Proceedings of the Project Review, Geo-Mathematical Imaging Group (Purdue University, West Lafayette IN),
Vol. 1 (2008) pp. 23-55.

NONLINEAR APPROXIMATION OF FUNCTIONS BY SUMS OF WAVE PACKETS*

FREDRIK ANDERSSON[†], MARCUS CARLSSON[‡], AND MAARTEN V. DE HOOP[§]

Abstract. We consider the problem of approximating functions by sums of wave packets. Our objective is to find sparse decompositions of image functions, over a finite range of scales. We also address the naturally connected task of approximating the wavefront set, computationally. We formulate the problem in terms of Hankel operators, Hankel matrices and their low-rank approximations, and develop an algebraic structure associated with the decomposition of functions into wave packets.

Key words. dyadic parabolic decomposition, nonlinear approximation, wave packets, Hankel operators

1. Introduction.

We consider the problem of approximating functions that arise in wave-equation imaging [12, 30] by sums of wave packets. Images are generated from (initial) data representing waves scattered off unknown boundaries, essentially, by backpropagation, and, by realistic acquisition design, are bandlimited in frequency. A natural candidate to represent images of these boundaries, and the data, is the frame of curvelets [21]. This frame can be viewed as a multiscale pyramid with certain directions and positions at each length scale. Representations of images with discontinuities along smooth (C^2) edges using curvelets are optimally sparse [11]; however, the functions that we consider here, do not contain sharp edges. Our objective is to find sparse decompositions of image, or data, functions over a finite range of scales, while honoring the dyadic parabolic decomposition of phase space underlying the construction of the frame of curvelets. We also address the naturally connected task of approximating the wavefront set, computationally, of these functions. Our approach is built on the work of Beylkin and Monzón [7, 8].

The common point of departure for decompositions of the type mentioned above consists of Fourier transforming the image function, and subjecting the result to a dyadic parabolic decomposition. One obtains functions in frequency, supported on the wedges (or boxes) that tile phase space in accordance with the dyadic parabolic decomposition (see Figure 1), by multiplication with appropriately chosen compactly supported window functions. The squares of these window functions form a partition of unity. Essentially, by Fourier series expansion of each such function defined on a wedge, one obtains a decomposition of the original image function with respect to a frame of curvelets.

We carry over the decomposition of functions with respect to a frame of curvelets to a nonlinear approximation by wave packets approach (for nonlinear wavelet approximations, see [13] and references therein). The objective is to approximate a function by very few packets. Non-linear approximations of functions, or signals, in one variable with wavelets have been in widespread use. The key concept here is to select the elements on which the function is projected, adaptively, that is, in a fashion depending on the function. In this spirit, pursuit algorithms [24] select approximation elements among given, redundant dictionaries of atoms. These dictionaries can contain multiple frames. The extent of such a dictionary can lead to a “combinatorial explosion” [24], however; pursuit algorithms reduce the computational complexity while searching for suboptimal approximations [24]. Matching pursuit [25], for example, reduces the computational complexity by a greedy strategy. In our approach, we go beyond the use of given dictionaries, and adapt the “location” and “shape” of wave packets extracted from a frame of curvelets, to the function to be decomposed, avoiding the “combinatorial explosion”.

Essentially, in our approach, we replace the Fourier series expansion of each function (in frequency) defined on a wedge in the dyadic parabolic decomposition in the following way. We begin with sampling the Fourier transform of the original image function at points that lie on uniform, oriented, grids tied to the wedges, such that the discrete inverse Fourier transform guarantees a prescribed accuracy. Given the values of a component function on a grid over a wedge, and a prescribed accuracy, we then construct a sum,

*This research was supported by the members, BP, ConocoPhillips, ExxonMobil, StatoilHydro and Total, of the Geo-Mathematical Imaging Group.

[†] Centre for Mathematical Sciences, Lund University, Box 118, SE-22100, Lund, Sweden, and Center for Computational and Applied Mathematics, Purdue University, 150 N. University Street, West Lafayette IN 47907, USA.

[‡] Centre for Mathematical Sciences, Lund University, Box 118, SE-22100, Lund, Sweden.

[§] Center for Computational and Applied Mathematics, Purdue University, 150 N. University Street, West Lafayette IN 47907, USA (mdehoop@purdue.edu).

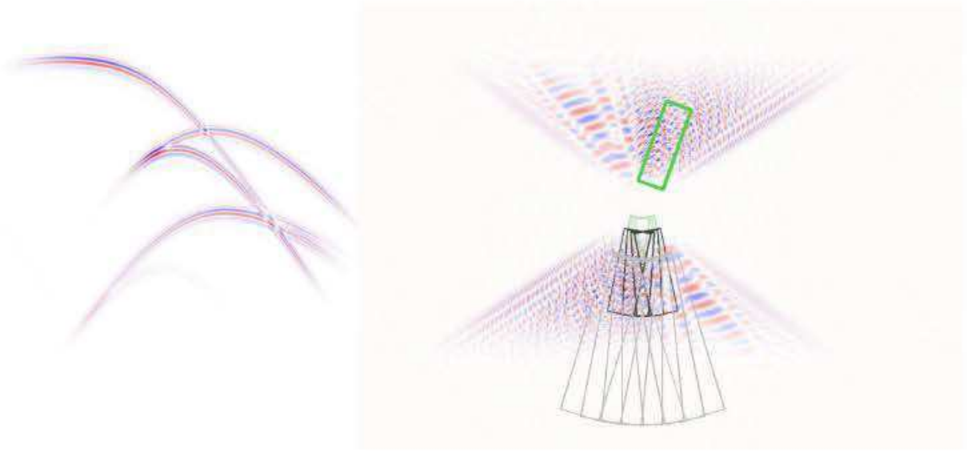

FIG. 1. *Left: synthetic (seismic) image used throughout this paper; right: the dyadic parabolic decomposition, and Fourier transformed image (real part). The particular box* $(B_{\nu,k})$ *of the dyadic parabolic decomposition that is used in the numerical experiments illustrated in Figures 2-6 is plotted in green.*

with minimal number of terms, of powers of complex nodes that approximates these function values. The construction consists of finding the complex nodes and, subsequently, obtaining the complex weights. The complex nodes introduce an algebraic structure underlying the decomposition into adaptive wave packets. By writing the complex nodes in the form of complex exponentials, an analogy with Fourier bases becomes apparent. It is tempting to view the complex nodes as (multi-dimensional) resonances and the wave packets as particles with momentum. (Indeed, in the one-dimensional counterpart, one can establish a connection with spectral estimation in signal processing [26].)

We reformulate the problem of finding the above mentioned complex nodes and complex weights in terms of Hankel operators. We form a block Hankel matrix from the function values to be approximated. The nodes and weights then define a low-rank block Hankel approximation of the original block Hankel matrix. To find these nodes and weights, we use properties of the Takagi factorization of the block Hankel matrix, while the error of the approximation is controlled by the con-eigenvalues arising in this factorization. The nodes appear as common roots of complex polynomials generated by the corresponding con-eigenvectors. As in the work of Beylkin and Monzón, this approach can be viewed as a finite-dimensional version of the Adamjan, Arov and Kreĭn theory [1, 2, 3].

We carry out numerical experiments with synthetic and field (seismic) data. We demonstrate that most of the complex weights appear to be smaller in magnitude than the prescribed accuracy, and the associated terms in the decomposition can be omitted. The sparsity of our decomposition does not rely on asymptotic estimates. Moreover, the process of adapting the wave packets reveals a bandlimited analogue of the wavefront set of an image function. In computations, for a given grid (that is, number of points), one has to select an appropriate con-eigenvalue of the block Hankel matrix to achieve a desired accuracy of the approximation.

Our applications pertain, but are not restricted, to seismic data and images. Decompositions of the type discussed above can, indeed, be exploited in the process of wave-equation imaging itself [18]. The generalized Radon transform admits a sparse matrix representation using curvelets [28, 17] yielding the notions of partial reconstruction – with data decomposed into wave packets – and illumination analysis – with image decomposed into wave packets. Moreover, it is possible to prove a concentration of curvelets, or wave packets, result for parametrices of hyperbolic evolution equations [4]; wave-equation imaging is composed of solutions to such equations. The decomposition developed in this paper has its roots in the theory of coherent wave packets and Fourier integral operators (Cordoba and Fefferman [14]) and contains elements of the dyadic parabolic decomposition of Fourier integral operators (Stein [29]) and the technique of parabolic cutoffs in the treatment of Fourier integral operators (the class $I^{p,l}$) with certain singular symbols (Greenleaf and Uhlmann [20]). The concept of parabolic cutoffs goes back to Boutet de Monvel [9]. Curvelets were first introduced in the parametrix construction for the wave equation with $C^{1,1}$ coefficients (Smith [27, 28]). Furthermore, the frame of curvelets and the associated curvelet transform can be related to the Fourier-Bros-Iagolnitzer (FBI) transform (Bros and Iagolnitzer [10]).

The outline of the paper is as follows. In Section 2 we summarize the dyadic parabolic decomposition of phase space, and introduce the decomposition into, and non-linear approximation by sums of wave packets and the relevant discretization and sampling (Section 2.2). In Section 3 we reformulate the non-linear approximation of sampled functions in phase space in terms of matrices. In Section 4 we develop the general principle and theory to carry out the program by expressing the approximation problem in terms of Hankel operators. The issue of developing a method for finding the complex nodes is postponed until Section 5. There exist pathological block Hankel matrices (and underlying functions) for which the construction of complex nodes proposed in Section 5 does not work; we characterize these in Section 5.3. In Section 6 we give a method of reducing the complexity of computations pertaining to finding the complex nodes. We conclude with some numerical experiments on real data in Section 7.

2. Dyadic parabolic decomposition of phase space. We summarize the method of second microlocalization, a tiling of phase space used throughout this paper. We begin with describing the dyadic parabolic decomposition and restrict ourselves to the two-dimensional case. We define boxes

$$B_k = \left[\xi_k' - \frac{L_k'}{2}, \xi_k' + \frac{L_k'}{2}\right] \times \left[-\frac{L_k''}{2}, \frac{L_k''}{2}\right],$$

where the centers ξ_k', and the side lengths L_k' and L_k'', satisfy the parabolic scaling condition,

$$\xi_k' \sim 2^k, \quad L_k' \sim 2^k, \quad L_k'' \sim 2^{k/2}, \quad \text{as } k \to \infty.$$

For $k = 0$, B_0 is a box centered at $\xi_0' = 0$, with $L_k' = L_k''$.

We will cover the plane $\mathbb{R}^2$ by a set of boxes indexed by two variables, k and ν, where ν (for each k) takes values in a set of $N_k = \lfloor 2^{k/2} \rfloor$ unit vectors distributed uniformly over the unit circle; we adhere to the convention that $e_1 = (1, 0)$ is one of these vectors. For each ν, k we let $\Theta_{\nu,k}$ denote the rotation operator on $\mathbb{R}^2$ such that $\Theta_{\nu,k}\nu = e_1$. The boxes $B_{\nu,k}$ are defined by

$$B_{\nu,k} = \Theta_{\nu,k}^{-1}(B_k).$$

The center of the box $B_{\nu,k}$ is $\xi_{\nu,k}^{\mathrm{cent}}$ with direction ν and $|\xi_{\nu,k}^{\mathrm{cent}}| \approx 2^k$.

We introduce smooth functions $\widehat{\chi}_k(\xi) \geq 0$ that vanish outside B_k, and set

$$\widehat{\chi}_{\nu,k}(\xi) = \widehat{\chi}_k(\Theta_{\nu,k}\xi).$$

Particular $\widehat{\chi}_k$ can be constructed such that

$$|\widehat{\chi}_0(\xi)|^2 + \sum_{k \geq 1} \sum_{\nu} |\widehat{\chi}_{\nu,k}(\xi)|^2 = 1, \tag{2.1}$$

yielding a partition of unity, while

$$|\langle \nu, \partial_\xi \rangle^j \, \partial_\xi^\alpha \widehat{\chi}_{\nu,k}(\xi)| \leq C_{j,\alpha} \, |\xi|^{-(j+|\alpha|/2)}, \tag{2.2}$$

in which the constants, $C_{j,\alpha}$, are independent of ν, k.

We define

$$\widehat{\varphi}_{\nu,k}(\xi) = \rho_k^{-1/2} \, \widehat{\chi}_{\nu,k}(\xi), \tag{2.3}$$

with $\rho_k = |B_k| = L_k' \, L_k'' \sim 2^{3k/2}$. The functions $\varphi_{\nu,k}$ satisfy estimates of the type

$$|\varphi_{\nu,k}(x)| \leq C_N 2^{3k/4} \, (\, 2^k |\langle \nu, x \rangle| + 2^{k/2} |x| \,)^{-N}. \tag{2.4}$$

The $\varphi_{\nu,k}$'s generate a tight frame of curvelets (in L^2) and lead to a transform pair that is described below; we will, here, depart from such a transform pair and develop a (non-linear approximation) decomposition where the translations of $\varphi_{\nu,k}$ are found optimally with regards to the function to be decomposed.

The frame representation mentioned above arises as follows. Following the partition of unity, one considers $\widehat{u}(\xi)\,|\widehat{\chi}_{\nu,k}(\xi)|^2$ and expands $\widehat{u}(\xi)\,\widehat{\chi}_{\nu,k}(\xi)$ in a Fourier series on its support, $B_{\nu,k}$, generating an orthonormal basis $\exp[-2\pi\mathrm{i}\langle x_j^{\nu,k},\xi\rangle]$. Here,

$$x_j^{\nu,k} = \Theta_{\nu,k}^{-1} D_k^{-1} X_j, \tag{2.5}$$

in which

$$X_j := (j_1, j_2),$$

while capturing the scaling of B_k in the dilation matrix

$$D_k = \begin{pmatrix} L_k' & 0 \\ 0 & L_k'' \end{pmatrix}.$$

The Fourier series expansion leads to the introduction of translates of the cutoff functions, $\varphi_{\nu,k}(x - x_j^{\nu,k})$. With the multi-index notation $\gamma = (x_j^{\nu,k}, \nu, k)$, we set $\varphi_\gamma(x) = \varphi_{\nu,k}(x - x_j^{\nu,k})$, or

$$\widehat{\varphi}_\gamma(\xi) = \rho_k^{-1/2}\,\widehat{\chi}_{\nu,k}(\xi)\,\exp[-2\pi\mathrm{i}\langle x_j^{\nu,k},\xi\rangle], \quad k \geq 1 \tag{2.6}$$

for scale k, orientation ν and a location $x_j^{\nu,k}$. The coefficients in the Fourier series expansion of $\widehat{u}(\xi)\,\widehat{\chi}_{\nu,k}(\xi)$ can be written in the form of an inner product

$$u_\gamma = \int u(x)\overline{\varphi_\gamma(x)}\,\mathrm{d}x, \tag{2.7}$$

and it follows that

$$\widehat{u}(\xi)\,|\widehat{\chi}_{\nu,k}(\xi)|^2 = \sum_{\gamma':\ k'=k,\ \nu'=\nu} u_{\gamma'}\widehat{\varphi}_{\gamma'}(\xi), \tag{2.8}$$

so that (cf. (2.1))

$$u(x) = \sum_\gamma u_\gamma \varphi_\gamma(x). \tag{2.9}$$

Equations (2.7) and (2.9) define a curvelet transform pair.

In (2.6 we can make the translation in frequency explicit and introduce

$$\widehat{\varphi}_\gamma^{\mathrm{cent}}(\xi) = \rho_k^{-1/2}\,\widehat{\chi}_{\nu,k}(\xi)\,\exp[-2\pi\mathrm{i}\langle x_j^{\nu,k},\xi - \xi_{\nu,k}^{\mathrm{cent}}\rangle], \tag{2.10}$$

simply affecting the coefficients according to

$$u_\gamma^{\mathrm{cent}} = u_\gamma \exp[-2\pi\mathrm{i}\langle x_j^{\nu,k},\xi_{\nu,k}^{\mathrm{cent}}\rangle], \quad \gamma = (x_j^{\nu,k},\nu,k)$$

cf. (2.8), where $\langle x_j^{\nu,k},\xi_{\nu,k}^{\mathrm{cent}}\rangle = (L_k')^{-1}|\xi_{\nu,k}^{\mathrm{cent}}|\,j_1$.

2.1. Decomposition into wave packets. We seek to develop a non-linear approximation approach to obtain sparse representations of functions u. We will differ from the above strategy by chosing $x_j^{\nu,k}$ to suit the function u to be approximated while letting the coordinates of these points to become complex. Thus we replace $x_j^{\nu,k}$ by $x_j^{\nu,k} + \mathrm{i}y_j^{\nu,k} \in \mathbb{C}^2$. We introduce functions

$$\widehat{\psi}_{j,\nu,k}(\xi) = \rho_{j,\nu,k}^{-1/2}\,|\widehat{\chi}_{\nu,k}(\xi)|^2\,\exp[-2\pi\mathrm{i}\langle x_j^{\nu,k} + \mathrm{i}y_j^{\nu,k},\xi - \xi_{\nu,k}^{\mathrm{cent}}\rangle]; \tag{2.11}$$

$\rho_{j,\nu,k}$ is determined by the normalization, $\|\psi_{j,\nu,k}\|_2 = 1$. Setting $\widehat{u_{\nu,k}}(\xi) = \widehat{u}(\xi)\,|\widehat{\chi}_{\nu,k}(\xi)|^2$, we then replace identity (2.8) by the approximation

$$\widehat{u_{\nu,k}}(\xi) \approx \sum_{j=1}^{N_{\nu,k}} c_j^{\nu,k} \widehat{\psi}_{j,\nu,k}(\xi) := \widehat{u_{\nu,k}^{N_{\nu,k}}}(\xi). \tag{2.12}$$

The aim is to find, for each pair ν, k, for given $\epsilon > 0$, an $N_{\nu,k}$ and points $x_j^{\nu,k} + \mathrm{i}y_j^{\nu,k}$, such that

$$\|u_{\nu,k}^{N_{\nu,k}} - u_{\nu,k}\|_2 < C_k\,\epsilon, \quad \text{with } C_k = \mathcal{O}(2^{-k/2}k^{-1-\delta}),\ \delta > 0. \tag{2.13}$$

Then $\sum_{\nu,k}\sum_{j=1}^{N_{\nu,k}} c_j^{\nu,k}\psi_{j,\nu,k}(x)$ yields a non-linear approximation of $u(x)$. As in the frame expansion, we could have expanded $\widehat{u_{\nu,k}}$ in a Fourier series resulting in an exact expansion with $y_j^{\nu,k} = 0$, $x_j^{\nu,k}$ lying on the lattice in (2.5), and $\rho_{j,\nu,k} = \rho_k$; here, we have introduced an alternative expansion with prescribed (and finite) accuracy ϵ. Instead of locking the positions of the packets to pre-assigned values, we have let them vary freely. In this way, we will be able to accurately capture the locations of singularities of u in the decomposition. Moreover, the exponential factor $\exp[2\pi\,\langle y_j^{\nu,k}, \xi - \xi_{\nu,k}^{\mathrm{cent}}\rangle]$, will tune the shape of the wavepackets. In this respect, we note that it has been established [11, 4] that the coefficients u_γ decay rapidly (super-algebraically) away from singularities; this result is essentially asymptotic. We develop, here, a representation or decomposition of a function, with singularities, which is sparse over a finite range of scales.

2.2. Discretization. We construct quadratures using rotated grids with respect to the partitioning functions $\widehat{\chi}_{\nu,k}$. We briefly review the necessary notation, and refer to [4] for details.

We normalize coordinates x such that u is supported on the disc, $\mathcal{D} = \{x \in \mathbb{R}^2\ :\ |x| < \frac{1}{2}\}$. We begin with sampling (2.8) in accordance with discretizing its inverse Fourier transform, that is,

$$u_{\nu,k}(x) \approx \frac{1}{\tau_k'\tau_k''}\sum_l \widehat{u}_{\nu,k}(\eta_l^{\nu,k})\exp[2\pi\mathrm{i}\langle x, \eta_l^{\nu,k}\rangle], \tag{2.14}$$

where the points $\eta_l^{\nu,k}$ are chosen from a regular lattice: The points are obtained from the finite set

$$\Xi^k = \left\{ (l_1, l_2) \in \mathbb{Z}^2 \ \middle|\ -\frac{M_k'}{2} \le l_1 \le \frac{M_k'}{2}, -\frac{M_k''}{2} \le l_2 \le \frac{M_k''}{2} \right\},$$

the elements of which are denoted by Ξ_l^k (in analogy with the notation X_j), and are given by

$$\eta_l^{\nu,k} = \xi_{\nu,k}^{\mathrm{cent}} + \Theta_{\nu,k}^{-1} S_k^{-1}\Xi_l^k.$$

Here, the parameters $M_k = (M_k', M_k'')$ are even natural numbers with $M_k' > L_k'$ and $M_k'' > L_k''$, while $\tau_k' = M_k'/L_k'$ and $\tau_k'' = M_k''/L_k''$ are the *oversampling* factors. The dilation matrix S_k is defined as

$$S_k = \begin{pmatrix} \tau_k' & 0 \\ 0 & \tau_k'' \end{pmatrix} = \begin{pmatrix} \dfrac{M_k'}{L_k'} & 0 \\ 0 & \dfrac{M_k''}{L_k''} \end{pmatrix}.$$

In practice, u is given in the form of an evenly sampled function on a covering of $\mathcal{D}$. We apply a USFFT [19, 6] to be able to evaluate the Fourier transform, $\widehat{u}$, at unequally spaced points $\eta_l^{\nu,k}$ jointly for all ν, k.

Upon discretization (2.14), the coefficients $c_j^{\nu,k}$ in (2.12) get replaced by coefficients $b_j^{\nu,k}$, say:

$$\widehat{u}_{\nu,k}(\eta_l^{\nu,k}) \approx \sum_{j=1}^{N_{\nu,k}} b_j^{\nu,k}\widehat{\psi}_{j,\nu,k}(\eta_l^{\nu,k}). \tag{2.15}$$

The exponential factor in $\widehat{\psi}_{j,\nu,k}(\eta_l^{\nu,k})$ (cf. (2.10)) can be written in the form

$$\begin{aligned}(2.16)\quad \exp[-2\pi\mathrm{i}\langle x_j^{\nu,k}+\mathrm{i}y_j^{\nu,k},\eta_{\mathbf{l}}^{\nu,k}-\xi_{\nu,k}^{\mathrm{cent}}\rangle] &= (z_j^{\nu,k})^{\mathbf{l}} = (z_{j;1}^{\nu,k})^{l_1}(z_{j;2}^{\nu,k})^{l_2},\\ z_{j;1,2}^{\nu,k} &= \exp[-2\pi\mathrm{i}\,(S_k^{-1}\Theta_{\nu,k}(x_j^{\nu,k}+\mathrm{i}y_j^{\nu,k}))_{1,2}],\end{aligned}$$

where we have adapted the notation to $\mathbf{l}=(l_1,l_2)$, replacing l, for the remainder of the paper. Setting $a_j^{\nu,k}=\rho_{j,\nu,k}^{-1/2}b_j^{\nu,k}$,

$$(2.17)\qquad \widehat{u}(\eta_{\mathbf{l}}^{\nu,k})\,|\widehat{\chi}_{\nu,k}(\eta_{\mathbf{l}}^{\nu,k})|^2 \approx \sum_{j=1}^{N_{\nu,k}} a_j^{\nu,k}\,|\widehat{\chi}_{\nu,k}(\eta_{\mathbf{l}}^{\nu,k})|^2(z_j^{\nu,k})^{\mathbf{l}}.$$

We will refer to the $z_j^{\nu,k}$ as "quadrature nodes". The objective is to find nodes, such that a desired accuracy is reached in (2.17) with minimal $N_{\nu,k}$. In this framework, not only the weights, $a_j^{\nu,k}$, and $N_{\nu,k}$ but also the nodes will depend on $u_{\nu,k}$.

To find the nodes, we consider the approximation

$$(2.18)\qquad \widehat{u}(\eta_{\mathbf{l}}^{\nu,k}) \approx \sum_{j\in J} a_j^{\nu,k}\,(z_j^{\nu,k})^{\mathbf{l}}.$$

which, upon multiplication by $|\widehat{\chi}_{\nu,k}(\eta_{\mathbf{l}}^{\nu,k})|^2$, attains the form of (2.17) with $N_{\nu,k}<|J|$; here, we note that all $\eta_{\mathbf{l}}^{\nu,k}$ are contained in $B_{\nu,k}$. This strategy is motivated by the following scenario. Suppose that u consists of a few point scatterers, then $\widehat{u}$ will naturally decompose into a sum of few exponentials, whereas $\widehat{u_{\nu,k}}$ would not. In our approximation strategy, we determine the nodes, using (2.18), first, and then compute the weights, using (2.17), while reducing the number of nodes to $N_{\nu,k}$.

In Sections 3-6 we focus on finding the nodes, and hence use approximation (2.18). We focus on a single ν,k box of the dyadic parabolic decomposition, and hence we will suppress the subscripts and superscripts relating to such a box. We set $M_k'=2m_1$ and $M_k''=2m_2$, and introduce the matrix $f(l_1+m_1,l_2+m_2)=\widehat{u}(\eta_{\mathbf{l}}^{\nu,k})\,I_{B_{\nu,k}}(\eta_{\mathbf{l}}^{\nu,k})=\widehat{u}(\eta_{\mathbf{l}}^{\nu,k})$, $-m_1\le l_1\le m_1$, $-m_2\le l_2\le m_2$, that is

$$(2.19)\qquad f=\begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0,2m_2)\\ f(1,0) & f(1,1) & \dots & f(1,2m_2)\\ \vdots & \vdots & \ddots & \vdots\\ f(2m_1,0) & f(2m_1,1) & \dots & f(2m_1,2m_2)\end{pmatrix}.$$

(We will also use the notation $F(\xi)=\widehat{u}(\xi)\,I_{B_{\nu,k}}(\xi)$.)

3. Sparse decompositions of two-dimensional data. We introduce some notation. We use bold letters for multi-indices, $\mathbf{m}=(m_1,m_2)$ with $m_1,m_2\in\mathbb{N}$. Let $\mathbb{M}_{\mathbf{m}}$ be the space of $(m_1+1)\times(m_2+1)$ matrices with complex entries $(a_{\mathbf{i}})_{\mathbf{0}\le\mathbf{i}\le\mathbf{m}}$, where $\mathbf{0}$ stands for $(0,0)$ and $\mathbf{i}\le\mathbf{m}$ means that $i_1\le m_1$ and $i_2\le m_2$. Given a point $z=(z_1,z_2)\in\mathbb{C}^2$, we denote by $\boxed{z}$ the element in $\mathbb{M}_{2\mathbf{m}}$ given by

$$(3.1)\qquad \boxed{z}=\begin{pmatrix} 1 & z_2 & \dots & z_2^{2m_2}\\ z_1 & z_1z_2 & \dots & z_1z_2^{2m_2}\\ \vdots & \vdots & \ddots & \vdots\\ z_1^{2m_1} & z_1^{2m_1}z_2 & \dots & z_1^{2m_1}z_2^{2m_2}\end{pmatrix}.$$

The matrix formulation analogue of (2.18) can be stated as follows: Consider $f\in\mathbb{M}_{2\mathbf{m}}$ as in (2.19); we seek an approximation of the form

$$(3.2)\qquad f\approx\sum_{j=1}^{N} a_j\,\boxed{z_j},$$

with $a_j \in \mathbb{C}$. In this process, the key objective is, for given $\mathbf{m}$, to find points $\{z_j\}_{j=1}^N$ such that

$$\|f - \mathit{Proj}_{\mathcal{Z}} f\| \leq \epsilon,$$

with $0 < \epsilon \ll 1$ and $N = N(\epsilon) \ll (2m_1+1)(2m_2+1)$, while $\mathit{Proj}_{\mathcal{Z}}$ stands for the orthogonal projection onto the subspace $\mathcal{Z} = \operatorname{span}\{\boxed{z_j}\}$ of $\mathbb{M}_{2\mathbf{m}}$. We develop such approximations in Sections 4-6. Once the set $\{z_j\}_{j=1}^N$ has been obtained, we determine the a_j's in principle by solving the normal equations corresponding with the linear system (3.2) subjected to the windowing as in (2.17).

In the case of functions in one variable, a method to obtain approximations of the form (3.2) has been developed by Beylkin and Monzón [8], which we will now describe. After sampling, a function defines a vector f in $\mathbb{C}^{2m+1}$, if $2m+1$ is the number of sample points. We will consider both column and row vectors as elements of $\mathbb{C}^{2m+1}$ and matrices as operators on this space in the usual way. One begins by forming the Hankel matrix,

$$H_f = \begin{pmatrix} f(0) & f(1) & \dots & f(m) \\ f(1) & f(2) & \dots & f(m+1) \\ \vdots & \vdots & \ddots & \vdots \\ f(m) & f(m+1) & \dots & f(2m) \end{pmatrix}.$$

By the Takagi factorization [22], H_f has $m+1$ con-eigenvectors $u_1, u_2, \dots, u_{m+1} \in \mathbb{C}^{m+1}$ associated with con-eigenvalues $\sigma_1 \geq \dots \geq \sigma_{m+1} \geq 0$. That is,

$$H_f u_n = \sigma_n \overline{u_n}, \quad n = 1, \dots, m+1, \tag{3.3}$$

where the bar denotes complex conjugation of the entries in u_n. Alternatively, this can be expressed as

$$H_f = \overline{U} \Sigma U^*, \tag{3.4}$$

where $U = (u_1, \dots, u_{m+1})$ is unitary and

$$\Sigma = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_{m+1} \end{pmatrix}.$$

If $D_{u_n} \in \mathbb{M}_{m,2m}$ denotes the "difference" operator given by

$$D_{u_n} = \begin{pmatrix} u_n(0) & u_n(1) & \dots & u_n(m) & 0 & \dots & 0 \\ 0 & u_n(0) & u_n(1) & \dots & u_n(m) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \\ 0 & \dots & 0 & u_n(0) & u_n(1) & \dots & u_n(m) \end{pmatrix},$$

then equation (3.3) can be rewritten in the form

$$D_{u_n} f = \sigma_n \overline{u_n}, \quad n = 1, \dots, m+1. \tag{3.5}$$

We now fix n. We let $\dot{u}_n \in \mathbb{C}^{2m+1}$ be the first row in D_{u_n}, and let S denote the "shift" operator on $\mathbb{C}^{2m+1}$, defined as

$$S(a(0), a(1), \dots, a(2m)) = (0, a(0), \dots, a(2m-1)) \quad \text{for } a \in \mathbb{C}^{2m+1}.$$

It is easy to see that the vectors $\dot{u}_n, S\dot{u}_n, \dots, S^m \dot{u}_n$ are linearly independent; the range of D_{u_n} is $\operatorname{span}\{S^k \dot{u}_n\}_{k=0}^m$. Furthermore, we form the polynomial $P_{u_n}(z) = \sum_{p=0}^m u_n(p) z^p$. Generically, P_{u_n} will have m distinct roots, which we assume is the case, and label them z_j, $j = 1, \dots, m$. We form vectors, $\boxed{z_j} \in \mathbb{C}^{2m+1}$, according to $\boxed{z_j} = (1, z_j, \dots, z_j^{2m})$. It follows that $\langle \boxed{z_j}, \overline{S^k \dot{u}_n} \rangle = z_j^k P_{u_n}(z_j) = 0$ [1], and hence the complex conjugates of $S^k \dot{u}_n$ are orthogonal to $\operatorname{span}\{\boxed{z_j}\}_{j=1}^m$; arguing by dimension, we obtain

$$\mathbb{C}^{2m+1} = \operatorname{span}\left\{ \boxed{z_j} \right\}_{j=1}^m \oplus \operatorname{span}\left\{\overline{S^k \dot{u}_n}\right\}_{k=0}^m, \tag{3.6}$$

which implies that

$$\operatorname{Ker} D_{u_n} = \operatorname{span}\{\boxed{z_j}\}_{j=1}^m.$$

We write $\mathcal{Z}_n = \operatorname{span}\{\boxed{z_j}\}_{j=1}^m$.

The equation (cf. (3.5))

$$D_{u_n} y = \sigma_n \overline{u_n} \tag{3.7}$$

clearly has many solutions. Let y_p denote the solution of this equation with minimal norm, that is, y_p is the (unique) solution that is orthogonal to $\operatorname{Ker} D_{u_n}$. By extending D_{u_n} to a circular matrix $\dot{D}_{u_n}$ and subjecting the equation, $\dot{D}_{u_n} y = \sigma_n \overline{\dot{u}_n}$, corresponding with (3.7), to the finite Fourier transform, it is straightforward to show [2] that there exists a solution y' to (3.7) such that $\|y'\| = \|\sigma_n \overline{u_n}\| = \sigma_n$. Hence,

$$\|y_p\| \leq \sigma_n. \tag{3.8}$$

Clearly, $f - y' \in \mathcal{Z}_n$. Using (3.6), we deduce that y_p is the orthogonal projection of f onto $\operatorname{span}\{\overline{S^k \dot{u}_n}\}_{k=0}^m$. With (3.8) it follows that the distance from f to the subspace $\mathcal{Z}_n$ is less than or equal to σ_n. This is essentially a reformulation of [8, Theorem 2]. One obtains the approximation

$$f \approx \sum_{j=1}^m a_j \boxed{z_j} \tag{3.9}$$

with error estimate σ_n. Given the number of sample points, $2m+1$, the best approximation is obtained by letting $n = m+1$, because σ_{m+1} is the smallest con-eigenvalue of H_f. Moreover, if the vector f originates from sampling a piecewise continuous function, F, on a closed interval, $[0,1]$ say ($f(k) = \frac{1}{\sqrt{2m+1}} F(\frac{k}{2m})$ for $k = 0, 1, \dots, 2m$), then $\lim_{m \to \infty} \sigma_{m+1} = 0$. This demonstrates convergence of the approximation method.

However, m is not a small number compared to $2m+1$, the original dimension of f, and hence the approximation with $n = m+1$ cannot be sparse in general. In fact, up to this point, the development is essentially a slight improvement of Prony's method, which prescribes that one should first append two numbers to the vector f such that H_f becomes singular (then $\sigma_{m+2} = 0$), which then by the above arguments implies that $f \in \mathcal{Z}_{m+2}$; thus, an exact representation of f in the form (3.9) with m replaced by $m+1$ can be obtained. However, Prony's method is known to be unstable and therefore of limited use in practice [8, Section 2.3].

The key observation made by Beylkin and Monzón lies in discovering that the number of significant terms in the approximation (3.9) is approximately equal to the index n (if we assume that f comes from sampling some continuous function). They also observe a rapid decay of the con-eigenvalues of H_f for given m. Thus, we can choose an $n \ll m+1$ such that σ_n is small and obtain an approximation

$$f \approx \sum_{j=1}^n a_j \boxed{z_j}, \tag{3.10}$$

with (cf. (3.2)) $N = n$ and error of the same size as σ_n:

$$\|f - \sum_{j=1}^n a_j \boxed{z_j}\| \approx \sigma_n.$$

[1] We topologize $\mathbb{C}^{2m+1}$ with the standard scalar product.

[2] See Lemma 4.5 for a proof in two dimensions.

4. Quadrature nodes: General principle. In this section, we analyze and begin to extend approximations of the type (3.10) to two dimensions, thus considering functions in two variables. We let $f \in \mathbb{M}_{2\mathbf{m}}$ be as in (2.19). We develop the necessary preparation for an explicit construction of non-linear approximations, given in the next section, while making use of a single con-eigenvalue con-eigenvector pair as in the previous section. We arrive at a description of approximations in terms of quadrature nodes that are constrained to lie in a specific variety. The explicit construction of quadrature nodes, however, requires a refinement of this method and is addressed in Section 5.

We will throughout treat $\mathbb{M}_{2\mathbf{m}}$ as a Hilbert space with the usual scalar product $\langle u, v\rangle = \sum_{\mathbf{i}} u(\mathbf{i})\overline{v(\mathbf{i})}$ and we will assume that $\|f\| = 1$. We first form the operator $H_f : \mathbb{M}_{\mathbf{m}} \to \mathbb{M}_{\mathbf{m}}$ given by

$$(H_f u)(\mathbf{i}) = \sum_{\mathbf{0}\leq\mathbf{j}\leq\mathbf{m}} f(\mathbf{i}+\mathbf{j})u(\mathbf{j}).$$

Let $e_{\mathbf{i}}$ be the standard basis in $\mathbb{M}_{\mathbf{m}}$, that is, $e_{\mathbf{i}}(\mathbf{j}) = 1$ if $\mathbf{i} = \mathbf{j}$ and zero otherwise. If we order this basis in some arbitrary fashion, then H_f is represented by a symmetric matrix, as the following calculation demonstrates:

$$\langle H_f e_{\mathbf{i}}, e_{\mathbf{j}}\rangle = f(\mathbf{i}+\mathbf{j}) = \langle H_f e_{\mathbf{j}}, e_{\mathbf{i}}\rangle.$$

But then, by the Takagi factorization, H_f has con-eigenvalues $\sigma_1 \geq \ldots \geq \sigma_{(m_1+1)(m_2+1)} \geq 0$ and corresponding con-eigenvectors $u_n \in \mathbb{M}_{\mathbf{m}}$. That is,

$$H_f u_n = \sigma_n \overline{u_n}, \quad n = 1, \ldots, (m_1+1)(m_2+1). \tag{4.1}$$

As in [8], the error in our approximations will depend on σ_n for chosen n. Thus the decay of σ_n with n is important in this context.

PROPOSITION 4.1. *Let $\mathbf{m} = (m, m)$ with $m \geq 9$ is an odd number. Let $F \in C^1([0,1]^2)$ be given and let $f \in \mathbb{M}_{2\mathbf{m}}$ be sampled on an equally spaced grid according to $f(\mathbf{i}) = \frac{1}{2m}F(\frac{i_1}{2m}, \frac{i_2}{2m})$. Then*

$$\sigma_n \leq \frac{\|F'\|_\infty}{5m}$$

for all $n \geq \frac{(m+1)^2}{2}$.

Proof. For $0 \leq j_1 \leq (\frac{m-1}{2})$, $0 \leq j_2 \leq m$ we define $b_{\mathbf{j}} = \frac{1}{\sqrt{2}}(e_{(2j_1,j_2)} - e_{(2j_1+1,j_2)}) \in \mathbb{M}_{\mathbf{m}}$. These vectors form an orthonormal set that span a subspace $\mathcal{M}$ of dimension $\frac{m+1}{2}(m+1)$. We note that

$$\Big|H_f(b_{\mathbf{j}})(\mathbf{k})\Big| = \frac{1}{\sqrt{2}2m}\Big|F\Big(\frac{k_1+2j_1}{2m}, \frac{k_2+j_2}{2m}\Big) - F\Big(\frac{k_1+2j_1+1}{2m}, \frac{k_2+j_2}{2m}\Big)\Big| \leq \frac{\|F'\|_\infty}{\sqrt{2}2^2m^2}$$

so that

$$\|H_f(b_{\mathbf{j}})\|^2 \leq \sum_{\mathbf{0}\leq\mathbf{k}\leq\mathbf{m}} \frac{\|F'\|_\infty^2}{2^5m^4} = \frac{(m+1)^2\|F'\|_\infty^2}{2^5m^4}.$$

As $\{b_{\mathbf{j}}\}$ is an orthonormal set and $2^{2.5}m/(m+1) > 5$, the operator norm of $H_f|_{\mathcal{M}}$ is less than $\frac{\|F'\|_\infty}{5m}$, where $H_f|_{\mathcal{M}}$ denotes the operator H_f restricted to $\mathcal{M}$. We recall that $\{\sigma_n\}_{n=1}^{(m+1)^2}$ is just the set of singular values for H_f, and therefore (given $k \in \mathbb{N}$) we have

$$\sigma_{(m+1)^2-k} = \inf\Big\{\|H_f|_{\mathcal{N}}\| : \mathcal{N} \subset \mathbb{M}_{\mathbf{m}} \text{ is a linear subspace with } \dim\mathcal{N} = k\Big\}, \tag{4.2}$$

where $\|\cdot\|$ denotes the operator norm. But then we have that

$$\sigma_{\frac{(m+1)^2}{2}} = \sigma_{(m+1)^2-\frac{m+1}{2}(m+1)} \leq \frac{\|F'\|_\infty}{5m},$$

from the statement in the proposition follows. □

We note that the requirement that $m_1 = m_2$ in the proposition can be removed.

We now return to the problem of finding quadrature nodes for a fixed $f \in \mathbb{M}_{\mathbf{2m}}$. For $\mathbf{0} \leq \mathbf{j} \leq \mathbf{m}$ let $S_{\mathbf{j}}$ denote the cyclic shift operator with index $\mathbf{j}$ in $\mathbb{M}_{\mathbf{2m}}$, that is, for $a \in \mathbb{M}_{\mathbf{2m}}$ we have

$$S_{\mathbf{j}} a(\mathbf{i}) = a[\mathbf{i} - \mathbf{j}]_{\mathbf{2m+1}},$$

where $\mathbf{1} = (1,1)$ and

$$[\mathbf{i} - \mathbf{j}]_{\mathbf{2m+1}} = \big((i_1 - j_1) \mod (2m_1 + 1),\, (i_2 - j_2) \mod (2m_2 + 1)\big).$$

Loosely speaking, the operator $S_{\mathbf{j}}$ takes a matrix a, moves it j_1 times downwards, j_2 times to the right and fills up the empty spaces by the terms that have been "pushed out".

Let $\dot{u}_n$ be the element in $\mathbb{M}_{\mathbf{2m}}$ formed by adding zeros to the right and below the matrix $u_n \in \mathbb{M}_{\mathbf{m}}$, and let $\mathcal{R}_n \subset \mathbb{M}_{\mathbf{2m}}$ be the subspace given by

$$\mathcal{R}_n = \operatorname{span}\{S_{\mathbf{i}} \overline{\dot{u}_n} : \mathbf{0} \leq \mathbf{i} \leq \mathbf{m}\}. \tag{4.3}$$

Let P_{u_n} be the polynomial given by

$$P_{u_n}(z) = \sum_{\mathbf{0} \leq \mathbf{i} \leq \mathbf{m}} u_n(\mathbf{i}) z^{\mathbf{i}},$$

where $z = (z_1, z_2) \in \mathbb{C}^2$ and $z^{\mathbf{i}} = z_1^{i_1} z_2^{i_2}$. Let $V(P_u)$ denote the algebraic variety $\{z \in \mathbb{C}^2 : P_u(z) = 0\}$, and set

$$\mathcal{Z}_n = \operatorname{span}\{\boxed{z} : z \in V(P_{u_n})\},$$

(cf. (3.1) for the notation). Finally, we form the "partial difference" operator $D_{u_n} : \mathbb{M}_{\mathbf{2m}} \to \mathbb{M}_{\mathbf{m}}$ according to

$$(D_{u_n} y)(\mathbf{i}) = \sum_{\mathbf{0} \leq \mathbf{j} \leq \mathbf{m}} y(\mathbf{i} + \mathbf{j}) u_n(\mathbf{j}). \tag{4.4}$$

We note the identity

$$(D_{u_n} y)(\mathbf{i}) = \langle y, S_{\mathbf{i}} \overline{\dot{u}_n} \rangle, \tag{4.5}$$

and that

$$D_{u_n} f = H_f u_n = \sigma_n \overline{u_n}. \tag{4.6}$$

LEMMA 4.2. *The operator* $D_{u_n}^* : \mathbb{M}_{\mathbf{m}} \to \mathbb{M}_{\mathbf{2m}}$ *satisfies*

$$P_{D_{u_n}^* v} = P_{\overline{u_n}} P_v,$$

where $D_{u_n}^*$ *denotes the adjoint of* D_{u_n}. *In particular,* $\mathcal{R}_n = \operatorname{Ran} D_{u_n}^*$.

Proof. By direct evaluation, we find that

$$P_{D_{u_n}^* v}(z) = \langle D_{u_n}^* v, \overline{\overline{\boxed{z}}} \rangle = \langle v, D_{u_n} \overline{\overline{\boxed{z}}} \rangle = \sum_{\mathbf{0} \leq \mathbf{i} \leq \mathbf{m}} v(\mathbf{i}) \overline{\overline{z}^{\mathbf{i}} P_{u_n}(\overline{z})} = P_v(z) P_{\overline{u_n}}(z),$$

using (4.4). The second statement is an immediate consequence. □

We recall that a polynomial on $\mathbb{C}^2$ is called reduced if all of its factors are distinct. Analogously to equation (3.6) we have:

PROPOSITION 4.3. *Assume that P_{u_n} is reduced and that $u_n(\mathbf{m}) \neq 0$.*[3] *Then*

$$\mathbb{M}_{\mathbf{2m}} = \mathcal{R}_n \oplus \mathcal{Z}_n,$$

and $\mathcal{Z}_n = \operatorname{Ker} D_{u_n}$.

Proof. Given any $z \in \mathbb{C}^2$ we have

$$(D_{u_n} \boxed{z})(\mathbf{i}) = \langle \boxed{z}, S_{\mathbf{i}} \overline{\dot{u}_n} \rangle = z^{\mathbf{i}} P_{u_n}(z);$$

hence, $\mathcal{R}_n \perp \mathcal{Z}_n$, and $\mathcal{Z}_n \subset \operatorname{Ker} D_{u_n}$. Let $y \in \mathbb{M}_{\mathbf{m}} \ominus (\mathcal{R}_n \oplus \mathcal{Z}_n)$. As $\langle \boxed{z}, y \rangle = P_{\overline{y}}(z)$ we infer that $V(P_{\overline{y}}) \supset V(P_{u_n})$. By Hilbert's Nullstellensatz and Proposition 9, Ch. 4.2 in [15], we deduce that

$$P_{\overline{y}} = P_{u_n} P_v$$

for some $v \in \mathbb{M}_{\mathbf{m}}$. But then

$$\overline{y} = \sum_{\mathbf{0} \leq \mathbf{i} \leq \mathbf{m}} v(\mathbf{i}) S_{\mathbf{i}} \dot{u}_n,$$

whence $y \in \mathcal{R}_n$, which is a contradiction. This concludes the proof that $\mathbb{M}_{\mathbf{2m}} = \mathcal{R}_n \oplus \mathcal{Z}_n$. To argue that $\mathcal{Z}_n = \operatorname{Ker} D_{u_n}$, we recall that $\operatorname{Ker} D_{u_n} = \mathbb{M}_{\mathbf{2m}} \ominus \operatorname{Ran} D^*_{u_n} = \mathbb{M}_{\mathbf{2m}} \ominus \mathcal{R}_n$ where the second equality follows by Lemma 4.2. □

As the vectors $S_{\mathbf{i}} \dot{u}_n$, $\mathbf{0} \leq \mathbf{i} \leq \mathbf{m}$, are linearly independent, we obtain the following corollary

COROLLARY 4.4. *Under the assumptions of Proposition 4.3, we have*

$$\dim \mathcal{R}_n = (m_1 + 1)(m_2 + 1)$$

and

$$\dim \mathcal{Z}_n = 3m_1 m_2 + m_1 + m_2.$$

Proposition 4.3 and equation (4.6) show that if $y_p \in \mathbb{M}_{\mathbf{2m}}$ denotes the smallest (in $\mathbb{M}_{\mathbf{2m}}$-norm) solution to

$$D_{u_n} y = \sigma_n \overline{u_n},$$

then y_p is the orthogonal projection of f onto $\mathcal{R}_n$. Moreover, $\|y_p\|$ is the distance from f to $\mathcal{Z}_n$. We now show that

$$\|y_p\| \leq \sigma_n,$$

in analogy with (3.8). Let $\gamma = (e^{\frac{2\pi \mathrm{i}}{2m_1+1}}, e^{\frac{2\pi \mathrm{i}}{2m_2+1}}) \in \mathbb{C}^2$ and set $C = 1/\sqrt{(2m_1+1)(2m_2+1)}$. We define the discrete Fourier transform in two variables, $\mathcal{F} : \mathbb{M}_{\mathbf{2m}} \to \mathbb{M}_{\mathbf{2m}}$, as

$$(\mathcal{F} y)(\mathbf{i}) = C P_y(\gamma^{-\mathbf{i}}).$$

We recall that this is a unitary operator with inverse transform given by $(\mathcal{F}^{-1} y)(\mathbf{i}) = C P_y(\gamma^{\mathbf{i}})$.

Using that $(D_{u_n} y)(\mathbf{i}) = \langle y, S_{\mathbf{i}} \overline{\dot{u}_n} \rangle$, we extend D_{u_n} to an operator $\dot{D}_{u_n} : \mathbb{M}_{\mathbf{2m}} \to \mathbb{M}_{\mathbf{2m}}$ by letting $\mathbf{0} \leq \mathbf{i} \leq 2\mathbf{m}$:

$$(\dot{D}_{u_n} y)(\mathbf{i}) = \langle y, S_{\mathbf{i}} \overline{\dot{u}_n} \rangle.$$

[3] This corresponds to the assumption that P_{u_n} has m distinct roots in the one-dimensional case.

Given $a, b \in \mathbb{M}_{\mathbf{2m}}$, we define the matrix $a \odot b$ by componentwise multiplication. Let $d_n = \mathcal{F}\dot{u}_n$. By explicit calculation, if follows that

$$\mathcal{F}^{-1}\dot{D}_{u_n}\mathcal{F}a = C^{-1}d_n \odot a, \quad \text{for all } a \in \mathbb{M}_{\mathbf{2m}}. \tag{4.7}$$

LEMMA 4.5. *The equation*

$$D_{u_n}y = \sigma_n\overline{u_n} \tag{4.8}$$

has a solution, y', with $\|y'\| = \sigma_n$.

Proof. We consider the extended equation $\sigma_n\overline{\dot{u}_n} = \dot{D}_{u_n}y$. Any solution y is clearly also a solution to the original equation (4.8). Using (4.7), we transform the equation to

$$\sigma_n\overline{d_n} = \sigma_n\overline{\mathcal{F}\dot{u}_n} = \mathcal{F}^{-1}(\sigma_n\overline{\dot{u}_n}) = \mathcal{F}^{-1}\dot{D}_{u_n}\mathcal{F}\mathcal{F}^{-1}y = C^{-1}d_n \odot \mathcal{F}^{-1}y.$$

This equation is easily seen to have the solution $\mathcal{F}^{-1}y'(\mathbf{i}) = \sigma_n C\frac{\overline{d(\mathbf{i})}}{d(\mathbf{i})}$, and

$$\|y'\| = \|\mathcal{F}^{-1}y'\| = \sigma_n C\sqrt{\sum_{\mathbf{0}\leq\mathbf{i}\leq\mathbf{2m}} 1} = \sigma_n C\,\frac{1}{C} = \sigma_n,$$

as desired. □

COROLLARY 4.6. *Generically, we have $\|f - Proj_{\mathcal{Z}_n}f\| \leq \sigma_n$, where $Proj_{\mathcal{Z}_n}$ denotes the orthogonal projection onto $\mathcal{Z}_n$.*

Proof. If we assume that f and n are such that Proposition 4.3 applies to u_n, then the statement follows immediately from Lemma 4.5. What we mean with "holds generically" will be specified in the remark below. The proof that Proposition 4.3 indeed applies generically is given in Appendix B. □

REMARK 4.7. *Let $\mathbb{S}(\mathbb{M}_{\mathbf{2m}})$ denote the unit sphere in the linear space $\mathbb{M}_{\mathbf{2m}}$; we recall that f is assumed to be a fixed element in $\mathbb{S}(\mathbb{M}_{\mathbf{2m}})$. By the expression "In the generic case .." in the above Corollary, we mean that the subset in $\mathbb{S}(\mathbb{M}_{\mathbf{2m}})$ of elements for which the result or statement is not applicable is of zero area-measure. If f arises as a measurement from some experiment, this means that we have assumed that the probability measure is absolutely continuous with respect to area-measure.*

In summary, let us consider a given f, and let us choose an n such that σ_n is "small". Corollary 4.4 states that, in the generic case, $\dim \mathcal{Z}_n = 3m_1m_2 + m_1 + m_2$. Moreover, Corollary 4.6 states that

$$\|f - Proj_{\mathcal{Z}_n}f\| \leq \sigma_n.$$

Hence, if we choose points $\{z_j\}_{j=1}^N \in \mathbb{C}^2$ in $V(P_{u_n})$ such that $\{\boxed{z_j}\}$ is a basis for $\mathcal{Z}_n$, we have found an approximation of f of the type (3.2) with $N = 3m_1m_2 + m_1 + m_2$ and accuracy $\leq \sigma_n$. Moreover, Proposition 4.1 implies that if f originates from sampling a C^1 function, we can make σ_n arbitrarily small upon choosing $\mathbf{m}$, and n, sufficiently large.

So far, we have not addressed the issue of how to find points $\{z_j\}_{j=1}^N \in \mathbb{C}^2$ in $V(P_{u_n})$ that yield a basis for $\mathcal{Z}_n$. In the next section, we further develop the theory to arrive at a construction of an appropriate basis. Furthermore, $N = 3m_1m_2 + m_1 + m_2$ is not a small number. We provide insight in how to reduce N in Section 6.

5. Quadrature nodes: An explicit construction.

A natural approach to address the ambiguity in choice of the points $\{z_j\}$ is to constrain them to the intersection of two $\mathcal{Z}_n$'s. For any pair $n, n' \leq \dim \mathbb{M}_{\mathbf{m}}$, $n' \neq n$, we expect that

$$\mathcal{Z}_n \cap \mathcal{Z}_{n'} = \operatorname{span}\{\boxed{z} : z \in V(P_{u_n}, P_{u_{n'}})\},$$

where $V(P_{u_n}, P_{u_{n'}}) = \{z \in \mathbb{C}^2 : P_{u_n}(z) = 0,\ P_{u_{n'}}(z) = 0\}$ is, in the generic case, a finite set. Based on Corollary 4.6, one might conjecture that $\|f - Proj_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}}f\|$ will be small if $\sigma_n, \sigma_{n'}$ are sufficiently small, achieved by choosing n, n' appropriately. In the following two subsections, we show that *(i)* In the generic case, P_{u_n} and $P_{u_{n'}}$ have exactly $2m_1m_2$ common solutions $\{z_j\}_{j=1}^{2m_1m_2}$, and

$$\mathcal{Z}_n \cap \mathcal{Z}_{n'} = \operatorname{span}\{\boxed{z_j}\}_{j=1}^{2m_1m_2}.$$

(ii) It is possible that $f \perp \mathcal{Z}_n \cap \mathcal{Z}_{n'}$, that is, $\mathit{Proj}_{\mathcal{Z}_n\cap\mathcal{Z}_{n'}} f = 0$, also with the largest $n = (m_1+1)(m_2+1) - 1$ and $n' = (m_1+1)(m_2+1)$. In this case, the error of the approximation is $\|f\|$.

5.1. Proof of *(i)*. We will use the notation $\begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}$ for the operator from $\mathbb{M}_{2\mathbf{m}}$ into $\mathbb{M}_{\mathbf{m}} \oplus \mathbb{M}_{\mathbf{m}} = \mathbb{M}_{\mathbf{m}}^2$ given by

$$\begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}(g) = \begin{pmatrix} D_{u_n}g \\ D_{u_{n'}}g \end{pmatrix}, \quad g \in \mathbb{M}_{2\mathbf{m}}.$$

We will use both 2×1-matrices and 1×2-matrices to denote the elements in $\mathbb{M}_{\mathbf{m}}^2$. We note that $\begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}^* : \mathbb{M}_{\mathbf{m}}^2 \to \mathbb{M}_{2\mathbf{m}}$, is given by

$$\begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}^* \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = D_{u_n}^* h_1 + D_{u_{n'}}^* h_2, \quad h_1, h_2 \in \mathbb{M}_{\mathbf{m}}, \tag{5.1}$$

and that f satisfies

$$\begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}(f) = \begin{pmatrix} \sigma_n\overline{u_n} \\ \sigma_{n'}\overline{u_{n'}} \end{pmatrix},$$

while $\left\|\begin{pmatrix} \sigma_n\overline{u_n} \\ \sigma_{n'}\overline{u_{n'}} \end{pmatrix}\right\| = \sqrt{\sigma_n^2 + \sigma_{n'}^2}$ (in the usual product topology). However, in contrast to the results in the previous section, we will show that the smallest solution, y_p, to the equation

$$\begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}(y) = \begin{pmatrix} \sigma_n\overline{u_n} \\ \sigma_{n'}\overline{u_{n'}} \end{pmatrix},$$

is not always of the same magnitude as $\sqrt{\sigma_n^2 + \sigma_{n'}^2}$.

We begin with determining for which $(\sigma_1, \sigma_2) \in \mathbb{C}^2$ the equation

$$\begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix}(y) = \begin{pmatrix} \sigma_1\overline{v_1} \\ \sigma_2\overline{v_2} \end{pmatrix}, \tag{5.2}$$

for given $v_1, v_2 \in \mathbb{M}_{\mathbf{m}}$, has a solution, thereby answering the question of which pairs $v_1, v_2 \in \mathbb{M}_{\mathbf{m}}$ are con-eigenvectors to some operator H_f. Given any polynomial $p(z) = \sum_{\mathbf{i}} a_{\mathbf{i}} z^{\mathbf{i}}$ on $\mathbb{C}^2$, we let $\deg p$ denote the smallest multi-index such that $\mathbf{i} \leq \deg p$ for all $\mathbf{i}$ with $a_{\mathbf{i}} \neq 0$.

PROPOSITION 5.1. *Let $v_1, v_2 \in \mathbb{M}_{\mathbf{m}}$ such that $v_1(\mathbf{m}) \neq 0 \neq v_2(\mathbf{m})$*[4] *be given and let q be a greatest common divisor of $P_{\overline{v_1}}, P_{\overline{v_2}}$. Let $d_i \in \mathbb{M}_{\mathbf{m}}$ be such that $P_{\overline{v_i}} = P_{d_i}q$ for $i = 1, 2$. Form the matrixes $A_1, A_2 \in \mathbb{M}_{\deg q}$ by*

$$A_1(\mathbf{i}) = \langle \overline{v_1}, S_{\mathbf{i}} d_2 \rangle \quad \textit{and} \quad A_2(\mathbf{i}) = \langle \overline{v_2}, S_{\mathbf{i}} d_1 \rangle.$$

Then equation (5.2) has a solution if and only if $\sigma_1 A_1 - \sigma_2 A_2 = 0$. In particular, if A_1 and A_2 are linearly independent, then equation (5.2) is solvable only for $\sigma_1 = \sigma_2 = 0$.

Proof. By Lemma 4.2 an element $(w_1, w_2) \in \mathbb{M}_{\mathbf{m}}^2$ lies in Ker $\begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix}^*$ if and only if

$$P_{\overline{v_1}} P_{w_1} + P_{\overline{v_2}} P_{w_2} = q(P_{d_1}P_{w_1} + P_{d_2}P_{w_2}) = 0.$$

[4] This assumption simplifies the statement of the proposition but can easily be removed.

But two polynomials are equal if and only if they have the same factors, which implies that (w_1, w_2) lies in Ker $\begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix}^*$ if and only if

$$\left\{ \begin{array}{lcl} P_{w_1} & = & P_{d_2} P_r \\ P_{w_2} & = & -P_{d_1} P_r \end{array} \right. \tag{5.3}$$

for some $r \in \mathbb{M}_{\deg q}$. Thus

$$(\sigma_1 \overline{v_1}, \sigma_2 \overline{v_2}) \in \operatorname{Ran} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix}$$

if and only if

$$\langle (\sigma_1 \overline{v_1}, \sigma_2 \overline{v_2}), (w_1, w_2) \rangle = 0 \tag{5.4}$$

for all $(w_1, w_2) \in \mathbb{M}_{\mathbf{m}}^2$ satisfying (5.3). The left-hand side of (5.4) equals

$$\sigma_1 \sum_{\mathbf{0} \leq \mathbf{i} \leq \deg q} \overline{r(\mathbf{i})} A_1(\mathbf{i}) - \sigma_2 \sum_{\mathbf{0} \leq \mathbf{i} \leq \deg q} \overline{r(\mathbf{i})} A_2(\mathbf{i}) = \Big\langle \sigma_1 A_1 - \sigma_2 A_2, r \Big\rangle,$$

which is zero for all $r \in \mathbb{M}_{\deg q}$ if and only if $\sigma_1 A_1 - \sigma_2 A_2 = 0$. □

It is not difficult to construct examples of $v_1, v_2 \in \mathbb{M}_{\mathbf{m}}$ in the proof above, for which $\langle v_1, v_2 \rangle = 0$ and $q \neq 1$ such that one of the following holds: (1) $A_1 = A_2 = 0$, or (2) $A_1,\ A_2$ are non-zero but linearly dependent, or (3) $A_1,\ A_2$ are linearly independent. When $q = 1$ the situation much simpler, as the following corollary shows.

COROLLARY 5.2. *In the generic case we have $q = 1$. Then*

$$\operatorname{Ran} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix} = \left\{ \begin{pmatrix} -\overline{v_2} \\ \overline{v_1} \end{pmatrix} \right\}^{\perp}$$

In particular, if we assume that $\langle v_1, v_2 \rangle = 0$, then the equation

$$\begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix} (y) = \begin{pmatrix} \sigma_1 \overline{v_1} \\ \sigma_2 \overline{v_2} \end{pmatrix}$$

has a solution for every pair $(\sigma_1, \sigma_2) \in \mathbb{C}^2$.

Proof. The first equality follows by the calculations in the proof of Proposition 5.1, and the second statement is an immediate consequence. The proof that $q = 1$ holds generically is given in Appendix B. □

Proposition 5.1 and Corollary 5.2 describe which pairs $(v_1, v_2) \in \mathbb{M}_{\mathbf{m}}^2$ can appear as con-eigenvectors to some H_f. To prove claim *(i)* made in the beginning of this section, we first need the following lemma. We will temporarily use the notation $\mathcal{Z}_i = \operatorname{span}\{\boxed{z} : z \in V(P_{v_i})\}$ and $\mathcal{R}_i = \operatorname{span}\{S_{\mathbf{j}} \overline{v_i} : \mathbf{0} \leq \mathbf{j} \leq \mathbf{m}\}$ for $i = 1, 2$ although this conflicts with the previous notation.

LEMMA 5.3. *Let $v_1, v_2 \in \mathbb{M}_{\mathbf{m}}$ be such that*

(i) $v_1(\mathbf{m}) \neq 0$ *and* $v_2(\mathbf{m}) \neq 0$*,*

(ii) the polynomials $c_1(z_1) = \sum_{k_1=0}^{m_1} v_1(k_1, m_2) z_1^{k_1}$ *and* $c_2(z_1) = \sum_{k_1=0}^{m_1} v_2(k_1, m_2) z_1^{k_1}$ *have no common roots,*

(iii) the polynomials $d_1(z_2) = \sum_{k_2=0}^{m_2} v_1(m_1, k_2) z_2^{k_2}$ *and* $d_2(z_2) = \sum_{k_2=0}^{m_2} v_2(m_1, k_2) z_2^{k_2}$ *have no common roots.*

Then, for all $w \in \mathbb{M}_{2\mathbf{m}}$ we have that $P_w \in \langle P_{v_1}, P_{v_2} \rangle$ if and only if $\overline{w} \in \mathcal{R}_1 + \mathcal{R}_2$.

Proof. The "if" part is immediate by Lemma 4.2. Conversely, let $w \in \mathbb{M}_{2\mathbf{m}}$ be such that $P_w \in \langle P_{v_1}, P_{v_2} \rangle$. Then

$$P_w = P_{v_1} g_1 - P_{v_2} g_2 \tag{5.5}$$

for some $g_1, g_2 \in \mathbb{C}[z_1, z_2]$. First assume that $g_1(z_1, z_2) = \sum_{k=0}^{a} h_k^1(z_2) z_1^k$ with $a > m_1$ and $h_a^1 \neq 0$. By *(i)* it then follows that g_2 also can be written as $g_2(z_1, z_2) = \sum_{k=0}^{a} h_k^2(z_2) z_1^k$ and that

$$h_a^1 d_1 = h_a^2 d_2.$$

By *(iii)* it follows that there is some $p \in \mathbb{C}[z_2]$ such that $h_a^1 = p d_2$ and $h_a^2 = p d_1$. By setting

$$\begin{aligned} &g_1'(z_1, z_2) = g_1(z_1, z_2) - z_1^{a-m_1} p(z_2) P_{v_2}(z_1, z_2) \\ &\Big(= \sum_{k=0}^{a-1} h_k^1(z_2) z_1^k + p(z_2) d_2(z_2) z_1^a - z_1^{a-m_1} p(z_2) \big(d_2(z_2) z_1^{m_1} + \text{lower powers of } z_1 \big) \Big) \end{aligned}$$

and

$$g_2'(z_1, z_2) = g_2(z_1, z_2) - z_1^{a-m_1} p(z_2) P_{v_1}(z_1, z_2),$$

we obtain a new pair such that $f = P_{v_1} g_1' - P_{v_2} g_2'$ with the property that g_j' can be written as $g_j'(z_1, z_2) = \sum_{k=0}^{a-1} h_k^j{}'(z_2) z_1^k$, $j = 1, 2$.

We now repeat the argument with the variables interchanged. It becomes important to note that the corresponding h_k^j's have degree $\leq m_1$ and that this property is preserved in the inductive process. We omit the details. At the end we conclude that we can assume that g_1 and g_2 are such that there are $w_1, w_2 \in \mathbb{M}_{\mathbf{m}}$ with $g_j = P_{w_j}$, which proves that $\overline{w} \in \mathcal{R}_1 + \mathcal{R}_2$, as desired. □

THEOREM 5.4. *In the generic case we have*

$$\operatorname{Ker} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix} = \mathcal{Z}_1 \cap \mathcal{Z}_2 \quad \textit{and} \quad \dim \mathcal{Z}_1 \cap \mathcal{Z}_2 = 2 m_1 m_2.$$

Moreover, $\# V(P_{v_1}, P_{v_2}) = 2 m_1 m_2$ *and*

$$\mathcal{Z}_1 \cap \mathcal{Z}_2 = \operatorname{span}\{ \boxed{z} : z \in V(P_{v_1}, P_{v_2}) \}.$$

Proof. Bernstein's theorem [5] implies that $\# V(P_{v_1}, P_{v_2}) = 2 m_1 m_2$ holds generically. We will assume that this is the case as well as that P_{v_1} and P_{v_2} are irreducible, that $\langle P_{v_1}, P_{v_2} \rangle$ is a radical ideal, and that Lemma 5.3 applies. The proof that this holds generically is postponed until Appendix B.

It is immediate by Proposition 4.3 and the assumption that P_{v_1} and P_{v_2} are irreducible, that $\operatorname{Ker} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix} = \mathcal{Z}_1 \cap \mathcal{Z}_2$. By Corollary 5.2 we have that

$$\dim \operatorname{Ran} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix} = 2(m_1 + 1)(m_2 + 1) - 1,$$

which implies that

$$\dim \operatorname{Ker} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix} = \dim \mathbb{M}_{2\mathbf{m}} - \dim \operatorname{Ran} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix} = 2 m_1 m_2,$$

and, hence, the first part is proved. As mentioned above, $\# V(P_{v_1}, P_{v_2}) = 2 m_1 m_2$ holds generically by Bernstein's theorem and, clearly, we have

$$\operatorname{span}\{ \boxed{z} : z \in V(P_{v_1}, P_{v_2}) \} \subset \mathcal{Z}_1 \cap \mathcal{Z}_2.$$

To conclude the proof, it suffices to show that

$$\operatorname{span}\{ \boxed{z} : z \in V(P_{v_1}, P_{v_2}) \}^{\perp} \subset (\mathcal{Z}_1 \cap \mathcal{Z}_2)^{\perp}. \tag{5.6}$$

We note that

$$(\mathcal{Z}_1 \cap \mathcal{Z}_2)^{\perp} = \left(\operatorname{Ker}\begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix}\right)^{\perp} = \operatorname{Ran}\begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix}^* = \mathcal{R}_1 + \mathcal{R}_2,$$

where the last equality follows by by Lemma 4.2. To prove (5.6), assume that $w \in \operatorname{span}\{\boxed{z} : z \in V(P_{v_1}, P_{v_2})\}^{\perp}$ is arbitrary. Then $P_{\overline{w}}(z) = 0$ for all $z \in V(P_{v_1}, P_{v_2})$ which in turn, via Hilbert's Nullstellensatz, implies that $P_{\overline{w}} \in \sqrt{\langle P_{v_1}, P_{v_2}\rangle} = \langle P_{v_1}, P_{v_2}\rangle$, where the last equality follows as $\langle P_{v_1}, P_{v_2}\rangle$ is assumed to be radical. By Lemma 5.3 this implies that $\overline{w} \in \mathcal{R}_1 + \mathcal{R}_2$, as desired. □

Pairs $(v_1, v_2) \in \mathbb{M}_{\mathbf{m}}$ such that Theorem 5.4 applies will in the further analysis be referred to as *proper*.

If we set $v_1 = u_n$ and $v_2 = u_{n'}$ then Theorem 5.4 proves claim *(i)* made in the beginning of this section, although it remains to be shown that a generic f gives rise to a proper pair $(u_n, u_{n'})$. This is the content of the next theorem, the proof of which is given in Appendix B.

THEOREM 5.5. *Given a generic $f \in \mathbb{M}_{2\mathbf{m}}$, the pair $(u_n, u_{n'})$ is proper for all $n \neq n'$. In particular, we generically have $\dim \mathcal{Z}_n \cap \mathcal{Z}_{n'} = 2m_1m_2$ and*

$$\mathcal{Z}_n \cap \mathcal{Z}_{n'} = \operatorname{span}\{\boxed{z_j}\}_{j=1}^{2m_1m_2},$$

where $\{z_j\}_{j=1}^{2m_1m_2} = V(P_{u_n}, P_{u_{n'}})$.

Once the z_j's have been found, *Proj*$_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f$ can be calculated by solving the normal equations to the linear system

$$f \approx \sum_{j=1}^{2m_1m_2} a_j \boxed{z_j}.$$

At this point it is straightforward to calculate the actual error $\|f - \mathit{Proj}_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f\|$, but it is advantageous to have an estimate of the error that does not involve computing the z_j's. In the next section we shall therefore give an upper bound for the error in terms of $u_n, u_{n'}$ that is simple to calculate.

5.2. The error $\|f - \mathit{Proj}_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f\|$.

Let $s_1 \geq s_2 \geq \ldots \geq s_{(2m_1+1)(2m_2+1)}$ be the singular values of the operator $\begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}$.

PROPOSITION 5.6. *Given $f \in \mathbb{M}_{2\mathbf{m}}$ such that Theorem 5.4 applies. Set $a = \dim \mathbb{M}_{2\mathbf{m}} - 2m_1m_2 - 1$. Then*

$$\|f - \mathit{Proj}_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f\| \leq \frac{\sqrt{\sigma_n^2 + \sigma_{n'}^2}}{s_a}.$$

Proof. For simplicity of notation we set $A = \begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}$. By standard operator theory, we can write $A = V\Sigma U$ where $U : \mathbb{M}_{2\mathbf{m}} \to \mathbb{C}^{(2m_1+1)(2m_2+1)}$ is unitary, $\Sigma : \mathbb{C}^{(2m_1+1)(2m_2+1)} \to \mathbb{C}^{(2m_1+1)(2m_2+1)}$ is a diagonal matrix with $s_1, s_2 \ldots, s_{(2m_1+1)(2m_2+1)}$ on the diagonal (in that order) and $V : \mathbb{C}^{(2m_1+1)(2m_2+1)} \to \mathbb{M}_{\mathbf{m}}^2$ is an isometry when restricted to $\operatorname{Ran}\Sigma$. Moreover $U(\operatorname{Ker} A) = \operatorname{Ker}\Sigma$ which, as $\dim \operatorname{Ker} A = \dim \mathcal{Z}_1 \cap \mathcal{Z}_2 = 2m_1m_2$ and $a = (2m_1+1)(2m_2+1) - 2m_1m_2 - 1$, implies that the last $2m_1m_2$ s_j's are zero and $s_a > 0$.

Because $y_p \perp \mathcal{Z}_n \cap \mathcal{Z}_{n'}$ and $(\sigma_n u_n, \sigma_{n'} u_{n'}) \in \operatorname{Ran} A$ we get that the equation $(\sigma_n u_n, \sigma_{n'} u_{n'}) = Ay_p$ is equivalent to

$$x = \Sigma U y_p,$$

where x is such that $Vx = (\sigma_n u_n, \sigma_{n'} u_{n'})$. We also have $\|x\| = \|(\sigma_n u_n, \sigma_{n'} u_{n'})\| = \sqrt{\sigma_n^2 + \sigma_{n'}^2}$ and $\|Uy_p\| = \|y_p\|$. By the above remarks it follows that Uy_p is orthogonal to $\operatorname{Ker}\Sigma$; since s_a is the first non-zero singular value, we get that

$$\|x\| \geq s_a \|Uy_p\|;$$

but then

$$\sqrt{\sigma_n^2 + \sigma_{n'}^2} \geq s_a \|y_p\|,$$

as desired. □

5.3. Statement *(ii)*.. In this subsection we set $n = (m_1+1)(m_2+1)$ and $n' = (m_1+1)(m_2+1) - 1$, because, intuitively, this choice should yield the least error. We construct a matrix $f \in \mathbb{M}_{\mathbf{m}}$ of function values, such that Theorem 5.5 applies, but

$$\mathit{Proj}_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f = 0.$$

The construction does not yield a "small" error. The results in this section are independent of the further developments and may therefore be skipped.

We recall that, given $v \in \mathbb{M}_{\mathbf{m}}$, $\dot{v} \in \mathbb{M}_{2\mathbf{m}}$ is the matrix with v in the upper left corner and zeroes everywhere else.

LEMMA 5.7. *Given a proper pair (v_1, v_2) and any $(\sigma_1, \sigma_2) \in \mathbb{C}^2$, let $\{z_j\}_{j=1}^{2m_1 m_2}$ be an enumeration of $V(P_{v_1}, P_{v_2})$. Then the quadratic linear system*

$$\left\{ \begin{array}{rll} \sigma_1 \overline{v_1(\mathbf{i})} & = \langle y_p, S_{\mathbf{i}} \overline{\dot{v}_1} \rangle, & \mathbf{0} \leq \mathbf{i} \leq \mathbf{m} \\ \sigma_2 \overline{v_2(\mathbf{i})} & = \langle y_p, S_{\mathbf{i}} \overline{\dot{v}_2} \rangle, & \mathbf{0} \leq \mathbf{i} \leq \mathbf{m}, \mathbf{i} \neq \mathbf{m} \\ 0 & = \langle y_p, \boxed{\overline{z_i}} \rangle, & i = 1, 2, .., 2m_1 m_2 \end{array} \right. \tag{5.7}$$

has a unique solution y_p, which is the smallest solution to (5.2).

Proof. Theorem 5.4 implies that the system of equations

$$\left\{ \begin{array}{rl} \sigma_1 \overline{v_1} = & D_{v_1} y_p \\ \sigma_2 \overline{v_2} = & D_{v_2} y_p \\ \operatorname{Ker} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix} \perp & y_p \end{array} \right. \tag{5.8}$$

has a solution, and that $y_p \perp \operatorname{Ker} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix}$ is equivalent with the third set of equations in (5.7). Moreover, the first set of equations in (5.7) is equivalent with $\sigma_1 \overline{v_1} = D_{v_1} y_p$ by the definition of D_{v_1}. The system (5.8) consists of $(2m_1+1)(2m_2+1)+1$ equations. The fact that

$$-P_{v_2} P_{v_1} + P_{v_1} P_{v_2} = 0$$

together with the assumption that $v_1(\mathbf{m}) \neq 0$ shows that $S_{\mathbf{m}} \dot{v}_2$ is linearly dependent on the set

$$\left\{ S_{\mathbf{i}} \overline{\dot{v}_1} : \mathbf{0} \leq \mathbf{i} \leq \mathbf{m} \right\} \bigcup \left\{ S_{\mathbf{i}} \overline{\dot{v}_2} : \mathbf{0} \leq \mathbf{i} \leq \mathbf{m} \text{ but } \mathbf{i} \neq \mathbf{m} \right\},$$

and therefore the "$\mathbf{m}$*th*" equation of $\sigma_2 \overline{v_2} = D_{v_2} y_p$ can be removed from the system (5.8) without changing the set of solutions. □

We note that the number of equations in (5.8) is equal to the number of unknowns in (elements of) y_p, and thus we conclude that given a proper pair (v_1, v_2) and any (σ_1, σ_2), we can calculate the smallest solution y_p to (5.2) by inverting a matrix representing this system of equations.

We observe that the ratio

$$\frac{\|y_p\|}{\|(\sigma_1 v_1, \sigma_2 v_2)\|} \tag{5.9}$$

in Lemma 5.7 can get arbitrarily large even for proper pairs (v_1, v_2). To see this, select $v_1, v_2 \in \mathbb{M}_{\mathbf{m}}$ such that (5.2), for $\sigma_1 = \sigma_2 = 1$, has no solution. Proposition 5.1 guarantees that this is possible. Then we take a sequence (v_1^i, v_2^i) of proper pairs that converge to (v_1, v_2) and let (y_p^i) be the corresponding sequence of smallest solutions to the system

$$\begin{pmatrix} \overline{v_1^i} \\ \overline{v_2^i} \end{pmatrix} = \begin{pmatrix} D_{v_1^i} \\ D_{v_2^i} \end{pmatrix} (y_p^i).$$

If the sequence $(\|y_p^i\|)$ were to be bounded, then we could take a convergent subsequence with limit y_p^∞, and by continuity this would solve the equation

$$\begin{pmatrix} \overline{v_1} \\ \overline{v_2} \end{pmatrix} = \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix} (y_p^\infty),$$

which contradicts Proposition 5.1.

We now have the necessary ingredients to generate the aforementioned existence of an $f \in \mathbb{M}_{\mathbf{m}}$ such that $\mathit{Proj}_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f = 0$. The idea is the following:

- Set $\sigma_n = 1$, $\sigma_{n'} = 2$, and pick $v_1, v_2 \in \mathbb{M}_{\mathbf{m}}$ such that
  - (a) $\|v_1\| = \|v_2\| = 1$ and $v_1 \perp v_2$
  - (b) P_{v_1} and P_{v_2} have a common factor and
  - (c) the by Proposition 5.1 associated matrices A_1 and A_2 satisfy $A_1 - 2A_2 \neq 0$.
- Alter v_1 and v_2 slightly to get two new vectors $v_1{}'$ and $v_2{}'$ that are proper, satisfy (a), and such that the corresponding smallest solution y_p to (5.2) is large.
- Use Lemma 5.7 to calculate y_p and set $f = y_p / \|y_p\|$.
- By the construction, $(1, v_1)$ and $(2, v_2)$ are con-eigenpairs to the operator $H_{y_p} = \|y_p\| H_f$. Check if $1/\|y_p\|$ and $2/\|y_p\|$ are the two smallest con-eigenvalues to H_f. If so, set $u_n = v_1$ and $u_{n'} = v_2$ and note that by the construction of y_p we have $f \perp \mathcal{Z}_n \cap \mathcal{Z}_{n'}$. If not, try again.

We have constructed an explicit counterexample, $f \in \mathbb{M}_{4,4}$, in Appendix A.

5.4. Numerical experiments. We carry out some numerical experiments. To facilitate these, we rephrase the results obtained so far in this section. Given $f \in \mathbb{M}_{2\mathbf{m}}$ we seek approximations of the form

$$f \approx \sum_{j=1}^{N} a_j \boxed{z_j}$$

with $N \ll (2m_1 + 1)(2m_2 + 1) \approx 4m_1 m_2$. In the generic case, given any n, n', we may assume that the con-eigenvectors $u_n, u_{n'}$ to H_f are proper. We first note that

$$\begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix} (f) = \begin{pmatrix} \sigma_n \overline{u_n} \\ \sigma_{n'} \overline{u_{n'}} \end{pmatrix} \tag{5.10}$$

and decompose f as

$$f = \mathit{Proj}_K f + \mathit{Proj}_{K^\perp} f, \tag{5.11}$$

with $K = \operatorname{Ker} \begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}$. It then turns out that, generically, $K = \mathcal{Z}_n \cap \mathcal{Z}_{n'}$, which is spanned by

$$\{\boxed{z_j}\}_{j=1}^{2m_1 m_2} = \{\boxed{z} : \ z \in V(P_{u_n}, P_{u_{n'}})\},$$

while $\mathit{Proj}_{K^\perp} f = y_p$ where y_p is the smallest solution to (5.10) considered as an equation with f as unknown. Thus (5.11) can be written

$$f = \mathit{Proj}_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f + y_p$$

where the error

$$\|f - \mathit{Proj}_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f\| = \|y_p\|$$

is estimated in Section 5.2.

Our computational methodology can be summarized as follows:

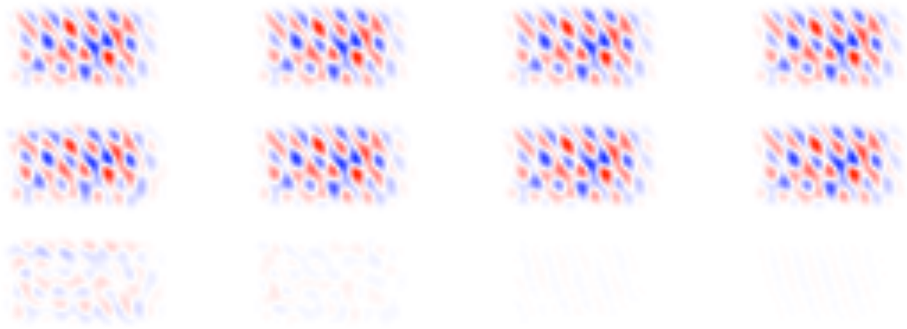

FIG. 2. *Original data and reconstructions on the box depicted in Figure 1, right. The panels (from left to right) display reconstruction using n = 8, 16, 32, and 64 quadrature nodes; $n' = n + 1$ and $2m_1m_2 = 2 \cdot 64 \cdot 32 = 4096$. In each panel the top shows the original and the middle the reconstruction. The residual is shown at the bottom.*

1. Sample Fourier transform data on boxes $B_{\nu,k}$ according to (2.17).
2. For each box, compute a Takagi factorization of H_f, cf. (3.3).
3. Choose con-eigenvectors u_n, $u_{n'}$ with σ_n, $\sigma_{n'}$ corresponding to the desired level of accuracy.
4. Compute the quadrature nodes $\{z_j\}_{j=1}^{2m_1m_2} = V(P_{u_n}, P_{u_{n'}})$.
5. Obtain the weights a_j from the linear system

$$f \approx \sum_{j=1}^{2m_1m_2} a_j \; \boxed{z_j} \; .$$

6. Restrict the roots z_j to those with associated weights a_j above the desired approximation level (ϵ).

As mentioned in Section 2.2, (1) can be realized by using the USFFT. Moreover, there exist fast algorithms to compute the Takagi factorization of H_f in (2) (see, for example, [23]). The delicate issue in (3) is which (n, n') and corresponding $(\sigma_n, \sigma_{n'})$ to choose. They should be chosen to ensure the desired level of accuracy, ϵ, in steps (5) and (6). However, in the absence of a sharp estimate, some scanning over (n, n') will be required. The computationally most demanding component in our approach is step (4), accurately solving the relevant algebraic problem. Step (5) comprises solving a system of normal equations. However, because of the possible occurrence of nodes that lie far away from the complex unit circles, some level of regularization becomes necessary to solve the relevant linear system of equations. We have opted for Tychonov regularization with a term $\alpha \sum_{j=1}^{2m_1m_2} |a_j|^2$ in the minimization thus promoting sparsity – one could employ an ℓ^1 minimization technique [16] instead. We find that, typically, in step (6) roughly n terms need to be kept; this is in accordance with what was observed for the one-dimensional result in [8].

We carry some numerical experiments, using the image in Figure 1, left. This image was generated by a wave equation and represents modelled seismic data, and contains caustics. Figure 1, right, shows the real part of the Fourier transform of the image; we consider the data on the box $(B_{\nu,k})$ depicted in this figure. In Figure 2, top row, we show $\mathrm{Re}\{\widehat{u}_{\nu,k}(\eta_1^{\nu,k})\}$, revealing the multiplication by $|\widehat{\chi}_{\nu,k}(\eta_1^{\nu,k})|^2$. The middle row shows the reconstructions using 8, 16, 32 and 64 ($= n$) quadrature nodes. In the bottom row, we show the residuals between the original and reconstructed sampled functions, confirming the convergence of our approach. We show the same results but in logrithmic scale in Figure 3, which demonstrates that, indeed, using (roughly) n terms in the approximation is a proper guiding principle. Moreover, the number of nodes (here, n) used in the reconstructions is much smaller than the total number of nodes, here $2m_1m_2 = 4096$, as desired.

In Figure 4, the partial reconstructions corresponding to the data on the specific box are illustrated. This figure captures how much the shapes of the wave packets have been adjusted to the data. Naturally, these reconstructions are restricted to a single scale, and hence represent only part of the frequency content of the original data. In Figure 5 we illustrate how our approach gleans information about the wavefront set of the

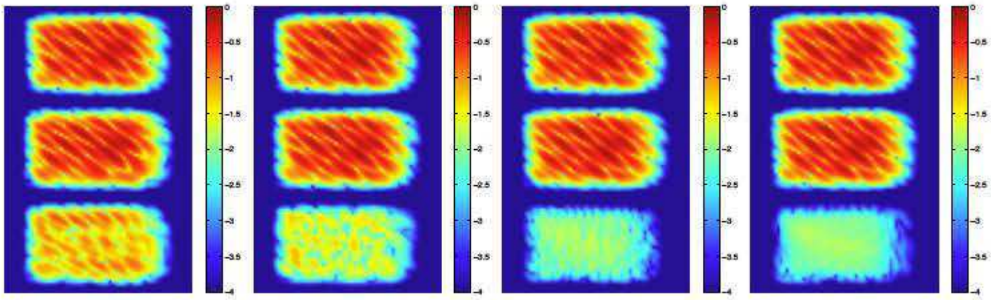

FIG. 3. *The counterpart of Figure 2 in logarithmic scale. The relevant con-eigenvalues are:* $\sigma_8 = 0.048817$, $\sigma_9 = 0.043787$; $\sigma_{16} = 0.013516$, $\sigma_{17} = 0.012902$; $\sigma_{32} = 0.001735$, $\sigma_{33} = 0.001618$; $\sigma_{64} = 0.000104$, $\sigma_{65} = 0.000102$.

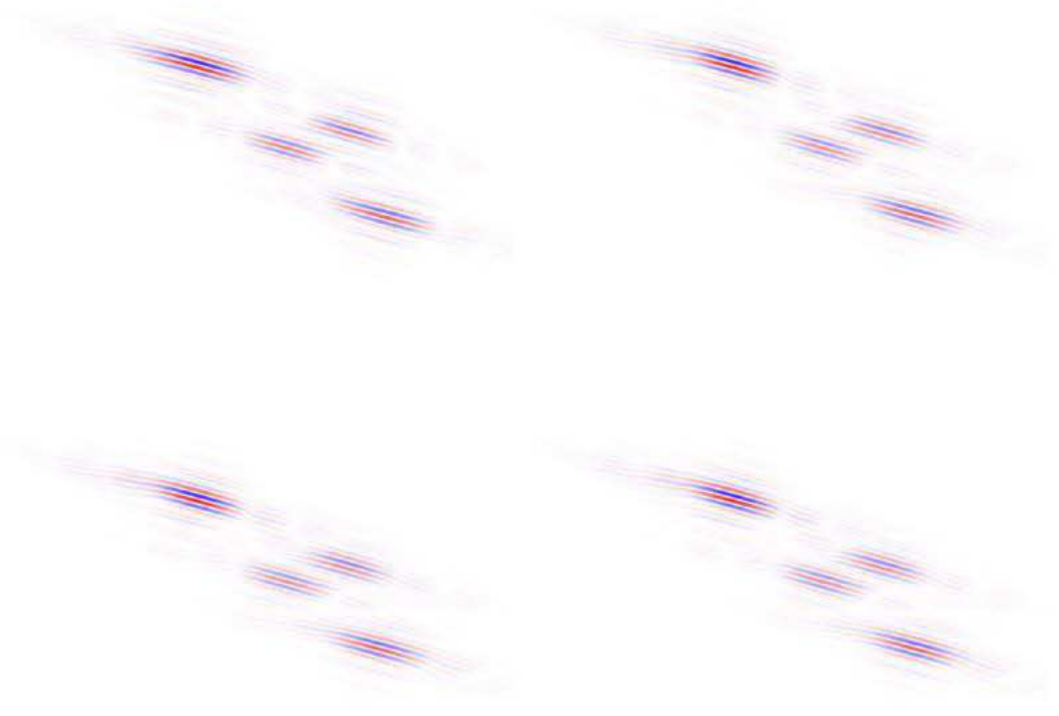

FIG. 4. *Partial reconstructions (from a single box, and hence single scale) in space. Top left:* $n = 8$*; top right:* $n = 16$*; bottom left:* $n = 32$*; bottom right:* $n = 64$*. In particular, we observe the improved localization with increasing* n.

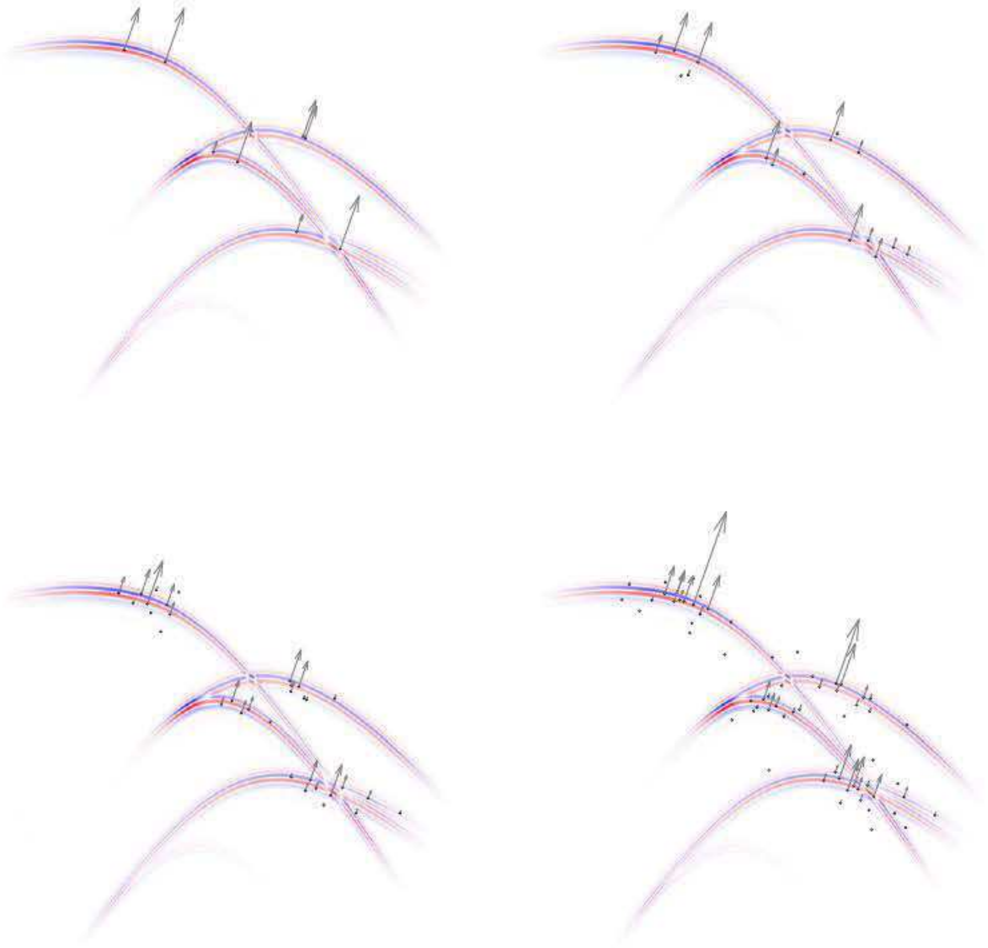

FIG. 5. *The positions* $(x_j^{\nu,k})$ *of the nodes are indicated by the blue dots, and the magnitude of the weights,* $|a_j|$, *are illustrated by the lengths of the black arrows. For the direction of the black arrows, we take the* ν *defining the box* $B_{\nu,k}$ *(in Figure 1). Top left:* $n =$ *8; top right:* $n =$ *16; bottom left:* $n =$ *32; bottom right:* $n =$ *64.*

image. In analogy with the curvelets (cf. (2.6)), we identify the imaginary parts of the complex logarithms of $z_j = z_j^{\nu,k}$ with positions (translations) $x_j^{\nu,k}$ through (2.16). The magnitudes of the weights, $|a_j|$, are illustrated by the lengths of the black arrows; the arrows point in the direction of ν.

6. Quadrature nodes: Reduction of polynomial degrees. A direct implementation of approach developed in the previous section has two drawbacks:

(*i*) it requires that we calculate more roots $\{z_j\}$ than we need for the approximation;
(*ii*) the polynomials P_{u_n} and $P_{u_{n'}}$ have high degrees.

Indeed, the number of samples represented by f is about $4m_1m_2$ while the number of roots are $2m_1m_2$, although numerical experiments show that we need far fewer terms in practice. The second point is a drawback primarily because of the complexity of the algorithm with increasing degrees for finding the $\{z_j\}$.

We shall bypass both issues in this section. The idea is most naturally explained in the one-dimensional setting of Section 3. There, $f \in \mathbb{C}^{2m+1}$, H_f is a square Hankel matrix

$$H_f = \begin{pmatrix} f(0) & f(1) & \dots & f(m) \\ f(1) & f(2) & \dots & f(m+1) \\ \vdots & \vdots & \ddots & \vdots \\ f(m) & f(m+1) & \dots & f(2m) \end{pmatrix},$$

and the singular value decomposition can be chosen such that $H_f = \overline{U}\Sigma U^*$, where the column vectors in U are the con-eigenvectors. This means that

$$H_f u_n = D_{u_n} f = \sigma_n \overline{u_n}.$$

The only point where this particular form of the singular value decomposition is used is in the proof that $D_{u_n} y = \sigma_n \overline{u_n}$ has a solution with norm less than σ_n, which implies that

$$\|f - \mathit{Proj}_{\mathcal{Z}_n} f\| \leq \sigma_n. \tag{6.1}$$

Now, if we were to write $2m = p + q$ with $q > p$ and follow the same strategy as above, with the operator

$$H_f^p = \begin{pmatrix} f(0) & \dots & f(p) \\ f(1) & \dots & f(p+1) \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ f(q) & \dots & f(p+q) \end{pmatrix}$$

we would loose the con-eigenvalue structure of the singular value decomposition and therefore the estimate (6.1), but the remaining parts of the argument would go through. More specifically, we carry out a standard singular value decomposition

$$H_f^p = V\Sigma U^*$$

and let $u_1, \dots, u_{p+1}$ respectively $v_1, \dots, v_{p+1}$ be the column vectors of U and V, respectively. We then define $D_{u_n}^p : \mathbb{C}^{2m+1} \to \mathbb{C}^q$ by

$$D_{u_n}^p = \begin{pmatrix} u_n(0) & u_n(1) & \dots & u_n(p) & 0 & \dots & 0 \\ 0 & u_n(0) & u_n(1) & \dots & u_n(p) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & & \\ 0 & \dots & 0 & u_n(0) & u_n(1) & \dots & u_n(p) \end{pmatrix}$$

and note that

$$H_f^p(u_n) = D_{u_n}^p f = \sigma_n v_n.$$

Then, with $\{z_k\}_{k=0}^p = V(P_{u_n})$, we can approximate f in $\mathcal{Z}_n^p = \operatorname{span}\{z_k\}_{k=0}^p$, with the only adjustment that the estimate (6.1) has to be replaced with an estimate involving the singular values of $D_{u_n}^p$ as well. However, both the size of P_{u_n} as well as $\mathcal{Z}_n^p$ were reduced by the above method.

We return to the two-dimensional case. We observed in Lemma 4.5 that (6.1) continues to hold in two variables, but that while approximating f in $\mathcal{Z}_n \cap \mathcal{Z}_{n'}$ we need to accept a weaker estimate of the form

$$\|f - \mathit{Proj}_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f\| \leq \frac{\sqrt{\sigma_n^2 + \sigma_{n'}^2}}{s_a},$$

where s_a is a singular value of $\begin{pmatrix} D_{u_n} \\ D_{u_{n'}} \end{pmatrix}$ (compare Proposition 5.6 and Section 5.3).

The approach based on "non-square H_f" carries naturally over to two variables. Set $2\mathbf{m} = \mathbf{p} + \mathbf{q}$ with $\mathbf{q} > \mathbf{p}$, define $H_f^{\mathbf{p}} : \mathbb{M}_{\mathbf{p}} \to \mathbb{M}_{\mathbf{q}}$ by

$$(H_f^{\mathbf{p}} u)(\mathbf{i}) = \sum_{\mathbf{0} \leq \mathbf{j} \leq \mathbf{p}} f(\mathbf{i} + \mathbf{j}) u(\mathbf{j}), \quad \mathbf{0} \leq \mathbf{i} \leq \mathbf{q}$$

and $D_u^{\mathbf{p}} : \mathbb{M}_{2\mathbf{m}} \to \mathbb{M}_{\mathbf{q}}$ by

$$(D_u^{\mathbf{p}} y)(\mathbf{i}) = \sum_{\mathbf{0} \leq \mathbf{j} \leq \mathbf{p}} y(\mathbf{i} + \mathbf{j}) u(\mathbf{j}), \quad \mathbf{0} \leq \mathbf{i} \leq \mathbf{q}.$$

We do a singular value decomposition for the operator $H_f^{\mathbf{p}}$ and get orthonormal vectors $u_1, \ldots, u_{(p_1+1)(p_2+1)} \in \mathbb{M}_{\mathbf{p}}$, orthonormal vectors $v_1, \ldots, u_{(p_1+1)(p_2+1)} \in \mathbb{M}_{\mathbf{q}}$ and singular values $\sigma_1, \ldots, \sigma_{(p_1+1)(p_2+1)}$ such that

$$\sigma_n v_n = H_f^{\mathbf{p}} u_n = D_{u_n}^{\mathbf{p}} f.$$

We consider n to be fixed. Let $\dot{u}_n$ be the element in $\mathbb{M}_{2\mathbf{m}}$ formed by adding zeros to the right and below the matrix $u_n \in \mathbb{M}_{\mathbf{p}}$, let $\mathcal{R}_n \subset \mathbb{M}_{2\mathbf{m}}$ be the subspace

$$\mathcal{R}_n = \operatorname{span}\{S_{\mathbf{j}} \overline{\dot{u}_n} : \mathbf{0} \leq \mathbf{j} \leq \mathbf{q}\}.$$

Let P_{u_n} be the polynomial given by $P_{u_n}(z) = \sum_{\mathbf{0} \leq \mathbf{j} \leq \mathbf{p}} u_n(\mathbf{j}) z^{\mathbf{j}}$ and set

$$\mathcal{Z}_n = \operatorname{span}\{\boxed{z} : z \in V(P_{u_n})\}.$$

The following proposition summarizes Proposition 4.3 and Corollary 4.4 in this more general setting. The proof is identical, hence it is omitted.

PROPOSITION 6.1. *Assume that P_{u_n} is reduced and that $u_n(\mathbf{m}) \neq 0$. Then*

$$\mathbb{M}_{2\mathbf{m}} = \mathcal{R}_n \oplus \mathcal{Z}_n$$

and $\mathcal{Z}_n = \operatorname{Ker} D_{u_n}$. In particular, $\dim \mathcal{Z}_n = \dim \mathbb{M}_{2\mathbf{m}} - \dim \mathbb{M}_{\mathbf{q}}$.

Next, we address the essential part of Theorem 5.5:

THEOREM 6.2. *Given a generic $f \in \mathbb{M}_{2\mathbf{m}}$ and any n, n' we have $\# V(P_{v_n}, P_{v_{n'}}) = 2p_1p_2$ and*

$$\mathcal{Z}_n \cap \mathcal{Z}_{n'} = \operatorname{span}\{\boxed{z} : z \in V(P_{u_n}, P_{u_{n'}})\}.$$

Proof. We prove in Appendix B that in the generic case, Theorem 5.4 applies to $u_n, u_{n'}$ (with $\mathbf{p} = \mathbf{m}$). The theorem then says that $\# V(P_{v_n}, P_{v_{n'}}) = 2p_1p_2$ and that $\{\boxed{z} : z \in V(P_{u_n}, P_{u_{n'}})\}$ is a linearly independent set, with the difference that in that notation we have $\boxed{z} \in \mathbb{M}_{2\mathbf{p}}$. In the notation of this section we have $\boxed{z} \in \mathbb{M}_{2\mathbf{m}}$, so as $\mathbf{p} \leq \mathbf{m}$ we conclude that the set $\{\boxed{z} : z \in V(P_{u_n}, P_{u_{n'}})\} \subset \mathbb{M}_{2\mathbf{m}}$ is linearly independent. It remains to prove that $\dim \mathcal{Z}_n \cap \mathcal{Z}_{n'} = 2p_1p_2$, which follows by the calculation

$$\dim \operatorname{Ran} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix}^* = 2 \dim \mathbb{M}_{\mathbf{q}} - \dim \operatorname{Ker} \begin{pmatrix} D_{v_1} \\ D_{v_2} \end{pmatrix}^* =$$
$$= 2 \dim \mathbb{M}_{\mathbf{q}} - (q_1 + 1 - p_1)(q_2 + 1 - p_2) = \ldots = \dim \mathbb{M}_{2\mathbf{m}} - 2p_1p_2$$

where we have used similar methods as in the proof of Proposition 5.1. □

Finally, we have the error estimate

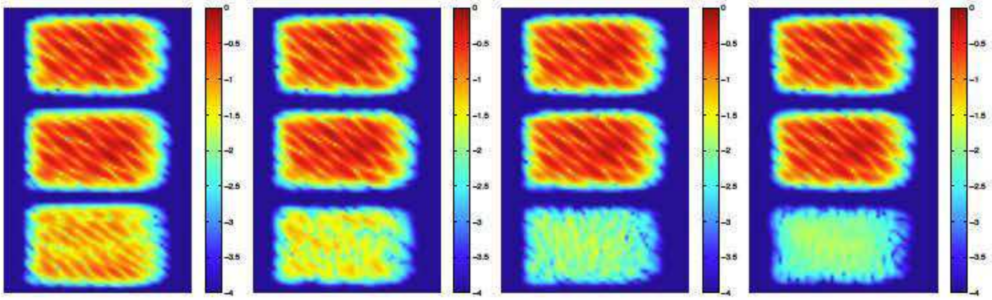

FIG. 6. *Original data and reconstructions, in logarithmic scale, for the box depicted in Figure 1, right. The panels (from left to right) display reconstructions based on a non-square Hankel matrix* ($p_1/q_1 = p_2/q_2 = 1/7$) *using* n *=8, 16, 32, and 64 quadrature nodes;* $2p_1p_2 = 256$. *In each panel the top shows the original, the middle the reconstruction, and the bottom the residual. Compare Figure 3.*

PROPOSITION 6.3. *Given* $f \in \mathbb{M}_{\mathbf{2m}}$ *such that Theorem 6.2 applies, set* $a = \dim \mathbb{M}_{2m} - 2p_1p_2 - 1$. *Then*

$$\|f - \mathit{Proj}_{\mathcal{Z}_n \cap \mathcal{Z}_{n'}} f\| \leq \frac{\sqrt{\sigma_n^2 + \sigma_{n'}^2}}{s_a}.$$

We conclude this section with numerical experiments using, again, the synthetic image shown in Figure 1, left, and the same ν, k that were used throughout Subsection 5.4. In Figure 6, top row, we show, in logarithmic scale, $\mathrm{Re}\{\widehat{u}_{\nu,k}(\eta_1^{\nu,k})\}$, The middle row shows the reconstructions using 8, 16, 32 and 64 ($= n$) quadrature nodes. In the bottom row, we show the residuals between the original and reconstructed sampled functions, confirming, again the convergence of our approach. Again, using (roughly) n terms in the approximation appears to be a working guiding principle. We used $p_{1,2} = m_{1,2}/4$ and $q_{1,2} = 2m_{1,2} - p_{1,2} = 7m_{1,2}/4$ with $m_1 = 64$ and $m_2 = 32$; hence, the total number of nodes was $2p_1p_2 = 256$.

7. Data application.

In this section, we demonstrate the performance of our approach in the presence of noise, and in the case of field seismic data. We begin with adding random noise (25 %) to the image shown in Figure 1; the noisy image is illustrated in the fequency domain, in Figure 7. We repeat the steps followed to generate an approximation by sums of wave packets illustrated in Figure 6. The result is shown in Figure 8. Upon comparing – and observing the similarity between – the top right and the bottom right panels, we notice that the approximation procedure has the effect, and is capable, of *denoising*. In Figure 9 we illustrate how our approach gleans information about the wavefront set of the image in the presence of noise.

We then proceed with applying our procedure to a so-called stacked seismic reflection data section, extracted from TotalFinaElf's L7D survey acquired in the North Sea, and shown in Figure 10, left. The approximation by sums of wave packets is illustrated in Figure 10, right. The effective compression rate in this particular result is about 50. Indeed, collectively, with the results illustrated in Figures 3, 6 and 10, we confirm the *compression* capability of our approach.

Acknowledgments. The authors would like to thank Dr. Uli Walther for many fruitful discussions.

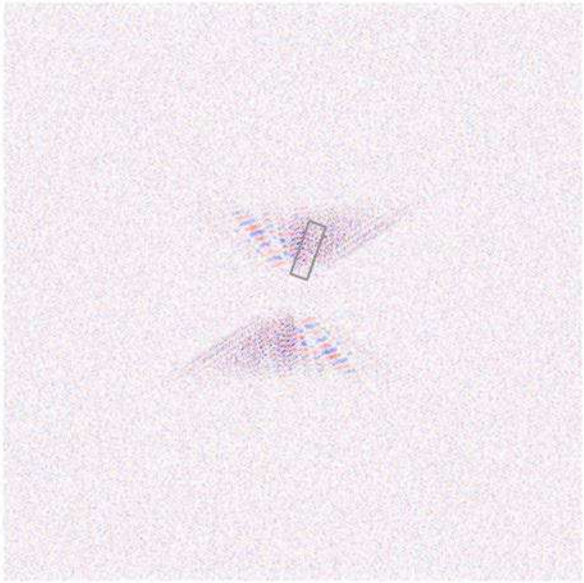

FIG. 7. *The data of Figure 1, left, contaminated with additive random noise. Counterpart of Figure 1, right.*

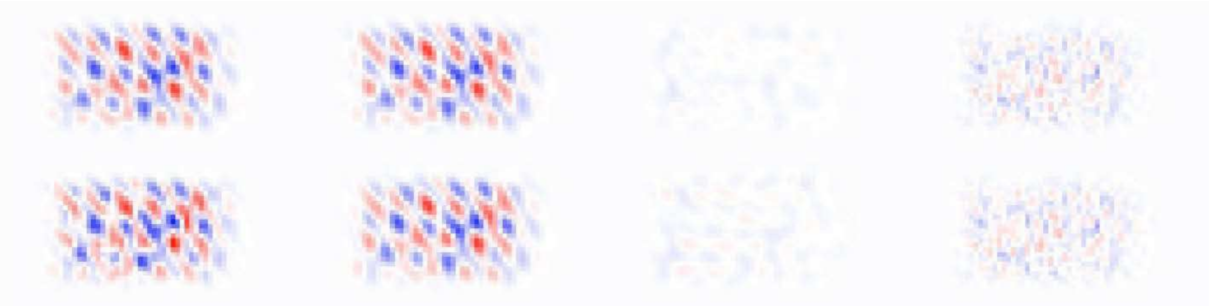

FIG. 8. *Top row, from left to right: Original windowed, Fourier transformed image (real part), reconstruction (approximation) using the original, difference between the original and the reconstruction using the original, real part of the difference between the windowed, Fourier transformed noisy image and the reconstruction using the noisy image ("denoising"); bottom row, from left to right: Windowed, Fourier transformed noisy image (real part), reconstruction (approximation) using the noisy image, difference between the original and the reconstruction using the noisy image, real part of the difference between the original windowed, Fourier transformed image and the windowed, Fourier transformed noisy image;* $n = 16$*. Counterpart of Figure 2.*

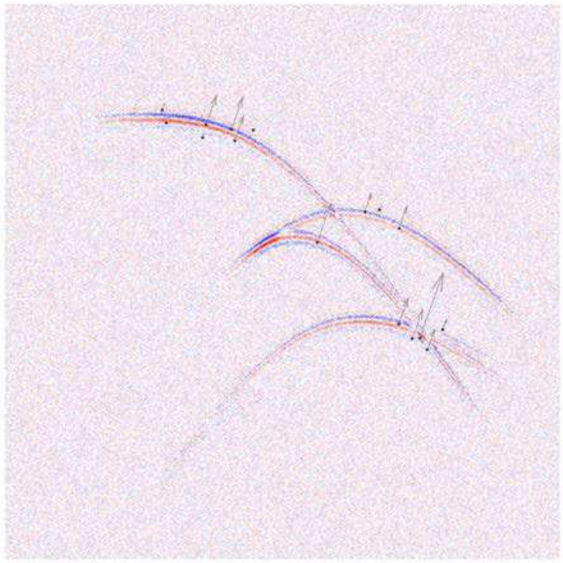

FIG. 9. *Counterpart of Figure 5 starting from the noisy image illustrated in Figure 7; $n = 16$. The nodes are stable under random perturbations.*

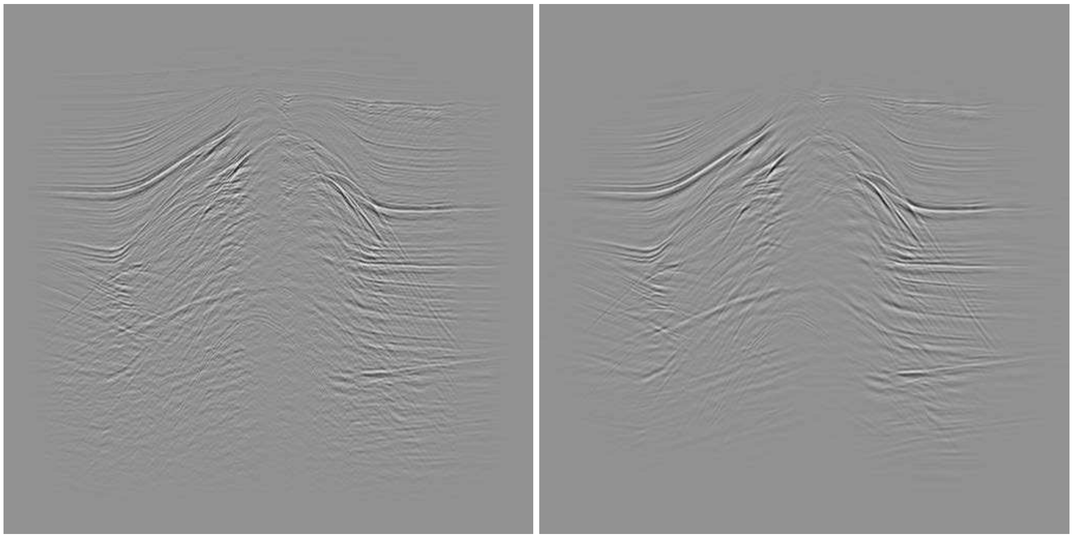

FIG. 10. *A stacked exploration seismic (North Sea) data section (512^2 samples; left). Right: reconstruction with* 5130 *packets (accounting for oversampling) with an effective compression rate of about* 50.

Appendix A. An example of a degenerate case: Failure of convergence.

Using the procedure outlined at the end of section 5.3 we obtain the following explicit example of an $f \in \mathbb{M}_{(4,4)}$ such that

$$f \perp \operatorname{span}\{\boxed{z} : z \in V(P_{u_n}, V(P_{u_{n'}})\},$$

where σ_n and $\sigma_{n'}$ are the smallest con-eigenvalues.

EXAMPLE A.1. Let $m_1 = m_2 = 2$ and take $v_1, v_2 \in \mathbb{M}_{(2,2)}$ such that

$$P_{v_1}(x, y) = (x - 1)(x(y^2 + 2y - 8) + (y^2 + 3))/n_1$$

and

$$P_{v_1}(x, y) = (x - 1)(x(2y^2 - 3y - 1) + (5y^2 - 1))/n_2$$

where $n_1 = 14.2$ and $n_2 = 6.29$ are constants so that $\|v_1\| = \|v_2\| = 1$. Recall Proposition 5.1. Setting $q(x, y) = (x - 1)$ we easily obtain that $A_1, A_2 \in \mathbb{M}_{(1,0)}$ are given by

$$A_1 = \begin{pmatrix} -7 \\ -7 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 7 \\ 7 \end{pmatrix}$$

We now look for v_1', v_2' of the form

$$v_1'^{\,\epsilon} = \begin{pmatrix} -3 & 0 & -1 \\ 11 & -2 & 0 \\ -8 & 2 & 1+\epsilon \end{pmatrix} /n_1{}^{\epsilon}$$

$$v_2'^{\,\epsilon} = \begin{pmatrix} 1 & 0 & -5 \\ 0 & 3+\epsilon & 3 \\ -1 & -3 & 2 \end{pmatrix} /n_2{}^{\epsilon}$$

where $n_1{}^{\epsilon}$ and $n_2{}^{\epsilon}$ are normalizing constants. Note that $v_1'^{\,\epsilon} \perp v_2'^{\,\epsilon}$ and that $v_1'^{\,0} = v_1$, $v_2'^{\,0} = v_2$. It turns out that the value $\epsilon = 0.02$ yields a proper pair and that the corresponding solution y_p to the equation system (5.7) is such that the singular values of the operator H_{y_p} are

$$\sigma_9 = 1,\ \sigma_8 = 2,\ \sigma_7' = 2.70, \ldots,\ \sigma_1' = 289.$$

Thus setting

$$f = \frac{y_p}{\|y_p\|} = \frac{y_p}{205} = 10^{-2} \begin{pmatrix} -5.57 & 2.84 & 19.07 & 25.66 & -27.10 \\ 2.18 & -16.37 & 16.52 & 19.90 & -31.32 \\ 2.58 & -25.04 & 14.62 & 17.36 & -29.44 \\ 0.59 & -29.15 & 12.66 & 15.71 & -37.24 \\ 0.21 & -29.94 & 14.99 & 19.30 & -11.35 \end{pmatrix}$$

we obtain an $f \in \mathbb{M}_{44}$ with lowest con-eigenvalues $\sigma_n = 1/205$, $\sigma_{n'} = 2/205$ and corresponding con-eigenvectors $u_n = \overset{\circ\ 0.02}{v_1}$, $u_{n'} = \overset{\circ\ 0.02}{v_2}$. In particular, by the construction we get

$$f \perp \operatorname{span}\{\boxed{z} : z \in V(P_{u_n}, P_{u_{n'}})\}.$$

Appendix B. The notion of "holds generically".

In algebraic geometry, a statement S is often said to hold generically if the set where it fails is contained in a proper algebraic variety. More precisely, if S concerns u for $u \in \mathbb{M}_{\mathbf{m}}$, say, then S holds generically if there are polynomials $q_1, \ldots, q_n \in \mathbb{C}[\mathbb{M}_{\mathbf{m}}]$ such that S is true whenever $u \notin V(q_1, \ldots, q_n)$. ($\mathbb{C}[\mathbb{M}_{\mathbf{m}}]$ denotes all polynomials with the entries in $\mathbb{M}_{\mathbf{m}}$ as variables).

In this paper we use the definition that a statement S holds generically if the set where it fails has zero Lebesgue measure. It is not hard to see that this is a weaker definition, i.e. all proper algebraic varieties have Lebesgue measure zero. On any linear space we will denote the Lebesgue measure by L. [5]

Let $A_{\mathbf{m}}$ denote the area-measure on $\mathbb{S}(\mathbb{M}_{\mathbf{m}})$ - the unit sphere in $\mathbb{M}_{\mathbf{m}}$. In case the statement S concerns only u with u normalized, i.e. $u \in \mathbb{S}(\mathbb{M}_{\mathbf{m}})$, we will say that it holds generically if the set where it fails has zero $A_{\mathbf{m}}-$measure.

The following proposition shows that Proposition 4.3, Corollaries 4.4, 4.6 and 5.2 hold generically.

PROPOSITION B.1. *Given $u \in \mathbb{S}(\mathbb{M}_{\mathbf{m}})$ (or $u \in \mathbb{M}_{\mathbf{m}}$), we generically have that P_u is irreducible and $u(\mathbf{m}) \neq 0$.*

Proof. The second statement is trivial so we will only prove the first. Set $(\mathbb{M}_{\mathbf{m}})_{ir} = \{u \in \mathbb{M}_{\mathbf{m}} : P_u \text{ is irreducible}\}$. Clearly, $u \in \mathbb{S}(\mathbb{M}_{\mathbf{m}})$ is such that P_u is irreducible if and only if it the same holds for αu for all $\alpha \in \mathbb{C} \setminus \{0\}$, and therefore it suffices to show that $L((\mathbb{M}_{\mathbf{m}})_{ir}) = 0$. Given $u \in (\mathbb{M}_{\mathbf{m}})_{ir}^c$ there exists a $\mathbf{0} < \mathbf{k} < \mathbf{m}$ and $u_1 \in \mathbb{M}_{\mathbf{k}}$, $u_2 \in \mathbb{M}_{\mathbf{m-k}}$ such that $P_u = P_{u_1}P_{u_2}$. If we hold $\mathbf{k}$ fixed and consider the entries in u_1, u_2 as variables, it is easily seen that the set of such u's is the image of a $\mathbb{M}_{\mathbf{m}}$-valued polynomial on $\mathbb{C}^{\dim \mathbb{M}_{\mathbf{k}} + \dim \mathbb{M}_{\mathbf{m-k}}}$. As $\dim \mathbb{M}_{\mathbf{k}} + \dim \mathbb{M}_{\mathbf{m-k}} < \dim \mathbb{M}_{\mathbf{m}}$, it follows by Theorem 1, Section 3.3 [15] that such a set is contained in a proper algebraic variety, and therefore it has Lebesgue measure zero. □

We now consider $(u_1, u_2) \in (\mathbb{M}_{\mathbf{m}})^2$. The first objective is to prove that in the generic case, $\langle P_{u_1}, P_{u_2} \rangle$ is a radical ideal. For $w \in \mathbb{C}^2$ let $E_w : \mathbb{C}[z_1, z_2] \to \mathbb{C}$ denote the functional of evaluation at w.

LEMMA B.2. *Let $I \subset \mathbb{C}[z_1, z_2]$ be an ideal such that $\sqrt{I} = E_{\mathbf{0}}$. Then I is radical if and only if it contains elements p_1, p_2 of the form*

$$p_1(z_1, z_2) = z_1 + \{\textit{monomials with degree } \geq 2\}$$
$$p_2(z_1, z_2) = z_2 + \{\textit{monomials with degree } \geq 2\}.$$

Proof. The "only if"-direction is obvious, so we will only prove the "if"-part. By Hilberts Nullstellensatz there is an $n \in \mathbb{N}$ such that $z_1^n \in I$ and $z_2^n \in I$, and hence $E_{\mathbf{0}}^{2n} \subset I$. Given any $p \in E_{\mathbf{0}}$, it is easy to see that we can find a $q \in E_{\mathbf{0}}$ such that $p - q(p_1(z), p_2(z)) \in E_{\mathbf{0}}^{2n} \subset I$, which completes the proof. □

The next proposition together with Proposition B.1 and Bernstein's theorem implies that Theorem 5.4 indeed holds generically. (Technically, we also have to show that Lemma 5.3 holds generically, but this is almost immediate so we omit the argument.)

PROPOSITION B.3. *Given $u_1, u_2 \in \mathbb{S}(\mathbb{M}_{\mathbf{m}})$, (or $u_1, u_2 \in \mathbb{M}_{\mathbf{m}}$), the ideal $\langle P_{u_1}, P_{u_2} \rangle$ is radical in the generic case.*

Proof. Like in Proposition B.1 it suffices to prove the statement for $u_1, u_2 \in \mathbb{M}_{\mathbf{m}}$. By Proposition B.1 we may assume that P_{u_1} and P_{u_2} have no common factor. By the Lasker-Noether theorem we can then write

$$\langle P_{u_1}, P_{u_2} \rangle = \cap_{j=1}^N I_j$$

where $\sqrt{I_j} = E_{w_j}$ for some $w_j \in \mathbb{C}^2$. Given $g_1, g_2 \in \mathbb{C}[z_1, z_2]$ let $Res(g_1, g_2, z_1)$ denote the residual of g_1 and g_2 considered as polynomials in the variable z_1 with coefficients in $\mathbb{C}[z_2]$ (see [15]). Set $g(z_2) = Res(P_{u_1}, P_{u_2}, z_1)$ and recall that $g \in \langle P_{u_1}, P_{u_2} \rangle$ (Section 3.6, Proposition 1 [15]). g has a multiple zero if and only if

$$Res(g, \frac{d}{dz_2}g, z_2) = 0.$$

Written out this means that a number of polynomials on $\mathbb{M}_{\mathbf{m}}^2$ should vanish for (u_1, u_2), and therefore it generically holds that g has no multiple zero. Switching positions of z_1 and z_2 and applying Lemma B.2 we deduce that $I_j = E_{w_j}$ for each j. Hence $\langle P_{u_1}, P_{u_2} \rangle = \cap_{j=1}^N E_{w_j}$ which yields the desired result. □

[5] This definition is imprecise, but we ignore this as we will only be interested in sets of measure zero.

We will now prove that Theorem 5.5 holds generically, and for this we will need several lemmas.

LEMMA B.4. *Given a generic $f \in \mathbb{M}_{\mathbf{2m}}$, all con-eigenvectors u_n are such that P_{u_n} is irreducible.*

Proof. For all $\mathbf{0} \leq \mathbf{k} \leq \mathbf{m}$, let $\mathcal{A}_{\mathbf{k}}$ be the set of all $f \in \mathbb{M}_{\mathbf{2m}}$ such that there exists an n with the property that $P_{u_n} = q_1 q_2$ where $q_1, q_2 \in \mathbb{C}[z_1, z_2]$ and $\deg q_1 \leq \mathbf{k}$, $\deg q_2 \leq \mathbf{m} - \mathbf{k}$. To complete the proof it clearly suffices to show that each $\mathcal{A}_{\mathbf{k}}$ has zero Lebesgue measure for fixed $\mathbf{k}$.

By standard arguments from integration theory, it follows that if $\beta : \mathbb{R}^{N_1} \to \mathbb{R}^{N_2}$ is a differentiable function and $N_2 > N_1$, then $\operatorname{Im}\beta = \{\beta(y) : y \in \mathbb{R}^{N_1}\}$ has zero Lebesgue measure. We will prove the lemma by showing that there exists a sequence of functions $\beta_j : \mathbb{R}^{N_1} \to \mathbb{C}^{\dim \mathbb{M}_{\mathbf{2m}}}$, ($j = 1, \ldots, \infty$), with $N_1 < 2 \dim \mathbb{M}_{\mathbf{2m}}$ and $\mathcal{A}_{\mathbf{k}} \subset \cup_{j=1}^{\infty} \operatorname{Im}\beta_j$.

For any $v \in \mathbb{M}_{\mathbf{k}}$ and $w \in \mathbb{M}_{\mathbf{m}-\mathbf{k}}$, define $u = u(v, w) \in \mathbb{M}_{\mathbf{m}}$ by requiring that $P_u = P_v P_w$. Recall that u is a con-eigenvector to an $f \in \mathbb{M}_{\mathbf{2m}}$ if and only if

$$\sigma \overline{u} = D_u f \tag{B.1}$$

for some $\sigma \geq 0$. By Lemma 4.2 we have that D_u^* is injective so

$$\dim \operatorname{Ker} D_u = \dim \mathbb{M}_{\mathbf{2m}} - \dim \mathbb{M}_{\mathbf{m}}.$$

Let $(v_0, w_0) \in \mathbb{M}_{\mathbf{k}} \times \mathbb{M}_{\mathbf{m}-\mathbf{k}}$ be fixed but arbitrary and denote $u(v_0, w_0)$ by u_0. Set $M = \dim \mathbb{M}_{\mathbf{2m}} - \dim \mathbb{M}_{\mathbf{m}}$ and choose a basis $e_1, \ldots, e_M$ for $\operatorname{Ker} D_{u_0}$. For any $a \in \mathbb{C}^M$, $u(v, w) \in \mathbb{M}_{\mathbf{m}}$ and $\sigma \in \mathbb{R}$ consider the following equation-system with $f \in \mathbb{M}_{\mathbf{2m}}$ as unknown:

$$\left\{ \begin{array}{ll} \sigma \overline{u} & = D_u f \\ a(k) & = \langle f, e_k \rangle, \qquad k = 1, \ldots, M \end{array} \right. \tag{B.2}$$

If we order the $\dim \mathbb{M}_{\mathbf{2m}}$ equations in (B.2) and identify $\mathbb{M}_{\mathbf{2m}}$ with $\mathbb{C}^{\dim \mathbb{M}_{\mathbf{2m}}}$ via a unitary map U, this can be written as $\alpha_0 = \Lambda_0 U f$, where $\alpha_0 \in \mathbb{C}^{\dim \mathbb{M}_{\mathbf{2m}}}$ and Λ^0 is an $\dim \mathbb{M}_{\mathbf{2m}} \times \dim \mathbb{M}_{\mathbf{2m}}$-matrix. We will consider α_0 in the natural way as a function of $(v, w, \sigma, a) \in \mathbb{M}_{\mathbf{k}} \times \mathbb{M}_{\mathbf{m}-\mathbf{k}} \times \mathbb{R} \times \mathbb{C}^M$, and Λ_0 as a function of $(v, w) \in \mathbb{M}_{\mathbf{k}} \times \mathbb{M}_{\mathbf{m}-\mathbf{k}}$. For (v, w) in an open neighborhood $\mathcal{O}_0$ of (v_0, w_0), Λ_0 is invertible, so if f satisfies (B.1) for some $\sigma \geq 0$ and $u = u(v, w)$ with $(v, w) \in \mathcal{O}_0$ then $f = U^{-1}\Lambda_0^{-1}(v, w)\alpha_0(v, w, \sigma, a)$ for some $a \in \mathbb{C}^M$. Define $\beta_0 : \mathcal{O}_0 \times \mathbb{R} \times \mathbb{C}^M \to \mathbb{M}_{\mathbf{2m}}$ via $\beta_0 = U^{-1}(\Lambda_0(v, w))^{-1}\alpha_0(v, w, \sigma, a)$.

By a compactness argument, we can choose a sequence $\big((v_j, w_j)\big)_{j=1}^{\infty}$ in $\mathbb{M}_{\mathbf{k}} \times \mathbb{M}_{\mathbf{m}-\mathbf{k}}$ such that the corresponding sets $\mathcal{O}_j$ cover $\mathbb{M}_{\mathbf{k}} \times \mathbb{M}_{\mathbf{m}-\mathbf{k}}$. If $f \in \mathcal{A}_{\mathbf{k}}$, then there exists some j such that f satisfies (B.2) for some $(v, w, \sigma, a) \in \mathcal{O}_j \times \mathbb{R} \times \mathbb{C}^M$, and hence

$$\mathcal{A}_{\mathbf{k}} \subset \cup_{j=1}^{\infty} \operatorname{Im}\beta_j.$$

Moreover, each β_j is defined on $\mathcal{O}_j \times \mathbb{R} \times \mathbb{C}^M$ which is an open subset of $\mathbb{M}_{\mathbf{k}} \times \mathbb{M}_{\mathbf{m}-\mathbf{k}} \times \mathbb{R} \times \mathbb{C}^M$, which clearly can be identified with $\mathbb{R}^{N_1}$ where

$$\begin{aligned} N_1 &= 2\dim \mathbb{M}_{\mathbf{k}} + 2\dim \mathbb{M}_{\mathbf{m}-\mathbf{k}} + 1 + 2M = \\ &= 2\dim \mathbb{M}_{\mathbf{k}} + 2\dim \mathbb{M}_{\mathbf{m}-\mathbf{k}} + 1 + 2(\dim \mathbb{M}_{\mathbf{2m}} - \dim \mathbb{M}_{\mathbf{m}}) = \\ &= 2\dim \mathbb{M}_{\mathbf{2m}} + 1 - 2(\dim \mathbb{M}_{\mathbf{m}} - \dim \mathbb{M}_{\mathbf{k}} - \dim \mathbb{M}_{\mathbf{m}-\mathbf{k}}) \end{aligned}$$

The last parenthesis is always a positive integer, and hence $N_1 < 2\dim \mathbb{M}_{\mathbf{2m}}$. By the remarks in the beginning, the proof is complete. □

Let *ort* denote the polynomial $ort(u_1, u_2) = \sum_{\mathbf{0} \leq \mathbf{k} \leq \mathbf{m}} u_1(\mathbf{k}) u_2(\mathbf{k})$ on $\mathbb{M}_{\mathbf{m}}^2$.

LEMMA B.5. *There exists a set of polynomials $\{p_m : 0 \leq m \leq M\}$ on $\mathbb{M}_{\mathbf{m}}^2$ such that*

(i) $V(p_0, \ldots, p_M)$ contains the set of all non-proper pairs.

(ii) $\forall m\ \exists d_1(m), d_2(m) \in \mathbb{N}$ such that $p_m(ru_1, su_2) = r^{d_1(m)} s^{d_2(m)} p_m(u_1, u_2)$

(iii) $\forall m$, $ort \nmid p_m$, (i.e. ort does not divide p_m).

Proof. Using the proof of Propositions B.1 and B.3, Bernstein's theorem and basic algebraic geometry it is clear that there exists a non-zero polynomial q such that $V(q)$ contains the set of non-proper pairs. If q has properties (ii) and (iii) we set $p_0 = q$ and $M = 0$. Otherwise, we define the polynomials $\{p'_{(a,b)}\}_{(a,b)\in\mathbb{N}^2}$ on $\mathbb{M}^2_{\mathbf{m}}$ via the identity

$$q(ru_1, su_2) = \sum_{(a,b)} p'_{(a,b)}(u_1, u_2) r^a s^b$$

for all $(u_1, u_2) \in \mathbb{M}^2_{\mathbf{m}}$ and $r, s \in \mathbb{C}$. Clearly the amount of non-zero $p'_{(a,b)}$ are finite and each $p'_{(a,b)}$ satisfies (ii). Rename the non-zero $p'_{(a,b)}$'s to $p'_0, \ldots, p'_M$. We claim that the set of non-proper pairs is contained in the variety $V(p'_0, \ldots, p'_M)$. Assume not and let (v_1, v_2) be a non-proper pair such that $p'_m(v_1, v_2) \neq 0$ for some m, $0 \leq m \leq M$. Then we can find a $r, s \in \mathbb{C}$ such that $q(rv_1, sv_2) \neq 0$, which is a contradiction because the pair (rv_1, rv_2) is also non-proper.

For $m = 0 \ldots, M$ let n_m be numbers such that $p'_m = ort^{n_m} p_m$ where $ort \nmid p_m$. Clearly each p_m satisfy (ii) as well. The proof is complete if we show that the set of non-proper pairs is contained in the variety $V(p_0, \ldots, p_M)$. To see this, assume the contrary and let (u_1, u_2) be a non-proper pair such that $p_m(u_1, u_2) \neq 0$ for some m. A short argument shows that we may choose a non-zero $r \in \mathbb{C}$ such that the pair $(u_1, u_2 + ru_1)$ is non-proper and $p_m(u_1, u_2 + ru_1) \neq 0$, (in case u_2 is reducible one might need to take $(u_1 + ru_2, u_2)$ instead). But u_1 and $u_2 + ru_1$ are not orthogonal and therefore

$$p'_m(u_1, u_2 + ru_1) = ort^{n_m}(u_1, u_2 + ru_1) p_m(u_1, u_2 + ru_1) \neq 0,$$

which is a contradiction. □

Let $\mathcal{B} \subset \mathbb{S}(\mathbb{M}_{\mathbf{m}})^2$ denote the set of orthonormal pairs, i.e.

$$\mathcal{B} = \{(u_1, u_2) \in \mathbb{S}(\mathbb{M}_{\mathbf{m}})^2 : u_1 \perp u_2\}.$$

Let $\mathcal{B}_{ir}$ be the set of pairs (u_1, u_2) in $\mathcal{B}$ such that both P_{u_1} and P_{u_2} are irreducible, and let $\mathcal{B}_{np}$ be the set of pairs (u_1, u_2) in $\mathcal{B}$ that are not proper. $\mathcal{B}$ can be considered as a differentiable manifold with real dimension $4 \dim \mathbb{M}_{\mathbf{m}} - 4$. In order to simplify notation, instead of defining the hole atlas we will specify one typical chart $T : \mathbb{R}^{4 \dim \mathbb{M}_{\mathbf{m}} - 4} \to \mathcal{B}$. This is done as follows:

Define $S : \mathbb{C}^{\dim \mathbb{M}_{\mathbf{m}}} \to \mathbb{M}_{\mathbf{m}}$ via

$$S(c) = \begin{pmatrix} c(0) & c(1) & \ldots & c(m_1) \\ c(m_1+1) & \ddots & \ldots & c(2m_1+1) \\ \vdots & \ddots & \vdots & \vdots \\ c((m_1+1)m_2) & \ldots & \ldots & c(\dim \mathbb{M}_{\mathbf{m}} - 1) \end{pmatrix}.$$

Define $B : \mathbb{C}^{2\dim \mathbb{M}_{\mathbf{m}} - 3} \to \mathbb{C}$ via $B(c) = \overline{c(0)} + \sum_{k=1}^{\dim \mathbb{M}_{\mathbf{m}} - 2} \overline{c(k)} c(\dim \mathbb{M}_{\mathbf{m}} - 2 + k)$ and $V_1, V_2 : \mathbb{C}^{2\dim \mathbb{M}_{\mathbf{m}} - 3} \to \mathbb{C}^{\dim \mathbb{M}_{\mathbf{m}}}$ via

$$\begin{aligned} V_1(c) &= (1, c(0), c(1), \ldots, c(\dim \mathbb{M}_{\mathbf{m}} - 2)) \\ V_2(c) &= (-B(c), 1, c(\dim \mathbb{M}_{\mathbf{m}} - 1), \ldots, c(2\dim \mathbb{M}_{\mathbf{m}} - 4)) \end{aligned}$$

Note that B was defined such that $V_1(c) \perp V_2(c)$ for all $c \in \mathbb{C}^{2\dim \mathbb{M}_{\mathbf{m}} - 3}$. Finally define $T : \mathbb{C}^{2\dim \mathbb{M}_{\mathbf{m}} - 3} \times \mathbb{R} \times \mathbb{R} \to \mathcal{B}$ via

$$T(c, r_1, r_2) = \Big(\frac{(1 + \mathrm{i} r_1) S(V_1(c))}{\|(1 + \mathrm{i} r_1) S(V_1(c))\|}, \frac{(1 + \mathrm{i} r_2) S(V_2(c))}{\|(1 + \mathrm{i} r_2) S(V_2(c))\|} \Big). \tag{B.3}$$

By identifying $\mathbb{C}$ with $\mathbb{R}^2$ as usual we will regard T as a map from $\mathbb{R}^{4\dim \mathbb{M}_{\mathbf{m}} - 4}$ into $\mathcal{B}$. It is easy to see that T is injective and that it covers most of $\mathcal{B}$. By changing the definition of S one obtains other maps like T that together form an atlas for $\mathcal{B}$, as desired. We denote this collection of maps by $\{T_l\}_{l=1}^N$ where N is some finite number.

LEMMA B.6. *Let $\{T_l\}_{l=1}^N$ be as above and let l be fixed. Then $T_l^{-1}(\mathcal{B}_{np})$ has zero Lebesgue measure.*

Proof. Assume for simplicity that $T_l = T$; the chart defined via (B.3). Let $p_1, \dots, p_M$ be the polynomials given by Lemma B.5. Then $\mathcal{B}_{np} \subset V(p_1, \dots, p_M)$ by (i). By (ii) and the definition of T it is easy to see that for each p_m there is a polynomial q_m on $\mathbb{R}^{4\dim \mathbb{M}_{\mathbf{m}}-4}$ such that

$$p_m(T(r)) = \frac{q_m(r)}{n_1(r)^{d_1(m)} n_2(r)^{d_2(m)}},$$

where $n_1(r), n_2(r)$ are factors coming from the first and second denominator in (B.3). Moreover, q_m is not identically zero because this would imply that $\mathcal{B} \subset V(p_m)$ which by (ii) would imply that p_m annihilates all orthogonal pairs in $\mathbb{M}_{\mathbf{m}}^2$ which by basic algebraic geometry implies that ort divides p_m, thus contradicting (iii). Hence $T^{-1}(\mathcal{B}_{np}) \subset V(q_1, \dots, q_M)$, and it is easy to see that each proper algebraic variety has zero Lebesgue measure. ▯

PROPOSITION B.7. *Given a generic $f \in \mathbb{M}_{2\mathbf{m}}$, all pairs of con-eigenvectors $u_n,\ u_{n'}$ are proper.*

Proof. Let $\mathcal{A}$ be the set of f's in $\mathbb{M}_{2\mathbf{m}}$ with the property that all con-eigenvectors are irreducible. Let $\{T_l\}_{l=1}^N$ be the atlas for $\mathcal{B}$ defined after (B.3), and for each l, let $\mathcal{A}^l$ be the subset of f's in $\mathcal{A}$ such that one pair of con-eigenvectors lie in $\operatorname{Im} T_l$. Finally, let $\mathcal{A}^l_{np}$ be the subset of f's in $\mathcal{A}^l$ such that at least one pair of con-eigenvectors in $\operatorname{Im} T_l$ is not proper. In order to prove the proposition it suffices to show that $L(\mathcal{A}^l_{np}) = 0$ for a fixed l, by virtue of Lemma B.4 and the fact that the number of charts T_l is finite. Consider u_1 and u_2 as functions on $\mathbb{R}^{4\dim \mathbb{M}_{\mathbf{m}}-4}$ defined by $\big(u_1(r), u_2(r)\big) = T_l(r)$. For $r \in T_l^{-1}(\mathcal{B}_{ir})$ we have, (using the same arguments as in Proposition 5.1), that

$$\dim \operatorname{Ker} \begin{pmatrix} D_{u_1(r)} \\ D_{u_2(r)} \end{pmatrix}^* = 1$$

which implies that

$$\dim \operatorname{Ran} \begin{pmatrix} D_{u_1(r)} \\ D_{u_2(r)} \end{pmatrix} = 2\dim \mathbb{M}_{\mathbf{m}} - 1.$$

Set $d_{ran} = 2\dim \mathbb{M}_{\mathbf{m}} - 1$ and $d_{ker} = \dim \mathbb{M}_{2\mathbf{m}} - d_{ran}$.

Let $r_0 \in \mathbb{R}^{4\dim \mathbb{M}_{\mathbf{m}}-4}$ such that $T_l(r_0) \in \mathcal{B}_{ir}$ be fixed but arbitrary. Choose a basis $e_1, \dots, e_{d_{ker}}$ for $\operatorname{Ker} \begin{pmatrix} D_{u_1(r_0)} \\ D_{u_2(r_0)} \end{pmatrix}$ and another basis $e'_1, \dots, e'_{d_{ran}}$ for $\operatorname{Ran} \begin{pmatrix} D_{u_1(r_0)} \\ D_{u_2(r_0)} \end{pmatrix}$. Let $\Omega_0 \subset \mathbb{R}^{4\dim \mathbb{M}_{\mathbf{m}}-4}$ be a neighborhood of r_0 such that for all $r \in \Omega_0$ the following conditions hold:

$$\begin{aligned} &\{e_1, \dots, e_{d_{ker}}\}^\perp \cap \operatorname{Ker} \begin{pmatrix} D_{u_1(r)} \\ D_{u_2(r)} \end{pmatrix} = \{0\} \\ &\{e'_1, \dots, e'_{d_{ran}}\}^\perp \cap \operatorname{Ran} \begin{pmatrix} D_{u_1(r)} \\ D_{u_2(r)} \end{pmatrix} = \{0\} \\ &T_l(r) \in \mathcal{B}_{ir}. \end{aligned}$$

We omit the argument which proves that such a neighborhood always can be found. Given $r \in \Omega_0$, $b \in \mathbb{C}^{d_{ker}}$ and $(\sigma_1, \sigma_2) \in \mathbb{R}^+ \times \mathbb{R}^+$ consider the following equation-system with $f \in \mathbb{M}_{2\mathbf{m}}$ as unknown:

$$\text{(B.4)} \qquad \begin{cases} \left\langle \begin{pmatrix} \sigma_1 \overline{u_1(r)} \\ \sigma_2 \overline{u_2(r)} \end{pmatrix}, e'_k \right\rangle &= \left\langle \begin{pmatrix} D_{u_1(r)} \\ D_{u_2(r)} \end{pmatrix} f, e'_k \right\rangle, \qquad k = 1, \dots, d_{ran} \\ \qquad\qquad b(k) &= \langle f, e_k \rangle, \qquad\qquad\qquad\quad k = 1, \dots, d_{ker} \end{cases}$$

The equation-system consists of $d_{ran} + d_{ker} = \dim \mathbb{M}_{2\mathbf{m}}$ equations. If we order them then the left hand side defines a vector in $\mathbb{C}^{\dim \mathbb{M}_{2\mathbf{m}}}$ in a natural way, which we will consider as a function of $(r, b, \sigma_1, \sigma_2) \in \Omega_0 \times \mathbb{C}^{d_{ker}} \times \mathbb{R}^+ \times \mathbb{R}^+$ and denote by α_0. If we let $U : \mathbb{M}_{2\mathbf{m}} \to \mathbb{C}^{\dim \mathbb{M}_{2\mathbf{m}}}$ denote any fixed unitary operator, then the equation-system (B.4) can be written as

$$\alpha_0 = \Lambda_0 U f,$$

where Λ_0 is an $\dim \mathbb{M}_{\mathbf{2m}} \times \dim \mathbb{M}_{\mathbf{2m}}$-matrix that we consider as a function of $r \in \Omega_0$. Note that $\Lambda_0(a)$ always is invertible, by the choice of Ω_0. Define the C^∞-function $\beta_0 : \Omega_0 \times \mathbb{C}^{d_{ker}} \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{M}_{\mathbf{2m}}$ via

$$\beta_0(r, b, \sigma_1, \sigma_2) = U^{-1}\big(\Lambda_0(r)\big)^{-1}\alpha_0(r, b, \sigma_1, \sigma_2).$$

Recall that the above construction was made for an arbitrary point $r_0 \in \mathbb{R}^{4\dim \mathbb{M}_{\mathbf{m}}-4}$ such that $T_l(r_0) \in \mathcal{B}_{ir}$. Given another point r_1 with the same properties we will associate to it the function β_1 and the set Ω_1 defined analogously. A short argument which we omit shows that $\mathcal{B}_{ir}$ is an open set. Therefore, by a simple compactness argument, we may choose a sequence of points $r_0, r_1, \ldots \in \mathbb{R}^{4\dim \mathbb{M}_{\mathbf{m}}-4}$ such that

$$\cup_{j=0}^{\infty}\Omega_j \supset T_l^{-1}(\mathcal{B}_{ir}). \tag{B.5}$$

Now, given any $f \in \mathcal{A}^l$ then by definition there exists a pair of con-eigenvectors $(v_1, v_2) \in \operatorname{Im} T_l \cap \mathcal{B}_{ir}$. Let (σ_1, σ_2) be the corresponding con-eigenvalues. By (B.5) we may pick a j such that there exists an $r \in \Omega_j$ such that $v_1 = u_1(r)$ and $v_2 = u_2(r)$. It is not hard to see that

$$f = \beta_j\big(r, b, \sigma_1, \sigma_2\big)$$

for some choice of $b \in \mathbb{C}^{d_{ker}}$. Hence

$$\mathcal{A}^l_{np} \subset \cup_{j=1}^{\infty}\beta_j\Big(\big(\Omega_j \cap T_l^{-1}(\mathcal{B}_{np})\big) \times \mathbb{C}^{d_{ker}} \times \mathbb{R}^+ \times \mathbb{R}^+\Big)$$

To conclude the proof we thus have to show that

$$L\Big(\beta_j\Big(\big(\Omega_j \cap T_l^{-1}(\mathcal{B}_{np})\big) \times \mathbb{C}^{d_{ker}} \times \mathbb{R}^+ \times \mathbb{R}^+\Big)\Big) = 0 \tag{B.6}$$

for each fixed value of l and j. By Lemma B.6 it follows that

$$L\Big(\big(\Omega_j \cap T_l^{-1}(\mathcal{B}_{np})\big) \times \mathbb{C}^{d_{ker}} \times \mathbb{R}^+ \times \mathbb{R}^+\Big) = 0 \tag{B.7}$$

and moreover,

$$\big(\Omega_j \cap T_l^{-1}(\mathcal{B}_{np})\big) \times \mathbb{C}^{d_{ker}} \times \mathbb{R}^+ \times \mathbb{R}^+ \subset \mathbb{R}^{4\dim \mathbb{M}_{\mathbf{m}}-4} \times \mathbb{C}^{d_{ker}} \times \mathbb{R} \times \mathbb{R}$$

which has real dimension

$$4\dim \mathbb{M}_{\mathbf{m}} - 4 + 2d_{ker} + 2 = 4\dim \mathbb{M}_{\mathbf{m}} - 4 + 2(\dim \mathbb{M}_{\mathbf{2m}} - (2\dim \mathbb{M}_{\mathbf{m}} - 1)) + 2 = 2\dim \mathbb{M}_{\mathbf{2m}}.$$

Hence β_j is a differentiable function between spaces of equal dimension. The identity (B.6) now follows by (B.7) and the fact that such a function maps sets of Lebesgue measure zero to sets of Lebesgue measure zero. □

It remains to show that Theorem 6.2 holds generically. A proof can be constructed using the above ideas, but we will omit this.

REFERENCES

[1] V.M. Adamjan, D.Z. Arov, M.G. Kreĭn, Infinite Hankel matrices and generalized Carathéodory-Fejér and I. Schur problems, Funkcional. Anal. i Priložen, 2 (4) (1968), pp. 1-17.

[2] V.M. Adamjan, D.Z. Arov, M.G. Kreĭn, Infinite Hankel matrices and generalized problems of Carathéodory-Fejér and F. Riesz, Funkcional. Anal. i Priložen, 2 (1) (1968), pp. 1-19.

[3] V.M. Adamjan, D.Z. Arov, M.G. Kreĭn, Analytic properties of the Schmidt pairs of a Hankel operator and the generalized Schur-Takagi problem, Mat. Sb. (N.S.), 86 (128) (1971), pp. 34-75.

[4] F. Andersson, M.V. de Hoop, H.F. Smith, G. Uhlmann, Multi-scale approach to hyperbolic evolution equations with limited smoothness, Comm. Partial Differential Equations (2008), in print.

[5] D.N. Bernstein, The number of roots of a system of equations, Funkcional. Anal. i Priložen, 9 (1975), pp. 1-4.

[6] G. Beylkin, On the fast Fourier transform of functions with singularities, Applied and Computational Harmonic Analysis, 2 (1995), pp. 363-381.

[7] G. Beylkin, L. Monzón, On generalized Gaussian quadratures for exponentials and their Applications, Applied and Computational Harmonic Analysis, 12 (2002), pp. 332-373.

[8] G. Beylkin, L. Monzón, On approximation of functions by exponential sums, Applied and Computational Harmonic Analysis, 19 (2005), pp. 17-48.

[9] L. Boutet de Monvel, Hypoelliptic operators with double characteristics and related pseudodifferential operators, Comm. Pure Appl. Math., 27 (1974), pp. 585-639.

[10] J. Bros, D. Iagolnitzer, Support essentiel et structure analytique des distributions, Séminaire Goulaouic-Lions-Schwartz, exp. no. 19 (1975-1976).

[11] E.J. Candès, D.L. Donoho, New tight frame of curvelets and optimal representations of objects with piecewise-C^2 singularities, Comm. Pure Appl. Math., 57 (2002), pp. 219-266.

[12] J.F. Claerbout, Imaging the Earth's Interior. Blackwell Scientific Publications, Oxford, 1985.

[13] A. Cohen, Applied and computational aspects of nonlinear wavelet approximation, In: Multivariate Approximation and Applications, N. Dyn, D. Leviatan, D. Levin, A. Pinkus (eds.), Cambridge University Press, Cambrige, 2001, pp. 188-212.

[14] A. Córdoba, C. Fefferman, Wave packets and Fourier integral operators, Comm. Partial Differential Equations, 3-11 (1978), pp. 979-1005.

[15] D. Cox, J. Little, D. O'Shea, Ideals, Varieties, and Algorithms. Springer-Verlag, New York, 1997.

[16] I. Daubechies, M. Defrise, C. de Mol, An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Comm. Pure Appl. Math., 57 (2004), pp. 1413-1541.

[17] M.V. de Hoop, H.F. Smith, G. Uhlmann, R.D. van der Hilst, Seismic imaging with the generalized Radon transform and double beamforming: A curvelet transform perspective, submitted (2008).

[18] H. Douma, M.V. de Hoop, Leading order seismic imaging using curvelets, Geophysics, 72 (2007), pp. S231-S248.

[19] A. Dutt, V. Rokhlin, Fast Fourier transforms for nonequispaced data, SIAM Journal on Scientific Computing, 14 (1993), pp. 1368-1393.

[20] A. Greenleaf, G. Uhlmann, Estimates for singular Radon transforms and pseudodifferential operators with singular symbols, J. Funct. Anal., 89 (1990), pp. 202-232.

[21] G. Hennenfent, F.J. Herrmann, Seismic denoising with non-uniformly sampled curvelets, Comp. in Sci. and Eng., 8 (2006), pp. 16-25.

[22] R.A. Horn, C.R. Johnson, Matrix Analysis. Cambridge University Press, Cambridge, 1990.

[23] F.T. Luk, S. Qiao, A fast singular value algorithm for Hankel matrices, In: Fast Algorithms for Structured Matrices: Theory and Applications, Contemporary Mathematics, 323, V. Olshevsky (ed.), American Mathematical Society, 2003, pp. 169-17.

[24] S. Mallat, A Wavelet Tour of Signal Processing. Academic Press, San Diego, 1998.

[25] S. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries, IEEE Transactions on Signal Processing, 41 (12) (1993), pp. 3397-3415.

[26] V.F. Pisarenko, The retrieval of harmonics from a covariance function, Geophys. J. R. Astr. Soc., 33 (1973), pp. 347-366.

[27] H.F. Smith, A Hardy space for Fourier integral operators, Jour. Geom. Anal., 8 (1998), pp. 629-653.

[28] H.F. Smith, A parametrix construction for wave equations with $C^{1,1}$ coefficients, Ann. Inst. Fourier, Grenoble, 48 (1998), pp. 797-835.

[29] E.M. Stein, Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press, Princeton, 1993.

[30] C.C. Stolk, M.V. de Hoop, Seismic inverse scattering in the downward continuation approach, Wave Motion, 43 (2006), pp. 579-598.