Further results on multidimensional rational covariance extension with application to texture generation

Axel Ringh, Johan Karlsson and Anders Lindquist

***Abstract*— The rational covariance extension problem is a moment problem with several important applications in systems and control as, for example, in identification, estimation, and signal analysis. Here we consider the multidimensional counterpart and present new results for the well-posedness of the problem. We apply the theory to texture generation by modeling the texture as the output of a Wiener system. The static nonlinearity in the Wiener system is assumed to be a thresholding function and we identify both the linear dynamical system and the thresholding parameter.**

I. Introduction

Moment problems with rationality constraints are ubiquitous in the areas of systems, control and signal processing. One important example is the *rational covariance extension problem*. First posed by R.E. Kalman [32] in 1981, this problem can be stated as follows: Given a finite covariance sequence $c := (c_0, \dots, c_n)$, determine all infinite extensions $c_{n+1}, c_{n+2}, \dots$ such that

$$\Phi(e^{i\theta}) = \sum_{k=-\infty}^{\infty} c_k e^{-ik\theta} \tag{1}$$

is a positive rational function of degree bounded by $2n$. The reason for calling this a covariance extension problem is that a function Φ of the form (1) can be regarded as the spectral density of a zero-mean, stationary stochastic process $\{y_t;\, t \in \mathbb{Z}\}$ with covariance lags $c_k = E[y_{t+k}\bar{y}_t]$ [41, Sections 3.2-3.3], [50, Section 1.3]. Here $^-$ denotes complex conjugation. Moreover, it is well-known that the rational covariance extension problem is equivalent to a truncated trigonometric moment problem with a certain complexity constraint on the solution [41, Section 12.5]. In fact, the problem amounts to determining all coercive spectral densities $\Phi(e^{i\theta}) = P(e^{i\theta})/Q(e^{i\theta})$ such that

$$c_k = \int_{\mathbb{T}} e^{ik\theta}\Phi(e^{i\theta})\frac{d\theta}{2\pi},$$

where $\mathbb{T} = [-\pi, \pi]$, and where P and Q are trigonometric polynomials of degree at most n.

This work was supported by the Swedish Research Council (VR) grant 2014-5870, the Swedish Foundation of Strategic Research (SSF) grant AM13-0049, and the ACCESS Linnaeus Center at KTH Royal Institute of Technology.

A. Ringh and J. Karlsson are with the Division of Optimization and Systems Theory, Department of Mathematics, and the ACCESS Linnaeus Center, KTH Royal Institute of Technology, Stockholm, Sweden. Email: aringh@kth.se, johan.karlsson@math.kth.se. A. Lindquist is with the Department of Automation and School of Mathematics, Shanghai Jiao Tong University, Shanghai, China, and the Division of Optimization and Systems Theory, Department of Mathematics, KTH Royal Institute of Technology, Stockholm, Sweden. Email: alq@kth.se.

The rational covariance extension problem was partially solved in 1983 by T.T. Georgiou [22], [23], who proved that to each positive covariance sequence c and positive numerator polynomial P there is a corresponding denominator polynomial Q such that $\Phi = P/Q$ matches the covariances and conjectured that this correspondence is unique. This conjecture was then proved in [12], where it was also established that this (complete) parameterization is smooth, i.e., a diffeomorphism.

Since then, this and similar problems has been extensively studied in the literature [7], [8], [10], [17], [18], [24], [40], [43], [44], [53], and this research has provided the stimulus for research in the general theory for scalar moment problems [9], [11], [27]. Moreover a number of multivariate counterparts, i.e., when Φ is a matrix-valued spectral density, have also been solved [5], [20], [26], [39], [45], [47]. All this work is connected to dynamical systems that depends on one variable, typically representing time. However, many problems in spectral estimation, signal processing, system identification, and image processing are inherently multidimensional [6]. Multidimensional systems theory has been applied to many different problems, for example pollution models [21], agricultural models [3], [52], texture modeling [35], and image processing [16]. Therefore, more recently, interest has also been directed towards a multidimensional version of the rational covariance extension problem [25], [26], [34], [48], [49], a problem which is linked to earlier work on maximum entropy solutions [36]–[38].

The focus of the present paper is on the multidimensional rational covariance extension problem, which can be posed as a complexity-constrained, multidimensional, truncated, trigonometric moment problem. In Section II we define the problem and present results from the literature, mainly from [49]. Section III is devoted to well-posedness of this inverse problem and contains some new results reported here for the first time. Finally, in Section IV we consider an example related to Wiener system identification and texture generation.

II. The multidimensional rational covariance extension problem

The multidimensional rational covariance extension problem is an inverse problem. To formally define it, let $\Lambda \subset \mathbb{Z}^d$ be a finite index set such that $0 \in \Lambda$ and $-\Lambda = \Lambda$, and let

$$c := [c_{\mathbf{k}} \mid \mathbf{k} := (k_1, \dots, k_d) \in \Lambda] \tag{2}$$

be a set of known covariances, which are complex numbers with the symmetry $c_{-\mathbf{k}} = \bar{c}_{\mathbf{k}}$. Also let $|\Lambda|$ denote

the cardinality of the index set Λ. The problem amounts to parametrizing a certain family of nonnegative bounded measures $d\mu$ on $\mathbb{T}^d$ such that

$$c_{\mathbf{k}} = \int_{\mathbb{T}^d} e^{i(\mathbf{k},\boldsymbol{\theta})} d\mu(\boldsymbol{\theta}) \quad \text{for all } \mathbf{k} \in \Lambda, \tag{3}$$

where $(\mathbf{k},\boldsymbol{\theta}) := \sum_{j=1}^d k_j\theta_j$. Like in the one-dimensional case, the reason for referring to this problem as a covariance extension problem is that the c in (2) can be interpreted as a set of covariance lags $c_{\mathbf{k}} := E[y_{\mathbf{t}+\mathbf{k}}\bar{y}_{\mathbf{t}}]$ of a discrete-time, zero-mean, homogeneous[1] stochastic process $\{y_{\mathbf{t}};\, \mathbf{t} \in \mathbb{Z}^d\}$.

Now, let

$$d\mu(\boldsymbol{\theta}) = \Phi(e^{i\boldsymbol{\theta}})dm(\boldsymbol{\theta}) + d\hat{\mu}(\boldsymbol{\theta}), \tag{4}$$

where Φdm is the absolutely continuous and $d\hat{\mu}$ the singular part in the Lebesgue decomposition of $d\mu$ and $dm(\boldsymbol{\theta}) := (1/2\pi)^d \prod_{j=1}^d d\theta_j$ is the (normalized) Lebesgue measure on $\mathbb{T}^d$. In general, if a solution to (3) exists, there are infinitely many measures $d\mu$ satisfying the equation. We wish to parametrize the family of measures for which the spectral density takes the form

$$\Phi(e^{i\boldsymbol{\theta}}) = \frac{P(e^{i\boldsymbol{\theta}})}{Q(e^{i\boldsymbol{\theta}})}, \quad p, q \in \bar{\mathfrak{P}}_+ \backslash \{0\}. \tag{5}$$

Here $\mathfrak{P}_+$ is the set of coefficients $p := [p_{\mathbf{k}} \mid \mathbf{k} \in \Lambda]$ corresponding to trigonometric polynomials

$$P(e^{i\boldsymbol{\theta}}) = \sum_{\mathbf{k}\in\Lambda} p_{\mathbf{k}} e^{-i(\mathbf{k},\boldsymbol{\theta})} \tag{6}$$

that are positive for all $\boldsymbol{\theta} \in \mathbb{T}^d$. The set $\mathfrak{P}_+$ is in fact a convex cone, and by $\bar{\mathfrak{P}}_+$ and $\partial\mathfrak{P}_+$ we will denote the closure and the boundary $\bar{\mathfrak{P}}_+ \backslash \mathfrak{P}_+$, respectively. It is then easily verified that $\partial\mathfrak{P}_+$ is the subset of all $p \in \bar{\mathfrak{P}}_+$ such that the corresponding nonnegative trigonometric polynomial $P(e^{i\boldsymbol{\theta}})$ is zero in at least one point. In this context we also introduce the dual cone of $\bar{\mathfrak{P}}_+$, called $\bar{\mathfrak{C}}_+$, the interior of which is given by

$$\mathfrak{C}_+ := \left\{ c \mid \langle c, p\rangle > 0, \quad \text{for all } p \in \bar{\mathfrak{P}}_+ \setminus \{0\} \right\}.$$

Here, $\langle c, p\rangle = \sum_{\mathbf{k}\in\Lambda} c_{\mathbf{k}}\bar{p}_{\mathbf{k}}$ denotes the complex inner product. The boundary $\partial\mathfrak{C}_+$ of $\mathfrak{C}_+$ consists of all $c \in \bar{\mathfrak{C}}_+$ such that $\langle c, p\rangle = 0$ for some $p \in \bar{\mathfrak{P}}_+ \setminus \{0\}$. The cone $\bar{\mathfrak{C}}_+$ plays an important role in the theory. In particular, it is easily seen that, for any c satisfying (3) for some nonnegative measure $d\mu$, we have

$$\langle c, p\rangle = \int_{\mathbb{T}^d} P(e^{i\boldsymbol{\theta}})d\mu \geq 0$$

for all $p \in \bar{\mathfrak{P}}_+$, and hence $c \in \bar{\mathfrak{C}}_+$ (cf. [34, Proposition 2.2]). In fact, the converse statement is also true [34, Theorem 2.3].

In view of the complexity constraint (5), parametrizing the rational family of solutions to (3) is a non-convex problem. However, formulating the inverse problem as an optimization problem, we can use a regularizing functional that turns out to promote rational solutions. To this end, let $p \in \bar{\mathfrak{P}}_+ \setminus \{0\}$ and consider the functional

$$\mathbb{I}_p(d\mu) = \int_{\mathbb{T}^d} P(e^{i\boldsymbol{\theta}}) \log \frac{P(e^{i\boldsymbol{\theta}})}{\Phi(e^{i\boldsymbol{\theta}})}\, dm(\boldsymbol{\theta}), \tag{7}$$

which is the Kullback-Leibler divergence between the two measures Pdm and $d\mu = \Phi dm + d\hat{\mu}$ [27], [49]. The primal problem then amounts to minimizing (7) subject to (3), i.e.,

$$\begin{aligned} &\min_{d\mu\geq 0} \quad \int_{\mathbb{T}^d} P(e^{i\boldsymbol{\theta}}) \log \frac{P(e^{i\boldsymbol{\theta}})}{\Phi(e^{i\boldsymbol{\theta}})}\, dm(\boldsymbol{\theta}), \\ &\text{subject to} \quad c_{\mathbf{k}} = \int_{\mathbb{T}^d} e^{i(\mathbf{k},\boldsymbol{\theta})} d\mu(\boldsymbol{\theta}) \quad \text{for all } \mathbf{k} \in \Lambda. \end{aligned} \tag{8}$$

From this, one can readily derive the Lagrangian dual functional obtained by relaxing the equality constraint, which takes the form

$$\mathbb{J}_p(q) = \langle c, q\rangle - \int_{\mathbb{T}^d} P(e^{i\boldsymbol{\theta}}) \log Q(e^{i\boldsymbol{\theta}})dm, \tag{9}$$

and the corresponding dual optimization problem is

$$\begin{aligned} &\min_{q} \quad \langle c, q\rangle - \int_{\mathbb{T}^d} P(e^{i\boldsymbol{\theta}}) \log Q(e^{i\boldsymbol{\theta}})dm, \\ &\text{subject to} \quad q \in \bar{\mathfrak{P}}_+. \end{aligned} \tag{10}$$

Using the primal-dual pair of optimization problems (8) and (10), we get the following (complete) characterization of the multidimensional rational covariance extension problem.

Theorem 1 ([49, Theorem 2.1]): For every $(c, p) \in \mathfrak{C}_+ \times (\bar{\mathfrak{P}}_+ \setminus \{0\})$ the functional (9) is strictly convex, and (10) has a unique minimizer $\hat{q} \in \bar{\mathfrak{P}}_+ \setminus \{0\}$. Moreover, there exists a unique $\hat{c} \in \partial\mathfrak{C}_+$ and a nonnegative singular measure $d\hat{\mu}$ with support $\operatorname{supp}(d\hat{\mu}) \subseteq \{\boldsymbol{\theta} \in \mathbb{T}^d \mid \hat{Q}(e^{i\boldsymbol{\theta}}) = 0\}$ such that

$$c_{\mathbf{k}} = \int_{\mathbb{T}^d} e^{i(\mathbf{k},\boldsymbol{\theta})} \left(\frac{P}{\hat{Q}} dm + d\hat{\mu} \right) \quad \text{for all } \mathbf{k} \in \Lambda$$

and

$$\hat{c}_{\mathbf{k}} = \int_{\mathbb{T}^d} e^{i(\mathbf{k},\boldsymbol{\theta})} d\hat{\mu}, \text{ for all } \mathbf{k} \in \Lambda.$$

For any such $d\hat{\mu}$, the measure

$$d\mu(\boldsymbol{\theta}) = (P(e^{i\boldsymbol{\theta}})/\hat{Q}(e^{i\boldsymbol{\theta}}))dm(\boldsymbol{\theta}) + d\hat{\mu}(\boldsymbol{\theta})$$

is an optimal solution to (8). Moreover, $d\hat{\mu}$ can be chosen with support in at most $|\Lambda| - 1$ points.

In view of Theorem 1, the optimality conditions for (10) can be summarized as follows.

Corollary 2 ([34, Corollary 5.3]): Let $c \in \mathfrak{C}_+$. Then $\hat{q}$ is an optimal solution to (10) if and only if

$$\begin{aligned} &\hat{q} \in \bar{\mathfrak{P}}_+, \qquad \hat{c} \in \partial\mathfrak{C}_+, \qquad \langle \hat{c}, \hat{q}\rangle = 0 \\ &c_{\mathbf{k}} = \int_{\mathbb{T}^d} e^{i(\mathbf{k},\boldsymbol{\theta})} \frac{P}{\hat{Q}} dm + \hat{c}_{\mathbf{k}} \text{ for all } \mathbf{k} \in \Lambda. \end{aligned}$$

[1] That is, $E[y_{\mathbf{t}+\mathbf{k}}\bar{y}_{\mathbf{t}}]$ is independent of $\mathbf{t}$ for all $\mathbf{k}$. Homogeneity generalizes stationarity for $d = 1$ to the multidimensional case $d > 1$.

III. Well-posedness of the problem

Since (8) is an inverse problem, we are not only interested in the existence of a solution but also in the question of whether (8) is well-posed. Theorem 1 guarantees that for $(c,p) \in \mathfrak{C}_+ \times (\bar{\mathfrak{P}}_+ \setminus \{0\})$ a unique solution $(\hat{q},\hat{c})$ exists. It remains to investigate how this solution depends on the parameters of the problem, i.e., on the tuple (c,p). To this end, from Propositions 7.3 and 7.4 in [34], we first have the following result.

Proposition 3: Let c, p and $\hat{q}$ be as in Theorem 1. Then the map $(c,p) \mapsto \hat{q}$ is continuous.

Next, consider the continuity of the map $(c,p) \mapsto \hat{c}$. For $d \leq 2$, note that $\int_{\mathbb{T}^d} Q^{-1} dm = \infty$ for all $q \in \partial\mathfrak{P}_+$. Hence, when $p \in \mathfrak{P}_+$ and thus P is strictly positive on all of $\mathbb{T}^d$, the gradient of (9) will be $-\infty$ on the boundary $\partial\mathfrak{P}_+$. Therefore, the optimal solution $\hat{q}$ is pushed into $\mathfrak{P}_+$, and in view of Theorem 1 we can thus guarantee that $\hat{c} = 0$.

Proposition 4 ([49, Corollary 2.3]): Suppose that $d \leq 2$. Then, for any $c \in \mathfrak{C}_+$ and $p \in \mathfrak{P}_+$ there exists a $q \in \mathfrak{P}_+$ such that $d\mu = (P/Q)dm$ satisfies (3). Moreover this q is the unique solution to (10).

Hence, since $\hat{c} = 0$ for $d \leq 2$ and $p \in \mathfrak{P}_+$, the continuity of $\mathfrak{C}_+ \times \mathfrak{P}_+ \ni (c,p) \mapsto \hat{c}$ is trivial in this case. For later reference we formulate this result in the following proposition.

Proposition 5: Let c, p, $\hat{q}$ and $\hat{c}$ be as in Theorem 1. Then, for $d \leq 2$ and all $(c,p) \in \mathfrak{C}_+ \times \mathfrak{P}_+$, the mapping $(c,p) \to (\hat{q},\hat{c})$ is continuous.

Remark 6: In the case $d \leq 2$ in turns out that the result of Proposition 3 can be strengthened. In fact, for a fixed $p \in \mathfrak{P}_+$, the map $\mathfrak{C}_+ \ni c \mapsto \hat{q} \in \mathfrak{Q}_+ \subset \mathfrak{P}_+$ is a diffeomorphism onto its image $\mathfrak{Q}_+$ [49, Theorem 4.4]. Moreover, for fixed $c \in \mathfrak{C}_+$ the map $\mathfrak{P}_+ \ni p \mapsto \hat{q} \in \mathfrak{Q}_+$ is also a diffeomorphism [49, Theorem 4.5] .

However, the case $d \geq 3$ turns out to be trickier, since we can have $\hat{c} \neq 0$ although $p \in \mathfrak{P}_+$. In the following subsection we will investigate the continuity of the map $(c,p) \mapsto \hat{c}$ for the case $d \geq 3$ under certain conditions.

A. New results on well-posedness

In view of Propositions 3 and 5, the problem (8) is well-posed for $d \leq 2$ and $p \in \mathfrak{P}_+$. However, the well-posedness in $\hat{c}$ is in some sense trivial, since in this case we have $\hat{c} = 0$ for all $p \in \mathfrak{P}_+$ and $c \in \mathfrak{C}_+$. For $d \geq 3$ this is not the case, as we can have $\hat{q} \in \partial\mathfrak{P}_+$ even when $p \in \mathfrak{P}_+$ and thus might have $\hat{c} \neq 0$. In this subsection we extend the well-posedness result of Proposition 5 to hold also in some cases where $d \geq 3$.

However, we begin by noting that the condition $p \in \mathfrak{P}_+$ in general cannot be weakened to $p \in \bar{\mathfrak{P}}_+ \setminus \{0\}$, which the following one-dimensional example illustrates (cf. [34, Example 3.8]).

Example 7: Let

$$c = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \int_{-\pi}^{\pi} \begin{bmatrix} e^{-i\theta} \\ 1 \\ e^{i\theta} \end{bmatrix} (2dm + d\nu_0),$$

where $dm = d\theta/2\pi$, and $d\nu_0$ is the singular measure $\delta_0(\theta)d\theta$ with support in $\theta = 0$. Since $d\mu := 2dm + d\nu_0$ is positive, $c \in \bar{\mathfrak{C}}_+$. Moreover, since the Toeplitz matrix $T_c := [c_{j-\ell}]_{j,\ell}$ is positive definite, i.e.,

$$T_c = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} > 0,$$

we have $c \in \mathfrak{C}_+$ (see, e.g., [40, p. 2853]). Thus we know that for each $p \in \mathfrak{P}_+$ we have a unique $\hat{q} \in \mathfrak{P}_+$ such that $P/\hat{Q}$ matches c, and hence $\hat{c} = 0$ (Proposition 4). However, for $p = 2(-1,2,-1)^T$ we have that $\hat{q} = (-1,2,-1)^T$ and $\hat{c} = (1,1,1)^T$ (Corollary 2). Then, for the sequence (p_k), where $p_k = 2(-1, 2+1/k, -1)^T \in \mathfrak{P}_+$, we have $\hat{c}_k = 0$, so

$$\lim_{k\to\infty} \hat{c}_k = \lim_{k\to\infty} \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T \neq \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T,$$

which shows that the mapping $p \mapsto \hat{c}$ is not continuous.

One way to try to establish continuity of the map $(c,p) \mapsto \hat{c}$ is to try to use the already established continuity from (c,p) to $\hat{q}$ in Proposition 3. From the KKT conditions in Corollary 2 we have in particular that

$$c_{\mathbf{k}} = \int_{\mathbb{T}^d} e^{i(\mathbf{k},\boldsymbol{\theta})} \frac{P}{\hat{Q}} dm + \hat{c}_{\mathbf{k}} \text{ for all } \mathbf{k} \in \Lambda,$$

and hence continuity of $\hat{c}$ would follow if $\int_{\mathbb{T}^d} P\hat{Q}^{-1} dm$ is continuous in $(c,p,\hat{q})$, whenever the integral is finite. If $p \in \mathfrak{P}_+$, this follows from the continuity of the map $\hat{q} \mapsto \hat{Q}^{-1}$ in $L_1(\mathbb{T}^d)$. Note here that we know from Example 7 that the condition $p \in \mathfrak{P}_+$ is actually needed. In fact, if $p \in \partial\mathfrak{P}_+$ we may have "pole-zero cancellations" in $P/\hat{Q}$ (cf. [34, Example 5.10]), and then $\int_{\mathbb{T}^d} P/\hat{Q} dm$ may be finite even if $\hat{Q}^{-1} \notin L_1(\mathbb{T}^d)$.

For the case $d \leq 2$, the continuity of the map $\hat{q} \mapsto \hat{Q}^{-1}$ in $L_1(\mathbb{T}^d)$ is trivial, since, if $\int_{\mathbb{T}^d} \hat{Q}^{-1} dm$ is finite, then $\hat{q} \in \mathfrak{P}_+$ and $\hat{Q}$ is bounded away from zero (cf. Propositions 4 and 5). However, for the case $d \geq 3$ the optimal $\hat{q}$ may belong to the boundary $\partial\mathfrak{P}_+$, i.e., $\hat{Q}$ is zero in some point. Here we show L_1 continuity of $\hat{q} \mapsto \hat{Q}^{-1}$ for certain cases. The proof is deferred to the appendix.

Proposition 8: For $d \geq 3$, let $\hat{q} \in \bar{\mathfrak{P}}_+$ and suppose that the Hessian $\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}} \hat{Q}$ is positive definite in each point where $\hat{Q}$ is zero. Then $\hat{Q}^{-1} \in L_1(\mathbb{T}^d)$ and the mapping from the coefficient vector $q \in \bar{\mathfrak{P}}_+$ to Q^{-1} is L_1 continuous in the point $\hat{q}$.

From Propositions 8 and 5 the following continuity result follows directly.

Corollary 9: Let c, p, $\hat{q}$, and $\hat{c}$ be as in Theorem 1. For all $c \in \mathfrak{C}_+$ and all $p \in \mathfrak{P}_+$, the mapping $(c,p) \to (\hat{q},\hat{c})$ is continuous in any point (c,p) in which $\hat{Q}$ is strictly positive or in which the Hessian $\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}} \hat{Q}$ is positive definite in each point where $\hat{Q}$ is zero.

IV. Example in texture generation

Wiener systems form a class of nonlinear dynamical systems that consist of a linear dynamic part composed with a static nonlinearity, as in Figure 1. They belong to a class of so called block-oriented systems, which has a long history

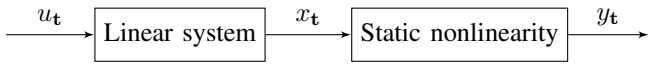


Fig. 1. A Wiener system with thresholding as static nonlinearity.

[4], and applications are found in many areas of science and engineering [2]. A lot of research has been done in the area of identification of Wiener systems, see, e.g, [31] and references therein, and the area is still active [1], [42], [51].

In this example we shall use Wiener systems to model and generate textures. The idea of using dynamical systems for modeling of textures and images is not new and has been considered in, e.g., [13], [45]. The setup we present here is motivated by [19], where thresholded Gaussian random fields are used to model porous materials for design of surface structures in pharmaceutical film coatings.

To this end, we let $\{x_{\mathbf{t}};\, \mathbf{t} \in \mathbb{Z}^d\}$ be the stationary output of a linear system with Gaussian white noise input $\{u_{\mathbf{t}};\, \mathbf{t} \in \mathbb{Z}^d\}$, and let $y_{\mathbf{t}} = f(x_{\mathbf{t}})$ where f is the static nonlinearity

$$f(x) = \begin{cases} 1 & x > \tau \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

where the thresholding parameter τ is assumed to be unknown. We assume that $u_{\mathbf{t}}$ is a zero-mean process, and hence $x_{\mathbf{t}}$ is also a zero-mean Gaussian process, which we assume to be normalized $c_{\mathbf{0}} := E[x_{\mathbf{t}}^2] = 1$. Due to these assumptions, the output $y_{\mathbf{t}}$ of the static nonlinearity has mean

$$E[y_{\mathbf{t}}] = P(y_{\mathbf{t}} = 1) = 1 - P(x_{\mathbf{t}} \leq \tau) = 1 - \phi(\tau), \tag{12}$$

where $\phi(\tau)$ is the Gaussian cumulative distribution function

$$\phi(\tau) = \int_{-\infty}^{\tau} \frac{1}{\sqrt{2\pi}} \exp(-s^2/2) ds.$$

A. Covariances of thresholded Gaussian variables

Next we consider the relation between the covariances of the input $x_{\mathbf{t}}$ and those of the output $y_{\mathbf{t}}$, respectively, and use this to estimate the covariances of the process $x_{\mathbf{t}}$ [2], [46].

To this end, let x_1, $x_2 \in N(0,1)$ be two jointly Gaussian stochastic variables and set $y_\ell = f(x_\ell)$, for $\ell = 1, 2$, where $f : \mathbb{R} \to \mathbb{R}$ is a given function. In addition, let ρ and r be the covariances $\rho := E[x_1 x_2]$ and $r := E[y_1 y_2] - E[y_1]E[y_2]$, respectively. We are interested in the relation between ρ and r, and to this end we introduce $R := E[y_1 y_2]$. Now note that R is related to the covariance ρ via [46, Equation 21] (see also [2, p. 32]), i.e.,

$$\frac{\partial R}{\partial \rho} = \int_{\mathbb{R}^2} \frac{\exp\left(-\frac{x_1^2 + x_2^2 - 2\rho x_1 x_2}{2(1-\rho^2)}\right)}{2\pi\sqrt{1-\rho^2}} f'(x_1) f'(x_2) dx_1 dx_2.$$

In our case $f(x)$ is given by (11), and thus $f'(x) = \delta_\tau(x)$ is a Dirac delta function at τ. Therefore

$$\frac{\partial R}{\partial \rho} = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{\tau^2}{1+\rho}\right),$$

and from this it follows that

$$R(\rho) = b + \int_0^\rho \frac{1}{2\pi\sqrt{1-s^2}} \exp\left(-\frac{\tau^2}{1+s}\right) ds,$$

for some constant b. In order to determine b, first note that $\rho = 0$ implies that x_1 and x_2 are uncorrelated, and hence independent, since the joint distribution is Gaussian. This in turn means that y_1 and y_2 are independent, since f is a static function, and hence we get

$$b = R(0) = E[y_1 y_2] = E[y_1]E[y_2].$$

Therefore r can be expressed as

$$\begin{aligned} r &= R(\rho) - E[y_1]E[y_2] \\ &= \int_0^\rho \frac{1}{2\pi\sqrt{1-s^2}} \exp\left(-\frac{\tau^2}{1+s}\right) ds. \end{aligned} \tag{13}$$

The integrand is well-defined for $-1 < \rho < 1$, and the integral converges for all values in the closed interval $[-1, 1]$. Moreover, the integrand is strictly positive on $(-1, 1)$ and by the inverse function theorem this transformation is invertible.

B. Estimating the linear part of the Wiener system

By using the inverse of (13) we can estimate the covariances $c_{\mathbf{k}} := E[x_{\mathbf{t}+\mathbf{k}} x_{\mathbf{t}}]$ from estimates of the covariances $r_{\mathbf{k}} := E[y_{\mathbf{t}+\mathbf{k}} y_{\mathbf{t}}] - E[y_{\mathbf{t}+\mathbf{k}}]E[y_{\mathbf{t}}]$. Note however that (13) depends on the threshold parameter τ, which is assumed to be unknown. In order to estimate τ we use (12), which gives $\tau_{\text{est}} = \phi^{-1}(1 - E[y_{\mathbf{t}}])$. Having estimates of the covariances $c_{\mathbf{k}}$, we can now appeal to Theorem 1 in order to estimate a rational spectrum for $x_{\mathbf{t}}$.

Given this rational spectral density we want to recover a linear dynamical system corresponding to the spectrum. In the one-dimensional case, $d = 1$, this is always possible by spectral factorization, since the spectral density can be written as a sum-of-one-square

$$\Phi(e^{i\theta}) = \frac{P(e^{i\theta})}{Q(e^{i\theta})} = \frac{|b(e^{i\theta})|^2}{|a(e^{i\theta})|^2}.$$

However in higher dimensions this is in general not possible [15]. For a strictly positive spectrum, a factorization as a sum-of-several-squares is always possible [14]

$$\Phi(e^{i\boldsymbol{\theta}}) = \frac{P(e^{i\boldsymbol{\theta}})}{Q(e^{i\boldsymbol{\theta}})} = \frac{\sum_{k=1}^{\ell} |b_k(e^{i\boldsymbol{\theta}})|^2}{\sum_{k=1}^{m} |a_k(e^{i\boldsymbol{\theta}})|^2}.$$

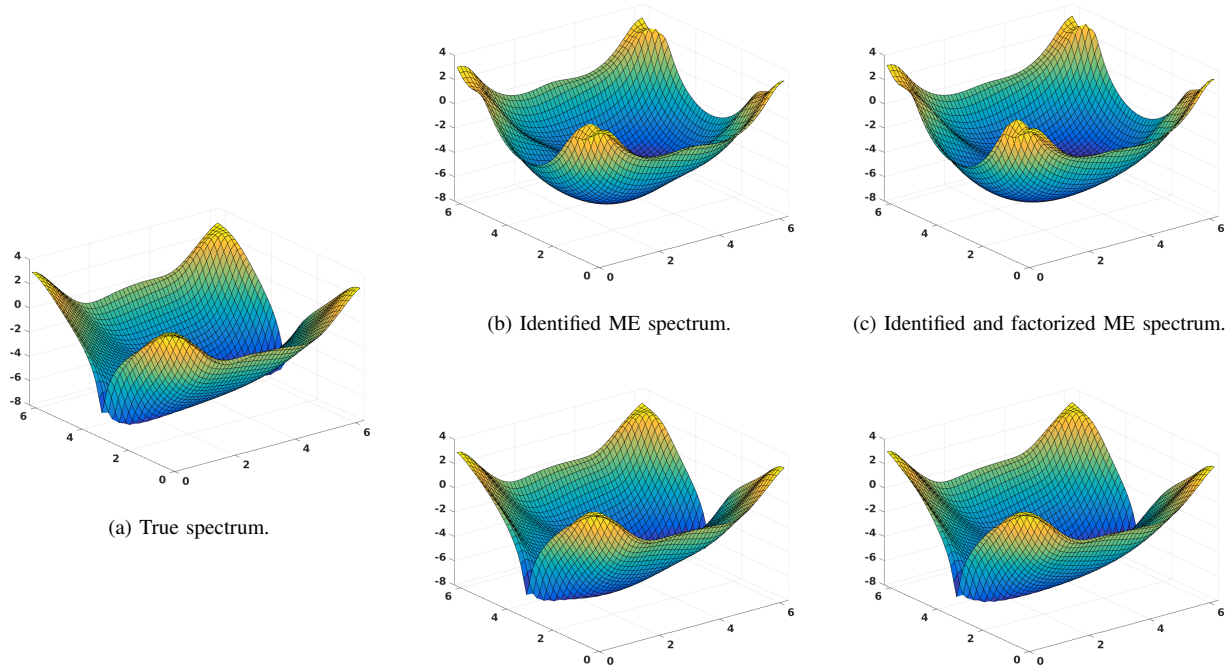
However, for $m > 1$ the interpretation of this in terms of a dynamical system is not clear to the authors. We therefore resort to a heuristic and apply the factorization procedure in [28, Theorem 1.1.1] although some of the conditions required to ensure the existence of a spectral factor may not be met (cf. [49, Section 7]).

The complete procedure for identifying the Wiener system with thresholding as static nonlinearity is summarized in Algorithm 1.

C. Simulation results

Next we test the procedure outlined above on simulated data. To this end, we consider the two-dimensional recursive filter with transfer function given by

$$\frac{b(e^{i\theta_1}, e^{i\theta_2})}{a(e^{i\theta_1}, e^{i\theta_2})} = \frac{\sum_{\mathbf{k}\in\Lambda_+} b_{\mathbf{k}} e^{-i(\mathbf{k},\boldsymbol{\theta})}}{\sum_{\mathbf{k}\in\Lambda_+} a_{\mathbf{k}} e^{-i(\mathbf{k},\boldsymbol{\theta})}},$$



(a) True spectrum.

(b) Identified ME spectrum.

(c) Identified and factorized ME spectrum.

(d) Identified spectrum with true p.

(e) Identified and factorized spectrum with true p.

Fig. 2. Log-plot of the true spectrum and the identified spectra, both before and after factorization.

Algorithm 1

Input: $(y_{\mathbf{t}})$

1: Estimate threshold parameter: $\tau_{\text{est}} = \phi^{-1}(1 - E[y_{\mathbf{t}}])$
2: Estimate covariances: $r_{\mathbf{k}} := E[y_{\mathbf{t}+\mathbf{k}} y_{\mathbf{t}}] - E[y_{\mathbf{t}+\mathbf{k}}]E[y_{\mathbf{t}}]$
3: Compute covariances $c_{\mathbf{k}} := E[x_{\mathbf{t}+\mathbf{k}} x_{\mathbf{t}}]$ by using (13)
4: Estimate a rational spectrum using Theorem 1
5: Apply the factorization procedure in [28, Theorem 1.1.1]

Output: τ_{est}, coefficients for the linear dynamical system

where $\Lambda_+ = \{(k_1, k_2) \in \mathbb{Z}^2 \mid 0 \leq k_1 \leq 2, 0 \leq k_2 \leq 2\}$ and the coefficients are given by $b_{(k_1,k_2)} = B_{k_1+1,k_2+1}$ and $a_{(k_1,k_2)} = A_{k_1+1,k_2+1}$, where

$$B = \begin{bmatrix} 0.75 & -0.2 & 0.05 \\ 0.2 & 0.3 & 0.05 \\ -0.05 & -0.05 & 0.1 \end{bmatrix}, \; A = \begin{bmatrix} 3.6623 & -4.0222 & 0.9987 \\ -4.0939 & 4.8705 & -1.1913 \\ 1.2018 & -1.3539 & 0.2155 \end{bmatrix}.$$

The threshold parameter in (11) is set to $\tau = 0.06$.

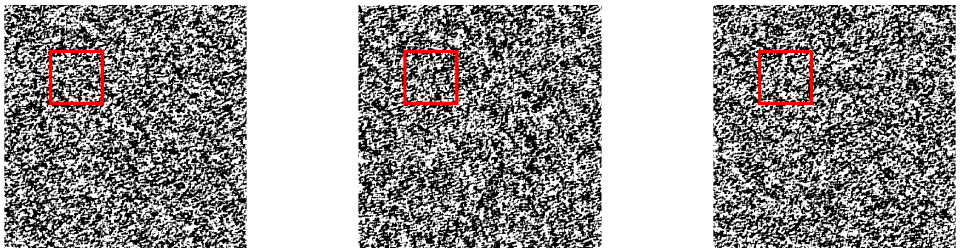
The system is simulated with Gaussian white noise as input, and 500×500 samples are taken as output. These samples are used to estimate the threshold parameter, which gives the estimate $\tau_{\text{est}} = 0.0570$. Moreover, they are used to estimate covaraiances $r_{\mathbf{k}}$ on a grid $\Lambda = \{(k_1, k_2) \in \mathbb{Z}^2 \mid |k_1| \leq 3, |k_2| \leq 3\}$. Note that this grid Λ does not agree with the true degree of the linear system. From the estimated covariances $(r_{\mathbf{k}})$ we determine the covariances $(c_{\mathbf{k}})$, which are then used in the optimization problem (10). We compute the solution with two different P, the first one being $P \equiv 1$, which corresponds to the maximum entropy (ME) solution, and the second one being $P = P_{\text{true}}$, i.e., the trigonometric polynomial corresponding to the filter b. The optimization problems are solved using the CVX toolbox in Matlab [29], [30]. The corresponding spectra obtained are shown in Figure 2. As can be seen in the figure, using the true P gives a better agreement with the true spectrum, shown in Figure 2a, which indicates that an appropriate tuning of p can improve the fit. Although there are methods in the literature on how to do simultaneously estimation of p and q [7], [17], [33], [48], [49], the question on how to best select p is still open.

After estimating the spectra, we compute estimates of filter coefficients for the autoregressive part of the linear system, and the corresponding estimates are

$$A_{\text{ME}} = \begin{bmatrix} 4.1270 & -3.8799 & 0.3572 & 0.2297 \\ -5.4210 & 4.0752 & 0.4412 & -0.2174 \\ 2.4057 & -0.0926 & -1.7157 & 0.1816 \\ -0.4199 & -0.6931 & 0.9018 & -0.1010 \end{bmatrix}$$

$$A_{\text{True } P} = \begin{bmatrix} 3.7207 & -4.3079 & 1.3210 & -0.0861 \\ -4.2527 & 5.4070 & -1.6585 & 0.0364 \\ 1.3381 & -1.6108 & 0.1836 & 0.2351 \\ -0.0562 & 0.0019 & 0.2183 & -0.2145 \end{bmatrix}$$

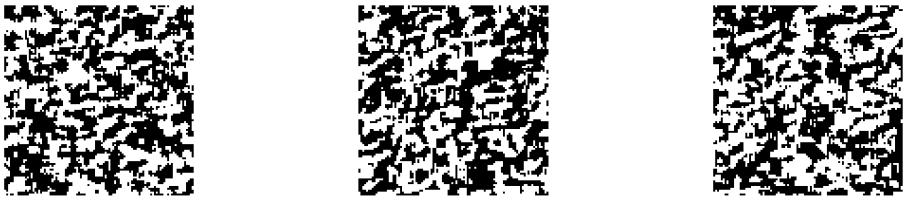
Using these filter coefficients, together with the corresponding filter coefficients for the moving-average part, we simulate the estimated Wiener system. The corresponding generated textures are shown in Figure 3. Visually, the generated textures seem to have similar structures. However, by comparing the covariances, which are shown in Figure 4, it can be seen that the texture generated by the filter obtained using the true p matches the higher order covariances considerably better.



(a) Output of true system.

(b) Output of identified maximum entropy system.

(c) Output of identified system with true p.

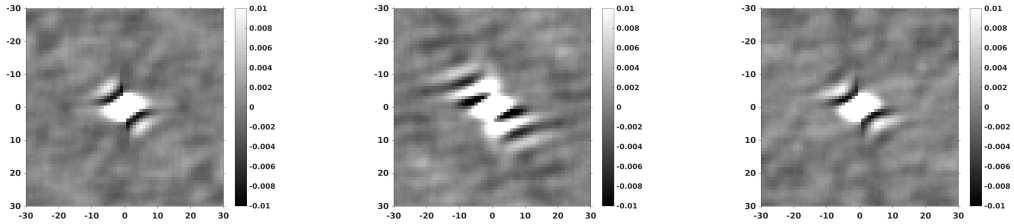


(d) Close-up of Figure 3a.

(e) Close-up of Figure 3b.

(f) Close-up of Figure 3c.

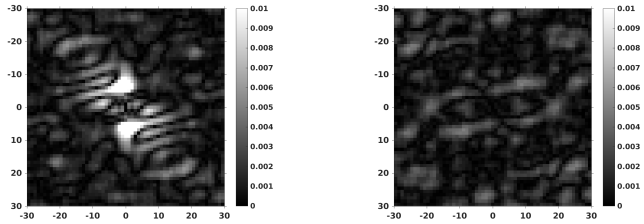
Fig. 3. Output of the true and the identified systems. Figures 3a - 3c show 500×500 samples, and Figures 3d - 3f show 100×100 samples.



(a) Covariances $r_{\mathbf{k}}$ of texture in Figure 3a.

(b) Covariances $r_{\mathbf{k}}^{\mathrm{ME}}$ of texture in Figure 3b.

(c) Covariances $r_{\mathbf{k}}^{\mathrm{True}\ P}$ of texture in Figure 3c.



(d) Absolute error $|r_{\mathbf{k}} - r_{\mathbf{k}}^{\mathrm{ME}}|$.

(e) Absolute error $|r_{\mathbf{k}} - r_{\mathbf{k}}^{\mathrm{True}\ P}|$.

Fig. 4. Covariances and covariance errors for the textures. Here $\mathbf{k} = (k_1, k_2)$ where the x-axis corresponds to k_1 and the y-axis corresponds to k_2.

V. Conclusion and Future Work

In this paper we continue our work on the multidimensional rational covariance extension initiated in [49]. We develop theory for identification of Wiener systems with applications to texture generation, and present new results on the well-posedness of the problem in dimension $d \geq 3$. However, a complete such characterization of well-posedness is still missing (see, e.g., Example 12 in the Appendix).

Another remaining issue is that spectral factorization typically is not possible for multidimensional spectral densities, and therefore we have resorted to an heuristic approach to approximate factorization. An alternative framework that avoids this problem is to model the texture as an output of an one-dimensional vector valued process as in [13] and [45]. This framework is not symmetric with respect to the coordinate-axes since it assumes stationarity only in one direction, a feature which may or may not be desirable depending on the application at hand.

Appendix

Proof of Proposition 8: The proof will be carried out using a sequence of lemmas. First we bound the integral of

Q^{-1} over the ball $B_\rho(\boldsymbol{\theta}_0) := \{\boldsymbol{\theta} \in \mathbb{T}^d \mid \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|_2 \leq \rho\}$, where the bound only depends on the Hessian of Q.

Lemma 10: Let $d \geq 3$ and $q \in \bar{\mathfrak{P}}_+$. If the Hessian $\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}} Q(e^{i\boldsymbol{\theta}}) \geq \gamma I > 0$, for $\boldsymbol{\theta} \in B_\rho(\boldsymbol{\theta}_0)$, then

$$\int_{B_\rho(\boldsymbol{\theta}_0)} Q(e^{i\boldsymbol{\theta}})^{-1} dm(\boldsymbol{\theta}) \leq \rho^{d-2}/\gamma.$$

Proof: By integrating the inequality twice, we see that $Q(e^{i\boldsymbol{\theta}}) \geq \gamma\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|^2/2$ for $\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta} \in B_\rho(\boldsymbol{\theta}_0)} Q(e^{i\boldsymbol{\theta}})$. By radial symmetry, the integral of $(\gamma\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|^2/2)^{-1}$ is in turn bounded by the integral of $(\gamma\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|^2/2)^{-1}$. Therefore,

$$\int_{B_\rho(\boldsymbol{\theta}_0)} Q(e^{i\boldsymbol{\theta}})^{-1} dm(\boldsymbol{\theta}) \leq \int_{B_\rho(\boldsymbol{\theta}_0)} \frac{2}{\gamma\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|^2} dm(\boldsymbol{\theta}) \leq \frac{\rho^{d-2}}{\gamma},$$

using basic approximations in spherical coordinates. ■

Secondly we show that if q satisfies the condition in Lemma 10, then any polynomial sufficiently close to q is well-behaved.

Lemma 11: Let $d \geq 3$, $q \in \bar{\mathfrak{P}}_+$, and assume that the Hessian $\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}} Q(e^{i\boldsymbol{\theta}})$ is positive definite in the zero $\boldsymbol{\theta}_0$ of Q. Further, let $q_k \in \bar{\mathfrak{P}}_+$ for $k \in \mathbb{N}$ such that $q_k \to q$ as $k \to \infty$. Then for any $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ and a $\rho > 0$ such that

$$\int_{B_\rho(\boldsymbol{\theta}_0)} Q(e^{i\boldsymbol{\theta}})^{-1} dm(\boldsymbol{\theta}) \leq \epsilon \quad \text{and}$$
$$\int_{B_\rho(\boldsymbol{\theta}_0)} Q_k(e^{i\boldsymbol{\theta}})^{-1} dm(\boldsymbol{\theta}) \leq \epsilon$$

for all $k \geq N$.

Proof: Let γ be such that $\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}} Q(e^{i\boldsymbol{\theta}})|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \geq 3\gamma I$, and let $\rho_1 > 0$ be such that $\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}} Q(e^{i\boldsymbol{\theta}}) \geq 2\gamma I$ for $\boldsymbol{\theta} \in B_{\rho_1}(\boldsymbol{\theta}_0)$. This is always possible since Q is C^∞ and hence the second derivatives are continuous. Next, let $\rho = \min(\rho_1, (\epsilon\gamma)^{1/(d-2)})$ and select an N such that $\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}} Q_k(e^{i\boldsymbol{\theta}}) \geq \gamma I$ for $\boldsymbol{\theta} \in B_\rho(\boldsymbol{\theta}_0)$ holds for all $k \geq N$. Such an N exists since $\nabla_{\boldsymbol{\theta}\boldsymbol{\theta}} Q_k(e^{i\boldsymbol{\theta}}) \to \nabla_{\boldsymbol{\theta}\boldsymbol{\theta}} Q(e^{i\boldsymbol{\theta}}) \geq 2\gamma I$ on $B_\rho(\boldsymbol{\theta}_0)$. From Lemma 10 it then follows that

$$\int_{B_\rho(\boldsymbol{\theta}_0)} Q(e^{i\boldsymbol{\theta}})^{-1} dm(\boldsymbol{\theta}) \leq \rho^{d-2}/\gamma \leq \epsilon$$
$$\int_{B_\rho(\boldsymbol{\theta}_0)} Q_k(e^{i\boldsymbol{\theta}})^{-1} dm(\boldsymbol{\theta}) \leq \rho^{d-2}/\gamma \leq \epsilon,$$

which proves the lemma. ■

Next, continuing the proof of Proposition 8, we use the fact that the integrals of $\hat{Q}^{-1}$ and Q_k^{-1} in a neighborhood of the zero set of $\hat{Q}$ can be made arbitrary small. Further, the convergence is uniform on the complement of this set, and hence convergence of the integrals will follow. To this end, let the sequence $(q_k) \subset \bar{\mathfrak{P}}_+$ converge to $\hat{q}$. For any $\epsilon > 0$ we need to show that there is an $N \in \mathbb{N}$ such that $\|Q_k^{-1} - \hat{Q}^{-1}\|_1 < \epsilon$ for all $k > N$. Also note that $\hat{Q}$ has finitely many zeros. To see this, assume that this is not so. By compactness of $\mathbb{T}^d$ the zeros have an accumulation point. However, this is contradicted by the fact that the Hessian is positive definite in each zero of $\hat{Q}$ and hence any zero of $\hat{Q}$ is isolated. Using Lemma 11 there is a $\rho > 0$ and an $N_1 \in \mathbb{N}$ such that

$$\int_{\cup_\ell B_\rho(\boldsymbol{\theta}_\ell)} \hat{Q}(e^{i\boldsymbol{\theta}})^{-1} dm(\boldsymbol{\theta}) \leq \epsilon/3 \quad \text{and}$$
$$\int_{\cup_\ell B_\rho(\boldsymbol{\theta}_\ell)} Q_k(e^{i\boldsymbol{\theta}})^{-1} dm(\boldsymbol{\theta}) \leq \epsilon/3$$

for all $k > N_1$. Since $Q_k \to \hat{Q}$ uniformly and $\hat{Q} > 0$ on $\mathbb{T}^d \setminus \cup_\ell B_\rho(\boldsymbol{\theta}_\ell)$, there is an N_2 such that $\|\hat{Q}^{-1} - Q_k^{-1}\|_{L_1(\mathbb{T}^d \setminus \cup_\ell B_\rho(\boldsymbol{\theta}_\ell))} < \epsilon/3$ for all $k > N_2$. The result now follows since, for $k > N := \max(N_1, N_2)$, we have

$$\begin{aligned}\|Q_k^{-1} - \hat{Q}^{-1}\|_1 &\leq \|Q_k^{-1} - \hat{Q}^{-1}\|_{L_1(\mathbb{T}^d \setminus \cup_\ell B_\rho(\boldsymbol{\theta}_\ell))} \\ &\quad + \|\hat{Q}^{-1}\|_{L_1(\cup_\ell B_\rho(\boldsymbol{\theta}_\ell))} + \|Q_k^{-1}\|_{L_1(\cup_\ell B_\rho(\boldsymbol{\theta}_\ell))} \\ &\leq \epsilon,\end{aligned}$$

which shows the continuity in the point $\hat{q}$. ■

Finally, we note that there are cases, even for $d = 3$, where the conditions in Corollary 9 are not satisfied and hence the corollary does not apply. In these cases the question of well-posedness is still open. The following is an example of this.

Example 12: Let $d = 3$, $Q \in \bar{\mathfrak{P}}_+$, and let the integer $n \geq 2$. Then, assuming that

$$Q(e^{i\boldsymbol{\theta}}) \geq \theta_1^2 + \theta_2^2 + \theta_3^{2n},$$

we have $\|Q^{-1}\|_{L_1(\mathbb{T}^3)} < \infty$. To see this, first note that Q^{-1} is unbounded only at the origin. Therefore $\|Q^{-1}\|_{L_1(\mathbb{T}^3)} < \infty$ if $\|Q^{-1}\|_{L_1(B_\rho)} < \infty$ for some $1 > \rho > 0$, where $B_\rho := B_\rho(0)$. A variable change into spherical coordinates[2] gives

$$\begin{aligned}\int_{B_\rho} Q(e^{i\boldsymbol{\theta}})^{-1} dm(\boldsymbol{\theta}) &\leq \int_{B_\rho} (\theta_1^2 + \theta_2^2 + \theta_3^{2n})^{-1} dm(\boldsymbol{\theta}) \\ &= \int_{-\pi/2}^{\pi/2} \int_0^\rho \frac{|\sin(\varphi_1)| dr d\varphi_1/(2\pi)^2}{\sin(\varphi_1)^2 + r^{2(n-1)}\cos(\varphi_1)^{2n}}.\end{aligned}$$

Note that the integrand is uniformly bounded outside the set $|\varphi_1| < \varepsilon$ for any $1 > \varepsilon > 0$. However due to symmetry it is enough to consider the set $\mathcal{S} = \{(r, \varphi_1) \mid 0 \leq \varphi_1 \leq \epsilon,\ r \in [0, \rho]\}$. Moreover, inside $\mathcal{S}$ we have $\alpha_1\varphi_1 \leq \sin(\varphi_1) \leq \alpha_2\varphi_1$ and $\cos(\varphi_1) \geq \alpha_3$ for some positive constants $\alpha_1, \alpha_2, \alpha_3$. Therefore $\|Q^{-1}\|_{L_1(B_\rho)}$ is finite if the following integral is finite:

$$\begin{aligned}&\int_0^\rho \left(\int_0^\epsilon \frac{\varphi_1 d\varphi_1}{\varphi_1^2 + r^{2(n-1)}} \right) dr = \int_0^\rho \left(\frac{1}{2} \log\left(1 + \frac{\epsilon^2}{r^{2(n-1)}}\right) \right) dr \\ &\leq \int_0^\rho \left(\frac{1}{2} \log\left(\frac{2}{r^{2(n-1)}}\right) \right) dr \\ &= \int_0^\rho \left(\frac{1}{2} \left(\log 2 - 2(n-1)\log r\right) \right) dr < \infty.\end{aligned}$$

This shows that the integral of Q^{-1} over $\mathbb{T}^3$ is finite.

References

[1] M.R. Abdalmoaty and H. Hjalmarsson. A simulated maximum likelihood method for estimation of stochastic Wiener systems. In *IEEE 55th Conf. Decision and Control*, pp. 3060–3065. IEEE, 2016.

[2] J.S. Bendat. *Nonlinear systems techniques and applications*. Wiley, 1998.

[2] Let $(\theta_1, \theta_2, \theta_3) = (r\sin(\varphi_1)\cos(\varphi_2),\ r\sin(\varphi_1)\sin(\varphi_2),\ r\cos(\varphi_1))$ and hence $dm = (r^2|\sin(\varphi_1)|/(2\pi)^3) dr d\varphi_1 d\varphi_2$.

[3] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *J. Royal Statistical Society. Series B*, pp. 192–236, 1974.

[4] S.A. Billings. Identification of nonlinear systems-a survey. In *IEE Proc. D-Control Theory and Appl.*, vol. 127, pp. 272–285. IET, 1980.

[5] A. Blomqvist, A. Lindquist, and R. Nagamune. Matrix-valued Nevanlinna-Pick interpolation with complexity constraint: an optimization approach. *IEEE Trans. Autom. Contr.*, 48(12):2172–2190, 2003.

[6] N.K. Bose. *Multidimensional Systems Theory and Applications*. Kluwer Academic Publishers, second edition, 2003.

[7] C.I. Byrnes, P. Enqvist, and A. Lindquist. Identifiability and well-posedness of shaping-filter parameterizations: A global analysis approach. *SIAM J. Control and Optimization*, 41(1):23–59, 2002.

[8] C.I. Byrnes, T.T. Georgiou, and A. Lindquist. A new approach to spectral estimation: a tunable high-resolution spectral estimator. *IEEE Transactions on Signal Processing*, 48(11):3189–3205, 2000.

[9] C.I. Byrnes, T.T. Georgiou, and A. Lindquist. A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint. *IEEE Transactions on Automatic Control*, 46(6):822–839, 2001.

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

[11] C.I. Byrnes and A. Lindquist. The generalized moment problem with complexity constraint. *Integral Equations and Operator Theory*, 56(2):163–180, 2006.

[12] C.I. Byrnes, A. Lindquist, S.V. Gusev, and A.S. Matveev. A complete parameterization of all positive rational extensions of a covariance sequence. *IEEE Trans. Autom. Contr.*, 40(11):1841–1857, 1995.

[13] A. Chiuso, A. Ferrante, and G. Picci. Reciprocal realization and modeling of textured images. In *44th IEEE Conf. Decision and Control and European Control Conf.*, pp. 6059–6064, Dec 2005.

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

[15] B. Dumitrescu. *Positive Trigonometric Polynomials and Signal Processing Applications*. Springer, Berlin, 2007.

[16] M.P. Ekstrom. *Digital image processing techniques*. Academic Press, 1984.

[17] P. Enqvist. A convex optimization approach to ARMA(n,m) model design from covariance and cepstral data. *SIAM Journal on Control and Optimization*, 43(3):1011–1036, 2004.

[18] P. Enqvist and E. Avventi. Approximative covariance interpolation with a quadratic penalty. In *46th IEEE Conf. Decision and Control*, pp. 4275–4280, 2007.

[19] S. Eriksson Barman. Gaussian random field based models for the porous structure of pharmaceutical film coatings. In *Acta Stereologica [En ligne], Proceedings ICSIA, 14th ICSIA abstracts*, 2015. `http://popups.ulg.ac.be/0351-580X/index.php?id=3775`.

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

[21] E. Fornasini. A 2-D systems approach to river pollution modelling. *Multidimensional Systems and Signal Processing*, 2(3):233–265, 1991.

[22] T.T. Georgiou. Partial Realization of Covariance Sequences. Center for Mathematical Systems Theory, Univerisity of Florida. PhD thesis, 1983.

[23] T.T. Georgiou. Realization of power spectra from partial covariance sequences. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35(4):438–449, 1987.

[24] T.T. Georgiou. The interpolation problem with a degree constraint. *IEEE Transactions on Automatic Control*, 44(3):631–635, 1999.

[25] T.T. Georgiou. Solution of the general moment problem via a one-parameter imbedding. *IEEE Transactions on Automatic Control*, 50(6):811–826, 2005.

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

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

[28] J. S. Geronimo and H. J. Woerdeman. Positive extensions, Fejér-Riesz factorization and autoregressive filters in two variables. *Annals of Mathematics*, 160(3):839–906, 2004.

[29] M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, *Recent Advances in Learning and Control*, volume 371 of *Lecture Notes in Control and Information Sciences*, pages 95–110. Springer-Verlag, London, 2008.

[30] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, Mar. 2014.

[31] W. Greblicki. Nonparametric identification of Wiener systems. *IEEE Transactions on information theory*, 38(5):1487–1493, 1992.

[32] R.E. Kalman. Realization of covariance sequences. In *Toeplitz memorial conference*, 1981. Tel Aviv, Israel.

[33] J. Karlsson, T.T. Georgiou and A. Lindquist. The Inverse Problem of Analytic Interpolation With Degree Constraint and Weight Selection for Control Synthesis. *IEEE Transactions on Automatic Control*, 55(2): 405–418, 2010.

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

[35] R. Kashyap and P. Lapsa. Synthesis and estimation of random fields using long-correlation models. *IEEE transactions on pattern analysis and machine intelligence*, PAMI-6(6):800–809, 1984.

[36] S.W. Lang and J.H. McClellan. Spectral estimation for sensor arrays. In *Proceedings of the First ASSP Workshop on Spectral Estimation*, pages 3.2.1–3.2.7, 1981.

[37] S.W. Lang and J.H. McClellan. Multidimensional MEM spectral estimation. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 30(6):880–887, 1982.

[38] S.W. Lang and J.H. McClellan. Spectral estimation for sensor arrays. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 31(2):349–358, 1983.

[39] A. Lindquist, C. Masiero, and G. Picci. On the multivariate circulant rational covariance extension problem. In *IEEE 52nd Annual Conference on Decision and Control (CDC)*, pages 7155–7161, 2013.

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

[41] A. Lindquist and G. Picci. *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification*, volume 1 of *Series in Contemporary Mathematics*. Springer-Verlag Berlin Heidelberg, 2015.

[42] F. Lindsten, T.B. Schön, and M.I. Jordan. Bayesian semiparametric Wiener system identification. *Automatica*, 49(7):2053–2063, 2013.

[43] H.I. Nurdin. New results on the rational covariance extension problem with degree constraint. *Systems & Control Letters*, 55(7):530 – 537, 2006.

[44] M. Pavon and A. Ferrante. On the geometry of maximum entropy problems. *SIAM Review*, 55(3):415–439, 2013.

[45] G. Picci and F.P. Carli. Modelling and simulation of images by reciprocal processes. In *Tenth international conference on Computer Modelling and Simulation, UKSIM*, pages 513–518, 2008.

[46] R. Price. A useful theorem for nonlinear devices having gaussian inputs. *IRE Transactions on Information Theory*, 4(2):69–72, 1958.

[47] F. Ramponi, A. Ferrante, and M. Pavon. A globally convergent matricial algorithm for multivariate spectral estimation. *IEEE Transactions on Automatic Control*, 54(10):2376–2388, 2009.

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

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

[50] P. Stoica and R. Moses. *Introduction to Spectral Analysis*. Prentice-Hall, Upper Saddle River, N.J., 1997.

[51] B. Wahlberg, J. Welsh, and L. Ljung. Identification of stochastic Wiener systems using indirect inference. *IFAC-PapersOnLine*, 48(28):620 – 625, 2015. 17th IFAC Symposium on System Identification SYSID 2015.

[52] P. Whittle. On stationary processes in the plane. *Biometrika*, pages 434–449, 1954.

[53] M. Zorzi. Rational approximations of spectral densities based on the alpha divergence. *Mathematics of Control, Signals, and Systems*, 26(2):259–278, 2014.