

KERNEL DENSITY ESTIMATION VIA DIFFUSION AND THE COMPLEX EXPONENTIALS APPROXIMATION PROBLEM

BY

PIERO BARONE

Istituto per le Applicazioni del Calcolo “M. Picone”, C.N.R., Via dei Taurini 19, 00185 Rome, Italy

Abstract. A kernel method is proposed to estimate the condensed density of the generalized eigenvalues of pencils of Hankel matrices whose elements have a joint non-central Gaussian distribution with nonidentical covariance. These pencils arise when the complex exponentials approximation problem is considered in Gaussian noise. Several moments problems can be formulated in this framework, and the estimation of the condensed density above is the main critical step for their solution. It is shown that the condensed density satisfies approximately a diffusion equation, which allows us to estimate an optimal bandwidth. It is proved by simulation that good results can be obtained even when the signal-to-noise ratio is so small that other methods fail.

Introduction. Many difficult moments problems such as the trigonometric, the complex, and the Hausdorff ones can be formulated as the complex exponentials approximation problem (CEAP), which can be stated as follows, denoting random variables by bold characters. Given a uniformly sampled signal made up of a linear combination of complex exponentials

$$s_k = \sum_{j=1}^{p^*} c_j \xi_j^k, \quad (0.1)$$

where $c_j, \xi_j \in \mathbb{C}$, let us assume to know an even number $n = 2p$, $p \geq p^*$ of noisy samples

$$\mathbf{d}_k = s_k + \epsilon_k, \quad k = 0, 1, 2, \dots, n - 1$$

where ϵ_k is a complex Gaussian, zero mean, white noise, with finite known variance σ^2 . The CEAP problem consists of estimating p^*, c_j, ξ_j , $j = 1, \dots, p^*$. This is a well-known ill-posed inverse problem often addressed in the literature; see, e.g., [15, 17, 18, 23]. More specifically, in [12, 13], a method based on the solution of many CEAP problems was developed to reconstruct high resolution Magnetic Resonance (MR) images from a small set of data, cutting the permanence time of a patient in the MR machine by

Received June 20, 2012.

2010 *Mathematics Subject Classification.* Primary 62G07, 41A30.

Key words and phrases. Condensed density, random matrices, parabolic PDE.

E-mail address: p.barone@iac.cnr.it, piero.barone@gmail.com

©2014 Brown University

Reverts to public domain 28 years from publication

a factor of about three and improving spatial resolution at the same time. The basic idea was that a statistical approach based on replicated measurements could overcome the bad reputation of standard methods to solve the CEAP. We then developed several techniques to generate pseudosamples from the only measured sample and exploited the strong nonlinearity of the critical parameters ξ_j to build powerful estimation methods [4–11, 20, 22, 24]. In this paper a new method is proposed to estimate a density function, related to the nonlinear parameters ξ_j , which turns out to be the basic stone of all the above quoted methods.

We notice that, in the noiseless case and when $p = p^*$, the parameters ξ_j are the generalized eigenvalues of the pencil (U_1, U_0) where U_1 and U_0 are Hankel matrices defined as

$$U_0 = \begin{bmatrix} s_0 & s_1 & \dots & s_{p-1} \\ s_1 & s_2 & \dots & s_p \\ \cdot & \cdot & \dots & \cdot \\ s_{p-1} & s_p & \dots & s_{n-2} \end{bmatrix}, \quad U_1 = \begin{bmatrix} s_1 & s_2 & \dots & s_p \\ s_2 & s_3 & \dots & s_{p+1} \\ \cdot & \cdot & \dots & \cdot \\ s_p & s_{p+1} & \dots & s_{n-1} \end{bmatrix}. \tag{0.2}$$

If we define \mathbf{U}_0 and \mathbf{U}_1 as U_0 and U_1 but start from \mathbf{d}_k , $k = 0, \dots, n-1$, it is evident that the generalized eigenvalues of the pencil $(\mathbf{U}_1, \mathbf{U}_0)$ provide information about the location in the complex plane of the generalized eigenvalues ξ_j , $j = 1 \dots, p$, whose estimation is the most difficult part of CEAP.

To make precise this statement, let us consider a $p \times p$ random matrix \mathbf{U} and denote by $\{\xi_j, j = 1, \dots, p\}$ its eigenvalues which form a set of exchangeable random variables. Their marginal density $h(z)$, $z \in \mathbb{C}$, also called condensed density [21], is the expected value of the (random) normalized counting measure on the zeros of \mathbf{U} , i.e.,

$$h(z) = \frac{1}{p} E \left[\sum_{j=1}^p \delta(z - \xi_j) \right]$$

or, equivalently, for all Borel sets $A \subset \mathbb{C}$,

$$\int_A h(z) dz = \frac{1}{p} \sum_{j=1}^p Prob(\xi_j \in A).$$

If a pencil $\mathbf{U} = (\mathbf{U}_1, \mathbf{U}_0)$ of random matrices is considered, the condensed density of its generalized eigenvalues can be computed by the formula above. Its relative maxima provide information about the location in the complex plane of the generalized eigenvalues ξ_j , $j = 1 \dots, p$. In [4–6], the use of this function for solving several moments problems was illustrated. However, when the signal-to-noise ratio (SNR) measured, e.g., by $SNR = \min_j \frac{|\xi_j|^2}{\sigma^2}$ is low, the computation of the condensed density is very difficult even if we assume to have replicated observations as we do in the following. The main problem is that many relative maxima related to noise are present in the condensed density which are slightly less intense than those related to the true signals. As we are interested only in signal-related relative maxima of the condensed density, we look for an estimation method with noise-filtering abilities.

The aim of the paper is to show that it is possible to devise a kernel density estimation method which does have such properties. In [2], the same project was developed for

the real exponentials approximation problem which is better known as the exponential analysis problem. The idea was to use a method proposed by Botev et al. [14], based on a class of kernels which satisfy a diffusion PDE, which allows us to automatically estimate the optimal bandwidth. It was shown that the generalized eigenvalues of the pencil \mathbf{U} can be approximated by the ratio of Gaussian variables whose density satisfies a specific diffusion PDE. In the limit for $SNR \downarrow \infty$, this PDE belongs to the class considered in [14] and their method can be used.

In the following, a different approach is proposed. An explicit expression of the condensed density is derived and approximated by the Laplace method. Then an anisotropic diffusion PDE is derived which admits the approximated condensed density as a solution in the limit for $\sigma \downarrow 0$. Finally, an optimal bandwidth is derived on the same lines of the Botev et al. results. We can then propose an estimator which has better filtering abilities than the standard one based on Gaussian kernel.

The paper is organized as follows. In Section 1, the condensed density of the generalized eigenvalues of \mathbf{U} is derived. In Section 2, its Laplace approximation is computed. In Section 3, the anisotropic diffusion PDE is derived. In Section 4, the kernel estimator and the optimal bandwidth are computed. In Section 5, the proposed algorithm is illustrated. Finally in Section 6, two numerical examples are discussed. The Appendix contains the proof of most theorems and lemmas.

1. The condensed density of the generalized eigenvalues of U . From [6, eq.(7)] the condensed density of the generalized eigenvalues of \mathbf{U} is

$$h_n(z, \sigma) = \frac{2}{n} E \left[\sum_{j=1}^{n/2} \delta(z - \xi_j) \right] = \frac{2}{n} \sum_{j=1}^{n/2} E [\delta(z - \xi_j)] = \frac{2}{n} \sum_{j=1}^{n/2} h_n^{(j)}(z, \sigma)$$

where

$$h_n^{(j)}(z, \sigma) = \frac{1}{(\pi\sigma^2)^n} \int_{\mathbb{C}^{n/2-1}} \int_{\mathbb{C}^{n/2}} J_C^*(\underline{\zeta}^{(j)}, z, \underline{\gamma}) e^{-\frac{1}{\sigma^2} \sum_{k=0}^{n-1} |\sum_{h \neq j}^{1, n/2} \gamma_h \zeta_h^k + \gamma_j z^k - s_k|^2} d\underline{\zeta}^{(j)} d\underline{\gamma},$$

$\underline{\zeta}^{(j)} = \{\zeta_h, h \neq j\}$, and

$$J_C^*(\underline{\zeta}^{(j)}, z, \underline{\gamma}) = \begin{cases} \gamma & \text{if } n = 2, \\ (-1)^{n/2} \prod_{j=1}^{1, n/2} \gamma_j \prod_{r < h, r, h \neq j} (\zeta_r - \zeta_h)^4 \prod_{r \neq j} (\zeta_r - z)^4 & \text{if } n \geq 4. \end{cases}$$

In the following Lemmas we give simpler forms and properties of the condensed density.

LEMMA 1.1.

$$h_n^{(j)}(z, \sigma) = \frac{1}{(\pi\sigma^2)^n} \int_{\mathbb{C}^{n/2-1}} \int_{\mathbb{C}^{n/2}} J_C^* e^{-\frac{1}{\sigma^2} [(\underline{\gamma} - \underline{\mu}_j)^H Q_j (\underline{\gamma} - \underline{\mu}_j) + \nu_j]} d\underline{\zeta}^{(j)} d\underline{\gamma},$$

$$J_C^* = J_C^*(\underline{\zeta}^{(j)}, z, \underline{\gamma}), \quad \underline{\mu}_j = \underline{\mu}_j(\underline{\zeta}^{(j)}, z), \quad Q_j = Q_j(\underline{\zeta}^{(j)}, z), \quad \nu_j = \nu_j(\underline{\zeta}^{(j)}, z),$$

and

$$x_{hk}^{(j)} = \begin{cases} \zeta_h^{k-1}, & h \neq j \\ z^{k-1}, & h = j \end{cases}, \quad \underline{s} = [s_0, \dots, s_{n-1}],$$

$$X_j(h, k) = \bar{x}_{hk}^{(j)}, \quad q_j = X_j \underline{s}, \quad X_j \in \mathbb{C}^{n/2 \times n}, \quad Q_j = X_j X_j^H \in \mathbb{C}^{n/2 \times n/2},$$

$$\underline{\mu}_j = Q_j^{-1} \underline{q}_j, \quad \nu_j = \sum_{k=0}^{n-1} |s_k|^2 - \underline{\mu}_j^H Q_j \underline{\mu}_j = \underline{s}^H (I_n - X_j^H (X_j X_j^H)^{-1} X_j) \underline{s}.$$

LEMMA 1.2. If

$$X(h, k) = \bar{\xi}_h^{k-1}, \quad \underline{s} = X^H \underline{c}, \quad \nu(\underline{\xi}) = \underline{s}^H (I_n - X^H (X X^H)^{-1} X) \underline{s},$$

then $\nu(\underline{\xi}) = 0$.

Proof. $\nu(\underline{\xi}) = \underline{s}^H (I_n - X^H (X X^H)^{-1} X) \underline{s} = \underline{c}^H X (I_n - X^H (X X^H)^{-1} X) X^H \underline{c} = 0$.

LEMMA 1.3. When $n = 2$, the condensed density is given by

$$h_2^{(1)}(z, \sigma) = e^{-\rho \frac{|z - \xi_1|^2}{1 + |z|^2}} \left(\frac{\rho |1 + \bar{z} \xi_1|^2}{\pi (1 + |z|^2)^3} + \frac{1}{\pi (1 + |z|^2)^2} \right)$$

where $\rho = \frac{|c_1|^2}{\sigma^2}$ denotes the SNR.

LEMMA 1.4. When $n > 2$,

$$h_n^{(j)}(z, \sigma) = \int_{\mathbb{R}^{n-2}} f_j(\underline{\zeta}^{(j)}, z) \frac{\prod_{r < h, r, h \neq j} |\zeta_r - \zeta_h|^8 \prod_{r \neq j} |\zeta_r - z|^8 e^{-\frac{1}{\sigma^2} \nu_j(\underline{\zeta}^{(j)}, z)} d\Re \underline{\zeta}^{(j)} d\Im \underline{\zeta}^{(j)}}{(\pi \sigma^2)^{n/2} |\tilde{Q}_j(\underline{\zeta}^{(j)}, z)|^{\frac{1}{2}}}$$

where

$$f_j(\underline{\zeta}^{(j)}, z, \sigma) = \mathbb{E} \left[\prod_{i=1}^{1, n/2} \tilde{\gamma}^T A_i \tilde{\gamma} \right], \quad A_i = I_2 \otimes \underline{e}_i \underline{e}_i^T.$$

\mathbb{E} denotes the expectation with respect to the Gaussian density $N(\underline{\tilde{\mu}}_j, \Sigma_j)$ where $\Sigma_j = \frac{\sigma^2}{2} \tilde{Q}_j^{-1}$,

$$\tilde{Q}_j = \begin{bmatrix} \Re Q_j & -\Im Q_j \\ \Im Q_j & \Re Q_j \end{bmatrix},$$

and $\tilde{\gamma}, \underline{\tilde{\mu}}_j$ are obtained by stacking the real and imaginary parts of $\underline{\gamma}$ and $\underline{\mu}_j$ respectively.

LEMMA 1.5.

$$f_j(\underline{\zeta}^{(j)}, z, \sigma) = P_{n/2}(\sigma^2, z) = \sum_{h=0}^{n/2} \beta_h(\underline{\zeta}^{(j)}, z) \frac{\sigma^{2h}}{2^h D_j^{n-h}}$$

where $\beta_h(\underline{\zeta}^{(j)}, z)$ are positive polynomials and $D_j = \det(\tilde{Q}_j)$. Moreover,

$$\beta_0(\underline{\zeta}^{(j)}, z) = \prod_{i=1}^{n/2} \tilde{q}_j^T \hat{Q}_j A_i \hat{Q}_j \tilde{q}_j,$$

and $\beta_{n/2}(\underline{\zeta}^{(j)}, z)$ is the only coefficient that does not depend on \underline{s} .

COROLLARY 1.6. When $n > 2$,

$$h_n^{(j)}(\tilde{z}, \sigma) = \int_{\mathbb{R}^{n-2}} g_j(\tilde{\zeta}^{(j)}, \tilde{z}, \sigma) e^{-\frac{1}{\sigma^2} \nu_j(\tilde{\zeta}^{(j)}, \tilde{z})} d\tilde{\zeta}^{(j)} \tag{1.1}$$

where \tilde{z} and $\tilde{\zeta}^{(j)}$ are the vectors obtained by stacking the real and imaginary parts of z and $\zeta^{(j)}$ respectively, and

$$g_j(\tilde{\zeta}^{(j)}, \tilde{z}, \sigma) = \frac{1}{\sigma^n \pi^{n/2}} \sum_{k=0}^{n/2} \frac{\sigma^{2k}}{2^k} \beta_k(\tilde{\zeta}^{(j)}, \tilde{z}) \frac{\prod_{r<h;r,h \neq j} |\tilde{\zeta}_r - \tilde{\zeta}_h|^8 \prod_{r \neq j} |\tilde{\zeta}_r - \tilde{z}|^8}{D_j^{n-k+\frac{1}{2}}(\tilde{\zeta}^{(j)}, \tilde{z})}.$$

COROLLARY 1.7. When $n > 2, \forall j = 1, \dots, n/2,$

$$\begin{aligned} h_n^{(j)}(\tilde{z}, \infty) &= h_n^{(1)}(\tilde{z}, \infty) \\ &= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^{n-2}} \beta_{n/2}(\tilde{\zeta}^{(1)}, \tilde{z}) \frac{\prod_{r<h;r,h \neq 1} |\tilde{\zeta}_r - \tilde{\zeta}_h|^8 \prod_{r \neq 1} |\tilde{\zeta}_r - \tilde{z}|^8}{D_1^{\frac{n+1}{2}}(\tilde{\zeta}^{(1)}, \tilde{z})} d\tilde{\zeta}^{(1)}. \end{aligned}$$

Moreover, this is also the condensed density obtained when $\underline{s} = \underline{0}$, i.e., when $\underline{c} = \underline{0}$ and it is circularly symmetric, i.e., it depends only on $|\tilde{z}|^2$.

2. The Laplace approximation of the condensed density. The expression of the condensed density given in Corollary 1.6 can be approximated in the limit for $\sigma \downarrow 0$ as follows.

THEOREM 2.1. In the limit for $\sigma \downarrow 0$ for $\tilde{z} \in \mathcal{N}_j$ where \mathcal{N}_j is a neighbor of ξ_j , we have

$$h_n^{(j)}(\tilde{z}, \sigma) \approx \hat{h}_n^{(j)}(\tilde{z}, \sigma) = \sigma^{-2} G_j(\hat{\zeta}_j^{(j)}, \tilde{z}) e^{-\frac{1}{\sigma^2} \nu_j(\hat{\zeta}_j^{(j)}, \tilde{z})} \tag{2.1}$$

where

$$\begin{aligned} G_j(\hat{\zeta}_j^{(j)}, \tilde{z}) &= (2\pi)^{\frac{n}{2}-1} |H_j(\hat{\zeta}_j, \tilde{z})|^{-\frac{1}{2}} K_j(\hat{\zeta}_j, \tilde{z}) \\ K_j(\tilde{\zeta}^{(j)}, \tilde{z}) &= \frac{1}{\pi^{\frac{n}{2}}} \beta_0(\tilde{\zeta}^{(j)}, \tilde{z}) \frac{\prod_{r<h;r,h \neq j} |\tilde{\zeta}_r - \tilde{\zeta}_h|^8 \prod_{r \neq j} |\tilde{\zeta}_r - \tilde{z}|^8}{D_j^{n+\frac{1}{2}}(\tilde{\zeta}^{(j)}, \tilde{z})}. \end{aligned}$$

H_j is the Hessian of $\nu_j, \hat{\zeta}_j$ is the unique minimum of $\nu_j(\tilde{\zeta}^{(j)}, \xi_j)$ in a neighbor \mathcal{D}_j of $\{\xi_h, h \neq j\}$, and $\beta_0(\tilde{\zeta}^{(j)}, \tilde{z})$ is given in Lemma 1.5.

3. The diffusion equation. In [6, Th. 2] it was proved that $h_n(z, \sigma)$ converges weakly to the positive measure $\frac{2}{n} \sum_{h=1}^p \delta(z - \xi_h)$ when $\sigma \downarrow 0$. By the definition of $h_n^{(j)}(z, \sigma)$, the same proof implies that $h_n^{(j)}(z, \sigma)$ converges weakly to $\delta(z - \xi_j)$. Moreover, from Corollary 1.2 we know that for $\sigma \downarrow \infty, h_n^{(j)}(z, \sigma)$ converges to a density independent of ξ . Therefore we can guess that when σ moves from 0 to $\infty, h_n^{(j)}(z, \sigma)$ diffuses from an atomic measure centered in ξ_j to a measure circularly symmetric w.r. to zero. We then look for a diffusion equation which admits $\hat{h}_n^{(j)}(z, \sigma)$ as solution for $\sigma \downarrow 0$.

Let us define $z = x + iy$ and consider the anisotropic diffusion $h_t = L[h]$ where

$$L[h](x, y, t) = \operatorname{div} \left[a_j(x, y) \nabla \left(\frac{h(x, y, t)}{p(x, y)} \right) \right], \tag{3.1}$$

$p(x, y) = h_n^{(j)}(x, y, \infty)$ is the stationary probability density, and $a_j(x, y) > 0$ is the unknown diffusion coefficient. Substituting $h(x, y, t)$ in the equation above with $\hat{h}_n^{(j)}(x, y, \sigma)$,

$t = \sigma^2$ where now σ is considered as a variable (not a fixed known value), and dropping the indices j, n and the variables (x, y) , we get

$$\frac{C e^{-\frac{\nu}{t}}}{p^3} \left(\frac{E_1}{t} + \frac{E_2}{t^2} + \frac{E_3}{t^3} \right) = 0 \tag{3.2}$$

where

$$\begin{aligned} E_1 &= -a_y G_y p^2 - a_x G_x p^2 + a_y G p_y p + a_x G p_x p + 2ap (G_y p_y + G_x p_x) \\ &\quad - a (G_{yy} + G_{xx}) p^2 + aG (p_{yy} + p_{xx}) p - 2aG (p_y^2 + p_x^2), \\ E_2 &= a_y G p^2 \nu_y + a_x G p^2 \nu_x + 2aG_y p^2 \nu_y + 2aG_x p^2 \nu_x - 2aG p_y p \nu_y \\ &\quad - 2aG p_x p \nu_x + aG p^2 (\nu_{yy} + \nu_{xx}) - G p^3, \\ E_3 &= -aG p^2 \nu_y^2 - aG p^2 \nu_x^2 + G p^3 \nu. \end{aligned}$$

In the limit for $t \downarrow 0$, the dominant term on the left side of the equation is $\frac{C e^{-\frac{\nu}{t}}}{p^3} \frac{E_3}{t^3}$; therefore, the equation is approximately satisfied when $E_3 = 0$ or, equivalently, when

$$a(x, y) = \frac{p(x, y)\nu(x, y)}{\nu_x(x, y)^2 + \nu_y(x, y)^2} > 0.$$

In the following Lemma we prove that the Csiszár distance between $h(x, y, t)$ and $p(x, y)$, defined as

$$\mathcal{D}(h, p) = \int_{\mathbb{R}^2} p(x, y) \Psi \left(\frac{h(x, y, t)}{p(x, y)} \right) dx dy, \quad \Psi \in \mathcal{C}^2 : \mathbb{R}^+ \rightarrow \mathbb{R}^+, \quad \Psi''(\cdot) > 0, \quad \Psi'(1) = 0,$$

is a monotonic decreasing function of t ; therefore, $h(x, y, t)$ tends monotonically to $p(x, y)$ when $t \downarrow \infty$.

LEMMA 3.1.

$$\frac{\partial \mathcal{D}(h, p)}{\partial t} = - \int_{\mathbb{R}^2} a(x, y) \Psi'' \left(\frac{h(x, y, t)}{p(x, y)} \right) \left\| \nabla \left(\frac{h(x, y, t)}{p(x, y)} \right) \right\|_2^2 dx dy < 0.$$

4. The kernel estimator. Given a sample of size R of the data

$$\underline{\mathbf{d}}^{(r)} = [\mathbf{d}_1^{(r)}, \dots, \mathbf{d}_n^{(r)}], \quad r = 1, \dots, R$$

where $\mathbb{E}[\mathbf{d}_k^{(r)}] = s_k$, we consider the kernel estimator of $h_n^{(j)}(z, t)$ with bandwidth t given by

$$\mathbf{h}_n^{(j)}(x, y, t) = \frac{1}{R} \sum_{r=1}^R \Phi_j(x, y, \zeta_j(r); t)$$

where $\Phi_j(x, y, \zeta_j(r); t)$ is a solution of the diffusion equation (3.1) at time t with initial condition $\delta(z - \zeta_j(r))$ and $\zeta_j(r), j = 1, \dots, n/2$ are the generalized eigenvalues obtained from the data $\underline{\mathbf{d}}^{(r)}$. Hence $\mathbf{h}_n^{(j)}(x, y, t)$ is a solution of the diffusion equation (3.1) at time t with initial condition

$$\mathbf{E}_j(x, y) = \frac{1}{R} \sum_{r=1}^R \delta(z - \zeta_j(r)). \tag{4.1}$$

Therefore $\mathbf{E}_j(x, y)$ is the empirical distribution of the generalized eigenvalue ζ_j and

$$\mathbb{E}[\mathbf{E}_j(z)] = \frac{1}{R} \sum_{r=1}^R \int_{\mathbb{C}} \delta(z - \zeta) h_n^{(j)}(\zeta) d\zeta = h_n^{(j)}(z)$$

where $h_n^{(j)}(z)$ is the unknown true density corresponding to the known fixed value of σ .

In order to find the optimal bandwidth, we need the form of the kernel $\Phi_j(x, y, \mathbf{d}^{(r)}; t)$ for $t \downarrow 0$. By construction, dropping the index r , we have

$$\Phi_j(x, y, \zeta; t) = \hat{h}_n^{(j)}(x, y, t) = t^{-1} G_j(x, y, \zeta) e^{-\frac{1}{t} \nu_j(x, y, \zeta)}, \quad t \downarrow 0.$$

Associated to the anisotropic diffusion (3.1), there exists a Markov process \mathbf{X}_t whose transition probabilities are given, when $t \downarrow 0$, by $\Phi_j(z, \zeta; t)$. Moreover, when the initial density of \mathbf{X}_0 is $\delta(z - \zeta)$, the density of \mathbf{X}_t is ([16, eq. (5.1), Ch.X.5])

$$\Phi_j(z, \zeta; t) = \int_{\mathbb{C}} \Phi_j(z, u; t) \delta(u - \zeta) du.$$

Therefore the kernel $\Phi_j(z, \zeta; t)$, $\forall \zeta$ and $t > 0$, satisfies the forward equation

$$\begin{cases} \frac{\partial \Phi_j}{\partial t}(z, \zeta; t) - L[\Phi_j(z, \zeta; t)] = 0 \\ \Phi_j(z, \zeta; 0) = \delta(z - \zeta) \end{cases} \quad (4.2)$$

Moreover, the conditional expectation of $\delta(u - \mathbf{X}_t)$ on the hypothesis that $\mathbf{X}_0 = z$ is ([16, eq. (4.5), Ch.X.4])

$$\Phi_j(z, \zeta; t) = \int_{\mathbb{C}} \Phi_j(u, \zeta; t) \delta(u - z) du$$

and therefore the kernel $\Phi_j(z, \zeta; t)$, $\forall z$ and $t > 0$, satisfies the backward equation

$$\begin{cases} \frac{\partial \Phi_j}{\partial t}(z, \zeta; t) - L^*[\Phi_j(z, \zeta; t)] = 0 \\ \Phi_j(z, \zeta; 0) = \delta(z - \zeta) \end{cases} \quad (4.3)$$

where

$$L[\Phi_j(z, \zeta; t)] = \operatorname{div}_z \left[a_j(z) \nabla_z \left(\frac{\Phi_j(z, \zeta; t)}{p_j(z)} \right) \right]$$

i.e.

$$L[\Phi_j(z, \zeta; t)] = a_j(z) \Delta_z \left[\frac{\Phi_j(z, \zeta; t)}{p_j(z)} \right] + (\nabla_z a_j(z))^T \nabla_z \left[\frac{\Phi_j(z, \zeta; t)}{p_j(z)} \right]$$

and

$$L^*[\Phi_j(z, \zeta; t)] = \Delta_\zeta \left[a_j(\zeta) \frac{\Phi_j(z, \zeta; t)}{p_j(\zeta)} \right] - \operatorname{div}_\zeta \left[\frac{\Phi_j(z, \zeta; t)}{p_j(\zeta)} \nabla_\zeta a_j(\zeta) \right]$$

is the adjoint operator of L , and div_z and ∇_z denote respectively the divergence and the gradient operators w.r. to the variable z .

The mean integrated squared error (MISE) criterion to determine an optimal bandwidth t is given, dropping the index n , by

$$MISE_{\mathbf{h}^{(j)}}(t) = \mathbb{E}_{h^{(j)}} \int_{\mathbb{C}} [\mathbf{h}^{(j)}(z, t) - h^{(j)}(z, t)]^2 dz$$

or

$$MISE_{\mathbf{h}^{(j)}}(t) = \int_{\mathbb{C}} [\mathbb{E}_{h^{(j)}}\{\mathbf{h}^{(j)}(z, t)\} - h^{(j)}(z)]^2 dz + \int_{\mathbb{C}} Var_h[\mathbf{h}^{(j)}(z, t)] dz.$$

Following Botev et al. [14], we have

THEOREM 4.1.

$$\int_{\mathbb{C}} [\mathbb{E}_{h^{(j)}}\{\mathbf{h}^{(j)}(z, t)\} - h^{(j)}(z)]^2 dz \approx t^2 \|L[h^{(j)}]\|^2$$

$$\int_{\mathbb{C}} Var_h[\mathbf{h}^{(j)}(z, t)] dz \approx \frac{1}{2Rt} \mathbb{E}_{h^{(j)}}[G_j].$$

The MISE is then

$$MISE_{\mathbf{h}^{(j)}}(t) \approx t^2 \|L[h^{(j)}]\|^2 + \frac{1}{2Rt} \mathbb{E}_{h^{(j)}}[G_j]$$

which has a unique real positive minimum in

$$t_j = \sqrt[3]{\frac{\mathbb{E}_{h^{(j)}}[G_j]}{4R\|L[h^{(j)}]\|^2}}. \tag{4.4}$$

The optimal kernel estimator of $h_n^{(j)}(z, t)$ is then given by

$$\mathbf{h}_n^{(j)}(z, t_j) = \frac{1}{R} \sum_{r=1}^R \Phi_j(z, \zeta_j(r); t_j),$$

and the optimal kernel estimator of $h_n(z, t)$ is given by

$$\mathbf{h}_n(z, t_1, \dots, t_{n/2}) = \frac{2}{n} \sum_{j=1}^{n/2} \mathbf{h}_n^{(j)}(z, t_j). \tag{4.5}$$

5. The algorithm. In the following, we assume that the generalized eigenvalues $\zeta_j^{(r)}$, $j = 1 \dots n/2$ of the pencils $[U_1^{(r)}, U_0^{(r)}]$ —where the Hankel matrices $U_0^{(r)}, U_1^{(r)}$ are based on $d_k^{(r)}$ —have been computed for each $r = 1, \dots, R$ and clustered in such a way that for each $j = 1 \dots n/2$ the j -th cluster is the set $\{\zeta_j^{(r)}, r = 1, \dots, R\}$ whose elements are independent realizations of the r.v. ζ_j . The k-means method [19] can be used to solve the clustering problem.

In order to use the optimal kernel estimator given in eqs. 4.4 and 4.5, for each $j = 1 \dots n/2$ and $r = 1, \dots, R$, we need to solve the initial value problem obtained by eq. 3.1 in the limit for $\sigma \downarrow 0$ with initial condition given in eq. 4.1. By using the transformation $\tilde{h}(x, y, t) = \frac{h(x, y, t)}{p(x, y, t)}$, the equation $h_t = L[h]$ can be rewritten as $\tilde{h}_t = \frac{1}{p} \operatorname{div} [a \nabla(\tilde{h})]$. Therefore the initial value problems are

$$\begin{cases} \frac{\partial \Phi_{jr}}{\partial t} = \frac{1}{p} \left(\Delta \Phi_{jr} + a_x \frac{\partial \Phi_{jr}}{\partial x} + a_y \frac{\partial \Phi_{jr}}{\partial y} \right) \\ \Phi_{jr}(x, y, 0) = \frac{E_j(x, y)}{p(x, y)} \end{cases} \tag{5.1}$$

where Δ is the Laplacian operator, $E_j(x, y)$ is the empirical distribution of the generalized eigenvalue ζ_j , and

$$a(x, y; j, r) = \frac{p(x, y)\nu_{jr}(x, y)}{\left(\frac{\partial \nu_{jr}}{\partial x}(x, y)\right)^2 + \left(\frac{\partial \nu_{jr}}{\partial y}(x, y)\right)^2}, \quad \nu_{jr}(x, y) = \hat{\underline{s}}^H (I_n - X_{jr}^H (X_{jr} X_{jr}^H)^{-1} X_j) \hat{\underline{s}}$$

where

$$\hat{\underline{s}} = \frac{1}{R} \sum_{r=1}^R \underline{d}^{(r)}, \quad X_{jr}(h, k) = \bar{x}_{hk}^{(jr)}, \quad x_{hk}^{(jr)} = \begin{cases} \left(\zeta_h^{(r)}\right)^{k-1}, & h \neq j \\ z^{k-1}, & h = j \end{cases}.$$

The density $p(x, y) = h_n^{(j)}(x, y, \infty)$ is circularly symmetric (Cor. 1.7). Moreover, a closed form model of its modulus for each n can be found in [7].

The initial value problems were solved by a collocation method described in [25] in a $m_x \times m_y$ non-uniform grid. For each (j, r) the solution was approximated by the tensor product of Chebyshev polynomials in each spatial variable. Fast Fourier transform was used to compute the spatial derivatives. The resulting non-linear ODE system

$$\underline{w}'(t) = F(t, \underline{w}(t)), \quad \underline{w}(0) = \underline{e}, \quad \underline{w}(t), \underline{e} \in \mathbb{R}^{m_x m_y}$$

was then solved by MATLAB's built-in function *ode45.m*, where F is the discretized right-hand side of Eq. 5.1 and \underline{e} is the discretization of $E_j(x, y)$. The method is fast and stable provided that some spurious oscillations of the spatial derivatives close to the border of the integration region are filtered out. This task is accomplished by multiplying the derivatives by the function

$$F(x, y) = [\arctan\{(\tilde{x} + \gamma)/\phi\} - \arctan\{(\tilde{x} - \gamma)/\phi\}][\arctan\{(\tilde{y} + \gamma)/\phi\} - \arctan\{(\tilde{y} - \gamma)/\phi\}]$$

where

$$\tilde{x} = \pi(x - x_{min})/(x_{max} - x_{min}) - \pi/2, \quad \tilde{y} = \pi(y - y_{min})/(y_{max} - y_{min}) - \pi/2$$

and γ and ϕ are suitable positive parameters.

To compute the optimal bandwidth, $\mathbb{E}_{h^{(j)}}[G_j]$ is estimated by the sample mean of $G_j(\zeta, z)$; i.e., if the computed generalized eigenvalues are denoted by $\zeta_j^{(r)}$, then

$$\begin{aligned} \mathbb{E}_{h^{(j)}}[G_j] &\approx \frac{1}{R} \sum_{r=1}^R \int_{\mathbb{R}^2} G_j(\zeta_j^{(r)}, x, y) \hat{h}_n^{(j)}(\zeta_j^{(r)}, x, y) dx dy \\ &\approx \frac{t_j}{R} \sum_{r=1}^R \sum_{h=1}^{m_x} \sum_{k=1}^{m_y} [\Phi_{jr}(x_h, y_k, t_j)]^2 e^{\frac{\nu_{jr}(\zeta_j^{(r)}, x_h, y_k)}{t_j}} \delta_x(h) \delta_y(k) \end{aligned}$$

where $\delta_x(h) = x_h - x_{h-1}$, $\delta_x(1) = \delta_x(2)$, $\delta_y(k) = y_k - y_{k-1}$, and $\delta_y(1) = \delta_y(2)$. Moreover,

$$\|L[h_n^{(j)}]\|^2 \approx \left\| \frac{\partial \Phi_{jr}}{\partial t} \right\|^2 \approx \frac{1}{R} \sum_{r=1}^R \sum_{h=1}^{m_x} \sum_{k=1}^{m_y} \left[\frac{\partial \Phi_{jr}(x_h, y_k, \zeta_j^{(r)}, t_j)}{\partial t} \right]^2 \delta_x(h) \delta_y(k).$$

To compute $\mathbb{E}_{h^{(j)}}[G_j]$ and $\|L[h_n^{(j)}]\|^2$, we first need an estimate of t_j which can be provided e.g. by the variance \hat{t}_j of the generalized eigenvalues in each cluster.

6. Numerical results. In order to appreciate the advantages of the proposed kernel estimator, two numerical experiments were performed with different SNRs. $R = 10$ independent realizations $d_k^{(r)} = s_k + \epsilon_k^{(r)}$, $k = 1, \dots, n$, $r = 1, \dots, R$ of the r.v. \mathbf{d}_k were generated from the complex exponentials model with $p^* = 5$ components given by

$$\underline{\xi} = [e^{-0.1-i2\pi 0.3}, e^{-0.05-i2\pi 0.28}, e^{-0.0001+i2\pi 0.2}, e^{-0.0001+i2\pi 0.21}, e^{-0.3-i2\pi 0.35}]$$

$$\underline{c} = [6, 3, 1, 1, 20], \quad n = 74, \quad p = 37, \quad \sigma = 1, \quad \sigma = 3.$$

We notice that the frequencies of the 4th and 5th components are closer than the Nyquist frequency if $n < 1/(0.21 - 0.20) = 100$. Therefore, a super-resolution problem has to be solved. To speed up the computations, we limit the analysis of the condensed density to two regions containing respectively the first and second components and the third and fourth ones. The fifth component is isolated with a large amplitude; therefore, it is easy to identify even if its decay is fast. The considered regions are the rectangles defined by $\Omega = \Omega_x \times \Omega_y$ where $\Omega_x = [-0.8, 0.4]$, $\Omega_y = [-1.4, -0.4]$ for the first region and $\Omega_x = [-0.1, 0.6]$, $\Omega_y = [0.5, 1.3]$ for the second one. A mesh of size $m_x = 64$, $m_y = 64$ was considered. The values $\gamma = 1.6$, $\phi = 0.02$ of the derivatives filter were selected by trial and error.

In order to apply the proposed method, a pilot density estimate was first computed by the closed form approximation method given in [4]:

$$\hat{h}(z) \propto \sum_{r=1}^N \sum_{k=1}^p \hat{\Delta} \left\{ \Psi \left[\left(\frac{R_{kk}^{(r)}(z)^2}{\sigma^2 \beta} + 1 \right) \right] \right\}$$

where $R^{(r)}(z)$ is the R -factor of the QR factorization of the matrix $U_1^{(r)} - zU_0^{(r)}$ and $\beta = 5n\sigma^2$ [3, Prop. 6]. Then the generalized eigenvalues of the pencils $[U_1^{(r)}, U_0^{(r)}]$, $r = 1, \dots, R$ were pooled and the k-means method of clustering was applied with the number of clusters equal to the number of relative maxima of the pilot estimate.

For comparison, a Gaussian kernel estimate of density was also computed by Algorithm 1 in [14, App. E].

In Figure 1, the results obtained for the first region when $\sigma = 1$ are plotted: the empirical density (top left), the pilot density computed by the closed form approximation method (top right), the Gaussian kernel estimate (bottom left) and the result obtained by the proposed method (bottom right). In Figure 2, the same results are plotted for the second region. In Figures 3 and 4, the results obtained when $\sigma = 3$ are reported. The positions of the true complex exponentials are marked by a cross. It can be noticed that the proposed method is able to identify the two true complex exponentials even in the worst case, filtering out most of the spurious peaks of the empirical condensed density. Even if the location of the peaks is not perfect, it is the only method which is able to provide a reasonable solution to the super-resolution problem in the second region, and also for the smallest SNR considered.

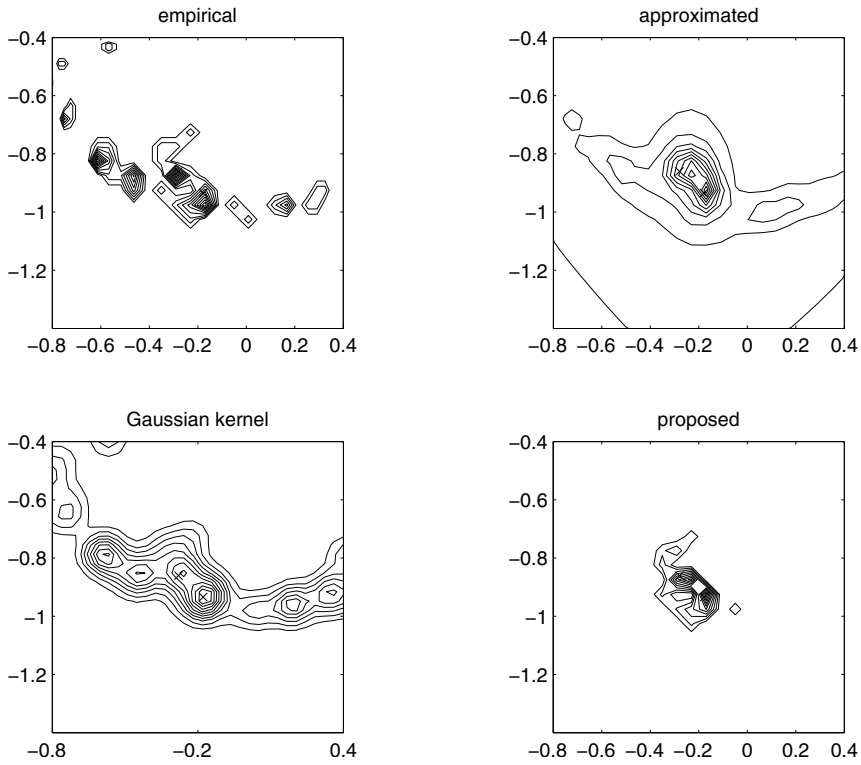


FIG. 1. First region, $\sigma = 1$. The empirical density (top left), the pilot density computed by the closed form approximation method (top right), the Gaussian kernel estimate (bottom left) and the result obtained by the proposed method (bottom right).

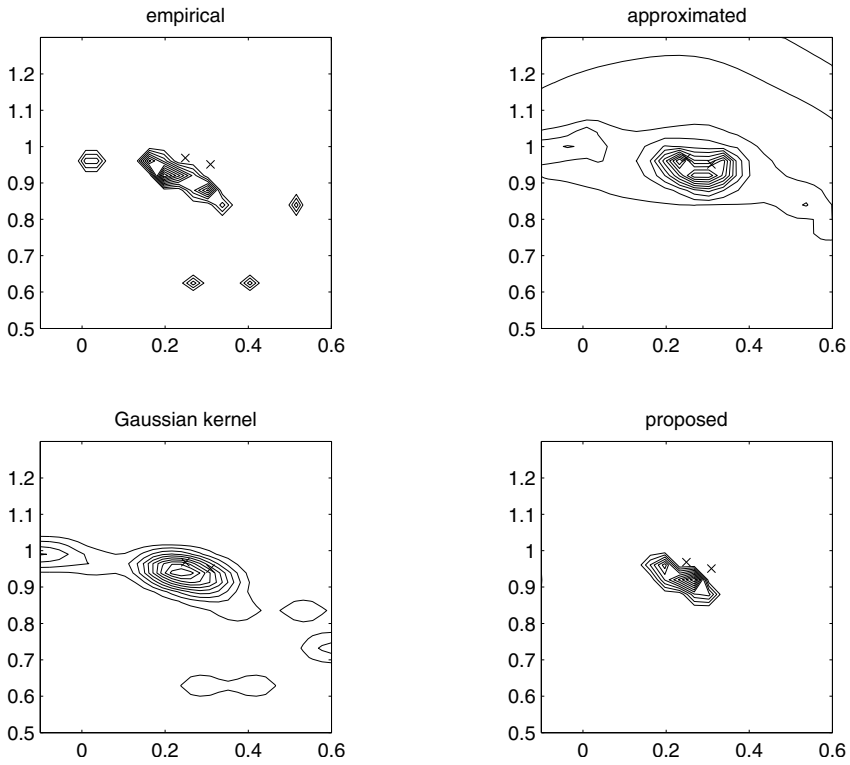


FIG. 2. Second region, $\sigma = 1$. The empirical density (top left), the pilot density computed by the closed form approximation method (top right), the Gaussian kernel estimate (bottom left) and the result obtained by the proposed method (bottom right).

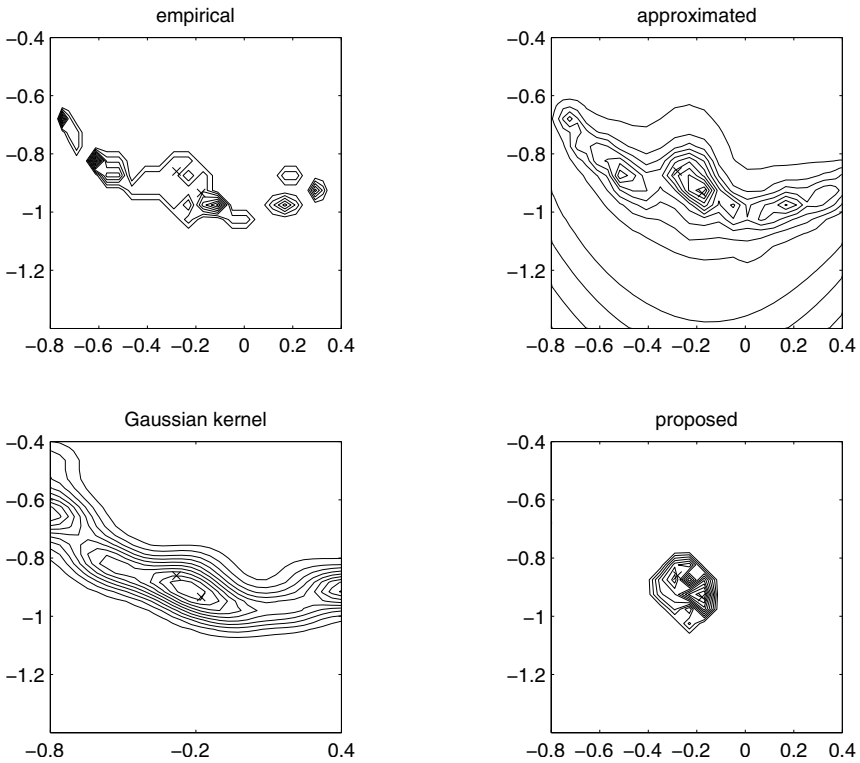


FIG. 3. First region, $\sigma = 3$. The empirical density (top left), the pilot density computed by the closed form approximation method (top right), the Gaussian kernel estimate (bottom left) and the result obtained by the proposed method (bottom right).

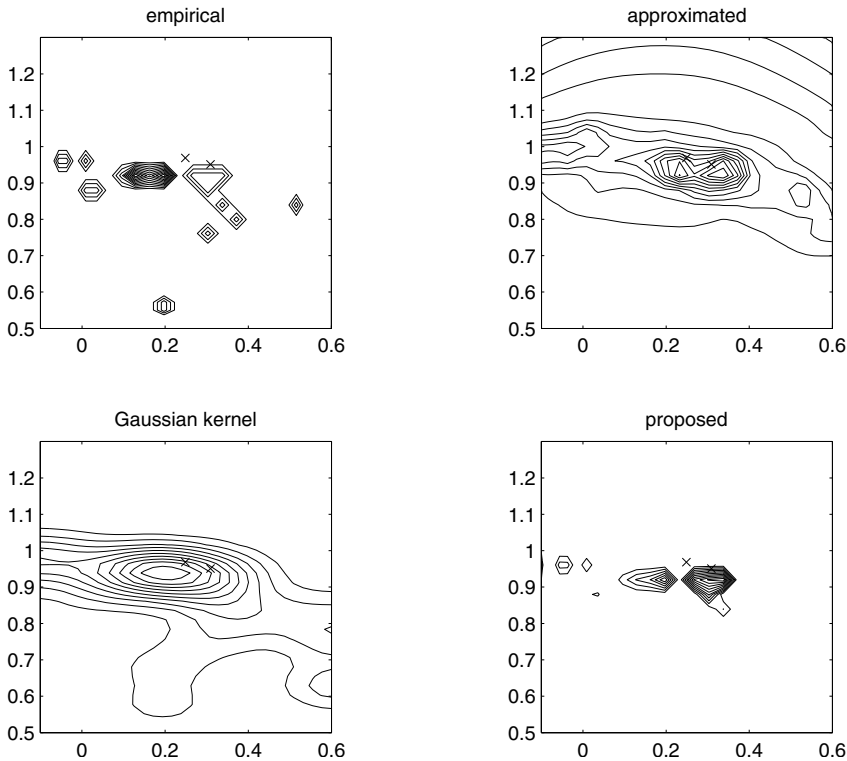


FIG. 4. Second region, $\sigma = 3$. The empirical density (top left), the pilot density computed by the closed form approximation method (top right), the Gaussian kernel estimate (bottom left) and the result obtained by the proposed method (bottom right).

Appendix.

Proof of Lemma 1.1. Let us find $\underline{\mu}_j \in \mathbb{C}^{n/2}$, $Q_j \in \mathbb{C}^{n/2 \times n/2}$ and $\nu_j \in \mathbb{C}$ such that

$$\sum_{k=0}^{n-1} \left| \sum_{h \neq j}^{1, n/2} \gamma_h \zeta_h^k + \gamma_j z^k - s_k \right|^2 = (\underline{\gamma} - \underline{\mu}_j)^H Q_j (\underline{\gamma} - \underline{\mu}_j) + \nu_j.$$

We have

$$\sum_{k=0}^{n-1} \left| \sum_{h \neq j}^{1, n/2} \gamma_h \zeta_h^k + \gamma_j z^k - s_k \right|^2 = \sum_{k=0}^{n-1} \left| \sum_h \gamma_h x_{hk}^{(j)} - s_k \right|^2.$$

Choosing $\underline{\mu}_j, Q_j$ and ν_j as defined in the text of the Lemma, we get the result.

Proof of Lemma 1.3. From Lemma 1.1 we have

$$h_2^{(1)}(z, \sigma) = \frac{1}{(\pi\sigma^2)^2} \int_{\mathbb{C}} \gamma e^{-\frac{1}{\sigma^2} [(\gamma - \mu_1)^H Q_1 (\gamma - \mu_1) + \nu_1]} d\gamma$$

where $X_1 = [1 \ \bar{z}]$, $Q_1 = 1 + |z|^2$, $\mu_1 = \frac{s_0 + \bar{z}s_1}{1 + |z|^2}$, $\nu_1 = |s_0|^2 + |s_1|^2 - \frac{|s_0 + \bar{z}s_1|^2}{1 + |z|^2}$; therefore

$$\begin{aligned} h_2^{(1)}(z, \sigma) &= \frac{1}{(\pi\sigma^2)^2} e^{-\frac{1}{\sigma^2}\nu_1} \int_{\mathbb{C}} \gamma e^{-\frac{(\gamma-\mu_1)^H(\gamma-\mu_1)(1+|z|^2)}{\sigma^2}} d\gamma \\ &= \frac{1}{(\pi\sigma^2)^2} e^{-\frac{1}{\sigma^2}\nu_1} \int_{\mathbb{R}^2} |\gamma|^2 e^{-\frac{(\gamma-\mu_1)^H(\gamma-\mu_1)(1+|z|^2)}{\sigma^2}} d\Re\gamma d\Im\gamma \\ &= \frac{1}{(\pi\sigma^2)^2} e^{-\frac{1}{\sigma^2}\nu_1} \frac{\pi\sigma^2}{1 + |z|^2} \left(|\mu_1|^2 + \frac{\sigma^2}{1 + |z|^2} \right). \end{aligned}$$

If $\xi = \frac{s_1}{s_0}$, then

$$\nu_1 = |s_0|^2 \frac{|z - \xi|^2}{1 + |z|^2}, \quad |\mu_1|^2 = |s_0|^2 \frac{|1 + \bar{z}\xi|^2}{(1 + |z|^2)^2}$$

and

$$h_2^{(1)}(z, \sigma) = e^{-\frac{|s_0|^2}{\sigma^2} \frac{|z - \xi|^2}{1 + |z|^2}} \left(\frac{|s_0|^2 |1 + \bar{z}\xi|^2}{\pi\sigma^2(1 + |z|^2)^3} + \frac{1}{\pi(1 + |z|^2)^2} \right).$$

But $s_0 = c_1$ and $s_1 = c_1\xi_1$; hence $\xi_1 = \xi$, and we get the thesis.

Proof of Lemma 1.4. By considering the vector $\tilde{\mu}_j$ obtained by stacking the real and imaginary parts of $\underline{\mu}_j$ and the real isomorph \tilde{Q}_j of the matrix Q_j , and remembering that the Jacobian with respect to the real and imaginary part of a complex variable is $J_R = |J_C|^2$, we get from Lemma 1.1 for $n > 2$

$$\begin{aligned} h_n^{(j)}(z, \sigma) &= \frac{1}{(\pi\sigma^2)^n} \int_{\mathbb{R}^{n-2}} \int_{\mathbb{R}^n} |J_C^*|^2 e^{-\frac{1}{\sigma^2}[(\tilde{\gamma} - \tilde{\mu}_j)^H \tilde{Q}_j (\tilde{\gamma} - \tilde{\mu}_j) + \nu_j]} d\tilde{\gamma} d\Re\underline{\zeta}^{(j)} d\Im\underline{\zeta}^{(j)} \\ &= \frac{1}{(\pi\sigma^2)^n} \int_{\mathbb{R}^{n-2}} \int_{\mathbb{R}^n} \prod_{j=1}^{1, n/2} |\gamma_j|^2 \prod_{r < h; r, h \neq j} |\zeta_r - \zeta_h|^8 \\ &\quad \cdot \prod_{r \neq j} |\zeta_r - z|^8 e^{-\frac{1}{\sigma^2}[(\tilde{\gamma} - \tilde{\mu}_j)^H \tilde{Q}_j (\tilde{\gamma} - \tilde{\mu}_j) + \nu_j]} d\tilde{\gamma} d\Re\underline{\zeta}^{(j)} d\Im\underline{\zeta}^{(j)}. \end{aligned}$$

By defining

$$f_j(\underline{\zeta}^{(j)}, z, \sigma) = \frac{1}{(\pi\sigma^2)^{n/2} |\tilde{Q}_j^{-1}|^{\frac{1}{2}}} \int_{\mathbb{R}^n} \left(\prod_{i=1}^{1, n/2} \tilde{\gamma}^T A_i \tilde{\gamma} \right) e^{-\frac{1}{\sigma^2}(\tilde{\gamma} - \tilde{\mu}_j)^H \tilde{Q}_j (\tilde{\gamma} - \tilde{\mu}_j)} d\tilde{\gamma}$$

and noticing that

$$|\gamma_i|^2 = \tilde{\gamma}^T A_i \tilde{\gamma},$$

we have

$$h_n^{(j)}(z, \sigma) = \int_{\mathbb{R}^{n-2}} f_j(\underline{\zeta}^{(j)}, z) \frac{\prod_{r < h; r, h \neq j} |\zeta_r - \zeta_h|^8 \prod_{r \neq j} |\zeta_r - z|^8}{(\pi\sigma^2)^{n/2} |\tilde{Q}_j(\underline{\zeta}^{(j)}, z)|^{\frac{1}{2}}} e^{-\frac{1}{\sigma^2}\nu_j} d\Re\underline{\zeta}^{(j)} d\Im\underline{\zeta}^{(j)},$$

and the thesis follows.

Proof of Lemma 1.5. By considering the normalized vectors $\underline{y}_j = \Sigma_j^{-\frac{1}{2}} \tilde{\gamma}$, we have $\underline{y}_j \sim N(\underline{m}_j, I)$ where

$$\underline{m}_j = \Sigma_j^{-\frac{1}{2}} \underline{\mu}_j = \frac{\sqrt{2}}{\sigma} \tilde{Q}_j^{1/2} \tilde{Q}_j^{-1} \tilde{q}_j = \frac{\sqrt{2}}{\sigma} \tilde{Q}_j^{-1/2} \tilde{q}_j = \frac{\sqrt{2}}{\sigma D_j^{1/2}} \hat{Q}_j^{1/2} \tilde{q}_j$$

where $D_j = \det(\tilde{Q}_j)$ and $\hat{Q}_j = \text{adj}(\tilde{Q}_j)$. Moreover,

$$\tilde{\gamma}^H A_i \tilde{\gamma} = (\Sigma_j^{-\frac{1}{2}} \tilde{\gamma})^T \Sigma_j^{\frac{1}{2}} A_i \Sigma_j^{\frac{1}{2}} \Sigma_j^{-\frac{1}{2}} \tilde{\gamma} = \underline{y}_j^T \Sigma_j^{\frac{1}{2}} A_i \Sigma_j^{\frac{1}{2}} \underline{y}_j = \underline{y}_j^T B_i \underline{y}_j$$

where

$$B_i = \Sigma_j^{\frac{1}{2}} A_i \Sigma_j^{\frac{1}{2}} = \frac{\sigma^2}{2} \tilde{Q}_j^{-1/2} A_i \tilde{Q}_j^{-1/2} = \frac{\sigma^2}{2 D_j} \hat{Q}_j^{1/2} A_i \hat{Q}_j^{1/2}.$$

From [1, Th. 1], denoting by \mathcal{Q}_i the quadratic form $\underline{y}_j^T B_i \underline{y}_j$, we have the recursion

$$f_j(\zeta^{(j)}, z, \sigma) = E \left[\prod_{i=1}^{1, n/2} \mathcal{Q}_i \right] = \sum_{i=0}^{n/2-1} 2^i \sum_{j_1=2}^{n/2} \cdots \sum_{j_i=2}^{n/2} \left(g_{j_1 \dots j_i} E \left[\frac{\mathcal{Q}_2 \cdots \mathcal{Q}_{n/2}}{\mathcal{Q}_{j_1} \cdots \mathcal{Q}_{j_i}} \right] \right)$$

where for $i = 0$, $g = \underline{m}^T B_1 \underline{m} + \text{tr}(B_1) = E[\mathcal{Q}_1]$ and for $i > 0$, $j_1 \neq j_2 \neq \dots \neq j_i$ and

$$g_{j_1 \dots j_i} = \underline{m}^T (B_1 B_{j_1} \dots B_{j_i} + B_{j_1} B_1 B_{j_2} \dots B_{j_i} + \dots + B_{j_1} B_{j_2} \dots B_{j_i} B_1) \underline{m} + \text{tr}(B_1 B_{j_1} \dots B_{j_i}).$$

But then

$$g_{j_1 \dots j_i} = \frac{\sigma^{2(i-1)}}{2^{i-1} D_j^{i+1}} \tilde{q}_j^T \mathcal{A} \tilde{q}_j + \frac{\sigma^{2i}}{2^i D_j^i} \text{tr}(A_1 \hat{Q}_j A_{j_1} \hat{Q}_j \dots A_{j_i} \hat{Q}_j)$$

$$g_{j_1 \dots j_i} = \frac{\sigma^{2i}}{2^i D_j^i} \left(\frac{2}{\sigma^2 D_j} \hat{Q}_j^{1/2} \tilde{q}_j^T \mathcal{A} \tilde{q}_j + \text{tr}(A_1 \hat{Q}_j A_{j_1} \hat{Q}_j \dots A_{j_i} \hat{Q}_j) \right)$$

where

$$\mathcal{A} = \hat{Q}_j A_1 \hat{Q}_j A_{j_1} \hat{Q}_j \dots \hat{Q}_j A_{j_i} \hat{Q}_j + \hat{Q}_j A_{j_1} \hat{Q}_j A_1 \hat{Q}_j A_{j_2} \hat{Q}_j \dots A_{j_i} \hat{Q}_j + \dots + \hat{Q}_j A_{j_1} \hat{Q}_j \dots A_{j_i} \hat{Q}_j A_1 \hat{Q}_j.$$

We have

$$E[\mathcal{Q}_1] = \frac{1}{D_j^2} \tilde{q}_j^T \hat{Q}_j A_1 \hat{Q}_j \tilde{q}_j + \frac{\sigma^2}{2 D_j} \text{tr}(A_1 \hat{Q}_j)$$

$$E[\mathcal{Q}_1 \mathcal{Q}_2] = E[\mathcal{Q}_1] E[\mathcal{Q}_2] + 4 \underline{m}_j^T B_1 B_2 \underline{m}_j + 2 \text{tr}(B_1 B_2)$$

$$= \frac{1}{D_j^4} \tilde{q}_j^T \hat{Q}_j A_1 \hat{Q}_j \tilde{q}_j \tilde{q}_j^T \hat{Q}_j A_2 \hat{Q}_j \tilde{q}_j$$

$$+ \frac{\sigma^2}{2 D_j^3} \left(\tilde{q}_j^T \hat{Q}_j A_1 \hat{Q}_j \tilde{q}_j \text{tr}(A_2 \hat{Q}_j) + \tilde{q}_j^T \hat{Q}_j A_2 \hat{Q}_j \tilde{q}_j \text{tr}(A_1 \hat{Q}_j) + \tilde{q}_j^T \hat{Q}_j A_i \hat{Q}_j A_i \hat{Q}_j \tilde{q}_j \right)$$

$$+ \frac{\sigma^4}{4 D_j^2} \left(\text{tr}(A_1 \hat{Q}_j) \text{tr}(A_2 \hat{Q}_j) + \text{tr}(A_1 \hat{Q}_j A_2 \hat{Q}_j) \right)$$

and, in general,

$$f_j(\zeta^{(j)}, z, \sigma) = P_{n/2}(\sigma^2, z) = \sum_{h=0}^{n/2} \beta_h(\zeta^{(j)}, z) \frac{\sigma^{2h}}{2^h D_j^{n-h}}$$

where $\beta_h(\underline{\zeta}^{(j)}, z)$ are positive polynomials. Moreover, $\beta_{n/2}(\underline{\zeta}^{(j)}, z)$ is the only coefficient that does not depend on \tilde{q}_j and therefore it does not depend on \underline{s} .

Proof of Corollary 1.6. By Lemma 1.4,

$$h_n^{(j)}(z, \sigma) = \int_{\mathbb{R}^{n-2}} f_j(\underline{\zeta}^{(j)}, z) \frac{\prod_{r<h;r,h\neq j} |\zeta_r - \zeta_h|^8 \prod_{r\neq j} |\zeta_r - z|^8 e^{-\frac{1}{\sigma^2} \nu_j} d\Re \underline{\zeta}^{(j)} d\Im \underline{\zeta}^{(j)}}{(\pi \sigma^2)^{n/2} |\tilde{Q}_j(\underline{\zeta}^{(j)}, z)|^{\frac{1}{2}}} =$$

$$\frac{1}{\sigma^n \pi^{n/2}} \sum_{k=0}^{n/2} \frac{\sigma^{2k}}{2^k} \int_{\mathbb{R}^{n-2}} \beta_k(\underline{\zeta}^{(j)}, z) \frac{\prod_{r<h;r,h\neq j} |\zeta_r - \zeta_h|^8 \prod_{r\neq j} |\zeta_r - z|^8 e^{-\frac{1}{\sigma^2} \nu_j} d\Re \underline{\zeta}^{(j)} d\Im \underline{\zeta}^{(j)}}{D_j^{n-k+\frac{1}{2}}}.$$

Noting that $|\zeta_r - \zeta_h|^8 = |\tilde{\zeta}_r - \tilde{\zeta}_h|^8$ and $|\zeta_r - z|^8 = |\tilde{\zeta}_r - \tilde{z}|^8$, we get the thesis.

Proof of Corollary 1.7. Noticing that in the definition of $g_j(\tilde{\underline{\zeta}}^{(j)}, \tilde{z}, \sigma)$, when $\sigma \downarrow \infty$, all terms vanish but the last one, and we get the first part of the thesis. By Lemma 1.5 $\beta_{n/2}(\tilde{\underline{\zeta}}^{(j)}, \tilde{z})$ is the only coefficient which does not depend on \underline{s} . Therefore, when $\underline{s} = \underline{0}$, Equation 1.1 reduces to $h_n^{(j)}(\tilde{z}, \infty)$. Finally, by symmetry, the condensed density does not depend on j ; therefore, all $h_n^{(j)}(\tilde{z}, \infty)$ must be equal. Moreover, in [7], it was proved that, when $\underline{s} = \underline{0}$, it depends only on $|\tilde{z}|^2$.

Proof of Theorem 2.1. We recall that if

$$I = \int_{\underline{y} \in \mathcal{D}} K(\underline{y}) e^{-\lambda \nu(\underline{y})} d\underline{y}, \quad \mathcal{D} \text{ open set } \subset \mathbb{R}^d, \quad \lambda \in \mathbb{R}^+$$

and $\nu(\underline{y})$ has a unique minimum in $\overline{\mathcal{D}}$ and this minimum occurs at a stationary point $\hat{\underline{y}}$ of $\nu(\underline{y})$, then the Laplace's approximation to I is given by

$$\tilde{I} = (2\pi)^{\frac{d}{2}} \lambda^{-\frac{d}{2}} |H(\hat{\underline{y}})|^{-\frac{1}{2}} K(\hat{\underline{y}}) e^{-\lambda \nu(\hat{\underline{y}})}$$

where $H(\underline{y})$ is the Hessian of ν .

We know that $\nu_j \geq 0$ and, by Lemma 1.2, $\underline{\xi}$ is the only vector such that $\nu_j(\underline{\xi}) = 0$. Therefore by continuity, $\nu_j(\tilde{\underline{\zeta}}^{(j)}, \xi_j)$ has a unique minimum $\hat{\underline{\zeta}}_j$ in a neighbor \mathcal{D}_j of $\{\xi_h, h \neq j\}$. Moreover, from Corollary 1.6, we notice that the dominant term in the sum defining $g_j(\tilde{\underline{\zeta}}^{(j)}, \tilde{z}, \sigma)$ when $\sigma \downarrow 0$ is the first one; therefore, in this case

$$g_j(\tilde{\underline{\zeta}}^{(j)}, \tilde{z}, \sigma) \approx \frac{1}{\sigma^n \pi^{n/2}} \beta_0(\tilde{\underline{\zeta}}^{(j)}, \tilde{z}) \frac{\prod_{r<h;r,h\neq j} |\tilde{\zeta}_r - \tilde{\zeta}_h|^8 \prod_{r\neq j} |\tilde{\zeta}_r - \tilde{z}|^8}{D_j^{n+\frac{1}{2}}(\tilde{\underline{\zeta}}^{(j)}, \tilde{z})} = \frac{1}{\sigma^n} K_j(\tilde{\underline{\zeta}}^{(j)}, \tilde{z}).$$

Then by using Laplace's approximation with $\lambda = \frac{1}{\sigma^2}$ and $d = n - 2$, we have, for $\tilde{z} \in \mathcal{N}_j$ where \mathcal{N}_j is a neighbor of ξ_j ,

$$\hat{h}_n^{(j)}(\tilde{z}, \sigma) = \frac{1}{\sigma^n} \int_{\mathcal{D}_j} K_j(\tilde{\underline{\zeta}}^{(j)}, \tilde{z}) e^{-\frac{1}{\sigma^2} \nu_j(\tilde{\underline{\zeta}}^{(j)}, \tilde{z})} d\tilde{\underline{\zeta}}^{(j)}$$

$$\approx \sigma^{-2} (2\pi)^{\frac{n}{2}-1} |H_j(\hat{\underline{\zeta}}_j, \tilde{z})|^{-\frac{1}{2}} K_j(\hat{\underline{\zeta}}_j, \tilde{z}) e^{-\frac{1}{\sigma^2} \nu_j(\hat{\underline{\zeta}}_j, \tilde{z})} \tag{6.1}$$

where H_j is the Hessian of ν_j . For simplicity, we will denote this approximation by the same symbol $\hat{h}_n^{(j)}(z, \sigma)$. Let us define

$$G_j(\hat{\underline{\zeta}}_j^{(j)}, \tilde{z}) = (2\pi)^{\frac{n}{2}-1} |H_j(\hat{\underline{\zeta}}_j, \tilde{z})|^{-\frac{1}{2}} K_j(\hat{\underline{\zeta}}_j, \tilde{z});$$

then

$$\hat{h}_n^{(j)}(\tilde{z}, \sigma) = \sigma^{-2} G_j(\hat{\zeta}^{(j)}, \tilde{z}) e^{-\frac{1}{\sigma^2} \nu_j(\hat{\zeta}^{(j)}, \tilde{z})}. \tag{6.2}$$

Proof of Lemma 3.1.

$$\begin{aligned} \frac{\partial \mathcal{D}(h, p)}{\partial t} &= \int_{\mathbb{R}^2} p(x, y) \frac{\partial}{\partial t} \left[\Psi \left(\frac{h(x, y, t)}{p(x, y)} \right) \right] dx dy = \\ &= \int_{\mathbb{R}^2} \Psi' \left(\frac{h(x, y, t)}{p(x, y)} \right) h_t(x, y) dx dy = \int_{\mathbb{R}^2} \Psi' \left(\frac{h(x, y, t)}{p(x, y)} \right) \operatorname{div} \left[a(x, y) \nabla \left(\frac{h(x, y, t)}{p(x, y)} \right) \right] dx dy. \end{aligned}$$

Integrating by parts we get

$$\begin{aligned} \frac{\partial \mathcal{D}(h, p)}{\partial t} &= \int_{\mathbb{R}^2} \operatorname{div} \left[\Psi' \left(\frac{h(x, y, t)}{p(x, y)} \right) a(x, y) \nabla \left(\frac{h(x, y, t)}{p(x, y)} \right) \right] dx dy \\ &\quad - \int_{\mathbb{R}^2} \nabla \left[\Psi' \left(\frac{h(x, y, t)}{p(x, y)} \right) \right] \cdot a(x, y) \nabla \left(\frac{h(x, y, t)}{p(x, y)} \right) dx dy, \end{aligned}$$

where \cdot denotes scalar product. By the divergence theorem, the first integral is zero because $\Psi'(1) = 0$. Moreover,

$$\nabla \left[\Psi' \left(\frac{h(x, y, t)}{p(x, y)} \right) \right] = \psi'' \left(\frac{h(x, y, t)}{p(x, y)} \right) \nabla \left(\frac{h(x, y, t)}{p(x, y)} \right).$$

Proof of Theorem 4.1.

$$\frac{\partial \mathbb{E}_{h^{(j)}} \{ \mathbf{h}^{(j)}(z, t) \}}{\partial t} = \int_{\mathbb{C}} \frac{\partial \Phi_j}{\partial t}(z, \zeta; t) h^{(j)}(\zeta) d\zeta = \int_{\mathbb{C}} L^* [\Phi_j(z, \zeta; t)] h^{(j)}(\zeta) d\zeta.$$

By definition of adjoint operator, taking into account that $\lim_{z \rightarrow \infty} \Phi_j(z, \zeta; t) = 0$, we get for $t \downarrow 0$

$$\frac{\partial \mathbb{E}_{h^{(j)}} \{ \mathbf{h}^{(j)}(z, t) \}}{\partial t} = \int_{\mathbb{C}} L^* [\Phi_j(z, \zeta; t)] h^{(j)}(\zeta) d\zeta = \int_{\mathbb{C}} \Phi_j(z, \zeta; t) L[h^{(j)}(\zeta)] d\zeta.$$

But $\lim_{t \rightarrow 0} \Phi_j(z, \zeta; t) = \delta(z - \zeta)$. Hence

$$\int_{\mathbb{C}} \Phi_j(z, \zeta; t) L[h^{(j)}(\zeta)] d\zeta \approx L[h^{(j)}(z)], \quad t \downarrow 0.$$

By considering the first-order Taylor series approximation of $\mathbb{E}_{h^{(j)}} \{ \mathbf{h}^{(j)}(z, t) \}$, we get

$$\mathbb{E}_{h^{(j)}} \{ \mathbf{h}^{(j)}(z, t) \} = \mathbb{E}_{h^{(j)}} \{ \mathbf{h}^{(j)}(z, 0) \} + t \frac{\partial \mathbb{E}_{h^{(j)}} \{ \mathbf{h}^{(j)}(z, t) \}}{\partial t} \Big|_{t=0} + O(t^2),$$

but

$$\mathbb{E}_{h^{(j)}} \{ \mathbf{h}^{(j)}(z, 0) \} = \int_{\mathbb{C}} \Phi_j(z, \zeta; 0) h^{(j)}(\zeta) d\zeta = \int_{\mathbb{C}} \delta(z - \zeta) h^{(j)}(\zeta) d\zeta = h^{(j)}(z).$$

Hence

$$\mathbb{E}_{h^{(j)}} \{ \mathbf{h}^{(j)}(z, t) \} = h^{(j)}(z) + t L[h^{(j)}(z)] + O(t^2)$$

and

$$\int_{\mathbb{C}} [\mathbb{E}_{h^{(j)}} \{ \mathbf{h}^{(j)}(z, t) \} - h^{(j)}(z)]^2 dz \approx \int_{\mathbb{C}} [t L[h^{(j)}(z)]]^2 dz = t^2 \|L[h^{(j)}]\|^2.$$

For approximating the integrated variance, let us consider first the second moment

$$\mathbb{E}_{h^{(j)}} \{ \Phi_j(z, \zeta; t)^2 \} = \int_{\mathbb{C}} \Phi_j(z, \zeta; t)^2 h^{(j)}(\zeta) d\zeta = \int_{\mathbb{C}} t^{-2} G_j(z, \zeta)^2 e^{-\frac{2}{t} \nu_j(z, \zeta)} h^{(j)}(\zeta) d\zeta.$$

But it was proved in [6, Th. 2] that

$$\lim_{t \rightarrow 0} \frac{2}{t} G_j(z, \zeta) e^{-\frac{2}{t} \nu_j(z, \zeta)} = \delta(z - \zeta).$$

Hence, for $t \downarrow 0$,

$$\mathbb{E}_{h^{(j)}} \{ \Phi_j(z, \zeta; t)^2 \} \approx \frac{1}{2} \int_{\mathbb{C}} t^{-1} G_j(z, \zeta) \delta(z - \zeta) h^{(j)}(\zeta) d\zeta = \frac{1}{2} t^{-1} G_j(z, z) h^{(j)}(z).$$

As $\zeta_j(r)$ are independent $\forall r$, it follows that

$$\begin{aligned} \text{Var}_h[\mathbf{h}^{(j)}(z, t)] &= \text{Var}_h \left[\frac{1}{R} \sum_{r=1}^R \Phi_j(x, y, \zeta_j(r); t) \right] = \frac{1}{R^2} \sum_{r=1}^R \text{Var}_h [\Phi_j(x, y, \zeta_j(r); t)] \\ &= \frac{1}{R} \mathbb{E}_{h^{(j)}} [\Phi_j(x, y, \zeta_j(r); t)^2] - \frac{1}{R} (\mathbb{E}_{h^{(j)}} [\Phi_j(x, y, \zeta_j(r); t)])^2 \\ &\approx \frac{1}{2Rt} G_j(z, z) h^{(j)}(z) - (h^{(j)}(z) + tL[h^{(j)}(z)])^2 \approx \frac{1}{2Rt} G_j(z, z) h^{(j)}(z) \end{aligned}$$

because for $t \downarrow 0$ the second term is negligible w.r. to the first one and

$$\int_{\mathbb{C}} \text{Var}_h[\mathbf{h}^{(j)}(z, t)] dz \approx \frac{1}{2Rt} \int_{\mathbb{C}} G_j(z, z) h^{(j)}(z) dz = \frac{1}{2Rt} \mathbb{E}_{h^{(j)}} [G_j].$$

REFERENCES

- [1] Y. Bao and A. Ullah, *Expectation of quadratic forms in normal and nonnormal variables with applications*, J. Statist. Plann. Inference **140** (2010), no. 5, 1193–1205, DOI 10.1016/j.jspi.2009.11.002. MR2581122 (2011c:62169)
- [2] P. Barone, *A diffusion equation for the density of the ratio of two jointly distributed Gaussian variables and the exponential analysis problem*, SIAM J. Sci. Comput. **34** (2012), no. 2, A1053–A1078, DOI 10.1137/110835323. MR2914315
- [3] P. Barone, *A black box method for solving the complex exponentials approximation problem* arXiv:0905.4602v2 (2012), Digital Signal Processing 23 (2013), 49–64.
- [4] P. Barone, *On the condensed density of the generalized eigenvalues of pencils of Gaussian random matrices and applications*, J. Multivariate Anal. **111** (2012), 160–173, DOI 10.1016/j.jmva.2012.05.009. MR2944412
- [5] P. Barone, *Estimation of a new stochastic transform for solving the complex exponentials approximation problem: Computational aspects and applications*, Digital Signal Processing 20,3 (2010), 724–735.
- [6] P. Barone, *A new transform for solving the noisy complex exponentials approximation problem*, J. Approx. Theory **155** (2008), no. 1, 1–27, DOI 10.1016/j.jat.2008.04.007. MR2474704 (2010a:62177)
- [7] P. Barone, *On the distribution of poles of Padé approximants to the Z-transform of complex Gaussian white noise*, J. Approx. Theory **132** (2005), no. 2, 224–240, DOI 10.1016/j.jat.2004.10.014. MR2118518 (2005h:41028)
- [8] P. Barone, *Random matrices in Magnetic Resonance signal processing*, The 8-th SIAM Conference on Applied Linear Algebra (2003).
- [9] P. Barone and R. March, *A novel class of Padé based method in spectral analysis*, J. Comput. Methods Sci. Eng., 1 (2001), 185–211.
- [10] P. Barone and R. March, *Some properties of the asymptotic location of poles of Padé approximants to noisy rational functions, relevant for modal analysis*, IEEE Trans. Signal Process. **46** (1998), no. 9, 2448–2457, DOI 10.1109/78.709533. MR1665253 (99m:41023)
- [11] P. Barone and A. Ramponi, *A new estimation method in modal analysis*, IEEE Trans. Signal Process. **48** (2000), no. 4, 1002–1014, DOI 10.1109/78.827534. MR1765795
- [12] P. Barone and G. Sebastiani, *A new method for Magnetic Resonance image reconstruction with short acquisition time and truncation artifact reduction*, IEEE Trans. Medical Imaging, 11 (1992), 250–259.

- [13] P. Barone and G. Sebastiani, *System for reproducing high-resolution magnetic resonance images with a short measurement time*, US Patent n.5.646.529, Jul. 8, 1997.
- [14] Z. I. Botev, J. F. Grotowski, and D. P. Kroese, *Kernel density estimation via diffusion*, *Ann. Statist.* **38** (2010), no. 5, 2916–2957, DOI 10.1214/10-AOS799. MR2722460 (2011k:62098)
- [15] D. Donoho, *Superresolution via sparsity constraints*, *SIAM J. Math. Anal.* **23** (1992), no. 5, 1309–1331, DOI 10.1137/0523074. MR1177792 (93f:42023)
- [16] W. Feller, *An introduction to probability theory and its applications. Vol. II*, John Wiley & Sons Inc., New York, 1966. MR0210154 (35 #1048)
- [17] G. H. Golub, P. Milanfar, and J. Varah, *A stable numerical method for inverting shape from moments*, *SIAM J. Sci. Comput.* **21** (1999/00), no. 4, 1222–1243 (electronic), DOI 10.1137/S1064827597328315. MR1740393 (2000m:65048)
- [18] P. Henrici, *Applied and computational complex analysis*, Wiley-Interscience [John Wiley & Sons], New York, 1974. Volume 1: Power series—integration—conformal mapping—location of zeros; Pure and Applied Mathematics. MR0372162 (51 #8378)
- [19] J. MacQueen, *Some methods for classification and analysis of multivariate observations*, *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability* (Berkeley, Calif., 1965/66), Univ. California Press, Berkeley, Calif., 1967, pp. Vol. I: Statistics, pp. 281–297. MR0214227 (35 #5078)
- [20] R. March and P. Barone, *Application of the Padé method to solving the noisy trigonometric moment problem: Some initial results*, *SIAM J. Appl. Math.* **58** (1998), no. 1, 324–343 (electronic), DOI 10.1137/S0036139995268742. MR1610064 (99h:44009)
- [21] J. M. Hammersley, *The zeros of a random polynomial*, *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. II*, University of California Press, Berkeley and Los Angeles, 1956, pp. 89–111. MR0084888 (18,941c)
- [22] R. March and P. Barone, *Reconstruction of a piecewise constant function from noisy Fourier coefficients by Padé method*, *SIAM J. Appl. Math.* **60** (2000), no. 4, 1137–1156 (electronic), DOI 10.1137/S0036139998333841. MR1760030 (2001b:41018)
- [23] M. R. Osborne and G. K. Smyth, *A modified Prony algorithm for exponential function fitting*, *SIAM J. Sci. Comput.* **16** (1995), no. 1, 119–138, DOI 10.1137/0916008. MR1311681 (95i:62059)
- [24] V. Viti, C. Petrucci and P. Barone, *Prony methods in NMR spectroscopy*, *International Journal of Imaging Systems and Technology*, **8** (1997), 565–571.
- [25] J. A. C. Weideman and S. C. Reddy, *A MATLAB differentiation matrix suite*, *ACM Trans. Math. Software* **26** (2000), no. 4, 465–519, DOI 10.1145/365723.365727. MR1939962 (2003g:65004)