Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 54-7 (2021) 553–558

A Generalized Multidimensional Circulant Rational Covariance and Cepstral Extension Problem ⋆

Bin Zhu *, **Mattia Zorzi** **

* *School of Intelligent Systems Engineering, Sun Yat-sen University, Waihuan East Road 132, 510006 Guangzhou, China (e-mail: zhub26@mail.sysu.edu.cn)*

** *Department of Information Engineering, University of Padova, Via Gradenigo 6/B, 35131 Padova, Italy (e-mail: zorzimat@dei.unipd.it)*

Abstract: We propose a general scenario to estimate the spectral density of an homogeneous random field from its moments. More precisely, we consider a multidimensional rational covariance and cepstral extension problem. The latter is usually solved by searching the spectral density maximizing the entropy rate while matching the moments. The generality of our mathematical formulation can be seen from the employed entropic index as well as the definition of cepstral coefficients. We characterize the solution in the circulant case. Finally, we apply our theory to a 2-d system identification problem.

Keywords: Multidimensional circulant covariance extension, generalized cepstral matching, alpha-entropy, convex optimization.

1. INTRODUCTION

Homogeneous random fields represent a traditional tool in the engineering community; for instance, they are employed for target parameter estimation in automotive radars, see Engels (2014); Zhu et al. (2019). In this paper, we consider a zero-mean real-valued Gaussian homogeneous random field $y(\mathbf{t})$, where $\mathbf{t} = (t_1, \dots, t_d)$ is a vector-valued index belonging to the discrete domain $\mathbb{Z}^d$, with $d \in \mathbb{N}_+$. Its covariance $c_{\mathbf{k}} := \mathbb{E}[\, y(\mathbf{t}+\mathbf{k})y(\mathbf{t})\,]$ depends only on the index difference $\mathbf{k} = (k_1, \dots, k_d) \in \mathbb{Z}^d$. In the 1-dimensional case ($d = 1$), this is also known as a (wide-sense) stationary process. By the spectral representation theorem (cf. e.g., Yaglom, 1957), the covariances are Fourier coefficients of a nonnegative spectral measure $\mathrm{d}\gamma$ on $\mathbb{T}^d := (-\pi, \pi]^d$. In practice, we often have to infer the unknown spectral measure from a number of covariances typically estimated from a finite-size realization of the underlying random field. More precisely, let $\Lambda \subset \mathbb{Z}^d$ be a specific index set that has a finite cardinality, contains the all-zero index, and is symmetric with respect to the origin, i.e. $\mathbf{k} \in \Lambda$ implies $-\mathbf{k} \in \Lambda$. Here, we consider Λ as the multidimensional box $\{\mathbf{k} \in \mathbb{Z}^d \,:\, |k_j| \le n_j,\ j = 1, \dots, d\}$ where n_j's are given integers. Assuming that the spectral measure $\mathrm{d}\gamma$ is *absolutely continuous* with respect to the Lebesgue measure, we have the (trigonometric) moment equations

$$c_{\mathbf{k}} = \int_{\mathbb{T}^d} e^{i\langle \mathbf{k},\, \boldsymbol{\theta}\rangle} \Phi(e^{i\boldsymbol{\theta}}) \mathrm{d}\mu(\boldsymbol{\theta}) \quad \text{for all } \mathbf{k} \in \Lambda, \tag{1}$$

where $\boldsymbol{\theta} = (\theta_1, \dots, \theta_d) \in \mathbb{T}^d$, $\langle \mathbf{k}, \boldsymbol{\theta}\rangle := \sum_j k_j\theta_j$ is the Euclidean inner product, $\mathrm{d}\mu(\boldsymbol{\theta}) = (2\pi)^{-d} \prod_{j=1}^d \mathrm{d}\theta_j$ is the normalized Lebesgue measure on $\mathbb{T}^d$, $e^{i\boldsymbol{\theta}}$ is a shorthand notation for $(e^{i\theta_1}, \dots, e^{i\theta_d})$, and $\Phi(e^{i\boldsymbol{\theta}})$ is the power spectral density (i.e. the Radon-Nikodym derivative of $\mathrm{d}\gamma$ with respect to $\mathrm{d}\mu$). However, the problem to find Φ satisfying (1) is typically ill-posed since in general, it has multiple solutions if one solution exists.

In order to remedy such ill-posedness, a standard approach in the literature known as *Rational Covariance Extension* (RCE) is to reformulate the problem as a constrained optimization one. As one of the forerunners, Byrnes et al. (2001) proposed a *weighted maximum entropy* formulation in the case of $d = 1$. It turns out that, under certain technical conditions, the solution to the above problem has the form of a rational function $\Phi(e^{i\boldsymbol{\theta}}) = P(e^{i\boldsymbol{\theta}})/Q(e^{i\boldsymbol{\theta}})$ where P, Q are positive Laurent polynomials and Q is uniquely determined by P. Thus, one can obtain a complete parametrization of rational solutions of bounded degree to the moment equations (1). It is worth noting that several extensions in the 1-d case have been proposed, see for instance Georgiou and Lindquist (2003); Ferrante et al. (2008); Zorzi (2015); Zhu (2020), while few extensions to the multidimensional case are available, see Georgiou (2006). The importance to obtain a rational spectral density lies in the observation that one can build a finite-dimensional linear stochastic system via spectral factorization of the resulting spectrum, whose output is equivalent to the original random process in the sense that they have the same covariance function. For this reason,

⋆ This work was supported by the SID project "A Multidimensional and Multivariate Moment Problem Theory for Target Parameter Estimation in Automotive Radars" (ZORZ_SID19_01) funded by the Department of Information Engineering of the University of Padova. B. Zhu was also partially supported by the "Hundred-Talent Program" of the Sun Yat-sen University.

10.1016/j.ifacol.2021.08.418

the RCE paradigm can be viewed as an identification or modeling problem.

In the above formulation, the polynomial P gives spectral zeros and it must be properly chosen for a particular application. In the special case of $P \equiv 1$, Burg's maximum entropy principle can be recovered, which yields an all-pole model. One way to choose such P is to use additional data in the form of *cepstral coefficients*, which are defined as follows:

$$m_{\mathbf{k}} = \int_{\mathbb{T}^d} e^{i\langle \mathbf{k}, \boldsymbol{\theta}\rangle} \log \Phi(e^{i\boldsymbol{\theta}}) \mathrm{d}\mu(\boldsymbol{\theta}) \quad \text{for all } \mathbf{k} \in \Lambda_0, \tag{2}$$

where the index set $\Lambda_0 = \Lambda \setminus \{\mathbf{0}\}$. Notice that the cepstral coefficient at index $\mathbf{0}$ does not play a role in the analysis, and hence it is excluded at the beginning. We can then set up a constrained maximum entropy problem (Ringh et al., 2016, 2018):

$$\max_{\Phi \geq 0} \int_{\mathbb{T}^d} \log \Phi(e^{i\boldsymbol{\theta}}) \mathrm{d}\mu(\boldsymbol{\theta}) \quad \text{s.t. (1) and (2).} \tag{3}$$

The latter is referred to as the multidimensional rational covariance and cepstral extension problem. If some technical conditions are satisfied, such a problem admits a unique solution having the form $\Phi(e^{i\boldsymbol{\theta}}) = P(e^{i\boldsymbol{\theta}})/Q(e^{i\boldsymbol{\theta}})$ with the constant term of the numerator $p_{\mathbf{0}} = 1$ fixed.

The aim of this paper is to generalize the above framework in two respects. First, we employ the α-entropy as the the objective functional which is more general than the one used in (3). Second, we generalize the constraint (2) accordingly. We analyze the existence of a solution to the corresponding problem by means of duality theory. It turns out that such a solution does not always exist. On the other hand, it is always possible to find an approximate solution which solves a regularized version of the dual problem. Finally, we apply this generalized theory to solve a 2-d system identification problem.

2. GENERALIZED PROBLEM

Before introducing our general problem formulation, we need to generalize the concept of cepstral coefficients.

Definition 1. Given the power spectral density $\Phi(e^{i\boldsymbol{\theta}})$ and a real number α such that $0 < \alpha \leq 1$, the generalized cepstral coefficients are defined as (Tokuda et al., 1990)

$$m_{\alpha,\mathbf{k}} = \begin{cases} \dfrac{1}{\alpha} \displaystyle\int_{\mathbb{T}^d} e^{i\langle \mathbf{k}, \boldsymbol{\theta}\rangle} \Phi(e^{i\boldsymbol{\theta}})^\alpha \mathrm{d}\mu(\boldsymbol{\theta}) & \text{if } \mathbf{k} \neq \mathbf{0}, \\ \dfrac{1}{\alpha} \left(\displaystyle\int_{\mathbb{T}^d} \Phi(e^{i\boldsymbol{\theta}})^\alpha \mathrm{d}\mu(\boldsymbol{\theta}) - 1 \right) & \text{if } \mathbf{k} = \mathbf{0}. \end{cases} \tag{4}$$

Extending by continuity the above definition we have

$$\lim_{\alpha \to 0} m_{\alpha,\mathbf{k}} = \int_{\mathbb{T}^d} e^{i\langle \mathbf{k}, \boldsymbol{\theta}\rangle} \log \Phi(e^{i\boldsymbol{\theta}}) \mathrm{d}\mu(\boldsymbol{\theta}),$$

that is, we recover the standard cepstral coefficients. To ease the notation, we drop the subscript α and just write $m_{\mathbf{k}}$ instead. The idea is to generalize the problem in (3) by using the definition above for the cepstral coefficients. Accordingly, given two sets of real numbers $\mathbf{c} = \{c_{\mathbf{k}}\}_{\mathbf{k} \in \Lambda}$ such that $c_{\mathbf{k}} = c_{-\mathbf{k}}$, and $\mathbf{m} = \{m_{\mathbf{k}}\}_{\mathbf{k} \in \Lambda_0}$ such that $m_{\mathbf{k}} = m_{-\mathbf{k}}$, we want to find a rational function $\Phi : \mathbb{T}^d \to \mathbb{R}_+$ such that (1) and

$$m_{\mathbf{k}} = \frac{1}{\alpha} \int_{\mathbb{T}^d} e^{i\langle \mathbf{k}, \boldsymbol{\theta}\rangle} \Phi(e^{i\boldsymbol{\theta}})^\alpha \mathrm{d}\mu(\boldsymbol{\theta}) \quad \text{for all } \mathbf{k} \in \Lambda_0 \tag{5}$$

hold.

At this point we need to introduce a suitable definition of entropy which is "consistent" with the definition of the generalized cepstral coefficients in (4). An entropic index should measure how close the spectral density is to normalized white noise, i.e. a random field having constant spectral density equal to one. To this purpose, we consider the alpha divergence between two spectral densities Φ and Ψ (Zorzi, 2014):

$$\mathbb{D}_\alpha(\Phi \| \Psi) = \int_{\mathbb{T}^d} \left(\frac{1}{\alpha(\alpha-1)} \Phi(e^{i\boldsymbol{\theta}})^\alpha \Psi(e^{i\boldsymbol{\theta}})^{1-\alpha} + \frac{1}{1-\alpha} \Phi(e^{i\boldsymbol{\theta}}) + \frac{1}{\alpha} \Psi(e^{i\boldsymbol{\theta}}) \right) \mathrm{d}\mu(\boldsymbol{\theta})$$

where $\alpha \in \mathbb{R} \setminus \{0, 1\}$. Then, given the power spectral density $\Phi(e^{i\boldsymbol{\theta}})$ and $\alpha \in \mathbb{R} \setminus \{0, 1\}$, we define as alpha-entropy:

$$\begin{aligned} \mathbb{H}_\alpha(\Phi) &= -\mathbb{D}_\alpha(\Phi \| 1) + \frac{1}{1-\alpha} \left(\int_{\mathbb{T}^d} \Phi(e^{i\boldsymbol{\theta}}) \mathrm{d}\mu(\boldsymbol{\theta}) - 1 \right) \\ &= \frac{1}{\alpha(\alpha-1)} \left(1 - \int_{\mathbb{T}^d} \Phi(e^{i\boldsymbol{\theta}})^\alpha \mathrm{d}\mu(\boldsymbol{\theta}) \right). \end{aligned} \tag{6}$$

The latter can be extended by continuity to $\alpha = 0$:

$$\lim_{\alpha \to 0} \mathbb{H}_\alpha(\Phi) = \int_{\mathbb{T}^d} \log \Phi(e^{i\boldsymbol{\theta}}) \mathrm{d}\mu(\boldsymbol{\theta})$$

which corresponds to the usual entropy rate. Indeed, $\int_{\mathbb{T}^d} \Phi \mathrm{d}\mu$ is fixed by the constraint (1) with $\mathbf{k} = \mathbf{0}$. Hence, the constant term $\int_{\mathbb{T}^d} \Phi \mathrm{d}\mu - 1$ plays no role in our problem. Combining the alpha-entropy defined in (6) and the generalized cepstral coefficients defined in (4), we obtain the generalized problem

$$\max_{\Phi \geq 0} \mathbb{H}_\alpha(\Phi) \quad \text{s.t. (1) and (5)} \tag{7}$$

which incorporates (3) as a limit case ($\alpha \to 0$). In the case where $\alpha \to 1$, it is not difficult to see that (5) boils down to (1).

Although the optimization problem in (7) looks appealing, one may encounter significant technical difficulties in its analysis due to multiple dimensionality. As noticed by Ringh et al. (2015) and Zhu et al. (2021), the situation is simpler if one considers the discrete version of the problem, which at the same time, yields a computable theory. In particular, the fast Fourier transform (FFT) can be used to compute the Fourier integrals on a discrete grid. In addition, the discrete spectrum can be viewed as a sampling in the frequency domain, which corresponds to a *periodic* homogeneous field (in the "space" domain) as explained in Zhu et al. (2021). Since the covariance matrix of a stationary periodic process in the 1-d case has a *circulant* structure, the discrete problem is also referred to as Circulant Rational Covariance Extension. Since in practice we always have a finite collection of observations of the random field, we may just adopt the mathematical idealization that the data come from a periodic field, provided that the period is sufficiently large, see Carli et al. (2011); Lindquist and Picci (2013).

In order to state the discrete problem, we need to set some notations first. Let $\mathbf{N}$ denote the vector $(N_1, N_2, \ldots, N_d) \in \mathbb{N}_+^d$, where the component N_j stands for the number of equal-length partitions of the interval $(-\pi, \pi]$ in the j-th dimension of $\mathbb{T}^d$. Next, define the index set

$\mathbb{Z}^d_{\mathbf{N}} := \{\boldsymbol{\ell} = (\ell_1, \dots, \ell_d) : 0 \le \ell_j \le N_j - 1,\ j = 1, \dots, d\}$, whose cardinality is $|\mathbf{N}| := \prod_{j=1}^d N_j$. We can now discretize the domain $\mathbb{T}^d$ as $\mathbb{T}^d_{\mathbf{N}} := \{(\frac{2\pi}{N_1}\ell_1, \dots, \frac{2\pi}{N_d}\ell_d) : \boldsymbol{\ell} \in \mathbb{Z}^d_{\mathbf{N}}\}$, so the number of grid points is also $|\mathbf{N}|$. Moreover, let us introduce the symbol $\boldsymbol{\zeta}_{\boldsymbol{\ell}} := (\zeta_{\ell_1}, \dots, \zeta_{\ell_d})$ for a point in the discrete d-torus with $\zeta_{\ell_j} = e^{i2\pi\ell_j/N_j}$ and define $\boldsymbol{\zeta}^{\mathbf{k}}_{\boldsymbol{\ell}} := \prod_{j=1}^d \zeta^{k_j}_{\ell_j}$, which is just the complex exponential function $e^{i\langle \mathbf{k},\, \boldsymbol{\theta}\rangle}$ evaluated at a grid point in $\mathbb{T}^d_{\mathbf{N}}$. We can now define a discrete measure with equal mass on the grid points in $\mathbb{T}^d_{\mathbf{N}}$ as

$$\mathrm{d}\eta(\boldsymbol{\theta}) = \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \delta(\theta_1 - \frac{2\pi}{N_1}\ell_1, \dots, \theta_d - \frac{2\pi}{N_d}\ell_d) \prod_{j=1}^d \frac{\mathrm{d}\theta_j}{N_j}. \quad (8)$$

One can easily verify that integrals against $\mathrm{d}\eta$ are essentially Riemann sums. The discrete version of the optimization problem (7) results when we replace the normalized Lebesgue measure $\mathrm{d}\mu$ with the discrete measure $\mathrm{d}\eta$. In addition, for reasons that will be clear in the next section, we take the parametrization $\alpha = 1 - \nu^{-1}$ with $\nu \in \mathbb{N}_+$ and $\nu > 1$. So the optimization problem in (7) becomes

$$\begin{aligned} \max_{\Phi\ge 0} \quad & \frac{\nu^2}{(\nu-1)} \left(\frac{1}{|\mathbf{N}|} \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\frac{\nu-1}{\nu}} - 1 \right) \\ \text{s.t.} \quad & c_{\mathbf{k}} = \frac{1}{|\mathbf{N}|} \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \boldsymbol{\zeta}^{\mathbf{k}}_{\boldsymbol{\ell}} \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \quad \forall \mathbf{k} \in \Lambda, \\ & m_{\mathbf{k}} = \frac{\nu}{\nu-1} \frac{1}{|\mathbf{N}|} \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \boldsymbol{\zeta}^{\mathbf{k}}_{\boldsymbol{\ell}} \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\frac{\nu-1}{\nu}} \quad \forall \mathbf{k} \in \Lambda_0. \end{aligned} \quad (9)$$

This will be the focus of investigation in the rest of this paper.

3. DUAL ANALYSIS

The aim of this section is to characterize the rational solution (if it exists) to Problem (9) by means of the dual analysis. Thus, we consider the objective function in (9) multiplied by ν^{-1} and discard the constant term not depending on Φ. Thus, the Lagrangian is

$$\begin{aligned} \mathcal{L}_\nu(\Phi, P, Q) = & \frac{\nu}{(\nu-1)|\mathbf{N}|} \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\frac{\nu-1}{\nu}} \\ & - \frac{1}{|\mathbf{N}|} \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) + \langle \mathbf{q},\, \mathbf{c}\rangle - \langle \mathbf{p},\, \mathbf{m}\rangle \end{aligned} \quad (10)$$

where $\mathbf{p} = \{p_{\mathbf{k}}\}_{\mathbf{k}\in\Lambda}$ and $\mathbf{q} = \{q_{\mathbf{k}}\}_{\mathbf{k}\in\Lambda}$ are the Lagrange multipliers such that $p_{\mathbf{0}} = 1$, $p_{\mathbf{k}} = p_{-\mathbf{k}} \in \mathbb{R}$ and $q_{\mathbf{k}} = q_{-\mathbf{k}} \in \mathbb{R}$. Moreover,

$$P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = \sum_{\mathbf{k}\in\Lambda} p_{\mathbf{k}} \boldsymbol{\zeta}^{-\mathbf{k}}_{\boldsymbol{\ell}}, \quad Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = \sum_{\mathbf{k}\in\Lambda} q_{\mathbf{k}} \boldsymbol{\zeta}^{-\mathbf{k}}_{\boldsymbol{\ell}} \quad (11)$$

are Laurent polynomials. The inner product between two multisequences is denoted by $\langle \mathbf{q},\, \mathbf{c}\rangle := \sum_{\mathbf{k}\in\Lambda} q_{\mathbf{k}} c_{\mathbf{k}}$, and in $\langle \mathbf{p},\, \mathbf{m}\rangle$ we set $m_{\mathbf{0}}$ equal to an arbitrary but fixed number. Notice that we take the integer $\nu \ge 2$ so that the fraction $\nu/(\nu-1) > 0$. Next we fix the Lagrange multipliers $\mathbf{q}$ and $\mathbf{p}$, or equivalently the polynomials Q and P, and consider the problem

$$\sup_{\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})\ge 0\ \forall \boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \mathcal{L}_\nu(\Phi, P, Q).$$

The last two inner products in (10) do not depend on Φ, and thus can be ignored in the sup problem. There are a number of cases to analyze. We first claim that in order for the supremum to be finite, it is necessary that $Q \ge 0$ on $\mathbb{T}^d_{\mathbf{N}}$ (more precisely, on the d-torus). To see this, consider the following two cases:

(i) If $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) < 0$ for some $\boldsymbol{\ell} \in \mathbb{Z}^d_{\mathbf{N}}$ and $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \ge 0$, then we can take $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \to +\infty$ and $\Phi = 0$ elsewhere on $\mathbb{T}^d_{\mathbf{N}}$, which leads to

$$\begin{aligned} \mathcal{L}_\nu(\Phi, P, Q) = & \frac{\nu}{(\nu-1)|\mathbf{N}|} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\frac{\nu-1}{\nu}} \\ & - \frac{1}{|\mathbf{N}|} Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) + \cdots \to +\infty. \end{aligned} \quad (12)$$

(ii) If $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) < 0$ for some $\boldsymbol{\ell} \in \mathbb{Z}^d_{\mathbf{N}}$ and $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) < 0$, we can take the same choice of Φ as in the previous point. By comparison of the first two terms on the right-hand side of (12)

$$\frac{\frac{\nu}{(\nu-1)|\mathbf{N}|} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\frac{\nu-1}{\nu}}}{\frac{1}{|\mathbf{N}|} Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})} = \frac{\nu P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{(\nu-1) Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})} \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{-\frac{1}{\nu}} \to 0,$$

we conclude again that $\mathcal{L}_\nu(\Phi, P, Q) \to +\infty$.

Thus, we have to assume $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \ge 0$ for all $\boldsymbol{\ell} \in \mathbb{Z}^d_{\mathbf{N}}$. Next, we discuss different situations concerning the polynomial P. Suppose $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) < 0$ for some index $\boldsymbol{\ell}$. Then the summand in the Lagrangian

$$\frac{1}{|\mathbf{N}|} \left[\frac{\nu}{\nu-1} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\frac{\nu-1}{\nu}} - Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \right] \le 0. \quad (13)$$

In other words, $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$ maximizes the Lagrangian. Therefore, it is equivalent to consider the problem

$$\sup_{\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})\ge 0\ \forall \boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \mathcal{L}'_\nu(\Phi, P, Q)$$

where the modified Lagrangian is

$$\begin{aligned} \mathcal{L}'_\nu(\Phi, P, Q) := \sum_{\boldsymbol{\ell}:\, P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})\ge 0} \Big[& \frac{\nu}{\nu-1} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\frac{\nu-1}{\nu}} \\ & - Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Big]. \end{aligned}$$

We divide our discussion into the following four situations.

(i) Points $\boldsymbol{\zeta}_{\boldsymbol{\ell}}$ where $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$. In this case, $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})$ has no effects on $\mathcal{L}'_\nu(\Phi, P, Q)$ and can be set arbitrarily.
(ii) Points $\boldsymbol{\zeta}_{\boldsymbol{\ell}}$ where $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$ and $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0$. The reasoning is the same as that in (13) so one should set $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$.
(iii) Points $\boldsymbol{\zeta}_{\boldsymbol{\ell}}$ where $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0$ and $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$. In this case, taking $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \to +\infty$ and $\Phi = 0$ elsewhere gives a value of the Lagrangian equal to $+\infty$. In other words, if $\mathcal{L}'_\nu(\Phi, P, Q)$ is finite, we must have $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0 \implies Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0$. Such a condition can also be rephrased as $\mathrm{supp}(Q) \supset \{\boldsymbol{\zeta}_{\boldsymbol{\ell}} : P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0\}$ where $\mathrm{supp}(\cdot)$ denotes the support of a function.
(iv) Points $\boldsymbol{\zeta}_{\boldsymbol{\ell}}$ where $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0$ and $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0$. For a feasible direction $\delta\Phi$ such that $\Phi + \varepsilon\delta\Phi \ge 0$ on $\mathbb{T}^d_{\mathbf{N}}$ for sufficiently small $\varepsilon > 0$, let us compute the directional derivative of the modified Lagrangian

$$\begin{aligned} & \delta\mathcal{L}'_\nu(\Phi, P, Q; \delta\Phi) \\ & = \sum_{\boldsymbol{\ell}:\, P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})>0} \left[\left(P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{-\frac{1}{\nu}} - Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \right) \delta\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \right]. \end{aligned}$$

Clearly, the stationary point of the modified Lagrangian is such that $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{-\frac{1}{\nu}} - Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$ which yields

$$\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = \left(\frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right)^{\nu}. \tag{14}$$

Notice that this is a well defined function since we have started from the condition $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0$.

It is not difficult to show that $\mathcal{L}'_{\nu}(\Phi, P, Q)$ is a concave function in Φ (via reasoning on each summand) so that a stationary point is indeed a maximizer. To summarize, the supremum of the Lagrangian is attained at the function

$$\Phi_{\nu}(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = \begin{cases} 0 & \text{if } P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \leq 0 \\ (P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})/Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}))^{\nu} & \text{if } P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0. \end{cases} \tag{15}$$

Plugging the functional form Φ_{ν} into the original Lagrangian (10), we obtain the dual function

$$J_{\nu}(P,Q) = \frac{1}{(\nu-1)|\mathbf{N}|} \sum_{\boldsymbol{\ell}:\, P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})>0} \frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\nu}}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\nu-1}} + \langle \mathbf{q},\, \mathbf{c}\rangle - \langle \mathbf{p},\, \mathbf{m}\rangle. \tag{16}$$

The dual problem is

$$\min_{P,Q}\ J_{\nu}(P,Q) \quad \text{s.t. } (P,Q) \in \mathscr{L} \tag{17}$$

with feasible set

$$\mathscr{L} := \{(P,Q)\ :\ Q \geq 0,\ \operatorname{supp}(Q) \supset \{\boldsymbol{\zeta}_{\boldsymbol{\ell}} : P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0\}\}. \tag{18}$$

Let us proceed to compute the first variations of J_{ν} at any *interior point* (P,Q) along directions δP and δQ:

$$\begin{aligned} &\delta J_{\nu}(P,Q;\delta Q) \\ &= \sum_{\mathbf{k}\in\Lambda} \delta q_{\mathbf{k}} \left[-\frac{1}{|\mathbf{N}|} \sum_{\boldsymbol{\ell}:\, P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})>0} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\mathbf{k}} \left(\frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right)^{\nu} + c_{\mathbf{k}} \right], \\ &\delta J_{\nu}(P,Q;\delta P) \\ &= \sum_{\mathbf{k}\in\Lambda} \delta p_{\mathbf{k}} \left[\frac{\nu}{(\nu-1)|\mathbf{N}|} \sum_{\boldsymbol{\ell}:\, P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})>0} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\mathbf{k}} \left(\frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right)^{\nu-1} - m_{\mathbf{k}} \right]. \end{aligned}$$

Imposing the derivatives to vanish in any feasible direction, we arrive at the conditions of covariance and generalized cepstral matching:

$$\begin{aligned} c_{\mathbf{k}} &= \frac{1}{|\mathbf{N}|} \sum_{\boldsymbol{\ell}:\, P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})>0} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\mathbf{k}} \left(\frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right)^{\nu} \quad \forall \mathbf{k} \in \Lambda, \\ m_{\mathbf{k}} &= \frac{\nu}{(\nu-1)|\mathbf{N}|} \sum_{\boldsymbol{\ell}:\, P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})>0} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\mathbf{k}} \left(\frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right)^{\nu-1} \quad \forall \mathbf{k} \in \Lambda_0. \end{aligned} \tag{19}$$

Therefore, if the dual problem (17) has an interior-point minimizer, it must be a stationary point where we have both covariance and generalized cepstral matching for the function Φ_{ν} in (15). This means that Φ_{ν} is the unique solution to Problem (9). Moreover, Φ_{ν} is rational because $\nu \in \mathbb{N}_+$. The latter observation explains why we chose $\nu \in \mathbb{N}_+$. However, the existence of such a minimizer seems rather nontrivial to prove given the shape of the feasible set $\mathscr{L}$ in (18), since in general the polynomial P is allowed to take negative values on some grid points. On the other hand, if the optimal $(\hat{P}, \hat{Q})$ is such that $\hat{P}(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0$ for all $\boldsymbol{\ell} \in \mathbb{Z}^d_{\mathbf{N}}$, then we automatically have $\hat{Q} > 0$ at all the grid points, and hence $(\hat{P}, \hat{Q})$ is in the interior of $\mathscr{L}$. Drawing inspiration from Enqvist (2004), we tackle this issue by considering the regularized dual function

$$J_{\nu,\lambda}(P,Q) := J_{\nu}(P,Q) - \frac{\lambda}{|\mathbf{N}|} \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \log P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})$$

where $\lambda > 0$ is the regularization parameter and the penalty term can be equivalently written as the integral of $-\log P(e^{i\boldsymbol{\theta}})$ on $\mathbb{T}^d$ against the discrete measure $\mathrm{d}\eta(\boldsymbol{\theta})$. Due to the presence of the logarithm, the polynomial P is forced to take positive values on all the grid points $\boldsymbol{\zeta}_{\boldsymbol{\ell}}$. As a consequence, the feasible set of the regularized problem is

$$\mathscr{L}_{\lambda} := \{(P,Q)\ :\ Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0 \text{ and } P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0\ \forall \boldsymbol{\ell} \in \mathbb{Z}^d_{\mathbf{N}}\},$$

which is significantly simplified with respect to (18). The regularized dual problem is just

$$\min_{P,Q}\ J_{\nu,\lambda}(P,Q) \quad \text{s.t. } (P,Q) \in \mathscr{L}_{\lambda}. \tag{20}$$

Let us first check the behavior of the function $J_{\nu,\lambda}$ near the boundary of $\mathscr{L}_{\lambda}$ which is composed of polynomial pairs (P,Q) such that $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$ or $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$ for some $\boldsymbol{\ell} \in \mathbb{Z}^d_{\mathbf{N}}$. Consider a sequence (P_k, Q_k) that tends to a point (P,Q) on $\partial\mathscr{L}_{\lambda}$. There are two cases.

(i) If the limit point is such that $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$ for some grid point, then, since all the summands in the first term of (16) are positive, we have (as $k \to \infty$)

$$\begin{aligned} J_{\nu,\lambda}(P_k,Q_k) &> \langle \mathbf{q}_k,\, \mathbf{c}\rangle - \langle \mathbf{p}_k,\, \mathbf{m}\rangle \\ &\quad - \frac{\lambda}{|\mathbf{N}|} \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \log P_k(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \to +\infty. \end{aligned}$$

(ii) If the limit point is such that $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0$ for all $\boldsymbol{\ell} \in \mathbb{Z}^d_{\mathbf{N}}$, then it must happen that $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$ for some grid point. It follows again that $J_{\nu,\lambda}(P_k,Q_k) \to +\infty$ because at least one term in the sum $\frac{P_k(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\nu}}{Q_k(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^{\nu-1}}$ blows up while the inner products and the sum of $\log P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})$ remain bounded.

In view of the reasoning above, for a sufficiently large real number β, any point in the sublevel set of the regularized dual function

$$J^{-1}_{\nu,\lambda}(-\infty, \beta] := \{(P,Q) \in \mathscr{L}_{\lambda}\ :\ J_{\nu,\lambda}(P,Q) \leq \beta\}$$

must be away from $\partial\mathscr{L}_{\lambda}$. It is not difficult to prove that $J_{\nu,\lambda}$ is strictly convex. Hence, if an *interior-point* solution to (20) exists, say $(\hat{P}, \hat{Q})$, then it is a stationary point of $J_{\nu,\lambda}$ which is also unique. The stationarity condition implies that the spectral density $\hat{\Phi} = (\hat{P}/\hat{Q})^{\nu}$ satisfies the covariance constraints in (19). The cepstral constraints however, are not exactly satisfied. Indeed, we have

$$m_{\mathbf{k}} + \varepsilon_{\mathbf{k}} = \frac{\nu}{(\nu-1)|\mathbf{N}|} \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\mathbf{k}} \left(\frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right)^{\nu-1} \quad \forall \mathbf{k} \in \Lambda_0, \tag{21}$$

where the error term is

$$\varepsilon_{\mathbf{k}} = \frac{\lambda}{|\mathbf{N}|} \sum_{\boldsymbol{\ell}\in\mathbb{Z}^d_{\mathbf{N}}} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\mathbf{k}} \frac{1}{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}. \tag{22}$$

In other words, $\hat{\Phi} = (\hat{P}/\hat{Q})^{\nu}$ fulfills covariance matching and approximate generalized cepstral matching with error $\varepsilon_{\mathbf{k}}$ for each index $\mathbf{k} \in \Lambda_0$. As seen from (22), the regularization parameter λ controls the cepstral matching error: the smaller λ is, the better the cepstral approximation is (at the price that the solution could be very close to $\partial\mathscr{L}_{\lambda}$ and thus difficult to compute numerically with a gradient method).

Finally, we report the following result, whose proof is omitted due to the limit space constraint.

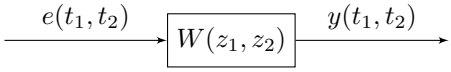
Theorem 1. Assume that there exists a function Φ_0 defined on the discrete d-torus such that $\Phi_0(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) > 0$ for all $\boldsymbol{\ell} \in \mathbb{Z}_{\mathbf{N}}^d$ and the covariances admit a representation

$$c_{\mathbf{k}} = \frac{1}{|\mathbf{N}|} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_{\mathbf{N}}^d} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\mathbf{k}} \Phi_0(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \quad \forall \mathbf{k} \in \Lambda.$$

Then, the regularized dual problem (20) admits a unique solution $(\hat{P}, \hat{Q})$ in the open set $\mathscr{L}_\lambda$ such that the function $\hat{\Phi} = (\hat{P}/\hat{Q})^\nu$ defined on the discrete grid achieves covariance matching and approximate generalized cepstral matching as detailed in (21).

4. APPLICATION IN 2-D SYSTEM IDENTIFICATION

We apply our theory to the identification of a 2-dimensional linear stochastic systems in the same fashion as that of Ringh et al. (2016, Section 7).



Consider a 2-d linear time-invariant (LTI) system described by the transfer function $W(z_1, z_2)$, which is excited by a white noise process/field $e(t_1, t_2)$ and produces an output process $y(t_1, t_2)$, see the picture above. Furthermore, let the true system W have the structure that corresponds to our optimal spectrum (14), namely

$$W(\mathbf{z}) = \left[\frac{b(\mathbf{z})}{a(\mathbf{z})}\right]^\nu = \left[\frac{\sum_{\mathbf{k}\in\Lambda_+} b_{\mathbf{k}} \mathbf{z}^{-\mathbf{k}}}{\sum_{\mathbf{k}\in\Lambda_+} a_{\mathbf{k}} \mathbf{z}^{-\mathbf{k}}}\right]^\nu$$

where $\Lambda_+ := \{(k_1, k_2) : 0 \le k_1, k_2 \le 1\}$, (z_1, z_2) is abbreviated as $\mathbf{z}$, and $\mathbf{z}^{\mathbf{k}}$ stands for $z_1^{k_1} z_2^{k_2}$. If the white noise input has unit variance, then the spectral density of the output process y is $(P/Q)^\nu$ with $P(e^{i\boldsymbol{\theta}}) = |b(e^{i\boldsymbol{\theta}})|^2$, $Q(e^{i\boldsymbol{\theta}}) = |a(e^{i\boldsymbol{\theta}})|^2$. We fix the integer $\nu = 2$. For simplicity, we also impose a separable form $(1 - \alpha_1 z_1^{-1})(1 - \alpha_2 z_2^{-1})$ on the polynomials a, b and take $|\alpha_j| < 1$, for $j = 1, 2$, so that the transfer function W is stable and minimum-phase (under the engineering convention). The system parameters can be assigned via $[a_{0,0}\ a_{0,1}\ a_{1,0}\ a_{1,1}] = [1\ -\alpha_2\ -\alpha_1\ \alpha_1\alpha_2]$. It is convenient to collect the 2-d system parameters into matrices. In our particular example, we have

$$A = \begin{bmatrix} 1 & -0.7 \\ -0.5 & 0.35 \end{bmatrix}, \quad B = \begin{bmatrix} 0.6696 & -0.5357 \\ -0.4018 & 0.3214 \end{bmatrix}, \qquad (23)$$

where $a_{k_1,k_2} = [A]_{k_1+1,k_2+1}$ and similar for b_{k_1,k_2}. Notice that the constraint $p_{\mathbf{0}} = 1$ translates into a normalization condition $\|B\|_{\mathrm{F}} = 1$, where $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm. The true spectrum Φ defined on the 2-d domain is shown in Fig. 1(left).

Next, we demonstrate how to recover the system matrices A and B from the covariances and generalized cepstral coefficients of the output process y with the indices in the set $\Lambda = \{(k_1, k_2) : -1 \le k_1, k_2 \le 1\}$. In our framework, such a procedure can be divided into three steps:

(i) Evaluate the power spectrum of the output process y on a discrete grid and compute $c_{\mathbf{k}}$'s and $m_{\mathbf{k}}$'s with $\mathbf{k} \in \Lambda$;
(ii) Solve the dual optimization problem (20), given $c_{\mathbf{k}}$'s and $m_{\mathbf{k}}$'s in Step (i) and a regularization parameter $\lambda > 0$;
(iii) Factorize the optimal spectrum $\hat{\Phi}$ in order to obtain a transfer function $\hat{W}(z_1, z_2)$.

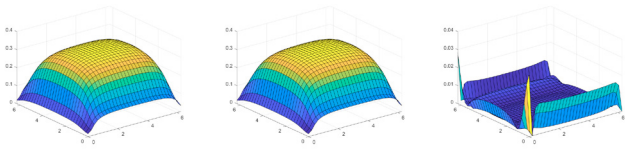


Fig. 1. *Left.* The true spectrum Φ of y. *Center.* Estimated spectrum $\hat{\Phi}$ via solving the regularized dual optimization problem with $\lambda = 10^{-10}$ and *true* covariances - generalized cepstral coefficients. *Right.* The pointwise relative error.

Since our theory developed in the previous sections builds a *rational* spectral density from data, the spectral factorization in Step (iii) reduces to that of polynomials. An immediate remark is that the factorization of a positive Laurent trigonometric polynomial into *one square* is in general impossible in the case of several variables [1] (which we call "multidimensional"). Fortunately in the 2-d case, Geronimo and Woerdeman (2004, Theorems 1.1.1 & 1.1.3) have given a sufficient and necessary condition to check such factorability and an explicit formula to compute the factors when the factorization is possible. The nontrivial part of the condition states that a certain reduced covariance matrix should have a specific low rank. Notice also that a factor $a(e^{i\boldsymbol{\theta}})$ can nonetheless be computed from a positive Laurent polynomial $P(e^{i\boldsymbol{\theta}})$ even when the aforementioned rank condition is not met. However, as discussed in Ringh et al. (2016, Section 7), such an operation results in a large difference between $P(e^{i\boldsymbol{\theta}})$ and $|a(e^{i\boldsymbol{\theta}})|^2$.

For Step (i), we compute the covariances and the generalized cepstral coefficients using the true spectrum evaluated on a 30×30 grid in the frequency domain $\mathbb{T}^2$. Due to the symmetry, we only need to evaluate $\mathbf{c} = [c_{0,0}\ c_{0,1}\ c_{1,-1}\ c_{1,0}\ c_{1,1}]$, $\mathbf{m} = [m_{0,1}\ m_{1,-1}\ m_{1,0}\ m_{1,1}]$ which are put in lexicographic ordering. Clearly, the variables $\mathbf{p}$ and $\mathbf{q}$ can be arranged in the same way so that we have 9 real variables in total. Then the regularized dual problem (20) with $\lambda = 10^{-10}$ (very weak regularization) is solved using a gradient descent algorithm initialized at $q_{00} = 1$ and the rest variables equal to 0, which corresponds to constant polynomials $P = Q \equiv 1$. The iterations terminate when the norm of the gradient is less than 10^{-4}. The optimal spectrum $\hat{\Phi}$ returned by the solver is plotted in Fig. 1(center). Let us define the pointwise relative error as $|\hat{\Phi}(e^{i\boldsymbol{\theta}}) - \Phi(e^{i\boldsymbol{\theta}})|/\Phi(e^{i\boldsymbol{\theta}})$ and plot it in Fig. 1(right). One can see that the errors are visible (up to a maximum of 3.44%) near the boundaries of the domain, i.e., the axes $\theta_1 = 0$ and $\theta_2 = 0$, while in the interior, the error plot is quite flat. As a complement, we also compute the cumulative relative error on the whole grid $\|\hat{\Phi} - \Phi\|_{\mathrm{F}}/\|\Phi\|_{\mathrm{F}} = 0.22\%$, meaning that the two spectra are practically indistinguishable.

[1] It has been shown that sum-of-squares factorization is always possible for multivariate Laurent polynomials that are strictly positive on the multi-torus (Dritschel, 2004). However, the factors in general have degrees larger than the original Laurent polynomial.

Given the optimal polynomials $\hat{P}$ and $\hat{Q}$, we proceed to run the factorization algorithm in Geronimo and Woerdeman (2004, Theorems 1.1.1 & 1.1.3) and compute factors $\hat{a}$ and $\hat{b}$. The numerical values of their coefficients are reported in the matrices

$$\hat{A} = \begin{bmatrix} 1.0029 & -0.7021 \\ -0.4955 & 0.3461 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 0.6716 & -0.5377 \\ -0.3989 & 0.3188 \end{bmatrix}.$$

After a comparison with (23), we can conclude that the system parameters are recovered with very small errors, which is an expected result since the true model coincides with the solution form of our optimization problem.

Finally, we observe that in many practical situations, the covariances $c_{\mathbf{k}}$ and the cepstral coefficients $m_{\mathbf{k}}$ are unknown, and instead we can only measure the output process $y(t_1, t_2)$ at a finite number of "time" indices. In plain words, we have to infer the transfer function W from the samples of the output, say a regularly spaced dataset $\mathcal{Y} = \{y(t_1, t_2) \,:\, 0 \leq t_1 \leq N_1 - 1,\ 0 \leq t_2 \leq N_2 - 1\}$. In that case, Step (i) of the previous procedure needs to be changed as the covariances and cepstral coefficients now must be estimated from $\mathcal{Y}$. In such a scenario, it is possible to adopt similar ideas of those ones in Zhu et al. (2021) and Stoica and Moses (2005, Chapter 2) in order to estimate $c_{\mathbf{k}}$'s and $m_{\mathbf{k}}$'s satisfying the assumptions of Theorem 1.

5. CONCLUSIONS

We proposed a general framework for the multidimensional rational covariance and cepstral extension problem. The idea is to generalize the definition of entropy and cepstral coefficient. We showed that in the circulant case this problem admits a rational solution matching the covariances while cepstral matching is approximate. Finally, we applied our theory to solve a 2-d system identification problem. Our future research direction is to characterize the solution in the non-circulant scenario. As highlighted by Karlsson et al. (2016), when $d \geq 3$, it is not even guaranteed that the regularized version of Problem (3) admits a *rational* solution. Our conjecture is that, given $d \geq 3$, it is always possible to find a $\nu \in \mathbb{N}_+$ sufficiently large for which our generalized framework provides a rational spectral density matching the $c_{\mathbf{k}}$'s while the $m_{\mathbf{k}}$'s are matched approximately.

REFERENCES

Byrnes, C.I., Gusev, S.V., and Lindquist, A. (2001). From finite covariance windows to modeling filters: A convex optimization approach. *SIAM Review*, 43(4), 645–675.

Carli, F.P., Ferrante, A., Pavon, M., and Picci, G. (2011). A maximum entropy solution of the covariance extension problem for reciprocal processes. *IEEE Transactions on Automatic Control*, 56(9), 1999–2012.

Dritschel, M.A. (2004). On factorization of trigonometric polynomials. *Integral Equations and Operator Theory*, 49(1), 11–42.

Engels, F. (2014). Target shape estimation using an automotive radar. In *Smart Mobile In-Vehicle Systems: Next Generation Advancements*, chapter 16, 271–290. Springer Science+Business Media.

Enqvist, P. (2004). 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.

Ferrante, A., Pavon, M., and Ramponi, F. (2008). Hellinger versus Kullback-Leibler multivariable spectrum approximation. *IEEE Transactions on Automatic Control*, 53(4), 954–967.

Georgiou, T.T. and Lindquist, A. (2003). Kullback–Leibler approximation of spectral density functions. *IEEE Transactions on Information Theory*, 49(11), 2910–2917.

Georgiou, T. (2006). Relative entropy and the multivariable multidimensional moment problem. *IEEE Transactions on Information Theory*, 52(3), 1052–1066.

Geronimo, J.S. and Woerdeman, H.J. (2004). Positive extensions, Fejér-Riesz factorization and autoregressive filters in two variables. *Annals of Mathematics*, 839–906.

Karlsson, J., Lindquist, A., and Ringh, A. (2016). The multidimensional moment problem with complexity constraint. *Integral Equations and Operator Theory*, 84(3), 395–418.

Lindquist, A. and Picci, G. (2013). The circulant rational covariance extension problem: The complete solution. *IEEE Transactions on Automatic Control*, 58(11), 2848–2861.

Ringh, A., Karlsson, J., and Lindquist, A. (2015). The multidimensional circulant rational covariance extension problem: Solutions and applications in image compression. In *54th Annual Conference on Decision and Control (CDC)*, 5320–5327. IEEE.

Ringh, A., Karlsson, J., and Lindquist, A. (2016). Multidimensional rational covariance extension with applications to spectral estimation and image compression. *SIAM Journal on Control and Optimization*, 54(4), 1950–1982.

Ringh, A., Karlsson, J., and Lindquist, A. (2018). Multidimensional rational covariance extension with approximate covariance matching. *SIAM Journal on Control and Optimization*, 56(2), 913–944.

Stoica, P. and Moses, R. (2005). *Spectral Analysis of Signals*. Pearson Prentice Hall, Upper Saddle River, NJ.

Tokuda, K., Kobayashi, T., and Imai, S. (1990). Generalized cepstral analysis of speech-unified approach to LPC and cepstral method. In *First International Conference on Spoken Language Processing*.

Yaglom, A.M. (1957). Some classes of random fields in n-dimensional space, related to stationary random processes. *Theory of Probability and Its Applications*, 2(3), 273–320.

Zhu, B. (2020). On the well-posedness of a parametric spectral estimation problem and its numerical solution. *IEEE Transactions on Automatic Control*, 65(3), 1089–1099.

Zhu, B., Ferrante, A., Karlsson, J., and Zorzi, M. (2019). Fusion of sensors data in automotive radar systems: A spectral estimation approach. In *IEEE 58th Conference on Decision and Control (CDC 2019)*, 5088–5093. IEEE.

Zhu, B., Ferrante, A., Karlsson, J., and Zorzi, M. (2021). M^2-spectral estimation: A relative entropy approach. *Automatica*, 125, 109404.

Zorzi, M. (2014). Rational approximations of spectral densities based on the Alpha divergence. *Math. Control Signals Systems*, 26(2), 259–278.

Zorzi, M. (2015). An interpretation of the dual problem of the THREE-like approaches. *Automatica*, 62, 87–92.