

Automatic deconvolution in 4Pi-microscopy with variable phase

Giuseppe Vicidomini, Roman Schmidt, Alexander Egner,
Stefan W. Hell, and Andreas Schönle*

Max-Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

[*aschoen@gwdg.de](mailto:aschoen@gwdg.de)

Abstract: 4Pi-microscopy doubles the aperture of the imaging system by coherent addition of the wavefronts for illumination and/or detection through opposing objective lenses. This improves the axial resolution 3-7 fold, but the raw data usually features ghost images which have to be removed by image reconstruction. This straightforward procedure is sometimes precluded by imperfect alignment of the instrument or a specimen with strong variations of its refractive index, because the image formation process now depends on the space-variant phase difference between the counter-propagating wavefronts. Here we present a computationally fast method of parametric blind deconvolution that allows for automatic and robust simultaneous estimation of both the object and the phase function in such cases. We verify the performance of our approach on both synthetic and real data. Because the method does not require a-priori knowledge of the phase function it is a major step towards reliable 4Pi-imaging and automatic image restoration by non-expert users.

© 2010 Optical Society of America

OCIS codes: (180.2520) Fluorescence microscopy; (180.6900) Three-dimensional microscopy; (100.1830) Deconvolution; (100.1455) Blind deconvolution.

References and links

1. S. W. Hell, "Double-scanning confocal microscope," European Patent (ISSN 0491289) (1990).
2. C. Cremer and T. Cremer, "Considerations on a laser-scanning-microscope with high resolution and depth of field," *Microsc. Acta* **81**(1), 31–44 (1978).
3. S. W. Hell and A. Schönle, "Nanoscale resolution in far-field fluorescence microscopy," in *Science of Microscopy*, P. W. Hawkes and J. C. H. Spence, eds. (Springer, 2007), pp. 790–834.
4. M. Nagorni and S. W. Hell, "Coherent use of opposing lenses for axial resolution increase in fluorescence microscopy. I. Comparative study of concepts," *J. Opt. Soc. Am. A* **18**(1), 36–48 (2001).
5. G. Vicidomini, S. W. Hell, and A. Schönle, "Automatic deconvolution of 4Pi-microscopy data with arbitrary phase," *Opt. Lett.* **34**(22), 3583–3585 (2009).
6. C. M. Blanca, J. Bewersdorf, and S. W. Hell, "Determination of the unknown phase difference in 4Pi-confocal microscopy through the image intensity," *Opt. Commun.* **206**(4–6), 281–285 (2002).
7. S. W. Hell, C. M. Blanca, and J. Bewersdorf, "Phase determination in interference-based superresolving microscopes through critical frequency analysis," *Opt. Lett.* **27**(11), 888–890 (2002).
8. D. Baddeley, C. Carl, and C. Cremer, "4Pi microscopy deconvolution with a variable point-spread function," *Appl. Opt.* **45**(27), 7056–7064 (2006).
9. T. Staudt, M. C. Lang, R. Medda, J. Engelhardt, and S. W. Hell, "2,2prime-Thiodiethanol: A new water soluble mounting medium for high resolution optical microscopy," *Microsc. Res. Tech.* **70**, 1–9 (2007).
10. M. C. Lang, T. Staudt, J. Engelhardt, and S. W. Hell, "4Pi microscopy with negligible sidelobes," *New J. Phys.* **10**, 043041 (2008).
11. M. Schrader, K. Bahlmann, G. Giese, and S. W. Hell, "4Pi-Confocal Imaging in Fixed Biological Specimens," *Biophys. J.* **75**(4), 1659–1668 (1998).

12. J. Enderlein, "Theoretical study of detection of a dipole emitter through an objective with high numerical aperture," *Opt. Lett.* **25**(9), 634–636 (2000).
13. B. Richards and E. Wolf, "Electromagnetic Diffraction in Optical Systems. II. Structure of the Image Field in an Aplanatic System," *Proc. R. Soc. Lond. A* **253**(1274), 358–379 (1959).
14. M. Bertero, P. Boccacci, G. Desiderà, and G. Vicidomini, "Image deblurring with Poisson data: from cells to galaxies," *Inverse Probl.* **25**, 123006 (2009).
15. A. Egner, M. Schrader, and S. W. Hell, "Refractive index mismatch induced intensity and phase variations in fluorescence confocal, multiphoton and 4Pi-microscopy," *Opt. Commun.* **153**(4–6), 211–217 (1998).
16. A. Egner, S. Verrier, A. Goroshkov, H.-D. Soling, and S. W. Hell, "4Pi-microscopy of the Golgi apparatus in live mammalian cells," *J. Struct. Biol.* **147**(1), 70–76 (2004).
17. G. Vicidomini, P. Boccacci, A. Diaspro, and M. Bertero, "Application of the split-gradient method to 3D image deconvolution in fluorescence microscopy," *J. Microsc.* **234**(1), 47–61 (2008).
18. L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astron. J.* **79**, 745–754 (1974).
19. W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration," *J. Opt. Soc. Am.* **62**(1), 55–59 (1972).
20. C. Kelley, *Iterative Method for Optimization*, vol. 18 (SIAM, Philadelphia, 1999).
21. D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via l_1 minimization," *Proc. Natl. Acad. Sci. U.S.A.* **100**(5), 2197–2202 (2003).
22. H. Lantéri, M. Roche, O. Cuevas, and C. Aime, "A general method to devise maximum-likelihood signal restoration multiplicative algorithms with non-negativity constraints," *Signal Process.* **81**, 945–974 (2001).
23. I. Csizsár, "Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problem," *Ann. Stat.* **19**, 2032–2066 (1991).

1. Introduction

By utilizing high aperture angles, modern objective lenses provide lateral resolution at the diffraction limit. However, due to the asymmetry of an imaging system featuring just a single lens, the axial resolution is several times worse than the lateral resolution. By employing two opposing lenses and coherent illumination and/or detection from both sides, the family of 4Pi microscopes [1] doubles the aperture used for imaging, thus creating and/or detecting a total wavefront that comes closer to a complete spherical wavefront. Speculating about the use of the full solid angle in focusing, a paper of 1978 suggested that appropriate illumination from all directions would allow focusing the light virtually to a point, or at least much sharper than the diffraction limit, thus bringing about infinite spatial resolution in far-field microscopy [2]. This conjecture was based on the assumption that by holographically reconstructing the wavefronts originating from a nanosized light source (and reversing their propagation direction) a focal spot of the same nanosized dimension as the source could be created. This assumption clearly does not hold, because only the propagating part of the field but not the evanescent waves around the original source can be reconstructed. This is in fact the very essence of the diffraction limit: high spatial frequencies cannot be retrieved from light emitted by an object or encoded in the light being focused because they decrease exponentially with increasing distance from the light source regardless of the focusing geometry. The conjecture is thus invalid and in contrast, a 4Pi microscope as introduced [1] and known today relies on a total wavefront that is far from being completely spherical. Moreover, the goal is not to overcome the diffraction limit, but to provide a 3-7 fold increase in axial resolution that is still within the realm of diffraction. This important gain in axial resolution is achieved due to the alleviation of the asymmetry inherent to focusing from a single side only. By enlarging the total solid focusing angle as much as it is technically possible with aplanatic objective lenses, this arrangement pushes the far-field optical resolution in all three dimensions towards the diffraction limit.

The fact that the converging or collected total wavefront of the two-lens system is not spherical leads to side-lobes in both the illumination intensity pattern (I-PSF) of a 4Pi type A or C microscope and the detection PSF (D-PSF) of a 4Pi type B or C microscope and thus also in the effective PSF (E-PSF) of the system [3]. In the data, these lobes produce characteristic replica of the object which have to be removed from the final image by deconvolution.

Reliable and artifact-free image reconstruction is currently only possible if the lobe height is sufficiently low ($< 50\%$ of the main maximum for constructive interference) [4] and, because the lobe pattern depends on the relative phase of the opposing beam-paths, if the relative phase of the two beam paths is either constant [5–7] or if it can be estimated *a-priori* at every point of the object [8].

Importantly, the lobe height is readily controlled by employing e.g. two-photon excitation (TPE), confocal detection or using coherent detection in a type C arrangement [3]. In fact, employing an embedding medium of high refractive index, e.g. the water-miscible TDE (2,2-Thiodiethanol) [9], oil immersion and very high numerical apertures (NA) has resulted in virtually lobe-free TPE 4Pi microscopy [10]. Unfortunately only a sub-class of fixed samples allows this combination of measures and the phase-problem remains for all other applications. The *a-priori* estimation of an unknown phase relies on isolated, axially thin objects which have to be distributed throughout a sufficiently sparse sample [8]. In most application it is therefore currently mandatory to match the refractive indices of the specimen and the embedding medium and to either accurately monitor or compensate for the phase variation introduced by the axial scanning process [11]. While this combination of measures is often feasible, it is time-consuming, adds experimental complexity and restricts the choice of possible samples: For specimen which intrinsically feature a non-negligible and unknown variation of refractive index, the relative phase and thus the form of the E-PSF will always vary throughout the object rendering reliable image interpretation based on existing methods virtually impossible. We have therefore developed a parametric-blind deconvolution method that allows for simultaneous estimation of both, the unknown, spatially varying phase and the object, without prior knowledge of either. Based on a maximum *a-posteriori* approach, it allows for a space-variant relative phase function (RPF) $\phi(\mathbf{r})$ which is either assumed to be smooth or approximated by a linear combination of a few polynomial basis functions. Both approaches are tested on both synthetic and real data and compared. We verified the applicability of our method to one-photon confocal 4Pi type A microscopy, because this particular configuration is experimentally most simple, least demanding on sample brightness and stability and commercially available and because the data analysis is more challenging than in the case of TPE or coherent detection which result in lower side lobes and therefore less ill-posed problems.

2. Imaging model

Since imaging in the scanning 4Pi fluorescence microscope is incoherent, the relation between the expected image $\bar{g}(\mathbf{r})$ and the object $f(\mathbf{s})$ in the absence of noise can be described by the following equation

$$\bar{g}(\mathbf{r}) = \int h_{\text{eff}}(\mathbf{r}, \mathbf{s}) f(\mathbf{s}) d\mathbf{s} \quad (1)$$

where h_{eff} is the 4Pi E-PSF, \mathbf{s} the position in the sample and \mathbf{r} is the position in the image back projected into the sample, i.e. the scanning position. For space-invariant PSFs where $h_{\text{eff}}(\mathbf{r}, \mathbf{s}) = h_{\text{eff}}(\mathbf{r} - \mathbf{s})$ Eq. (1) is a convolution and the forward operator in common restoration techniques can be efficiently implemented using fast Fourier transforms (FFT). Here, the space variant phase function precludes this by rendering the form of the PSF position dependent. Fortunately the special form of the 4Pi E-PSF allows its decomposition into a ϕ -dependent linear combination of a number of phase-independent and thus space-invariant convolution operators which can then be efficiently implemented using FFT algorithms. This is either achieved by using a convenient approximation of the PSF [8] or by deriving an adequate representation of the PSF using vectorial focusing theory [5]. Importantly, this approach also allows re-calculation of the forward operator for a given object but various phase functions without re-evaluation of the convolutions. In a previous publication we have used the latter technique to efficiently esti-

mate both the object and a constant global phase directly from the image [5]. Here we extend the approach by replacing the constant global phase by the space-variant RPF. The form of the E-PSF of a 4Pi fluorescence microscope of type A using one photon excitation is given by

$$h_{\text{eff}}(\mathbf{r}, \mathbf{s}) = |\mathbf{E}_1(\mathbf{r} - \mathbf{s}) + \exp(i\phi(\mathbf{r}))\mathbf{E}_2(\mathbf{r} - \mathbf{s})|^2 h_{\text{det}}(\mathbf{r} - \mathbf{s}) = h_{\text{eff}}(\mathbf{r} - \mathbf{s}, \phi(\mathbf{r})) \quad (2)$$

where $\mathbf{E}_{1,2}$ denotes the counter-propagating focal fields, with $\mathbf{E}_2(\mathbf{r}) = \mathbf{M}\mathbf{E}_1(\mathbf{M}\mathbf{r})$. The transformation matrix \mathbf{M} accounts for the counter-propagation of the beams and h_{det} is the D-PSF. In the case of confocal detection, the D-PSF is given by the integral of the field created by a dipole in the image plane over the finite-sized pinhole [12] and $\mathbf{E}_{1,2}$ are readily evaluated using standard vectorial focusing theory [13]. The forward operator H is thus defined by

$$\bar{\mathbf{g}}(\mathbf{r}) = [H(\phi)f](\mathbf{r}) = \int h_{\text{eff}}(\mathbf{r} - \mathbf{s}, \phi(\mathbf{r}))f(\mathbf{s})d\mathbf{s}. \quad (3)$$

Where the operator $H(\phi)f$ depends nonlinearly on the phase function ϕ but linearly on f for any given ϕ . Using Eq. (2) it can be written as

$$H(\phi)f = A_0f + \cos(\phi)A_1f - \sin(\phi)A_2f \quad (4)$$

$$H'(\phi)f = -\sin(\phi)A_1f - \cos(\phi)A_2f \quad (5)$$

where we have also introduced its derivative with respect to the phase, H' . The convolution operators are independent of the RPF and given by

$$A_0f = [(|\mathbf{E}_1|^2 + |\mathbf{E}_2|^2)h_{\text{det}}] \otimes f \quad (6)$$

$$A_1f = [2\Re(\mathbf{E}_1^*\mathbf{E}_2)h_{\text{det}}] \otimes f \quad (7)$$

$$A_2f = [2\Im(\mathbf{E}_1^*\mathbf{E}_2)h_{\text{det}}] \otimes f. \quad (8)$$

They are readily computed using the Fourier space multiplication technique requiring a few FFTs. A similar decomposition using five convolutions involving fourth powers of the fields is readily derived for two-photon excitation, where the I-PSF depends on the squared intensity. We note that the forward operator proposed by Baddeley and coworkers [8] requires the same number of convolutions but is based on an approximation of the E-PSF that becomes inaccurate for high numerical apertures.

Images are usually acquired on a regular 3-dimensional grid using photon counting devices. If we identify each voxel by its index \mathbf{n} and denote the discretized object, expected image and RPF by the vectors \mathbf{f} , $\bar{\mathbf{g}}$ and ϕ , we can write Eqs. (4) and (3) as

$$\bar{\mathbf{g}} = \mathbf{H}(\phi)\mathbf{f} + \mathbf{b} = \mathbf{A}_0\mathbf{f} + \cos(\phi)\mathbf{A}_1\mathbf{f} - \sin(\phi)\mathbf{A}_2\mathbf{f} + \mathbf{b}. \quad (9)$$

where the \mathbf{A}_j are the convolution matrices associated with the convolution operators A_j . Here and in all subsequent equations multiplication and division of one vector by another is meant voxel by voxel. We introduced also a possible additional background contribution \mathbf{b} . The measurement process is dominated by shot noise and count rates are usually in the range of zero to a few hundred photons per voxel. The measured value $\mathbf{g}(\mathbf{n})$ at a voxel \mathbf{n} is therefore the realization of an independent Poisson random variable with its expectation value given by $\bar{\mathbf{g}}(\mathbf{n})$. During reconstruction we maximize the *a-posteriori* probability of observing the specimen \mathbf{f} and RPF ϕ given the image \mathbf{g} . It is given by

$$P(\mathbf{f}, \phi | \mathbf{g}) = P(\mathbf{g} | \mathbf{f}, \phi)P(\mathbf{f})P(\phi) / P(\mathbf{g}). \quad (10)$$

$P(\phi)$, $P(\mathbf{g})$ and $P(\mathbf{f})$ denote the *a-priori* probability distributions for the phase, for the image and for the specimen, respectively and the probability of observing \mathbf{g} is given by

$$P(\mathbf{g} | \mathbf{f}, \phi) = \prod_{\mathbf{n}} \text{poi}[\mathbf{g}(\mathbf{n}) | \bar{\mathbf{g}}(\mathbf{n})] \quad (11)$$

where $\text{poi}(g|\lambda) = \lambda^g \exp(-\lambda)/g!$. We assume that the random variables associated with the object and the RPF are Gibbs random fields and thus

$$P(\mathbf{x}) = \exp[-\mu_x J_x(\mathbf{x})]/Z, \quad (12)$$

where $\mathbf{x} = \mathbf{f}, \boldsymbol{\phi}$ for object and RPF respectively, $\mu_x > 0$ is the regularization parameter, Z is the partition function, and J_x is the penalization function, or simply prior. Maximization of the *a-posteriori* probability distribution with respect to \mathbf{f} and $\boldsymbol{\phi}$ is then equivalent to minimization of its negative logarithm, the so called cost function

$$J(\mathbf{f}, \boldsymbol{\phi}|\mathbf{g}) = J_o(\mathbf{f}, \boldsymbol{\phi}|\mathbf{g}) + \mu_f J_f(\mathbf{f}) + \mu_\phi J_\phi(\boldsymbol{\phi}), \quad (13)$$

where constant terms are omitted and the likelihood term J_o is given by

$$J_o(\mathbf{f}, \boldsymbol{\phi}|\mathbf{g}) = \sum_{\mathbf{n}} \bar{\mathbf{g}}(\mathbf{n}) - \mathbf{g}(\mathbf{n}) \ln \bar{\mathbf{g}}(\mathbf{n}). \quad (14)$$

The regularization parameters are usually adjusted empirically and the penalization functions J_f, J_ϕ , used to introduce *a-priori* knowledge about the object and RPF into the reconstruction, have to be chosen carefully. This kind of maximum a-posteriori (MAP) approach is well established in microscopy and therefore, suitable object penalization functions for typical objects have been proposed in the literature (for a recent review see [14] and references therein). The most common ones are Tikhonov regularization $J_f(f) = \|f\|_2^2/2$ which imposes an energy constraint on the restored object, the H^1 case $J_f(f) = \|\nabla f\|_2^2/2$ that favors smooth objects and total variation (TV) regularization $J_f(f) = \|\nabla f\|_1$, for objects with sharp edges and areas of (nearly) constant intensity. Here, $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the usual ℓ_1 and ℓ_2 norms. For the microtubule images analyzed in this work we found that classical Tikhonov penalization resulted in visually appealing reconstructed objects. As a side-benefit, its implementation has no significant impact on calculation time when compared to the unconstrained case $J_f(f) = 0$.

The key to successful simultaneous estimation of phase and object is the appropriate choice of the prior for the phase, J_ϕ . It has been shown that the RPF induced solely by refractive index mismatch between mounting medium and immersion medium depends linearly on the axial scanning position for moderate sample thickness [15]. While contributions from the object, e.g. small organelles and the nucleus break this simple dependence, the RPF usually remains smooth and never exhibits phase jumps [11, 16]. In this work we therefore explore two approaches for the penalization function, J_ϕ , based on either allowing for an arbitrary but smooth RPF or for such relative phase functions that can be described by a linear combination of a few suitable polynomials.

3. Arbitrary, smooth phase function

Here, we allow for an arbitrary phase function and hence we call this approach FPBR (full parametric blind reconstruction). Regularization is achieved by favoring smooth RPFs using H^1 penalization. To discretize the H^1 penalization function we assume, for simplicity, that the $N_x \times N_y \times N_z$ array $\boldsymbol{\phi}$ is expanded by taking $\boldsymbol{\phi}(N_x + 1, n_y, n_z) = \boldsymbol{\phi}(N_x, n_y, n_z)$, $\boldsymbol{\phi}(n_x, N_y + 1, n_z) = \boldsymbol{\phi}(n_x, N_y, n_z)$ and $\boldsymbol{\phi}(n_x, n_y, N_z + 1) = \boldsymbol{\phi}(n_x, n_y, N_z)$. Under this assumption we write the discrete version of the H^1 penalization function in the following form

$$J_\phi(\boldsymbol{\phi}) = \frac{1}{2} \sum_{\mathbf{n}} \sum_{\mathbf{m} \in \eta_{\mathbf{n}}} [\boldsymbol{\phi}(\mathbf{m}) - \boldsymbol{\phi}(\mathbf{n})]^2, \quad (15)$$

where $\eta_{\mathbf{n}} = \{(n_x + 1, n_y, n_z), (n_x, n_y + 1, n_z), (n_x, n_y, n_z + 1)\}$ are the upper direct neighboring voxels of \mathbf{n} , $\boldsymbol{\phi}(\mathbf{n}) \in \mathbb{R}$ and J_ϕ is calculated independently of the interpretation of $\boldsymbol{\phi}$ as a phase

during minimization. We note that RPF's that differ by constant offsets of multiples of 2π are obviously equivalent and identify the same solution making this a proper prior. To minimize the resulting cost function $J(\mathbf{f}, \boldsymbol{\phi} | \mathbf{g})$ simultaneously in \mathbf{f} and $\boldsymbol{\phi}$ we use a similar alternating minimization method (AMM) as in our previous work [5]. For the l -th cycle we have

$$\mathbf{f}_{l+1} = \arg \min_{\mathbf{f} \geq 0} J(\mathbf{f}, \boldsymbol{\phi}_l | \mathbf{g}) \quad (16)$$

$$\boldsymbol{\phi}_{l+1} = \arg \min_{\boldsymbol{\phi}} J(\mathbf{f}_{l+1}, \boldsymbol{\phi} | \mathbf{g}) \quad (17)$$

where Eq. (16) adds non-negativity of the object as a constraint. Thus, during each cycle first the RPF is kept fixed and then the object is kept fixed while the other is updated by minimizing the cost function w.r.t. to it. Therefore, $J(\mathbf{f}_l, \boldsymbol{\phi}_l | \mathbf{g})$ is guaranteed to decrease as the number of cycles l increases. It is important to point out that, while convex in \mathbf{f} for fixed RPF, $J(\mathbf{f}, \boldsymbol{\phi} | \mathbf{g})$ is non-convex as a function of both variables and therefore, convergence of the AMM to a global minimum is not guaranteed. The quality of the local minimum found may therefore depend on the initial guess of the RPF. We will see that a suitable strategy can be devised to ensure visually appealing restored objects. For the solution of Eq. (16) we chose the split-gradient-method (SGM) [17] due to its robustness, the simplicity of its implementation and its capability to enforce the non-negativity constraint in a natural fashion. In the case of Poisson data and Tikhonov penalization the SGM iterations are given by (see appendix for a derivation):

$$\mathbf{f}_{l,i+1} = \frac{\mathbf{f}_{l,i}}{\mathbf{H}(\boldsymbol{\phi}^l)^T \mathbf{1} + \mu_f \mathbf{f}_{l,i}} \left(\mathbf{H}(\boldsymbol{\phi}^l)^T \frac{\mathbf{g}}{\mathbf{H}(\boldsymbol{\phi}^l) \mathbf{f}_{l,i} + \mathbf{b}} \right). \quad (18)$$

For $\mu_f = 0$ the SGM algorithm is equivalent to the well-known Richardson-Lucy [18, 19] algorithm for the linear operator $\mathbf{H}(\boldsymbol{\phi})$. We note that a similar iterative algorithm can be obtained for all the object penalization functions mentioned above. For the RPF update we adopt a quasi-Newton method (QNM) [20] defined by the iteration

$$\boldsymbol{\phi}_{l,i+1} = \boldsymbol{\phi}_{l,i} - \tau_{l,i} \boldsymbol{\rho}(\boldsymbol{\phi}_{l,i}). \quad (19)$$

The descent direction $\boldsymbol{\rho}$ is given by

$$\boldsymbol{\rho}(\boldsymbol{\phi}_{l,i}) = (\mathbf{C}_{l,i})^{-1} \nabla_{\boldsymbol{\phi}} [J(\mathbf{f}_{l+1}, \boldsymbol{\phi} | \mathbf{g})] (\boldsymbol{\phi}_{l,i}) \quad (20)$$

where $\mathbf{C}_{l,i}$ is the Hessian matrix of the function $J(\mathbf{f}_{l+1}, \boldsymbol{\phi} | \mathbf{g})$ w.r.t. $\boldsymbol{\phi}$. An analytical expression for the gradient is derived in the appendix while an approximation of \mathbf{C} can be calculated using the Broyden, Fletcher, Goldfarb and Shanno (BFGS) formula [20]. However, due to the large size of $\boldsymbol{\phi}$ on the order of a megavoxel, we implemented the limited-BFGS approach where \mathbf{C}^{-1} is never explicitly formed or stored. Rather, a direct estimate for the descent direction $\boldsymbol{\rho}$ is obtained by maintaining a history of past updates (< 10) of $\boldsymbol{\phi}$ and the gradient $\nabla_{\boldsymbol{\phi}} [J(\mathbf{f}_{l+1}, \boldsymbol{\phi} | \mathbf{g})]$ [20]. The step size $\tau_{l,i}$ can be efficiently determined by the Armijo line-search method [20], because the several calculations of $J(\mathbf{f}_{l+1}, \boldsymbol{\phi} | \mathbf{g})$ necessary for various RPFs do not require re-evaluation of the FFTs due to the decomposition of H . As outlined below, suitable stopping criteria for both the SGM and QNM parts of the cycle have to be defined.

4. Sparsity-based approach

Voxel-by-voxel estimation of the RPF has the disadvantage that the number of parameters to estimate (one parameter per voxel) is very large effectively slowing down the convergence. By using a sparsity-based approach, an emerging and powerful method in image processing [21], the number of free parameters can be significantly reduced. It is based on the observation,

that many natural signals can be represented as a linear combination of a few (sparse) well chosen basis functions. We assume that given a suitable set of basis functions ϕ_p , any RPF is approximated well by a sparse linear combination of them. Accordingly we will refer to this approach as SPBR (sparse parametric blind reconstruction). We can write

$$\phi \simeq \mathbf{D}\boldsymbol{\psi} = \sum_{\mathbf{p}} \psi_{\mathbf{p}} \phi_{\mathbf{p}} \quad (21)$$

with $\|\boldsymbol{\psi}\|_0 \ll N$, N being the number of voxels. We call $\boldsymbol{\psi}$ the RPF's coefficients vector and the operator \mathbf{D} the dictionary and $\|\cdot\|_0$ is the ℓ_0 pseudo-norm, which counts the number of nonzero elements in a vector. The choice of dictionary plays a crucial role since it implies strong assumptions about the behavior of the RPF. With ϕ being a smooth function, several choices of smooth basis functions (which need not be orthogonal) seem reasonable, including (steerable) wavelets and curvelets. In this work we selected products of polynomials in normalized cartesian coordinates $\mathbf{r} = (x, y, z)$

$$\phi_{\mathbf{p}}(\mathbf{r}) = x^{p_1} y^{p_2} z^{p_3} \quad (22)$$

with $p_1 + p_2 + p_3 \leq N_{\phi}$ and the normalization of the coordinates was chosen such that the longest edge has length 1. The optimum normalization and dictionary size, i.e. the choice of N_{ϕ} largely depends on prior knowledge about the specimen. Choosing a large dictionary may increase computational complexity and slow convergence while small dictionaries may not be able to model a realistic RPF. In the limiting case where the phase variation is solely induced by refractive index mismatch between mounting and immersion medium the RPF is known to be linear and we could set N_{ϕ} to 1. For the experimental data analyzed in this work $N_{\phi} \leq 6$ (i.e. a dictionary size of $N_{\psi} \leq 84$) proved adequate.

The cost function is now given by $J(\mathbf{f}, \phi | \mathbf{g}) = J(\mathbf{f}, \mathbf{D}\boldsymbol{\psi} | \mathbf{g})$ and has to be minimized w.r.t \mathbf{f} and the component vector $\boldsymbol{\psi}$. Sparsity could be enforced using the ℓ_0 pseudo-norm penalization function, i.e. $J_{\phi}(\phi) = \|\boldsymbol{\psi}\|_0$ but its minimization results in a hard numeric problem. It was shown that for several problems, a sparse solution can also be obtained by replacing the ℓ_0 pseudo-norm with the ℓ_1 norm, i.e. $J_{\phi}(\phi) = \|\boldsymbol{\psi}\|_1$ [21]. Motivated by this work, we implemented the minimization using the same AMM as above, simply replacing the large vector $\phi_{l,i}$ by the component vector $\boldsymbol{\psi}_{l,i}$. An analytical expression of $\nabla_{\boldsymbol{\psi}}[J(\mathbf{f}_{l+1}, \mathbf{D}\boldsymbol{\psi} | \mathbf{g})]$ is given in the appendix and the Hessian matrix $\mathbf{C}_{l,i}$ w.r.t. to the much smaller vector $\boldsymbol{\psi}$ can now be directly implemented using the BFGS method.

4.1. Stopping criteria and initialization

During each cycle, the SGM algorithm was stopped when the relative change during one iteration dropped below a certain threshold ϵ_{SGM} , i.e.

$$\|\mathbf{f}_{l,i+1} - \mathbf{f}_{l,i}\|_2 \leq \epsilon_{\text{SGM}} \|\mathbf{f}_{l,i}\|_2. \quad (23)$$

To stop the QNM algorithm it turns out to be much safer to terminate the iteration when the norm of the gradient is sufficiently small compared to the norm of the gradient at the initialization value [20]:

$$\|\nabla_x[J(\mathbf{f}_{l+1}, \phi | \mathbf{g})](\mathbf{x}_{l,i})\|_2 \leq \epsilon_{\text{QNM}} \|\nabla_x[J(\mathbf{f}_{l+1}, \phi | \mathbf{g})](\mathbf{x}_{l,0})\|_2, \quad (24)$$

where $\mathbf{x} = \phi, \boldsymbol{\psi}$ for the FPBR and SPBR approach respectively. The choice of the two parameters ϵ_{SGM} and ϵ_{QNM} was based on the following consideration. At early AMM cycles the object estimate changes significantly from one SGM iteration to the next, and thus it is desirable to update the RPF after just a few updates to the object. However, because the object estimate is changing rapidly, the RPF estimate changes significantly every time it is updated, and thus

making the tolerance ϵ_{QNM} too small at the early AMM cycles would waste computation time. Therefore, we choose the two parameters ϵ_{SGM} and ϵ_{QNM} as a function of the AMM cycle. We therefore start the AMM algorithm with somewhat loose tolerance values which are tightened in later cycles. From the test cases that we ran, we found that the combination of tolerance values $\epsilon_{\text{SGM}} = 10^{\max(-6, -\lfloor l/25 \rfloor - 1)}$ and $\epsilon_{\text{QNM}} = 10^{\max(-4, -\lfloor l/25 \rfloor - 2)}$ gave a good compromise between the computation time required and the quality of the results obtained. To stop the AMM algorithm we used a criterion based on the difference of the values of the cost function between two consecutive AMM cycles

$$|J(\mathbf{f}_{l+1}, \boldsymbol{\phi}_{l+1} | \mathbf{g}) - J(\mathbf{f}_l, \boldsymbol{\phi}_l | \mathbf{g})| \leq \epsilon_{\text{AMM}} |J(\mathbf{f}_l, \boldsymbol{\phi}_l | \mathbf{g})| \quad (25)$$

A typical value for ϵ_{AMM} is 10^{-8} for simulation and 10^{-7} for real data. In all examples we initialized \mathbf{f}_0 with a constant vector and $\boldsymbol{\phi}_0$ with a constant RPF.

5. Validation on simulated data

Both PBR approaches were tested on simulated data and an example is shown in Fig. 1. We used a 2-dimensional (xz) image in order to speed up evaluation and parameter optimization and created a test pattern that is reminiscent of different subcellular structures, i.e. microtubules, Golgi and mitochondria. Using a strongly varying cubic RPF ($N_\phi = 3$) we obtained the simulated noise-free image using the forward operator from Eq. (9), added a constant background and subsequently applied a Poisson noise perturbation. For quantitative comparison of the quality of the restored objects, we use the Kullback-Leibler divergence (D_{KL}) of the estimate \mathbf{f}_E from the original \mathbf{f}

$$D_{\text{KL}}(\mathbf{f}, \mathbf{f}_E) = \sum_{\mathbf{n}} \{ \mathbf{f}(\mathbf{n}) \ln[\mathbf{f}(\mathbf{n}) - \mathbf{f}_E(\mathbf{n})] + [\mathbf{f}_E(\mathbf{n}) - \mathbf{f}(\mathbf{n})] \}, \quad (26)$$

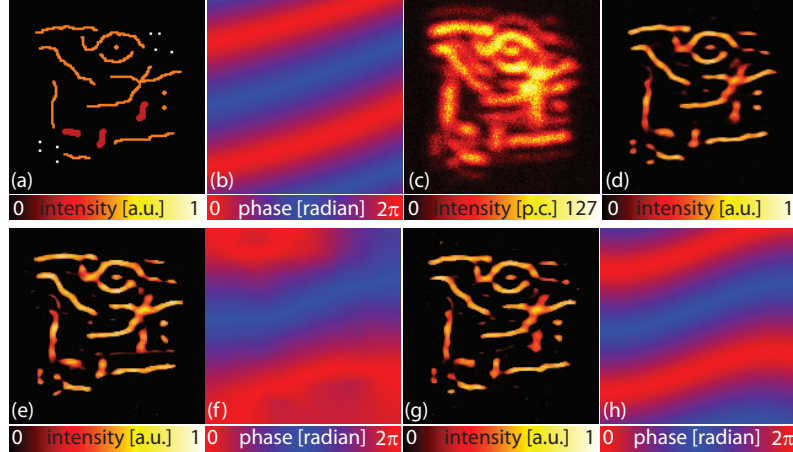


Fig. 1. Image restoration on simulated data (qualitative results): (a) phantom (136x136 pixels, pixel size 20x20nm in both lateral and axial direction); (b) cubic RPF; (c) simulated image obtained using the following parameters: objective NA 1.46 oil (RI 1.515), excitation wavelength $\lambda_{ex} = 632$ nm, emission wavelength $\lambda_{em} = 680$ nm, pinhole 0.5 AU, constant background of 2 photon counts and expected photon counts of 100 in the bright pixel; (d) restored object (the RPF is known *a-priori*); (e) restored object and (f) reconstructed RPF for FPBR; (g) restored object (h) and reconstructed RPF for SPBR.

which is the best measure in the presence of a non-negativity constraint [23]. The quality of the reconstructed RPFs was assessed using a modified Euclidean distance (D_E) between the original RPF ϕ and the estimate ϕ_E

$$D_E(\phi, \phi_E) = \|\exp(i\mathbf{T}\phi_E) - \exp(i\mathbf{T}\phi)\|_2, \quad (27)$$

where the exponential takes into account for the phase wrapping and \mathbf{T} is a mask setting all voxels to zero where the estimated object \mathbf{f}_E is below a very low threshold of 10^{-6} of its maximum. This ensures that only regions with non-negligible contribution to the image formation are used for the calculation of D_E .

The proposed algorithms were implemented in MATLAB. For comparison, we first tested the performance of the SGM algorithm separately from the AMM algorithm, i.e. we assumed that the RPF is known *a-priori* and we estimated only the object. Figures 1(a)–1(d) shows the results for $\mu_f = 10^{-5}$ and demonstrates the well known ability of the SGM algorithm to generate artefact-free estimate of the object and to compensate for phase variation. Moreover, noise was suppressed and structures are well preserved. We then compared these results to the estimates obtained by using the proposed AMM algorithms to simultaneously reconstruct the RPF and the object. Keeping the same μ_f we used $\mu_\phi = 50$ for FPBR and $\mu_\phi = 10^{-3}$, $N_\phi = 4$ for SPBR. Figures 1(e)–1(h), and 2 compare the results obtained using FPBR and SPBR: Both approaches generated suitable restored objects and reliably reconstructed RPFs. However, an accurate visual inspection of the restored objects spotlights a slightly superior performance of the SPBR respect to FPBR. Interference artifacts that were partially removed by FPBR were completely removed by SPBR and consequently this approach also reaches better values for the Kullback-Leibler divergence D_{KL} and for D_E . Because the cost function is convex in \mathbf{f} for a given RPF the difference between the two results is rooted in their different ability to approach the global optimum for the RPF. Since we had assumed a polynomial RPF the superior performance of SPBR is not surprising given the fact that the number of free parameters is much lower and the model allows for a perfect representation of the RPF using the basis functions. Please also note that the algorithms cannot be expected to yield a good estimate of the RPF in regions far away from the object structures; a fact also reflected in the calculation of D_E .

An important aspect is the difference of convergence rates for the two methods. As seen from Fig. 2 the reduction of free parameters in the SPBR approach leads to convergence after less cycles of the AMM as expected. For large data sets the calculation of the updated object will usually dominate due to the $\mathcal{O}(N \log(N))$ complexity of the FFTs and therefore this advantage

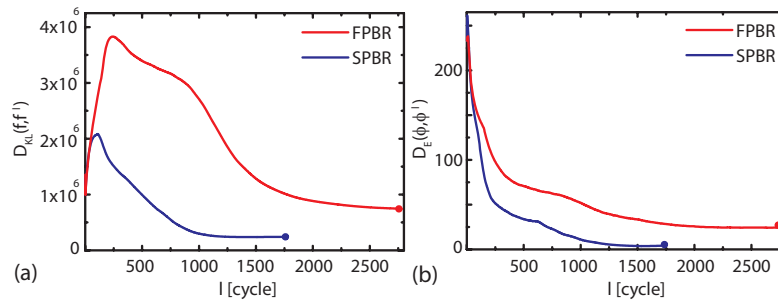


Fig. 2. Image restoration on simulated data (quantitative results): (a) Kullback-Leibler divergences between test phantom and reconstructed object as a function of the number of cycles for FPBR vs. SPBR; (b) Euclidean distances between reconstructed and assumed RPF as a function of the number of cycles for FPBR and SPBR.

directly translates into faster convergence in terms of calculation time. In our case the SPBR approach was usually twice as fast. However we note that since the two approaches implement different algorithms for the update of the RPF with different regularization parameters and stopping criteria, their speed is co-determined by the choice of parameters and the time needed for each RPF calculation. The latter is of complexity $\mathcal{O}(N)$ for APBR and $\mathcal{O}(N_\psi N)$ for SBPR. For smaller images and/or (much) larger dictionaries this advantage will therefore become smaller and eventually vanish.

We also note that the local quality of the restoration may slightly vary with the RPF, an effect rooted in the dependence of the frequency content of the PSF on the local phase. This very small effect can be completely eliminated by obtaining several images with different relative phases [5].

6. Validation on experimental data

In order to test our method on real data we employed a standard beam scanning confocal microscope (Leica Microsystems, Mannheim, Germany) to which we attached a custom build 4Pi

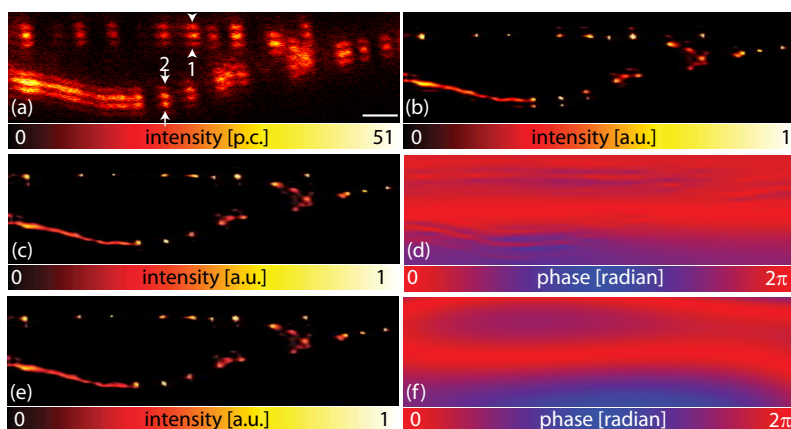


Fig. 3. Image restoration on experimental data: (a) raw xz image of microtubules wrapping around the nucleus in a Vero cell (186×352 pixels); (b) deconvolved data assuming constant RPF; (c) restored object and (d) reconstructed RPF for FPBR; (e) restored object and (f) reconstructed RPF for SPBR. Scale bar $1\mu\text{m}$.

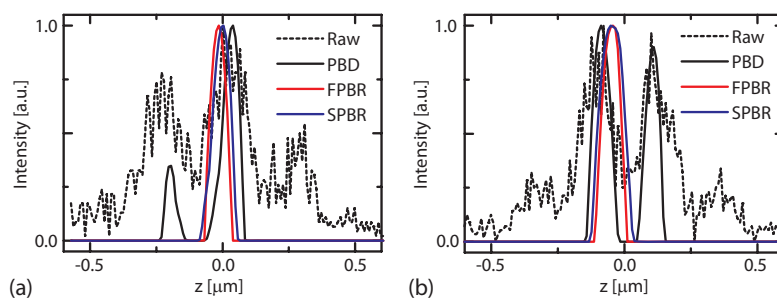


Fig. 4. Intensity profiles of the restored objects taken at the positions indicated by arrows in Fig. 3(a). (a) intensity profile at site 1; (b) intensity profile at site 2.

module to mount two opposing objective lenses (HCX PL APO 100 \times , 1.46 NA OIL CORR; Leica Microsystems). A driven-piezo mirror controls the difference in pathlength between both cavity arms. Three-dimensional imaging was performed on *Vero* cells. Cells were immunolabelled with a monoclonal mouse antiserum directed against the alpha-tubulin (Sigma-Aldrich, St. Louis, MO, USA), the building-block of microtubules. The primary antibodies were detected with secondary antibodies (sheep anti-mouse; Jackson ImmunoResearch Laboratories, West Grove; PA; USA) custom labelled with the fluorophore KK114. Samples were mounted in the TDE (2,2'-Thiodiethanol, Sigma-Aldrich, St. Louis, MO) water-miscible mounting medium, in order to adjust the refractive index of the sample to ~ 1.5 [9]. To provide a more challenging test situation, index matching was not performed completely, so that a mismatch between nucleus and medium remained. Additionally, a linear phase shift along z was introduced by changing the relative length of the two cavity arms while scanning axially. Parameters were chosen such that constructive interference on the top of the xz slice [Fig. 3(a), region of interest 1] and destructive interference on the bottom [Fig. 3(a), region of interest 2] were observed, rendering the RPF a combination of externally-induced and sample-induced phase shifts. To obtain a genuine 3D structure with a thickness of a few micrometers, imaging of microtubules wrapping the nucleus was performed [Fig. 3(a)]. Fluorescence after one-photon excitation of the fluorophore KK114 at 635nm was detected over the wavelength range from 660nm to 700nm by a photon-counting avalanche photodiode (APD) (PerkinElmer, Waltham, MA, USA). The confocal pinhole was set to 0.5 AU. We started by validating the method on a single 2D axial image. We used $\mu_f = 10^{-4}$, $\mu_\phi = 25$ for FPBR and $\mu_\phi = 10^{-3}$, $N_\phi = 3$ for SPBR. A constant background of 2 counts was estimated directly from the images.

Figure 3(b) shows the results obtained using the parametric blind deconvolution (PBD) assuming a single, constant phase as devised in our earlier work [5]. Since the algorithm cannot account for the varying phase, it is not able to completely remove the interference artifacts. On the other hand, both PBR methods generate interference artifact-free object estimates. As expected, the reconstructed RPFs show a change from destructive to constructive interference along the axial coordinate. To better appreciate the ability of the method to remove sidelobes Fig. 4 shows intensity profiles from the sites indicated by arrows in Fig. 3. The artifacts remaining after constant phase PBR are clearly removed by both variable phase PBR approaches. While visual inspection shows some difference between FPBR and SPBR, most prominently in their reconstructed RPFs, the differences in the restored objects are not significant. As before, the main difference between the two methods is the rate of convergence. While SPBR requires ~ 190 cycles to reach the stopping criterion, FPBR requires ~ 450 cycles and is 3 times slower. We also performed SPBR with larger dictionaries ($N_\phi \leq 8$) but did not observe significant improvements.

Finally, we applied our methods to 3D data-stacks. Since SBPR proved to converge faster and yielded comparable or better results in our previous analysis we restricted ourselves to this method. We also used a simple trick to speed up reconstruction in the more complex 3D case. We first applied the SBPR algorithm to a dense xz slice of the 3D data-stack using $\epsilon_{\text{AMM}} = 10^{-5}$ and used the resulting coefficient vector ψ to initialize the 3D reconstruction Fig. 5 shows the raw data and the restored object along with the reconstructed RPF. Importantly, also on experimental data, the AMM algorithm converges to appealing results even though the cost function is not convex in ϕ or ψ .

7. Discussion and conclusion

While they feature the highest axial resolution in diffraction limited imaging, microscopes that use the nearly spherical wavefront afforded by two opposing lenses usually suffer from sidelobe artifacts in their raw image data. Reliable image restoration is therefore mandatory for their

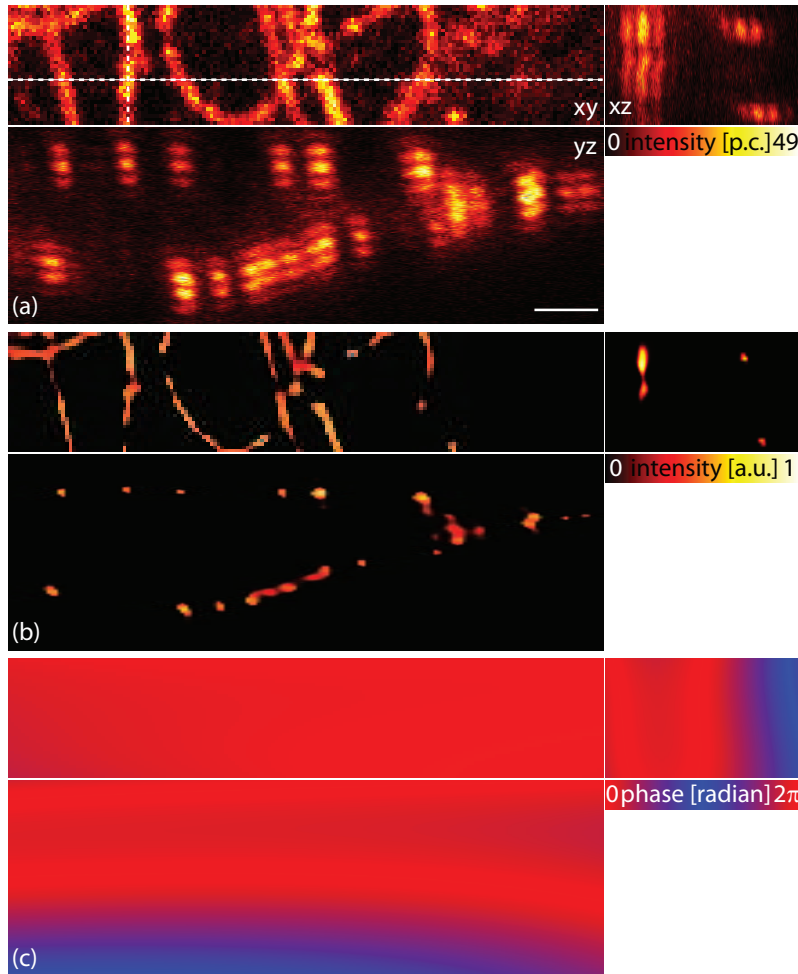


Fig. 5. 3D Parametric blind restoration: (a) raw data ($228 \times 334 \times 40$ voxels); (b) restored object and (c) reconstructed RPF for SPBR. The dotted lines indicate the position of the xz- and yz-slices shown for each of the stacks. Scale bar $1 \mu\text{m}$.

routine application. In 4Pi-microscopy imaging conditions can often be tuned such that the PSF is space-invariant or its spatial dependence is known, allowing for the application of presently available techniques. However, such adjustments may be a major obstacle or even impossible in many real-life applications. To date the data acquired in such experiments could either not be used at all or had to be reconstructed by trying to manually estimate the RPF and repeatedly performing deconvolutions. Usually the most appealing recovery will then be chosen for further analysis introducing a subjective component into the interpretation of the results. We have addressed this problem by developing a parametric blind restoration method based on a maximum a-posteriori approach which uses as much of the available *a-priori* information about the noise, the object and the RPF as possible. While our method is not guaranteed to converge to a global optimum due to the non-convex nature of the resulting cost function our tests have shown that the algorithm gracefully converges to visual appealing restored objects. By assuming a phase function represented by a sparse linear combination of polynomials rather than an arbitrary one

we were able to speed up convergence while yielding similar or better results, both on experimental and simulated data. The method proposed here might be improved or accelerated by the use of new minimization algorithms, not using the AMM approach or by initializing the RPF with a more educated guess before starting the iteration. The problem of possible convergence to local minima could be addressed by systematically re-starting the algorithm with different RPFs or by implementing a scheme of simulated annealing. Also, more suitable penalization functions or dictionaries for the relative phase function could emerge. One could also imagine extending the 4Pi microscope to simultaneously map fluorescence and changes in the refractive index, e.g. by detecting differential interference contrast. Using an AMM approach very similar to the one presented here, this additional data could help to significantly improve the reconstruction of the RPF e.g. by imposing local restraints on its derivative. However, even in its present implementation the algorithm performs well on experimental 4Pi data and should be an important step towards more routine application of this method also to live cells by making data acquisition less demanding and faster. Finally a similar approach might prove useful when deconvolving raw data from other microscopy modalities when their E-PSF is position dependent, e.g. in the presence of sample induced aberrations. To this end a decomposition of the forward operator could be based on approximation of the E-PSF by a Fourier series to enable fast re-calculation with various forms of the PSF.

Appendix

Split-gradient-method (SGM) algorithm

Given ϕ , \mathbf{g} and \mathbf{b} we will derive an iterative algorithm to reconstruct the object function \mathbf{f} . We decompose the gradient of the cost function $J(\mathbf{f}, \phi | \mathbf{g})$ with respect to \mathbf{f} as follows

$$-\nabla_{\mathbf{f}}[J_o(\mathbf{f}, \phi | \mathbf{g})](\mathbf{f}) = U_o(\mathbf{f} | \phi, \mathbf{g}) - V_o(\mathbf{f} | \phi, \mathbf{g}) \quad (28)$$

$$-\nabla_{\mathbf{f}}[J_f(\mathbf{f})](\mathbf{f}) = U_f(\mathbf{f}) - V_f(\mathbf{f}), \quad (29)$$

where $U_o(\mathbf{f}; \phi, \mathbf{g})$ and $V_o(\mathbf{f}; \phi, \mathbf{g})$ are positive vectors depending on \mathbf{f} and $U_f(\mathbf{f})$ and $V_f(\mathbf{f})$ are non-negative vectors depending on \mathbf{f} . It is obvious that such a decomposition always exists but is not unique. The general form of an iteration step in the multiplicative form, as described in [22], is then

$$\mathbf{f}_{i+1} = \mathbf{f}_i \left(\frac{U_o(\mathbf{f}_i | \phi, \mathbf{g}) + \mu_r U_f(\mathbf{f}_i)}{V_o(\mathbf{f}_i | \phi, \mathbf{g}) + \mu_r V_f(\mathbf{f}_i)} \right). \quad (30)$$

It is important to remark that all the iterates are automatically non-negative if the initial guess \mathbf{f}_0 is non-negative. Now, from Eq. (13) we obtain

$$\nabla_{\mathbf{f}}[J_o(\mathbf{f}, \phi | \mathbf{g})](\mathbf{f}) = \mathbf{H}(\phi)^T \mathbf{1} - \mathbf{H}(\phi)^T \mathbf{g} [\mathbf{H}(\phi) \mathbf{f} + \mathbf{b}]^{-1}, \quad (31)$$

where $\mathbf{H}(\phi)^T$ is the transpose of the forward operator $\mathbf{H}(\phi)$ and $\mathbf{1}$ is the vector whose entries are all equal to 1. Since all elements of $\mathbf{H}(\phi)$ are non-negative for non-negative \mathbf{f} , a natural choice for U_o and V_o is

$$U_o(\mathbf{f} | \phi, \mathbf{g}) = \mathbf{H}(\phi)^T \frac{\mathbf{g}}{\mathbf{H}(\phi) \mathbf{f} + \mathbf{b}} \quad (32)$$

$$V_o(\mathbf{f} | \phi, \mathbf{g}) = \mathbf{H}(\phi)^T \mathbf{1}. \quad (33)$$

The gradient with respect to \mathbf{f} of the Tikhonov penalty term $J_f(\mathbf{f}) = \|\mathbf{f}\|_2^2$ is given by

$$\nabla_{\mathbf{f}}[J_f(\mathbf{f})](\mathbf{f}) = \mathbf{f}, \quad (34)$$

and therefore, we can choose the decomposition $U_f(\mathbf{f}) = \mathbf{0}$ $V_f(\mathbf{f}) = \mathbf{f}$. By substitution of the gradients decompositions above into Eq. (30) we obtain Eq. (18).

Quasi-Newton-method (QNM) algorithm

Given \mathbf{f} , \mathbf{g} and \mathbf{b} the QNM method is used to update the relative phase function ϕ directly (FPBR) or through the component vector ψ (SPBR). In both cases the evaluation of the descent direction [see Eq. (20)] requires computation of the gradient of the cost function. In FPBR, the gradient of the cost function respect to ϕ is given by

$$\nabla_{\phi} [J(\mathbf{f}, \phi; \mathbf{g})](\phi) = \mathbf{H}'(\phi)\mathbf{f} - \frac{\mathbf{g}}{\mathbf{H}(\phi)\mathbf{f} + \mathbf{b}} \mathbf{H}'(\phi)\mathbf{f} + \mu_{\phi} \mathbf{L}\phi, \quad (35)$$

where the derivative of the forward operator, $\mathbf{H}'(\phi)\mathbf{f} = \nabla_{\phi} [\mathbf{H}(\phi)\mathbf{f}](\phi)$ is given by Eq. (5). The operator \mathbf{L} is given by

$$(\mathbf{L}\phi)(\mathbf{n}) = \sum_{\mathbf{m} \in \eta_{\mathbf{n}}} \{\phi(\mathbf{m}) - \phi(\mathbf{n})\}, \quad (36)$$

and $\eta_{\mathbf{n}} = \{(n_x \pm 1, n_y, n_z), (n_x, n_y \pm 1, n_z), (n_x, n_y, n_z \pm 1)\}$ is the set of direct neighbor voxels of \mathbf{n} .

In SPBR the gradient of the cost function with respect to ψ can also be directly calculated. For the p th component we obtain

$$\partial_{\psi_p} [J(\mathbf{f}, \mathbf{D}\psi; \mathbf{g})] = \sum_{\mathbf{n}} \left[\mathbf{H}'_p(\mathbf{D}\psi)\mathbf{f} - \frac{\mathbf{g}}{\mathbf{H}(\mathbf{D}\psi)\mathbf{f} + \mathbf{b}} \mathbf{H}'_p(\mathbf{D}\psi)\mathbf{f} \right](\mathbf{n}) + \mu_{\psi} \psi_p / [\psi_p^2 + \delta]^{1/2} \quad (37)$$

with

$$\mathbf{H}'_p(\mathbf{D}\psi)\mathbf{f} = \partial_{\psi_p} [\mathbf{H}(\mathbf{D}\psi)\mathbf{f}]. \quad (38)$$

In Eq. (37) the approximation $\|\psi\|_1 \simeq \sum_p (\psi_p^2 + \delta)^{1/2}$ is used for removing the singularity of the derivative at $\psi_p = 0$. δ can be set to the precision of the machine. Using the decomposition of \mathbf{H} , its component-wise derivative with respect to ψ_p is given by

$$\mathbf{H}'_p(\mathbf{D}\psi)\mathbf{f} = -\sin(\mathbf{D}\psi)\phi_p \mathbf{A}_1 \mathbf{f} - \cos(\mathbf{D}\psi)\phi_p \mathbf{A}_2 \mathbf{f}. \quad (39)$$

in analogy to Eq. (5).

Acknowledgment

We thank Johann Engelhardt, Marcel Leutenegger, Dorothea Hahn, Robert Stück and Thorsten Hohage for useful discussions, Kirill Kolmakov and Vladimir Belov for providing us with the new fluorescent dye KK114, Christian Wurm for preparing the biological sample. We also thank Jan Keller for providing the MATLAB code of electric field computation. This work has been supported by the German Federal Ministry of Education and Research through the project INVERS.