Available online at www.sciencedirect.com

ScienceDirect

Journal of Approximation Theory 257 (2020) 105450

Journal of Approximation Theory

www.elsevier.com/locate/jat

Full Length Article

Calculating the spectral factorization and outer functions by sampling-based approximations—Fundamental limitations☆

Holger Boche[a,b], Volker Pohl[a,*]

[a] *Technical University of Munich, Chair of Theoretical Information Technology, Arcisstraße 21, 80333 Munich, Germany*

[b] *Munich Center for Quantum Science and Technology (MCQST), Schellingstraße 4, 80799 Munich, Germany*

Received 6 March 2019; received in revised form 6 April 2020; accepted 29 May 2020
Available online 3 June 2020

Communicated by K. Groechenig

Abstract

This paper considers the problem of approximating the spectral factor of continuous spectral densities with finite Dirichlet energy based on finitely many samples of these spectral densities. Although there exists a closed form expression for the spectral factor, this formula shows a very complicated behavior because of the non-linear dependency of the spectral factor from spectral density and because of a singular integral in this expression. Therefore approximation methods are usually applied to calculate the spectral factor.

It is shown that there exists no sampling-based method which depends continuously on the samples and which is able to approximate the spectral factor for all densities in this set. Instead, to any sampling-based approximation method there exists a large set of spectral densities so that the approximation method does not converge to the spectral factor for every spectral density in this set as the number of available sampling points is increased. The paper will also show that the same results hold for sampling-based algorithms for the calculation of the outer function in the theory of Hardy spaces.

☆ This work was partly supported by Germany's Excellence Strategy, EXC-2111-390814868. H. Boche is grateful for the support of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within the Gottfried Wilhelm Leibniz-program under Grant BO 1734/20-1. V. Pohl acknowledges the support of the DFG under Grant PO 1347/3-1. The material in this paper was presented in part at the 58th IEEE Conference on Decision and Control (CDC), Nice, France, Dec. 2019.

* Corresponding author.
E-mail addresses: boche@tum.de (H. Boche), volker.pohl@tum.de (V. Pohl).

https://doi.org/10.1016/j.jat.2020.105450

MSC: 41A65; 47A68; 65D15; 94A20

Keywords: Approximation; Sampling; Spectral factorization; Outer functions; Finite energy; Uniform recovery

Contents

1. Introduction

Spectral factorization is an important operation in many areas of signal processing and engineering and it is closely related to the inner–outer factorization in Hardy space theory. Best known is certainly its application in the Wiener–Kolmogorov theory of smoothing and prediction for stationary time series, which was initialized by the seminal works of Kolmogorov and Wiener [28,44,45] and which found multiple generalizations and extensions in different areas of science and engineering such as in estimation and filtering [13,20,46], in robust and optimal control [39,48], or in the theory of stochastic realizations [30], to mention only a few. Moreover, spectral factorization plays a key role in the definition and determination of several metrics (e.g. Riemannian and Finsler) between spectral densities [4,26] which are applied in numerous signal processing and control applications.

Let ϕ be a *spectral density*, i.e. a non-negative real function on the *unit circle* $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ satisfying the *Szegő condition* $\log \phi \in L^1(\mathbb{T})$ (also known as *Paley–Wiener condition*). Then *spectral factorization* is the process of writing ϕ uniquely as a product of a function $\phi_+(z)$ and the functions $\phi_-(z) = \overline{\phi_+(1/\overline{z})}$, i.e.

$$\phi(\zeta) = \phi_+(\zeta)\,\phi_-(\zeta) = \left|\phi_+(\zeta)\right|^2 \qquad \text{for all } \zeta \in \mathbb{T}\,. \tag{1}$$

Therein, the *spectral factor*[1] is an analytic function for all z in the *unit disk* $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ and satisfies $\phi_+(z) \neq 0$ for all $z \in \mathbb{D}$.

[1] Sometimes, one considers factorizations where the spectral factor ϕ_+ is not required to be analytic and non-zero in $\mathbb{D}$. Then the factorization is no longer unique and a complete characterization of all such spectral factors for rational spectral densities can be found in [3].

The problem of determining the spectral factor ϕ_+ of a given spectral density ϕ arises in many different applications such as estimation and prediction, communications, signal processing, and control theory [27,39,43,46]. In particular, it is a fundamental operation in the design and determination of the so called *Wiener filter*, frequently used in almost every area of signal processing [14,19,21,22]. For all these applications, it is important to have effective design methods for the Wiener filter with a corresponding analysis of the stability and robustness with respect to uncertainty in the spectral factors [6]. Technically, there exists a closed form expression for the *spectral factorization mapping* $\mathrm{S} : \phi \mapsto \phi_+$, given by

$$\phi_+(z) = (\mathrm{S}\phi)(z) = \exp\left(\frac{1}{4\pi}\int_{-\pi}^{\pi} \log \phi(\mathrm{e}^{\mathrm{i}\tau}) \frac{\mathrm{e}^{\mathrm{i}\tau}+z}{\mathrm{e}^{\mathrm{i}\tau}-z}\mathrm{d}\tau\right), \qquad z \in \mathbb{D}\,. \tag{2}$$

Nevertheless, because of the apparent non-linear relation between ϕ and its spectral factor ϕ_+ and because of the singular integral in (2), this mapping shows a very complicated behavior [1,5,7,8,23,24]. Therefore (2) is rarely used for calculating ϕ_+ but different classes of approximation algorithms were developed for calculating the spectral factor [2,18,37,47].

Nowadays all such algorithms are assumed to run on a digital computer. Therefore only finitely many input numbers can be processed and so any algorithm for calculating the spectral factor ϕ_+ on a digital computer has to be *sampling-based*, i.e. it can use only finitely many samples $\{\phi(\zeta_n)\}_{n=1}^N$ of ϕ, taken on a certain sampling set $\mathcal{T}_N = \{\zeta_n\}_{n=1}^N \subset \mathbb{T}$. Then, we do not expect that the algorithm determines the exact spectral factor ϕ_+ but only an approximation $\mathrm{S}_N(\phi)$. However, we do expect that the approximation error can be made arbitrarily small by increasing the size N of the sampling set $\mathcal{T}_N$. Thus we expect that

$$\lim_{N\to\infty} \|\phi_+ - \mathrm{S}_N(\phi)\|_{\mathscr{D}} = 0 \qquad \text{for all } \phi \in \mathscr{D}\,, \tag{3}$$

wherein $\mathscr{D}$ stands for the set of spectral densities under consideration and the approximation is considered with respect to an appropriate norm $\|\cdot\|_{\mathscr{D}}$. This paper asks the following question: Is it possible to find an appropriate sequence $\{\mathcal{T}_N\}_{N\in\mathbb{N}} \subset \mathbb{T}$ of sampling sets together with a corresponding sequence $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ of sampling-based approximation operators such that (3) is satisfied? This problem is investigated for spectral densities in the set $\mathscr{D}_+$ of strictly positive and continuous spectral densities of finite Dirichlet energy with the property that ϕ_+ is continuous on the closed unit disk $\overline{\mathbb{D}} = \mathbb{D}\cup\mathbb{T}$. In other words, we consider this problem on a set of spectral densities with very good analytic properties (continuous densities of finite energy where even the spectral factor is continuous). Despite these simple assumptions, we are going to show that there exists no sequence $\boldsymbol{S} = \{\mathrm{S}_N\}_{N\in\mathbb{N}}$ of sampling-based approximation operators such that (3) is satisfied for all $\phi \in \mathscr{D}_+$ and with respect to the uniform norm. Actually, it is shown that to every sequence $\boldsymbol{S}$ there exists a large set of spectral densities ϕ such that $\|\phi_+ - \mathrm{S}_N(\phi)\|_{\mathscr{D}}$ does not converges to zero as $N \to \infty$

Spectral factorization is closely related to the so-called *inner–outer factorization* which plays a central role in the theory of Hardy spaces [17,29]. It states that any function in a Hardy space $\mathcal{H}^p(\mathbb{D})$, $1 \le p \le \infty$, of the unit disk can be factorized as $f = f_o f_i$ with a so-called *inner function* f_i and an *outer function* f_o. For the outer function f_o, there exists a closed form expression, almost identical to Eq. (2) for the spectral factor. This similarity will allow us to derive in the second part of the paper a result showing that there exists no sampling-based approximation procedure which is able determine the outer function for all continuous, positive functions of finite Dirichlet energy.

Since Shannon's seminal paper [38] on sampling of bandlimited functions, sampling theory plays a fundamental role in applied mathematics and signal processing [16,25,32,35,41] with

applications in image processing [40], graph theory [15], estimation and prediction of stochastic processes [31], to mention only a few fields. Despite its far-reaching importance, this paper is going to show that there exist fundamental limits for the applicability of sampling-based methods. Specifically, we will show that the calculation of the spectral factor and the outer function from samples of the spectral density is impossible, in general.

The paper is structured as follows. Section 2 gives a very short overview of the used notation and signal spaces. In Section 3, sampling-based methods for the spectral factorization are investigated. To this end, basic properties of the spectral factorization mapping are recaptured and an appropriate set of spectral densities, $\mathscr{D}_+$, is defined. Afterward, an axiomatic characterization of sampling-based approximation methods is introduced and it is shown that in this class of approximation methods there exists no approximation method which is able to calculate the spectral factor for all densities in the class $\mathscr{D}_+$. Section 4 translates the results of Section 3 to similar statements for the calculation of the outer function. The paper closes with a short summary and a discussion of further mathematical and practical properties of the spectral factorization in Section 5.

2. Notation and function spaces

The Banach space of all functions continuous on $\mathbb{T}$ with norm $\|f\|_\infty = \max_{\zeta\in\mathbb{T}} |f(\zeta)|$ is denoted by $\mathcal{C}(\mathbb{T})$, and $L^p(\mathbb{T})$ with $1 \leq p \leq \infty$ stands for the usual Banach space of p-integrable functions on $\mathbb{T}$. The *Fourier coefficients* of any $f \in L^1(\mathbb{T})$ are given by

$$c_n(f) = \tfrac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}\theta})\, \mathrm{e}^{\mathrm{i}n\theta}\, \mathrm{d}\theta\,, \qquad n \in \mathbb{Z}\,,$$

and it is well known that for $1 < p < \infty$ every $f \in L^p(\mathbb{T})$ can be recovered from its Fourier coefficients by means of the *Fourier series*

$$f(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n\in\mathbb{Z}} c_n(f)\, \mathrm{e}^{\mathrm{i}n\theta}\,, \qquad \theta \in [-\pi, \pi)\,,$$

which converges in the norm of $L^p(\mathbb{T})$ and uniformly on $\mathbb{T}$. With every $f \in L^p(\mathbb{T})$, one associates its *conjugate function* $\widetilde{f} \in L^p(\mathbb{T})$ given by

$$\widetilde{f}(\mathrm{e}^{\mathrm{i}\theta}) = (\mathrm{H}f)(\mathrm{e}^{\mathrm{i}\theta}) = -\mathrm{i} \sum_{n\in\mathbb{Z}} \operatorname{sgn}(n)\, c_n(f)\, \mathrm{e}^{\mathrm{i}n\theta}\,, \qquad \theta \in [-\pi, \pi)\,, \tag{4}$$

with the usual signum function given by $\operatorname{sgn}(n) = n/|n|$ for $n \neq 0$ and $\operatorname{sgn}(0) = 0$. The mapping $\mathrm{H} : f \mapsto \widetilde{f}$ is said to be the *Hilbert transform* and we note that it is also common to call $\widetilde{f}$ the Hilbert transform of f.

We will also need the *Hardy spaces* $\mathcal{H}^p(\mathbb{D})$ on the unit disk [17,29]. For every $1 \leq p < \infty$, the Banach space $\mathcal{H}^p(\mathbb{D})$ is the set of all functions f analytic in $\mathbb{D}$ and which satisfy

$$\|f\|_p \coloneqq \left(\sup_{0<r<1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| f(r\mathrm{e}^{\mathrm{i}\theta}) \right|^p \mathrm{d}\theta \right)^{1/p} < \infty\,,$$

and $H^\infty(\mathbb{D})$ denotes the set of all bounded analytic functions on $\mathbb{D}$ with norm $\|f\|_\infty = \sup_{z\in\mathbb{D}} |f(z)|$. The *disk algebra* $\mathcal{A}(\mathbb{D}) = \mathcal{H}^\infty(\mathbb{D}) \cap \mathcal{C}(\mathbb{T})$ contains all functions analytic in $\mathbb{D}$ and continuous in the closed unit disk $\overline{\mathbb{D}}$, equipped with the norm $\|f\|_{\mathcal{A}(\mathbb{D})} = \max_{z\in\overline{\mathbb{D}}} |f(z)|$. For every $f \in \mathcal{H}^p(\mathbb{D})$ the radial limit $\lim_{r\to\infty} f(r\mathrm{e}^{\mathrm{i}\theta})$ exists almost everywhere on $\mathbb{T}$ and defines so a function in $L^p(\mathbb{T})$. Therefore $\mathcal{H}^p(\mathbb{D})$ can be identified with the closed subspace $\{f \in L^p(\mathbb{T}) : c_n(f) = 0 \text{ for all } n < 0\}$ of $L^p(\mathbb{T})$. We also note that a function f in $\mathbb{D}$ is said to be an *outer function* if there exists a $\varphi \in L^1(\mathbb{T})$ such that

$$f(z) = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log \phi(\mathrm{e}^{\mathrm{i}\tau}) \frac{\mathrm{e}^{\mathrm{i}\tau} + z}{\mathrm{e}^{\mathrm{i}\tau} - z}\, \mathrm{d}\tau \right) \qquad \text{for all } z \in \mathbb{D}\,.$$

Note that this definition implies immediately that any outer function f is analytic in $\mathbb{D}$ with $f(z) \neq 0$ for all $z \in \mathbb{D}$ and it shows that every spectral factor (2) is an outer function.

In signal processing, the L^2-norm is usually interpreted as the *energy* of a function. Nevertheless, in physics a stronger notion of energy is often applied which will also be adapted in this paper. Namely, we require that all of our spectral densities have finite *Dirichlet energy* which is equal to the square of the seminorm

$$\|f\|_{\mathrm{E}} = \left(\sum_{n\in\mathbb{Z}} |n|\ |c_n(f)|^2\right)^{1/2} . \tag{5}$$

Based on this seminorm, the Dirichlet space $\mathcal{D}_{\mathcal{C}}$ is defined to be the set of all continuous functions on $\mathbb{T}$ with finite Dirichlet energy, i.e.

$$\mathcal{D}_{\mathcal{C}} = \left\{ f \in \mathcal{C}(\mathbb{T}) \ :\ \|f\|_{\mathrm{E}} < +\infty \right\} \quad \text{with norm} \quad \|f\|_{\mathcal{D}_{\mathcal{C}}} = \max\left(\|f\|_{\infty}, \|f\|_{\mathrm{E}}\right) .$$

The so defined $\mathcal{D}_{\mathcal{C}}$ is a Banach space and we refer to papers like [10,36] for a detailed discussion on the relation between $\|f\|_{\mathrm{E}}$ and the energy of certain physical quantities. We would also like to point out that the norm in $\mathcal{D}_{\mathcal{C}}$ is equal to the norm in the Sobolev space $H^{1/2}(\mathbb{T}) = W^{1/2,2}(\mathbb{T})$.

3. Spectral factorization — Divergence of sampling-based approximation methods

This section considers the problem of approximating the spectral factor ϕ_+ of a spectral density ϕ based on the given samples of ϕ. The first subsection provides a very short review of basic properties of the spectral factorization and the second subsection defines precisely the set $\mathscr{D}_+$ of spectral densities on which this operation is considered. Afterwards, Section 3.3 gives an axiomatic description of the approximation algorithms which are considered. The last subsection will show that for every arbitrary sampling-based approximation algorithm which falls in our axiomatic framework, there always exist spectral densities for which the algorithm does not converge to the desired spectral factor.

3.1. Definition and basic properties

Let $\phi \in \mathcal{C}(\mathbb{T})$ be a real function so that $\phi(\zeta) \geq c_0 > 0$ for all $\zeta \in \mathbb{T}$. Then its *spectral factor* ϕ_+ is given by (2). Let us define the function u by $u(\zeta) = \log \phi(\zeta)$ for all $\zeta \in \mathbb{T}$ and write $z = r\mathrm{e}^{\mathrm{i}\theta} \in \mathbb{D}$ with $r \in [0, 1)$ and $\theta \in [-\pi, \pi)$. Then, by separating the kernel in (2) into is real and imaginary part, (2) can be rewritten as

$$\phi_+(r\mathrm{e}^{\mathrm{i}\theta}) = \exp\left(\tfrac{1}{2}\,(\mathrm{P}_r u)\,(\mathrm{e}^{\mathrm{i}\theta})\right) \cdot \exp\left(\mathrm{i}\tfrac{1}{2}\,(\mathrm{Q}_r u)\,(\mathrm{e}^{\mathrm{i}\theta})\right) , \tag{6}$$

with the usual *Poisson-* and *conjugate Poisson integral* given by

$$(\mathrm{P}_r u)\,(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} u(\mathrm{e}^{\mathrm{i}\tau})\,\frac{1-r^2}{1-2r\cos(\theta-\tau)+r^2}\,\mathrm{d}\tau$$

$$\text{and} \qquad (\mathrm{Q}_r u)\,(\mathrm{e}^{\mathrm{i}\theta}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} u(\mathrm{e}^{\mathrm{i}\tau})\,\frac{2\,r\sin(\theta-\tau)}{1-2r\cos(\theta-\tau)+r^2}\,\mathrm{d}\tau ,$$

respectively. Moreover, we notice that $\big(\mathrm{Q}_r u\big)(\mathrm{e}^{\mathrm{i}\theta}) = \big(\mathrm{P}_r \widetilde{u}\big)(\mathrm{e}^{\mathrm{i}\theta})$ for $\theta \in [-\pi, \pi)$, wherein $\widetilde{u}$ is the conjugate function (4) of u. Moreover, for continuous functions $u \in \mathcal{C}(\mathbb{T})$, it is well known [17] that

$$\lim_{r\to 1} (\mathrm{P}_r u)\,(\mathrm{e}^{\mathrm{i}\theta}) = u(\mathrm{e}^{\mathrm{i}\theta}) \qquad \text{for all } \theta \in [-\pi, \pi) . \tag{7}$$

So if $u \in \mathcal{C}(\mathbb{T})$ has the property that $\widetilde{u} \in \mathcal{C}(\mathbb{T})$ then (6) and (7) show that

$$\phi_+(\zeta) = \exp\left(\tfrac{1}{2}\left[u(\zeta) + \mathrm{i}\,\widetilde{u}(\zeta)\right]\right) , \qquad \zeta \in \mathbb{T} \tag{8}$$

is a continuous function on $\mathbb{T}$ and in particular that $\phi_+ \in \mathcal{A}(\mathbb{D})$.

We emphasize again that the spectral factorization mapping (2) is a highly non-linear operator showing a very complicated behavior on several Banach spaces. Since this paper studies the possibility of calculating the spectral factor on a digital computer, we consider only signal spaces for which the spectral factorization mapping itself shows a very good behavior. Then it is shown that even though we consider only nice spectral densities, there still exists no sampling-based algorithm which is able to determine the spectral factor for all these nice spectral densities.

3.2. Sets of spectral densities

We consider the spectral factorization for densities ϕ in the Dirichlet space $\mathcal{D}_{\mathcal{C}}$ of continuous functions with finite Dirichlet energy. Additionally, in order that the spectral factorization is well defined, one has to require that $\log \phi \in L^1(\mathbb{T})$. Here we make the even stronger assumption that ϕ is strictly positive on $\mathbb{T}$, i.e. we consider the set

$$\mathscr{D} := \left\{\phi \in \mathcal{D}_{\mathcal{C}} \ : \ \min\nolimits_{\zeta\in\mathbb{T}} \phi(\zeta) \geq c_0 > 0\right\}$$

with a certain positive constant c_0. However, there is a disadvantage in working with $\mathscr{D}$, namely $\mathscr{D}$ is not a linear space. For this reason, we consider the set

$$\log(\mathscr{D}) := \{u = \log(\phi) \ : \ \phi \in \mathscr{D}\} , \tag{9}$$

i.e. the set of all functions u defined by $u(\zeta) = \log \phi(\zeta)$, $\zeta \in \mathbb{T}$, for some $\phi \in \mathscr{D}$. The following lemma (whose proof can be found in Appendix A.2) shows that (9) is a Banach space, namely the space $\mathcal{D}_{\mathcal{C}}$.

Lemma 3.1. *Let $\phi \in \mathscr{D}$ be arbitrary and let $u \in \log(\mathscr{D})$ be defined by $u(\zeta) = \log \phi(\zeta)$ for all $\zeta \in \mathbb{T}$. Then $u \in \mathcal{D}_{\mathcal{C}}$.*

Conversely, let $u \in \mathcal{D}_{\mathcal{C}}$ be arbitrary and let $\phi \in \exp(\mathcal{D}_{\mathcal{C}})$ be defined by $\phi(\zeta) = \exp u(\zeta)$ for all $\zeta \in \mathbb{T}$. Then $\phi \in \mathscr{D}$.

So there is a one-to-one correspondence between the set $\mathscr{D}$ of positive, continuous spectral densities of finite Dirichlet energy and the Banach space $\mathcal{D}_{\mathcal{C}}$ of continuous functions of finite Dirichlet energy, i.e.

$$\mathscr{D} = \exp(\mathcal{D}_{\mathcal{C}}) \qquad \text{and} \qquad \mathcal{D}_{\mathcal{C}} = \log(\mathscr{D}) . \tag{10}$$

Nevertheless, we will restrict the set of spectral densities even further by considering only densities ϕ which have the additional property that their spectral factor ϕ_+ is again a continuous function, i.e. we consider spectral densities in the set

$$\mathscr{D}_+ := \{\phi \in \mathscr{D} \ : \ \phi_+ \in \mathcal{A}(\mathbb{D})\} . \tag{11}$$

As discussed in Section 3.1, the additional requirement $\phi_+ \in \mathcal{A}(\mathbb{D})$ is satisfied if and only if the corresponding function $u = \log \phi$ has the property that $\widetilde{u} \in \mathcal{C}(\mathbb{T})$. So it follows from (10) that

$$\mathscr{D}_+ = \exp(\mathcal{B}_0) \qquad \text{and} \qquad \mathcal{B}_0 = \log(\mathscr{D}_+)$$

wherein $\mathcal{B}_0$ is the separable Banach space

$$\mathcal{B}_0 = \left\{u \in \mathcal{D}_{\mathcal{C}} \ : \ \widetilde{u} \in \mathcal{C}(\mathbb{T})\right\} \qquad \text{with norm} \qquad \|u\|_{\mathcal{B}_0} = \max\left(\|u\|_\infty \ , \ \|\widetilde{u}\|_\infty \ , \ \|u\|_{\mathrm{E}}\right).$$

3.3. Approximation sequences

Next, we give a precise definition of the sampling-based algorithms for the calculation of the spectral factor investigated in this paper. These algorithms are characterized by four simple axioms given in the following Definition.

Definition 3.2. Let $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ be a sequence of operators $\mathrm{S}_N : \mathscr{D}_+ \to \mathcal{A}(\mathbb{D})$. We say that $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ is a *sampling-based approximation procedure for the spectral factorization* if it has the following properties:

(A) To every $N \in \mathbb{N}$ there exists a finite subset $\mathcal{T}_N \subset \mathbb{T}$ such that for arbitrary $\phi_1, \phi_2 \in \mathscr{D}_+$

$$\begin{aligned} & \phi_1(\zeta_n) = \phi_2(\zeta_n) && \text{for all } \zeta_n \in \mathcal{T}_N \\ \text{implies} \quad & \big(\mathrm{S}_N\phi_1\big)(z) = \big(\mathrm{S}_N\phi_2\big)(z) && \text{for all } z \in \overline{\mathbb{D}}\,. \end{aligned}$$

(B) There exists a dense subset $\mathcal{M} \subset \mathcal{B}_0$ such that

$$\lim_{N\to\infty} \|\phi_+ - \mathrm{S}_N(\phi)\|_{\mathcal{A}(\mathbb{D})} = 0 \qquad \text{for all } \phi \in \exp(\mathcal{M})\,,$$

i.e. for all $\phi = \exp(u)$ with $u \in \mathcal{M}$.

(C) $\mathrm{S}_N(\phi)$ is an outer function for every $N \in \mathbb{N}$ and for each $\phi \in \mathscr{D}_+$.

(D) Let $\widetilde{\mathrm{S}}_N : \mathcal{B}_0 \to \mathcal{A}(\mathbb{D})$ be defined by $\widetilde{\mathrm{S}}_N(u) = \mathrm{S}_N(\exp u)$ for every $u \in \mathcal{B}_0$. Then $\widetilde{\mathrm{S}}_N$ is a continuous mapping for every $N \in \mathbb{N}$, i.e. if $\{u_n\}_{n\in\mathbb{N}} \subset \mathcal{B}_0$ is a convergent sequence with limit $u \in \mathcal{B}_0$ then

$$\lim_{n\to\infty} \left\| \mathrm{S}_N\big[\exp(u)\big] - \mathrm{S}_N\big[\exp(u_n)\big] \right\|_{\mathcal{A}(\mathbb{D})} = 0 \qquad \text{for all } N \in \mathbb{N}\,.$$

Remark 3.1. We emphasize that Definition 3.2 makes no assumption on the linearity of the operators S_N. Some or even all of them might be non-linear.

Property (A) is the key assumption. It requires that each approximation $\mathrm{S}_N(\phi)$ of ϕ_+ is calculated based on finitely many samples of ϕ. Property (B) requires that the approximation procedure converges at least for a certain subset of $\mathscr{D}_+$. Since we look for approximation procedures which converge for all $\phi \in \mathscr{D}_+$, this property is effectively no restriction but a necessary requirement. Property (C) demands that $\mathrm{S}_N(\phi)$ is an outer function. Since ϕ_+ is an outer function, this property is a very natural assumption requiring that the outer function ϕ_+ is approximated by outer functions $\mathrm{S}_N(\phi)$. Finally, Property (D) makes an assumption on the continuity of the used operators. Also this is a natural requirement for any robust approximation procedure.

Note that Property (D) of Definition 3.2 is necessary because we allow for non-linear approximation operators S_N. Indeed, assume S_N would be linear. Since S_N is also sampling-based, as required by Property (A), the operators S_N would have necessarily the form $(\mathrm{S}_N\phi)(z) = \sum_n \phi(\zeta_n)\, p_n(z)$ with certain functions $p_n \in \mathcal{A}(\mathbb{D})$. Then for every fixed $N \in \mathbb{N}$, S_N would be bounded and so also continuous. However, since the spectral factorization mapping S itself is non-linear, it seems not reasonable to allow only for linear approximation operators S_N. Consequently, Property (D) is an important assumption to assure that every S_N is stable

with respect to errors in the given spectral density. Moreover, we refer to [9] where a similar axiomatic approach was applied to characterize sampling-based approximations for the Hilbert transform. Since the Hilbert transform is linear, it was assumed there that each approximation operator S_N is linear and consequently no Axiom (D) was required in this work.

From a practical point of view, the four properties of Definition 3.2 are effectively no restrictions. Basically every algorithm for the spectral factorization which should be implemented on a digital computer has to satisfy these properties. For illustration, the following example gives a simple practical (linear) procedure for calculating the spectral factor and which satisfies the four properties of Definition 3.2.

Example 1. Let $\phi \in \mathscr{D}_+$ be an arbitrary spectral density and assume the samples $\{\phi(\zeta_n) : \zeta_n \in \mathcal{T}_N\}$ of ϕ are given. Then consider the following two step procedure for a linear sampling-based spectral factorization algorithm.

1. First, one determines an approximation ϕ_N of ϕ by interpolating the given samples by a spline of certain degree. For this spline approximation holds $\lim_{N\to\infty} \|\phi - \phi_N\|_\infty = 0$, and since ϕ is strictly positive also ϕ_N is strictly positive [42].
2. Because of the positivity of ϕ_N, one can determine the spectral factor $(\phi_N)_+$ of the spline ϕ_N using standard algorithms for polynomial spectral factorization [37].

This way, one obtains a (linear) algorithm which determines an approximation $(\phi_N)_+$ of the spectral factor ϕ_+, and it is easily seen that this approximation procedure satisfies the four properties of Definition 3.2. Moreover, applying other interpolation methods in the first step of the described procedure (e.g. trigonometric interpolation) yields many other sampling-based approximation methods for the spectral factorization.

3.4. Divergence of sampling-based spectral factorization algorithms

Is it possible to find a sampling-based method $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ for calculating the spectral factor which satisfies the properties of Definition 3.2 and which converges to the spectral factor ϕ_+ for all $\phi \in \mathscr{D}_+$? The following main theorem shows that the answer to this question is negative.

Theorem 3.3. *Let $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ be an arbitrary sampling-based approximation method for the spectral factorization as defined in Definition 3.2. Then the set $\mathcal{R}_0 \subset \mathcal{B}_0$ of all $u \in \mathcal{B}_0$ such that for $\phi = \exp(u)$*

$$\limsup_{N\to\infty} \|\phi_+ - \mathrm{S}_N(\phi)\|_{\mathcal{A}(\mathbb{D})} > 0 \tag{12}$$

holds, is a residual set in $\mathcal{B}_0$.

Remark 3.2. Note that (12) does not imply that the approximation error $\|\phi_+ - \mathrm{S}_N(\phi)\|_{\mathcal{A}(\mathbb{D})}$ diverges as $N \to \infty$. In fact, the proof of Theorem 3.3 will show that actually only the imaginary part of $\phi_+ - \mathrm{S}_N(\phi)$ diverges as $N \to \infty$ implying that the error $\|\phi_+ - \mathrm{S}_N(\phi)\|_{\mathcal{A}(\mathbb{D})}$ oscillates but remains bounded as $N \to \infty$.

Theorem 3.3 shows that for any sampling-based approximation method $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ of the spectral factorization there always exists a density $\phi \in \mathscr{D}_+$ such that $\mathrm{S}_N(\phi)$ does not converge to the desired spectral factor ϕ_+. Moreover, it shows that the set of spectral densities for which $\mathrm{S}_N(\phi)$ does not converge is large in the following sense: Let $\mathcal{D}_0$ be the set of all $\phi \in \mathscr{D}_+$ satisfying (12), then $\log(\mathcal{D}_0)$ is a residual set in $\mathcal{B}_0$.

It is also worth to notice that the theorem considers the convergence only with respect to the uniform norm. Since $\phi \in \mathscr{D}_+$ is a continuous density of finite Dirichlet energy, it would be natural to require convergence with respect to the norm $\|\cdot\| = \max\left(\|\cdot\|_{\mathcal{A}(\mathbb{D})}, \|\cdot\|_{\mathrm{E}}\right)$. However, Theorem 3.3 shows that all sampling-based algorithms diverge already in the much weaker norm $\|\cdot\|_{\mathcal{A}(\mathbb{D})}$. Similarly, if one considers $\{\mathrm{S}_N\}$ as operators $\mathcal{C}(\mathbb{T}) \to \mathcal{A}(\mathbb{D})$ then Theorem 3.3 shows that any sampling-based approximation method diverges (in the norm of $\mathcal{A}(\mathbb{D})$) even on the subset of all positive densities $\phi \in \mathcal{C}(\mathbb{T})$ with finite Dirichlet energy. So there exists a fortiori no sampling-based method which converges for all positive continuous spectral densities.

The proof of Theorem 3.3 is based on a similar result for sampling-based approximation methods for the Hilbert transform H. For ease of reference, we restate this result from [11] here in the form as needed for the subsequent proof of Theorem 3.3.

Theorem 3.4. *Let $\{\mathrm{H}_N\}_{N\in\mathbb{N}}$ be a sequence of continuous operators $\mathrm{H}_N : \mathcal{B}_0 \to \mathcal{C}(\mathbb{T})$ which satisfies the following two properties*

(I) For every $N \in \mathbb{N}$ there exists finite subset $\mathcal{T}_N \subset \mathbb{T}$ such that for all $u_1, u_2 \in \mathcal{B}_0$

$$\begin{aligned} & u_1(\zeta_n) = u_2(\zeta_n) && \text{for all } \zeta_n \in \mathcal{T}_N \\ \text{implies} \quad & (\mathrm{H}_N u_1)(\zeta) = (\mathrm{H} u_2)(\zeta) && \text{for all } \zeta \in \mathbb{T}\,. \end{aligned}$$

(II) There exists a subset $\mathcal{M} \in \mathcal{B}_0$ such that $\lim_{N\to\infty} \|\mathrm{H}u - \mathrm{H}_N(u)\|_\infty = 0$ for all $u \in \mathcal{M}$.

Then the set $\left\{u \in \mathcal{B}_0 \;:\; \limsup_{N\to\infty} \|\mathrm{H}u - \mathrm{H}_N(u)\|_\infty > 0\right\}$ is a residual set in $\mathcal{B}_0$.

Proof (*Theorem* 3.3)**.** Let $u \in \mathcal{B}_0$ be arbitrary with the corresponding density $\phi = \exp(u) \in \mathscr{D}_+$, and assume that

$$\lim_{N\to\infty} \|\phi_+ - \mathrm{S}_N(\phi)\|_{\mathcal{A}(\mathbb{D})} = 0\,. \tag{13}$$

Since $u \in \mathcal{B}_0$, we know that $u, \widetilde{u} \in \mathcal{C}(\mathbb{T})$ and so (8) shows that $U := \log \phi_+ \in \mathcal{A}(\mathbb{D})$. By the definition of the operators S_N, it is clear that $\mathrm{S}_N(\phi) \in \mathcal{A}(\mathbb{D})$, and $\phi \in \mathscr{D}_+$ implies that there exists and $c_0 > 0$ such that $\min_{\zeta\in\mathbb{T}} \phi(\zeta) = c_0$ and so $\min_{\zeta\in\mathbb{T}} u(\zeta) = \log c_0$. Then the minimum principle for analytic functions in $\mathbb{D}$ and (8) show that

$$|\phi_+(z)| \geq \min_{\zeta\in\mathbb{T}} \exp\left(\tfrac{1}{2}u(\zeta)\right) = \exp\Big(\tfrac{1}{2}\min_{\zeta\in\mathbb{T}} u(\zeta)\Big) = \sqrt{c_0} \qquad \text{for all } z \in \overline{\mathbb{D}}\,.$$

So if we set $g = \phi_+$ and $g_N = \mathrm{S}_N(\phi)$, these functions satisfy the conditions of Lemma A.3 in the Appendix and we get from (13)

$$\lim_{N\to\infty} \min_{|z|\leq 1} |U(z) - \log \mathrm{S}_N\phi(z)| = 0\,.$$

Similarly, applying the maximum principle for analytic functions yields

$$\lim_{N\to\infty} \max_{|z|=1} |U(z) - \log \mathrm{S}_N\phi(z)| = 0\,. \tag{14}$$

According to (8), U can be written on $\mathbb{T}$ as $U(\zeta) = \log\phi_+(\zeta) = \frac{1}{2}\big[u(\zeta) + \mathrm{i}\,\widetilde{u}(\zeta)\big]$ for all $\zeta \in \mathbb{T}$. In a similar way, we separate $\log \mathrm{S}_N(\phi)$ into its real- and imaginary part and write

$$\log\big(\mathrm{S}_N\phi\big)(\zeta) = \log\big(\mathrm{S}_N\big[\exp(u)\big]\big)(\zeta) = (\mathrm{A}_N u)(\zeta) + \mathrm{i}\,(\mathrm{B}_N u)(\zeta)\,, \qquad \zeta \in \mathbb{T}\,.$$

Because S_N is a mapping $\mathcal{B}_0 \to \mathcal{A}(\mathbb{D})$, the so defined operators A_N and B_N are mappings $\mathcal{B}_0 \to \mathcal{C}(\mathbb{T})$. We verify next that the sequence $\{\mathrm{B}_N\}_{N\in\mathbb{N}}$ satisfies the conditions of Theorem 3.4.

According to Property (A), the operator S_N uses only the values $\{\phi(\zeta_n) : \zeta_n \in \mathcal{T}_N\}$ of the spectral density ϕ to calculate the approximation $\mathrm{S}_N(\phi)$. Consequently, the operators A_N and B_N are concentrated on the values $\{u(\zeta_n) = \log \phi(\zeta_n) : \zeta_n \in \mathcal{T}_N\}$ showing that $\{\mathrm{B}_N\}_{N\in\mathbb{N}}$ satisfies Condition (I) of Theorem 3.4.

Since $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ satisfies Property (B), (13) is satisfied for all $\phi \in \exp(\mathcal{M})$ where $\mathcal{M}$ is a dense subset of $\mathcal{B}_0$. Consequently, (14) implies for the corresponding imaginary part

$$\lim_{N\to\infty} \left\| \tfrac{1}{2}\widetilde{u} - \mathrm{B}_N(u) \right\|_\infty = 0 \qquad \text{for all } u \in \mathcal{M} \tag{15}$$

showing that $\{\mathrm{B}_N\}_{N\in\mathbb{N}}$ satisfies Condition (II) of Theorem 3.4.

Assume $\{u_n\}_{n\in\mathbb{N}} \subset \mathcal{B}_0$ is a convergent sequence such that $\lim_{n\to\infty} \|u - u_n\|_{\mathcal{B}_0} = 0$ for some $u \in \mathcal{B}_0$. Then the continuity property (D) of the sequence $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ implies

$$\lim_{n\to\infty} \|\mathrm{B}_N(u) - \mathrm{B}_N(u_n)\|_\infty = 0 \qquad \text{for every } N \in \mathbb{N}\,.$$

Showing that $\mathrm{B}_N : \mathcal{B}_0 \to \mathcal{C}(\mathbb{T})$ is continuous for every $N \in \mathbb{N}$.

Thus $\{\mathrm{B}_N\}_{N\in\mathbb{N}}$ satisfies the conditions of Theorem 3.4 and so this theorem implies that the set of all $u \in \mathcal{B}_0$ satisfying (15) can only be a set of first category. Consequently, also the set of all $u \in \mathcal{B}_0$ for which $\phi = \exp(u)$ satisfies (13) can only be of first category. So the statement of the theorem is proved. ■

4. Sampling-based approximation methods for calculating the outer function

In the theory of Hardy spaces, the so called *inner–outer factorization* plays a prominent role. Since this factorization is very closely related to the spectral factorization, we can translate Theorem 3.3 into a similar divergence result for the calculation of the outer function.

4.1. Outer functions and minimal phase systems

For $1 < p \leq \infty$, let $f \in \mathcal{H}^p(\mathbb{D})$ be arbitrary. Then there exists a so-called *inner function* $f_i \in \mathcal{H}^\infty(\mathbb{D})$ with $|f(\zeta)| = 1$ for almost every $\zeta \in \mathbb{T}$ and an *outer function* $f_o \in \mathcal{H}^p(\mathbb{D})$ such that $f(z) = f_i(z) f_o(z)$ for $z \in \mathbb{D}$. Therein, the outer function is given explicitly by

$$f_o(z) = \exp\left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \log\left|f(\mathrm{e}^{\mathrm{i}\tau})\right| \frac{\mathrm{e}^{\mathrm{i}\tau} + z}{\mathrm{e}^{\mathrm{i}\tau} - z}\, \mathrm{d}\tau \right)\,, \qquad z \in \mathbb{D}\,, \tag{16}$$

and it satisfies $f_o(\zeta) = |f(\zeta)|$ for almost all $\zeta \in \mathbb{T}$. The inner–outer factorization of f is unique up to a unitary factor and (16) is well defined as long as f satisfies the Szegő condition $\log|f| \in L^1(\mathbb{T})$ [17,29].

This factorization plays also a considerable important role in signal- and system theory [20, 33]. There it is applied for decomposing a causal linear system into a *minimum phase* part (the outer function) and an *allpass* part (the inner function). As an application, we mention the following standard filter design problem. Assume we want to design the transfer function $f(\mathrm{e}^{\mathrm{i}\theta})$ of a linear digital filter with a predefined amplitude response $F(\mathrm{e}^{\mathrm{i}\theta}) = \left|f(\mathrm{e}^{\mathrm{i}\theta})\right|$. If the filter should additionally be causal, then the solution to this problem can be written as $f(z) = f_i(z) f_o(z)$ with the outer function f_o, given in (16), completely determined by the predefined amplitude response F. The allpass part (the inner function) can then be chosen arbitrary to influence other properties of the filter but without changing the prescribed amplitude response of the filter. The outer-function of f is completely determined by the amplitude

$F(\zeta) = |f(\zeta)|$, $\zeta \in \mathbb{T}$ and we will write

$$F_o(z) = \exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} \log F(\mathrm{e}^{\mathrm{i}\tau})\,\frac{\mathrm{e}^{\mathrm{i}\tau}+z}{\mathrm{e}^{\mathrm{i}\tau}-z}\,\mathrm{d}\tau\right), \qquad z \in \mathbb{D}, \tag{17}$$

for the outer function associated with a given amplitude response F.

4.2. Divergence of sampling-based algorithms

Working on digital computers, the calculation of the outer function F_o has to be based on finitely many samples $\{F(\zeta_n)\}_{n=1}^N$ of the given amplitude response F, taken on a sampling set $\mathcal{T}_N = \{\zeta_n\}_{n=1}^N \subset \mathbb{T}$. Then we look for algorithms T_N which determine an approximation $F_N = \mathrm{T}_N(F)$ of F_o, based on the samples on $\mathcal{T}_N$ and such that F_N converges to F_o as the number N of available sampling points tends to infinity. As for the spectral factorization, we ask whether there exist such sampling-based approximation methods $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ which converge for all positive continuous amplitude responses F of finite energy.

The outer function (16), respective (17), is well defined if $\log|f| = \log F \in L^1(\mathbb{T})$, and this condition is satisfied for all $f \in \mathcal{H}^p(\mathbb{D})$ with $1 \le p \le \infty$ and in particular for all $f \in \mathcal{A}(\mathbb{D})$ (see, e.g., [17]). Nevertheless, as in the case of the spectral factorization, we make an even stronger assumption on the predefined amplitude response F, namely that F is a spectral density in the set $\mathscr{D}_+$, defined in (11). With the given amplitude response $F \in \mathscr{D}_+$, we associate the corresponding density $\phi(\zeta) = F(\zeta)^2 = |f(\zeta)|^2$ and it is not hard to see that $\phi \in \mathscr{D}_+$ if and only if $F \in \mathscr{D}_+$. Then it follows from (2) and (16) that the outer function F_o, which corresponds to the given amplitude response F, is equal to the spectral factor of the associated density ϕ, i.e.

$$F_o(z) = \phi_+(z) = (\mathrm{S}\phi)(z) = (\mathrm{S}F)(z) = \left(\mathrm{S}\left[|f|^2\right]\right)(z), \qquad z \in \mathbb{D}.$$

Thus the outer function can be calculated using any spectral factorization algorithm with input $\phi = F^2 = |f|^2$. So we can take over the axiomatic framework for sampling-based approximation methods from Definition 3.2 for approximation methods for the outer function.

Definition 4.1. A sequence $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ of operators $\mathrm{T}_N : \mathscr{D}_+ \to \mathcal{A}(\mathbb{D})$ is said to be a *sampling-based approximation procedure for the outer function* if every T_N can be written as

$$\mathrm{T}_N(F) = \mathrm{S}_N(F^2), \qquad N \in \mathbb{N}$$

with a sampling-based approximation procedure $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ as given in Definition 3.2.

Then, it comes as little surprise that Theorem 3.3 holds almost literally for approximation methods for the outer function.

Theorem 4.2. *Let $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ be a sampling-based approximation method for the outer function according Definition 4.1. Then there exists a residual set $\mathcal{R}_0 \subset \mathcal{B}_0$ such that for every $F = \exp(u)$ with $u \in \mathcal{R}_0$, one has*

$$\limsup_{N\to\infty} \|F_o - \mathrm{T}_N(F)\|_{\mathcal{A}(\mathbb{D})} > 0. \tag{18}$$

Thus for every sampling-based method $\{\mathrm{T}_N\}_{N\in\mathbb{N}}$ for calculating the outer function there always exist amplitude responses $F \in \mathscr{D}_+$ such that $\mathrm{T}_N(F)$ does not converge to the desired

F_o. Moreover, the corresponding divergence set $\mathcal{D}_0 = \{F \in \mathscr{D}_+ : F \text{ satisfies (18)}\}$ is large in the sense that $\log(\mathcal{D}_0)$ is a residual set in $\mathcal{B}_0$. So there exists no sampling-based algorithm which is able to calculate the outer function for all functions in the disk algebra $\mathcal{A}(\mathbb{D})$. This statement is even true for the subset of all those $f \in \mathcal{A}(\mathbb{D})$ for which the amplitude of its boundary function $F(\zeta) = |f(\zeta)|$, $\zeta \in \mathbb{T}$, belongs to $\mathscr{D}_+$, i.e. for which F is continuous, strictly positive, has finite energy, and for which f_o is known to be continuous in $\overline{\mathbb{D}}$. Even for these very nice Hardy space functions, no sampling-based algorithm for calculating f_o does exist.

It might be worth to notice that our assumptions on the boundary function $F(\zeta) = |f(\zeta)|$, $\zeta \in \mathbb{T}$, imply also certain restrictions on the inner function f_i of f. Recall that any inner function can be written as the product of a *Blaschke product* and of a *singular function* [17, Chap. II.6]. Since we assume that $|f(\zeta)| \geq c_0 > 0$ for all $\zeta \in \mathbb{T}$, and since f is analytic in $\mathbb{D}$, it follows that f can only have finitely many zeros in $\mathbb{D}$. So the Blaschke product associated with f_i is finite. Moreover, since F is assumed to be continuous, also f and f_i are continuous on $\mathbb{T}$. However, since a singular function is not continuous, we see that the singular factor of f_i has to be equal to zero. So overall, the inner function is just a finite Blaschke product under our assumptions on the boundary function of f.

5. Summary and discussion

Theorem 3.3 shows that there is no sampling-based approximation method $\{\mathrm{S}_N\}_{N\in\mathbb{N}}$ which is able to calculate the spectral factor ϕ_+ for all spectral densities ϕ in the set $\mathscr{D}_+$ of positive, continuous densities of finite Dirichlet energy with a continuous spectral factor. Using the one-to-one correspondence (10) between the set $\mathscr{D}_+$ and the Banach space $\mathcal{B}_0$, it even follows that the divergence set $\mathcal{D}_0 = \left\{\phi \in \mathscr{D}_+ : \lim_{N\to\infty} \|\phi_+ - \mathrm{S}_N(\phi)\|_{\mathcal{A}(\mathbb{D})} > 0\right\}$ is large in the sense that $\log(\mathcal{D}_0)$ is a residual set in $\mathcal{B}_0$.

In the theory of Hardy spaces of the unit disk as well as in filter design problems in engineering applications, the outer function plays a central role. It is given by a function on the unit circle in a very similar way as the spectral factor is given be the spectral density. Consequently, our main Theorem 3.3 could readily translated into a corresponding result for calculating the outer function using sampling-based approximation algorithms. It turns out that also for this problem there exists no sampling-based method which converges for all continuous positive functions of finite Dirichlet energy.

Finally, we would like to note that the spectral factorization mapping poses quite a number of further very interesting properties with consequences for practical applications. For example, [12] investigated the *Turing computability* of the spectral factor ϕ_+ for computable continuous spectral densities ϕ. Turing computability is of utmost importance for practical applications, because only if the spectral factor ϕ_+ is Turing computable, it can be calculated numerically on digital hardware like computers, digital signal processors (DSP), or field-programmable gate arrays (FPGAs). In [12] a particular strong result of the non Turing computability of the spectral factorization was proved. In fact, [12] constructed a computable, continuous, and strictly positive spectral density ϕ such that $\phi_+(1)$ is not a computable complex number. In other words, there exists no Turing machine (i.e. an numerical algorithm on a digital computer) which is able to compute the complex number $\phi_+(1)$. Nevertheless, (1) implies that the spectral factor of any strictly positive, computable spectral density ϕ always satisfies $\left|\phi_+(\mathrm{e}^{\mathrm{i}\theta})\right| = \left|\phi(\mathrm{e}^{\mathrm{i}\theta})\right|^{1/2}$ for all $\theta \in [-\pi, \pi)$. Therefore $\|\phi_+\|_{\mathcal{A}(\mathbb{D})} = \|\phi\|^{1/2}_{C(\mathbb{T})}$ is always a computable number. However, ϕ_+ cannot be effectively approximated in the $\mathcal{A}(\mathbb{D})$-norm by elementary computable functions (e.g. by computable sequences polynomials with rational

coefficients). This implies that the spectral factorization is (to the best of our knowledge) the first example of practical relevance showing that the *First-Main Theorem* of Pour-El and Richards [34] for the computability of linear operators can generally not be extended to non-linear operators. We refer also to the proof and discussion of [12, Theorem III.1] for more details on the First-Main Theorem of Pour-El and Richards. It is also interesting to note that the question of the (non-) Turing computability of spectral factors for spectral densities in the scale of Sobolev spaces could be completely characterized (cf. [12, Theorem III.4]). Unfortunately, it turned out that the Sobolev space $H^{1/2}(\mathbb{T})$, which is of particular importance in practical application, contains a large set of computable densities ϕ for which the spectral factor ϕ_+ is not computable.

CRediT authorship contribution statement

Holger Boche: Conceptualization, Methodology, Funding acquisition, Project administration. **Volker Pohl:** Investigation, Writing - original draft, Writing - review & editing.

Appendix. Proofs of auxiliary lemmas

This appendix presents the proof of the important Lemma 3.1. Before that, we state two auxiliary lemmas needed for proving Lemma 3.1 and we give a technical result (Lemma A.3) needed in the proof of Theorem 3.3.

A.1. Auxiliary lemmas

The proof of Lemma 3.1 is based on a seminorm, defined for every $f \in \mathcal{D}_{\mathcal{C}}$ by

$$\|f\|_{\Delta} = \sqrt{\frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{\tau^2} \left| f(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - f(\mathrm{e}^{\mathrm{i}\theta}) \right|^2 \mathrm{d}\theta \, \mathrm{d}\tau} \, . \tag{A.1}$$

The following lemma shows that (A.1) is equivalent to the seminorm $\|\cdot\|_{\mathrm{E}}$ defined in (5).

Lemma A.1. *There exists a constant* $0 < C_{\Delta} < 1$ *such that*

$$C_{\Delta} \, \|f\|_{\mathrm{E}} \leq \|f\|_{\Delta} \leq \|f\|_{\mathrm{E}} \qquad \textit{for all } f \in \mathcal{D}_{\mathcal{C}} \, . \tag{A.2}$$

Proof. Let $f \in \mathcal{D}_{\mathcal{C}} \subset \mathcal{C}(\mathbb{T})$ be arbitrary with Fourier coefficients $\{c_n(f)\}_{n\in\mathbb{Z}}$. Then for arbitrary $\theta, \tau \in \mathbb{T}$, we have $f(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - f(\mathrm{e}^{\mathrm{i}\theta}) = \sum_{n\in\mathbb{Z}} c_n(f) \mathrm{e}^{\mathrm{i}n\theta} \left[\mathrm{e}^{\mathrm{i}n\tau} - 1\right]$ and Parseval's theorem yields

$$\tfrac{1}{2\pi} \int_{-\pi}^{\pi} \left| f(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - f(\mathrm{e}^{\mathrm{i}\theta}) \right|^2 \mathrm{d}\theta = \sum_{n\in\mathbb{Z}} |c_n(f)|^2 \left| \mathrm{e}^{\mathrm{i}n\tau} - 1 \right|^2 \, .$$

Dividing this equation by $2\pi\tau^2$ and integrating over τ yields

$$\begin{aligned} \|f\|_{\Delta}^2 &= \sum_{n\in\mathbb{Z}} |c_n(f)|^2 \, \tfrac{1}{2\pi} \int_{-\pi}^{\pi} \left| \tfrac{\mathrm{e}^{\mathrm{i}n\tau} - 1}{\tau} \right|^2 \mathrm{d}\tau = \sum_{n\in\mathbb{Z}} |c_n(f)|^2 \, \tfrac{1}{2\pi} \int_{-\pi}^{\pi} \left| \tfrac{\sin(n\,\tau/2)}{\tau/2} \right|^2 \mathrm{d}\tau \\ &= \sum_{n\in\mathbb{Z}} |n| \, |c_n(f)|^2 \left(\tfrac{1}{\pi} \int_{-n\pi/2}^{n\pi/2} \left| \tfrac{\sin(\omega)}{\omega} \right|^2 \mathrm{d}\omega \right) . \end{aligned} \tag{A.3}$$

Moreover, the factor on the right hand side satisfies the inequalities

$$C_{\Delta} := \tfrac{1}{\pi} \int_{-\pi/2}^{\pi/2} \left| \tfrac{\sin(\omega)}{\omega} \right|^2 \mathrm{d}\omega \leq \tfrac{1}{\pi} \int_{-n\pi/2}^{n\pi/2} \left| \tfrac{\sin(\omega)}{\omega} \right|^2 \mathrm{d}\omega < \tfrac{1}{\pi} \int_{-\infty}^{\infty} \left| \tfrac{\sin(\omega)}{\omega} \right|^2 \mathrm{d}\omega = 1 \, .$$

Inserting this lower and upper bound into (A.3) yields (A.2). ■

Also the following technically simple lemma is needed in the proof of Lemma 3.1 .

Lemma A.2. *Let $x_0 \in (-1, 0)$ be arbitrary, then*

$$|\log(1+x)| \leq \frac{|x|}{|x_0|} \log\left(\frac{1}{1+x_0}\right) \qquad \text{for all } x \geq x_0 \,. \tag{A.4}$$

Proof. To simplify notation, we set $\psi(x) := \log(1+x)$ for all $x > -1$. By the concavity of ψ, a Taylor series expansion of ψ around $x = 0$ shows that

$$\psi(x) = \log(1+x) \leq x \qquad \text{for all } x > -1 \,. \tag{A.5}$$

Let $x_0 \in (-1, 0)$ be arbitrary but fixed. Then (A.5) implies $-\log(1+x_0) = \log \frac{1}{1+x_0} \geq -x_0 = |x_0|$, showing that $\frac{1}{|x_0|} \log \frac{1}{1+x_0} \geq 1$. Combining this relation with (A.5), one obtains

$$\tfrac{x}{|x_0|} \log \tfrac{1}{1+x_0} \geq x \geq \log(1+x) \qquad \text{for all } x > -1 \,. \tag{A.6}$$

If $x \geq 0$ then (A.6) yields immediately (A.4) because then $|x| = x$ and $|\log(1+x)| = \log(1+x)$.

We still have to consider the case $x \in [x_0, 0)$. Since ψ is concave, we know that $\psi(tx_0) \geq t\psi(x_0) + (1-t)\psi(0) = t\psi(x_0)$ for all $t \in [0, 1]$. Substitute $x = tx_0$ yields

$$\log(1+x) \geq \tfrac{x}{x_0} \log(1+x_0) \quad \text{for all } x \in [x_0, 0] \,. \tag{A.7}$$

Since $x < 0$, we have $\log(1+x) = -|\log(1+x)|$, and because also $x_0 < 0$, Inequality (A.7) yields

$$-|\log(1+x)| \geq -\tfrac{|x|}{|x_0|} \log \tfrac{1}{1+x_0} \qquad \text{for all } x \in [x_0, 0) \,.$$

Multiplying both sides with -1, yields (A.4). ■

The following simple fact is needed in the proof of Theorem 3.3.

Lemma A.3. *Let $g \in \mathcal{A}(\mathbb{D})$ be such that there exists a $c_0 > 0$ such that $|g(z)| \geq c_0$ for all $z \in \overline{\mathbb{D}}$, and let $\{g_N\}_{N\in\mathbb{N}} \subset \mathcal{A}(\mathbb{D})$ be such that*

$$\lim_{N\to\infty} \|g - g_N\|_{\mathcal{A}(\mathbb{D})} = 0 \,. \tag{A.8}$$

Then $\lim_{N\to\infty} \|\log g - \log g_N\|_{\mathcal{A}(\mathbb{D})} = 0$.

Proof. We define for every $N \in \mathbb{N}$ the function φ_N by $\varphi_N(z) = g_N(z)/g(z)$ for all $z \in \overline{\mathbb{D}}$. Since $g(z) \neq 0$ for all $z \in \overline{\mathbb{D}}$, each φ_N belongs to $\mathcal{A}(\mathbb{D})$. Moreover, since $|g(z) - g_N(z)| = |g(z)|\,|1 - \varphi_N(z)| \geq c_0\,|1 - \varphi_N(z)|$ for every $z \in \overline{\mathbb{D}}$, assumption (A.8) implies

$$\lim_{N\to\infty} \|1 - \varphi_N\|_{\mathcal{A}(\mathbb{D})} = 0 \,. \tag{A.9}$$

So there exists an $N_0 \in \mathbb{N}$ such that for every $N \geq N_0$ always $|1 - \varphi_N(z)| < 1/2$ is satisfied. Therefore $|\varphi_N(z)| < 1/2$ for all $N \geq N_0$ and for every $z \in \overline{\mathbb{D}}$. By the continuity of the logarithm, we thus have $\log \varphi_N \in \mathcal{A}(\mathbb{D})$ for all $N \geq N_0$ and so (A.9) implies

$$\lim_{N\to\infty} \|\log \varphi_N\|_{\mathcal{A}(\mathbb{D})} = \lim_{N\to\infty} \|\log g_N - \log g\|_{\mathcal{A}(\mathbb{D})} = 0 \,,$$

which finishes the proof. ■

A.2. Proof of Lemma 3.1

We start by proving the first part of the lemma. To this end, let $\phi \in \mathscr{D}$ be arbitrary and let $c_0 > 0$ be the constant so that

$$\phi(\zeta) \geq c_0 \qquad \text{for all } \zeta \in \mathbb{T} \,. \tag{A.10}$$

Since log is a continuous function, it follows that $u \in \mathcal{C}(\mathbb{T})$ as long as ϕ is continuous and strictly positive. We still have to prove that $\|\phi\|_{\mathrm{E}} < \infty$ implies $\|u\|_{\mathrm{E}} < \infty$. This is done by determine the seminorm $\|u\|_{\Delta}$, introduced in (A.1), and by applying Lemma A.1. To this end, we consider for arbitrary $\theta, \tau \in [-\pi, \pi)$ the expression $\left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right| = \left|\log \frac{\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)})}{\phi(\mathrm{e}^{\mathrm{i}\theta})}\right|$. Since log is monotonically increasing, (A.10) implies

$$\log \frac{c_0}{\|\phi\|_\infty} \leq \log\left(\frac{\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)})}{\phi(\mathrm{e}^{\mathrm{i}\theta})}\right) \leq \log \frac{\|\phi\|_\infty}{c_0} \,. \tag{A.11}$$

Combining both inequalities, we have thus

$$\left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right| = \left|\log\left(\frac{\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)})}{\phi(\mathrm{e}^{\mathrm{i}\theta})}\right)\right| \leq \log \frac{\|\phi\|_\infty}{c_0} \,.$$

Next, we notice that

$$\begin{aligned} \log\left(\frac{\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)})}{\phi(\mathrm{e}^{\mathrm{i}\theta})}\right) &= \log\left(1 + \frac{\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - \phi(\mathrm{e}^{\mathrm{i}\theta})}{\phi(\mathrm{e}^{\mathrm{i}\theta})}\right) \geq \log\left(\frac{c_0}{\|\phi\|_\infty}\right) \\ &= \log\left(1 + \frac{c_0 - \|\phi\|_\infty}{\|\phi\|_\infty}\right) \end{aligned} \tag{A.12}$$

using the first inequality in (A.11) to obtain the right hand side. Now we set

$$x_0 := \frac{c_0 - \|\phi\|_\infty}{\|\phi\|_\infty} = -1 + \frac{c_0}{\|\phi\|_\infty} \qquad \text{and} \qquad x := \frac{\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - \phi(\mathrm{e}^{\mathrm{i}\theta})}{\phi(\mathrm{e}^{\mathrm{i}\theta})} \,.$$

By these definitions, it is clear that $-1 < x_0 < 0$ and (A.12) becomes $\log(1+x) \geq \log(1+x_0)$, showing that $x \geq x_0$. With the so defined x and x_0, Lemma A.2 implies for arbitrary $\theta, \tau \in [-\pi, \pi)$

$$\begin{aligned} \left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right| &= |\log(1+x)| \leq \frac{1}{|x_0|} \log\left(\frac{1}{1+x_0}\right) |x| \\ &= \frac{\|\phi\|_\infty}{\|\phi\|_\infty - c_0} \log\left(\frac{\|\phi\|_\infty}{c_0}\right) |x| \\ &= K_1 \left|\frac{\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - \phi(\mathrm{e}^{\mathrm{i}\theta})}{\phi(\mathrm{e}^{\mathrm{i}\theta})}\right| \leq \frac{K_1}{c_0} \left|\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - \phi(\mathrm{e}^{\mathrm{i}\theta})\right| \end{aligned}$$

with the finite constant $K_1 = \frac{\|\phi\|_\infty}{\|\phi\|_\infty - c_0} \log\left(\frac{\|\phi\|_\infty}{c_0}\right)$. In view of (A.1), this last inequality implies certainly $\|f\|_{\Delta} \leq \frac{K_1}{c_0} \|\phi\|_{\Delta}$ and applying Lemma A.1 yields finally $\|u\|_{\mathrm{E}} \leq \frac{K_1}{c_0 C_\Delta} \|\phi\|_{\mathrm{E}}$ showing that $\|\phi\|_{\mathrm{E}} < \infty$ implies $\|u\|_{\mathrm{E}} < \infty$.

To prove the second part of the lemma, we notice at the beginning that because exp is a continuous function from $\mathbb{R}$ onto the set $\mathbb{R}_+$ of non-negative real numbers, it is clear that $\phi \in \mathcal{C}(\mathbb{T})$ and that there exists $c_0 \geq \exp(-\|u\|_\infty) > 0$ so that (A.10) is satisfied. We still have to show that $\|u\|_{\mathrm{E}} < \infty$ implies $\|\phi\|_{\mathrm{E}} < \infty$. This is done using the seminorm (A.1) and Lemma A.1. Let $\theta, \tau \in [-\pi, \pi)$ be arbitrary. Then

$$\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - \phi(\mathrm{e}^{\mathrm{i}\theta}) = \exp\left[u(\mathrm{e}^{\mathrm{i}\theta})\right] \left(\exp\left[u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right] - 1\right)$$

and consequently

$$\left|\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - \phi(\mathrm{e}^{\mathrm{i}\theta})\right| \leq \exp\bigl(\|u\|_\infty\bigr) \cdot \left|\exp\left[u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right] - 1\right| . \tag{A.13}$$

Next, we use the Taylor series expansion of the exp-function to derive an upper bound on the last factor in (A.13). Together with the triangle inequality, we get

$$\begin{aligned}
\left|\exp\bigl[u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\bigr] - 1\right| &\leq \textstyle\sum_{k=1}^{\infty} \frac{1}{k!} \left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right|^k \\
&= \left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right| \textstyle\sum_{k=1}^{\infty} \frac{1}{k!} \left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right|^{k-1} \\
&\leq \left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right| \textstyle\sum_{k=1}^{\infty} \frac{1}{k!} \bigl(2\,\|u\|_\infty\bigr)^{k-1} \\
&= \frac{\left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right|}{2\|u\|_\infty} \textstyle\sum_{k=1}^{\infty} \frac{1}{k!} \bigl(2\,\|u\|_\infty\bigr)^{k} \\
&= \left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right| \frac{\exp(2\|u\|_\infty)-1}{2\|u\|_\infty} .
\end{aligned} \tag{A.14}$$

Inserting the last inequality into (A.13) gives

$$\left|\phi(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - \phi(\mathrm{e}^{\mathrm{i}\theta})\right| \leq C(u)\, \left|u(\mathrm{e}^{\mathrm{i}(\theta+\tau)}) - u(\mathrm{e}^{\mathrm{i}\theta})\right| \qquad \text{for all } \theta, \tau \in [-\pi, \pi)$$

with the positive constant $C(u) := \exp(\|u\|_\infty) \frac{\exp(2\|u\|_\infty)-1}{2\|u\|_\infty}$. This inequality implies immediately a corresponding inequality for the seminorms (A.1) of ϕ and u, and together with (A.2), we get

$$C_\Delta\ \|\phi\|_{\mathrm{E}} \leq \|\phi\|_\Delta \leq C(u)\ \|u\|_\Delta \leq C(u)\ \|u\|_{\mathrm{E}} .$$

So if $u \in \mathcal{D}_{\mathcal{C}}$, i.e. if $\|u\|_{\mathrm{E}} < +\infty$ the last inequality implies $\|\phi\|_{\mathrm{E}} < +\infty$, i.e. $\phi \in \mathcal{D}_{\mathcal{C}}$. ■

References

[1] B. Anderson, Continuity of the spectral factorization operation, Math. Appl. Comput. 4 (2) (1985) 139–156.
[2] B. Anderson, K. Hitz, N. Diem, Recursive algorithm for spectral factorization, IEEE Trans. Circuits Syst. 21 (6) (1974) 742–750.
[3] G. Baggio, A. Ferrante, Parametrization of minimal spectral factors of discrete-time rational spectral densities, IEEE Trans. Automat. Control 64 (1) (2019) 396–403.
[4] G. Baggio, A. Ferrante, R. Sepulchre, Conal distances between rational spectral densities, IEEE Trans. Automat. Control 64 (5) (2019) 1148–1857.
[5] S. Barclay, Continuity of the spectral factorization mapping, J. Lond. Math. Soc. (2) 70 (2004) 763–779.
[6] H. Boche, V. Pohl, Structural properties of the wiener filter – stability, smoothness properties, and FIR approximation behavior, IEEE Trans. Inform. Theory 52 (9) (2006) 4272–4282.
[7] H. Boche, V. Pohl, Continuity versus boundedness of the spectral factorization mapping, Studia Math. 189 (2) (2008) 131–145.
[8] H. Boche, V. Pohl, On the boundedness of the spectral factorization mapping on decomposing Banach algebras, SIAM J. Control Optim. 47 (3) (2008) 1415–1427.
[9] H. Boche, V. Pohl, On the strong divergence of Hilbert transform approximations and a problem of Ul'yanov, J. Approx. Theory 204 (2016) 34–60.
[10] H. Boche, V. Pohl, CaLculating the Hilbert transform on spaces with energy concentration: Convergence and divergence regions, IEEE Trans. Inform. Theory 65 (1) (2019) 586–603.
[11] H. Boche, V. Pohl, Investigations on the approximability and computability of the Hilbert transform with applications, Appl. Comput. Harmon. Anal. 48 (2) (2020) 706–730.
[12] H. Boche, V. Pohl, On the algorithmic solvability of spectral factorization and applications, IEEE Trans. Inform. Theory 66 (2020) http://dx.doi.org/10.1109/TIT.2020.2968028, in press, online available.
[13] H.W. Bode, C.E. Shannon, A simplified derivation of linear least square smoothing and prediction theory, Proc. IRE 38 (1950) 417–425.
[14] J. Chen, J. Benesty, Y. Huang, S. Doclo, New insights into the noise reduction wiener filter, IEEE Trans. Audio, Speech, Language Process. 14 (4) (2006) 1218–1234.

[15] S. Chen, R. Varma, A. Sandryhaila, J. Kovačević, Discrete signal processing on graphs: Sampling theory, IEEE Trans. Signal Process. 63 (24) (2015) 6510–6523.
[16] Y.C. Eldar, T. Michaeli, Beyond bandlimited sampling: Nonlinearities, smoothness and sparsity, IEEE Signal Process. Mag. 26 (3) (2009) 48–68.
[17] J.B. Garnett, Bounded Analytic Functions, Academic Press, New York, 1981.
[18] T.T. Georgiou, P.P. Khargonekar, Spectral factorization of matrix-valued functions using interpolation theory, IEEE Trans. Circuit. Syst. 36 (4) (1989) 568–574.
[19] R. Hardie, A fast image super-resolution algorithm using an adaptive wiener filter, IEEE Trans. Image Process. 16 (12) (2007) 2953–2964.
[20] M.H. Hayes, Statistical Digital Signal Processing and Modeling, John Wiley & Sons, Inc., New York, etc., 1996.
[21] P. Hoeher, S. Kaiser, P. Robertson, Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering, in: IEEE Intern. Conf. Acoustics, Speech, and Signal Process., vol. 3, 1997-04 pp. 1845–1848.
[22] M.L. Honig, J.S. Goldstein, Adaptive reduced-rank interference suppression based on the multistage wiener filter, IEEE Trans. Commun. 50 (6) (2002) 986–994.
[23] B. Jacob, J.R. Partington, On the boundedness and continuity of the spectral factorization mapping, SIAM J. Control Optim. 40 (1) (2001) 88–106.
[24] B. Jacob, J. Winkin, H. Zwart, Continuity of the spectral factorization on a vertical strip, Systems Control Lett. 37 (1999) 183–192.
[25] A.J. Jerri, The Shannon sampling theorem – its various extensions and applications: A tutorial review, Proc. IEEE 65 (11) (1977) 1565–1596.
[26] X. Jiang, L. Ning, T.T. Georgiou, Distances and Riemannian metrics for multivariate spectral densities, IEEE Trans. Automat. Control 57 (7) (2012) 1723–1735.
[27] T. Kailath, A.H. Sayed, B. Hassibi, Linear Estimation, Prentice-Hall, Inc., Upper Saddle River, NJ, 2000.
[28] A.N. Kolmogorov, Sur l'interpolation et extrapolation des suites stationnaires, C.R. Acad. Sci. Paris 208 (1939) 2043–2045.
[29] P. Koosis, Introduction to H_p Spaces, Cambridge University Press, Cambridge, 1980.
[30] A. Lindquist, G. Picci, Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification, Springer Verlag, Berlin, Heidelberg, 2015.
[31] E. Masry, Polynomial interpolation and prediction of continuous-time processes from random samples, IEEE Trans. Inform. Theory 43 (2) (1997) 776–783.
[32] A.V. Oppenheim, A.S. Willsky, S.H. Nawab, Signals and Systems, Prentice-Hall, Upper Saddle River, NJ, 1997.
[33] V. Pohl, H. Boche, Advanced topics in system and signal theory: A mathematical approach, Foundations in Signal Processing, Communications and Networking, Springer-Verlag, Berlin, 2009.
[34] M.B. Pour-El, J.I. Richards, Computability in Analysis and Physics, Springer-Verlag, Berlin, 1989.
[35] J.G. Proakis, D.G. Maolakis, Digital signal processing: Principles, algorithms and applications, McMillan, New York, 1992.
[36] W.T. Ross, The classical Dirichlet space, Contemp. Math. 393 (2006) 171–197.
[37] A.H. Sayed, T. Kailath, A survey of spectral factorization methods, Numer. Linear Algebra Appl. 8 (2001) 467–496.
[38] C.E. Shannon, Communication in the presence of noise, Proc. IRE 37 (1) (1949) 10–21.
[39] O.J. Staffans, Quadratic optimal control of stable systems through spectral factorization, Math. Control Signals Systems 8 (1995) 167–197.
[40] H.J. Trusseli, M.S. Kulkarni, Sampling and processing of color signals, IEEE Trans. Image Process. 5 (4) (1996) 677–681.
[41] M. Unser, Sampling – 50 years after Shannon, Proc. IEEE 88 (4) (2000) 569–587.
[42] M. Unser, A. Aldroubi, M. Eden, B-spline signal processing. I. Theory, IEEE Trans. Signal Process. 41 (2) (1993) 821–833.
[43] K.S. Vastola, H.V. Poor, Robust Wiener-Kolmogorov theory, IEEE Trans. Inform. Theory IT-30 (2) (1984) 316–327.
[44] N. Wiener, Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications, The MIT Press and John Wiley & Sons, Inc., Cambridge, 1949.
[45] N. Wiener, E. Hopf, Über eine Klasse singulärer Integralgleichungen, Sitzungsberichte der Königlich Preußische Akademie der Wissenschaften zu Berlin (1931) 696–706.

[46] N. Wiener, P. Masani, The prediction theory of multivariate stochastic processes, I. The regularity condition, Acta Math. 98 (1957) 111–150.
[47] D.C. Youla, N.N. Kazanjian, Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle, IEEE Trans. Circuits Syst. 25 (2) (1978) 57–69.
[48] K. Zhou, J.C. Doyle, K. Glover, Robust and Optimal Control, Prentice Hall, Englewood Cliffs, NJ, USA, 1996.