Model-based System Engineering for Powertrain Systems Optimization

Sandra Hamze

UNIVERSITÉ GRENOBLE ALPES

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE ALPES

Spécialité : **Automatique-Productique**

Arrêté ministériel : 7 août 2006

Présentée par
Sandra HAMZE

Thèse dirigée par **Emmanuel WITRANT** et codirigée par **Delphine BRESCH-PIETRI** et **Vincent TALON**
préparée au sein du
laboratoire Grenoble Images Parole Signal Automatique (GIPSA-lab)
dans **l'école doctorale Electronique, Electrotechnique, Automatique, Traitement du Signal (EEATS)**
en collaboration avec **Renault s.a.s**

Optimisation Multi-objectifs Inter-systèmes des Groupes Motopropulseurs

Model-based System Engineering for Powertrain Systems Optimization
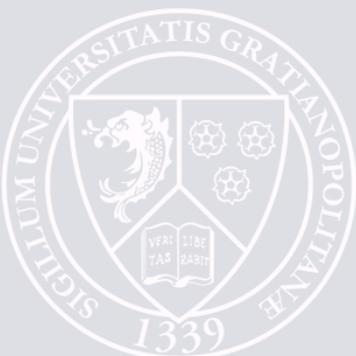
Thèse soutenue publiquement le
5 Novembre 2019,
devant le jury composé de:

Michel DAMBRINE
UVHC-Ensiame, Examinateur, Président du jury
Sorin OLARU
CentraleSupélec, Rapporteur
Yann CHAMAILLARD
Université d'Orléans, Rapporteur
Emmanuel WITRANT
GIPSA-lab, Directeur de thèse
Delphine BRESCH-PIETRI
CAS-MINES ParisTech, Encadrante de thèse
Vincent TALON
Renault s.a.s, Encadrant de thèse

UNIVERSITÉ DE GRENOBLE ALPES

ÉCOLE DOCTORALE EEATS
Electronique, Electrotechnique, Automatique, Traitement du Signal

T H È S E

pour obtenir le titre de

docteur en Automatique

de l'Université de Grenoble Alpes
Mention : Automatique-Productique

Présentée et soutenue par

Sandra HAMZE

Optimisation Multi-objectifs Inter-systèmes des Groupes Motopropulseurs

Model-based System Engineering for Powertrain Systems Optimization

Thèse dirigée par Emmanuel Witrant

préparée au laboratoire Grenoble Images Parole Signal Automatique (GIPSA-lab)

soutenue le 5 Novembre 2019

Jury :

Président et examinateur :	Michel DAMBRINE	-	UVHC-Ensiame
Rapporteur :	Sorin OLARU	-	CentraleSupélec
Rapporteur :	Yann CHAMAILLARD	-	Université d'Orléans
Directeur :	Emmanuel WITRANT	-	GIPSA-lab
Encadrants :	Delphine BRESCH-PIETRI	-	CAS-MINES ParisTech
	Vincent TALON	-	Renault s.a.s

Acknowledgement

This PhD thesis is the fruit of three years of hard work and dedication. It wouldn't have been feasible without the support of many to whom I extend my sincere gratitude.

First of all, I extend my utmost gratitude to my director Mr. Emmanuel WITRANT and my co-director Miss Delphine BRESCH-PIETRI from GIPSA-lab for their worthful efforts in orienting the work done throughout the PhD, in providing insight and expertise that greatly assisted the workflow, and in proofreading this manuscipt. On the other hand, I am immensely grateful to Mr. Vincent TALON from Renault s.a.s for granting me the opportunity of doing this PhD, and for taking my hand throughout the industrial part of the realized work. It is their knowledge and experience that helped put my work on the right track.

Besides, I take this opportunity to thank Mr. Sorin OLARU, Mr. Yann CHAMAILLARD, and Mr. Michel DAMBRINE for their participation in the jury, and for their efforts in carefully examining this manuscript and reporting the necessary amendments.

I would also like to send my sincere thanks to my colleagues in Renault with whom I shared a friendly atmosphere and an excellent working environment during the first year of my PhD and to my friends in GIPSA-lab who provided an enjoyable multicultural ambience which I enjoyed during the second and third years of the PhD.

Last but not least, I share the credit of my work with my parents, my sister Carol, and my fiancé Hadi. Their encouragement and morale helped me overcome the work stress and pushed me forward all the way through these three years of hard work.

It is only taking me a moment to thank you all, but your thoughtfulness I will remember a long time!

Contents

Contents v

List of Figures

List of Tables

Nomenclature

ARMAX   *Auto-Regressive Moving-Average eXogenous*

ARX   *Auto-Regressive with eXogenous input*

CLF   *Control-Lyapunov Function*

CVA   *Canonical Variable Algorithm*

ECU   *Engine Control Unit*

EGR   *Exhaust Gas Return/Recirculation*

HCCI   *Homogeneous Charge Compression Ignition*

HIL   *Hardware-In-the-Loop*

ICE   *Internal Combustion Engine*

INCOSE   *International Council on Systems Engineering*

IO   *Input-Output*

IVM   *Instrumental Variable Methods*

LPV   *Linear Parameter-Varying*

LQG   *Linear Quadratic Gaussian*

LQR   *Linear Quadratic Regulator*

LTI   *Linear Time-Invariant*

LTV   *Linear Time-Varying*

MBSE   *Model-Based Systems Engineering*

MVEM   *Mean-Value Engine Model*

MIL   *Model-In-the-Loop*

MOESP   *Multivariable Output-Error State-SPace*

MPC   *Model Predictive Control*

NEDC   *New European Driving Cycle*

NRMSE   *Normalized Root Mean Square Error*

N4SID   *Numerical algorithms for Subspace State-Space System IDentification*

ODE   *Ordinary Differential Equation*

PDE   *Partial Differential Equation*

PI   *Proportional Integral*

PID   *Proportional Integral Derivative*

PIL   *Processor-In-the-Loop*

PPC   *Partially Premixed Combustion*

PRBS   *Pseudo-Random Binary Signal*

PWA   *Piece-Wise Affine*

PWARX   *PieceWise Auto-Regressive eXogenous*

PWM   *Pulse Width Modulation*

RDE   *Real Driving Emissions*

SD   *Secure Digital*

SIL   *Software-In-the-Loop*

SMC   *Sliding Mode Control*

SPI   *Serial Peripheral Interface*

SS   *State-Space*

SSARX   *Sub-Space identification method using an ARX estimation based algorithm to compute weighting*

TD   *Technical Definition*

WLTC   *Worldwide harmonized Light vehicles Test Cycle*

EDO   *Equation aux Dérivées Ordinaires*

EDP   *Equation aux Dérivées Partielles*

GMP   *Groupe MotoPropulseur*

Contexte et Objectif de la Thèse

Chapter Summary

1.1 Ingénierie des systèmes à base de modèles

"L'ingénierie des systèmes à base de modèles (MBSE) est l'application formalisée de la modélisation à l'appui des exigences, de la conception, de l'analyse, de la vérification et de la validation des systèmes, dès la phase de conception jusqu'à la fin des phases de développement et du cycle de vie." [FGS07]

Bien que les gens utilisent des modèles dans l'ingénierie des systèmes depuis des décennies, les modèles de systèmes n'ont pas encore suffisamment remplacé les documents dans les processus d'ingénierie et les communications jusqu'à présent. Pour aller dans ce sens, l'International Council on Systems Engineering (INCOSE) a introduit, en 2007, l'approche MBSE dans un cadre bien défini, accompagné d'une feuille de route qui établit une vision pour l'avenir des pratiques en ingénierie système [FGS07].

Aujourd'hui, après une dizaine d'années de l'initiative MBSE de INCOSE, MBSE s'applique déjà à un nombre considérable de domaines (électrique, mécanique, etc. . .), implique de nombreux types de modèles (statiques, variables dans le temps, déterministes, stochastiques, etc...) et de techniques de modélisation (Computer Aided Design, Analytical modeling, etc...) et offre de nombreux avantages [Ber11], tels que:

- l'amélioration de la productivité et de l'abordabilité de l'ingénierie des systèmes;

- la réduction du coût de la conception du système et la limitation de ses risques en vérifiant les performances et les exigences du système dès le début du processus;

- l'anticipation des changements et des améliorations du système avant la mise en œuvre, c'est-à-dire avant que le changement ne devienne coûteux;

- l'amélioration de la qualité et de la fiabilité des systèmes conçus;

- la fourniture de multiples alternatives de conception en fournissant une rétroaction rapide sur les décisions de conception;

- la promotion de l'interopérabilité et de l'intégration des modèles tout au long du cycle de vie du système, ce qui ouvre la porte à la réutilisation des modèles et qui permet un environnement de conception unifié, de la phase d'identification des exigences du système à la phase de test et de validation.

1.2   Le groupe motopropulseur

Le groupe motopropulseur (GMP) d'une automobile est l'ensemble des composants qui transforment la puissance de son moteur en mouvement. Bien que la définition demeure inchangée, que le moteur soit un moteur à combustion interne, électrique, hybride ou d'un autre type, les composants du GMP peuvent varier. Dans les véhicules traditionnels équipés d'un moteur à combustion interne, le moteur, la transmission, l'arbre de transmission, les différentiels, les essieux et tout ce qui se trouve entre le moteur et les roues sont les composantes qui composent le GMP.

La composante du GMP qui fait l'objet d'une visite exclusive dans cette
thèse est le moteur. En plus de considérer le moteur dans son ensemble,
cette étude tient également compte de deux sous-composantes du mo-
teur: le papillon électrique et le circuit d'air du moteur, en particulier
l'échangeur de chaleur et le recirculateur des gaz d'échappement EGR.
Les chapitres suivants traiteront plus en détail de ces sous-composantes.

1.3 Ingénierie des systèmes à base de modèles pour l'optimisation des systèmes du groupe motopropulseur

En raison de ses mérites évidents, MBSE a été reçu par l'industrie au-
tomobile à bras ouverts. L'industrie est consciente de la nécessité de
transformer l'ingénierie basée sur le texte en ingénierie basée sur le
modèle, car la première est incapable de faire face à la complexité
croissante des systèmes automobiles et aux exigences de plus en plus
strictes des gouvernements et des clients. De la réduction des émissions
à l'économie de carburant en passant par la performance supérieure, les
défis actuels et futurs sont nombreux, et l'évolution du développement
automobile est impérative.

1.3.1 Conception d'un système de commande basé sur un modèle

L'un des piliers de l'évolution des concepts de conception automobile
est l'optimisation et le contrôle basés sur des modèles. Avec les modèles
orientés commande à portée de main, la conception de commande basée
sur modèle entre en jeu et apporte toute une liste d'avantages [Mic+10],
tels que:

- la possibilité d'évaluer le comportement et la robustesse du sys-
  tème en boucle fermée au début du cycle de conception. De nom-
  breuses architectures de contrôle peuvent donc être testées sur le
  modèle avant la génération de tout logiciel ou la construction de
  tout hardware, ce qui minimise les coûts d'ingénierie et le temps
  de conception;

- l'assistance à la définition des exigences et spécifications des nou-

veaux systèmes moteurs (définitions techniques des moteurs, exigences de contrôle, exigences de mise au point, etc...), en offrant la possibilité de simuler des systèmes en boucle fermée dans des cycles de conduite normatifs et réels;

- l'aide à la conception du contrôle en fournissant un cadre pour vérifier l'interaction entre le modèle du système et le contrôleur, que ce soit sous la forme d'un code logiciel ou une fois intégré à un microcontrôleur;

- la réduction des coûts des moteurs en permettant l'intégration des observateurs.

Le modèle en V, illustré à la Figure 1.1, est un outil graphique utilisé pour décomposer les différentes étapes de la conception du système de contrôle, depuis l'établissement des exigences du système et la construction de son modèle jusqu'à la validation et la mise au point de la conception sur le hardware prévu.
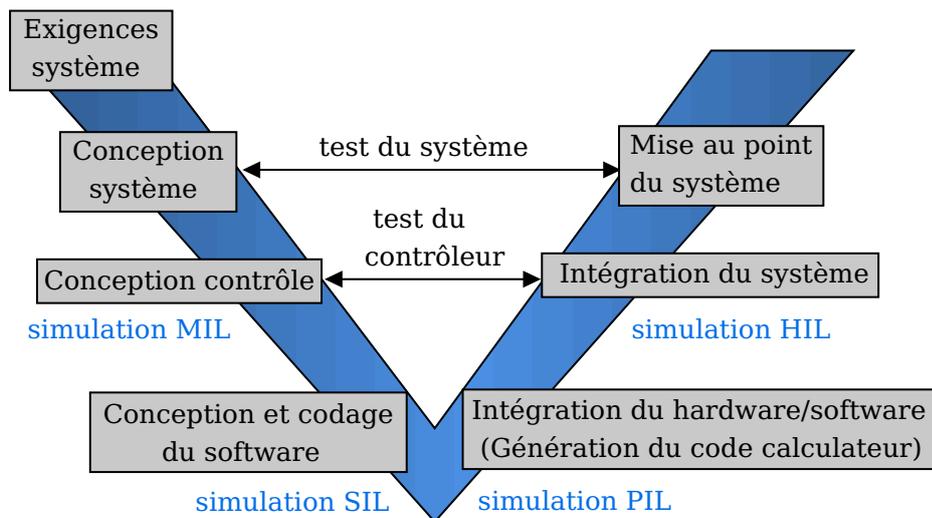
Figure 1.1: Cycle de développement des systèmes de contrôle

Les travaux de cette thèse se déroulent dans la phase de conception du contrôle. Dans cette phase, également appelée phase MIL (Model-in-the-Loop), les modèles des systèmes orientés commande sont simulés en boucle fermée avec leurs conceptions de contrôle-commande correspondantes. Les systèmes en boucle fermée qui en résultent sont ensuite validés et vérifiés pour répondre aux exigences nécessaires pendant les cycles de conduite réels. Après la phase MIL, les phases suivantes

reprennent et complètent le processus de développement du système de contrôle, notamment en intégrant le logiciel (SIL), le processeur (PIL) et le hardware dans la boucle (HIL). Dans ces phases, les lois de contrôle sont générées sous forme de codes intégrables, embarquées sur un microprocesseur, validées et calibrées sur le système réel ou à l'aide de simulateurs en temps réel.

1.3.2 Optimisation des systèmes du groupe motopropulseur

La réduction de la consommation de carburant, la conformité aux normes de contrôle de la pollution, les coûts de fabrication et d'autres défis industriels importants obligent Renault et d'autres acteurs de l'industrie automobile à développer des GMPs de plus en plus complexes. Sur ces GMPs, le nombre d'actionneurs, de briques technologiques et donc de combinaisons de commandes explose. Par exemple, sur les GMPs équipés de moteurs diesel, la combinaison des technologies moteur et post-traitement connaît une croissance exponentielle. Déterminer en début de cycle de développement la bonne definition technique du GMP parfaitement adapté aux normes et aux besoins des différents pays devient extrêmement difficile avec les méthodologies actuelles. C'est la raison pour laquelle l'optimisation des systèmes du GMP est d'une importance capitale.

Aujourd'hui, l'optimisation de la definition technique d'un GMP ne se fait plus sur des cycles stabilisés mais sur des cycles normatifs, comme le New European Driving Cycle (NEDC) et le Worldwide Harmonized Light Vehicles Test Cycle (WLTC), et sur des cycles de conduite réels, comme le cycle Real Driving Emissions (RDE). Elle est réalisée sur un système en boucle avec un contrôleur dont l'objectif est de suivre certaines consignes. L'optimiseur agira donc à la fois sur les macroprogrammes du GMP et sur les paramètres du contrôleur, notamment sur les cartographies permettant de définir les consignes des grandeurs à contrôler.

Du point de vue de l'automatique, cela nécessite deux éléments essentiels:

1. Un controlleur qui exécute des cycles de conduite sur le simulateur de GMP sur lequel les paramètres du controlleur et les cartographies de consignes peuvent être modifiés ;

2. Un simulateur de GMP permettant d'accéder le plus possible à la physique tout en assurant un temps de simulation acceptable pour l'optimisation des algorithmes.

Sur cette base, cette thèse développera des méthodologies et testera des algorithmes d'optimisation qui seront utilisés pour identifier et calibrer les modèles de systèmes et les paramètres de contrôleurs, qui sont susceptibles d'être intégrés dans les outils et processus de définir les definitions techniques du GMP tout en respectant diverses contraintes, telles que la performance énergétique et le respect des normes antipollution.

1.4   Objectif de la thèse

Après avoir placé cette thèse dans l'étape MIL de la conception du système de contrôle-commande à base de modèles, et après avoir énoncé son objectif général en matière d'optimisation des systèmes du GMP, nous précisons ici son objectif précis et les principaux axes de travail autour desquels s'articule le travail de cette thèse.

Aujourd'hui, les ingénieurs de Renault construisent les modèles de systèmes de GMPs sur des plates-formes de modélisation graphique et de simulation. Ces plates-formes sont largement utilisées par les constructeurs de moteurs car elles permettent une modélisation simple et rapide de modèles nonlinéaires précis. Du point de vue de l'ingénierie de l'Alliance Renault-Nissan, de nombreuses activités d'ingénierie sont liées à ces modèles, c'est pourquoi ils sont disponibles pour presque tous les moteurs de l'Alliance. Ces modèles sont souvent associés à la conception de commande à dérivée intégrale proportionnelle (PID), qui reste la technique de commande la plus utilisée dans l'industrie automobile en raison de sa simplicité et de sa capacité à traiter un large éventail de problèmes de commande. La linéarisation des modèles autour de chaque point de fonctionnement pourrait être un moyen de faciliter le réglage du régulateur PID, ou même de le remplacer par un régulateur linéaire. Cependant, cette solution n'est pas simple en raison de la complexité de la dynamique conduisant à des goulots d'étranglement numériques et informatiques. C'est pourquoi l'intérêt principal de Renault dans ce travail de recherche est de déterminer des procédures efficaces d'identification des modèles linéaires pour les systèmes du GMP en utilisant uniquement leurs entrées et sorties, chaque fois que cela est

possible.

Au-delà de l'identification et de la mise au point de modèles linéaires de systèmes du GMP, Renault s'intéresse également au problème des retards dans ces systèmes. Les retards sont très répandus dans les GMPs, et les ingénieurs du controle sont confrontés à ces retards dans de nombreuses applications de commande du GMP, telles que la commande EGR, la régulation du régime de ralenti (ISC) et la régulation du rapport air/carburant. Les retards, s'ils ne sont pas pris en compte dans la conception de la commande du GMP, peuvent avoir des effets négatifs sur ses performances. Pour Renault, il est essentiel de trouver des techniques fonctionnelles pour traiter ces retards dans ses futurs GMPs. C'est l'un des nombreux piliers qui permettront à Renault de maîtriser et d'exceller dans la conception du contrôle du GMP, qui est un élément clé pour maintenir et améliorer la qualité et la compétitivité de ses véhicules. Afin de répondre aux attentes de Renault dans cette thèse et d'apporter des solutions innovantes aux problèmes abordés, les travaux de thèse peuvent être structurés en trois parties décrites ci-dessous, où les développements théoriques attendus sont associés aux réalisations industrielles demandées.

1.5 Structure de la thèse

En plus de ce chapitre d'introduction, cette thèse comprend quatre autres chapitres. Notez que la notation utilisée dans chaque chapitre est limitée au chapitre lui-même et ne s'applique pas aux autres chapitres.

Chapitre 3: Identification des Modèles Boîte-Noire pour des Moteurs à Essence à l'Aide des Méthodes Subspace

Ce chapitre définit une méthodologie qui permet l'identification automatique dans le domaine temporel d'un modèle state-space (SS) linéaire d'un moteur essence, à partir d'un modèle à valeur moyenne nonlinéaire (MVEM). Sachant que chaque constructeur automobile propose une large gamme de moteurs et que les definitions techniques de chacun de ces moteurs sont en constante évolution, l'avantage fondamental de la méthodologie proposée est qu'elle permet de passer d'une definition technique à une autre sans avoir à définir une nouvelle technique d'identification du système linéaire à chaque changement. L'approche adoptée dans ce chapitre suggère la réalisation d'une identification linéaire du modèle SS dans le domaine temporel à l'aide de méthodes

subspace. Comme le modèle est une boîte noire, seules ses entrées et sorties sont accessibles. Il est excité en utilisant des signaux d'entrée avec des propriétés de corrélation et de fréquence adéquates, et les sorties correspondantes sont enregistrées. En utilisant cet ensemble d'entrées et de sorties, des méthodes subspace ainsi que des méthodes de minimisation des erreurs de prédiction sont utilisées pour estimer un modèle linéaire SS à chaque point de fonctionnement. Ce chapitre conclut sur l'ordre des modèles et les paramètres d'algorithme les plus pratiques pour l'application de ces méthodes d'identification sur les moteurs essence.

Chapitre 4: Identification et Commande du Modèle Linéaire du Papillon Electrique

Ce chapitre présente une méthodologie pour identifier les modèles SS et ARX linéaires à paramètres variants (LPV) d'un papillon électrique à partir d'un modèle nonlinéaire de ce dernier. Le modèle nonlinéaire, dont les paramètres de variation sont l'angle du papillon et sa vitesse, est d'abord écrit sous la forme d'un système LPV SS utilisant les équations physiques. Ce modèle donne un aperçu des éléments à utiliser dans le vecteur de régression, la deuxième étape étant d'identifier le système à l'aide d'une régression lineaire en minimisant un critère des moindres carrés. Pour contrôler les modèles linéaires, un régulateur quadratique linéaire (LQR) est conçu séparément pour chacun des systèmes LPV identifiés, équipé d'un feedforward de perturbation et d'un feedforward de référence pour faire face à la perturbation sur l'entrée existant dans les modèles à paramètres variables du papillon et au problème de suivi de référence, respectivement. Le régulateur est evalué sur le modèle nonlinéaire dont le modèle linéaire est extrait. Cette approche se révèle intéressante essentiellement grâce au modèle ARX, qui permet de contourner les modèles mathématiques SS dans les futures identifications des modèles linéaires de papillons électriques. La crédibilité du modèle ARX est vérifiée par rapport à un autre modèle nonlinéaire et par rapport à un banc d'essai de papillon électrique réel.

Chapitre 5: Identification et Contrôle des Modèles de Transport dans la Chaine d'Air du Moteur

Les phénomènes de transport sous leurs différentes formes sont abondants dans la chaine d'air du moteur. Ce chapitre aborde deux de ces phénomènes de transport, le transport d'énergie (chaleur) et le transport de masse, en présentant leurs modèles orientés commande et la conception de la commande du modèle de transport de masse. Pour dé-

montrer les phénomènes de transport de chaleur et de masse sur des composants la chaine d'air, l'échangeur de chaleur et l'EGR sont choisis respectivement. Le chapitre est donc divisé en deux parties:

- La première partie présente l'identification des flux advectifs présents dans l'échangeur de chaleur. Les phénomènes d'échange de chaleur qui se produisent dans les tuyaux d'un échangeur de chaleur peuvent être modélisés sous forme d'équations aux dérivées partielles (EDP) hyperboliques de premier ordre. La reformulation de ces équations sous la forme d'un système à retard préserve la dimensionnalité infinie du système, tout en diminuant sa complexité mathématique. Ce chapitre propose, à l'aide d'une technique de moyennage spatiale et de la méthode des caractéristiques, une modélisation des températures du flux dans la chaine d'air sous forme d'un système à retard. Une méthode d'optimisation par descente de gradient est proposée pour estimer les paramètres de ce système à retard, en utilisant des mesures de température sur les bords d'une section de la chaine d'air. L'intérêt de cette approche est mis en évidence par les données expérimentales obtenues sur un banc d'essai d'échangeur de chaleur.

- La deuxième partie présente le modèle orienté commande et la conception de commande du phénomène de transport du rapport des gaz brûlés (BGR), qui se produit dans le trajet d'admission d'un moteur à combustion interne (MCI), en raison de la réorientation des gaz brûlés vers le trajet d'admission par l'EGR basse pression. Inspiré du modèle à retard présenté dans la première partie du chapitre, et basé sur un modèle AMESim® nonlinéaire du moteur, le BGR du collecteur d'admission est modélisé comme un modèle SS à retard de sortie, ou bien comme un système couplé EDO-EDP, qui prend en compte le délai entre le moment où les gaz brûlés quittent le collecteur d'échappement et celui où ils sont réadmis dans le collecteur d'admission. En plus de leur retard de transport de masse, les BGRs sur le trajet d'admission sont également soumis à des contraintes d'inégalité d'état et d'entrée. L'objectif du problème de contrôle est de suivre un profil de sortie de référence du BGR dans le collecteur d'admission, en tenant compte du retard de transport et des contraintes d'état (sortie) et d'entrée du système. Dans ce but, deux approches de contrôle optimal indirecte sont mises en œuvre et comparées, l'approche discretize-then-optimize et l'approche optimize-then-discretize. Pour tenir compte des contraintes d'inégalité d'état, les deux méthodes sont

équipées de techniques d'optimisation sous contraintes telles que les méthodes Lagrangienne augmentée et UZAWA. Les conditions d'optimalité nécessaires sont formulées, dans chacun des deux cas, et les équations résultantes sont résolues numériquement à l'aide de la méthode de la descente du gradient projeté, qui assure la non-violation des contraintes d'inégalité sur l'entrée. La nouveauté du travail effectué dans ce chapitre réside dans la prise en compte des contraintes du système et de la dimensionnalité infinie du phénomène de transport de masse qui le régit. Les mérites du modèle à retard et de la conception de contrôle basée sur le modèle sont illustrés sur le modèle nonlinéaire AMESim® sur lequel le modèle mathématique est basé.

Chapitre 6: Conclusion et Perspectives d'Avenir

Le présent chapitre résume les travaux réalisés dans le cadre de cette thèse. Il indique les orientations des travaux futurs possibles et mentionne certains problèmes en suspens qui pourraient révéler des solutions intéressantes à long terme.

Context and Objective of the Thesis

Chapter Summary

2.1 Model-based systems engineering

"Model-based systems engineering (MBSE) is the formalized application of modeling to support system requirements, design, analysis, verification and validation activities, beginning in the conceptual design phase and continuing throughout development and later life cycle phases." [FGS07]

Although people have been using models in system engineering for decades, yet system models haven't sufficiently replaced documents in engineering processes and communications, so far. To push forward in this direction, the International Council on Systems Engineering (INCOSE) introduced, in 2007, the MBSE approach in a well-defined framework, accompanied by a roadmap which establishes a vision for the future of system engineering practices [FGS07].

Today, after around a decade of the INCOSE MBSE initiative, MBSE already applies to a considerable number of fields (electrical, mechanical, etc...), involves many sorts of models (static, time-varying, deterministic, stochastic, etc...) and modeling techniques (Computer Aided Design, Analytical modeling, etc...), and offers numerous advantages [Ber11], such as:

- improving the productivity and affordability of systems engineering;

- reducing the cost of system design and limiting its risks by verifying the system's performance and requirements early in the process;

- anticipating system changes and improvements before implementation, i.e. before the change becomes costly;

- ameliorating the quality and reliability of designed systems;

- providing multiple design alternatives by providing rapid feedback on design decisions;

- promoting model interoperability and integration throughout the system life cycle, which opens the door to model reuse and allows a unified design environment starting from the system requirements identification phase until the testing and validation phase.

2.2   The powertrain

The powertrain of an automobile is the set of components which transform the power of its engine into motion. Although the definition remains unchanged, whether the engine is an internal combustion engine (ICE), electric, hybrid, or of other type, the components of the powertrain may vary. In traditional vehicles powered by an ICE, the engine, transmission, driveshaft, differentials, axles, and whatever lies between the engine and the wheels, are the components that make up the powertrain.

The component of the powertrain that is exclusively visited in this thesis is the engine. In addition to considering the engine as a whole, this thesis also considers two sub-components of the engine: the electric throttle and the engine air-path, in particular the heat exchanger and

the exhaust gas recirculation (EGR). The following chapters will discuss these sub-components in further details.

2.3 Model-based systems engineering for powertrain systems optimization

Due to its obvious merits, MBSE was received by the automobile industry with open arms. The industry is realizing the need to transform the text-based engineering to model-based engineering, because the former is unable to cope with the ever-increasing complexity of the automotive systems and with the more and more stringent demands of governments and customers. From reduced emissions to fuel economy to superior performance, the current and future challenges are numerous, and the evolution in automotive development is imperative.

2.3.1 Model-based control system design

A key pillar in evolving automotive design concepts is model-based optimization and control. With control-oriented models at hand, model-based control design comes into the picture, and brings in a whole list of benefits [Mic+10], such as:

- the possibility to assess the closed-loop system's behavior and robustness in the beginning of the design cycle. Many possible control architectures can therefore be tested on the model before generating any software or building any hardware, which minimizes the engineering costs and design time;

- assistance in defining new engines' system requirements and specifications (engines' technical definitions (TDs), control requirements, calibration requirements, etc...), by offering the possibility of closed-loop system simulation in normative and real driving cycles;

- assistance in control design by providing a framework to verify the interaction between the plant model and the controller whether in the form of software code or once embedded on a microcontroller;

- reducing engine costs by allowing the integration of observer design.

The V-model, shown in Figure 2.1, is a graphical tool used to break-down the different stages of the control system design, from setting the requirements of the system and building its model until validating and calibrating the design on the intended hardware.
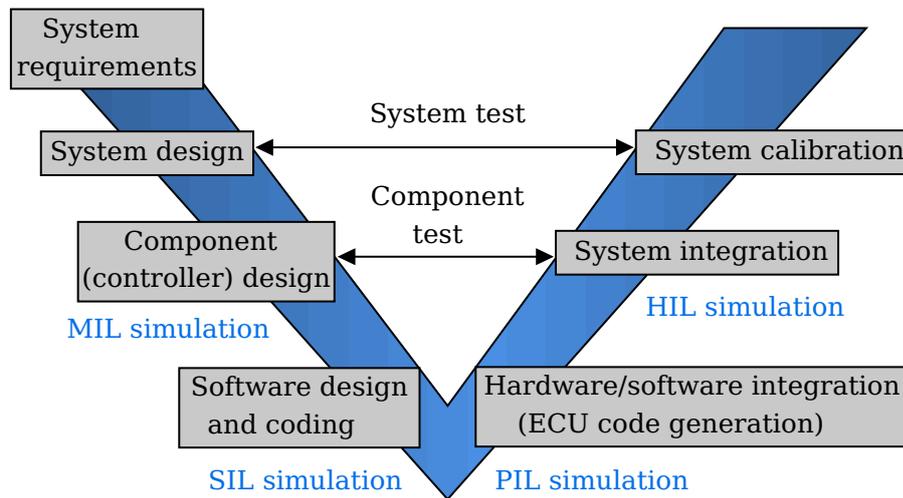
Figure 2.1: Control system development lifecycle

The work in this thesis takes place in the controller design phase. In this phase, also called the model-in-the-loop (MIL) phase, control-oriented plant models are simulated in closed-loop with their corresponding control designs. The resulting closed-loop systems are then validated and verified to meet the necessary requirements during real driving cycles. Following the MIL phase, the subsequent phases take over and complete the control system development process, notably by integrating the software (SIL), the processor (PIL), and the hardware in the loop (HIL). In these phases, the control laws are generated in the form of embeddable codes, embarked on a microprocessor, and validated and calibrated on the actual system or using real-time simulators.

2.3.2 Powertrain systems optimization

Reduced fuel consumption, compliance with pollution control standards, manufacturing costs, and other strong industrial challenges are forcing Renault and other actors in the automotive industry to develop increasingly complex powertrains. On these powertrains, the number of actuators, technological bricks and therefore control combinations explode. For example, on powertrains equipped with diesel engines, the combina-

tion of engine and post-treatment technologies is growing exponentially. Determining at the beginning of the development cycle, the right powertrain TD which is perfectly in line with the standards and needs of the different countries becomes extremely difficult with current methodologies. This is the reason why powertrain systems optimization is of core significance.

Nowadays, optimization of the TD of a powertrain is no longer done on stabilized but on normative cycles, such as the New European Driving Cycle (NEDC) and the Worldwide Harmonized Light Vehicles Test Cycle (WLTC), and on real driving cycles, such as Real Driving Emissions cycle (RDE). It is performed on a looped system with a controller whose objective is to follow certain setpoints. The optimizer will therefore act on both, the powertrain macro sizes and the controller parameters, especially the maps allowing to define the setpoints of the quantities to be controlled.

From an automatic control point of view, this prerequires two essential elements:

1. A controller which runs driving cycles on the powertrain simulator on which the controllers' parameters and the setpoint maps can be modified;

2. A powertrain simulator to access as much as possible of the physics while ensuring acceptable simulation time for optimization algorithms.

On this basis, this thesis will develop methodologies and test optimization algorithms that will be used to identify and calibrate system models and controller parameters, that are susceptible to be embedded in the tools and processes for defining powertrain TDs while respecting various constraints, such as energy performance and compliance with pollution control standards.

2.4 Objective of the thesis

After having placed this thesis in the MIL step of the model-based control system design, and having stated its general objective with regard to powertrain systems optimization, we hereby specify its precise objec-

tive and the principal work axes around which the work in this thesis revolves.

Today, Renault engineers build the powertrain systems models on graphical modeling and simulation platforms. These platforms are widely used by engine manufacturers because they allow easy and fast modeling of accurate nonlinear models. From the Renault-Nissan Alliance engineering perspective, many of the engineering activities are attached to these models, which is why they are available for almost all the Alliance engines. These models are often associated with proportional integral derivative (PID) control design, which is still the most widely used control technique in the automotive industry due to its simplicity and ability to handle a broad range of control problems. A way to facilitate the tuning of the PID controller, or even to replace it by a linear controller, could be the linearization of the models around each operating point. However, this solution is not straightforward due to the complexity of the dynamics leading to numerical and computational bottlenecks. Therefore, the main interest for Renault in this research work is to determine efficient linear model identification procedures for powertrain systems using only their inputs and outputs, whenever that is possible.

In addition to identifying and calibrating linear models of powertrain systems, another main interest for Renault in this work is tackling the problem of time-delays in powertrain systems. Time delays are widespread in powertrains, and control engineers confront these delays in many of the powertrain control applications, such as EGR control, idle speed control (ISC), and air-to-fuel ratio control. Time delays, if not accounted for in the powertrain control design, can have disadvantageous effects on the powertrain performance. For Renault, it is substantial to find functional techniques for treating these delays in its future powertrains. This is one of the many pillars that will enable Renault to master and excel in powertrain control design, which is a key element in maintaining and improving the quality and competitiveness of its future vehicles. In order to meet the expectations of Renault in this thesis and provide innovative solutions to the problems under consideration, the thesis work can be structured in the three parts described below, where the expected theoretical developments are associated with the requested industrial achievements.

2.5   Thesis structure

In addition to this introductory chapter, this thesis comprises four other chapters. Note that the notation used in each chapter is restricted to the chapter itself and doesn't carry on to other chapters.

Chapter 3: Black-box Gasoline Engine Identification using Subspace Methods

This chapter defines a methodology that permits automatic time-domain identification of a linear state-space (SS) model of a gasoline engine starting from a nonlinear mean-value engine model (MVEM). Knowing that each automobile manufacturer proposes a broad series of engines and that the TDs of each of these engines is ever-changing, the basic advantage of the proposed methodology is that it allows switching from one TD to another without having to define a new linear system identification technique with every switch. The approach adopted in this chapter suggests carrying out a linear SS model identification in time domain using subspace methods. Because the model is a black-box, only its inputs and outputs are accessible. It is excited using input signals with adequate correlation and frequency properties, and the corresponding outputs are registered. Using this set of inputs and outputs, subspace methods along with prediction error minimization methods are used to estimate a linear SS model at each operating point. This chapter concludes on the model order and the algorithm parameters most convenient for the application of these identification methods on gasoline engines.

Chapter 4: Electric Throttle Linear Model Identification and Control

This chapter presents a methodology for identifying SS and autoregressive with exogenous input (ARX) linear parameter-varying (LPV) models of an electric throttle starting from a nonlinear model of the latter. The nonlinear model, whose scheduling parameters are the angle of the throttle and its speed, is first written in the form of an LPV SS system using the physical equations. This model highlights the throttle's discontinuities and nonlinearities, and gives insight about the elements to be used in the regression vector, whereby the second step is to identify the system using a linear regression by minimizing a least squares criterion. To control the linear models, a linear quadratic regulator (LQR) controller is designed separately for each one of the identified LPV systems, equipped with disturbance and reference feedforwards to cope with the

input disturbance existing in the parameter-varying models of the throttle and with the reference tracking problem respectively. The regulator is evaluated on the nonlinear model from which the linear model was extracted. This approach reveals interesting basically due to the ARX model, which allows bypassing the SS mathematical models in future electric throttle linear model identifications. The credibility of the ARX model is verified with respect to another nonlinear model and with respect to a real electric throttle test-bench.

Chapter 5: Identification and Control of Transport Models in the Engine Air-path

Transport phenomena in their different forms are abundant in the engine air-path. This chapter touches on two of these transport phenomena, energy (heat) transport and mass transport, by presenting their control-oriented models and the control design of the mass transport model. To demonstrate the heat and mass transport phenomena on air-path components, the heat exchanger and the EGR are chosen respectively. Therefore, the chapter is divided in two parts:

- The first part presents the identification of advective flows present in the heat exchanger. The heat exchange phenomena occurring in the pipes of a heat exchanger can be modeled as first-order hyperbolic partial differential equations (PDEs). Reformulating these equations as a time-delay system preserves the infinite-dimensional property of the system, yet decreases its mathematical complexity. Using a space-averaging technique and the method of characteristics, this chapter proposes a time-delay system modeling of the flow temperatures in the air-path. A gradient-descent optimization method is proposed to estimate the parameters of this time-delay system, using boundary measurements of temperature in an air-path section. The interest of this approach is emphasized with experimental data obtained from a heat exchanger test-bench.

- The second part presents the control-oriented model and control design of the burned gas ratio (BGR) transport phenomenon, witnessed in the intake path of an ICE, due to the redirection of burned gases to the intake path by the low-pressure EGR. Inspired by the time-delay model presented in the first part of the chapter, and based on a nonlinear AMESim$^{®}$ model of the engine, the BGR in the intake manifold is modeled as a SS output time-delay model, or alternatively as an ODE-PDE coupled system, that take into account the time delay between the moment at which the combusted

gases leave the exhaust manifold and that at which they are read-mitted in the intake manifold. In addition to their mass transport delay, the BGRs in the intake path are also subject to state and in-put inequality constraints. The objective of the control problem is to track a reference output profile of the BGR in the intake man-ifold, taking into account the transport delay and the state (out-put) and input constraints of the system. In this aim, two indirect optimal control approaches are implemented and compared, the discretize-then-optimize approach and the optimize-then-discretize approach. To account for the state inequality constraints, both methods are equipped with techniques for constrained optimiza-tion such as the augmented Lagrangian and the UZAWA methods. The necessary conditions of optimality are formulated, in each of both cases, and the resulting equations are solved numerically using the projected gradient-descent method, which ensures the non-violation of the input inequality constraints. The novelty of the work done in this chapter lies in considering the system's con-straints and the infinite-dimensionality of the mass transport phe-nomenon governing it. The merits of the time-delay model and the model-based control design are illustrated on the nonlinear AMESim$^{®}$ model on which the mathematical model is based.

Chapter 6: Conclusions and Future Perspectives

This chapter summarizes the work done in this thesis. It points out directions of possible future work, and mentions some open problems that might reveal interesting to tackle in the long-term.

Chapter 3

Black-box Gasoline Engine Identification using Subspace Methods

Chapter Summary

The gasoline engine, also called petrol engine, is an ICE equipped with a spark-ignition system used to ignite gasoline in a cylinder containing a piston. The ignition of the fuel causes an explosive combustion inside the cylinder, thereby pushing the piston downwards and initiating the rotation of the crankshaft connected to it.

This engine dates back to 1876 when a German engineer named Nikolaus August Otto developed what is known as the "Otto cycle engine".

Nine years later, in 1885, Gottlieb Daimler reshaped this engine, which was originally intended for stationary use, and adapted it for transportation purposes [SK03].

An invention which revolutionized transportation back then, the gasoline engine still fits in the transportation scheme of nowadays. In addition to its satisfactory performance, lately, the international opinion rose up against diesel engines, which are usually considered more polluting than their gasoline analogues, leading to a more stringent economic policy. So as the popularity of the diesel engine declines, and the era of the full electric engines hardly begins, the gasoline engine reigns. This is why further study, such as this one, is being conducted on gasoline engines. In this study, the engine at hand is the Renault engine H5Ft400. It will be described later in further details.

A nonlinear MVEM of the H5Ft400 engine is built by Renault engineers on LMS Imagine.Lab AMESim®, an integrated simulation platform, which provides a graphical modeling interface. As previously mentioned in chapters 1 and 2, it is of considerable importance to by-pass the nonlinear models of the powertrain systems and find innovative methodologies for identifying linear models of these systems. This gives rise to the objective of this chapter, which is to find a methodology to identify a linear model of the H5Ft400 engine.

State of the art In the literature, identification of ICE linear models has been addressed in more than one reference, and different types of models and identification methods have been used. In [WDR07]; [WDR06]; [WRL08], the authors modeled a diesel engine in the form of LPV SS and LPV input-output (IO) models, whose parameters were identified using the recursive least-squares algorithm [WDR07] and using linear algorithms originating from classical identification methods [WDR07]. [WRL08] and [WDR07] used, as scheduling parameters for the LPV model, the engine speed and indicators of the load, such as the boost pressure [WRL08], due to the absence of the engine torque measurement. A SS representation was also employed by [Ser14] to model a partially premixed combustion (PPC) engine, whereas it used ARX and auto-regressive moving-average exogenous (ARMAX) representations to model a diesel engine. Whilst the above mentioned works used time-domain modeling, [KHO17] used frequency-domain modeling to model the diesel engine in the form of a transfer function and identify its parameters using nonparametric and parametric identification procedures.

Similar kinds of models and identification methods where used for gasoline engines, where [TAW12] utilized the least-squares approach to identify an LPV linear regression model of the engine, a model which is then transformed into the LPV SS form. [MS12] did the same transformation from LPV IO to LPV SS, but he used the downhill simplex method, which is a nonlinear iterative optimization method, for the identification of the LPV IO model. Both [TAW12] and [MS12] chose the engine speed and the intake manifold pressure as scheduling parameters for their LPV models. The gasoline engine was also modeled in the frequency-domain, where [ABMST94] used standard identification techniques to identify a transfer function of the model. Finally, opposing the current of black-box modeling, [EH14] modeled the gasoline engine as a physics-based mathematical nonlinear model.

Moreover, whilst [TAW12], which models only the engine intake manifold, limited the model input to the opening of the throttle valve , [EH14] and [MS12] added another input which is the opening of (or the mass flow through) the waste-gate. As for the output, [TAW12] and [MS12] selected the air charge of the cylinders, whereas the model presented by [EH14] output the boost pressure, intake and exhaust manifold pressures, and the turbocharger rotational speed.

Proposed method Regarding the H5Ft400 engine, whose scheme is shown in Figure 3.1, and which we consider from the air-chain point of view, a first reflection on the problem suggests basing the foreseen linear model of the engine on a mathematical model inspired from the physical laws governing the air-chain dynamics, i.e. analytical linearization of the nonlinear model presented in [EH14]. However, this linearization is not straightforward for different reasons. First of all, parameters such as the compressor and turbine flow rates, compressor and turbine isentropic efficiencies, the volumetric efficiency, and the amount of energy transferred to the exhaust gas are data maps, neither explicit nor smooth mathematical equations. Second, the mass flow rate through the throttle and the wastegate is a function of the valve's effective opening surface, which is highly nonlinear, and a function of the pressure ratio around the actuator which is a maximum function, i.e. an undifferentiable function which hinders the calculation of the Jacobian matrix, thus hindering the linearization process. As a consequence, this work direction is dropped, and the black-box identification approach is adopted.

The black-box identification approach is advantageous for numerous reasons. In addition to the fact that it avoids the limitations of the
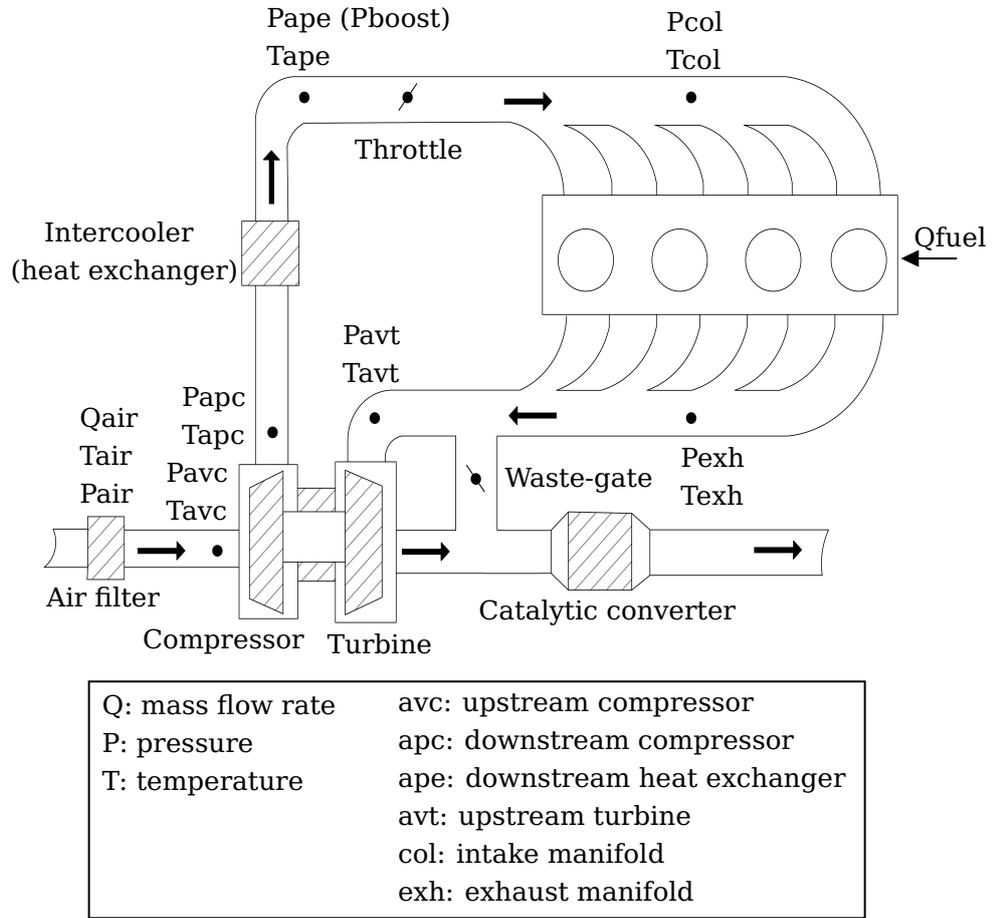
Figure 3.1: H5Ft400 engine scheme

linearization process, mentioned above, consider the case of a co-simulation between AMESim®, MATLAB®, and Simulink®, which will happen to be the case in this work. The presence of such a co-simulation environment deprives any solver from having access to the states of the AMESim® model which cannot be communicated through a Simulink s-function. This implies that the only accessible data are the inputs and the outputs of the system, and therefore gives rise to the objective of this chapter which is to identify a linear SS model of the H5Ft400 gasoline engine at each of its operating points, using only its inputs and outputs, i.e. identify a black-box linear model of the H5Ft400 engine.

Therefore, we formulate the identification problem as follows. Inspired by the LPV models in the literature, we propose to identify a discrete linear model of the H5Ft400 engine at each of its operating points be-

longing to the (engine speed, mean effective pressure) map. Because the engine torque is not measurable, it is represented, in the engine map, by the mean effective pressure which varies in direct proportionality with it, according to the equation [Hey+88]

$$PME = \frac{2\pi n_r T}{n_c V_d} \tag{3.1}$$

where PME is the mean effective pressure (Pa), n_r is the number of crankshaft rotations for a complete engine cycle, T is the torque (Nm), n_c is the number of cylinders, and V_d is the cylinder volume displacement (m^3). Note that the mean effective pressure intended in this case is the brake mean effective pressure, which is equal to the indicated mean effective pressure minus the friction mean effective pressure. In other words, it is the mean effective pressure that takes into account the engine efficiency and represents the actual output of the engine at the crankshaft. For the sake of simplicity, we denote brake mean effective pressure as PME, and we point out that, in this case, T in equation (3.1) is equal to the effective torque not the indicated torque. Moreover, for the H5Ft400 engine, $n_r = 2$ (because it is a 4-stroke engine), $n_c = 4$, and $V_d = 1200 \, \text{cm}^3$.

The openings of the throttle and waste-gate, represented by their pulse width modulations, denoted respectively as PWM$_{th}$ and PWM$_{wg}$ (th with respect to throttle; wg with respect to waste-gate), are chosen as inputs to the linear models. The intake manifold pressure P_{col} and the pressure after the heat exchanger considered as the boost pressure P_{boost}, are chosen as outputs, in addition to the pressure upstream the turbocharger P_{avt}, and the turbo speed W_t. Section 3.2 details this choice of inputs and outputs.

The input profiles are sent to the nonlinear MVEM gasoline engine model built on AMESim$^{\circledR}$, and the corresponding outputs are registered. Although they are simulated outputs, they play the role of measured outputs in this identification problem. In other words, it is with these outputs that we compare the eventual outputs of the identified linear model, and check the validity of the identification results.

Concerning the identification algorithm, the subspace method is considered for the estimation of the SS model matrices, due to the merits it offers. First of all, we do not need to know the structure of the model beforehand, which eliminates the need for any prior parameterization, and

it is suitable for multi-variable systems. To reinforce the subspace algorithm and improve the quality of the estimation, it is implemented in collaboration with the prediction-error minimization method [VV01]. The initial estimate of the SS system matrices is determined using the subspace identification method, before the prediction-error method takes over to further minimize the error between the output of the linear estimate and that of the nonlinear AMESim® model.

The novelty of this approach is that it investigates the use of the subspace methods in the identification of gasoline engine air-path. Although [BSJTJ06] solicited subspace methods for the identification of an homogeneous charge compression ignition (HCCI) engine, but generally speaking, subspace methods are still under-investigated in this domain.

We hereby detail the steps of the identification, which is carried out in the time-domain.

Step 1: Excitation of the AMESim® model using disturbed input signals (disturbed PWM_{th} and PWM_{wg}) and registration of the corresponding outputs (P_{col}, P_{boost}, P_{avt}, and W_t). Note that the disturbance used to perturb the inputs is a multi-sine wave [KHO17], chosen because it can cover a wide frequency range. Also note that the multi-sine signal disturbing PWM_{th} is of very low correlation with that disturbing PWM_{wg}.

Step 2: Identification of a discrete linear SS model at each of the engine's operating points, characterized by the speed and mean effective pressure. Recall that the identification algorithm is a combination of subspace methods and prediction-error minimization method.

Step 3: Once all the models at all the operating points are identified, they are subject to validation criteria to verify their quality. If they respect the validation criteria, the models are then validated on another set of input-output data.

Chapter structure This chapter is organized as follows. Section 3.1 describes the H5Ft engine AMESim® model and points out its important aspects. Section 3.2 defines the inputs and outputs of the linear model to be identified, and justifies their choice. Section 3.3 presents the identification's preparation phase, i.e. the excitation of the system using special disturbance signals. Section 3.4 briefs the subspace methods and details the identification process and its results. Consequently, section 3.5 evaluates the identification results, first by subjecting them to validation criteria, and second by validating the identified linear sys-

tem on a new set of data. Finally, section 3.6 concludes the chapter and points out its limitations and possibilities of future work.

3.1 H5FT400 nonlinear model

The H5Ft400 is a gasoline engine made by Renault. As shown in Figure 3.1, it has four cylinders with a displacement volume of $1200\,\mathrm{cm}^3$, turbocharger with waste gate, variable-valve-timing intake and exhaust systems, and a direct injection system.

Renault engineers built the nonlinear model of this engine, shown in Figure 3.2, on AMESim®. The model simulation is launched using MATLAB® which communicates with the AMESim® platform via an s-function of the adjoint Simulink® file. This structure results in a co-simulation environment between MATLAB®, Simulink®, and AMESim®.
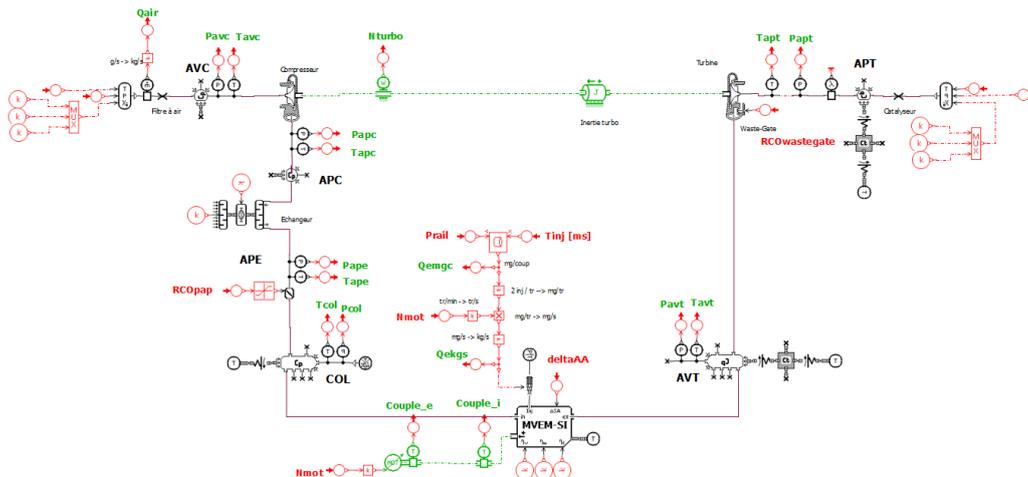
Figure 3.2: Nonlinear model of the H5Ft400 engine built on AMESim®

The engine model takes in numerous inputs, among which are the atmospheric temperature, atmospheric pressure, the variable-valve timings during intake and exhaust, rail pressure, injection time, spark advance, engine speed, and most importantly PWM_{wg} and PWM_{th}. On the other hand, it gives many outputs, among which are the pressures and temperatures at different positions of the engine, the air-fuel ratio, the torque, the turbo speed, and the fuel and air flows.

Two PID controllers are designed on Simulink® to regulate PWM_{wg} and

PWM_{th}. The first controller takes P_{boost} as feedback from the model, and regulates the value of PWM_{wg} while minimizing the error on P_{boost}, whereas the second one takes P_{col} as feedback, and outputs the desired value of PWM_{th} while minimizing the error on P_{col}. Note that the error is the difference between the output of the AMESim® model and the setpoint value of the concerned parameter, where the setpoint value is the value predefined by experts for each operating point in the (speed, PME) ISO field.
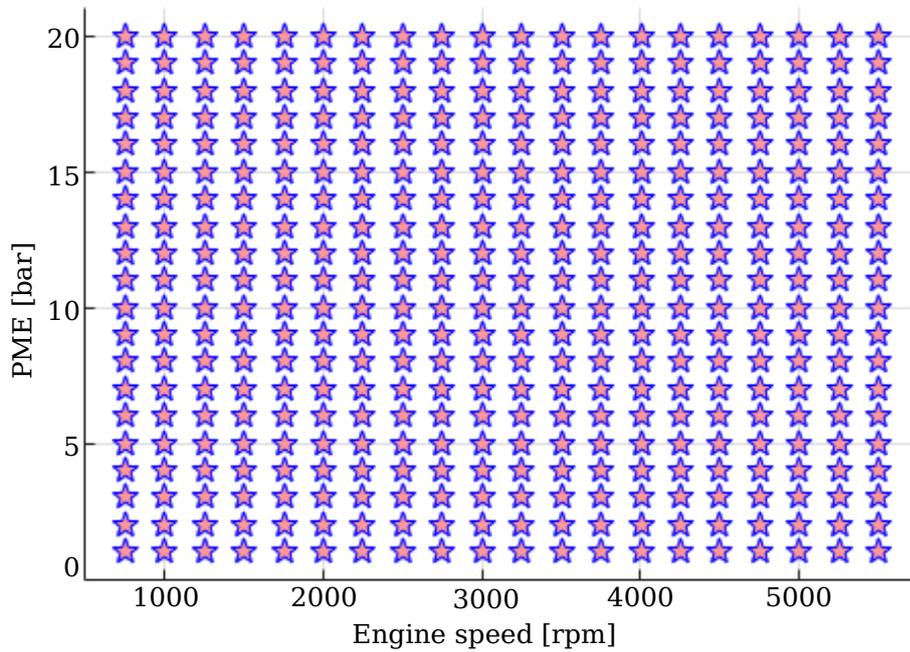
Figure 3.3: ISO field covering the engine operating range

An ISO field, shown in Figure 3.3, is a (speed, PME) cartography of around 300 points which covers the full functioning range of an engine. Generally speaking, the engine speed ranges from $750\,\mathrm{rpm}$ to $5500\,\mathrm{rpm}$ (with a step of $250\,\mathrm{rpm}$) and the PME ranges from $1\,\mathrm{bar}$ to $20\,\mathrm{bar}$ (with a step of $1\,\mathrm{bar}$). In the case of the H5FT400 engine, we consider the $4200\,\mathrm{rpm}$ speed range, to which corresponds a number of points whose PME ranges from low ($\approx -1\,\mathrm{bar}$) to high ($\approx 18\,\mathrm{bar}$).

3.2   Linear state-space model inputs and outputs

As mentioned earlier, the linear SS model declares PWM_{th} and PWM_{wg} as inputs, and P_{col}, P_{boost}, P_{avt}, and W_t as outputs. PWM_{th} and PWM_{wg} are selected as inputs because they are the control inputs when it comes to regulating the flows and pressures in the air-path of a turbocharged gasoline engine. PWM_{th} has a direct impact on P_{col} and PWM_{wg} has a direct impact on P_{boost}. The relevance of these variables for the control of the engine air-path justifies their choice for the identification problem.

The relationship between the inputs and the outputs can be depicted clearly if 2 functioning zones of the air-path are distinguished, for a particular engine speed.

- Zone of low engine torque: In this zone, the waste-gate is fully-open and hence uncontrolled, and the throttle is therefore the only actuator whose excitation can affect the air-path. Upon excitation of the throttle, P_{col}, which is less than $1\,\text{bar}$ in this case, is altered. Consequently, the cylinder filling, the temperature T_{avt} and pressure P_{avt} before the turbine, the turbo speed W_t, and the boost pressure P_{boost} are all affected. However, as P_{col} is relatively low, its effect on the other parameters and on the turbocharger is minimal, so we can consider that in this zone, the engine is on atmospheric functioning.

- Zone of high engine torque: In this zone, the throttle is fully-open and hence uncontrolled, and the waste-gate is therefore the only actuator whose excitation can affect the air-path. Upon excitation of the waste-gate, the temperature T_{avt} and pressure P_{avt} before the turbine are affected, which alters the turbo speed W_t, and eventually P_{boost}, and P_{col} which varies between $1\,\text{bar}$ and around $2.8\,\text{bar}$ in this case.

On one hand, this distinction shows that in order to well excite the system throughout all its operating range, both actuators, the throttle and the waste-gate, have to be triggered and subject to disturbance. On the other hand, this distinction points out that in the low-torque zone, P_{col} is the most important output which the estimated model should be able to reproduce, and the three other outputs are considered secondary, whereas in the high-torque zone, all four outputs P_{boost}, P_{col}, P_{avt}, and W_t are important and the estimated model should be able to reproduce

them all. This relative importance of the outputs is taken into consideration in the identification process described in section 3.4.2.

3.3   Experiment design

This section details the preparations that precede the identification phase.   It describes the nature of the generated disturbance signals, shows the correlation between the disturbance signal referring to PWM_{wg} and that referring to PWM_{th}, shows the spectral analysis done on the disturbance signals to verify their frequency content, and describes the construction of the disturbed input signals and the corresponding outputs of the system.

3.3.1   Disturbance signals

The major concern when generating an excitation disturbance signal for an identification problem is how persistently this signal excites the system to be identified. It has to be sufficiently rich to be able to generate informative input-output identification data. According to [Lju99], identifying a linear system requires an input signal with a customized input spectrum and a low crest factor. Moreover, the fact that the engine system admits two inputs adds an additional requirement, which is the input signal's ability to be generated in two sequences with minimum correlation. To meet these requirements, we survey a number of popular excitation signals, such as the Gaussian white noise, random binary signal (RBS), pseudo-random binary signal (PRBS), sine signal, square signal, triangular signal, saw-tooth signal, chirp signal (also called swept sinusoids), and multi-sines signal. Each of these signals presents certain advantages and disadvantages. We invite the reader to refer to [Lju99]; [VNJP17]; [ALMO12], among others, for detailed records on their definition, merits, and demerits.

Consider the Gaussian white noise, for example. On one hand, it uniformly solicits all the frequencies, but on the other hand, it has a high crest factor. On the contrary, RBS (Random Binary Signal) and the white noise-like PRBS, have a very low crest factor and result in uncorrelated sequences, but the RBS has an uncontrollable spectrum and the PRBS has to be generated in full length to reveal its nice frequency and correlation properties. Moving on to the sine signal, it is also inconvenient for

exciting the H5FT400 engine system because it solicits a unique fundamental frequency. Considering the sine-based signals, in particular the square, triangular, and saw-tooth signals, their spectral analysis shows a principal spike at their fundamental frequency and smaller spikes at harmonic frequencies, which are nothing but multiples of the fundamental frequency. So these signals are also an expression of a single frequency and are unsuitable inputs for the identification problem, which is why we continue investigating other signals, such as the chirp signal. The chirp signal is a very interesting signal because of its favorable features. First, the excited frequency band can be assigned by the user, and second, it has a low crest factor, as low as that of the sine signal which is equal to $\sqrt{2} = 1.414$. The major drawbacks of the chirp signal are: its power spectrum will not be null at frequencies falling outside the specified frequency band [Lju99] and it doesn't allow selecting particular frequencies within a frequency band. The energy is distributed almost uniformly over the whole band instead [ALMO12]. The signal which avoids this drawback is the multi-sines signal. The multi-sines signal is a combination of multiple sine waves with multiple frequencies. It can be customized to cover particular frequencies or wide and precise ranges of frequencies (frequency bands), and it has shown to ensure very good estimates of the system at the solicited frequencies. Another merit of the multi-sine signal is that the phase of the signal can be manipulated which gives the user control over the correlation between different generated signals. However, the multi-sine signal is not free of drawbacks as it is usually characterized with a high crest factor especially when all the summed sines are in phase with each other [Lju99]. Fortunately, there exists a solution for this problem which is the Schroader-phased multisine [Sch70]. It is a sum of sinusoids with variable frequencies and phases selected in a way that ensures a low crest factor. That being said, the multi-sines signal is compliant with the problem requirements, which is why it is chosen as the excitation input of the system. This input is denoted as $dist$ from "disturbance", and it is a function of time t:

$$dist(t) = \sum_{k=1}^{d} a_k cos(2\pi f_k t + \phi_k) \tag{3.2}$$

where d is the number of summed sine waves, a_k is the amplitude of sine wave k (%), f_k is the frequency of sine wave k (Hz), and ϕ_k is the phase of sine wave k (rad)

Two multi-sines disturbance signals, shown in Figure 3.4, are created: the first one is added to the steady-state value of PWM_{wg} and denoted by $dist_{wg}$, while the second one is added to the steady-state value of PWM_{th} and denoted by $dist_{th}$. $dist_{wg}$ and $dist_{th}$ are created with a frequency f ranging between $0.01\,\text{Hz}$ and $5\,\text{Hz}$, the former with a step of $0.01\,\text{Hz}$ and the latter with a step of $0.015\,\text{Hz}$. Hence, $d_{wg} = 500$ and $d_{th} = 333$.

The amplitude a of both disturbance signals is set to 3% before being divided by the maximum amplitude of the multi-sines, which guarantees that the signals are limited to an amplitude of 3%.

The phase ϕ of both disturbance signals is calculated as

$$\phi_k = \frac{-k(k-1)}{d}\pi \tag{3.3}$$

To verify that this phase ensures a low peak-to-peak amplitude, the crest factor is calculated, and $dist_{wg}$ is found to have a crest factor equal to 1.67 and $dist_{th}$ a crest factor equal to 1.88. Knowing that the lowest possible crest factor, corresponding to the PRBS signal with zero mean is equal to 1, and that corresponding to the sine signal is equal to 1.414, it can be noticed that the disturbances' crest factors are not high, which means that the Schroader phase attained its objective.

3.3.2 Disturbance signals correlation study

To improve the efficiency of the identification, the disturbed PWM_{wg} and PWM_{th} signals must have the minimum possible correlation. To study the correlation between $dist_{wg}$ and $dist_{th}$, the correlation coefficient r is used.

$$r(dist_{wg}, dist_{th}) = \frac{cov(dist_{wg}, dist_{th})}{std_{dist_{wg}} * std_{dist_{th}}} \tag{3.4}$$

where $cov(dist_{wg}, dist_{th})$ is the covariance of $dist_{wg}$ and $dist_{th}$, and $std_{dist_{wg}}$ and $std_{dist_{wg}}$ are the respective standard deviations of $dist_{wg}$ and $dist_{th}$.

This correlation coefficient reflects the linear correlation between the two disturbance signals, i.e. the correlation when the signals are not shifted from each other. $r(dist_{wg}, dist_{th})$ is a constant equal to 0.0175, which indicates that the two signals are almost not correlated.
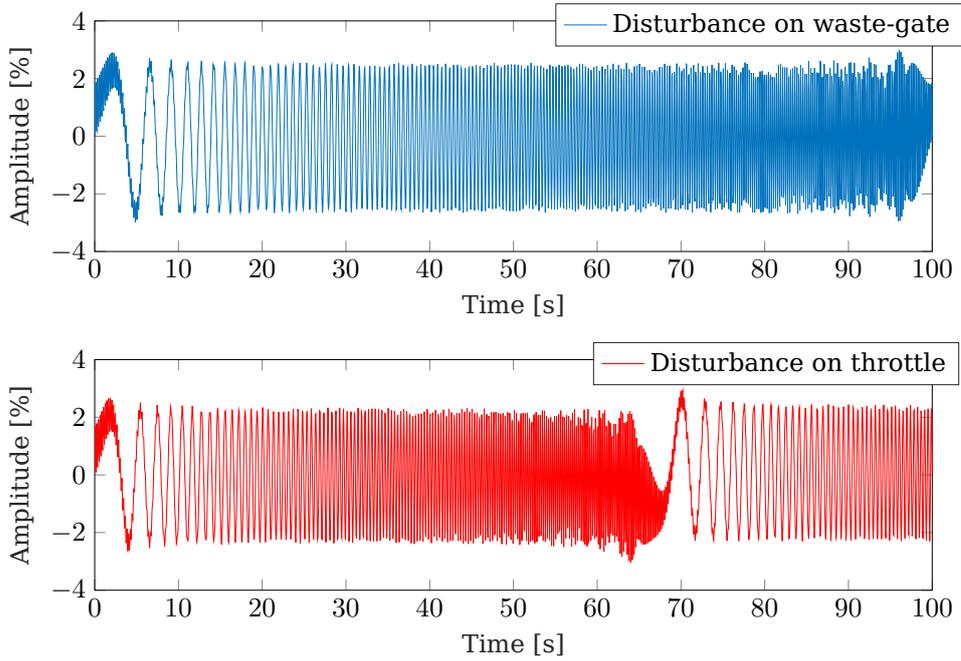
Figure 3.4: Disturbance signals added to PWM_{wg} and PWM_{th}

3.3.3 Disturbance signals spectral analysis

To verify the frequency content of the input signals, i.e. of the disturbances added to the input signals, a spectral analysis is carried out. Because the chosen disturbance is a multi-sines signal which is built manually, it is clear that the frequency range goes from $0.01\,\mathrm{Hz}$ up to $5\,\mathrm{Hz}$, as shown in Figure 3.5. However, for the sake of generality, the spectral analysis is presented. It is used, in case the disturbance signal is random or automatically generated, to verify its frequency content.

The spectral analysis is done by calculating the 1-sided power spectral density, denoted by PSD,

$$PSD = \frac{2|FFT(signal)|^2}{L_s F_s} \tag{3.5}$$

where the signal is either $dist_{wg}$ or $dist_{th}$, L_s is its length, $F_s = \frac{1}{T_s}$ is its sampling frequency, T_s is its sampling time equal to $0.001\,\mathrm{s}$, and FFT is its Fast Fourier transform, which is a vector of complex numbers whose magnitude is considered here. Note that the power spectral density is measured in $power/Hz = amplitude^2/Hz$, where the amplitude of the
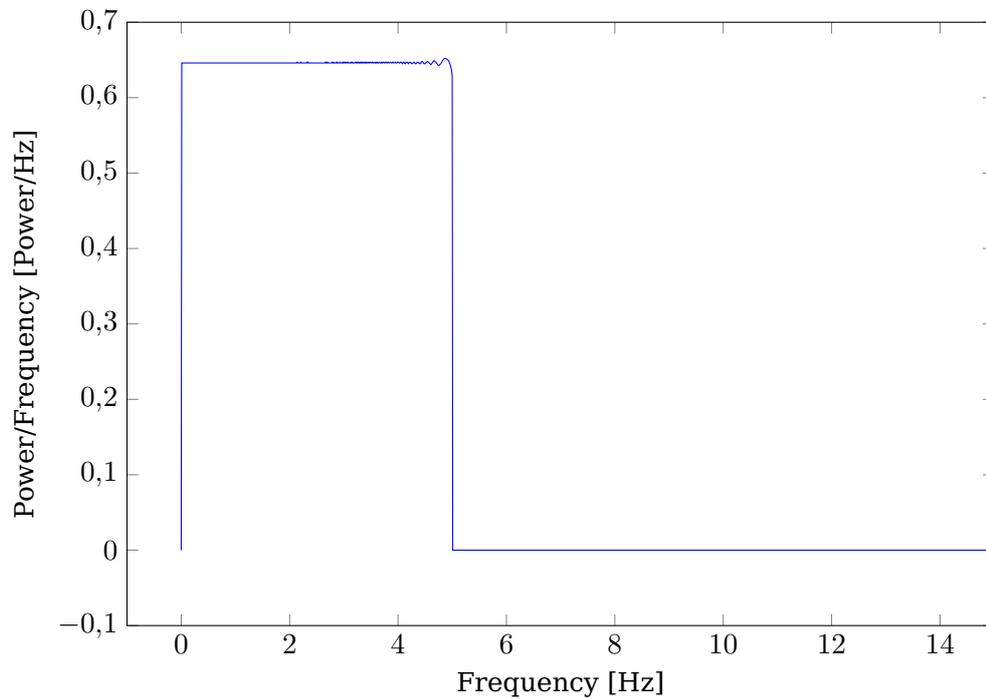
Figure 3.5: Spectrum of the disturbance signals

disturbance signal is measured in %, because it represents the PWM of the throttle and waste-gate valves.

3.3.4   Disturbed inputs and outputs

The disturbed inputs are formed as follows. The 2 input signals PWM_{wg} and PWM_{th} are stabilized by their corresponding PID controllers. Once they become stable, after approximately 10 seconds, their final values are fixed, the PID controllers stop working, and the respective disturbances $dist_{wg}$ and $dist_{th}$ are added to the signals. From that moment on, the system runs in open-loop. The four outputs of the engine model are collected by the end of the simulation. Figure 3.6 shows a sample of the simulation inputs and outputs at operating point $(2000\,\text{rpm}, 11.35\,\text{bar})$.
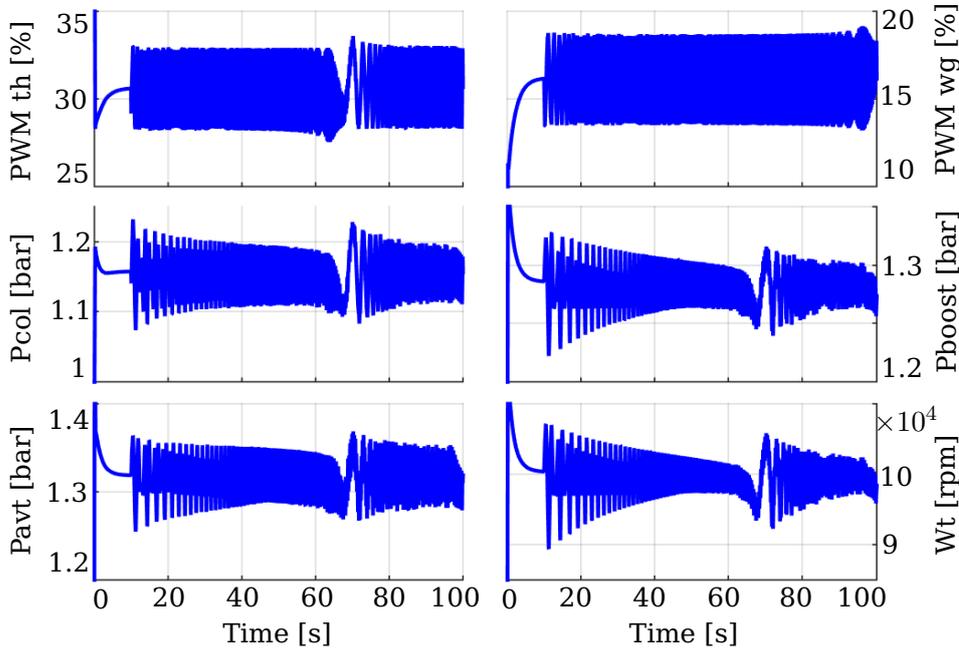
Figure 3.6: Inputs and outputs at operating point $(2000\,\mathrm{rpm}, 11.35\,\mathrm{bar})$
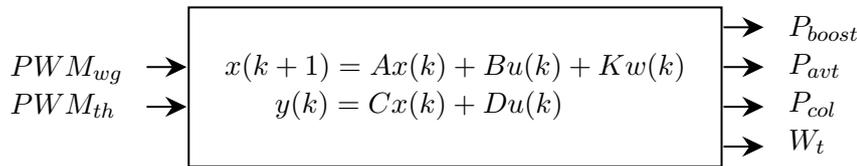
Figure 3.7: Discrete black-box SS model

3.4 State-space linear model identification

In this section, we identify, at each operating point (speed, PME), a discrete black-box SS model whose inputs and outputs are shown in Figure 3.7, where A, B, C, D and K are the SS system matrices, $x(k)$ is the system state, $u(k) = \begin{bmatrix} \mathrm{PWM}_{wg} & \mathrm{PWM}_{th} \end{bmatrix}^T$ is the system input, $y(k) = \begin{bmatrix} P_{boost} & P_{avt} & P_{col} & W_t \end{bmatrix}^T$ is the system output, $w(k)$ is the process noise, and $t = kT_s$ is the time sampled with a time step $T_s = 1\,\mathrm{ms}$, using $k \in \mathbb{N}_0$. The following subsections give a briefing about the subspace algorithm, describe the identification procedure, and detail the identification results.

3.4.1 Subspace methods

For the sake of completeness, and to be able to later specify the parameters of the identification algorithm, the basic steps of the subspace method [Lju99] are summed up briefly down below.

Consider a SS model whose matrices are to be estimated:

$$
\begin{aligned}
x(t+1) &= Ax(t) + Bu(t) + Kw(t) \\
y(t) &= Cx(t) + Du(t) + v(t)
\end{aligned}
\tag{3.6}
$$

where A, B, C, D, and K are the SS system matrices, $x(t)$ is the system state of dimension n (where n is the order of the system), $u(t)$ is the system input, $y(t)$ is the system output of dimension p, $w(t)$ and $v(t)$ are the respective process and measurement noise, and t is the time. The identification is carried according to the following procedure:

1. Using input-output data, form a matrix $G = \frac{1}{N}\mathbb{Y}\pi_{\mathbb{U}^T}^{\perp}\phi^T$ where $\mathbb{U}$ is a vector of inputs, $\mathbb{Y}$ is a vector of outputs, $\pi_{\mathbb{U}^T}^{\perp}$ is the orthogonal projection of $\mathbb{U}$, ϕ is a matrix defined such that the measurement noise is eliminated, and N is a constant used for normalization.

2. From G, which is a $pr \times n^*$ dimensional matrix (where r is the maximal prediction horizon and $n^* = n$ in case of known system order and $n^* \geq n$ in case of unknown system order), construct $\widehat{G} = W_1 G W_2$ where W_1 and W_2 are weighting matrices to be selected. In order to get the rank n of G, it is subjected to a singular-value decomposition $\widehat{G} = ZSV^T$, where Z and V are orthonormal matrices. Only the most significant singular values in S are retained while the rest are set to 0, which results in a new decomposition $\widehat{G} = Z_1 S_1 V_1^T$ in another basis. In other words, the rank n of G is the number of nonzero Eigen values of S.

3. Form the observability matrix $\widehat{\mathcal{O}}_r = W_1^{-1} Z_1 R$ where R is an arbitrary full rank matrix that determines the coordinate basis for the state representation. Knowing that $O_r = \begin{bmatrix} C & CA & \dots & CA^{r-1} \end{bmatrix}^T$, determine the estimates $\widehat{A}$ and $\widehat{C}$ of the SS system matrices, such that $\widehat{C} = \widehat{\mathcal{O}}_r(1:p, 1:n)$ and $\widehat{\mathcal{O}}_r(p+1:pr, 1:n) = \mathcal{O}_r(1:p(r-1), 1:n)\widehat{A}$. The notation used for matrix indexing is such that a matrix $F(m:n, o:p)$ is a matrix obtained from matrix F by extracting from it rows $m, m+1, ..., n$ and columns $o, o+1, ..., p$.

4. After calculating $\widehat{A}$ and $\widehat{C}$, the estimated matrices of the SS system $\widehat{B}$ and $\widehat{D}$, in addition to the estimated initial state $\widehat{x_0}$, are calculated by solving a linear regression problem

$$\underset{B,D,x_0}{\arg\min} \frac{1}{N} \sum_{t=1}^{N} \|y(t) - \widehat{C}(qI - \widehat{A})^{-1}Bu(t) - Du(t) - \widehat{C}(qI - \widehat{A})^{-1}x_0\delta(t)\|^2$$

This problem aims at minimizing the least-squares criterion of the error $y - \widehat{y}$ where $\widehat{y} = \widehat{C}(qI - \widehat{A})^{-1}Bu(t) + Du(t) + \widehat{C}(qI - \widehat{A})^{-1}x_0\delta(t)$ is the estimated output, $\widehat{C}(qI - \widehat{A})^{-1}x_0\delta(t)$ is the term corresponding to the initial state x_0, q is a shift operator, I is an identity matrix, and $\delta(t)$ is the unit pulse at time 0.

An additional step could be added to form the noise model, i.e. calculate $w(t) = \widehat{x}(t+1) - \widehat{A}\widehat{x}(t) - \widehat{B}u(t)$ and $v(t) = y(t) - \widehat{C}\widehat{x}(t) - \widehat{D}u(t)$

3.4.2 Identification algorithm

The sequential procedure used for the identification of the discrete-time linear SS model of the gasoline engine air-path is elaborated below.

1. Choose the operating point (speed, PME), where speed = $\{2000\,\mathrm{rpm}, 4200\,\mathrm{rpm}, 6000\,\mathrm{rpm}\}$ and PME corresponds to the mean effective pressure (between $-1\,\mathrm{bar}$ and $18\,\mathrm{bar}$).

2. Define the sampling time $Ts = 0.001\,\mathrm{s}$, the input vector $u = \begin{bmatrix} \mathrm{PWM}_{wg} & \mathrm{PWM}_{th} \end{bmatrix}$ and the output vector $y = \begin{bmatrix} P_{boost} & P_{avt} & P_{col} & W_t \cdot 10^{-5} \end{bmatrix}$.

3. Specify the order of the system (4, 6, or 8): the minimum order is set to 4 because the system has four outputs.

4. Specify the options of the MATLAB® solver "ssest", which is a combination of subspace and prediction-error minimization methods. The subspace method described in subsection 3.4.1 includes different parameters whose values play an important role in its efficiency and quality of its estimation. The MATLAB® function "ssestOptions" enables the user to alter the algorithm's modifiable parameters. While the default values of some parameters are retained, other parameters, shown in Table 3.1, are modified.

5. Estimate the SS system matrices using the "ssest" identification method of MATLAB$^{®}$. This method implements, as a first step, the non-iterative subspace algorithm to provide an initial estimate of the SS system. Consequently, as a second step, it implements the prediction-error minimization algorithm, which launches an iterative search to refine the black-box estimation.

6. Evaluate the goodness of fit between the estimated and the simulated system using the normalized root mean square error (NRMSE) of MATLAB$^{®}$, whose values vary between $-\inf$ (bad fit) to 1 (perfect fit). Note that when P_{col} is less than $1\,\mathrm{bar}$, it is the only output whose fit is taken into consideration because in this case, the waste-gate is fully-open and hence not controlled, which means that the output P_{boost} is of little importance. Otherwise, the fits of the two basic outputs, the P_{col} and P_{boost}, in addition to those of P_{avt} and W_t, are also taken into consideration.

7. Repeat steps 3 to 6 until all the possible combinations (system order - solver options), shown in Table 3.2, are tested.

8. After testing all 108 combinations (36 combinations of solver options $\times$ 3 model orders (4, 6, and 8)), determine that which yields the best fit and check the stability of the estimated models.

• Disturbance model: For systems in time-domain, the disturbance parameter is not necessarily null, which is why the user can choose or not to estimate the K matrix in equation (3.6). This parameter is modified in MATLAB$^{®}$ using the option "DisturbanceModel" which has two different versions: "none", which implies $K = 0$, and "estimate", which implies that the K matrix is a free parameter to be estimated.

• Iterative numerical search method: This option specifies the numerical method used to estimate the parameters of the SS model. It is modified in MATLAB$^{®}$ using the option "SearchMethod". Although a particular search method can be chosen to perform the iterative search, but we choose to specify it as "auto", which sequentially employs different line search methods at each iteration, and continues in the first descent direction which yields the lowest estimation cost. The line search methods are implemented in the following order: the subspace Gauss-Newton least squares, the Levenberg-Marquardt least squares, the adaptive subspace Gauss-Newton, and the steepest descent least squares. Although we don't switch manually among the numerical methods, but we mention this option in Table 3.2 to remind the reader that the search switches among different numerical methods, and doesn't rely on a single one.

• The weighting matrices W_1 and W_2: Different algorithms such as MOESP [Ver94], N4SID [VODM94], CVA [Lar90], IVM [Vib95], etc... use different choices of weights.

$\ast$ MOESP $W_1 = I$ $W_2 = (\frac{1}{N}\phi\pi_{\mathbb{U}^T}^{\perp}\phi^T)^{-1}\phi\pi_{\mathbb{U}^T}^{\perp}$

$\ast$ CVA $W_1 = (\frac{1}{N}\mathbb{Y}\pi_{\mathbb{U}^T}^{\perp}\mathbb{Y})^{-\frac{1}{2}}$ $W_2 = (\frac{1}{N}\phi\pi_{\mathbb{U}^T}^{\perp}\phi^T)^{-\frac{1}{2}}$

$\ast$ N4SID $W_1 = I$ $W_2 = (\frac{1}{N}\phi\pi_{\mathbb{U}^T}^{\perp}\phi^T)^{-1}\phi$

This parameter is modified in MATLAB$^{®}$ using the option "N4Weight" which has four different versions: "MOESP", "CVA", "SSARX" [Jan03], and "auto: automatic choice between MOESP and CVA". Despite the existence of a unifying theorem for MOESP, N4SID, and CVA [VODM95], nevertheless the choice of the weighting matrices is not straightforward, which makes them one of the simulation options to be shuffled and tested one by one.

● Initial state of system: This parameter is modified in MATLAB$^{®}$ using the option "InitialState" which can take six different forms, among which three are shuffled to estimate the best initial state: "zero" where the initial state is set to zero, "estimate" where the initial state is estimated as any other parameter, and "backcast" where the initial state is estimated using the best least squares fit.

● Focus: It determines how to estimate the error between the measured and the modeled outputs, hence acting as a weighting filter in the loss function. It is modified in MATLAB$^{®}$ using the option "Focus", which can take two forms. The first form is called "simulation" where the solver focuses on finding a good fit on the current simulation step, uses the current input to generate the model response, and minimizes the simulation error between measured and simulated outputs during estimation. The second form is called "stability" where the solver focuses on finding a good fit on the following simulation step and minimizes the prediction error between measured and predicted outputs. Note that we use MATLAB2013b$^{®}$, where the "stability" option is the same as the "prediction" option, available in MATLAB2018b$^{®}$, but with stability enforcement.

Table 3.1: Options of the "ssest" algorithm

3.4.3 Identification results

From the major objective of this work, which is to bypass the nonlinear AMESim$^{®}$ model of the engine and identify automatically a black-box linear model of the engine, the minor objectives of the identification process are deduced and listed below.

- Verify the adequacy and effectiveness of the experiment design;

- test different model orders and find that which best suits each operating point or, if possible, that which suits all operating points;

- test different combinations of solver options and find that which best identifies the model parameters at each operating point or, if possible, that which can identify reliable models at all operating points;

Combination	1	2	3	4	5	6	7	8	9
Disturbance	none	none	none	none	none	none	none	none	none
Initial state	zero	zero	zero	zero	zero	zero	estimate	estimate	estimate
N4Weight	MOESP	MOESP	CVA	CVA	SSARX	SSARX	MOESP	MOESP	CVA
Focus	simulation	stability	simulation	stability	simulation	stability	simulation	stability	simulation
Search method	auto	auto	auto	auto	auto	auto	auto	auto	auto
Combination	10	11	12	13	14	15	16	17	18
Disturbance	none	none	none	none	none	none	none	none	none
Initial state	estimate	estimate	estimate	backcast	backcast	backcast	backcast	backcast	backcast
N4Weight	CVA	SSARX	SSARX	MOESP	MOESP	CVA	CVA	SSARX	SSARX
Focus	stability	simulation	stability	simulation	stability	simulation	stability	simulation	stability
Search method	auto	auto	auto	auto	auto	auto	auto	auto	auto
Combination	19	20	21	22	23	24	25	26	27
Disturbance	estimate	estimate	estimate	estimate	estimate	estimate	estimate	estimate	estimate
Initial state	zero	zero	zero	zero	zero	zero	estimate	estimate	estimate
N4Weight	MOESP	MOESP	CVA	CVA	SSARX	SSARX	MOESP	MOESP	CVA
Focus	simulation	stability	simulation	stability	simulation	stability	simulation	stability	simulation
Search method	auto	auto	auto	auto	auto	auto	auto	auto	auto
Combination	28	29	30	31	32	33	34	35	36
Disturbance	estimate	estimate	estimate	estimate	estimate	estimate	estimate	estimate	estimate
Initial state	estimate	estimate	estimate	backcast	backcast	backcast	backcast	backcast	backcast
N4Weight	CVA	SSARX	SSARX	MOESP	MOESP	CVA	CVA	SSARX	SSARX
Focus	stability	simulation	stability	simulation	stability	simulation	stability	simulation	stability
Search method	auto	auto	auto	auto	auto	auto	auto	auto	auto

Table 3.2: Combinations of the solver options

PME	Model order					
	4		6		8	
	NRMSE	#	NRMSE	#	NRMSE	#
−1	0.86	31	0.88	35	**0.89**	**13**
0.86	0.78	13	0.79	9	0.86	17
2.21	0.95	9	0.92	33	**0.97**	**17**
3.54	0.85	9	0.89	17	**0.91**	**11**
4.78	**0.95**	**9**	0.93	35	0.94	33
6.30	**0.89**	**7**	0.88	17	0.81	11
7.98	0.65	13	**0.83**	**17**	0.81	9
9.74	0.64	34	0.65	35	**0.66**	**35**
11.76	**0.60**	**9**	0.55	11	0.45	11
13.5	0.34	35	**0.4**	**35**	−0.38	17
15.3	0.49	9	**0.59**	**33**	0.50	17
16.84	0.27	35	0.18	9	**0.47**	**9**
17.84	0.44	11	0.42	17	**0.45**	**33**
18.51	**0.44**	**35**	0.28	11	0.39	9
18.98	0.49	17	0.42	7	**0.76**	**11**
18.94	0.49	15	0.42	7	**0.73**	**11**

Table 3.3: Identification results at $4200\,\mathrm{rpm}$

- study the degree of similarity between the identified models, especially those corresponding to adjacent operating points.

For a global view of the results discussed in this subsection, we refer to Table 3.3. It presents the operating points considered in the identification process. They are points in the $4200\,\mathrm{rpm}$ speed range whose PME ranges from $-1\,\mathrm{bar}$ to $18.94\,\mathrm{bar}$. They are good representatives of the engine in both zones, the low-torque zone (from $-1\,\mathrm{bar}$ to $9.74\,\mathrm{bar}$) and the high-torque zone (from $11.76\,\mathrm{bar}$ to $18.94\,\mathrm{bar}$). This table also shows the NRMSE of the identified models with respect to their PME versus three model orders (4, 6, and 8), and shows the corresponding number of the chosen combination (#), from Table 3.2. The bold cells highlighted in light grey correspond to the highest NRMSE for a particular operating point, therefore pointing out its best model order and the best combination which can be used for its identification.

Note that the NRMSE values in Table 3.3 are calculated as follows. In the low-torque zone, they are calculated according to the level of fit of P_{col}, as it is the only output of concern in this zone, whereas in the high-

	Min	Max	Mean	Standard deviation
P_{boost} [bar]	0.93	2.18	1.39	0.06
P_{col} [bar]	0.15	2.15	1.14	0.06
P_{avt} [bar]	1	3.44	1.75	0.07
W_t [rpm]	626.97	226060	108790	8034.2

Table 3.4: Statistical properties of the outputs at $4200\,\mathrm{rpm}$

torque zone, they are calculated according to the levels of fit of all four outputs, P_{col}, P_{boost}, P_{avt}, and W_t. Table 3.4 is given as an indication on the statistical properties of the four outputs in the $4200\,\mathrm{rpm}$ speed range, which might be useful for controller tuning, in case the presented model is to be exploited for control purposes.

3.4.3.1 Adequacy of the experiment design

Starting off with the first minor objective, we inspect the NRMSE of the fit at each operating point and present a selection of graphical results. The better the level of fit between the outputs simulated by the AMESim® model and those generated by the identified model, the better the adequacy and effectiveness of the experiment design. The following results will show that the excitation process has variant effects on the outputs depending on whether its a low-torque or a high-torque zone, and leads to better identification results in the low-torque zone where P_{col} is the major concerned output.

Table 3.3 shows the levels of fit of P_{col} at a set of operating points in the low-torque zone ($-1\,\mathrm{bar} \leq PME \leq 9.74\,\mathrm{bar}$), and Figure 3.8 shows a sample of the graphical results in this zone, corresponding to $PME = -1\,\mathrm{bar}$. These results show high fits between the simulated and the estimated output P_{col}, with an average NRMSE of 0.87, calculated from the bold cells highlighted in light grey.

The fact that the estimated outputs show such a high fit with the simulated outputs in terms of tendency and amplitude, indicates, in addition to the efficiency of the estimation technique, that the input PWM$_{th}$ well-excites the system. This points out the good frequency coverage of this input signal, and therefore the adequate choice of the multi-sines as an excitation input. Notice also, from Figure 3.6, the linear relationship between the input PWM$_{th}$ and the output P_{col}, which also justifies why the linear model is well-suited for the data in the low-torque zone.
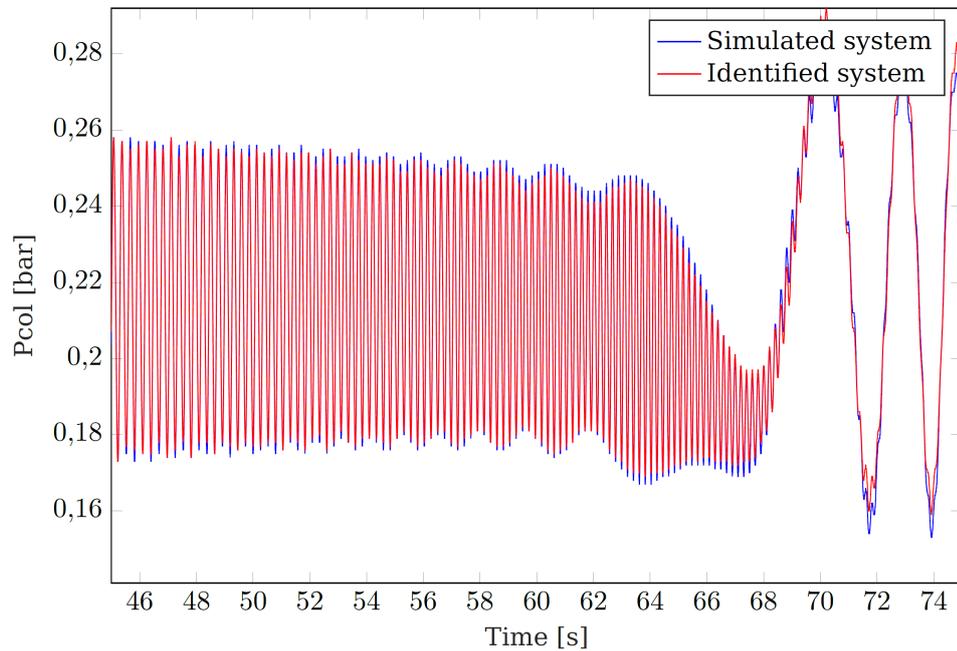
Figure 3.8: Identification results at operating point $(4200\,\mathrm{rpm}, -1\,\mathrm{bar})$

On the other hand, Table 3.3 shows the levels of fit of the four outputs at a set of operating points in the high-torque zone ($11.76\,\mathrm{bar} \leq PME \leq 18.9401\,\mathrm{bar}$), and Figure 3.9 shows a sample of the graphical results in this zone, corresponding to $PME = 11.76\,\mathrm{bar}$. In contrast to the results in the $2000\,\mathrm{rpm}$ speed range, which present an average NRMSE equal to 0.74 in the high-torque zone and good fits in terms of amplitude as well as tendency (see for example Figure 3.10), these results in the $4200\,\mathrm{rpm}$ speed range present a lower fit between the simulated and the estimated outputs, with an average NRMSE of 0.55, calculated from the bold cells highlighted in light grey. Also, though the tendencies of the simulated outputs are still followed by the estimated outputs, except that their incongruous amplitudes shed light on the difficulty of isolating the effects of the 2 inputs, PWM_{th} and P_{col}, on the outputs. That is to say, contrary to the case when $P_{col} \leq 1\,\mathrm{bar}$ where the throttle is the only valve in charge of the air flow, here the throttle and waste-gate are both active and are simultaneously affecting the four outputs, which renders the excitation process more complicated.
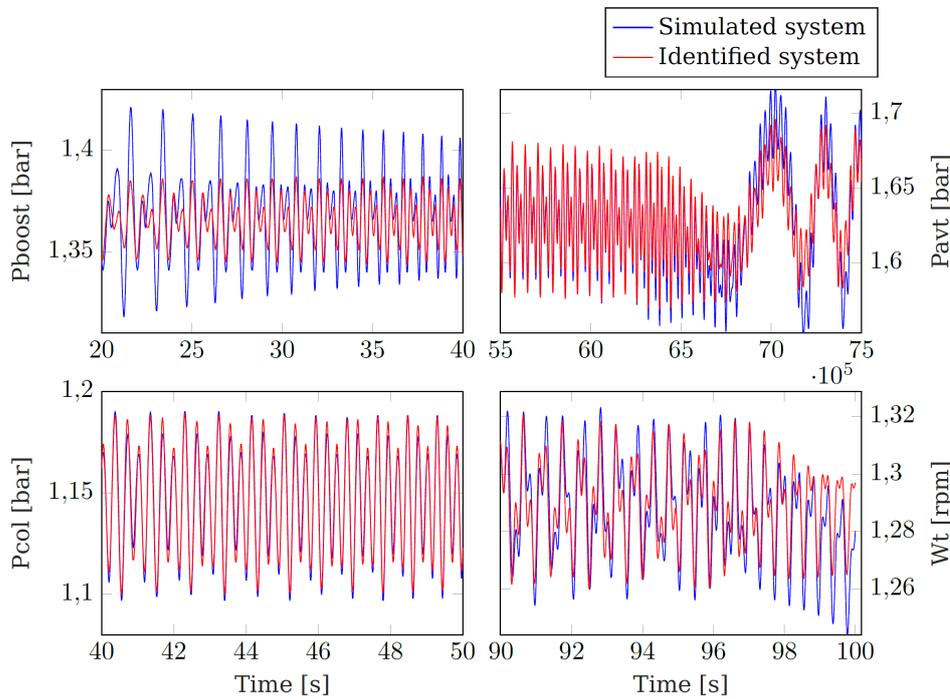
Figure 3.9: Identification results at operating point $(4200\,\mathrm{rpm}, 11.76\,\mathrm{bar})$

3.4.3.2 Model orders

Moving on to the second minor objective of the identification, three model orders were tested, 4, 6, and 8, as shown in Table 3.3. The bold cells highlighted in light grey correspond to the highest NRMSE for a particular operating point, and therefore correspond to the chosen order for the estimated system. For a more global view of the results, Table 3.5 shows the repartition (in %) of model orders of estimated systems in the same speed range. It shows that 25% of the estimated systems are of order 4, 18.75% are of order 6, and 56.25% are of order 8. From these 2 tables, the following conclusions can be drawn out:

- Although order 8 is able to best represent the biggest number of models, nevertheless lower orders such as 4 and 6 prove, for some operating points, to be better representatives.

- There is no relationship between the model order and the position of the operating point in the ISO field. For a certain engine speed, operating points with low or high PME might have system models with order 4, 6, or 8.
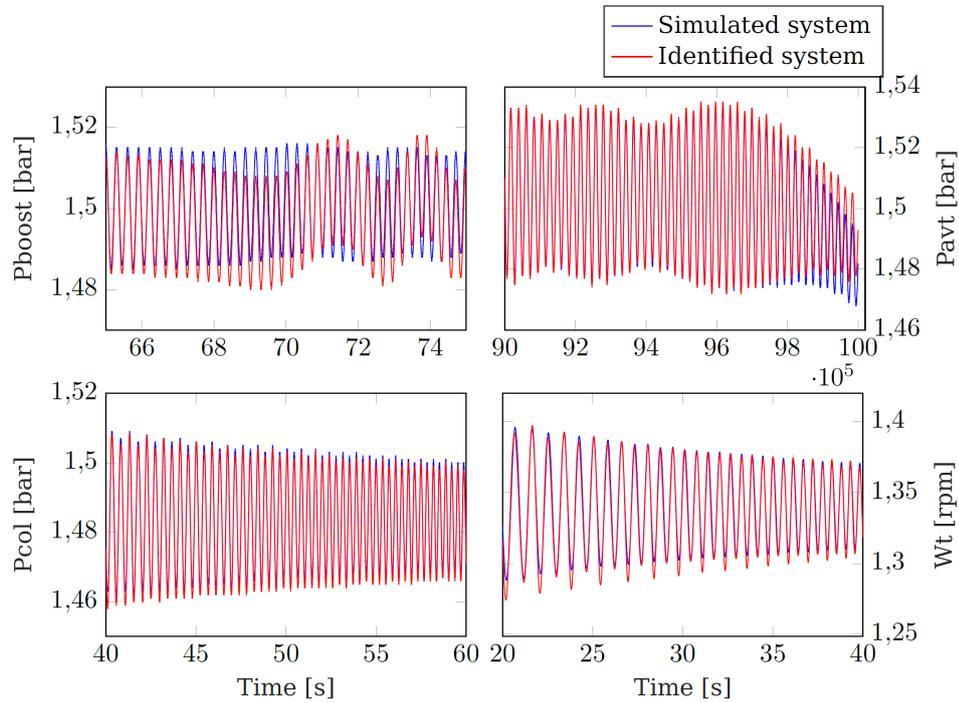
Figure 3.10: Identification results at operating point $(2000\,\mathrm{rpm}, 14.87\,\mathrm{bar})$

Model order	Percentage
4	25
6	18.75
8	56.25

Table 3.5: Repartition of model orders of estimated systems at $4200\,\mathrm{rpm}$

- Generally speaking, the impact of the model order is less evident on low-torque points ($-1 \leq PME \leq 9.74$) than on high-torque points ($11.76 \leq PME \leq 18.94$), where changing the order causes a considerable change in the % of fit.

These observations lead to the conclusion that there is no single preferable model order to represent the engine on all operating points. And in case we would like to impose a unique order for a group of operating points, this is more likely to work on low-torque points than on high-torque points.

3.4.3.3 Combinations of solver options

In addition to testing different model orders, the identification algorithm also tested 36 different combinations of solver options, shown in Table 3.2, in an aim to find that or those which yield the best models. The bold cells highlighted in light grey correspond to the highest NRMSE for a particular operating point, and therefore correspond to the best combination for this point. For a more global view of the results, Table 3.6 shows the repartition (in %) of the best solver options combinations for operating points in the same speed range. It shows that 6.25% of the estimated systems were identified using the solver options combination 7, 18.75% were identified using the solver options combination 9, and so on.

A few remarks can be made in this context, regarding the solver options that were part of the best combinations and those that were never solicited.

- Both forms of the option "Disturbance", "none" and "estimate", were used.

- The "estimate" and "backcast" forms of the option "InitialState" were used, but "zero" was not.

- The weighting forms "MOESP", "CVA", and "SSARX" of the option "N4Weight" were all used.

- The "Focus" option only took the "simulation" form, never the "stability" form.

As to whether there is a relationship between the order of the model and the chosen combinations, Table 3.3 shows that the answer is no, because different combinations are chosen for different operating points under the same order. Even considering the best combinations (bold and highlighted in light grey), it can be noticed that operating points with different model orders might share the same combination.

From the side of the PME, the relationship between the PME and the chosen combinations is also unclear. For most of the operating points, three different combinations are chosen for different order models.

That being said, we conclude that there is no one best combination of solver options that is always able to estimate the best model on all oper-

ating points. However, we can still notice that, out of 36 tested combinations, only 7 were actually used. This conclusion is time-saving, because the number of combinations that need to be tested, for three model orders, in future black-box identifications of linear SS models of a gasoline engine air-path, reduces from 108 combinations to 21 combinations, thus shrinking the testing time by 81%. In this context, we would like to point out the importance of testing different combinations in the first identification attempt, because testing more combinations increases the chance of finding those which are able to identify the models with the best fits. This can be shown clearly in Table 3.7, which shows that the more the tested combinations, the better the average NRMSE of the estimated models.

Combination	Percentage
7	6.25
9	18.75
11	18.75
13	6.25
17	18.75
33	12.5
35	18.75

Table 3.6: Repartition of solver options combinations at $4200\,\mathrm{rpm}$

Number of tested combinations	8	9	36
NRMSE	0.27	0.42	0.71

Table 3.7: Effect of testing different number of combinations

3.4.3.4 Degree of similarity between the identified models

The degree of similarity between the identified models is another aspect which merits being inspected, because it shows how close from each other the identified models are, and therefore to what extent it is possible to group these models in batches, where each batch groups a number of operating points that share the same model. This grouping is particularly interesting in the control phase which usually follows the identification.

To study the similarity between models, it is intuitive to start by comparing the matrices A, B, C, and D of the identified SS systems. However, comparing the system matrices among different operating points is not

straightforward for two reasons. The major reason is because the identified models are not necessarily of the same order. Some are of order 4, while others are of orders 6 and 8. The second reason is because, even for the same model order, identifying the SS systems, without setting the same C matrix for all operating points, would result in systems with different states. For example, $A(1,1)$ in system 1 would not necessarily correspond to $A(1,1)$ in system 2 even if both systems have the same order. To find other means of comparison between models and to prove that the same system can have different SS matrices, a well-known example, the DC motor, is chosen for demonstration.
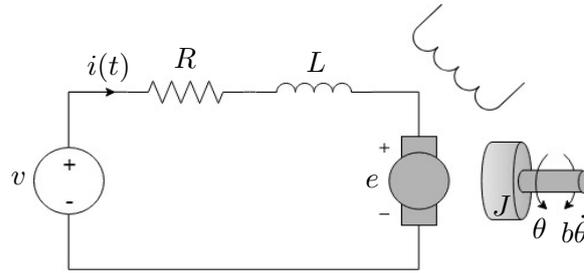
Figure 3.11: DC motor scheme

The DC motor, whose scheme is shown in Figure 3.11, is a combination of a voltage source giving a voltage V (V), a resistance R (ω), an inductance L (H), and a rotor of inertia J ($\mathrm{kg\,m}^{-2}$) propelled by an electromotive force e, which is subject to a magnetic field. The electromotive force e is represented by M which is the electromotive force constant or the motor torque constant ($\mathrm{V\,rad}^{-1}\,\mathrm{s}^{-1}$), and the motor viscous friction is represented by a constant b ($\mathrm{Nm\,s}$). Let θ be the rotational angle of the rotor (rad), $\dot{\theta}$ be its rotational speed ($\mathrm{rad\,s}^{-1}$) and i be the current (A). Based on Newton's 2nd law and Kirchhoff's voltage law, this system can be described in the following SS representation:

$$\frac{d}{dt}\begin{bmatrix} \dot{\theta} \\ i \end{bmatrix} = \begin{bmatrix} -\frac{b}{J} & \frac{M}{J} \\ -\frac{M}{L} & -\frac{R}{L} \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ i \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{L} \end{bmatrix} V \tag{3.7}$$

$$z = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \dot{\theta} \\ i \end{bmatrix} \tag{3.8}$$

Considering Table 3.8, showing the numerical values of the DC motor parameters, the SS system matrices become

R	$1\,\omega$	b	$0.1\,\mathrm{Nm\,s}$
L	$0.5\,\mathrm{H}$	M	$0.01\,\mathrm{V\,rad^{-1}\,s^{-1}}$
J	$0.01\,\mathrm{kg\,m^{-2}}$		

Table 3.8: Numerical values of the DC motor parameters

$$A_{sim} = \begin{bmatrix} -10 & 1 \\ -0.02 & -2 \end{bmatrix} \quad B_{sim} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad C_{sim} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad D_{sim} = 0$$

Let $SS_{sim} = (A_{sim},\ B_{sim},\ C_{sim},\ D_{sim})$ be the simulated SS model. Its input V and output $z = \dot{\theta}$ are collected and given to the "ssest" estimator which estimates a new SS system denoted by $SS_{est} = (A_{est}, B_{est},\ C_{est},\ D_{est})$.

$$A_{est} = \begin{bmatrix} -12.22 & 1,064 \\ -14.41 & -0,4883 \end{bmatrix} \quad B_{est} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad C_{est} = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad D_{est} = 0$$

Note that $B_{est} = B_{sim}$, $C_{est} = C_{sim}$, and $D_{est} = D_{sim}$, and this was imposed in the options of the solver. Even under such constraints, the solver will not estimate the same A_{est} matrix as A_{sim}.

Nevertheless, Figure 3.12 shows that the output generated by the simulated model and that generated by the identified model overlap perfectly, despite the inequality of the SS systems. To understand the origin of the difference between the simulated and identified systems matrices, we conduct different tests, where one test parameter is altered at a time. Below is a nonexclusive list of some parameters which, once changed, change the estimated system matrices with them.

- System simulator: For example, if we build the DC motor system in the form of physical equations on MATLAB® and build another one graphically on AMESim® with the same parameters, then extract the input-output data from each of them, and give them to the estimator, the estimated systems will have the same output but different matrices.

- Sampling time/discretization scheme: For example, consider a unique DC motor mathematical system built on MATLAB®. If we extract two sets of input-output data from this system, such that the only difference between them is their sampling time/discretization scheme, and use each of these sets to estimate
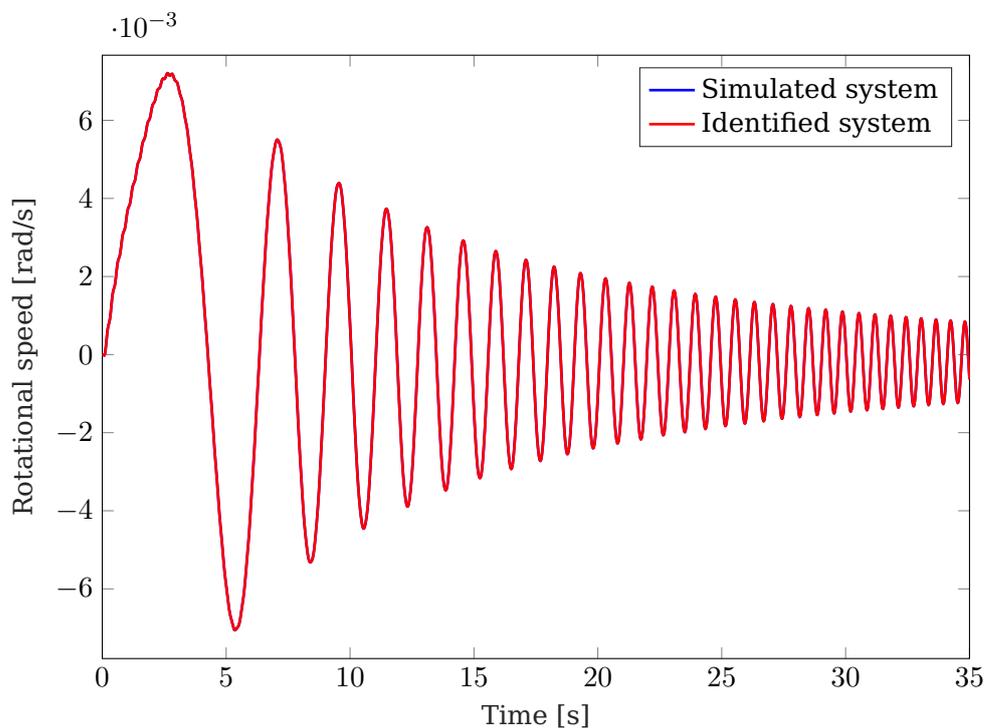
Figure 3.12: Rotational speed of the DC motor

a SS system, the estimated systems will again have the same output but different matrices.

- Prediction-error minimization algorithms: As mentioned in subsection 3.4.2, the subspace method's basic role is to provide an initial estimate of the system, which will in turn be the starting point of the iterative prediction-error minimization algorithm. It is possible, however, to specify an initial system manually. But even if we specify the original mathematical system as an initial system, the prediction-error minimization algorithm will estimate another SS system, with same output but different SS matrices.

- Weighting schemes and other subspace solver options: Changing these options for the same initial system, sampling time, discretization scheme, and prediction-error minimization solver will change the estimated SS matrices, though for the same output.

These different factors point out the difficulty of estimating models with exact same matrices, which makes model comparison a cumbersome task. Nevertheless, if the simulated model and the identified model of

the DC motor do not share the same SS matrices, they do share how-
ever the same Eigen values, -10 and -2, and the same DC gain, 0.0999.
This observation is expected, because the Eigen values are the system's
characteristic roots and the DC gain is the value of the transfer func-
tion, relating the inputs to the outputs of the system, evaluated as time
tends to infinity. This means that they are good representatives of the
dynamics governing the system and its input-output response at steady-
state. This implies that if two systems behave the same way, although
their matrices may be different, their Eigen values and DC gains will be
the same. Therefore, Eigen values and DC gains can be used to confirm
similarity between SS models whose matrices are not equal.

Going back to the identification of the gasoline engine, the Eigen values
and DC gains of systems, at different operating points in the $4200\,\mathrm{rpm}$
speed range, are shown respectively in Figures 3.13 and 3.14.

Based on these figures, the comparison of the Eigen values and DC gains
leads to the following observations:

- No system coincides with another system in all its Eigen values or
  its DC gains, even those belonging to adjacent operating points;

- a couple of operating points such as ($16.8354\,\mathrm{bar}$ - $17.8407\,\mathrm{bar}$)
  or ($4.7852\,\mathrm{bar}$ - $18.5058\,\mathrm{bar}$) coincide in some Eigen values, and
  a couple others such as ($2.2110\,\mathrm{bar}$ - $16.8354\,\mathrm{bar}$) or ($3.5382\,\mathrm{bar}$ -
  $18.9401\,\mathrm{bar}$) coincide in some DC gains, but except for the points
  ($16.8354\,\mathrm{bar}$ - $17.8407\,\mathrm{bar}$), these operating points are not adjacent
  points in the ISO field;

- operating points which coincide in some Eigen values are not the
  same as those which coincide in DC gains;

- even points which perfectly coincide with other points in some
  Eigen values and DC gains have other Eigen values and DC gains
  which are completely isolated from other points.

These observations point out the fact that the degree of similarity be-
tween the identified SS models is almost null, which means that group-
ing models in sorts of batches is not possible at this stage.

Finally, to study the stability of the identified discrete-time systems,
the magnitudes of the Eigen values, shown in Figure 3.13, are plot-
ted against the unit circle, as shown in Figure 3.15. The fact that all
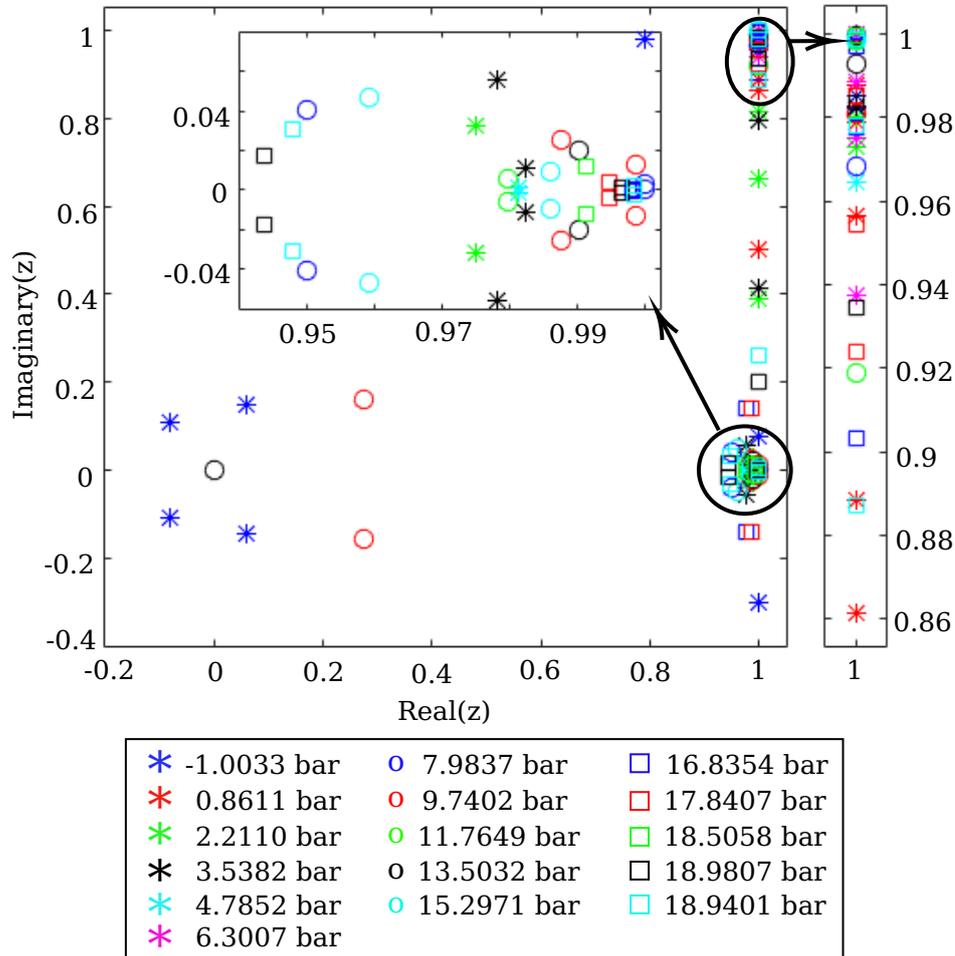
Figure 3.13: Eigen values of identified systems at $4200\,\mathrm{rpm}$

the magnitudes fall outside the unit circle is an indicator of the systems' instability. This instability makes sense because the input-output data used for identification are collected from the open-loop AMESim® model, which is only stable on certain operating points, but exhibits exponential divergence on others. Especially in the high-torque zone, if the waste-gate is not controlled, the engine model will experience unstable behavior. Note that the AMESim® system is excited in open-loop, because according to [Lju99], if the data given to subspace methods is acquired in closed-loop, they will produce biased estimates. Therefore, to avoid this bias, we perform an open-loop identification while bearing in mind the need to complement the identified system with a stabilizing
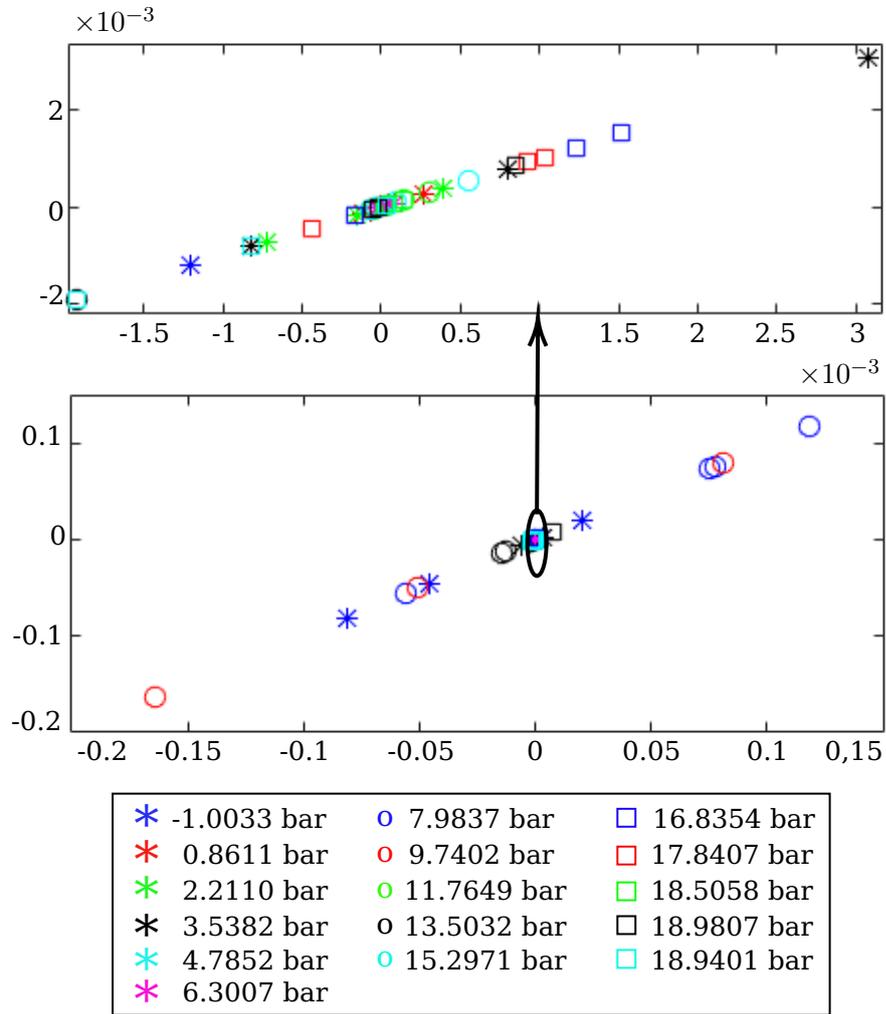
Figure 3.14: DC gains of identified systems at $4200\,\mathrm{rpm}$

controller.

3.5 State-space linear model validation

3.5.1 Validation criteria

The SS system identified using the identification algorithm detailed earlier has four outputs: P_{col}, P_{boost}, P_{avt}, and W_t. To verify that the identified SS system produces outputs that fit well the simulated outputs
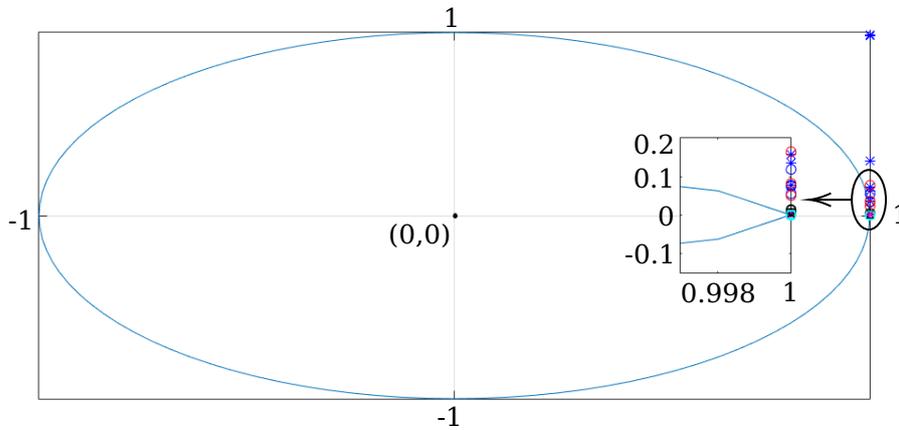
Figure 3.15: Stability analysis of identified systems at $4200\,\mathrm{rpm}$

issued by the nonlinear AMESim® model, the margin of allowed error between the identified SS system outputs and the simulated outputs is defined and the admissible zone is shown in yellow in Figure 3.16, where y represents the identified output and x represents the simulated output.

For all four outputs, the margin of error is limited to 5% when the pressures P_{col}, P_{boost}, and P_{avt} are greater than or equal to $50\,\mathrm{mbar}$, or when the turbo speed is greater than or equal to $5000\,\mathrm{rpm}$. In this case, the admissible zone is located between the two straight lines $y = 1.05x$ and $y = 0.95x$.

When the pressures are less than $50\,\mathrm{mbar}$, or when the turbo speed is less than $5000\,\mathrm{rpm}$, the margin of error is maintained at 5%, but the 5% in this case is the admissible error corresponding to $50\,\mathrm{mbar}$ or $5000\,\mathrm{rpm}$ respectively. Concerning the pressures, the admissible zone is hence located between the two straight lines $y = x + 0.00238$ and $y = x - 0.0025$. As for the turbo speed, the admissible zone is located between the two straight lines $y = x + 238.0952$ and $y = x - 250$.

Figure 3.17 shows that all four outputs of the identified models for operating points belonging to $4200\,\mathrm{rpm}$ speeds respect the validation criteria depicted in Figure 3.16, which makes these models eligible for validation on another set of data.
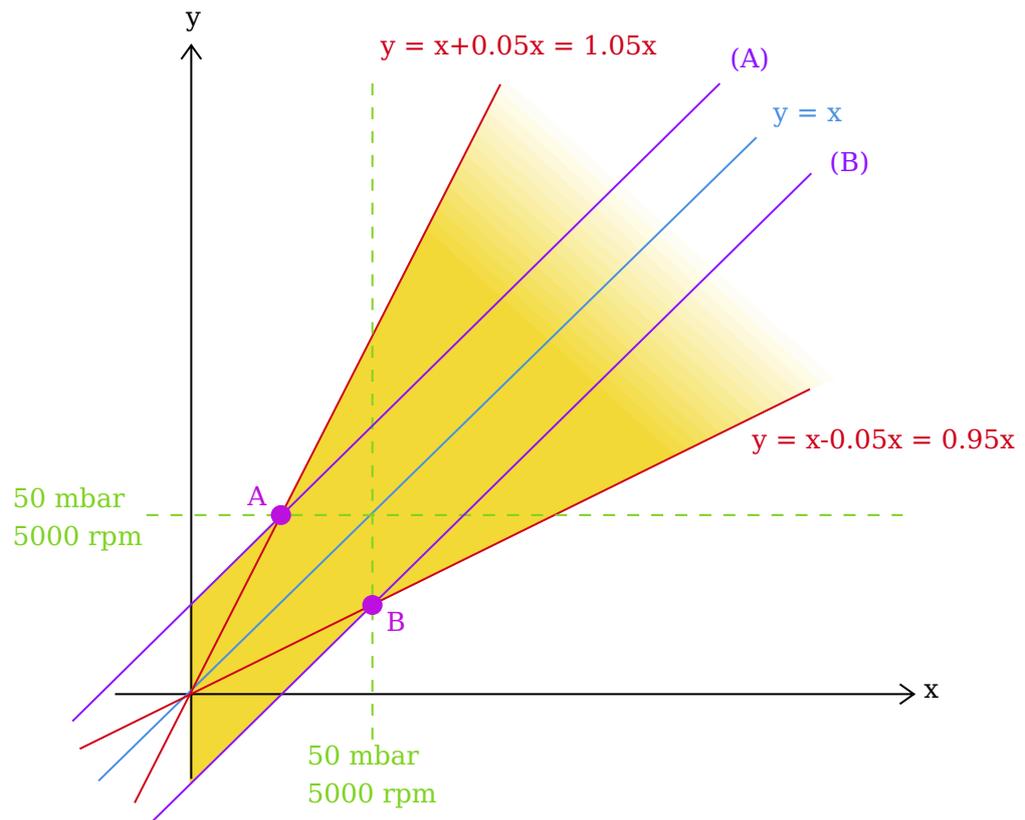
Figure 3.16: Criteria used for validation of identification results
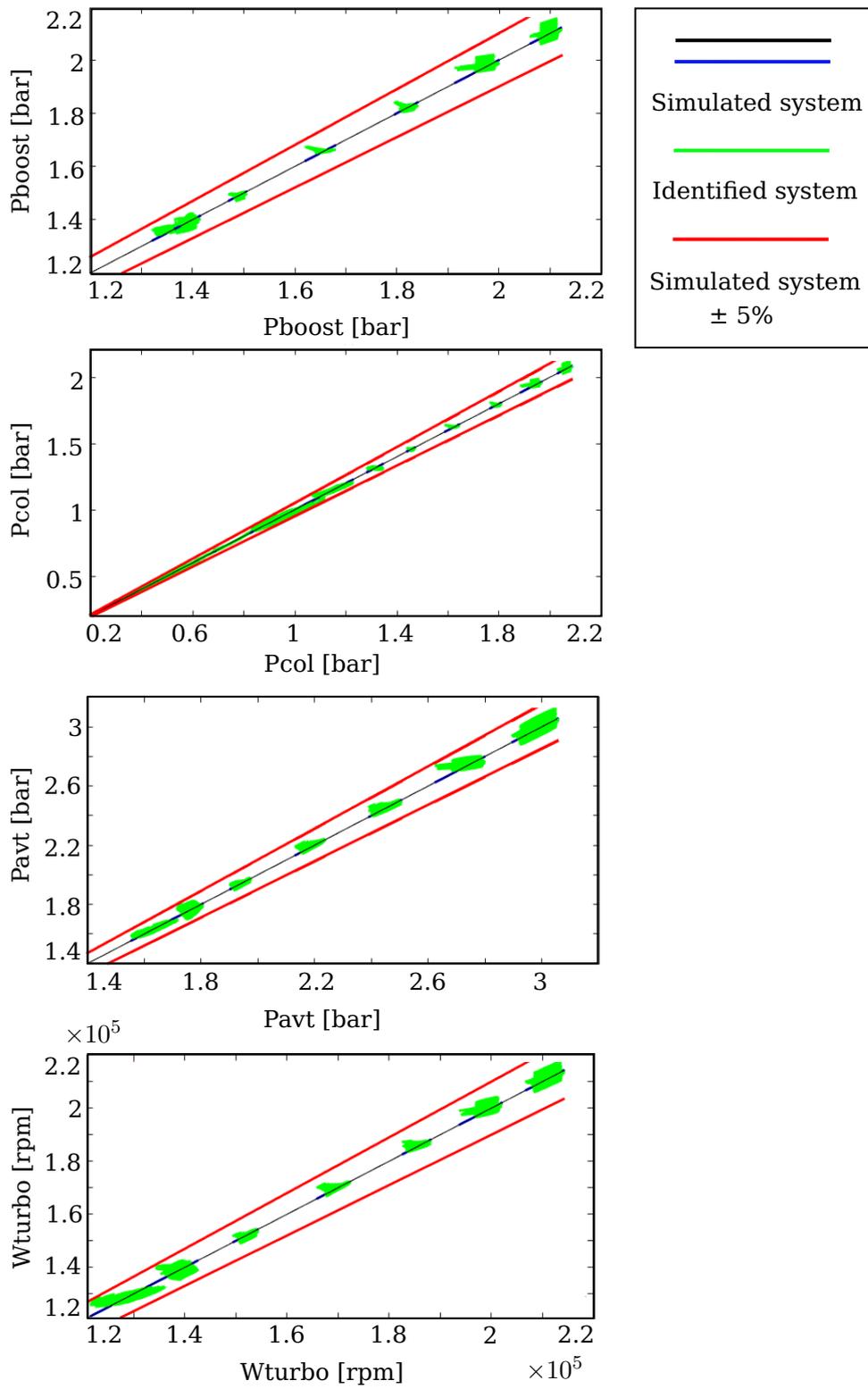
Figure 3.17: Identification results at $4200\,\mathrm{rpm}$ subject to validation criteria
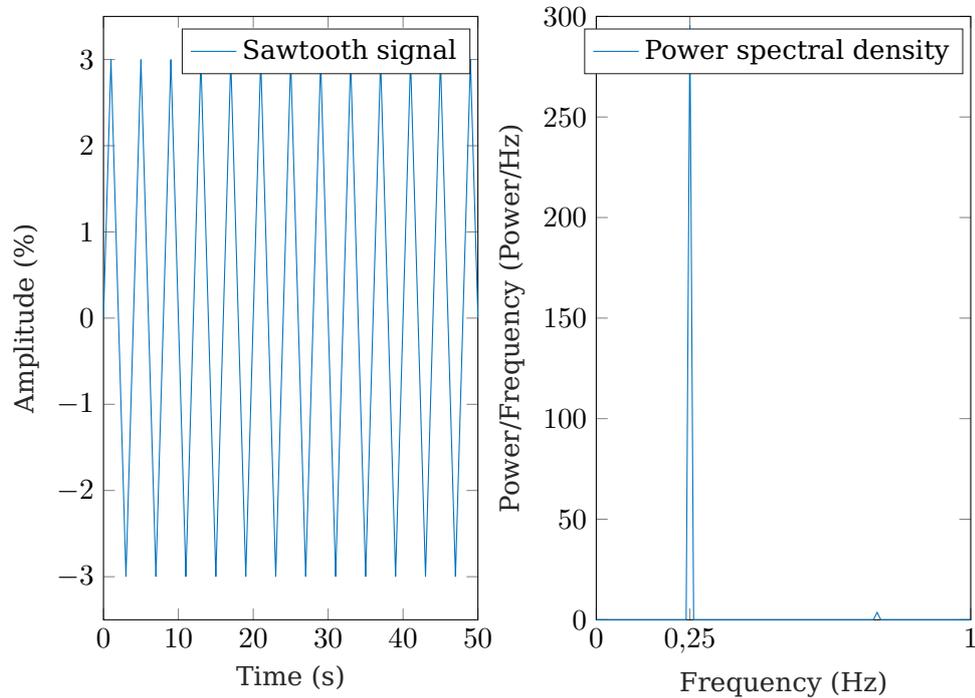
Figure 3.18: Sawtooth signal and its power spectral density

3.5.2   Validation results

To validate the identified models, they are subject to input signals of the triangular and PRBS types, and their outputs are compared to those of the AMESim® model. The amplitude of the triangular and PRBS signals varies between -3% and 3% of the throttle's or waste-gate's steady-state value of the simulation. They are sampled with a rate of $1\,\mathrm{ms}$. The triangular signal solicits only one frequency equal to $0.25\,\mathrm{Hz}$, and the PRBS signal solicits frequencies in the range of $[0,1]$ Hz. Figures 3.18 and 3.19 show the respective plots of these signals along-side their power spectral densities.

To show a sample of the validation results corresponding to the triangular and the PRBS input signals, we refer to Figures 3.20 and 3.21 in the low-torque zone, and to Figures 3.22 and 3.23 in the high-torque zone. For a more global view of the results, Figure 3.24 shows the validation results, corresponding to the triangular signal, subject to the validation criteria described in subsection 3.5.1, and Figure 3.25 shows the validation results, corresponding to the PRBS signal, subject to the validation criteria. In the low-torque zone, P_{col} is shown to respect the 5% margin
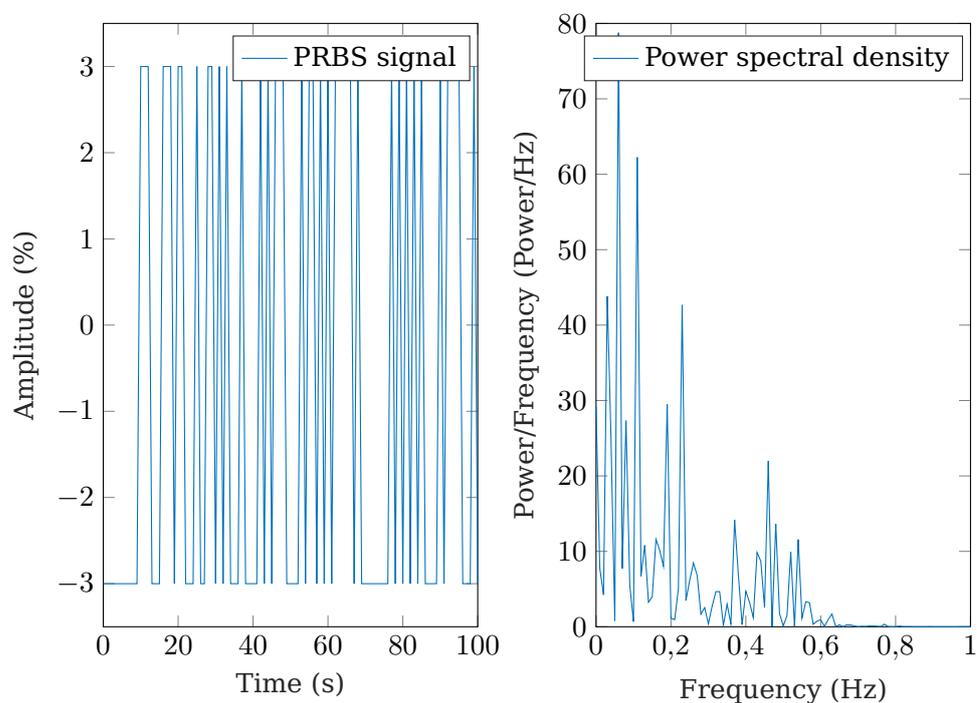
Figure 3.19: PRBS signal and its power spectral density

error, whereas in the high-torque zone, the four outputs are shown to respect 5% margin error only on the last 2 operating points of the ISO field in the $4200\,\mathrm{rpm}$ speed range. This is in conformity with the identification results, shown in Table 3.3, regarding the efficiency of the identification technique in identifying the gasoline engine at operating points of low torque, and its limited efficiency at operating points of high torque. Nevertheless, the models identified at high torque, might be sufficient for exploitation for control purposes.
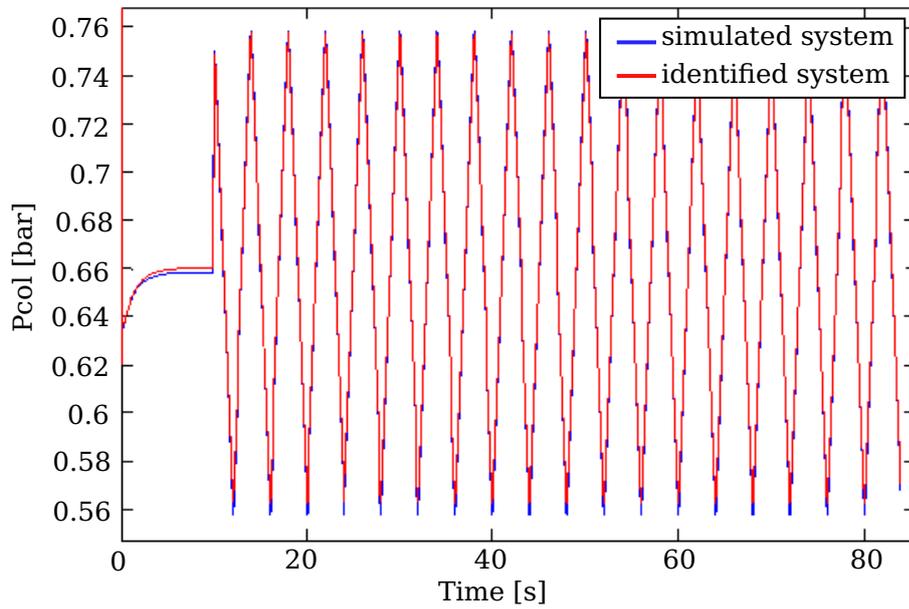
Figure 3.20: P_{col} resulting from a triangular signal at operating point $(4200\,\mathrm{rpm}, 4.7852\,\mathrm{bar})$
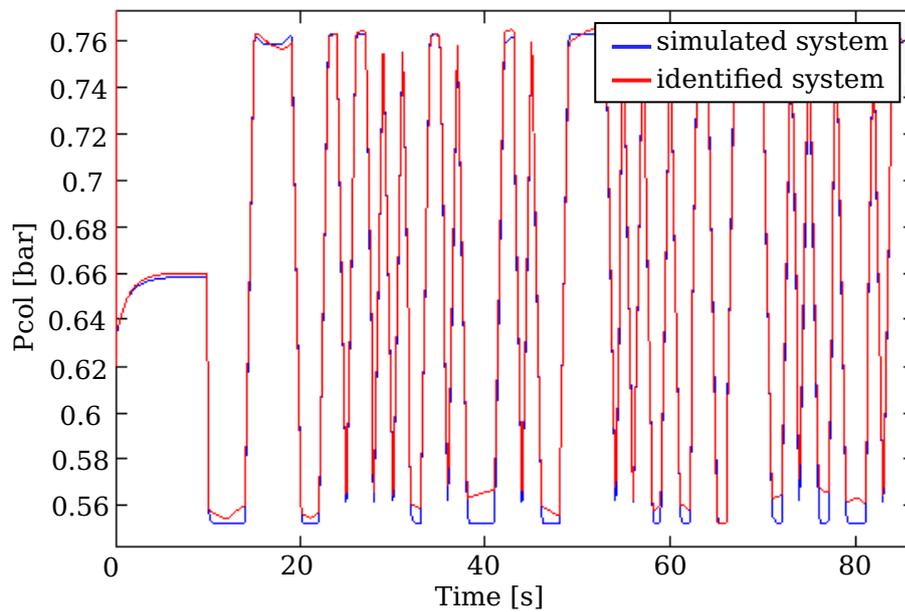
Figure 3.21: P_{col} resulting from a PRBS signal at operating point $(4200\,\mathrm{rpm}, 4.7852\,\mathrm{bar})$
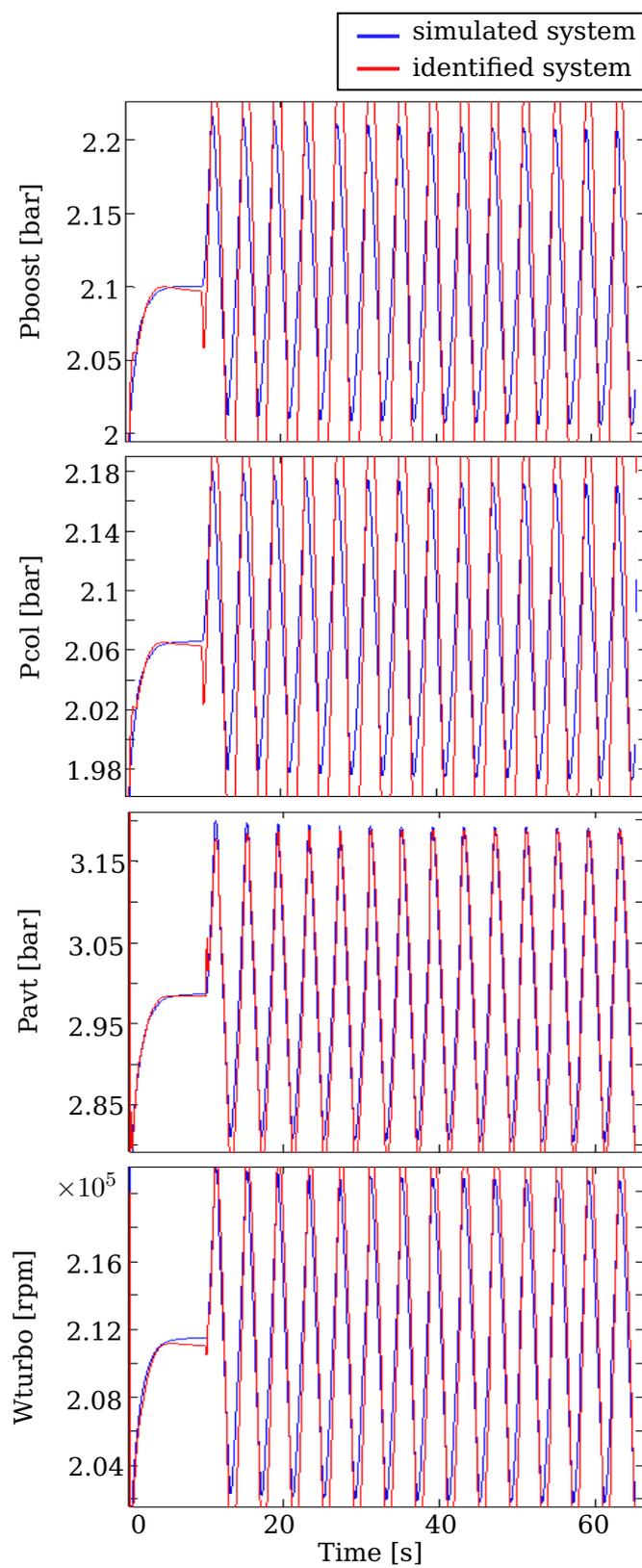
Figure 3.22: Validation results corresponding to the triangular signal at operating point (4200 rpm, 18.98 bar)
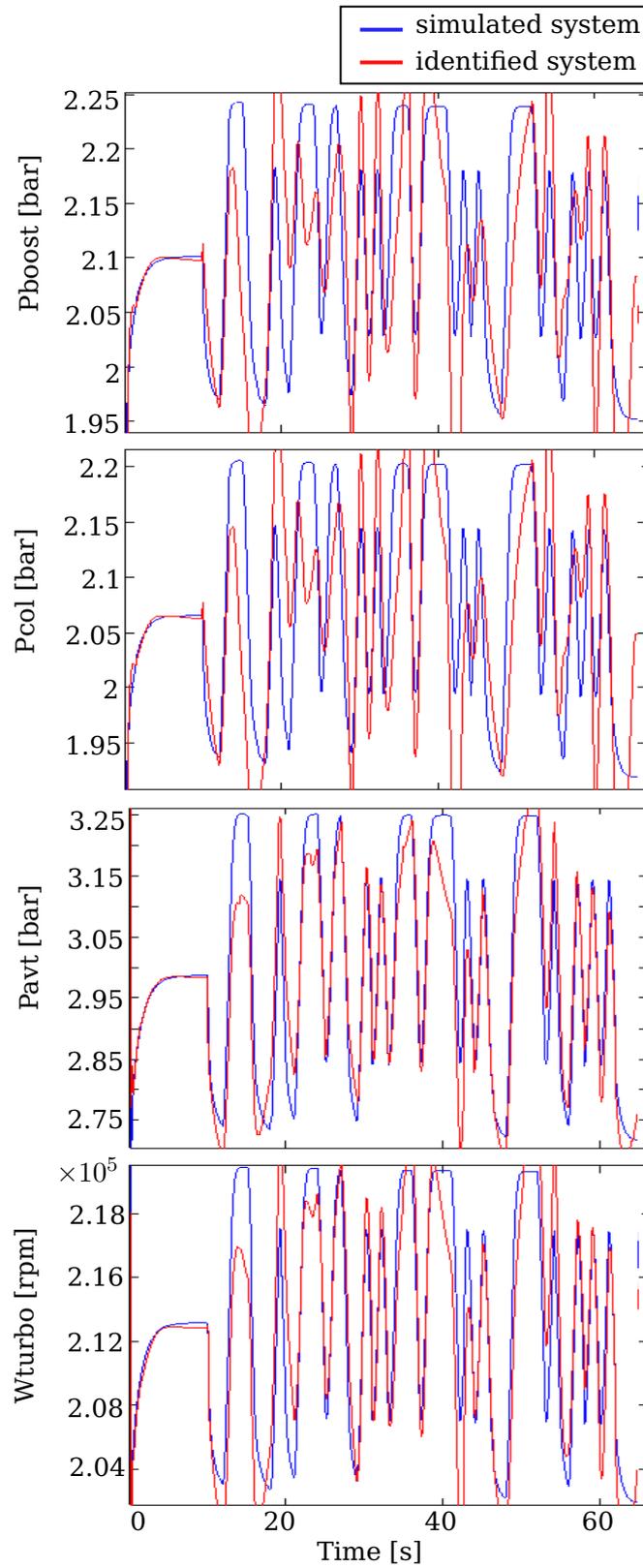
Figure 3.23: Validation results corresponding to the PRBS signal at operating point (4200 rpm, 18.98 bar)
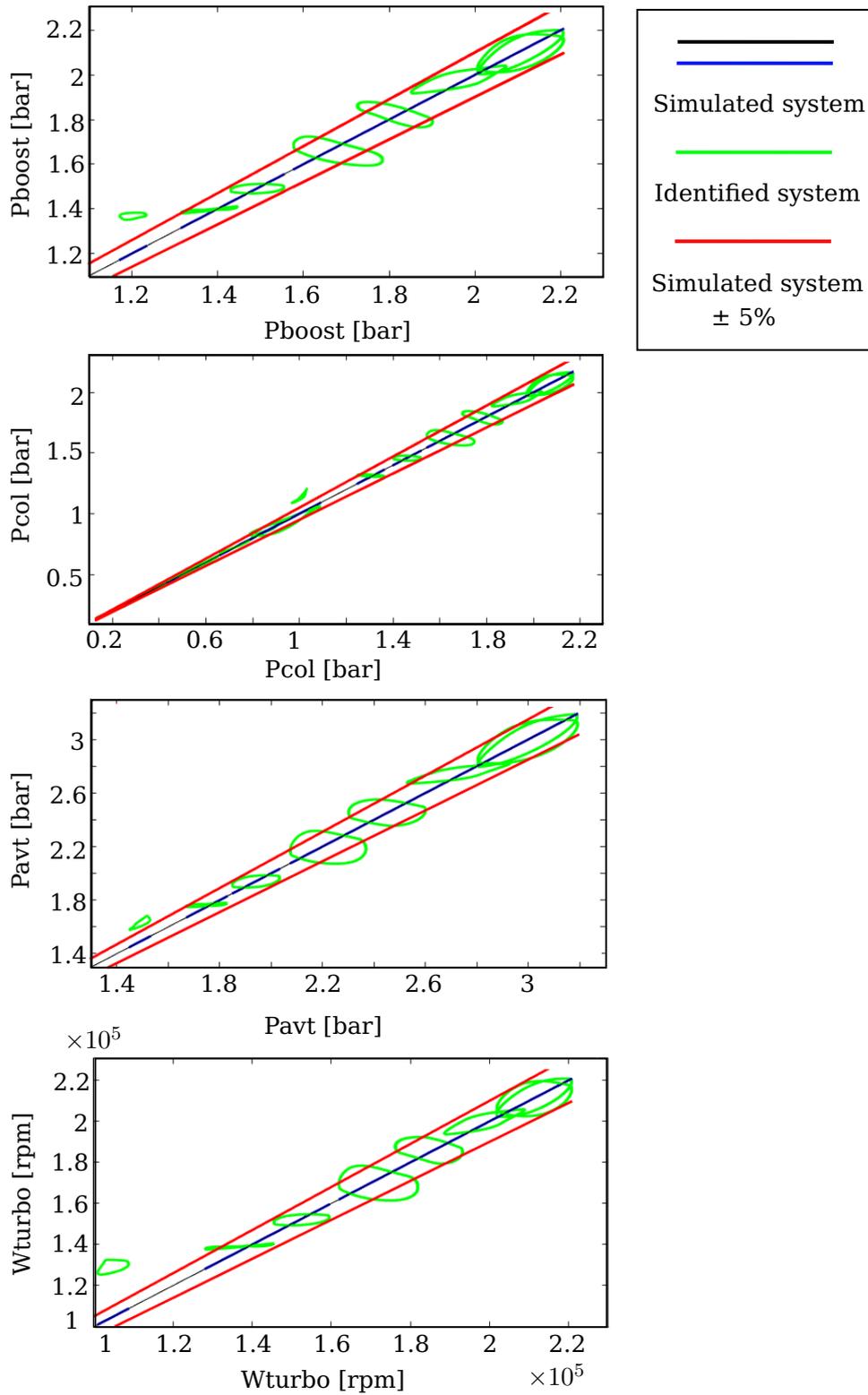
Figure 3.24: Validation results corresponding to the triangular signal at 4200 rpm subject to validation criteria
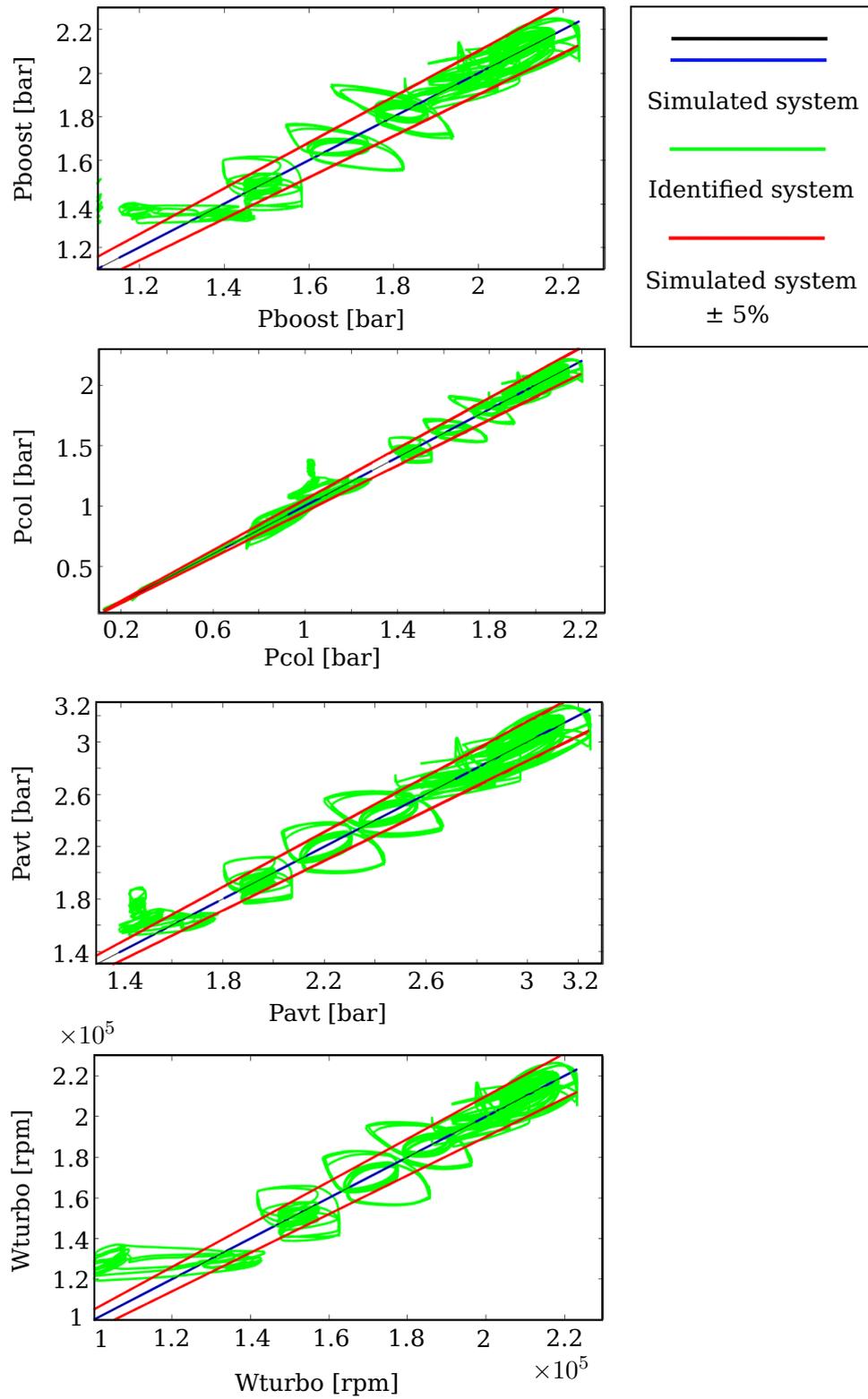
Figure 3.25: Validation results corresponding to the PRBS signal at 4200 rpm subject to validation criteria

3.6 Conclusion and future perspectives

In this chapter, a methodology for gasoline engine identification is proposed. A linear SS model of the engine is identified at a set of operating points in the $4200\,\mathrm{rpm}$ speed range, with mean effective pressure varying from low to high. Multi-sine disturbance signals with good frequency properties are added to well-excite the almost uncorrelated system inputs, PWM_{th} and PWM_{wg}. These inputs along with their corresponding registered outputs, P_{col}, P_{boost}, P_{avt}, and W_t, are fed into the subspace and prediction-error minimization algorithms, which in their turn identify, at each operating point, a linear SS model of the engine. Different model orders are tested without leading to any preference, as to the best model order to be used in future gasoline engine identifications. Different solver option combinations are also tested, and out of 36 combinations, the solver showed preference for only 7, which brings about a time-save of 81% in similar future identifications.

This study, which is one of the few that approach identification of the gasoline engine air-path from the subspace algorithm point of view, proves being highly efficient in identifying operating points were $Pcol <$ $1\,\mathrm{bar}$, yet proves being less efficient for points in the high-torque zone. This deficiency presents the first scope of improvement to this work. Another scope of improvement would be controlling the identified model, and testing its performance in closed-loop. Moreover, given that while some estimated models are of order 4, others are of order 6 or 8, another issue that could be addressed is the model reduction, which in turn opens the door to model comparison and probably model grouping. Finally, the experiments used for identification and validation are ISO experiments being held at specific (speed, mean effective pressure) operating points. A natural extension would be to validate the method on transient tests, preferably real driving cycles.

Electric Throttle Linear Model Identification and Control

The electric throttle, shown in Figure 4.1, is an important component of the powertrain system of nowadays automotive vehicles whose advanced features, such as cruise control, are at the core of an enjoyable driving experience. As shown in the scheme of Figure 4.2, the electric throttle mainly consists of a throttle valve, a DC motor, two return springs, a "low" limit switch spring and a "high" limit switch spring, reduction gears, and two position sensors used to ensure an accurate reading of

the throttle's angular position. Its job is to regulate the amount of air entering an ICE, hence playing an important role in its performance, fuel economy, and pollution aspect.
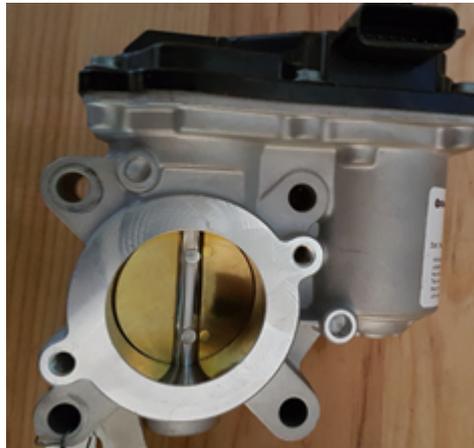
Figure 4.1: Electric throttle
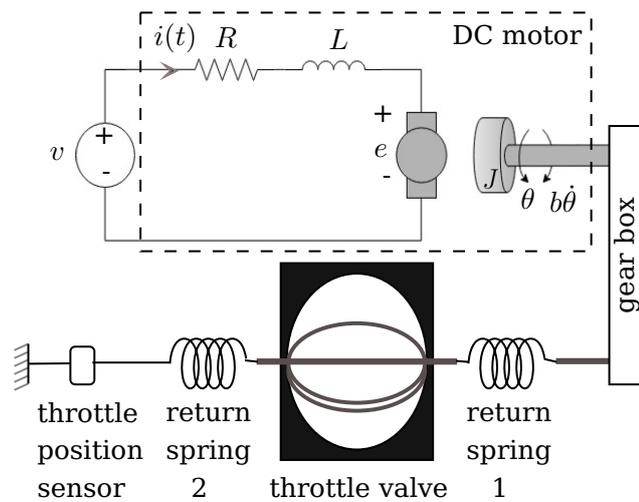
Figure 4.2: Electric throttle scheme

A reliable throttle model is the foundation of its design and control, because it permits to test and improve the design early enough in the design lifecycle to save significant amounts of time and money. However, the electric throttle presents a modeling challenge, despite its simplicity, due to its nonlinearities, which find their origins in the static friction torque, the Coulomb friction torque, the preload torque, and in its dual spring system, comprised of a "low" and a "high" spring, which function selectively under the limp-home position and above it. The limp-home

position is the default angular position taken by the throttle when the engine control unit (ECU) stops sending it control inputs, after it detects abnormal or nonexistent sensor readings, or in case of a communication error between them. At this position, the throttle admits in the intake manifold a small amount of gas, just enough to maintain a minimum level of power and to keep the engine running. By doing so, it puts the engine and the car on safe mode and ensures a certain level of security in case the throttle control fails in the middle of the highway, for example. When the angular position of the throttle flap descends below the limp-home position, the "low" spring becomes active to resist further closure of the valve, and maintain it at least at the limp-home position. The "high" spring, on the other hand, is always active, independently of the valve's angular position. This spring constantly provides a torque to close or to resist the opening of the throttle.

Despite the challenges, building a good model of an electric throttle is worth the effort because of its importance in the overall engine control cycle, from the MIL to the HIL phase. Moreover, the electric throttle along with the other actuators of the electric drive chain, such as the high-pressure and low-pressure EGR valves and the air-intake valve in diesel engines, share the same control strategy and the same structure of one or two return springs, a reduction gear, and a valve. Having accurate models and easily-calibrated controllers for these actuators is crucial in the calibration of the gasoline engines subject to the upcoming euro7 emission norm. In this chapter, for the sake of linearity itself and the advantages that accompany it, we are particularly interested in identifying a linear model of the electric throttle. Given an open-loop AMESim$^{®}$ model of the electric throttle, the model is written in a linear form, and then linearly controlled, to follow a predefined set-point or reference trajectory. The linear control strategy applied on the AMESim$^{®}$ model can hence be transfered to the real throttle and tested in later phases of control design. The methodology of deriving the electric throttle's linear model and controlling it is described in this chapter. It is generic and can therefore be extended to other similar actuators such as the EGR valve.

State of the art Electric throttle identification and control is a topic whose deliberation dates back to the late 1900s. Almost 20 years of research have already been conducted on the subject, which makes an extensive literature regarding the types of models, identification techniques, and control strategies.

Regarding the models of electric throttles in cars and motorbikes, different sorts of models can be found: Mathematical, black-box, piece-wise affine (PWA), piecewise auto-regressive exogenous (PWARX), ARMAX, SS, LPV models, etc... Most of the works in the literature present linear and nonlinear mathematical models of the electric throttle based on the physical equations representing the functioning of the DC motor and the spring(s), take for example [AK+17]; [BPP05]; [PCS13]. Black-box models also had their share of electric throttle; see for example [DPPJH04] and [CTSF11] where an linear time-invariant (LTI) nominal model was experimentally identified in the frequency domain. PWA models are widely spread as well. [VBMPP06], for example, presents discrete-time models of the static and dynamic friction present in the throttle. In [VMP05], a clustering-based method was used to identify a PWARX model. LPV modeling can be seen in [TE09] with an LPV mathematical model and in [ZYZ15] with a discrete-time physics-based model where the time-varying parameters are the vehicle battery voltage and the friction torque, which is a function of the throttle's position. In [Yan04], six different linear models were identified at six different operating points of the throttle. In [GB00], a mathematical model is established for the linear parts and more than one ARMAX model, distributed below and above the limp-home position, are estimated by MATLAB for the nonlinear parts. [PD11] presents a linear SS model, but the addition of a nonlinear friction model, results in a nonlinear model of the throttle. On the other hand, avoiding all the identification difficulties, [JSK09] presents a model-free control technique.

As for the model identification techniques, [BPP05]; [JZS14]; [PCS13] use an online estimator of the friction, preload torque, and spring nonlinearities. In [BPP05], however, the online estimator is neural-network based. [BNS06], on the other hand, uses a state observer enhanced by an output-error technique where the model is extracted from the closed-loop identification carried on with a proportional controller. [NSHO06] relies on nonlinear optimization and genetic algorithms to estimate the system parameters. Fuzzy system modeling is used in [YW09] where Taylor expansion is used to linearize an input-output model. In contrast to the online estimation techniques, [DPPJH04]; [PDJP06] perform an offline identification which is carried out separately for the linear part representing the DC motor and for the nonlinear part representing the spring limp-home position and the static and dynamic frictions. [DP-PJH04], however, is distinguished by a multi-step identification approach that characterizes the dynamics of the linear throttle depending on its physical form, and another single-step black-box identification approach

that captures the part of the model which has no physical significance.

When it comes to control, PID is one of the most popular forms of electric throttle control; take for example [AK+17]; [Li+12]. Sliding mode control (SMC) has also been used [OHP01]; [PD11], in addition to H_2/H_∞ [ZYZ15], and other types of controllers. Optimal control also has its share of electric throttle control. For instance, an optimal linear quadratic Gaussian (LQG) controller was used with another proportional controller to compensate for the nonlinearities observed in the throttle system [KSHKS96]. LQG optimal control with integral action applied on a SS linear model was also compared with RST control applied on an ARMAX model [GB00]. In case of LQG and RST control, the control law applied below the limp-home position is different from that applied above it. Also, an adaptive scheme is followed in [GB00], where three methods, among which is recursive least-squares, are used to estimate the parameters of the nonlinear limp-home position.

Proposed methodology In a first step, this chapter proposes to use the physical equations corresponding to the AMESim® model to form an LPV SS model of the electric throttle, taking into account the throttle's nonlinearities, such as limp-home position nonlinearity. The LPV SS model captures the most important aspects of the AMESim® model without considering unnecessary details and functionalities that would complicate the linear system without offering additional accuracy. Consequently, the LPV SS system is used as a basis leading to the formation of a regression vector, employed later in the identification of an LPV ARX model of the throttle. In a second step, the identified ARX and SS systems are subject to LQR control design, which features reference and disturbance feedforwards, and which is later verified on the original open-loop AMESim® model. In addition to these simulation results, experimental results are demonstrated on a real throttle test-bench, to which the LPV model is compared and adapted acccordingly.

Note that from the identification point of view, although the PWARX identification of the electric throttle was suggested in [VMP05], the advantage of the work proposed in this chapter is the reduced number of sub-models. Furthermore, the state of the art shows that mathematical modeling and black-box modeling are both used in the literature to model the electric throttle, but they are usually used separately not in collaboration with one another, because mathematical modeling is used when the physical equations of the system are known and black-box modeling is used in the opposite case, when these equations are not

known. In [DPPJH04], for example, though both techniques were used together, but each of them was used separately to identify a different part of the model. In this work, to take advantage of the merits of both techniques and provide a model which implies the physical laws governing the electric throttle's functioning, we propose to use the physical equations that we know about the throttle to set the structure of the black-box model. That is to say, we use the physical equations to define the regression vector of the ARX model. Once this is done, all we need is the system input and output to do a system identification and find the model parameters. Even if a new throttle model with different parameters has to be identified, the ARX model regressor remains unchanged, and the new throttle's input and output are plugged in to identify the new linear model. Combining both techniques exploits the capability of black-box modeling to capture the obscure aspects of the electric throttle functioning, and results in a physically meaningful black-box model all at the same time. Consequently, the conversion of the ARX model to a SS model results in a 2nd order model which has a physical meaning, unlike the numerically identified SS model in [GB00].

From the control point of view, a PID controller could have been directly used to control the nonlinear AMESim$^®$ models. But taking into consideration that PID control has been widely explored with the electric throttle, it is interesting to explore other control techniques as well. LQR control is chosen for this application for one main reason. In addition to its numerous well-known advantages, an LQR controller has one tuning parameter (the Q/R weighting ratio). This interesting feature allows rapid switching from one throttle model to another, without having to do time-consuming controller calibrations each time. This is essential in the automobile industry, taking into consideration the frequent changes in the engines' TDs. Finally, another contribution of this work is the systematic methodology proposed for converting an AMESim$^®$ nonlinear model into a linear model and controlling it, a methodology which can be applied on other actuators' models, similar in structure to the electric throttle.

Chapter structure This chapter is organized as follows. Section 4.1 details the physical equations governing the functioning of the electric throttle and the scheduling parameters used to split the functioning zones. Section 4.2 describes the throttle nonlinear AMESim$^®$ model, and discusses the identification of the LPV models. Section 4.3 presents the electric throttle test-bench and identifies the LPV model corresponding to it. Section 4.4 presents the LQR controller associated with the

initial LPV models, and the application of the control design on the initial nonlinear model of the throttle. Section 4.5 describes the overall proposed methodology to present an AMESim® nonlinear model of the throttle in a linear form and control it, and section 4.6 concludes the chapter.

4.1 Modeling an electric throttle with an LPV Model

To understand the electric throttle's nonlinear behavior and get insight about its different functioning zones, the mathematical equations, lying behind its components, are inspected, by analyzing the reference model provided by the engineers from Renault and referring to the documentation of AMESim®. The underlying equations, representing the torques stemming from the different model components shown in Figure 4.2, govern the electric throttle's functioning and shed light on its nonlinearities. Four torques are listed below: the motor torque, the torques stemming from the "high" and "low" return springs, and the friction torque applied on the motion.

- The motor torque is:

$$T_{motor} = K \cdot i \cdot \eta_{GB} \qquad (4.1)$$

  where K is the motor torque constant ($\mathrm{N\,m\,A^{-1}}$), i is the electric current (A), and η_{GB} is the gear box efficiency.

- The torque stemming from the "high" spring is:

$$T_{high} = T_h + T_{vh} + T_{ph} \qquad (4.2)$$

  where $T_h = ktor \cdot \theta$ is the "high" spring torque ($\mathrm{N\,m}$), $T_{vh} = rtor \cdot \dot{\theta}$ is the viscous friction torque associated with the "high" spring ($\mathrm{N\,m}$), and T_{ph} is the constant "high" spring preload torque ($\mathrm{N\,m}$).

  The constant $ktor$ is the "high" spring stiffness ($\mathrm{N\,m\,rad^{-1}}$), θ is the angular displacement (rad) of the throttle flap, $rtor$ is the viscous friction constant associated with the "high" spring ($\mathrm{N\,m\,s\,rad^{-1}}$), and $\dot{\theta}$ is the angular speed ($\mathrm{rad\,s^{-1}}$) of the throttle flap with respect to time.

- The torque stemming from the "low" spring is:

$$T_{low} = \begin{cases} T_{contact} & , if\, T_{contact} \leq T_{pl} \\ T_{spring} & , else \end{cases} \tag{4.3}$$

where
$$T_{contact} = \begin{cases} -kint(\theta - \theta_0) - reff \cdot \dot{\theta} & , if\, \theta < \theta_0 \\ 0 & , if\, \theta \geq \theta_0 \end{cases} \text{ is the contact}$$
torque (N m), and

$$T_{spring} = \begin{cases} T_l + T_{vl} + T_{pl} & , if\, \theta < \theta_0 \\ 0 & , if\, \theta \geq \theta_0 \end{cases} \text{ is the spring torque (N m),}$$

where $T_l = kcont \cdot \theta$ is the "low" spring torque (N m), $T_{vl} = rcont \cdot \dot{\theta}$ is the viscous friction torque associated with the "low" spring (N m), and $T_{pl} = kint \cdot disth$ is the constant "low" spring preload torque (N m).

The constant $kint$ is the contact stiffness (N m rad−1), θ_0 is the initial displacement or the limp-home position (rad), $kcont$ is the "low" spring stiffness (N m rad−1), $rcont$ is the viscous friction constant associated with the "low" spring (N m s rad^{-1}), $disth$ is the angular displacement threshold to reach the "low" spring preload (rad), and $reff$ is the effective contact damping coefficient (N m s rad^{-1}) calculated as
$$reff = r(1 - \exp\frac{-(\theta - \theta_0)}{e})$$

where r is the damping to reach the preload (N m s rad^{-1}) and e is the limit penetration distance for full damping (rad).

Note that the conditions $T_{contact} \leq T_{pl}$ and $T_{contact} > T_{pl}$ in equation (4.3) are respectively equivalent to $0 < \Delta\theta \leq disth$ and $\Delta\theta > disth$ where $\Delta\theta = -\theta + \theta_0$. Figure 4.3 shows how the stiffness changes from contact to spring stiffness as $\Delta\theta$ reaches the angular displacement threshold $disth$.

- AMESim$^{\circledR}$ provides different ways for modeling the friction torque, from static models, such as the Coulomb friction model, to dynamic models, such as the Dahl, reset-integrator, or the LuGre model. Because the AMESim$^{\circledR}$ models used in this work solicit two different friction models, the static Coulomb friction model and the dynamic reset-integrator model, the calculation of the friction torque will be addressed from both perspectives.
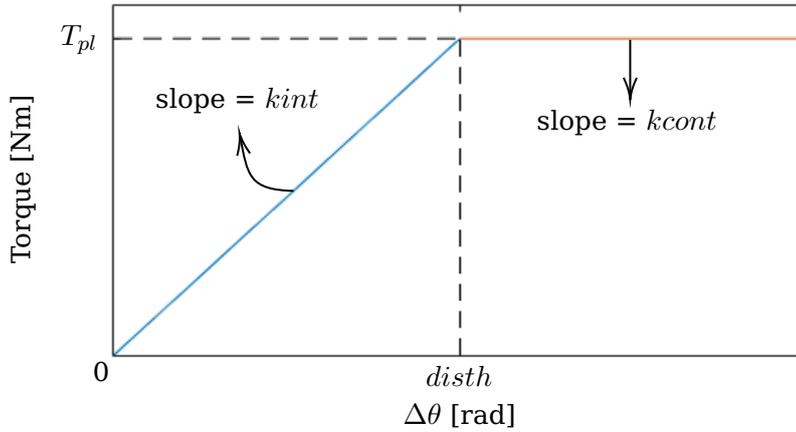
Figure 4.3: Low spring: contact vs spring torque

∗ In case of the Coulomb friction model, the torque stemming from the friction component is:

$$T_{friction} = T_c \text{sgn}(\dot{\theta}) + b\dot{\theta} \tag{4.4}$$

where T_c is the constant Coulomb friction torque (N m) and b is the motor viscous friction constant (N m s rad^{-1}).

Note that equation (4.4) represents the dynamic frictions included in the Coulomb friction model: the Coulomb friction which is the friction that opposes a body's motion with a constant force regardless of its velocity and contact area, and the viscous friction which is a linear function of the body's velocity. From Figure 4.4, which shows a scheme of the Coulomb and viscous friction torques, notice the discontinuity of the model and the uncertainty of the Coulomb torque at $\dot{\theta} = 0$, as it can take any value in the interval $[-T_c, T_c]$. In addition to these two dynamic frictions, the Coulomb friction model also allows modeling the static friction force, otherwise called stiction, which represents the sticking of two bodies as they tend to slip over each other but do not move ($\dot{\theta} = 0$). Because stiction is neglected in the original AMESim® model , it is not detailed any further here.

∗ In case of the reset-integrator model, the torque stemming from the friction component is:

$$T_{friction} = \begin{cases} T_{stick} + T_{vs} & , if \; |\theta_{diff}| < dtrel \\ T_c \text{sgn}(\dot{\theta}) & , if \; |\theta_{diff}| \geq dtrel \end{cases} \tag{4.5}$$
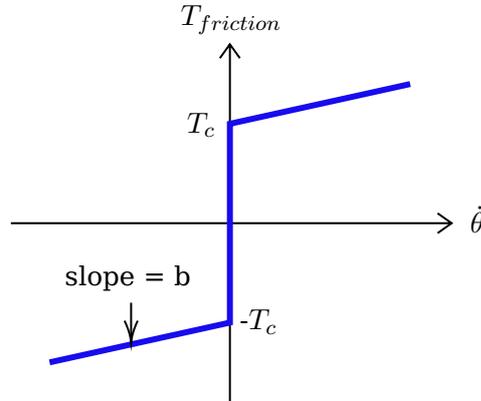
Figure 4.4: Friction torque as a function of the angular velocity for the Coulomb friction model

where $T_{stick} = krel \cdot \theta_{diff}$ (N m), and $T_{vs} = rtors \cdot \dot{\theta}/N$ is the equivalent viscous friction torque during stiction (N m).

The constant $rtors$ is the equivalent viscous friction constant during stiction (N m s rad^{-1}), N is the gear ratio, $krel$ is the stick stiffness (N m rad-1), and $dtrel$ is the stick displacement threshold (rad). θ_{diff} is an internal variable of the friction model standing for the relative differential angle, between the two ports of the inertial component, during stiction. The dynamics of θ_{diff} are calculated as:

$$\frac{d\theta_{diff}}{dt} = \begin{cases} 0 & , if\,(\dot{\theta} < 0\,and\,\theta_{diff} \leq -dtrel) \\ 0 & , if\,(\dot{\theta} > 0\,and\,\theta_{diff} \geq dtrel) \\ \frac{\dot{\theta}}{N} & , else \end{cases}$$

The integrator implied by this internal variable θ_{diff} is what gives this reset-integrator friction model its name. During stiction phase, the integrator input is set equal to $\dot{\theta}$ (divided by N to account for the gear ratio), implying an "integrator" action, and it is "reset" to zero once θ_{diff} reaches the stick displacement threshold $dtrel$ and starts sliding during the Coulomb friction phase. Figure 4.5 shows a scheme of the Coulomb and static friction torques, T_c and T_s, with respect to θ and $\dot{\theta}$. $T_s = krel \cdot dtrel$ is the constant stiction torque (N m) which acts as an upper limit for $|T_{friction}|$. The continuous transition, shown in Figure 4.5, from T_s to T_c is due to the Stribeck effect, which is not detailed here because it is considered negligible in the AMESim® model used in this work.
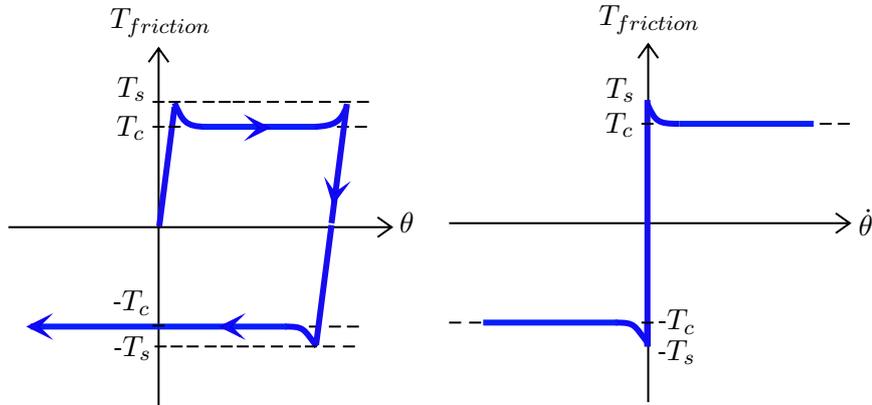
Figure 4.5: Friction torque as a function of the angular displacement (left) and as a function of the rotary velocity (right) for the reset-integrator friction model

The above equations reveal the origins of the throttle nonlinearities. Let's first consider equations (4.1), (4.2), and (4.3). While the motor torque, "high" and "low" spring torques, and "high" and "low" viscous friction torques are respectively linear functions of the current i, the relative angular displacement θ, and the relative angular velocity $\dot{\theta}$, the "high" and "low" springs preload torques reveal the first sources of non-linearity. The contact torque in equation (4.3) shows a nonlinearity due to the limp-home position θ_0. Finally, the Coulomb and stiction torques in equations (4.4) and (4.5) reflect other nonlinearities due to the sign function of the angular velocity $\dot{\theta}$ and due to the torque saturation imposed by T_s.

In addition to the nonlinearities, the equations above reveal the different functioning zones of the throttle, shown in Figure 4.6. The first zone definition, shown in the top table of Figure 4.6, is with respect to θ (zones 1, 2, and 3). Indeed, while T_{motor} and T_{high} are common for zones 1, 2, and 3, T_{low} is calculated differently in each of these zones. When $\theta_0 \leq \theta \leq \theta_{max}$ (zone 3), T_{low} is nonexistent, whereas when $\theta_{min} \leq \theta < \theta_0 - disth$ (zone 1) and $\theta_0 - disth \leq \theta < \theta_0$ (zone 2), the motor is subject to T_{low}, coming in the form of T_{spring} and $T_{contact}$, respectively. Note that the springs never allow the throttle to go beyond its limit positions θ_{min} and θ_{max} to avoid pushing the limits of the DC motor.

Considering $T_{friction}$ in case of the reset-integrator model, it is calculated differently in each of the 5 zones A, B, C, D, and E, shown in the middle table of Figure 4.6, independently of the value of θ. This table

shows the second zone split, which is with respect to θ_{diff} and $T_{friction}$. Note that selecting θ_{diff} among the scheduling parameters is permissible because it is a function of θ, which in turn is a measurable parameter. On the other hand, in case of the Coulomb friction model, $T_{friction}$ is calculated differently in each of the two zones F and G shown in the bottom table of Figure 4.6. This table shows the second zone split, which is with respect to $\dot{\theta}$.

Therefore, by crisscrossing zones $1-3$ with zones $A-E$, when $T_{friction}$ is calculated using the reset-integrator model, 15 functioning zones of the electric throttle can be distinguished, and by crisscrossing zones $1-3$ with zones $F-G$, when $T_{friction}$ is calculated using the Coulomb friction model, 6 functioning zones of the electric throttle can be distinguished. Therefore, the number of functioning zones of an electric throttle's LPV model depends on the type and complexity of the models chosen to depict the phenomena taking place in the throttle. The 6-zone and the 15-zone split schemes presented here will be projected respectively on two different AMESim® models and their modeling efficiencies will be compared.

The above-mentioned physical equations can be used to write the AMESim® model in the form of an LPV SS model. The initial AMESim® model provided by the Renault engineers can be written as a 4-zones LPV SS model, two of these zones lying below the limp-home position and two lying above it. This LPV SS model takes as scheduling parameters the position and speed of the throttle. This model is used to form an LPV ARX model and to design an LQR controller for the AMESim® model. In the rest of the chapter, this AMESim® model is referred to as AMESim® model A.

Unfortunately, upon the late arrival of an electric throttle test-bench, after the above proposed identification and control methodology had already been applied, verification of the AMESim® model, with respect to the test-bench, turned out to be unsuccessful. Upon giving the AMESim® model and the test-bench the same input, they resulted in different outputs. This is why another AMESim® model is designed, one which is sure to be valid with respect to the test-bench, and the ARX system identification proposed above is executed again, this time based on the new AMESim® model. In the rest of the chapter, this AMESim® model is referred to as AMESim® model B.

The next sections detail the AMESim® nonlinear models A and B, the functioning zones in which they fall and therefore the equations leading
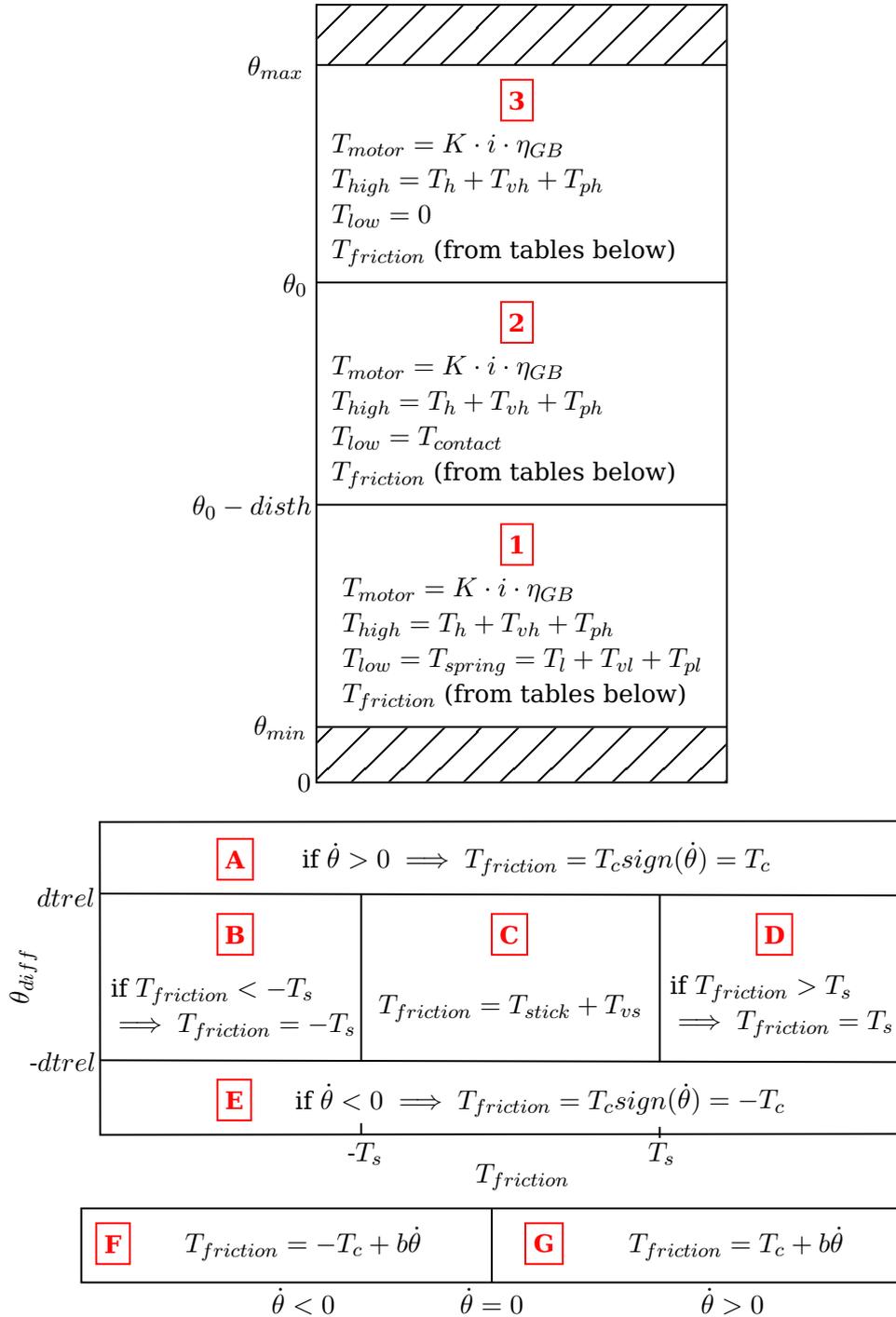
Figure 4.6: Functioning zones of the electric throttle

to their respective LPV SS models A_0, A, and B, and the identifications leading to their respective LPV ARX models A and B.

4.2 Identification of an electric throttle model and simulation results

4.2.1 Nonlinear AMESim® model A and its nonlinearities

The model of the throttle built on the simulation platform LMS Imagine.Lab AMESim® is shown in Figure 4.7. It is modeled as a combination of the throttle flap, a DC motor, 2 springs, a "low" return spring and a "high" return spring, reduction gears, and a group of sensors. No separate component is dedicated for modeling the friction. Instead, the inertial component is parameterized in way that the friction on the movement of throttle is taken into account. The control of the throttle is built in Simulink®, which accesses AMESim® model A through an S-function, thus resulting in a co-simulation environment launched by MATLAB.
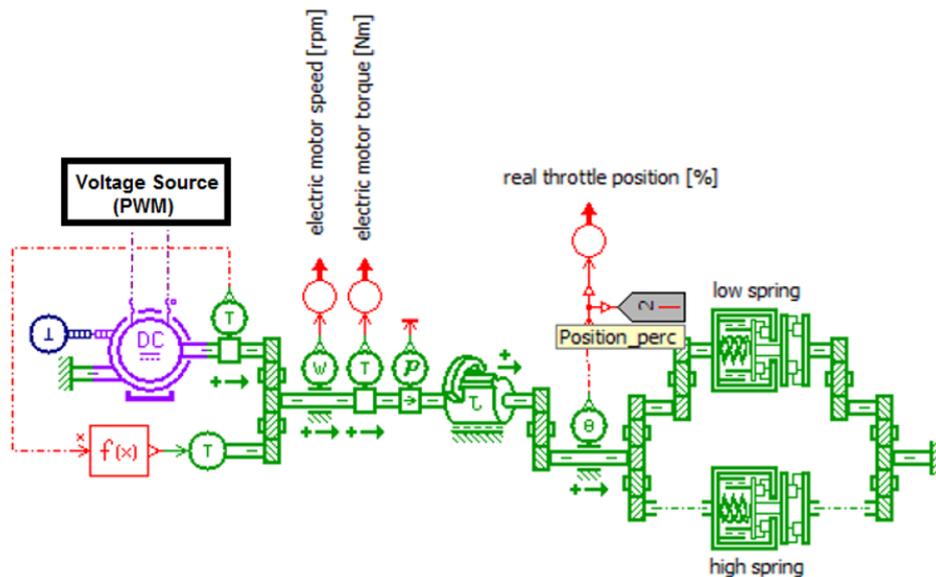
Figure 4.7: AMESim® nonlinear model A of the electric throttle

The throttle takes a PWM (%) as input and gives an angular position (%) as output. Many of its states are accessible: power (W), torque (N m), speed (RPM), voltage (V), and current (A).

This nonlinear model built in AMESim® was validated by Renault engineers with respect to a real throttle test-bench present in Renault. A set of experiments was conducted in open-loop on the real throttle at different temperatures, where a voltage (PWM) signal was sent to the throttle, and its corresponding angular positions were recorded over time, resulting in a set of measured input and output profiles. The AMESim® open-loop simulation of a concatenation of some of these experiments is shown in Figure 4.8.
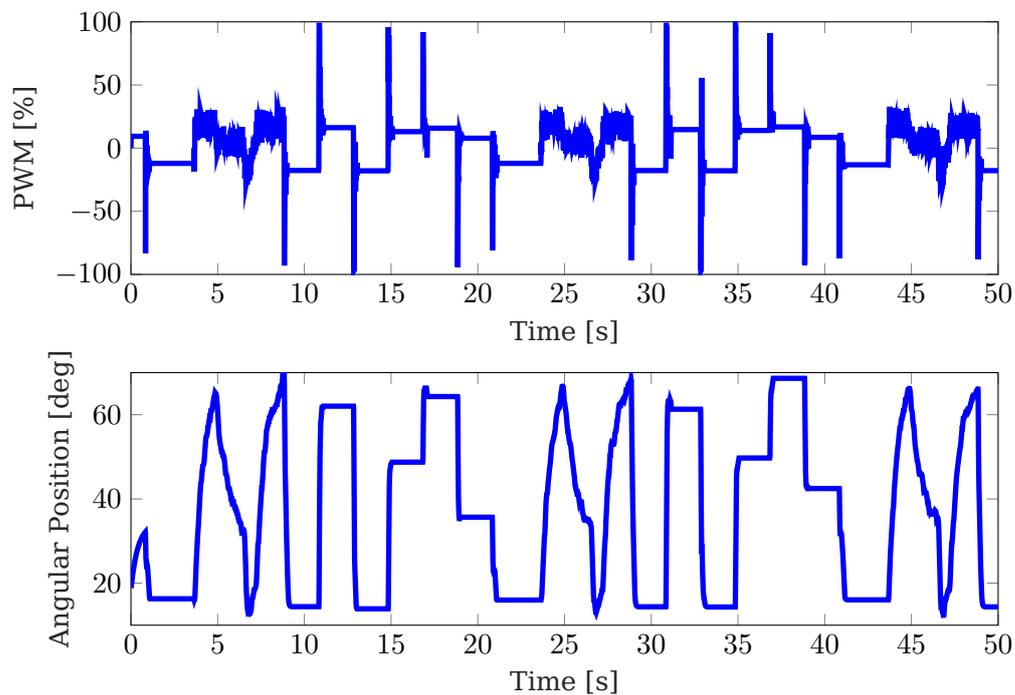
Figure 4.8: Input and output signals of AMESim® nonlinear open-loop model A

LQR control design requires a linear model, which brings us to the next step: finding the linear equivalent of AMESim® nonlinear model A, thus the need to understand the origin of the model nonlinearities prior to any linearization attempt. This model presents a high nonlinearity mainly due to the springs' preload torques, the springs' torques at the limp-home position, and Coulomb friction torque, whose value is directly associated to the sign of the angular speed. Due to these nonlinearities, the system is represented as an LPV system with four sets of parameters in four functioning zones, as shown in the next subsection.

4.2.2 Linear-parameter-varying state-space model A

As mentioned earlier, AMESim$^{®}$ model A admits the Coulomb friction model for the calculation of $T_{friction}$, which means that its functioning is restricted to 6 zones: 1F, 1G, 2F, 2G, 3F, and 3G (referring to Figure 4.6). This means that writing this model in the form of an LPV model requires only two scheduling parameters: θ and $\dot{\theta}$, the respective position and speed of the throttle.

A few remarks can be made in this context about the functioning zones:

- to simplify the linear model and minimize the number of functioning zones, the contact torque is neglected and $disth$ is assumed equal to 0, which leads to merging zones 1 and 2 in the first table of Figure 4.6 and considering that $T_{low} = T_{spring}$ all the time.

- the Coulomb friction model presents a discontinuity of the friction torque at $\dot{\theta} = 0$ and does not explicitly specify the value of the Coulomb torque at this speed. For simulation reasons, however, we assign the value of T_c to the Coulomb friction torque at $\dot{\theta} = 0$, as if it falls in zone G where $\dot{\theta} > 0$.

Taking these remarks into account, the 6 zones reduce to four zones, which are split at the points presenting the nonlinearities originating from the limp-home position and the Coulomb friction torque, as follows:

- zone $1 - 2$F: $\theta_{min} \leq \theta < \theta_0$ and $\dot{\theta} < 0$

- zone $1 - 2$G: $\theta_{min} \leq \theta < \theta_0$ and $\dot{\theta} \geq 0$

- zone 3F: $\theta_0 \leq \theta \leq \theta_{max}$ and $\dot{\theta} < 0$

- zone 3G: $\theta_0 \leq \theta \leq \theta_{max}$ and $\dot{\theta} \geq 0$

Two other remarks can be made concerning particularities in AMESim$^{®}$ model A:

- in zones 1 and 2 (i.e. below the limp-home position), only the "low" limit spring is active. The "high" spring is active only in zone 3 (i.e. above the limp-home position), which means that $T_{high} = 0$ in these zones.

- viscous friction is considered at the level of the inertial component (DC motor) following the Coulomb friction model, not on the level of the springs, which means that $T_{vl} = T_{vh} = 0$.

To build the LPV SS model of the electric throttle, the first equation to consider is Newton's Second Law of rotation.

$$J\ddot{\theta}_{motor} = \sum T = T_{motor} + NT_{high} + NT_{low} - T_{friction} \qquad (4.6)$$

where J is the moment of inertia of the motor ($\mathrm{kg\,m^2}$) and $\ddot{\theta}_{motor}$ is the angular acceleration of the motor with respect to time ($\mathrm{rad\,s^{-2}}$).

$\theta_{motor} = \frac{\theta}{N}$ (θ being the angular displacement of the throttle which is on the same level as the springs) and $T_{on\,the\,motor\,level} = NT_{on\,the\,spring\,level}$. Therefore, T_{high} and T_{low} are multiplied by the gear ratio N, as they are generated on the level of the springs, not on the level of the DC motor to which we apply Newton's Second Law.

Equation (4.6) takes two different forms depending on the functioning zone.

In zones $1 - 2F$ and $1 - 2G$:

$$J\ddot{\theta}_{motor} = T_{motor} + NT_{low} - T_{friction}$$

$$= Ki \cdot \eta_{GB} + Nkcont\Delta\theta + NT_{pl} - T_c\mathrm{sgn}(\dot{\theta}) - \frac{b\dot{\theta}}{N}$$

In zones 3F and 3G:

$$J\ddot{\theta}_{motor} = T_{motor} + NT_{high} - T_{friction}$$

$$= Ki \cdot \eta_{GB} + Nktor\Delta\theta - NT_{ph} - T_c\mathrm{sgn}(\dot{\theta}) - \frac{b\dot{\theta}}{N}$$

The second equation to consider is Kirchhoff's voltage law.

$$V = K\dot{\theta} + L\frac{di}{dt} + Ri \qquad (4.7)$$

where V is the electric voltage (V), R is the electric resistance (Ω), and L is the electric inductance (H).

J	$3.29 \times 10^{-6}\,\mathrm{kg\,m^2}$	T_{ph}	$0.05\,\mathrm{N\,m}$
N	0.04566	T_{pl}	$0.05\,\mathrm{N\,m}$
K	$0.021\,\mathrm{N\,m\,A^{-1}}$	$ktor$	$0.55\,\mathrm{N\,m\,rad{-}1}$
R	$2\,\Omega$	$kcont$	$5.16\,\mathrm{N\,m\,rad{-}1}$
L	$0.001\,25\,\mathrm{H}$	θ_{min}	$5°$
η_{GB}	0.95	θ_{max}	$95°$
b	$0.0002\,\mathrm{N\,m\,s\,rad^{-1}}$	θ_0	$18°$
T_c	$0.005\,\mathrm{N\,m}$		

Table 4.1: Parameters of linear model A

To build the linear SS model A_0, consider the state vector $X_0 = \begin{bmatrix} \theta & \dot{\theta} & i \end{bmatrix}^T$, the output $Y = \theta$, the input $U = V$, and the state disturbance B_{0d} where B_{0d} is a matrix containing the parameters relative to θ_0, the preload torque, and Coulomb's friction torque.

The SS system can hence be expressed as

$$\dot{X}_0 = A_0 X_0 + B_0 U + B_{0d}$$
$$Y = C_0 X_0 + DU \tag{4.8}$$

where $C_0 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$, $D = 0$, and A_0, B_0, and B_{0d} are the system matrices whose parameters vary with the variation of the functioning zone following equations (4.6) and (4.7):

In zones $1 - 2F$ and $1 - 2G$:

$$A_0 = \begin{bmatrix} 0 & 1 & 0 \\ \frac{-Nkcont}{J} & \frac{-b}{NJ} & \frac{K\eta_{GB}}{J} \\ 0 & \frac{-K}{L} & \frac{-R}{L} \end{bmatrix} \quad B_0 = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix} \quad B_{0d} = \begin{bmatrix} 0 \\ \frac{Nkcont\theta_0 + NT_{pl} - T_c\mathrm{sgn}(\dot{\theta})}{J} \\ 0 \end{bmatrix}$$

In zones 3F and 3G:

$$A_0 = \begin{bmatrix} 0 & 1 & 0 \\ \frac{-Nktor}{J} & \frac{-b}{NJ} & \frac{K\eta_{GB}}{J} \\ 0 & \frac{-K}{L} & \frac{-R}{L} \end{bmatrix} \quad B_0 = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{L} \end{bmatrix} \quad B_{0d} = \begin{bmatrix} 0 \\ \frac{Nktor\theta_0 - NT_{ph} - T_c\mathrm{sgn}(\dot{\theta})}{J} \\ 0 \end{bmatrix}$$

Table 4.1 shows the numerical values of the constants used in the resulting LPV SS system.

The validity of LPV SS model A_0 depends on its level of fit with AMESim® nonlinear model A from which it was extracted. To study this level of fit, the same voltage signal sent to AMESim® nonlinear model A is sent to

the linear model, and their corresponding angular positions are registered over time. The results are shown in Figure 4.9.
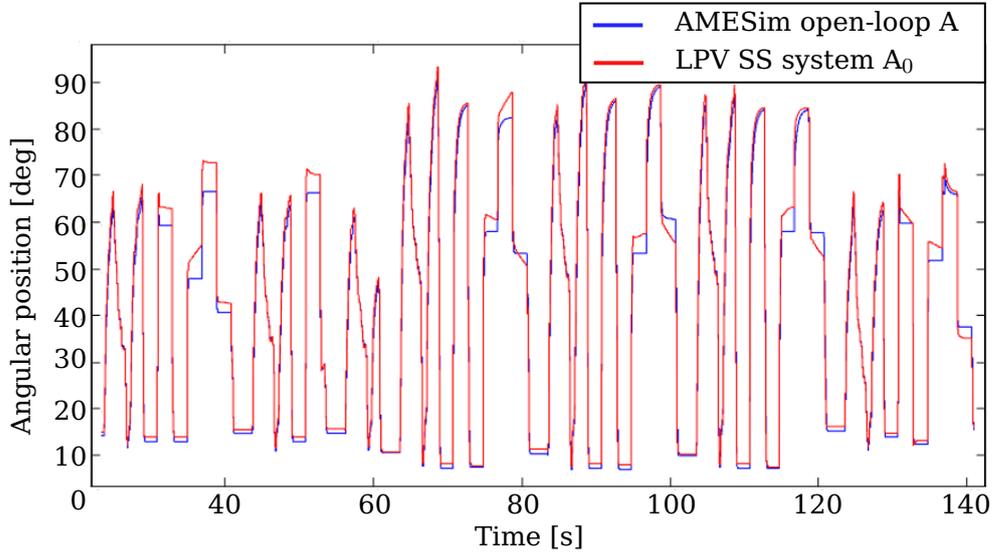
Figure 4.9: Throttle angular position of AMESim® nonlinear model A and linear SS model A_0

Figure 4.9 shows a good fit between the two models, confirmed by an NRMSE equal to 0.921. However, expressing the electric throttle system in the form of SS calls for preliminary knowledge of the physical equations governing the system, thereby eliminating the ability to do an identification by solely relying on the input and the output signals. Nevertheless, SS model A_0 has some advantages. First, it provides the reference regressor structure for the identification of ARX model A described in subsection 4.2.3. Second, it allows to assess the credibility of the ARX identification by comparing the identified parameters with the initial parameters given by the physics-based SS model.

4.2.3 Identification of linear ARX model A

The identified ARX model is a linear regression model of the predictor form [Lju99]

$$\hat{y}(t|\xi) = \phi^T(t)\xi \tag{4.9}$$

where $\hat{y}$ is the predicted output, ϕ is the regression vector expressed as

$$\phi(t) = \begin{bmatrix} \theta(t) & \theta(t-1) & V & \text{sgn}(\dot{\theta}) & 1 \end{bmatrix}^T \tag{4.10}$$

ξ is the vector of parameters to be identified and t is the time.

As previously mentioned, the regression vector ϕ inherits its elements from the discretization of equations (4.6) and (4.7), where $\ddot{\theta}$ and $\dot{\theta}$ in equation (4.6) are discretized using the central difference scheme, and the motor inductance L in equation (4.7) is considered equal to zero due to its negligible impact.

The objective function to be minimized is the prediction error

$$\epsilon(t, \xi) = y(t) - \phi^T(t)\xi \tag{4.11}$$

where y is the output of AMESim® nonlinear model A. The objective function to be minimized is the least squares criterion

$$V_N(\xi, Z^N) = \frac{1}{N} \sum_{t=1}^{N} \frac{1}{2} [y(t) - \phi^T(t)\xi]^2 \tag{4.12}$$

where the time t varies between 1 and N, and Z^N is the data set containing the inputs (in this case V) and the outputs (in this case θ). The linear regression least-squares method is chosen because of the merits it offers. Minimizing the least-squares criterion (4.12) with the linear regression (4.9) is a convex optimization problem, for which we can compute the analytical solution as [Lju99]:

$$\xi_N^{LS} = arg\,min V_N(\xi, Z^N) = R^{-1}(N) * f(N) \tag{4.13}$$

where $R^{-1}(N) = \frac{1}{N} \sum_{t=1}^{N} \phi(t)\phi^T(t)$ and $f(N) = \frac{1}{N} \sum_{t=1}^{N} \phi(t)y(t)$

We suppose that the information content in our input-output data is sufficiently rich to obtain accurate parameters. Four ARX models are identified, one for each of the four functioning zones. The ξ identified vectors are shown in Table 4.2, each corresponding to one of the ARX models.

To verify to what extent ARX model A comprises physical meaning, in comparison to SS mathematical model A_0, these vectors are compared to the initial ξ vectors, shown in Table 4.3, corresponding to the first SS formulation in zones $1-2F$, $1-2G$ and zones 3F, 3G.

Zone $1-2F$	Zone $1-2G$
$\xi_{1-2F} = \begin{bmatrix} 1.9846 \\ -0.9847 \\ 1.577 \cdot 10^{-6} \\ 2.1686 \cdot 10^{-9} \\ 1.2739 \cdot 10^{-5} \end{bmatrix}^T$	$\xi_{1-2G} = \begin{bmatrix} 1.9838 \\ -0.9839 \\ 1.5853 \cdot 10^{-6} \\ -9.8839 \cdot 10^{-7} \\ 1.1535 \cdot 10^{-5} \end{bmatrix}^T$
Zone 3F	Zone 3G
$\xi_{3F} = \begin{bmatrix} 1.9841 \\ -0.9841 \\ 1.6131 \cdot 10^{-6} \\ -2.1601 \cdot 10^{-6} \\ -5.5758 \cdot 10^{-7} \end{bmatrix}^T$	$\xi_{3G} = \begin{bmatrix} 1.9883 \\ -0.9883 \\ 1.1874 \cdot 10^{-6} \\ -5.3499 \cdot 10^{-7} \\ 4.6316 \cdot 10^{-7} \end{bmatrix}^T$

Table 4.2: Least-squares estimates ξ of ARX model A

Zones $1-2F$ and $1-2G$	Zones 3F and 3G
$\xi_{1-2F,1-2G} = \begin{bmatrix} 1.9872 \\ -0.9879 \\ 3.1915 \cdot 10^{-5} \\ -1.5198 \cdot 10^{-5} \\ 2.4443 \cdot 10^{-4} \end{bmatrix}^T$	$\xi_{3F,3G} = \begin{bmatrix} 1.9872 \\ -0.9873 \\ 3.1915 \cdot 10^{-5} \\ -1.5198 \cdot 10^{-5} \\ 1.8374 \cdot 10^{-5} \end{bmatrix}^T$

Table 4.3: Least-squares estimates ξ of SS model A_0

The vectors stemming from ARX model A identification mostly fall in the same range as those corresponding to the SS formulation. Moreover, Figure 4.10 shows how well the output of the identified ARX model A fits the output of AMESim® nonlinear model A. In fact, this model shows a higher level of fit with AMESim® nonlinear model A compared to SS model A_0 because the NRMSE increases from 0.921 in the case of SS to 0.9672 in the case of ARX. This can be interpreted by the fact that the SS model explicitly implies the physical laws presented in the mathematical equations, whereas the ARX model is not restricted to these equations and therefore keeps room for expressing obscure physical phenomena. The proximity of the ξ vectors from those of SS model A_0 in addition to the good level of fit shown in Figure 4.10, both prove that ARX model A can in fact be utilized to identify a linear model of an electric throttle just by taking into account its inputs and outputs. This is substantial and helpful because it implies that it is possible to accelerate the modeling process. In other words, it is possible to surpass the AMESim® model by identifying immediately a linear model using inputs and outputs coming from a test-bench experiment. This process does not only save time, but it also makes room for linear control techniques in electric throttle applications.
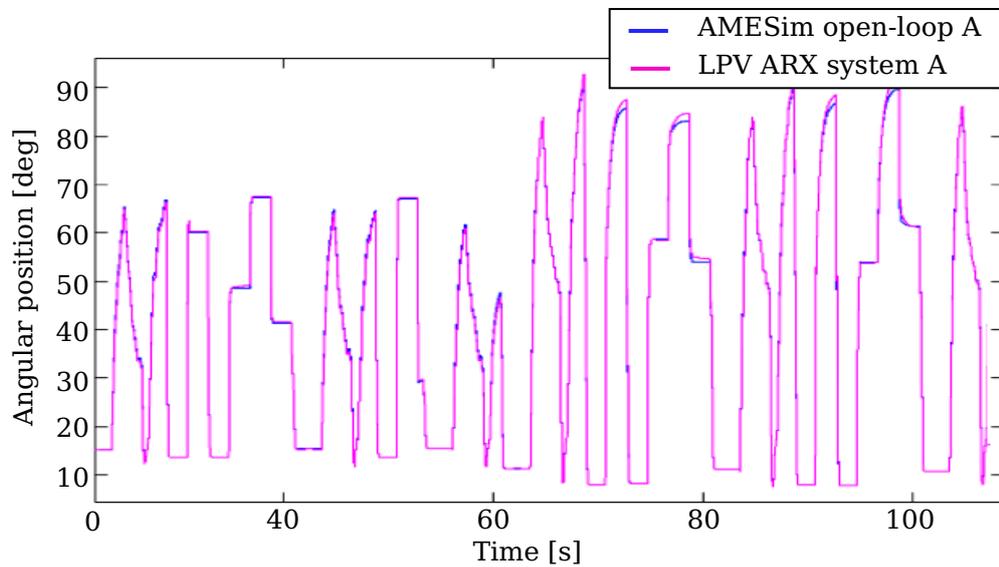
Figure 4.10: Throttle angular position of identified LPV ARX model A

In view of the criticality of the electric throttle's limp-home position, ARX model A is also verified on a simulation shown in Figure 4.11 where the position of the throttle varies slightly around the limp-home posi-

tion, lying at around $18°$. A good fit between AMESim® model A and ARX model A can be noted and verified with an NRMSE equal to 0.9520. Also, the passages through limp-home position do not show any blatant discontinuities.
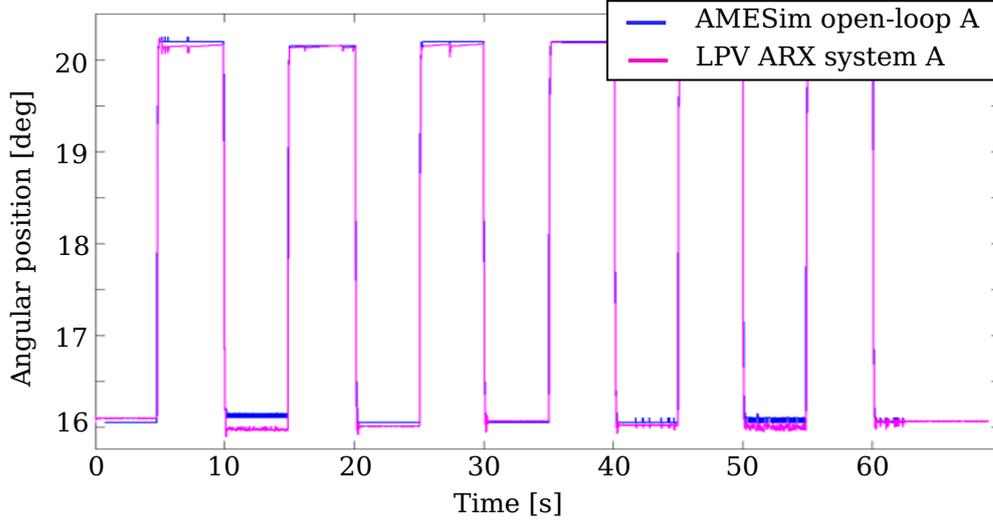
Figure 4.11: Around-limp-home throttle angular position of AMESim® nonlinear model A and identified LPV ARX model A

Finally, to facilitate the control design (presented in section 4.4), ARX model A is converted to a SS model, referred to as "SS model A". Forward Euler method is applied to transform the discrete model into a continuous one, and the third state i is omitted from the SS vector because the inductance L is negligible, and can hence be approximated by 0. This approximation removes from (4.7) the term corresponding to the dynamics of the current.

The resulting SS systems have the following form:

$$\dot{X} = AX + BU + B_d$$
$$Y = CX + DU \tag{4.14}$$

where $X = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix}$, $A = \begin{bmatrix} 0 & 1 \\ \frac{\xi_{i1}+\xi_{i2}-1}{dt^2} & \frac{\xi_{i1-2}}{dt} \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ \frac{\xi_{i3}}{dt^2} \end{bmatrix}$,

$B_d = \begin{bmatrix} 0 \\ \frac{\xi_{i4}\text{sgn}(\dot{\theta})+\xi_{i5}}{dt^2} \end{bmatrix}$, $C = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $D = 0$, and i is the model number

($i = 1 - 2F$, $1 - 2G$, $3F$, $3G$).

4.3 Electric throttle test-bench and experimental results

The work in the previous section motivated us to build a test-bench of the electric throttle, shown in Figure 4.12, to allow testing the methodology and the model presented earlier. The test-bench consists of different components: the H4Bt Renault engine electric throttle from Continental Edison (orange), an Arduino Mega R3 2560 microcontroller board (violet), an L298N H-bridge motor driver (green), an secure digital (SD) card (blue), a current sensing resistor with an RC filter (grey), and a power supply (black). The connections between these components are depicted in Figure 4.13.
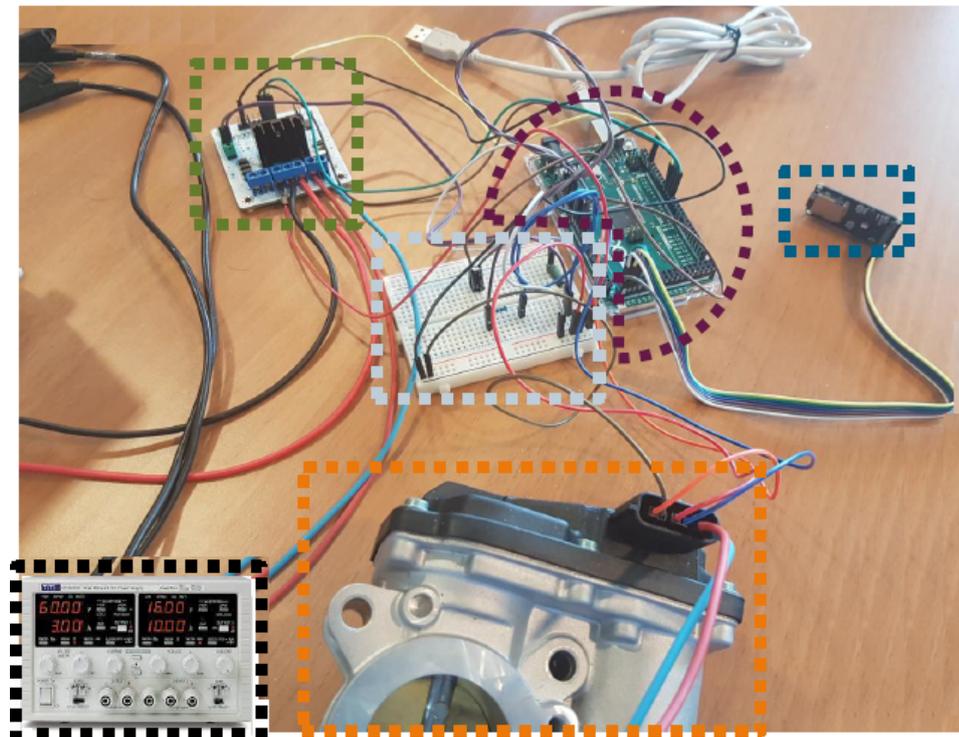
Figure 4.12: Electric throttle test-bench

In order to actuate the opening of the throttle valve, the microcontroller sends the H-bridge motor driver two logical values specifying the direction of rotation of the throttle motor, and an input voltage between 0 and 5 volts specifying the valve's angular position. The throttle valve motor needs to be supplied with an input voltage between 0 and 12 volts. Therefore, taking into consideration that the L298N H-bridge motor driver
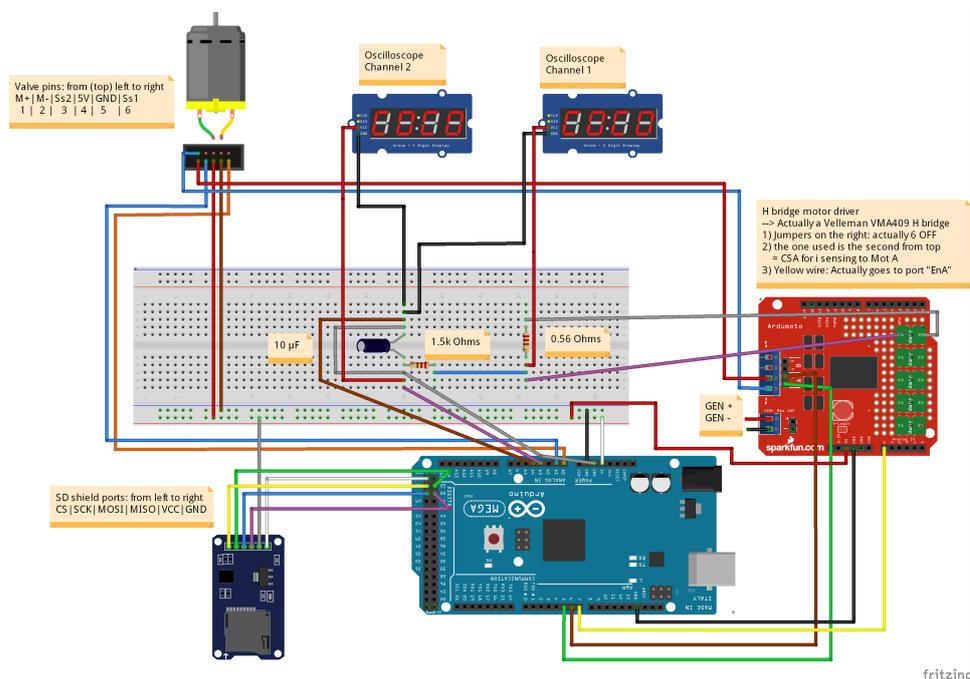
Figure 4.13: Connections of the electric throttle test-bench

induces a voltage drop of approximately $2\,\mathrm{V}$, a generator supplying a voltage of $14.5\,\mathrm{V}$ is used. The output angular position of the throttle valve, between 0 and 90 degrees, is measured by means of the throttle sensors which receive their $5\,\mathrm{V}$ power supply from the microcontroller board. The input and output data are saved on the SD card which communicates with the microcontroller using the serial peripheral interface (SPI) communication protocol. In addition, a current sensing resistor is used to sense, at a frequency of about $2\,\mathrm{kHz}$, the current going through the motor. This high frequency is incompatible with that of a real-time simulation, which is why an RC filter is plugged onto the sensing resistor to ensure an appropriate sampling. With a resistance $R = 1.5\,\mathrm{k\Omega}$ and a capacitance $C = 10\,\mathrm{\mu F}$, the frequency of the RC filter is equal to $f_c = 1/(2\pi RC) \approx 10\,\mathrm{Hz}$. This frequency is well-suited for the real-time simulation and the SD card logging frequency ranging between $20\,\mathrm{Hz}$ to $100\,\mathrm{Hz}$.

4.3.1 Nonlinear AMESim® model B and its nonlinearities

Though AMESim® model A was verified by Renault engineers on a test-bench in Renault, its evaluation on the test-bench described above re-

vealed unsuccessful, basically for two reasons. First, it does not cover
all the nonlinearities of the real throttle test-bench, as shown in subsec-
tion 4.2.2. Second, it assumes that the "high" spring is only active above
the limp-home position and not everywhere in the functioning zone. For
these reasons, a new AMESim® model, referred to as AMESim® model
B, is designed. It is the adjusted or rectified version of model A, compat-
ible with the throttle bench.
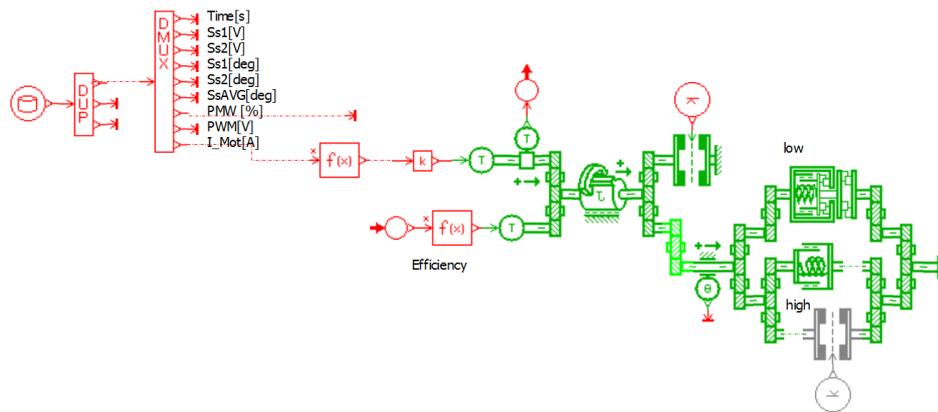
Figure 4.14: AMESim® model B

AMESim® model B, shown in Figure 4.14, is slightly different from
AMESim® model A, presented earlier. The differences between both
models are summed up in the list below.

- The principal components of AMESim® model B are the 2-sided
  inertia, the gearbox, the throttle flap, the "low" spring, and the
  "high" spring. The friction associated with the inertial component
  (to its upper-right) and that associated with the "high" spring (be-
  low it) are put in separate components rather than being embed-
  ded in the components themselves, which accounts for two extra
  components compared to AMESim® model A.

- AMESim® model B models the electric throttle from a pure me-
  chanical point of view. Hence, as opposed to AMESim® model A,
  AMESim® model B does not contain a DC motor. It also does not
  take into consideration the electrical parts of the test-bench such
  as the current sensing resistor and the RC filter.

- AMESim® model B considers the viscous friction at the level of
  the "high" and "low" return springs, whereas AMESim® model A
  considers it at the level of the inertial component.

- In contrast to AMESim$^®$ model A which restricts the functioning of the "high" spring to the zone where $\theta \geq \theta_0$, AMESim$^®$ model B features a "high" spring that is equally active below and above the limp-home position.

- The contact torque corresponding to the "low" spring is not neglected in AMESim$^®$ model B, as is the case in AMESim$^®$ model A. However, the "low" spring torque is not visited, which means that it spans zones 2 and 3 from Figure 4.6.

- Whereas AMESim$^®$ model A considers the Coulomb friction model, AMESim$^®$ model B considers the reset-integrator friction model, which means that it spans zones A, B, C, D, and E from Figure 4.6.

- The span of functioning zones covered by AMESim$^®$ model B makes it more detailed than AMESim$^®$ model A, covering 10 zones of instead of 4.

AMESim$^®$ model B and the electric throttle test-bench introduced earlier show an important proximity in their open-loop results, shown in Figure 4.15. When the same input voltage is given to model B and to the test-bench, the output position simulated by model B shows being very close to that measured on the real throttle. Note that AMESim$^®$ model B was compared with the test-bench using a scenario different than that shown in Figure 4.8, as the latter is not testable on the test-bench, because the input profile contains high frequencies that are incompatible with a real-time simulation, and the output profile features partial opening positions of the valve. The valve's opening profile gives insight about its frictional behavior and high nonlinearity, originating from the "high" and "low" springs preload torques, the contact torque, the limp-home position θ_0, and the Coulomb torque. Starting from the limp-home position, the throttle valve shows no change in position as the voltage increases, until it gets close to $7\,\text{V}$. At this threshold, it opens all the way from the limp-home position to its maximum opening position, and it maintains this position despite further increase in the voltage input. The same scenario repeats itself with the decrease of the voltage input. As soon as the input decreases below $5\,\text{V}$, the valve returns back to the limp-home position, thus showing that the throttle valve in the test-bench cannot open partially.

As mentioned earlier, AMESim$^®$ model B calculates $T_{friction}$ using the reset-integrator model, which implies that its functioning should normally span 15 zones: 1A-1E, 2A-2E, and 3A-3E (referring to Figure 4.6).
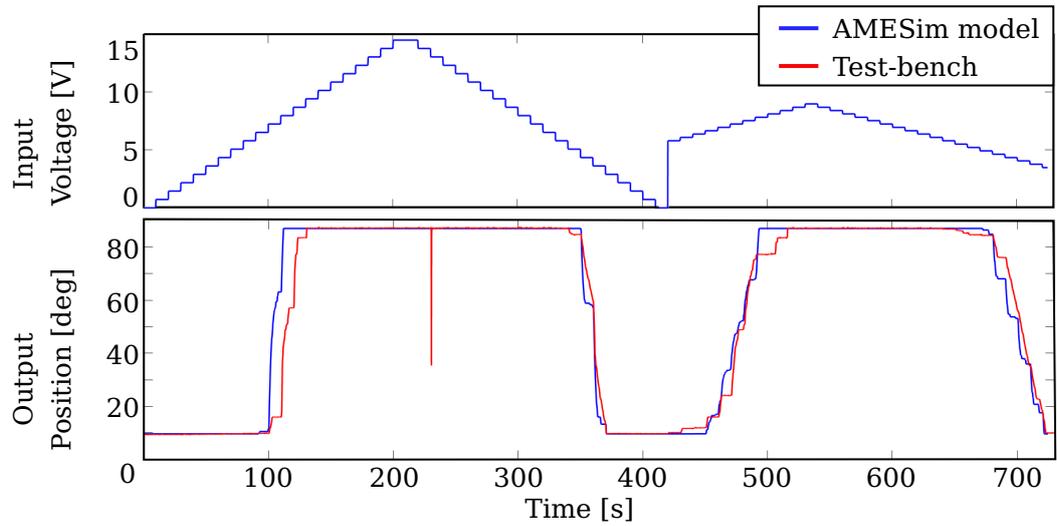
Figure 4.15: Voltage input (upper plot) and position output (lower plot) of the electric throttle

This means that writing this model in the form of an LPV model requires four scheduling parameters: θ, $\dot{\theta}$, θ_{diff}, and $T_{friction}$.

A remark can be made in this context about the functioning zones of AMESim® model B, whose parameters are shown in Table 4.4. When $\theta < \theta_0$, the contact torque attained during all the scenarios used to validate AMESim® model B lies between $0\,\mathrm{N\,m}$ and $0.27\,\mathrm{N\,m}$, as shown in the lower plot of Figure 4.18 (showing the "low" spring torque). Hence, its absolute value never exceeds $T_{pl} = 1.36\,\mathrm{N\,m}$, which implies that T_{spring} is never solicited. This is also exposed in Figure 4.16, which shows, between the blue dashed lines, the range of torque in which falls $T_{contact}$.

This remark points out the fact that AMESim® model B does not solicit zone 1, and is restricted to zones 2 and 3. Consequently, the total number of zones visited by this model reduces from 15 to 10. These 10 zones are split at the points presenting the nonlinearities originating from the limp-home position, the Coulomb friction torque, the contact torque, and the stiction torque, as follows:

- zone 2A: $\theta_0 - disth \leq \theta < \theta_0$, $\dot{\theta} > 0$, and $\theta_{diff} \geq dtrel$

- zone 2B: $\theta_0 - disth \leq \theta < \theta_0$, $|\theta_{diff}| < dtrel$, and $T_{friction} < -T_s$

- zone 2C: $\theta_0 - disth \leq \theta < \theta_0$, $|\theta_{diff}| < dtrel$, and $|T_{friction}| \leq T_s$

- zone 2D: $\theta_0 - disth \leq \theta < \theta_0$, $|\theta_{diff}| < dtrel$, and $T_{friction} > T_s$

J	$3.29 \times 10^{-6}\,\text{kg}\,\text{m}^2$	$kint$	$77\,922.26\,\text{N}\,\text{m}\,\text{rad}{-}1$
N	0.0456	$kcont$	$0.086\,\text{N}\,\text{m}\,\text{rad}{-}1$
K	$0.0235\,\text{N}\,\text{m}\,\text{A}^{-1}$	$rcont$	$0\,\text{N}\,\text{m}\,\text{s}\,\text{rad}^{-1}$
η_{GB}	0.95	r	$5.73 \times 10^{-6}\,\text{N}\,\text{m}\,\text{s}\,\text{rad}^{-1}$
T_c	$0.004\,\text{N}\,\text{m}$	e	$1.745 \cdot 10^{-5}\,\text{rad}$
T_s	$0.0048\,\text{N}\,\text{m}$	$disth$	$1.745 \times 10^{-5}\,\text{rad}$
T_{pl}	$1.36\,\text{N}\,\text{m}$	$krel$	$2.75\,\text{N}\,\text{m}\,\text{rad}{-}1$
T_{ph}	$0.259\,\text{N}\,\text{m}$	$dtrel$	$0.0017\,\text{rad}$
$ktor$	$0.056\,\text{N}\,\text{m}\,\text{rad}{-}1$	$rtors$	$57.29\,\text{N}\,\text{m}\,\text{s}\,\text{rad}^{-1}$
$rtor$	$0.0955\,\text{N}\,\text{m}\,\text{s}\,\text{rad}^{-1}$	θ_0	$0.1675\,\text{rad}$

Table 4.4: AMESim® model B parameters
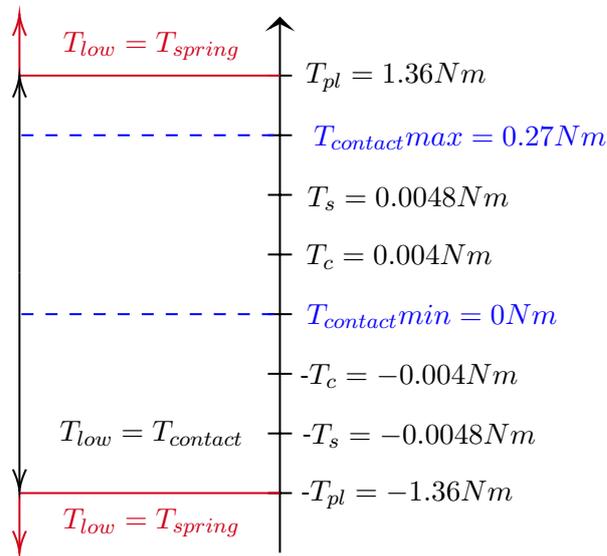
Figure 4.16: Torque range of $T_{contact}$

Component	Torque	Equation	Figure
Inertia	T_{motor}	(4.1)	4.17
High spring	T_{high}	(4.2)	4.18
Low spring	T_{low}	(4.3)	4.18
Friction on motor	$T_{friction}$	(4.5)	4.19
Friction on high spring	$T_{friction}$	(4.5)	(deactivated)

Table 4.5: Torque calculation of AMESim$^{®}$ model B

- zone 2E: $\theta_0 - disth \leq \theta < \theta_0$ and $\dot{\theta} < 0$, and $\theta_{diff} \leq -dtrel$

- zone 3A: $\theta_0 \leq \theta \leq \theta_{max}$, $\dot{\theta} > 0$, and $\theta_{diff} \geq dtrel$

- zone 3B: $\theta_0 \leq \theta \leq \theta_{max}$, $|\theta_{diff}| < dtrel$, and $T_{friction} < -T_s$

- zone 3C: $\theta_0 \leq \theta \leq \theta_{max}$, $|\theta_{diff}| < dtrel$, and $|T_{friction}| \leq T_s$

- zone 3D: $\theta_0 \leq \theta \leq \theta_{max}$, $|\theta_{diff}| < dtrel$, and $T_{friction} > T_s$

- zone 3E: $\theta_0 \leq \theta \leq \theta_{max}$ and $\dot{\theta} < 0$, and $\theta_{diff} \leq -dtrel$

For what follows, consider Figure 4.14 showing AMESim$^{®}$ model B. We discuss the mathematical equations, representing the torques stemming from the different components in this model, the total torque applied on the motor, in addition to the applicability of Newton's Second Law on the inertial component.

Starting with the mathematical equations, Table 4.5 shows the component name, its torque, equation, and the figure showing the plot of the torque. Note that T_{motor} is the torque to the left of the inertial component. The torque stemming from the "high" spring T_{high} and that stemming from the "low" spring T_{low} are the spring torques calculated at the level of the throttle. The torque stemming from the friction component $T_{friction}$ is that resulting from stiction and Coulomb friction and calculated at the level of the inertial component. $T_{friction}$ corresponding to the high spring is equal to zero because, although the friction component is there, it is deactivated and has no effect.

Notice from the middle plot of Figure 4.18 how T_{high} is always active independently of the value of θ, whereas T_{low} exists only when $\theta < \theta_0$. Concerning $T_{friction}$, Figure 4.19 shows how the friction torque becomes saturated as soon as its absolute value reaches $T_s = 0.0048\,\mathrm{N\,m}$. It also shows, though not clearly, how the absolute value of the friction

torque becomes equal to the Coulomb torque as the absolute value of θ_{diff} reaches $dtrel$ at certain points.
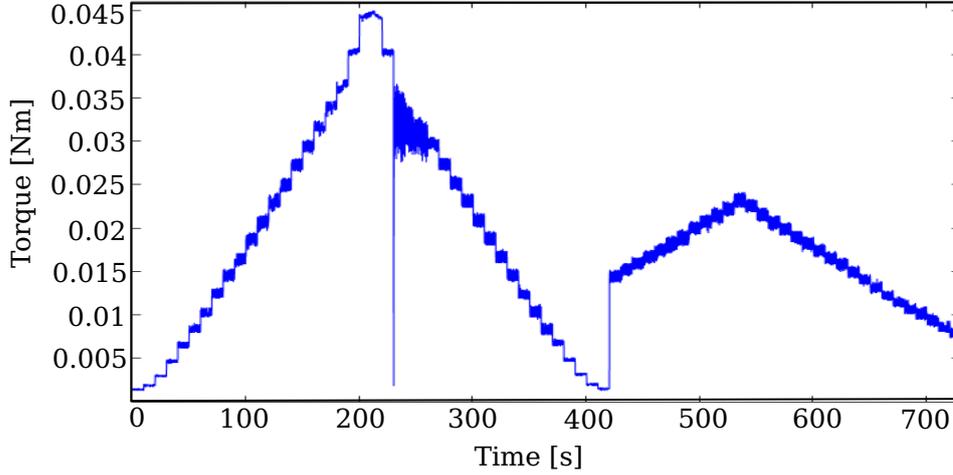
Figure 4.17: Motor Torque

The total torque stemming from the two springs and the friction component combined can be calculated as:

$$T_{springsfriction} = \begin{cases} (T_{high} - T_{low})N + T_{friction} & , if\ \theta < \theta_0 \\ T_{high}N + T_{friction} & , if\ \theta \geq \theta_0 \end{cases} \quad (4.15)$$

$T_{springsfriction}$ is the torque to the right of the inertial component, and its plot is shown in Figure 4.20. T_{high} and T_{low} are multiplied by the gear ratio N because they are generated on the level of the springs.

Knowing the torque T_{motor} applied on the inertial component from the left and $T_{springsfriction}$ applied on the inertial component from the right, the calculation of the angular acceleration is supposed to be trivial. However, the link between the sum of torques and the angular acceleration of the motor is not straightforward, because of the saturability of Newton's Second Law of Rotation.

To demonstrate the saturability of Newton's Second Law on the inertial component of AMESim® model B, imagine the inertial component isolated from the rest of the model, and consider only the torques applied on it from the left T_{motor} and from the right $T_{springsfriction}$. Figure 4.21 shows that the angular acceleration of the motor computed automatically by AMESim® model B is not equal to that computed using Newton's Second Law $\ddot{\theta} = (T_{motor} - T_{springsfriction})/J$.
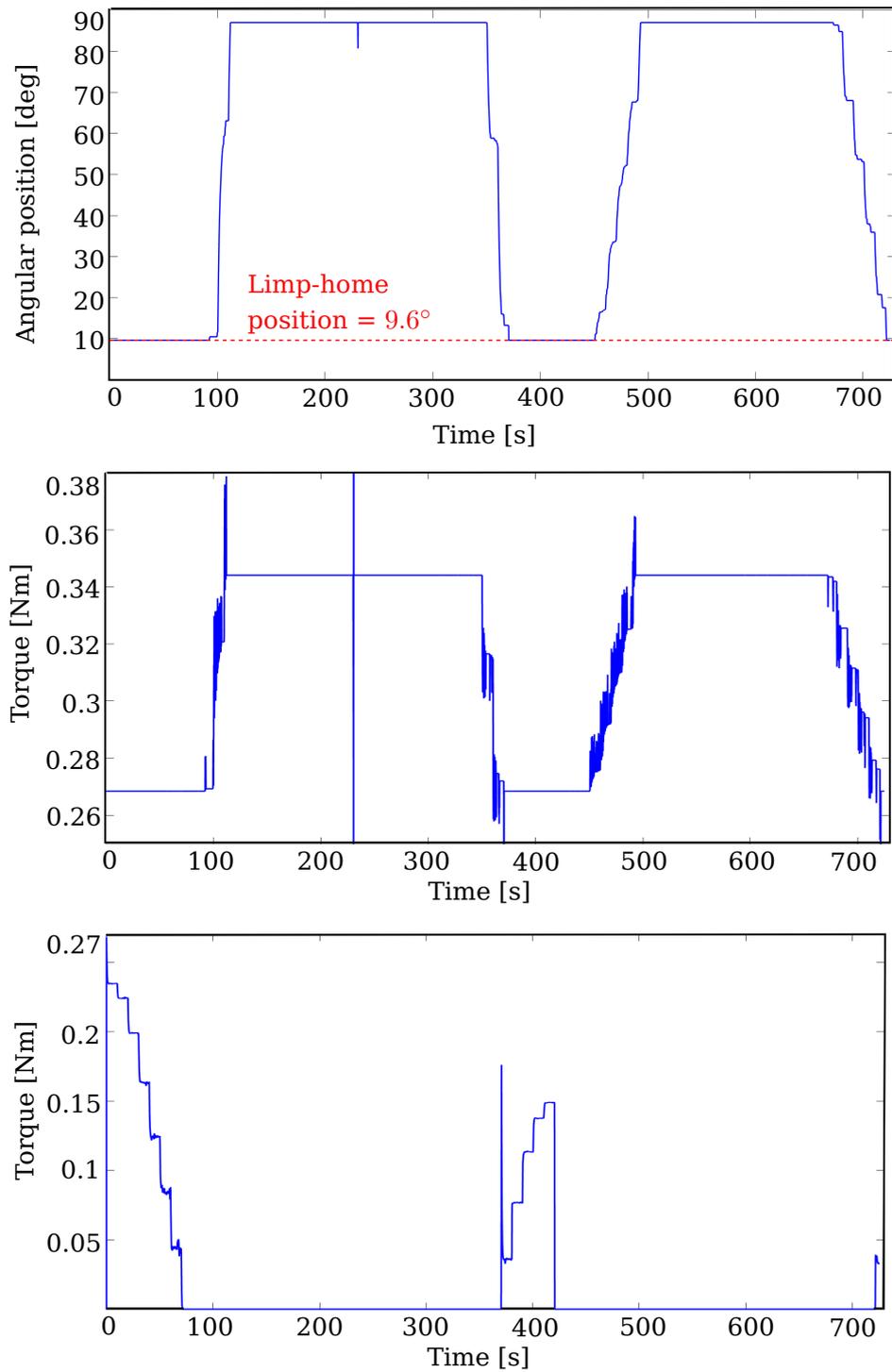
Figure 4.18: Throttle angular position (upper plot); Torque stemming from the "high" spring (middle plot); Torque stemming from the "low" spring (lower plot)
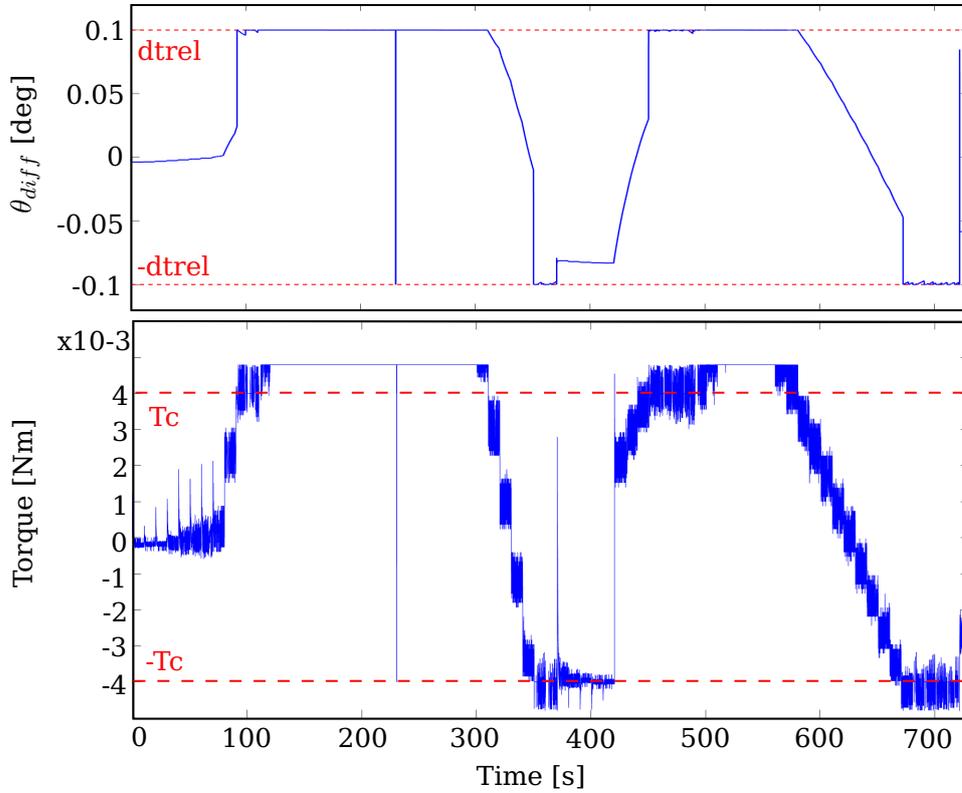
Figure 4.19: Torque stemming from the friction component

The reason behind this inequality is the ideal endstop system of the inertial component, which works in a 2-step process:

- Step 1: Ignoring the effect of dry friction, the angular displacement θ_{motor} and velocity $\dot{\theta}_{motor}$ of the motor are computed. If the angular displacement reaches an endstop, i.e. attempts to exceed its limit values $x_{min} = 0°$ and $x_{max} = 1905°$, it is saturated and its angular velocity is forced at $0\,\mathrm{rad\,s^{-1}}$. Note that $x_{max} = 1905°$ on the level of the motor is translated into $\theta_{max} = 87°$ on the level of the throttle due to the gearbox of ratio N.

- Step 2: If none of the endstops is reached, the angular displacement θ_{motor}, speed $\dot{\theta}_{motor}$, and acceleration $\ddot{\theta}_{motor}$ of the motor are computed, taking into account the dry friction. The following equation shows the computation of $\ddot{\theta}_{motor}$ and the conditions of applicability of Newton's Second Law:
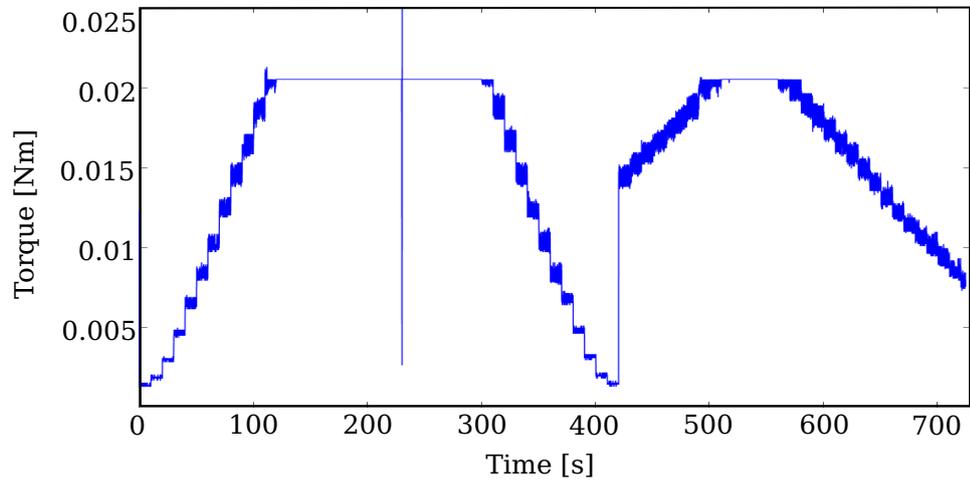
Figure 4.20: Torque stemming from the two springs and the friction component combined
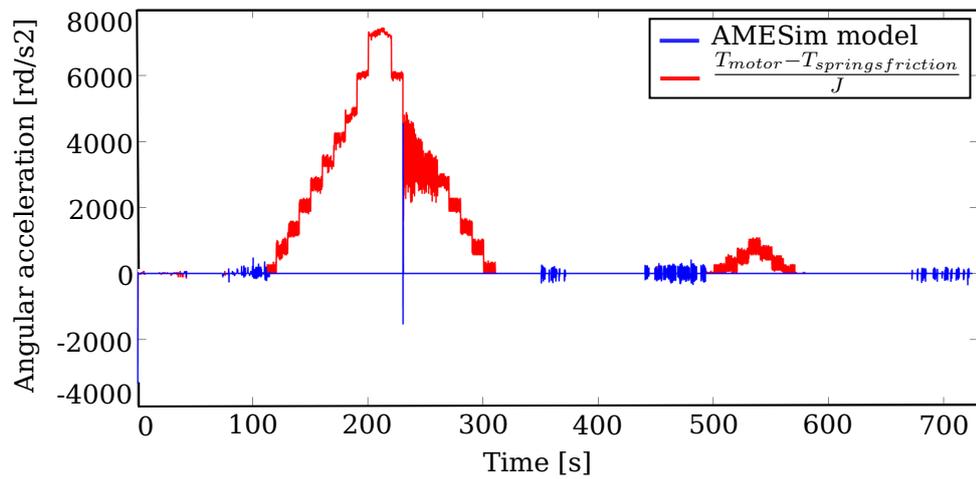
Figure 4.21: Angular acceleration of the motor

$$\ddot{\theta}_{motor} = \begin{cases} \frac{T_{motor} - T_{springs friction}}{J} & , if \, |T_{motor} - T_{springs friction}| > T_s \\ & or \, \dot{\theta}_{motor} \neq 0 \\ 0 & , if \, |T_{motor} - T_{springs friction}| \leq T_s \\ & and \, \dot{\theta}_{motor} = 0 \end{cases}$$

(4.16)

This fact highlights an additional merit of ARX modeling, which short-cuts the calculation of θ without being hindered by the saturability of Newton's Second Law or by the numerical manipulations carried out by the AMESim® simulator during the conversion of $\ddot{\theta}$ to θ. ARX model B corresponding to the test-bench is detailed in the next subsection.

4.3.2 Identification of linear ARX model B

The mathematical equations governing the dynamics of AMESim® model B can be used to conclude two things: the number and repartition of the functioning zones, and the elements of the regression vector. These elements are extracted from each torque as shown in the list below.

- From T_{motor}, extract the element i. However, i and the input voltage V are directly related together, either by virtue of Kirchoff's law, or the current sensing resistor. This is why i will be replaced by V in the regression vector, as it is the input of the system.

- From T_{high}, extract θ, $\dot{\theta}$, and 1 to stand for the constant term.

- From T_{low}, extract θ, $\dot{\theta}$, $\exp(\theta)$, and 1.

- From $T_{friction}$, extract θ_{diff}, $\dot{\theta}$, and 1. Taking into account that θ_{diff} is either a constant or a function of θ, it is excluded from the list of chosen elements.

By discretizing $\ddot{\theta}$ and $\dot{\theta}$ using the central difference scheme, the regression vector can be written as:

$$\phi(t) = \begin{bmatrix} \theta(t) & \theta(t-1) & V & exp(\theta(t)) & 1 \end{bmatrix}^T$$

(4.17)

Using the same estimation technique presented in section 4.2.3, 10 ARX models are identified, one for each of the 10 functioning zones. The values of θ, $\dot{\theta}$, θ_{diff}, and $T_{friction}$ used to split the zones are those issued by AMESim® model B. These identifications give rise to the ξ vectors presented in Table 4.6.

The experimental data used for identification is presented in Figure 4.15, which shows in the upper plot the voltage input that is given to the real throttle as well as to AMESim® model B, and shows in the lower plot the position output of the real throttle (red) and of AMESim® model B (blue). To assess the ARX system identification, Figure 4.22 shows, as well as the position outputs of AMESim® model B and the test-bench, the position output of the identified ARX model B (green). The goodness of fit between the output of ARX model B and that of the electric throttle test-bench is expressed with an NRMSE equal to 0.7993. This is considered a good fit, taking into consideration that ARX model B is extracted from AMESim® model B, whose own output compared with that of the test-bench has an NRMSE equal to 0.7925.
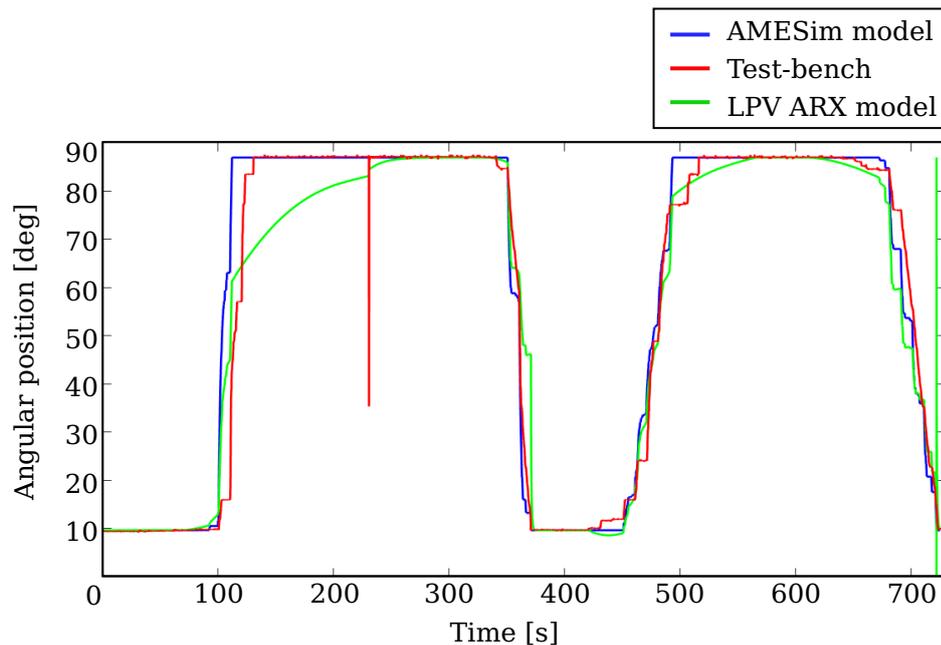
Figure 4.22: Voltage input and position output of the electric throttle

Note that this experiment does not solicit all the functioning zones of the electric throttle. Figure 4.23 shows the repartition of the functioning zones throughout the experiment. Only zones 2C, 3A, 3B, 3C, 3D, and 3E are solicited. In other words, when $\theta \geq \theta_0$, all the torque calculation

Zone 2C	Zone 3A
$\xi_{2C} = \begin{bmatrix} 0.18259 \cdot 10^{-1} \\ 0.86371 \cdot 10^{-1} \\ 0.67548 \cdot 10^{-10} \\ -0.39618 \cdot 10^{-2} \\ 0.39086 \cdot 10^{-2} \end{bmatrix}^T$	$\xi_{3A} = \begin{bmatrix} 1.9641 \\ -0.96414 \cdot 10^{-1} \\ 0.18531 \cdot 10^{-5} \\ 0.25441 \cdot 10^{-6} \\ -0.11326 \cdot 10^{-4} \end{bmatrix}^T$
Zone 3E	Zone 3C
$\xi_{3E} = \begin{bmatrix} 1.9763 \\ -0.97631 \cdot 10^{-1} \\ -0.40508 \cdot 10^{-6} \\ -0.88021 \cdot 10^{-6} \\ 0.95418 \cdot 10^{-6} \end{bmatrix}^T$	$\xi_{3C} = \begin{bmatrix} 1.8989 \\ -0.89886 \cdot 10^{-1} \\ 0.92807 \cdot 10^{-8} \\ -0.24116 \cdot 10^{-6} \\ 0.15978 \cdot 10^{-6} \end{bmatrix}^T$
Zone 3D	Zone 3B
$\xi_{3D} = \begin{bmatrix} 1.5010 \\ -0.49089 \cdot 10^{-1} \\ 0.27359 \cdot 10^{-13} \\ -0.40596 \cdot 10^{-3} \\ 0.31450 \cdot 10^{-3} \end{bmatrix}^T$	$\xi_{3B} = \begin{bmatrix} -0.15427 \cdot 10^{-6} \\ 1.9461 \\ 0.17212 \cdot 10^{-2} \\ -0.10042 \cdot 10^{-1} \\ -0.99782 \cdot 10^{-2} \end{bmatrix}^T$

Table 4.6: Least-squares estimates ξ of ARX model B

possibilities are visited, which is not the case when $\theta < \theta_0$.
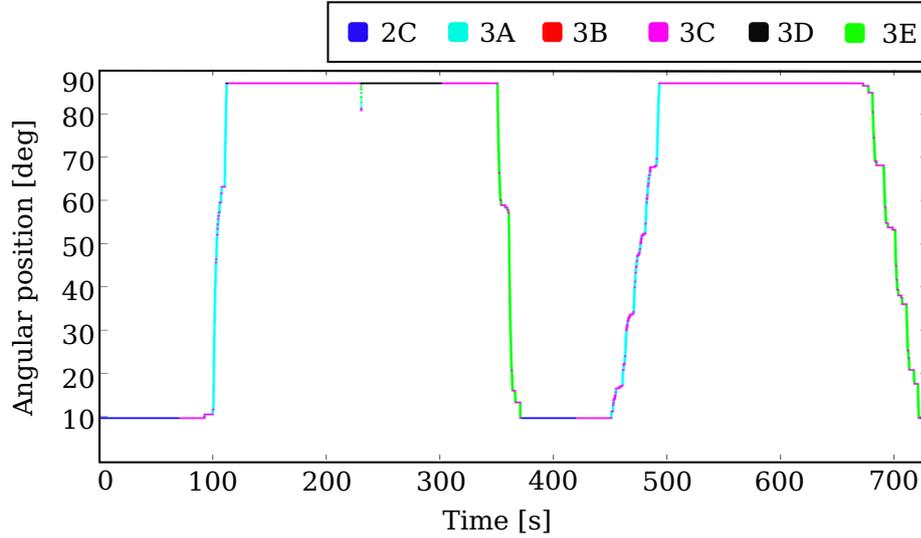
Figure 4.23: Repartition of the functioning zones during the reference experiment

Moreover, notice how the angular position generated by AMESim® model B, as well as that generated by the linear ARX model B never exceed $\theta_{max} = 87°$. The fact that the outputs are strictly limited to this position is an indication that both models respect the physical limits of the electric throttle.

Reversing the central difference discretization in ARX model B, the ARX discrete model B can be transformed into SS model B expressed as:

$$\begin{bmatrix} \dot{\theta} \\ \ddot{\theta} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{\xi_{i1}+\xi_{i2}-1}{dt^2} & \frac{\xi_{i1-2}}{dt} \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\xi_{i3}}{dt^2} \end{bmatrix} V + \begin{bmatrix} 0 \\ \frac{\xi_{i4}\exp(\theta)+\xi_{i5}}{dt^2} \end{bmatrix} 1 \qquad (4.18)$$

where i is the model number ($i = $ 2C, 3A, 3B, 3C, 3D, 3E).

4.4 Control of the throttle LPV model

In this section, the theory of the infinite-horizon LQR is briefly presented, the LQR controller initially designed on linear SS model A is detailed and evaluated on AMESim® model A.

4.4.1 Linear quadratic regulator with reference and disturbance feedforwards

Using SS model A, a model-based LQR controller is designed to control the voltage input of the throttle, which in turn affects the angular position of the flap. The objective of the controller is to drive the output of the LPV system towards the reference output. In addition to the tracking objective, the controller should also compensate for the process noise which appears in the term B_d. For these two reasons, the designed LQR is equipped with reference feedforward and disturbance feedforward.

Considering system (4.14), the objective of the LQR is to minimize the cost function

$$J = \frac{1}{2} \int_0^\infty [(X - X_{ref})^T Q (X - X_{ref}) + U^T R U] dt \qquad (4.19)$$

where X_{ref} is the reference state $\begin{bmatrix} \theta_{ref} & \dot{\theta}_{ref} \end{bmatrix}^T$, θ_{ref} is the reference angular position (rad), Q is the state weighting matrix, and R is the input weighting matrix.

The Hamiltonian can be expressed as

$$H = \frac{1}{2}(X - X_{ref})^T Q (X - X_{ref}) + \frac{1}{2} U^T R U + p^T \dot{X} \qquad (4.20)$$

where p is the co-state vector and p^* is its optimal value. Note that the "$*$" is used to denote the optimal value of a parameter in the sequel.

The necessary optimality conditions can be expressed as [Kir04]

- $\dot{p}^* = \frac{\partial H}{\partial X}$

- $\frac{\partial H}{\partial U} = 0$

Therefore,

$$U^* = -R^{-1}B^T p^* = -R^{-1}B^T (KX^* + S_1 + S_2) \qquad (4.21)$$

where K is the solution of the differential Ricatti equation,

$$\dot{K} + KA - KBR^{-1}B^T K + A^T K + Q = 0 \qquad (4.22)$$

S_1 is the reference feedforward term, obtained from the following equation,

$$\dot{S}_1 + A^T S_1 - KBR^{-1}B^T S_1 - QX_{ref} = 0 \qquad (4.23)$$

and S_2 is the disturbance feedforward term and the solution of

$$\dot{S}_2 + A^T S_2 - KBR^{-1}B^T S_2 + KB_d = 0 \qquad (4.24)$$

Because the horizon is considered to be infinite, the dynamics of K, S_1, and S_2 can be considered null as time goes to infinity, i.e. $\dot{K} \to 0$, $\dot{S}_1 \to 0$, and $\dot{S}_2 \to 0$. Let J be the cost function expressed as:

$$J = \frac{1}{2} \int_0^{t_f} [(Y - Y_{ref})^T W (Y - Y_{ref}) + U^T RU] dt \qquad (4.25)$$

where W is a weighting constant, which implies that $Q = C^T W C$.

4.4.2 Control of linear model A

The control is designed on LPV SS model A converted from ARX model A in a way that for each of the models, a controller is designed separately. The controllers' parameters are:

$Q = C^T W C$ where $W = 30000$

$R = I$ where I is a (2×2) identity matrix

The solutions of equation (4.22) are

$$K_{1-2G} = \begin{bmatrix} 154 & 0.7 \end{bmatrix} \quad K_{1-2F} = \begin{bmatrix} 152.3 & 0.7 \end{bmatrix}$$

$$K_{3G} = \begin{bmatrix} 171 & 1 \end{bmatrix} \qquad K_{3F} = \begin{bmatrix} 170.9 & 0.8 \end{bmatrix}$$

The solutions of equation (4.23) are

$$S_{1\,1-2G} = \begin{bmatrix} -297 & 0 \\ -1 & 0 \end{bmatrix} \quad S_{1\,1-2F} = \begin{bmatrix} -291.7 & 0 \\ -1 & 0 \end{bmatrix}$$

$$S_{1\,3G} = \begin{bmatrix} -339.8 & 0 \\ -1.5 & 0 \end{bmatrix} \quad S_{1\,3F} = \begin{bmatrix} -304.2 & 0 \\ -1.1 & 0 \end{bmatrix}$$

The solutions of equation (4.24) are

$$S_{2\,1-2G} = \begin{bmatrix} 5.4 \\ 0.04 \end{bmatrix} \quad S_{2\,1-2F} = \begin{bmatrix} 6.1 \\ 0.05 \end{bmatrix}$$

$$S_{2\,3G} = \begin{bmatrix} -0.06 \\ 0 \end{bmatrix} \quad S_{2\,3F} = \begin{bmatrix} 0.9 \\ 0 \end{bmatrix}$$

The objective of the control is to drive the output of linear SS model A, which fits almost perfectly the output of AMESim$^{\circledR}$ model A (shown in blue in Figure 4.24) toward the reference output, also called the set-point (shown in red in Figure 4.24). When the control takes effect, this objective is attained. This can be shown by the almost exact superposition of the green curve over the red one.
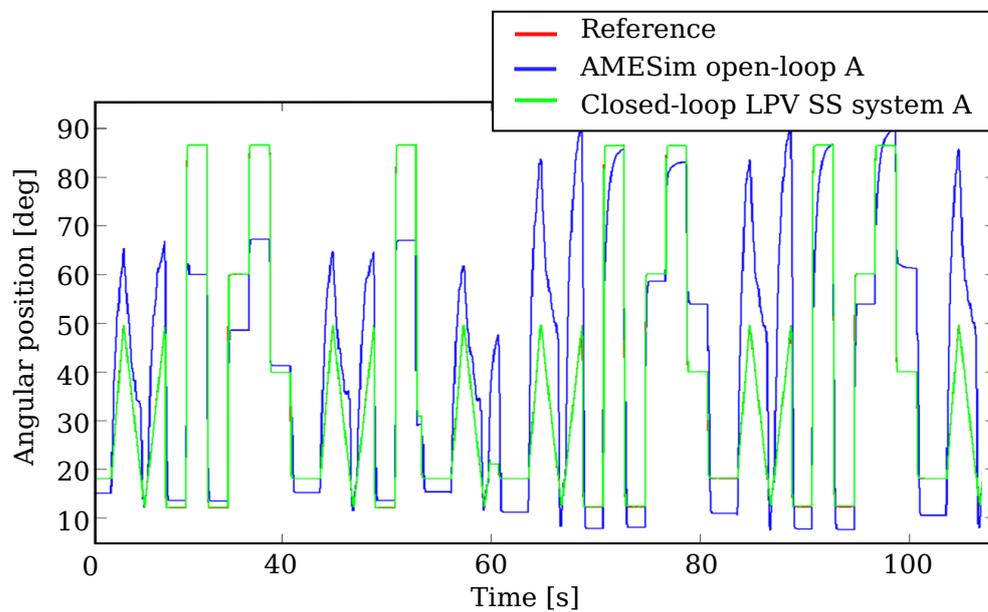
Figure 4.24: Throttle angular position of the closed-loop LPV SS system A

4.4.3 Control of AMESim® model A

The control designed on the identified linear SS model A is evaluated on AMESim® nonlinear model A with an objective of driving its output toward the reference. Knowing that the level of fit between the open-loop AMESim® model and the set-point was expressed in an NRMSE of 0.4443, Figure 4.25 shows that the objective is well fulfilled after applying the LQR control with reference and disturbance feedforwards, as the NRMSE rises up to 0.7653 in the closed-loop case.
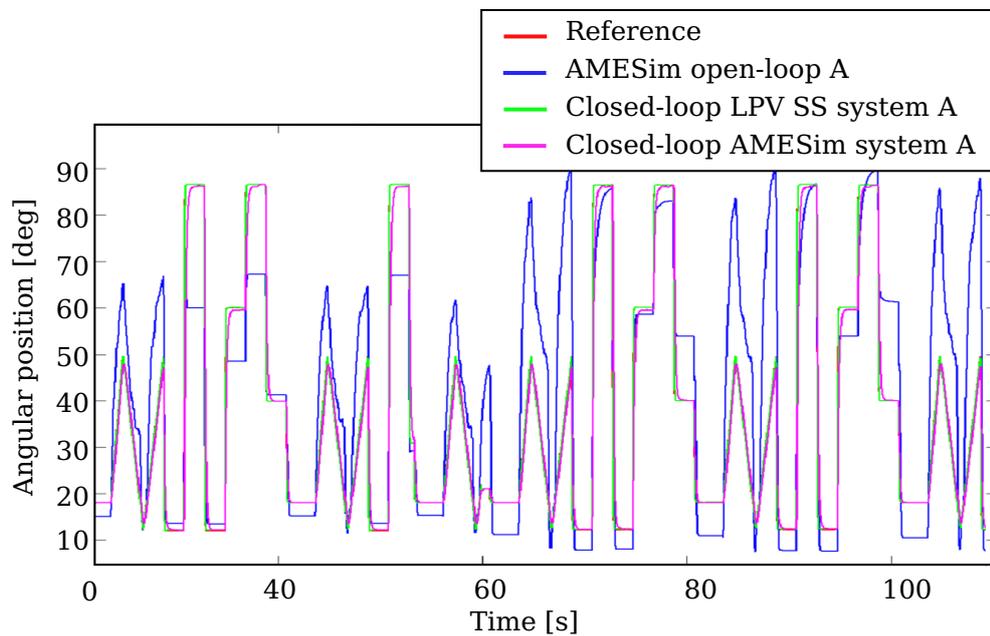
Figure 4.25: Throttle angular position of the closed-loop AMESim® nonlinear model A

4.5 Methodology of identification and control

In this section, the methodology of controlling a nonlinear AMESim® model of an electric throttle by means of linear optimal control is described in the flowchart shown in Figure 4.26. Note, however, that there are two prerequisites before identifying the linear model of the throttle and controlling it.

- Test-bench experiments: it is preferable to have 2 sets of test-

bench experiments, one representing a series of throttle open-
ing/closing motion in different conditions of temperature, ampli-
tude, frequency, etc. . . , and another representing a series of throt-
tle opening/closing motion around the limp-home position.

- A valid AMESim$^{®}$ model of the electric throttle.

Concerning the flowchart, a number of key points are noted.

- Study AMESim$^{®}$ model means studying the components of the
  model, the physical equations attributed to each of them, simu-
  lation properties, numerical manipulations, and whatever features
  necessary for constructing a reliable primary physical model of the
  electric throttle.

- When the SS model output does not follow its reference trajectory,
  the controller is reconsidered. Reconsidering the controller means
  re-tuning the weights Q and R of the performance function. It
  might also mean reconsidering the structure of the LQR controller
  such as the addition or the elimination of reference or disturbance
  feedforwards, or the addition of an error integral term to ensure a
  better reference tracking.

- If we are sure of the regression vector to use, the SS model might
  be avoided altogether along with its controller. In this case, we can
  go directly from step 1 of the flowchart (Study AMESim$^{®}$ model)
  to step 8 (Identify the ARX model), design a discrete LQR con-
  troller, and validate it on the AMESim$^{®}$ model. The advantage of
  this method is that it is a shortcut to the final step, but its applica-
  tion requires prior knowledge of the model.

- Although the SS model is a good representative of the AMESim$^{®}$
  model and its controller suffices to drive the output of the
  AMESim$^{®}$ model towards the reference output, it is still worthy
  to convert the SS into an ARX model, basically because when the
  SS and the ARX models are compared to the AMESim$^{®}$ model, the
  identified ARX model appears to be more accurate, as it is more ca-
  pable of capturing the model nonlinearities. The SS hence serves
  as a guideline for the identification of the ARX model.

- When this methodology is applied to an actuator that is different
  from the electric throttle, the number of models and controllers
  vary depending on the nonlinearities and discontinuities governing

the functioning of the actuator, i.e. the number of its functioning zones.

4.6 Conclusion and future perspectives

This chapter presents a methodology for linear ARX model identification and reference tracking control of an electric throttle. From the identification point of view, the mathematical LPV SS system that corresponds to the throttle's nonlinear model is presented, and is used to define the regression vector of an ARX model. The SS model and the identified ARX model have outputs which are very close to those of the nonlinear model, which points out the usefulness of using a mathematical model to form a black-box model that not only captures the model nonlinearities, but can be used independently from the mathematical model in future identifications. This methodology is also verified on another electric throttle nonlinear model, where a new ARX model is identified, and is shown to be compatible with a real throttle-test-bench. From the control point of view, an LQR controller is designed to control the linear SS, ARX, and nonlinear models. This method, which can be applied to other actuators similar to the electric throttle, splits the functioning of the electric throttle into 4 or 10 zones (depending on the chosen friction model), which results in an easily-implemented LPV model-controller structure.
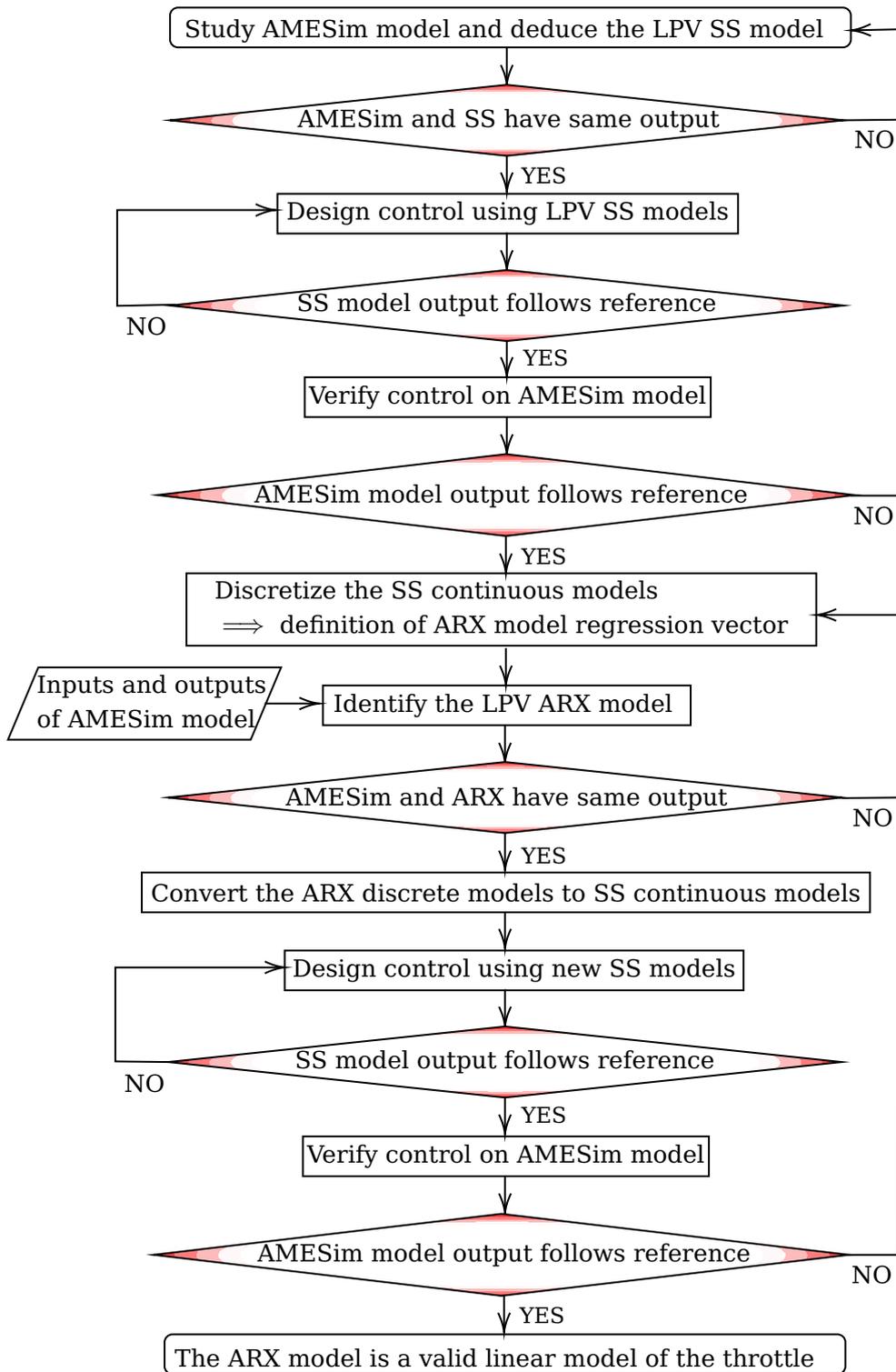
Figure 4.26: Flowchart describing the process of linearization and control of AMESim® model of electric throttle

Identification and Control of Transport Models in the Engine Air-path

Chapter Summary

The engine air-path, whose scheme is shown in Figure 5.1, is a combination of components, valves, and tubes, whose job is to collect air and guide it to the engine cylinders, and to clear away exhaust gases to the external environment. The air-path is critical to the function of the engine, because it provides oxygen to its combustion process and removes

away its harmful emissions. The cleaner, more continuous, and well-calibrated the air flowing through and away from the engine cylinders is, the better the mileage, power, and performance of the vehicle. The engine air-path consists of the air intake and the air exhaust systems.

The air-path shown in Figure 5.1 corresponds to the H5Ft400 engine. On the left of the figure is the air intake system composed of different components. The cold air intake marks the beginning of the air intake system as it pulls fresh air from outside. The introduced fresh air passes through a screen and an air filter which filters out dirt and other finer particles, and continues in the air intake tube via the compressor and the heat exchanger, which decreases the temperature of compressed air. After passing through the throttle body which regulates its mass flow rate Q_{air}, the air reaches the intake manifold from which it is directed to the combustion chambers via the intake valves. To regulate the amount of fuel to be injected into the cylinders, P_{col} in the intake manifold is measured by means of a MAP sensor. On the other hand, the different components composing the air exhaust system are shown on the right of the figure. The opening of the exhaust valves expels air from the combustion chambers to the exhaust manifold, connecting the cylinders to the exhaust piping. In the exhaust piping, part of the exhaust gases spins up the turbine, which in turn spins the compressor wheel in the air intake system. The harmful emissions in these gases are treated using the catalytic converter, before being ejected out of the exhaust system through the tailpipe. The other part of the exhaust emissions that doesn't pass through the turbine and the rest of the exhaust line, either bypasses the turbine via the wastegate or is sent back to the air intake path through the EGR system. The EGR system of the H5Ft400 engine is a low-pressure EGR. It is one of different types of EGR, the low-pressure EGR where the exhaust gases are recirculated from downstream of the turbine to upstream of the compressor, the high-pressure EGR where the exhaust gases are recirculated from upstream of the turbine to downstream of the compressor, and the dual-loop EGR where the low-pressure and high-pressure EGR coexist.

Transport phenomena exist everywhere in the engine air-path in their different forms: momentum transport, mass transport, and energy transport otherwise called heat transport. Two of these transport phenomena are addressed in this chapter. Heat transport is tackled from the heat exchanger point of view, while mass transport is tackled from the perspective of the EGR.

Pape (Pboost)
Tape

Pcol
Tcol

Throttle

Heat exchanger
(intercooler)

Qfuel

Pavt
Tavt

Qair
Tair
Pair

Papc
Tapc
Pavc
Tavc

Pexh
Texh

Waste-gate

Air filter

EGR

Compressor Turbine

Catalytic
converter

Heat exchanger

Q: mass flow rate	avc: upstream compressor
P: pressure	apc: downstream compressor
T: temperature	ape: downstream heat exchanger
	avt: upstream turbine
	col: intake manifold
	exh: exhaust manifold

Figure 5.1: H5Ft400 engine air-path scheme (with EGR)

Heat transport model in a heat exchanger

A heat exchanger is a device used to transfer heat among fluids without bringing them in direct contact. Different types of heat exchangers are used in different places of ICEs, to ensure adequate cooling and favorable operating conditions that allow ICEs to work more efficiently and prevent their failure. An intercooler is a particular type of heat exchanger, implemented in gasoline and diesel turbocharged ICEs, such as the H5Ft, to cool down the hot compressed air coming from the turbocharger. When intake air is compresed, it simultaneously experiences a temperature boost and density drop. The temperature boost becomes

disadvantageous after a certain threshold because it overheats the engine, therefore leading to premature ignitions and engine knock. The density drop is also unhealthy because it implies less oxygen content per unit volume, translated as a decrease in power production and a deterioration of combustion efficiency. By removing the waste energy in the compressed gas, the intercooler reduces the chances of knock and increases the volumetric efficiency of the ICE.

The heat exchangers in the air-path of the H5Ft engine are shown in Figure 5.1. Their heat transport model will exclusively include advective heat transport as radiative heat transport is considered small. The advective flows take place in the tubular structures making up the heat exchanger. They affect its temperature dynamics in a way that can be modeled as a first order hyperbolic PDE [BAB14]:

$$h\frac{dS_a}{dx}\Delta T = \rho S_a C_p \frac{\partial T}{\partial t} + Q C_p \frac{\partial T}{\partial x} \tag{5.1}$$

This equation illustrates the total advective heat transfer rate $\left(h\frac{dS_a}{dx}\Delta T\right)$ as the summation of the rate of heat transfer of the pipe in time $\left(\rho S_a C_p \frac{\partial T}{\partial t}\right)$ and the rate of heat transfer to the wall of the pipe $\left(Q C_p \frac{\partial T}{\partial x}\right)$. $\rho(\mathrm{kg\,m^{-3}})$, $C_p(\mathrm{J\,K^{-1}\,kg^{-1}})$, $Q(\mathrm{kg\,s^{-1}})$, $h(\mathrm{J\,s^{-1}\,m^{-2}\,K^{-1}})$, $S_a(\mathrm{m^2})$, and $\Delta T(\mathrm{K})$ represent the density of the fluid, specific heat, mass flow rate, advection exchange coefficient, heat transfer surface area, and the temperature gradient respectively.

Though different types of heat exchangers and intercoolers are used in engine air-paths, this work focuses in particular on tubular heat exchangers, which are very common devices in industry and very good examples of advective heat exchange phenomena. They permit heat exchange between two fluids separated by a wall and circulating in opposite or, less often, in similar directions.

Objective In the absence of an optimal operation of heat exchangers and intercoolers in the air-path and elsewhere in the engine, negative impacts on the engine performance can be recorded. For example, a dysfunctional intercooler in the intake air-path will result in volumetric efficiency loss and will increase knock in gasoline engines. A dysfunctional heat exchanger in the exhaust pipeline, on the other hand, will affect the turbocharger performance and degrade the quality of emissions expeled by the catalytic converter. Therefore, regulating the heat transport phenomena taking place in these heat exchangers is a must

in modern automobiles which should satisfy exigent criteria of performance, driving comfort, polluting emissions, etc... This regulation is done using simulation tools and temperature models that allow simulating and analyzing the thermal behavior of the gases circulating in their pipelines, and controlling it all at the same time. Therefore the general objective of this chapter can be divided in two parts: the first consists of setting the structure of the heat transport model in a heat exchanger and determining its mathematical equations, and the second consists of performing a parametric identification of the model to uncover the optimal values of its parameters.

Regarding the model structure, equation (5.1) shows that heat transport phenomena (in particular advective flows) occuring in a heat exchanger can be modeled as first-order hyperbolic PDEs. Infinite-dimensional, these hyperbolic PDEs accurately describe the behavior of these physical phenomena, yet recasting them as a time-delay system can be interesting as it preserves their infinite-dimensional property but also decreases their mathematical complexity, hence easing the design of a real-time control strategy.

Regarding the parametric identification of the model, it begins with an informative experiment design to generate interesting input-output data, in this case signifying the boundary temperature measurements in the tubular heat exchanger. The identification algorithm aims at identifying the optimal parameters of the heat exchanger's mathematical model, while minimizing, as much as possible, an error function expressing the difference between the output temperature of the experiment and that of the model. A robust parametric identification is important for several reasons. First of all, it results in a satisfactory error and a reliable model, which is valid on experiments other than that used for its identification, and which is able to reproduce the heat transport dynamics of a real heat exchanger. Also, the advection exchange coefficient is usually not well-known, difficult to identify, and changes as the heat exchanger ages in time. Hence, a robust technique for parametric identification of a heat-exchanger allows adapting its model to accomodate the changes it experiences over the course of time. Such a simple, easily-parameterized, yet accurate model is a cornerstone for analyzing a heat exchanger's thermal behavior and making it accessible to control design.

State of the art Multi-flow transport and advective flows phenomena, such as those encountered in a heat exchanger, although previously

modeled as a first-order transfer function featuring delay [TL60], are better described by two first-order hyperbolic PDEs which better implicate the physical balance laws (see [BAB14]). In the case of a tubular heat exchanger, for instance, two hyperbolic PDEs, which take into consideration the temperature evolution of both fluids (the one in the internal tube as well as the one in the external tube), are far more realistic than a single equation covering only the internal temperature evolution.

In order to deal with hyperbolic PDE equations modeling the temperature dynamics, different methods were presented in the literature. [BKO11] for example replaces the PDEs representing a heat exchanger with sets of nonlinear (ODE) with delayed inputs, each of these sets attributed to a section of the heat exchanger. Laplace transformation and numerical solutions of the PDE were also proposed in [ASGY04]; [Rom84]; [RW92]. However, these numerical solutions exhibit convergence and stability issues. [DS16] on the other hand uses an analytical approach and a physical lumping approach with time delay at certain positions of the heat exchanger in order to deal with the flow and wall dynamics. However, the drawback of this methodology is its relative complexity order due to the division of the exchanger in cells.

Proposed approach In this chapter, we represent the advective flows in a heat exchanger as coupled first-order hyperbolic PDEs. We propose to reformulate them as delay equations using the method of characteristics (see for example [Str04] or [KK14], [NRWS16] for an application in fluid networks, and [CK68] for an application in power lines). In order to ease this reformulation, we decouple the equations by using spatial-lumping. We then use a gradient-descent optimization method to estimate the parameters of the corresponding time-delay system (dynamics coefficients and time-delay), using boundary measurements of temperature in the heat exchanger. The interest of this technique is illustrated on experimental data obtained on a test-bench.

Mass transport model in the EGR

The low-pressure EGR system, denoted as LP-EGR, is a system which draws off some of the exhaust gases, cools them down by means of a heat exchanger, and redirects them back into the air intake system. In addition to that they absorb the combustion heat, the cooled down exhaust gases, having been burned already, are of low oxygen content, and therefore do not participate in the combustion process. This causes a reduction in the combustion speed and temperature, which is translated in a limited production of NOx, the product of a reaction between

nitrogen and oxygen catalyzed by high temperature conditions. This is essential in the case of diesel engines, where controlling and decreasing the production of nitrogen oxides is extremely important for meeting the upcoming emission standards. In gasoline engines such as the H5Ft400, the EGR is not implemented to treat nitrogen oxides in the first place, though it does. The technology basically aims at promoting fuel economy and improving knock resistance, which comes by as a natural effect of the reduction in peak in-cylinder temperatures. Though the benefits of EGR in diesel and gasoline engines do not coincide, its functioning in both cases is governed by the same physical principles and mathematical equations, detailed in the rest of this chapter.

During the transport of exhaust air all the way back to the intake manifold, it experiences two distinct compositions:

- Stage of pure exhaust gas: this stage exists prior to the EGR valve, and the burned-gas ratio (BGR) with respect to the overall gas mixture is equal to 100% because the exhaust gas is not yet mixed with fresh air.

- Stage of impure exhaust gas: after the EGR, the exhaust gas gets mixed with fresh air arriving from the outside environment, which decreases the BGR below 100%. Note that in gasoline engines, such as the H5Ft400, the BGR doesn't exceed 15% or 20%, in contrary to diesel engines, which operate with excess air and with a BGR reaching 50%.

Overlooking the dynamics of dilution in the volume downstream the EGR valve and assuming that fresh air and recirculated gases blend together instantaneously, the BGR is calculated as

$$BGR = \frac{Q_{EGR}}{Q_{EGR} + Q_{air}} \tag{5.2}$$

where Q_{EGR} is the mass flow rate of the recirculated exhaust air admitted through the EGR valve to the air intake line, and Q_{air} is the mass flow rate of the cold fresh air.

Understanding the dynamics of the BGR is fundamental for its control, which is why we hereby present the mathematical equations governing the physics behind the mixture of burned gases and fresh air, as it propagates in the intake line, shown in Figure 5.2, which is an extract of Figure 5.1. The burned gas fraction along the line evolves progressively

with time, and its dynamics vary according to the intake line components it is traversing. Two kinds of components can be distinguished: tube sections, where gas transport takes place, and control volumes, where gas mixing takes place [CB13].
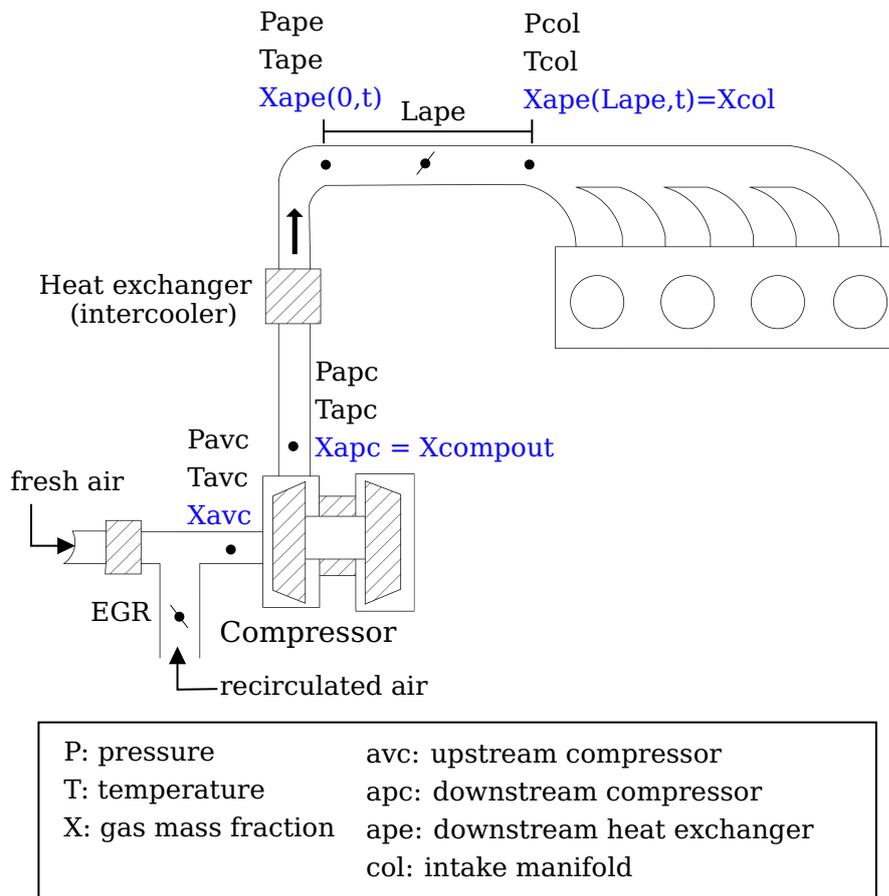
Figure 5.2: Air intake path scheme

Let $X(x, t)$ be the BGR at time t for a spatial coordinate $x \in [0, L]$. Its dynamics can be modeled using a first-order hyperbolic PDE expressed as [CB13]:

$$\partial_t X(x, t) + u(x, t)\partial_x X(x, t) = 0 \tag{5.3}$$

where $X(0, t) = X_{in}(t)$ and $X(x, 0)$ are the respective boundary and initial conditions, and $u(x, t)$ is a propagation speed.

Knowing the boundary condition $X_{in}(t)$ at the entrance of the tube section, and considering $u(x, t) = u(t)$ (with a plug-flow assumption, consid-

ering the fluid as incompressible), the burned gas fraction at the exit of the tube section can be calculated, using equation (5.3) and the method of characteristics, as [WN10]:

$$X(L,t) = X(t - \tau_f(t)) \tag{5.4}$$

where $\tau_f(t)$ is the time-varying time delay due to the transport of the gas inside the tube section, calculated as [CB13]:

$$\tau_f(t) \approx \frac{P(t)V_{tube}}{RT(t)Q(t)} \tag{5.5}$$

where V_{tube} is the volume of the tube section (m^3), $R = 287.058$ $(\mathrm{J\,kg^{-1}\,K^{-1}})$ is the gas constant, $T(t)$ is the temperature (K), $Q(t)$ is the mass flow rate $(\mathrm{kg\,s^{-1}})$, and $P(t)$ is the pressure (Pa) of the mixture at time t.

Equation (5.4) shows that the transport of the BGR inside the tube is, in fact, not immediate and subject to time delay. Taking into account this transport time is important in the case of LP-EGR, as the distance traveled by the recirculated gas is relatively long, and neglecting it could lead to severe performance degradation, especially during transient phases.

In a control volume, such as the intake manifold, by ignoring the air fraction dynamics with respect to space, the volume-average air fraction dynamics with respect to time can be formulated as a 0-D model, implying the law of conservation of mass and expressed as [CB13]:

$$\dot{X}_{cv} = \frac{RT_{cv}}{P_{cv}V_{cv}}[\sum_{in=1}^{l}(X_{in}Q_{in}) - X_{cv}\sum_{out=1}^{o}(Q_{out})] \tag{5.6}$$

where X_{cv}, T_{cv} (K), P_{cv} (Pa), and V_{cv} (m^3) are the control volume's respective BGR, temperature, pressure, and volume, X_{in} and Q_{in} $(\mathrm{kg\,s^{-1}})$ are the respective BGR and mass flow rate of the gas coming into the control volume, and Q_{out} $(\mathrm{kg\,s^{-1}})$ is the mass flow rate of the gas going out of the control volume.

Objective The amount of exhaust gas being recirculated has a huge impact on whether the EGR is functioning in favor of the engine overall performance or against it, which is why controlling the BGR is essential

to procure its benefits and avoid its drawbacks. For example, increasing the BGR raises the temperature of the intake mixture, leading to better vaporization of the fuel, better homogeneity of the mixture in the intake manifold, and therefore better combustion. This is highly recommended in cold weather, or directly after starting the engine to accelerate its warm-up. Nevertheless, the BGR can't be increased unlimitedly, because the higher the BGR, the more dilute the mixture in the intake manifold is. This mixture is hence poor in fuel and oxygen, and therefore low in power, which is inconvenient when high or full power is in demand [Mca33]. So the BGR ratio can be advantageous or disadvantageous depending on the operating conditions of the engine, the driver's commands and expectations, the external environment, and several other factors. Taking into account all these factors, experts define a cartography that maps different engine operating conditions to setpoints or reference values of X_{col}, the exhaust gas mass fraction in the intake manifold.

Therefore, the purpose of EGR control is to manipulate the opening angle of the EGR valve such that X_{col} attains its desired reference value, at the current operating conditions, while taking into account the transport delay of the intake gas and respecting the physical constraints of the valve, all at the same time.

State of the art Before tackling the modeling of mass fraction dynamics and how the literature dealt with the hyperbolic PDEs modeling them, note that all kinds of EGR were visited in former studies. However, LP-EGR has been less investigated than HP-EGR (where HP stands for high pressure), which was more popular in the early days of the EGR technology, and Dual-EGR, which recently started being popular. Note also that most of the studies conducted on EGR modeling and control considered diesel engines, and only a few considered gasoline engines. However, this fact is of little importance given that the EGR operates on the same principles in gasoline and diesel engines.

Modeling: Different kinds of models were used in the literature to model the engine air-path and the EGR system. Most of the studies relied on mathematical and MVEM nonlinear models, inspired from the physical principles, to model the air-path and EGR system dynamics [VNMKSWC98]; [JK98]; [UCKC00]; [AFGG03]; [Wan08]; [GMC09]; [YLSCHJC09]; [YW11]. Among these studies, many of them linearized the physical models [ADR95]; [VNMKSCW98]; [LK98]; [SKF00], sometimes as LPV models [JG03]; [ODR07]; [PS08] or linear time-varying

(LTV) models [WSKB17], with or without model reduction. Nevertheless, these models did not explicitly take the transport delay into account, unlike [BPLCP12] and [CB13], where the delay due to the mass transport phenomena resulted in SS formulations, describing the dynamics of X_{col} with time-varying input and state delays, respectively. The model in [BPLCP12] uses the input-delay model, describing $\dot{X}_{col}$ by means of a composition balance equation, for control purposes. The control input is the burned gas ratio downstream the EGR valve, delayed by the summation of the times needed to transport the gas from downstream the EGR valve to the compressor, from downstream the compressor to the intercooler, and from downstream the intercooler to the intake manifold. From the input, the EGR mass flow rate can be deduced, and the conversion of the latter to the effective valve opening is done using the Saint-Venant equation. The model in [CB13], on the other hand, is used for observation purposes. It divides the air intake path into control volumes, in which the air fraction dynamics are modeled as an ODE, and into tube sections, in which the air fraction dynamics are modeled as first-order hyperbolic PDEs, which can be reformulated as a time-delay system by means of the method of characteristics. The delay in this case corresponds to the transport of the gas from downstream the compressor to the intercooler, inside the intercooler (considered here as a tube section), and from downstream the intercooler to the intake manifold.

Control: Different control techniques have been used in the literature to control the EGR valve and the burned gas ratio in the intake manifold; take for example model predictive control (MPC) [ODR07], PID [XMJB10], LQG [PS08], SMC [UCKC00], feed-forward and PI feedback control [YLSCHJC09], control-Lyapunov function (CLF) [JK98], among others. Despite the efficacy of these techniques in controlling the air or burned gas fraction in the intake line, yet they suffer from a major drawback, which is their reliance on 0-D, instead of 1-D, air-path models. By doing so, they neglect the time delays induced by the transport of the gas mixture in the tubes of the air-path, which might be justifiable in the case of HP-EGR, but not in the case of LP-EGR and Dual-loop EGR, where ignoring the long travel distance of the recirculated gases and its corresponding transport delay might have serious impacts on the efficiency of the air-path control design, and consequently on the overall engine performance. To avoid this drawback, automakers and researchers started integrating 1-D air-path models in the engine control design. In [BPLCP12], a prediction-based trajectory tracking control is designed on a time-varying input delay model to determine the intake BGR for gasoline engines. In [CB13], where air fraction transport phe-

nomena in a LP-EGR-equipped diesel engine are modeled by means of a cascade of first-order LPV hyperbolic systems with dynamics associated with the boundary conditions, Lyapunov-based techniques and matrix inequalities are employed to find sufficient conditions for the exponential stabilization and observation of this class of systems. Overall, 1-D model-based air-path control design is a topic which is still under-investigated.

Proposed approach The transport of the BGR in the air intake path is modeled, as per [CB13], as a set of coupled ODEs and PDEs attributed respectively to a succession of control volumes and tube sections. In this work, we consider the average BGR dynamics in the tube downstream the compressor, considered as a control volume, and model it using a 0-D model. The part of the air intake path extending from upstream the heat exchanger until the intake manifold is considered as a tube section and modeled using a 1-D model. This results in a simplified ODE-PDE coupled system, whose one-dimensional part is discretized, for control purposes, by means of the method of lines [Ros71]. To track a reference profile of X_{col}, the EGR system is subject to two indirect optimal control approaches: discretize-then-optimize and optimize-then discretize [HPUU08]. In the former, the PDE is discretized first, leading to a delay-free traditional SS system, and the calculation of the necessary conditions of optimality follows afterwards. In the latter, on the other hand, the necessary conditions of optimality stem directly from the ODE-PDE coupled system, and then the discretization takes place to solve the resulting boundary-value problem. In both cases, however, because the BGR is limited between 0 and 100, the Augmented Lagrangian method is employed to ensure that the controlled output respects this constraint. The resulting system of equations representing the necessary optimality conditions is solved numerically using the projected gradient-descent method, which makes sure that the constraints on the control input X_{avc} (which is the BGR upstream the compressor) are respected. Although constraining the control problem is not a novel idea in the field of EGR control (take for example [CCP06]), but explicitly accounting for the input and output constraints while considering an infinite-dimensional air-path model that accounts for its transport delays, is, to the best of our knowledge, novel. The merits of using these control techniques are evaluated on the nonlinear AMESim® model from which the mathematical model was initially extracted.

Chapter structure The remainder of this chapter is organized as follows. Section 5.1 focuses on model identification of a heat exchanger.

It presents the hyperbolic PDEs modeling the temperature dynamics, details their reformulation as a time-delay system, formulates the time-delay system parameter estimation problem, presents the gradient-descent identification technique, and discusses the merits of this approach on experimental data obtained from a test-bench. Section 5.2 presents a delay linear system modeling of the EGR mass fraction. It also presents the formulation of the EGR contol problem and recapitulates the Augmented Lagrangian method which is at the core of the considered indirect optimal control methods, in their turn detailed in the same section, along with their simulation results. Finally, section 5.3 concludes the chapter and points out future work perspectives.

5.1 Identification of advective flows time-delay model in a heat exchanger

5.1.1 PDE model and problem formulation
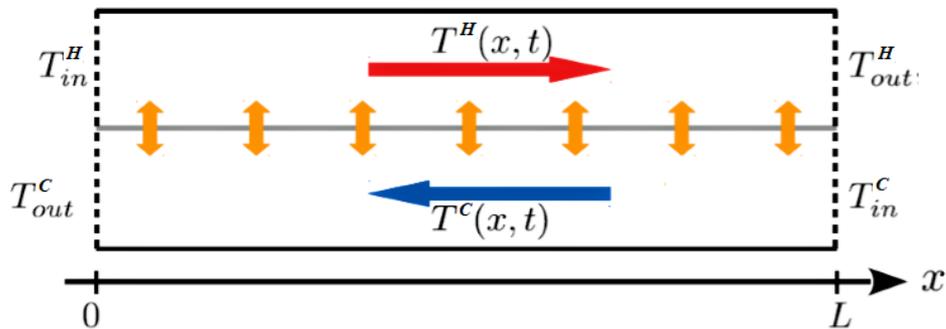
Figure 5.3: Schematic representation of a section of the air-path

To study the temperature dynamics of the advective flows, the heat exchanger's tubular structures are assumed to be split in two separate volumes separated by an internal wall. A schematic representation of this section is shown in Figure 5.3. It shows two volumes of length L separated by a wall. In the top volume, a hot fluid of distributed temperature T^H circulates in the rightward direction and, in the bottom volume, a cold fluid of distributed temperature T^C circulates in the leftward direction. Heat exchange takes place between the two fluids through the wall separating the two tubes. T_{in}^γ ($\gamma = (H, C)$) is the temperature of the fluid at the entrance of the tube and T_{out}^γ is its temperature at the exit.

The following assumptions from [ZWB17] are considered:

- the flow is one-dimensional (along an axis called x), with the hot fluid direction considered as positive.

- the heat diffusion through the tubes is negligible;

- advection between the two fluids is the only heat exchange taking place;

- the external walls of the tubular structure are adiabatic and no heat exchange between it and its surrounding environment takes place;

- the heat transfer coefficient is constant.

Under these assumptions, the temperature dynamics in the heat exchanger's tubular structure can be represented as the following system of coupled hyperbolic first order PDEs [BAB14][ZWB17]:

$$\frac{\partial T^H(x,t)}{\partial t} + c_1 \frac{\partial T^H(x,t)}{\partial x} = -d_1(T^H(x,t) - T^C(x,t)) \tag{5.7a}$$

$$\frac{\partial T^C(x,t)}{\partial t} - c_2 \frac{\partial T^C(x,t)}{\partial x} = d_2(T^H(x,t) - T^C(x,t)) \tag{5.7b}$$

where the parameters c_1, d_1, c_2, and d_2 are constant positive coefficients. x represents the spatial coordinate which spans an interval $[0, L]$, and t represents the time coordinate which spans an interval $[t_0, \infty)$. The corresponding boundary and initial conditions are

$T^H(0,t) = T_{in}^H(t)$ and $T^C(L,t) = T_{in}^C(t)$

$T^H(x,0) = T_0^H(x)$ and $T^C(x,0) = T_0^C(x)$

It can be noticed that the transport and heat exchange parameters c_1, d_1, c_2, and d_2 are highly uncertain (as d_1 and d_2 depend on the advection exchange coefficient which reveals troublesome to determine in practice). Therefore, the aim is to obtain an estimation technique of these parameters using only boundary measurements of the temperatures $T_{in}^H(t)$, $T_{in}^C(t)$, $T^H(L,t)$, and $T^C(0,t)$. With this aim in view, we propose to derive a time-delay approximation of this model.

5.1.2   From PDE to time-delay equations

Consider the coupled first-order hyperbolic PDE system (5.7) representing the heat exchange dynamics. In order to decouple these equations, the space-averaging technique presented in [NRWS16] is used to replace $T^C(x,t)$ in equation (5.7a) by the average

$$\frac{T_{in}^C(t)+T_{out}^C(t)}{2} = \frac{T^C(L,t)+T^C(0,t)}{2} \triangleq \frac{f(t)}{2}$$

and to replace $T^H(x,t)$ in equation (5.7b) by the average

$$\frac{T_{in}^H(t)+T_{out}^H(t)}{2} = \frac{T^H(0,t)+T^H(L,t)}{2} \triangleq \frac{g(t)}{2}$$

Let $T_x^H = \frac{\partial T^H(x,t)}{\partial x}$ and $T_t^H = \frac{\partial T^H(x,t)}{\partial t}$ (and similarly for T^C). The system dynamics can hence be expressed as

$$\widetilde{T}_t^H + c_1 \widetilde{T}_x^H = -d_1 \widetilde{T}^H(x,t) + \frac{d_1}{2}f(t) \tag{5.8a}$$

$$\widetilde{T}_t^C - c_2 \widetilde{T}_x^C = -d_2 \widetilde{T}^C(x,t) + \frac{d_2}{2}g(t) \tag{5.8b}$$

where $\widetilde{T}$ is another notation of temperature, used to differentiate between the PDEs (5.7) showing the temperature dynamics before the space-averaging assumption and the PDEs (5.8) showing the temperature dynamics after the space-averaging assumption. The corresponding boundary and initial conditions are

$\widetilde{T}^H(0,t) = \widetilde{T}_{in}^H(t)$ and $\widetilde{T}^C(L,t) = \widetilde{T}_{in}^C(t)$

$\widetilde{T}^H(x,0) = \widetilde{T}_0^H(x)$ and $\widetilde{T}^C(x,0) = \widetilde{T}_0^C(x)$

As the two equations (5.8a) and (5.8b) belong to the same class, only the first equation is discussed hereafter.

Following the method of characteristics presented in [Eva10], consider a point $(x,t) \in [0,L] \times [t_0, \infty)$ and define the function

$$z(\theta) = \widetilde{T}^H(c_1\theta + x, \theta + t) \tag{5.9}$$

$C = \{x(\theta), t(\theta), z(\theta)\}$ is the characteristic curve which passes through point (x, t) and satisfies the ODEs:

$$\frac{dx(\theta)}{d\theta} = c_1 \tag{5.10a}$$

$$\frac{dt(\theta)}{d\theta} = 1 \tag{5.10b}$$

$$\begin{aligned}
\frac{dz(\theta)}{d\theta} &= c_1 \widetilde{T}_x^H(c_1\theta + x, \theta + t) + \widetilde{T}_t^H(c_1\theta + x, \theta + t) \\
&= -d_1 \widetilde{T}^H(c_1\theta + x, \theta + t) + \frac{d_1}{2} f(\theta + t) \\
&= -d_1 z(\theta) + \frac{d_1}{2} f(\theta + t)
\end{aligned} \tag{5.10c}$$

To solve equation (5.10c), consider first the solution of the homogeneous equation:

$$\frac{dz(\theta)}{d\theta} + d_1 z(\theta) = 0 \tag{5.11}$$

which is $z(\theta) = z(0)e^{-d_1\theta}$, where $z(0) = z(\theta = 0) = \widetilde{T}^H(x, t)$. Note that $\widetilde{T}^H(0, \theta + t)$ corresponds to $c_1\theta + x = 0$, i.e. $\theta = \frac{-x}{c_1}$, which implies that $z(\theta) = z(\frac{-x}{c_1}) = \widetilde{T}^H(0, t - \frac{x}{c_1}) = \widetilde{T}_{in}^H(t - \frac{x}{c_1})$.

Therefore, the solution of the homogeneous equation is

$$\widetilde{T}^H(x, t) = \widetilde{T}_{in}^H\left(t - \frac{x}{c_1}\right) e^{-d_1 \frac{x}{c_1}} \tag{5.12}$$

Hence, for $t \geq \frac{x}{c_1}$, the solution of the nonhomogeneous equation (5.10) is:

$$\begin{aligned}
\widetilde{T}^H(x, t) = &\widetilde{T}_{in}^H\left(t - \frac{x}{c_1}\right) e^{-d_1 \frac{x}{c_1}} \\
&+ \frac{d_1}{2} e^{-d_1 \frac{x}{c_1}} \int_0^{\frac{x}{c_1}} f\left(\eta + t - \frac{x}{c_1}\right) e^{d_1\eta} d\eta
\end{aligned} \tag{5.13}$$

By evaluating (5.13) at $x = L$, the output temperature $\widetilde{T}^H(L, t)$ is thus obtained as

$$\widetilde{T}^H(L,t) = \widetilde{T}^H_{in}\left(t - \frac{L}{c_1}\right)e^{-d_1\frac{L}{c_1}}$$
$$+ \frac{d_1}{2}e^{-d_1\frac{L}{c_1}}\int_0^{\frac{L}{c_1}} f\left(\eta + t - \frac{L}{c_1}\right)e^{d_1\eta}d\eta \tag{5.14}$$

where $\widetilde{T}^H_{in}\left(t - \frac{L}{c_1}\right) = \widetilde{T}^H_{in}(0)$ when $t \leq \frac{L}{c_1}$.

Similarly, from (5.8b), one obtains

$$\widetilde{T}^C(x,t) = \widetilde{T}^C_{in}\left(t - \frac{L-x}{c_2}\right)e^{-d_2\frac{L-x}{c_2}}$$
$$+ \frac{d_2}{2}e^{-d_2\frac{L-x}{c_2}}\int_0^{\frac{L-x}{c_2}} g\left(\eta + t - \frac{L-x}{c_2}\right)e^{d_2\eta}d\eta \tag{5.15}$$

By evaluating (5.15) at the other boundary of the cold tube at $x = 0$, the output temperature $\widetilde{T}^C(0,t)$ is obtained as

$$\widetilde{T}^C(0,t) = \widetilde{T}^C_{in}\left(t - \frac{L}{c_2}\right)e^{-d_2\frac{L}{c_2}}$$
$$+ \frac{d_2}{2}e^{-d_2\frac{L}{c_2}}\int_0^{\frac{L}{c_2}} g\left(\eta + t - \frac{L}{c_2}\right)e^{d_2\eta}d\eta \tag{5.16}$$

Thus, (5.13)-(5.16) are time-delay equations depending only on past values of the boundary measurements.

5.1.3   Estimation of the parameters of time-delay equations

The time-delay equations (5.14) and (5.16) are now used to identify the vector $\kappa = \begin{bmatrix} c_1 & d_1 & c_2 & d_2 \end{bmatrix}$, and hence the time delays $\frac{L}{c_1}$ and $\frac{L}{c_2}$. The proposed optimization method is a gradient-descent algorithm, such as the one proposed in [WGDW06] for the optimal control of time-delay systems. The objective is to minimize over time the square of the error ε

$$\varepsilon = \begin{bmatrix} T^H{}_{measured} - T^H{}_{model}(L,\cdot) \\ T^C{}_{measured} - T^C{}_{model}(0,\cdot) \end{bmatrix} \tag{5.17}$$

which is defined as the difference between the test-bench output temperatures (measured temperatures) and the ones provided by the model (5.14) and (5.16).

The gradient-descent method is summarized below. Its objective is to find the optimal vector κ^* that minimizes the cost function defined as

$$J_0(\hat{\kappa}, t) = \frac{1}{2t_f} \int_0^{t_f} \varepsilon(\hat{\kappa}, t)^T Q_0 \varepsilon(\hat{\kappa}, t) dt \tag{5.18}$$

where t_f is a given time horizon, Q_0 is a weighting matrix, and $\hat{\kappa}$ is an estimate of the parameter vector κ, whose derivative is defined as

$$\dot{\hat{\kappa}} = -\alpha(\hat{\kappa}, t) \nabla J_0(\hat{\kappa}, t) \tag{5.19}$$

The gradient is computed as

$$\nabla J_0(\hat{\kappa}, t) = -\frac{1}{t_f} \int_0^{t_f} \varepsilon(\hat{\kappa}, t)^T Q_0 S(\hat{\kappa}, t) dt \tag{5.20}$$

where $S = \begin{bmatrix} \frac{\partial y}{\partial c_1} & \frac{\partial y}{\partial d_1} & \frac{\partial y}{\partial c_2} & \frac{\partial y}{\partial d_2} \end{bmatrix}$ is the sensitivity function, $y = [T^H(L,t) \; T^C(0,t)]^T$, and $\alpha(\hat{\kappa}, t)$ is the step obtained from Newton's method [MNT99] and is defined as

$$\alpha(\hat{\kappa}, t) = \left(\frac{1}{t_f} \int_0^{t_f} S(\hat{\kappa}, t)^T Q_0 S(\hat{\kappa}, t) dt + \upsilon I \right)^{-1} \tag{5.21}$$

with υ a positive constant acting as a tuning parameter.

By definition, one has

$$S_1 = \frac{\partial y}{\partial c_1} = \begin{bmatrix} \frac{\partial \widetilde{T}^H(L,t)}{\partial c_1} \\ 0 \end{bmatrix} \quad S_2 = \frac{\partial y}{\partial d_1} = \begin{bmatrix} \frac{\partial \widetilde{T}^H(L,t)}{\partial d_1} \\ 0 \end{bmatrix}$$
$$S_3 = \frac{\partial y}{\partial c_2} = \begin{bmatrix} 0 \\ \frac{\partial \widetilde{T}^C(0,t)}{\partial c_2} \end{bmatrix} \quad S_4 = \frac{\partial y}{\partial d_2} = \begin{bmatrix} 0 \\ \frac{\partial \widetilde{T}^C(0,t)}{\partial d_2} \end{bmatrix} \tag{5.22}$$

in which

$$\frac{\partial \widetilde{T}^H(L,t)}{\partial c_1} = -\frac{Ld_1}{2c_1{}^2} f(t) + \frac{L}{c_1{}^2} e^{-d_1 \frac{L}{c_1}} \times$$

$$\left(\frac{d\widetilde{T}_{in}^H \left(t - \frac{L}{c_1} \right)}{dt} + d_1 \widetilde{T}_{in}^H \left(t - \frac{L}{c_1} \right) \right.$$

$$+ \frac{d_1^2}{2} \int_0^{\frac{L}{c_1}} f \left(\eta + t - \frac{L}{c_1} \right) e^{d_1 \eta} d\eta$$

$$\left. + \frac{d_1}{2} \int_0^{\frac{L}{c_1}} \frac{df \left(\eta + t - \frac{L}{c_1} \right)}{dt} e^{d_1 \eta} d\eta \right) \tag{5.23}$$

$$\frac{\partial \widetilde{T}^H(L,t)}{\partial d_1} = \widetilde{T}_{in}^H \left(t - \frac{L}{c_1} \right) \left(\frac{-L}{c_1} \right) e^{-d_1 \frac{L}{c_1}}$$

$$+ \frac{c_1 - Ld_1}{2c_1} e^{-d_1 \frac{L}{c_1}} \int_0^{\frac{L}{c_1}} f \left(\eta + t - \frac{L}{c_1} \right) e^{d_1 \eta} d\eta \tag{5.24}$$

$$+ \frac{d_1}{2} e^{-d_1 \frac{L}{c_1}} \int_0^{\frac{L}{c_1}} f \left(\eta + t - \frac{L}{c_1} \right) (\eta) e^{d_1 \eta} d\eta$$

and S_3 and S_4 have the same form as S_1 and S_2 respectively, except that c_1 is replaced by c_2, d_1 is replaced by d_2, and $f(\cdot)$ is replaced by $g(\cdot)$.

In practice, the integrals in the sensitivities are estimated using the trapezoidal rule, and the time derivatives $\frac{df(.)}{dt}$ and $\frac{\widetilde{T}_{in}^H(.)}{dt}$ are approximated using a Backward Euler discretization. Subsection 5.1.5 shows that the effect of this approximation is not substantial and that the sensitivity resulting from the experimental measurements' noise is quite low.

5.1.4 Experimental setup

The proposed approach has been evaluated on a heat exchanger test-bench available in GIPSA-lab, Grenoble, France. As shown in Figure 5.4, it consists of a hot tank (equipped with an immersed heater which heats the water up to a target temperature), a cold tank, and a heat exchanger which is a tubular structure made of two concentric tubes, an external tube through which the hot fluid circulates and an internal tube through which the cold fluid circulates. The surfaces of the heat exchanger tubes are designed to maximize the turbulence of the flows. In addition to that, it also consists of hot fluid and cold fluid pumps, manual and automatic
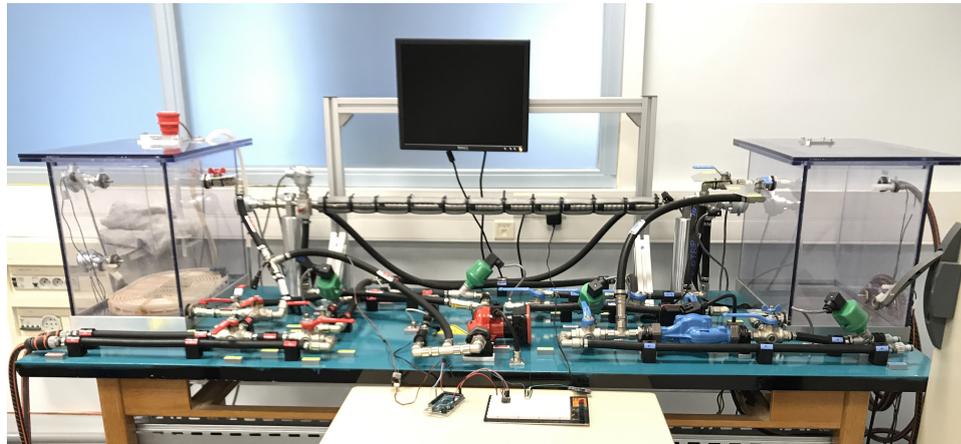
Figure 5.4: Heat exchanger test-bench in GIPSA-lab

valves to direct the flows, and various sensors of temperature, pressure, and volumetric flow rate. The sensors provide the measurements of the temperature and the volumetric flow rate of each of the fluids at the entrances and exits of the exchanger. A fresh water source and the immersed heater allow controlling the temperatures at the entrances of the exchanger.

In addition, the test-bench is equipped with a target PC used for data acquisition and a host PC through which the user can communicate with the target PC using the XPC Target of Simulink.

The exchanger is dimensioned according to the values provided in Table 5.1. Hence, for a hot fluid temperature of $70°C$ and a flow rate of $0.2\,\mathrm{m^3\,h^{-1}}$, a temperature difference of $10°C$ between the input and the output is expected, i.e. a theoretical heat capacity of $2.3\,\mathrm{kW}$.

	Flow Rate	T_{in}	T_{out}	Reynold's nb	Flow Type
Cold	$0.2\,\mathrm{m^3\,h^{-1}}$	$20°C$	$30°C$	8000	Turbulent
Hot	$0.2\,\mathrm{m^3\,h^{-1}}$	$70°C$	$60°C$	5500	Transitional

Table 5.1: Nominal thermal specifications of heat exchanger test-bench

5.1.5 Experimental results

Two sets of experiments are used, one for identification and another for validation. The identification experiment is held at around $T_{in}^H = 40°C$,

and volume flow rates of $\dot{q} = 260\,\text{L}\,\text{hr}^{-1}$. The validation experiment is a concatenation of two experiments held at a temperature $T_{in}^H = 60°C$, and mass flow rates of $\dot{q} = 260\,\text{L}\,\text{hr}^{-1}$. The difference between the first and the second validation experiment is that, in the first experiment, a disturbance is applied on the entrance of the hot circuit (T_{in}^H) and no disturbance is applied on the cold circuit, whereas the second experiment consists in the reverse, that is a disturbance is applied on the entrance of the cold circuit (T_{in}^C) and no disturbance is applied on the hot circuit.

The optimization problem is set with the following parameters:

- Initial set of parameters
  $$\kappa_0 = \begin{bmatrix} 0.01 & 0.001 & 0.01 & 0.001 \end{bmatrix}$$

- $Q_0 = \begin{bmatrix} 4 & 0 \\ 0 & 10 \end{bmatrix}$

  $Q_{0_{11}}$ and $Q_{0_{22}}$ are not taken equal because, in the experimental setup, the hot tube is the external tube. Therefore, the hot temperature can be subject to unmodelled exchange with the external environment despite the isolation. Consequently, a higher weight is attributed to the cold temperature error.

- $\upsilon = 10^{-9}$

- $\|\nabla J_0\| \leq 10^{-6}$ is a stop condition of the optimization process.

The primary concern is to show that the time-delay model is able to reproduce the same output as the PDE system. For this sake, a first estimation of the transport and exchange parameters is done using the discretized PDE system and the gradient-descent estimation method derived previously. The estimated parameters are shown in Table 5.2.

c_1 $(\text{m}\,\text{s}^{-1})$	d_1 $(\text{W}\,\text{J}^{-1}\text{m}^{-1})$	c_2 $(\text{m}\,\text{s}^{-1})$	d_2 $(\text{W}\,\text{J}^{-1}\text{m}^{-1})$
0.12144	0.023693	0.17998	0.029412

Table 5.2: Parameters estimated with the PDE model

Using these parameters, the PDE system and the time-delay system are simulated with the measured input temperatures and compared with the outflows measurements provided by the test-bench experiments. Figure 5.5 corresponds to the identification experiment, with Figure 5.6 showing its corresponding absolute error. Figures 5.7 and 5.8 show the same results on the validation experiments. The mean-squared averaged error

$$\frac{1}{N}\sum_{i=1}^{N}(T_{measured}^{\gamma} - T_{model}^{\gamma})^2 \qquad (5.25)$$

corresponding to the experiments simulated with the parameters estimated from the PDE model is summarized in Table 5.3.
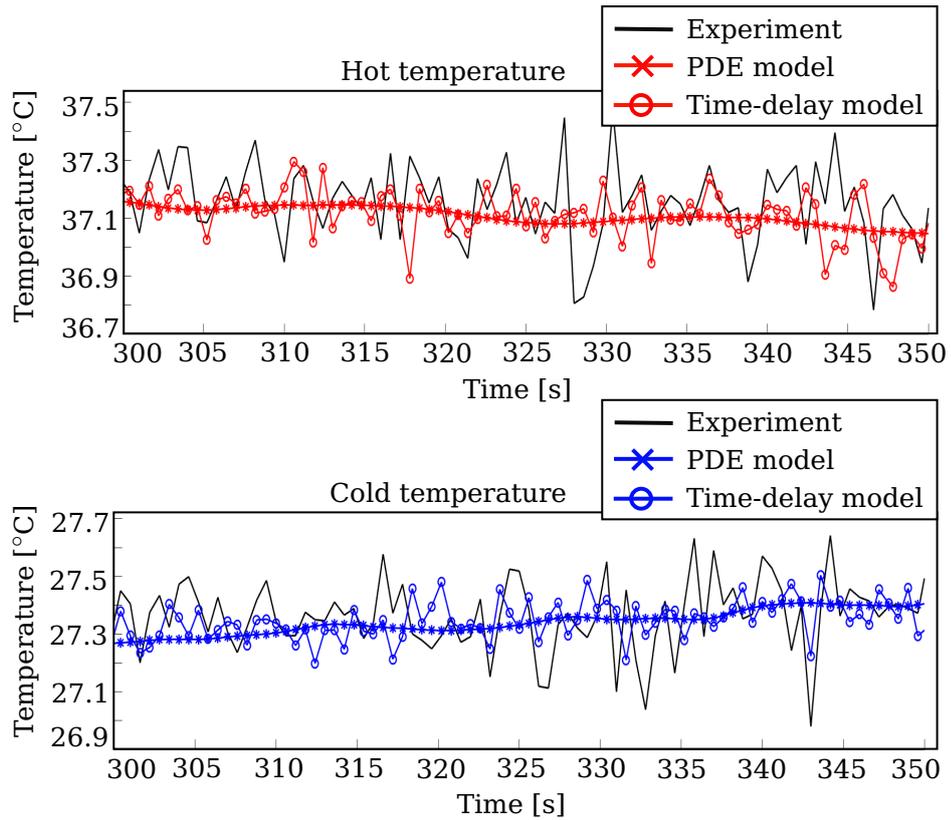
Figure 5.5: Comparison of PDE and time-delay models with respect to the experimental data - identification experiment

The simulation results show the credibility of the PDE and the time-delay models as the error separating them from the experimental data is low, especially in the identification experiment. In addition to that, the error separating each of the PDE and the time-delay models from the experiment is very close indicating a very slight difference between them, which illustrates the adequacy of the averaging assumption.

The following step is to identify the system using the time-delay model, i.e. equations (5.14) and (5.16). The results are shown in Table 5.4.
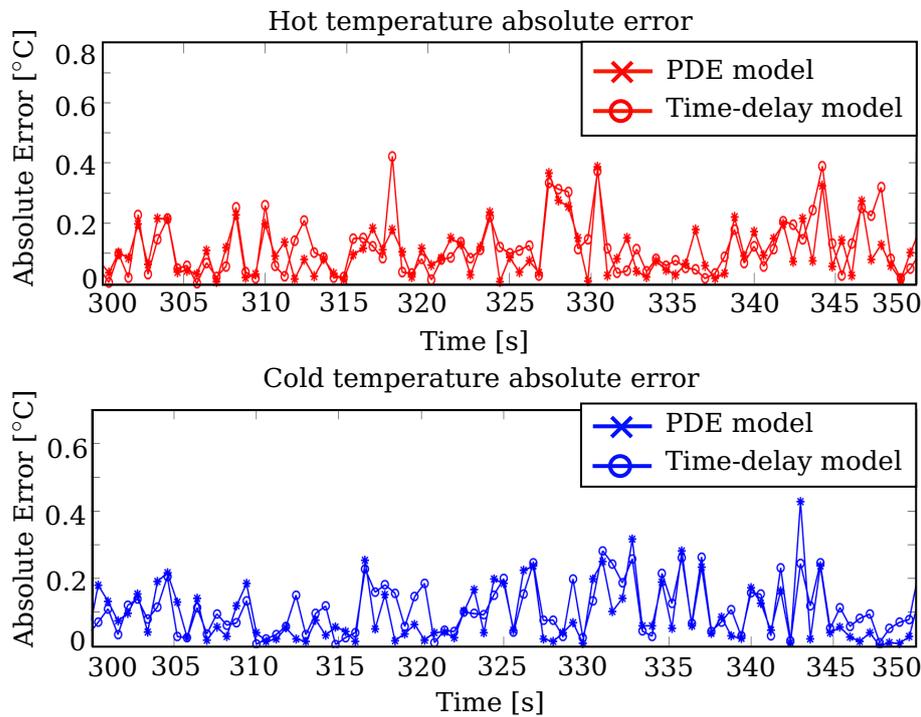
Figure 5.6: Absolute errors of PDE and time-delay models with respect to the experimental data - identification experiment

Although the estimated parameters are quite different from those estimated in the case of the PDE system, the simulation results tend to imply that this new set of parameters is also valid. Indeed simulating the PDE model and the time-delay model using these parameters, the results shown in Figure 5.9 for the identification experiment and those shown in Figure 5.11 for the validation experiment are obtained. The mean-squared averaged error corresponding to the experiments simulated with the parameters estimated from the time-delay model is summarized in Table 5.5.

These results show that the two models have a high level of similarity and a good level of fit with the experimental data. This can also be seen in Table 5.5 and Figures 5.10 and 5.12 which show low absolute errors in the case of the identification experiment as well as the validation experiment.

In addition to that, comparing Tables 5.3 and 5.5 shows that the mean-squared errors obtained on the identification experiment with the parameters estimated from the PDE system are almost equal to those
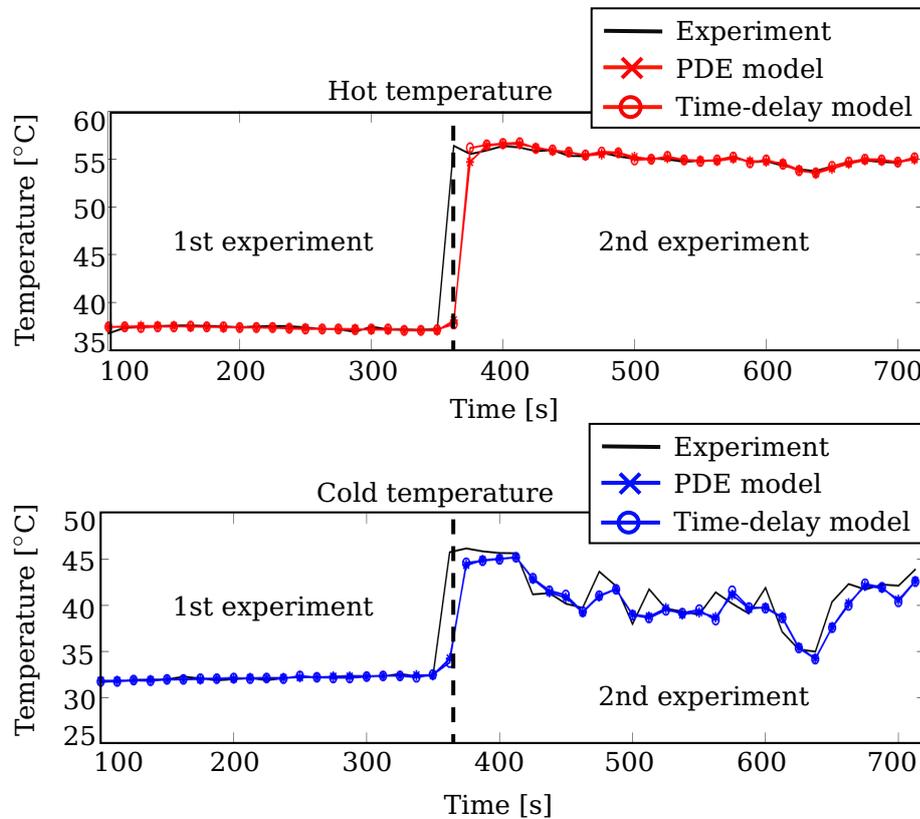
Figure 5.7: Comparison of PDE and time-delay models with respect to the experimental data - validation experiment

recorded on the same experiment but with the parameters estimated from the time-delay system. Although this implies that the estimation algorithm works similarly on both models, the mean-squared errors corresponding to the validation experiment are much better in the case when the estimation is done using the time-delay model. In addition to that, a time-save of 98% can be noted when launching the estimation algorithm using the time-delay system. This implies that the time-delay system is more adequate and more computationally efficient than the PDE model to be used for estimation using the gradient-descent algorithm.

Finally, Tables 5.2 and 5.4 indicate that the parameter vector estimated using the PDE model is far from being equal to that estimated using the time-delay model. However, it can be noted that the ratio $\frac{c_1}{d_1} = 5.12$ from Table 5.2 is almost equal to the ratio $\frac{c_1}{d_1} = 5.06$ from Table 5.4. The same remark can also be attributed to the ratio $\frac{c_2}{d_2}$. This might be explained by the simulation results depicted in Figures 5.5 and 5.9 showing that
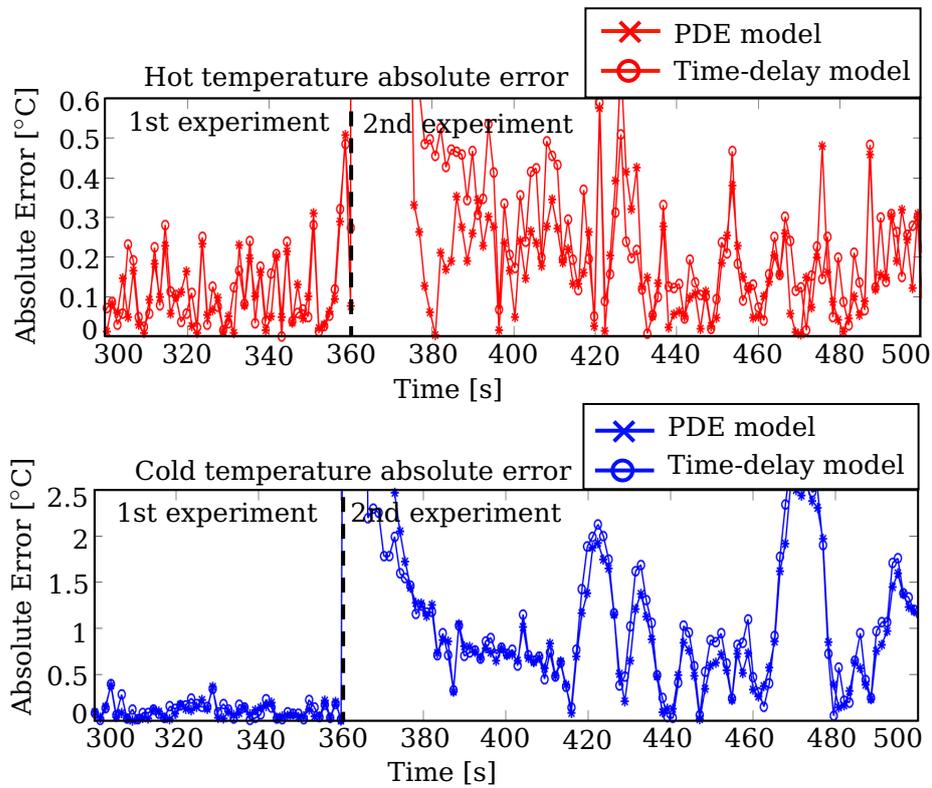
Figure 5.8: Absolute errors of PDE and time-delay models with respect to the experimental data - validation experiment

the PDE system and the time-delay model represent more accurately the steady-state behavior of the model rather than its transient state.

Experiment set	Identification		Validation	
Fluid	Hot	Cold	Hot	Cold
PDE	0.0234	0.019	3.3333	2.3533
Time-delay	0.0275	0.0237	4.3190	2.5425

Table 5.3: Mean-squared averaged error

c_1 $(\mathrm{m\,s^{-1}})$	d_1 $(\mathrm{W\,J^{-1}m^{-1}})$	c_2 $(\mathrm{m\,s^{-1}})$	d_2 $(\mathrm{W\,J^{-1}m^{-1}})$
15.9021	3.14237	16.1599	2.59372

Table 5.4: Parameters estimated with the time-delay model
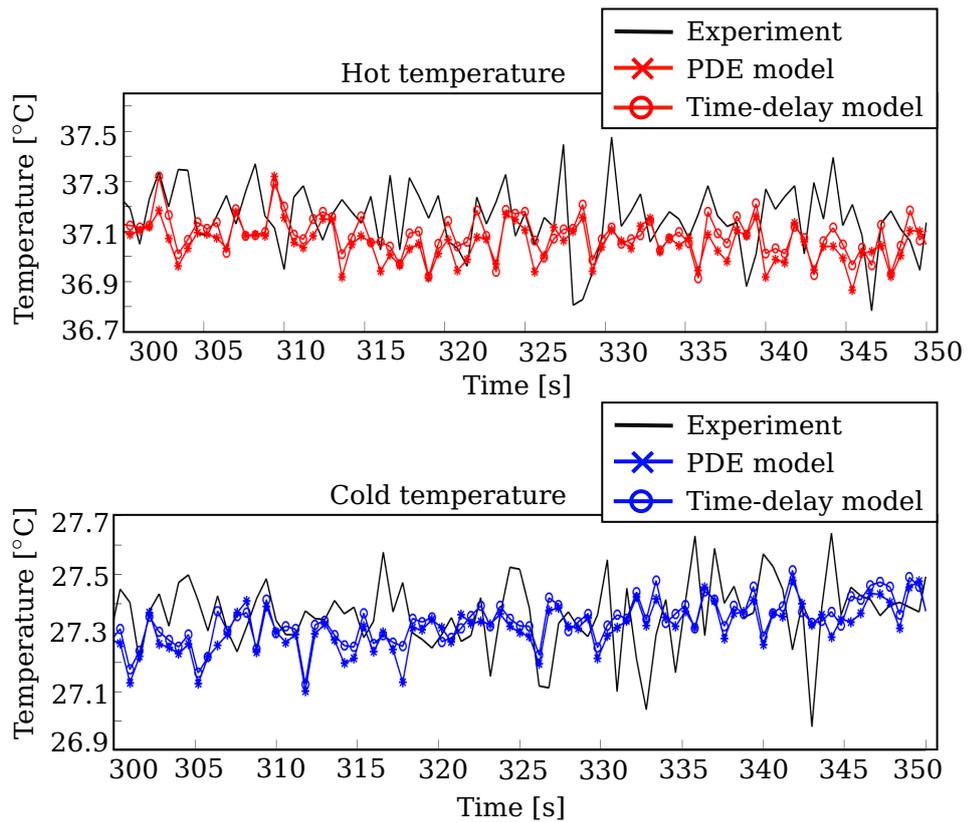
Figure 5.9: Comparison of PDE and time-delay models with respect to the experimental data - identification experiment
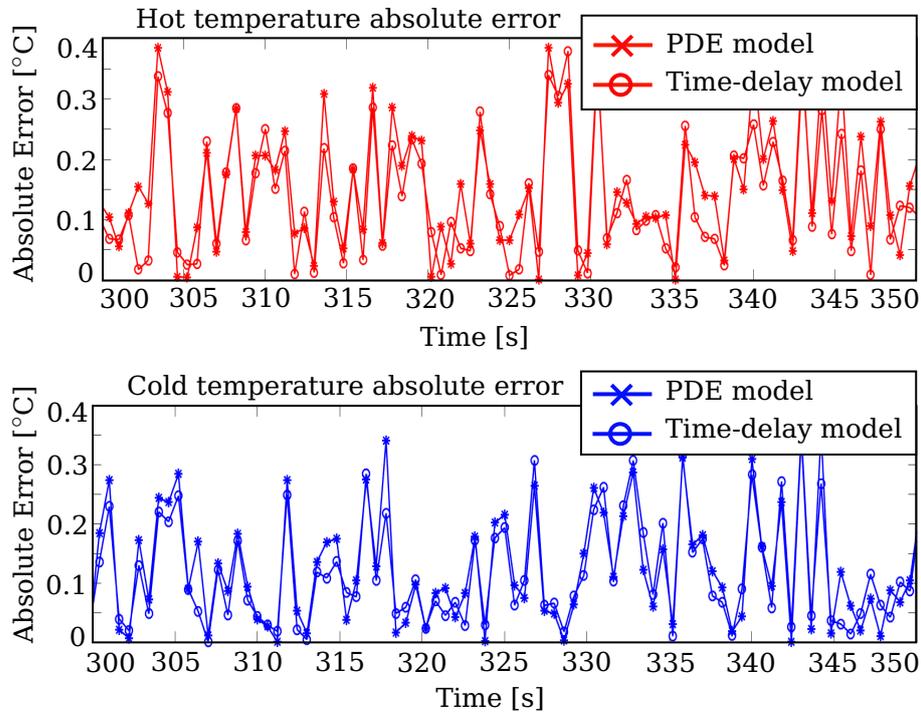
Figure 5.10: Absolute errors of PDE and time-delay models with respect to the experimental data - identification experiment

Experiment set	Identification		Validation	
Fluid	Hot	Cold	Hot	Cold
PDE	0.0357	0.0259	0.0655	0.3947
Time-delay	0.0335	0.0252	0.0862	0.3921

Table 5.5: Mean-squared averaged error
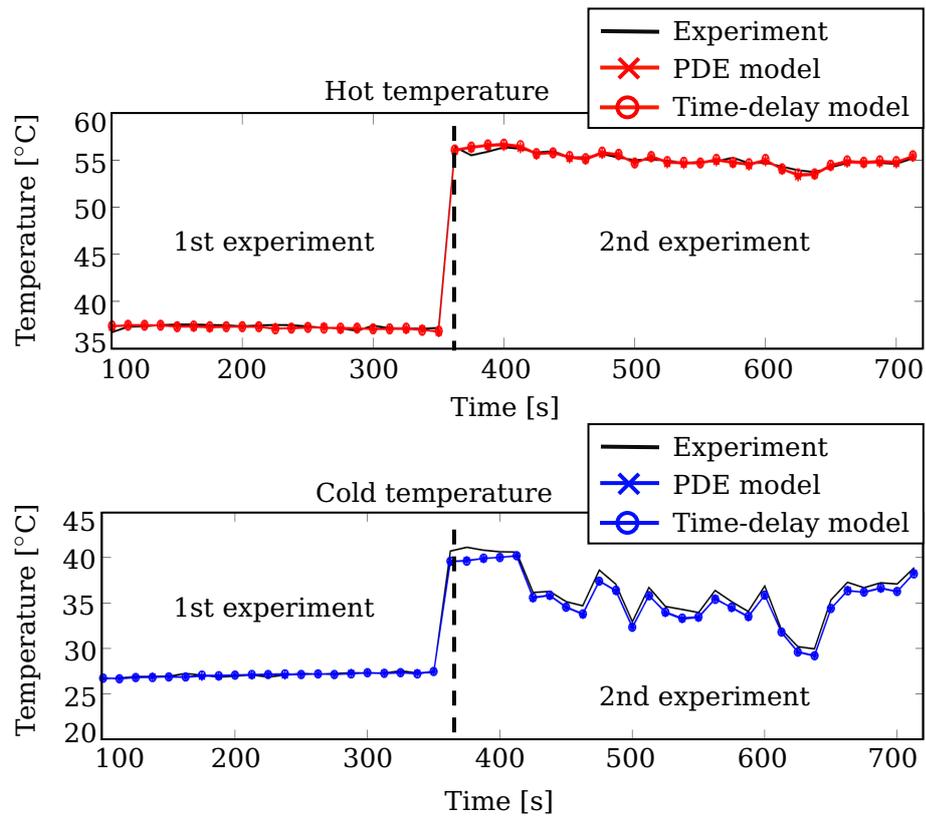
Figure 5.11: Comparison of PDE and time-delay models with respect to the experimental data - validation experiment
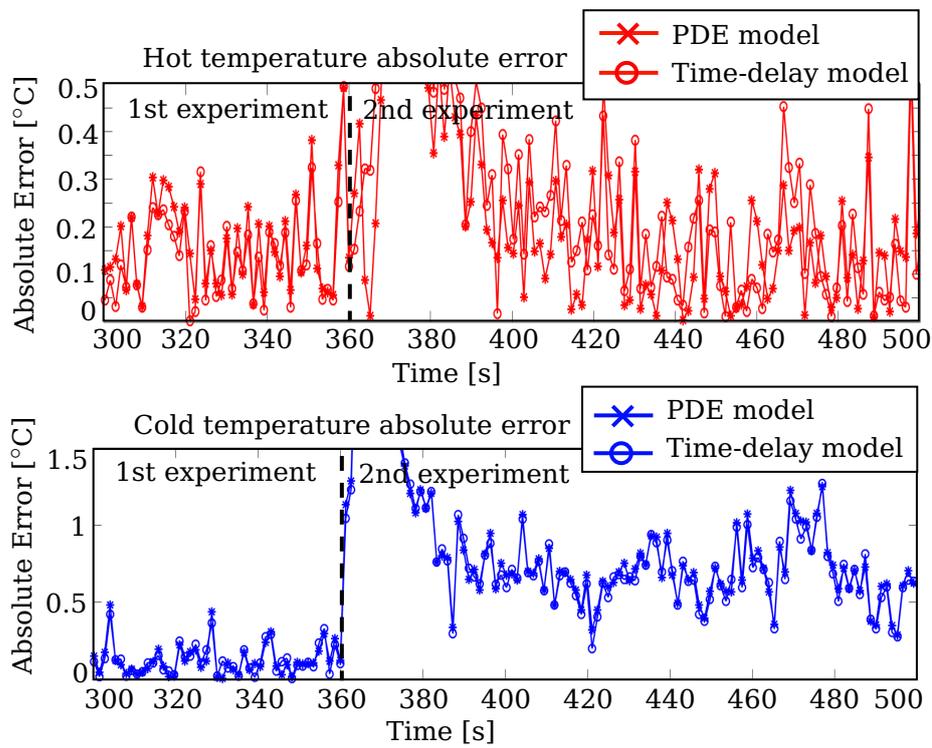
Figure 5.12: Absolute errors of PDE and time-delay models with respect to the experimental data - validation experiment

5.2 Optimal control of mass-transport time-delay model in an EGR

In this section, we turn our attention to another transport problem, which is optimal control of the LP-EGR time-delay system, whose linear form is hereby presented.

5.2.1 EGR linear state-space model

The schematic presentation of the section of the air intake path extending from the LP-EGR valve to the intake manifold is shown in Figure 5.2. It shows the fresh air coming from the external environment mixing with the recirculated gases, admitted to the intake path through the EGR valve, in the volume upstream the compressor (AVC). The mixture of gases then traverses the compressor, the volume downstream the compressor (APC), the heat exchanger (HE), and the tube section following it (APE). Note how AVC and APC are considered as control volumes, where mere gas mixing occurs, whereas HE and APE are considered as tube sections. $X(x,t)_n$ and $X(t)_m$ are the respective burned gas fractions in the tube sections $n = \{HE, APE\}$ and control volumes $m = \{COL, APC, AVC\}$. x represents the spatial coordinate which spans an interval $[0, L_n]$ (L_n being the length of tube section n), and t represents the time coordinate which spans an interval $[t_0, \infty)$.

To calculate the burned gas fractions in the intake path, we consider the following assumptions, inspired from [CB13]:

- The control volumes in the EGR path are large compared to their length, which means that the mass transport phenomenon is negligible compared to the mixing phenomenon, which justifies modeling the gas mixing in these volumes as ODEs.

- the gas mixing in the heat exchanger of the intake path is considered negligible, and the heat exchanger is entirely considered as a tube section;

- the mixing dynamics in the intake manifold are negligible;

From the first two assumptions, and from equation (5.6) describing the BGR dynamics in a control volume, the BGR dynamics in the control

volume downstream the compressor can be written as:

$$\dot{X}_{apc} = \alpha_{apc}[-Q_{engine}X_{apc}(t) + Q_{engine}X_{avc}(t)] \tag{5.26}$$

where $\alpha_{apc} = \frac{RT_{apc}}{P_{apc}V_{apc}}$, X_{apc} (%), T_{apc} (K), P_{apc} (Pa), and V_{apc} (m^3) are the respective BGR, temperature, pressure, and volume in the control volume downstream the compressor, and Q_{engine} (kg s^{-1}) is the mass flow rate of the gas flowing in the intake air-path, equal to $(Q_{air} + Q_{EGR})$ (assuming that the speed dynamics of the particles are fast in comparison to the dynamics of the BGR).

According to [BPLCP12], the control input to the model describing the dynamics of X_{col} is X_{avc}, the burned gas ratio downstream the EGR valve, delayed by the transport time in the intake path. It is chosen as control input because, once calculated, it can be translated to EGR mass flow rate by virtue of equation (5.2), and then to effective valve opening by virtue of the Saint-Venant equation.

The third assumption implies that the dynamics of X_{col} are not expressed as an ODE, as is the case for control volumes. Instead, $X_{col}(t)$ is supposed to be equal to $X_{ape}(L_{ape}, t)$, calculated at the very end of the tube section downstream the heat exchanger. This assumption, along with equation (5.4) accounting for the transport of the BGR in the heat exchanger and in the tube section following it, can be formulated as:

$$X_{col}(t) = X_{ape}(L_{ape}, t) = X_{apc}(t - \tau_{total}(t)) \tag{5.27}$$

where X_{col} (%) and X_{ape} (%) are the respective BGRs in the intake manifold and in the tube downstream the heat exchanger, $\tau_{total}(t) = \tau_{ape}(t) + \tau_{HE}(t)$, $\tau_{ape}(t) = \frac{P_{ape}(t)V_{ape}}{RT_{ape}(t)Q_{engine}(t)}$ and $\tau_{HE}(t) = \frac{P_{HE}(t)V_{HE}}{RT_{HE}(t)Q_{engine}(t)}$ are the respective delays, measured in seconds, due to the transport of the gas in the tube downstream the heat exchanger and in the heat exchanger itself, L_{ape} (m), T_{ape} (K), P_{ape} (Pa), and V_{ape} (m^3) are the respective length, temperature, pressure, and volume of the tube downstream the heat exchanger, and T_{HE} (K), P_{HE} (Pa), and V_{HE} (m^3) are the respective temperature, pressure, and volume of the heat exchanger.

Therefore, the resulting SS model is an output delay model expressed as:

$$\begin{aligned} \dot{\mathcal{X}}(t) &= A(t)\mathcal{X}(t) + B(t)U(t) \\ Y &= \mathcal{X}(t - \tau_{total}(t)) \end{aligned} \tag{5.28}$$

where $A(t) = -\alpha_{apc}Q_{engine}$, $B(t) = \alpha_{apc}Q_{engine}$, $\mathcal{X}(t) = X_{apc}(t)$, $U(t) = X_{avc}(t)$, and $Y(t) = X_{col}(t) = X_{apc}(t - \tau_{total}(t))$.

Note that, alternatively, in view of (5.3), these dynamics can be formulated as the following ODE-PDE system:

$$
\begin{aligned}
\dot{\mathcal{X}}(t) &= A(t)\mathcal{X}(t) + B(t)U(t) \\
\partial_t X(x,t) + u(t)\partial_x X(x,t) &= 0 \\
X(0,t) &= \mathcal{X}(t) \\
Y(t) &= X(1,t)
\end{aligned}
\tag{5.29}
$$

where $u(t) = \frac{1}{\tau_{total}(t)}$ is the BGR propagation speed, assuming a normalized length $L_{ape} = 1$.

Figure 5.13 shows X_{col} generated by the AMESim® model and by model (5.28). The small difference between the two plots is an indicator of the adequacy of the assumptions.
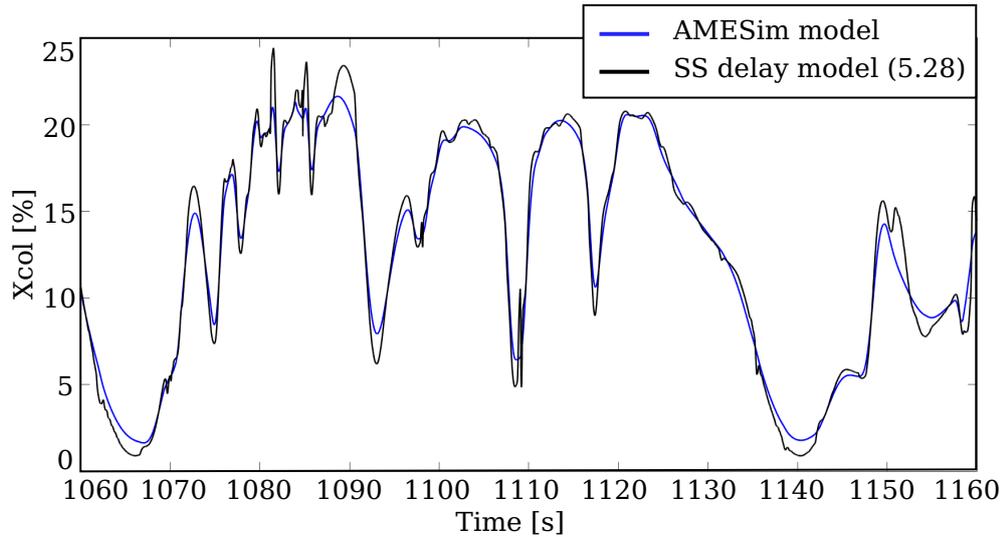
Figure 5.13: Xcol from AMESim® and the linear SS delay models

5.2.2   Control problem formulation

The objective of controlling the EGR is to drive the exhaust gas mass fraction in the intake manifold $X_{col}(t)$ to its setpoint $r(t)$, using minimum control effort U. Therefore, an objective function to minimize, with respect to U, can be expressed as:

$$J_1(t) = \frac{1}{2}\|X_{col}(T) - r(T)\|_{L_1}^2 + \frac{1}{2}\int_0^T \|X_{col}(t) - r(t)\|_{Q_1}^2 + \|U(t)\|_{R_1}^2 dt \quad (5.30)$$

where $\|\cdot\|_\beta$ (with $\beta^T = \beta$ a positive matrix) denotes the weighted Euclidean norm, $L_1 \in \mathbb{R}_{\geq 0}$, $Q_1 \in \mathbb{R}_{\geq 0}$, and $R_1 \in \mathbb{R}_{>0}$ are the matrix weights used to manage the trade-off between the different minimization objectives implied in the cost function J_1, and T is the final time.

This optimization problem is subject to the infinite-dimensional dynamics described in equation (5.28) or (5.29), and to constraints on the input $U(t) = X_{avc}(t)$ and on the output/state $X_{col}(t)$, because these quantities represent gas mass fractions whose values vary between 0% and 100%. This leads to the inequality constraints

$$0 \leq X_{col} \leq 100$$
$$0 \leq U(t) \leq 100$$

Based on the above, we formulated a constrained optimization problem of a time-delay system. Before addressing the application of the indirect optimal control methods, the next subsections give a general overview of their principles, with a focus on the Augmented Lagrangian method, and the UZAWA method, which are used to manage the input and state constraints.

5.2.3 Indirect optimal control methods

Optimal control is a control strategy that aims at finding an admissible control law, while optimizing a performance criterion, which is a function of state and control variables. Because solving the optimal control problem analytically is unlikely, two numerical methods were developed: the direct method and the indirect method. Firstly developed for finite-dimensional systems, these methods were later extended to infinite-dimensional systems. This section gives a briefing about the indirect methods, states the necessary conditions of optimality in the finite as well as in the infinite-dimensional case, and points out the main references in the field. The particularities concerning the constraints are left to the following section 5.2.4 which discusses the techniques used in this work to manage the state and output constraints.

The direct methods discretize the original control problem to recast it as a finite-dimensional problem, which often results in a computationally expensive-to-solve nonlinear constrained optimization problem, to be solved using nonlinear programming algorithms. [Bet10] can be referred to, for a detailed explanation of the direct methods and the algorithms they use to solve the control problem.

On the other hand, the indirect methods, which are the methods adopted in the sequel for the air-path control problem, solve the optimal control problem, in continuous time, using Pontryagin Minimum Principle [PBGM62]. Although Pontryagin Minimum Principle provides only the necessary (but not sufficient) conditions for optimality, using the calculus of variations, it often reduces the computational load compared to the direct methods. For a detailed explanation of the indirect methods and the necessary conditions for optimality in the cases of unconstrained admissible controls and states, constrained controls, and constrained states, one can refer to [Kir04], and for an extension to the case of constrained controls and states, one can refer to [SW77].

Both methods, though fundamentally developed for finite-dimensional state dynamics, yet they found their way through infinite-dimensional state dynamics, thanks to the works of [Lio71], [AT81], [Trö10], [LY95], and others. As already mentioned in the introduction, the air-path control problem is approached from two different perspectives: discretize-then-optimize and optimize-then-discretize. In the discretize-then-optimize method, the discretization of the system transforms the infinite-dimensional PDE into a finite-dimensional system of ODEs, before the calculation of the first-order optimality conditions takes place. This leads to the definition of a standard optimal control problem, unlike the optimize-then-discretize method which preserves the infinite-dimensional system dynamics, as it calculates the necessary optimality conditions before discretizing the PDE. As both methods are addressed in this work, we proceed by stating the necessary conditions of optimality in each of both cases.

- Necessary optimality conditions in case of finite-dimensional dynamics

The necessary conditions, corresponding to the finite-dimensional case, are mentioned briefly below, by referring to [Kir04], which provides a thorough demonstration of their origins.

Let s be the state of the system, U be its control input, d be the function

describing the dynamics, J be the cost function, $\mathcal{F}$ be the terminal cost, $\mathcal{G}$ be the cost-to-go, and t be the time falling in the interval $[t_0, t_f]$.

As a first step, the admissible state and control input are considered to be unbounded. We also assume that the initial state and time s_0 and t_0 are specified, the final time t_f is specified, whereas the final state s_f can be specified or free.

After discretization, the dynamics of the system can be described as:

$$\dot{s}(t) = d(s(t), U(t), t) \tag{5.31}$$

The objective of the controller is to find the optimal control input U^* that minimizes the cost function

$$J(U) = \mathcal{F}(s(t_f), t_f) + \int_{t_0}^{t_f} \mathcal{G}(s(t), U(t), t) dt \tag{5.32}$$

For this purpose, define the Hamiltonian

$$\mathcal{H}(s(t), U(t), p(t), t) = \mathcal{G}(s(t), U(t), t) + p^T(t) d(s(t), U(t), t) \tag{5.33}$$

where $p(t)$ is the adjoint state or co-state.

For all $t \in [t_0, t_f]$, the necessary conditions of optimality can be written as:

$$\dot{s}^*(t) = \frac{\partial \mathcal{H}}{\partial p}(s^*(t), U^*(t), p^*(t), t) \tag{5.34a}$$

$$\dot{p}^*(t) = -\frac{\partial \mathcal{H}}{\partial s}(s^*(t), U^*(t), p^*(t), t) \tag{5.34b}$$

$$0 = \frac{\partial \mathcal{H}}{\partial U}(s^*(t), U^*(t), p^*(t), t) \tag{5.34c}$$

$$\left[\frac{\partial \mathcal{F}}{\partial s}(s^*(t_f), t_f) - p^*(t_f)\right]^T \delta s_f + \left[\mathcal{H}(s^*(t_f), U^*(t_f), p^*(t_f), t_f) + \frac{\partial \mathcal{F}}{\partial t}(s^*(t_f), t_f)\right] \delta t_f = 0 \tag{5.35}$$

where s^*, U^*, and p^* are the optimal state, control input and co-state, respectively.

Note that the necessary conditions (5.34a), (5.34b), and (5.35) remain unchanged in case of a constrained input. Condition (5.34c), however, becomes

$$\mathcal{H}(s^*(t), U^*(t), p^*(t), t) \leq \mathcal{H}(s^*(t), U(t), p^*(t), t) \tag{5.36}$$

for all $t \in [t_0, t_f]$ and for all admissible $U(t)$. Equation (5.36) is, in fact, the Pontryagin Minimum Principle.

Considering the air-path control problem, the final time is specified ($t_f = T$) implying that $\delta t_f = 0$, and the final state is free implying that $\delta s_f \neq 0$. Therefore, in this case, (5.35) reduces to $p^*(T) = \frac{\partial \mathcal{F}}{\partial s}(s^*(T), T)$.

- Necessary optimality conditions in case of infinite-dimensional dynamics

For the calculation of the necessary conditions in the infinite-dimensional case, we refer to the Lagrangian-based adjoint method, from [HPUU08], which is a general optimal control approach that applies to linear and nonlinear PDEs. We consider minimizing a cost function subject to infinite-dimensional equality constraints, such as PDEs representing the system dynamics. We also consider the boundary-control case, because it is representative of the EGR control problem.

Let s be the state of the system, s_r be a bounded reference to be tracked, U be the system's control input, Z be a partial derivative operator, H be a bounded linear operator, q and r be strictly positive weights, x be the spatial coordinate falling in the interval $[0, 1]$, and t be the time coordinate falling in the interval $[t_0, t_f]$.

The dynamics of the system can be described by the following PDE:

$$\partial_t s(x, t) = Z(s(x, t)) \tag{5.37}$$

where $s(x, t_0) = s_0(x)$ is the initial condition, and $G_1(s(0, t)) = U(t), G_1(s(1, t)) = 0$ is one of two possible boundary-control scenarios.

The objective of the controller is to find the optimal control input U^* which minimizes the cost function

$$
\begin{aligned}
J(U) = &\int_{t_0}^{t_f} \int_0^1 |H(s(x, t) - s_r(x, t))|^2 dx dt + \int_{t_0}^{t_f} r|U(t)|^2 dt \\
&+ \int_0^1 q|H(s(x, t_f) - s_r(x, t_f))|^2 dx
\end{aligned}
\tag{5.38}
$$

For this purpose, define the Lagrangian

$$\mathcal{L}(s, U, p, \lambda) = \int_{t_0}^{t_f} \int_0^1 |H(s(x,t) - s_r(x,t))|^2 dx dt + \int_{t_0}^{t_f} r|U(t)|^2 dt$$

$$+ \int_0^1 q|H(s(x,t_f) - s_r(x,t_f))|^2 dx + \int_{t_0}^{t_f} < \lambda(t), (U(t) - G_1(s(0,t))) > dt$$

$$+ \int_{t_0}^{t_f} \int_0^1 < p(x,t), \partial_t s(x,t) - Z(s(x,t)) > dx dt$$

$$(5.39)$$

where $p(x,t)$ is the adjoint state, $\lambda(t)$ is the Lagrange multiplier associated to the boundary control $G_1(s(0,t)) = U(t)$, and $< \cdot, \cdot >$ denotes a scalar product.

The necessary conditions for optimality can be written as:

$$L_U(s^*, U^*, p^*, \lambda^*) = 0 \qquad (5.40a)$$
$$L_s(s^*, U^*, p^*, \lambda^*) = 0 \qquad (5.40b)$$
$$L_p(s^*, U^*, p^*, \lambda^*) = 0 \qquad (5.40c)$$
$$L_\lambda(s^*, U^*, p^*, \lambda^*) = 0 \qquad (5.40d)$$
$$p^*(x, t_f) = 2qH(s^*(x, t_f) - s_r(x, t_f)) \qquad (5.40e)$$
$$s^*(x, t_0) = s_0(x) \qquad (5.40f)$$

where the index of L denotes its partial derivatives with respect to the variable.

5.2.4 Augmented Lagrangian method

The literature proposes different methods for constrained optimization, such as the penalty methods, the primal approach, or the primal-dual approach (Lagrangian method). The dual approach (Augmented Lagrangian method), which is the approach we use in this work, shares some features with these methods. Following is a preface which discusses some of these ideas before attacking the Augmented Lagrangian method. Starting off by the penalty methods, they are methods which approximate the constrained optimization problem by an unconstrained optimization problem by adding to the objective function a penalty function, which is a term that penalizes the violation of the constraints, thus resulting in an augmented objective function. Here is an example of an exterior penalty function $P(\mu)$ used in the case of inequality constraint $Z_i(\mu) \leq 0$ for $i = 1$.

- When $Z_i(\mu) \leq 0 : P(\mu) = 0$ because the constraint is satisfied.

- When $Z_i(\mu) > 0 : P(\mu) > 0$ because the constraint is violated. Take for example $P(\mu) = Z_i(\mu)$.

Therefore, the penalty function can be formulated as $P(\mu) = max(0, Z_i(\mu))$ or as $P(\mu) = max(0, Z_i(\mu))^p$ (where p is an integer) for a more severe penalization. Moving on to the primal-dual method, as indicated in its name, it associates to the primal problem, which aims at minimizing the original objective function, a dual problem, which aims at maximizing a Lagrange dual function. To illustrate the Lagrangian method in case of inequality constraints, we consider the following standard constrained optimization problem:

$$
\begin{aligned}
minimize \quad & M(\mu) \\
s.t. \quad & Z_i(\mu) \leq 0, \quad i = 1, 2, ...m
\end{aligned}
\tag{5.41}
$$

In what follows, we define the Lagrangian, Lagrange dual function, and Lagrange dual problem corresponding to the primal minimization problem (5.41), by referring to [BV04].

The Lagrangian $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is defined as:

$$
\mathcal{L}(\mu, \lambda) = M(\mu) + \sum_{i=1}^{m} \lambda_i Z_i(\mu)
\tag{5.42}
$$

where $\lambda \geq 0$ is the dual variable or Lagrange multiplier vector, composed of λ_i corresponding to the inequality constraints $Z_i(\mu) \leq 0$.

The Lagrange dual function $\Theta : \mathbb{R}^m \to \mathbb{R}$ is defined as:

$$
\Theta(\lambda) = \min_{\mu} \mathcal{L}(\mu, \lambda)
\tag{5.43}
$$

An important property of the dual function $\Theta(\lambda)$ is that, for $\lambda \geq 0$, it is a lower bound on $M(\mu)$ for all feasible values of μ, which implies that it is also a lower bound of the optimal value M^* of problem (5.41).

The Lagrange dual problem is the optimization problem defined as:

$$
\begin{aligned}
maximize \quad & \Theta(\lambda) \\
s.t. \quad & \lambda \geq 0
\end{aligned}
\tag{5.44}
$$

For each $\lambda \geq 0$, there exists a dual function $\Theta(\lambda)$, i.e. there exists a lower bound for M^*. Nevertheless, a lower bound is only meaningful

when $\lambda \geq 0$ and $\Theta(\lambda) > -\infty$. In this case, the Lagrange multiplier λ is said to be dual feasible (feasible for the dual problem (5.44)). The idea behind the Lagrangian dual problem is to find $\lambda \geq 0$ which results in the best lower bound $\Theta(\lambda)$, i.e. to find the dual optimal Lagrange multiplier λ^*.

Always referring to [BV04], the concept of weak and strong Lagrangian duality is briefly introduced. To do so, the optimal solution M^* of the primal problem (5.41) can be formulated as:

$$M^* = \min_{\mu} \max_{\lambda \geq 0} \mathcal{L}(\mu, \lambda) \tag{5.45}$$

This is justifiable because
$\max_{\lambda \geq 0} \mathcal{L}(\mu, \lambda) = \max_{\lambda \geq 0} \left(M(\mu) + \sum_{i=1}^{m} \lambda_i Z_i(\mu) \right)$ is equal to ∞ when the constraints are not respected, but it is equal to $M(\mu)$ when $Z_i(\mu) \leq 0$.

Also, from (5.43) and (5.44), the optimal solution Θ^* of the dual problem (5.44) can be formulated as:

$$\Theta^* = \max_{\lambda \geq 0} \min_{\mu} \mathcal{L}(\mu, \lambda) \tag{5.46}$$

We talk of weak duality when

$$\max_{\lambda \geq 0} \min_{\mu} \mathcal{L}(\mu, \lambda) \leq \min_{\mu} \max_{\lambda \geq 0} \mathcal{L}(\mu, \lambda)$$

and we talk of strong duality when

$$\max_{\lambda \geq 0} \min_{\mu} \mathcal{L}(\mu, \lambda) = \min_{\mu} \max_{\lambda \geq 0} \mathcal{L}(\mu, \lambda)$$

The primal-dual optimal point (μ^*, λ^*) satisfying the strong duality condition is called a saddle-point for the Lagrangian.

The Augmented Lagrangian approach, proposed by [Hes69] and [Pow69], dates back to the year 1969. Similarly to penalty methods, it transforms a constrained optimization problem to an unconstrained optimization problem by adding a penalty term penalizing the constraint violation. However, it considers the Lagrangian, which is why it is also called the "Method of Multipliers" or the "Penalty-Multiplier Method", and is considered as an advantageous method which merges many of the merits of both methods, the penalty and the local duality methods.

Since its first introduction in 1969, different variants of the Augmented Lagrangian method have been introduced to improve the quality of the solution or to solve different kinds of constrained optimization problems; take for example the alternating direction method of multipliers

(ADMM) or the generalized Augmented Lagrangians (GAL) method. In the rest of this subsection, however, we stick to the classical Augmented Lagrangian method considering the case of equality constraints, presented in [Hes69]. Since we are dealing with inequality constraints, we show how a simple manipulation of the constraint inequality function can transform the inequality constraint problem into an equality constraint problem, upon which the classical method can be applied.

By referring to [Hes69], consider the equality constrained problem:

$$
\begin{aligned}
minimize \quad & M(\mu) \\
s.t. \quad & e(\mu) = 0
\end{aligned}
\tag{5.47}
$$

where $M : \mathbb{R}^n \to \mathbb{R}$ and $e : \mathbb{R}^n \to \mathbb{R}$ are assumed to be twice continuously differentiable functions. To solve the minimization problem (5.47), the Augmented Lagrangian method suggests using the augmented Lagrangian function

$$
\mathcal{L}_a(\mu, \lambda) = M(\mu) + \lambda^T e(\mu) + \frac{c}{2}|e(\mu)|^2
\tag{5.48}
$$

where c is a positive constant chosen sufficiently large, and λ is the Lagrange multiplier whose appropriate value is to be computed. The reason behind this suggestion is that if μ^* is a minimum point of the function $\mathcal{L}_a$ respecting the equality constraint $e(\mu^*) = 0$, then, according to [Hes69], μ^* is also a minimum to M and respects the equality constraint $e = 0$. In other words, μ^* is also a solution of problem (5.47).

To compute the appropriate value of λ starting from an initial estimate $\lambda^{(0)}$, the following update law is used:

$$
\lambda^{(k+1)} = \lambda^{(k)} + c^{(k)} e(\mu^{(k)})
\tag{5.49}
$$

where $0 < c^{(k)} \leq c$ and $\mu^{(k)} = \underset{\mu}{\operatorname{argmin}}\, \mathcal{L}_a(\mu, \lambda^{(k)})$.

Not only is this method useful for minimization problems with equality constraints, but it is also able to deal with certain inequality constraints. To illustrate the incorporation of inequality constraints in the Augmented Lagrangian method, we reconsider the minimization problem (5.41), taking $i = 1$ for simplicity. Under certain assumptions [LY08]; [Ber76], this minimization problem can be written as an equality-constrained minimization problem, similar to problem (5.47), where

$$
e_i(\mu) = Z_i(\mu) + u_i = 0
\tag{5.50}
$$

u_i being a slack variable equal or greater than zero. By doing so, the augmented Lagrangian in case of equality constraints (5.48) can be reformulated as:

$$\mathcal{L}_a(\mu, \lambda) = M(\mu) + \sum_{i=1}^{m} \{\lambda_i^T \left[Z_i(\mu) + u_i\right] + \frac{c}{2} \left[Z_i(\mu) + u_i\right]^2\} \tag{5.51}$$

The dual function can therefore be written as:

$$\Theta(\lambda) = \min_{u \geq 0, \mu} \mathcal{L}_a(\mu, \lambda) = \min_{u \geq 0, \mu} \{M(\mu) + \sum_{i=1}^{m} \{\lambda_i^T \left[Z_i(\mu) + u_i\right] + \frac{c}{2} \left[Z_i(\mu) + u_i\right]^2\}\}$$
$$\tag{5.52}$$

The first minimization of the dual function $\Theta(\lambda)$ is with respect to $u \geq 0$, and then its minimization with respect to μ follows. It is shown in [LY08] and [Ber76] that minimizing $\Theta(\lambda)$ with respect to $u \geq 0$ yields an optimal value of u_i equal to:

$$u_i = max\{0, -Z_i(\mu) - \frac{\lambda_i}{c}\} \tag{5.53}$$

Substituting for this u_i in equation (5.52) allows its reformulation as [LY08]; [Ber76]; [Ius99]:

$$\Theta(\lambda) = \min_{\mu} \{M(\mu) + \frac{1}{2c} \sum_{i=1}^{m} \{[max\{0, \lambda_i + cZ_i(\mu)\}]^2 - \lambda_i^2\}\} \tag{5.54}$$

The minimization with respect to μ now takes place as per equation (5.54), yielding a value $\mu^{(k)}$ for each iteration k. Finally, in view of equations (5.49), (5.50), and (5.53), the appropriate value of λ_i can be computed as:

$$\lambda_i^{(k+1)} = max\{0, \lambda_i^{(k)} + cZ_i(\mu^{(k+1)})\} \tag{5.55}$$

This iterative solution of the constrained optimization problem is, in fact, the Uzawa algorithm. The Uzawa algorithm was first introduced in [Uza58] and adapted to the augmented Lagrangian in [FG83] to become the Augmented Lagrangian Uzawa Method (ALUM), which is nothing but the Uzawa algorithm applied to the saddle-point problem. ALUM can be seen as the projected gradient-descent method applied to the constrained dual problem, to maximize the dual function (5.52) for $\lambda \geq 0$. It substitutes the constrained dual problem by a sequence of unconstrained problems, and solves them iteratively in three basic steps:

- Choose an initial nonnegative Lagrange multiplier $\lambda^{(0)}$.

- Calculate $\mu^{(k+1)}$, solution of the unconstrained problem (5.54).

- Update the Lagrange multiplier $\lambda^{(k+1)}$, as per equation (5.55).

5.2.5   Indirect method: discretize-then-optimize approach

In this approach, discretization of the system precedes the calculation of the necessary optimality conditions. Its objective is to transform the system into a delay-free system to which we can attribute delay-free control. System (5.28) is an output delay system with a time-varying delay equal to $\tau_{total}(t)$. In what follows, for simplicity, we consider a time-invariant delay, denoted as τ_{total}, which is the average of the time-varying delays $\tau_{total}(t)$. To transform system (5.28) into a delay-free SS system, we use equation (5.29) and discretize the transport PDE along the x-direction, using the method of lines [Ros71]. Figure 5.14 shows the tube section's discretization points, equal to $x = idw$, and their corresponding values, $X_i(t) = X(x = idw, t)$. $X_0(t) = X(0, t) = X_{apc}(t) = \mathcal{X}(t)$ is the BGR at the entrance of the tube section and $X_N(t) = X(1, t) = X_{apc}(t - \tau_{total}) = \mathcal{X}(t - \tau_{total}) = Y(t)$ is the BGR at the exit of the tube section of length $L = 1$.
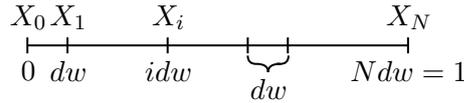
Figure 5.14: Discretization of $X(x, t)$ along the x-direction

This discretization scheme results in a vector of discretization points $\mathcal{Z} = \begin{bmatrix} X_1 & X_2 & \dots & X_N \end{bmatrix}^T \in \mathbb{R}_{\geq 0}^N$, upon which we discretize the transport PDE in equation (5.29), in time using an explicit forward Euler scheme, and in space using an implicit backward discretization scheme:

$$\dot{X}_i(t) = -\frac{1}{\tau_{total}} \frac{X_i - X_{i-1}}{dw} \tag{5.56}$$

For $i = 1$: $\dot{X}_1 = -\frac{1}{\tau_{total}dw}(X_1 - X_0) = -\frac{1}{\tau_{total}dw}(X_1 - \mathcal{X}(t))$

For $i = 2$: $\dot{X}_2 = -\frac{1}{\tau_{total}dw}(X_2 - X_1)$

Discretizing the system forward in time and backward in space results in a discretization scheme which is conditionally stable. This was shown in Von Neumann's stability analysis [CFN50], where the stability condition is such that $0 \leq \sigma = \frac{u\Delta t}{\Delta w} \leq 1$, u being the positive speed of the wave equation, Δt being the time discretization step, and Δw being the space discretization step. Because we are discretizing a hyperbolic PDE, σ is also called the Courant-Friedrichs-Lewy (CFL) number, which first appeared in [CFL28] in 1927. In case of the EGR model, $u = \frac{1}{\tau_{total}} = 10$,

$\Delta t = 0.01$, and $\Delta w = 0.1$, which implies that $\sigma = 1$, thus ensuring the stability of the discretization scheme.

Consequently, using (5.56), the discretized version of the transport PDE in equation (5.29) can be expressed as:

$$\dot{\mathcal{Z}}(t) = F \cdot \mathcal{Z}(t) + G \cdot \mathcal{X}(t) \tag{5.57}$$

where $\quad F = \begin{bmatrix} -\frac{1}{\tau_{total}dw} & & & \\ \frac{1}{\tau_{total}dw} & \ddots & & \mathbf{0} \\ & \ddots & \ddots & \\ \mathbf{0} & & \frac{1}{\tau_{total}dw} & -\frac{1}{\tau_{total}dw} \end{bmatrix} \in \mathbb{R}^{N \times N}$

and $\quad G = \begin{bmatrix} \frac{1}{\tau_{total}dw} & 0 & \cdots & 0 \end{bmatrix}^T \in \mathbb{R}^N_{\geq 0}$.

As a result of equations (5.28) and (5.57), the delay-free time-varying SS model describing the dynamics of the system can be written as:

$$\dot{\mathcal{X}\mathcal{Z}}(t) = \Lambda(t) \cdot \mathcal{X}\mathcal{Z}(t) + \Xi(t) \cdot U(t) \tag{5.58}$$

where $\mathcal{X}\mathcal{Z}(t) = \begin{bmatrix} \mathcal{X}(t) & \mathcal{Z}(t)^T \end{bmatrix}^T \in \mathbb{R}^{N+1}_{\geq 0}$, $\mathcal{X}\mathcal{Z}(0) = \mathcal{X}\mathcal{Z}_0$ is the state vector at $t = 0$, $\Lambda(t) = \begin{bmatrix} A(t) & 0_{1 \times N} \\ G & F \end{bmatrix} \in \mathbb{R}^{(N+1) \times (N+1)}$, $\Xi(t) = \begin{bmatrix} B(t) \\ 0_{N \times 1} \end{bmatrix} \in \mathbb{R}^{N+1}$, and $0_{a \times b}$ is an $(a \times b)$ zero matrix.

At this stage, the availability of the delay-free model allows for a delay-free control design. Based on equation (5.58), the first step of this control design is to reformulate the objective function (5.30) as:

$$J_2(t) = \frac{1}{2}\|\mathcal{X}\mathcal{Z}_{N+1}(T) - r(T)\|^2_{L_2} + \frac{1}{2} \int_0^T (\|\mathcal{X}\mathcal{Z}_{N+1}(t) - r(t)\|^2_{Q_2} + \|U(t)\|^2_{R_2})dt \tag{5.59}$$

where $L_2 \in \mathbb{R}_{\geq 0}$, $Q_2 \in \mathbb{R}_{\geq 0}$, and $R_2 \in \mathbb{R}_{>0}$ are the weights of the cost function J_2,

and to reformulate the inequality constraints as:

$$0 \leq \mathcal{X}\mathcal{Z}_{N+1}(t) \leq 100$$
$$0 \leq U(t) \leq 100$$

While the constraints on $U(t)$ are taken care of using the projected gradient descent method, $\mathcal{X}\mathcal{Z}_{N+1}$ on the other hand is considered in the state-constraint vector $I \in \mathbb{R}^2_{\leq 0}$, therefore expressed as:

$$I(\mathcal{X}\mathcal{Z}(t))) = \begin{bmatrix} \mathcal{X}\mathcal{Z}_{N+1}(t) - 100 \\ -\mathcal{X}\mathcal{Z}_{N+1}(t) \end{bmatrix}$$

In order to calculate the necessary conditions of optimality, the augmented Hamiltonian is introduced, and is expressed as:

$$\mathcal{H}(t) = \frac{1}{2}\|\mathcal{X}\mathcal{Z}_{N+1}(t) - r(t)\|^2_{Q_2} + \frac{1}{2}\|U(t)\|^2_{R_2} + p^T(t)(\Lambda(t) \cdot \mathcal{X}\mathcal{Z}(t) + \Xi(t) \cdot U(t))$$
$$+ \zeta_c(\lambda, I(\mathcal{X}\mathcal{Z}(t)))$$

(5.60)

where $p(t) \in \mathbb{R}^{N+1}$ is the co-state vector,
and $\zeta_c(\lambda, I(\mathcal{X}\mathcal{Z}(t))) = \frac{1}{2c}(\|max(0_{2\times1}, \lambda + cI(\mathcal{X}\mathcal{Z}(t)))\|^2 - \|\lambda\|^2) \in \mathbb{R}^2_{\geq0}$,
$\lambda \in \mathbb{R}^2_{\geq0}$ is the Lagrange multiplier, and c is a positive scalar.

From the augmented Hamiltonian, the necessary conditions of optimality can be derived as:

$$\frac{\partial\mathcal{H}(t)}{\partial p(t)} = \dot{\mathcal{X}\mathcal{Z}}(t) = \Lambda(t) \cdot \mathcal{X}\mathcal{Z}(t) + \Xi(t) \cdot U(t)$$

(5.61)

$$\frac{\partial\mathcal{H}(t)}{\partial \mathcal{X}\mathcal{Z}(t)} = -\dot{p}(t) = Q_2(\mathcal{X}\mathcal{Z}_{N+1}(t) - r(t)) + \Lambda(t)^T p(t) + \frac{\partial\zeta_c(\lambda, I(\mathcal{X}\mathcal{Z}(t)))}{\partial\mathcal{X}\mathcal{Z}(t)}$$

(5.62)

$$\mathcal{H}(\mathcal{X}\mathcal{Z}^*(t), U^*(t), p^*(t), \lambda^*(t), t) \leq \mathcal{H}(\mathcal{X}\mathcal{Z}^*(t), U(t), p^*(t), \lambda^*(t), t)$$

(5.63)

for all $t \in [0, T]$ and for all admissible $U(t)$.

Note that

$$\frac{\partial\zeta_c(\lambda, I(\mathcal{X}\mathcal{Z}(t)))}{\partial\mathcal{X}\mathcal{Z}(t)}$$
$$= \frac{\partial I(\mathcal{X}\mathcal{Z}(t))}{\partial\mathcal{X}\mathcal{Z}(t)} \cdot \frac{\partial\zeta_c(\lambda, I(\mathcal{X}\mathcal{Z}(t)))}{\partial I(\mathcal{X}\mathcal{Z}(t)))}$$
$$= \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 1 & -1 \end{bmatrix} max(0_{2\times1}, \lambda + cI(\mathcal{X}\mathcal{Z}(t)))$$

$$\frac{\partial\mathcal{H}(t)}{\partial U(t)} = U(t)^T R_2 + p(t)^T \Xi(t)$$

(5.64)

$$\frac{\partial \mathcal{H}(t)}{\partial \lambda} = \frac{\partial \zeta_c(\lambda, I(\mathcal{XZ}(t)))}{\partial \lambda} = \frac{1}{c}(max(0_{2\times 1}, \lambda + cI(\mathcal{XZ}(t))) - \lambda) \qquad (5.65)$$

Uzawa method is used to solve this problem numerically, and the algorithm will execute the following steps:

1. Choose arbitrary values of $\lambda^{(0)} \in \mathbb{R}_{\geq 0}$ and $U(t)^{(0)}$, where the subscript $k \in [0, \infty)$ refers to the number of iteration.

2. Calculate $\mathcal{XZ}(t)$ using (5.58) and $\mathcal{XZ}(0)$

3. Calculate $p(t)$ using (5.62) and $p(T) = L_2(\mathcal{XZ}_{N+1}(T) - r(T))$

4. Update $U(t)$

$$U^{(k+1)}(t) = U^{(k)}(t) - v_1 \cdot \frac{\partial \mathcal{H}(t)}{\partial U(t)}$$

   where $\frac{\partial \mathcal{H}(t)}{\partial U(t)}$ is obtained from (5.64) and v_1 is a positive constant.

5. Saturate $U^{(k+1)}(t)$ by projecting it on the feasible set.

$$U^{(k+1)}(t) = \begin{cases} 0 & , U^{(k+1)}(t) < 0 \\ U^{(k+1)}(t) & , 0 \leq U^{(k+1)}(t) \leq 100 \\ 100 & , U^{(k+1)}(t) > 100 \end{cases}$$

6. For each $\lambda^{(k)}$, verify that $\|U(t)^{(k+1)} - U(t)^{(k)}\| \leq \epsilon$, where ϵ is a positive constant. If condition is true, go to step 7. Otherwise, go back to step 2.

7. Update λ

$$\lambda^{(k+1)} = \lambda^{(k)} + v_2 \cdot \frac{\partial \mathcal{H}(t)}{\partial \lambda}$$
$$= \lambda^{(k)} + \frac{v_2}{c}(max(0_{2\times 1}, \lambda^{(k)} + cI(\mathcal{XZ}(t))) - \lambda^{(k)})$$

   where v_2 is a positive constant. This update law reduces to equation (5.55) when $v_2 = c$.

8. Verify that $\|\lambda^{(k+1)} - \lambda^{(k)}\| \leq \epsilon$. If condition is true, terminate. Otherwise, go back to step 3.

5.2.6 Indirect method: optimize-then-discretize approach

In this approach, following equation (5.29), the objective function (5.30) can be reformulated as:

$$J_3(t) = \frac{1}{2}\|X(1,T) - r(T)\|_{L_3}^2 + \frac{1}{2}\int_0^T \left(\|X(1,t) - r(t)\|_{Q_3}^2 + \|U(t)\|_{R_3}^2\right) dt$$
(5.66)

where $L_3 \in \mathbb{R}_{\geq 0}$, $Q_3 \in \mathbb{R}_{\geq 0}$, and $R_3 \in \mathbb{R}_{>0}$ are the weights of the cost function J_3,

and the inequality constraints can be reformulated as:

$$0 \leq X(1,t) \leq 100$$
$$0 \leq U(t) \leq 100$$

Similar to the discretize-then-optimize approach, the constraints on $U(t)$ are taken care of using the projected gradient descent method, whereas the constraints on $X(1,t)$ are considered in the state-constraint vector $I(X(1,t)) \in \mathbb{R}_{\leq 0}^2$, expressed as: $I(X(1,t)) = \begin{bmatrix} X(1,t) - 100 \\ -X(1,t) \end{bmatrix}$

To calculate the necessary optimality conditions, the augmented Lagrangian is introduced, and is expressed as:

$$\mathcal{L}(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x,t), p_1(t), p_2(t))$$
$$= \frac{1}{2}\|X(1,T) - r(T)\|_{L_3}^2 + \frac{1}{2}\int_0^T \left(\|X(1,t) - r(t)\|_{Q_3}^2 + \|U(t)\|_{R_3}^2\right) dt$$
$$+ \int_0^T p_1^T(t)\left(A(t)\mathcal{X}(t) + B(t)U(t) - \dot{\mathcal{X}}(t)\right) dt$$
$$+ \int_0^T \int_0^1 p_2^T(x,t)\left(\partial_t X(x,t) + \frac{1}{\tau_{total}}\partial_x X(x,t)\right) dxdt$$
$$+ \int_0^T \zeta_c(\lambda, I(X(1,t)))dt$$
(5.67)

where $p_1(t) \in \mathbb{R}$ and $p_2(x,t) \in \mathbb{R}$ are the Lagrange multipliers associated to the ODE and PDE dynamics respectively,
$\zeta_c(\lambda, I(X(1,t))) = \frac{1}{2c}(\|max(0_{2\times 1}, \lambda + cI(X(1,t)))\|^2 - \|\lambda\|^2)$, $\lambda \in \mathbb{R}_{\geq 0}^2$ is the Lagrange multiplier, and c is a positive scalar.

To ease the calculation of the necessary optimality conditions, we first expand and clarify some terms in the Lagrangian and reformulate its double integral terms using integration by parts.

$$\int_0^T -p_1^T(t)\dot{\mathcal{X}}(t)dt$$

$$= -p_1^T(t)\mathcal{X}(t)\big|_0^T + \int_0^T \mathcal{X}(t)\dot{p_1}^T(t)dt$$

$$= -p_1^T(T)\mathcal{X}(T) + p_1^T(0)\mathcal{X}(0) + \int_0^T \frac{dp_1^T(t)}{dt}\mathcal{X}(t)dt$$

$$\int_0^T \int_0^1 p_2^T(x,t)\partial_t X(x,t)dxdt$$

$$= \int_0^1 \left(p_2^T(x,t)X(x,t)\big|_0^T - \int_0^T X(x,t)\partial_t p_2^T(x,t)dt \right) dx$$

$$= \int_0^1 \left(p_2^T(x,t)X(x,T) - p_2^T(x,0)X(x,0) \right) dx - \int_0^1 \int_0^T X(x,t)\partial_t p_2^T(x,t)dtdx$$

$$\int_0^T \int_0^1 p_2^T(x,t)\frac{1}{\tau_{total}}\partial_x X(x,t)dxdt$$

$$= \int_0^T \left(p_2^T(x,t)\frac{1}{\tau_{total}}X(x,t)\big|_0^1 - \int_0^1 X(x,t)\frac{1}{\tau_{total}}\partial_x p_2^T(x,t)dx \right) dt$$

$$= \int_0^T \left(p_2^T(1,t)\frac{1}{\tau_{total}}X(1,t) - p_2^T(0,t)\frac{1}{\tau_{total}}X(0,t) \right) dt - \int_0^T \int_0^1 \frac{1}{\tau_{total}}X(x,t)\partial_x p_2^T(x,t)dxdt$$

Consequently, the Lagrangian can be expressed as:

$$\mathcal{L}(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x,t), p_1(t), p_2(t))$$

$$= \frac{1}{2}\|X(1,T) - r(T)\|_{L_3}^2 + \frac{1}{2}\int_0^T \left(\|X(1,t) - r(t)\|_{Q_3}^2 + \|U(t)\|_{R_3}^2 \right) dt$$

$$+ \int_0^T p_1^T(t)\left(A(t)\mathcal{X}(t) + B(t)U(t) \right) dt + p_1^T(0)\mathcal{X}(0) - p_1^T(T)\mathcal{X}(T)$$

$$+ \int_0^T \frac{dp_1^T(t)}{dt}\mathcal{X}(t)dt + \int_0^1 \left(p_2^T(x,t)X(x,T) - p_2^T(x,0)X(x,0) \right) dx$$

$$- \int_0^T \int_0^1 \left(\frac{1}{\tau_{total}}X(x,t)\partial_x p_2^T(x,t) + X(x,t)\partial_t p_2^T(x,t) \right) dxdt$$

$$+ \int_0^T \left(p_2^T(1,t)\frac{1}{\tau_{total}}X(1,t) - p_2^T(0,t)\frac{1}{\tau_{total}}\mathcal{X}(t) \right) dt$$

$$+ \int_0^T \zeta_c(\lambda, I(X(1,t)))dt$$

$$(5.68)$$

From the augmented Lagrangian, the necessary conditions of optimality can be derived by studying the variation of $\mathcal{L}(\cdot)$ with respect to $U(t)$, $\mathcal{X}(t)$, and $X(x,t)$, denoted respectively as $\mathcal{L}_U(\cdot)$, $\mathcal{L}_\mathcal{X}(\cdot)$, and $\mathcal{L}_X(\cdot)$.

$$\mathcal{L}_U(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x,t), p_1(t), p_2(t)) = \int_0^T \left(U(t)^T R_3 + B^T(t)p_1(t) \right) \delta U(t) dt$$

$$\mathcal{L}_\mathcal{X}(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x,t), p_1(t), p_2(t))$$
$$= \int_0^T \left(A^T(t)p_1(t) - \frac{1}{\tau_{total}} p_2^T(0,t) + \frac{dp_1^T(t)}{dt} \right) \delta\mathcal{X}(t) dt - p_1^T(T)\delta\mathcal{X}(T) - p_1^T(0)\delta\mathcal{X}(0)$$

$$\mathcal{L}_X(\mathcal{X}(t), \dot{\mathcal{X}}(t), U(t), X(x,t), p_1(t), p_2(t))$$
$$= L_3 \left(X(1,T) - r(T) \right) \delta X(1,T) + \int_0^1 \left(p_2^T(x,T)\delta X(x,T) - p_2^T(x,0)\delta X(x,0) \right) dx$$
$$+ \int_0^T \left(Q_3 \left(X(1,t) - r(t) \right) \delta X(1,t) + \frac{1}{\tau_{total}} p_2^T(1,t)\delta X(1,t) + \frac{\partial\zeta_c(\lambda, I(X(1,t)))}{\partial X(1,t)} \delta X(1,t) \right) dt$$
$$- \int_0^T \int_0^1 \left(\frac{1}{\tau_{total}} \partial_x p_2^T(x,t)\delta X(x,t) + \partial_t p_2^T(x,t)\delta X(x,t) \right) dxdt$$

where

$$\frac{\partial\zeta_c(\lambda, I(X(1,t)))}{\partial X(1,t)} = \frac{\partial\zeta_c(\lambda, I(X(1,t)))^T}{\partial I(X(1,t)))} \cdot \frac{\partial I(X(1,t))}{\partial X(1,t)}$$
$$= max(0_{2\times1}, \lambda + cI(X(1,t)))^T \begin{bmatrix} 1 & -1 \end{bmatrix}^T$$
$$= max(0, \lambda_1 + cX(1,t) - 100c) - max(0, \lambda_2 - cX(1,t))$$

To obtain the necessary optimality conditions, these variations are set to zero, while admissible (i.e. $\delta\mathcal{X}(0) = \delta X(x,0) = 0$), which gives:

- $\delta U(t) \quad = 0 \implies U(t)^T R_3 + B^T(t)p_1(t) = 0 \Rightarrow U(t) = -R_3^{-1}B(t)^T(t)p_1(t)$

- $\delta \mathcal{X}(t) \quad = 0 \implies \dfrac{dp_1(t)}{dt} + A^T(t)p_1(t) = \dfrac{1}{\tau_{total}}p_2(0,t)$

- $\delta \mathcal{X}(T) \quad = 0 \implies -p_1(T) = 0 \Rightarrow p_1(T) = 0$

- $\delta X(x,t) = 0 \implies \partial_t p_2(x,t) + \dfrac{1}{\tau_{total}}\partial_x p_2(x,t) = 0$

- $\delta X(x,T) = 0 \implies p_2(x,T) = 0$

- $\delta X(1,t) \; = 0 \implies Q_3\left(X(1,t) - r(t)\right) + \dfrac{1}{\tau_{total}}p_2(1,t)$

$$+ max(0, \lambda_1 + cX(1,t) - 100c) - max(0, \lambda_2 - cX(1,t)) = 0$$

$$\implies p_2(1,t) = \tau_{total}Q_3\left(X(1,t) - r(t)\right)$$

$$+ \tau_{total}\left(max(0, \lambda_1 + cX(1,t) - 100c) - max(0, \lambda_2 - cX(1,t))\right)$$

- $\delta X(1,T) = 0 \implies X(1,T) - r(T) = 0 \implies X(1,T) = r(T)$

For the sake of clarity, we hereby sum up the boundary-value problem resulting from the system dynamics and the necessary optimality conditions.

Equations describing the dynamics of the system:

$$\partial_t X(x,t) + \frac{1}{\tau_{total}}\partial_x X(x,t) = 0 \qquad\qquad ((5.29)\text{ revisited})$$

Initial condition $\qquad\quad : X(x,0)$
Boundary conditions $\quad : X(0,t) \text{ and } X(1,t)$
Terminal condition $\qquad : X(1,T) = r(T)$

$$\dot{\mathcal{X}} = A(t)\mathcal{X}(t) - B(t)R_3^{-1}B^T(t)p_1(t) \qquad\qquad (5.69)$$

Initial condition $\qquad\quad : \mathcal{X}(0) = \mathcal{X}_0$

Equations describing the dynamics of the adjoint states:

$$\frac{dp_1(t)}{dt} + A^T(t)p_1(t) = \frac{1}{\tau_{total}}p_2(0,t) \qquad\qquad (5.70)$$

Terminal condition $\qquad : p_1(T) = 0$

$$\partial_t p_2(x,t) + \frac{1}{\tau_{total}}\partial_x p_2(x,t) = 0 \qquad\qquad (5.71)$$

Terminal condition $\qquad : p_2(x,T) = 0$
Boundary condition $\quad : p_2(1,t) = \tau_{total}Q_3\left(X(1,t) - r(t)\right)$
$\qquad\qquad +\tau_{total}\left(max(0, \lambda_1 + cX(1,t) - 100c) - max(0, \lambda_2 - cX(1,t))\right)$

To solve this boundary-value problem, we discretize the PDEs in equations (5.29) and (5.71). Concerning the PDE in equation (5.29), as in the previous section, it is discretized forward in time and backward in space, upon equation (5.56), resulting in its discretized version (5.57). Concerning equation (5.71), it is discretized, along the x-direction, using the method of lines, in a discretization scheme similar to that shown in Figure 5.14 (p instead of X), where $p_{2_i}(t) = p_2(x = idw, t)$, $p_{2_0}(t) = p_2(0, t)$ is taken at the entrance of the tube section, and $p_{2_N}(t) = p_2(1, t)$ is taken at the exit of the tube section of length $L = 1$. This discretization scheme results in a vector of discretization points $W = \begin{bmatrix} p_{2_0} & p_{2_1} & \cdots & p_{2_{N-1}} \end{bmatrix}^T$, upon which we discretize equation (5.71) in time using a backward Euler scheme, and in space using a forward discretization scheme:

$$\dot{p}_{2_i}(t) = -\frac{1}{\tau_{total}} \frac{p_{2_{i+1}} - p_{2_i}}{dw} \tag{5.72}$$

For $i = 0$: $\dot{p}_{2_0} = -\frac{1}{\tau_{total}dw}(p_{2_1} - p_{2_0})$

For $i = N - 1$: $\dot{p}_{2_{N-1}} = -\frac{1}{\tau_{total}dw}(p_{2_N} - p_{2_{N-1}})$

This discretization scheme is used in this case because p_{2_N} is the known boundary condition and it is used to calculate all the other values p_{2_i}. Consequently, the discretized version of equation (5.71) can be expressed as:

$$\dot{W}(t) = KW(t) + Dp_2^T(1, t) \tag{5.73}$$

where $K = \begin{bmatrix} \frac{1}{\tau_{total}dw} & -\frac{1}{\tau_{total}dw} & & \mathbf{0} \\ & \ddots & \ddots & \\ & & \ddots & -\frac{1}{\tau_{total}dw} \\ \mathbf{0} & & & \frac{1}{\tau_{total}dw} \end{bmatrix} \in \mathbb{R}^{N \times N}$

and $D = \begin{bmatrix} 0 & \cdots & 0 & \frac{-1}{\tau_{total}dw} \end{bmatrix}^T \in \mathbb{R}^N_{\leq 0}$.

Now that the discretization of the PDEs is done, the boundary-value problem can be reformulated. By concatenating equations (5.57) and (5.69), and concatenating equations (5.70) and (5.73), the boundary-value problem can be expressed as:

$$\dot{\mathcal{X}\mathcal{Z}}(t) = \Lambda(t) \cdot \mathcal{X}\mathcal{Z}(t) + \Upsilon(t) \cdot PP(t) \tag{5.74}$$

$$\dot{PP}(t) = \Gamma(t) \cdot PP(t) + \Omega \cdot p_2(1, t) \tag{5.75}$$

where $\mathcal{XZ}$ and Λ are given in equation (5.58),
$PP(t) = \begin{bmatrix} p_1(t) & W(t)^T \end{bmatrix}^T$, $\Omega = \begin{bmatrix} 0 & D^T \end{bmatrix}^T$,

$$\Upsilon(t) = \begin{bmatrix} -B(t)R_3^{-1}B(t)^T & 0_{1\times N} \\ 0_{N\times 1} & 0_{N\times N} \end{bmatrix}, \text{ and}$$

$$\Gamma(t) = \begin{bmatrix} -A^T(t) & \begin{bmatrix} \frac{1}{\tau_{total}} & 0_{1\times N-1} \end{bmatrix} \\ 0_{N\times 1} & K \end{bmatrix}.$$

By referring to the boundary condition $p_2(1, t)$ corresponding to equation (5.71), (5.75) becomes

$$
\begin{aligned}
\dot{PP}(t) &= \Gamma(t) \cdot PP(t) - \Omega \cdot \tau_{total} \cdot Q_3 \cdot X(1, t) + \Omega \cdot \tau_{total} \cdot Q_3 \cdot r(t) \\
&\quad - \Omega \cdot \tau_{total} \cdot max\left(0, \lambda_1 + cX(1, t) - 100c\right) + \Omega \cdot \tau_{total} \cdot max\left(0, \lambda_2 - cX(1, t)\right) \\
&= \Gamma(t) \cdot PP(t) - \Omega \cdot \tau_{total} \cdot Q_3 \cdot E \cdot \mathcal{XZ}(t) + \Omega \cdot \tau_{total} \cdot Q_3 \cdot r(t) \\
&\quad - \Omega \cdot \tau_{total} \cdot max\left(0, \lambda_1 + c \cdot E \cdot \mathcal{XZ}(t) - 100c\right) + \Omega \cdot \tau_{total} \cdot max\left(0, \lambda_2 - c \cdot E \cdot \mathcal{XZ}(t)\right)
\end{aligned}
$$

where $E = \begin{bmatrix} 0_{1\times N} & 1 \end{bmatrix}$, and $PP(T) = \begin{bmatrix} p_1(T) & W(T) \end{bmatrix}^T = \begin{bmatrix} 0 & 0_{N\times 1} \end{bmatrix}^T$ is the terminal condition.

To solve this problem numerically, UZAWA method, which was used earlier, is implemented. The algorithm executes the following steps.

1. Choose arbitrary values of $\lambda^{(0)} \in \mathbb{R}_+$ and $U(t)^{(0)}$

2. Calculate $\mathcal{XZ}(t)$ using (5.74) and $\mathcal{XZ}(0)$

3. Calculate $PP(t)$ using (5.75) and $PP(T)$

4. Update $U(t)$

$$U^{(k+1)}(t) = U^{(k)}(t) - v_1 \cdot \frac{\partial \mathcal{L}(\cdot)}{\partial U(t)}$$

where $\frac{\partial \mathcal{L}(\cdot)}{\partial U(t)}$ is obtained from $\mathcal{L}_U(\cdot)$ as $U(t)R_3 + B^T(t)p_1(t)$ and v_1 is a positive constant.

5. Saturate $U^{(k+1)}(t)$ by projecting it on the feasible set.

$$U^{(k+1)}(t) = \begin{cases} 0 & , U^{(k+1)}(t) < 0 \\ U^{(k+1)}(t) & , 0 \leq U^{(k+1)}(t) \leq 100 \\ 100 & , U^{(k+1)}(t) > 100 \end{cases}$$

6. For each $\lambda^{(k)}$, verify that $\|U(t)^{(k+1)} - U(t)^{(k)}\| \leq \epsilon$, where ϵ is a positive constant. If condition is true, go to step 7. Otherwise, go back to step 2.

7. Update λ

$$\lambda^{(k+1)} = \lambda^{(k)} + \upsilon_2 \cdot \frac{\partial \mathcal{L}(\cdot)}{\partial \lambda}$$
$$= \lambda^{(k)} + \frac{\upsilon_2}{c} \left(max(0_{2\times 1}, \lambda^{(k)} + cI(X(1,t))) - \lambda^{(k)} \right)$$

where υ_2 is a positive constant. This update law reduces to equation (5.55) when $\upsilon_2 = c$.

8. Verify that $\|\lambda^{(k+1)} - \lambda^{(k)}\| \leq \epsilon$. If condition is true, terminate. Otherwise, go back to step 3.

5.2.7 Simulation results

5.2.7.1 EGR linear model control

The control designed on the EGR linear model results in a closed-loop system whose input and output are shown in Figure 5.15, for the discretize-then-optimize approach, and in Figure 5.16, for the optimize-then-discretize approach. Both approaches share the same solver parameters, shown in Table 5.6.

It can be noticed from Figures 5.15 and 5.16 that starting from the same initial input $U^{(0)} = X_{avc} = 50\%$, both methods calculate the same optimal

λ_0	$\begin{bmatrix} 0.01 & 0.01 \end{bmatrix}^T$	c	0.1
ϵ	0.002	$L_{2,3}$	50
υ_1	0.01	$Q_{2,3}$	50
υ_2	0.01	$R_{2,3}$	0.01

Table 5.6: Parameters of the solver
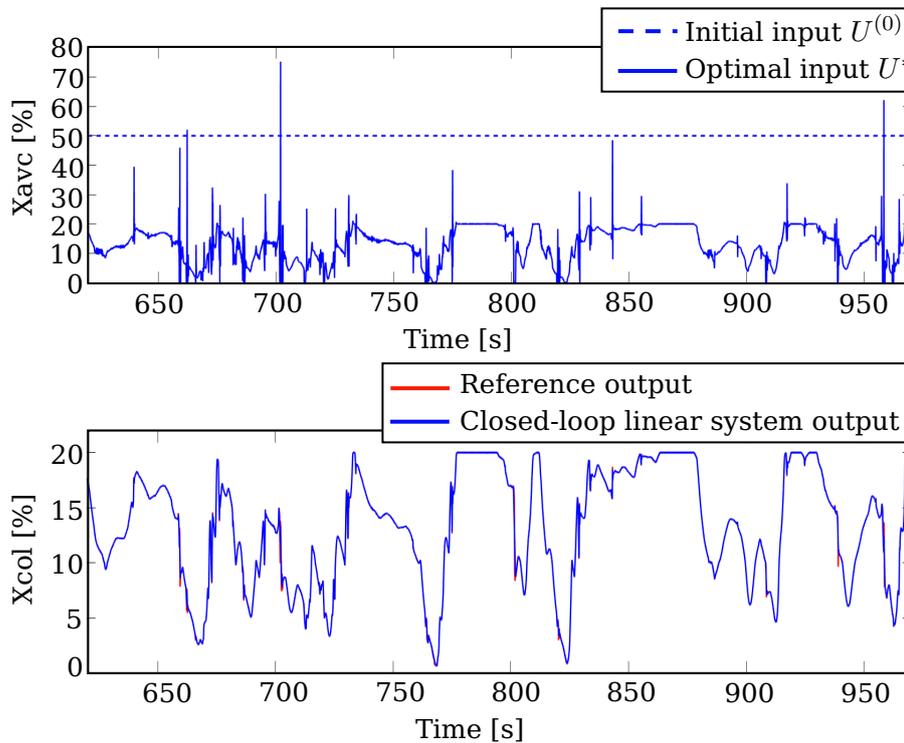
Figure 5.15: Discretize-then-optimize approach: closed-loop system input and output

input $U^* = X_{avc}^*$ and result in the same output X_{col}, which perfectly tracks the reference.

While some authors prefer the discretize-then-optimize approach because the gradient stems directly from the original cost function, others prefer the optimize-then-discretize approach, because it is more precise and less sensible to the choice of the numerical solver, as the discretization comes in after the calculation of the necessary conditions of optimality. The EGR control problem, however, doesn't show a preference for any of both approaches, in terms of merits as well as demerits. From the positive point of view, the discretization in both cases seems to retain the precision conveyed by the infinite-dimensionality of the control problem. From the negative point of view, both approaches are computationally expensive. Their iterative nature prolongs the time needed to calculate an optimal solution, which makes them not practical for an implementation on an ECU, unless subject to more efficient faster numerical solvers, such as those proposed in [HHP94] for large-scale optimization problems.
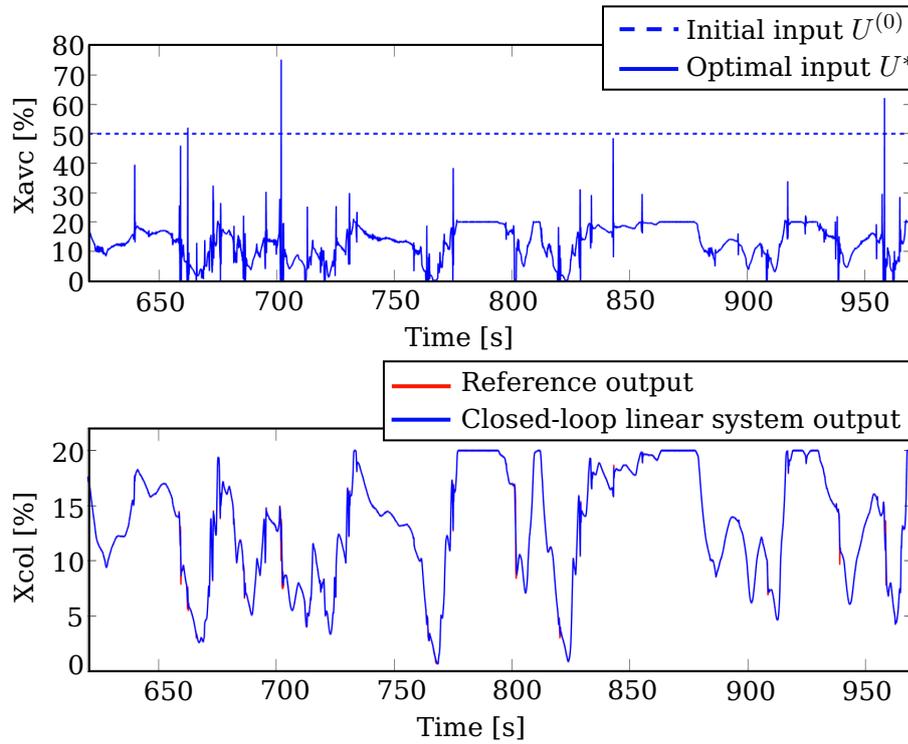
Figure 5.16: Optimize-then-discretize approach: closed-loop system input and output

5.2.7.2 AMESim® nonlinear model control

The control input in the AMESim® nonlinear model is the EGR duty cycle (%). Therefore, verifying the optimal control law on the nonlinear AMESim® model requires converting the optimal control input X^*_{avc} to EGR duty cycle (%). This conversion is a 4-step process:

- Convert X^*_{avc} to Q_{EGR} using the ODE:

$$\dot{X}_{avc} = \alpha_{avc} \left[- \left(Q_{EGR} + Q_{air} \right) X_{avc} + Q_{EGR} X_{em} \right] \tag{5.76}$$

where $\alpha_{avc} = \frac{RT_{avc}}{P_{avc}V_{avc}}$, T_{avc} (K), P_{avc} (Pa), and V_{avc} (m³) are the respective temperature, pressure, and volume upstream the compressor, and X_{em} is the BGR of the gas arriving from the exhaust manifold. This ODE formulation stems from the fact that the volume upstream the compressor is considered as a control volume.

S_{EGR} (mm^2)	0	25	125	210	250	250.01	250.015	250.02	250.025
Duty cycle (%)	0	10	20	30	40	50	60	70	100

Table 5.7: Look-up table matching S_{EGR} and EGR duty cycle

- Convert Q_{EGR} to S_{EGR} using the Saint-Venant equation, which calculates the mass flow rate as a function of the pressure drop across the EGR valve, as follows:

$$Q_{EGR} = \frac{S_{EGR}C_qC_mP_{up}}{\sqrt{T_{up}}} \tag{5.77}$$

where S_{EGR} (m^2) is the cross-sectional surface area of the EGR opening, C_q is the flow coefficient equal to 1, P_{up} (Pa) and T_{up} (K) are the respective pressure and temperature upstream the EGR valve, and C_m is the mass flow parameter expressed, by assuming a subsonic flow [Hey+88], as:

$$C_m = \sqrt{\frac{2\gamma_g}{R(\gamma_g-1)}\left(\left(\frac{P_{down}}{P_{up}}\right)^{\frac{2}{\gamma_g}} - \left(\frac{P_{down}}{P_{up}}\right)^{\frac{\gamma_g+1}{\gamma_g}}\right)}$$

where γ_g is the specific heat ratio of the gas, and P_{down} (Pa) is the pressure downstream the EGR valve.

- Saturate S_{EGR} between $S_{min} = 0$ and $S_{max} = 250.025$ mm^2 to make sure that the physical limits of the actuator are respected, and smooth it using a median filter to avoid the peaks and outliers.

- Convert S_{EGR} (mm^2) to EGR duty cycle (%) using look-up Table 5.7.

Processing the optimal input X_{avc}^* in this way to obtain the optimal EGR duty cycle, results in an eventually smoother input X_{avc}, shown in Figure 5.17 along with the system output, which well tracks the reference output. Note that the output falling between $t = 100$ s and $t = 150$ s corresponds to a stop phase in which the engine is not running. Therefore, tracking the reference output in this phase is senseless.

The nonlinear AMESim$^®$ model was initially controlled using a proportional integral (PI) controller, with proportional and integral gains equal to 1. The input of the controller is the error between the current X_{col} and its reference value, and the output of the controller is the EGR duty cycle, whose value is saturated between 0 and 100 to avoid exceeding the
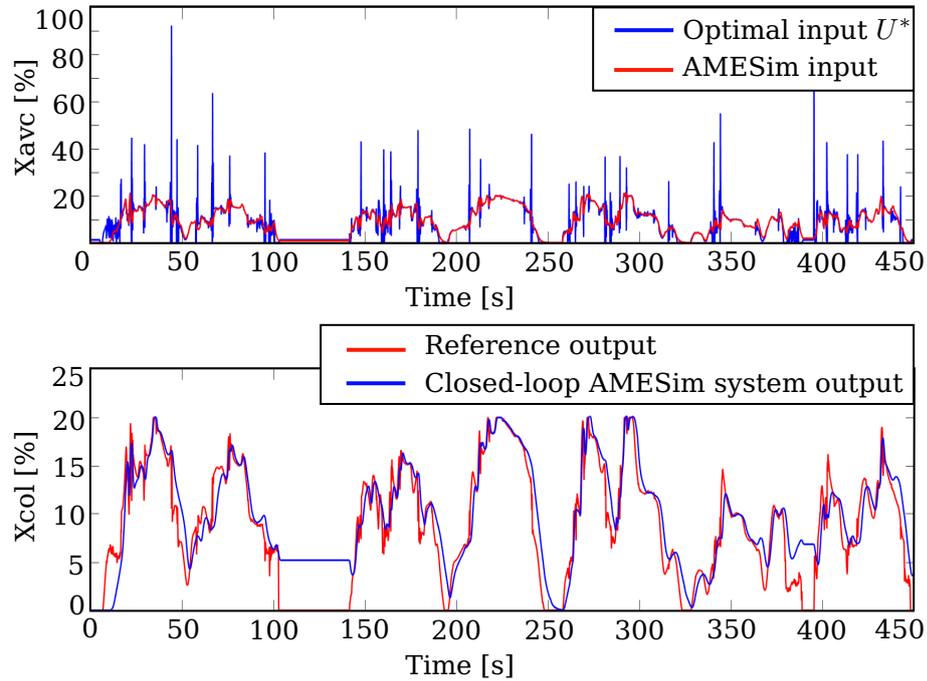
Figure 5.17: AMESim® model input and output

physical limits of the valve. Figure 5.18 helps comparing the optimal controllers proposed in this chapter with the PI controller, by showing the output X_{col}, the mass flow through the EGR Q_{EGR}, and the EGR duty cycle in both cases. The output plot shows that, while both controllers are able to track the reference output, the optimal controllers feature less overshoot and transient oscillations than the PI controller.

5.3 Conclusion and future perspectives

Heat and mass transport phenomena, taking place in the pipes of the engine air-path, such as the heat exchanger and the EGR system, can be modeled as first-order hyperbolic PDEs of infinite-dimensional nature. A space-averaging technique is proposed in this chapter to allow decoupling the hyperbolic PDEs representing advective flows in a heat exchanger and reformulating them as a time-delay system. The merits of this technique are illustrated on experimental data and replacing the original PDEs by the time-delay model for identification showed increased accuracy of the identified parameters and enhanced computa-
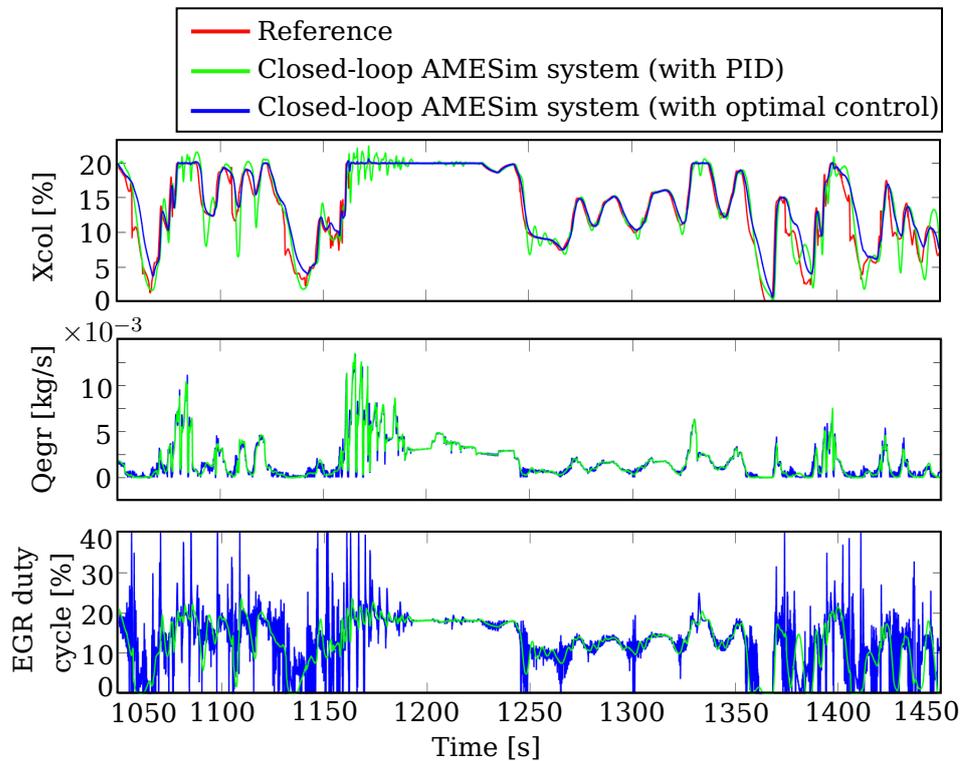
Figure 5.18: AMESim® model output and EGR mass flow rate and duty cycle

tional efficiency. Testing this computationally-efficient estimator on a car heat exchanger would be an interesting sequel of this work. From the EGR perspective, the BGR mass transport in the intake manifold is also described by a time-delay model, in this case exploited for control purposes. Two indirect optimal control approaches, taking into account the infinite-dimensional nature of the model, are compared: discretize-then-optimize and optimize-then-discretize. To account for the system's input and state constraints, the controllers are equipped with constraint management techniques such as the Augmented Lagrangian Uzawa method, and their merits are demonstrated on the original AMESim® model. To deal with the computationally expensive optimization process, a future scope of this research work lies in exploring more efficient simulation and numerical tools. Also, replacing the time-invariant delay by a time-variant delay would enhance the control performance by making it more realistic.

Conclusions and Perspectives

The control development lifecycle in modern automotive industry is model-based, which is why accurate reliable models of engine actuators are fundamental for successful engine control design. Traditional control techniques, such as the PID, have long been used in the automotive industry, and automotive engineers have well experimented them. They know their strengths and weaknesses, and they know for a fact that they are insufficient for addressing many of today's advanced control problems. Nevertheless, introducing to the automotive industry identification and control techniques never-before tested and evaluated on real engine models is a blunt hasty step, which puts the industry under the risk of losing lots of time and money. This is where this work and similar works come in to prepare the field for the introduction of such techniques, and to evaluate the applicability of novel methods to certain powertrain systems.

The research covered by this PhD thesis is two-fold as it addresses the questions of system identification as well as system control. The first topic of interest is linear black-box model identification. Linear black-box models are becoming increasingly important in the automotive industry for two reasons. First, concerning the black-box property, the ever-changing TDs of engines provoke an ever-lasting need for models of these engines. Building numerous sophisticated engine models is an exhausting task, which requires a huge effort and delays the control development process. Linearity of the model is another concern for the industry, because it is the cornerstone for integrating advanced linear control techniques in the design lifecycle. In this context, chapter 3 of this thesis evaluates the ability of subspace methods to identify a black-box linear SS model of the engine air-path, and chapter 4 proposes two black-box linear ARX models of the throttle, identified using the least-squares method. The second topic of interest is modeling engine air-path components featuring gas transport as time-delay models and controlling them. Gas transport and advective flows in the engine air-path inevitably induce time-delay. Automakers know that developing air-path control without accounting for this delay is less likely to yield an optimal control law. In this context, chapter 5 of this thesis proposes time-delay modeling and control schemes for the advective flows and mass trans-

port phenomena taking place in the engine air-path. The conclusions and future perspectives of each chapter are summarized hereafter.

In chapter 3, we identify black-box linear SS models of the H5Ft400 engine air-path at a set of operating points whose torque ranges from low to high, using subspace methods and prediction error minimization algorithms. Excellent models were identified in the low-torque zone, and good ones were identified in the high-torque zone. This chapter sheds light on the adjustable parameters of the solvers, and specifies the most successful combinations of parameter values that yield the best estimates. Some interesting extensions to this work would be:

- understanding and addressing the hysterisis phenomenon observed in the outputs of the AMESim® model upon exciting the throttle or the waste-gate in opening/closing sequences, which might be handy in improving the quality of the estimation;

- subjecting the identified SS systems to model reduction, resulting in equal model orders on all operating points of a certain ISO speed. By doing so, studying the level of similarity of the identified models becomes more straightforward, which opens the door to grouping the operating points that share similar models (attributing a single model for more than one operating point) and reduces the number of submodels for a particular ISO speed;

- replacing the iterative prediction error minimization algorithms by other optimization algorithms, preferably non-iterative, to accelerate the system identification process and to refine the initial estimate produced by the subspace methods in a more efficient way;

- evaluating the identification method proposed in this chapter when it comes to identifying the air-path dynamics of a diesel engine, whose main inputs would be the waste-gate and the high or low-pressure EGR and whose respective main outputs would be P_{boost} and the air flow rate in the intake path;

- exploiting the air-path model for open-loop and closed-loop control purposes to regulate the throttle and waste-gate valves and attain desired values of pressure in the air-path, and carrying on the control design process through the MIL, SIL, PIL, HIL, and calibration phases.

In chapter 4, we inspire from the mathematical model of the electric throttle to set the structure of a black-box linear regressor ARX model,

identify it using the least-squares method, and control it using an LQR controller. The identified model, which is verified with respect to a non-linear AMESim® model and a real throttle test-bench, is an LPV model which highlights the throttle's nonlinearities and discontinuities as it varies from one functioning zone to another. Below is a non exhaustive list of possible future work directions:

- improving the AMESim® model and test bench of the electric throttle to cover the torque range where $T_{low} = T_{spring}$. The validity of the identified ARX model can then be verified on the whole functioning range of the electric throttle;

- testing the identification and control methodology on other actuators, similar in physical structure to the electric throttle, such as the diesel engine air-intake valve;

- evaluating the linear control design of the identified SS and ARX models, firstly on the electric throttle test bench, and eventually on-board the vehicle after integrating all the steps of the control development lifecycle.

In chapter 5, we use the method of characteristics and spatial lumping to convert a system of first-order hyperbolic equations representing advective flows in a heat exchanger into a time-delay system, and we use the gradient-descent method to identify the parameters of the time-delay model. Exploiting this model to identify a heat exchanger test-bench shows being more efficient than the PDE system in terms of accuracy and computational efficiency. A future perspective to this work would be:

- understanding the nature of the advection exchange coefficient and how it changes with time. Using a time-varying instead of a constant advection exchange coefficient is more realistic and might increase the accuracy of the identified transport and heat exchange parameters;

- going a step forward toward real heat exchangers by considering the vertical flow of the liquid in the heat exchanger tube and dropping the assumption that advection is the only heat transfer taking place by considering other forms of heat transfer such as heat diffusion through the tubes of exchanger and that with the external environment;

- replacing the liquid water circulating in one of the heat exchanger tubes by an air-fuel mixture representative of the gas circulating in the pipes of an ICE, and replacing the liquid water circulating in the other tube by a liquid or gas coolant commonly used in ICE intercoolers. Although this significantly increases the complexity of the identification problem because it raises questions regarding gas mixing, homogeneity, and values of the gas constants such as density or specific heat as the gas evolves in the heat exchanger tube, but this long-term perspective is a natural extension to this work if the identification method is intended for engine air-path exchangers;

- considering the piping structure of a real intercooler which, regardless of its type, is certainly more complex than the single-tube structure considered in this work. Herein, the identified time-delay model only uses the input temperature at the tube boundary to predict its output temperature. Conserving this property with a complex piping structure involving elbows and other bent segments is not natural, and the intuitive approach that comes to mind is discretizing the structure into straight tube segments where the output temperature of a certain segment is considered the input temperature of the following one;

To tackle the control of time-delay systems, we adopt the EGR system example, which presents a constrained optimization problem. To solve this problem, we refer to indirect optimal control methods, the augmented Lagrangian method, and Uzawa algorithm. Very good tracking results were observed as a result of this method, which is worth extending in the following directions:

- controlling the SS model subject to time-varying delay in the output instead of constant delay, which would improve the output tracking without impacting the application of the indirect optimal control methods;

- testing alternative efficient numerical methods, other than the iterative time-consuming projected gradient-descent method, to accelerate the computation of the optimal law and increase its chance of real-time implementation on board an engine ECU.

- extending the optimal controller to the case of dual-EGR, by integrating into the LP-EGR time-delay system an HP-EGR model, with

or without time-delay as the distance traveled by the burned gases in case of HP-EGR is relatively short.

- assessing the performance of the optimal controller on a real LP-EGR system, by persuing all the phases of the V-cycle.

A global long-term perspective of the work done in this PhD is to obtain an assembly of credible linear models of all the systems composing the powertrain, accompanied with well-defined practical identification methods, and easily-calibrated controllers. The desirable characteristics of this assembly would be:

- its accessibility to rapid identification and simulation;

- its ability to reproduce the behavior of a real powertrain, just like the AMESim$^{\circledR}$ models used today;

- its adaptability and openness to welcome rapidly any change in the TD of the powertrain, without having to reconstruct the assembly from scratch.

The automotive industry, continuously pressured by manufacturing standards and marketing deadlines, has little time for testing new system identification and control techniques and updating what it has already been employing for years. The industry is in fact experimenting new techniques, such as artificial neural networks for system identification. Another example is MPC, which was lately welcomed by the automotive industry for implementation on real engines. Despite these promising experiences, however, the gap between theoretical progress and automotive applications remains wide. Many identification and control techniques that are very successful in theory remain unexplored in automotive applications, and are shelved because the industrialists do not have the time or the will to invest in them. The work done in this PhD is an investment in this direction, aiming at bridging this gap between real automotive applications, such as throttle and EGR control, and successful theories on system identification and control. In other words, inventing new control theories or new techniques for system identification is not in the scope of this work. It rather focuses on experimenting some already existing theories and techniques on real-life problems in the automotive industry, and evaluating their eligibility for implementation on an ECU. It addresses the automotive industrialists, and directs their attention to these identification and control methods, provides them with a preliminary assessment of their merits, demerits, and limitations, and enables

them to judge their worthiness to pass from the MIL phase to the SIL, PIL, and HIL phases of the control development lifecycle.

Further research on similar topics is necessary to open up the eyes of the industry on the control theories developed in the academic world, and to give these theories a chance of being integrated in future automotive technologies. It is our job as researchers and engineers not to let successful theories fade from the memories because they hadn't been put to good use. It is our job to bring them to the light. Theory shines in the light of industrial applications.

Bibliography

[ABMST94]   Maurizio Abate, B Ross Barmish, Carlos Murillo-Sánchez, and Roberto Tempo. "Application of some new tools to robust stability analysis of spark ignition engines: A case study". In: *IEEE Transactions on Control Systems Technology* 2.1 (1994), pp. 22–30.

[ADR95]   Alois Amstutz and Luigi R Del Re. "EGO sensor based robust output control of EGR in diesel engines". In: *IEEE Transactions on Control Systems Technology* 3.1 (1995), pp. 39–48.

[AFGG03]   Markus Ammann, Nicholas Patrick Fekete, Lino Guzzella, and Adolf Hermann Glattfelder. "Model-based control of the VGT and EGR in a turbocharged common-rail Diesel engine: theory and passenger car implementation". In: *SAE transactions* (2003), pp. 527–538.

[AK+17]   Bragadeshwarn Ashok, Ramesh Kumar, et al. "An integrated pedal follower and torque based approach for electronic throttle control in a motorcycle engine". In: *Engineering Journal* 21.1 (2017), pp. 63–80.

[ALMO12]   Paul Annus, Raul Land, Mart Min, and Jaan Ojarand. "Simple signals for system identification". In: *Fourier Transform-Signal Processing*. IntechOpen, 2012.

[ASGY04]   Sorour Alotaibi, Mihir Sen, Bill Goodwine, and K.T. Yang. "Controllability of Cross-Flow Heat Exchangers". In: *International Journal of Heat and Mass Transfer* 47 (2004), pp. 913–924.

[AT81]   Nasir Uddin Ahmed and Kok Lay Teo. *Optimal control of distributed parameter systems*. Elsevier Science Inc., 1981.

[BAB14]   Francois Bonne, Mazen Alamir, and Patrick Bonnay. "Nonlinear observer of the thermal loads applied to the helium bath of a cryogenic Joule-Thompson cycle". In: *Journal of Process Control* 24 (2014), pp. 73–80.

[Ber11]   Jeff Bergenthal. "Final Report Model Based Engineering (MBE) Subcommittee". In: *NDIA Systems Engineering Division-M&S Committee* (2011).

177

[Ber76]   Dimitri P Bertsekas. "On penalty and multiplier methods for constrained minimization". In: *SIAM Journal on Control and Optimization* 14.2 (1976), pp. 216–235.

[Bet10]   John T Betts. *Practical methods for optimal control and estimation using nonlinear programming*. Vol. 19. SIAM, 2010.

[BKO11]   Monika Bakosova, Michal Kacur, and Juraj Oravec. "Control of a Tubular Heat Exchanger". In: *18th International Conference on Process Control*. Tatranska Lomnica, Slovakia, 2011, pp. 338–343.

[BNS06]   Alessandro Beghi, Lorenzo Nardo, and Marco Stevanato. "Observer-based discrete-time sliding mode throttle control for drive-by-wire operation of a racing motorcycle engine". In: *IEEE Transactions on Control Systems Technology* 14.4 (2006), pp. 767–775.

[BPLCP12]   Delphine Bresch-Pietri, Thomas Leroy, Jonathan Chauvin, and Nicolas Petit. "Prediction-based trajectory tracking of external gas recirculation for turbocharged SI engines". In: *2012 American Control Conference (ACC)*. IEEE. 2012, pp. 5718–5724.

[BPP05]   Miroslav Barić, Ivan Petrović, and Nedjeljko Perić. "Neural network-based sliding mode control of electronic throttle". In: *Engineering Applications of Artificial Intelligence* 18.8 (2005), pp. 951–961.

[BSJTJ06]   Johan Bengtssonv, Petter Strandhv, Rolf Johansson, Per Tunestal, and Bengt Johansson. "Model predictive control of homogeneous charge compression ignition (HCCI) engine dynamics". In: *2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control*. IEEE. 2006, pp. 1675–1680.

[BV04]   Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[CB13]   Felipe Castillo Buenaventura. "Modélisation et Contrôle de la Boucle d'Air des Moteurs Diesel pour Euro 7". PhD thesis. Université de Grenoble, 2013.

[CCP06] Jonathan Chauvin, Gilles Corde, and Nicolas Petit. "Constrained motion planning for the airpath of a Diesel HCCI engine". In: *Proceedings of the 45th IEEE conference on decision and control*. IEEE. 2006, pp. 3589–3596.

[CFL28] Richard Courant, Kurt Friedrichs, and Hans Lewy. "On the partial difference equations of mathematical physics". In: *Mathematische Annalen* 100.1 (1928), pp. 32–74.

[CFN50] Jules G Charney, Ragnar Fjörtoft, and J von Neumann. "Numerical integration of the barotropic vorticity equation". In: *Tellus* 2.4 (1950), pp. 237–254.

[CK68] Kenneth L Cooke and David K Krumme. "Differential-difference equations and nonlinear initial-boundary value problems for linear hyperbolic partial differential equations". In: *Journal of Mathematical Analysis and Applications* 24.2 (1968), pp. 372–387.

[CTSF11] Matteo Corno, Mara Tanelli, Sergio M Savaresi, and Luca Fabbri. "Design and validation of a gain-scheduled controller for the electronic throttle body in ride-by-wire racing motorcycles". In: *IEEE Transactions on Control Systems Technology* 19.1 (2011), pp. 18–30.

[DPPJH04] Josko Deur, Danijel Pavkovic, Nedjeljko Peric, Martin Jansz, and Davor Hrovat. "An electronic throttle control strategy including compensation of friction and limp-home effects". In: *IEEE Transactions on Industry Applications* 40.3 (2004), pp. 821–834.

[DS16] Dragutin Debeljkovic and Goran Simeunovic. "Modelling, Simulation and Dynamic Analysis of the Time Delay Model of the Recuperative Heat Exchange". In: *Tehnika-Masinstivo* 65.3 (2016), pp. 407–413.

[EH14] Jamil El Hadef. "Approche quasi-systématique du contrôle de la chaîne d'air des moteurs suralimentés, basée sur la commande prédictive non linéaire explicite". PhD thesis. Université d'Orléans, 2014.

[Eva10] Lawrence C. Evans. *Partial Differential Equations*. Second edition. Vol. 19. Providence, Rhode Island: American Mathematical Society, 2010.

[FG83]   Michel Fortin and Roland Glowinski. *Augmented Lagrangian methods: applications to the numerical solution of boundary-value problems*. Amsterdam: North-Holland, 1983.

[FGS07]   Sanford Friedenthal, Regina Griego, and Mark Sampson. "INCOSE model based systems engineering (MBSE) initiative". In: *INCOSE 2007 Symposium*. Vol. 11. 2007.

[GB00]   Johan Gagner and Rickard Bondesson. "Adaptive Real-time Control of Nonlinear Throttle Unit". In: *MSc Theses* (2000).

[GMC09]   Olivier Grondin, Philippe Moulin, and Jonathan Chauvin. "Control of a turbocharged Diesel engine fitted with high pressure and low pressure exhaust gas recirculation systems". In: *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*. IEEE. 2009, pp. 6582–6589.

[Hes69]   Magnus R Hestenes. "Multiplier and gradient methods". In: *Journal of optimization theory and applications* 4.5 (1969), pp. 303–320.

[Hey+88]   John B Heywood et al. *Internal combustion engine fundamentals*. Mcgraw-hill New York, 1988.

[HHP94]   William W Hager, Donald W Hearn, and Panos M Pardalos. *Large scale optimization: state of the art*. Springer US, 1994.

[HPUU08]   Michael Hinze, René Pinnau, Michael Ulbrich, and Stefan Ulbrich. *Optimization with PDE constraints*. Vol. 23. Springer Science & Business Media, 2008.

[Ius99]   Alfredo N. Iusem. "Augmented Lagrangian methods and proximal point methods for convex optimization". In: *Investigación Operativa* 8.11-49 (1999), p. 7.

[Jan03]   Magnus Jansson. "Subspace identification and ARX modeling". In: *Proceedings of SYSID 2003* (2003).

[JG03]   Merten Jung and Keith Glover. "Control-oriented linear parameter-varying modelling of a turbocharged diesel engine". In: *Proceedings of 2003 IEEE Conference on Control Applications, 2003. CCA 2003*. Vol. 1. IEEE. 2003, pp. 155–160.

[JK98]   Mrdjan Jankovic and I Kolmanovsky. "Robust nonlinear controller for turbocharged diesel engines". In: *Proceedings of the 1998 American Control Conference. ACC (IEEE Cat. No. 98CH36207)*. Vol. 3. IEEE. 1998, pp. 1389–1394.

[JSK09]   Shugang Jiang, Michael H Smith, and James Kitchen. *Optimization of PID control for engine electronic throttle system using iterative feedback tuning*. Tech. rep. SAE Technical Paper, 2009.

[JZS14]   Xiaohong Jiao, Jiangyan Zhang, and Tielong Shen. "An adaptive servo control strategy for automotive electronic throttle and experimental validation". In: *IEEE Transactions on Industrial Electronics* 61.11 (2014), pp. 6275–6284.

[KHO17]   Thijs van Keulen, Lars Huijben, and Tom Oomen. "Identification of control-relevant diesel engine models using a local linear parametric approach". In: *IFAC-PapersOnLine* 50.1 (2017), pp. 7836–7841.

[Kir04]   Donald E Kirk. *Optimal control theory: an introduction*. Mineola, New York: Dover Publications, 2004.

[KK14]   Iasson Karafyllis and Miroslav Krstic. "On the relation of delay equations to first-order hyperbolic partial differential equations". In: *ESAIM:COCV* 20.3 (2014), pp. 894–923.

[KSHKS96]   Asaka Kitahara, Akiko Sato, Masatoshi Hoshino, I Kurihara, and Seichi Shin. "LQG based electronic throttle control with a two degree of freedom structure". In: *Proceedings of 35th IEEE Conference on Decision and Control*. Vol. 2. IEEE. 1996, pp. 1785–1788.

[Lar90]   Wallace E Larimore. "Canonical variate analysis in identification, filtering, and adaptive control". In: *29th IEEE Conference on Decision and Control*. IEEE. 1990, pp. 596–604.

[Li+12]   Ling Li et al. "Simulation on Electronic Throttle Control System Based on D2P". In: *Applied Mechanics and Materials*. Vol. 190. Trans Tech Publ. 2012, pp. 313–316.

[Lio71]   Jacques Louis Lions. *Optimal Control of Systems Governed by Partial Differential Equations (Grundlehren der Mathematischen Wissenschaften)*. First edition. Vol. 170. Springer-Verlag Berlin Heidelberg, 1971.

[Lju99]   Lennart Ljung. *System identification: theory for the user*. Second edition. Prentice-hall, 1999.

[LK98]   Michael Larsen and Petar Kokotovic. "Passivation design for a turbocharged diesel engine model". In: *Proceedings of the 37th IEEE Conference on Decision and Control (Cat. No. 98CH36171)*. Vol. 2. IEEE. 1998, pp. 1535–1540.

[LY08]   David G Luenberger and Yinyu Ye. *Linear and Nonlinear Programming*. Fourth edition. Springer, 2008.

[LY95]   Xunjing Li and Jiongmin Yong. *Optimal control theory for infinite dimensional systems*. Birkhäuser Basel, 1995.

[Mca33]   William H Mcadams. *Method of controlling recycling of exhaust gas in internal combustion engines*. US Patent 1,916,325. 1933.

[Mic+10]   Lawrence Michaels et al. *Model-based systems engineering and control system development via virtual hardware-in-the-loop simulation*. Tech. rep. SAE Technical Paper, 2010.

[MNT99]   Kaj Madsen, Hans Bruun Nielsen, and Ole Tingleff. *Methods for non-linear least squares problems*. Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 1999.

[MS12]   Javad Mohammadpour and Carsten W Scherer. *Control of linear parameter varying systems with applications*. Springer Science & Business Media, 2012.

[NRWS16]   David Novella-Rodriguez, Emmanuel Witrant, and Olivier Sename. "Control-Oriented Modeling of Fluid Networks: A Time-Delay Approach". In: *Recent Results on Nonlinear Delay Control Systems*. Springer, 2016, pp. 275–289.

[NSHO06]   Kazushi Nakano, Umerujan Sawut, Kohji Higuchi, and Yasusuke Okajima. "Modelling and observer-based sliding-mode control of electronic throttle systems". In: *ECTI Transactions on Electrical Eng., Electronics, and Communications* 4.1 (2006), pp. 22–28.

[ODR07]   Peter Ortner and Luigi Del Re. "Predictive control of a diesel engine air path". In: *IEEE transactions on control systems technology* 15.3 (2007), pp. 449–456.

[OHP01]   Urnit Ozguner, Sulgi Hong, and Yaodong Pan. "Discrete-time sliding mode control of electronic throttle valve". In: *Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No.01CH37228)*. Vol. 2. IEEE. 2001, pp. 1819–1824.

[PBGM62]   Lev Semenovich Pontryagin, Vladimir Grigorevich Boltyanski, Revaz Valerianovich Gamkrelidze, and Evgenij Frolovic Mishchenko. *The Mathematical Theory of Optimal Processes*. New York: Interscience Publishers, 1962.

[PCS13]   Giulio Panzani, Matteo Corno, and Sergio Matteo Savaresi. "On adaptive electronic throttle control for sport motorcycles". In: *Control Engineering Practice* 21.1 (2013), pp. 42–53.

[PD11]   Jacob L Pedersen and Stephen J Dodds. "A comparison of two robust control techniques for throttle valve control subject to nonlinear friction". In: *Advances in Computing and Technology, University of East London* (2011), pp. 91–99.

[PDJP06]   Danijel Pavković, Joško Deur, Martin Jansz, and Nedjeljko Perić. "Adaptive control of automotive electronic throttle". In: *Control Engineering Practice* 14.2 (2006), pp. 121–136.

[Pow69]   Michael JD Powell. "A method for nonlinear constraints in minimization problems". In: *Optimization* (1969), pp. 283–298.

[PS08]   Alexandros Plianos and Richard Stobart. *Modeling and control of diesel engines equipped with a two-stage turbo-system*. Tech. rep. SAE Technical paper, 2008.

[Rom84]   Fred E. Romie. "Transient Response of the Counter-Flow Heat Exchanger". In: *ASME J. Heat Transfer* 106.3 (1984), pp. 620–626.

[Ros71]   Dale W. Ross. "Controller design for time lag systems via a quadratic criterion". In: *IEEE Transactions on Automatic Control* 16.6 (1971), pp. 664–672.

[RW92]   Yimin Xuan Roetzel Wilfried. "Transient Response of Parallel and Counter-Flow Heat Exchangers". In: *J. Heat Transfer* 114.2 (1992), pp. 510–512.

[Sch70]   Manfred Schroeder. "Synthesis of low-peak-factor signals and binary sequences with low autocorrelation (Corresp.)" In: *IEEE Transactions on Information Theory* 16.1 (1970), pp. 85–89.

[Ser14]   Daniel Blasco Serrano. *Combustion Engine Identification and Control*. Department of Automatic Control, Lund University, 2014.

[SK03]   Osama Suleiman Khayal. "The History of the Internal Combustion Engine". In: *Annals of the Faculty of Engineering Hunedoara* (2003), pp. 145–148.

[SKF00]   Anna G Stefanopoulou, Ilya Kolmanovsky, and James S Freudenberg. "Control of variable geometry turbocharged diesel engines for reduced emissions". In: *IEEE transactions on control systems technology* 8.4 (2000), pp. 733–745.

[Str04]   John C Strikwerda. *Finite difference schemes and partial differential equations*. Second edition. Vol. 88. Philadelphia, USA: SIAM, 2004.

[SW77]   Andrew P Sage and Chelsea C White. *Optimum systems control*. Second edition. Englewood Cliffs, New Jersey: Prentice Hall, 1977.

[TAW12]   Roland Tóth, Hossam Seddik Abbas, and Herbert Werner. "On the state-space realization of LPV input-output models: Practical approaches". In: *IEEE transactions on control systems technology* 20.1 (2012), pp. 139–153.

[TE09]   Andreas Thomasson and Lars Eriksson. "Model-based throttle control using static compensators and IMC based PID-design". In: *IFAC workshop on engine and powertrain control, simulation and modeling*. 2009.

[TL60]   Herman Thal-Larsen. "Dynamics of heat exchangers and their models". In: *Journal of Basic Engineering* 82.2 (1960), pp. 489–500.

[Trö10]   Fredi Tröltzsch. *Optimal control of partial differential equations: theory, methods, and applications.* Vol. 112. American Mathematical Soc., 2010.

[UCKC00]   Vadim Utkin, Hao-Chi Chang, Ilya Kolmanovsky, and Jeffrey A Cook. "Sliding mode control for variable geometry turbocharged diesel engines". In: *Proceedings of the 2000 American Control Conference. ACC (IEEE Cat. No. 00CH36334).* Vol. 1. 6. IEEE. 2000, pp. 584–588.

[Uza58]   Hirofumi Uzawa. "Iterative methods for concave programming". In: *Studies in linear and nonlinear programming* 6 (1958), pp. 154–165.

[VBMPP06]   Mario Vašak, Mato Baotić, Manfred Morari, Ivan Petrović, and Nedjeljko Perić. "Constrained optimal control of an electronic throttle". In: *International Journal of Control* 79.05 (2006), pp. 465–478.

[Ver94]   Michel Verhaegen. "Identification of descriptor systems and periodic linear, time-varying systems". In: *IFAC Proceedings Volumes* 27.8 (1994), pp. 443–448.

[Vib95]   Mats Viberg. "Subspace-based methods for the identification of linear time-invariant systems". In: *Automatica* 31.12 (1995), pp. 1835–1851.

[VMP05]   Mario Vasak, Luka Mladenovic, and Nedjeljko Peric. "Clustering-based identification of a piecewise affine electronic throttle model". In: *31st Annual Conference of IEEE Industrial Electronics Society, 2005. IECON 2005.* IEEE. 2005, 6–pp.

[VNJP17]   Jouni Vuojolainen, Niko Nevaranta, Rafal Jastrzebski, and Olli Pyrhonen. "Comparison of Excitation Signals in Active Magnetic Bearing System Identification". In: *Modeling, Identification and Control* 38.1 (2017), pp. 1–11.

[VNMKSCW98] Michiel Van Nieuwstadt, Paul Moraal, Ilya Kolmanovsky, Anna Stefanopoulou, M Criddle, and P Wood. "A Comparison of SISO and MIMO Designs for EGR-VGT Control of a High-Speed Diesel Engine". In:

IFAC Workshop on Adavances in Automotive Control, Mohican State Park, Ohio, USA. 1998.

[VNMKSWC98]   Michiel Van Nieuwstadt, Paul Moraal, Ilya Kolmanovskyt, Anna Stefanopoulou, P Wood, and M Criddle. "Decentralized and multivariable designs for egr-vgt control of a diesel engine". In: *IFAC Proceedings Volumes* 31.1 (1998), pp. 189–194.

[VODM94]   Peter Van Overschee and Bart De Moor. "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems". In: *Automatica* 30.1 (1994), pp. 75–93.

[VODM95]   Peter Van Overschee and Bart De Moor. "A unifying theorem for three subspace system identification algorithms". In: *Automatica* 31.12 (1995), pp. 1853–1864.

[VV01]   Vincent Verdult and Michel Verhaegen. "Identification of multivariable LPV state space systems by local gradient search". In: *2001 European Control Conference (ECC)*. IEEE. 2001, pp. 3675–3680.

[Wan08]   Junmin Wang. "Hybrid robust air-path control for diesel engines operating conventional and low temperature combustion modes". In: *IEEE Transactions on Control Systems Technology* 16.6 (2008), pp. 1138–1151.

[WDR06]   Xiukun Wei and Luigi Del Re. "On persistent excitation for parameter estimation of quasi-LPV systems and its application in modeling of diesel engine torque". In: *IFAC Proceedings Volumes* 39.1 (2006), pp. 517–522.

[WDR07]   Xiukun Wei and Luigi Del Re. "Gain Scheduled H_∞ Control for Air Path Systems of Diesel Engines Using LPV Techniques". In: *IEEE transactions on control systems technology* 15.3 (2007), pp. 406–415.

[WGDW06]   Emmanuel Witrant, Didier Georges, and C Canudas De Wit. "Optimal Control Design for the Stabilization of Network Controlled Systems". In: *Proceedings of the 2006 American Control Conference*. IEEE. Minneapolis, Minnesota, USA, 2006, pp. 2777–2782.

[WN10]   Emmanuel Witrant and Silviu-Iulian Niculescu. "Modeling and Control of Large Convective Flows with Time-delays". In: *Mathematics in Engineering, Science and Aerospace* 1.2 (2010), pp. 191–205.

[WRL08]   Xiukun Wei, Luigi del Re, and Lihua Liu. "Air path identification of diesel engines by LPV techniques for gain scheduled control". In: *Mathematical and Computer Modelling of Dynamical Systems* 14.6 (2008), pp. 495–513.

[WSKB17]   Ashley P Wiese, Anna G Stefanopoulou, Amey Y Karnik, and Julia H Buckland. "Model Predictive Control for Low Pressure Exhaust Gas Recirculation with Scavenging". In: *2017 American Control Conference (ACC)*. IEEE. 2017, pp. 3638–3643.

[XMJB10]   Yang Xiaolong, Huang Ming, Liu Jingping, and Liu Biao. "A robust EGR control system for gasoline engine using PID". In: *2010 International Conference on Optoelectronics and Image Processing*. Vol. 1. IEEE. 2010, pp. 116–119.

[Yan04]   Chang Yang. *Model-based analysis and tuning of electronic throttle controllers*. Tech. rep. SAE Technical Paper, 2004.

[YLSCHJC09] Bilal Youssef, Guenael Le Solliec, Gilles Corde, Olivier Hayat, Christophe Jabeur, and Pierre Olivier Calendini. "Low pressure EGR control for a turbocharged diesel HCCI engine". In: *2009 IEEE Control Applications,(CCA) & Intelligent Control,(ISIC)*. IEEE. 2009, pp. 346–351.

[YW09]   Xiaofang Yuan and Yaonan Wang. "A novel electronic-throttle-valve controller based on approximate model method". In: *IEEE Transactions on Industrial Electronics* 56.3 (2009), pp. 883–890.

[YW11]   Fengjun Yan and Junmin Wang. "Control of dual loop EGR air-path systems for advanced combustion diesel engines by a singular perturbation methodology". In: *Proceedings of the 2011 American Control Conference*. IEEE. 2011, pp. 1561–1566.

[ZWB17]   Fairouz Zobiri, Emmanuel Witrant, and Francois Bonne. "PDE Observer Design for Counter-Current Heat Flows in a Heat-Exchanger". In: *IFAC-PapersOnline* 50.1 (2017), pp. 7127–7132.

[ZYZ15]   Shupeng Zhang, Jie J Yang, and Guoming G Zhu. "LPV Modeling and Mixed Constrained H_2/H_∞ Control of an Electronic Throttle". In: *IEEE/ASME Transactions on Mechatronics* 20.5 (2015), pp. 2120–2132.

Résumé — Pour faire face à la complexité croissante des systèmes automobiles et aux exigences de plus en plus strictes des gouvernements et des clients, la pratique de l'ingénierie dans l'industrie automobile se transforme de l'ingénierie basée sur le texte à l'ingénierie basée sur le modèle. Le cycle de développement d'un système de contrôle basé sur un modèle décompose les différentes étapes de la conception du système de contrôle, depuis l'établissement des exigences du système et la construction de son modèle jusqu'à la validation et la calibration de la conception sur le hardware prévu. Les travaux de cette thèse se déroulent au cours de la phase de conception du contrôleur, également appelée model-in-the-loop (MIL). Dans cette phase, les modèles orientés commande sont simulés en boucle fermée avec leurs conceptions de commande correspondantes. Après la phase MIL, les phases suivantes prennent en charge et complètent le processus de développement du système de contrôle, notamment en intégrant le software (SIL), le processeur (PIL) et le hardware dans la boucle (HIL).

Le groupe motopropulseur (GMP) d'une automobile est l'ensemble des composants qui transforment la puissance de son moteur en mouvement. Le composant du GMP qui fait l'objet d'une visite exclusive dans le cadre de cette thèse est le moteur, et le principal sujet d'intérêt est l'optimisation des systèmes du GMP, qui est d'une importance capitale pour les raisons suivantes. La réduction de la consommation de carburant, le respect des normes antipollution, les coûts de fabrication et d'autres défis industriels importants obligent Renault et d'autres acteurs de l'industrie automobile à développer des motorisations de plus en plus complexes. Sur ces GMPs, le nombre d'actionneurs, de briques technologiques et donc de combinaisons de commandes explose. Déterminer au début du cycle de développement, la définition technique (DT) du GMP en parfaite adéquation avec les normes et les besoins des différents pays devient extrêmement difficile avec les méthodologies actuelles, et évoque la nécessité d'optimiser les systèmes du GMP, qui nécessite deux éléments essentiels. Le premier est un contrôleur qui exécute des cycles de conduite sur le simulateur de GMP sur lequel les paramètres des contrôleurs et les cartographies de consignes peuvent être modifiés ; et le deuxieme est un simulateur de GMP pour accéder au plus grand nombre possible de données physiques tout en garantissant un temps de simulation acceptable pour les algorithmes d'optimisation.

De cette base émerge l'objectif général de cette thèse, qui est de simplifier la DT d'un GMP en développant des méthodologies et en testant des algorithmes d'optimisation pour identifier et calibrer les modèles de systèmes et les paramètres de contrôle. Voici les objectifs précis et les principaux axes de travail autour desquels s'articule le travail de cette thèse.

Le premier intérêt principal de Renault dans ce travail de recherche est de déterminer des procédures efficaces d'identification de modèles de boîtes noires linéaires pour les systèmes de GMPs. Aujourd'hui, les ingénieurs de Renault construisent les modèles de systèmes de GMPs sur des plateformes de modélisation graphique et de simulation, comme AMESim®. Ces plates-formes sont largement utilisées par les constructeurs de moteurs car elles permettent une modélisation simple et rapide de modèles non linéaires précis. Du point de vue de l'ingénierie de l'Alliance Renault-Nissan, de nombreuses activités d'ingénierie sont liées à ces modèles, c'est pourquoi ils sont disponibles pour presque tous les moteurs de l'Alliance. Ces modèles sont souvent associés à la commande proportionnelle-dérivée-intégrale (PID), qui reste la technique de commande la plus utilisée dans l'industrie automobile en raison de sa simplicité et de sa capacité à traiter un large éventail de problèmes de commande. Un moyen de faciliter le réglage du contrôleur PID, ou même de le remplacer par un contrôleur linéaire, pourrait être la linéarisation des modèles autour de chaque point de fonctionnement. Cependant, cette solution n'est pas simple en raison de la complexité de la dynamique conduisant à des goulots d'étranglement numériques et informatiques. Renault s'intéresse non seulement à la propriété linéaire des modèles identifiés, mais aussi à leur aspect boîte-noire. C'est parce que l'évolution constante des moteurs provoque un besoin constant de modèles de moteurs, ce qui nécessite d'énormes efforts d'ingénierie et retarde le processus de développement des commandes. De plus, l'environnement de co-simulation entre AMESim® et MATLAB-Simulink® empêche tout solveur d'avoir accès aux états du modèle AMESim® qui sont difficiles à communiquer via une s-fonction. Cet objectif précis est abordé dans les chapitres 3 et 4 de la thèse, qui traitent de l'identification d'un modèle linéaire boîte-noire d'un trajet d'air d'un moteur à essence par des méthodes subspace (chapitre 3), et d'un papillon électrique par la méthode des moindres carrés (chapitre 4).

Au-delà de l'identification et de la calibration de modèles linéaires de systèmes de GMPs, Renault s'intéresse également à la problématique

des retards dans ces GMPs. Les retards sont très répandus dans les GMPs, et les ingénieurs de la régulation sont confrontés à ces retards dans de nombreuses applications de commande du GMP, telles que la commande de la recirculation des gaz d'échappement (EGR) et la régulation du rapport air/carburant. L'intégration des retards dans les modèles de GMPs améliore la conception de la commande du GMP, ce qui se traduit par de meilleures performances du GMP et donc par des véhicules plus compétitifs. Cet objectif précis est abordé dans le chapitre 5 de la thèse, qui traite de l'identification et du contrôle des modèles de transport dans la chaine d'air du moteur. Cette partie sera divisée en deux sous-parties, la première qui est axée sur l'identification et se concentre sur le transport d'énergie (chaleur) dans un échangeur de chaleur, et la seconde qui est axée sur le contrôle et se concentre sur le transport de masse dans un système EGR.

En plus du chapitre d'introduction, la thèse comprend quatre autres chapitres, dont le contenu est résumé dans la suite.

Le chapitre 3, intitulé "Identification des Modèles Boîte-Noire pour des Moteurs à Essence à l'Aide des Méthodes Subspace", définit une méthodologie qui permet l'identification automatique dans le domaine temporel d'un modèle linéaire à boîte noire type state-space (SS) d'un moteur à essence à partir d'un modèle à valeur moyenne non linéaire (MVEM) construit sur AMESim®. Le modèle linéaire SS déclare les pourcentages d'ouverture des vannes papillon et wastegate comme entrées, et la pression d'admission et la pression de suralimentation comme sorties en plus de la pression en amont de la turbine et de la vitesse du turbo. La méthodologie consiste à exciter le système MVEM en utilisant des signaux d'entrée de type multisines en raison des avantages qu'ils offrent en termes de faible facteur de crête, de gamme de fréquence personnalisable et de faible corrélation linéaire entre les deux signaux d'entrée. En utilisant l'ensemble d'entrées et de sorties obtenu, des méthodes subspace sont utilisées pour identifier une estimation initiale du modèle SS à chaque point de fonctionnement caractérisé par le régime moteur et la pression effective moyenne de frein, qui varie en proportion directe avec le couple d'une zone à faible couple vers une zone à couple élevé. Ces méthodes sont renforcées par des algorithmes de minimisation des erreurs de prédiction afin d'améliorer la qualité de l'estimation et de minimiser davantage l'erreur entre la sortie de l'estimation linéaire et celle du modèle non linéaire AMESim®. Ce chapitre conclut sur l'ordre des modèles et les paramètres d'algorithme les plus pratiques pour l'application de ces méthodes d'identification sur

les moteurs à essence, ce qui réduit le temps nécessaire pour tester les options du solveur de 81% à. Il conclut également sur l'efficacité de la technique d'identification pour identifier le moteur à essence aux points de fonctionnement à faible couple, et sur l'adéquation des multisines comme entrée d'excitation dans cette zone de couple. D'autre part, l'efficacité limitée de la méthodologie aux points de fonctionnement à couple élevé met en évidence la difficulté de dissocier les effets du papillon et du wastegate sur les sorties dans cette zone de couple. C'est-à-dire, contrairement au cas où la pression dans le collecteur d'admission est inférieure à 1 bar où le papillon est la seule vanne en charge du débit d'air, ici le papillon et le wastegate sont tous deux actifs et affectent simultanément les quatre sorties, ce qui rend le processus d'excitation plus complexe. Enfin, sachant que chaque constructeur automobile propose une large gamme de moteurs et que les DT de chacun de ces moteurs sont en constante évolution, l'avantage fondamental de la méthodologie proposée est qu'elle permet de passer d'une DT à une autre sans avoir à définir une nouvelle technique d'identification linéaire à chaque changement.

Le chapitre 4, intitulé "Identification et Commande du Modèle Linéaire du Papillon Electrique", présente une méthodologie pour identifier des modèles linéaires à paramètres variants (LPV) type SS et auto-régressifs avec entrée exogène (ARX) d'un papillon électrique à partir d'un modèle non linéaire du dernier. Le modèle non linéaire, dont les paramètres de variation sont l'angle du papillon et sa vitesse, est d'abord écrit sous la forme d'un système LPV SS en utilisant les équations physiques qui régissent sa dynamique. Ce modèle met en évidence les discontinuités et les non-linéarités du papillon et donne un aperçu des éléments à utiliser dans le vecteur de régression, la deuxième étape étant d'identifier le système par régression linéaire en minimisant un critère des moindres carrés. Pour contrôler les modèles linéaires, un régulateur quadratique linéaire (LQR) est conçu séparément pour chacun des systèmes LPV identifiés, équipé d'un feedforward de perturbation et de référence pour faire face aux perturbations d'entrée existant dans les modèles LPV du papillon et au problème du suivi de référence respectivement. Le régulateur est évalué sur le modèle non linéaire d'où le modèle linéaire a été extrait. Cette approche se révèle intéressante fondamentalement grâce au modèle ARX, qui permet de contourner les modèles mathématiques SS dans les futures identifications de modèles linéaires de papillons électriques, et dont la crédibilité est vérifiée par rapport à un autre modèle non linéaire et par rapport à un véritable banc de test de papillon électrique. Aussi, ce chapitre montre que le nombre idéal de sous-

modèles LPV représentatifs du papillon électrique dépend des choix de modélisation et du niveau de précision requis. Enfin, le fait d'utiliser les équations physiques que nous connaissons du papillon pour définir la structure du modèle boîte-noire implique de combiner les mérites de la modélisation mathématique et de la modélisation boîte-noire en fournissant un modèle qui implique les lois physiques régissant le papillon électrique, mais qui capture ses aspects obscurs et ne nécessite que des entrées et des sorties du système pour l'identifier.

Le chapitre 5, intitulé "Identification et Contrôle des Modèles de Transport dans la Chaine d'Air du Moteur", traite des phénomènes de transport dans la chaine d'air du moteur. Il aborde deux de ces phénomènes de transport, le transport d'énergie et le transport de masse, en présentant leurs modèles orientés commande et la conception de la commande du modèle de transport de masse. Pour démontrer les phénomènes de transport de chaleur et de masse sur les composants de la chaine d'air, l'échangeur de chaleur et l'EGR sont choisis respectivement. Ce chapitre est divisé en deux parties:

- La première partie présente l'identification des flux advectifs présents dans l'échangeur de chaleur. Les phénomènes d'échange thermique qui se produisent dans les tuyaux d'un échangeur de chaleur peuvent être modélisés sous forme d'équations aux dérivées partielles hyperboliques du premier ordre (EDPs). La reformulation de ces équations sous la forme d'un système à retard préserve la propriété infinie du système, tout en diminuant sa complexité mathématique. Ce chapitre propose, à l'aide d'une technique de moyennage spatiale et de la méthode des caractéristiques, une modélisation des températures du flux dans la chaine d'air sous forme d'un système à retard. Une méthode d'optimisation par descente de gradient est proposée pour estimer les paramètres de ce système à retard, en utilisant des mesures de température sur les bords d'une section de la chaine d'air. L'intérêt de cette approche est mis en évidence par les données expérimentales obtenues sur un banc d'essai d'échangeur de chaleur. En comparant le modèle EDP et le modèle à retard lorsqu'il est exploité à cette fin d'identification, le modèle à retard est aussi précis que le modèle EDP, mais beaucoup plus efficace sur le plan informatique avec un gain de temps égal à 98%.

- La deuxième partie présente le modèle orienté commande et la conception de commande du phénomène de transport du rapport

des gaz brûlés (BGR), qui se produit dans le trajet d'admission d'un moteur, en raison de la réorientation des gaz brûlés vers le trajet d'admission par l'EGR basse pression. Inspiré du modèle à retard présenté dans la première partie du chapitre, et basé sur un modèle AMESim® non linéaire du moteur, le BGR du collecteur d'admission est modélisé comme un modèle SS à retard de sortie ou comme un système couplé EDO-EDP qui prend en compte le retard entre le moment auquel les gaz brulés quittent le collecteur d'échappement et celui auquel ils sont réadmis dans le collecteur d'admission. En plus de leur retard de transport de masse, les BGRs sur le trajet d'admission sont également soumis à des contraintes d'inégalité d'état et d'entrée. L'objectif du problème de contrôle est de suivre un profil de sortie de référence du BGR dans le collecteur d'admission, en tenant compte du retard de transport et des contraintes d'état (sortie) et d'entrée du système. Dans ce but, deux approches de contrôle optimal indirecte sont mises en oeuvre et comparées, l'approche discretize-then-optimize et l'approche optimize-then-discretize. Pour tenir compte des contraintes d'inégalité d'état, les deux méthodes sont équipées de techniques d'optimisation sous contraintes telles que les méthodes Lagrangienne augmentée et UZAWA. Les conditions d'optimalité nécessaires sont formulées, dans chacun des deux cas, et les équations résultantes sont résolues numériquement à l'aide de la méthode de la descente du gradient projeté, qui assure la non-violation des contraintes d'inégalité sur l'entrée. La nouveauté du travail effectué dans ce chapitre réside dans la prise en compte des contraintes du système et de la dimensionnalité infinie du phénomène de transport de masse qui le régit. Les mérites du modèle à retard et de la conception de contrôle basée sur le modèle sont illustrés sur le modèle nonlinéaire AMESim® sur lequel le modèle mathématique est basé.

Enfin, le dernier chapitre, intitulé "Conclusions et Perspectives d'Avenir", résume les travaux réalisés dans le cadre de cette thèse, indique les orientations des travaux futurs possibles et réaffirme l'importance de ces travaux pour combler l'écart entre les applications automobiles réelles et les théories réussies sur l'identification et le contrôle des systèmes.

Summary — To cope with the ever-increasing complexity of the automotive systems and the more and more stringent demands of governments and customers, engineering practice in the automotive industry is transforming from text-based engineering to model-based engineering. The model-based control system development lifecycle breaks down the different stages of the control system design, from setting the requirements of the system and building its model until validating and calibrating the design on the intended hardware. The work in this thesis takes place in the controller design phase, which is also called the model-in-the-loop (MIL) phase. In this phase, control-oriented plant models are simulated in closed-loop with their corresponding control designs. Following the MIL phase, the subsequent phases take over and complete the control system development process, notably by integrating the software (SIL), the processor (PIL), and the hardware in the loop (HIL).

The powertrain of an automobile is the set of components which transform the power of its engine into motion. The component of the powertrain that is exclusively visited in this thesis is the engine, and the main topic of interest is powertrain systems optimization, which is of core significance due to the following reasons. Reduced fuel consumption, compliance with pollution control standards, manufacturing costs, and other strong industrial challenges are forcing Renault and other actors in the automotive industry to develop increasingly complex powertrains. On these powertrains, the number of actuators, technological bricks and therefore control combinations explode. Determining at the beginning of the development cycle, the right powertrain technical definition (TD) which is perfectly in line with the standards and needs of the different countries becomes extremely difficult with current methodologies, and evokes the need for powertrain systems optimization, which prerequires two essential elements. The first is a controller which runs driving cycles on the powertrain simulator on which the controllers' parameters and the setpoint maps can be modified; and the second is a powertrain simulator to access as much as possible of the physics while ensuring acceptable simulation time for optimization algorithms. From this basis emerges the general objective of this thesis, which is to simplify the TD of a powertrain by developing methodologies and testing optimization algorithms to identify and calibrate system models and controller parameters. In what follows are the precise objectives and the principal work axes around which the work in this thesis revolves.

The first main interest for Renault in this research work is to determine

efficient linear black-box model identification procedures for powertrain systems. Today, Renault engineers build the powertrain systems models on graphical modeling and simulation platforms, such as AMESim®. These platforms are widely used by engine manufacturers because they allow easy and fast modeling of accurate nonlinear models. From the Renault-Nissan Alliance engineering perspective, many of the engineering activities are attached to these models, which is why they are available for almost all the Alliance engines. These models are often associated with proportional-integral-derivative (PID) control design, which is still the most widely used control technique in the automotive industry due to its simplicity and ability to handle a broad range of control problems. A way to facilitate the tuning of the PID controller, or even to replace it by a linear controller, could be the linearization of the models around each operating point. However, this solution is not straightforward due to the complexity of the dynamics leading to numerical and computational bottlenecks. Renault is not only interested in the linear property of the identified models, but in their black-box aspect as well. This is because the ever-changing TDs of engines provoke an everlasting need for engine models which requires huge engineering effort and delays the control development process. Also, the co-simulation environment between AMESim® and MATLAB-Simulink® deprives any solver from having access to the states of the AMESim® model which are hard to communicate through a Simulink® s-function. This precise objective is addressed in chapters 3 and 4 of the thesis, which discuss linear black-box model identification of a gasoline engine air-path using subspace methods (chapter 3), and of an electric throttle using least-squares method (chapter 4).

In addition to identifying and calibrating linear models of powertrain systems, another main interest for Renault in this work is tackling the problem of time-delays in powertrain systems. Time delays are widespread in powertrains, and control engineers confront these delays in many powertrain control applications, such as exhaust gas recirculation (EGR) control, idle speed control (ISC), and air-to-fuel ratio control. Integrating time delays in powertrain models improves powertrain control design, which leads to better powertrain performance, and therefore more competitive vehicles. This precise objective is addressed in chapter 5 of the thesis, which discusses identification and control of transport models in the engine air-path. This part will be split in two sub-parts, the first which is identification-oriented and focusing on energy(heat) transport in a heat-exchanger, and the second which is control-oriented and focusing on mass transport in an EGR system.

In addition to the introductory chapter, the thesis comprises four other chapters, whose content is summarized in the sequel.

Chapter 3, entitled as "Black-box Gasoline Engine Identification using Subspace Methods", defines a methodology that permits automatic time-domain identification of a linear black-box state-space (SS) model of a gasoline engine starting from a nonlinear mean-value engine model (MVEM) built on AMESim®. The linear SS model declares the percentage openings of the throttle and waste-gate valves as inputs, and the intake manifold pressure and the boost pressure as outputs in addition to the pressure upstream the turbine and the turbo speed. The methodology consists of exciting the MVEM system using input signals of multisines type due to the merits they offer in terms of low crest factor, customizable frequency range, and low linear correlation between the two input signals. Using the resulting set of inputs and outputs, subspace methods are used to identify an initial estimate of the SS model at each operating point characterized by the engine speed and the break mean effective pressure, which varies in direct proportionality with the torque from a low-torque zone to a high-torque zone. These methods are reinforced by prediction error minimization algorithms in order to improve the quality of the estimation and further minimize the error between the output of the linear estimate and that of the nonlinear AMESim® model. This chapter concludes on the model order and the algorithm parameters most convenient for the application of these identification methods on gasoline engines, which decreases the time needed to test the solver options by 81%. It also concludes on the efficiency of the identification technique in identifying the gasoline engine at operating points of low torque, and the adequacy of the multi-sines as an excitation input in this torque zone. On the other hand, the limited efficiency of the methodology at operating points of high torque highlights the difficulty of dissociating the effects of the throttle and the waste-gate on the outputs in this torque zone. That is to say, contrary to the case when the pressure in the intake manifold is less than 1 bar where the throttle is the only valve in charge of the air flow, here the throttle and waste-gate are both active and are simultaneously affecting the four outputs, which renders the excitation process more complicated. Finally, knowing that each automobile manufacturer proposes a broad series of engines and that the TDs of each of these engines is ever-changing, the basic advantage of the proposed methodology is that it allows switching from one TD to another without having to define a new linear system identification technique with every switch.

Chapter 4, entitled as "Electric Throttle Linear Model Identification and Control", presents a methodology for identifying SS and auto-regressive with exogenous input (ARX) linear parameter-varying (LPV) models of an electric throttle starting from a nonlinear model of the latter. The nonlinear model, whose scheduling parameters are the angle of the throttle and its speed, is first written in the form of an LPV SS system using the physical equations governing its dynamics. This model highlights the throttle's discontinuities and nonlinearities, and gives insight about the elements to be used in the regression vector, whereby the second step is to identify the system using linear regression by minimizing a least-squares criterion. To control the linear models, a linear quadratic regulator (LQR) controller is designed separately for each one of the identified LPV systems, equipped with disturbance and reference feedforwards to cope with the input disturbance existing in the LPV models of the throttle and with the reference tracking problem respectively. The regulator is evaluated on the nonlinear model from which the linear model was extracted. This approach reveals interesting basically due to the ARX model, which allows bypassing the SS mathematical models in future electric throttle linear model identifications, and whose credibility is verified with respect to another nonlinear model and with respect to a real electric throttle test-bench. Also, this chapter shows that the ideal number of LPV submodels representative of the electric throttle depends on the modeling choices and the level of precision required. Finally, the fact that we use the physical equations that we know about the throttle to set the structure of the black-box model implies combining the merits of mathematical modeling and black-box modeling by providing a model which implies the physical laws governing the electric throttle, yet captures its obscure aspects and only needs the system's inputs and outputs to be identified.

Chapter 5, entitled as "Identification and Control of Transport Models in the Engine Air-path", addresses transport phenomena in the engine air-path. It touches on two of these transport phenomena, energy transport and mass transport, by presenting their control-oriented models and the control design of the mass transport model. To demonstrate the heat and mass transport phenomena on air-path components, the heat exchanger and the EGR are chosen respectively. This chapter is divided in two parts:

- The first part presents the identification of advective flows present in the heat exchanger. Heat exchange phenomena occurring in the pipes of a heat exchanger can be modeled as first-order hyperbolic

partial differential equations (PDEs). Reformulating these equations as a time-delay system preserves the infinite-dimensional property of the system, yet decreases its mathematical complexity. Using a space-averaging technique and the method of characteristics, this chapter proposes a time-delay system modeling of the flow temperatures in the air-path. A gradient-descent optimization method is proposed to estimate the parameters of this time-delay system, using boundary measurements of temperature in an air-path section. The interest of this approach is emphasized with experimental data obtained from a heat exchanger test-bench. By comparing the PDE model and the time-delay model when exploited for this identification purpose, the time-delay model shows being as accurate as the PDE model, yet much more computationally efficient with a time-save equal to 98%.

- The second part presents the control-oriented model and control design of the burned gas ratio (BGR) transport phenomenon, witnessed in the intake path of an engine, due to the redirection of burned gases to the intake path by the low-pressure EGR. Inspired by the time-delay model presented in the first part of the chapter, and based on a nonlinear AMESim® model of the engine, the BGR in the intake manifold is modeled as a SS output time-delay model, or alternatively as an ODE-PDE coupled system, that take into account the time delay between the moment at which the combusted gases leave the exhaust manifold and that at which they are readmitted in the intake manifold. In addition to their mass transport delay, the BGRs in the intake path are also subject to state and input inequality constraints. The objective of the control problem is to track a reference output profile of the BGR in the intake manifold, taking into account the transport delay and the state (output) and input constraints of the system. In this aim, two indirect optimal control approaches are implemented and compared, the discretize-then-optimize approach and the optimize-then-discretize approach. To account for the state inequality constraints, both methods are equipped with techniques for constrained optimization such as the augmented Lagrangian and the UZAWA methods. The necessary conditions of optimality are formulated, in each of both cases, and the resulting equations are solved numerically using the projected gradient-descent method, which ensures the non-violation of the input inequality constraints. The novelty of the work done in this chapter lies in considering the system's constraints and the infinite-dimensionality of the mass transport phe-

nomenon governing it. The merits of the time-delay model and the model-based control design are illustrated on the nonlinear AMESim® model on which the mathematical model is based.

Finally, the last chapter entitled as "Conclusions and Future Perspectives", summarizes the work done in this thesis, points out directions of possible future work, and restates the importance of this work in bridging the gap between real automotive applications and successful theories on system identification and control.

Résumé — L'optimisation des systèmes de groupes motopropulseurs (GMP) dans les automobiles modernes repose sur l'ingénierie des systèmes basée sur des modèles pour faire face à la complexité croissante des systèmes automobiles et aux exigences de conception des commandes. Deux prérequis à l'optimisation du GMP basée sur le modèle sont le simulateur de GMP et la conception des commandes, qui assurent un fonctionnement satisfaisant du GMP pendant les cycles de conduite. Cette thèse s'articule autour de ces deux prérequis et fait donc partie de la phase de model-in-the-loop du cycle de développement du contrôle-commande. Elle vise d'abord à identifier les modèles de systèmes de GMP orientés commande, en particulier les modèles linéaires de boîtes noires en raison des avantages qu'ils présentent en termes d'accessibilité à la conception de la commande linéaire et de facilité d'intégration des modifications dans la définition technique du système de GMP. Elle vise également à identifier et à contrôler les systèmes de GMP à retard de transport car l'intégration du retard dans le modèle et la conception des commandes est cruciale pour la précision du modèle et l'optimalité de son contrôle. En se basant sur ces prémisses, nous abordons le GMP du point de vue de la chaîne d'air du moteur. Nous identifions d'abord un modèle linéaire de boîte noire type state-space (SS) de la chaîne d'air d'un moteur essence, en utilisant un algorithme d'identification basé sur les méthodes subspace. Différents ordres de modèles et paramètres d'algorithmes sont testés et ceux qui donnent les meilleurs résultats d'identification et de validation sont mis en évidence, ce qui conduit à un gain de temps de 85% pour des identifications futures similaires. Bien que cette partie considère la chaîne d'air dans son ensemble, le reste du travail se concentre sur certains de ses composants, notamment le papillon électrique, l'échangeur de chaleur et la recirculation des gaz d'échappement (EGR). En ce qui concerne le papillon électrique, nous nous inspirons des lois physiques régissant le fonctionnement de ce dernier pour construire un modèle mathématique linéaire à paramètres variables (LPV) type SS, qui sert à définir la structure vectorielle de régression du modèle LPV boîte noire type ARX, qui est représentatif d'un banc d'essais du papillon électrique et reflète ses nonlinéarités et discontinuités en variant entres ses zones fonctionnelles. Pour traiter les questions des délais de transport de chaleur et de transport de masse dans la chaîne d'air du moteur, nous nous référons respectivement à l'échangeur de chaleur et à l'EGR. La refonte des équations aux dérivées partielles (EDP) hyperboliques

202

à dimensions infinies décrivant ces phénomènes de transport sous la forme d'un système à retard facilite l'identification du système adjoint et la conception du contrôle. Pour ce faire, une technique de moyennage spatiale et la méthode des caractéristiques sont utilisées pour découpler les EDP hyperboliques décrivant les flux advectifs dans un échangeur de chaleur et pour les reformuler en un système à retard temporel. Réduire l'erreur entre la température de sortie du modèle et celle d'un banc d'essai d'échangeur de chaleur, c'est ce que recherche la méthode à descente de gradient utilisée pour identifier les paramètres du système à retard, qui dépasse les EDP en termes de précision d'identification et de temps de calcul. D'autre part, l'EGR est abordé sous l'angle du contrôle et les EDP décrivant le phénomène du transport de masse dans sa structure tubulaire sont refondus en un système SS soumis à un retard de sortie. Pour réguler le rapport de gaz brûlé dans le gaz d'admission, la quantité de gaz recirculé est contrôlée à l'aide de deux approches de contrôle optimale indirecte qui prennent en compte la nature infinie du modèle. Ces méthodes sont accompagnées de la méthode du Lagrangien Augmenté et la méthode d'Uzawa pour garantir le respect des contraintes d'entrée et d'état, permettant ainsi d'obtenir un contrôleur plus performant que le PID initialement existant. D'une manière générale, cette thèse se situe à mi-chemin entre le secteur académique et le secteur industriel. En évaluant l'éligibilité de l'intégration des théories existants d'identification et de contrôle des systèmes dans des applications automobiles réelles, elle met en évidence les avantages et les inconvénients de ces théories et ouvre de nouvelles perspectives dans le domaine de l'optimisation des systèmes du GMP basée sur les modèles.

Mots clés: Ingénierie système basée sur un modèle, optimisation du groupe motopropulseur, chaîne d'air du moteur, identification du système, systèmes à retard, contrôle optimal

Abstract — Powertrain systems optimization in modern automobiles relies on model-based systems engineering to cope with the increasingly complex automotive systems and challenging control design requirements. Two prerequisites for model-based powertrain optimization are the powertrain simulator and the control design, which ensures a desirable powertrain operation during driving cycles. This thesis revolves around these two prerequisites, and therefore belongs to the model-in-the-loop phase of the control development lifecycle. It first aims at identifying control-oriented powertrain systems models, particularly linear black-box models because of the merits they present in terms of accessibility to linear control design and facility of integrating changes in the powertrain system technical definition. It also aims at identifying and controlling powertrain systems featuring transport time delay because integrating the delay in the model and control design is crucial on the former's system representability and on the latter's optimality. Based on these premises, we address the powertrain from the engine air-path perspective. We first identify a linear black-box state-space (SS) model of a gasoline engine air-path, using an identification algorithm based on the subspace methods. Different model orders and algorithm parameters are tested and those yielding the best identification and validation results are made clear, which leads to an 85% time gain in future similar identifications. While this part considers the air-path as a whole, the rest of the work focuses on specific air-path components, notably the electric throttle, the heat-exchanger, and the exhaust gas recirculation (EGR). Regarding the electric throttle, we inspire from the physical laws governing the throttle functioning to construct a linear-parameter-varying (LPV) mathematical SS model, which serves to set the regression vector structure of the LPV black-box ARX model, which is representative of an electric throttle test bench and reflects its nonlinearities and discontinuities as it varies from one functioning zone to another. To address the questions of heat and mass transport time delays in the engine air-path, we refer to the heat exchanger and the EGR respectively. Recasting the infinite-dimensional hyperbolic partial differential equations (PDEs), describing these transport phenomena, as a time-delay system facilitates the adjoint system identification and control design. To that end, a space-averaging technique and the method of characteristics are used to decouple the hyperbolic PDEs describing the advective flows in a heat exchanger, and to reformulate them as a time-delay system. Reducing the error between the output temperature of the model and that of a heat exchanger test-bench is what seeks

the gradient-descent method used to identify the parameters of the time-delay system, which surpasses the PDEs in terms of identification accuracy and computational efficiency. On the other hand, the EGR is addressed from a control-oriented perspective, and the PDEs describing the mass transport phenomenon in its tubular structure are recast as a SS system subject to output delay. To regulate the burned gas ratio in the intake gas, the amount of recirculated gas is controlled using two indirect optimal control approaches that take into account the model's infinite-dimensional nature. These methods are accompanied with the Augmented Lagrangian Uzawa method to guarantee the respect of the input and state constraints, thus resulting in a controller of superior performance than the initially existing PID. Generally speaking, this thesis is located half-way between the academic sector and the industrial one. By evaluating the eligibility of integrating existing system identification and control theories in real automotive applications, it highlights the merits and demerits of these theories and opens up new prospects in the domain of model-based powertrain systems optimization.