2015 IEEE 54th Annual Conference on Decision and Control (CDC)
December 15-18, 2015. Osaka, Japan

The Multidimensional Circulant Rational Covariance Extension Problem: Solutions and Applications in Image Compression

Axel Ringh, Johan Karlsson and Anders Lindquist

Abstract— **Rational functions play a fundamental role in systems engineering for modelling, identification, and control applications. In this paper we extend the framework by Lindquist and Picci for obtaining such models from the circulant trigonometric moment problems, from the one-dimensional to the multidimensional setting in the sense that the spectrum domain is multidimensional. We consider solutions to weighted entropy functionals, and show that all rational solutions of certain bounded degree can be characterized by these. We also consider identification of spectra based on simultaneous covariance and cepstral matching, and apply this theory for image compression. This provides an approximation procedure for moment problems where the moment integral is over a multidimensional domain, and is also a step towards a realization theory for random fields.**

I. Introduction

In 1981 R.E. Kalman posed the so called *rational covariance extension problem* (RCEP) [25]: Given a finite covariance sequence $c_0, \dots, c_n$, determine all extensions $c_{n+1}, c_{n+2}, \dots$ to an infinite sequence such that

$$\Phi(e^{i\theta}) = \sum_{k=-\infty}^{\infty} c_k e^{-ik\theta}, \quad \theta \in \mathbb{T} := (-\pi, \pi]$$

is a non-negative rational function of degree bounded by n, i.e., of the form $\Phi(e^{i\theta}) = P(e^{i\theta})/Q(e^{i\theta})$ where $P(e^{i\theta})$ and $Q(e^{i\theta})$ are non-negative trigonometric polynomials of degree less than or equal to n. Finite-dimensional systems are naturally represented as rational functions and this inverse problem is important in systems theory for estimation and realization of low degree systems [36].

The problem was partially solved in 1983, when T.T. Georgiou [19] proved that to each positive covariance sequence and non-negative numerator polynomial P, there exists a rational covariance extension of the sought form $\Phi = P/Q$. He also conjectured that this extension is unique and that it gives a complete parameterization of all rational extensions. This became a long standing conjecture, and was proved first in [11]. This led to an approach based on convex optimization [9], where the extension Φ is obtained as the maximizer of a generalized entropy functional:

$$\begin{aligned} &\max \quad \int_{\mathbb{T}} P(e^{i\theta}) \log \Phi(e^{i\theta}) d\theta \\ &\text{subject to} \quad c_k = \frac{1}{2\pi} \int_{\mathbb{T}} e^{ik\theta} \Phi(e^{i\theta}) d\theta, \ \text{for } k = 0, 1, \dots, n. \end{aligned}$$

This approach have been extensively studied in the one-dimensional setting, where the domain (in this case $\mathbb{T}$) is a one dimensional set [4]–[6], [12], [16], [20], [35], [40], [44], [46]. The approach has also been generalized to a quite complete theory for scalar moment problems [7], [8], [10], [23] and a number of matrix valued counterparts have been solved [1], [18], [22], [34], [41], [42], [50].

In the recent paper [35], Lindquist and Picci studied a discrete version of the one-dimensional RCEP where the covariances now satisfy a moment condition

$$c_k = \frac{1}{N} \sum_{j=1}^{N} \zeta_j^k \Phi(\zeta_j), \quad k = 0, 1, \dots, n,$$

where $\zeta_j = e^{ij\frac{2\pi}{N}}$, $j = 1, 2, \dots, N$. These are covariances of a periodic stationary process with period N, and the corresponding covariance extension problem is called the *circulant rational covariance extension problem*.

In this paper the multidimensional version of this problem is considered, and we extend the theory developed in [35] to the case where the domain of the process as well as the spectrum is naturally embedded in d dimensions (in this case $\mathbb{Z}^d$ and $\mathbb{T}^d$). This turns out to be a nontrivial extension leading to technical complication, as we shall illustrate by examples. Moreover, we also provide a correction of a result in [35] for the case that P has a zero on the unit circle. In this case a function with support in the zero set of P may have to be added, something that was overlooked in [35].

Many spectral estimation problems, such as problems in radar, sonar, and medical imaging, are essentially multidimensional covariance extension problems, but it is also of interest for modelling multidimensional reciprocal processes, random Markov fields, and imaging (cf. [14], [33]). A considerable amount of research has been done, for example Woods [49], Ekstrom and Woods [15], Dickinson [13], Lang and McClellan [28]–[31], [37], [38], and Lev-Ari *et al.* [32], to mention a few. In many of these areas it seems natural to consider rational models. Nevertheless, the multidimensional version of the RCEP has only been considered at a few instances [21], [22], [45].

The outline of this work is as follows: in Section II we review some background material and set up notation. In Section III we derive the main result for covariance matching and

This work was supported by the Swedish Research Council, the Swedish Foundation for Strategic Research and the The Center for Industrial and Applied Mathematics, KTH.

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 Department of Mathematics, Shanghai Jiao Tong University, Shanghai, China, and the Center for Industrial and Applied Mathematics (CIAM) and the ACCESS Linnaeus Center, KTH Royal Institute of Technology, Stockholm, Sweden. Email: alq@kth.se.

978-1-4799-7886-1/15/$31.00 ©2015 IEEE

characterize the optimal solutions to the weighted entropy functional. In Section IV we consider simultaneous matching of covariance and cepstral coefficients, and in Section V we give examples of how the theory can be applied in image compression.

II. Background and notation

Consider the multidimensional, discrete-time, zero-mean, and homogeneous stochastic process $y(t) \in \mathbb{C}$, defined for $t \in \mathbb{Z}^d$. Homogeneity of the process implies that covariances $c_k := E(\, y(t+k)\overline{y(t)}\,)$ are invariant with "time" $t \in \mathbb{Z}^d$. The power spectrum, $d\mu$, represents the energy distribution across frequency of the signal, and is the non-negative measure on $\mathbb{T}^d$ whose Fourier coefficients are the covariances

$$c_{\boldsymbol{k}} = \int_{\mathbb{T}^d} e^{i(\boldsymbol{k},\boldsymbol{\theta})} d\mu(e^{i\boldsymbol{\theta}}),$$

where $\boldsymbol{k} := (k_1, \dots, k_d) \in \mathbb{Z}^d$, $\boldsymbol{\theta} := (\theta_1, \dots, \theta_d) \in \mathbb{T}^d$, $(\boldsymbol{k}, \boldsymbol{\theta}) := \sum_{j=1}^d k_j \theta_j$, and $e^{i\boldsymbol{\theta}} = (e^{i\theta_1}, \dots, e^{i\theta_d})$ (cf. [45]). For N-periodic processes, where $N := (N_1, \dots, N_d) \in \mathbb{Z}^d$ represent the period in each direction, the corresponding covariances are periodic as well (i.e., the covariance matrix is d-level circulant). Therefore, the support of the spectrum $d\mu$ belongs to the rectangular grid $\mathbb{T}_N^d = \{(\ell_1 \frac{2\pi}{N_1}, \dots, \ell_d \frac{2\pi}{N_d}) : \boldsymbol{\ell} \in \mathbb{Z}_N^d\}$ where

$$\mathbb{Z}_N^d = \{(\ell_1, \dots, \ell_d) : 0 \le \ell_j \le N_j - 1, j = 1, \dots, d\}.$$

We therefore represent the spectrum by $\Phi(\boldsymbol{\zeta_\ell}) := \mu(\boldsymbol{\zeta_\ell})|N|$, where

$$\boldsymbol{\zeta_\ell} := (e^{i\ell_1 \frac{2\pi}{N_1}}, \dots, e^{i\ell_d \frac{2\pi}{N_d}}),$$

which correspond to the energy in $\boldsymbol{\theta} \in \mathbb{T}_N^d$ and where $|N| = \prod_{j=1}^d N_j$ is a normalizing constant. We also define $\boldsymbol{\zeta_\ell^k} = \prod_{j=1}^d \zeta_{\ell_j}^{k_j}$. Then by definition the covariances are now the inverse discrete Fourier transform of the spectrum Φ:

$$c_{\boldsymbol{k}} = \frac{1}{|N|} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_N^d} \boldsymbol{\zeta_\ell^k} \Phi(\boldsymbol{\zeta_\ell}). \tag{1}$$

A central problem in signal analysis is the inverse problem of recovering the power spectrum Φ based on a finite set of known covariances. This inverse problem is a key component in many signal processing techniques and plays a fundamental role in prediction, analysis, and modelling of signals [48]. In this setting we consider finite covariance sequences $\{c_{\boldsymbol{k}}\}_{\boldsymbol{k} \in \Lambda}$ where $\Lambda \subset \mathbb{Z}^d$. In many applications, the indices of the covariance sequence is the rectangular set $\Lambda = \{(k_1, \dots, k_d) \in \mathbb{Z}^d : |k_j| \le n_j, j = 1, \dots, d\}$, but the theory holds for any index set such that $\boldsymbol{0} \in \Lambda$ and $-\Lambda = \Lambda$.[1] We denote the number of elements in Λ with $|\Lambda|$, and with $n_j := \max\{|k_j| : \boldsymbol{k} \in \Lambda\}$ we denote the highest index in dimension j.

Let $\mathfrak{P}$ be the set of all multidimensional trigonometric polynomials associated with the index set Λ:

$$P(e^{i\boldsymbol{\theta}}) = \sum_{\boldsymbol{k} \in \Lambda} p_{\boldsymbol{k}} e^{-i(\boldsymbol{k},\boldsymbol{\theta})}, \quad p_{-\boldsymbol{k}} = \bar{p}_{\boldsymbol{k}}.$$

Note that the only index sets of interest are the index sets Λ such that the monomials are linearly independent on $\mathbb{T}_N^d$. A sufficient condition for this is given in the following lemma, which is proved in the Appendix, and we assume that this condition holds throughout the rest of this paper.

Lemma 1: Let $2n_j < N_j$ for $j = 1, \dots, d$. Then a polynomial in $\mathfrak{P}$ cannot vanish in all the points $\boldsymbol{\zeta_\ell}$ with $\boldsymbol{\ell} \in \mathbb{Z}_N^d$ unless it is the zero-polynomial.

Next, we define the convex cone of positive polynomials

$$\mathfrak{P}_+(N) = \{P \in \mathfrak{P} : P(\boldsymbol{\zeta_\ell}) > 0 \text{ for } \boldsymbol{\ell} \in \mathbb{Z}_N^d\},$$

and the closure, $\bar{\mathfrak{P}}_+(N)$, consists of all trigonometric polynomials in $\mathfrak{P}$ that are non-negative on $\mathbb{T}_N^d$, i.e., in the points $\boldsymbol{\zeta_\ell}$ for $\boldsymbol{\ell} \in \mathbb{Z}_N^d$. We also define the interior of the dual cone

$$\mathfrak{C}_+(N) = \{c : \langle c, p \rangle > 0, \forall P \in \bar{\mathfrak{P}}_+(N) \setminus \{0\}\},$$

where the inner product is $\langle c, p \rangle = \sum_{\boldsymbol{k} \in \Lambda} c_{\boldsymbol{k}} \bar{p}_{\boldsymbol{k}}$. Let $\bar{\mathfrak{C}}_+(N)$ be the closure of $\mathfrak{C}_+(N)$, and let $\partial\mathfrak{P}_+(N)$ and $\partial\mathfrak{C}_+(N)$ be the boundaries of $\mathfrak{P}_+(N)$ and $\mathfrak{C}_+(N)$, respectively.

The dual cone characterizes the existence of a spectrum Φ that satisfies (1) for a given covariance c. In fact, Farkas Lemma implies[2] that *exactly* one of the following holds:

i) $\exists\, p \in \mathfrak{P}$ such that $P(\boldsymbol{\zeta_\ell}) \ge 0$ for all $\boldsymbol{\ell} \in \mathbb{Z}_N^d$, and $\langle c, p \rangle < 0$

ii) $\exists\, \Phi \ge 0$ such that (1) holds for all $\boldsymbol{k} \in \Lambda$.

Now note that i) holds if and only if $c \notin \bar{\mathfrak{C}}_+(N)$. Therefore, there exists a spectrum Φ that matches the covariance sequence c if and only if $c \in \bar{\mathfrak{C}}_+(N)$. Furthermore, if c belongs to the interior $\mathfrak{C}_+(N)$ then there exists a strictly positive matching spectrum Φ. To see this, let c^0 be the covariance sequence corresponding to the constant spectrum $\Phi \equiv 1$ via (1). Then since c belong to the interior dual cone $\mathfrak{C}_+(N)$ there exists $\varepsilon > 0$ such that $\tilde{c} = c - \varepsilon c^0 \in \bar{\mathfrak{C}}_+(N)$. The spectrum $\varepsilon + \tilde{\Phi}$, where $\tilde{\Phi}$ is a spectrum that matches $c - \varepsilon c^0$, is now a strictly positive spectrum that satisfy (1) for $\boldsymbol{k} \in \Lambda$.

III. The Multidimensional rational covariance extension problem

In this paper we study the structure of solutions to generalized maximum entropy problems and how such convex optimization problems can be used for obtaining rational solutions to the multidimensional trigonometric moment problem. Given such a covariance sequence we seek spectra Φ that satisfy the covariance constraints (1) for $\boldsymbol{k} \in \Lambda$, and that are non-negative rational trigonometric functions:

$$\Phi(\boldsymbol{\zeta_\ell}) = \frac{P(\boldsymbol{\zeta_\ell})}{Q(\boldsymbol{\zeta_\ell})}, \quad \text{where } P, Q \in \bar{\mathfrak{P}}_+(N) \setminus \{0\}.$$

The generalized maximum entropy problem we consider is an entropy functional of the following form:

$$\max_{\Phi \ge 0} \quad \frac{1}{|N|} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_N^d} P(\boldsymbol{\zeta_\ell}) \log \Phi(\boldsymbol{\zeta_\ell}) \tag{2}$$
$$\text{subject to} \quad (1) \text{ for } \boldsymbol{k} \in \Lambda.$$

[1] The relation $-\Lambda = \Lambda$ comes from the fact that $c_{-\boldsymbol{k}} = \bar{c}_{\boldsymbol{k}}$, and because of this Λ will have an odd number of elements.

[2] To see this, consider the real and imaginary parts separately and use for example [2, Page 263].

Note that in the multidimensional case, limits such as $P\log Q$ and P/Q may not be well defined. In the problems and derivations we therefore define the expressions $P\log Q$, P/Q, and P/Q^2 to be zero whenever $P=0$.

Although one could approach the primal problem (2) directly, it is often more convenient to work with the dual. This objective function takes the form

$$\mathbb{J}_P(Q) = \langle c, q\rangle - \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta_\ell})\log Q(\boldsymbol{\zeta_\ell}), \tag{3}$$

and the optimization problem is given by

$$\min_{Q\in\bar{\mathfrak{P}}_+(N)} \quad \mathbb{J}_P(Q). \tag{4}$$

Theorem 2: For every $P\in\bar{\mathfrak{P}}_+(N)\setminus\{0\}$ and $c\in\mathfrak{C}_+(N)$ the dual optimization problem (4) is convex and has a solution $\hat{Q}\in\bar{\mathfrak{P}}_+(N)\setminus\{0\}$. Moreover, there also exist a positive function $\hat{\mu}$, with support $\operatorname{supp}(\hat{\mu})\subseteq\{\boldsymbol{\zeta_\ell} : \hat{Q}(\boldsymbol{\zeta_\ell})=0, \boldsymbol{\ell}\in\mathbb{Z}_N^d\}$, such that $\hat{\Phi}=P/\hat{Q}+\hat{\mu}$ is optimal to (2). This $\hat{\mu}$ might not be unique, but uniquely defines a covariance sequence

$$\hat{c}_k = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} \boldsymbol{\zeta_\ell^k}\hat{\mu}(\boldsymbol{\zeta_\ell}), \text{ for all } \boldsymbol{k}\in\Lambda.$$

Furthermore, $\hat{c}$ belongs to $\partial\mathfrak{C}_+(N)$, and $\hat{\mu}$ can always be chosen so that it has support in at most $|\Lambda|-1$ points.

If we restrict the choice of P to $P\in\mathfrak{P}_+(N)$ we can say more about the solution.

Corollary 3: For every $c\in\mathfrak{C}_+(N)$ and $P\in\mathfrak{P}_+(N)$ there exists a unique $\hat{Q}\in\mathfrak{P}_+(N)$ such that $\hat{\Phi}=P/\hat{Q}$ satisfies (1). Moreover, both (2) and (4) are strictly convex optimization problems and their respective solutions are $\hat{\Phi}$ and $\hat{Q}$.

Note that this corollary is only valid for $P\in\mathfrak{P}_+(N)$, while Theorem 2 holds for all $P\in\bar{\mathfrak{P}}_+(N)\setminus\{0\}$. The reason for the difference between P in $\bar{\mathfrak{P}}_+(N)$ or in $\mathfrak{P}_+(N)$ is that if $P\in\mathfrak{P}_+(N)$ then $\mathbb{J}_P(Q)=\infty$ whenever Q is on the boundary $\partial\mathfrak{P}_+(N)$ and the optimal solution Q will not be attained on $\partial\mathfrak{P}_+(N)$. However, if $P\in\partial\mathfrak{P}_+(N)$ the optimal Q may belong to $\partial\mathfrak{P}_+(N)$ in which case Q only have zeros in a subset of the zeros of P and the sum is finite. This subtle difference will become more clear by the proof of Theorem 2 and Corollary 3, to which the remaining of this section will be devoted.

A. The primal problem

For a given $P\in\bar{\mathfrak{P}}_+(N)\setminus\{0\}$ and $c\in\mathfrak{C}_+(N)$, consider the primal problem (2) where Φ is a non-negative function defined on $\mathbb{T}_N^d$. Denoting the objective function by $\mathbb{I}_P$, the second directional derivative, given by

$$\partial^2\mathbb{I}_P(\Phi;\delta\Phi) = -\frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\frac{P(\boldsymbol{\zeta_\ell})}{\Phi(\boldsymbol{\zeta_\ell})^2}(\delta\Phi(\boldsymbol{\zeta_\ell}))^2, \tag{5}$$

is non-positive and thus the maximization problem is convex. For $P\in\mathfrak{P}_+(N)$ we see that (5) vanish if and only if $\delta\Phi\equiv 0$, and thus the problem is strictly convex in this case.

Since $c\in\mathfrak{C}_+(N)$ there exists a strictly positive Φ that is feasible for (2) (see section II), hence Slater's condition is satisfied [2, Page 226], ensuring strong duality and that the dual problem achieves a minimum.

B. Lagrangian relaxation of the problem

The Lagrangian of the primal problem (2) is given by

$$\begin{aligned}\mathcal{L}_P(\Phi,Q) &= \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta_\ell})\log\Phi(\boldsymbol{\zeta_\ell}) \\ &\quad + \sum_{\boldsymbol{k}\in\Lambda}\bar{q}_{\boldsymbol{k}}\Big(c_{\boldsymbol{k}} - \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\boldsymbol{\zeta_\ell^k}\Phi(\boldsymbol{\zeta_\ell})\Big) \\ &= \langle c,q\rangle + \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\big(P(\boldsymbol{\zeta_\ell})\log\Phi(\boldsymbol{\zeta_\ell}) - Q(\boldsymbol{\zeta_\ell})\Phi(\boldsymbol{\zeta_\ell})\big)\end{aligned} \tag{6}$$

where $\bar{q}_{\boldsymbol{k}}$, for $\boldsymbol{k}\in\Lambda$, are the Lagrangian multipliers.

We seek a saddle point to this problem, maximizing over Φ and minimizing over Q. Examining (6) we see that the dual function $\sup_{\Phi\geq 0}\mathcal{L}_P(\Phi,Q)$ is finite only if $Q\in\bar{\mathfrak{P}}_+(N)\setminus\{0\}$, since otherwise we can let $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})\to\infty$ in some point where $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})\leq 0$ and get $\sup_{\Phi\geq 0}\mathcal{L}_P(\Phi,Q)=\infty$.

Two other things can also be noticed from the Lagrangian, when maximizing over Φ. First: the expression is only finite for Φ which is zero in some point, if we have $P=0$ in this point as well. Therefore, we can not have $\Phi=0$ unless $P=0$. Second: the supremum is only finite for $Q=0$ in some point, if we have $P=0$ in the same point.

Now we consider the directional derivative

$$\begin{aligned}\delta\mathcal{L}_P(\Phi,Q;\delta\Phi) &= \lim_{\varepsilon\to 0}\frac{\mathcal{L}_P(\Phi+\varepsilon\delta\Phi,Q)-\mathcal{L}_P(\Phi,Q)}{\varepsilon} \\ &= \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}(P(\boldsymbol{\zeta_\ell})\frac{1}{\Phi(\boldsymbol{\zeta_\ell})} - Q(\boldsymbol{\zeta_\ell}))\delta\Phi(\boldsymbol{\zeta_\ell})\end{aligned}$$

for any direction $\delta\Phi$ such that $\Phi+\epsilon\delta\Phi\geq 0$ for some $\epsilon>0$. For an optimal point this should be less than or equal to zero for all feasible directions $\delta\Phi$. We now need to analyse this in different situations.

1) In the case $P>0$ we must have, as noted before, $\Phi>0$ in an optimal point. This means that all directions $\delta\Phi$ are feasible, and thus we need to have

$$P(\boldsymbol{\zeta_\ell})\frac{1}{\Phi(\boldsymbol{\zeta_\ell})} - Q(\boldsymbol{\zeta_\ell}) = 0 \quad\Longrightarrow\quad \Phi(\boldsymbol{\zeta_\ell}) = \frac{P(\boldsymbol{\zeta_\ell})}{Q(\boldsymbol{\zeta_\ell})}$$

2) For the case when $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})=0$ in some $\boldsymbol{\zeta}_{\boldsymbol{\ell}_0}\in\mathbb{Z}_N^d$, we first consider the expression for the Lagrangian (6). Note that $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})\log\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})=0$ in the first term in the sum. Hence if $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})>0$, the second term gives that the pair (Φ,Q) can only be a saddle point if $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})=0$. This means that in a stationary point we only have $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})>0$ if $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})=0$. Moreover, if $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})=0$ we need to have $\delta\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})\geq 0$. The corresponding term in the derivative then reads $-Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})\delta\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})\leq 0$, which is true since $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})\geq 0$.

Summarizing this Φ must thus take the following form:

$$\Phi(\boldsymbol{\zeta_\ell}) = \begin{cases}\dfrac{P(\boldsymbol{\zeta_\ell})}{Q(\boldsymbol{\zeta_\ell})} & \text{if } Q(\boldsymbol{\zeta_\ell})>0, \\ \text{arbitrary} & \text{if } Q(\boldsymbol{\zeta_\ell})=0.\end{cases} \tag{7}$$

C. The dual problem

Using (7) we get that the dual function takes the form

$$\sup_{\Phi\geq 0} \mathcal{L}_P(\Phi, Q) = \mathbb{J}_P(Q) + \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})(\log P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) - 1)$$

where

$$\mathbb{J}_P(Q) = \langle c, q\rangle - \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})\log Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})$$

and where by definition $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})/Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0}) = 0$ if $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0}) = 0$, regardless of the value of $Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}_0})$. Now since $|N|^{-1}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\log P - 1)$ is independent of Q, $\mathbb{J}_P(Q)$ and $\sup_{\Phi\geq 0}\mathcal{L}_P(\Phi, Q)$ obtains their minima at the same point Q. Therefore we can take $\mathbb{J}_P(Q)$ to be the dual function, resulting in the dual optimization problem (4).

From duality theory the dual problem is convex [2, Page 216]. Forming the second directional derivative gives

$$\partial^2\mathbb{J}_P(Q; \delta Q) = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} \frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^2}(\delta Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}))^2, \tag{8}$$

and the dual is thus strictly convex if $P \in \mathfrak{P}_+(N)$.

D. Complementarity

We now introduce $\hat{\mu}$ as the part of Φ which, according to (7), is not given by $P/\hat{Q}$. What remains to prove is that this $\hat{\mu}$ defines a unique covariance $\hat{c} \in \partial\mathfrak{C}_+(N)$, and that it can be chosen with mass in at most $|\Lambda| - 1$ points. In order to do this we consider the components of $\hat{c}$. These are given by

$$\hat{c}_{\boldsymbol{k}} := \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\hat{\mu}(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\left(\hat{\Phi}(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) - \frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{\hat{Q}(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right),$$

which belong to $\bar{\mathfrak{C}}_+(N)$ since $\hat{\mu}$ is non-negative. From the last expression we can see that $\hat{c}$ is in fact unique, although $\hat{\mu}$ might not be. To see this we first note that $\hat{\Phi}$ matches the covariance sequence c. Secondly we note from (8) that directions which are potentially not strictly convex all have components only in points where $P = 0$. Hence the value of $P/\hat{Q}$ does not change in these directions.

Moreover, for $\hat{q}$ we get that

$$\langle\hat{c}, \hat{q}\rangle = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\hat{Q}\hat{\mu} = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\hat{Q}(\boldsymbol{\zeta}_{\boldsymbol{\ell}})\left(\hat{\Phi}(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) - \frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{\hat{Q}(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right).$$

Since $\hat{\Phi}$ has the form given in (7), we get that this expression is zero. Thus $\hat{c} \in \partial\mathfrak{C}_+(N)$. However the representation theorem in [29] says that for all $c \in \partial\mathfrak{C}_+(N)$ there exists a discrete representation with support in at most $|\Lambda| - 1$ points, which completes the proof of Theorem 2 and Corollary 3. ■

IV. Covariance and Cepstral Matching

Theorem 2 and Corollary 3 parametrize all multidimensional rational solutions that matches a given set of covariances. Comparing to the maximum entropy (ME) solution (cf. [3]), a better dynamical range can be obtained by encompassing *a priori* information of the problem through the choice of P. However how to select P is a non-trivial problem. Example of methods proposed for selecting P in one dimension are for example based on inverse problems [17], [26], [27], or based on simultanious matching of covariances and cepstral coefficients [4], [5], [16], [35], [39]. In this section we extend the later to the circulant, multidimensional setting.

Given a spectrum Φ, the (real) cepstrum is defined as the (real) logarithm of of the spectrum, $\log\Phi$. The cepstral coefficients, $m_{\boldsymbol{k}}$, are the Fourier coefficients of the cepstrum:

$$m_{\boldsymbol{k}} = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\log\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}), \tag{9}$$

for $\boldsymbol{k} \in \mathbb{Z}^d$ [35, and references therein]. Given both covariances and a set of cepstral coefficients, we can use this extra information to simultaneously estimate the spectral poles and zeros. This is done by maximizing the (unweighted) entropy subject to constraints on matching the covariances and the cepstral coefficients

$$\max_{\Phi\geq 0} \quad \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\log\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \tag{10}$$

$$\text{subject to} \quad (1) \text{ for } \boldsymbol{k}\in\Lambda, \text{ and } (9) \text{ for } \boldsymbol{k}\in\Lambda\backslash\{0\}.$$

Using this, and introducing the set

$$\mathfrak{P}_{+,\circ}(N) := \{P \in \mathfrak{P}_+(N) \mid p_{\boldsymbol{0}} = 1\},$$

we can state the results as follows.

Theorem 4: Given a $c \in \mathfrak{C}_+(N)$ and any sequence $\{m_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\Lambda}$, such that $m_{\boldsymbol{0}} \in \mathbb{R}$ and $m_{-\boldsymbol{k}} = \bar{m}_{\boldsymbol{k}}$, there exist a solution $(\hat{P}, \hat{Q}) \in \bar{\mathfrak{P}}_{+,\circ}(N)\times\bar{\mathfrak{P}}_+(N)$ to the convex problem

$$\min_{P,Q} \ \langle c, q\rangle - \langle m, p\rangle + \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})\log\left(\frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right) \tag{11}$$

$$\text{subject to} \quad P \in \bar{\mathfrak{P}}_{+,\circ}(N),\ Q \in \bar{\mathfrak{P}}_+(N).$$

If any solution $(\hat{P}, \hat{Q})$ belongs to $\mathfrak{P}_{+,\circ}(N)\times\mathfrak{P}_+(N)$ then $\hat{\Phi} = \hat{P}/\hat{Q}$ is also an optimal solution to the primal problem (10), and thus fulfils covariance and cepstral matching.

Proof: Considering the covariance matching constraint (1) of the primal problem (10), for $\boldsymbol{k} = \boldsymbol{0}$, we see that for all $\boldsymbol{\ell}\in\mathbb{Z}_N^d$ we must have $\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \leq |N|c_{\boldsymbol{0}}$. The problem is thus bounded, and since the objective function is continuous and strictly concave (cf. proof of Theorem 2) the problem has an optimal solution if there exists a feasible point.

By relaxing both equality constraints we get the Lagrangian

$$\begin{aligned}\tilde{\mathcal{L}}(\Phi, P, Q) = &\frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\log\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \\ &+ \sum_{\boldsymbol{k}\in\Lambda}\bar{q}_{\boldsymbol{k}}\left(c_{\boldsymbol{k}} - \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}})\right) \\ &+ \sum_{\substack{\boldsymbol{k}\in\Lambda\\ \boldsymbol{k}\neq\boldsymbol{0}}}\bar{p}_{\boldsymbol{k}}\left(\frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\log\Phi(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) - m_{\boldsymbol{k}}\right)\end{aligned}$$

where $\bar{q}_{\boldsymbol{k}}$ and $\bar{p}_{\boldsymbol{k}}$ are Lagrangian multipliers. Note that this expression does not contain $p_{\boldsymbol{0}}$ or $m_{\boldsymbol{0}}$. Hence we can introduce $p_{\mathbf{0}}$ fixed to 1 and an arbitrary but fixed $m_{\mathbf{0}} \in \mathbb{R}$, without altering the problem. Rearranging terms we get the equivalent Lagrangian

$$\mathcal{L}(\Phi, P, Q) = \langle c, q\rangle - \langle m, p\rangle - \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} Q(\boldsymbol{\zeta_\ell})\Phi(\boldsymbol{\zeta_\ell}) + \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta_\ell})\log\Phi(\boldsymbol{\zeta_\ell}).$$

As before, $\sup_{\Phi\geq 0}\mathcal{L}(\Phi, P, Q)$ is only finite if we restrict Q to the cone $\bar{\mathfrak{P}}_+(N)$, and similarly we need to restrict P to the set $\bar{\mathfrak{P}}_{+,\circ}(N)$.

Considering the directional derivative of $\mathcal{L}$ with respect to Φ, we again get the expression

$$\delta\mathcal{L}(\Phi, P, Q; \delta\Phi) = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\left(\frac{P(\boldsymbol{\zeta_\ell})}{\Phi(\boldsymbol{\zeta_\ell})} - Q(\boldsymbol{\zeta_\ell})\right)\delta\Phi(\boldsymbol{\zeta_\ell}).$$

In order for this to be non-positive for all feasible directions $\delta\Phi$, similar analysis gives that we must have

$$\Phi(\boldsymbol{\zeta_\ell}) = \begin{cases}\dfrac{P(\boldsymbol{\zeta_\ell})}{Q(\boldsymbol{\zeta_\ell})} & \text{if } P(\boldsymbol{\zeta_\ell}) > 0,\\ \text{arbitrary} & \text{if } P(\boldsymbol{\zeta_\ell}) = 0.\end{cases}$$

This gives the dual functional

$$\sup_{\Phi\geq 0}\mathcal{L}(\Phi, P, Q) = \mathbb{J}(P, Q) - \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta_\ell}) \tag{12}$$

where

$$\mathbb{J}(P, Q) = \langle c, q\rangle - \langle m, p\rangle + \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta_\ell})\log\left(\frac{P(\boldsymbol{\zeta_\ell})}{Q(\boldsymbol{\zeta_\ell})}\right). \tag{13}$$

A closer look at the last term of (12) shows that

$$\frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta_\ell}) = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\sum_{\boldsymbol{k}\in\Lambda} p_{\boldsymbol{k}}\boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}} = \sum_{\boldsymbol{k}\in\Lambda} p_{\boldsymbol{k}}\prod_{1\leq j\leq d}\frac{1}{N_j}\sum_{\ell_j\in\mathbb{Z}_{N_j}} e^{ik_j\ell_j/N_j} = p_{\mathbf{0}} = 1, \tag{14}$$

since all of these sums vanish, except for $\boldsymbol{k} = \mathbf{0}$. The last term is thus a constant, and hence we can take (13) as the dual objective function, which gives us the dual problem in (11). Again, since it is the dual it is convex.

In order to ensure existence of a minimizer to the dual problem, we need to show that (13) is lower semi-continuous and that it has compact sublevel sets. This follows from the following lemmas, which are proved in the Appendix.

Lemma 5: Given $c \in \mathfrak{C}_+(N)$ and any sequence $\{m_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\Lambda}$, where $m_{\mathbf{0}} \in \mathbb{R}$ and $m_{-\boldsymbol{k}} = \bar{m}_{\boldsymbol{k}}$, $\mathbb{J}(P, Q)$ is a lower semi-continuous function on $\bar{\mathfrak{P}}_{+,\circ}(N)\times\bar{\mathfrak{P}}_+(N)\setminus\{0\}$.

Lemma 6: The sublevel sets of $\mathbb{J}(P, Q)$ are compact.

Now we use the Wirtinger derivative (see e.g., [35, Page 2853]) and form the partial derivative of $\mathbb{J}(P, Q)$ with respect to both $\bar{q}_{\boldsymbol{k}}$ and $\bar{p}_{\boldsymbol{k}}$. This gives

$$\frac{\partial\mathbb{J}(P, Q)}{\partial\bar{q}_{\boldsymbol{k}}} = c_{\boldsymbol{k}} - \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\frac{P(\boldsymbol{\zeta_\ell})}{Q(\boldsymbol{\zeta_\ell})}, \tag{15a}$$

$$\frac{\partial\mathbb{J}(P, Q)}{\partial\bar{p}_{\boldsymbol{k}}} = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\log\left(\frac{P(\boldsymbol{\zeta_\ell})}{Q(\boldsymbol{\zeta_\ell})}\right) - m_{\boldsymbol{k}}, \tag{15b}$$

where (15a) is valid for $\boldsymbol{k} \in \Lambda$ and (15b) is valid for $\boldsymbol{k} \in \Lambda\setminus\{0\}$, and where we in (15b) used a similar result as in (14). From this we see that if the optimal solution is in the interior, i.e., if $(\hat{P}, \hat{Q}) \in \mathfrak{P}_{+,\circ}(N)\times\mathfrak{P}_+(N)$, and thus a stationary point to $\mathbb{J}(P, Q)$, then the spectrum $\hat{\Phi} = \hat{P}/\hat{Q}$ fulfil both the covariance matching (1) and the cepstral matching (9). ■

As can be seen in the above proof, the stationarity of $\mathbb{J}(P, Q)$ in Q gives covariance matching and the stationarity in P gives cepstral matching. Therefore we can only guarantee matching for a solution in the interior $\mathfrak{P}_{+,\circ}(N)\times\mathfrak{P}_+(N)$. However for $\hat{P} \in \partial\mathfrak{P}_{+,\circ}(N)$ we cannot guarantee that a solution $\hat{Q}$ belongs to the interior $\mathfrak{P}_+(N)$ (cf. Theorem 2 and Corollary 3), and thus it is not possible to guarantee covariance matching. This subtle fact has been overlooked in [34], [35], where it is stated that also when $\hat{P} \in \partial\mathfrak{P}_{+,\circ}(N)$ we would have $\hat{Q} \in \mathfrak{P}_+(N)$, which would guarantee covariance matching.

A. Regularizing the problem

The motivation for considering simultaneous covariance and cepstral matching was to obtain a rational spectrum $\Phi = P/Q$ that matches the covariances, but without having to provide the prior P. However, the solution to (11) cannot be guaranteed to give a spectrum that satisfies the covariance matching (1). In order to remedy this we consider the Enqvist regularized problem [16], which has the objective function

$$\mathbb{J}_\lambda(P, Q) = \langle c, q\rangle - \langle m, p\rangle + \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta_\ell})\log\left(\frac{P(\boldsymbol{\zeta_\ell})}{Q(\boldsymbol{\zeta_\ell})}\right) - \lambda\frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\log P(\boldsymbol{\zeta_\ell}),$$

and where $\lambda \in (0, \infty)$ is the regularization parameter. This will be infinite for all $P \in \partial\mathfrak{P}_{+,\circ}(N)$, and hence the optimal solution is not obtained here. Moreover with this regularization the optimization problem becomes strictly convex, and hence we have a unique solution.

Theorem 7: Given $c \in \mathfrak{C}_+(N)$ and a sequence $\{m_{\boldsymbol{k}}\}_{\boldsymbol{k}\in\Lambda}$, where $m_{\mathbf{0}} \in \mathbb{R}$ and $m_{-\boldsymbol{k}} = \bar{m}_{\boldsymbol{k}}$, for all $\lambda > 0$ there exist a unique solution $(\hat{P}, \hat{Q})$ to the strictly convex problem

$$\begin{aligned}\min \quad & \mathbb{J}_\lambda(P, Q)\\ \text{subject to} \quad & P \in \mathfrak{P}_{+,\circ}(N),\ Q \in \mathfrak{P}_+(N).\end{aligned} \tag{16}$$

For $\hat{\Phi} = \hat{P}/\hat{Q}$ we have that $\hat{\Phi}$ fulfils the covariance matching (1), and approximately fulfils the cepstral matching (9) via

$$m_{\boldsymbol{k}} + \varepsilon_{\boldsymbol{k}} = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\hat{\Phi}(\boldsymbol{\zeta_\ell}),\quad \varepsilon_{\boldsymbol{k}} = \lambda\frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d}\boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\frac{1}{\hat{P}(\boldsymbol{\zeta_\ell})}.$$

Proof: All of the results in the theorem follows from Theorem 4, together with the discussion in the section leading up to it, except the exact and approximate matching of (1) and (9), and the strict convexity.

To get the covariance and approximate cepstral matching, note that the partial derivative with respect to $\bar{q}_{\boldsymbol{k}}$ is identical to (15a), and the derivative with respect to $\bar{p}_{\boldsymbol{k}}$ given by

$$\frac{\partial \mathbb{J}_\lambda(P,Q)}{\partial \bar{p}_{\boldsymbol{k}}} = \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}\left(\log\left(\frac{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right) - \frac{\lambda}{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right) - m_{\boldsymbol{k}}.$$

To show strict convexity we note that the second $(\delta P, \delta Q)$-directional derivative of $\mathbb{J}_\lambda$ is given by (cf. [35, Theorem 8])

$$\frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})\left(\frac{\delta P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})} - \frac{\delta Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}{Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}})}\right)^2 + \frac{\lambda \delta P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^2}{P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})^2}.$$

Since both terms are non-negative, they both need to be zero in order for the second derivative to vanish. However since $P > 0$ in the optimal point this implies that $\delta P \equiv 0$. From this we get that the first term becomes $\delta Q^2 P/Q^2$ and in the same way we must thus have $\delta Q \equiv 0$. ■

V. Application in Image compression

In this section we consider an application of the two-dimensional, periodic RCEP in compression of black-and-white images. The main idea is to approximate the image with a rational spectrum and thereby represent the image by fewer parameters. In fact, choosing n_1 and n_2 much smaller than the respective dimensions N_1 and N_2 of the image leads to a significant reduction in number of parameters describing the image.

Both the ME spectrum and the solution resulting from regularized covariance and cepstral matching are determined and then compared. However, the cepstrum is not well-defined if the discrete spectrum is zero in one of the grid points. Therefore, instead of determining the covariances and cepstral coefficients directly from the image, we transform the image, denoted by Ψ, to $\Phi = e^{\Psi}$. Then, as Ψ is real, Φ is real and positive for all discrete frequencies, and $\Psi = \log \Phi$. Then $\hat{\Phi}$ is obtained from Theorem 7 with $c_{\boldsymbol{k}}$ and $m_{\boldsymbol{k}}$ given by (1) and (9), and is hence a low-degree rational approximation of Ψ. In the implementation we use the real sequences of covariances and cepstral coefficients corresponding to extending the image by symmetric mirroring (i.e., using the discrete cosine transform, [43, Section 4.2]). Also note that covariances and cepstral coefficients of Φ can be computed as the inverse 2D-FFT of e^{Ψ} and Ψ respectively.

This method is then applied to two different images: (i) the Shepp-Logan phantom, often used in medical imaging [47], of size 256×256 pixels, and (ii) the classical Lenna image, often used in the image processing literature. To get a fair comparison between the solution to (16) and the ME solution, we compensate for the fact that the latter has approximately half as many parameters by letting it have a degree $\sqrt{2}$ higher (rounded up).

The original image of the Shepp-Logan phantom is shown in Fig. 1a, whereas the corresponding compressed images using covariance and cepstral matching with $n_1 + 1 = n_2 + 1 = 30$ and the ME solution with degree $n_1 + 1 = n_2 + 1 = 45 \approx \sqrt{2} \cdot 30$ are shown in in Fig. 1b and Fig. 1c, respectively. Hence we have reduced the original $256^2 = 65536$ parameters to $2 \cdot 30^2 = 1800$ parameters.

The original Lenna image is 512×512 pixels and shown in Fig. 2a. The regularized cepstral matching solution with $n_1 + 1 = n_2 + 1 = 60$, leading to a reduction of the number of parameters from 262144 to $2 \cdot 60^2 = 7200$, is shown in Fig. 2b. The compression by the ME solution with $n_1 + 1 = n_2 + 1 = 85 \approx \sqrt{2} \cdot 60$ is shown in Fig. 2c.

Compression by cepstral matching is clearly better for the Shepp-Logan phantom, whereas it is harder to decide which method is better for the Lenna image. In fact, in this case the ME compression has more ringing artifacts but is less blurred than the cepstral compression. This is probably due to the fact that there are fewer sharp transitions in pixel values in Fig. 1a. By placing poles and zero close to each other, these transitions can be more efficiently compressed. However, in the Lenna image there is a trade-off between having spectral zeros or matching higher frequencies.

Appendix

Proof of Lemma 1: To prove Lemma 1, we use the following theorem that is a special case of Theorem 16.8 in [24] which build on Hilbert's Nullstellensatz for multivariate polynomials.

Theorem 8 ([24] Theorem 16.8): Let $a(z_1, \ldots, z_d)$ be a non-zero complex polynomial in $z_1, \ldots, z_d$ and let the degree of $a(z_1, \ldots, z_d)$ in z_j be n_j for $j = 1, \ldots, d$. Let $S_1, \ldots, S_d$ be finite subsets of $\mathbb{C}$ with $|S_j| \geq n_i + 1$, for $i = 1, \ldots, d$, then $a(z) \neq 0$ for at least one point $z = (z_1, \ldots, z_d)$ in $S_1 \times \ldots \times S_d$.

Note that if $2n_j < N_j$ for $j = 1, \ldots, d$, then

$$P(e^{i\boldsymbol{\theta}}) = \left(\prod_{j=1}^{d} e^{-in_j\theta_j}\right)\sum_{\boldsymbol{k}\in\Lambda} p_{\boldsymbol{k}} \exp\left(i\sum_{j=1}^{d}(k_j + n_j)\theta_j\right) \quad (17)$$

The first factor in (17) has unit magnitude, hence if $P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 0$ for all $\boldsymbol{\ell} \in \mathbb{Z}_N^d$ then the second factor in (17) must vanish as well in all the points $\boldsymbol{\zeta}_{\boldsymbol{\ell}}$. Now, by Theorem 8 this factor is zero if and only if $P \equiv 0$, and the proof is complete. ■

Proof of Lemma 5: Lower semi-continuity of $\mathbb{J}(P,Q)$ follows since $x\log(x/y)$ is lower semi-continuous for $x, y \geq 0$. The only point where $x\log(x/y)$ is not continuous is $x = y = 0$ (in which the value is defined to be 0). Since $x\log(x/y) \geq -y\exp(-1)$, we have that $\liminf_{x,y\to 0} x\log(x/y) \geq 0$ and consequently lower semi-continuity follows. ■

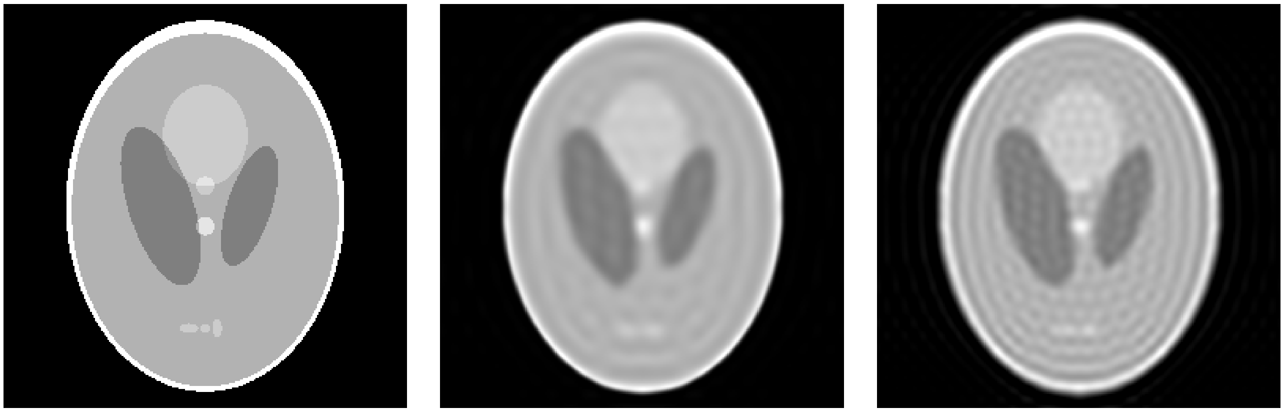
To prove Lemma 6 we need the two following results.

Lemma 9 ([10]): For a fixed $c \in \mathfrak{C}_+(N)$, there exists $\varepsilon > 0$ such that for every $(P,Q) \in \bar{\mathfrak{P}}_+(N)\setminus\{0\} \times \bar{\mathfrak{P}}_+(N)\setminus\{0\}$

$$\mathbb{J}_P(Q) \geq \varepsilon\|Q\|_\infty - \frac{1}{|N|}\sum_{\boldsymbol{\ell}\in\mathbb{Z}_N^d} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})\, \log\left(\|Q\|_\infty\right).$$

Proof of Lemma 9: Follows verbatim the proof of Proposition 2.1 in [10] if the integral $t \in [a,b]$ is replaced by the sum over $\boldsymbol{\zeta}_{\boldsymbol{\ell}}$ for $\boldsymbol{\ell} \in \mathbb{Z}_N^d$. ■

Lemma 10: For all trigonometric polynomials $P \in \bar{\mathfrak{P}}_+(N)$ we have that $|p_{\boldsymbol{k}}| \leq p_{\mathbf{0}}$, $\boldsymbol{k} \in \Lambda$.



(a) Original. (b) Cepstral matching, $n = 30$ and $\lambda = 10^{-2}$. (c) ME-solution, $n = 45$.

Fig. 1. Compression of the Shepp-Logan phantom, with a compression rate of about 97%.

(a) Original. (b) Cepstral matching, $n = 60$ and $\lambda = 10^{-2}$. (c) ME-solution, $n = 85$.

Fig. 2. Compression of the Lenna image, with a compression rate of about 97%.

Proof: This is proved by the fact that

$$|p_{\boldsymbol{k}}| = \left| \frac{1}{|N|} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_N^d} \boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \right| \leq \frac{1}{|N|} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_N^d} |\boldsymbol{\zeta}_{\boldsymbol{\ell}}^{\boldsymbol{k}}| \, |P(\boldsymbol{\zeta}_{\boldsymbol{\ell}})| = p_{\mathbf{0}}.$$

The last step follows from (14) and the fact that $P \geq 0$. ■

Proof of Lemma 6: The sublevel sets, $\mathbb{J}^{-1}(-\infty, r]$ for any $r \in \mathbb{R}$, are the $(P, Q) \in \bar{\mathfrak{P}}_{+,\circ}(N) \times \bar{\mathfrak{P}}_{+}(N)$ such that

$$r \geq \mathbb{J}(P, Q),$$

and to show that these are compact we start by splitting the objective function into two parts:

$$\mathbb{J}_1(P, Q) = \langle c, q \rangle - \frac{1}{|N|} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_N^d} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \log Q(\boldsymbol{\zeta}_{\boldsymbol{\ell}}),$$

$$\mathbb{J}_2(P) = -\langle m, p \rangle + \frac{1}{|N|} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_N^d} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \log P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}).$$

From Lemma 9 we get that

$$\mathbb{J}_1(P, Q) \geq \varepsilon \|Q\|_\infty - \frac{1}{|N|} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_N^d} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) \, \log(\|Q\|_\infty),$$

and from (14) we know that $|N|^{-1} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_N^d} P(\boldsymbol{\zeta}_{\boldsymbol{\ell}}) = 1$.

Turing the attention to $\mathbb{J}_2(P)$, we will show that it is bounded from below. To see this we first note that since $x \log(x) \geq -1/e$ we have

$$\frac{1}{|N|} \sum_{\boldsymbol{\ell} \in \mathbb{Z}_N^d} P \log(P) \geq -\frac{1}{e},$$

and this term is bounded from below. To bound the term $-\langle m, p \rangle$ from below we note that

$$|\langle m, p \rangle| = |\sum_{\boldsymbol{k} \in \Lambda} \bar{m}_{\boldsymbol{k}} p_{\boldsymbol{k}}| \leq |\Lambda| \|m\|_\infty \|p\|_\infty. \tag{18}$$

However from Lemma 10 we get that $\|p\|_\infty = p_{\mathbf{0}} = 1$, and thus $\langle m, p \rangle$ is bounded, hence we have $\mathbb{J}_2(P) \geq -|\Lambda| \|m\|_\infty - 1/e =: \rho > -\infty$. Using this we get that

$$r - \rho \geq \mathbb{J}_1(P, Q) \geq \varepsilon \|Q\|_\infty + \log(\|Q\|_\infty)$$

and comparing linear and logarithmic growth we see that the set is bounded both from above and below. Since it is the sublevel set of a lower semi-continuous function (Lemma 5) it will be closed, and hence it is compact. ■

REFERENCES

[1] A. Blomqvist, A. Lindquist, and R. Nagamune. Matrix-valued Nevanlinna-Pick interpolation with complexity constraint: an optimization approach. *IEEE Transactions on Automatic Control*, 48(12):2172–2190, 2003.

[2] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, 2004.

[3] J.P. Burg. Maximum entropy spectral analysis. In *Proceedings of the 37th Meeting Society of Exploration Geophysicists*, 1967.

[4] C.I. Byrnes, P. Enqvist, and A. Lindquist. Cepstral coefficients, covariance lags, and pole-zero models for finite data strings. *IEEE Transactions on Signal Processing*, 49(4):677–693, 2001.

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

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

[7] 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.

[8] C.I. Byrnes, T.T. Georgiou, A. Lindquist, and A. Megretski. Generalized interpolation in h-infinity with a complexity constraint. *Transactions of the American Mathematical Society*, 358(3):965–987, 2006.

[9] C.I. Byrnes, S.V. Gusev, and A. Lindquist. A convex optimization approach to the rational covariance extension problem. *SIAM Journal on Control and Optimization*, 37(1):211–229, 1998.

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

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

[12] F.P. Carli and T.T. Georgiou. On the covariance completion problem under a circulant structure. *IEEE Transactions on Automatic Control*, 56(4):918–922, 2011.

[13] B. Dickinson. Two-dimensional markov spectrum estimates need not exist. *IEEE Transactions on Information Theory*, 26(1):120–121, 1980.

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

[15] M.P. Ekstrom and J.W. Woods. Two-dimensional spectral factorization with applications in recursive digital filtering. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 24(2):115–128, 1976.

[16] 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.

[17] G. Fanizza. *Modeling and Model Reduction by Analytic Interpolation and Optimization*. PhD thesis, 2008. Optimization and systems theory, Department of mathematics, KTH Royal Institue of Technology.

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

[19] T.T. Georgiou. *Partial Realization of Covariance Sequences*. PhD thesis, 1983. Center for mathematical systems theory, Univeristy of Florida.

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

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

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

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

[24] S. Jukna. *Extremal combinatorics: with applications in computer science*. Springer Science & Business Media, 2011.

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

[26] 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.

[27] J. Karlsson and A. Lindquist. Stability-preserving rational approximation subject to interpolation constraints. *IEEE Transactions on Automatic Control*, 53(7):1724–1730, 2008.

[28] 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.

[29] S.W. Lang and J.H. McClellan. The extension of Pisarenko's method to multiple dimensions. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '82.*, volume 7, pages 125–128, May 1982.

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

[31] 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.

[32] H. Lev-Ari, S. Parker, and T. Kailath. Multidimensional maximum-entropy covariance extension. *IEEE Transactions on Information Theory*, 35(3):497–508, 1989.

[33] B.C. Levy, R. Frezza, and A.J. Krener. Modeling and estimation of discrete-time gaussian reciprocal processes. *IEEE Transactions on Automatic Control*, 35(9):1013–1023, 1990.

[34] 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.

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

[36] A. Lindquist and G. Picci. *Linear stochastic systems: A geometric approach to modeling, estimation and identification*. Springer-Verlag Berlin Heidelberg, 2015.

[37] J.H. McClellan and S.W. Lang. Mulit-dimensional MEM spectral estimation. In *Proceedings of the Institute of Acoustics "Spectral Analysis and its Use in Underwater Acoustics": Underwater Acoustics Group Conference, Imperial College, London, 29-30 April 1982*, pages 10.1–10.8, 1982.

[38] J.H. McClellan and S.W. Lang. Duality for multidimensional MEM spectral analysis. *Communications, Radar and Signal Processing, IEE Proceedings F*, 130(3):230–235, April 1983.

[39] B.R. Musicus and A.M. Kabel. Maximum entropy pole-zero estimation. Technical Report 510, Research Laboratory of Electronics, Massachusetts Institute of Technology, August 1985.

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

[41] 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.

[42] 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.

[43] K.R. Rao and P. Yip. *Discrete cosine transform: algorithms, advantages, applications*. Academic press, San Diego, C.A., 1990.

[44] A. Ringh and J. Karlsson. A fast solver for the circulant rational covariance extension problem. In *European Control Conference (ECC)*, pages 727–733, July 2015.

[45] A. Ringh, J. Karlsson, and A. Lindquist. Multidimensional rational covariance extension with applications to image compression. *arXiv preprint arXiv:1507.01430*, 2015.

[46] A. Ringh and A. Lindquist. Spectral estimation of periodic and skew periodic random signals and approximation of spectral densities. In *33rd Chinese Control Conference (CCC)*, pages 5322–5327, 2014.

[47] L.A Shepp and B.F. Logan. The Fourier reconstruction of a head section. *IEEE Transactions on Nuclear Science*, 21(3):21–43, 1974.

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

[49] J.W Woods. Two-dimensional Markov spectral estimation. *IEEE Transactions on Information Theory*, 22(5):552–559, 1976.

[50] M. Zorzi. A new family of high-resolution multivariate spectral estimators. *IEEE Transactions on Automatic Control*, 59(4):892–904, 2014.