



Total variation and spatial iteration-based 3D structured illumination microscopy

MINGXUAN CAI,^{1,6} HONGFEI ZHU,^{2,6} YILE SUN,¹ LU YIN,¹
FANGHUI XU,¹ HANMENG WU,¹ XIANG HAO,¹ RENJIE ZHOU,²
CUIFANG KUANG,^{1,3,4,5,*} AND XU LIU^{1,3}

¹State Key Laboratory of Modern Optical Instrumentation, College of Optical Science and Engineering, Zhejiang University, Hangzhou 310027, China

²Department of Biomedical Engineering, The Chinese University of Hong Kong, Hongkong, China

³Research Center for Intelligent Chips and Devices, Zhejiang Lab, Hangzhou 311121, China

⁴China Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan 030006, China

⁵Ningbo Research Institute, Zhejiang University, Ningbo 315100, China

⁶These authors contributed equally to this work

*cfkuang@zju.edu.cn

Abstract: Three-dimensional structured illumination microscopy (3D-SIM) plays an essential role in biological volumetric imaging with the capabilities of improving lateral and axial resolution. However, the traditional linear 3D algorithm is sensitive to noise and generates artifacts, while the low temporal resolution hinders live-cell imaging. In this paper, we propose a novel 3D-SIM algorithm based on total variation (TV) and fast iterative shrinkage threshold algorithm (FISTA), termed TV-FISTA-SIM. Compared to conventional algorithms, TV-FISTA-SIM achieves higher reconstruction fidelity with the least artifacts, even when the signal-to-noise ratio (SNR) is as low as 5 dB, and a faster reconstruction rate. Through simulation, we have verified that TV-FISTA-SIM can effectively reduce the amount of required data with less deterioration. Moreover, we demonstrate TV-FISTA-SIM for high-quality multi-color 3D super-resolution imaging, which can be potentially applied to live-cell imaging applications.

© 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

With the advantage of chemical specificity, optical fluorescence microscopes have been widely used in the field of biological research and medical diagnosis. However, due to the optical diffraction limit, the resolution of a conventional optical microscope is constrained to around 200 nm. To circumvent the resolution limit, numerous super-resolution imaging techniques have been developed since the 1990s, including Stimulated Emission Depletion (STED) [1], Structured Illumination Microscopy (SIM) [2–4], Photoactivated Localization Microscopy (PALM) [5,6], and Stochastic Optical Reconstruction Microscopy (STORM) [7,8]. STORM and PALM, classified as Single-Molecule Localization Microscopy (SMLM), are based on localizing each emitter in each sparsely labeled image. As many images are required, the imaging acquisition period is inevitably prolonged, making live-cell imaging difficult. On the other hand, STED relies on using a doughnut beam to deplete and reshape the Point Spread Function (PSF). Though the resolution of STED is improved by utilizing the doughnut beam, photobleaching and phototoxicity are also introduced. In contrast, SIM is popular in cell bioimaging, especially in live-cell imaging, because of its short imaging duration and relatively low excitation intensity.

In recent years, three-dimensional SIM (3D-SIM) has become a much more desirable imaging modality for biomedical studies [9–11], and improving the axial resolution is a core target for achieving high-quality 3D imaging. Compared with 2D-SIM, another excitation beam is introduced in 3D-SIM, and the spatial fringes are formed by the interference of three excitation beams, thus extending the Optical Transfer Function (OTF) in the axial direction [12]. To improve

the axial resolution, researchers have combined 4Pi structure to double the light collecting angle, such as 4Pi confocal fluorescence microscope [13], image interference microscopy (I^2M) [14], and an advanced method combined I^2M with incoherent interference illumination (I^5M) [15]. A new type of wide-field light microscopy, I^5S , which means the combination of the lateral performance of SIM and the axial performance of I^5M , has been recently applied to 3D-SIM [16]. By interfering with six beams in the object space, the OTF in I^5S is further extended in the axial direction, thus offering an axial resolution improvement when compared with conventional 3D-SIM using a single objective lens, thus providing more biological applications.

For SIM reconstruction, many linear reconstruction algorithms have been proposed [2,17–20]. However, the fidelity of those algorithms is often challenged [19,21] due to the reconstruction artifacts caused by post-processing algorithms [17,22,23]. In the meantime, the axial scanning process with a certain number of spatial fringes shifting will take a relatively long duration which is not conducive for live-cell imaging, although it is indispensable in 3D-SIM. It has been reported that the number of images needed in 2D-SIM can be reduced to four to shorten the duration of image acquisition [24], but the image quality will deteriorate significantly, and the concept has not been transferred to volumetric imaging. Deep Learning is also a plausible method for the reconstruction of SIM [25–28]. However, the performance is excessively dependent on ground truth, and transfer learning is inevitable for different samples and systems. The nonlinear gradient descent structured illumination microscopy (NGD-SIM) [29], proposed by us earlier, provides a way to reconstruct 3D-SIM images with less data. However, the reconstruction speed is too slow, and the performance is not that ideal when in high-noise conditions.

To date, though excellent achievements have been made in 2D-SIM [17,19], there is still a lack of 3D-SIM algorithms that are efficient enough to reduce the requirement of the image number and obtain a better image quality at the same time. Here we propose a novel 3D-SIM algorithm based on total variation (TV) [30] and fast iterative shrinkage threshold algorithm (FISTA) [31], called TV-FISTA-SIM. TV-FISTA-SIM realizes faster convergence speed and reduces its convergence complexity to $O(1/t^2)$ [30]. Under the condition of reduced raw image data and high noise, TV-FISTA-SIM can still recover high-quality super-resolution images due to its breakneck speed of convergence and the constraints of total variation regularization, enabling the possibility of live-cell 3D-SIM imaging. In addition, we publicly released the codes to benefit the community and expect the proposed TV-FISTA-SIM would expand the application scope of 3D-SIM in the future.

2. Principle

2.1. 3D-SIM reconstruction

In 3D-SIM, fluorescent samples are illuminated with spatially modulated patterns varying in lateral and axial dimensions. If fluorophore density $S(\mathbf{r})$ is excited by an illumination pattern $I(\mathbf{r})$, the resulting emission distribution $E(\mathbf{r})$ is

$$E(\mathbf{r}) = S(\mathbf{r})I(\mathbf{r}). \quad (1)$$

Considering the point spread function (PSF) of the imaging system, the obtained data $D(\mathbf{r})$ can be represented as a convolution of PSF with the product of illumination pattern $I(\mathbf{r})$ and fluorophore density $S(\mathbf{r})$ as:

$$D(\mathbf{r}) = [I(\mathbf{r}) \cdot S(\mathbf{r})] \otimes PSF(\mathbf{r}). \quad (2)$$

Transforming Eq. (2) to the Fourier domain, the spectrum of the images can be formulated as:

$$\tilde{D}(\mathbf{k}) = [\tilde{I}(\mathbf{k}) \otimes \tilde{S}(\mathbf{k})] \cdot OTF(\mathbf{k}), \quad (3)$$

where $OTF(\mathbf{k})$ represents the optical transfer function of the microscopy system. The existence of OTF blocks the high-frequency information, thus constraining the spatial resolution of

conventional wide-field microscopy. In 3D-SIM, the illumination pattern $I(\mathbf{r})$ can be represented as the sum of spatial components of different harmonic functions [3],

$$I(\mathbf{r}_{xy}, z) = \sum_m I_m(z) J_m(\mathbf{r}_{xy}), \quad (4)$$

where $m = 0, \pm 1, \pm 2$, $I_m(z)$ and $J_m(\mathbf{r}_{xy})$ denote two individual functions axially and laterally. The illumination patterns are kept fixed to the focal plane. When obtaining raw 3D-SIM data, the sample region of interest is placed at the focal plane. Therefore, the depth-resolved emission density can be represented as:

$$E(\mathbf{r}) = \sum_{z'} \sum_m S(\mathbf{r}) I_m(z' - z) J_m(\mathbf{r}_{xy}), \quad (5)$$

where z and z' refer to the axial position of the sample and illumination pattern, respectively. Thus, the spectrum of detected images can be transformed from Eq. (3) as:

$$\tilde{D}(\mathbf{k}) = \sum_m [\tilde{S}(\mathbf{k}) \otimes \tilde{J}_m(\mathbf{k}_{xy})] \cdot OTF_m(\mathbf{k}), \quad (6)$$

where $OTF_m(\mathbf{k})$ denotes of the Fourier transformation of $PSF(\mathbf{r}) \cdot I_m(z)$ in order of m . Due to the assume that $J_m(\mathbf{r}_{xy})$ and $I_m(z)$ can be represented as the harmonic functions [3] like $J_m(\mathbf{r}_{xy}) = e^{im(2\pi p r_{xy} + \varphi)}$, Eq. (6) can be formulated as:

$$\tilde{D}(\mathbf{k}) = \sum_m e^{im\varphi} \tilde{S}(\mathbf{k} - m\mathbf{p}) \cdot OTF_m(\mathbf{k}). \quad (7)$$

This shows that every raw 3D-SIM image can be understood as the sum of different orders of frequency components.

In practice, 3D-SIM generally requires periodic illumination patterns from three directions, and five phase shifts are required in every direction. Therefore, a total of fifteen raw images are required for reconstructing one slice in a volume. With the periodic illuminating patterns, undetectable high-frequency information will be shifted into the observation region and collected by the objective lens. Then, using a frequency-domain reconstruction algorithm in conventional 3D-SIM, different frequency components are moved back to their original positions to achieve an extended spectrum, thus improving the overall spatial resolution.

2.2. Total variation minimization

Considering a linear inverse problem: $\mathbf{y} = \mathbf{Hx} + \mathbf{n}$, where $\mathbf{x} \in \mathbb{R}^N$ is the expected target image calculated from the observation $\mathbf{y} \in \mathbb{R}^M$ that contains unknown noise, $\mathbf{H} \in \mathbb{R}^{M \times N}$ is the transfer function of a system, and $\mathbf{n} \in \mathbb{R}^M$ is the noise (Gaussian noise by default). When this problem is ill-conditioned, the standard way of estimating the target image is to solve the following optimization problem:

$$F(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x}), \quad (8)$$

where $f(\mathbf{x}) \triangleq \frac{1}{2} \|\mathbf{y} - \mathbf{Hx}\|_{l_2}^2$ is the traditional loss function of the l_2 norm and $g(\mathbf{x})$ denotes the regularization term, defined as [30]:

$$g(\mathbf{x}) = \lambda \|\mathbf{Dx}\|_{l_1} = \lambda \sum_{n=1}^N \sum_{d=1}^D |[\mathbf{D}_d \mathbf{x}]_n|, \quad (9)$$

where N is the number of pixels in the image, d represents the dimensionality of the image ($d = 3$ in volumetric images), \mathbf{D}_d denotes the dimensionality difference while ensuring proper edge conditions within the image, and $\|\cdot\|_{l_1}$ denotes the l_1 norm. Total variation regularization can effectively stabilize the solution of ill-posed inverse problems [30,32] and has proved successful in a wide range of applications in sparse recovery of images from incomplete and corrupted measurements [33–35].

2.3. Fast iterative shrinkage thresholding algorithm (FISTA)

Due to the existence of non-smooth regularization term in the loss function, one of the ideal methods of solving Eq. (7) is Iterative Shrinkage Thresholding Algorithm (ISTA) [36,37]:

$$\mathbf{x}^t \leftarrow p_L(\mathbf{x}^{t-1} - \gamma_t \nabla D(\mathbf{x}^{t-1})), \quad (10)$$

where the gradient can be expressed as $\nabla D(\mathbf{x}) = \mathbf{H}^T(\mathbf{Hx} - \mathbf{y})$, the step size is taken as $\gamma_t = 1/L$ with L being the Lipschitz constant to ensure the convergence [31], and the proximal operator is defined as $p_L(z) \stackrel{d}{=} \underset{x \in \mathbb{R}^N}{\operatorname{argmin}} \left\{ \frac{1}{L} g(x) + \frac{1}{2} \|x - z\|_2^2 \right\}$. For a standard ISTA algorithm, the complexity of the convergence rate is $O(1/t)$ [30].

Compared with ISTA, the FISTA algorithm can significantly accelerate the speed of convergence [31]. The procedure of iteration is as follows:

$$\mathbf{x}^t \leftarrow p_L(\mathbf{u}^{t-1} - \gamma_t \nabla D(\mathbf{u}^{t-1})), \quad (11)$$

$$q_t \leftarrow (1 + \sqrt{1 + 4q_{t-1}^2})/2, \quad (12)$$

$$\mathbf{u}^t \leftarrow \mathbf{x}^t + ((q_{t-1} - 1)/q_t)(\mathbf{x}^t - \mathbf{x}^{t-1}), \quad (13)$$

where $\mathbf{u}^0 = \mathbf{x}^0$ and $q_0 = 1$. When the step size is invariant, the final convergence rate can reach $O(1/t^2)$ [31], which significantly speeds up the image reconstruction, especially for 3D images, so it is crucial for solving the linear inverse problem with a large amount of data.

In TV calculation, one should particularly note the nested iteration problem, which is likely to cause extremely slow reconstruction when dealing with large-scale imaging problems in 3D computational imaging. Therefore, the optimization for FISTA is essential. The fast-parallel proximal algorithm of the FISTA counterpart is [30]:

$$z^t \leftarrow \mathbf{u}^{t-1} - \gamma_t \mathbf{H}^T(\mathbf{H}\mathbf{u}^{t-1} - \mathbf{y}), \quad (14)$$

$$\mathbf{x}^t \leftarrow \frac{1}{K} \sum_{k=1}^K \mathbf{W}_k^T \Gamma(\mathbf{W}_k z^t; \sqrt{2}K\lambda\gamma_t), \quad (15)$$

$$q_t \leftarrow (1 + \sqrt{1 + 4q_{t-1}^2})/2, \quad (16)$$

$$\mathbf{u}^t \leftarrow \mathbf{x}^t + ((q_{t-1} - 1)/q_t)(\mathbf{x}^t - \mathbf{x}^{t-1}), \quad (17)$$

where $\mathbf{u}^0 = \mathbf{x}^0$, $q_0 = 1$, $\Gamma(\mathbf{y}; \varepsilon) = \max(|\mathbf{y}| - \varepsilon, 0) \frac{\mathbf{y}}{|\mathbf{y}|}$, $\gamma = 2\sqrt{2}D$, $K = 2D$, λ denotes the coefficient of regularization term and, as mentioned above, γ_t is the step size, D represents the dimensionality of the image. \mathbf{W}_k is expressed as an orthogonal wavelet decomposition of the image, and the speed of convergence of the fast parallel proximal algorithm can be proved to be still $O(1/t^2)$, but effectively help to solve the nested iteration problem [30].

2.4. TV-FISTA-SIM

We propose an iterative spatial reconstruction algorithm termed TV-FISTA-SIM, which significantly accelerates and realizes a high-quality 3D-SIM reconstruction. We regard the process of 3D-SIM reconstruction as a minimum optimization problem for the loss function with total variational regularization term:

$$F = f + g, \quad (18)$$

$$f = \left\| \sum_m (J_m(\mathbf{r}_{xy}) S_{est}(\mathbf{r})) \otimes H_m(\mathbf{r}) - D(\mathbf{r}) \right\|_2^2, \quad (19)$$

where g denotes the total variation regularization term in 3D form. We will apply the TV-FISTA-SIM algorithm to reconstruct 3D-SIM images, and the partial derivation of the loss function

concerning the estimation of the sample $S_{est}(\mathbf{r})$ is [29]:

$$\frac{\partial f}{\partial S_{est}} = \sum_{m'} \left(\left(\sum_m (J_m(\mathbf{r}_{xy})S_{est}(\mathbf{r})) \otimes H_m(\mathbf{r}) - D(\mathbf{r}) \right) \otimes H_{m'}(\mathbf{r}) \right) \cdot J_{m'}(\mathbf{r}_{xy}). \quad (20)$$

Equation (14) can be formulated as:

$$\mathbf{z}^t \leftarrow \mathbf{u}^{t-1} - \gamma_t \frac{\partial f}{\partial S_{est}}. \quad (21)$$

To ensure the convergence in the iterative process of the TV-FISTA-SIM algorithm, we take the step size as the reciprocal of the Lipschitz constant. We can solve the Lipschitz constant related to the preset microscopy system through the Power Iteration algorithm [28] combined with the 3D-SIM imaging model.

Algorithm 1: Lipschitz constant in 3D-SIM

Input: System parameters $H_m(\mathbf{r})$, $I_m(z)$, iteration number it_num

Output: Lipschitz constant $L_{\theta,\phi}$

```

1 Initialize:  $\mathbf{X}_{\theta,\phi}^0$  ;
2 for  $i = \theta_1, \theta_2, \theta_3$  do
3   for  $j = \phi_1, \phi_2, \phi_3, \phi_4, \phi_5$  do
4     for  $k = 1 : it_n$  do
5        $\mathbf{X}_{\theta,\phi}^0 \leftarrow \sum_m (J_m(\mathbf{r}_{xy}) \mathbf{X}_{\theta,\phi}^0 \otimes H_m(\mathbf{r}))$  ;
6     end
7      $\mathbf{X}_{\theta,\phi}^1 \leftarrow \sum_m (J_m(\mathbf{r}_{xy}) \mathbf{X}_{\theta,\phi}^0 \otimes H_m(\mathbf{r}))$  ;
8      $L_{\theta,\phi} \leftarrow \sum (\mathbf{X}_{\theta,\phi}^1 \cdot \mathbf{X}_{\theta,\phi}^0) / \sum (\mathbf{X}_{\theta,\phi}^0 \cdot \mathbf{X}_{\theta,\phi}^0)$  ;
9   end
10 end

```

The overall reconstruction algorithm of TV-FISTA-SIM can be represented as follows.

Algorithm 2: TV-FISTA-SIM

Input: System parameters $H_m(\mathbf{r})$, $I_m(z)$, system OTF $\tilde{H}(\mathbf{k})$, Lipschitz constant $L_{\theta,\phi}$,

regularization coefficient λ , iteration number it_n , raw SIM images $D_{\theta,\phi}(\mathbf{r})$

Output: Reconstructed 3D SIM image $D_{SIM}(\mathbf{r})$

```

1 Initialize:  $\mathbf{x}^0, \mathbf{u}^0 \leftarrow \sum_{\theta,\phi} D_{\theta,\phi}(\mathbf{r}), q_0 \leftarrow 1$ ;
2 for  $k = 1 : it_n$  do
3   for  $i = D_{\theta,\phi}(\mathbf{r}) - \theta_1, \theta_2, \theta_3$  do
4     for  $j = D_{\theta,\phi}(\mathbf{r}) - \phi_1, \phi_2, \phi_3, \phi_4, \phi_5$  do
5        $\mathbf{z}^t \leftarrow \mathbf{x}^{t-1} - \gamma_t \frac{\partial f}{\partial \mathbf{x}^{t-1}}$ ;
6        $\mathbf{u}^t \leftarrow \frac{1}{K} \sum_{k=1}^K \mathbf{W}_k^T \Gamma(\mathbf{W}_k \mathbf{z}^t; \sqrt{2K} \lambda \gamma_t)$ ;
7        $q_t \leftarrow (1 + \sqrt{1 + 4q_{t-1}^2}) / 2$ ;
8        $\mathbf{x}^t \leftarrow \mathbf{u}^t + ((q_{t-1} - 1) / q_t) (\mathbf{u}^t - \mathbf{u}^{t-1})$ ;
9     end
10   end
11 end
12 Result:  $D_{SIM}(\mathbf{r}) \leftarrow \mathbf{x}^t$ 

```

2.5. Principle for data reduction

The TV-FISTA-SIM can achieve high-quality reconstruction results with a small amount of original data but close to the results with complete data, which greatly improves the temporal resolution of the system. The data reduction strategy is by reducing the number of images with different phases in each illumination direction.

The low data dependence of TV-FISTA-SIM mainly comes from two aspects. First, with interference patterns, the high-frequency components of the object are down-modulated into the support domain of OTF. There contains no additional high-frequency information in images with different numbers of phase shifts in a single illumination direction. The conventional linear reconstruction algorithm obtains five high-frequency components with different orders by solving the linear equations and then stitches these components together through the Wiener Filter-based SIM reconstruction algorithm (Wiener-SIM) [3,19,20]. However, linear equations will be underdetermined if the number of acquired images is less than five. By solving underdetermined equations as an optimization problem and adding regularization constraints (to ensure the uniqueness of the solution), we can still obtain the corresponding high-frequency components with a low error, which is included in the iterations of the TV-FISTA-SIM algorithm. Second, the high-frequency information in different illumination directions has overlapping areas in the Fourier domain. Thus, data redundancy exists in these areas, and the actual number of constraints will be twice the number of phases (i.e., even if there are only three phases, these overlapping

areas will have six constraints, more than five, to make an overdetermined system of equations). As pointed out in [38,39] regarding the Fourier Ptychography Microscopy (FPM), the iterative results relate to the overlap rate between different apertures. The data redundancy in these overlapping areas provides constraints for the convergence of the iterative algorithm. Due to redundancy, reducing the amount of data with different interference phases is feasible.

3. Simulation results and discussion

Simulated data has been used to verify the validity of TV-FISTA-SIM, including sparse fluorescent beads and non-sparse biological samples. Several parameters of the microscopy system are preset as follows: excitation wavelength $\lambda_{ex} = 561$ nm, emission wavelength $\lambda_{em} = 590$ nm, numerical aperture NA = 1.49, the refractive index of the embedding medium n = 1.518, the angle between each side beam direction and the optical axis $\theta = 60^\circ$; The pixel size is set to 10 nm in the sample of sparse fluorescent beads and 60 nm in the continuous biological samples. Under the 3D-SIM imaging process, fifteen raw volumetric image stacks are simulated with different three-dimensional illumination patterns. For the best axial resolution for verification and comparison, the I^5S imaging model is adopted. The algorithm can also be applied to the general 3D-SIM model, which is much simpler than I^5S .

After reconstruction, the super-resolution volumetric results are normalized, and we utilize Structure Similarity Index Measure (SSIM), and root-mean-squared error (RMSE) to quantify and evaluate the 3D image reconstruction effect:

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}, \quad (22)$$

where m denotes the number of pixels in the volumetric sample, y_i refers to the value of the i-th pixel in the reconstructed image stack, and \hat{y}_i refers to the value of the i-th pixel in the ground truth.

3.1. Robustness of the TV-FISTA-SIM to noises

A primary motivation of this study is the potential capability of TV-FISTA-SIM for addressing low-SNR reconstruction problems in 3D-SIM. Traditional SIM reconstruction requires high-SNR images to ensure high-quality reconstruction results [19,27,28], which may easily cause photobleaching and phototoxicity and limits the further application of live-cell imaging.

To verify the robustness of the algorithm, we first apply TV-FISTA-SIM to sparse samples with different noise levels. The simulated sample volume consists of thousands of 10-nm fluorescent beads sparsely seeded at random 3D locations, and the product of the sample and the illumination patterns are convolved with the 3D PSF calculated with the preset microscope parameters. Then, the raw volumes are corrupted with Gaussian noise in different noisy conditions, ranging from SNR 5 dB to SNR 25 dB, to simulate raw 3D-SIM image volumes.

Under different SNR conditions, sparse fluorescent beads samples are reconstructed using TV-FISTA-SIM in 100 iterations. Typical reconstruction results of a single slice from sample volumes with TV-FISTA-SIM are displayed in Fig. 1(a)-(e) with corresponding wide-field images, and the statistics are shown in Fig. 2. The results show that TV-FISTA-SIM performs well laterally and axially, and higher SNR results in better reconstruction performance. Importantly, even when fed with poor data whose SNR is as low as 5 dB in Fig. 1(a), the reconstruction result of TV-FISTA-SIM is very similar to the high SNR conditions and contains no apparent artifacts, indicating the great potential of TV-FISTA-SIM in 3D-SIM reconstruction under low SNR conditions.

To more accurately verify the strength of TV-FISTA-SIM in high-noise situations, we compare the reconstruction results using TV-FISTA-SIM with Wiener-SIM and NGD-SIM. Under the

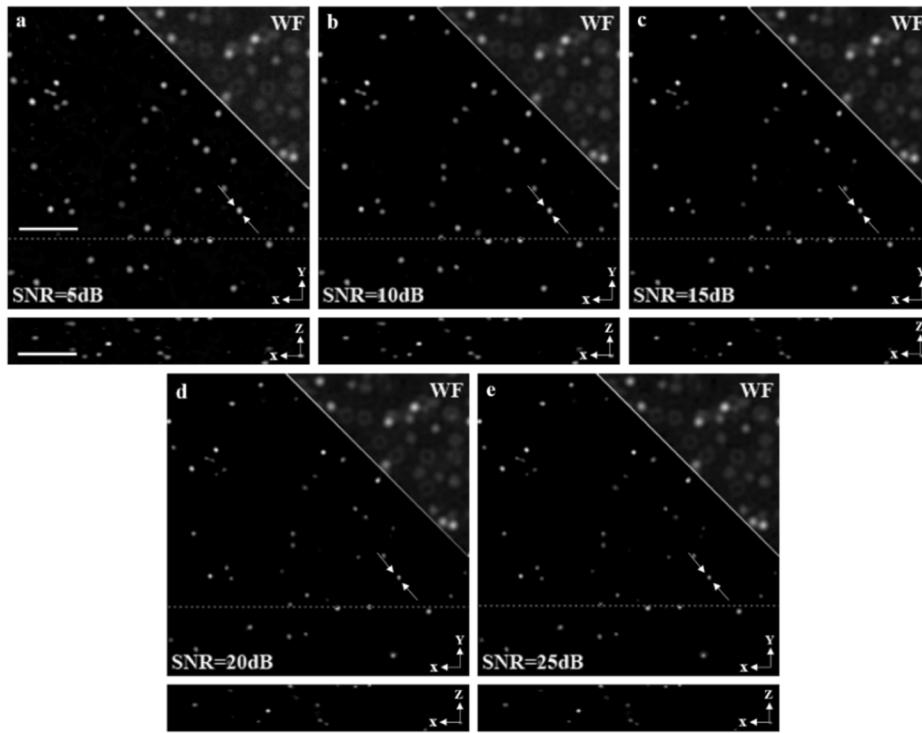


Fig. 1. Simulation results of fluorescent beads using TV-FISTA-SIM. (a)-(e) The wide-field (WF) images and TV-FISTA-SIM reconstructed images at different SNR. The lateral views (upper) show one single slice of the image volumes, and the axial views (lower) represent the single slice indicated by the white dashed line in (a)-(e). Scale bar: 500 nm.

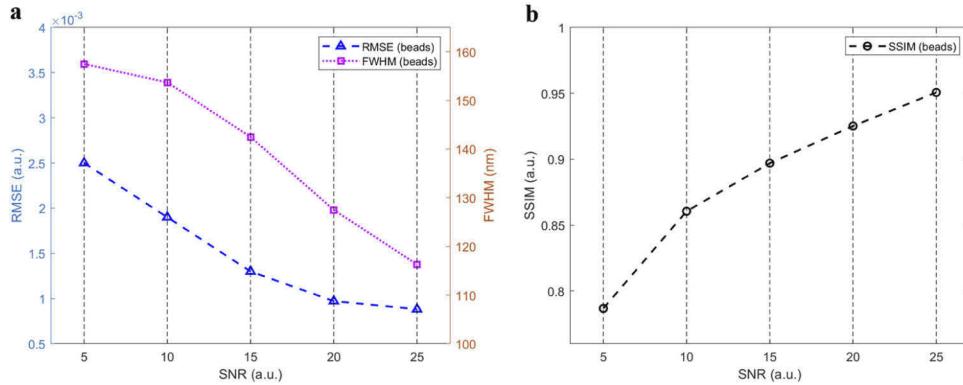


Fig. 2. (a) The statistical quantification of Root Mean Squared Error (RMSE, blue) and Full Width at Half Maximum (FWHM, magenta) and (b) Structure Similarity Index Measure (SSIM) indicated by white arrows of reconstructed 3D images v.s. SNR in Fig. 1(a)-(e).

condition of SNR 5 dB, the wide-field image, Wiener-SIM reconstruction result, TV-FISTA-SIM reconstruction result in 100 iterations, and NGD-SIM reconstruction result in 400 iterations are displayed in Fig. 3(a)-(d). The coefficients of total variation regularizer are set to zero in the reconstruction of TV-FISTA-SIM. The X-Y profiles along the white line and the Z profiles along the red line indicated in Fig. 3(a)-(d) are shown in Fig. 3(e) and Fig. 3(f), respectively. The

reconstruction using TV-FISTA-SIM has the highest fidelity reconstruction and contains the least artifacts. Furthermore, the axial resolution of wide-field, Wiener-SIM, NGD-SIM, and TV-FISTA-SIM can be quantified in FWHM as 215.1 nm, 139.1 nm, 122.9 nm, and 112.5 nm accordingly, which proves TV-FISTA-SIM delivers the most optimal reconstruction in extreme high-noise conditions.

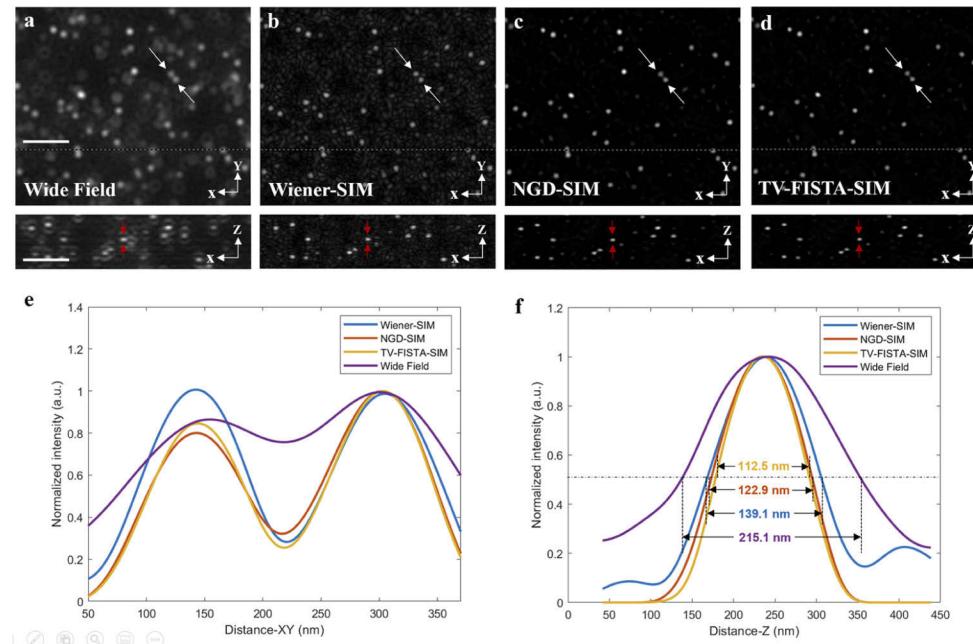


Fig. 3. Simulation results of fluorescent beads using different 3D-SIM reconstruction algorithms. (a)-(d) One single slice from raw wide-field image volume in SNR 5 dB (a), 3D-SIM image volume reconstructed with conventional Wiener-SIM (b), NGD-SIM (c), and TV-FISTA-SIM (d). Lateral (upper) and axial (lower) cross-sections are presented, respectively. (e) The X-Y intensity profile of the line indicated by white arrows in (a)-(d). (f) The Z intensity profile of the line indicated by red arrows in (a)-(d). Scale bar: 500 nm.

For non-sparse biological samples, TV-FISTA-SIM also shows great reconstruction abilities under high-noise imaging conditions. At a SNR 5 dB, Fig. 4(a) displays one single slice from a wide-field volumetric microtubules sample with excitation wavelength $\lambda_{ex} = 561$ nm, Fig. 4(b)-(d) show the same slice after reconstruction with different 3D-SIM algorithms. The Wiener-SIM reconstruction in Fig. 4(b) shows low resolution and artifacts under high-noise conditions, indicating a failed reconstruction. The NGD-SIM algorithm with 600 iterations in Fig. 4(c) can achieve high-resolution reconstruction in the case of high noise, but some artifacts and fractures still show its poor detail reconstruction.

In addition, the reconstruction results using TV-FISTA-SIM in Fig. 4(d) have the highest resolution, the least artifacts, and the best image continuity. To quantify the reconstruction performance of the three algorithms on microtubules, in Fig. 5, we plotted the RMSE and SSIM from the reconstruction of the volumetric microtubule samples with different levels of noise, where it can be clearly seen that the TV-FISTA-SIM has the lowest RMSE and the highest SSIM over all SNR conditions.

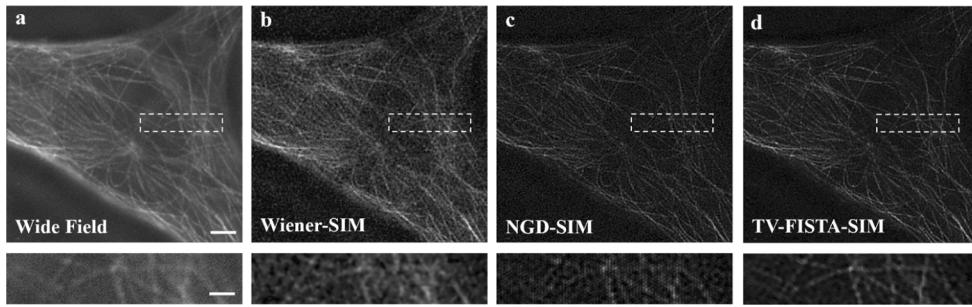


Fig. 4. Comparisons of microtubules between Wide Field (a), Wiener-SIM (b), NGD-SIM (c), and TV-FISTA-SIM (d) under SNR 5. Scale bar: 5 μm , 1.5 μm (zoom-in regions of Fig. 4(a)-(d)).

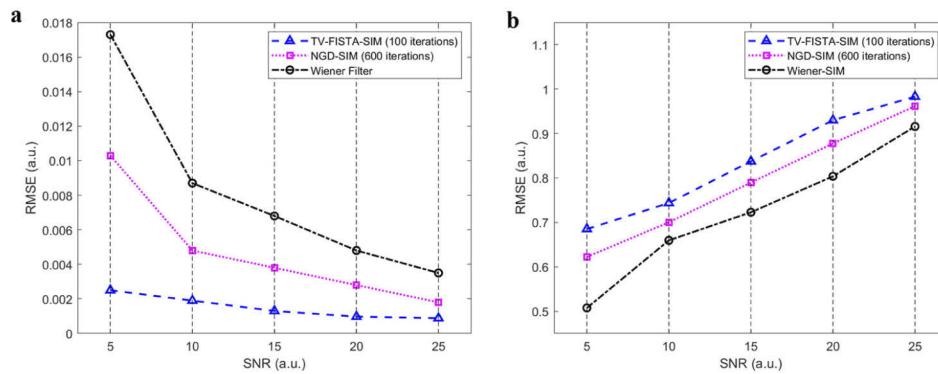


Fig. 5. (a) The root-mean-squared error (RMSE) and (b) Structure Similarity Index Measure (SSIM) of the reconstructed 3D microtubule images vs SNR by algorithms of TV-FISTA-SIM (100 iterations, blue), NGD-SIM (600 iterations, magenta), and Wiener-SIM (black).

3.2. High-quality reconstruction in reduced data

Apart from reconstructing high-quality images from low SNR raw data due to less exposure time, reducing the required data for acquisition and reconstruction is also essential. However, with less data, the reconstruction quality inevitably degrades. TV-FISTA-SIM shows better reconstruction than other algorithms, but it is vital to improving the effect further. The methods of regularization are often used to solve ill-posed inverse problems. In order to reduce the impact of the degradation caused by less data, we utilize the total variation regularizer in the optimization process, which significantly improves the reconstruction quality.

We apply the TV-FISTA-SIM algorithm with a total-variation coefficient greater than zero to reconstruct the 3D microtubule images with SNR as low as 5 dB and phase number as few as 1, the most extreme reconstruction condition. Figure 6(a)-(d) shows the reconstruction results under different total variation coefficients. With the addition of a total variation regularizer, a significant improvement can be seen in the image reconstruction. Furthermore, we use RMSE to quantitatively evaluate the performance of TV-FISTA-SIM in reduced data, in which numbers of phase reduce from five to one, with different total variation coefficients in Fig. 6(e). The overall RMSE curves in Fig. 6(e) display a downward trend within a specific range. Therefore, we can easily find the optimal total variation coefficient corresponding to the minimum RMSE, and at that point, the volumetric reconstruction effect will also present the optimal results. Furthermore, with different SNR conditions, the reconstruction quality could be improved by total variation

regularizer as well. With the combined evaluation of the RMSE and SSIM, similar enhancements for 3D reconstruction are shown in Fig. 6(f) and Fig. 6(g).

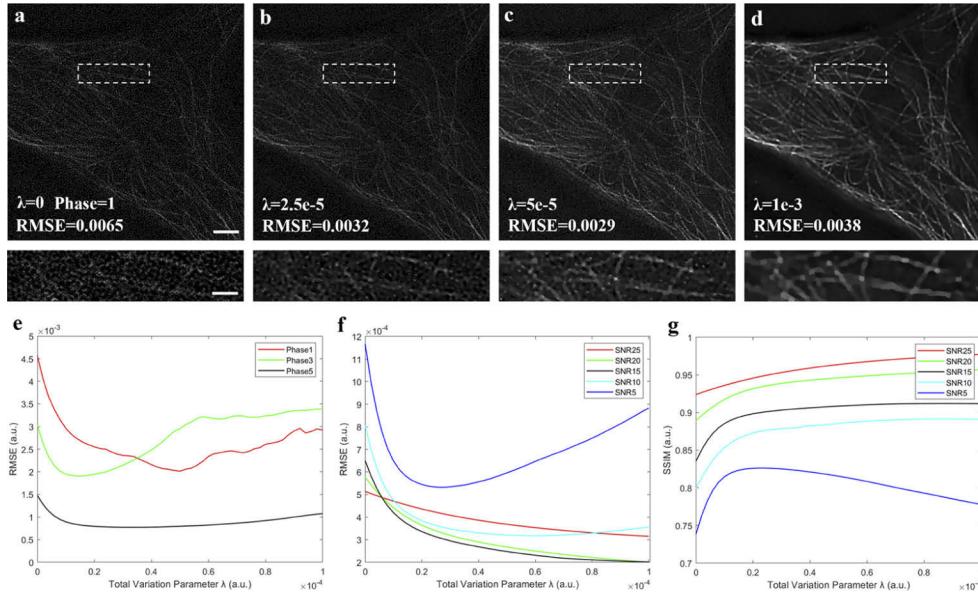


Fig. 6. (a)-(d) Comparisons of one single slice from reconstructed microtubules volumes under SNR 5 with different total variation parameters. The phase number in single direction is set to one, meaning total three images for reconstruction. Scale bar: 5 μm , 1.5 μm (zoom-in regions of Fig. 6(a)-(d)). (e) The RMSE under SNR 5 with different total variation parameters. Each curve represents different phase numbers used for reconstruction. (f) The RMSE with different total variation parameters. (g) The SSIM with different total variation parameters. Each curve in (f) and (g) represents different input SNR conditions with full data.

It should be noted that total variation can be understood as a constraint and an introduction of sparsity on images, in which the value of the total variation coefficient will play a key role. If the coefficient is too small, the constraint is insufficient, and the reconstruction may not be improved significantly (Fig. 6(b)). On the contrary, if the value is too large because the total variation helps maintain image continuity and smooth the image, it is likely to result in pixels connected (Fig. 6(d)), affecting the resolution and reconstruction. Additionally, the constraints of the regularizer depends on the actual condition and prior information, such as the types of samples, the noisy condition, the percentage of input sparsity, etc. Therefore, balancing the reconstruction and the actual imaging effects is necessary when adjusting total variation coefficients, which is like the parameter adjustment strategy in other SIM algorithms [19,29]. We generally recommend that users select appropriate parameters between 1e-6 and 1e-3 according to the actual situation.

3.3. Convergence rate of algorithms

In order to quantify the convergence rate and single iteration rate of TV-FISTA-SIM, a comparison with other similar gradient descent algorithms, NGD-SIM, and GD-SIM, is done. All algorithms were coded in MATLAB2020a and executed on the same device (Intel Xeon Gold 5120 CPU, 2.2GHz, NVIDIA Quadro P6000, RAM 128GB). The test sample is a simulated microtubules volume with the pixel size at 50 nm, which has also been chosen for simulated reconstruction in Ref. [29] and publicly released.

For the convergence rate, because TV-FISTA-SIM has the theoretical convergence rate $O(1/t^2)$, which shows a substantial advantage over NGD-SIM and GD-SIM. Figure 7 shows the convergence rate of the three algorithms evaluated by RMSE. In order to reach a RMSE of 1.65e-3, TV-FISTA-SIM only requires 259 iterations, while NGD-SIM needs 912 iterations and GD-SIM 2819 iterations. With the reduction of RMSE, the advantages of TV-FISTA-SIM will become more and more evident. The zoom-in region in Fig. 7 shows that the convergence rate of TV-FISTA-SIM before the first 50 iterations is slightly lower than NGD-SIM and GD-SIM, which is caused by the difference between the iteration processes but does not affect the conclusion. In terms of single iteration rate, TV-FISTA-SIM needs 2.34 s and NGD-SIM 1.51 s on average. A clearer comparison of reconstruction speed is shown in Table 1.

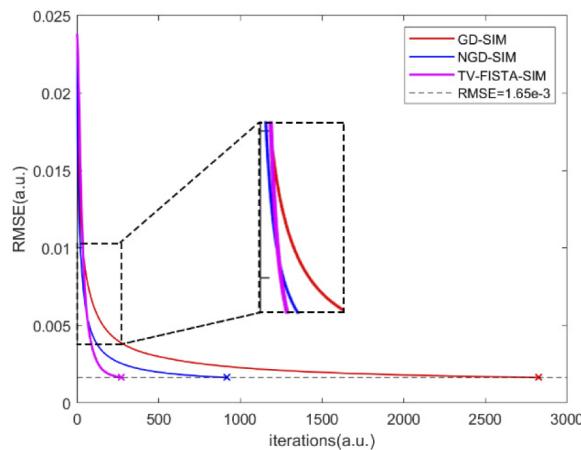


Fig. 7. Performances of reconstruction by GD-SIM, NGD-SIM, and TV-FISTA-SIM with different number of iterations. The performances of the reconstructed microtubules are measured by Root Mean Squared Error (RMSE).

Table 1. Reconstruction speed of different iterative algorithms for 3D-SIM in Fig. 7

	GD-SIM	NGD-SIM	TV-FISTA-SIM
Single Iteration Time (s)	1.31	1.51	2.34
Iteration Number	2819	912	259
Reconstruction Time (min)	61.5	23.0	10.1

3.4. Multi-color 3D imaging

Multi-color three-dimensional imaging is of great significance for observing biological structures and interaction mechanisms of cells. To verify the reconstruction effect of TV-FISTA-SIM on actual 3D samples with different excitation wavelengths and organelles, the 3D imaging samples have been tested in simulation. Figure 8 shows three reconstructed biological structures using Wiener-SIM, NGD-SIM and TV-FISTA-SIM and their wide-field images with the intensity-projection of the volumes, including mitochondrion (488 nm), microtubules (561 nm), and microfilaments (640 nm). As shown in Fig. 8, the performances of the NGD-SIM show a greater resolution and less background noise than Wiener-SIM but remain visible diagonal streak artifacts. Moreover, the statistical analysis in Tables 2 and 3 quantifies that the TV-FISTA-SIM typically realizes the best imaging performances in terms of all three wavelength illuminated samples with different structures, which demonstrates its reconstruction superiority. To better display the

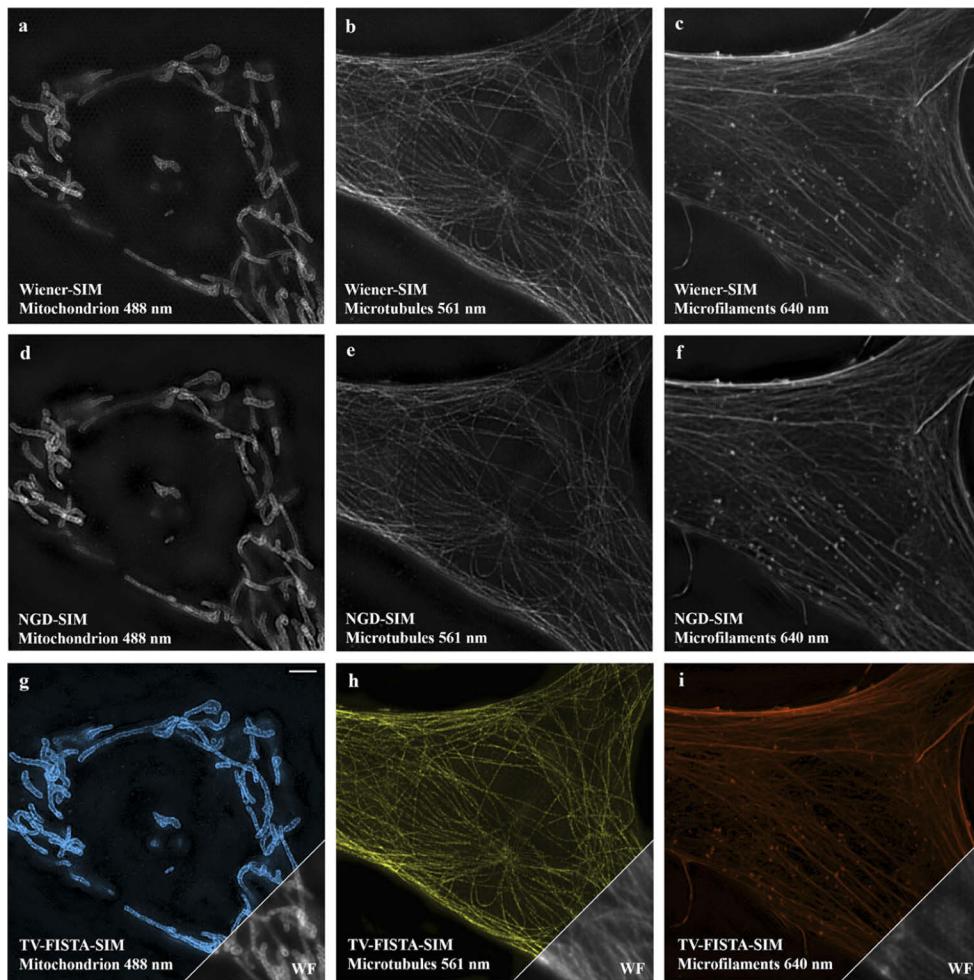


Fig. 8. Comparison of multi-color 3D-SIM reconstruction via Wiener-SIM, NGD-SIM and TV-FISTA-SIM with the intensity-projection of the volumes. (a-c) Reconstruction results of the Wiener-SIM, (d-f) NGD-SIM and (g-i) TV-FISTA-SIM of the Mitochondrion (488 nm, first column), Microtubules (561 nm, second column) and Microfilaments (640 nm, third column). Scale bar: 5 μ m.

Table 2. The RMSE of multi-color 3D-SIM reconstruction with different algorithms in Fig. 8

	Wiener-SIM	NGD-SIM	TV-FISTA-SIM
Mitochondrion (488 nm)	2.01e-3	1.80e-3	5.53e-4
Microtubules (561 nm)	3.32e-3	1.06e-3	2.47e-4
Microfilaments (640 nm)	2.43e-3	2.0e-3	4.24e-4

achievements of multi-color imaging with TV-FISTA-SIM, the Fig. 9 contains the reconstructed 3D sample (left) with the color-coded depth and transparency-characterized intensity information and the original wide-field image (right). We can clearly see that compared with the wide-field image; the 3D sample reconstructed by TV-FISTA-SIM has significantly improved resolution. This shows that TV-FISTA-SIM performs strongly in simulated multi-color 3D imaging as well, which will greatly expand the scope of application for the proposed algorithm in the future.

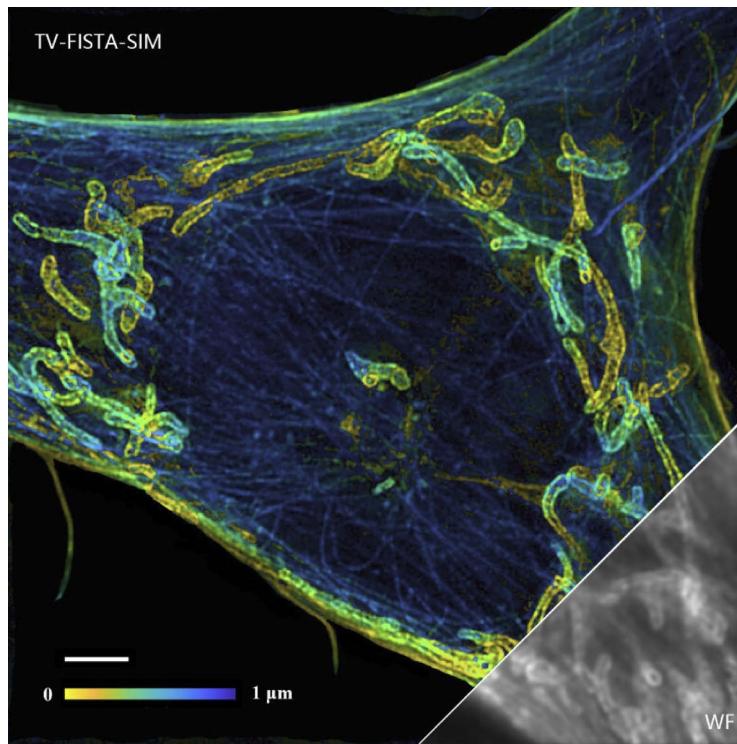


Fig. 9. Simulation of 3D imaging results of mitochondrion (488nm), microtubules (561nm), and microfilaments (640nm). Reconstructed 3D imaging results using TV-FISTA-SIM (left). The color encodes axial positions. Wide Field image with the intensity-projection of the volumes (right). Scale bar: 5 μm .

Table 3. The SSIM of multi-color 3D-SIM reconstruction with different algorithms in Fig. 8

	Wiener-SIM	NGD-SIM	TV-FISTA-SIM
Mitochondrion (488 nm)	0.8120	0.9288	0.9845
Microtubules (561 nm)	0.8413	0.9474	0.9887
Microfilaments (640 nm)	0.8693	0.9610	0.9842

4. Conclusion

We have proposed the TV-FISTA-SIM method as a novel 3D-SIM reconstruction algorithm and demonstrated it for high-quality volumetric reconstruction. Compared with the conventional 3D-SIM algorithms, TV-FISTA-SIM can achieve higher resolution with fewer artifacts under high-noise conditions. Moreover, TV-FISTA-SIM shows excellent potential to realize high-fidelity reconstruction with the reduction of original data, making it promising for real-time 3D-SIM imaging. TV-FISTA-SIM also has a faster convergence rate, which significantly improves the possibility of the algorithm into practical applications. More importantly, we verify that the TV-FISTA-SIM remains high-quality reconstruction in the case of 3D multi-color biological organelles. The proposed method can be applied to many illumination fringes, including sinusoidal and speckle patterns. In the future, we will explore the application of the TV-FISTA-SIM in the actual imaging scenes. We expect that our method can inspire further investigations and applications of 3D-SIM.

Funding. National Natural Science Foundation of China (61735017, 61827825, 62125504); Natural Science Foundation of Zhejiang Province (LD21F050002); Key Research and Development Program of Zhejiang Province (2020C01116); Zhejiang Lab (2020MC0AE01); Zhejiang Provincial Ten Thousand Plan for Young Top Talents (2020R52001); China Postdoctoral Science Foundation (2021M702798); Research Grants Council, University Grants Committee General Research Fund (14209521).

Acknowledgments. We would like to thank Xin Liu (Zhejiang University) for the support and kind assistant; Nansen Zhou (The Chinese University of Hong Kong) and Andreas Zheng (The Chinese University of Hong Kong) for the advise in organization of the manuscript.

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

Data availability. Data underlying the results presented in this paper are available in Ref. [40].

References

1. S. W. Hell and J. Wichmann, "Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy," *Opt. Lett.* **19**(11), 780–782 (1994).
2. M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**(2), 82–87 (2000).
3. M. G. L. Gustafsson, L. Shao, P. M. Carlton, C. J. R. Wang, I. N. Golubovskaya, W. Z. Cande, D. A. Agard, and J. W. Sedat, "Three-Dimensional Resolution Doubling in Wide-Field Fluorescence Microscopy by Structured Illumination," *Biophys. J.* **94**(12), 4957–4970 (2008).
4. F. Ströhl and C. F. Kamiński, "Frontiers in structured illumination microscopy," *Optica* **3**(6), 667 (2016).
5. E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science* **313**(5793), 1642–1645 (2006).
6. S. Manley, J. M. Gillette, G. H. Patterson, H. Shroff, H. F. Hess, E. Betzig, and J. Lippincott-Schwartz, "High-density mapping of single-molecule trajectories with photoactivated localization microscopy," *Nat. Methods* **5**(2), 155–157 (2008).
7. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods* **3**(10), 793–796 (2006).
8. B. Dudok, L. Barna, M. Ledri, S. I. Szabó, E. Szabadits, B. Pintér, S. G. Woodhams, C. M. Henstridge, G. Y. Balla, R. Nyilas, C. Varga, S.-H. Lee, M. Matolcsi, J. Cervenak, I. Kacsikovics, M. Watanabe, C. Sagheddu, M. Melis, M. Pistis, I. Soltesz, and I. Katona, "Cell-specific STORM super-resolution imaging reveals nanoscale organization of cannabinoid signaling," *Nat. Neurosci.* **18**(1), 75–86 (2015).
9. C. Morgan, J. A. Fozard, M. Hartley, I. R. Henderson, K. Bomblies, and M. Howard, "Diffusion-mediated HEI10 coarsening can explain meiotic crossover positioning in *Arabidopsis*," *Nat. Commun.* **12**(1), 4674 (2021).
10. C. Merigiano, R. Burla, M. L. Torre, S. D. Giudice, H. Teo, C. W. Liew, A. Chojnowski, W. I. Goh, Y. Olmos, K. Maccaroni, M. Giubettini, I. Chiolo, J. G. Carlton, D. Raimondo, F. Verni, C. L. Stewart, D. Rhodes, G. D. Wright, B. E. Burke, and I. Saggio, "AKTIP interacts with ESCRT I and is needed for the recruitment of ESCRT III subunits to the midbody," *PLoS Genet.* **17**(8), e1009757 (2021).
11. L. Xie and Z. Liu, "Single-cell imaging of genome organization and dynamics," *Mol. Syst. Biol.* **17**(7), e9653 (2021).
12. M. R. Arnison and C. J. R. Sheppard, "A 3D vectorial optical transfer function suitable for arbitrary pupil functions," *Opt. Commun.* **211**(1-6), 53–63 (2002).
13. S. Hell and E. H. K. Stelzer, "Properties of a 4Pi confocal fluorescence microscope," *J. Opt. Soc. Am. A* **9**(12), 2159 (1992).
14. M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, "3D widefield microscopy with two objective lenses: experimental verification of improved axial resolution," *Three-Dimensional Microscopy: Image Acquisition and Processing II* **2655**, 62–66 (1996).
15. M. G. Gustafsson, D. A. Agard, and J. W. Sedat, "I5M: 3D widefield light microscopy with better than 100 nm axial resolution," *J. Microsc. (Oxford, U. K.)* **195**(1), 10–16 (1999).
16. L. Shao, B. Isaac, S. Uzawa, D. A. Agard, J. W. Sedat, and M. G. L. Gustafsson, "I5S: Wide-Field Light Microscopy with 100-nm-Scale Resolution in Three Dimensions," *Biophys. J.* **94**(12), 4971–4983 (2008).
17. X. Huang, J. Fan, L. Li, H. Liu, R. Wu, Y. Wu, L. Wei, H. Mao, A. Lal, P. Xi, L. Tang, Y. Zhang, Y. Liu, S. Tan, and L. Chen, "Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy," *Nat. Biotechnol.* **36**(5), 451–459 (2018).
18. S. Tu, Q. Liu, X. Liu, W. Liu, Z. Zhang, T. Luo, C. Kuang, X. Liu, and X. Hao, "Fast reconstruction algorithm for structured illumination microscopy," *Opt. Lett.* **45**(6), 1567–1570 (2020).
19. G. Wen, S. Li, L. Wang, X. Chen, Z. Sun, Y. Liang, X. Jin, Y. Xing, Y. Jiu, Y. Tang, and H. Li, "High-fidelity structured illumination microscopy by point-spread-function engineering," *Light: Sci. Appl.* **10**(1), 70 (2021).
20. A. Lal, C. Shan, and P. Xi, "Structured illumination microscopy image reconstruction algorithm," *IEEE J. Select. Topics Quantum Electron.* **22**(4), 50–63 (2016).
21. S. J. Sahl, F. Balzarotti, J. Keller-Findeisen, M. Leutenegger, V. Westphal, A. Egner, F. Lavoie-Cardinal, A. Chmyrov, T. Grotjohann, and S. Jakobs, "Comment on "Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics"," *Science* **352**(6285), 527 (2016).

22. L. H. Schaefer, D. Schuster, and J. Schaffer, "Structured illumination microscopy: artefact analysis and reduction utilizing a parameter optimization approach," *J. Microsc. (Oxford, U. K.)* **216**(2), 165–174 (2004).
23. J. Pospíšil, K. Fliegel, and M. Klíma, "Analysis of image reconstruction artifacts in structured illumination microscopy," in *Applications of Digital Image Processing XL* (SPIE, 2017), Vol. 10396, pp. 627–638.
24. A. Lal, X. Huang, and P. Xi, "A frequency domain reconstruction of SIM image using four raw images," in *Optics in Health Care and Biomedical Optics VII* (SPIE, 2016), Vol. 10024, pp. 104–116.
25. Y. Zhang, X. Chen, B. Li, S. Jiang, T. Zhang, X. Zhang, X. Yuan, P. Qin, G. Zheng, and X. Ji, "Accelerated Phase Shifting for Structured Illumination Microscopy based on Deep Learning," *IEEE Trans. Comput. Imaging* **7**, 1–12 (2021).
26. L. Jin, B. Liu, F. Zhao, S. Hahn, B. Dong, R. Song, T. C. Elston, Y. Xu, and K. M. Hahn, "Deep learning enables structured illumination microscopy with low light levels and enhanced speed," *Nat. Commun.* **11**(1), 1934 (2020).
27. D. Li, Q. Dai, C. Qiao, X. Chen, S. Zhang, D. Li, and Y. Guo, "3D Structured Illumination Microscopy via Channel Attention Generative Adversarial Network," *IEEE J. Select. Topics Quantum Electron.* **27**(4), 1 (2021).
28. C. Qiao, D. Li, Y. Guo, C. Liu, T. Jiang, Q. Dai, and D. Li, "Evaluation and development of deep neural networks for image super-resolution in optical microscopy," *Nat. Methods* **18**(2), 194–202 (2021).
29. H. Zhu, H. Zhu, Y. Sun, Y. Sun, L. Yin, J. Han, M. Cai, Q. Liu, C. Kuang, C. Kuang, C. Kuang, C. Kuang, X. Hao, X. Liu, and X. Liu, "3D super-resolution microscopy based on nonlinear gradient descent structured illumination," *Opt. Express* **29**(14), 21428–21443 (2021).
30. U. S. Kamilov, "A Parallel Proximal Algorithm for Anisotropic Total Variation Minimization," *IEEE Trans. on Image Process.* **26**(2), 539–548 (2017).
31. A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *SIAM J. Imaging Sci.* **2**(1), 183–202 (2009).
32. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D* **60**(1-4), 259–268 (1992).
33. J. P. Oliveira, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Adaptive total variation image deblurring: A majorization-minimization approach," *Signal Processing* **89**(9), 1683–1693 (2009).
34. E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory* **52**(2), 489–509 (2006).
35. M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, "Fast Image Recovery Using Variable Splitting and Constrained Optimization," *IEEE Trans. on Image Process* **19**(9), 2345–2356 (2010).
36. I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," arXiv:math/0307152 (2003).
37. J. M. Bioucas-Dias and M. A. T. Figueiredo, "A New TwIST: Two-Step Iterative Shrinkage/Thresholding Algorithms for Image Restoration," *IEEE Trans. on Image Process* **16**(12), 2992–3004 (2007).
38. G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nat. Photonics* **7**(9), 739–745 (2013).
39. K. Guo, S. Dong, P. Nanda, and G. Zheng, "Optimization of sampling pattern and the design of Fourier ptychographic illuminator," *Opt. Express* **23**(5), 6171–6180 (2015).
40. M. Cai, H. Zhu, Y. Sun, L. Yin, F. Xu, H. Wu, X. Hao, R. Zhou, C. Kuang, and X. Liu, "TV-FISTA-SIM, The TV-FISTA-SIM project," Github, 2022, <https://github.com/zjucmx/TV-FISTA-SIM>