UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Novel results on the factorization and estimation of spectral densities

Ph.D. Candidate Giacomo Baggio

Ph.D. Advisor Prof. Augusto Ferrante

Coordinator Prof. Andrea Neviani

Ph.D. School in Information Engineering • Science & Information Technology • XXX cycle

this page is intentionally left blank

Version 1.0 Padova, January 14, 2018

Minor updates to the present manuscript may occasionally occur.
Please refer to the latest version available at: http://www.dei.unipd.it/~baggio.

this page is intentionally left blank

To $\mathcal{N}$.

this page is intentionally left blank

Abstract

THIS dissertation is divided into two main parts. The first part is concerned with one of the most classical and central problems in Systems and Control Theory, namely the factorization of rational matrix-valued spectral densities, commonly known as the *spectral factorization problem*. Spectral factorization is a fundamental tool for the solution of a variety of problems involving second-order statistics and quadratic cost functions in control, estimation, signal processing and communications. It can be thought of as the frequency-domain counterpart of the ubiquitous Algebraic Riccati Equation and it is intimately connected with the celebrated Kálmán–Yakubovich–Popov Lemma and, therefore, to passivity theory. Here, we provide a rather in-depth and comprehensive analysis of this problem in the discrete-time setting, a scenario which is becoming increasingly pervasive in control applications. The starting point in our analysis is a general spectral factorization result in the same vein of Dante C. Youla. Building on this fundamental result, we then investigate some key issues related to minimality and parametrization of minimal spectral factors of a given spectral density. To conclude, we show how to extend some of the ideas and results to the more general indefinite or J-spectral factorization problem, a technique of paramount importance in robust control and estimation theory.

In the second part of the dissertation, we consider the problem of estimating a spectral density from a finite set of measurements. Following the Byrnes–Georgiou–Lindquist THREE (Tunable High REsolution Estimation) paradigm, we look at spectral estimation as an optimization problem subjected to a generalized moment constraint. In this framework, we examine the global convergence of an efficient algorithm for the estimation of scalar spectral densities that hinges on the Kullback–Leibler criterion. We then move to the multivariate setting by addressing the delicate issue of existence of solutions to a parametric spectral estimation problem. Eventually, we study the geometry of the space of spectral densities by revisiting two natural distances defined in cones for the case of rational spectra. These new distances are used to formulate a "robust" version of THREE-like spectral estimation.

Sommario

La tesi è divisa in due parti. La prima parte riguarda uno dei problemi più importanti e classici della Teoria dei Sistemi e del Controllo, ossia la fattorizzazione di densità spettrali razionali a valori matriciali, meglio conosciuto come *problema di fattorizzazione spettrale*. Quest'ultimo rappresenta uno strumento fondamentale per la soluzione di una vasta gamma di problemi riguardanti statistiche del secondo ordine e funzioni costo quadratiche nella teoria del controllo, della stima, dell'elaborazione di segnali e delle comunicazioni. Il problema di fattorizzazione spettrale può essere visto come la controparte nel dominio della frequenza della soluzione di un'Equazione Algebrica di Riccati ed è strettamente connesso con il famoso Lemma di Kálmán–Yakubovich–Popov e, di conseguenza, con la teoria dei sistemi passivi. Questa prima parte fornisce un'analisi approfondita e completa del problema di fattorizzazione spettrale nel caso a tempo discreto, uno scenario sempre più diffuso nelle applicazioni del controllo. Il punto di partenza della nostra analisi è un risultato generale sulla fattorizzazione spettrale che si ispira ad un approccio ideato da Dante C. Youla. Basandoci su questo risultato, esaminiamo quindi alcuni aspetti chiave legati alla minimalità e alla parametrizzazione dei fattori spettrali minimi di una data densità spettrale. Per concludere, mostriamo come estendere alcuni idee e risultati al caso più generale di fattorizzazione spettrale indefinita o fattorizzazione J-spettrale, una tecnica di importanza primaria nella teoria del controllo e della stima robusta.

Nella seconda parte della tesi, consideriamo il problema della stima di una densità spettrale incognita a partire da un insieme finito di misure. Seguendo l'approccio THREE (Tunable High REsolution Estimation) di Byrnes, Georgiou, e Lindquist, interpretiamo il problema di stima spettrale come un problema di ottimizzazione soggetto ad un vincolo sui momenti generalizzato. In questo contesto, studiamo la convergenza globale di un algoritmo efficiente per la stima di densità spettrali scalari basata sul criterio di Kullback–Leibler. Successivamente, ci spostiamo ad analizzare il caso multivariato, considerando un problema estremamente delicato riguardante l'esistenza di una soluzione ad un problema di stima parametrico. Infine, analizziamo la geometria dello spazio delle densità spettrali rivisitando due distanze naturali definite su coni per il caso di spettri razionali. Queste nuove distanze verranno utilizzate per formulare un problema di stima spettrale "robusta" simile all'approccio THREE.

Contents

Ouverture

THE present dissertation focuses on two tightly connected problems which lie at the heart of the engineering and mathematical sciences. Both problems share a long and illustrious history and have been extensively studied and applied in many areas of Systems and Control Theory.

The first one is the classical problem of *spectral factorization*. This problem undoubtedly represents one of the cornerstones of Systems and Control Theory. Starting from the seminal works Kolmogorov [1939], Wiener [1949], it is almost impossible to acknowledge all the works and all the branches of Systems and Control Theory in which this classical problem has been studied and/or successfully applied. For the time being, we will be content with mentioning that this problem has proved to be fundamental in stochastic realization theory, Linear-Quadratic (LQ) optimal control, and $\mathscr{H}_\infty$ robust control. In spite of being a classical and largely investigated problem, quite surprisingly, there are still a number of aspects of this problem which have not satisfactorily been addressed in the literature. Their dissection will constitute Part I of the present dissertation. This part starts off with a general overview on the problem in Chapter 1. Then, inspired by the celebrated work Youla [1961], in Chapter 2-3-4, we discuss a general result on the solution of the multivariate spectral factorization problem in discrete-time and we elucidate some issues related to uniqueness and parametrization of minimal spectral factors. Chapter 5 is devoted to present some results concerning the more general J-spectral factorization problem. Ultimately, in Chapter 6, we draw the conclusions and list some open problems.

Part I is mostly based on the following works:

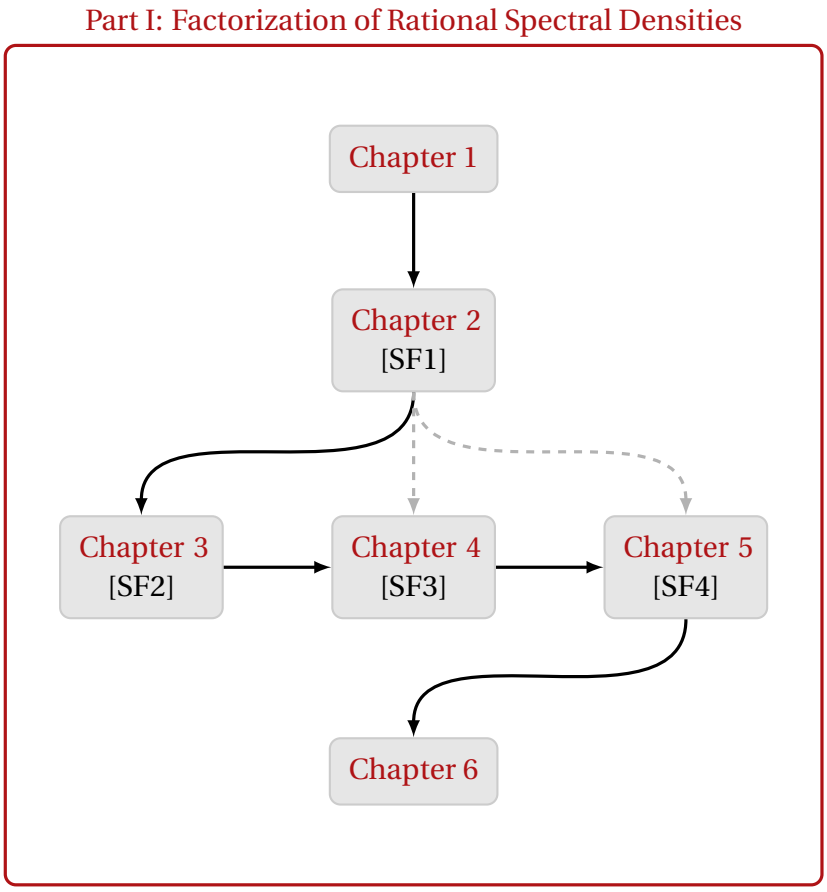
[SF1] G. Baggio and A. Ferrante. On the factorization of rational discrete-time spectral densities. *IEEE Transactions on Automatic Control*, 61(4):969–981, 2016a

[SF2] G. Baggio and A. Ferrante. On minimal spectral factors with zeroes and poles lying on prescribed regions. *IEEE Transactions on Automatic Control*, 61(8): 2251–2255, 2016b

[SF3] G. Baggio and A. Ferrante. Parametrization of minimal spectral factors of discrete-time rational spectral densities. *Submitted for publication*, 2017. [arXiv preprint: `arXiv:1609.02711`]

[SF4] G. Baggio. On minimal discrete-time J-spectral factorizations. *In preparation*, 2017c

The following diagram depicts the dependencies between different chapters of Part I together with the related works: A filled black arrow indicate the suggested "natural" flow, whereas a dashed gray arrow a possible alternative flow.



The second problem is, *de facto*, a class of problems which falls under the name of *generalized moment problems*. As for spectral factorization, (generalized) moments problems have a deep and rich history rooted in the early works of the great Russian mathematician Čebyšëv and his students, who exploited them in connection with the classical Central Limit Theorem towards the end of the 19$^{\text{th}}$ Century, *cf.* [Kreĭn and Nudel'man, 1977, pp. 166–171]. However, it is only in recent times that these problems have been profitably employed in several areas of Systems and Control Theory, such as interpolation theory, robust control, and approximation of covariances and spectral densities. The merit for this must be attributed to the Byrnes–Georgiou–Lindquist school who first brought to light the central role played by these problems in Control Theory, *cf.* Byrnes and Lindquist [2008] and references therein. Among the plethora of applications of generalized moment problems, in Part II of the dissertation we will focus on the subclass of problems dealing with *estimation of spectral densities*. After a brief introduction on generalized moment problems contained in Chapter 1, we review in Chapter 2 the particular problem of scalar Kullback–Leibler approximation of spectral densities treated in Georgiou and Lindquist [2003]. Chapter 3 studies the convergence properties of an efficient algorithm for the solution of the latter estimation problem, proposed in Pavon and Ferrante [2006]. In Chapter 4, we focus on a certain multivariate parametric extension of the above spectral estimation problem, establishing a general existence result. In Chapter 5, motivated by the issue of reducing "artifacts" in spectral estimation procedures, we investigate a new class of metrics of Finsler type in the space of spectral densities. Eventually, Chapter 6 contains some concluding remarks and future research directions.

Part II is mostly based on the following works:

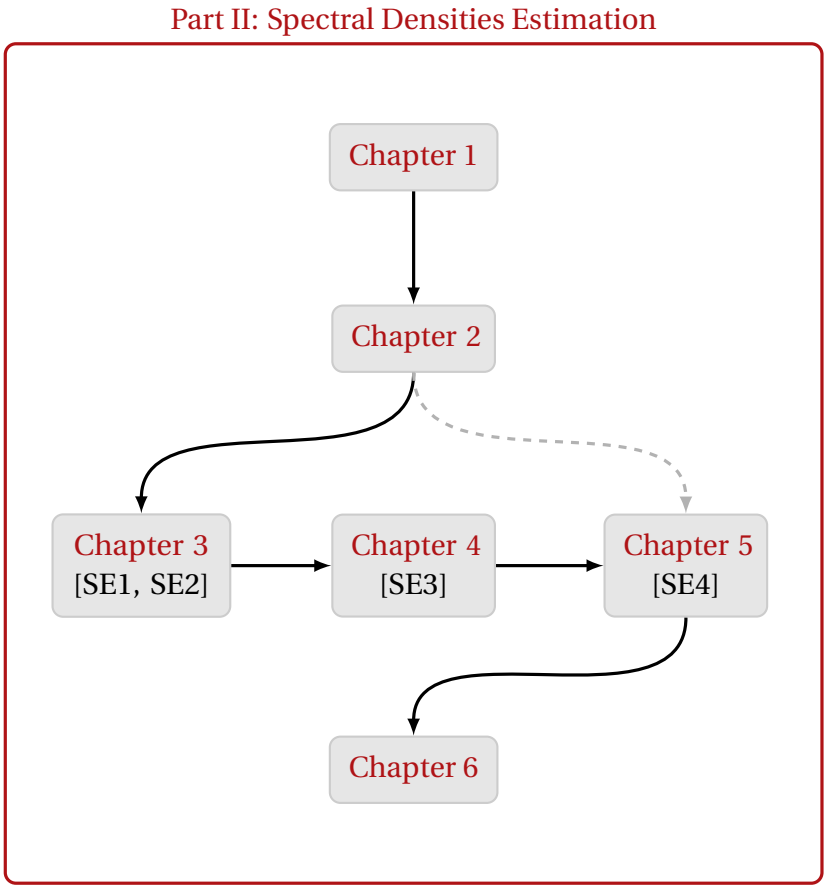
[SE1] G. Baggio. A global convergence analysis of the Pavon–Ferrante algorithm for spectral estimation. *IEEE Transactions on Automatic Control (Conditionally Accepted)*, 2017a. [arXiv preprint: arXiv:1612.03570]

[SE2] G. Baggio. On the convergence of a matricial fixed-point iteration connected with spectral estimation. *Proceedings of the 20th IFAC World Congress*, pages 7415–7420, 2017b

[SE3] B. Zhu and G. Baggio. On the existence of a solution to a spectral estimation problem *à la* Byrnes–Georgiou–Lindquist. *Submitted for publication*, 2017. [arXiv preprint: arXiv:1709.09012]

[SE4] G. Baggio, A. Ferrante, and R. Sepulchre. Finslerian metrics in the cone of spectral densities. *Submitted for publication*, 2017. [arXiv preprint: arXiv:1708.02818]

The following diagram depicts the dependencies between chapters of Part II. The meaning of the arrows is the same as before.



List of symbols and conventions

Symbol	Description	Page
$\emptyset$	Empty set	
$\oplus$	Direct sum of vector spaces	
dim	Dimension of a vector space	
$\mathbb{N}$	Set of natural numbers $\{0,1,2,3,\ldots\}$	
$\mathbb{Z}$	Ring of integers $\{\ldots,-2,-1,0,1,2,\ldots\}$	
$\mathbb{R}$	Field of real numbers	
$\mathbb{R}_0$	Non-zero real numbers $\mathbb{R}\setminus\{0\}$	
$\mathbb{R}_{\geq 0}$	Non-negative real numbers	
$\mathbb{R}_{>0}$	(Strictly) positive real numbers	
$\mathbb{C}$	Field of complex numbers	
$\mathbb{C}_0$	Non-zero complex numbers $\mathbb{C}\setminus\{0\}$	
$\overline{\mathbb{C}}$	Extended complex plane $\mathbb{C}\cup\{\infty\}$	
$\mathbb{T}$	Unit circle in the complex plane	
$\overline{a}$	Complex conjugate of $a\in\mathbb{C}$	
$\mathrm{Im}(a)$	Imaginary part of $a\in\mathbb{C}$	
$\mathrm{Re}(a)$	Real part of $a\in\mathbb{C}$	
$\lvert a\rvert$	Modulus of $a\in\mathbb{C}$	
$\mathbb{F}^{n\times m}$	Set of $n\times m$ matrices with entries in the field $\mathbb{F}$	
$\mathrm{GL}(n,\mathbb{F})$	Group of $n\times n$ invertible matrices with values in the field $\mathbb{F}$	
$\mathbb{S}_n$	Set of symmetric $n\times n$ matrices	
$\mathbb{S}_{+,n}$	Set of symmetric positive definite $n\times n$ matrices	
$\mathbb{H}_n$	Set of Hermitian $n\times n$ matrices	
$\mathbb{H}_{+,n}$	Set of Hermitian positive definite $n\times n$ matrices	
$\mathfrak{S}_n$	Set of symmetric positive semi-definite unit-trace $n\times n$ matrices	
I_n	$n\times n$ identity matrix (subscript is omitted if the dimension n is clear by the context)	
J	Signature matrix, *i.e.*, diagonal matrix with entries ± 1	
$\mathbf{0}_{n,m}$	$n\times m$ zero matrix (subscript is omitted if the dimensions n and m are clear by the context)	

Symbol	Description	Page
$\mathbf{0}_n$	$n \times n$ zero matrix (subscript is omitted if the dimension n is clear by the context)	
$[A]_{ij}$	Entry at (i, j) of matrix A	
$[A]_{i:j,k:h}$	Sub-matrix of A obtained by extracting the rows from index i to index j ($i \leq j$) of A and the columns from index k to index h ($k \leq h$) of A	
$\overline{A}$	Entry-wise complex conjugate of matrix A	
$A^\top$	Transpose of matrix A	
$A^{-\top}$	Transpose of matrix inverse, $A^{-\top} = (A^{-1})^\top$	
A^*	Hermitian transpose of matrix A, $A^* = \overline{A}^\top$	
A^{-1}	Inverse of matrix A	
A^+	Moore–Penrose pseudoinverse of A	
A^{-L}	Left inverse of matrix A	
A^{-R}	Right inverse of matrix A	
det	Determinant of a matrix	
rank	Rank of a matrix	
ker	Kernel or null space of a matrix	
im	Image or range of a matrix	
diag$[a_1, \ldots, a_n]$	Diagonal matrix with entries $a_1, \ldots, a_n$	
tr	Trace of a matrix	
$\|\cdot\|_{\mathrm{F}}$	Frobenius norm of a matrix	
span$\{x_1, \ldots, x_n\}$	Subspace spanned by vectors $x_1, \ldots, x_n$	
$\mathbb{F}[z]$	Ring of polynomials over the field $\mathbb{F}$	
deg	Degree of polynomial	
$\mathbb{F}[z, z^{-1}]$	Ring of Laurent polynomials (L-polynomials) over the field $\mathbb{F}$	
max deg	Maximum-degree of L-polynomial	
min deg	Minimum-degree of L-polynomial	
$\mathbb{F}(z)$	Set of rational functions over the field $\mathbb{F}$	
$v_\alpha(f)$	Valuation at $\alpha \in \mathbb{C}$ of $f(z) \in \mathbb{F}(z)$	36
$\mathbb{F}[z]^{n\times m}$	Set of polynomial $n \times m$ matrices over the field $\mathbb{F}$	
$\mathbb{F}[z, z^{-1}]^{n\times m}$	Set of Laurent polynomial (L-polynomial) $n \times m$ matrices over the field $\mathbb{F}$	
$\mathbb{F}(z)^{n\times m}$	Set of rational $n \times m$ matrices over the field $\mathbb{F}$	
$\mathbb{F}_*(z)^{n\times n}$	Set of rational $n \times n$ matrices over the field $\mathbb{F}$ of full rank on $\mathbb{T}$	
rk	Normal rank of a rational matrix, *i.e.*, rank almost everywhere in $\mathbb{C}$	

Symbol	Description	Page
$G^*(z)$	$G^*(z) := G^\top(1/z)$ if $G(z) \in \mathbb{R}(z)^{n\times m}$, $G^*(z) := \overline{G}^\top(1/\bar{z})$ if $G(z) \in \mathbb{C}(z)^{n\times m}$	
$G^{-*}(z)$	$G^{-*}(z) := [G^\top(1/z)]^{-1}$ if $G(z) \in \mathbb{R}(z)^{n\times m}$, $G^{-*}(z) := [\overline{G}^\top(1/\bar{z})]^{-1}$ if $G(z) \in \mathbb{C}(z)^{n\times m}$	
$G^{-L}(z)$	$G^{-L}(z) := [G^\top(1/z)]^{-L}$, $G(z) \in \mathbb{R}(z)^{n\times m}$	
$G^{-R}(z)$	$G^{-R}(z) := [G^\top(1/z)]^{-R}$, $G(z) \in \mathbb{R}(z)^{n\times m}$	
G^{hc}	Highest-column-degree coefficient matrix of $G \in \mathbb{R}[z,z^{-1}]^{m\times n}$	37
G^{lc}	Lowest-column-degree coefficient matrix of $G \in \mathbb{R}[z,z^{-1}]^{m\times n}$	37
G^{hr}	Highest-row-degree coefficient matrix of $G \in \mathbb{R}[z,z^{-1}]^{m\times n}$	37
G^{lr}	Lowest-row-degree coefficient matrix of $G \in \mathbb{R}[z,z^{-1}]^{m\times n}$	37
$\delta_p(G;p_i)$	Degree of pole $p_i \in \mathbb{C}$ of $G(z) \in \mathbb{R}(z)^{m\times n}$	39
$\delta_z(G;\alpha)$	Degree of zero $\alpha \in \mathbb{C}$ of $G(z) \in \mathbb{R}(z)^{m\times n}$	60
$\delta_M(G)$	McMillan degree of $G(z) \in \mathbb{R}(z)^{m\times n}$	39
$\|\cdot\|_{\mathcal{H}_2}$	$\mathcal{H}_2$-norm of a discrete-time transfer function	171
$\|\cdot\|_{\mathcal{H}_\infty}$	$\mathcal{H}_\infty$-norm of a discrete-time transfer function	142
$\mathscr{C}^k(X)$	Class of k-times differentiable functions with continuous derivatives on X (the domain X will be omitted if clear by the context)	
$\mathscr{C}(\mathbb{T},\mathbb{H}_m)$	Space of $m\times m$ Hermitian matrix-valued continuous function on the unit circle	121
$\deg_T$	Topological degree of a function	150
$\mathscr{S}^n_{\mathrm{b,rat}}(\mathbb{T})$	Space of $n\times n$ bounded rational matrix-valued discrete-time spectral densities or spectra	165
$\mathscr{S}^n_{\mathrm{rat}}(\mathbb{T})$	Space of $n\times n$ rational matrix-valued discrete-time spectral densities or spectra	25
$\mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$	Space of $n\times n$ rational matrix discrete-time J-spectral densities or J-spectra	29
$\mathscr{S}^n(\mathbb{T})$	Space of $n\times n$ matrix-valued discrete-time spectral densities or spectra (thought of as bounded positive operators on an Hilbert space)	165
$\mathscr{S}^m_+(\mathbb{T})$	Space of $m\times m$ matrix-valued discrete-time coercive and bounded spectral densities	117
$\mathbb{E}\{\cdot\}$	Expected value of a random variable/vector	

Throughout the dissertation, with a slight abuse of notation, when we say that a rational function $f(z)$ is analytic in a region $\mathbb{T}$ of the complex plane that is *not* open, we mean that $f(z)$ does not have poles in $\mathbb{T}$. In the case of a rational $f(z)$ this abuse does not cause any problems: In fact, $f(z)$ can have only finitely many poles so that there exists a larger open region $\mathbb{T}_\varepsilon \supset \mathbb{T}$ in which $f(z)$ is indeed analytic. For example, if $f(z)$ is rational and does not have poles on the unit circle, we say that $f(z)$ is analytic on the unit circle in place of $f(z)$ is analytic on an open annulus containing the unit circle. Notice that such an annulus does indeed exist. Finally, in what follows, the abbreviation a.e. will stand for almost everywhere. For instance, a property that holds "a.e. in $\mathbb{C}$" or "for $z \in \mathbb{C}$ a.e." means that this property holds for all but a finite number of points in $\mathbb{C}$.

Part I

Factorization of rational spectral densities

1. Introduction

THE purpose of this chapter is to introduce and motivate the topics we will deal with in the first part of the dissertation, namely the "standard" spectral factorization problem and its "indefinite" extension, commonly known under the name of J-spectral factorization. We will briefly outline the historical developments, the areas of Systems and Control Theory wherein these problems naturally emerge, and the contributions provided by the present dissertation to these problems.

1.1 Spectral factorization

WE denote by $\mathscr{S}^n_{\text{rat}}(\mathbb{T})$ the set of real-valued discrete-time $n \times n$ spectral density functions, *i.e.*, the set of rational matrix-valued functions $\Phi\colon \mathbb{T} \to \mathbb{R}^{n\times n}$ such that $\Phi(z) \geq 0$ for all $z \in \mathbb{T}$ for which Φ is defined (if $n = 1$ we will use the shorthand $\mathscr{S}_{\text{rat}}(\mathbb{T}) := \mathscr{S}^1_{\text{rat}}(\mathbb{T})$). In its discrete-time rational form, the spectral factorization problem can be stated as follows.

Problem 1.1 ((Discrete-time) rational spectral factorization)**.** Given a rational spectral density $\Phi \in \mathscr{S}^n_{\text{rat}}(\mathbb{T})$ find a factorization of the form

$$\Phi(z) = W^*(z)W(z), \quad z \in \overline{\mathbb{C}} \text{ a.e.}, \tag{1.1}$$

where $W(z) \in \mathbb{R}(z)^{\ell\times n}$ and $W^*(z) := W^\top(1/z)$.

The rational matrix-valued function W in (1.1) is called a *spectral factor* of Φ. Depending on the considered application, one is usually interested in spectral factors featuring some additional properties. For instance, in stochastic realization theory, the sought-for spectral factor is often required to be *(stochastically) minimal*, *i.e.*, having the smallest possible McMillan degree[1]. In the solution of optimal filtering problems, instead, the interest is for spectral factors that are analytic in a prescribed region of the complex region with (generalized) inverse analytic in another (possibly different) region. In particular, if the latter two regions coincide with $\{z \in \overline{\mathbb{C}} : |z| > 1\}$, then the spectral factor is called *minimum-phase* or *outer* or *canonical* spectral factor. This spectral factor yields the solution of the causal optimal filtering problem. When dealing instead with *acausal* optimal filtering problems, the sought-for spectral factor has prescribed analyticity regions that differ from those of the canonical one, see, *e.g.*, Lindquist and Picci [1991], Picci and Pinzoni [1994], Ferrante and Picci [2014].

[1] Such degree is equal to one half of the McMillan degree of the spectral density. The McMillan degree quantifies the "complexity" of a transfer function and will be defined in Section 2.2.

As briefly mentioned in the initial chapter of this dissertation, the spectral factorization problem is a classical and largely investigated problem in Systems and Control Theory. More precisely, spectral factorization is the common denominator of a circle of ideas including LQ optimization methods, passivity theory, positivity, second-order stationary stochastic processes and Riccati equations, and, as such, it seems fair to say that this problem is one of the pillars of modern Systems and Control Theory.

The origins of this mathematical tool dates back to the 1940s, when Kolmogorov and Wiener, independently of each other, introduce and study the scalar version of the spectral factorization problem in order to obtain a frequency-domain solution to optimal filtering problems for the discrete-time case[2] Kolmogorov [1939] and continuous-time case (the well-known *Wiener-Hopf technique*) Wiener [1949]. Besides optimal filtering, estimation and smoothing (see, for instance, Anderson and Moore [1979] for a comprehensive overview), the spectral factorization problem has found many interesting applications, for instance, in:

[2] In the scalar discrete-time polynomial case spectral factorization is also known as the Fejér–Riesz Lemma Ephremidze et al. [2009].

1) LQ optimal control theory Willems [1971], Callier and Winkin [1992], Stoorvogel and Saberi [1998], Swigart and Lall [2014], Aksikas et al. [2007], Gu et al. [2006], Ferrante and Ntogramatzidis [2013],

2) Stochastic realization Anderson [1969], Faurre et al. [1979], Ruckebusch [1980], Picci [1976], Lindquist and Picci [1979, 1985, 1991], Ferrante [1994], Lindquist et al. [1995], Lindquist and Picci [2015],

3) Circuit synthesis theory Newcomb [1966], Anderson [1999], Anderson and Vongpanitlerd [2006], Youla [2015],

4) Interpolation theory Byrnes et al. [2001b], Blomqvist et al. [2003], Byrnes et al. [2006],

5) Dynamical networks and graphical models identification Avventi et al. [2013], Materassi and Innocenti [2010], Materassi and Salapaka [2012], Hayden et al. [2017], Weerts et al. [2017],

6) Spectral estimation Ferrante et al. [2008], Ramponi et al. [2009], Avventi [2011], Ferrante et al. [2012a] (this point will be discussed in greater detail in Part II),

7) Generalized factor analysis Deistler et al. [2010], Anderson and Deistler [2009], Anderson et al. [2012], and

8) Passivity, from the classical positive-real systems theory Khalil [1996], Brogliato et al. [2007] to the more recent negative-imaginary systems theory Petersen and Lanzon [2010], Ferrante et al. [2017].

With reference to the continuous-time case, the existence of an outer solution to the multivariate spectral factorization problem was firstly proved by Dante C. Youla in his famous paper Youla [1961]. In that paper, Youla presents an ingenious technique, which exploits the *Smith–McMillan canonical form* of rational matrices, to construct (and hence prove the existence of) the outer spectral factor of a given spectral density.

Far from being merely of historical interest, Youla's method is relevant for, at least, three reasons:

1) It does not require *any* assumption on the starting spectral density, allowing, in particular, for spectral densities that are improper, normal rank deficient, and with zeros and poles on the imaginary axis.

2) It always leads to the computation of a (stochastically) minimal spectral factor, that is a spectral factor of minimal "complexity".

3) It allows to easily modify the region of analyticity of both the obtained spectral factor and its (right) inverse. This feature has proved to be fundamental, for instance, in Pandolfi [2001], Ferrante and Pandolfi [2002], where Youla's method is exploited to compute a spectral factor with prescribed analyticity region in order to weaken the standard assumptions for the solvability of the classical Positive Real lemma (or Kálmán–Yakubovich–Popov (KYP) lemma) equations.

After Youla's work, other more efficient numerical methods have been proposed for the computation of the outer spectral factor both in the continuous and discrete time case. Most of them are based on Newton-like optimization routines, or on state-space methods which are, in turn, linked to the solution of suitable Algebraic Riccati Equations (AREs) (and, hence, to Kálmán filtering theory). It is worth stressing that these techniques typically require a number of restrictive assumptions on the given spectral density. In particular, almost all the proposed methods, with a few notable exceptions, do not apply to spectral densities that are normal rank deficient, a situation of utmost practical relevance in high-dimensional scenarios, see, *e.g.*, Anderson et al. [2012]. Among the plethora of approaches it is worth mentioning the works Tunnicliffe-Wilson [1972], Rissanen [1973], Youla and Kazanjian [1978], Ran [1982], Van Der Schaft and Willems [1984], Ježek and Kučera [1985], Callier [1985], Georgiou [1988], Clements and Glover [1989], Chen and Francis [1989], Ran and Rodman [1990], Clements [1993], Wimmer [1997], Clements et al. [1997], Bini et al. [2003], Orchard and Wilson [2003], Janashia et al. [2011], Boettcher and Halwass [2013] (see also the survey Sayed and Kailath [2001] and the comprehensive monographs Bart et al. [1979, 2007, 2011] and references therein), and, in particular, the more recent Oară and Varga [2000], Oară [2005] which provide numerically efficient procedures for the solution of the continuous- and discrete-time problems under no additional *ad hoc* assumption besides that of starting from an available pre-factorization.

In Chapter 2, inspired by Youla [1961], we establish a discrete-time spectral factorization result that appears to be, to the best of our knowledge, the *most general* hitherto available. Specifically, we show that, given an *arbitrary* spectral density $\Phi(z)$, and two *arbitrary* regions featuring a geometry compatible with spectral factorization, $\Phi(z)$ admits a spectral factorization of the form in Equation (1.1), where the poles and zeros of $W(z)$ lie on the prescribed regions. The proof is constructive and gives, as a by-product, stochastic minimality of the spectral factor, which is, as previously pointed out, a crucial feature in stochastic realization theory and one of the key aspects in our analysis.

Other interesting problems connected with spectral factorization, especially in the context of stochastic realization theory, are those concerning *uniqueness of minimal spectral factors* and *parametrization of minimal spectral factors*.

From Youla [1961], it is well-known that the outer spectral factor is always minimal and *essentially unique*, *i.e.*, unique modulo left multiplication by constant orthogonal matrices. This results extends trivially to the case in which the analyticity region of the spectral factor and that of its inverse coincide. At a first sight, it is however not clear whether minimal spectral factors are still essentially unique when prescribing non-overlapping analyticity regions of the factor and of its inverse. For instance, w.r.t. the discrete-time case, assume that the desired analyticity region of the spectral factor is $\{z \in \overline{\mathbb{C}} : |z| > 1\}$, and that of its inverse is $\{z \in \overline{\mathbb{C}} : |z| < 1\}$. In this case, it seems possible that there may exist a non-constant (rational) all-pass transformation[3] that cancels out a pole inside the unit disk and, at the same time, adds the same pole to a given minimal spectral factor satisfying the desired analyticity requirements. Hence, in case it exists, such a transformation can, in principle, preserve the McMillan degree of the spectral factor. In Chapter 3, exploiting a result on minimal factorization of rational all-pass functions established in Alpay and Gohberg [1988], we will prove that this is actually not possible! This in turn implies that *all* minimal spectral factors with prescribed analyticity regions are essentially unique.

[3] A (rational) all-pass function is defined as a rational matrix $U \in \mathbb{R}(z)^{n\times n}$ s.t. $U(z)U^*(z) = I_n$ for all $z \in \mathbb{T}$ (see Definition 2.2).

In Chapter 4 we will deal with the problem of "efficiently" parametrizing the whole set of minimal spectral factors of a given spectral density. This problem has been treated in many papers Finesso and Picci [1982], Ran [1982], Ferrante et al. [1993], Ferrante [1994], Pavon [1994], Fuhrmann [1995], Ran [1995], Ferrante [1997a,b], Fuhrmann and Gombani [1998, 2000], Petersen and Ran [2001, 2002a,b], Ferrante [2005], to cite just a few key contributions. However, to the best of our knowledge, all the available results concern the continuous-time case and crucially rely on some restrictive assumptions on the spectral density, so that a general result (especially, in the discrete-time setting) has been missing so far. In this chapter, we fill this gap by providing a completely general discrete-time parametrization result in terms of

the all-pass divisors of a suitable "phase" function. The proof hinges on an elegant result of Ferrante and Picci [2017], in which the set of all-pass divisors of a given all-pass function is parametrized both algebraically (in terms of solutions of a certain Algebraic Riccati Equation) and geometrically (in terms of invariant subspaces of a certain matrix).

As a final remark, we point out that, in all the aforementioned chapters (with the only exception of Chapter 4) we will consider a factorization of the form $\Phi(z) = W^*(z)W(z)$, as in Problem 1.1, which corresponds to the solution of optimal control and circuit synthesis problems. All the results we present are, however, easily adaptable to obtain a dual counterpart for the factorization of the form $\Phi(z) = W(z)W^*(z)$. The latter is the natural factorization associated to the representation of second-order stationary stochastic processes and hence to filtering and estimation problems. If not clear by the context, we will use the terms right spectral factor and left spectral factor to distinguish the spectral factor that are solution of Problem 1.1 or of the aforesaid "dual" problem, respectively.

1.2 J-spectral factorization

DISPENSING with the assumption of positivity of the spectral density leads to a more general version of Problem 1.1, the J-spectral factorization problem. Here, the symbol $\mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$ will stand for the set of real-valued discrete-time $n \times n$ J-spectral density functions, *i.e.*, the set of rational matrix-valued functions $\Phi\colon \mathbb{T} \to \mathbb{R}^{n\times n}$ such that $\Phi(z)$ has constant signature[4] for all $z \in \mathbb{T}$ for which Φ is defined.

[4] We recall that the signature of a matrix is the difference between the number of positive and negative eigenvalues, counted with multiplicity.

Problem 1.2 ((Discrete-time) rational J-spectral factorization)**.** Given a rational J-spectral density $\Phi \in \mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$ find a factorization of the form

$$\Phi(z) = W^*(z)JW(z), \quad z \in \overline{\mathbb{C}} \text{ a.e.}, \tag{1.2}$$

where J is a signature matrix and $W(z) \in \mathbb{R}(z)^{\ell\times n}$.

By analogy with the "standard" (*i.e.*, positive semi-definite) case, the function W in (1.2) will be termed *J-spectral factor* of Φ. Furthermore, the definitions of (stochastically) minimal, outer, *etc.*, spectral factor can be extended in a straightforward manner to this general case.

The J-spectral factorization problem occurs naturally in a wide variety of problems in Systems and Control Theory, such as, to mention just a few relevant areas,

1) $\mathscr{H}_\infty$ control theory and LQ optimal games Yakubovich [1971], Banker [1972], Ball and Cohen [1987], Green et al. [1990], Ball et al. [1991], Takaba et al. [1994], Meinsma [1994], Stoorvogel [1996], Kwakernaak [1996], Ball and van der Schaft [1996], Hassibi et al. [1999], Iftime [2002],

2) Robust estimation and smoothing Colaneri et al. [1998], Colaneri and Ferrante [2002, 2004, 2006],

3) Optimal Hankel-norm model reduction Ball and Ran [1987a,b],

4) Stability analysis with Integral Quadratic Constraints (IQC) Goh [1996], Seiler [2015], Carrasco and Seiler [2015], Wang et al. [2016], Hu et al. [2016].

In the past decades, many works have addressed the J-spectral factorization problem, most of them considering the continuous-time case. More precisely, we can distinguish two main lines of research: The first one focuses on the existence of (minimal) J-spectral factorizations, whereas the second one on the numerical computation of J-spectral factorizations.

With reference to the first research topic, it is well-known that a J-spectral factorization is not always guaranteed to exist for a general para-Hermitian rational matrix-valued function[5]. In addition, even if such a factorization exists, a *minimal* one may fail to exist (see, *e.g.*, Clements [2000] for the continuous-time case and Colaneri and Ferrante [2006] for the discrete-time case). This is one of the most counterintuitive features of J-spectral factorization, when compared to the "standard" spectral factorization, for which, as mentioned before, a minimal factorization always exists. Conditions for the existence of a J-spectral factorization, and, in particular, for the existence of a minimal one have been investigated in several papers. A general existence condition, without any restriction on the analyticity regions of the J-spectral factor and its inverse, is presented in Gohberg et al. [1982][6]. This result states that an arbitrary J-spectral factorization exists if and only if the considered para-Hermitian matrix-valued function is a J-spectral density, *i.e.*, it has constant signature on the imaginary axis in the continuous time case, or on the unit circle in the discrete-time case. The main result of Ran [2003] is a necessary and sufficient condition for the existence of an outer J-spectral factor in terms of state-space realizations. Another existence criterion, stated in the frequency domain, is proposed in Meinsma [1995]. In Ran and Rodman [1991], the authors give a condition for the existence of an arbitrary J-spectral factorization which is minimal everywhere with the exception of one pre-selected point. In a connected work [Ran and Zizler, 1997, Sec. 6] it has been shown that a minimal factorization *generically* exist. Further, a condition for the existence of a complete set of minimal J-spectral factors is put forward in Karelin et al. [2001]. Finally, papers Petersen and Ran [2003], Lerer et al. [2004] address the existence and parametrization of minimal J-spectral factors in the non-square case. It is worth remarking that all the above-cited works on the existence of a minimal J-spectral factorization deal with the continuous-time case and adopt some facilitating assumptions on the J-spectral density, such as full normal rankness[7], properness, and absence of zeros/poles on the imaginary axis.

[5] A *para-Hermitian* matrix-valued function is a matrix-valued function which takes Hermitian values on the imaginary axis in the continuous-time case, or on the unit circle in the discrete-time case.

[6] See also Đoković [1993] for a shorter proof and an extension to the discrete-time case.

[7] The property of having full normal rank is often called *regularity*.

With reference to the numerical computation of J-spectral factors in continuous-time we refer to Shaked and Yaesh [1992], Kwakernaak and Sebek [1994], Aliev and Larin [1997], Park and Kailath [1997], Trentelman and Rapisarda [1999], Clements [2000], Stefanovski [2003], Zhong [2005], Zúñiga and Henrion [2006], Oară and Andrei [2011]. Many of these works address the polynomial case and adopt the same restrictive assumptions previously described. The discrete-time case does not seem to have received much attention, despite its practical relevance in applications. In fact, to the best of our knowledge, the only works dealing with this setting are Aliev and Larin [1998], Stefanovski [2004], Oară and Marinică [2013]. In the first two references, several efficient factorization algorithms for (Laurent) polynomial J-spectra are discussed. Whereas, in the third work, Riccati-based procedures for solving problems related to J-spectral factorization (such as the J-lossless factorization problem) are presented. In particular, these procedures apply in full generality, on the proviso that an arbitrary starting factorization is available. Nevertheless, these works apparently do not touch two important issues connected with the problem, namely the minimality properties of the factorization and the possibility of computing J-spectral factors with prescribed analyticity regions which differs from the "canonical" ones. These are aspects of crucial importance, for instance, in robust acausal estimation, see, *e.g.*, Colaneri and Ferrante [2006].

Chapter 5 studies the existence of *minimal* discrete-time J-spectral factorizations with *prescribed analyticity regions* of the factor and of its (right) inverse. We do so under mild assumptions on the given J-spectral density. In fact, we only require that the latter function has no poles/zeros on the unit circle, allowing, in particular, for improper and/or rank deficient J-spectral densities. Our main result consists of an easily verifiable condition for the existence of such a minimal J-spectral factor. This condition is given in terms of a decomposition of a Laurent polynomial matrix naturally connected to the starting J-spectral density. In addition, whenever such a condition is met, we provide an iterative algorithm for the calculation of the desired J-spectral factor. Remarkably, the proposed algorithm can be seen as a generalization of the discrete-time "standard" spectral factorization algorithm described in Chapter 2.

2. A general spectral factorization result *à la* Youla

In this chapter, we consider an *arbitrary* matrix-valued, rational spectral density $\Phi(z)$ and we show, with a constructive proof, that $\Phi(z)$ admits a minimal spectral factorization as in Problem 1.1. Further, the spectral factor and its (right) inverse are analytic in regions that may be selected with the *only* constraint that they satisfy some symplectic-type conditions. By suitably selecting the analyticity regions, this extremely general result particularizes into a corollary that may be viewed as the discrete-time counterpart of the matrix factorization method devised by Youla in his celebrated work Youla [1961]. Surprisingly, the discrete-time counterpart of this fundamental result has been so far missing. The reason could be due to the difficulty of deriving a result that parallels the Oono–Yasuura algorithm Oono and Yasuura [1954] that constitutes a crucial step in Youla's work.

The present chapter is based on Baggio and Ferrante [2016a] and is organized as follows. In Section 2.1, we introduce the problem and state our main result, after a few preliminary definitions. In Section 2.2, we review some notions from polynomial and rational matrix theory. Section 2.3 is devoted to collect a number of auxiliary results. In Section 2.4, we derive the proof of our main result and present some by-products of our theory. Section 2.5 shows a numerical example of the proposed factorization algorithm.

2.1 Problem formulation and main result

The object of our study is the discrete-time spectral factorization problem previously introduced in Problem 1.1: Given a arbitrary rational matrix-valued spectral density $\Phi(z) \in \mathscr{S}^n_{\text{rat}}(\mathbb{T})$, *i.e.*, a rational matrix-valued function that takes positive semidefinite values on all the points of $\mathbb{T}$ where is defined, we are interested in finding a factorization of the form

$$\Phi(z) = W^*(z)W(z),$$

with $W(z)$ being a rational spectral factor. Clearly, this problem admits many solutions. For control applications we are interested in solutions featuring some additional properties. Typical requirements are minimal "complexity" (as measured by the McMillan degree of $W(z)$), full row-rank of $W(z)$, and the fact that the poles and/or the zeros of $W(z)$ lie in certain regions of the complex plane. The most general kind of analyticity regions compatible with the problem are the following ones.

Definition 2.1 ((Weakly) Unmixed-symplectic region)**.** A set $\mathscr{A} \subset \overline{\mathbb{C}}$ is *unmixed-symplectic* if

$$\mathscr{A} \cup \mathscr{A}^* = \overline{\mathbb{C}} \setminus \mathbb{T}, \text{ and } \mathscr{A} \cap \mathscr{A}^* = \emptyset,$$

where $\mathscr{A}^* := \{z : z^{-1} \in \mathscr{A}\}$. The set $\mathscr{A} \subset \overline{\mathbb{C}}$ is *weakly unmixed-symplectic* if

$$\mathscr{A} \cup \mathscr{A}^* = \overline{\mathbb{C}}, \text{ and } \mathscr{A} \cap \mathscr{A}^* = \mathbb{T}.$$

Remark 2.1. The reason for the term "symplectic" is that $\mathscr{A}$ and $\mathscr{A}^*$ are symmetric with respect to the unit circle, a type of symmetry induced by symplectic property, see, *e.g.*, Ferrante and Levy [1998]. In this spirit, the corresponding property in continuous-time, where $\mathscr{A}^* := \{z : -z \in \mathscr{A}\}$, $\mathscr{A} \cup \mathscr{A}^*$ is the whole complex plane with the exception of the imaginary axis and $\mathscr{A} \cap \mathscr{A}^* = \emptyset$, could be called "unmixed-Hamiltonian".

The following definition will be used in what follows.

Definition 2.2 (All-pass or para-unitary matrix)**.** A rational matrix $G(z) \in \mathbb{R}(z)^{n\times n}$ is said to be *all-pass* or *para-unitary* if

$$G^*(z)G(z) = G(z)G^*(z) = I_n.$$

Remark 2.2. Notice that an all-pass matrix is unitary in the ordinary sense on the unit circle. The term all-pass is mostly used in control and estimation theory, whereas the term para-unitary matrix in network synthesis theory. Moreover, observe that the previous definition can be extended to complex rational matrices $G(z) \in \mathbb{C}(z)^{n\times n}$ which are unitary on $\mathbb{T}$ by letting $G^*(z) := \overline{G}^\top(1/\overline{z})$. These matrices are called paraconjugate-unitary or simply all-pass, as in the real case. We will exploit this "complex" extension in the next chapter.

We are now ready for our main result, whose proof is deferred to Section 2.4.

Theorem 2.1. *Let $\Phi(z) \in \mathscr{S}^n_{\text{rat}}(\mathbb{T})$ of normal rank $\text{rk}(\Phi) = r \neq 0$. Let $\mathscr{A}_p$ and $\mathscr{A}_z$ be two unmixed-symplectic sets. Then, there exists a function $W(z) \in \mathbb{R}(z)^{r\times n}$ such that*

1) *$\Phi(z) = W^*(z)W(z)$.*
2) *$W(z)$ is analytic in $\mathscr{A}_p$ and its right inverse $W^{-R}(z)$ is analytic in $\mathscr{A}_z$.*
3) *$W(z)$ is* stochastically minimal, i.e., *the McMillan degree of $W(z)$ is a half of the McMillan degree of $\Phi(z)$.*

Moreover,

4) *If $\mathscr{A}_p = \mathscr{A}_z$ then $W(z)$ satisfying points 1), and 2) is unique up to a constant, orthogonal matrix multiplier on the left,* i.e., *if $W_1(z)$ also satisfies points 1), and 2) then $W_1(z) = TW(z)$ where $T \in \mathbb{R}^{r\times r}$ is orthogonal. Therefore, if $\mathscr{A}_p = \mathscr{A}_z$, points 1) and 2) imply point 3).*

5) *If* $\Phi(z) = L^*(z)L(z)$ *is any factorization in which* $L(z) \in \mathbb{R}(z)^{r\times n}$ *is analytic in* $\mathscr{A}_z$, *then* $L(z) = V(z)W(z)$, $V(z) \in \mathbb{R}(z)^{r\times r}$ *being an all-pass matrix analytic in* $\mathscr{A}_z$. *Moreover, given an arbitrary all-pass matrix* $V(z) \in \mathbb{R}(z)^{r\times r}$ *being analytic in* $\mathscr{A}_p$, $L(z) := V(z)W(z)$ *is analytic in* $\mathscr{A}_p$ *and satisfies* $\Phi(z) = L^*(z)L(z)$, *so that, if* $\mathscr{A}_p = \mathscr{A}_z =: \mathscr{A}$ *then* $\Phi(z) = L^*(z)L(z)$ *is a factorization in which* $L(z) \in \mathbb{R}(z)^{r\times n}$ *is analytic in* $\mathscr{A}$ *if and only if* $L(z) = V(z)W(z)$, $V(z) \in \mathbb{R}(z)^{r\times r}$ *being an all-pass matrix analytic in* $\mathscr{A}$.

6) *If* $\Phi(z)$ *is analytic on the unit circle, then points 1)-5) still hold even if* $\mathscr{A}_p$ *is weakly unmixed-symplectic.*

7) *If* $\Phi(z)$ *is analytic on the unit circle and the rank of* $\Phi(z)$ *is constant on the unit circle, then points 1)-5) still hold even if* $\mathscr{A}_p$ *and/or* $\mathscr{A}_z$ *are weakly unmixed-symplectic.*

Of course, the most common requirement in Control Theory is that $W(z)$ is outer which corresponds to setting $\mathscr{A}_p = \mathscr{A}_z = \{z \in \overline{\mathbb{C}} : |z| > 1\}$ in the general case, $\mathscr{A}_p = \{z \in \overline{\mathbb{C}} : |z| \geq 1\}$ and $\mathscr{A}_z = \{z \in \overline{\mathbb{C}} : |z| > 1\}$ in the case when $\Phi(z)$ is analytic on the unit circle and $\mathscr{A}_p = \mathscr{A}_z = \{z \in \overline{\mathbb{C}} : |z| \geq 1\}$ when $\Phi(z)$ is analytic on the unit circle and the rank of $\Phi(z)$ is constant on the unit circle. This particular case of the previous result corresponds to the following result whose first six points are the discrete-time counterpart of the celebrated Youla's theorem [Youla, 1961, Thm. 2].

Theorem 2.2. *Let* $\Phi(z) \in \mathscr{S}^n_{\text{rat}}(\mathbb{T})$ *of normal rank* $\text{rk}(\Phi) = r \neq 0$. *Then, there exists a matrix* $W(z) \in \mathbb{R}(z)^{r\times n}$ *such that*

1) $\Phi(z) = W^*(z)W(z)$.

2) $W(z)$ *and its right inverse* $W^{-R}(z)$ *are both analytic in* $\{z \in \overline{\mathbb{C}} : |z| > 1\}$.

3) $W(z)$ *is unique up to a constant, orthogonal matrix multiplier on the left,* i.e., *if* $W_1(z)$ *also satisfies points 1) and 2), then* $W_1(z) = TW(z)$ *where* $T \in \mathbb{R}^{r\times r}$ *is orthogonal.*

4) *Any factorization of the form* $\Phi(z) = L^*(z)L(z)$ *in which* $L(z) \in \mathbb{R}(z)^{r\times n}$ *is analytic in* $\{z \in \overline{\mathbb{C}} : |z| > 1\}$, *has the form* $L(z) = V(z)W(z)$, *where* $V(z) \in \mathbb{R}(z)^{r\times r}$ *is an all-pass matrix analytic in* $\{z \in \overline{\mathbb{C}} : |z| > 1\}$. *Conversely, any* $L(z) = V(z)W(z)$, *where* $V(z) \in \mathbb{R}(z)^{r\times r}$ *is an all-pass matrix analytic in* $\{z \in \overline{\mathbb{C}} : |z| > 1\}$, *is a spectral factor of* $W(z)$ *analytic in* $\{z \in \overline{\mathbb{C}} : |z| > 1\}$.

5) *If* $\Phi(z)$ *is analytic on the unit circle, then* $W(z)$ *is analytic in* $\{z \in \overline{\mathbb{C}} : |z| \geq 1\}$.

6) *If* $\Phi(z)$ *is analytic on the unit circle and the rank of* $\Phi(z)$ *is constant on the unit circle, then* $W(z)$ *and its right inverse* $W^{-R}(z)$ *are both analytic in* $\{z \in \overline{\mathbb{C}} : |z| \geq 1\}$.

7) $W(z)$ *satisfying points 1) and 2) is* stochastically minimal, i.e., *the McMillan degree of* $W(z)$ *is a half of the McMillan degree of* $\Phi(z)$.

Remark 2.3. Notice that the hypothesis $\mathrm{rk}(\Phi) \neq 0$ of the previous results is only assumed to rule out the trivial case of an identically zero spectrum $\Phi(z)$ for which the only spectral factorizations clearly correspond to $W(z) = \mathbf{0}_{m,n}$, with m being arbitrary, so that, in this case, $W(z)$ cannot be chosen to be full row-rank.

Remark 2.4. A natural question which may have occurred to the reader at this point is the following one: *Why not to use a "brute-force" approach that exploits the well-known bilinear transformation technique to recover Youla's result [Youla, 1961, Thm. 2] in discrete-time?* Actually, there are many reasons to prefer a more "genuine" approach entirely conceived in the discrete-time setting. A main reason is that the bilinear transformation is not defined for every point in the complex plane and, therefore, this prevents the application of this technique in full generality. Moreover, w.r.t. the constructive procedure leading to the discrete-time result, using the bilinear transformation approach unnecessarily complicates the resulting formulas. Ultimately, it seems to be a well-accepted fact in the Systems and Control community that, as stated in [Zhou et al., 1996, p. 559] (and pointed out in many papers, *e.g.*, Ferrante and Picci [2017]),

> "[...] It is generally more appealing to give derivations in the [discrete-time] coordinates of the original data, also algorithm may be more reliable if generated for the specific model class."

2.2 Background on rational matrix-valued functions

For the benefit of the reader, we try to outline here a self-contained summary of the main notions and results from polynomial and rational matrix theory that will be used in the rest of this chapter (and in the rest of this part, as well).

Let $f(z) = p(z)/q(z) \in \mathbb{R}(z)$, $q(z) \neq 0$, be a nonzero rational function. We can always write $f(z)$ in the form

$$f(z) = \frac{n(z)}{d(z)}(z-\alpha)^{\nu}, \quad \forall\, \alpha \in \mathbb{C},$$

where ν is an integer and $n(z), d(z) \in \mathbb{R}[z]$ are nonzero polynomials such that $n(\alpha) \neq 0$ and $d(\alpha) \neq 0$. The integer ν is called *valuation of f(z) at* α and we denote it with the symbol $\nu_\alpha(f)$. The valuation of $f(z)$ at infinity is defined as $\nu_\infty(f) := \deg q(z) - \deg p(z)$, where $\deg(\cdot)$ denotes the degree of a polynomial. If $f(z)$ is the null function, by convention, $\nu_\alpha(f) = +\infty$ for every $\alpha \in \overline{\mathbb{C}}$. If $\nu_\alpha(f) < 0$, then $\alpha \in \overline{\mathbb{C}}$ is called a *pole* of $f(z)$ of multiplicity $-\nu_\alpha(f)$. If $\nu_\alpha(f) > 0$, then $\alpha \in \overline{\mathbb{C}}$ is called a *zero* of $f(z)$ of multiplicity $\nu_\alpha(f)$. The rational function $f(z)$ is said to be *proper* if $\nu_\infty(f) \geq 0$, *strictly proper* if $\nu_\infty(f) > 0$.

A polynomial matrix $G(z) \in \mathbb{R}[z]^{m\times n}$ is said to be *unimodular* if it has a polynomial inverse (either left, right or both). Similarly, a L-polynomial matrix $G(z) \in \mathbb{R}[z, z^{-1}]^{m\times n}$ is said to be *L-unimodular* if it has a L-polynomial inverse (either left, right or both). A square polynomial matrix $G(z) \in \mathbb{R}[z]^{n\times n}$ is unimodular if and only if its determinant is a nonzero constant $\alpha \in \mathbb{R}_0$. On the other hand, a square L-polynomial matrix $G(z) \in \mathbb{R}[z, z^{-1}]^{n\times n}$ is L-unimodular if and only if its determinant is a nonzero monomial αz^k, $\alpha \in \mathbb{R}_0$, $k \in \mathbb{Z}$.

Consider now a nonzero real L-polynomial vector $\mathbf{v}(z) \in \mathbb{R}[z, z^{-1}]^p$. We can write it as

$$\mathbf{v}(z) = \mathbf{v}_k z^k + \mathbf{v}_{k+1} z^{k+1} + \cdots + \mathbf{v}_{K-1} z^{K-1} + \mathbf{v}_K z^K,$$

with $\mathbf{v}_k$ and $\mathbf{v}_K$, $k \le K$, nonzero vectors in $\mathbb{R}^p$. We say that the integer k is the *minimum-degree* of $\mathbf{v}(z)$, written min deg $\mathbf{v}$, while the integer K is the *maximum-degree* of $\mathbf{v}(z)$, written max deg $\mathbf{v}$.[8]

[8] If $\mathbf{v}(z)$ is the zero vector, then min deg $\mathbf{v}$ and max deg $\mathbf{v}$ are set to $-\infty$, by convention.

Let $G(z) \in \mathbb{R}[z, z^{-1}]^{m\times n}$ and let k_i and K_i be the minimum- and maximum-degree of the i-th column of $G(z)$, for all $i = 1\ldots, m$. We define the *highest-column-degree coefficient matrix* of $G(z)$ as the constant matrix $G^{\mathrm{hc}} \in \mathbb{R}^{m\times n}$ whose i-th column consists of the coefficients of the monomials z^{K_i} in the same column of $G(z)$. Further, we define the *lowest-column-degree coefficient matrix* of $G(z)$ as the constant matrix $G^{\mathrm{lc}} \in \mathbb{R}^{m\times n}$ whose i-th column consists of the coefficients of the monomials z^{k_i} in the same column of $G(z)$. By considering, instead of the columns, the rows of $G(z)$ we can define, along the same lines in the above, the *highest-row-degree coefficient matrix* of $G(z)$, $G^{\mathrm{hr}} \in \mathbb{R}^{m\times n}$, and the *lowest-row-degree coefficient matrix* of $G(z)$, $G^{\mathrm{lr}} \in \mathbb{R}^{m\times n}$.

A classical result in rational matrix theory is the following [Youla, 1961, Sec. 2].

Theorem 2.3 (Smith–McMillan canonical form)**.** *Let $G(z) \in \mathbb{R}(z)^{m\times n}$ and let* rk$(G) = r$. *There exist unimodular matrices $U(z) \in \mathbb{R}[z]^{m\times r}$ and $V(z) \in \mathbb{R}[z]^{r\times n}$ such that*

$$D(z) := U(z)G(z)V(z) = \operatorname{diag}\left[\frac{\varepsilon_1(z)}{\psi_1(z)}, \frac{\varepsilon_2(z)}{\psi_2(z)}, \ldots, \frac{\varepsilon_r(z)}{\psi_r(z)}\right], \tag{2.1}$$

where $\varepsilon_1(z), \varepsilon_2(z), \ldots, \varepsilon_r(z), \psi_1(z), \psi_2(z), \ldots, \psi_r(z) \in \mathbb{R}[z]$ are monic polynomials satisfying the conditions: (i) $\varepsilon_i(z)$ and $\psi_i(z)$ are relatively prime, $i = 1, 2, \ldots, r$, (ii) $\varepsilon_i(z) \mid \varepsilon_{i+1}(z)$ and $\psi_{i+1}(z) \mid \psi_i(z)$, $i = 1, 2, \ldots, r-1$.[9]

[9] The notation $p(z) \mid q(z)$, with $p(z), q(z) \in \mathbb{R}[z]$, means that $p(z)$ divides $q(z)$.

The rational matrix $D(z)$ in (2.1) is known as the *Smith–McMillan canonical form* of $G(z)$. (In general, we say that a rational matrix is *canonic* if it is of the form in (2.1) and satisfies the conditions of the above theorem.) The (finite) zeros of $G(z)$ coincide with the zeros of $\varepsilon_r(z)$ and the (finite) poles of $G(z)$ with the zeros of $\psi_1(z)$. Note that, unlike what happens in the scalar case, the set of zeros and poles of a rational matrix may not be disjoint.

Remark 2.5. Usually the Smith–McMillan canonical form of $G(z)$ is defined as the quasi-diagonal form (see, *e.g.* , [Kailath, 1980, Ch. 6, §5])

$$\tilde{D}(z) = \begin{bmatrix} D(z) & \mathbf{0}_{r,n-r} \\ \mathbf{0}_{m-r,r} & \mathbf{0}_{m-r,n-r} \end{bmatrix},$$

and the unimodular matrices to obtain this form are taken to be square, namely $\tilde{U}(z) \in \mathbb{R}[z]^{m\times m}$ and $\tilde{V}(z) \in \mathbb{R}[z]^{n\times n}$. It is however apparent that the decomposition Theorem 2.3 is equivalent to the previous one. Indeed it suffices to define $U(z) = \tilde{U}(z)[I_r \mid \mathbf{0}_{m-r}]^\top$ and $V(z) = [I_r \mid \mathbf{0}_{n-r}]^\top \tilde{V}(z)$ and notice that these two matrices are again unimodular since they are polynomial and their left and right inverse, respectively, are polynomial as well.

Remark 2.6. Theorem 2.3 extends also to the case of complex matrix-valued rational functions, *i.e.*, elements of $\mathbb{C}(z)^{m\times n}$. This generalization will be used in Chapter 3.

Let $G(z) \in \mathbb{R}(z)^{m\times n}$ and write $G(z) = C(z)D(z)F(z)$, where $D(z)$ is the Smith–McMillan form of $G(z)$ and $C(z), F(z)$ are unimodular matrices. If $\mathrm{rk}(G) = m = n$, then the inverse of $G(z)$ has the form

$$G^{-1}(z) = F^{-1}(z)D^{-1}(z)C^{-1}(z)$$

and $D^{-1}(z)$ coincides with the Smith–McMillan canonical form of $G^{-1}(z)$, up to a permutation of the diagonal elements. Therefore, the poles of $G^{-1}(z)$ are exactly the zeros of $G(z)$. Likewise, if $G(z)$ has normal rank m (n), there always exists a right (left) inverse of $G(z)$ such that the poles of $G^{-R}(z)$ ($G^{-L}(z)$) coincide with the zeros of $G(z)$.[10] Indeed, we may take

$$G^{-R}(z) = F^{-R}(z)D^{-1}(z)C^{-1}(z), \tag{2.2}$$
$$G^{-L}(z) = F^{-1}(z)D^{-1}(z)C^{-L}(z). \tag{2.3}$$

[10] The latter fact is not true for all the right/left inverses of $G(z)$, since, in general, the zeros of $G(z)$ are among the poles of all such inverses, see [Kailath, 1980, Ch. 6, §5, Ex. 14].

In the following, we consider only right and left inverses of the form (2.2) and (2.3), respectively.

Let $\alpha_1, \alpha_2, \ldots, \alpha_t$ be the (finite) zeros and (finite) poles of $G(z)$. We can write the Smith–McMillan canonical form of $G(z)$ as

$$\mathrm{diag}\Big[(z-\alpha_1)^{\nu_1^{(1)}} \cdots (z-\alpha_t)^{\nu_t^{(1)}}, \ldots, (z-\alpha_1)^{\nu_1^{(r)}} \cdots (z-\alpha_t)^{\nu_t^{(r)}}\Big].$$

The integer exponents $\nu_i^{(1)} \le \nu_i^{(2)} \le \cdots \le \nu_i^{(r)}$, appearing in the above expression, are called the *structural indices* of $G(z)$ at α_i and they are used to represent the zero-pole structure at α_i of $G(z)$. To obtain the zero-pole structure at infinity of $G(z)$, we can proceed as follows. We make a change of variable, $z \mapsto \lambda^{-1}$, and compute the Smith–McMillan form of $G(\lambda^{-1})$, then the structural indices of $G(\lambda^{-1})$ at $\lambda = 0$ will give the

set of structural indices of $G(z)$ at $z=\infty$. Lastly, if $p_1,\dots,p_h$ are the distinct poles (the pole at infinity included) of $G(z)$, we recall that the *McMillan degree* of $G(z)$ can be defined as (see, *e.g.*, [Kailath, 1980, Ch. 6, §5])

$$\delta_M(G) := \sum_{i=1}^{h} \delta_p(G; p_i), \tag{2.4}$$

where $\delta_p(G;p_i)$ is the degree of the pole p_i, *i.e.*, the largest multiplicity that p_i possesses as a pole of any minor[11] of $G(z)$. In particular, if $D(z)$ in (2.1) is the Smith–McMillan form of $G(z)$ and $G(z)$ has no pole at infinity then $\delta_p(G;p_i)=\delta_p(D;p_i)$ for all $i=1,\dots,h$, which, in turn, yields $\delta_M(G)=\delta_M(D)=\sum_{i=1}^{r}\deg\psi_i(z)$.

[11] We recall that a *minor* of a matrix A is the determinant of some smaller square matrix, cut down from A by removing one or more of its rows, indexed by tuple $\mathbf{i}$, or columns, indexed by tuple $\mathbf{j}$. If, furthermore, $\mathbf{i}=\mathbf{j}$ the minor is called a *principal minor* of A.

2.3 Preliminary analysis

In this section, we collect a set of lemmata which we will exploit in the constructive proof of Theorem 2.1.

Lemma 2.1. *A matrix $G(z)\in\mathbb{R}(z)^{m\times n}$ is analytic in $\mathbb{C}_0$ together with its inverse (either right, left or both) if and only if it is a L-unimodular polynomial matrix.*

Proof. If $G(z)$ is L-unimodular, then $G(z)$ has an inverse (either left, right or both) which is L-polynomial. Hence, the only possible finite zeros/poles of $G(z)$ are located at $z=0$. This, in turn, implies that $G(z)$ must be analytic together with its inverse in $\mathbb{C}_0$.

Vice versa, suppose that $G(z)$ is analytic with its inverse in $\mathbb{C}_0$. Firstly, we notice that the existence of a left or right inverse for $G(z)$ implies that the normal rank of $G(z)$ is either $r=n$ or $r=m$, respectively. Without loss of generality, we can suppose that $r=n$. By virtue of Theorem 2.3, we can write $G(z)=C(z)D(z)F(z)$, where $C(z)\in\mathbb{R}[z]^{m\times n}$, $F(z)\in\mathbb{R}[z]^{n\times n}$ are unimodular (and, *a fortiori*, L-unimodular) polynomial matrices, respectively, and $D(z)\in\mathbb{R}(z)^{n\times n}$ is diagonal, canonic of the form

$$D(z)=\operatorname{diag}\left[\frac{\varepsilon_1(z)}{\psi_1(z)},\frac{\varepsilon_2(z)}{\psi_2(z)},\dots,\frac{\varepsilon_n(z)}{\psi_n(z)}\right].$$

The analyticity of $G(z)$ in $\mathbb{C}_0$ implies that all $\psi_i(z)\in\mathbb{R}[z]$, $i=1,\dots,n$, are nonzero monomials. The Smith–McMillan canonical form of $G^{-L}(z)$ is given by

$$\operatorname{diag}\left[\frac{\psi_n(z)}{\varepsilon_n(z)},\frac{\psi_{n-1}(z)}{\varepsilon_{n-1}(z)},\dots,\frac{\psi_1(z)}{\varepsilon_1(z)}\right].$$

Hence, the analyticity of $G^{-L}(z)$ in $\mathbb{C}_0$ implies that all $\varepsilon_i(z)\in\mathbb{R}[z]$, $i=1,\dots,n$, are nonzero monomials. Therefore, $D(z)$ is a L-unimodular polynomial matrix. Since $G(z)=C(z)D(z)F(z)$ is the product of three L-unimodular polynomial matrices, $G(z)$ must be a L-unimodular polynomial matrix. □

Lemma 2.2. *Let $\mathscr{A} \subset \overline{\mathbb{C}}$ be an unmixed-symplectic set. An all-pass matrix $G(z) \in \mathbb{R}(z)^{n\times n}$ analytic in $\mathscr{A}$ with inverse analytic in $\mathscr{A}$ is a constant orthogonal matrix.*

Proof. The analyticity of the inverse of $G(z)$ in $\mathscr{A}$ implies that of $G(z^{-1})$ in the same region, and therefore that of $G(z)$ in $\mathscr{A}^*$. We also notice that in the unit circle it holds $G^*(e^{j\omega})G(e^{j\omega}) = G^\top(e^{-j\omega})G(e^{j\omega}) = I_n$, $\forall\,\omega \in [0, 2\pi)$, and we can write out the diagonal elements in expanded form as

$$\sum_{i=1}^{n} |[G(e^{j\omega})]_{ik}|^2 = 1, \quad \forall\, k = 1, \ldots, n,\ \forall\,\omega \in [0, 2\pi).$$

The latter equation implies that

$$|[G(e^{j\omega})]_{ik}| \leq 1, \quad \forall\, i, k = 1, \ldots, n,\ \forall\,\omega \in [0, 2\pi),$$

and, therefore, we proved the analyticity of $G(z)$ on the unit circle. By Definition 2.1 of unmixed-symplectic set, it follows that $G(z)$ is analytic on the entire extended complex plane. This means that $G(z)$ is analytic and bounded in $\mathbb{C}$. Hence, we can apply Liouville's Theorem [Lang, 1985, Ch. V, Thm. 1.4] and conclude that $G(z)$ must be a constant orthogonal matrix. □

Remark 2.7. With the usual choice $\mathscr{A} = \{z \in \overline{\mathbb{C}} : |z| > 1\}$, the previous lemma reads as follows: An all-pass matrix $G(z) \in \mathbb{R}(z)^{n\times n}$ analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$ with inverse analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$ is a constant orthogonal matrix.

Definition 2.3 (Left-standard factorization)**.** Let $G(z) \in \mathbb{R}(z)^{m\times n}$ and let $\mathrm{rk}(G) = r$. A decomposition of the form $G(z) = A(z)\Delta(z)B(z)$ is called a *left-standard factorization* if

1) $\Delta(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal and analytic with its inverse in $\{z \in \mathbb{C}_0 : |z| \neq 1\}$;

2) $A(z) \in \mathbb{R}(z)^{m\times r}$ is analytic together with its left inverse in $\{z \in \mathbb{C}_0 : |z| \leq 1\}$;

3) $B(z) \in \mathbb{R}(z)^{r\times n}$ is analytic together with its right inverse in $\{z \in \mathbb{C} : |z| \geq 1\}$.

Remark 2.8. If, in Definition 2.3, $A(z)$ and $B(z)$ are interchanged, we have a *right-standard factorization*. Hence, it follows that any left-standard factorization of $G(z)$ generates a right-standard factorization of $G^\top(z)$, $G^{-1}(z)$ (if $G(z)$ is nonsingular), $G(z^{-1})$, *e.g.*, in the first case we have $G^\top(z) = B^\top(z)\Delta(z)A^\top(z)$.

Lemma 2.3. *Any rational matrix $G(z) \in \mathbb{R}(z)^{m\times n}$ of normal rank $\mathrm{rk}(G) = r$ admits a left-standard factorization.*

Proof. In view of Theorem 2.3, we can write $G(z) = C(z)D(z)F(z)$, where $C(z) \in \mathbb{R}[z]^{m\times r}$, $F(z) \in \mathbb{R}[z]^{r\times n}$ are unimodular polynomial matrices and $D(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal and canonic of the form

$$D(z) = \operatorname{diag}\left[\frac{\varepsilon_1(z)}{\psi_1(z)}, \frac{\varepsilon_2(z)}{\psi_2(z)}, \ldots, \frac{\varepsilon_r(z)}{\psi_r(z)}\right].$$

We factor $\varepsilon_i(z) \in \mathbb{R}[z]$ and $\psi_i(z) \in \mathbb{R}[z]$, $i = 1, \ldots, r$, in $D(z)$ into the product of three polynomials: the first without zeros in $\{z \in \mathbb{C} : |z| \leq 1\}$, the second without zeros in $\{z \in \mathbb{C} : |z| \neq 1\}$ and the third without zeros in $\{z \in \mathbb{C} : |z| \geq 1\}$. Thus, it is possible to write

$$D(z) = D_-(z)\Delta(z)D_+(z),$$

where $D_-(z)$ and its inverse are analytic in $\{z \in \mathbb{C} : |z| \leq 1\}$, $\Delta(z)$ and its inverse in $\{z \in \mathbb{C} : |z| \neq 1\}$ and $D_+(z)$ and its inverse in $\{z \in \mathbb{C} : |z| \geq 1\}$. Eventually, by choosing $A(z) := C(z)D_-(z)$ and $B(z) := D_+(z)F(z)$, we have that $G(z) = A(z)\Delta(z)B(z)$ is a left-standard factorization of $G(z)$. □

Left-standard factorizations are not unique. Indeed, any two decompositions are connected as follows.

Lemma 2.4. *Let $G(z) \in \mathbb{R}(z)^{m\times n}$ be a rational matrix of normal rank $\operatorname{rk}(G) = r$ and let $G(z) = A(z)\Delta(z)B(z) = A_1(z)\Delta_1(z)B_1(z)$ be two left-standard factorizations of $G(z)$. Then,*

$$A_1(z) = A(z)M^{-1}(z), \quad B_1(z) = N(z)B(z),$$

where $M(z) \in \mathbb{R}[z, z^{-1}]^{r\times r}$ and $N(z) \in \mathbb{R}[z, z^{-1}]^{r\times r}$ are two L-unimodular polynomial matrices such that

$$M(z)\Delta(z)N^{-1}(z) = \Delta_1(z). \tag{2.5}$$

Proof. By assumption,

$$G(z) = A(z)\Delta(z)B(z) = A_1(z)\Delta_1(z)B_1(z)$$

which, in turn, implies

$$\Delta_1^{-1}(z)A_1^{-L}(z)A(z)\Delta(z) = B_1(z)B^{-R}(z). \tag{2.6}$$

By Definition 2.3 of left-standard factorization, the right-hand side of (2.6) is analytic in $\{z \in \mathbb{C} : |z| \geq 1\}$, while the left-hand side of (2.6) in $\{z \in \mathbb{C}_0 : |z| < 1\}$. Therefore, it follows that $B_1(z)B^{-R}(z)$ is analytic in $\mathbb{C}_0$. Moreover, the inverse of $B_1(z)B^{-R}(z)$ satisfies

$$[B_1(z)B^{-R}(z)]^{-1} = \Delta^{-1}(z)[A_1^{-L}(z)A(z)]^{-1}\Delta_1(z)$$

and is also analytic in $\mathbb{C}_0$. Thus, by Lemma 2.1, $N(z) := B_1(z)B^{-R}(z)$ must be a L-unimodular matrix. Similarly, $M(z) := A_1^{-L}(z)A(z)$ is a L-unimodular matrix. Finally, a rearrangement of (2.6) yields (2.5). □

Remark 2.9. Notice that, by replacing the word "left-standard" with the word "right-standard" in Lemma 2.3 and Lemma 2.4, we obtain, by minor modifications in the proofs, a right-standard counterpart of Lemma 2.3 and Lemma 2.4.

Let $\Phi(z) \in \mathbb{R}(z)^{n\times n}$ be a (discrete-time) para-Hermitian matrix, *i.e.*, a rational matrix which takes Hermitian values on the unit circle, of normal rank $\mathrm{rk}(\Phi) = r$ and let $\Phi(z) = A(z)\Delta(z)B(z)$ be a left-standard factorization of $\Phi(z)$. We have that

$$\Phi(z) = \Phi^*(z) = B^*(z)\Delta^*(z)A^*(z)$$

is also a left-standard factorization of $\Phi(z)$. In particular, $\Delta^*(z)$ is equal to $\Delta(z)$, except for multiplication of suitable monomials of the form $\pm z^{k_i}$ in its diagonal entries, *i.e.*,

$$\Delta^*(z) = \Sigma(z)\Delta(z),$$

where

$$\Sigma(z) = \mathrm{diag}\,[e_1(z), e_2(z), \ldots, e_r(z)] \tag{2.7}$$

and $e_i(z) = \pm z^{k_i},\ k_i \in \mathbb{Z},\ i = 1,\ldots,r$. By invoking Lemma 2.4, we can write

$$A^*(z) = N(z)B(z), \quad B^*(z) = A(z)M^{-1}(z), \tag{2.8}$$

where $N(z)$, $M(z) \in \mathbb{R}[z, z^{-1}]^{r\times r}$ are L-unimodular matrices.

The following two lemmata are used to establish a further characterization of a rational matrix when it is positive semi-definite upon the unit circle.

Lemma 2.5. *Let $G(z) \in \mathbb{R}(z)^{n\times n}$ and let $\mathbb{Y}$ be a region of the complex plane such that*

1) $G(z)$ is Hermitian on $\mathbb{Y}$;

2) $\mathbf{x}^\top G(\lambda)\mathbf{x} \geq 0$, $\forall\, \mathbf{x} \in \mathbb{R}^n$ and $\forall\, \lambda \in \tilde{\mathbb{Y}} \subseteq \mathbb{Y}$ for which $G(\lambda)$ has finite entries.

Let $D(z) \in \mathbb{R}(z)^{r\times r}$ be the Smith–McMillan canonical form of $G(z)$ and denote by $g_{\mathbf{ij}}^{(\ell)}$ and $d_{\mathbf{ij}}^{(\ell)}$ the $\ell \times \ell$ minors ($1 \leq \ell \leq r$) of the rational matrices $G(z)$ and $D(z)$, respectively, obtained by selecting those rows and columns whose indices appear in the ordered ℓ-tuples $\mathbf{i}$ and $\mathbf{j}$, respectively. Then,

$$\min_{\mathbf{i}} v_\alpha(d_{\mathbf{ii}}^{(\ell)}) = \min_{\mathbf{i}} v_\alpha(g_{\mathbf{ii}}^{(\ell)}), \quad \forall \alpha \in \mathbb{Y}.$$

Proof. Firstly, we recall that for any rational matrix $G(z)$ it holds

$$\min_{\mathbf{i}} v_\alpha(d_{\mathbf{ii}}^{(\ell)}) = \min_{\mathbf{ij}} v_\alpha(d_{\mathbf{ij}}^{(\ell)}) = \min_{\mathbf{ij}} v_\alpha(g_{\mathbf{ij}}^{(\ell)}), \quad \forall \alpha \in \mathbb{C}.$$

The latter result is well-known and is presented, for instance, as an exercise in [Kailath, 1980, Ch. 6, §5, Ex. 6]. Hence, it remains to prove that

$$\min_{\mathbf{ij}} v_\alpha(g_{\mathbf{ij}}^{(\ell)}) = \min_{\mathbf{i}} v_\alpha(g_{\mathbf{ii}}^{(\ell)}), \quad \forall \alpha \in \mathbb{Y}. \tag{2.9}$$

Since $G(z)$ is Hermitian positive semi-definite on the region $\tilde{\mathbb{Y}}$, it admits a decomposition of the form $G(\lambda) = W(\lambda)\overline{W(\lambda)}^{\top}$ for all $\lambda \in \tilde{\mathbb{Y}}$. By applying the Binet–Cauchy Theorem (see [Gantmacher, 1959, Vol. I, Ch. 1, §2]), we have

$$g_{\mathbf{ij}}^{(\ell)}(\lambda) = \sum_{\mathbf{h}} w_{\mathbf{ih}}^{(\ell)}(\lambda)\,\overline{w_{\mathbf{jh}}^{(\ell)}(\lambda)}, \quad \forall\, \lambda \in \tilde{\mathbb{Y}}, \tag{2.10}$$

$$g_{\mathbf{ii}}^{(\ell)}(\lambda) = \sum_{\mathbf{h}} w_{\mathbf{ih}}^{(\ell)}(\lambda)\,\overline{w_{\mathbf{ih}}^{(\ell)}(\lambda)} = \sum_{\mathbf{h}} \left| w_{\mathbf{ih}}^{(\ell)}(\lambda) \right|^2, \quad \forall\, \lambda \in \tilde{\mathbb{Y}}, \tag{2.11}$$

where $g_{\mathbf{ij}}^{(\ell)}(\lambda)$ and $w_{\mathbf{ij}}^{(\ell)}(\lambda)$ denote the $\ell \times \ell$ minors of matrices $G(\lambda)$ and $W(\lambda)$, obtained by selecting those rows and columns whose indices appear in the ordered ℓ-tuples $\mathbf{i}$ and $\mathbf{j}$, respectively. Moreover, in both the summations (2.10)-(2.11), $\mathbf{h} := (h_1, \ldots, h_\ell)$, $1 \le h_1 < \cdots < h_\ell \le n$, runs through all such multi-indices. By using Cauchy–Schwarz inequality and (2.10)-(2.11), we have

$$\begin{aligned}
\left| g_{\mathbf{ij}}^{(\ell)}(\lambda) \right| &= \left| \sum_{\mathbf{h}} w_{\mathbf{ih}}^{(\ell)}(\lambda)\,\overline{w_{\mathbf{jh}}^{(\ell)}(\lambda)} \right| \\
&\le \left(\sum_{\mathbf{h}} \left| w_{\mathbf{ih}}^{(\ell)}(\lambda) \right|^2 \sum_{\mathbf{h}} \left| w_{\mathbf{jh}}^{(\ell)}(\lambda) \right|^2 \right)^{1/2} \\
&= \left(g_{\mathbf{ii}}^{(\ell)}(\lambda)\, g_{\mathbf{jj}}^{(\ell)}(\lambda) \right)^{1/2} \\
&\le \max\left\{ g_{\mathbf{ii}}^{(\ell)}(\lambda), g_{\mathbf{jj}}^{(\ell)}(\lambda) \right\}, \quad \forall\, \lambda \in \tilde{\mathbb{Y}}.
\end{aligned} \tag{2.12}$$

The latter inequality implies that for every zero $\alpha \in \mathbb{Y}$ of multiplicity k of a minor of $G(z)$, there exists at least one principal minor of $G(z)$ which has the same α either as a zero of multiplicity less than or equal to k or a pole of multiplicity greater than or equal to 0. Similarly, inequality (2.12) implies also that for every pole $\alpha \in \mathbb{Y}$ of multiplicity k of a minor of $G(z)$, there exists at least one principal minor of $G(z)$ which has the same pole of multiplicity greater than or equal to k. Therefore, we conclude that (2.9) holds. □

Lemma 2.6. *Let $\Phi(z) \in \mathscr{S}_{\mathrm{rat}}^n(\mathbb{T})$ of normal rank $\mathrm{rk}(\Phi) = r$ and let $D(z) \in \mathbb{R}(z)^{r\times r}$ be its Smith–McMillan canonical form. Then, the zeros and poles on the unit circle of the diagonal elements of $D(z)$ have even multiplicity.*

Proof. Firstly, we assume that the numerators and denominators of all entries in $\Phi(z)$ are relatively prime polynomials. Let $\alpha_1 = e^{j\omega_1}$, $\alpha_2 = e^{j\omega_2}, \ldots, \alpha_t = e^{j\omega_t}$, be the zeros/poles on the unit circle of $\Phi(z)$ and let $\nu_i^{(1)}, \nu_i^{(2)}, \ldots, \nu_i^{(r)}$, $(\nu_i^{(1)} \le \nu_i^{(2)} \le \cdots \le \nu_i^{(r)})$, be the structural indices of $\Phi(z)$ at α_i, $i = 1, \ldots, t$. Since $\Phi(z)$ is positive semi-definite on the unit circle, one can directly verify that the zeros and poles on the unit circle of the principal minors of $\Phi(z)$ must have even multiplicity. Now, we are in position to

apply Lemma 2.5. By considering the minors of order $\ell = 1$, it follows that $\nu_i^{(1)}$ is even for all $i = 1, 2, \dots, t$. Similarly, by considering the minors of order $\ell = 2$ in Lemma 2.5, it follows that $\nu_i^{(1)} + \nu_i^{(2)}$ is even for all $i = 1, 2, \dots, t$. Since $\nu_i^{(1)}$ is even, then also $\nu_i^{(2)}$ must be even for all $i = 1, 2, \dots, t$. Iterating the argument, we conclude that every zero/pole on the unit circle of the diagonal elements of $D(z)$ has even multiplicity. □

Remark 2.10. Lemma 2.5 can also be used to obtain an alternative proof of [Youla, 1961, Lemma 4, point 2], which represents the continuous-time counterpart of Lemma 2.6.

Lemma 2.7. *Let $\Phi(z) \in \mathscr{S}_{\text{rat}}^n(\mathbb{T})$ of normal rank $\operatorname{rk}(\Phi) = r$ and let $D(z) \in \mathbb{R}(z)^{r\times r}$ be its Smith–McMillan canonical form. Then $D(z)$ can be written as*

$$D(z) = \Sigma(z)\Lambda^*(z)\Theta^*(z)\Theta(z)\Lambda(z) \tag{2.13}$$

where $\Lambda(z)$ is diagonal, canonic and analytic with its inverse in $\{z \in \mathbb{C} : |z| \geq 1\}$ and, if $z = 0$ is either a zero, pole or both of $D(z)$, $\Lambda(z)$ has the same structural indices at $z = 0$ of $D(z)$; $\Theta(z)$ is diagonal, canonic and analytic with its inverse in $\mathbb{C} \setminus \mathbb{T}$; $\Sigma(z)$ has the form

$$\Sigma(z) = \operatorname{diag}[e_1(z), e_2(z), \dots, e_r(z)], \tag{2.14}$$

with $e_i(z) = \alpha_i z^{k_i}$, $\alpha_i \in \mathbb{R}_0$, $k_i \in \mathbb{Z}$, $i = 1, \dots, r$.

Proof. By direct computation, we obtain

$$D^*(z) = \Sigma'(z)\bar{D}(z), \tag{2.15}$$

where $\bar{D}(z)$ is canonic and $\Sigma'(z)$ is a diagonal matrix with elements αz^k, $\alpha \in \mathbb{R}_0$, $k \in \mathbb{Z}$, on its diagonal. Since $\Phi(z)$ is a spectrum, we can write

$$\Phi(z) = C(z)D(z)F(z) = F^*(z)D^*(z)C^*(z) = \Phi^*(z),$$

The matrices $F(z) \in \mathbb{R}[z]^{n\times r}$, $C(z) \in \mathbb{R}[z]^{r\times n}$, are unimodular, while $F^*(z)$, $C^*(z)$ are L-unimodular. By Lemma 2.1, $F(z)$, $C(z)$, $F^*(z)$, $C^*(z)$ are analytic in $\mathbb{C}_0$ with their inverses. Thus, we have (see [Kailath, 1980, Ch. 6, §5, Ex. 6])

$$\min_{\mathbf{i}} v_\alpha(d_{\mathbf{ii}}^{(\ell)}) = \min_{\mathbf{i}} v_\alpha(d_{\mathbf{ii}}^{*(\ell)}), \quad \forall \alpha \in \mathbb{C}_0, \ \forall \ell : 1 \leq \ell \leq r,$$

where $d_{\mathbf{ii}}^{(\ell)}$ and $d_{\mathbf{ii}}^{*(\ell)}$ denote the $\ell \times \ell$ minors of $D(z)$ and $D^*(z)$, respectively, obtained by selecting those rows and columns whose indices appear in the ordered ℓ-tuple $\mathbf{i}$. The previous equation implies that, for every $\alpha \in \mathbb{C}_0$, being either a pole, zero or both of $D(z)$, $D^*(z)$ has the same structural indices at α of $D(z)$. Therefore, since by (2.15) $\bar{D}(z)$ is canonic, it follows that

$$D^*(z) = \Sigma''(z)D(z)$$

where $\Sigma''(z)$ is diagonal with elements αz^k, $\alpha \in \mathbb{R}_0$, $k \in \mathbb{Z}$, on its diagonal. This means that any zero/pole at $\alpha \in \mathbb{C}_0$ in the diagonal terms of $D(z)$ is accompanied by a zero/pole at α^{-1}, and we can always write $D(z)$ as

$$D(z) = \Sigma_1(z)\Lambda^*(z)\Delta(z)\Lambda(z), \tag{2.16}$$

where $\Sigma_1(z)$ is diagonal with elements αz^k, $\alpha \in \mathbb{R}_0$, $k \in \mathbb{Z}$, on its diagonal; $\Lambda(z)$ and $\Delta(z)$ are diagonal, canonic and analytic with their inverse in $\{z \in \mathbb{C} : |z| \geq 1\}$ and $\mathbb{C} \setminus \mathbb{T}$, respectively. Moreover, if $z = 0$ is either a pole, zero or both of $D(z)$, $\Lambda(z)$ possesses the same structural indices at $z = 0$ of $D(z)$. As a matter of fact, let $\alpha_{i,k}$, $i = 1, \ldots, p_k$, and $\beta_{j,k}$, $j = 1, \ldots, q_k$, be the zeros and poles, respectively, in $\{z \in \mathbb{C}_0 : |z| < 1\}$ of $[D(z)]_{kk}$ and let $h_k \in \mathbb{Z}$ be the valuation at $z = 0$ of $[D(z)]_{kk}$. We can write, for all $k = 1, \ldots, r$,

$$\begin{aligned}
[D(z)]_{kk} &= z^{h_k} \frac{\prod_{i=1}^{p_k}(z - \alpha_{i,k}^{-1})(z - \alpha_{i,k})}{\prod_{j=1}^{q_k}(z - \beta_{j,k}^{-1})(z - \beta_{j,k})}[\Delta(z)]_{kk} \\
&= \underbrace{\gamma_k \frac{z^{h_k}}{z^{q_k - p_k}}}_{[\Sigma_1(z)]_{kk}} \underbrace{z^{-h_k} \frac{\prod_{i=1}^{p_k}(z^{-1} - \alpha_{i,k})}{\prod_{j=1}^{q_k}(z^{-1} - \beta_{j,k})}}_{[\Lambda^*(z)]_{kk}} [\Delta(z)]_{kk} \underbrace{z^{h_k} \frac{\prod_{i=1}^{p_k}(z - \alpha_{i,k})}{\prod_{j=1}^{q_k}(z - \beta_{j,k})}}_{[\Lambda(z)]_{kk}}
\end{aligned}$$

with $\gamma_k := (-1)^{q_k - p_k} \frac{\prod_{j=1}^{q_k} \beta_{j,k}}{\prod_{i=1}^{p_k} \alpha_{i,k}}$.

Now, by exploiting Lemma 2.6, $\Delta(z)$ can be written as

$$\Delta(z) = \Theta^2(z) = \Sigma_2(z)\Theta^*(z)\Theta(z),$$

with $\Sigma_1(z)$ diagonal with elements $\pm z^k$, $k \in \mathbb{Z}$, on its diagonal and $\Theta(z)$ diagonal, canonic and analytic together with its inverse in $\mathbb{C} \setminus \mathbb{T}$. Finally, we can rearrange $D(z)$ in the form

$$D(z) = \Sigma(z)\Lambda^*(z)\Theta^*(z)\Theta(z)\Lambda(z),$$

where $\Sigma(z) := \Sigma_1(z)\Sigma_2(z)$ has the form in (2.14). □

To conclude this section, we report below another useful result.

Lemma 2.8. *Let $\Psi(z) \in \mathscr{S}^r_{\text{rat}}(\mathbb{T})$ be L-unimodular and strictly positive definite on the unit circle. Then, Ψ^{hc} is nonsingular if and only if $\Psi(z)$ is a constant matrix.*

Proof. If $\Psi(z)$ is a constant matrix then $\Psi^{\text{hc}} = \Psi(z)$ is nonsingular, by definition of L-unimodular matrix.

Conversely, assume that Ψ^{hc} is nonsingular. Let us denote by $K_i \in \mathbb{Z}$, $i = 1, \ldots, r$, the maximum-degree of the i-th column of $\Psi(z)$ and by $k_i \in \mathbb{Z}$, $i = 1, \ldots, r$, the minimum-degree of the i-th row of $\Psi(z)$. Since $\Psi(z) = \Psi^*(z)$, we have that $\det \Psi(z)$ is a nonzero real constant and

$$K_i = -k_i, \quad i = 1, \ldots, r. \tag{2.17}$$

Moreover, since $\Psi(z)$ is positive definite on the unit circle, the diagonal elements of $\Psi(z)$ cannot be equal to zero and, therefore, $K_i \geq 0$, $i = 1,\ldots,r$. Actually, the nonsingularity of Ψ^{hc} yields

$$K_i = 0, \quad i = 1,\ldots,r, \tag{2.18}$$

otherwise one can check, by exploiting the Leibniz formula for determinants, that the maximum-degree of $\det\Psi(z)$ would be strictly positive; but this is not possible since, as noticed above, $\det\Psi(z)$ is a nonzero real constant and so $\max\deg(\det\Psi(z)) = 0$. By (2.18), all the entries of $\Psi(z)$ must have maximum-degree less than or equal to zero. But, by (2.17), $k_i = -K_i$ for all $i = 1,\ldots,r$, and so (2.18) also implies that all the entries of $\Psi(z)$ must have minimum-degree greater than or equal to zero. We conclude that

$$\max\deg[\Psi(z)]_{ij} = \min\deg[\Psi(z)]_{ij} = 0, \quad i,\, j = 1,\ldots,r,$$

and, therefore, $\Psi(z)$ must be a constant matrix. □

2.4 Proof of the main theorem

We are now ready to prove the main result of this chapter. For the sake of clarity and readability, we first prove the special case of Theorem 2.2 and we then proceed to the proof of our general Theorem 2.1.

Proof of Theorem 2.2. We first prove statement 3). Let $W(z)$ and $W_1(z)$ be two matrices satisfying 1) and 2). Then,

$$W^*(z)W(z) = W_1^*(z)W_1(z). \tag{2.19}$$

The latter equation implies $V^*(z)V(z) = I_r$, where $V(z) := W_1(z)W^{-R}(z)$ is analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$. Thus, $V(z) \in \mathbb{R}(z)^{r\times r}$ is an all-pass matrix analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$. Moreover, we have that $\Delta_1(z) := W_1(z) - V(z)W(z) = W_1(z)[I_n - W^{-R}(z)W(z)]$ satisfies

$$\begin{aligned}\Delta_1^*(z)\Delta_1(z) &= [I_n - W^*(z)W^{-R*}(z)]W_1^*(z)W_1(z)[I_n - W^{-R}(z)W(z)] \\ &= [I_n - W^*(z)W^{-R*}(z)]W^*(z)W(z)[I_n - W^{-R}(z)W(z)] = 0,\end{aligned} \tag{2.20}$$

so that

$$W_1(z) = V(z)W(z) \tag{2.21}$$

yielding that $V^{-1}(z) = W(z)W_1^{-R}(z)$ is analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$. In view of Lemma 2.2, we conclude that $V(z)$ is a constant orthogonal matrix.

Consider now statement 4) and let $\Phi(z) = L^*(z)L(z)$ where $L(z) \in \mathbb{R}(z)^{n\times r}$ is analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$. In this case, we can write

$$L^*(z)L(z) = W^*(z)W(z).$$

The latter equation implies $V^*(z)V(z) = I_r$, where $V(z) := L(z)W^{-R}(z)$ and $W(z) \in \mathbb{R}(z)^{r\times n}$ is a rational matrix satisfying 1) and 2). Since $L(z)$ and $W^{-R}(z)$ are both analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$, then $V(z) \in \mathbb{R}(z)^{r\times r}$ is an all-pass matrix analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$. The same computation that led to (2.21) now gives $L(z) = V(z)W(z)$.

Now, we provide a constructive proof of statements 1) and 2), which represent the core of the Theorem. The procedure is divided in four steps.

Step 1. Reduce $\Phi(z)$ to the Smith–McMillan canonical form. By using the same standard procedure described, for instance, in [Youla, 1961, Thm. 2], we arrive at

$$\Phi(z) = C(z)D(z)F(z), \tag{2.22}$$

where $C(z) \in \mathbb{R}[z]^{n\times r}$, $F(z) \in \mathbb{R}[z]^{r\times n}$ are unimodular polynomial matrices and $D(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal and canonic.

Step 2. According to Lemma 2.7, we can write $D(z)$ in the form

$$D(z) = \Sigma(z)\Lambda^*(z)\tilde{\Delta}(z)\Lambda(z), \tag{2.23}$$

where:

1) $\Lambda(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal, canonic and analytic together with $\Lambda^{-1}(z)$ in $\{z \in \mathbb{C} : |z| \geq 1\}$ and, if $z = 0$ is either a pole, zero or both of $D(z)$, $\Lambda(z)$ possesses the same structural indices at $z = 0$ of $D(z)$;

2) $\tilde{\Delta}(z) := \Theta^*(z)\Theta(z) = \tilde{\Delta}^*(z)$, where $\Theta(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal, canonic and analytic together with $\Theta^{-1}(z)$ in $\{z \in \mathbb{C} : |z| \neq 1\}$;

3) $\Sigma(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal of the form

$$\Sigma(z) = \operatorname{diag}[e_1(z), e_2(z), \ldots, e_r(z)],$$

where $e_i(z) = \alpha_i z^{k_i}$, $\alpha_i \in \mathbb{R}_0$, $k_i \in \mathbb{Z}$, $i = 1, \ldots, r$.

Let us define

$$A(z) := C(z)\Sigma(z)\Lambda^*(z), \quad B(z) := \Lambda(z)F(z).$$

We have that $\Phi(z) = A(z)\tilde{\Delta}(z)B(z)$ is a left-standard factorization of $\Phi(z)$.

Step 3. Let $I(z) := B^{-R}(z)\Theta^{-1}(z)$. By (2.8), we have $A^*(z) = N(z)B(z)$ and, therefore,

$$\begin{aligned} &I^*(z)\Phi(z)I(z) = I^*(z)\Phi^*(z)I(z) \\ &= \Theta^{-*}(z)B^{-R*}(z)B^*(z)\tilde{\Delta}^*(z)N(z)B(z)B^{-R}(z)\Theta^{-1}(z) \\ &= \Theta^{-*}(z)\Theta^*(z)\Theta(z)N(z)\Theta^{-1}(z) \\ &= \Theta(z)N(z)\Theta^{-1}(z) =: \Psi(z), \end{aligned} \tag{2.24}$$

where $N(z) = A^*(z)B^{-R}(z) \in \mathbb{R}[z, z^{-1}]^{r\times r}$ is a L-unimodular matrix. By (2.24), $\Psi(z)$ is a spectral density. Actually a good deal more is true. We notice that $A(z)\tilde{\Delta}(z)B(z)$ and $B^*(z)\tilde{\Delta}(z)A^*(z)$ are two left-standard factorizations of $\Phi(z)$. Hence, by replacing $\Delta_1(z)$ with $\tilde{\Delta}(z) = \tilde{\Delta}^*(z)$ in (2.5), we obtain

$$\tilde{\Delta}(z)N(z)\tilde{\Delta}^{-1}(z) = M(z), \tag{2.25}$$

where $M(z) \in \mathbb{R}[z, z^{-1}]$ is L-unimodular. Since $\tilde{\Delta}(z) = \Theta^*(z)\Theta(z)$ is diagonal and

$$\Theta(z) := \operatorname{diag}[\theta_1(z), \ldots, \theta_r(z)]$$

is canonic, (2.25) implies that $[N(z)]_{ij}$ is divisible by the L-polynomial $[\tilde{\Delta}(z)]_{jj}/[\tilde{\Delta}(z)]_{ii}$, $j \geq i$. But

$$[\tilde{\Delta}(z)]_{ii} = \theta_i^*(z)\theta_i(z) = \theta_i(1/z)\theta_i(z) = \pm z^{k_i}\theta_i^2(z),$$

where $k_i \in \mathbb{Z}$, $i = 1, \ldots, r$. Hence, $[N(z)]_{ij}$ must be divisible by the polynomial

$$f_{ij}^2(z) := \frac{\theta_j^2(z)}{\theta_i^2(z)}, \quad j \geq i,$$

and, *a fortiori*, by

$$f_{ij}(z) = \frac{\theta_j(z)}{\theta_i(z)}, \quad j \geq i.$$

This suffices to establish that $\Psi(z)$ is L-polynomial. Actually, by (2.24), it follows that $\Psi(z)$ has determinant which is a real nonzero constant. Hence, $\Psi(z)$ is L-unimodular and positive definite on the unit circle. The problem is now reduced to that of finding a factorization of $\Psi(z)$ of the form

$$\Psi(z) = P^*(z)P(z), \tag{2.26}$$

where $P(z) \in \mathbb{R}[z]^{r\times r}$ is a unimodular polynomial matrix. After this is achieved, the desired factorization for $\Phi(z)$ is obtained as $\Phi(z) = W^*(z)W(z)$ with

$$\begin{aligned} W(z) :&= P(z)\Theta(z)B(z) \\ &= P(z)\Theta(z)\Lambda(z)F(z) \\ &= P(z)D_+(z)F(z), \end{aligned} \tag{2.27}$$

where we have defined $D_+(z) := \Theta(z)\Lambda(z)$. Indeed, by straightforward algebra,

$$\begin{aligned} W^*(z)W(z) &= B^*(z)\Theta^*(z)P^*(z)P(z)\Theta(z)B(z) \\ &= B^*(z)\tilde{\Delta}(z)N(z)B(z) \\ &= B^*(z)\tilde{\Delta}(z)A^*(z) \\ &= \Phi^*(z) = \Phi(z). \end{aligned}$$

Step 4. We illustrate an algorithm which provides a factorization of L-unimodular spectral density $\Psi(z) = \Psi^*(z) \in \mathscr{S}^r_{\text{rat}}(\mathbb{T})$ positive definite on the unit circle into the product $P^*(z)P(z)$, where $P(z)$ is a unimodular polynomial matrix.

The algorithm can be thought of as a procedure that parallels the Oono–Yasuura algorithm Oono and Yasuura [1954], which constitutes a fundamental step in Youla's work Youla [1961], and consists of the following two steps. First of all, we define $\Psi_1(z) := \Psi(z)$ and denote by $h \in \mathbb{N}$ the loop counter of the algorithm, which is initially set to $h \leftarrow 1$.

1) Let $K_i \in \mathbb{Z}$, $i = 1, \dots, r$, be the maximum-degree of the i-th column of $\Psi_h(z)$ and $k_i \in \mathbb{Z}$, $i = 1, \dots, r$, be the minimum-degree of the i-th row of $\Psi_h(z)$. Consider the highest-column-degree coefficient matrix of $\Psi_h(z)$, denoted by Ψ_h^{hc}, and the lowest-row-degree coefficient matrix of $\Psi_h(z)$, denoted by Ψ_h^{lr}. As noticed in the proof of Lemma 2.8, the positive nature of $\Psi_h(z)$ implies that $K_i \geq 0$ for all $i = 1, \dots, r$. Moreover, the para-Hermitianity of $\Psi_h(z)$ implies that $\Psi_h^{\text{hc}} = (\Psi_h^{\text{lr}})^\top$ which, in turn, yields $K_i = -k_i$ for all $i = 1, \dots, r$.

   By Lemma 2.8, it follows that Ψ_h^{hc} is nonsingular if and only if $\Psi_h(z)$ is a constant matrix. If $\Psi_h(z)$ is a constant matrix, we set $\bar{h} := h$ and go to step 2). If this is not the case, we calculate a nonzero vector $\mathbf{v}_h := [v_1 \; v_2 \; \dots \; v_r]^\top \in \mathbb{R}^r$ such that $\Psi_h^{\text{hc}} \mathbf{v}_h = \mathbf{0}$. Let us define the *active index set*

$$\mathscr{I}_h := \{ i : v_i \neq 0 \}$$

   and the *highest maximum-degree active index set*, $\mathscr{M}_h \subset \mathscr{I}_h$,

$$\mathscr{M}_h := \{ i \in \mathscr{I}_h : \; K_i \geq K_j, \; \forall \, j \in \mathscr{I}_h \}.$$

   We pick an index $p \in \mathscr{M}_h$. Then, we define the polynomial matrix

$$\Omega_h^{-1}(z) := \begin{bmatrix} 1 & \cdots & 0 & \overset{\text{column } p}{\overset{\downarrow}{\frac{v_1}{v_p} z^{K_p - K_1}}} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots & & & 0 \\ \vdots & & 1 & \frac{v_{p-1}}{v_p} z^{K_p - K_{p-1}} & & & \vdots \\ \vdots & & & 1 & & & \vdots \\ \vdots & & & \frac{v_{p+1}}{v_p} z^{K_p - K_{p+1}} & 1 & & \vdots \\ 0 & & & \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \frac{v_r}{v_p} z^{K_p - K_r} & 0 & \cdots & 1 \end{bmatrix}. \tag{2.28}$$

Notice that the entry at (i,p) of $\Omega_h^{-1}(z)$ has the form

$$\frac{v_i}{v_p} z^{K_p - K_i} = \alpha_i z^{\delta_i}, \quad i = 1, \ldots, r, \tag{2.29}$$

with $\alpha_i := v_i / v_p \in \mathbb{R}$ and $\delta_i := K_p - K_i \geq 0$. In fact, if $K_i > K_p$, then $v_i = 0$ and so $\alpha_i = 0$. By (2.28), $\det \Omega_h^{-1}(z) = 1$ and, therefore, $\Omega_h^{-1}(z) \in \mathbb{R}[z]^{r \times r}$ is a unimodular polynomial matrix. By operating the transformation

$$\Psi_{h+1}(z) := \Omega_h^{-*}(z) \Psi_h(z) \Omega_h^{-1}(z),$$

we obtain a new positive definite matrix $\Psi_{h+1}(z)$ with the same determinant of $\Psi_h(z)$. Furthermore, the maximum-degree of the p-th column of $\Psi_{h+1}(z)$ is lower than K_p, while the maximum-degree of the i-th column, $i \neq p$, is not greater than K_i.

This fact needs a detailed explanation. If we post-multiply $\Psi_h(z)$ by $\Omega_h^{-1}(z)$, we obtain a matrix of the form

$$\begin{aligned} \Psi'_h(z) &:= \Psi_h(z) \Omega_h^{-1}(z) \\ &= \left[\begin{array}{c|c|c} [\Psi_h(z)]_{1:r,1:p-1} & \boldsymbol{\psi}_h(z) & [\Psi_h(z)]_{1:r,p+1:r} \end{array} \right], \end{aligned}$$

where all the L-polynomials in the p-th column vector

$$\boldsymbol{\psi}_h(z) = [\Psi_h(z)]_{1:r,p:p} + \sum_{i \neq p} \alpha_i z^{\delta_i} [\Psi_h(z)]_{1:r,i:i} \tag{2.30}$$

have maximum-degree lower than K_p, since $\Psi_h^{\mathrm{hc}} \mathbf{v}_h = \mathbf{0}$, and minimum-degree which satisfies

$$\min \deg [\boldsymbol{\psi}_h(z)]_i \geq k_i = -K_i, \quad i = 1, \ldots, r, \tag{2.31}$$

since in (2.30) $\delta_i \geq 0$, for all i such that $\alpha_i \neq 0$. Now, by pre-multiplying $\Psi'_h(z)$ by $\Omega_h^{-*}(z)$, the resulting matrix $\Psi_{h+1}(z)$ can be written in the form

$$\begin{aligned} \Psi_{h+1}(z) &= \Omega_h^{-*}(z) \Psi_h(z) \Omega_h^{-1}(z) \\ &= \left[\begin{array}{c|c|c} [\Psi_h(z)]_{1:p-1,1:p-1} & \boldsymbol{\psi}'_{h+1}(z) & [\Psi_h(z)]_{1:p-1,p+1:r} \\ \hline \boldsymbol{\psi}'^{\top}_{h+1}(z^{-1}) & \psi''_{h+1}(z) & \boldsymbol{\psi}'''^{\top}_{h+1}(z^{-1}) \\ \hline [\Psi_h(z)]_{p+1:r,1:p-1} & \boldsymbol{\psi}'''_{h+1}(z) & [\Psi_h(z)]_{p+1:r,p+1:r} \end{array} \right], \end{aligned}$$

where the p-th column vector

$$\left[\begin{array}{c|c|c} \boldsymbol{\psi}'^{\top}_{h+1}(z) & \psi''_{h+1}(z) & \boldsymbol{\psi}'''^{\top}_{h+1}(z) \end{array} \right]^{\top}$$

differs from $\boldsymbol{\psi}_h(z)$ only for the value of the p-th entry $\psi''_{h+1}(z)$. Moreover, the maximum-degree of $\psi''_{h+1}(z)$ cannot increase after the operation is performed, since

$$\psi''_{h+1}(z) = [\boldsymbol{\psi}_h(z)]_p + \sum_{i \neq p} \alpha_i z^{-\delta_i} [\boldsymbol{\psi}_h(z)]_i,$$

and, by (2.29), $\delta_i \geq 0$, for all i such that $\alpha_i \neq 0$. We conclude that all the L-polynomials in the p-th column of $\Psi_{h+1}(z)$ have maximum-degree lower than K_p, while, by (2.31), the maximum-degree of all the other columns does not increase. We notice also that, since $\Psi_{h+1}(z) = \Psi^*_{h+1}(z)$, all the L-polynomials in the p-th row of $\Psi_{h+1}(z)$ have minimum-degree greater than $k_p = -K_p$, while the minimum-degree of all the other rows does not decrease. Eventually, we update the value of the loop counter h by setting $h \leftarrow h+1$ and return to step 1).

2) Since $\Psi_{\bar{h}} \in \mathbb{R}^{r \times r}$ is positive definite, we can always factorize it into the product $\Psi_{\bar{h}} = C^\top C$ where $C \in \mathbb{R}^{r \times r}$, by using standard techniques such as the Cholesky decomposition (see, *e.g.*, [Golub and Van Loan, 1996, Ch. 4]). Finally, we have constructed a polynomial unimodular matrix

$$P(z) = C\Omega_{\bar{h}-1}(z)\Omega_{\bar{h}-2}(z)\cdots\Omega_1(z).$$

such that $\Psi(z) = P^*(z)P(z)$.

It is worthwhile noticing that the iterative procedure of step 1) is always brought to an end (after a maximum of $K_1 + \cdots + K_p$ iterations) since at the h-th iteration the maximum-degree of a column of $\Psi_h(z)$ is reduced at least by one, while the maximum-degree of all the other columns does not increase.

To complete the proof of statements 1) and 2), we notice that, by construction, the rational matrix $W(z)$, as defined in (2.27), and its right inverse are analytic in $\{z \in \mathbb{C} : |z| > 1\}$. Moreover, we recall that, if $z = 0$ is either a pole, zero or both of $D(z)$, $D_+(z)$ and $D(z)$ have the same zero-pole structure at $z = 0$. Now, suppose, by contradiction, that $W(z)$ has a pole at $z = \infty$. Then $W^*(z)$ has a pole at $z = 0$. But, since $\Phi(z) = W^*(z)W(z)$, it follows that

$$\begin{aligned} W^*(z) &= \Phi(z)W^{-R}(z) \\ &= C(z)D(z)F(z)F^{-R}(z)D_+^{-1}(z)P^{-1}(z) \\ &= C(z)D(z)D_+^{-1}(z)P^{-1}(z) \\ &= C(z)D_-(z)P^{-1}(z), \end{aligned} \tag{2.32}$$

where $D_-(z) := D(z)D_+^{-1}(z)$ has no pole at $z = 0$. Since $P^{-1}(z)$ and $C(z)$ are unimodular matrices, in view of (2.32), also $W^*(z)$ has no pole at $z = 0$. Hence, the contradiction. We conclude that $W(z)$ has no pole at infinity. Finally, by following a similar argument, it can be verified that also $W^{-R}(z)$ has no pole at infinity.

Now consider statement 5). If $\Phi(z)$ is analytic on the unit circle, then $\Theta(z)$ does not possess any finite pole. This, in turn, implies that $D_+(z) = \Theta(z)\Lambda(z)$ is analytic in $\{z \in \overline{\mathbb{C}} : |z| \geq 1\}$. Thus, $W(z)$, as defined in (2.27), is also analytic in the same region.

As for point 6), the additional assumption that the rank of $\Phi(z)$ is constant on the unit circle implies that $\Theta(z)$ does not possess any finite zero. Thus, $\Theta(z) = I_r$ and, by (2.27),

$$W^{-R}(z) = F^{-R}(z)\Lambda^{-1}(z)P^{-1}(z)$$

is analytic in $\{z \in \overline{\mathbb{C}} : |z| \geq 1\}$. Hence, $W(z)$ and its right inverse $W^{-R}(z)$ are both analytic in $\{z \in \overline{\mathbb{C}} : |z| \geq 1\}$.

Lastly, consider point 7). As shown in (2.23), the Smith-McMillan canonical form of $\Phi(z)$, $D(z)$, is connected to that of $W(z)$, $D_+(z) = \Theta(z)\Lambda(z)$, by

$$D(z) = \Sigma(z)D_+^*(z)D_+(z), \tag{2.33}$$

where $\Sigma(z) \in \mathbb{R}(z)^{r\times r}$ is a diagonal matrix with elements $\alpha_i z^{k_i}$, $\alpha_i \in \mathbb{R}_0$, $k_i \in \mathbb{Z}$, on its diagonal. Let $p_1, \ldots, p_h$ be the nonzero finite poles of $\Phi(z)$. By (2.33), it follows that

$$\delta_p(\Phi; p_i) = \begin{cases} \delta_p(W; p_i) & \text{if } |p_i| < 1, \\ 2\delta_p(W; p_i) & \text{if } |p_i| = 1, \\ \delta_p(W; 1/p_i) & \text{if } |p_i| > 1. \end{cases} \tag{2.34}$$

Moreover, if $p \in \overline{\mathbb{C}}$ is a pole of $\Phi(z)$ of degree $\delta_p(\Phi; p)$ then also $1/p$ is a pole of $\Phi(z)$ of the same degree and if $p \in \overline{\mathbb{C}}$ is not a pole of $\Phi(z)$ then neither p nor $1/p$ are poles of $W(z)$. Thus, we have

$$\begin{aligned} \sum_{i=1}^{h} \delta_p(\Phi; p_i) &= \sum_{i:|p_i|<1} \delta_p(W; p_i) + \sum_{i:|p_i|>1} \delta_p(W; 1/p_i) + \sum_{i:|p_i|=1} 2\delta_p(W; p_i) \\ &= 2 \sum_{i:|p_i|\leq 1} \delta_p(W; p_i) \end{aligned} \tag{2.35}$$

By (2.4), the McMillan degree of a rational matrix equals the sum of the degrees of all its poles, including the pole at infinity. If $\Phi(z)$ has no pole at infinity, then (2.35) directly yields $\delta_M(\Phi) = 2\delta_M(W)$. Otherwise, assume that $\Phi(z)$ has a pole at infinity. Since $W(z)$ and $\Phi(z)$ have the same structural indices at $z = 0$ and $W(z)$ has no pole at $z = \infty$, it follows that

$$\delta_p(\Phi; \infty) = \delta_p(\Phi; 0) = \delta_p(W; 0) \quad \text{and} \quad \delta_p(W; \infty) = 0. \tag{2.36}$$

Therefore, by equations (2.35) and (2.36),

$$\begin{aligned} \delta_M(\Phi) &= \sum_{i=1}^{h} \delta_p(\Phi; p_i) + \delta_p(\Phi; 0) + \delta_p(\Phi; \infty) \\ &= 2 \sum_{i:|p_i|\leq 1} \delta_p(W; p_i) + 2\delta_p(W; 0) = 2\delta_M(W). \end{aligned}$$

This completes the proof. □

We are now ready to prove our main Theorem 2.1. Many of the ideas for this proof can be elaborated from those of the proof of Theorem 2.2.

Proof of Theorem 2.1. We first show how to modify the constructive procedure used in the proof of Theorem 2.2 in order to obtain a spectral factor $W(z)$ which satisfies points 1) and 2). With reference to step 2 in the proof of Theorem 2.2, we rearrange the Smith–McMillan form of $\Phi(z)$ as

$$D(z) = \Sigma(z)\Lambda(z)\tilde{\Delta}(z)\Lambda(z), \tag{2.37}$$

where the only difference with respect to the decomposition in (2.23) is that here $\Lambda(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal, canonic and analytic in $\mathscr{A}_p \setminus \{\infty\}$ with $\Lambda^{-1}(z)$ analytic in $\mathscr{A}_z \setminus \{\infty\}$. Moreover, if $0 \notin \mathscr{A}_p$ and $z = 0$ is a pole of $D(z)$, then $\Lambda(z)$ has the same negative structural indices at $z = 0$ of $\Phi(z)$, and if $0 \notin \mathscr{A}_z$ and $z = 0$ is a zero of $D(z)$, then $\Lambda(z)$ has the same positive structural indices at $z = 0$ of $\Phi(z)$.

Now, to apply the procedure described in the proof of Theorem 2.2, it suffices to prove that for any choice of the unmixed-symplectic sets $\mathscr{A}_p$ and $\mathscr{A}_z$, the spectral density $\Psi(z)$, as defined in (2.24), is still L-unimodular. With reference to the notation introduced in the proof of Theorem 2.2, $\Psi(z)$ can be written as

$$\begin{aligned}\Psi(z) &= \Theta(z)N(z)\Theta^{-1}(z)\\ &= \Theta(z)A^*(z)B^{-R}(z)\Theta^{-1}(z)\\ &= \Theta(z)\Lambda(z)\Sigma^*(z)C^*(z)F^{-R}(z)\Lambda^{-1}(z)\Theta^{-1}(z)\\ &= \Sigma^*(z)D_+(z)\Xi(z)D_+^{-1}(z),\end{aligned} \tag{2.38}$$

where we have defined $\Xi(z) := C^*(z)F^{-R}(z) \in \mathbb{R}[z, z^{-1}]^{r\times r}$ which is L-unimodular and whose structure does not depend upon the choice of $\mathscr{A}_p$ and $\mathscr{A}_z$. Moreover, in this case, $D_+(z) = \Theta(z)\Lambda(z)$ is diagonal, canonic and analytic in $\mathscr{A}_p \setminus \{\infty\}$ with inverse analytic in $\mathscr{A}_z \setminus \{\infty\}$. Let us first consider the standard choice $\mathscr{A}_p = \mathscr{A}_z = \{z \in \overline{\mathbb{C}} : |z| > 1\}$. In the proof of Theorem 2.2, we have shown that $\Psi(z)$ is L-unimodular. Since $D_+(z)$ is diagonal and canonic and $\Sigma^*(z)$ is L-unimodular, by (2.38), it follows that $[\Xi(z)]_{ij} \in \mathbb{R}[z, z^{-1}]$ must be divisible[12] by the polynomial

$$p_{ij}(z) := \frac{[D_+(z)]_{jj}}{[D_+(z)]_{ii}}, \quad j \geq i.$$

[12] Here the concept of divisibility is the one associated to the ring of L-polynomials.

On the other hand, let us consider the opposite choice $\mathscr{A}_p = \mathscr{A}_z = \{z \in \mathbb{C} : |z| < 1\}$. By using the right-standard counterpart of Lemma 2.4 and by following *verbatim* the argument used in step 3 of Theorem 2.2, it can be proven that $\Psi(z)$ is still L-unimodular. Hence, by (2.38), $[\Xi(s)]_{ij}$ must be also divisible by the L-polynomial $p_{ij}(z^{-1})$, $j \geq i$. Therefore, $[\Xi(z)]_{ij}$ must be divisible by the L-polynomial

$$q_{ij}(z) := p_{ij}(z)p_{ij}(z^{-1}), \quad j \geq i.$$

Since, for any choice of the unmixed-symplectic sets $\mathscr{A}_p$ and $\mathscr{A}_z$, the factors of $[D_+(z)]_{jj}[D_+(z)]_{ii}^{-1}$, $j \geq i$, are contained in the ones of $q_{ij}(z)$, then $[\Xi(z)]_{ij}$ must be divisible by the polynomial $[D_+(z)]_{jj}[D_+(z)]_{ii}^{-1}$, $j \geq i$, for any choice of $\mathscr{A}_p$ and $\mathscr{A}_z$. We conclude that $\Psi(z)$ must be a L-polynomial matrix for any choice of $\mathscr{A}_p$ and $\mathscr{A}_z$. But, since $\Psi(z)$ is a spectral density, $\det\Psi(z)$ is a real constant, hence $\Psi(z)$ is L-unimodular.

To prove point 3) we need to show that the McMillan degree of the spectral factor $W(z)$ just obtained equals one half of the McMillan degree of $\Phi(z)$. To this aim, we can follow the same lines of the proof of point 7) of Theorem 2.2. In fact, we can define $\mathscr{A}_{p,1} := \mathscr{A}_p \setminus (\mathbb{T} \cup \{0, \infty\})$ and partition $\mathbb{C}_0$ as

$$\mathbb{C}_0 = \{z \in \mathbb{C} : 1/z \in \mathscr{A}_{p,1}\} \cup \mathbb{T} \cup \mathscr{A}_{p,1}$$

and replace equation (2.34) with the more general expression for the degree of the pole p_i of $\Phi(z)$

$$\delta_p(\Phi; p_i) = \begin{cases} \delta_p(W; p_i) & \text{if } 1/p_i \in \mathscr{A}_{p,1}, \\ 2\delta_p(W; p_i) & \text{if } |p_i| = 1, \\ \delta_p(W; 1/p_i) & \text{if } p_i \in \mathscr{A}_{p,1}. \end{cases}$$

The rest of the proof remains the same.

The proof of point 4) is very similar to that of point 3) of Theorem 2.2. The only difference is that the all-pass matrix function $V(z) := W_1(z)W^{-R}(z)$ and its inverse are not analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$ but they are analytic in $\mathscr{A}_p$, so that Lemma 2.2 still applies.

As for point 5), we define $V(z) := L(z)W^{-R}(z)$ which is clearly all-pass and analytic in $\mathscr{A}_z$, and the same computation that led to (2.21), gives $L(z) = V(z)W(z)$. On the other hand, if $V(z)$ is all-pass and analytic in $\mathscr{A}_p$, then it is immediate to check that $L(z) := V(z)W(z)$ is a spectral factor of $\Phi(z)$ and is analytic in $\mathscr{A}_p$ as well.

The proof of points 6) and 7) is exactly the same as that of points 5) and 6) of Theorem 2.2. □

2.4.1 Corollaries

To conclude this section, we present two straightforward yet interesting corollaries of Theorem 2.2. The first is a complete parametrization of the set of all spectral factors of a given spectrum.

Corollary 2.1. *Let $\Phi(z)$ be a given spectrum and $W(z)$ be any spectral factor satisfying conditions 1) and 2) of Theorem 2.2. Let $L(z) \in \mathbb{R}(z)^{m\times n}$, then $\Phi(z) = L^*(z)L(z)$ if and only if*

$$L(z) = V(z)\begin{bmatrix} I_r \\ \hline \mathbf{0}_{m-r,r} \end{bmatrix} W(z),$$

where $V(z) \in \mathbb{R}(z)^{m\times m}$ is an arbitrary all-pass matrix and $r = \mathrm{rk}(\Phi)$.

Proof. By repeating an argument used in points 3) and 4) of Theorem 2.2, we have that $L(z) = U(z)W(z)$, with $U(z) \in \mathbb{R}(z)^{m\times r}$ a rational matrix satisfying $U^*(z)U(z) = I_r$. If we choose $V(z) \in \mathbb{R}(z)^{m\times m}$ to be any all-pass matrix with $U(z)$ incorporated into its first r columns, *i.e.*,

$$U(z) = V(z)\left[\begin{array}{c} I_r \\ \hline \mathbf{0}_{m-r,r} \end{array}\right],$$

we conclude. □

The next result characterizes the spectral factors of L-polynomial spectra.

Corollary 2.2. *Let $\Phi(z)$ be a spectrum and $W(z)$ be the spectral factor provided in the (constructive) proof of Theorem 2.1. Assume that $\Phi(z)$ is L-polynomial. If $\infty \in \mathscr{A}_p$, then $W(z)$ is polynomial in z^{-1} (so that $W^*(z)$ is polynomial in z). Otherwise, $0 \in \mathscr{A}_p$ and $W(z)$ is polynomial in z (so that $W^*(z)$ is polynomial in z^{-1}).*

Proof. We consider only the case of $\infty \in \mathscr{A}_p$, the other being similar. If $\Phi(z)$ is L-polynomial, then the only finite pole it may possess is located at $z = 0$. Since $W(z)$ does not have the pole at infinity, $W(z)$ must be polynomial in z^{-1}. The latter fact, in turn, implies that $W^*(z)$ must be a polynomial matrix. □

2.5 A numerical example

In this section, we will illustrate an application to stochastic realization of the algorithm used in the constructive proof of Theorem 2.1. To this aim, let us consider a purely non-deterministic, second-order process $\{y(t)\}_{t\in\mathbb{Z}}$ whose spectral density is

$$\Phi(z) = \begin{bmatrix} \frac{-2z+6-2z^{-1}}{-2z+5-2z^{-1}} & z-1 & z-1 \\ z^{-1}-1 & -z+2-z^{-1} & -z+2-z^{-1} \\ z^{-1}-1 & -z+2-z^{-1} & -z+2-z^{-1} \end{bmatrix}.$$

We want to compute a stochastically minimal, anti-causal realization of $\{y(t)\}_{t\in\mathbb{Z}}$ having all its zeros in the (closed) unit disk. Since our method has been developed to compute a spectral factorization in the form $\Phi(z) = W^\top(z^{-1})W(z)$, this requirement corresponds to the choice $\mathscr{A}_z := \{z \in \mathbb{C} : |z| < 1\}$ and $\mathscr{A}_p := \{z \in \mathbb{C} : |z| > 1\}$. Notice that $\Phi(z)$ is not proper, it features a zero on the unit circle and is (normal) rank deficient, namely $\mathrm{rk}(\Phi) = 2$.

We now apply step-by-step the proposed factorization algorithm in order to compute a spectral factor $W(z) \in \mathbb{R}(z)^{2\times 3}$ analytic in $\mathscr{A}_p$ with right inverse analytic in $\mathscr{A}_z$.

Step 1. The Smith–McMillan canonical form of $\Phi(z)$ is given by

$$D(z) = \begin{bmatrix} \frac{1}{z(z-2)(z-1/2)} & 0 \\ 0 & z(z-1)^2 \end{bmatrix},$$

$\Phi(z)$ can be decomposed as

$$\Phi(z) = C(z)D(z)F(z),$$

where $C(z) \in \mathbb{R}[z]^{3\times 2}$ and $F(z) \in \mathbb{R}[z]^{2\times 3}$ are unimodular matrices.

Step 2. The matrices $\Lambda(z)$, $\Theta(z)$ and $\Sigma(z)$ defined in (2.37) have the form

$$\Lambda(z) = \begin{bmatrix} \frac{1}{z(z-\frac{1}{2})} & 0 \\ 0 & 1 \end{bmatrix}, \ \Theta(z) = \begin{bmatrix} 1 & 0 \\ 0 & z-1 \end{bmatrix}, \ \Sigma(z) = \begin{bmatrix} -\frac{1}{2z^2} & 0 \\ 0 & -z^2 \end{bmatrix}.$$

Note that $\Lambda(z)$ is analytic in $\mathscr{A}_p \setminus \{\infty\}$ with inverse analytic in $\mathscr{A}_z$. Let $A(z) = C(z)\Sigma(z)\Lambda^*(z)$, $B(z) = \Lambda(z)F(z)$.

Step 3. The matrix $\Psi(z) = \Theta(z)^{-1}N(z)\Theta(z)$, with $N(z) = A^*(z)B^{-R}(z)$, is given by

$$\begin{aligned}\Psi(z) &= \Theta(z)^{-1}N(z)\Theta(z) \\ &= \begin{bmatrix} -\frac{1}{2}z+\frac{3}{2}-\frac{1}{2}z^{-1} & -\frac{9}{4}z^3+\frac{25}{2}z^2-\frac{43}{2}z+\frac{43}{4}+\frac{1}{2}z^{-1} \\ \frac{1}{2}z+\frac{43}{4}-\frac{43}{2}z^{-1}+\frac{25}{2}z^{-2}-\frac{9}{4}z^{-3} & \psi_{22}(z) \end{bmatrix}.\end{aligned}$$

where $\psi_{22}(z) := \frac{9}{4}z^3 + \frac{341}{8}z^2 - \frac{1747}{8}z + \frac{2780}{8} - \frac{1747}{8}z^{-1} + \frac{341}{8}z^{-2} + \frac{9}{4}z^{-3}$. It is worth noting that $\Psi(z)$ is para-Hermitian, L-unimodular and positive definite upon the unit circle.

Step 4. Let $\Psi_1(z) := \Psi(z)$. The highest-column-degree coefficient matrix of $\Psi_1(z)$ is

$$\Psi_1^{\mathrm{hc}} = \begin{bmatrix} -\frac{1}{2} & -\frac{9}{4} \\ \frac{1}{2} & \frac{9}{4} \end{bmatrix}.$$

Since Ψ_1^{hc} is singular, we calculate a nonzero vector $\mathbf{v}_1 \in \ker \Psi_1^{\mathrm{hc}}$. One such a vector is given, for instance, by $\mathbf{v}_1 = [9, \ -2]^\top$. The highest maximum-degree active index set is $\mathscr{M}_1 = \{2\}$, we construct the unimodular matrix $\Omega_1^{-1}(z)$ of the form (2.28)

$$\Omega_1^{-1}(z) = \begin{bmatrix} 1 & -\frac{9}{2}z^2 \\ 0 & 1 \end{bmatrix}$$

in order to reduce the maximum degree of the second column of $\Psi_1(z)$,

$$\begin{aligned}\Psi_2(z) &= \Omega_1^{-*}(z)\Psi_1(z)\Omega_1^{-1}(z) \\ &= \begin{bmatrix} -\frac{1}{2}z+\frac{3}{2}-\frac{1}{2}z^{-1} & \frac{23}{4}z^2-\frac{77}{4}z+\frac{43}{4}+\frac{1}{2}z^{-1} \\ \frac{1}{2}z+\frac{43}{4}-\frac{77}{4}z^{-1}+\frac{23}{4}z^{-2} & -\frac{23}{4}z^2-\frac{973}{4}z+\frac{2123}{4}-\frac{973}{4}z^{-1}-\frac{23}{4}z^{-2} \end{bmatrix}.\end{aligned}$$

Since Ψ_2^{hc} is singular, we repeat the previous step. In this case, we have $\mathbf{v}_2 = [23, \ 2]^\top \in \ker \Psi_2^{\mathrm{hc}}$, $\mathscr{M}_2 = \{2\}$, and

$$\Omega_2^{-1}(z) = \begin{bmatrix} 1 & \frac{23}{2}z \\ 0 & 1 \end{bmatrix}.$$

Hence, we compute the reduced matrix

$$\begin{aligned}\Psi_3(z) &= \Omega_2^{-*}(z)\Psi_2(z)\Omega_2^{-1}(z)\\ &= \begin{bmatrix} -\frac{1}{2}z+\frac{3}{2}-\frac{1}{2}z^{-1} & -2z+5+\frac{1}{2}z^{-1} \\ \frac{1}{2}z+5-2z^{-1} & 2z+21+2z^{-1} \end{bmatrix}.\end{aligned}$$

Actually, Ψ_3^{hc} is singular. In this case, $\mathbf{v}_3 = [-4,\ 1]^\top \in \ker \Psi_3^{\mathrm{hc}}$, $\mathscr{M}_3 = \{2\}$,

$$\Omega_3^{-1}(z) = \begin{bmatrix} 1 & -4 \\ 0 & 1 \end{bmatrix}$$

and we obtain

$$\begin{aligned}\Psi_4(z) &= \Omega_3^{-*}(z)\Psi_3(z)\Omega_3^{-1}(z)\\ &= \begin{bmatrix} -\frac{1}{2}z+\frac{3}{2}-\frac{1}{2}z^{-1} & -1+\frac{5}{2}z^{-1} \\ \frac{5}{2}z-1 & 5 \end{bmatrix}.\end{aligned}$$

Yet another iteration is required; indeed Ψ_4^{hc} is singular. Thus, we proceed by computing $\mathbf{v}_4 = [-2\ 1]^\top \in \ker \Psi_4^{\mathrm{hc}}$, $\mathscr{M}_3 = \{1\}$,

$$\Omega_4^{-1}(z) = \begin{bmatrix} 1 & 0 \\ -\frac{1}{2}z & 1 \end{bmatrix}$$

and eventually we arrive at

$$\Psi_5 = \Omega_4^{-*}(z)\Psi_4(z)\Omega_4^{-1}(z) = \begin{bmatrix} \frac{1}{4} & -1 \\ -1 & 5 \end{bmatrix}.$$

The latter matrix is constant and positive definite; therefore it admits a Cholesky factorization

$$\Psi_5 = C^\top C, \quad C = \begin{bmatrix} \frac{1}{2} & -2 \\ 0 & 1 \end{bmatrix}.$$

The fourth step of the algorithm is concluded, since we found a factorization $\Psi(z) = P^*(z)P(z)$, with $P(z)$ unimodular of the form

$$\begin{aligned}P(z) &= C\Omega_4(z)\Omega_3(z)\Omega_2(z)\Omega_1(z)\\ &= \begin{bmatrix} -z+\frac{1}{2} & -\frac{1}{4}z\left(18z^2-55z+39\right) \\ \frac{1}{2}z & \frac{1}{4}\left(9z^3-23z^2+8z+4\right) \end{bmatrix}.\end{aligned}$$

Finally, we have that

$$W(z) = P(z)\Theta(z)B(z) = \begin{bmatrix} -\frac{1}{z} & \frac{1}{z}-1 & \frac{1}{z}-1 \\ \frac{1}{2z-1} & 0 & 0 \end{bmatrix}.$$

is a stochastically minimal spectral factor of $\Phi(z)$ analytic in $\mathscr{A}_p$ with right inverse analytic in $\mathscr{A}_z$.

Therefore the sought-for realization symbolically is given by

$$y(t) = W^*(z)e(t)$$

with $e(t)$ being white noise.

3. On the uniqueness of spectral factors

From the general discrete-time spectral factorization result of the previous chapter, a very interesting question arises. In fact, when the chosen unmixed-symplectic regions are such that $\mathscr{A}_p \equiv \mathscr{A}_z \equiv \mathscr{A}$, we have seen that the corresponding spectral factor is essentially unique (namely, unique up to multiplication on the left side by a constant orthogonal matrix), as described in the proof of Theorem 2.1, point 4): The key idea here is that, starting from a reference spectral factor $W(z)$, a second spectral factor $W_1(z)$ must be of the form $W_1(z) = Q(z)W(z)$ with $Q(z)$ being all-pass so that if $Q(z)$ has a pole in p, it necessarily has a zero in $1/p$; therefore, for any non-constant $Q(z)$, either $W_1(z)$ or its right-inverse is no longer analytic in $\mathscr{A}$. On the contrary, when $\mathscr{A}_p \neq \mathscr{A}_z$, we can easily obtain a spectral factor $W_1(z)$ with the prescribed analyticity properties by selecting an all-pass function $Q(z)$ featuring poles in $\overline{\mathbb{C}} \setminus \mathscr{A}_p$ and zeros in $\overline{\mathbb{C}} \setminus \mathscr{A}_z$. Thus, there appears to be an inherent ambiguity in the choice of the spectral factor in this general case. In this chapter, we show that this is in not the case if we further impose that the spectral factor has minimal complexity, as measured by its McMillan degree. In fact, we will show that, under this assumption, for any choice of the analyticity regions $\mathscr{A}_p$ and $\mathscr{A}_z$, the spectral factor is essentially unique.

The present chapter, which is adapted from Baggio and Ferrante [2016b], is divided in two main sections. In the first one, Section 3.1, we review an elegant and profound result on the parametrization of rational all-pass functions established in Alpay and Gohberg [1988] which will play a key role in the proof of our main result. Further, in the same section, we will present other ancillary results on the cancellation of zeros and poles in the product of rational matrices. The second section, Section 3.2, contains the proof of essential uniqueness of minimal spectral factors, and, therefore, represents the core of the chapter.

3.1 Ancillary results

The problem of cancelling part of the poles or zeros of a rational matrix-valued function via multiplication with another one (especially in the case in which such a function is all-pass) has been deeply investigated in circuit and systems theory. This problem has been originally addressed in Belevitch [1968] and more elaborately in Dewilde and Vandewalle [1975], Vandewalle and Dewilde [1977], Van Dooren [1990]. For some more recent advances on this topic we refer to Dym and Nevo [2005a,b], Oară and Andrei [2009], Oară and Sabău [2009]. In this section, we will present some results in this direction, which are instrumental for the proof of the main theorem.

We first explicitly state what we mean for degree of a pole and zero of a rational matrix-valued function. Let $G(z) \in \mathbb{C}(z)^{m\times n}$ and consider the Smith–McMillan form of $G(z)$, say

$$D(z) = \operatorname{diag}\left[\frac{\varepsilon_1(z)}{\psi_1(z)}, \frac{\varepsilon_2(z)}{\psi_2(z)}, \ldots, \frac{\varepsilon_r(z)}{\psi_r(z)}\right].$$

As recalled in the previous chapter, the (finite) zeros of $G(z)$ coincide with the roots of the numerator of the $\varepsilon_i(z)$'s, whereas the (finite) poles of $G(z)$ with the zeros of the roots of the $\psi(z)$'s. The degree of a pole and zero at $\alpha \in \mathbb{C}$ (denoted by $\delta_p(G;\alpha)$ and $\delta_z(G;\alpha)$, respectively) is equal to the sum of the degrees of the zero at α of all the $\psi_i(z)$ and of all the $\varepsilon_i(z)$, respectively.[13] If $G(z)$ has no pole (zero) at α, we let $\delta_p(G;\alpha) = 0$ ($\delta_z(G;\alpha) = 0$, respectively).

[13] If $\alpha = \infty$, then we can consider the mapping $z \mapsto \lambda^{-1}$ and the definition still applies by considering the degree of the pole/zero at $\lambda = 0$ of $G(\lambda)$.

Following Anderson and Gevers [1981], we now give a precise definition of pole/zero cancellation in the product of rational matrix-valued function.

Definition 3.1. Let $G(z) \in \mathbb{C}(z)^{m\times n}$, $H(z) \in \mathbb{C}(z)^{n\times p}$ and $\alpha \in \overline{\mathbb{C}}$. We say that in the product $G(z)H(z)$ there is:

1) a pole cancellation at α if $\delta_p(GH;\alpha) < \delta_p(G;\alpha) + \delta_p(H;\alpha)$;

2) a zero cancellation at α if $\delta_z(GH;\alpha) < \delta_z(G;\alpha) + \delta_z(H;\alpha)$;

3) a zero-pole cancellation at α if both conditions 1) and 2) are met.

Remark 3.1. If $\operatorname{rk}(G) = \operatorname{rk}(H) = n$ then a zero or pole cancellation at α in the product $G(z)H(z)$ always corresponds to a zero-pole cancellation at α. A proof of this fact is given below (Lemma 3.1).

However, in general, 1) and 2) are not equivalent. Indeed, consider for instance the product

$$G(z)H(z) = \begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}\frac{2z+3}{(z+1)(z+2)} \\ \frac{1}{z+2}\end{bmatrix} = \frac{1}{z+1}$$

and observe that there is a pole cancellation at -2 which does not correspond to a zero-pole cancellation at -2.

Lemma 3.1. *Let $G(z) \in \mathbb{C}(z)^{n\times r}$ and $H(z) \in \mathbb{C}(z)^{r\times m}$ with $\operatorname{rk}(G) = \operatorname{rk}(H) = r$. If $G(z)H(z)$ has a zero or pole cancellation at $\alpha \in \mathbb{C}$, then $G(z)H(z)$ has a zero-pole cancellation at α.*

Proof. Assume that $G(z)H(z)$ has a pole cancellation at $\alpha \in \mathbb{C}$ (the proof for the case of a zero cancellation at $\alpha \in \mathbb{C}$ goes along the same lines).

Let $D(z), D'(z) \in \mathbb{C}(z)^{r\times r}$ be the Smith–McMillan canonical form of $G(z)$, $H(z)$, respectively. We can write

$$G(z) = C(z)D(z)F(z) \quad \text{and} \quad H(z) = C'(z)D'(z)F'(z)$$

with $F(z), C'(z) \in \mathbb{C}[z]^{r\times r}$, $C(z) \in \mathbb{C}[z]^{n\times r}$ and $F'(z) \in \mathbb{C}[z]^{r\times m}$ unimodular matrices. Hence the product $G(z)H(z)$ can be written as

$$G(z)H(z) = C(z)D(z)F(z)C'(z)D'(z)F'(z)$$

where $M(z) := F(z)C'(z) \in \mathbb{C}[z]^{r\times r}$ is unimodular. Notice that, by virtue of the unimodularity of $C(z)$ and $F'(z)$, the Smith–McMillan canonical form of $G(z)H(z)$, denoted by $\Delta(z)$, coincides with that of $D(z)M(z)D'(z)$ (see [Kailath, 1980, Ex. 6.5-6]). Moreover observe that, since $\mathrm{rk}(G) = \mathrm{rk}(H) = r$, then $\mathrm{rk}(DMD') = r$. Therefore, by taking determinants, we have

$$\begin{aligned}\det\Delta(z) &= c\det D(z)\det D'(z)\\ &= c\frac{n(z)}{d(z)}\frac{(z-\alpha)^{\delta_z(G;\alpha)+\delta_z(H;\alpha)}}{(z-\alpha)^{\delta_p(G;\alpha)+\delta_p(H;\alpha)}}\end{aligned} \tag{3.1}$$

with $n(z)$ and $d(z)$ relatively prime polynomials s.t. $n(\alpha)\neq 0$, $d(\alpha)\neq 0$, and $c\in\mathbb{C},\ c\neq 0$. On the other hand, since $\Delta(z)$ is the Smith–McMillan canonical form of $G(z)H(z)$, we get

$$\det\Delta(z) = c\frac{n(z)}{d(z)}\frac{(z-\alpha)^{\delta_z(GH;\alpha)}}{(z-\alpha)^{\delta_p(GH;\alpha)}}. \tag{3.2}$$

Hence, a comparison of (3.1) and (3.2) yields

$$\delta_p(GH;\alpha)-\delta_p(G;\alpha)-\delta_p(H;\alpha) = \delta_z(GH;\alpha)-\delta_z(G;\alpha)-\delta_z(H;\alpha). \tag{3.3}$$

Since, by assumption, $G(z)H(z)$ has a pole cancellation at α, the left-hand side of (3.3) is strictly negative. This in turn implies that the right-hand side of (3.3) is strictly negative, *i.e.*, $G(z)H(z)$ has a zero cancellation at α. From this fact the thesis readily follows. □

Lemma 3.2. *Let $G(z)\in\mathbb{C}(z)^{n\times r}$ and $H(z)\in\mathbb{C}(z)^{r\times m}$ with $\mathrm{rk}(G)=\mathrm{rk}(H)=r$. If $G(z)$ and $H(z)$ have no zeros at $\alpha\in\mathbb{C}$ then*

$$\delta_p(GH;\alpha) = \delta_p(G;\alpha)+\delta_p(H;\alpha).$$

Proof. By following *verbatim* the first part of the proof of Lemma 3.1, we arrive at the expression

$$\begin{aligned}\det\Delta(z) &= \det D(z)\det M(z)\det D'(z)\\ &= c\det D(z)\det D'(z),\quad c\in\mathbb{C},\ c\neq 0.\end{aligned} \tag{3.4}$$

Since by assumption $G(z)$ and $H(z)$ have no zero at α, then $D(z)$ and $D'(z)$ have no zero at α. Furthermore, $\Delta(z)$ has no zero at α. This fact can be seen by taking the inverse of $\Delta(z)$, namely

$$\Delta^{-1}(z) = D'^{-1}(z)M^{-1}(z)D^{-1}(z),$$

and by noting that the latter has no pole at α, since the entries of $D^{-1}(z)$, $M^{-1}(z)$ and $D'^{-1}(z)$ do not have any pole at α. This in turn implies that $\delta_p(D;\alpha)$, $\delta_p(D';\alpha)$ and $\delta_p(\Delta;\alpha)$ coincide with the degree of the pole at α in $\det D(z)$, $\det D'(z)$ and $\det \Delta(z)$, respectively. Hence, by summing up all the previous considerations, we get

$$\begin{aligned}\delta_p(GH;\alpha) &= \delta_p(\Delta;\alpha)\\ &= \delta_p(\det\Delta;\alpha)\\ &\overset{(3.4)}{=} \delta_p(\det D;\alpha) + \delta_p(\det D';\alpha)\\ &= \delta_p(D;\alpha) + \delta_p(D';\alpha)\\ &= \delta_p(G;\alpha) + \delta_p(H;\alpha),\end{aligned}$$

which concludes the proof. □

Remark 3.2. Notice that Lemma 3.1 and Lemma 3.2 still hold when $\alpha = \infty$. As a matter of fact, in this case, we can apply the change of variable $z \mapsto \lambda^{-1}$ and then consider the (degree of the) zero/pole at $\lambda = 0$ in $G(\lambda)$ and $H(\lambda)$.

We conclude this preliminary section with a useful characterization of the class of (complex) all-pass rational matrices provided by the following lemma established in [Alpay and Gohberg, 1988, Thm. 3.12].

Lemma 3.3. *Let $V(z) \in \mathbb{C}(z)^{r\times r}$, $\delta_M(V) = n$, and let $\{\alpha_i\}_{i=1}^n$ be the poles of $V(z)$ counted with multiplicity, then $V(z)$ is all-pass if and only if it can be written as*

$$V(z) = UU_1(z)U_2(z)\cdots U_n(z), \tag{3.5}$$

with $U \in \mathbb{C}^{r\times r}$ being constant unitary and[14]

$$U_i(z) := I_r + \left(\frac{1-\overline{\alpha}_i z}{z-\alpha_i} - 1\right)P_i, \quad \alpha_i \in \overline{\mathbb{C}} \setminus \mathbb{T}, \tag{3.6}$$

with $P_i \in \mathbb{C}^{r\times r}$ being an orthogonal rank-one projection. Moreover, the product in the right-hand side of (3.5) is minimal, i.e., *$\delta_M(V) = \delta_M(U_1) + \cdots + \delta_M(U_n)$.*

[14] We adopt the convention $\frac{1-\overline{\alpha}_i z}{z-\alpha_i} =: z$ if $\alpha_i = \infty$.

Remark 3.3. Given any decomposition of a all-pass matrix $V(z)$ of the form in (3.6), we have that:

1) Every factor $U_i(z)$ in (3.6) is all-pass. Indeed, by direct computation:

$$\begin{aligned} U_i^*(z)U_i(z) &= I_r + \left(\frac{1-\alpha_i z^{-1}}{z^{-1}-\overline{\alpha}_i} - 1\right)P_i + \left(\frac{1-\overline{\alpha}_i z}{z-\alpha_i} - 1\right)P_i \\ &\qquad + \left(\frac{1-\alpha_i z^{-1}}{z^{-1}-\overline{\alpha}_i} - 1\right)\left(\frac{1-\overline{\alpha}_i z}{z-\alpha_i} - 1\right)P_i \\ &= I_r - 2P_i + \frac{1-\alpha_i z^{-1}}{z^{-1}-\overline{\alpha}_i}P_i + \frac{1-\overline{\alpha}_i z}{z-\alpha_i}P_i \\ &\qquad + 2P_i - \frac{1-\alpha_i z^{-1}}{z^{-1}-\overline{\alpha}_i}P_i - \frac{1-\overline{\alpha}_i z}{z-\overline{\alpha}_i}P_i \\ &= I_r. \end{aligned}$$

2) Every pole at α_i of $V(z)$ of degree d_i is accompanied by a zero of $V(z)$ at $1/\overline{\alpha}_i$ of the same degree. In particular, if $\alpha_i \neq \infty$, the Smith–McMillan canonical form of $U_i(z)$ in (3.6) is given by

$$\operatorname{diag}\left[\frac{1}{z-\alpha_i}, 1, \ldots, 1, z - \frac{1}{\overline{\alpha}_i}\right].$$

3) Since the decomposition is minimal and $\delta_M(U_i) = 1$, $i = 1, 2, \ldots, n$, it follows that

$$\delta_M(V) = \sum_{i=1}^{n} \delta_M(U_i) = n.$$

4) Since the orthogonal rank-one projection P_i in (3.6) can be written as $P_i = v_i v_i^*$ with $v_i \in \mathbb{C}^r$ s.t. $\|v_i\|^2 = v_i^* v_i = 1$, it holds[15]

$$\begin{aligned} \det U_i(z) &= \det\left[I_r + \left(\frac{1-\overline{\alpha}_i z}{z-\alpha_i} - 1\right)P_i\right] \\ &= \det\left[I_r + \left(\frac{1-\overline{\alpha}_i z}{z-\alpha_i} - 1\right)v_i v_i^*\right] \\ &= \det\left[1 + v_i^*\left(\frac{1-\overline{\alpha}_i z}{z-\alpha_i} - 1\right)v_i\right] \\ &= \frac{1-\overline{\alpha}_i z}{z-\alpha_i}. \end{aligned}$$

[15] Here, we exploit the fact that if $A \in \mathbb{C}^{n\times r}$ and $B \in \mathbb{C}^{r\times n}$, $\det(I_n + AB) = \det(I_r + BA)$.

3.2 The main theorem

THE following theorem is the main result of this chapter. It states that all minimal spectral factors of a general spectral density are essentially unique.

Theorem 3.1. *Let* $\Phi(z) \in \mathbb{R}(z)^{n\times n}$ *be a spectrum with* $\mathrm{rk}(\Phi) = r \leq n$, $r \neq 0$. *Let* $W(z), W_1(z) \in \mathbb{R}(z)^{r\times n}$ *be such that*

1) $W(z)$ *and* $W_1(z)$ *are spectral factors of* $\Phi(z)$, i.e., $\Phi(z) = W^*(z)W(z) = W_1^*(z)W_1(z)$;

2) $W(z), W_1(z)$ *are analytic in* $\mathscr{A}_p$ *and* $W^{-R}(z), W_1^{-R}(z)$ *are analytic in* $\mathscr{A}_z$, *where* $\mathscr{A}_p, \mathscr{A}_z$ *are weakly unmixed-symplectic regions;*

3) $W(z)$ *and* $W_1(z)$ *are stochastically minimal,* i.e., $\delta_M(W) = \delta_M(W_1) = \frac{1}{2}\delta_M(\Phi)$.

Then, $W_1(z) = TW(z)$ *with* $T \in \mathbb{R}^{r\times r}$ *constant orthogonal.*

Proof. Before illustrating the details of the proof, we outline the key steps in order to provide a "navigation chart" that may help the reader.

1) We consider the all-pass function $T(z)$ satisfying $W_1(z) = T(z)W(z)$ and we show that $T(z)$ must have no poles and zeros in the region $\mathscr{A}_p \cap \mathscr{A}_z$.

2) We then assume by contradiction, that $T(z)$ is non-constant and, more precisely, that $T(z)$ possesses poles both in $\mathscr{A}_p \setminus \mathscr{A}_z$ and in $\mathscr{A}_z \setminus \mathscr{A}_p$.

3) We decompose $T(z)$ according to Lemma 3.3 and, by exploiting the properties of this decomposition, we show that for each pole $\alpha \in \mathscr{A}_p \setminus \mathscr{A}_z$ of $T(z)$ there is a zero-pole cancellation both at α and at $1/\overline{\alpha}$ in the product $T(z)W(z)$. Hence, we arrive at the contradiction that there exists a spectral factor of $\Phi(z)$, say $\tilde{W}(z)$, such that $\delta_M(\tilde{W}) < \frac{1}{2}\delta_M(\Phi)$. Since this is not possible, we conclude that $T(z)$ must have no poles in the region $\mathscr{A}_p \setminus \mathscr{A}_z$.

4) Finally, we exploit the fact that, by point 3), $W(z)$ and $W_1(z)$ are stochastically minimal spectral factors to conclude that $T(z)$ must have no poles in the region $\mathscr{A}_z \setminus \mathscr{A}_p$. This implies that $T(z)$ is a constant and orthogonal matrix.

We now describe the details. Consider the matrix

$$T(z) := W_1(z)W^{-R}(z).$$

By taking into account Property 1), it is immediate to see that $T^*(z)T(z) = I$, *i.e.*, that $T(z)$ is all-pass. Moreover, since the inverse of $T(z)$ is given by

$$T^*(z) = T^{-1}(z) = W(z)W_1^{-R}(z),$$

it follows that $T(z)$ is analytic with its inverse in $\mathscr{A}_z \cap \mathscr{A}_p$. Now observe that

$$W_1(z) = T(z)W(z). \tag{3.7}$$

To see this, set $Z(z) := W_1(z) - T(z)W(z)$. By employing again Property 1), it is immediate to see that $Z^*(z)Z(z) = 0$ so that $Z(z)$ is identically zero in the unit circle and, eventually, $Z(z) = 0$. We need to show that $T(z)$ is constant.

Assume, *ab absurdo*, that $T(z)$ has McMillan degree d with poles $\alpha_1, \ldots, \alpha_n$ of degree $m_1, \ldots, m_n$ ($d = m_1 + \cdots + m_n$), respectively, such that $\alpha_1, \ldots, \alpha_t \in \mathscr{A}_p \setminus \mathscr{A}_z$ and $\alpha_{t+1}, \ldots, \alpha_n \in \mathscr{A}_z \setminus \mathscr{A}_p$. In what follows we assume that $\alpha_i \neq \infty$ for $i = 1, \ldots, n$. As a matter of fact, if this is not the case, we can always find a suitable Möbius transformation $z \mapsto f(z)$ such that $T(f(z))$ has only finite poles. Thus, by considering this transformation, the argument in the proof still applies.

By exploiting Lemma 3.3, we can decompose $T(z)$ as

$$T(z) = UU_{\alpha_1,1}(z)\cdots U_{\alpha_1,m_1-1}(z)U_{\alpha_n}(z)\cdots U_{\alpha_2}(z)U_{\alpha_1,m_1}(z), \tag{3.8}$$

with $U \in \mathbb{C}^{r\times r}$ constant unitary and

$$U_{\alpha_i,j}(z) := I_r + \left(\frac{1-\overline{\alpha}_i z}{z-\alpha_i} - 1\right) P_{i,j}, \tag{3.9}$$

$$U_{\alpha_i}(z) := U_{\alpha_i,1}(z)\cdots U_{\alpha_i,m_i}(z), \tag{3.10}$$

with $i = 1, \ldots, n$, $j = m_1, \ldots, m_n$, and $P_{i,j} \in \mathbb{C}^{r\times r}$ being an orthogonal rank-one projection.

Now, we can rearrange (3.7) in the form

$$U^*_{\alpha_1,m_1-1}(z)\cdots U^*_{\alpha_1,1}(z)U^*W_1(z) = U_{\alpha_n}(z)\cdots U_{\alpha_2}(z)U_{\alpha_1,m_1}(z)W(z). \tag{3.11}$$

Notice that the left-hand side of (3.11) is analytic in $\mathscr{A}_p \setminus \mathscr{A}_z$ with (right) inverse analytic in $\mathscr{A}_z \setminus \mathscr{A}_p$. It follows that the right-hand side of (3.11) must be analytic in $\mathscr{A}_p \setminus \mathscr{A}_z$ with (right) inverse analytic in $\mathscr{A}_z \setminus \mathscr{A}_p$. By rewriting the right-hand side of (3.11) in a more explicit way, we obtain

$$\begin{aligned}
&U_{\alpha_n}(z)\cdots U_{\alpha_2}(z)U_{\alpha_1,m_1}(z)W(z) = \\
&\overset{(3.9)}{=} U_{\alpha_n}(z)\cdots U_{\alpha_2}(z)\left(I_r - P_{1,m_1} + \frac{1-\overline{\alpha}_1 z}{z-\alpha_1}P_{1,m_1}\right)W(z) \\
&= U_{\alpha_n}(z)\cdots U_{\alpha_2}(z)\frac{1-\overline{\alpha}_1 z}{z-\alpha_1}P_{1,m_1}W(z) + \Delta(z)
\end{aligned}$$

where $\Delta(z) := U_{\alpha_n}(z)\cdots U_{\alpha_2}(z)(I_r - P_{1,m_1})W(z)$ has no pole at α_1. In fact, $\alpha_1 \in \mathscr{A}_p \setminus \mathscr{A}_z$ so that $W(z)$ does not have a pole at α_1. The minimality of the factorization of $T(z)$ in (3.8) implies that

$$\left(\frac{1-\overline{\alpha}_1 z}{z-\alpha_1}P_{1,m_1}\right)W(z)$$

must have a zero-pole cancellation at α_1. This fact needs a detailed explanation.

First, define $U_{\text{res}} := U_{\alpha_n}(z)\cdots U_{\alpha_2}(z)$ and notice that, since the factorization of $T(z)$ in (3.8) is minimal, the matrix

$$R(z) := U_{\text{res}}(z)\frac{1-\overline{\alpha}_1 z}{z-\alpha_1}P_{1,m_1}$$

has a pole at α_1. In fact, a pole cancellation at α_1 in $R(z)$ would imply a pole cancellation at α_1 in $U_{\alpha_n}(z)\cdots U_{\alpha_2}(z)U_{\alpha_1,m_1}(z)$, yielding that the degree of the pole α_1 in $T(z)$ is less than m_1. However, this is not possible since, by Lemma 3.3, the factorization in (3.8) is minimal. Now, since P_{1,m_1} is an orthogonal rank-one projection, there exists a unitary matrix $Q \in \mathbb{C}^{r\times r}$ such that

$$Q^* P_{1,m_1} Q = \text{diag}[1,0,\ldots,0].$$

Since Q is constant and nonsingular, also

$$\begin{aligned}\tilde{R}(z) := R(z)Q &= U_{\text{res}}(z)\frac{1-\overline{\alpha}_1 z}{z-\alpha_1}QQ^*P_{1,m_1}Q \\ &= U_{\text{res}}(z)Q\text{diag}\left[\frac{1-\overline{\alpha}_1 z}{z-\alpha_1},0,\ldots,0\right] \end{aligned} \tag{3.12}$$

has a pole at α_1. More in detail, at least one entry in the first column of $\tilde{R}(z)$ possesses a pole at α_1, while all the other columns are identically zero. Now consider $A(z) := R(z)W(z)$

$$A(z) := R(z)W(z) = \tilde{R}(z)\tilde{W}(z), \tag{3.13}$$

where $\tilde{W}(z) := Q^*W(z)$. As already observed, $A(z)$ is analytic in $\mathscr{A}_p \setminus \mathscr{A}_z$. Therefore, by taking into account that at least one entry in the first column of $\tilde{R}(z)$ possesses a pole at α_1, while all the other columns are identically zero, it is immediate that every element in the first row of $\tilde{W}(z)$ has a zero at α_1 or is identically zero. Then,

$$Q\,\text{diag}\left[\frac{1-\overline{\alpha}_1 z}{z-\alpha_1},0,\ldots,0\right]\tilde{W}(z) = \left(\frac{1-\overline{\alpha}_1 z}{z-\alpha_1}P_{1,m_1}\right)W(z)$$

has no pole at α_1. This implies that also $U_{\alpha_1,m_1}(z)W(z)$ has no pole at α_1 so that in the product $U_{\alpha_1,m_1}(z)W(z)$ there is a pole cancellation at α_1. Eventually, since $U_{\alpha_1,m_1}(z)$ has full (column-)rank and $W(z)$ has full row-rank, by Lemma 3.1, we can conclude that in the product $U_{\alpha_1,m_1}(z)W(z)$ there is a zero-pole cancellation at α_1.

By replacing (3.11) with

$$W_1^{-R}(z)UU_{\alpha_1}(z)\cdots U_{\alpha_1,m_1-1}(z) = W^{-R}(z)U^*_{\alpha_1,m_1}(z)U^*_{\alpha_2}(z)\cdots U^*_{\alpha_n}(z),$$

we can repeat almost *verbatim* the previous argument in order to conclude that $W^{-R}(z)U^*_{\alpha_1,m_1}(z)$ must have a zero-pole cancellation at $1/\overline{\alpha}_1$, or, equivalently, $U_{\alpha_1,m_1}(z)\cdot W(z)$ must have a zero-pole cancellation at $1/\overline{\alpha}_1$.

The zero-pole cancellations at α_1 and at $1/\overline{\alpha}_1$ in the product $U_{\alpha_1,m_1}(z)W(z)$ imply that

$$\delta_M(U_{\alpha_1,m_1}W) < \delta_M(W) = \frac{1}{2}\delta_M(\Phi).$$

Indeed, let $p_1,\dots,p_h \in \overline{\mathbb{C}}$ be the poles of $W(z)$ s.t. $p_i \neq 1/\overline{\alpha}_1$ for all $i = 1,\dots,h$. Since $U_{\alpha_1,m_1}(z)$ is analytic together with its inverse in $\overline{\mathbb{C}} \setminus \{\alpha_1, 1/\overline{\alpha}_1\}$, it holds

$$\delta_p(W;p_i) = \delta_p(U_{\alpha_1,m_1}W;p_i)$$

for all $i = 1,\dots,h$. Moreover, by the zero-pole cancellations: (i) $U_{\alpha_1,m_1}(z)W(z)$ has no pole at α_1, and (ii) $\delta_p(U_{\alpha_1,m_1}W;1/\overline{\alpha}_1) < \delta_p(W;1/\overline{\alpha}_1)$. Therefore

$$\begin{aligned}\delta_M(U_{\alpha_1,m_1}W) &= \sum_{i=1}^{h}\delta_p(W;p_i) + \delta_p(U_{\alpha_1,m_1}W;1/\overline{\alpha}_1)\\ &< \sum_{i=1}^{h}\delta_p(W;p_i) + \delta_p(W;1/\overline{\alpha}_1) = \delta_M(W).\end{aligned}$$

But this is clearly not possible since, by point 3), $W(z)$ is a stochastically minimal spectral factor. Therefore, $U_{\alpha_1,m_1}(z)$ must be a constant unitary matrix.

The previous reasoning still applies for all the other factors of $T(z)$ having a pole at α_i, $i = 1,\dots,t$, yielding that $m_i = 0$ for all $i = 1,\dots,t$, *i.e.*, $T(z)$ has no poles at α_i, $i = 1,\dots,t$.

It remains to show that $T(z)$ has no pole at $\alpha_{t+1},\dots,\alpha_n$. To this aim, we have

$$W_1(z) = T(z)W(z)$$

and since all the poles of $T(z)$ lie in $\mathscr{A}_z \setminus \mathscr{A}_p$, by Lemma 3.2, we have $\delta_p(W_1;\alpha_i) = \delta_p(T;\alpha_i) + \delta_p(W;\alpha_i)$ for all $i = t+1,\dots,n$, while for all the other poles p_i, $i = 1,\dots,h$, of $W(z)$, $\delta_p(W_1;p_i) = \delta_p(W;p_i)$. This implies that

$$\begin{aligned}\delta_M(W_1) &= \sum_{i=t+1}^{n}\delta_p(T;\alpha_i) + \sum_{i=t+1}^{n}\delta_p(W;\alpha_i) + \sum_{i=1}^{h}\delta_p(W;p_i)\\ &> \sum_{i=t+1}^{n}\delta_p(W;\alpha_i) + \sum_{i=1}^{h}\delta_p(W;p_i)\\ &= \delta_M(W),\end{aligned}$$

which, by virtue of the stochastic minimality of $W_1(z)$, leads to a contradiction. Hence $T(z)$ must have no poles at $\alpha_{t+1},\dots,\alpha_n$.

To conclude, we have shown that $T(z)$ has no poles and hence no zeros, due to the fact that $T(z)$ is an all-pass matrix. Therefore, since it has real entries, $T(z)$ must be a constant orthogonal matrix. □

4. Parametrization of minimal spectral factors

ANOTHER natural question arising from the general factorization result of Chapter 2 concerns the existence of an "efficient" and "compact" parametrization of the whole set of minimal spectral factors of a general rational spectral density. More precisely, this question asks whether all minimal spectral factors of a general discrete-time spectral density can be parametrized in terms of the *all-pass divisors* of a suitable *all-pass function*. Such a parametrization is particularly practical and effective in that the set of all-pass divisors of an all-pass function has been in-depth investigated and well-characterized in the literature from algebraic, geometric and computational viewpoints, see, *e.g.*, Glover [1984], Green and Anderson [1987], Alpay and Gohberg [1988], Fuhrmann and Hoffmann [1997], Michaletzky [1998], Picci [2013], Ferrante and Picci [2017].

As briefly mentioned in the introduction of this part, the main motivation for parametrizing the minimal spectral factors of a rational spectral density comes from stochastic realization theory Lindquist and Picci [1979, 1991]: Any second-order discrete-time purely nondeterministic stationary process $\{y(t)\}$ can be represented as the output of a linear state-space of the form

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) \\ y(t) = Cx(t) + Du(t) \end{cases}$$

where A, B, C, D are matrices of suitable dimensions and $\{u(t)\}$ is a white noise process. This representation is called *stochastic realization* of the process $\{y(t)\}$ and gives a very powerful model for processing and estimation of $\{y(t)\}$. Up to uninteresting changes of basis, stochastic realizations of minimal complexity are in one-to-one correspondence with minimal spectral factors $W(z)$ of the spectral density $\Phi(z)$ of $\{y(t)\}$. For this reason, the problem of parametrizing the minimal spectral factors of $\Phi(z)$ is crucial for the analysis and synthesis of different models of a given stochastic process.

The present chapter builds upon the results of Baggio and Ferrante [2017] and is outlined as follows. We begin by formally introducing the parametrization problem in Section 4.1. Next, Section 4.2 collects some results from Ferrante and Picci [2017] which provide a geometric and algebraic parametrization of the all-pass divisors of a given discrete-time all-pass function. Section 4.3 contains the statement and proof of the main result. This result provides, to the best of our knowledge, the most general parametrization of minimal spectral factors of a rational spectral density, in the sense that it applies to a *completely general* rational spectral density. Finally, in Section 4.4 we apply our main result to a concrete example.

4.1 Preliminary definitions and problem formulation

HERE we first give some preliminary definitions and then we introduce the parametrization problem.

Definition 4.1 (Coercive spectral density)**.** A spectral density $\Phi \in \mathscr{S}^n_{\text{rat}}(\mathbb{T})$ is said to be *coercive* if it positive definite on $\mathbb{T}$.

Remark 4.1. From the above definition it immediately follows that a coercive spectral density cannot have poles and zeros on the unit circle.

Definition 4.2 ((Left and right) all-pass divisor)**.** Given two all-pass functions $G_\ell(z)$ and $G_r(z)$, if

$$\delta_M(G_\ell(z)) + \delta_M(G_r(z)) = \delta_M(G_\ell(z)G_r(z)),$$

then $G_\ell(z)$ and $G_r(z)$ are said to be, respectively, *left all-pass divisor* and *right all-pass divisor* of $G(z) := G_\ell(z)G_r(z)$.

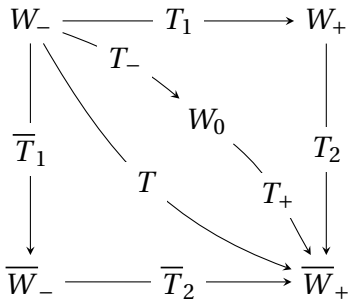
Consider now a rational spectral density $\Phi(z) \in \mathscr{S}^m_{\text{rat}}(\mathbb{T})$ of normal rank $\text{rk}(\Phi) = r \leq m$. Using a "dual" version of the results of Chapter 2, we know that $\Phi(z)$ admits a factorization of the form

$$\Phi(z) = W(z)W^*(z), \tag{4.1}$$

where $W(z) \in \mathbb{R}(z)^{m\times r}$ is a (left) spectral factor of $W(z)$. We can identify four "extremal" minimal (left) spectral factors of $\Phi(z)$, namely:

- $W_-(z)$ analytic with its (left) inverse in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$ (outer spectral factor).
- $W_+(z)$ analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$ with (left) inverse analytic in $\{z \in \overline{\mathbb{C}} : |z| < 1\}$ (maximum-phase spectral factor).
- $\overline{W}_-(z)$ analytic in $\{z \in \overline{\mathbb{C}} : |z| < 1\}$ with (left) inverse analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$ (conjugate maximum-phase spectral factor).
- $\overline{W}_+(z)$ analytic with its (left) inverse in $\{z \in \overline{\mathbb{C}} : |z| < 1\}$ (conjugate outer spectral factor).

These four spectral factors are connected by suitable all-pass transformations as depicted in the commutative diagram below, where an arrow indicates post-multiplication with the labelled object, *e.g.*, $W_+(z) = W_-(z)T_1(z)$.



In the next section, we will show that all minimal spectral factors are connected to $W_-(z)$ by transformations which correspond to the left all-pass divisors of the all-pass function $T(z) := W_-^{-L}(z)\overline{W}_+(z)$. We call $T(z)$ *conjugate phase function* associated with the spectral density $\Phi(z)$, since it can be regarded as the conjugate version of the well-known phase function $\overline{W}_+^{-L}(z)W_-(z)$, which is of crucial importance in stochastic realization theory [Lindquist and Picci, 2015, p. 326].

Remark 4.2. In the above formulation of the parametrization problem, it is worth pointing out two facts:

1) In contrast to what done in Chapter 2 and Chapter 3, we considered in Equation (4.1) the "dual" version of the spectral factorization problem as introduced in Problem 1.1. This is due to the fact this version is more naturally studied in stochastic realization theory, theory which provides the main motivation for our result. Clearly, the result we present in the next section can be easily adapted to the right spectral factors of a given rational spectral densities.

2) It is worth noting that a parametrization of the minimal spectral factors of $\Phi(z)$ in terms of the all-pass divisors of the phase function, in place of the conjugate phase function $T(z)$, can be obtained by minor modifications of the proof presented in the next section. In stating our main result, we chose to consider the conjugate phase function since it seems more natural to refer the parametrization to the minimum-phase spectral factor $W_-(z)$, due to its key role in stochastic modeling and filtering.

4.2 Auxiliary lemmata from Ferrante and Picci [2017]

To make this chapter more self-contained, in this section we state, without proof, some elegant results established in Ferrante and Picci [2017] which are essential for the proof of the main result. Interestingly, one of these results (Lemma 4.2) establishes a particularly elegant and profound characterization of the set of all-pass divisors of a given biproper all-pass function from both an *algebraic* viewpoint, namely in terms of solutions of a certain set of matrix equations that are in turn connected

to solutions of a suitable ARE, and *geometric* viewpoint, namely in terms of invariant subspaces of a certain matrix.

Lemma 4.1 ([Ferrante and Picci, 2017, Theorem 2.1, point 3)]). *Let $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$, $D \in \mathbb{R}^{m\times m}$ be given. If there exists $Q = Q^\top$ satisfying*

$$\begin{cases} A^\top QA - Q = C^\top C, \\ C^\top D - A^\top QB = \mathbf{0}, \\ D^\top D - B^\top QB = I, \end{cases}$$

then $K(z) := C(zI - A)^{-1}B + D$ is a (discrete-time) all-pass function.

Lemma 4.2 ([Ferrante and Picci, 2017, Corollary 4.1, Corollary 3.2, Remarks 3.1 and 3.2]). *Let $K(z) := C(zI - A)^{-1}B + D$ be a minimal realization of a rational biproper (discrete-time) all-pass function. Let $P_0 = P_0^\top$ be the solution of*

$$\begin{cases} AP_0A^\top - P_0 = BB^\top, \\ BD^\top - AP_0C^\top = \mathbf{0}, \\ DD^\top - CP_0C^\top = I. \end{cases}$$

Then, for each $P = [\Pi P_0^{-1}\Pi]^+$, where Π varies among the orthogonal projectors on A-invariant subspaces, the function

$$K_L(z) := C(zI - A)^{-1}G + L, \tag{4.2}$$

with

$$L := (I + CPC^\top)^{1/2}, \quad G := APC^\top L^{-\top}, \tag{4.3}$$

is a (non-minimal) realization of the left all-pass divisor of $K(z)$. Conversely, any left all-pass divisor of $K(z)$ is given by (4.2)-(4.3), up to multiplication from the right side by a constant orthogonal matrix.

4.3 The main result

THE next theorem provides a complete parametrization of all the minimal spectral factors of a completely general spectral density.

Theorem 4.1. *Let $\Phi(z) \in \mathscr{S}^m_{\text{rat}}(\mathbb{T})$ be a spectral density of normal rank $\text{rk}(\Phi) = r \le m$. Let $W_-(z)$ be the outer spectral factor of $\Phi(z)$ and $\overline{W}_+(z)$ be the conjugate outer spectral factor of $\Phi(z)$. Let $T(z) := W_-^{-L}(z)\overline{W}_+(z)$. Let $\mathscr{W}$ be the set of minimal spectral factors of $\Phi(z)$. Then*

$$\mathscr{W} = \left\{ W_-(z)T_\ell(z) : \ T_\ell(z) \text{ is a left all-pass divisor of } T(z) \right\}.$$

Proof. The proof is rather lengthy and involved, so that, for the sake of clarity, is divided in three main parts:

I. First, we show that, without loss of generality, we can restrict the attention to spectral densities that do not have poles/zeros at infinity (set of *biproper* spectral densities).

II. Then, we prove the statement of the theorem for the set of *coercive* spectral densities.

III. Finally, we show how the latter result can be extended to spectral densities that have poles/zeros on the unit circle and/or are normal rank deficient.

Part I. As far as the first part is concerned, suppose that $\Phi(z)$ has a pole/zero at infinity and consider a Möbius transformation $\lambda\colon \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ mapping z in $\lambda(z) = \frac{z-a}{1-az}$, where $a \in \mathbb{R}$ is such that $|a| < 1$ and $1/a$ does not coincide with a pole/zero of $\Phi(z)$. The inverse of this map has the same structure and maps $\lambda \mapsto z(\lambda) = \frac{\lambda+a}{1+a\lambda}$. We observe that:

1) $|\lambda(z)| = 1$ (resp. $|\lambda(z)| > 1$, $|\lambda(z)| < 1$) if and only if $|z| = 1$ (resp. $|z| > 1$, $|z| < 1$);

2) $[\lambda(z)]^{-1} = \lambda(z^{-1})$;

3) $\lambda(\infty) = -1/a$ and $\lambda(z) = \infty$ if and only if $z = 1/a$;

4) If $F(z) \in \mathbb{R}(z)^{k\times h}$ and $G(\lambda) = F(z(\lambda))$ then $\delta_M(F(z)) = \delta_M(G(\lambda))$ [Bart et al., 1979, p. 83];

As a consequence

1) $\Psi(\lambda) := \Phi(z(\lambda))$ is a coercive spectral density;

2) since $1/a$ is not a pole/zero of $\Phi(z)$ then $\Psi(\lambda)$ has no pole/zero at infinity;

3) the outer spectral factor $V_-(\lambda)$ (resp. conjugate outer spectral factor $\overline{V}_+(\lambda)$) of $\Psi(\lambda)$ is given by $V_-(\lambda) = W_-(z(\lambda))$ (resp. $\overline{V}_+(\lambda) = \overline{W}_+(z(\lambda))$);

4) the conjugate phase function $\mathscr{T}(\lambda)$ associated with $\Psi(\lambda)$ is given by $\mathscr{T}(\lambda) = T(z(\lambda))$ and $\mathscr{T}_\ell(\lambda)$ is a left all-pass divisor of $\mathscr{T}(\lambda)$ if and only if $\mathscr{T}_\ell(\lambda) = T_\ell(z(\lambda))$, where $T_\ell(z)$ is a left all-pass divisor of $T(z)$;

5) $V(\lambda)$ is a spectral factor of $\Psi(\lambda)$ if and only if $V(\lambda) = W(z(\lambda))$ where $W(z)$ is a spectral factor of $\Phi(z)$. Moreover, in this case, $\delta_M(V(\lambda)) = \delta_M(W(z))$.

Due to these facts, we can apply the argument presented in what follows to $\Psi(\lambda)$ and then transform back $\lambda(z) \mapsto z$ to recover the desired parametrization for the original spectrum $\Phi(z)$.

Part II. We now proceed with the second part of the proof. In this part, we will show that the theorem holds for the set of *coercive* spectral densities. To this end, assume that $\Phi(z)$ is coercive. In the light of Part I, we can also assume that $\Phi(z)$ is biproper, *i.e.*, it has no pole/zero at infinity. We first show that if $T_\ell(z)$ is a left all-pass divisor of $T(z)$ then $W(z) := W_-(z)T_\ell(z)$ is a minimal spectral factor of $\Phi(z)$. To this end it is clearly sufficient to show that the McMillan degree of $W(z) := W_-(z)T_\ell(z)$ equals the McMillan degree of $W_-(z)$ (which, in turn, is one half of the McMillan degree of the spectral density $\Phi(z)$). To prove this fact, we start from a minimal realization of the outer spectral factor $W_-(z)$:

$$W_-(z) = C(zI - A)^{-1}B + D \tag{4.4}$$

and we follow five steps:

1) We compute a realization of the all-pass function

$$T_1(z) := [W_-(z)]^{-1}W_+(z)$$

in terms of the quadruple A, B, C, D.

2) We compute a realization of the all-pass function

$$T_2(z) := [W_+(z)]^{-1}\overline{W}_+(z)$$

in terms of the quadruple A, B, C, D.

3) We compute a realization of the conjugate phase function $T(z) = T_1(z)T_2(z)$ again in terms of the quadruple A, B, C, D.

4) We use the results of Ferrante and Picci [2017], recalled in Section 4.2, that provide an explicit expression parametrizing the all-pass divisors of a given all-pass function; in this way, we have an expression of $T_\ell(z)$ in terms of the original data A, B, C, D and of a free parameter.

5) We compute the product $W_-(z)T_\ell(z)$ and show that its McMillan degree equals the McMillan degree of $W_-(z)$.

1) Let us consider a minimal realization (4.4) of $W_-(z)$ and let n be the McMillan degree of $W_-(z)$, *i.e.*, the dimension of the matrix A. Let

$$\Gamma := A - BD^{-1}C \tag{4.5}$$

be the zero matrix of $W_-(z)$ and recall that Γ is non-singular and all its eigenvalues have modulus smaller than 1. Moreover, it is worth noticing that the invertibility of D in (4.5) follows from the fact that $\Phi(z)$ is assumed to have no pole/zero at infinity. We now show that

$$T_1(z) := [W_-(z)]^{-1} W_+(z) = H_1(zI-\Gamma)^{-1}G_1 + U_1 \tag{4.6}$$

where

$$\begin{aligned} H_1 &:= D^{-1}C \\ U_1 &:= [I + H_1 X^{-1} H_1^\top]^{1/2} \\ G_1 &:= \Gamma X^{-1} H_1^\top U_1^{-1} \end{aligned}$$

and X is the solution of the Stein equation

$$\Gamma^\top X \Gamma = X + H_1^\top H_1. \tag{4.7}$$

Before proving (4.6), notice that: (i) (A,C) and hence (Γ, H_1) is observable so that X is negative definite. In view of (4.7), this implies that $X + H_1^\top H_1$ is negative definite as well so that $I + H_1 X^{-1} H_1^\top > 0$ and hence U_1 and G_1 are well defined. (ii) By direct computations we get $G_1 G_1^\top = \Gamma X^{-1} \Gamma^\top - X^{-1}$ so that (4.6) is a minimal realization. To prove (4.6), we show that: (i) the right-hand side of (4.6) is all-pass and (ii) the product $W_-(z)[H_1(zI-\Gamma)^{-1}G_1 + U_1]$ has a realization with the same state matrix A of $W_-(z)$ and with zero matrix similar to $\Gamma^{-\top}$. As for (i), it is a matter of direct computation to show that

$$H_1^\top U_1 = \Gamma^\top X G_1, \quad U_1^\top U_1 = I + G_1^\top X G_1.$$

These conditions, together with (4.7), guarantee that the right-hand side of (4.6) is all-pass, *cf.* Lemma 4.1. As for (ii), by taking into account that $BH_1 = BD^{-1}C = A - \Gamma = (zI-\Gamma) - (zI-A)$, we can easily see that $W_-(z)[H_1(zI-\Gamma)^{-1}G_1 + U_1] = C(zI-A)^{-1}B_+ + D_+$, where

$$\begin{aligned} B_+ &:= BU_1 + \Gamma X^{-1} H_1^\top U_1^{-1}, \\ D_+ &:= DU_1. \end{aligned}$$

Hence, its zero matrix is easily seen to be

$$\begin{aligned} \Gamma_+ &:= A - B_+ D_+^{-1} C \\ &= \Gamma - \Gamma X^{-1} H_1^\top [I + H_1 X^{-1} H_1^\top]^{-1} D^{-1} C \\ &= \Gamma (X^{-1} - X^{-1} H_1^\top [I + H_1 X^{-1} H_1^\top]^{-1} H_1 X^{-1}) X \\ &= \Gamma (X + H_1 H_1^\top)^{-1} X = X^{-1} \Gamma^{-\top} X. \end{aligned}$$

In conclusion, $T_1(z)$ is given by the right-hand side of (4.6) and

$$W_+(z) = C(zI-A)^{-1}B_+ + D_+. \tag{4.8}$$

2) We now show that

$$T_2(z) := [W_+(z)]^{-1}\overline{W}_+(z) = H_2(zI - A^{-\top})^{-1}G_2 + U_2 \tag{4.9}$$

where $H_2 := B_+^\top A^{-\top}$, $U_2 := [I + H_2 Y^{-1} H_2^\top]^{1/2}$, $G_2 := A^{-\top} Y^{-1} H_2^\top U_2^{-1}$ and Y is the solution of the Stein equation

$$A^{-1} Y A^{-\top} = Y + H_2^\top H_2. \tag{4.10}$$

Notice that from (stochastic) minimality of $W_+(z)$ it follows that the realization (4.8) is minimal. We can therefore use the same argument used in point 1) to see that: Y is positive definite, U_2 and G_2 are well defined, U_2 is invertible and (4.9) is a minimal realization. To prove (4.9), we show that: (i) the right-hand side of (4.9) is all-pass and (ii) the product $W_+(z)[H_2(zI - A^{-\top})^{-1}G_2 + U_2]$ has a realization with state matrix given by $A^{-\top}$ and with zero matrix similar to $\Gamma^{-\top}$. As for (i), it is a matter of direct computation to show that

$$H_2^\top U_2 = A^{-1} Y G_2, \quad U_2^\top U_2 = I + G_2^\top Y G_2.$$

These conditions, together with (4.10) guarantee that the right-hand side of (4.9) is all-pass by virtue of Lemma 4.1.

As for (ii), by taking into account that $B_+ H_2 = A H_2^\top H_2 = A(A^{-1} Y A^{-\top} - Y) = (zI - A)Y - Y(zI - A^{-\top})$, we can easily see that $W_+(z)[H_2(zI - A^{-\top})^{-1}G_2 + U_2] = D_+ U_2 + \overline{C}_+(sI - A^{-\top})G_2 + N(z)$, where

$$\overline{C}_+ := CY + D_+ H_2 \tag{4.11}$$

and $N(z) := C(zI - A)^{-1}(B_+ U_2 - Y G_2)$; it is now a matter of direct computation to show that $B_+ U_2 - Y G_2 = 0$ so that $N(z) = 0$. The zero matrix of the product $W_+(z)[H_2(zI - A^{-\top})^{-1}G_2 + U_2]$ is thus

$$\begin{aligned}
\overline{\Gamma}_+ &:= A^{-\top} - G_2 (D_+ U_2)^{-1} \overline{C}_+ \\
&= A^{-\top}(I - Y^{-1} H_2^\top [I + H_2 Y^{-1} H_2^\top]^{-1} H_2) - A^{-\top} Y^{-1} H_2^\top U_2^{-2} U_1^{-1} D^{-1} C Y \\
&= A^{-\top}(Y + H_2^\top H_2)^{-1} Y - Y^{-1}(A + A H_2^\top H_2 Y^{-1}) H_2^\top U_2^{-2} U_1^{-1} D^{-1} C Y \\
&= Y^{-1} A Y - Y^{-1} A H_2^\top U_2^2 U_2^{-2} U_1^{-1} D^{-1} C Y \\
&= Y^{-1}(A - A H_2^\top U_1^{-1} D^{-1} C) Y \\
&= Y^{-1} \Gamma_+ Y.
\end{aligned}$$

In conclusion, $T_2(z)$ is given by the right-hand side of (4.9).

Before proceeding to the next point, we need to establish a formula linking X and Y. First observe that taking the inverse of (4.7) and employing the Sherman–Morrison–Woodbury identity[16] we get

[16] Also known as the matrix inversion lemma.

$$X^{-1} = \Gamma X^{-1} \Gamma^\top - \Gamma X^{-1} H_1^\top (I + H_1 X^{-1} H_1^\top)^{-1} H_1 X^{-1} \Gamma^\top.$$

Moreover, equation (4.10) can be rewritten as

$$Y = AYA^\top + B_+B_+^\top. \tag{4.12}$$

By direct computation, we get $B_+B_+^\top = BB^\top + AX^{-1}A^\top - X^{-1}$ which, plugged in (4.12), gives the identity

$$Z = BB^\top + AZA^\top, \tag{4.13}$$

where $Z := Y + X^{-1}$. Notice that by reachability of (A,B), Z is invertible.

3) It is now immediate to compute the following realization of $T(z) = T_1(z)T_2(z)$: $T(z) = \bar{\mathscr{C}}(zI - \bar{\mathscr{A}})^{-1}\bar{\mathscr{B}} + \mathscr{D}$, where

$$\bar{\mathscr{A}} := \begin{bmatrix} \Gamma & G_1H_2 \\ \mathbf{0} & A^{-\top} \end{bmatrix}, \quad \bar{\mathscr{B}} := \begin{bmatrix} G_1U_2 \\ G_2 \end{bmatrix},$$
$$\bar{\mathscr{C}} := [H_1 \mid U_1H_2], \quad \text{and} \quad \mathscr{D} := U_1U_2.$$

By direct computation it is easy to see that

$$-\Gamma X^{-1} + X^{-1}A^{-\top} + G_1H_2 = \mathbf{0}$$

so that we can perform a change of basis in the state space of $T(z)$ induced by the transformation

$$\bar{T} = \begin{bmatrix} I & -X^{-1} \\ \mathbf{0} & I \end{bmatrix}$$

in such a way that

$$T(z) = \mathscr{C}(zI - \mathscr{A})^{-1}\mathscr{B} + \mathscr{D} \tag{4.14}$$

with

$$\mathscr{A} := \bar{T}^{-1}\bar{\mathscr{A}}\bar{T} = \begin{bmatrix} \Gamma & \mathbf{0} \\ \mathbf{0} & A^{-\top} \end{bmatrix},$$
$$\mathscr{B} := \bar{T}^{-1}\bar{\mathscr{B}} = \begin{bmatrix} G_1U_2 + X^{-1}G_2 \\ G_2 \end{bmatrix},$$

and

$$\mathscr{C} := \bar{\mathscr{C}}\bar{T} = [H_1 \mid U_1H_2 - H_1X^{-1}] = [H_1 \mid B^\top A^{-\top}].$$

Thus, is apparent that $(\mathscr{A},\mathscr{C})$ is observable. By using a dual argument it is not difficult to see that $(\mathscr{A},\mathscr{B})$ is reachable so that (4.14) is a minimal realization.

Now define

$$\mathscr{P}_0 := \begin{bmatrix} X^{-1} + X^{-1}Y^{-1}X^{-1} & X^{-1}Y^{-1} \\ Y^{-1}X^{-1} & Y^{-1} \end{bmatrix}$$

and observe that $\mathscr{P}_0$ is invertible; in fact,

$$\mathscr{P}_0^{-1} = \begin{bmatrix} X & -I \\ -I & Y + X^{-1} \end{bmatrix} = \begin{bmatrix} X & -I \\ -I & Z \end{bmatrix}.$$

By long but direct computations, we see that the following relations hold

$$\begin{cases} \mathscr{A}\mathscr{P}_0\mathscr{A}^\top - \mathscr{P}_0 = \mathscr{B}\mathscr{B}^\top \\ \mathscr{A}\mathscr{P}_0\mathscr{C}^\top = \mathscr{B}\mathscr{D}^\top \\ I + \mathscr{C}\mathscr{P}_0\mathscr{C}^\top = \mathscr{D}\mathscr{D}^\top \end{cases} \tag{4.15}$$

Similarly, we get

$$\mathscr{A}^\top\mathscr{P}_0^{-1}\mathscr{A} - \mathscr{P}_0^{-1} = \mathscr{C}^\top\mathscr{C}. \tag{4.16}$$

4) We are now in position to apply Lemma 4.2. In fact, we have a minimal realization (4.14) of the all-pass function $T(z)$ and an explicit expression of the unique solution $\mathscr{P}_0$ of the corresponding linear matrix equation (4.15). By Lemma 4.2, $T_\ell(z)$ is a left all-pass divisor of $T(z)$ if and only if it has the form

$$T_\ell(z) = [\mathscr{C}(zI - \mathscr{A})^{-1}\mathscr{B}_\mathscr{P} + \mathscr{D}_\mathscr{P}]\mathscr{O}$$

where

$$\begin{aligned} \mathscr{B}_\mathscr{P} &:= \mathscr{A}\mathscr{P}\mathscr{C}^\top(I + \mathscr{C}\mathscr{P}\mathscr{C}^\top)^{-1/2}, \\ \mathscr{D}_\mathscr{P} &:= (I + \mathscr{C}\mathscr{P}\mathscr{C}^\top)^{1/2}, \end{aligned} \tag{4.17}$$

$\mathscr{O}$ is an arbitrary orthogonal matrix and $\mathscr{P}$ is of the form

$$\mathscr{P} = \left[\Pi\mathscr{P}_0^{-1}\Pi\right]^+ \tag{4.18}$$

with Π (the parameter of the parametrization) varying among the orthogonal projectors on $\mathscr{A}$-invariant subspaces. Notice that $\mathscr{A}$ is block-diagonal and its blocks Γ and $A^{-\top}$ have disjoint spectra: $\sigma(\Gamma) \cap \sigma(A^{-\top}) = \emptyset$. Hence the invariant subspaces of $\mathscr{A}$ have the form

$$\operatorname{im}\begin{bmatrix} V_\gamma & \mathbf{0} \\ \mathbf{0} & V_a \end{bmatrix}$$

where $\operatorname{im}(V_\gamma)$ is a Γ-invariant subspace and $\operatorname{im}(V_a)$ is a $A^{-\top}$-invariant subspace. Notice that the arbitrary orthogonal matrix $\mathscr{O}$ does not influence the McMillan degree of product $W_-(z)T_\ell(z)$ so that, without loss of generality, from now on we set $\mathscr{O} = I$.

Now, let

$$V := \begin{bmatrix} V_\gamma & \mathbf{0} \\ \mathbf{0} & V_a \end{bmatrix}$$

be a matrix whose columns are a basis for an arbitrary $\mathscr{A}$-invariant subspace. Let $\tilde{T} := [V \mid \tilde{W}]$, where $\tilde{W}$ is such that $\tilde{T}$ is invertible. A change of basis on $\mathscr{A}$ induced by $\tilde{T}$ elicits a block-triangular structure

$$\tilde{T}^{-1}\mathscr{A}\tilde{T} = \begin{bmatrix} F_1 & \tilde{F}_{12} \\ \mathbf{0} & F_2 \end{bmatrix},$$

where, in turn, F_1 has the block-diagonal structure

$$F_1 = \begin{bmatrix} \Gamma_1 & \mathbf{0} \\ \mathbf{0} & A_1^{-\top} \end{bmatrix}. \tag{4.19}$$

We partition now $Q := (\tilde{T}^{-1}\mathscr{P}_0\tilde{T}^{-\top})^{-1}$ conformably as

$$\begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^\top & Q_2 \end{bmatrix}.$$

As shown in [Ferrante and Picci, 2017, Lemma 4.1], Q_1 is invertible so that we can set

$$\bar{T} := \begin{bmatrix} I & \mathbf{0} \\ -Q_{12}^\top Q_1^{-1} & I \end{bmatrix}$$

and we have $\bar{T}Q\bar{T}^\top = \mathrm{diag}[Q_1, Q_2 - Q_{12}^\top Q_1^{-1} Q_{12}]$. Therefore, by defining

$$T := \tilde{T}\bar{T}^\top = [V \mid W] = \begin{bmatrix} V_\gamma & \mathbf{0} & W_1 \\ \mathbf{0} & V_a & W_2 \end{bmatrix}, \tag{4.20}$$

(where W, partitioned conformably with V in two blocks, is a new completion of V to an invertible matrix) we have that $T^{-1}\mathscr{P}_0 T^{-\top}$ has the following block diagonal structure:

$$T^{-1}\mathscr{P}_0 T^{-\top} = \begin{bmatrix} P_1 & \mathbf{0} \\ \mathbf{0} & P_2 \end{bmatrix}.$$

Moreover, the structure of $T^{-1}\mathscr{A}T$ is easily seen to be

$$T^{-1}\mathscr{A}T = \begin{bmatrix} F_1 & F_{12} \\ 0 & F_2 \end{bmatrix}. \tag{4.21}$$

Finally by using (4.18) and observing that in the new basis Π (the orthogonal projector on $\mathrm{im}(V)$) is given by

$$\Pi = \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

we have

$$T^{-1}\mathscr{P}T^{-\top} = \begin{bmatrix} P_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{4.22}$$

5) we are now ready to compute the spectral factor $W(z) := W_-(z)T_\ell(z)$ and show that its McMillan degree is n. To this end, we first define n_γ and n_a to be the number of columns of V_γ and V_a, respectively. By direct computation, we see that

$$B\mathscr{C} = [BH_1 \mid BB^\top A^{-\top}] = [BD^{-1}C \mid ZA^{-\top} - AZ]$$

or

$$B\mathscr{C} = [(zI-\Gamma)-(zI-A) \mid (zI-A)Z - Z(zI-A^{-\top})].$$

Moreover,

$$D\mathscr{C} = [C \mid DB^{\top}A^{-\top}].$$

It is now easy to see that

$$W(z) := W_{-}(z)T_{\ell}(z) = D\mathscr{D}_{\mathscr{P}} + W_1(z) + W_2(z)$$

where

$$W_1(z) := (CZ + DB^{\top}A^{-\top})(zI - A^{-\top})^{-1}[\mathbf{0} \mid I]\mathscr{B}_{\mathscr{P}}$$

and

$$W_2(z) := C(zI-A)^{-1}([I \mid -Z]\mathscr{B}_{\mathscr{P}} + B\mathscr{D}_{\mathscr{P}}).$$

Thus, $\delta_M(W(z)) = \delta_M(W_1(z)) + \delta_M(W_2(z))$. To compute the McMillan degree $\delta_M(W_1(z))$ of $W_1(z)$, consider the term

$$\begin{aligned}
M &:= [\mathbf{0} \mid I]\mathscr{B}_{\mathscr{P}} \\
&= [\mathbf{0} \mid I]\mathscr{A}\mathscr{P}\mathscr{C}^{\top}\mathscr{D}_{\mathscr{P}}^{-1} \\
&= [\mathbf{0} \mid I]TT^{-1}\mathscr{A}TT^{-1}\mathscr{P}T^{\top}T^{-\top}\mathscr{C}^{\top}\mathscr{D}_{\mathscr{P}}^{-1} \\
&= [\mathbf{0} \mid V_a \mid W_2]T^{-1}\mathscr{A}TT^{-1}\mathscr{P}T^{\top}T^{-\top}\mathscr{C}^{\top}\mathscr{D}_{\mathscr{P}}^{-1} \\
&= [\mathbf{0} \mid V_a \mid W_2]\begin{bmatrix} F_1 & F_{12} \\ \mathbf{0} & F_2 \end{bmatrix}\begin{bmatrix} P_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}T^{-\top}\mathscr{C}^{\top}\mathscr{D}_{\mathscr{P}}^{-1} \\
&= [\mathbf{0} \mid V_a \mid W_2]\begin{bmatrix} F_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} P_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}T^{-\top}\mathscr{C}^{\top}\mathscr{D}_{\mathscr{P}}^{-1} \\
&= [\mathbf{0} \mid V_a \mid \mathbf{0}]\begin{bmatrix} F_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} P_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}T^{-\top}\mathscr{C}^{\top}\mathscr{D}_{\mathscr{P}}^{-1} \\
&= [\mathbf{0} \mid V_aA_1^{-\top} \mid \mathbf{0}]\begin{bmatrix} P_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}T^{-\top}\mathscr{C}^{\top}\mathscr{D}_{\mathscr{P}}^{-1} \\
&= V_aA_1^{-\top}E_2,
\end{aligned}$$

where E_2 is the second block rows in the partition of

$$\begin{bmatrix} P_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}T^{-\top}\mathscr{C}^{\top}\mathscr{D}_{\mathscr{P}}^{-1}$$

in three block rows, consistently with the partition $[\mathbf{0} \mid V_aA_1^{-\top} \mid \mathbf{0}]$. From (4.19), (4.20) and (4.21), it immediately follows that $A^{-\top}V_a = V_aA_1^{-\top}$ so that $(zI - A^{-\top})^{-1}V_a = V_a(zI - A_1^{-\top})^{-1}$. Thus

$$W_1(z) = (CZ + DB^{\top}A^{-\top})V_a(zI - A_1^{-\top})^{-1}A_1^{-\top}E_2.$$

Hence, $\delta_M(W_1(z)) \leq n_a = \dim(A_1)$.

To compute the McMillan degree of $W_2(z)$, we analyze the term $N := [I \mid -Z]\mathscr{B}_{\mathscr{P}} + B\mathscr{D}_{\mathscr{P}}$. It can be rewritten as

$$\begin{aligned} N &= [\mathbf{0} \mid -I]\mathscr{P}_0^{-1}\mathscr{A}\,\mathscr{P}\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1} + B\mathscr{D}_{\mathscr{P}} \\ &= [\mathbf{0} \mid -I](\mathscr{A}^{-\top}\mathscr{P}_0^{-1} + \mathscr{A}^{-\top}\mathscr{C}^\top\mathscr{C})\mathscr{P}\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1} + B\mathscr{D}_{\mathscr{P}} \end{aligned}$$

where, for the last equality we exploited (4.16). By direct computation, we get $[\mathbf{0} \mid -I]\mathscr{A}^{-\top}\mathscr{C}^\top = -B$, so that we easily obtain

$$\begin{aligned} [\mathbf{0} \mid -I]\mathscr{A}^{-\top}\mathscr{C}^\top\mathscr{C}\,\mathscr{P}\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1} + B\mathscr{D}_{\mathscr{P}} &= -B(\mathscr{C}\,\mathscr{P}\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1} - \mathscr{D}_{\mathscr{P}}) \\ &= -B(\mathscr{C}\,\mathscr{P}\mathscr{C}^\top - \mathscr{D}_{\mathscr{P}}^2)\mathscr{D}_{\mathscr{P}}^{-1} \\ &= B\mathscr{D}_{\mathscr{P}}^{-1} = [\mathbf{0} \mid I]\mathscr{A}^{-\top}\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1}. \end{aligned}$$

Therefore,

$$N = [\mathbf{0} \mid I]\mathscr{A}^{-\top}(I - \mathscr{P}_0^{-1}\mathscr{P})\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1}.$$

We now use the change of basis in (4.20) and observe that

$$\begin{aligned} I - \mathscr{P}_0^{-1}\mathscr{P} &= T^{-\top}T^\top - T^{-\top}T^\top\mathscr{P}_0^{-1}TT^{-1}\mathscr{P}T^{-\top}T^\top \\ &= T^{-\top}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}T^\top. \end{aligned}$$

Therefore

$$\begin{aligned} N &= [\mathbf{0} \mid I]T^{-\top}T^\top\mathscr{A}^{-\top}T^{-\top}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}T^\top\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1} \\ &= [\mathbf{0} \mid I]T^{-\top}\begin{bmatrix} F_1^{-\top} & \mathbf{0} \\ \star & F_2^{-\top} \end{bmatrix}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}T^\top\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1} \\ &= [\mathbf{0} \mid I]T^{-\top}\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & F_2^{-\top} \end{bmatrix}T^\top\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1}. \end{aligned}$$

Partition now $T^{-\top}$ conformably with T as

$$T^{-\top} = \begin{bmatrix} K_{11}^\top & K_{21}^\top & K_{31}^\top \\ K_{12}^\top & K_{22}^\top & K_{32}^\top \end{bmatrix}$$

so that

$$N = [0 \mid K_{32}^\top F_2^{-\top}]T^\top\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1} = K_{32}^\top F_2^{-\top}[T^\top\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1}]_2$$

where $[T^\top\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1}]_2$ denotes the second block rows of $T^\top\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1}$. From (4.21) it immediately follows that $AK_{32}^\top = K_{32}^\top F_2^{-\top}$ so that $(zI-A)^{-1}K_{32}^\top = K_{32}^\top(zI-F_2^{-\top})^{-1}$. Thus

$$W_2(z) = CK_{32}^\top(zI - F_2^{-\top})^{-1}F_2^{-\top}[T^\top\mathscr{C}^\top\mathscr{D}_{\mathscr{P}}^{-1}]_2. \tag{4.23}$$

Hence, $\delta_M(W_2(z)) \le 2n - n_a - n_\gamma = \dim(F_2)$. To reduce this bound consider the observability matrix of realization (4.23):

$$\begin{bmatrix} CK_{32}^\top \\ CK_{32}^\top F_2^{-\top} \\ CK_{32}^\top (F_2^{-\top})^2 \\ \vdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} K_{32}^\top$$

whose kernel (the un-observable subspace of the realization (4.23)) is the kernel of $K_{32}^\top$. Hence

$$\delta_M(W_2(z)) \le 2n - n_a - n_\gamma - \dim(\ker(K_{32}^\top)).$$

To find $\dim(\ker(K_{32}^\top))$, notice that $K_{32} \in \mathbb{R}^{(2n-n_a-n_\gamma)\times n}$ and from $T^{-1}T = I$ we immediately get $K_{32}^\top V_a = \mathbf{0}$. Let $\tilde{V}_a \in \mathbb{R}^{n\times(n-n_a)}$ be a matrix whose columns complete the columns of V_a to a basis of $\mathbb{R}^n$ so that $[V_a \mid \tilde{V}_a] \in \mathbb{R}^{n\times n}$ is nonsingular. We have

$$\operatorname{rank}(K_{32}) = \operatorname{rank}(K_{32}[V_a \mid \tilde{V}_a]) = \operatorname{rank}([\mathbf{0} \mid K_{32}\tilde{V}_a]) = \operatorname{rank}(K_{32}\tilde{V}_a) \le n - n_a$$

because $K_{32}\tilde{V}_a \in \mathbb{R}^{(2n-n_a-n_\gamma)\times(n-n_a)}$. By recalling that $\ker(K_{32}^\top) = [\operatorname{im}(K_{32})]^\perp$, we have $\dim(\ker(K_{32}^\top)) = 2n - n_a - n_\gamma - \dim(\operatorname{im}(K_{32})) = 2n - n_a - n_\gamma - \operatorname{rank}(K_{32}) \ge n - n_\gamma$. Thus,

$$\delta_M(W_2(z)) \le 2n - n_a - n_\gamma - \dim(\ker(K_{32}^\top)) \le n - n_a.$$

In conclusion,

$$\delta_M(W(z)) = \delta_M(W_1(z)) + \delta_M(W_2(z)) \le n_a + n - n_a = n,$$

and hence $\delta_M(W(z)) = n$ since n is the minimal McMillan degree for a spectral factor of $\Phi(z)$.

We now show the opposite direction, namely that if $W_0(z)$ is a minimal spectral factor of $\Phi(z)$ then $T_-(z) := [W_-(z)]^{-1}W_0(z)$ is a left all-pass divisor of the conjugate phase function $T(z)$. Clearly $T_-(z)$ is all-pass and by defining the all-pass function $T_+(z) := [W_0(z)]^{-1}\overline{W}_+(z)$, we have $T(z) = T_-(z)T_+(z)$ Therefore, we only need to show that $\delta_M(T_-(z)) + \delta_M(T_+(z)) = \delta_M(T(z))$. Since we have already seen that (4.14) is a minimal realization of $T(z)$, so that $\delta_M(T(z)) = 2n$, we need to show that $\delta_M(T_-(z)) + \delta_M(T_+(z)) = 2n$. But the McMillan degree of the product of two rational function is no larger than the sum of the McMillan degrees of the two factors, thus we only need to show that

$$\delta_M(T_-(z)) + \delta_M(T_+(z)) \le 2n. \tag{4.24}$$

To this aim, let us consider a minimal realization

$$W_0(z) = C_0(zI - A_0)^{-1}B_0 + D_0 \tag{4.25}$$

and let $\Gamma_0 := A_0 - B_0 D_0^{-1} C_0$ be the corresponding zero matrix. Notice that by the assumptions on $\Phi(z)$, A_0, Γ_0 and D_0 are invertible and none of the eigenvalues of A_0 and Γ_0 have unitary modulus. We consider two different basis in the state space of $W_0(z)$: one in which

$$\Gamma_0 = \begin{bmatrix} \Gamma_u & \mathbf{0} \\ \mathbf{0} & \Gamma_s \end{bmatrix}, \tag{4.26}$$

and the other in which

$$A_0 = \begin{bmatrix} A_u & \mathbf{0} \\ \mathbf{0} & A_s \end{bmatrix},$$

where all the eigenvalues of Γ_u and A_u have modulus larger than 1 and all the eigenvalues of Γ_s and A_s have modulus smaller than 1. Let γ_1, $\gamma_2 = n - \gamma_1$, a_1 and $a_2 = n - a_1$ be the dimensions of the matrices Γ_u, Γ_s, A_u and A_s, respectively. To conclude, we show that $\delta_M(T_-(z)) \leq \gamma_1 + a_1$ and $\delta_M(T_+(z)) \leq \gamma_2 + a_2$.

Let us consider (4.25) and the basis in which (4.26) holds. Partition C_0 conformably as $C_0 = [C_u \mid C_s]$. Notice that observability of (A_0, C_0) implies observability of the pair (Γ_0, C_0) and, in turn, observability of the pair (Γ_u, C_u). Thus, equation

$$\Gamma_u^\top X_u \Gamma_u = X_u + C_u^\top D_0^{-\top} D_0^{-1} C_u \tag{4.27}$$

admits a unique solution X_u that is positive definite and hence invertible. Hence, $U_- := [I + D_0^{-1} C_u X_u^{-1} C_u^\top D_0^{-\top}]^{1/2}$ is well defined and invertible. Let

$$X_1 := \begin{bmatrix} X_u^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

and consider the function

$$T_{1-}(z) := D_0^{-1} C_0 (zI - \Gamma_0)^{-1} G_- + U_-, \tag{4.28}$$

where $G_- := \Gamma_0 X_1 C_0^\top D_0^{-\top} U_-^{-1}$. Notice that G_- can be rewritten as

$$G_- = \begin{bmatrix} \Gamma_u X_u^{-1} C_u^\top D_0^{-\top} U_-^{-1} \\ \mathbf{0} \end{bmatrix}$$

so that $T_{1-}(z)$ may also be realized as $T_{1-}(z) := C_-(zI - \Gamma_u)^{-1} B_- + U_-$ where, $C_- := D_0^{-1} C_u$ and $B_- := \Gamma_u X_u^{-1} C_u^\top D_0^{-\top} U_-^{-1}$. It is now easy to see that $T_{1-}(z)$ is all-pass. In fact, by direct computation we see that $C_-^\top U_- = \Gamma_u^\top X_u B_-$ and $U_-^\top U_- = I + B_-^\top X_u B_-$ which together with (4.27) imply that $T_{1-}(z)$ is all-pass Lemma 4.1. In addition, since we have derived a realization whose state matrix is Γ_u, clearly $\delta_M(T_{1-}(z)) \leq \gamma_1$. Finally, since U_- is invertible, $T_{1-}^{-1}(z)$ is also a proper all-pass function with McMillan degree $\delta_M(T_{1-}^{-1}(z)) = \delta_M(T_{1-}(z)) \leq \gamma_1$.

We now compute $W_{0-}(z) := W_0(z)T_{1-}(z)$ which is a spectral factor of $\Phi(z)$ because $T_{1-}(z)$ is all-pass. By taking into account that $B_0D_0^{-1}C_0 = (zI-\Gamma_0)-(zI-A_0)$ a direct computation yields

$$W_{0-}(z) = C_0(zI-A_0)^{-1}B_{0-} + D_{0-},$$

where $B_{0-} := B_0U_- + G_-$ and $D_{0-} := D_0U_-$. The zero matrix Γ_- of $W_{0-}(z)$ is given by

$$\begin{aligned}\Gamma_- &= A_0 - B_{0-}D_{0-}^{-1}C_0 = \Gamma_0 - G_-U_-^{-1}D_0^{-1}C_0 \\ &= \begin{bmatrix} \Gamma_u - \Gamma_u X_u^{-1}C_u^\top[D_0D_0^\top + C_0X_1C_0^\top]^{-1}C_u & \mathbf{0} \\ \mathbf{0} & \Gamma_s \end{bmatrix}\end{aligned}$$

and, in view of (4.27),

$$\Gamma_- = \begin{bmatrix} X_u^{-1}\Gamma_u^{-\top}X_u & \mathbf{0} \\ \mathbf{0} & \Gamma_s \end{bmatrix}.$$

Thus all the zeros of $W_{0-}(z)$ have modulus smaller than 1.

In conclusion, there exists a proper all-pass function $T_{1-}(z)$ with $\delta_M(T_{1-}^{-1}(z)) = \delta_M(T_{1-}(z)) \le \gamma_1$, such that $W_{0-}(z) := W_0(z)T_{1-}(z)$ is a spectral factor of $\Phi(z)$ having (i) the same state matrix of $W_0(z)$ and (ii) all its zeros inside the unit disk.

Now we consider $V_0(z) := W_{0-}^{-*}(z)$ which has a realization with state matrix similar to $\Gamma_-^{-\top}$ and zero matrix of the form

$$\begin{bmatrix} A_u^{-\top} & \mathbf{0} \\ \mathbf{0} & A_s^{-\top} \end{bmatrix},$$

where A_u has dimension a_1 and all its eigenvalues have modulus larger than 1 and A_s has dimension $n-a_1$ and all its eigenvalues have modulus smaller than 1. We can apply to $V_0(z)$ the same procedure that led from $W_0(z)$ to $W_{0-}(z)$ and we conclude that there exists a proper all-pass function $T_{2-}(z)$ with $\delta_M(T_{2-}^{-1}(z)) = \delta_M(T_{2-}(z)) \le a_1$, such that $V_0(z)T_{2-}(z)$ has (i) the same state matrix of $V_0(z)$ and (ii) all its zeros outside the unit disk. Hence, $W_{0-}(z)T_{2-}(z) = [V_0(z)T_{2-}(z)]^{-*}$ is a spectral factor of $\Phi(z)$ having both its poles and its zeros all inside the unit disk. Hence $W_0(z)T_{1-}(z)T_{2-}(z) = W_{0-}(z)T_{2-}(z) = W_-(z)$ so that $T_-(z) := W_-^{-1}(z)W_0(z) = [T_{1-}(z)T_{2-}(z)]^{-1}$ which proves that $\delta_M(T_-(z)) \le \gamma_1 + a_1$.

The same argument, this time referred to the lower blocks A_s and Γ_s, now yields $\delta_M(T_+(z)) \le \gamma_2 + a_2 = 2n - (\gamma_1 + a_1)$ and hence (4.24).

Part III. In the last part of the proof, we will show that the result proved in Part II of the proof can be extended to the case of general spectral densities. First, by virtue of Part I of the proof, we can suppose that $\Phi(z)$ is biproper, *i.e.*, it has no pole/zero at infinity.

In the remaining part of the proof, we will show that:

1) $W(z) \in \mathbb{R}(z)^{m\times r}$ is a minimal spectral factor of $\Phi(z)$ if and only if it can be written as $W(z) = W_o(z)V(z)$ where $W_o(z)$ is an $n \times r$ fixed factor that is biproper with

zeros/poles in the unit circle and, possibly, in $z = 0$, and $V(z)$ varies among the $r \times r$ biproper minimal spectral factors of a given coercive $r \times r$ spectral density $\Psi(z)$.

2) By letting $V_-(z)$ denote the minimum phase stable spectral factor of $\Psi(z)$, any biproper minimal spectral factor of $\Psi(z)$ can be written as $V(z) = V_-(z)T_\ell(z)$ with $T_\ell(z)$ being a left all pass divisor of $T(z) = W_-^{-L}(z)\overline{W}(z)_+$.

With reference to point 1), we first notice that the spectral density $\Phi(z)$ can be written as $\Phi(z) = F(z)D(z)C(z)$ where $F(z) \in \mathbb{R}[z]^{n\times r}$ and $C(z) \in \mathbb{R}[z]^{r\times n}$ are unimodular matrices and $D(z) \in \mathbb{R}(z)^{r\times r}$ is the Smith–McMillan canonical form of $\Phi(z)$. The minimum-phase spectral factor of $\Phi(z)$ has the form (up to post-multiplication by constant orthogonal matrices)[17] $W_-(z) = F(z)\Theta(z)\Lambda_-(z)P_-(z)$, where $\Theta(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal and has finite poles/zeros on the unit circle, $\Lambda_-(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal and has all the finite (strictly) stable poles/zeros of $\Phi(z)$ in its diagonal, and $P_-(z) \in \mathbb{R}[z, z^{-1}]^{r\times r}$ is a suitable unimodular matrix. Consider the product $F(z)\Theta(z)$ which has poles/zeros in the unit circle and in $z = \infty$. This product can be factorized as

$$F(z)\Theta(z) = G_-(z)\Delta(z)G_+(z),$$

where $G_-(z) \in \mathbb{R}(z)^{n\times r}$ is biproper and of full column normal rank, $\Delta(z) \in \mathbb{R}(z)^{r\times r}$ is diagonal with monomials of the form z^{κ_i}, $\kappa_i \in \mathbb{Z}$, in its diagonal, and $G_+(z) \in \mathbb{R}[z, z^{-1}]^{r\times r}$ is unimodular. The previous factorization is known as a left Wiener–Hopf factorization at infinity, *cf.* Fuhrmann and Willems [1979]. Notice that $G_-(z)$ must have poles/zeros in the unit circle or in $z = 0$ only. This follows from the fact that (i) $\Delta(z)G_+(z)$ can have poles/zeros in $z = 0$ and in $z = \infty$ only, and (ii) the product $G_-(z)\Delta(z)G_+(z) = F(z)\Theta(z)$ has poles/zeros in the unit circle and in $z = \infty$. We define $W_o(z) := G_-(z)$, $\tilde{V}(z) := \Delta(z)G_+(z)\Lambda_-(z)P_-(z)$, and $\Psi(z) := \tilde{V}(z)\tilde{V}^*(z)$. Notice that $\Psi(z)$ has no zeros/poles in the unit circle and it has full normal rank, that is, it is coercive. Let $\{\beta_i\}_{i=1}^q$ denote the poles in the unit circle of $\Phi(z)$ and $\{\alpha_i\}_{i=1}^p$ denote the remaining poles of $\Phi(z)$. Note that, by construction, it holds

$$\delta_p(\Phi;\alpha_i) = \delta_p(\Psi;\alpha_i),\ i = 1,\dots,p, \tag{4.29}$$

$$\delta_p(W_o;\beta_i) = \frac{1}{2}\delta_p(\Phi;\beta_i),\ i = 1,\dots,q. \tag{4.30}$$

Observe also that, besides the poles $\{\alpha_i\}_{i=1}^p$, $\Psi(z)$ can possess additional poles only in $z = 0$ and $z = \infty$.

Now, let $V(z)$ be any biproper minimal spectral factor of $\Psi(z)$, and consider

$$W(z) := W_o(z)V(z).$$

Since $V(z)$ is taken to be minimal and biproper then, all its non-zero poles are in $\{\alpha_i\}_{i=1}^p$ and they satisfy $\sum_{i=1}^p \delta_p(V;\alpha_i) = \frac{1}{2}\sum_{i=1}^p \delta_p(\Psi;\alpha_i)$ [Bart et al., 2007, p. 163].

[17] This immediately follows from the constructive procedure described in the proof of Theorem 2.1.

Now, notice that $W(z)$ is a spectral factor of $\Phi(z)$ that is again biproper, since $W_o(z)$ and $V(z)$ are so. Moreover $W(z)$ has no pole/zero in $z=0$. This follows from the fact that (i) $\Phi(z)$ is biproper by assumption, and (ii) $W^*(z)$ has no pole/zero in $z=0$. In view of the previous observations and of (4.29),

$$\begin{aligned}\delta_M(W) &= \sum_{i=1}^{p}\delta_p(V;\alpha_i)+\sum_{i=1}^{q}\delta_p(W_o;\beta_i)\\ &= \frac{1}{2}\sum_{i=1}^{p}\delta_p(\Psi;\alpha_i)+\frac{1}{2}\sum_{i=1}^{q}\delta_p(\Phi;\beta_i)\\ &= \frac{1}{2}\sum_{i=1}^{p}\delta_p(\Phi;\alpha_i)+\frac{1}{2}\sum_{i=1}^{q}\delta_p(\Phi;\beta_i)=\frac{\delta_M(\Phi)}{2},\end{aligned}$$

i.e., $W(z)$ is a minimal spectral factor of $\Phi(z)$.

Conversely, let $W(z)$ be any minimal spectral factor of $\Phi(z)$ and let $V_-(z)$ be the stable minimum-phase spectral factor of $\Psi(z)$. As shown before, $W_o(z)V_-(z)$ is a minimal spectral factor of $\Phi(z)$, which in this case coincides with the stable minimum-phase one, that is $W_-(z)=W_o(z)V_-(z)$. It holds

$$W(z)=W_-(z)U(z)=W_o(z)V_-(z)U(z),$$

for a suitable all-pass matrix $U(z)\in\mathbb{R}(z)^{r\times r}$. Now observe that:

1) W has poles only in $\{\alpha_i\}_{i=1}^{p}\cup\{\beta_i\}_{i=1}^{q}$ since it is minimal,

2) U cannot have poles/zeros in the unit circle since it is all-pass, *cf.* Lemma 2.2,

3) W_o has poles/zeros in the unit circle and in $z=0$ only.

These three facts together imply that the non-zero poles of $V_-(z)U(z)$ belong to $\{\alpha_i\}_{i=1}^{p}$ and they satisfy $\delta_p(W;\alpha_i)=\delta_p(V_-U;\alpha_i)$, $i=1,\dots,p$. Moreover, we have that

$$\sum_{i=1}^{p}\delta_p(V_-U;\alpha_i)=\sum_{i=1}^{p}\delta_p(W;\alpha_i)=\frac{1}{2}\sum_{i=1}^{p}\delta_p(\Phi;\alpha_i)\overset{(4.29)}{=}\frac{1}{2}\sum_{i=1}^{p}\delta_p(\Psi;\alpha_i). \tag{4.31}$$

Finally, we notice that $U(z)$ cannot have poles/zeros in $z=0$ and $z=\infty$ (*i.e.*, $U(z)$ must be biproper), otherwise $W(z)=W_-(z)U(z)$ would have a pole in $z=0$ or $z=\infty$ and consequently, in view of the biproperness of $\Phi(z)$, it would not be minimal. This implies that $\delta_p(V_-U;0)=\delta_p(\Psi;0)$. The latter observation together with (4.31) yields

$$\begin{aligned}\delta_M(V_-U) &= \sum_{i=1}^{p}\delta_p(V_-U;\alpha_i)+\delta_p(V_-U;0)\\ &= \frac{1}{2}\sum_{i=1}^{p}\delta_p(\Psi;\alpha_i)+\delta_p(\Psi;0)=\frac{1}{2}\delta_M(\Psi),\end{aligned}$$

that is the product $V_-(z)U(z)$ is a biproper minimal spectral factor of $\Psi(z)$.

We now address point 2) We first notice that

$$\begin{aligned} T(z) &= W_-^{-L}(z)\overline{W}_+(z) = V_-^{-1}(z)W_o^{-L}(z)W_o(z)\overline{V}_{+,0}(z) \\ &= V_-^{-1}(z)\overline{V}_{+,0}(z), \end{aligned} \tag{4.32}$$

where $\overline{V}_{+,0}(z)$ denotes the minimal biproper spectral factor of $\Psi(z)$ having unstable poles/zeros with the only exception for those in $z=0$. Since $\Psi(z)$ is coercive, we can apply the result proved in Part II of the proof and conclude that any minimal spectral factor $V_\ell(z)$ of $\Psi(z)$ (and in particular the biproper ones) can be written in the form

$$V_\ell(z) = V_-(z)T_\ell(z),$$

with $T_\ell(z)$ being a left all pass divisor of $\overline{T}(z) := V_-^{-1}(z)\overline{V}_+(z)$, *i.e.*, $\overline{T}(z) = T_\ell(z)T_{\ell,r}(z)$ with $\delta_M(\overline{T}) = \delta_M(T_\ell) + \delta_M(T_{\ell,r})$, where in this case $\overline{V}_+(z)$ denotes the conjugate outer spectral factor of $\Psi(z)$. In particular, when applied to $\overline{V}_{+,0}(z)$ the latter result reads as

$$\overline{V}_{+,0}(z) = V_-(z)T(z),$$

where $T(z)$, as defined in (4.32), must be a left all pass divisor of $\overline{T}(z)$, *i.e.*, $\overline{T}(z) = T(z)T_r(z)$ with $\delta_M(\overline{T}) = \delta_M(T) + \delta_M(T_r)$. We claim that $T_r(z)$ can have poles/zeros in $z=0$ and $z=\infty$ only. To prove this fact, we decompose $\Psi(z)$ in its Smith–McMillan form $\Psi(z) = \tilde{F}(z)\tilde{D}(z)\tilde{C}(z)$, where $\tilde{F}(z)$ and $\tilde{C}(z)$ are $r\times r$ unimodular matrices and $\tilde{D}(z)$ is the Smith–McMillan canonical form of $\Psi(z)$. Using the dual counterpart of the results proved in Chapter 2 it follows that (up to post-multiplication by orthogonal matrices) $\overline{V}_{+,0}(z)$ and $\overline{V}_+(z)$ have the form $\overline{V}_{+,0}(z) = \tilde{F}(z)\bar{\Lambda}_{+,0}(z)P_{+,0}(z)$, $\overline{V}_+(z) = \tilde{F}(z)\bar{\Lambda}_+(z)P_+(z)$, where $\bar{\Lambda}_{+,0}(z)$ and $\bar{\Lambda}_+(z)$ are the diagonal parts of $\tilde{D}(z)$ containing the zeros/poles in $\{z\in\mathbb{C}:|z|>1\cup z=0\}$ and in $\{z\in\mathbb{C}:|z|>1\}$, respectively, and $P_+(z)$ and $P_{+,0}(z)$ are unimodular matrices. In the light of this fact, we have that

$$T_r(z) = \overline{V}_+^{-1}(z)\overline{V}_{+,0}(z) = P_+(z)^{-1}\bar{\Lambda}_+^{-1}(z)\bar{\Lambda}_{+,0}(z)P_{+,0}(z),$$

is a matrix whose poles are in $z=0$ or $z=\infty$ only.

Now, for every biproper minimal spectral factor $V_\ell(z)$ of $\Psi(z)$, we have

$$V_\ell(z) = V_-(z)T_\ell(z), \quad T(z) = T_\ell(z)T_{r,0,\ell}(z),$$

where $T_{r,0,\ell}(z) := T_{\ell,r}(z)T_r^{-1}(z)$. Since $T_\ell(z)$ and $T(z)$ have no pole/zero in $z=0$ and $z=\infty$, it follows that $T_{r,0,\ell}(z)$ must be biproper, so that all the zeros/poles in $z=0$ and $z=\infty$ must cancel out in the product $T_{\ell,r}(z)T_r^{-1}(z)$. Let $\{\gamma\}_{i=1}^t$ be the

poles of $\overline{T}(z)$ different from 0 and ∞. In view of the minimality of the factorization $\overline{T}(z) = T_\ell(z)T_{\ell,r}(z)$, we have

$$\delta_p(\overline{T};\gamma_i) = \delta_p(T_\ell;\gamma_i) + \delta_p(T_{\ell,r};\gamma_i), \quad i = 1,\dots,t,$$

cf. [Bart et al., 2007, p. 163]. Since $T_r^{-1}(z)$ has poles only in $z = 0$ and in $z = \infty$ and $T(z)$, $T_{r,0,\ell}(z)$ are biproper, it follows that (i) $T(z) = \overline{T}(z)T_r^{-1}(z)$ has the same poles (and pole degrees) of $\overline{T}(z)$ except for those in 0 and ∞, and (ii) $T_{r,0,\ell}(z) := T_{\ell,r}(z)T_r^{-1}(z)$ has the same poles (and pole degrees) of $T_{\ell,r}(z)$ except for those in 0 and ∞. This implies

$$\begin{aligned}\delta_M(T) &= \sum_{i=1}^{t} \delta_p(\overline{T};\gamma_i) = \sum_{i=1}^{t} \delta_p(T_\ell;\gamma_i) + \delta_p(T_{\ell,r};\gamma_i) \\ &= \sum_{i=1}^{t} \delta_p(T_\ell;\gamma_i) + \delta_p(T_{\ell,r,0};\gamma_i) = \delta_M(T_\ell) + \delta_M(T_{\ell,r,0}).\end{aligned}$$

Therefore, any minimal biproper spectral factor of $\Psi(z)$ can be written as $V_\ell(z) = V_-(z)T_\ell(z)$ with $T_\ell(z)$ being a left all pass divisor of $T(z)$. Eventually, by virtue of the one-to-one relation between the biproper minimal spectral factors of $\Psi(z)$ and the minimal spectral factors of $\Phi(z)$ the latter result applies to the minimal spectral factors of $\Phi(z)$ as well. □

4.4 An illustrative example

To demonstrate the usefulness (and non-triviality) of our main result, here we apply Theorem 4.1 to a concrete example arising from stochastic realization theory. To this end, consider a zero-mean purely nondeterministic second-order stationary process $\{y(t)\}$ whose spectral density is

$$\Phi(z) = \frac{1}{z^2 - \frac{5}{2}z + 1}\begin{bmatrix} 3z^2 + \frac{17}{8}z + 3 & 0 \\ 0 & \frac{2}{3}z^2 - \frac{20}{9}z + \frac{2}{3} \end{bmatrix}.$$

We want to compute *all* the minimum "complexity" (i.e., with minimal McMillan degree) dynamical models for the process $\{y(t)\}$. One possible model is the following *minimum-phase* model that can be computed by standard procedures:

$$\begin{cases} x(t+1) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \end{cases}$$

where

$$A = \frac{1}{2}I_2,\ B = I_2,\ C = \begin{bmatrix} 1/4 & 0 \\ 0 & 1/6 \end{bmatrix},\ D = I_2, \tag{4.33}$$

and with $\{u(t)\}$ being a white noise process. This is, however, just one possible choice: to obtain all the models of minimum complexity we can apply our result as follows.

Following the second part of the proof of Theorem 4.1, we have that the conjugate phase function $T(z)$ admits the following minimal state-space realization

$$T(z) = \mathscr{C}(zI_4 - \mathscr{A})^{-1}\mathscr{B} + \mathscr{D}.$$

where

$$\mathscr{A} = \begin{bmatrix} 1/4 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}, \mathscr{B} = \begin{bmatrix} -15/14 & 0 \\ 0 & -16/15 \\ -3/7 & 0 \\ 0 & -3/10 \end{bmatrix},$$
$$\mathscr{C} = \begin{bmatrix} 1/4 & 0 & 2 & 0 \\ 0 & 1/6 & 0 & 2 \end{bmatrix}, \mathscr{D} = \begin{bmatrix} 1/2 & 0 \\ 0 & 2/3 \end{bmatrix}.$$

By invoking Lemma 4.2, we have that there is a one-to-one correspondence between left all-pass divisors of $T(z)$ and invariant subspaces of $\mathscr{A}$. It is immediate to see that the $\mathscr{A}$-invariant subspaces can be classified in four "categories", namely

$$\{0\} \oplus \mathscr{V}, \text{ span}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \oplus \mathscr{V}, \text{ span}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \oplus \mathscr{V}, \mathbb{R}^2 \oplus \mathscr{V},$$

where $\mathscr{V}$ is any subspace of $\mathbb{R}^2$. Thus, in this case, it follows that there are four "categories" of left all-pass divisors of $T(z)$ and to each of these categories there corresponds an *infinite* number of (essentially unique) minimal spectral factors of $\Phi(z)$.[18] More precisely, the left all-pass divisors corresponding to the above categories can be computed as in (4.2)-(4.3) and the corresponding (essentially unique) spectral factors are recovered by pre-multiplying the minimal spectral factor $W_-(z)$, described by the state-space representation (4.33), by the particularly chosen all-pass divisor. We explicitly show how to parametrize the minimal spectral factors belonging to the first category, the others being similar. Thus, we focus on the following category of $\mathscr{A}$-invariant subspaces

$$\{0\} \oplus \mathscr{V}.$$

We can further divide this category into three sub-categories, according to their dimensions:

1) $\{0\} \oplus \{0\}$.

2) $\{0\} \oplus \mathbb{R}^2$.

3) $\mathscr{W}_\theta := \{0\} \oplus \text{span}\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}, \theta \in [0, \pi)$.

[18] Notice that in, contrast to the multivariate case, in the scalar case the number of (essentially unique) minimal spectral factors is always finite!

In the first case $T_\ell = I_2$ and the corresponding spectral factor is the minimum-phase one, $W_-(z)$. In the second and third case, we first compute $\mathscr{P}_0^{-1}$ according to Equation (4.15) in the proof

$$\mathscr{P}_0^{-1} = \begin{bmatrix} 1/15 & 0 & -1 & 0 \\ 0 & 1/32 & 0 & -1 \\ -1 & 0 & 4/3 & 0 \\ 0 & -1 & 0 & 4/3 \end{bmatrix}.$$

Then, for the second case, we have

$$\mathscr{P} = \left[\Pi_2 \mathscr{P}_0^{-1} \Pi_2\right]^+ = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 3/4 & 0 \\ 0 & 0 & 0 & 3/4 \end{bmatrix},$$

where Π_2 is the orthogonal projection onto $\{0\} \oplus \mathbb{R}^2$. In view of Equation (4.17), this yields the (essentially unique) all-pass divisor

$$\overline{T}_1(z) = \mathscr{C}_1 (z I_2 - \mathscr{A}_1)^{-1} \mathscr{B}_1 + \mathscr{D}_1,$$

with

$$\mathscr{A}_1 = 2I_2,\ \mathscr{B}_1 = \frac{3}{2} I_2,\ \mathscr{C}_1 = 2I_2,\ \mathscr{D}_1 = 2I_2.$$

To this all-pass divisor there corresponds the (essentially unique) unstable minimum-phase spectral factor

$$\overline{W}_-(z) = W_-(z)\overline{T}_1(z) = \overline{C}(z I_2 - \overline{A})^{-1}\overline{B} + \overline{D},$$

where

$$\overline{A} = 2I_2,\ \overline{B} = \begin{bmatrix} -4/5 & 8/5 \\ -8/5 & -4/5 \end{bmatrix},\ \overline{C} = \begin{bmatrix} -7/8 & -7/4 \\ 5/3 & -5/6 \end{bmatrix},\ \overline{D} = 2I_2.$$

In the third case, we have

$$\mathscr{P}_\theta = \left[\Pi_{\mathscr{W}_\theta} \mathscr{P}_0^{-1} \Pi_{\mathscr{W}_\theta}\right]^+ = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{3}{4}\cos^2\theta & \frac{3}{4}\cos\theta\sin\theta \\ 0 & 0 & \frac{3}{4}\cos\theta\sin\theta & \frac{3}{4}\sin^2\theta \end{bmatrix},$$

where $\Pi_{\mathscr{W}_\theta}$ denotes the orthogonal projection onto $\mathscr{W}_\theta$. This yields an entire family of (essentially unique) all-pass divisors parametrized by θ

$$T_\theta(z) = \mathscr{C}_{1,\theta}(z-2)^{-1}\mathscr{B}_{1,\theta} + \mathscr{D}_{1,\theta},$$

with

$$\mathscr{B}_{1,\theta} = \begin{bmatrix} \cos\theta & \sin\theta \end{bmatrix}, \; \mathscr{C}_{1,\theta} = 3\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \; \mathscr{D}_{1,\theta} = \begin{bmatrix} 1+\cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & 1+\sin^2\theta \end{bmatrix}.$$

As before, to this family of all-pass divisors there corresponds the family of (essentially unique) minimal spectral factors

$$W_\theta(z) = W_-(z)\,T_\theta(z) = C_\theta(zI_2 - A_\theta)^{-1}B_\theta + D_\theta, \quad \theta \in [0,\pi),$$

where $(A_\theta, B_\theta, C_\theta, D_\theta)$ is a minimal realization of $W_\theta(z)$ that admits the following explicit form

$$A_\theta = \begin{bmatrix} \frac{8\sin^2\theta+2}{\sin^2\theta+4} & -\frac{3\sqrt{5}\sin 2\theta}{\cos 2\theta-9} \\ -\frac{3\sqrt{5}\sin 2\theta}{\cos 2\theta-9} & -\frac{11\cos 2\theta+21}{2(\cos 2\theta-9)} \end{bmatrix}, \; B_\theta = \begin{bmatrix} \frac{3\cos\theta\sin\theta}{\sqrt{\sin^2\theta+4}} & \frac{3\sin^2\theta+2}{\sqrt{\sin^2\theta+4}} \\ \frac{\sqrt{\frac{5}{2}}(\cos 2\theta+3)}{\sqrt{9-\cos 2\theta}} & \frac{\sqrt{\frac{5}{2}}\cos\theta\sin\theta\sqrt{9-\cos 2\theta}}{\sin^2\theta+4} \end{bmatrix},$$

$$C_\theta = \begin{bmatrix} \frac{3\cos\theta\sin\theta}{\sqrt{\sin^2\theta+4}} & \frac{\sqrt{9-\cos 2\theta}(23\cos 2\theta+33)}{8\sqrt{10}(\sin^2\theta+4)} \\ \frac{9\sin^2\theta+1}{3\sqrt{\sin^2\theta+4}} & \frac{7\sqrt{\frac{5}{2}}\cos\theta\sin\theta\sqrt{9-\cos 2\theta}}{6(\sin^2\theta+4)} \end{bmatrix}, \; D_\theta = \begin{bmatrix} 1+\cos^2\theta & \cos\theta\sin\theta \\ \cos\theta\sin\theta & 1+\sin^2\theta \end{bmatrix}.$$

Similarly, we can obtain a parametrization of each one of the other three classes of spectral factors and in such a way we get four classes of spectral factors accounting for *all* the models of minimal complexity for $\{y(t)\}$. Notice, that each class contains infinitely many (essentially different) spectral factors so that we have parametrized all the *infinitely many* models of minimal complexity for $\{y(t)\}$.

5. Existence of minimal J-spectral factorizations

In the last three chapters, we have addressed and scrupulously dissected the spectral factorization problem for an arbitrary discrete-time spectral density. The present chapter studies a natural extension of this problem, the so-called J-spectral factorization problem. As mentioned in the introductive Chapter 1, the latter problem has found application in a wide range of areas in Systems and Control Theory, most notably in $\mathscr{H}_\infty$ control theory, see Francis [1987], Ionescu et al. [1999], [Bart et al., 2011, Part VII] and references therein. Indeed, in the frequency-domain approach to $\mathscr{H}_\infty$ control, the design of the controller essentially boils down to the solution of two J-spectral factorization problems, one for the state-feedback gain and one for the state-estimator gain Green et al. [1990]. From a technical viewpoint, J-spectral factorization is far more challenging than standard spectral factorization problem due to the presence of a number counterintuitive features occurring only when the to-be-factorized J-spectral density is *not* positive semi-definite on the unit circle. One (and perhaps the most interesting) of these counterintuitive features concerns the minimality properties of the factorization. From Chapter 2, we know that a minimal standard spectral factorization with prescribed poles/zeros regions always exists. Quite surprisingly, this is not the case for J-spectral factorization! This counterintuitive fact has been pointed out in Ran and Rodman [1991], Clements [2000], Colaneri and Ferrante [2006] with reference to both the continuous and discrete time case. For instance, one can check, by direct computation or using the results we will discuss in this chapter, that the following J-spectral density

$$\Phi_a(z) = \begin{bmatrix} 0 & z+a \\ z^{-1}+a & 0 \end{bmatrix}, \quad |a| < 1,$$

does not admit, for all $|a| < 1$, an outer J-spectral factor which is *(stochastically) minimal*, namely whose McMillan degree equals one. On the other hand, a minimal spectral factor of $\Phi_a(z)$ that is *not* outer does exist. As a matter of fact,

$$W_a(z) = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & z+a \\ 1 & -z-a \end{bmatrix}, \quad |a| < 1,$$

satisfies $\Phi_a(z) = W_a^*(z) J W_a(z)$ (here $J = \operatorname{diag}[1, -1]$), is minimal but possesses a pole at $z = \infty$ so that it is *not* analytic in $\{z \in \overline{\mathbb{C}} : |z| > 1\}$.

The present chapter aims at elucidating the issue of existence of *minimal* discrete-time J-spectral factorization of a general J-spectral density, under the only assumption that the latter does not possess poles/zeros on the unit circle. Our principal result,

which, in particular, applies to non-regular and improper J-spectral densities, is a condition for the existence of a minimal J-spectral factorization with prescribed analyticity regions expressed in terms of a standard decomposition of a connected L-unimodular J-spectral density. Furthermore, in case such a factorization exists, we describe an iterative procedure for the computation of the desired minimal J-spectral factor. This chapter goes along the lines of Baggio [2017c] and has the following structure. First, we formally introduce the J-spectral factorization problem in Section 5.1. Then, we list in Section 5.2 a number of auxiliary results; one of these results, in particular, provides a useful decomposition for L-unimodular J-spectral densities. The main existence result is presented in Section 5.3. To conclude, in Section 5.4, we illustrate the applicability of our result by means of a numerical example.

5.1 Problem formulation and standing assumptions

Consider a discrete-time rational J-spectral density $\Phi(z) \in \mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$. Following the arguments in Gohberg et al. [1982], Đoković [1993] applied to the discrete-time case, it can be shown that a rational matrix-valued function admits a J-spectral factorization as in Problem 1.2 if and only if this function is a J-spectral density. Thus, $\Phi(z)$ can be J-spectrally factorized as

$$\Phi(z) = W^*(z) J W(z),$$

for a suitable signature matrix J. As in the standard spectral factorization problem, the J-spectral factor $W(z) \in \mathbb{R}(z)^{m\times n}$ is often required to fulfill some additional analyticity constraints. For instance, in $\mathscr{H}_\infty$ control and robust causal estimation, the sought-for J-spectral factor must possess zeros/poles only in the open unit disk (outer J-spectral factor) or, in acausal robust filtering problems, the pole region is fixed by the system dynamics while the zero region is the open unit disk, *cf.* Colaneri and Ferrante [2006]. Henceforth, we consider the most general analyticity regions compatible with the factorization problem, namely the (weakly) unmixed-symplectic regions previously introduced in Definition 2.1. In addition, throughout this chapter, we will make use of the following assumptions.

Assumption 5.1. $\Phi(z) \in \mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$ has no pole and zero on the unit circle.

Assumption 5.2. $\Phi(z) \in \mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$ is biproper, *i.e.*, it does not possess any pole/zero at $z = 0$ and $z = \infty$.

Remark 5.1. While Assumption 5.1 is required for technical reasons and is quite fundamental in the derivation of the main result, Assumption 5.2 is a "fictious" assumption, in the sense that all the results in this chapter are valid even in case this assumption is not met. The latter assumption will be however helpful in order to simplify some proofs. Appendix A shows how to extend the results of the chapter to the non-biproper case.

5.2 Preliminary analysis

THIS section contains some preliminary lemmata which are instrumental for the proof of the main theorem in the next section. Most of these results can be seen as suitable extensions of corresponding results of Chapter 2.

Lemma 5.1. *Let $\Phi(z) \in \mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$ be of normal rank $\mathrm{rk}(\Phi) = r$ and let $D(z) \in \mathbb{R}(z)^{r\times r}$ be its Smith–McMillan canonical form. Moreover, let $\mathscr{A}_p$ and $\mathscr{A}_z$ denote two weakly unmixed-symplectic regions of the complex plane. Then $D(z)$ can be written as*

$$D(z) = \Sigma(z)\Lambda^*(z)\Lambda(z) \tag{5.1}$$

where

- *$\Lambda(z)$ is diagonal and has as (finite) poles and zeros those of $D(z)$ lying on $\overline{\mathbb{C}} \setminus \mathscr{A}_p$ and $\overline{\mathbb{C}} \setminus \mathscr{A}_z$, respectively,*
- *$\Sigma(z)$ is diagonal with non-zero monomials z^k, $k \in \mathbb{Z}$, on its diagonal.*

Proof. It directly follows from Lemma 2.7, using Assumption 5.1. □

Lemma 5.2. *Let $\Phi(z) \in \mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$ be of normal rank $\mathrm{rk}(\Phi) = r$. Let $\Phi(z) = C(z)D(z)F(z)$, with $C(z) \in \mathbb{R}[z]^{n\times r}$, $F(z) \in \mathbb{R}[z]^{r\times n}$ unimodular and $D(z) \in \mathbb{R}(z)^{r\times r}$, be the Smith–McMillan decomposition of $\Phi(z)$. Furthermore, consider the decomposition of $D(z)$ in Lemma 5.1. Then,*

$$\Psi(z) := \Lambda(z)\Sigma^*(z)C^*(z)F^{-R}(z)\Lambda^{-1}(z) \tag{5.2}$$

is an $r \times r$ L-unimodular J-spectral density. Moreover, $\Psi(z)$ and $\Phi(z)$ have the same signature on the unit circle.

Proof. It follows by applying *verbatim* the arguments of Theorem 2.1, bearing in mind that under Assumption 5.1 we have $\Theta(z) = I_r$. □

Lemma 5.3. *Let $\Psi(z) \in \mathscr{S}^r_{\mathrm{rat},J}(\mathbb{T})$ be a non-constant L-unimodular J-spectral density of normal rank $\mathrm{rk}(\Psi) = r$, and let $\Psi^{\mathrm{hc}} \in \mathbb{R}^{r\times r}$ denote the highest-column-degree coefficient matrix of $\Psi(z)$. If Ψ^{hc} is non-singular, then $\Psi(z)$ has (at least) one zero entry on its diagonal.*

Proof. Since $\Psi(z)$ is L-unimodular and para-Hermitian we have that $\det\Psi(z)$ is a non-zero real constant. Let us denote by $K_i \in \mathbb{Z}$, $i = 1,\dots,r$, the maximum-degree of the i-th column of $\Psi(z)$ and suppose, by contradiction, that Ψ^{hc} is non-singular and $\Psi(z)$ has no diagonal element equal to zero. In view of the latter fact, it follows that $K_i \geq 0$, for all $i = 1,\dots,r$. Moreover, by assumption, $\Psi(z)$ is not constant, thus there exists at least one index $j \in \{1,\dots,r\}$ such that $K_j > 0$. But then, in view of Leibniz formula for determinants and the fact that Ψ^{hc} is non-singular, it would be $\max\deg(\det\Psi(z)) > 0$, which contradicts the fact that $\det\Psi(z)$ is a constant. Hence, the contradiction. □

The following result states that an L-unimodular J-spectral density can be always decomposed in a special form via unimodular transformations. The proof is constructive and provides an iterative procedure leading to this form.

Lemma 5.4. *Let $\Psi(z) \in \mathscr{S}^r_{\mathrm{rat},J}(\mathbb{T})$ be an L-unimodular J-spectral density. Then $\Psi(z)$ can be decomposed as*

$$\Psi(z) = U^*(z)\Delta(z)U(z), \tag{5.3}$$

where $U(z) \in \mathbb{R}[z]^{n\times n}$ is a unimodular polynomial matrix and $\Delta(z) \in \mathbb{R}[z,z^{-1}]^{n\times n}$ has the block form

$$\Delta(z) = \left[\begin{array}{c|c|c} \mathbf{0} & H(z) & \mathbf{0} \\ \hline H^*(z) & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \bar{J} \end{array}\right] \tag{5.4}$$

with $\bar{J} \in \mathbb{R}^{(n-2p)\times(n-2p)}$ is a constant signature matrix and

$$H(z) = \begin{bmatrix} z^{k_1} & 0 & 0 & \cdots & 0 \\ h_{21}(z) & z^{k_2} & 0 & \cdots & 0 \\ h_{31}(z) & h_{32}(z) & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & z^{k_{p-1}} & 0 \\ h_{p1}(z) & h_{p2}(z) & \cdots & h_{p(p-1)}(z) & z^{k_p} \end{bmatrix} \tag{5.5}$$

belongs to $\mathbb{R}[z,z^{-1}]^{p\times p}$, where $0 < k_1 \le k_2 \le \cdots \le k_p$ and each $h_{ij}(z) \in \mathbb{R}[z]$ is such that either $\min\deg h_{ij} > k_j$ *and* $\max\deg h_{ij} < k_i$, *or $h_{ij} = 0$.*

Proof. The proof is constructive: We will describe a procedure which allows to decompose $\Psi(z)$ as in Equation (5.3). This procedure can be seen as an extension to the indefinite case of the algorithm developed in step 4 of the proof of statements 1) and 2) of Theorem 2.2 for the unimodular factorization of an L-unimodular spectral density and consists of two main "blocks". Firstly, we denote by $\Psi_h(z)$, $h \in \mathbb{N}_{>0}$, the matrix obtained after the h-th iteration of the procedure. We set $\Psi_1(z) := \Psi(z)$ and $h \leftarrow 1$. Moreover, we denote by $K_i \in \mathbb{Z}$, $i = 1,\dots,r$, the maximum-degree of the i-th column of $\Psi_h(z)$ and we denote by Ψ_h^{hc} the highest-column-degree coefficient matrix of $\Psi_h(z)$. In what follows, we illustrate the two main blocks of the aforementioned procedure.

Block 1: Ψ_h^{hc} singular. In case Ψ_h^{hc} is singular, we compute a non-zero vector $\mathbf{v}_h = [v_1\ v_2\ \dots\ v_r]^\top \in \mathbb{R}^r$ such that $\Psi_h^{\mathrm{hc}}\mathbf{v}_h = \mathbf{0}$. Next, we choose an index $p \in \mathscr{M}$, where $\mathscr{M} = \{i \in \mathscr{I} \,:\, K_i \ge K_j,\ \forall j \in \mathscr{I}\}$ is the highest maximum-degree active index set and $\mathscr{I} = \{i \,:\, v_i \neq 0\}$ the active index set, as defined in the proof of Theorem 2.2. Hence, we

calculate the unimodular matrix

$$\Omega_h^{-1}(z) := \begin{array}{c} \text{column } p \\ \downarrow \end{array} \begin{bmatrix} 1 & \cdots & 0 & \frac{\nu_1}{\nu_p} z^{K_p - K_1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots & & & 0 \\ \vdots & & 1 & \frac{\nu_{p-1}}{\nu_p} z^{K_p - K_{p-1}} & & & \vdots \\ \vdots & & & 1 & & & \vdots \\ \vdots & & & \frac{\nu_{p+1}}{\nu_p} z^{K_p - K_{p+1}} & 1 & & \vdots \\ 0 & & & \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \frac{\nu_r}{\nu_p} z^{K_p - K_r} & 0 & \cdots & 1 \end{bmatrix}.$$

By defining

$$\Psi_{h+1}(z) := \Omega_h^{-*}(z) \Psi_h(z) \Omega_h^{-1}(z),$$

we obtain a new L-unimodular J-spectral density $\Psi_{h+1}(z)$ Moreover, the maximum-degree of the p-th column of $\Psi_{h+1}(z)$ is lower than K_p, while the maximum-degree of the i-th column, $i \neq p$, is not greater than K_i, as shown in the proof of Theorem 2.2. Eventually, we set $h \leftarrow h+1$ and start a new iteration.

Block 2: Ψ_h^{hc} non-singular. In case Ψ_h^{hc} is non-singular, it must be that either $\Psi_h(z)$ is a constant matrix or, from Lemma 5.3, $\Psi_h(z)$ has (at least) one diagonal entry equal to zero. In the first case, a factorization of $\Psi(z)$ of the form in Equation (5.3) with Δ being a signature matrix is readily obtained via eigen-decomposition, after normalization of eigenvalues to ± 1. In the second case, we proceed as follows.

First, by suitable symmetric row and column permutations, we can bring $\Psi_h(z)$ into the block matrix form

$$\Psi_h^{(1)}(z) := \left[\begin{array}{c|c} 0 & \Psi_{h,21}^{(1)}(z) \\ \hline \Psi_{h,21}^{(1)*}(z) & \Psi_{h,22}'(z) \end{array}\right],$$

where $\Psi_{h,21}^{(1)}(z) \in \mathbb{R}[z, z^{-1}]^{1\times(n-1)}$ and $\Psi_{h,22}^{(1)}(z) \in \mathbb{R}[z, z^{-1}]^{(n-1)\times(n-1)}$. Then, by post-multiplying $\Psi_{h,21}^{(1)}(z)$ by a suitable unimodular matrix $V(z) \in \mathbb{R}[z]^{(n-1)\times(n-1)}$, we can reduce it to the row Hermite form [Kailath, 1980, pp. 375-376],

$$\Psi_{h,21}^{(1)}(z) V(z) = \left[\begin{array}{c|c} h_{11}(z) & \mathbf{0} \end{array}\right], \quad h_{11}(z) \in \mathbb{R}[z, z^{-1}].$$

Next, let us define

$$V_1(z) := \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & V(z) \end{array}\right] \in \mathbb{R}[z]^{r\times r}. \tag{5.6}$$

We have

$$\begin{aligned}\Psi_h^{(2)}(z) &:= V_1^*(z)\Psi_h^{(1)}(z)V_1(z)\\ &= \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & V^*(z) \end{array}\right] \Psi_h'(z) \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & V(z) \end{array}\right]\\ &= \left[\begin{array}{c|c|c} 0 & h_{11}(z) & \mathbf{0} \\ \hline h_{11}^*(z) & \Psi_{h,22}^{(2)}(z) & \Psi_{h,23}^{(2)}(z) \\ \hline \mathbf{0} & \Psi_{h,23}^{(2)*}(z) & \Psi_{h,33}^{(2)}(z) \end{array}\right], \end{aligned} \tag{5.7}$$

where we have defined

$$\left[\begin{array}{c|c} \Psi_{h,22}^{(2)}(z) & \Psi_{h,23}^{(2)}(z) \\ \hline \Psi_{h,23}^{(2)*}(z) & \Psi_{h,33}^{(2)}(z) \end{array}\right] := V_1^*(z)\Psi_{h,22}^{(1)}(z)V_1(z). \tag{5.8}$$

The determinant of the matrix in (5.7) is given by

$$\det \Psi_h^{(2)}(z) = -h_{11}(z)h_{11}^*(z)\det \Psi_{h,33}^{(2)}(z). \tag{5.9}$$

Notice that $\det \Psi_h^{(2)}(z)$ is a non-zero constant, since $\Psi_h^{(2)}(z)$ is L-unimodular and para-Hermitian. Therefore, Equation (5.9) implies that $h_{11}(z)$ and $\Psi_{h,33}^{(2)}(z)$ are also L-unimodular, so that $h_{11}(z) = z^{\pm k_1}$, $k_1 > 0$. Finally, by using a unimodular transformation $V_2(z) \in \mathbb{R}[z, z^{-1}]^{r\times r}$ which clears the entry $\Psi_{h,22}^{(2)}(z)$ and, if needed, rearrange the 1st and 2nd column of $\Psi_h^{(2)}(z)$, we can transform $\Psi_h^{(2)}(z)$ into

$$\begin{aligned}\Psi_h^{(3)}(z) &:= V_2^*(z)\Psi_h^{(2)}(z)V_2(z)\\ &= \left[\begin{array}{c|c|c} 0 & z^{k_1} & \mathbf{0} \\ \hline z^{-k_1} & 0 & \Psi_{h,23}^{(3)}(z) \\ \hline \mathbf{0} & \Psi_{h,23}^{(3)*}(z) & \Psi_{h,33}^{(3)}(z) \end{array}\right]. \end{aligned} \tag{5.10}$$

Next, as observed before, $\Psi_{h,33}^{(3)}(z)$ is again an L-unimodular para-Hermitian matrix of full normal rank. Hence, we can apply the entire procedure so far outlined, this time referred to the bottom diagonal block of the previously obtained matrix. By applying recursively this operation, after a finite number of steps, we end up with a matrix of the form

$$\Psi_h^{(\ell)}(z) = \left[\begin{array}{c|c|c|c|c|c|c} 0 & z^{k_1} & 0 & 0 & 0 & \cdots & \mathbf{0} \\ \hline z^{-k_1} & 0 & \Psi_{h,23}^{(\ell)}(z) & \Psi_{h,24}^{(\ell)}(z) & \Psi_{h,25}^{(\ell)}(z) & \cdots & \Psi_{h,2\ell}^{(\ell)}(z) \\ \hline 0 & \Psi_{h,23}^{(\ell)*}(z) & 0 & z^{k_2} & 0 & \cdots & \mathbf{0} \\ \hline 0 & \Psi_{h,24}^{(\ell)*}(z) & z^{-k_2} & 0 & \Psi_{h,45}^{(\ell)}(z) & \cdots & \Psi_{h,4\ell}^{(\ell)}(z) \\ \hline 0 & \Psi_{h,25}^{(\ell)*}(z) & 0 & \Psi_{h,45}^{(\ell)*}(z) & 0 & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \Psi_{h,2\ell}^{(\ell)*}(z) & \mathbf{0} & \Psi_{h,4\ell}^{(\ell)*}(z) & \mathbf{0} & \cdots & \bar{J} \end{array}\right].$$

At this point, we notice that we can apply a sequence of polynomial elementary transformations in order to:

1) clear the non-zero entries above of $\bar{J}$ and their conjugates (entries highlighted in gray in the previous equation).

2) clear the non-zero entries above each z^{k_i} and their conjugates (in red in the previous equation),

3) reduce as much as possible the max degree of the non-zero entries above each z^{-k_i} and their conjugates (in blue in the previous equation) and increase as much as possible the min degree of the same entries, or, if possible, clear these entries. In this way, each conjugate entry $\Psi^{(\ell)*}_{2j,i+2}$ in blue has either max degree strictly smaller than k_i and min degree strictly greater than k_j, or is identically zero.

Denoting by $V_\ell \in \mathbb{R}[z]^{n\times n}$ the unimodular matrix obtained as the product of the previously described elementary transformations, we get

$$\Psi^{(\ell+1)}_h(z) := V_\ell^* \Psi^{(\ell)}_h(z) V_\ell = \left[\begin{array}{c|c|c|c|c|c|c} 0 & z^{k_1} & 0 & 0 & 0 & \cdots & \mathbf{0} \\ \hline z^{-k_1} & 0 & h^*_{21} & 0 & h^*_{31} & \cdots & \mathbf{0} \\ \hline 0 & h_{21} & 0 & z^{k_2} & 0 & \cdots & \mathbf{0} \\ \hline 0 & 0 & z^{-k_2} & 0 & h^*_{32} & \cdots & \mathbf{0} \\ \hline 0 & h_{31} & 0 & h_{32} & 0 & \cdots & \mathbf{0} \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \bar{J} \end{array}\right],$$

where we let $h_{ij} := \Psi^{(\ell)*}_{2j,i+2} \in \mathbb{R}[z]$. Eventually, a symmetric permutation of columns and rows in the previous expression yields the desired decomposition. □

Remark 5.2. It is worth observing that the matrix $\Delta(z)$ obtained in the decomposition of Lemma 5.4 is typically not unique. As a matter of fact, consider the following L-unimodular J-spectral density

$$\Psi(z) = \begin{bmatrix} 0 & 0 & z & 0 \\ 0 & 0 & z^2 & z^3 \\ z^{-1} & z^{-2} & 0 & 0 \\ 0 & z^{-3} & 0 & 0 \end{bmatrix}.$$

This matrix is already decomposed as in Lemma 5.4 with $\bar{J} = 0$ and $H(z) = \left[\begin{smallmatrix} z & 0 \\ z^2 & z^3 \end{smallmatrix}\right]$, however by applying the unimodular transformation

$$U(z) := \begin{bmatrix} -1 & 0 & 0 & 0 \\ z & 1 & 0 & 0 \\ 0 & 0 & -z & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

on the latter, we obtain

$$\Psi'(z) := U^*(z)\Psi(z)U(z) = \begin{bmatrix} 0 & 0 & z^2 & 0 \\ 0 & 0 & 0 & z^2 \\ z^{-2} & 0 & 0 & 0 \\ 0 & z^{-2} & 0 & 0 \end{bmatrix},$$

which is again in the desired form of Lemma 5.4. This lack of uniqueness, however, it is not crucial for the main result of the next section.

5.3 The main existence result

In this section, we state and prove the main result of this chapter. This result provides a necessary and sufficient condition for the existence of a minimal J-spectral factor with prescribed zeros and poles regions, as well as a procedure for computing the latter spectral factor whenever it exists.

Theorem 5.1. *Let $\Phi(z) \in \mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$ be a J-spectral density of normal rank $\mathrm{rk}(\Phi) = r > 0$ and consider a corresponding L-unimodular $\Psi(z) \in \mathscr{S}^r_{\mathrm{rat},J}(\mathbb{T})$, as defined in Lemma 5.2. Take any decomposition $\Psi(z) = U^*(z)\Delta(z)U(z)$ as in Lemma 5.4. Then $\Phi(z)$ admits a minimal J-spectral factorization*

$$\Phi(z) = W^*(z)JW(z), \tag{5.11}$$

with $W(z) \in \mathbb{R}(z)^{r\times n}$ analytic in a weakly unmixed-symplectic region $\mathscr{A}_p$ with (right) inverse analytic in another (possibly different) weakly unmixed-symplectic region $\mathscr{A}_z$, if and only if $\Delta(z) = J$.

Proof. We first address sufficiency, namely we prove that if $\Delta(z) = J$ then there exists a minimal J-spectral factor of $\Phi(z)$ with the desired analyticity properties. To this end, it suffices to notice that if $\Delta(z)$ is constant then one can apply *verbatim* the iterative procedure outlined in the proofs of Theorem 2.1 and Theorem 2.2 for the calculation of a minimal spectral factor of a "standard" spectral density (*i.e.*, for $J = I$). Hence, following the same procedure, we end up with a desired J-spectral factor of the form

$$W(z) = P(z)D_+(z)F(z),$$

where $P(z)$ is unimodular and such that $P^*(z)JP(z) = \Psi(z)$, $D_+(z)$ is the part of the Smith–McMillan form of $\Phi(z)$ containing all the finite poles and zeros of $\Phi(z)$ lying in $\overline{\mathbb{C}} \setminus \mathscr{A}_p$ and $\overline{\mathbb{C}} \setminus \mathscr{A}_z$, respectively, and $F(z)$ is the same unimodular matrix in the definition of $\Psi(z)$ in Lemma 5.2. We point out that the calculation of $P(z)$ can be performed via the algorithm described in the proof of Lemma 5.4, which, in the case $\Delta(z) = J$, coincides exactly with the algorithm described in step 4 of the proof of statements 1) and 2) of Theorem 2.2.

Next we address necessity, namely we will prove that if $\Psi(z)$ does not admit a decomposition with $\Delta = J$ then there does not exist a minimal J-spectral factor of $\Phi(z)$ having the desired analyticity properties. In view of Lemma 5.4, we know that $\Psi(z)$ can be decomposed in the form $\Psi(z) = U^*(z)\Delta(z)U(z)$, with $U(z)$ unimodular, and

$$\Delta(z) = \left[\begin{array}{c|c|c} \mathbf{0} & H(z) & \mathbf{0} \\ \hline H^*(z) & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \bar{J} \end{array}\right] \in \mathbb{R}(z)^{r\times r} \tag{5.12}$$

with $\bar{J}$ being a signature matrix and

$$H(z) = \begin{bmatrix} z^{k_1} & 0 & 0 & \cdots & 0 \\ h_{21}(z) & z^{k_2} & 0 & \cdots & 0 \\ h_{31}(z) & h_{32}(z) & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & z^{k_{p-1}} & 0 \\ h_{p1}(z) & h_{p2}(z) & \cdots & h_{p(p-1)}(z) & z^{k_p} \end{bmatrix}, \tag{5.13}$$

where $0 < k_1 \le k_2 \le \cdots \le k_p$ and each $h_{ij}(z) \in \mathbb{R}[z]$ is such that either $\min\deg h_{ij} > k_j$ and $\max\deg h_{ij} < k_i$, or $h_{ij} = 0$. Notice that, even if $\Delta(z)$ is not constant, the latter can be still factorized in the form $\Delta(z) = V^*(z)JV(z)$ where $V(z)$ is an L-unimodular matrix having either a pole at $z = 0$ or a zero at $z = 0$ (or both) and J is a suitable signature matrix. In order to see this fact, consider, for instance, the choice

$$V(z) = \frac{1}{\sqrt{2}}\left[\begin{array}{c|c|c} I & H(z) & \mathbf{0} \\ \hline I & -H(z) & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \sqrt{2}I \end{array}\right], \qquad J = \left[\begin{array}{c|c|c} I & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & -I & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \bar{J} \end{array}\right]. \tag{5.14}$$

Hence, from the proof of Theorem 2.1, it holds that $V(z)D_+(z)F(z)$, where $D_+(z)$ is diagonal and contains all the finite poles and zeros of $\Phi(z)$ lying in $\overline{\mathbb{C}} \setminus \mathscr{A}_p$ and $\overline{\mathbb{C}} \setminus \mathscr{A}_z$, respectively, is a (typically non-minimal) J-spectral factor of $\Phi(z)$. In general, all the J-spectral factors of $\Phi(z)$ have the form

$$W(z) = P(z)D_+(z)F(z) \tag{5.15}$$

where $P(z) \in \mathbb{R}(z)^{r\times r}$ is a (typically non-unimodular) rational matrix satisfying $\Psi(z) = P^*(z)JP(z)$. Further, from the Smith–McMillan decomposition $\Phi(z) = C(z)D(z)F(z)$ and the equality $W^*(z) = \Phi(z)W^{-R}(z)J$, it follows that

$$W^*(z) = C(z)D_-(z)P^{-1}(z)J, \tag{5.16}$$

where $C(z)$ is unimodular and $D_-(z) := D(z)D_+^{-1}(z)$ is diagonal and contains all the finite poles and zeros of $\Phi(z)$ lying in the regions $\mathscr{A}_p$ and $\mathscr{A}_z$, respectively.

We now claim that if $W(z)$ is a minimal J-spectral factor of $\Phi(z)$ with analyticity regions $\mathscr{A}_p$ and $\mathscr{A}_z$, then it must necessarily be of the form $W(z) = P(z)D_+(z)F(z)$ with $P(z)$ unimodular and such that $P^*(z)JP(z) = \Psi(z)$. In order to prove the claim we first show that, under the previous assumptions on $W(z)$, $P(z)$ must necessarily be L-unimodular. To this end, consider any J-spectral factor $W(z) = V(z)D_+(z)F(z)$ with $V(z)$ L-unimodular and such that $V^*(z)JV(z) = \Psi(z)$[19] and let $\widetilde{W}(z)$ be a minimal J-spectral factor of $\Phi(z)$ with analyticity regions $\mathscr{A}_p$ and $\mathscr{A}_z$. Let

[19] Recall that such a spectral factor always exists. A concrete example is given in Equation (5.14).

$$\widetilde{W}(z) = \widetilde{P}(z)\widetilde{D}_+(z)\widetilde{F}(z)$$

be the Smith–McMillan decomposition of $\widetilde{W}(z)$ and note that $\widetilde{W}(z)$ and $W(z)$ are related by a J-all-pass transformation $Q(z)$, namely $\widetilde{W}(z) = Q(z)W(z)$ with $Q^*(z)JQ(z) = J$. By rearranging the equality $\Phi(z) = C(z)D(z)F(z) = \widetilde{W}^*(z)J\widetilde{W}(z)$, we have

$$C(z)D_-(z) = \widetilde{W}^*(z)J\widetilde{W}(z)F^{-R}(z)D_+^{-1}(z). \tag{5.17}$$

From [Bart et al., 2007, p. 163], it follows that

$$\delta(\widetilde{W}^*;\alpha) + \delta(\widetilde{W};\alpha) = \delta(\Phi;\alpha), \quad \forall \alpha \in \overline{\mathbb{C}} \tag{5.18}$$

since $\widetilde{W}(z)$ is minimal. In view of this fact and since the left-hand side of (5.17) is analytic in $\overline{\mathbb{C}} \setminus \{\mathscr{A}_p \cup \infty\}$, all the non-zero poles of $\widetilde{W}(z)$ belonging to $\overline{\mathbb{C}} \setminus \{\mathscr{A}_p \cup \infty\}$ must cancel out in the product $\widetilde{W}(z)F^{-R}(z)D_+^{-1}(z)$. In fact, suppose, by contradiction, that the product $\widetilde{W}(z)F^{-R}(z)D_+^{-1}(z)$ possesses a non-zero pole $\alpha \in \overline{\mathbb{C}} \setminus \{\mathscr{A}_p \cup \infty\}$. Then it must necessarily be that a pole at α in one of the entries of $\widetilde{P}^*(z)J\widetilde{P}(z)\widetilde{D}_+(z)\widetilde{F}(z)F^{-R}(z)$ must reduce its degree in the product $\widetilde{D}_+^*(z)\widetilde{P}^*(z)J\widetilde{P}(z)\widetilde{D}_+(z)\widetilde{F}(z)F^{-R}(z)$. Due to the fact that $C(z)D_-(z)$ has no pole at α, this in turn implies that a zero at α in one of the diagonal entries $\widetilde{D}_+^*(z)$ must cancel out in the latter product. By virtue of Lemma 3.1, it then follows that the product $\widetilde{D}_+^*(z)\widetilde{P}^*(z)\widetilde{P}(z)\widetilde{D}_+(z)\widetilde{F}(z)F^{-R}(z)$, and hence $\widetilde{W}^*(z)J\widetilde{W}(z)$ since $\widetilde{F}(z)$ is unimodular and $\widetilde{F}(z)F^{-R}(z)$ is L-unimodular, has a pole cancellation at α. This fact contradicts Equation (5.18), implying that the factorization $\widetilde{W}^*(z)J\widetilde{W}(z)$ is not minimal.

Thus, we have that

$$P(z) := Q(z)V(z) = \widetilde{W}(z)F^{-R}(z)D_+^{-1}(z)$$

is L-unimodular, so that we can write $\widetilde{W}(z) = P(z)D_+(z)F(z)$ with $P(z)$ L-unimodular, as required.

Now notice that, if $P(z)$ is L-unimodular and not unimodular then it must possess either a zero at $z = 0$ or a pole at $z = 0$ (or both). In view of (5.15)-(5.16) and the fact that $\Phi(z)$ is biproper by assumption, this implies that $W(z)$ has an additional pole either at $z = 0$ or at $z = \infty$. Consequently, $W(z)$ is not minimal since Equation (5.18) is not met for $\alpha = 0$ and/or $\alpha = \infty$. We conclude that in order to have a minimal factorization $\Phi(z) = W^*(z)JW(z)$, $\Psi(z)$ must admit a factorization with $P(z)$ unimodular.

So, it remains to prove that, if $\Delta(z)$ is not constant, it is not possible to factorize $\Delta(z)$ in the form $\Delta(z) = P^*(z)JP(z)$, with $P(z)$ unimodular. Suppose that $\Delta(z)$ has the form in (5.12). Let $P(z) := \sum_{i=\ell_{\min}}^{\ell_{\max}} P_i z^i$, where P_i are $r \times r$ constant real matrices. If $P(z)$ is unimodular and satisfies $\Delta(z) = P^*(z)JP(z)$, then it follows that

1) $\ell_{\min} \geq 0$, otherwise $P(z)$ would have a pole at $z = 0$.

2) $\ell_{\max} \leq k_p$, otherwise $P(z)$ would have a zero at $z = 0$, since $P_i^\top JP_0 = \mathbf{0}$, for $i > k_p$, which implies that either $P_i = \mathbf{0}$, for $i > k_p$, or P_0 is singular.

Now assume, by contradiction, that $P(z) = \sum_{i=0}^{k_p} P_i z^i$ has no zero at $z = 0$. This implies that $P_0 \in \mathbb{R}^{r\times r}$ is non-singular. Since $\Delta(z) = P^*(z)JP(z)$, the following hold

$$\begin{aligned}
P_0^\top JP_{k_p} &= E_{k_p} \\
P_1^\top JP_{k_p} + P_0^\top JP_{k_{p-1}} &= E_{k_{p-1}} \\
P_2^\top JP_{k_p} + P_1^\top JP_{k_{p-1}} + P_0^\top JP_{k_{p-2}} &= E_{k_{p-2}} \\
\vdots \qquad\qquad & \vdots \\
P_{k_{p-1}}^\top JP_{k_p} + P_{k_{p-2}}^\top JP_{k_{p-1}} + \cdots + P_0^\top JP_1 &= E_1 \\
P_{k_p}^\top JP_{k_p} + P_{k_{p-1}}^\top JP_{k_{p-1}} + \cdots + P_0^\top JP_0 &= E_0
\end{aligned}$$

where, in view of (5.12) and (5.13), each E_i, has the form

$$E_0 = \left[\begin{array}{c|c|c} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \bar{J} \end{array}\right], \quad E_i = \left[\begin{array}{c|c|c} \mathbf{0} & \tilde{E}_i & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} & \mathbf{0} \end{array}\right], \quad i =, 1, \ldots, k_p,$$

and the non-zero entries of $\tilde{E}_i$ correspond to the coefficients of monomials z^i in $H(z)$.

Now since P_0 is invertible we can rewrite the above system of equations as

$$\begin{aligned}
P_{k_p} &= JP_0^{-\top}E_{k_p} \\
P_{k_{p-1}} &= JP_0^{-\top}E_{k_{p-1}} - JP_0^{-\top}P_1^{\top}JP_{k_p} \\
P_{k_{p-2}} &= JP_0^{-\top}E_{k_{p-2}} - JP_0^{-\top}P_1^{\top}JP_{k_{p-1}} - JP_0^{-\top}P_2^{\top}JP_{k_p} \\
\vdots &= \qquad\qquad \vdots \\
P_1 &= JP_0^{-\top}E_1 - JP_0^{-\top}P_{p-1}^{\top}JP_{k_1} - \cdots - JP_0^{-\top}P_{p-1}^{\top}JP_{k_p} \\
P_0 &= JP_0^{-\top}E_0 - JP_0^{-\top}P_1^{\top}JP_1 - \cdots - JP_0^{-\top}P_{k_p}^{\top}JP_{k_p}.
\end{aligned} \tag{5.19}$$

At this point, from the first k_p equations it follows that each P_i, $i = 1,\dots,k_p$ can be written as

$$P_i = \sum_{j=i}^{k_p} \tilde{P}_{i,j}E_j,$$

for suitably defined matrices $\tilde{P}_{i,j} \in \mathbb{R}^{r\times r}$. By using the previous expression of P_i and recalling the block form of the E_i's, it can be seen that the last equation in (5.19) yields

$$P_0 = \left[\begin{array}{c|c|c} \mathbf{0} & \tilde{P}_{0,12} & \tilde{P}_{0,13} \\ \hline \mathbf{0} & \tilde{P}_{0,22} & \tilde{P}_{0,23} \\ \hline \mathbf{0} & \tilde{P}_{0,32} & \tilde{P}_{0,33} \end{array}\right]$$

for suitable matrices $\tilde{P}_{0,ij}$, $i = 1,2,3$, $j = 2,3$. But this implies that P_0 is singular. Consequently $P(z)$ has a zero at $z = 0$, which implies that $P(z)$ is not unimodular. Hence, we get a contradiction. □

Remark 5.3. Observe that Theorem 5.1 and Lemma 5.4 together provide a procedure for checking the existence of a minimal J-spectral factorization with prescribed poles/zeros regions. As a matter of fact, it suffices to compute a decomposition of the L-unimodular matrix $\Psi(z)$ via the algorithm described in the proof of Lemma 5.4 and then check whether this decomposition yields $\Delta = J$ or not. Furthermore, in case such a minimal factorization exists, the desired J-spectral factor can be readily computed as described in the proof of Theorem 5.1.

Remark 5.4. It is worth noticing that, when a minimal factorization featuring the desired analyticity properties does not exist, then, by following the procedure in the above proof, we can still recover a factorization which is minimal and with the desired analyticity properties *with the only exception for the points* $z = 0$ *and* $z = \infty$. Indeed, the J-spectral factor be computed as $W(z) = V(z)D_+(z)F(z)$, where $V(z)$ is the L-unimodular matrix defined in Equation (5.14).

5.4 A numerical example

We present here an example to elucidate the applicability of the main result, namely Theorem 5.1.

Let us consider the following 3×3 J-spectral density depending on a parameter $\varepsilon\in\mathbb{R}_0$

$$\Phi_\varepsilon(z)=\begin{bmatrix} \frac{z+2+z^{-1}}{(z-2)(z^{-1}-2)} & \frac{1+z^{-1}}{(z-2)(z^{-1}-2)}+\varepsilon & \frac{1+z}{(z-2)(z^{-1}-2)}-\varepsilon \\ \frac{1+z}{(z-2)(z^{-1}-2)}+\varepsilon & \frac{1}{(z-2)(z^{-1}-2)}+\varepsilon & \frac{z}{(z-2)(z^{-1}-2)} \\ \frac{1+z^{-1}}{(z-2)(z^{-1}-2)}-\varepsilon & \frac{z^{-1}}{(z-2)(z^{-1}-2)} & \frac{1}{(z-2)(z^{-1}-2)}-\varepsilon \end{bmatrix},$$

We want to check the existence of a minimal outer J-spectral factor, *i.e.*, a minimal J-spectral factor corresponding to the choice $\mathscr{A}_p=\mathscr{A}_z=\{z\in\overline{\mathbb{C}}:|z|>1\}$, and whenever the latter exists we want to compute it. First, one may check that $\Phi_\varepsilon(z)$ is not of full normal rank, namely $\operatorname{rk}(\Phi_\varepsilon)=2$, and that it satisfies Assumption 5.1 and Assumption 5.2. Indeed, the Smith–McMillan decomposition (*cf.* Theorem 2.3) of $\Phi_\varepsilon(z)$ is given by

$$\Phi_\varepsilon(z)=C(z)D(z)F(z),$$

with $D(z)$ canonical of the form

$$D(z)=\begin{bmatrix} \frac{1}{(z-2)(z-1/2)} & 0 \\ 0 & (z-2)(z-1/2) \end{bmatrix},$$

and $C_\varepsilon(z)$ and $F_\varepsilon(z)$ unimodular matrices which may be taken to be

$$C_\varepsilon(z)=\begin{bmatrix} \varepsilon z^2-\frac{1+5\varepsilon}{2}z+\varepsilon-\frac{1}{2} & 1 \\ \varepsilon z^2-\frac{1+5\varepsilon}{2}z+\varepsilon & 1 \\ -\frac{1}{2} & 0 \end{bmatrix},$$

$$F_\varepsilon(z)=\begin{bmatrix} 2\varepsilon z^2+(1-5\varepsilon)z+2\varepsilon+1 & 1 & 2\varepsilon z^2+(1-5\varepsilon)z+2\varepsilon \\ -2\varepsilon^2 & 0 & -2\varepsilon^2 \end{bmatrix}.$$

In view of the choice of the analyticity regions, w.r.t. the decomposition of Lemma 5.1, we have

$$\Lambda(z)=\begin{bmatrix} \frac{1}{z-\frac{1}{2}} & 0 \\ 0 & z-\frac{1}{2} \end{bmatrix},\quad \Sigma(z)=\begin{bmatrix} -\frac{1}{2z} & 0 \\ 0 & -2z \end{bmatrix},$$

so that, we obtain the following L-unimodular J-spectral density $\Psi(z)$ of Lemma 5.2

$$\Psi_\varepsilon(z)=\begin{bmatrix} -\frac{1}{2}\varepsilon z+\frac{1+5\varepsilon}{4}-\frac{1}{2}\varepsilon z^{-1} & -2z^{-1}+2+\frac{1}{2}z \\ -2z+2-\frac{1}{2}z^{-1} & -\frac{2}{\varepsilon}z-\frac{1-5\varepsilon}{\varepsilon^2}-\frac{2}{\varepsilon}z^{-1} \end{bmatrix}.$$

Next, we decompose Ψ_ε as in Lemma 5.4 using the iterative procedure described in the proof of the same lemma. We distinguish three cases:

- $\varepsilon > 1/3$: we get $\Psi_\varepsilon(z) = U_\varepsilon^*(z) J U_\varepsilon(z)$ with $J = \operatorname{diag}[1,-1]$ and $U_\varepsilon(z)$ unimodular of the form

$$U_\varepsilon(z) = \frac{1}{\sqrt{\varepsilon - \frac{1}{3}}} \begin{bmatrix} \frac{3\varepsilon - 6\varepsilon z + 27\varepsilon^2 - 4}{18\varepsilon} & \frac{3\varepsilon z - 6\varepsilon + 2}{3\varepsilon} \\ -\frac{1}{9\varepsilon} & \frac{1}{3\varepsilon} \end{bmatrix}.$$

  Hence, by virtue of Theorem 5.1, the desired minimal J-spectral factorization exists and is given by $\Phi_\varepsilon(z) = W_\varepsilon^*(z) J W_\varepsilon(z)$ with

$$\begin{aligned} W_\varepsilon(z) &= U_\varepsilon(z)\Lambda(z)F_\varepsilon(z) \\ &= \frac{1}{\sqrt{\varepsilon - \frac{1}{3}}} \begin{bmatrix} \frac{\varepsilon z + \frac{1}{6} z - \frac{1}{2}\varepsilon + \frac{1}{6}}{z - \frac{1}{2}} & \frac{\varepsilon z - \frac{1}{2}\varepsilon + \frac{1}{6}}{z - \frac{1}{2}} & \frac{\frac{1}{6} z}{z - \frac{1}{2}} \\ \frac{\frac{1}{3} z - \varepsilon z + \frac{1}{2}\varepsilon + \frac{1}{3}}{z - \frac{1}{2}} & \frac{\frac{1}{3}}{z - \frac{1}{2}} & \frac{\frac{1}{3} z - \varepsilon z + \frac{1}{2}\varepsilon}{z - \frac{1}{2}} \end{bmatrix}. \end{aligned}$$

- $\varepsilon < 1/3$, $\varepsilon \neq 0$: we get $\Psi_\varepsilon(z) = U_\varepsilon^*(z) J U_\varepsilon(z)$ with $J = \operatorname{diag}[1,-1]$ and $U_\varepsilon(z)$ unimodular of the form

$$U_\varepsilon(z) = \frac{1}{\sqrt{\frac{1}{3} - \varepsilon}} \begin{bmatrix} -\frac{1}{9\varepsilon} & \frac{1}{3\varepsilon} \\ -\frac{3\varepsilon - 6\varepsilon z + 27\varepsilon^2 - 4}{18\varepsilon} & -\frac{3\varepsilon z - 6\varepsilon + 2}{3\varepsilon} \end{bmatrix}.$$

  Hence, by virtue of Theorem 5.1, also in this case the desired minimal J-spectral factorization exists and is given by $\Phi_\varepsilon(z) = W_\varepsilon^*(z) J W_\varepsilon(z)$ with

$$\begin{aligned} W_\varepsilon(z) &= U_\varepsilon(z)\Lambda(z)F_\varepsilon(z) \\ &= \frac{1}{\sqrt{\frac{1}{3} - \varepsilon}} \begin{bmatrix} \frac{\frac{1}{3} z - \varepsilon z + \frac{1}{2}\varepsilon + \frac{1}{3}}{z - \frac{1}{2}} & \frac{\frac{1}{3}}{z - \frac{1}{2}} & \frac{\frac{1}{3} z - \varepsilon z + \frac{1}{2}\varepsilon}{z - \frac{1}{2}} \\ \frac{-\varepsilon z - \frac{1}{6} z + \frac{1}{2}\varepsilon - \frac{1}{6}}{z - \frac{1}{2}} & \frac{-\varepsilon z + \frac{1}{2}\varepsilon - \frac{1}{6}}{z - \frac{1}{2}} & \frac{-\frac{1}{6} z}{z - \frac{1}{2}} \end{bmatrix}. \end{aligned}$$

- $\varepsilon = 1/3$: we get $\Psi_\varepsilon(z) = U_{1/3}^*(z) \Delta(z) U_{1/3}(z)$ with

$$\Delta(z) = \begin{bmatrix} 0 & z \\ z^{-1} & 0 \end{bmatrix},$$

  and $U_{1/3}(z)$ unimodular of the form

$$U_{1/3}(z) = \begin{bmatrix} \frac{7}{2} - z & 3z - 6 \\ -\frac{1}{3} & 1 \end{bmatrix}.$$

  Since $\Delta(z)$ is not constant, by Theorem 5.1, the desired minimal J-spectral factorization does *not* exist.

6. Conclusions

SPECTRAL factorization and its indefinite extension, J-spectral factorization, are ubiquitous in Systems and Control Theory. In this part of the dissertation, we have investigated several aspects of these two problems. In Chapter 2, we presented a procedure for the factorization of a discrete-time spectral density with prescribed analyticity regions in the *most general setting*. We then analyzed, in the same general setting, the connected problems of uniqueness and parametrization of minimal spectral factors in Chapter 3 and Chapter 4, respectively. Finally, in Chapter 5, we addressed the J-spectral factorization problem, deriving a result on the existence of minimal J-spectral factorization under mild assumptions.

In the proofs of the main results of these chapters, we used a combination of state-space and polynomial methods. It is worth remarking that, in contrast to the state-space approach, the employment of technical tools from polynomial/rational matrix theory, such as the Smith–McMillan form, seems to be essential in order to rule out some unnecessary yet hard-to-remove assumptions, such as regularity or properness. This demonstrates how these tools, which have been extensively exploited in the past and then partially forgotten by the advent of the more elegant state-space formalism, can still be very useful.

Directions for future research regard, in particular, the J-spectral factorization problem. A first direction is to dispense with the assumption of absence of zeros/poles on the unit circle, extending in this way the main result of Chapter 5 to arbitrary J-spectral densities. The critical issue here is that, in contrast to standard spectral factorization (*cf.* Lemma 2.6), a J-spectral density can have poles and zeros on the unit circle that do not appear with even multiplicity in the invariant Smith–McMillan functions. This prevents the application of tools and ideas similar to those employed for the standard spectral factorization in Chapter 1. As a concrete example of a J-spectral density exhibiting this critical feature, consider the following function

$$\Phi(z) = \begin{bmatrix} 0 & \frac{z-1}{z+1} \\ -\frac{z-1}{z+1} & 0 \end{bmatrix}.$$

A (right) J-spectral factor of $\Phi(z)$ with $J = \operatorname{diag}[1, -1]$ exists and is given by

$$W(z) := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & \frac{z-1}{z+1} \\ 1 & -\frac{z-1}{z+1} \end{bmatrix}.$$

However, the Smith–McMillan canonical form of $\Phi(z)$ is easily seen to be of the form

$$D(z) = \begin{bmatrix} \frac{z-1}{z+1} & 0 \\ 0 & \frac{z-1}{z+1} \end{bmatrix}$$

so that its diagonal terms have a zero at $z = 1$ and a pole at $z = -1$ of odd multiplicity.

A second compelling direction is to prove or disprove the existence of those J-spectral factors with prescribed poles/zeros regions that *cannot* be minimal. In case such spectral factors exist, a subsequent objective would be to calculate and/or provide meaningful upper bounds on the McMillan degree of these "pathological" J-spectral factors. In order to tackle these problems via the approach used in Chapter 5, the first step would be to investigate the structure of non-minimal J-factorizations of the L-unimodular J-spectral density $\Psi(z)$ introduced in Lemma 5.2.

A. Ruling out the biproperness assumption

In this appendix, we will show that Assumption 5.2 in Chapter 5 can be made without any loss of generality, since all the results of Chapter 5 can be extended to the non-biproper case.

Let $\Phi(z) \in \mathscr{S}^n_{\mathrm{rat},J}(\mathbb{T})$ be a J-spectral density satisfying Assumption 5.1 in Chapter 5 and with prescribed weakly unmixed-symplectic analyticity regions of the J-spectral factor and of its (right) inverse, say $\mathscr{A}_p$ and $\mathscr{A}_z$, respectively. To extend the results of Chapter 5 to the case of non-biproper $\Phi(z)$, we follow a reasoning similar to that used in the first part of the proof of Theorem 4.1 in Chapter 4. More precisely, suppose that $\Phi(z)$ has either a pole or zero at infinity (or both) and consider a Möbius transformation $\lambda\colon \overline{\mathbb{C}} \to \overline{\mathbb{C}}$ mapping z in $\lambda(z) = \frac{z-a}{1-az}$, where $a \in \mathbb{R}$ is such that $a \neq 0$ and $1/a$ does not coincide with a pole/zero of $\Phi(z)$. The inverse of this map has the same structure and maps $\lambda \mapsto z(\lambda) = \frac{\lambda+a}{1+a\lambda}$. Define $\widetilde{\mathscr{A}_p} := \{\lambda(z) \in \overline{\mathbb{C}} : z \in \mathscr{A}_p\}$ and $\widetilde{\mathscr{A}_z} := \{\lambda(z) \in \overline{\mathbb{C}} : z \in \mathscr{A}_z\}$. We observe that:

1) $1/\lambda(z) = \lambda(1/z)$;
2) $|\lambda(z)| = 1$ (resp. $\lambda(z) \in \widetilde{\mathscr{A}_p}$, $\lambda(z) \in \widetilde{\mathscr{A}_z}$) if and only if $|z| = 1$ (resp. $z \in \mathscr{A}_p$, $z \in \mathscr{A}_z$);
3) $\lambda(\infty) = -1/a$ and $\lambda(z) = \infty$ if and only if $z = 1/a$;
4) If $F(z) \in \mathbb{R}(z)^{k\times h}$ and $G(\lambda) = F(z(\lambda))$ then $\delta_M(F(z)) = \delta_M(G(\lambda))$, see [Bart et al., 1979, p. 83];

As a consequence,

1) $\widetilde{\mathscr{A}_p}$ and $\widetilde{\mathscr{A}_z}$ are unmixed-symplectic regions;
2) $\widetilde{\Phi}(\lambda) := \Phi(z(\lambda))$ is a para-Hermitian matrix satisfying Assumption 5.1;
3) since $1/a$ is not a pole/zero of $\Phi(z)$, then $\widetilde{\Phi}(\lambda)$ is biproper;
4) $\widetilde{W}(z)$ is a minimal J-spectral factor of $\widetilde{\Phi}(\lambda)$ analytic in $\widetilde{\mathscr{A}_p}$ with (right) inverse analytic in $\widetilde{\mathscr{A}_z}$ if and only if $W(z)$ is a minimal J-spectral factor of $\Phi(\lambda)$ analytic in $\mathscr{A}_p$ with (right) inverse analytic in $\mathscr{A}_z$;

Thus, it follows that we can apply all the results presented in Chapter 5 to $\widetilde{\Phi}(\lambda)$ with corresponding analyticity regions $\widetilde{\mathscr{A}_p}$, $\widetilde{\mathscr{A}_z}$, and then transform back $\lambda(z) \mapsto z$ to recover the same results for the starting $\Phi(z)$.

Part II

Spectral densities estimation

1. Introduction

ESTIMATING the spectral density of a stochastic process from a finite set of measurement is a central problem in many applications of signal/image processing, control theory, econometrics, and bioengineering. Over the past fifty years, many approaches have been developed to tackle this problem, ranging from parametric to non-parametric techniques (we refer the interested reader to McClellan [1982], Stoica and Moses [1997], Pillai and Shim [2012] for a comprehensive and detailed treatment).

More recently, a novel extremely powerful paradigm for spectral estimation has been advocated by Byrnes, Georgiou, and Lindquist and termed THREE (Tunable High REsolution Estimation) Byrnes et al. [2000] (see also Georgiou [2001], Byrnes et al. [2001b], Georgiou and Lindquist [2003]). In its general form, this paradigm consists of recasting the estimation problem as an optimization problem subjected to a generalized moment constraint. This approach may be viewed as a (substantial) extension of classical Burg-like maximum entropy methods Burg [1975], Jaynes [1982] (see also Pavon and Ferrante [2013] for a unifying geometric picture). As a side comment, we point out that the use of maximum entropy methods is long-established in estimation and statistics since they are based on a very simple and natural rationale, namely, using Dempster's words [Dempster, 1972, p. 161],

> "[...] the principle of seeking maximum entropy is a principle of seeking maximum simplicity of explanation."

Remarkably, THREE-like estimators can be tuned in order to achieve higher resolution within prescribed frequency bands and they typically outperform standard estimation techniques, such as periodogram and AR methods, when short observation records are available. Furthermore, these estimators have been successfully applied in several applied contexts such as speech processing and recognition Byrnes et al. [2007] and tissue temperature sensing Amini et al. [2005].

The THREE paradigm fits into the framework of *generalized moment problems*, *cf.* Akhiezer [1965], Shohat and Tamarkin [1943], Akhiezer and Kreĭn [1962], Kreĭn and Nudel'man [1977]. In the scalar case, given a sequence of complex numbers, $(c_0, c_1, \dots, c_n)$ and a basis, $(\alpha_1, \alpha_2, \dots, \alpha_n)$, of a (finite-dimensional) subspace of the Banach space of complex-valued continuous functions defined on the real interval $[a, b]$, the generalized moment problem consists of finding a positive measure $\mathrm{d}\mu$ such

that

$$\int_a^b \alpha_k(t)\,\mathrm{d}\mu(t) = c_k, \quad k = 0, 1, \ldots, n-1.$$

This problem has a notable history. For the case $\alpha_k(t) = t^k$, it was studied by Russian scholar Čebyšëv in Čebyšëv [1874] and his students, in particular Markov and Lyapunov, as a tool to prove some important limit theorems in probability theory. On a finite interval this problem is known as the Hausdorff moment problem and was solved by Hausdorff for an infinite sequence of moments in Hausdorff [1921]. For the case of infinite sequence of moments on an infinite interval, the problem is known as the Hamburger moment problem Hamburger [1920], while on the semi-infinite interval $[0, \infty)$ it is called the Stieltjes moment problem Stieltjes [1894]. In the latter work, the term "moment problem", borrowed from mechanics, was firstly introduced in the context of studying the analytic behavior of continued fractions.

Since these seminal contributions, the moment problem, together with its modifications and generalizations, has been investigated by a number of illustrious mathematicians such as Carathéodory, Schur, Toeplitz, Nevanlinna, Pick, and many others, and has been influential in the development of modern analysis (we refer to [Kreĭn and Nudel'man, 1977, pp. 166–171], Grenander and Szegö [1958], Kjeldsen [1993], Byrnes and Lindquist [2008] for an accurate historical account).

Generalized moment problems have found fruitful applications in Systems and Control Theory as well. In the majority of these applications, the emphasis has been put on solutions of the moment problem which are computable and feature low "complexity". This has fostered in the past twenty years a large research effort, pioneered by Byrnes, Georgiou, Lindquist and co-workers, on the computation of *rational* solutions of particular moment problems featuring *bounded* McMillan degree.

The mathematical tools used to tackle this class of problems are very profound and beautiful: They embrace differential geometry, global inverse function theorems, analytic interpolation, convex optimization, topological methods, *etc.* Byrnes and Lindquist [2003], Georgiou [2005, 2006], Byrnes and Lindquist [2006, 2007, 2009], Karlsson et al. [2016]. Besides the aforementioned THREE approach to spectral estimation, some remarkable applications include the covariance extension problem Kálmán [1982], Georgiou [1983, 1987a], Byrnes et al. [1995], Byrnes and Lindquist [1997], Byrnes et al. [1998], Carli et al. [2011], Lindquist and Picci [2013], Lindquist et al. [2013], interpolation and robust $\mathscr{H}_\infty$ control Georgiou [1987b, 1999], Byrnes et al. [2001b], Blomqvist et al. [2003], Byrnes et al. [2006], Georgiou and Lindquist [2006], stochastic modelling and identification Byrnes et al. [2001c,a], Enqvist [2004], Georgiou and Lindquist [2008], Avventi et al. [2013], to mention just a few references.

In this part of the dissertation, we will analyze several problems arising from the THREE framework. We begin by outlining, in Chapter 2, a generalized version of the THREE paradigm. Following Georgiou and Lindquist [2003], we will dwell on the "Kullback–Leibler" scalar case and, specifically, on the numerical aspects of the approach. In particular, one of the main drawbacks of the gradient-based algorithm of Georgiou and Lindquist [2003] yielding to the solution of this problem relies on its numerical instability. In order to tackle this issue, in Pavon and Ferrante [2006], the authors presented an alternative numerically robust method based on a fixed-point iteration in the space of unit-trace positive semi-definite matrices. In that paper, a proof of the convergence of the algorithm, although conjectured and confirmed by a large number of numerical evidences, was missing. Five years later, the work Ferrante et al. [2011] provided a partial answer to this conjecture, proving local convergence to the (closure of the set of) positive definite fixed points of the iteration by exploiting the center manifold theory. In Chapter 3 we will address instead the issue of *global* convergence of the Pavon–Ferrante iteration. Chapter 4 concerns a particular instance of parametric THREE-like spectral estimation. More precisely, a parametric family of estimators was introduced in Ferrante et al. [2010] that appears favourable from a computational viewpoint. In the same paper, the authors proved the existence of a solution to the problem only for a special case of prior. In this chapter, we will present instead a *general* existence result which applies to any prior and whose proof hinges on topological degree theory. To conclude, in Chapter 5, we will introduce and analyze some new distances in the space of spectral densities arising from Finsler geometry. In particular, we will focus the attention on several "appealing" properties enjoyed by these distances and how the latter may be used to define a "robust" version of THREE-like spectral estimation.

2. THREE-like spectral estimation

THE aim of this chapter is to illustrate the THREE approach for spectral estimation firstly introduced by Byrnes, Georgiou, and Lindquist in Byrnes et al. [2000]. In Section 2.1, we outline a multivariate version of the latter approach which allows for the handling of an additional prior information. This generalized THREE approach has been discussed in Georgiou and Lindquist [2003] for the scalar case and, *e.g.*, in Georgiou [2005], Ferrante et al. [2008] for the multivariate case. According to this formulation, the estimation problem is recast as an approximation problem in the space of spectral densities under a generalized moment constraint. As in every approximation problem, a suitable "similarity" criterion in the space of spectral densities must be selected. Subsection 2.1.1 deals with the delicate issue of existence of solutions to the latter problem. In Section 2.2, we follow the treatment in Georgiou and Lindquist [2003] and we stick to the scalar case in which the chosen measure coincides with the Kullback–Leibler divergence. In particular, in Subsection 2.2.1 we outline the derivation of the dual problem, which is is finite-dimensional (Lagrange multipliers are indeed Hermitian matrices), and, in Subsection 2.2.2, we briefly discuss some issues arising from the numerical solution of the dual problem. The latter subsection provides the main motivation for the problem that we will address in the next chapter.

2.1 Problem formulation

LET $\{y(t)\}_{t\in\mathbb{Z}}$ be an m-dimensional zero-mean second-order purely nondeterministic stationary stochastic process and suppose we possess a finite observation record of the process, say $\{y(t_k)\}_{k=1}^N$. The task is to estimate the unknown spectral density $\Phi \in \mathscr{S}_+^m(\mathbb{T})$ of the process $\{y(t)\}_{t\in\mathbb{Z}}$ from the data $\{y(t_k)\}_{k=1}^N$. Henceforth, we will denote by $\mathscr{S}_+^m(\mathbb{T})$ the space of $m \times m$ functions that are well-defined and positive definite on $\mathbb{T}$, that is the space of coercive and bounded discrete-time spectral densities (when $m = 1$, we let $\mathscr{S}_+^1(\mathbb{T}) := \mathscr{S}_+(\mathbb{T})$ to simplify the notation). A THREE-like approach hinges on the following four ingredients:

1) A rational filter $G(z)$ to process the data. Here, $G(z)$ models a bank of filters of the form

$$G(z) = (zI - A)^{-1}B, \quad A \in \mathbb{C}^{n\times n}, \ B \in \mathbb{C}^{n\times m}, \tag{2.1}$$

where A is Schur stable, *i.e.*, all the eigenvalues of A are strictly inside $\mathbb{T}$, the pair (A, B) is reachable, and B is of full column rank ($n \geq m$).

2) An estimate $\Sigma \in \mathbb{H}_n$ of the steady-state covariance of the output state $\{x(t)\}_{t\in\mathbb{Z}}$ of the filter whose dynamics is described by

$$x(t+1) = Ax(t) + By(t), \quad t \in \mathbb{Z}.$$

3) An *a priori* estimate $\Psi \in \mathscr{S}_+^m(\mathbb{T})$ of the spectral density Φ.

4) A distance-like function between spectral densities $d\colon \mathscr{S}_+^m(\mathbb{T}) \times \mathscr{S}_+^m(\mathbb{T}) \to \mathbb{R}_{\geq 0}$.

On the one hand, the possibility of selecting the filter bank (2.1) makes this approach very flexible. In fact, $G(z)$ can be designed by the user in order to guarantee higher resolution within certain frequency bands. On the other hand, the same filter also imposes a *generalized moment constraint* on the available covariance estimate of the form

$$\int G\Phi G^* := \int_{-\pi}^{\pi} G(e^{j\theta})\Phi(e^{j\theta})G^*(e^{j\theta})\frac{\mathrm{d}\theta}{2\pi} = \Sigma. \tag{2.2}$$

Here, and throughout the remaining of the dissertation, we will use the above shorthand where integration, unless otherwise specified, takes place on $[-\pi,\pi]$ with respect to normalized Lebesgue measure $\mathrm{d}\theta/2\pi$. Within this setting, the estimation problem boils down to solving the following constrained approximation problem.

Problem 2.1 (THREE-like estimation)**.** Let $\Psi \in \mathscr{S}_+^m(\mathbb{T})$ and $\Sigma \in \mathbb{H}_{+,n}$. Find $\hat{\Phi} \in \mathscr{S}_+^m(\mathbb{T})$ that solves

$$\min_{\Phi\in\mathscr{K}} d(\Phi,\Psi),$$

where

$$\mathscr{K} := \left\{ \Phi \in \mathscr{S}_+^m(\mathbb{T}) : \int G\Phi G^* = \Sigma \right\}.$$

Remark 2.1. It is worth noting that the estimation of the covariance matrix Σ from the filtered data $\{x(t_i)\}_{i=1}^N$ in the above setting is a most delicate issue that has been addressed in Ramponi et al. [2009], Ferrante et al. [2012b], Zorzi and Ferrante [2012]. In the first paper the problem was solved using a projection-based technique, in the second paper via a maximum-entropy method, and in the third one by means of an approach that exploits an estimate of the covariance lags sequence of the process $\{y(t)\}_{t\in\mathbb{Z}}$.

Remark 2.2. The index $d(\cdot,\cdot)$ plays a key role in the above problem. Indeed, a suitable choice of this index should lead to a tractable solution of the corresponding dual problem of Problem 2.1 and the latter solution should ideally be rational and of low complexity (as measured by the McMillan degree), at least when the prior Ψ is such. We just mention here that this choice is particularly crucial in the multivariate case. This point will be discussed in greater detail in Chapter 4 and Chapter 5.

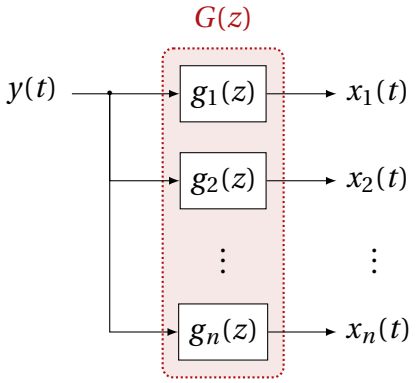
Remark 2.3. The *a priori* estimate Ψ is usually taken to be a coarse estimate of the to-be-estimated spectrum Φ. In case this information is not available, the most reasonable choice is to set $\Psi = I_m$.

The existence of $\Phi \in \mathscr{S}_+^m(\mathbb{T})$ satisfying (2.2) and, hence the existence of solutions to Problem 2.1, is a non-trivial issue that may be viewed as a (multivariate) generalized moment problem. Before addressing this issue, we illustrate how we can recover several important problems of Systems and Control Theory as special cases of Problem 2.1.

Example 2.1 (Covariance extension problem). Consider the scalar case $m = 1$ and pick

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

With this choice, the k-th component of the filter $G(z)$ is $g_k(z) = z^{n-k}$ (see also the block diagram below).



In addition, let

$$\Sigma = \begin{bmatrix} c_0 & c_1 & 0 & c_{n-1} \\ \bar{c}_1 & c_0 & \cdots & c_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{c}_{n-1} & \bar{c}_{n-2} & \cdots & c_0 \end{bmatrix}$$

where

$$c_k := \mathbb{E}\{y(k)\bar{y}(n-k)\}$$

which corresponds to the k-th covariance lag of $\{y(t)\}_{t\in\mathbb{Z}}$. Therefore, Problem 2.1 consists in finding a suitable extension $\{c_k\}_{k>n}$ of the partial covariance lags sequence

$\{c_k\}_{k=0}^{n-1}$ such that

$$\Phi(e^{j\theta}) = \sum_{k=-\infty}^{\infty} c_k e^{-j\theta k}, \quad \forall\, \theta \in [-\pi, \pi].$$

This is the maximum-entropy *covariance extension problem* Georgiou [1987a], Byrnes and Lindquist [1997], Byrnes et al. [1995, 1997, 1998], Carli et al. [2011], Lindquist and Picci [2013], Lindquist et al. [2013]. It is known that the set of spectral densities consistent with the data is nonempty if $\Sigma \geq 0$ and contains infinitely many elements if $\Sigma > 0$.

Example 2.2 (Nevanlinna–Pick interpolation). For the sake of simplicity, let us suppose again that $\{y(t)\}_{t\in\mathbb{Z}}$ is scalar and consider

$$A = \begin{bmatrix} p_1 & 0 & 0 & \cdots & 0 \\ 0 & p_2 & 0 & \cdots & 0 \\ 0 & 0 & \ddots & & \vdots \\ \vdots & \vdots & & p_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & p_n \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}, \quad |p_i| < 1,\ i = 1, \dots, n.$$

In view of this choice, the k-th element of the filter bank is

$$g_k(z) = \frac{z}{z - p_k}, \quad |p_k| < 1.$$

The state covariance can be shown to have the form of a Pick matrix

$$\Sigma = \begin{bmatrix} \frac{w_1+\bar{w}_1}{1-p_1\bar{p}_1} & \frac{w_1+\bar{w}_2}{1-p_1\bar{p}_2} & \cdots & \frac{w_1+\bar{w}_n}{1-p_1\bar{p}_n} \\ \frac{w_2+\bar{w}_1}{1-p_2\bar{p}_1} & \frac{w_2+\bar{w}_2}{1-p_2\bar{p}_2} & \cdots & \frac{w_2+\bar{w}_n}{1-p_2\bar{p}_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{w_n+\bar{w}_1}{1-p_n\bar{p}_1} & \frac{w_n+\bar{w}_2}{1-p_n\bar{p}_2} & \cdots & \frac{w_n+\bar{w}_n}{1-p_n\bar{p}_n} \end{bmatrix}$$

with

$$w_k = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{e^{-j\theta} + p_k}{e^{j\theta} - p_k} \Phi(e^{j\theta})\, \mathrm{d}\theta, \quad k = 1, \dots, n.$$

In this case, Problem 2.1 turns into a classical *Nevanlinna–Pick interpolation problem* (see, *e.g.*, [Doyle et al., 1992, Chap. 9] for a general overview, and Georgiou [1999], Byrnes et al. [2001b, 2006] for further results). As in the previous example, the set of solutions is nonempty if $\Sigma \geq 0$ and contains infinitely many elements if $\Sigma > 0$.

2.1.1 Existence of solutions

The first issue to worry about in solving Problem 2.1 concerns the existence of solutions. To this end, let $\mathscr{C}(\mathbb{T},\mathbb{H}_m)$ be the space of $m \times m$ Hermitian matrix-valued continuous function and consider the following linear operator

$$\begin{aligned} \Gamma\colon \mathscr{C}(\mathbb{T},\mathbb{H}_m) &\to \mathbb{H}_n \\ \Phi &\mapsto \int G\Phi G^*. \end{aligned} \tag{2.3}$$

The following proposition (which is an assembly of results from Georgiou [2002], Ferrante et al. [2008]) provides a series of equivalent conditions for the existence of a solution to Problem 2.1.

Proposition 2.1. *The following conditions are equivalent:*

1) *The set of solutions to Problem 2.1 is nonempty.*

2) *There exists $H \in \mathbb{C}^{m\times n}$ such that*

$$\Sigma - A\Sigma A^* = BH + H^*B^*.$$

3) *The following rank condition holds*

$$\operatorname{rank}\begin{bmatrix} \Sigma - A\Sigma A^* & B \\ B^* & \mathbf{0} \end{bmatrix} = 2m.$$

4) *It holds $\Sigma \in \operatorname{Range}\Gamma$.*

5) *The following relation holds*

$$(I_n - \Pi_{B'})(I_n - A'A'^*)(I_n - \Pi_{B'}) = 0,$$

where $B' := \Sigma^{-1/2}B$, $A' := \Sigma^{-1/2}A\Sigma^{-1/2}$, and $\Pi_{B'} := B'(B'^*B')^{-1}B'^*$ denotes the orthogonal projection onto $\operatorname{im}(B')$.

2.2 The scalar Kullback–Leibler case

In this section, following Georgiou and Lindquist [2003], we focus on the particular case of Problem 2.1 in which:

1) the spectral densities are scalar ($m = 1$), and

2) $d(\cdot,\cdot)$ is chosen to be the Kullback–Leibler divergence, namely

$$d(\Phi,\Psi) \equiv \mathbb{D}(\Phi\|\Psi) := \int \Psi \log\frac{\Psi}{\Phi}.$$

This distance-like function arises in information-theoretic settings where probability distributions are evaluated in terms of their entropy or the amount of information they contain [Cover and Thomas, 2012, Ch. 2]. It is known under a variety names such as *relative entropy, cross entropy, information divergence, information gain, etc.*, and applied in many areas of information theory, statistics, probability and estimation theory. With this choice, Problem 2.1 reads as follows.

Problem 2.2 (THREE-like Kullback–Leibler estimation)**.** Let $\Psi \in \mathscr{S}_+(\mathbb{T})$ and $\Sigma \in \mathbb{H}_n$, $\Sigma > 0$. Find $\hat{\Phi} \in \mathscr{S}_+(\mathbb{T})$ that solves

$$\min_{\Phi \in \mathscr{K}} \mathbb{D}(\Phi, \Psi),$$

where

$$\mathscr{K} := \left\{ \Phi \in \mathscr{S}_+(\mathbb{T}) : \int G \Phi G^* = \Sigma \right\}.$$

[20] It can be shown that this condition is met for sure if A is singular.

Remark 2.4. Notice that, if the spectra Φ and Ψ have the same zero-th moment[20], *i.e.*,

$$\int \Phi = \int \Psi,$$

then $\mathbb{D}(\Phi \| \Psi) \geq 0$. The justification of the choice of $\mathbb{D}(\Phi \| \Psi)$ even for spectra that do not satisfy the previous condition is discussed in [Georgiou and Lindquist, 2003, Sec. III]. This choice basically hinges on the possibility of rescaling Ψ. This implies that in the corresponding optimization problem we are approximating the "shape" of the *a priori* spectral density Ψ. We point out that in several applications requiring the discrimination of two spectra, *e.g.*, in speech processing, what really matters is indeed the "shape" of the spectra rather then their relative scaling. In recent year, this aspect has stimulated the search for sensible *projective* distances in the cone of spectral densities Martin [2000], Georgiou [2007a,b]. This point will be also discussed in Chapter 5.

Remark 2.5. In the majority of applications arising from statistics, probability, information theory, *etc.*, minimizing $\mathbb{D}(\Phi \| \Psi)$ instead of $\mathbb{D}(\Psi \| \Phi)$ is rather unusual. Here the choice is justified by the fact that it leads to a more tractable problem (as we shall see in what follows), and, more importantly, with this choice the maximum entropy estimate can be recovered as a particular case ($\Psi = 1$).

Remark 2.6. As described in [Ferrante et al., 2011, Remark 2.2], without loss of generality, we can assume that $\Sigma = I_n$ and $\int \Psi = 1$. As a matter of fact, if $\Sigma \neq I_n$, then we can replace (A, B) with $(\Sigma^{-1/2} A \Sigma^{-1/2}, \Sigma^{-1/2} B)$ to get an equivalent problem where $\Sigma = I_n$. Likewise, if $\int \Psi = \alpha \neq 1$, we can replace Ψ with Ψ/α and G with $\sqrt{\alpha} G$ to obtain an equivalent version of the problem in which $\int \Psi = 1$.

2.2.1 Variational analysis and dual problem

We outline here the variational analysis of Georgiou and Lindquist [2003] leading to the optimal conditions of Problem 2.2 and to its dual (we refer also to Pavon and Ferrante [2005, 2006], Ferrante et al. [2011] where some additional details and differences are spelled out). In what follows, we will suppose that the spectral densities are continuous functions and $\Sigma = I_n$ (*cf.* Remark 2.6).

Let us define

$$\mathscr{L}_+ := \{\Lambda \in \mathbb{H}_n : G^*\Lambda G > 0,\ \forall \theta \in [-\pi, \pi]\}.$$

Now pick $\Lambda \in \mathscr{L}_+$ and consider the *Lagrangian functional*

$$\begin{aligned} L(\Phi, \Lambda) &= \mathbb{D}(\Psi \| \Phi) + \operatorname{tr}\left(\Lambda\left(\int G\Phi G^* - I_n\right)\right) \\ &= \mathbb{D}(\Psi \| \Phi) + \int G^*\Lambda G\Phi - \operatorname{tr}\Lambda. \end{aligned}$$

Next, we consider the unconstrained minimization of $L(\Phi, \Lambda)$, namely

$$\min_{\Phi \in \mathscr{S}_+(\mathbb{T})} L(\Phi, \Lambda). \tag{2.4}$$

Notice that, for each fixed Λ, $\Phi \mapsto L(\Phi, \Lambda)$ is a strictly convex function. Hence, the problem in (2.4) is a unconstrained convex optimization problem. In addition, we have the following optimality condition [Georgiou and Lindquist, 2003, Theorem 5].

Theorem 2.1. *Suppose that the one of the feasibility conditions in Proposition 2.1 is met and $\hat{\Lambda} \in \mathscr{L}_+$ satisfies*

$$\int G \frac{\Psi}{G^*\hat{\Lambda}G} G^* = I_n. \tag{2.5}$$

Then

$$\hat{\Phi} = \frac{\Psi}{G^*\hat{\Lambda}G} \tag{2.6}$$

is the unique solution of Problem 2.2.

Remark 2.7. Observe that, in view of the above theorem, the solution of the approximation is unique and has a form parametrized by a suitable Hermitian matrix $\hat{\Lambda}$. Nevertheless, this does not implies that the matrix $\hat{\Lambda}$ leading to this solution is unique. In fact, it is apparent that all the Hermitian matrices $\hat{\Lambda}'$ satisfying

$$G^*(\hat{\Lambda} - \hat{\Lambda}')G = 0, \quad \forall \theta \in [-\pi, \pi],$$

yield the same solution of the approximation problem. The previous condition is equivalent to assert that $\hat{\Lambda} - \hat{\Lambda}' \in (\operatorname{Range}\Gamma)^\perp$, where $(\operatorname{Range}\Gamma)^\perp$ is the orthogonal

complement (w.r.t. the standard trace inner product in $\mathbb{H}_n$) of the range of the linear operator Γ as defined in (2.3). The latter can be expressed by [Ferrante et al., 2008, Sec. IV]

$$(\text{Range}\,\Gamma)^{\perp} = \{X \in \mathbb{H}_n : G^* X G = 0,\ \forall \theta \in [-\pi, \pi]\}. \tag{2.7}$$

Thus, the set of all solutions of Problem 2.2 is the affine space $\{\hat{\Lambda} + X,\ X \in (\text{Range}\,\Gamma)^{\perp}\}$.

Thanks to Theorem 2.1 the original approximation problem is now reduced to finding $\Lambda \in \mathscr{L}_+$ satisfying (2.5). This can be accomplished via duality theory. Consider the *dual functional*

$$\Lambda \mapsto \inf_{\Phi \in \mathscr{S}_+(\mathbb{T})} L(\Phi, \Lambda).$$

For $\Lambda \in \mathscr{L}_+'$, the latter takes the form

$$\Lambda \mapsto L\left(\frac{\Psi}{G^*\Lambda G}, \Lambda\right) = \int \Psi \log G^* \Lambda G - \text{tr}(\Lambda) + \int \Psi.$$

Next, consider the maximization of the dual functional over the set $\mathscr{L}_+$. By defining

$$\mathbb{J}(\Lambda) := -\int \Psi \log G^* \Lambda G + \text{tr}(\Lambda), \tag{2.8}$$

the *dual problem* is then equivalent to

$$\min_{\Lambda \in \mathscr{L}_+} \mathbb{J}(\Lambda). \tag{2.9}$$

This problem is again a convex optimization problem. In Georgiou and Lindquist [2003], Λ is further restricted to belong to $\text{Range}\,\Gamma$. On this restricted domain, the dual functional becomes strictly convex and $\mathbb{J}(\cdot)$ has a unique minimum point as shown in [Georgiou and Lindquist, 2003, Sec. V] (see also Ferrante et al. [2007] for a different proof of this fact).

2.2.2 Some numerical considerations

When restricted to $\tilde{\mathscr{L}}_+ := \mathscr{L}_+ \cap \text{Range}\,\Gamma$ the dual functional in $\mathbb{J}(\cdot)$ is strictly convex, so that the dual problem in (2.9) can be solved using an iterative algorithm based on Netwon's method. In order to apply this method, however, a suitable parametrization of the elements in $\tilde{\mathscr{L}}_+$ that preserves global convexity is needed. In what follows we briefly outline the parametrization considered in [Georgiou and Lindquist, 2003, Sec. VII].

Given $\Lambda \in \tilde{\mathscr{L}}_+$, let M be the unique solution of the Lyapunov equation

$$M = A^* M A + \Lambda.$$

It holds

$$G^*\Lambda G = \left(G - \frac{1}{2}B\right)^* K + K^*\left(G - \frac{1}{2}B\right),$$

where $K := MB$. Let $\mathscr{K}_+$ the space of all $K \in \mathbb{C}^{n\times 1}$ such that $G^*\Lambda G > 0$ on $\mathbb{T}$. The dual problem (2.9) can be shown to be equivalent to minimizing the convex functional

$$\hat{\mathbb{J}}(K) := HK + K^*H^* - \int \Psi \log\left(\left(G - \frac{1}{2}B\right)^* K + K^*\left(G - \frac{1}{2}B\right)\right)$$

over $\mathscr{K}_+$, where $H \in \mathbb{C}^{1\times n}$ is the matrix satisfying the existence condition 2) of Proposition 2.1. The above-defined functional $\hat{\mathbb{J}} : \mathscr{K}_+ \to \mathbb{R}$ is strictly convex, and one can apply a Newton-like method as the one described in Byrnes et al. [2000]. However, using this parametrization, the gradient of the dual functional is unbounded in the neighborhood of the boundary of $\mathscr{K}_+$ and this causes serious numerical difficulties in practical implementations.

To conclude, we mention that an alternative parametrization has been proposed at the end of [Georgiou and Lindquist, 2003, Sec. VII] which is better behaved on the boundary but destroys global convexity of the problem. In this case, an homotopy continuation method similarly to the one of Enqvist [2001], Nagamune [2003] can be used to solve the problem.

3. On the convergence of the Pavon–Ferrante algorithm

As mentioned in the previous chapter, gradient-based numerical methods for the solution of the scalar Kullback–Leibler estimation problem (Problem 2.2) are severely affected by a number of numerical issues. These are mainly due to the unboundedness of the gradient around the boundary of the set wherein to look for a solution. This implies that such methods typically do not behave numerically well and may require a significant number of back-stepping iterations before reaching convergence. To enhance numerical stability and reduce computational burden, in Pavon and Ferrante [2005, 2006] an alternative iterative method based on a fixed-point iteration in the space of trace-one positive semi-definite matrices was introduced. On the one hand, in case of rational prior, the Pavon–Ferrante algorithm can be efficiently implemented in a numerically robust way via the solution of an Algebraic Riccati Equation and a Lyapunov equation [Ferrante et al., 2011, Sec. IV]. On the other hand, in spite of a huge amount of numerical evidences, showing the global convergence of the algorithm to a prescribed set of fixed points which provide the solution of Problem 2.2 has revealed to be an highly non-trivial challenge. This open problem has been subsequently invesigated in Ferrante et al. [2007, 2011]. In the first work, the authors showed that the iteration can be seen as a modified gradient descent method with fixed step size. In the second work, a rather tortuous yet enlightening proof of local convergence towards the prescribed set of fixed points of the iteration was established. Both these results however were not completely satisfactory since they do not provide an answer to the more challenging question of *global* convergence of the iteration. In this chapter, we fill this gap by analyzing the global convergence properties of the Pavon–Ferrante iteration. Specifically, after introducing the algorithm in Section 3.1 and collecting some preliminary results in Section 3.2, Section 3.3 presents a proof of *global* convergence of the Pavon–Ferrante algorithm to one of its fixed point. In Section 3.4 we then discuss how to modify the algorithm in order to guarantee convergence to a *positive definite* fixed point. Importantly, such a fixed point is always guaranteed to yield the solution of Problem 2.2. We conclude by illustrating a numerical example that demonstrates the effectiveness of the proposed approach.

Except for the last section, this chapter is based on the material in Baggio [2017a]. We also mention that a global convergence analysis for a "discretized" version of the Pavon–Ferrante iteration has been discussed in Baggio [2017b].

3.1 The algorithm and its properties

THE Pavon–Ferrante algorithm, as introduced in Pavon and Ferrante [2005, 2006], is a fixed-point iteration of the form

$$\Lambda_{k+1} = \Theta(\Lambda_k) := \int \Lambda_k^{1/2} G \left(\frac{\Psi}{G^* \Lambda_k G} \right) G^* \Lambda_k^{1/2}, \quad k \in \mathbb{Z},\ k \geq 0, \tag{3.1}$$

where the initialization is taken to be a positive definite Λ_0 in $\mathfrak{S}_n$ and the latter symbol denote the set of unit-trace positive semi-definite matrices. Let us define

$$\mathscr{M} := \{ \Lambda \in \mathfrak{S}_n : G^*(e^{j\theta}) \Lambda G(e^{j\theta}) > 0,\ \forall \theta \in [-\pi, \pi) \}, \tag{3.2}$$

$$\mathscr{M}_+ := \{ \Lambda \in \mathscr{M} : \Lambda > 0 \} \subset \mathscr{M}. \tag{3.3}$$

Iteration (3.1) features several interesting properties. One of these properties is that the map $\Theta(\cdot)$ preserves positivity and unit-trace as described by the following result.

Proposition 3.1 ([Pavon and Ferrante, 2006, Thm 4.1])**.** *Iteration* $\Theta(\cdot)$ *maps elements of* $\mathscr{M}$ *(*$\mathscr{M}_+$*) into elements of* $\mathscr{M}$ *(*$\mathscr{M}_+$*, respectively).*

Another fundamental property is the following one. If iteration (3.1) converges to a *positive definite* fixed point of $\Theta(\cdot)$, say $\hat{\Lambda} > 0$, then

1) $G^*(e^{j\theta}) \hat{\Lambda} G(e^{j\theta}) > 0$, $\forall\, \theta \in [-\pi, \pi)$, and

2) by multiplying Equation (3.1) on both sides by $\hat{\Lambda}^{-1/2}$,

$$\int G \frac{\Psi}{G^* \hat{\Lambda} G} G^* = I_n. \tag{3.4}$$

In view of Theorem 2.1, the previous two points readily imply that such a $\hat{\Lambda}$ yields the solution of Problem 2.2 via Equation (2.6). Importantly, such a fixed point always exists. In fact, let $\mathscr{S}$ denote the space of $\Lambda \in \mathbb{H}_n$ satisfying the optimality conditions in Theorem 2.1. Then, we have the following result.

Proposition 3.2 ([Ferrante et al., 2011, Thm 3.2])**.** *The set of positive definite fixed points of iteration* (3.1) *is a nonempty open convex set* $\mathscr{P}$ *of the space* $\mathscr{S}$*.*

To conclude, we remark that the positive definite ones are not the only fixed points of iteration (3.1) which provide the solution to Problem 2.2. Indeed, in the closure of $\mathscr{P}$ there exist singular elements which still satisfy the aforementioned optimality conditions and, therefore, solve Problem 2.2 via (2.6), see [Ferrante et al., 2011, Sec. II-B]. In general, however, singular fixed points are not guaranteed to satisfy the required optimality conditions and, therefore, to solve Problem 2.2 via (2.6).

We conclude this section with a characterization of a class of singular fixed points of $\Theta(\cdot)$ and with two concrete examples of fixed points that illustrate the above described "special" situations.

Proposition 3.3 ([Pavon and Ferrante, 2006, Prop. 4.3])**.** *Let $x \in \mathbb{C}^n$, $x \neq 0$, and let $P_x := (xx^*)/(x^*x)$ denote the orthogonal projection rank-one projection onto* span$\{x\}$, *$x \in \mathbb{C}^n$. Then $P_x \in \mathfrak{S}_n$ is a fixed point of $\Theta(\cdot)$.*

Example 3.1. Consider $\Psi = 1$, $G(e^{j\theta}) = (e^{j\theta} I - A)^{-1} B$, where

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Note that A is (strictly) Schur stable and the pair (A, B) is reachable. Observe also that the feasibility condition in Proposition 2.1, point 3), is met. Indeed, by taking $H = [0, 1/2]$, it holds $I_2 - AA^* = BH + H^*B^*$. Let $P_x := (xx^*)/(x^*x)$ denote the orthogonal projection onto span$\{x\}$, $x \in \mathbb{C}^n$, and consider $x = [4, 3]^\top$. From Proposition 3.3 P_x is a fixed point of the iteration, however

$$\int \frac{GG^*}{G^*P_xG} = \begin{bmatrix} \frac{25}{7} & -\frac{75}{28} \\ -\frac{75}{28} & \frac{25}{7} \end{bmatrix} \neq I_2$$

so that condition (3.4) is not met and therefore P_x does not lead to the solution of Problem 2.2.

Example 3.2. Consider the quantities defined in the previous example and $x = [1, 0]^\top$. In this case, it holds

$$\int \frac{GG^*}{G^*P_xG} = I_2$$

so that condition (3.4) is met although the fixed point P_x is singular. Moreover, in this case, it holds $G(e^{j\theta})^* P_x G(e^{j\theta}) > 0$ for all $\theta \in [-\pi, \pi)$. From these two facts it follows that P_x yields the solution of the spectral estimation problem via Equation (2.6).

3.2 Auxiliary results

In this section, we collect some auxiliary results which will be used in the proof of the main theorem presented in the next section. The first result is a consequence of Jensen's inequality.

Lemma 3.1. *Let $X \subseteq \mathbb{R}$. Consider an integrable function $f\colon X \to \mathbb{R}_{>0}$ and an integrable function $w\colon X \to \mathbb{R}_{>0}$ satisfying $\int_X w(x)\,\mathrm{d}x = 1$, then*

$$\log \int_X w(x) f(x)\,\mathrm{d}x \geq \int_X w(x) \log f(x)\,\mathrm{d}x,$$

and the equality is attained if and only if $f(x)$ is constant a.e. on X.

Proof. It follows from Jensen's inequality [Cover and Thomas, 2012, Thm. 2.6.2] applied to the strictly concave function $\log(\cdot)$. □

Another ancillary lemma is stated and proved below.

Lemma 3.2. *Let $X \subseteq \mathbb{R}$. Let $w\colon X \to \mathbb{R}$ and $f\colon X \times X \to \mathbb{R}_{\geq 0}$ be integrable functions with $f(\cdot,\cdot)$ symmetric,* i.e., *$f(x,y) = f(y,x)$ for all $x, y \in X$. Then it holds*

$$\int_X \int_X w(x)^2 f(x,y)\,\mathrm{d}x\,\mathrm{d}y \geq \int_X \int_X w(x)\,w(y) f(x,y)\,\mathrm{d}x\,\mathrm{d}y. \tag{3.5}$$

Proof. Thanks to the symmetry of $f(\cdot,\cdot)$, we can rewrite the left hand-side of (3.5) as

$$\begin{aligned}\int_X \int_X w(x)^2 f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_X \int_X w(y)^2 f(y,x)\,\mathrm{d}y\,\mathrm{d}x \\ &= \int_X \int_X w(y)^2 f(x,y)\,\mathrm{d}x\,\mathrm{d}y,\end{aligned}$$

so that

$$\int_X \int_X w(x)^2 f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_X \int_X \left(\frac{w(x)^2 + w(y)^2}{2}\right) f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Now, since $w(x)^2 + w(y)^2 \geq 2\,w(x)\,w(y)$, due to the fact that $(w(x) - w(y))^2 \geq 0$, for all $x, y \in X$, the claim follows. □

Next we focus the attention on the cost functional $\mathbb{J}(\cdot)$, as defined in Equation (2.8). This function will play a key role in the convergence analysis of the next section.

Lemma 3.3. *$\mathbb{J}(\cdot)$ is a convex, bounded and continuous function on $\mathfrak{S}_n$.*

Proof. We first note that convexity readily follows from the fact that $-\log(\cdot)$ is convex. Next, in view of the stability of A and reachability of the pair (A, B), for all $\Lambda \in \mathfrak{S}_n$, $G^*\Lambda G$ is a nonzero rational spectral density analytic on (an open annulus containing) $\mathbb{T}$. This in turn implies that $\log G^*\Lambda G$ is integrable on $\mathbb{T}$, see, *e.g.*, [Rozanov, 1967, p. 64]. Since Ψ is bounded on $\mathbb{T}$, $\Psi \log G^*\Lambda G$ is again integrable on $\mathbb{T}$. This in turn implies that $\mathbb{J}(\cdot)$ is bounded on $\mathfrak{S}_n$. Now let $\bar{\Lambda} \in \mathfrak{S}_n$ and consider any sequence $\{\Lambda_k\}_{k\geq 0}$, in $\mathfrak{S}_n$ such that $\lim_{k\to\infty} \Lambda_k = \bar{\Lambda}$. The corresponding sequence $\{G^*\Lambda_k G\}_{k\geq 0}$ is composed of nonzero rational spectral densities analytic on $\mathbb{T}$ and such that $\lim_{k\to\infty} G^*\Lambda_k G = G^*\bar{\Lambda}G$ uniformly on $\mathbb{T}$, where the limit $G^*\bar{\Lambda}G$ is a spectral density as before. In fact, by denoting with $g_{i,\theta}$ the i-th entry of $G(e^{j\theta})$ and by $\lambda_{k,ij}$, $\bar{\lambda}_{ij}$ the (i,j)-th entry of Λ_k, $\bar{\Lambda}$,

respectively, uniform convergence follows from the fact that $\lim_{k\to\infty}\Lambda_k = \bar{\Lambda}$ and

$$\begin{aligned}\left|G^*(e^{j\theta})\Lambda_k G(e^{j\theta}) - G^*(e^{j\theta})\bar{\Lambda}G(e^{j\theta})\right| &= \left|\sum_{i,j}\lambda_{k,ij}g^*_{i,\theta}g_{j,\theta} - \sum_{i,j}\bar{\lambda}_{ij}g^*_{i,\theta}g_{j,\theta}\right| \\ &= \left|\sum_{i,j}(\lambda_{k,ij}-\bar{\lambda}_{ij})g^*_{i,\theta}g_{j,\theta}\right| \\ &\le \sum_{i,j}\left|\lambda_{k,ij}-\bar{\lambda}_{ij}\right|\,|g^*_{i,\theta}g_{j,\theta}| \qquad (3.6)\\ &\le G_{\max}\sum_{i,j}\left|\lambda_{k,ij}-\bar{\lambda}_{ij}\right|, \quad \forall\,\theta\in[-\pi,\pi),\end{aligned}$$

where $G_{\max} := \max_{i,j,\theta}|g^*_{i,\theta}g_{j,\theta}| < \infty$ since each $g_{i,\theta}$ is analytic and bounded on $\mathbb{T}$ and (3.6) uses the triangle inequality. Hence, from [Nurdin, 2006, Corollary 4.6], it follows that the sequence $\{\log G^*\Lambda_k G\}_{k\ge 0}$ is uniformly integrable on $\mathbb{T}$. Eventually, since Ψ is bounded on $\mathbb{T}$, $\{\Psi\log G^*\Lambda_k G\}_{k\ge 0}$ is again uniformly integrable, so that by Vitali's convergence theorem [Rudin, 1987, p. 133] it holds

$$\begin{aligned}\lim_{k\to\infty}\mathbb{J}(\Lambda_k) &= 1 - \lim_{k\to\infty}\int\Psi\log G^*\Lambda_k G \\ &= 1 - \int\lim_{k\to\infty}\Psi\log G^*\Lambda_k G \\ &= 1 - \int\Psi\log G^*\bar{\Lambda}G = \mathbb{J}(\bar{\Lambda}),\end{aligned}$$

which proves continuity. □

We observe that $\mathbb{J}(\cdot)$ is not injective on $\mathfrak{S}_n$. As a matter of fact, consider the orthogonal complement (w.r.t. the trace inner product in $\mathbb{H}_n$) of $\operatorname{Range}\Gamma$, where the linear operator Γ has been defined in Equation (2.3). In view of Equation (2.7), given $\Lambda\in\mathfrak{S}_n$ and any non-zero matrix $\Lambda^\perp\in(\operatorname{Range}\Gamma)^\perp$ such that $\Lambda+\Lambda^\perp\ge 0$, it follows that $\mathbb{J}(\Lambda) = \mathbb{J}(\Lambda+\Lambda^\perp)$. In addition, $\Lambda+\Lambda^\perp\in\mathfrak{S}_n$, since every element in $(\operatorname{Range}\Gamma)^\perp$ is traceless [Ferrante et al., 2011, Sec. II].

At this point, we analyze the behavior of $\mathbb{J}(\cdot)$ in the region of the boundary of $\mathfrak{S}_n$ defined by

$$\mathcal{N} := \left\{\Lambda\in\mathfrak{S}_n : \exists\bar{\theta}\in[-\pi,\pi) \text{ s.t. } G^*(e^{j\bar{\theta}})\Lambda G(e^{j\bar{\theta}}) = 0\right\}. \qquad (3.7)$$

The following lemma provides a useful result in this regard.

Lemma 3.4. *For all $\bar{\Lambda}\in\mathcal{N}$, the (right-sided) directional derivative of $\mathbb{J}(\cdot)$ at $\bar{\Lambda}$ along any direction $\Delta\bar{\Lambda}\in\mathbb{H}_n$ such that $\bar{\Lambda}+\Delta\bar{\Lambda}\in\mathscr{M}$ takes the value $-\infty$.*

Proof. Let $\bar{\Lambda} \in \mathcal{N}$. First, we note that $\bar{\Lambda} + \varepsilon\Delta\bar{\Lambda} \in \mathcal{M}$ for all $\Delta\bar{\Lambda} \in \mathbb{H}_n$ such that $\bar{\Lambda} + \Delta\bar{\Lambda} \in \mathcal{M}$, and for all $\varepsilon \in (0,1]$. The (right-sided) directional derivative of $\mathbb{J}(\cdot)$ at $\bar{\Lambda}$ in the direction $\Delta\bar{\Lambda}$ is given by

$$\begin{aligned}
\nabla\mathbb{J}(\bar{\Lambda};\Delta\bar{\Lambda}) &:= \lim_{\varepsilon\to 0^+} \frac{\mathbb{J}(\bar{\Lambda}+\varepsilon\Delta\bar{\Lambda}) - \mathbb{J}(\bar{\Lambda})}{\varepsilon} \\
&= \lim_{\varepsilon\to 0^+} \left(\frac{1}{\varepsilon}\operatorname{tr}(\bar{\Lambda}+\varepsilon\Delta\bar{\Lambda}) - \frac{1}{\varepsilon}\operatorname{tr}(\bar{\Lambda}) - \frac{1}{\varepsilon}\int \Psi \log \frac{G^*(\bar{\Lambda}+\varepsilon\Delta\bar{\Lambda})G}{G^*\bar{\Lambda}G} \right) \\
&= -\lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}\int \Psi \log\left(1+\varepsilon\frac{G^*\Delta\bar{\Lambda}G}{G^*\bar{\Lambda}G}\right) \\
&= -\int \Psi \frac{G^*\Delta\bar{\Lambda}G}{G^*\bar{\Lambda}G},
\end{aligned}$$

where we exploited the fact that $\operatorname{tr}(\Delta\bar{\Lambda}) = 0$, and, in the last step, the Taylor expansion of $\log(1+x)$. Eventually, since (i) $\bar{\Lambda} \in \mathcal{N}$, and (ii) $\Delta\bar{\Lambda}$ is such that $\bar{\Lambda} + \Delta\bar{\Lambda} \in \mathcal{M}$, there exists (at least) a frequency $\bar{\theta} \in [-\pi,\pi)$ such that $G^*(e^{j\bar{\theta}})\bar{\Lambda}G(e^{j\bar{\theta}}) = 0$ and $G^*(e^{j\bar{\theta}})\Delta\bar{\Lambda}G(e^{j\bar{\theta}}) \neq 0$. Since $\Psi(e^{j\theta}) > 0$ for all $\theta \in [-\pi,\pi)$, this in turn implies

$$\int \Psi \frac{G^*\Delta\bar{\Lambda}G}{G^*\bar{\Lambda}G} = \infty,$$

which yields the thesis. □

Remark 3.1. Lemma 3.4 provides a characterization of the elements of $\mathcal{N}$ in terms of directional derivatives of $\mathbb{J}(\cdot)$ along directions belonging to the subset of $\mathfrak{S}_n$ given by $\mathcal{M}$. Notice that this result typically does not characterize *all* directional derivatives of $\mathbb{J}(\cdot)$ evaluated at elements in $\mathcal{N}$ along directions pointing to $\mathfrak{S}_n$. However, for a particular subset of $\mathcal{N}$ this is indeed the case. Let $x \in \mathbb{C}^n$, $\|x\| = 1$, and let $P_x := xx^*$ denote the orthogonal projection onto the subspace spanned by x. Consider the following subset of $\mathcal{N}$

$$\begin{aligned}
\mathcal{N}_0 := \Big\{ & P_{\bar{x}} \in \mathfrak{S}_n : \exists \bar{\theta} \in [-\pi,\pi),\ \bar{x} \in \mathbb{C}^n,\ \|\bar{x}\| = 1,\ \text{s.t.} \\
& \text{(i)}\ \left\langle \bar{x}, G(e^{j\bar{\theta}}) \right\rangle = 0,\ \text{and} \\
& \text{(ii)}\ \left\langle x, G(e^{j\bar{\theta}}) \right\rangle \neq 0,\ \forall\, x \in \mathbb{C}^n,\ \|x\| = 1,\ x \neq \bar{x} \Big\},
\end{aligned} \tag{3.8}$$

where $\langle A, B\rangle = \operatorname{tr}(AB^*)$ for $A, B \in \mathbb{C}^{n\times n}$. By exploiting the same argument of the proof of Lemma 3.4, it follows that the directional derivative of $\mathbb{J}(\cdot)$ evaluated at $\bar{\Lambda} \in \mathcal{N}_0$ along *any* direction $\Delta\bar{\Lambda} \in \mathbb{H}_n \setminus \{0\}$ such that $\bar{\Lambda} + \Delta\bar{\Lambda} \in \mathfrak{S}_n$ is unbounded below. Moreover, it is worth noticing that, for the particular case $n = 2$, it holds $\mathcal{N}_0 \equiv \mathcal{N}$.

To conclude this section, we recall the discrete-time version of LaSalle's invariance principle, whose proof can be found in [LaSalle, 1986, Prop. 2.6] or, with reference to the continuous-time case, in [Khalil, 1996, Theorem 4.4].

Proposition 3.4 (Discrete-time LaSalle's invariance principle)**.** *Consider a discrete-time system*

$$x(t+1) = f(x(t)), \quad x(0) \in \mathscr{X},\ t \geq 0,$$

where $f : \mathscr{X} \to \mathscr{X}$ is continuous and $\mathscr{X}$ is an invariant and compact set. Suppose $V(\cdot)$ is a continuous function of $x \in \mathscr{X}$, bounded below and satisfying

$$\Delta V(x) := V(f(x)) - V(x) \leq 0, \quad \forall x \in \mathscr{X},$$

that is $V(x)$ is non-increasing along (forward) trajectories of the dynamics. Then any trajectory converges to the largest invariant subset $\mathscr{I}$ contained in

$$\mathscr{E} := \{x \in \mathscr{X} : \Delta V(x) = 0\}.$$

3.3 The main convergence result

In this section, we present the main results of this chapter. The first result (Theorem 3.1) states that the cost function (2.8) is always non-increasing along the trajectories of (3.1). This result provides a positive answer to a conjecture raised in the conclusive part of Ferrante et al. [2011] and paves the way for a global convergence proof based on LaSalle's principle.

Theorem 3.1. *For every $\Lambda \in \mathfrak{S}_n$ it holds*

$$\Delta \mathbb{J}(\Lambda) := \mathbb{J}(\Theta(\Lambda)) - \mathbb{J}(\Lambda) \leq 0, \tag{3.9}$$

where $\mathbb{J}(\cdot)$ has been defined in (2.8). Moreover $\Delta \mathbb{J}(\Lambda) = 0$ if and only if $\Theta(\Lambda) = \Lambda + \Lambda^{\perp}$, with $\Lambda^{\perp} \in (\operatorname{Range} \Gamma)^{\perp}$, where Γ is the linear operator defined in Equation (2.3).

Proof. For the sake of clarity, in what follows we will use the following shorthand: Given a matrix-valued function defined on the unit circle we let $F_\theta := F(e^{j\theta})$. By plugging the expression of $\Theta(\cdot)$ into (3.9), we get

$$\begin{aligned} & \Delta \mathbb{J}(\Lambda) = \mathbb{J}(\Theta(\Lambda)) - \mathbb{J}(\Lambda) \leq 0 \\ \iff & \int \Psi_\theta \log G_\theta^* \Theta(\Lambda) G_\theta - \int \Psi_\theta \log G_\theta^* \Lambda G_\theta \geq 0 \\ \iff & \int \Psi_\theta \log f_\theta \geq 0, \end{aligned} \tag{3.10}$$

where $f_\theta := \frac{G_\theta^* \Theta(\Lambda) G_\theta}{G_\theta^* \Lambda G_\theta}$ is well-defined and strictly positive on $\mathbb{T}$, since $\Theta(\Lambda)$ has the same rank and kernel of $\Lambda \in \mathfrak{S}_n$, *cf.* [Ferrante et al., 2011, Prop. 2.1].

Firstly, we notice that the following inequality holds

$$\int \Psi_\theta \log f_\theta = -2 \int \Psi_\theta \log f_\theta^{-1/2} \geq -2 \log \int \Psi_\theta f_\theta^{-1/2}, \tag{3.11}$$

which is a consequence of Lemma 3.1.

Secondly, by defining $\Pi_\theta := \frac{\Lambda^{1/4} G_\theta G_\theta^* \Lambda^{1/4}}{G_\theta^* \Lambda G_\theta}$, we have the following chain of equations

$$\begin{aligned}
1 &= \int \Psi_\theta f_\theta^{-1} f_\theta = \int \int f_\theta^{-1} \Psi_\theta \Psi_\omega \langle \Pi_\theta, \Pi_\omega \rangle && (3.12)\\
&\geq \int \int f_\theta^{-1/2} f_\omega^{-1/2} \Psi_\theta \Psi_\omega \langle \Pi_\theta, \Pi_\omega \rangle && (3.13)\\
&= \left\langle \int \Psi_\theta f_\theta^{-1/2} \Pi_\theta, \int \Psi_\omega f_\omega^{-1/2} \Pi_\omega \right\rangle \\
&= \|\Lambda^{1/2}\|_{\mathrm{F}}^2 \left\| \int \Psi_\theta f_\theta^{-1/2} \Pi_\theta \right\|_{\mathrm{F}}^2 && (3.14)\\
&\geq \left| \left\langle \Lambda^{1/2}, \int \Psi_\theta f_\theta^{-1/2} \Pi_\theta \right\rangle \right|^2 && (3.15)\\
&= \left(\int \Psi_\theta f_\theta^{-1/2} \right)^2 && (3.16)
\end{aligned}$$

where

- Equation (3.12) follows by noticing that $f_\theta = \int \Psi_\omega \langle \Pi_\theta, \Pi_\omega \rangle$,
- Equation (3.13) uses Lemma 3.2 applied to the symmetric function $\Psi_\theta \Psi_\omega \langle \Pi_\theta, \Pi_\omega \rangle$,
- Equation (3.14) exploits the fact that $\|\Lambda^{1/2}\|_{\mathrm{F}}^2 = \mathrm{tr}(\Lambda) = 1$,
- Equation (3.15) follows from Cauchy–Schwarz inequality, and
- Equation (3.16) uses the fact that $\langle \Lambda^{1/2}, \Pi_\theta \rangle = 1$.

Eventually, a combination of the two sets of inequalities (3.11) and (3.12)-(3.16) yields $\int \Psi_\theta \log f_\theta \geq 0$ which, in turn, implies $\Delta \mathbb{J}(\Lambda) \leq 0$, in view of equivalence (3.10).

Now it remains to prove that we attain equality in (3.9) if only if $\Lambda \in \mathfrak{S}_n$ is such that $\Theta(\Lambda) = \Lambda + \Lambda^\perp$ with $\Lambda^\perp \in (\mathrm{Range}\,\Gamma)^\perp$. In view of the definition of $(\mathrm{Range}\,\Gamma)^\perp$ given in Equation (2.7), the "if" part becomes straightforward. Indeed, if $\Theta(\Lambda) = \Lambda + \Lambda^\perp$, we have

$$\begin{aligned}
\Delta \mathbb{J}(\Lambda) &= \int \Psi_\theta \log \frac{G_\theta^* \Theta(\Lambda) G_\theta}{G_\theta^* \Lambda G_\theta} = \int \Psi_\theta \log \frac{G_\theta^* (\Lambda + \Lambda^\perp) G_\theta}{G_\theta^* \Lambda G_\theta} \\
&= \int \Psi_\theta \log \frac{G_\theta^* \Lambda G_\theta}{G_\theta^* \Lambda G_\theta} = 0.
\end{aligned}$$

So it remains to prove the "only if" part, *i.e.*, if equality in (3.9) is attained for Λ then $\Theta(\Lambda) = \Lambda + \Lambda^{\perp}$ with $\Lambda^{\perp} \in (\mathrm{Range}\,\Gamma)^{\perp}$. To this end, we notice that a necessary condition for (3.9) to hold with equality is to have (3.11) satisfied with equality. By Lemma 3.1, this implies that the function f_θ is constant for every $\theta \in [-\pi, \pi)$, namely

$$f_\theta = \frac{G_\theta^* \Theta(\Lambda) G_\theta}{G_\theta^* \Lambda G_\theta} = \kappa, \quad \forall\, \theta \in [-\pi, \pi),$$

where $\kappa > 0$ is a real constant. Now equality in (3.9) is attained (if and) only if $\kappa = 1$, and therefore we have that

$$G_\theta^* \Theta(\Lambda) G_\theta = G_\theta^* \Lambda G_\theta, \quad \forall\, \theta \in [-\pi, \pi).$$

From the latter equation and by definition of $(\mathrm{Range}\,\Gamma)^{\perp}$ in Equation (2.7), it follows that $\Theta(\Lambda) = \Lambda + \Lambda^{\perp}$, $\Lambda^{\perp} \in (\mathrm{Range}\,\Gamma)^{\perp}$. This completes the proof. □

The following theorem is based on the previous one and states that iteration (3.1) always converges to one of the fixed points of $\Theta(\cdot)$.

Theorem 3.2. *The trajectories generated by iteration* (3.1) *converge for all* $\Lambda_0 \in \mathfrak{S}_n$ *to elements belonging to* $\mathscr{F} := \{\Lambda \in \mathfrak{S}_n : \Theta(\Lambda) = \Lambda\}$.

Proof. The proof consists of an application of the discrete-time version of LaSalle's invariance principle (Proposition 3.4). The natural candidate Lyapunov function $V(\cdot)$ of Proposition 3.4 is given in this case by $\mathbb{J}(\Lambda)$ which is continuous and bounded for every $\Lambda \in \mathfrak{S}_n$ (Lemma 3.3), and, by virtue of Theorem 3.1, non-increasing along the (forward) trajectories of the dynamics (3.1). Hence, by LaSalle's invariance principle, we have that the (forward) trajectories generated by iteration (3.1) converges to the largest invariant set $\mathscr{I}$ contained in

$$\mathscr{E} := \{\Lambda \in \mathfrak{S}_n : \Delta\mathbb{J}(\Lambda) = 0\}.$$

Therefore, it remains to show that the trajectories in $\mathscr{I}$ consist of fixed points of $\Theta(\cdot)$ only, that is, $\mathscr{I} \equiv \mathscr{F}$. To this end, by Theorem 3.1, we know that the elements $\Lambda \in \mathscr{E}$ satisfy the condition

$$\Theta(\Lambda) = \Lambda + \Lambda^{\perp}, \tag{3.17}$$

with $\Lambda^{\perp} \in (\mathrm{Range}\,\Gamma)^{\perp}$. In view of the latter constraint on the dynamics (3.1) and the definition of $(\mathrm{Range}\,\Gamma)^{\perp}$ in Equation (2.7), it follows that any trajectory belonging to $\mathscr{E}$ must obey to the recurrence relation

$$\Lambda_{k+1} = \Lambda_k^{1/2} M \Lambda_k^{1/2}, \quad \Lambda_0 \in \mathscr{E},\ k \geq 0, \tag{3.18}$$

where

$$M := \int \Psi \frac{GG^*}{G^* \Lambda_0 G},$$

depends on the initial condition Λ_0 only, in view of Equation (3.17). Now, since (3.18) must generate unit trace trajectories starting from any $\Lambda_0 \in \mathfrak{S}_n$, we have $\mathrm{tr}(\Lambda_0) = \mathrm{tr}(\Lambda_1) = \mathrm{tr}(\Lambda_2) = 1$. By exploiting the cyclic property and the linearity of the trace, this in turn implies that

$$\begin{aligned}
&\mathrm{tr}(\Lambda_0) - 2\mathrm{tr}(\Lambda_1) + \mathrm{tr}(\Lambda_2) = \mathrm{tr}(\Lambda_0) - 2\mathrm{tr}(\Lambda_0^{1/2} M \Lambda_0^{1/2}) + \mathrm{tr}(M\Lambda_0^{1/2} M \Lambda_0^{1/2}) \\
&= \mathrm{tr}(\Lambda_0) - \mathrm{tr}(\Lambda_0^{3/4} M \Lambda_0^{1/4}) - \mathrm{tr}(\Lambda_0^{1/4} M \Lambda_0^{3/4}) + \mathrm{tr}(\Lambda_0^{1/4} M \Lambda_0^{1/2} M \Lambda_0^{1/4}) \\
&= \mathrm{tr}\left(\Lambda_0^{1/4}(I - M)\Lambda_0^{1/4}\right)^2 \\
&= \left\| \Lambda_0^{1/4}(I - M)\Lambda_0^{1/4} \right\|_{\mathrm{F}}^2 = 0.
\end{aligned}$$

The previous equation is satisfied if and only if $\Lambda_0^{1/4}(I - M)\Lambda_0^{1/4} = 0$, or, equivalently, if and only if

$$\Lambda_0 = \Lambda_0^{1/2} M \Lambda_0^{1/2}.$$

From the latter equation and Equation (3.18) it readily follows that Λ_0 must be a fixed point of $\Theta(\cdot)$. This ends the proof. □

As a final result of this section, we characterize a whole family of fixed points of $\Theta(\cdot)$ that are *not* asymptotically stable. The following result provides a partial answer to another conjecture of [Ferrante et al., 2011, Sec. V] claiming that orthogonal rank-one projections which do not belong to the closure of the set of positive definite fixed points $\mathscr{P}$ are unstable equilibrium points of $\Theta(\cdot)$.

Proposition 3.5. *The set $\mathscr{N}_0$ defined in Equation (3.8) consists of fixed points that are not asymptotically stable for the dynamics* (3.1).

Proof. Let $\bar{\Lambda} \in \mathscr{N}_0$. Notice that $\bar{\Lambda}$ is a rank-one orthogonal projection so that $\bar{\Lambda}$ is a fixed point of $\Theta(\cdot)$ by Proposition 3.3. Now observe that, in view of Lemma 3.4 and Remark 3.1, all the (right-sided) directional derivatives at $\mathbb{J}(\bar{\Lambda})$ along directions pointing to $\mathfrak{S}_n$ take the value $-\infty$. This implies that in a sufficiently small neighbourhood U of $\bar{\Lambda}$, it holds $\mathbb{J}(\bar{\Lambda}) > \mathbb{J}(\Lambda)$, $\forall\, \Lambda \in U \cap \mathfrak{S}_n$, $\Lambda \neq \bar{\Lambda}$. In light of this, the claim follows from the fact that $\mathbb{J}(\cdot)$ is non-increasing along trajectories of the dynamics (3.1) (Theorem 3.1). □

3.4 A modified version of the algorithm

THE main result of the previous section (Theorem 3.2) showed that the Pavon–Ferrante algorithm globally converges to one of its fixed points in $\mathfrak{S}_n$. This result however does not tell whether this fixed point also leads to the solution of Problem 2.2. As previously described, this is automatically verified if the fixed point is positive definite.

The purpose of this section is to introduce and analyze the convergence properties of a new algorithm that can be seen as a modified version of the Pavon–Ferrante iteration. The proposed iteration is proved to be globally convergent to a *positive definite* fixed point of the Pavon–Ferrante agorithm. Consequently, it always globally converges to a point which yields the solution of Problem 2.2. Further, we show that this algorithm is amenable to an efficient and numerically robust implementation in case the prior is rational.

3.4.1 Preliminaries

The following preliminary results will be crucial for the proof in the next section.

Proposition 3.6. *Suppose that the sequence* $\{\Lambda_k\}_{k\geq 0}$ *generated by iteration* (3.1) *converges to an element* $\bar{\Lambda} \in \mathscr{N}$. *Then there exists* $N > 0$ *such that for every* $k \geq N$ *there is a* $\hat{\Lambda}_k \in \mathfrak{S}_n$ *satisfying*

$$\hat{\Lambda}_k \geq \mu I, \quad \mu > 0, \quad \textit{and} \tag{3.19}$$

$$\mathbb{J}(\hat{\Lambda}_k) - \mathbb{J}(\Lambda_k) \leq \nu, \quad \nu < 0. \tag{3.20}$$

Proof. Given $\Lambda \in \mathfrak{S}_n$, we define[21]

$$\theta_{\min} := \arg\min_{\theta} G(e^{j\theta})^* \Lambda G(e^{j\theta}).$$

[21] Observe that $G^*\Lambda G$ could have several minima (even an uncountable number, for instance if $G^*\Lambda G$ is flat). If this is actually the case, we pick one of these minima.

Let $\Pi_{\mathscr{N}} \in \mathbb{H}_n$ denote the orthogonal projection onto span$\{G(e^{j\theta_{\min}})\}$. We have that

$$\Lambda_{\mathscr{N}} := \frac{\Pi_{\mathscr{N}}^{\perp} \Lambda \Pi_{\mathscr{N}}^{\perp}}{\operatorname{tr}(\Pi_{\mathscr{N}}^{\perp} \Lambda \Pi_{\mathscr{N}}^{\perp})},$$

with $\Pi_{\mathscr{N}}^{\perp} := I - \Pi_{\mathscr{N}}$, is such that $G(e^{j\theta_{\min}})^* \Lambda_{\mathscr{N}} G(e^{j\theta_{\min}}) = 0$, and therefore $\Lambda_{\mathscr{N}} \in \mathscr{N}$. Associate to any element of the sequence Λ_k, $k \geq 0$, a corresponding projected term $\Lambda_{k,\mathscr{N}}$, $k \geq 0$, as above. Since the sequence $\{\Lambda_k\}_{k\geq 0} \xrightarrow{k\to\infty} \bar{\Lambda} \in \mathscr{N}$ and $\{G^*\Lambda G\}_{k\geq 0}$ uniformly converges to the continuous spectral density $G^*\bar{\Lambda}G$ featuring a finite number of zeros on $\mathbb{T}$, it follows that

$$\{\Lambda_{k,\mathscr{N}}\}_{k\geq 0} \xrightarrow{k\to\infty} \bar{\Lambda} \in \mathscr{N},$$

that is, $\bar{\Lambda}$ is an accumulation point for $\{\Lambda_{k,\mathscr{N}}\}_{k\geq 0}$.

Next, by virtue of Lemma 3.4, we have that the (right-sided) directional derivative of the cost functional $\mathbb{J}(\cdot)$ along directions evaluated at every element of $\mathscr{N}$ and pointing to $\mathscr{M}$ takes the value $-\infty$. Since $\mathbb{J}(\cdot)$ is decreasing along trajectories generated by the iteration, this implies that given $\Lambda \in \mathscr{N}$, for all $\gamma \in (0,1]$, $\hat{\Lambda} := \Lambda + \gamma\left(\frac{I}{n} - \Lambda\right) \geq$

$\mu_{\Lambda,\gamma} I$, so that there always exists $\bar{\gamma} \in (0,1]$ such that $\mathbb{J}(\hat{\Lambda}) - \mathbb{J}(\Lambda) \leq \nu_{\lambda,\bar{\gamma}}$. Further, since $\mathcal{N}$ is compact, we can define

$$\mu := \min_{\Lambda \in \mathcal{N}} \mu_{\Lambda,\bar{\gamma}} > 0, \quad \bar{\nu} := \min_{\Lambda \in \mathcal{N}} \nu_{\Lambda,\bar{\gamma}} < 0;$$

Since $\bar{\Lambda}$ is an accumulation point for both $\{\Lambda_k\}_{k\geq 0}$ and $\{\Lambda_{\mathcal{N},k}\}_{k\geq 0}$ and $\mathbb{J}(\cdot)$ is a continuous function, for all $\varepsilon > 0$ there exists N such that for all $k \geq N$

$$|\mathbb{J}(\Lambda_k) - \mathbb{J}(\Lambda_{\mathcal{N},k})| < \varepsilon.$$

By defining $\hat{\Lambda}_k = \Lambda_{\mathcal{N},k} + \bar{\gamma}\left(\frac{I}{n} - \Lambda_{\mathcal{N},k}\right)$ and $\nu = \bar{\nu} + \varepsilon$, we have that (3.19)-(3.20) are met. □

Remark 3.2. As described in the proof of the previous proposition the term $\hat{\Lambda}_k$ can be computed as

$$\hat{\Lambda}_k = \Lambda_{\mathcal{N},k} + \gamma\left(\frac{I}{n} - \Lambda_{\mathcal{N},k}\right),$$

where the step size $\gamma > 0$ must be properly chosen, in order to satisfy conditions (3.19)-(3.20).

In order to select γ, one can adopt the following iterative method: Start from a fixed pre-selected value of γ and then halve it until the following condition is met

$$\mathbb{J}(\hat{\Lambda}_k) - \mathbb{J}(\Lambda_k) = \int \Psi \log \Upsilon < 0, \tag{3.21}$$

where $\Upsilon := \frac{G^* \Lambda_k G}{G^* \hat{\Lambda}_k G}$ is a rational function analytic on (an open annulus containing) $\mathbb{T}$. Notice that, since $\mathbb{J}(\Lambda)$ is a continuous, convex function of Λ, a sufficient condition for (3.21) to hold is that the (right-sided) directional derivative of $\mathbb{J}(\cdot)$ evaluated at $\hat{\Lambda}_k$, along the direction pointing to Λ_k is positive, namely

$$\nabla \mathbb{J}(\hat{\Lambda}_k; \Lambda_k) := \lim_{\varepsilon \to 0^+} \frac{\mathbb{J}(\hat{\Lambda}_k + \varepsilon(\Lambda_k - \hat{\Lambda}_k)) - \mathbb{J}(\hat{\Lambda}_k)}{\varepsilon} > 0. \tag{3.22}$$

An explicit calculation of the latter derivative yields (*cf.* the proof of Lemma 3.4)

$$\nabla \mathbb{J}(\hat{\Lambda}_k; \Lambda_k) = 1 - \int \Psi \Upsilon$$

Hence, checking (3.22) boils down to check whether $\int \Psi \Upsilon < 1$. In case of rational Ψ, this can be efficiently carried out by exploiting the residue theorem [Ahlfors, 1953, p. 155]. Indeed, in this case, it holds

$$\int \Psi \Upsilon = \sum_{i:|p_i|<1} \mathrm{Res}_i, \tag{3.23}$$

with Res_i being the residue of the pole p_i of the rational function $\Psi(z)\Upsilon(z)/z$.

Proposition 3.7. *For every $\Lambda \in \mathscr{M}_c$ where $\mathscr{M}_c$ is any compact set in $\mathscr{M}$ there exists $\Lambda^\perp \in (\operatorname{Range}\Gamma)^\perp$ such that $\hat{\Lambda} := \Lambda + \Lambda^\perp \geq \varepsilon_c I$, $\varepsilon_c > 0$, and $\mathbb{J}(\hat{\Lambda}) = \mathbb{J}(\Lambda)$.*

Proof. As done in the proof of Theorem 3.1, for the sake of clarity, in what follows we will use the shorthand $F_\theta := F(e^{j\theta})$ for a matrix-valued function defined on $\mathbb{T}$. The proof goes along the lines of [Ferrante et al., 2011, Lemma 3.1 and Theorem 3.2] and is constructive. Considers $\Lambda \in \mathscr{M}_c$. From [Ferrante et al., 2011, Lemma 3.1], there exists a vector $v \in \mathbb{C}^{n\times 1}$ such that

$$G_\theta^* \Lambda G_\theta = G_\theta^* v v^* G_\theta, \quad \forall\, \theta \in [-\pi, \pi). \tag{3.24}$$

On the unit circle $G_\theta^* \Lambda G_\theta$ is continuous and (strictly) positive and $G_\theta^* G_\theta$ is continuous. Thus,

$$\begin{aligned} \mu_\Lambda &:= \min_{\theta\in[-\pi,\pi)} G_\theta^* \Lambda G_\theta > 0, \\ \nu &:= \max_{\theta\in[-\pi,\pi)} G_\theta^* G_\theta < \infty. \end{aligned}$$

By defining $\varepsilon_\Lambda := \mu_\Lambda / \kappa\nu > 0$, $\kappa > 2$, we have

$$\begin{aligned} G_\theta^* \left(\frac{1}{2}\Lambda - \varepsilon_\Lambda I\right) G_\theta &\geq \frac{\mu_\Lambda}{2} - \frac{\mu_\Lambda}{\kappa\nu}\nu \\ &= \frac{\kappa - 2}{2\kappa}\mu_\Lambda > 0, \end{aligned}$$

for all $\theta \in [-\pi, \pi)$. Hence, by invoking again [Ferrante et al., 2011, Lemma 3.1], there exists a vector $\bar{v} \in \mathbb{C}^{n\times 1}$ such that for all $\theta \in [-\pi, \pi)$

$$G_\theta^* \left(\frac{1}{2}\Lambda - \varepsilon_\Lambda I\right) G_\theta = G_\theta^* \bar{v} \bar{v}^* G_\theta. \tag{3.25}$$

Consequently, for all $\theta \in [-\pi, \pi)$, it holds

$$\begin{aligned} G_\theta^* \Lambda G_\theta &= G_\theta^* \left(\frac{1}{2}\Lambda + \varepsilon_\Lambda I\right) G_\theta + G_\theta^* \left(\frac{1}{2}\Lambda - \varepsilon_\Lambda I\right) G_\theta \\ &= G_\theta^* \left(\frac{1}{2} v v^* + \varepsilon_\Lambda I\right) G_\theta + G_\theta^* \bar{v} \bar{v}^* G_\theta \\ &= G_\theta^* \left(\frac{1}{2} v v^* + \varepsilon_\Lambda I + \bar{v} \bar{v}^*\right) G_\theta. \end{aligned}$$

Then, by defining

$$\Lambda^\perp := \frac{1}{2} v v^* + \varepsilon_\Lambda I + \bar{v} \bar{v}^* - \Lambda,$$

we have $\Lambda^\perp \in (\operatorname{Range}\Gamma)^\perp$ and $\Lambda + \Lambda^\perp \geq \varepsilon_\Lambda I$. Eventually, we can define

$$\varepsilon_c := \min_{\Lambda\in\mathscr{M}_c} \varepsilon_\Lambda,$$

and the latter is strictly greater than zero in view of the compactness of $\mathscr{M}_c$. □

Remark 3.3. It is worth observing that, once calculated the value of ε_Λ, the vectors v and $\bar{v}$ in (3.24) and (3.25) can be computed in closed-form through the stabilizing solution of a Riccati equation, as described in the proof of [Ferrante et al., 2011, Lemma 3.1]. This follows from the fact that for any $\Lambda \in \mathscr{M}$, the following (right) spectral factorization holds (*cf.* [Ferrante et al., 2011, Lemma A.1])

$$G^*(z)\Lambda G(z) = W_\Lambda^*(z)W_\Lambda(z),$$

where $W_\Lambda(z)$ can be explicitly expressed in the form

$$W_\Lambda(z) = (B^*PB)^{-1/2}B^*P(A(zI-A)^{-1}+I)B \tag{3.26}$$

and $P \in \mathbb{H}_n$ is the stabilizing solution of the discrete-time algebraic Riccati equation

$$\Pi = A^*\Pi A - A^*\Pi B(B^*\Pi B)^{-1}B^*\Pi A + \Lambda. \tag{3.27}$$

From this fact, it is a matter of direct computation to check that the vector

$$v := PB(B^*PB)^{-1/2} \in \mathbb{C}^n, \tag{3.28}$$

is such that $G^*(e^{j\theta})vv^*G(e^{j\theta}) = G^*(e^{j\theta})\Lambda G(e^{j\theta})$ for all $\theta \in [-\pi,\pi)$ (see also [Ferrante et al., 2011, Lemma 3.1]).

3.4.2 The new algorithm

The structure of the proposed algorithm goes as follows. We run iteration (3.1) until either we reach a limit point that is positive definite or a point Λ_k that is very close to be singular. In the first case we terminate the procedure, while in the second case we distinguish two situations, namely

1) if Λ_k is sufficiently close to $\mathscr{N}$, we compute $\hat{\Lambda}_k$ as in Proposition 3.6. Such $\hat{\Lambda}_k$ is bounded away from singularity and is such that the cost functional $\mathbb{J}(\cdot)$ strictly decreases. Next, we restart the iteration (3.1) initialized at $\hat{\Lambda}_k$;

2) otherwise, Λ_k is converging to an element contained in $\mathscr{M}$. In this case, we compute a term $\hat{\Lambda}_k$ as in Proposition 3.7 and we restart iteration (3.1) with initialization $\hat{\Lambda}_k$.

The next key result asserts that the proposed algorithm is globally convergent to a positive definite fixed point.

Theorem 3.3. *For every initialization $\Lambda_0 \in \mathfrak{S}_n$, the above described algorithm converges to a positive definite fixed point of $\Theta(\cdot)$.*

Proof. From Theorem 3.2 we know that the fixed-point iteration converges to a fixed point of $\Theta(\cdot)$ in $\mathfrak{S}_n$. Clearly, if the iteration automatically converges to a positive definite fixed point, the thesis immediately follows. If the iteration converges to a singular fixed point, the latter belongs either to $\mathscr{N}$ or to $\mathscr{M}$. In the first case, by virtue of Proposition 3.6, when Λ_k is sufficiently close to $\mathscr{N}$, we can always find a term $\hat{\Lambda}$ such that $\hat{\Lambda}$ is bounded away from being singular and the cost functional $\mathbb{J}(\cdot)$ is decreasing by a quantity that is bounded away from zero. In the second case, the iteration is converging to a point that belongs to a compact subset of $\mathscr{M}$ defined by the complement in $\mathfrak{S}_n$ of a suitable open neighbourhood of $\mathscr{N}$. Hence, by virtue of Proposition 3.7, we can find a term $\hat{\Lambda}$ which is bounded away from singularity and leaves invariant the cost functional $\mathbb{J}(\cdot)$. In either case, by restarting the iteration with the new initialization $\hat{\Lambda}$, we have that $\mathbb{J}(\cdot)$ is always decreasing along the trajectories of the algorithm and, at the same time, is bounded away from the boundary of $\mathfrak{S}_n$. Hence, asymptotic convergence to a positive definite fixed point of $\Theta(\cdot)$ is guaranteed by LaSalle's invariance principle, as described in the proof of Theorem 3.2. □

The numerical analysis outlined in [Ferrante et al., 2011, Sec. IV] for the case of a rational prior spectral density Ψ and the observations made in Remark 3.2 and Remark 3.3 can be exploited to design an efficient numerical implementation of the above introduced algorithm. In what follows, we will illustrate this implementation, which is also schematically reported in Algorithm 1 at the end of this subsection.

As initialization step, we pick any $n \times n$ unit-trace positive definite matrix, and we fix to some small values the parameters δ_1, δ_2, $\varepsilon > 0$. We suppose that Ψ is rational. At each step of the algorithm, we evaluate the condition $\|\Theta(\Lambda) - \Lambda\|_{\mathrm{F}} \geq \delta_1$. If the latter condition is met, we apply the Pavon–Ferrante iteration (3.1) as it is. Otherwise, we have two possible cases. If the current iterate Λ is "sufficiently" bounded away from singularity ($\Lambda \geq \delta_2 I$), then we stop the algorithm. If not, we employ one of the following "correction" strategies and then restart the iteration:

- In case the iteration is approaching the region $\mathscr{N}$, *i.e.*, $\min_\theta G^* \Lambda G < \varepsilon$, then we compute a correction term $\hat{\Lambda}$ as in Proposition 3.6 and Remark 3.2, namely

$$\hat{\Lambda} = \Lambda_{\mathscr{N}} + \gamma \left(\frac{I}{n} - \Lambda_{\mathscr{N}} \right), \quad \gamma > 0,$$

  where

$$\Lambda_{\mathscr{N}} := \frac{\Pi_{\mathscr{N}}^{\perp} \Lambda \Pi_{\mathscr{N}}^{\perp}}{\operatorname{tr}(\Pi_{\mathscr{N}}^{\perp} \Lambda \Pi_{\mathscr{N}}^{\perp})}, \tag{3.29}$$

  and $\Pi_{\mathscr{N}}$ denotes the orthogonal projection onto $\operatorname{span}\{G(e^{j\theta_{\min}})\}$, where $\theta_{\min}$ is a frequency minimizing the coercive spectral density $G^* \Lambda G$.

- Otherwise, the iteration is approaching the region $\mathscr{M}$, so that we add a correction term $\Lambda^{\perp}$ to the current iterate of the form described in Proposition 3.7 and Remark 3.3, namely

$$\Lambda^{\perp} := \frac{1}{2} v v^* + \varepsilon_{\Lambda} I + \bar{v}\bar{v}^* - \Lambda, \tag{3.30}$$

with v, $\bar{v} \in \mathbb{C}^n$ such that, for all $\theta \in [-\pi, \pi)$,

$$G^* v v^* G = G^* \Lambda G, \tag{3.31}$$

$$G^* \bar{v}\bar{v}^* G = G^* \left(\frac{1}{2}\Lambda - \varepsilon_{\Lambda} I \right) G, \tag{3.32}$$

and

$$\varepsilon_{\Lambda} = \frac{\| W_{\Lambda}^{-1}(z) \|_{\mathscr{H}_\infty}^{-2}}{\kappa \| W_I(z) \|_{\mathscr{H}_\infty}^2}, \tag{3.33}$$

where $\kappa > 2$, $W_{\#}(z)$ denotes the minimum-phase (right) spectral factor of $G^*(\#)G$, as defined in (3.26), and $\|\cdot\|_{\mathscr{H}_\infty}$ denotes the $\mathscr{H}_\infty$-norm of a discrete-time transfer function. We recall that for a stable discrete-time transfer function $T(z)$ the latter norm is defined as

$$\| T(z) \|_{\mathscr{H}_\infty} := \sup_{z \in \mathbb{C} : |z| > 1} \sigma_{\max}(T(z)),$$

where $\sigma_{\max}(\cdot)$ denotes the maximum singular value of a matrix.

We now analyze more in detail the numerical properties of the above outlined implementation. As shown in [Ferrante et al., 2011, Sec. IV], running the Pavon–Ferrante iteration (line 3 of Algorithm 1) essentially requires the calculation of:

1) the stabilizing solution of the ARE of order n in (3.27);

2) the solution of a discrete-time Lyapunov equation of order $n + n_\Psi$, where n_Ψ denote the state-space dimension of a minimal spectral factor of Ψ_θ.

The first correction action in lines 6–15 of Algorithm 1 boils down to the computation of:

1) the following $\mathscr{H}_\infty$-norm

$$\min_{\theta} G^* \Lambda G = \| W_{\Lambda}^{-1}(z) \|_{\mathscr{H}_\infty}^{-2},$$

where $W_\Lambda(z)$ is the minimum-phase spectral factor in (3.26) (which in turn can be computed via the solution of ARE (3.27)), and a corresponding minimum frequency $\theta_{\min}$. This can be performed by means of several efficient numerical routines, see, *e.g.*, Boyd et al. [1989], Bruinsma and Steinbuch [1990];

2) the evaluation of the condition at line 10, which in turn can be reduced to the computation of the integral in (3.23), as described in Remark 3.2.

The second correction action of lines 17–19 of Algorithm 1 needs the computation of:

1) the value of ε_Λ in (3.33) that, in essence, consists of the computation of two AREs of the form in (3.27), and two $\mathscr{H}_\infty$-norms;

2) the vectors v and $\bar{v}$ that have the form in (3.28) and whose calculation requires the stabilizing solution of two AREs of the form in (3.27) (one of which has been already computed in the previous step).

In view of the previous analysis, we argue that all the required operations can be implemented by numerically efficient and robust routines.

Algorithm 1 Modified Pavon–Ferrante algorithm

1: Pick $\Lambda \in \mathfrak{S}_n$, $\Lambda > 0$, $\kappa > 2$, and sufficiently small $\varepsilon, \delta_1, \delta_2 \in \mathbb{R}_{>0}$.
2: **while** $\|\Theta(\Lambda) - \Lambda\|_F \geq \delta_1$ **do**
3: Set $\Lambda \leftarrow \Theta(\Lambda)$
4: **end while**
5: **if** $\Lambda \not> \delta_2 I$ **then**
6: Set $\gamma \in \mathbb{R}_{>0}$
7: **if** $\min_\theta G^* \Lambda G < \varepsilon$ **then**
8: Compute $\Lambda_{\mathscr{N}}$ as in (3.29)
9: Set $\hat{\Lambda} \leftarrow \Lambda_{\mathscr{N}} + \gamma(\frac{I}{n} - \Lambda_{\mathscr{N}})$
10: **while** $\mathbb{J}(\hat{\Lambda}) - \mathbb{J}(\Lambda) \geq 0$ **do**
11: Set $\gamma \leftarrow \gamma/2$
12: Set $\hat{\Lambda} \leftarrow \Lambda_{\mathscr{N}} + \gamma(\frac{I}{n} - \Lambda_{\mathscr{N}})$
13: **end while**
14: Set $\Lambda \leftarrow \hat{\Lambda}$
15: **go to** 2
16: **else**
17: Compute $\Lambda^\perp$ as in (3.30)
18: Set $\Lambda \leftarrow \Lambda + \Lambda^\perp$
19: **go to** 2
20: **end if**
21: **end if**
22: **return** Λ

3.4.3 A numerical example

We illustrate here a very simple example which demonstrates the effectiveness of the algorithm devised in the previous section. We consider the following 2×2 setup

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \Psi(z) = \frac{1.25}{(z+1.5)(z^{-1}+1.5)}. \tag{3.34}$$

Notice that A is Schur stable, (A, B) is a reachable pair, and $\Psi(e^{j\theta}) > 0$ for $\theta \in [-\pi, \pi]$. Further, it may be checked that the feasibility conditions in Proposition 2.1 are met, hence the solution to Problem 2.2 exists.

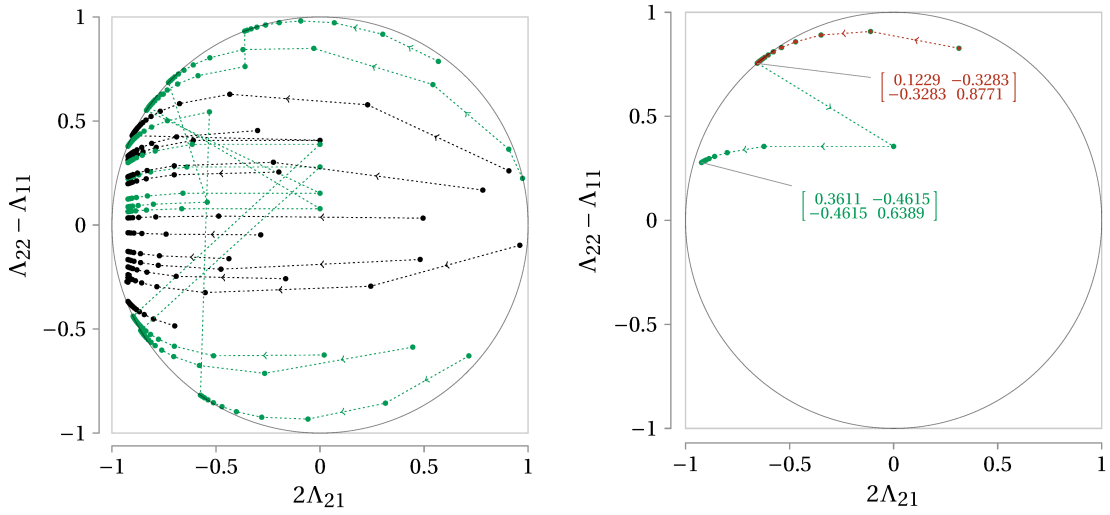


Figure 3.1: Examples of application of Algorithm 1 using the setup in (3.34), and values of the parameter $\kappa = 4$, $\varepsilon_1 = 10^{-4}$, $\delta_1 = 10^{-6}$, and $\delta_2 = 10^{-4}$. Each filled dot represents an iterate of the algorithm intialized at a random real positive definite unit-trace matrix.

We ran Algorithm 1 using the above setup starting from different real-valued positive definite unit-trace initializations.[22] The obtained results are shown in Figure 3.1. In these plots, we exploit the isomorphism

$$\varphi\colon \begin{bmatrix} a & b \\ b & c \end{bmatrix} \mapsto \frac{1}{\sqrt{2}} \begin{bmatrix} 2b \\ c-a \\ c+a \end{bmatrix}$$

which allows to identify the space of positive semi-definite 2×2 real matrices, namely

[22] If the initialization is a real matrix, it can be shown that Algorithm 1 generates real-valued trajectories.

$\mathbb{S}_2$, with the cone in $\mathbb{R}^3$

$$\left\{(x,y,z)\in\mathbb{R}^3 : z \geq \sqrt{x^2+y^2}\right\}.$$

Thus, the circle in the plots represents the section of this cone corresponding to unit-trace matrices, after a suitable normalization.

In the left plot, the black trajectories represent those trajectories that converge to a positive definite fixed point without requiring a "correction" action, whereas the green ones are those in which, at least, one "correction" step is applied. In the right plot, one particular trajectory is plotted. In this case, the red trajectory is obtained by running the Pavon–Ferrante algorithm without any "correction" step, and the green one by running Algorithm 1 using the same initialization. In the former case the iteration numerically converges to a "bad" singular matrix, in the sense that this matrix does *not* yield the solution of Problem 2.2. In the latter case the iteration converges to a positive definite matrix which automatically leads to the solution of Problem 2.2.

We stress that convergence to a "bad" singular fixed point for the Pavon–Ferrante iteration could be due either to a numerical issue[23] or to the fact that the iteration indeed is indeed not globally convergent to the (closure of the) set of positive definite fixed points. This fact represents still an open problem. In either case, by using the new algorithm proposed in the previous subsection convergence to a positive definite fixed point is always guaranteed, and this comes at a very little extra computational cost.

[23] More precisely, one possible issue could be due to numerical round-off errors in the computation of the square root of a nearly singular positive definite matrix.

4. Multivariate parametric estimation: An existence result

In contrast to the scalar case, the analysis and solution of THREE-like *multivariate* spectral estimation problems represent a much more difficult challenge. Indeed, in this case the feasibility of Problem 2.1 strongly depends on the selected distance-like index. We mention, in particular, the works Georgiou [2006], where a multivariate extension of the Kullback–Leibler divergence, the quantum relative entropy, is considered; Ferrante et al. [2008], Ramponi et al. [2009], which deal with a sensible generalization of the Hellinger distance; and Ferrante et al. [2012a], Georgiou and Lindquist [2017], where the selected distance index coincides with the multivariate Itakura–Saito distance, and the latter is shown to be the "correct" generalization of the Kullback–Leibler distance for the case of stochastic processes. It is worth remarking that the latter two approaches lead to rational solutions with bounded McMillan degree when the prior is rational. Further, Zorzi [2014a] and Zorzi [2015] introduce two more general frameworks based on the notion of Beta and Tau divergence families, wherein the multivariate Kullback–Leibler and Itakura–Saito distance can be recovered as special cases.

In Ferrante et al. [2010] an alternative approach to the problem was pursued. In that paper, the authors considered a *parametric* version of the THREE-like approach that does not need the definition of a distance index. The proposed parametric family was mainly motivated by the fact that: (i) for the particular case of no prior information ($\Psi = I$) the maximum-entropy solution belongs to this class, and (ii) the computation of a solution of the problem can be carried out by means of a multivariate version of the efficient fixed-point iteration discussed in the previous chapter. However, the result in Ferrante et al. [2010] was not satisfactory because the authors only showed that a parametric solution of the problem exists when the prior Ψ has a very special structure. In fact, this provides the motivation of the current chapter. As a continuation of the work in Ferrante et al. [2010], here we will show that a solution to the parametric spectral estimation problem exists given *any* fixed matrix-valued prior density that is bounded and coercive.

The main machinery behind our existence proof is the *topological degree theory* from non-linear analysis. As an historical remark, Georgiou was the first to apply the degree theory to rational covariance extension Georgiou [1983, 1987a,b] to show existence of a solution, and it was further developed by Byrnes, Lindquist, and coworkers Byrnes et al. [1995] to prove the uniqueness and well-posedness. These theories were established before the discovery of the cost function in the optimization framework Byrnes et al. [2001b, 1998, 2001c], which was later called generalized entropy criterion.

The present chapter builds upon Zhu and Baggio [2017] and is structured as follows. In Section 4.1 we introduce the parametric family of multivariate estimators described in Ferrante et al. [2010] and formulate the existence problem. In Section 4.2, we review some notions of topological degree theory that will be instrumental for the proof the main existence result presented in Section 4.3. To conclude, Section 4.4 illustrates the application of the main result to the special case of covariance extension.

4.1 A parametric family of estimators

THE problem setting is exactly the same of that discussed in Chapter 2. Let us define the set

$$\mathscr{L}_+ := \{\Lambda \in \mathbb{H}_n \,:\, G^*(z)\Lambda G(z) > 0,\ \forall z \in \mathbb{T}\}, \tag{4.1}$$

where $G(z)$ denotes the filter bank defined in (2.1). Notice that, by the continuous dependence of eigenvalues on the matrix entries, one can verify that $\mathscr{L}_+$ is an open subset of $\mathbb{H}_n$. Before introducing the parametric family of estimators that will be the object of our investigation we need the following lemma.[24]

[24] This result is stated without proof in Ferrante et al. [2010]. An actual proof of it, with a different notation, can be found in [Avventi, 2011, Sec. A.5.1].

Lemma 4.1 ([Ferrante et al., 2010, Lemma 11.4.1])**.** *Consider the filter bank $G(z)$ as defined in (2.1) and let $\Lambda \in \mathscr{L}_+$. Then, the discrete-time algebraic Riccati equation*

$$\Pi = A^*\Pi A - A^*\Pi B(B^*\Pi B)^{-1}B^*\Pi A + \Lambda,$$

admits a unique stabilizing solution $P_\Lambda \in \mathbb{H}_n$. The corresponding matrix $B^*P_\Lambda B$ is positive definite and the spectrum of closed loop matrix

$$Z_\Lambda := A - B(B^*P_\Lambda B)^{-1}B^*P_\Lambda A$$

lays in the open unit disk. Let L_Λ be the unique (upper triangular) right Cholesky factor of $B^*P_\Lambda B$, so that it holds $B^*P_\Lambda B = L_\Lambda^* L_\Lambda$. The following (right) spectral factorization holds

$$G^*\Lambda G = W_\Lambda^* W_\Lambda \tag{4.2}$$

where

$$W_\Lambda(z) := L_\Lambda^{-*}B^*P_\Lambda A(zI - A)^{-1}B + L_\Lambda,$$

is the unique stable and minimum-phase (right) spectral factor of $G^*\Lambda G$ such that $W_\Lambda(\infty)$ is lower triangular.

The following class of spectral estimator was introduced in Ferrante et al. [2010]

$$\Phi_\Lambda := W_\Lambda^{-1}\Psi W_\Lambda^{-*}, \quad \Lambda \in \mathscr{L}_+, \tag{4.3}$$

where $\Psi \in \mathscr{S}_+^m(\mathbb{T})$ is a coercive spectral density encoding prior information.

As mentioned in the beginning of the chapter, the main motivation behind this choice is that this class of estimators is suitable for the implementation of the following multivariate version of the Pavon–Ferrante algorithm

$$\Lambda_{k+1} = \int \Lambda_k^{1/2} G(W_{\Lambda_k}^{-1} \Psi W_{\Lambda_k}^{-*}) G^* \Lambda_k^{1/2}, \quad \Lambda_0 > 0, \tag{4.4}$$

which has proved to feature several interesting numerical properties. Additionally, the considered parametric class contains the maximum-entropy estimator when $\Psi = I$.

The first delicate issue to address when considering the class of estimators (4.3) concerns the existence of an estimate of this class compatible with the covariance data $\Sigma \in \operatorname{Range}\Gamma$ generated as in Chapter 2. This problem formally reads as follows.

Problem 4.1. Given the rational filter bank $G(z)$ in (2.1). Let $\Sigma \in \operatorname{Range}_+\Gamma$, where

$$\operatorname{Range}_+\Gamma := \operatorname{Range}\Gamma \cap \mathbb{H}_{+,n}, \tag{4.5}$$

and $\Psi \in \mathscr{S}_+^m(\mathbb{T})$. Find $\Lambda \in \mathscr{L}_+$ such that

$$\Phi_\Lambda := W_\Lambda^{-1} \Psi W_\Lambda^{-*}$$

satisfies

$$\int G \Phi_\Lambda G^* = \Sigma. \tag{4.6}$$

Define $\mathscr{L}_+^\Gamma := \mathscr{L}_+ \cap \operatorname{Range}\Gamma$, and consider the map

$$\begin{aligned} \omega \colon \mathscr{L}_+^\Gamma &\to \operatorname{Range}_+\Gamma \\ \Lambda &\mapsto \int G \Phi_\Lambda G^*. \end{aligned} \tag{4.7}$$

As indicated in Ferrante et al. [2010] and will be clear in the next section, this is a continuous map between open subsets of the linear space $\operatorname{Range}\Gamma$, and Problem 4.1 is feasible if and only if ω is surjective. Theorem 11.4.3 in Ferrante et al. [2010] guarantees such surjectivity when the prior is a scalar density times a positive definite matrix. In the remaining of this chapter, we shall extend that result to accommodate a generic matrix spectral density Ψ.

4.2 A short review of topological degree theory

THE proof of our main result relies on the notion of topological degree of a continuous map. The degree theory forms an important part of differential topology and is closely related to fixed-point theory, *cf.* [Outerelo and Ruiz, 2009, Ch. I] for a rather informative historical account. In particular, the degree theory is a powerful

tool to prove existence of a solution to a system of non-linear equations. There are several versions of the theory for different types of maps. Although the maps that we consider in this chapter are between open subsets of the Euclidean space, we shall use the more general degree theory for continuous maps between smooth, connected, boundary-less manifolds. Some salient points of the theory are reviewed below.

We mainly follow the lines of [Outerelo and Ruiz, 2009, Ch. III]. Suppose $U, V \subset \mathbb{R}^n$ are open and connected, and $f\colon U \to V$ is a proper $\mathscr{C}^1$ function. Recall that f is called *proper* if the preimage of every compact set in V is compact in U. Our major concern is solvability of the equation

$$f(x) = y. \tag{4.8}$$

A point $y \in V$ is called a regular value of f if either

(i) for any $x \in f^{-1}(y)$, $\det f'(x) \neq 0$ or

(ii) $f^{-1}(y)$ is empty.

Here $f^{-1}(y)$ denotes the preimage of y under f, *i.e.*, the set

$$\{x \in U : f(x) = y\},$$

and $f'(x)$ denotes the Jacobian matrix of f evaluated at x. Let y be a regular value of type (i), the degree of f at y is defined as

$$\deg_T(f, y) := \sum_{f(x)=y} \operatorname{sign} \det f'(x), \tag{4.9}$$

where the sign function

$$\operatorname{sign}(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$$

and not defined at 0.

Throughout this chapter, properness will be a crucial property of our function. Since f is proper, one can show that the preimage $f^{-1}(y)$ is finite following the classical inverse function theorem, and hence the sum above is well defined. For regular values of type (ii), we set $\deg_T(f, y) = 0$. Moreover, the set of regular values is dense in V by Sard–Brown Theorem [Outerelo and Ruiz, 2009, p. 63]. Further properties of the degree related to our problem are listed below:

- The degree of f at y does not depend on the choice of regular value. Therefore, we can define the degree of f as

  $$\deg_T(f) = \deg_T(f, y)$$

  for any regular value y.

- If $\deg_T(f) \neq 0$, then for any $y \in V$, there exists $x \in U$ such that $f(x) = y$, that is, the map f is surjective. A proof of this fact can be found in [Byrnes et al., 1995, p. 1849].

- Homotopy invariance. If $H\colon U \times [0,1] \to V$, $(x,t) \mapsto y$ is jointly continuous in (x,t) and proper, then $\deg_T(H_t, y)$ is defined and independent of $t \in [0,1]$. Here $H_t\colon U \to V$ is defined by $H_t(x) = H(x,t)$.

One important point of theory is that degree can be defined for continuous functions through approximation by smooth functions [Outerelo and Ruiz, 2009, Proposition and Definition 3.1, p. 111], and (4.9) is just a way of computing it in the special case of $\mathscr{C}^1$ [Schwartz, 1969, Remark p. 71]. In particular, the homotopy invariance of the degree holds in the continuous case [Outerelo and Ruiz, 2009, Proposition 3.4, p. 112].

4.3 A general existence result

Our main theorem will be preceded by some lemmata. Take $\Psi = I$ the identity matrix, and the map ω would reduce to

$$\begin{aligned} \tilde{\omega}\colon \mathscr{L}_+^\Gamma &\to \operatorname{Range}_+ \Gamma \\ \Lambda &\mapsto \int G(G^*\Lambda G)^{-1}G^*. \end{aligned} \tag{4.10}$$

Lemma 4.2. *The map $\tilde{\omega}$ is continuously differentiable.*

Proof. The map

$$\begin{aligned} \operatorname{inv}\colon \mathrm{GL}(n,\mathbb{C}) &\to \mathrm{GL}(n,\mathbb{C}) \\ X &\mapsto X^{-1} \end{aligned} \tag{4.11}$$

is smooth, which follows from Cramer's rule in linear algebra. Hence, the function $\tilde{F}_\Lambda(e^{j\theta}) := G(G^*\Lambda G)^{-1}G^*$ inside the integral of (4.10) is also smooth in Λ. Moreover, since G is a rational function, all the partial derivatives of $\tilde{F}_\Lambda(e^{j\theta})$ with respect to Λ are continuous in θ (and Λ). Then by Leibniz's rule for differentiation under the integral sign, partial derivatives of $\tilde{\omega}$ of all orders exist.

Next, we show that the first order partial derivatives are continuous. For the time being, let us consider the map $\tilde{\omega}$ defined on $\mathscr{L}_+$. (We made the domain restricted to the intersection with $\operatorname{Range}\Gamma$ out of the consideration of dimensionality.) From Brookes [2011], the directional derivative of the map (4.11) at $X \in \mathrm{GL}(n,\mathbb{C})$ along direction ΔX is given by

$$\nabla \operatorname{inv}(X; \delta X) = -X^{-1}(\Delta X)X^{-1}.$$

Using this fact, the directional derivative of $\tilde{\omega}$ at $\Lambda \in \mathscr{L}_+$ along direction $\Delta\Lambda$ is

$$\nabla\tilde{\omega}(\Lambda;\Delta\Lambda) = -\int G(G^*\Lambda G)^{-1}(G^*\Delta\Lambda G)(G^*\Lambda G)^{-1}G^* \tag{4.12}$$

such that $\delta\Lambda \in \mathbb{H}_n$ and $\Lambda + \Delta\Lambda$ stays in $\mathscr{L}_+$. Let us denote the integrand in (4.12) by $\tilde{F}_{\Lambda,\Delta\Lambda}(e^{j\theta})$. For a fixed $\Delta\Lambda$, one can see that the directional derivative $\nabla\tilde{\omega}(\Lambda;\Delta\Lambda)$ is continuous w.r.t. Λ. To see this fact, let a sequence $\{\Lambda_k\}_{k\geq 1} \subset \mathscr{L}_+$ converge to some $\bar{\Lambda} \in \mathscr{L}_+$ as $k \to \infty$. Notice that, the corresponding sequence of matrix-valued functions $\{G^*\Lambda_k G\}_{k\geq 1}$ is such that $G^*(e^{j\theta})\Lambda_k G(e^{j\theta}) > 0$ for all $\theta \in [-\pi,\pi]$ and $k \geq 1$. Since the eigenvalues of a continuous matrix-valued function $F\colon [a,b] \to \mathbb{C}^{n\times n}$, $\theta \mapsto F(\theta)$, depend continuously on θ (see, for instance, [Bhatia, 2013, Cor. VI.1.6]), we have that $G^*\Lambda_k G \geq \mu_k I$ where

$$\mu_k := \min_\theta \lambda_{\min}\Big(G^*(e^{j\theta})\Lambda_k G(e^{j\theta})\Big) > 0$$

and $\lambda_{\min}(\cdot)$ denotes the smallest eigenvalue. Further, since the sequence $\{\Lambda_k\}_{k\geq 1}$ converges to an element $\bar{\Lambda} \in \mathscr{L}_+$, then $\{G^*\Lambda_k G\}_{k\geq 1}$ converges uniformly to the function $G^*(e^{j\theta})\bar{\Lambda}G(e^{j\theta})$ which is positive definite for all $\theta \in [-\pi,\pi]$. Hence, there exists $\mu > 0$ such that $\mu_k \geq \mu$ for all k. On the other hand, since $\Delta\Lambda$ is fixed, it must hold that $G^*\Delta\Lambda G \leq MI$, where

$$M := \max_\theta \lambda_{\max}\Big(G^*(e^{j\theta})\Delta\Lambda G(e^{j\theta})\Big).$$

Here $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of a matrix. Therefore, we have

$$\tilde{F}_{\Lambda_k,\Delta\Lambda} \leq M\mu^{-2}GG^*, \ \forall k.$$

Moreover,

$$\big|[\tilde{F}_{\Lambda_k,\Delta\Lambda}]_{i\ell}\big| \leq M\mu^{-2}G_{\max}, \quad \forall\, k \geq 1, \ \forall\, i,\ell,$$

where $G_{\max} := \max_{\theta,i,\ell}|[GG^*]_{i\ell}| < \infty$ since the entries of $G(e^{j\theta})G^*(e^{j\theta})$ are continuous functions of θ, analytic in an open annulus containing the unit circle. Hence, by Lebesgue's dominated convergence theorem [Rudin, 1987, p. 26], we have

$$\lim_{k\to\infty} \nabla\tilde{\omega}(\Lambda;\delta\Lambda) = -\int \lim_{k\to\infty} \tilde{F}_{\Lambda_k,\Delta\Lambda} = \nabla\tilde{\omega}(\bar{\Lambda};\Delta\Lambda).$$

Partial derivatives can then be recovered by the operation $\langle\Delta\Lambda_1, \nabla\tilde{\omega}(\Lambda;\Delta\Lambda_2)\rangle$ by choosing $\Delta\Lambda_k$, $k = 1,2$ to be the standard basis matrices of $\mathbb{H}_n$, where the notation $\langle M_1, M_2\rangle := \mathrm{tr}(M_1M_2)$ denotes the standard inner product in $\mathbb{H}_n$. In this way, one can see that every partial derivative of $\tilde{\omega}$ is continuous in Λ. □

Lemma 4.3. *The map*

$$\begin{aligned} H\colon \mathscr{L}_+^\Gamma \times [0,1] &\to \operatorname{Range}_+\Gamma \\ (\Lambda, t) &\mapsto \int G\Phi_{\Lambda,t}G^*. \end{aligned} \tag{4.13}$$

is a proper continuous homotopy between ω and $\tilde{\omega}$, where

$$\Phi_{\Lambda,t} := W_\Lambda^{-1}[\,t\Psi + (1-t)I\,]W_\Lambda^{-*}. \tag{4.14}$$

Proof. By definition we need to show two things, namely that H is jointly continuous in Λ and t and that H is proper. In order to prove joint continuity, we first prove that the spectral factor $W_\Lambda(z)$ is a continuous function of $\Lambda \in \mathscr{L}_+^\Gamma$ for all $\theta \in [-\pi,\pi]$.[25] To this end we first recall that, in view of Lemma 4.1, the spectral factor $W_\Lambda(z)$ has the form

$$W_\Lambda(z) = L_\Lambda^{-*}B^*P_\Lambda A(zI-A)^{-1}B + L_\Lambda,$$

where P_Λ is the unique stabilizing solution of the following Discrete-time Algebraic Riccati Equation (DARE)

$$\Pi = A^*\Pi A - A^*\Pi B(B^*\Pi B)^{-1}B^*\Pi A + \Lambda, \tag{4.15}$$

and L_Λ is the (right) Cholesky factor of $B^*P_\Lambda B > 0$. Hence, since the Cholesky factor is a continuous function of a positive definite matrix argument [Schatzman, 2002, Lemma 12.1.6], it suffices to show that P_Λ is a continuous function of $\Lambda \in \mathscr{L}_+^\Gamma$. This follows from the fact that P_Λ is associated to the stable subspace of an extended symplectic pencil [Ionescu et al., 1999, Thm. 5.5.1]. More precisely, consider the following extended symplectic matrix pencil associated to DARE (4.15)

$$z\begin{bmatrix} I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -A^* & \mathbf{0} \\ \mathbf{0} & -B^* & \mathbf{0} \end{bmatrix} - \begin{bmatrix} A & \mathbf{0} & B \\ \Lambda & -I & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad z \in \mathbb{C}. \tag{4.16}$$

From [Avventi, 2011, Thm. A.5.1], The above pencil admits a stable deflating subspace[26] $\mathscr{V}$ of dimension n and any matrix $V \in \mathbb{C}^{(2n+m)\times n}$ generating $\mathscr{V}$ can be partitioned as

$$V = \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix},$$

where $V_1 \in \mathbb{C}^{n\times n}$ invertible, $V_2 \in \mathbb{C}^{n\times n}$, and $V_3 \in \mathbb{C}^{m\times n}$. In light of his fact, the stabilizing solution P_Λ of DARE (4.15) can be written as

$$P_\Lambda = V_2V_1^{-1},$$

[25] An alternative proof of this claim can be found in [Avventi, 2011, Thm. A.5.5], where the author proved the stronger fact that the function $\Lambda \in \mathscr{L}_+^\Gamma \mapsto W_\Lambda$ is an homeomorphism.

[26] A subspace $\mathscr{V}$ of a matrix pencil $zM - N$ is said to be *deflating* if $\dim(M\mathscr{V} + N\mathscr{V}) = \dim(\mathscr{V})$.

Finally, since the extended symplectic pencil (4.16) has no generalized eigenvalues on the unit circle, the stable and antistable generalized eigenvalues are strictly separated and the existence of a continuous basis for the stable deflating subspace and hence the continuous dependence on Λ of the stabilizing solution P_Λ of DARE (4.15) can be found, for instance, in Stewart [1973].

Consider now

$$\Phi_{\Lambda,t}(e^{j\theta}) = W_\Lambda^{-1}(e^{j\theta})[\, t\Psi(e^{j\theta}) + (1-t)I\,]W_\Lambda^{-*}(e^{j\theta}).$$

As a linear combination in $t \in [0,1]$ of continuous functions of Λ, $\Phi_{\Lambda,t}(e^{j\theta})$ is jointly continuous w.r.t. $t \in [0,1]$ and $\Lambda \in \mathscr{L}_+^\Gamma$, for all $\theta \in [-\pi,\pi]$.

Next we need to show the continuity together with the integral. Consider any sequence $\{(\Lambda_k, t_k)\}_{k\geq 1} \subset \mathscr{L}_+^\Gamma \times [0,1]$ such that $\lim_{k\to\infty} t_k = \bar{t} \in [0,1]$ and $\lim_{k\to\infty} \Lambda_k = \bar{\Lambda} \in \mathscr{L}_+^\Gamma$. By the same line of reasoning as in the proof of Lemma 4.2, there exists $\mu > 0$ such that $G^*\Lambda_k G \geq \mu I$, $\forall k$. Therefore, it holds that

$$\begin{aligned} G\Phi_{\Lambda_k,t_k}G^* &\leq K_{\max} G(G^*\Lambda_k G)^{-1}G^* \\ &\leq K_{\max}\mu^{-1}GG^*, \quad \forall k \geq 1, \end{aligned}$$

where $K_{\max} := \max_{t,\theta} \lambda_{\max}(t\Psi + (1-t)I) < \infty$ since Ψ is bounded. The rest of the argument is also similar. Given the joint continuity of $\Phi_{\Lambda,t}$ in Λ and t, one can show that the following limit holds

$$\lim_{k\to\infty} \int G\Phi_{\Lambda_k,t_k}G^* = \int \lim_{k\to\infty} G\Phi_{\Lambda_k,t_k}G^* = \int G\Phi_{\bar{\Lambda},\bar{t}}G^*.$$

This proves joint continuity of H in t and Λ.

Once we have joint continuity, the properness is not difficult to prove. In fact, let $K \subset \mathrm{Range}_+\Gamma$ be a compact subset, and we next show that the set

$$H^{-1}(K) := \left\{(\Lambda, t) \in \mathscr{L}_+^\Gamma \times [0,1] : H(\Lambda, t) \in K\right\}$$

is compact. The argument is essentially the same as the proof of Theorem 11.4.1 of Ferrante et al. [2010]. Since our setting is finite-dimensional, a set being compact is equivalent to being closed and bounded. If $H^{-1}(K)$ is unbounded, one can then find a sequence $\{(\Lambda_k, t_k)\} \subset H^{-1}(K)$ such that $\|(\Lambda_k, t_k)\| \to \infty$ as $k \to \infty$, which necessarily implies $\|\Lambda_k\| \to \infty$. However, in this case $H(\Lambda_k, t_k)$ would tend to be singular, which contradicts the premise of K being compact. This proves the boundedness.

To prove the closedness, if a sequence $\{(\Lambda_k, t_k)\}$ in $H^{-1}(K)$ converges to (Λ, t), then Λ cannot be on the boundary of $\mathscr{L}_+$, otherwise $\|H(\Lambda_k, t_k)\| \to \infty$, which again

contradicts the compactness of K. To see the latter fact, notice that

$$\begin{aligned} H(\Lambda_k, t_k) &= \int G\Phi_{\Lambda_k,t}G^* \\ &= \int GW_{\Lambda_k}^{-1}[\,t\Psi + (1-t)I\,]W_{\Lambda_k}^{-*}G^* \\ &\geq K_{\min}\int G(G^*\Lambda_k G)^{-1}G^*, \end{aligned}$$

where $K_{\min} := \min_{t,\theta}\lambda_{\min}\left(t\Psi(e^{j\theta}) + (1-t)I\right) > 0$ since Ψ is coercive. Now if $\{\Lambda_k\}$ approaches the boundary of $\mathscr{L}_+$, $\partial\mathscr{L}_+$, then $G^*(e^{j\theta})\Lambda_k G(e^{j\theta})$ tends to be singular for some θ. Since G has rank m on $\mathbb{T}$, this in turn implies that $\|H(\Lambda_k, t_k)\| \to \infty$ as $k \to \infty$. Therefore, by the joint continuity of H, $(\Lambda, t) \in H^{-1}(K)$. This concludes the proof of properness. □

Theorem 4.1. *The map ω is surjective.*

Proof. Given the second listed property of the degree, the claim follows directly if we can show that

$$\deg_T(\omega) \neq 0.$$

We notice first that ω is proper by Theorem 11.4.1 from Ferrante et al. [2010], and thus the degree is well defined. By Lemma 4.3 and the homotopy invariance of the degree,

$$\deg_T(\omega) = \deg_T(\tilde{\omega}).$$

As a consequence of Sard–Brown theorem [Outerelo and Ruiz, 2009, p. 63], the codomain $\mathrm{Range}_+\Gamma$ must contain a regular value of $\tilde{\omega}$ since it has positive $\mathrm{Range}\,\Gamma$-Lebesgue measure. By Lemma 4.2, the $\mathscr{C}^1$ degree (4.9) of $\tilde{\omega}$ at a regular value is well-defined. Meanwhile, from Theorem 11.4.2 of Ferrante et al. [2010], we know that $\tilde{\omega}$ is bijective. Therefore, we must have

$$\deg_T(\tilde{\omega}) \neq 0,$$

and this concludes the proof. □

4.4 The special case of covariance extension

In this conclusive section, we shall particularize the main result of the previous section to the special case of multivariate covariance extension. To this aim, consider the matrix pair (A, B):

$$A = \begin{bmatrix} 0 & I_m & 0 & \cdots & 0 \\ 0 & 0 & I_m & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I_m \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I_m \end{bmatrix}. \tag{4.17}$$

Here each block in A or B is of $m \times m$ and A is a $(p+1) \times (p+1)$ block matrix while B is a $(p+1)$-block column vector. It is easy to verify that in this case

$$G(z) = (zI - A)^{-1}B = \begin{bmatrix} z^{-p-1}I_m \\ z^{-p}I_m \\ \vdots \\ z^{-2}I_m \\ z^{-1}I_m \end{bmatrix}, \tag{4.18}$$

Symbolically, the steady state vector is given by

$$x(t) = G(z)y(t) = \begin{bmatrix} y(t-p-1) \\ \vdots \\ y(t-2) \\ y(t-1) \end{bmatrix}, \tag{4.19}$$

and the covariance matrix Σ has a block-Toeplitz structure, *i.e.*,

$$\Sigma = \begin{bmatrix} C_0 & C_1^* & C_2^* & \cdots & C_p^* \\ C_1 & C_0 & C_1^* & \cdots & C_{p-1}^* \\ C_2 & C_1 & C_0 & \cdots & C_{p-2}^* \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ C_p & C_{p-1} & \cdots & C_1 & C_0 \end{bmatrix}, \tag{4.20}$$

with $C_k := \mathbb{E}\{y(t+k)y(t)^*\} \in \mathbb{C}^{m\times m}$. In fact, the constraint (4.6) is equivalent to the set of *moment equations*

$$\int_{-\pi}^{\pi} e^{jk\theta}\Phi(e^{j\theta})\frac{d\theta}{2\pi} = C_k, \quad k = 0, 1, \ldots, p. \tag{4.21}$$

To find a spectral density Φ satisfying (4.21) is the classical *(multivariate) covariance extension problem*. It is known that a solution exists if $\Sigma \geq 0$ and there are infinitely many solutions if $\Sigma > 0$.

Given $\Lambda \in \mathscr{L}_+$ and $G(z)$ in (4.18), $G^*\Lambda G$ is now an L-polynomial matrix that takes positive definite values on the unit circle. Let us take

$$Q(z) := \sum_{k=-p}^{p} Q_k z^k \equiv G^*\Lambda G, \quad Q_{-k} = Q_k^* \in \mathbb{C}^{m\times m}. \tag{4.22}$$

Then according, for instance, to the main result of Chapter 2 in Part I, $Q(z)$ admits a spectral factorization

$$Q(z) = D^*(z)D(z), \tag{4.23}$$

where $D(z) = \sum_{k=0}^{p} D_k z^{-k}$ is a $m \times m$ matrix polynomial (with negative powers) and the scalar polynomial $\det D(z)$ has all its roots strictly inside the unit circle. We shall call such $D(z)$ Schur.[27] Therefore, the outer spectral factor in (4.2) is just

$$W_\Lambda(z) \equiv D(z). \tag{4.24}$$

[27] Moreover, one can make such spectral factor unique if the constant matrix coefficient D_0 is required to be lower triangular with real and positive diagonal elements.

We have the following corollary of Theorem 4.1.

Corollary 4.1. *Given a finite matrix covariance sequence $C_0, C_1, \ldots, C_n$, for any $\Psi \in \mathscr{S}_+^m(\mathbb{T})$, there exists a Schur polynomial $D(z)$ such that the spectral density*

$$\Phi := D^{-1} \Psi D^{-*} \tag{4.25}$$

satisfies the moment equations (4.6). The polynomial $D(z)$ is a right Schur spectral factor of $G^* \Lambda G$ for some $\Lambda \in \mathscr{L}_+^\Gamma$.

In particular, when taking $\Psi(z) = N(z)N^*(z)$ with $N(z) = \sum_{k=0}^{p} N_k z^{-k}$, $N_k \in \mathbb{C}^{m \times m}$, which is the spectral density of a moving-average process, the spectral density Φ in (4.25) would correspond to an m-dimensional vector ARMA process

$$\sum_{k=0}^{p} D_k \, y(t-k) = \sum_{k=0}^{p} N_k \, w(t-k), \quad t \in \mathbb{Z}, \tag{4.26}$$

and we recover the main result of [Georgiou, 1983, Sec. V] under a more general setting.

5. Finsler geometries in the cone of spectral densities

RECENTLY, the design and analysis of distances between spectral densities (or, equivalently, between stationary stochastic processes) have witnessed a renewed interest in the control and signal processing community, *cf.* Martin [2000], De Cock and De Moor [2002], Georgiou and Lindquist [2003], Georgiou [2006, 2007a,b,c], Georgiou et al. [2009], Jiang et al. [2012b], Ferrante et al. [2008, 2012a], Basseville [2013], Afsari and Vidal [2013, 2014], Zorzi [2014a,b, 2015] and references therein. This interest has been primarily triggered by the large number of applications in which the problem of measuring the discrepancy between spectral densities represents a key aspect. Besides the previously described THREE-like spectral estimation methods, these applications include speech processing Gray et al. [1980], Jiang et al. [2012a], time-series clustering Boets et al. [2005], Lauwers and De Moor [2017], and model reduction of linear (stochastic) systems Afsari and Vidal [2017], to cite just a few relevant ones.

When designing a distance between points of a given space with the aim of solving computational engineering problems, one has to take care of aspects concerning the *modelling*, *mathematical*, and *computational* features of the to-be-designed distance.

Modelling considerations consist of endowing the distance with some suitable *invariance* properties, which ensure that the distance is consistent with what is modelled. For instance, for the case of covariance (*i.e.*, positive definite) matrices such an invariant property coincides with the so-called congruence or affine invariance, whereas for rational spectral densities, we will show that *filtering invariance* emerges as a natural invariance property. This invariance property usually takes the form of an invariant group action on the space, which leads to the notion of *homogeneous space*. More generally, the study of invariants of homogeneous spaces lies at the heart of modern geometry. Indeed, as Felix Klein stated in the Erlangen program [Klein, 1893, p. 67], the principal aim of geometry is

> "[...] to investigate those properties of the images belonging to the manifold which remain unchanged under the transformations in the group."[28]

From a mathematical viewpoint, another standard requirement is to equip the underlying space with a *differential metric structure* such that the distance between two points corresponds to the length of a minimal geodesic. This is especially relevant when dealing with optimization problems on finite and infinite dimensional manifolds, and therefore for problems involving approximation, smoothing, and averaging of spectral densities, *e.g.*, in the context of speech morphing Jiang et al. [2008a,b,

[28] *"[...] Es ist eine Mannigfaltigkeit und in derselben eine Transformationsgruppe gegeben; man soll die der Mannigfaltigkeit angehörigen Gebilde hinsichtlich solcher Eigenschaften untersuchen, die durch die Transformationen der Gruppe nicht geändert werden."*

2012a]. The classical framework is *Riemannian* geometry, in which the differential structure involves an inner product on the tangent space. The present chapter focuses on the broader framework of *Finsler* geometry, where the differential structure only requires a norm.

Numerical considerations include the existence of an algorithmic framework to perform the calculations necessary to the considered engineering problem, starting with the evaluation of the distance itself. These are of primary importance in high-dimensional problems, and, *a fortiori*, so for "infinite-dimensional" objects such as spectral densities.

The starting point in this chapter is to acknowledge that the space of spectral densities is a *cone* and to revisit two classical and, in a sense, natural distances that have been studied in cones: The part metric (often called Thompson metric) and the projective metric introduced by Hilbert. Applying this classical framework to the space of *rational* spectral densities, which seems novel, we show that the resulting distances have a number of particularly desirable properties:

1) they endow the cone of spectral densities with a Finsler geometry with explicit geodesics,

2) they are filtering invariant, and

3) efficient algorithms exist to compute them.

The layout of this chapter is as follows. Section 5.1 briefly reviews some standard definitions and results from differential geometry and some properties of cones. In the same section, an overview of Hilbert and Thompson geometries on cones will be provided. In Section 5.2, we revisit the Hilbert and Thompson distances for the case of rational multivariate spectral densities and we illustrate the properties of such distances. Eventually, in Section 5.3, we discuss a promising application of the proposed distances concerning a robust version of THREE estimation. The chapter is mostly based on Baggio et al. [2017].

5.1 Preliminaries

In this section, we give some background definitions/results from metric and differential geometry, theory of cones, and, ultimately, we introduce the Hilbert and Thompson metrics on cones. A comprehensive overview of the topics treated in this chapter can be found, for instance, in the monographs Burago et al. [2001], Absil et al. [2009], Deng [2012], Lemmens and Nussbaum [2012].

5.1.1 Metric and geodesic spaces

A metric space consists of a pair (X, d), where X is a set and $d\colon X \times X \to \mathbb{R}_{\geq 0}$ is a function, called *metric* or *distance* function, such that the following hold for all $x, y, z \in X$:

1) $d(x, y) \geq 0$ (*non-negativity*),

2) $d(x, y) = 0$ if and only if $x = y$ (*definiteness*),

3) $d(x, y) = d(y, x)$ (*symmetry*),

4) $d(x, z) \leq d(x, y) + d(y, z)$ (*triangle inequality*).

If $d(\cdot, \cdot)$ satisfies properties 1) and 2) only , then $d(\cdot, \cdot)$ is said to be a *divergence*. A *complete* metric space is a metric space in which every Cauchy sequence is convergent. A *geodesic path* (or, simply, a *geodesic*) in a metric space (X, d) is a path $\gamma\colon [a, b] \to X$ that is distance-preserving, that is, such that $|\gamma(t_1) - \gamma(t_1)| = |t_1 - t_2|$ for all t_1 and t_2 in X. A metric space (X, d) is said to be *geodesic* if given two arbitrary points in X there exists a geodesic path that joins them. If the geodesic connecting the two points is unique then (X, d) is said to be *uniquely geodesic*.

5.1.2 Geometry of homogeneous spaces

An *homogeneous space* is a differentiable manifold with a *transitive* action of a Lie group G,[29] whence it has a representation as a quotient $M = G/H$ with H a closed subgroup of G. A homogeneous space $M = G/H$ is said to be *reductive* if there exists a decomposition $\mathfrak{g} = \mathfrak{m} + \mathfrak{h}$ (direct sum) such that $\mathrm{Ad}_H(\mathfrak{m}) := \{hAh^{-1} : h \in H,\, A \in \mathfrak{m}\}$, where $\mathfrak{g}$ and $\mathfrak{h}$ are the Lie algebras of G and H, respectively. The most relevant property of a reductive homogeneous space is the existence of a (unique, up to multiplication by a constant) H-invariant bilinear form $\langle \cdot, \cdot \rangle_x$ defined in the tangent space of every point $x \in G/H$. Equipped with such a family of inner products on the tangent space G/H becomes a *Riemannian manifold*. The length of any $\mathscr{C}^1$ path $\gamma\colon [a, b] \to G/H$ is defined as

$$\ell(\gamma) := \int_a^b \langle \gamma'(t), \gamma'(t) \rangle_{\gamma(t)} \, \mathrm{d}t.$$

and the geodesic distance between any two elements of the space is equal to the infimum of the lengths of $\mathscr{C}^1$ paths in G/H joining them, namely

$$d(x, y) := \inf\{\ell(\gamma) : \gamma \in \mathscr{C}^1[x, y]\}, \tag{5.1}$$

where $\mathscr{C}^1[x, y]$ denotes the set of all differentiable paths $\gamma\colon [a, b] \to \mathring{\mathscr{K}}$ such that $\gamma(0) = x$ and $\gamma(1) = y$. With this distance $(G/H, d)$ becomes a complete geodesic

[29] We recall that a Lie group is a smooth manifold with an additional, compatible structure of group.

metric space. In general, one can define a family of H-invariant norms $\|\cdot\|_x$ in the tangent space of every $x \in G/H$. This yields a so-called *Finsler manifold*. As before, one can define the length of any $\mathscr{C}^1$ path $\gamma\colon [a,b] \to G/H$ as

$$\ell(\gamma) := \int_a^b \left\|\gamma'(t)\right\|_{\gamma(t)} \mathrm{d}t.$$

and the geodesic distance between any two elements as in (5.1).

5.1.3 Cones in Banach spaces

A subset $\mathscr{K}$ of a real Banach space $\mathscr{B}$ with norm $\|\cdot\|_{\mathscr{B}}$ is a *cone* if for each $x \in \mathscr{K}$ and positive scalars λ, the product λx is in $\mathscr{K}$. A cone $\mathscr{K}$ is said to be

- *closed* if $\mathscr{K}$ is a closed subset of $\mathscr{B}$;
- *(strictly) convex* if $\mathscr{K} + \mathscr{K} \subseteq \mathscr{K}$ and $\lambda\mathscr{K} \subseteq \mathscr{K}$ ($\lambda\mathscr{K} \subset \mathscr{K}$, resp.) for positive scalars λ;
- *solid* if the interior of $\mathscr{K}$, here denoted by $\mathring{\mathscr{K}}$, is non-empty;
- *pointed* if $\mathscr{K} \cap -\mathscr{K} = \{0\}$.

A cone that satisfies all the above listed properties will be called *proper*. A proper cone $\mathscr{K}$ induces a partial ordering $\leq_{\mathscr{K}}$ on $\mathscr{B}$ by

$$x \leq_{\mathscr{K}} y \iff y - x \in \mathscr{K}.$$

For $x, y \in \mathscr{K}$, we say that y dominates x if there exists $\beta > 0$ such that $x \leq_{\mathscr{K}} \beta y$. We write $x \sim_{\mathscr{K}} y$ if y dominates x, and x dominates y. The relation $\sim_{\mathscr{K}}$ is an equivalence relation on $\mathscr{K}$. The corresponding equivalence classes are called *parts* or *components* of $\mathscr{K}$.

5.1.4 Hilbert and Thompson geometries

Let $\mathscr{K}$ be a proper cone. Given two elements $x, y \in \mathscr{K} \setminus \{0\}$, we define the following quantities

$$M(x,y) := \inf\{\lambda : x \leq_{\mathscr{K}} \lambda y\}, \tag{5.2}$$

if the set is non-empty, and $M(x,y) := \infty$ otherwise, and

$$m(x,y) := \sup\{\mu : \mu y \leq_{\mathscr{K}} x\} = \frac{1}{M(y,x)}. \tag{5.3}$$

Definition 5.1 (Bushell [1973], Thompson [1963])**.** The *Hilbert (projective) metric* and the *Thompson (part) metric* between elements $x, y \in \mathcal{K} \setminus \{0\}$ are defined respectively by

$$d_H(x,y) := \log \frac{M(x,y)}{m(x,y)}, \tag{5.4}$$

$$d_T(x,y) := \log \max\{M(x,y), M(y,x)\}, \tag{5.5}$$

if $x \sim_{\mathcal{K}} y$, and $d_H(x,y) = d_T(x,y) := \infty$, otherwise.

As a simple example, consider $\mathcal{B} = \mathbb{R}^n$ and $\mathcal{K}$ to be the positive orthant of $\mathbb{R}^n$, *i.e.*, $\mathcal{K} := \{(x_1, \ldots, x_n) : x_i \geq 0, 1 \leq i \leq n\}$. In this case, for $x, y \in \mathring{\mathcal{K}}$, it holds

$$M(x,y) = \max_i \{x_i / y_i\},$$
$$m(x,y) = \min_i \{x_i / y_i\},$$

so that Hilbert and Thompson metrics on $\mathring{\mathcal{K}}$ read, respectively, as

$$d_H(x,y) = \log \frac{\max_i \{x_i/y_i\}}{\min_i \{x_i/y_i\}},$$
$$d_T(x,y) = \log \max \left\{ \max_i \{x_i/y_i\}, \max_i \{y_i/x_i\} \right\}.$$

Thompson metric is a *bona fide* distance on each part of the cone $\mathcal{K}$ (and, in particular, on the interior $\mathring{\mathcal{K}}$). The main result in Thompson [1963] states that each part of $\mathcal{K}$ is a *complete* metric space with respect to this metric provided that $\mathcal{K}$ is normal, *i.e.*, there exists $\gamma > 0$ such that $\|x\|_{\mathcal{B}} \leq \gamma \|y\|_{\mathcal{B}}$ holds whenever $0 \leq_{\mathcal{K}} x \leq_{\mathcal{K}} y$. Hilbert metric is a distance between *rays* in each part of $\mathcal{K}$: $d_H(x,y) = 0$, $x, y \in \mathcal{K}$, $x \sim_{\mathcal{K}} y$, if and only if $x = \lambda y$ with $\lambda > 0$.

Hilbert and Thompson metrics have been of great interest to analysts, especially for their contractive properties. As a matter of fact, many naturally occurring maps in analysis, both linear and non-linear, are either non-expansive or contractive with respect to these metrics Bushell [1973], Liverani and Wojtkowski [1994], Lemmens and Nussbaum [2012]. Moreover, it has been proven that among all projective distances d on $\mathcal{K}$ for which the positive linear transformations are contractive w.r.t. d, Hilbert metric is the one with the best possible contraction ratio Kohlberg and Pratt [1982].

Importantly, Thompson and Hilbert metric endow the cone with a structure of Finsler manifold Nussbaum [1994]. In the finite-dimensional case, the interior of the cone $\mathcal{K}$ defines an n-dimensional manifold and the tangent space at each point may be identified with $\mathbb{R}^n$ and the Finsler structure is recovered by defining the norm

$$\|v\|_x^T := \inf\{\alpha > 0 : -\alpha x \leq_{\mathcal{K}} v \leq_{\mathcal{K}} \alpha x\} \tag{5.6}$$

on the tangent space at each point $x \in \mathring{\mathscr{K}}$. Thompson distance between two points in $\mathring{\mathscr{K}}$ is then recovered by minimizing the lengths of a $\mathscr{C}^1$ curve over all $\mathscr{C}^1$ paths connecting the points as in (5.1). Hilbert metric is obtained along the same lines when the norm above is replaced by the semi-norm

$$\|v\|_x^H := M(v,x) - m(v,x).$$

The Finslerian nature of Hilbert and Thompson geometries allows for the definition of minimal geodesics connecting two points in the interior of the cone $\mathscr{K}$. Differently from the Riemannian framework, minimal geodesics connecting two points are usually not unique [Lemmens and Nussbaum, 2012, Ch. 2]. A class of minimal geodesics w.r.t. both the Thompson and Hilbert metric connecting $x, y \in \mathring{\mathscr{K}}$, is given by $\varphi\colon [0,1] \to \mathring{\mathscr{K}}$,

$$\varphi(t) = \begin{cases} \left(\frac{\beta^t - \alpha^t}{\beta - \alpha}\right) y + \left(\frac{\beta\alpha^t - \alpha\beta^t}{\beta - \alpha}\right) x, & \text{if } \beta \neq \alpha, \\ \alpha^t x, & \text{if } \beta = \alpha, \end{cases} \tag{5.7}$$

where $\beta := M(y,x)$ and $\alpha := m(y,x)$. These are projective straight lines. It is interesting to note that if $\mathscr{K}$ is strictly convex then the curves in (5.7) are the only minimal geodesics w.r.t. the Hilbert metric Nussbaum and Walsh [2004]. Clearly, there might exist more suitable choices of minimal geodesics, depending on the particular cone considered.

In particular, if $\mathscr{K}$ is the cone of $n \times n$ positive semi-definite Hermitian matrices, then the Thompson distance between $X, Y \in \mathring{\mathscr{K}} \equiv \mathbb{H}_{+,n}$ is given by

$$d_T(X,Y) = \log\max\left\{\lambda_{\max}(XY^{-1}), \lambda_{\max}(X^{-1}Y)\right\},$$

where $\lambda_{\max}(\cdot)$ denotes the maximum eigenvalue, and a natural class of minimal geodesics connecting X to Y, w.r.t. Thompson metric, is given by $\varphi_T\colon [0,1] \to \mathbb{H}_{+,n}$,

$$\varphi_T(t) = X^{1/2}(X^{-1/2}YX^{-1/2})^t X^{1/2}. \tag{5.8}$$

This is precisely the (unique, up to a re-parametrization) Riemannian geodesic of $\mathbb{H}_{+,n}$ connecting X to Y, see, *e.g.*, [Bhatia, 2009, Thm. 6.1.6]. Furthermore, in this case, the Hilbert distance reads as

$$d_H(X,Y) = \log \frac{\lambda_{\max}(XY^{-1})}{\lambda_{\min}(XY^{-1})},$$

and

$$\varphi_H(t) = \frac{\varphi_T(t)}{\operatorname{tr}(\varphi_T(t))}, \quad 0 \leq t \leq 1, \tag{5.9}$$

is a minimal geodesic connecting $X, Y \in \mathbb{H}_{+,n}$, $\mathrm{tr}(X) = \mathrm{tr}(Y) = 1$, w.r.t. Hilbert metric [Lemmens and Nussbaum, 2012, Prop. 2.6.8].

Finally, we remark that the Finslerian framework so far discussed for the case of finite-dimensional spaces applies without any substantial change to the case of infinite-dimensional manifolds of bounded positive self-adjoint operators on an Hilbert space. For further details on this generalization we refer to the works by Corach and co-workers Corach et al. [1993, 1994], and, in particular, to Corach and Maestripieri [1999, 2000].

5.2 Finsler distances in the cone of spectral densities

Let $\mathscr{L}_2^n[-\pi,\pi]$ be the space of n-dimensional vector-valued functions on $\mathbb{T}$ that are square integrable w.r.t. the normalized Lebesgue measure. The space $\mathscr{L}_2^n[-\pi,\pi]$ endowed with the inner product $\langle f, g\rangle_2 := \int_{-\pi}^{\pi} f(e^{j\theta})^* g(e^{j\theta}) \frac{\mathrm{d}\theta}{2\pi}$, $f, g \in \mathscr{L}_2^n[-\pi,\pi]$, forms an Hilbert space. In what follows, we denote by $\mathscr{S}^n(\mathbb{T})$ the cone of $n \times n$ bounded positive self-adjoint operators on $\mathscr{L}_2^n[-\pi,\pi]$, namely,

$$\mathscr{S}^n(\mathbb{T}) := \big\{\Phi\colon \mathbb{T} \to \mathbb{C}^{n\times n} \text{ s.t. } \Phi(e^{j\theta}) = \Phi(e^{j\theta})^*,\ \forall e^{j\theta} \in \mathbb{T},\\ \text{and } \langle f, \Phi f\rangle_2 \geq 0,\ \forall f \in \mathscr{L}_2^n[-\pi,\pi]\big\}.$$

Elements of $\mathscr{S}^n(\mathbb{T})$ will be thought of as $n \times n$ real discrete-time *spectral densities* (in case $n = 1$ we use the shorthand $\mathscr{S}(\mathbb{T}) := \mathscr{S}^1(\mathbb{T})$). Further, we use the symbol $\mathscr{S}^n_{\mathrm{b,rat}}(\mathbb{T})$ to denote the set of rational $n \times n$ spectral densities that are bounded (that is, do not have poles) on $\mathbb{T}$. Notice that $\mathscr{S}^n_{\mathrm{b,rat}}(\mathbb{T}) \subset \mathscr{S}^n(\mathbb{T})$.

In view of the above definitions, the framework outlined in the previous section provides Finslerian distances in the cone $\mathscr{S}^n(\mathbb{T})$, and, therefore, in the space of (bounded) *rational* spectral densities $\mathscr{S}^n_{\mathrm{b,rat}}(\mathbb{T})$. Interestingly, it turns out that in the latter case the expressions of Thompson and Hilbert metric are connected with the classical spectral factorization problem discussed in Part I of the dissertation. Before giving the explicit expressions of these distances, we recall that the $\mathscr{L}_\infty$-norm of a matrix-valued function G which is bounded on $\mathbb{T}$ is defined as Zhou et al. [1996]

$$\|G\|_{\mathscr{L}_\infty} := \operatorname{ess\,sup}_{\theta\in[-\pi,\pi)} \sigma_{\max}(G(e^{j\theta})),$$

where $\sigma_{\max}(\cdot)$ denotes the maximum singular value of a matrix and $\operatorname{ess\,sup}_{x\in X} f(x)$ the essential supremum of a measurable function f on X, that is the supremum of f almost everywhere on X. If, in addition, G is analytic in (an open set containing) the complement of the open unit disk, it holds

$$\|G\|_{\mathscr{L}_\infty} \equiv \|G\|_{\mathscr{H}_\infty} = \sup_{z\in\mathbb{C}\colon |z|>1} \sigma_{\max}(G(z)).$$

Theorem 5.1. *Consider two full normal rank spectral densities $\Phi_1, \Phi_2 \in \mathcal{S}^n_{\text{b,rat}}(\mathbb{T}) \subset \mathcal{S}^n(\mathbb{T})$ and let $W_1, W_2 \in \mathbb{R}(z)^{n\times n}$ denote the corresponding minimum-phase spectral factors. If $W_2^{-1}W_1$ has no zero/pole on $\mathbb{T}$, then the Hilbert and Thompson metrics between Φ_1 and Φ_2 are given, respectively, by*

$$d_H(\Phi_1,\Phi_2) = \log \left\|W_2^{-1}W_1\right\|^2_{\mathscr{H}_\infty} \left\|W_1^{-1}W_2\right\|^2_{\mathscr{H}_\infty},$$
$$d_T(\Phi_1,\Phi_2) = \log \max\left\{\left\|W_2^{-1}W_1\right\|^2_{\mathscr{H}_\infty}, \left\|W_1^{-1}W_2\right\|^2_{\mathscr{H}_\infty}\right\}.$$

Otherwise, it holds $d_H(\Phi_1,\Phi_2) = d_T(\Phi_1,\Phi_2) = \infty$.

Proof. In view of the definition of $M(\cdot,\cdot)$ in (5.2), for any full normal rank $\Phi_1, \Phi_2 \in \mathcal{S}^n_{\text{b,rat}}(\mathbb{T})$, it holds

$$\begin{aligned} M(\Phi_1,\Phi_2) &= \inf\left\{\lambda \in \mathbb{R} : \Phi_1(e^{j\theta}) \le \lambda\Phi_2(e^{j\theta}), \theta \in [-\pi,\pi]\right\} \\ &= \inf\left\{\lambda \in \mathbb{R} : \Phi_2^{-\frac{1}{2}}(e^{j\theta})\Phi_1(e^{j\theta})\Phi_2^{-\frac{1}{2}}(e^{j\theta}) \le \lambda I_n, \theta \in [-\pi,\pi]\right\} \\ &= \left\|\Phi_2^{-\frac{1}{2}}\Phi_1\Phi_2^{-\frac{1}{2}}\right\|_{\mathscr{L}_\infty}, \end{aligned} \tag{5.10}$$

if $\Phi_2^{-\frac{1}{2}}\Phi_1\Phi_2^{-\frac{1}{2}}$ is analytic on $\mathbb{T}$, and $M(\Phi_1,\Phi_2) = \infty$ otherwise. In order to deal with rational matrix-valued functions we can replace, without affecting the value of $M(\Phi_1,\Phi_2)$, the square root $\Phi_2^{1/2}$ in the latter expression with the minimum-phase spectral factor $W_2 \in \mathbb{R}^{n\times n}(z)$ of Φ_2. (In fact, $\Phi_2^{1/2}U = W_2$, where U is a suitable $n \times n$ matrix all-pass function). Therefore, Equation (5.10) becomes

$$\begin{aligned} M(\Phi_1,\Phi_2) &= \left\|\Phi_2^{-\frac{1}{2}}\Phi_1\Phi_2^{-\frac{1}{2}}\right\|_{\mathscr{L}_\infty} \\ &= \left\|W_2^{-1}\Phi_1 W_2^{-*}\right\|_{\mathscr{L}_\infty} \\ &= \left\|W_2^{-1}W_1W_1^*W_2^{-*}\right\|_{\mathscr{L}_\infty} \\ &= \left\|W_2^{-1}W_1\right\|^2_{\mathscr{L}_\infty}, \end{aligned}$$

if $W_2^{-1}W_1$ has no pole on $\mathbb{T}$, and $M(\Phi_1,\Phi_2) = \infty$ otherwise. Further, if $W_2^{-1}W_1$ has no pole on $\mathbb{T}$, $W_2^{-1}W_1$ is analytic in (an open set containing) the complement of the (open) unit disk, so that we have

$$M(\Phi_1,\Phi_2) = \left\|W_2^{-1}W_1\right\|^2_{\mathscr{H}_\infty},$$

where we have replaced the $\mathscr{L}_\infty$-norm with the $\mathscr{H}_\infty$-norm. Similarly, we have

$$M(\Phi_2,\Phi_1) = \left\|W_1^{-1}W_2\right\|^2_{\mathscr{H}_\infty},$$

if $W_1^{-1}W_2$ has no pole on $\mathbb{T}$, or, equivalently, if $W_2^{-1}W_1$ has no zero on $\mathbb{T}$, and $M(\Phi_2,\Phi_1) = \infty$ otherwise.

Eventually, observing that $m(\Phi_1,\Phi_2) = M(\Phi_2,\Phi_1)^{-1}$, a substitution of the values of $M(\Phi_1,\Phi_2)$ and $M(\Phi_2,\Phi_1)$ into the expressions of Hilbert and Thompson metrics in Definition 5.1 yields the thesis. □

Theorem 5.1 shows that the computation of Hilbert and Thompson metrics in the cone of rational spectral densities essentially consists of:

1) the calculation of the minimum-phase spectral factors W_1 and W_2, and

2) the calculation of the $\mathscr{H}_\infty$-norm of the "ratio" of the latter spectral factors.

Remarkably, these two steps represent two extensively studied problems in Systems and Control Theory and several algorithms are available in the literature to perform these steps. More specifically, as discussed in Part I of the dissertation there exist several efficient routines for the calculation of minimum-phase spectral factors based on the solution of suitable Stein and Riccati equations, see, *e.g.*, Oară [2005]. Whereas, an efficient method for computing $\mathscr{H}_\infty/\mathscr{L}_\infty$-norms of rational matrix-valued transfer functions is the Boyd–Balakrishnan–Bruinsma–Steinbuch method Bruinsma and Steinbuch [1990], Boyd and Balakrishnan [1990] which is based on an iterative bisection-like algorithm and leads to quadratic convergence.[30] Moreover, an upper bound to these norms can be found by inspecting the eigenvalues of the symplectic matrix associated with the state-space representation of the system [Zhou et al., 1996, Lemma 21.10].

[30] The approach is formulated in the continuous-time case, however there exist variants of this algorithm for computing the discrete-time $\mathscr{H}_\infty$-norm, based on computing eigenvalues of symplectic instead of Hamiltonian matrices, see, *e.g.*, Genin et al. [1998].

Remark 5.1. The proof of Theorem 5.1 shows that the expressions of the Hilbert and Thompson metric still hold if we replace the canonical (*i.e.*, minimum-phase) spectral factors of the two spectra Φ_1, Φ_2 with any other spectral factor of Φ_1, Φ_2 (*i.e.*, spectral factors not necessarily analytic in the complement of the unit disk and with analytic inverse in the complement of the closure of the unit disk). The important difference is that, in this case, the $\mathscr{H}_\infty$-norm must be replaced by the $\mathscr{L}_\infty$-norm.

Remark 5.2. As discussed in the previous section, the difference between Hilbert and Thompson metric consists of the fact that the Thompson metric is a *bona fide* distance on each part of a cone (and, in particular, on its interior), while Hilbert metric is a distance between rays in each part of a cone. It is worth remarking that, for the case of spectral densities, projective invariance has proved to be a desirable property since in many applications, such as speech processing, the shape of the spectral densities rather than their relative scalings is the discriminative feature Afsari and Vidal [2014], Georgiou [2007a].

The following result provides a class of interesting and computable minimal geodesics w.r.t. Hilbert and Thompson metrics.

Proposition 5.1. *A minimal geodesic in $\mathscr{S}^n(\mathbb{T})$ connecting $\Phi_1, \Phi_2 \in \mathring{\mathscr{S}}^n(\mathbb{T})$ w.r.t. Thompson metric is given by*

$$\varphi_T(t) = W_1(W_1^{-1}\Phi_2 W_1^{-*})^t W_1^*, \tag{5.11}$$

where $W_1 \in \mathbb{R}^{n\times n}(z)$ is the minimum-phase spectral factor of Φ_1. A minimal geodesic in $\mathscr{S}^n(\mathbb{T})$ between $\Phi_1, \Phi_2 \in \mathring{\mathscr{S}}^n_{\text{b,rat}}(\mathbb{T})$ s.t. $\int_{-\pi}^{\pi} \operatorname{tr}(\Phi_1(e^{j\theta}))\frac{\mathrm{d}\theta}{2\pi} = \int_{-\pi}^{\pi} \operatorname{tr}(\Phi_2(e^{j\theta}))\frac{\mathrm{d}\theta}{2\pi} = 1$, w.r.t. Hilbert metric is given by

$$\varphi_H(t) = \frac{W_1(W_1^{-1}\Phi_2 W_1^{-*})^t W_1^*}{\int_{-\pi}^{\pi} \operatorname{tr}(W_1(W_1^{-1}\Phi_2 W_1^{-*})^t W_1^*)\frac{\mathrm{d}\theta}{2\pi}}. \tag{5.12}$$

Proof. The result follows from a generalization of the expressions (5.8)-(5.9) of the Thompson and Hilbert geodesics in the constant positive definite case to the case of bounded positive operators on an Hilbert space. A detailed derivation of this generalization w.r.t. Thompson metric can be found in Corach et al. [1993, 1994]. As a consequence, by adapting the expressions (5.8)-(5.9) to spectral densities in $\mathscr{S}^n(\mathbb{T})$ and by using the fact that $\Phi_1^{1/2} = W_1 U$ with U being an $n \times n$ matrix all-pass function, Equation (5.11) and Equation (5.12) are recovered. □

Remark 5.3. The expressions of Hilbert and Thompson metrics in Theorem 5.1 and of the geodesics in Proposition 5.1 apply also to the case of general non-rational spectral densities in $\mathscr{S}^n(\mathbb{T})$.[31] In this case, however, one issue that arises is that the distance between almost identical spectral densities can be made arbitrarily large. With reference to the scalar case, this occurs when one of the two spectral densities exhibits a very sharp and narrow frequency peak. For the sake of illustration, consider the two scalar spectral densities in $\mathscr{S}(\mathbb{T})$

$$\phi_1(e^{j\theta}) = 1, \quad \phi_{2,\varepsilon}(e^{j\theta}) = \begin{cases} \varepsilon^{-1} & \text{if } |\theta| \le \varepsilon, \\ 1 & \text{otherwise,} \end{cases} \tag{5.13}$$

where $\theta \in [-\pi, \pi]$ and $\varepsilon > 0$. It can be seen that, for $\varepsilon \to 0$, $d_H(\phi_1, \phi_{2,\varepsilon}) \to \infty$ and $d_T(\phi_1, \phi_{2,\varepsilon}) \to \infty$, in spite of the fact that the two spectral densities are identical with the only exception of a neighbourhood of the frequency $\theta = 0$ (see also Figure 5.1). Importantly, when restricting the attention to spectral densities that are "sufficiently regular", *e.g.*, those belonging to the space of rational spectral densities with bounded McMillan degree, these pathological cases are ruled out.

[31] Notice that in case the minimum-phase spectral factors of Φ_1, Φ_2 do not exist, the expressions in Theorem 5.1 and Proposition 5.1 still hold by replacing the minimum-phase spectral factors with the corresponding matrix square roots $\Phi_1^{1/2}$, $\Phi_2^{1/2}$.

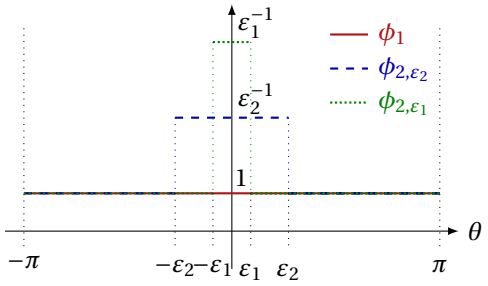


Figure 5.1: Qualitative plot of the spectral densities in Equation (5.13) for two values of ε, *i.e.*, $\varepsilon_2 > \varepsilon_1 > 0$.

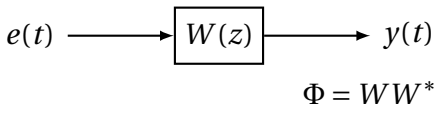
5.2.1 Filtering invariance

Let $\mathbb{R}_*(z)^{n\times n}$ denote the set of $n \times n$ real matrix-valued rational functions of full rank on $\mathbb{T}$. The above-introduced distances possess the following important property:

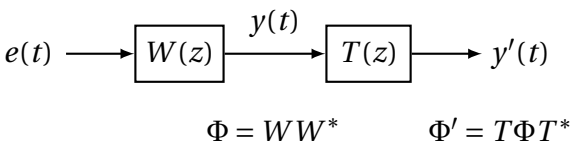
$$\forall\, T \in \mathbb{R}_*(z)^{n\times n}:\ d(\Phi_1,\Phi_2) = d(T\Phi_1 T^*, T\Phi_2 T^*). \tag{5.14}$$

This property readily follows from the definition of Hilbert and Thompson distances and the expression of $M(\Phi_1,\Phi_2)$ in (5.10). Since the set $\mathbb{R}_*(z)^{n\times n}$ defines a group, the mapping $\Phi \mapsto T\Phi T^*$ defines a congruence group action of $\mathbb{R}_*(z)^{n\times n}$ on the set of rational spectral densities. This group action is transitive, that is, any rational spectral density can be obtained by acting on the identity element.

A metric that satisfies (5.14) can be said to be *filtering invariant* because of the following statistical interpretation. Any spectral density Φ with minimum-phase spectral factor W can be identified to a n-dimensional zero-mean second-order stationary purely nondeterministic stochastic process $\{y(t)\}_{t\in\mathbb{Z}}$ generated by filtering a white noise process $\{e(t)\}_{t\in\mathbb{Z}}$ through W (see also the block diagram representation below).



The action $\Phi \mapsto T\Phi T^*$ has therefore the interpretation of filtering the process with the linear time-invariant filter $T \in \mathbb{R}_*(z)^{n\times n}$, as pictorially shown below.



Likewise, property (5.14) has the interpretation that the distance between two spectral densities, or, equivalently, two zero-mean second-order stationary purely nondeterministic stochastic processes, is unchanged when the two processes are filtered by the same filter. Any filtering invariant metric is entirely specified by defining the distance to identity. Furthermore, one has $d(\Phi, I) = d(\Phi^{-1}, I)$. In other words, the distance is a *distortion* measure.

Filtering invariance is a fundamental property of classical metrics. In the scalar case, $\phi_1, \phi_2 \in \mathscr{S}_{\mathrm{b,rat}}(\mathbb{T})$, the log spectral deviation Gray et al. [1980]

$$\left(\int_{-\pi}^{\pi} \left\| \log \frac{\phi_1}{\phi_2} \right\|^2 \frac{\mathrm{d}\theta}{2\pi} \right)^{1/2}$$

is an early example of filtering invariant distortion measure. The recent work Jiang et al. [2012b] shows that the multivariate generalization

$$\begin{aligned} d_R(\Phi_1, \Phi_2) &= \left(\int_{-\pi}^{\pi} \left\| \log \Phi_1^{-1/2} \Phi_2 \Phi_1^{-1/2} \right\|_{\mathrm{F}}^2 \frac{\mathrm{d}\theta}{2\pi} \right)^{1/2} \\ &= \left(\int_{-\pi}^{\pi} \left\| \log W_1^{-1} \Phi_2 W_1^{-*} \right\|_{\mathrm{F}}^2 \frac{\mathrm{d}\theta}{2\pi} \right)^{1/2} \end{aligned} \tag{5.15}$$

is the unique Riemannian distance that is filtering invariant. This metric is a natural generalization of the affine-invariant metric between positive definite matrices. Affine invariance corresponds to filtering invariance in the static case: The congruence group action reduces to an action of the general linear group. The metric is in this case a distance between n-dimensional zero-mean second-order random vectors, and the invariance property is an invariance with respect to an affine change of coordinates. The importance of this invariance property in the context of estimation problems has been emphasized for instance in Smith [2005]. In Jiang et al. [2012b], filtering invariance emerges as a natural property when measuring the "flatness" of innovations processes. Filtering invariance is also a leading prerequisite in the work Martin [2000], whose resulting metric, which applies to scalar spectral densities $\phi_1, \phi_2 \in \mathscr{S}_{\mathrm{b,rat}}(\mathbb{T})$, can be written as

$$d_M(\phi_1, \phi_2) = \left(\int_{-\pi}^{\pi} \left(\mathfrak{D}^{\frac{1}{2}} \log \frac{\phi_1}{\phi_2} \right)^2 \frac{\mathrm{d}\theta}{2\pi} \right)^{1/2},$$

where $\mathfrak{D}^{\lambda}$, $\lambda > 0$, is the fractional derivative operator in the frequency domain.

The Riemannian distance (5.15) and the Thompson metric introduced in Theorem 5.1 are thus close relatives: They are *bona fide* distances which satisfy filtering invariance and endow the cone of spectral densities with a differential metric structure. The first one induces a Riemannian structure through an invariant inner product (that reduces to the standard inner product at identity), while the second induces a

Finslerian structure through the invariant norm (5.6). Both distances depend on the same log spectral quantity frequency-wise, but the Riemannian distance results in a two-norm of that frequency-domain function, whereas the Finsler distance results in an infinite-norm.

A merit of the Finslerian distance over its Riemannian relative is at the computational level. The calculation of the Riemannian distance requires the frequency-wise computation of the matrix logarithm of $W_1^{-1}\Phi_2 W_1^{-*}$, an operation which appears numerically challenging. In contrast, the calculation of the Thompson metric involves the computation of minimum-phase spectral factors and $\mathscr{H}_\infty$-norms, for which efficient numerical algorithms are available. It is worth observing that the Thompson geodesic in Proposition 5.1 coincides with the (unique, up to a re-parametrization) Riemannian geodesic between spectral densities.

For completeness, it should be mentioned that one way of overcoming the computational burden of the Riemannian distance is to replace it with a divergence measure. In the static case, Kullback–Leibler divergence approximates the Riemannian distance up to third order. In the dynamic case, the paper Jiang et al. [2012b] considers quadratic approximations of divergence measures. In the rational case, one such quantity takes the form

$$d_F(\Phi_1,\Phi_2) = \|W_2^{-1}W_1\|_{\mathscr{H}_2}^2 + \|W_1^{-1}W_2\|_{\mathscr{H}_2}^2 - 2n,$$

where $\|\cdot\|_{\mathscr{H}_2}$ denotes the $\mathscr{H}_2$-norm of a discrete-time transfer function [Zhou et al., 1996, Sec. 4.3]. It is not a distance (in fact, it does not obey the triangle inequality) but it provides a tractable quadratic approximation of the Riemannian distance.

5.3 Towards robust spectral estimation

CONSIDER the THREE-like estimation problem introduced in Chapter 2. In this setting, the spectral estimation problem reduces to a constrained optimization problem of the form in Problem 2.1. As previously pointed out, one crucial aspect in this problem concerns the choice of the distance index $d(\cdot,\cdot)$ to minimize. Many works have addressed Problem 2.1 using various distance-like functions both in the scalar and multivariate case. A common feature of all those distance-like functions is that they involve the two-norm of a frequency-wise quantity defined on the unit circle.

Choosing the Finsler distances of this chapter in place of the distance $d(\cdot,\cdot)$ in Problem 2.1 could lead to a *robust* version of the THREE estimation problem. A main motivation for this modified formulation concerns the reduction of *artifacts* in the solution of Problem 2.1. The presence of artifacts is an issue that affects many of the spectral estimation methods proposed in the literature, see, for instance, the discussion in [Ferrante et al., 2012a, Sec. VII-B]. Artifacts are usually present in the form

of high and narrow frequency peaks in the spectral estimate $\hat{\Phi}$. In light of Remark 5.3, the advantage of using either Thompson and Hilbert distance in Problem Problem 2.1 consists of the fact that these artifacts are highly penalized by these distances, and, consequently, they should not appear in the optimal spectral estimate $\hat{\Phi}$.

More generally, the optimization of $\mathscr{H}_\infty$-norms in place of or in complement to $\mathscr{H}_2$-norms has been a very frutiful direction of research in linear system theory. Building upon this heritage, we envision that the distances introduced in the present chapter could also open novel avenues in robust statistical estimation.

6. Conclusions

In this second part of the dissertation, we have addressed a few problems arising from the THREE-like approach to spectral estimation. We first reviewed, in Chapter 2, the scalar estimation problem that hinges on the Kullback–Leibler criterion. We then focused on a particularly efficient numerical approach for the solution of the latter problem, proposed in Pavon and Ferrante [2006]. This approach is based on a fixed-point iteration. Complementing some previous (partial) attempts in the literature Ferrante et al. [2007, 2011], in Chapter 3, we provided a proof a *global* convergence of the iteration to one of its fixed points. Furthermore, we discussed a modification of the Pavon–Ferrante iteration, for which convergence to the set of positive definite fixed points is guaranteed. Importantly, such fixed points always yield the solution of the Kullback–Leibler spectral estimation problem. In Chapter 4, we considered the parametric approach to multivariate spectral estimation introduced in Ferrante et al. [2010]. Our main contribution there has been a proof of the existence of a solution to the problem for *any* (bounded and coercive) prior spectral density. This considerably extends the main result of Ferrante et al. [2010]. Ultimately, in Chapter 5, we addressed the problem of geometrizing the space of multivariate spectral densities. We showed that a *natural* class of distances on this space is that arising from a *Finsler geometry* framework. Notable representatives of this class are the Hilbert and Thompson metrics. In the end of the chapter, using these distances, we formulated an open problem regarding a *robust* version of the THREE-like spectral estimation paradigm.

Besides the latter very interesting open problem, other aspects that are currently under investigation are related to the topics touched in Chapter 3 and Chapter 4.

With reference to Chapter 3, a question which remains unanswered concerns the global convergence of the Pavon–Ferrante algorithm to the *closure* of the the set of positive definite fixed points of the iteration. On the one hand, numerical simulations suggest that there do exist some particular cases in which the latter fact is not verified (see the numerical example in Subsection 3.4.3). On the other hand, it is also true that this could be due to numerical issues in the computation of the matrix square root which occur when the sequence generated by the iteration approaches the boundary of the set of unit trace positive semi-definite matrices. In the light of this observation, it would be interesting to provide a formal argument to either prove or disprove the above conjecture.

Regarding Chapter 4, an immediate question concerns the *uniqueness* of the solution to the parametric estimation problem and, more strongly, *well-posedness* of the problem. Following previous work on rational covariance extension Byrnes et al.

[1995], we intend to pursue the uniqueness problem in the frame of the global inverse function theorem now attributed to Hadamard. More precisely, we aim at applying the following result[32] to the map $\omega\colon \mathscr{L}_+^\Gamma \to \mathrm{Range}_+\Gamma$, as defined in (4.7).

[32] We refer to Gordon [1972] and also [Krantz and Parks, 2013, p. 127] for a slightly weaker version of the theorem.

Theorem 6.1 (Hadamard)**.** *Let M_1 and M_2 be connected, oriented, boundary-less n-dimensional manifolds of class $\mathscr{C}^1$, and suppose that M_2 is simply connected. Then a $\mathscr{C}^1$ map $f\colon M_1 \to M_2$ is a diffeomorphism if and only if f is proper and the Jacobian determinant of f never vanishes.*

To this end, we first need to check whether the codomain of ω is simply connected. This is provided by the following simple propositon.

Proposition 6.1. *The set* $\mathrm{Range}_+\Gamma$ *defined in* (4.5) *is simply connected.*

Proof. By definition [Krantz and Parks, 2013, p. 127], we need to show that: Whenever $f\colon [0,1] \to \mathrm{Range}_+\Gamma$ is a closed curve, i.e., f is continuous with $f(0) = f(1) = \Sigma$, there exists a continuous function $F\colon [0,1]\times[0,1] \to \mathrm{Range}_+\Gamma$ such that

(i) $F(t,0) = f(t)$, for all $t \in [0,1]$,

(ii) $F(0,u) = F(1,u) = \Sigma$, for all $u \in [0,1]$, and

(iii) $F(t,1) = \Sigma$, for all $t \in [0,1]$.

One can easily verify that $F(t,u) := (1-u)f(t) + u\Sigma$ is the desired function. □

Given Theorem 6.1 and the above proposition, the open question becomes: Is the map ω continuously differentiable? If so, how to compute its Jacobian?

Finally, another research direction concerns the computation of a solution to the estimation problem via the "multivariate" Pavon–Ferrante algorithm in Equation (4.4). In this case, a compelling question is how to extend the "scalar" global convergence result presented in Chapter 3 to this more general case.

Summary and outlook

THE present dissertation dealt with two different, yet intimately connected, problems. The first problem, spectral factorization, constitutes the backbone of many areas of Systems and Control Theory. In the first part of the dissertation, we focused on the discrete-time setting, and, without introducing any facilitating assumptions, we derived a general spectral factorization result. Our approach has been inspired by the well-known Youla's method and relies mainly on tools from polynomial matrix theory. Furthermore, we analyzed a number of intriguing issues stemming from this result, namely

- uniqueness and parameterization of minimal spectral factors, and
- extension to the indefinite "J-spectral" case.

We believe that our results provide a rather deep and comprehensive "road-map" on spectral factorization and related problems.

The second problem we considered is the so-called THREE-like spectral estimation problem introduced by Byrnes, Georgiou, and Lindquist at the beginning of the 2000s. This problem, albeit quite recent, is rooted in the celebrated and widely investigated theory of (generalized) moment problems. In the second part of the dissertation, we addressed a number of questions, mostly of technical nature, arising from the THREE-like framework. Specifically, these questions concerned:

- the convergence of an efficient numerical estimation procedure,
- the existence of solutions to a parametric multivariate estimation problem,
- the geometrization of the space of spectral densities.

Besides a large amount of advanced (linear-)algebraic techniques, in the derivation of our results, we used a combination of a variety of powerful mathematical tools arising from non-linear systems theory, differential topology and geometry. From our study, several open problems have emerged. These offer interesting and worthwhile avenues for future research.

Bibliography

P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009. [Cited on page 160]

B. Afsari and R. Vidal. Group action induced distances on spaces of high-dimensional linear stochastic processes. In *Geometric Science of Information*, pages 425–432. Springer, 2013. [Cited on page 159]

B. Afsari and R. Vidal. Distances on spaces of high-dimensional linear stochastic processes: A survey. In *Geometric Theory of Information*, pages 219–242. Springer, 2014. [Cited on pages 159 and 167]

B. Afsari and R. Vidal. Bundle reduction and the alignment distance on spaces of state-space LTI systems. *IEEE Transactions on Automatic Control (to appear)*, 2017. [Cited on page 159]

L. V. Ahlfors. *Complex analysis: an introduction to the theory of analytic functions of one complex variable*. McGraw-Hill, New York, London, 3rd edition, 1953. [Cited on page 138]

N. I. Akhiezer. *The classical moment problem and some related questions in analysis*. Oliver & Boyd, Edinburgh, 1965. [Cited on page 113]

N. I. Akhiezer and M. G. Kreĭn. *Some questions in the theory of moments*, volume 2 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1962. [Cited on page 113]

I. Aksikas, J. J. Winkin, and D. Dochain. Optimal LQ-feedback regulation of a non-isothermal plug flow reactor model by spectral factorization. *IEEE Transactions on Automatic Control*, 52(7):1179–1193, 2007. [Cited on page 26]

F. A. Aliev and V. B. Larin. Algorithm of J-spectral factorization of polynomial matrices. *Automatica*, 33(12):2179–2182, 1997. [Cited on page 31]

F. A. Aliev and V. B. Larin. *Optimization of linear control systems: Analytical methods and computational algorithms*. Stability and Control: Theory, Methods and Applications. CRC Press, 1998. [Cited on page 31]

D. Alpay and I. Gohberg. Unitary rational matrix functions. In *Topics in interpolation theory of rational matrix-valued functions*, pages 175–222. Springer, 1988. [Cited on pages 28, 59, 62, and 69]

A. N. Amini, E. S. Ebbini, and T. T. Georgiou. Noninvasive estimation of tissue temperature via high-resolution spectral analysis techniques. *IEEE Transactions on Biomedical Engineering*, 52(2):221–228, 2005. [Cited on page 113]

B. D. O. Anderson. The inverse problem of stationary covariance generation. *Journal of Statistical Physics*, 1(1):133–147, 1969. [Cited on page 26]

B. D. O. Anderson. Riccati equations, network theory and Brune synthesis: Old solutions for contemporary problems. In *Dynamical systems, control, coding, computer vision*, pages 1–25. Springer, 1999. [Cited on page 26]

B. D. O. Anderson and M. Deistler. Properties of zero-free spectral matrices. *IEEE Transactions on Automatic Control*, 54(10):2365–2375, 2009. [Cited on page 26]

B. D. O. Anderson and M. Gevers. On multivariable pole-zero cancellations and the stability of feedback systems. *IEEE Transactions on Circuits and Systems*, 28(8): 830–833, 1981. [Cited on page 60]

B. D. O. Anderson and J. B. Moore. *Optimal filtering*. Prentice-Hall, Englewood Cliffs, NJ, 1979. [Cited on page 26]

B. D. O. Anderson and S. Vongpanitlerd. *Network analysis and synthesis: A modern systems theory approach*. Dover Books on Engineering Series. Dover Publications, 2006. [Cited on page 26]

B. D. O. Anderson, M. Deistler, W. Chen, and A. Filler. Autoregressive models of singular spectral matrices. *Automatica*, 48(11):2843–2849, 2012. [Cited on pages 26 and 27]

E. Avventi. *Spectral moment problems: generalizations, implementation and tuning*. PhD thesis, KTH Royal Institute of Technology, 2011. [Cited on pages 26, 148, and 153]

E. Avventi, A. Lindquist, and B. Wahlberg. ARMA identification of graphical models. *IEEE Transactions on Automatic Control*, 58(5):1167–1178, 2013. [Cited on pages 26 and 114]

G. Baggio. A global convergence analysis of the Pavon–Ferrante algorithm for spectral estimation. *IEEE Transactions on Automatic Control (Conditionally Accepted)*, 2017a. [arXiv preprint: `arXiv:1612.03570`]. [Cited on page 127]

G. Baggio. On the convergence of a matricial fixed-point iteration connected with spectral estimation. *Proceedings of the 20th IFAC World Congress*, pages 7415–7420, 2017b. [Cited on page 127]

G. Baggio. On minimal discrete-time J-spectral factorizations. *In preparation*, 2017c. [Cited on page 94]

G. Baggio and A. Ferrante. On the factorization of rational discrete-time spectral densities. *IEEE Transactions on Automatic Control*, 61(4):969–981, 2016a. [Cited on page 33]

G. Baggio and A. Ferrante. On minimal spectral factors with zeroes and poles lying on prescribed regions. *IEEE Transactions on Automatic Control*, 61(8):2251–2255, 2016b. [Cited on page 59]

G. Baggio and A. Ferrante. Parametrization of minimal spectral factors of discrete-time rational spectral densities. *Submitted for publication*, 2017. [arXiv preprint: arXiv:1609.02711]. [Cited on page 69]

G. Baggio, A. Ferrante, and R. Sepulchre. Finslerian metrics in the cone of spectral densities. *Submitted for publication*, 2017. [arXiv preprint: arXiv:1708.02818]. [Cited on page 160]

J. A. Ball and N. Cohen. Sensitivity minimization in an H^∞ norm: Parametrization of all suboptimal solutions. *International Journal of Control*, 46(3):785–816, 1987. [Cited on page 29]

J. A. Ball and A. C. M. Ran. Optimal Hankel norm model reductions and Wiener–Hopf factorization I: The canonical case. *SIAM Journal on Control and Optimization*, 25 (2):362–382, 1987a. [Cited on page 30]

J. A. Ball and A. C. M. Ran. Optimal Hankel norm model reductions and Wiener–Hopf factorization II: The non-canonical case. *Integral Equations and Operator Theory*, 10(3):416–436, 1987b. [Cited on page 30]

J. A. Ball and A. J. van der Schaft. J-inner-outer factorization, J-spectral factorization, and robust control for nonlinear systems. *IEEE Transactions on Automatic Control*, 41(3):379–392, 1996. [Cited on page 29]

J. A. Ball, J. W. Helton, and M. Verma. A factorization principle for stabilization of linear control systems. *International Journal of Robust and Nonlinear Control*, 1(4): 229–294, 1991. [Cited on page 29]

M. Banker. Linear stationary quadratic games. In *Proceedings of the 1972 IEEE Conference on Decision and Control (CDC) and 11th Symposium on Adaptive Processes*, pages 193–197, 1972. [Cited on page 29]

H. Bart, I. Gohberg, M. Kaashoek, and A. C. M. Ran. *Factorization of matrix and operator functions: The state space method*, volume 178 of *Linear Operators and Linear Systems*. Birkhäuser, Basel, 2007. [Cited on pages 27, 85, 88, and 102]

H. Bart, I. Gohberg, M. Kaashoek, and A. C. M. Ran. *A state space approach to canonical factorization with applications*, volume 200 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel, 2011. [Cited on pages 27 and 93]

M. Bart, I. Gohberg, and M. Kaashoek. *Minimal factorization of matrix and operator functions*, volume 1 of *Operator Theory: Advances and Applications*. Birkhäuser, Basel, 1979. [Cited on pages 27, 73, and 109]

M. Basseville. Divergence measures for statistical data processing – an annotated bibliography. *Signal Processing*, 93(4):621–633, 2013. [Cited on page 159]

V. Belevitch. *Classical network theory*. Holden-Day, San Francisco, 1968. [Cited on page 59]

R. Bhatia. *Positive definite matrices*. Princeton University Press, 2009. [Cited on page 164]

R. Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2013. [Cited on page 152]

D. A. Bini, G. Fiorentino, L. Gemignani, and B. Meini. Effective fast algorithms for polynomial spectral factorization. *Numerical Algorithms*, 34(2):217–227, 2003. [Cited on page 27]

A. Blomqvist, A. Lindquist, and R. Nagamune. Matrix-valued Nevanlinna–Pick interpolation with complexity constraint: An optimization approach. *IEEE Transactions on Automatic Control*, 48(12):2172–2190, 2003. [Cited on pages 26 and 114]

J. Boets, K. De Cock, B. De Moor, and M. Espinoza. Clustering time series, subspace identification and cepstral distances. *Communications in Information & Systems*, 5 (1):69–96, 2005. [Cited on page 159]

A. Boettcher and M. Halwass. A Newton method for canonical Wiener–Hopf and spectral factorization of matrix polynomials. *Electronic Journal of Linear Algebra*, 26(1):59, 2013. [Cited on page 27]

S. Boyd and V. Balakrishnan. A regularity result for the singular values of a transfer matrix and a quadratically convergent algorithm for computing its L_∞-norm. *Systems & Control Letters*, 15(1):1–7, 1990. [Cited on page 167]

S. Boyd, V. Balakrishnan, and P. Kabamba. A bisection method for computing the H_∞ norm of a transfer matrix and related problems. *Mathematics of Control, Signals and Systems*, 2(3):207–219, 1989. [Cited on page 142]

B. Brogliato, R. Lozano, B. Maschke, and O. Egeland. *Dissipative systems analysis and control*. Springer-Verlag, London, 2007. [Cited on page 26]

M. Brookes. The matrix reference manual. [http://www.ee.ic.ac.uk/hp/staff/dmb/-matrix/intro.html], 2011. Accessed: 2017-10-10. [Cited on page 151]

N. A. Bruinsma and M. Steinbuch. A fast algorithm to compute the H_∞-norm of a transfer function matrix. *Systems & Control Letters*, 14(4):287–293, 1990. [Cited on pages 142 and 167]

D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society Providence, RI, 2001. [Cited on page 160]

J. P. Burg. *Maximum entropy spectral analysis*. PhD thesis, Stanford Univ., Stanford, California, 1975. [Cited on page 113]

P. J. Bushell. Hilbert's metric and positive contraction mappings in a Banach space. *Archive for Rational Mechanics and Analysis*, 52(4):330–338, 1973. [Cited on page 163]

C. I. Byrnes and A. Lindquist. On the partial stochastic realization problem. *IEEE Transactions on Automatic Control*, 42(8):1049–1070, 1997. [Cited on pages 114 and 120]

C. I. Byrnes and A. Lindquist. A convex optimization approach to generalized moment problems. In *Control and modeling of complex systems*, pages 3–21. Springer, 2003. [Cited on page 114]

C. I. Byrnes and A. Lindquist. The generalized moment problem with complexity constraint. *Integral Equations and Operator Theory*, 56(2):163–180, 2006. [Cited on page 114]

C. I. Byrnes and A. Lindquist. Interior point solutions of variational problems and global inverse function theorems. *International Journal of Robust and Nonlinear Control*, 17(5-6):463–481, 2007. [Cited on page 114]

C. I. Byrnes and A. Lindquist. Important moments in systems and control. *SIAM Journal on Control and Optimization*, 47(5):2458–2469, 2008. [Cited on pages 17 and 114]

C. I. Byrnes and A. Lindquist. The moment problem for rational measures: Convexity in the spirit of Krein. In *Modern Analysis and Applications*, pages 157–169. Springer, 2009. [Cited on page 114]

C. I. Byrnes, A. Lindquist, S. V. Gusev, and A. S. Matveev. A complete parameterization of all positive rational extensions of a covariance sequence. *IEEE Transactions on Automatic Control*, 40(11):1841–1857, 1995. [Cited on pages 114, 120, 147, 151, and 173]

C. I. Byrnes, H. J. Landau, and A. Lindquist. On the well-posedness of the rational covariance extension problem. In *Current and Future Directions in Applied Mathematics*, pages 83–108. Springer, 1997. [Cited on page 120]

C. I. Byrnes, S. V. Gusev, and A. Lindquist. A convex optimization approach to the rational covariance extension problem. *SIAM Journal on Control and Optimization*, 37(1):211–229, 1998. [Cited on pages 114, 120, and 147]

C. I. Byrnes, T. T. Georgiou, and A. Lindquist. A new approach to spectral estimation: A tunable high-resolution spectral estimator. *IEEE Transactions on Signal Processing*, 48(11):3189–3205, 2000. [Cited on pages 113, 117, and 125]

C. I. Byrnes, P. Enqvist, and A. Lindquist. Cepstral coefficients, covariance lags, and pole-zero models for finite data strings. *IEEE Transactions on Signal Processing*, 49 (4):677–693, 2001a. [Cited on page 114]

C. I. Byrnes, T. T. Georgiou, and A. Lindquist. A generalized entropy criterion for Nevanlinna–Pick interpolation with degree constraint. *IEEE Transactions on Automatic Control*, 46(6):822–839, 2001b. [Cited on pages 26, 113, 114, 120, and 147]

C. I. Byrnes, S. V. Gusev, and A. Lindquist. From finite covariance windows to modeling filters: A convex optimization approach. *SIAM Review*, 43(4):645–675, 2001c. [Cited on pages 114 and 147]

C. I. Byrnes, T. T. Georgiou, A. Lindquist, and A. Megretski. Generalized interpolation in H^∞ with a complexity constraint. *Transactions of the American Mathematical Society*, 358(3):965–987, 2006. [Cited on pages 26, 114, and 120]

C. I. Byrnes, A: Lindquist, and T. T. Georgiou. Method and apparatus for speaker verification using a tunable high-resolution spectral estimator, June 19 2007. US Patent 7,233,898. [Cited on page 113]

F. Callier. On polynomial matrix spectral factorization by symmetric extraction. *IEEE Transactions on Automatic Control*, 30(5):453–464, 1985. [Cited on page 27]

F. M. Callier and J. Winkin. LQ-optimal control of infinite-dimensional systems by spectral factorization. *Automatica*, 28(4):757–770, 1992. [Cited on page 26]

F. P. Carli, A. Ferrante, M. Pavon, and G. Picci. A maximum entropy solution of the covariance extension problem for reciprocal processes. *IEEE Transactions on Automatic Control*, 56(9):1999–2012, 2011. [Cited on pages 114 and 120]

J. Carrasco and P. Seiler. Integral quadratic constraint theorem: A topological separation approach. In *Proceedings of the 54th IEEE Conference on Decision and Control (CDC)*, pages 5701–5706, 2015. [Cited on page 30]

P. Čebyšëv. Sur les valeurs limites des intégrales. *Journal de Mathématiques Pures et Appliquées*, pages 157–160, 1874. [Cited on page 114]

T. Chen and B. A. Francis. Spectral and inner-outer factorizations of rational matrices. *SIAM Journal on Matrix Analysis and Applications*, 10(1):1–17, 1989. [Cited on page 27]

D. J. Clements. Rational spectral factorization using state-space methods. *Systems & Control Letters*, 20(5):335–343, 1993. [Cited on page 27]

D. J. Clements. A state-space approach to indefinite spectral factorization. *SIAM Journal on Matrix Analysis and Applications*, 21(3):743–767, 2000. [Cited on pages 30, 31, and 93]

D. J. Clements and K. Glover. Spectral factorization via Hermitian pencils. *Linear Algebra and its Applications*, 122:797–846, 1989. [Cited on page 27]

D. J. Clements, B. D. O. Anderson, A. J. Laub, and J. B. Matson. Spectral factorization with imaginary-axis zeros. *Linear Algebra and its applications*, 250:225–252, 1997. [Cited on page 27]

P. Colaneri and A. Ferrante. A J-spectral factorization approach for $\mathscr{H}_\infty$ estimation problems in discrete time. *IEEE Transactions on Automatic Control*, 47(12):2108–2113, 2002. [Cited on page 30]

P. Colaneri and A. Ferrante. Algebraic Riccati equation and J-spectral factorization for $\mathscr{H}_\infty$ estimation. *Systems & Control Letters*, 51(5):383–393, 2004. [Cited on page 30]

P. Colaneri and A. Ferrante. Algebraic Riccati equation and J-spectral factorization for $\mathscr{H}_\infty$ smoothing and deconvolution. *SIAM Journal on Control and Optimization*, 45 (1):123–145, 2006. [Cited on pages 30, 31, 93, and 94]

P. Colaneri, M. Maroni, and U. Shaked. H_∞ prediction and smoothing for discrete-time systems: A J-spectral factorization approach. In *Proceedings of the 37th IEEE Conference on Decision and Control (CDC)*, pages 2836–2842, 1998. [Cited on page 30]

G. Corach and A. L. Maestripieri. Differential and metrical structure of positive operators. *Positivity*, 3(4):297–315, 1999. [Cited on page 165]

G. Corach and A. L. Maestripieri. Differential geometry on Thompson's components of positive operators. *Reports on Mathematical Physics*, 45(1):23–37, 2000. [Cited on page 165]

G. Corach, H. Porta, and L. Recht. Geodesics and operator means in the space of positive operators. *International Journal of Mathematics*, 4(02):193–202, 1993. [Cited on pages 165 and 168]

G. Corach, H. Porta, and L. Recht. Convexity of the geodesic distance on spaces of positive operators. *Illinois Journal of Mathematics*, 38(1):87–94, 1994. [Cited on pages 165 and 168]

T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley & Sons, 2nd edition, 2012. [Cited on pages 122 and 129]

K. De Cock and B. De Moor. Subspace angles between ARMA models. *Systems & Control Letters*, 46(4):265–270, 2002. [Cited on page 159]

M. Deistler, B. D. O. Anderson, A. Filler, Ch. Zinner, and W. Chen. Generalized linear dynamic factor models: An approach via singular autoregressions. *European Journal of Control*, 16(3):211–224, 2010. [Cited on page 26]

A. P. Dempster. Covariance selection. *Biometrics*, pages 157–175, 1972. [Cited on page 113]

S. Deng. *Homogeneous Finsler Spaces*. Springer Monographs in Mathematics. Springer, New York, 2012. [Cited on page 160]

P. Dewilde and J. Vandewalle. On the factorization of a nonsingular rational matrix. *IEEE Transactions on Circuits and Systems*, 22(8):637–645, 1975. [Cited on page 59]

J. C. Doyle, B. A. Francis, and A. R. Tannenbaum. *Feedback control theory*. Macmillan, New York, 1992. [Cited on page 120]

H. Dym and S. Nevo. Zero cancellation. *Linear Algebra and its Applications*, 404:1–26, 2005a. [Cited on page 59]

H. Dym and S. Nevo. Pole cancellation. *Linear Algebra and its Applications*, 404:27–57, 2005b. [Cited on page 59]

P. Enqvist. A homotopy approach to rational covariance extension with degree constraint. *International Journal of Applied Mathematics and Computer Science*, 11: 1173–1201, 2001. [Cited on page 125]

P. Enqvist. A convex optimization approach to ARMA (n, m) model design from covariance and cepstral data. *SIAM Journal on Control and Optimization*, 43(3): 1011–1036, 2004. [Cited on page 114]

L. Ephremidze, G. Janashia, and E. Lagvilava. A simple proof of the matrix-valued Fejér–Riesz theorem. *Journal of Fourier Analysis and Applications*, 15(1):124–127, 2009. [Cited on page 26]

P. Faurre, M. Clerget, and F. Germain. *Opérateurs rationnels positifs: Application à l'hyperstabilité et aux processus aléatoires*. Dunod, Paris, 1979. [Cited on page 26]

A. Ferrante. A parameterization of minimal stochastic realizations. *IEEE Transactions on Automatic Control*, 39(10):2122–2126, 1994. [Cited on pages 26 and 28]

A. Ferrante. A parametrization of the minimal square spectral factors of a nonrational spectral density. *Journal of Mathematical Systems Estimation and Control*, 7:197–226, 1997a. [Cited on page 28]

A. Ferrante. A homeomorphic characterization of minimal spectral factors. *SIAM Journal on Control and Optimization*, 35(5):1508–1523, 1997b. [Cited on page 28]

A. Ferrante. Minimal representations of continuous-time processes having spectral density with zeros in the extended imaginary axis. *Systems & Control Letters*, 54(5): 511–520, 2005. [Cited on page 28]

A. Ferrante and B. C. Levy. Canonical form of symplectic matrix pencils. *Linear Algebra and its Applications*, 274(1-3):259–300, 1998. [Cited on page 34]

A. Ferrante and L. Ntogramatzidis. The generalised discrete algebraic Riccati equation in linear-quadratic optimal control. *Automatica*, 49(2):471–478, 2013. [Cited on page 26]

A. Ferrante and L. Pandolfi. On the solvability of the positive real lemma equations. *Systems & Control Letters*, 47(3):211–219, 2002. [Cited on page 27]

A. Ferrante and G. Picci. A complete LMI/Riccati theory from stochastic modeling. In *Proceedings of 21st International Symposium on Mathematical Theory of Networks and Systems (MTNS), Groningen*, pages 1367–1374, 2014. [Cited on page 25]

A. Ferrante and G. Picci. Representation and factorization of discrete-time rational all-pass functions. *IEEE Transactions on Automatic Control*, 62(7):3262–3276, 2017. [Cited on pages 11, 29, 36, 69, 71, 72, 74, and 79]

A. Ferrante, Gy. Michaletzky, and M. Pavon. Parametrization of all minimal square spectral factors. *Systems & Control Letters*, 21(3):249–254, 1993. [Cited on page 28]

A. Ferrante, M. Pavon, and F. Ramponi. Further results on the Byrnes–Georgiou–Lindquist generalized moment problem. In *Modeling, Estimation and Control*, pages 73–83. Springer, Berlin Heidelberg, 2007. [Cited on pages 124, 127, and 173]

A. Ferrante, M. Pavon, and F. Ramponi. Hellinger versus Kullback–Leibler multivariable spectrum approximation. *IEEE Transactions on Automatic Control*, 53(4):954–967, 2008. [Cited on pages 26, 117, 121, 124, 147, and 159]

A. Ferrante, M. Pavon, and M. Zorzi. Application of a global inverse function theorem of Byrnes and Lindquist to a multivariable moment problem with complexity constraint. In *Three Decades of Progress in Control Sciences*, pages 153–167. Springer, Berlin Heidelberg, 2010. [Cited on pages 115, 147, 148, 149, 154, 155, and 173]

A. Ferrante, F. Ramponi, and F. Ticozzi. On the convergence of an efficient algorithm for Kullback–Leibler approximation of spectral densities. *IEEE Transactions on Automatic Control*, 56(3):506–515, 2011. [Cited on pages 115, 122, 123, 127, 128, 131, 133, 136, 139, 140, 141, 142, and 173]

A. Ferrante, C. Masiero, and M. Pavon. Time and spectral domain relative entropy: A new approach to multivariate spectral estimation. *IEEE Transactions on Automatic Control*, 57(10):2561–2575, 2012a. [Cited on pages 26, 147, 159, and 171]

A. Ferrante, M. Pavon, and M. Zorzi. A maximum entropy enhancement for a family of high-resolution spectral estimators. *IEEE Transactions on Automatic Control*, 57(2): 318–329, 2012b. [Cited on page 118]

A. Ferrante, A. Lanzon, and L. Ntogramatzidis. Discrete-time negative imaginary systems. *Automatica*, 79:1–10, 2017. [Cited on page 26]

L. Finesso and G. Picci. A characterization of minimal square spectral factors. *IEEE Transactions on Automatic Control*, 27(1):122–127, 1982. [Cited on page 28]

B. A. Francis. *A Course in $\mathcal{H}_\infty$ Control Theory*, volume 88. Springer-Verlag, Berlin Heidelberg, 1987. [Cited on page 93]

P. A. Fuhrmann. On the characterization and parametrization of minimal spectral factors. *Journal of Mathematical Systems, Estimation, and Control*, 5(4):383–444, 1995. [Cited on page 28]

P. A. Fuhrmann and A. Gombani. On a Hardy space approach to the analysis of spectral factors. *International Journal of Control*, 71(2):277–357, 1998. [Cited on page 28]

P. A. Fuhrmann and A. Gombani. On the Lyapunov equation, coinvariant subspaces and some problems related to spectral factorizations. *International Journal of Control*, 73(12):1129–1159, 2000. [Cited on page 28]

P. A. Fuhrmann and J. Hoffmann. Factorization theory for stable, discrete-time inner functions. *Journal of Mathematical Systems Estimation and Control*, pages 383–400, 1997. [Cited on page 69]

P. A. Fuhrmann and J. C. Willems. Factorization indices at infinity for rational matrix functions. *Integral Equations and Operator Theory*, 2(3):287–301, 1979. [Cited on page 85]

F. R. Gantmacher. *The theory of matrices*, volume I-II. AMS Chelsea Publishing Company, New York, 1959. [Cited on page 43]

Y. Genin, P. Van Dooren, and V. Vermaut. Convergence of the calculation of H_∞-norm and related questions. In *Proceedings of 13th International Symposium on Mathematical Theory of Networks and Systems (MTNS), Padova*, pages 429–432, 1998. [Cited on page 167]

T. T. Georgiou. *Partial realization of covariance sequences*. PhD thesis, Univ. Florida, Gainesville, 1983. [Cited on pages 114, 147, and 157]

T. T. Georgiou. Realization of power spectra from partial covariance sequences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(4):438–449, 1987a. [Cited on pages 114, 120, and 147]

T. T. Georgiou. A topological approach to Nevanlinna–Pick interpolation. *SIAM Journal on Mathematical Analysis*, 18(5):1248–1260, 1987b. [Cited on pages 114 and 147]

T. T. Georgiou. On a Schur-algorithm based approach to spectral factorization: State-space formulae. *Systems & Control Letters*, 10(2):123–129, 1988. [Cited on page 27]

T. T. Georgiou. The interpolation problem with a degree constraint. *IEEE Transactions on Automatic Control*, 44(3):631–635, 1999. [Cited on pages 114 and 120]

T. T. Georgiou. Spectral estimation via selective harmonic amplification. *IEEE Transactions on Automatic Control*, 46(1):29–42, 2001. [Cited on page 113]

T. T. Georgiou. The structure of state covariances and its relation to the power spectrum of the input. *IEEE Transactions on Automatic Control*, 47(7):1056–1066, 2002. [Cited on page 121]

T. T. Georgiou. Solution of the general moment problem via a one-parameter imbedding. *IEEE Transactions on Automatic Control*, 50(6):811–826, 2005. [Cited on pages 114 and 117]

T. T. Georgiou. Relative entropy and the multivariable multidimensional moment problem. *IEEE Transactions on Information Theory*, 52(3):1052–1066, 2006. [Cited on pages 114, 147, and 159]

T. T. Georgiou. Distances and Riemannian metrics for spectral density functions. *IEEE Transactions on Signal Processing*, 55(8):3995–4003, 2007a. [Cited on pages 122, 159, and 167]

T. T. Georgiou. An intrinsic metric for power spectral density functions. *IEEE Signal Processing Letters*, 14(8):561–563, 2007b. [Cited on pages 122 and 159]

T. T. Georgiou. The meaning of distances in spectral analysis. http://www.ieeecss-oll.org/lectures/2007/meaning-distances-spectral-analysis, 2007c. Plenary lecture, 46th IEEE Conference on Decision and Control, New Orleans, USA. [Cited on page 159]

T. T. Georgiou and A. Lindquist. Kullback–Leibler approximation of spectral density functions. *IEEE Transactions on Information Theory*, 49(11):2910–2917, 2003. [Cited on pages 17, 113, 115, 117, 121, 122, 123, 124, 125, and 159]

T. T. Georgiou and A. Lindquist. Remarks on control design with degree constraint. *IEEE Transactions on Automatic Control*, 51(7):1150–1156, 2006. [Cited on page 114]

T. T. Georgiou and A. Lindquist. A convex optimization approach to ARMA modeling. *IEEE Transactions on Automatic Control*, 53(5):1108–1119, 2008. [Cited on page 114]

T. T. Georgiou and A. Lindquist. Likelihood analysis of power spectra and generalized moment problems. *IEEE Transactions on Automatic Control*, 62(9):4580–4592, 2017. [Cited on page 147]

T. T. Georgiou, J. Karlsson, and M. S. Takyar. Metrics for power spectra: An axiomatic approach. *IEEE Transactions on Signal Processing*, 57(3):859–867, 2009. [Cited on page 159]

K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L_∞-error bounds. *International Journal of Control*, 39(6):1115–1193, 1984. [Cited on page 69]

K.-C. Goh. Structure and factorization of quadratic constraints for robustness analysis. In *Proceedings of the 35th IEEE Conference on Decision and Control (CDC)*, pages 4649–4654, 1996. [Cited on page 30]

I. Gohberg, P. Lancaster, and L. Rodman. Factorization of selfadjoint matrix polynomials with constant signature. *Linear and Multilinear Algebra*, 11(3):209–224, 1982. [Cited on pages 30 and 94]

G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins University Press, Baltimore, MD, USA, 1996. [Cited on page 51]

W. B. Gordon. On the diffeomorphisms of Euclidean space. *The American Mathematical Monthly*, 79(7):755–759, 1972. [Cited on page 174]

R. Gray, A. Buzo, A. Gray, and Y. Matsuyama. Distortion measures for speech processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4):367–376, 1980. [Cited on pages 159 and 170]

M. Green and B. D. O. Anderson. Factorization of all-pass matrix functions. *International Journal of Control*, 45(1):215–235, 1987. [Cited on page 69]

M. Green, K. Glover, D. Limebeer, and J. Doyle. A J-spectral factorization approach to $\mathscr{H}_\infty$ control. *SIAM Journal on Control and Optimization*, 28(6):1350–1371, 1990. [Cited on pages 29 and 93]

U. Grenander and G. Szegö. *Toeplitz forms and their applications*. University of California Press, Berkeley, CA, 1958. [Cited on page 114]

G. Gu, X.-R. Cao, and H. Badr. Generalized LQR control and Kalman filtering with relations to computations of inner-outer and spectral factorizations. *IEEE Transactions on Automatic Control*, 51(4):595–605, 2006. [Cited on page 26]

H. Hamburger. Über eine Erweiterung des Stieltjesschen Momentenproblems. *Mathematische Annalen*, 81(2):235–319, 1920. [Cited on page 114]

B. Hassibi, A. H. Sayed, and T. Kailath. *Indefinite-Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories*, volume 16. Society for Industrial and Applied Mathematics (SIAM), 1999. [Cited on page 29]

F. Hausdorff. Summationsmethoden und Momentfolgen. I, II. *Mathematische Zeitschrift*, 9(1):74–109, 280–299, 1921. [Cited on page 114]

D. Hayden, Y. Yuan, and J. Gonçalves. Network identifiability from intrinsic noise. *IEEE Transactions on Automatic Control*, 62(8):3717–3728, 2017. [Cited on page 26]

B. Hu, M. J. Lacerda, and P. Seiler. Robustness analysis of uncertain discrete-time systems with dissipation inequalities and integral quadratic constraints. *International Journal of Robust and Nonlinear Control*, 2016. [Cited on page 30]

O. V. Iftime. *A J-spectral factorization approach to H_∞ control problems*. PhD thesis, Universiteit Twente, 2002. [Cited on page 29]

V. Ionescu, C. Oară, and M. Weiss. *Generalized Riccati theory and robust control: A Popov function approach*. John Wiley, 1999. [Cited on pages 93 and 153]

G. Janashia, E. Lagvilava, and L. Ephremidze. A new method of matrix spectral factorization. *IEEE Transactions on information theory*, 57(4):2318–2326, 2011. [Cited on page 27]

E. T. Jaynes. On the rationale of maximum-entropy methods. *Proceedings of the IEEE*, 70(9):939–952, 1982. [Cited on page 113]

J. Ježek and V. Kučera. Efficient algorithm for matrix spectral factorization. *Automatica*, 21(6):663–669, 1985. [Cited on page 27]

X. Jiang, Z.-Q. Luo, and T. T. Georgiou. Power spectral geodesics and tracking. In *Proceedings of the 47th IEEE Conference on Decision and Control (CDC)*, pages 1315–1319, 2008a. [Cited on page 159]

X. Jiang, S. Takyar, and T. T. Georgiou. Metrics and morphing of power spectra. In *Recent Advances in Learning and Control*, pages 125–135. Springer, 2008b. [Cited on page 159]

X. Jiang, Z.-Q. Luo, and T. T. Georgiou. Geometric methods for spectral analysis. *IEEE Transactions on Signal Processing*, 60(3):1064–1074, 2012a. [Cited on pages 159 and 160]

X. Jiang, L. Ning, and T. T. Georgiou. Distances and Riemannian metrics for multivariate spectral densities. *IEEE Transactions on Automatic Control*, 57(7):1723–1735, 2012b. [Cited on pages 159, 170, and 171]

T. Kailath. *Linear systems*, volume 156. Prentice-Hall, Englewood Cliffs, NJ, 1980. [Cited on pages 38, 39, 42, 44, 61, and 97]

R. E. Kálmán. Realization of covariance sequences. In *Toeplitz Centennial*, pages 331–342. Springer, 1982. [Cited on page 114]

I. Karelin, L. Lerer, and A. C. M. Ran. *J*-symmetric factorizations and algebraic Riccati equations. In *Recent Advances in Operator Theory*, pages 319–360. Springer, 2001. [Cited on page 30]

J. Karlsson, A. Lindquist, and A. Ringh. The multidimensional moment problem with complexity constraint. *Integral Equations and Operator Theory*, 84(3):395–418, 2016. [Cited on page 114]

H. K. Khalil. *Nonlinear systems*. Prentice-Hall, Upper Saddle River, NJ, 1996. [Cited on pages 26 and 133]

T. H. Kjeldsen. The early history of the moment problem. *Historia Mathematica*, 20 (1):19–44, 1993. [Cited on page 114]

F. Klein. Vergleichende Betrachtungen über neuere geometrische Forschungen. *Mathematische Annalen*, 43(1):63–100, 1893. [Cited on page 159]

E. Kohlberg and J. W. Pratt. The contraction mapping approach to the Perron-Frobenius theory: Why Hilbert's metric? *Mathematics of Operations Research*, 7(2):198–210, 1982. [Cited on page 163]

A. N. Kolmogorov. Sur l'interpolation et extrapolation des suites stationnaires. *Comptes rendus de l'Académie des Sciences*, 208:2043–2045, 1939. [Cited on pages 15 and 26]

S. G. Krantz and H. R. Parks. *The implicit function theorem: History, theory, and applications*. Modern Birkhäuser Classics. Birkhäuser, Basel, 2013. [Cited on page 174]

M. G. Kreĭn and A. A. Nudel'man. *The Markov moment problem and extremal problems*, volume 50 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1977. [Cited on pages 17, 113, and 114]

H. Kwakernaak. Frequency domain solution of the standard H_∞ problem. In *Polynomial methods for control systems design*, pages 57–105. Springer, 1996. [Cited on page 29]

H. Kwakernaak and M. Sebek. Polynomial J-spectral factorization. *IEEE Transactions on Automatic Control*, 39(2):315–328, 1994. [Cited on page 31]

S. Lang. *Complex analysis*. Graduate texts in mathematics. Springer-Verlag, New York, 1985. [Cited on page 40]

J. P. LaSalle. *The Stability and Control of Discrete Processes*, volume 62 of *Applied Mathematical Sciences*. Springer Verlag, New York, 1986. [Cited on page 133]

O. Lauwers and B. De Moor. A time series distance measure for efficient clustering of input/output signals by their underlying dynamics. *IEEE Control Systems Letters*, 1 (2):286–291, 2017. [Cited on page 159]

B. Lemmens and R. Nussbaum. *Nonlinear Perron-Frobenius Theory*, volume 189 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2012. [Cited on pages 160, 163, 164, and 165]

L. Lerer, M. A. Petersen, and A. C. M. Ran. Existence of minimal nonsquare J-symmetric factorizations for self-adjoint rational matrix functions. *Linear Algebra and its Applications*, 379:159–178, 2004. [Cited on page 30]

A. Lindquist and G. Picci. On the stochastic realization problem. *SIAM Journal on Control and Optimization*, 17(3):365–389, 1979. [Cited on pages 26 and 69]

A. Lindquist and G. Picci. Realization theory for multivariate stationary Gaussian processes. *SIAM Journal on Control and optimization*, 23(6):809–857, 1985. [Cited on page 26]

A. Lindquist and G. Picci. A geometric approach to modelling and estimation of linear stochastic systems. *Journal of Mathematical Systems, Estimation and Control*, 1(3): 241–333, 1991. [Cited on pages 25, 26, and 69]

A. Lindquist and G. Picci. The circulant rational covariance extension problem: The complete solution. *IEEE Transactions on Automatic Control*, 58(11):2848–2861, 2013. [Cited on pages 114 and 120]

A. Lindquist and G. Picci. *Linear Stochastic Systems*, volume 1 of *Series in Contemporary Mathematics*. Springer-Verlag, Berlin Heidelberg, 2015. [Cited on pages 26 and 71]

A. Lindquist, Gy. Michaletzky, and G. Picci. Zeros of spectral factors, the geometry of splitting subspaces, and the algebraic Riccati inequality. *SIAM Journal on Control and Optimization*, 33(2):365–401, 1995. [Cited on page 26]

A. Lindquist, C. Masiero, and G. Picci. On the multivariate circulant rational covariance extension problem. In *Proceedings of the 52nd IEEE Conference on Decision and Control (CDC)*, pages 7155–7161, 2013. [Cited on pages 114 and 120]

C. Liverani and M. P. Wojtkowski. Generalization of the Hilbert metric to the space of positive definite matrices. *Pacific Journal of Mathematics*, 166(2):339–355, 1994. [Cited on page 163]

R. J. Martin. A metric for ARMA processes. *IEEE Transactions on Signal Processing*, 48 (4):1164–1170, 2000. [Cited on pages 122, 159, and 170]

D. Materassi and G. Innocenti. Topological identification in networks of dynamical systems. *IEEE Transactions on Automatic Control*, 55(8):1860–1871, 2010. [Cited on page 26]

D. Materassi and M. V. Salapaka. On the problem of reconstructing an unknown topology via locality properties of the Wiener filter. *IEEE Transactions on Automatic Control*, 57(7):1765–1777, 2012. [Cited on page 26]

J. H. McClellan. Multidimensional spectral estimation. *Proceedings of the IEEE*, 70(9): 1029–1039, 1982. [Cited on page 113]

G. Meinsma. Polynomial solutions to H_∞ problems. *International Journal of Robust and Nonlinear Control*, 4(3):323–351, 1994. [Cited on page 29]

G. Meinsma. J-spectral factorization and equalizing vectors. *Systems & Control Letters*, 25(4):243–249, 1995. [Cited on page 30]

Gy. Michaletzky. A note on the factorization of discrete time inner functions. *Journal of Mathematical Systems Estimation and Control*, 8(4):479–482, 1998. [Cited on page 69]

R. Nagamune. A robust solver using a continuation method for Nevanlinna–Pick interpolation with degree constraint. *IEEE Transactions on Automatic Control*, 48 (1):113–117, 2003. [Cited on page 125]

R. W. Newcomb. *Linear multiport synthesis*. McGraw-Hill, New York, 1966. [Cited on page 26]

H. I. Nurdin. Spectral factorization of a class of matrix-valued spectral densities. *SIAM Journal on Control and Optimization*, 45(5):1801–1821, 2006. [Cited on page 131]

R. D. Nussbaum. Finsler structures for the part metric and Hilbert's projective metric and applications to ordinary differential equations. *Differential and Integral Equations*, 7(5-6):1649–1707, 1994. [Cited on page 163]

R. D. Nussbaum and C. Walsh. A metric inequality for the Thompson and Hilbert geometries. *Journal of Inequalities in Pure and Applied Mathematics*, 5(3), 2004. [Cited on page 164]

C. Oară. Constructive solutions to spectral and inner–outer factorizations with respect to the disk. *Automatica*, 41(11):1855–1866, 2005. [Cited on pages 27 and 167]

C. Oară and R. Andrei. Zero cancellation for general rational matrix functions. *Linear Algebra and its Applications*, 431(11):2000–2023, 2009. [Cited on page 59]

C. Oară and R. Andrei. Computation of the general (J, J')-lossless factorization. *IEEE Transactions on Automatic Control*, 56(3):710–717, 2011. [Cited on page 31]

C. Oară and R. Marinică. J factorizations of a general discrete-time system. *Automatica*, 49(7):2221–2228, 2013. [Cited on page 31]

C. Oară and Ş. Sabău. Minimal indices cancellation and rank revealing factorizations for rational matrix functions. *Linear Algebra and its Applications*, 431(10):1785–1814, 2009. [Cited on page 59]

C. Oară and A. Varga. Computation of general inner-outer and spectral factorizations. *IEEE Transactions on Automatic Control*, 45(12):2307–2325, 2000. [Cited on page 27]

D. Ž. Đoković. Factorization of hermitian matrix polynomials with constant signature. *Linear Algebra and its Applications*, 194:85–90, 1993. [Cited on pages 30 and 94]

Y. Oono and K. Yasuura. Synthesis of finite passive $2n$-terminal networks with prescribed scattering matrices. *Memoirs of the Faculty of Engineering, Kyushu University*, 14:125–177, 1954. [Cited on pages 33 and 49]

H. J. Orchard and A. N. Wilson. On the computation of a minimum-phase spectral factor. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 50(3):365–375, 2003. [Cited on page 27]

E. Outerelo and J. M. Ruiz. *Mapping degree theory*, volume 108. American Mathematical Society, Providence, RI, 2009. [Cited on pages 149, 150, 151, and 155]

L. Pandolfi. An observation on the positive real lemma. *Journal of Mathematical Analysis and Applications*, 255(2):480–490, 2001. [Cited on page 27]

P. Park and T. Kailath. Dual J-spectral factorization. *International Journal of Control*, 68(2):287–296, 1997. [Cited on page 31]

M. Pavon. On the parametrization of non-square spectral factors. In *U. Helmke, R. Mennicken, J. Sauer (Eds.), Systems and Networks: Mathematical Theory and Applications*, volume 79, pages 413–416. Akademie Verlag, Berlin, 1994. [Cited on page 28]

M. Pavon and A. Ferrante. A new algorithm for Kullback–Leibler approximation of spectral densities. In *Proceedings of the 44th IEEE Conference on Decision and Control (CDC)*, pages 7332–7337, 2005. [Cited on pages 123, 127, and 128]

M. Pavon and A. Ferrante. On the Georgiou–Lindquist approach to constrained Kullback–Leibler approximation of spectral densities. *IEEE Transactions on Automatic Control*, 51(4):639–644, 2006. [Cited on pages 17, 115, 123, 127, 128, 129, and 173]

M. Pavon and A. Ferrante. On the geometry of maximum entropy problems. *SIAM Review*, 55(3):415–439, 2013. [Cited on page 113]

I. R. Petersen and A. Lanzon. Feedback control of negative-imaginary systems. *IEEE Control Systems*, 30(5):54–72, 2010. [Cited on page 26]

M. A. Petersen and A. C. M. Ran. Minimal square spectral factors via triples. *SIAM Journal on Matrix Analysis and Applications*, 22(4):1222–1244, 2001. [Cited on page 28]

M. A. Petersen and A. C. M. Ran. Minimal nonsquare spectral factors. *Linear Algebra and its Applications*, 351:553–565, 2002a. [Cited on page 28]

M. A. Petersen and A. C. M. Ran. Nonsquare spectral factors via factorizations of a unitary function. *Linear Algebra and its Applications*, 351:567–583, 2002b. [Cited on page 28]

M. A. Petersen and A. C. M. Ran. Minimal nonsquare J-spectral factorization, generalized Bezoutians and common zeros for rational matrix functions. *Integral Equations and Operator Theory*, 47(2):197–216, 2003. [Cited on page 30]

G. Picci. Stochastic realization of Gaussian processes. *Proceedings of the IEEE*, 64(1): 112–122, 1976. [Cited on page 26]

G. Picci. Some remarks on discrete-time unstable spectral factorization. In *Mathematical System Theory, Festschrift in Honor of Uwe Helmke on the Occasion of His Sixtieth Birthday*, pages 301–310. K. Hüper and J. Trumpf Editors, Springer, Berlin, 2013. [Cited on page 69]

G. Picci and S. Pinzoni. Acausal models and balanced realizations of stationary processes. *Linear Algebra and its Applications*, 205:997–1043, 1994. [Cited on page 25]

S. U. Pillai and T. I. Shim. *Spectrum estimation and system identification*. Springer-Verlag, New York, 2012. [Cited on page 113]

F. Ramponi, A. Ferrante, and M. Pavon. A globally convergent matricial algorithm for multivariate spectral estimation. *IEEE Transactions on Automatic Control*, 54(10): 2376–2388, 2009. [Cited on pages 26, 118, and 147]

A. Ran and P. Zizler. On self-adjoint matrix polynomials with constant signature. *Linear Algebra and its Applications*, 259:133–153, 1997. [Cited on page 30]

A. C. M. Ran. Minimal factorization of selfadjoint rational matrix functions. *Integral Equations and Operator Theory*, 5(1):850–869, 1982. [Cited on pages 27 and 28]

A. C. M. Ran. Minimal square spectral factors. *Systems & Control Letters*, 24(5):307–316, 1995. [Cited on page 28]

A. C. M. Ran. Necessary and sufficient conditions for existence of J-spectral factorization for para-Hermitian rational matrix functions. *Automatica*, 39(11):1935–1939, 2003. [Cited on page 30]

A. C. M. Ran and L. Rodman. Stable invariant Lagrangian subspaces: factorization of symmetric rational matrix functions and other applications. *Linear Algebra and its Applications*, 137:575–620, 1990. [Cited on page 27]

A. C. M. Ran and L. Rodman. On symmetric factorizations of rational matrix functions. *Linear and Multilinear Algebra*, 29(3-4):243–261, 1991. [Cited on pages 30 and 93]

J. Rissanen. Algorithms for triangular decomposition of block Hankel and Toeplitz matrices with application to factoring positive matrix polynomials. *Mathematics of Computation*, 27(121):147–154, 1973. [Cited on page 27]

Yu. A. Rozanov. *Stationary random processes*. Holden-Day, San Francisco, 1967. [Cited on page 130]

G. Ruckebusch. Théorie géométrique de la Représentation Markovienne. *Annales de l'Institut Henri Poincaré*, 16(3):225–297, 1980. [Cited on page 26]

W. Rudin. *Real and complex analysis*. McGraw-Hill, New York, 1987. [Cited on pages 131 and 152]

A. H. Sayed and T. Kailath. A survey of spectral factorization methods. *Numerical Linear Algebra with Applications*, 8(6-7):467–496, 2001. [Cited on page 27]

M. Schatzman. *Numerical analysis: a mathematical introduction*. Oxford University Press, 2002. [Cited on page 153]

J. T. Schwartz. *Nonlinear functional analysis*. Lecture Notes in Pure and Applied Mathematics. CRC Press, 1969. [Cited on page 151]

P. Seiler. Stability analysis with dissipation inequalities and integral quadratic constraints. *IEEE Transactions on Automatic Control*, 60(6):1704–1709, 2015. [Cited on page 30]

U. Shaked and I. Yaesh. A simple method for deriving J-spectral factors. *IEEE Transactions on Automatic Control*, 37(6):891–895, 1992. [Cited on page 31]

J. A. Shohat and J. D. Tamarkin. *The problem of moments*, volume 1 of *Mathematical Surveys and Monographs*. American Mathematical Society, New York, 1943. [Cited on page 113]

S. T. Smith. Covariance, subspace, and intrinsic Cramér–Rao bounds. *IEEE Transactions on Signal Processing*, 53(5):1610–1630, 2005. [Cited on page 170]

J. Stefanovski. Polynomial J-spectral factorization in minimal state space. *Automatica*, 39(11):1893–1901, 2003. [Cited on page 31]

J. Stefanovski. Discrete J-spectral factorization of possibly singular polynomial matrices. *Systems & Control Letters*, 53(2):127–140, 2004. [Cited on page 31]

G. W. Stewart. Error and perturbation bounds for subspaces associated with certain eigenvalue problems. *SIAM Review*, 15(4):727–764, 1973. [Cited on page 154]

T. J. Stieltjes. Recherches sur les fractions continues. In *Annales de la Faculté des Sciences de Toulouse*, volume 8, pages A1–A47, 1894. [Cited on page 114]

P. Stoica and R. L. Moses. *Introduction to spectral analysis*, volume 1. Prentice-Hall, Upper Saddle River, NJ, 1997. [Cited on page 113]

A. A. Stoorvogel. Stabilizing solutions of the H_∞ algebraic Riccati equation. *Linear Algebra and its Applications*, 240:153–172, 1996. [Cited on page 29]

A. A. Stoorvogel and A. Saberi. The discrete algebraic Riccati equation and linear matrix inequality. *Linear Algebra and its Applications*, 274(1-3):317–365, 1998. [Cited on page 26]

J. Swigart and S. Lall. Optimal controller synthesis for decentralized systems over graphs via spectral factorization. *IEEE Transactions on Automatic Control*, 59(9): 2311–2323, 2014. [Cited on page 26]

K. Takaba, N. Morihira, and T. Katayama. $\mathbf{H}_\infty$ control for descriptor systems – a J-spectral factorization approach –. In *Proceedings of the 33rd IEEE Conference on Decision and Control (CDC)*, pages 2251–2256, 1994. [Cited on page 29]

A. C. Thompson. On certain contraction mappings in a partially ordered vector space. *Proceedings of the American Mathematical Society*, 14(3):438–443, 1963. [Cited on page 163]

H. L. Trentelman and P. Rapisarda. New algorithms for polynomial J-spectral factorization. *Mathematics of Control, Signals and Systems*, 12(1):24–61, 1999. [Cited on page 31]

G. Tunnicliffe-Wilson. The factorization of matricial spectral densities. *SIAM Journal on Applied Mathematics*, 23(4):420–426, 1972. [Cited on page 27]

A. J. Van Der Schaft and J. C. Willems. A new procedure for stochastic realization of spectral density matrices. *SIAM Journal on Control and Optimization*, 22(6): 845–855, 1984. [Cited on page 27]

P. Van Dooren. Rational and polynomial matrix factorizations via recursive pole-zero cancellation. *Linear Algebra and its Applications*, 137:663–697, 1990. [Cited on page 59]

J. Vandewalle and P. Dewilde. On the irreducible cascade synthesis of a system with a real rational transfer matrix. *IEEE Transactions on Circuits and Systems*, 24(9): 481–494, 1977. [Cited on page 59]

S. Wang, H. Pfifer, and P. Seiler. Robust synthesis for linear parameter varying systems using integral quadratic constraints. *Automatica*, 68:111–118, 2016. [Cited on page 30]

H. H. M. Weerts, P. M. J. Van den Hof, and A. G. Dankers. Identifiability of linear dynamic networks. 2017. [arXiv preprint: arXiv:1609.00864]. [Cited on page 26]

N. Wiener. *Extrapolation, interpolation, and smoothing of stationary time series*. Wiley, New York, 1949. [Cited on pages 15 and 26]

J. Willems. Least squares stationary optimal control and the algebraic Riccati equation. *IEEE Transactions on Automatic Control*, 16(6):621–634, 1971. [Cited on page 26]

H. K. Wimmer. Spectral factorizations with unmixed pole and zero sets. *Automatica*, 33(2):235–238, 1997. [Cited on page 27]

V. A. Yakubovich. On synthesis of optimal controls in a linear differential game on a finite time interval with quadratic payoff. *Soviet Mathematics Doklady*, 12(5): 1454–1458, 1971. [Cited on page 29]

D. C. Youla. On the factorization of rational matrices. *IRE Transactions on Information Theory*, 7(3):172–189, 1961. [Cited on pages 15, 27, 28, 33, 35, 36, 37, 44, 47, and 49]

D. C. Youla. *Theory and Synthesis of Linear Passive Time-Invariant Networks*. Cambridge University Press, 2015. [Cited on page 26]

D. C. Youla and N. Kazanjian. Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle. *IEEE Transactions on Circuits and Systems*, 25(2):57–69, 1978. [Cited on page 27]

Q.-C. Zhong. J-spectral factorization of regular para-Hermitian transfer matrices. *Automatica*, 41(7):1289–1293, 2005. [Cited on page 31]

K. Zhou, J. C. Doyle, and K. Glover. *Robust and optimal control*. Prentice-Hall, Englewood Cliffs, NJ, 1996. [Cited on pages 36, 165, 167, and 171]

B. Zhu and G. Baggio. On the existence of a solution to a spectral estimation problem *à la* Byrnes–Georgiou–Lindquist. *Submitted for publication*, 2017. [arXiv preprint: arXiv:1709.09012]. [Cited on page 148]

M. Zorzi. A new family of high-resolution multivariate spectral estimators. *IEEE Transactions on Automatic Control*, 59(4):892–904, 2014a. [Cited on pages 147 and 159]

M. Zorzi. Rational approximations of spectral densities based on the alpha divergence. *Mathematics of Control, Signals, and Systems*, 26(2):259–278, 2014b. [Cited on page 159]

M. Zorzi. Multivariate spectral estimation based on the concept of optimal prediction. *IEEE Transactions on Automatic Control*, 60(6):1647–1652, 2015. [Cited on pages 147 and 159]

M. Zorzi and A. Ferrante. On the estimation of structured covariance matrices. *Automatica*, 48(9):2145–2151, 2012. [Cited on page 118]

J. C. Zúñiga and D. Henrion. A Toeplitz algorithm for polynomial J-spectral factorization. *Automatica*, 42(7):1085–1093, 2006. [Cited on page 31]

Acknowledgements

This two-hundred(-and-a-little-more)-page manuscript comes at the end of an intense and exciting (yet sometimes tortuous and frustrating) journey — the Ph.D. For all the (admittedly, rather modest) achievements collected along this journey, my deepest gratitude goes to the people who guided, accompanied and supported me during these three years:

- Augusto Ferrante, my advisor, for his constant encouragement and admirable dedication to research. You are the profound thinker (not just an extraordinarily skilled mathematician) that I've always dreamed to be.
- Sandro Zampieri, (my co-advisor, I would say) for his constructive criticism and healthy realism. If I will survive in the (irreversibly deteriorating, *alas!*) Academia environment the credit is certainly yours!
- Rodolphe Sepulchre, for his intellectual curiosity, contagious enthusiasm, and warm hospitality during my visit in Cambridge.
- All the other formidable scholars with whom I collaborated and/or interact. Among them, a special mention goes to Francesco Ticozzi and Fulvio Forni (two exceptional minds!).
- $\mathcal{N}$., to whom this manuscript is dedicated, for a 1962 b/w Sicilian movie, an hot chocolate turned into spritz, and an endless number of other *"cose piccole e serene"*. For standing at my side in both the hard and good times, it is almost impossible to put into words how much I owe you.
- All the officemates (in Padova and Cambridge) for the time spent together laughing, sharing ideas and frustrations. In particular, I enormously thank all the office-329&330's (ex-)mates: It will be extremely hard to find such a wonderful working environment!
- My family, for offering me the opportunity of deepening my knowledge and for supporting me for all these years.

Index