

ESPRIT for multidimensional general grids.

Fredrik Andersson^a, Marcus Carlsson^b

^a*Institute of Geophysics, ETH-Zurich, Sonneggstrasse 5, 8092 Zurich, Switzerland*

^b*Centre for Mathematical Sciences, Lund University, Box 118, SE-22100, Lund, Sweden*

Abstract

We present a method for complex frequency estimation in several variables, extending the classical one-dimensional ESPRIT-algorithm, and consider how to work with data sampled on non-standard domains, i.e going beyond multi-rectangles.

Keywords: complex frequency estimation, Hankel, finite rank, Kronecker theorem, sums of exponentials

2010 MSC: 15B05, 41A63, 42A10.

1. Introduction

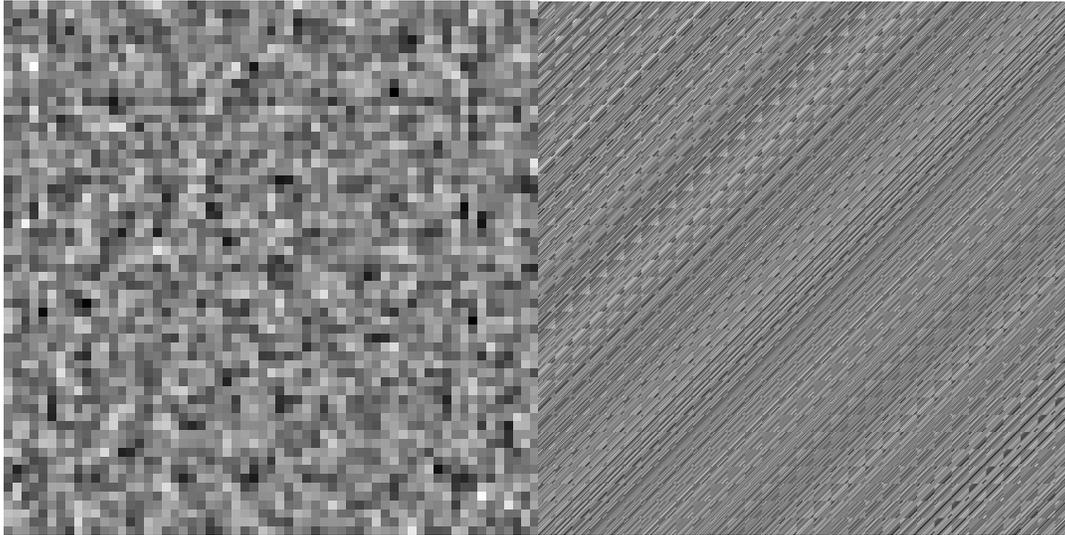
Given equidistant measurements of a signal f and a model order K we seek to approximate f by an exponential sum

$$f(x) \approx \sum_{k=1}^K c_k e^{\zeta_k x}, \quad c_k, \zeta_k \in \mathbb{C}. \quad (1.1)$$

The complex parameters ζ_k play the role of frequencies in the case when they lie on the imaginary axis, and the problem is therefore often referred to as *frequency estimation*. Once the parameters ζ_k are found, the subsequent retrieval of the coefficients c_k is a linear problem which can be solved by the least squares method. In this paper, we are interested in multivariate version of this problem.

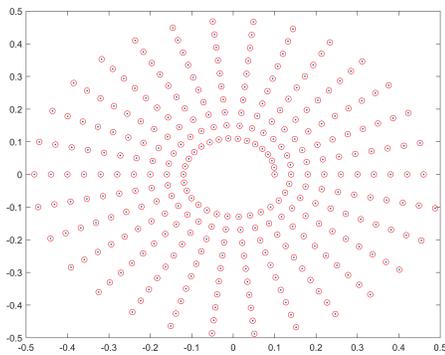
The frequency estimation problem goes far back in time. Already in 1795, it was discussed by Prony [12]. In the noise free case, i.e. when f already is of the form given by the right hand side of (1.1), it is possible to exactly recover both the complex frequencies and the corresponding coefficients. The approximation problem is of great importance, and several different approaches and methods have been devoted to solving it. Classical methods include ESPRIT [27], MUSIC [31], Pisarenkos method [22], and the matrix pencil method, see e.g. [16] which also provides a good account on the history of the method and precursors, as well as the recent contribution [25] which finds a unifying framework for these methods. From the perspective of functional analysis, a lot of theoretic results were obtained by Adamjan, Arov and Krein (AAK) [1]. These connections were further exploited in [7] for the construction of optimal quadrature nodes. The connection between approximation theory and the results by Adamjan, Arov and Krein is discussed in more detail in [6].

Email addresses: fredrik.andersson@erdw.ethz.ch (Fredrik Andersson), marcus.carlsson@math.lu.se (Marcus Carlsson)

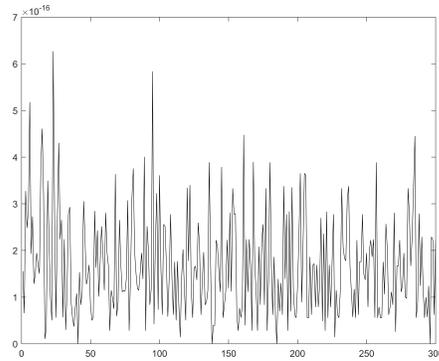


(a) Real part of f .

(b) The corresponding block-Hankel matrix



(c) True 2d frequencies in blue dots and estimated frequencies in red circles



(d) Error between estimated and true frequencies

Figure 1: Two dimensional example where a function f of two variables is a linear combination of 300 exponential functions, which is sampled on a 61×61 -lattice (i.e. $N = 31$).

There have been several suggestions on how to extend the one-dimensional results to several variables, c.f. [5, 10, 11, 13, 14, 17, 18, 21, 26, 28, 30, 35, 36]. See [15] for an account of earlier attempts. The multidimensional case poses several difficulties; Let us assume that f is of the form

$$f(x) = \sum_{k=1}^K c_k e^{\zeta_k \cdot x}, \quad c_k \in \mathbb{C}, \quad \zeta_k \in \mathbb{C}^d, \quad (1.2)$$

(possibly distorted by noise) and that we wish to retrieve the frequencies $\zeta_k = (\zeta_{k,1}, \dots, \zeta_{k,d})$,

where $\zeta_k \cdot x = \sum_{j=1}^d \zeta_{k,j} x_j$. We suppose for the moment that f has been sampled on an integer multi-cube $\{1, \dots, 2N - 1\}^d$. A common engineering approach is to use some reformulation of the two dimensional problem to one-dimensional problems and from there recover the two dimensional components, see e.g. the recent articles [34, 24]. This causes a pairing problem that can lead to difficulties in practice, depending on the particular application at hand. An advantage with these type of methods is that they can be fast since they only sample data along lines.

To be more precise, one may average over all dimensions but one, say the p :th one, (or extract a “fiber” in dimension p from the data) and then use a 1-d technique for estimating $\zeta_{1,p}, \dots, \zeta_{K,p}$, and repeat this for the remaining dimensions. However, first of all this limits the amount of frequencies we are theoretically able to retrieve to $N - 1$ (in the noise free case, using e.g. ESPRIT, MUSIC or Prony’s method), and secondly it is not clear how to pair the Kd frequencies $\zeta_{k,p}$ into K multi-frequencies $\zeta_k = (\zeta_{k,1}, \dots, \zeta_{k,d})$. Note that there are K^d possible combinations, whereas only K are sought. Also, these methods do not use the full data set and hence as an estimator they are likely to be more sensitive to noise, than a method which employs the full data set in the estimation. A related algorithm, which brings the idea of only using samples along certain lines to its extreme, is considered in [23] where it is shown that one can (theoretically) get away with $3K - 1$ samples in total, in order to retrieve K multi-frequencies (for $d = 2$).

The focus in this article is different, we suppose that f has been measured on a cube or more generally some irregular domain in \mathbb{R}^d , and propose a method which uses the full data set to retrieve the multi-frequencies and which avoids the pairing problem. This method is an extension of the ESPRIT-method. It was first introduced in [26] in the 2d square case, and has recently been extended to several variables independently by [29] and [33]. These articles also contain analysis of noise sensitivity and the former proposes accelerated versions. We review this method in its basic form in Sections 2 (one variable) and 3 (several variables).

The main contribution of the present work is to extend this method to work with data sampled on non-standard domains. This problem is also discussed in the 2d-setting in [13, 32]. We also show how to realize block Hankel operators as multivariable summing operators, thereby providing a connection with theoretical results connecting the rank of these operators with functions of the form (1.2), provided in [4] and [2].

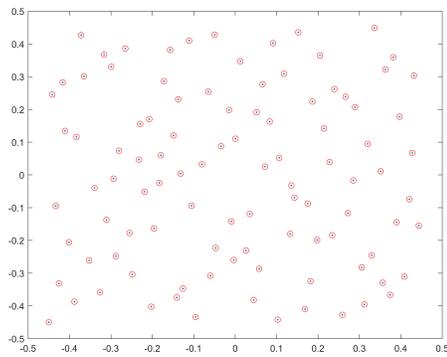
We now discuss a few examples to illustrate the pros and cons of the method suggested here. Suppose first that f is sampled on a 19^3 data-tensor, i.e. $N = 10$ and $d = 3$. As it turns out, we are able to estimate $(N - 1) \cdot N^{d-1}$ distinct multi-frequencies, avoiding degenerate cases. With the method considered here we may then retrieve 900 (randomly distributed) complex frequencies (assuming no noise), as opposed to 9 using reduction to one-dimensional techniques. By “degenerate cases” we for example rule out situations where many datapoints align parallel with one of the coordinate axes. For applications where this is the case the method considered here may not work. This is further discussed in Section 3, without making any rigorous attempt at formalizing what is meant by degenerate cases.

We now consider the example shown in Figure 1, still in the multi-cube setting. The underlying function f is in this case constructed as a linear combination of 300 purely oscillatory exponential functions with frequencies distributed along a spiral as depicted in panel c). The coefficients are chosen randomly, and f is sampled on a 61 by 61 square grid (so $N = 31$ and $d = 2$). The real part of f is shown in panel a). The block-Hankel matrix that is generated from f is shown in panel b). We have followed Algorithm 2 to recover the underlying frequencies and they are recovered at machine precision error as shown in panel d).

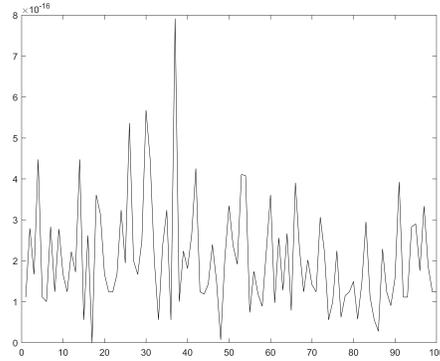
Using reduction to one-dimensional techniques as discussed earlier, the maximum amount of



(a) Real part of a 2d general domain data set constructed as a linear combination of 100 exponential functions.



(b) True 2d frequencies in blue dots and estimated frequencies in red circles.



(c) Error between estimated and true frequencies.

Figure 2: Two dimensional general domain example.

frequencies that could have been retrieved from this data set is 30. With the present method it is $30 \cdot 31 = 930$ (see (3.3)) compared to the total amount of measurements $61^2 = 3721$. Or viewed from a different angle, to recover the 300 frequencies with reduction to one-dimensional techniques, we would theoretically need to have at least 601 measurements in each direction, and then it is not an easy problem to stably recover the 300 components in each dimension. Even if this is overcome, one then has $300 \cdot 300 = 90000$ possible multi-frequencies to pick from, and 89700 has to be discarded, which may take some time. We refer again to [34, 24] for two recent approaches on how to do this, while underlining that these methods may well be superior (and faster) in situations when fewer frequencies are sought and dense sampling along certain lines is not an obstacle in practice.

We now turn our focus to an example using the general domain Hankel structure, introduced in Section 5, which allows us to work with samples of f on domains with different shapes. In this case, we have used 100 exponential functions to generate the function f . This function is then sampled on a semi circular-like domain Ω shown in Figure 2 a). Figure 2 b) shows the distribution of the underlying frequencies, and Figure 2 c) shows the reconstruction error in determining these frequencies. Again, the error is down at machine precision.

More examples are given in the numerical section, where we also test how the method performs with noise present, but we wish to underline that the main contribution of this paper is the deduction of an algorithm that is capable of exact retrieval in the noise-free case on general domains. Its performance as an estimator will be investigated elsewhere.

The paper is organized as follows, in Section 2 we go over the essentials of the classical ESPRIT algorithm in one variable, in Section 3 we extend this to the multi-cube case, using block Hankel matrices, and show how to solve the pairing problem. We then consider how to make this work for data sampled on non-cubical domains. This relies on so called *general domain Hankel matrices*, which are introduced in Section 4 and 5. The subsequent extension of the ESPRIT-type algorithm from Section 3 is given in Section 6. Numerical examples are given in Section 7.

2. One-dimensional ESPRIT

This section is intended as a quick review of ESPRIT in the simplest possible setting, to provide intuition for the more advanced versions introduced in later sections. For example we only work with square Hankel matrices (which we have found works best in practice), although non-square Hankel matrices will be considered in Section 5 and 6 in the multivariable case. The study of ESPRIT and related algorithms in 1 dimension is of course a field in itself, and we refer e.g. to [25] for a recent contribution with more flexibility.

Let us start by assuming that f is a given one-dimensional function which is sampled at integer points, and that

$$f(x) = \sum_{k=1}^K c_k e^{\zeta_k x},$$

for distinct values of ζ_k , $k = 1, \dots, K$. A *Hankel* matrix is a matrix H which has constant values on its anti-diagonals. It is easy to see that each Hankel matrix can be generated from a function h by setting $H(m, n) = h(m + n)$. Let us now consider the case where a (square $N \times N$) Hankel matrix H is generated by the function f above, where $N > K$.

Let Λ denote the Vandermonde matrix generated by the numbers e^{ζ_k} , i.e., let

$$\Lambda(j, k) = e^{\zeta_k j}, \quad 1 \leq j \leq N, \quad 1 \leq k \leq K$$

and let Λ_k denote the columns of Λ . For the elements of the Hankel matrix H it thus holds that

$$H(m, n) = f(m + n) = \sum_{k=1}^K c_k e^{\zeta_k(m+n)} = \sum_{k=1}^K c_k e^{\zeta_k m} e^{\zeta_k n},$$

implying that H can be written as

$$H = \sum_{k=1}^K c_k \Lambda_k \Lambda_k^T = \Lambda \text{diag}(c) \Lambda^T,$$

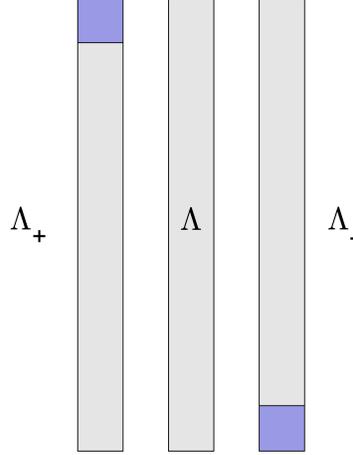


Figure 3: Illustration of the notation Λ_+ and Λ_- . For Λ_+ the first row is deleted, and for Λ_- the last row is deleted. The elements marked in blue are to be deleted.

where $c = (c_1, \dots, c_K)$. From this observation it is clear that the rank of H is K .

Now, by the singular value decomposition it follows that we can write $H = U\Sigma V^T$ where U and V are $N \times K$ matrices and Σ a $K \times K$ diagonal matrix containing the non-zero singular values of H . It then holds that the corresponding singular vectors U (or V) are linear combinations of the columns of Λ , i.e.,

$$U = \Lambda B,$$

where B is an invertible $K \times K$ matrix.

Let $D = \text{diag}(e^{\zeta_1}, \dots, e^{\zeta_K})$. Let us also introduce the notation Λ_+ for the matrix that is obtained by deleting the first row from the matrix Λ , and similarly, let Λ_- denote the matrix that is obtained by deleting the last row from Λ . There is a simple relationship between Λ_+ and Λ_- , namely

$$\Lambda_+ = \Lambda_- D.$$

Moreover, it holds that

$$\begin{aligned} U_+ &= \Lambda_+ B = \Lambda_- D B \\ U_- &= \Lambda_- B \end{aligned}$$

Since $U^* U = I$ and $N > K$, it typically holds that U_- is injective. In particular it then has a left inverse U_-^\dagger given by

$$U_-^\dagger = (U_-^* U_-)^{-1} U_-^* = (B^* \Lambda_-^* \Lambda_- B)^{-1} B^* \Lambda_-^* = B^{-1} (\Lambda_-^* \Lambda_-)^{-1} \Lambda_-^*.$$

Let us now consider

$$A = U_-^\dagger U_+ = B^{-1} (\Lambda_-^* \Lambda_-)^{-1} \Lambda_-^* \Lambda_- D B = B^{-1} D B.$$

It follows that the columns of B^{-1} are eigenvectors of the matrix A with e^{ζ_k} being the corresponding eigenvalues. This implies that the exponentials e^{ζ_k} can be recovered from H by diagonalizing the matrix A , and this is the essence of the famous ESPRIT-method.

Algorithm 1 One-dimensional ESPRIT

- 1: Form Hankel matrix $H(m, n) = f(m + n)$ from samples of f .
 - 2: Compute singular value decomposition $H = U\Sigma V^T$, retaining only non-zero singular values and corresponding vectors.
 - 3: Form U_+ and U_- by deleting the first and last rows from U , respectively.
 - 4: Form $A = (U_-^* U_+)^{-1} U_-^* U_+$.
 - 5: Diagonalize $A = B^{-1} \text{diag}(\lambda_1, \dots, \lambda_K) B$ by making an eigenvalue decomposition of A .
 - 6: Recover $\zeta_k = \log(\lambda_k)$
-

We end by making a remark about the connection between Hankel matrices with low-rank and functions that are sums of exponential functions. For every function f being a sum of K (distinct) exponential functions the rank of the corresponding Hankel matrix is K (given sufficiently many samples). The reverse state will typically be true, but there are exceptions. These exceptions are of degenerate character, and therefore this problem is often discarded in practice. For a longer discussion of these issues, see Section 2 and 11 in [4].

3. Multi-dimensional ESPRIT for block-Hankel matrices

We now consider the case where f is a given d -dimensional function which is sampled at integer points, and that

$$f(x) = \sum_{k=1}^K c_k e^{\zeta_k \cdot x},$$

where $\zeta_k = (\zeta_{k,1}, \dots, \zeta_{k,d})$ and $x = (x_1, \dots, x_d)$. A d -dimensional Hankel operator can be viewed as a linear operator on the tensor product $\mathbb{C}^N \otimes \dots \otimes \mathbb{C}^N$ whose entry corresponding to multi-indices $m = (m_1, \dots, m_d)$ and $n = (n_1, \dots, n_d)$ equals $f(m + n)$. This in particular entails that samples of f needs to be available on a $(2N - 1)^d$ -dimensional grid. The restriction that we use the same amount of sample points in each direction is made for simplicity, the rectangular case is a special case of the more general setup in Section 6.

If we identify $\mathbb{C}^N \otimes \dots \otimes \mathbb{C}^N$ with \mathbb{C}^{N^d} using the reverse lexicographical order, i.e. by identifying entry m in the former with

$$m_1 + m_2 N + m_3 N^2 + \dots + m_d N^{d-1} \quad (0 \leq m_j < N) \quad (3.1)$$

in the latter, the corresponding operator can be realized as a block-Hankel matrix, which we denote by \mathbf{H} . Analogously, given multi-index m we will write \mathbf{m} for the number (3.1), and if u is an element in $\mathbb{C}^N \otimes \dots \otimes \mathbb{C}^N$ we write \mathbf{u} for its vectorized version in \mathbb{C}^{N^d} .

Just as for the one-dimensional case it can be shown that if Λ is the Vandermonde matrix generated by the vectors ζ_k , i.e.,

$$\Lambda(\mathbf{m}, k) = e^{\zeta_k \cdot \mathbf{m}},$$

then

$$\mathbf{H} = \sum_{k=1}^K c_k \mathbf{\Lambda}_k \mathbf{\Lambda}_k^T = \mathbf{\Lambda} \text{diag}(c) \mathbf{\Lambda}^T,$$

and in particular \mathbf{H} and $\mathbf{\Lambda}$ typically have rank K . If, due to insufficient sampling or a particular alignment of the ζ_k 's, this affirmation is false, then the method described below will not apply, at least without further refinements. For example, maximal rank (i.e. K) can be achieved by rotation of the grid and/or by increasing N , which is further discussed in Section 6. Other tricks to deal with this issue is found in Section III of [26]. Below we assume that problems related to rank deficiency or higher multiplicity of eigenvalues are absent, and discuss how to treat such issues towards the end of the section.

We thus assume that \mathbf{H} and $\mathbf{\Lambda}$ has rank K and write $\mathbf{H} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$ for the singular value decomposition where we omit singular values that are zero and corresponding singular vectors, so in particular $\mathbf{\Sigma}$ is a diagonal $K \times K$ -matrix. As in the one-dimensional case it then holds that the singular vectors in \mathbf{U} (or \mathbf{V}) are linear combinations of the columns of $\mathbf{\Lambda}$, i.e.,

$$\mathbf{U} = \mathbf{\Lambda} \mathbf{B}, \tag{3.2}$$

where \mathbf{B} is an invertible $K \times K$ matrix like before.

We now generalize the previous operation of deleting the first and last row respectively to several dimensions. By $\mathbf{\Lambda}_{p+}$ we denote the matrix that is obtained by deleting all entries of $\mathbf{\Lambda}$ with indices \mathbf{m} corresponding to multi-indices m of the form $(m_1, \dots, 1, \dots, m_d)$, i.e., all the first elements with respect to dimension p . Similarly, we denote by $\mathbf{\Lambda}_{p-}$ the matrix that is obtained by deleting from $\mathbf{\Lambda}$ all the elements related to multi-indices $(m_1, \dots, N_p, \dots, m_d)$, i.e., all the last elements with respect to dimension p . Let

$$D_p = \text{diag}(e^{\zeta^{1,p}}, \dots, e^{\zeta^{K,p}})$$

and note that it in each dimension p we have

$$\mathbf{\Lambda}_{p+} = \mathbf{\Lambda}_{p-} D_p.$$

Moreover, it holds that

$$\begin{aligned} \mathbf{U}_{p+} &= \mathbf{\Lambda}_{p+} \mathbf{B} = \mathbf{\Lambda}_{p-} D_p \mathbf{B} \\ \mathbf{U}_{p-} &= \mathbf{\Lambda}_{p-} \mathbf{B} \end{aligned}$$

Let us now again consider

$$\begin{aligned} A_p &= \mathbf{U}_{p-}^\dagger \mathbf{U}_{p+} = (\mathbf{U}_{p-}^* \mathbf{U}_{p-})^{-1} \mathbf{U}_{p-}^* \mathbf{U}_{p+} \\ &= (\mathbf{B}^* \mathbf{\Lambda}_{p-}^* \mathbf{\Lambda}_{p-} \mathbf{B})^{-1} \mathbf{B}^* \mathbf{\Lambda}_{p-}^* \mathbf{\Lambda}_{p-} D_p \mathbf{B} \\ &= \mathbf{B}^{-1} (\mathbf{\Lambda}_{p-}^* \mathbf{\Lambda}_{p-})^{-1} (\mathbf{B}^*)^{-1} \mathbf{B}^* \mathbf{\Lambda}_{p-}^* \mathbf{\Lambda}_{p-} D_p \mathbf{B} \\ &= \mathbf{B}^{-1} D_p \mathbf{B}. \end{aligned}$$

The remarkable observation here is that \mathbf{B}^{-1} *simultaneously* diagonalizes all the matrices A_p , and that the eigenvalues to A_p are $(e^{\zeta^{1,p}}, \dots, e^{\zeta^{K,p}})$ in the correct order independent of p . This implies that all the exponentials $e^{\zeta^{k,1}}$, $k = 1, \dots, K$, can be recovered from H_f by diagonalization of A_1 , and then we may use *the same* eigenvectors for diagonalization of the remaining A_p 's, and in this

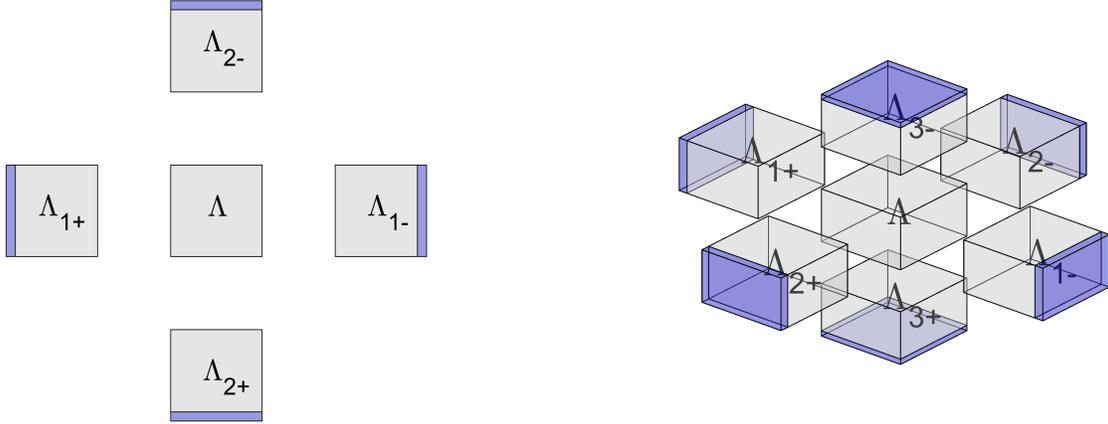


Figure 4: Illustration of the notation Λ_{p+} and Λ_{p-} . For Λ_{p+} the first elements in dimension p are deleted, while for Λ_{p-} the last elements in dimension p are deleted. The elements marked in blue are to be deleted. The left panel shows a two-dimensional case, and in the right panel a three-dimensional case is shown.

way obtain the multi-frequencies $\zeta_k = (\zeta_{k,1}, \dots, \zeta_{k,d})$ for $k = 1, \dots, K$ directly without any need for a pairing procedure. This phenomenon was first observed in [26] in the two-dimensional case, although it is not very clear to see. A more accessible account, which extends [26] to the higher dimensional case, is given in [29] and independently in [33]. To summarize, the algorithm in its most simple form reads as follows:

Algorithm 2 d -dimensional ESPRIT

- 1: Form the d -block Hankel matrix $\mathbf{H}(\mathbf{m}, \mathbf{n}) = f(m+n)$ from samples of f .
 - 2: Compute the singular value decomposition $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^T$.
 - 3: **for** $p = 1, \dots, d$ **do**
 - 4: Form \mathbf{U}_{p+} and \mathbf{U}_{p-} by deleting first and last elements in dimension p , respectively.
 - 5: Form $A_p = (\mathbf{U}_{p-}^* \mathbf{U}_{p-})^{-1} \mathbf{U}_{p-}^* \mathbf{U}_{p+}$.
 - 6: **end for**
 - 7: Diagonalize $A_1 = B^{-1} D_1 B$ by making an eigenvalue decomposition of A_1 .
 - 8: **for** $p = 1, \dots, d$ **do**
 - 9: Compute the diagonal matrices $D_p = \text{diag}(\lambda_{1,p}, \dots, \lambda_{K,p}) = B A_p B^{-1}$.
 - 10: Recover $\zeta_{k,p} = \log(\lambda_{k,p})$ (which are automatically correctly paired).
 - 11: **end for**
-

We now discuss problems that may arise in the above approach. The algorithm works as stated as long as $(\mathbf{\Lambda}_{p-})^* \mathbf{\Lambda}_{p-}$ is invertible for all p , and moreover we need that no eigenvalue in D_1 has multiplicity higher than 1 in order for the eigenvectors in B to be correctly determined. The first limitation leads to the restriction

$$K \leq N^{d-1}(N-1), \quad (3.3)$$

since $\mathbf{\Lambda}_{p-}$ needs to have fewer columns than rows. This restriction was mentioned already in the introduction and also appears in Lemma 2 of [29], where it is noted that this condition holds generically if the frequencies are sampled at random. In the same article, a linear combination step

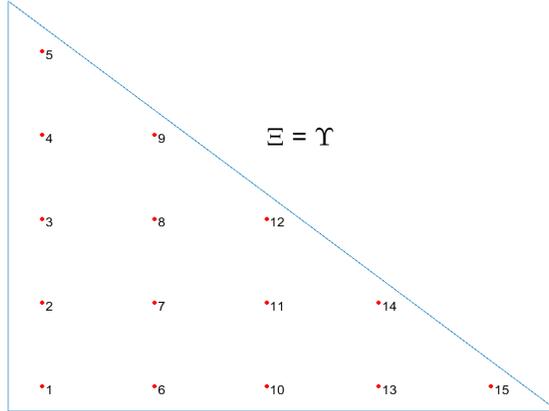


Figure 5: Samples of a triangular domain with corresponding ordering. If the vertical axis is considered to be the first coordinate, then this ordering is the reverse lexicographical ordering.

is also applied to avoid problems with multiplicity. Indeed, suppose for simplicity that $d = 2$ and that D_1 does not have distinct eigenvalues. Then

$$\alpha A_1 + \beta A_2 = B^{-1}(\alpha D_1 + \beta D_2)B \quad (3.4)$$

has distinct eigenvalues for most choices of α and β , so the problem can be circumvented by randomly choosing α, β and compute B from the above linear combination. This trick is also employed in [24], albeit for a different algorithm.

A final remark on Algorithm 2 concerning time-complexity. In [29] it is noted that Algorithm 2 can be slow since performing an SVD on a large matrix is time consuming. It is suggested to circumvent this problem by using “truncated SVD” which computes the K first singular vectors based on a variant of Lanczos algorithm. We note that similar improvements can be applied to the Algorithm 3 presented below, but we do not follow these threads here.

This completes the multidimensional version of ESPRIT in the case when data f is sampled on a regular multi-cube. Next we address the more general setting of data measured on regular grids with various shapes, which is the main contribution of this paper, presented in Algorithm 3. For this we need to introduce so called general domain Hankel matrices, by viewing block Hankel matrices as multi-dimensional summing operators.

4. The summation operator formalism in one variable

Given a subset $\Upsilon \subset \mathbb{N}$, let $\ell^2(\Upsilon)$ be the “sequences” indexed by Υ and equipped with the standard ℓ^2 -norm. If $\Upsilon = \{1, 2, \dots, N\}$, then $\ell^2(\Upsilon)$ reduces to the standard \mathbb{C}^N . Note that we may define a classical Hankel matrix H given by the sequence $f = (f_2, f_3, \dots, f_{2N})$ as the operator with domain and codomain equal to $\ell^2(\{1, \dots, N\})$, given by the summation formula

$$(H_f(a))_n = \sum_{m \in \{1, \dots, N\}} f_{m+n} a_m, \quad n \in \{1, \dots, N\} \quad (4.1)$$

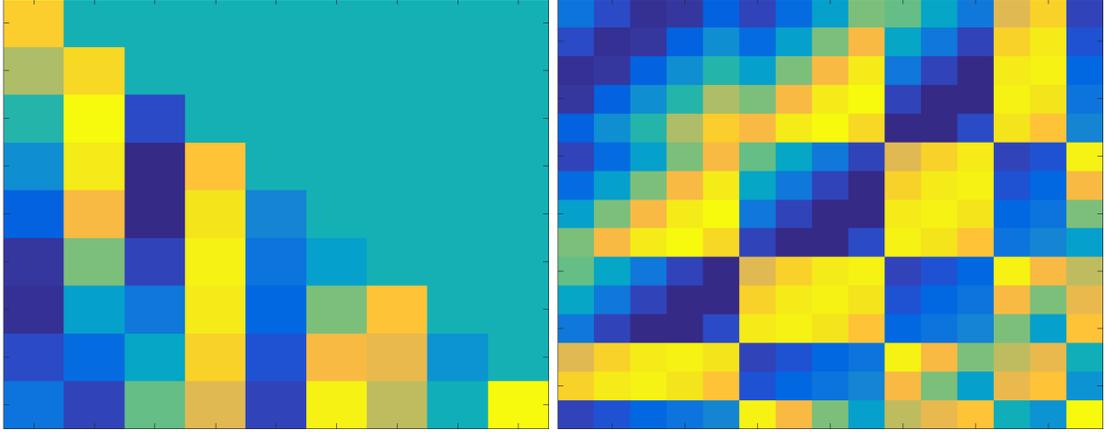


Figure 6: a) Function defined on $\Omega = \Xi + \Upsilon$ with $\Xi = \Upsilon$ as in Figure 5. b) The corresponding general domain Hankel matrix, using the ordering from Figure 5.

where a represents any element in $\ell^2(\{1, \dots, N\})$. This suggests the following generalization; Let $M, N \in \mathbb{N}$ such that $M + N = 2P$ be given, and consider $H_{f,M,N} : \ell^2(\{1, \dots, M\}) \rightarrow \ell^2(\{1, \dots, N\})$ given by

$$(H_{f,M,N}(a))_n = \sum_{m \in \{1, \dots, M\}} f_{m+n} a_m, \quad n \in \{1, \dots, N\}. \quad (4.2)$$

We remark that

$$\{1, \dots, M\} + \{1, \dots, N\} = \{2, \dots, 2P\}, \quad (4.3)$$

where the latter is the grid on which f is sampled. A moments thought reveals that $H_{f,M,N}$ equals the $N \times M$ Hankel matrix given by f . For example, if $f = (1, 2, 3, 4, 5)$, then $P = 3$ and we can pick e.g. $M = 4$ and $N = 2$. We then have

$$H_{f,4,2} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \end{pmatrix}. \quad (4.4)$$

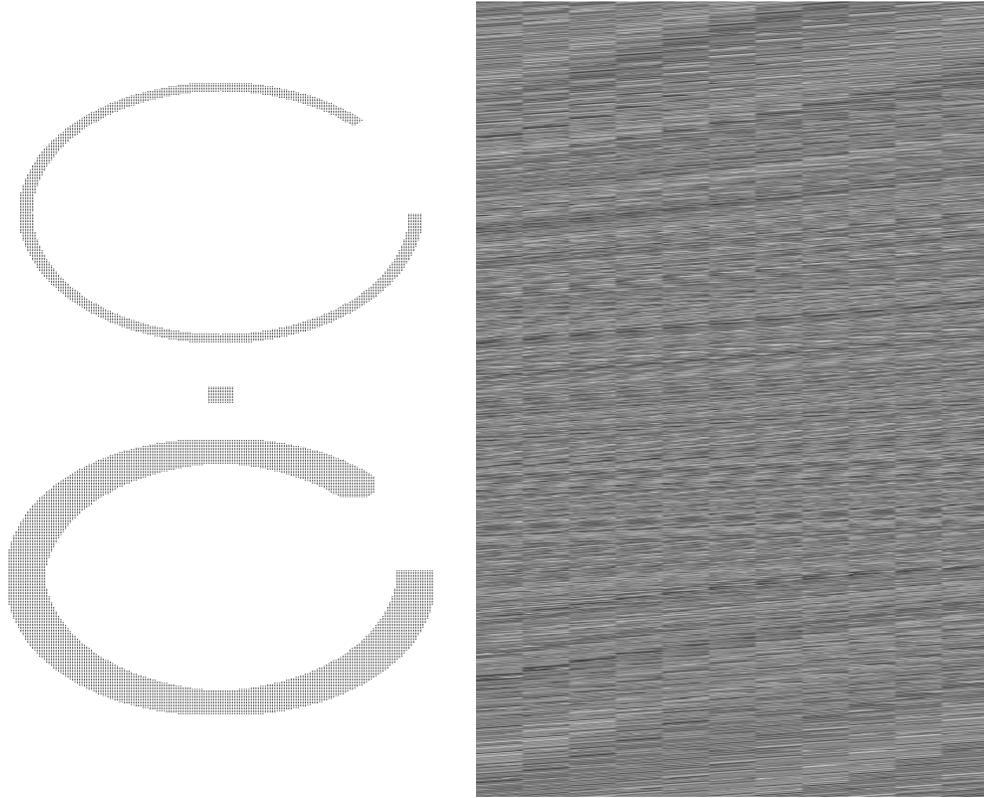
As long as M and N are larger than K , it is possible to perform ESPRIT on these rectangular Hankel matrices as well, although square ones seems to be preferred.

5. General domain Hankel operators

General domain Hankel integral operators were introduced in [2] (albeit under the name truncated correlation operators) and their discretizations were studied in [4] where the term *general domain Hankel matrix* was coined. The discrete versions have been studied earlier in e.g. [20] under the name quasi-Hankel matrices. We briefly revisit their construction, which is explained in greater detail in Section 4 of [4].

Let Ξ, Υ be any bounded subsets of \mathbb{Z}^d , and in analogy with (4.3) set

$$\Upsilon + \Xi = \Omega, \quad (5.1)$$



(a) General domain setup with Υ on top, Ξ in the middle and Ω in the bottom. The Ξ set is a 11×11 grid
(b) The corresponding general domain Hankel matrix

Figure 7: Two dimensional general domain example.

as in Figure 5 and 7 a). Suppose we are interested in a function f on some domain in \mathbb{R}^d , and that Υ and Ξ are chosen such that $\{j_1\Delta x_1, \dots, j_d\Delta x_d\} : j \in \Omega\}$ cover the domain where f is defined, for some choice of sampling length $\Delta x_1, \dots, \Delta x_d$. The function f thus gives rise to a *multidimensional sequence*

$$f_j = f(j_1\Delta x_1, \dots, j_d\Delta x_d), \quad j \in \Omega, \quad (5.2)$$

or more formally a function in $\ell^2(\Omega)$, see Figure 6 a). We will refer to such functions as *md-sequences*, to distinguish them from ordinary sequences, (i.e. vectors in \mathbb{C}^N). In analogy with (4.2), the md-sequence f gives rise to a corresponding general domain Hankel operator

$$(H_{f, \Upsilon, \Xi}(g))_n = \sum_{m \in \Upsilon} f_{m+n} g_m, \quad n \in \Xi, \quad (5.3)$$

where g is an arbitrary md-sequence on Υ . When Υ, Ξ are clear from the context or irrelevant, we drop them from the notation.

We may of course represent g as a vector by ordering the entries (as in Figure 5 for a triangle). More precisely, by picking any bijection

$$o_y : \{1, \dots, |\Upsilon|\} \rightarrow \Upsilon \quad (5.4)$$

(where $|\Upsilon|$ denotes the amount of elements in the set Υ), we can identify g with the vector \mathbf{g} given by

$$(\mathbf{g}_j)_{j=1}^{|\Upsilon|} = g(o_y(j)).$$

Letting o_x be an analogous bijection for Ξ , it is clear that $H_{f,\Upsilon,\Xi}$ can be represented as a $|\Xi| \times |\Upsilon|$ -matrix, where the (n, m) 'th element is $f(o_x(n) + o_y(m))$, see Figure 6 b) and 7 b). Such matrices will be called *general domain Hankel matrices* and denoted $\mathbf{H}_{f,\Upsilon,\Xi}$, letting the bijections o_x and o_y be implicit in the notation (usually we will use the reverse lexicographical order as in Section 3). In particular, if we set

$$\Upsilon = \Xi = \{1, \dots, N\}^d,$$

we retrieve in this way the block Hankel matrices discussed in Section 3. Note that in the one dimensional case, $H_{f,\Upsilon,\Xi}$ and $\mathbf{H}_{f,\Upsilon,\Xi}$ are virtually the same thing, whereas this is not the case in several variables. The former is an operator acting on md-sequences in $\ell^2(\Upsilon)$ to md-sequences in $\ell^2(\Xi)$, the latter is a matrix representation of the former that can be used e.g. for computer implementations.

An example where $\Xi = \Upsilon$ (and hence also Ω) are triangles is shown in Figures 5 and 6. The domains $\Xi = \Upsilon$ are shown in Figures 5 with a particular ordering. Ω is then a triangle of side-length 9, and a function f on Ω is shown in Figure 6 a). The corresponding general domain Hankel matrix, given the ordering from Figure 5, is shown in b). Note especially how columns 3, 4 and 5 are visible and show up as rectangular Hankel matrices of different sizes. Columns 3 and 5 both give rise to a (blue) square Hankel matrix on the “block-structure diagonal”, which also show up as rectangular Hankel matrices off the “block-diagonal”. Column 4 generates the (yellow) Hankel matrices that are never on the diagonal, corresponding to the fact that every second number in a 1d Hankel matrix does not show up on the diagonal.

Figure 7 a) shows how a semi-circular domain Ω is constructed from the two domains Υ and Ξ . Note that the set Ξ is a rather small domain - a grid of size 11×11 . The size of Ξ will determine how many frequencies that can be recovered, in this case $11 \times 10 = 110$, (which follows by the same argument that led to formula (3.3)). The general domain Hankel matrix that is constructed from f is shown in Figure 7 b).

6. Frequency retrieval; general domain ESPRIT

As earlier, we assume that we have samples of a function f of the form

$$f(x) = \sum_{k=1}^K c_k e^{\zeta_k \cdot x}, \quad (6.1)$$

where the axes have been scaled so that we sample at integer points. However, we now assume that samples are available only on Ω which is a domain of the form $\Xi + \Upsilon$ for some domains Ξ and Υ , where we only require that Ξ is convex (in the sense that it arise as the discretization of a

convex domain). Given $\lambda \in \mathbb{C}^d$ we let $\Lambda_\Omega(\lambda)$ denote the md-sequence λ^j for $j \in \Omega$, where we use multi-index notation $\lambda^j = \lambda_1^{j_1} \lambda_2^{j_2} \dots \lambda_d^{j_d}$. Rephrased, we suppose that f can be written

$$f = \sum_{k=1}^K c_k \Lambda_\Omega(\lambda_k), \quad \lambda_k = (e^{\zeta_{k,1}}, \dots, e^{\zeta_{k,d}}). \quad (6.2)$$

It is then easy to see that

$$H_f(g) = \sum_{k=1}^K c_k H_{\Lambda_\Omega(\lambda_k)}(g) = \sum_{k=1}^K c_k \Lambda_\Xi(\lambda_k) \langle g, \overline{\Lambda_\Upsilon(\lambda_k)} \rangle, \quad (6.3)$$

which in particular shows that H_f is a rank K operator, assuming of course that the c_k 's are non-zero, the λ_k 's are distinct and that $\{\Lambda_\Xi(\lambda_k)\}_{k=1}^K$ and $\{\Lambda_\Upsilon(\lambda_k)\}_{k=1}^K$ form linearly independent sets.

A formal investigation of when this condition holds becomes very involved, but is possible given certain assumptions on Ξ and Υ , we refer to the work of B. Mourrain et al. [20, 8]. Another viewpoint is to consider general domain Hankel operators as discretizations of corresponding integral operators, whose rank structure is easier to characterize, we refer to the articles [2, 4] (by the authors) as well as the recent contribution [19] by B. Mourrain. In particular, it is well known that any set of exponential functions with different exponents always is linearly independent, as functions on whatever fixed open set in \mathbb{R}^n . The condition that $\{\Lambda_\Upsilon(\lambda_k)\}_{k=1}^K$ be linearly independent will thus always be satisfied given that the sampling is dense enough. We refer to Section 6 of [4] for more information on passing between discrete (summing) and continuous (integral) general domain Hankel operators. In particular, technical conditions (see (6.3)) on the boundary of the domains in \mathbb{R}^d which correspond to Υ and Ξ are given so that the sampled md-vectors converge, in a sense made precise, to the corresponding exponential functions (see the proof of Theorem 6.4 in [4]). In the remainder of this paper, we assume that the linear independence condition is fulfilled.

Letting o_x be the reverse lexicographical ordering on Ξ , we identify md-sequences u in $\ell^2(\Xi)$ with sequences (vectors) \mathbf{u} in $\ell^2(|\Xi|)$, as explained in Section 5. The SVD of \mathbf{H}_f thus gives rise to singular vectors $\mathbf{u}_k, \mathbf{v}_k$, whose md-sequence counterparts satisfy

$$H_f(v_k) = \sigma_k u_k, \quad H_f^*(u_k) = \sigma_k v_k.$$

Since $\{u_k\}_{k=1}^K$ span the range of H_f , i.e. the span of $\{\Lambda_\Xi(\lambda_k)\}_{k=1}^K$ by (6.3), it follows that we can write

$$u_j = \sum_{k=1}^K b_{j,k} \Lambda_\Xi(\lambda_k) \quad (6.4)$$

where the numbers $b_{j,k}$ form a square $K \times K$ invertible matrix. Let \mathbf{U} be the matrix with columns $\mathbf{u}_1, \dots, \mathbf{u}_K$ and let $\mathbf{\Lambda}$ be the matrix with the columns $\mathbf{\Lambda}_\Xi(\lambda_1), \dots, \mathbf{\Lambda}_\Xi(\lambda_K)$. The relation (6.4) can then be expressed

$$\mathbf{U} = \mathbf{\Lambda} \mathbf{B} \quad (6.5)$$

(compare with (3.2)).

By a ‘‘fiber’’ in Ξ we refer to a subset obtained by freezing all variables but one. Since we have assumed that the grid Ξ is the discretization of a convex domain, we have that each fiber in the first dimension is of the form $\{M_1, \dots, M_2\} \times \{m_2\} \times \dots \times \{m_d\}$, where $M_1 \leq M_2$ depend on the

“frozen” variables m_2, \dots, m_d . We denote the grid that arises by removing the first (respectively last) element of each such fiber by $\Xi_{+,1}$ (respectively $\Xi_{-,1}$), where we assume that the sampling has been done so that no fiber consists of a singleton. Analogous definitions/assumptions apply to the other variables, yielding grids $\Xi_{\pm,2}, \dots, \Xi_{\pm,d}$.

Now, given a fixed dimension p and a md-sequence w in $\ell^2(\Xi)$, we let $w_{\pm,p}$ be the md-sequence restricted to the grid $\Xi_{\pm,p}$. Moreover, given a matrix like \mathbf{U} , whose columns are given by the vectorizations of the md-sequences u_k in $\ell^2(\Xi)$, we denote by $\mathbf{U}_{+,p}$, (resp. $\mathbf{U}_{-,p}$) the matrix formed by the vectorized md-sequences $u_{k,+,p} \in \ell^2(\Xi_{+,p})$ (resp. $u_{k,-,p} \in \ell^2(\Xi_{-,p})$). Equation (6.5) then implies that

$$\mathbf{U}_{\pm,p} = \mathbf{\Lambda}_{\pm,p} B. \quad (6.6)$$

Setting $D_p = \text{diag}(e^{\zeta_{1,p}}, \dots, e^{\zeta_{K,p}})$, we also have $\mathbf{\Lambda}_{+,p} = \mathbf{\Lambda}_{-,p} D_p$ which combined implies that

$$\mathbf{U}_{+,p} = \mathbf{\Lambda}_{-,p} D_p B. \quad (6.7)$$

In analogy with the computations in Section 3 we have

$$A_p := (\mathbf{U}_{-,p}^* \mathbf{U}_{-,p})^{-1} \mathbf{U}_{-,p}^* \mathbf{U}_{+,p} = (B^* \mathbf{\Lambda}_{-,p}^* \mathbf{\Lambda}_{-,p} B)^{-1} B^* \mathbf{\Lambda}_{-,p}^* \mathbf{\Lambda}_{-,p} D_p B = B^{-1} D_p B$$

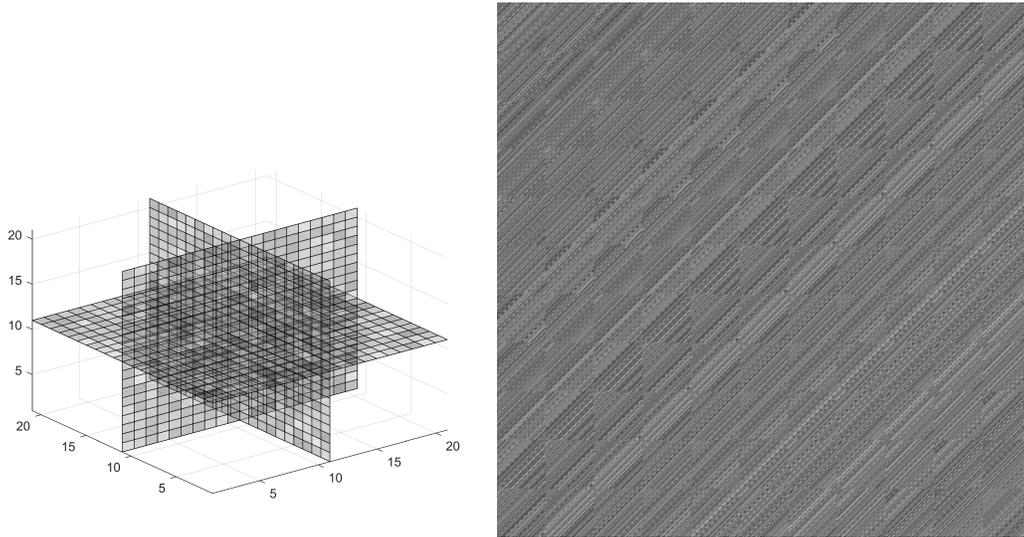
so we can retrieve the desired frequencies $\lambda_{k,p} = e^{\zeta_{k,p}}$ by computing A_p and diagonalize it. Since all matrices A_p are diagonalized by the same matrix B , the issue with grouping of the complex frequencies is easily solved just as in the previous case. The algorithm, which we call general domain ESPRIT, is summarized in Algorithm 3.

Algorithm 3 General domain ESPRIT

- 1: Order the elements in Ξ and Υ .
 - 2: Form the general domain Hankel matrix $\mathbf{H}(\mathbf{m}, \mathbf{n}) = f(m+n)$ from samples of f , where e.g. \mathbf{m} is the order of the multi-index m in Ξ .
 - 3: Compute the singular value decomposition $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$.
 - 4: **for** $p = 1, \dots, d$ **do**
 - 5: Form \mathbf{U}_{p+} and \mathbf{U}_{p-} by deleting first and last elements in each fiber in the p :th coordinate, respectively.
 - 6: Form $A_p = (\mathbf{U}_{p-}^* \mathbf{U}_{p-})^{-1} \mathbf{U}_{p-}^* \mathbf{U}_{p+}$.
 - 7: **end for**
 - 8: Diagonalize $A_1 = B^{-1} D_1 B$ by making an eigenvalue decomposition of A_1 .
 - 9: **for** $p = 1, \dots, d$ **do**
 - 10: Compute the diagonal matrices $D_p = \text{diag}(\lambda_{1,p}, \dots, \lambda_{K,p}) = B A_p B^{-1}$.
 - 11: Recover $\zeta_{k,p} = \log(\lambda_{k,p})$ (which are automatically correctly paired).
 - 12: **end for**
-

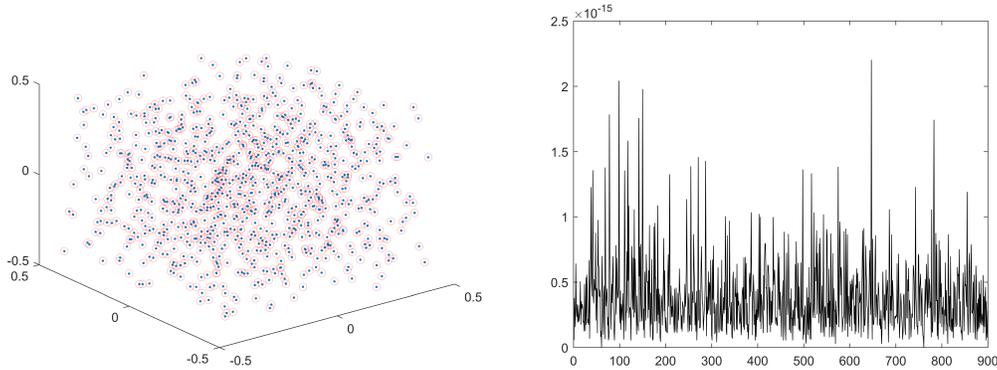
7. Numerical examples

To illustrate the methods discussed we will perform a number of numerical simulations. We have already shown Figure 1 and 2, in which 300 and 100 frequencies in 2d were retrieved up to machine precision, using Algorithm 2 (in the first case) and Algorithm 3 (in the second case). We continue here with a 3d example in the absence of noise, and then end by briefly discussing what



(a) Real part of a 3d data set sampled $21 \times 21 \times 21$ lattice ($N = 11$).

(b) The corresponding block-Hankel matrix



(c) True 3d frequencies in blue dots and estimated frequencies in red circles

(d) Error between estimated and true frequencies

Figure 8: Three dimensional example with data being a linear combination of 900 exponential functions.

happens in the presence of noise. Note that due to the non-linearity of the problem, this can behave quite differently in different situations.

The layout of Figure 8 is similar to that of Figure 1. In this case 900 purely oscillatory exponential functions were used to generate a function f sampled on a grid of size $21 \times 21 \times 21$. In contrast to Figure 1, the frequencies are now randomly distributed. The real part of f is shown in panel a), the (real part of the) corresponding block-Hankel matrix is shown in panel b), the distribution of the exponentials in panel c) and the frequency reconstruction error is shown in panel d). According to (3.3), the maximal amount of frequencies we may retrieve in this situation is $11^2 * 10 = 1210$, to

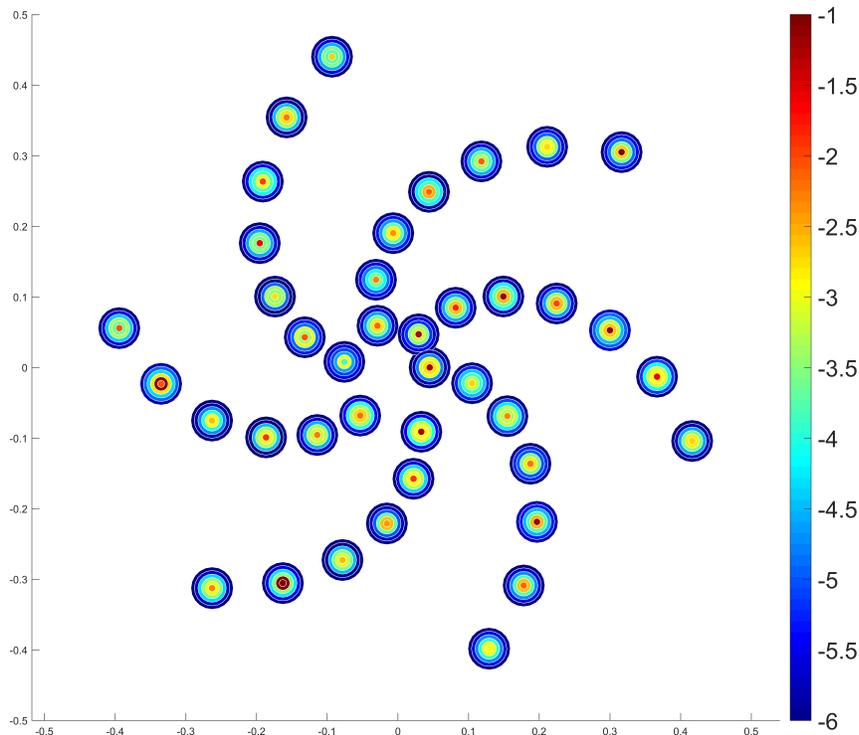


Figure 9: Illustration of error in the estimation of frequencies. The size of the discs illustrate the estimation error at each frequency for the corresponding noise level in \log_{10} -scale.

be compared with the total available data points $(21)^3 = 9261$.

Finally, we briefly illustrate the impact of noise on the proposed algorithms. However, we underline that the purpose of the present article is to provide an algorithm that correctly retrieves complex frequencies for functions of the form (6.1) in the absence of noise. If noise is present it is no longer true that the different A_p 's share eigenvectors, and hence choosing the eigenvectors of A_1 (step 8 in Algorithm 3) to diagonalize the rest becomes ad hoc. One way around this is to use some method to find the best simultaneous diagonalization of $\{A_p\}_{p=1}^d$ (see e.g. [9]), another is to use tools from optimization to preprocess the function f so that it does become of the form (6.1) (see e.g. [3]).

The scatter plot in Figure 9 shows the distribution of 40 frequencies in a case of purely oscillatory exponential functions. These are used to generate a function f that is then sampled on a 41×41 grid. In addition, different levels of normally distributed noise is added to the original data. Six noise levels are chosen so that the ratio between the ℓ^2 norm of the noise and the noise-free data is 10^0 , 10^{-5} , 10^{-1} , 10^{-2} , 10^{-3} , and 10^{-4} , respectively. Each noise level is portrayed with a particular color, according to the list in 10. The impact that the noise have on the singular values of the

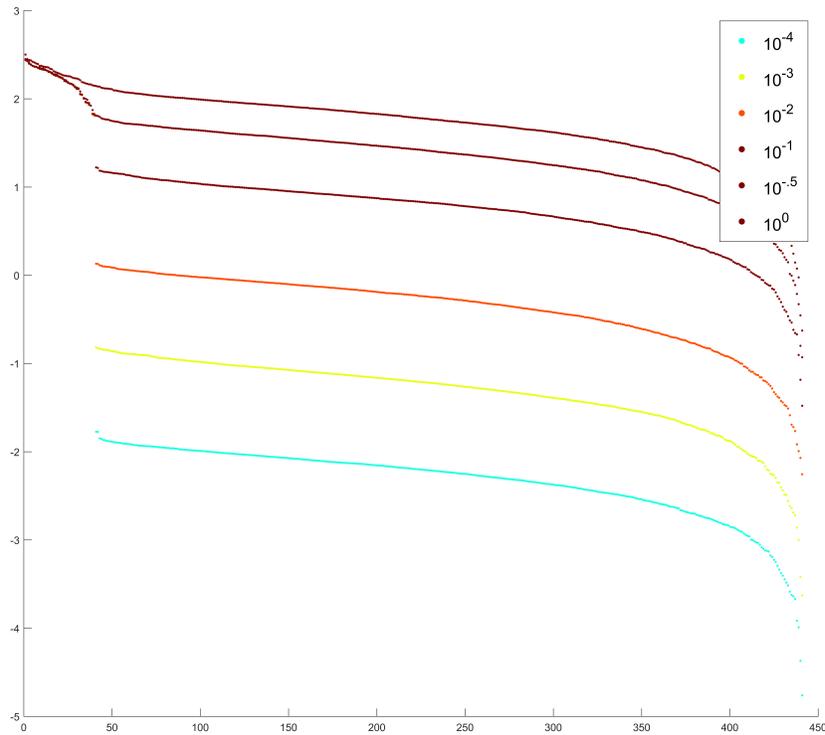


Figure 10: The impact on the singular values for different noise levels.

corresponding block-Hankel matrix is also shown in Figure 10. Recall that the noise free block Hankel matrix has rank 40, i.e. 40 non-zero singular values. When low noise is present this shows up as a jump in the magnitude of the singular values, clearly visible in plot 10. The first 40 singular values for e.g. yellow and orange are covered by the brown ones, therefore not visible. We can see that for the four lower levels, the impact of noise does not affect the original distribution on the singular values much, whereas for the highest noise level it affects essentially all singular values, and for the second highest level, it is just starting to have an impact. From this, we would expect to see a high error for the highest noise level, and a relatively small error for the four smallest noise levels. To illustrate the effect on the individual frequency nodes, each of the nodes in Figure 9 have 6 circles around it. The color of each one of these circles show the error level in logarithmic scale according to the colorbar. Here we can see that the impact is pretty much as could be expected, with a low error for the lower noise levels and increasing for higher noise levels.

8. Conclusions

We have shown how to extract the underlying multi-dimensional frequencies from data that is constructed as linear combinations of (oscillating) exponential functions both for rectangular domains and more general domains. The approach does not require a pairing of one-dimensional frequencies components.

References

- [1] Vadim M Adamjan, Damir Z Arov, and MG Kreĭn. Analytic properties of schmidt pairs for a hankel operator and the generalized schur-takagi problem. *Sbornik: Mathematics*, 15(1):31–73, 1971.
- [2] Fredrik Andersson and Marcus Carlsson. On general domain truncated correlation and convolution operators with finite rank. *Integr. Eq. Op. Th.*, 82(3), 2015.
- [3] Fredrik Andersson and Marcus Carlsson. Fixed-point algorithms for frequency estimation and structured low rank approximation. *Applied and Computational Harmonic Analysis*, 2017.
- [4] Fredrik Andersson and Marcus Carlsson. On the structure of positive semi-definite finite rank general domain Hankel and Toeplitz operators in several variables. *Complex Analysis and Operator Theory*, 11(4):755–784, 2017.
- [5] Fredrik Andersson, Marcus Carlsson, and Maarten de Hoop. Nonlinear approximation of functions in two dimensions by sums of exponential functions. *Applied and Computational Harmonic Analysis*, 29(2):156–181, 2010.
- [6] Fredrik Andersson, Marcus Carlsson, and V Maarten. Sparse approximation of functions using sums of exponentials and aak theory. *Journal of Approximation Theory*, 163(2):213–248, 2011.
- [7] Gregory Beylkin and Lucas Monzón. On approximation of functions by exponential sums. *Applied and Computational Harmonic Analysis*, 19(1):17–48, 2005.
- [8] Jerome Brachat, Pierre Comon, Bernard Mourrain, and Elias Tsigaridas. Symmetric tensor decomposition. *Linear Algebra and its Applications*, 433(11-12):1851–1872, 2010.
- [9] Jean-Francois Cardoso and Antoine Souloumiac. Jacobi angles for simultaneous diagonalization. *SIAM journal on matrix analysis and applications*, 17(1):161–164, 1996.
- [10] Annie Cuyt. Multivariate Padé-approximants. *Journal of mathematical analysis and applications*, 96(1):283–293, 1983.
- [11] Annie Cuyt and Wen-shin Lee. Sparse interpolation of multivariate rational functions. *Theoretical Computer Science*, 412(16):1445–1456, 2011.
- [12] Baron Gaspard Riche De Prony. Essai expérimental et analytique: sur les lois de la dilatabilité de fluides élastique et sur celles de la force expansive de la vapeur de l’alkool, a différentes températures. *Journal de l’École polytechnique*, 1(22):24–76, 1795.

- [13] Nina Golyandina, Anton Korobeynikov, Alex Shlemov, and Konstantin Usevich. Multivariate and 2d extensions of singular spectrum analysis with the rssa package. *Journal of Statistical Software*, page 41, 2014.
- [14] Jouhayna Harmouch, Houssam Khalil, and Bernard Mourrain. Structured low rank decomposition of multivariate hankel matrices. *Linear Algebra and its Applications*, 2017.
- [15] Yingbo Hua. Estimating two-dimensional frequencies by matrix enhancement and matrix pencil. *IEEE Transactions on Signal Processing*, 40(9):2267–2280, 1992.
- [16] Yingbo Hua and Tapan K Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(5):814–824, 1990.
- [17] Sun-Yuan Kung, K Si Arun, and DV Bhaskar Rao. State-space and singular-value decomposition-based approximation methods for the harmonic retrieval problem. *JOSA*, 73(12):1799–1811, 1983.
- [18] Stefan Kunis, Thomas Peter, Tim Römer, and Ulrich von der Ohe. A multivariate generalization of Prony’s method. *Linear Algebra and its Applications*, 490:31–47, 2016.
- [19] Bernard Mourrain. Polynomial–exponential decomposition from moments. *Foundations of Computational Mathematics*, pages 1–58, 2016.
- [20] Bernard Mourrain and Victor Y Pan. Multivariate polynomials, duality, and structured matrices. *Journal of complexity*, 16(1):110–180, 2000.
- [21] Thomas Peter, Gerlind Plonka, and Robert Schaback. Reconstruction of multivariate signals via Prony’s method. *Proc. Appl. Math. Mech., to appear*.
- [22] Vladilen F Pisarenko. The retrieval of harmonics from a covariance function. *Geophysical Journal International*, 33(3):347–366, 1973.
- [23] Gerlind Plonka and Marius Wischerhoff. How many fourier samples are needed for real function reconstruction? *Journal of Applied Mathematics and Computing*, 42(1-2):117–137, 2013.
- [24] Daniel Potts and Manfred Tasche. Parameter estimation for multivariate exponential sums. *Electronic Transactions on Numerical Analysis*, 40:204–224, 2013.
- [25] Daniel Potts and Manfred Tasche. Parameter estimation for nonincreasing exponential sums by prony-like methods. *Linear Algebra and its Applications*, 439(4):1024–1039, 2013.
- [26] Stephanie Rouquette and Mohamed Najim. Estimation of frequencies and damping factors by two-dimensional esprit type methods. *IEEE Transactions on signal processing*, 49(1):237–245, 2001.
- [27] Richard Roy and Thomas Kailath. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on acoustics, speech, and signal processing*, 37(7):984–995, 1989.

- [28] Joseph J Sacchini, William M Steedly, and Randolph L Moses. Two-dimensional Prony modeling and parameter estimation. *IEEE Transactions on signal processing*, 41(11):3127–3137, 1993.
- [29] Souleymen Sahnoun, Konstantin Usevich, and Pierre Comon. Multidimensional harmonic retrieval by nd esprit: Algorithm, computations and perturbation analysis. 2017.
- [30] Tomas Sauer. Prony’s method in several variables. *Numerische Mathematik*, pages 1–28, 2016.
- [31] Ralph Schmidt. Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation*, 34(3):276–280, 1986.
- [32] A. Shlemov and N. Golyandina. Shaped extension of singular spectrum analysis. In *Proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems (MTNS 2014)*, July 7-11, 2014, Groningen, The Netherlands, 2014.
- [33] Jens Steinwandt, Florian Roemer, Martin Haardt, and Giovanni Del Galdo. Performance analysis of multi-dimensional esprit-type algorithms for arbitrary and strictly non-circular sources with spatial smoothing. *IEEE Transactions on Signal Processing*, 65(9):2262–2276, 2017.
- [34] Johan Swärd, Stefan Ingi Adalbjörnsson, and Andreas Jakobsson. Computationally efficient estimation of multi-dimensional spectral lines. In *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pages 4885–4889. IEEE, 2016.
- [35] Filiep Vanpoucke, Marc Moonen, and Yannick Berthoumieu. An efficient subspace algorithm for 2-d harmonic retrieval. In *Acoustics, Speech, and Signal Processing, 1994. ICASSP-94., 1994 IEEE International Conference on*, volume 4, pages IV–461. IEEE, 1994.
- [36] Yi Zhou, Da-zheng Feng, and Jian-qiang Liu. A novel algorithm for two-dimensional frequency estimation. *Signal Processing*, 87(1):1–12, 2007.