A RANDOMIZED MULTIVARIATE MATRIX PENCIL METHOD FOR SUPERRESOLUTION MICROSCOPY

MARTIN EHLER, STEFAN KUNIS, THOMAS PETER, AND CHRISTIAN RICHTER

Abstract. The matrix pencil method is an eigenvalue based approach for the parameter identification of sparse exponential sums. We derive a reconstruction algorithm for multivariate exponential sums that is based on simultaneous diagonalization. Randomization is used and quantified to reduce the simultaneous diagonalization to the eigendecomposition of a single random matrix. To verify feasibility, the algorithm is applied to synthetic and experimental fluorescence microscopy data.

1. Introduction

Many imaging and data analysis problems in the applied sciences lead to the numerical task of parameter identification in exponential sums $\sum_{j=1}^{M} c_j \mathrm{e}^{-2\pi i \langle t_j, \cdot \rangle}$. For sparse exponential sums, i.e., for small M, Prony's method enables the identification of its parameters $\{t_j\}_{j=1}^{M} \subset \mathbb{R}^d$ and contributions $\{c_j\}_{j=1}^{M} \subset \mathbb{C}$ from relatively few sampling values, see e.g. [13, 14] and references therein.

The most feasible implementations for $d = 1$ are based on the eigenvalue analysis of the associated Prony matrix, see e.g. [2, 11]. The principles of the multivariate setting have been examined in [8, 7, 1, 10], for instance, but associated numerical schemes have not been extensively studied yet.

Here, we shall develop a numerical scheme based on a randomized multivariate matrix pencil method. We construct matrices $S_1, \ldots, S_d$ from the sampling values, so that their simultaneous diagonalization yields the parameters $\{t_j\}_{j=1}^{M}$. Since $S_1, \ldots, S_d$ are not normal, standard numerical algorithms for simultaneous diagonalization are not available, cf. [3, 4, 5, 6]. To circumvent this problem, we derive the joint eigenbasis from the eigendecomposition of a single matrix that is a random linear combination of $S_1, \ldots, S_d$. While [1] diagonalizes S_1 and hopes for simple eigenvalues, the recent paper [10, Alg. 3.1] also uses the above random linear combination and argues that generically the eigenvalues are simple. Here, we describe the situation in more detail and quantify the influence of the minimal separation of $\{t_j\}_{j=1}^{M}$ on the eigendecomposition of the random matrix.

To check on its feasibility, our methodology is applied to analyze fluorescence microscopy images. We cast the problem of locating protein markers as a parameter identification in exponential sums. Due to its analytic roots, Prony's method enables the identification of locations at the subpixel scale, sometimes referred to as superresolution fluorescence microscopy, cf. [15]. The results on experimental fluorescence images show that our scheme is numerically feasible.

2010 *Mathematics Subject Classification.* 65T40, 42C15, 30E05, 65F30.

Key words and phrases. frequency analysis, spectral analysis, exponential sum, moment problem, super-resolution.

The outline is as follows: In Section 2 we develop our numerical scheme. The approach of simultaneous diagonalization to identify $\{t_j\}_{j=1}^M$ is presented in Section 2.1. The problem of simultaneous diagonalization is reduced to the diagonalization of a single random matrix in Section 2.2, where we examine the influence of the minimal separation of the parameters $\{t_j\}_{j=1}^M$. Our new scheme is applied to synthetic and to experimental fluorescence microscopy data in Section 3.

2. Reconstruction of sparse exponential sums from samples

Let $\{t_j\}_{j=1}^M \subset [0,1)^d$ always denote M pairwise different d-dimensional parameters and consider the exponential sum

$$f(k) = \sum_{j=1}^{M} c_j \mathrm{e}^{-2\pi \mathrm{i}\langle t_j, k\rangle}, \quad k \in \mathbb{Z}^d, \tag{2.1}$$

with nonzero coefficients $\{c_j\}_{j=1}^M \subset \mathbb{C}\backslash\{0\}$. Our aim is to identify the parameters $\{t_j\}_{j=1}^M$ and coefficients $\{c_j\}_{j=1}^M$ from sampling values $\{f(k)\}_{k\in I}$ with suitable $I \subset \mathbb{Z}^d$.

2.1. Reconstruction by simultaneous diagonalization.

For $n \in \mathbb{N}$, let $I_n := \{0,\dots,n\}^d$ and select a fixed ordering of the elements in I_n. Knowledge of the sampling values of f on the set difference $I := I_{n+1} - I_n$ enables us to build the matrices

$$T := (f(k-l))_{k,l\in I_n}, \qquad T_\ell := (f(k-l+e_\ell))_{k,l\in I_n}, \quad \ell = 1,\dots,d.$$

If T has rank M, then we compute the reduced singular value decomposition

$$T = U\Sigma V^*,$$

where $\Sigma \in \mathbb{R}^{M\times M}$ is positive definite and $U \in \mathbb{C}^{N\times M}$, $V \in \mathbb{C}^{M\times N}$ satisfy $U^*U = V^*V = \mathrm{id} \in \mathbb{R}^{M\times M}$ with $N := \#I_n = (n+1)^d$. Therefore, we can define the set of $M \times M$ matrices

$$S_\ell := U^* T_\ell V \Sigma^{-1}, \quad \ell = 1,\dots,d. \tag{2.2}$$

These matrices turn out to be simultaneous diagonalizable, cf. Theorem 2.1, which shall enable us to identify the vectors $\{t_j\}_{j=1}^M$.

In the following theorem, K_d denotes an absolute constant that only depends on d and is further specified in [7, 8]. We also make use of

$$z_j := \mathrm{e}^{-2\pi i t_j} := (\mathrm{e}^{-2\pi i t_{j,1}}, \dots, \mathrm{e}^{-2\pi i t_{j,d}}), \quad j = 1,\dots,M,$$

so that it is sufficient to reconstruct $\{z_j\}_{j=1}^M$ in order to identify $\{t_j\}_{j=1}^M$.

Theorem 2.1. *If $n \geq \frac{K_d}{\min_{i\neq j}\|z_i - z_j\|}$, then T has rank M and $S_1,\dots,S_d$ are simultaneously diagonalizable. Furthermore, any regular matrix W that simultaneously diagonalizes $S_1,\dots,S_d$ yields a permutation τ on $\{1,\dots,M\}$ such that*

$$W^{-1} S_\ell W = \mathrm{diag}(\langle z_{\tau(1)}, e_\ell\rangle, \dots, \langle z_{\tau(M)}, e_\ell\rangle), \quad \ell = 1,\dots,d.$$

Proof. According to [8], T always admits the factorization

$$T = A^* D A, \tag{2.3}$$

where A is the $M \times N$ multivariate complex Vandermonde matrix

$$A = \left(z_j^k\right)_{\substack{j=1,\dots,M \\ k\in I_n}},$$

and $D = \operatorname{diag}(c_1, \dots, c_M)$. The condition on n implies that A has full rank M, cf. [7, 8]. Hence, T has indeed rank M since all $c_1, \dots, c_M$ are nonzero.

We also deduce the factorization

$$T_\ell = A^* D_\ell A, \quad \ell = 1, \dots, d,$$

where the diagonal matrix D_ℓ is given by

$$D_\ell := \operatorname{diag}(c_1 \langle z_1, e_\ell \rangle, \dots, c_M \langle z_M, e_\ell \rangle), \quad \ell = 1, \dots, d.$$

We shall now check that the specific matrix $W_0 := (AU)^*$ (which is not accessible to us) simultaneously diagonalizes $S_1, \dots, S_d$. Indeed, by inserting the definitions, we obtain

$$W_0^{-1} S_\ell W_0 = (AU)^{-*} U^* A^* D_\ell A V \Sigma^{-1} (AU)^*.$$

Note that the reduced singular value decomposition implies that both matrices, AU and AV, are regular. Since $\Sigma = U^* T V = U^* A^* D A V$, we deduce $\Sigma^{-1} = (AV)^{-1} D^{-1} (AU)^{-*}$, which implies

$$W_0^{-1} S_\ell W_0 = D_\ell D^{-1} = \operatorname{diag}(\langle z_1, e_\ell \rangle, \dots, \langle z_M, e_\ell \rangle), \quad \ell = 1, \dots, d,$$

so that W_0 simultaneously diagonalizes $S_1, \dots, S_d$. Note that W_0 also diagonalizes any complex linear combination

$$C_\mu := \sum_{\ell=1}^{d} \overline{\mu}_\ell S_\ell, \quad \mu \in \mathbb{C}^d. \tag{2.4}$$

Because of

$$W_0^{-1} C_\mu W_0 = \operatorname{diag}\left(\sum_{\ell=1}^{d} \bar{\mu}_\ell \langle z_1, e_\ell \rangle, \dots, \sum_{\ell=1}^{d} \bar{\mu}_\ell \langle z_1, e_\ell \rangle \right),$$

the eigenvalues $\lambda_1(\mu), \dots, \lambda_M(\mu)$ of C_μ are

$$\lambda_j(\mu) = \langle z_j, \mu \rangle$$

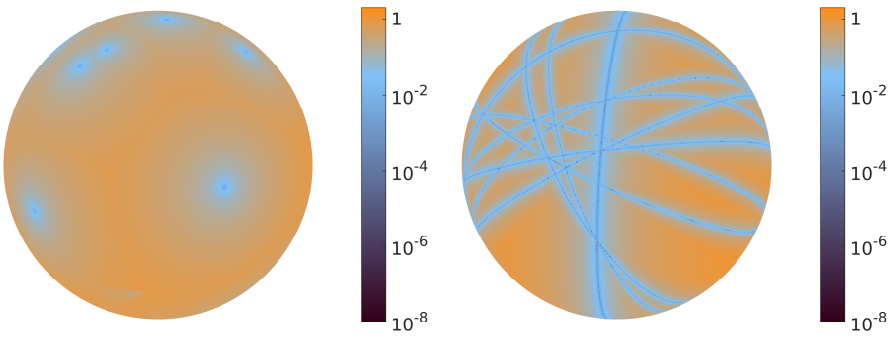
with the ordering induced by W_0. Since $\{t_j\}_{j=1}^M$ are pairwise different, so are $\{z_j\}_{j=1}^M$, and, hence, there is $\tilde{\mu} \in \mathbb{S}_{\mathbb{C}}^{d-1} = \{x \in \mathbb{C}^d : \|x\| = 1\}$ such that $\langle z_i - z_j, \tilde{\mu} \rangle \neq 0$ for all $i \neq j$ and thus $\{\lambda_j(\tilde{\mu})\}_{j=1}^M$ are pairwise different. In other words, all eigenspaces of $C_{\tilde{\mu}}$ are 1-dimensional.

Any matrix $W = (w_1, \dots, w_M)$ that simultaneously diagonalizes $S_1, \dots, S_d$ also diagonalizes $C_{\tilde{\mu}}$. Thus, there is a permutation τ such that $w_{\tau(i)}$ spans the same space as the i-th column of W_0, which concludes the proof. ∎

According to Theorem 2.1, the diagonalization of S_ℓ encodes the ℓ-th entry of a permutation of the vectors $\{z_j\}_{j=1}^M$. We require simultaneous diagonalization to ensure that these entries are associated to the same permutation across all $\ell = 1, \dots, d$.

In general, the matrices $S_1, \dots, S_d$ are not normal. Therefore, the numerical task of simultaneous diagonalization is difficult and many simultaneous diagonalization algorithms in the literature are not suitable, cf. [3, 4, 5, 6]. We attempt to circumvent such issues by using C_μ from (2.4), which shall enable us to restrict our diagonalization efforts to a single matrix:

Corollary 2.2. *If $\mu \in \mathbb{C}^d$ is such that $\lambda_1(\mu), \dots, \lambda_M(\mu)$ are pairwise different, then any matrix W that diagonalizes C_μ also simultaneously diagonalizes $S_1, \dots, S_d$.*



(a) $S_1, S_2 \in \mathbb{C}^{5\times 5}$, $\mu \in \mathbb{S}^1_{\mathbb{C}}$ (b) $d = 3$, $M = 5$, and $\mu \in \mathbb{S}^2$

FIGURE 2.1. Visualization of the smallest distance of any two eigenvalues of C_μ.

Proof. The matrices $C_\mu, S_1, \dots, S_d$ are simultaneously diagonalizable. The same arguments as in the proof of Theorem 2.1 imply the assertion. ∎

According to Corollary 2.2 we aim to find $\mu \in \mathbb{C}^d$ such that $\lambda_1(\mu), \dots, \lambda_M(\mu)$ are pairwise different. For a nonzero vector $z \in \mathbb{C}^d$, let $z^\perp$ denote the $d-1$-dimensional linear subspace of $\mathbb{C}^d$ orthogonal to z. The proof of Theorem 2.1 reveals that

$$\{\mu \in \mathbb{C}^d : \lambda_1(\mu), \dots, \lambda_M(\mu) \text{ are pairwise different}\} = \mathbb{C}^d \setminus \bigcup_{i\neq j} (z_i - z_j)^\perp \tag{2.5}$$

Hence, this set is the entire $\mathbb{C}^d$ except for at most $\binom{M}{2}$ many $(d-1)$-dimensional subspaces.

Example 2.3. *Let $d = 2$, $M = 5$, and choose $t_1, \dots, t_5 \in [0,1)^2$ randomly. We construct $S_1, S_2 \in \mathbb{C}^{5\times 5}$ by (2.2). Thus, we choose $\mu = (\mu_1, \mu_2)^\top \in \mathbb{S}^1_{\mathbb{C}}$ and construct $C_\mu = \mu_1 S_1 + \mu_2 S_2$. According to (2.5) we expect $\binom{5}{2} = 10$ great circles on $\mathbb{S}^1_{\mathbb{C}}$, with the property that choosing a μ from one of those great circles results in a C_μ, that has at least one eigenspace of dimension larger than one. For $\xi \in \mathbb{C}$, with $\|\xi\| = 1$ we get $C_{\mu\xi} = \xi\,(\mu_1 S_1 + \mu_2 S_2)$. This shows that the multiplication of C_μ by a global phase ξ does not change the pairwise differences of the eigenvalues of C_μ and therefore we can use Hopf fibration, to identify great circles on $\mathbb{S}^1_{\mathbb{C}}$ with a single point on $\mathbb{S}^2$, for visualization. Indeed we can observe that the minimal distance of any two eigenvalues of C_μ is nonzero on $\mathbb{S}^2$ except for 10 points, see Figure 2.1(a). Note, that we only see 8 of those 10 points in 2.1(a), the other 2 are on the back side of the sphere.*

For visual illustration of the expected great circles, we now switch to the real case and choose $d = 3$, $M = 5$, and restrict μ to the real sphere $\mathbb{S}^2$. In Figure 2.1(b) we see 10 great circles on $\mathbb{S}^2$, for which C_μ has eigenspaces of dimension larger than one. Observe that away from those great circles, the minimal distance of any two eigenvalues of C_μ rapidly increases.

Remark 2.4. *Our approach to simultaneous diagonalization of $S_1, \dots, S_d$ suggested in Corollary 2.2 requires our present setting, in which $\{z_j\}_{j=1}^M$ are pairwise different. It does not apply to the problem of simultaneous diagonalization in general.*

2.2. **Simultaneous diagonalization by random linear combinations.** The present section is dedicated to quantify the difference $\lambda_i(\mu) - \lambda_j(\mu)$ in relation to the difference $z_i - z_j$. If $\mu \in \mathbb{S}^{d-1}_{\mathbb{C}}$ is a random vector, distributed according to the unitarily invariant probability measure on $\mathbb{S}^{d-1}_{\mathbb{C}}$, then

$$\mathbb{E}|\lambda_i(\mu) - \lambda_j(\mu)| = \frac{1}{\sqrt{d}}\|z_i - z_j\|.$$

The following result provides a more quantitative analysis:

Theorem 2.5. *Let $i \neq j$ be fixed and suppose $\epsilon \in [0,1]$. If $\mu \in \mathbb{S}^{d-1}_{\mathbb{C}}$ is a random vector, distributed according to the unitarily invariant probability measure on $\mathbb{S}^{d-1}_{\mathbb{C}}$, then the probability that*

$$|\lambda_i(\mu) - \lambda_j(\mu)| < \epsilon\|z_i - z_j\| \tag{2.6}$$

holds is at most $2\sqrt{\frac{d}{\pi}}\epsilon$.

Theorem 2.5 immediately implies that the probability that any of the inequalities

$$|\lambda_i(\mu) - \lambda_j(\mu)| \geq \epsilon\|z_i - z_j\|, \quad \forall i \neq j, \tag{2.7}$$

is violated is at most $\binom{M}{2}2\sqrt{\frac{d}{\pi}}\epsilon$. In other words, if we select about M^2 many independent μ, then the probability that (2.7) fails is at most of the order ϵ.

Proof of Theorem 2.5. The complex sphere $\mathbb{S}^{d-1}_{\mathbb{C}}$ admits the standard identification with the real sphere $\mathbb{S}^{2d-1}$ by $x \mapsto \binom{\mathrm{Re}(x)}{\mathrm{Im}(x)}$, and $\binom{\mathrm{Re}(\mu)}{\mathrm{Im}(\mu)}$ is distributed according to the orthogonal invariant probability measure on $\mathbb{S}^{2d-1}$, the latter being the standard normalized surface measure.

Let $y := \frac{z_i - z_j}{\|z_i - z_j\|} \in \mathbb{S}^{d-1}_{\mathbb{C}}$, so that $|\lambda_i(\mu) - \lambda_j(\mu)|/\|z_i - z_j\| = |\langle y, \mu\rangle|$. Since

$$\left|\left\langle \binom{\mathrm{Re}(y)}{\mathrm{Im}(y)}, \binom{\mathrm{Re}(\mu)}{\mathrm{Im}(\mu)} \right\rangle\right| = |\,\mathrm{Re}\,(\langle y, \mu\rangle)| \leq |\langle y, \mu\rangle|, \tag{2.8}$$

we obtain an upper bound by simply considering

$$\left|\left\langle \binom{\mathrm{Re}(y)}{\mathrm{Im}(y)}, \binom{\mathrm{Re}(\mu)}{\mathrm{Im}(\mu)} \right\rangle\right| \leq \epsilon. \tag{2.9}$$

Due to the orthogonal invariance of the surface measure on $\mathbb{S}^{2d-1}$, the distribution of the left-hand-side in (2.9) does not depend on the special choice of $y \in \mathbb{S}^{d-1}_{\mathbb{C}}$, so that we can simply assume that $\binom{\mathrm{Re}(y)}{\mathrm{Im}(y)}$ is the north pole. The inequality (2.9) reduces to $-\epsilon \leq \mathrm{Re}(\mu_1) \leq \epsilon$, hence, describes the complement of two opposing spherical caps in $\mathbb{S}^{2d-1}$. This "equatorial band" has measure

$$1 - \mathcal{I}_{[1-\epsilon^2]}(d - \frac{1}{2}, \frac{1}{2}) = \mathcal{I}_{[\epsilon^2]}(\frac{1}{2}, d - \frac{1}{2}),$$

see, for instance, [9], where $\mathcal{I}_{[x]}(a,b)$ is the cumulative distribution function of the Beta distribution, i.e.,

$$\mathcal{I}_{[x]}(a,b) = \frac{\int_0^x t^{a-1}(1-t)^{b-1}dt}{\mathrm{Beta}(a,b)}, \qquad \mathrm{Beta}(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

For $d = 1$, we observe

$$\mathcal{I}_{[\epsilon^2]}(1/2, 1/2) = \frac{2\arcsin(\epsilon)}{\pi} \leq \frac{2}{\sqrt{\pi}}\epsilon.$$

Suppose now $d \geq 2$ and define

$$f(x) := 2\sqrt{x} - \mathcal{I}_{[x]}(1/2, d-1/2)\,\mathrm{Beta}(1/2, d-1/2).$$

A short calculation yields that its derivative satisfies

$$f'(x) = \frac{1-(1-x)^{d-3/2}}{\sqrt{x}} \geq 0, \quad x \in [0,1].$$

Since $f(0) = 0$, we obtain

$$\mathcal{I}_{[\epsilon^2]}(\frac{1}{2}, d-\frac{1}{2}) \leq \frac{2\epsilon}{\mathrm{Beta}(1/2, d-1/2)}, \quad \epsilon \in [0,1]. \tag{2.10}$$

The observation $1/\,\mathrm{Beta}(1/2, d-1/2) \leq \sqrt{d/\pi}$ concludes the proof. ∎

Remark 2.6. *A short calculation leads to*

$$\mathcal{I}_{[\epsilon^2]}(\frac{1}{2}, d-\frac{1}{2}) = \frac{2}{\pi}\Big[\arcsin(\epsilon) + \epsilon \sum_{k=2}^{d} \frac{4^{k-2}(k-2)!^2}{(2k-3)(2k-4)!}(1-\epsilon^2)^{k-3/2}\Big].$$

One then deduces directly that, for fixed d and small ϵ, the term $\mathcal{I}_{[\epsilon^2]}(\frac{1}{2}, d-\frac{1}{2})$ is of the order ϵ.

Theorem 2.1, Corollary 2.2, and Theorem 2.5 enable us to determine $z_{\tau(1)}, \ldots, z_{\tau(M)}$. The actual parameters $t_{\tau(j)}$ are computed as the principal values of $\log(z_{\tau(j)})$. The coefficients $c_{\tau(1)}, \ldots, c_{\tau(M)}$ can be determined by solving the linear system $T = A^* D A$ for $D = \mathrm{diag}(c_{\tau(1)}, \ldots, c_{\tau(M)})$ by the least squares method. We have summarized these steps in Algorithm 1.

Algorithm 1 Prony's method using the multivariate matrix pencil approach

1: **input** $f(k)$, $k \in I$.
2: Compute the reduced singular value decomposition of T.
3: Build the matrices $S_1, \ldots, S_d$.
4: Choose random $\mu \in \mathbb{S}_{\mathbb{C}}^{d-1}$ and compute a matrix W that diagonalizes C_μ.
5: Use W to simultaneously diagonalize $S_1, \ldots, S_d$ and reconstruct $z_{\tau(1)}, \ldots, z_{\tau(M)}$.
6: Compute $t_{\tau(j)}$ as the principal value of $\log(z_{\tau(j)})$, $j = 1, \ldots, M$.
7: Solve $\mathrm{argmin}_c \, \|A^* c - f\|_2$ to recover $c_{\tau(1)}, \ldots, c_{\tau(M)}$.
8: **return** $t_{\tau(1)}, \ldots, t_{\tau(M)}$ and $c_{\tau(1)}, \ldots, c_{\tau(M)}$.

3. Application in superresolution microscopy

3.1. **Mathematical model.** In fluorescence microscopy one puts a fluorescence marker on proteins and stimulates them with a laser. In accordance with the fluorescent microscope's resolution limits, proteins are modeled as point sources, cf. [15], so that the probe is considered a tempered distribution

$$G = \sum_{j=1}^{M} c_j \delta_{t_j}, \tag{3.1}$$

on $\mathbb{R}^d$, where $\{t_j\}_{j=1}^M \subset [0,1)^d$ is associated to the protein locations and δ_{t_j} denotes the Dirac delta function with center t_j. Let $\mathcal{F}$ denote the Fourier transform on the space of tempered distributions on $\mathbb{R}^d$. Then $\mathcal{F}(G)$ is an exponential sum

$$\mathcal{F}(G) = \sum_{j=1}^{M} c_j \mathrm{e}^{-2\pi i \langle t_j, \cdot \rangle}. \tag{3.2}$$

The actual measurements g are the convolution of G with some smooth and sufficiently fast decaying function φ,

$$g = G * \varphi = \sum_{j=1}^{M} c_j \varphi(\cdot - t_j).$$

Usually, φ is modeled as a Gaussian with known parameters determined by the camera system.

In order to determine the locations $\{t_j\}_{j=1}^M$ and the contributions $\{c_j\}_{j=1}^M$, suppose we have access to the Fourier transform of the measurements,

$$\mathcal{F}(g) = \mathcal{F}(G)\mathcal{F}(\varphi).$$

Since φ is known, let us also assume that we have access to $\mathcal{F}(\varphi)$. If φ is a Gaussian, for instance, we know $\mathcal{F}(\varphi)$ analytically. We now look for some sampling set $I \subset \mathbb{Z}^d$, where $\mathcal{F}(\varphi)$ does not vanish, and are able to determine the right-hand-side of

$$\mathcal{F}(G)(k) = \mathcal{F}(g)(k)/\mathcal{F}(\varphi)(k), \quad k \in I. \tag{3.3}$$

Combining (3.2) with (3.3) leads to the sampling problem (2.1) discussed in the previous sections, i.e.,

$$\sum_{j=1}^{M} c_j \mathrm{e}^{-2\pi i \langle t_j, k \rangle} = f(k), \qquad k \in I, \tag{3.4}$$

with $f(k) := \mathcal{F}(g)(k)/\mathcal{F}(\varphi)(k)$. The parameters $\{t_j\}_{j=1}^M$ and $\{c_j\}_{j=1}^M$ can now be determined by Algorithm 1 in principle. Note that the above derivations in this section have also been used in [12] in combination with the univariate Prony's method.

In practice though, we are not able to numerically compute the Fourier transform of g directly, so that the right-hand-side of (3.4) is not readily available. Aiming at the application of the discrete Fourier transform (DFT), we recognize that sufficient decay of φ implies $g \in L^1(\mathbb{R}^d)$, so that its periodization

$$g_{\mathrm{per}} := \sum_{l \in \mathbb{Z}^d} g(\cdot + l)$$

converges pointwise almost everywhere towards a function $g_{\mathrm{per}} \in L^1(\mathbb{T}^d)$, where $\mathbb{T}^d \simeq [0,1)^d$ is the d-dimensional torus. Let $\hat{g}_{\mathrm{per}}(k)$ denote the k-th Fourier coefficient of g_{per}. The Poisson formula yields

$$\mathcal{F}(g)(k) = \hat{g}_{\mathrm{per}}(k), \quad k \in I.$$

Thus, (3.4) can be evaluated by first computing the periodization g_{per}, so that its Fourier coefficients yield

$$\sum_{j=1}^{M} c_j \mathrm{e}^{-2\pi \mathrm{i} \langle t_j, k \rangle} = \hat{g}_{\mathrm{per}}(k)/\mathcal{F}(\varphi)(k), \qquad k \in I. \tag{3.5}$$

Numerically, the DFT enables the approximation of the Fourier coefficients $\hat{g}_{\mathrm{per}}(k)$, $k \in I$, from samples of g_{per}.

It should be mentioned that all numerical experiments were realized in Python on an Intel i7, 8GByte, 3GHz, macOS 10.12.

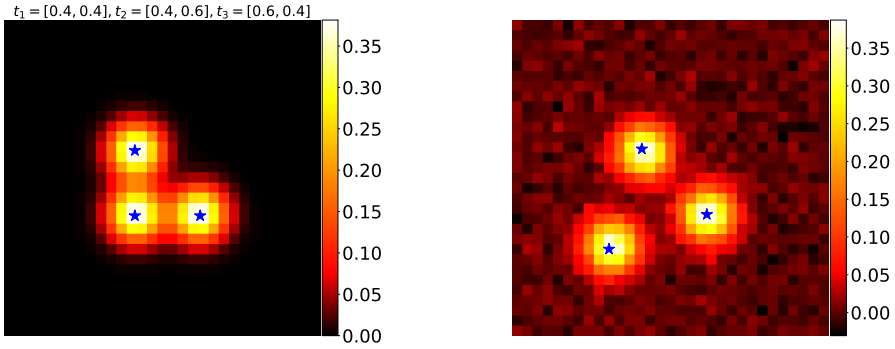
3.2. **Numerical results on synthetic data.** In our numerical experiments, we shall apply an implementation of the DFT to compute the discrete Fourier transform of samples of g_{per}. The sampling rate of g and hence g_{per} is determined by the pixel resolution. For both, synthetic and experimental fluorescence microscopy data, we choose $\varphi(\cdot) = \mathrm{e}^{-b\|\cdot\|^2}$ with adjusted parameter b derived from the camera system. Therefore, the values $\mathcal{F}(\varphi)$ are even available in analytic form.

Our analysis is first used on synthetic data in Figure 3.1 with

$$\begin{aligned} t_1 &= \left(\tfrac{2}{5}, \tfrac{2}{5}\right), & c_1 &= 1, & b &= 150, \\ t_2 &= \left(\tfrac{2}{5}, \tfrac{3}{5}\right), & c_2 &= 1, & & \\ t_3 &= \left(\tfrac{3}{5}, \tfrac{2}{5}\right), & c_3 &= 1. & & \end{aligned}$$

The measurements g are first exact and in a second experiment corrupted by additive Gaussian noise with a signal to noise ratio of SNR $= 2.554$, cf. Figure 3.1. For our computations we choose, if not stated otherwise, $n = 4$, so that $I = \{-4, \dots, 5\}^2$ and T is an $N \times N$ Toeplitz matrix with $N = 25$. These matrix dimensions show that our methodology is numerically feasible. By examining significant drops in the singular values of T, we determine M being 3 for the synthetic data. The reconstructed locations $\tilde{t}_1, \tilde{t}_2, \tilde{t}_3$ satisfy $\|t_j - \tilde{t}_j\| \leq 1.88 \cdot 10^{-3}$, for $i = 1, 2, 3$, in the noisy regime, and coincide with the correct locations up to machine precision in the noise-free regime, see Figure 3.1. It is important to note that our approach does not require the parameters $\{t_j\}_{j=1}^M$ to lie on the pixel grid. The pixel grid is only used to approximate $\hat{g}_{\mathrm{per}}(k)$, $k \in I$, by the DCT to determine the right-hand-side in (3.5). Indeed, the locations that we compute do not lie on the pixel grid, so we are identifying locations on the subpixel level. This is an important advantage we gain by making our computations in the Fourier domain. Figure 3.2 shows the difference between true locations $t_1 = 0.44, t_2 = 0.56$ of two one dimensional Gaussians, compared to the local maxima of their sum. For illustration purpose we use a one dimensional scenario in Figure 3.2. Even though this effect is negligible when $\|t_1 - t_2\|_2 \gg 0$, it would entail miscalculations when the positions t_1, t_2 of two proteins are close to each other. Consider a movie, where each frame is a picture as in Figure 1(b) and the found locations t_j are used to compute movement speeds of each protein. Then one would falsely compute an accelerated attraction and a longer contact phase of two approaching proteins if this effect is not considered.

To illustrate potential numerical issues when the measurements are corrupted by noise, i.e., when $\tilde{g} := g + \varepsilon$ is measured in place of g, we show in Figure 3.3 the real-parts of $\hat{g}_{\mathrm{per}}(k)$, $\hat{\tilde{g}}_{\mathrm{per}}(k)$, approximated by the DFT and $\mathcal{F}(\varphi)(k) = \hat{\varphi}_{\mathrm{per}}(k)$, as well as the respective ratios on a line $k_1 = 0$ and $k_2 = -15, \dots, 15$. Even though we are dealing with images of the size 31×31 pixels, the frequency data of the noisy ratio $\hat{\tilde{g}}_{\mathrm{per}}(k)/\hat{\varphi}_{\mathrm{per}}(k)$ seems only reliable close to the center. While $\hat{\varphi}_{\mathrm{per}}(k)$ decays with growing k, the noise keeps $\hat{\tilde{g}}_{\mathrm{per}}(k)$ from decaying, so that the ratio becomes unreasonably large. Therefore, we must restrict n depending on the noise level, and $n = 4$ seems to work in our synthetic data with fixed SNR as well as



(a) Blue stars indicate the three identified locations within noiseless synthetic data.

(b) Good location identification within synthetic data corrupted by additive Gaussian noise with SNR = 2.554.

FIGURE 3.1. In noiseless synthetic data and in the presence of additive Gaussian noise in spatial domain, our proposed algorithm manages to find the locations t_1, t_2, t_3 with reasonable accuracy.

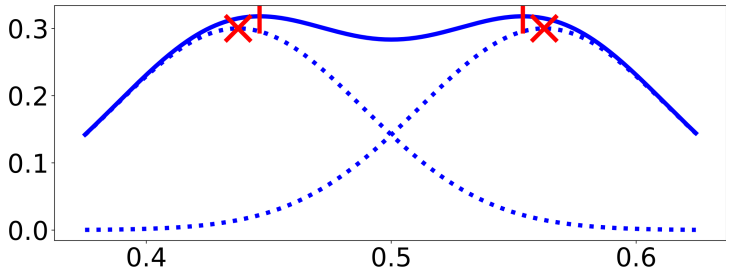


FIGURE 3.2. The red crosses show the true location of $t_1 = 0.44, t_2 = 0.56$ of two one-dimensional Gaussians, each depicted as a dotted line. The red bars however, show the local maxima of the sum of these gaussians and this sum is shown in a continuous line.

in our fluorescence microscopy data. Figure 3.4 shows the ratios $\hat{g}_{\text{per}}(k)/\hat{\varphi}(k)$ for $k \in \{-4, \ldots, 5\}^2$.

Theorem 2.1 requires n to be larger if the minimal separation distance

$$q := \min_{j \neq i} \|z_j - z_i\|$$

becomes smaller. In Figure 3.5 we illustrate this relation by two examples with noisy synthetic data, one for $q_1 = 0.283$ and the other for $q_2 = 0.057$. For $n = 1$ and $n = 4$, the locations can still be recovered reasonably well for q_1. In the case q_2, the choice $n = 1$ fails to recover the locations that are close to each other but $n = 4$ is successful.

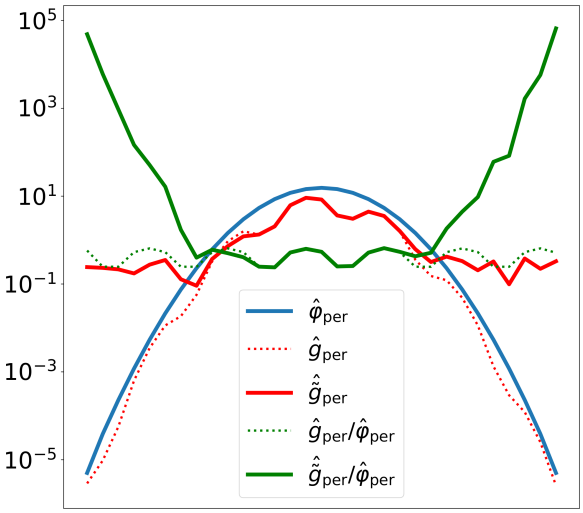
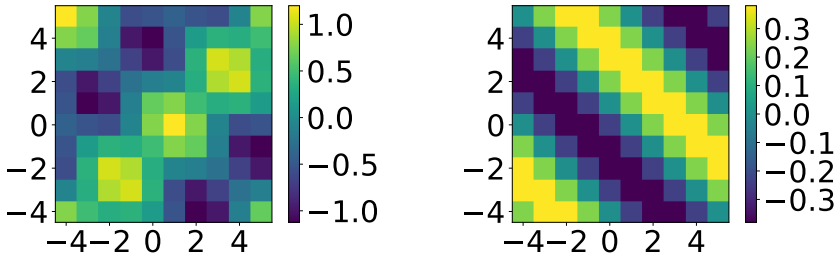


FIGURE 3.3. The horizontal axis corresponds to $k_1 = 0$ and $k_2 = -15, \ldots, 15$. The decay of the Fourier coefficients $\hat{\tilde{g}}_{\mathrm{per}}$ stagnates in the presence of noise, so that the ratio $\hat{\tilde{g}}_{\mathrm{per}}(k)/\hat{\varphi}_{\mathrm{per}}(k)$ is unbounded away from the center.

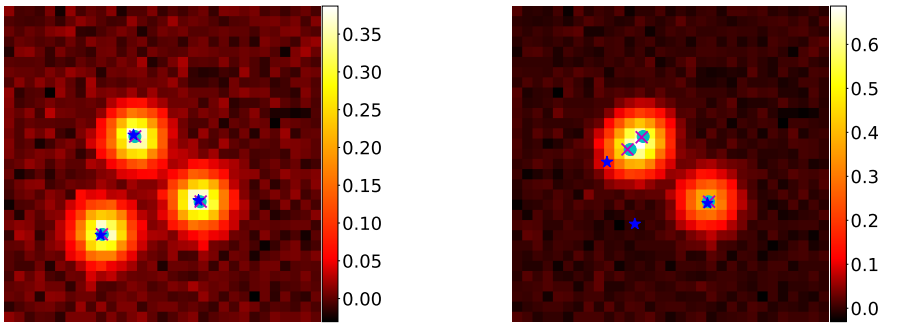


(a) Real part (b) Imaginary part

FIGURE 3.4. $\hat{g}_{\mathrm{per}}(k)/\hat{\varphi}_{\mathrm{per}}(k)$ on $k \in \{-4, \ldots, 5\}^2$.

3.3. **Numerical results on fluorescence microscopy data.** The cell-surface receptor IFNAR2 (type I interferon beta-subunit) of living cells was labelled with biofunctionalized quantum dots (QD605, Cat. No. Q21501MP, Invitrogen [16]). These nanoparticles are small in size (hydrodynamic radius of 15-21 nm) but show an extraordinary high fluorescence signal. Single-molecule imaging was done on an inverted TIRF (total internal reflection fluorescence) microscope (Olympus IX71) with a scientific grade digital camera (Hamamatsu ORCA Flash 4.0). After optical magnification (150xTIRF objective UAPO; NA, 1.45; Olympus) and pixel-binning the final pixel size in the image plane was calculated to be 87 nm. To achieve a high signal-to-noise ratio the signal integration time was set to 32 ms.

The decay of the singular values of T with $n = 4$ for the experimental fluorescence microscopy data in Figure 3.6(a) suggest $M = 8$. This yields $C_\mu, S_1, S_2 \in \mathbb{C}^{8\times 8}$



(a) $\min_{i\neq j} \|z_i - z_j\| = 0.283$: locations are recovered with error margins $\leq 7.1 \cdot 10^{-3}$ and $\leq 2.8 \cdot 10^{-3}$ for $n = 1$ and $n = 4$, respectively.

(b) $\min_{i\neq j} \|z_i - z_j\| = 0.057$: $n = 1$ fails. Locations are correctly recovered for $n = 4$ with error $\leq 1 \cdot 10^{-2}$.

FIGURE 3.5. Noisy synthetic data with SNR = 2.554. The light blue circles show the true locations t_1, t_2, t_3. The blue stars show the reconstruction with $n = 1$, the magenta crosses show the reconstruction with $n = 4$. In accordance with the "spirit" of the requirements on n in Theorem 2.1, well-separated true locations allow for small n. If locations are not well-separated, then $n = 1$ fails but the choice $n = 4$ enables reconstruction.

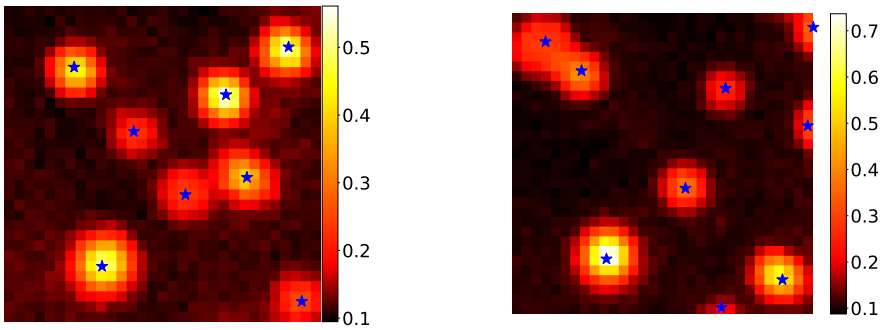
and our algorithm finds the parameters t_j, c_j, $j = 1, \ldots, 8$, in less than a millisecond. Note in Figure 3.6(b) that our algorithm, somewhat surprisingly, successfully identifies proteins at the boundary of the image, even though one would expect artifacts due to periodization issues. However, those identified translations close to the boundary are not very reliable and will need a post- or pre-processing step in a more elaborate analysis in practice.

CONCLUSION

We proposed an algorithm that finds multivariate frequencies out of structured samples of a finite sum of multivariate exponentials. Our proposed algorithm is a multivariate generalization of a matrix pencil method and is based on simultaneous diagonalization of a pencil of non-normal matrices. We also studied a method to simultaneously diagonalize the occurring non-normal matrices by analyzing random linear combinations. Randomness was also quantified in relation to the minimal separation of the exponential parameters. We successfully tested our algorithm on experimental data from fluorescence microscopy.

ACKNOWLEDGEMENTS

The authors have been partially funded by WWTF through project VRG12-009, by DAAD through P.R.I.M.E. 57338904, by FWF project P30148, and by DFG-SFB944.



(a) (b)

FIGURE 3.6. Experimental data with blue stars marking identified locations.

REFERENCES

[1] F. Andersson and M. Carlsson, *ESPRIT for multidimensional general grids* arXiv:1705.07892 (2017).
[2] R. Beinert and G. Plonka, *Sparse phase retrieval of one-dimensional signals by Prony's method*, Frontiers of Applied Mathematics and Statistics **3** (2017), no. 5.
[3] A. Bunse-Gerstner, R. Byers, and V. Mehrmann, *Numerical methods for simultaneous diagonalization*, SIAM J. Matrix Anal. Appl. **14** (1993), no. 4, 927–949.
[4] J.-F. Cardoso and A. Souloumiac, *Jacobi angles for simultaneous diagonalization*, SIAM J. Matrix Anal. Appl. **17** (1996), no. 1, 161–164.
[5] G. H. Golub and C. F. Van Loan, *Matrix computations*, Johns Hopkins Studies in the Mathematical Sciences, The Johns Hopkins University Press, 1996.
[6] D. Kressner, *Numerical methods for general and structured eigenvalue problems*, Springer, 2005.
[7] S. Kunis, H. M. Möller, T. Peter, and U. von der Ohe, *Prony's method under an almost sharp multivariate Ingham inequality*, J. Fourier Anal. Appl. **to appear** (2018), 1–13.
[8] S. Kunis, T. Peter, T. Römer, and U. von der Ohe, *A multivariate generalization of Prony's method*, Lin. Alg. Appl. **490** (2016), 31–47.
[9] S. Li, *Concise formulas for the area and volume of a hyperspherical cap*, Asian J. Math. Stat. **4** (2011), no. 1, 66–70.
[10] B. Mourrain, *Polynomial–Exponential Decomposition From Moments*, Found. Comput. Math. **to appear** (2018), 1–58.
[11] T. Peter and G. Plonka, *A generalized Prony method for reconstruction of sparse sums of eigenfunctions of linear operators*, Inverse Problems **29** (2013), no. 2, 025001.
[12] T. Peter, D. Potts, and M. Tasche, *Nonlinear approximation by sums of exponentials and translates*, SIAM J. Sci. Comput., **33** (2011), no. 4, 1920–1947.
[13] D. Potts and M. Tasche, *Parameter estimation for exponential sums by approximate Prony method*, Signal Process. **90** (2010), 1631–1642.
[14] ———, *Parameter estimation for nonincreasing exponential sums by Prony-like methods*, Linear Algebra Appl. **439** (2013), no. 4, 1024-1039.
[15] V. Studer, J. Bobinc, M. Chahid, H. S. Mousavi, E. Candès, and M. Dahan, *Compressive fluorescence microscopy for biological and hyperspectral imaging*, Proc. Nat. Acad. Sci. **109** (2012), no. 26, E1679–E1687.
[16] C. You, S. Wilmes, C. P. Richter, O. Beutel, D. Liße, and J. Piehler, *Electrostatically Controlled Quantum Dot Monofunctionalization for Interrogating the Dynamics of Protein Complexes in Living Cells*, ACS Chemical Biology **8** (2013), no. 2, 320–326.

(M. Ehler) University of Vienna, Department of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna
E-mail address: martin.ehler@univie.ac.at

(S. Kunis) Universität Osnabrück, Institut für Mathematik, Albrechtstr. 28a, D-49076 Osnabrück
E-mail address: stefan.kunis@uni-osnabrueck.de

(T. Peter) University of Vienna, Department of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna
E-mail address: petert53@univie.ac.at

(C. Richter) Universität Osnabrück, Institut für Biologie, Barbarastr. 11, D-49076 Osnabrück
E-mail address: christian.richter@biologie.uni-osnabrueck.de