SPITFIR(e): A supermaneuverable algorithm for fast denoising and deconvolution of 3D fluorescence microscopy and videos

Sylvain Prigent, Hoai-Nam Nguyen, Ludovic Leconte, Cesar Augusto Valades-Cruz, Bassam Hajj, Jean Salamero, Charles Kervrann

SPITFIR(e): A supermaneuverable algorithm for restoring 2D-3D fluorescence images and videos, and background subtraction

Sylvain Prigent[1,2,⋆], Hoai-Nam Nguyen [1,2,⋆], Ludovic Leconte[1,2], Cesar Augusto Valades-Cruz[1,2], Bassam Hajj[3], Jean Salamero[1,2], Charles Kervrann[1,2,∗]

[1] SERPICO Project-Team, Inria Centre Rennes-Bretagne Atlantique, 35042, Rennes Cedex, France
[2] SERPICO/STED Team, UMR144 CNRS Institut Curie, PSL Research University, Sorbonne Universités, 75005, Paris, France
[3] Laboratoire Physico-Chimie, Institut Curie, PSL Research University, Sorbonne Universités, CNRS UMR168, 75005, Paris, France

[⋆] S. Prigent and H.-N. Nguyen contributed equally to this work.
[∗] Corresponding author: charles.kervrann@inria.fr

Abstract

While fluorescent microscopy imaging has become the spearhead of modern biology as it is able to generate long-term videos depicting 4D nanoscale cell behaviors, it is still limited by the optical aberrations and the photon budget available in the specimen and to some extend to photo-toxicity. A direct consequence is the necessity to develop flexible and "off-road" algorithms in order to recover structural details and improve spatial resolution, which is critical when pushing the illumination to the low levels in order to limit photo-damages. Moreover, as the processing of very large temporal series of images considerably slows down the analysis, special attention must be paid to the feasibility and scalability of the developed restoration algorithms. To address these specifications, we present a very flexible method designed to restore 2D-3D+Time fluorescent images and subtract undesirable out-of-focus background. We assume that the images are sparse and piece-wise smooth, and are corrupted by mixed Poisson-Gaussian noise. To recover the unknown image, we consider a novel convex and non-quadratic regularizer Sparse Hessian Variation) defined as the mixed norms which gathers image intensity and spatial second-order derivatives. This resulting restoration algorithm named SPITFIR(e) (SParse fIT for Fluorescence Image Restoration) utilizes the primal-dual optimization principle for energy minimization and can be used to process large images acquired with varied fluorescence microscopy modalities. It is nearly parameter-free as the practitioner needs only to specify the amount of desired sparsity (weak, moderate, high). Experimental results in lattice light sheet, stimulated emission depletion, multifocus microscopy, spinning disk confocal, and wide-field microscopy demonstrate the generic ability of the SPITFIR(e) algorithm to efficiently reduce noise and blur, and to subtract undesirable fluorescent background, while avoiding the emergence of deconvolution artifacts.

Fluorescence microscopy provides a very powerful framework to biologists for observing, analyzing, and studying specific fluorescently-tagged structures and biological processes at very high spatial and temporal resolutions. In general, the specimen is first labelled by fluorescent molecules before being excited by the illumination light of a given wavelength; upon the excitation, fluorophore re-emits light of relatively longer wavelength which is then collected by photosensitive sensors to form the digitized image of the observed sample. Despite number of advantages, there are two major limitations of fluorescence microscopy. The first limitation is the presence of photon (shot) noise in acquired images. Shot noise is mainly due to the quantum nature of light, implying that the arrival of a photon on a sensor is a random event and thus the number of incident photons over a period of time is a random variable depending on the brightness of the light source. Moreover, experimentally in biological samples the signal-to-noise-ratio (SNR) is usually very low because low dose of illumination light is required to avoid photo-bleaching (i.e., progressive fading of the emission intensity) of fluorescent molecules and to preserve specimen integrity (photo-toxicity) [1;2]. Additionally, the quality of acquired fluorescence images is worsened by the blurring effects induced by several factors such as excitation wavelength, immersion medium refraction, specimen thickness, and the limited aperture of the objective which results in light diffraction through the optical system. The diffraction phenomenon implies that light emitted by an infinitely small point source appears wider at the focal plan and spreads into a specific pattern called "point spread function" (PSF).

Noise and blur not only degrade the image quality in terms of overall visualization but also have a negative influence on specimen analysis, including detection and segmentation of biological structures. To improve the quality of images acquired by fluorescence microscopes, restoration (deconvolution and/or denoising) is then frequently applied as pre-processing step before quantitative analysis, and amounts to recovering the original unknown image u from the observed noisy and blurred image f represented as follows: $f = \mathcal{T}(h * u)$, where h denotes the 2D or 3D spatial response of the device representing the blur related to the optical system (e.g., PSF) assumed to be linear shift-invariant, $*$ denotes the convolution operator, and $\mathcal{T}$ is a degradation operator modeling the measurement noise. If the image is not blurred but mainly corrupted by noise, h is represented by the Dirac (or impulse) function and the restoration problem translates into a denoising problem. Finally, the degradation operator (shot noise and readout noise) is usually modeled as a mixed Poisson-Gaussian distribution which is non-stationary and signal-dependent.

In the last twenty years, many restoration methods have been investigated in order to "inverse" the model in 2D-3D fluorescence microscopy. The most popular restoration approaches are linear methods (e.g., Wiener filtering). Despite the simplicity and low-computation-requirement, they usually do not restore fine image details at frequency components that are beyond the bandwidth of the PSF (i.e., the support of its Fourier transform). Another popular technique is the Richardson-Lucy (RL) algorithm [3;4;5]. While well-dedicated to Poisson noise removal, RL deconvolution creates structural noise and artifacts after a small number of iterations, which constitutes a major problem in fluorescence imaging [6]. Actually, the most flexible and robust deconvolution methods consist in minimizing an energy functional composed of a data fidelity term and a regularization term that encompasses prior knowledge (positivity, smoothness, sparsity,...) about the solution. The seminal Tikhonov-Miller (TM) approach [7;8;9] may be considered as the starting point to the development of regularization methods in fluorescence microscopy. While it was successful in many applications, the TM regularizer tends to produce blurred images since low gradient magnitudes are encouraged in the entire image, including at image discontinuities. To avoid over-smoothing caused by

quadratic functionals, non-quadratic regularizers have been first studied, especially the Total Variation (TV) regularizer that penalizes the L_1 norm of the gradient. Nevertheless, TV creates "stair-casing" effects, which are particular undesirable in fluorescence imaging [10;11]. To overcome this disadvantage, Lefkimmiatis *et al.* [12] recently proposed a family of convex regularizers built upon the matrix norm of the Hessian (Schatten norm), computed at each point of the image. This piece-wise smoothness promoting regularizer combined with a non-quadratic data term, was recently evaluated in fluorescence microscopy [13]. Nevertheless, this regularizer does not impose sparsity prior on the fluorescent signals which is also an important feature in fluorescence imaging. This idea was investigated by Arigovindan *et al.* [14] who reported good results 3D wide-field fluorescence images by using a non-convex sparsity-promoting regularizer; this regularizer (named here Log Sparse Hessian Variation (LHSV)) has been implemented into the ER-Decon deconvolution algorithm. Concurrently, other sparsity-promoting methods, focused on first-order differentiation-based regularizers have been investigated in medical imaging such as GraphNet (GN) [15;16;17] and Sparse Variation (SV) [18], and in image processing (e.g., TV-L_1 [19;20]). Nevertheless, the sparsity-promoting regularizers reported in Table 1 have limitations. They were designed for 2D medical imaging [15;16;17;18] or 3D microscopy [14;21] but never for both, are based on the first-order [15;16;17;18] or second derivatives [14;21], are formulated as non-convex [14] or convex [21] optimization problems which are now efficiently solved by performant algorithms such as ADMM [22;23] and FISTA [24].

To overcome the main limitations of previous methods, we introduce a flexible method (SPITFIR(e)) to deconvolve and/or denoise a large range of 2D/3D fluorescence microscopy images and videos (Supplementary Fig. 1). SPITFIR(e) is able to handle low signal-to-noise ratios, high-dimensional images (2D/3D/4D multispectral images), and characteristics of microscopy set-ups, especially if the PSF is not well known. The resulting approach includes a novel 2D/3D/4D sparsity promoting regularizer, can adapt to most if not all fluorescence techniques without fine-tuning of parameters, and shows robustness to different sources of noise and degradations. The SPITFIR(e) parameters are automatically adjusted from the image contents and the practitioner needs only to specify if the target image is "highly sparse", "moderately sparse" or "weakly sparse". We validated the performance of our approach and compared SPITFIR(e) to competitive deconvolution methods on several datasets, and we confirmed the results on real images acquired with lattice light sheet (LLSM), stimulated emission depletion (STED), multi-focus (MFM), spinning disk confocal (CM), and wide-field microscopy (WF). For the evaluation on different imaging modes, we have chosen to apply SPITFIR(e) on a limited number of structures relatively difficult to resolve or for which the photon budget is critical, i.e., mitochondria, microtubules and biomolecules in motion in living cells.

Results

Overview of SPITFIR(e). We present SPITFIR(e), an algorithm well grounded in the regularization theory for inverse problems, that can robustly recover information from general multidimensional noisy and blurry fluorescence images. This algorithm is able to adapt to several if not all microscopy modalities as illustrated in Figs. 1-3, and to different sources of noise and blur. The practitioner needs only to set the desired level of sparsity, which amounts here to selecting if the image is *"highly sparse"*, *"moderately sparse"* or *"weakly sparse"*. A "weakly" sparse image contains clutter and complex contents. Unlike the traditional methods requiring the tedious manual adjustment of the parameter that balances the data fidelity term and the prior term, SPITFIR(e) includes a self-tuning and scale-invariant technique for auto-
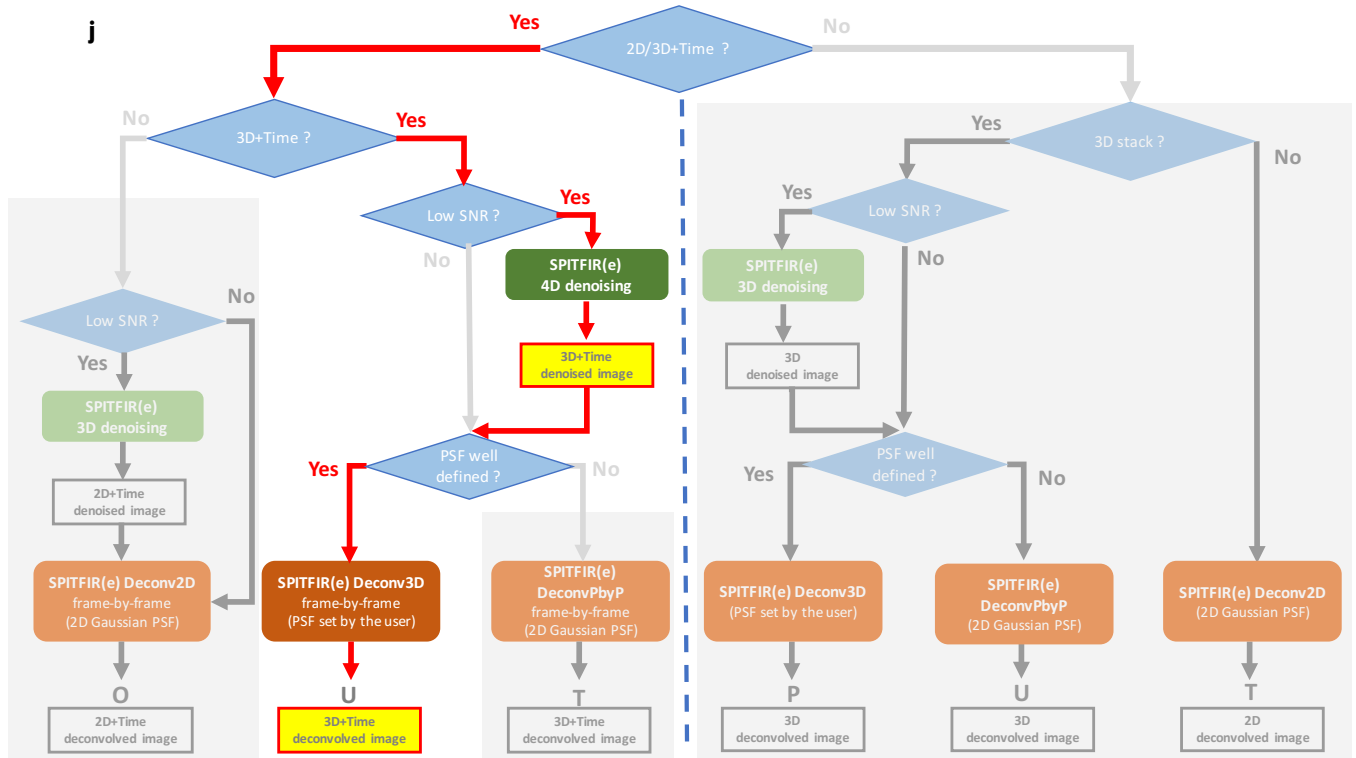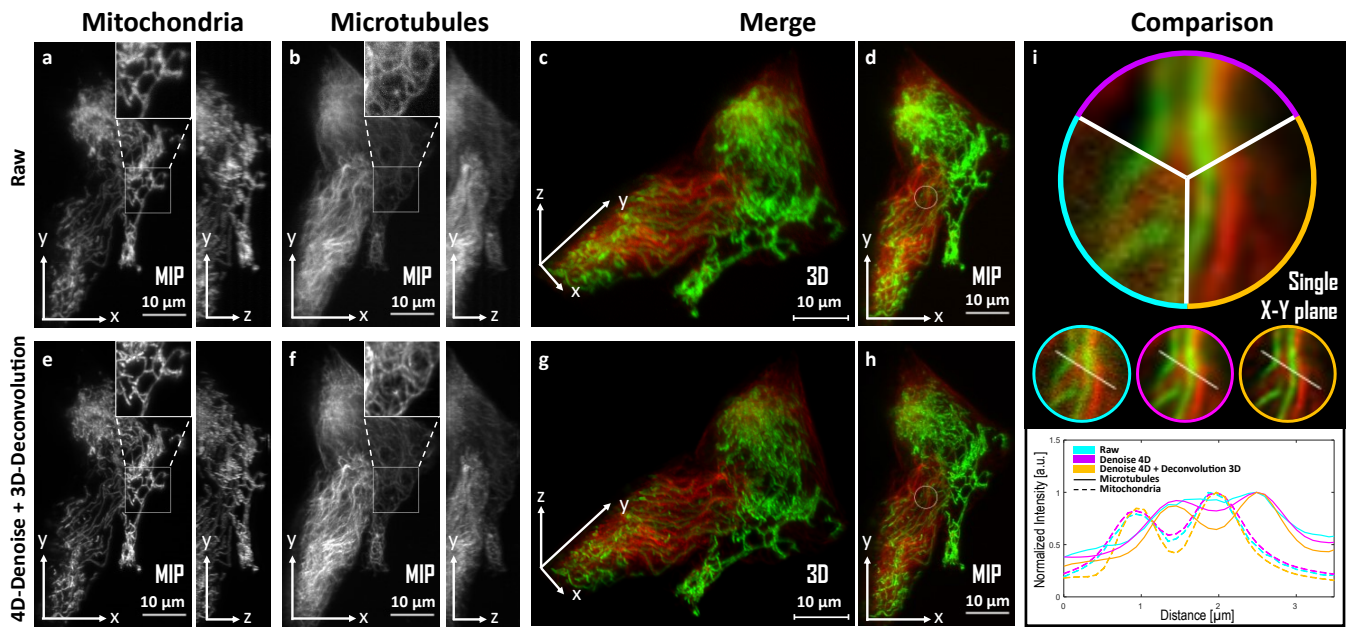
Fig. 1: SPITFIR(e) for 4D-denoising, background subtraction, and 3D-deconvolution on LLSM images. 60 planes 3D volumes of live RPE1 cells double stained with PKMR for Mitochondria (**a,e**) and with Tubulin Tracker^{TM} Deep Red for Microtubules (**b,f**) were acquired within $2.2s$ per stack using lattice light-sheet microscopy. MIP of representative raw (see Material and Methods section) images for Mitochondria and Microtubules, respectively, before (**a,b**) and after (**e, f**) SPITFIR(e) 4D denoising and 3D deconvolution (3D Gaussian PSF, $\sigma_{xy} = 1.5$ pixels and $\sigma_z = 1.0$ pixel). Insets are zoomed area illustrating SPITFIR(e) improvement in spatial resolution and SNR. 3D angular views and MIP of composite images before (**c,d**) and after (**g,h**) SPITFIR(e) treatment (Red for Microtubules; Green for Mitochondria). 4D denoising SPITFIR(e) image improvement shown on a single x-y plane (**i**) zoomed from the inset indicated in (**d, h**). Raw image is indicated as a blue lined sector in the upper part and in the left circle, beneath; 4D denoised image is similarly indicated in magenta, while the 4D denoised + 3D deconvolved image is in yellow. Intensity line profiles were measured as indicated in the 3 circles and plotted for each processing step and both Mitochondria and Microtubules at the lower part of (**i**). **j**, Workflow for image restoration extracted from the global flow chart in Supplementary Fig. 1. Scale bars are indicated in the bottom right corner (**a-h**). Related videos are shown as Supplementary movies S1 and S2.

2

matic adaptation (Methods). The minimization of the sum of the two convex terms is originally implemented in a computationally efficient way with a recent version of proximal-splitting optimization algorithm[25;26] (Methods and Supplementary Notes (Appendix B)). In what follows, we show that SPITFIR(e) can restore 3D fluorescence images under low excitation light excitation conditions by chaining denoising and deconvolution with the same algorithm with very small computing times compared to other competitive methods[13;21]. A "plane-by-plane" (PbyP) deconvolution strategy is also included for fast deconvolution if the 3D PSF is not well known. It turns out that the decomposition into a series of 2D decoupled deconvolution problems provided very good results in a small computing time. The nearly parameter-free SPITFIR(e) algorithm has been implemented both in CPU and GPU, is interfaced with Napari (https://www.napari-hub.org/plugins/napari-sdeconv) and can be downloaded for free here: https://github.com/sylvainprigent/simglib.

SPITFIR(e) is flexible and adapts to every fluorescence microscopy modality. SPITFIR(e) includes several strategies to adapt to microscopy specificities, and to particular spatial and temporal acquisition conditions (Supplementary Fig. 1). For instance, the practitioner can apply the conventional 3D deconvolution strategy if the PSF is perfectly known. If this is not the case, the PbyP strategy which consists in separately deconvolving each XY section of the 3D stack given a 2D Gaussian PSF model, can be applied. Surprisingly, the PbyP strategy could produce better visual results and is faster than direct 3D deconvolution with an imperfect or approximately measured 3D PSF. Meanwhile, in the case of low-photon regimes or low-exposure times, the results may be significantly improved by applying the "denoise-before-deconvolve" approach. For instance, a 3D+Time (4D) volume can be denoised as a whole, while the planes of each 3D stack are independently deconvolved with a 2D Gaussian PSF model. An exemplary demonstration of the flexibility of SPITFIR(e) is given by the experiments performed with a lattice light sheet microscope[27] (LLSM) which allows to acquire 3D images for several hours with limited photobleaching. In LLSM imaging, a sheet of light inclined about 32 degrees illuminate one slice of the sample while an objective oriented at 90 degrees from the illumination slice allows to observe and image the illuminated slice. Moving the illumination slice allows then to create a 3D image. As the illumination angle makes the raw image stack skewed, post-processing is required to create a 3D "deskewed" stack. Consequently, deconvolving 3D LLSM images is not straightforward as the PSF is not well characterized and the deskew operation affects the noise statistics. We then design the following SPITFIR(e) deconvolution pipeline according to the flowchart illustrated in Supplementary Fig. 1. We hypothesized here that chaining the denoising and deconvolution steps could provide better restorations results. So, first we denoise the 3D+Time image sequence (4D denoising) with SPITFIR(e) (the PSF function h is a Dirac function). Then we deconvolve each 3D stack separately with SPITFIR(e) by using an anisotropic 3D Gaussian PSF model. The settings of the PSF are chosen given the image resolution along the XY and Z axes.

We illustrate the efficiency of the "denoise-before-deconvolve" approach using mitochondria (Fig. 1). Live imaging mitochondria and their interactions with other intracellular components at high spatial resolution in 3D and high temporal frequency, is known to be challenging. Beyond the usual optical limits, labeled mitochondria are very sensitive to light[28], which could induce artifactual fluorescence signals such as exaltation, probes dissociation and finally fragmentation of mitochondria and cell death. This may impair accurate 3D+Time observation of their behaviors in normal and stress conditions. For this reasons, imaging mitochondria dynamics and structural features will serve as one of the biological thread in the following parts of the

manuscript. We here linked 3D LLSM at low illumination regimes to efficiently preserve fluorescence from photobleaching and consequently from phototoxic effects, with 4D denoising and 3D image deconvolution (3D Gaussian PSF, $\sigma_{xy} = 1.5$ pixels and $\sigma_z = 1.0$ pixel) by using SPITFIR(e). We imaged live RPE1 cells in full volume (Fig. 1 labeled for mitochondria (Fig. 1(**a**) and (**e**)) and microtubules (Fig. 1(**b**) and (**f**)). A reconstituted comparison of the successive steps of the SPITFIR(e) workflow is shown in Fig. 1(**i**). For the sake of visual clarity, only one single plane of the zoomed area of a stack is shown, (as indicated in Fig. 1(**d**) and (**h**)) and clock wise decomposed from the raw image to full processing. More visual representations can be appreciated in supplementary Movies 1 and 2, in 3D+Time along both XY and Z axes. Interestingly, chaining denoising and deconvolution with SPITFIR(e) is particularly visible on the microtubule staining, while improvement in resolution is more obvious on mitochondria (compare insets of zoomed area between Fig. 1(**b**) and (**f**) and between (**a**) and (**e**)). Line intensity plots through the center of the fluorescent filamentous structure demonstrate the recovering of the signal from the noise (Fig. 1(**i**) bottom part). Furthermore, the potential of SPITFIR(e) in LLSM imaging is also illustrated on a volume of 58 planes depicting endosomes stained with the AP2 proteins (Supplementary Fig. 2) known to be involved in the early steps of endocytosis[29]. In this experiment, we compared both the 3D deconvolution and PbyP deconvolution strategies with a 3D Gaussian model ($\sigma_{xy} = 1.0$ pixel and $\sigma_z = 0.8$ pixel) and a 2D Gaussian model ($\sigma_{xy} = 1.0$ pixel), respectively. The line intensity plots (**b**) through the center of the fluorescent spots demonstrates that SPITFIR(e) with the BbyP strategy preserves the intensity peaks, and better narrows the full-width-half-maximum (FWMH) when compared to RL[3] and HR-C3D[13]. These two case-studies indicate the ability of SPITFIR(e) to recover general shapes (i.e., filaments and point-like objects) in noisy and blurred fluorescence images.

Meanwhile, we also evaluated the flexibility of SPITFIR(e) on spinning disk confocal microscopy images. Spinning disk microscopes with many pinholes allows a high speed acquisition of images and therefore is compatible with acquisition regimes corresponding to live cell imaging. Nevertheless, time exposure must be reduced to capture fast intracellular events, such as biomolecules dynamics. Supplementary Fig. 3 shows the maximum intensity projection along the Z-axis of a stack depicting Rab5 proteins stained with Green Fluorescence Protein (GFP) close to endosomes within a cell. The image is blurred and noisy, and contains small spots at the resolution limit of the microscope against a dark background. Supplementary Fig. 4 shows the maximum intensity projection along the Z-axis of a stack depicting mCherry-actin filaments, that is thin tubular-like structures. This image is apparently not as sparse as the other image. These two examples are challenging because the PSF is approximately known and the image contents are very different. The PSF is assumed to be approximated by a 3D and a 2D Gaussian models. In each figure, we reported the results obtained with four strategies (3D deconvolution, PbyP deconvolution, 3D denoising + 3D deconvolution, 3D denoising + PbyP deconvolution) and different levels of sparsity. Applying denoising beforehand enables here to better enhance weak fluorescent signals while removing blur and background. From Supplementary Figs. 3-4, it turns out that the images are visually best results on the GFP-Rab5 and mCherry-actin images when the sparsity is "high" and "moderate", respectively

With SPITFIR(e), we have also improved the resolution limits of current STED microscopy. This nanoscopy is one of the very few light microscopy techniques allowing to solve mitochondria cristae organization, as illustrated here in live cells (Fig. 2). STED microscopy provides subdiffraction resolution while preserving useful aspects of fluorescence microscopy, such as optical sectioning, and molecular
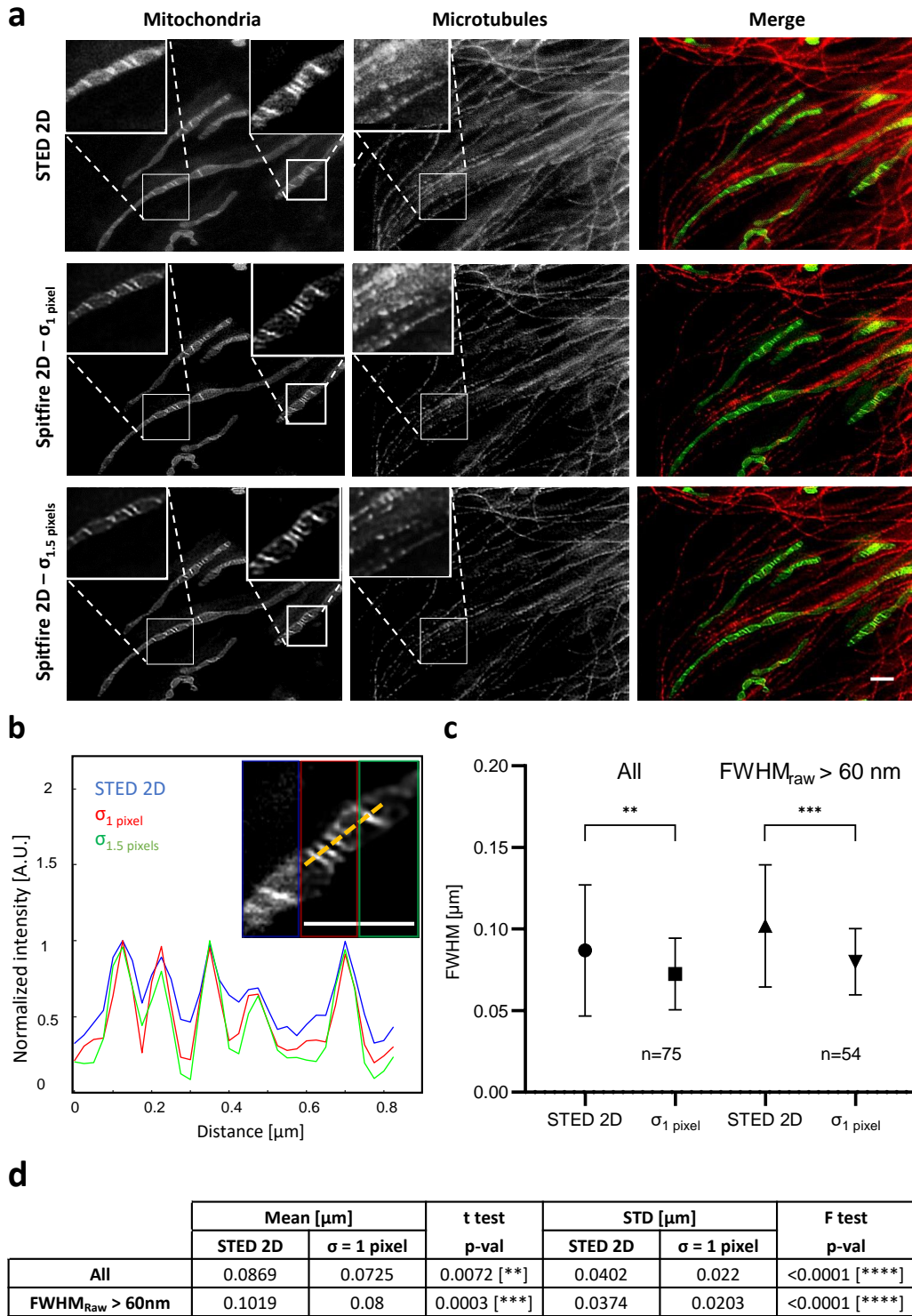
Fig. 2: **SPITFIR(e) increases the signal-to-noise of low-light images in STED imaging**. Live RPE1 cells double stained with PKMO (mitochondria) and Tubulin TrackerTM Deep Red (microtubules) were imaged by STED nanoscopy before and after being processed by SPITFIR(e) (**a**). Partial overview of 2D STED raw data of RPE1 cells for mitochondria, microtubules and their superimposed images are shown from left to right (**a**, upper panel). 2D STED images were improved by SPITFIR(e) using two different 2D Gaussian model $\sigma_{xy} = 1.0$ pixel (**a**, middle panel) and $\sigma_{xy} = 1.5$ pixels (**a**, lower panel). Insets (1 and 2) show zoomed area where SPITFIR(e) image quality improvement is illustrated by comparison of two processing conditions (middle and lower panel) with the same area in the raw STED image. Inset(**b**) shows a magnified composite image where the 2D STED raw part is indicated as a blue lined rectangle and 2D denoised + 2D deconvolved parts are indicated in red ($\sigma_{xy} = 1.0$ pixel) and green ($\sigma_{xy} = 1.5$ pixels). (**b**) Normalized intensity line profiles were measured in the insets (**b**) for cristae regions. Yellow line indicated in the inset (**b5**) serve to identify fluorescence profiles. Pixel size is equal to 25 nm. Scale bars: 1 μm in (**a**) and (**b**). Line profiles of individual cristae (n=75, from 8 different acquisitions) were fitted using a Gaussian model. Full width half maximum (FWHM) was estimated on raw STED and 2D denoised + 2D deconvolved with $\sigma_{xy} = 1.0$ pixel (**c**). SPITFIR(e) improvement is shown separately for all the line profiles analyzed and the line profiles from bigger cristae (FWHM$_{raw}$ = 60 nm), indication of lower SNR. Statistics analysis is including in (**d**). It indicates the improvement of average resolution (unpaired two-sample t-test) and variance of cristae resolution (F-test).

4

specificity and sensitivity. The PSF shape depends on several factors such as optics, laser power, fluorescence response of the specimen, and thermal drift which can add deformation by sheering. In Fig. 2, we deconvolved with SPITFIR(e) two 2D images depicting mitochondria and their alignment on microtubules, by applying SPITFIR(e) 2D with a 2D Gaussian PSF model and two different bandwidths σ_{xy} of size 1.5 pixels and 1.0 pixel respectively, respectively (Fig. 2(**a**)). We selected two regions of interest (ROI) in the mitochondria channel to better appreciate the restoration results in (Fig. 2(**a, b**)). Normalized fluorescence intensity line profiles, positioned at two distinct cristae regions of mitochondria (Fig. 2(**b**)) show the dual improvement of SPITFIR(e) with both an increase in the SNR and to a lower extend, an enhanced lateral resolution. Two simulated PSFs sizes were proposed to evaluate the capacity of SPITFIR(e) to improve the image resolution. The first PSF ($\sigma_{xy} = 1.0$ pixel) was similar than the expected experimental PSF of our microscope and the second PSF ($\sigma_{xy} = 1.5$ pixels) corresponds to the situation where there is an issue of defocusing. Clearly in both cases we observe an image improvement of the two individual cristae line profiles. The fitting of a Gaussian model to intensity profiles ($n = 75$, from 8 different acquisitions) allowed us to estimate the resolution at FWHM of raw and restored images (**b-c**). The obtained resolution is statistically improved when SPITFIR(e) is applied (with a smaller variance), as reported in **d**, and agrees well with the reported electron microscopy results[30].

SPITFIR(e) reliably restores weak fluorescent signals under extreme low-light conditions. Additional demonstrations of the ability of SPITFIR(e) to retrieve the high spatial-frequency information under extreme light conditions in 3D multifocus microscopy and spinning disk confocal microscopy, are presented below. Here, we focus on the performance limits of 3D deconvolution knowing that it is possible to improve the results by denoising the images beforehand as described above.

First, we tested SPITFIR(e) to visualize the 3D dynamics of mitochondria in living cells under extreme low-light conditions with multifocus (MFM) for different SNR values by reducing the excitation light dose (see Fig. 3(**a**)). In this experiment, each 3D stack was restored with a 3D Gaussian PSF model. The results on 2D plane (6th plane in (**a**)) are shown in Fig. 3 for six different SNRs values using the same sample. We considered three different levels of sparsity: "highly sparse" , "moderately sparse", and "weakly sparse". In Fig. 3(**b**) (bottom), we reported the corresponding Fourier Shell Correlation curves computed by selecting the image with the higher SNR value (SNR 6) as reference image. We can notice that for SNR 5 we get a good correlation in high frequency and the restored images are visually very similar to images with SNR 6. For SNR 4 to SNR 2, we can observe that SPITFIR(e) gradually looses correlation in high frequencies, and the reconstructed images are visually less accurate. For very poor SNR conditions (SNR 1), the images are very noisy and SPITFIR(e) (3D deconvolution) was not able to recover the signals and the high frequencies correlation is close to zero. Meanwhile, the images obtained by applying RL deconvolution did not allow us to retrieve information at SNR 4 and below. Moreover, artifacts and spatial distortions appear in the final image reconstructed when using RL deconvolution. Considering a sparse-promoting regularization helped here reveal structural details at lower SNR values. Overall, the most satisfying results with SPITFIR(e) are obtained with "moderate" sparsity.

Next, we evaluated the performance of SPITFIR(e) on 3D spinning disk confocal images acquired with different speeds and excitation light doses to capture mast movements and reduce photobleaching and photo-toxicity, respectively. To assess the overall sensitivity of SPITFIR(e), 4D images were captured at variable excitation light levels, reducing light intensity by a factor of 2 and 4. When we reduce by

half the laser power (raw data denoted $I_{1/2}$ in Supplementary Fig. 5), the bleaching effect is less visible and the SNR is lower. Furthermore, when we reduce the laser power by factor 4 (raw data denoted $I_{1/4}$ in Supplementary Fig. 5) we do not observe bleaching effect anymore. In Supplementary Fig. 5 (**b**), the three time-lapse temporal series display reduced bleaching curves which correlates with the excitation light intensity levels and SNR values. Nevertheless, the reduction factor has the undesirable consequence of lowering the signal-to-noise ratio (SNR) of the image, resulting in a poor-quality image which is also degraded by optics. We applied the proposed SPITFIR(e) restoration confocal pipeline on the three 3D+Time images I, $I_{1/2}$, $I_{1/4}$. The results shown in Fig. 5 are named SPITFIR(e)(I), SPITFIR(e)($I_{1/2}$), SPITFIR(e)($I_{1/4}$). We can notice that even with a low SNR ($I_{1/4}$) we obtain a good enough quality restored images for spot detection. We conclude that that even at low illumination doses and low SNR, we can obtain a good enough quality restored images with SPITFIR(e), which is helpful to preserve the integrity of samples, and till get exploitable data for quantitative analysis.

SPITFIR(e) quantitatively outperforms state-of-the-art algorithms. In fluorescence microscopy, restoration algorithms were dedicated to remove Poisson-Gaussian noise originating from the low-photons regimes (Poisson noise) and dark current induced by the electronic imaging detectors (Gaussian noise). In Methods (and Supplementary Notes (Appendix A)), we formally demonstrate that a quadratic fidelity term combined with any regularization term is optimal to handle Poisson-Gaussian noise. This new result allows us to apply SPITFIR(e) on images acquired with low light doses (Fig. 3 and Supplementary Fig. 5) and to quantitatively compare state-of-the-art deconvolution algorithms mainly developed for Gaussian noise removal, and applied here to fluorescence microscopy images.

First, we fairly benchmarked the performance of a large collection of competitive deconvolution algorithms given in Table 1, including the popular deconvolution methods such as RL, iterative constrained Tikhonov-Miller (ICTM), Gold-Meinel (GM)[32], and applied on artificially noisy and blurred 2D images. We considered four ground-truth images shown in Supplementary Table 2 which were blurred with 2D Gaussian PSF model with different standard deviation values σ_{xy}. A Gaussian noise with zero mean and variance τ^2 was also added to these images in order to generate observed noisy and blurry data. In Supplementary Table 2, we reported the best possible PSNR values for each method (with optimal parameter adjustment) and the true PSF size σ_{xy} and noise variance τ^2. In this experiment, RL and GM algorithm were often ranked at the end of the benchmark, while SPITFIR(e) provided the best PSNR values in most of cases. In terms of visual assessment, we noticed noise amplification with RL (Fig. 6(**c**) and GM (Fig. 6 (**d**), resulting in the apparition of undesired high-intensity pixels (i.e., the "night sky" effect) and unrealistic reconstructed structures since the high-frequency components were not correctly restored. ICTM produced deconvolution results without noise amplification artifacts but tends to over-smooth image details due to the quadratic nature of Tikhonov-Miller penalty, while TV (**f**), TV-L_1 (**g**) and SV (**l**) regularizers generate artificial sharper edges (staircasing effects), respectively. HV (**m**) generates visually more pleasant deconvolution results. Regarding smooth-approximation based regularizations, LSHV (**j**) provide results with restored details which are slightly sharper than those obtained with ICTM (**e**), but noise is not sufficiently removed. Finally, SPITFIR(e) (SHV, **n**) avoids over-smoothing and over-sharpening effect while preserving details and removing noise in the background. These quantitative results were confirmed on 3D microscopy images by considering additional performance criteria.

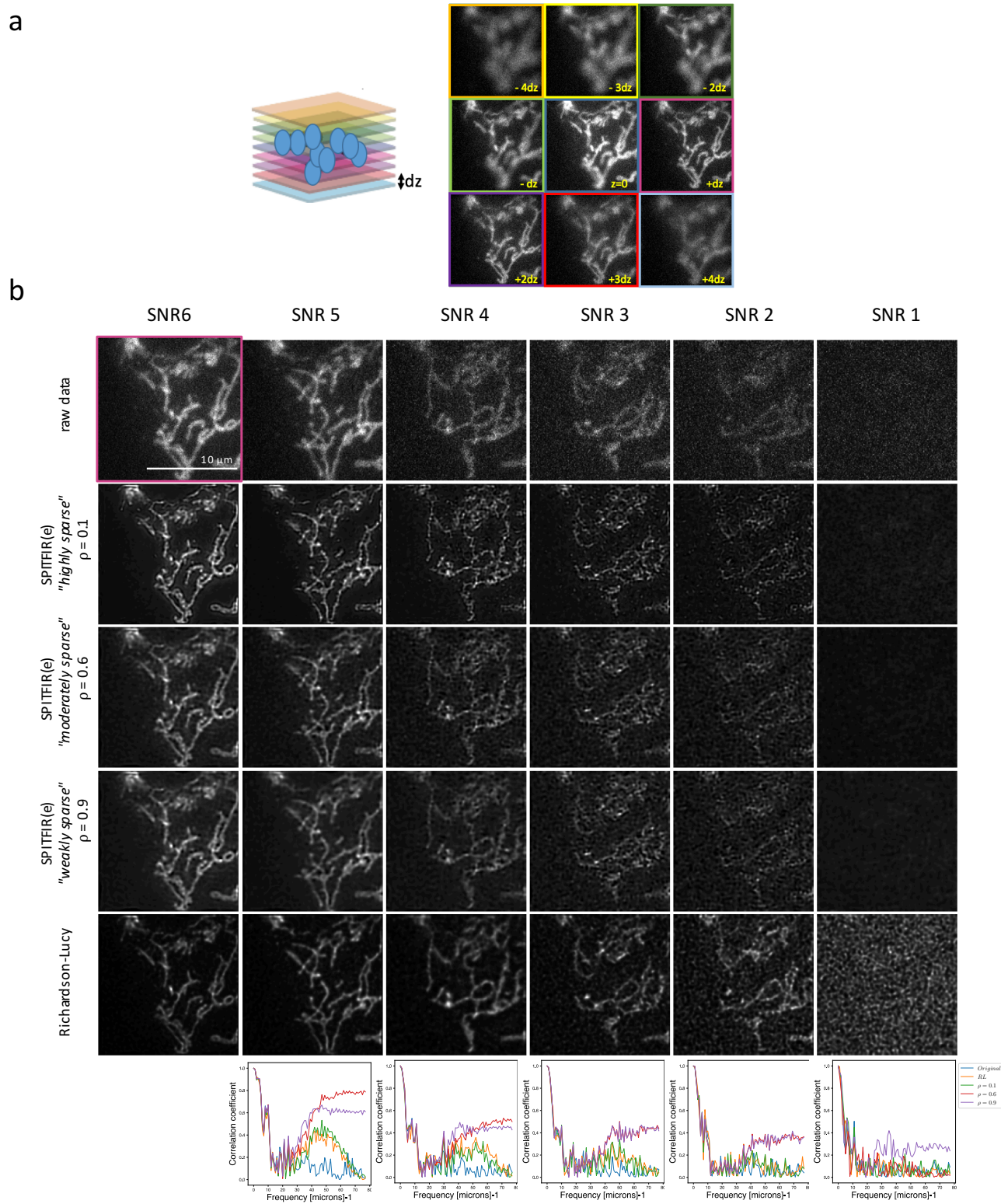As the implementation in 3D of most of aforementioned regular-

Fig. 3: **A comparison of wide-field multifocus microscopy image reconstruction with different amounts of sparsity and SNR values.** (**a**) MFM allows simultaneous acquisition of a 3D stack of nine images equally spaced with (dz = 330 nm) focal 2D images (dx, dy = 120 nm) with a frame rate up to 100 images per second, thus covering the whole volume of mitochondria in a single 3D image[31]. The set-up is based on a custom diffraction grating that forms multiple focus-shifted images followed by a chromatic correction module. (**b**) The image grid display the deconvolution results on the 6th plane for each SNR value (+dz in **a**)) of the wide-field MFM stack depicting mitochondria in U2OS cells transfected with TOM20 (translocase of outer mitochondrial membrane) fused to GFP (GFP-TOM20). The 3D stack has been deconvolved with 3D Gaussian PSF model ($\sigma_{xy} = 1.5$ pixels and $\sigma_z = 1.5$ pixel) and with three different levels of sparsity ("weak", "moderate", and "high"). The plots in the bottom are the Fourier Shell Correlation of the reconstructed image for each SNR value, by considering the higher SNR value (SNR 6) as reference image. Scale bar: 10 μm.

izers in Table 1 is not always possible, we focused on the Hessian-regularized deconvolution algorithm [13] (HR-C3D) as it was demonstrated to outperform previous software such as Huygens, Microvolution [33], and ER-Decon (LSHV) [14]. RL deconvolution is also considered and serves here as a baseline method. A first comparison of SPITFIR(e), HR-C3D and RL applied to LLSM imaging (see Supplementary Fig. 2) has been previously described. For the sake of objectivity, we present here the results obtained on a three-channel volume (22 planes) depicting actin (green channel, stained with Alexa Fluor 488 phalloidin), nucleus (blue challenge, stained with DAPI) and mitochondria (red channel, stained with MitoTracker Red CMXRos), acquired with a wide-field microscope and already used for algorithm evaluation by Ikoma *et al.* [13]. In Supplementary Fig. 7, we report the normalized root-mean-square error (NRMSE), global Pearson correlation coefficients (RSP), PSNR values, and Fourier Shell Correlation (FSC) curves [14], between the restored wide-field images obtained with the two competing methods (SPITFIR(e), HR-C3D) and the reference spinning disk confocal image deconvolved with the Huygens software [13], for each channel. We carefully followed the recommendations given in ref [13] to fairly normalize the deconvolved images with respect to the L_2 norm of the deconvolved confocal image (with offset correction). The good performance of SPITFIR(e) is first confirmed by the PSNR/NMSE values which are slightly higher/smaller than those obtained with HR-C3D [13]. The resolution estimated from FSC curves is also higher (nucleus, actin) or comparable (mitochondria). This is especially demonstrative at high frequencies in the case of actin filaments and cell nucleus. SPITFIR(e) produces visually smoother images than HR-C3D but more correlated to the spinning disk confocal microscopy images

Finally, we conducted experiments on a 3D confocal microscopy and compared the results of RL, SPITFIR(e) and HR-C3D to those produced by CARE [34], a supervised deep-learning-based deconvolution method (Supplementary Fig. 8). As CARE is learned from a training set of dedicated biological structures, we evaluate the algorithms on the stack provided by the authors [34] and depicting the envelopes of nuclei stained with GFP-LAP2b (developing eye of zebrafish (Daniorerio) embryos. We focused on the nuclear envelop as this channel contains more structural details. In Supplementary Figs. 8 we display for the sake of visual assessment the restoration of 5th plane of an anisotropic 3D stack composed 18 images [34]. The reconstruction quality is visibly superior with CARE as expected. Nevertheless SPITFIRE(e) with 3D Gaussian PSF ($\sigma_{xy} = 2.0$ pixels, $\sigma_z = 0.5$ pixel) and Gibson-Lanni model [35] (truncated to 3 planes around the center of the PSF) was able to provide satisfactory results. The main difference are visible on small filaments better enhanced with CARE. When we plot an intensity profile (**b**), we can see that it is the SPITFIR(e) method that gives the best compromise between pick thickness and rebound effect around the contours. The results can be visually improved with SPITFIR(e) if the volume is denoised beforehand and deconvolved with the PbyP strategy (Supplementary Fig. 8 (**a**), right).

Discussion

An important challenge in live fluorescence imaging is to limit photodamages on the observed sample by reducing the dose of excitation light, while at the same time collecting enough photons to produce informative images for quantitative analysis. To overcome the difficulties of acquisition of 2D+Time and 3D+Time images at low excitation dose and boost the signal corrupted by blur and different sources of noise (shot noise, readout noise), we developed SPITFIR(e), a supermanoeuvrable and easy-to-use solution able to generate spatially consistent and high-quality fluorescence images. The approach is based on the fast minimization of a convex energy composed a novel

sparse-promoting regularization term and a Poisson-Gaussian-aware quadratic fidelity term. Unlike previous methods which were tested on a few microscopy modalities [13;14;21] and need a careful adjustment of regularization parameters to obtain high-quality restoration results, SPITFIR(e) is a nearly automatized method which adapts to any microscopy technique, and image dimensionality. With this unique algorithm, the practitioner can switch from 4D image denoising to 3D deconvolution, or can chain the two processing tasks to obtain high-quality images. He only needs to specify the level of desired sparsity ("low", "moderate", "high") and to supply the measured PSF or the parameters of the 2D or 3D theoretical PSF (e.g., Gaussian) model.

We demonstrated on experimental data that SPITFIR(e) extracted high frequency information and achieved significant improvement in the output quality even with a poor calibrated PSF. Nevertheless, SPITFIR(e) could miss some details in noisy images taken at low excitation light intensity or short exposure time. As the lower light intensities approaches the limits of ability of SPITFIR(e) to recover information reliably, a successful strategy with consists in chaining denoising (3D or 4D) and deconvolution with a 2D or 3D PSF model to make the labeled biomolecules more discernible from noise. In presence of Poisson shot noise and readout noise, our results demonstrated that the proposed strategy is robust and able to boost signals reliably, and confirms that the quadratic fidelity term is suitable for Poisson-Gaussian noise as theoretically proved (Methods).

We have showed that SPITFIR(e) substantially reduces noise and provides better high-quality images than other currently available methods over a wide variety of experimental conditions. In the particular case of LLSM, which is designed already for long lasting fast imaging of 3D large and close to isotropic volumes, SPITFIR(e) shows all its power, since it not only improves resolution and reduces noise, but still allows an increase in the imaging rates by permitting to reduce the illumination dose. SPITFIR(e) also achieved high resolution in subdiffracted and superresolution microscopy under low-light conditions, without perturbing the signal amplitudes. For two-dimensional live-cell STED imaging, line intensity profiles along mitochondria cristae showed that SPITFIR(e) reached an improved lateral resolution of order 1.5. Little is known about the actual dynamics of mitochondrial cristae in the multiple physiological processes to which mitochondria activity is required. This was mainly due to the visualization challenge of cristae structure in living cells in the past, both due to the lack of optical resolution and to the deleterious effect of light on mitochondria. Hessian-SIM modality [36] or STED nanoscopy [37;38] are the only approaches so far, able to achieve this goal. While only unique 2D STED images of live mitochondria are shown here, SPITFIR(e) applied on STED images might be of strong interest in this context since it allows to restore these fine structures even when depletion is incomplete and scanning imaging at very low illumination, making live relatively long term imaging accessible.

More generally, unlike smoothness regularizers, our sparse-promoting regularizer is able to reveal small details in the raw data, while preserving signal peaks and removing background, with no discernible artifacts in the final restored images. We also demonstrated that on 2D/3D benchmarks that SPITFIRE(e) outperforms existing deconvolution algorithms. Performance has been quantitatively studied in terms of PSNR values on artificially degraded 2D images and in terms of FRC on 3D experimental data. Finally, we showed that SPITFIR(e) is able to produce comparable results to supervised machine-learning techniques [34]. Unlike content-dependent methods based on deep-learning algorithms requiring many training images, SPITFIR(e) is an unsupervised method which utilizes sparsity and smoothness as priori knowledge and adapts to any image contents.

In conclusion, our results on experimental data endorse that SPITFIR(e) may be considered as a "swiss-knife", able to han-

dle different samples captured under various fluorescence micro-scopes, adapt to various sources of signal degradation, image contents, and variable signal-to-noise ratios. This algorithm does not require a well-calibrated PSF and is easily controlled by choosing three levels of desired sparsity, that is "high", "medium", and "low". We are convinced that SPITFIR(e) will be very helpful to push the spatiotemporal resolution limit of sub-diffracted and superresolution fluorescence microscopy techniques and will better resolve 3D structures and dynamics of biological components observed at low excitation light intensity to reduce artifacts in live cells. Our method was implemented as an Napari plugin (`https://www.napari-hub.org/plugins/napari-sdeconv`) for ease of use. It is also available as open source C++ code (CPU and GPU).

Methods

Variational model for image restoration. An observed 2D/3D image $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}, d = 2, 3$ is a blurred and noisy version of the underlying true image $u : \Omega \to \mathbb{R}$ (i.e., $u \in \mathbb{R}^\Omega$). A general approach for restoring an image consists in finding the minimizer of an energy functional, i.e.,

$$\hat{u} = \arg \min_{u \in \mathbb{R}^\Omega} F(f, u; h, \mathcal{T}) + \lambda R_\rho(u) \qquad (1)$$

where $F(f, u; h, \mathcal{T})$ and $R_\rho(u)$ are the data fidelity term and the regularization term, respectively, and $\lambda > 0$ and $\rho \in [0, 1]$ are the parameters controlling the amount of smoothness and sparsity in the image u. The fidelity data term is the distance between the restored image u and the observed image f where h is the PSF function and $\mathcal{T}$ is the degradation operator that encompasses Poisson and Gaussian noise. In what follows, we give the explicit expressions of the two terms.

Data fidelity term. A fidelity data term is generally derived from the general formation model, for instance dedicated to low photon counts and low-light regimes in fluorescence microscopy [13]. In our approach, $u \in \Omega$ is assumed to be corrupted by a mixed Poisson-Gaussian. The local noise variance is then represented at a given location $\mathbf{x} \in \Omega$ as follows [39;40;60] (see details in Supplementary Notes (Appendix A)):

$$V_{noise}(\mathbf{x}) \triangleq \alpha^{-1} v(\mathbf{x}) + \tau^2 = \mathbb{E}[(f(\mathbf{x}) - (h * u)(\mathbf{x}))^2], \qquad (2)$$

where f is the noisy image, $v = h * u$, $\mathbb{E}[\cdot]$ denotes the mathematical expectation, $\alpha > 0$ is the quantization factor of the photodetector, and $\tau^2 > 0$ represents the Gaussian noise variance. If we assume that the average intensities in v and f are preserved, it follows that:

$$\int_\Omega \alpha^{-1} v(\mathbf{x}) + \tau^2 \, d\mathbf{x} = \int_\Omega \alpha^{-1} f(\mathbf{x}) + \tau^2 \, d\mathbf{x} \triangleq \alpha^{-1} \|f\|_1 + |\Omega| \tau^2$$

$$= \int_\Omega \mathbb{E}[(f(\mathbf{x}) - (h * u)(\mathbf{x}))^2] d\mathbf{x} \qquad (3)$$

$$= \int\int_\Omega (f(\mathbf{x}) - (h * u')(\mathbf{x}))^2 p(u'(\mathbf{x})) \, d\mathbf{x} \, du'(\mathbf{x})$$

where $p(u'(\mathbf{x}))$ represents the probability distribution of u. If we assume the following prior $p(u'(\mathbf{x})) = \mathbf{1}[f(\mathbf{x}) = ((h * u)(\mathbf{x}))]$ where $\mathbf{1}[\cdot]$ is the indicator function, we finally obtain:

$$\int_\Omega (f(\mathbf{x}) - (h * u)(\mathbf{x}))^2 d\mathbf{x} = \|f - (h * u)\|_2^2 = \alpha^{-1} \|f\|_1 + |\Omega| \tau^2. \quad (4)$$

Starting from the seminal paper [41], the restored image is found by solving the following optimization problem (Poisson-Gaussian noise):

$$\min_u R(u) \quad \text{subject to} \quad \|f - (h * u)\|_2^2 = \alpha^{-1} \|f\|_1 + |\Omega| \tau^2. \qquad (5)$$

It turns out that (5) is a constrained formulation with an equality constraint which is not convex. The corresponding formulation with inequality constraint is

$$\min_u R(u) \quad \text{subject to} \quad \|f - (h * u)\|_2^2 \leq \alpha^{-1} \|f\|_1 + |\Omega| \tau^2. \qquad (6)$$

It has been established that, under additional assumptions, that (5) and (6) are equivalent. A Lagrange formulation can be then derived since the the right-hand side of the inequality does not depend on the unknown image u:

$$\min_u R(u) + \lambda' \|f - (h * u)\|_2^2 \Leftrightarrow \min_u \|f - (h * u)\|_2^2 + \lambda R(u). \quad (7)$$

where the parameter $\lambda' > 0$ balances the two energy terms with $\lambda' = \lambda^{-1} \neq 0$ (see Supplementary Notes (Appendix A) for details). In conclusion, a quadratic fidelity term $F(\cdot)$ is appropriate for Gaussian-noise removal provided it is combined with regularizer $R(\cdot)$.

Regularization term. Commonly-used regularizers in image restoration have the following form:

$$R(u) = \int_\Omega \phi(Du(\mathbf{x})) \, d\mathbf{x}, \qquad (8)$$

where D is a linear operator (called "regularization operator") used to control the spatial distribution of u and $\phi(\cdot)$ is a positive potential function usually related to a norm distance. A typical example is the Tikhonov-Miller regularization [8] defined as the squared norm of the gradient of u: $\phi(Du(\mathbf{x})) = \|\nabla u(\mathbf{x})\|_2^2$.

We propose a convex sparsity-promoting regularizer based on second-order derivatives to avoid the emergence of undesirable stair-casing effects. The Sparse Hessian Variation (SHV) regularizer is defined as follows

$$\text{SHV}_\rho(u) = \int_\Omega \underbrace{\sqrt{\rho^2 \|\mathcal{H}u(\mathbf{x})\|_F^2 + (1 - \rho)^2 u(\mathbf{x})^2}}_{\|D_{2,\rho} u(\mathbf{x})\|_2} \, d\mathbf{x}, \qquad (9)$$

where $\|\cdot\|_F$ denotes the matrix Frobenius norm and $\rho \in [0, 1]$ is the weighting parameter. Moreover, since the Frobenius norm of the Hessian matrix is equal to the Euclidean norm of its vectorized version, the operator $D_{2,\rho} \in \mathbb{R}^{10}$ is defined (for $d = 3$) as:

$$D_{2,\rho} u(\mathbf{x}) \triangleq \left((1 - \rho)u(\mathbf{x}), \rho \frac{\partial^2 u(\mathbf{x})}{\partial_{xx}}, \rho \frac{\partial^2 u(\mathbf{x})}{\partial_{yy}}, \rho \frac{\partial^2 u(\mathbf{x})}{\partial_{zz}}, \rho \frac{\partial^2 u(\mathbf{x})}{\partial_{xy}}, \right.$$

$$\left. \rho \frac{\partial^2 u(\mathbf{x})}{\partial_{xz}}, \rho \frac{\partial^2 u(\mathbf{x})}{\partial_{yz}}, \rho \frac{\partial^2 u(\mathbf{x})}{\partial_{yx}}, \rho \frac{\partial^2 u(\mathbf{x})}{\partial_{zx}}, \rho \frac{\partial^2 u(\mathbf{x})}{\partial_{zy}} \right)^T. \quad (10)$$

The idea behind this combination is to sparsify jointly the spatial distribution of image intensities and the second order derivatives of the image to encourage smooth variations between spatially-contiguous non-zero regions of the underlying image. The resulting image contain bright objects against a large smooth dark background, and no spurious edge. Unlike LSHV [14], SHV is convex and more flexible as it involves a parameter $\rho \in [0, 1]$ which is helpful to eliminate or preserve background. In SPITFIR(e), we consider only three values to control the amount of sparsity: $\rho = 0.1$ ("sparse"), $\rho = 0.6$ ("moderately sparse"), and $\rho = 0.9$ ("weakly sparse").

Discrete formulation and optimization. We present here the discretization setting required to derive the optimization algorithm. Let us consider a sampling grid of the following form in 3D ($d = 3$) (expressions in 2D are similar).

$$\Lambda = \mathbb{Z}^3 \cap \Omega = \{1, 2, \dots, N_x\} \times \{1, 2, \dots, N_y\} \times \{1, 2, \dots, N_z\}.$$

The observed noisy and blurry image f is represented by its digitized (discrete) version as follows: $f = \mathcal{T}(Hu)$, where $f, u \in \mathbb{R}_+^N$ with $N = N_x \times N_y \times N_z$, $H \in \mathbb{R}^{N \times N}$ is a matrix that models the point spread function of the microscope in the discrete setting, and $\mathcal{T}$ is degradation operator. Both u and f are assumed to be non negative. In the discrete setting, the blurring operator H corresponds to a discrete convolution which can be efficiently computed by using fast Fourier transform (FFT) [42;43;44;45]. To discretize $D_{2,\rho}$, we use standard finite differences for the Hessian operators with Neumann conditions on image boundaries. The 2D-3D deconvolution problem is then defined in the discrete setting as the minimizer of the following energy:

$$E(u) = \frac{1}{2} \|Hu - f\|_2^2 + \lambda \|D_{2,\rho}u\|_2 + \imath_{\mathcal{C}}(u), \qquad (11)$$

where $\lambda > 0$ is the regularization parameter and $\imath_{\mathcal{C}}$ is the characteristic function of the convex subset $\mathcal{C}$ defined as: $\imath_{\mathcal{C}} = 0$ if $u \in \mathcal{C}$ and $+\infty$ otherwise, is very helpful to impose positivity constraint on the solution.

The optimization problem (11) is convex since the underlying energy functional is defined as the sum of convex terms, but it is non-smooth. To minimize the energy (11), several algorithms have been investigated, including FISTA [46], ADMM [22;23] and MM [47] algorithms. Instead of applying the aforementioned optimization methods, we investigated a first-order method to

Name	Regularizer $\phi(Du(\mathbf{x}))$	Differentiation	Convexity order	Smoothness
TV-L$_{1,\rho}$ [19]	$\rho\|\nabla u(\mathbf{x})\|_2 + (1-\rho)|u(\mathbf{x})|\, d\mathbf{x}$	1	Convex	Non smooth
GraphNet (GN$_\rho$) [15]	$\rho\|\nabla u(\mathbf{x})\|_2^2 + (1-\rho)|u(\mathbf{x})|$	1	Convex	Non smooth
Log Sparse Hessian Variation (LSHV$_\epsilon$) [14]	$\log(\epsilon + u^2(\mathbf{x}) + \|\mathcal{H}u(\mathbf{x})\|_F^2)$	2	Non convex	Smooth
Sparse Variation (SV$_\rho$) [18]	$\sqrt{\rho^2\|\nabla u(\mathbf{x})\|_2^2 + (1-\rho)^2 u(\mathbf{x})^2}$	1	Convex	Non smooth

Table 1: Properties of regularizers used for restoring fluorescence microscopy images.

SPITFIR(e) 200 iterations	Image size	CPU Intel Xeon (2.8 GHz, 4 threads)	GPU Nvidia Quadro M2000	Nvidia Tesla K80
2D Denoising	512×512	349 ms	345 ms	278 ms
3D Denoising	$512 \times 512 \times 22$	8 s 74 ms	3 s 78 ms	1s 277ms
4d Denoising	$512 \times 512 \times 14 \times 32$	8 min 23 s 789 ms	1 min 40 s 98 ms	45 s 744 ms
2D Deconvolution	512×512	767 ms	550 ms	428 ms
3D Deconvolution	$512 \times 512 \times 22$	23 s 574 ms	5 s 786 ms	2 s 713 ms

Table 2: Computing times of SPITFIR(e) (200 iterations for energy minimization).

minimize the sum of convex functions, based on the proximal splitting approaches [25;26;48;49;50;51;52;53]. The main idea consists in splitting the original problem into several simple sub-problems in the way that each single function of the sum can be processed separately (see details in Supplementary 1C). These operators are well-suited for large-scale problems arising in signal and image processing, because they only exploit first-order information of the function and thus enable fast and efficient computation. To solve the minimization problem (11), we adapted the full splitting approach described in [25;26]. The key idea is to evaluate the gradient, proximity and linear operators individually in order to avoid implicit operations such as inner loops or inverse of linear operators.

Self-tuning of the regularization parameter. The adjustment of the regularization parameter λ in (11) is generally crucial in most image restoration algorithms, and may be time consuming. A non optimal choice of this parameter may over-smooth object borders, suppress structural details, generate artifacts or weakly reduce noise. Therefore, we investigated a self-tuning strategy which is based on the comparison of restorations and performance criteria over a range of parameters λ. The practical issue is to automatically tune this parameter on a case-by-case basis to get the best performance as possible, from the input noisy image. To address this issue, we propose a self-tuning approach based on the analysis of one or several informative regions of interest (ROI) in the input noisy image. The ROI can be the entire image or a $p \times p$ patch containing significant fluorescent signal and randomly drawn or manually selected by the user in the input image.

Instead of applying the principle of cross-validation, we investigated here an approach based on the minimax principle which avoids the selection of low and high values for the regularization parameter [54]. First the energy (1) is rewritten as follows:

$$E_{\lambda'}(f,u) = (1-\lambda')F(f,u;h,\mathcal{T}) + \lambda' R_\rho(u) \quad (12)$$

where $\lambda' \in]0,1[$ is a positive constant such that $\lambda' = \lambda(1+\lambda)^{-1}$. The Min-Max principle amounts to finding $u^\star$ that minimizes, over all λ' values, the maximum of $E_{\lambda'}(f,u)$:

$$
\begin{aligned}
(u^\star, {\lambda'}^\star) &= \min_u \max_{\lambda'} E_{\lambda'}(f,u) \quad (13)\\
&= \max_{\lambda'} \min_u E_{\lambda'}(f,u) \stackrel{\triangle}{=} \max_{\lambda'} E_{\lambda'}(f,u^\star).
\end{aligned}
$$

The optimization problem may be solved by examining the finite set of solutions $E_{\lambda'}(f,u^\star)$ and selecting the λ' valued associated to the highest energy $E_{\lambda'}(f,u^\star)$. The good news is that Gennert and Yuille [54] demonstrated that there exists a unique value λ' that maximizes $E_{\lambda'}(f,u^\star)$ (convexity of $E_{\lambda'}(f,u^\star)$). Hence, by initializing $\lambda' = 0.5$ and considering a finite set $\{\lambda'_m\}_{m=1}^M$ of increasing values such that $\lambda'_{\lfloor M/2\rfloor} = 0.5$, a fast search technique can be developed which consists in restoring the image with

$\lambda' = \lambda'_{\lfloor M/2\rfloor+1}$ and $\lambda'_{\lfloor M/2\rfloor-1}$ and keeping the λ' corresponding to the highest energy $E_{\lambda'}(f,u^\star)$. The raw image is then restored with a highest (or lowest value) of λ', and the SPITFIR(e) procedure is repeatedly applied on raw data (with $\lambda = \lambda'(1-\lambda')^{-1}$) until the energy $E_{\lambda'}(f,u^\star)$ decreases (or increases).

In practice, the level of sparsity ("sparse", "moderately sparse", "weakly sparse") is first selected by the practitioner beforehand. The parameter λ' is then estimated by applying the aforementioned optimization procedure to a single or multiple 2D ROIs (or patches) located in informative areas, that is in areas containing meaningful fluorescent signals, as illustrated in Supplementary Fig. 9.

Implementation details and code availability

Images have been processed on classical workstation for CPU calculation (CPU Intel Xeon 2.8GHz 4 threads). We use this setup since it is the most common setup in biology labs. For high speed processing we developed a GPU version of SPITFIR(e). We run the GPU on two types of GPU, a Nvidia Quadro M2000 that can be found in most workstations and a Nvidia Tesla K80 for computing grid. Typical computing times (CPU and GPU) for processing 2D-3D-4D images are reported in the table below: These computing times are small and take into account the full data processing steps (including data normalization and data copy to the GPU device). The code can be downloaded for free from the GitLab website (https://github.com/sylvainprigent/simglib) along with accompanying documentation.

Data source

See description in Appendix C (Microsopy and Cell Cultures), Supplementary Information.

References

[1] Carlton, P. *et al.* Fast live simultaneous multi-wavelength 4-dimensional optical microscopy. *Proc. Natl. Acad. Sci. USA* **107**, 16016–16022 (2010).

[2] Boulanger, J. *et al.* Patch-based nonlocal functional for denoising fluorescence microscopy image sequences. *IEEE Trans. Medical Imaging* **29**, 442–454 (2010).

[3] Richardson, W. H. Bayesian-based iterative method of image restoration. *J. Optical Society of America* **62**, 55–59 (1972). URL http://www.osapublishing.org/abstract.cfm?URI=josa-62-1-55.

[4] Lucy, L. B. An iterative technique for the rectification of observed distributions. *The Astronomical J.* **79**, 745–754 (1974).

[5] Shepp, L. A. & Vardi, Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Medical Imaging* **1**, 113–122 (1982).

[6] Sibarita, J.-B. *Deconvolution Microscopy*, 201–243 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005). URL https://doi.org/10.1007/b102215.

[7] van der Voort, H. T. M. & Strasters, K. C. Restoration of confocal images for quantitative image analysis. *J. Microscopy* **178**, 165–181 (1995). URL http://dx.doi.org/10.1111/j.1365-2818.1995.tb03593.x.

[8] van Kempen, G. M. P., van der Voort, H. T. M., Bauman, J. G. J. & Strasters, K. C. Comparing maximum likelihood estimation and constrained Tikhonov-Miller restoration. *IEEE Engineering in Medicine and Biology Magazine* **15**, 76–83 (1996).

[9] van Kempen, G. M. P., van Vliet, L. J., Verveer, P. J. & van der Voort, H. T. M. A quantitative comparison of image restoration methods for confocal microscopy. *J. Microscopy* **185**, 354–365 (1997). URL http://dx.doi.org/10.1046/j.1365-2818.1997.d01-629.x.

[10] Dey, N. *et al.* ichardson-lucy algorithm with total variation regularization for 3d confocal microscope deconvolution. *Microsc. Res. Tech.* **69**, 260–266 (2006).

[11] Hom, E. *et al.* Aida: An adaptive image deconvolution algorithm with application to multi-frame and three-dimensional data. *J. Opt. Soc. Am. A* **24**, 1580–1600 (2007).

[12] Lefkimmiatis, S., Ward, J. P. & Unser, M. Hessian Schatten-norm regularization for linear inverse problems. *IEEE Trans. Image Processing* **22**, 1873–1888 (2013).

[13] Ikoma, H., Broxton, M., Kudo, T. & Wetzstein, G. A convex 3d deconvolution algorithm for low photon count fluorescence imaging. *Scientific Reports* **8**, 111489 (2018).

[14] Arigovindan, M. *et al.* High-resolution restoration of 3D structures from widefield images with extreme low signal-to-noise-ratio. *Proc. Natl. Acad. Sci. USA* **110**, 17344–17349 (2013). URL http://www.pnas.org/content/110/43/17344.abstract. http://www.pnas.org/content/110/43/17344.full.pdf.

[15] Ng, B., Vahdat, A., Hamarneh, G. & Abugharbieh, R. Generalized sparse classifiers for decoding cognitive states in fMRI. In Wang, F., Yan, P., Suzuki, K. & Shen, D. (eds.) *ProceedingsInt. MICCAI Workshop on Machine Learning in Medical Imaging (MLMI)*, 108–115 (Beijing, China, 2010).

[16] Kandel, B. M., Wolk, D. A., Gee, J. C. & Avants, B. Predicting cognitive data from medical images using sparse linear regression. In Gee, J. C., Joshi, S., Pohl, K. M., Wells, W. M. & Zöllei, L. (eds.) *Int. Conf. Information Processing in Medical Imaging (IPMI 2013)*, 86–97 (Asilomar, CA, USA, 2013).

[17] Grosenick, L., Klingenberg, B., Katovich, K., Knutson, B. & Taylor, J. E. Interpretable whole-brain prediction analysis with GraphNet. *NeuroImage* **72**, 304–321 (2013). URL http://www.sciencedirect.com/science/article/pii/S1053811912012487.

[18] Eickenberg, M., Dohmatob, E., Thirion, B. & Varoquaux, G. Grouping total variation and sparsity: Statistical learning with segmenting penalties. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 685–693 (Munich, Germany, 2015).

[19] Chan, T. F. & Esedoglu, S. Aspects of total variation regularized L1 function approximation. *SIAM J. Applied Mathematics* **65**, 1817–1837 (2005). URL https://doi.org/10.1137/040604297.

[20] Michel, V., Gramfort, A., Varoquaux, G., Eger, E. & Thirion, B. Total variation regularization for fMRI-based prediction of behavior. *IEEE Trans. Medical Imaging* **30**, 1328–1340 (2011).

[21] Zhao, W. *et al.* Sparse deconvolution improves the resolution of live-cell super-resolution fluorescence microscopy. *Nat. Biotechnol.* (2021).

[22] Afonso, M., Bioucas-Dias, J. & Figueiredo, M. An augmented lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Trans. Image Processing* **20**, 681–695 (2011).

[23] Boyd, S., Parikh, N., Chu, E., Peleato, B. & Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **3**, 1–122 (2007).

[24] Beck, A. & Teboulle, M. Fast iterative shrinkage-thresholding algorithm

[25] Condat, L. A primal–dual splitting method for convex optimization involving lipschitzian, proximable and linear composite terms. *J. Optimization Theory and Applications* **158**, 460–479 (2013). URL http://dx.doi.org/10.1007/s10957-012-0245-9.

[26] Condat, L. A generic proximal algorithm for convex optimization – application to total variation minimization. *IEEE Signal Processing Letters* **21**, 985–989 (2014).

[27] Chen, B.-C. & et al. Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Nature* **346**, 1257998–1, 1257998–12 (2014).

[28] Yang, Z. *et al.* Cyclooctatetraene-conjugated cyanine mitochondrial probes minimize phototoxicity in fluorescence and nanoscopic imaging. *Chem. Sci.* **11**, 8506–8516 (2020). URL http://dx.doi.org/10.1039/D0SC02837A.

[29] Renard, H. *et al.* Endophilin-a3 and galectin-8 control the clathrin-independent endocytosis of cd166. *Nat. Commun* **11** (2020).

[30] Lea, P. & Hollenberg, M. Mitochondrial structure revealed by high-resolution scanning electron microscopy. *Am J Anat.* **184** (1989).

[31] Hajj, B. *et al.* Whole-cell, multicolor superresolution imaging using volumetric multifocus microscopy. *Proc. Natl. Acad. Sci. USA* **111**, 17480–17485 (2014).

[32] Sage, D. *et al.* DeconvolutionLab2: an open-source software for deconvolution microscopy. *Methods* **115**, 28–41 (2017).

[33] Bruce, M. & Butte, M. Real-time gpu-based 3d deconvolution. *Opt. Express* **21**, 4766–4773 (2013).

[34] Weigert, M., Schmidt, U. & et al., T. B. Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nat. Methods* **15**, 1090–1097 (2018).

[35] Kirshner, H., Aguet, F., Sage, D. & Unser, M. 3-D PSF fitting for fluorescence microscopy: Implementation and localization application. *Journal of Microscopy* **249**, 13–25 (2013).

[36] Huang, X. *et al.* Fast, long-term, super-resolution imaging with hessian structured illumination microscopy. *Nat. Biotechnol.* **36**, 451–459 (2018).

[37] Till, S., Roesch, A., Riedel, D. & Jakobs, S. Live-cell sted nanoscopy of mitochondrial cristae. *Sci. Rep.* **9** (2019).

[38] Yang, X. *et al.* Mitochondrial dynamics quantitatively revealed by sted nanoscopy with an enhanced squaraine variant probe. *Nat. Commun* **11** (2020).

[39] Mugnier, L., Fusco, T. & Conan, J.-M. Mistral: a myopic edge-preserving image restoration method, with application to astronomical adaptive-optics-corrected long-exposure images. *J. Opt. Soc. Am. A* **21**, 1841–1854 (2004).

[40] Soulez, F., Denis, L., Tourneur, Y. & Thiébaut, É. Blind deconvolution of 3D data in wide field fluorescence microscopy. In *Int. Symp. Biomedical Imaging (ISBI)* (Barcelone, Spain, 2012).

[41] Rudin, L. I., Osher, S. & Fatemi, E. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena* **60**, 259 – 268 (1992). URL http://www.sciencedirect.com/science/article/pii/016727899290242F.

[42] Heideman, M., Johnson, D. & Burrus, C. Gauss and the history of the fast Fourier transform. *IEEE ASSP Magazine* **1**, 14–21 (1984).

[43] Van Loan, C. *Computational Frameworks for the Fast Fourier Transform* (Society for Industrial and Applied Mathematics, 1992). URL http://epubs.siam.org/doi/abs/10.1137/1.9781611970999. http://epubs.siam.org/doi/pdf/10.1137/1.9781611970999.

[44] Frigo, M. & Johnson, S. G. FFTW: an adaptive software architecture for the FFT. In *IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, vol. 3, 1381–1384 (1998).

[45] Frigo, M. & Johnson, S. G. The design and implementation of FFTW3. *Proceedings of the IEEE* **93**, 216–231 (2005).

[46] Beck, A. & Teboulle, M. Fast gradient-based algorithm for constrained total variation image denoising and deblurring problems. *IEEE Trans. Image Processing* **18**, 2419–2434 (2009).

[47] Figueiredo, M., Bioucas-Dis, J. & Nowack, R. Majorization-

minimization algorithms for wavelet-based image restoration. *IEEE Trans. Image Processing* **16**, 2980–2991 (2007).

[48] Mercier, B. *Lectures on Topics in Finite Element Solution of Elliptic Problems* (Springer-Verlag Berlin Heidelberg, 1979).

[49] Eckstein, J. & Bertsekas, D. P. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming* **55**, 293–318 (1992). URL `http://dx.doi.org/10.1007/BF01581204`.

[50] Combettes, P. L. Solving monotone inclusions via compositions of nonexpansive averaged operators. *Optimization* **53**, 475–504 (2004).

[51] Combettes, P. L. & Wajs, V. R. Signal recovery by proximal forward-backward splitting. *SIAM J. Multiscale Modeling and Simulation* **4**, 1168–1200 (2005). URL `https://doi.org/10.1137/050626090`.

[52] Combettes, P. L. & Pesquet, J.-C. *Proximal Splitting Methods in Signal Processing*, 185–212 (Springer New York, New York, NY, 2011). URL `http://dx.doi.org/10.1007/978-1-4419-9569-8_10`.

[53] Chambolle, A. & Pock, T. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging and Vision* **40**, 120–145 (2011). URL `http://dx.doi.org/10.1007/s10851-010-0251-1`.

[54] Gennert, M. & Yuille, A. Determining the optimal weights in multiple objective function optmization. In *Proc. IEEE Int. Conf. Comp. Vision (ICCV)* (1988).

[55] Boulanger, J. *et al.* Fast high-resolution 3d total internal reflection fluorescence microscopy by incidence angle scanning and azimuthal averaging. *Proc. Natl. Acad. Sci. USA* **111**, 17164–17169 (2014).

[56] Gidon, A. *et al.* Rab11A/MyosinVb/Rab11-FIP2 complex frames two late recycling steps of langerin from erc to plasma membrane. *Traffic* **13** (2012).

[57] Meinel, E. S. Origins of linear and nonlinear recursive restoration algorithms. *J. Optical Society of America A* **3**, 787–799 (1986). URL `http://josaa.osa.org/abstract.cfm?URI=josaa-3-6-787`.

[58] Nikolova, M. Weakly constrained minimization: Application to the estimation of images and signals involving constant regions. *J. Math. Imaging and Vision* **21**, 155–175 (2004). URL `https://doi.org/10.1023/B:JMIV.0000035180.40477.bd`.

[59] Good, I. J. & Gaskins, R. A. Nonparametric roughness penalties for probability densities. *Biometrika* **58**, 255–277 (1971).

[60] Foi, A., Trimeche, M., Katkovnik, V. & Egiazarian, K. Practical Poissonian-Gaussian noise modeling and fitting for single-image rawdata. *IEEE Transactions on Image Processing* **17**, 1737–1754 (2008).

[61] Combettes, P. L., Dũng, . & Vũ, B. C. Proximity for sums of composite functions. *J. Math. Analysis and Applications* **380**, 680 – 688 (2011). URL `http://www.sciencedirect.com/science/article/pii/S0022247X11002137`.

[62] Combettes, P. L., Condat, L., Pesquet, J. C. & Vũ, B. C. A forward-backward view of some primal-dual optimization methods in image recovery. In *IEEE Int. Conf. Image Processing (ICIP)*, 4141–4145 (2014).

[63] Moreau, J. J. Proximité et dualité dans un espace hilbertien. *Bulletin de la Société Mathématique de France* **93**, 273–299 (1965). URL `http://www.numdam.org/item?id=BSMF_1965__93__273_0`.

Author contributions

C.K. devised the project and the main conceptual ideas, supervised the project and was in charge of overall direction and planning. C.K., H.-N.N. and S.P. designed and implemented the presented SPITFIR(e) method, in discussions with J.S., C.A. V.-C. and L.L. J.S. C.A. V.-C. and L.L. provided the real datasets (LLSM, STED, CM) datasets and conceived experiments and implemented the analysis workflows applied to real images. B.H. provided the real MFM datasets. H.-N.N. and S.P. performed experiments on artificial datasets and real images (WF, MFM, LLSM, CM). C.K., H.-N.N. and S.P. co-wrote the manuscript. All authors provided critical feedback and helped shape the research, analysis and manuscript.

Competing interests

The authors declare no competing interests.

SUPPLEMENTARY INFORMATION

SPITFIR(e): A supermaneuverable algorithm for restoring 2D-3D fluorescence images and videos, and background subtraction

Sylvain Prigent[1,2,★], Hoai-Nam Nguyen [1,2,★], Ludovic Leconte[1,2], Cesar Augusto Valades-Cruz[1,2], Bassam Hajj[3], Jean Salamero[1,2], Charles Kervrann[1,2,*]

[1] SERPICO Project-Team, Inria Centre Rennes-Bretagne Atlantique, 35042, Rennes Cedex, France
[2] SERPICO/STED Team, UMR144 CNRS Institut Curie, PSL Research University, Sorbonne Universités, 75005, Paris, France
[3] Laboratoire Physico-Chimie, Institut Curie, PSL Research University, Sorbonne Universités, CNRS UMR168, 75005, Paris, France

[★] S. Prigent and H.-N. Nguyen contributed equally to this work.
[*] Corresponding author: charles.kervrann@inria.fr
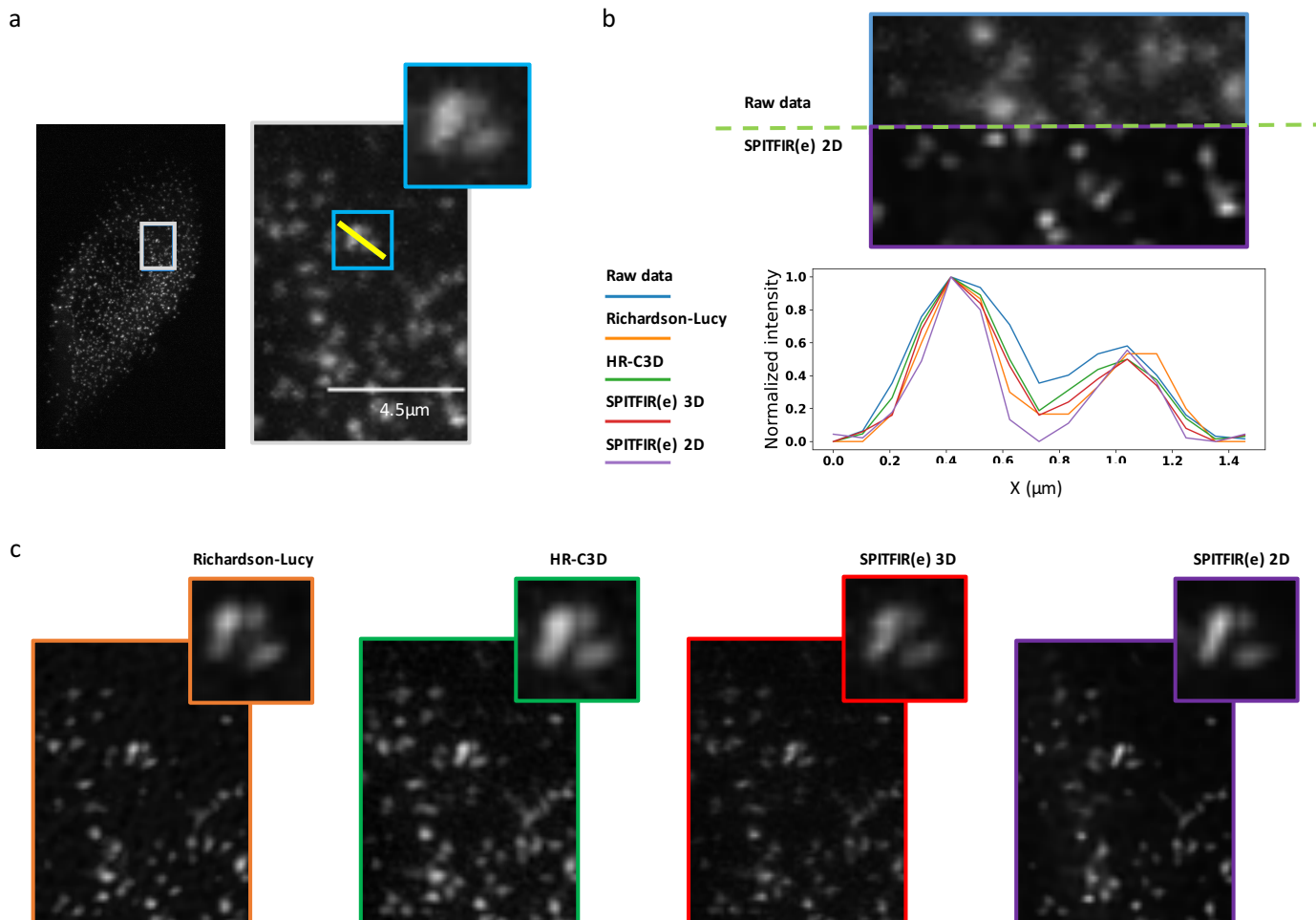
SUPPLEMENTARY VIDEOS

Suppementary Movie 1. Composite movie corresponding to the time series of insets in Fig. 2 **a, b, e** and **f**. Upper part from left to right, microtubules, mitochondria and superimposed MIP data from LLSM after deskew treatment. Lower part, similar disposition after SPITFIR(e) 4D denoising and 3D deconvolution. Only 1:45 minutes of a 4 minutes series is shown. Scale bar: 5 μm.

Supplementary Movie 2. Multi-Angle view of the 3D of superimposed image stacks from the same time series as in Supplementary Movie 1. Upper part data from LLSM after deskew treatment, lower part after SPITFIR(e) 4D denoising and 3D deconvolution. Scale Bar = 5 μm.
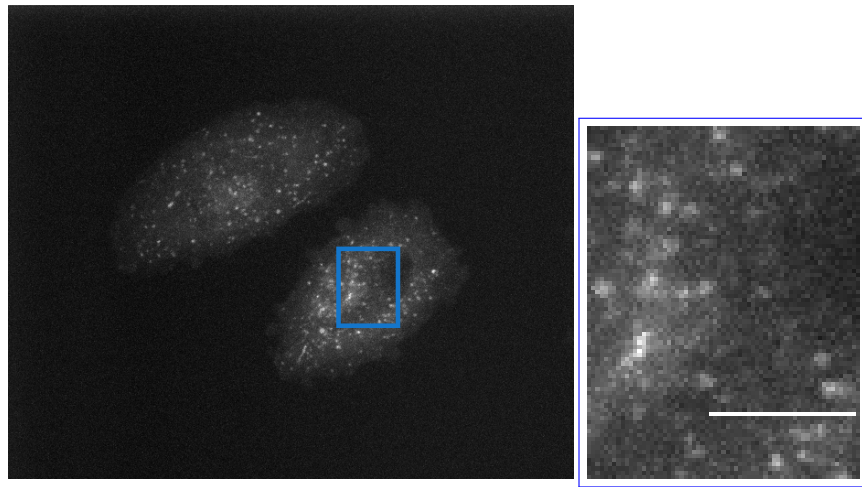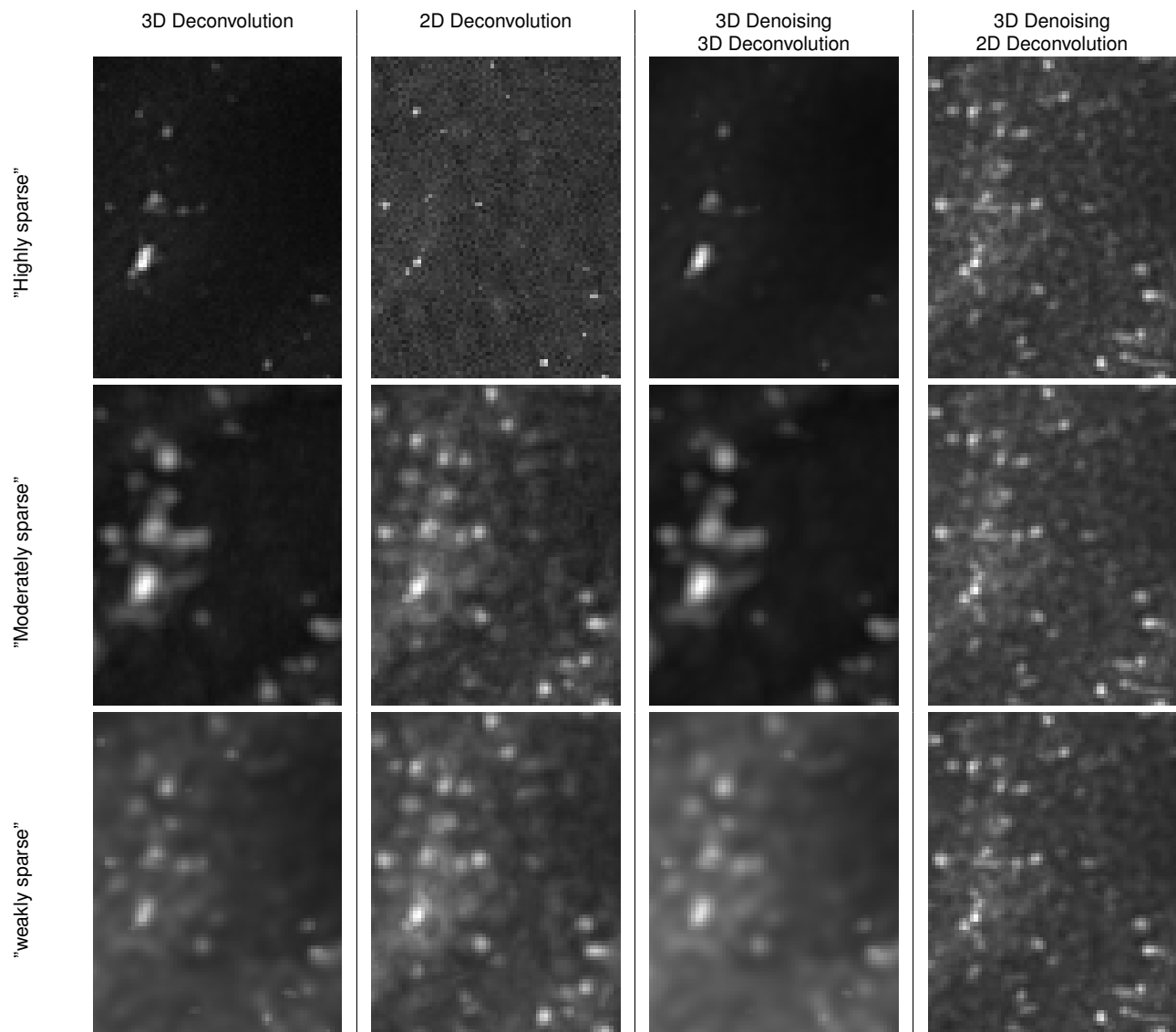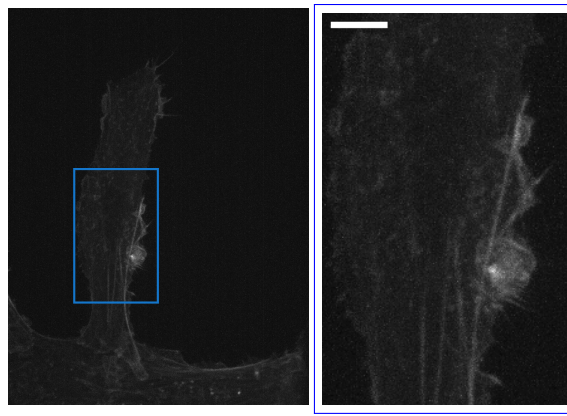
SUPPLEMENTARY FIGURES
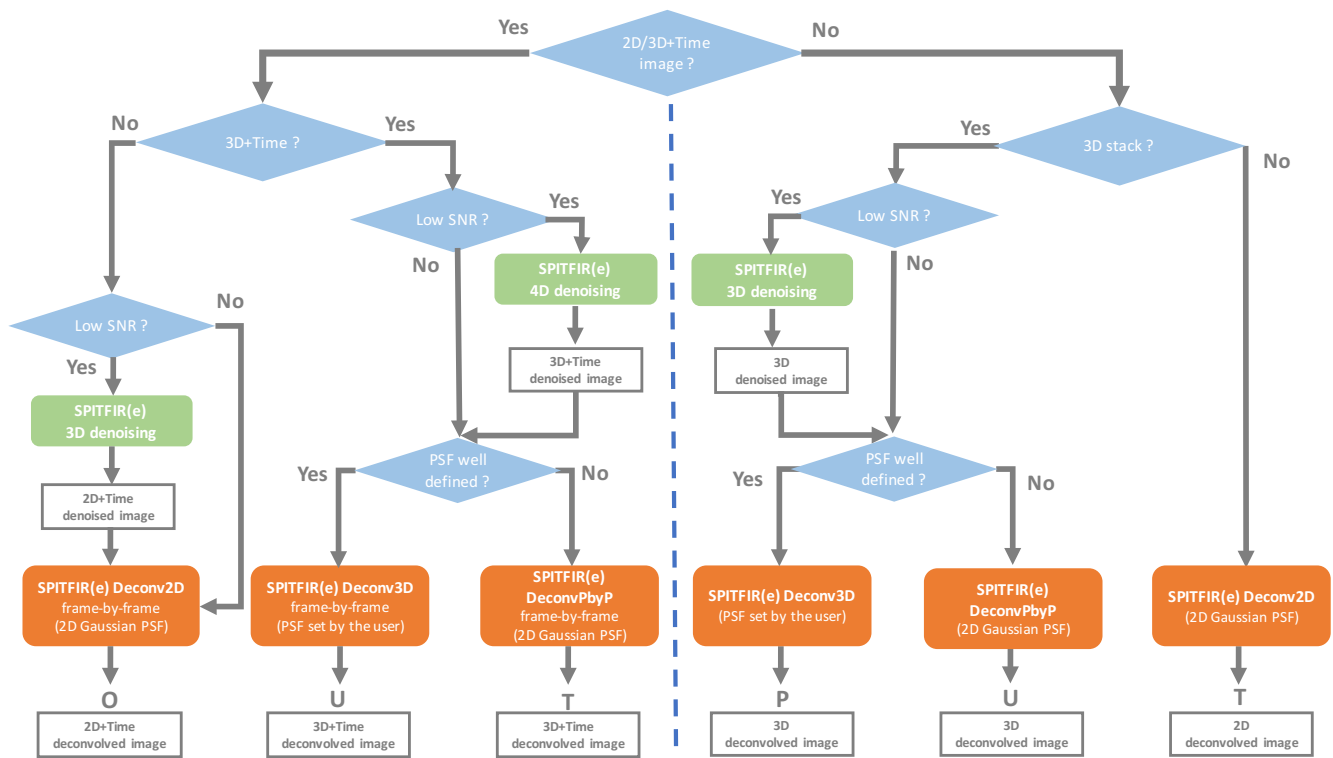
Supplementary Fig. 1: **SPITFIR(e) overview.** The flow chart describes the sequence of restoration steps applied to 3D, 2D+Time, or 3D+Time images based on responses of end-users.

Supplementary Fig. 2: **Deconvolution of a 3D LLSM image single plane image depicting σ subunit-eGFP of the AP2 complex (scale bar: 4.5 μm, source: CNRS UMR144 Institut Curie) with SPITFIR(e) and HR-C3D.** (**a**) Original image (raw data) extracted from a 3D volume of 58 planes. The pixel size is 104 nm along the XY directions and 325 nm along the z axis. (**b**) The intensity profiles for the four methods (Richardson-Lucy, HR-C3D, SPITFIR(e) 3D and SPITFIR(e) 2D) are measured along the yellow segment superimposed within a ROI (red window) in (**a**). (**c**) Images obtained with the four different deconvolution algorithms. SPITFIR(e) 3D used a 3D Gaussian PSF model with parameters $\sigma_{xy} = 1.0$ pixel and $\sigma_z = 0.8$ pixel. SPITFIR(e) 2D used a 2D Gaussian PSF with parameter $\sigma_{xy} = 1.0$ pixel. Scale bar: 4.5 μm.

a

b

Supplementary Fig. 3: Deconvolution of 3D spinning disk confocal microscopy image. **a**, Display of a 2D plane depicting Hela cells expressing GFP-Rab5 proteins (scale bar: 4 μm). **b**, Results with SPITFIR(e) with different amounts of sparsity ("weak" ($\rho = 0.9$), "moderate ($\rho = 0.6$), "high" ($\rho = 0.1$), automatic selection of the regularization parameter, and different strategies with SPITFIR(e): 2D Deconvolution, 3D Decovolution, 3D Denoising + 3D Deconvolution, 3D Denoising + 2D deconvolution ($\sigma_{xy} = 1.5$ pixels and $\sigma_z = 0.5$ pixel).

a

b

| 3D Deconvolution | 2D Deconvolution | 3D Denoising 3D Deconvolution | 3D Denoising 2D Deconvolution |

"Highly sparse"

"Moderately sparse"

"weakly sparse"
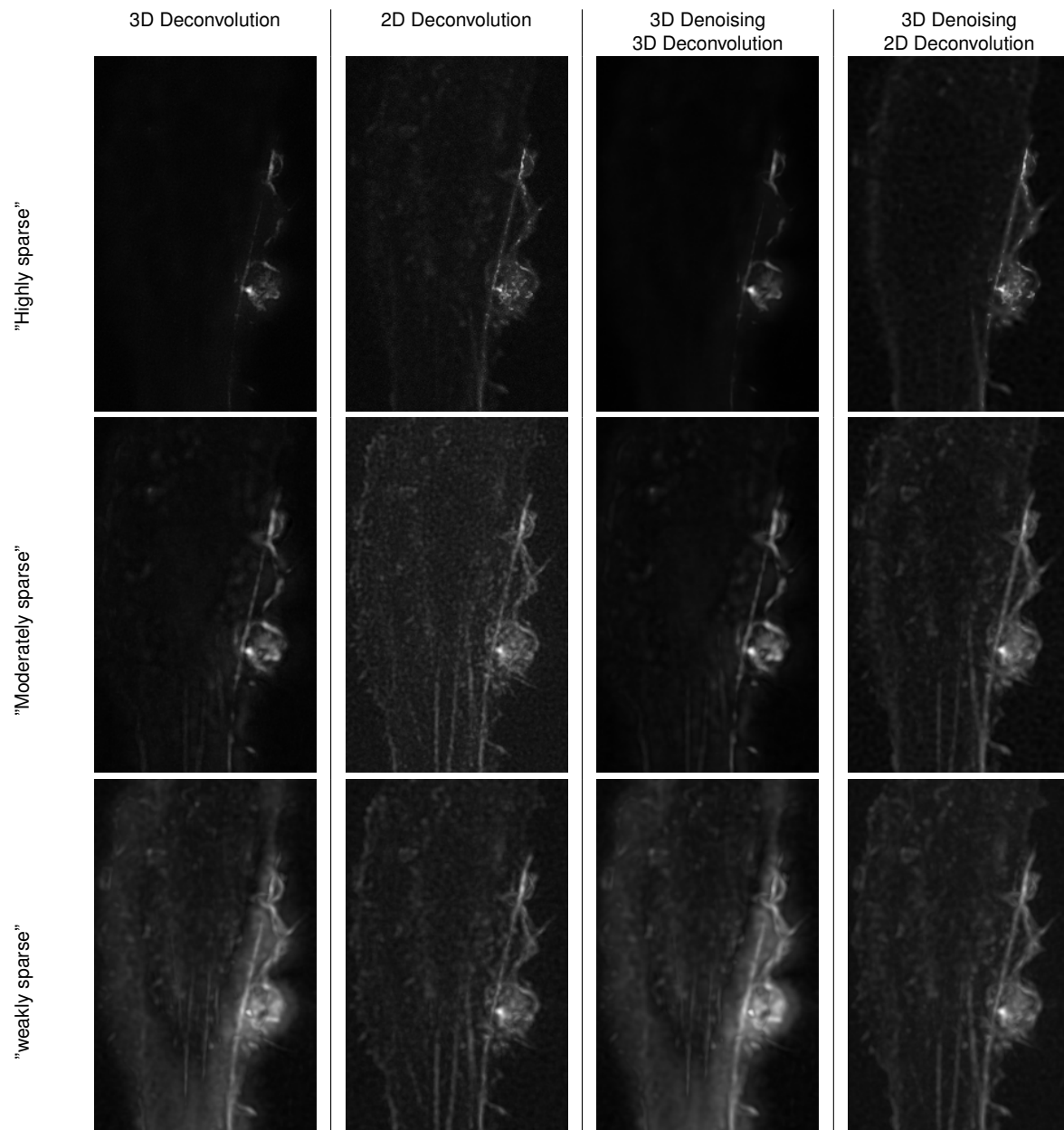
Supplementary Fig. 4: Deconvolution of 3D spinning disk confocal microscopy image. **a**, Image depicting mCherry-LifeAct in RPE1 cells (scale bar: 5 μm). **b**, Results with SPITFIR(e) with different levels of sparsity ("weak" ($\rho = 0.9$), "moderate ($\rho = 0.6$), "high" ($\rho = 0.1$), automatic selection of the regularization parameter and different strategies with SPITFIR(e): 2D Deconvolution, 3D Decovolution, 3D Denoising + 3D Deconvolution, 3D Denoising + 2D deconvolution ($\sigma_{xy} = 1.5$ pixels and $\sigma_z = 0.5$ pixel).

18
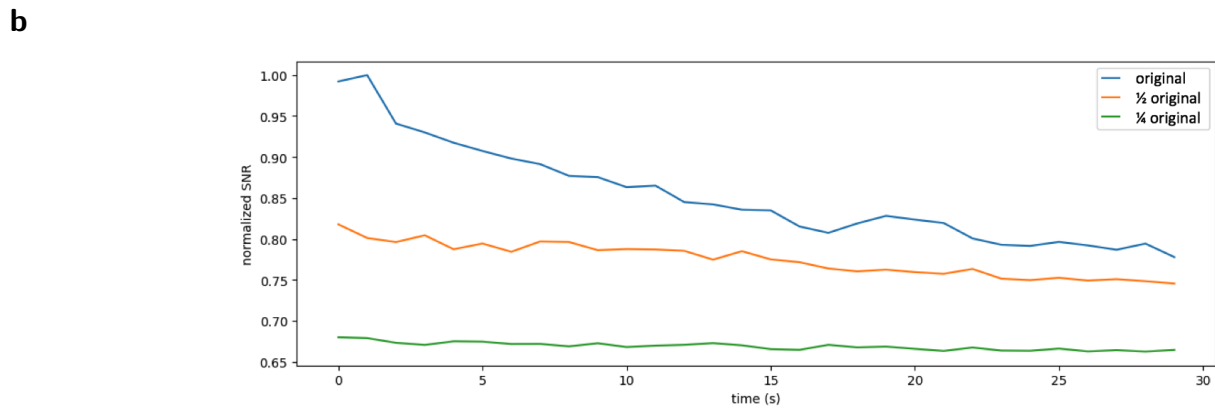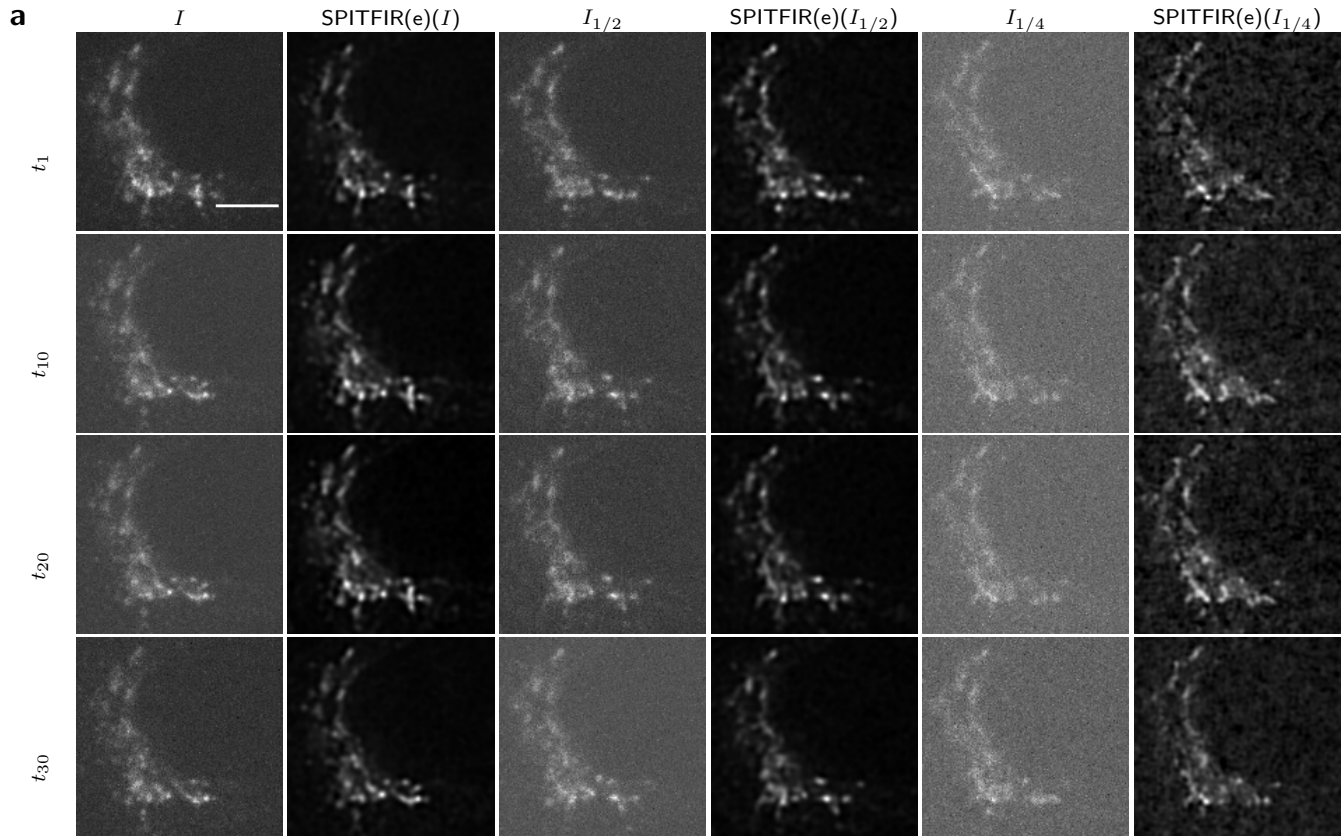
Supplementary Fig. 5: **SPITFIR(e) (4D denoising + 3D deconvolution) applied to spinning disk confocal microscopy images. a**, Data are temporal series (30 time points) of 3D stacks composed 14 planes each depicting CD-M6PR-eGFP Hela cells. The sample is imaged 3 times with different doses of illumination light. The figure depicts the 6th plane extracted from the 3D stack of a region of interest (ROI) at time t_1, t_{10}, t_{20}, and t_{30}. Scale bar = 4.5 μm. The I image is illuminated with high dose of illumination light to get an image with a high SNR. The $I_{1/2}$ image is illuminated with half power compared to I. The $I_{1/4}$ image is obtained by dividing the dose of illumination light by a factor 4 compared to I. We applied SPITFIR(e) (4D denoising + 3D deconvolution with a Gibson-Lanni PSF model truncated to 3 planes around the center of the PSF and the acquisition parameters) to the three 4D volumes. The results are denoted SPITFIR(e)(I), SPITFIR(e)($I_{1/2}$) and SPITFIR(e)($I_{1/4}$), respectively. **b**, The plot shows the SNR of the three original movies over time.
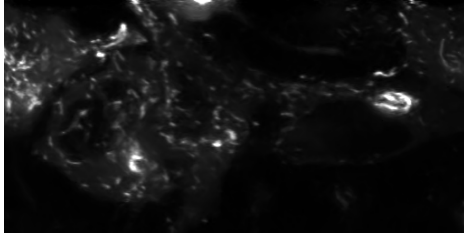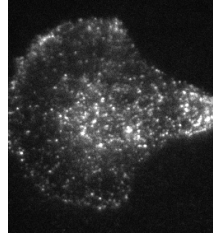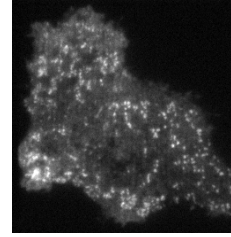
Image #1 (512×256) (CM) Image #2 (227×238) (TIRFM) Image #3 (256×283) (TIRFM) Image #4 (1054×1028) (SIM)
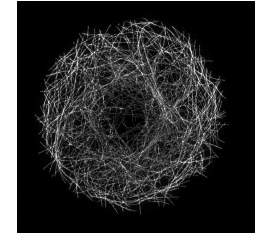
		RL	GM	ICTM	TV	TV-L_1	GN	TVH	LSHV	GR	SV	HV	SHV (SPITFIR(e))
Image #1 (CM)	$\sigma_{xy} = 1.0$												
	$\tau = 0.01$	30.01	23.73	39.62	40.21	40.23	39.71	36.66	35.48	36.01	**42.13**	41.78	**42.15**
	$\tau = 0.02$	25.12	19.89	37.04	38.04	38.06	37.10	32.98	32.75	32.93	39.45	39.34	**39.55**
	$\tau = 0.04$	20.52	15.89	34.40	35.76	35.76	34.43	28.56	28.52	28.47	**36.68**	36.43	36.61
	$\sigma_{xy} = 1.25$												
	$\tau = 0.01$	31.02	28.92	38.52	38.88	38.86	38.61	35.88	33.96	34.55	**40.84**	40.42	**40.87**
	$\tau = 0.02$	25.95	24.11	36.35	36.81	36.80	36.41	32.56	31.88	32.20	**38.65**	38.28	38.57
	$\tau = 0.04$	21.02	19.36	33.90	34.87	34.90	33.99	28.42	28.17	28.27	**36.09**	35.80	35.96
	$\sigma_{xy} = 1.5$												
	$\tau = 0.01$	31.85	32.89	37.50	37.95	37.89	37.61	35.20	32.70	33.28	**39.87**	39.12	39.61
	$\tau = 0.02$	26.86	27.92	35.51	36.04	36.03	35.57	32.15	31.05	31.44	**37.78**	37.19	37.69
	$\tau = 0.04$	21.67	22.65	33.25	33.86	33.86	33.34	28.23	27.73	27.84	**35.34**	34.83	35.08
Image #2 (TIRFM)	$\sigma_{xy} = 1.0$												
	$\tau = 0.01$	27.61	10.66	31.08	30.92	**31.29**	31.00	30.73	29.79	30.73	30.92	31.01	31.01
	$\tau = 0.02$	22.81	29.64	29.79	29.77	**30.01**	29.69	28.72	28.56	28.65	29.83	29.89	**30.01**
	$\tau = 0.04$	17.93	27.50	28.27	28.40	28.43	28.21	25.75	25.75	25.76	28.60	28.63	**28.78**
	$\sigma_{xy} = 1.25$												
	$\tau = 0.01$	28.26	17.07	29.96	29.78	**30.13**	30.00	29.73	28.42	29.75	30.01	29.87	30.01
	$\tau = 0.02$	23.94	15.37	28.99	28.94	**29.23**	28.95	27.94	27.53	27.92	29.17	29.02	29.16
	$\tau = 0.04$	18.98	12.13	27.64	27.68	27.77	27.58	25.22	25.21	25.22	27.97	27.85	**28.09**
	$\sigma_{xy} = 1.5$												
	$\tau = 0.01$	28.46	23.29	28.97	28.80	28.91	27.38	28.93	**29.09**	29.07	**29.09**	28.90	**29.08**
	$\tau = 0.02$	24.76	20.61	28.17	28.10	28.38	28.19	27.35	26.72	27.29	**28.44**	28.04	28.29
	$\tau = 0.04$	19.94	16.54	26.98	26.97	27.10	26.91	24.81	24.74	24.78	27.34	27.08	**27.35**
Image #3 (TIRFM)	$\sigma_{xy} = 1.0$												
	$\tau = 0.01$	28.19	13.56	33.30	33.05	**33.40**	33.22	32.64	32.16	32.68	33.07	33.27	33.27
	$\tau = 0.02$	23.21	11.75	**31.96**	31.50	31.74	31.89	30.27	30.26	30.26	31.49	31.83	31.82
	$\tau = 0.04$	18.26	28.65	30.40	30.20	30.23	30.34	26.86	26.85	26.87	30.20	**30.50**	30.51
	$\sigma_{xy} = 1.25$												
	$\tau = 0.01$	29.11	19.47	32.42	32.14	**32.49**	32.28	31.86	30.99	31.83	32.22	32.34	32.34
	$\tau = 0.02$	24.31	16.88	**31.24**	30.82	31.19	31.15	29.67	29.51	29.63	30.81	31.11	31.14
	$\tau = 0.04$	19.28	12.99	**29.76**	29.42	29.51	29.68	26.47	26.46	26.46	29.60	29.68	**29.76**
	$\sigma_{xy} = 1.5$												
	$\tau = 0.01$	29.69	25.03	**31.51**	31.36	31.62	31.39	31.13	29.98	31.12	31.36	**31.48**	31.48
	$\tau = 0.02$	25.31	21.46	**30.60**	30.18	30.55	30.49	29.15	28.80	29.14	30.33	30.42	30.40
	$\tau = 0.04$	20.23	16.86	**29.26**	28.86	28.98	29.19	26.17	26.15	26.08	**29.21**	29.05	29.24
Image #4 (SIM)	$\sigma_{xy} = 1.0$												
	$\tau = 0.01$	31.24	21.49	36.57	35.87	35.93	37.06	36.19	33.74	36.14	37.89	37.62	**38.21**
	$\tau = 0.02$	25.89	19.17	33.95	33.76	33.77	34.42	32.60	31.89	32.60	35.19	34.99	**35.54**
	$\tau = 0.04$	20.80	15.66	31.33	31.41	31.44	31.79	28.41	28.31	27.66	32.48	32.31	**32.70**
	$\sigma_{xy} = 1.25$												
	$\tau = 0.01$	32.13	28.63	34.79	34.55	34.56	35.35	34.58	31.46	34.53	36.08	35.66	**36.29**
	$\tau = 0.02$	26.82	24.75	32.55	32.36	32.40	33.09	31.67	30.33	31.64	33.79	33.49	**34.07**
	$\tau = 0.04$	21.45	19.94	30.29	30.29	30.35	30.67	28.03	27.64	27.35	31.41	31.13	**31.54**
	$\sigma_{xy} = 1.5$												
	$\tau = 0.01$	32.34	33.04	33.14	33.36	33.40	33.70	33.08	29.82	33.09	34.47	34.00	**34.66**
	$\tau = 0.02$	27.55	28.78	31.27	31.19	31.25	31.80	30.67	29.03	30.66	32.48	32.08	**32.64**
	$\tau = 0.04$	22.15	23.57	29.36	29.37	29.46	29.88	27.60	26.94	27.04	30.44	30.08	**30.54**

Supplementary Table 2: **PSNR scores of deconvolution methods applied to four 2D images (confocal (CM), TIRFM, SIM).** The best scores (± 0.05 dB) are in bold style. Images #1-#3 depict fluorescently tagged proteins (Rab11-mCherry, Langerin-YFP, TfnR-pHluorin (Transferin-receptor)) corresponding to small bright spots over a dark background, observed in spinning confocal microscopy (CM) (Image #1, courtesy of PICT-IBiSA imaging platform) and total internal reflection fluorescence microscopy (TIRFM) (Images #2 and #3) respectively, previously acquired [55;56]. Image #4 depicts microtubules observed in Structured Illumination Microscopy (SIM) with high resolution (up to 100 nm) This image (CIL 36147) is taken from the Cell Image Library (CIL) (http://www.cellimagelibrary.org). The images were deconcolved with the following methods: Richardson-Lucy (RL)[3] algorithm, Gold-Meinel (GM)[57] algorithm, iterative constrained Tikhonov-Miller (ICTM)[7] algorithm, Total Variation (TV)[41] regularizer, Total-Variation-L1 (TV-L1)[19], GraphNet (GN)[15] regularizer, Total Variation-Huber (TVH)[58] regularizer, Log Sparse Hessian Variation (LSHV)[14] regularizer, Good's roughness (GR)[59] regularizer, Sparse Variation (SV)[18] regularizer, Hessian Variation (HV)[12] regularizer, and Sparse Hessian Variation (SHV) (our SPITFIR(e) method) regularizer.

20
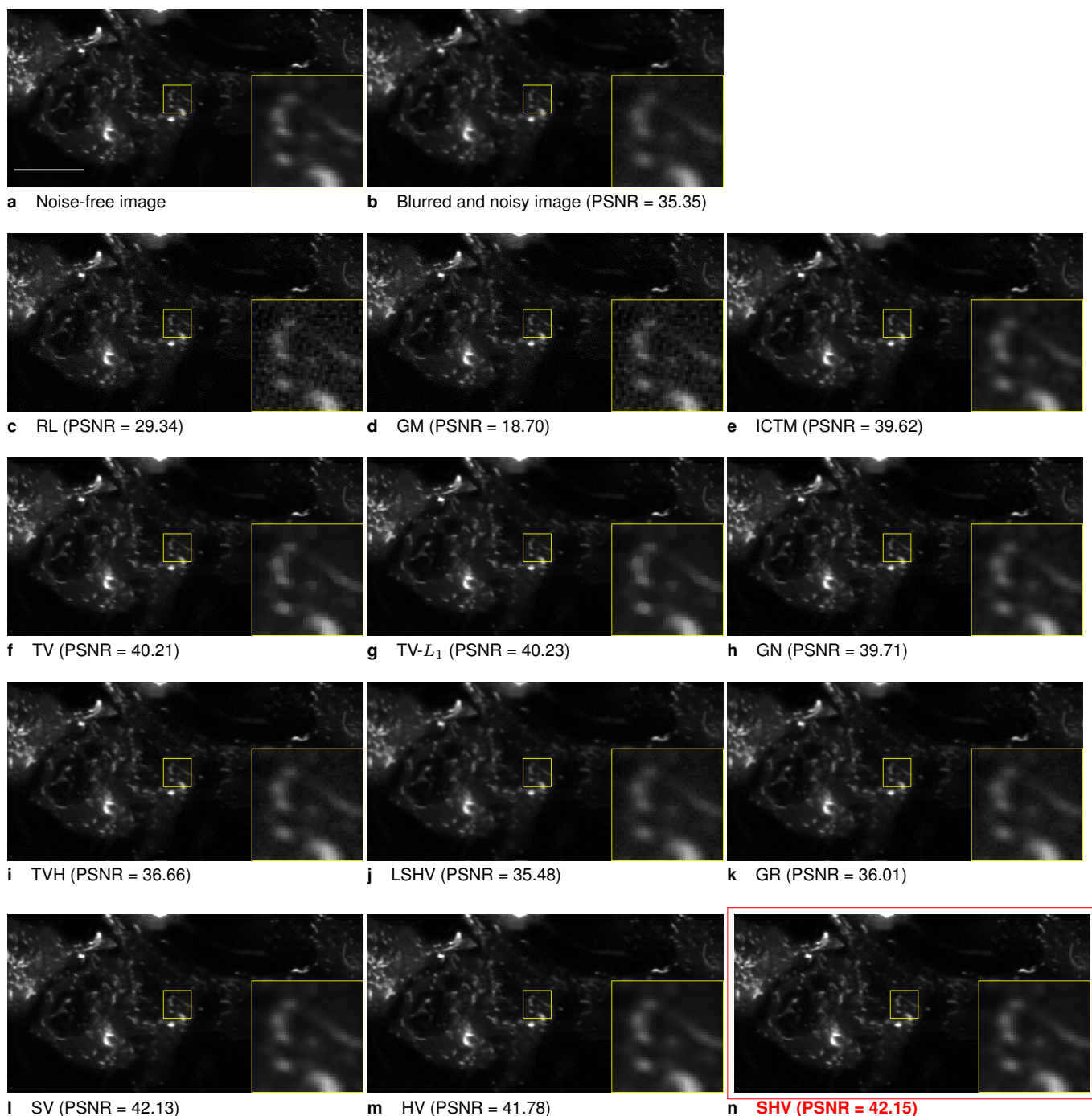
Supplementary Fig. 6: Deconvolution of a confocal microscopy image with different methods. a, The original 512×256 pixels image (Image #1) (source: CNRS UMR144 Institut Curie) depicting Rab11a-mCherry protein in M10 cells, was previously acquired in confocal spinning-disk microscopy. Scale bar: 10 μm. **b** Artificially corrupted image by a Gaussian PSF $\sigma_{xy} = 1.0$ pixel and Gaussian white noise with standard deviation $\tau = 0.01$ (**b**). **c**, image deconvolved with the Richardson-Lucy (RL)[3] algorithm. **d**, image deconvolved with the Gold-Meinel (GM)[57] algorithm. **e**, image deconvolved with the iterative constrained Tikhonov-Miller (ICTM)[7] algorithm. **f**, image deconvolved with the Total Variation (TV)[41] regularizer. **g**, image deconvolved with the Total-Variation-L1 (TV-L1)[19]. **h**, image deconvolved with the GraphNet (GN)[15] regularizer. **i**, image deconvolved with the Total Variation-Huber (TVH)[58] regularizer. **j**, image deconvolved with the Log Sparse Hessian Variation (LSHV)[14] regularizer. **k**, image deconvolved with the Good's roughness (GR)[59] regularizer. **l**, image deconvolved with the Sparse Variation (SV)[18] regulariser. **m**, image deconvolved with the Hessian Variation (HV)[12] regularizer. **n**, image deconvolved with the Sparse Hessian Variation (SHV) (our SPITFIR(e) method) regularizer (red box). Zoom-in views are displayed in yellow windows for comparison in details.

21
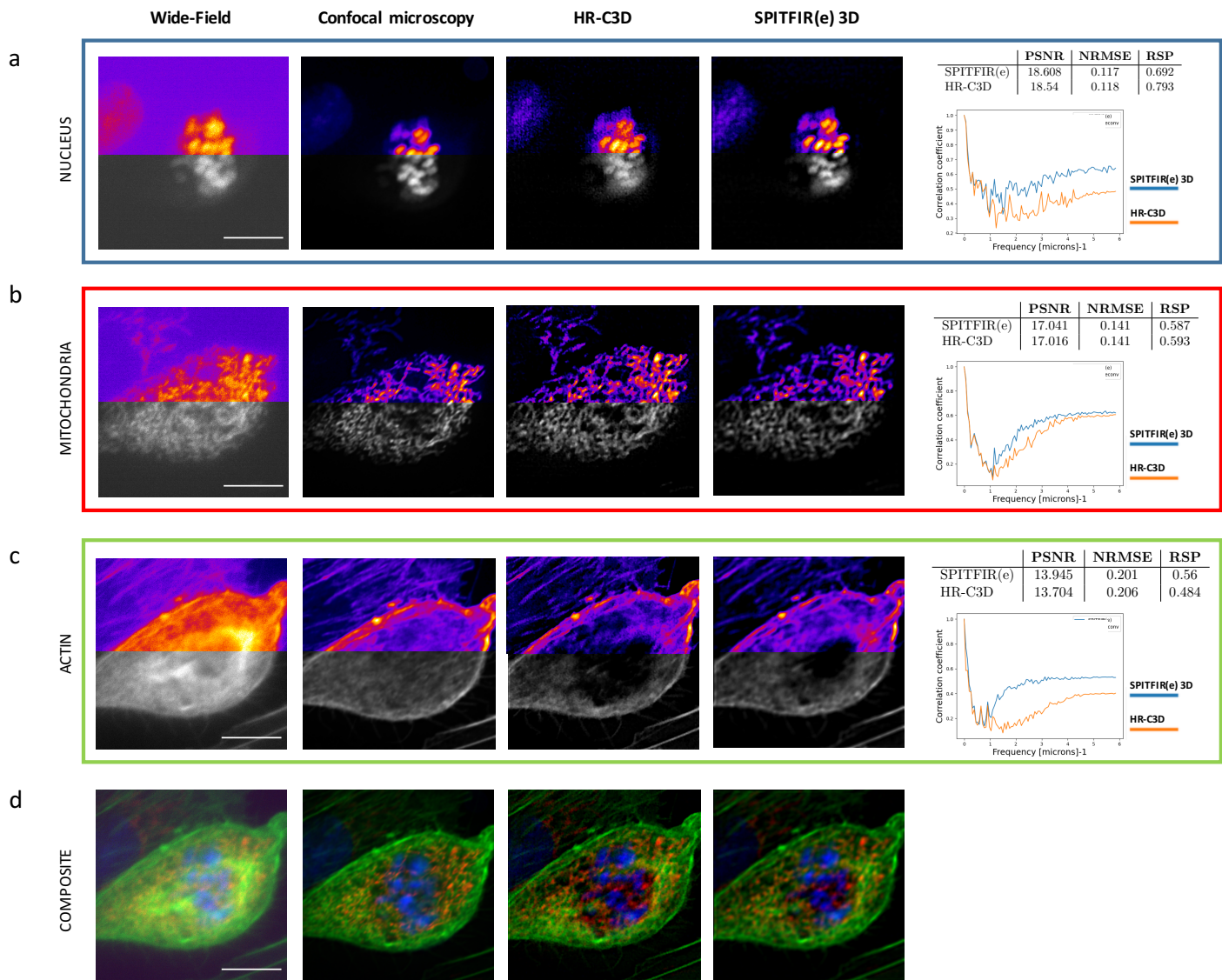
Supplementary Fig. 7: A comparison of SPITFIR(e) 3D and HR-C3D[13] applied to a 3D wide-field microscopy image. Display of 2D (512×512 pixels) images extracted from a stack composed of 22 planes depicting nucleus stained with DAPI (**a**, blue channel, 11th plane), mitochondria staind with MitoTracker Red CMXRos (**b**, red channel, 6th plane), and actin staind with lexa Fluor 488 phalloidin (**c**, green channel, 4th plane) (see description in [13]). **d**, Display of the wide-field composite image (11th plane, merging of the three channels), the corresponding confocal image deconvolved with Huygens software (see [13]) which will serve as high-resolution reference image, and the deconvolved images with HR-C3D and SPITFIR(e) 3D, from left to right, respectively. The three plots show the Fourier Shell Correlation (FSC) between the wide-field image deconvolved with the two competing methods and the reference (spinning disk) confocal image. Each plot corresponds to one color channel. The Peak-Signal-to-Noise-Ratio (PSNR), the Normalized Root Mean Square Error (NRMSE) and the global Pearson. correlation coefficient (RSP) are reported in Tables. Scale bar: 10 μm.
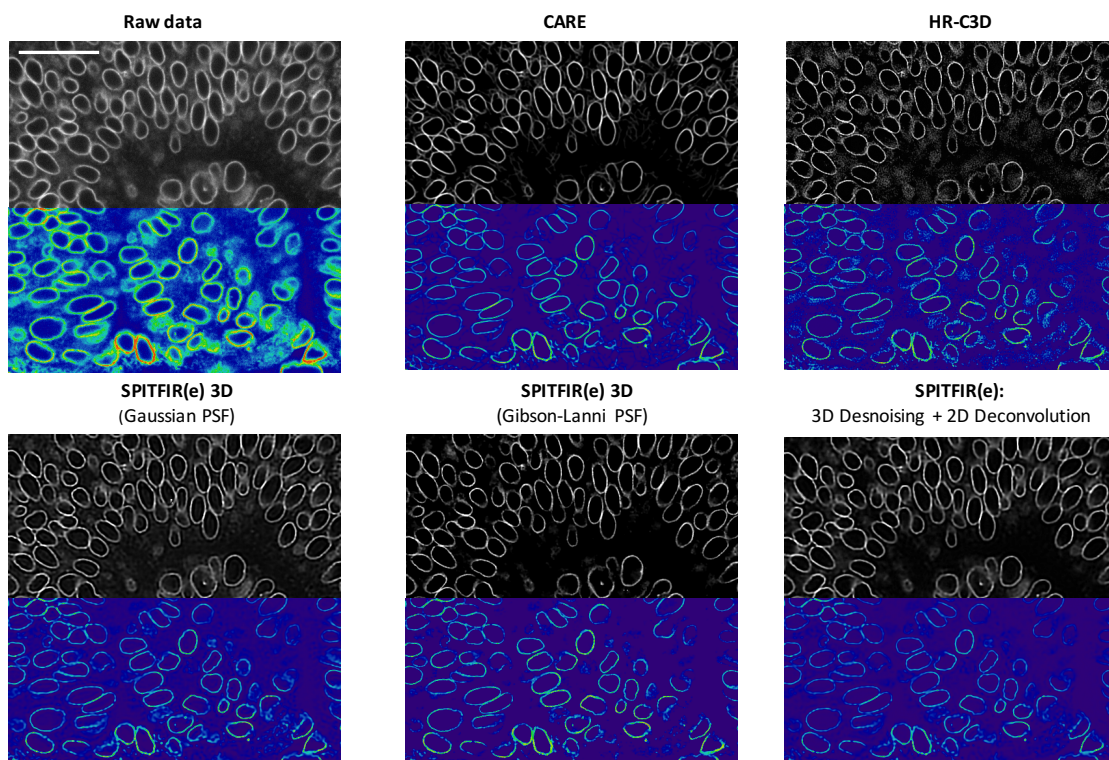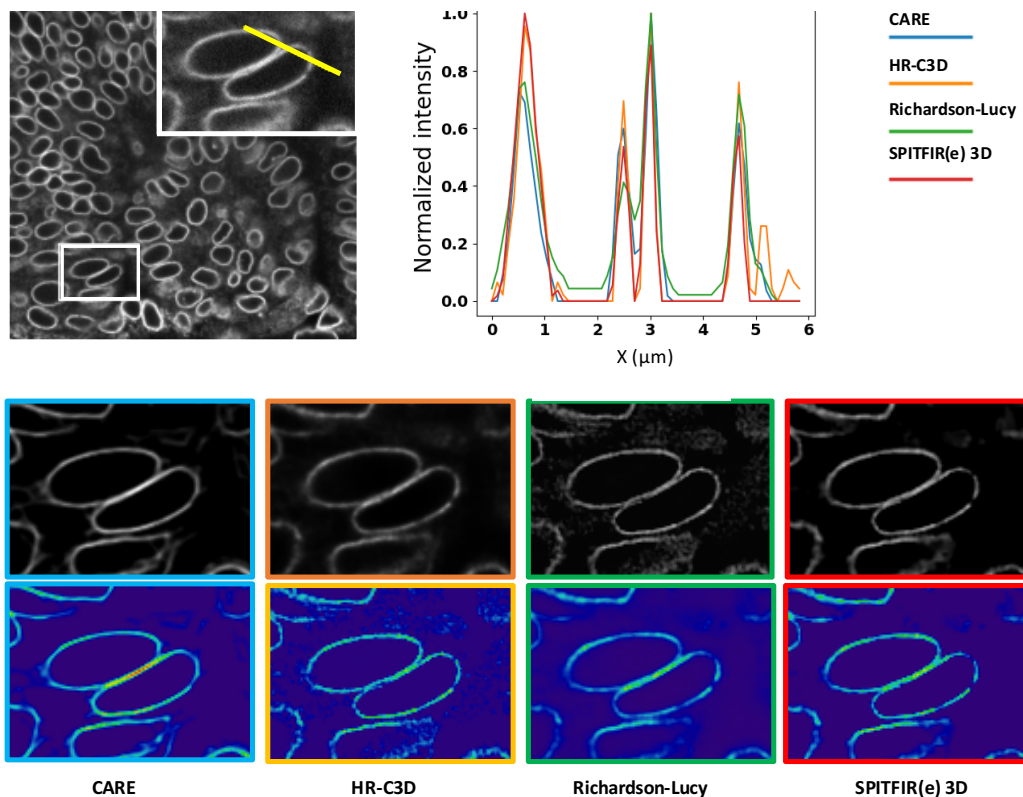
Supplementary Fig. 8**:   Deconvolution of a 3D spinning disk confocal microscopy image depicting the envelopes of nuclei stained with GFP-LAP2b (developing eye of zebrafish (Daniorerio) embryos, source:**[34]**) (scale bar: 25** μ**m). a,** Comparison of RL, HR-C3D, CARE and SPITFIR(e) with a Gibson-Lanni PSF modeld and a 3D Gaussian PSF model. We display the results for the 5th plane. **b,** The plots show the intensity profile for each 3D deconvolution method along the line drawn in yellow in ROI (white rectangle). Bottom, Zoom-in views (5th plane) of deconvolution results obtained with CARE, RL, HR-C3D and SPITFIR(e) 3D (Gibson-Lanni (generated from [34]: 60X/1.3-NA objective, 488 nm, z step = 2 μm) or 3D Gaussian PSF model ($\sigma_{xy} = 2.0$ pixels and $\sigma_z = 0.5$ pixel).
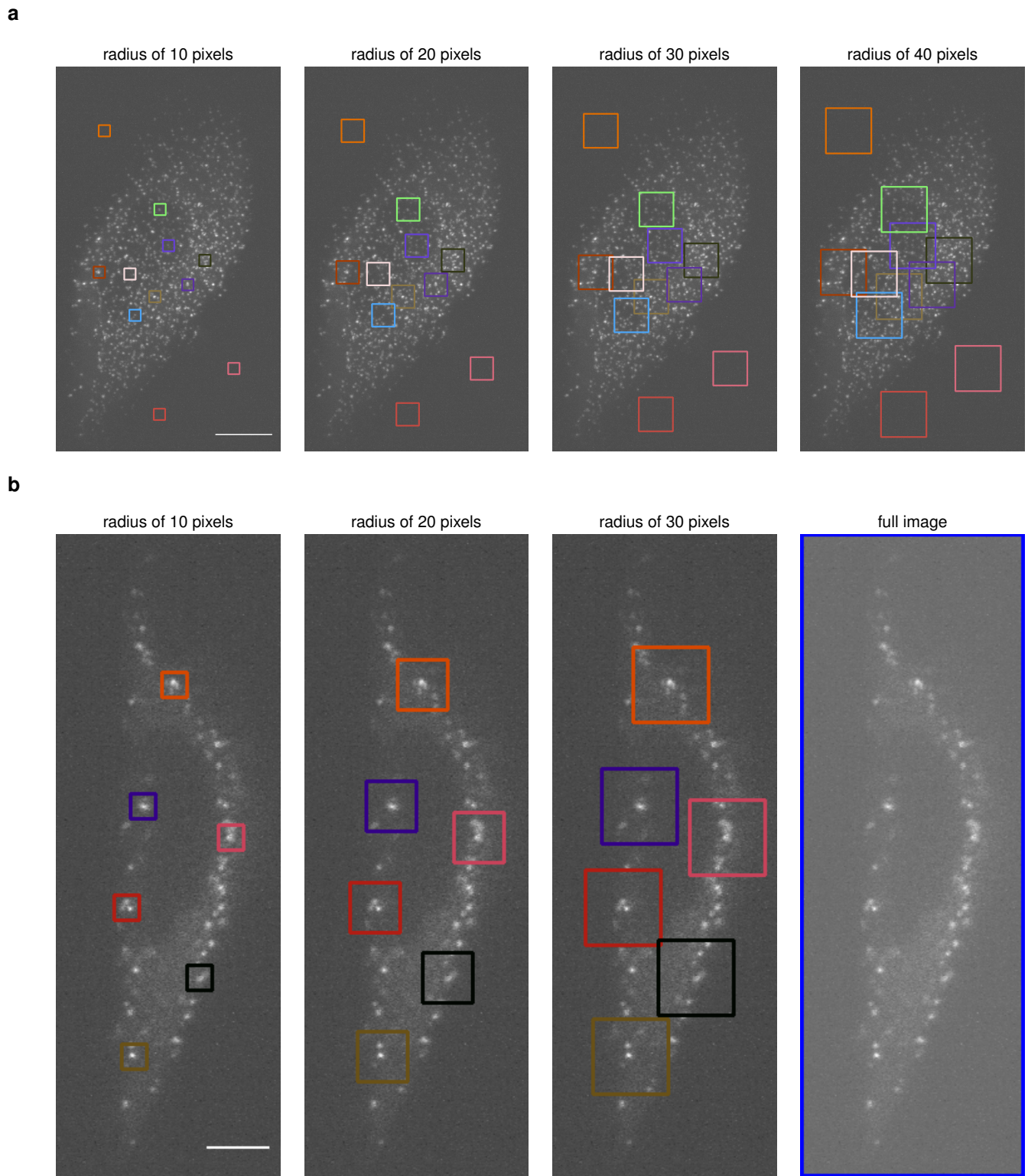
Supplementary Fig. 9: Automatic selection of λ parameter using the minimax principle applied to ROIs with variable radius. a, Display of ROIs with radius of 10, 20, 30 and 40 pixels superimposed on the maximum projection image of 3D LLSM image depicting σ subunit-eGFP of the AP2 complex in Hela cells. Scale bar: 10 μm. **b**, Display of ROIs with radius of 10, 20, and 30 pixels superimposed on the maximum projection image of 3D LLSM image depicting σ subunit-eGFP of the AP2 complex in Hela cells (Orientation XY based on Fig. 1). Scale bar: 5 μm.

SUPPLEMENTARY NOTES

A. Definition of the energy for images corrupted by Poisson-Gaussian noise

A fidelity data term is generally derived from the general formation model, for instance dedicated to low photon counts and low-light regimes in fluorescence microscopy (e.g., [13]). We formally demonstrate below that a conventional quadratic fidelity term, which is optimal when the images are corrupted by additive white Gaussian noise, is also appropriate for mixed Poisson-Gaussian noise.

Calculation of noise variance

Let us consider the following observation model[60]:

$$f(\mathbf{x}) = v(\mathbf{x}) + \aleph(v(\mathbf{x})) + \varepsilon(\mathbf{x}), \tag{14}$$

with $v(\mathbf{x}) = (h * u)(\mathbf{x})$, where $\mathbf{x} \in \Omega$ is the pixel position in the domain Ω, $\aleph$ and ε represent the signal-dependent Poisson noise component and the zero-mean Gaussian noise component respectively, such that:

$$\alpha(v(\mathbf{x}) + \aleph(v(\mathbf{x}))) \sim \mathcal{P}(\alpha v(\mathbf{x})), \quad \varepsilon(\mathbf{x}) \sim \mathcal{N}(0, \tau^2), \tag{15}$$

$\alpha > 0$ is the quantization factor of the photodetector, and $\tau^2 > 0$ represents the Gaussian noise variance. According to the properties of Poisson distribution ($\mathbb{E}[\cdot]$ and $\mathrm{Var}[\cdot]$ denote the mathematical expectation and the variance respectively), we have

$$\begin{cases} \mathbb{E}[\alpha(v(\mathbf{x}) + \aleph(v(\mathbf{x})))]] &= \alpha v(\mathbf{x}) = \mathrm{Var}[\alpha(v(\mathbf{x}) + \aleph(v(\mathbf{x})))] = \alpha^2 \mathrm{Var}[\aleph(v(\mathbf{x}))] \\ \mathbb{E}[\alpha(v(\mathbf{x}) + \aleph(v(\mathbf{x})))]] &= \alpha v(\mathbf{x}) + \alpha \mathbb{E}[\aleph(v(\mathbf{x}))], \end{cases} \tag{16}$$

which yields to $\mathbb{E}[\aleph(v(\mathbf{x}))] = 0$ and $\mathrm{Var}[\aleph(v(\mathbf{x}))] = v(\mathbf{x})/\alpha$. Hence,

$$\mathbb{E}[f(\mathbf{x})] = v(\mathbf{x}) + \underbrace{\mathbb{E}[\aleph(v(\mathbf{x}))]}_{= 0} + \underbrace{\mathbb{E}[\varepsilon(\mathbf{x})]}_{= 0}, \tag{17}$$

and the overall variance of $f(\mathbf{x})$ is the sum of $\mathrm{Var}[\aleph(v(\mathbf{x}))]$ and $\mathrm{Var}[\varepsilon(\mathbf{x})] = \tau^2$:

$$\mathrm{Var}[f(\mathbf{x})] = \mathbb{E}[(f(\mathbf{x}) - \mathbb{E}[f(\mathbf{x})])^2] = \mathbb{E}[(f(\mathbf{x}) - v(\mathbf{x}))^2] = v(\mathbf{x})/\alpha + \tau^2. \tag{18}$$

We conclude that $v(\mathbf{x})$ is defined as the expected value of the noisy observations $f(\mathbf{x})$ and $\mathrm{Var}[f(\mathbf{x})] = \alpha^{-1} v(\mathbf{x}) + \tau^2$ is the noise variance [39;40;60].

Definition of energy model

Assume an image $v = h * u$ corrupted by a mixed Poisson-Gaussian and denote f the noisy image. As explained above, the local noise variance can be represented at a given location $\mathbf{x} \in \Omega$ as:

$$V_{noise}(\mathbf{x}) \triangleq \mathbb{E}[(f(\mathbf{x}) - (h * u)(\mathbf{x}))^2] = \alpha^{-1} v(\mathbf{x}) + \tau^2. \tag{19}$$

Let us assume that the average intensity is preserved, that is

$$\int_\Omega v(\mathbf{x}) \, d\mathbf{x} = \int_\Omega f(\mathbf{x}) \, d\mathbf{x}. \tag{20}$$

It follows that, by pre-multiplying the intensities on both sides of the previous equation by a factor α^{-1} and adding the constant value τ^2, we have:

$$\int_\Omega \alpha^{-1} v(\mathbf{x}) + \tau^2 \, d\mathbf{x} = \int_\Omega \alpha^{-1} f(\mathbf{x}) + \tau^2 \, d\mathbf{x}. \tag{21}$$

Meanwhile, by integrating (19) over the image domain Ω, we get

$$\int_\Omega \mathbb{E}[(f(\mathbf{x}) - (h * u)(\mathbf{x}))^2] d\mathbf{x} = \int_\Omega \alpha^{-1} v(\mathbf{x}) + \tau^2 d\mathbf{x}, \tag{22}$$

$$\int \int_\Omega (f(\mathbf{x}) - (h * u')(\mathbf{x}))^2 p(u'(\mathbf{x})) \, d\mathbf{x} \, du'(\mathbf{x}) = \int_\Omega \alpha^{-1} v(\mathbf{x}) + \tau^2 d\mathbf{x}, \tag{23}$$

where $p(u'(\mathbf{x}))$ represents the probability distribution of u. Assume the following prior $p(u'(\mathbf{x})) = \mathbf{1}[f(\mathbf{x}) = ((h * u)(\mathbf{x}))]$ where $\mathbf{1}[\cdot]$ is the indicator function. Hence, we have by using (21)

$$\int_\Omega (f(\mathbf{x}) - (h * u)(\mathbf{x}))^2 d\mathbf{x} = \int_\Omega \alpha^{-1} v(\mathbf{x}) + \tau^2 d\mathbf{x}, \tag{24}$$

$$= \int_\Omega \alpha^{-1} f(\mathbf{x}) + \tau^2 d\mathbf{x}, \tag{25}$$

$$\|f - (h * u)\|_2^2 = \alpha^{-1}\|f\|_1 + |\Omega|\tau^2. \tag{26}$$

Now, starting from the seminal paper[41], the restored image is found by solving the following optimization problem (Poisson-Gaussian noise):

$$\min_u R(u) \quad \text{subject to} \quad \|f - (h * u)\|_2^2 = \alpha^{-1}\|f\|_1 + |\Omega|\tau^2. \tag{27}$$

It turns out that (27) is a constrained formulation with an equality constraint which is not convex. The corresponding formulation with inequality constraint is

$$\min_u R(u) \quad \text{subject to} \quad \|f - (h * u)\|_2^2 \leq \alpha^{-1}\|f\|_1 + |\Omega|\tau^2. \tag{28}$$

It has been established that, under additional assumptions, that (**??**) and (28) are equivalent. A Lagrange formulation can be then derived since the the right-hand side of the inequality does not depend on the unknown image u:

$$\min_u R(u) + \lambda'\|f - (h * u)\|_2^2, \tag{29}$$

where the parameter $\lambda' > 0$ balances the two energy terms and is an unknown function of the Poisson-Gaussian noise variance. The Karush-Tucker conditions guarantees that (30) and (28) are equivalent for a particular choice of λ'. Finally, if we set $\lambda' = \lambda^{-1} \neq 0$, we can equivalently reformulate the minimization problem as:

$$\min_u \|f - (h * u)\|_2^2 + \lambda R(u). \tag{30}$$

B. Discrete formulation and optimization

Discrete formulation

The observed noisy and blurry image f is represented by its digitized (discrete) version and not by its continuously defined counterpart. The continuous model is not appropriate for discrete images even though the estimation of the continuous image u from discrete samples of f is in principle possible. Consider a discrete formulation by assuming that the images u and f are non negative and sampled according to the sampling grid

$$\Lambda = \mathbb{Z}^2 \cap \Omega = \{1, 2, \ldots, N_x\} \times \{1, 2, \ldots, N_y\} \times \{1, 2, \ldots, N_z\}. \tag{31}$$

The observed noisy and blurry image f is represented by its digitized (discrete) version as follows:

$$f = \mathcal{T}(h * u)$$

where $f, u \in \mathbb{R}_+^N$ with $N = N_x \times N_y \times N_z$, $H \in \mathbb{R}^{N \times N}$ is a matrix that models the point spread function of the microscope in the discrete setting, and $\mathcal{T}$ is the degradation operator. For a coordinate $(i, j, k) \in \Lambda$, we denote by $u_{i,j,k}$ (resp. $f_{i,j,k}$) the value of u (resp. f) at position $(i, j, k) \in \Lambda$. A discrete version of these images are therefore given by $\{u_{i,j,k}\}_{1 \leq i \leq N_x, 1 \leq j \leq N_y, 1 \leq k \leq N_z}$ and $\{f_{i,j,k}\}_{1 \leq i \leq N_x, 1 \leq j \leq N_y, 1 \leq k \leq N_z}$. Denote $\mathcal{X} = \mathbb{R}^N$ with $N = N_x \times N_y \times N_z$, a finite dimensional vector space equipped with a standard inner (scalar) product

$$\langle w, w' \rangle_{\mathcal{X}} = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \sum_{k=1}^{N_z} w_{i,j,k} w'_{i,j,k}. \tag{32}$$

The induced norm by the defined inner product is given by

$$\|w\|_{\mathcal{X}_2} = \sqrt{\langle w, w \rangle_{\mathcal{X}}} = \left(\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \sum_{k=1}^{N_z} w_{i,j,k}^2 \right)^{\frac{1}{2}}. \tag{33}$$

In the discrete setting, the blurring operator H corresponds to a discrete convolution which can be efficiently computed by using fast Fourier transform (FFT)[42;43;44;45]. To discretize $D_{2,\rho}$, we use standard finite differences to approximate the second derivatives along the three dimensions of the volume at voxel (i, j, k), with Neumann conditions on image boundaries:

$$(D_{2,\rho} u)_{i,j,k} = \begin{pmatrix} (1-\rho)u_{i,j,k} \\ \rho(\Delta_{1,1}^{(2)} u)_{i,j,k} \\ \rho(\Delta_{2,2}^{(2)} u)_{i,j,k} \\ \rho(\Delta_{3,3}^{(2)} u)_{i,j,k} \\ \rho(\Delta_{1,2}^{(2)} u)_{i,j,k} \\ \rho(\Delta_{1,3}^{(2)} u)_{i,j,k} \\ \rho(\Delta_{2,3}^{(2)} u)_{i,j,k} \\ \rho(\Delta_{2,1}^{(2)} u)_{i,j,k} \\ \rho(\Delta_{3,1}^{(2)} u)_{i,j,k} \\ \rho(\Delta_{3,2}^{(2)} u)_{i,j,k} \end{pmatrix} \in \mathbb{R}^{10}, \tag{34}$$

where

$$(\Delta_{1,1}^{(2)} u)_{i,j,k} = \begin{cases} u_{i+1,j,k} - 2u_{i,j,k} + u_{i-1,j,k} & \text{if } 1 < i < N_x, \\ 0 & \text{otherwise,} \end{cases}$$

$$(\Delta_{2,2}^{(2)} u)_{i,j,k} = \begin{cases} u_{i,j+1,k} - 2u_{i,j,k} + u_{i,j-1,k} & \text{if } 1 < j < N_y, \\ 0 & \text{otherwise,} \end{cases}$$

$$(\Delta_{3,3}^{(2)} u)_{i,j,k} = \begin{cases} \delta^2(u_{i,j,k+1} - 2u_{i,j,k} + u_{i,j,k-1}) & \text{if } 1 < k < N_z, \\ 0 & \text{otherwise,} \end{cases}$$

$$(\Delta_{1,2}^{(2)} u)_{i,j,k} = (\Delta_{2,1}^{(2)} u)_{i,j,k} = \begin{cases} (u_{i+1,j+1,k} - u_{i+1,j,k} - u_{i,j+1,k} + u_{i,j,k}) & \text{if } i < N_x, j < N_y \text{ and } k < N_z, \\ 0 & \text{otherwise.} \end{cases}$$

$$(\Delta_{1,3}^{(2)} u)_{i,j,k} = (\Delta_{3,1}^{(2)} u)_{i,j,k} = \begin{cases} \delta(u_{i+1,j,k+1} - u_{i+1,j,k} - u_{i,j,k+1} + u_{i,j,k}) & \text{if } i < N_x, j < N_y \text{ and } k < N_z, \\ 0 & \text{otherwise.} \end{cases}$$

$$(\Delta_{2,3}^{(2)} u)_{i,j,k} = (\Delta_{3,2}^{(2)} u)_{i,j,k} = \begin{cases} \delta(u_{i,j+1,k+1} - u_{i,j+1,k} - u_{i,j,k+1} + u_{i,j,k}) & \text{if } i < N_x, j < N_y \text{ and } k < N_z, \\ 0 & \text{otherwise.} \end{cases}$$

and δ is the ratio of the lateral-to-axial step sizes.

The discrete operators can be used to define the discrete SHV regularizer as:

$$\|D_{2,\rho}u\|_2 = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \sum_{k=1}^{N_z} \|(D_{2,\rho}u)_{i,j}\|_2, \tag{36}$$

where the L_1-norm acts now on the discrete domain Λ. The 3D deconvolution problem is defined in the discrete setting as the minimizer of the following energy:

$$E(u) = \frac{1}{2}\|Hu - f\|_2^2 + \lambda\|D_{2,\rho}u\|_2 + \imath_{\mathcal{C}}, \tag{37}$$

where $\lambda > 0$ is the regularization parameter and $\imath_{\mathcal{C}}$ is the indicator of a convex set $\mathcal{C}$ such as:

$$\mathcal{C} = \{u : \underline{u}_{i,j,k} \leq u_{i,j,k} \leq u_{\max}, \ 1 \leq i \leq N_x, \ 1 \leq j \leq N_y, \ 1 \leq k \leq N_z\}, \tag{38}$$

where the upper bound $u_{\max} > 0$ is the maximal intensity value allowed and $\underline{u}_{i,j,k} \geq 0$ is an estimated lower bound of the pixel intensity which is spatially varying and then adapted to each pixel location. The spatially varying constraint on the lower bound of pixel intensity not only guarantees positivity of the solution but also helps to avoid over-sparsifying effect. In our experiment, for the sake of simplicity, we use the following lower bound $\underline{u}_{i,j,k} = \max(0, \tilde{f}_{i,j,k} - c\tau)$ where $c > 0$ and $\tilde{f}$ is a smoothed version of the observed noisy image f by a low-pass (Gaussian) filter.

Energy minimization and splitting algorithms

We notice that the objective function

$$E(u) = \frac{1}{2}\|Hu - f\|_2^2 + \lambda\|D_{2,\rho}u)\|_2 + \imath_{\mathcal{C}}, \tag{39}$$

is a sum of linear composite functions as $u \mapsto \sum_{n=1}^{3} \mathscr{F}_m(L_m u)$, where each $\mathscr{F}_m$ is a convex function and each L_m is a linear bounded operator. Formally, we can write $\mathscr{F}_1 = \imath_{\mathcal{C}}$, $L_1 = \mathrm{Id}$, $\mathscr{F}_2 = \lambda\|\cdot\|_2$, $L_2 = D_{2,\rho}$ and $\mathscr{F}_3 = \frac{1}{2}\|\cdot\|^2$, $L_3 = H(\cdot) - f$. Generic primal-dual proximal approaches can be used to minimize this linear combination of convex functions as proposed in[61;62], but it is not optimal since the smoothness of the quadratic terms $\frac{1}{2}\|Hu - f\|_2$ is not exploited. In order to solve the problem (39), the design of an appropriate algorithm requires therefore to take into account the specific form of the corresponding energy, i.e., the sum of a simple convex function $\mathscr{F} = \imath_{\mathcal{C}}$, a more sophisticated composite function $\mathscr{G} \circ L = \lambda\|D_{2,\rho}(\cdot)\|_2$ (here, $\mathscr{G} = \lambda\|\cdot\|_2$ and $L = D_{2,\rho}$) and a smooth function $\mathscr{H} = \frac{1}{2}\|H(\cdot) - f\|_2$.

In what follows, we present a first-order method to minimize the sum of convex functions based on the proximal splitting approaches[25;26;48;49;50;51;52;53]. It consists in decomposing (splitting) the original problem into several simple sub-problems in the way that each single function of the sum can be processed separately. Indeed, smooth function involves its gradient operator, while non-smooth function implies its Moreau proximity operator[63]. These operators are well-suited for large-scale problems arising in signal and image processing, because they only exploit first-order information of the function and thus enable fast and efficient computation.

Let us recall first that the proximity operator of a convex function $\mathcal{J} : \mathbb{R}^N \to \mathbb{R}$ is defined as

$$\mathrm{prox}_{\tau\mathcal{J}}(u) = \arg\min_{w \in \mathbb{R}^N} \ \mathcal{J}(w) + \frac{1}{2\tau}\|u - w\|_2^2, \tag{40}$$

where $\tau > 0$ is a control parameter. From this definition, it easy to verify that the proximity operator of the function $\mathscr{F}(u) = \imath_{\mathcal{C}}(u)$ is nothing else than the projection onto the convex subset $\mathcal{C}$ as the following

$$\begin{aligned} \mathrm{prox}_{\tau\mathscr{F}}(u) &= \arg\min_{w \in \mathbb{R}^N} \ \imath_{\mathcal{C}}(w) + \frac{1}{2\tau}\|u - w\|_2^2 \\ &= \arg\min_{w \in \mathcal{C}} \ \|u - w\|_2^2. \end{aligned} \tag{41}$$

If we denote $\mathrm{proj}_{\mathcal{C}}$ the projection operator on $\mathcal{C}$, its closed-form expression is given by

$$\left(\mathrm{proj}_{\mathcal{C}}(u)\right)_{i,j,k} = \max(\underline{u}_{i,j,k}, \min(u_{\max}, u_{i,j,k})). \tag{42}$$

Moreover, the quadratic function $\mathscr{H}(u) = \frac{1}{2}\|Hu - f\|_2^2$ possesses an analytic form for its associated proximity operator

$$\begin{aligned} \mathrm{prox}_{\tau\mathscr{H}}(u) &= \arg\min_{w \in \mathbb{R}^N} \ \frac{1}{2}\|Hw - f\|_2^2 + \frac{1}{2\tau}\|u - w\|_2^2 \\ &= (\tau H^\top H + \mathrm{Id})^{-1}(\tau H^\top Hf + u), \end{aligned} \tag{43}$$

where the symbol $^\top$ denotes the adjoint of a linear operator and $H^\top : \mathbb{R}^N \to \mathbb{R}^N$ satisfies $\langle Hw, w'\rangle_{\mathbb{R}^N} = \langle w, H^\top w'\rangle_{\mathbb{R}^N}$. The evaluation of $\mathrm{prox}_{\tau\mathscr{H}}(u)$ corresponds to the inverse of a linear system that is not always possible in practice due to the high dimensionality of the problem. For this reason, the optimization methods which involve the gradient of $\mathscr{H}$ are more appropriate since they do not require any inverse operator. In the comparison with $\mathscr{F}$ and $\mathscr{H}$, the calculation of the proximity operator in the case of the composite function $\mathscr{G} \circ L(u) = \lambda\|D_{2,\rho}u\|_2$ is theoretically possible but is challenging because of the presence of $D_{2,\rho}$ which is not diagonal.

To solve the minimization problem (39), we adopt the full splitting approach described in [25;26]. The key idea of this approach is to evaluate the gradient, proximity and linear operators individually in order to avoid implicit operations such as inner loops or inverse of linear operators. Accordingly, only "simple" computations are considered such as the gradient $\nabla\mathscr{H}$, the proximity operator of $\mathscr{F}$ and $\mathscr{G}$, the linear mapping L and its adjoint operators $L^\top$. The corresponding proximal algorithm for the problem is written under the following general form at iteration ℓ:

$$u^{(\ell+1)} = \mathrm{prox}_{\gamma\mathscr{F}}\left\{u^{(\ell)} - \gamma\left(L^\top z^{(\ell)} + \nabla\mathscr{H}(u^{(\ell)})\right)\right\}, \tag{44a}$$

$$z^{(\ell+1)} = \mathrm{prox}_{\zeta\mathscr{G}^\star}\left\{z^{(\ell)} + \zeta L\left(2u^{(\ell+1)} - u^{(\ell)}\right)\right\}, \tag{44b}$$

where $\gamma, \zeta > 0$ are proximal parameters, $\mathscr{G}^\star$ denotes the Legendre-Fenchel conjugate of $\mathscr{G}$ and its proximity operator $\mathrm{prox}_{\zeta\mathscr{G}^\star}$ can be directly computed from $\mathrm{prox}_{\frac{\mathscr{G}}{\zeta}}$ by using the Moreau's identity $v = \zeta\,\mathrm{prox}_{\frac{\mathscr{G}}{\zeta}}\left(\frac{v}{\zeta}\right) + \mathrm{prox}_{\zeta\mathscr{G}^\star}(v)$. Following [25;26], to guarantee the convergence of the proposed algorithm, the parameters γ and ζ must fulfill the condition

$$\gamma\left(\frac{1}{2} + \zeta\|L^\top L\|\right) < 1, \tag{45}$$

where $\|\cdot\|$ denotes the operator norm. The proofs of the convergence can be found in [25]. We also note that the proposed algorithm belongs to the class of primal-dual algorithms which provide not only the solution of the primal problem (a.k.a. the original minimization problem) but also the solution of its dual problem.

Since the closed-form of $\mathrm{prox}_{\gamma\mathscr{F}}$ is already given, it remains to define the analytic expression of other terms in (44a) and (44b). We start with the gradient of the quadratic function $\mathscr{H}$ which is straightforwardly obtained by

$$\nabla\mathscr{H}(u) = H^\top(Hu - f). \tag{46}$$

Next, we notice that the regularization operator $L = D_{2,\rho}$ is a linear mapping, then its adjoint operator $L^\top = D_{2,\rho}^\top$ is defined using the equation $\langle u, D_{2,\rho}^\top v\rangle = \langle D_{2,\rho}u, v\rangle$, which implies ($d = 3$)

$$\begin{aligned}\left(D_{2,\rho}^{~\top}v\right)_{i,j,k} =~& (1-\rho)v_{i,j,k}^1 + \rho[(\Delta_{1,1}^{(2)})^\top v^2]_{i,j,k} + \rho[(\Delta_{2,2}^{(2)})^\top v^3]_{i,j,k} + \rho[(\Delta_{3,3}^{(2)})^\top v^4]_{i,j,k} \\ &+ \rho[(\Delta_{1,2}^{(2)})^\top v^5]_{i,j,k} + \rho[(\Delta_{1,3}^{(2)})^\top v^6]_{i,j,k} + \rho[(\Delta_{2,3}^{(2)})^\top v^7]_{i,j,k},\end{aligned} \tag{47}$$

where the involving adjoint operators are given below:

$$[(\Delta_{1,1}^{(2)})^\top w]_{i,j,k} = \begin{cases} w_{i+1,j,k} - 2w_{i,j,k} + w_{i-1,j,k} & \text{if } 1 < i < N_x \\ 0 & \text{otherwise} \end{cases}$$

$$[(\Delta_{2,2}^{(2)})^\top w]_{i,j,k} = \begin{cases} w_{i,j+1,k} - 2w_{i,j,k} + w_{i,j-1,k} & \text{if } 1 < j < N_y \\ 0 & \text{otherwise} \end{cases}$$

$$[(\Delta_{3,3}^{(2)})^\top w]_{i,j,k} = \begin{cases} w_{i,j,k+1} - 2w_{i,j,k} + w_{i,j,k-1} & \text{if } 1 < k < N_z \\ 0 & \text{otherwise} \end{cases}$$

$$[(\Delta_{1,2}^{(2)})^\top w]_{i,j,k} = [(\Delta_{2,1}^{(2)})^\top w]_{i,j,k} = \begin{cases} w_{i,j,k} & \text{if } i = 1, j = 1, 1 \le k \le N_z \\ w_{i,j,k} - w_{i,j-1,k} & \text{if } i = 1, 1 < j < N_y, 1 \le k \le N_z; \\ -w_{i,j-1,k} & \text{if } i = 1, j = N_y, 1 \le k \le N_z; \\ \\ w_{i,j,k} - w_{i-1,j,k} & \text{if } 1 < i < N_x, j = 1, 1 \le k \le N_z; \\ w_{i,j,k} - w_{i-1,j,k} - w_{i,j-1,k} + w_{i-1,j-1,k} & \text{if } 1 < i < N_x, 1 < j < N_y, 1 \le k \le N_z; \\ -w_{i,j-1,k} + w_{i-1,j-1,k} & \text{if } 1 < i < N_x, j = N_y, 1 \le k \le N_z; \\ \\ -w_{i-1,j,k} & \text{if } i = N_x, j = 1, 1 \le k \le N_z; \\ -w_{i-1,j,k} + w_{i-1,j-1,k} & \text{if } i = N_x, 1 < j < N_y, 1 \le k \le N_z; \\ w_{i-1,j-1,k} & \text{if } i = N_x, j = N_y, 1 \le k \le N_z. \end{cases}$$

$$[(\Delta^{(2)}_{1,3})^\top w]_{i,j,k} = [(\Delta^{(2)}_{3,1})^\top w]_{i,j,k} = \begin{cases} w_{i,j,k} & \text{if } i = 1, k = 1, 1 \le j \le N_y \\ w_{i,j,k} - w_{i,j,k-1} & \text{if } i = 1, 1 < k < N_z, 1 \le j \le N_y; \\ -w_{i,j,k-1} & \text{if } i = 1, k = N_z, 1 \le j \le N_y; \\ \\ w_{i,j,k} - w_{i-1,j,k} & \text{if } 1 < i < N_x, k = 1, 1 \le j \le N_y; \\ w_{i,j,k} - w_{i-1,j,k} - w_{i,j,k-1} + w_{i-1,j,k-1} & \text{if } 1 < i < N_x, 1 < k < N_z, 1 \le j \le N_y; \\ -w_{i,j,k-1} + w_{i-1,j,k-1} & \text{if } 1 < i < N_x, k = N_z, 1 \le j \le N_y; \\ \\ -w_{i-1,j,k} & \text{if } i = N_x, k = 1, 1 \le j \le N_y; \\ -w_{i-1,j,k} + w_{i-1,j,k-1} & \text{if } i = N_x, 1 < k < N_z, 1 \le j \le N_y; \\ w_{i-1,j,k-1} & \text{if } i = N_x, k = N_z, 1 \le j \le N_y. \end{cases}$$

$$[(\Delta^{(2)}_{2,3})^\top w]_{i,j,k} = [(\Delta^{(2)}_{3,2})^\top w]_{i,j,k} = \begin{cases} w_{i,j,k} & \text{if } j = 1, k = 1, 1 \le i \le N_x \\ w_{i,j,k} - w_{i,j,k-1} & \text{if } j = 1, 1 < k < N_z, 1 \le i \le N_x; \\ -w_{i,j,k-1} & \text{if } j = 1, k = N_z, 1 \le i \le N_x; \\ \\ w_{i,j,k} - w_{i,j-1,k} & \text{if } 1 < j < N_y, k = 1, 1 \le i \le N_x; \\ w_{i,j,k} - w_{i,j-1,k} - w_{i,j,k-1} + w_{i,j-1,k-1} & \text{if } 1 < j < N_y, 1 < k < N_z, 1 \le i \le N_x; \\ -w_{i,j,k-1} + w_{i,j-1,k-1} & \text{if } 1 < j < N_y, k = N_z, 1 \le i \le N_x; \\ \\ -w_{i,j-1,k} & \text{if } j = N_y, k = 1, 1 \le i \le N_x; \\ -w_{i,j-1,k} + w_{i,j-1,k-1} & \text{if } j = N_y, 1 < k < N_z, 1 \le i \le N_x; \\ w_{i,j-1,k-1} & \text{if } j = N_y, k = N_z, 1 \le i \le N_x. \end{cases}$$

From equations (47), one can deduce the following upper bound:

$$\|D_{2,\rho}{}^\top D_{2,\rho}\| \le (1-\rho)^2 + 64\rho^2 \, , \tag{48a}$$

that are used for choosing the proximal parameters γ and ζ according to (45).

The last term we want to deal with is the proximity operator $\mathrm{prox}_{\zeta \mathscr{G}^\star}$. We also note that the proposed primal-dual algorithm does not necessitate evaluating the proximity operator of the composite function $\mathscr{G} \circ L$ as in the case of generic proximal algorithms, but only $\mathrm{prox}_{\frac{\mathscr{G}}{\zeta}}$ is required. Since $\mathscr{G}$ is related to the mixed norm whose the proximity operator is defined as:

$$\big(\mathrm{prox}_{\tau\|\cdot\|_2}(v) \big)^\kappa_{i,j,k} = \max(0, \|L_{i,j,k}v\|_2 - \tau) \frac{v^\kappa_{i,j,k}}{\|L_{i,j,k}v\|_2} \, , \tag{49}$$

where $L_{i,j,k} : v \in \mathcal{X}^{10} \mapsto (v^\kappa_{i,j,k})_{1 \le \kappa \le 10} \in \mathbb{R}^{10}$ is a linear operator ($d = 3$). By using the Moreau's identity, we obtain the closed-form expression of $\mathrm{prox}_{\zeta \mathscr{G}^\star}$ as the following:

$$\big(\mathrm{prox}_{\zeta \mathscr{G}^\star}(v) \big)^\kappa_{i,j,k} = \frac{v^\kappa_{i,j,k}}{\max\left(1, \frac{\|L_{i,j,k}v\|_2}{\lambda}\right)} \, , \tag{50}$$

which shows that $\mathrm{prox}_{\zeta \mathscr{G}^\star}$ is independent from ζ and moreover it is an pointwise operator. These properties allow therefore fast and efficient computation by exploiting the intrinsic parallelism of multicore processors.

Energy model for 4D denoising

A temporal series of 3D noisy images $f : \Omega \subset \mathbb{R}^4 \to \mathbb{R}$ is a noisy version of the underlying true 3D image sequence $u : \Omega \to \mathbb{R}$ modeled as follows: $f = \mathcal{T}(u)$. It follows that the denoising problem is formulated as the minimization of an energy functional defined as

$$\int_\Omega (u(\mathbf{x}) - f(\mathbf{x}))^2 \ + \ \lambda\|D_{2,\rho}u(\mathbf{x})\|_2 \, d\mathbf{x} \ + \ \imath_{\boldsymbol{\mathcal{C}}}(u) \tag{51}$$

where

$$\begin{aligned} D_{2,\rho}u(\mathbf{x}) \ := \ \Big(& (1-\rho)u(\mathbf{x}), \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{xx}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{yy}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{zz}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{tt}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{xy}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{xz}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{xt}}, \\ & \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{yx}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{yz}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{yt}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{zx}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{zy}}, \ \rho\frac{\partial^2 u(\mathbf{x})}{\partial_{zt}} \Big) \in \mathbb{R}^{14} \, . \end{aligned} \tag{52}$$

To solve the optimization problem, we split the original problem (51) into several simple sub-problems as explained above. The resulting individual functions involved in the sum are then minimized separately. The algorithm is expected to be faster since FFT is not required in the implementation for image denoising. Nevertheless, the amount of data is much more larger since the whole 4D image sequence is denoised at once. In practice, a very long image sequence can be segmented into sub-sequences with a small overlap, and is processed independently and in parallel. The overlapping images can be averaged to reduce possible artifacts at the end.

C. Microscopy and cell cultures

Cell culture fluorescence labeling.
The hTERT-immortalized RPE1 cells (Human Retinal Pigment Epithelial Cell) were purchased from ATCC (CRL-4000). RPE1 cells within 30 passage number were used and grown in DMEM-F12 medium without phenol red (Life Technologies) supplemented with 10% (vol/vol) fetal bovine serum (FBS) at 37°C in humidified atmosphere with 5% CO2. For LLSM (Lattice Light Sheet Microscopy) imaging, cells were seeded 5 to 6 hours before image acquisition at 150.103 per well of a 6 well plate, containing 3 to 4 coverslips with a 5mm diameter, treated as previously described in (Chen *et al.* 2012)[1]. Live cells were stained for microtubule and mitochondria using the Tubulin TrackerTM DR (Tub-Tracker Deep Red, ThermoFisher Scientific) probe (Ex. 652nm/Em. 669) at 10 ng/ml and 100ng/ml of PKMR (PK Mito Red), a cyclooctatetraene-conjugated-cyanine-3 dye (Ex. 549/Em. 569) kindly provided by Dr. Z. Chen from Peking University (Yang *et al.* 2020)[2], respectively. Cell labeling was performed in two steps. Cells were first incubated for 15min with Tubulin Tracker DR and washed twice with LLSM medium (DMEM/F-12, without phenol red, with 1% BSA, 0.01% penicillin–streptomycin, 1mM pyruvate, and 20mM HEPES). Coverslips were then transferred to the lattice light-sheet microscope (LLSM) sample holder and inserted into the imaging chamber containing 6 ml of the later medium and 0.1 mg/ml of PKMR.

Previously described Hela cells expressing Rab5-eGFP CRISP genome edited SUM 159 breast carcinoma cell line, using eGFP-tagged σ unit of the AP2 complex (Adaptor Protein complex) were kindly provided by Dr. Tomas Kirchhausen (HMS, Boston, MS, USA) and CD-M6PR-EGFP expressing Hela cell line, by Dr Bernard Hoflack (TU-Dresden, Dresden, Germany). Human retinal pigment epithelial cells (RPE1) expressing mCherry-LifeAct were kindly provided by Dr Manuel Thery (Vignaud *et al.* 2012)[3]. All these cells were cultured, prepared and imaged with the LLSM as described above.

For STED imaging RPE1 cells were plated on glass bottom Petri dishes (μ-dish 35mm-High,1.5H, Ibidi, GmbH, Grafelfing, Germany). Microtubule labeling was performed as for LLSM imaging, while mitochondria were labeled by diluting 16000 $\times$ a 1mM stock solution of PKMITO^{TM} Orange [gifted by Dr. Z. Chen from Peking University and Genvivo Biotech, Beijing, PR of China] in DMEM/F12 medium for a final 62.5 nM, for 15 minutes at 37°C before image acquisition.

Previously acquired MFM datasets of U2OS10 cells transfected with TOM20 (translocase of outer mitochondrial membrane) fused to GFP (GFP-TOM20) (Hajj *et al.* 2014)[4] were used in this study.

Previously acquired datasets of M10 cells, stably expressing Langerin tagged with eYFP, or Rab11a tagged with mCherry (Gidon *et al.* 2012)[5]) or transiently transfected with PhLuorin Transferrin-Receptor (TfR) (Boulanger *et al.* 2014)[6]) were also exploited in this study.

Microscopy.
Lattice light-sheet microscopy (LLSM):   Lattice light-sheet microscopy was done on a commercialized version of a previously described setup Chen *et al.* 2012)[1] from 3i (Denver, USA). Cells were scanned incrementally through a 20 μm long light sheet in 600 nm steps using a fast piezoelectric flexure stage equivalent to $\simeq$ 325 nm with respect to the detection objective and were imaged using a sCMOS camera (Orca-Flash 4.0; Hamamatsu, Bridgewater, NJ). Excitation was achieved with 560- or 642-nm diode lasers (MPB Communications) at 10% and 15% acousto-optic tunable filter transmittance with 50 mW and 100 mW respectively (initial box power) through an excitation objective (Special Optics 28.6$\times$ 0.7 NA 3.74-mm water-dipping lens) and detected via a Nikon CFI Apo LWD 25$\times$ 1.1 NA water-dipping objective with a 2.5$\times$ tube lens with a final pixel size of 104 nm. Lattice light-sheet imaging was performed using an excitation pattern of outer NA equal to 0.55 and inner NA equal to 0.493. Composite volumetric datasets were obtained using $\simeq$10 ms exposure/slice/channel at a time resolution of 2.2 s per total cell volume (about 60 planes). Fifty-six time points were acquired within 3 to 4 minutes (total raw data, only 1:45 min were analyzed for the movie). Acquired data were deskewed, a necessary step to realigned image frames, using LLSpy, a python library to facilitate lattice light sheet data processing (copyright to T. Lambert, Harvard Medical School, Boston, USA; https://github.com/tlambert03/LLSpy). Deskewed images are then considered as Raw images. Napari, a multi-dimensional image viewer for Python (https://github.com/napari/napari), was used for 3D rendering. Maximum intensity projections were generated using ImageJ/Fiji 1.53c (Schindelin *et al.*, 2012)[7]. Intensity profiles were plotted using Matlab2019b. Related supplementary movies S1 and S2 were generated using Napari and napari-animation plugin (`https://github.com/napari/napari-animation`).

Multifocus microscopy (MFM):   MultiFocus microscopy (MFM)[31] allows simultaneous acquisition of a 3D stack of nine equally spaced focal 2D images with a frame rate up to 100 volumes per second, thus covering the whole volume of a nucleus or cell

[1] B.-C. Chen *et al.*, *Science*, 346(6208): 1257998 (2014). (doi:10.1126/science.1257998)

[2] Z. Yang *et al.*, *Chem. Sci.*, 11:8506 (2020). (doi:10.1039/d0sc02837a)

[3] Vignaud *et al.* J. Cell Sci. 125(9): 2134-2140 (2012). (doi: 10.1242/jcs.104901)

[4] B. Hajj *et al.* Proc. Natl. Acad. Sci. USA, 111:17480-17485 (2014). (doi: 10.1073/pnas.1412396111)

[5] A. Gidon *et al.*, Traffic, 13(6):815-833 (2012). (doi:10.1111/j.1600-0854.2012.01354.x)

[6] J. Boulanger *et al.*, Proc. Natl. Acad. Sci. US, 111(8): 7164-17169 (2014). (doi:10.1073/pnas.1414106111)

[7] Schindelin *et al.*, *Nat. Methods*, 9(7), 676–682 (2012). (doi:10.1038/nmeth.2019)

components in a single exposure[31]. The set-up is based on a diffraction grating that forms multiple focus-shifted images followed by a chromatic correction module. This imaging technique is especially designed for high-speed imaging, and is able to capture the signal of weak fluorescent samples such as single fluorophores.

Stimulated emission depletion microscopy (STED):   Image acquisition was performed with a STEDYCON module (Abberior Instruments, Göttingen, Germany) mounted at the camera port TCS SP8 STED microscope (Leica, Manheim, Germany) with a HC PL APO C2S ×100 oil objective (1.4 NA) used in 2D mode. Depletion was performed with a 775 nm pulsed laser 1-7 ns at 80% (413 mW at the objective lens position). Labeled mitochondria and microtubules were imaged with excitation at a wavelength of 594 nm and 640 nm, respectively. Nominal laser powers were adjusted at 7% for the 594 nm laser (3.15 μW at the objective lens position) and at 10% for the 640 nm (26 mW at the objective lens position). The time-gated fluorescence detection was done on a detector single photon counting avalanche photodiode between 578-627 nm and 650-700 nm with a pinhole settled at 1.1 AU. Imaging was executed with the 9 lines accumulation. Pixel dwell time was 10 μs and a pixel size of 25 nm × 25 nm. Images were generated using ImageJ/Fiji 1.53c3. Line profiles plots were measured with ImageJ. Then line profiles were fitted to a Gaussian model using Matlab. Finally, Full width at half maximum (FWHM) is estimated from the fitting results.

Other Imaging modalities: Spinning disk confocal microscopy (CM) and Ring Total Internal Reflection microscopy (RING-TIRFM) were either used as described before (Gidon *et al.* 2012)[3], (Pecot *et al.* 2018)[8] or previously acquired datasets ((Boulanger *et al.* 2014)[4]; (Basset *et al.* 2015)[9]) reused, in this study.

[8]T. Pécot *et al.*, eLife, 7:e32311 (2018). (doi:10.7554/eLife.32311)

[9]A. Basset *et al.*, *IEEE Trans. Image Process.*, 24(11):4512-4527 (2015). (doi:10.1109/TIP.2015.2450996)

D. Quantitative evaluation and method comparison on 2D artificially noisy and blurred fluorescence images

We demonstrate that SPITFIR(e) (based on Sparse Hessian Variation) provides the best PSNR (Peak Signal to Noise Ratio) values when the ground truth and the PSF are known, when compared to other competitive methods reported in Table 1. As the implementation in 3D of several regularizers in Table 1 is not always possible, we conducted fair experiments on 2D images shown in Supplementary Table 2. We included the performance of classical deconvolution methods such as Richardson-Lucy (RL), iterative constrained Tikhonov-Miller (ICTM), Gold-Meinel (GM))[32] on the four tested images. The four ground-truth images in Supplementary Table 2 were normalized in the range $[0, 1]$ and further blurred by considering a Gaussian point spread function (PSF) with different standard deviation values σ_{xy}. A Gaussian noise with zero mean and variance τ^2 was also added to these images in order to generate observed noisy and blurry data. Most of competing methods are known to perform better if the images are corrupted by Gaussian noise. Note that, in the case of RL algorithm, which is originally designed to deal with Poisson noise, the degraded images are re-scaled to the original dynamics of the underlying reference images. The RL deconvolution results are then re-normalized for a fair comparison with those obtained by the other methods. Moreover, before applying GM, the noisy images are smoothed with a Gaussian filter as pre-processing step because this algorithm assumes that the noise is negligible. In these experiments, we consider three different PSF sizes which correspond to $\sigma_{xy} \in \{1, 1.25, 1.5\}$ and three distinct noise levels $\tau \in \{0.01, 0.02, 0.04\}$ (intensities are in the range $[0, 1]$).

The parameters involved in each method have been adjusted in order to get the highest PSNR values for each individual image in Supplementary Table 2. We can notice that the SHV-based deconvolution approach outperforms the competing methods in most cases (highest PSNR scores in 21 over 36 cases). In these cases, SV and TV-L_1, which are respectively the second (12/36) and the third (7/36) position on the benchmark, achieve slightly lower PSNR scores while comparing with SHV. In contrast to these results, the non-regularization methods such as RL and GM algorithm are often ranked at the end of the benchmark. Last but not least, the GR measure that we thought a good alternative for TV, provides disappointing deconvolution outcomes in terms of PSNR values on artificially corrupted images.

In terms of visual assessment, SPITFIR(e) avoid over-smoothing and over-sharpening effect while preserving details and removing noise in the background, as illustrated on Image #1 in Supplementary Fig. 6. In Supplementary Fig. 6(**c**)-(**d**), we notice noise amplification with RL (**c**) and GM (**d**), resulting in the apparition of undesired high-intensity pixels (i.e. the "night sky" effect) and unrealistic reconstructed structures since the high-frequency components were not correctly restored. Unlike RL and GM algorithms, ICTM (**e**) and GraphNet (**h**) produce deconvolution results without noise amplification artifacts, but they tend to over-smooth (blur) image details due to the quadratic nature of Tikhonov-Miller penalty. In contrast to these over-smoothed results, those obtained with TV (**f**), TV-L_1 (**g**) and SV (**l**) regularizers generate artificial sharper edges (staircasing effects), respectively. As depicted in Supplementary Fig. 6(**m**), HV which promotes piecewise linear reconstructions, provides visually more pleasant deconvolution results. Similar visual results are obtained with SPITFIR(e) (SHV, (**n**)), which enforce the co-localization of high-intensity pixels and non-zero directional derivatives, and allow to obtain non-zero regions with eventually high-magnitude gradients. It results in better reconstruction of the image foreground, in which large variation between pixels at the object center and those at the object boundary may occur. Regarding smooth-approximation based regularizations, TVH (**i**) and LSHV (**j**) provide results with restored details which are slightly sharper than those obtained with ICTM (**e**), but noise is not sufficiently removed. It not only disturbs the visual effect but also makes the detection of small or low-contrast objects more difficult. In our opinion, this issue is due to the smooth nature of the approximation for small values of directional derivatives, resulting in noisy reconstructed gradients. As consequence, the PSNR values obtained with these two regularizers are sharply inferior to those obtained with SV, HV and SHV. In the comparison with TVH (**i**) and LSHV (**j**), the smooth version of GR (**k**) yields visual similar deconvolution results. However, GR (**k**) suffer from severe "white pixel" artifacts in the background region. These "white pixels" are discarded to fairly compute the PSNR values. Overall, SHV (SPITFIR(e)) (**n**) provided the most satisfying results in this comparison experiment.