Soft thresholding schemes for multiple signal classification algorithm

SEBASTIAN ACUÑA,[1,*] **IDA S. OPSTAD,**[1] **FRED GODTLIEBSEN,**[1] **BALPREET SINGH AHLUWALIA,**[1,2] **AND KRISHNA AGARWAL**[1]

[1]*Department of Physics and Technology, UiT The Arctic University of Norway, NO-9037 Tromsø, Norway*
[2]*Department of Clinical Science, Intervention & Technology, Karolinska Institute, 17177 Stockholm, Sweden*
*sebastian.acuna@uit.no

Abstract: Multiple signal classification algorithm (MUSICAL) exploits temporal fluctuations in fluorescence intensity to perform super-resolution microscopy by computing the value of a super-resolving indicator function across a fine sample grid. A key step in the algorithm is the separation of the measurements into signal and noise subspaces, based on a single user-specified parameter called the threshold. The resulting image is strongly sensitive to this parameter and the subjectivity arising from multiple practical factors makes it difficult to determine the right rule of selection. We address this issue by proposing soft thresholding schemes derived from a new generalized framework for indicator function design. We show that the new schemes significantly alleviate the subjectivity and sensitivity of hard thresholding while retaining the super-resolution ability. We also evaluate the trade-off between resolution and contrast and the out-of-focus light rejection using the various indicator functions. Through this, we create significant new insights into the use and further optimization of MUSICAL for a wide range of practical scenarios.

1. Introduction

Conventional optical microscopy is limited in resolution due to diffraction of light. The need to overcome this limit has given rise to super-resolution microscopy techniques, also called optical nanoscopy. Among these techniques, structured illumination microscopy (SIM) [1] allows a lateral resolution enhancement by a factor of 2 over the optical resolution limit, stimulated emission depletion (STED) microscopy [2] and single molecule localization (SML) [3–5] can achieve resolutions close to 20 nm, and MINFLUX [6] which combines concepts of SML and STED to achieve even 2 nm resolution. However, these techniques require expensive and complex setups as in SIM and STED, or a high amount of light dose and long acquisition time in SML.

A new family of purely computational nanoscopy techniques has also emerged, where statistical analysis of spatio-temporal fluctuations of fluorescence intensity arising from the photokinetic properties of fluorescent molecules is used for super-resolution. Examples include super-resolution optical fluctuation imaging (SOFI) [7], super-resolution radial fluctuations (SRRF) [8], multiple signal classification algorithm (MUSICAL) [9], entropy-based super-resolution imaging (ESI) [10], 3B [11], spatial covariance reconstructive (SCORE) super-resolution fluorescence microscopy [12], super-resolution with auto-correlation two-step deconvolution (SACD) [13] and sparsity-based super-resolution correlation microscopy (SPARCOM) [14].

Most of these methods assume that the emitters are stationary and that the photokinetic properties do not change during the image acquisition. MUSICAL, on the contrary, simply exploits the spatio-temporal variations as acquired on a temporal stack of fluorescence microscopy images, irrespective of whether they arise from photokinetics or movement of fluorophores. It decomposes the stack into a set of vectors called eigenimages (details in section 2.2) and separates them into two orthogonal spaces called the signal and the noise subspaces. These are used into a special function called the *indicator function* that exploits the fact that the spatial-temporal

distribution of fluorophores is encoded in eigenimages through the point spread function (PSF) of the microscope. It is designed to have a high value at an emitter location and lower otherwise, which enables super-resolution when applied to a grid finer than the original microscopy image.

MUSICAL uses three user-specified control parameters that have a bearing on the reconstructed nanoscopy image: a parameter α determines the effective contrast, a sub-pixelation parameter determines the size of the finer grid, and a threshold parameter determines how the eigenimages are split between the signal and noise subspaces. While α and sub-pixelation prominently determine the aesthetic of the nanoscopy image, the threshold parameter has a significant role in the nature and scale of details that get reconstructed. The problem is that its selection depends on multiple factors such as fluorophore's density and photokinetics, sample's geometry, out of focus light, signal to noise ratio, etc. in a complicated and non-obvious manner. Even though some rules of thumb for threshold selection have been reported [9,15,16], it still remains as the most non-intuitive parameter of MUSICAL as its value is subjective and depends mostly on the user's experience. The problem of threshold selection is illustrated in Fig. 1, where for the same sample (Fig. 1(a)), a single value does not produce the best possible reconstruction for the entire sample. This figure also illustrates the process of threshold selection, which starts with a plot similar to Fig. 1(b) from where the user selects a value such as −0.5 (point A) or −0.3 (point B). The name of the points come from the rules used to pick them (more details in section 2.4) and as it can be seen from the results on region 1(Fig. 1(c)) and 2 (Fig. 1(d)), the results vary between regions. For example, if we consider the indicated ring in each case, for region 1 the ring is only recovered for value B, while is reconstructed better for value A in region 2 as the ring appears clearer.
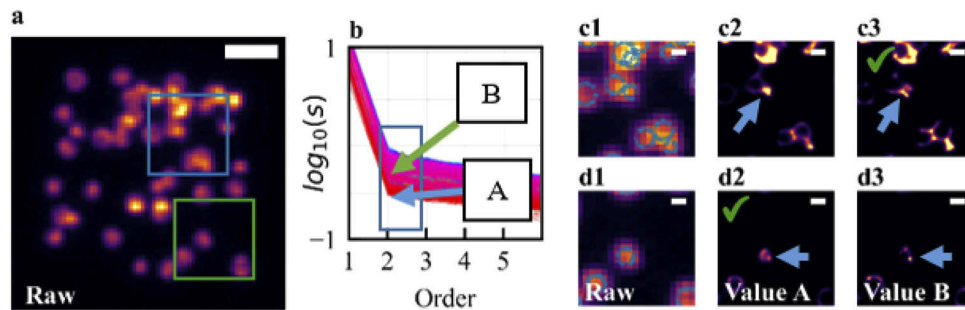

Fig. 1. Effect of MUSICAL's threshold on image reconstruction. A sample consisting of rings in the focal plane allows to see the effect of two different thresholds in the final reconstruction (a green tick marks the best one). The sample consists of 500 frames. (a) Mean image with 2 regions marked with rectangles. Scale bar 1μm. (b) Plot of the singular values [15] used by the user to pick a threshold. (c),(d) MUSICAL results in regions 1 and 2, respectively. Scale bars 200 nm. (c1),(d1) show the mean images with the actual emitters on top in blue color.

This work addresses the problem of the subjectivity of threshold selection and the sensitivity of MUSICAL's reconstructions to it. This is achieved in the following manner:

- We have scrutinized the effect of threshold selection and identified the root cause of threshold sensitivity.

- We have proposed a generalized form of indicator function design that carry two new families of indicator functions.

- We have evaluated quantitatively and qualitatively the comparative advantages of the new indicator functions while generating new insights into practically useful properties such as resolution-contrast trade-offs, out-of-focus light rejection, and dynamic range utilization.

The outline of the paper is as follows. section 2 presents the theory behind MUSICAL and the role of the threshold in image reconstruction. The new generalized framework for the MUSICAL indicator function design is presented in section 3 section 4 presents the results and insights on a variety of simulated and experimental data. The conclusions are summarized in section 5.

2. Background

2.1. Imaging model

Let's consider a sample composed of point-like blinking photon emitters, i.e. individual fluorophore molecules. In this case, blinking means that the number of photons emitted by each particle varies over time, owing to the photokinetics of the fluorophores [17], with no assumption on temporal sparsity. A stack of images taken of such sample from a diffraction-limited system over T time-steps is expressed in a matrix form as $\mathbf{A} = [\mathbf{a}(1)\dots\mathbf{a}(T)]$, where each column vector $\mathbf{a}(t)$ contains the intensity measured by a set of sensing elements (i.e. camera pixels) at time step t. For N emitters, a single image $\mathbf{a}(t)$ can be modelled as follows:

$$\mathbf{a}(t) = \sum_{n=1}^{N} \mathbf{g}(\mathbf{r}_n(t))s_n(t). \tag{1}$$

In this model, $\mathbf{g}(\mathbf{r})$ represents the PSF of the microscope and corresponds to the intensity produced by an emitter located at $\mathbf{r}$ as measured on the sensor. Finally, $s_n(t)$ is the number of photons produced by the nth emitter during the time step t, which is a random variable resulting from the photokinetics of the fluorophore. For simplicity, we here only consider stationary emitters, $r_n(t) = r_n$. Then, Eq. (1) can be written as the matrix equation $\mathbf{a}(t) = \mathbf{G}\mathbf{s}(t)$ by constructing two new matrices $\mathbf{G} = [\mathbf{g}(\mathbf{r}_1)\dots\mathbf{g}(\mathbf{r}_N)]$ and $\mathbf{s}(t) = [s_1(t)\dots s_N(t)]^T$. Therefore, every single image is a linear combination of the columns of $\mathbf{G}$, weighted by the photon emissions of emitters.

2.2. Key concept of MUSICAL

MUSICAL involves a sliding window operation, with window defined as a crop of the image stack. Processing a single window returns a super-resolved version of it, and the final MUSICAL reconstruction is built by overlaying and stitching all super-resolved windows together. The size of the window corresponds to the approximate size of the main lobe of the PSF, which is estimated from the wavelength of emission and numerical aperture of the imaging system. While performing nanoscopy using MUSICAL on a single window, the observed data is decomposed into two orthogonal subspaces. In this paper, we use notation associated with singular value decomposition such as used in [9]. Accordingly, the matrix $\mathbf{A}$ is decomposed as $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$, where $\mathbf{U}$ contains the basis vectors $\mathbf{u}_i$. These vectors are called eigenimages and contain the spatial information of the sample, while their corresponding singular values σ_i in the diagonal of matrix $\mathbf{S}$ is a measure of its statistical significance. Since matrix $\mathbf{V}$ is not used, MUSICAL can benefit of the relation between SVD and Eigenvalue decomposition as $\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ (as shown in [18] for computational efficiency), where $\mathbf{\Lambda}$ contains the eigenvalues λ_i instead of the singular values. Since both values are related ($\lambda_i = \sigma_i^2$), the discussion below can be generally applicable irrespective of whether singular value or eigenvalue decomposition is used. The two key concepts of MUSICAL are presented below.

Key concept 1 (KC1): MUSICAL separates the basis $\mathbf{U}$ into two subspaces, namely the signal subspace $\mathcal{S}$ and the orthogonal complementary noise subspace $\mathcal{N}$ using a threshold σ_0 such that:

$$\mathbf{u}_i \in \begin{cases} \mathcal{S} & \text{if } \sigma_i \geq \sigma_0 \\ \mathcal{N} & \text{otherwise.} \end{cases} \tag{2}$$

Key concept 2 (KC2): If the threshold is chosen such that the subspace $\mathcal{N}$ indeed contains contribution from only the noise, then the subspace $\mathbf{S}$ contains $\mathbf{g}(\mathbf{r}_n)$ for every emitter and $\mathcal{N}$ is

devoid of them. Consequently, using the orthogonality of $\mathcal{S}$ and $\mathcal{N}$, $\mathbf{g}(\mathbf{r}_n)$ are also orthogonal to $\mathcal{N}$. This property is used to design the indicator function $f(\mathbf{r})$, discussed in detail in section 2.3.

In the ideal case, the matrix $\mathbf{A}$ is rank deficient (i.e., it contains some zero singular values), which happens when the number of emitters is smaller than the number of sensor elements. In such case, the threshold is simply $\sigma_0 = 0$ and the signal subspace $\mathcal{S}$ is formed by the eigenimages with non-zero singular values while the noise subspace $\mathcal{N}$ by the ones with value zero. However, real samples are composed of a large number of emitters such that the matrix $\mathbf{A}$ is full ranked. In addition, real microscopy data contains noise, such that none of the singular values are strictly zero. Therefore, σ_0 is chosen on a case-by-case basis. Intuitively, the eigenvectors associated with large eigenvalues represent the more statistically prominent structures in the stack as compared to the other eigenvectors.

The key concept 2: Since the space spanned by $\mathbf{g}(\mathbf{r}_i)$ and the eigenimages of the signal space are the same, the noise space is also orthogonal to $\mathbf{g}(\mathbf{r}_i)$, i.e. when the PSF is evaluated at an emitter's location. This property is exploited in designing indicator function that indicates the presence of emitters.

2.3. Indicator function

Let us consider an arbitrary test point $\mathbf{r}_{test}$ in the sample region. Its image is given by the vector $\mathbf{g}(\mathbf{r})$, which can be represented as $\mathbf{g}(\mathbf{r}) = \mathbf{g}_{\mathcal{S}}(\mathbf{r}) + \mathbf{g}_{\mathcal{N}}(\mathbf{r})$, where $\mathbf{g}_{\mathcal{S}}(\mathbf{r})$ and $\mathbf{g}_{\mathcal{N}}(\mathbf{r})$ are the projections of $\mathbf{g}(\mathbf{r})$ onto the signal and noise spaces. The magnitudes of these projections are computed as:

$$||\mathbf{g}_{\mathcal{S}}(\mathbf{r}_{test})|| = \sqrt{\sum_{\mathbf{u}_i \in \mathcal{S}} g_i{}^2} \quad ; \quad ||\mathbf{g}_{\mathcal{N}}(\mathbf{r}_{test})|| = \sqrt{\sum_{\mathbf{u}_i \in \mathcal{N}} g_i{}^2} \quad \text{where} \quad g_i = |\mathbf{g}(\mathbf{r}_{test}) \cdot \mathbf{u}_i|. \quad (3)$$

Using KC2 and the projections in Eq. (3), MUSICAL constructs the following indicator function:

$$f(\mathbf{r}_{test}) = \left(\frac{||\mathbf{g}_{\mathcal{S}}(\mathbf{r}_{test})||}{||\mathbf{g}_{\mathcal{N}}(\mathbf{r}_{test})||} \right)^{\alpha}. \quad (4)$$

This indicator function generates a large value when $\mathbf{r}_{test}$ is at the location of the emitters, since $||\mathbf{g}_{\mathcal{N}}(\mathbf{r}_{test})||$ is zero (KC2). For this it is important that the threshold value σ_0 for defining the signal and noise subspaces is chosen appropriately.

2.4. Threshold selection and its associated problem

The current practice of threshold selection starts with a plot of the singular values of the microscopy image stack, such as shown in Fig. 1(b). This plot is made by computing the singular values for each window and then plotting them as lines in logarithmic scale, with the x-axis showing the order when sorted decreasingly. Then, the user specifies a threshold σ_0 based on observations derived from this plot with particular interest in the inflection or knee point (indicated with a blue rectangle) observed at the second singular value of every window. The rule of thumb reported in the original MUSICAL article [9], referred to as Rule A here, involves choosing roughly the threshold as the smallest second singular value across all the windows, i.e.

$$\text{Rule A: } \sigma_0 = \min(\sigma_2; \forall \text{windows}). \quad (5)$$

According to another rule reported in [16], referred here as Rule B, the threshold is roughly selected as the center of the span of the second singular values for all the windows, i.e.

$$\text{Rule B: } \sigma_0 = \frac{\min(\sigma_2; \forall \text{windows}) + \max(\sigma_2; \forall \text{windows})}{2}. \quad (6)$$

Both rules assume that the eigenimages can be clearly separated and therefore, they constitute hard thresholding schemes. An illustrative example is shown in Fig. 2(a) where the classification

and corresponding weighing is shown for Rule B in comparison to Rule A. In more practical terms, a higher value of threshold, decreases the cardinality of the signal subspace across the entire images as shown in the cardinality maps (c-map) of Fig. 2(b). These maps encode the number of eigenimages in the signal subspace for each window as an intensity value, which is reason why the the c-map B (corresponding to Rule B) looks dimmer. For even further insight, we show in Fig. 2(c) and (d), the actual singular values and eigenimages projection respectively for several orders. Something interesting to note, is that foreground regions generally have higher eigenvalues than the background for any given order. However, the projections of g_i shown Fig. 2(d) follow two different trends. First, the value of g_i observe an inversion in the pattern as the order increases. The lower orders have high values of g_i in the foreground relative to the background, and vice versa for the higher orders. This reversal of trend for higher orders (low singular values) gets exploited in the indicator function as their contribution to the denominator is small. Second, generally speaking, eigenimages with lower eigenvalues (i.e. higher orders) are associated with higher spatial frequency components [19], an effect that is more prominent for the first few orders. These observations will be further discussed later in relation with the new indicator functions.
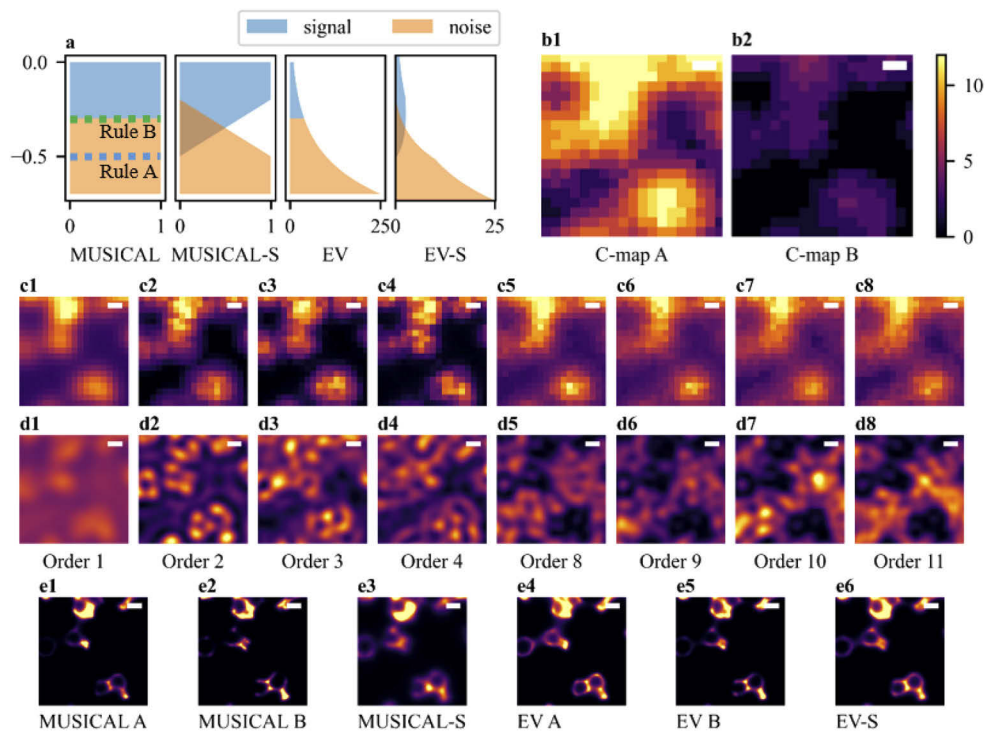

Fig. 2. MUSICAL's singular values and eigenimages. (a) Weighting coefficients for the different methods. The rules are labeled on top to show how the weights change depending on the threshold. In this example, rule A means a threshold of −0.5 while rule B means −0.3. For soft methods, the noise and signal are used both in numerator and denominator of the indicator function. (b) Cardinality maps that indicate the number of eigenimages picked as signal for the two rules of selection for every windows (each pixel corresponds to one window). (c), (d) Heatmap of singular values (c) and projection of different eigenimages (d) sorted by increasing order. (e) Results of using different indicator functions. The scale bars are 200 nm.

Since Rule B decreases the cardinality of the signal space, it also discards noise much more effectively. However, information related to the actual structure may also be relegated as noise, which increases the value of denominator in the foreground and compromises resolution. On the other hand, and following the example of Fig. 2, Rule A includes eigenimages up to order 11 in the foreground. Yet, as noted in Fig. 2(d), the g_i of the foreground is smaller than the background in orders 10 and 11. Therefore, their exclusion from the noise subspace is not useful for the foreground. Even more: the inclusion of higher g_i corresponding to the background due to these orders may increase the background artifacts. Therefore, a trade-off is involved in either rule and the manifestation of this trade-off varies from case-to-case.

In practice, even the most experienced bioimaging user of MUSICAL may not know what to expect from the sample being imaged. Moreover, one could argue that more candidate rules can be designed from analysis of the histogram of the second singular values, such as explored in [18]. However, in all these cases, a fundamental limitation is that they all imply hard thresholding: eigenimages included in the signal subspace are considered in a hard solely as representing the structure and the eigenimages included in the noise subspace are considered in solely as representing the noise. In reality, the presence of noise implies that each eigenimage is corrupted. Therefore, a perfect separation of the eigenimages into a signal or noise subspace is not possible.

3. New indicator function design

Here, we consider two solutions to the problem mentioned previously:

- Eigenvalue (EV) weighing: the magnitude of the eigenvalue is included in the indicator function. It keeps the hard thresholding but softens the effect of first few eigenimages that are classified as noise but may have structural information, as shown in Fig. 2(a).

- Soft-thresholding: hard thresholding is removed using a weighing function for each eigenimage. It can be added to MUSICAL and EV, with the new methods abbreviated MUSICAL-S and EV-S, respectively.

We begin with a generalized form of indicator function, which allows for more flexibility in its design and that paves the path for developing the indicator functions for the above identified solutions:

$$f(\mathbf{r}_{test}) = \left(\sqrt{\frac{\sum_{i=1}^{N} a_i g_i^2}{\sum_{i=1}^{N} b_i g_i^2}} \right)^{\alpha} . \tag{7}$$

Here, a_i and b_i are the design parameters of the indicator function. This generalization can be readily adapted to the original MUSICAL indicator function given by Eq. (4) by using the following assignments for a_i and b_i, with $a_i + b_i = 1$:

$$a_i = \begin{cases} 1 & \text{if } \mathbf{u}_i \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} \qquad b_i = \begin{cases} 1 & \text{if } \mathbf{u}_i \in \mathcal{N} \\ 0 & \text{otherwise} \end{cases} . \tag{8}$$

3.1. Indicator function with eigenvalue (EV) weighing

This indicator function follows $a_i + b_i = \lambda^{-1}$ and is defined as:

$$a_i = \begin{cases} \lambda_i^{-1} & \text{if } \mathbf{u}_i \in \mathcal{S} \\ 0 & \text{otherwise} \end{cases} \qquad b_i = \begin{cases} \lambda_i^{-1} & \text{if } \mathbf{u}_i \in \mathcal{N} \\ 0 & \text{otherwise} \end{cases} . \tag{9}$$

This indicator function design retains both KC1 and KC2, but does not use the Euclidean projections $||\mathbf{g}_{\mathcal{S}}(\mathbf{r}_{\text{test}})||$ and $||\mathbf{g}_{\mathcal{N}}(\mathbf{r}_{\text{test}})||$ on the signal and noise subspaces. It instead weighs the

projections on individual eigenimages g_i according to the inverse of its singular value. This is graphically illustrated in Fig. 2(a) for rule B.

We present further insight into the EV indicator function using Fig. 2(c-d). Due to multiplication with the inverse of eigenvalues, g_i for any order gets amplified for background regions and attenuated for the foreground regions. The foreground attenuation helps the higher order eigenimages to better support the resolution and the lower orders in reducing the dynamic range of the nanoscopy image on the higher side. The background amplification helps the higher order eigenimages to better suppress the background artifacts and the lower orders in reducing the dynamic range on the lower side. Thereon, the effect of hard thresholding is still present. Nonetheless, a significant softening is achieved as described next. Consider the orders that which are assigned to the signal subspace using Rule A but to the noise subspace using Rule B, however when treated using EV indicator function. When included in the signal subspace, they reduce the dynamic range of the original version of MUSICAL. This is significant because when using rule A, the original indicator function of MUSICAL generally supports better resolution but has extremely high indicator function values for few foreground pixels in nanoscopy image. This results into some pixels being highly saturated in the MUSICAL images and dynamic range of the image is not well-utilized, as reported in the supplement of the original MUSICAL paper [9]. On the other hand, when included in the noise subspace they help in improving the resolution, which may have been compromised in rule B as discussed before in section 2.4.

The proposed indicator function is inspired from the EV formulation reported previously for inverse source problems, for example in [20,21]. The similarity between these previously reported EV formulations and the one proposed here is limited to the denominator component of Eq. (7) when combined with b_i defined in Eq. (9). The use of the signal subspace in the numerator and the application of the indicator function on one sliding window at a time are unique to the MUSICAL architecture, first reported in [9] while incorporation of EV weighing in a_i in Eq. (9) is proposed for the first time here.

3.2. Indicator function with soft threshold (MUSICAL-S)

An alternative approach is to use continuous functions for a_i and b_i. Our proposed function is defined in Eq. (10) and graphically illustrated in Fig. 2(a).

$$a_i(x) = \begin{cases} 1 & \text{if } sigma_i \geq \sigma_{\max} \\ 0 & \text{if } sigma_i \leq \sigma_{\min} \\ \frac{\log_{10} \sigma_i - \log_{10} \sigma_{\min}}{\log_{10} \sigma_{\max} - \log_{10} \sigma_{\min}} & \text{otherwise} \end{cases} \quad ; \quad b_i(x) = 1 - a_i(x). \quad (10)$$

We enforce that $a_i + b_i = 1$. In this equation $\sigma_{\max}$ and $\sigma_{\min}$ must be given. We pick those to be the maximum and minimum of the second singular values across all the windows in the image, since the change in the trend is strongly evident in the second singular values. We note that the choice also obviates the need of user specification of $\sigma_{\min}$ and $\sigma_{\max}$. Thus, this approach includes both soft and automatic thresholding properties.

The design of indicator function as suggested above implies that the signal and noise subspaces are no longer orthogonal. There are some eigenimages ($\mathbf{u}_i; \sigma_i \geq \sigma_{\max}$) that are relegated to the signal subspace $\mathcal{S}$ with full confidence ($a_i = 1, b_i = 0$). Similarly, there are some eigenimages ($\mathbf{u}_i; \sigma_i \leq \sigma_{\min}$) that are relegated to the noise subspace $\mathcal{N}$ with full confidence ($a_i = 0, b_i = 1$). For the remaining eigenimages, it is understood that these eigenimages may have non-negligible signal information even while being significantly affected by noise. Therefore, these eigenimages are relegated to signal and noise space with reduced confidence (indicated by non-extreme values of a_i and b_i) while the behaviour and roles of the g_i for lowest and the highest orders are unambiguous. The role of some intermediate orders such as order 8 is not explicit. When included in the signal space, it does not contribute resolution in foreground but may help in

pushing the lower limit of the dynamic range up by enhancing the background. On the other hand, when included in the noise space, they may degrade the resolution but also pull the upper limit of the dynamic range down. By including them in reduced proportions in both denominators and numerators, we expect to strike a balance between the positive and negative aspects of their inclusion in the signal and noise subspaces and thereby achieve a softening effect of the threshold.

We note that this design is a significant deviation from the key concepts of MUSICAL. Since the signal and noise spaces now share some eigenimages, the KC1 defined in section 2.2 does not apply. Furthermore, the KC2 has to be redefined as follows:

Redefined KC2: If the signal and noise subspaces are suitably defined, then the b_i weighted projection of $\mathbf{g}(\mathbf{r}_n) \forall n$ on the noise subspace $\mathcal{N}$ is small, which allows the denominator in Eq. (7) to be small and the overall indicator function to be high at the emitter locations.

3.3. Indicator function for EV with soft-threshold (EV-S)

The concept of soft-thresholding can be integrated in EV as well, as shown below:

$$
a_i(x) = \begin{cases} \lambda_i^{-1} & \text{if } sigma_i \geq \sigma_{max} \\ 0 & \text{if } sigma_i \leq \sigma_{min} \\ \lambda_i^{-1} \left(\dfrac{\log_{10} \sigma_i - \log_{10} \sigma_{min}}{\log_{10} \sigma_{max} - \log_{10} \sigma_{min}} \right) & \text{otherwise} \end{cases} \quad ; \quad b_i(x) = \lambda_i^{-1} - a_i(x).
$$

(11)

As before, $a_i + b_i = \lambda^{-1}$. Its illustration in Fig. 2(a) shows that EV-S is the softest indicator function, allowing smoother transition in a_i and b_i. While EV alleviates the sensitivity to the threshold, MUSICAL-S removes the need for user-specified threshold and reduces the sensitivity to noise in the overlapping region in Fig. 2(a). Since EV-S combines traits from both, it is expected to demonstrate reduced sensitivity and soft thresholding.

3.4. Discussion of the proposed generalized framework

Through the generalized indicator function, we have allowed for the creation of families of MUSICAL algorithms based on some design rules. Specifically, two families have been created.

Family based on coefficient relationship: These are defined based on the relationship between a_i and b_i. The family $a_i + b_i = 1$ may be considered the original MUSICAL family while the family $a_i + b_i = \lambda_i^{-1}$ as the EV family. Other families may also be designed similarly.

Family of coefficient continuity: These are defined based on the individual characteristics of a_i and b_i. For example, in the family of hard threshold, the curve corresponding to a_i experiences an abrupt jump at the threshold σ_0, and the signal and noise spaces are disjoint. In the family of soft threshold, the signal and noise subspaces are no longer orthogonal (KC1 does not apply) but the eigenimages in the overlapping space of signal and noise subspaces are weighted log-linearly. Other approaches may be designed for choosing the overlapping region or designing the weights, leading to other families of algorithms.

Two final notes on the newly defined indicator functions:

No resolution enhancement expected: The aim of these new indicator functions is not resolution enhancement. In fact, no indicator function is expected to enhance or deteriorate resolution in particular. The only expected effect is a minor trade-off in resolution and contrast arising from different treatments of eigenimages in the signal and noise subspaces, as seen in Fig. 2(e).

Removing subjectivity through automatic thresholding: An important implication of the soft automatic indicator functions is that the subjectivity in threshold selection as well as the dependence on heuristics is removed. Outloook for further customization by an advanced user is possible, for example through a different choice of σ_{min} and σ_{max}, or a_i and b_i.

4. Results

We performed the following studies to compare the performances of the newly proposed indicator functions with the original indicator functions of MUSICAL:

- Quantitative analysis: We compare resolution and contrast for the indicator functions. We consider the effect of intensity fluctuations (determined by the photon emission on/off time) and the signal to background ratio (SBR). 2D samples are used so that other effects such as out-of-focus light do not affect the quantitative results.

- Qualitative analysis: Structures in biological samples are simulated to study how the different indicator functions deal with realistic 3D structures and out-of-focus light.

- Results on experimental data: We show comparative results using experimental data of a diverse set of samples. The dynamic range coverage is also investigated.

We consider a total of six indicator functions referred to as MUSICAL A, MUSICAL B, MUSICAL-S, EV A, EV B and EV-S. The value of $\alpha = 4$ is used as recommended in [9] and 10 subpixels per pixel are considered sufficient for the investigations presented in this article.

Simulation methods for studies 1-3: The simulation involves the following steps: simulating the geometry of the structures, the emitters distribution over the structures, and the photokinetics of each emitter using the photoemission model of [22], applying the Gibson-Lanni PSF [23,24] of the microscope to simulate the raw noise-free image stacks, and then simulating the noise in raw image stack using the noise model in [9]. We have simulated images for an epifluorescence microscope of numerical aperture 1.42, pixel size 80 nm, and 665 nm emission wavelength. Thereby, the theoretical resolution limit is ~285 nm using Rayleigh criterion [25]. For each sample, 500 frames were generated with 10 ms of exposure time per frame. Duty cycle is the ratio of the average time a fluorophore emits light (τ_{on}) to the total cycle ($\tau_{on} + \tau_{off}$). Photobleaching is not considered. Where not mentioned explicitly, SBR of 4 and duty cycle of 5% are utilized.

4.1. Quantitative analysis

Study of resolution: This study is performed on a sample comprising two crossing lines with an angle of 60° as shown in Fig. 3(a), where the emitters are uniformly distributed at a density of 500 per μm. The results of Fig. 3(b) show the distance from the crossing point at which the 2 lines can be differentiated according to the Rayleigh criterion [25] for different values of the duty cycle and SBR. The diffraction limited resolution is shown as a blue line for reference.

The resolution is estimated by computing the ratio between valley and peaks across a moving vertical section (shown in the inset of Fig. 3(b)), starting from the intersection point ($x = 0$). Let $\mathbf{l}(x)$ denote the image's intensities across a vertical line passing by x. We compute the ratio:

$$r(x) = \frac{v(\mathbf{l}(x))}{min(p_1(\mathbf{l}(x)), p_2(\mathbf{l}(x)))}. \tag{12}$$

Here, v and p_i are functions that return the minimum and maximum intensity value respectively value around the expected valley and peaks i (bottom and top peaks). The reported resolution is given as the minimum value x at which $r(x) \leq 0.835$. The range considered for duty cycle goes from 0.1 % (comparable to single molecule localization microscopy data) up to 50%. This last case corresponds to a highly dense spatio-temporal emission situation, which is a challenging situation for most fluctuations based techniques [26]. Further, we consider SBR from 2 to 6, where SBR 2 is considered quite poor. From Fig. 3, the first observation is that all the methods provide resolution enhancement by a factor between 2 and 3. When comparing rules, we find that the best results are obtained using rule B for both MUSICAL and EV. Considering hard thresholding and comparing the MUSICAL and EV families, MUSICAL takes the lead in general.
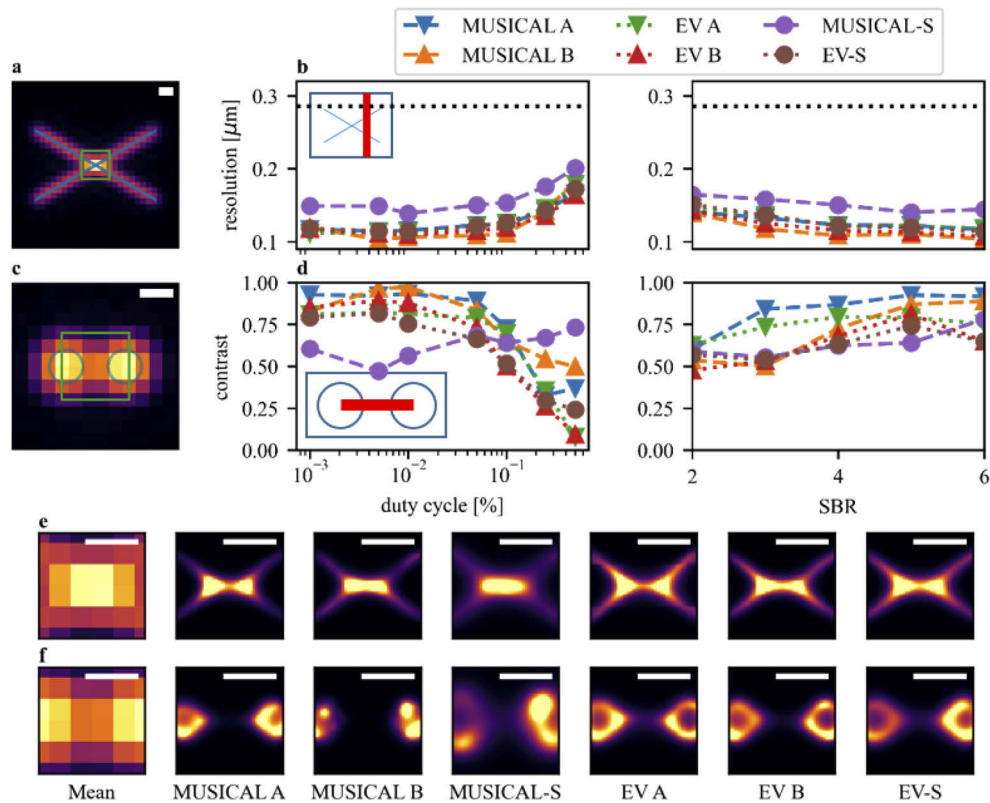
Fig. 3. Resolution and contrast study. (a), (c) Studied samples with the groundtruth in blue and green square shown magnified in panels (e), (f) for qualitative illustration. (b), (d) Resolution and contrast, respectively, for different methods at different duty cycles and SBR. Each data point in Fig. (b), (d) corresponds to the mean value over 100 independent simulations. Dashed line in (b) shows the resolution limit. Scale bar is 200 nm.

Among soft-threshold methods, EV-S performs comparable to MUSICAL and EV while but MUSICAL-S displays the worst performance among all of them. The poorer resolution of MUSICAL-S is also observed in Fig. 3(c) where the borders of the central cross appear diffused compared to the other methods, producing a lower resolution. Lastly, we note that duty cycles below 10% and SBR higher than 4 do not provide significant improvement in resolution.

Study of contrast: We here consider a sample containing emitters distributed uniformly across the perimeter of two circles of diameter 200 nm (density of 500 per μm) at a distance of 150 nm between edges. The mean image and groundtruth are shown in Fig. 3(c). The contrast is defined as $c = 1 - r(\mathbf{l}(0))$, where $r(x)$ is defined in Eq. (12) and $\mathbf{l}(0)$ represents the intensities across the horizontal section between circles (see inset of Fig. 3(d)). Higher values of c indicate better contrast. The results in Fig. 3(d) indicate that the contrast for all the indicator functions deteriorates as the duty cycle increases, with a larger slope after 10%. The only exception is MUSICAL-S which shows low sensitivity to duty cycle and has relatively flat contrast for all duty cycles. In terms of SBR, similar to the line sample, the curves become flat for SBR ≥ 4. When comparing rules, rule A achieves marginally better contrast both in MUSICAL and EV, with MUSICAL performing the best. The soft-threshold methods showed all similar and poorer results. An additional study on the effect of different noise distributions is included in Supplement 1.

Summary of the quantitative study: The soft-thresholding schemes provide resolution and contrast similar to the other hard thresholding schemes. In general, EV-S outperformed MUSICAL-S in terms of resolution and contrast due to the weighting scheme by including singular values. Additionally, from the results we observe that above a SBR of 4 and and duty cycle of 10%, no improvement in resolution or contrast were achieved.

4.2. Qualitative examples

Vesicles with surface labeling (Fig. 4(a1-a8)): Four vesicles of different sizes are simulated as spheres of diameters 150, 200, 250 and 300 nm. Membrane labeling is simulated by distributing emitters on surface with density 800 per μm^2. The vesicles are placed such that their centers are at the focal plane. Larger vesicles result with certain portions out of focus. In MUSICAL A (Fig. 4(a3)), the smallest and therefore dimmer structure (top-left vesicle) is almost invisible. This is not the case for MUSICAL B (Fig. 4(a4)) which is explained by a lower cardinality of the signal space and therefore better discarding of noise. In the case of MUSICAL-S (Fig. 4(a5)) the structure is even more visible, and the remaining vesicles looks more uniform. No clear difference is observed across EV reconstructions (Fig. 4(a6, a7) and (a8)). It is of interest to observe that for the largest vesicle (300 nm in diameter), the out of focus light is rejected by all the methods, producing a hollow in the middle that is not visible for the microscopy image. Below that size, the entire vesicle can be considered to be in focus. We observe that even if the size of the circle reconstructed is proportional to the original radius, this does not match the correct size when using full width at half maximum (FWHM).

Microtubules with background debris (Fig. 4(b1-b8)): Microtubules are fiber-like polymers of tubulin monomers. Fluorescent dyes label the monomers, which may be present as freely dispersing in addition to microtubule fiber [27]. This results in fluorescent debris, which is generally unwelcome in reconstruction. For this sample, we simulated fibers of 30 nm in diameter, with dyes distributed randomly across their surface at a linear density of 800 emitters per μm. Additionally, debris is added as single emitters distributed randomly with a volumetric density of 1000 emitters per μm^3. Both, microtubule-bound and free emitters, are assumed to have the same photo-kinetics. The geometry consists of three crossing lines forming an inverted triangle when seen from the top (Fig. 4(b2)). In the top-left and bottom corner, the structures meet in the focal plane, while the microtubules meeting at the top-right corner are both out of focus and separated by 500 nm in axial direction. The spatial distribution can be seen in Fig. 4(b1).

The results show that debris is absent in reconstructions regardless of the method while the overall structure is well reconstructed. A difference in the reconstruction of regions with out-of-focus structures is noticeable: MUSICAL A , B and S (Fig. 4(b3, b4) and (b5) respectively) achieve a better rejection than their EV counterparts (Fig. 4(b6, b7) and (b8)). However, the corners of the triangle which are the brightest points do not allow a clearer reconstruction of the strands due to the dynamic range problem of MUSICAL reported in [9] and discussed in section 3. In this sense, a better reconstruction is obtained by EV. It shows that a desired level of rejection of out-of-focus structures could determine the choice between MUSICAL and EV.

Mitochondrion (Fig. 4(c1-c8)): Mitochondria are tubular structures with diameters close to or larger than the diffraction limit. Even if such structures are in focus, the sometimes large diameters combined with the dynamic nature of these organelles in living samples, causes large portions of the samples to be out of focus in realistic microscopy experiments. Here, we consider an example of three mitochondria with diameter 300 nm and a density of emitters of 3000 per μm^2 on the outer membrane. Each mitochondrion is in a different plane, with the left most mitochondrion in the focal plane. The top one is in the plane 300 nm above (closer to the coverslip) and the last one is 300 nm below the focal plane (further from the coverslip). As in the case of microtubules, the out of focus rejection is prominent for all methods with MUSICAL obtaining the strongest rejection of out of focus signal. Only the portions in the focal plane are
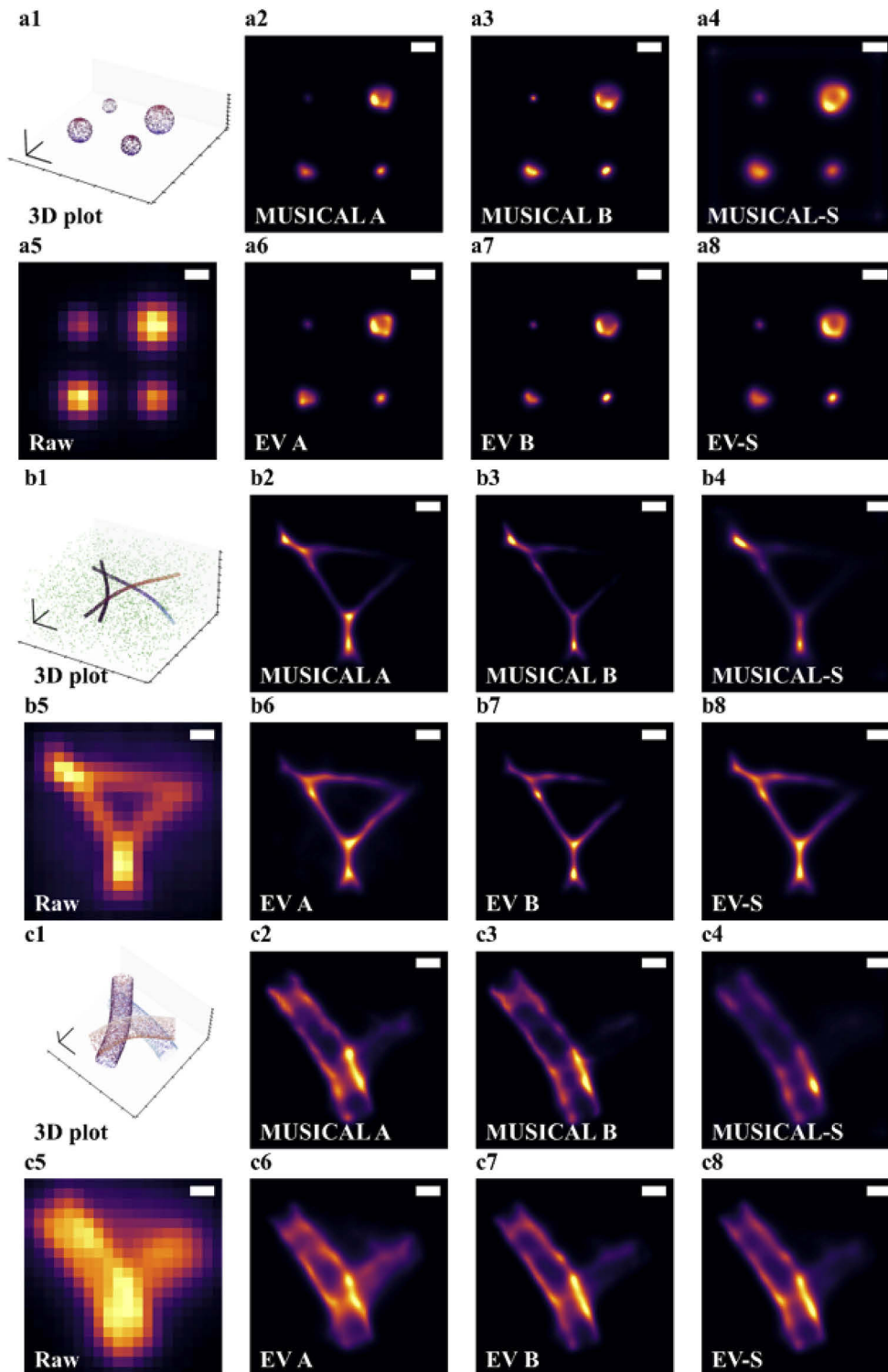
Fig. 4. Qualitative study on synthetic samples. Three different samples are shown. (a) 4 vesicles of different size. (b) Three crossing microtubules with debris. (c) The crossing mithondria. All scale bars are 300 nm.
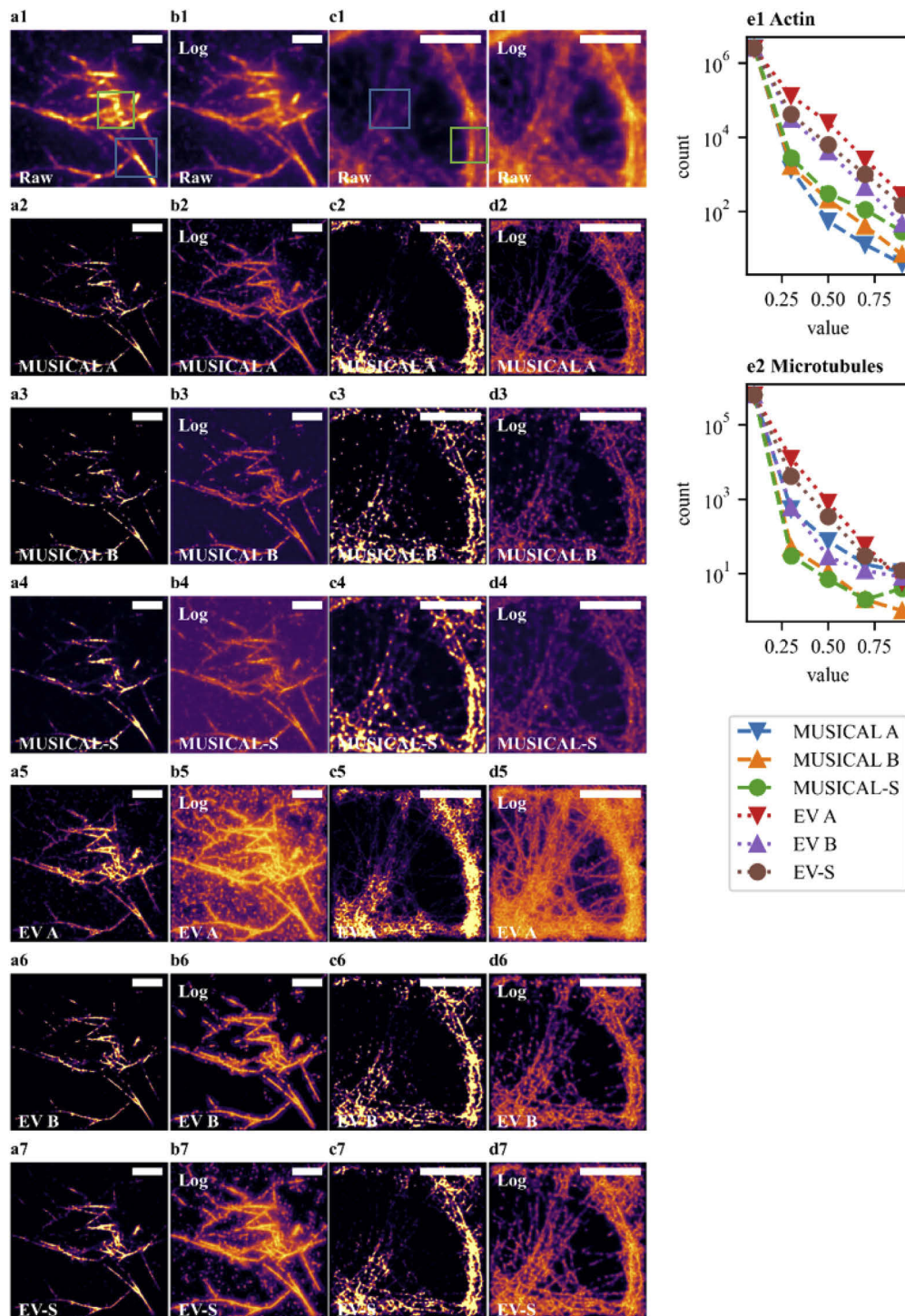
Fig. 5. Results on experimental data. The results are displayed in two scales: linear (a) and (c) and logarithmic with base 10 (b) and (d). Blue and green boxes show regions of interest. The scale bar is 2 μm in every image. (a), (b). Results for actin (invitro). (c), (d) Results on microtubules. (e) Comparative plot of the intensities obtained from the histogram of each normalized image by its maximum.

reconstructed with good contrast. Further, the structure away from the coverslip is rejected less effectively than the structure above, which can be explained by the asymmetry of the PSF.

Summary of this study: The MUSICAL indicator functions were found to perform stronger out-of-focus rejection than the EV indicator functions.

4.3. Results on experimental data

The methods were here tested on real microscopy data of filamentous actin and microtubules in fixed cells. The samples used for testing the methods correspond to invitro actin filaments [9] and microtubules in fixed cells [28]. 500 frames were used for each reconstruction.

Invitro actin filaments (Fig. 5(a,b)): This sample is thin in the sense that that all the structures can be considered in focus. We marked two regions of interest (blue and green boxes) where super-resolution can give better insight of the structure than conventional microscopy. In the blue box (close to bottom), the bifurcation is clearly visible using the different indicator functions, and almost no difference is observed in terms of structure. However, the region in the green box is reconstructed better in EV as they show a better distribution of the intensities. This allows to observe the entire network of strands without saturating the colors in other regions and is mainly attributed to the contrast enhancement due to the softening effect of EV. On the other hand in the images in logarithmic scale (Fig. 5(b)), we observe how EV introduces artifacts in the background, which is reduced when using a larger threshold such as the one used with rule B. In the case of the soft thresholding methods, EV-S looks crisper and more defined that MUSICAL-S, while providing better contrast between foreground and background. The pixel distribution is plotted in log scale in Fig. 5(a8), where it can be clearly observed how EV produces an intensity distribution that is higher in the middle tones, making better use of the dynamic range.

Microtubules in fixed cells (Fig. 5(c,d)): This sample of fixed cell is different from the actin sample in the sense that some structures are expected to be out-of-focus. For images in linear scale (Fig. 5(c1-c7)), the methods display similar performance, except for EV A. In particular, the blue region (left box) presents a high degree of artifacts, where it is difficult to visualize individual strands due to high saturation. The same occurs in the green region, where EV A recovers just one single structure, while all the remaining methods manage to recover 2 strands. This illustrates an example of the minor trade-off between resolution and contrast. The better contrast and visibility, as well as poorer rejection of out-of-focus structures is strongly evident in the log scale (Fig. 5(d)). Between MUSICAL-S and EV-S, the latter produces a better result by achieving better contrast and definition. Both methods recover the dim structures in the right-most box.

5. Conclusion

Through a generalized framework for MUSICAL indicator function design, we have proposed new indicator function families and specific indicator function designs to address problem of hard threshold in MUSICAL. The EV family of indicator functions soften the effect of threshold and provide better utilization of the dynamic range of MUSICAL. This is generally achieved at no cost of resolution but poorer rejection of out-of-focus light as compared to the MUSICAL family of indicator functions. Further, a soft threshold family is proposed that does not define signal and noise subspaces of MUSICAL as strictly orthogonal, and allows an overlap between them. Therefore, it removes the concept of hard thresholding, however, altering the key concepts of MUSICAL. While MUSICAL-S indicate poorer resolution, the EV-S indicator function which combines the traits of both EV and soft-threshold families shows consistently good results across a wide variety of quantitative, qualitative, and experimental studies. Through this work, we widen the horizon for MUSICAL in two important aspects. First, the sensitivity and subjective choice of threshold is removed which makes it easier to use. Second, it opens exciting avenues for further development of fluctuations based super-resolution algorithms in general, and MUSICAL

in particular. In the future, it will be interesting to design customized indicator functions for challenging scenarios such as dynamic samples where the different methods can be tested.

Funding

Disclosures

The authors declare that there are no conflicts of interest related to this article.

See Supplement 1 for supporting content.

References

1. M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," J. Microsc. **198**(2), 82–87 (2000).
2. S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," Opt. Lett. **19**(11), 780–782 (1994).
3. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," Science **313**(5793), 1642–1645 (2006).
4. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," Nat. Methods **3**(10), 793–796 (2006).
5. J. Schnitzbauer, M. T. Strauss, T. Schlichthaerle, F. Schueder, and R. Jungmann, "Super-resolution microscopy with DNA-PAINT," Nat. Protoc. **12**(6), 1198–1228 (2017).
6. K. C. Gwosch, J. K. Pape, F. Balzarotti, P. Hoess, J. Ellenberg, J. Ries, and S. W. Hell, "MINFLUX nanoscopy delivers 3d multicolor nanometer resolution in cells," Nat. Methods **17**, 217–224 (2020).
7. T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, "Fast, background-free, 3d super-resolution optical fluctuation imaging (SOFI)," Proc. Natl. Acad. Sci. **106**(52), 22287–22292 (2009).
8. N. Gustafsson, S. Culley, G. Ashdown, D. M. Owen, P. M. Pereira, and R. Henriques, "Fast live-cell conventional fluorophore nanoscopy with imagej through super-resolution radial fluctuations," Nat. Commun. **7**(1), 12471 (2016).
9. K. Agarwal and R. Macháň, "Multiple signal classification algorithm for super-resolution fluorescence microscopy," Nat. Commun. **7**(1), 13752 (2016).
10. I. Yahiatene, S. Hennig, M. Müller, and T. Huser, "Entropy-based super-resolution imaging (ESI): From disorder to fine detail," ACS Photonics **2**(8), 1049–1056 (2015).
11. Y. S. Hu, X. Nan, P. Sengupta, J. Lippincott-Schwartz, and H. Cang, "Accelerating 3B single-molecule super-resolution microscopy with cloud computing," Nat. Methods **10**(2), 96–97 (2013).
12. Y. Deng, M. Sun, P.-H. Lin, J. Ma, and J. W. Shaevitz, "Spatial covariance reconstructive (SCORE) super-resolution fluorescence microscopy," PLoS One **9**(4), e94807 (2014).
13. W. Zhao, J. W. Liu, C. Kong, Y. Zhao, C. Guo, C. Liu, X. Ding, X. Ding, J. Tan, and H. Li, "Faster super-resolution imaging with auto-correlation two-step deconvolution," arXiv preprint arXiv:1809.07410 (2018).
14. O. Solomon, M. Mutzafi, M. Segev, and Y. C. Eldar, "Sparsity-based super-resolution microscopy from correlation information," Opt. Express **26**(14), 18238–18269 (2018).
15. S. Acuña, F. Ströhl, I. S. Opstad, B. S. Ahluwalia, and K. Agarwal, "MusiJ: an ImageJ plugin for video nanoscopy," Biomed. Opt. Express **11**(5), 2548–2559 (2020).
16. I. S. Opstad, F. Ströhl, Å. B. Birgisdottir, S. Maldonado, T. Kalstad, T. Myrmel, K. Agarwal, and B. S. Ahluwalia, "Adaptive fluctuation imaging captures rapid subcellular dynamics," in *Advances in Microscopic Imaging II*, vol. 11076 (SPIE - International Society for Optics and Photonics,USA, 2019), p. 110761W.
17. G. T. Dempsey, J. C. Vaughan, K. H. Chen, M. Bates, and X. Zhuang, "Evaluation of fluorophores for optimal performance in localization-based super-resolution imaging," Nat. Methods **8**(12), 1027–1036 (2011).
18. S. A. Acuña-Maldonado, "Multiple Signal Classification Algorithm: computational time reduction and pattern recognition applications," Master's thesis, UiT The Arctic University of Norway, Tromsø, Norway (2019).
19. K. Agarwal and D. K. Prasad, "Eigen-analysis reveals components supporting super-resolution imaging of blinking fluorophores," Sci. Rep. **7**(1), 4445 (2017).
20. V. Madisetti, *The Digital Signal Processing Handbook* (CRC, USA, 2009), 2nd ed.
21. D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques* (Simon & Schuster, USA, 1992).

22. A. Girsault, T. Lukes, A. Sharipov, S. Geissbuehler, M. Leutenegger, W. Vandenberg, P. Dedecker, J. Hofkens, and T. Lasser, "SOFI simulation tool: A software package for simulating and testing super-resolution optical fluctuation imaging," PLoS One **11**(9), e0161602 (2016).

23. S. F. Gibson and F. Lanni, "Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy," J. Opt. Soc. Am. A **9**(1), 154–166 (1992).

24. J. Li, F. Xue, and T. Blu, "Fast and accurate three-dimensional point spread function computation for fluorescence microscopy," J. Opt. Soc. Am. A **34**(6), 1029–1034 (2017).

25. L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University, 2012), 2nd ed.

26. I. S. Opstad, S. Acuña, L. E. V. Hernandez, J. Cauzzo, N. Škalko Basnet, B. S. Ahluwalia, and K. Agarwal, "Fluorescence fluctuations-based super-resolution microscopy techniques: an experimental comparative study," arXiv preprint arXiv:2008.09195 (2020).

27. E. Nogales, "Structural insights into microtubule function," Annu. Rev. Biophys. Biomol. Struct. **30**(1), 397–420 (2001).

28. K. Agarwal, R. Macháň, and D. K. Prasad, "Non-heuristic automatic techniques for overcoming low signal-to-noise-ratio bias of localization microscopy and multiple signal classification algorithm," Sci. Rep. **8**(1), 4988 (2018).