,:}(z) \mid \Delta \mathbf{W}_{i,:}(z) \text{ is the } i^{\text{th}} \text{ row of } \Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}_0}\}$$

We denote similarly by $\Delta \mathbf{W}_{y,i}(z)$ and $\Delta \mathbf{W}_{u,i}(z)$ the i^{th} row of $\Delta \mathbf{W}_y(z)$ and $\Delta \mathbf{W}_u(z)$ respectively. Then, the data set \mathcal{Z}^∞ is informative w.r.t. the MIMO ARMAX model structure \mathcal{M} if and only if, for all $i = 1, \dots, n_y$, the following property holds for all $(\Delta \mathbf{W}_{y,i}(z), \Delta \mathbf{W}_{u,i}(z)) \in \Delta \mathbf{W}_{\mathbf{0},i}$:

$$\bar{E} [|\Delta \mathbf{W}_{u,i}(z) \mathbf{u}(t)|^2] = 0 \implies \Delta \mathbf{W}_{u,i}(z) \equiv \mathbf{0}_{1 \times n_u} \quad (\text{B.7})$$

Proof. From Theorem B.1, the data are informative if the property (B.6) holds for all $\Delta \mathbf{W}(z) \in \Delta \mathbf{W}_{\mathbf{0}}$ with $\Delta \mathbf{W}_{\mathbf{0}}$ as defined in (B.5). We are going to prove that the latter is equivalent to the fact that (B.7) holds for all $\Delta \mathbf{W}_i(z) \in \Delta \mathbf{W}_{\mathbf{0},i}$ for all $i = 1, \dots, n_y$ with $\Delta \mathbf{W}_{\mathbf{0},i}$ as defined in the theorem statement.

First, it is straightforward to prove that the right-hand side of (B.6) is equivalent to $\Delta \mathbf{W}_{u,i}(z) \equiv \mathbf{0}_{1 \times n_u} \quad \forall i \in \llbracket 1, n_y \rrbracket$.

Let us now focus on the left-hand side of (B.6). It is successively equivalent to

$$\begin{aligned} & \sum_{i=1}^{n_y} \bar{E} [|\Delta \mathbf{W}_{u,i}(z) \mathbf{u}(t)|^2] = 0 \\ \Leftrightarrow & \quad \bar{E} [|\Delta \mathbf{W}_{u,i}(z) \mathbf{u}(t)|^2] = 0 \quad \forall i \in \llbracket 1, n_y \rrbracket \end{aligned}$$

Consequently, the property (B.6) is equivalent to

$$\bar{E} [|\Delta \mathbf{W}_{u,i}(z) \mathbf{u}(t)|^2] = 0 \quad \forall i \in \llbracket 1, n_y \rrbracket \implies \Delta \mathbf{W}_{u,i}(z) \equiv \mathbf{0}_{1 \times n_u} \quad \forall i \in \llbracket 1, n_y \rrbracket$$

Let us observe that $\Delta \mathbf{W}_{u,i}(z) = C_i^{-1'} \mathbf{B}'_i - C_i^{-1''} \mathbf{B}''_i$ where both notations \mathbf{B} and C_i are defined in Section 2.5. Noting that there are no common parameters in $\Delta \mathbf{W}_{u,i}(z)$ and $\Delta \mathbf{W}_{u,i'}(z)$ ($i' \neq i$), the fact that (B.6) holds for all $\Delta \mathbf{W}(z) \in \Delta \mathbf{W}_{\mathbf{0}}$ is therefore equivalent to the fact that (B.7) holds for all $\Delta \mathbf{W}_i(z) \in \Delta \mathbf{W}_{\mathbf{0},i}$ and for all $i = 1, \dots, n_y$. This completes thus the proof. \square

Theorem B.1 is important since it allows one to check whether an input vector \mathbf{u} yields informative data for the identification of a MIMO ARMAX system \mathcal{S} (see (2.2)) by successively checking that this input vector yields informative data for the identification of the n_y MISO ARMAX systems \mathcal{S}_i in (2.3) (see Section 2.1). To see this, let us define, based on \mathcal{Z}^∞ in (2.10), the sets $\mathcal{Z}_i^\infty = \{\mathbf{x}_i(t) = (y_i(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, +\infty\}$ with y_i as given in (2.3). Then, we observe that verifying (B.7) for all $\Delta \mathbf{W}_i(z) \in \Delta \mathbf{W}_{\mathbf{0},i}$ is equivalent to verifying that \mathcal{Z}_i^∞ is informative with respect to the MISO model structure \mathcal{M}_i

$$\mathcal{M}_i = \{(\mathbf{G}_i(z, \boldsymbol{\theta}_i), H_i(z, \boldsymbol{\theta}_i)) \mid \boldsymbol{\theta}_i \in \mathcal{D}_{\theta,i}\}$$

where \mathbf{G}_i is the i^{th} row of \mathbf{G} and $\mathcal{D}_{\theta,i}$ is the set that restricts $\boldsymbol{\theta}_i$ to those values such that $\mathbf{G}_i(z, \boldsymbol{\theta}_i)$ is stable and $H_i(z, \boldsymbol{\theta}_i)$ is stable and inversely stable.

In the sequel of this appendix, we will focus on deriving results for data informativity w.r.t. MISO ARMAX model structures.

B.5.2 Notation for MISO ARMAX model structures

In a MISO ARMAX model structure, we have that

$$\begin{aligned}\mathbf{G}(z, \boldsymbol{\theta}) &= \mathbf{B}(z, \tilde{\boldsymbol{\theta}})/A(z, \boldsymbol{\eta}) \\ H(z, \boldsymbol{\theta}) &= C(z, \boldsymbol{\eta})/A(z, \boldsymbol{\eta})\end{aligned}$$

where $\mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ is a vector consisting of n_u independently parametrized polynomials $B_k(z, \tilde{\boldsymbol{\theta}}_k)$ defined as in Section 3.3.1 ($k = 1, \dots, n_u$) and where $A(z, \boldsymbol{\eta})$ and $C(z, \boldsymbol{\eta})$ are monic independently parametrized polynomials. As in Section 3.3.1, the parameter vector $\boldsymbol{\theta} \in \mathbb{R}^{\tilde{n}}$ is the concatenation of $\tilde{\boldsymbol{\theta}}_k$ ($k = 1, \dots, n_u$), i.e., $\boldsymbol{\theta} = (\tilde{\boldsymbol{\theta}}_1^T, \tilde{\boldsymbol{\theta}}_2^T, \dots, \tilde{\boldsymbol{\theta}}_{n_u}^T)^T$ and where \tilde{n} is the dimension of $\tilde{\boldsymbol{\theta}}$. Let us also introduce the following notation for $A(z, \boldsymbol{\eta})$ and $C(z, \boldsymbol{\eta})$

$$\begin{aligned}C(z, \boldsymbol{\eta}) &= 1 + \sum_{m=1}^{\deg(C)} \eta_m z^{-m} \\ A(z, \boldsymbol{\eta}) &= 1 + \sum_{m=1}^{\deg(A)} \eta_{(m+\deg(C))} z^{-m}\end{aligned}$$

where η_m denotes the m^{th} entry of $\boldsymbol{\eta}$.

B.5.3 Result

Theorem B.3 (MISO ARMAX informativity and regressor persistency). *Consider the data set \mathcal{Z}^∞ obtained by applying an input vector \mathbf{u} on a MISO ARMAX system (see Definition 2.2). Consider a full-order MISO ARMAX model structure \mathcal{M} for this MISO system and the notations introduced in Section B.5.2. Then, the data set \mathcal{Z}^∞ is informative with respect to \mathcal{M} if the regressor $\boldsymbol{\phi}_{\mathbf{u}}$ defined below is PE:*

$$\boldsymbol{\phi}_{\mathbf{u}}(t) = \begin{pmatrix} \phi_{u_1}(t) \\ \vdots \\ \phi_{u_{n_u}}(t) \end{pmatrix} \quad \text{with} \quad \phi_{u_k}(t) = \begin{pmatrix} u_k(t - \rho_k) \\ \vdots \\ u_k(t - \mu_k) \end{pmatrix} \quad (\text{B.8})$$

The dimension ν of the regressor $\boldsymbol{\phi}_{\mathbf{u}}$ in (B.8) is given by $\nu = n_u + \sum_{k=1}^{n_u} (\mu_k - \rho_k)$ (the dimension of ϕ_{u_k} being equal to $\mu_k - \rho_k + 1$ ($k = 1, \dots, n_u$)) and the scalars μ_k ($k = 1, \dots, n_u$) in (B.8) are given by $\mu_k = \deg(B_k) + \min(\deg(A), \deg(C))$.

B.5.4 Proof of Theorem B.3

First note that Theorem B.1 also applies to the case $n_y = 1$. The data are informative if and only if the property (B.6) holds for all $(\Delta W_y(z), \Delta \mathbf{W}_u(z)) \in \boldsymbol{\Delta}_{\mathbf{W}_0}$. For this

proof, we are going to deduce two sufficient conditions that guarantee (B.6) for all $\Delta \mathbf{W}(z) \in \mathbf{\Delta W}_0$ by using the fact that $\Delta W_y(z) \equiv 0$. For MISO ARMAX model structures, the latter implies

$$\frac{C(z, \boldsymbol{\eta}')}{A(z, \boldsymbol{\eta}')} = \frac{C(z, \boldsymbol{\eta}'')}{A(z, \boldsymbol{\eta}'')} \quad (\text{B.9})$$

First sufficient condition for (B.6)

For MISO ARMAX model structures, $\Delta \mathbf{W}_u(z)$ can directly be rewritten as follows

$$\begin{aligned} \Delta \mathbf{W}_u(z) &= \frac{1}{C(z, \boldsymbol{\eta}')} \mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - \frac{1}{C(z, \boldsymbol{\eta}'')} \mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \\ &= \frac{1}{C(z, \boldsymbol{\eta}')C(z, \boldsymbol{\eta}'')} \left(C(z, \boldsymbol{\eta}'') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - C(z, \boldsymbol{\eta}') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \right) \end{aligned} \quad (\text{B.10})$$

Therefore, the left-hand side of (B.6) can be recast as

$$\bar{E} \left[\frac{1}{C(z, \boldsymbol{\eta}')C(z, \boldsymbol{\eta}'')} \left(C(z, \boldsymbol{\eta}'') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - C(z, \boldsymbol{\eta}') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \right) \mathbf{u}(t) \right]^2 = 0 \quad (\text{B.11})$$

Since $\boldsymbol{\theta}' = (\tilde{\boldsymbol{\theta}}'^T, \boldsymbol{\eta}'^T)^T$ and $\boldsymbol{\theta}'' = (\tilde{\boldsymbol{\theta}}''^T, \boldsymbol{\eta}''^T)^T$ belong to \mathcal{D}_θ , the transfer functions $C(z, \boldsymbol{\eta}')$ and $C(z, \boldsymbol{\eta}'')$ are both stable and inversely stable and so does $C(z, \boldsymbol{\eta}')C(z, \boldsymbol{\eta}'')$. Consequently, from Lemma B.1, (B.11) is equivalent to

$$\bar{E} \left[\left(C(z, \boldsymbol{\eta}'') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - C(z, \boldsymbol{\eta}') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \right) \mathbf{u}(t) \right]^2 = 0$$

The term $C(z, \boldsymbol{\eta}'') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - C(z, \boldsymbol{\eta}') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}'')$ is a row vector of n_u polynomials where each entry $C(z, \boldsymbol{\eta}'') B_k(z, \tilde{\boldsymbol{\theta}}'_k) - C(z, \boldsymbol{\eta}') B_k(z, \tilde{\boldsymbol{\theta}}''_k)$ has a delay of ρ_k and a degree of $\deg(B_k) + \deg(C)$. The latter has an expression of the form

$$C(z, \boldsymbol{\eta}'') B_k(z, \tilde{\boldsymbol{\theta}}'_k) - C(z, \boldsymbol{\eta}') B_k(z, \tilde{\boldsymbol{\theta}}''_k) = \sum_{h=\rho_k}^{\deg(B_k)+\deg(C)} \delta_{k,[C]}^{(h)} z^{-h}$$

where the coefficients $\delta_{k,[C]}^{(h)}$ are known polynomial functions of $\boldsymbol{\theta}'$ and $\boldsymbol{\theta}''$. Combining the above elements yields

$$\left(C(z, \boldsymbol{\eta}'') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - C(z, \boldsymbol{\eta}') \mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \right) \mathbf{u}(t) = \boldsymbol{\phi}_{\mathbf{u},[C]}^T(t) \boldsymbol{\delta}_{[C]} \quad (\text{B.12})$$

with $\boldsymbol{\delta}_{[C]}$ the vector of dimension $\nu_{[C]}$ consisting of the concatenation of all $\delta_{k,[C]}^{(h)}$ and $\boldsymbol{\phi}_{\mathbf{u},[C]}$ the regressor defined as follows

$$\boldsymbol{\phi}_{\mathbf{u},[C]}(t) = \begin{pmatrix} \phi_{u_1,[C]}(t) \\ \vdots \\ \phi_{u_{n_u},[C]}(t) \end{pmatrix} \quad \text{with} \quad \phi_{u_k,[C]}(t) = \begin{pmatrix} u_k(t - \rho_k) \\ \vdots \\ u_k(t - \deg(B_k) - \deg(C)) \end{pmatrix}$$

Moreover, the right-hand side of (B.6) is equivalent to

$$C(z, \boldsymbol{\eta}'')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - C(z, \boldsymbol{\eta}')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \equiv \mathbf{0}_{1 \times n_u}$$

which in turn is equivalent to $\boldsymbol{\delta}_{[C]} = \mathbf{0}$.

Denote by $\mathbf{D}_{[C]}$ the set of all $\boldsymbol{\delta}_{[C]}$ such that (B.12) holds for all $\Delta \mathbf{W}(z) \in \boldsymbol{\Delta}_{\mathbf{W}_0}$. Then, (B.6) holds for all $\Delta \mathbf{W}(z) \in \boldsymbol{\Delta}_{\mathbf{W}_0}$ if the following property holds for all $\boldsymbol{\delta}_{[C]} \in \mathbf{D}_{[C]}$

$$\boldsymbol{\delta}_{[C]}^T \bar{E} \left[\boldsymbol{\phi}_{\mathbf{u},[C]}(t) \boldsymbol{\phi}_{\mathbf{u},[C]}^T(t) \right] \boldsymbol{\delta}_{[C]} = 0 \implies \boldsymbol{\delta}_{[C]} = \mathbf{0} \quad (\text{B.13})$$

If $\boldsymbol{\phi}_{\mathbf{u},[C]}$ is PE, then (B.13) holds for all $\boldsymbol{\delta}_{[C]} \in \mathbf{D}_{[C]}$ and so the data are informative. Since it is clear that $\mathbf{D}_{[C]} \neq \mathbb{R}^{\nu_{[C]}}$ because $\boldsymbol{\eta}'$ and $\boldsymbol{\eta}''$ do not cover the whole real space \mathbb{R}^{n_η} due to stability restriction, we cannot conclude that the persistency of $\boldsymbol{\phi}_{\mathbf{u},[C]}$ is necessary for the data informativity.

Second sufficient condition for (B.6)

For the second condition, we will use (B.9). By multiplying (B.9) by $A(z, \boldsymbol{\eta}'')$, we obtain

$$C(z, \boldsymbol{\eta}'') = \frac{A(z, \boldsymbol{\eta}'')C(z, \boldsymbol{\eta}')}{A(z, \boldsymbol{\eta}')} \quad (\text{B.14})$$

By injecting (B.14) into the expression of $\Delta \mathbf{W}_u(z)$ in (B.10), we obtain

$$\Delta \mathbf{W}_u(z) = \frac{1}{C(z, \boldsymbol{\eta}')} \mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - \frac{A(z, \boldsymbol{\eta}')}{A(z, \boldsymbol{\eta}'')C(z, \boldsymbol{\eta}')} \mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \quad (\text{B.15})$$

By factorizing by $1/(A(z, \boldsymbol{\eta}'')C(z, \boldsymbol{\eta}'))$, (B.15) can be recast as

$$\Delta \mathbf{W}_u(z) = \frac{1}{A(z, \boldsymbol{\eta}'')C(z, \boldsymbol{\eta}')} \left(A(z, \boldsymbol{\eta}'')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - A(z, \boldsymbol{\eta}')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \right)$$

Therefore, the left-hand side of (B.6) can be recast as

$$\bar{E} \left[\frac{1}{A(z, \boldsymbol{\eta}'')C(z, \boldsymbol{\eta}')} \left(A(z, \boldsymbol{\eta}'')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - A(z, \boldsymbol{\eta}')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \right) \mathbf{u}(t) \right]^2 = 0 \quad (\text{B.16})$$

Since $\boldsymbol{\theta}' = (\tilde{\boldsymbol{\theta}}'^T, \boldsymbol{\eta}'^T)^T$ and $\boldsymbol{\theta}'' = (\tilde{\boldsymbol{\theta}}''^T, \boldsymbol{\eta}''^T)^T$ belong to \mathcal{D}_θ , the transfer functions $C(z, \boldsymbol{\eta}')$ and $A(z, \boldsymbol{\eta}'')$ are both stable and inversely stable and so does the transfer $A(z, \boldsymbol{\eta}'')C(z, \boldsymbol{\eta}')$. Consequently, from Lemma B.1, (B.16) is equivalent to

$$\bar{E} \left[\left(A(z, \boldsymbol{\eta}'')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - A(z, \boldsymbol{\eta}')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \right) \mathbf{u}(t) \right]^2 = 0$$

The term $A(z, \boldsymbol{\eta}'')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - A(z, \boldsymbol{\eta}')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}'')$ is a row vector of n_u polynomials where each entry $A(z, \boldsymbol{\eta}'')B_k(z, \tilde{\boldsymbol{\theta}}'_k) - A(z, \boldsymbol{\eta}')B_k(z, \tilde{\boldsymbol{\theta}}''_k)$ has a delay of ρ_k and a degree of $\deg(B_k) + \deg(A)$. The latter has an expression of the form

$$A(z, \boldsymbol{\eta}'')B_k(z, \tilde{\boldsymbol{\theta}}'_k) - A(z, \boldsymbol{\eta}')B_k(z, \tilde{\boldsymbol{\theta}}''_k) = \sum_{h=\rho_k}^{\deg(B_k)+\deg(A)} \delta_{k,[A]}^{(h)} z^{-h}$$

where the coefficients $\delta_{k,[A]}^{(h)}$ are known polynomial functions of $\boldsymbol{\theta}'$ and $\boldsymbol{\theta}''$. Combining the above elements yields

$$\left(A(z, \boldsymbol{\eta}'')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - A(z, \boldsymbol{\eta}')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \right) \mathbf{u}(t) = \boldsymbol{\phi}_{\mathbf{u},[A]}^T(t) \boldsymbol{\delta}_{[A]}$$

with $\boldsymbol{\delta}_{[A]}$ the vector of dimension $\nu_{[A]}$ consisting of the concatenation of all $\delta_{k,[A]}^{(h)}$ and $\boldsymbol{\phi}_{\mathbf{u},[A]}$ the regressor defined as follows

$$\boldsymbol{\phi}_{\mathbf{u},[A]}(t) = \begin{pmatrix} \phi_{u_1,[A]}(t) \\ \vdots \\ \phi_{u_{n_u},[A]}(t) \end{pmatrix} \quad \text{with} \quad \phi_{u_k,[A]}(t) = \begin{pmatrix} u_k(t - \rho_k) \\ \vdots \\ u_k(t - \deg(B_k) - \deg(A)) \end{pmatrix}$$

Moreover, the right-hand side of (B.6) is equivalent to

$$A(z, \boldsymbol{\eta}'')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}') - A(z, \boldsymbol{\eta}')\mathbf{B}(z, \tilde{\boldsymbol{\theta}}'') \equiv \mathbf{0}_{1 \times n_u}$$

which in turn is equivalent to $\boldsymbol{\delta}_{[A]} = \mathbf{0}$.

Denote by $\mathbf{D}_{[A]}$ the set of all $\boldsymbol{\delta}_{[A]}$ such that (B.12) holds for all $\Delta \mathbf{W}(z) \in \boldsymbol{\Delta} \mathbf{w}_0$. Then, (B.6) holds for all $\Delta \mathbf{W}(z) \in \boldsymbol{\Delta} \mathbf{w}_0$ if the following property holds for all $\boldsymbol{\delta}_{[A]} \in \mathbf{D}_{[A]}$

$$\boldsymbol{\delta}_{[A]}^T \bar{E} \left[\boldsymbol{\phi}_{\mathbf{u},[A]}(t) \boldsymbol{\phi}_{\mathbf{u},[A]}^T(t) \right] \boldsymbol{\delta}_{[A]} = 0 \implies \boldsymbol{\delta}_{[A]} = \mathbf{0} \quad (\text{B.17})$$

If $\boldsymbol{\phi}_{\mathbf{u},[A]}$ is PE, then (B.17) holds for all $\boldsymbol{\delta}_{[A]} \in \mathbf{D}_{[A]}$ and so the data are informative. Since it is clear that $\mathbf{D}_{[A]} \neq \mathbb{R}^{\nu_{[A]}}$ because $\boldsymbol{\eta}'$ and $\boldsymbol{\eta}''$ do not cover the whole real space \mathbb{R}^{n_η} due to stability restriction, the persistency of $\boldsymbol{\phi}_{\mathbf{u},[A]}$ is not necessary for the data informativity.

The least conservative condition

In both previous paragraphs, we have developed two sufficient conditions for the data informativity w.r.t. MISO ARMAX: the persistency of $\boldsymbol{\phi}_{\mathbf{u},[C]}$ and the persistency of $\boldsymbol{\phi}_{\mathbf{u},[A]}$.

Depending on the dimension of C and the dimension of A , one of these two regressors will contain all the elements of the other one and additional elements. Since only one

of the two regressors have to be PE for the data informativity, we can restrict attention to the regressor with the smaller dimension, i.e., the predictor (B.8) with μ_k as defined in the statement of the theorem. \square

Remark B.1. Like for the other model structures, if we apply Theorem B.3 to the SISO case ($n_u = 1$), we observe that the condition for data informativity is that the regressor $\phi_u = \phi_{u_1}$ of dimension $\nu = \deg(B_1) - \rho_1 + \min(\deg(A), \deg(C)) + 1$ is PE. Using Lemma 2.4 (see Section 2.7), this condition is equivalent to the fact that the power spectrum $\Phi_u(\omega)$ of $u = u_1$ is nonzero in at least $\nu = \deg(B_1) - \rho_1 + \min(\deg(A), \deg(C)) + 1$ frequencies in $]-\pi, \pi]$. We then retrieve the results for data informativity with respect to ARMAX model structures in the SISO case [GBM08].

B.6 Proof of the persistency of $\varphi^{(sin)}$

To prove the persistency of $\varphi^{(sin)}$ in (3.7), we will need the following lemma.

Lemma B.2. Consider the discrete-time complex-valued signal x given by

$$x(t) = e^{j\omega t} \quad t \in \mathbb{N}^*$$

where $\omega \in \mathbb{R}$. Then,

$$\bar{E}[x(t)] = \begin{cases} 0 & \text{if } \omega \neq 0 \text{ modulo } 2\pi. \\ 1 & \text{if } \omega = 0 \text{ modulo } 2\pi. \end{cases}$$

Proof. We have to prove that $\bar{E}[x(t)]$ is well-defined by proving that

$$\frac{1}{N} \sum_{t=1}^N E[x(t)]$$

converges to 0 when $\omega = 0$ modulo 2π and to 1 when $\omega \neq 0$ modulo 2π when $N \rightarrow +\infty$. Since x is deterministic, we have $E[x(t)] = x(t)$. Let us prove then that

$$\frac{1}{N} \sum_{t=1}^N x(t)$$

converges to 0 when $\omega = 0$ modulo 2π and to 1 when $\omega \neq 0$ modulo 2π when $N \rightarrow +\infty$. For all $N \in \mathbb{N}^*$, we have that

$$\begin{aligned} \frac{1}{N} \sum_{t=1}^N x(t) &= \frac{1}{N} \sum_{t=1}^N (e^{j\omega})^t \\ &= \begin{cases} \frac{1}{N} \frac{1 - e^{jN\omega}}{1 - e^{j\omega}} & \text{if } \omega \neq 0 \text{ modulo } 2\pi. \\ \frac{1}{N} \sum_{t=1}^N 1 & \text{if } \omega = 0 \text{ modulo } 2\pi. \end{cases} \\ &= \begin{cases} \frac{1}{N} \frac{1 - e^{-j\omega} - e^{jN\omega} + e^{j(N-1)\omega}}{2 - 2\cos(\omega)} & \text{if } \omega \neq 0 \text{ modulo } 2\pi. \\ 1 & \text{if } \omega = 0 \text{ modulo } 2\pi. \end{cases} \end{aligned}$$

The result is readily obtained for $\omega = 0$ modulo 2π . Let us focus on the case where $\omega \neq 0$ modulo 2π . In order to conclude that $\bar{E}[x(t)]$ is equal 0, we have to prove that the complex-valued sequence $(a_N)_{N \in \mathbb{N}^*}$ converges to 0 with a_N given by

$$a_N = \frac{1}{N} \frac{1 - e^{-j\omega} - e^{jN\omega} + e^{j(N-1)\omega}}{2 - 2 \cos(\omega)} \quad (\text{B.18})$$

The latter is equivalent to the fact that the sequence $(b_N)_{N \in \mathbb{N}^*} = (|a_N|)_{N \in \mathbb{N}^*}$ converges to 0 with a_N as defined in (B.18). Let us prove that it is the case.

For all $N \in \mathbb{N}^*$, we have that

$$b_N = \frac{1}{N} \left| \frac{1 - e^{-j\omega} - e^{jN\omega} + e^{j(N-1)\omega}}{2 - 2 \cos(\omega)} \right| \leq \frac{1}{N} \frac{4}{2 - 2 \cos(\omega)} \leq \frac{1}{N} \frac{2}{1 - \cos(\omega)}$$

Denote by c_N the term $c_N = 1/N \times 2/(1 - \cos(\omega))$. The sequence $(c_N)_{N \in \mathbb{N}^*}$ converges to 0 when $N \rightarrow +\infty$. Consequently, from the squeezing theorem, we conclude that the sequence $(b_N)_{N \in \mathbb{N}^*}$ converges to 0. Therefore, the sequence $(a_N)_{N \in \mathbb{N}^*}$ also converges and its limit $\bar{E}[x(t)]$ is equal to 0, which concludes the proof. \square

Let us now prove the persistency of $\varphi^{(sin)}$ in (3.7). For this purpose, let us calculate $\bar{E}[\varphi^{(sin)}(t)\varphi^{(sin)*}(t)]$:

$$\begin{aligned} & \bar{E}[\varphi^{(sin)}(t)\varphi^{(sin)*}(t)] \\ &= \frac{1}{4} \begin{pmatrix} 1 & e^{j2\omega_1 t} & \dots & \dots & e^{j(\omega_1 - \omega_s)t} & e^{j(\omega_1 + \omega_s)t} \\ e^{-j2\omega_1 t} & 1 & \dots & \dots & e^{-j(\omega_1 + \omega_s)t} & e^{-j(\omega_1 - \omega_s)t} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ e^{-j(\omega_1 - \omega_s)t} & e^{j(\omega_1 + \omega_s)t} & \dots & \dots & 1 & e^{j2\omega_s t} \\ e^{-j(\omega_1 + \omega_s)t} & e^{j(\omega_1 - \omega_s)t} & \dots & \dots & e^{-j2\omega_s t} & 1 \end{pmatrix} \end{aligned}$$

Since all ω_l ($l = 1, \dots, s$) are different and belong to $]0, \pi[$, then $\omega_l \pm \omega_{l'} \neq 0$ modulo 2π for all $l = 1, \dots, s$ and $l' = 1, \dots, s$. Consequently, from Lemma B.2, we have that $\bar{E}[\varphi^{(sin)}(t)\varphi^{(sin)*}(t)] = \frac{1}{4}\mathbf{I}_{2s} > 0$, which concludes the proof for the persistency of $\varphi^{(sin)}$. \square

B.7 Proof of Lemma 3.5

Let us consider Theorem 3.1 for the MISO case ($n_y = 1$) and let us define, for each $k = 1, \dots, n_u$, the set

$$\Delta \mathbf{X}_k = \left\{ \Delta X_k(z) = X_k(z, \tilde{\theta}'_k) - X_k(z, \tilde{\theta}''_k) \mid \tilde{\theta}'_k \text{ and } \tilde{\theta}''_k \in \mathcal{D}_{\tilde{\theta}_k} \right\}$$

with $\mathcal{D}_{\tilde{\theta}_k} = \left\{ \tilde{\theta}_k \mid X_k(z, \tilde{\theta}_k) \text{ is stable} \right\}$ and $X_k(z, \tilde{\theta}_k) = B_k(z, \tilde{\theta}_k)$ (ARX/FIR case) or $X_k(z, \tilde{\theta}_k) = G_k(z, \tilde{\theta}_k)$ (BJ/OE case).

Now, because of condition (ii), let us observe that the scalar inputs u_k are all mutually uncorrelated. Consequently, the condition (3.2) can be successively rewritten as:

$$\begin{aligned} \sum_{k=1}^{n_u} \bar{E} [(\Delta X_k(z)u_k(t))^2] = 0 &\implies \Delta X_k \equiv 0 \quad \forall k \in \llbracket 1, n_u \rrbracket \\ \Leftrightarrow \bar{E} [(\Delta X_k(z)u_k(t))^2] = 0 \quad \forall k \in \llbracket 1, n_u \rrbracket &\implies \Delta X_k \equiv 0 \quad \forall k \in \llbracket 1, n_u \rrbracket \end{aligned} \quad (\text{B.19})$$

Since they are no common parameters in ΔX_k and ΔX_m ($m \neq k$), we have that (B.19) is also equivalent to the fact that the following condition holds for each $k = 1, \dots, n_u$:

$$\bar{E} [(\Delta X_k(z)u_k(t))^2] = 0 \implies \Delta X_k \equiv 0 \quad (\text{B.20})$$

From the above reasoning and Theorem 3.1, \mathbf{u} yields informative data with respect to \mathcal{M} if and only if, for each $k = 1, \dots, n_u$, (B.20) holds for all $\Delta X_k \in \mathbf{\Delta X}_k$.

By Theorem 3.1 applied to the SISO case ($n_u = 1, n_y = 1$) for each individual k , that (B.20) holds for all $\Delta X_k \in \mathbf{\Delta X}_k$ is equivalent to the fact that u_k yields informative data for a SISO model structure described¹ by $B_k(z, \tilde{\theta}_k)$ (in the ARX/FIR case) and by $G_k(z, \tilde{\theta}_k)$ (in the BJ/OE case). Lemma 3.5 then follows from the classical necessary and sufficient conditions for data informativity in the SISO case. These results in the SISO case are given in [GBM08], but are also reminded below Theorem 3.3. \square

B.8 Lemma 3.5 for the ARMAX model structure

Lemma B.3. *Consider the data set \mathcal{Z}^∞ in (2.10) obtained by applying a multisine input vector \mathbf{u} (see (2.14)) on a MISO ARMAX system with n_u inputs. Consider a full-order MISO ARMAX model structure \mathcal{M} for this MISO system and the notations introduced in Appendix B.5.2. Then, the data set \mathcal{Z}^∞ is informative with respect to \mathcal{M} if, for each $k = 1, \dots, n_u$,*

(i) *the scalars input u_k contains a number s_k of nonzero sinusoids that satisfies*

$$s_k \geq \frac{\deg(B_k) - \rho_k + \min(\deg(A), \deg(C)) + 1}{2}$$

(ii) *the s_k frequencies of the sinusoids in u_k are different from the s_m frequencies of the sinusoids in all u_m ($m \neq k$), i.e., the signals u_k are mutually uncorrelated.*

¹The matrices A and H do not play any role for data informativity (Theorem 3.1).

Moreover, if (i) and (ii) are respected, \mathcal{Z}^∞ is informative whatever the values of the s_k (nonzero) frequencies, the s_k (nonzero) amplitudes and the s_k phase shifts defining u_k ($k = 1, \dots, n_u$).

Proof. We are going to show that the corresponding input regressor $\phi_{\mathbf{u}}$ is PE to prove the data informativity by virtue of Theorem B.3 in Appendix B.5. Because of condition (ii), let us observe that the scalar inputs u_k are all mutually uncorrelated. Consequently, all ϕ_{u_k} are all mutually uncorrelated and so the matrix $\bar{E}[\phi_{\mathbf{u}}(t)\phi_{\mathbf{u}}^T(t)]$ is of the form

$$\bar{E}[\phi_{\mathbf{u}}(t)\phi_{\mathbf{u}}^T(t)] = \text{bdiag} \left(\bar{E}[\phi_{u_1}(t)\phi_{u_1}^T(t)], \dots, \bar{E}[\phi_{u_{n_u}}(t)\phi_{u_{n_u}}^T(t)] \right)$$

Therefore, we have that $\bar{E}[\phi_{\mathbf{u}}(t)\phi_{\mathbf{u}}^T(t)] \succ 0$ (i.e., $\phi_{\mathbf{u}}$ is PE) is equivalent to the fact that $\bar{E}[\phi_{u_k}(t)\phi_{u_k}^T(t)] \succ 0 \forall k \in \llbracket 1, n_u \rrbracket$ (i.e., each ϕ_{u_k} is PE). We know that, for each $k = 1, \dots, n_u$, the PSD of each u_k is nonzero in at least $\deg(B_k) - \rho_k + \min(\deg(A), \deg(C)) + 1$ different frequencies (condition (i)) which is equal to the dimension of the regressor ϕ_{u_k} . Therefore, from Lemma 2.4, each ϕ_{u_k} is PE, which concludes the proof. \square

Appendix for Chapter 4

Contents

C.1	Proof of Theorem 4.1	189
C.2	Proof of Lemma 2.2	190
C.3	Proof of Theorem 4.2	191
C.4	Proof of Lemma 4.1	191
C.5	Proof of Theorem 4.3	193
C.6	Orders of the polynomials $\Upsilon_{y,ii}(z)$ and $\Upsilon_{u,ik}(z)$	193

C.1 Proof of Theorem 4.1

The proof is adapted from the one in [BGM10]. The idea is to prove that the left-hand sides of (2.12) and (4.1) are equivalent and that the same holds for the right-hand sides. Note that the latter is straightforward since $\Delta\mathbf{W}(z) \equiv \mathbf{0}_{n_y \times (n_y + n_u)} \Leftrightarrow (\Delta\mathbf{W}_y(z), \Delta\mathbf{W}_u(z)) \equiv (\mathbf{0}_{n_y \times n_y}, \mathbf{0}_{n_y \times n_u})$.

First, combining (2.2) and (2.13) leads to

$$\mathbf{x}(t) = \begin{pmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{pmatrix} = \begin{pmatrix} \mathbf{\Gamma}_y(z)\mathbf{G}_0(z)\mathbf{r}(t) + \mathbf{\Gamma}_y(z)\mathbf{H}_0(z)\mathbf{e}(t) \\ \mathbf{\Gamma}_u(z)\mathbf{r}(t) - \mathbf{\Gamma}_u(z)\mathbf{K}(z)\mathbf{H}_0(z)\mathbf{e}(t) \end{pmatrix} \quad (\text{C.1})$$

where $\mathbf{\Gamma}_u(z)$ and $\mathbf{\Gamma}_y(z)$ are respectively the input and output sensitivity transfer function matrices of the closed-loop depicted in Figure 2.1, defined respectively by $\mathbf{\Gamma}_u(z) = (\mathbf{I}_{n_u} + \mathbf{K}(z)\mathbf{G}_0(z))^{-1}$ and $\mathbf{\Gamma}_y(z) = (\mathbf{I}_{n_y} + \mathbf{G}_0(z)\mathbf{K}(z))^{-1}$. Using (C.1) and the independence assumption between \mathbf{r} and \mathbf{e} , $\bar{E} [|\Delta\mathbf{W}(z)\mathbf{x}(t)|^2] = 0$ leads to

$$\begin{cases} \bar{E} [|\Delta\mathbf{W}_y(z)\mathbf{\Gamma}_y(z) - \Delta\mathbf{W}_u(z)\mathbf{\Gamma}_u(z)\mathbf{K}(z)\mathbf{H}_0(z)\mathbf{e}(t)|^2] = 0 \\ \bar{E} [|\Delta\mathbf{W}_u(z)\mathbf{\Gamma}_u(z) + \Delta\mathbf{W}_y(z)\mathbf{\Gamma}_y(z)\mathbf{G}_0(z)\mathbf{r}(t)|^2] = 0 \end{cases} \quad (\text{C.2})$$

We are going to prove that (C.2) is equivalent to the left-hand side of (4.1). Let us first prove that the first equation of (C.2) is equivalent to $\Delta \mathbf{W}_y(z) - \Delta \mathbf{W}_u(z) \mathbf{K}(z) \equiv \mathbf{0}_{n_y \times n_y}$. Since $\bar{E} [\mathbf{e}(t) \mathbf{e}^T(t)] \succ 0$ and $\mathbf{H}_0(z)$ is stable and inversely stable, we have that the power spectrum matrix $\bar{\Phi}_{\tilde{\mathbf{v}}}$ of $\tilde{\mathbf{v}}(t) = \mathbf{H}_0(z) \mathbf{e}(t)$ is strictly positive definite at all frequencies. Consequently, the first equation of (C.2) is equivalent to

$$\Delta \mathbf{W}_y(z) \Gamma_y(z) - \Delta \mathbf{W}_u(z) \Gamma_u(z) \mathbf{K}(z) \equiv \mathbf{0}_{n_y \times n_y} \quad (\text{C.3})$$

From the push-through rule [SP05, Chapter 3], $\Gamma_u(z) \mathbf{K}(z) = \mathbf{K}(z) \Gamma_y(z)$. Consequently, (C.3) is equivalent to

$$(\Delta \mathbf{W}_y(z) - \Delta \mathbf{W}_u(z) \mathbf{K}(z)) \Gamma_y(z) \equiv \mathbf{0}_{n_y \times n_y}$$

Finally, by post-multiplying by $\Gamma_y^{-1}(z)$, we obtain

$$\Delta \mathbf{W}_y(z) - \Delta \mathbf{W}_u(z) \mathbf{K}(z) \equiv \mathbf{0}_{n_y \times n_y} \quad (\text{C.4})$$

which is the first equation of the left-hand side of (4.1).

Let us prove that the second equation of (C.2) is equivalent to $\bar{E} [|\Delta \mathbf{W}_u(z) \mathbf{r}(t)|^2] = 0$. Combining the second equation of (C.2) with (C.4) leads to

$$\bar{E} [|\Delta \mathbf{W}_u(z) (\Gamma_u(z) + \mathbf{K}(z) \Gamma_y(z) \mathbf{G}_0(z)) \mathbf{r}(t)|^2] = 0 \quad (\text{C.5})$$

From the push-through rule [SP05, Chapter 3], $\Gamma_y(z) \mathbf{G}_0(z) = \mathbf{G}_0(z) \Gamma_u(z)$. Consequently, (C.5) is equivalent to

$$\bar{E} [|\Delta \mathbf{W}_u(z) (\mathbf{I}_{n_u} + \mathbf{K}(z) \mathbf{G}_0(z)) \Gamma_u(z) \mathbf{r}(t)|^2] = 0$$

Finally, by observing¹ that $\mathbf{I}_{n_u} + \mathbf{K}(z) \mathbf{G}_0(z) = \Gamma_u^{-1}(z)$, we obtain that the second equation of (C.2) is equivalent to $\bar{E} [|\Delta \mathbf{W}_u(z) \mathbf{r}(t)|^2] = 0$. The proof follows. \square

C.2 Proof of Lemma 2.2

From Parseval's theorem, the term $\bar{E} [|\Delta \mathbf{W}_u(z) \mathbf{r}(t)|^2] = 0$ in the left-hand side of (4.1) is equivalent to

$$\text{tr} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \Delta \mathbf{W}_u(e^{j\omega}) \bar{\Phi}_{\mathbf{r}}(\omega) \Delta \mathbf{W}_u^*(e^{j\omega}) d\omega \right) = 0$$

where tr is the trace operator. Consequently, if $\bar{\Phi}_{\mathbf{r}}(\omega) \succ 0$ for almost all ω , then $\bar{E} [|\Delta \mathbf{W}_u(z) \mathbf{r}(t)|^2] = 0$ implies $\Delta \mathbf{W}_u \equiv \mathbf{0}_{n_y \times n_u}$. When $\Delta \mathbf{W}_u \equiv \mathbf{0}_{n_y \times n_u}$, the first equation of the left-hand side of (4.1) is equivalent to $\Delta \mathbf{W}_y \equiv \mathbf{0}_{n_y \times n_y}$. Consequently, we have data informativity when $\bar{\Phi}_{\mathbf{r}}(\omega) \succ 0 \forall \omega$ (or at almost all ω). \square

¹And by assuming that $\Gamma_u(z)$ does not filter out any part of \mathbf{r} .

C.3 Proof of Theorem 4.2

We are going to prove that the property (4.1) for all $\Delta \mathbf{W}(z) = (\Delta \mathbf{W}_y(z), \Delta \mathbf{W}_u(z)) \in \Delta_{\mathbf{W}}$ (with $\Delta_{\mathbf{W}}$ as defined in (2.11)) is equivalent to the fact that the property (4.3) holds for all $\Delta \mathbf{W}_{i:}(z) = (\Delta \mathbf{W}_{y,i:}(z), \Delta \mathbf{W}_{u,i:}(z)) \in \Delta_{\mathbf{W},i}$ (with $\Delta_{\mathbf{W}}$ as defined in (4.2)) and for all $i = 1, \dots, n_y$.

Let us first note that $\Delta \mathbf{W}(z) \equiv \mathbf{0}_{n_y \times (n_y + n_u)}$ is equivalent to $\Delta \mathbf{W}_{i:}(z) \equiv \mathbf{0}_{1 \times (n_y + n_u)}$ $\forall i \in \llbracket 1, n_y \rrbracket$. Secondly, it is also straightforward to see that $\Delta \mathbf{W}_y(z) - \Delta \mathbf{W}_u(z) \mathbf{K}(z) \equiv \mathbf{0}_{n_y \times n_y}$ is equivalent to $\Delta \mathbf{W}_{y,i:}(z) - \Delta \mathbf{W}_{u,i:}(z) \mathbf{K}(z) \equiv \mathbf{0}_{1 \times n_y}$ for all $i \in \llbracket 1, n_y \rrbracket$. Finally, let us prove that $\bar{E} [\|\Delta \mathbf{W}_u(z) \mathbf{r}(t)\|^2] = 0$ is equivalent to $\bar{E} [\|\Delta \mathbf{W}_{u,i:}(z) \mathbf{r}(t)\|^2] = 0$ for all $i \in \llbracket 1, n_y \rrbracket$. For this purpose, observe that the term $\bar{E} [\|\Delta \mathbf{W}_u(z) \mathbf{r}(t)\|^2] = 0$ can be recast as follows

$$\bar{E} [\|\Delta \mathbf{W}_u(z) \mathbf{r}(t)\|^2] = \sum_{i=1}^{n_y} \bar{E} [\|\Delta \mathbf{W}_{u,i:}(z) \mathbf{r}(t)\|^2]$$

Since the term $\bar{E} [\|\Delta \mathbf{W}_{u,i:}(z) \mathbf{r}(t)\|^2]$ is nonnegative ($i = 1, \dots, n_y$), we have indeed that $\bar{E} [\|\Delta \mathbf{W}_u(z) \mathbf{r}(t)\|^2] = 0$ is equivalent to $\bar{E} [\|\Delta \mathbf{W}_{u,i:}(z) \mathbf{r}(t)\|^2] = 0$ for all $i = 1, \dots, n_y$.

We have thus proven that the property (4.1) for all $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}}$ is equivalent to

$$\begin{cases} \Delta \mathbf{W}_{y,i:}(z) - \Delta \mathbf{W}_{u,i:}(z) \mathbf{K}(z) \equiv \mathbf{0} \\ \bar{E} [\|\Delta \mathbf{W}_{u,i:}(z) \mathbf{r}(t)\|^2] = 0 \end{cases} \quad (i = 1, \dots, n_y) \quad (\text{C.6})$$

$$\implies \Delta \mathbf{W}_{i:}(z) = (\Delta \mathbf{W}_{y,i:}(z), \Delta \mathbf{W}_{u,i:}(z)) = \mathbf{0} \quad (i = 1, \dots, n_y)$$

for all $\Delta \mathbf{W}_{i:}(z) = (\Delta \mathbf{W}_{y,i:}(z), \Delta \mathbf{W}_{u,i:}(z)) \in \Delta_{\mathbf{W},i}$ ($i = 1, \dots, n_y$).

Using the parametrization introduced in Section 2.5, we observe that there are no common parameters in $\Delta \mathbf{W}_{i:}(z) = (\Delta \mathbf{W}_{y,i:}(z), \Delta \mathbf{W}_{u,i:}(z))$ and $\Delta \mathbf{W}_{j:}(z) = (\Delta \mathbf{W}_{u,j:}(z), \Delta \mathbf{W}_{y,j:}(z))$ ($j \neq i$). Therefore, (C.6) is equivalent to the fact that property (4.3) holds for all $\Delta \mathbf{W}_{i:}(z) = (\Delta \mathbf{W}_{y,i:}(z), \Delta \mathbf{W}_{u,i:}(z)) \in \Delta_{\mathbf{W},i}$ and for all $i = 1, \dots, n_y$, completing the proof. \square

C.4 Proof of Lemma 4.1

For the left factorization of $(\Delta \mathbf{W}_{y,i:}(z), -\Delta \mathbf{W}_{u,i:}(z)) \in \Delta_{\mathbf{W},i}$, we will consider the one that consists on putting all entries of $(\Delta \mathbf{W}_{y,i:}(z), -\Delta \mathbf{W}_{u,i:}(z))$ on the same denominators.

Let us first make the following observation. The matrix $\mathbf{H}(z, \theta)$ is diagonal for all $\theta \in \mathcal{D}_\theta$ and so only the i^{th} entry of the $1 \times n_y$ row vector $\Delta \mathbf{W}_{y,i:}(z)$ of transfer functions is possibly nonzero. Let us denote $\Delta W_{y,ii}(z)$ the i^{th} entry of $\Delta \mathbf{W}_{y,i:}(z)$. Observe first that $\Delta W_{y,ii}(z) = W_{y,ii}(z, \theta') - W_{y,ii}(z, \theta'') = 1 - H_i^{-1}(z, \theta') - (1 - H_i^{-1}(z, \theta'')) = H_i^{-1}(z, \theta'') - H_i^{-1}(z, \theta')$.

Let us first study the case of the ARMAX model structure (see (2.17)). For all $\Delta \mathbf{W}_{i,:}(z) = (\Delta \mathbf{W}_{y,i}(z), \Delta \mathbf{W}_{u,i}(z)) \in \mathbf{\Delta W}_{\mathbf{i}}$, we have that

$$\begin{cases} \Delta W_{y,ii}(z) &= C_i^{-1}(z, \boldsymbol{\eta}'') A_i(z, \boldsymbol{\eta}'') - C_i^{-1}(z, \boldsymbol{\eta}') A_i(z, \boldsymbol{\eta}') \\ -\Delta \mathbf{W}_{u,i}(z) &= C_i^{-1}(z, \boldsymbol{\eta}'') \mathbf{B}_{i,:}(z, \tilde{\boldsymbol{\theta}}'') - C_i^{-1}(z, \boldsymbol{\eta}') \mathbf{B}_{i,:}(z, \tilde{\boldsymbol{\theta}}') \end{cases}$$

Therefore, by posing $Q_i = C_i' C_i''$, we obtain the factorization proposed in the lemma. Since ARX and FIR model structures are special cases of ARMAX model structures with $C_i(z, \boldsymbol{\eta}) = 1$ and $C_i(z, \boldsymbol{\eta}) = A_i(z, \boldsymbol{\eta}) = 1$ ($i = 1, \dots, n_y$) respectively, we also obtain the factorization proposed in the lemma for these two model structures.

Let us now consider the BJ model structures given in (2.18). For all $\Delta \mathbf{W}_{i,:}(z) = (\Delta \mathbf{W}_{y,i}(z), \Delta \mathbf{W}_{u,i}(z)) \in \mathbf{\Delta W}_{\mathbf{i}}$, we have that

$$\begin{cases} \Delta W_{y,ii}(z) &= C_i^{-1}(z, \boldsymbol{\eta}'') D_i(z, \boldsymbol{\eta}'') - C_i^{-1}(z, \boldsymbol{\eta}') D_i(z, \boldsymbol{\eta}') \\ -\Delta \mathbf{W}_{u,i}(z) &= C_i^{-1}(z, \boldsymbol{\eta}'') D_i(z, \boldsymbol{\eta}'') \mathbf{G}_{i,:}(z, \tilde{\boldsymbol{\theta}}'') - C_i^{-1}(z, \boldsymbol{\eta}') D_i(z, \boldsymbol{\eta}') \mathbf{G}_{i,:}(z, \tilde{\boldsymbol{\theta}}') \end{cases}$$

where $\mathbf{G}_{i,:}(z, \tilde{\boldsymbol{\theta}}')$ and $\mathbf{G}_{i,:}(z, \tilde{\boldsymbol{\theta}}'')$ are respectively the i^{th} row of $\mathbf{G}(z, \tilde{\boldsymbol{\theta}}')$ and $\mathbf{G}(z, \tilde{\boldsymbol{\theta}}'')$. Since $C_i(z, \boldsymbol{\eta}')$, $C_i(z, \boldsymbol{\eta}'')$, $D_i(z, \boldsymbol{\eta}')$ and $D_i(z, \boldsymbol{\eta}'')$ are scalars, we have that

$$(\Delta W_{y,ii}, -\Delta \mathbf{W}_{u,i}) = C_i^{-1'} C_i^{-1''} (C_i' D_i'' - C_i'' D_i', C_i' D_i'' \mathbf{G}_{i,:}'' - C_i'' D_i' \mathbf{G}_{i,:}') \quad (\text{C.7})$$

Let us put each entry of $\mathbf{G}_{i,:}(z, \tilde{\boldsymbol{\theta}}')$ and $\mathbf{G}_{i,:}(z, \tilde{\boldsymbol{\theta}}'')$ on the same denominator as follows

$$\begin{aligned} \mathbf{G}_{i,:}(z, \tilde{\boldsymbol{\theta}}') &= \mathcal{F}_i^{-1}(z, \tilde{\boldsymbol{\theta}}') \boldsymbol{\mathcal{G}}_i(z, \tilde{\boldsymbol{\theta}}') \\ \mathbf{G}_{i,:}(z, \tilde{\boldsymbol{\theta}}'') &= \mathcal{F}_i^{-1}(z, \tilde{\boldsymbol{\theta}}'') \boldsymbol{\mathcal{G}}_i(z, \tilde{\boldsymbol{\theta}}'') \end{aligned}$$

where $\mathcal{F}_i(z, \tilde{\boldsymbol{\theta}}') = \prod_{k=1}^{n_u} F_{ik}'$ and $\boldsymbol{\mathcal{G}}_i(z, \tilde{\boldsymbol{\theta}}') = (\mathcal{G}_{i,k}(z, \tilde{\boldsymbol{\theta}}'))_{k \in \llbracket 1, n_u \rrbracket}$ is a row vector of n_u polynomials of the form $\mathcal{G}_{i,k}(z, \tilde{\boldsymbol{\theta}}') = B_{ik}(z, \tilde{\boldsymbol{\theta}}') \prod_{l=1, l \neq k}^{n_u} F_{il}(z, \tilde{\boldsymbol{\theta}}')$. The polynomial $\mathcal{F}_i(z, \tilde{\boldsymbol{\theta}}'')$ and the matrix $\boldsymbol{\mathcal{G}}_i(z, \tilde{\boldsymbol{\theta}}'')$ are defined similarly.

Since $\mathcal{F}_i(z, \tilde{\boldsymbol{\theta}}')$ and $\mathcal{F}_i(z, \tilde{\boldsymbol{\theta}}'')$ are scalar polynomials, (C.7) can be recast as follows

$$\begin{aligned} \Delta W_{y,ii} &= Q_i^{-1} \Upsilon_{y,ii} = Q_i^{-1} (\mathcal{F}_i' \mathcal{F}_i'' C_i' D_i'' - \mathcal{F}_i' \mathcal{F}_i'' C_i'' D_i') \\ -\Delta \mathbf{W}_{u,i} &= Q_i^{-1} \boldsymbol{\Upsilon}_{u,i} = Q_i^{-1} (C_i' D_i'' \mathcal{F}_i' \boldsymbol{\mathcal{G}}_i'' - C_i'' D_i' \mathcal{F}_i'' \boldsymbol{\mathcal{G}}_i') \end{aligned}$$

where $Q_i = C_i' C_i'' \mathcal{F}_i' \mathcal{F}_i''$. For OE, $C_i' = C_i'' = D_i' = D_i'' = 1$. Hence, for OE model structures, $\Upsilon_{y,ii} \equiv 0$ and $\boldsymbol{\Upsilon}_{u,i}$ is the one defined in the statement. This concludes the proof. \square

C.5 Proof of Theorem 4.3

We will prove that the property (4.3) is equivalent to (4.8). First, $(\Upsilon_{y,ii}, \Upsilon_{u,i}) \equiv (0, \mathbf{0}_{1 \times n_u})$ is equivalent to $(\Delta \mathbf{W}_{y,i}, \Delta \mathbf{W}_{u,i}) = (\mathbf{0}_{1 \times n_y}, \mathbf{0}_{1 \times n_u})$ since Q_i is invertible and that all $\Delta \mathbf{W}_{y,ij}$ with $j \neq i$ are equal to 0. Hence the right-hand sides of (4.3) and (4.8) are equivalent.

Secondly, the equation $\Delta \mathbf{W}_{y,i}(z) - \Delta \mathbf{W}_{u,i}(z)\mathbf{K}(z) = \mathbf{0}_{1 \times n_y}$ combined with the right-factorization of $\mathbf{K}(z) = \mathbf{N}(z)\mathbf{V}^{-1}(z)$ can be rewritten as follows

$$(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z)) \begin{pmatrix} \mathbf{I}_{n_y} \\ \mathbf{N}(z)\mathbf{V}^{-1}(z) \end{pmatrix} \equiv \mathbf{0}_{1 \times n_y}$$

By post-multiplying by $\mathbf{V}(z)$ and by pre-multiplying by $Q_i(z)$, the latter is equivalent to

$$(\Upsilon_{y,i}(z), \Upsilon_{u,i}(z)) \begin{pmatrix} \mathbf{V}(z) \\ \mathbf{N}(z) \end{pmatrix} \equiv \mathbf{0}_{1 \times n_y}$$

By observing that all entries $\Upsilon_{y,ij}$ with $j \neq i$ of the row vector $\Upsilon_{y,i}$ are equal to 0, the latter is equivalent to

$$(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \underbrace{\begin{pmatrix} \mathbf{V}_i(z) \\ \mathbf{N}(z) \end{pmatrix}}_{\Xi_i(z)} \equiv \mathbf{0}_{1 \times n_y}$$

which is the first equation of the left-hand side of (4.8).

Finally, $\bar{E}[|\Delta \mathbf{W}_{u,i}(z)\mathbf{r}(t)|^2] = 0$ is equivalent to $\bar{E}[|Q_i^{-1}(z)\Upsilon_{u,i}(z)\mathbf{r}(t)|^2] = 0$. Since $Q_i(z)$ is a FIR (stable) and $Q_i^{-1}(z)$ is stable, the latter is equivalent to the equation $\bar{E}[|Q_i(z)Q_i^{-1}(z)\Upsilon_{u,i}(z)\mathbf{r}(t)|^2] = 0$ which is in turn equivalent to $\bar{E}[|\Upsilon_{u,i}(z)\mathbf{r}(t)|^2] = 0$ (see Lemma B.1), which is the desired result. \square

C.6 Orders of the polynomials $\Upsilon_{y,ii}(z)$ and $\Upsilon_{u,ik}(z)$

When nonzero (i.e., for the BJ/ARX/ARMAX case), $\Upsilon_{y,ii}(z)$ can be rewritten as in (4.9) with $\mu_{y,i}$ given by (see Lemma 4.1):

- $\mu_{y,i} = \deg(A_i)$ for the ARX case.
- $\mu_{y,i} = \deg(A_i) + \deg(C_i)$ for the ARMAX case.
- $\mu_{y,i} = \deg(C_i) + \deg(D_i) + 2 \sum_{k=1}^{n_u} \deg(F_{ik})$ for the BJ case.

The polynomials $\Upsilon_{u,ik}(z)$ can be written as in (4.10) with ρ_{ik} as defined in Section 2.5 and with $\mu_{u,ik}$ given by (see Lemma 4.1):

- $\mu_{u,ik} = \deg(B_{ik})$ for the FIR/ARX case.
- $\mu_{u,ik} = \deg(B_{ik}) + \deg(C_i)$ for the ARMAX case.
- $\mu_{u,ik} = \deg(B_{ik}) + \deg(F_{ik}) + 2 \sum_{l=1, l \neq k}^{n_u} \deg(F_{il})$ for the OE case.
- $\mu_{u,ik} = \deg(C_i) + \deg(D_i) + \deg(B_{ik}) + \deg(F_{ik}) + 2 \sum_{l=1, l \neq k}^{n_u} \deg(F_{il})$ for the BJ case.

Appendix for Chapter 5

Contents

D.1 Proof of Lemma 5.1	195
D.2 Proof of Theorem 5.1	197

D.1 Proof of Lemma 5.1

Let us first observe that the multisine u_2 can be decomposed as follows

$$u_2 = u_2^{(\parallel u_1)} + u_2^{(\perp u_1)}$$

where

- $u_2^{(\parallel u_1)}$ is the multisine whose PSD shares the same frequencies with the PSD of u_1 , i.e., the multisine part of u_2 that is correlated to u_1 . The PSD of $u_2^{(\parallel u_1)}$ is nonzero in $m_{1,2}$ frequencies.
- $u_2^{(\perp u_1)}$ is the multisine whose PSD does not share any frequency with the PSD of u_1 , i.e., the multisine part of u_2 that is not correlated to u_1 (and to $u_2^{(\parallel u_1)}$ too by construction). The PSD of $u_2^{(\perp u_1)}$ is nonzero in $m_2 - m_{1,2}$ frequencies.

Consider now Theorem 3.1 with $n_u = 2$ and $n_y = 1$. Since $u_2^{(\perp u_1)}$ is uncorrelated to u_1 and to $u_2^{(\parallel u_1)}$, the left-hand side of (3.2) is equivalent to the following equation system, for all $\Delta \mathbf{X}(z) = (\Delta X_1(z), \Delta X_2(z)) \in \Delta \mathbf{X}$,

$$\begin{cases} \bar{E} \left[\|\Delta X_1(z)u_1(t) + \Delta X_2(z)u_2^{(\parallel u_1)}(t)\|^2 \right] = 0 \\ \bar{E} \left[\|\Delta X_2(z)u_2^{(\perp u_1)}(t)\|^2 \right] = 0 \end{cases} \quad (\text{D.1})$$

We have to prove that (D.1) implies $\Delta \mathbf{X} = (\Delta X_1, \Delta X_2) \equiv \mathbf{0}_{1 \times 2}$.

Let us prove the lemma for BJ/OE model structures. The proof will follow for ARX/FIR model structures. The term $\Delta X_2(z)$ in the second equation of (D.1) can be recast as

$$\Delta X_2(z) = \frac{B_2(z, \tilde{\boldsymbol{\theta}}'_2)}{F_2(z, \tilde{\boldsymbol{\theta}}'_2)} - \frac{B_2(z, \tilde{\boldsymbol{\theta}}''_2)}{F_2(z, \tilde{\boldsymbol{\theta}}''_2)} = \frac{B_2(z, \tilde{\boldsymbol{\theta}}'_2)F_2(z, \tilde{\boldsymbol{\theta}}''_2) - B_2(z, \tilde{\boldsymbol{\theta}}''_2)F_2(z, \tilde{\boldsymbol{\theta}}'_2)}{F_2(z, \tilde{\boldsymbol{\theta}}'_2)F_2(z, \tilde{\boldsymbol{\theta}}''_2)}$$

By injecting the latter in the second equation of (D.1), we obtain

$$\bar{E} \left[\left(\left(\frac{B_2(z, \tilde{\boldsymbol{\theta}}'_2)F_2(z, \tilde{\boldsymbol{\theta}}''_2) - B_2(z, \tilde{\boldsymbol{\theta}}''_2)F_2(z, \tilde{\boldsymbol{\theta}}'_2)}{F_2(z, \tilde{\boldsymbol{\theta}}'_2)F_2(z, \tilde{\boldsymbol{\theta}}''_2)} \right) u_2^{(\perp u_1)}(t) \right)^2 \right] = 0$$

Since $F_2(z, \tilde{\boldsymbol{\theta}}'_2)F_2(z, \tilde{\boldsymbol{\theta}}''_2)$ is stable (product of two FIR filters) and inversely stable (stability restriction of the set \mathcal{D}_θ), by virtue of Lemma B.1, the latter is equivalent to

$$\bar{E} \left[\left(\left(B_2(z, \tilde{\boldsymbol{\theta}}'_2)F_2(z, \tilde{\boldsymbol{\theta}}''_2) - B_2(z, \tilde{\boldsymbol{\theta}}''_2)F_2(z, \tilde{\boldsymbol{\theta}}'_2) \right) u_2^{(\perp u_1)}(t) \right)^2 \right] = 0$$

By denoting $\boldsymbol{\delta}_2$ the vector containing the polynomial coefficients of $B_2(z, \tilde{\boldsymbol{\theta}}'_2)F_2(z, \tilde{\boldsymbol{\theta}}''_2) - B_2(z, \tilde{\boldsymbol{\theta}}''_2)F_2(z, \tilde{\boldsymbol{\theta}}'_2)$, the latter can be rewritten as follows

$$\boldsymbol{\delta}_2^T \bar{E} \left[\boldsymbol{\phi}_{u_2^{(\perp u_1)}}(t) \boldsymbol{\phi}_{u_2^{(\perp u_1)}}^T(t) \right] \boldsymbol{\delta}_2 = 0 \quad (\text{D.2})$$

where $\boldsymbol{\phi}_{u_2^{(\perp u_1)}}$ is defined by

$$\boldsymbol{\phi}_{u_2^{(\perp u_1)}}(t) = \begin{pmatrix} u_2^{(\perp u_1)}(t - \rho_2) \\ u_2^{(\perp u_1)}(t - \rho_2 - 1) \\ \vdots \\ u_2^{(\perp u_1)}(t - \varpi_2) \end{pmatrix}$$

with ϖ_2 as defined in the lemma statement. The regressor $\boldsymbol{\phi}_{u_2^{(\perp u_1)}}$ is PE since the PSD of $u_2^{(\perp u_1)}$ is nonzero in at least $\varpi_2 - \rho_2 + 1$ different frequencies in $]-\pi, \pi]$ (Lemma 2.4). Consequently, (D.2) implies that $\boldsymbol{\delta}_2 = \mathbf{0}$ which in turn implies that $B_2(z, \tilde{\boldsymbol{\theta}}'_2)F_2(z, \tilde{\boldsymbol{\theta}}''_2) - B_2(z, \tilde{\boldsymbol{\theta}}''_2)F_2(z, \tilde{\boldsymbol{\theta}}'_2) \equiv 0$. This leads to $\Delta X_2(z) \equiv 0$. By injecting the latter in the first equation of (D.1), we obtain $\bar{E} [|\Delta X_1(z)u_1(t)|^2] = 0$.

By rewriting $\bar{E} [|\Delta X_1(z)u_1(t)|^2] = 0$ as we did for $\bar{E} [|\Delta X_2(z)u_2^{(\perp u_1)}(t)|^2] = 0$, we obtain

$$\boldsymbol{\delta}_1^T \bar{E} [\boldsymbol{\phi}_{u_1}(t) \boldsymbol{\phi}_{u_1}^T(t)] \boldsymbol{\delta}_1 = 0$$

where ϕ_{u_1} is defined by

$$\phi_{u_1}(t) = \begin{pmatrix} u_1(t - \rho_1) \\ u_1(t - \rho_1 - 1) \\ \vdots \\ u_1(t - \varpi_1) \end{pmatrix}$$

with ϖ_1 as defined in the lemma statement and δ_1 the vector containing all the polynomial coefficients of $B_1(z, \tilde{\theta}'_1)F_1(z, \tilde{\theta}''_1) - B_1(z, \tilde{\theta}''_1)F_1(z, \tilde{\theta}'_1)$. The PSD of u_1 is nonzero in at least $\varpi_1 - \rho_1 + 1$ frequencies in $]-\pi, \pi]$. Therefore, from Lemma 2.4, ϕ_{u_1} is PE. Consequently, we can state that, with such an input u_1 , $\bar{E} [|\Delta X_1(z)u_1(t)|^2] = 0$ implies that $\delta_1 = \mathbf{0}$ which in turn implies that $\Delta X_1(z) \equiv 0$. The proof follows from Theorem 3.1. \square

D.2 Proof of Theorem 5.1

First, since u is a white Gaussian noise with a nonzero variance, u^n is also a white noise with a nonzero variance. To prove Theorem 5.1, we will prove that the power spectrum $\Phi_{\mathbf{u}}(\omega)$ is positive definite at almost all ω and so we will conclude the proof by virtue of Lemma 2.1. To show that $\Phi_{\mathbf{u}}(\omega)$ is positive definite at almost all ω , we will derive a formal expression of $\Phi_{\mathbf{u}}(\omega)$ and then to prove that its determinant is nonzero for almost all ω . For this purpose, we will need the expectation value of $u^n(t)$ when $u(t)$ is zero-mean Gaussian. It is given in the next lemma.

Lemma D.1. [PP02, Chapter 5]. *Consider a zero-mean Gaussian variable X and denote by σ^2 its variance. Then, for all $n \in \mathbb{N}^*$,*

$$E[X^n] = \begin{cases} \sigma^n(n-1)!! & \text{when } n \text{ is even.} \\ 0 & \text{when } n \text{ is odd.} \end{cases}$$

where the operator $!!$ is defined for odd integer p by $p!! = p \times (p-2) \times (p-4) \times \dots \times 3 \times 1$.

Let us now calculate $\Phi_{\mathbf{u}}(\omega)$ by taking the Fourier transform of the correlation matrix $\mathbf{R}_{\mathbf{u}}(\tau)$ given by

$$\mathbf{R}_{\mathbf{u}}(\tau) = \begin{pmatrix} \bar{E}[u(t)u(t-\tau)] & \bar{E}[u(t)u^n(t-\tau)] \\ \bar{E}[u^n(t)u(t-\tau)] & \bar{E}[u^n(t)u^n(t-\tau)] \end{pmatrix}$$

Since u and u^n are white noises and so stationary signals, the operator \bar{E} can be replaced by the expectation operator E in the above correlation matrix expression. Let us denote $\sigma_u^2 = E[u^2(t)]$. By using the fact that u and u^n are white and that u is zero-mean Gaussian, we have that, from Lemma D.1,

$$\begin{aligned} E[u(t)u(t-\tau)] &= \sigma_u^2 \delta(\tau) \\ E[u(t)u^n(t-\tau)] &= \begin{cases} 0 & \text{when } n \text{ is even.} \\ n!! \sigma_u^{n+1} \delta(\tau) & \text{when } n \text{ is odd.} \end{cases} \\ E[u^n(t)u(t-\tau)] &= E[u(t)u^n(t+\tau)] \end{aligned}$$

For the calculation of $E[u^n(t)u^n(t - \tau)]$, we give here the details since it is not as simple as the previous ones.

First, when $\tau \neq 0$, we have that $E[u^n(t)u^n(t - \tau)] = E[u^n(t)]E[u^n(t - \tau)]$ since u^n is a white noise. If n is odd, $E[u^n(t)] = E[u^n(t - \tau)] = 0$ from Lemma D.1 and so $E[u^n(t)u^n(t - \tau)] = 0$. If n is even, then $E[u^n(t)] = E[u^n(t - \tau)] = \sigma_u^n(n - 1)!!$ from Lemma D.1 and so $E[u^n(t)u^n(t - \tau)] = (\sigma_u^n(n - 1)!!)^2 = \sigma_u^{2n}((n - 1)!!)^2$.

For $\tau = 0$, we have that $E[u^n(t)u^n(t - \tau)] = E[u^{2n}(t)] = \sigma_u^{2n}(2n - 1)!!$ from Lemma D.1.

We deduce the expression for $E[u^n(t)u^n(t - \tau)]$:

- when n is even

$$E[u^n(t)u^n(t - \tau)] = \sigma_u^{2n} [(2n - 1)!! - ((n - 1)!!)^2] \delta(\tau) + \sigma_u^{2n} ((n - 1)!!)^2$$

- when n is odd

$$E[u^n(t)u^n(t - \tau)] = \sigma_u^{2n}(2n - 1)!!\delta(\tau)$$

where δ is the Dirac distribution.

By taking the Fourier transform of $\mathbf{R}_u(\tau)$, the power spectrum matrix Φ_u is given by

$$2\pi \begin{cases} \begin{pmatrix} 0 & 0 \\ 0 & \sigma_u^{2n}(2n - 1)!! \end{pmatrix} \delta(\omega) + \begin{pmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_u^{2n} [(2n - 1)!! - ((n - 1)!!)^2] \end{pmatrix} & \text{when } n \text{ is even.} \\ \begin{pmatrix} \sigma_u^2 & \sigma_u^{n+1}n!! \\ \sigma_u^{n+1}n!! & \sigma_u^{2n}(2n - 1)!! \end{pmatrix} & \text{when } n \text{ is odd.} \end{cases}$$

Since the power spectrum matrix $\Phi_u(\omega)$ is positive semidefinite at all frequencies ω , let us prove that the determinant of $\Phi_u(\omega)$ is nonzero for almost all ω to prove that it is strictly positive definite at almost all frequencies ω . When

- n is even, for all frequencies $\omega \neq 0$

$$\det(\Phi_u(\omega)) = 4\pi^2 \sigma_u^{2n+2} [(2n - 1)!! - ((n - 1)!!)^2]$$

- n is odd, for all frequencies ω

$$\det(\Phi_u(\omega)) = 4\pi^2 \sigma_u^{2n+2} [(2n - 1)!! - (n!!)^2]$$

Let us prove that $(2n-1)!! - ((n-1)!!)^2 \neq 0$ by proving that $(2n-1)!! > ((n-1)!!)^2$ when n is nonzero and even. The proof that $(2n-1)!! > (n!)^2$ when n is odd is based on the same principle. Let us first observe that

$$\begin{aligned} \frac{(2n-1)!!}{((n-1)!!)^2} &= \frac{(2n-1) \times (2n-3) \times \cdots \times (n+1) \times (n-1)!!}{((n-1)!!)^2} \\ &= \frac{(2n-1) \times (2n-3) \times \cdots \times (n+1)}{(n-1) \times \cdots \times 3 \times 1} \end{aligned}$$

The numerator and the denominator of the latter are the product of $n/2$ factors. Since the minimal factor of the numerator is strictly greater than the maximal factor of the denominator for $n \geq 2$, then $(2n-1)!! > ((n-1)!!)^2$. Therefore, $\det(\Phi_{\mathbf{u}}(\omega)) > 0$ for almost all ω . With Lemma 2.1, the conclusion follows. \square

Appendix for Chapter 6

Contents

E.1	Inertial effects for objects in rotation	201
E.1.1	Coriolis effect	201
E.1.2	Centrifugal effect	202
E.1.3	Euler effect	202

E.1 Inertial effects for objects in rotation

Let us consider an object M with a mass m in motion in a frame of reference $(0, x, y, z)$ that is rotating w.r.t. a Galilean frame of reference around the z -axis with an angular rate vector $\vec{\Omega}$. This situation is depicted in Figure E.1. The object M is then subjected to three inertial forces that will be described in the following paragraphs.

E.1.1 Coriolis effect

The first force is the Coriolis one denoted \vec{f}_{cor} and it is given by

$$\vec{f}_{cor}(t) = 2m\vec{\Omega}(t) \wedge \frac{d\vec{OM}}{dt}(t)$$

where \wedge is the vector product.

The Coriolis force has a significant effect on heavy objects with high speed. For instance, this phenomenon explains the difference of the rotation direction of cyclones between the North and the South hemisphere. This effect is also responsible of the faster wear of one rail of a railway for high-speed trains moving along a North-South axis. In the North hemisphere, the East rail is the one that wears the fastest. For the MEMS gyroscope, since both masses are not heavy, the gyroscope is designed such

that both masses have lightly damped resonant dynamics in order to increase the term $d|\overrightarrow{OM}(t)|/dt$ which in turn increases the Coriolis force on the sense mass.

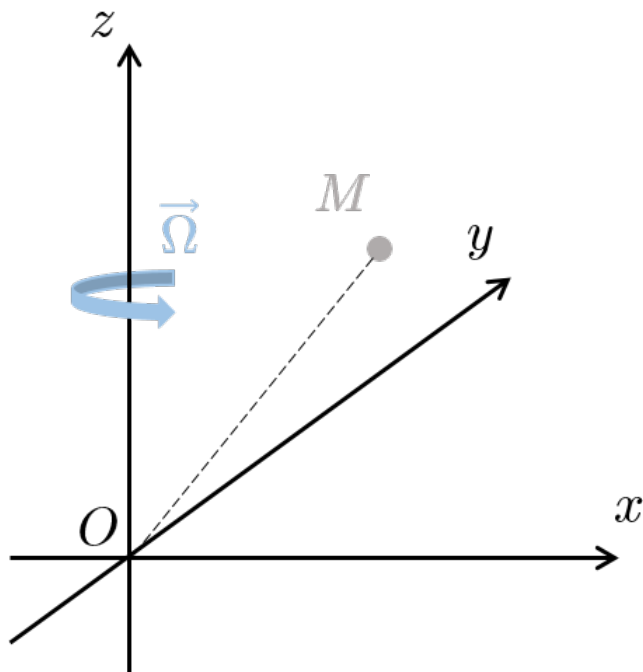


Figure E.1: An object M moving into a rotating frame of reference.

E.1.2 Centrifugal effect

The second force is the centrifugal one denoted \vec{f}_{cen} and it is given by

$$\vec{f}_{cen}(t) = m \vec{\Omega}(t) \wedge \vec{\Omega}(t) \wedge \overrightarrow{OM}(t)$$

E.1.3 Euler effect

The last inertial effect is the Euler one denoted \vec{f}_{eul} and it is given by

$$\vec{f}_{eul}(t) = m \frac{d\vec{\Omega}}{dt}(t) \wedge \overrightarrow{OM}(t)$$

The centrifugal and Euler effects are negligible for small displacements $|\overrightarrow{OM}(t)|$. It is the case for the MEMS gyroscope since $|\overrightarrow{OM}(t)|$ has an magnitude order of 10^{-6} m.

Appendix for Chapter 7

Contents

F.1	Comparison between SISO-by-SISO approach and MISO identification	204
F.2	Hankel-singular-value-based reduction	208
F.2.1	Order 8	208
F.2.2	Order 6	210
F.3	Initialization method for the grey-box modeling	213
F.3.1	Transformation of the state-space model into a transfer function matrix	213
F.3.2	Considered method for the initialization	215
F.4	Bode of the tracking controller $\mathbf{K}(z)$	216

F.1 Comparison between SISO-by-SISO approach and MISO identification

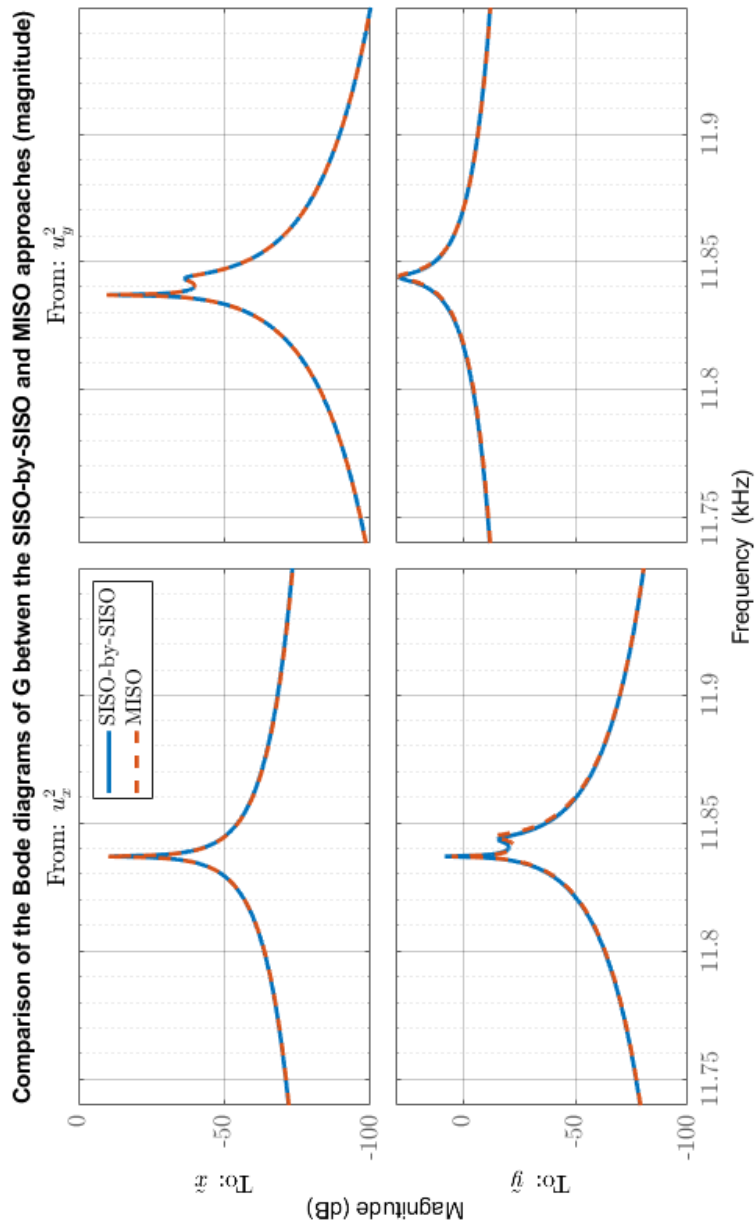


Figure F.1: Bode diagram magnitudes of the identified \mathbf{G} from the SISO-by-SISO approach (solid blue line) and the identified \mathbf{G} from the MISO identifications (orange dashed line).

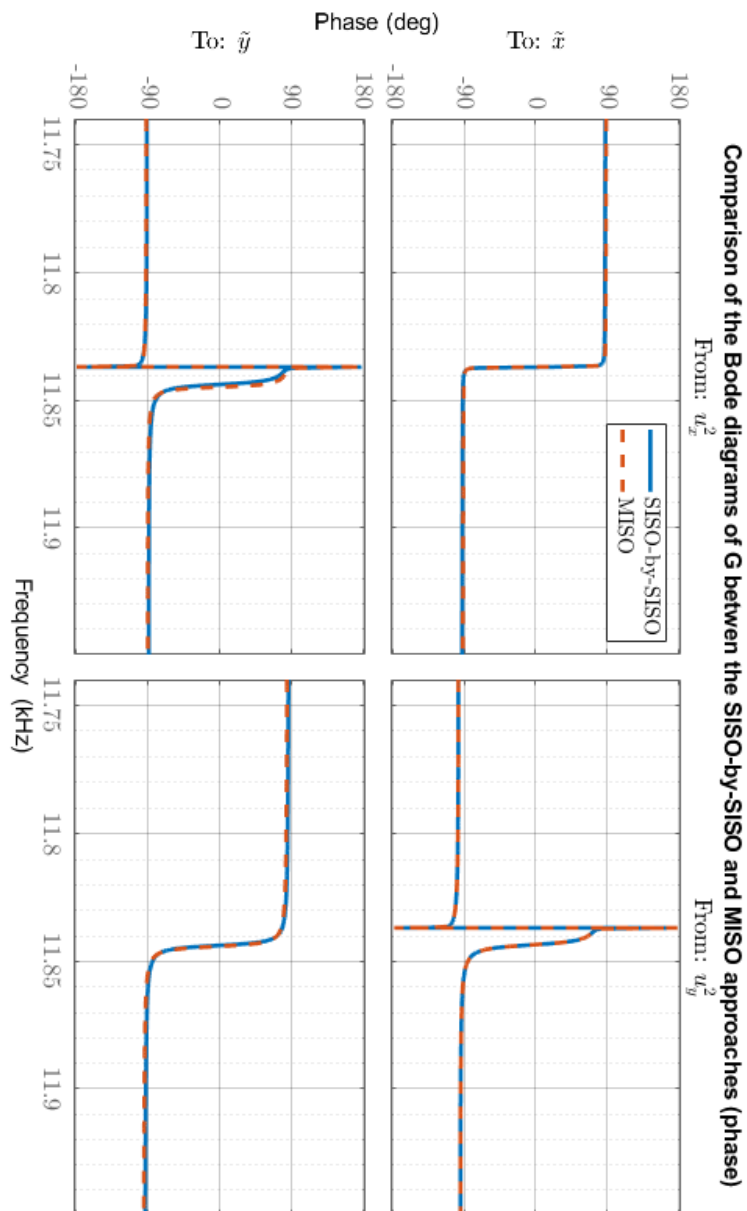


Figure F.2: Bode diagram phases of the identified \mathbf{G} from the SISO-by-SISO approach (solid blue line) and the identified \mathbf{G} from the MISO identifications (orange dashed line).

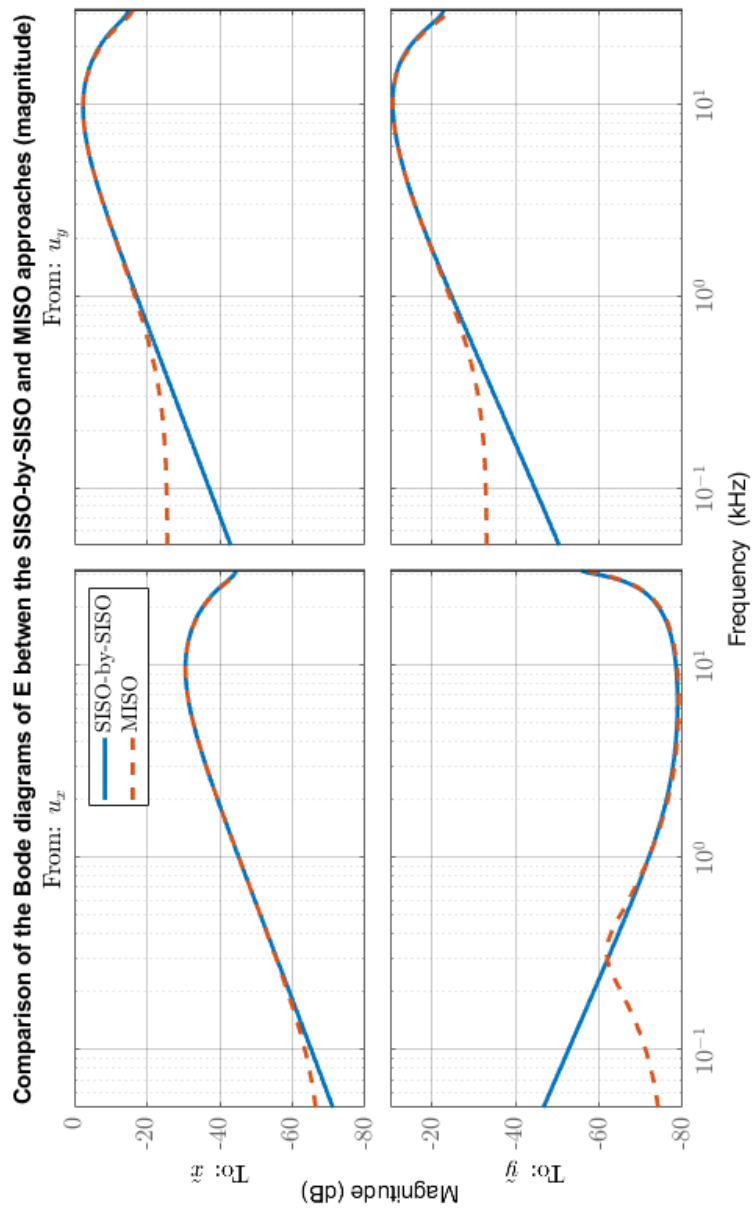


Figure F.3: Bode diagram magnitudes of the identified \mathbf{Q} from the SISO-by-SISO approach (solid blue line) and the identified \mathbf{Q} from the MISO identifications (orange dashed line).

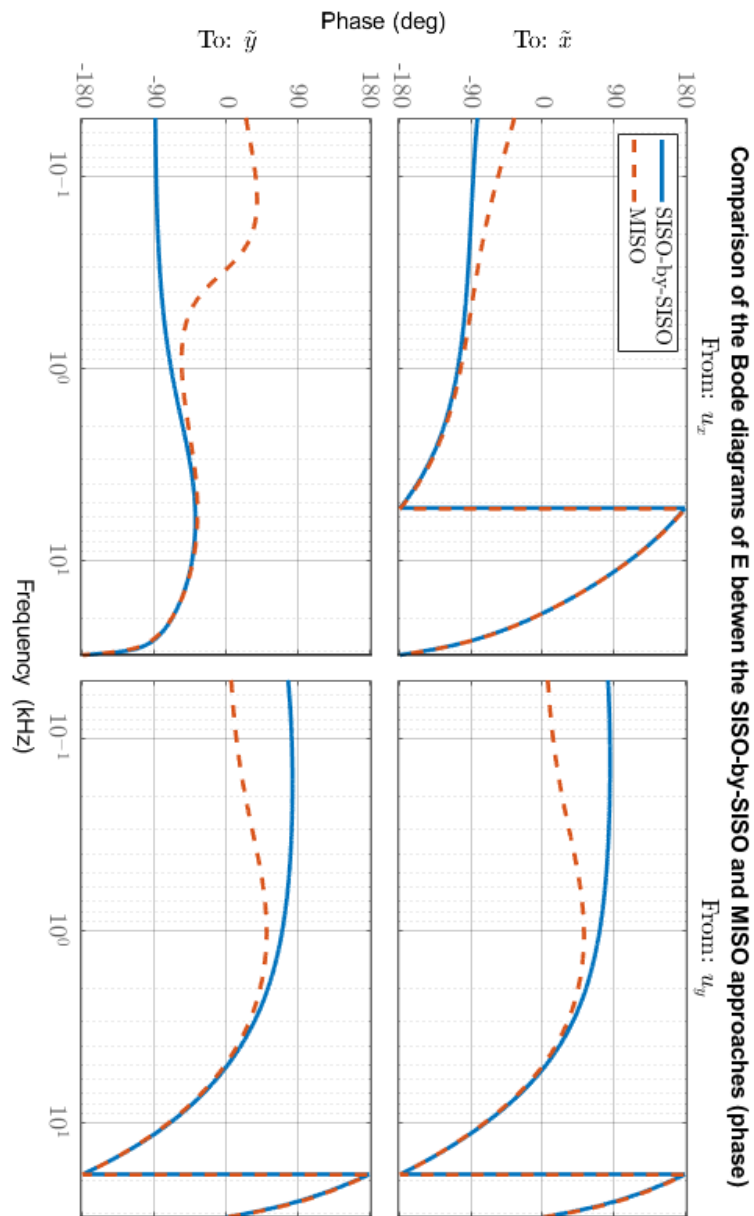


Figure F.4: Bode diagram phases of the identified \mathbf{Q} from the SISO-by-SISO approach (solid blue line) and the identified \mathbf{Q} from the MISO identifications (orange dashed line).

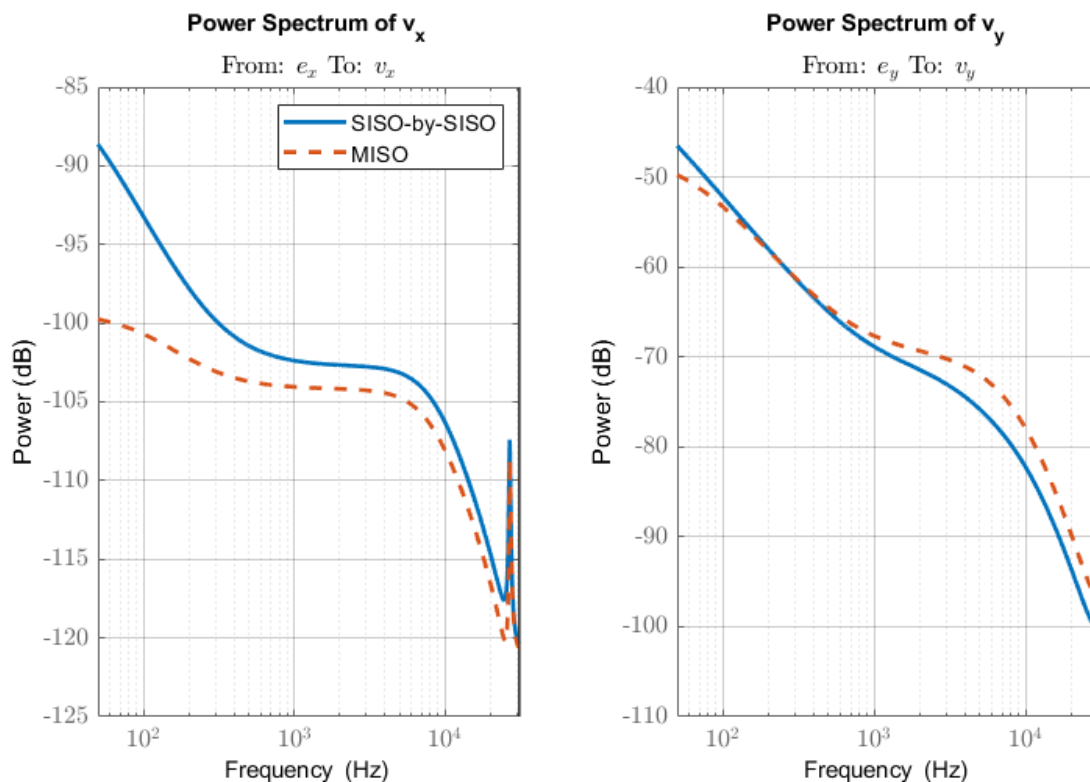


Figure F.5: Comparison of the DSPs of the measurement noises v_x (left) and v_y (right) between the SISO-by-SISO approach (blue solid line) from the MISO identifications (orange dotted line).

F.2 Hankel-singular-value-based reduction

In this method, we compute the Hankel singular values of the balanced state space of $\mathbf{G}(z, \tilde{\theta}_N)$ and we truncate the states by keeping the ones with the highest Hankel singular values (corresponding to the states with the highest energy). We have tried different truncation order: 10, 8, 6 and 4. For the last three truncation orders, we observed some mismatches between the identified black-box model and the Hankel-singular-value-based reduced-order models. In the next two paragraphs, we show the results obtained for a truncation order equal to 8 and 6.

F.2.1 Order 8

The Figures F.6 and F.7 give the Bode diagram magnitudes and phases respectively of the reduced-order model based on Hankel singular values of order 8 and the identified black-box model of order 12. For the truncation order 8, we observe that both models have the same Bode diagrams for the direct transfers and for G_{xy} . However, a gain and phase mismatch around the resonance frequencies can be seen for G_{xy} .

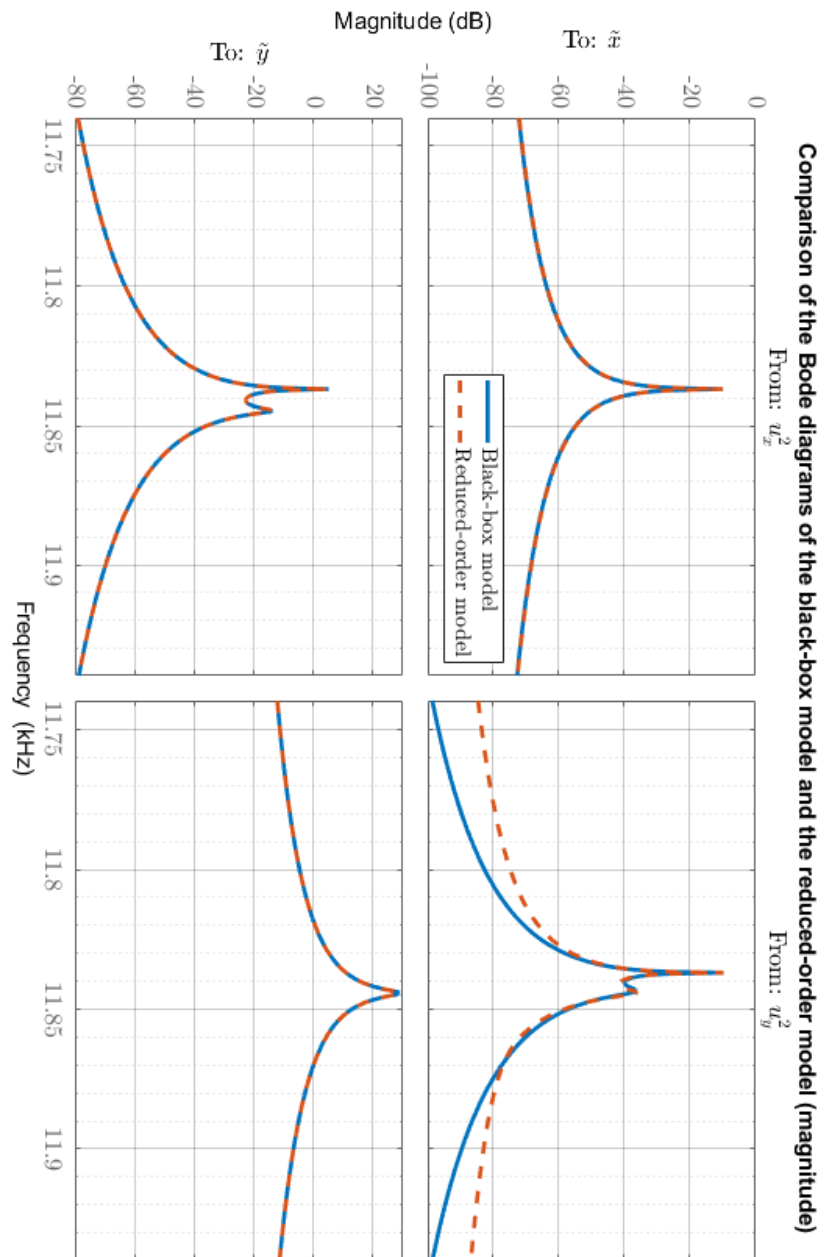


Figure F.6: Bode diagram magnitudes of the black-box model (blue solid line) and of the reduced-order model based on Hankel singular values of order 8 (orange dotted line).

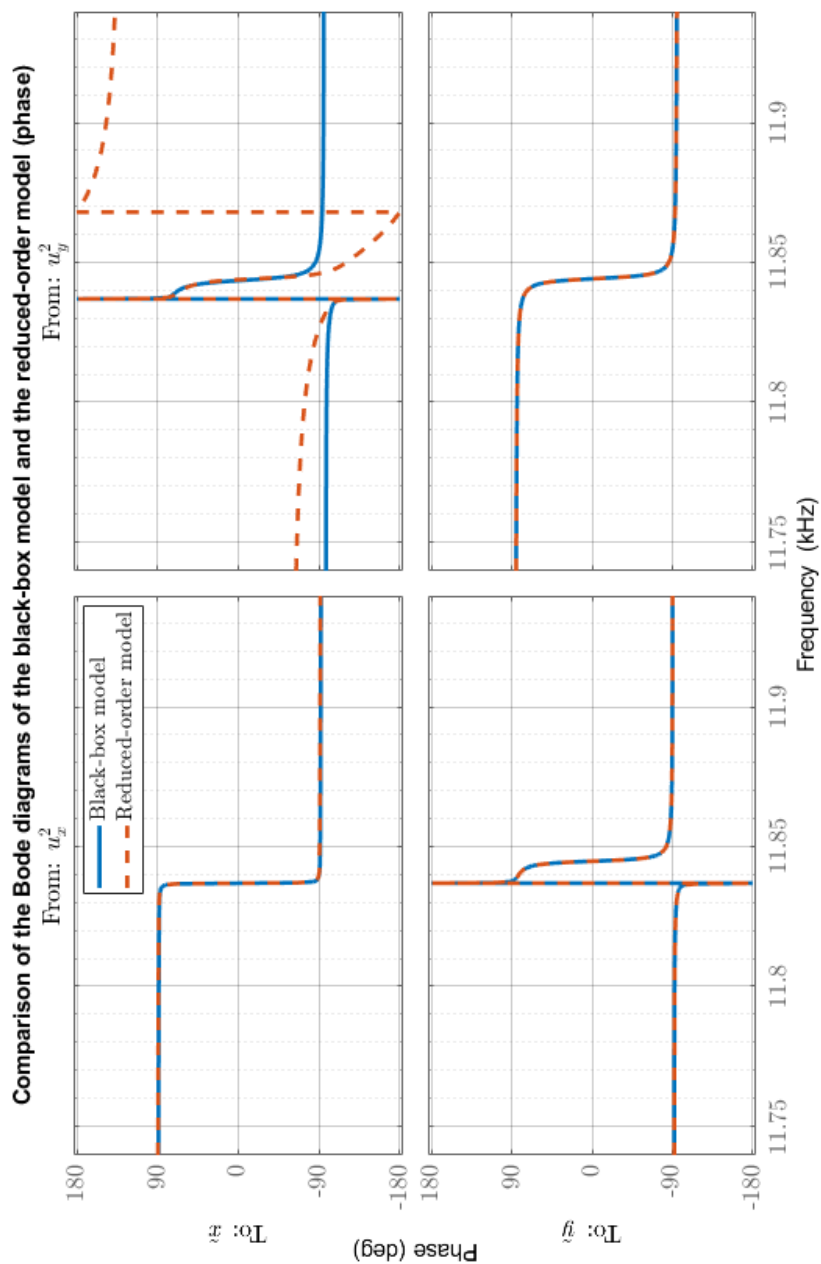


Figure F.7: Bode diagram phases of the black-box model (blue solid line) and of the reduced-order model based on Hankel singular values of order 8 (orange dotted line).

F.2.2 Order 6

The Figures F.8 and F.9 give the Bode diagram magnitudes and phases respectively of the reduced-order model based on Hankel singular values of order 6 and the identified black-box model of order 12. For the truncation order 6, the Bode diagram of both transfers G_{xy} and G_{yx} have large discrepancies around the resonance frequencies.

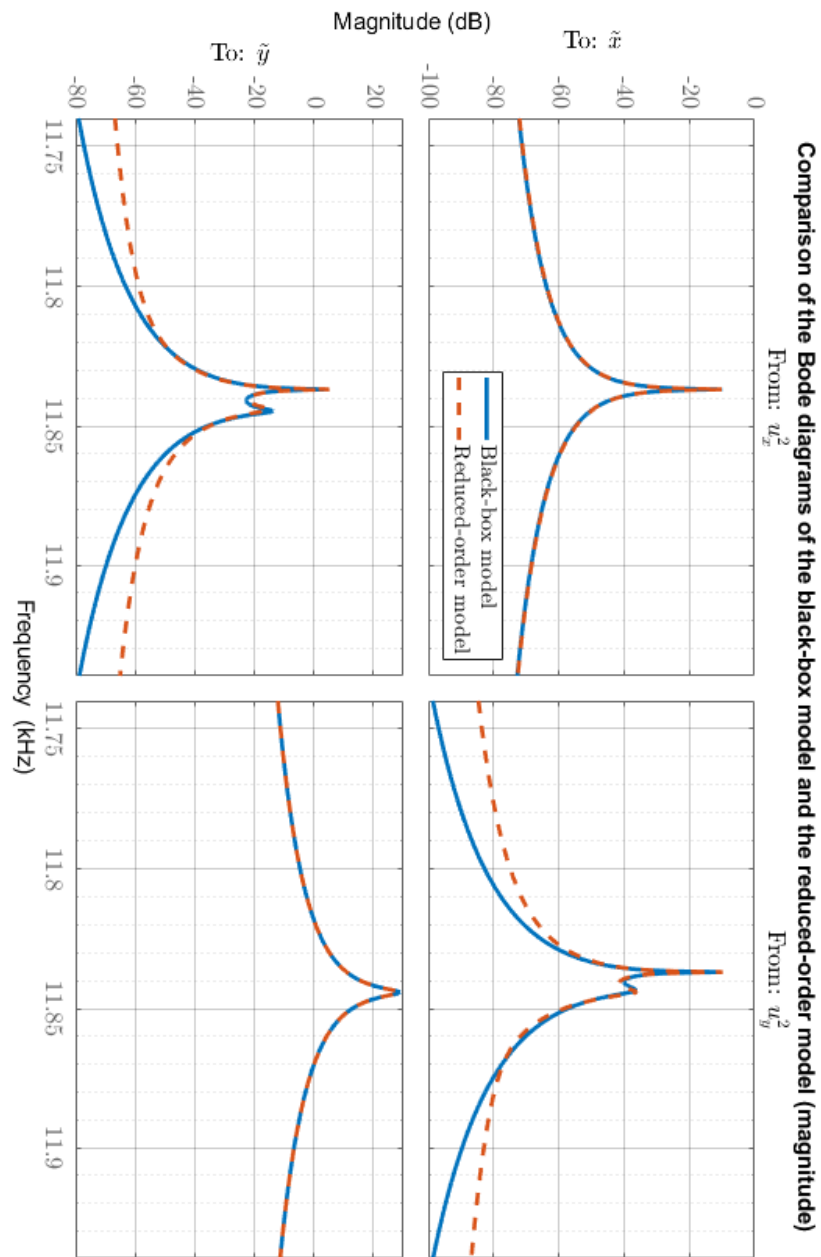


Figure F.8: Bode diagram magnitudes of the identified black-box model (blue solid line) and of the reduced-order model based on Hankel singular values of order 6 (orange dotted line).

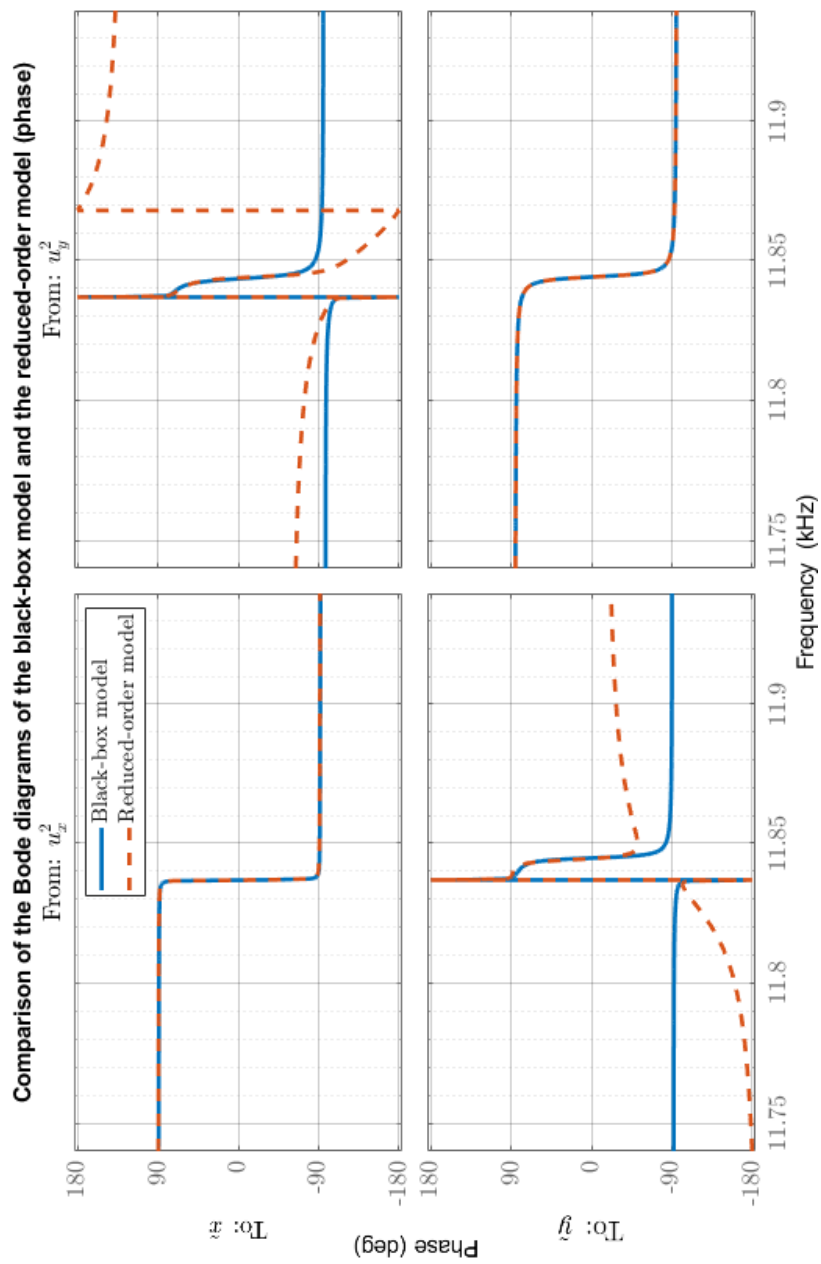


Figure F.9: Comparison of Bode diagram phases of the identified black-box model (blue solid line) and of the reduced-order model based on Hankel singular values of order 6 (orange dotted line).

F.3 Initialization method for the grey-box modeling

The idea for the initialization of the grey-box identification is to find the grey-box parameters that match the properties of the identified resonances in $\hat{\mathbf{G}}$.

F.3.1 Transformation of the state-space model into a transfer function matrix

Let us rewrite the state space model given in (7.18) into a continuous time transfer function matrix $\mathbf{G}_{gb} = \mathbf{G}_{mea}\mathbf{G}_{mec}$ where \mathbf{G}_{mea} and \mathbf{G}_{mec} are the continuous time transfer function matrices such that

- the input of \mathbf{G}_{mea} is $(x_{mec}, y_{mec})^T$ and its output is $(\tilde{x}_{mec}, \tilde{y}_{mec})^T$. This transfer function matrix is obtained by transforming the measurement equations (7.17) into the Laplace domain.
- the input of \mathbf{G}_{mec} is $(u_x^2, u_y^2)^T$ and its output is $(x_{mec}, y_{mec})^T$. This transfer function matrix is obtained by transforming the mechanical equations (7.15) (coupled with the actuation) into the Laplace domain.

Let us denote by p the Laplace variable and define the four entries of \mathbf{G}_{gb} as follows

$$\mathbf{G}_{gb}(p) = \begin{pmatrix} G_{gb,xx}(p) & G_{gb,xy}(p) \\ G_{gb,yx}(p) & G_{gb,yy}(p) \end{pmatrix}$$

By using the Laplace variable on the state-space equations, we have that

$$\mathbf{G}_{mea}(p) = \begin{pmatrix} \frac{\beta_x}{\frac{p}{\lambda_x} + 1} & 0 \\ 0 & \frac{\beta_y}{\frac{p}{\lambda_y} + 1} \end{pmatrix}$$

and

$$\mathbf{G}_{mec}(p) = \begin{pmatrix} G_{mec,xx}(p) & G_{mec,xy}(p) \\ G_{mec,yx}(p) & G_{mec,yy}(p) \end{pmatrix}$$

where

$$\begin{aligned}
 G_{mec,xx}(p) &= \frac{\bar{\alpha}_x(p^2 + \bar{d}_{yy}p + \bar{k}_{yy})}{(p^2 + \bar{d}_{yy}p + \bar{k}_{yy})(p^2 + \bar{d}_{xx}p + \bar{k}_{xx}) - (\bar{d}_{yx}p + \bar{k}_{yx})(\bar{d}_{xy}p + \bar{k}_{xy})} \\
 G_{mec,xy}(p) &= \frac{-\bar{\alpha}_y(\bar{d}_{xy}p + \bar{k}_{xy})}{(p^2 + \bar{d}_{yy}p + \bar{k}_{yy})(p^2 + \bar{d}_{xx}p + \bar{k}_{xx}) - (\bar{d}_{yx}p + \bar{k}_{yx})(\bar{d}_{xy}p + \bar{k}_{xy})} \\
 G_{mec,yx}(p) &= \frac{-\bar{\alpha}_x(\bar{d}_{yx}p + \bar{k}_{yx})}{(p^2 + \bar{d}_{yy}p + \bar{k}_{yy})(p^2 + \bar{d}_{xx}p + \bar{k}_{xx}) - (\bar{d}_{yx}p + \bar{k}_{yx})(\bar{d}_{xy}p + \bar{k}_{xy})} \\
 G_{mec,yy}(p) &= \frac{\bar{\alpha}_y(p^2 + \bar{d}_{xx}p + \bar{k}_{xx})}{(p^2 + \bar{d}_{yy}p + \bar{k}_{yy})(p^2 + \bar{d}_{xx}p + \bar{k}_{xx}) - (\bar{d}_{yx}p + \bar{k}_{yx})(\bar{d}_{xy}p + \bar{k}_{xy})}
 \end{aligned}$$

Let us denote by $d_{mec}(p)$ the common denominator of the transfer functions in $\mathbf{G}_{mec}(p)$. We observe that each transfer function in $\mathbf{G}_{gb}(p)$ has an order of 5 and should have two resonances (since $1/d_{mec}(p)$ does). An additional anti-resonance should also be observed for the transfer functions $\mathbf{G}_{gb,xx}(p)$ and $\mathbf{G}_{gb,yy}(p)$, which is not the case in practice. To circumvent this problem, for the initialization determination, we will assume that the common denominator d_{mec} is such that

$$d_{mec}(p) \approx (p^2 + \bar{d}_{yy}p + \bar{k}_{yy})(p^2 + \bar{d}_{xx}p + \bar{k}_{xx}) \quad (\text{F.1})$$

Therefore, with this assumption, the entries of $\mathbf{G}_{mec}(p)$ become

$$\begin{aligned}
 G_{mec,xx}(p) &= \frac{\bar{\alpha}_x}{p^2 + \bar{d}_{xx}p + \bar{k}_{xx}} \\
 G_{mec,xy}(p) &= \frac{-\bar{\alpha}_y(\bar{d}_{xy}p + \bar{k}_{xy})}{(p^2 + \bar{d}_{xx}p + \bar{k}_{xx})(p^2 + \bar{d}_{yy}p + \bar{k}_{yy})} \\
 G_{mec,yx}(p) &= \frac{-\bar{\alpha}_x(\bar{d}_{yx}p + \bar{k}_{yx})}{(p^2 + \bar{d}_{xx}p + \bar{k}_{xx})(p^2 + \bar{d}_{yy}p + \bar{k}_{yy})} \\
 G_{mec,yy}(p) &= \frac{\bar{\alpha}_y}{p^2 + \bar{d}_{yy}p + \bar{k}_{yy}}
 \end{aligned}$$

In this case, $\mathbf{G}_{gb,xx}(p)$ and $\mathbf{G}_{gb,yy}(p)$ have only one resonance and do not have an anti-resonance. This fact and the denominator assumption are verified in our black-box identification since we observe only one resonance in the identified transfer functions $G_{xx}(z, \hat{\boldsymbol{\theta}}_{xx,N})$ and $G_{yy}(z, \hat{\boldsymbol{\theta}}_{yy,N})$ and there is no anti-resonance. We then deduce an

expression for the entries of $\mathbf{G}_{gb}(p)$ with the denominator assumption given in in (F.1)

$$\begin{aligned}
 G_{gb,xx}(p) &= \frac{\bar{\alpha}_x \beta_x}{(p^2 + \bar{d}_{xx}p + \bar{k}_{xx}) \left(\frac{p}{\lambda_x} + 1 \right)} \\
 G_{gb,xy}(p) &= \frac{-\bar{\alpha}_y \beta_x (\bar{d}_{xy}p + \bar{k}_{xy})}{(p^2 + \bar{d}_{xx}p + \bar{k}_{xx})(p^2 + \bar{d}_{yy}p + \bar{k}_{yy}) \left(\frac{p}{\lambda_x} + 1 \right)} \\
 G_{gb,yx}(p) &= \frac{-\bar{\alpha}_x \beta_y (\bar{d}_{yx}p + \bar{k}_{yx})}{(p^2 + \bar{d}_{xx}p + \bar{k}_{xx})(p^2 + \bar{d}_{yy}p + \bar{k}_{yy}) \left(\frac{p}{\lambda_y} + 1 \right)} \\
 G_{gb,yy}(p) &= \frac{\bar{\alpha}_y \beta_y}{(p^2 + \bar{d}_{yy}p + \bar{k}_{yy}) \left(\frac{p}{\lambda_y} + 1 \right)}
 \end{aligned}$$

The next step should be to convert the continuous-time transfer function matrix $\mathbf{G}_{gb}(p)$ into a discrete-time transfer function and then match the coefficients with the ones of the identified transfer functions in $\hat{\mathbf{G}}(z)$. However, the unstable zeros of the identified transfers $G_{xx}(z, \hat{\boldsymbol{\theta}}_{N,xx})$ and $G_{yy}(z, \hat{\boldsymbol{\theta}}_{N,yy})$ are not described by the grey-box model structure which prevents us to determine the grey-box parameters by matching the polynomial coefficients.

F.3.2 Considered method for the initialization

The solution that we will consider is to find the grey-box parameters that match the properties of the identified transfer functions in $\hat{\mathbf{G}}$ such as the static gain, the resonance frequencies and the quality factors.

- resonance frequencies: from the grey-box model structures and due to the fact that the quality factors are high, we have that $\omega_{0,x} \approx \sqrt{\bar{k}_{xx}}$ and $\omega_{0,y} \approx \sqrt{\bar{k}_{yy}}$. We then deduce an initial value for \bar{k}_{xx} and \bar{k}_{yy} from the resonance frequencies identified from $\hat{\mathbf{G}}$.
- quality factors: from the identified resonances of $G_{xx}(z, \hat{\boldsymbol{\theta}}_{N,xx})$ and $G_{yy}(z, \hat{\boldsymbol{\theta}}_{N,yy})$ and the corresponding grey-box transfer functions $G_{gb,xx}(p)$ and $G_{gb,yy}(p)$, we have the following relations for the quality factors q_x for the drive mode and q_y for the sense mode:

$$\begin{cases} q_x &= \frac{\bar{k}_{xx}}{\bar{d}_{xx}\omega_{0,x}} \\ q_y &= \frac{\bar{k}_{yy}}{\bar{d}_{yy}\omega_{0,y}} \end{cases}$$

We deduce an initial estimate for \bar{d}_{xx} and \bar{d}_{yy} .

- static gain: from the grey-box model structure and by denoting by a_{xx} , a_{xy} , a_{yx} and a_{yy} the static gains of the identified transfer functions $G_{xx}(z, \hat{\boldsymbol{\theta}}_{N,xx})$, $G_{xy}(z, \hat{\boldsymbol{\theta}}_{N,xy})$, $G_{yx}(z, \hat{\boldsymbol{\theta}}_{N,yx})$ and $G_{yy}(z, \hat{\boldsymbol{\theta}}_{N,yy})$ respectively, we should have that

$$\left\{ \begin{array}{l} a_{xx} = \frac{\bar{\alpha}_x \beta_x}{\bar{k}_{xx}} \\ a_{xy} = -\frac{\bar{\alpha}_y \beta_x \bar{k}_{xy}}{\bar{k}_{xx} \bar{k}_{yy}} \\ a_{yx} = -\frac{\bar{\alpha}_x \beta_y \bar{k}_{yx}}{\bar{k}_{xx} \bar{k}_{yy}} \\ a_{yy} = \frac{\bar{\alpha}_y \beta_y}{\bar{k}_{yy}} \end{array} \right.$$

Unfortunately, we can find an infinite number of possibilities for the six parameters $\bar{\alpha}_x$, $\bar{\alpha}_y$, β_x , β_y , \bar{k}_{xy} and \bar{k}_{yx} . To solve the problem, we will fix two grey-box parameters. In our case, we will fix β_x and β_y since we know that the displacements of both masses have an order of magnitude of 10^{-6} m and the mechanical output voltage has an order of magnitude of 10^{-4} V for the drive mode and 1 V for the sense mode. Therefore, we will fix $\beta_x = 10^2$ V/m and $\beta_y = 10^6$ V/m. From these fixed parameters and the identified static gains a_{xx} , a_{xy} , a_{yx} and a_{yy} , we can uniquely deduce the parameters $\bar{\alpha}_x$, $\bar{\alpha}_y$, \bar{k}_{yx} and \bar{k}_{xy} .

- for the effect of the electronic card, we have to choose some initial values for both cut-off frequencies λ_x and λ_y . This first-order effect might come from the anti-aliasing filters with a cut-off frequency of $2\pi \times 15000$ rad/s. Therefore, we will choose $\lambda_x = \lambda_y = 2\pi \times 15000$.
- finally, to estimate \bar{d}_{yx} , we choose one frequency ω' and we compute the frequency response of the corresponding black-box model at this frequency, i.e., we compute $G_{yx}(e^{j\omega'}, \hat{\boldsymbol{\theta}}_{N,yx})$. Then, we find the value of \bar{d}_{yx} such that $G_{gb,yx}(e^{j\omega'}, \bar{d}_{yx}) = G_{yx}(e^{j\omega'}, \hat{\boldsymbol{\theta}}_{N,yx})$ holds. The same approach holds for \bar{d}_{xy} with the transfer functions $G_{gb,yx}$ and $G_{yx}(z, \hat{\boldsymbol{\theta}}_{N,yx})$. In both cases, the frequency ω' has been chosen equal to the identified resonance frequency $\omega_{0,x}$.

Remark F.1. *At this point, we can wonder if an optimization with **greyest** is useful since we can estimate initial values of the parameters that match some properties of the identified model $\hat{\mathbf{G}}$. However, for the cut-off frequencies λ_x and λ_y , we made an assumption for the values. Moreover, the determination of the parameters was based on an assumption for the denominator d_{mec} made in (F.1).*

F.4 Bode of the tracking controller $\mathbf{K}(z)$

In this section, we give the plots of the magnitude and the phase of the Bode diagram of the MIMO tracking controller $\mathbf{K}(z)$ in Figures F.10 and F.11 and respectively.

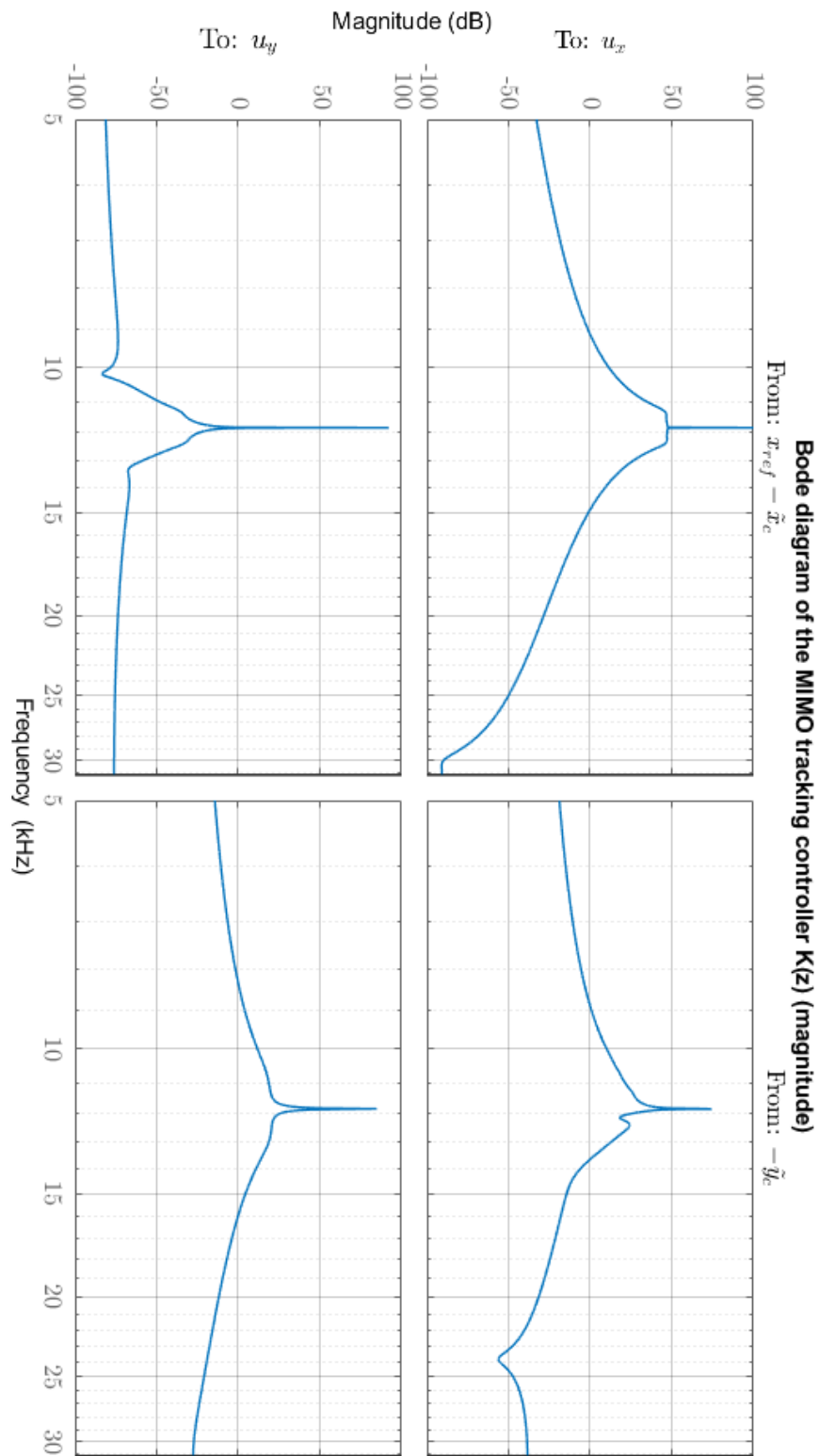


Figure F.10: Bode diagram magnitudes of the MIMO tracking controller $\mathbf{K}(z)$.

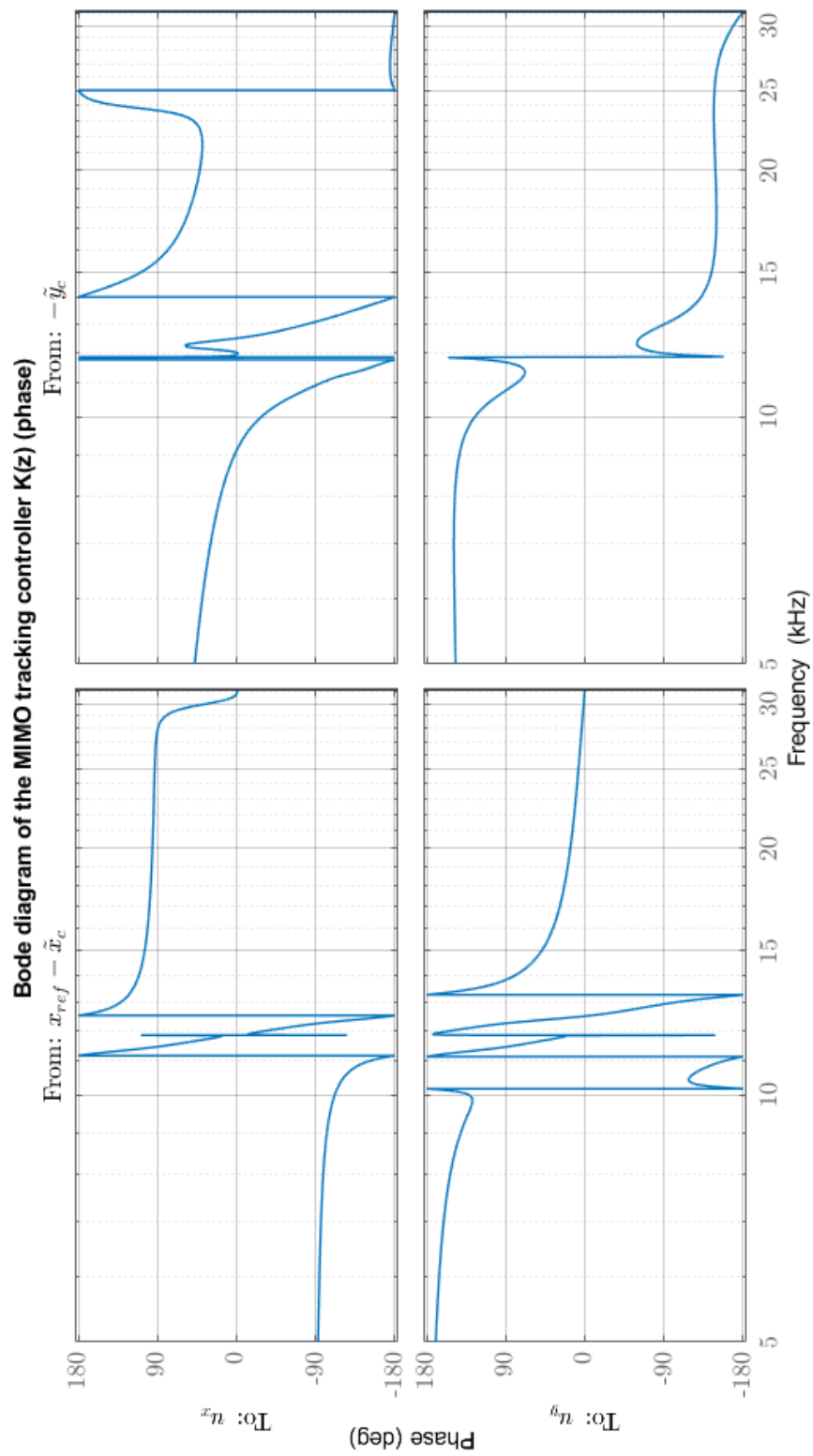


Figure F.11: Bode diagram phases of the MIMO tracking controller $\mathbf{K}(z)$.

Long résumé de la thèse

Long résumé : Introduction de la thèse

Contents

G.1	Identification des systèmes	221
G.2	Consistance d'un estimateur pour l'identification des systèmes	223
G.3	État de l'art sur l'informativité des données et premières questions de recherche	224
G.4	Et pour les systèmes non-linéaires identifiables avec la méthode Prediction Error?	226
G.5	Étude pratique de la thèse : gyroscope MEMS	227
G.5.1	Présentation du gyroscope MEMS et des défis de contrôle	227
G.5.2	Objectives et présentation du projet Next4MEMS	229
G.5.3	État de l'art sur l'identification des gyroscopes MEMS	230
G.6	Organisation du long résumé des contributions	232

Dans ce chapitre, nous résumons l'introduction de la thèse.

G.1 Identification des systèmes

Les modèles mathématiques des dynamiques d'un système ont une place cruciale dans beaucoup de domaines des sciences de l'ingénieur tels que la chimie, la mécanique, l'électromagnétisme, etc. Ces modèles ont une multitude de fonctions :

- synthèse de correcteur : dans beaucoup d'applications telles que la mécanique, les dynamiques d'un système doivent satisfaire des performances telles que la stabilité, la rapidité, etc. Pour cela, un correcteur est synthétisé pour changer

le comportement du système afin qu'il ait le comportement souhaité. Ainsi, un modèle mathématique est nécessaire pour faire la synthèse de ce correcteur.

- prédiction : quand les dynamiques futures d'un système doivent être prédites, un modèle peut servir pour faire cette prédiction. Un exemple est la prédiction météo [Vas01].
- diagnostic : pour assurer la fiabilité d'un système, nous cherchons à détecter les pannes que le système peut subir. Dans ce cas là, il s'agit de détecter toutes déviations du comportement du système par rapport au comportement souhaité (nominal). Pour faire cette comparaison, nous pouvons utiliser un modèle du comportement nominal du système [Din08].
- rétro-ingénierie : un modèle peut aussi être utilisé pour déterminer des dynamiques intrinsèques au système qui sont inaccessibles pour l'utilisateur. C'est par exemple le cas de l'étude des rétroactions entre les composants d'une cellule en biologie [Kha08, TA05].

Les performances du correcteur, de la prédiction, etc dépendent souvent de la qualité du modèle. Il est alors très important de garantir une modélisation précise des dynamiques du système.

Il existe trois classes de méthodes de modélisation. Pour la première classe, le modèle est obtenu uniquement à partir des équations de la physique (e.g., les lois de la mécanique de Newton, les lois de l'électromagnétisme de Maxwell, etc). Cette classe de modélisation s'appelle modélisation boîte blanche. À l'opposé, dans la seconde classe de méthodes de modélisation appelée boîte noire, le modèle est complètement déterminé à partir d'un jeu de données expérimentales collectées sur le système. Enfin, la dernière classe de méthodes de modélisation est un mélange des deux autres classes précédentes dont le nom est boîte grise. Ici, la modélisation s'appuie sur les équations de la physique dont certains paramètres inconnus sont estimés à partir d'un jeu de données expérimentales.

L'identification des systèmes est le domaine scientifique qui consiste à déterminer un modèle à partir de données expérimentales. Ce domaine regroupe les méthodes de modélisation boîte grise et boîte noire. Pour chaque identification, l'utilisateur doit choisir trois ingrédients. Le premier ingrédient est la structure de modèle qui est un ensemble de modèles candidats qui peuvent être, par exemple, boîte grise/boîte noire, paramétrés/non-paramétrés, linéaires/non-linéaires, etc. Le deuxième ingrédient est le jeu de données expérimentales récoltées. Le dernier ingrédient est le critère d'identification qui permet de classer tous les modèles candidats de la structure de modèle et de déterminer le modèle optimal qui explique aux mieux les données récoltées par rapport au critère d'identification. L'identification peut être faite en boucle ouverte ou en boucle fermée. Dans le dernier cas, il est souvent nécessaire d'ajouter un signal d'excitation externe à la loi de commande.

Pour garantir un modèle précis du système, il est important que la structure de modèle puisse modéliser toutes les dynamiques du système. Lorsque c'est le cas, il existe un modèle dans la structure de modèle qui décrit exactement toutes les dynamiques du système. L'objectif est alors de déterminer ce modèle, appelé *vrai modèle*. Lorsque la structure de modèle est paramétrée, le vecteur de paramètres lié au vrai modèle est appelé *vrai vecteur de paramètres*. Par conséquent, l'objectif de l'identification des systèmes dans ce cas est de déterminer ce vrai modèle (ou le vrai vecteur de paramètres pour les structures de modèle paramétrées).

Cependant, à cause de la présence de bruit dans les données récoltées, il est impossible de faire cette estimation de façon exacte puisque le nombre de données et la puissance d'excitation sont finies en pratique. Ainsi, le but est de déterminer l'estimée la plus précise possible. Pour cela, il est nécessaire de garantir une propriété de l'estimateur lié au critère d'identification choisi : sa consistance. Nous expliquons cette notion dans le prochain paragraphe.

G.2 Consistance d'un estimateur pour l'identification des systèmes

La consistance est la propriété fondamentale qui garantit que le modèle identifié (resp. le vecteur de paramètres identifié pour les structures de modèle paramétrées) converge vers le vrai modèle (resp. le vrai vecteur de paramètres pour les structures de modèle paramétrées) lorsque le nombre de données tend vers l'infini. Si cette propriété n'est pas garantie, alors le modèle identifié ne sera pas précis, même si nous augmentons le nombre de données. Elle dépend directement du choix des trois ingrédients pour l'identification.

Dans cette thèse, nous considérons la méthode d'identification de type boîte noire par erreur de prédiction (*Prediction Error* en anglais) avec un critère des moindres carrés pour le critère d'identification. Les structures de modèle considérées dans cette thèse sont linéaires temps invariant (*Linear Time Invariant* ou *LTI* en anglais) de type fonctions de transfert dont les coefficients polynomiaux des numérateurs et des dénominateurs n'ont pas forcément de signification physique. Dans cette méthode, le vecteur de paramètres (qui est constitué des coefficients polynomiaux mentionnés précédemment) identifié est la solution d'un problème d'optimisation d'un coût des moindres carrés sur la différence entre la sortie mesurée et le prédictor construit à partir des données et de la structure de modèle. Cette différence s'appelle l'erreur de prédiction. La consistance de l'estimateur de cette méthode a été étudiée premièrement pendant les années 1970 [Lju76a, Lju76b]. La plupart des travaux se sont intéressés au cas de l'identification directe en boucle fermée avec différentes hypothèses sur l'excitation et le correcteur [SGL75, SLG76, AG82, GLS77]. Cependant ces résultats ne donnent que des conditions suffisantes et non nécessaires.

Il faudra alors attendre les travaux de Ljung [Lju99] pour développer une condition nécessaire et suffisante pour la consistance : l'estimée est consistante si et seulement

si le vrai vecteur de paramètres est l'unique minimiseur du critère asymptotique des moindres carrés, i.e., le critère d'identification obtenu quand le nombre de données tend vers l'infini. Pour garantir l'unicité du minimiseur, il faut respecter deux conditions : l'identifiabilité de la structure de modèle et l'informativité des données. L'identifiabilité est la propriété qui garantit que deux vecteurs de paramètres différents ne donnent pas le même modèle. Elle a été largement étudiée et des structures de modèle linéaires avec des propriétés d'identifiabilité fortes ont été développées telles que Finite Impulse Response (FIR, *Réponse Impulsionnelle Finie* en français), ARX, ARMAX, Output-Error (OE) et Box-Jenkins (BJ).

L'informativité des données permet de garantir que deux modèles ne donnent pas le même prédictor. Combiné avec l'identifiabilité, deux vecteurs de paramètre ne donnent pas le même prédictor et donc un seul vecteur de paramètres minimise le critère d'identification asymptotique : c'est le vrai vecteur de paramètre. L'informativité est donc une propriété cruciale à garantir puisque, sans celle-ci, l'estimée obtenue serait alors imprécise. À l'inverse, lorsque l'informativité est garantie, nous pouvons atteindre n'importe quel niveau de précision sur les incertitudes des paramètres, à condition que le nombre de données puisse être augmenté. Du fait de son rôle crucial dans l'identification, nous nous focalisons sur cette propriété dans cette dissertation. Dans le prochain paragraphe, nous donnons un état de l'art de cette propriété pour la méthode par erreur de prédiction avec un critère des moindres carrés.

G.3 État de l'art sur l'informativité des données et premières questions de recherche

L'informativité des données a été largement étudiée pour l'identification des systèmes linéaires une-entrée une sortie (*single-input single-output* ou *SISO* en anglais) en boucle ouverte et boucle fermée avec les structures de modèle classiques (FIR, ARX, ARMAX, OE, BJ). Les conditions développées dans ce cas sont basées sur la notion de richesse du signal d'excitation utilisée pendant l'identification. Ce signal est riche d'ordre ν si et seulement si sa densité spectrale de puissance est non-nulle en ν pulsations de l'intervalle $]-\pi, \pi]$ [Lju99, SS89]. De cette notion de richesse ont été développées des conditions nécessaires et suffisantes pour l'informativité des données dans les cas boucle ouverte et boucle fermée dans [GBM08, GBBM09]. Plus précisément, ces travaux ont déterminé un ordre minimal de richesse, dépendant de la complexité de la structure de modèle, que le signal d'excitation doit respecter pour garantir l'informativité des données. En boucle fermée, cet ordre minimal (pour l'excitation externe dans ce cas là) dépend aussi de la complexité du correcteur qui contribue lui aussi à l'informativité des données via la rétroaction du bruit de mesure. Par ailleurs, si le correcteur est suffisamment complexe, l'informativité des données est alors garantie sans excitation externe (*costless identification* en anglais).

Alors que l'informativité des données est un sujet mature pour le cas SISO dans la littérature, nous ne pouvons pas dire de même pour les systèmes plusieurs-entrées

plusieurs-sorties (*multiple-inputs multiple-outputs* ou *MIMO* en anglais). Les études dans la littérature sur ce sujet sont principalement focalisées sur le développement de conditions pour vérifier qu'un correcteur MIMO est suffisamment complexe pour garantir l'informativité des données pour une identification sans excitation externe. Dans [BGM10], une condition nécessaire et une autre suffisante ont été développées pour n'importe quelle structure de modèle linéaire composée de matrices de fonctions de transfert. Pour les structures de modèle plusieurs-entrées une-sortie (*multiple-inputs single-output* ou *MISO* en anglais) de type ARMAX, le papier [YZ18] donne une condition nécessaire et suffisante mais sous des hypothèses fortes pour le correcteur et la structure de modèle ARMAX.

L'identification sans excitation externe est un cas très particulier. Dans la majorité des cas des identifications en boucle fermée, une excitation externe doit être ajoutée. Pour le cas boucle ouverte, il est clair qu'une excitation sur le vecteur d'entrée est nécessaire pour l'identification. La seule condition disponible dans la littérature qui nous permet de vérifier l'informativité des données dans le cas où il y a une excitation est basée sur la matrice densité spectrale de puissance du vecteur d'excitation : les données sont informatives si cette matrice est définie positive pour quasiment toutes les pulsations de l'intervalle $]-\pi, \pi]$ [BGM10]. Cependant, cette condition est suffisante et parfois très restrictive. Par exemple, cette condition ne sera jamais garantie pour une excitation de type sinusoïdale alors que ce type d'excitation pourrait être le seul permis par les limitations physiques du système. Ainsi, nous nous intéressons dans cette thèse à développer des conditions moins conservatrices, i.e., nous allons considérer ces deux premières questions de recherche dans cette dissertation

Question de Recherche 1

Pour un système MIMO LTI, une structure de modèle et une excitation donnés, comment pouvons nous vérifier que les données récoltées sur le système sont informatives pour une identification en *boucle ouverte* ?

Question de Recherche 2

Pour un système MIMO LTI, une structure de modèle, un correcteur MIMO LTI stabilisant le système en boucle fermée et une excitation externe donnés, comment pouvons nous vérifier que les données récoltées sur le système sont informatives pour une identification *directe en boucle fermée* ?

L'étude des deux questions de recherche est restreinte dans cette thèse aux structures de modèle qui sont des généralisations naturelles des structures classiques FIR, ARX, ARMAX, OE, BJ du cas SISO. Pour ces structures de modèle, nous prouvons que la vérification de l'informativité des données peut se faire sortie-par-sortie pour les deux configurations d'identification.

Dans le cas boucle ouverte (Question de Recherche 1), cette analyse d'informativité sortie-par-sortie se simplifie alors en l'étude de l'informativité pour l'identification en boucle ouverte des systèmes MISO. Inspiré du cas SISO [GBBM09], une condition suffisante¹ pour vérifier l'informativité des données dans le cas MISO est la persistance d'excitation d'un vecteur de signaux appelé régresseur. Ce régresseur contient des versions retardées des signaux du vecteur d'entrée. Ensuite, nous développons des conditions nécessaires et suffisantes pour vérifier la persistance d'excitation pour deux types d'excitation pour le vecteur d'entrée : multisinus et bruits blancs filtrés. Dans le dernier cas, le vecteur d'entrée \mathbf{u} du système MISO considéré est de la forme $\mathbf{u}(t) = \mathbf{M}(z)\mathbf{v}(t)$ où $\mathbf{M}(z)$ est une matrice de fonctions de transfert stables et $\mathbf{v}(t)$ est un vecteur de bruits blancs indépendants. Pour les deux types d'excitation considérés, la persistance du régresseur est vérifiée si et seulement si une certaine matrice est de rang plein ligne. Cette matrice dépend de la complexité de la structure de modèle et de la paramétrisation du vecteur d'entrée, i.e., des amplitudes, des déphasages et des pulsations dans le cas multisinus et des coefficients des filtres de $\mathbf{M}(z)$ dans le cas bruits blancs filtrés. Nous cherchons ensuite des conditions sur les deux types d'excitation pour garantir cette propriété de rang plein ligne. Ceci va nous conduire dans le cas multisinus à une condition nécessaire sur le nombre de pulsations. Dans le cas bruit blanc, cette condition nécessaire se traduit par une complexité minimale de la matrice de fonctions de transfert $\mathbf{M}(z)$.

Pour le cas boucle fermée (Question de Recherche 2), la vérification de l'informativité sortie-par-sortie se traduit aussi par la vérification du rang plein ligne d'une certaine matrice. Cette matrice dépend des coefficients du correcteur, de la paramétrisation de l'excitation externe et de la complexité de la structure de modèle. Nous verrons aussi qu'elle sépare clairement la contribution du correcteur et de l'excitation externe sur l'informativité des données.

G.4 Et pour les systèmes non-linéaires identifiables avec la méthode Prediction Error ?

Dans les études présentées dans le paragraphe précédent et dans la littérature, l'informativité des données n'a été étudiée que pour l'identification des systèmes LTI alors que la plupart des systèmes réels ont des dynamiques non-linéaires. Heureusement, la méthode d'identification par erreur de prédiction peut être utilisée pour identifier certaines classes de systèmes non-linéaires tels que les systèmes bloc-orientés avec des

¹Pour les structures de modèle ARX et FIR, cette condition est aussi nécessaire.

non-linéarités statiques (Hammerstein/Wiener [GB10, ST17]). De notre connaissance de la littérature, il n'y a pas de résultats sur l'informativité des données pour l'identification des systèmes Hammerstein/Wiener avec la méthode d'identification Prediction Error. Dans cette dissertation, nous allons nous intéresser à l'identification d'une classe particulière de systèmes Hammerstein donnée par

$$y(t) = G_{0,1}(z)u(t) + G_{0,2}(z)u^n(t) + H_0(z)e(t) \quad (\text{G.1})$$

où u est l'entrée, y la sortie, $n \in \mathbb{N}^*$, e un bruit blanc et z l'opérateur d'avance en temps discret. Les termes $G_{0,1}(z)$, $G_{0,2}(z)$ et $H_0(z)$ sont trois fonctions de transfert à temps discret à identifier. En choisissant une structure de modèle non-linéaire qui permet de modéliser toutes les dynamiques de ces 3 fonctions de transferts, il nous faut alors garantir l'informativité des données par rapport à cette structure de modèle. Ainsi, la troisième question de recherche que nous allons considérer est la suivante

Question de Recherche 3

Pour la classe de système Hammerstein de la forme

$$y(t) = G_{0,1}(z)u(t) + G_{0,2}(z)u^n(t) + H_0(z)e(t) \quad n \in \mathbb{N}^*$$

et pour une structure de modèle non-linéaire et une excitation données, comment pouvons nous vérifier que les données récoltées sur ce système sont informatives pour une identification en boucle ouverte avec la méthode Prediction Error ?

L'idée qui est suivie pour résoudre cette question de recherche est basée sur le reformulation du système (G.1) en un système MISO LTI avec deux entrées données par $u_1 = u$ et $u_2 = u^n$. Nous pouvons alors appliquer les résultats développés pour la Question de Recherche 1 à cette étude. Nous considérerons deux types d'excitation pour u : multisinus et bruit blanc gaussien. Bien que le système non-linéaire décrit par (G.1) est une structure très particulière, elle permet de modéliser les dynamiques principales d'un système réel qui est le gyroscope inertiel MEMS, présenté dans la prochaine section.

G.5 Étude pratique de la thèse : gyroscope MEMS

G.5.1 Présentation du gyroscope MEMS et des défis de contrôle

En plus de la partie théorique décrite précédemment, cette thèse se concentre sur un problème pratique : l'identification d'un gyroscope MEMS [Apo16, Kem11]. Il s'agit d'un capteur inertiel qui mesure la vitesse de rotation autour d'un axe donné et dont le principe est basé sur l'effet Coriolis. Un gyroscope (GYPRO®) est illustré en Figure 1.3.

Parmi tous les types de gyroscopes qui existent, le gyroscope MEMS a de nombreux avantages comme sa petite taille, sa faible consommation d'énergie et son bas prix.

Au niveau de l'architecture, un gyroscope MEMS est composé de deux masses, chacune vibrant dans une direction donnée de sorte que les deux directions de vibrations soient orthogonales. Les dynamiques de ces vibrations ont un comportement résonant (une pulsation de résonance pour chaque masse) et ces vibrations sont couplées par la force de Coriolis qui agit sur chaque masse et qui est proportionnelle à la vitesse de rotation. L'idée générale de mesure de la vitesse de rotation est basée sur une estimation de l'effet Coriolis. Pour ceci, nous avons besoin de faire vibrer les masses et de mesurer leur déplacement. Ainsi, une instrumentation de type électrostatique est souvent utilisée pour exciter et mesurer le déplacement des deux masses. L'actuation de chaque masse se fait via des tensions d'excitation (une par masse) et qui sont donc les entrées du gyroscope MEMS. Les sorties de gyroscope MEMS sont aussi deux tensions qui sont une image du déplacement des masses (une tension par masse). Le gyroscope MEMS est donc un système MIMO avec deux entrées et deux sorties.

Le gyroscope MEMS souffre d'imperfections de fabrications qui créent des dynamiques couplées entre les deux masses, c'est à dire le mouvement d'une masse génère un mouvement sur l'autre masse même sans effet Coriolis. Les deux phénomènes responsables de ces couplages sont l'anisoélasticité et l'effet non-proportionnel de l'amortissement [Sau08]. Ces dynamiques croisées peuvent être modélisées par des coefficients "croisés" de raideurs et d'amortissement. De plus, l'instrumentation électrostatique apporte elle aussi des dynamiques supplémentaires aux dynamiques mécaniques du capteur, parasitant la mesure de la vitesse de rotation si ces dynamiques additionnelles ne sont pas prises en compte lors de cette mesure. Par exemple, un effet de couplage capacitif entre les tensions d'excitation et de sortie est présent [AS05] et peut dégrader fortement la mesure de la vitesse de rotation si non pris en compte dans le procédé d'estimation de cette vitesse [LWZ⁺17]. Ce phénomène sera appelé couplage capacitif parasite dans cette dissertation. Par ailleurs, une non-linéarité carrée sur les tensions d'excitation est présente puisque les forces électrostatiques qui agissent sur les masses sont proportionnelles au carré des tensions d'excitation [Sau08].

Pour remédier à ces problèmes, le gyroscope est placé en boucle fermée avec un correcteur MIMO afin d'améliorer la précision de mesure de la vitesse de rotation. Les objectifs de contrôle sont choisis tels que l'estimation de la vitesse angulaire se fasse en 3 étapes. Premièrement, le déplacement d'une des deux masses (appelée masse drive dans ce résumé) doit suivre une consigne sinusoïdale. Par ailleurs, la pulsation de cette consigne sinusoïdale doit être celle de la pulsation de résonance de la masse drive afin de diminuer la puissance du signal de commande (tension d'excitation de la masse drive) qui assure l'asservissement sinusoïdal de la masse drive. Deuxièmement, le mouvement de l'autre masse (appelée masse sense dans ce résumé) généré par l'effet Coriolis doit être supprimé par le correcteur. Comme la force de Coriolis peut être vue comme une perturbation entrée sur la masse sense, le signal de contrôle de la masse sense (tension d'excitation de

la masse sense) est alors une image de la force de Coriolis. Ainsi, une estimation de cette force peut se faire via la commande de la masse sense. Enfin, en dernière étape, nous pouvons déduire la vitesse de rotation de l'estimation de la force de Coriolis puisque cette dernière est proportionnelle à la vitesse de rotation. Plusieurs travaux disponibles dans la littérature ont développé des correcteurs linéaires pour assurer le contrôle du gyroscope MEMS tels que des correcteurs PI, PID, etc [Sau08, EMK12, PTKH09] ou encore des correcteurs non-linéaires par modes glissants [FB09b, FB09a, Fei10, ZDLG08].

G.5.2 Objectives et présentation du projet Next4MEMS

Les méthodes de contrôle citées précédemment permettent d'obtenir des estimées relativement précises de la vitesse de rotation. Cependant, nous pouvons nous demander si nous pouvons encore améliorer la précision afin d'avoir des capteurs gyroscopiques qui sont plus compétitifs face aux capteurs optiques. C'est le but du projet Next4MEMS financé par Bpifrance² dans lequel s'inscrit cette thèse. Le projet Next4MEMS est un consortium de trois entreprises privées développant des capteurs inertiels (Tronics/TDK, iXblue, Asygn) et de deux laboratoires publics (ONERA³ et Ampère). Le but du laboratoire Ampère dans le projet est le développement d'une méthodologie pour la synthèse d'un correcteur optimal en vue d'une implémentation sur le gyroscope MEMS. Quatre doctorants du laboratoire Ampère participent au projet :

- Fabrício SAGGIN (encadré par Xavier BOMBOIS, Gérard SCORLETTI et Anton KORNIENKO). Il travaille sur la synthèse de correcteur et la méthode qu'il considère dans sa thèse est la synthèse \mathcal{H}_∞ . Cette méthode fonctionne bien pour les systèmes MIMO avec des dynamiques croisées comme pour le gyroscope MEMS.
- Kévin COLIN (encadré par Xavier BOMBOIS, Laurent BAKO et Anton KORNIENKO). La synthèse du correcteur doit se faire par rapport à un modèle du gyroscope. Comme mentionné en Section G.1, pour obtenir des performances optimales, il est important d'avoir un modèle précis des dynamiques du gyroscope. C'est ce problème qui va nous intéresser en seconde partie de cette thèse en développant une méthodologie d'identification.
- Jorge AYALA (encadré par Gérard SCORLETTI et Anton KORNIENKO). Il travaille sur la partie analyse de robustesse qui lie la modélisation et la synthèse de correcteurs. Le correcteur est basé sur un modèle des dynamiques du système. Cependant, à cause des incertitudes sur les paramètres estimés et les phénomènes non-modélisés, les performances du correcteur peuvent se dégrader lorsqu'il est implémenté sur le vrai système. Ainsi, son objectif dans sa thèse est de développer des méthodes pour évaluer cette dégradation de performances.
- Federico MORELLI (encadré par Xavier BOMBOIS, Laurent BAKO et Anton KORNIENKO). Dans les objectifs de contrôle, la pulsation de la consigne sinusoïdale

²Banque publique française d'investissement.

³Office National d'Études et de Recherches Aérospatiales.

pour la masse drive doit être égale à sa pulsation de résonance. Cependant, les propriétés mécaniques des masses du gyroscope changent en fonction de la température modifiant les pulsations de résonance. Une des objectifs de Federico MORELLI dans sa thèse est de développer des méthodes pour asservir la pulsation de la consigne à la pulsation de résonance de la masse drive. Deux méthodes sont considérées dans ses travaux : l'extremum seeking [ZO11] et l'identification récursive [LS83].

Comme cette thèse se focalise sur l'identification du gyroscope MEMS, nous donnons un état de l'art sur ce sujet dans le prochain paragraphe.

G.5.3 État de l'art sur l'identification des gyroscopes MEMS

Les premiers travaux sur l'identification des gyroscopes MEMS peuvent être trouvés dans [MGH01, CHM03] où des méthodes d'identification en boucle ouverte et boucle fermée avec des structures de modèle de type MIMO ARMAX ont été appliquées avec des excitations de type sweep. Cependant l'architecture de gyroscope étudié dans [MGH01, CHM03] ne correspond pas à l'architecture du gyroscope considéré dans cette thèse puisque le gyroscope de [MGH01, CHM03] considère des vibrations angulaires des masses alors que celui du projet Next4MEMS repose sur des vibrations linéaires des masses.

Pour le type de gyroscope considéré dans le projet, la plupart des papiers étudie la modélisation des dynamiques mécaniques du gyroscope. Dans [PS02], l'attention est portée sur les termes croisés de raideur et trois méthodes sont développées pour les estimer. Cependant les termes croisés d'amortissement ne sont pas considérés dans cette étude ce qui limite son utilisation pour une identification d'un gyroscope réel. Dans [PS04], tous ces termes sont modélisés avec une méthode fréquentielle alors que les méthodes développés dans [FY11, FY12] sont basées sur une estimation en ligne avec un observateur en boucle ouverte et en boucle fermée.

Bien que ces travaux donnent plusieurs méthodes de modélisation, elles ne modélisent que les dynamiques mécaniques et ne considèrent pas les effets supplémentaires causés par l'instrumentation du gyroscope. Par exemple, le couplage capacitif parasite mentionné en Section G.5.1 n'est pas considéré. À notre connaissance, le seul papier qui identifie ce phénomène est [AS05]. Il s'agit d'une identification dans le domaine fréquentielle qui se base sur une linéarisation de la non-linéarité de type carré (mentionnée en Section G.5.1) pour des excitations sinusoïdales. Bien que cette approche semble donner un modèle précis des dynamiques du gyroscope, ce dernier n'est valable que sur une courte plage de pulsations alors que nous cherchons justement à obtenir un modèle valable sur une plus large plage de pulsations pour la synthèse du correcteur.

Enfin, dans tous les travaux cités sur l'identification des gyroscopes MEMS, aucun ne considère une estimation des densités spectrales de puissance des bruits de mesure sur

chaque tension de sortie alors que ces dernières sont importantes pour la synthèse du correcteur \mathcal{H}_∞ . Ainsi, les structures de modèle et les méthodes d'identification présentes dans la littérature ne semblent pas être adaptées à nos besoins de modélisation. Ainsi, dans cette thèse, nous étudions cette dernière question de recherche

Question de Recherche 4

Pour la synthèse optimale d'un correcteur pour le MEMS gyroscope, comment pouvons nous développer une structure de modèle complète qui permet de modéliser les dynamiques principales du gyroscope MEMS, à savoir les dynamiques mécaniques, le couplage capacitif parasite et les bruits de mesure? À partir de cette structure de modèle, comment pouvons nous développer une méthode d'identification pour une modélisation précise de ces trois types de dynamiques sans linéarisation ?

Pour développer la structure de modèle complète, nous utilisons les principes de la physique. Cette structure de modèle est de type bloc-orientée et peut être identifiée efficacement avec la méthode Prediction Error. Nous considérons alors cette méthode pour développer la méthodologie d'identification pour répondre à la deuxième partie de la Question de Recherche 4.

Comme nous l'avons déjà expliqué en Section G.5.1, le gyroscope MEMS est un système MIMO avec deux entrées et deux sorties. Par conséquent, il serait logique de faire une expérience d'identification pour laquelle les deux entrées sont excitées et les deux sorties sont mesurées simultanément. Cependant, en réalité, certains dynamiques peuvent être cachées par d'autres, causant des imprécisions lors de l'estimation du modèle. Par exemple, les dynamiques mécaniques croisées peuvent être cachées par les dynamiques mécaniques directes. Pour résoudre ce problème, il suffit de voir qu'un système MIMO avec deux entrées et deux sorties peut être vu comme un ensemble de quatre sous-systèmes avec une entrée et une sortie (les quatre blocs "Subsystem" illustrés dans la Figure 1.4). L'idée est d'exciter une entrée (par exemple l'entrée u_1 de la Figure 1.4) et ensuite d'identifier les deux sous-systèmes correspondants (Subsystems 1 et 3 de la Figure 1.4). Ensuite, nous excitons l'autre entrée (u_2 de la Figure 1.4) et ensuite nous identifions les deux sous-systèmes correspondants (Subsystems 2 et 4 de la Figure 1.4).

De plus, les dynamiques des sous-systèmes sont de la forme (G.1) avec $n = 2$. Ainsi, pour l'identification, nous pouvons utiliser la même astuce considérée pour répondre à la Question de Recherche 3 en transformant chaque sous-système en un système MISO LTI avec deux entrées. Afin de garantir une estimée précise des dynamiques de chaque MISO,

il nous faut choisir des tensions d'excitations telles que les données sont informatives par rapport à la structure de modèle choisie. Les résultats obtenus pour répondre à la Question de Recherche 3 seront alors utiles dans ce cas pratique. Comme mentionné en Section G.2, avec l'informativité des données, nous pouvons rendre les incertitudes sur les paramètres identifiés aussi petites que nous voulons, à *condition* d'augmenter le nombre de données. Cependant, pour certains choix d'excitation, ceci peut demander un nombre trop grand de données ce qui n'est pas réalisable en pratique. Pour obtenir des faibles niveaux d'incertitudes sur les paramètres identifiés pour un nombre de données raisonnable, il nous faut aussi augmenter le rapport signal sur bruit (*signal-to-noise ratio* ou *SNR* en anglais) des sorties mesurées.

Même si le rapport signal sur bruit est augmenté et les données sont informatives pour l'identification des quatre sous-systèmes, les quatre identifications MISO LTI sont des problèmes d'optimisation non-linéaires complexes et donc une estimée initiale est requise pour initialiser les algorithmes d'optimisation. Pour déterminer ces estimées initiales, nous proposons une approche qui repose sur une série d'identifications SISO de chaque fonction de transfert impliqué dans les quatre structures de modèle MISO considérées pour la modélisation du gyroscope MEMS.

G.6 Organisation du long résumé des contributions

Ce long résumé de la thèse va être organisé de la façon suivante

- en Annexe H, nous donnons les bases de l'identification par Prediction Error pour résumer les résultats théoriques de la thèse (résumé du Chapitre 2).
- en Annexe I, nous résumons les résultats donnés dans les Chapitres 3, 4 et 5 qui répondent respectivement aux Questions de Recherche 1, 2 et 3.
- en Annexe J, nous présentons le MEMS et nous développons une structure de modèle complète pour la modélisation des trois dynamiques principales du gyroscope MEMS (résumé du Chapitre 6).
- en Annexe K, nous résumons la méthode d'identification des dynamiques principales du gyroscope MEMS (résumé du Chapitre 7).
- en Annexe L, nous donnons la conclusion de la thèse et des ouvertures pour des recherches futures (résumé du Chapitre 8).

Long résumé : L'identification par erreur de prédiction

Contents

H.1	Les signaux quasi-stationnaires à temps discret	233
H.2	La Méthode Prediction Error pour l'identification des systèmes MIMO en boucle ouverte	234
H.3	Identification en boucle fermée et consistance	237
H.4	Structures de modèles considérées dans cette thèse	238
H.5	Résultats sur l'informativité des données dans le cas MIMO présents dans la littérature	240
H.6	Persistance d'excitation de vecteurs de signaux quasi-stationnaires .	242
H.7	Conclusion de l'Annexe H	243

Dans cette annexe, nous donnons les notions essentielles pour la présentation des contributions théoriques de la thèse liées aux Questions de Recherche 1, 2 et 3.

H.1 Les signaux quasi-stationnaires à temps discret

Dans la partie théorique de la thèse, nous considérerons des vecteurs de signaux à temps discret. Nous allons supposer que les signaux sont échantillonnés avec une période d'échantillonnage constante T_s . Dans ce cas, un vecteur de signaux \mathbf{x} peut être vu comme un ensemble de valeurs $\{\mathbf{x}(tT_s)\}_{t \in \mathbb{N}^*}$. Cependant, pour simplifier les notations, nous enlèverons l'indication de la période d'échantillonnage en normalisant le temps, c'est à dire nous verrons les vecteurs de signaux à temps discret \mathbf{x} comme un ensemble de valeurs $\{\mathbf{x}(t)\}_{t \in \mathbb{N}^*}$.

Pour la méthode par erreur de prédiction, nous avons besoin de définir la classe des vecteurs de signaux quasi-stationnaires à temps discret. Pour cela, nous définissons

l'opérateur \bar{E} pour les vecteurs de signaux à temps discret tels que la limite ci-dessous soit définie

$$\bar{E}[\mathbf{x}(t)] = \lim_{N \rightarrow +\infty} \frac{1}{N} \sum_{t=1}^N E[\mathbf{x}(t)] \quad (\text{H.1})$$

où E est l'opérateur d'espérance (moyenne). Ainsi, avec l'opérateur \bar{E} , nous définissons la classe des vecteurs de signaux quasi-stationnaires

Définition H.1. *Un vecteur de signaux à temps discret \mathbf{x} composé de n signaux x_1, \dots, x_n à valeur réelle ou complexe est quasi-stationnaire si les deux conditions suivantes sont vérifiées*

- (a) *le terme $\bar{E}[\mathbf{x}(t)]$ est bien défini et est égal à une constante.*
- (b) *la matrice suivante est bien définie et est une fonction qui ne dépend que de $\tau \in \mathbb{Z}$*

$$\bar{E}[\mathbf{x}(t)\mathbf{x}^T(t-\tau)] = \begin{pmatrix} \bar{E}[x_1(t)x_1(t-\tau)] & \bar{E}[x_1(t)x_2(t-\tau)] & \cdots & \bar{E}[x_1(t)x_n(t-\tau)] \\ \bar{E}[x_2(t)x_1(t-\tau)] & \bar{E}[x_2(t)x_2(t-\tau)] & \cdots & \bar{E}[x_2(t)x_n(t-\tau)] \\ \vdots & \vdots & \ddots & \vdots \\ \bar{E}[x_n(t)x_1(t-\tau)] & \bar{E}[x_n(t)x_2(t-\tau)] & \cdots & \bar{E}[x_n(t)x_n(t-\tau)] \end{pmatrix}$$

L'ensemble des signaux quasi-stationnaires contient les signaux multisinus déterministes et les signaux stochastiques de type bruits blancs filtrés. Dans cette dissertation, nous considérerons ces deux types de signaux quasi-stationnaires.

H.2 La Méthode Prediction Error pour l'identification des systèmes MIMO en boucle ouverte

Soit \mathcal{S} un système MIMO avec un vecteur d'entrée $\mathbf{u} \in \mathbb{R}^{n_u}$ et un vecteur de sortie $\mathbf{y} \in \mathbb{R}^{n_y}$ défini par

$$\mathcal{S} : \quad \mathbf{y}(t) = \mathbf{G}_0(z)\mathbf{u}(t) + \mathbf{H}_0(z)\mathbf{e}(t) \quad (\text{H.2})$$

où $\mathbf{G}_0(z)$ est une matrice stable de fonctions de transfert de dimension $n_y \times n_u$, $\mathbf{H}_0(z)$ une matrice monique¹, stable et inversement stable de fonctions de transfert de dimension $n_y \times n_y$, $\mathbf{e} \in \mathbb{R}^{n_y}$ est un vecteur de bruits blancs tel que $\bar{E}[\mathbf{e}(t)\mathbf{e}^T(t)] = \mathbf{\Sigma}_0 \succ 0$. Nous allons aussi faire l'hypothèse suivante :

Hypothèse H.1. *La matrice de fonctions de transfert $\mathbf{H}_0(z)$ est supposée être diagonale, i.e.,*

$$\mathbf{H}_0(z) = \text{diag}(H_{0,1}(z), \dots, H_{0,n_y}(z))$$

où *diag* est la notation introduite dans la section Notations (voir page xx).

¹C'est à dire $\mathbf{H}_0(z = \infty) = 1$.

H.2. La Méthode Prediction Error pour l'identification des systèmes MIMO en boucle ouverte

Grâce à l'Hypothèse H.1, le système \mathcal{S} donné par (H.2) peut être vu comme un ensemble de n_y systèmes MISO donnés par

$$\mathcal{S}_i : \quad y_i(t) = \mathbf{G}_{0,i}(z)\mathbf{u}(t) + H_{0,i}(z)e_i(t) \quad (i = 1, \dots, n_y) \quad (\text{H.3})$$

où $\mathbf{G}_{0,i}(z)$ est la i -ième ligne de $\mathbf{G}_0(z)$. Les sorties peuvent cependant être corrélées via la matrice de covariance Σ_0 du vecteur de bruits blancs \mathbf{e} .

Quand le système \mathcal{S} donné par (H.2) est en boucle ouverte, le vecteur de signaux qui va être excité pour l'expérience d'identification est le vecteur d'entrée \mathbf{u} et il est supposé être indépendant de \mathbf{e} . Nous allons considérer deux types d'excitation pour \mathbf{u} .

Dans le premier type, les entrées u_k ($k = 1, \dots, n_u$) du vecteur d'entrée \mathbf{u} sont des multisinus qui sont la somme de s sinusoïdes de pulsations différentes $\omega_l \in]0, \pi[$ ($l = 1, \dots, s$), i.e.,

$$u_k(t) = \sum_{l=1}^s \lambda_{kl} \cos(\omega_l t + \psi_{kl}) \quad (k = 1, \dots, n_u). \quad (\text{H.4})$$

où λ_{kl} et ψ_{kl} sont respectivement l'amplitude et le déphasage de la sinusoïde à la pulsation ω_l de l'entrée u_k . Certaines amplitudes λ_{kl} peuvent être égales à 0 pour quelques valeurs de $k = 1, \dots, n_u$ mais, pour chaque $l = 1, \dots, s$, il existe une valeur de k telle que $\lambda_{kl} \neq 0$. Nous noterons par s_k le nombre de sinusoïdes à amplitude non-nulle du signal u_k ($s_k \leq s \forall k$).

Dans le second type d'excitation, \mathbf{u} est donné par $\mathbf{u}(t) = \mathbf{M}(z)\mathbf{v}(t)$ avec $\mathbf{M}(z)$ une matrice stable de fonctions de transfert $\mathbf{M}(z) = (M_{kq}(z))_{(k,q) \in \llbracket 1, n_u \rrbracket \times \llbracket 1, f \rrbracket}$ et \mathbf{v} un vecteur de f bruits blancs indépendants v_q ($q = 1, \dots, f$), i.e., $\mathbf{v} = (v_1, \dots, v_f)^T$. Ainsi, chaque signal u_k de \mathbf{u} est donné par

$$u_k(t) = \sum_{q=1}^f M_{kq}(z)v_q(t) \quad (k = 1, \dots, n_u). \quad (\text{H.5})$$

Le nombre f de signaux qui composent \mathbf{v} n'est pas nécessairement égal à n_u et certaines fonctions de transfert $M_{kq}(z)$ peuvent être nulles, mais, pour chaque $q = 1, \dots, f$, il existe une valeur de $k = 1, \dots, n_u$ telle que $M_{kq}(z) \neq 0$.

Pour identifier le système \mathcal{S} , nous considérons une structure de modèle qui modélise toutes les dynamiques de \mathcal{S} et paramétrée comme suit

$$\mathcal{M} = \{(\mathbf{G}(z, \boldsymbol{\theta}), \mathbf{H}(z, \boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \mathcal{D}_\theta\}$$

où $\boldsymbol{\theta} \in \mathcal{D}_\theta \subset \mathbb{R}^n$ est le vecteur de paramètre. Comme cette structure de modèle peut modéliser toutes les dynamiques de \mathcal{S} (nous parlerons alors de structure de modèle *d'ordre complet*), il existe alors un vrai vecteur de paramètres $\boldsymbol{\theta}_0 \in \mathcal{D}_\theta$ tel que

$(\mathbf{G}(z, \boldsymbol{\theta}_0), \mathbf{H}(z, \boldsymbol{\theta}_0)) = (\mathbf{G}_0(z), \mathbf{H}_0(z))$. L'ensemble \mathcal{D}_θ contient les vecteurs de paramètres tels que $\mathbf{G}(z, \boldsymbol{\theta})$ soit stable et $\mathbf{H}(z, \boldsymbol{\theta})$ soit monique, stable et inversement stable. De plus la matrice $\mathbf{H}(z, \boldsymbol{\theta})$ est diagonale, conséquence de l'Hypothèse H.1. Nous allons aussi supposer que cette structure de modèle est globalement identifiable en $\boldsymbol{\theta}_0$, i.e., pour tout $\boldsymbol{\theta} \in \mathcal{D}_\theta$,

$$(\mathbf{G}(z, \boldsymbol{\theta}), \mathbf{H}(z, \boldsymbol{\theta})) = (\mathbf{G}(z, \boldsymbol{\theta}_0), \mathbf{H}(z, \boldsymbol{\theta}_0)) \implies \boldsymbol{\theta} = \boldsymbol{\theta}_0$$

Pour chaque modèle $(\mathbf{G}(z, \boldsymbol{\theta}), \mathbf{H}(z, \boldsymbol{\theta})) \in \mathcal{M}$, nous définissons le prédicteur $\hat{\mathbf{y}}(t, \boldsymbol{\theta})$ à l'horizon 1 pour la sortie \mathbf{y} à l'instant t par

$$\hat{\mathbf{y}}(t, \boldsymbol{\theta}) = \mathbf{W}_y(z, \boldsymbol{\theta})\mathbf{y}(t) + \mathbf{W}_u(z, \boldsymbol{\theta})\mathbf{u}(t) = \mathbf{W}(z, \boldsymbol{\theta})\mathbf{x}(t)$$

où

$$\mathbf{W}_y(z, \boldsymbol{\theta}) = \mathbf{I}_{n_y} - \mathbf{H}^{-1}(z, \boldsymbol{\theta}) \quad (\text{H.6})$$

$$\mathbf{W}_u(z, \boldsymbol{\theta}) = \mathbf{H}^{-1}(z, \boldsymbol{\theta})\mathbf{G}(z, \boldsymbol{\theta}) \quad (\text{H.7})$$

$$\mathbf{W}(z, \boldsymbol{\theta}) = (\mathbf{W}_y(z, \boldsymbol{\theta}), \mathbf{W}_u(z, \boldsymbol{\theta}))$$

$$\mathbf{x}(t) = \begin{pmatrix} \mathbf{y}(t) \\ \mathbf{u}(t) \end{pmatrix}$$

À partir d'un jeu de données \mathcal{Z}^N récoltées sur le système \mathcal{S} et défini par

$$\mathcal{Z}^N = \left\{ \mathbf{x}(t) = (\mathbf{y}^T(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, N \right\}$$

nous définissons le vecteur optimal de paramètre $\hat{\boldsymbol{\theta}}_N$ qui minimise le critère $V_N(\boldsymbol{\theta})$ des moindres carrés sur l'erreur de prédiction $\boldsymbol{\epsilon}(t, \boldsymbol{\theta}) = \mathbf{y}(t) - \hat{\mathbf{y}}(t, \boldsymbol{\theta})$:

$$\hat{\boldsymbol{\theta}}_N = \arg \min_{\boldsymbol{\theta} \in \mathcal{D}_\theta} V_N(\boldsymbol{\theta}, \mathcal{Z}^N) \quad (\text{H.8})$$

$$V_N(\boldsymbol{\theta}, \mathcal{Z}^N) = \frac{1}{N} \sum_{t=1}^N \boldsymbol{\epsilon}^T(t, \boldsymbol{\theta}) \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\epsilon}(t, \boldsymbol{\theta})$$

où la matrice de covariance $\boldsymbol{\Sigma}_0$ est supposée être connue mais peut être estimée simultanément avec $\hat{\boldsymbol{\theta}}_N$ (voir, e.g., [Lju99, Chapitre 15]). Nous voulons que $\hat{\boldsymbol{\theta}}_N$ soit une estimée consistante de $\boldsymbol{\theta}_0$, i.e., $\hat{\boldsymbol{\theta}}_N$ doit converger vers $\boldsymbol{\theta}_0$ avec une probabilité de 1 lorsque $N \rightarrow +\infty$. Cette propriété est cruciale puisqu'elle permet de rendre l'erreur de modélisation $\hat{\boldsymbol{\theta}}_N - \boldsymbol{\theta}_0$ aussi petite que voulue en choisissant un nombre de données N suffisamment grand.

Dans [Lju99, Chapitre 8], il a été prouvé que la consistance est obtenue lorsque le vrai vecteur de paramètres $\boldsymbol{\theta}_0$ est l'unique minimiseur du critère asymptotique donné par $\bar{E} [\boldsymbol{\epsilon}^T(t, \boldsymbol{\theta}) \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\epsilon}(t, \boldsymbol{\theta})]$, ou de façon équivalente, lorsque $\boldsymbol{\theta}_0$ est l'unique vecteur de paramètres $\boldsymbol{\theta}$ pour lequel $\boldsymbol{\epsilon}(t, \boldsymbol{\theta}) = \mathbf{e}(t)$. Pour garantir cette unicité du minimiseur, le vecteur d'entrée \mathbf{u} choisi pour la génération des données doit garantir l'informativité de ces données, définie ci-dessous :

Définition H.2 (Informativité des données [BGM10, Lju99]). *Considérons le cadre d'identification défini ci-dessus. Supposons qu'un vecteur d'entrée quasi-stationnaire \mathbf{u} est appliqué au système \mathcal{S} donné par (H.2) de l'instant $t = 1$ à l'instant $t = +\infty$ générant le jeu de données \mathcal{Z}^∞ suivant*

$$\mathcal{Z}^\infty = \left\{ \mathbf{x}(t) = (\mathbf{y}^T(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, +\infty \right\} \quad (\text{H.9})$$

Soit une structure de modèle \mathcal{M} générant le prédicteur $\hat{\mathbf{y}}(t, \boldsymbol{\theta}) = \mathbf{W}(z, \boldsymbol{\theta})\mathbf{x}(t)$. Définissons aussi l'ensemble suivant

$$\Delta_{\mathbf{W}} = \left\{ \Delta \mathbf{W}(z) = \mathbf{W}(z, \boldsymbol{\theta}') - \mathbf{W}(z, \boldsymbol{\theta}'') \mid \boldsymbol{\theta}' \text{ et } \boldsymbol{\theta}'' \text{ dans } \mathcal{D}_\theta \right\} \quad (\text{H.10})$$

Alors, le jeu de données \mathcal{Z}^∞ est informatif par rapport à la structure de modèle \mathcal{M} lorsque, pour tout $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}}$, nous avons

$$\bar{E} [\|\Delta \mathbf{W}(z)\mathbf{x}(t)\|^2] = 0 \implies \Delta \mathbf{W}(z) \equiv \mathbf{0}_{n_y \times (n_y + n_u)} \quad (\text{H.11})$$

où $\Delta \mathbf{W}(z) \equiv \mathbf{0}_{n_y \times (n_y + n_u)}$ signifie que $\Delta \mathbf{W}(e^{j\omega}) = \mathbf{0}_{n_y \times (n_y + n_u)}$ pour quasiment toutes les pulsations ω et où \bar{E} est l'opérateur défini par (H.1).

D'après la Définition H.2, si le vecteur d'entrée \mathbf{u} génère des données informatives, alors nous avons que $\boldsymbol{\epsilon}(t, \boldsymbol{\theta}) = \mathbf{e}(t)$ si et seulement si la propriété suivante est respectée : $(\mathbf{G}(z, \boldsymbol{\theta}), \mathbf{H}(z, \boldsymbol{\theta})) = (\mathbf{G}(z, \boldsymbol{\theta}_0), \mathbf{H}(z, \boldsymbol{\theta}_0))$. Comme la structure de modèle est supposée être globalement identifiable en $\boldsymbol{\theta}_0$, alors la dernière propriété implique que $\boldsymbol{\theta}_0$ est l'unique vecteur de paramètres $\boldsymbol{\theta}$ pour lequel $\boldsymbol{\epsilon}(t, \boldsymbol{\theta}) = \mathbf{e}(t)$, et donc nous obtenons la consistance de $\hat{\boldsymbol{\theta}}_N$.

H.3 Identification en boucle fermée et consistance

L'identification peut aussi se faire en boucle fermée. Dans ce cas là, le système \mathcal{S} donné par (H.2) est opéré en boucle fermée avec un correcteur MIMO LTI stabilisant la boucle (voir Figure 2.1). Un vecteur d'excitation externe $\mathbf{r} \in \mathbb{R}^{n_u}$ peut être rajouté au vecteur de commande pour des besoins d'identification. Dans ce cas là, l'entrée \mathbf{u} appliquée au système \mathcal{S} est donnée par

$$\mathbf{u}(t) = -\mathbf{K}(z)\mathbf{y}(t) + \mathbf{r}(t) \quad (\text{H.12})$$

Les vecteurs de signaux \mathbf{r} et \mathbf{e} sont supposés être indépendants et nous allons considérer les mêmes types d'excitation qu'en boucle ouverte, i.e., multisinus et bruits blancs filtrés. Pour le besoin de référer à ce type d'excitation en boucle fermée et éviter une confusion avec les excitations en boucle ouverte, nous répétons les équations de ces deux types d'excitations pour \mathbf{r}

$$r_k(t) = \sum_{l=1}^s \lambda_{kl} \cos(\omega_l t + \psi_{kl}) \quad (k = 1, \dots, n_u). \quad (\text{H.13})$$

$$r_k(t) = \sum_{q=1}^f M_{kq}(z)v_q(t) \quad (k = 1, \dots, n_u). \quad (\text{H.14})$$

En appliquant un vecteur d'excitation externe \mathbf{r} sur le système en boucle fermée de l'instant $t = 1$ à l'instant $t = N$, nous collectons le jeu de données entrée/sortie suivant

$$\mathcal{Z}_{cl}^N = \left\{ \mathbf{x}(t) = (\mathbf{y}^T(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, N \right\}$$

et nous identifions une estimée $\hat{\boldsymbol{\theta}}_N$ de $\boldsymbol{\theta}_0$ en considérant le même critère d'identification (H.8) utilisé en boucle ouverte et la même structure de modèle \mathcal{M} . Si nous ajoutons l'hypothèse que les fonctions de transfert de $\mathbf{G}_0(z)\mathbf{K}(z)$ contiennent un retard (i.e., elles sont proportionnelles à z^{-1}), la consistance de l'estimée de $\hat{\boldsymbol{\theta}}_N$ est aussi garantie lorsque \mathcal{M} est globalement identifiable en $\boldsymbol{\theta}_0$ et que le vecteur d'excitation externe \mathbf{r} génère l'informativité des données. La définition de l'informativité des données en boucle fermée est presque la même que celle en boucle ouverte. La seule différence se trouve dans la façon de collecter le jeu de données.

Définition H.3 (Informativité des données en boucle fermée [BGM10, Lju99]). *Considérons le cadre d'identification défini en Section H.2. Supposons qu'un vecteur d'excitation externe \mathbf{r} excite le système \mathcal{S} en boucle fermée avec le correcteur $\mathbf{K}(z)$ comme illustré en Figure 2.1 de $t = 1$ à $t = +\infty$ et supposons que nous collectons le jeu de données suivant*

$$\mathcal{Z}_{cl}^\infty = \left\{ \mathbf{x}(t) = (\mathbf{y}^T(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, +\infty \right\}$$

Soit une structure de modèle \mathcal{M} générant le prédicteur $\hat{\mathbf{y}}(t, \boldsymbol{\theta}) = \mathbf{W}(z, \boldsymbol{\theta})\mathbf{x}(t)$. Le jeu de données \mathcal{Z}^∞ est informatif par rapport à la structure de modèle \mathcal{M} si (H.11) est assurée pour tout $\Delta\mathbf{W}(z) \in \boldsymbol{\Delta}_{\mathbf{W}}$ (où $\boldsymbol{\Delta}_{\mathbf{W}}$ et $\Delta\mathbf{W}(z)$ sont ceux définis en Définition 2.2).

H.4 Structures de modèles considérées dans cette thèse

Les structures de modèle \mathcal{M} qui vont être considérées sont des extensions naturelles des structures de modèle classiques utilisées pour l'identification des systèmes SISO, i.e., les structures de modèle ARX, FIR, ARMAX, BJ et OE. Ces extensions sont définies dans cette section. Pour cela, décomposons le vecteur de paramètres $\boldsymbol{\theta}$ comme suit

$$\boldsymbol{\theta} = \begin{pmatrix} \tilde{\boldsymbol{\theta}} \\ \boldsymbol{\eta} \end{pmatrix}$$

où $\tilde{\boldsymbol{\theta}}$ est le vecteur contenant les paramètres uniquement présents dans $\mathbf{G}(z, \boldsymbol{\theta})$ et $\boldsymbol{\eta}$ le vecteur contenant le reste des paramètres.

ARMAX : Les structures de modèle ARMAX sont définies par

$$\begin{cases} \mathbf{G}(z, \boldsymbol{\theta}) &= \mathbf{A}^{-1}(z, \boldsymbol{\eta})\mathbf{B}(z, \tilde{\boldsymbol{\theta}}) \\ \mathbf{H}(z, \boldsymbol{\theta}) &= \mathbf{A}^{-1}(z, \boldsymbol{\eta})\mathbf{C}(z, \boldsymbol{\eta}) \end{cases}$$

où $\mathbf{A}(z, \boldsymbol{\eta})$ et $\mathbf{C}(z, \boldsymbol{\eta})$ sont des matrices diagonales, polynomiales et moniques de dimension $n_y \times n_y$ et $\mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ est une matrice polynomiale de dimension $n_y \times n_u$. Les

structures de modèle ARX et FIR sont des cas particuliers de la structure ARMAX avec $\mathbf{C}(z, \boldsymbol{\eta}) = \mathbf{I}_{n_y}$ dans le cas ARX et avec $\mathbf{C}(z, \boldsymbol{\eta}) = \mathbf{A}(z, \boldsymbol{\eta}) = \mathbf{I}_{n_y}$ dans le cas FIR.

Le matrice $\mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ est composée de $n_y n_u$ polynômes indépendamment paramétrés $B_{ik}(z, \tilde{\boldsymbol{\theta}}) = B_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) = z^{-\rho_{ik}} \dot{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ ($i = 1, \dots, n_y, k = 1, \dots, n_u$) où les retards ρ_{ik} peuvent être différents et chaque $\dot{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ est un polynôme avec un retard égal à 0 paramétré par un vecteur de paramètres $\tilde{\boldsymbol{\theta}}_{ik}$. La matrice diagonale $H(z, \boldsymbol{\eta})$ est composée de n_y fonctions de transfert indépendamment paramétrées $H_i(z, \boldsymbol{\eta}_i)$ dans sa diagonale avec $H_i(z, \boldsymbol{\eta}_i) = C_i(z, \boldsymbol{\eta}_i)/A_i(z, \boldsymbol{\eta}_i)$ où $C_i(z, \boldsymbol{\eta}_i)$ et $A_i(z, \boldsymbol{\eta}_i)$ sont des polynômes moniques paramétrés par un vecteur de paramètres $\boldsymbol{\eta}_i$. Les coefficients polynomiaux de $\dot{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ sont les paramètres de $\tilde{\boldsymbol{\theta}}_{ik}$ alors que les coefficients polynomiaux de $A_i(z, \boldsymbol{\eta}_i)$ et $C_i(z, \boldsymbol{\eta}_i)$ sont ceux de $\boldsymbol{\eta}_i$. Par conséquent, nous avons que

$$\dot{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) = \tilde{\theta}_{ik,1} + \sum_{m=1}^{\deg(\dot{B}_{ik})} \tilde{\theta}_{ik,(m+1)} z^{-m}$$

$$C_i(z, \boldsymbol{\eta}_i) = 1 + \sum_{m=1}^{\deg(C_i)} \eta_{i,m} z^{-m}$$

$$A_i(z, \boldsymbol{\eta}_i) = 1 + \sum_{m=1}^{\deg(A_i)} \eta_{i,(m+\deg(C_i))} z^{-m}$$

où $\tilde{\theta}_{ik,m}$ est le m -ième élément de $\tilde{\boldsymbol{\theta}}_{ik}$ et $\eta_{i,m}$ le m -ième élément de $\boldsymbol{\eta}_i$. Le nombre de paramètres à identifier dans \dot{B}_{ik} , C_i et A_i est respectivement égal à $\deg(\dot{B}_{ik}) + 1$, $\deg(C_i)$ et $\deg(A_i)$.

BJ : Pour les structures de modèle BJ, les matrices $\mathbf{G}(z, \boldsymbol{\theta})$ et $\mathbf{H}(z, \boldsymbol{\theta})$ sont indépendamment paramétrées. Elles sont alors données par

$$\begin{cases} \mathbf{G}(z, \boldsymbol{\theta}) &= \mathbf{G}(z, \tilde{\boldsymbol{\theta}}) \\ \mathbf{H}(z, \boldsymbol{\theta}) &= \mathbf{H}(z, \boldsymbol{\eta}) \end{cases}$$

La structure de modèle OE est un cas particulier de la structure de modèle BJ avec $\mathbf{H}(z, \boldsymbol{\eta}) = \mathbf{I}_{n_y}$.

La matrice $\mathbf{G}(z, \tilde{\boldsymbol{\theta}})$ est composée de $n_y n_u$ fonctions de transfert indépendamment paramétrées $G_{ik}(z, \tilde{\boldsymbol{\theta}}) = G_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) = z^{-\rho_{ik}} \dot{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})/F_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ où $\dot{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ et $F_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ sont des polynômes paramétrés par un vecteur de paramètres $\tilde{\boldsymbol{\theta}}_{ik}$. De plus, $F_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ est monique et $\dot{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ a un retard égal à zéro. La matrice diagonale $H(z, \boldsymbol{\eta})$ est composée de n_y fonctions de transfert indépendamment paramétrées $H_i(z, \boldsymbol{\eta}_i)$ dans sa diagonale avec $H_i(z, \boldsymbol{\eta}_i) = C_i(z, \boldsymbol{\eta}_i)/D_i(z, \boldsymbol{\eta}_i)$ où $C_i(z, \boldsymbol{\eta}_i)$ et $D_i(z, \boldsymbol{\eta}_i)$ sont des polynômes moniques paramétrés par un vecteur de paramètres $\boldsymbol{\eta}_i$. Les coefficients polynomiaux de $\dot{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ et de $F_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik})$ sont les paramètres de $\tilde{\boldsymbol{\theta}}_{ik}$ alors que ceux de

$C_i(z, \boldsymbol{\eta}_i)$ et de $D_i(z, \boldsymbol{\eta}_i)$ sont les paramètres de $\boldsymbol{\eta}_i$. Par conséquent, nous avons que :

$$\begin{aligned}\hat{B}_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) &= \tilde{\theta}_{ik,1} + \sum_{m=1}^{\deg(\hat{B}_{ik})} \tilde{\theta}_{ik,(m+1)} z^{-m} \\ F_{ik}(z, \tilde{\boldsymbol{\theta}}_{ik}) &= 1 + \sum_{m=1}^{\deg(F_{ik})} \tilde{\theta}_{ik,(m+\deg(\hat{B}_{ik})+1)} z^{-m} \\ C_i(z, \boldsymbol{\eta}_i) &= 1 + \sum_{m=1}^{\deg(C_i)} \eta_{i,m} z^{-m} \\ D_i(z, \boldsymbol{\eta}_i) &= 1 + \sum_{m=1}^{\deg(D_i)} \eta_{i,(m+\deg(C_i))} z^{-m}\end{aligned}$$

où $\tilde{\theta}_{ik,m}$ est le m -ième élément de $\tilde{\boldsymbol{\theta}}_{ik}$ et $\eta_{i,m}$ le m -ième élément de $\boldsymbol{\eta}_i$. Le nombre de paramètres à identifier présents dans \hat{B}_{ik} , F_{ik} , D_i et dans C_i est alors égal à $\deg(\hat{B}_{ik})+1$, $\deg(F_{ik})$, $\deg(D_i)$ et à $\deg(C_i)$ respectivement.

Pour toutes ces structures de modèle, le vecteur de paramètres $\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{\tilde{n}}$ est la concaténation des vecteurs $\tilde{\boldsymbol{\theta}}_{ik}$ ($i = 1, \dots, n_y$, $k = 1, \dots, n_u$), i.e., $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\theta}}_{11}^T, \tilde{\boldsymbol{\theta}}_{12}^T, \dots, \tilde{\boldsymbol{\theta}}_{1n_u}^T, \dots, \tilde{\boldsymbol{\theta}}_{n_y 1}^T, \tilde{\boldsymbol{\theta}}_{n_y 2}^T, \dots, \tilde{\boldsymbol{\theta}}_{n_y n_u}^T)^T$. Le vecteur de paramètres $\boldsymbol{\eta} \in \mathbb{R}^{n_\eta}$ résulte aussi de la concaténation des vecteurs $\boldsymbol{\eta}_i$ ($i = 1, \dots, n_y$), i.e., $\boldsymbol{\eta} = (\boldsymbol{\eta}_1^T, \boldsymbol{\eta}_2^T, \dots, \boldsymbol{\eta}_{n_y}^T)^T$.

Les structures de modèle MIMO ARX/FIR sont globalement identifiables en n'importe quel vrai vecteur de paramètres $\boldsymbol{\theta}_0 \in \mathcal{D}_\theta$. Pour les autres structures de modèle MIMO (ARMAX/BJ/OE), elles sont globalement identifiables en $\boldsymbol{\theta}_0$ si $\boldsymbol{\theta}_0$ ne génère pas des compensations pôles/zéros pour les fonctions de transfert de ces structures.

Notations d'indices : Dans la suite, nous réserverons l'indice $i \in \llbracket 1, n_y \rrbracket$ pour référer à la sortie y_i de \mathbf{y} alors que l'indice $k \in \llbracket 1, n_u \rrbracket$ sera dédié pour référer à l'entrée u_k de \mathbf{u} .

H.5 Résultats sur l'informativité des données dans le cas MIMO présents dans la littérature

Comme mentionné dans l'introduction (Annexe H), il y a très peu de résultats sur l'informativité des données dans le cas MIMO présents dans la littérature. Parmi les résultats existants, nous avons ces deux lemmes suivants (un pour la cas boucle ouverte et l'autre pour le cas boucle fermée)

Lemme H.1. *Considérons la Définition H.2 et le jeu infini de données \mathcal{Z}^∞ correspondant à une expérience en boucle ouverte avec une excitation \mathbf{u} . Si la matrice densité spectrale de puissance $\boldsymbol{\Phi}_{\mathbf{u}}(\omega)$ de \mathbf{u} est strictement définie positive pour quasiment toutes les pulsations ω , alors le jeu de données \mathcal{Z}^∞ est informatif par rapport à n'importe quelle structure de modèle, indépendamment de son type et de sa complexité.*

Lemme H.2. *Considérons la Définition H.3 et le jeu infini de données \mathcal{Z}_{cl}^∞ correspondant à une expérience en boucle ouverte avec une excitation externe \mathbf{r} . Si la matrice densité spectrale de puissance $\Phi_{\mathbf{r}}(\omega)$ de \mathbf{r} est strictement définie positive pour quasiment toutes les pulsations ω , alors le jeu de données \mathcal{Z}_{cl}^∞ est informatif par rapport à n'importe quelle structure de modèle, indépendamment de son type et de sa complexité.*

La preuve du Lemme H.2 peut être trouvée dans [BGM10] et le Lemme H.1 est un cas spécial du Lemme H.2 lorsque $\mathbf{K} = \mathbf{0}_{n_u \times n_y}$ et $\mathbf{r} = \mathbf{u}$.

D'après ces lemmes, les vecteurs d'excitation \mathbf{u} (resp. \mathbf{r}) tels que $\Phi_{\mathbf{u}}(\omega) \succ 0$ pour quasiment toutes les pulsations ω de l'intervalle $] -\pi, \pi]$ (resp. $\Phi_{\mathbf{r}}(\omega) \succ 0$ pour quasiment toutes les pulsations ω de l'intervalle $] -\pi, \pi]$) garantiront toujours l'informativité des données et donc des estimées consistantes $\hat{\theta}_N$.

Cette propriété de positivité de la matrice densité spectrale de puissance est satisfaite pour les vecteurs d'excitation de type bruits blancs filtrés tels que $f \geq n_u$ et $\mathbf{M}(e^{j\omega})$ soit de rang plein ligne pour quasiment toutes les pulsations ω de l'intervalle $] -\pi, \pi]$. Ce résultat est donné dans le théorème suivant pour le cas boucle ouverte (le cas boucle fermée étant similaire).

Théorème H.1. *Considérons Définition H.2 et le jeu infini de données \mathcal{Z}^∞ correspondant à une expérience en boucle ouverte avec une excitation \mathbf{u} . Si \mathbf{u} est du type (2.5) avec $f \geq n_u$ et $\text{rank}(\mathbf{M}(e^{j\omega})) = n_u$ pour quasiment toutes les pulsations ω de l'intervalle $] -\pi, \pi]$, alors le jeu de données \mathcal{Z}^∞ est informatif par rapport à n'importe quelle structure de modèle, indépendamment de son type et de sa complexité.*

Démonstration. Quand \mathbf{u} est du type (H.5), sa matrice densité spectrale de puissance $\Phi_{\mathbf{u}}(\omega)$ vaut alors $\Phi_{\mathbf{u}}(\omega) = \mathbf{M}(e^{j\omega})\Sigma_v\mathbf{M}^*(e^{j\omega})$ où $\Sigma_v = E[\mathbf{v}(t)\mathbf{v}^T(t)]$ est une matrice définie positive. Avec l'hypothèse sur $\mathbf{M}(e^{j\omega})$ dans l'énoncé du théorème, nous observons alors que $\Phi_{\mathbf{u}}(\omega)$ est positive définie pour quasiment toutes les pulsations ω de l'intervalle $] -\pi, \pi]$. Comme montré dans le Lemme H.1, un vecteur d'entrée \mathbf{u} satisfaisant cette propriété garantit l'informativité des données par rapport à n'importe quelle structure de modèle, indépendamment de son type et de sa complexité. \square

Un cas particulier du Théorème H.1 est lorsque $f = n_u$ et $\mathbf{M}(z)$ est une matrice diagonale puisque, dans ce cas, $\mathbf{M}(e^{j\omega})$ sera toujours de rang plein ligne pour quasiment toutes les pulsations ω . Ce cas particulier d'excitation est assez classique dans la pratique : il correspond en effet au cas où toutes les entrées sont excitées avec des bruits blancs filtrés indépendants. Cependant, dans certains cas, les limites de l'expérience peuvent entraver ce choix classique et nous pouvons être confrontés à une excitation avec des bruits blancs filtrés générés avec $f < n_u$ bruits blancs indépendants. Dans ce cas, $\det(\Phi_{\mathbf{u}}(\omega)) = 0$ à toutes les pulsations ω et nous ne pouvons plus utiliser le Lemme H.1 pour prouver l'informativité des données. Cependant, dans le résumé de

l'Annexe I, nous verrons que beaucoup de ces excitations particulières peuvent conduire à l'informativité des données.

Le cas de l'excitation multisinus est également un cas où les Lemmes H.1 et H.2 ne peuvent être d'aucune aide. En effet, toutes les excitations multisinus du type (H.4) en boucle ouverte et (H.13) en boucle fermée ont une matrice densité spectrale de puissance dont le déterminant est égal à zéro à toutes les pulsations (preuve en Annexe A.1). Nous montrerons néanmoins dans le résumé de l'Annexe I que de nombreuses excitations multisinus peuvent donner des données informatives.

Dans [BGM10], il existe un autre résultat qui concerne l'informativité des données en boucle fermée lorsque l'excitation \mathbf{r} est nulle. En effet, comme il est également mentionné dans le résumé de l'introduction, la consistance peut parfois être assurée par la seule excitation due à la perturbation $\mathbf{H}_0(z)\mathbf{e}(t)$ renvoyée par le correcteur $\mathbf{K}(z)$. Ce résultat n'est pas rappelé dans le résumé mais peut être trouvé dans cette thèse. Pour cela, se référer au Théorème 2.2. Un résultat similaire concernant uniquement la structure du modèle ARMAX est donné dans [YZ18].

En Annexe I, nous proposerons une condition alternative qui est basée sur une autre approche que celle présentée dans [BGM10]. Cette approche résumée en Annexe I aura l'avantage de pouvoir être étendue au cas classique où le vecteur d'excitation \mathbf{r} est non-nul. Nous pouvons montrer également par un exemple que notre condition alternative pour le cas $\mathbf{r} = \mathbf{0}_{n_u \times 1}$ peut être moins conservatrice que celle du Théorème 2.2.

H.6 Persistance d'excitation de vecteurs de signaux quasi-stationnaires

Une notion importante qui sera nécessaire pour résumer les conditions d'informativité des données dans le cas MIMO de cette thèse est la notion de persistance d'excitation d'un vecteur de signaux. La persistance d'excitation d'un vecteur de signaux (à valeur complexe ou réelle) est définie comme suit :

Définition H.4 (Persistance d'excitation). *Un vecteur de signaux quasi-stationnaires à valeur réelle ou complexe ϕ est excitant de façon persistante (nous garderons l'acronyme anglais PE pour persistently exciting dans la suite) si et seulement si $\bar{E}[\phi(t)\phi^*(t)] \succ 0$.*

Nous avons le résultat suivant concernant les vecteurs de signaux quasi-stationnaires à valeur réelle.

Lemme H.3. *Considérons un vecteur ϕ contenant p signaux quasi-stationnaires ϕ_h à valeur réelle. Les trois propositions suivantes sont équivalentes :*

- (a) ϕ est PE.

$$(b) \bar{E}[\phi(t)\phi^T(t)] \succ 0.$$

$$(c) \forall \alpha \in \mathbb{R}^p, \bar{E}[(\alpha^T \phi(t))^2] = 0 \text{ si et seulement si } \alpha = \mathbf{0}_{p \times 1}.$$

Démonstration. (a) \Leftrightarrow (b) : ceci résulte d'une application simple de la Définition H.4 pour un vecteur de signaux à valeur réelle.

$$(b) \Leftrightarrow (c) : \bar{E}[\phi(t)\phi^T(t)] \succ 0 \text{ est équivalent au fait que, pour tout } \alpha \in \mathbb{R}^p,$$

$$\alpha^T \bar{E}[\phi(t)\phi^T(t)] \alpha = 0 \Leftrightarrow \alpha = \mathbf{0}_{p \times 1}$$

qui est à son tour équivalent à la condition (c). \square

Remarque H.1. La condition (c) du Lemme H.3 signifie que les signaux ϕ_h de ϕ sont linéairement indépendants.

De la Définition H.4 et du Lemme H.3, nous avons également la propriété que, pour tout $m \in \mathbb{N}$, $\phi(t - m)$ est PE si et seulement si $\phi(t)$ est PE et la propriété que, si ϕ est PE, alors tous les sous-vecteurs de ϕ sont également PE (voir Lemme 3.1 du Chapitre 3).

Finalement, considérons un vecteur de signaux ϕ_u composé d'éléments qui dépendent d'un signal scalaire quasi-stationnaire u

$$\phi_u(t) = \begin{pmatrix} u(t - \rho) \\ u(t - \rho - 1) \\ \vdots \\ u(t - \mu) \end{pmatrix} \quad (\text{H.15})$$

pour des valeurs arbitraires de $\rho \in \mathbb{N}$ et de $\mu \in \mathbb{N}$ tel que $\mu \geq \rho$. Un tel vecteur de signaux est appelé régresseur. La dimension de ϕ_u est $\nu = \mu - \rho + 1$. Nous avons le résultat suivant reliant la persistance d'excitation de ϕ_u et la densité spectrale de puissance de u :

Lemme H.4 ([GBM08, Lju99]). *Considérons le régresseur ϕ_u défini par (H.15) avec un signal quasis-stationnaire u . Le vecteur ϕ_u est PE si et seulement si la densité spectrale de puissance $\Phi_u(\omega)$ de u a une amplitude non-nulle en $\nu = \mu - \rho + 1$ pulsations de l'intervalle $]-\pi, \pi]$.*

H.7 Conclusion de l'Annexe H

Dans cette annexe, nous avons présenté des notions générales sur la méthode Prediction Error pour l'identification en boucle ouverte et en boucle fermée directe des systèmes MIMO. Nous avons introduit la définition de l'informativité des données pour les deux configurations. Les structures modèles et les excitations qui sont considérées dans cette

thèse ont également été données. Sur la base de ces notions, nous pouvons résumer les résultats théoriques principaux développés dans cette thèse liés aux Questions de Recherche 1, 2 et 3 données en introduction (Annexe G).

Long résumé : Résumé des Chapitres 3, 4 et 5

Contents

I.1	Résumé du Chapitre 3 (Question de Recherche 1)	246
I.1.1	Simplification de l'étude MIMO en étude MISO	246
I.1.2	Informativité des données pour les systèmes MISO	248
I.1.3	Persistance d'excitation de régresseur ϕ_u	250
I.1.4	Multisinus générant la persistance d'excitation du régresseur	253
I.1.5	Bruits blancs filtrés générant la persistance d'excitation du régresseur	254
I.1.6	Autres conditions de vérification de l'informativité des don- nées non basées sur la persistance d'excitation du régresseur ϕ_u	254
I.1.7	Approche de vérification de l'informativité des données pour l'identification en boucle ouverte des systèmes MIMO	255
I.1.8	Conclusion sur le résumé des résultats du Chapitre 3	256
I.2	Résumé du Chapitre 4 (Question de Recherche 2)	256
I.2.1	Simplification de l'étude	257
I.2.2	Réécriture du côté gauche de (I.15)	258
I.2.3	Réécriture du côté gauche de (I.20)	260
I.2.4	Résultat principal et interprétation	263
I.2.5	Approche de vérification de l'informativité des données pour l'identification en boucle fermée des systèmes MIMO	265
I.2.6	Conclusion sur le résumé des résultats du Chapitre 4	266
I.3	Résumé du Chapitre 5 (Question de Recherche 3)	266
I.3.1	Système non-linéaire considéré, structure de modèle et infor- mativité des données	266
I.3.2	Excitation bruit blanc gaussien	268
I.3.3	Excitation multisinus	269

I.3.4	Méthode pour le calcul de m_2	270
I.3.5	Résultat sur l'informativité des données sans calcul de m_2 .	270
I.3.6	Approche de vérification de l'informativité des données pour l'identification en boucle ouverte du système Hammerstein parallèle donné par (I.27)	271
I.3.7	Conclusion sur le résumé des résultats du Chapitre 5	272

Dans cette annexe, nous résumons les résultats des Chapitres 3, 4 et 5 liés aux Questions de Recherche 1, 2 et 3 respectivement.

I.1 Résumé du Chapitre 3 (Question de Recherche 1)

Dans cette section, le système \mathcal{S} défini par (H.2) est identifié en boucle ouverte dans une des structures de modèle définies dans la Section H.4 avec une excitation \mathbf{u} du type (H.4) ou (H.5). Les résultats donnés dans cette section répondent à la Question de Recherche 1.

I.1.1 Simplification de l'étude MIMO en étude MISO

Avec l'hypothèse d'indépendance entre le vecteur de bruits blancs \mathbf{e} et le vecteur d'entrée \mathbf{u} , nous avons le premier résultat suivant :

Théorème I.1. *Considérons la Définition H.2 et les notations de la Section H.4 pour les structures de modèle \mathcal{M} ARX, FIR, BJ et OE. Définissons l'ensemble*

$$\Delta_{\mathbf{X}} = \{ \Delta \mathbf{X}(z) = \mathbf{X}(z, \boldsymbol{\theta}') - \mathbf{X}(z, \boldsymbol{\theta}'') \mid \boldsymbol{\theta}' \text{ and } \boldsymbol{\theta}'' \text{ in } \mathcal{D}_{\theta} \} \quad (\text{I.1})$$

où $\mathbf{X}(z, \boldsymbol{\theta})$ est défini par

- $\mathbf{X}(z, \boldsymbol{\theta}) = \mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ for ARX/FIR model structure.
- $\mathbf{X}(z, \boldsymbol{\theta}) = \mathbf{G}(z, \tilde{\boldsymbol{\theta}})$ for BJ/OE model structures.

Alors, le jeu de données \mathcal{Z}^{∞} introduit dans la définition H.2 est informatif par rapport à la structure de modèle \mathcal{M} , si et seulement si, pour tout $\Delta \mathbf{X}(z) \in \Delta_{\mathbf{X}}$, la propriété suivante est obtenue

$$\bar{E}[|\Delta \mathbf{X}(z) \mathbf{u}(t)|^2] = 0 \implies \Delta \mathbf{X}(z) \equiv \mathbf{0}_{n_y \times n_u}$$

Démonstration. Voir Annexe B.1 pour la preuve. □

Pour les structures MIMO de type ARMAX, nous ne présenterons pas un résumé de ces dernières puisque le résultat n'est pas simple à montrer. Ce dernier se trouve en Annexe B.5.

Avec la paramétrisation particulière des structures de modèle considérées dans cette dissertation, nous montrons à partir du Théorème I.1 que l'informativité des données peut se vérifier sortie-par-sortie. Ceci est présenté dans le théorème suivant :

Théorème I.2. *Considérons la Définition H.2 et le Théorème I.1 et définissons, à partir de $\Delta_{\mathbf{X}}$ défini par (I.1), les ensembles*

$$\Delta_{\mathbf{X},i} = \{\Delta \mathbf{X}_{i,:}(z) \mid \Delta \mathbf{X}_{i,:}(z) \text{ est la } i\text{-ième ligne de } \Delta \mathbf{X}(z) \in \Delta_{\mathbf{X}}\}$$

pour tous les $i = 1, \dots, n_y$. Alors, le jeu de données \mathcal{Z}^∞ introduit dans la définition H.2 est informatif par rapport à la structure du modèle \mathcal{M} , si et seulement si, pour tous les $i = 1, \dots, n_y$, la propriété suivante est obtenue pour tout $\Delta \mathbf{X}_{i,:}(z) \in \Delta_{\mathbf{X},i}$

$$\bar{E}[|\Delta \mathbf{X}_{i,:}(z) \mathbf{u}(t)|^2] = 0 \implies \Delta \mathbf{X}_{i,:}(z) \equiv \mathbf{0}_{1 \times n_u} \quad (\text{I.2})$$

Démonstration. Voir Annexe B.3 pour la preuve. □

Les Théorèmes I.1 et I.2 sont importants car ils permettent de vérifier si un vecteur d'entrée \mathbf{u} fournit des données informatives pour l'identification d'un système MIMO \mathcal{S} (voir (H.2)) en vérifiant successivement que ce vecteur d'entrée donne des données informatives pour l'identification des n_y systèmes MISO \mathcal{S}_i (H.3) (voir Section H.2). Pour s'en rendre compte, définissons, à partir de \mathcal{Z}^∞ donné par (H.9), les jeux infinis de données

$$\mathcal{Z}_i^\infty = \{\mathbf{x}_i(t) = (y_i(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, \infty\}$$

avec y_i comme indiqué dans (H.3). Ensuite, en utilisant le Théorème I.1, nous observons que la vérification (I.2) pour tout $\Delta \mathbf{X}_{i,:}(z) \in \Delta_{\mathbf{X},i}$ équivaut à vérifier que \mathcal{Z}_i^∞ est informatif par rapport à la structure du modèle MISO \mathcal{M}_i

$$\mathcal{M}_i = \{(\mathbf{G}_{i,:}(z, \boldsymbol{\theta}_i), H_i(z, \boldsymbol{\theta}_i)) \mid \boldsymbol{\theta}_i \in \mathcal{D}_{\boldsymbol{\theta},i}\}$$

où $\mathbf{G}_{i,:}$ est la i -ième ligne de \mathbf{G} et $\mathcal{D}_{\boldsymbol{\theta},i}$ est l'ensemble des vecteurs de paramètres $\boldsymbol{\theta}_i$ tels que $\mathbf{G}_{i,:}(z, \boldsymbol{\theta}_i)$ soit stable et $H_i(z, \boldsymbol{\theta}_i)$ soit stable et inversement stable. Pour résumer, afin de vérifier que \mathcal{Z}^∞ est informatif par rapport à \mathcal{M} , nous avons à vérifier que, pour tous les $i = 1, \dots, n_y$, \mathcal{Z}_i^∞ est informatif par rapport à la structure de modèle MISO \mathcal{M}_i .

Ainsi, dans le paragraphe suivant, nous résumons les résultats sur l'informativité des données pour les systèmes MISO.

I.1.2 Informativité des données pour les systèmes MISO

Par conséquent, nous ne considérerons désormais que les structures de modèle MISO du type \mathcal{M}_i . Redéfinissons quelques notations relatives aux vecteurs ligne $\mathbf{X}_i(z)$ correspondant à la structure du modèle MISO \mathcal{M}_i . Pour une structure de modèle MISO \mathcal{M}_i telle que définie à la fin du paragraphe précédent et qui peut être dérivée des structures MIMO de la Section H.4, $\mathbf{X}_i(z, \boldsymbol{\theta}_i)$ est donné par :

- $\mathbf{X}_i(z, \tilde{\boldsymbol{\theta}}_i) = \mathbf{B}_i(z, \tilde{\boldsymbol{\theta}}_i)$ pour les structures de modèle ARX/FIR.
- $\mathbf{X}_i(z, \tilde{\boldsymbol{\theta}}_i) = \mathbf{G}_i(z, \tilde{\boldsymbol{\theta}}_i)$ pour les structures de modèle OE/BJ.

Par la suite, nous supprimerons les indices i et i :. Il nous faut aussi donner quelques notations sur les fonctions de transfert des structures de modèle MISO considérées.

Dans une structure de modèle MISO ARX/FIR, nous avons donc $\mathbf{X}(z, \tilde{\boldsymbol{\theta}}) = \mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ où $\mathbf{B}(z, \tilde{\boldsymbol{\theta}})$ est un vecteur ligne constitué de n_u polynômes paramétrés indépendamment $B_k(z, \tilde{\boldsymbol{\theta}}_k)$ ($k = 1, \dots, n_u$). Dans le cas BJ/OE, $\mathbf{X}(z, \tilde{\boldsymbol{\theta}}) = \mathbf{G}(z, \tilde{\boldsymbol{\theta}})$ où $\mathbf{G}(z, \tilde{\boldsymbol{\theta}})$ est un vecteur ligne constitué de n_u fonctions de transfert indépendamment paramétrées $G_k(z, \tilde{\boldsymbol{\theta}}_k)$. Dans les deux cas, le vecteur de paramètres $\tilde{\boldsymbol{\theta}} \in \mathbb{R}^{\tilde{n}}$ est la concaténation des $\tilde{\boldsymbol{\theta}}_k$ ($k = 1, \dots, n_u$), c'est-à-dire $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\theta}}_1^T, \tilde{\boldsymbol{\theta}}_2^T, \dots, \tilde{\boldsymbol{\theta}}_{n_u}^T)^T$ et où \tilde{n} est la dimension de $\tilde{\boldsymbol{\theta}}$. Introduisons également la notation suivante pour chaque entrée de ces vecteurs :

$$B_k(z, \tilde{\boldsymbol{\theta}}_k) = z^{-\rho_k} \mathring{B}_k(z, \tilde{\boldsymbol{\theta}}_k)$$

$$G_k(z, \tilde{\boldsymbol{\theta}}_k) = z^{-\rho_k} \frac{\mathring{B}_k(z, \tilde{\boldsymbol{\theta}}_k)}{F_k(z, \tilde{\boldsymbol{\theta}}_k)}$$

où les retards ρ_k peuvent être différents pour chaque $k \in \llbracket 1, n_u \rrbracket$ et où \mathring{B}_k et F_k sont tous deux des polynômes (F_k est d'ailleurs un polynôme monique et \mathring{B}_k a un retard nul). Dans la suite, nous indiquerons le degré de ces polynômes par $\deg(\mathring{B}_k)$ et $\deg(F_k)$, respectivement. Les coefficients des polynômes \mathring{B}_k et F_k sont les paramètres dans $\tilde{\boldsymbol{\theta}}_k$. Par conséquent, nous avons

$$\mathring{B}_k(z, \tilde{\boldsymbol{\theta}}_k) = \tilde{\theta}_{k,1} + \sum_{m=1}^{\deg(\mathring{B}_k)} \tilde{\theta}_{k,(m+1)} z^{-m}$$

$$F_k(z, \tilde{\boldsymbol{\theta}}_k) = 1 + \sum_{m=1}^{\deg(F_k)} \tilde{\theta}_{k,(m+\deg(\mathring{B}_k)+1)} z^{-m}$$

où $\tilde{\theta}_{k,m}$ désigne le m -ième élément de $\tilde{\boldsymbol{\theta}}_k$. Le nombre de paramètres à identifier dans \mathring{B}_k et F_k est donc égal à $\deg(\mathring{B}_k) + 1$ et $\deg(F_k)$ respectivement.

A partir de ces notations et en utilisant la notion de persistance d'excitation définie en Section H.6, nous pouvons donner un résultat fondamental pour l'informativité des données.

Théorème I.3 (Informativité MISO et persistance de régresseur). *Considérons le jeu de données \mathcal{Z}^∞ obtenu en appliquant une excitation \mathbf{u} sur un système MISO (voir Définition H.2). Considérons une structure de modèle \mathcal{M} pour ce système MISO et les notations introduites au-dessus de ce théorème. Alors, le jeu de données \mathcal{Z}^∞ est informatif par rapport à \mathcal{M} si le régresseur $\phi_{\mathbf{u}}$ défini ci-dessous est PE :*

$$\phi_{\mathbf{u}}(t) = \begin{pmatrix} \phi_{u_1}(t) \\ \vdots \\ \phi_{u_{n_u}}(t) \end{pmatrix} \quad \text{with} \quad \phi_{u_k}(t) = \begin{pmatrix} u_k(t - \rho_k) \\ \vdots \\ u_k(t - \mu_k) \end{pmatrix} \quad (\text{I.3})$$

La dimension ν du régresseur $\phi_{\mathbf{u}}$ défini par (I.3) est donné par $\nu = n_u + \sum_{k=1}^{n_u} (\mu_k - \rho_k)$ (la dimension de ϕ_{u_k} étant égal à $\mu_k - \rho_k + 1$ ($k = 1, \dots, n_u$)) et les scalaires μ_k ($k = 1, \dots, n_u$) dans (I.3) sont définis par :

- $\mu_k = \deg(B_k)$ pour les structures de modèle ARX/FIR.
- $\mu_k = \deg(B_k) + \deg(F_k) + 2 \sum_{m=1, m \neq k}^{n_u} \deg(F_m)$ pour les structures de modèle BJ/OE.

Pour les structures de modèle ARX/FIR, la persistance de $\phi_{\mathbf{u}}$ n'est pas seulement une condition suffisante pour l'informativité des données mais aussi une condition nécessaire.

Démonstration. Voir Annexe B.4 pour la preuve. □

En appliquant le Théorème I.3 au cas SISO ($n_u = 1$), nous observons alors que l'informativité des données est garantie si le régresseur $\phi_{\mathbf{u}} = \phi_{u_1}$ de dimension $\nu = \mu_1 - \rho_1 + 1$ est PE. En utilisant le Lemme H.4, la condition est équivalente au fait que la densité spectrale de puissance de u_1 a une amplitude non-nulle pour au moins $\nu = \mu_1 - \rho_1 + 1$ pulsations appartenant à l'intervalle $]-\pi, \pi]$. En calculant μ_1 avec $n_u = 1$, nous retrouvons les résultats classiques de l'informativité des données par rapport aux structures de modèle SISO classiquement utilisées [GBM08]. De plus, dans le cas SISO, la persistance de ϕ_{u_1} est une condition qui est aussi nécessaire pour l'informativité des données, alors que dans le cas MISO ceci n'est plus le cas pour les structures de modèle BJ/OE. Cependant, l'exemple donné en Section 3.8 montre que la condition de persistance d'excitation n'est pas trop conservatrice.

Remarque I.1. *En Annexe B.5, nous prouvons aussi que l'informativité des données dans le cas ARMAX peut être vérifié en analysant la persistance d'excitation d'un certain régresseur. Comme dans le cas BJ/OE, ce résultat est une condition suffisante pour l'informativité des données.*

Ainsi, afin de vérifier l'informativité des données pour une structure de modèle MISO de type ARX, ARMAX, FIR, BJ ou OE, il nous faut vérifier la persistance d'excitation d'un régresseur bien défini dont la complexité dépend de celle de la structure de modèle considérée. La persistance d'excitation de ce régresseur est une condition nécessaire et suffisante pour les structures de modèle ARX et FIR et seulement suffisante pour les structures de modèle BJ, ARMAX et FIR. Nous résumons les résultats développés sur la persistance d'excitation de régresseurs $\phi_{\mathbf{u}}$ dans le prochain paragraphe.

I.1.3 Persistance d'excitation de régresseur $\phi_{\mathbf{u}}$

Dans ce paragraphe, nous donnons les résultats importants pour la persistance d'excitation pour n'importe quel régresseur $\phi_{\mathbf{u}}$ de la forme donné par (I.3) lorsque \mathbf{u} est un vecteur de multisinus donnés par (H.4) ou un vecteur de bruits blancs filtrés donnés par (H.5). Le résultat fondamental qui va être donné ici pour la vérification de la persistance de $\phi_{\mathbf{u}}$ est basé sur le lemme suivant

Lemme I.1. *Considérons un vecteur de signaux quasi-stationnaires ϕ donné par $\phi(t) = \mathbf{T}\varphi(t)$ où \mathbf{T} est une matrice complexe, déterministe et indépendante du temps et φ est un vecteur de signaux quasi-stationnaires à valeur complexe qui a la propriété d'être PE. Alors, ϕ est PE si et seulement si \mathbf{T} est de rang plein ligne.* \square

Démonstration. En raison de la Définition H.4, le lemme sera prouvé si nous montrons que $\bar{E}[\phi(t)\phi^*(t)] = \mathbf{T}\bar{E}[\varphi(t)\varphi^*(t)]\mathbf{T}^* \succ 0$. Puisque φ est PE, nous avons $\bar{E}[\varphi(t)\varphi^*(t)] \succ 0$. Par conséquent, une matrice \mathbf{T} de rang plein ligne est effectivement une condition nécessaire et suffisante pour avoir $\bar{E}[\phi(t)\phi^*(t)] \succ 0$. \square

Ainsi, l'idée de l'approche pour le développement d'une condition nécessaire et suffisante est de réécrire le régresseur $\phi_{\mathbf{u}}$ de la forme $\phi_{\mathbf{u}} = \mathbf{T}\varphi$ lorsque \mathbf{u} est un vecteur de multisinus donnés par (H.4) ou un vecteur de bruits blancs filtrés donnés par (H.5). Pour ces deux cas, nous pouvons réécrire $\phi_{\mathbf{u}}$ de la façon suivante

$$\phi_{\mathbf{u}}(t) = \mathbf{T}^{(sin)}\varphi^{(sin)}(t) \quad (\text{I.4})$$

$$\phi_{\mathbf{u}}(t) = \mathbf{T}^{(fwn)}w^{-1}(z)\varphi^{(fwn)}(t) \quad (\text{I.5})$$

où (I.4) est la réécriture de $\phi_{\mathbf{u}}$ quand \mathbf{u} est de type multisinus et (I.5) est la réécriture de $\phi_{\mathbf{u}}$ quand \mathbf{u} est de type bruits blancs filtrés. Le polynôme $w(z)$ est le plus petit multiple commun des dénominateurs de $\mathbf{M}(z)$. Dans la Section 3.3.3, nous montrons comment construire $\mathbf{T}^{(sin)}$, $\varphi^{(sin)}$, $\mathbf{T}^{(fwn)}$ et $\varphi^{(fwn)}$. Ici, nous ne donnerons que les expressions.

Pour le cas multisinus, la matrice $\mathbf{T}^{(sin)}$ est une matrice bloc-colonne de dimension $\nu \times 2s$ et de la forme

$$\mathbf{T}^{(sin)} = \begin{pmatrix} \mathbf{T}_1^{(sin)} \\ \vdots \\ \mathbf{T}_{n_u}^{(sin)} \end{pmatrix} \quad (\text{I.6})$$

où les matrices $\mathbf{T}_k^{(sin)}$ de dimension $\nu_k \times 2s$ sont données par

$$\mathbf{T}_k^{(sin)} = \begin{pmatrix} \bar{\lambda}_{k1} e^{-j\rho_k \omega_1} & \bar{\lambda}_{k1}^* e^{j\rho_k \omega_1} & \cdots & \bar{\lambda}_{ks} e^{-j\rho_k \omega_s} & \bar{\lambda}_{ks}^* e^{j\rho_k \omega_s} \\ \bar{\lambda}_{k1} e^{-j(\rho_k+1)\omega_1} & \bar{\lambda}_{k1}^* e^{j(\rho_k+1)\omega_1} & \cdots & \bar{\lambda}_{ks} e^{-j(\rho_k+1)\omega_s} & \bar{\lambda}_{ks}^* e^{j(\rho_k+1)\omega_s} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \bar{\lambda}_{k1} e^{-j\mu_k \omega_1} & \bar{\lambda}_{k1}^* e^{j\mu_k \omega_1} & \cdots & \bar{\lambda}_{ks} e^{-j\mu_k \omega_s} & \bar{\lambda}_{ks}^* e^{j\mu_k \omega_s} \end{pmatrix} \quad (\text{I.7})$$

où $\bar{\lambda}_{kl} = \lambda_{kl} e^{j\psi_{kl}}$ (phaseur). Le vecteur de signaux $\varphi^{(sin)}$ de dimension $2s$ est quant à lui défini par

$$\varphi^{(sin)}(t) = \begin{pmatrix} e^{j\omega_1 t} \\ e^{-j\omega_1 t} \\ \vdots \\ \vdots \\ e^{j\omega_s t} \\ e^{-j\omega_s t} \end{pmatrix} \quad (\text{I.8})$$

Pour le cas bruits blancs filtrés $\mathbf{u}(t) = \mathbf{M}(z)\mathbf{v}(t)$, il nous faut d'abord mettre toutes les fonctions de transfert de $\mathbf{M}(z)$ sur le plus petit dénominateur commun $w(z)$ et nous notons $\mathbf{L}(z)$ la matrice de filtres FIR obtenue qui est de même dimension que $\mathbf{M}(z)$. Pour chaque FIR $L_{kq}(z)$ ($k = 1, \dots, n_u$, $q = 1, \dots, f$) de $\mathbf{L}(z) = (L_{kq})_{(k,q) \in \llbracket 1, n_u \rrbracket \times \llbracket 1, f \rrbracket}$, nous noterons par d_{kq} son retard et par n_{kq} son degré. Nous définissons les scalaires suivants

$$\begin{cases} \alpha_q &= \min_{k \in \llbracket 1, n_u \rrbracket} d_{kq} + \rho_k \\ \beta_q &= \max_{k \in \llbracket 1, n_u \rrbracket} n_{kq} + \mu_k \end{cases}$$

Finalement, nous allons considérer la notation suivante

$$\xi = f + \sum_{q=1}^f (\beta_q - \alpha_q)$$

À partir de ces notations, nous pouvons alors donner une expression formelle de la matrice $\mathbf{T}^{(fwn)}$. Cette dernière est une matrice bloc de dimension $\nu \times \xi$ et est de la forme

$$\mathbf{T}^{(fwn)} = \begin{pmatrix} \mathbf{T}_{11}^{(fwn)} & \cdots & \mathbf{T}_{1f}^{(fwn)} \\ \vdots & \cdots & \vdots \\ \mathbf{T}_{n_u 1}^{(fwn)} & \cdots & \mathbf{T}_{n_u f}^{(fwn)} \end{pmatrix} \quad (\text{I.9})$$

où les matrices $\mathbf{T}_{kq}^{(fwn)}$ ($k = 1, \dots, n_u$, $q = 1, \dots, f$) sont de dimension $(\mu_k - \rho_k + 1) \times (\beta_q - \alpha_q + 1)$ et sont définies par

$$\mathbf{T}_{kq}^{(fwn)} = \left(\mathbf{0}_{(\mu_k - \rho_k + 1) \times (d_{kq} + \rho_k - \alpha_q)} \quad \tilde{\mathbf{T}}_{kq}^{(fwn)} \quad \mathbf{0}_{(\mu_k - \rho_k + 1) \times (\beta_q - n_{kq} - \mu_k)} \right) \quad (\text{I.10})$$

avec $\tilde{\mathbf{T}}_{kq}^{(fwn)}$ définie par

$$\tilde{\mathbf{T}}_{kq}^{(fwn)} = \begin{pmatrix} l_{kq}^{(d_{kq})} & \dots & \dots & \dots & l_{kq}^{(n_{kq})} & \dots & \mathbf{0} \\ & \ddots & \dots & \dots & \dots & \ddots & \\ \mathbf{0} & & l_{kq}^{(d_{kq})} & \dots & \dots & \dots & l_{kq}^{(n_{kq})} \end{pmatrix}$$

Le vecteur de signaux $\varphi^{(fwn)}$ est quant à lui défini par

$$\varphi^{(fwn)}(t) = \begin{pmatrix} \mathbf{v}_1(t) \\ \vdots \\ \mathbf{v}_f(t) \end{pmatrix} \quad \text{avec} \quad \mathbf{v}_q(t) = \begin{pmatrix} v_q(t - \alpha_q) \\ v_q(t - \alpha_q - 1) \\ \vdots \\ v_q(t - \beta_q) \end{pmatrix} \quad (\text{I.11})$$

Nous avons alors le résultat principal sur la persistance d'excitation et donc sur l'informativité des données pour la Question de Recherche 1.

Théorème I.4. *Considérons le régresseur $\phi_{\mathbf{u}}$ de dimension ν (voir Théorème I.3) et les expressions (I.4) et (I.5) pour ce régresseur correspondant respectivement au cas multisinus (voir (H.4)) et au cas bruits blancs filtrés (voir (H.5)). Alors, $\phi_{\mathbf{u}}$ est PE si et seulement si*

- $\mathbf{T}^{(sin)}$ est de rang plein ligne (i.e., $\text{rank}(\mathbf{T}^{(sin)}) = \nu$) pour le cas multisinus.
- $\mathbf{T}^{(fwn)}$ est de rang plein ligne (i.e., $\text{rank}(\mathbf{T}^{(fwn)}) = \nu$) pour le cas bruits blancs filtrés.

Démonstration. Examinons d'abord le cas multisinus. Nous observons que le vecteur de signaux $\varphi^{(sin)}$ dans (I.8) est (toujours) PE puisque $\bar{E}[\varphi^{(sin)}(t)\varphi^{(sin)*}(t)] = \frac{1}{4} \cdot \mathbf{I}_{2s} \succ 0$ (voir Annexe B.6). Par conséquent, le résultat découle d'une application directe du Lemme I.1. Considérons maintenant le cas bruits blancs filtrés. Nous observons que $\varphi^{(fwn)}$ dans (I.11) est (toujours) PE puisqu'il contient des éléments de séquences indépendantes de bruits blancs. Par ailleurs, comme $w^{-1}(z)$ est stable, nous avons que $w^{-1}(z)\varphi^{(fwn)}$ est PE. En effet, pour ce filtre $w(z)$, il est simple de montrer à l'aide de la condition (c) de Lemme H.3 que $\varphi^{(fwn)}$ est PE si et seulement si $w^{-1}(z)\varphi^{(fwn)}$ est PE. Par conséquent, une application de Lemme I.1 donne l'équivalence entre une matrice de rang plein ligne $\mathbf{T}^{(fwn)}$ et le fait que $\phi_{\mathbf{u}}(t)$ est PE. \square

Il est facile de construire $\mathbf{T}^{(sin)}$ et $\mathbf{T}^{(fwn)}$ et de vérifier si le signal d'entrée correspondant produira un $\phi_{\mathbf{u}}$ excitant de façon persistante (PE) et donc des données informatives pour la structure de modèle considérée \mathcal{M} . En effet, ces matrices dépendent des paramètres des structures du modèle ρ_k et μ_k et de la paramétrisation du vecteur d'entrée.

Cette paramétrisation du vecteur d'entrée est caractérisée par les amplitudes λ_{kl} , les déphasages ψ_{kl} et les pulsations ω_l (pour le cas multisinus) et par les coefficients des filtres de la matrice $\mathbf{M}(z)$ (pour le cas bruits blancs filtrés).

Le Théorème I.4 nous donne donc un outil pour vérifier si un vecteur d'entrée donné donnera des données informatives pour une structure de modèle donnée. Cependant, ce théorème ne donne pas beaucoup d'indices sur les propriétés que le vecteur d'entrée doit satisfaire pour donner des matrices $\mathbf{T}^{(sin)}$ et $\mathbf{T}^{(fwn)}$ de rang plein ligne. Nous résumons ces indices dans les deux prochains paragraphes.

I.1.4 Multisinus générant la persistance d'excitation du régresseur

Dans cette section, nous donnons deux propriétés importantes¹ que le vecteur d'excitation \mathbf{u} de type multisinus (H.4) doit garantir afin d'avoir la persistance de $\phi_{\mathbf{u}}$. Elles sont données dans le lemme suivant

Lemme I.2. *Considérons un régresseur $\phi_{\mathbf{u}}$ de dimension ν (voir Théorème I.3). Supposons que le signal d'entrée correspondant \mathbf{u} de dimension n_u est une multisinus contenant s pulsations différentes (voir (H.4)). Pour chaque entrée u_k de \mathbf{u} , rappelons que nous notons par s_k le nombre de pulsations pour lesquelles l'amplitude correspondante est non-nulle. Alors, pour que $\phi_{\mathbf{u}}$ soit PE, il est nécessaire d'avoir*

$$s \geq \frac{\nu}{2} \quad (\text{I.12})$$

$$s_k \geq \frac{\mu_k - \rho_k + 1}{2} \quad \forall k \in \llbracket 1, n_u \rrbracket \quad (\text{I.13})$$

Démonstration. Pour (I.13), la preuve est dans la Section 3.3.2 en dessous du Lemme 3.1. Pour (I.12), le nombre de colonnes dans $\mathbf{T}^{(sin)}$ est de $2s$ et elles sont toutes non-nulles (voir Remarque 3.3) tandis que son nombre de lignes est donné par ν . Par conséquent, il est clair que (I.12) est une condition nécessaire pour que $\mathbf{T}^{(sin)}$ soit de rang plein ligne. \square

Dans le cas SISO, l'informativité des données se résume également à la persistance d'un régresseur. En utilisant le Lemme H.4 dans la Section H.6 pour une entrée scalaire multisinus, nous voyons que l'informativité des données dans le cas SISO sera atteinte si et seulement si l'entrée scalaire multisinus contient un nombre de pulsations/sinusoïdes supérieur ou égal à la moitié de la dimension du régresseur correspondant à la structure du modèle considérée. Nous avons une analogie intéressante dans le cas MISO. Pour que $\phi_{\mathbf{u}}$ soit PE, le nombre s de pulsations dans le vecteur d'entrée multisinus doit être supérieur ou égal à la moitié de la dimension du régresseur $\phi_{\mathbf{u}}$ et le nombre s_k de pulsations dans chaque entrée scalaire u_k de ce vecteur d'entrée doit également être supérieur ou égal à la moitié de la dimension de la partie correspondante ϕ_{u_k} de $\phi_{\mathbf{u}}$.

¹D'autres conditions ont été développées et sont présentes dans la Section 3.4.

La principale différence est toutefois que la combinaison de ces conditions n'est pas suffisante pour garantir que $\phi_{\mathbf{u}}$ soit PE pour le vecteur d'entrée \mathbf{u} . Dans l'Exemple 3.4, nous illustrons le fait que ces conditions ne sont pas suffisantes.

I.1.5 Bruits blancs filtrés générant la persistance d'excitation du régresseur

Dans ce paragraphe, nous donnons la condition nécessaire développée pour la persistance d'excitation de $\phi_{\mathbf{u}}$ quand \mathbf{u} est un vecteur de bruits blancs filtrés de la forme (H.5). La voici :

Lemme I.3. *Considérons un régresseur $\phi_{\mathbf{u}}$ de dimension ν (voir Théorème I.3) et qui correspond à un vecteur d'entrée \mathbf{u} comme dans (H.5). Supposons que la matrice $\mathbf{T}^{(fwn)}$ de dimension $\nu \times \xi$ correspondant à ce vecteur d'entrée (voir Section I.1.3) comporte χ colonnes non-nulles ($\chi \leq \xi$). Alors, pour que $\phi_{\mathbf{u}}$ soit PE, il est nécessaire d'avoir*

$$\chi \geq \nu$$

Démonstration. Le nombre de colonnes non-nulles dans $\mathbf{T}^{(fwn)}$ est χ alors que son nombre de lignes est donné par ν . Par conséquent, il est nécessaire que $\chi \geq \nu$ pour que $\mathbf{T}^{(fwn)}$ soit de rang plein ligne. \square

Comme la persistance d'excitation du régresseur $\phi_{\mathbf{u}}$ n'est pas une condition nécessaire pour l'informativité des données dans les cas BJ/OE, nous donnons dans le prochain paragraphe des résultats qui garantissent l'informativité sans pour autant garantir nécessairement la persistance du régresseur $\phi_{\mathbf{u}}$.

I.1.6 Autres conditions de vérification de l'informativité des données non basées sur la persistance d'excitation du régresseur $\phi_{\mathbf{u}}$

Dans certains cas particuliers pour l'entrée \mathbf{u} , nous n'avons pas besoin d'analyser la persistance d'excitation du régresseur $\phi_{\mathbf{u}}$ du Théorème I.3 pour vérifier si ce vecteur d'entrée fournit des données informatives.

Pour le cas du bruit blanc filtré, une telle situation est décrite dans le Théorème H.1 (Section H.5). Pour le cas multisinus, nous avons ce résultat suivant lorsque les entrées u_k sont mutuellement indépendantes.

Lemme I.4. *Considérons le jeu de données \mathcal{Z}^∞ défini par (H.9) obtenu en appliquant un vecteur d'entrée multisinus \mathbf{u} (voir (H.4)) sur un système MISO avec des entrées n_u . Considérons une structure de modèle d'ordre complet \mathcal{M} pour ce système MISO et les notations introduites dans la Section I.1.2. Alors, l'ensemble de données \mathcal{Z}^∞ est informatif par rapport à \mathcal{M} si, pour chaque $k = 1, \dots, n_u$,*

(i) les entrées u_k contiennent un nombre s_k de sinusoides à amplitude non-nulle qui satisfont :

- $s_k \geq \frac{\deg(B_k) - \rho_k + 1}{2}$ quand \mathcal{M} est ARX et FIR.
- $s_k \geq \frac{\deg(B_k) - \rho_k + \deg(F_k) + 1}{2}$ quand \mathcal{M} est BJ et OE.

(ii) les s_k pulsations des sinusoides de u_k sont différentes des s_m pulsations des sinusoides de toutes les entrées u_m ($m \neq k$), i.e., les entrées u_k sont mutuellement indépendantes.

De plus, si (i) et (ii) sont satisfaites, \mathcal{Z}^∞ est informatif pour n'importe quelle valeur des s_k pulsations (non-nulles), des s_k amplitudes (non-nulles) et des s_k déphasages définissant u_k ($k = 1, \dots, n_u$).

Démonstration. Voir Annexe B.7 pour la preuve. □

I.1.7 Approche de vérification de l'informativité des données pour l'identification en boucle ouverte des systèmes MIMO

Dans les paragraphes précédents, nous avons résumé les résultats sur l'informativité des données pour l'identification en boucle ouverte des systèmes MISO. Dans cette section, nous donnons une procédure que l'utilisateur peut suivre pour vérifier l'informativité des données en boucle ouverte à partir d'un vecteur d'entrée \mathbf{u} et d'une structure de modèle MIMO \mathcal{M} .

Premièrement, si \mathbf{u} est du type (H.5) avec $f \geq n_u$ et avec $\text{rank}(\mathbf{M}(e^{j\omega})) = n_u$ pour quasiment toutes les pulsations ω de l'intervalle $]-\pi, \pi]$, alors nous avons l'informativité des données en vertu du Lemme H.1 de la Section H.5 ($\Phi_{\mathbf{u}}(\omega) \succ 0$ pour quasiment toutes les pulsations ω de l'intervalle $]-\pi, \pi]$).

Si ce n'est pas le cas, comme indiqué en dessous du Théorème I.2, nous devons vérifier, pour chaque $i = 1, \dots, n_y$, l'informativité de \mathcal{Z}_i^∞ par rapport aux structures n_y du modèle MISO \mathcal{M}_i qui peuvent être dérivées de \mathcal{M} . Pour chaque $i = 1, \dots, n_y$, cela peut être fait en utilisant les résultats présentés dans les Sections I.1.2, I.1.4, I.1.5, I.1.6.

Dans les deux paragraphes suivants, nous donnons une procédure pour le cas multisinus et une autre pour le cas bruits blancs filtrés lorsque nous ne pouvons pas conclure avec Lemme H.1.

Tout d'abord, pour les vecteurs d'entrée multisinus \mathbf{u} du type (H.4), si \mathbf{u} respecte les conditions de Lemme I.3, alors \mathcal{Z}_i^∞ est informatif. Si ce n'est pas le cas, nous construi-

sons la matrice $\mathbf{T}^{(sin)}$ correspondant à la structure du modèle MISO \mathcal{M}_i et nous vérifions si elle est de rang plein ligne. Notons que $\mathbf{T}^{(sin)}$ ne peut être de rang plein ligne que si (I.13) et (I.12) sont satisfaits. Si $\mathbf{T}^{(sin)}$ est de rang plein ligne, alors \mathcal{Z}_i^∞ est informatif par rapport à \mathcal{M}_i .

Pour les excitations de type bruits blancs filtrés \mathbf{u} (H.5), nous construisons $\mathbf{T}^{(fwn)}$ correspondant à la structure du modèle MISO \mathcal{M}_i et nous vérifions si cette matrice est de rang plein ligne. Notez que cela ne peut être le cas que si la condition du Lemme I.2 est satisfaite. Si $\mathbf{T}^{(fwn)}$ est un rang plein ligne, alors \mathcal{Z}_i^∞ est informatif par rapport à \mathcal{M}_i .

Bien entendu, il peut arriver que, pour au moins un i (i.e., une sortie), la matrice $\mathbf{T}^{(fwn)}$ (ou $\mathbf{T}^{(sin)}$ dans le cas multisinus) ne soit pas de rang plein ligne. Alors, si \mathcal{M} est ARX ou FIR, nous savons que \mathcal{Z}_i^∞ n'est pas informatif par rapport à \mathcal{M}_i et donc que \mathcal{Z}^∞ n'est pas informatif par rapport à \mathcal{M} . Si \mathcal{M} est BJ, ARMAX ou OE, alors nous ne pouvons conclure ni que les données sont informatives ni que les données ne sont pas informatives (puisque le Théorème I.3 et l'Annexe B.5 ne donnent que des conditions suffisantes pour ces structures de modèle).

Les résultats présentés dans cette section sont illustrés dans un exemple présent dans la Section 3.8.

I.1.8 Conclusion sur le résumé des résultats du Chapitre 3

Le but de cette section était de résumer les résultats développés dans le Chapitre 3 pour répondre à la Question de Recherche 1, i.e., des conditions permettant de vérifier l'informativité des données pour l'identification en boucle ouverte des structures de modèle MIMO et ce pour des excitations \mathbf{u} de type bruits blancs filtrés et multisinus. Le résultat principal rappelé dans cette section est la vérification de l'informativité des données en analysant le fait qu'une certaine matrice est de rang plein ligne pour chaque sortie. Cette matrice dépend de la complexité de la structure du modèle et de la paramétrisation de l'excitation. Cette nouvelle condition est clairement moins conservatrice que la seule condition disponible dans la littérature et donnée dans Lemme H.1. Nous avons également obtenu des résultats qui permettent de vérifier l'informativité des données sans vérifier la persistance d'excitation du régresseur puisque cette condition est conservatrice pour les structures des modèles BJ/OE. Sur la base des résultats rappelés dans cette section, nous verrons dans la prochaine section que nous pouvons aborder la Question de Recherche 2.

I.2 Résumé du Chapitre 4 (Question de Recherche 2)

Dans cette section, nous résumons les résultats du Chapitre 4 répondant à la Question de Recherche 2. Ainsi, nous allons considérer que le système \mathcal{S} défini par (H.2) est identifié en boucle fermée dans une des structures de modèle définies dans la Section H.4.

Le correcteur linéaire $\mathbf{K}(z)$ utilisé pour la boucle fermée stabilise le système et une excitation externe \mathbf{r} peut être ajoutée à la loi de commande telle que le vecteur d'entrée \mathbf{u} soit donné par (H.12).

I.2.1 Simplification de l'étude

Avec l'hypothèse d'indépendance entre \mathbf{r} et \mathbf{e} , nous avons le résultat suivant concernant l'informativité des données dans le cas boucle fermée

Théorème I.5. *Considérons la Définition H.3 et, en particulier, le jeu de données \mathcal{Z}_{cl}^∞ obtenu par l'application du signal d'excitation externe \mathbf{r} au système \mathcal{S} en boucle fermée représenté dans la Figure 2.1. Considérons également une des structures de modèle \mathcal{M} définies dans la Section H.4. Rappelons que \mathbf{r} et \mathbf{e} sont supposés être indépendants. Pour chaque $\Delta\mathbf{W}(z) \in \Delta\mathbf{W}$ (voir (H.10) dans la Définition H.3), nous définissons de manière similaire $\Delta\mathbf{W}_y(z)$ et $\Delta\mathbf{W}_u(z)$ (voir (H.6) et (H.7) respectivement). Alors le jeu de données \mathcal{Z}_{cl}^∞ défini par (H.9) est informatif par rapport à la structure du modèle \mathcal{M} si et seulement si, pour tout $\Delta\mathbf{W}(z) = (\Delta\mathbf{W}_y(z), \Delta\mathbf{W}_u(z)) \in \Delta\mathbf{W}$, la propriété suivante est obtenue*

$$\begin{cases} \Delta\mathbf{W}_y(z) - \Delta\mathbf{W}_u(z)\mathbf{K}(z) & \equiv \mathbf{0}_{n_y \times n_y} \\ \bar{E} [|\Delta\mathbf{W}_u(z)\mathbf{r}(t)|^2] & = 0 \end{cases} \implies (\Delta\mathbf{W}_y, \Delta\mathbf{W}_u) \equiv (\mathbf{0}_{n_y \times n_y}, \mathbf{0}_{n_y \times n_u})$$

Démonstration. Voir l'Annexe C.1 pour la preuve. □

Le Théorème I.5 montre la contribution individuelle du correcteur $\mathbf{K}(z)$ et de l'excitation externe \mathbf{r} sur l'informativité des données. Comme dans le cas boucle ouverte avec le Théorème I.2, nous avons un théorème similaire dans le cas boucle fermée sur la simplification de la vérification de l'informativité des données

Théorème I.6. *Considérons la Définition H.3, le Théorème I.5 et une des structures de modèle \mathcal{M} définies dans la Section H.4. À partir de l'ensemble $\Delta\mathbf{W}$ (voir (H.10) de la Définition H.3), définissons les ensembles*

$$\Delta\mathbf{W}_{\mathbf{i}} = \{\Delta\mathbf{W}_{i,:}(z) \mid \Delta\mathbf{W}_{i,:}(z) \text{ est la } i\text{-ième ligne de } \Delta\mathbf{W}(z) \in \Delta\mathbf{W}\} \quad (\text{I.14})$$

pour tous les $i = 1, \dots, n_y$. Pour chaque $\Delta\mathbf{W}_{i,:}(z) \in \Delta\mathbf{W}_{\mathbf{i}}$, nous définissons de façon similaire $\Delta\mathbf{W}_{u,i}(z)$ et $\Delta\mathbf{W}_{y,i}(z)$ qui sont les i -ièmes lignes de $\Delta\mathbf{W}_u(z)$ et $\Delta\mathbf{W}_y(z)$ respectivement. Alors, le jeu de données \mathcal{Z}_{cl}^∞ est informatif par rapport à la structure de modèle \mathcal{M} si et seulement si, pour tous les $i = 1, \dots, n_y$, la propriété suivante est obtenue pour tout $\Delta\mathbf{W}_{i,:}(z) = (\Delta\mathbf{W}_{y,i}(z), \Delta\mathbf{W}_{u,i}(z)) \in \Delta\mathbf{W}_{\mathbf{i}}$:

$$\begin{cases} \Delta\mathbf{W}_{y,i}(z) - \Delta\mathbf{W}_{u,i}(z)\mathbf{K}(z) & \equiv \mathbf{0}_{1 \times n_y} \\ \bar{E} [|\Delta\mathbf{W}_{u,i}(z)\mathbf{r}(t)|^2] & = 0 \end{cases} \implies (\Delta\mathbf{W}_{y,i}, \Delta\mathbf{W}_{u,i}) \equiv (\mathbf{0}_{1 \times n_y}, \mathbf{0}_{1 \times n_u}) \quad (\text{I.15})$$

Démonstration. Voir l'Annexe C.3 pour la preuve. □

Le Théorème I.6 permet de simplifier la vérification de l'informativité des données : elle peut se faire sortie-par-sortie². Par conséquent, dans la suite, nous considérerons un i arbitraire et limiterons l'attention à (I.15) pour ce i particulier.

Pour simplifier la présentation dans la suite, nous utiliserons souvent des notations plus courtes pour chaque fonction de transfert de la forme $M(z)$, $M(z, \theta')$ et $M(z, \theta'')$ en enlevant l'argument et en les désignant respectivement par M , M' et M'' .

I.2.2 Réécriture du côté gauche de (I.15)

Afin de donner le résultat principal du Chapitre 4, il nous faut réécrire le côté gauche de (I.15), qui est composé de deux équations avec des inconnues rationnelles $\Delta \mathbf{W}_{y,i}(z)$ et $\Delta \mathbf{W}_{u,i}(z)$, en deux équations avec des inconnues polynomiales. Ceci nous permettra alors de donner la condition principale pour la vérification de l'informativité des données. Pour cela, observons que la première équation du côté gauche de (I.15) peut s'écrire de la façon suivante

$$(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z)) \begin{pmatrix} \mathbf{I}_{n_y} \\ \mathbf{K}(z) \end{pmatrix} \equiv \mathbf{0}_{1 \times n_y} \quad (\text{I.16})$$

En considérant une factorisation à droite du correcteur $\mathbf{K}(z)$ et une factorisation à gauche du terme $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$, nous prouverons que nous pouvons transformer (I.16) en une forme polynomiale. De plus, nous verrons que cette factorisation à gauche de $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ nous permettra de transformer la deuxième équation du côté gauche de (I.15) en une forme polynomiale.

Donnons d'abord une factorisation à gauche pour la matrice rationnelle $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ qui doit être valable pour tout $\Delta \mathbf{W}_i(z) = (\Delta \mathbf{W}_{y,i}(z), \Delta \mathbf{W}_{u,i}(z)) \in \Delta \mathbf{W}_i$ avec $\Delta \mathbf{W}_i$ l'ensemble défini par (I.14). La factorisation qui est considéré est obtenue en mettant toutes les entrées de $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ sur le plus petit commun dénominateur. Cette factorisation est donnée dans le lemme suivant.

Lemme I.5. *Considérons les structures de modèle définies dans la Section H.4. Pour ces structures de modèle, la matrice bloc rationnelle $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ peut être factorisée à gauche en $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z)) = Q_i^{-1}(z)(\Upsilon_{y,i}(z), \Upsilon_{u,i}(z))$ où le vecteur ligne polynomial $\Upsilon_{y,i}(z)$ de dimension n_y a toutes ses entrées égales à 0 sauf peut-être la i -ième entrée, notée par $\Upsilon_{y,ii}(z)$. Le vecteur ligne polynomial $\Upsilon_{u,i}(z)$ est de dimension n_u . Les polynômes scalaires $Q_i(z)$ et $\Upsilon_{y,ii}(z)$ et le vecteur ligne polynomial $\Upsilon_{u,i}(z)$ sont donnés par*

- pour FIR : $Q_i = 1$, $\Upsilon_{y,ii} = 0$, $\Upsilon_{u,i} = \mathbf{B}_i'' - \mathbf{B}_i'$.

²Toutefois, nous n'avons pas la même simplification que dans le cas de la boucle ouverte. Dans le cas de la boucle ouverte, nous pouvons simplifier l'étude de l'informativité des données par rapport à une structure de modèle MIMO en considérant les structures de modèle MISO correspondantes. En effet, notons ici que la première équation du côté gauche de (I.15) contient toujours le correcteur MIMO $\mathbf{K}(z)$.

- pour ARX : $Q_i = 1$, $\Upsilon_{y,ii} = A_i'' - A_i'$, $\Upsilon_{u,i} = \mathbf{B}_i'' - \mathbf{B}_i'$.
- pour ARMAX : $Q_i = C_i' C_i''$, $\Upsilon_{y,ii} = C_i' A_i'' - C_i'' A_i'$, $\Upsilon_{u,i} = C_i' \mathbf{B}_i'' - C_i'' \mathbf{B}_i'$.
- pour OE : $Q_i = \prod_{k=1}^{n_u} F_{ik}' F_{ik}''$, $\Upsilon_{y,ii} = 0$ and $\Upsilon_{u,i} = (\Upsilon_{u,ik})_{k \in \llbracket 1, n_u \rrbracket}$ with

$$\Upsilon_{u,ik} = (B_{ik}'' F_{ik}' - B_{ik}' F_{ik}'') \prod_{l=1, l \neq k}^{n_u} F_{il}' F_{il}''$$

- pour BJ : $Q_i = C_i' C_i'' \prod_{k=1}^{n_u} F_{ik}' F_{ik}''$, $\Upsilon_{y,ii} = D_i'' C_i' \prod_{k=1}^{n_u} F_{ik}' F_{ik}'' - D_i' C_i'' \prod_{k=1}^{n_u} F_{ik}' F_{ik}''$ and $\Upsilon_{u,i} = (\Upsilon_{u,ik})_{k \in \llbracket 1, n_u \rrbracket}$ with

$$\Upsilon_{u,ik} = D_i'' C_i' B_{ik}'' F_{ik}' \prod_{l=1, l \neq k}^{n_u} F_{il}' F_{il}'' - D_i' C_i'' B_{ik}' F_{ik}'' \prod_{l=1, l \neq k}^{n_u} F_{il}' F_{il}''$$

Démonstration. Voir l'Annexe C.4 pour la preuve. □

Pour le correcteur $\mathbf{K}(z)$, nous allons considérer la factorisation à droite consistant à mettre toutes les entrées de $\mathbf{K}(z)$ sur le plus petit commun dénominateur $J(z)$

$$\mathbf{K}(z) = \mathbf{N}(z) \mathbf{V}^{-1}(z) \quad (\text{I.17})$$

où $\mathbf{N}(z)$ est une matrice composé de FIR de dimension $n_u \times n_y$ et où $\mathbf{V}(z)$ est donnée par $\mathbf{V}(z) = \underbrace{\text{diag}(J(z), \dots, J(z))}_{n_y \text{ times}}$.

En se basant sur la factorisation de $\mathbf{K}(z)$ donnée par (I.17) et celle du Lemme I.5, nous pouvons transformer (I.15) en une forme de matrice polynomiale. Pour ce faire, en se basant sur l'ensemble $\Delta_{\mathbf{W},i}$ défini par (I.14) ($i = 1, \dots, n_y$), nous définissons l'ensemble \mathbf{Y}_i composé de toutes les matrices polynomiales $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))$ obtenu en considérant la factorisation à gauche de $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ donnée dans le Lemme I.5, i.e.,

$$\begin{aligned} \mathbf{Y}_i = \{ & (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \mid (\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z)) = Q_i^{-1}(z) (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \\ & \text{où } \Delta \mathbf{W}_{i:}(z) = (\Delta \mathbf{W}_{y,i:}(z), \Delta \mathbf{W}_{u,i:}(z)) \in \Delta_{\mathbf{W},i} \text{ et } Q_i(z) \text{ est défini} \\ & \text{dans le Lemme I.5} \} \end{aligned} \quad (\text{I.18})$$

Théorème I.7. *Considérons la Définition H.3, le Théorème I.6 et une des structures de modèle \mathcal{M} définies dans la Section H.4. Considérons la factorisation de $\mathbf{K}(z)$ donnée par (I.17). Pour tous les $i = 1, \dots, n_y$, notons*

$$\Xi_i(z) = \begin{pmatrix} \mathbf{V}_{i:}(z) \\ \mathbf{N}(z) \end{pmatrix} \quad (\text{I.19})$$

où $\mathbf{V}_{i:}(z)$ est la i -ième ligne de $\mathbf{V}(z)$.

Pour chaque $\Delta \mathbf{W}_i(z) = (\Delta \mathbf{W}_{y,i}(z), \Delta \mathbf{W}_{u,i}(z)) \in \Delta_{\mathbf{W},i}$, considérons la factorisation à gauche $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z)) = Q_i^{-1}(z)(\Upsilon_{y,i}(z), \Upsilon_{u,i}(z))$ donnée dans le Lemme I.5 et notons par $\Upsilon_{y,ii}$ la i -ième entrée du vecteur ligne polynomial $\Upsilon_{y,i}$. Alors, le jeu de données \mathcal{Z}_{cl}^∞ donné par (H.9) est informatif par rapport à la structure de modèle \mathcal{M} si et seulement si, pour tous les $i = 1, \dots, n_y$, la propriété suivante est satisfaite

$$\begin{cases} (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \Xi_i(z) \equiv \mathbf{0} \\ \bar{E} [||\Upsilon_{u,i}(z)r(t)||^2] = 0 \end{cases} \implies (\Upsilon_{y,ii}, \Upsilon_{u,i}) \equiv (0, \mathbf{0}_{1 \times n_u}) \quad (\text{I.20})$$

pour tout $(\Upsilon_{y,ii}, \Upsilon_{u,i}) \in \mathbf{Y}_i$ où \mathbf{Y}_i est défini par (I.18).

Démonstration. Voir l'Annexe C.5 pour la preuve. \square

Remarque I.2. La factorisation de $(\Delta \mathbf{W}_{y,i}(z), -\Delta \mathbf{W}_{u,i}(z))$ et $\mathbf{K}(z)$ n'est bien sûr pas unique. Les factorisations choisies ne sont donc qu'un des choix possibles. Toutefois, il est important de noter que ce choix particulier n'introduit aucun conservatisme puisque la condition d'informativité des données du Théorème I.6 et celle du Théorème I.7 sont équivalentes.

I.2.3 Réécriture du côté gauche de (I.20)

Nous avons montré dans le précédent paragraphe comment transformer (I.15) en une condition à inconnues polynomiales (I.20) où les inconnues sont le polynôme $\Upsilon_{y,ii}(z)$ et le vecteur ligne $\Upsilon_{u,i}(z)$ de polynômes. Afin de donner le résultat principal du Chapitre 4, nous avons besoin de réécrire (I.20) dans une forme adéquate. Pour cela, introduisons quelques notations supplémentaires.

Pour chaque polynôme $\Upsilon_{u,ik}(z)$ ($k = 1, \dots, n_u$) du vecteur ligne $\Upsilon_{u,i}(z)$ et pour le polynôme $\Upsilon_{y,ii}(z)$, nous pouvons écrire

$$\begin{aligned} \Upsilon_{y,ii}(z) &= \tilde{\delta}_{y,i}^T \mathbf{Z}_{y,i}(z) \\ \Upsilon_{u,ik}(z) &= \tilde{\delta}_{u,ik}^T \mathbf{Z}_{u,ik}(z) \end{aligned}$$

où $\mathbf{Z}_{y,i}(z) = (z^{-1}, \dots, z^{-\mu_{y,i}})^T$, $\mathbf{Z}_{u,ik}(z) = (z^{-\rho_{ik}}, \dots, z^{-\mu_{u,ik}})^T$ et $\tilde{\delta}_{y,i}$, $\tilde{\delta}_{u,ik}$ sont des vecteurs de coefficients. Les valeurs de $\mu_{y,i}$, $\mu_{u,ik}$ ($i = 1, \dots, n_y$, $k = 1, \dots, n_u$) en fonction du type de structure du modèle et de la complexité de la structure du modèle sont données en Annexe C.6. Nous pouvons alors écrire que

$$\Upsilon_{u,i}(z) = \tilde{\delta}_{u,i}^T \mathbf{Z}_{u,i}(z)$$

où $\tilde{\delta}_{u,i}^T = (\tilde{\delta}_{u,i1}^T, \dots, \tilde{\delta}_{u,in_u}^T)$ et $\mathbf{Z}_{u,i}(z) = \text{bdiag}(\mathbf{Z}_{u,i1}(z), \dots, \mathbf{Z}_{u,in_u}(z))$ où la notation bdiag est définie dans la Section Notations en page xx. Nous pouvons finalement écrire que

$$(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) = \tilde{\delta}_i^T \text{bdiag}(\mathbf{Z}_{y,i}(z), \mathbf{Z}_{u,i}(z)) \quad (\text{I.21})$$

où $\tilde{\delta}_i^T = (\tilde{\delta}_{y,i}^T, \tilde{\delta}_{u,i}^T)$. Nous noterons par ζ_i et $\zeta_{u,i}$ les dimensions de $\tilde{\delta}_i$ et $\tilde{\delta}_{u,i}$ respectivement. Ces dimensions peuvent être facilement déduites à partir des entiers $\mu_{y,i}$,

$\mu_{u,ik}$ ($i = 1, \dots, n_y$, $k = 1, \dots, n_u$) donnés en Annexe C.6. Introduisons également l'ensemble

$$\mathbf{D}_i = \{\tilde{\boldsymbol{\delta}}_i \mid (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) = \tilde{\boldsymbol{\delta}}_i^T \text{bdiag}(\mathbf{Z}_{y,i}(z), \mathbf{Z}_{u,i}(z)) \in \mathbf{Y}_i\} \quad (\text{I.22})$$

Notons que, sauf dans le cas FIR où $\mathbf{D}_i = \mathbb{R}^{\zeta_i}$, \mathbf{D}_i est un sous-ensemble de \mathbb{R}^{ζ_i} . Il convient de noter qu'un tel ensemble \mathbf{D}_i a également été considéré dans le cas de la boucle ouverte (voir Annexe B.4).

Réécriture de la première équation du côté gauche de (I.20) : Commençons par réécrire la première équation du côté gauche de (I.20), i.e., $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))\boldsymbol{\Xi}_i(z) \equiv \mathbf{0}$. De l'expression de $\boldsymbol{\Xi}_i(z)$ donné par (I.19) et de celle de $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))$ donnée par (I.21), nous pouvons réécrire cette équation comme donné ci-dessous

$$\underbrace{\begin{pmatrix} \tilde{\boldsymbol{\delta}}_{y,i}^T & \tilde{\boldsymbol{\delta}}_{u,i}^T \end{pmatrix}}_{\tilde{\boldsymbol{\delta}}_i^T} \underbrace{\begin{pmatrix} \mathbf{Z}_{y,i}(z)\mathbf{V}_{i:}(z) \\ \mathbf{Z}_{u,i}(z)\mathbf{N}(z) \end{pmatrix}}_{\mathbf{U}^{(i)}(z)} = \mathbf{0} \quad (\text{I.23})$$

En notant respectivement par $\gamma_{i,min}$ et $\gamma_{i,max}$ le minimum des retards et le maximum des degrés des entrées de $\mathbf{U}^{(i)}(z)$, nous pouvons réécrire (I.23) comme ci-dessous

$$\underbrace{\begin{pmatrix} \tilde{\boldsymbol{\delta}}_{y,i}^T & \tilde{\boldsymbol{\delta}}_{u,i}^T \end{pmatrix}}_{\tilde{\boldsymbol{\delta}}_i^T} \begin{pmatrix} \mathbf{R}^{(i)} \\ \mathbf{S}^{(i)} \end{pmatrix} (\mathbf{I}_{n_y} \otimes \mathbf{Z}_{U^{(i)}}(z)) = \mathbf{0} \quad (\text{I.24})$$

où $\mathbf{Z}_{U^{(i)}}(z) = (z^{-\gamma_{i,min}}, \dots, z^{-\gamma_{i,max}})^T$. Comme (I.24) est une matrice de polynômes qui doivent tous être nuls, il est clair que (I.24) est équivalent à

$$\underbrace{\begin{pmatrix} \tilde{\boldsymbol{\delta}}_{y,i}^T & \tilde{\boldsymbol{\delta}}_{u,i}^T \end{pmatrix}}_{\tilde{\boldsymbol{\delta}}_i^T} \begin{pmatrix} \mathbf{R}^{(i)} \\ \mathbf{S}^{(i)} \end{pmatrix} = \mathbf{0}$$

Lorsque la structure de modèle est du type OE ou FIR, nous avons que $\Upsilon_{y,ii}(z) = 0$ (Lemme I.5) et donc (I.25) devient dans ce cas

$$\tilde{\boldsymbol{\delta}}_{u,i}^T \mathbf{S}^{(i)} = \mathbf{0}$$

Ainsi, la première équation du côté gauche de (I.20) peut se réécrire de façon équivalente sous la forme suivante

$$\left\{ \begin{array}{l} \underbrace{\begin{pmatrix} \tilde{\boldsymbol{\delta}}_{y,i}^T & \tilde{\boldsymbol{\delta}}_{u,i}^T \end{pmatrix}}_{\tilde{\boldsymbol{\delta}}_i^T} \begin{pmatrix} \mathbf{R}^{(i)} \\ \mathbf{S}^{(i)} \end{pmatrix} = \mathbf{0} \quad (\text{cas BJ/ARX/ARMAX}) \\ \tilde{\boldsymbol{\delta}}_{u,i}^T \mathbf{S}^{(i)} = \mathbf{0} \quad (\text{cas OE/FIR}) \end{array} \right. \quad (\text{I.25})$$

Plus de détails et des exemples sur cette réécriture sont donnés dans le Chapitre 4.

Réécriture de la deuxième équation du côté gauche de (I.20) : Ici, nous réécrivons l'équation $\bar{E} [|\Upsilon_{u,i}(z)\mathbf{r}(t)|^2] = 0$. Cette équation a déjà été étudiée dans le Chapitre 3 (et résumé dans la Section I.1). L'idée est d'introduire un régresseur comme illustré ci-dessous

$$\begin{aligned}\Upsilon_{u,i}(z)\mathbf{r}(t) &= \sum_{k=1}^{n_u} \Upsilon_{u,ik}(z)r_k(t) \\ &= \sum_{k=1}^{n_u} \tilde{\delta}_{u,ik}^T \phi_{r_k,i}(t) \\ &= \underbrace{\left(\tilde{\delta}_{u,i1}^T \quad \tilde{\delta}_{u,i2}^T \quad \cdots \quad \tilde{\delta}_{u,in_u}^T \right)}_{\tilde{\delta}_{u,i}^T} \phi_{\mathbf{r},i}(t)\end{aligned}$$

où

$$\phi_{\mathbf{r},i}(t) = \begin{pmatrix} \phi_{r_1,i}(t) \\ \phi_{r_2,i}(t) \\ \vdots \\ \phi_{r_{n_u},i}(t) \end{pmatrix} \quad \text{with} \quad \phi_{r_k,i}(t) = \begin{pmatrix} r_k(t - \rho_{ik}) \\ r_k(t - \rho_{ik} - 1) \\ \vdots \\ r_k(t - \mu_{u,ik}) \end{pmatrix}$$

Ensuite, nous avons vu que nous pouvons réécrire le régresseur $\phi_{\mathbf{r},i}(t)$ sous la forme $\phi_{\mathbf{r},i}(t) = \mathbf{T}^{(i)}\varphi_i(t)$ où $\mathbf{T}^{(i)}$ est une matrice à valeur réelle ou complexe déterministe et indépendante du temps et φ_i un vecteur de signaux qui est PE. Ainsi, l'équation $\bar{E} [|\Upsilon_{u,i}(z)\mathbf{r}(t)|^2] = 0$ devient $\tilde{\delta}_{u,i}^T \bar{E} [\phi_{\mathbf{r},i}(t)\phi_{\mathbf{r},i}^*(t)] \tilde{\delta}_{u,i} = 0$ ce qui devient à son tour équivalent à $\tilde{\delta}_{u,i}^T \mathbf{T}^{(i)} \bar{E} [\varphi_i(t)\varphi_i^*(t)] \mathbf{T}^{(i)*} \tilde{\delta}_{u,i} = 0$. Comme φ_i est PE, nous avons que $\bar{E} [\varphi_i(t)\varphi_i^*(t)] \succ 0$ ce qui permet de conclure que $\tilde{\delta}_{u,i}^T \mathbf{T}^{(i)} \bar{E} [\varphi_i(t)\varphi_i^*(t)] \mathbf{T}^{(i)*} \tilde{\delta}_{u,i} = 0$ est équivalent à

$$\tilde{\delta}_{u,i}^T \mathbf{T}^{(i)} = 0 \tag{I.26}$$

Lorsque \mathbf{r} est un multisinus du type (H.13), la matrice $\mathbf{T}^{(i)}$ est égale à $\mathbf{T}^{(sin)}$ donnée par (I.6)-(I.7) mais en remplaçant les scalaires ρ_k et μ_k par ρ_{ik} et $\mu_{u,ik}$ respectivement ($k = 1, \dots, n_u$). Lorsque \mathbf{r} est du type bruits blancs filtrés (H.14), la matrice $\mathbf{T}^{(i)}$ est égale à $\mathbf{T}^{(fwn)}$ donnée par (I.9)-(I.10) mais en remplaçant aussi les scalaires ρ_k et μ_k par ρ_{ik} et $\mu_{u,ik}$ respectivement ($k = 1, \dots, n_u$).

A partir des deux réécritures (I.25) et (I.26), nous pouvons donner le résultat principal qui répond à la Question de Recherche 2. Il est donné dans le paragraphe suivant.

I.2.4 Résultat principal et interprétation

Théorème I.8. *Considérons la Définition H.3 et, en particulier, le jeu de données \mathcal{Z}_{cl}^∞ obtenu par l'application du signal d'excitation \mathbf{r} au système \mathcal{S} en boucle fermée de la Figure 2.1. Considérons le Théorème I.7 et une des structures de modèle \mathcal{M} définies dans la Section H.4. Considérons enfin les notations introduites dans la Section I.2.3 et, en particulier, les équations (I.25) et (I.26) qui sont respectivement équivalentes à la première et deuxième équation du côté gauche de (I.20). Le jeu de données \mathcal{Z}_{cl}^∞ est informatif par rapport à \mathcal{M} si, pour chaque $i = 1, \dots, n_y$, au moins une des matrices $\mathbf{P}^{(a,i)}$, $\mathbf{P}^{(b,i)}$ et $\mathbf{P}^{(c,i)}$ définies ci-dessous est de rang plein ligne.*

- (a) La matrice $\mathbf{P}^{(a,i)}$ est égale à $\begin{pmatrix} \mathbf{R}^{(i)} \\ \mathbf{S}^{(i)} \end{pmatrix}$ pour le cas BJ/AMAX/ARX et à $\mathbf{S}^{(i)}$ pour le cas OE/FIR.
- (b) La matrice $\mathbf{P}^{(b,i)}$ est égale à $\mathbf{T}^{(i)}$.
- (c) La matrice $\mathbf{P}^{(c,i)}$ est égale à $\begin{pmatrix} \mathbf{R}^{(i)} & \mathbf{0} \\ \mathbf{S}^{(i)} & \mathbf{T}^{(i)} \end{pmatrix}$ pour le cas BJ/AMAX/ARX et à $\begin{pmatrix} \mathbf{S}^{(i)} & \mathbf{T}^{(i)} \end{pmatrix}$ pour le cas OE/FIR.

De plus, lorsque $\mathbf{r} = \mathbf{0}_{n_u \times 1}$, le jeu de données \mathcal{Z}_{cl}^∞ est informatif par rapport à \mathcal{M} si, pour chaque $i = 1, \dots, n_y$, la matrice $\mathbf{P}^{(a,i)}$ définie au-dessus est de rang plein ligne.

Démonstration. Nous allons prouver le théorème pour les structures de modèle du type BJ/ARX/ARMAX. La preuve pour le cas OE/FIR peut être obtenue en utilisant la même argumentation. Notons d'abord que $\tilde{\boldsymbol{\delta}}_i = (\tilde{\boldsymbol{\delta}}_{y,i}^T, \tilde{\boldsymbol{\delta}}_{u,i}^T)^T$ défini par (I.25)-(I.26) est contraint d'appartenir à l'ensemble \mathbf{D}_i avec \mathbf{D}_i comme défini dans (I.22). Lorsque $\tilde{\boldsymbol{\delta}}_i \in \mathbf{D}_i$, une matrice $\mathbf{P}^{(a,i)}$ de rang plein ligne est une condition suffisante³ pour que $\tilde{\boldsymbol{\delta}}_i = \mathbf{0}$ soit la solution unique de (I.25). Notons également que $\tilde{\boldsymbol{\delta}}_i = \mathbf{0}$ est équivalent à $(\Upsilon_{y,ii}, \Upsilon_{u,i}) = (0, \mathbf{0})$. Par conséquent, en utilisant l'équivalence rappelée dans l'énoncé du théorème, lorsque $\mathbf{P}^{(a,i)}$ est de rang plein ligne, nous avons aussi que $(\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z))\Xi_i(z) \equiv \mathbf{0} \implies (\Upsilon_{y,ii}, \Upsilon_{u,i}) \equiv (0, \mathbf{0})$ pour tout $(\Upsilon_{y,ii}, \Upsilon_{u,i}) \in \mathbf{Y}_i$ avec \mathbf{Y}_i comme défini par (I.18). Il est clair, d'après le Théorème I.7, que ce dernier implique que \mathcal{Z}_{cl}^∞ est informatif (en particulier lorsque $\mathbf{r} = \mathbf{0}_{n_u \times 1}$).

Notons maintenant que (I.25) et (I.26) peuvent être combinées en $\tilde{\boldsymbol{\delta}}_i^T \mathbf{P}^{(c,i)} = \mathbf{0}$. Si $\mathbf{P}^{(c,i)}$ est de rang plein ligne, alors $\tilde{\boldsymbol{\delta}}_i^T \mathbf{P}^{(c,i)} = \mathbf{0}$ implique $\tilde{\boldsymbol{\delta}}_i = \mathbf{0}$. Par conséquent, en utilisant l'équivalence rappelée dans l'énoncé du théorème, lorsque $\mathbf{P}^{(c,i)}$ est de rang

³Si \mathbf{D}_i était égal à \mathbb{R}^{c_i} , alors ce serait une condition nécessaire et suffisante.

plein ligne, nous avons aussi que (I.20) est garantie pour tous les $(\Upsilon_{y,ii}, \Upsilon_{u,i}) \in \mathbf{Y}_i$. Il est clair, selon le Théorème I.7, que ce dernier implique que \mathcal{Z}_{cl}^∞ est informatif.

Enfin, si $\mathbf{P}^{(b,i)}$ est de rang plein ligne, l'équation (I.26) implique $\tilde{\delta}_{u,i} = \mathbf{0}$. Par conséquent, en utilisant l'équivalence rappelée dans l'énoncé du théorème, lorsque $\mathbf{P}^{(b,i)}$ est de rang plein ligne, nous avons également cela : $\bar{E} [||\Upsilon_{u,i}(z)\mathbf{r}(t)||^2] = 0 \implies \Upsilon_{u,i} \equiv \mathbf{0}$ pour tous $\Upsilon_{u,i}(z) \in \{\Upsilon_{u,i}(z) \mid (\Upsilon_{y,ii}(z), \Upsilon_{u,i}(z)) \in \mathbf{Y}_i\}$. Puisque $\Upsilon_{u,i} \equiv \mathbf{0}$, la première équation du côté gauche de (I.20) se réduit à $\Upsilon_{y,ii}(z)J(z) \equiv 0$ où $J(z)$ est le polynôme défini après (I.17). Ce dernier implique toujours que $\Upsilon_{y,ii} \equiv 0$. Par conséquent, en vertu du Théorème I.7, nous avons également l'informativité des données lorsque $\mathbf{P}^{(b,i)}$ est de rang plein ligne. \square

Le Théorème I.8 montre que l'on peut vérifier si une excitation externe donnée \mathbf{r} fournit des données informatives en vérifiant, pour chaque sortie, si l'une des trois matrices $\mathbf{P}^{(a,i)}$, $\mathbf{P}^{(b,i)}$ ou $\mathbf{P}^{(c,i)}$ est de rang plein ligne.

En particulier, si $\mathbf{P}^{(a,i)}$ est de rang plein ligne pour chaque sortie i , alors l'ensemble de données \mathcal{Z}_{cl}^∞ fournira des données informatives même si l'excitation externe \mathbf{r} est égale à zéro.

Pour que $\mathbf{P}^{(a,i)}$ ait un rang plein ligne, le nombre de ses colonnes doit être supérieur à son nombre de lignes ζ_i . Pour que cette condition nécessaire soit remplie, le correcteur $\mathbf{K}(z)$ doit avoir une complexité suffisante.

Même si $\mathbf{P}^{(a,i)}$ n'est pas de rang plein ligne, nous pouvons bien sûr obtenir des données informatives en ajoutant une excitation non-nulle \mathbf{r} du type (H.13) ou (H.14). Dans ce cas, l'informativité des données peut être vérifiée en vérifiant si $\mathbf{P}^{(c,i)}$ est de rang plein ligne. On observe que $\mathbf{P}^{(a,i)}$ et $\mathbf{P}^{(c,i)}$ ont le même nombre ζ_i de lignes, mais le nombre de colonnes de $\mathbf{P}^{(c,i)}$ est supérieure à celle de $\mathbf{P}^{(a,i)}$ (en raison de la matrice $\mathbf{T}^{(i)}$ liée à l'excitation externe). Par conséquent, même si $\mathbf{P}^{(a,i)}$ n'est pas de rang plein ligne, l'ajout de colonnes supplémentaires peut compenser l'insuffisance de rang de $\mathbf{P}^{(a,i)}$ et rendre $\mathbf{P}^{(c,i)}$ de rang plein ligne. En particulier, une excitation multisinus du type (H.13) avec s sinusoïdes différentes ajoute $2s$ de colonnes à celles déjà présentes dans $\mathbf{P}^{(a,i)}$. Pour une excitation par bruits blancs filtrés du type (H.14), nous avons, comme indiqué dans la Section I.1.5, la relation suivante : plus le nombre f de bruits blancs est grand dans \mathbf{v} et plus la complexité des filtres dans $\mathbf{M}(z)$ est grande, plus le nombre de colonnes ajoutées à $\mathbf{P}^{(a,i)}$ est grand.

La considération ci-dessus permet de déduire une condition pour garantir l'informativité des données lorsque $\mathbf{P}^{(a,i)}$ n'est pas de rang plein ligne pour chaque sortie i . La richesse de l'excitation (par exemple, le nombre de sinusoïdes dans le cas d'un multisinus) doit être suffisante pour garantir que, pour chaque sortie i , le nombre de colonnes dans $\mathbf{T}^{(i)}$ est supérieur à la différence entre le nombre ζ_i de lignes de $\mathbf{P}^{(a,i)}$ et le rang de

$\mathbf{P}^{(a,i)}$. Si c'est le cas, il faut tout de même vérifier que $\mathbf{P}^{(c,i)}$ est de rang plein ligne pour chaque i car une insuffisance de rang peut se produire en raison de certaines combinaisons malheureuses de coefficients du correcteur et de la paramétrisation de l'excitation externe (amplitudes, pulsations et déphasages pour le cas multisinus et coefficients des filtres dans $\mathbf{M}(z)$ pour le cas bruits blancs filtrés). Un phénomène similaire a également été observé dans le cas de la boucle ouverte au Chapitre 3.

Dans le cas de la boucle ouverte, nous avons montré que notre condition d'informativité des données MIMO (Théorème I.3) est similaire à la condition d'informativité des données SISO de [GBM08] lorsque nous utilisons le Théorème I.3 avec $n_u = n_y = 1$. Cela a été possible grâce au fait que la persistance d'excitation du régresseur était à l'origine des résultats des cas SISO et MIMO. Dans le cas en boucle fermée, cette comparaison est beaucoup plus compliquée puisque la vérification du rang matriciel du Théorème I.8 est difficile à mettre en relation avec les conditions sur les degrés du numérateur et du dénominateur du correcteur et sur les degrés des polynômes impliqués de la structure du modèle dans le cas du SISO. Cependant, nous pouvons encore faire quelques comparaisons entre le cas MIMO et le cas SISO. Ceci est illustré dans l'Exemple 4.2 dans lequel nous appliquons le Théorème I.8 pour vérifier l'informativité des données pour l'identification d'un système SISO ARMAX en boucle fermée avec deux correcteurs SISO. Il a été illustrée que la condition du Théorème I.8 peut être conservatrice pour les structures du modèle SISO ARMAX.

I.2.5 Approche de vérification de l'informativité des données pour l'identification en boucle fermée des systèmes MIMO

Dans cette section, nous avons résumé les résultats sur l'informativité des données pour l'identification directe en boucle fermée des systèmes MIMO développés dans le Chapitre 4. Dans ce paragraphe, nous faisons une synthèse des résultats en donnant une procédure que l'utilisateur peut suivre pour vérifier l'informativité des données avec un correcteur $\mathbf{K}(z)$, une excitation externe non-nulle \mathbf{r} et une structure de modèle MIMO \mathcal{M} .

Premièrement, si le vecteur d'excitation externe est du type (H.14) avec $f \geq n_u$ et avec $\text{rank}(\mathbf{M}(e^{j\omega})) = n_u$ pour presque toutes les pulsations ω de l'intervalle $]-\pi, \pi]$, alors nous avons l'informativité des données en utilisant le Lemme H.2 de la Section H.5.

Si ce n'est pas le cas, nous devons utiliser les résultats résumés dans cette section. Premièrement, pour toutes les sorties $i = 1, \dots, n_y$, nous construisons et calculons le rang des matrices $\mathbf{P}^{(a,i)}$. Si elles sont toutes de rang plein ligne, nous concluons que le correcteur $\mathbf{K}(z)$ garantit seul l'informativité des données et que les données sont informatives même en présence d'une excitation externe. Toutefois, si la matrice $\mathbf{P}^{(a,i)}$ n'est pas de rang plein ligne pour certaines sorties i , nous construisons les matrices correspondantes $\mathbf{P}^{(c,i)}$ et nous vérifions qu'elles sont toutes de rang plein ligne. Si c'est le cas, alors le jeu de données \mathcal{Z}_d^∞ est informatif par rapport à \mathcal{M} . Sinon, nous ne pouvons pas conclure sur l'informativité des données, sauf pour les structures de modèle

FIR pour lesquelles la condition de rang est une condition nécessaire et suffisante.

I.2.6 Conclusion sur le résumé des résultats du Chapitre 4

Par rapport à la Question de Recherche 2, nous avons résumé dans cette section les résultats du Chapitre 4 qui consistent en une condition principale permettant de vérifier si un vecteur d'excitation donné \mathbf{r} du type (H.13)-(H.14) donne des données informatives lorsque le système fonctionne en boucle fermée avec un correcteur $\mathbf{K}(z)$ LTI MIMO donné. Cette condition est similaire à celle développée au Chapitre 3 puisqu'elle consiste à vérifier si une matrice est de rang plein ligne pour chaque sortie. Cette matrice dépend de la complexité de la structure du modèle, des coefficients du correcteur et de la paramétrisation de l'excitation externe. Dans la littérature, seul le cas sans excitation externe (*costless* en anglais) a été pris en compte pour le développement des résultats sur l'informativité des données des systèmes MIMO. Le seul résultat qui existe dans la littérature pour des excitations externes non-nulles est la positivité définie de la matrice densité spectrale de puissance de \mathbf{r} qui est trop restrictive. La principale contribution de notre condition est que nous pouvons vérifier l'informativité des données pour les excitations externes \mathbf{r} pour lesquelles la matrice densité spectrale de puissance de \mathbf{r} dégénère à toutes les pulsations. Cette condition sépare aussi clairement la contribution du correcteur et de l'excitation externe sur l'informativité des données.

Après avoir longuement étudié le cas linéaire dans les Chapitres 3 et 4, nous résumons les résultats sur le cas non-linéaire du Chapitre 5 (Question de Recherche 3) en étudiant l'informativité des données pour l'identification d'une classe particulière de systèmes de type Hammerstein parallèle.

I.3 Résumé du Chapitre 5 (Question de Recherche 3)

I.3.1 Système non-linéaire considéré, structure de modèle et informativité des données

Le système qui est considéré dans le Chapitre 5 et dans ce résumé est un système SISO non-linéaire d'entrée u et de sortie y défini par

$$\mathcal{S} : \quad y(t) = G_{0,1}(z)u(t) + G_{0,2}(z)u^n(t) + H_0(z)e(t) \quad (\text{I.27})$$

où $G_{0,1}(z)$ et $G_{0,2}(z)$ sont des fonctions de transfert stables, $H_0(z)$ est une fonction de transfert stable, inversement stable et monique, e est un bruit blanc et n est entier qui sera supposé connu et tel que $n \geq 2$. Nous supposons que u et e sont indépendants (c'est-à-dire que l'expérience d'identification est réalisée en boucle ouverte). Par conséquent, u^n et e sont également indépendants. La Figure 5.1 illustre le système \mathcal{S} considéré dans (I.27).

Afin d'exploiter les résultats du Chapitre 3 (résumés en Section I.1), nous pouvons réécrire (I.27) comme un système MISO linéaire avec un vecteur d'entrée $\mathbf{u} = (u, u^n)^T$

$$\mathcal{S} : \quad y(t) = \mathbf{G}_0(z)\mathbf{u}(t) + H_0(z)e(t)$$

où $\mathbf{G}_0(z) = (G_{0,1}(z), G_{0,2}(z))$.

Il peut donc être identifié dans la structure du modèle MISO (linéaire) suivante

$$\mathcal{M} = \{(\mathbf{G}(z, \boldsymbol{\theta}), H(z, \boldsymbol{\theta})) \mid \boldsymbol{\theta} \in \mathcal{D}_\theta\} \quad (\text{I.28})$$

avec $\mathbf{G}(z, \boldsymbol{\theta}) = (G_1(z, \boldsymbol{\theta}), G_2(z, \boldsymbol{\theta}))$ où $G_1(z, \boldsymbol{\theta})$, $G_2(z, \boldsymbol{\theta})$ et $H(z, \boldsymbol{\theta})$ sont des fonctions de transfert paramétrées par $\boldsymbol{\theta}$. L'ensemble \mathcal{D}_θ contient les vecteurs de paramètres $\boldsymbol{\theta}$ tels que $G_1(z, \boldsymbol{\theta})$ et $G_2(z, \boldsymbol{\theta})$ soient stables et $H(z, \boldsymbol{\theta})$ soit stable, inversement stable et monique. Nous considérerons que la structure du modèle dans (I.28) est d'ordre complet⁴ et est globalement identifiable en $\boldsymbol{\theta}_0$. Enfin, nous limiterons notre attention aux structures MISO ayant la forme ARX, FIR, BJ, OE afin de garder la présentation simple (voir la Section I.1.2 pour les notations de ces structures de modèle particulières⁵).

Adaptons légèrement la Définition H.2 afin de définir formellement l'informativité des données pour le problème examiné dans le Chapitre 5 et résumé dans cette section.

Définition I.1. *Considérons le cadre défini ci-dessus et supposons qu'une entrée quasi-stationnaire u soit appliqué au système réel \mathcal{S} défini par (I.27) de $t = 1$ à $t = +\infty$, ce qui donne l'ensemble de données infini \mathcal{Z}^∞ :*

$$\mathcal{Z}^\infty = \left\{ \mathbf{x}(t) = (y(t), \mathbf{u}^T(t))^T \mid t = 1, \dots, +\infty \right\} \quad (\text{I.29})$$

où $\mathbf{u} = (u, u^n)^T$. Considérons une structure de modèle MISO \mathcal{M} avec $n_u = 2$ générant le prédicteur $\hat{y}(t, \boldsymbol{\theta}) = \mathbf{W}(z, \boldsymbol{\theta})\mathbf{x}(t)$. Définissons aussi l'ensemble suivant

$$\Delta_{\mathbf{W}} = \{ \Delta \mathbf{W}(z) = \mathbf{W}(z, \boldsymbol{\theta}') - \mathbf{W}(z, \boldsymbol{\theta}'') \mid \boldsymbol{\theta}' \text{ and } \boldsymbol{\theta}'' \text{ in } \mathcal{D}_\theta \}$$

Le jeu de données \mathcal{Z}^∞ est dit informatif par rapport à la structure de modèle \mathcal{M} quand, pour tout $\Delta \mathbf{W}(z) \in \Delta_{\mathbf{W}}$, nous avons

$$\bar{E} [\|\Delta \mathbf{W}(z)\mathbf{x}(t)\|^2] = 0 \implies \Delta \mathbf{W}(z) \equiv \mathbf{0}_{1 \times 3}$$

Comme nous voulons une estimation consistante de $(G_{0,1}(z), G_{0,2}(z), H_0(z))$, l'objectif du Chapitre 5 est de développer des résultats⁶ afin de vérifier si une expérience donnée avec une excitation u donnera des données informatives \mathcal{Z}^∞ (voir (I.29)).

À cette fin, nous utiliserons les résultats de la littérature et ceux que nous avons déjà développés pour les systèmes MISO dans le Chapitre 3. En effet, ces résultats sont valables pour des vecteurs d'entrée arbitraires \mathbf{u} et donc aussi pour des vecteurs d'entrée du type $\mathbf{u} = (u, u^n)^T$. Ces résultats peuvent être trouvés dans le Lemme H.1 (positivité de la matrice de la densité spectrale de puissance d'entrée), Théorème I.1 (informativité

⁴i.e., $\exists \boldsymbol{\theta}_0$ tel que $(\mathbf{G}(z, \boldsymbol{\theta}_0), H(z, \boldsymbol{\theta}_0)) = (\mathbf{G}_0(z), \mathbf{H}_0(z))$.

⁵Notons que nous avons dans ce cas $n_u = 2$.

⁶Dans cette section, nous allons faire un résumé de ces résultats.

pour les structures de modèle MIMO en boucle ouverte) et Théorème I.3 (informativité et persistance d'excitation du régresseur). Dans le prochain lemme, nous donnons un résultat supplémentaire qui sera utile dans la suite et qui porte sur l'informativité des données par rapport aux structures de modèle MISO pour une identification en boucle ouverte.

Lemme I.6. *Considérons le MISO LTI \mathcal{S} défini par (H.2) avec deux entrées ($n_u = 2$) et qui est identifié dans une structure de modèle MISO telle que définie dans la Section I.1.2. Supposons que le vecteur d'entrée $\mathbf{u} = (u_1, u_2)^T$ est un multisinus appliqué au système réel \mathcal{S} . Notons par Φ_{u_1} (resp. Φ_{u_2}) la densité spectrale de puissance (DSP) de u_1 (resp. u_2). Supposons que Φ_{u_1} (resp. Φ_{u_2}) soit non-nulle en m_1 (resp. m_2) pulsations appartenant à l'intervalle $]-\pi, \pi]$. Enfin, supposons que Φ_{u_1} et Φ_{u_2} sont simultanément non-nulles en $m_{1,2}$ pulsations dans l'intervalle $]-\pi, \pi]$ ($m_{1,2}$ peut être égal à zéro). Alors, le jeu de données \mathcal{Z}^∞ est informatif par rapport à \mathcal{M} si $m_1 \geq \varpi_1 - \rho_1 + 1$ et $m_2 - m_{1,2} \geq \varpi_2 - \rho_2 + 1$ où ρ_1 et ρ_2 sont les retards de $G_1(z, \boldsymbol{\theta})$ et $G_2(z, \boldsymbol{\theta})$ respectivement. Les scalaires ϖ_k ($k = 1, 2$) sont donnés par*

- $\varpi_k = \deg(B_k)$ lorsque \mathcal{M} est ARX/FIR.
- $\varpi_k = \deg(B_k) + \deg(F_k)$ lorsque \mathcal{M} est OE/BJ.

Démonstration. Voir l'Annexe D.1 pour la preuve. □

Comme déjà mentionné, Lemme I.6 et les résultats du Chapitre 3 (résumé dans la Section I.1) se rapportent à un vecteur d'entrée arbitraire \mathbf{u} avec $n_u = 2$. Ils peuvent donc également être utilisés dans le cas où le vecteur d'entrée $\mathbf{u} = (u_1, u_2)^T$ est de la forme $u_1 = u$ et $u_2 = u^n$. Ce fait sera utilisé dans la suite pour donner les résultats de l'informativité des données pour le système non-linéaire de type Hammerstein \mathcal{S} donné par (5.1). Nous considérerons deux types d'excitation : bruit blanc gaussien et multisinus. Dans le paragraphe suivant, nous analysons le cas de l'excitation bruit blanc gaussien.

I.3.2 Excitation bruit blanc gaussien

Lorsque u est un bruit blanc gaussien, nous avons toujours l'informativité des données comme donné dans le résultat suivant :

Theorem I.1. *Considérons le système \mathcal{S} donné par (I.27) identifié dans une structure du modèle \mathcal{M} donnée par (I.28) avec $\mathbf{u} = (u, u^n)^T$. Considérons également la Définition I.1. Supposons que nous appliquons un bruit blanc gaussien de moyenne nulle u au système \mathcal{S} . Alors, le jeu de données \mathcal{Z}^∞ avec $\mathbf{u} = (u, u^n)^T$ est toujours informatif, indépendamment de la complexité de la structure du modèle.*

Démonstration. Voir l'Annexe D.2 pour la preuve. □

Ce type d'excitation stochastique est intéressant car il permet d'identifier tout système du type (I.27). Considérons maintenant le cas où l'entrée u du système (I.27) est un multisinus.

I.3.3 Excitation multisinus

Dans ce paragraphe, u est un multisinus de la forme

$$u(t) = \sum_{l=1}^{s_1} \lambda_l \cos(\omega_l t + \psi_l) \quad (\text{I.30})$$

où λ_l et ψ_l sont respectivement l'amplitude et le déphasage de la sinusoïde à la pulsation ω_l appartenant à l'intervalle $]0, \pi]$. Comme u est un multisinus, u^n est aussi un multisinus. Nous pouvons donc vérifier l'informativité des données en vérifiant la persistance d'excitation du régresseur ϕ_u associée à la structure de modèle \mathcal{M} . Nous pouvons faire ceci en construisant la matrice $\mathbf{T}^{(sin)}$ et vérifier son rang. Or, ceci nous contraint à calculer les amplitudes, fréquences et déphasages de u^n ce qui peut être lourd en terme de calculs lorsque la puissance de non-linéarité n et le nombre s_1 de sinusoïdes augmentent. Nous proposons alors des résultats plus simples dans cette section.

Ce résultat est basé sur le fait que $\mathbf{u} = (u, u^n)^T$ est un vecteur multisinus dont la seconde entrée $u_2 = u^n$ a une densité spectrale de puissance plus riche en contenu fréquentiel que $u_1 = u$. Nous pouvons alors exploiter le Lemme I.6 :

Théorème I.9. *Considérons que le système \mathcal{S} donné par (I.27) est excité avec le multisinus u donné par (I.30) avec $\omega_l \in]0, \pi[$ ($l = 1, \dots, s_1$). Définissons m_2 et $m_{1,2}$ comme dans le Lemme I.6 pour $\Phi_{u_1} = \Phi_u$ et $\Phi_{u_2} = \Phi_{u^n}$. Alors, le jeu de données \mathcal{Z}^∞ défini par (I.29) est informatif par rapport à \mathcal{M} (voir (I.28)) si $2s_1 \geq \varpi_1 - \rho_1 + 1$ et si $m_2 - m_{1,2} \geq \varpi_2 - \rho_2 + 1$ où ϖ_1 et ϖ_2 sont définis dans le Lemme I.6.*

Démonstration. Conséquence du Lemme I.6. □

Pour utiliser la condition (suffisante) d'informativité des données du Théorème I.9, il suffit de savoir combien de pulsations sont présentes dans u^n et de les comparer avec celles de u (voir (I.30)).

À cette fin, nous pouvons calculer la transformée de Fourier rapide de la séquence $\{u^n(t) \mid t = 1, \dots, N'\}$ pour une valeur suffisamment importante de N' . Une autre procédure à cette fin est donnée dans le paragraphe suivant. Dans cette procédure, nous supposons que $\lambda_l > 0$ et $\psi_l = 0$ ($l = 1, \dots, s_1$) pour le multisinus (I.30). Voir la Remarque I.3 dans la suite pour plus de détails.

I.3.4 Méthode pour le calcul de m_2

Considérons donc (I.30) avec $\lambda_l > 0$, $\psi_l = 0$ et $\omega_l \in]0, \pi[$ ($l = 1, \dots, s_1$) et observons que $u^n(t) = u^{n-1}(t)u(t)$, $u^{n-1}(t) = u^{n-2}(t)u(t)$, \dots , $u^2(t) = u(t)u(t)$. Par conséquent, pour obtenir les sinusoides de u^n , nous pouvons d'abord déterminer celles de u^2 à partir de celles de u , puis celles de u^3 à partir de celles de u^2 , \dots , celles de u^{n-1} à partir de celles de u^{n-2} et enfin celles de u^n à partir de celles de u^{n-1} .

Voyons maintenant comment obtenir les pulsations de u^p à partir de celles de u^{p-1} ($p = 2, \dots, n$). Pour cela, nous savons que u^{p-1} est un multisinus et donc $u^p(t) = u^{p-1}(t)u(t)$ peut s'écrire comme la somme des produits de chaque sinusoides de u par chaque sinusoides de u^{p-1} . En utilisant le fait que $\cos(\alpha)\cos(\beta) = 1/2(\cos(\alpha + \beta) + \cos(\alpha - \beta)) \quad \forall(\alpha, \beta) \in \mathbb{R}^2$, tous les produits de cette somme peuvent être écrits comme la somme de deux sinusoides.

En faisant cela, nous voyons que l'ensemble des pulsations dans u^p peut être déterminé en ajoutant et en soustrayant ω_l ($l = 1, \dots, s_1$) aux pulsations de chaque sinusoides présente dans u^{p-1} . Bien entendu, dans l'ensemble obtenu, il faut supprimer les doublons (par exemple, la pulsation $-\omega'$ est équivalente à la pulsation ω' et la pulsation $\omega' + 2k\pi$ avec $k \in \mathbb{Z}$ est équivalente à ω'). On obtient alors un ensemble de pulsations dans l'intervalle $[0, \pi]$. Désignons par s_2 le nombre de pulsations dans cet ensemble. Alors, m_2 est donné par

- $m_2 = 2s_2$ si l'ensemble ne contient ni 0 ni π .
- $m_2 = 2s_2 - 1$ si l'ensemble contient 0 ou π .
- $m_2 = 2s_2 - 2$ si l'ensemble contient 0 et π .

Remarque I.3. *Comme déjà mentionné, la procédure ci-dessus n'est pas garantie de calculer la bonne valeur de m_2 si les déphasages ne sont pas tous égaux à zéro. Pour le montrer, considérons le multisinus : $u(t) = 1/2 \cos(0, 4t) + \cos(0, 6t) + \cos(0.5t + \pi/2)$ et appliquons la procédure ci-dessus pour $n = 2$: nous obtenons l'ensemble des pulsations 0, 0, 1, 0, 2, 0, 8, 0, 9, 1, 1, 1 et 1, 2. Cependant, la pulsation 1 n'apparaît pas dans u^2 en raison de l'annulation des termes sinusoidaux à cette pulsation. Notons cependant qu'une telle situation est rare dans la pratique et que nous pouvons donc utiliser avec précaution la procédure de cette section lorsque les déphasages sont non-nuls. Néanmoins, une vérification à l'aide de la transformée de Fourier rapide peut toujours être utile. \square*

I.3.5 Résultat sur l'informativité des données sans calcul de m_2

Dans certains cas particuliers, nous pouvons vérifier l'informativité des données sans calculer m_2 . Un de ces cas particuliers est donné dans le résultat suivant

Lemme I.7. *Considérons que l'excitation de (I.27) (avec $n \geq 2$) est donnée par $u(t) = \sum_{l=1}^{s_1} \lambda_l \cos(\omega_l t)$ avec $\lambda_l > 0$ et $\omega_l \in]0, \pi[$ ($l = 1, \dots, s_1$) et tel que les s_1 pulsations $n\omega_l$ soient*

- (i) *toutes différentes entre elles (modulo 2π).*
- (ii) *toutes différentes de π et de 0 (modulo 2π).*
- (iii) *toutes différentes de chaque pulsation ω_h ($h = 1, \dots, s_1$) (modulo 2π).*

Alors, le jeu de données \mathcal{Z}^∞ est informatif par rapport à \mathcal{M} si

$$s_1 \geq \frac{1}{2} \max(\varpi_1 - \rho_1 + 1, \varpi_2 - \rho_2 + 1)$$

où ϖ_1 et ϖ_2 sont définis dans le Lemme I.6.

Démonstration. En suivant la procédure de la Section I.3.4, nous observons que $n\omega_1, \dots, n\omega_{s_1}$ sont des pulsations du multisinus u^n . Comme ces s_1 pulsations positives sont toutes différentes (modulo 2π) et différentes de π (modulo 2π), cela implique que la DSP de u^n sera non-nulle à au moins $2s_1$ de pulsations en $]-\pi, \pi[$, c'est-à-dire $m_2 \geq 2s_1$. Du fait que les s_1 pulsations $n\omega_l$ sont (modulo 2π) différentes des pulsations ω_h ($h = 1, \dots, s_1$) de u , nous avons aussi que $m_2 - m_{1,2} \geq 2s_1$. Le résultat découle alors du Théorème I.9. \square

I.3.6 Approche de vérification de l'informativité des données pour l'identification en boucle ouverte du système Hammerstein parallèle donné par (I.27)

Résumons maintenant les différents résultats du Chapitre 5 en donnant une approche générale pour vérifier l'informativité des données par rapport à \mathcal{M} . Si u est un bruit blanc gaussien, le jeu de données \mathcal{Z}^∞ est informatif, indépendamment de la complexité de la structure du modèle.

Concentrons-nous maintenant sur le cas multisinus. Tout d'abord, nous vérifions que $s_1 \geq \varpi_1 - \rho_1 + 1$. Si ce n'est pas le cas, l'ensemble de données \mathcal{Z}^∞ n'est certainement pas informatif par rapport à \mathcal{M} . Si $2s_1 \geq \varpi_1 - \rho_1 + 1$, nous pouvons vérifier si la condition (conservatrice) du Lemme I.7 est satisfaite. Si ce n'est pas le cas, nous calculons m_2 et $m_{1,2}$ (en utilisant l'approche avec la transformée de Fourier rapide ou la procédure de la Section I.3.4) et nous vérifions la condition (moins conservatrice) de Théorème I.9. Si cette condition n'est toujours pas validée, nous devons calculer l'expression complète de u^n (avec toutes les amplitudes et déphasages) et utiliser les résultats du Chapitre 3 sur cette expression et celle de u .

Cette procédure est illustrée avec un exemple dans la Section 5.6.

I.3.7 Conclusion sur le résumé des résultats du Chapitre 5

L'objectif du Chapitre 5, dont la présente section est le résumé, était de répondre à la Question de Recherche 3 en étudiant l'informativité des données pour une classe particulière de systèmes de type Hammerstein parallèle donnée par (I.27) avec la méthode Prediction Error. L'idée qui a été suivie dans le chapitre était de réécrire ce système non-linéaire en un système MISO LTI avec deux entrées $u_1 = u$ et $u_2 = u^n$. Cela nous a permis d'utiliser les résultats précédents développés dans le Chapitre 3 tels que le Théorème I.1. Ensuite, nous avons analysé l'informativité des données pour deux types d'excitations : le bruit blanc gaussien et le multisinus. L'excitation bruit blanc gaussien est un choix efficace puisqu'elle garantit toujours des données informatives pour ce type de systèmes Hammerstein parallèle. Lorsque nous optons pour une excitation multisinus, une vérification simple peut être effectuée avec les Lemmes I.6 et I.7 qui ont été développés afin de simplifier la vérification de l'informativité des données.

Sur la base de cette étude, nous verrons plus tard que certains résultats développés dans ce chapitre seront utiles pour aborder la deuxième partie pratique de la thèse : l'identification d'un gyroscope MEMS. Un résumé de la présentation de ce système réel est présenté dans le chapitre suivant.

Long résumé : Gyroscopie MEMS et objectifs de modélisation

Contents

J.1	Présentation de l'architecture du gyroscopie MEMS	274
J.2	Principe général de mesure de la vitesse angulaire Ω	274
J.3	Dynamiques principales du gyroscopie MEMS	275
J.3.1	Dynamiques des deux masses	275
J.3.2	Instrumentation de type capacités interdigitées du gyroscopie MEMS	276
J.3.3	Dynamiques additionnelles	278
J.4	Correcteur et mesure de la vitesse angulaire	279
J.4.1	Objectifs de contrôle	279
J.4.2	Déduction de la vitesse angulaire à partir des correcteurs	280
J.5	Conclusion sur le résumé du Chapitre 6	281

Dans cette annexe, nous donnons une brève présentation du gyroscopie MEMS afin de pouvoir résumer les résultats du Chapitre 7. Ainsi, cette annexe consiste en un résumé du Chapitre 6. La but de cette annexe est de présenter le gyroscopie MEMS, ses objectifs de contrôle et de modélisation. Nous allons aussi répondre à la première partie de la Question de Recherche 4 en déterminant une structure de modèle plus globale pour l'identification des dynamiques principales du gyroscopie MEMS à partir des équations de la physique.

Notations : nous noterons par t_c le temps continu alors que t correspondra toujours à l'entier qui indexe les signaux à temps discret $x(t)$ (Voir Section H.1). Pour un signal x en temps continu, les notations \dot{x} et \ddot{x} correspondent à sa première et deuxième dérivée par rapport au temps t_c .

J.1 Présentation de l'architecture du gyroscope MEMS

Le gyroscope MEMS (*Micro Electro-Mechanical Structure* en anglais) est un capteur inertiel utilisé pour mesurer une vitesse angulaire (ou vitesse de rotation) d'un objet autour d'un axe donné. Ici, nous allons considérer un mouvement de rotation autour de l'axe z . Ce micro-capteur est constitué d'une structure fixée à l'objet dont nous mesurons la vitesse angulaire. Fixée à cette structure par des micro-poutres à base de silicone, une première masse m_x oscille le long de l'axe x . Une deuxième masse m_y , oscillant le long de l'axe y , est fixée à la masse m_x avec des micro-poutres à base de silicone. Ces micro-poutres agissent mécaniquement comme des ressorts et des amortisseurs. Elles sont conçues de telle sorte que les deux masses forment des oscillateurs légèrement amortis et nous indiquerons par $\omega_{0,x}$ (resp. $\omega_{0,y}$) la pulsation de résonance de m_x (resp. m_y). La vitesse angulaire autour de l'axe z est désignée par Ω et nous souhaitons la mesurer. La Figure 6.1 représente un schéma du gyroscope MEMS.

Chaque masse peut être actionnée par des forces externes : $\vec{f}_x(t_c)$ pour la masse m_x et $\vec{f}_y(t_c)$ pour la masse m_y . L'instrumentation générant ces forces et mesurant les positions des masses (notées x_{mec} pour la masse m_x et y_{mec} pour la masse m_y) est expliquée plus tard. Dans la section suivante, nous donnons l'idée générale sur la façon de déduire la vitesse angulaire

J.2 Principe général de mesure de la vitesse angulaire Ω

Dans ce paragraphe, nous présentons brièvement la manière de déduire la vitesse angulaire Ω . Une présentation plus technique sera donnée plus tard dans la Section J.4.2. Lorsque $\Omega \neq 0$, les deux masses sont soumises à un effet inertiel qui est l'effet de Coriolis. Cet effet est expliqué dans l'Annexe E.1. La force de Coriolis sur la masse m_x est $f_{cor,x} = 2m_y\Omega\dot{y}_{mec}$ et celle sur la masse m_y est $f_{cor,y} = -2m_x\Omega\dot{x}_{mec}$. Il est important de noter que la force de Coriolis est proportionnelle à la vitesse angulaire Ω . Ce fait sera exploité ultérieurement pour son estimation. L'idée de la déduction de la vitesse angulaire Ω est basée sur l'estimation de l'effet de Coriolis sur la masse m_y , c'est-à-dire le terme $f_{cor,y} = -2m_x\Omega\dot{x}_{mec}$. Ce terme explique les rôles des deux masses :

- Masse m_x : pour que le terme $-2m_x\Omega\dot{x}_{mec}$ existe, il est nécessaire d'avoir une vitesse \dot{x}_{mec} non-nulle. Ainsi, le rôle de la masse m_x est la génération de ce terme de Coriolis par son mouvement. Cette masse s'appelle dans la littérature la masse *drive*.
- Masse m_y : nous estimons le terme de Coriolis $-2m_x\Omega\dot{x}_{mec}$ à partir du mouvement de la masse m_y . Cette masse s'appelle dans la littérature la masse *sense*.

Pour être plus précis, l'idée générale de la déduction de la vitesse angulaire est basée sur trois étapes générales. Premièrement, la masse drive m_x est excitée afin d'avoir une

vitesse non-nulle \dot{x}_{mec} . Deuxièmement, le terme de Coriolis $f_{cor,y} = -2m_x\Omega\dot{x}_{mec}$ est estimé à partir du mouvement de la masse sense m_y . Enfin, la troisième étape consiste à déduire Ω à partir de l'estimation de $f_{cor,y} = -2m_x\Omega\dot{x}_{mec}$. Nous donnerons plus tard une explication plus technique sur la manière dont ces trois étapes sont réalisées en pratique (voir Section J.4.2). Afin de répondre à la première partie de la Question de Recherche 4, nous nous sommes concentrés dans le Chapitre 6 à déterminer une structure de modèle plus globale pour la dynamique du gyroscope MEMS que celles qui ont été proposées dans la littérature. Afin de pouvoir donner ce résultat dans ce résumé, nous présentons, dans la section suivante, l'instrumentation et les principales dynamiques du gyroscope MEMS qui sont considérées dans cette thèse.

J.3 Dynamiques principales du gyroscope MEMS

J.3.1 Dynamiques des deux masses

Quand $\Omega \neq 0$, la deuxième loi de Newton sur chaque masse donne¹ :

$$\begin{aligned} m_x\ddot{x}_{mec}(t_c) + d_{xx}\dot{x}_{mec}(t_c) + k_{xx}x_{mec}(t_c) + d_{xy}\dot{y}_{mec}(t_c) + k_{xy}y_{mec}(t_c) \\ = f_x(t_c) + 2m_y\Omega(t_c)\dot{y}_{mec}(t_c) \end{aligned} \quad (J.1)$$

$$\begin{aligned} m_y\ddot{y}_{mec}(t_c) + d_{yy}\dot{y}_{mec}(t_c) + k_{yy}y_{mec}(t_c) + d_{yx}\dot{x}_{mec}(t_c) + k_{yx}x_{mec}(t_c) \\ = f_y(t_c) - 2m_x\Omega(t_c)\dot{x}_{mec}(t_c) \end{aligned} \quad (J.2)$$

où k_{xx} , k_{xy} , k_{yx} , k_{yy} sont des coefficients de raideur et d_{xx} , d_{xy} , d_{yx} , d_{yy} sont des coefficients d'amortissement. Les coefficients de couplage croisé k_{xy} , k_{yx} , d_{xy} et d_{yx} sont dus à des imperfections de fabrication qui provoquent une anisoélasticité (k_{xy} et k_{yx}) et un amortissement non-proportionnel (d_{xy} et d_{yx}). L'effet de Coriolis est également un couplage croisé des deux masses lorsque $\Omega \neq 0$. Un schéma-bloc de ces dynamiques mécaniques est donné dans la Figure 6.2. Lorsque $\Omega = 0$, les dynamiques des deux masses peuvent être considérées comme un système MIMO où le vecteur d'entrée est $(f_x(t_c), f_y(t_c))^T$ et le vecteur de sortie est $(x_{mec}(t_c), y_{mec}(t_c))^T$.

Cependant, dans la pratique, les forces externes f_x et f_y ne sont pas les entrées réelles du MEMS, de même que x_{mec} et y_{mec} ne sont pas les sorties réelles mesurées. En effet, nous devons tenir compte de l'instrumentation des deux masses qui ont été présentées précédemment. Celle considérée dans le gyroscope de cette étude est décrite dans le paragraphe suivant.

¹Puisque les déplacements des masses ont un ordre de grandeur microscopique, nous avons supposé que nous pouvons négliger les forces centrifuges et d'Euler (voir l'Annexe E.1 pour plus de détails). L'effet de la gravité est également négligé en raison de l'effet d'échelle.

J.3.2 Instrumentation de type capacités interdigitées du gyroscope MEMS

Les mouvements des deux masses sont générés et mesurés à l'aide d'une instrumentation électrostatique² utilisant des capacités interdigitées. Les capacités interdigitées sont des actionneurs et des capteurs microscopiques constitués de deux peignes (structures de doigts) électriquement conducteurs qui sont interdigités, c'est-à-dire que les dents d'un peigne sont opposées aux trous de l'autre peigne. Un peigne est fixé à une structure et l'autre est en mouvement. Une capacité interdigitée est illustrée dans la Figure 6.3. La principale propriété avantageuse de cette architecture est que la capacité générée entre les deux peignes est proportionnelle à la distance qui les sépare (distance d dans la Figure 6.3). Cependant, cette propriété n'est valable que pour de petites distances. C'est pourquoi les capacités interdigitées sont souvent utilisées dans des applications à l'échelle micro et nanoscopique comme le gyroscope MEMS.

Dans le gyroscope MEMS, il y a quatre capacités interdigitées au total, deux par masse. Pour chaque masse, une capacité interdigitée est utilisée pour générer la force de mise en vibration et l'autre pour la mesure du déplacement. Pour le mode drive, pour chaque capacité interdigitée, le peigne statique est fixé sur la structure et le peigne mobile est fixé sur la masse drive. Pour le mode sense, pour chaque capacité interdigitée, le peigne statique est fixé sur la masse drive et le peigne mobile est fixé sur la masse sense. Les capacités interdigitées sont placées comme illustré dans la Figure 6.4.

Pour la mise en vibration, nous excitons le gyroscope MEMS avec une entrée à temps discret $u_x(t)$ pour le mode drive (resp. $u_y(t)$ pour le mode sense) qui est convertie en une tension $V_x(t_c)$ (resp. $V_y(t_c)$) par un convertisseur numérique-analogique qui est un bloqueur d'ordre zéro. La tension V_x (resp. V_y) est appliquée entre les deux peignes de la capacités interdigitée dédié à l'excitation de la masse drive (resp. de la masse sense) (voir Figure 6.3 avec la tension V). Il en résulte des forces électrostatiques f_x et f_y proportionnelles au carré de la tension correspondante, c'est-à-dire

$$\begin{aligned}f_x(t_c) &= \alpha_x V_x^2(t_c) \\f_y(t_c) &= \alpha_y V_y^2(t_c)\end{aligned}$$

où α_x et α_y sont des constantes³ qui dépendent des propriétés électrostatiques et géométriques des doigts des peignes. Nous pouvons observer qu'il y a une non-linéarité carrée à l'entrée.

Les signaux en temps discret u_x et u_y constituent les entrées du gyroscope MEMS et produisent les tensions d'excitation V_x et V_y à travers le bloqueur d'ordre zéro. Ainsi,

²Il existe d'autres moyens d'actionner et de mesurer les déplacements des masses. Dans [YAN98] une instrumentation piézoélectrique est considérée.

³Tant que les déplacements des masses sont maintenus faibles, ces constantes ne dépendent pas du temps et des déplacements des masses. Cet avantage justifie que ce type d'actionneur soit souvent utilisé pour les applications de micro-échelle.

les tensions d'entrée sont données par

$$\begin{aligned} V_x(t_c) &= u_x(t) \quad \text{pour } tT_s \leq t_c < (t+1)T_s \\ V_y(t_c) &= u_y(t) \quad \text{pour } tT_s \leq t_c < (t+1)T_s \end{aligned}$$

où T_s est la période d'échantillonnage. La pulsation d'échantillonnage $\omega_s = 2\pi/T_s$ est égale à $2\pi \times 62500$ rad/s dans le gyroscope considéré dans cette étude.

Avec le bloqueur d'ordre zéro, les tensions V_x et V_y ont une forme en escalier. Ainsi, du fait que la non-linéarité carrée est statique, nous avons que

$$f_x(t_c) = \alpha_x u_x^2(t) \quad \text{pour } tT_s \leq t_c < (t+1)T_s \quad (\text{J.3})$$

$$f_y(t_c) = \alpha_y u_y^2(t) \quad \text{pour } tT_s \leq t_c < (t+1)T_s \quad (\text{J.4})$$

Pour la mesure des deux déplacements $x_{mec}(t_c)$ et $y_{mec}(t_c)$, nous utilisons les capacités interdigitées dédiées à la mesure comme illustrés dans la Figure 6.4 et nous exploitons le fait que la capacité entre les deux peignes est proportionnelle à la distance séparant ces peignes (distance d dans la figure 6.3). Si nous pouvons mesurer cette capacité, alors la distance d sera connue à une constante près. C'est cette idée qui est exploitée dans le gyroscope MEMS pour mesurer les deux déplacements $x_{mec}(t_c)$ et $y_{mec}(t_c)$.

À cette fin, chaque capacité interdigitée dédiée à la mesure (voir Figure 6.4) est placée dans un circuit électronique⁴ qui convertit proportionnellement la capacité en une tension. Cette tension est également proportionnelle à la capacité et donc proportionnelle au déplacement de la masse correspondante. Ensuite, ces tensions sont amplifiées, filtrées à travers un filtre anti-repliement de premier ordre de pulsation de coupure $2\pi \times 15000$ rad/s et converties dans le domaine du temps discret avec un convertisseur analogique-numérique. Par conséquent, les sorties du convertisseur analogique-numérique sont deux signaux à temps discret qui seront désignés par $\tilde{x}_{mec}(t)$ et $\tilde{y}_{mec}(t)$ et qui sont supposés être proportionnels à $x_{mec}(t_c = tT_s)$ et $y_{mec}(t_c = tT_s)$ respectivement, c'est-à-dire

$$\tilde{x}_{mec}(t) = \beta_x x_{mec}(t) \quad (\text{J.5})$$

$$\tilde{y}_{mec}(t) = \beta_y y_{mec}(t) \quad (\text{J.6})$$

où β_x et β_y sont des constantes qui sont fonction des propriétés électrostatiques et géométriques des doigts des peignes, des propriétés du circuit de conversion électronique et des amplificateurs. Dans la Figure 6.5, nous donnons un schéma des dynamiques mécaniques instrumentées du gyroscope MEMS.

Par conséquent, les sorties des dynamiques mécaniques instrumentées sont les signaux en temps discret $\tilde{x}_{mec}(t)$ et $\tilde{y}_{mec}(t)$. Puisque nous avons vu que les entrées de ces dynamiques sont u_x^2 et u_y^2 , alors les dynamiques mécaniques peuvent être représentées par

⁴Le gyroscope MEMS est embarqué sur une carte électronique qui permet de gérer les conversions analogique-numérique, d'implanter un correcteur, etc.

une matrice $\mathbf{G}_0(z)$ de fonctions de transfert en temps discret MIMO de dimension 2×2 lorsque $\Omega = 0$, c'est-à-dire

$$\begin{pmatrix} \tilde{x}_{mec}(t) \\ \tilde{y}_{mec}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} G_{0,xx}(z) & G_{0,xy}(z) \\ G_{0,yx}(z) & G_{0,yy}(z) \end{pmatrix}}_{\mathbf{G}_0(z)} \begin{pmatrix} u_x^2(t) \\ u_y^2(t) \end{pmatrix}$$

Les fonctions de transfert dans $\mathbf{G}_0(z)$ dépendent des paramètres de (J.1) et (J.2) mais aussi des constantes α_x , α_y , β_x et β_y . Chaque fonction de transfert $G_{0,xx}(z)$ et $G_{0,yy}(z)$ a un comportement résonant légèrement amorti à la pulsation de résonance $\omega_{0,x}$ pour $G_{0,xx}(z)$ et à la pulsation de résonance $\omega_{0,y}$ pour $G_{0,yy}(z)$. Les fonctions de transfert $G_{0,xy}(z)$ et $G_{0,yx}(z)$ ont deux résonances légèrement amorties aux deux pulsations de résonance $\omega_{0,x}$ et $\omega_{0,y}$.

Toutefois, l'utilisation de l'instrumentation capacitive crée une dynamique supplémentaire et les sorties réelles mesurées ne sont donc pas égales à $\tilde{x}_{mec}(t)$ et $\tilde{y}_{mec}(t)$. Ces dynamiques supplémentaires sont présentées dans le paragraphe suivant.

J.3.3 Dynamiques additionnelles

Dans ce paragraphe, nous présenterons quelques dynamiques supplémentaires impliquées dans le gyroscope MEMS en raison de l'utilisation de l'instrumentation capacitive. Nous verrons que les sorties mesurées des MEMS sont deux signaux à temps discret \tilde{x} et \tilde{y} qui sont chacun la somme de trois signaux :

$$\begin{aligned} \tilde{x} &= \tilde{x}_{mec} + \tilde{x}_{par} + v_x \\ \tilde{y} &= \tilde{y}_{mec} + \tilde{y}_{par} + v_y \end{aligned}$$

où \tilde{x}_{mec} et \tilde{y}_{mec} sont les sorties mécaniques données respectivement par (6.5) et par (6.6). Dans ce paragraphe, nous expliquons ce qui se cache derrière les signaux \tilde{x}_{par} , v_x , \tilde{y}_{par} et v_y .

Le phénomène derrière les signaux \tilde{x}_{par} et \tilde{y}_{par} est un couplage capacitif parasite qui se crée entre le vecteur de tensions $(V_x, V_y)^T$ et le vecteur de sortie $(\tilde{x}, \tilde{y})^T$. Ce couplage est le résultat des interactions des champs électriques entre les capacités interdigitées dédiées à l'excitation et celles dédiées à la mesure. Dans la littérature [AS05], ce phénomène est modélisé par un pont linéaire résistance-capacité en série entre les vecteurs $(V_x, V_y)^T$ et $(\tilde{x}, \tilde{y})^T$ et ce pont se traduit par un effet filtrage passe-haut. Du point de vue systémique, les dynamiques de cet effet parasite peuvent être modélisées par

$$\begin{pmatrix} \tilde{x}_{par}(t) \\ \tilde{y}_{par}(t) \end{pmatrix} = \underbrace{\begin{pmatrix} Q_{0,xx}(z) & Q_{0,xy}(z) \\ Q_{0,yx}(z) & Q_{0,yy}(z) \end{pmatrix}}_{\mathbf{Q}_0(z)} \begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix}$$

où $\mathbf{Q}_0(z)$ est une matrice de fonctions de transfert de dimension 2×2 .

La troisième et dernière dynamique considérée pour la modélisation est le bruit de mesure. Dans cette thèse, ce bruit est modélisé comme deux perturbations stochastiques v_x et v_y qui s'ajoutent respectivement aux signaux $\tilde{x}_{mec} + \tilde{x}_{par}$ et $\tilde{y}_{mec} + \tilde{y}_{par}$. Comme nous allons le voir plus tard, il est important d'estimer la matrice densité spectrale de puissance du vecteur de bruits $\mathbf{v} = (v_x, v_y)^T$. Pour cela, nous allons supposer que les deux bruits de mesure v_x et v_y sont générés par

$$\begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} = \underbrace{\begin{pmatrix} H_{0,x}(z) & 0 \\ 0 & H_{0,y}(z) \end{pmatrix}}_{\mathbf{H}_0(z)} \begin{pmatrix} e_x(t) \\ e_y(t) \end{pmatrix}$$

où $(e_x, e_y)^T$ est un vecteur de bruits blancs.

D'autres phénomènes existent tels que la dépendance des dynamiques mécaniques avec la température, l'effet Spring-softening [Sau08], saturation des amplificateurs, etc. Cependant, il est possible de maîtriser leurs effets. Par exemple, pour l'effet Spring-Softening, la composante DC des tensions V_x et V_y doit rester constante pour maîtriser cet effet [Sau08].

Ainsi, lorsque $\Omega = 0$ et à température fixée⁵, les trois types de dynamiques différentes peuvent se combiner dans un modèle plus complet par rapport à ce qui a été proposé dans la littérature, répondant à la première partie de la Question de Recherche 4 :

$$\begin{aligned} \begin{pmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{pmatrix} &= \begin{pmatrix} Q_{0,xx}(z) & Q_{0,xy}(z) \\ Q_{0,yx}(z) & Q_{0,yy}(z) \end{pmatrix} \begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix} + \begin{pmatrix} G_{0,xx}(z) & G_{0,xy}(z) \\ G_{0,yx}(z) & G_{0,yy}(z) \end{pmatrix} \begin{pmatrix} u_x^2(t) \\ u_y^2(t) \end{pmatrix} + \mathbf{v}(t) \\ \mathbf{v}(t) &= \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} = \begin{pmatrix} H_{0,x}(z) & 0 \\ 0 & H_{0,y}(z) \end{pmatrix} \begin{pmatrix} e_x(t) \\ e_y(t) \end{pmatrix} \end{aligned} \quad (\text{J.7})$$

Dans le prochain paragraphe, nous montrons en quoi il est important d'avoir une estimée précise des fonctions de transfert de (J.7).

J.4 Correcteur et mesure de la vitesse angulaire

J.4.1 Objectifs de contrôle

Pour déduire la vitesse de rotation Ω , le capteur est souvent placé en boucle fermée comme illustré dans la Figure 6.9. L'objectif principal de ce correcteur est de faire en sorte que le vecteur de sortie mécanique $(\tilde{x}_{mec}, \tilde{y}_{mec})^T$ soit asservi au vecteur de consignes $(x_{ref}, 0)^T$ où x_{ref} est un signal sinusoïdal à une pulsation notée ω_{ref} . Puisque $\tilde{x}_{mec} = \beta_x x_{mec}$ et $\tilde{y}_{mec} = \beta_y y_{mec}$, alors nous aurons $x_{mec} = x_{ref}/\beta_x$ et $y_{mec} = 0$. Nous verrons dans le paragraphe suivant comment ces objectifs peuvent nous permettre de déduire la vitesse angulaire Ω .

⁵Pour cela, le gyroscope MEMS et sa carte électronique sont placés dans une chambre thermique dont la température intérieure est régulée à une consigne de température constante.

Afin de garantir ces objectifs, l'architecture du correcteur qui a été proposée par Fabrício SAGGIN est donnée dans la Figure 6.10. Ce correcteur a deux degrés de liberté qui sont les deux blocs "Tracking Controller $\mathbf{K}(z)$ " et "Parasitic Compensation $\mathbf{K}_{comp}(z)$ ". Les deux blocs seront synthétisés indépendamment.

Le bloc "Parasitic Compensation $\mathbf{K}_{comp}(z)$ " a pour but de compenser les signaux parasites \tilde{x}_{par} et \tilde{y}_{par} des sorties mesurées \tilde{x} et \tilde{y} . Les signaux compensés sont notés par \tilde{x}_c pour la sortie drive et \tilde{y}_c pour la sortie sense. En désignant par $\mathbf{K}_{comp}(z)$ la matrice de fonction de transfert de dimension 2×2 utilisée pour cette compensation, les signaux compensés \tilde{x}_c et \tilde{y}_c sont obtenus de la façon suivante

$$\begin{pmatrix} \tilde{x}_c(t) \\ \tilde{y}_c(t) \end{pmatrix} = \begin{pmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{pmatrix} + \mathbf{K}_{comp}(z) \begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix}$$

À partir de (J.7), l'expression optimale pour $\mathbf{K}_{comp}(z)$ afin de garantir $\tilde{x}_c(t) = \tilde{x}_{mec} + v_x$ et $\tilde{y}_c(t) = \tilde{y}_{mec} + v_y$ est $\mathbf{K}_{comp}(z) = -\mathbf{Q}_0(z)$. Par conséquent, si nous avons une estimation précise $\hat{\mathbf{Q}}(z)$ de $\mathbf{Q}_0(z)$, le couplage capacitif parasite sera bien compensé en considérant $\mathbf{K}_{comp}(z) = \hat{\mathbf{Q}}(z)$.

Le correcteur d'asservissement $\mathbf{K}(z)$ doit assurer un rejet approprié des bruits v_x et v_y et le suivi de la référence $(x_{ref}, 0)^T$. Avec le correcteur d'asservissement, les commandes u_x et u_y sont données par

$$\begin{pmatrix} u_x(t) \\ u_y(t) \end{pmatrix} = \mathbf{K}(z) \begin{pmatrix} x_{ref}(t) - \tilde{x}_c(t) \\ -\tilde{y}_c(t) \end{pmatrix}$$

Afin d'augmenter l'amplitude de la vibration de la masse drive m_x avec un effort de contrôle minimal u_x , la pulsation ω_{ref} du signal de référence sinusoïdal x_{ref} doit être égale à la pulsation de résonance $\omega_{0,x}$.

Si $\mathbf{K}(z)$ et $\mathbf{K}_{comp}(z)$ atteignent leurs objectifs respectifs, nous aurons la propriété suivante : $x_{mec} \approx \tilde{x}_{ref}$ et $\tilde{y}_{mec} \approx 0$. Dans le paragraphe suivant, nous expliquons comment déduire la vitesse angulaire Ω lorsque le gyroscope MEMS fonctionne comme décrit ci-dessus.

J.4.2 Dédution de la vitesse angulaire à partir des correcteurs

Nous supposons que les objectifs de $\mathbf{K}(z)$ et $\mathbf{K}_{comp}(z)$ sont atteints et donc que $\tilde{x}_{mec} = x_{ref}$ et $\tilde{y}_{mec} = 0$. Étant donné que la dynamique mécanique est la seule dynamique restante, considérons les équations (J.1) et (J.2). Dans ces équations, x_{mec} aura une forme sinusoïdale et sera proportionnelle à x_{ref} et y_{mec} sera égal à zéro. Par conséquent, la force f_y générée par la capacité interdiguée correspondante est donnée par :

$$f_y = 2m_x\Omega\dot{x}_{mec} + d_{yx}\dot{x}_{mec} + k_{yx}x_{mec} \quad (\text{J.8})$$

Comme $f_y = \alpha_y u_y^2$ et $\tilde{x}_{mec} = \beta_x x_{mec} = x_{ref}$, (J.8) devient

$$u_y^2 = \underbrace{(a\Omega + b)}_{\hat{\Omega}} \dot{x}_{ref} + \frac{k_{yx}}{\alpha_y \beta_x} x_{ref}$$

où $a = 2m_x/(\alpha_y \beta_x)$ et $b = d_{yx}/(\alpha_y \beta_x)$.

En raison de la forme sinusoïdale de x_{ref} , si nous démodulons⁶ u_y^2 avec \dot{x}_{ref} , nous obtenons $\hat{\Omega}$. Ensuite, si nous connaissons précisément les valeurs de d_{yx} , m_x , α_y et β_x , nous pouvons déduire Ω . Dans le Chapitre 7 et comme il le sera résumé dans l'Annexe K, nous pouvons estimer ces variables à une constante donnée. Notons qu'il est également possible d'estimer directement les deux constantes a et b en effectuant une calibration du capteur avec différentes valeurs de vitesses angulaires connues Ω .

Il est à noter que le principe de la déduction de la vitesse angulaire par le gyroscope MEMS est similaire au principe de la transmission par radio-fréquence (RF). En effet, dans la transmission RF, un signal $q(t_c)$ que nous voulons transmettre est d'abord modulé par un signal porteur sinusoïdal $q_{car}(t_c)$ et ensuite envoyé par un canal. Ensuite, ce signal modulé $q(t_c)q_{car}(t_c)$ est reçu par une antenne et il est démodulé pour récupérer le signal initial $q(t_c)$. Pour le gyroscope MEMS, nous avons que $q(t_c) = \hat{\Omega}(t_c)$ et $q_{car}(t_c) = \dot{x}_{ref}(t_c)$. La transmission est due à l'effet de Coriolis et l'antenne est la masse sense. La Figure 6.11 illustre cette analogie.

J.5 Conclusion sur le résumé du Chapitre 6

Dans cette annexe, nous avons présenté le gyroscope MEMS et les défis de contrôle autour de ce capteur. Nous avons aussi déterminé un modèle plus global pour les dynamiques principales du gyroscope MEMS, répondant à la première partie de la Question de Recherche 4. Afin de synthétiser un correcteur pour avoir une estimée précise de la vitesse angulaire, il nous faut déterminer des estimées précises des fonctions de transfert de (J.7). Pour cela, nous avons proposé une méthode d'identification des phénomènes liés à ces fonctions de transfert dans le Chapitre 7 et dont le résumé se trouve dans la prochaine annexe.

⁶Dans [Sau08], la technique de démodulation utilisée est la démodulation synchrone. C'est la même technique que celle considérée dans le gyroscope MEMS de cette étude. Voir [Sau08] pour plus de détails sur la démodulation synchrone.

Long résumé : Résumé du Chapitre 7

Contents

K.1	Méthode d'identification	284
K.1.1	Simplification	284
K.1.2	Structures de modèle	285
K.1.3	Jeux de données	285
K.1.4	Estimateurs	286
K.2	Limitations de l'expérimentation et informations a priori	286
K.3	Initialisation : approche SISO-par-SISO	288
K.3.1	Résumé de la méthode	288
K.3.2	Résumé des résultats obtenus avec l'approche SISO-par-SISO et de leur interprétation	289
K.4	Identifications MISO et résultats	290
K.4.1	Choix des entrées pour les identifications MISO	290
K.4.2	Étude de l'informativité des données	291
K.4.3	Résultats des identifications MISO	292
K.5	Réduction d'ordre du modèle par identification boîte grise	293
K.6	Conclusion sur le résumé des résultats du Chapitre 7	296

Dans cette annexe, nous résumons la méthode d'identification des fonctions de transfert de (J.7), c'est à dire, nous répondons à la deuxième partie de la Question de Recherche 4. Pour les résultats associés au correcteur synthétisé sur les modèles identifiés de ces fonctions de transfert, le lecteur est invité à lire le Chapitre 7.

K.1 Méthode d'identification

K.1.1 Simplification

Le système représenté par (J.7) est un système MIMO non-linéaire de type Hammerstein avec une non-linéarité carrée (statique) et avec deux entrées u_x et u_y et deux sorties \tilde{x} et \tilde{y} . Toutefois, les dynamiques liées à certaines fonctions de transfert telles que $G_{0,xx}(z)$ et $G_{0,yy}(z)$ peuvent être dominantes par rapport à celles des autres fonctions de transfert telles que $G_{0,xy}(z)$ et $G_{0,yx}(z)$ et donc des modèles inexacts peuvent être obtenus pour $G_{0,xy}(z)$ et $G_{0,yx}(z)$ si nous effectuons une identification MIMO en excitant le gyroscope MEMS avec des entrées non nulles u_x et u_y en même temps. Nous allons donc simplifier l'identification comme suit :

- nous excitons seulement u_x en fixant u_y à 0. Dans ce cas, (J.7) devient

$$\begin{pmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{pmatrix} = \begin{pmatrix} Q_{0,xx}(z) \\ Q_{0,yx}(z) \end{pmatrix} u_x(t) + \begin{pmatrix} G_{0,xx}(z) \\ G_{0,yx}(z) \end{pmatrix} u_x^2(t) + \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} \quad (\text{K.1})$$

- nous excitons seulement u_y en fixant u_x à 0. Dans ce cas, (J.7) devient

$$\begin{pmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{pmatrix} = \begin{pmatrix} Q_{0,xy}(z) \\ Q_{0,yy}(z) \end{pmatrix} u_y(t) + \begin{pmatrix} G_{0,xy}(z) \\ G_{0,yy}(z) \end{pmatrix} u_y^2(t) + \begin{pmatrix} v_x(t) \\ v_y(t) \end{pmatrix} \quad (\text{K.2})$$

Nous observons que (K.1) est composé de deux systèmes non-linéaires de type Hammerstein parallèle \mathcal{S}_{xx} et \mathcal{S}_{yx} une-entrée une-sortie donnés par

$$\mathcal{S}_{xx} : \quad \tilde{x}(t) = Q_{0,xx}(z)u_x(t) + G_{0,xx}(z)u_x^2(t) + H_{0,x}(z)e_x(t) \quad (\text{K.3})$$

$$\mathcal{S}_{yx} : \quad \tilde{y}(t) = Q_{0,yx}(z)u_x(t) + G_{0,yx}(z)u_x^2(t) + H_{0,y}(z)e_y(t) \quad (\text{K.4})$$

Ainsi, lorsque u_x est excité seulement, nous devons identifier \mathcal{S}_{xx} et \mathcal{S}_{yx} .

La même remarque peut être faite pour (K.2), composé des systèmes non-linéaires de type Hammerstein parallèle \mathcal{S}_{xy} et \mathcal{S}_{yy} une-entrée une-sortie donnés par

$$\mathcal{S}_{xy} : \quad \tilde{x}(t) = Q_{0,xy}(z)u_y(t) + G_{0,xy}(z)u_y^2(t) + H_{0,x}(z)e_x(t)$$

$$\mathcal{S}_{yy} : \quad \tilde{y}(t) = Q_{0,yy}(z)u_y(t) + G_{0,yy}(z)u_y^2(t) + H_{0,y}(z)e_y(t)$$

Ces 4 systèmes correspondent aux 4 sous-systèmes mentionnés dans l'introduction (Annexe G). Par ailleurs, observons que ces sous-systèmes appartiennent à la classe des systèmes non-linéaires étudiés dans le Chapitre 5 avec $n = 2$ (dont le résumé est en Annexe I.3).

Ainsi, afin de traiter l'identification de ces 4 sous-systèmes, nous allons considérer la même astuce utilisée en Annexe I.3, c'est à dire nous allons les réécrire en systèmes

MISO LTI comme ci-dessous

$$\mathcal{S}_{xx} : \quad \tilde{x}(t) = (Q_{0,xx}(z) \quad G_{0,xx}(z)) \mathbf{u}_x(t) + H_{0,x}(z)e_x(t)$$

$$\mathcal{S}_{yx} : \quad \tilde{y}(t) = (Q_{0,yx}(z) \quad G_{0,yx}(z)) \mathbf{u}_x(t) + H_{0,y}(z)e_y(t)$$

$$\mathcal{S}_{xy} : \quad \tilde{x}(t) = (Q_{0,xy}(z) \quad G_{0,xy}(z)) \mathbf{u}_y(t) + H_{0,x}(z)e_x(t)$$

$$\mathcal{S}_{yy} : \quad \tilde{y}(t) = (Q_{0,yy}(z) \quad G_{0,yy}(z)) \mathbf{u}_y(t) + H_{0,y}(z)e_y(t)$$

où $\mathbf{u}_x(t) = (u_x, u_x^2)^T$ et $\mathbf{u}_y(t) = (u_y, u_y^2)^T$.

K.1.2 Structures de modèle

Afin d'identifier les fonctions de transfert \mathcal{S}_{xx} et \mathcal{S}_{yx} , nous allons considérer ces deux structures de modèle MISO de type Box-Jenkins

$$\mathcal{M}_{xx} = \left\{ (Q_{xx}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}}), G_{xx}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}}), H_x(z, \boldsymbol{\eta}_x)) \mid \boldsymbol{\theta}_{xx} = (\tilde{\boldsymbol{\theta}}_{Q_{xx}}^T, \tilde{\boldsymbol{\theta}}_{G_{xx}}^T, \boldsymbol{\eta}_x^T)^T \in \mathcal{D}_{\boldsymbol{\theta}_{xx}} \right\}$$

$$\mathcal{M}_{yx} = \left\{ (Q_{yx}(z, \tilde{\boldsymbol{\theta}}_{Q_{yx}}), G_{yx}(z, \tilde{\boldsymbol{\theta}}_{G_{yx}}), H_y(z, \boldsymbol{\eta}_y)) \mid \boldsymbol{\theta}_{yx} = (\tilde{\boldsymbol{\theta}}_{Q_{yx}}^T, \tilde{\boldsymbol{\theta}}_{G_{yx}}^T, \boldsymbol{\eta}_y^T)^T \in \mathcal{D}_{\boldsymbol{\theta}_{yx}} \right\}$$

où les ensembles $\mathcal{D}_{\boldsymbol{\theta}_{xx}}$ et $\mathcal{D}_{\boldsymbol{\theta}_{yx}}$ sont définis de façon similaire comme dans les annexes précédentes.

Puisque \mathcal{M}_{xx} est une structure de modèle BJ, les fonctions de transfert $Q_{xx}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}})$, $G_{xx}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}})$ et $H_x(z, \boldsymbol{\eta}_x)$ sont données par

$$Q_{xx}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}}) = z^{-\rho_{Q_{xx}}} \frac{\mathring{B}_{Q_{xx}}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}})}{F_{Q_{xx}}(z, \tilde{\boldsymbol{\theta}}_{Q_{xx}})}$$

$$G_{xx}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}}) = z^{-\rho_{G_{xx}}} \frac{\mathring{B}_{G_{xx}}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}})}{F_{G_{xx}}(z, \tilde{\boldsymbol{\theta}}_{G_{xx}})}$$

$$H_x(z, \boldsymbol{\eta}_x) = \frac{C_x(z, \boldsymbol{\eta}_x)}{D_x(z, \boldsymbol{\eta}_x)}$$

où $\mathring{B}_{Q_{xx}}$ et $\mathring{B}_{G_{xx}}$ sont des polynômes à retard nul et où $F_{Q_{xx}}$, $F_{G_{xx}}$, C_x et D_x sont des polynômes moniques. Les entiers $\rho_{Q_{xx}}$ et $\rho_{G_{xx}}$ sont respectivement les retards de Q_{xx} et G_{xx} . Nous définissons de manière similaire les polynômes $\mathring{B}_{Q_{yx}}$, $\mathring{B}_{G_{yx}}$, $F_{Q_{yx}}$, $F_{G_{yx}}$, C_y et D_y pour les fonctions de transfert Q_{yx} , G_{yx} et H_y dans la structure de modèle \mathcal{M}_{yx} . Nous définissons de façon similaire ces notations pour les structures de modèle \mathcal{M}_{xy} et \mathcal{M}_{yy} .

K.1.3 Jeux de données

En appliquant une excitation quasi-stationnaire u_x au gyroscope MEMS de $t = 1$ à $t = N$ tout en fixant u_y à 0, nous collectons simultanément les jeux de données suivants

$$\mathcal{Z}_{xx}^N = \{(\tilde{x}(t), \mathbf{u}_x^T(t))^T \mid t = 1, \dots, N\}$$

$$\mathcal{Z}_{yx}^N = \{(\tilde{y}(t), \mathbf{u}_x^T(t))^T \mid t = 1, \dots, N\}$$

Nous définissons de façon similaire les jeux de données \mathcal{Z}_{xy}^N et \mathcal{Z}_{yy}^N en appliquant une excitation quasi-stationnaire u_y au gyroscope MEMS de $t = 1$ à $t = N$ tout en fixant u_x à 0.

A partir de ces jeux de données et des structures de modèle précédemment introduites, pour chaque $(Q_{xx}(z, \tilde{\theta}_{Q_{xx}}), G_{xx}(z, \tilde{\theta}_{G_{xx}}), H_x(z, \boldsymbol{\eta}_x)) \in \mathcal{M}_{xx}$ et chaque $(Q_{yx}(z, \tilde{\theta}_{Q_{yx}}), G_{yx}(z, \tilde{\theta}_{G_{yx}}), H_y(z, \boldsymbol{\eta}_y)) \in \mathcal{M}_{yx}$, nous construisons les prédicteurs suivant

$$\begin{aligned}\hat{x}(t, \boldsymbol{\theta}_{xx}) &= H_x^{-1}(z, \boldsymbol{\eta}_x)(Q_{xx}(z, \tilde{\theta}_{Q_{xx}}), G_{xx}(z, \tilde{\theta}_{G_{xx}})\mathbf{u}_x(t) + (1 - H_x^{-1}(z, \boldsymbol{\eta}_x))\tilde{x}(t)) \\ \hat{y}(t, \boldsymbol{\theta}_{yx}) &= H_y^{-1}(z, \boldsymbol{\eta}_y)(Q_{yx}(z, \tilde{\theta}_{Q_{yx}}), G_{yx}(z, \tilde{\theta}_{G_{yx}})\mathbf{u}_x(t) + (1 - H_y^{-1}(z, \boldsymbol{\eta}_y))\tilde{y}(t))\end{aligned}$$

K.1.4 Estimateurs

Nous pouvons alors définir les vecteurs de paramètres optimaux $\hat{\boldsymbol{\theta}}_{N,xx}$ et $\hat{\boldsymbol{\theta}}_{N,yx}$ donnés par

$$\hat{\boldsymbol{\theta}}_{N,xx} = \arg \min_{\boldsymbol{\theta}_{xx} \in \mathcal{D}_{\boldsymbol{\theta}_{xx}}} \frac{1}{N} \sum_{t=1}^N (\tilde{x}(t) - \hat{x}(t, \boldsymbol{\theta}_{xx}))^2 \quad (\text{K.5})$$

$$\hat{\boldsymbol{\theta}}_{N,yx} = \arg \min_{\boldsymbol{\theta}_{yx} \in \mathcal{D}_{\boldsymbol{\theta}_{yx}}} \frac{1}{N} \sum_{t=1}^N (\tilde{y}(t) - \hat{y}(t, \boldsymbol{\theta}_{yx}))^2 \quad (\text{K.6})$$

Nous définissons de la même manière les vecteurs de paramètres optimaux $\hat{\boldsymbol{\theta}}_{N,xy}$ et $\hat{\boldsymbol{\theta}}_{N,yy}$ à partir des jeux de données \mathcal{Z}_{xy}^N et \mathcal{Z}_{yy}^N et des structures du modèle \mathcal{M}_{xy} et \mathcal{M}_{yy} . Afin de garantir un modèle précis du gyroscope MEMS, nous nous demandons comment concevoir l'excitation u_x pour garantir des estimations précises $\hat{\boldsymbol{\theta}}_{N,xx}$ et $\hat{\boldsymbol{\theta}}_{N,yx}$. Le même problème se pose pour \mathcal{M}_{xy} et \mathcal{M}_{yy} avec la conception de u_y .

Dans la prochaine section, nous donnons les limitations de l'expérimentation pour le choix des entrées u_x et u_y et nous discutons rapidement du choix de ces entrées pour les identifications MISO de \mathcal{S}_{xx} , \mathcal{S}_{xy} , \mathcal{S}_{yx} et \mathcal{S}_{yy} .

K.2 Limitations de l'expérimentation et informations a priori

Avant tout, nous devons donner les limitations de l'expérimentation imposées par la carte électronique. Pour le gyroscope MEMS considéré dans cette étude, la tension d'entrée u_x doit être positive, ne peut pas dépasser 2V et doit avoir une composante continue (DC) u_{DC} constante¹ et, dans notre cas, égale à 1V, soit

$$u_x(t) = u_{DC} + \bar{u}_x(t)$$

¹La tension continue doit être maintenue constante pour éviter l'effet Spring-softening qui fait varier les deux pulsations de résonances lorsque cette constante est modifiée [Sau08].

où $-1V \leq \bar{u}_x(t) \leq 1V$. Ceci est aussi vrai pour u_y . Ainsi, pour la synthèse des excitations, nous devons nous focaliser sur les signaux \bar{u}_x et \bar{u}_y .

Afin d'obtenir des modèles identifiés avec des incertitudes relativement petites, il est important de maximiser le rapport signal sur bruit. Par conséquent, nous devrions concevoir \bar{u}_x de manière à ce qu'il ait une puissance maximale et qu'il garantisse $-1V \leq \bar{u}_x(t) \leq 1V$. Cependant, il ne suffit pas de maximiser la puissance d'entrée en fonction des limitations d'entrée pour que ce signal soit approprié pour l'identification. En effet, le choix de \bar{u}_x doit également se baser sur une optimisation des densités spectrales de puissance de u_x et u_x^2 pour les deux sorties \tilde{x} et \tilde{y} . Pour cela, nous disposons de quelques informations *a priori*

- les dynamiques de $G_{0,xx}$ et $G_{0,yy}$ sont caractérisées par une résonance à la pulsation de résonance $\omega_{0,x}$ pour $G_{0,xx}$ et $\omega_{0,y}$ pour $G_{0,yy}$. Les fonctions de transfert croisées $G_{0,xy}$ et $G_{0,yx}$ ont deux résonances chacune, aux deux pulsations de résonance $\omega_{0,x}$ et $\omega_{0,y}$.
- pour une température de 20°C, $\omega_{0,x} \approx 2\pi \times 11836.8$ rad/s et $\omega_{0,y} \approx 2\pi \times 11843.4$ rad/s.
- pour une température de 20°C, le facteur de qualité de la résonance du mode drive a un ordre de grandeur de 10^5 et celui du mode sense a un ordre de grandeur de 10^4 .
- les fonctions de transfert de \mathbf{Q}_0 sont des filtres passe-haut.

Ainsi la fonction de transfert $G_{0,yx}$ doit être excitée aux pulsations de résonance $\omega_{0,x}$ et $\omega_{0,y}$ tandis que $G_{0,xx}$ doit être excitée à $\omega_{0,x}$. Par conséquent, pour identifier $G_{0,xx}$ et $G_{0,yx}$ avec une expérience unique, le spectre de puissance de l'entrée des deux fonctions de transfert (qui est u_x^2) doit être non-nul aux pulsations proches de $\omega_{0,x}$ et $\omega_{0,y}$. Pour les fonctions de transfert $Q_{0,xx}$ et $Q_{0,yx}$, les dynamiques sont distribuées sur une gamme de pulsations plus large que $G_{0,xx}$ et $G_{0,yx}$. Par conséquent, l'entrée des deux fonctions de transfert $Q_{0,xx}$ et $Q_{0,yx}$ (qui est u_x) devrait les exciter sur une large gamme de pulsations, de préférence en haute fréquence puisque les deux fonctions de transfert sont des filtres passe-haut. Nous devons vérifier si ce choix d'entrée donne l'informativité des données pour les quatre identifications MISO.

Puisque $\hat{\theta}_{N,xx}$, $\hat{\theta}_{N,yx}$, $\hat{\theta}_{N,xy}$ et $\hat{\theta}_{N,yy}$ sont des solutions à des problèmes complexes d'optimisation non-linéaire, nous devons trouver d'abord des estimées initiales relativement précises de chaque fonction de transfert dans \mathcal{S}_{xx} , \mathcal{S}_{yx} , \mathcal{S}_{xy} et \mathcal{S}_{yy} . Pour cela, nous avons développée une approche appelée SISO-par-SISO. Dans cette approche, comme par exemple pour (K.5) et (K.6), nous utiliserons le fait que la non-linéarité de l'entrée est connue.

K.3 Initialisation : approche SISO-par-SISO

K.3.1 Résumé de la méthode

L'idée de l'estimation initiale des fonctions de transfert des sous-systèmes \mathcal{S}_{xx} , \mathcal{S}_{xy} , \mathcal{S}_{yx} et \mathcal{S}_{yy} est d'isoler les dynamiques liées à une seule fonction de transfert afin de les identifier dans une structure de modèle SISO. Nous ne ferons que résumer cette approche dans cette section. Pour tout le formalisme mathématique, le lecteur est invité à lire la Section 7.3. Les grandes lignes de cette méthode (appelée SISO-par-SISO) sont données ci-dessous.

- Identification des fonctions de transfert de $\mathbf{G}_0(z)$** : Tout d'abord, nous considérons une excitation multisinus pour \bar{u}_x dont les pulsations sont choisies proche des pulsations $\omega_{0,x}/2$ et $\omega_{0,y}/2$. Grâce à la non-linéarité carrée, le signal d'entrée u_x^2 de $G_{0,xx}$ et $G_{0,yx}$ est un multisinus avec plus de pulsations que u_x et dont certaines se trouvent autour de $\omega_{0,x}$ et $\omega_{0,y}$. Nous collectons les données de sorties de \tilde{x} et \tilde{y} et nous calculons les données du signal u_x^2 . Les signaux \tilde{x} , \tilde{y} et u_x^2 sont ensuite filtrés avec un filtre passe-bande Butterworth d'ordre 6 qui ne garde qu'une mince bande de pulsations contenant les pulsations $\omega_{0,x}$ et $\omega_{0,y}$. Ceci est justifié par le fait que les dynamiques de $G_{0,xx}$ et $G_{0,yx}$ sont concentrées dans cette bande de pulsation. Ainsi, les modèles pour $G_{0,xx}(z)$ et $G_{0,yx}(z)$ (notés $G_{xx,init}(z)$ et $G_{yx,init}(z)$ respectivement) sont identifiées en utilisant comme sorties les signaux de sorties \tilde{x} et \tilde{y} filtrés obtenus et comme entrée le signal u_x^2 aussi filtré. En faisant de même avec u_y , nous pouvons modéliser de façon similaire les fonctions de transfert $G_{0,xy}(z)$ et $G_{0,yy}(z)$. Nous considérons une structure de modèle SISO de type OE pour chaque identification.
- Identification des fonctions de transfert de $\mathbf{E}_0(z)$** : Ensuite, nous considérons pour \bar{u}_x une séquence binaire pseudo-aléatoire (SBPA). Nous collectons ensuite les données de sorties de \tilde{x} et \tilde{y} . Après la collecte des données, nous générons les signaux \tilde{x}_{sim} et \tilde{y}_{sim} en filtrant le signal u_x^2 (calculé à partir de \bar{u}_x) avec les modèles précédemment identifiés $G_{xx,init}(z)$ et $G_{yx,init}(z)$ respectivement, i.e., $\tilde{x}_{sim} = G_{xx,init}(z)u_x^2(t)$ et $\tilde{y}_{sim} = G_{yx,init}(z)u_x^2(t)$ en utilisant la même signal d'excitation u_x qui a servi pour l'expérience. Nous construisons ensuite les signaux $\tilde{x} - \tilde{x}_{sim}$ et $\tilde{y} - \tilde{y}_{sim}$ qui seront considérés comme sorties pour l'identification de $E_{0,xx}$ et $E_{0,yx}$ respectivement. Ceci permet de supprimer les dynamiques mécaniques des sorties. L'entrée pour l'identification est u_x pour les deux modèles. L'identification de $E_{0,xy}$ et $E_{0,yy}$ se fait de façon similaire en considérant une excitation SBPA pour u_y . Nous considérons une structure de modèle SISO de type OE pour chaque identification.
- Identification des fonctions de transfert de $\mathbf{H}_0(z)$** : Enfin, en n'excitant ni u_x ni u_y , nous récoltons les données des sorties \tilde{x} et \tilde{y} pour la modélisation des fonctions de transfert des bruits de mesures $H_{0,x}(z)$ et $H_{0,y}(z)$ respectivement. Nous considérons une structure de modèle SISO de type ARMA pour chaque identification.

Le degré de chaque polynôme impliqué dans les structures de modèle est déterminé de telle sorte que la structure de modèle soit d'ordre complet et soit de complexité minimale. Ceci permet d'éviter de possibles compensations pôles/zéros au vrai vecteur de paramètre afin de garantir l'identifiabilité des structures de modèle. Ces degrés sont donnés dans les Tables 7.1, 7.2 et 7.3.

K.3.2 Résumé des résultats obtenus avec l'approche SISO-par-SISO et de leur interprétation

Les diagrammes de Bode de tous les modèles identifiés avec l'approche dite SISO-par-SISO sont donnés dans les Figures 7.1, 7.2 et 7.3. Examinons ces diagrammes de Bode, en commençant par ceux qui correspondent au couplage capacitif parasite (voir Figure 7.2). Comme prévu, les modèles identifiés pour $Q_{0,xx}$, $Q_{0,xy}$ et $Q_{0,yy}$ présentent tous un effet passe-haut. Nous observons également que le gain de ces modèles identifiés est important et c'est particulièrement le cas dans la gamme des pulsations de résonance du gyroscope (environ $2\pi \times 11836.8$ rad/s). Il semble qu'il n'y ait pratiquement aucun couplage capacitif du mode drive vers le mode sense mais cette estimation n'est pas fiable (voir Section 7.3). Enfin, il semble qu'il y ait un effet passe-bas en haute fréquence. Cela peut être dû à une dynamique supplémentaire provenant de la carte électronique.

Deuxièmement, interprétons les résultats des fonctions de transfert des dynamiques mécaniques identifiées (voir Figure 7.1). Premièrement, nous pouvons observer un comportement de résonance dans toutes les fonctions de transfert. Pour $G_{0,xy}$ et $G_{0,yx}$, nous pouvons observer deux résonances. Pour les termes directs $G_{0,xx}$ et $G_{0,yy}$ une seule résonance est visible. Les facteurs de qualité des résonances des modèles des fonctions de transfert $G_{0,xx}$ et $G_{0,yy}$ ont le même ordre de grandeur que ceux donnés par le fabricant du gyroscope. Pour les fonctions de transfert directs $G_{0,xx}$ et $G_{0,yy}$, un zéro instable a été identifié mais il ne correspond pas à un phénomène physique. Il pourrait provenir de la dynamique de la carte électronique. Cependant, il est important de noter que la validité des modèles de $\mathbf{G}_0(z)$ ne peut être garantie que dans la gamme de pulsations $2\pi \times [11750, 11950]$ rad/s en raison du pré-filtrage des données utilisé dans l'approche SISO-par-SISO. La validité des modèles de $\mathbf{Q}_0(z)$ est aussi restreinte puisque les modèles de $\mathbf{G}_0(z)$ ont été utilisés pour la détermination des fonctions de transfert de couplage capacitif parasite.

Enfin, pour la mesure du bruit (voir Figure 7.3), nous observons une pente de 20dB/décade dans la gamme des basses pulsations, qui est l'effet du bruit de scintillement (flicker noise en anglais). Ensuite, nous remarquons une partie constante dans le modèle identifié de $H_{0,x}$ et qui est due au bruit mécano-thermique. Elle n'est cependant pas observée dans $H_{0,y}$. Un petit pic est observé dans les haute fréquence dans H_x et une pente de -40 dB/décade est présente dans les haute fréquence pour les deux PSD. Des travaux supplémentaires devraient être effectués pour déterminer l'interprétation physique de ce pic et de la pente de -40 dB/décade en haute fréquence. Pour cette dernière, la dynamique supplémentaire provenant de la carte électronique que nous avons observé dans les modèles identifiés de \mathbf{Q}_0 pourrait en être la raison.

Il est clair que nous ne pouvons pas utiliser ces modèles pour la conception du correcteur, car nous avons quelques problèmes de validité avec les modèles des fonctions de transfert dans \mathbf{G}_0 et \mathbf{Q}_0 . Afin de contourner ce problème de validité, nous allons procéder à l'identification MISO expliquée dans la Section K.1 sans procéder à un pré-filtrage des données. Comme nous le verrons dans la section suivante, les modèles obtenus à l'aide de cette approche MISO seront très proches de ceux obtenus à l'aide de l'approche SISO-par-SISO. Toutefois, comme l'excitation d'entrée utilisée dans l'identification MISO sera plus large (ce n'était pas le cas, par exemple, pour l'identification des fonctions de transfert en \mathbf{G}), nous aurons davantage confiance dans les modèles obtenus.

K.4 Identifications MISO et résultats

En utilisant les critères d'identification (K.5) et (K.6), nous pouvons obtenir des modèles pour les sous-systèmes (K.3) et (K.4) correspondant au gyroscope MEMS (J.7) lorsque $u_y = 0$. Pour cela, nous devons réaliser une expérience en appliquant un signal u_x à (J.7) tout en gardant $u_y = 0$. Le choix du signal d'excitation u_x sera décrit dans la suite. La procédure d'optimisation pour résoudre les critères (K.5) et (K.6) sera initialisée avec les modèles obtenus dans la section précédente avec la méthode SISO-par-SISO. Par conséquent, la complexité des fonctions de transfert (i.e., leur degré et retard) des 4 structures de modèle MISO qui seront utilisées dans les critères (K.5) et (K.6) est la même que celle des fonctions de transfert des structures de modèle SISO utilisées lors de l'identification SISO-par-SISO. Ces degrés sont donnés dans les Tables 7.1, 7.2 et 7.3.

Nous ne décrivons que l'identification MISO des sous-systèmes (K.3) et (K.4) avec $u_x \neq 0$ et $u_y = 0$ (voir (K.5) et (K.6)). Une approche similaire peut être adoptée pour l'identification de \mathcal{S}_{xy} et \mathcal{S}_{yy} , c'est-à-dire lorsque $u_x = 0$ et $u_y \neq 0$.

K.4.1 Choix des entrées pour les identifications MISO

Premièrement, il est important que le spectre de u_x^2 soit non-nul dans le voisinage direct de $\omega_{0,x}$ et $\omega_{0,y}$ afin de bien exciter les résonances. Par conséquent, comme dans la section K.3, nous choisirons pour l'excitation multisinus \bar{u}_x des pulsations d'excitation autour de $\omega_{0,x}/2$ et $\omega_{0,y}/2$ comme dans la méthode SISO-par-SISO afin d'exciter uniquement les effets mécaniques autour des pulsations de résonance $\omega_{0,x}$ et $\omega_{0,y}$ grâce à la non-linéarité carrée. Deuxièmement, pour le couplage parasite, il est également important d'exciter en basse et haute fréquence puisque le couplage parasite a également une dynamique dans ces gammes de pulsations. À cette fin, l'excitation multisinus \bar{u}_x considérée est constituée de 250 pulsations réparties dans l'intervalle $[0, \omega_s/2]$. Elles sont regroupées en 6 ensembles et réparties régulièrement dans chaque ensemble comme suit

- Ensemble 1 (autour de $\omega_{0,x}/2$) : de $2\pi \times 5918.415$ jusqu'à $2\pi \times 5918.49$ avec un pas de $2\pi \times 0.0005$. Au total, il y a 151 pulsations.

- Ensemble 2 (entre $\omega_{0,x}/2$ et $\omega_{0,y}/2$) : de $2\pi \times 5918.5$ jusqu'à $2\pi \times 5921$ avec un pas de $2\pi \times 0.5$. Au total, il y a 6 pulsations.
- Ensemble 3 (autour de $\omega_{0,y}/2$) : de $2\pi \times 5921.5$ jusqu'à $2\pi \times 5922.2$ avec un pas de $2\pi \times 0.005$. Au total, il y a 141 pulsations.
- Ensemble 4 (basse pulsation) : de $2\pi \times 500$ jusqu'à $2\pi \times 5000$ avec un pas de $2\pi \times 100$. Au total, il y a 46 pulsations.
- Ensemble 5 (après $\omega_{0,y}/2$ et avant $\omega_{0,x}$) : de $2\pi \times 6000$ jusqu'à $2\pi \times 11000$ avec un pas de $2\pi \times 1000$. Au total, il y a 11 pulsations.
- Ensemble 6 (après $\omega_{0,y}$ et jusqu'à $\omega_s/2$) : de $2\pi \times 12000$ jusqu'à $2\pi \times 31000$ avec un pas de $2\pi \times 1000$. Au total, il y a 20 pulsations.

Cette distribution de pulsations est illustrée dans la Figure 7.4.

Pour le choix des amplitudes, nous décidons de choisir la même amplitude pour toutes les sinusoïdes. Pour les déphasages, nous les choisissons de façon aléatoire afin de diminuer le facteur de crête du signal multisinus \bar{u}_x pour ensuite maximiser sa puissance par rapport aux limitations expérimentales discutées en Section K.2. Pour cela, nous faisons 1000 tirages aléatoires des déphasages, chacun choisi de façon indépendante par rapport aux autres et chacun suivant une loi uniforme entre 0 et 2π . Puis, pour chaque tirage, nous simulons le multisinus obtenu de $t = 1$ à $t = N = 62500$ et nous déterminons la valeur absolue maximale. À partir de ces 1000 tirages, nous choisissons pour le multisinus \bar{u}_x les déphasages qui proviennent du tirage donnant la plus petite des valeurs absolues maximales obtenues. Le signal est ensuite normalisée par cette valeur absolue maximale afin de satisfaire $-1V \leq \bar{u}_x(t) \leq 1V \forall t = 1, \dots, N$. Enfin, le signal u_x est obtenu en ajoutant la composante continue DC u_{DC} (voir (K.2)). Bien que ce choix d'entrée permet d'augmenter la puissance maximale afin d'obtenir de faibles incertitudes sur les paramètres identifiés, l'informativité des données doit être garantie. L'étude de l'informativité des données est faite dans le paragraphe suivant.

K.4.2 Étude de l'informativité des données

Il est intéressant de noter que les résultats du Chapitre 5 (résumés dans l'Annexe I.3) nous permettent de vérifier que l'excitation choisie u_x donnera des données informatives pour les critères d'identification (K.5) et (K.6).

Premièrement, dans u_x , nous avons $s_1 = 250$ de pulsations non nulles et une valeur moyenne non-nulle. Par conséquent, la densité spectrale de puissance de u_x est non nulle aux pulsations $m_1 = 2s_1 + 1 = 501$.

Nous ne pouvons pas utiliser le Lemme I.7 avec $n = 2$ pour conclure sur l'informativité des données puisque nous avons quelques pulsations et leur double dans l'ensemble des pulsations, violant la condition (iii) du Lemme I.7. Par exemple, les deux pulsations $2\pi \times$

1000 et $2\pi \times 2000$ font partie de l'ensemble des pulsations d'excitation. Essayons alors d'utiliser le Lemme I.6 en calculant m_2 et $m_{1,2}$. Pour ce calcul, nous pourrions utiliser la procédure donnée dans la Section I.3.4. Cependant, même si toutes les amplitudes sont positives, les déphasages ne sont pas égaux à 0. Par conséquent, nous calculons la transformée de Fourier rapide de u_x^2 avec un grand nombre de données N' ($N' = 10^7$) afin de donner une borne inférieure pour m_2 . Nous calculons également la transformée de Fourier rapide de u_x afin de déterminer une borne inférieure pour $m_{1,2}$. L'observation de la transformée de Fourier rapide nous permet de conclure que $m_2 - m_{1,2} \geq 100$. Par conséquent, avec cette excitation, nous avons

$$\begin{aligned} \deg(B_{Q_{xx}}) + \deg(F_{Q_{xx}}) &\leq 501 = m_1 \\ \deg(B_{G_{xx}}) + \deg(F_{G_{xx}}) &\leq 100 \leq m_2 - m_{1,2} \end{aligned}$$

et que

$$\begin{aligned} \deg(B_{Q_{yx}}) + \deg(F_{Q_{yx}}) &\leq 501 = m_1 \\ \deg(B_{G_{yx}}) + \deg(F_{G_{yx}}) &\leq 100 \leq m_2 - m_{1,2} \end{aligned}$$

où les degrés des polynômes impliqués dans les inégalités précédentes ont été obtenus avec l'approche SISO-par-SISO et sont dans les Tables 7.1 et 7.2.

Par conséquent, d'après le Lemme I.6, l'excitation choisie u_x donnera des données informatives pour les critères d'identification (7.7) et (7.8).

K.4.3 Résultats des identifications MISO

En réalisant une expérience avec $u_y = 0$ et le signal u_x présenté dans la Section K.4.1 de $t = 1$ à $t = N = 62500$, nous déterminons les solutions des critères d'identification (K.5) et (K.6), en initialisant ces problèmes d'optimisation avec les modèles obtenus dans la Section K.3. Cette expérience d'identification conduit à des modèles pour les fonctions de transfert dans \mathcal{S}_{xx} et \mathcal{S}_{yx} . De même, en utilisant une expérience avec $u_x = 0$ et un signal u_y ayant la même expression que le signal u_x décrit dans la Section K.4.1, nous pouvons déterminer des modèles pour les fonctions de transfert dans \mathcal{S}_{xy} et \mathcal{S}_{yy} .

Notez que, même si les critères d'identification MISO sont initialisés en utilisant les modèles obtenus dans la Section K.3, nous avons pu observer quelques problèmes de minima locaux puisque les modèles identifiés de $G_{0,xy}$ et $G_{0,yx}$ nous ont donné des résultats absurdes (pulsations de résonance éloignées de $\omega_{0,x}$ et $\omega_{0,y}$). Nous avons alors contraint chaque paramètre de $\theta_{G_{xy}}$ et $\theta_{G_{yx}}$ de se situer entre -5% et $+5\%$ de leur valeur initiale (identifiée).

Les diagrammes de Bode des modèles obtenus comparés à ceux de l'approche SISO-par-SISO sont donnés dans les Figures F.1 et F.2 en Annexe F.1. Comme nous l'avons déjà mentionné, nous observons que les modèles identifiés à l'aide de l'approche MISO sont très proches des modèles initiaux obtenus à l'aide de l'approche SISO-par-SISO. Pour G_{xy} et G_{yx} , cela s'explique certainement par l'optimisation sous contrainte.

Nous pouvons alors synthétiser les correcteurs $\mathbf{K}_{comp}(z)$ et $\mathbf{K}(z)$. Dans le Chapitre 7, nous montrons que le correcteur $\mathbf{K}_{comp}(z)$ dimensionné à partir du modèle identifié de $\mathbf{Q}_0(z)$ supprime bien l'effet de couplage capacitif parasite. Cependant, pour la synthèse de $\mathbf{K}(z)$ avec la méthode \mathcal{H}_∞ sur la base des modèles identifiés de $\mathbf{G}_0(z)$ et $\mathbf{H}_0(z)$, nous obtenons un ordre trop important pour le correcteur $\mathbf{K}(z)$ pour permettre son implémentation sur la carte électronique. La raison provient de l'ordre du modèle identifié de $\mathbf{G}_0(z)$: il est d'ordre 12 alors que cette matrice de fonctions de transfert est supposée modéliser la vibration de deux masses. L'ordre important du modèle est causé par la méthode d'identification qui a été développée est qui identifie chaque sous-système de façon indépendante. Ainsi, les liens physiques entre les fonctions de transfert de $\mathbf{G}_0(z)$ ne sont pas identifiés ce qui explique l'ordre important obtenu. Dans le prochain paragraphe, nous donnons une méthode de réduction d'ordre du modèle.

K.5 Réduction d'ordre du modèle par identification boîte grise

L'approche qui sera considérée pour réduire l'ordre du modèle des dynamiques mécaniques est une identification de type boîte grise de ces dynamiques en s'inspirant des équations de la physique, données en Annexe J. La première motivation de cette approche est qu'elle permettra de réduire l'ordre du modèle des vibrations des deux masses. La deuxième motivation est l'intérêt pratique d'avoir un modèle où les paramètres ont une signification physique.

Pour rappel, les dynamiques mécaniques sont régies par les équations (J.1) et (J.2). Pour l'actionnement, les équations correspondantes à prendre en compte sont (J.3) et (J.4). Pour la mesure des déplacements, nous devrions considérer les équations (J.5) et (J.6). Or, il semble qu'il y ait des dynamiques additionnelles provenant de la carte électronique. En effet en haute fréquence, un effet filtrage passe-bas est visible dans les fonctions de transfert du bruit de mesure et du couplage capacitif parasite. Ainsi, au lieu de considérer (J.5) et (J.6), nous devrions considérer un effet filtrage passe-bas pour ces équations, c'est à dire

$$\frac{1}{\lambda_x} \dot{\tilde{x}}(t_c) + \tilde{x}(t_c) = \beta_x x_m(t_c) \quad (\text{K.7})$$

$$\frac{1}{\lambda_y} \dot{\tilde{y}}(t_c) + \tilde{y}(t_c) = \beta_y y_m(t_c) \quad (\text{K.8})$$

Cet effet passe-bas pourrait provenir du filtre anti-repliement de pulsation de coupure égale à $2\pi \times 15000$ rad/s.

Afin de réaliser l'identification boîte grise avec la fonction MATLAB `greyest`, il nous faut décrire les dynamiques mécaniques instrumentées entre le vecteur d'entrée $(u_x^2, u_y^2)^T$ et le vecteur de sortie $(\tilde{x}, \tilde{y})^T$ dans une représentation d'état. Pour cela, nous introdui-

sons le vecteur d'état suivant

$$\mathbf{w}(t_c) = \begin{pmatrix} x_{mec}(t_c) \\ \dot{x}_{mec}(t_c) \\ y_{mec}(t_c) \\ \dot{y}_{mec}(t_c) \\ \tilde{x}(t_c) \\ \tilde{y}(t_c) \end{pmatrix}$$

En combinant les équations (J.1), (J.2), (J.3), (J.4), (K.7) et (K.8), nous pouvons écrire les dynamiques mécaniques instrumentées en représentation d'état de la façon suivante

$$\begin{cases} \dot{\mathbf{w}}(t_c) = \mathbf{A}_w \mathbf{w}(t_c) + \mathbf{B}_w \begin{pmatrix} u_x^2(t_c) \\ u_y^2(t_c) \end{pmatrix} \\ \begin{pmatrix} \tilde{x}(t_c) \\ \tilde{y}(t_c) \end{pmatrix} = \mathbf{C}_w \mathbf{w}(t_c) \end{cases} \quad (\text{K.9})$$

où

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{k_{xx}}{m_x} & -\frac{d_{xx}}{m_x} & -\frac{k_{xy}}{m_x} & -\frac{d_{xy}}{m_x} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -\frac{k_{yx}}{m_y} & -\frac{d_{yx}}{m_y} & -\frac{k_{yy}}{m_y} & -\frac{d_{yy}}{m_y} & 0 & 0 \\ \lambda_x \beta_x & 0 & 0 & 0 & -\lambda_x & 0 \\ 0 & 0 & \lambda_y \beta_y & 0 & 0 & -\lambda_y \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} 0 & 0 \\ \frac{\alpha_x}{m_x} & 0 \\ 0 & 0 \\ 0 & \frac{\alpha_y}{m_y} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Il est à noter que l'ordre du modèle de la représentation d'état (K.9) est de 6. Il nous faut identifier les paramètres physiques dans ces matrices. Or, dans l'état actuel, nous rencontrons un problème d'identifiabilité en raison de la présence des divisions par les masses m_x et m_y . Ainsi, il nous fait considérer ces nouveaux paramètres pour régler le

problème d'identifiabilité :

$$\begin{array}{ll}
 \bar{k}_{xx} = k_{xx}/m_x & \bar{d}_{xx} = d_{xx}/m_x \\
 \bar{k}_{xy} = k_{xy}/m_x & \bar{d}_{xy} = d_{xy}/m_x \\
 \bar{k}_{yy} = k_{yy}/m_y & \bar{d}_{yy} = d_{yy}/m_y \\
 \bar{k}_{yx} = k_{yx}/m_y & \bar{d}_{yx} = d_{yx}/m_y \\
 \bar{\alpha}_x = \alpha_x/m_x & \bar{\alpha}_y = \alpha_y/m_y
 \end{array}$$

Il nous faut maintenant générer des données expérimentales pour identifier les paramètres physiques. La première possibilité est de synthétiser des signaux d'entrées u_x et u_y puis de réaliser une expérience d'identification sur le gyroscope MEMS. Cependant, ici, nous voulons juste réduire la complexité d'un modèle déjà identifié (modèle boîte noire de $\mathbf{G}_0(z)$ identifié avec 4 identifications MISO) et qui est précis. Ainsi, au lieu de générer expérimentalement les données, nous pouvons directement les simuler à partir de ce modèle. Ceci a plusieurs avantages. Premièrement, les données obtenues ne sont pas bruitées. Deuxièmement, nous n'avons pas à considérer la non-linéarité de type carré dans la synthèse des excitations. Finalement, nous n'avons pas à satisfaire les contraintes d'excitation de la carte électronique (tensions maximales/minimales, tension DC fixe, etc). Les entrées u_x^2 et u_y^2 qui seront alors considérées sont des bruits blancs gaussiens indépendants.

Même si il n'y a pas de bruits de mesure, nous pouvons toujours rencontrer des problèmes numériques de type minima locaux puisque l'optimisation est un problème non-linéaire. Il nous faut alors déterminer une estimée initiale précise des paramètres. L'idée est d'exploiter le modèle boîte noire identifié avec les identifications MISO. La méthode d'initialisation est décrite en Annexe F.3. L'idée suivie pour une estimation initiale des paramètres est de les choisir de telle façon que le modèle boîte grise ait les mêmes propriétés des résonances (gain statique, pulsations de résonances, facteurs de qualité, etc) que celles du modèle boîte noire.

Les Figures 7.10 et 7.11 donnent respectivement l'amplitude et la phase de la réponse fréquentielle du modèle boîte d'ordre 6 (qui a été discrétisé avec l'hypothèse d'un bloqueur d'ordre zéro en utilisant la fonction `c2d` de MATLAB) et le modèle boîte noire d'ordre 12. Les deux diagrammes de Bode sont très proches pour toutes les fonctions de transfert, suggérant que la réduction d'ordre par modélisation boîte grise fonctionne bien. Toutefois, les paramètres identifiés λ_x and λ_y sont proches de $2\pi \times 7000$ rad/s et non de $2\pi \times 15000$ rad/s. Ainsi, l'effet passe-bas ne semble pas être causé par le filtre anti-repliement. Bien que cela ne soit pas nécessaire pour la synthèse du correcteur $\mathbf{K}(z)$, il est toujours intéressant de déterminer sa nature et ceci est envisagé dans des travaux futurs.

A partir de ce modèle, le correcteur $\mathbf{K}(z)$ est synthétisé avec la méthode \mathcal{H}_∞ . Comme illustré dans le Chapitre 7, l'asservissement de la masse drive à la consigne sinusoïdale

x_{ref}) semble être performant. Pour le mode sense, il nous faut faire des expériences avec des vitesses angulaires Ω non-nulles pour vérifier les performances du contrôle sur ce mode. Cela est prévu dans des travaux futurs.

K.6 Conclusion sur le résumé des résultats du Chapitre 7

Dans cette annexe, nous avons résumé l'approche développée pour l'identification des dynamiques principales d'un gyroscope MEMS afin de concevoir un correcteur, c'est-à-dire répondre à la deuxième partie de la Question de Recherche 4. Cette approche a été basée sur l'identification MISO de 4 sous-systèmes. Cependant, une initialisation était nécessaire afin d'éviter un problème de minima locaux lors de ces identifications MISO et nous avons développé l'approche SISO-par-SISO pour y répondre. Ensuite, un signal a été conçu afin d'augmenter le rapport signal sur bruit et de garantir l'informativité des données simultanément, ceci dans l'unique but de garantir des estimations précises. Presque toutes les fonctions de transfert sont identifiées avec précision et un correcteur a pu être conçu basé sur ces modèles.

Comme expliqué au Chapitre 6 (résumé dans l'Annexe J), le correcteur dispose de deux degrés de liberté. Pour le bloc "Parasitic Compensation $\mathbf{K}_{comp}(z)$ " de la Figure 6.10, nous avons vu que le couplage capacitif parasite est bien compensé en utilisant le modèle identifié de $\mathbf{Q}_0(z)$. Pour le bloc "Tracking Controller $\mathbf{K}(z)$ ", nous avons dû réduire la complexité du modèle de $\mathbf{G}_0(z)$ afin d'éviter un correcteur d'ordre important pour pouvoir l'implémenter. À cette fin, nous avons envisagé une identification par boîte grise mais où les données sont simulées à partir du modèle identifié de $\mathbf{G}_0(z)$. Cela a permis une réduction importante de l'ordre du modèle avec les mêmes performances de modélisation et un correcteur a été conçu avec l'approche \mathcal{H}_∞ et basé sur ce modèle d'ordre réduit. Nous avons observé que les objectifs d'asservissement sont bien atteints pour le mode drive.

Il semble alors que la Question de Recherche 4 ait été résolue pour les objectifs de contrôle du mode drive. Bien sûr, il reste encore quelques expériences à faire, comme des expériences de vitesse angulaire non-nulle, afin de vérifier les performances du correcteur sur la précision de la mesure de la vitesse angulaire et les objectifs de contrôle du mode sense.

Dans l'annexe suivante, nous donnons la conclusion de cette thèse et quelques pistes pour des travaux futurs sont fournies.

Long résumé : Conclusion de la thèse

Contents

L.1	Conclusion sur les travaux réalisés pendant la thèse	297
L.2	Ouvertures pour des recherches futures	299
L.2.1	Restrictions sur les structures de modèle considérées	299
L.2.2	Conservatisme des conditions développées et erreurs numériques possibles sur le calcul du rang d'une matrice	300
L.2.3	Informativité des données pour les systèmes non-linéaires identifiables avec la méthode Prediction Error	300
L.2.4	De l'analyse de l'informativité des données à la synthèse d'excitation optimale	301
L.2.5	Gyroscope MEMS : incertitudes, analyse de la vitesse angulaire et dépendance avec la température	301

Dans ce chapitre, nous donnons la conclusion de ce mémoire. Nous donnons aussi quelques pistes de recherche pour des travaux futurs.

L.1 Conclusion sur les travaux réalisés pendant la thèse

Dans cette thèse, nous avons d'abord étudié l'informativité des données par rapport aux structures de modèle MIMO linéaires qui sont des généralisations naturelles des structures de modèle SISO classiques : FIR, ARX, ARMAX, OE et BJ. Nous avons considéré deux cas qui sont l'identification en boucle ouverte (Question de recherche 1) et l'identification directe en boucle fermée (Question de recherche 2). Dans les deux cas, les types d'excitation qui ont été pris en compte sont les bruits blancs filtrés et les multisinus, qui sont deux types d'excitation couramment utilisés dans l'identification

par Prediction Error. Cette étude a été motivée par l'absence de conditions pas trop conservatrices pour la vérification de l'informativité des données pour l'identification des systèmes MIMO. Nous avons vu dans les Chapitres 3 et 4 (résumés dans les Annexes I.1 et I.2) que l'informativité des données peut être vérifiée en analysant le rang d'une matrice qui dépend de la complexité de la structure du modèle, du correcteur (pour les expériences en boucle fermée) et de la paramétrisation de l'excitation. Même si cette nouvelle condition s'accompagne d'une dose de conservatisme pour la plupart des structures de modèle MIMO considérées, elle est toujours moins conservatrice que celles développées dans la littérature, sauf pour les correcteurs ayant des hypothèses particulières pour l'identification en boucle fermée des systèmes MISO ARMAX sans excitation externe [YZ18]. Les conditions développées dans cette thèse n'ont pas besoin de faire ces hypothèses/restrictions. Par conséquent, cette nouvelle condition de vérification de l'informativité des données répond bien aux Questions de recherche 1 et 2.

Toutefois, l'étude précédente ne s'intéresse qu'aux systèmes linéaires, alors que la plupart des systèmes de la vie réelle présentent certains comportements non-linéaires qui peuvent être identifiés avec précision dans le cadre de l'identification par erreur de prédiction. C'est pourquoi, dans le Chapitre 5 (résumé dans l'Annexe I.3), nous avons étudié l'informativité des données pour une classe particulière de systèmes non-linéaires donnée par (I.27) (Question de recherche 3). Cette étude particulière a été motivée par le système réel étudié dans le cadre de cette thèse : le gyroscope MEMS. Ses dynamiques principales peuvent être modélisées avec une telle structure. Afin de répondre à la Question de recherche 3, nous avons vu que le système (I.27) peut être réécrit comme un système linéaire MISO avec 2 entrées : $u_1 = u$ et $u_2 = u^n$. Par conséquent, les résultats développés dans le Chapitre 3 peuvent être utilisés pour vérifier l'informativité des données pour l'identification d'un tel système non-linéaire. Cependant, dans le cas où n est important et où l'excitation u est de type multisinus, le résultat principal développé au Chapitre 3 (persistance d'excitation d'un régresseur particulier) peut nécessiter un coût de calcul élevé pour vérifier l'informativité des données puisque nous devons calculer toutes les amplitudes, les déphasages et les pulsations du multisinus u^n . Par conséquent, des résultats supplémentaires nécessitant moins de calculs (mais avec plus de conservatisme) que ceux obtenus dans le Chapitre 3 ont été développés dans le Chapitre 5 pour vérifier l'informativité des données dans ce cas.

Enfin, dans la partie pratique de cette thèse, nous avons considéré le problème de l'identification du gyroscope MEMS. Nous avons vu dans le Chapitre 6 (résumé dans l'Annexe J) qu'il est important de concevoir un correcteur optimal pour le gyroscope MEMS afin d'augmenter la précision de l'estimation de la vitesse angulaire. Par conséquent, un modèle précis des dynamiques principales du gyroscope MEMS est nécessaire. Cependant, la littérature ne propose ni une structure de modèle complète ni une méthode d'identification pour identifier les trois principales dynamiques du gyroscope MEMS qui sont le couplage mécanique, le couplage capacitif parasite et le bruit de mesure (Question de recherche 4). C'est pourquoi, dans le Chapitre 6, nous avons développé

une structure de modèle paramétrique complète pour l'identification de ces trois phénomènes (première partie de la Question de Recherche 4). Dans le Chapitre 7 (résumé dans l'Annexe K), une méthode d'identification a été développée afin d'estimer tous les fonctions de transfert dans cette structure modèle (deuxième partie de la Question de Recherche 4). Sur la base du modèle identifié du gyroscopes MEMS, un correcteur à deux degrés de liberté a été conçu par Fabrício SAGGIN, un autre doctorant travaillant dans le projet Next4MEMS. Nous avons pu observer de grandes performances pour les objectifs d'asservissement et de rejet du couplage capacitif parasite. Par conséquent, pour les objectifs de contrôle considérés dans la thèse, notre méthodologie d'identification semble répondre à la Question de Recherche 4. Cependant, nous devons encore analyser la qualité de l'estimation de la vitesse angulaire avec ce correcteur.

L.2 Ouvertures pour des recherches futures

L.2.1 Restrictions sur les structures de modèle considérées

Dans les Chapitres 3 et 4, nous avons considéré les systèmes MIMO pour lesquels la matrice de fonction sde transfert du bruit $\mathbf{H}_0(z)$ a été supposée diagonale. C'est la première limitation de l'étude présentée dans ce mémoire. Heureusement, pour le cas de la boucle ouverte et pour les structures de modèle FIR, ARX, OE et BJ, nous pouvons encore relier l'informativité des données à la persistance des mêmes régresseurs ϕ_u données dans le Théorème I.3 même si la matrice $\mathbf{H}_0(z)$ n'est pas diagonale. Cela a été prouvé dans [CBBM20b]. Toutefois, pour les structures du modèle ARMAX, cette hypothèse est nécessaire pour obtenir le résultat de l'Annexe B.5.

Nous avons également considéré que tous les numérateurs et dénominateurs des entrées de $\mathbf{G}(z, \theta)$ et $\mathbf{H}(z, \theta)$ ne partagent pas de paramètres communs (à l'exception des paramètres dans le vecteur $\boldsymbol{\eta}$ pour les structures des modèles ARX et ARMAX). Toutefois, pour certains systèmes réels, les dynamiques peuvent être, par exemple, modélisées dans une structure modèle pour laquelle les fonctions de transfert ont certains paramètres communs. C'est le cas du gyroscope MEMS puisque, en convertissant la représentation d'état boîte grise en une matrice de fonctions de transfert (voir l'Annexe F.3), nous pouvons observer que les fonctions de transfert ont certains paramètres communs. Dans ces cas là, l'identification d'un système au sein d'une structure de modèle d'ordre complet avec une paramétrisation commune réduit le nombre de paramètres, ce qui par la suite diminue la variance des paramètres identifiés. De plus, une paramétrisation commune implique souvent des liens physiques entre les fonctions de transfert du système, également importants à identifier. Par conséquent, l'étude de l'informativité des données avec des paramètres communs est un sujet intéressant mais aussi un important défi à relever dans le futur.

L.2.2 Conservatisme des conditions développées et erreurs numériques possibles sur le calcul du rang d'une matrice

Comme expliqué dans les Chapitres 3 et 4 (résumés en Annexes I.1 et I.2), nous pouvons pour le moment conclure que le critère de vérification du rang développé dans cette thèse est une condition suffisante pour ARMAX, OE et BJ en boucle ouverte et pour ARX, ARMAX, OE et BJ en boucle fermée. Par exemple, dans le cas de la boucle fermée, nous ne pouvons pas facilement conclure que la vérification du rang de $\mathbf{P}^{(c,i)}$ donnée dans le Théorème I.8 est nécessaire puisque les ensembles \mathbf{D}_i dans (I.22) ne couvrent pas tout l'espace \mathbb{R}^{ζ_i} , sauf pour les structures du modèle FIR pour lesquelles nous avons $\mathbf{D}_i = \mathbb{R}^{\zeta_i}$. Mais nous pouvons encore en déduire une condition nécessaire et suffisante. Comme mentionné dans la Remarque 4.3, dans le cas d'une boucle fermée avec une excitation externe non nulle ($\mathbf{r} \neq \mathbf{0}_{n_u \times 1}$), une condition nécessaire et suffisante pour l'informativité des données est que, pour chaque sortie $i = 1, \dots, n_y$, l'intersection du noyau de $\mathbf{P}^{(c,i)T}$ et de l'ensemble \mathbf{D}_i est égale à $\mathbf{0}$. Toutefois, cela nécessite une étude géométrique des ensembles \mathbf{D}_i qui sera effectuée dans des travaux futurs. Un raisonnement similaire s'applique au cas de la boucle ouverte où les ensembles correspondants \mathbf{D}_i sont examinés dans la preuve du Théorème I.3.

Le deuxième problème est lié au fait que la nouvelle condition développée dans cette thèse est basée sur une vérification de rang d'une certaine matrice qui dépend de la paramétrisation de l'excitation, du correcteur (pour les expériences en boucle fermée) et de la complexité de la structure du modèle. Cependant, comme mentionné dans la Remarque 3.5, un calcul de rang peut être inexact. Des travaux supplémentaires devraient être effectués afin de contourner ce problème.

L.2.3 Informativité des données pour les systèmes non-linéaires identifiables avec la méthode Prediction Error

Dans le Chapitre 5 (résumé dans l'Annexe I.3), nous avons examiné une classe très particulière de systèmes Hammerstein donnée par (I.27). Le choix de cette restriction a été motivé par l'identification du gyroscope MEMS dont les dynamiques principales peuvent être modélisées au sein de telles structures de modèle avec $n = 2$. Cependant, cette structure de modèle non-linéaire particulière est assez restrictive pour l'étude de l'informativité des données collectées sur d'autres systèmes réels non-linéaires. Nous pouvons alors nous demander comment généraliser l'étude de l'informativité des données pour des structures de modèle parallèles de Hammerstein plus complexes avec différents types de non-linéarités statiques d'entrée telles que des polynômes, des non-linéarités de type « dead-zone », des saturations, etc. Nous pouvons également envisager d'autres types de structures de modèle parallèles bloc-orientées comme Wiener, Hammerstein-Wiener, Wiener Hammerstein, etc [GB10, ST17].

L.2.4 De l'analyse de l'informativité des données à la synthèse d'excitation optimale

Comme mentionné dans le Chapitre 1 (résumé dans l'Annexe G), l'informativité des données est une propriété cruciale pour obtenir un modèle précis, à condition que le nombre de données N puisse être augmenté. En effet, l'informativité des données (combinée à l'identifiabilité globale en θ_0) ne garantit que la convergence du vecteur paramètre identifié $\hat{\theta}_N$ vers le vrai vecteur paramètre θ_0 . Cependant, pour un nombre de données N fixé, toutes les excitations qui donnent des données informatives par rapport à la structure du modèle ne fournira pas les mêmes incertitudes sur les paramètres identifiés (c'est-à-dire la même matrice de covariance qui est une mesure des incertitudes). Pour certains choix (malheureux) d'excitations qui donnent des données informatives, nous pourrions avoir besoin d'augmenter très largement le nombre de données N pour garantir un niveau d'incertitudes relativement petit, ce qui n'est pas idéal dans la pratique. C'est pourquoi, dans le Chapitre 7, nous avons conçu les excitations pour l'identification du gyroscope MEMS de telle sorte que les données soient informatives et que le rapport signal sur bruit soit augmenté simultanément afin d'obtenir de petites variations sur les paramètres identifiés pour un nombre raisonnable de données N .

Par conséquent, l'étape suivante après l'étude de l'informativité des données est naturellement le choix de l'excitation optimale qui garantit des niveaux d'incertitude souhaités : c'est par exemple le cas de la synthèse d'entrée optimale où l'entrée synthétisée doit être celle qui a une puissance minimale et qui satisfait certaines contraintes d'incertitudes souhaitées [BSG⁺06, BBHS08, JH05].

L'étude fournie dans cette thèse peut aider dans le choix de la paramétrisation de l'entrée afin de réaliser cette optimisation. En effet, l'informativité des données implique la positivité définie de la matrice de covariance, qui est une condition nécessaire pour réaliser la synthèse d'entrée optimale. Par exemple, pour une structure de modèle MISO ARX avec excitation de type multisinus en boucle ouverte, nous avons vu qu'il y a un nombre nécessaire de sinusoides qui doit être garanti pour l'informativité des données. En d'autres termes, pour la conception optimale de l'excitation multisinus, il est nécessaire de satisfaire cette condition de richesse, sinon l'optimisation ne donnera jamais une solution pour laquelle la matrice de covariance est définie positive. Des travaux futurs devraient être menés pour relier les résultats développés sur l'informativité des données pour l'identification des systèmes MIMO à la synthèse d'expérience optimale pour leur identification.

L.2.5 Gyroscope MEMS : incertitudes, analyse de la vitesse angulaire et dépendance avec la température

Pour le gyroscope MEMS considéré dans cette thèse, un correcteur MIMO a été conçu sur la base du modèle identifié. Les objectifs d'asservissement du mode drive, présentés au Chapitre 6 (et résumés dans l'Annexe J), ont été évalués et vérifiés. Cependant, dans ces objectifs de contrôle, il n'y a pas de spécification sur l'estimation de la vitesse

angulaire. Dans les travaux futurs, il sera important de vérifier la qualité de la mesure de la vitesse angulaire obtenue avec ce correcteur MIMO. Ce type d'expériences sera également utile pour vérifier les performances d'asservissement du mode sense.

Dans le Chapitre 7 (résumé en Annexe K), nous avons développé un modèle pour les trois principales dynamiques du gyroscope MEMS. Cependant, un modèle seul n'est pas suffisant pour la synthèse d'un correcteur. Il est également important de prendre en compte les incertitudes sur les paramètres identifiés. Comme nous avons conçu l'excitation multisinus de manière à augmenter le rapport signal sur bruit, nous pouvons nous attendre à de petites incertitudes. Heureusement, avec la méthode Prediction Error, nous pouvons estimer les incertitudes¹ pour une probabilité donnée α pour vérifier ce fait [BSG⁺06]. Nous obtenons alors un ensemble de modèles (ensemble incertain) pour lequel nous savons que le vrai modèle appartient avec une probabilité de α . Une première étude sur les fonctions de transfert directes G_{xx} , E_{xx} , G_{yy} et E_{yy} et avec $\alpha = 0,99$ a montré que ces incertitudes sont effectivement relativement faibles. Toutefois, des travaux supplémentaires devraient être menés pour vérifier que les incertitudes sur toutes les fonctions de transfert identifiées sont suffisamment faibles pour garantir que l'ensemble des performances du correcteur (objectifs de contrôle du mode drive *et* du mode sense) sont toujours valables pour tous les modèles de l'ensemble incertain et donc pour le système réel (inconnu). Cette étude est l'essence même du domaine scientifique appelé l'analyse de robustesse. Dans le projet Next4MEMS, cette partie est étudiée par Jorge AYALA, un des quatre doctorants travaillant dans le projet Next4MEMS.

Comme mentionné précédemment, le correcteur MIMO a été conçu sur la base de l'identification des trois principaux phénomènes impliqués dans le gyroscope MEMS : le transfert mécanique, le couplage capacitif parasite et le bruit de mesure. Cependant, il existe d'autres dynamiques qui peuvent détériorer les performances de ce correcteur MIMO. Par exemple, les dynamiques du transfert mécanique dépendent fortement de la température. Plus précisément, les deux résonances $\omega_{0,x}$ et $\omega_{0,y}$ varient lorsque la température change en raison des variations de la raideur des micro-poutres siliconées et utilisées pour fixer les deux masses du MEMS (voir Chapitre 6). Dans le Chapitre 7, nous avons considéré une température fixe (20 °Celsius) pour le gyroscope MEMS et donc des résonances fixes $\omega_{0,x}$ et $\omega_{0,y}$. Cette dépendance avec la température n'a donc pas posé de problèmes lors de la procédure d'identification présentée au Chapitre 7. Cependant, le but du projet Next4MEMS est d'utiliser ces capteurs pour des applications aérospatiales pour lesquelles la température varie beaucoup. De plus, comme mentionné au Chapitre 6, il est important que la pulsation ω_{ref} de la consigne x_{ref} soit égale à $\omega_{0,x}$. Pour vérifier que le correcteur MIMO LTI calculé peut gérer ce problème de dépendance avec la température, une analyse de robustesse a été effectuée par Jorge AYALA pour différentes valeurs de $\omega_{0,x}$ et $\omega_{0,y}$ (c'est-à-dire pour différentes températures). Il a montré que les performances (comme la stabilité de la boucle) sont détériorées avec le correcteur MIMO LTI conçu à la température de 20 °C (celle présentée dans le Chapitre 7) pour toute la gamme de température $[-40, 80]^{\circ}\text{C}$, qui est dans l'ensemble des

¹si la consistance de l'estimateur est garantie.

températures considérées dans le cahier des charges du gyroscope MEMS du projet.

Pour résoudre ce problème, la possibilité qui est suivie par Fabrício SAGGIN est de concevoir un correcteur MIMO linéaire paramètre variant (LPV) dont le paramètre variable est $\omega_{0,x}$ (et non la température puisqu'il n'y a pas de thermomètre implémentée dans la carte électronique embarquant le gyroscope MEMS). Pour la synthèse du correcteur LPV, nous avons besoin de deux ingrédients : un modèle LPV du transfert mécanique et une méthode pour estimer la pulsation de résonance $\omega_{0,x}$ afin d'adapter le correcteur et la pulsation ω_{ref} du signal de référence x_{ref} (qui doit correspondre à $\omega_{0,x}$). Ce dernier problème est étudié par Federico MORELLI, un autre doctorant participant au projet Next4MEMS. Deux approches sont envisagées dans ses travaux : l'extremum seeking [ZO11] pour l'adaptation du signal de référence (ou consigne) et l'identification récursive [LS83] pour l'identification en ligne de $\omega_{0,x}$ afin d'adapter ω_{ref} et/ou les paramètres du correcteur.

Pour obtenir un modèle LPV pour le transfert mécanique, nous pouvons considérer l'approche locale [T6t10, GBH⁺18, LGPW14, MREL18] afin d'utiliser la méthode d'identification développée dans le Chapitre 7. L'idée de cette approche est en deux étapes. Tout d'abord, nous calculons différents modèles d'ordre réduit boîte grise en suivant l'approche du Chapitre 7 (c'est-à-dire, premièrement une identification boîte noire et deuxièmement une réduction boîte grise du modèle boîte noire) pour différentes températures. Ensuite, nous déterminons les fonctions qui décrivent l'évolution de chaque paramètre du modèle de la boîte grise par rapport à la pulsation de résonance $\omega_{0,x}$ en effectuant, par exemple, une interpolation polynomiale pour chaque paramètre. Cette possibilité sera étudiée dans de futurs travaux.

AUTORISATION DE SOUTENANCE

Vu les dispositions de l'arrêté du 25 mai 2016,

Vu la demande du directeur de thèse

Monsieur X. BOMBOIS

et les rapports de

M. J. SCHOUKENS

Professeur Emérite - Vrije Universiteit Brussel - Department of Engineering Technology (INDI) -
4th floor - Pleinaan 5 - 1050 Brussels - Belgique

et de

M. G. MERCERE

Maître de Conférences HDR - LIAS-ENSIP - Bât B25 - 2 rue Pierre Brousse - TSA 41105 -
86073 Poitiers cedex 9

Monsieur COLIN Kévin

est autorisé à soutenir une thèse pour l'obtention du grade de **DOCTEUR**

Ecole doctorale ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE

Fait à Ecully, le 26 août 2020

P/Le directeur de l'E.C.L.
Le directeur des Etudes

