Optimizing 3D multiphoton fluorescence microscopy

Ido Kaminer,† Jonathan Nemirovsky,† and Mordechai Segev*

Physics Department and Solid State Institute, Technion, Haifa 32000, Israel
**Corresponding author: msegev@tx.technion.ac.il*

Received July 2, 2013; revised August 24, 2013; accepted August 25, 2013;
posted August 28, 2013 (Doc. ID 192757); published September 30, 2013

We present a new optimization concept for 3D multiphoton fluorescence microscopy by finding the optimal excitation beam giving rise to the smallest possible light-emitting volume or the highest possible signal to noise ratio (SNR). © 2013 Optical Society of America

OCIS codes: (190.0190) Nonlinear optics; (190.4180) Multiphoton processes; (180.2520) Fluorescence microscopy; (180.4315) Nonlinear microscopy; (180.6900) Three-dimensional microscopy.
http://dx.doi.org/10.1364/OL.38.003945

Scanning multiphoton fluorescence microscopy (MFM) is a powerful tool used for in-depth, high resolution three-dimensional (3D) sectioning [1]. In life sciences, current techniques allow a noninvasive, *in vivo*, laser microscopy, without staining, with sub-wavelength 3D resolution and millimeters penetration [1,2]. In addition, MFM is also used for nondestructive inspection of various products and specimens (see, e.g., [3,4]), rewritable high density 3D optical storage [5,6], and 3D multiphoton lithography [7]. In MFM, images are formed by scanning samples point-by-point with a focused laser beam ("pump"), giving rise to multiphoton absorption, which populates a higher energy level from which light fluorescence occurs. While two-photon excitation microscopy is perhaps most commonly used [1,5–7], the three-photon case is widely used as well [2–4]. In both cases, higher frequency fluorescence induced by the laser light is detected and collected to form a digital 3D image. The basic advantages of these nonlinear techniques over the linear techniques are associated with single-photon absorption are two-fold: (1) the absence of linear absorption allows much larger penetration depth into the specimen. (2) MFM facilitates sectioning: one can scan the specimen plane by plane. The reason is that the multiphoton absorption coefficient depends on the light intensity: when the laser beam has its minimum width at some spot in a given plane—multiphoton fluorescence at that individual spot are much stronger than from anywhere else in the volume. Moreover, since several photons combine to create a single photon, the effective excitation profile, and with it the fluorescent region around that spot, are narrower than the intensity profile of the laser beam. Yet, even so, the resolution of the 3D image is always limited by the volume occupied by laser spot size at its focus. Many attempts have been made to find the smallest possible laser spot size. Most of these used the paraxial approximation. Especially important was [8], optimizing the spot size of single photon absorption. Of course, in MFM, multiphoton processes have a smaller effective spot size, due to their dependence on nonlinear intensity. Indeed, paraxial resolution estimates for MFM were performed for Gaussian beams, see, e.g., [9,10]. But the smallest spot size obtained in modern microscopes is never limited by paraxiality. Clearly, methodical optimization of a specifically engineered beam (amplitude and phase) could yield a substantial additional reduction of the effective nonlinear spot size, and a reduction in the nonlinear volume underlying the 3D resolution of MFM.

Here, we introduce a new approach for optimizing both the nonlinear effective 3D volume for MFM and the signal to noise ratio (SNR). Our approach is based on precise solutions of the Helmholtz equation (no paraxial approximation), accounting for the nonlinear effects of multiphoton absorption. The effective volume and the total fluorescence noise are quantified by a tailored figure of merit (FOM) which naturally includes all the "out of focus" signal contributions. By including these, the advantage of the MFM technique is fully revealed. We demonstrate this approach for two-photon and three-photon microscopy, and test it for various values of numerical aperture (NA). In all cases, we compare our results to the best case Gaussian beam with the same NA. We find a specific design of a single laser beam undergoing multiphoton absorption, for which the SNR is improved by 45%. Yet, the full strength of our algorithm is illustrated when we find a specific beam that reduces the on-axis secondary peaks by a factor of 200, which increases the on-axis SNR by a factor of 70. Likewise, for three-photon microscopy, our approach yields an illumination volume 2.5 times smaller than the best Gaussian beam of unity NA. Finally, we discuss the next steps needed to make our new approach applicable for the full EM problem described by the Maxwell equation.

Given an optical field $E(x,y,z)$ obeying the Helmholtz equation, we define the following family of FOMs

$$S = \frac{\iiint f(x,y,z)|E(x,y,z)|^{2n}\mathrm{d}V}{\int |E(x,y,z)|^{2n}\mathrm{d}V}. \tag{1}$$

The choice of $f(x,y,z)$ defines what parameter we wish to optimize. For example, $f(x,y,z)$ can be a 3D step function ("window") in case we wish to optimize the overall SNR. Alternatively, we can choose $f = x^2 + y^2 + z^2$ for optimizing the size of the illumination volume, since this FOM is the variance of the light intensity. In all of the above, the FOM is calculated for the fluorescence light emitted following n-photon excitation—this is why the intensity is taken with a power of $2n$. To find the optimal beam, we minimize S under the constraint that $E(x,y,z)$ satisfies the Helmholtz equation—which, assuming sources relying on propagating (nonevanescent) waves, amounts to requiring the Fourier transform of $E(x,y,z)$

to reside on the surface of a sphere of radius $k = \omega/c$ in k-space. For a single forward-propagating beam, the solution is confined (in k-space) to a half-sphere. Additional constraints may be included depending on specific design. Importantly, a threshold exists at $n > 2$ (for $n = 2$ we suggest to take $f = \sqrt{(x^2 + y^2 + z^2)}$ [11]). In what follows, we demonstrate our method by presenting three cases of optimization: of the total SNR, of the on-axis SNR for $n = 2$, and of the illumination volume for $n = 3$.

In most cases, the optical NA is restricted to some value. This is incorporated into our method by restricting the field in Fourier space to reside within a certain spatial angle which translates to a slice of the k-sphere. For NA = 1 (assuming air as background material) the field is restricted to the k-hemisphere in Fourier space. Hence, we optimize the beam shape by starting from the convenient initial choice of a half-sphere of equal distribution, use the FOM [Eq. (1)] to quantify the S parameter, and then find the optimal improvement iteratively through a numerical procedure. To do that, we introduce a new algorithm for the optimization of the spot size. This numerical method combines a variational approach and an infinite-dimensional gradient descent, which is nevertheless computationally easy. The procedure is demonstrated in Fig. 1 for the simplest case of no additional constraints. To explain the optimization procedure, first note that any angularly symmetric beam can be described by an integral expression for the electric field (our initial choice corresponds to $w = 1$)

$$E(r,z) = \int_0^{\pi/2} w(k_\theta) e^{ikz\cos(k_\theta)} J_0(kr\sin(k_\theta))\sin(k_\theta)\mathrm{d}k_\theta. \tag{2}$$

This defines the field for NA = 1 (in air); the only change needed for a smaller NA is changing the $\pi/2$ into a smaller angle. One can directly extend this into materials with permittivity different than 1. Note that the expression in Eq. (2) is an exact solution of the (nonparaxial) Helmholtz equation for any choice of the function w. For the optimization procedure, we substitute Eq. (2) into Eq. (1) and add an infinitesimally small variation in w, expanding S to first-order in w. Finally, we define the optimal variation to be the one that has the strongest gradient in S, i.e., decrease (increase) the spot volume (the SNR) as much as possible. One can prove that this happens when

$$\Delta w(k_\theta) = \iint E|E|^{2n-2} J_0(kr\sin(k_\theta)) e^{-ikz\cos(k_\theta)} \times (f(r,z) - S) r\mathrm{d}r\mathrm{d}z. \tag{3}$$

Here lies the strength of our algorithm: a single calculation of Eq. (3) is sufficient to update the entire field, adding the best possible variation, even though this variation is found in an infinite-dimension phase space. The algorithm has three steps: (1) calculate the current parameter S with the FOM defined in Eq. (1), (2) use it in Eq. (3) to calculate Δw, and (3) update w by adding Δw multiplied by a small number which sets the "step size," and then normalize the beam power. It is convenient to begin with an initial uniform distribution ($w = 1$ for all k_θ). The condition for stopping the algorithm is not a zero gradient, but a gradient which is approximately proportional to the value of w in the same iteration. To gain some intuition about the function we optimize, w, note that it is related to the Fourier structure of the excitation beam. One can design a setup where the spot plane ($z = 0$) is in the focal plane of a high NA lens. The required mask at the focal plane is azimuthally symmetric and only depends on the radius k_r as $\mathrm{mask}(k_r) = w(\arcsin(k_r/k)/\sqrt{(k^2 - k_r^2)})$.

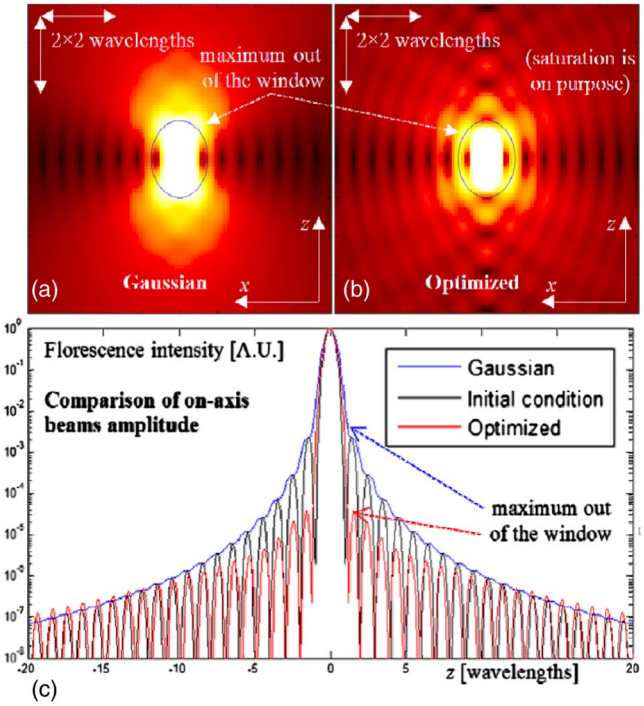

Fig. 1. Demonstration of SNR optimizations for $n = 2$ and NA = 1. (a) and (b) Amplitudes of the MFM spot of an optimal Gaussian beam and the optimized beam found by our algorithm, respectively. The beams are normalized to have the same "signal" strength, meaning that the numerator of Eq. (1) is the same. Note that the side lobes of the optimized beam are smaller, mainly along the z axis. The blue ellipses mark the edge of the FOM window, with diameters of 2×1.5 wavelengths. The maximum fluorescence intensity outside the window (marked by dashed arrows) is reduced by 54%. The figures present the normalized intensity, and the axes are measured in units of the optical wavelength. The axes are z (vertical) and x (horizontal). (c) Comparison between the on-axis MFM side lobes of the optimal Gaussian beam (blue), the beam of equal angular distribution (black), and our optimized solution (red). The plots are normalized to have the same maximum peak amplitude. The maximum fluorescence intensity outside the window (marked by dashed arrows) is reduced by a factor of 200!

Our first example is the optimization of the SNR in a two-photon fluorescence process ($n = 2$) for NA = 1, and thus we aim to decrease the unwanted side-lobes. We take the FOM to be an elliptic window of diameter 2×1.5 wavelengths, so that f equals 1 inside the window and zero outside. This way, S measures the signal strength (numerator) divided by the total fluorescence intensity (denominator). The SNR equals $S/(1 - S)$, and since S is between 0 and 1, it turns out that maximizing S also

maximizes the SNR. Our optimization result is presented in Figs. 1(a) and 1(b), showing an SNR improvement of 45%: from 4.49 in the best Gaussian case to 6.54 in the optimized case. It is important to note that our initial condition of equal distribution already achieves most of this improvement, by having an SNR of 6.05.

Our second example optimizes the SNR in sectioning microscopy. To do that, we change the FOM to include only on-axis fluorescence light, so that we improve the spot along the z axis by minimizing the on-axis secondary peaks, which usually appear both before and after the main peak. This is done by changing the integration of Eq. (1) to be a 1D integral along z at $r = 0$, and choosing f to be a window function with a width of two wavelengths centered around the highest peak ($z = 0$). The SNR is calculated as before. The result of this optimization is presented in Fig. 1(c), demonstrating an improvement of a factor of 70. In this example, the SNR for the natural choice of a Gaussian beam is ~101, whereas the best SNR found through our procedure is 7007. Of course, a better fit of the window size can further improve the results. Unlike the previous examples, here the initial conditions of equal angular distribution in k-space are far from optimum, having an SNR of only 293. This emphasizes that our procedure is successful even when the initial conditions are far from the best result. Consequently, while for some FOMs the performances of the Gaussian beam can only be slightly improved, other FOMs are significantly improved, highlighting the fact that the beam of equal distribution is not only suboptimal, but also sometimes very far from optimum.

In our last example, the goal is to optimize the illumination volume for three-photon microscopy. Hence, we take $f = x^2 + y^2 + z^2$, which yields the illumination spot volume to be smaller than the best Gaussian beam by a factor of 2.3 (In nonparaxial optics the limit case of a Gaussian with the thinnest waist is actually an Airy-disk beam [12,13].). Compare the volumes Fig. 2(a) (the Gaussian beam) and Fig. 2(b) (the optimal beam). As before, the initial condition for our algorithm is taken as a uniform k-space distribution beam, which in this case is already much better than the natural choice of a Gaussian beam. In fact, the optimized illumination volume is smaller by only 7%, with respect to the choice of initial condition. For a smaller NA values, the improvement (with respect to the initial choice) made by our algorithm can go beyond 20%.

The algorithm is illustrated in Fig. 2(c), which displays w as a function of k_θ for several iterations, for the results presented in Figs. 1(a) and 1(b). It is easily seen that the optimal k-space distribution is very different than the initial choice of uniform distribution (red line). In particular, the field at steeper angles (close to the limit of $\pi/2$ for NA = 1) should be stronger to achieve a smaller spot. We find a similar trend favoring the steep angles (for the field) for any value of NA, and for the other FOMs above. Figure 2(d) shows the difference between the amplitudes of the optimal beam [Fig. 2(b)] and the beam of uniform k-space distribution. The orange part around the center emphasizes that the spot size is indeed smaller, but this comes at the expense of increasing the adjacent peaks on the z axis (yellow spots above and below the center). This clarifies the physical limitation on the improvement in the spot size: there is a trade-off between squeezing the main spot and increasing the side-lobes. The result is of course much more impressive when compared to Gaussian beams. Nevertheless, we choose to compare the result to beams of uniform k-space distribution since it highlights the counterintuitive result: against expectations, the beam of uniform k-space distribution does not give the optimal spot size (or the optimal SNR) for any value of NA in MFM (expect for a perfect 4Pi setup).

It is important to also discuss the limits of our method. First, the procedure stops at the first local extremum and while we conjecture this is also the global minimum—so far we were not able to prove it. This emerges from the

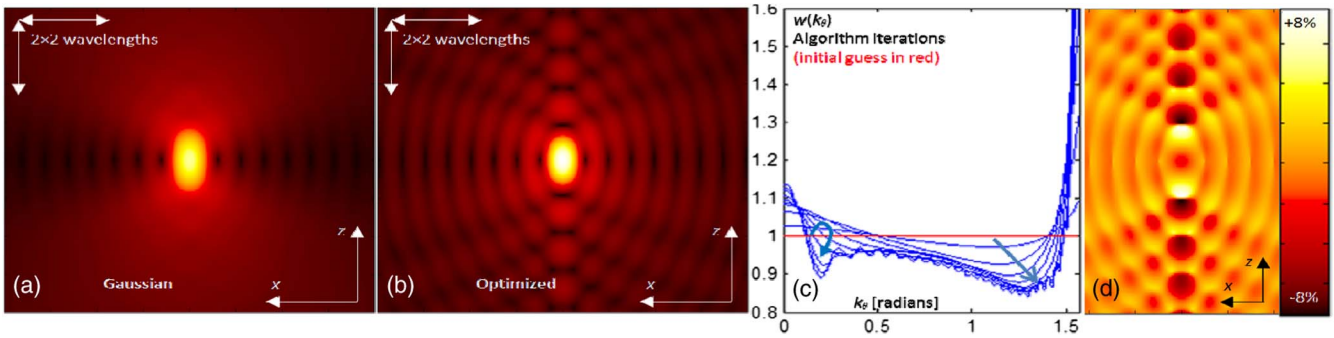

Fig. 2. Demonstration of the procedure for minimizing the illumination volume, for $n = 3$ and NA = 1. (a) and (b) Amplitudes of the MFM spot of an optimal Gaussian beam and the optimized beam found by our algorithm, respectively. The beams are normalized so that the numerator of Eq. (1) is the same. The two subfigures have the same color scale to emphasize that the optimized illumination spot is smaller and more intense. The volumes are calculated as $S^{3/2}$ and measured in wavelengths. (c) Several iterations in the optimization process, showing the angular k-space distribution $w(k_\theta)$ in subsequent steps in the algorithm (plotting every 2nd step until the procedure converges). The initial distribution is uniform (red line), with later steps having a steeper curve around the right boundary of $\pi/2$. (d) The difference between amplitude of the initial condition (uniform angular k-space distribution) and amplitude of the optimal spot. Positive (bright) values are where the optimized beam has smaller amplitude, plotted in % of the maximum amplitudes, and in the same length scales of the two left subfigures. This highlights pros and cons of the optimized solution: the width of the optimized beam is narrower mostly along the z axis, but the peak intensity is slightly smaller and it contains side-lobes.

fact that our functional is not convex for any of our FOMs. Second, we do not know whether the optimization procedure is always stable. The evolution of $w(k_\theta)$ around the desired extremum might strongly depend on the exact initial conditions for some FOMs. To test our procedure numerically, we add ~20% white noise to the initial w, and find that, for the above mentioned FOMs, the optimization still converges to similar results. The noise is even slowly decreasing, emphasizing the stability of the procedure. This does not promise stability for other situations, yet the stability is not crucial for the algorithm to converge to an improved beam, and might even improve performances by dodging local extremum points. For example, the factor of improvement for the illumination spot minimization (Fig. 1) is increased from 2.3 to more than 2.5, although adding 20% noise only makes the initial conditions worse, and the factor of improvement for the on-axis SNR is increased to 73 (final SNR of 7430), although the SNR in the initial condition is now 269 instead of 293. In other words, the performance of our algorithm improves when the initial conditions include higher noise (this finding could possibly be related to the phenomenon of stochastic resonance [14]). Most importantly, having an unstable optimization procedure does not reflect on the stability of the derived beam. On the contrary, the dynamics of the light in MFM systems is always stable since the nonlinear fluorescence light has a negligible effect on the fundamental harmonic that has created it.

To make our algorithm applicable to the state-of-the-art microscopy techniques, two important developments are still needed. First, a FOM that optimizes both the illumination spot size and the SNR should be designed. The simplest idea is to take a linear combination of the variance ($f = x^2 + y^2 + z^2$) and the inside-out window ($1 - f$). However, better designs might be possible. This calls for further study with extensive numerical testing. Second, our current algorithm handles only scalar fields, i.e., solutions of the Helmholtz equation. One should really optimize the vector field obeying Maxwell's equations: the field should be expressed via a 3D vector $\mathbf{w}$, and therefore S should be written as a function of $\mathbf{w}$, while also accounting for the tensor nature of the nonlinearity. Then, a variation approach should be used to extract the vector $\Delta\mathbf{w}$ and update $\mathbf{w}$ in each step. This way, the optimization would be performed on a vector field and the iterative algorithm would involve the whole vector being updated in parallel. While this adds some computing steps, the efficiency of the optimization process is of the same order.

Looking ahead, our method is general and is not limited to the parameter constraints of the present system: it could be generalized to different types of nonlinear processes, and be incorporated in leading experimental techniques as the STED [15,16], or be used for optimization of the family of abruptly autofocusing waves [17–19]. Moreover, the same approach would let us optimize spatiotemporal fields. For example, optimization of time-dependent EM fields, composed of several frequencies, would be achieved by confining the k-space field to a sphere of a width given by the frequency range. Finally, utilizing the polarization of the EM field is known to significantly improve the resolution of linear microscopy [20,21]. Our optimization algorithm can be applied for the full Maxwell equations in time and space, thus giving the optimal space–time volume for MFM.

This work was supported by an Advanced Grant from the European Research Council (ERC) and by the Israeli Focal Technology Area (FTA) on Nanophotonics for Detection.

†Authors with equal contribution.

References

1. W. Denk, J. Strickler, and W. Webb, Science **248**, 73 (1990).
2. Y. Barad, H. Eisenberg, M. Horowitz, and Y. Silberberg, Appl. Phys. Lett. **70**, 922 (1997).
3. S. Maiti, J. B. Shear, R. M. Williams, W. R. Zipfel, and W. W. Webb, Science **275**, 530 (1997).
4. D. Yelin and Y. Silberberg, Opt. Express **5**, 169 (1999).
5. D. A. Parthenopoulos and P. M. Rentzepis, Science **245**, 843 (1989).
6. B. H. Cumpston, S. P. Ananthavel, S. Barlow, D. L. Dyer, J. E. Ehrlich, L. L. Erskine, A. A. Heikal, S. M. Kuebler, I.-Y. Sandy Lee, D. McCord-Maughon, J. Qin, H. Röckel, M. Rumi, X.-L. Wu, S. R. Marder, and J. W. Perry, Nature **398**, 51 (1999).
7. W. Haske, V. W. Chen, J. M. Hales, W. Dong, S. Barlow, S. R. Marder, and J. W. Perry, Opt. Express **15**, 3426 (2007).
8. T. R. M. Sales, Phys. Rev. Lett. **81**, 3844 (1998).
9. M. Gu, Opt. Lett. **21**, 988 (1996).
10. J. Squier and M. Muller, Rev. Sci. Inst. **72**, 2855 (2001).
11. https://www.researchgate.net/publication/256120919_Supplementary_material_for_the_paper_Optimizing_3D_multiphoton_fluorescence_microscopy?ev=srch_pub.
12. G. B. Airy, Trans. Cambridge Philos. Soc. **5**, 283 (1835).
13. V. N. Mahajan, J. Opt. Soc. Am. A **3**, 470 (1986).
14. L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 223 (1998).
15. S. W. Hell and J. Wichmann, Opt. Lett. **19**, 780 (1994).
16. R. Schmidt, C. A. Wurm, S. Jakobs, J. Engelhardt, A. Egner, and S. W. Hell, Nat. Methods **5**, 539 (2008).
17. N. K. Efremidis and D. Christodoulides, Opt. Lett. **35**, 4045 (2010).
18. D. G. Papazoglou, N. K. Efremidis, D. N. Christodoul-ides, and S. Tzortzakis, Opt. Lett. **36**, 1842 (2011).
19. P. Zhang, J. Prakash, Z. Zhang, M. S. Mills, N. K. Efremidis, D. N. Christodoulides, and Z. Chen, Opt. Lett. **36**, 2883 (2011).
20. R. Dorn, S. Quabis, and G. Leuchs, Phys. Rev. Lett. **91**, 233901 (2003).
21. K. Youngworth and T. Brown, Opt. Express **7**, 77 (2000).