

# Matrices polynomiales et égalisation de canal

Sylvie Icart

▶ **To cite this version:**

Sylvie Icart. Matrices polynomiales et égalisation de canal. Traitement du signal et de l'image [eess.SP]. Université Nice Sophia Antipolis, 2013. tel-00805547

**HAL Id: tel-00805547**
**https://theses.hal.science/tel-00805547**

Submitted on 28 Mar 2013

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

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







# Matrices polynomiales et égalisation de canal

Sylvie ICART

Mémoire en vue de l'obtention d'une

Habilitation à Diriger des Recherches

Université de Nice-Sophia Antipolis
Laboratoire I3S - CNRS
Polytech'Nice-Sophia Département Electronique
ED STIC

$1^{er}$ mars 2013

| | | | |
|---|---|---|---|
| Président : | Jean-François | LAFAY | |
| Rapporteurs : | Vladimir | KUČERA | MIAS |
| | Éric | MOREAU | ISITV |
| | Martine | OLIVI | INRIA |
| Examinateurs : | Pierre | COMON | GIPSA Lab |
| | Jean-François | LAFAY | IRCCYN |

# Remerciements

L'Habilitation à Diriger des Recherches est un long chemin, merci à ceux qui m'ont accompagnée.

Merci à Vladimir Kučera, spécialiste des matrices polynomiales et rationnelles de m'avoir fait l'honneur d'être rapporteur de cette thèse et d'être venu assister à la soutenance malgré ses nombreuses occupations.

Merci à Eric Moreau d'avoir accepté de rapporter sur mon travail après avoir accepté de rapporter sur les travaux des doctorants que j'ai encadrés. Je souhaite que nous ayons encore l'occasion de travailler souvent ensemble.

Merci à Martine Olivi d'avoir accepté d'être rapporteur. Même si nous sommes proches géographiquement et thématiquement, nous n'avions jamais eu l'opportunité de travailler ensemble. J'espère que cette HDR ouvre la voie à une longue et fructueuse collaboration.
Je suis ravie d'avoir un automaticien, un traiteur du signal et une mathématicienne comme rapporteurs, mon travail se situant à la frontière de ces trois disciplines.

Merci à Pierre Comon pour toutes ces années de collaboration et d'encouragements. Il est parti vers d'autres horizons, mais les moyens techniques actuels vont nous permettre, à n'en pas douter, de poursuivre nos travaux.

Merci à Jean-François Lafay qui a présidé le jury de cette HDR et avec qui j'ai fait mes premiers pas en recherche. Il m'a donné son goût pour l'automatique, les cours bien écrits et pour les démonstrations rigoureuses. Le temps a (trop) vite passé.

Merci à Viviane, Micheline et Marie-Pierre pour leur gentillesse, leur sourire et leur efficacité, ainsi qu'à toute l'équipe administrative et technique qui nous permet de travailler dans de bonnes conditions à l'I3S.

Merci à Gérard Favier qui a été le premier lecteur de ce manuscrit pour ses remarques pertinentes. Et merci à tous ceux de notre équipe-thème-projet avec qui je partage le thé venu des 5 continents.

La pause de midi est un moyen de refaire le monde, tout en courant plus ou moins vite avec l'équipe I3S-INRIA-assimilés. Merci en particulier à notre "coach" Jean-Claude Bermond, grâce à ses conseils avisés j'ai pu faire moins de 4 heures au marathon et me décider à m'inscrire à l'HDR. Merci à Laure pour ses blagues de course et pour son amitié et à Philippe qui apprivoise les canards.

Merci à ma sœur et mes filles Laurane et Solène - mon or et mon soleil - qui m'ont soutenue à la fin de ce parcours, et à mon frère qui m'aide toujours à faire beaucoup de poussière.

Merci enfin à Jean-Pierre avec qui je n'ai jamais fait de recherche mais avec qui j'ai beaucoup trouvé.

*A ma mère, qui aurait aimé lire ce mémoire*
*et à mon père qui en aurait été fier*

# Table des matières


**1 Introduction** **5**

**2 Curriculum vitæ** **7**
2.1 Formation . . . . . 7
2.2 Expérience professionnelle . . . . . 7
2.3 Activités d'Enseignement . . . . . 8
2.4 Fonctions d'intérêt collectif . . . . . 11
2.5 Activités de Recherche . . . . . 13

**3 Matrices L-polynomiales et égalisation** **17**
3.1 Polynômes de Laurent . . . . . 17
3.1.1 Notations . . . . . 18
3.1.2 Propriétés et définitions . . . . . 18
3.2 Matrices polynomiales de Laurent . . . . . 24
3.2.1 Forme de Smith des matrices polynomiales de Laurent 24
3.2.2 Ordre d'une matrice polynomiale de Laurent . . . . . 28
3.2.3 Para-conjugaison et forme hermitienne . . . . . 32
3.2.4 Matrice L-polynomiale para-hermitienne . . . . . 33
3.2.5 Matrice L-polynomiale para-unitaire . . . . . 39
3.3 Diagonalisation d'une matrice polynomiale . . . . . 45
3.3.1 Diagonalisation d'une matrice polynomiale normale . . 45
3.3.2 Valeurs singulières d'une fonction continue sur le cercle à valeurs dans $\mathbb{C}^{n\times n}$ . . . . . 46
3.3.3 Diagonalisation d'une matrice $\boldsymbol{A}\boldsymbol{A}^H$ continue sur le cercle . . . . . 48
3.3.4 Presque-diagonalisation d'une matrice para-hermitienne 49
3.4 Matrices polynomiales et égalisation . . . . . 52
3.4.1 Problème de l'égalisation aveugle de canaux de communication . . . . . 52
3.4.2 Mélanges instantanés . . . . . 53
3.4.3 Mélanges convolutifs FIR . . . . . 53

3.4.4 Algorithmes de diagonalisation de L-matrices . . . . . 54

**4 Conclusion et perspectives** **57**

**5 Bibliographie** **61**
5.1 Publications . . . . . 61
5.2 Références . . . . . 64

**A Quelques définitions classiques** **69**

**B Publications choisies** **73**

# Nomenclature

| | |
|---|---|
| $\mathbb{R}$ | ensemble des nombres réels |
| $\mathbb{C}$ | ensemble des nombres complexes |
| $\mathbb{C}^*$ | $\mathbb{C} \setminus \{0\}$ |
| $\mathcal{C}$ | cercle unité |
| $\jmath$ | unité imaginaire $\jmath^2 = -1$ |
| $a$ | scalaire (réel ou complexe) |
| $\underline{v}$ | tableau à $n$ lignes et une colonne (vecteur) |
| $\boldsymbol{A}$ | matrice |
| $\mathbf{0}$ | matrice nulle |
| $\boldsymbol{I}$ | matrice identité |
| $\boldsymbol{I}_n$ | matrice identité de taille $n \times n$ |
| $a^*$ | conjugué de $a$ |
| $a_{ij}$ | élément de la i$^{\text{e}}$ ligne et j$^{\text{e}}$ colonne de la matrice $\boldsymbol{A}$ |
| $\boldsymbol{A}^H$ | matrice $\boldsymbol{A}$ transposée conjuguée |
| $\widetilde{\boldsymbol{A}}(z)$ | matrice para-conjuguée de $\boldsymbol{A}$ définie par $\widetilde{\boldsymbol{A}}(z) = \boldsymbol{A}^H(\frac{1}{z^*}), \forall z \in \mathbb{C}^*$ |
| $\mathbb{C}[z]$ | anneau des polynômes en l'indéterminée $z$ et à coefficients dans $\mathbb{C}$ |
| $\mathbb{C}[z^{-1}]$ | anneau des polynômes en l'indéterminée $z^{-1}$ et à coefficients dans $\mathbb{C}$ |
| $\mathbb{C}[z, z^{-1}]$ | anneau des polynômes de Laurent en l'indéterminée $z$ ie des polynômes ayant des puissances positives et négatives de $z$ appelés aussi L-polynômes |
| $\mathcal{U}$ | groupe des éléments inversibles de $\mathbb{C}[z, z^{-1}]$ (monômes non nuls) |
| $\mathbb{C}^n[z, z^{-1}]$ | ensemble des tableaux à $n$ lignes et une colonne dont les éléments sont des polynômes de Laurent |

$\mathbb{C}^{n\times n}[z, z^{-1}]$ ensemble des matrices $n \times n$ dont les éléments sont des polynômes de Laurent, ou encore matrices L-polynomiales

$\widetilde{p}(z)$ polynôme para-conjugué de $p$ défini par $\widetilde{p}(z) = p^*(\frac{1}{z^*}), \forall z \in \mathbb{C}^*$

$\deg(p)$ degré en la variable $z$ d'un polynôme $p \in \mathbb{C}[z]$ ou en la variable $z^{-1}$ d'un polynôme $p \in \mathbb{C}[z^{-1}]$

$d(p)$ L-degré d'un polynôme $p \in \mathbb{C}[z, z^{-1}]$, soit degré en la variable $z$ du polynôme associé à $p(z)$

$N$ degré de MacMillan d'une matrice $\in \mathbb{C}^{n\times n}[z^{-1}]$

$\overset{LS}{\sim}$ équivalence au sens de Smith de deux matrices L-polynomiales

$\mathbb{C}(z)$ corps des fractions rationnelles en l'indéterminée $z$ et à coefficients dans $\mathbb{C}$

$\mathbb{C}_s(z)$ anneau des fractions rationnelles stables (pôles à l'intérieur du disque unité)

$\mathcal{C}(\mathbb{R})$ ensemble des fonctions continues définies sur le cercle $\mathcal{C}$ à valeurs dans $\mathbb{R}$

$\mathcal{C}_1(\mathbb{C}^n)$ ensemble des fonctions continues $f$ définies sur le cercle $\mathcal{C}$ à valeurs dans $\mathbb{C}^n$ tq $f(\Omega)f(\Omega)^H = 1, \forall \Omega \in \mathcal{C}$

$1_{\mathcal{C}}$ fonction tq $1_{\mathcal{C}}(\Omega) = 1, \forall \Omega \in \mathcal{C}$

$0_{\mathcal{C}}$ fonction tq $0_{\mathcal{C}}(\Omega) = 0, \forall \Omega \in \mathcal{C}$

# Chapitre 1

# Introduction

Dans la théorie des systèmes linéaires multivariables, les matrices interviennent naturellement afin d'exprimer de manière synthétique les liens unissant les signaux d'entrée et de sortie. Dans le cadre des systèmes linéaires à temps continu invariants dans le temps, les transformées de Laplace des signaux d'entrée et de sortie sont reliées par une matrice de transfert dont chaque élément est une fraction rationnelle de la variable de Laplace : on manipule alors des matrices rationnelles. Une fraction rationnelle est le quotient de deux polynômes, c'est également le cas pour les matrices rationnelles, mais les notions de numérateur et de dénominateur sont un peu plus délicates, notamment car le produit matriciel n'est pas commutatif et que la notion de primarité entre deux facteurs polynomiaux matriciels n'est pas aussi simple que dans le cas scalaire où elle se résume à l'absence de zéro commun [54].

Dans le cas des systèmes linéaires multivariables à temps discret, les transformées en $z$ des signaux d'entrée et de sortie sont reliées par une matrice de transfert qui est une matrice rationnelle dans le cas des systèmes à réponse impulsionnelle infinie et une matrice polynomiale en $z^{-1}$ dans le cas des systèmes à réponse impulsionnelle finie.

Les propriétés de stabilité, causalité, observabilité, commandabilité, minimalité des systèmes sont ainsi bien connues, les références en la matière étant les livres de Kailath [54] pour les systèmes à temps continu en automatique et Vaidyanathan [80] pour les systèmes à temps discret en traitement du signal. Les matrices polynomiales et rationnelles ont tout d'abord été utilisées pour les problèmes d'automatique et de commande, voir par exemple les travaux de Kučera [55], Vidyasagar [81], ou le travail collectif [73], elles interviennent dorénavant aussi en télécommunication [77, 80, 53] notamment via les méthodes de type sous-espaces pour l'identification aveugle [64, 50, 61].

Dans ce mémoire, nous nous focaliserons sur un type de matrices particulier : les matrices polynomiales de Laurent, dont les éléments sont des

polynômes de Laurent, c'est à dire des polynômes avec des puissances positives et négatives de la variable $z$. Ce type de polynômes ne peut être associé à un filtre causal mais il se rencontre notamment lorsqu'on étudie le spectre de signaux à temps discret en sortie de filtre à réponse impulsionnelle finie. Nous commencerons par présenter les propriétés des polynômes de Laurent, puis des matrices polynomiales de Laurent. Nous définirons notamment la L-forme de Smith qui est une extension de la forme de Smith classique, et donnerons une définition précise du degré et l'ordre de ces matrices (notions parfois confondues dans la littérature).

Nous étudierons plus particulièrement les matrices para-hermitiennes et para-unitaires qui sont des matrices respectivement égales à leur matrice para-conjuguée ou dont l'inverse est égale à la para-conjuguée. Nous nous attacherons à développer leurs propriétés particulières en terme de degré notamment, et de factorisation.

Lors de l'étude des systèmes et en traitement du signal, de nombreuses factorisations de matrices à coefficients constants interviennent : factorisations QR (à l'aide d'une matrice orthogonale et d'une matrice triangulaire), LU (à l'aide de deux matrices triangulaires : une inférieure et une supérieure), SVD (décompositions en valeurs singulières à l'aide de deux matrices unitaires), EVD (décompositions en valeurs propres-vecteurs propres) [49]. En particulier, le théorème spectral montre que toute matrice hermitienne est diagonalisable à l'aide d'une matrice unitaire, c'est-à-dire que les matrices intervenant dans l'EVD sont des matrices unitaires. La factorisation de Cholesky d'une matrice hermitienne définie positive se fait quant à elle à l'aide d'une matrice triangulaire et de sa transposée conjuguée. Ces factorisations ne peuvent pas s'étendre simplement aux matrices polynomiales car les coefficients de ces matrices n'appartiennent pas à un corps mais à un anneau (celui des polynômes de Laurent). De plus, certaines propriétés, comme par exemple la positivité, ne peuvent s'entendre que sur le cercle unité (lorsque $z = e^{j\omega}$). Nous montrerons que dans le cas général, une décomposition EVD dont tous les termes sont polynomiaux pour une matrice para-hermitienne définie positive sur le cercle unité n'existe pas, mais qu'on peut presque-diagonaliser ces matrices à l'aide de matrices para-unitaires continues sur le cercle unité. Enfin, nous montrerons quel rôle jouent les factorisations des matrices para-unitaires dans l'égalisation aveugle de systèmes convolutifs multivariables, suivant en cela les travaux de McWhriter et al [35, 62, 44].

# Chapitre 2

# Curriculum vitæ

Sylvie ICART
née le 6 juin 1965, 2 enfants
Maître de Conférences, 61e section depuis octobre 1991
Laboratoire I3S Informatique Signaux Systèmes, Pôle SIS
Polytech'Nice-Sophia Antipolis, Département Electronique

## 2.1 Formation

- Thèse de doctorat, spécialité Automatique et Informatique Industrielle «Etude du découplage ligne par ligne avec stabilité : le cas des systèmes linéaires» au Laboratoire d'Automatique de Nantes, juillet 1990, directeur de thèse : J.F. Lafay, rapporteurs : J. Descusse (LAN), J.M. Dion (LAG), V. Kucera (UTIA CSAV), membre : B. Espiau (INRIA), mention Très Honorable.
- DEA en Automatique et Informatique Industrielle à l'Ecole Nationale Supérieure de Mécanique de Nantes, «Quelques résultats sur le problème de Morgan», octobre 1987 (mention Bien).
- Ingénieur ENSM, spécialité Automatique, juin 1987 (mention Très Bien).
- Classes Préparatoires M' au lycée Masséna de Nice, 1982-84.
- Baccalauréat série C au lycée Amiral de Grasse, 1982 (mention Très Bien).

## 2.2 Expérience professionnelle

- Maître de Conférences à l'Université de Nice-Sophia Antipolis :
  UFR Sciences, Département EEA de 1991 à 1999,
  Polytech'Nice-Sophia, département Electronique, anciennement Ecole

Supérieure d'Ingénieurs de Nice-Sophia Antipolis (ESINSA) depuis 1999. Membre du Laboratoire I3S depuis 1991.
Congé parental à temps partiel (80%) de septembre 2001 à août 2003.
Congé pour recherches (CRCT) de septembre 2011 à août 2012.

- Boursière post-doctorale à l'INRIA Sophia–Antipolis, projet PRISME, Etude de la commande d'un bras de robot flexible, en 1990-91.
- Stagiaire à l'Aérospatiale de Cannes, service Attitude Orbite Contrôle Satellite, découverte du contrôle en entreprise, août 1991.
- Doctorante au LAN, Contrôle de systèmes multivariables, en 1987-90.
- Moniteur de l'Enseignement Supérieur à l'ENSM en 1989-90.
- Vacataire à l'ENSM en électronique et automatique, à l'IUT GEII en automatique et à l'UER Sciences Economiques en mathématiques de 1987 à 1989.
- Stagiaire au Centre de Mathématiques Appliquées de l'Ecole des Mines de Paris à Sophia-Antipolis : Etude d'une bille sur un plan incliné (2 mois 1/2 en 1986).

## 2.3 Activités d'Enseignement

Depuis mes débuts en tant qu'enseignant-chercheur, j'ai toujours veillé à enseigner à différents niveaux : du premier cycle au master, en cours, travaux dirigés et travaux pratiques, et à essayer de transmettre à mes étudiants mon enthousiasme pour les mathématiques appliquées à l'ingénierie.

Par exemple, enseigner à des étudiants à bac +1 des outils mathématiques pour l'ingénieur permet de mieux appréhender les acquis de nos étudiants tant d'un point de vue mathématiques que d'un point de vue méthode de travail. En ce qui concerne les travaux pratiques, il me paraît essentiel que les étudiants ne travaillent pas uniquement en simulation avec des logiciels dédiés (tels Matlab ou Scilab) mais également sur des maquettes pédagogiques. Même si cela est notablement plus exigeant d'un point de vue préparation et entretien, ceci permet aux étudiants de se confronter à des problèmes réels.

Plutôt que de dresser une liste exhaustive des enseignements que j'ai effectués, j'ai choisi d'indiquer les grandes thématiques que j'ai abordées [1]. La liste des enseignements actuels ainsi que les documents pédagogiques afférents sont disponibles à l'adresse : http ://users.polytech.unice.fr/∼icart.

**Mathématiques** : Lorsque j'ai été nommée maître de conférences à l'université de Nice, j'ai eu la responsabilité des cours d'algèbre linéaire de 1ère année à l'ESINSA durant 2 ans (45h Cours, 45h TD). L'école ayant

1. les horaires indiquées sont ceux par étudiant, le nombre de groupes allant de 2 à 8 en TD et de 6 à 10 en TP.

ouvert la même année que celle où j'ai été recrutée, j'ai eu la liberté d'innover en illustrant, par exemple, les notions matricielles sur des exemples physiques concrets.

Toujours dans ce souci de relier mathématiques et ingénierie et suite au travail d'un groupe de réflexion, il a été décidé d'introduire en 1999 un nouveau cours : "Outils Mathématiques pour l'Ingénieur" (OMI) en 1ère année. J'ai participé à l'élaboration du programme de ce cours. J'ai la responsabilité de la partie géométrie de cette matière depuis lors (12h de cours, 12h de TD), et adapte son contenu en fonction des programmes de terminale. En raison de l'augmentation des flux d'étudiants, je n'assure plus seule ces enseignements, mais travaille maintenant en équipe avec des vacataires.

**Logique et Electronique** : Parallèlement à ces enseignements plutôt théoriques, j'ai effectué des TP d'Electronique en ESINSA 1 (utilisation d'oscilloscopes et de GBF, porte ET, OU, diode, multiplexeur, 15h TP) ainsi que des TP de logique en Maîtrise EEA (bascules, systèmes synchrones et asynchrones 30h TP) lorsque j'étais en poste à l'UFR Sciences.

**Traitement du Signal** : J'ai tout d'abord eu la responsabilité des cours de Signaux en 2e année de DEUG SPI (12h Cours, 12h de TD). Ce cours était orienté sur l'étude des réponses transitoires de systèmes physiques simples à l'aide de la transformée de Laplace.

Lors du renouvellement de la maquette de la maîtrise EEA en 1996, il a été décidé d'introduire le traitement du signal dans ce cursus. J'ai eu la responsabilité de ces Cours, TD et TP en licence (transformée de Fourier, corrélation, spectre, filtrage) puis en maîtrise (filtrage linéaire des signaux déterministes et stochastiques à temps discret). Pour illustrer les notions de fréquence sur des signaux réels, j'ai choisi de traiter des signaux de parole à l'aide du logiciel Matlab. J'ai effectué ces enseignements (16h Cours, 16h TD) durant 2 ans, avant ma mutation à l'ESINSA.

Depuis 1999, j'ai la responsabilité des Cours, TD et TP de Signaux en deuxième année du premier cycle ou PeiP2 (15h Cours, 15h TD, 15h TP) : systèmes linéaires du premier et du second ordre, notions de filtre passe-haut et passe-bas. Nous traitons également des signaux sonores. J'assure le cours ainsi qu'une partie des TD et TP et coordonne ces enseignements pour lesquels, suivant les années, interviennent des vacataires, des moniteurs de l'enseignement supérieur (DCCE) ou des ATER (10 groupes de TP). Depuis que le premier cycle est devenu un Parcours des écoles d'ingénieurs Polytech, les étudiants ne se destinent plus uniquement à l'électronique au sens large, donc même si les théma-

tiques abordées sont les mêmes, la finalité du module est différente. En effet, ce cours d'introduction au traitement du signal et à la théorie des systèmes linéaires ne sera suivi pour la plupart des étudiants d'aucun approfondissement.

**Automatique** : Après avoir fait des TP puis des TD et enfin des cours sur les systèmes asservis linéaires continus et les systèmes d'état en maîtrise EEA, j'enseigne actuellement l'automatique continue et discrète au département Electronique de Polytech'Nice-Sophia.

En 1ère année de cycle ingénieur (Elec3), je participe aux TD (23h TD) et ai la responsabilité des TP (16.5h TP, 6 groupes). Comme indiqué dans le préambule, les TP d'asservissement s'effectuent sur des maquettes pédagogiques. Les étudiants peuvent ainsi se rendre compte que l'automatique "ce n'est pas que des maths", voir la différence entre simulation et réalité, et appréhender les limites entre théorie et pratique. J'effectue ces enseignements pour moitié avec des ATER qui changent chaque année.

En Elec4, j'ai la responsabilité des cours et TD d'Automatique (23h Cours, 23h TD) : lois de commande par retour d'état, correcteurs numériques à avance et retard de phase, PID. J'ai également eu la responsabilité des TP d'Automatique en Elec4 durant plusieurs années (1999-2007).

Depuis 2009, j'assure la mise à niveau en automatique d'étudiants chinois en programme Mundus au niveau bac+3. Ces cours de 20 heures sont une nouvelle expérience pédagogique car les étudiants n'ont ni la même culture scientifique ni les mêmes méthodes de travail que nos étudiants habituels.

Les étudiants au département Electronique n'ont pas vocation à devenir des spécialistes de l'automatique. Le but de cet enseignement est qu'ils puissent en appréhender les notions fondamentales (boucle ouverte, boucle fermée, contre-réaction, problème de stabilité, interactions entre analogique et numérique...) et qu'ils soient aptes à saisir les problématiques associées, ce que j'ai pu observer notamment lors de leurs stages de fin d'études.

**Systèmes multivariables** : L'étude des systèmes multivariables ne se fait qu'au niveau Master 2 puisque c'est à la fois un problème compliqué et pointu. C'est pourquoi je l'ai enseigné en DEA ARAVIS, puis au Master SICOM (15h Cours durant 6 ans). L'étude de ces systèmes est illustrée par des problèmes d'automatique (commande) mais également de télécommunications lorsque plusieurs émetteurs et plusieurs récepteurs travaillent simultanément : il s'agit alors d'égalisation multivariable. Ce dernier point était en phase avec mes travaux de recherche

actuels. Depuis la refonte des masters en 2007, l'étude des systèmes multivariables n'est plus enseignée localement.

**Activités pédagogiques annexes** : Enfin, je participe régulièrement aux Journées Portes Ouvertes ainsi qu'au Sophia Forum et aux entretiens Polytech qui nous permettent de recruter nos futurs étudiants. J'encadre aussi tous les ans des stagiaires en entreprise au niveau technicien ou ingénieurs (3 à 6 par an).

## 2.4 Fonctions d'intérêt collectif

Avant de détailler les fonctions d'intérêt collectif auxquelles j'ai consacré beaucoup de temps et d'énergie, voici, par ordre chronologique, la liste des fonctions d'intérêt collectif qui m'ont été confiées ou pour lesquelles j'ai été élue, au sein du laboratoire, de l'école ou de l'université :

- Membre du conseil du laboratoire I3S de février 93 à septembre 94.
- Membre suppléante de la commission de spécialistes 61e section de l'UNS de 1998 à 2002.
- Responsable des stages techniciens au Département Electronique de Polytech'Nice-Sophia d'octobre 1999 à juin 2009.
- Coresponsable de l'option Traitement Numérique du signal au Département Electronique de Polytech de 2000 à 2003.
- Membre titulaire de la commission de spécialistes 61e section de septembre 2002 à juillet 2008.
- Membre suppléant du Comité des Projets de l'I3S de 2003 à 2005.
- Assesseur de la commission de spécialistes 61e section, mars 2005 à juillet 2008 (fin des CS).
- Membre du Conseil d'Ecole de Polytech de 2005 à 2010.
- Membre du jury du concours ITRF session 2006 niveau Ingenieur BAP Electronicien.
- Responsable de l'option TNS de 2006 à 2009.
- Membre du Conseil du Département Réseaux et Télécom de l'IUT Nice Côte d'Azur de 2007 à 2011.
- Responsable des stages de fin d'études au Département Electronique depuis juin 2009.
- Membre du vivier de la section 61 de l'université de Nice-Sophia Antipolis depuis février 2009.
- Membre du Comité de Sélection du poste MCF n°184, EPU Nice-Sophia Antipolis, avril-mai 2009.
- Membre du Comité de Sélection du poste MCF n°167, IUT GEII Nice, mai 2011.

– Membre du Conseil de Département Electronique de Polytech'Nice-Sophia depuis juin 2011.

**Responsabilité des Stages Techniciens (1999-2009)**

Les stages techniciens concernent 70 étudiants. D'une durée de 8 semaines minimum, ils ont lieu en fin d'Elec4. La responsabilité des stages est un mélange de travail administratif (établissement des conventions, envoi et suivi des conventions), de contact avec le milieu d'accueil (évaluation des stages), de suivi pédagogique (validation des stages) et d'organisation (choix des jurys de soutenance, répartition des rapports de stage aux relecteurs, planning des jurys et enfin compilation des notes données par les entreprises, les relecteurs et le jury de soutenance). Pendant les 10 années où j'ai été responsable de ces stages, nous ne disposions au département Electronique d'aucun support administratif, c'était donc une activité très chronophage.

Le stage technicien est une étape importante dans la vie de nos étudiants. Leur premier stage, ou stage ouvrier, se fait majoritairement dans un domaine fort éloigné de l'électronique. Pour beaucoup de nos élève-ingénieurs, le stage technicien est donc le premier contact avec le milieu industriel dans lequel ils évolueront durant leur future carrière. C'est aussi pour les enseignants, un indicateur des tendances actuelles et un outil leur permettant de faire évoluer le contenu des enseignements, notamment via les exemples d'application.

**Responsabilité de l'option TNS (2006-09)**

L'option TNS au département Elec s'étend sur 2 semestres (plus un semestre de stage). La responsabilité consiste en l'établissement des emplois du temps, en la recherche d'intervenants académiques et industriels (plus d'une vingtaine) et à la coordination des différents enseignements. Enfin, il faut adapter les maquettes pédagogiques afin de coller au mieux à cette spécialité en constante évolution.

Nous avions également un partenariat privilégié avec la société Texas Instrument, je m'occupais du Prix Texas ainsi que de l'organisation de la rencontre annuelle du programme Elite de TI qui s'adressait à tous les enseignants-chercheurs du domaine à l'UNS (IUT, UFR Sciences, Polytech).

Parallèlement, et à cause d'un réel sous-encadrement, je m'occupais de recenser tous les enseignements et intervenants en automatique et traitement du signal au département ainsi qu'en premier cycle. Ceci a conduit notamment à l'établissement d'une fiche de poste d'un professeur en 61e section (poste pourvu en janv. 2009) et à celle d'un poste de maître de conférences (poste pourvu en sept. 2009).

**Responsabilité des stages de fin d'études (2009-)**

Depuis septembre 2009, j'ai pris la responsabilité des stages de fin d'études au département Electronique. Ces stages d'une durée de 6 mois concernent une soixantaine d'étudiants.

Comme pour les stages techniciens, cette responsabilité est multiple mais s'étend sur une plus grande partie de l'année universitaire : jurys, plannings, diffusion des offres de stage, conseils aux étudiants, affectation des tuteurs, conventions, relations avec les entreprises afin de pérenniser les accueils de stagiaires. En parallèle, je m'occupe du devenir des étudiants nouvellement diplômés. Concernant la partie administrative, nous bénéficions depuis 2 ans d'une assistante à temps partiel, ce qui a allégé mon travail purement administratif. De plus, grâce à notre expérience, j'ai pu établir un cahier des charges puis testé la gestion des conventions en ligne. Le site est maintenant opérationnel et est utilisé depuis 2 ans par tous les départements de Polytech (http ://conventions.polytech.unice.fr/). J'utilise également la plateforme Jalon de l'université pour gérer les dépôts en ligne des stages.

L'importance des stages de fin d'études au sein d'une école d'ingénieurs n'est plus à démontrer. La mutation de nos étudiants d'élèves en ingénieurs grâce à leurs acquis, leur travail et leur capacité à mener à bien un projet est le but d'une formation. Les relations que nous tissons avec les entreprises et les laboratoires lors de ces stages sont essentielles. Elles nous permettent d'adapter le contenu des enseignements via une forte interaction entre le responsable des stages, les responsables d'option et les maîtres de stage en entreprise.

## 2.5 Activités de Recherche

Le point commun de tous mes travaux de recherche est l'utilisation de l'algèbre linéaire avec une emphase particulière sur les méthodes utilisant des matrices polynomiales. Dans cette partie, les références citées sont celles dont je suis co-auteur.

Après avoir fait de la commande multivariable en thèse [1, 5, 6, 7, 27, 28], puis de l'automatique pour la robotique en post-doctorat [29, 8], je me suis orientée vers le traitement du signal et l'égalisation multivariable.

J'ai participé à l'encadrement de la thèse d'Olivier Lemoine, dirigée par Gérard Alengrin et soutenue en octobre 1995 sur la détection de signaux non stationnaires par représentation temps-fréquence [9, 10].

La participation au GDR Traitement du Signal [30] m'a fait découvrir un nouvel axe de recherche : l'égalisation aveugle multicapteurs, c'est-à-dire

de systèmes pour lesquels on ne dispose pas de séquence d'apprentissage pour l'identification mais dont on connaît les propriétés statistiques. J'ai ainsi encadré seule deux stages de DEA. Roland Gautier a travaillé sur la séparation de mélanges convolutifs de février à juillet 1995 ; ce travail a donné lieu à une publication à ICASSP'96 [11]. Le stage de DEA de Ansi Ansori portait quant à lui sur la simulation d'une chaîne de communication à l'aide du logiciel Matlab, de février à juin 1996.

Parallèlement, dans notre groupe de travail local, nous nous sommes intéressés aux problèmes d'égalisation de systèmes monovariables lorsque l'erreur est bornée (ce qui est une hypothèse tout à fait réaliste !), en tenant compte du caractère fini de l'alphabet [12, 13].

Puis, nous avons travaillé avec Joël LeRoux sur l'égalisation de systèmes à l'aide d'un critère exprimé dans le domaine spectral et applicable à des signaux complexes [14].

J'ai ensuite co-dirigé avec Pierre Comon la thèse de Ludwig Rota : «Egalisation aveugle de systèmes multi-utilisateurs» (soutenue le 15 décembre 2004). Les systèmes étudiés comportent plusieurs émetteurs et plusieurs capteurs. Les signaux émis sont supposés inconnus mais indépendants. Le canal de transmission est linéaire stationnaire invariant dans le temps et à réponse impulsionnelle finie : sa matrice de transfert est alors une matrice polynomiale. Si les observations sont blanchies, cette matrice est de plus para-unitaire. Nous nous servons des propriétés des matrices para-unitaires et d'un critère de constraste basé sur les statistiques d'ordre supérieur à deux des observations pour construire un égaliseur permettant de retrouver les sources. Ces travaux ont donné lieu à trois communications dans des conférences internationales [17, 15, 16]. L'extension des derniers résultats est parue en mars 2009 dans la revue Signal Processing [2].

De février 2007 à juin 2010, j'ai co-encadré avec Luc Deneire et Lieven De Lathauwer (de l'université de Louvain en Belgique) la thèse de Mikael Sorensen dans le cadre du programme européen Marie Curie Est Signal : «Tensor Tools with Application in Signal Processing», thèse soutenue le 14 juin 2010. Ces travaux portent sur la diagonalisation de tenseurs et le problème se pose différemment selon que l'on a des tenseurs symétriques ou non. En effet, l'utilisation des moments d'ordre supérieur à deux pour l'étude des systèmes multivariables, fait intervenir non plus seulement des matrices mais également des tenseurs (algèbre multilinéaire). Les compétences en algèbre multilinéaire de plusieurs chercheurs au sein de l'équipe Signal de l'I3S nous ont permis de mener à bien des études pointues sur ce sujet. Ces travaux ont donné lieu à trois communications dans des conférences internationales [18, 19, 20] ainsi qu'à un article dans la revue Signal Processing [4] et un à paraître dans SIMAX [3].

Actuellement, j'étudie plus particulièrement les propriétés structurelles des matrices polynomiales de Laurent. Ce travail a donné lieu à un article de conférence [21] et va être développé dans les sections suivantes.

Enfin, je suis assez régulièrement relectrice pour des conférences (ICASSP, ITS, ICC, EUSIPCO) ou des revues (Automatica, IEEE Transaction on Signal Processing, Signal Processing, International Journal of Adaptive Control and Signal Processing, Springer Signal Image and Video Processing).

## Encadrements et jurys

### Encadrement

- Thèse O. Lemoine «Détection de signaux non stationnaires par représentation temps-fréquence», dirigée par G. Alengrin, octobre 1995 (30%).
- DEA R. Gautier : «Séparation de mélanges convolutifs», juillet 1995 (100%).
- DEA A. Ansori «Simulation d'une chaîne de communication à l'aide du logiciel Matlab», juin 1996 (100%).
- Thèse L. Rota «Egalisation aveugle de systèmes multi-utilisateurs», décembre 2004, co-direction avec P. Comon (50%).
- Thèse M. Sorensen «Tensor Tools with Application in Signal Processing», juin 2010, co-direction avec L. Deneire et L. De Lathauwer (30%).

O. Lemoine est actuellement chef de projets à Thalès, R. Gautier maître de conférences à Brest, L. Rota ingénieur d'études chez Orange Business Services et M. Sorensen est en post-doctorat à U.K. Leuven.

### Participation à des jurys de thèse

- Ludwig Rota «Egalisation aveugle de systèmes multi-utilisateurs», 15 décembre 2004.
- Mikael Sorensen «Tensor Tools with Application in Signal Processing», le 14 juin 2010.
- Thomas Bouilloc «Applications de décompositions tensorielles à l'identification de modèles de Volterra et aux systèmes de communication MIMO-CDMA», 9 Décembre 2011.

## Publications

La liste de mes publications se trouve en page 61.

# Chapitre 3

# Matrices L-polynomiales et égalisation

Comme indiqué précédemment, nous allons plus particulièrement développer dans ce mémoire les résultats de nos derniers travaux sur les matrices polynomiales de Laurent, appelées encore matrices L-polynomiales et leur application en égalisation aveugle multivariable. Nous ferons d'abord des rappels sur les polynômes de Laurent puis exposerons différentes propriétés des matrices polynomiales de Laurent notamment en terme de décomposition et de factorisation. Les notions d'hermitianité et d'unitarité s'étendent aux matrices polynomiales, on parle alors de matrice para-hermitienne et de matrice para-unitaire ; mais les décompositions en valeurs singulières (PSVD) ou en valeurs-propres, vecteurs propres (PEVD) ne peuvent s'étendre car les éléments des matrices considérées n'appartiennent pas à un corps. Nous montrerons cependant que l'on peut approximer de telles décompositions à l'aide de fonctions continues sur le cercle unité, sans toutefois pouvoir imposer aux différents facteurs d'être polynomiaux.

Enfin, nous montrerons quel rôle jouent les matrices para-unitaires et para-hermitiennes dans l'égalisation aveugle de mélanges convolutifs à réponse impulsionnelle finie. [1]

## 3.1 Polynômes de Laurent

Les notations et dénominations sur les polynômes diffèrent d'un ouvrage à l'autre et suivant le domaine d'application considéré. Dans cette partie, nous présentons les notations et définitions utilisées dans la suite de ce document

1. On aurait pu commencer par le problème qui a motivé ces travaux, mais il aurait fallu pour cela introduire de nombreuses notations qui nuiraient à la lecture.

et qui seront utiles pour l'études des matrices polynomiales de Laurent. Une annexe page 69 présente les autres définitions rencontrées dans la littérature.

### 3.1.1 Notations

On note $\mathbb{C}[z]$ l'anneau des polynômes [2] en l'indéterminée $z$ et à coefficients dans $\mathbb{C}$ :

$$p(z) \in \mathbb{C}[z] \text{ ssi } p(z) = \sum_{i=0}^{n} p_i z^i \text{ avec } n \in \mathbb{N}, p_i \in \mathbb{C}$$

Le degré de $p$ est noté $\deg(p)$ et correspond au plus grand entier $i$ tel que $p_i \neq 0$.

On note $\mathbb{C}_s[z]$ l'ensemble des polynômes dont les racines sont toutes de module strictement inférieur à 1, ces polynômes sont à minimum de phase strict.

On note $\mathbb{C}[z^{-1}]$ l'anneau des polynômes en l'indéterminée $z^{-1}$ et à coefficients dans $\mathbb{C}$ :

$$p(z) \in \mathbb{C}[z^{-1}] \text{ ssi } p(z) = \sum_{i=0}^{n} p_i z^{-i} \text{ avec } n \in \mathbb{N}, p_i \in \mathbb{C}$$

Le degré de $p$ est aussi noté $\deg(p)$ et correspond au plus grand entier $i$ tel que $p_i \neq 0$.

On note $\mathbb{C}[z, z^{-1}]$ l'anneau des polynômes de Laurent en l'indéterminée $z$ c'est à dire des polynômes ayant des puissances positives et négatives de $z$ :

$$p(z) \in \mathbb{C}[z, z^{-1}] \text{ ssi } p(z) = \sum_{i=m}^{n} p_i z^i \text{ avec } m, n \in \mathbb{Z}, p_i \in \mathbb{C}.$$

On appellera par la suite ces polynômes des L-polynômes.

### 3.1.2 Propriétés et définitions

**Propriété 1.** *L'ensemble des polynômes de Laurent* $(\mathbb{C}[z, z^{-1}], +, .)$ *est un anneau intègre.*

$(\mathbb{C}[z, z^{-1}], +)$ est un groupe commutatif : $+$ est une loi interne, commutative, associative, l'élément neutre est le polynôme nul, l'opposé de $p(z)$ est $-p(z)$.

2. par souci de clarté, nous noterons la plupart du temps un polynôme en faisant apparaître explicitement la variable associée, c'est à dire on notera le polynôme $p$ comme $p(z)$, $p(z^{-1})$ est alors la fonction obtenue en substituant $z$ par $z^{-1}$.

La multiplication est associative et possède un élément neutre $p(z) = z^0 = 1$, elle est distributive par rapport à l'addition, il n'y a pas de diviseur de zéro donc $(\mathbb{C}[z, z^{-1}], +, .)$ est un anneau intègre. □

L'ensemble de ses éléments inversibles (qui est donc un groupe commutatif) est défini par $\mathcal{U} = \{p(z) \in \mathbb{C}[z, z^{-1}], p(z) = az^\alpha \text{ avec } a \in \mathbb{C}^*, \alpha \in \mathbb{Z}\}$, ce sont des monômes.

Deux L-polynômes sont associés s'ils sont égaux à une multiplication par un élément de $\mathcal{U}$ près. Comme pour un polynôme "classique" auquel on peut associer un polynôme unitaire (ou monique c'est à dire dont le coefficient du terme de plus haut degré est égal à 1) on peut associer à tout L-polynôme un polynôme L-unitaire :

**Définition 2.** *On dira qu'un polynôme est L-unitaire s'il est unitaire et qu'il n'admet pas zéro comme racine.*

Soit $p(z) \in \mathbb{C}[z, z^{-1}]$, posons $p(z) = \sum_{i=m}^{n} p_i z^i$ avec $p_i \in \mathbb{C}, m, n \in \mathbb{Z}, m \leq n, p_m$ et $p_n$ non nuls. On peut donc factoriser
$p(z) = p_n z^m \sum_{i=m}^{n} \frac{p_i}{p_n} z^{i-m} = p_n z^m \sum_{i=0}^{n-m} \pi_i z^i = p_n z^m \pi(z)$ avec $\pi(z) \in \mathbb{C}[z]$ et L-unitaire car par définition $\pi_0 = \frac{p_m}{p_n} \neq 0$ et $\pi_{n-m} = \frac{p_n}{p_n} = 1$. Notons que le polynôme $\pi$ associé à $p$ ainsi défini est unique.

Pour montrer que l'anneau des polynômes de Laurent est euclidien, on a besoin de définir un stathme $d$, c'est à dire une application de $\mathbb{C}[z, z^{-1}] \setminus \{0\}$ dans $\mathbb{N}$ telle que $d(pq) \leq d(p)d(q)$ pour tout L-polynômes $p, q$ non nuls et auquel on peut associer une division euclidienne.

**Définition 3.** *Soit $p(z) \in \mathbb{C}[z, z^{-1}]$, $p(z) = \sum_{i=m}^{n} p_i z^i$ avec $p_i \in \mathbb{C}, m, n \in \mathbb{Z}, m \leq n, p_m$ et $p_n$ non nuls. Posons*

$$p(z) = p_n z^m \pi(z) \tag{3.1}$$

*$\pi(z) \in \mathbb{C}[z]$ est un polynôme de degré $n-m$ ($\geq 0$ par définition). On appellera L-degré d'un L-polynôme $p$ le degré du polynôme associé $\pi$, on notera ce L-degré $d(p)$.*

Notons que la multiplication par un monôme ne modifie pas le degré ainsi défini et que le L-degré ne coïncide pas avec le degré lorsqu'on est sur $\mathbb{C}[z]$ : par exemple $p(z) = z^2 + z^4 = z^2(1 + z^2)$ a un degré 4 sur $\mathbb{C}[z]$ mais un L-degré de 2.

Enfin, pour un polynôme $p$ appartenant à $\mathbb{C}[z]$, on a $\deg(p) \geq d(p)$ avec égalité ssi $z = 0$ n'est pas racine de $p$. Soit $p \in \mathbb{C}[z]$ tel que $\deg(p) = n$, on peut donc écrire :

- $p(z) = z^0 \sum_{i=0}^{n} p_i z^i$ avec $p_0 \neq 0$ si 0 n'est pas racine de $p$, par définition du L-degré on a donc $d(p) = n = \deg(p)$

- si 0 est racine de $p$, alors on écrit $p(z) = z^m \pi(z)$ avec $\pi$ qui n'admet pas de racine en 0. On a alors $\deg(p) = m + \deg(\pi) = m + d(p)$ d'après la définition du L-degré.

Remarque : le L-degré ainsi défini est identique à l'ordre effectif utilisé en Télécom, notamment par McWhriter [44].

**Propriété 4.** $(\mathbb{C}[z, z^{-1}], +, .)$ *est un anneau euclidien.*

Soient $a(z)$ et $b(z)$ deux L-polynômes quelconques avec $b \neq 0$. Montrons qu'il existe deux L-polynômes $q(z)$ et $r(z)$ tels que

$$a(z) = b(z)q(z) + r(z) \text{ avec } r = 0 \text{ ou } d(r) < d(b)$$

Posons $a(z) = z^{m_a}\alpha(z), \alpha(z) \in \mathbb{C}[z], \deg(\alpha) = d(a)$, et $b(z) = z^{m_b}\beta(z), \beta(z) \in \mathbb{C}[z]$, $\deg(\beta) = d(b)$.

$\mathbb{C}[z]$ étant un anneau euclidien, on peut faire la division euclidienne de $\alpha$ par $\beta$ :

$$\exists\ \kappa, \rho \in \mathbb{C}[z] \text{ tq } \alpha(z) = \beta(z)\kappa(z) + \rho(z), \rho = 0 \text{ ou } \deg(\rho) < \deg(\beta)$$

Par définition de $\alpha$ et $\beta$, on a donc :

$$z^{-m_a}a(z) = z^{-m_b}b(z)\kappa(z) + \rho(z), \text{ soit } a(z) = b(z)\{z^{m_a - m_b}\kappa(z)\} + z^{m_a}\rho(z)$$

$z^{m_a - m_b}\kappa(z)$ et $z^{m_a}\rho(z)$ sont des L-polynômes et $d(z^{m_a}\rho(z)) = d(\rho)$ par définition du L-degré. Si $\rho$ est non nul, $d(b) = \deg(\beta) > \deg(\rho) \geq d(\rho)$. En posant $q(z) = z^{m_a - m_b}\kappa(z)$ et $r(z) = z^{m_a}\rho(z)$, on obtient bien la division euclidienne de $a$ par $b$, avec $d(r) < d(b)$. Si $\rho = 0$ alors on a $r = 0$. □

Rappelons qu'un idéal $\mathcal{I}$ d'un anneau $\mathcal{R}$ est un sous groupe additif tel que $\forall x \in \mathcal{R}$ et $y \in \mathcal{I}$, $xy \in \mathcal{I}$ et $yx \in \mathcal{I}$. Il est principal s'il peut-être généré par un seul élément.

$\mathbb{C}[z, z^{-1}]$ étant euclidien, il est principal (principal ideal domain PID). Comme l'anneau est principal on peut définir les notions de plus grand commun diviseur (pgcd) et de plus petit commun multiple (ppcm) d'une famille de L-polynômes. Un plus grand diviseur n'étant défini qu'à une multiplication par un élément inversible près, on imposera à ce pgcd d'être L-normalisé :

**Définition 5.** *Soit* $\{p_1, \ldots, p_n\}$ *une famille d'éléments de* $\mathbb{C}[z, z^{-1}]$. *On appelle plus grand diviseur commun des* $p_i$ *le polynôme* $p \in \mathbb{C}[z]$ *tel que* $p$ *divise* $p_i$ *pour tout* $i$ *et* $p$ *est L-unitaire.*

**Propriété 6.** $(\mathbb{C}[z, z^{-1}], +, .)$ *est une algèbre sur* $\mathbb{C}$.

Cette propriété est évidente : si $p(z)$ et $q(z)$ sont des L-polynômes et $a$ et $b$ sont des nombres complexes, $ap(z)$ est un L-polynôme, $a(p(z)+q(z)) = ap(z)+aq(z)$, $(a+b)p(z) = ap(z)+bp(z)$, $a(bp(z)) = (ab)p(z)$ et enfin $1p(z) = p(z)$. □

**Définition 7.** *Soit $p(z) \in \mathbb{C}[z, z^{-1}]$ un L-polynôme, le polynôme para-conjugué de $p(z)$ est le L-polynôme $\widetilde{p}(z) \in \mathbb{C}[z, z^{-1}]$ défini*[3] *par*

$$\widetilde{p}(z) = p^*(\frac{1}{z^*}), \forall z \in \mathbb{C}^*$$

La para-conjugaison est une involution sur l'algèbre $(\mathbb{C}[z, z^{-1}], +, .)$ puisque pour tous L-polynômes $p, p_1, p_2$ on a $\widetilde{\widetilde{p}} = p$, $\widetilde{p_1 + p_2} = \widetilde{p_1} + \widetilde{p_2}$, $\widetilde{p_1 p_2} = \widetilde{p_2}\widetilde{p_1}$ et $\forall \lambda \in \mathbb{C}, \widetilde{\lambda p} = \lambda^* \widetilde{p}$.

Lorsqu'on se restreint au cercle unité, c'est à dire si $z = e^{\jmath\omega}, \omega \in \mathbb{R}$, alors $\frac{1}{z^*} = z$ et $\widetilde{p}(e^{\jmath\omega}) = p^*(e^{\jmath\omega})$. En particulier, $p(e^{\jmath\omega})\widetilde{p}(e^{\jmath\omega}) = |p(e^{\jmath\omega})|^2 \in \mathbb{R}^+, \forall\omega$.

Remarquons que si $p(z) \in \mathbb{C}[z]$ alors $\widetilde{p}(z) \in \mathbb{C}[z^{-1}]$.

**Propriété 8.** *Le L-degré est identique pour un L-polynôme et son para-conjugué.*

Soit $p(z) = \sum_{i=m}^{n} p_i z^i$, $p_m$ et $p_n$ non nuls, $m \leq n$, alors par définition $\widetilde{p}(z) = \sum_{i=m}^{n} p_i^* z^{-i} = z^{-n} \sum_{i=m}^{n} p_i^* z^{n-i}$, soit en faisant le changement d'indice $l = n - i$, $\widetilde{p}(z) = z^{-n} \sum_{l=0}^{n-m} p_{n-l}^* z^l$. $p_n^*$ et $p_m^*$ étant non nuls par hypothèse, le L-degré de $\widetilde{p}(z)$ est égal à $n - m$ qui est aussi par définition le L-degré de $p$. Cette démonstration vaut pour $m, n \in \mathbb{N}$ ou $\mathbb{Z}$, c'est à dire pour un polynôme ou un polynôme de Laurent. □

Remarquons enfin que le polynôme $\gamma(z)$ associé au polynôme para-conjugué $\widetilde{p}(z)$ de $p(z)$ est le polynôme réciproque du polynôme $\pi(z)$ associé à $p(z)$ c'est-à-dire que $\pi(z)$ et $\gamma(z)$ sont des polynômes de même degré $(n - m)$ vérifiant $\gamma(z) = z^{n-m}\widetilde{\pi}(z)$.
En effet, posons $p(z) = \sum_{i=m}^{n} p_i z^i = z^m \sum_{i=0}^{n-m} \pi_i z^i$ avec $\pi_i = p_{i+m}$. On a $\widetilde{p}(z) = z^{-n} \sum_{l=0}^{n-m} p_{n-l}^* z^l = z^{-n} \sum_{l=0}^{n-m} \gamma_l z^l$, soit $\gamma_l = p_{n-l}^*$.
$\pi(z)$ étant un polynôme de degré $n - m$, son polynôme réciproque est
$\pi_*(z) = z^{n-m}\widetilde{\pi}(z) = z^{n-m}\pi^*(\frac{1}{z^*}) = z^{n-m} \sum_{i=0}^{n-m} \pi_i^* z^{-i} = \sum_{i=0}^{n-m} \pi_i^* z^{n-m-i} = \sum_{i=0}^{n-m} p_{i+m}^* z^{n-m-i}$, soit en faisant le changement d'indice $l = n - m - i$, $\pi_*(z) = z^{n-m} \sum_{l=0}^{n-m} p_{n-l}^* z^l$, soit $\pi_*(z) = \gamma(z)$.

Si $\pi(z)$ et $\gamma(z)$ sont définis comme les polynômes L-unitaires associés à $p(z)$ et $\widetilde{p}(z)$, la relation est un peu plus complexe : on a $p_m^* \gamma(z) = p_n^* \pi_*(z)$.

3. lorsque $p(z) \in \mathbb{C}[z]$, on trouve également une définition différente pour que le polynôme para-conjugué soit aussi un polynôme appartenant à $\mathbb{C}[z]$, cf Annexe page 69.

**Définition 9.** *Soit $p$ un L-polynôme, on dira que $p$ est para-hermitien ssi il est égal à son para-conjugué, c'est à dire $p = \widetilde{p}$ ou encore $p(z) = p^*(\frac{1}{z^*})\ \forall z \in \mathbb{C}^*$.*

**Propriété 10.** *Si $p$ est un L-polynôme para-hermitien, alors $p(z) = \sum_{i=-d}^{d} p_i z^i$ avec $p_{-i} = p_i^*$, et le L-degré de $p$ est pair.*

Soit $p(z) = \sum_{i=m}^{n} p_i z^i$, $p_m$ et $p_n$ non nuls, alors $\widetilde{p}(z) = \sum_{i=m}^{n} p_i^* z^{-i} = \sum_{i=-n}^{-m} p_{-i}^* z^i$, d'où pour avoir égalité des polynômes $n = -m$ et $p_{-i} = p_i^*$ pour $i = 1$ à $n$ et $p_0 = p_0^*$. Il suffit ensuite d'écrire $p(z) = z^{-d} \sum_{i=0}^{2d} p_{i-d} z^i$ pour avoir le L-degré égal à $2d$. □

Les polynômes para-hermitiens vérifient également la propriété suivante, également connue sous le nom de Théorème de Fejér-Riesz :

**Propriété 11.** *Soit $p(z)$ un L-polynôme para-hermitien strictement positif sur le cercle alors on peut factoriser [75] :*

$$p(z) = l(z)\widetilde{l}(z) \text{ avec } l(z) \in \mathbb{C}_s[z] \text{ à minimum de phase strict.}$$

Remarquons tout d'abord qu'un polynôme para-hermitien a toujours une valeur réelle sur le cercle, puisque $\widetilde{p}(e^{\jmath\omega}) = p^*(e^{\jmath\omega})$. De plus, les zéros d'un polynôme para-hermitien sont symétriques par rapport au cercle unité : si $z_0$ est zéro de $p(z)$, alors $\frac{1}{z_0^*}$ est également zéro de $p(z)$. En effet, comme $p(z) = \widetilde{p}(z)$, on a $p(z_0) = 0 = p^*(\frac{1}{z_0^*})$, soit $p(\frac{1}{z_0^*}) = 0$. Supposons de plus que ce polynôme est strictement positif sur le cercle, c'est à dire $p(e^{\jmath\omega}) > 0\ \forall\omega$, ce qui implique, en particulier, que $p(z)$ n'a pas de racine sur le cercle. Posons $p(z) = \alpha \prod_{i=1}^{d} (1 - a_i^{-1} z)(1 - a_i^{-1*} z^{-1})$, où les $a_i$ sont les zéros à l'intérieur du cercle unité. $p(e^{\jmath\omega}) = \alpha \prod_{i=1}^{d} |1 - a_i^{-1} e^{\jmath\omega}|^2 > 0$ implique que $\alpha$ est un nombre réel positif. En posant $l(z) = \sqrt{\alpha} \prod_{i=1}^{d} (1 - a_i^{-1} z)$, on a immédiatement la factorisation désirée. □

**Définition 12.** *Soit $p$ un L-polynôme, on dira que $p$ est para-unitaire ssi $p(z)\widetilde{p}(z) = 1\ \forall z \in \mathbb{C}^*$. On a $p(z) = e^{\jmath\theta} z^\alpha, \theta \in \mathbb{R}, \alpha \in \mathbb{Z}$.*

D'après leur définition, les L-polynômes para-unitaires sont des éléments inversibles de l'anneau des L-polynômes, ce sont donc des monômes. Pour que $p(z)\widetilde{p}(z) = 1$, il faut de plus que la constante du monôme soit de module 1, soit, $p(z)$ est para-unitaire ssi $p(z) = e^{\jmath\theta} z^\alpha, \theta \in \mathbb{R}, \alpha \in \mathbb{Z}$. □

**Propriété 13.** *L'algèbre des polynômes de Laurent sur le cercle unité et à valeurs dans $\mathbb{C}$ est dense dans l'ensemble des fonctions continues sur le cercle à valeurs dans $\mathbb{C}$.*

Soit $\mathcal{A}$ l'algèbre des polynômes de Laurent sur le cercle unité $\mathcal{C}$ à valeurs dans $\mathbb{C}$ :

$$\mathcal{A} = \{p : \mathcal{C} \to \mathbb{C} \text{ tq } p(z) = \sum_{i=m}^{n} p_i z^i \text{ avec } m, n \in \mathbb{Z}, p_i \in \mathbb{C} \text{ et } |z| = 1\}$$

Sur le cercle, $\frac{1}{z^*} = z$ et donc si $p(z) \in \mathcal{A}, \widetilde{p}(z) \in \mathcal{A}$, soit $\mathcal{A}$ est auto-adjoint. De plus, $\mathcal{A}$ sépare les points du cercle, c'est à dire que $\forall x, y \in \mathcal{C}$ il existe $p(z) \in \mathcal{A}$ tel que $p(x) \neq p(y)$. Enfin, le cercle unité est un espace métrique compact. D'après le théorème de Stone-Weierstrass, on peut donc en déduire que $\mathcal{A}$ est dense dans l'ensemble des fonctions continues de $\mathcal{C}$ à valeurs dans $\mathbb{C}$ pour la norme infinie $||f|| = \sup_{x \in \mathcal{C}} |f(x)|$. □

Toute fonction continue et $2\pi$-périodique sur $\mathbb{R}$ à valeurs dans $\mathbb{C}$ peut donc être approchée par des polynômes trigonométriques.

## 3.2 Matrices polynomiales de Laurent

L'anneau des polynômes de Laurent $(\mathbb{C}[z, z^{-1}], +, .)$ n'est pas un corps car tous ses éléments ne sont pas inversibles donc l'ensemble des tableaux à $n$ lignes et une colonne dont les entrées (composantes) sont des éléments de $\mathbb{C}[z, z^{-1}]$ n'est pas un espace vectoriel mais un module sur $\mathbb{C}[z, z^{-1}]$. Cet ensemble est noté $\mathbb{C}^n[z, z^{-1}]$ et ses éléments $\underline{v}(z)$. Par abus de langage, et pour simplifier la lecture, nous appellerons toutefois ces éléments des vecteurs L-polynomiaux.

Remarquons enfin que l'ensemble des tableaux $\underline{v}(z) \in \mathbb{C}^n[z, z^{-1}]$ dont les composantes sont des L-polynômes para-hermitiens est un-sous groupe de $\mathbb{C}^n[z, z^{-1}]$ mais pas un sous-module (car le produit d'un L-polynôme quelconque par un L-polynôme para-hermitien n'est pas para-hermitien).

L'ensemble des matrices $n \times n$ dont les entrées sont des L-polynômes est un anneau noté $\mathbb{C}^{n\times n}[z, z^{-1}]$. Ses éléments sont appelés des matrices polynomiales de Laurent ou encore matrices L-polynomiales. $\mathbb{C}^{n\times n}[z]$ est quant à lui l'anneau des matrices dont les composantes sont des éléments de $\mathbb{C}[z]$, c'est à dire des matrice polynomiales "classiques".

**Définition 14.** *Les éléments inversibles de l'anneau des matrices L-polynomiales $\mathbb{C}^{n\times n}[z, z^{-1}]$ sont appelés matrices L-unimodulaires, ce sont des matrices dont le déterminant est un élément de $\mathcal{U}$, c'est à dire un monôme non nul.*

Comme pour les matrices polynomiales, on peut définir des opérations élémentaires sur les matrices polynomiales de Laurent : multiplier une ligne (resp. une colonne) par un L-polynôme non nul, échanger deux lignes (resp. colonnes), ajouter à une ligne (resp. colonne) le produit d'une autre ligne (resp. colonne) par un L-polynôme. Ces opérations élémentaires sur les lignes correspondent à des multiplications à gauche par une matrice L-unimodulaire et celles sur les colonnes à une multiplication à droite par une matrice L-unimodulaire.

### 3.2.1 Forme de Smith des matrices polynomiales de Laurent

Comme l'anneau des polynômes de Laurent est un anneau principal, on peut définir la forme de Smith normale [36] d'une matrice à coefficients dans $\mathbb{C}[z, z^{-1}]$.

**Définition 15.** *Soit $\boldsymbol{M}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$, alors il existe $\boldsymbol{U}_1(z)$ et $\boldsymbol{U}_2(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ L-unimodulaires telles que*

$$\boldsymbol{U}_1(z)\boldsymbol{M}(z)\boldsymbol{U}_2(z) = \boldsymbol{\Lambda}(z)$$

$$\boldsymbol{\Lambda}(z) = \begin{bmatrix} \lambda_1(z) & & & \\ & \ddots & & 0 \\ & & \lambda_r(z) & \\ & 0 & & 0 \end{bmatrix}$$

- $\lambda_i(z)$ uniques à une multiplication par un monôme près
- $\lambda_i(z)$ *divise* $\lambda_{i+1}(z)$.
- $r$ *est le rang (normal) de la matrice* $\boldsymbol{M}(z)$.

$\boldsymbol{\Lambda}(z)$ *est appelée forme normale de Smith de* $\boldsymbol{M}(z)$, *les* $\lambda_i(z)$ *sont appelés facteurs invariants ou diviseurs élémentaires.*

Afin de définir ces facteurs invariants de manière unique, on va leur imposer, comme pour les pgcd, d'être L-unitaires.

**Définition 16.** *Soit* $\boldsymbol{M}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$. *On appelle L-forme de Smith de* $\boldsymbol{M}(z)$ *la matrice polynomiale diagonale* $\boldsymbol{\Lambda}(z) \in \mathbb{C}^{n\times n}[z]$ *telle que*

$$\boldsymbol{U}_1(z)\boldsymbol{M}(z)\boldsymbol{U}_2(z) = \boldsymbol{\Lambda}(z)$$

*où* $\boldsymbol{U}_1(z)$ *et* $\boldsymbol{U}_2(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ *sont L-unimodulaires et*

$$\boldsymbol{\Lambda}(z) = \begin{bmatrix} \mu_1(z) & & & \\ & \ddots & & 0 \\ & & \mu_r(z) & \\ & 0 & & 0 \end{bmatrix}$$

*avec* $\mu_i(z)$ *divise* $\mu_{i+1}(z)$ *et est L-unitaire.*
*Les* $\mu_i(z) \in \mathbb{C}[z]$ *sont appelés L-polynômes invariants de* $\boldsymbol{M}(z)$.

Sur l'anneau des matrices polynomiales "classiques" $\mathbb{C}^{n\times n}[z]$, il existe une forme de Smith bien connue [54] qui fait intervenir des matrices unimodulaires dont le déterminant est une constante non nulle. Dans le cas d'une matrice carrée de rang plein, on a le résultat suivant :

Soit $\boldsymbol{M}(z) \in \mathbb{C}^{n\times n}[z]$, alors $\exists \boldsymbol{V}_1(z)$ et $\boldsymbol{V}_2(z) \in \mathbb{C}^{n\times n}[z]$ unimodulaires telles que

$$\boldsymbol{V}_1(z)\boldsymbol{M}(z)\boldsymbol{V}_2(z) = \mathrm{diag}\{\lambda_i(z)\}$$

les $\lambda_i(z)$ sont les polynômes invariants de $\boldsymbol{M}(z)$, ils sont uniques si on les suppose normalisés (monic) et vérifient $\lambda_i(z)$ divise $\lambda_{i+1}(z)$. La matrice $\mathrm{diag}\{\lambda_i(z)\}$ est appelée forme de Smith de $\boldsymbol{M}(z)$.

Il est facile de montrer que si $\boldsymbol{M}(z) \in \mathbb{C}^{n\times n}[z]$ les L-polynômes invariants coïncident avec les polynômes invariants lorsque zéro n'est pas racine de ces polynômes.

**Propriété 17.** *Si $\boldsymbol{M}(z)$ est une matrice L-polynomiale de rang plein, alors son déterminant est à une multiplication par un monôme près, le produit de ses L-polynômes invariants.*

Par définition de la L-forme de Smith de $\boldsymbol{M}(z)$, on a

$$\det \boldsymbol{\Lambda}(z) = \prod_i \mu_i(z) = \det \boldsymbol{U}_1(z) \det \boldsymbol{M}(z) \det \boldsymbol{U}_2(z)$$

avec $\det \boldsymbol{U}_i(z)$ qui sont des monômes car les $\boldsymbol{U}_i$ sont L-unimodulaires. □

On peut définir l'équivalence au sens de Smith de deux matrices L-polynomiales :

**Définition 18.** *Soient $\boldsymbol{M}(z)$ et $\boldsymbol{P}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$, alors $\boldsymbol{M}(z) \overset{LS}{\sim} \boldsymbol{P}(z)$, ssi elles ont même L-forme de Smith, c'est à dire ssi il existe $\boldsymbol{U}_1(z)$ et $\boldsymbol{U}_2(z)$ L-unimodulaires telles que*

$$\boldsymbol{M}(z) = \boldsymbol{U}_1(z)\boldsymbol{P}(z)\boldsymbol{U}_2(z)$$

Pour éviter la confusion avec l'équivalence au sens de Smith "classique" (ie à des matrices unimodulaires près), on appellera cette équivalence, l'équivalence au sens de L-Smith ou L-équivalence au sens de Smith afin de rappeler que les égalités ont lieu à des matrices polynomiales de Laurent L-unimodulaires près.

**Propriété 19.** *Les matrices L-unimodulaires sont L-équivalentes au sens de Smith à la matrice identité.*

En effet, le déterminant d'une matrice L-unimodulaire est égal à un monôme, mais aussi au produit de ses L-polynômes invariants et d'un monôme. Donc, ces L-polynômes sont des monômes, mais par définition, les L-polynômes invariants sont L-unitaires, ils ne peuvent donc être égaux qu'à $1 : \mu_i(z) = 1, \forall z$. □

**Propriété 20.** *Si deux matrices sont L-équivalentes au sens de Smith, alors les pgcd des mineurs d'ordre $j$ de ces matrices sont égaux.*

Soient $\boldsymbol{A}(z)$ et $\boldsymbol{B}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ telles que $\boldsymbol{A}(z) \overset{LS}{\sim} \boldsymbol{B}(z)$, il existe $\boldsymbol{U}(z)$ et $\boldsymbol{V}(z)$ L-unimodulaires telles que $\boldsymbol{B}(z) = \boldsymbol{U}(z)\boldsymbol{A}(z)\boldsymbol{V}(z)$. Soit $b(z)$ un mineur d'ordre j de $\boldsymbol{B}(z)$, d'après le formule de Binet-Cauchy, on a :

$$b(z) = \sum_i u_i(z)a_i(z)v_i(z) \tag{3.2}$$

où $u_i, a_i, v_i$ sont des mineurs d'ordre $j$ de $\boldsymbol{U}$, $\boldsymbol{A}$ et $\boldsymbol{V}$.

Soit $\Delta_j(z)$ le pgcd de tous les mineurs d'ordre $j$ de $\boldsymbol{A}$. D'après la relation (3.2), $\Delta_j(z)$ divise $b(z)$ et ceci quelque soit le mineur d'ordre $j$ de $\boldsymbol{B}(z)$, donc $\Delta_j(z)$ divise le pgcd des mineurs d'ordre $j$ de $\boldsymbol{B}(z)$ que l'on va noter $\delta_j(z)$. En écrivant maintenant $\boldsymbol{A}(z) = \boldsymbol{U}^{-1}(z)\boldsymbol{B}(z)\boldsymbol{V}^{-1}(z)$, par le même raisonnement, on peut en déduire que $\delta_j(z)$ divise $\Delta_j(z)$. Ces polynômes étant L-unitaires, on peut en déduire qu'ils sont égaux. □

**Propriété 21.** *Soit $\boldsymbol{M}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ une matrice L-polynomiale, soient $\mu_i(z)$ ses L-polynômes invariants. Soit $\Delta_j(z)$ le pgcd des mineurs d'ordre $j$ de $\boldsymbol{M}(z)$, alors*

$$\mu_j(z) = \frac{\Delta_j(z)}{\Delta_{j-1}(z)}$$

D'après la propriété 20, le pgcd des mineurs d'ordre $j$ de $\boldsymbol{M}(z)$ est égal au pgcd des mineurs d'ordre $j$ de sa L-forme de Smith $\boldsymbol{\Lambda}(z)$. $\boldsymbol{\Lambda}(z)$ étant une matrice diagonale telle que $\mu_i(z)$ divise $\mu_{i+1}(z)$, le pgcd des mineurs d'ordre $j$ de $\boldsymbol{\Lambda}(z)$ est $\prod_{i=1}^{j}\mu_i(z)$, soit $\Delta_j(z) = \prod_{i=1}^{j}\mu_i(z)$ et de même, $\Delta_{j-1}(z) = \prod_{i=1}^{j-1}\mu_i(z)$ d'où $\Delta_j(z) = \Delta_{j-1}(z)\mu_j(z)$. Là encore, tous les polynômes considérés sont L-unitaires et n'admettent donc aucun zéro en zéro ou l'infini. □

**Exemple**

Soient $\boldsymbol{A}(z) = \begin{bmatrix} z & 0 \\ 0 & z \end{bmatrix}$, $\boldsymbol{B}(z) = \begin{bmatrix} z & 0 \\ z & z \end{bmatrix}$ et $\boldsymbol{C}(z) = \begin{bmatrix} z & 0 \\ z-1 & z \end{bmatrix}$ 3 matrices L-unimodulaires, alors on a $\boldsymbol{A}(z) \overset{S}{\sim} \boldsymbol{B}(z) \overset{S}{\sim} zI$ mais $\boldsymbol{C}(z) \overset{S}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & z^2 \end{bmatrix}$ ($z = 0$ ne fait pas perdre de rang de la même manière pour $\boldsymbol{B}(z)$ et $\boldsymbol{C}(z)$) mais $\boldsymbol{A}(z) \overset{LS}{\sim} \boldsymbol{B}(z) \overset{LS}{\sim} \boldsymbol{C}(z) \overset{LS}{\sim} \boldsymbol{I}$ puisqu'elles sont L-unimodulaires !
On peut vérifier que $\boldsymbol{U}_1(z)\boldsymbol{C}(z)\boldsymbol{U}_2(z) = \boldsymbol{I}$ avec $\boldsymbol{U}_1(z) = \begin{bmatrix} 1 & -1 \\ -z^{-1}+z^{-2} & z^{-1} \end{bmatrix}$ et $\boldsymbol{U}_2(z) = \begin{bmatrix} 1 & z \\ 0 & 1 \end{bmatrix}$ par exemple.

Remarquons que la somme des degrés des L-polynômes invariants de $\boldsymbol{C}(z)$ est égale à 0. En effet, la somme des degrés des L-polynômes invariants comptabilise tous les zéros sauf ceux qui sont en zéro ou l'infini. Le propre des matrices L-unimodulaires étant de n'avoir que des zéros en zéro ou l'infini, il est donc logique que cette somme vaille zéro pour $\boldsymbol{C}(z)$. Précisons enfin que pour l'application considérée (cf page 52) on déconvolue à une matrice de permutation et une matrice diagonale de retards près, les retards sont des zéros en l'infini, qui n'apparaissent donc pas dans les L-polynômes invariants, et n'influent donc pas sur la somme de ces degrés.

### 3.2.2 Ordre d'une matrice polynomiale de Laurent

Avant de définir l'ordre d'une matrice L-polynomiale, nous allons brièvement rappeler ce qu'est la forme de Smith-McMillan d'une matrice rationnelle carrée de rang plein, sachant que cette forme existe pour toute matrice rationnelle [54]. Soit $\mathbb{C}^{n\times n}(z)$ l'anneau des matrices rationnelles, c'est-à-dire des matrices dont les entrées sont des fractions rationnelles en l'indéterminée $z$.

Soit $\boldsymbol{R}(z) \in \mathbb{C}^{n\times n}(z)$ une matrice rationnelle propre de rang plein dont les coefficients $r_{ij}(z)$ sont des fractions rationnelles propres sous forme réduite (pas de simplification pôle-zéro), soit $d(z)$ le ppcm normalisé des dénominateurs des $r_{ij}(z)$. Posons $\boldsymbol{N}(z) = \boldsymbol{R}(z)d(z)$. Par définition de $d(z)$, $\boldsymbol{N}(z) \in \mathbb{C}^{n\times n}[z]$ et admet donc une forme de Smith. Soit $\boldsymbol{\Lambda}(z)$ la forme de Smith de $\boldsymbol{N}(z)$, il existe $\boldsymbol{U}_1(z)$ et $\boldsymbol{U}_2(z) \in \mathbb{C}^{n\times n}[z]$ unimodulaires telles que

$$\boldsymbol{U}_1(z)\boldsymbol{N}(z)\boldsymbol{U}_2(z) = \boldsymbol{\Lambda}(z)$$

d'où

$$\boldsymbol{U}_1(z)\boldsymbol{R}(z)\boldsymbol{U}_2(z) = \frac{\boldsymbol{\Lambda}(z)}{d(z)} = \text{ diag}\left\{\frac{\lambda_i(z)}{d(z)}\right\}$$

Posons $\dfrac{\lambda_i(z)}{d(z)} = \dfrac{\varepsilon_i(z)}{\psi_i(z)}$ avec $\varepsilon_i(z)$ et $\psi_i(z)$ premiers entre eux (et normalisés). Alors

$$\boldsymbol{U}_1(z)\boldsymbol{R}(z)\boldsymbol{U}_2(z) = \boldsymbol{M}(z) = \text{diag}\left\{\frac{\varepsilon_i(z)}{\psi_i(z)}\right\} \tag{3.3}$$

$\boldsymbol{M}(z)$ est appelé forme de Smith-McMillan de $\boldsymbol{R}(z)$.

Les polynômes $\varepsilon_i(z)$ et $\psi_i(z)$ sont uniques et sont tels que $\varepsilon_i(z)$ divise $\varepsilon_{i+1}(z)$, $\psi_{i+1}(z)$ divise $\psi_i(z)$ (car $\lambda_i(z)$ divise $\lambda_{i+1}(z)$) et $\psi_1(z) = d(z)$. Les racines de $\varepsilon_i(z)$ et de $\psi_i(z)$ sont respectivement les zéros et les pôles de la matrice $\boldsymbol{R}(z)$.

**Définition 22.** *Soit $\boldsymbol{R}(z) \in \mathbb{C}^{n\times n}(z)$ une matrice rationnelle carrée de rang plein. On appelle degré de McMillan de $\boldsymbol{R}(z)$*

$$N = \sum_{i=1}^{n} deg(\psi_i(z))$$

*le degré de McMillan d'une matrice rationnelle correspond donc au nombre de ses pôles (avec multiplicité).*

**Exemple**

Soit $\boldsymbol{R}(z) = \begin{bmatrix} 1 & 0 \\ 1 - z^{-1} & 1 \end{bmatrix}$. Posons $\boldsymbol{N}(z) = z\boldsymbol{R}(z) = \begin{bmatrix} z & 0 \\ z - 1 & z \end{bmatrix}$, la forme de Smith de $\boldsymbol{N}(z)$, obtenue à l'aide des pgcd des mineurs, est $\begin{bmatrix} 1 & 0 \\ 0 & z^2 \end{bmatrix}$ et la forme de Smith McMillan de $\boldsymbol{R}(z)$ est donc $\boldsymbol{M}(z) = \begin{bmatrix} \frac{1}{z} & 0 \\ 0 & z \end{bmatrix}$, soit $\varepsilon_1(z) = 1$, $\varepsilon_2(z) = z$, $\psi_1(z) = z$, $\psi_2(z) = 1$. Le degré de McMillan est donc $N = 1$.

Remarquons que $\boldsymbol{R}(z)$ possède des zéros et des pôles en zéro, et que la forme de Smith-McMillan de $\boldsymbol{R}(z)$ n'est pas causale alors que $\boldsymbol{R}(z)$ l'est. □

Comme on a défini un L-degré pour les L-polynômes, il peut être intéressant de définir un degré ou ordre pour une matrice L-polynomiale car ceci n'est pas toujours clairement défini dans la littérature concernant l'égalisation. Rappelons que pour les matrices polynomiales et rationnelles classiques on distingue [54, 80] :

i. l'ordre d'une matrice polynomiale :
Si $\boldsymbol{Q}(z) \in \mathbb{C}^{n\times n}[z]$ avec $\boldsymbol{Q}(z) = \sum_{k=0}^{l} \boldsymbol{Q}_k z^k$, alors on appelle ordre de $\boldsymbol{Q}$ le plus grand indice $k$ tel que $\boldsymbol{Q}_k$ est différent de la matrice nulle.
Si $\boldsymbol{Q}(z) \in \mathbb{C}^{n\times n}[z^{-1}]$ avec $\boldsymbol{Q}(z) = \sum_{k=0}^{l} \boldsymbol{Q}_k z^{-k}$, alors on appelle ordre de $\boldsymbol{Q}$ le plus grand indice $k$ tel que $\boldsymbol{Q}_k$ est différent de la matrice nulle.
$l + 1$ est parfois appelé longueur de $\boldsymbol{Q}$.

ii. le terme de degré est réservé aux matrices rationnelles et est alors défini comme le degré de McMillan de cette matrice.

Dans le cas d'une matrice rationnelle propre, on montre que ce degré correspond à la dimension minimale d'une réalisation d'état associée. Dans le cas des systèmes à temps discret, le degré d'une matrice rationnelle causale $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}(z)$ correspond donc au nombre minimum de retards nécessaires pour implémenter le système ayant pour matrice de transfert $\boldsymbol{H}(z)$.

Si on dispose d'une factorisation polynomiale à gauche irréductible de $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}(z)$ : $\boldsymbol{H}(z) = \boldsymbol{Q}^{-1}(z)\boldsymbol{P}(z)$ avec $\boldsymbol{P}(z)$ et $\boldsymbol{Q}(z) \in \mathbb{C}^{n\times n}[z]$ des matrices polynomiales premières entre elles à gauche, alors le degré du déterminant de $\boldsymbol{Q}(z)$ (dénominateur de la factorisation irréductible) est égal au degré de la matrice $\boldsymbol{H}(z)$, ce qui se comprend car on peut déduire aisément de la forme de Smith-McMillan une factorisation construite à partir des matrices unimodulaires et des $\varepsilon_i/\psi_i$ définis dans (3.3) [54]. Il est à noter que $\det \boldsymbol{Q}(z) \in \mathbb{C}[z]$ et que les matrices $\boldsymbol{P}(z)$ et $\boldsymbol{Q}(z)$ ne doivent pas être exprimées comme des matrices de la variable $z^{-1}$.

En définissant le degré d'une fraction rationnelle irréductible propre comme étant le degré de son dénominateur on obtient pour une matrice rationnelle de rang plein [80] :

$$\deg(\det \boldsymbol{H}(z)) \leq N$$

En effet, en tenant compte de (3.3), avec $\boldsymbol{U}_i(z)$ unimodulaires donc de déterminant constant non nul, on obtient :
$\deg(\det \boldsymbol{H}(z)) \leq \sum_{i=1}^{n} \deg(\psi_i(z)) = N$.
Dans le cas d'un système à réponse impulsionnelle finie (FIR), $\boldsymbol{H}(z) = \sum_{k=0}^{l} \boldsymbol{H}_k z^{-k}$, on a donc besoin d'au moins $l$ retards pour implémenter $\boldsymbol{H}(z)$, soit

$$l \leq N$$

De plus, $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}[z^{-1}]$, mais peut également être vue comme un élément de $\mathbb{C}^{n\times n}(z)$ dont tous les pôles sont en zéro (ou encore dont tous les zéros sont en l'infini), les dénominateurs de la forme de Smith -McMillan sont donc de la forme $\psi_i(z) = z^{n_i}, n_i \in \mathbb{N}$ et $n_{i+1} \leq n_i$ et $N = \sum_i n_i$.

Considérons maintenant le cas des matrices L-polynomiales.

**Définition 23.** *Soit $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$, $\boldsymbol{H}(z) = \sum_{k=m}^{p} \boldsymbol{H}_k z^k$ avec $\boldsymbol{H}_k \in \mathbb{C}^{n\times n}, m, p \in \mathbb{Z}, m \leq p, \boldsymbol{H}_m$ et $\boldsymbol{H}_p$ non égaux à la matrice nulle. Posons*

$$\boldsymbol{H}(z) = z^m \sum_{i=0}^{p-m} \Xi_i z^i = z^m \boldsymbol{\Xi}(z)$$

*avec $\boldsymbol{\Xi}(z) \in \mathbb{C}^{n\times n}[z]$. On appelle L-ordre de la matrice $\boldsymbol{H}$ l'ordre de la matrice (polynomiale) associée $\boldsymbol{\Xi}(z)$, soit ici $p - m$.*

Comme pour la définition du L-degré d'un polynôme, il faut distinguer le L-ordre de l'ordre pour une matrice polynomiale : ceux-ci ne coïncident pas si on peut mettre $z^\alpha$ ($\alpha \in \mathbb{Z}$) en facteur de la matrice. La définition du L-ordre correspond à celle de l'ordre utilisée dans [44] et plus généralement dans les communications numériques où l'on travaille souvent à des retards près. Donc, sauf en cas de confusion possible (ie lorsqu'on manipule des matrices polynomiales dont le coefficient matriciel du terme de degré zéro est nul) et afin d'alléger un peu la lecture, on utilisera le terme ordre pour une matrice L-polynomiale.

Remarquons que la L-forme de Smith de $\boldsymbol{H}(z)$ est reliée à la forme de Smith de $\boldsymbol{\Xi}(z)$ : si on L-normalise les polynômes invariants de $\boldsymbol{\Xi}(z)$ (c'est-à-dire si on supprime les zéros en zéro puisqu'ils sont déjà monic), on retrouve les L-polynômes invariants de $\boldsymbol{H}(z)$.

**Exemples**

Revenons à l'exemple précédent : $\boldsymbol{R}(z) = \begin{bmatrix} 1 & 0 \\ 1-z^{-1} & 1 \end{bmatrix} = \boldsymbol{R}_0 + z^{-1}\boldsymbol{R}_1$ $(m = 0, p = 1)$, son ordre est $1(= p)$, son L-ordre est $1(= p-m)$ et son degré de McMillan $N = 1$.

Considérons $\boldsymbol{A}(z) = \begin{bmatrix} z^{-1} & 0 \\ 0 & z^{-1} \end{bmatrix} = \boldsymbol{0} + z^{-1}\boldsymbol{I}$. L'ordre de $\boldsymbol{A}(z)$ est 1, son L-ordre est 0. Sa forme de Smith-McMillan est $\boldsymbol{M}(z) = \boldsymbol{A}(z) = z^{-1}\boldsymbol{I}$ et son degré de MacMillan est donc $N = 2$. □

La notion de degré d'une matrice n'étant définie que pour des matrices causales, on va associer à chaque matrice L-polynomiale une matrice causale (FIR).

**Définition 24.** *Soit* $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}], \boldsymbol{H}(z) = \sum_{k=m}^{p} \boldsymbol{H}_k z^k$ *avec* $m, p \in \mathbb{Z}, m \leq p, \boldsymbol{H}_m$ *et* $\boldsymbol{H}_p$ *non égaux à la matrice nulle. Posons*

$$\boldsymbol{H}(z) = z^p \sum_{i=0}^{p-m} \boldsymbol{H}_{p-i} z^{-i} = z^p \overline{\overline{\boldsymbol{H}}}(z)$$

*avec* $\overline{\overline{\boldsymbol{H}}}(z) \in \mathbb{C}^{n\times n}[z^{-1}]$ *donc causale. On définit le L-degré de* $\boldsymbol{H}(z)$ *comme le degré de McMillan de la matrice causale associée* $\overline{\overline{\boldsymbol{H}}}(z)$.

Le L-degré ainsi défini correspond au nombre minimum de retards (ou la dimension minimale de l'état associé au système), nécessaires pour implémenter le système "décalé" c'est à dire le système causal associé.

Par définition, $\boldsymbol{\Xi}(z) = z^{p-m}\overline{\overline{\boldsymbol{H}}}(z)$. La forme de Smith-McMillan de $\overline{\overline{\boldsymbol{H}}}(z)$ est donc $\frac{1}{z^{p-m}}\boldsymbol{S}(z)$ où $\boldsymbol{S}(z)$ est la forme de Smith de $\boldsymbol{\Xi}(z)$.

**Exemple**

Soit $\boldsymbol{C}(z) = \begin{bmatrix} z & 0 \\ z-1 & z \end{bmatrix} = \boldsymbol{C}_0 + \boldsymbol{C}_1 z$, son ordre est 1. On a $\boldsymbol{C}(z) = z\boldsymbol{R}(z)$ où $\boldsymbol{R}(z)$ est celui de l'exemple page 29, et donc le L-degré de $\boldsymbol{C}(z)$ est 1.

Soit $\boldsymbol{H}(z) = \begin{bmatrix} 1 & 1 \\ 1 & z^{-1}+z \end{bmatrix} = \boldsymbol{H}_{-1}z^{-1} + \boldsymbol{H}_0 + \boldsymbol{H}_1 z$, son ordre est 2. $\overline{\overline{\boldsymbol{H}}}(z) = z^{-1}\boldsymbol{H}(z) = \begin{bmatrix} z^{-1} & z^{-1} \\ z^{-1} & 1+z^{-2} \end{bmatrix} = z^{-2}\begin{bmatrix} z & z \\ z & 1+z^2 \end{bmatrix}$. La forme de Smith-McMillan de $\overline{\overline{\boldsymbol{H}}}(z)$ est $z^{-2}\begin{bmatrix} 1 & 0 \\ 0 & z(1+z^2)-z^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{z^2} & 0 \\ 0 & \frac{1-z+z^2}{z} \end{bmatrix}$, le L-degré de $\boldsymbol{H}(z)$ est donc 3 (et est supérieur à l'ordre).

### 3.2.3 Para-conjugaison et forme hermitienne

Soit $f$ une fonction, alors on définit la para-conjuguée de $f$ et on note $\sigma(f) = \widetilde{f}$ la fonction telle que $\widetilde{f}(x) = f^H(\frac{1}{x^*})$ où $^H$ désigne la transposée conjuguée et $^*$ la conjugaison. On a de manière évidente $\sigma(\sigma(f)) = f$, c'est à dire que la para-conjugaison est une involution.

Notons que si $\underline{v} \in \mathbb{C}^{n\times 1}[z, z^{-1}]$, alors $\widetilde{\underline{v}} \in \mathbb{C}^{1\times n}[z, z^{-1}]$ car on conjugue les coefficients des L-polynômes mais pas la variable $z$. Il en va de même pour les matrices L-polynomiales. Pour toute matrice $\boldsymbol{M}(z) \in \mathbb{C}^{n\times m}[z, z^{-1}]$, on note $\widetilde{\boldsymbol{M}}(z)$ sa matrice para-conjuguée définie par $\widetilde{\boldsymbol{M}}(z) = \boldsymbol{M}^H(\frac{1}{z^*}), \forall z \in \mathbb{C}^*$. On a donc $\widetilde{\boldsymbol{M}}(z) \in \mathbb{C}^{m\times n}[z, z^{-1}]$.

On peut associer à cette involution une forme sesqui-linéaire sur $\mathbb{C}^{n\times 1}[z, z^{-1}]$, qui n'est pas un produit scalaire au sens usuel puisque les "scalaires" n'appartiennent pas à un corps mais à un module.

**Définition 25.** *Soient $\underline{u}, \underline{v} \in \mathbb{C}^{n\times 1}[z, z^{-1}]$ alors on définit le produit de $\underline{u}$ par $\underline{v}$, $\langle \underline{u}, \underline{v}\rangle \in \mathbb{C}[z, z^{-1}]$ :*

$$\langle \underline{u}(z), \underline{v}(z)\rangle = \widetilde{\underline{u}}(z)\underline{v}(z) = \underline{u}^H(\frac{1}{z^*})\underline{v}(z), \forall z \in \mathbb{C}^*$$

**Propriété 26.** *$\langle ., .\rangle$ est une application linéaire à droite, sesqui-linéaire à gauche et hermitienne relativement à la para-conjugaison.*

Soient $\lambda \in \mathbb{C}[z, z^{-1}]$ et $\underline{u}, \underline{u}_1, \underline{u}_2, \underline{v}, \underline{v}_1, \underline{v}_2 \in \mathbb{C}^{n\times 1}[z, z^{-1}]$ :

- $\langle \underline{u}(z), \underline{v}_1(z) + \lambda(z)\underline{v}_2(z)\rangle = \langle \underline{u}(z), \underline{v}_1(z)\rangle + \lambda(z)\langle \underline{u}(z), \underline{v}_2(z)\rangle$
- $\langle \underline{u}_1(z) + \lambda(z)\underline{u}_2(z), \underline{v}(z)\rangle = \langle \underline{u}_1(z), \underline{v}(z)\rangle + \widetilde{\lambda}(z)\langle \underline{u}_2(z), \underline{v}(z)\rangle$
- $\widetilde{\langle \underline{u}(z), \underline{v}(z)\rangle} = \underline{u}^H\widetilde{(\frac{1}{z^*})\underline{v}(z)} = \underline{v}^H(\frac{1}{z^*})\underline{u}(z) = \langle \underline{v}(z), \underline{u}(z)\rangle$, soit le produit est hermitien relativement à la para-conjugaison.

On déduit de cette dernière propriété que $p(z) = \langle \underline{u}(z), \underline{u}(z)\rangle$ est un L-polynôme égal à son para-conjugué, donc qu'il est est para-hermitien.

Si on décompose $\underline{u}(z) = \sum_i u_i(z)\underline{e}_i$ où $\underline{e}_i$ est la base canonique classique de $\mathbb{C}^n$, alors $\langle \underline{u}(z), \underline{u}(z)\rangle = \sum_i u_i(z)u_i^*(\frac{1}{z^*})$. Le produit d'un vecteur L-polynomial par lui même est donc un L-polynôme. Ce produit ne peut donc pas induire une norme car la norme d'un vecteur est un nombre réel positif.

Examinons maintenant ce qu'il en est lorsqu'on se restreint au cercle unité, c'est à dire on pose $z = e^{\jmath\omega}, \omega \in \mathbb{R}$. Dans ce cas $\frac{1}{z^*} = z$ et $\widetilde{u}(e^{\jmath\omega}) = u^*(e^{\jmath\omega})$ et ainsi $\langle \underline{u}(e^{\jmath\omega}), \underline{u}(e^{\jmath\omega})\rangle = \sum_i u_i(e^{\jmath\omega})u_i^*(e^{\jmath\omega}) = \sum_i |u_i(e^{\jmath\omega})|^2 \geq 0\ \forall\omega \in \mathbb{R}$.

Posons $|| \underline{u}(e^{\jmath\omega})||_L^2 = \langle \underline{u}(e^{\jmath\omega}), \underline{u}(e^{\jmath\omega})\rangle$, alors
- $|| \underline{u}(e^{\jmath\omega})||_L = 0 \Rightarrow \underline{u}(e^{\jmath\omega}) = \underline{0}\ \forall\omega \in \mathbb{R}$.
- Soit $\lambda(z) \in \mathbb{C}[z, z^{-1}]$, alors $\langle \lambda(e^{\jmath\omega})\underline{u}(e^{\jmath\omega}), \lambda(e^{\jmath\omega})\underline{u}(e^{\jmath\omega})\rangle = \lambda(e^{\jmath\omega})\lambda^*(e^{\jmath\omega})\langle \underline{u}(e^{\jmath\omega}), \underline{u}(e^{\jmath\omega})\rangle$ soit $\langle \lambda(e^{\jmath\omega})\underline{u}(e^{\jmath\omega}), \lambda(e^{\jmath\omega})\underline{u}(e^{\jmath\omega})\rangle = |\lambda(e^{\jmath\omega})|^2 || \underline{u}(e^{\jmath\omega})||_L^2$.
- $\langle \underline{u}(e^{\jmath\omega})+\underline{v}(e^{\jmath\omega}), \underline{u}(e^{\jmath\omega})+\underline{v}(e^{\jmath\omega})\rangle = \sum_i |u_i(e^{\jmath\omega})+v_i(e^{\jmath\omega})|^2 \leq \sum_i |u_i(e^{\jmath\omega})|^2+|v_i(e^{\jmath\omega})|^2$ soit $\langle \underline{u}(e^{\jmath\omega}) + \underline{v}(e^{\jmath\omega}), \underline{u}(e^{\jmath\omega}) + \underline{v}(e^{\jmath\omega})\rangle \leq || \underline{u}(e^{\jmath\omega})||_L^2 + || \underline{v}(e^{\jmath\omega})||_L^2$.

Pour $\omega$ donné, $|| \underline{u}(e^{\jmath\omega})||_L$ est donc une norme.

**Propriété 27.** *Soit $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$, alors*

$$\langle \boldsymbol{H}(z)\underline{u}(z), \underline{v}(z)\rangle = \langle \underline{u}(z), \widetilde{\boldsymbol{H}}(z)\underline{v}(z)\rangle$$

Il suffit d'appliquer la propriété 26 :

$$\langle H(z)\underline{u}(z), \underline{v}(z)\rangle = \underline{u}^H(\frac{1}{z^*})\boldsymbol{H}^H(\frac{1}{z^*})\underline{v}(z) = \widetilde{\underline{u}}(z)\boldsymbol{H}^H(\frac{1}{z^*})\underline{v}(z) = \langle \underline{u}(z), \boldsymbol{H}^H(\frac{1}{z^*})\underline{v}(z)\rangle \quad \square$$

### 3.2.4 Matrice L-polynomiale para-hermitienne

La notion de matrice hermitienne sur $\mathbb{C}^{n\times n}$ s'étend aux matrices L-polynomiales, on appelle alors ces matrices des matrices para-hermitiennes [44].

**Définition 28.** *Soit $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ une matrice L-polynomiale, $\boldsymbol{H}(z)$ est para-hermitienne ssi*

$$\boldsymbol{H}(z) = \boldsymbol{H}^H(\frac{1}{z^*}), \forall z \in \mathbb{C}^*$$

Une matrice para-hermitienne n'est donc pas une matrice causale, sauf si elle est constante (et donc hermitienne), c'est pourquoi ces matrices ont été peu étudiées. Elle interviennent cependant naturellement lorsqu'on étudie les transformées en $z$ des matrices de corrélation (spectre), c'est pourquoi nous allons en présenter quelques propriétés originales. La première concerne les relations entre les différents coefficients (matriciels) de la matrice :

**Propriété 29.** *Si $\boldsymbol{H}(z)$ est L-polynomiale para-hermitienne, alors*

$$\boldsymbol{H}(z) = \sum_{k=-d}^{d} \boldsymbol{H}_k z^k \text{ avec } \boldsymbol{H}_0 \text{ hermitienne et } \boldsymbol{H}_{-k} = \boldsymbol{H}_k^H, \forall k$$

*et l'ordre de $\boldsymbol{H}(z)$ est pair.*

En effet, $\boldsymbol{H}(z) = \sum_{k=-d}^{d} \boldsymbol{H}_k z^k = \boldsymbol{H}^H(\frac{1}{z^*}) = \left(\sum_{k=-d}^{d} \boldsymbol{H}_k z^{-k*}\right)^H = \sum_{k=-d}^{d} \boldsymbol{H}_k^H z^{-k} = \sum_{k=d}^{-d} \boldsymbol{H}_{-k}^H z^k$.

Il suffit ensuite d'écrire $\boldsymbol{H}(z) = z^{-d} \sum_{k=0}^{2d} \boldsymbol{H}_{k-d} z^k$, $\boldsymbol{H}_{-d}$ et $\boldsymbol{H}_d$ étant non nuls par définition de $d$, l'ordre de $\boldsymbol{H}(z)$ est $2d$. □

Remarquons que les seules matrices polynomiales (en $z$ ou $z^{-1}$) para-hermitiennes sont donc des matrices constantes hermitiennes : $\boldsymbol{H}(z) = \boldsymbol{H}_0 = \boldsymbol{H}_0^H$.

**Propriété 30.** *Les L-polynômes invariants d'une matrice L-polynomiale para-hermitienne sont des polynômes auto-inversifs c'est à dire qu'ils vérifient*

$$\mu_i(z) = e^{\jmath\theta_i} z^{m_i} \widetilde{\mu}_i(z)$$

*où $\theta_i \in \mathbb{R}$ et $m_i$ est le degré du $i^e$L-polynôme invariant.*

Soit $\boldsymbol{H} \in \mathbb{C}^{n\times n}[z, z^{-1}]$ une matrice para-hermitienne et soit $\boldsymbol{\Lambda}$ la L-forme de Smith de $\boldsymbol{H}(z)$, alors il existe $\boldsymbol{U}_1(z)$ et $\boldsymbol{U}_2(z) \in \mathbb{C}[z, z^{-1}]$ L-unimodulaires telles que $\boldsymbol{H}(z) = \boldsymbol{U}_1(z)\boldsymbol{\Lambda}(z)\boldsymbol{U}_2(z)$, d'où $\boldsymbol{H}^H(\frac{1}{z^*}) = \boldsymbol{U}_2^H(\frac{1}{z^*})\boldsymbol{\Lambda}^H(\frac{1}{z^*})\boldsymbol{U}_1^H(\frac{1}{z^*})$. Par définition, $\det \boldsymbol{U}_1(z) \in \mathcal{U}$ donc $\det \boldsymbol{U}_1^H(\frac{1}{z^*})$ est également un monôme, soit $\widetilde{\boldsymbol{U}_1}$ est L-unimodulaire. Il en est de même pour $\boldsymbol{U}_2$. On peut donc en déduire que la L-forme de Smith de $\widetilde{\boldsymbol{H}}(z)$ est L-équivalente à $\boldsymbol{\Lambda}^H(\frac{1}{z^*})$. Ces matrices ne sont cependant pas égales car l'une appartient à $\mathbb{C}^{n\times n}[z]$ et l'autre à $\mathbb{C}^{n\times n}[z^{-1}]$.

Soit donc $\boldsymbol{\Lambda}(z) = \mathrm{diag}\{\mu_i(z)\}$ avec $\mu_i(z)$ L-unitaire et $\mu_i(z)$ divise $\mu_{i+1}(z)$. Posons $\mu_i(z) = \sum_{k=0}^{m_i-1} \mu_{ik} z^k + z^{m_i}$ avec $\mu_{i0} \neq 0$ puisque $\mu_i$ est L-normalisé. On a $\widetilde{\mu}_i(z) = \sum_{k=0}^{m_i-1} \mu_{ik}^* z^{-k} + z^{-m_i} = \mu_{i0}^* z^{-m_i} \nu_i(z)$ où $\nu_i(z)$ est un polynôme L-normalisé de degré $m_i = \deg(\mu_i)$. $\nu_i(z)$ est le polynôme réciproque normalisé de $\mu_i$ (la normalisation entraîne le coefficient $\mu_{i0}^*$). Or, $\mu_i(z)$ divise $\mu_{i+1}(z)$, il existe donc $\beta_i \in \mathbb{C}[z]$ tel que $\mu_{i+1}(z) = \mu_i(z)\beta_i(z)$ d'où $\widetilde{\mu}_{i+1}(z) = \widetilde{\mu}_i(z)\widetilde{\beta}_i(z) = \mu_{i0}^* z^{-m_i} \nu_i(z) \beta_{i0}^* z^{-b_i} \alpha_i(z)$ où $\alpha_i(z)$ est le polynôme L-normalisé associé à $\beta_i(z)$. Or $\widetilde{\mu}_{i+1}(z) = \mu_{(i+1)0}^* z^{-m_{i+1}} \nu_{i+1}(z)$, d'où $\nu_{i+1}(z) = \mu_{(i+1)0}^{-1*} \mu_{i0}^* z^{m_{i+1}-m_i-b_i} \alpha_i(z)\nu_i(z)$, soit $\nu_i(z)$ divise $\nu_{i+1}(z)$. Par définition des $\nu_i$, ces polynômes sont L-normalisés et ils vérifient que $\nu_i(z)$ divise $\nu_{i+1}(z)$. La matrice $\mathrm{diag}\{\nu_i(z)\}$ est donc la L-forme de Smith de $\boldsymbol{H}(z)$. Par unicité de cette forme, on obtient donc $\nu_i(z) = \mu_i(z) \forall i$, soit par définition des $\nu_i(z)$ : $\mu_i(z) = \mu_{i0}^{-1*} z^{m_i} \widetilde{\mu}_i(z)$. Si on regarde les coefficients de $z^0$ de cette équation, on obtient $\mu_{i0} = \mu_{i0}^{-1*}$, soit $|\mu_{i0}| = 1$. □

La terminologie de polynômes auto-inversif a été empruntée à [68]. Nous détaillons en annexe les différentes définitions des polynômes para-réciproques et des polynômes para-conjugués avec lesquels ils peuvent parfois être confondus.

**Propriété 31.** *Si $\boldsymbol{H}(z)$ est une matrice para-hermitienne alors les matrices polynomiale $\boldsymbol{\Xi}(z)$ et FIR associées $\overline{\overline{\boldsymbol{H}}}(z)$ sont para-conjuguées.*

$\boldsymbol{H}(z)$ étant para-hermitienne, d'après la propriété 29, on peut écrire

$$\boldsymbol{H}(z) = \boldsymbol{H}_0 + \sum_{k=1}^{d} \boldsymbol{H}_k z^{-k} + \sum_{k=1}^{d} \boldsymbol{H}_k^H z^k$$

Par ailleurs, avec les notations des définitions 23 et 24, on a $m = -d$ et $p = d$ d'où $\boldsymbol{\Xi}(z) = z^d \boldsymbol{H}(z)$ et $\overline{\overline{\boldsymbol{H}}}(z) = z^{-d}\boldsymbol{H}(z)$,
soit $\boldsymbol{\Xi}(z) = \boldsymbol{H}_0 z^d + \sum_{k=1}^{d} \boldsymbol{H}_k z^{-k+d} + \sum_{k=1}^{d} \boldsymbol{H}_k^H z^{k+d}$
et $\overline{\overline{\boldsymbol{H}}}(z) = \boldsymbol{H}_0 z^{-d} + \sum_{k=1}^{d} \boldsymbol{H}_k z^{-k-d} + \sum_{k=1}^{d} \boldsymbol{H}_k^H z^{k-d}$ . D'où $\widetilde{\boldsymbol{\Xi}}(z) = \boldsymbol{\Xi}^H(\frac{1}{z^*}) = \boldsymbol{H}_0^H z^{-d} + \sum_{k=1}^{d} \boldsymbol{H}_k^H z^{k-d} + \sum_{k=1}^{d} \boldsymbol{H}_k z^{-k-d}$, soit puisque $\boldsymbol{H}_0$ est hermitienne, $\widetilde{\boldsymbol{\Xi}}(z) = \overline{\overline{\boldsymbol{H}}}(z)$. □

Notons que ni $\boldsymbol{\Xi}(z)$, ni $\overline{\overline{\boldsymbol{H}}}(z)$ ne sont elles-mêmes para-hermitiennes puisqu'elles sont polynomiales (en $z$ et $z^{-1}$).

**Propriété 32.** *Soit $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ une matrice L-polynomiale para-hermitienne, soient $\lambda(z)$ un L-polynôme de $\mathbb{C}[z, z^{-1}]$ et $\underline{u}(z)$ un vecteur L-polynomial non nul de $\mathbb{C}^{n\times 1}[z, z^{-1}]$ vérifiant $\boldsymbol{H}(z)\underline{u}(z) = \lambda(z)\underline{u}(z)$, alors $\lambda(z)$ est para-hermitien.*
*Supposons de plus que $\underline{v}$ et $\mu$ sont tels que $\boldsymbol{H}(z)\underline{v}(z) = \mu(z)\underline{v}(z)$ avec $\lambda \neq \mu$, alors $\langle \underline{u}(z), \underline{v}(z)\rangle = 0$.*

En effet, d'après la propriété 27, et comme $\boldsymbol{H}$ est para-hermitienne : $\langle \boldsymbol{H}(z)\underline{u}(z), \underline{u}(z)\rangle = \langle \underline{u}(z), \widetilde{\boldsymbol{H}}(z)\underline{u}(z)\rangle = \langle \underline{u}(z), \boldsymbol{H}(z)\underline{u}(z)\rangle$. Par définition de $\underline{u}(z)$ et $\lambda(z)$, on a donc $\langle \lambda(z)\underline{u}(z), \underline{u}(z)\rangle = \langle \underline{u}(z), \lambda(z)\underline{u}(z)\rangle$, soit à l'aide des propriétés 26 : $\widetilde{\lambda}(z)\langle \underline{u}(z), \underline{u}(z)\rangle = \lambda(z)\langle \underline{u}(z), \underline{u}(z)\rangle$ et comme par hypothèse $\langle \underline{u}(z), \underline{u}(z)\rangle \neq 0$, on en déduit que $\widetilde{\lambda}(z) = \lambda(z)$, soit $\lambda(z)$ est para-hermitien.

Par définition de $\underline{v}$ et $\mu$ on a $\langle \boldsymbol{H}(z)\underline{u}(z), \underline{v}(z)\rangle = \langle \underline{u}(z), \boldsymbol{H}(z)\underline{v}(z)\rangle = \langle \lambda(z)\underline{u}(z), \underline{v}(z)\rangle = \langle \underline{u}(z), \mu(z)\underline{v}(z)\rangle$, soit $(\widetilde{\lambda}(z) - \mu(z))\langle \underline{u}(z), \underline{v}(z)\rangle = 0$. Si les L-polynômes $\lambda$ et $\mu$ sont distincts, alors comme $\lambda = \widetilde{\lambda}$, on a $\widetilde{\lambda}(z) \neq \mu(z)$, on peut en déduire que les vecteurs $\underline{u}$ et $\underline{v}$ sont orthogonaux au sens du produit défini page 32. □

**Exemple**

Soit $\boldsymbol{H}(z) = \begin{bmatrix} 0 & z \\ \frac{1}{z} & 0 \end{bmatrix}$, $\boldsymbol{H}(z) = \boldsymbol{H}^H(\frac{1}{z^*})$, $\boldsymbol{H}(z)$ est para-hermitienne mais ses coefficients ne le sont pas. $\det \boldsymbol{H}(z) = -1$ donc $\boldsymbol{H}$ est également L-unimodulaire (on ne peut pas dire unimodulaire car elle n'est pas polynomiale

en $z$). Il est facile de montrer que $\underline{u}_1(z) = \frac{\sqrt{2}}{2}\begin{bmatrix} z \\ 1 \end{bmatrix}$ est tel que $\boldsymbol{H}(z)\underline{u}_1(z) = \underline{u}_1(z)$. Prenons le vecteur $\underline{u}_2(z)$ orthogonal à $\underline{u}_1(z)$ ($\widetilde{\underline{u}}_1(z)\underline{u}_2(z) = 0$) et tel que $\widetilde{\underline{u}}_2(z)\underline{u}_2(z) = 1$ soit $\underline{u}_2(z) = \frac{\sqrt{2}}{2}\begin{bmatrix} 1 \\ -z^{-1} \end{bmatrix}$ alors on montre que la matrice $\boldsymbol{U}(z)$ formée des vecteurs $\underline{u}_1(z)$ et $\underline{u}_2(z)$ est telle que : $\boldsymbol{H}(z)\boldsymbol{U}(z) = \boldsymbol{U}(z)\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

Sachant qu'un polynôme para-hermitien peut se factoriser à l'aide d'un polynôme à minimum de phase (cf Théorème de Fejér-Riesz) et qu'une matrice hermitienne définie positive peut se factoriser en $\boldsymbol{H} = \boldsymbol{L}\boldsymbol{L}^H$ avec $\boldsymbol{L}$ triangulaire, une question se pose naturellement pour les matrices para-hermitiennes définies positives sur le cercle : peut-on les factoriser sous la forme $\boldsymbol{H}(z) = \boldsymbol{L}(z)\widetilde{\boldsymbol{L}}(z)$ où $\boldsymbol{L}(z)$ est une matrice triangulaire inférieure à minimum de phase ?

**Propriété 33.** *Si $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ est une matrice para-hermitienne définie positive sur $\mathcal{C}$, alors il existe une matrice rationnelle triangulaire inférieure $\boldsymbol{L}(z) \in \mathbb{C}^{n\times n}(z)$ dont les zéros[4] sont tous à l'intérieur du disque unité telle que*

$$\boldsymbol{H}(z) = \boldsymbol{L}(z)\widetilde{\boldsymbol{L}}(z)$$

Examinons le cas d'une matrice 2×2 et appliquons une méthode de type Cholesky pour faire cette factorisation. Supposons qu'une telle matrice existe, alors,

$$\boldsymbol{H}(z) = \boldsymbol{L}(z)\widetilde{\boldsymbol{L}}(z) = \begin{bmatrix} l_{11}(z) & 0 \\ l_{21}(z) & l_{22}(z) \end{bmatrix}\begin{bmatrix} \tilde{l}_{11}(z) & \tilde{l}_{21}(z) \\ 0 & \tilde{l}_{22}(z) \end{bmatrix}$$

d'où le système d'équations :

$$\begin{cases} h_{11}(z) = l_{11}(z)\tilde{l}_{11}(z) \\ h_{21}(z) = l_{21}(z)\tilde{l}_{11}(z) \\ h_{22}(z) = l_{21}(z)\tilde{l}_{21}(z) + l_{22}(z)\tilde{l}_{22}(z) \end{cases}$$

Par hypothèse, $\boldsymbol{H}(z) = \widetilde{\boldsymbol{H}}(z)$ et est définie positive sur le cercle, ce qui donne en particulier que $h_{11}(z) = \widetilde{h_{11}}(z)$ et est définie positif sur le cercle, soit $h_{11}(z)$ est un L-polynôme para-hermitien défini positif sur le cercle. D'après la propriété 11, on déduit de la première équation que $l_{11}(z) \in \mathbb{C}_s[z]$ est le facteur à minimum de phase de $h_{11}(z)$. De la deuxième équation, on déduit que $l_{21}(z) = \frac{h_{21}(z)}{\tilde{l}_{11}(z)}$. $l_{11}(z)$ étant à minimum de phase, $\tilde{l}_{11}(z)$ a ses racines à

4. $\boldsymbol{L}(z)$ étant triangulaire, les zéros de $\boldsymbol{L}(z)$ coïncident avec les zéros de ses éléments diagonaux $l_{ii}(z)$.

l'extérieur du disque unité (sauf s'il est constant). On a donc $l_{21}(z) \in \mathbb{C}(z)$ et instable. Enfin,

$$\begin{aligned} l_{22}(z)\tilde{l}_{22}(z) &= h_{22}(z) - l_{21}(z)\tilde{l}_{21}(z) = h_{22}(z) - \frac{h_{21}(z)\tilde{h}_{21}(z)}{l_{11}(z)\tilde{l}_{11}(z)} \\ &= \frac{h_{22}(z)h_{11}(z) - h_{21}(z)\tilde{h}_{21}(z)}{h_{11}(z)} \end{aligned} \tag{3.4}$$

mais $l_{22}(z)$ n'étant pas polynomial a priori, nous ne pouvons utiliser la factorisation à minimum de phase. Cependant le numérateur de (3.4) est le déterminant de $\boldsymbol{H}(z)$ et est donc para-hermitien, défini positif sur le cercle par hypothèse et admet ainsi une factorisation à minimum de phase : $\det \boldsymbol{H}(z) = \alpha(z)\tilde{\alpha}(z)$ avec $\alpha(z) \in \mathbb{C}_s[z]$. Ainsi,

$$l_{22}(z)\tilde{l}_{22}(z) = \frac{\alpha(z)\tilde{\alpha}(z)}{l_{11}(z)\tilde{l}_{11}(z)}$$

et on peut choisir

$$l_{22}(z) = \frac{\alpha(z)}{l_{11}(z)} \in \mathbb{C}_s(z)$$

par définition de $\alpha(z)$ et $l_{11}(z)$. On a ainsi trouvé une factorisation où tous les zéros de la matrice rationnelle $\boldsymbol{L}(z)$ sont à l'intérieur du disque unité, on ne peut garantir toutefois que les pôles sont aussi à l'intérieur du disque (ils ne le sont pas dès que cette matrice n'est pas polynomiale) et elle ne peut donc pas être qualifiée de factorisation à minimum de phase.

Dans le cas où $n \geq 2$, c'est un peu plus complexe. En paramétrant $\boldsymbol{H}(z) = \boldsymbol{L}(z)\widetilde{\boldsymbol{L}}(z)$ où $\boldsymbol{L}(z)$ est une matrice triangulaire inférieure, on obtient $h_{11}(z) = l_{11}(z)\tilde{l}_{11}(z)$ et donc $l_{11}(z)$ est le facteur à minimum de phase (polynomial) de $h_{11}(z)$.

Pour tout $j > 1$ : $l_{j1}(z) = \frac{h_{1j}(z)}{\tilde{l}_{11}(z)}$. Ainsi, $l_{j1}(z)$ est instable dès que $h_{11}(z)$ est différent d'une constante.

Pour $i > 1$ on montre que $l_{ii}(z) = \dfrac{\gamma_i(z)}{\alpha_i(z)}$ où $\alpha_i(z)$ est le ppcm des dénominateurs des $l_{ik}(z), k = 1$ à $i$, et $\gamma_i(z)$ est le facteur à minimum de phase du polynôme $\alpha_i(z)\tilde{\alpha}_i(z)(h_{ii}(z) - \sum_{k=1}^{i-1} l_{ik}(z)\tilde{l}_{ik}(z))$.

Enfin, pour $j > i$, on a $l_{ji}(z) = \dfrac{h_{ji}(z) - \sum_{k=1}^{i-1} l_{ik}(z)\tilde{l}_{jk}(z)}{\tilde{l}_{ii}(z)}$. □

Remarquons que si, par hypothèse, on sait déjà qu'il existe une factorisation de la matrice para-hermitienne de la forme $\boldsymbol{H}(z) = \boldsymbol{N}(z)\widetilde{\boldsymbol{N}}(z)$ avec $\boldsymbol{N}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ (ce qui est le cas d'une matrice spectrale, cf p. 53) alors, grâce au théorème de factorisation spectrale [55], on sait qu'il existe une factorisation de $\boldsymbol{H}(z)$ sous la forme $\boldsymbol{H}(z) = \boldsymbol{C}(z)\widetilde{\boldsymbol{C}}(z)$ où $\boldsymbol{C}(z) \in \mathbb{C}^{n\times n}[z^{-1}]$

est une matrice FIR Schur c'est-à-dire ayant tous ses zéros à l'intérieur du disque unité.

Si ce n'est pas les zéros qui nous intéressent lors de la factorisation mais les pôles, on peut de même, faire une factorisation stable d'une matrice para-hermitienne définie positive sur le cercle :

**Propriété 34.** *Si $\boldsymbol{H}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ est une matrice para-hermitienne définie positive sur $\mathcal{C}$, alors il existe une matrice rationnelle triangulaire inférieure $\boldsymbol{L}_s(z) \in \mathbb{C}^{n\times n}(z)$ dont tous les pôles sont à l'intérieur du disque unité telle que*

$$\boldsymbol{H}(z) = \boldsymbol{L}_s(z)\widetilde{\boldsymbol{L}_s}(z)$$

Le problème se pose différemment pour les zéros et pour les pôles. En effet, dans le cas d'une matrice rationnelle triangulaire, les zéros de la matrice sont les zéros des éléments de la diagonale, et donc pour garantir que les zéros sont bien placés, il suffit de regarder les éléments diagonaux alors que pour que $z_p$ soit pôle d'une matrice rationnelle, il suffit qu'il soit pôle d'un des éléments de la matrice.

Dans le cas où $n = 2$, avec les mêmes notations que précédemment, choisissons pour $l_{s11}(z)$ le facteur à maximum de phase de $h_{11}(z)$, soit $l_{s11}(z) = \tilde{l}_{11}(z)$. On a ensuite $l_{s21}(z) = \frac{h_{21}(z)}{\tilde{l}_{s11}(z)}$ qui par hypothèse est une fraction rationnelle stable puisque les racines de $\tilde{l}_{s11}(z)$ sont à l'intérieur du disque unité. Enfin, $l_{s22}(z) = \dfrac{\alpha(z)}{\tilde{l}_{11}(z)} \in \mathbb{C}_s(z)$. Par choix de $\alpha(z)$ ou $\tilde{\alpha}(z)$ comme numérateur de $l_{s22}(z)$, on peut également choisir les zéros de $l_{s22}(z)$ dans ou hors du disque. On a ainsi une factorisation de $\boldsymbol{H}(z)$ pour laquelle tous les pôles sont à l'intérieur du disque unité.

Dans le cas où $n \geq 2$, on va procéder de même que pour la propriété 33 :
Pour tout $j > 1$ : $l_{sj1}(z) = \frac{h_{1j}(z)}{\tilde{l}_{s11}(z)}$, soit $l_{sj1}(z)$ stable.

Pour $i > 1$ on montre que $l_{sii}(z) = \dfrac{\gamma_i(z)}{\alpha_i(z)}$ où $\alpha_i(z)$ est le ppcm des $l_{sik}(z), k = 1$ à $i$, et $\gamma_i(z)$ est le facteur à maximum de phase du polynôme $\alpha_i(z)\tilde{\alpha}_i(z)(h_{ii}(z) - \sum_{k=1}^{i-1} l_{sik}(z)\tilde{l}_{sik}(z))$.

Enfin, pour $j > i$, on a $l_{sji}(z) = \dfrac{h_{ji}(z) - \sum_{k=1}^{i-1} l_{sik}(z)\tilde{l}_{sjk}(z)}{\tilde{l}_{sii}(z)}$. Si tous les pôles de $l_{sji}$ sont à l'intérieur du disque unité, la matrice $\boldsymbol{L}_s$ ainsi obtenue est stable et telle que $\boldsymbol{H} = \boldsymbol{L}_s\widetilde{\boldsymbol{L}_s}$. Dans le cas contraire, supposons que $z_p$ soit pôle de degré $l$ de $l_{sji}$. Soit $\boldsymbol{D}$ la matrice diagonale définie par $d_{ij} = \delta_{ij}$ pour $i \neq j$ et $d_{ii}(z) = \frac{(1-z_p^{-1}z)^l}{(1-z_p^* z^{-1})^l}$ alors $\boldsymbol{H} = \boldsymbol{L}_s\boldsymbol{D}\widetilde{\boldsymbol{L}_s\boldsymbol{D}}$ et on obtient la factorisation désirée.

### 3.2.5 Matrice L-polynomiale para-unitaire

La notion de matrice unitaire sur $\mathbb{C}^{n\times n}$ s'étend aux matrices polynomiales [80] et aux matrices L-polynomiales [44], on appelle alors ces matrices des matrices para-unitaires. Ces matrices possèdent de nombreuses propriétés, utilisées notamment dans les bancs de filtre et les problèmes d'égalisation.

**Définition 35.** *Soit* $\boldsymbol{U} \in \mathbb{C}^{n\times n}[z, z^{-1}]$, $\boldsymbol{U}(z)$ *est para-unitaire ssi*

$$\boldsymbol{U}(z)\boldsymbol{U}^H(\frac{1}{z^*}) = \boldsymbol{U}^H(\frac{1}{z^*})\boldsymbol{U}(z) = \boldsymbol{I}, \forall z \in \mathbb{C}^*$$

**Exemple**

Soit $\boldsymbol{H}(z) = \begin{bmatrix} 0 & z \\ \frac{1}{z} & 0 \end{bmatrix}$, on a vu au paragraphe précédent que $\boldsymbol{H}(z) = \boldsymbol{H}^H(\frac{1}{z^*})$, soit $\boldsymbol{H}(z)$ est para-hermitienne. On a également $\boldsymbol{H}(z)\boldsymbol{H}^H(\frac{1}{z^*}) = \boldsymbol{I}$, cette matrice est donc également para-unitaire. Comme $\det \boldsymbol{H}(z) = -1$, $\boldsymbol{H}(z)$ est également L-unimodulaire.

**Propriété 36.** *Les colonnes d'une matrice para-unitaire sont orthonormées (au sens du* $\langle ., . \rangle$ *de la définition 25).*

Il suffit de développer : $\widetilde{\boldsymbol{U}}(z)\boldsymbol{U}(z) = \begin{bmatrix} \underline{u}_1^H(\frac{1}{z^*}) \\ \vdots \\ \underline{u}_n^H(\frac{1}{z^*}) \end{bmatrix} \begin{bmatrix} \underline{u}_1(z) & \dots & \underline{u}_n(z) \end{bmatrix} = \boldsymbol{I}$

et on obtient $\langle \underline{u}_i(z), \underline{u}_j(z) \rangle = \delta_{ij}$ □

Cette propriété, exprimée avec les coefficients temporels (c'est-à-dire avec les coefficients des polynômes considérés) est connue sous le nom de "Shifted Orthogonality Condition" de Turcajova (étude d'ondelettes) [79].

**Propriété 37.** *Les coefficients d'une matrice para-unitaire d'ordre* $l$, $\boldsymbol{U}(z) = z^m \sum_{k=0}^{l} \boldsymbol{U}_k z^k$ *(avec* $m \in \mathbb{Z}, \boldsymbol{U}_0$ *et* $\boldsymbol{U}_l$ *non nuls) vérifient :*

$$\sum_{k=0}^{l} \boldsymbol{U}_k \boldsymbol{U}_k^H = \boldsymbol{I} \tag{3.5}$$

$$\sum_{j=k}^{l} \boldsymbol{U}_j \boldsymbol{U}_{j-k}^H = \boldsymbol{0} \ \forall k \in \{1, \dots, l\} \tag{3.6}$$

*en particulier, si l'ordre de* $\boldsymbol{U}$ *est non nul, on a* $\boldsymbol{U}_0\boldsymbol{U}_l^H = \boldsymbol{0}$.

On développe

$$\begin{aligned}\boldsymbol{U}(z)\widetilde{\boldsymbol{U}}(z) &= z^m \left(\sum_{k=0}^{l} \boldsymbol{U}_k z^k\right) z^{-m} \left(\sum_{k=0}^{l} \boldsymbol{U}_k^H z^{-k}\right) \\ &= \sum_{k=0}^{l} \boldsymbol{U}_k \boldsymbol{U}_k^H + \sum_{k=1}^{l}\sum_{j=k}^{l} \left(\boldsymbol{U}_j \boldsymbol{U}_{j-k}^H z^k + (\boldsymbol{U}_j \boldsymbol{U}_{j-k}^H)^H z^{-k}\right)\end{aligned}$$

□

Soit $\mathbb{U}$ la matrice formée des coefficients de $z^{-m}\boldsymbol{U}(z)$ (ces coefficients sont ceux de la réponse impulsionnelle lorsque $\boldsymbol{U}(z)$ est FIR) : $\mathbb{U} = [\boldsymbol{U}_0\, \boldsymbol{U}_1 \ldots \boldsymbol{U}_l] \in \mathbb{C}^{n\times n(l+1)}$, alors (3.5) implique que $\mathbb{U}$ est une matrice semi-unitaire : $\mathbb{U}\mathbb{U}^H = \boldsymbol{I}$. Nous avons notamment exploité cette propriété dans un critère d'optimisation permettant de déterminer un égaliseur [4].

**Propriété 38.** *Toute matrice L-polynomiale para-unitaire est L-équivalente à la matrice identité.*

On utilise la propriété reliant le déterminant et les L-polynômes invariants : soient $\lambda_i$ les L-polynômes invariants de $\boldsymbol{U}(z)$ matrice para-unitaire, on a $\det \boldsymbol{U}(z) = cz^\alpha \prod_{i=1}^n \lambda_i(z)$, avec $c \in \mathbb{C}^*$, $\alpha \in \mathbb{Z}$, d'où $\det \widetilde{\boldsymbol{U}}(z) = c^* z^{-\alpha} \prod_{i=1}^n \widetilde{\lambda}_i(z)$. $\boldsymbol{U}(z)$ étant para-unitaire $\boldsymbol{U}(z)\widetilde{\boldsymbol{U}}(z) = \boldsymbol{I}$, d'où pour les déterminants : $cc^* z^\alpha \prod_{i=1}^n \lambda_i(z) z^{-\alpha} \prod_{i=1}^n \widetilde{\lambda}_i(z) = 1$, on en déduit donc que les $\lambda_i$ sont tous des éléments inversibles, c'est à dire des monômes et comme par définition ils sont L-unitaires, ils sont égaux à 1. □

**Corollaire 39.** *Le déterminant d'une matrice para-unitaire $\boldsymbol{U}(z) \in \mathbb{C}^{n\times n}[z, z^{-1}]$ est para-unitaire :*

$$\det \boldsymbol{U}(z) = cz^\alpha \ \textit{avec}\ |c| = 1 \ \textit{et}\ \alpha \in \mathbb{Z}.$$

A l'aide de la propriété 19 (L-équivalence d'une matrice L-unimodulaire à la matrice identité), on déduit immédiatement :

**Corollaire 40.** *Toute matrice para-unitaire est L-équivalente à une matrice L-unimodulaire.*

Il est bien connu qu'une matrice unitaire peut être décomposée en un produit de rotations planes [49]. Il existe une décomposition analogue pour les matrices para-unitaires, faisant intervenir des rotations et des matrices comportant un seul retard (matrices para-unitaires élémentaires [62]). Cette décomposition a été développée dans [80] pour le cas réel FIR et nous l'avons étendue au cas complexe FIR [2, 74].

**Propriété 41.** *Soit* $\boldsymbol{U}(z) \in \mathbb{C}^{n\times n}[z^{-1}]$ *une matrice para-unitaire FIR de degré de McMillan* $N$*, alors il existe des matrices unitaires* $\boldsymbol{Q}_i \in \mathbb{C}^{n\times n}$ *telles que*

$$\boldsymbol{U}(z) = \boldsymbol{Q}_0\boldsymbol{Z}(z)\boldsymbol{Q}_1 \dots \boldsymbol{Z}(z)\boldsymbol{Q}_{N-1}\boldsymbol{Z}(z)\boldsymbol{Q}_N \tag{3.7}$$

*avec*

$$\boldsymbol{Z}(z) = \begin{bmatrix} \boldsymbol{I}_{n-1} & 0 \\ 0 & z^{-1} \end{bmatrix} \in \mathbb{C}^{n\times n}[z^{-1}] \tag{3.8}$$

Chaque matrice unitaire $\boldsymbol{Q}_i$ peut ensuite se décomposer en produit de $M = n(n-1)/2$ rotations de Givens [49, p. 215] $\boldsymbol{U}(i, j, \theta, \phi)$, $1 \le i < j \le n$ et

$$\boldsymbol{U}(i,j,\theta,\phi) = \begin{bmatrix} \boldsymbol{I}_{n_1} & \vdots & \boldsymbol{0} & \vdots & \boldsymbol{0} \\ \cdots & \cos\theta & \cdots & \sin\theta e^{\jmath\phi} & \cdots \\ \boldsymbol{0} & \vdots & \boldsymbol{I}_{n_2} & \vdots & \boldsymbol{0} \\ \cdots & -\sin\theta e^{-\jmath\phi} & \cdots & \cos\theta & \cdots \\ \boldsymbol{0} & \vdots & \boldsymbol{0} & \vdots & \boldsymbol{I}_{n_3} \end{bmatrix} \begin{matrix} \\ i \\ \\ j \\ \\ \end{matrix} \tag{3.9}$$

avec $n_1 + n_2 + n_3 = n$, et on obtient ainsi une factorisation de matrice para-unitaire à l'aide de produits de matrices de rotations et de $\boldsymbol{Z}(z)$. Dans le cas $n = 2$, les matrices unitaires $\boldsymbol{Q}_i$ sont simplement des matrices de rotations.

On peut déduire de la propriété précédente que le degré de MacMillan d'une matrice para-unitaire FIR est égal au degré (en la variable $z^{-1}$ ) du déterminant (polynomial en $z^{-1}$) de cette matrice [80] alors que dans le cas d'une matrice FIR quelconque, le degré de MacMillan est supérieur ou égal au degré du déterminant. Gao [45] montre quant à lui, que le degré de MacMillan $N$ d'une matrice para-unitaire FIR appartenant à $\mathbb{C}^{n\times n}[z^{-1}]$ est relié à l'ordre $l$ de la matrice par $l \le N \le nl$.

**Exemple**

Soit $\boldsymbol{A}(z) = \begin{bmatrix} z^{-1} & 0 \\ 0 & z^{-1} \end{bmatrix} = \boldsymbol{0} + z^{-1}\boldsymbol{I}$. $\boldsymbol{A}(z)$ est FIR para-unitaire, d'ordre $l = 1$. Sa forme de Smith-McMillan est $\boldsymbol{SM}(z) = \boldsymbol{A}(z) = z^{-1}\boldsymbol{I}$ et son degré de McMillan est donc $N = 2$. Par contre, le L-ordre de $\boldsymbol{A}(z)$ est égal à l'ordre de $\boldsymbol{I}$, soit 0, qui est également le degré de McMillan de $\boldsymbol{I}$.

On dit qu'un système dont la matrice de transfert $\boldsymbol{H}(z)$ est une matrice rationnelle est un système sans perte [80] si la matrice de transfert est causale, stable et para-unitaire, c'est à dire si le degré des numérateurs des $h_{ij}(z)$ est inférieur au degré des dénominateurs, si les pôles des $h_{ij}(z)$ sont à

l'intérieur du disque unité et si $\boldsymbol{H}(z)\boldsymbol{H}^H(\frac{1}{z^*}) = \boldsymbol{H}^H(\frac{1}{z^*})\boldsymbol{H}(z) = \boldsymbol{I}, \forall z \in \mathbb{C}^*$. Les matrices L-polynomiales telles que $\inf_{\boldsymbol{H}_i \neq \boldsymbol{0}} i < 0$ ne sont donc pas des matrices sans perte car elles ne sont pas causales.

Mais à toute matrice L-polynomiale para-unitaire $\boldsymbol{H} \in \mathbb{C}^{n\times n}[z, z^{-1}]$, on peut associer une matrice FIR sans perte (donc para-unitaire) : posons $\boldsymbol{H}(z) = z^p\overline{\overline{\boldsymbol{H}}}(z)$ avec $\overline{\overline{\boldsymbol{H}}}_0 \neq \boldsymbol{0}, \overline{\overline{\boldsymbol{H}}}(z) \in \mathbb{C}^{n\times n}[z^{-1}]$ donc causale et stable (tous ses pôles sont en zéro). Il est immédiat que $\overline{\overline{\boldsymbol{H}}}(z)\overline{\overline{\boldsymbol{H}}}^H(\frac{1}{z^*}) = \boldsymbol{I}$ ssi $\boldsymbol{H}(z)\boldsymbol{H}^H(\frac{1}{z^*}) = \boldsymbol{I}$.

De même, on peut associer à toute matrice L-polynomiale para-unitaire une matrice polynomiale comme on l'a fait en définition 23, cette matrice étant également para-unitaire : $\boldsymbol{H}(z) = z^m\boldsymbol{\Xi}(z)$ avec $\boldsymbol{\Xi}(z) \in \mathbb{C}^{n\times n}[z], \boldsymbol{\Xi}_0 \neq \boldsymbol{0}$ et $\boldsymbol{\Xi}(z)^H(\frac{1}{z^*})\boldsymbol{\Xi}(z) = \boldsymbol{I}$.

Le fait que le coefficient du terme de degré 0 est non nul va nous permettre d'établir une propriété originale de l'ordre d'une matrice para-unitaire FIR (polynomiale en $z^{-1}$).

**Propriété 42.** *Soit $\boldsymbol{U}(z) \in \mathbb{C}^{n\times n}[z^{-1}]$ une matrice para-unitaire FIR telle que $\boldsymbol{U}_0 \neq 0$, alors le degré de Mac-Millan de $\boldsymbol{U}(z)$ est égal à l'ordre de $\boldsymbol{U}(z)$.*

Nous allons montrer cette propriété dans le cas où les coefficients sont réels et la dimension $n = 2$, mais tout porte à croire qu'elle est vraie pour une dimension quelconque. Commençons par établir le lemme suivant :

**Lemme 43.** *Soit $\boldsymbol{H}(z) \in \mathbb{R}^{2\times 2}[z^{-1}]$ une matrice para-unitaire FIR de degré $N$ définie par*

$$\boldsymbol{H}(z) = \boldsymbol{Z}(z)\ \boldsymbol{R}_1\ \boldsymbol{Z}(z)\ \boldsymbol{R}_2\ \ldots \boldsymbol{Z}(z)\ \ \boldsymbol{R}_{N-1}\boldsymbol{Z}(z) \tag{3.10}$$

*avec* $\boldsymbol{Z}(z) = \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix} = \boldsymbol{Z}_0 + \boldsymbol{Z}_1 z^{-1}$ *et* $\boldsymbol{R}_i = \begin{bmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{bmatrix}$, *alors*

$$\boldsymbol{H}(z) = \boldsymbol{H}_0 + \boldsymbol{H}_1 z^{-1} + \cdots + \boldsymbol{H}_N z^{-N}$$

*avec* $\boldsymbol{H}_0 = \begin{bmatrix} \prod_{k=1}^{N-1}\cos\theta_k & 0 \\ 0 & 0 \end{bmatrix}$ *et* $\boldsymbol{H}_N = \begin{bmatrix} 0 & 0 \\ 0 & \prod_{k=1}^{N-1}\cos\theta_k \end{bmatrix}$.

On montre ceci par récurrence : la propriété est vraie pour $N = 2$ : $\boldsymbol{H}(z) = \boldsymbol{Z}(z)\ \boldsymbol{R}_1\ \boldsymbol{Z}(z) = (\boldsymbol{Z}_0 + \boldsymbol{Z}_1 z^{-1})\boldsymbol{R}_1(\boldsymbol{Z}_0 + \boldsymbol{Z}_1 z^{-1}) = \boldsymbol{Z}_0\boldsymbol{R}_1\boldsymbol{Z}_0 + (\boldsymbol{Z}_0\boldsymbol{R}_1\boldsymbol{Z}_1 + \boldsymbol{Z}_1\boldsymbol{R}_1\boldsymbol{Z}_0)z^{-1} + \boldsymbol{Z}_1\boldsymbol{R}_1\boldsymbol{Z}_1 z^{-2}$ d'où, par identification des $\boldsymbol{H}_i$, et comme $\boldsymbol{R}_1\boldsymbol{Z}_0 = \begin{bmatrix} \cos\theta_i & 0 \\ -\sin\theta_i & 0 \end{bmatrix}$ et $\boldsymbol{R}_1\boldsymbol{Z}_1 = \begin{bmatrix} 0 & \sin\theta_i \\ 0 & \cos\theta_i \end{bmatrix}$, $\boldsymbol{H}_0 = \boldsymbol{Z}_0\boldsymbol{R}_1\boldsymbol{Z}_0 = \begin{bmatrix} \cos\theta_1 & 0 \\ 0 & 0 \end{bmatrix}$ et $\boldsymbol{H}_2 = \boldsymbol{Z}_1\boldsymbol{R}_1\boldsymbol{Z}_1 = \begin{bmatrix} 0 & 0 \\ 0 & \cos\theta_1 \end{bmatrix}$.

Soit maintenant $\boldsymbol{H}(z) = (\boldsymbol{Z}(z)\ \boldsymbol{R}_1\ \boldsymbol{Z}(z)\ \boldsymbol{R}_2\ \ldots \boldsymbol{Z}(z))\ \boldsymbol{R}_{N-1}\boldsymbol{Z}(z) = \boldsymbol{M}(z)\ \boldsymbol{R}_{N-1}(\boldsymbol{Z}_0 + \boldsymbol{Z}_1 z^{-1})$. $\boldsymbol{M}(z)$ est une matrice para-unitaire de degré $N-1$ de la forme (3.10) et qui s'écrit donc
$\boldsymbol{M}(z) = \begin{bmatrix} \prod_{k=1}^{N-2} \cos\theta_k & 0 \\ 0 & 0 \end{bmatrix} + \boldsymbol{M}_1 z^{-1} + \cdots + \begin{bmatrix} 0 & 0 \\ 0 & \prod_{k=1}^{N-2} \cos\theta_k \end{bmatrix} z^{-N+1}$.
Ainsi $\boldsymbol{H}_0 = \boldsymbol{M}_0\boldsymbol{R}_{N-1}\boldsymbol{Z}_0 = \begin{bmatrix} \prod_{k=1}^{N-2} \cos\theta_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\theta_{N-1} & 0 \\ -\sin\theta_{N-1} & 0 \end{bmatrix} = \begin{bmatrix} \prod_{k=1}^{N-1} \cos\theta_k & 0 \\ 0 & 0 \end{bmatrix}$
et $\boldsymbol{H}_N = \boldsymbol{M}_{N-1}\boldsymbol{R}_{N-1}\boldsymbol{Z}_1 = \begin{bmatrix} 0 & 0 \\ 0 & \prod_{k=1}^{N-2} \cos\theta_k \end{bmatrix} \begin{bmatrix} 0 & \sin\theta_{N-1} \\ 0 & \cos\theta_{N-1} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & \prod_{k=1}^{N-1} \cos\theta_k \end{bmatrix}$. □

Remarquons que l'hypothèse $\boldsymbol{H}_0 \neq \boldsymbol{0}$ est équivalente à $\cos\theta_k \neq 0\ \forall k$, elle même équivalente à $\boldsymbol{H}_N \neq \boldsymbol{0}$. De plus, on a bien $\boldsymbol{H}_0\boldsymbol{H}_N^T = \boldsymbol{0}$.

Prenons maintenant le cas d'une matrice para-unitaire FIR quelconque $\boldsymbol{U}(z)$. D'après la factorisation de Vaidyanathan [80, p. 729], $\boldsymbol{U}(z)$ s'écrit sous la forme : $\boldsymbol{U}(z) = \boldsymbol{R}_0\boldsymbol{H}(z)\boldsymbol{R}_N \begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$ avec $\boldsymbol{H}(z)$ de la forme (3.10), $\boldsymbol{R}_0$ et $\boldsymbol{R}_N$ des matrices de rotation et $N$ est le degré de MacMillan de $\boldsymbol{U}(z)$, soit $\boldsymbol{U}(z) = \boldsymbol{R}_0(\boldsymbol{H}_0 + \boldsymbol{H}_1 z^{-1} + \cdots + \boldsymbol{H}_N z^{-N})\boldsymbol{R}_N \begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$.

Le coefficient du terme en $z^0$ de $\boldsymbol{U}(z)$ est $\boldsymbol{U}_0 = \boldsymbol{R}_0\boldsymbol{H}_0\boldsymbol{R}_N \begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$ et celui du terme en $z^{-N}$ est $\boldsymbol{U}_N = \boldsymbol{R}_0\boldsymbol{H}_N\boldsymbol{R}_N \begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$. Les matrices $\boldsymbol{R}_0$, $\boldsymbol{R}_N$ et $\begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$ étant régulières, $\boldsymbol{H}_0 \neq \boldsymbol{0} \Leftrightarrow \boldsymbol{U}_0 \neq \boldsymbol{0}$ et $\boldsymbol{H}_N \neq \boldsymbol{0} \Leftrightarrow \boldsymbol{U}_N \neq \boldsymbol{0}$. On a ainsi montré que toute matrice paraunitaire $\boldsymbol{U}(z)$ telle que $\boldsymbol{U}_0 \neq \boldsymbol{0}$ s'écrit $\boldsymbol{U}(z) = \boldsymbol{U}_0 + \boldsymbol{U}_1 z^{-1} + \cdots + \boldsymbol{U}_N z^{-N}$ avec $\boldsymbol{U}_N \neq \boldsymbol{0}$, soit que l'ordre de $\boldsymbol{U}(z)$ est égal au degré de MacMillan de la matrice. □
Remarquons que l'on a bien (propriété 36) $\boldsymbol{U}_0\boldsymbol{U}_N^T = \boldsymbol{R}_0\boldsymbol{H}_0\boldsymbol{H}_N^T\boldsymbol{R}_0^T = \boldsymbol{0}$.

La propriété précédente nous permet de déterminer une factorisation des matrices L-polynomiales para-unitaires.

**Corollaire 44.** *Soit $\boldsymbol{U}(z) \in \mathbb{R}^{2\times 2}[z, z^{-1}]$ une matrice para-unitaire d'ordre $p-m$ : $\boldsymbol{U}(z) = \sum_{k=m}^{p} \boldsymbol{U}_k z^k$ alors le L-degré de $\boldsymbol{U}$ est égal à l'ordre de $\boldsymbol{U}$ et on peut factoriser*

$$\boldsymbol{U}(z) = z^p \boldsymbol{R}_0(z)\boldsymbol{Z}(z)\ldots\boldsymbol{Z}(z)\boldsymbol{R}_{p-m} \begin{bmatrix} 1 & 0 \\ 0 & \pm 1 \end{bmatrix}$$

*où $\boldsymbol{Z}(z) = \begin{bmatrix} 1 & 0 \\ 0 & z^{-1} \end{bmatrix}$ et les $\boldsymbol{R}_i$ sont des matrices de rotation.*

En effet, soit $\overline{\overline{\boldsymbol{U}}}(z)$ la matrice FIR associée à $\boldsymbol{U}(z)$ : $\overline{\overline{\boldsymbol{U}}}(z) = z^{-p}\boldsymbol{U}(z)$. $\overline{\overline{\boldsymbol{U}}}(z)$ étant une matrice para-unitaire, il existe une factorisation de $\overline{\overline{\boldsymbol{U}}}(z)$ comportant $\overline{\overline{N}}$ éléments où $\overline{\overline{N}}$ est le degré de McMillan de $\overline{\overline{\boldsymbol{U}}}(z)$. Le premier coefficient de $\overline{\overline{\boldsymbol{U}}}(z)$ étant non nul, on peut appliquer les résultats de la propriété 42, soit $\overline{\overline{N}} = p - m$, ce qui donne la factorisation désirée. □

Si les propriétés d'hermitianité et d'unitarité s'étendent aux cas des matrices polynomiales, il est séduisant de penser que l'on peut "diagonaliser " une matrice para-hermitienne à l'aide d'une matrice para-unitaire, comme on peut –théorème spectral- diagonaliser toute matrice hermitienne à l'aide d'une matrice unitaire. Les travaux de McWhriter et al [35, 62, 44] laissent en effet à penser que l'on peut étendre la décomposition en valeurs singulières (SVD) et en valeurs propres (EVD) aux matrices polynomiales. Leur algorithme donne une version approchée d'une telle décomposition.
La section suivante présente les résultats actuels sur cette conjoncture.

## 3.3 Diagonalisation d'une matrice polynomiale

Dans le cas de matrices polynomiales, le terme de valeur propre existe [48] : soit $\boldsymbol{A}(z) \in \mathbb{C}^{n\times n}[z]$, $\lambda_0$ est valeur propre de $\boldsymbol{A}(z)$ ssi elle est racine de l'équation

$$\det \boldsymbol{A}(\lambda_0) = 0$$

$\underline{x}_0$ est un vecteur propre de $\boldsymbol{A}(\lambda)$ associé à la valeur propre $\lambda_0$ si $\boldsymbol{A}(\lambda_0)\underline{x}_0 = \underline{0}$. On étend ainsi la notion de valeur propre comme racine du déterminant d'une matrice polynomiale d'ordre 1 ($\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$) à une matrice polynomiale d'ordre quelconque : $\det(\boldsymbol{A}_m\lambda^m + \cdots + \boldsymbol{A}_1\lambda + \boldsymbol{A}_0) = 0$. Les valeurs propres d'une matrice polynomiale ainsi définies sont donc des éléments de $\mathbb{C}$.

Si l'on désire diagonaliser une matrice polynomiale, c'est à dire , grâce à des opérations à droite et à gauche par des matrices polynomiales rendre cette matrice diagonale, ce n'est pas grâce aux valeurs propres et vecteurs propres ci-dessus définies car la matrice diagonale obtenue n'est pas une matrice constante, mais une matrice dépendant de la variable $z$ : on cherche des fonctions $\lambda(z)$ et des vecteurs $\underline{v}(z)$ tels que

$$\boldsymbol{A}(z)\underline{v}(z) = \lambda(z)\underline{v}(z) \tag{3.11}$$

Comme on l'a vu dans la section 3.2.3 page 35 si la matrice est para-hermitienne, et s'il existe des L-polynômes et des L-vecteurs vérifiant (3.11) alors les L-polynômes sont para-hermitiens. De plus si les L-polynômes sont différents, les vecteurs correspondants sont orthogonaux au sens du produit défini par $< \underline{u}(z), \underline{v}(z) >= \widetilde{\underline{u}}(z)\underline{v}(z)$. Le problème qui se pose est celui de l'existence des ces éléments puisqu'on ne travaille pas sur un espace vectoriel. Comme nous le verrons par la suite, il n'existe pas toujours de solution L-polynomiale, et donc les méthodes consistant à ramener un problème d'ordre $m$ et de dimension $n \times n$ en un problème d'ordre 1 (faisceau) et de dimension $nm \times nm$ [32] ne peuvent être utilisées. Le seul résultat général portant sur la diagonalisation de matrice polynomiale concerne, à notre connaissance, les matrices normales [65].

### 3.3.1 Diagonalisation d'une matrice polynomiale normale

Soit $\boldsymbol{P}(z) \in \mathbb{C}^{n\times n}[z]$ une matrice polynomiale $\boldsymbol{P}(z) = \sum_{i=0}^{l} \boldsymbol{P}_i z^i$. On dira que cette matrice est normale ssi la matrice $\boldsymbol{P}(z)$ est normale $\forall z \in \mathbb{C}$ : $\boldsymbol{P}^H(z)\boldsymbol{P}(z) = \boldsymbol{P}(z)\boldsymbol{P}^H(z)$. On montre dans [65] que si une matrice polynomiale est normale alors ses coefficients vérifient $\boldsymbol{P}_i\boldsymbol{P}_j^H = \boldsymbol{P}_j^H\boldsymbol{P}_i$ pour

tout $i, j = 0, 1, \ldots l$. En particulier, tous les coefficients $\boldsymbol{P}_i, i = 0, \ldots l$ sont des matrices normales.

**Propriété 45.** *Si une matrice polynomiale $\boldsymbol{P}(z) \in \mathbb{C}^{n\times n}[z]$ est normale, et s'il existe un $\mu_0 \in \mathbb{C}$ tel que $\boldsymbol{P}(\mu_0)$ a $n$ valeurs propres distinctes, alors il existe une matrice unitaire $\boldsymbol{U} \in \mathbb{C}^{n\times n}$ telle que*

$$\boldsymbol{U}^H \boldsymbol{P}(z)\boldsymbol{U} = \boldsymbol{\Lambda}(z) \text{ avec } \boldsymbol{\Lambda}(z) \text{ diagonale}$$

*c'est à dire que tous les coefficients matriciels sont simultanément diagonalisables par une matrice unitaire.*

Remarquons que l'hypothèse de normalité est très forte et que pour l'application qui nous intéresse, il n'y a aucune raison qu'elle soit vérifiée. Avant de nous intéresser aux valeurs propres éventuelles d'une matrice L-polynomiale, nous allons nous intéresser aux valeurs singulières. L'ensemble des polynômes de Laurent n'étant pas fermé, nous allons relâcher la contrainte pour l'existence de solution de valeurs singulières et nous allons nous intéresser au problème de l'existence des valeurs singulières pour des fonctions continues sur le cercle unité. Ensuite, nous nous servirons de la propriété de densité de l'ensemble des polynômes de Laurent dans l'ensemble des fonctions continues pour conclure que l'on peut approcher les valeurs singulières par des polynômes.

### 3.3.2 Valeurs singulières d'une fonction continue sur le cercle à valeurs dans $\mathbb{C}^{n\times n}$

**Propriété 46.** *Soit $\mathcal{C}_1(\mathbb{C}^n)$ l'ensemble des fonctions continues $\underline{w}$ de $\mathcal{C}$ à valeurs dans $\mathbb{C}^n$ telles que $\underline{w}(\Omega)^H \underline{w}(\Omega) = 1$ pour tout $\Omega \in \mathcal{C}$. Soit $\boldsymbol{A}(\Omega)$ une fonction continue de $\mathcal{C}$ à valeurs dans $\mathbb{C}^{n\times n}$. La borne supérieure*

$$\sup_{\underline{u}(\Omega),\underline{v}(\Omega)\in\mathcal{C}_1(\mathbb{C}^n)} \Re\left(\underline{u}(\Omega)^H \boldsymbol{A}(\Omega)\underline{v}(\Omega)\right) \tag{3.12}$$

*est atteinte en une fonction continue de $\mathcal{C}$ à valeurs dans $\mathbb{R}$ que nous appellerons $\sigma(\Omega)$.*

Notons $\mathcal{C}(\mathbb{R})$ l'ensemble des fonctions continues sur $\mathcal{C}$ à valeurs dans $\mathbb{R}$. Soit $\psi : \mathcal{C}_1(\mathbb{C}^n) \times \mathcal{C}_1(\mathbb{C}^n) \mapsto \mathcal{C}(\mathbb{R})$, définie pour tout $\Omega \in \mathcal{C}$ par

$$\psi(\underline{u}, \underline{v})(\Omega) = \Re\left(\underline{u}(\Omega)^H \boldsymbol{A}(\Omega)\underline{v}(\Omega)\right) \tag{3.13}$$

$\psi$ est une fonction continue, $\mathcal{C}_1(\mathbb{C}^n) \times \mathcal{C}_1(\mathbb{C}^n)$ est un compact car $\mathcal{C}_1(\mathbb{C}^n)$ est un compact (ses éléments étant des fonctions continues bornées) et que le

cercle $\mathcal{C}$ est un compact. D'après le théorème de Stone-Weierstrass, $\psi$ atteint ses bornes, soit : il existe $\underline{u}^\star \in \mathcal{C}_1(\mathbb{C}^n), \underline{v}^\star \in \mathcal{C}_1(\mathbb{C}^n)$ tels que

$$\sigma(\Omega) = \underline{u}^\star(\Omega)^H \boldsymbol{A}(\Omega)\underline{v}^\star(\Omega) = \max_{\underline{u},\underline{v}\in\mathcal{C}_1(\mathbb{C}^n)} \Re\left(\underline{u}(\Omega)^H \boldsymbol{A}(\Omega)\underline{v}(\Omega)\right)$$

et $\sigma \in \mathcal{C}(\mathbb{R})$ est continue. □

Intéressons-nous maintenant à ce qu'il est convenu d'appeler la deuxième valeur singulière d'une matrice continue $\boldsymbol{A}(\Omega)$ qui est une fonction continue de $\mathcal{C}$ à valeurs dans $\mathbb{C}^{n\times n}$. Comme pour le cas des matrices constantes, on va procéder par déflation :

**Propriété 47.** *Soit $\boldsymbol{A}_{(2)} = \boldsymbol{A} - \sigma_1\underline{v}_1\underline{u}_1^H$ où $(\sigma_1, \underline{v}_1, \underline{u}_1)$ est solution du problème d'optimisation (3.12). Soit $(\sigma_2, \underline{v}_2, \underline{u}_2)$ la solution du problème d'optimisation relatif à $\boldsymbol{A}_{(2)}$, ie $\sigma_2 = \underline{u}_2^H\boldsymbol{A}_{(2)}\underline{v}_2 = \max_{\underline{u},\underline{v}\in\mathcal{C}_1(\mathbb{C}^n)} \Re\left(\underline{u}^H\boldsymbol{A}_{(2)}\underline{v}\right)$ alors en notant $0_\mathcal{C}$ la fonction nulle et $1_\mathcal{C}$ la fonction unité sur le cercle, on a :*

*i.* $\forall \underline{v} \in \mathcal{C}_1(\mathbb{C}^n), \underline{v}^H\boldsymbol{A}_{(2)}\underline{u}_1 = 0_\mathcal{C}$.

*ii.* $\forall \underline{u} \in \mathcal{C}_1(\mathbb{C}^n), \underline{u}^H\boldsymbol{A}_{(2)}^H\underline{v}_1 = 0_\mathcal{C}$.

*iii.* $\underline{u}_1^H\underline{u}_2 = 0_\mathcal{C}$ *et* $\boldsymbol{A}\underline{u}_2 = \boldsymbol{A}_{(2)}\underline{u}_2 = \sigma_2\underline{u}_2$.

*iv.* $\underline{v}_1^H\underline{v}_2 = 0_\mathcal{C}$ *et* $\boldsymbol{A}^H\underline{v}_2 = \boldsymbol{A}_{(2)}^H\underline{v}_2 = \sigma_2\underline{u}_2$.

$\boldsymbol{A}_{(2)}$ étant une fonction continue, on peut lui appliquer les résultats de la propriété 46, à savoir $\underline{u}_2$ et $\underline{v}_2$ sont des éléments de $\mathcal{C}_1(\mathbb{C}^n)$ et $\sigma_2$ est une fonction continue de $\mathcal{C}$ à valeurs dans $\mathbb{R}$.

Soit $\underline{u} \in \mathcal{C}_1(\mathbb{C}^n)$, posons $\underline{u}(\Omega) = \alpha(\Omega)\underline{u}_1(\Omega) + \beta(\Omega)\underline{w}(\Omega), \forall\Omega \in \mathcal{C}$ avec $\alpha, \beta \in \mathcal{C}(\mathbb{R}), \underline{w} \in \mathcal{C}_1(\mathbb{C}^n)$ et $\underline{w}^H\underline{u}_1 = 0_\mathcal{C}$. On a :

$$\boldsymbol{A}_{(2)}\underline{u} = (\boldsymbol{A} - \sigma_1\underline{v}_1\underline{u}_1^H)(\alpha\underline{u}_1 + \beta\underline{w}) = \alpha\boldsymbol{A}\underline{u}_1 + \beta\boldsymbol{A}\underline{w} - \sigma_1\alpha\underline{v}_1\underline{u}_1^H\underline{u}_1 - \sigma_1\beta\underline{v}_1\underline{u}_1^H\underline{w}$$

Or $\underline{u}_1^H\underline{u}_1 = 1_\mathcal{C}$ et $\boldsymbol{A}\underline{u}_1 = \sigma_1\underline{v}_1$, d'où $\boldsymbol{A}_{(2)}\underline{u} = \beta\boldsymbol{A}\underline{w}$. En particulier, prenons $\underline{u} = \underline{u}_1$, dans ce cas $\underline{w} = \underline{0}$ et donc $\boldsymbol{A}_{(2)}\underline{u}_1 = 0_\mathcal{C}$, d'où, $\forall\underline{v} \in \mathcal{C}_1(\mathbb{C}^n), \underline{v}^H\boldsymbol{A}_{(2)}^H\underline{u}_1 = 0_\mathcal{C}$. Eu utilisant, $\boldsymbol{A}^H\underline{v}_1 = \sigma_1\underline{v}_1$, on obtient par le même raisonnement ii.

Choisissons maintenant $\underline{v} = \underline{v}_2$ dans i : $\underline{v}_2^H\boldsymbol{A}_{(2)}\underline{u}_1 = 0_\mathcal{C} = \sigma_2\underline{u}_2^H\underline{u}_1$, soit $\underline{u}_2^H\underline{u}_1 = 0_\mathcal{C}$. Finalement, $\boldsymbol{A}\underline{u}_2 = (\boldsymbol{A}_{(2)} + \sigma_1\underline{v}_1\underline{u}_1^H)\underline{u}_2 = \boldsymbol{A}_{(2)}\underline{u}_2 + \sigma_1\underline{v}_1\underline{u}_1^H\underline{u}_2 = \boldsymbol{A}_{(2)}\underline{u}_2$, et de même pour iv. □

**Théorème 1.** *Soit $\boldsymbol{A}(\Omega)$ une fonction continue de $\mathcal{C}$ à valeurs dans $\mathbb{C}^{n\times n}$, alors il existe $\boldsymbol{U}(\Omega)$ et $\boldsymbol{V}(\Omega)$ des fonctions continues de $\mathcal{C}$ à valeurs dans $\mathbb{C}^{n\times n}$ telles que $\boldsymbol{U}(\Omega)\boldsymbol{U}(\Omega)^H = \boldsymbol{I}$ et $\boldsymbol{V}(\Omega)\boldsymbol{V}(\Omega)^H = \boldsymbol{I}$ pour tout $\Omega$ dans $\mathcal{C}$ et*

$$\boldsymbol{A}(\Omega) = \boldsymbol{V}(\Omega)\boldsymbol{\Sigma}(\Omega)\boldsymbol{U}(\Omega)^H$$

*avec $\boldsymbol{\Sigma}(\Omega) = \mathrm{diag}\{\sigma_i(\Omega)\}$ et $\sigma_i \in \mathcal{C}(\mathbb{R}), i = 1$ à $n$.*

On définit $\boldsymbol{A}_{(i+1)} = \boldsymbol{A}_{(i)} - \sigma_i \underline{v}_i \underline{u}_i^H$ où $(\sigma_i, \underline{v}_i, \underline{u}_i)$ est solution du problème d'optimisation (3.12) relatif à $\boldsymbol{A}_{(i)}$, pour $i = 1$ à $n$, avec $\boldsymbol{A}_{(1)} = \boldsymbol{A}$. Il suffit alors de procéder par récurrence sur les $\boldsymbol{A}_{(i)}$ : la propriété est vraie pour $i = 1$ (propriété 46), et ensuite d'utiliser les résultats i à iv en remplaçant $\boldsymbol{A}_{(2)}$ par $\boldsymbol{A}_{(i+1)}$ (ce que nous n'avons pas fait pour alléger quelque peu les notations !) pour obtenir la décomposition désirée. □

### 3.3.3 Diagonalisation d'une matrice $\boldsymbol{A}\boldsymbol{A}^H$ continue sur le cercle

**Théorème 2.** *Soit $\boldsymbol{A}(\Omega)$ une fonction continue de $\mathcal{C}$ à valeurs dans $\mathbb{C}^{n\times n}$. Soit $\boldsymbol{H}(\Omega)$ la fonction continue définie par $\boldsymbol{H}(\Omega) = \boldsymbol{A}(\Omega)\boldsymbol{A}(\Omega)^H$, alors il existe une fonction continue $\boldsymbol{V}(\Omega)$ de $\mathcal{C}$ à valeurs dans $\mathbb{C}^{n\times n}$ telle que $\boldsymbol{V}(\Omega)\boldsymbol{V}(\Omega)^H = \boldsymbol{I}$ pour tout $\Omega$ dans $\mathcal{C}$ et*

$$\boldsymbol{H}(\Omega) = \boldsymbol{V}(\Omega)\boldsymbol{\Lambda}(\Omega)\boldsymbol{V}(\Omega)^H$$

*avec $\boldsymbol{\Lambda}(\Omega) = \mathrm{diag}\{\lambda_i(\Omega)\}, \lambda_i \in \mathcal{C}(\mathbb{R})$ et $\lambda_i \geq 0, i = 1$ à $n$.*

Soient $\boldsymbol{U}$ et $\boldsymbol{V}$ les matrices para-unitaires (sur le cercle) permettant la décomposition en valeurs singulières de $\boldsymbol{A}$ : $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{U}^H$. Par définition de $\boldsymbol{H}$ on a $\boldsymbol{H} = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{U}^H\boldsymbol{U}\boldsymbol{\Sigma}^H\boldsymbol{V}^H = \boldsymbol{V}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H\boldsymbol{V}^H$, d'où $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}\boldsymbol{\Sigma}^H$ est une matrice diagonale dont les éléments diagonaux sont les fonctions continues $\lambda_i = |\sigma_i|^2 \geq 0$. □

On a ainsi montré que toute matrice continue para-hermitienne définie positive sur le cercle est diagonalisable sur le cercle à l'aide d'une matrice para-unitaire continue. Peut-on pour autant en déduire que toute matrice L-polynomiale para-hermitienne est diagonalisable à l'aide d'une matrice L-polynomiale para-unitaire ? Le problème est que même si la matrice $\boldsymbol{A}$ est L-polynomiale, on ne peut garantir que les valeurs singulières définies par (46) sont elles-mêmes L-polynomiales car l'ensemble des polynômes de Laurent n'est pas fermé. Nous allons illustrer ceci à l'aide d'un exemple 2x2.

**Exemple de matrice L-polynomiale para-hermitienne qui n'admet pas de "valeurs propres" L-polynomiales**

Soit $\boldsymbol{H}(z) = \begin{bmatrix} 1 & 1 \\ 1 & -2z^{-1} + 6 - 2z \end{bmatrix} = \boldsymbol{H}^T(z^{-1})$, $\boldsymbol{H}(z)$ est une matrice para-hermitienne de déterminant : $\det \boldsymbol{H}(z) = -2z^{-1} + 5 - 2z = -2z^{-1}(1 - \frac{5}{2}z + z^2)$. $\det \boldsymbol{H}(e^{j\omega}) = -4\cos\omega + 5 \geq 0, \forall\omega$.

La L-forme de Smith de $\boldsymbol{H}(z)$ est $\boldsymbol{S}(z) = \begin{bmatrix} 1 & 0 \\ 0 & 1 - \frac{5}{2}z + z^2 \end{bmatrix}$ (calculé à partir des L-pgcd des mineurs).

Supposons qu'il existe $\boldsymbol{U}(z)$ L-polynomiale para-unitaire telle que

$$\boldsymbol{U}^H(\frac{1}{z^*})\boldsymbol{H}(z)\boldsymbol{U}(z) = \boldsymbol{\Lambda}(z)$$

D'après le corollaire 40, $\boldsymbol{U}(z)$ est également L-unimodulaire et $\boldsymbol{S}(z)$ et $\boldsymbol{\Lambda}(z)$ sont L-équivalentes. D'après la propriété 20, le pgcd normalisé des mineurs de ces deux matrices sont donc égaux.

Soit donc ici $pgcd(\lambda_1(z), \lambda_2(z)) = 1$ (ces deux L-polynômes sont premiers entre eux) et $\lambda_1(z)\lambda_2(z) = c'z^{\alpha'} \det \boldsymbol{H}(z), c' \in \mathbb{C}$ et $\alpha' \in \mathbb{Z}$. On a par ailleurs montré (propriété 32) que si une solution L-polynomiale existait alors les $\lambda_i$ étaient para-hermitiens, et donc si $z_0$ est racine du polynôme, $\frac{1}{z_0^*}$ est également racine. Les seules solutions possibles sont donc (à une permutation d'indice près) $\lambda_1(z) = cz^\alpha$ et $\lambda_2(z) = dz^\beta(1 - \frac{5}{2}z + z^2)$. $\lambda_1(z)$ para-hermitien implique que $cz^\alpha = c^*z^{-\alpha}$, soit que $\lambda_1(z) = c, c \in \mathbb{R}^*$.

On cherche donc si $\boldsymbol{H}(z)\underline{v}(z) = c\underline{v}(z), c \in \mathbb{R}^*$ admet une solution L-polynomiale en $\underline{v}(z)$ :

$$\boldsymbol{H}(z)\underline{v}(z) = c\underline{v}(z) \Leftrightarrow \begin{cases} (1-c)v_1(z) + v_2(z) = 0 \\ v_1(z) + (-2z^{-1} + 6 - c - 2z)v_2(z) = 0 \end{cases}$$

soit $v_1(z) + (c-1)(-2z^{-1} + 6 - c - 2z)v_1(z) = 0$, soit puisque $v_1(z) \neq 0$, $-2(c-1)z^{-1} + (-c^2 + 7c - 5) - 2(c-1)z = 0$ pour tout $z$, ce qui est impossible avec $c$ constant.

Une solution polynomiale n'existe pas dans le cas général, mais on peut montrer que les valeurs singulières et donc les valeurs propres d'une matrice para-hermitienne L-polynomiale peuvent être approchées par des L-polynômes.

### 3.3.4 Presque-diagonalisation d'une matrice para-hermitienne

**Lemme 48.** *Les vecteurs pour lesquels le maximum de la fonction $\psi$ définie en (3.13) est atteint peuvent être approchés par des vecteurs polynomiaux :*

$$\forall \varepsilon \in \mathbb{R}^+,\ \exists\ \underline{p}, \in \mathbb{C}^n[z, z^{-1}]\ tq\ \sup_{\Omega \in \mathcal{C}}\ \sup_{i=1 \grave{a}\ n} |u_i^\star(\Omega) - p_i(\Omega)| < \varepsilon$$
$$\forall \varepsilon \in \mathbb{R}^+,\ \exists\ \underline{q}, \in \mathbb{C}^n[z, z^{-1}]\ tq\ \sup_{\Omega \in \mathcal{C}}\ \sup_{i=1\ \grave{a}\ n} |v_i^\star(\Omega) - q_i(\Omega)| < \varepsilon$$

Chaque composante $u_i^\star(\Omega)$ du vecteur $\underline{u}^\star(\Omega)$ est une fonction continue de $\mathcal{C}$ sur $\mathbb{R}$. D'après la propriété 13, elle peut donc être approchée par un polynôme $p_i$. Soit $\underline{p}(\Omega)$ le vecteur polynomial ayant pour composantes les $p_i(\Omega)$, il est clair qu'il vérifie le lemme ci-dessus. On raisonne de même pour $\underline{v}^\star$ et $\underline{q}$. □

Les vecteurs polynomiaux ne sont pas uniques, non seulement parce que les $p_i$ ne le sont pas, mais aussi car si $\{\underline{p}, \underline{q}\}$ est solution, alors $\{\underline{p}e^{j\phi}, \underline{q}e^{-j\phi}\}$ avec $\phi \in \mathbb{R}$ l'est également.

**Propriété 49.** *Soit $\boldsymbol{A}(z)$ une matrice L-polynomiale appartenant à $\mathbb{C}^{n\times n}[z, z^{-1}]$. La fonction $\sigma$ pour laquelle $\psi$ atteint son maximum peut-être approchée par un polynôme :*

$$\forall \varepsilon \in \mathbb{R}^+,\ \exists s \in \mathbb{C}[z, z^{-1}]\ tq\ \sup_{\Omega\in\mathcal{C}} |\sigma(\Omega) - s(\Omega)| < \varepsilon$$

En effet, posons $s(\Omega) = \underline{p}(\Omega)\boldsymbol{A}(\Omega)\underline{q}(\Omega), \forall \Omega \in \mathcal{C}$, où $\underline{p}$ et $\underline{q}$ sont des approximations polynomiales (à la précision désirée) des solutions $\underline{u}^\star$ et $\underline{v}^\star$ du problème d'optimisation. Alors, en écrivant :

$$\sigma - s = \underline{u}^{\star H}\boldsymbol{A}\underline{v}^\star - \underline{p}^H\boldsymbol{A}\underline{q} = (\underline{u}^\star - \underline{p})^H\boldsymbol{A}(\underline{v}^\star - \underline{q}) + (\underline{u}^\star - \underline{p})^H\boldsymbol{A}\underline{q} + \underline{p}^H\boldsymbol{A}(\underline{v}^\star - \underline{q})$$

et en remarquant que $\boldsymbol{A}$ étant polynomiale elle est bornée sur $\mathcal{C}$, par choix de $\underline{p}$ et $\underline{q}$, on peut avoir $\sup_{\Omega\in\mathcal{C}} |\sigma(\Omega) - s(\Omega)|$ aussi petit que l'on veut. □

**Propriété 50.** *Soit $\boldsymbol{H}(z)$ une matrice L-polynomiale para-hermitienne, alors il existe $n$ vecteurs L-polynomiaux $\underline{v}_i(z)$ et $n$ L-polynômes $\lambda_i(z)$ tels que*

$$\boldsymbol{H}(z)\underline{v}_i(z) \sim \lambda_i(z)\underline{v}_i(z)\ avec\ \underline{v}_i(z)^H\underline{v}_j(z) \lesssim \delta_{ij}, \forall z \in \mathcal{C}$$

On utilise les deux propriétés précédentes pour chacune des valeurs singulières. La continuité étant uniforme, on obtient le résultat énoncé.

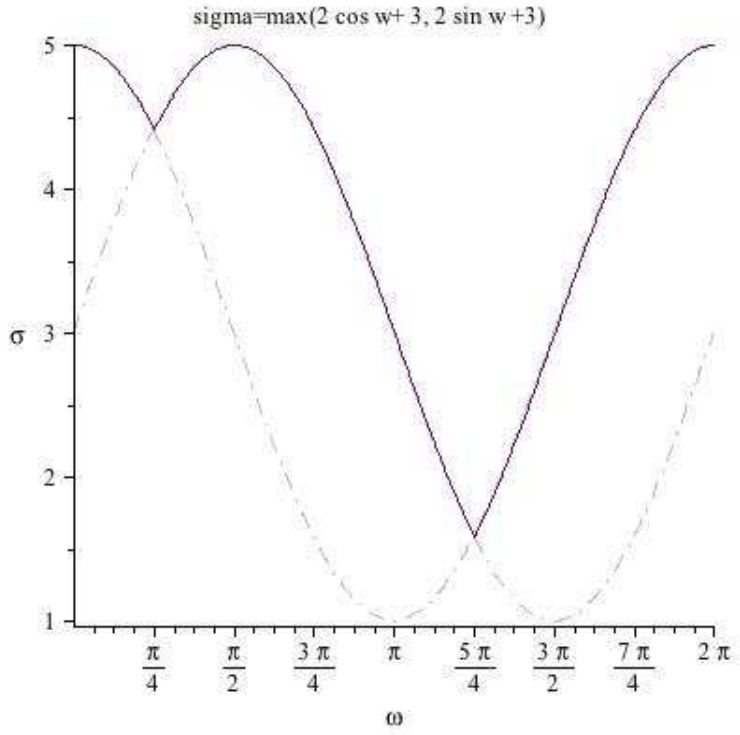
Remarquons que les propriétés de densité ne nous permettent pas d'avoir une information quelconque sur le L-degré du polynôme qui peut être très grand. Pour ceci, illustrons d'abord le problème de switch rencontré lorsqu'on utilise la maximisation.

### Exemple illustrant le problème de switch entre "valeurs propres"

Considérons la matrice diagonale (para-hermitienne) :
$\boldsymbol{A}(z) = \begin{bmatrix} z^{-1} + 3 + z & 0 \\ 0 & jz^{-1} + 2 - jz \end{bmatrix} = \boldsymbol{A}_1 z^{-1} + 3\boldsymbol{I} + \boldsymbol{A}_1^H z$, alors
$\boldsymbol{A}(e^{j\omega}) = \begin{bmatrix} 2\cos\omega + 3 & 0 \\ 0 & 2\sin\omega + 3 \end{bmatrix}$ (matrice définie positive $\forall\omega$).

Dans le cas d'une matrice diagonale 2×2, on peut montrer que
$\sigma(\boldsymbol{A}, \omega) = \max\{|a_{11}(e^{j\omega})|, |a_{22}(e^{j\omega})|\}$, ce qui nous donne ici
$\sigma(\boldsymbol{A}, \omega) = \max\{|2\cos\omega + 3|, |2\sin\omega + 3|\}$, soit

$\sigma(\boldsymbol{A}, \omega) = 3 + 2\cos\omega = 3 + e^{j\omega} + e^{-j\omega}$ pour $\omega \in [0, \pi/4]$
$\sigma(\boldsymbol{A}, \omega) = 3 + 2\sin\omega = 3 - j(e^{j\omega} - e^{-j\omega})$ pour $\omega \in [\pi/4, 5\pi/4]$
$\sigma(\boldsymbol{A}, \omega) = 3 + 2\cos\omega = 3 + e^{j\omega} + e^{-j\omega}$ pour $\omega \in [5\pi/4, 9\pi/4]$
...




La matrice $\boldsymbol{A}(e^{j\omega})$ étant diagonale, ses "valeurs propres" (au sens éléments de la diagonale puisque la matrice est diagonale et que $\boldsymbol{A}\underline{v} = \lambda\underline{v}$) sont clairement $3 + 2\cos\omega$ et $3 + 2\sin\omega$ pour tout $\omega$ ... et l'on a bien $\boldsymbol{A}(z)\begin{bmatrix}1\\0\end{bmatrix} = (z^{-1} + 3 + z)\begin{bmatrix}1\\0\end{bmatrix}$ et $\boldsymbol{A}(z)\begin{bmatrix}0\\1\end{bmatrix} = (jz^{-1} + 3 - jz)\begin{bmatrix}0\\1\end{bmatrix}$ $\forall z$.

Le problème de la valeur singulière qui switche est dû au fait qu'on veut le maximum pour chaque $\omega$, cela ne signifie pas que les valeurs propres elles-mêmes switchent... (problème de définition de la valeur singulière). La fonction $\sigma(\boldsymbol{A}, \omega)$ ainsi définie est bien continue, mais elle n'est pas dérivable. Le polynôme approchant $\sigma$ risque donc d'avoir un L-ordre très grand !

Le choix d'une solution garantissant une fonction la plus lisse possible pour $\sigma$ fera l'objet de travaux ultérieurs.

## 3.4 Matrices polynomiales et égalisation

Le problème qui a motivé nos travaux sur les matrices polynomiales est celui de l'égalisation aveugle de canal. Le but de cette partie n'est pas de faire un état de l'art sur l'égalisation aveugle mais de montrer où interviennent les matrices polynomiales dans cette problématique. Dans toute la suite, nous supposerons que le canal est linéaire et invariant dans le temps.

### 3.4.1 Problème de l'égalisation aveugle de canaux de communication

Le problème d'égalisation de canaux de communication est un problème classique en traitement du signal. Supposons que l'on ait un système MIMO comportant $n$ antennes émettrices et $p$ antennes réceptrices. En présence de multi-trajets et en présence de bruit $\underline{b}_k$, les $n$ signaux sources $s_i$ sont reliés aux $p$ signaux reçus $w_i$ par :

$$\underline{w}_k = \sum_j \boldsymbol{C}_j \underline{s}_{k-j} + \underline{b}_k \tag{3.14}$$

où $\{\boldsymbol{C}_k, k \in \mathbb{Z}, \boldsymbol{C}_k \in \mathbb{C}^{p\times n}\}$ sont les coefficients (matriciels) de la réponse impulsionnelle du canal, $\underline{s}_k$ est le vecteur source à l'instant $k$ dont les composantes (en bande de base) sont $s_i(k) \in \mathbb{C}, i = 1$ à $n$ , $\underline{w}_k$ le vecteur des observations dont les composantes sont $w_i(k) \in \mathbb{C}, i = 1$ à $p$ et $\underline{b}_k$ est le vecteur bruit à l'instant $k$.

Le problème de l'égalisation aveugle MIMO est de trouver un filtre appelé égaliseur qui, à partir des seules observations, permet de retrouver les signaux sources. Ce problème se pose lorsque les signaux sources sont inconnus, soit dans le cas de signaux interceptés, soit lorsqu'on supprime les séquences d'apprentissage afin de gagner du débit utile. Il est clair que la solution du problème ainsi posé n'est pas unique. Prenons la transformée en $z$ de l'équation (3.14) en l'absence de bruit :

$$\underline{w}(z) = \boldsymbol{C}(z)\underline{s}(z) \tag{3.15}$$

où $\boldsymbol{C}(z) = \sum_k \boldsymbol{C}_k z^{-k}$, $\underline{s}(z) = \sum_k \underline{s}_k z^{-k}$ et $\underline{w}(z) = \sum_k \underline{w}_k z^{-k}$.
Si $\{\boldsymbol{C}(z), \underline{s}(z)\}$ est solution du problème, alors $\{\boldsymbol{C}(z)\boldsymbol{T}(z), \boldsymbol{M}(z)\underline{s}(z)\}$ tel que $\boldsymbol{T}(z)\boldsymbol{M}(z) = \boldsymbol{I}$ l'est également.

Certaines hypothèses sont alors classiquement faites sur les signaux sources, ce sont des signaux indépendants, identiquement distribués, de variance unité et de moyenne nulle, une source au plus est gaussienne. La détermination de l'égaliseur peut ainsi se faire à une permutation, un facteur d'échelle, et de retards purs près.

### 3.4.2 Mélanges instantanés

Lorsque la réponse impulsionnelle du canal a un seul élément non nul, on se trouve dans le cas de mélanges instantanés. La transformée en $z$ de l'équation (3.14) se réduit à :

$$\underline{w}(z) = \boldsymbol{C}\underline{s}(z) \text{ avec } \boldsymbol{C} \in \mathbb{C}^{p\times n} \tag{3.16}$$

Avec les hypothèses d'indépendance précédentes sur les signaux sources, on se retrouve alors avec un problème d'Analyse en Composantes Indépendantes. Il est désormais bien connu que ce problème ne peut pas se résoudre à l'aide des statistiques d'ordre 2 mais nécessite l'utilisation des statistiques d'ordre supérieur [40, 42, 51]. En présence de bruit additif de densité de probabilité inconnue, l'égalisation peut ainsi se formaliser grâce à la minimisation d'un critère appelé contraste. Parmi les algorithmes proposés pour résoudre ce problème [42], citons notamment un algorithme permettant la diagonalisation conjointe des matrices cumulantes : Joint Approximation Diagonalisaton of Eigenmatrices (JADE) [39, 38], celui-ci permet de trouver une matrice unitaire qui rend l'ensemble des matrices cumulantes des observations le plus diagonales possible et donc les signaux en sortie de l'égaliseur le plus indépendants possibles. L'égalisation ne se fait alors qu'à une permutation près, une permutation n'ayant aucune influence sur l'indépendance.

### 3.4.3 Mélanges convolutifs FIR

Lorsque le canal est à réponse impulsionnelle finie, la somme $\sum_k \boldsymbol{C}_k z^{-k}$ comporte un nombre fini d'éléments et $\boldsymbol{C}(z)$ est une matrice polynomiale appartenant à $\mathbb{C}^{p\times n}[z^{-1}]$ : $\boldsymbol{C}(z) = \sum_{k=0}^{l} \boldsymbol{C}_k z^{-k}$. $l+1$ est parfois appelé longueur du canal (alors que $l$ est le degré de la matrice polynomiale en $z^{-1}$).

Intéressons-nous maintenant au spectre des signaux d'entrée $\boldsymbol{\Gamma}_s(z)$ et de sortie $\boldsymbol{\Gamma}_w(z)$ :

$$\boldsymbol{\Gamma}_w(z) = \boldsymbol{C}(z)\boldsymbol{\Gamma}_s(z)\boldsymbol{C}^H(\frac{1}{z^*})$$

Sous les hypothèses classiques (sources indépendantes, centrées et de variance unité), on a $\boldsymbol{\Gamma}_s(z) = \boldsymbol{I}_n$ et le spectre des signaux observés est donc :

$$\boldsymbol{\Gamma}_w(z) = \boldsymbol{C}(z)\boldsymbol{C}^H(\frac{1}{z^*}) = (\sum_{k=0}^{l} \boldsymbol{C}_k z^{-k})(\sum_{k=0}^{l} \boldsymbol{C}_k^H z^k) \tag{3.17}$$

$\boldsymbol{\Gamma}_w(z)$ est alors une matrice L-polynomiale appartenant à $\mathbb{C}^{p\times p}[z, z^{-1}]$. De plus, $\widetilde{\boldsymbol{\Gamma}}_w(z) = \boldsymbol{C}(z)\widetilde{\boldsymbol{C}}(z) = \boldsymbol{\Gamma}_w(z)$, soit $\boldsymbol{\Gamma}_w(z)$ est une matrice L-polynomiale para-hermitienne d'ordre $2l$.

Supposons maintenant que les signaux observés ont été blanchis à l'ordre 2, c'est à dire que les signaux $\underline{w}$ sont filtrés à l'aide de $\boldsymbol{B}(z)$ de façon à ce que les signaux $\underline{y}(z) = \boldsymbol{B}(z)\underline{w}(z)$ aient un spectre égal à l'identité. Le système filtre plus blanchisseur a pour fonction de transfert $\boldsymbol{D}(z) = \boldsymbol{B}(z)\boldsymbol{C}(z)$. Le spectre du signal en sortie du blanchisseur est donc $\boldsymbol{\Gamma}_y(z) = \boldsymbol{D}(z)\boldsymbol{\Gamma}_s(z)\widetilde{\boldsymbol{D}}(z) = \boldsymbol{I}$ par définition de $\boldsymbol{B}$ et donc le système canal plus blanchisseur est tel que $\boldsymbol{D}(z)\widetilde{\boldsymbol{D}}(z) = \boldsymbol{I}$, c'est à dire que $\boldsymbol{D}(z)$ est une matrice L-polynomiale para-unitaire.

La factorisation (3.7) des matrices para-unitaires pour l'égalisation de systèmes $2 \times 2$ a été tout d'abord utilisée dans [58] pour le cas réel et [59] pour le cas complexe mais l'approche temporelle remonte à 1990 [42]. Le filtrage d'un signal par une matrice unitaire ou une matrice retard du type (3.8) peut se traduire comme une relation entre les cumulants. Une méthode algébrique permet ensuite d'estimer une à une les rotations intervenant dans la factorisation.

Nous avons également exploité cette factorisation dans [2, 74] afin de déterminer un égaliseur à l'aide de l'algorithme PAFA : PAra-unitary FActorisation. Outre la factorisation de l'égaliseur para-unitaire, nous avons utilisé les cumulants d'ordre 4 des observations (blanchies) pour définir un critère de contraste [41, 40] que nous avons maximisé. Cette maximisation est effectuée à l'aide d'un algorithme itératif de type Jacobi avec des matrices de rotations de Givens. Les résultats obtenus ont été comparés à ceux de l'algorithme PAJOD (Partial Approximate JOint Diagonalization) de [43] et à sa version améliorée dans [4]. Cet algorithme, qui ne donne pas un égaliseur para-unitaire mais seulement semi-unitaire, utilise également une diagonalisation conjointe des matrices cumulantes.

Comme toutes les méthodes reposant sur des contrastes, l'égalisation ne se fait qu'à une matrice de permutation et une matrice diagonale de retards près.

### 3.4.4 Algorithmes de diagonalisation de L-matrices

McWhriter et co-auteurs [35, 62, 44] ou [76] se sont intéressés à la décomposition en valeurs singulières et à la diagonalisation de matrices polynomiales. Pour simplifier, comme il est possible de diagonaliser une matrice hermitienne constante à l'aide de rotations, on fait de même pour les matrices L-polynomiales para-hermitiennes mais avec des matrices para-unitaires élémentaires. Ces matrices sont de la forme :

$$\boldsymbol{G}(z) = \boldsymbol{U}(i, j, \theta, \phi)\boldsymbol{B}^{(j,t)}(z) \tag{3.18}$$

où $\boldsymbol{U}(i,j,\theta,\phi)$ est une matrice de rotation de Givens élémentaire définie en (3.9) et $\boldsymbol{B}^{(j,t)}(z)$ est une matrice de retard élémentaire :

$\boldsymbol{B}^{(j,t)}(z) = \begin{bmatrix} \boldsymbol{I}_{j-1} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & z^{-t} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{I}_{n-j} \end{bmatrix}$, $t$ pouvant être positif ou négatif.

Comme pour les algorithmes de Jacobi classiques, à chaque itération de l'algorithme, le but de la multiplication d'une matrice $\boldsymbol{R}(z)$ par une matrice élémentaire est d'éliminer le terme dominant des coefficients hors-diagonaux de $\boldsymbol{R}(z)$, mais cette fois, il s'agit de coefficients des L-polynômes situés en dehors de la diagonale : on a donc les indices $i$ et $j$ de (3.18) qui correspondent au polynôme $r_{ij}(z)$ et l'indice $t$ qui correspond à la puissance de $z$ à laquelle intervient ce terme dominant dans $r_{ij}(z)$. Ce faisant, le L-degré des élément diagonaux -qui correspondent aux "valeurs propres"- augmente. L'algorithme SBR2 (second-order Sequential Best Rotation) proposé dans [62] permet toutefois de limiter le degré des "valeurs propres" et ainsi de décorréler fortement des signaux : il suffit de filtrer les signaux reçus $\underline{w}_k$ par la matrice para-unitaire : $\boldsymbol{H}_l(z) = \boldsymbol{G}_l(z)\dots\boldsymbol{G}_2(z)\boldsymbol{G}_1(z)$ où les $\boldsymbol{G}_i(z)$ sont des matrices para-unitaires élémentaires et $l$ est déterminé par l'algorithme SBR2. Soit $\underline{v}(z) = \boldsymbol{H}_l(z)\underline{w}(z)$, alors le spectre de $\underline{v}_k$ est $\boldsymbol{\Gamma}_v(z) = \boldsymbol{H}(z)\boldsymbol{\Gamma}_v(z)\widetilde{\boldsymbol{H}}(z) \approxeq \boldsymbol{D}(z)$ où $\boldsymbol{D}(z)$ est une matrice diagonale. Une version Householder de cet algorithme a été développée dans [71].

L'exemple page 48 montre qu'une solution polynomiale exacte de

$$\boldsymbol{H}(z)\boldsymbol{\Gamma}_v(z)\widetilde{\boldsymbol{H}}(z) = \boldsymbol{\Lambda}(z)$$

où $\boldsymbol{\Lambda}(z)$ est une matrice diagonale polynomiale n'existe pas toujours. L'approximation en sortie du SBR2 n'est donc pas seulement due à la méthode itérative, contrairement à ce qu'il est parfois laissé entendre dans la littérature - voir par exemple [63] pour une extension de SBR2 aux quaternions ou pour l'extension de l'algorithme MUSIC aux signaux large-bande [33]- mais aussi à la structure de la matrice para-hermitienne que l'on cherche à diagonaliser.

# Chapitre 4

# Conclusion et perspectives

La diagonalisation des matrices polynomiales est un problème ouvert qui suscite actuellement un intérêt croissant. Notre contribution a été de poser clairement le problème, notamment en terme d'ensembles sur lesquels on travaille (l'anneau des polynômes de Laurent), de définition d'ordre et de degré lorsqu'on s'intéressait à des matrices non-causales, ce qui est le cas des matrices para-hermitiennes, et donc en particulier des matrices spectrales. On a ainsi pu montrer sur un exemple simple (2×2 d'ordre 2) qu'il n'existe pas toujours de solution exacte polynomiale à la diagonalisation d'une matrice para-hermitienne à l'aide de matrices para-unitaires. On a pu toutefois trouver une approximation continue sur le cercle permettant de presque-diagonaliser une matrice polynomiale para-hermitienne. Un exemple, là encore très simple, a montré que cette approximation, calculée à l'aide des valeurs singulières, pouvait donner un résultat d'ordre très élevé, non imposé par le système lui-même mais par la méthode de calcul et la définition adoptée. Il faut donc réfléchir à une stratégie pour le problème d'optimisation associé.

Dans ce mémoire, je n'ai pas développé les travaux auxquels j'ai participé sur les tenseurs, il aurait fallu pour cela rajouter deux chapitres sur les définitions et propriétés. Remarquons cependant que les matrices polynomiales peuvent également être vues comme des tenseurs d'ordre 3. En effet, soit $\boldsymbol{P}(z) = \sum_k \boldsymbol{P}_k z^k$, une matrice polynomiale, alors chaque coefficient de cette matrice est lui même un polynôme $p_{ij}(z) = \sum_k p_{ijk} z^k$. On peut donc associer à $\boldsymbol{P}(z)$ le tenseur $p_{ijk}$.

De nombreuses questions sur les propriétés des matrices polynomiales de Laurent restent actuellement ouvertes ainsi que sur leurs applications potentielles en traitement du signal. J'aimerais en citer ici quelques unes.

Il serait intéressant de trouver dans quelles conditions (sur la structure, les multiplicités ?) il existe une solution polynomiale au problème de la diago-

nalisation d'une matrice para-hermitienne par transformation para-unitaire. Lorsqu'une solution polynomiale existe, il faudrait s'assurer si ces conditions éventuelles sont vérifiées lors des applications de la PSVD [70, 33, 63]. Cependant, même si ces conditions ne le sont pas, nous avons bien conscience qu'une solution approchée peut donner des résultats satisfaisants.

Il existe des techniques d'accélération qu'on peut mettre en oeuvre lors de la diagonalisation de matrices hermitiennes (par exemple les techniques de décalage, consistant à retrancher une matrice proportionnelle à l'identité, de manière explicite ou implicite). Ces techniques peuvent-elles être transposées aux matrices para-hermitiennes ?

Pour l'étude des systèmes, de nombreuses propriétés peuvent être démontrées en utilisant soit les matrices polynomiales (numérateur, dénominateur), soit les matrices systèmes, soit une réalisation d'état (éventuellement minimale) associée, soit les matrices de Toeplitz ou de Hankel associées aux systèmes étudiés. L'information dont on dispose est la même, mais elle est présentée différemment suivant l'option choisie et à ce titre utilise des outils spécifiques. Il est fort probable qu'il en soit de même avec les tenseurs d'ordre 3 et les matrices de Laurent. Cette piste est intéressante à creuser et permettrait de faire le lien entre la PSVD (décomposition en valeur singulières de matrices polynomiales) [62] et la HOSVD (décompositions en valeurs singulières de tenseurs d'ordre supérieur) [60]. Ce lien n'est pas trivial : en effet, nous avons vu qu'une matrice para-unitaire est représentée par une collection de matrices unitaires assortie de contraintes supplémentaires. Il s'agit donc d'une décomposition de Tucker plus contrainte que la HOSVD.

Au delà des problèmes de diagonalisation, égalisation, définition, je pourrais pour conclure ce résumé de mes travaux faire le pont entre les acquis dûs à ma formation, mes recherches et mon enseignement. Au départ, l'automatique a été pour moi un sujet de prédilection : c'était un vaste champ de jeux intellectuel permettant de faire le lien entre des mathématiques et des applications, sans lesquelles les mathématiques perdent à mes yeux un peu de leur intérêt. L'enseignement m'a quant à lui permis d'essayer de faire passer des messages parfois complexes à l'aide d'exemples simples, un sacré enjeu de la pédagogie, demandant de la rigueur, de l'imagination et de l'énergie ! La recherche demande les mêmes exigences, et tout autant de temps de réflexion. L'encadrement d'étudiants niveau master ou doctorat est encore autre chose : il se situe à la frontière entre l'enseignement et la recherche. Ces étudiants ont une réelle motivation, leur enthousiasme est communicatif, même s'il faut parfois le modérer, et nous pousse à voir les problèmes différemment, leurs compétences sont source d'enrichissement mutuel.

Enfin, pouvoir concilier enseignement et recherche sans faire l'un au dé-

triment de l'autre (et réciproquement) ni au détriment de nos étudiants est un souci quotidien pour l'enseignante-chercheuse que je suis et pour nombre de mes collègues, mais c'est aussi un des charmes de notre métier.

# Chapitre 5

# Bibliographie

Cette section regroupe tout d'abord mes publications par catégorie puis les références citées dans le manuscrit.

## 5.1 Publications

### Revues avec comités de lecture


[1] S. Icart and J.F. Lafay. Design of a minimal precompensator for the decoupling problem. *Systems and Control Letters*, 10(5) :325 – 332, 1988.

[2] S. Icart, P. Comon, and L. Rota. Blind paraunitary equalization. *Signal Process.*, 89(3) :283–290, mars 2009.

[3] M. Sørensen, L. De Lathauwer, P. Comon, S. Icart, and L. Deneire. Canonical polyadic decomposition with a columnwise orthonormal factor matrix. *SIAM Journal on Matrix Analysis and Applications*, 33(4) :1190–1213, 2012.

[4] M. Sørensen, L. De Lathauwer, S. Icart, and L. Deneire. On Jacobi-type methods for blind equalization of paraunitary channels. *Signal Process.*, 92(3) :617–624, 2012.


### Conférences internationales avec comités de lecture


[5] S. Icart, J. F. Lafay, and M. Malabre. A unified study of the fixed modes of systems decoupled via regular static state feedback. In *Proceedings of Conference on New Trends in System Theory*, pages 425–432, Gênes, Italie, juillet 1990.

[6] J.F. Lafay, P. Zagalak, A. Herrera, and S. Icart. Structural results about the interactor. In *Proceedings of the 29th IEEE Conference on Decision and Control*, pages 1048 –1049 vol.2, dec 1990.

[7] S. Icart and J.F. Lafay. Decoupling with stability for a class of linear systems via static state feedback. In *European Control Conference*, pages 1288–1291, Grenoble, juillet 1991.

[8] S. Icart, J. Leblond, and C. Samson. Stabilizing law for a flexible arm in torsion. In *Proceedings of IEEE International Conference on Robotics and Automation*, pages 794 –799 vol.1, Nice, may 1992.

[9] O. Lemoine and S. Icart. Détection de cibles par méthode CFAC et transformée de gabor discrète. In *Proceedings of GRETSI'93*, pages 1243–46, Juan-lesPins, sept 1993.

[10] O. Lemoine, S. Icart, and C. Bérenguer. Detection of transient signals using time-frequency distributions. In *Proceedings of EUSIPCO'94*, pages 272–275, Edinburgh, Scotland, sept 1994.

[11] S. Icart and R. Gautier. Blind separation of convolutive mixtures using second and fourth order moments. In *Proceedings of ICASSP'96*, volume 5, pages 3018–21, Atlanta, Georgia, may 1996.

[12] S. Icart, J. LeRoux, L. Pronzato, E. Thierry, and A. A. Zhigljavsky. Blind equalization in presence of bounded errors. In *Proceedings of ICASSP'97*, volume 5, pages 3949–3952, Munich, apr 1997.

[13] S. Icart, J. LeRoux, L. Pronzato, E. Thierry, and A. A. Zhigljavsky. Blind equalization in presence of bounded errors with unknown bound. In *Proceedings of IEEE Workshop SPAWC'97*, pages 41–44, Paris, apr 1997.

[14] J. LeRoux and S. Icart. System frequency response phase estimation based on skewness or kurtosis maximization. In *Proceedings of ICASSP'00*, volume 1, pages 564–567, Turkey, june 2000.

[15] L. Rota, P. Comon, and S. Icart. Blind equalization of MIMO channels. In *Proceedings of SPAWC'03*, pages 314–318, Rome, Italy, june 2003.

[16] L. Rota, P. Comon, and S. Icart. Blind MIMO paraunitary equalizer. In *Proceedings of ICASSP'03*, volume 4, Hong-Kong, apr 2003.

[17] L. Rota, P. Comon, and S. Icart. Egalisation aveugle de mélanges convolutifs MIMO. In *Proceedings of GRETSI'03*, Paris, sept 2003.

[18] M. Sørensen, S. Icart, P. Comon, and L. Deneire. Gradient based Approximate Joint Diagonalization by Orthogonal Transforms. In *Proceedings of EUSIPCO'08*, pages session L11–1, Lausanne, Suisse, aug 2008.

[19] M. Sørensen, P. Comon, S. Icart, and L. Deneire. Approximate tensor diagonalization by invertible transforms. In *Proceedings of EUSIPCO'09*, page 4, Glasgow, Royaume-Uni, aug 2009.

[20] M. Sørensen, P. Comon, S. Icart, and L. Deneire. PARAFAC2 receivers for Orthogonal Space-Time Block Codes. In *Proceedings of ICC'09*, pages 1–5, Dresden, Allemagne, june 2009.

[21] S. Icart and P. Comon. Some properties of Laurent polynomial matrices. In *9th IMA International Conference on Mathematics in Signal Processing*, Birmingham, dec 2012.

## Rapports

[22] S. Icart. Quelques résultats sur le problème de morgan. DEA, Laboratoire d'Automatique de Nantes ENSM, 1987.

[23] S. Icart. *Etude du découplage ligne par ligne avec stabilité : cas des systèmes linéaires*. Thèse, ENSM, Nantes, juillet 1990.

[24] S. Icart, J. Leblond, and C. Samson. Some results on feedback stabilization of a one-link flexible arm. Rapport de recherche RR-1682, INRIA, 1992.

[25] L. Rota, P. Comon, and S. Icart. Blind MIMO equalization under paraunitary constraint. Research Report 2002.35, I3S, September 2002.

[26] S. Icart. Matrices polynomiales et égalisation de canal. Habilitation à Diriger des Recherches, UNS I3S, mars 2013.

## Communications sans acte

[27] S. Icart. Caractérisation géométrique des zéros d'interconnexion pour des systèmes commandables régulièrement découplables. GR Automatique, Pôle non linéaire, Systèmes linéaires multivariables et robustesse, février 1989.

[28] S. Icart. Une condition nécessaire et suffisante de découplage avec stabilité pour un système décalé. GR Automatique, Pôle Non linéaire, Systèmes linéaires multivariables et robustesse, mai 1990.

[29] S. Icart. Loi de commande stabilisante pour un bras de robot flexible. GR Automatique, Pôle SARTA, Systèmes à paramètres répartis, mai 1992.

[30] S. Icart. Quelques propriétés des matrices polynomiales et des matrices rationnelles. application à une réalisation minimale de la fonction de transfert. GR Traitement du Signal, Systèmes multivariables, mars 1994.

## 5.2 Références

[31] K. Abed-Meraim, P. Loubaton, and E. Moulines. A subspace algorithm for certain blind identification problems. *IEEE Transactions on Information Theory*, pages 499–511, March 1997.

[32] M. Al-Ammari and F. Tisseur. Hermitian matrix polynomials with real eigenvalues of definite type. Part I : Classification. *Linear Algebra and its Applications*, 436(10) :3954 – 3973, 2012.

[33] M. A. Alrmah, S. Weiss, and S. Lambotharan. An extension of the MUSIC algorithm to broadband scenarios using a polynomial eigenvalue decomposition. In *EUSIPCO*, Barcelona, Spain, 2011.

[34] G. Ancochea. Zeros of self-inversive polynomials. *Proceedings of the American Mathematical Society*, 4(6) :900–902, 1953.

[35] P. D. Baxter and J. Mc Whirter. Blind signal separation of convolutive mixtures. In *Second IMA Int. Conf. on Mathematics in Communications*, Lancaster UK, December 2002.

[36] P. B. Bhattacharya, S. K. Jain, and S. R. Nagpaul. *Basic Abstract Algebra*. Cambridge University Press, 1994.

[37] A. Bultheel and M. Van Barel. *Linear Algebra, Rational Approximation, and Orthogonal Polynomials*. Elsevier, 1997.

[38] J. F. Cardoso and A. Souloumiac. Blind beamforming for non-Gaussian signals. *IEEE Proceedings Part F*, 140(6) :362–370, December 1993.

[39] J. F. Cardoso and A. Souloumiac. Jacobi angles for simultaneous diagonalization. *SIAM J. Matrix. Anal. Appl.*, 17(1) :161–164, January 1996.

[40] P. Comon. Independent component analysis, a new concept ? *Signal Processing*, 36(3) :287 – 314, 1994.

[41] P. Comon. Contrasts for multichannel blind deconvolution. *Signal Processing Letters, IEEE*, 3(7) :209 –211, july 1996.

[42] P. Comon and C. Jutten, editors. *Handbook of Blind Source Separation : Independent Component Analysis and Applications*. Academic Press, 2010.

[43] P. Comon and L. Rota. Blind separation of independent sources from convolutive mixtures. *IEICE Trans. on Fundamentals of Elec. Com. Comput. Sciences*, E86-A(3) :550–563, March 2003.

[44] J.A. Foster, J.G. McWhirter, M.R. Davies, and J.A. Chambers. An algorithm for calculating the QR and singular value decompositions of polynomial matrices. *IEEE Transactions on Signal Processing*, 58(3) :1263 –1274, mar 2010.

[45] X. Gao, T. Q. Nguyen, and G. Strang. On Factorization of $M$-Channel Paraunitary Filterbanks. *IEEE Transaction on Signal Processing*, 49(7) :1433–1446, July 2001.

[46] Y. V. Genin. On polynomials nonnegative on the unit circle and related questions. *Linear Algebra and its Applications*, 247(0) :317 – 325, 1996.

[47] I. Gohberg, P. Lancaster, and L. Rodman. Spectral analysis of selfadjoint matrix polynomials. *The Annals of Mathematics*, 112(1) :pp. 33–71, 1980.

[48] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials*. Academic Press, 1982.

[49] G. H. Golub and C. F. Van Loan. *Matrix Computation*. The John Hopkins University Press, 3rd edition, 1996.

[50] A. Gorokhov and P. Loubaton. Subspace based techniques for blind separation of convolutive mixtures with temporally correlated sources. *IEEE Trans. Circuits and Systems*, 44 :813–820, September 1997.

[51] A. Hyvärinen and E. Oja. Independent component analysis : algorithms and applications. *Neural Networks*, 13(4–5) :411–430, 2000.

[52] S. Icart and P. Comon. Matrices polynomiales de Laurent. Research Report 2002.35, I3S, October 2012.

[53] Y. Inouye and R. Liu. A system-theoretic foundation for blind equalization of a FIR MIMO channel system. *IEEE Trans. on Circuits and Systems*, 49(4) :425–436, April 2002.

[54] T. Kailath. *Linear Systems*. Prentice Hall, 1980.

[55] V. Kučera. *Analysis and Design of Discrete Linear Control Systems*. Prentice Hall, 1991.

[56] P. Lancaster and L. Rodman. Canonical forms for hermitian matrix pairs under strict equivalence and congruence. *SIAM Review*, 47(3) :pp. 407–443, 2005.

[57] P. Lancaster and F. Tisseur. Hermitian quadratic matrix polynomials : Solvents and inverse problems. *Linear Algebra and its Applications*, 436(10) :4017 – 4026, 2012.

[58] L. De Lathauwer, B. De Moor, and J. Vandewalle. An Algebraic Approach to Blind MIMO Identification. In *Proc. 2nd Int. Workshop on Independent Component Analysis and Blind Signal Separation (ICA 2000)*, pages 211–214, Helsinki, Finland, June 19-22 2000.

[59] L. De Lathauwer, B. De Moor, and J. Vandewalle. An Algebraic Approach to the Blind Identification of Paraunitary Filters. In *Proc. of*

*IEEE Wireless Communications and Networking Conference*, pages 1162 –1165 vol.3, Chicago, USA, September 23-28 2000.

[60] L. De Lathauwer, B. De Moor, and J. Vandewalle. A multilinear singular value decomposition. *SIAM J. Matrix. Anal. Appl.*, 21(4) :1253–1278, 2000.

[61] P. Loubaton, E. Moulines, and P. A. Regalia. Subspace methods for blind identification and deconvolution. In Giannakis, Hua, Stoica, and Tong, editors, *Signal Processing Advances in Wireless and Mobile Communications*, chapter 3. Prentice-Hall, 2001.

[62] J.G. McWhirter, P. D. Baxter, T. Cooper, S. Redif, and J. Foster. An EVD algorithm for para-hermitian polynomial matrices. *IEEE Transactions on Signal Processing*, 55(5) :2158–2169, May 2007.

[63] G. M. Menanno and N. Le Bihan. Quaternion polynomial matrix diagonalization for the separation of polarized convolutive mixture. *Signal Processing*, 90(7) :2219 – 2231, 2010.

[64] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrague. Subspace methods for the blind identification of multichannel FIR filters. *IEEE Transactions on Signal Processing*, 43(2) :516–525, February 1995.

[65] H. Nakazato and P. Psarrakos. On the shape of numerical range of matrix polynomials. *Linear Algebra and its Applications*, 338(1-3) :105 – 123, 2001.

[66] P. J. O'Hara and R. S. Rodriguez. Some properties of self-inversive polynomials. *Proceedings of the American Mathematical Society*, 44(2) :331–335, June 1974.

[67] M. Olivi. *Paramétrage des matrices rationnelles intérieures et applications à la théorie des systèmes.* HDR, Université de Nice Sophia-Antipolis, October 2010.

[68] Q. I. Rahman and G. Schmeisser. *Analytic Theory of Polynomials*. London mathematical Society Monographs. Clarendon Press, 2002.

[69] H. Reddy and P. Rajan. Continued fraction expansion of 1-d complex discrete reactance functions with application to 2-d stability testing. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '87.*, volume 12, pages 1689 – 1691, apr 1987.

[70] S. Redif, J.G. McWhirter, and S. Weiss. Design of FIR paraunitary filter banks for subband coding using a polynomial eigenvalue decomposition. *Signal Processing, IEEE Transactions on*, 59(11) :5253 –5264, nov. 2011.

[71] S. Redif, S. Weiss, and J.G. McWhirter. An approximate polynomial matrix eigenvalue decomposition algorithm for para-hermitian matrices.

In *Signal Processing and Information Technology (ISSPIT), 2011 IEEE International Symposium on*, pages 421 –425, dec. 2011.

[72] P.A. Regalia and S.K. Mitra. An alternate cascade factorization for para-unitary matrices. *Circuits and Systems, IEEE Transactions on*, 38(11) :1373 –1376, nov 1991.

[73] J.P. Richard, editor. *Algèbre et analyse pour l'automatique*. Traité IC2. Hermes, Paris, 2001.

[74] L. Rota. *Egalisation aveugle de systèmes multi-utilisateurs*. Thèse, Université de Nice Sophia-Antipolis, Décembre 2004.

[75] A. H. Sayed and T. Kailath. A Survey of Spectral Factorization Methods. *Numer. Linear Algebra Appl.*, 8 :467–496, 2001.

[76] A. Tkacenko. Approximate eigenvalue decomposition of para-hermitian systems through successive FIR paraunitary transformations. In *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pages 4074 –4077, march 2010.

[77] L. Tong, G. Xu, B. Hassibi, and T. Kailath. Blind channel identification based on second-order statistics : a frequency-domain approach. *IEEE Transactions on Information Theory*, 41(1) :329 –334, jan 1995.

[78] L. Tong, G. Xu, and T. Kailath. Blind identification and equalization based on second-order statistics : a time domain approach. *IEEE Transactions on Information Theory*, 40(2) :340 –349, mar 1994.

[79] R. Turcajovà. Factorizations and construction of linear phase paraunitary filter banks and higher multiplicity wavelets. *Numerical Algorithms*, 8 :1–25, 1994.

[80] P. P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice-Hall, London, 1993.

[81] M. Vidyasagar. *Control Systems Synthesis : A Factorization Approach*, volume 8 of *North Holland System and Control Series*. MIT Press, Cambridge, MA, 1985.

# Annexe A

# Quelques définitions classiques

Le but de cette annexe est de préciser différentes définitions classiques trouvées dans la littérature sur les polynômes, certaines définitions étant proches mais non identiques ainsi que les notations afférentes.

**Polynôme à minimum de phase strict (cas discret FIR)** [80] : Soit $p(z) \in \mathbb{C}[z^{-1}]$, on dira que $p(z)$ est à minimum de phase strict si toutes ses racines sont à l'intérieur strictement du disque unité. Ces polynômes sont appelés également strictement Hurwitz. Lorsque les racines sont toutes à l'extérieur du disque ou sur le cercle unité, $p(z)$ est à maximum de phase.

**Polynôme Schur** [55] : Soit $p(z) \in \mathbb{C}[z]$, alors on dira que $p(z)$ est un polynôme Schur s'il n'a pas de racine de module strictement supérieur à 1. Remarquons que si $p(z) \in \mathbb{C}[z]$ est un polynôme Schur, alors soit $q(z) = p(z^{-1})$, $q(z) \in \mathbb{C}[z^{-1}]$ et est à maximum de phase.

**Polynôme para-conjugué d'un polynôme de Laurent** : Soit $p(z) \in \mathbb{C}[z, z^{-1}]$, alors on définit le polynôme para-conjugué de $p(z)$ comme le L-polynôme noté $\widetilde{p}(z) \in \mathbb{C}[z, z^{-1}]$ et défini par

$$\widetilde{p}(z) = p^*(\frac{1}{z^*}), \forall z \in \mathbb{C}^*$$

Notons que si $p(z) \in \mathbb{C}[z]$ alors $\widetilde{p}(z) \in \mathbb{C}[z^{-1}]$.

**Polynôme para-conjugué d'un polynôme de degré $n$ (cas discret)** [69] : Soit $p(z) \in \mathbb{C}[z]$ un polynôme de degré $n$. Alors on définit le polynôme para-conjugué de $p(z) = \sum_{i=0}^{n} p_i z^i$ comme le polynôme noté $p_*(z) \in \mathbb{C}[z]$ et défini par :

$$p_*(z) = z^n \sum_{i=0}^{n} p_i^* z^{-i} = z^n p^*(\frac{1}{z^*}) = z^n \widetilde{p}(z)$$

**Polynôme miroir d'un polynôme de degré $n$ (cas discret)** [46] : Soit $p(z) \in \mathbb{C}[z]$ un polynôme de degré $n$. Alors on définit le polynôme miroir de $p(z)$ (conjugate mirror image) comme le polynôme noté $\hat{p}(z) \in \mathbb{C}[z]$ et défini par $\hat{p}(z) = z^n p^*(\frac{1}{z^*}) = z^n \widetilde{p}(z)$.
Le polynôme miroir est donc identique au polynôme para-conjugué : les coefficients du polynôme para-conjugué sont les miroirs des coefficients du polynôme initial.

**Polynôme réciproque d'un polynôme de degré $n$** [67] : La définition est identique à celle du polynôme miroir ou du polynôme para-conjugué : $z^n p^*(\frac{1}{z^*})$, il est alors noté $\widetilde{p}(z)$, notation que nous ne retiendrons pas puisque dans ce document $\widetilde{p}(z) = p^*(\frac{1}{z^*})$.

**Polynôme para-conjugué d'un polynôme (cas continu)** [73] : Soit $p(s) \in \mathbb{C}[s]$, alors, dans le cas des systèmes à temps continu, on définit le polynôme para-conjugué de $p(s)$ par $\widetilde{p}(s) = p^*(-s^*)$. On précise parfois que c'est une para-conjugaison par rapport à l'axe imaginaire [37]. Notons que dans le cas continu, si $p(s) \in \mathbb{C}[s]$, alors $\widetilde{p}(s) \in \mathbb{C}[s]$.

**Polynôme para-pair (para-even)** [69] : On dira qu'un polynôme est para-pair s'il est égal à son para-conjugué.
Dans le cas continu, un polynôme para-pair vérifie $p(s) = p^*(-s^*)$.
Dans le cas discret FIR, $p(z) \in \mathbb{C}[z^{-1}]$, un polynôme para-pair vérifie donc $p(z) = z^n p^*(\frac{1}{z^*})$.
Si on s'intéresse aux polynômes de Laurent, alors la para-parité ( !) est identique à la para-hermitianité : $p(z) = \widetilde{p}(z) = p^*(\frac{1}{z^*})$.

**Polynôme para-impair (para-odd)** [69] : On dira qu'un polynôme est para-impair s'il est égal à l'opposé de son para-conjugué. Dans le cas continu, on a donc $p(s) = -p^*(-s^*)$ et dans le cas discret FIR, $p(z) = -p_*(z)$.

**Polynôme auto-inversif (self-inversive)** [34, 66, 68] : Un polynôme $p(z) \in \mathbb{C}[z]$ est self-inversive si l'ensemble de ses racines (avec multiplicité) est symétrique par rapport au cercle unité (c'est donc l'ensemble des racines qui est 'inversive'). Soit $p(z) = \sum_{i=0}^{n} p_i z^i, p_n \neq 0$, alors $p(z)$ est auto-inversif si $p_n p(z) = p_0 z^n \widetilde{p}(z)$.

**Polynôme para-réciproque** [46] : Soit $p(z) \in \mathbb{C}[z]$ un polynôme de degré $n$, alors $p$ est para-réciproque si $p(z) = \varepsilon z^n p^*(\frac{1}{z^*})$ $\forall z$ avec $\varepsilon$ une constante de module 1.

**Polynôme '$n$ self-réciprocal'** [68] : Soit $p(z)$ un polynôme $\in \mathbb{C}[z]$ de degré au plus $n$, alors $p$ est $n$ self-reciprocal si $p(z) = z^n p(\frac{1}{z})$.

1. il s'agit de définitions pour des séries formelles, où l'une des bornes de la sommation peut être infinie, nous les avons énoncées dans le cas d'un polynôme.

**Degré et ordre d'un polynôme [1]de Laurent** [37] : Soit $p(z) \in \mathbb{C}[z, z^{-1}]$ un L-polynôme, $p(z) = \sum_{i=m}^{n} p_i z^i$, $p_m p_n \neq 0, m \leq n$, alors $m$ est l'ordre du polynôme et $n$ est son degré. Avec cette définition, si le degré est négatif $p(z) \in \mathbb{C}[z^{-1}]$, et si l'ordre est non négatif $p(z) \in \mathbb{C}[z]$.

**Système sans perte (loss-less)** : Un système dont la matrice de transfert est une matrice rationnelle causale (chaque élément de la matrice est lui même causal c'est-à-dire que le degré de son numérateur est inférieur au degré de son dénominateur), stable (tous les pôles de la matrice sont dans le disque unité) et para-unitaire sur le cercle $\boldsymbol{H}(e^{j\omega})^H \boldsymbol{H}(e^{j\omega}) = \boldsymbol{I},\ \forall \omega \in \mathbb{R}$ est appelé système sans perte. Notons que grâce à la continuation analytique, si la propriété de para-unitarité est vraie sur le cercle, elle est vraie pour tout $z \in \mathbb{C}$.

**Fonction passe-tout (all-pass)** : Une fonction $h(z)$ de la variable complexe $z$ à valeurs dans $\mathbb{C}$ est passe-tout si son module est constant sur le cercle unité : $|h(e^{j\omega})| = c, c \in \mathbb{R}^{+*},\ \forall \omega \in \mathbb{R}$.

Dans le cas où $h$ est une fraction rationnelle, $h \in \mathbb{C}(z)$ alors en utilisant la continuation analytique, $\tilde{h}(z)h(z) = c^2$ pour $z = e^{j\omega}$ implique

$$\tilde{h}(z)h(z) = c^2,\ \forall z \in \mathbb{C}$$

On montre alors que si $\alpha$ est un pôle d'une fraction ractionnelle passe-tout alors $\dfrac{1}{\alpha^*}$ est un zéro, cette fonction peut donc s'écrire :

$$h(z) = \beta \prod_{k=1}^{N} \frac{-\alpha_k^* + z^{-1}}{1 - \alpha_k z^{-1}}, \beta \neq 0, \alpha_k \in \mathbb{C}$$

Une fonction passe-tout n'est pas nécessairement sans-perte car elle peut ne pas être stable.

# Annexe B

# Publications choisies

Ci-après les références [11, 2, 4], soit une communication à ICASSP et deux articles de la revue Signal Processing.

# BLIND SEPARATION OF CONVOLUTIVE MIXTURES USING SECOND AND FOURTH ORDER MOMENTS

*S. Icart and R. Gautier*

Lab. I3S, CNRS and GDR TdSI
Bât 4, 250 av. A. Einstein, 06560 Sophia Antipolis, France
icart@alto.unice.fr

**ABSTRACT**

This paper deals with the problem of blind identification of $p$-non gaussian inputs, $q$-outputs AR systems in the special case $p < q$. In this case, the identification problem is degenerated, therefore as the classical Levinson (Robinson) algorithm cannot be applied, we use the Inouye's method. As this procedure assumes that AR is normalized, it is necessary to split the problem in two parts: first, to estimate the convolutive mixture by means of linear prediction, and second, to estimate the instantaneous mixture. The first one requires second order moments and the second one, high order statistics. Numerical simulations are presented to show the influence of the conditioning of the instantaneous mixture in the identification problem in presence of noise.

## 1. Problem statement

Nowadays, to identify an unknown channel, one has to use a training sequence (*e.g.* GSM in mobile communication context). If one wants to increase the useful amount of information without any training sequence, one has to deal with blind identification. In a narrow band context, relations between emitting signals and received signals consist in a constant transfer matrix, and the problem of blind identification is reduced to the problem of instantaneous mixtures separation. Now, if one wants that several sources share the same band and takes into account possibly multipath, one gets the following model of convolutive mixtures.

So, let us consider a discrete time-series $y(n)$ of dimension $q$ which are the signals received on the $q$ sensors at time $n$.

$$y(n) = [H(z)]s(n) + w(n) \qquad (1)$$

where $s(n)$ is the $p$-vector of the p emitting sources signals at time $n$. The number of sources $p$ is assumed to be known. Theses sources are supposed to be mutually independent and non-gaussian. For sake of simplicity, we will assume that $E(s(n)s^T(m)) = I_p\delta(n-m)$ where $I_p$ denotes the $p \times p$ identity matrix. $w(n)$ is a white gaussian noise $E(w(n)w(m)^T) = \sigma^2\delta(n-m)I_q$. $H(z)$ is an unknown rational transfer matrix (function of $z^{-1}$).

In this context, we suppose that the number of sensors exceeds the number of emettors ($p < q$), this being relevant in numerous applications (*e.g.* mobile communication context). Moreover, we suppose that $rank(H(z)) = p$, $\forall z$, which implies a minimum-phase system and that the entries of $H(z)$ have no common zeros (that is, the channels are *sufficiently* different one from each other). This assumption is now rather classical. Moreover, we focus our attention to block identification (rather than adaptative identification).

## 2. Convolutive mixtures: a linear prediction approach

Suppose first, that there is no noise ($w(n) = 0$), the process defined by (1) is a vectorial ARMA process. As $H(z)$ is a rational matrix, it can be factorized as $H(z) = A(z)^{-1}B(z)$, where $A(z)$ and $B(z)$ are left coprime polynomial matrices [K80]. According to our assumptions, we have $rank(B(z)) = p\ \forall z$. In case of an AR model, $B(z)$ is a constant matrix. Let us decompose $A(z)$ as $\sum_{i=0}^{M} A_i z^{-i}$ where $M$ is the order of the AR model. In case of a normalized AR, one has $A_0 = I_q$.

We briefly recall the linear prediction approach: an innovation sequence of $y(n)$ is defined as

$$i(n) = y(n) - \{y(n)/sp\{y_l(n-k)/k \geq 1, l = 1, q\}\} \qquad (2)$$

where $y_l(n)$ stands for the l-th component of vector $y(n)$, / for orthogonal projection and $sp\{.\}$ for the Hilbert subspace generated by $\{.\}$. This innovation will be *normalized* if its covariance matrix equals $I_p$. Note that this normalized innovation is not unique: actually, if $i(n)$ is one, then every $Pi(n)$ with $P$ orthogonal is also normalized.

If the process is a normalized AR, we can identify $A_i$ by means of Yule-Walker generalized equations

using second order moments. Let us denote $R_i = E\{y(n)y^T(n-i)\}$ the autocorrelation of $y$ and let $\mathcal{R}_M$ be the following block-Toeplitz matrix

$$\mathcal{R}_M = \begin{pmatrix} R_0 & R_1 & \dots & R_M \\ R_1^T & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_1 \\ R_M^T & \dots & R_1^T & R_0 \end{pmatrix}$$

Let $\mathcal{A}_M = \begin{bmatrix} I & A_1 & \dots & A_M^M \end{bmatrix}$ where $A_i^M$ are the (matricial) coefficients of the AR estimated process, and let $\mathcal{D}_M = \begin{bmatrix} D_M & 0 & \dots & 0 \end{bmatrix}$ corresponds to the quadratic estimation error, then one has to solve the following equation:

$$\mathcal{A}_M \mathcal{R}_M = \mathcal{D}_M$$

In monovariable cases, the well-known Levinson algorithm allows to solve the problem without explicitly inversing $\mathcal{R}_M$ (exploiting backward and forward estimation errors). Robinson [WR65] extended this algorithm to multivariable cases. In both cases $\mathcal{R}_M$ is assumed to be full rank.

As $p < q$, the problem of linear prediction is degenerated: $y(n)$ is only generated by $p$ independent sources, so $rank(R_i) < q$, and the Levinson-Robinson algorithm cannot be applied. But, Inouye [I83] solves the problem in this case using pseudo-inverse instead of inverse. Stability of solution is guaranteed.

Here, $y(n) = [H(z)]s(n)$, with $H(z)$ minimum-phase and $E(s(n)s^T(m)) = I_p\delta(n-m)$, so $s(n)$ is an innovation sequence of $y(n)$ and $H(z)$ is the transfer matrix between a normalized innovation of $y(n)$ and $y(n)$.

For linear prediction, it is assumed that there is no noise, but it still runs in presence of noise: for this, let us first estimate the standard deviation of noise $\sigma^2$ as the smallest eigenvalue of the covariance matrix, and then proceed with the whitened correlation matrix. We then compute a predictor filter $G(z)$ which allows to extract $i(n)$.

If AR is not normalized, as $s(n)$ is a normalized innovation of $y(n)$, after having extracted $i(n)$ by linear prediction, we must extract independent components of instantaneous mixtures by solving $i(n) = Bs(n)$.

## 3. Separation of instantaneous mixtures

In this section, we recall the principles of the separation of instantaneous mixtures [CS94] [C91]. In this case, the transfer matrix is independent of time $H(z) = B$ where $B$ is a $q \times p$ constant matrix. As it is well known, identification is not possible with second order moments and requires 4th-order cumulants. As usual, this blind separation is up to permutation and phase on each input.

The problem of separation of instantaneous mixtures consists in "unmixing" the signal $s_l(n)$ from a block of $N$ observations $y(n)$, *i.e.* in the determination of a matrix $B'$ such that $e(n) = B'y(n)$ where each component of $e(n)$ only depends on one source (and possibly some additive noise). If $B$ is known, separation is achieved by taking $B' = B^\sharp$ where $B^\sharp$ is the pseudo-inverse of $B$ ($B^\sharp = (B^HB)^{-1}B^H$). If $B$ is unknown, one has to add the following assumptions: the source signals $s_i$ are mutually independent stationary process with non-zero 4-th order cumulant. Separation leads then to find a unitary matrix $V$ which maximises

$$C(V) = \sum_{i,k,l} |Cum(z_i, z_i, z_k, z_l)|^2$$

where $Cum$ is the 4-th order cumulant of the corresponding whitened process $z(t)$ with $e(t) = Vz(t)$ and $z$ spatially white. An estimation of $B$ can then be deduced from a jointly diagonalization of its most significant eigenpairs (for more details, see [CS94] and Jade algorithm herein).

## 4. Simulations

As constant modulus signals give good estimation for Jade algorithm [SC93], we choose as inputs signals real, white, null average, binary (-1,1) signals with standard deviation equals to one.

To begin, we run simulations without noise. We then compute a predictor filter $G(z)$ by Inouye's algorithm [I83], with $G(z)H(z) = B$ where $B$ is the instantaneous mixtures matrix. We then determine $B$ by means of Jade algorithm [CS94]. To overcome permutation and phase drawbacks, and to present a bit error ratio between the effective signal and the identified signals, we will need to compute the permutation matrix

$$F = Round(B^\sharp B') \tag{3}$$

where *Round* rounds to the nearest integer each component of the matrix, $B^\sharp$ is the pseudo inverse of $B$ and $B'$ is the estimation of $B$. In the following, we note $\gamma = |B^\sharp B'|$ where $|\ |$ takes the absolute value of each component of the matrix. If we take constant modulus signals, $B'$ is close to $B$, so $F$ is significant (close to $\gamma$). So, we filter the output data of Jade algorithm by $F$. *N.B. this is possible only because $B$ is known and this will be only used to quantify the bit error ratio.* Finally, according to the binary inputs, we choose as decision criterion the sign of $\hat{s}_i$, the estimation of each signal source.

We run simulations for a 3 sources, 4 captors system for two different systems: one with a well-conditioned

instantaneous mixtures matrix ($cond(B_w) = 1.35$) and second with an ill-conditioned one ($cond(B_i) = 5.13$).

$$B_w = \begin{pmatrix} 0.3 & 0.2 & 0.4 \\ 0.3 & 0.5 & 0.2 \\ 0.4 & -0.4 & 0.45 \\ -0.6 & 0.1 & 0.38 \end{pmatrix} \quad B_i = \begin{pmatrix} 0.3 & 0.2 & 0.1 \\ 0.2 & 0.5 & 0.4 \\ 0.1 & 0 & 0.2 \\ 0.5 & 0.1 & 0.4 \end{pmatrix}$$

The pole part of the AR is the same in both cases :

$$A_0 = I_4 \quad A_1 = A_2 = \begin{pmatrix} 0.2 & 0.1 & 0.2 & 0.3 \\ 0.3 & 0.4 & 0.5 & 0 \\ 0.4 & 0.1 & 0.2 & 0 \\ 0.2 & 0.4 & 0.2 & 0.1 \end{pmatrix}$$

For $N = 100$ samples, one gets (with $w$ subscript for well-conditioned matrix and $i$ subscript for ill-conditioned one):

$$\gamma_w = \begin{pmatrix} 0.0313 & 0.0513 & 0.9982 \\ 0.9962 & 0.0812 & 0.0330 \\ 0.0812 & 0.9954 & 0.0516 \end{pmatrix}$$

$$\gamma_i = \begin{pmatrix} 0.0207 & 0.0606 & 0.9979 \\ 0.9996 & 0.0197 & 0.0206 \\ 0.0216 & 0.9980 & 0.0601 \end{pmatrix}$$

As $N$ increases, $\gamma$ becomes closer to the real permutation matrix $F = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

Figures 1 and 2 give the average mean square error MSE of one source in function of the size of block $N$ *before decision* in both cases. After decision, the bit error ratio is zero either $B$ is well or bad conditioned; the identification is well done.

Let us now examine what happens in presence of noise. After having determined the noise standard deviation $\sigma^2$ as the smallest eigenvalue of the covariance matrix, the procedure is the same than above. Figures 3 and 4 present the bit error ratio for one source (for ten realizations) in function of

- the number of samples used for identification;
- the signal to noise ratio.

In both cases, the size of block has a small influence on the results. But, if when $B$ is well-conditioned, one gets quite acceptable results, one gets bad results for an ill-conditioned case. For example, the bit error ratio

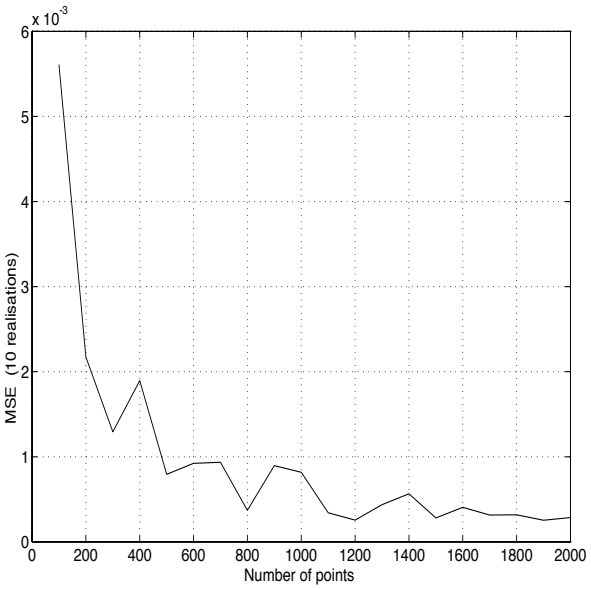



Figure 1: well-conditioned

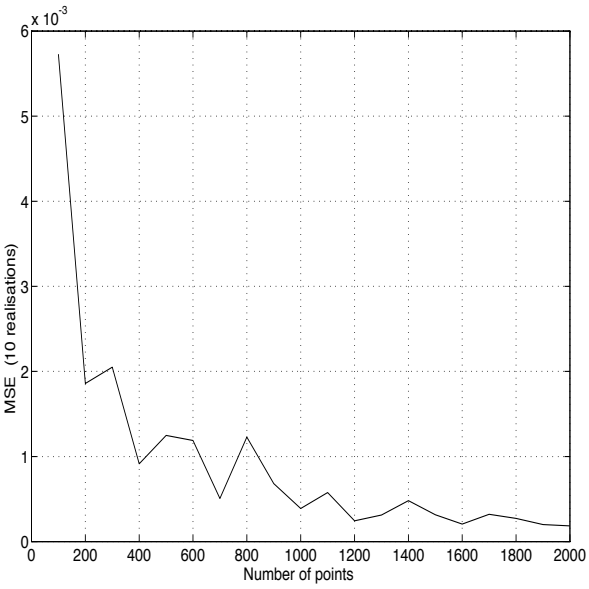



Figure 2: ill-conditioned

is about 5% for an SNR of 17dB and more than 25% for 10dB (in this case, one gets a null bit error ratio only for a SNR of 24dB!). The conditioning of $B$ has few influence when noise is zero, but is crucial in presence of noise: errors due to the noise are amplified.

## 5. Conclusion

In this paper, we have presented a new blind separation of convolutive mixtures method, using both 2nd and 4th order moments, cascading two existing algorithms. As the HOS are only required to determine few parameters, this method is not too expensive in time of calculus. We have performed simulations to illustrate the bit error ratio. It appears that the conditioning of the instantaneous mixtures matrix has a great importance in pre-treatment of the data to decrease number of errors. Further work will address the problem of overdetermination of the model order $M$.

## Acknowledgment

The authors want to thank Ph. Loubaton who is at the origin of this work in the framework of the working group GT4-multivariate systems- inside the GDR 134-CNRS.

## 6. References

[CS94] J.-F. Cardoso and A.Souloumiac, "Blind beamforming for non gaussian signals", in Proc. IEE, vol. 140, pp. 362-370, 1994.

[C91] P. Comon, "Independent component analysis", in Proc. Int. Workshop on Higher-Order Stat., Chamrousse, France, pp. 111-120, 1991.

[I83] Y. Inouye, "Modeling of Multichannel Time Series and Extrapolation of Matrix-Valued Autocorrelation Sequences", IEEE Trans. Acoust., Speech, and Signal Processing, vol. ASSP-31, no. 1, 1983.

[K80] T. Kailath, "Linear Systems", Prentice-Hall, 1980.

[SC93] A.Souloumiac and J.-F. Cardoso, "Performances en séparation de sources", in Proc. GRETSI, Juan-les-Pins, France, pp. 321-324, 1993.

[WR65] R. A. Wiggins and E. A. Robinson, "Recursive Solution to the Multichannel Filtering Problem", J. Geophys. Res., vol. 70, no. 8, pp. 1885-1891, 1965.

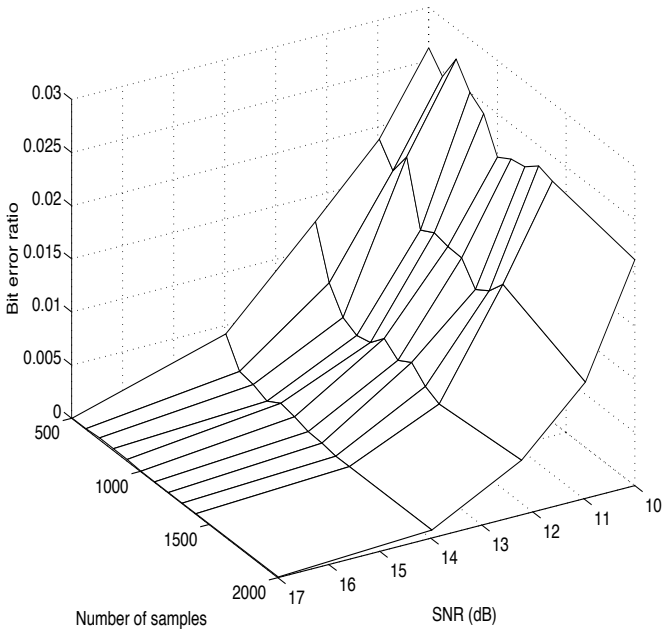



Figure 3: well-conditioned

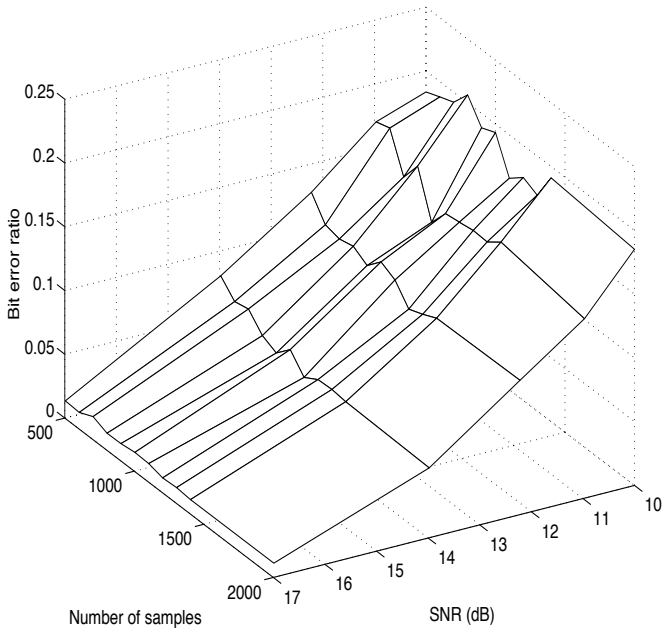



Figure 4: ill-conditioned

Signal Processing 89 (2009) 283–290

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

# Blind paraunitary equalization

Sylvie Icart [a,*], Pierre Comon [a], Ludwig Rota [b]

[a] Laboratoire I3S, CNRS UMR 6070, Université de Nice, 2000 route des Lucioles, BP 121, 06903 Sophia-Antipolis Cedex, France
[b] Orange Business Service, Sophia-Antipolis, France

ARTICLE INFO

*Article history:*
Received 26 October 2007
Received in revised form
2 June 2008
Accepted 28 August 2008
Available online 1 October 2008

*Keywords:*
Blind deconvolution
Blind equalization
Multiple-input/multiple-output (MIMO)
High order statistics (HOS)
Paraunitary filter
Pairwise cumulants

ABSTRACT

In this paper a blind multiple-input/multiple-output (MIMO) space–time equalizer is described, dedicated to convolutive mixtures when observations have been prewhitened. Filters preserving space–time whiteness are paraunitary; a parameterization of such filters with plane rotations is proposed. Theoretical developments then lead to a numerical algorithm that sweeps all pairs of delayed outputs. This algorithm involves the solution of a polynomial system in two unknowns, whose coefficients depend on the output cumulants. Simulations and performance of the numerical algorithm are reported.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

In actual digital communication systems, the equalization problem is solved using learning sequences. These sequences, known by transmitter and receiver, permit to estimate channel parameters. Nevertheless, learning sequences may be seen as reducing the throughput, because they occupy a non-negligible space in transmitted sequences. In other words, less actual information is transmitted, i.e. the useful data rate is lower than the system data rate. Next, in some situations, e.g. interception in electronic warfare, learning sequences exist but are not known to the receiver.

Blind separation methods for multiple-input multiple-output (MIMO) channels have raised an increasing interest for digital communications since they do not need learning sequences. Most blind MIMO equalization techniques use high order statistics (HOS) for separating signals [1–8]. This can be implicit through constant modulus [9–12] or constant power [13] criteria. Purely deterministic approaches also exist and exploit either the finite alphabet property [14–16] or the presence of another diversity in addition to time and space [17,18].

Our main contribution consists of a block algorithm dedicated to blind MIMO equalization and exploiting HOS of observations [19]. The particularity of our method is that it is based on a factorization of paraunitary filters [20], as suggested in [21,22]. The paraunitary condition is not very restrictive since prewhitening can always be performed in a first stage (in a non-unique manner) even if it is not always an obvious operation [22–25]. Moreover, the algorithm designed herein can be implemented either "on-line" or "off-line".

More precisely, our algorithm can run iterations on a single given block until convergence, or on the contrary, run a new iteration after each symbol arrival. Convergence of on-line algorithms is known to be much longer in terms of number of symbols required (typically several thousands of symbols).

Note that specific criteria [26,27] and algorithms [28] dedicated to paraunitary channels have already been proposed. Unfortunately, the paraunitary constraint is not accurately verified in the modeling of [27,28],

* Corresponding author. Tel.: +33 4 92 94 27 13.
*E-mail address:* Sylvie.Icart@unice.fr (S. Icart).

0165-1684/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2008.08.014

especially at low signal-to-noise ratios (SNR), because the channel is parameterized with a semi-unitary matrix.

The paper is organized as follows. In the second section, model and hypotheses are presented. The parameterization of MIMO paraunitary filters is introduced in Section 3, and input–output cumulant relations are established. A contrast criterion is proposed in Section 4, for finding equalizer parameters and separating source signals. An iterative algorithm built from previous theoretical results is described in the fifth section. The core of the algorithm is a Jacobi-type iteration, in which all possible Givens rotations in the parameterization of the paraunitary filter are swept until convergence. Finally, the performance for various SNR is illustrated in the last section.

## 2. Model and notations

Throughout the paper, $(\mathsf{T})$ stands for transposition, $(\mathsf{H})$ for conjugate transposition, $(*)$ for complex conjugation, and $\jmath=\sqrt{-1}$. Also denote by $\mathbb{Z}$ the set of integers, by $\mathbb{N}$ the subset of positive integers and by $z^{-1}$ the time-delay operator. Vectors and matrices are denoted with bold lowercase and bold uppercase letters, respectively. The entries of matrix $\boldsymbol{G}$ are denoted $G_{ij}$, where subscript $ij$ denotes the $i$-th row and the $j$-th column of $\boldsymbol{G}$.

Let us consider a digital communication system of $N$ antennas and $N$ receivers in a multipath environment. Let $\boldsymbol{s}(n)=(s_1(n),\dots,s_N(n))^{\mathsf{T}}$ denote the $N$-dimensional source vector of baseband complex signals at time $n$ and $\boldsymbol{w}(n)=(w_1(n),\dots,w_N(n))^{\mathsf{T}}$ the $N$-dimensional observation vector. Let $\boldsymbol{C}[z]$ be the transfer function of the linear time invariant (LTI) mixing channel and $\{\boldsymbol{C}(k),k=0,\dots,K\}$ be the sequence of $N\times N$ matrices of the complex finite impulse response (FIR) of the channel. We then have

$$\boldsymbol{w}(n)=\sum_{k=0}^{K}\boldsymbol{C}(k)\boldsymbol{s}(n-k), \tag{1}$$

where $K$ denotes the memory length of the channel. Also denote

$$\boldsymbol{C}[z]=\sum_{m=0}^{K}\boldsymbol{C}(m)z^{-m}. \tag{2}$$

The case of instantaneous mixtures is not considered in the present paper, since already addressed via pairwise processing by various authors since 1991. Thus, it will be assumed that $K>0$.

The multichannel blind deconvolution problem consists of finding an LTI filter $\boldsymbol{H}[z]$, i.e. the *equalizer*, in order to retrieve the $N$ input signals $s_i(n), i\in\{1,\dots,N\},\forall n\in\mathbb{Z}$, solely from the observation of the outputs $\boldsymbol{w}(n)$ of the unknown LTI channel $\boldsymbol{C}[z]$.

Let $\hat{\boldsymbol{s}}(n)=(\hat{s}_1(n),\dots,\hat{s}_N(n))^{\mathsf{T}}$ be the $N$-dimensional estimated source vector. This means, with the above notation:

$$\hat{\boldsymbol{s}}(n)=\sum_{l=0}^{L}\boldsymbol{H}(l)\boldsymbol{w}(n-l), \tag{3}$$

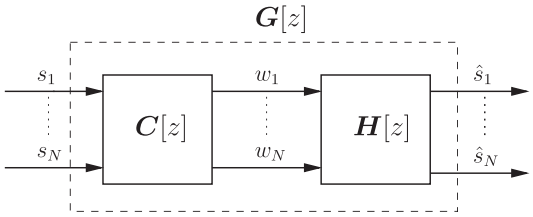



**Fig. 1.** Global system $\boldsymbol{G}$: sources $s_i$ are filtered by channel $\boldsymbol{C}[z]$ and observations $w_i$ are equalized by $\boldsymbol{H}[z]$.

where $L$ is the memory length of the equalizer, or, as a function of original sources themselves:

$$\hat{\boldsymbol{s}}(n)=\sum_{m=0}^{K+L}\boldsymbol{G}(m)\boldsymbol{s}(n-m), \tag{4}$$

where $\boldsymbol{G}[z]$ denotes the global filter $\boldsymbol{G}[z]=\boldsymbol{H}[z]\boldsymbol{C}[z]$ (see Fig. 1 above).

**Definition 1** (*Paraunitarity*). A square polynomial matrix $\boldsymbol{H}[z]\in\mathbb{C}^{N\times N}$ is said to be paraunitary [21] if

$$\boldsymbol{H}^{\mathsf{H}}[1/z^*]\boldsymbol{H}[z]=\boldsymbol{I}_N=\boldsymbol{H}[z]\boldsymbol{H}^{\mathsf{H}}[1/z^*], \tag{5}$$

where $\boldsymbol{I}_N\in\mathbb{R}^{N\times N}$ is the identity matrix.

In this paper, we assume the following hypotheses:

(H1) Inputs $s_i(n),\forall i\in\{1,\dots,N\},\forall n\in\mathbb{Z}$, are mutually independent and identically distributed (i.i.d.) zero-mean random processes, with unit variance.

(H2) Vector $\boldsymbol{s}(n)$ is stationary up to the considered order $r$, $r\geqslant 3$, i.e. $\forall i\in\{1,\dots,N\}$, the order-$r$ marginal cumulants,

$$\mathsf{C}_p^q[s_i]=\mathsf{Cum}[\underbrace{s_i(n),\dots,s_i(n)}_{p\text{ terms}},\underbrace{s_i^*(n),\dots,s_i^*(n)}_{q=r-p\text{ terms}}]$$

do not depend on $n$. For definitions of cumulants, refer to [29] and references therein.

(H3) At most one source has a zero marginal cumulant of order $r$.

(H4) $\boldsymbol{C}[z]$, $\boldsymbol{H}[z]$, and hence $\boldsymbol{G}[z]=\boldsymbol{H}[z]\boldsymbol{C}[z]$ are all paraunitary. Thus we have the global relation:

$$\boldsymbol{H}[z]\boldsymbol{C}[z]\boldsymbol{C}^{\mathsf{H}}[1/z^*]\boldsymbol{H}^{\mathsf{H}}[1/z^*]=\boldsymbol{I}_N. \tag{6}$$

**Remark 2.** One can always whiten the observations by using a filter that factorizes the second-order power spectrum, i.e. a classical prewhitening of the observations [22,23]. In other words, paraunitary filters are relevant after a space–time standardization of observations (second-order white with unit covariance). The way space–time whitening is implemented is out of the scope of this paper, and it is assumed that observations $\boldsymbol{w}(n)$ are space–time white.

As is now well known, statistical independence of $\hat{s}_i$ can only allow to blindly recover the sources up to a permutation matrix $\boldsymbol{P}$, and a diagonal delay filter $\boldsymbol{\Lambda}[z]$, so that $\boldsymbol{H}[z]\boldsymbol{C}[z]=\boldsymbol{\Lambda}[z]\boldsymbol{P}$.

## 3. Equalization

We introduce in this section the global parameterization of the paraunitary equalizer, which will allow in Section 5 an implementation in terms of Givens rotations.

### 3.1. Parameterization

From our observation model and from hypotheses described in Section 2, we have a first proposition:

**Proposition 3.** *Let $\boldsymbol{H}[z] \in \mathbb{C}^{N\times N}$ be an FIR paraunitary filter of Mac Millan degree*[1] *$\lambda_h$. Then matrix $\boldsymbol{H}[z]$ can be factorized (non-uniquely) as the product*:

$$\boldsymbol{H}[z] = \boldsymbol{A}[z]\boldsymbol{Q}\boldsymbol{B}[z], \tag{7}$$

*where $\boldsymbol{Q} \in \mathbb{C}^{N\times N}$ is unitary and $\boldsymbol{A}[z]$ and $\boldsymbol{B}[z]$ are FIR paraunitary filters of Mac Millan degree $\ell_A$ and $\ell_B$, respectively, with $0 \leqslant \ell_A \leqslant \lambda_h$, $0 \leqslant \ell_B \leqslant \lambda_h$, $\ell_A + \ell_B = \lambda_h$.*

**Proof.** The proof is easy since we can extend the paraunitary factorization of [30] to the $N$-dimensional complex case:

$$\boldsymbol{H}[z] = \boldsymbol{Q}_0\boldsymbol{Z}[z]\boldsymbol{Q}_1 \dots \boldsymbol{Z}[z]\boldsymbol{Q}_{\lambda_h-1}\boldsymbol{Z}[z]\boldsymbol{Q}_{\lambda_h}, \tag{8}$$

where $\boldsymbol{Q}_p \in \mathbb{C}^{N\times N}$ are unitary for $p \in \{0,\dots,\lambda_h\}$, $\lambda_h$ is the MacMillan degree of $\boldsymbol{H}[z]$, and $\boldsymbol{Z}[z]$ denotes the $N \times N$ diagonal matrix

$$\boldsymbol{Z}[z] = \begin{bmatrix} \boldsymbol{I}_{N-1} & 0 \\ 0 & z^{-1} \end{bmatrix}. \tag{9}$$

Thus, for $0 < \ell_A \leqslant p$, we have

$$\boldsymbol{A}[z] = \boldsymbol{Q}_0\boldsymbol{Z}[z] \dots \boldsymbol{Q}_{p-1}\boldsymbol{Z}[z], \tag{10}$$

and for $0 < \ell_B \leqslant \lambda_h - p$, we have

$$\boldsymbol{B}[z] = \boldsymbol{Z}[z]\boldsymbol{Q}_{p+1} \dots \boldsymbol{Z}[z]\boldsymbol{Q}_{\lambda_h}. \tag{11}$$

When $\ell_A = 0$ (respectively $\ell_B = 0$), we can replace $\boldsymbol{A}[z]$ (respectively $\boldsymbol{B}[z]$) by $\boldsymbol{I}_N$. Then, for a fixed $p$, any paraunitary filter of MacMillan degree $\lambda_h$ can be factorized like in (7). □

### 3.2. MIMO relations

In the remaining, we assume the following notation for fourth-order cumulants, e.g. cumulants of vector $\boldsymbol{w}$,

$$\Gamma^{\boldsymbol{w}}_{eg,fh}(\boldsymbol{v}) = \mathrm{Cum}[w_e(n-v_1), w_f^*(n-v_2), w_g(n-v_3), w_h^*(n-v_4)], \tag{12}$$

where $e,f,g,h$ take their values in $\{1,\dots,N\}$, and $v_i \in \mathbb{N}, \forall i \in \{1,\dots,4\}$.

Now, consider the following input–output relations for the convolutive model

$$\hat{s}_i(n) = \sum_{m=0}^{l_A}\sum_{r=1}^{N}\sum_{q=1}^{N} A_{iq}(m)Q_{qr}x_r(n-m), \tag{13}$$

and

$$x_r(n-m) = \sum_{l=0}^{l_B}\sum_{s=1}^{N} B_{rs}(l)w_s(n-m-l). \tag{14}$$

From (13) and thanks to the multilinearity property of cumulants, we can express the input–output relations between cumulants of input $\boldsymbol{x}$ and output $\hat{\boldsymbol{s}}$,

$$\Gamma^{\hat{\boldsymbol{s}}}_{ik,jl} = \sum_{abcd}\sum_{\boldsymbol{\tau}}\sum_{qrst} A_{iq}(\tau_1)A^*_{jr}(\tau_2)A_{ks}(\tau_3)A^*_{lt}(\tau_4)Q_{qa} \times Q^*_{rb}Q_{sc}Q^*_{td}\Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau}), \tag{15}$$

with $\boldsymbol{\tau} = (\tau_1,\tau_2,\tau_3,\tau_4)$. Since $\hat{\boldsymbol{s}}$ is the output of filter $\boldsymbol{A}[z]Q$, the range of each $\tau_i$ is $[0,\dots,\ell_A]$ and all indices $a,b,c,d,i,j,k,l,q,r,s,t$ take their values in $\{1,\dots,N\}$. Then, we can write a similar relation between cumulants of input $\boldsymbol{w}$ and output $\boldsymbol{x}$,

$$\Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau}) = \sum_{\boldsymbol{\rho}}\sum_{efgh} B_{ae}(\rho_1)B^*_{bf}(\rho_2)B_{cg}(\rho_3)B^*_{dh}(\rho_4)\Gamma^{\boldsymbol{w}}_{eg,fh}(\boldsymbol{\tau}+\boldsymbol{\rho}), \tag{16}$$

with $\boldsymbol{\rho} = (\rho_1,\rho_2,\rho_3,\rho_4)$. Here, the range of each $\rho_i$ is $[0,\dots,\ell_B]$ since inputs are filtered by $\boldsymbol{B}[z]$.

### 3.3. Construction of $\mathbf{Q}_p$

Matrices $\boldsymbol{Q}_p$ involved in the paraunitary factorization (8) are unitary. It is well known that any $N \times N$ unitary matrix can be written as a product of $M = N(N-1)/2$ Givens rotations [31, p. 215] $\boldsymbol{U}(i,j,\theta,\phi)$ with $1 \leqslant i < j \leqslant N$ and

$$\boldsymbol{U}(i,j,\theta,\phi) = \begin{bmatrix} \boldsymbol{I}_{n_1} & \vdots & \boldsymbol{0} & \vdots & \boldsymbol{0} & \\ \dots & \cos\theta & \dots & \sin\theta\,\mathrm{e}^{\jmath\phi} & \dots & i, \\ \boldsymbol{0} & \vdots & \boldsymbol{I}_{n_2} & \vdots & \boldsymbol{0} & \\ \dots & -\sin\theta\,\mathrm{e}^{-\jmath\phi} & \dots & \cos\theta & \dots & j, \\ \boldsymbol{0} & \vdots & \boldsymbol{0} & \vdots & \boldsymbol{I}_{n_3} & \end{bmatrix} \tag{17}$$

where $n_1 + n_2 + n_3 = N - 2$. These unitary rotations allow to develop a *Jacobi*-type iteration algorithm for equalizer (7).

Since we use Givens matrices, we can make two remarks.

**Remark 4.** For $N = 2$, unitary matrix $\boldsymbol{Q}_p$ consists of a single rotation

$$\boldsymbol{Q}_p = \begin{bmatrix} \cos\theta_p & \sin\theta_p\,\mathrm{e}^{\jmath\phi_p} \\ -\sin\theta_p\,\mathrm{e}^{-\jmath\phi_p} & \cos\theta_p \end{bmatrix}. \tag{18}$$

Then, only one pair of angles $(\theta_p,\phi_p)$ is needed for each $\boldsymbol{Q}_p$. Hence, when more than two signals are observed, it is also possible to use a deflation-type algorithm [32].

**Remark 5.** Proposition 3 tells that $\lambda_h + 1$ unitary matrices are used to obtain a complete factorization of the equalizer. For instance, the $2 \times 2$ length-1 paraunitary filter defined by $\mathrm{diag}(\boldsymbol{H}[z]) = [-z^{-1},\ z^{-1}]$ and 0 elsewhere, needs $\lambda_h = \deg[\det(\boldsymbol{H}[z])] = 2$ diagonal matrices $\boldsymbol{Z}[z]$.

[1] The Mac Millan degree of a paraunitary filter is also the degree of its determinant [21, p. 737].

Indeed, $\boldsymbol{H}[z]$ has the following factorization:

$$\boldsymbol{H}[z] = \boldsymbol{Q}_2 \cdot \boldsymbol{Z}[z] \cdot \boldsymbol{Q}_1 \cdot \boldsymbol{Z}[z] \cdot \mathrm{e}^{i\pi/2}, \tag{19}$$

where

$$\boldsymbol{Q}_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and

$$\boldsymbol{Q}_1 = \begin{bmatrix} 0 & J \\ J & 0 \end{bmatrix},$$

i.e. $(\theta_2 = \pi/2,\ \phi_2 = 0)$ and $(\theta_1 = \pi/2, \phi_1 = \pi/2)$. The exponential is part of the unknown matrix $\boldsymbol{\Lambda}[z]\boldsymbol{P}$, an indetermination inherent in the problem as explained in Section 2. It is also true for $L \geqslant 1$ since we can group all exponentials in a single matrix $\boldsymbol{\Lambda}[z]\boldsymbol{P}$. Thus, reflections, e.g. $\mathrm{diag}(\boldsymbol{Q}) = [1, -1]$, are covered by the parameterization of Proposition 3, thanks to the inherent indeterminacy. Parameterization (7) is thus complete.

## 4. Optimization criterion

Let us start this section with two key definitions. The first definition deals with the set of filters that do not affect the independence of sources, a key point of blind source separation methods. The second definition describes the main properties of a *contrast*, i.e. an optimization criterion for separating signals.

**Definition 6** (*Trivial filters*). The set $\mathscr{S}$ of source processes is characterized by assumptions, such as (H1). One defines the set $\mathscr{T}$ of trivial filters, as containing all filters that do not affect these assumptions. In other words, $\mathscr{S}$ is stable by the operation of $\mathscr{T}$. For instance, filters of the form $\boldsymbol{\Lambda}[z]\boldsymbol{P}$, where $\boldsymbol{P}$ is a permutation matrix, and $\boldsymbol{\Lambda}[z]$ a diagonal filter, do not affect mutual independence between components of $\boldsymbol{s}(n)$. If in addition $\boldsymbol{s}(n)$ is an i.i.d. non-Gaussian process, $\boldsymbol{\Lambda}[z]$ should contain only pure delays, integer multiples of the sampling period, and fixed complex factors; in other words, the entries of $\boldsymbol{\Lambda}[z]$ are of the form $\gamma z^k$, with $k \in \mathbb{Z}$ and $|\gamma| = 1$.

**Definition 7** (*Contrast*). Let $\mathscr{H}$ be a set of filters, and denote $\mathscr{H} \cdot \mathscr{S}$ the set of standardized linear processes obtained by letting filters of $\mathscr{H}$ operate on processes of $\mathscr{S}$. Denote $\boldsymbol{I}$ the identity matrix. An optimization criterion, $\Upsilon(\boldsymbol{H}; \boldsymbol{w})$, will be referred to as a contrast for $\boldsymbol{H} \in \mathscr{H}, \boldsymbol{w} \in \mathscr{H} \cdot S$, if it satisfies the three properties below [2]:

- *Invariance*: The contrast should not change within the set of acceptable solutions, which means that $\forall \boldsymbol{w} \in \mathscr{H} \cdot \mathscr{S}$, $\forall \boldsymbol{H} \in \mathscr{T}$ then $\Upsilon(\boldsymbol{H}; \boldsymbol{w}) = \Upsilon(\boldsymbol{I}; \boldsymbol{w})$.
- *Domination*: If sources are already separated, any filter should decrease the contrast, i.e. $\forall \boldsymbol{w} \in \mathscr{S}$, $\forall \boldsymbol{H} \in \mathscr{H}$, then $\Upsilon(\boldsymbol{H}; \boldsymbol{w}) \leqslant \Upsilon(\boldsymbol{I}; \boldsymbol{w})$.
- *Discrimination*: The maximum contrast should be reached only for filters linked to each other via trivial filters: $\forall \boldsymbol{w} \in \mathscr{S}$, $\Upsilon(\boldsymbol{H}; \boldsymbol{w}) = \Upsilon(\boldsymbol{I}; \boldsymbol{w}) \Rightarrow \boldsymbol{H} \in \mathscr{T}$.

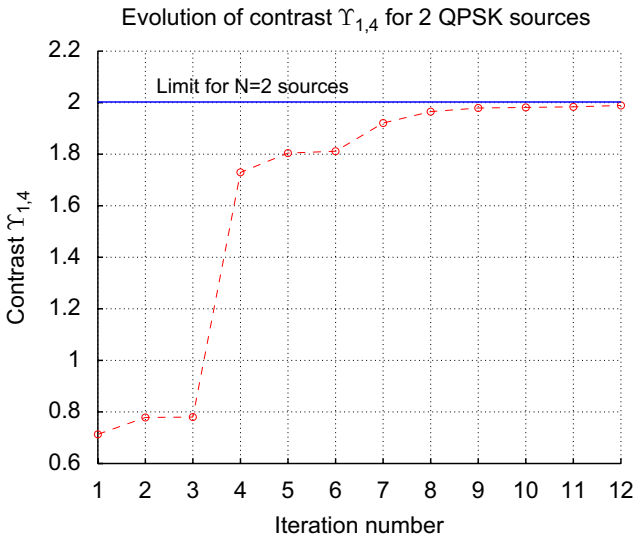



**Fig. 2.** Typical evolution of contrast $\Upsilon_{1,4}$ for 25 dB of SNR and length-3 equalizer.

Now, let us focus on the contrast used to carry out the equalization of the system. We can separate signals blindly by using an approach based on HOS. Indeed, since fourth-order cross-cumulants are a good measure of statistical independence, we can build a contrast based on these values [2,16]:

**Proposition 8.** *Consider sources having kurtosis of the same sign* $\varepsilon$, $\varepsilon = \pm 1$. *Then, the source separation can be performed solely from channel outputs by maximizing the following contrast*:

$$\Upsilon_{1,4} = \varepsilon \sum_{i=1}^{N} \Gamma^{\hat{\boldsymbol{s}}}_{ii,ii}, \tag{20}$$

*where cumulants of* $\hat{\boldsymbol{s}}$ *are defined by*

$$\Gamma^{\hat{\boldsymbol{s}}}_{ii,ii} = \sum_{abcd} \sum_{\tau} \sum_{qrst} A_{iq}(\tau_1) A^*_{ir}(\tau_2) A_{is}(\tau_3) A^*_{it}(\tau_4) Q_{qa} Q^*_{rb} \times Q_{sc} Q^*_{td} \Gamma^{\boldsymbol{x}}_{ac,bd}(\boldsymbol{\tau}). \tag{21}$$

*Hence, in order to estimate* $\boldsymbol{H}[z]$, *the optimization problem can be rewritten as*

$$\boldsymbol{H} = \mathrm{Arg} \max_{\mathbb{Q}} \Upsilon_{1,4}, \tag{22}$$

*where* $\mathbb{Q}$ *stands for the set of all Givens rotations* $\boldsymbol{Q}_p$, *of form* (17). *This set spans the whole group of unitary matrices.*

It has been proved in [2] that $\Upsilon_{1,4}$ is a contrast. We maximize $\Upsilon_{1,4}$ with respect to each pair $(\theta, \phi)$ in turn. The sequence of values of $\Upsilon_{1,4}$ obtained this way is monotonically increasing. Since it is also bounded above, it converges. Fig. 2 is an example of such a behavior.

## 5. A Jacobi-type algorithm

In this section we detail an iteration of the algorithm for one pair $(\theta, \phi)$ of Givens rotation. Then, as in *Jacobi sweeping* algorithms for matrices [31], we sweep $M = N(N-1)/2$ pairs for each unitary matrix $\boldsymbol{Q}_p$.

### 5.1. Processing a pair of outputs

For the sake of clarity, we drop index $p$. We have to find all pairs $(\theta,\phi)$ that maximize (20), the other pairs being fixed. To reach this goal, (21) is simplified firstly by expanding it, and secondly by collecting terms involving $\theta$ or $\phi$. In this manner, we obtain the output cumulants of $\hat{\boldsymbol{s}}$ as a linear combination of terms of the form $\cos^{\alpha}\theta\sin^{4-\alpha}\theta\,\mathrm{e}^{\jmath\eta\phi}$. More precisely

$$\begin{aligned}\Gamma^{\hat{\mathbf{s}}}_{ii,ii} = {} & \mathscr{K}^{(0)}_{(4)}\cos^4\theta + \mathscr{K}^{(1)}_{(3)}\cos^3\theta\,\sin\theta\,\mathrm{e}^{\jmath\phi} \\ & + \mathscr{K}^{(-1)}_{(3)}\cos^3\theta\sin\theta\,\mathrm{e}^{-\jmath\phi} + \mathscr{K}^{(0)}_{(2)}\cos^2\theta\sin^2\theta \\ & + \mathscr{K}^{(2)}_{(2)}\cos^2\theta\sin^2\theta\,\mathrm{e}^{2\jmath\phi} + \mathscr{K}^{(-2)}_{(2)}\cos^2\theta\sin^2\theta\,\mathrm{e}^{-2\jmath\phi} \\ & + \mathscr{K}^{(1)}_{(1)}\cos\theta\sin^3\theta\,\mathrm{e}^{\jmath\phi} + \mathscr{K}^{(-1)}_{(1)}\cos\theta\sin^3\theta\,\mathrm{e}^{-\jmath\phi} \\ & + \mathscr{K}^{(3)}_{(1)}\cos\theta\sin^3\theta\,\mathrm{e}^{3\jmath\phi} + \mathscr{K}^{(-3)}_{(1)}\cos\theta\sin^3\theta\,\mathrm{e}^{-3\jmath\phi} \\ & + \mathscr{K}^{(0)}_{(0)}\sin^4\theta + \mathscr{K}^{(4)}_{(0)}\sin^4\theta\,\mathrm{e}^{4\jmath\phi} \\ & + \mathscr{K}^{(-4)}_{(0)}\sin^4\theta\,\mathrm{e}^{-4\jmath\phi} + \mathscr{K}^{(2)}_{(0)}\sin^4\theta\,\mathrm{e}^{2\jmath\phi} \\ & + \mathscr{K}^{(-2)}_{(0)}\sin^4\theta\,\mathrm{e}^{-2\jmath\phi}, \end{aligned} \tag{23}$$

where

$$\mathscr{K}^{(\eta)}_{(\alpha)}\,\mathrm{e}^{\jmath\eta\phi} = \sum_{\boldsymbol{\tau}} A_{iq}(\tau_1)A^*_{ir}(\tau_2)A_{is}(\tau_3)A^*_{it}(\tau_4)\Gamma^{\mathbf{x}}_{ac,bd}(\boldsymbol{\tau}). \tag{24}$$

According to (18), (21) and (24), each coefficient $\mathscr{K}^{(\eta)}_{(\alpha)}$ depends on indices $a,b,c,d$ and $q,r,s,t$ and as the product $Q_{qa}Q^*_{rb}Q_{sc}Q^*_{td}$ gives an expression of the form $\cos^{\alpha}\theta\sin^{4-\alpha}\theta\mathrm{e}^{\jmath\eta\phi}$ with $\alpha\in\{0,\ldots,4\}$ and $\eta\in\{\pm1,\pm2,\pm3,\pm4\}$; let $\eta=2\beta+\alpha-4$, we get

$$\Gamma^{\hat{\mathbf{s}}}_{ii,ii} = \sum_{\alpha=0}^{4}\sum_{\beta=0}^{4-\alpha}\mathscr{K}^{(2\beta+\alpha-4)}_{(\alpha)}\cos^{\alpha}\theta\sin^{4-\alpha}\theta\,\mathrm{e}^{\jmath(2\beta+\alpha-4)\phi}. \tag{25}$$

**Remark 9.** Computing each $\Gamma^{\hat{\mathbf{s}}}_{ii,ii}$ without simplifications require a total of $(L+1)^4N^8$ loops (Eq. (21)). Thus, readers can find an implementation trick in Appendix A for reducing this number of loops to $((L+1)N)^4$. Moreover, the method detailed in Appendix A uses less memory for storing matrices. Only the cumulant tensor $\Gamma^{\mathbf{x}}_{ac,bd}(\boldsymbol{\tau})$, which is computed at the beginning of each loop, needs a large memory, i.e. up to $N^4(L+1)^4$ values to store.

Next, we make the change of variables: $\cos\phi=(1-t^2)/(1+t^2)$, $\sin\phi=2t/(1+t^2)$ with $t=\tan\phi/2$, and $\cos\theta=1/\sqrt{1+u^2}$, $\sin\theta=u/\sqrt{1+u^2}$ with $u=\tan\theta$. Thus, the numerator and the denominator of the polynomial obtained are both of degree 8 in variable $t$ and 4 in variable $u$. For maximizing contrast (20), we have to find roots of polynomial system (26), i.e. stationary points of $\Upsilon_{1,4}$, obtained by canceling the partial derivatives:

$$\left.\begin{aligned}\Phi_1(u,t) &\stackrel{\text{def}}{=} \frac{\partial\,\Upsilon_{1,4}}{\partial u}\\ \Phi_2(u,t) &\stackrel{\text{def}}{=} \frac{\partial\,\Upsilon_{1,4}}{\partial t}\end{aligned}\right\}. \tag{26}$$

After some simplifications, e.g. obvious root $u=0$ in $\Phi_2(u,t)$, polynomials obtained have different global degrees: degree 12 for $\Phi_1(u,t)$ (leading monomial is $t^8u^4$), and degree 11 for $\Phi_2(u,t)$ (leading monomial is $t^8u^3$). Then we consider only variable $u$ for $\Phi_1(u,t)$ and $\Phi_2(u,t)$, and we collect terms of same degree in $u$. Hence, system (26) can be rewritten

$$\left.\begin{aligned}\Phi_1(u,t) &= \sum_{k=0}^{4}\chi_{4-k}(t)\,u^k\\ \Phi_2(u,t) &= \sum_{k=0}^{3}\xi_{3-k}(t)\,u^k\end{aligned}\right\}. \tag{27}$$

The polynomial system (27) admits a solution if and only if the resultant, i.e. determinant of a Sylvester matrix, is null [33]. Thus, we obtain the determinant to solve

$$\begin{vmatrix}\xi_0(t) & 0 & 0 & 0 & \chi_0(t) & 0 & 0\\ \xi_1(t) & \xi_0(t) & 0 & 0 & \chi_1(t) & \chi_0(t) & 0\\ \xi_2(t) & \xi_1(t) & \xi_0(t) & 0 & \chi_2(t) & \chi_1(t) & \chi_0(t)\\ \xi_3(t) & \xi_2(t) & \xi_1(t) & \xi_0(t) & \chi_3(t) & \chi_2(t) & \chi_1(t)\\ 0 & \xi_3(t) & \xi_2(t) & \xi_1(t) & \chi_4(t) & \chi_3(t) & \chi_2(t)\\ 0 & 0 & \xi_3(t) & \xi_2(t) & 0 & \chi_4(t) & \chi_3(t)\\ 0 & 0 & 0 & \xi_3(t) & 0 & 0 & \chi_4(t)\end{vmatrix} = 0. \tag{28}$$

It turns out that this polynomial generally admits no more than 16 real roots. When all real roots are found, we have to plug them back in (23) for selecting the one leading to the *absolute* maximum of (22). It is not that hard to compute since powerful algorithms exist that can rapidly find roots of univariate polynomials. The complete algorithm can be summarized as follows.

**compute** $\Gamma^{\mathbf{w}}_{eg,fh}(\boldsymbol{\tau}+\boldsymbol{\rho})$
**set** $M$:=$N(N-1)/2$ and $T$:=$\lceil\sqrt{N\lambda_h}\,\rceil+1$
**set** $k$:=$0\ldots T$ and $p$:=$0\ldots\lambda_h$
**initialize** $\mathbf{Q}^{(k)}_p$ to identity (i.e all $(\theta,\phi)$ to zero) for all $k$ and $p$
**for** each $k$ **do**
  **for** each $p$ **do**
    $\boldsymbol{A}^{(k)}_p[z]=\mathbf{Q}^{(k)}_0\boldsymbol{Z}[z]\ldots\mathbf{Q}^{(k)}_{p-1}\boldsymbol{Z}[z]$
    $\boldsymbol{B}^{(k)}_p[z]=\boldsymbol{Z}[z]\mathbf{Q}^{(k-1)}_{p+1}\ldots\boldsymbol{Z}[z]\mathbf{Q}^{(k-1)}_{\lambda_h}$
    **compute** $\Gamma^{\mathbf{x}^{(k)}}_{ac,bd}(\boldsymbol{\tau})$ (with $x^{(k)}[z]=\boldsymbol{B}^{(k)}_p[z]\boldsymbol{w}[z]$)
    **search for** $\mathbf{Q}^{(k)}_p$ maximizing $\Upsilon_{1,4}$ (one obtains $M$ pairs of angles $(\theta,\phi)$ yielding $\mathbf{Q}^{(k)}_p$)
**return** $\boldsymbol{H}[z]=\mathbf{Q}^{(\mathrm{T})}_0\boldsymbol{Z}[z]\ldots\mathbf{Q}^{(\mathrm{T})}_p\boldsymbol{Z}[z]\mathbf{Q}^{(\mathrm{T})}_{p+1}\ldots\boldsymbol{Z}[z]\mathbf{Q}^{(\mathrm{T})}_{\lambda_h}$

### 5.2. Sweeping the pairs

Let us comment some points of this algorithm:

- The resulting tensor $\Gamma^{\mathbf{x}}_{ac,bd}(\boldsymbol{\tau})$ is composed of $N^2(L-\ell_B+1)^2$ matrices each of size $N(L-\ell_B+1)\times N(L-\ell_B+1)$. Hence, the algorithm needs to store $N^4(L-\ell_B+1)^4$ entries in memory. Simplifications are generally impossible since we need exact values for computing roots of polynomial system (27).
- In order to increase the precision of the angles, we suggest to execute $T=\lceil\sqrt{L}\,\rceil+1$ sweeps. Actually, the first angles computed are not well defined since all other angles are still null (set at stage **initialize**). Hence, when loop on each $k$ is repeated several times, angles are better estimated. This procedure reminds the well-known *Jacobi sweeping* widely used for eigenvalue computation.

**Remark 10.** In the case of a slowly time varying linear channel, the initialization step **initialize** can be modified since angles previously found give a better approximation than null values.

## 6. Computer results

For large data blocks, the performance of the algorithm described in Section 5 comes close to that of the MMSE equalizer, and the larger the block length, the larger the accuracy.

Algorithm has been tested on mixtures of length $K = 3$. We have generated 100 random channels. The equalizer has been determined for each channel with data blocks of 256 symbols or 512 symbols. Then, the equalizer has been tested with two different realizations of 5000 symbols each in order to compute the symbol error rate (SER). Thus, the minimal resolution is $(2 * 5000 * 100)^{-1} = 10^{-6}$. Fig. 3 shows median results of the 100 trials for QPSK signals, i.e. median is an estimate of average SER. This figure proves that algorithm works well on short data length since from only 256 symbols and with a noise of 16 dB, the median SER is below the minimal resolution.

For each trial, the MMSE solution is computed. Another basis for comparison is the distance (Frobenius norm) between the equalizer found and the zero-forcing (ZF) solution. In fact, the ZF solution for MIMO convolutive mixtures is simply the inverse of the channel. Then, when $K = L$, the global transfer function with the ZF equalizer $\boldsymbol{H}_{ZF}[z]$ is

$$\boldsymbol{G}[z] = \sum_n \sum_m \boldsymbol{C}(m)\boldsymbol{H}_{ZF}(n)z^{-(n-m)} = \boldsymbol{I}z^{-L}. \tag{29}$$

So, we just have to compare the global transfer function $\boldsymbol{G}[z]$ obtained with the paraunitary equalizer, to the identity matrix up to a multiplicative matrix $\boldsymbol{D}[z]$ of the form $\boldsymbol{\Lambda}[z]\boldsymbol{P}$, where $\boldsymbol{\Lambda}[z]$ is diagonal and $\boldsymbol{P}$ a permutation. In order to do this, we proceed by determining the best matrix $\boldsymbol{D}[z]$ as follows. Once $\boldsymbol{G}[z]$ is estimated, we store in an $N \times N(2L-1)$ array $\mathbb{G}$ by merely stacking the matrices one after the other. Next, we search for columns containing the entry of largest modulus in each row of $\mathbb{G}$. By comparing column indices, which have to be different modulo $N$, we find matrix $\boldsymbol{P}$ and $\boldsymbol{\Lambda}[z]$. Next, we normalize each row by its entry of largest modulus. The matrix obtained this way is denoted $\bar{\mathbb{G}}$. Then, we replace the $N$ previous entries of $\bar{\mathbb{G}}$ by zeros. This is the same as subtracting $\mathbb{D}$ to $\bar{\mathbb{G}}$, where $\mathbb{D}$ is made from $\boldsymbol{D}[z]$ with the same size as $\mathbb{G}$. Finally, we compute the Frobenius norm of the resulting matrix, i.e. $\|\bar{\mathbb{G}} - \mathbb{D}\|$. This value is the distance between the paraunitary equalizer found and the ZF solution. Medians of distances are reported in Fig. 4. We note that between 0 and 8 dB and for 512 symbols, the distance decreases rapidly from 1 to 0.25. Then, after 8 dB, the distance remains approximately constant.

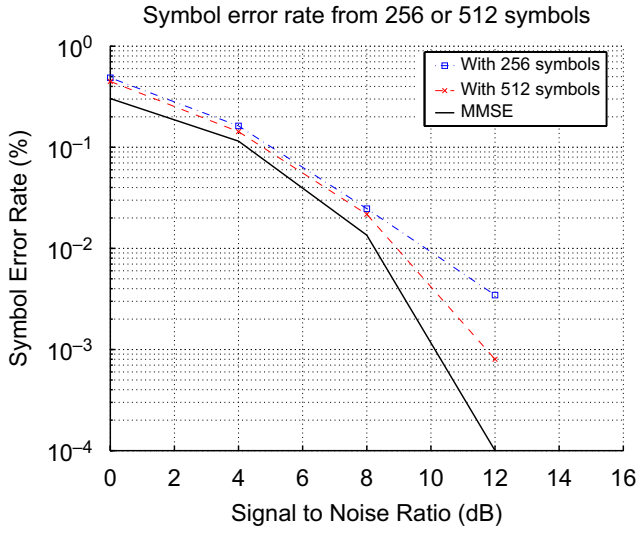



**Fig. 3.** Symbol error rate when a length-3 equalizer is built from blocks of 256 or 512 symbols. For both lengths, the median of the SER is below $10^{-6}$ after 12 dB.

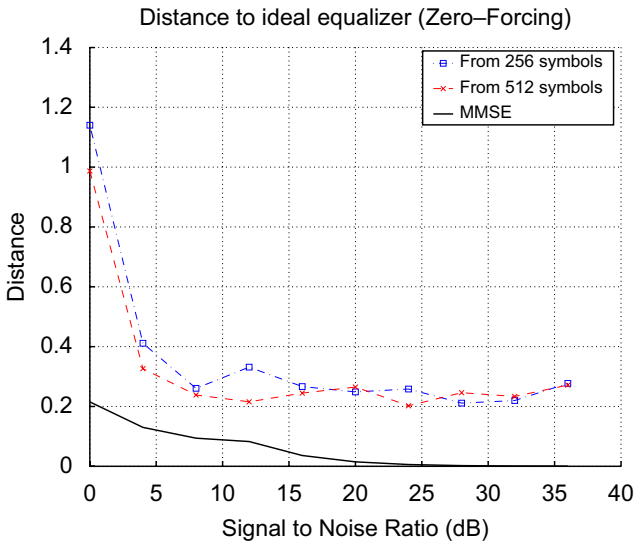



**Fig. 4.** Distance to the zero-forcing equalizer when a length-3 equalizer is built from blocks of 256 or 512 symbols. Solid line: distance between ZF and MSE.

## 7. Concluding remarks

The Blind MIMO Equalization problem can be reduced to a simpler one via a prior space–time second-order whitening, namely the equalization of a paraunitary channel. Thanks to this statement, we have proposed a parameterization of paraunitary equalizers, in order to carry out blind source separation, i.e. deconvolution without a learning sequence. Based on this parameterization, a numerical algorithm has been devised, which iteratively maximizes a contrast through successive sweeps of complex plane rotations, and eventually equalizes the channel. Moreover, we have demonstrated that the computation of the blind equalizer can be completed within a polynomial complexity, useful in burst-mode transmissions. Performances presented in the last section report an average SER of 0.35% at 12 dB of SNR, with data length of 256 symbols only. Other issues currently being addressed include robustness to channel length misadjustment, which are of prime importance as

pointed out in [34,35], and space–time whitening implementations.

## Appendix A. Complexity savings

In order to reduce the complexity of the algorithm, it is suggested to store each $\mathscr{K}$ in a matrix $\boldsymbol{T} \in \mathbb{C}^{N^4 \times N^4}$. This matrix is defined by

$$\boldsymbol{T} \stackrel{\text{def}}{=} \sum_{\tau,i} \boldsymbol{f}_i(\boldsymbol{\tau})\boldsymbol{c}(\boldsymbol{\tau})^{\mathsf{T}}, \tag{A.1}$$

where, for $N = 2$,

$$\boldsymbol{f}_i(\boldsymbol{\tau}) = \begin{bmatrix} f_i^{1111}(\tau_1,\tau_2,\tau_3,\tau_4) \\ f_i^{1112}(\tau_1,\tau_2,\tau_3,\tau_4) \\ f_i^{1121}(\tau_1,\tau_2,\tau_3,\tau_4) \\ \vdots \\ f_i^{2222}(\tau_1,\tau_2,\tau_3,\tau_4) \end{bmatrix}, \tag{A.2}$$

with $f_i^{qrst}(\tau_1,\tau_2,\tau_3,\tau_4) = A_{iq}(\tau_1)A_{ir}^*(\tau_2)A_{is}(\tau_3)A_{it}^*(\tau_4)$, and

$$\boldsymbol{c}(\boldsymbol{\tau}) = \begin{bmatrix} \Gamma^{\boldsymbol{x}}_{11,11}(\tau_1,\tau_2,\tau_3,\tau_4) \\ \Gamma^{\boldsymbol{x}}_{11,12}(\tau_1,\tau_2,\tau_3,\tau_4) \\ \Gamma^{\boldsymbol{x}}_{12,11}(\tau_1,\tau_2,\tau_3,\tau_4) \\ \vdots \\ \Gamma^{\boldsymbol{x}}_{22,22}(\tau_1,\tau_2,\tau_3,\tau_4) \end{bmatrix}. \tag{A.3}$$

For only two observations, we have $\boldsymbol{T} \in \mathbb{C}^{16\times 16}$. Hence we obtain the $N^8 = 256$ possible values of (21) by summing $L^4$ matrices of size $16 \times 16$. Then, we just have to define each $\mathscr{K}$ by summing entries of $\boldsymbol{T}$. Moreover, we see that elements of $\boldsymbol{f}_i(\boldsymbol{\tau})$ and $\boldsymbol{c}(\boldsymbol{\tau})$ have the same ordering, i.e. $q \equiv a, r \equiv b, s \equiv c, t \equiv d$. Hence, both vectors can be constructed thanks to the same loops. So, we reduce the number of loops by $N^4$, and hence the computation time.

We immediately see another simplification if we include indices $\tau_1,\tau_2,\tau_3,\tau_4$ in vectors $\boldsymbol{f}_i$ and $\boldsymbol{c}$, e.g. by Kronecker product. In this case, both vectors are of length $(N(L+1))^4$ and then $\boldsymbol{T}$ is $(N(L+1))^4 \times (N(L+1))^4$. So, we only have $(N(L+1))^4$ loops for building vectors $\boldsymbol{f}_i$ and $\boldsymbol{c}$, and then we have to make only one product $\boldsymbol{f}_i\boldsymbol{c}^{\mathsf{T}}$ for matrix $\boldsymbol{T}$. Nevertheless, the memory capacity is then the drawback of this method since more memory is used for storing matrix $\boldsymbol{T}$. As a consequence, this last simplification is probably interesting for small values of $L$.

## References


[1] D. Donoho, On minimum entropy deconvolution, in: A. Press (Ed.), Applied Time Series Analysis II, 1981, pp. 565–609.
[2] P. Comon, Contrasts for multichannel blind deconvolution, IEEE Signal Process. Lett. 3 (7) (July 1996) 209–211.
[3] J.K. Tugnait, Identification and deconvolution of multichannel nonGaussian processes using higher order statistics and inverse filter criteria, IEEE Trans. Signal Process. 45 (March 1997) 658–672.
[4] E. Moreau, J.C. Pesquet, Generalized contrasts for multichannel blind deconvolution of linear systems, IEEE Signal Process. Lett. 4 (6) (June 1997) 182–183.
[5] J.K. Tugnait, On blind separation of convolutive mixtures of independent linear signals in unknown additive noise, IEEE Trans. Signal Process. 46 (11) (1998) 3117–3123.
[6] Y. Inouye, R. Liu, Direct blind deconvolution of multiuser-multichannel system driven by temporally white source signals, in: Proceedings of ICASSP, vol. 5, Phoenix, AZ, USA, 1999, pp. 2889–2892.
[7] Z. Ding, T. Nguyen, Stationary points of a kurtosis maximization algorithm for blind signal separation and antenna beamforming, IEEE Trans. Signal Process. 48 (6) (June 2000) 1587–1596.
[8] N. Thirion-Moreau, E. Moreau, Generalized criteria for blind multivariate signal equalization, IEEE Signal Process. Lett. 9 (2) (February 2002) 72–74.
[9] I. Fijalkow, A. Touzni, J.R. Treichler, Fractionally spaced equalization using CMA: robustness to channel noise and lack of disparity, IEEE Trans. Signal Process. 45 (1) (January 1997) 56–66.
[10] A.J.V. der Veen, A. Paulraj, An analytical constant modulus algorithm, IEEE Trans. Signal Process. 44 (5) (May 1996) 1136–1155.
[11] C.B. Papadias, Globally convergent blind source separation based on a multiuser kurtosis maximization criterion, IEEE Trans. Signal Process. 48 (12) (December 2000) 3508–3519.
[12] A. Touzni, I. Filjakow, M.G. Larimore, J.R. Treichler, A globally convergent approach for blind MIMO adaptive deconvolution, IEEE Trans. Signal Process. 49 (6) (June 2001) 1166–1178.
[13] O. Grellier, P. Comon, B. Mourrain, P. Trebuchet, Analytical blind channel identification, IEEE Trans. Signal Process. 50 (9) (September 2002).
[14] D. Yellin, B. Porat, Blind identification of FIR systems excited by discrete-alphabet inputs, IEEE Trans. Signal Process. 41 (3) (1993) 1331–1339.
[15] S. Talwar, M. Viberg, A. Paulraj, Blind estimation of multiple co-channel digital signals arriving at an antenna array: part I, algorithms, IEEE Trans. Signal Process. 44 (5) (May 1996) 1184–1197.
[16] P. Comon, Contrasts, independent component analysis, and blind deconvolution, Int. J. Adapt. Control Signal Process. 18 (3) (April 2004) 225–243 (special issue).
[17] N.D. Sidiropoulos, G.B. Giannakis, R. Bro, Blind PARAFAC receivers for DS-CDMA systems, IEEE Trans. Signal Process. 48 (3) (March 2000) 810–823.
[18] D. Nion, L.D. Lathauwer, A block factor analysis based receiver for blind multi-user access in wireless communications, in: Proceedings of ICASSP, Toulouse, France, May 14–19, 2006, pp. 825–828.
[19] L. Rota, P. Comon, S. Icart, Blind equalization of MIMO channels, in: Proceedings of Signal Processing Advances in Wireless Communications, Rome, Italy, June 15–18, 2003.
[20] L. Rota, P. Comon, S. Icart, Blind MIMO Equalization under paraunitary constraint, I3S Report RR-2002-35-FR, Laboratoire I3S, September 2002.
[21] P.P. Vaidyanathan, Multirate Systems and Filter Banks, Prentice-Hall, London, 1993.
[22] Y. Inouye, R. Liu, A system-theoretic foundation for blind equalization of an FIR MIMO channel system, IEEE Trans. Circuits Systems 49 (4) (April 2002) 425–436.
[23] A.H. Sayed, T. Kailath, A survey of spectral factorization methods, Numer. Linear Algebra Appl. 8 (2001) 467–496.
[24] L.D. Lathauwer, B.D. Moor, J. Vandewalle, An algebraic approach to blind MIMO identification, in: Proceedings of the Second International Workshop on Independent Component Analysis and Blind Signal Separation (ICA 2000), Helsinki, Finland, June 19–22, 2000, pp. 211–214.
[25] L.D. Lathauwer, B.D. Moor, J. Vandewalle, An algebraic approach to the blind identification of paraunitary filters, in: Proceedings of IEEE Wireless Communications and Networking Conference, Chicago, USA, September 23–28, 2000.
[26] E. Moreau, A block algorithm for blind signal deconvolution, in: Proceedings of SPAWC'97, Paris, France, April 1997, pp. 93–96.
[27] P. Comon, E. Moreau, L. Rota, Blind separation of convolutive mixtures: a contrast-based joint diagonalization approach, in: Proceedings of the International Conference on Independent Component Analysis, San Diego, December 9–13, 2001, pp. 686–691.
[28] P. Comon, L. Rota, Blind separation of independent sources from convolutive mixtures, in: IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol. E86-A(3), March 2003, pp. 550–563.
[29] P. McCullagh, Tensor Methods in Statistics, Monographs on Statistics and Applied Probability, Chapman & Hall, London, 1987.
[30] X. Gao, T.Q. Nguyen, G. Strang, On factorization of $M$-channel paraunitary filterbanks, IEEE Trans. Signal Process. 49 (7) (July 2001) 1433–1446.

[31] G.H. Golub, C.F.V. Loan, Matrix Computation, third ed., John Hopkins University Press, Baltimore, MD, 1996.
[32] N. Delfosse, P. Loubaton, Adaptive blind separation of independent sources: a deflation approach, Signal Processing 45 (1995) 59–83.
[33] D. Cox, J. Little, D.O Shea, Using Algebraic Geometry, Springer, New York, 1998.
[34] Y.Y.T. Acar, A. Petropulu, Blind MIMO system estimation based on PARAFAC decomposition of tensors formed based hos of the system output, IEEE Trans. Signal Process. 54 (11) (2006) 4156–4168.
[35] B. Chen, A.P. Petropulu, Frequency domain blind MIMO system identification based on second- and higher-order statistics, IEEE Trans. Signal Process. 49 (8) (2001) 1677–1688.

Signal Processing 92 (2012) 617–624

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

# On Jacobi-type methods for blind equalization of paraunitary channels

Mikael Sørensen [a,*], Lieven De Lathauwer [a], Sylvie Icart [b], Luc Deneire [b]

[a] K.U.Leuven - E.E. Dept. (ESAT) - SCD-SISTA, Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium
[b] Les Algorithmes - Euclide-B, 06903 Sophia Antipolis, France

ARTICLE INFO

*Article history:*
Received 20 September 2010
Received in revised form
31 March 2011
Accepted 6 July 2011
Available online 19 July 2011

*Keywords:*
Blind equalization
Blind deconvolution
Tensors
Jacobi method

ABSTRACT

In this paper a study of the cumulant-based blind equalization algorithms PAJOD and PAFA is conducted. Both algorithms assume that the data have been pre-whitened and hence the problem reduces to the estimation of paraunitary channels. The main contribution of this paper is an efficient implementation of the PAJOD algorithm called PAJOD2. Second, a performance comparison between the PAJOD, PAJOD2 and PAFA algorithms is reported.

© 2011 Elsevier B.V. All rights reserved.

## 1. Introduction

Blind equalization of linear time-invariant MIMO channels refers to channel equalization techniques where only the observed signal is known. The observed signal is assumed to consist of an unknown convolutive mixture of input signals. The cumulant-based blind equalization algorithms Partial Approximate JOint Diagonalization (PAJOD) and PAraunitary FActorization (PAFA) were proposed in [1] and in [2], respectively. Both are based on contrast maximization and the working assumption for both algorithms is that the data have been pre-whitened. A method to perform pre-whitening has been proposed in [3].

Due to the pre-whitening the problem reduces to a search for a paraunitary equalizer. The PAFA algorithm looks for a paraunitary equalizer while the PAJOD algorithm only searches for a semi-unitary equalizer. The PAFA and PAJOD algorithms both consist of a Jacobi-type iteration where the Jacobi subproblem is solved by a computationally demanding resultant-based procedure. The PAFA algorithm requires the rooting of a 56th degree polynomial in each Jacobi subproblem. PAJOD requires the rooting of either a 3rd or 24th degree polynomial in each of its Jacobi subproblems, as will be explained later.

The main contribution of this paper is a more efficient implementation of the PAJOD algorithm called PAJOD2. Part of this work has been presented in [4]. The paraunitary equalizer PAFA fully takes the structure of the problem into account while the semi-unitary equalizer PAJOD only partially exploits the structure of the problem. However, the PAJOD algorithm is less computationally demanding than the PAFA algorithm. Hence, a comparison of the PAJOD, PAJOD2 and PAFA algorithms based on computer simulations will be reported.

The paper is organized as follows. First the notation used throughout the paper will be introduced. Since the algorithms are based on paraunitary filters and contrast optimization, a few basic notions about paraunitary filters and contrasts will be presented before discussing the system model. Next, in Section 2 a brief review of the PAJOD and PAFA algorithms is given. Furthermore, the more efficient implementation of the PAJOD algorithm called PAJOD2 will be presented. Section 3 will compare the PAJOD, PAJOD2 and PAFA methods based on computer

* Corresponding author.
*E-mail addresses:* Mikael.Sorensen@kuleuven-kortrijk.be (M. Sørensen), Lieven.DeLathauwer@kuleuven-kortrijk.be (L. De Lathauwer), icart@i3s.unice.fr (S. Icart), deneire@i3s.unice.fr (L. Deneire).

0165-1684/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.sigpro.2011.07.005

simulations. We end the paper with a conclusion in Section 4.

### 1.1. Notation

Let $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{C}[z]$ denote the set of natural, integer, real, complex numbers and the set of polynomials in $z$ with coefficients in $\mathbb{C}$, respectively. Furthermore, let $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^\dagger$, $\mathrm{Re}\{\cdot\}$, $\mathrm{Im}\{\cdot\}$ and $\|\cdot\|_F$ denote the conjugate, transpose, conjugate-transpose, pseudo-inverse, real part, imaginary part and the Frobenius norm of a matrix, respectively. The operator $\mathrm{diag}(\cdot)$ sets all the off-diagonal elements of a matrix equal to zero. Let $\mathbf{I}_R \in \mathbb{C}^{R\times R}$ denote the identity matrix. Given $\mathbf{A} \in \mathbb{C}^{m\times n}$, then $\mathbf{A}_{ij}$ denotes the $i$th row–$j$th column entry of $\mathbf{A}$. Finally, let $\mathbf{H}(z) = \sum_n \mathbf{H}(n)z^{-n}$.

### 1.2. Paraunitary filter

A filter $\mathbf{H}(z) = \sum_{l=0}^{L-1} \mathbf{H}(l)z^{-l} \in \mathbb{C}[z]^{R\times R}$ is paraunitary if $\mathbf{H}^H(1/z^*)\mathbf{H}(z) = \mathbf{I}_R$. The paraunitary filter $\mathbf{H}(z)$ satisfies the properties [5]:

- The inverse filter $\mathbf{H}^{-1}(z) = \mathbf{H}^H(1/z^*)$ is paraunitary.
- The channel impulse response matrix $\overline{\mathbf{H}} = [\mathbf{H}(0), \mathbf{H}(1), \ldots, \mathbf{H}(L-1)] \in \mathbb{C}^{R\times RL}$ is a semi-unitary matrix, i.e., $\overline{\mathbf{H}}\overline{\mathbf{H}}^H = \mathbf{I}_R$.

### 1.3. Contrast optimization

The notion of contrast optimization was introduced in [6] and applied in the framework of MIMO equalization in [7]. Let $\mathcal{H}$ and $\mathcal{S}$ be the set of paraunitary filters and the set of transmitted symbol sequences, respectively. Furthermore, let $\mathcal{H}\cdot\mathcal{S}$ denote the set of recovered symbol sequences and $\mathcal{T}$ denote the set of paraunitary equalizers that do not violate the working assumptions on $\mathcal{S}$ specified below. Moreover, let $\mathbf{I}$ denote the identity operator, then a function $\mathcal{J}(\mathbf{H};\mathbf{x})$ is called a contrast if it satisfies the properties [7]:

- *Invariance*: $\mathcal{J}(\mathbf{H};\mathbf{x}) = \mathcal{J}(\mathbf{I};\mathbf{x}), \forall \mathbf{H}\in\mathcal{T}, \forall \mathbf{x}\in\mathcal{H}\cdot\mathcal{S}$.
- *Domination*: $\mathcal{J}(\mathbf{H};\mathbf{x}) \le \mathcal{J}(\mathbf{I};\mathbf{x}), \forall \mathbf{H}\in\mathcal{H}, \forall \mathbf{x}\in\mathcal{S}$.
- *Discrimination*: $\mathcal{J}(\mathbf{H};\mathbf{x}) = \mathcal{J}(\mathbf{I};\mathbf{x}), \forall \mathbf{x}\in\mathcal{S} \Rightarrow \mathbf{H}\in\mathcal{T}$.

Under the assumption that there exists an equalizer that will fully recover the symbols, an equalizer corresponding to the global maximum of the contrast function is guaranteed to recover the symbol sequence.

#### 1.3.1. System model

Let $\mathbf{s}(n)$, $\mathbf{x}(n) \in \mathbb{C}^R$ be the symbol and observation vector at time instant $n \in \mathbb{N}$, respectively. Assume that $\mathbf{s}(n)$ and $\mathbf{x}(n)$ are related via

$$\mathbf{x}(n) = \sum_{k=0}^{K-1} \mathbf{F}(k)\mathbf{s}(n-k),$$

where $\mathbf{F}(k) \in \mathbb{C}^{R\times R}$, $k \in [0,K-1]$, is the channel impulse response of the paraunitary filter $\mathbf{F}(z)$. The problem is to estimate the symbol sequence $\{\mathbf{s}(n)\}$ from the observation sequence $\{\mathbf{x}(n)\}$. This is done by the equalizer $\mathbf{H}(z)$ such that

$$\mathbf{y}(n) = \sum_{l=0}^{L-1} \mathbf{H}(l)\mathbf{x}(n-l) = \sum_{l=0}^{L-1}\sum_{k=0}^{K-1} \mathbf{H}(l)\mathbf{F}(k)\mathbf{s}(n-l-k),$$

where $\mathbf{H}(l) \in \mathbb{C}^{R\times R}$, $l \in [0,L-1]$, are the channel impulse response coefficients of $\mathbf{H}(z)$ and $\mathbf{y}(n)$ is the recovered symbol vector at time instant $n$. Since $\mathbf{F}(z)$ is paraunitary, we know that the equalizer $\mathbf{H}(z)$ is also paraunitary and that $K=L$. In the PAJOD and PAFA methods, the working assumptions on the transmitted signal sequences are:

- $\mathbf{s}_r(n)$ are mutually independent i.i.d., zero-mean processes with unit-variance for all $r \in [1,R]$.
- $\mathbf{s}(n)$ is stationary up to order 4 and hence the marginal cumulants of order 4 do not depend on $n$.
- At most one source has zero marginal cumulant of order 4.

## 2. Paraunitary equalization algorithms

The PAJOD and PAFA blind paraunitary equalization algorithms will be briefly reviewed and the computationally improved version of PAJOD algorithm called PAJOD2 will be introduced.

### 2.1. PAJOD

In [1], the cumulants of the observed data are stored in a set of $RL\times RL$ matrices $\mathbf{M}(\mathbf{b},\gamma)$ in such a way that for a fixed pair $(\mathbf{b},\gamma) = ([b_1,b_2],[\gamma_1,\gamma_2])$ we have the relation

$$\mathbf{M}_{\alpha_1 R+a_1,\alpha_2 R+a_2}(\mathbf{b},\gamma) = \mathrm{Cum}[y_{a_1}(n-\alpha_1), y^*_{a_2}(n-\alpha_2), y_{b_1}(n-\gamma_1), y^*_{b_2}(n-\gamma_2)].$$

Moreover, in [1] it is shown that the function

$$\mathcal{J}_2^2 = \sum_{\mathbf{b},\gamma} \|\mathrm{diag}(\overline{\mathbf{H}}\mathbf{M}(\mathbf{b},\gamma)\overline{\mathbf{H}}^H)\|_F^2 \tag{1}$$

is a contrast function, where $\|\mathrm{diag}(\mathbf{A})\|_F^2 = \sum_i |\mathbf{A}_{ii}|^2$ and $\overline{\mathbf{H}} = [\mathbf{H}(0),\mathbf{H}(1),\ldots,\mathbf{H}(L-1)] \in \mathbb{C}^{R\times RL}$. Due to the paraunitary assumption on $\mathbf{H}(z)$, the matrix $\overline{\mathbf{H}}$ is a semi-unitary matrix, i.e., $\overline{\mathbf{H}}\overline{\mathbf{H}}^H = \mathbf{I}_R$.

#### 2.1.1. Jacobi procedure for semi-unitary matrices

To numerically find the semi-unitary matrix $\overline{\mathbf{H}}$ that will maximize the contrast (1) a Jacobi procedure was proposed in [1]. This procedure can be seen as a double extension of the JADE algorithm [8,9]. First, the unknown matrix is semi-unitary instead of unitary. Second, only the $R$ first diagonal entries are of interest.

A Jacobi procedure is based on the fact that any $RL\times RL$ unitary matrix with determinant equal to one can be parametrized as a product of Givens rotations [10]:

$$\mathbf{V} = \prod_{p=1}^{RL-1}\prod_{q=p+1}^{RL} \Theta[p,q]^H,$$

where $\Theta[p,q]$ is equal to the identity matrix, except for entries

$$\Theta_{pp}[p,q] = \Theta_{qq}[p,q] = \cos(\theta[p,q]),$$

$$\Theta_{qp}[p,q] = -\Theta_{pq}[p,q]^* = \sin(\theta[p,q])e^{i\phi[p,q]}, \theta[p,q], \phi[p,q] \in \mathbb{R}.$$

Let $\mathbf{V}$ denote the product of Givens matrices with the initial value $\mathbf{V}=\mathbf{I}_{RL}$. The updating rules are given by $\mathbf{M}(\mathbf{b},\gamma)\leftarrow\Theta[p,q]^H\mathbf{M}(\mathbf{b},\gamma)\Theta[p,q]$ and $\mathbf{V}\leftarrow\Theta[p,q]^H\mathbf{V}$. In the PAJOD algorithm the semi-unitary matrix $\overline{\mathbf{H}}$ is determined as the first $R$ rows of the unitary matrix $\mathbf{V}$. The Givens rotation matrix $\Theta[p,q]$ is chosen as the maximizer of

$$\begin{aligned}\mathcal{J}_2^2(p,q)&=\sum_{\mathbf{b},\gamma}\sum_{k=1}^{R}|(\Theta[p,q]^H\mathbf{M}(\mathbf{b},\gamma)\Theta[p,q])_{kk}|^2\\&=\sum_{\mathbf{b},\gamma}\sum_{k=1}^{R}\left|\sum_{\eta,\mu=1}^{RL}\Theta^*_{\eta k}[p,q]\Theta_{\mu k}[p,q]\mathbf{M}(\mathbf{b},\gamma)_{\eta\mu}\right|^2\\&=\sum_{\omega}\sum_{k=1}^{R}\left|\sum_{\eta,\mu=1}^{RL}\Theta^*_{\eta k}[p,q]\Theta_{\mu k}[p,q]\mathbf{M}^{(\omega)}_{\eta\mu}\right|^2.\end{aligned}\tag{2}$$

Notice that we denote $\mathbf{M}(\mathbf{b},\gamma)$ by $\mathbf{M}^{(\omega)}$ for simplicity. Since plane rotations where $p>R$ do not have any effect on the first $R$ rows of the matrices $\mathbf{M}^{(\omega)}$ only Givens rotations where $p\le R$ are considered. Furthermore one has to distinguish between the cases where $q\le R$ and $q>R$. The problems are illustrated in Fig. 1. When $q\le R$, then the aim of the Givens rotation matrix $\Theta[p,q]$ is to jointly diagonalize the set of matrices $\{\mathbf{M}^{(\omega)}\}$ by maximizing the entries $\mathbf{M}^{(\omega)}_{pp}$ and $\mathbf{M}^{(\omega)}_{qq}$ for all $\omega$ as illustrated in Fig. 1(a). On the other hand, when $q>R$, then the aim of the Givens rotation matrix $\Theta[p,q]$ reduces to jointly diagonalize the set of matrices $\{\mathbf{M}^{(\omega)}\}$ by maximizing the entry $\mathbf{M}^{(\omega)}_{pp}$ for all $\omega$ as illustrated in Fig. 1(b).

Let $\overline{\mathbf{M}}^{(\omega)}=\Theta[p,q]^H\mathbf{M}^{(\omega)}\Theta[p,q]$ and for notational convenience let $c=\cos(\theta[p,q])$ and $s=\sin(\theta[p,q])e^{i\phi[p,q]}$. Then for the case where $q\le R$ Eq. (2) is equal to

$$\mathcal{J}_2^2(p,q)|_{q\le R}=\sum_{\omega}|\overline{\mathbf{M}}^{(\omega)}_{pp}|^2+|\overline{\mathbf{M}}^{(\omega)}_{qq}|^2,\tag{3}$$

where

$$\overline{\mathbf{M}}^{(\omega)}_{pp}=[c\ \ s^*]\begin{bmatrix}\mathbf{M}^{(\omega)}_{pp}&\mathbf{M}^{(\omega)}_{pq}\\\mathbf{M}^{(\omega)}_{qp}&\mathbf{M}^{(\omega)}_{qq}\end{bmatrix}\begin{bmatrix}c\\s\end{bmatrix}$$

and

$$\overline{\mathbf{M}}^{(\omega)}_{qq}=[c\ \ -s]\begin{bmatrix}\mathbf{M}^{(\omega)}_{pp}&\mathbf{M}^{(\omega)}_{pq}\\\mathbf{M}^{(\omega)}_{qp}&\mathbf{M}^{(\omega)}_{qq}\end{bmatrix}\begin{bmatrix}c\\-s^*\end{bmatrix}.$$

The maximization problem (3) is equivalent to the JADE diagonalization problem and therefore the JADE algorithm [8] can be applied to solve this problem.

When $q>R$ only the first diagonal term should be maximized, and Eq. (2) reduces to

$$\mathcal{J}_2^2(p,q)|_{q>R}=\sum_{p}|\overline{\mathbf{M}}^{(\omega)}_{pp}|^2.\tag{4}$$

In [1] a resultant based [11] approach was taken to solve the maximization problem (4). It amounted to the rooting of a polynomial of degree 24 that has at most eight real roots.

An outline of the PAJOD algorithm is given in Algorithm 1.

**Algorithm 1.** Outline of the PAJOD procedure.

```
Estimate the cumulant tensor M(b,γ)
Initialize V = I_RL
Repeat until convergence
for p = 1 to R do
  for q = p+1 to RL do
    if q ≤ R then
      calculate optimal Θ[p,q] from J_2^2(p,q)|_{q≤R} according to [8].
    else
      calculate optimal Θ[p,q] from J_2^2(p,q)|_{q>R} according to [1].
    end if
    M^(k) ← Θ[p,q]^H M^(k) Θ[p,q]
    V ← Θ[p,q]^H V
  end for
end for
Set H = V(:,1:R)
```

### 2.2. PAJOD2

This section will introduce the PAJOD2 algorithm, which is a computationally improved version of the PAJOD algorithm. The overall algorithmic structure is still as in Algorithm 1. When $q\le R$, then the PAJOD2 algorithm will apply the JADE algorithm to solve the Jacobi subproblem, just as in the PAJOD case.

However, for the case $q>R$ a more efficient eigenvector-based approach will be proposed. In the derivation we

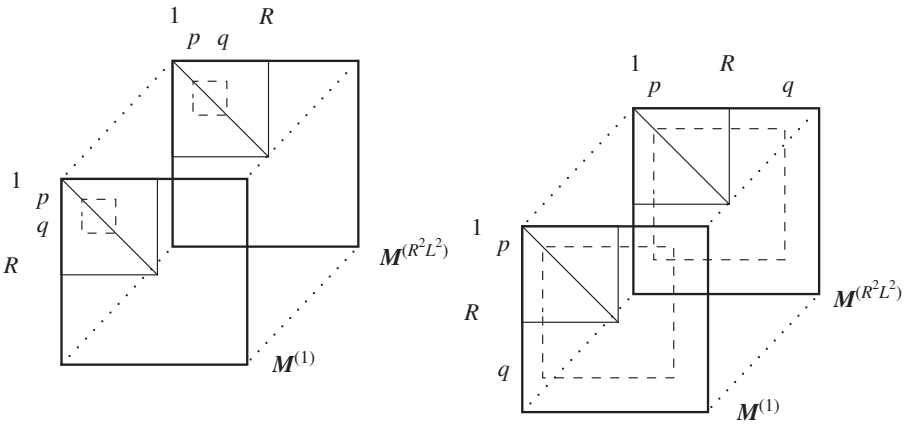



**Fig. 1.** The figures illustrate the PAJOD optimization problems for the case when $\mathbf{b}$ varies in $\{1,\ldots,R\}^2$ and $\gamma$ in $\{0,\ldots,L-1\}^2$. (a) $q\le R$. (b) $q>R$.

will make use of the trigonometric identities

$$2\cos^2(\theta) = (1+\cos(2\theta)),$$

$$2\sin^2(\theta) = (1-\cos(2\theta)),$$

$$2\cos(\theta)\sin(\theta) = \sin(2\theta),$$

$$\cos(2\phi) = \cos^2(\phi) - \sin^2(\phi),$$

$$\sin(2\phi) = 2\cos(\phi)\sin(\phi).$$

Let $\theta = \theta[p,q]$, $\phi = \phi[p,q]$, $\hat{c} = \cos(2\theta)$, $\hat{s} = \sin(2\theta)$, $\alpha^{(\omega)} = \mathbf{M}_{pp}^{(\omega)} - \mathbf{M}_{qq}^{(\omega)}$ and $\beta^{(\omega)} = \mathbf{M}_{pp}^{(\omega)} + \mathbf{M}_{qq}^{(\omega)}$, then Eq. (4) can be written as

$$\begin{aligned}\mathcal{J}_2^2(p,q)|_{q>R} = &\frac{1}{4}\sum_{\omega}|\beta^{(\omega)}|^2 + |\alpha^{(\omega)}|^2\hat{c}^2 + 2\,\mathrm{Re}\{\beta^{(\omega)*}\alpha^{(\omega)}\}\hat{c}\\ &+ (|\mathbf{M}_{pq}^{(\omega)}|^2 + |\mathbf{M}_{qp}^{(\omega)}|^2 + 2\mathrm{Re}\{\mathbf{M}_{pq}^{(\omega)*}\mathbf{M}_{qp}^{(\omega)}e^{i2\phi}\})\hat{s}^2\\ &+ 2\,\mathrm{Re}\{\alpha^{(\omega)}(\mathbf{M}_{pq}^{(\omega)*}e^{i\phi} + \mathbf{M}_{qp}^{(\omega)*}e^{-i\phi})\}\hat{c}\hat{s}\\ &- 2\,\mathrm{Re}\{\beta^{(\omega)*}(\mathbf{M}_{pq}^{(\omega)}e^{-i\phi} + \mathbf{M}_{qp}^{(\omega)}e^{i\phi})\}\hat{s}. \end{aligned} \tag{5}$$

By inspection of (5) we can identify the term independent of $\theta$ and $\phi$ as

$$k = \frac{1}{4}\sum_{\omega}|\beta^{(\omega)}|^2 = \frac{1}{4}\sum_{\omega}|\mathbf{M}_{pp}^{(\omega)} + \mathbf{M}_{qq}^{(\omega)}|^2. \tag{6}$$

The linear terms in the variables $\hat{c}$ and $\hat{s}$ of (5) can be written as

$$\begin{aligned} l(\hat{c},\hat{s}) &= \frac{1}{2}\sum_{\omega}\mathrm{Re}\{\beta^{(\omega)*}\alpha^{(\omega)}\}\hat{c} - \mathrm{Re}\{\beta^{(\omega)*}(\mathbf{M}_{pq}^{(\omega)}e^{-i\phi} + \mathbf{M}_{qp}^{(\omega)}e^{i\phi})\}\hat{s}\\ &= \frac{1}{2}\sum_{\omega}\mathrm{Re}\{\beta^{(\omega)*}\alpha^{(\omega)}\}\hat{c} - \mathrm{Re}\{\beta^{(\omega)*}(\mathbf{M}_{pq}^{(\omega)} + \mathbf{M}_{qp}^{(\omega)})\}\hat{s}\cos(\phi)\\ &\quad + \mathrm{Re}\{i\beta^{(\omega)*}(\mathbf{M}_{pq}^{(\omega)} - \mathbf{M}_{qp}^{(\omega)})\}\hat{s}\,\sin(\phi)\\ &= \sum_{\omega}\mathbf{g}^{(\omega)T}\mathbf{v}, \end{aligned} \tag{7}$$

where

$$\mathbf{v} = \begin{bmatrix} \cos(2\theta[p,q]) \\ \sin(2\theta[p,q])\cos(\phi[p,q]) \\ \sin(2\theta[p,q])\sin(\phi[p,q]) \end{bmatrix},$$

$$\mathbf{z}^{(\omega)} = \frac{1}{2}\begin{bmatrix} \mathbf{M}_{pp}^{(\omega)} - \mathbf{M}_{qq}^{(\omega)} \\ -(\mathbf{M}_{pq}^{(\omega)} + \mathbf{M}_{qp}^{(\omega)}) \\ i(\mathbf{M}_{pq}^{(\omega)} - \mathbf{M}_{qp}^{(\omega)}) \end{bmatrix},$$

$$\mathbf{g}^{(\omega)} = \mathrm{Re}\{(\mathbf{M}_{pp}^{(\omega)} + \mathbf{M}_{qq}^{(\omega)})^*\mathbf{z}^{(\omega)}\}.$$

The quadratic terms in the variables $\hat{c}$ and $\hat{s}$ of (5) will now be written as

$$\begin{aligned} q(\hat{c},\hat{s}) &= \frac{1}{4}\sum_{\omega}|\alpha^{(\omega)}|^2\hat{c}^2 + 2\,\mathrm{Re}\{\alpha^{(\omega)}(\mathbf{M}_{pq}^{(\omega)*}e^{i\phi} + \mathbf{M}_{qp}^{(\omega)*}e^{-i\phi})\}\hat{c}\hat{s}\\ &\quad + (|\mathbf{M}_{pq}^{(\omega)}|^2 + |\mathbf{M}_{qp}^{(\omega)}|^2 + 2\,\mathrm{Re}\{\mathbf{M}_{pq}^{(\omega)*}\mathbf{M}_{qp}^{(\omega)}e^{i2\phi}\})\hat{s}^2\\ &= \frac{1}{4}\sum_{\omega}|\alpha^{(\omega)}|^2\hat{c}^2 + (|\mathbf{M}_{pq}^{(\omega)}|^2 + |\mathbf{M}_{qp}^{(\omega)}|^2)\hat{s}^2(\cos^2(\phi)\\ &\quad + \sin^2(\phi)) + 2\,\mathrm{Re}\{\mathbf{M}_{pq}^{(\omega)*}\mathbf{M}_{qp}^{(\omega)}\}\hat{s}^2\cos(2\phi)\\ &\quad + 2\,\mathrm{Re}\{i\mathbf{M}_{pq}^{(\omega)*}\mathbf{M}_{qp}^{(\omega)}\}\hat{s}^2\sin(2\phi)\\ &\quad + 2\,\mathrm{Re}\{\alpha^{(\omega)}(\mathbf{M}_{pq}^{(\omega)} + \mathbf{M}_{qp}^{(\omega)})^*\}\hat{c}\hat{s}\,\cos(\phi)\\ &\quad + 2\,\mathrm{Re}\{i\alpha^{(\omega)}(\mathbf{M}_{pq}^{(\omega)} - \mathbf{M}_{qp}^{(\omega)})^*\}\hat{c}\hat{s}\,\sin(\phi)\\ &= \frac{1}{4}\sum_{\omega}|\alpha^{(\omega)}|^2\hat{c}^2 + (|\mathbf{M}_{pq}^{(\omega)}|^2 + |\mathbf{M}_{qp}^{(\omega)}|^2)\hat{s}^2(\cos^2(\phi)\\ &\quad + \sin^2(\phi)) + 2\,\mathrm{Re}\{\mathbf{M}_{pq}^{(\omega)*}\mathbf{M}_{qp}^{(\omega)}\}\hat{s}^2(\cos^2(\phi) - \sin^2(\phi))\\ &\quad + 4\,\mathrm{Re}\{i\mathbf{M}_{pq}^{(\omega)*}\mathbf{M}_{qp}^{(\omega)}\}\hat{s}^2\cos(\phi)\sin(\phi)\\ &\quad + 2\,\mathrm{Re}\{\alpha^{(\omega)}(\mathbf{M}_{pq}^{(\omega)} + \mathbf{M}_{qp}^{(\omega)})^*\}\hat{c}\hat{s}\cos(\phi)\\ &\quad + 2\,\mathrm{Re}\{\alpha^{(\omega)}(\mathbf{M}_{pq}^{(\omega)} - \mathbf{M}_{qp}^{(\omega)})^*\}\hat{c}\hat{s}\sin(\phi)\\ &= \mathbf{v}^T\sum_{\omega}\mathbf{G}^{(\omega)}\mathbf{v}, \end{aligned} \tag{8}$$

where $\mathbf{G}^{(\omega)} = \mathrm{Re}\{\mathbf{z}^{(\omega)}\mathbf{z}^{(\omega)H}\}$.

From Eqs. (6)–(8), Eq. (4) can be reformulated as

$$\mathcal{J}_2^2(p,q)|_{q>R} = \mathbf{v}^T\mathbf{G}\mathbf{v} + \mathbf{g}^T\mathbf{v} + k, \tag{9}$$

where $\mathbf{G} = \sum_{\omega}\mathbf{G}^{(\omega)}$ and $\mathbf{g} = \sum_{\omega}\mathbf{g}^{(\omega)}$. We should maximize (9) under the constraint $\|\mathbf{v}\|_F^2 = 1$.

Maximizing (9) subject to the constraint that $\|\mathbf{v}\|_F^2 = 1$ is a classical problem. In the context of source separation, it for instance appeared in [12]. Using the Lagrange multiplier method leads to

$$2(\mathbf{G} + \lambda\mathbf{I}_3)\mathbf{v} + \mathbf{g} = 0, \quad \lambda \in \mathbb{R}. \tag{10}$$

Assuming that $(\mathbf{G} + \lambda\mathbf{I}_3)^{-1}$ exists, we have

$$\mathbf{v} = -\tfrac{1}{2}(\mathbf{G} + \lambda\mathbf{I}_3)^{-1}\mathbf{g}.$$

Given the eigenvalue decomposition $\mathbf{G} = \mathbf{E}\Lambda\mathbf{E}^T$, we have

$$\|\mathbf{v}\|_F^2 = 1 \Leftrightarrow \frac{1}{4}\sum_{n=1}^{3}\frac{(\mathbf{E}_n^T\mathbf{g})^2}{(\Lambda_{nn} + \lambda)^2} = 1, \tag{11}$$

where $\mathbf{E}_n$ and $\Lambda_{nn}$ denote the $n$th eigenvector and eigenvalue of $\mathbf{G}$, respectively. From (11) one can deduce that the problem amounts to rooting a polynomial of degree 6 and thereafter selecting the root of which the corresponding $\mathbf{v}$ maximizes $\mathcal{J}_2^2(p,q)$.

If $(\mathbf{G} + \lambda\mathbf{I}_3)^{-1}$ does not exist,[1] which could occur if $\lambda = 0$ and $\mathbf{G}$ is singular or when $\lambda = -\Lambda_{nn}$ for some $n$, then we have to resort to (10) for the computation of $\mathbf{v}$ [12]:

$$\mathbf{v} = -\tfrac{1}{2}(\mathbf{G} - \Lambda_{nn}\mathbf{I}_3)^{\dagger}\mathbf{g} + c_n\mathbf{E}_n,$$

where $c_n$ is a real constant chosen such that $\|\mathbf{v}\|_F^2 = 1$ and $\mathcal{J}_2^2(p,q)$ is maximum. If it exists, it is given by

$$c_n = \mathrm{sign}(\mathbf{E}_n^T\mathbf{g})\sqrt{1 - \|(\mathbf{G} - \Lambda_{nn}\mathbf{I}_3)^{\dagger}\mathbf{g}\|^2/4}.$$

We emphasize that PAJOD and PAJOD2 compute the same filter. However, the steps taken when $p > q$ are different. An outline of the PAJOD2 algorithm is given in Algorithm 2. By comparing Algorithms 1 and 2 it is clear that the difference between the PAJOD and PAJOD2 algorithms is in the computation of the optimal $\Theta[p,q]$ from $\mathcal{J}_2^2(p,q)|_{q>R}$.

**Algorithm 2.** Outline of the PAJOD2 procedure.

Estimate the cumulant tensor $\mathcal{M}(\mathbf{b},\gamma)$
Initialize $\mathbb{V} = \mathbf{I}_{RL}$
Repeat until convergence
**for** $p = 1$ to $R$ **do**

[1] This case has not been observed in our simulations.

**for** $q=p+1$ to $RL$ **do**
  **if** $q \le R$ **then**
    calculate optimal $\Theta[p,q]$ from $\mathcal{J}_2^2(p,q)|_{q\le R}$ according to [8].
  **else**
    calculate optimal $\Theta[p,q]$ from $\mathcal{J}_2^2(p,q)|_{q>R}$ according to Section 2.2.
  **end if**
  $\mathbf{M}^{(k)} \leftarrow \Theta[p,q]^H \mathbf{M}^{(k)} \Theta[p,q]$
  $\mathbb{V} \leftarrow \Theta[p,q]^H \mathbb{V}$
**end for**
**end for**
Set $\mathbb{H} = \mathbb{V}(:,1:R)$

### 2.3. PAFA

The PAFA algorithm is based on the fact that any FIR paraunitary filter $\mathbf{H}(z) \in \mathbb{C}[z]^{R\times R}$ of length $L$ can be factorized as [5]

$$\mathbf{H}(z) = \mathbf{Q}^{(M)}\mathbf{Z}(z)\mathbf{Q}^{(M-1)}\mathbf{Z}(z)\cdots\mathbf{Z}(z)\mathbf{Q}^{(0)}, \tag{12}$$

where $M$ is the McMillan degree of $\mathbf{H}(z)$, $\mathbf{Q} \in \mathbb{C}^{R\times R}$ is unitary and $\mathbf{Z}(z) = \left[\begin{smallmatrix}\mathbf{I}_{R-1} & \mathbf{0}\\ \mathbf{0} & z^{-1}\end{smallmatrix}\right]$.

The factorization (12) fully takes the paraunitary structure of the problem into account while PAJOD does not. An approach that cyclically estimates one unitary matrix $\mathbf{Q}^{(m)}$ while the other matrices are fixed was proposed in [2].

We now describe how unitary matrix $\mathbf{Q}^{(m)}$ is estimated while the other matrices are fixed. First (12) is reformulated as

$$\mathbf{H}(z) = \mathbf{A}(z)\mathbf{Q}^{(m)}\mathbf{B}(z), \tag{13}$$

where $\mathbf{A}(z) = \mathbf{Q}^{(M)}\mathbf{Z}(z)\cdots\mathbf{Z}(z)\mathbf{Q}^{(m+1)}$ and $\mathbf{B}(z) = \mathbf{Q}^{(m-1)}\mathbf{Z}(z)\cdots\mathbf{Z}(z)\mathbf{Q}^{(0)}$ are of length $l_A \le L$ and $l_B \le L$, respectively, which satisfy $l_A + l_B = L$. The relation between the signal vector after the equalization of $\mathbf{y}(n)$ and the observation vector $\mathbf{x}(n)$ can now be written as

$$\mathbf{y}_r(n) = \sum_{q,s,t}\sum_{u,p}\mathbf{A}_{rq}(u)\mathbf{Q}^{(m)}_{qs}\mathbf{B}_{st}(p)\mathbf{x}_t(n-p-u),$$

where the range of the variables is $u \in \{0,\ldots,l_A\}$, $p \in \{0,\ldots,l_B\}$ and $r,q,s,t \in \{1,\ldots,R\}$. Let the output vector of the filter $\mathbf{B}(z)$ from Eq. (13) be denoted by $\mathbf{w}(n)$, then we also have the relation

$$\mathbf{y}_r(n) = \sum_{qs}\sum_{u}\mathbf{A}_{rq}(u)\mathbf{Q}^{(m)}_{qs}\mathbf{w}_s(n-u).$$

The fourth-order cumulants of $\mathbf{x}(n)$ will be denoted as

$$C_2^{2,\mathbf{x}}[r,\mathbf{k}] = \mathrm{Cum}[x_r(n-k_1), x_r^*(n-k_2), x_r(n-k_3), x_r^*(n-k_4)],$$

$$C_2^{2,\mathbf{x}}[\mathbf{r},\mathbf{k}] = \mathrm{Cum}[x_{r_1}(n-k_1), x_{r_2}^*(n-k_2), x_{r_3}(n-k_3), x_{r_4}^*(n-k_4)],$$

where $\mathbf{r} = [r_1,r_2,r_3,r_4]$ and $\mathbf{k} = [k_1,k_2,k_3,k_4]$. Under the assumption that the signs $\varepsilon = \pm 1$ of the kurtosis of the different signals are the same, it was shown in [7] that the function

$$\mathcal{J}_1^4 = \varepsilon\sum_{r=1}^{R} C_2^{2,\mathbf{y}}[r,\mathbf{0}] \tag{14}$$

is a contrast, where

$$C_2^{2,\mathbf{y}}[r,\mathbf{0}] = \sum_{\mathbf{s}}\sum_{\mathbf{u}}\sum_{\mathbf{q}}\mathbf{A}_{rq_1}(u_1)\mathbf{A}^*_{rq_2}(u_2)\mathbf{A}_{rq_3}(u_3)\mathbf{A}^*_{rq_4}(u_4)$$

$$\mathbf{Q}^{(m)}_{q_1s_1}\mathbf{Q}^{(m)*}_{q_2s_2}\mathbf{Q}^{(m)}_{q_3s_3}\mathbf{Q}^{(m)*}_{q_4s_4}C_2^{2,\mathbf{w}}[\mathbf{s},\mathbf{u}],$$

$$C_2^{2,\mathbf{w}}[\mathbf{s},\mathbf{u}] = \sum_{\mathbf{p}}\sum_{\mathbf{j}}\mathbf{B}_{s_1j_1}(p_1)\mathbf{B}^*_{s_2j_2}(p_2)\mathbf{B}_{s_3j_3}(p_3)\mathbf{B}^*_{s_4j_4}(p_4)C_2^{2,\mathbf{x}}[\mathbf{j},\mathbf{u}+\mathbf{p}]$$

and $\mathbf{s},\mathbf{q},\mathbf{j} \in \{1,\ldots,R\}^4$, $\mathbf{u} \in \{0,\ldots,l_A\}^4$ and $\mathbf{p} \in \{0,\ldots,l_B\}^4$.

For the two-channel case, $R=2$, the computation of the rotation matrix $\mathbf{Q}^{(m)}$ that will maximize (14) amounts to the rooting of a 56th degree polynomial [13] when a resultant-based procedure is employed. This polynomial generally does not admit more than 16 real roots, see [2,13] for further details. Hence, the PAFA algorithm must be considered as computationally more demanding than the PAJOD algorithm. For the case $R>2$ no solution for the PAFA problem was given in [2].

## 3. Simulation results

The algorithms PAJOD, PAJOD2 and PAFA will be tested on random 2-Input-2-Output channels of varying SNR. The transmitted data blocks consist each of a QPSK sequence of 512 symbols length. The paraunitary channel is generated, just as in [1], as follows:

$$\mathbf{F}(z) = \mathbf{R}(\phi_0,\theta_0)\prod_{m=1}^{K-1}\mathbf{Z}(z)\mathbf{R}(\phi_m,\theta_m),$$

where

$$\mathbf{Z}(z) = \begin{bmatrix}1 & 0\\ 0 & z^{-1}\end{bmatrix}, \quad \mathbf{R}(\phi_m,\theta_m) = \begin{bmatrix}\cos(\phi_m) & -\sin(\phi_m)e^{-i\theta_m}\\ \sin(\phi_m)e^{i\theta_m} & \cos(\phi_m)\end{bmatrix}$$

and the parameters $\phi_m$ and $\theta_m$ are drawn from a uniform distribution over $[0,2\pi)$. The filtered QPSK sequences of unit variance are perturbed by an additive white circular complex Gaussian noise with identity covariance matrix.

To see if the PAFA algorithm, which fully exploits the paraunitary structure of the channel, is better than the PAJOD algorithm, which only partially exploits the paraunitary structure of the channel, the performance will be based on the distance between the estimated and the true equalizer and the symbol error rate (SER). Furthermore, the computational complexity of the algorithms will also serve as a basis of comparison. To measure the elapsed time used to execute the algorithms in MATLAB, the built-in functions *tic*$(\cdot)$ and *toc*$(\cdot)$ are used.

Due to the inherent indeterminacies of the equalization problem, the equalizer can only be estimated up to a matrix of the form $\mathbf{D}(z) = \Lambda(z)\mathbf{P}$, where $\mathbf{P}$ is a permutation matrix and $\Lambda(z)$ is diagonal matrix with entries of the form $\lambda z^k$, $k \in \mathbb{Z}$. Let $\mathbf{G}$ be the $R \times \mathrm{R}(2L-1)$ matrix obtained by stacking the global impulse response matrices after each other, then a measure of how far we are from perfect equalization is given by

$$\mathrm{dist}(\mathbf{G}) = \min_{\mathbf{D}(z)}\|\mathbf{G}\mathbf{D}(z)\|_F.$$

A procedure to solve this problem was given in [14] and will not be repeated here.

### 3.1. Paraunitary channel with $K=L$

In the first simulation, the paraunitary channel and equalizer are of length $K=L=3$. The mean SER, time and dist values over 100 simulations can be seen in Fig. 2. From Fig. 2(b) it is observed that below 15 dB the PAFA algorithm seems to capture the channel structure better than the PAJOD algorithm. It is also clear that the PAJOD2 algorithm is cheaper than the PAJOD algorithm which in turn is cheaper than the PAFA method. We also observe a variation of the computation time with the noise level. The reason is that at low SNR the problem of extracting the paraunitary channel from the cumulants of the observed data is more costly since the paraunitary structure is more "hidden" in the noisy data.

### 3.2. Paraunitary channel with $K<L$

A second simulation was conducted in order to investigate the robustness of the paraunitary method against channel order mismatch. We consider the case when $K=3$ and $L=4$, i.e., the channel is overmodeled. The mean SER, time and dist values over 100 simulations can be seen in Fig. 3. Comparison with Fig. 2 shows that the algorithms can handle the mismatch quite well. The relative performance of these algorithms remains the same.

### 3.3. Paraunitary channel with varying length K

A third simulation was conducted in order to investigate the computation time of the algorithms as a function of the filter length. The paraunitary channel and equalizer are of the same length $K=L$. The SNR was fixed to 10 while the filter and channel length varied from 4 to 10 with a hop factor of 1.

The mean SER and computation time over 100 simulations can be seen in Fig. 4. Here it is clear that the computational complexity of the PAFA algorithm grows exponentially as a function of the filter length. We also note that the gap in computation time between the PAJOD and PAJOD2 methods decreases with increasing $K$. The reason is that when $K$ increases, then so does the time needed to compute the coefficients of the involved polynomials, while the actual cost of the rooting of the polynomials remains the same. However, the PAJOD method has to compute 24 coefficients while the PAJOD2 method only has to compute six coefficients. Consequently, a gap in computation time between the methods remains apparent. Secondly, the computation of the involved cumulant tensors also increases as illustrated in Fig. 5. In the case of slowly changing channels an adaptive estimation of the cumulant tensor is possible. Hence in that case a reduction of the computation time of the cumulants is possible.

## 4. Conclusion

In this paper the problem of blind equalization of paraunitary channels by means of Jacobi-type procedures was studied. After a review of the PAJOD and PAFA methods we proposed a computationally more efficient method called PAJOD2. The proposed method simplified

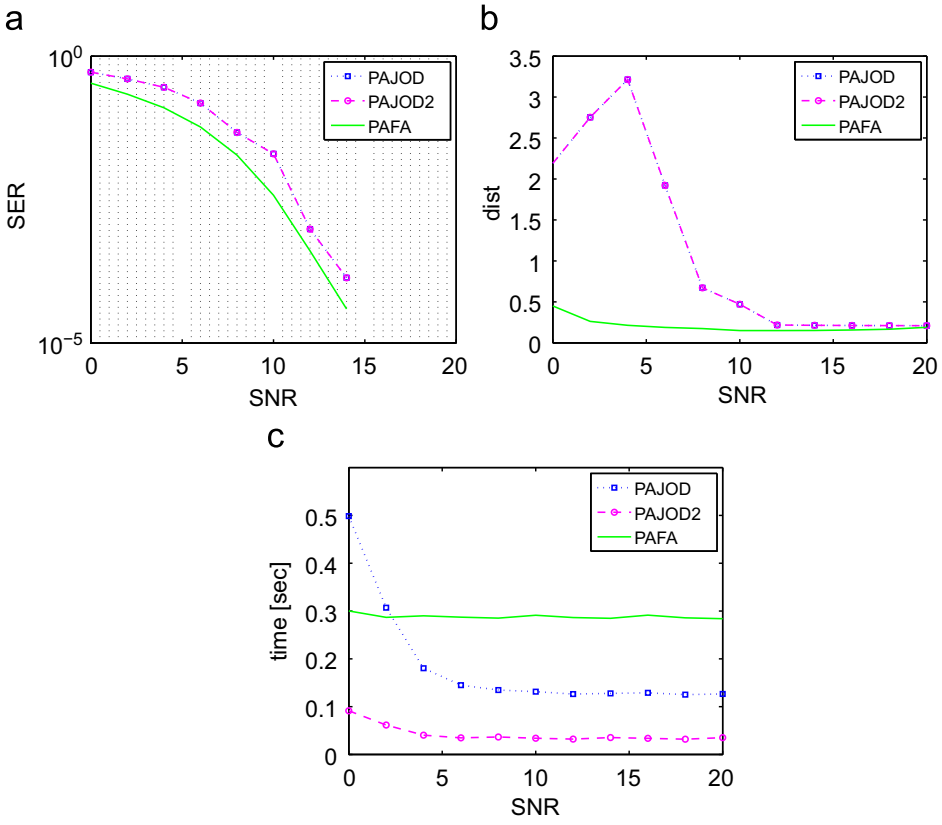



**Fig. 2.** Mean SER, dist and time values for the case when $K=L=3$. (a) Mean SER. (b) Mean dist. (c) Mean time.

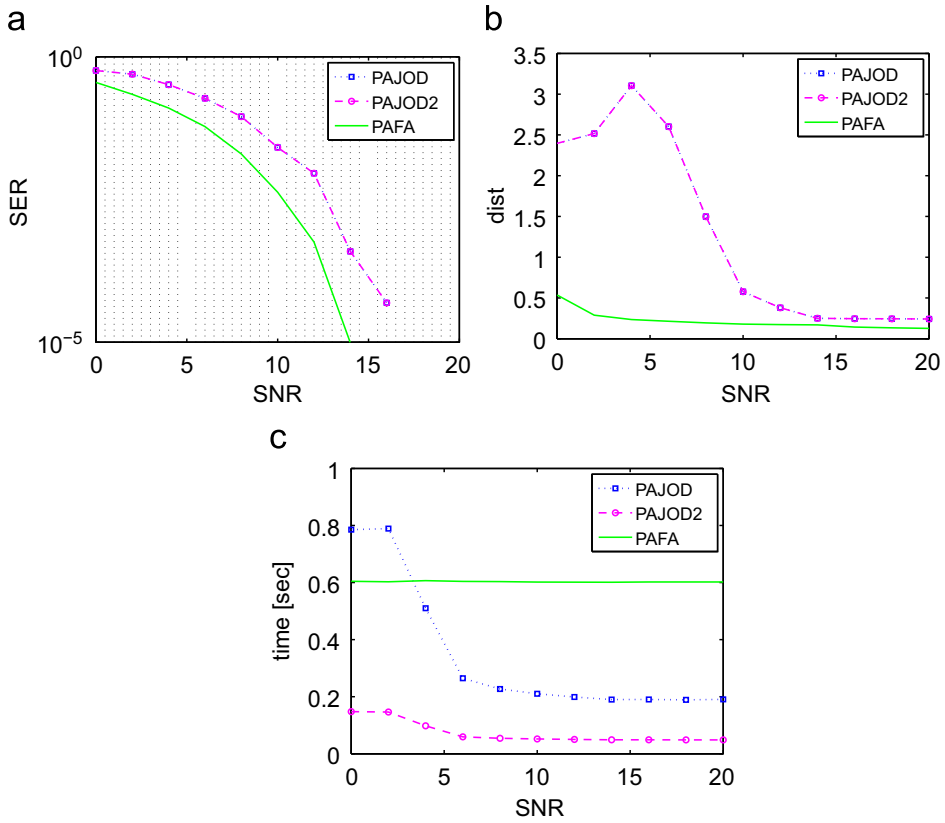



**Fig. 3.** Mean SER, dist and time values for the case when $K=3$ and $L=4$. (a) Mean SER. (b) Mean dist. (c) Mean time.

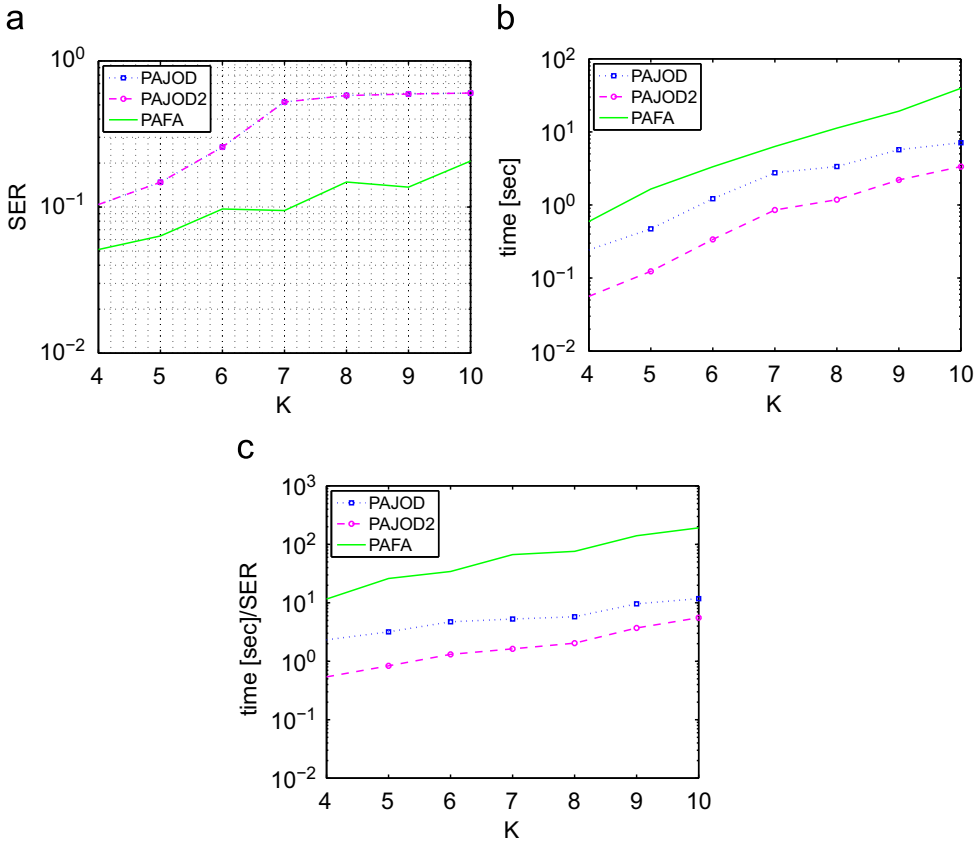



**Fig. 4.** Mean SER and computation time for filters and channels of varying length $K=L$. (a) Mean SER. (b) Mean dist. (c) Mean time.

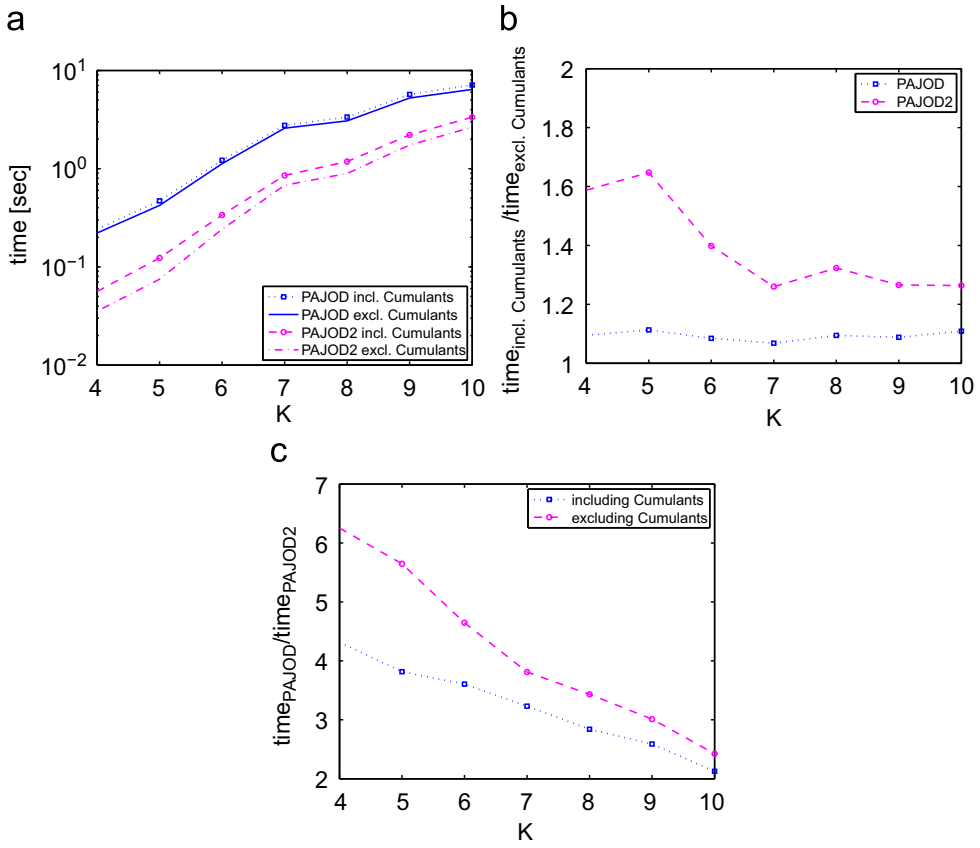



**Fig. 5.** Mean computation times for filters and channels of varying length $K=L$. (a) Mean time. (b) Mean $\text{time}_{\text{incl. Cumulants}}/\text{time}_{\text{excl. Cumulants}}$. (c) Mean $\text{time}_{\text{PAJOD}}/\text{time}_{\text{PAJOD2}}$.

the Jacobi-subproblem of the PAJOD method by going from the rooting of a 24th degree polynomial to the rooting of a polynomial of degree 6.

Simulation results indicate that for low SNR the PAFA method is more accurate than the PAJOD method. However, for high SNR both methods perform similarly. Such a comparison has not been conducted in the literature. Furthermore, computer simulations confirmed that the PAJOD2 method is faster than the PAJOD method.

## Acknowledgments

Mikael Sørensen and Lieven De Lathauwer would like to thank L. Rota for kindly providing the PAJOD and PAFA Matlab code. Research supported by: (1) Research Council K.U.Leuven: GOA-MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC), CIF1, STRT1/08/23, (2) F.W.O.: (a) project G.0427.10N, (b) Research Communities ICCoS, ANMMM and MLDM, (3) the Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007–2011), (4) EU: ERNSI.

## References

[1] P. Comon, L. Rota, Blind separation of independent sources from convolutive mixtures, IEICE Trans. Fund. (2003) 542–549.
[2] S. Icart, P. Comon, L. Rota, Blind paraunitary equalization, Signal Process. 89 (2009) 283–290.
[3] J.G. McWhirter, P.D. Baxter, T. Cooper, J. Foster, An EVD algorithm for para-Hermitian polynomial matrices, IEEE Trans. Signal Process. 5 (2007) 2158–2169.
[4] M. Sørensen, L. De Lathauwer, L. Deneire, An efficient Jacobi-type algorithm for blind equalization of paraunitary channels, in: Proceedings of the EUSIPCO 2009, pp. 510–514.
[5] P. Vaidyanathan, Multirate Systems and Filter Banks, Signal Processing Series, Prentice-Hall, New Jersey, 1993.
[6] D. Donoho, On minimum entropy deconvolution, Appl. Time Ser. Anal. II (1981) 565–608.
[7] P. Comon, Contrasts for multichannel blind deconvolution, IEEE Signal Process. Lett. 3 (1996) 209–211.
[8] J.-F. Cardoso, A. Souloumiac, Blind beamforming for non-Gaussian signals, IEE Proc. F 140 (1993) 362–370.
[9] J.-F. Cardoso, A. Souloumiac, Jacobi angles for simultaneous diagonalization, SIAM J. Math. Anal. Appl. 17 (1996) 161–164.
[10] G. Golub, C. Van Loan, Matrix Computations, 3rd ed. in 1996, The John Hopkins University Press, Baltimore, 1983.
[11] D. Cox, J. Little, D. O'Shea, Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 2nd ed. in 1996, Springer Verlag, 1992.
[12] A. Belouchrani, K. Abed-Meraim, Y. Hua, Jacobi-like algorithms for joint block diagonalization: applications to source localization, in: Proceedings of the IEEE International Workshop on Intelligent Signal Processing and Communications Systems (ISPACS), 1998.
[13] L. Rota, Egalisation aveugle de systemes multi-utilisateurs, Ph.D. Thesis, Laboratoire I3S, France, 2004.
[14] P. Comon, Blind Separation of Independent Sources from Convolutive Mixtures, A Partial Joint Diagonalization (PAJOD) Approach, Technical Report RR-2002-04-FR, I3S, 2002.