



3D super-resolution microscopy based on nonlinear gradient descent structured illumination

HONGFEI ZHU,^{1,4} YILE SUN,^{1,4} LU YIN,¹ JIAXIAO HAN,¹ MINGXUAN CAI,¹ QIULAN LIU,² CUIFANG KUANG,^{1,2,3,5} XIANG HAO,¹ AND XU LIU^{1,2}

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

²*Research Center for Intelligent Sensing, Zhejiang Lab, Hangzhou 311100, China*

³*Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, 030006, China*

⁴*These authors contributed equally to this work*

⁵*cfkuang@zju.edu.cn*

Abstract: Three-dimensional structured illumination microscopy (3D-SIM) is an essential tool for volumetric fluorescence imaging, which improves both axial and lateral resolution by down-modulating high-frequency information of the sample into the passband of optical transfer function (OTF). And when combining with the 4Pi structure, the performance of 3D-SIM can be further improved. The reconstruction results of generally used linear 3D algorithm, however, are lack of high-fidelity and prone to generate artifacts. In this paper, we proposed a novel iterative algorithm based on gradient descent combined with a nonlinear optimizer, which can be applied to all 3D-SIM setups (including I⁵S setup). We verified through simulation that the proposed solution, termed as nonlinear gradient descent structured illumination microscopy (NGD-SIM), achieves more fidelity results which can reach the limitation of theoretical resolution improvement of SIM. Moreover, it can be firmly validated on simulation that this algorithm can effectively reduce the amount of raw data in the case of sinusoidal-pattern illumination, i.e., the algorithm doesn't need five-step phase shifting; data with any number of phases can theoretically be reconstructed. Our method also provides the possibility to extend the application of sinusoidal-pattern illumination to any kind of interference fringe, which is generated by diversified types of illumination mode.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Since it was proposed and invented, optical fluorescence microscopy has become an irreplaceable method for biomedical researches. However, because of the diffraction limit, the resolution of conventional microscopes is constrained to around 200 nm, which is not sufficient for developing research requirements. In order to go beyond the diffraction limit, diversified solutions have been developed in the last two decades. These methods, such as photoactivated localization microscopy (PALM) [1,2] and stochastic optical reconstruction microscopy (STORM) [3,4], rely on statistical computations and repetitive shootings to surpass the diffraction limit. And stimulated emission depletion microscopy (STED) [5] is based on scanning confocal microscopy to “reshape” the point spread function (PSF) of the system. Compared with what is mentioned above, structured illumination microscopy (SIM) [6–8] plays an essential role in live-cell imaging. STORM and PALM are supposed to use thousands of images to localize each emitter, which makes it difficult to observe living cells, owing to long-time imaging procedure and drift. Furthermore, the light intensity required in STED is much higher than SIM, and it may cause severe photobleaching and

phototoxicity. In contrast, SIM takes much less time in the imaging step and uses significantly lower light to obtain images.

Due to the shape of Optical Transfer Function (OTF), which is confined by Numerical Aperture (NA) [9], in the spatial frequency domain, the lateral resolution is several times better than the axial resolution. Considering the limit of light collecting angle, which determines NA for a single objective lens, a second objective lens is used and placed in a symmetrical position with the first objective lens about the focal plane. By using two objective lenses, the light collecting angle is doubled, and the new OTF can be stretched in axial direction correspondingly. The method that uses two identical objective lenses was applied in fluorescence microscopy such as 4Pi-B confocal microscopy [10], I^2M [11] and I^5M [12]. For SIM, by combining I^5M and structured illumination, which is termed I^5S [13], the axial resolution can be improved more than twice; that is, the OTF will be further stretched in SIM.

In order to reconstruct super-resolution images from SIM, numerous algorithms have been reported [6,14–17]. But nearly most of them aim at solving two-dimensional problems. In three-dimensional imaging, the most commonly used algorithm is the Wiener filtering strategy (later, we will use Wiener-Filter SIM to refer to this linear 3D-SIM reconstruction algorithm) [13,18], which is not accurate enough and usually produces artifacts. In this manuscript, we report a novel algorithm called Nonlinear Gradient Descent SIM (NGD-SIM) by utilizing a gradient descent mechanism with a nonlinear optimizer, which is widely used in deep learning and many other applications. A similar gradient descent method has been applied in two-dimensional Fourier ptychographic microscopy called PEFP to estimate the information of Interference fringes [19]. Here we expand this idea to 3D-SIM to calculate 3D object iteratively. The nonlinear gradient descent algorithm will accelerate the convergence of the iterative process. The new method demonstrates that it can perform ideally in both sparse and continuous samples according to simulations. In the meantime, compared with former techniques, the resolution and quality of reconstructed images obtained by NGD-SIM can be improved to a better degree both laterally and axially. According to simulations on MATLAB, the best lateral and axial resolution can reach 102.5nm and 62.5nm respectively (comparing with Wiener-Filter SIM, which are 140nm and 85nm under the same empirical condition). Our algorithm, to some extent, can greatly reduce the amount of the raw image data used in reconstruction and improve the temporal resolution consequently.

2. Principle

2.1. Forward model of the imaging process

Three-dimensional structured illumination microscopy is a linear, translation-invariant (LTI) optical system, whose response can be described by Transfer Function (in optical system, it is named PSF) $H(r)$ and convolution theory. The observed diffraction-limited data $D(r)$ is the convolution of the pointwise product of the illumination pattern $I(r)$ and the emitting object $S(r)$ with the PSF.

$$D(r) = [I(r) \cdot S(r)] * H(r) \quad (1)$$

Where r denotes real-space coordinate vector and $*$ refers to convolution.

Owing to the moiré fringe effect, the illumination pattern $I(r)$ in Eq. (1) down-modulates the original undetectable high-frequency components of the object into the passband of the detection OTF support [20, 21]. And these high-frequency components can be straightforwardly extracted if the illumination pattern satisfies the following conditions [13, 18]: First, the illumination pattern can be decomposed into a sum of a finite number of components, each of which is separated into a production of an axial and a lateral function. The specific expression is as

follows:

$$I(\mathbf{r}) = \sum_m I_m(z) J_m(\mathbf{r}_{xy}) \quad (2)$$

where $m = 0, \pm 1, \pm 2$; $J_m(\mathbf{r}_{xy})$ and $I_m(z)$ denotes lateral and axial illumination intensity in the real-space domain respectively. Second, each lateral function should be a simple harmonic function (i.e., $J_m(\mathbf{r}_{xy})$ can be expressed as $e^{im(2\pi p_{xy} + \varphi)}$, which implied that $\tilde{J}_m(\mathbf{k}_{xy}) = \delta(\mathbf{k}_{xy} - m\mathbf{p})e^{im\varphi}$. Third, relative to the focal plane of the microscope, the illumination patterns are supposed to remain fixed when acquiring a volumetric data set by moving the region of interest of the sample to the focal plane along the z -axis. Under the above conditions, combining Eq. (1) and Eq. (2), The observed diffraction-limited data $D(\mathbf{r})$ can be rewritten as Eq. (3),

$$\begin{aligned} D(\mathbf{r}) &= \sum_m J_m(\mathbf{r}_{xy}) S(\mathbf{r}) * (I_m(z) H(\mathbf{r})) \\ &= \sum_m (J_m(\mathbf{r}_{xy}) S(\mathbf{r})) * H_m(\mathbf{r}) \end{aligned} \quad (3)$$

In Eq. (3), the specific form of $J_m(\mathbf{r}_{xy})$ has been given above, and here we give the form of $I_m(z)$. It needs to be clearly stated that $J_m(\mathbf{r}_{xy})$ has the same expressions for 3D-SIM and I^5S , while the expressions of $I_m(z)$ are slightly different. The expressions of $I_m(z)$ for 3D-SIM are shown in Eq. (4) and I^5S in Eq. (5),

$$\left\{ \begin{array}{l} I_0(z) = 3 \\ I_{\pm 1}(z) = e^{i2\pi p_{z1}z} + e^{-i2\pi p_{z1}z} \\ I_{\pm 2}(z) = 1 \end{array} \right\} \quad (4)$$

$$\left\{ \begin{array}{l} I_0(z) = 6 + 2e^{i2\pi(2p_{z2}-2p_{z1})z} + 2e^{-i2\pi(2p_{z2}-2p_{z1})z} + e^{i2\pi 2p_{z2}z} + e^{-i2\pi 2p_{z2}z} \\ I_{\pm 1}(z) = 2e^{i2\pi 2p_{z1}z} + 2e^{-i2\pi 2p_{z1}z} + 2e^{i2\pi(2p_{z2}-p_{z1})z} + 2e^{-i2\pi(2p_{z2}-p_{z1})z} \\ I_{\pm 2}(z) = 2 + e^{i2\pi(2p_{z2}-2p_{z1})z} + e^{-i2\pi(2p_{z2}-2p_{z1})z} \end{array} \right\} \quad (5)$$

where p_{z1} equals to $\frac{1}{\lambda_{ex}}(1 - \cos \theta)$ and p_{z2} equals to $\frac{1}{\lambda_{ex}}$. θ denotes the angle between the center beam and the side beam. λ_{ex} denotes wavelength of excitation light.

The Fourier transform of Eq. (3) takes the form:

$$\tilde{D}(\mathbf{k}) = \sum_m e^{im\varphi} \tilde{S}(\mathbf{k} - m\mathbf{p}) \cdot \tilde{O}_m(\mathbf{k}) = \sum_m e^{im\varphi} \tilde{D}_m(\mathbf{k}) \quad (6)$$

where \mathbf{k} denotes the coordinate vector in reciprocal space

Equation (6) reveals that the acquired volumetric data is a mixture of the different frequency bands of the sample, one for each index m . If these different frequency bands can be moved to the correct position, it is equivalent to artificially expanding the support of the original OTF (synthetic OTF). However, the range of the synthetic OTF support in the Fourier domain depends on the distribution of frequency shift points formed by $\tilde{I}(\mathbf{k})$ ($\tilde{I}(\mathbf{k})$ is the Fourier transform of the illumination pattern $I(\mathbf{r})$, which is expressed as the accumulation of a series of δ functions, and each δ function represents a frequency shift point). For the general 3D-SIM (three-beam illumination, single objective lens reception) and I^5S (symmetrically two-sided three-beam illumination, double objective lens reception), the illumination pattern $I(\mathbf{r})$ is different, specifically expressed as different expressions of $I_m(z)$ (refer to Eq. (4) and Eq. (5)). Therefore, I^5S will produce more frequency shift points in the Fourier domain than the 3D-SIM. At the same time, due to the dual objective lens reception, the detection OTF support of the I^5S will be extended in the k_z direction than the original OTF support of 3D-SIM. The above two effects combine to make the axial resolution of I^5S higher than that of 3D-SIM. The corresponding original OTF, synthesized OTF, and Fourier domain frequency shift points distribution comparison diagrams of 3D-SIM and I^5S are shown in Fig. 1.

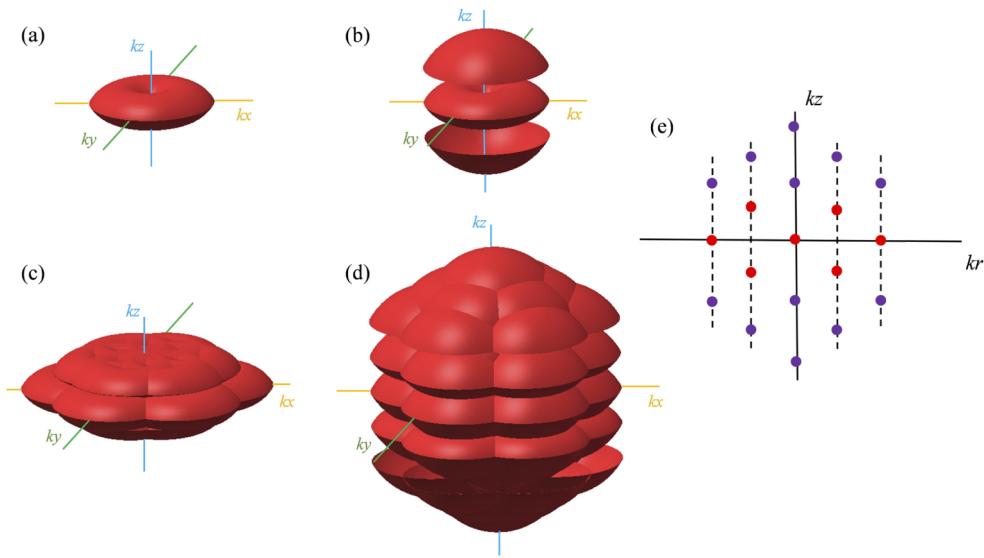


Fig. 1. OTF support by 3D-SIM and I^5S in 3D rendering. (a) The OTF support of conventional optical microscope. (b) The OTF support of I^5S , compared with OTF using a single objective lens, is stretched in the k_z axis. (c) Effective OTF support is produced by three-beam illumination in ordinary 3D-SIM. (d) The effective OTF support by illumination with six beams and detection through two opposing objective lenses. (e) The frequency shifts under the different illumination modes. The red dots denote extended frequency shift when using interference of three beams; The purple dots represent the extra frequency shifts for six beams by applying I^5S , which will never be present in ordinary 3D-SIM (Three-beam interference).

2.2. Sample reconstruction of Wiener-SIM

The conventional sample reconstruction method firstly separates these information components by acquiring additional images with varying illumination pattern phases $m\varphi$ to construct a set of five independent linear equations from which $\tilde{D}_m(\mathbf{k})$ can be extracted. Then $\tilde{D}_m(\mathbf{k})$ is moved back to their correct position, that is, $\pm mp$ and a linear and normalized Wiener filtering algorithm is utilized to stitch these components into a high-resolution image containing high-frequency information.

The Wiener-filter SIM stitched these components using Wiener filtering strategy and it can be realized by the following formula if $\tilde{D}_m(\mathbf{k})$ has been solved.

$$\tilde{S}(\mathbf{k}) = \frac{\sum_m \tilde{O}_m^*(\mathbf{k} + mp)\tilde{D}_m(\mathbf{k} + mp)}{\sum_m |\tilde{O}_{m'}(\mathbf{k} + mp)|^2 + \omega^2} A(\mathbf{k}) \quad (7)$$

Where $\tilde{S}(\mathbf{k})$ is the estimate of the true sample information and ω^2 is the Wiener parameter. $A(\mathbf{k})$ is an apodization function which is used to reduce the frequency components outside the synthesized OTF support to zero.

2.3. Sample reconstruction of NGD-SIM

Here we propose an iterative algorithm combined with a nonlinear optimizer based on gradient descent that does not require solving linear equations and can obtain a higher fidelity of reconstructed image while greatly reducing the volume of the raw data.

We regard the sample reconstruction process as a minimization problem and construct following loss function,

$$f = \left\| \sum_m (J_m(\mathbf{r}_{xy}) S_{est}(\mathbf{r})) * H_m(\mathbf{r}) - P(\mathbf{r}) \right\|^2 \quad (8)$$

where $S_{est}(\mathbf{r})$ denotes the estimation of the sample and $P(\mathbf{r})$ denotes the captured volumetric data; $\|\cdot\|^2$ is the square of the L2 norm.

The partial derivation of the loss function with respect to $S_{est}(\mathbf{r})$ is

$$\frac{\delta f}{\delta S_{est}} = \sum_{m'} \left[\sum_m (J_m(\mathbf{r}_{xy}) S_{est}(\mathbf{r}) * H_m(\mathbf{r}) - P(\mathbf{r})) * H_{m'}(\mathbf{r}) \right] \cdot J_{m'}(\mathbf{r}_{xy}) \quad (9)$$

According to Eq. (9), a conventional gradient descent algorithm can be used to update $S_{est}(\mathbf{r})$,

$$S_{est}^{\sim update}(\mathbf{r}) = S_{est}(\mathbf{r}) - k \frac{\delta f}{\delta S_{est}} \quad (10)$$

where k denotes the learning rate of which we recommend take the value as 1/3 when all physical parameters participating in the iteration have been normalized. In fact, the value of 1/3 is that we take the reciprocal of the Lipschitz constant L of the system (i.e. $1/L$) as the step size of the gradient descent. This choice is a conservative choice that must be able to converge. The explanation of Lipschitz constant and how to calculate Lipschitz constant are given in Appendix A.

To speed up the convergence rate, we learn from the strategy of machine learning and introduce Root Mean Square Propagation (RMSProp) [22] nonlinear optimizer into the gradient descent process. Then Eq. (7) will evolve into the following form,

$$V^{\sim update} = \beta \left(\frac{\delta f}{\delta S_{est}} \right)^2 + (1 - \beta)V \quad (11)$$

$$S_{est}^{\sim update}(\mathbf{r}) = S_{est}(\mathbf{r}) - \frac{\frac{\delta f}{\delta S_{est}}}{(V^{\sim update})^\alpha + \varepsilon} \quad (12)$$

where we recommend a value of 0.99 for parameter β and 0.1 for parameter α . ε is set to 10^{-7} by default and the initial value of V is set to zero. In fact, Eq. (11) calculates the cumulative sum of decayed gradient squared and β represents the decay rate. Equation (12) takes the reciprocal of the sum of decayed gradient squared as the final equivalent gradient which reduce the differences of update rate between different features, and uses an exponential factor α to control the overall update step size. Regarding the choice of parameters here, we refer to the choices in the original literature for β and ε , and we found that the α value (0.5) in the original literature does not make the algorithm converge. We enumerated the α parameters and chose 0.1 as the recommended value for α , which shows the fastest convergence rate. The effect of specific α parameters on the convergence speed and reconstruction results can be found in Appendix B. Refer to Appendix C for the comparison of convergence speed and reconstruction results between NGD-SIM and gradient descent SIM (GD-SIM) without a nonlinear optimizer.

Owing to the fact that the intensity of the sample is non-negative, a correction needs to be added at the end of each iteration.

$$S_{est}^{\sim update}(S_{est}^{\sim update} < 0) = 0 \quad (13)$$

The above reconstruction process indicates that the proposed algorithm does not require five-step phase shifting to acquire different frequency bands of the sample $\tilde{D}_m(\mathbf{k})$. In the simulation, the observed diffraction-limited data with only one phase is available to acquire better results than using Wiener Filter while more phase redundancy can improve the accuracy and noise tolerance of the reconstruction results.

3. Results and analysis of simulation

In order to validate the performance of our algorithm, several simulations have been done under the conditions of both sparse and continuous samples. In order to fit the practical experiments, relevant parameters are preset as close to the actual situations as possible: excitation wavelength $\lambda_{\text{ex}} = 640 \text{ nm}$, emission wavelength $\lambda_{\text{em}} = 670 \text{ nm}$, numerical aperture NA = 1.49, refractive index of the embedding medium n = 1.518, the aperture angle $\theta = 60^\circ$. The amount of phase shift each time $\varphi = \frac{2}{m}\pi$, where m is the total number of phase shifting. These parameters can define the wave number k and OTF in the frequency domain correspondingly. The pixel size in the simulated continuous samples is set as 50 nm which is compatible with the pixel size of EMCCD and in the sparse samples is set as 10 nm in order to get a better quantitative resolution for comparison. To comparing with ordinary Wiener Filter, the simulated pattern will have three orientations (0°, 60°, 120°) and five-phase shifts to fit the requirements of Wiener Filter. Hence up to 15 raw images are recorded for each 3D sample. Furthermore, we also simulated the situation when the number of phase shifting is from one to five to prove that NGD-SIM can effectively reduce the amount of raw data. To obtain best axial resolution, all the simulated data are made from I5S imaging model which is more complex than general 3D-SIM.

We first apply NGD-SIM (using different number of phases) and Wiener Filter to 87 simulated sparse fluorescent 10nm-diameter beads to demonstrate the reconstruction ability. The results are shown in Fig. 2. The lateral slices (Fig. 2(a2)-(f2)) and profiles (Fig. 2(g1), (g2)) illustrate a significant improvement in lateral resolution by using NGD-SIM. With increasing the number of phases used, the fidelity of NGD-SIM performs better. Comparing with the full width at half maximum (FWHM) of Wiener Filter which is 140 nm, the FWHM of NGD-SIM reaches 125 nm, 123 nm, 122.5 nm, 120 nm and 102.5 nm when applying 1 to 5 phases respectively; In addition, the axial slices (Fig. 2(a3)-(f3)) and profiles (Fig. 2(h1), (h2)) demonstrate that striking improvement in resolution can achieve in both axial and lateral direction. In traditional Wiener Filter, the FWHM of axial profile is around 85 nm, nonetheless, by using different number of phases, it can reach 70 nm, 64 nm, 67.5 nm, 62.5 nm and 62.5 nm respectively. It is also worth finding that with increasing the number of phases, the out-of-focus information can be suppressed, and when using 15 images these blurred patterns are totally removed, which denotes the axial resolution improvement. In Fig. 2(h2), When using 1 phase (3 images), the position of profile peak is deviated slightly (as depicted in the sky-blue curve). This is because when the number of phase shifts is too small, the process of iteration will produce errors in the calculation of the position of a single point. Therefore, though the 1-phase reconstruction result of NGD-SIM outperforms the result when using Wiener Filter, it is best to use at least 3 phases (9 images) in the image reconstruction procedure in the case of comprehensive consideration of resolution and data volume.

For continuous and complex 3D objects, the wide-field simulated microtubules are restored by both two techniques (Fig. 3). The wide field image shows severe out-of-focus blurring which obscures the sample (Fig. 3(a1)). By using Wiener Filter (Fig. 3(c1)-(c3)), the blurring can be removed effectively, but there are still many structures that remain unresolved. In contrast to Wiener Filter, NGD-SIM performs much better in both axial and lateral direction especially when applying 3 or more phases in this iterative algorithm (Fig. 3(d1)-(d3), (e1)-(e3)): the out-of-focus information and artifacts are perfectly removed and the thickness of microtubules is much slenderer (as shown in the yellow outlined regions). NGD-SIM outperforms Wiener Filter in the position where fluorescence is densely distributed. Our solution also shows a distinguishable function when using only 1 phase (3 images). Though it does recover a few details which are even better than using Wiener Filter, the continuity of the whole image is still destructed (Fig. 3(f3)). Therefore, after considering the tradeoff between the amount of data in the algorithm and the effect of image reconstruction, we recommend that at least 3 phases of raw

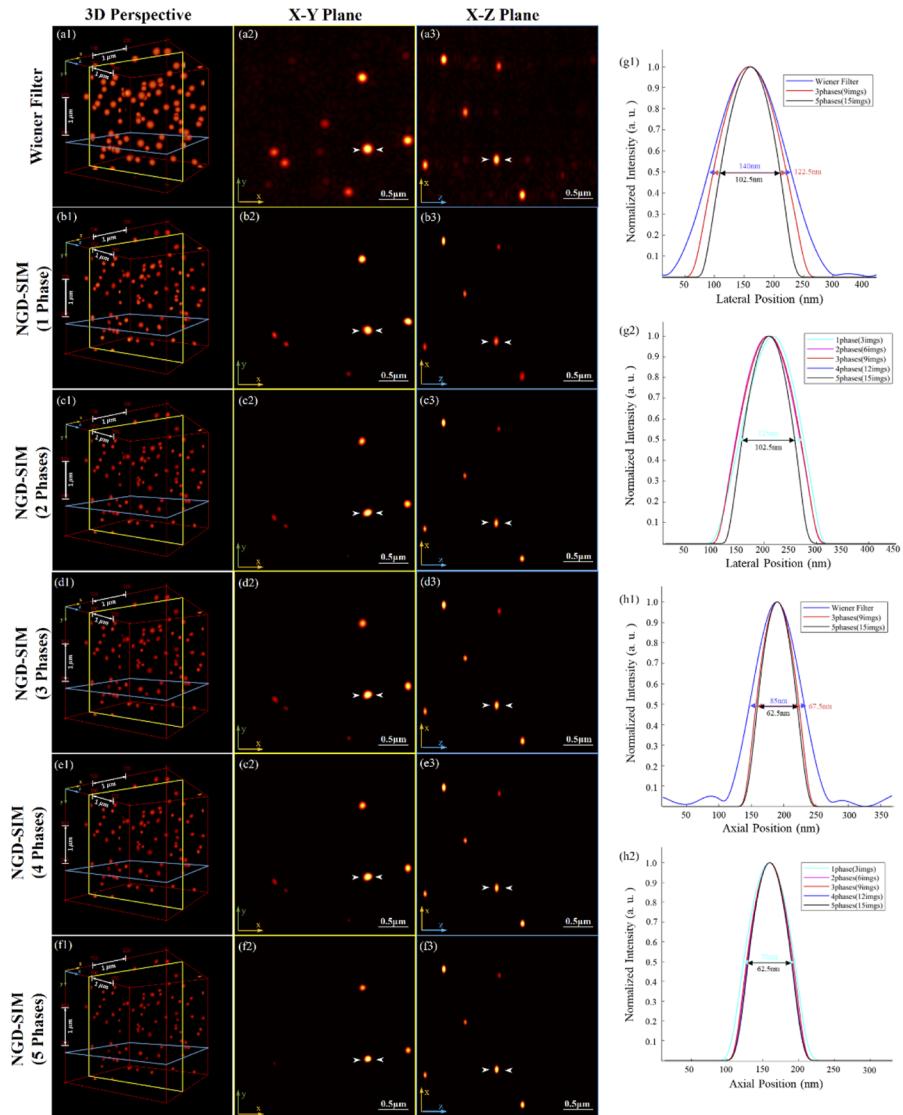


Fig. 2. Reconstruction results of simulated fluorescent beads of Wiener Filter and NGD-SIM using different numbers of phases (images). (a1)-(f1) 3D perspective of images reconstructed by Wiener Filter and NGD-SIM with 1–5 phases (3, 6, 9, 12, 15 images per slice) respectively. (a2)-(f2) single xy slices of corresponding 3D results (outlined by yellow planes in (a1)-(f1)) which are recovered by Wiener Filter and NGD-SIM with 1–5 phases (3, 6, 9, 12, 15 images per slice) respectively. (a3)-(f3) single xz slices of corresponding 3D results (outlined by blue planes of (a1)-(f1)) which are recovered by Wiener Filter and NGD-SIM with 1–5 phases (3, 6, 9, 12, 15 images per slice) respectively. In order to make the comparisons clearer and more intuitive, both the lateral profiles and axial profiles of single beads are divided into 2 groups: (g1) Lateral and (h1) axial profiles along the marked positions (the white arrows in (a2), (d2), (f2) and (a3), (d3), (f3)) by Wiener Filter (blue line), NGD-SIM with 3 phases (red line) and 5phases (black line) respectively; (g2) Lateral and (h2) axial profiles along the marked positions (the white arrows in (b2)-(f2) and (b3)-(f3)) by NGD-SIM with 1–5 phases (sapphire line, magenta line, red line, navy blue line and black line) respectively.

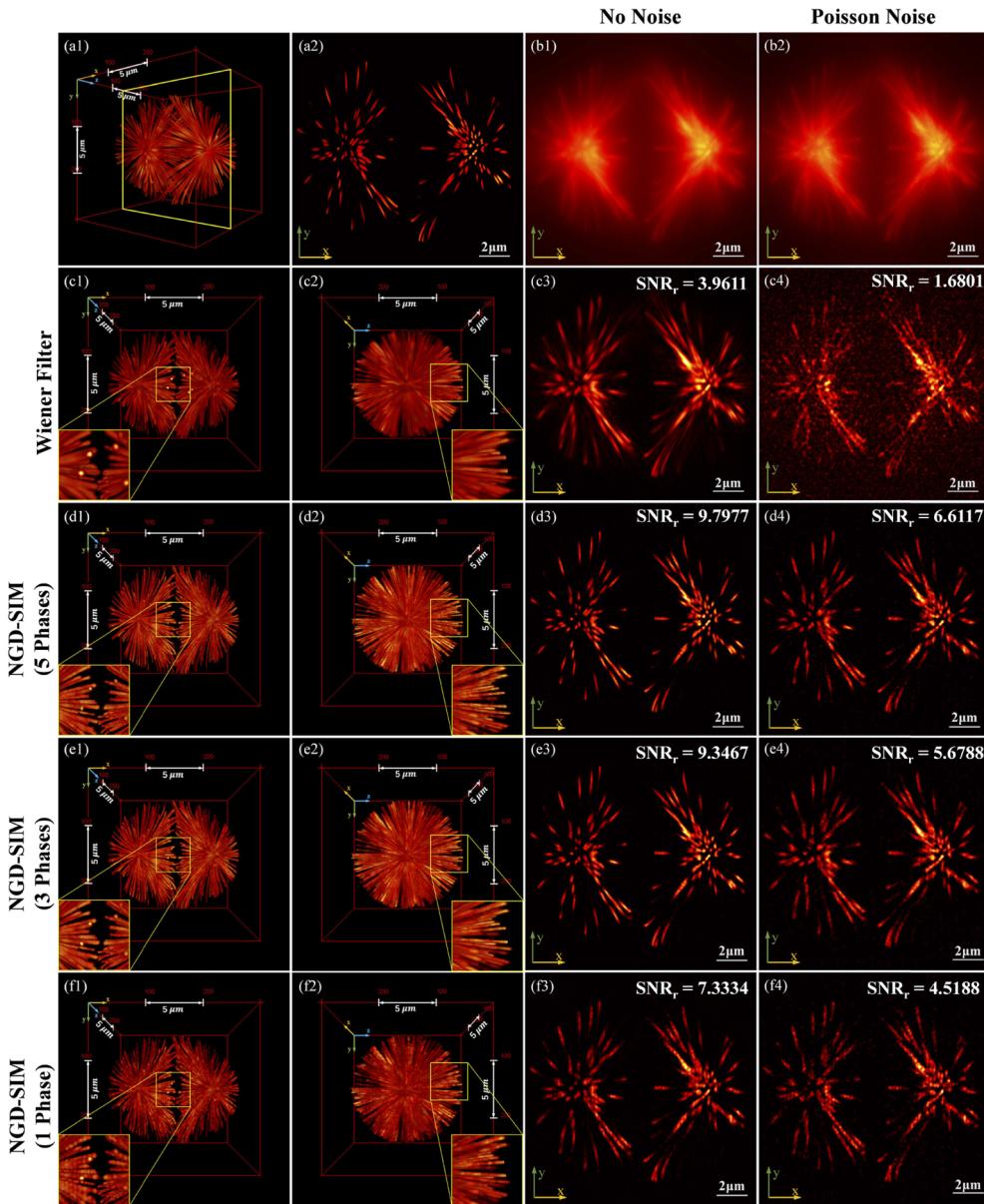


Fig. 3. Reconstruction results of a complex sample of Wiener Filter and NGD-SIM using different numbers of phases (images). (a1) 3D ground truth image. The yellow plane outlined in (a1) indicates the corresponding location of slices shown in (a2),(b1)-(b2), (c3)-(f3) and (c4)-(f4). (a2) Ground truth image of the yellow plane. (c1)-(c2), (d1)-(d2), (e1)-(e2), (f1)-(f2) Lateral and axial 3D perspectives of the images reconstructed by NGD-SIM using 5, 3 and 1 phases (15, 9 and 3 images respectively). The insets show expansions of the yellow outlined regions of the images. (b1), (c3)-(f3) Single xy slices of wide field image, Wiener Filter, NGD-SIM using different numbers of phases, respectively. (b2), (a4)-(f4) Single xy slices with Poisson Noise of wide field image, Wiener Filter, NGD-SIM using different numbers of phases, respectively. The SNR of recovered images (SNR_r) is marked on the upper right of each image.

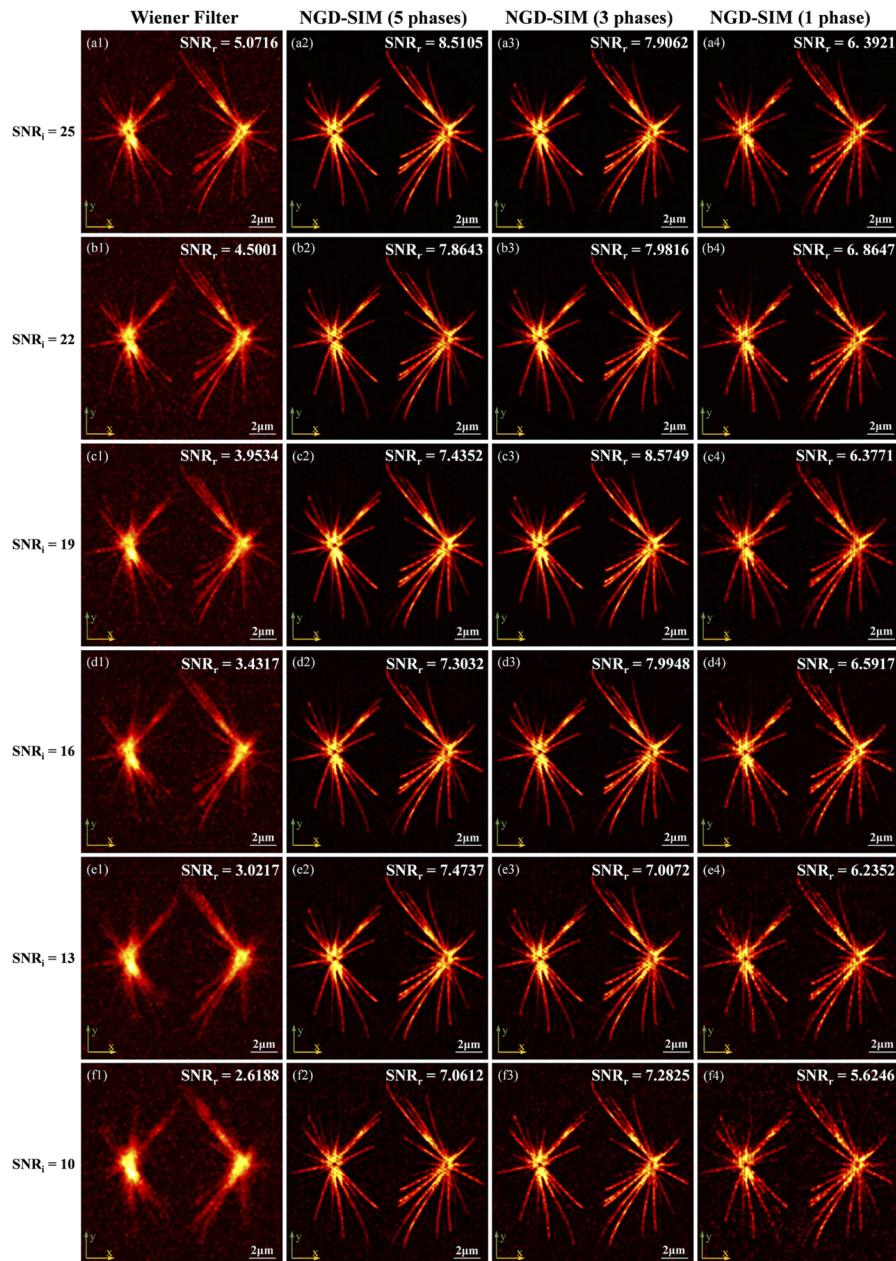


Fig. 4. single xy slices of reconstruction results of a complex sample by Wiener Filter and NGD-SIM using different phase (image) numbers under different degree of noises. The corresponding SNR of raw images (SNR_i) is marked on the far left. The SNR of recovered images (SNR_r) is marked on the upper right of each image. (a1)-(f1) Reconstruction results of Wiener Filter under different SNR_i (different degree of Gaussian noises). (a2)-(f2) Results of NGD-SIM reconstructed with 5 phases (15 images) under different SNR_i . (a3)-(f3) Results of NGD-SIM reconstructed with 3 phases (9 images) under different SNR_i . (a4)-(f4) Results of NGD-SIM reconstructed with 1 phase (3 images) under different SNR_i .

data should be used in the algorithm to process the image, so as to obtain reliable super-resolution results while reducing the amount of data used.

In addition, in order to test the robustness of the proposed algorithm against noise, Poisson noise, which is related to photon count and the system itself, is added to raw data stacks of microtubules. As illustrated in Fig. 3(b2), (c4)-(f4), the resolution and quality of wide-field images are degraded due to the noise. On the contrary, however, it is obvious that the novel algorithm can still significantly improve the image quality, which is much better than the traditional method. As for noise from the environment, different degree of Gaussian Noise is added to the whole image stacks of simulated microtubules and thus makes the raw data has different Signal to Noise Ratio (SNR_i). Figure 4 shows the reconstruction results of these raw data by Wiener Filter and NGD-SIM. It is obvious that no matter how the noise of raw data changes, the reconstruction results of NGD-SIM are always more superior and obtain higher SNR (SNR_r , as shown in Fig. 5) than using Wiener Filter and show stronger robustness to Gaussian Noise. Even the images processed with 1 phase are better than traditional Wiener Filter. What must be taken into account is the choice of parameter ω^2 in Wiener Filter, so we choose ω^2 which derives the best result to compare with NGD-SIM. The detail of the relation between ω^2 and reconstruction performances of Wiener Filter is shown in Appendix D (Fig. 11).

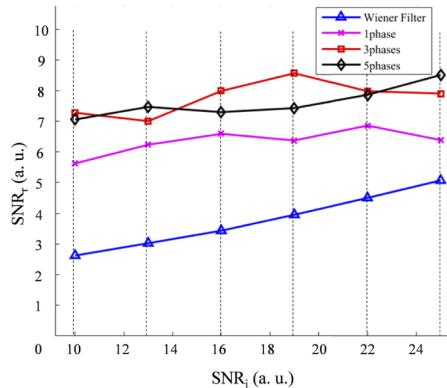


Fig. 5. Reconstruction quality by Wiener Filter and NGD-SIM, measured by SNR (marked as SNR_r), with different SNR of raw images (marked as SNR_i) and different numbers of phases (images).

To quantitatively measure the speed of NGD-SIM and Wiener Filter, both techniques were coded by MATLAB2020a and executed on the same device (Intel(R) Xeon(R) Gold 5120 CPU, 2.2 GHz, NVIDIA Quadro P6000, RAM 128GB). The size of a single frame image in the simulation is set as 301×301 and one 3D data will contain 301 frames. In order to meet the requirements of using GPU acceleration, we estimate that the memory of GPU needs at least 12G. The reconstruction process of Wiener Filter will take around 600s (10 minutes). Considering to achieve results with high-fidelity, we recommend presetting the number of iterations to 2000 when using NGD-SIM, which will take around 3000s (50 minutes), i.e., it will take 1.5s per iteration on average.

4. Conclusion and discussion

We proposed a novel iterative algorithm, namely NGD-SIM, based on gradient descent combined with a nonlinear optimizer called Root Mean Square Propagation (RMSprop), which allows us to reconstruct 3D raw data and obtain reconstruction images with higher fidelity for all kinds of 3D-SIM. Taking linear reconstruction algorithm for a reference, another distinguishing advantage of NGD-SIM is that it requires a smaller number of raw images but can achieve higher reconstruction resolution, which will greatly improve the imaging speed and temporal

resolution of the system in actual imaging. The proposed method is proven to be applicable for any illumination pattern, including both sinusoidal and speckle illumination. Therefore, it is entirely possible to further reduce the number of captured images if the illumination pattern is well-designed, like multi-spot patterns formed by the interference of four orthogonal light beams. We believe NGD-SIM will provide a more effective and precise method in any 3D structured illumination microscopic system or spatial frequency modulation super-resolution Imaging setup, including I^5S setup.

Appendix A: The definition and calculation of the Lipschitz constant

The definition of Lipschitz constant is as follows,

Assuming $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is a smooth convex and continuously differentiable function with Lipschitz continuous gradient $L(f)$:

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\| \leq L(f) \|\mathbf{x} - \mathbf{y}\| \quad \text{for every } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \quad (14)$$

Where $\|\cdot\|$ denotes L2 norm and $L(f) \geq 0$ is the Lipschitz constant of ∇f .

In fact, the reciprocal of the Lipschitz constant of ∇f was firstly taken as the step size in the Iterative Shrinkage/Thresholding algorithm (ISTA). The gradient descent algorithm and Iterative Shrinkage/Thresholding algorithm are similar in form (if the regularization term in ISTA is ignored, then the two algorithms are completely consistent). Literature [23] has proved in detail that ISTA will converge with the nonasymptotic global rate of convergence $O(1/k)$. For specific derivations, readers can refer to Literature [23], and we will not repeat them here. Thus, based on the above, we also take the step size of the gradient descent algorithm as $1/L$.

The Lipschitz constant of ∇f is calculated by Power Iteration algorithm and we formulate Power Iteration algorithm as follows:

Algorithm (Power Iteration)

Input: expression of ∇f , number of iterations N for Power Iteration algorithm

Initialization: $\mathbf{x} = 0$

For $n = 1, 2, \dots, N$

$$\mathbf{x} = \nabla f(\mathbf{x})$$

End

$$\mathbf{y} = \nabla f(\mathbf{x})$$

$$L(f) = \frac{\sum \mathbf{x} \cdot \mathbf{y}}{\sum \mathbf{x} \cdot \mathbf{x}} \quad (\sum A \text{ denotes the sum of all elements of matrix } A \text{ and } \cdot \text{ denotes pointwise product})$$

Output: $L(f)$ of ∇f

Appendix B: The influence of α parameter on the convergence speed and reconstruction results of NGD-SIM

Figure 6 illustrates the reconstruction results of NGD-SIM with different α by 2000 iterations. And Fig. 7 further shows the Mean Square Error of reconstruction results as a function of α and iteration number.

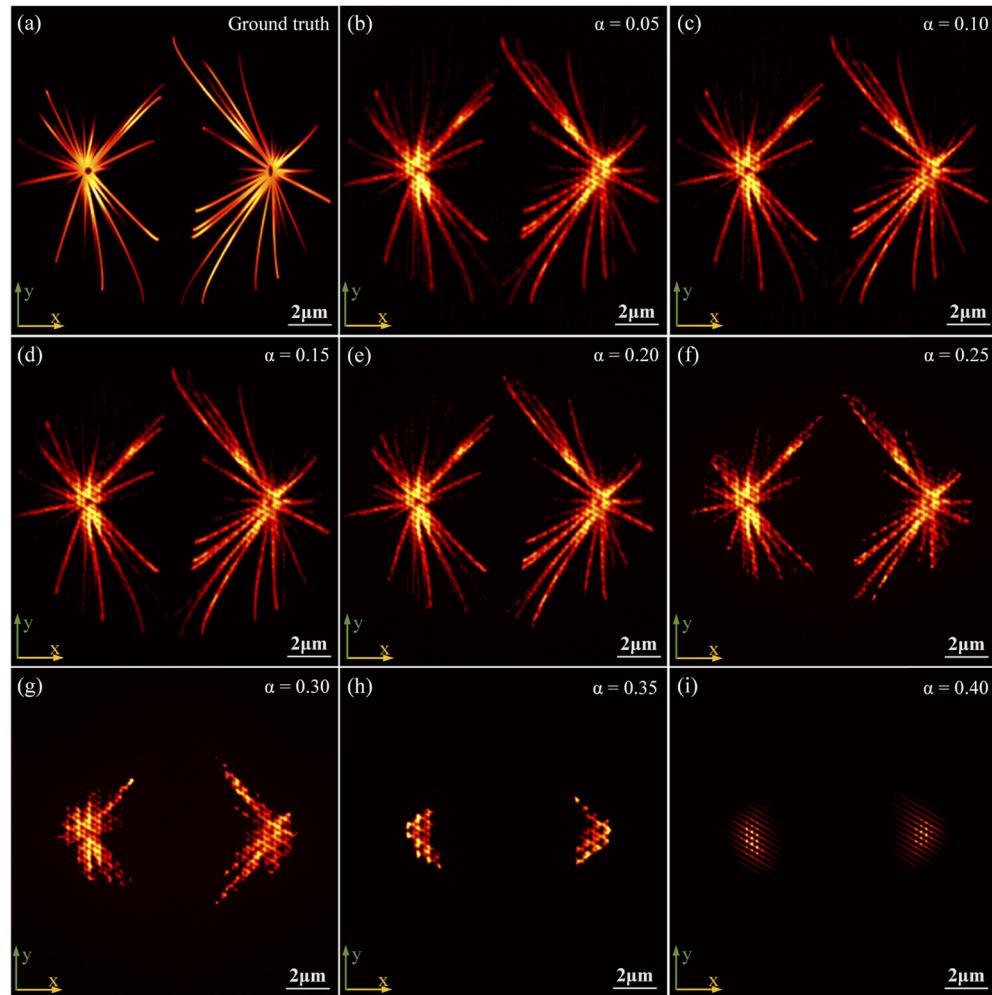


Fig. 6. Reconstruction results of NGD-SIM (5 phases, 15 images) with different α . (a) A xy slice of 3D groundtruth simulated sample. (b)-(i) Reconstruction results of NGD-SIM with different α by 2000 iterations. Corresponding α is marked in the upper right corner of each image.

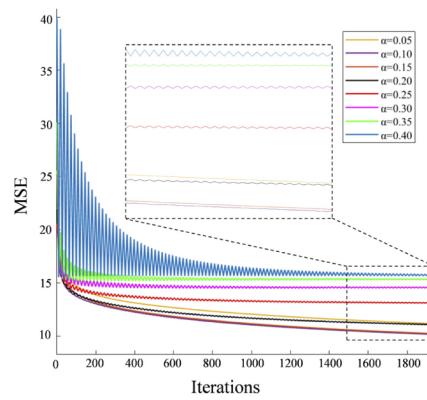


Fig. 7. Reconstruction performances with different α parameter and different number of iterations. The performances are valued by Mean Square Error (MSE). The black dotted box is an enlarged view of MSE under different α with the number of iterations around 1500–2000 times. When $\alpha = 0.1$, the reconstruction result is best.

Appendix C: comparison of convergence between NGD-SIM and ordinary gradient descent algorithm

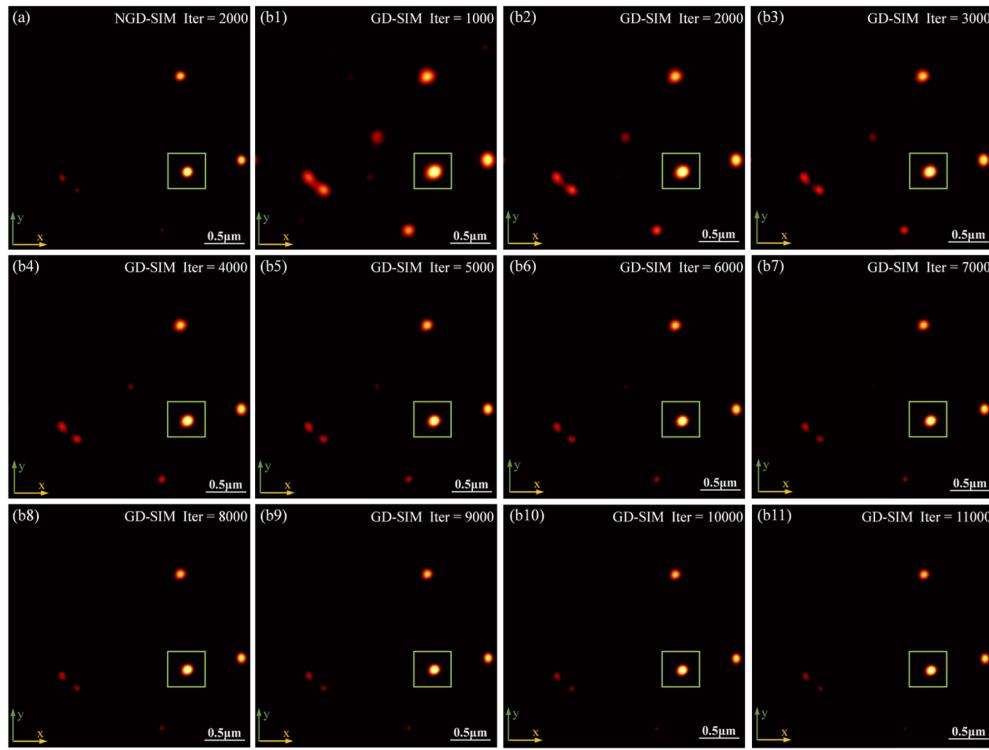


Fig. 8. Comparison of NGD-SIM and ordinary gradient descent algorithm (here marked as GD-SIM) for simulated fluorescent beads. (a) The result of NGD-SIM with 2000 iterations. (b1)-(b11) Results of GD-SIM with number of iterations from 1000 to 11000. The full width at half maximum (FWHM) of the beads in the green boxes under the influence of iterations is shown in Fig. 10.

Figure 8 and Fig. 9 are reconstruction results of NGD-SIM (2000 iterations) and ordinary gradient descent algorithm (GD-SIM, from 1000–12000 iterations). And Fig. 10 shows the MSE of reconstructed microtubules and FWHM of reconstructed beads as a function of iteration. It is obvious that NGD-SIM converges much fast than GD-SIM which denotes the better resolution and less processing time.

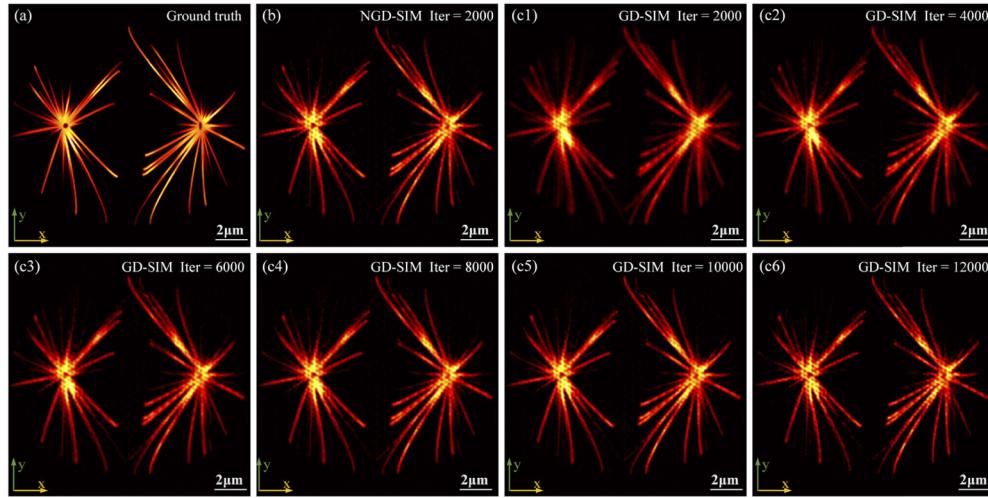


Fig. 9. Comparison of NGD-SIM and ordinary gradient descent algorithm (here marked as GD-SIM) for simulated microtubules. (a) a xy slice of 3D microtubules sample. (b) Results of NGD-SIM with 2000 iterations. (c1)-(c6) Results of GD-SIM with number of iterations from 2000 to 12000. The MSE of reconstructed results under the influence of iterations is shown in Fig. X.

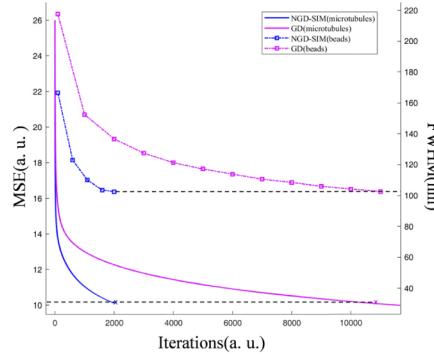


Fig. 10. Performances of reconstruction by NGD-SIM and GD-SIM with different number of iterations. The performances of reconstructed microtubules and fluorescent beads are measured by MSE and FWHM respectively. As shown in the black dotted lines, it is obvious that the reconstruction effect of 2000 NGD-SIM iterations is comparable to the counterpart with 12000 iterations for both simulated microtubules and fluorescent beads.

Appendix D: relation between parameter ω^2 and reconstruction results of Wiener Filter

The following Fig. 11 depicts the influence of ω^2 on the reconstruction performance of Wiener Filter. It is wise to choose appropriate ω^2 when processing raw data with different SNRⁱ.

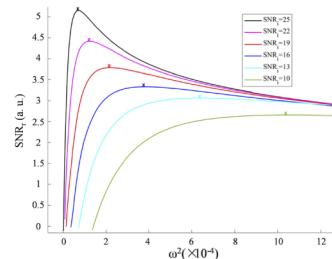


Fig. 11. Performances of reconstruction results by Wiener Filter with different parameter ω^2 and SNR of raw data (SNR_i). The results are evaluated by SNR (marked as SNR_r). The mark ‘x’ denotes the corresponding ω value when the SNR_r is the highest.

Funding. National Natural Science Foundation of China (61735017, 61827825); Natural Science Foundation of Zhejiang Province (LD21F050002); Zhejiang Provincial Ten Thousand Plan for Young Top Talents (2020R52001); Key Research and Development Program of Zhejiang Province (2020C01116); Fundamental Research Funds for the Central Universities (2019XZZX003-06, K20200132); Zhejiang Lab (2020MC0AE01).

Acknowledgements. We would like to thank Ruizhi Cao (Caltech) and Xin Liu (Zhejiang University) for the support and kind assistant.

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 not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. 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).
2. 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).
3. 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).
4. B. Dudok, L. Barna, M. Ledri, S. I. Szabó, E. Szabadits, B. Pintér, S. G. Woodhams, C. M. Henstridge, G. Y. Balla, and R. Nyilas, “Cell-specific STORM super-resolution imaging reveals nanoscale organization of cannabinoid signaling,” *Nature Neuroscience*.
5. 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).
6. M. G. Gustafsson, “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy,” *J. Microsc.* **198**(2), 82–87 (2000).
7. M. G. Gustafsson, L. Shao, P. M. Carlton, C. 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).
8. F. Strhl and C. F. Kaminski, “Frontiers in structured illumination microscopy,” *Optica* **3** (2016).
9. M. R. Arnison and C. J. Sheppard, “A 3D vectorial optical transfer function suitable for arbitrary pupil functions,” *Opt. Commun.* **211**(1-6), 53–63 (2002).
10. S. Hell and E. H. Stelzer, “Properties of a 4Pi confocal fluorescence microscope,” *J. Opt. Soc. Am. A* **9**(12), 2159–2166 (1992).
11. M. G. Gustafsson, D. A. Agard, and J. W. Sedat, “3D widefield microscopy with two objective lenses: experimental verification of improved axial resolution,” in *Three-Dimensional Microscopy: Image Acquisition and Processing III*, (International Society for Optics and Photonics, 1996), 62–66.
12. M. G. L. Gustafsson, D. A. Agard, and J. W. Sedat, “I5M: 3D widefield light microscopy with better than 100 nm axial resolution,” *J. Microsc.* **195**(1), 10–16 (1999).

13. L. Shao, B. Isaac, S. Uzawa, D. A. Agard, J. W. Sedat, and M. G. Gustafsson, "I5S: wide-field light microscopy with 100-nm-scale resolution in three dimensions," *Biophys. J.* **94**(12), 4971–4983 (2008).
14. L. Chen, X. Huang, J. Fan, L. Li, and S. Tan, "Fast, long-term super-resolution imaging with Hessian structured illumination microscopy," *Nature Biotechnology* 36(2018).
15. X. Hao, S. Tu, Q. Liu, X. Liu, and X. Liu, "Fast reconstruction algorithm for structured illumination microscopy," *Optics Letters* 45(2020).
16. L. H. Jin, B. Liu, F. Q. Zhao, S. Hahn, B. W. Dong, R. Y. Song, T. C. Elston, Y. K. Xu, and K. M. Hahn, "Deep learning enables structured illumination microscopy with low light levels and enhanced speed," *Nat Commun* 11(2020).
17. G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nat. Photonics* **7**(9), 739–745 (2013).
18. Q. Liu, Y. Yuan, W. Liu, X. Li, C. Kuang, X. Hao, and X. Liu, "Sub-60-nm 3D super-resolution imaging via saturated I5S," *Opt. Commun.* **473**, 125981 (2020).
19. R. Z. Cao, Y. H. Chen, W. J. Liu, D. Z. Zhu, C. F. Kuang, Y. K. Xu, and X. Liu, "Inverse matrix based phase estimation algorithm for structured illumination microscopy," *Biomed. Opt. Express* **9**(10), 5037–5051 (2018).
20. R. Heintzmann and T. Huser, "Super-Resolution Structured Illumination Microscopy," *Chem. Rev.* **117**(23), 13890–13908 (2017).
21. Yicong Wu and Hari Shroff, "Faster, sharper, and deeper: structured illumination microscopy for biological imaging," *Nature Methods* (2018).
22. T. Tieleman and G. Hinton, "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude," COURSERA: Neural networks for machine learning **4**, 26–31 (2012).
23. A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems," *SIAM J. Imaging Sci.* **2**(1), 183–202 (2009).