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.

Key words. frequency analysis, spectral analysis, exponential sum, moment problem, superresolution

AMS subject classifications. 65T40, 42C15, 30E05, 65F30

1. Introduction. Many imaging and data analysis problems in the applied sciences lead to the numerical task of parameter identification in exponential sums $\sum_{j=1}^{M} c_j e^{-2\pi i \langle t_j, \cdot \rangle}$. 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. The univariate case $d = 1$ is well understood (see, e.g., [18, 20] and the references therein), and the most feasible implementations are based on computing the zeros of the polynomial in the kernel of the associated moment matrix [2, 16] or on the classical matrix pencil method [9].

A multivariate generalization of Prony's method can be found in [11, 12, 22] and consists of a two-step procedure: 1) compute the kernel of a multivariate moment matrix and 2) compute the common zeros of the polynomials associated to the kernel. General purpose algorithms for the second step, in particular symbolic computations, are computationally quite expensive. Popular projection-based ideas reduce the multivariate parameter identification problem to a series of univariate problems (see, e.g., [5, 6, 19] and the references therein) and ask for sampling the exponential sum on specific lines—data which might not be available in applications.

In contrast, the matrix pencil method [9] has recently been generalized to the multivariate setting [1, 8, 15, 21] and realizes the sought parameters as rank reducing numbers of d singular matrix pencils. This approach constructs 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 the simultaneous diagonalization are not available; cf. [3, 4, 7, 10]. 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 algorithm [15, Alg. 3.1] and the algorithm introduced in [21] also use a random linear combination and argue that generically the eigenvalues are simple. Moreover, in [21] the authors analyze the influence of perturbations on their multivariate ESPRIT-method. Here, we describe the situation of using a random linear combination of $S_1, \ldots, S_d$ in more detail and quantify the influence of the

∗Received on June 18, 2018. Accepted January 30, 2019. Published online on March 19, 2019. Recommended by Daniel Potts.

†University of Vienna, Department of Mathematics, Oskar-Morgenstern-Platz 1, A-1090 Vienna
{martin.ehler, thomas.peter}@univie.ac.at.

‡Osnabrück University, Institute for Mathematics, Albrechtstr. 28a, D-49076 Osnabrück.

¶Osnabrück University, Research Center of Cellular Nanoanalytics, Barbarastr. 11, D-49076 Osnabrück
stefan.kunis@uni-osnabrueck.de.

§Osnabrück University, Institute for Biology, Barbarastr. 11, D-49076 Osnabrück
christian.richter@biologie.uni-osnabrueck.de.

minimal separation of the $\{t_j\}_{j=1}^M$ on the eigendecomposition of the random matrix. Similar to the stability analysis of the univariate matrix pencil method in [14], we expect our analysis to be useful in analysing the stability of the randomized multivariate matrix pencil method.

To verify 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. [23]. The results on experimental fluorescence images show that our scheme is numerically feasible.

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

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

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. Let $I_n := \{0, \ldots, n\}^d$ for $n \in \mathbb{N}$, 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

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

For $d = 1$, the classical matrix pencil method [9] realizes the nodes $\mathrm{e}^{-2\pi\mathrm{i}t_j}$ as rank reducing numbers of the singular matrix pencil $zT - T_1$. Similar to the recent approaches [1, 15, 21], we directly generalize this for $d > 1$ to a simultaneous diagonalization problem as follows. 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 a diagonal matrix with positive entries on its diagonal and $U, V \in \mathbb{C}^{N\times M}$ 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

$$(2.2) \qquad S_\ell := U^*T_\ell V\Sigma^{-1}, \qquad \ell = 1, \ldots, d,$$

which were recently also discussed in [8, Prop. 4.1]. These matrices turn out to be simultaneous diagonalizable, cf. Theorem 2.1, which shall enable us to identify the vectors $\{t_j\}_{j=1}^M$. We also make use of

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

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 $n \geq 2\pi d/\min_{i \neq j} \|z_i - z_j\|$, then T has rank M and $S_1, \ldots, S_d$ are simultaneously diagonalizable. Furthermore, any regular matrix W that simultaneously diagonalizes $S_1, \ldots, S_d$ yields a permutation τ on $\{1, \ldots, M\}$ such that*

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

Proof. According to [12], the matrix T always admits the factorization

$$T = A^* D A,$$

with the multivariate complex Vandermonde matrix

$$A = \left(z_j^k \right)_{\substack{j=1,\ldots,M \\ k \in I_n}} \in \mathbb{C}^{M \times N}$$

and $D = \mathrm{diag}(c_1, \ldots, c_M)$. The condition on n implies that A has full rank M; see [11, Cor. 2.5] also for possible improvements on the factor $2\pi d$. Together with $c_1, \ldots, c_M \neq 0$, this yields T having rank M. We also deduce the factorization

$$T_\ell = A^* D_\ell A, \qquad \ell = 1, \ldots, d,$$

where the diagonal matrix D_ℓ is given by

$$D_\ell := \mathrm{diag}(c_1 \langle z_1, e_\ell \rangle, \ldots, c_M \langle z_M, e_\ell \rangle), \qquad \ell = 1, \ldots, d.$$

We shall now verify that the specific matrix $W_0 := (AU)^*$ (which is not accessible to us) simultaneously diagonalizes $S_1, \ldots, S_d$. 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

$$W_0^{-1} S_\ell W_0 = D_\ell D^{-1} = \mathrm{diag}(\langle z_1, e_\ell \rangle, \ldots, \langle z_M, e_\ell \rangle), \qquad \ell = 1, \ldots, d,$$

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

$$(2.3) \qquad C_\mu := \sum_{\ell=1}^{d} \bar{\mu}_\ell S_\ell, \qquad \mu = (\mu_1, \ldots, \mu_d)^\top \in \mathbb{C}^d.$$

Because of

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

the eigenvalues $\lambda_1(\mu), \ldots, \lambda_M(\mu)$ 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 distinct, so are $\{z_j\}_{j=1}^{M}$. 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 $i \neq j$. Thus, $\{\lambda_j(\tilde{\mu})\}_{j=1}^{M}$ are pairwise distinct. In other words, all eigenspaces of $C_{\tilde{\mu}}$ are one-dimensional.

Any matrix $W = (w_1, \ldots, w_M)$ that simultaneously diagonalizes $S_1, \ldots, S_d$ 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. $\square$

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 the simultaneous diagonalization to ensure that these entries are associated to the same permutation across all $\ell = 1, \ldots, d$.

In general, the matrices $S_1, \ldots, S_d$ 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, 7, 10]. We attempt to circumvent such issues by using C_μ from (2.3), 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 $\lambda_1(\mu), \ldots, \lambda_M(\mu)$ are pairwise different, then any matrix W that diagonalizes C_μ also simultaneously diagonalizes $S_1, \ldots, S_d$.*

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

According to Corollary 2.2 we aim to find $\mu \in \mathbb{C}^d$ such that $\lambda_1(\mu), \ldots, \lambda_M(\mu)$ 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

$$(2.4) \qquad \left\{ \mu \in \mathbb{C}^d : \lambda_1(\mu), \ldots, \lambda_M(\mu) \text{ are pairwise different} \right\} = \mathbb{C}^d \setminus \bigcup_{i \neq j} (z_i - z_j)^\perp.$$

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 $t_1, \ldots, t_5 \in [0,1)^2$ randomly. We construct $S_1, S_2 \in \mathbb{C}^{5 \times 5}$ by (2.2). Thus, we choose $\mu = (\mu_1, \mu_2)^\top \in \mathbb{S}_{\mathbb{C}}^1$ and construct $C_\mu = \mu_1 S_1 + \mu_2 S_2$. According to (2.4) we expect $\binom{5}{2} = 10$ great circles on $\mathbb{S}_{\mathbb{C}}^1$ 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. Therefore we can use Hopf fibration, to identify great circles on $\mathbb{S}_{\mathbb{C}}^1$ 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 Figure 2.1(a), the other 2 being on the backside 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 the simultaneous diagonalization of $S_1, \ldots, S_d$ suggested in Corollary 2.2 requires our present setting, in which $\{z_j\}_{j=1}^M$ are pairwise different. Since it does not apply to the problem of simultaneous diagonalization in general, we refer to [15] for the case of nodes with multiplicities.

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 $z_i - z_j$. If $\mu \in \mathbb{S}_{\mathbb{C}}^{d-1}$ is a random vector distributed according to the unitarily invariant probability measure on $\mathbb{S}_{\mathbb{C}}^{d-1}$, then

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

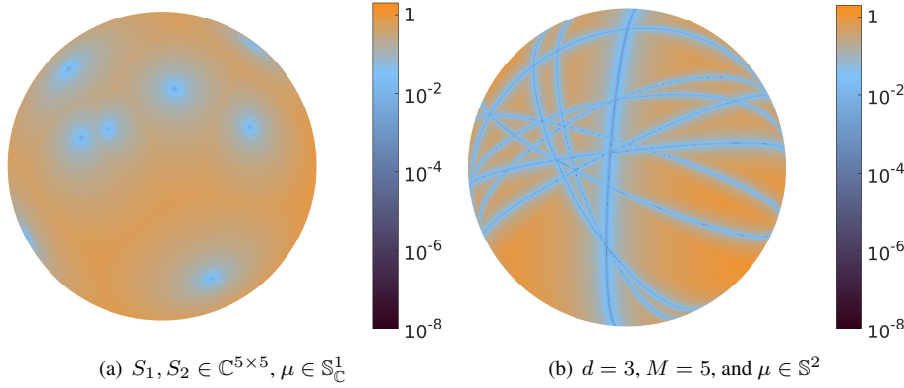The following result provides a more quantitative analysis:

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

FIG. 2.1. *Visualization of the smallest distance of any two eigenvalues of C_μ.*

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

(2.5)
$$|\lambda_i(\mu) - \lambda_j(\mu)| < \epsilon \|z_i - z_j\|$$

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

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

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

$$\left| \left\langle \begin{pmatrix} \mathrm{Re}(y) \\ \mathrm{Im}(y) \end{pmatrix}, \begin{pmatrix} \mathrm{Re}(x) \\ \mathrm{Im}(x) \end{pmatrix} \right\rangle \right| = |\mathrm{Re}(\langle y, \mu \rangle)| \leq |\langle y, \mu \rangle|,$$

we obtain an upper bound by simply considering

(2.6)
$$\left| \left\langle \begin{pmatrix} \mathrm{Re}(y) \\ \mathrm{Im}(y) \end{pmatrix}, \begin{pmatrix} \mathrm{Re}(\mu) \\ \mathrm{Im}(\mu) \end{pmatrix} \right\rangle \right| \leq \epsilon.$$

Due to the orthogonal invariance of the surface measure on $\mathbb{S}^{2d-1}$, the distribution of the left-hand side in (2.6) does not depend on the special choice of $y \in \mathbb{S}_{\mathbb{C}}^{d-1}$, so that we can simply assume that $\begin{pmatrix} \mathrm{Re}(y) \\ \mathrm{Im}(y) \end{pmatrix}$ is the north pole. The inequality (2.6) reduces to $-\epsilon \leq \mathrm{Re}(\mu_1) \leq \epsilon$, hence, it 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, [13], 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

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

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

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

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

REMARK 2.6. Theorem 2.5 immediately implies that the probability that any of the inequalities

$$|\lambda_i(\mu) - \lambda_j(\mu)| \geq \epsilon\|z_i - z_j\|, \qquad \forall i \neq j,$$

is violated is at most $\binom{M}{2}2\sqrt{\frac{d}{\pi}}\epsilon$. In other words, if we select about M^2 independent μ, then the probability that (2.5) fails is at most of the order ϵ.

Moreover, a short calculation leads to

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

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 the least-squares method applied to the linear system $T = A^*DA$, for $D = \mathrm{diag}(c_{\tau(1)}, \ldots, c_{\tau(M)})$. 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. [23], so that the considered probe is a tempered distribution

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

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

$$(3.1) \qquad \mathcal{F}(G) = \sum_{j=1}^{M} c_j e^{-2\pi i \langle t_j, \cdot \rangle}.$$

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, then 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

$$(3.2) \qquad \mathcal{F}(G)(k) = \mathcal{F}(g)(k)/\mathcal{F}(\varphi)(k), \qquad k \in I.$$

Combining (3.1) with (3.2) leads to the sampling problem (2.1) discussed in the previous sections, i.e.,

$$(3.3) \qquad \sum_{j=1}^{M} c_j e^{-2\pi i \langle t_j, k \rangle} = f(k), \qquad k \in I,$$

with $f(k) := \mathcal{F}(g)(k)/\mathcal{F}(\varphi)(k)$. In principle, the parameters $\{t_j\}_{j=1}^{M}$ and $\{c_j\}_{j=1}^{M}$ can now be determined by Algorithm 1. Note that the above derivations in this section have also been used in [17] 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 (3.3) is not readily available. Aiming at the application of the discrete Fourier transform (DFT), we recognize that a sufficient decay of φ implies $g \in L^1(\mathbb{R}^d)$, so that its periodization

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

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

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

Thus, (3.3) can be evaluated by first computing the periodization g_{per}, so that its Fourier coefficients yield

(3.4)
$$\sum_{j=1}^{M} c_j \mathrm{e}^{-2\pi\mathrm{i}\langle t_j, k \rangle} = \hat{g}_{\mathrm{per}}(k)/\mathcal{F}(\varphi)(k), \qquad k \in I.$$

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. Python codes are available at t-peter.com.

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 of 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 an adjusted parameter b derived from the camera system. Therefore, the values $\mathcal{F}(\varphi)$ are 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 at first exact and in a second experiment corrupted by additive Gaussian noise with a signal-to-noise ratio of $\mathrm{SNR} = 2.554$; 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 (3.4). 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 that we gain by making our computations in the Fourier domain. Figure 3.2 shows the difference between the 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 3.1(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.
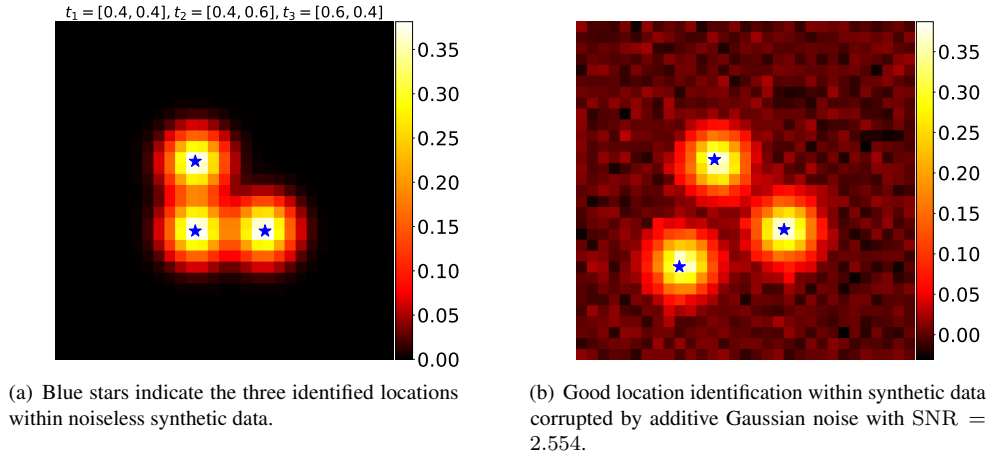
(a) Blue stars indicate the three identified locations within noiseless synthetic data.

(b) Good location identification within synthetic data corrupted by additive Gaussian noise with $\mathrm{SNR} = 2.554$.

FIG. 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.*
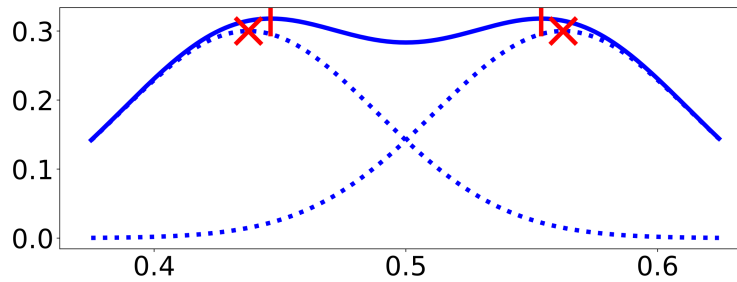

FIG. 3.2. *The red crosses show the true location of $t_1 = 0.44, t_2 = 0.56$ of two one-dimensional Gaussians, each depicted as a dotted line. The red bars, however, show the local maxima of the sum of these Gaussians, and this sum is shown in a continuous line.*

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 $k_2 = -15, \ldots, 15$. Even though we are dealing with images of size 31×31 pixels, the frequency data of the noisy ratio $\hat{\tilde{g}}_{\mathrm{per}}(k)/\hat{\varphi}_{\mathrm{per}}(k)$ seems reliable only 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 in our fluorescence microscopy data. Figure 3.4 shows the ratios $\hat{g}_{\mathrm{per}}(k)/\hat{\varphi}(k)$ 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.
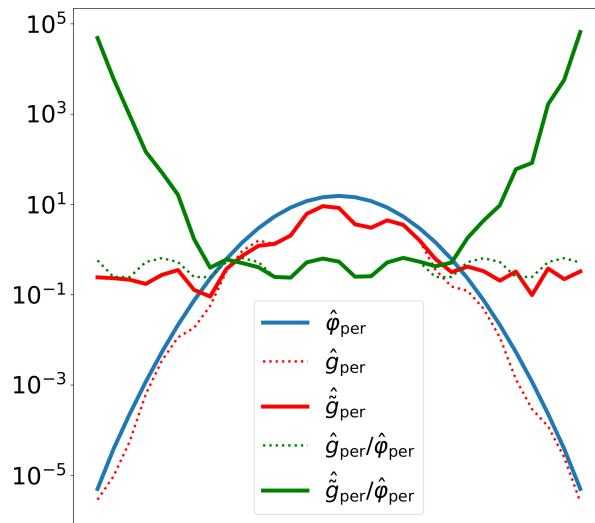
72 M. EHLER, S. KUNIS, T. PETER, AND C. RICHTER

FIG. 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.*
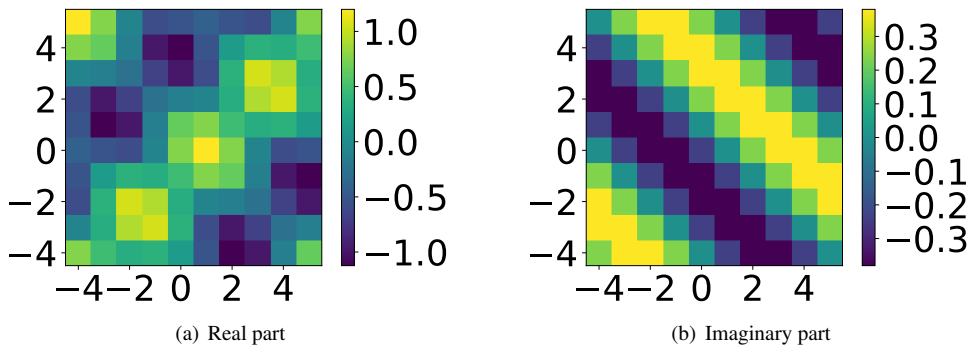

(a) Real part

(b) Imaginary part

FIG. 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 labeled with biofunctionalized quantum dots (QD605, Cat. No. Q21501MP, Invitrogen [24]). 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) suggests $M = 8$. This yields $C_\mu, S_1, S_2 \in \mathbb{C}^{8 \times 8}$, and our algorithm finds the parameters $t_j, c_j, j = 1, \ldots, 8$, in less than a millisecond. Note that in Figure 3.6(b), our algorithm, somewhat surprisingly, successfully identifies proteins at the
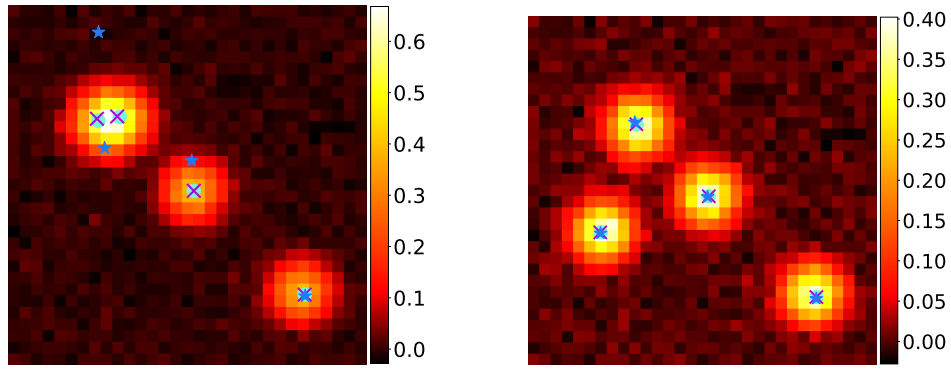
(a) $\min_{i \neq j} \|z_i - z_j\| = 0.283$: locations are recovered with error margins $\leq 8 \cdot 10^{-3}$ and $\leq 1.6 \cdot 10^{-3}$ for $n = 1$ and $n = 3$, respectively.

(b) $\min_{i \neq j} \|z_i - z_j\| = 0.059$: $n = 1$ fails. Locations are correctly recovered for $n = 3$ with error $\leq 7.1 \cdot 10^{-3}$.

FIG. 3.5. *Noisy synthetic data with* $\mathrm{SNR} = 2.554$. *The light blue circles show the true locations* t_1, t_2, t_3, t_4. *The blue stars show the reconstruction with* $n = 1$, *the magenta crosses show the reconstruction with* $n = 3$. *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* $n = 3 < M = 4$ *enables reconstruction.*
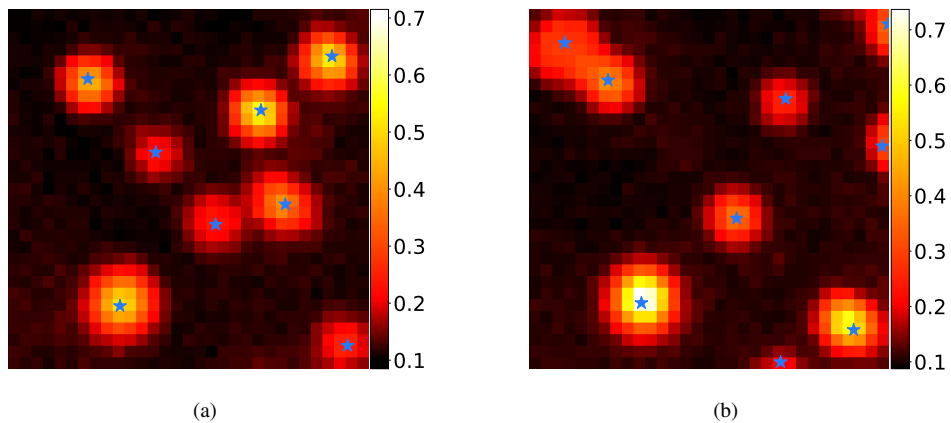

(a)

(b)

FIG. 3.6. *Experimental data with blue stars marking identified locations.*

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 a post- or pre-processing step in a more elaborate analysis will be needed 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.

REFERENCES

[1] F. ANDERSSON AND M. CARLSSON, *ESPRIT for multidimensional general grids*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 1470–1488.

[2] R. BEINERT AND G. PLONKA, *Sparse phase retrieval of one-dimensional signals by Prony's method*, Front. Appl. Math. Stat., 3 (2017), 5 pages.

[3] A. BUNSE-GERSTNER, R. BYERS, AND V. MEHRMANN, *Numerical methods for simultaneous diagonalization*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 927–949.

[4] J.-F. CARDOSO AND A. SOULOUMIAC, *Jacobi angles for simultaneous diagonalization*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 161–164.

[5] A. CUYT AND W.-S. LEE, *Multivariate exponential analysis from the minimal number of samples*, Adv. Comput. Math., 44 (2018), pp. 987–1002.

[6] B. DIEDERICHS AND A. ISKE, *Projection-based multivariate frequency estimation*, in Sampling Theory and Applications (SampTA), 2017 I, IEEE Conference Proceedings, Los Alamitos, 2017, pp. 360–363.

[7] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.

[8] J. HARMOUCH, H. KHALIL, AND B. MOURRAIN, *Structured low rank decomposition of multivariate Hankel matrices*, Linear Algebra Appl., 542 (2018), pp. 162–185.

[9] Y. HUA AND T. K. SARKAR, *Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise*, IEEE Trans. Acoust. Speech Signal Process., 38 (1990), pp. 814–824.

[10] D. KRESSNER, *Numerical Methods and Software for General and Structured Eigenvalue Problems*, Springer, Berlin, 2005.

[11] 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., 24 (2018), pp. 1306–1318.

[12] S. KUNIS, T. PETER, T. RÖMER, AND U. VON DER OHE, *A multivariate generalization of Prony's method*, Linear Algebra Appl., 490 (2016), pp. 31–47.

[13] S. LI, *Concise formulas for the area and volume of a hyperspherical cap*, Asian J. Math. Stat., 4 (2011), pp. 66–70.

[14] A. MOITRA, *Super-resolution, extremal functions and the condition number of Vandermonde matrices*, in STOC'15-Proceedings of the 2015 ACM Symposium on Theory of Computing, ACM, New York, 2015, pp. 821–830.

[15] B. MOURRAIN, *Polynomial-exponential decomposition from moments*, Found. Comput. Math., 18 (2018), pp. 1435–1492.

[16] T. PETER AND G. PLONKA, *A generalized Prony method for reconstruction of sparse sums of eigenfunctions of linear operators*, Inverse Problems, 29 (2013), Art. 025001, 21 pages.

[17] T. PETER, D. POTTS, AND M. TASCHE, *Nonlinear approximation by sums of exponentials and translates*, SIAM J. Sci. Comput., 33 (2011), pp. 1920–1947.

[18] D. POTTS AND M. TASCHE, *Parameter estimation for exponential sums by approximate Prony method*, Signal Process., 90 (2010), pp. 1631–1642.

[19] ———, *Parameter estimation for multivariate exponential sums*, Electron. Trans. Numer. Anal., 40 (2013), pp. 204–224.
    http://etna.ricam.oeaw.ac.at/vol.40.2013/pp204-224.dir/pp204-224.pdf

[20] ———, *Parameter estimation for nonincreasing exponential sums by Prony-like methods*, Linear Algebra Appl., 439 (2013), pp. 1024–1039.

[21] S. SAHNOUN, K. USEVICH, AND P. COMON, *Multidimensional ESPRIT for damped and undamped signals: algorithm, computations, and perturbation analysis*, IEEE Trans. Signal Process., 65 (2017), pp. 5897–5910.

[22] T. SAUER, *Prony's method in several variables*, Numer. Math., 136 (2017), pp. 411–438.

[23] V. STUDER, J. BOBIN, M. CHAHID, H. S. MOUSAVI, E. CANDES, AND M. DAHAN, *Compressive fluorescence microscopy for biological and hyperspectral imaging*, Proc. Natl. Acad. Sci. USA, 109 (2012), pp. E1679–E1687.

[24] 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 Chem. Bio., 8 (2012), pp. 320–326.