Proceedings of the 35th Chinese Control Conference
July 27-29, 2016, Chengdu, China

An Identification Approach to Image Deblurring

Bin Zhu[1], Anders Lindquist[2]

1. University of Padova, Padova, Italy
E-mail: niuniuneverdie@gmail.com

2. Shanghai Jiao Tong University, Shanghai, P. R. China,
and KTH Royal Institute of Technology, Stockholm, Sweden
E-mail: alq@kth.se

Abstract: In this paper we present a new method of reconstructing an image that undergoes a spatially invariant blurring process and is corrupted by noise. The methodology is based on a theory of multidimensional moment problems with rationality constraints. This can be seen as generalized spectral estimation with a finiteness condition, which in turn can be considered a problem in system identification. With noise it becomes an ill-posed deconvolution problem and needs regularization. A Newton solver is developed, and the algorithm is tested on two images under different boundary conditions. These preliminary results show that the proposed method could be a viable alternative to regularized least squares for image deblurring, although more work is needed to perfect the method.

Key Words: Multidimensional moment problem, Rationality constraints, Image deblurring, Deconvolution, Convex optimization

1 Introduction

Image deblurring is a deconvolution problem in two dimensions. It is well noted that the problem of deconvolution is ill-posed [1–3], and hence regularization is crucial. The deblurring problem is often formulated as a regularized least squares problem, such as Tikhonov regularization, which has a closed form solution. Other regularization methods include those exploiting partial derivatives [4], total-variation deblurring [5, 6], or penalized maximum likelihood [7].

Blurring a two-dimensional image $\Phi(x)$, $x \in K \subset \mathbb{R}^2$, can be modeled as a convolution integral

$$b(x) = \int_K \kappa(x-y)\Phi(y)dy, \tag{1}$$

where κ is a kernel function, called the point spread function (PSF). Deblurring amounts to the deconvolution of (1), i.e., to recover the original image Φ from the blurred image b.

If the blurred image is observed in discrete points $x_1, x_2, \dots, x_n$ like pixels, then (1) becomes a generalized two-dimensional moment problem

$$c_k = \int_K \alpha_k(x)\Phi(x)dx, \quad k = 1, 2, \dots, n, \tag{2}$$

where $c_k := b(x_k)$ and $\alpha_k(x) := \kappa(x_k - x)$, $k = 1, 2, \dots, n$. Here $\alpha_1, \alpha_2, \dots, \alpha_n$ are called *basis functions*. Reconstructing Φ from $c_1, c_2, \dots, c_n$ is an inverse problem, which may or may not have a solution. If it does, it will in general have infinitely many. To achieve compression of data, we impose the rationality constraint

$$\Phi(x) = \frac{P(x)}{Q(x)}, \tag{3}$$

where P and Q are nonnegative functions formed by linear combinations of the basis functions. This can be seen as a (generalized) two-dimensional spectral estimation problem with a finiteness condition, and hence as a two-dimensional identification problem [8]. (In fact, general basis functions, rather than trigonometric ones, are also used in system identification [9].) If (2) does not have a solution, which is the usual case, a regularized approximate solution need to be determined.

The one-dimensional moment problem with rationality constraint has been studied intensively during the last decades. It originated with the *rational covariance extension problem*, first formulated in [10]. In the present context this problem can be reformulated as follows. Given a sequence of covariance lags $c_0, c_2, \dots, c_n$ with a positive definite Toeplitz matrix, parameterize the family of all functions (3) defined on the unit circle $\mathbb{T}$ in the complex plane and satisfying the moment condition (2), where P and Q are symmetric trigonometric polynomials of degree at most n. The first result on this problem can be found in [11], where it was shown that there exists a solution for each choice of zeros of P, and it was conjectured that the assignment is unique. This conjecture was proved in [12], where it was shown the the complete parameterization is smooth, and hence solution can be tuned continuously. In [13, 14] it was shown that each solution is the unique solution of a pair of dual convex optimization problems. This leads to a long list of results with more general basis functions, among them [15–32].

More recently, these results were generalized to the multidimensional case [33] with applications to spectral estimation and image compression [34]. Related results can be found in [35]. It turns out that the early papers [36–38] contain results that are equivalent to some major results in [33, 34], but the basic idea of smooth parameterization is missing there.

In this paper, we apply the method of the moment problem with rationality constraint to image deblurring with the help of regularization. The paper is organized as follows. In Section 2, we briefly introduce the main result of the theory of multidimensional moment problem and in Section 3 regularized approximate solutions are determined for the case that the estimated moments contain errors. We consider the optimization problem for image deblurring in the framework of multidimensional moment problem in Section 4, and a Newton solver is developed. Finally, some implementation details of the proposed method are given in Section 5 along with two reconstructed images. These results are preliminary, and better methods to tune the solutions will be developed in future work.

2 The multidimensional moment problem

We start by reviewing some results in [33]. Let $\mathfrak{P}_+$ be the positive cone of vectors $p := (p_1, p_2, \dots, p_n)$ such that

$$P(x) = \sum_{k=1}^{n} p_k \alpha_k(x) > 0 \quad \text{for all } x \in K, \tag{4}$$

and let $\bar{\mathfrak{P}}_+$ be the closure of $\mathfrak{P}_+$ and $\partial\mathfrak{P}_+ := \bar{\mathfrak{P}}_+ \backslash \mathfrak{P}_+$ its boundary. Then, given a set of real numbers $c_1, c_2, \dots, c_n$, and linearly independent functions $\alpha_1, \alpha_2, \dots, \alpha_n$ defined on a compact subset $K \subset \mathbb{R}^d$, consider the problem to find solutions Φ to the moment condition (2) of the rational form (3), where $p, q \in \mathfrak{P}_+$. Here of course q is the vector of coefficients of Q. Next define the open dual cone $\mathfrak{C}_+$ of vectors $c := (c_1, c_2, \dots, c_n)$, i.e.,

$$\mathfrak{C}_+ = \left\{ c \mid \langle c, p \rangle = \sum_{k=1}^{n} c_k p_k > 0, \quad \forall p \in \bar{\mathfrak{P}}_+ \backslash \{0\} \right\}. \tag{5}$$

If the cone $\mathfrak{P}_+$ is nonempty and has the property

$$\int_K \frac{1}{Q} dx = \infty \text{ for all } q \in \partial\mathfrak{P}_+, \tag{6}$$

it follows from [33, Corollary 3.5] that the moment equations

$$c_k = \int_K \alpha_k \frac{P}{Q} dx, \quad k = 1, 2, \dots, n. \tag{7}$$

have a unique solution $q \in \mathfrak{P}_+$ for each $(c, p) \in \mathfrak{C}_+ \times \mathfrak{P}_+$. Moreover, the solution can be obtained by minimizing the strictly convex functional

$$\mathbb{J}_p^c(q) = \langle c, q \rangle - \int_K P \log Q dx, \tag{8}$$

over all $q \in \mathfrak{P}_+$. This is the dual of the optimization problem to maximize an entropy-like functional

$$\mathbb{I}_p(\Phi) = \int_K P(x) \log \Phi(x) dx \tag{9}$$

over all $\Phi \in \mathfrak{F}_+$ satisfying

$$\int_K \alpha_k(x)\Phi(x) dx = c_k, \quad k = 1, 2, \dots, n, \tag{10}$$

where $\mathfrak{F}_+$ is the class of positive functions in $L_1(K)$.

We note that maximizing (9) is equivalent to minimizing the Kullback-Leibler pseudo-distance given P

$$\mathbb{D}(P\|\Phi) = \int_K P(x) \log \frac{P(x)}{\Phi(x)} dx. \tag{11}$$

In fact,

$$\mathbb{D}(P\|\Phi) = \int_K P(x) \log P(x) dx - \mathbb{I}_p(\Phi). \tag{12}$$

From [33, Theorem 3.4] we have that the map sending $q \in \mathfrak{P}_+$ to $c \in \mathfrak{C}_+$ is a diffeomorphism, so the problem as stated above is well-posed.

3 Regularized approximation

In practice, the moments are often estimated from a finite number of data, for example, the ergodic estimates for the covariance lags, and they may not belong to the dual cone $\mathfrak{C}_+$, and then no solution exists. The problem may be ill-posed also for other reasons. When the data sequence is short, the estimates may contain large errors. Therefore, it is reasonable to match the estimated moments only approximately by allowing an error $d := (d_1, d_2, \dots, d_n)$ in the moment equations so that

$$c_k - \int_K \alpha_k \Phi dx = d_k, \quad k = 1, 2, \dots, n. \tag{13}$$

Then the problem is modified to minimize

$$\frac{1}{2}\|d\|^2 + \lambda \mathbb{D}(P\|\Phi), \tag{14}$$

subject to (13) for some suitable $\lambda > 0$. Here $\lambda\mathbb{D}(P\|\Phi)$ is a regularization term which makes the solution smooth. In view of (12), this problem can be reformulated as the problem to maximize

$$\mathbb{I}(\Phi, d) = \int_K P(x) \log \Phi(x) dx - \frac{1}{2\lambda}\|d\|^2 \tag{15}$$

subject to (13) over all Φ and d. Regularization problems of this type been considered in [39, 40]. Also see [41], where similar results are given.

We assume the condition (6) holds. Modifying the idea of [39, 40] to the setting of [33], we form the Lagrangian

$$\begin{aligned} L(\Phi, d, q) &= \mathbb{I}(\Phi, d) + \sum_{k=1}^{n} q_k \left(c_k - \int_K \alpha_k \Phi dx - d_k \right) \\ &= \int_K P \log \Phi dx - \int_K Q\Phi dx - \frac{1}{2\lambda} d^\top d + \langle c - d, q \rangle \end{aligned} \tag{16}$$

with the directional derivative

$$\delta L(\Phi, d, q; \delta\Phi, \delta d) = \int_K \left(\frac{P}{\Phi} - Q \right) \delta\Phi dx - (\lambda^{-1} d + q)^\top \delta d. \tag{17}$$

For stationarity we require that

$$\Phi = \frac{P}{Q} \quad \text{and} \quad d = -\lambda q, \tag{18}$$

which inserted into $L(\Phi, d, q)$ yields the dual functional

$$\varphi(q) = \mathbb{J}_p(q) + \int_K P(\log P - 1) dx, \tag{19}$$

where the last term is constant and

$$\mathbb{J}_p(q) = \frac{\lambda}{2}\langle q, q \rangle + \langle c, q \rangle - \int_K P \log Q dx. \tag{20}$$

Setting the gradient of $\mathbb{J}_p$ equal to zero, we obtain the moment equations with errors

$$\int_K \alpha_k \frac{P}{Q} dx = c_k + \lambda q_k, \quad k = 1, 2, \dots, n. \tag{21}$$

The regularization parameter λ controls how much error/noise is allowed in the solution. By choosing λ small, the error in the moment equation becomes small. In practice, however, it may be difficult for the algorithm to converge if λ is chosen too small.

We need to show that (21) actually has a solution, which would follow if (20) would have an interior minimum. It is easy to see that (20) is strictly convex.

Lemma 1. *The functional* (20) *has compact sublevel sets* $\mathbb{J}_p^{-1}(-\infty, r]$, $r \in \mathbb{R}$.

Proof. The sublevel set $\mathbb{J}_p^{-1}(-\infty, r]$ is closed, so it remains to prove that it is bounded, i.e., $\alpha = \|Q\|_\infty$ is bounded. Set $Q = \alpha\tilde{Q}$, where $\tilde{Q}(x) \leq 1$. Then we have

$$\mathbb{J}_p(q) = \frac{\lambda}{2}\langle\tilde{q},\tilde{q}\rangle\alpha^2 + \langle c,\tilde{q}\rangle\alpha - \int_K P dx \log\alpha - \int_K P\log\tilde{Q}dx \geq a_0\alpha^2 + a_1\alpha - a_2\log\alpha,$$

where $a_0 := \lambda\langle\tilde{q},\tilde{q}\rangle/2 > 0$, $a_1 := \langle c,\tilde{q}\rangle$ and $a_2 := \int_K P dx > 0$. Hence, if $q \in \mathbb{J}_p^{-1}(-\infty, r]$,

$$a_0\alpha^2 + a_1\alpha - a_2\log\alpha \leq r.$$

Comparing quadratic and logarithmic growth we see that α is bounded from above. Since $\log\alpha \to -\infty$ as $\alpha \to 0$, it is also bounded away from zero. □

Consequently, by strict convexity, (20) has a unique minimum. We have to rule out that this minimum is on the boundary of $\mathfrak{P}_+$. In other words, we need to establish that the minimal point is an interior point so that it satisfies the stationary condition (21).

Lemma 2. *The minimum point of* $\mathbb{J}_p$ *does not lie on the boundary.*

Proof. We proceed along the lines of [13, p.662]. Let $q \in \mathfrak{P}_+$ be arbitrary, and let q_0 be on the boundary. Set $\delta q = q - q_0$ and define $q_\mu = q_0 + \mu\delta q$. Since $q_\mu = \mu q + (1-\mu)q_0$ and $\mathfrak{P}_+$ is convex, it belongs to $\mathfrak{P}_+$ for all $\mu \in (0, 1]$. Next, calculate the directional derivative

$$\delta\mathbb{J}_p(q_\mu,\delta q) = \lambda\langle q_\mu,\delta q\rangle + \langle c,\delta q\rangle - \int_K \frac{P}{Q_\mu}\delta Q dx$$
$$= \langle c + \lambda q_\mu,\delta q\rangle - \int_K R_\mu dx, \text{ where } R_\mu := \frac{P}{Q_\mu}\delta Q.$$

Since

$$\frac{dR_\mu}{d\mu} = -P\frac{(Q-Q_0)^2}{Q_\mu^2} \leq 0,$$

R_μ is monotonically decreasing and converges to $R_0 = P(Q-Q_0)/Q_0$ as $\mu \to 0$. However, by condition (6), R_0 is not integrable, and hence $\delta\mathbb{J}(q_\mu,\delta q) \to -\infty$ as $\mu \to 0$. □

4 Application to image deblurring

We now return to the convolution equation

$$b(x) = \int_K \kappa(x-y)\Phi(y)dy, \tag{22}$$

introduced in Section 1, where κ is the point spread function (PSF), Φ is original image and b is the blurred image. Then setting $c_k := b(x_k)$ and $\alpha_k(x) := \kappa(x_k - x)$, we obtain the moment equations (2). We want to recover the object Φ from the blurred image b given the PSF κ.

After discretization, the blurring process is described by a linear transform plus some additive noise, i.e.,

$$\mathbf{b} = A\mathbf{x} + \eta. \tag{23}$$

Here we have introduced the bold lower-case letters $\mathbf{b}$ and $\mathbf{x}$ to denote the vectorized discretization of the bivariate functions $b(x)$ and $\Phi(x)$, respectively. The blurring matrix A is determined by the PSF and the boundary condition depending on our assumptions of how the picture would be continued outside the image [4, 42, 43].

As pointed out in [1], the continuous inverse problem (22) is ill-posed. Although the problem may become well-posed after discretization, the blurring matrix A is typically ill-conditioned. Due to the presence of the noise term η, the directly inverted solution is unacceptable from the physical point of view. Therefore, regularization must be introduced in order to produce a visually meaningful solution.

Noted that each row of the blurring matrix A is the discrete analogue to the basis function α_k in the formulation of the moment problem. As already mentioned, A is nonsingular although rather close to being singular, and hence its rows are linearly independent. Therefore, linear combination of the basis functions becomes matrix-vector multiplication

$$\mathbf{q} := \text{vec}(Q) = A^\top q, \tag{24}$$

where the matrix Q here is the discretization of the function $Q(x)$, and 'vec' denotes the vectorizing operation for the matrix. Due to the fact that the blurring matrix A is highly structured [4][5], evaluation of the multiplication can be obtained efficiently with 2-dimensional fast Fourier transform (FFT) or discrete cosine transform (DCT), depending on the boundary condition.

4.1 The optimization problem

Using the vectorized notation as in (23) and (24), the discretized objective functional corresponding to (8) can be written as

$$\mathbb{J}_p(q) = \mathbf{b}^\top q - \mathbf{p}^\top\log(A^\top q), \tag{25}$$

where $\mathbf{p}$ here is the discretized prior function P. The vector-valued $\log$ function denotes taking logarithm for each entry of the vector. The reconstructed image

$$\hat{\mathbf{x}} = \mathbf{p}./(A^\top q^*), \tag{26}$$

where q^* is the optimal solution that minimizes (25) and the operation './' means element-wise division.

Consider the vector-valued $\log$ function first. For a matrix $A \in \mathbb{R}^{n\times n}$ and a vector $x \in \mathbb{R}^n$,

$$(\log A^\top x)_i = \log(a_i^\top x),$$

where a_i is the i'th column of A. The elements of the first order derivative (Jacobian) of $\log A^\top x$ are given by

$$\left[\frac{d\log(A^\top x)}{dx}\right]_{ji} = \frac{\partial\log(a_j^\top x)}{\partial x_i} = \frac{a_{ij}}{a_j^\top x},$$

that is, the j'th row of the Jacobian matrix is $a_j^\top/(a_j^\top x)$, so we have

$$\frac{d\log(A^\top x)}{dx} = D_1(x)A^\top,$$

where $D_1(x) := \mathrm{diag}(1/a_j^\top x)$. Consequently,

$$\begin{aligned}\frac{d}{d\tau}\mathbb{J}_p(q+\tau v)|_{\tau=0} &= \mathbf{b}^\top v - \mathbf{p}^\top D_1(q)A^\top v \\ &= \langle \mathbf{b} - AD_1(q)\mathbf{p}, v\rangle,\end{aligned}$$

and therefore the gradient of $\mathbb{J}_p$ is given by

$$\nabla\mathbb{J}_p(q) = \mathbf{b} - AD_1(q)\mathbf{p}. \tag{27}$$

Similarly, for the computation of the Hessian, we form the following

$$\begin{aligned}\frac{\partial^2}{\partial\tau\partial\xi}\mathbb{J}_p(q+\tau v+\xi w)|_{\tau,\xi=0} &= \frac{\partial}{\partial\xi}\left[\mathbf{b}^\top v - \mathbf{p}^\top D_1(y)A^\top v\right] \\ &= \mathbf{p}^\top \mathrm{diag}\left[\frac{a_j^\top w}{(a_j^\top q)^2}\right]A^\top v,\end{aligned}$$

where $y = q + \tau v + \xi w$. We can rewrite

$$\mathbf{p}^\top \mathrm{diag}\left[\frac{a_j^\top w}{(a_j^\top q)^2}\right] = w^\top AD_2(\mathbf{p}, q)$$

in the last term, where $D_2(\mathbf{p}, q) := \mathrm{diag}(\mathbf{p}_j/(a_j^\top q)^2)$. We then have

$$\frac{\partial^2}{\partial\tau\partial\xi}\mathbb{J}_p(q+\tau v+\xi w)|_{\tau,\xi=0} = w^\top AD_2(\mathbf{p}, q)A^\top v.$$

Therefore, the formula for Hessian is

$$\nabla^2\mathbb{J}_p(q) = AD_2(\mathbf{p}, q)A^\top. \tag{28}$$

4.2 Choice of the prior P

Recall that the primal problem to maximize (9) subject to (10) is equivalent to minimizing the Kullback-Leibler divergence (11) subject to the same moment equations. Although the Kullback-Leibler divergence is not a metric, it can be used as a pseudo-distance. In $\mathbb{D}(P, \Phi)$ the function P could be regarded as a prior, and we want the Φ to be "as close as possible" to P in this sense. The choice of P considerably affects the quality of the solution. Choosing $P \equiv 1$ corresponds to no prior information, and the solution is referred to as the *maximum entropy* solution [8]. It is also demonstrated in the literature that the maximum entropy solution is often unsatisfactory. In the setting of image deblurring, the blurred image itself should serve as better prior information.

5 Numerical examples

For the image deblurring problem in the presence of noise we solve the regularized optimization problem to minimize

$$\min_{\mathbf{q}>0} \quad \mathbb{J}_p(q) = \mathbf{b}^\top q - \mathbf{p}^\top \log(A^\top q) + \frac{\lambda}{2}\|q\|^2. \tag{29}$$

The gradient (27) and Hessian (28) are modified a bit as

$$\nabla\mathbb{J}_p(q) = \mathbf{b} - AD_1(q)\mathbf{p} + \lambda q, \tag{30}$$

$$\nabla^2\mathbb{J}_p(q) = AD_2(\mathbf{p}, q)A^\top + \lambda I. \tag{31}$$

Newton's method [44] is used to solve the optimization problem (29).

Two images are chosen for the numerical test. One is the famous Lena with a resolution 256×256 and the other shows a part of the moon surface with a resolution 512×512. The blur type on the test images is out-of-focus and the PSF array is given below with radius $r = 15$:

$$\kappa_{ij} = \begin{cases} 1/(\pi r^2) & \text{if } (i-k)^2 + (j-l)^2 \le r^2, \\ 0 & \text{elsewhere,}\end{cases} \tag{32}$$

where (k, l) is the center of the PSF array. Moreover, a periodic boundary condition is assumed for the Lena image, while a reflexive boundary condition is chosen for the reconstruction of the moon image. The intensity of the noise is characterized by the signal-to-noise ratio (SNR), which is set as 40dB in the test.

The central part of Newton's method is to solve the linear system of equations

$$\nabla^2\mathbb{J}_p\Delta q = \nabla\mathbb{J}_p,$$

for the Newton direction Δq, and we use the conjugate gradient (CG) method [45, 46] to solve it iteratively. In each CG iteration, multiplication with the Hessian is evaluated with 4 two-dimensional FFTs/inverse FFTs (or DCTs), which makes this linear solver the major computational cost of the algorithm. To enforce the positivity constraint on $\mathbf{q} = \mathrm{vec}(Q)$ we restrict the step length τ of the line search in the Newton direction. In fact, we have in the Newton iteration

$$q_+ = q - \tau\Delta q,$$

and therefore

$$\mathbf{q}_+ = A^\top q_+ = A^\top q - \tau A^\top\Delta q = \mathbf{q} - \tau\Delta\mathbf{q},$$

where $\Delta\mathbf{q} := A^\top\Delta q$. The maximum step length is taken as

$$\tau_{\max} = \min\{\mathbf{q}_i/\Delta\mathbf{q}_i | \Delta\mathbf{q}_i > 0\}.$$

With the constraint $0 < \tau < \tau_{\max}$, various line search methods can be used.

The original image and the corresponding blurred one is depicted in Fig. 1 for the Lena image and in Fig. 3 for the moon image. The reconstructed images are shown in Fig. 2 and Fig. 4, respectively. For comparison we also compute the classical Tikhonov reconstruction, where the regularization parameter is chosen with generalized cross-validation (GCV).

In Fig. 2 we see that choosing the blurred image $\mathbf{b}$ as the prior indeed improves the reconstruction. Moreover, the solution of the regularized moment problem looks smoother compared with Tikhonov reconstruction without losing many details. In fact, some reconstruction artifacts are less pronounced. This can also be observed from Fig. 4. However, some work remains to perfect this method.

6 Open problems and future research

A question worth investigating is whether exchanging $\|d\|^2$ in (14) for a more general positive definite form $d^\top Wd$

Fig. 1: Lena: original sharp image and the blurred one

Fig. 2: Reconstructed images, Tikhonov method (left), and solutions of the moment problem, $p = 1,\ \lambda = 12$ (middle), and $p = b,\ \lambda = 0.11$ (right).

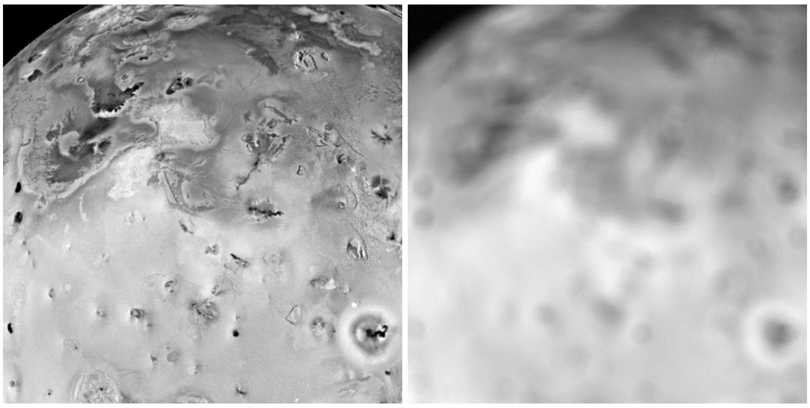

Fig. 3: Moon: original sharp image and the blurred one

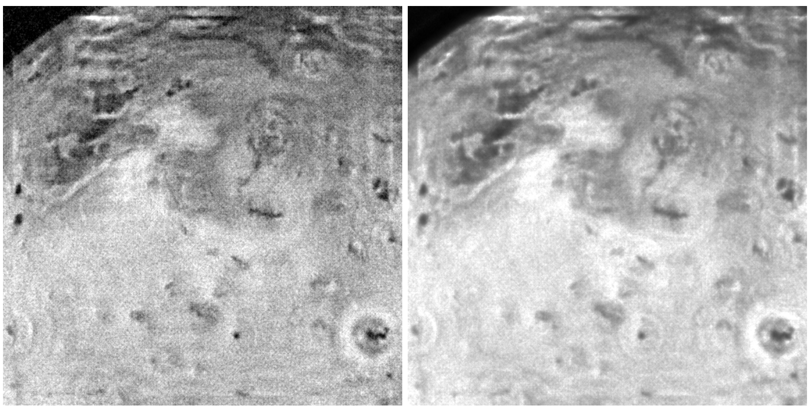

Fig. 4: Reconstructed images, Tikhonov method (left), and solutions of the moment problem with $p = b,\ \lambda = 0.4$ (right).

giving different weights to the error components could improve the reconstruction.

An obvious downside is that the number of basis functions is very high. One could investigate whether including a sparsity promoting regularization term in the cost function could improve numerics.

Instead of using the blurred image as a prior one could try to modify the procedure in the style of [19, 20] to use estimated logarithmic moments. How to actually construct such estimates is however an open question.

References

[1] M. Bertero, and P. Boccacci, *Introduction to Inverse Problems in Imaging*, IOP Publishing Ltd, 1998.

[2] D. Commenges, The deconvolution problem: fast algorithms including the preconditioned conjugate-gradient to compute a MAP estimator, *IEEE Trans. on Automatic Control*, 29(3): 229–243, 1984.

[3] T. F. Chan, and J. Shen, *Image processing and analysis: variational, PDE, wavelet, and stochastic methods*, Society for Industrial and Applied Mathematics, Philadelphia, 2005.

[4] P. C. Hansen, J. G. Nagy, and D. P. O'leary, *Deblurring Images: Matrices, Spectra, and Filtering*, Society for Industrial and Applied Mathematics, Philadelphia, 2006.

[5] C. R. Vogel, *Computational Methods for Inverse Problems*, Society for Industrial and Applied Mathematics, Philadelphia, 2002.

[6] T. F. Chan, G. H. Golub, and P. Mulet,A nonlinear primal-dual method for total variation-based image restoration, *SIAM J. Sci. Comput.*, 20(6): 1964–1977, 1999.

[7] M. Hanke, J. G. Nagy, and C. Vogel, Quasi-Newton approach to nonnegative image restorations, *Linear Algebra and its Applications*, 316: 223–236, 2000.

[8] A. Lindquist, and G. Picci, *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification*, Springer, Heidelberg, New York, Dordrecht and London, 2015.

[9] B. Wahlberg, Systems identification using Laguerre models, *IEEE Trans. Automatic Control*, 36: 551–562, 1991.

[10] R. E. Kalman, Realization of covariance sequences, *Proc. Toeplitz Memorial Conference*, Tel Aviv, Israel, 1981.

[11] T. T. Georgiou, Realization of power spectra from partial covariance sequences, *IEEE Trans. Acoustics, Speech and Signal Processing*, 35: 438–449, 1987.

[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. Automatic Control*, 40: 1841–1857, 1995.

[13] C. I. Byrnes, S. V. Gusev, and A. Lindquist, A convex optimization approach to the rational covariance extension problem, *SIAM J. Control and Opimization*, 37: 211–229, 1998.

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

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

[16] C. I. Byrnes, T. T. Georgiou, and A. Lindquist, A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint, *IEEE Trans. on automatic control*, 46(5): 822–839, 2001.

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

[18] C. I. Byrnes, and A. Lindquist, The moment problem for rational measures: convexity in the spirit of Krein, in *Modern Analysis and Application: Mark Krein Centenary Conference, Vol. I: Operator Theory and Related Topics*, Book Series: Operator Theory Advances and Applications, Vol. 190, 157–169, Birkhäuser, 2009.

[19] C. I. Byrnes, P. Enqvist, and A. Lindquist, Cepstral coefficients, covariance lags and pole-zero models for finite data strings, *IEEE Trans. on Signal Processing*, 50: 677–693, 2001.

[20] 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: 23–59, 2002.

[21] A. Blomqvist, A. Lindquist, and R. Nagamune, Matrix-valued Nevanlinna-Pick interpolation with complexity constraint: An optimization approach, *IEEE Trans. Autom. Contr.*, 48: 2172-2190, 2003.

[22] C. I. Byrnes, G. Fanizza, and A. Lindquist, A homotopy continuation solution of the covariance extension equation, in *New Directions and Applications in Control Theory*, 27–42, Springer Verlag, 2005.

[23] C. I. Byrnes, and A. Lindquist, Important moments in systems and control, *SIAM J. Control and Optimization*, 47(5): 2458–2469, 2008.

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

[25] T. T. Georgiou, and A. Lindquist, A convex optimization approach to ARMA modeling, *IEEE Trans. Automatic Control*, 53: 1108–1119, 2008.

[26] C. I. Byrnes, T. T. Georgiou, A. Lindquist, and A. Megretski, Generalized interpolation in H^∞ with a complexity constraint, *Trans. of the American Mathematical Society*, 358(3): 965–988, 2006.

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

[28] M. Pavon and A. Ferrante, On the Georgiou-Lindquist approach to constrained Kullback-Leibler approximation of spectral densities, *IEEE Trans. on Automatic Control*, 51(4): 639–644, 2006.

[29] A. Ferrante, M. Pavon, and M. Zorzi, A maximum entropy enhancement for a family of high-resolution spectral estimators, *IEEE Trans. on Automatic Control*, 57(2): 318–329, 2012.

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

[31] A. Lindquist, and G. Picci, The circulant rational covariance extension problem: the complete solution, *IEEE Trans. on automatic control*, 58(11): 2848–2861, 2013.

[32] A. Lindquist, C. Masiero, and G. Picci, On the multivariate circulant rational covariance extension problem, *Proc. 52st IEEE Conf. Decision and Control*, 7155–7161, 2013.

[33] J. Karlsson, A. Lindquist, and A. Ringh, The multidimensional moment problem with complexity constraint, ArXiv e-prints, 2015.

[34] A. Ringh, J. Karlsson, and A. Lindquist, Multidimensional rational covariance extension with applications to spectral estimation and image compression, ArXiv e-prints, 2015.

[35] T. T. Georgiou, Relative Entropy and the multi-variable multi-dimensional Moment Problem, *IEEE Trans. on Information Theory*, 52(3): 1052–1066, 2006.

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

[37] S. W. Lang, and J. H. McClellan, Multidimensional MEM

spectral estimation, *IEEE Trans. on Acoustics, Speech and Signal Processing*, 30: 880–887, 1982.

[38] J. H. McClellan and S. W. Lang, Multidimensional MEM spectral estimation, in *Proceedings of the Institute of Acoustics "Spectral Analysis and its Use in Underwater Acoustics": Underwater Acoustics Group Conference*, 10.1–10.8, Imperial College, London, 1982.

[39] P. Enqvist, and E. Avventi, Approximative covariance interpolation with a quadratic penalty, *Proceedings of the 46th IEEE Conference on Decision and Control*, 4275–4280, 2007.

[40] E. Avventi, and P. Enqvist, Approximative linear and logarithmic interpolation of spectra, chapter in Avventi's PhD thesis.

[41] A. Ringh, J. Karlsson, and A. Lindquist, Multidimensional rational covariance extension with approximate covariance matching, *Extended abstract for MTNS16*, Minneapolis, 2016.

[42] J. G. Nagy, K. Palmer, and L. Perrone, Iterative methods for image deblurring: a Matlab object-oriented approach, *Numerical Algorithm*, 36: 79–93, 2004.

[43] M. K. Ng, R. H. Chan, and W. C. Tang,A fast algorithm for deblurring models with Neumann boundary conditions, *SIAM J. Sci. Comput.*, 21(3): 851-866, 1999.

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

[45] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Publishing Company, 1996.

[46] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Society for Industrial and Applied Mathematics, Philadelphia, 1997.

[47] A. Ringh, and J. Karlsson, A fast solver for the circulant rational covariance extension problem, in *European Control Conference*, 727–733, 2015.