Adaptive optimization for axial multi-foci generation in multiphoton microscopy

XINYANG LI,[1,2,6] **YUANLONG ZHANG,**[1,6] **KAN LIU,**[1,3] **HAO XIE,**[1] **HAOQIAN WANG,**[2] **LINGJIE KONG,**[4,5,*] iD **AND QIONGHAI DAI**[1]

[1]*Department of Automation, Tsinghua University, Beijing 100084, China*
[2]*Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China*
[3]*Tsinghua-Berkeley Shenzhen Institute, Tsinghua University, Shenzhen 518055, China*
[4]*State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instrument, Tsinghua University, Beijing 100084, China*
[5]*IDG/McGovern Institute for Brain Research, Tsinghua University, Beijing 100084, China*
[6]*These authors contributed equally to this work*
[*]*konglj@tsinghua.edu.cn*

Abstract: To improve imaging speed, multifocal excitation is widely adopted as a parallel strategy in laser-scanning microscopy. Specifically, axial multifocal microscopy is popular in neuroscience as it enables functional imaging of neurons in multiple depths simultaneously. However, previous phase searching algorithms for axial multi-foci generation generally generate foci of uniform intensities, which cannot compensate the scattering-induced power loss in deep tissue and causes inhomogeneous excitation. Here, we propose a novel adaptive optimization-based phase-searching method (AdaPS) to generate axial multi-foci with arbitrary intensity modulations for scattering-induced loss compensation. By adopting Adaptive Moment Estimation (Adam) as the searching algorithm, our method could escape from unsatisfactory local minima and stably converge to the optimal phase pattern with errors at least an order of magnitude lower. We validate AdaPS through both numerical simulations and experiments and demonstrate that AdaPS could provide uniform multi-depth imaging in scattering phantom and enable high-fidelity multi-depth recordings of neural network dynamics in mouse brain in vivo.

1. Introduction

Laser scanning microscopy has achieved great success in imaging and laser writing process [1,2]. However, hindered by the innate inertia of the mechanical scanners, laser scanning microscopy could not steer the focus rapidly [3], which limits its applications in high speed scenarios. Compared to the standard single focus scanning microscopy, multifocal scanning microscopy enables parallel access to the samples and thus improves the volumetric rate by many folds. Multifocal scanning microscopy is now widely applied in fluorescence imaging [4], Raman scattering microscopy [5], photostimulation [6], super-resolution microscopy [7] and two-photon polymerization [8]. In terms of spatial distribution of the foci, multifocal scanning microscopy can be divided into three different types, that is lateral [9], axial [10] and hybrid [11]. Among them, axial multifocal microscopy is particularly popular and holds great potentials in functional imaging and optogenetic photo-stimulation in neuroscience. Compared with Bessel beam microscopy [12,13] which also extends the axial excitation range, multifocal microscopy is more energy-efficient and suffers less background because it has no sidelobes [12]. Recently, axial multi-foci has been used for investigating cross-layer neural activities [14] and stimulating neural ensembles [15].

To achieve axial multifocal excitation, there are generally two strategies. The first strategy is spatio-temporal multiplexing, *i.e.* splitting the beam into multiple beamlets spatially before the objective, then modulating the wavefront of each beamlet and guiding them through the objective

to form multiple foci [16]. Such a strategy necessitates several extra beam splitters to split the beam, which increases the system complexity and is lack of flexibility. The second strategy employs a spatial light modulator (SLM), conjugated to the entrance pupil of the objective to control the beam phase for axial multi-foci generation. Such a method is easy to set up and the focal distribution (including number, position, and intensity) are flexible to adjust by phase modulation, which is more versatile among various applications.

However, considering ideal focal distribution needs both amplitude and phase modulation, it is non-trivial to get such a phase for the generation of expected multi-foci using common phase-only SLMs. Some algorithms such as Gerchberg–Saxton algorithm (GS) and its modified version, known as weighted Gerchberg–Saxton algorithm (GSW) and phase-only Gerchberg–Saxton algorithm (GSP), could generate multi-foci in 3D space with relatively uniform intensity distribution [10,17,18]. Nevertheless, these algorithms could not flexibly control the intensities of each focus, which hampers the application of multi-foci in deep tissue imaging, where power loss of the excitation laser induced by tissue scattering should be considered. Due to the heterogeneous refractive index of neurons, glia and blood vessels, a tight Gaussian focus would lose its energy exponentially as the penetration depth increases according to the Beer-Lambert law [19]. As a consequence, when applying multi-foci generated by the aforementioned algorithms in functional imaging of neural circuits, signals from the deeper layers would degrade severely. The different signal-to-noise ratios (SNRs) between different layers would result in the omission or misidentification of neurons, as well as the distortion of fluorescent dynamics. Methods like amplitude Gerchberg-Saxton algorithm (GSA) can modulate multi-foci more accurately but it relies on complex modulation [10], which reduce the laser power utilization efficiency when adjusting the beam amplitude [20]. Other intuitive methods, for instance, tuning the coefficients of Zernike polynomials for each focus to control the energy, could be time-consuming [14].

Here we propose an adaptive optimization-based phase-searching method (AdaPS) to effectively search the phases for the generation of axial multi-foci with flexible intensity control. AdaPS is based on a framework of stochastic optimization and Adam optimizer [21] is adopted to guarantee fast and accurate convergence. We theoretically prove that AdaPS would work better than previous GSW-kind methods and further analyze the influences of three factors (focus number, focus interval, and scattering length) through statistical analysis. We also set up a multifocal two-photon microscope and compare AdaPS with previous algorithms to show its superior performances. We demonstrate the capability of AdaPS in uniform multi-depth imaging of scattering phantom and high-fidelity recordings of neural network dynamics in mouse brain *in vivo*.

2. Methods

In this section, we derive the proposed algorithm for the generation of axial multi-foci with flexible intensity control. We also derive the GSW algorithm with the same notations to show that AdaPS would guarantee better solutions theoretically. We start with setting up a model to describe the interactions between the phase at the objective entrance pupil and the corresponding field distribution in the focal space, as shown in Figs. 1(a) and (b). We assume the optical field at the entrance pupil of the objective could be written as

$$E_0(r_0, \ \theta_0) = e^{-\left(\frac{r_0}{\omega_0}\right)^2} e^{i\phi(r_0, \ \theta_0)},\tag{1}$$

where ω_0 is the incident beam waist and $(r_0, \ \theta_0)$ is the polar coordinate on the entrance pupil plane. Since the generated field distribution should be circular symmetric, we assume $\phi(r_0, \ \theta_0) = \phi(r_0)$. Through Fresnel propagation in low numerical aperture conditions, the optical field in the focal

space is

$$E_1(r_1, \ \theta_1) = \frac{e^{ikf}}{i\lambda f} \int_0^{2\pi} \int_0^\infty \ r_0 E_0 e^{-\frac{ik}{f} r_0 r_1 \cos(\theta_0 - \theta_1)} e^{\frac{ik}{2f}(r_0^2 + r_1^2)} \mathrm{d}r_0 \mathrm{d}\theta_0, \tag{2}$$

where $(r_1, \ \theta_1)$ is the coordinate on the objective plane, f is the focal length of the objective and $k = 2\pi/\lambda$ is the wave number. After adding the lens modulation term $\exp(ikr_1^2/2f)$, the field $E_1(r_1, \ \theta_1)$ propagates further to $f + \Delta z$ after the objective

$$\begin{aligned}
&E_2(r_2, \ \theta_2, \ \Delta z) \\
&= \frac{e^{ik(f+\Delta z)}}{i\lambda(f + \Delta z)} \int_0^{2\pi} \int_0^\infty r_1 E_1 e^{-\frac{ik}{(f+\Delta z)} r_1 r_2 \cos(\theta_1 - \theta_2)} e^{\frac{ikr_1^2}{2f}} e^{\frac{ik}{2(f+\Delta z)}(r_1^2 + r_2^2)} \mathrm{d}r_1 \mathrm{d}\theta_1.
\end{aligned} \tag{3}$$

By using Gaussian integral, $E_2(r_2, \ \theta_2, \ \Delta z)$ could be finally simplified as

$$E_2(r_2, \ \theta_2, \ \Delta z) = E_2(r_2, \ \Delta z) \propto \int_0^\infty r_0 e^{-\left(\frac{r_0}{\omega_0}\right)^2} e^{i\phi(r_0)} e^{ik_0\left(\Delta z \frac{r_0^2}{2f^2}\right)} J_0\left(-\frac{k_0 r_0 r_2}{f}\right) \mathrm{d}r_0, \tag{4}$$

where J_0 is the zero-order Bessel function of the first kind [22].

To get the discretization model for numerical simulation, we need to evaluate the integral at discrete points $\{r_{0, m}\}$ on the entrance pupil plane and $\{r_{2, m}\}$ on the focal plane. Besides, discretized functions in such discrete grids should faithfully approximate the value of the corresponding continuous function. For a finite spatial range $[0, \ R]$ on the focal plane, the sampling points on the pupil and focal plane should be

$$r_{0, m} = \frac{j_{0, m}}{R} \frac{f}{k_0}, \ m = 1, \ldots, M - 1,$$

$$r_{2, m} = \frac{j_{0, m}}{j_{0, M}} R, \ m = 1, \ldots, M - 1, \tag{5}$$

where $j_{0, m}$ is the m-th zeros of J_0 [23]. The forward model in Eq. (4) could be written as discrete Hankel transform [23]

$$\mathbf{E}_2(\Delta z) = \mathbf{Y}[\mathbf{p}(\Delta z) \odot \mathbf{\Phi}], \tag{6}$$

and elements of $\mathbf{Y}$, $\mathbf{p}$ and $\mathbf{\Phi}$ are

$$Y_{m, n} = \frac{2}{j_{0,M} J_1^2(j_{0, n})} J_0\left(\frac{j_{0, m} j_{0, n}}{j_{0, M}}\right),$$

$$p_m(\Delta z) = e^{-\left(\frac{r_{0, m}}{\omega_0}\right)^2} e^{ik_0\left(\Delta z \frac{r_{0, m}^2}{2f^2}\right)},$$

$$\Phi_m = e^{i\phi_m}. \tag{7}$$

Here $\odot$ means Hadamard product. Note that Y is an orthogonal matrix, which means $\mathbf{Y}^T \mathbf{Y} = \mathbf{Y}\mathbf{Y}^T = \mathbf{I}$.

Our goal is to generate L foci at axial positions $\{\Delta z_l\}_{l=1, \ldots, L}$, of which the corresponding intensities should be $\{u_l\}_{l=1, \ldots, L}$ (in two-photon scenario). In GSW, a desired intensity distribution

$\mathbf{E}_{target}(\Delta z)$ is used to control the iteration process. We could define $\mathbf{E}_{target}(\Delta z)$ as

$$\mathbf{E}_{target}(\Delta z_l) = \begin{cases} u_l^{1/4}, & for\ r_{2,m} = \frac{j_{0,1}}{j_{0,M}}R \\ 0, & others. \end{cases} \tag{8}$$

The GSW algorithm is then used to generate the desired intensity distribution by maximizing the energy efficiency, *i.e.* the total intensity at target axial positions $\{\Delta z_l\}_{l=1,\ldots,L}$

$$\underset{\Phi}{\text{argmin}} - \sum_{l=1}^{L} ||\mathbf{E}_2(\Delta z_l)||^4. \tag{9}$$

And meanwhile, the similarity between the intensities of generated foci and the target $\mathbf{E}_{target}(\Delta z)$ is also taken into consideration. In this step, a group of weights denoted by $\{w_l\}_{l=1,\ldots,L}$ are introduced to guarantee that the generated foci have the specified intensity distribution $\mathbf{E}_{target}(\Delta z_l)$. The axial GSW algorithm for multifocal generation could be written as

---
Algorithm 1. GSW algorithm

Initialize $w_l^{(0)} = 1, \phi_m^{(0)}$ is random within $[0, 2\pi), t = 0$.
while $t <$ T **do:**
  $t = t + 1$
  $w_l^{(t)} \leftarrow w_l^{(t-1)} \frac{\sum_{l=1}^{L} ||\mathbf{E}_2(\Delta z_l)||^4}{L ||\mathbf{E}_2(\Delta z_l)||^4}$
  for $m = 1, \ldots, M - 1$
    $\phi_m^{(t)} \leftarrow \arg\left[\sum_{l=1}^{L} Y^T \left(w_l^{(t)} \mathbf{E}_{target}(\Delta z_l) \odot \frac{\mathbf{E}_2(\Delta z_l)}{|\mathbf{E}_2(\Delta z_l)|}\right)\right],$
  end for
end while
return $\phi^{(t)}$

Because the problem in Eq. (9) is not convex [17], in each step (*i.e.* an iteration in **Algorithm 1**) the GSW algorithm alternatively fix $\Phi_1, \ldots, \Phi_{m-1}, \Phi_{m+1}, \ldots, \Phi_{M-1}$ and use gradient descent method to solve $\arg\min - \sum_{l=1}^{L} w_l ||\mathbf{E}_2(\Delta z_l)||^4$ for $m = 1, \ldots, M - 1$ [24]. The performance of such optimization methods severely relies on the initial value and is very easy to be stuck in local minima.

Different from GSW, we explicitly decouple the energy efficiency term and the distribution penalty term (the uniformity of weighted focal intensities) in the loss function

$$\underset{\Phi}{\text{argmin}}\, F(\phi) = - \sum_{l=1}^{L} ||\mathbf{E}_2(\Delta z_l)||^4 + \rho \frac{\max_l ||\mathbf{E}_2(\Delta z_l)||^4 / u_l}{\min_l ||\mathbf{E}_2(\Delta z_l)||^4 / u_l}. \tag{10}$$

The second term guarantees that after adding extra weights u_l to each focus, all the foci should have specified intensities. ρ is a hand-tuned constant to adjust regularization. The optimization problem is still non-convex. To avoid oscillation and being stuck in local minima, we use Adam, an adaptive optimization method that applies adaptive learning rates for each parameter, as well as estimating the lower-order moments to accelerate convergence [21]. The algorithm flow chart is shown below


Algorithm 2. Adam optimization algorithm

Initialize $\alpha, \beta_1, \beta_2, m_0, v_0, t, \epsilon$, T

while $t <$ T do:

$\quad t = t + 1$

$\quad g^{(t)} \leftarrow \nabla_\phi F^{(t)}\big(\phi^{(t)}\big)$

$\quad m^{(t)} \leftarrow \beta_1 m^{(t-1)} + (1 - \beta_1)g^{(t)}$

$\quad v^{(t)} \leftarrow \beta_2 v^{(t-1)} + (1 - \beta_2)(g^{(t)})^2$

$\quad \widehat{m}^{(t)} \leftarrow m^{(t)}/\big(1 - \beta_1^{(t)}\big)$

$\quad \widehat{v}^{(t)} \leftarrow v^{(t)}/\big(1 - \beta_2^{(t)}\big)$

$\quad \phi^{(t)} \leftarrow \phi^{(t-1)} - \alpha\widehat{m}^{(t)}/(\sqrt{\widehat{v}^{(t)}} + \epsilon)$

end while

return $\phi^{(t)}$

Here $\nabla_\phi F^{(t)}(\phi^{(t)})$ could be implemented by numerical gradients at t-th iteration. Thanks to the nature of adaptive optimization and the additional bias-corrected moments ($\widehat{m}^{(t)}$ and $\widehat{v}^{(t)}$), Adam optimizer could not only escape from local minima and saddle points but also guarantee a dramatically fast convergence [21,25], which leads AdaPS to a better solution compared to GSW theoretically.

The operating workflow of AdaPS is schematized in Fig. 1(c) and can be summarized as 3 steps. Firstly, the phase is randomly initialized and a Fresnel propagation-based forward model is applied to calculate the optical field near the focal plane of the objective [Eq. (6)]. Secondly, current optical field is compared with the target field and an error is calculated through a predefined loss function [Eq. (10)]. Finally, Adam optimizer is utilized to minimize the loss function (Algorithm 2). This process iterates repeatedly until the specified number of iterations is reached. Our code is made publicly available on Github (https://github.com/Xinyang-Li/AdaPS).
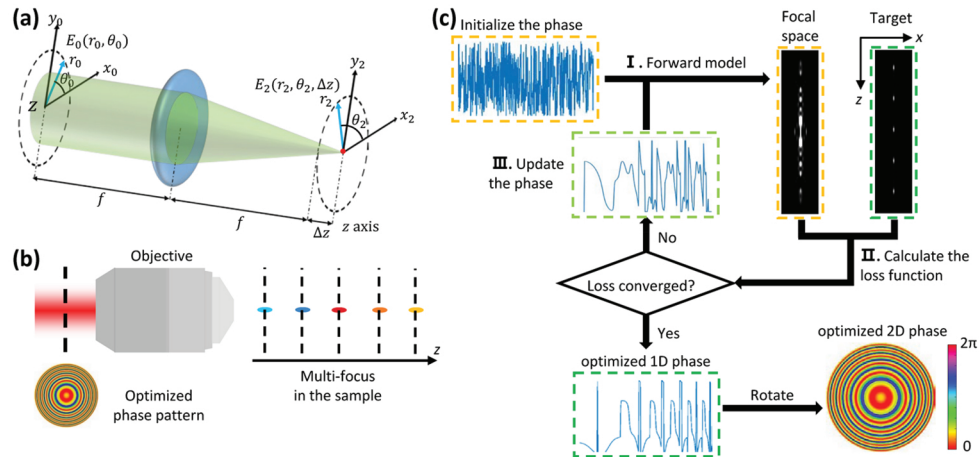
Fig. 1. Overview of the principles of AdaPS. (a) Schematic for calculating the light field near the focal plane of the objective. (b) An optimized phase pattern displayed on the SLM is conjugated to the entrance pupil plane of the objective, and then multi-foci is generated in different depths. (c) The workflow of our Adam-based technique: I. A random phase pattern is initialized and the corresponding light field in focal space is calculated through a forward model; II. A loss is calculated to represent the difference between current focal field and final target; III. Adam optimizer is applied to update the phase until the loss converges.
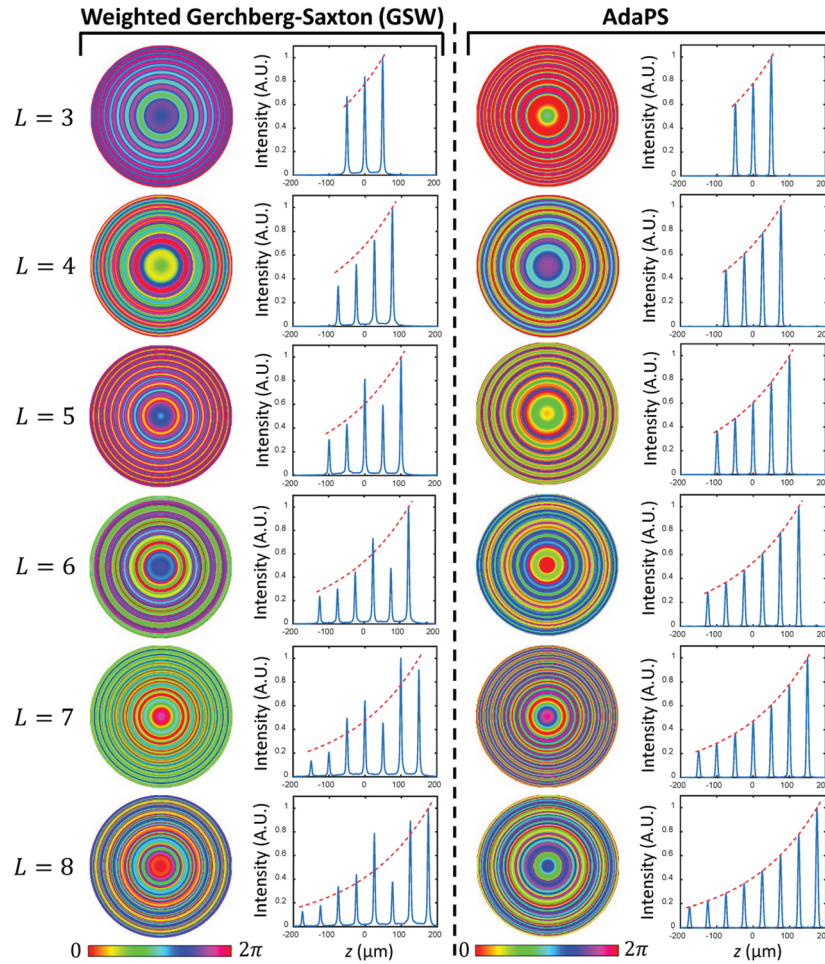
Fig. 2. Simulation results of multifocal generation with intensity modulation by GSW and our proposed method. For each method, the first column shows the generated phase, and the second column shows the corresponding axial intensity by two-photon excitation ($L = 3 \sim 8$, $\Delta z_l = 50$ µm, $s = 400$ µm). The red curve shows the target intensity distribution, generated by the Beer-Lambert law.

3. Simulations

To test the performance of AdaPS, we perform numerical simulations with different focus numbers (L), focus intervals (Δz_l), and scattering lengths (s) in two-photon scenario. The algorithms are implemented under Python 3.6.

For two-photon excitation in scattering media, the effective excitation power decays exponentially with the penetration depth, which is quantified by scattering length (or scattering mean-free-path) [26]. In the simulation, we suppose the scattering length $s = 400$ µm. With focus interval $\Delta z_l = 50$ µm, the target two-photon intensity of focus l (of total focus number L) is $u_l = e^{\frac{2(l-1)}{s} \Delta z_l}$. We then run Algorithm 1 and Algorithm 2 with focus number from 3 to 8 to check the results, as shown in Fig. 2. Since the performance of GSW is closely relevant to the initial phase, we conduct 30 independent trials repeatedly and select the one with minimal loss to show in Fig. 2. Under the same focus number and specified intensity requirement, focal distribution

generated by AdaPS could exactly match the expected exponential curves (red dashed curves), while GSW method has larger deviations. As the focus number increases, the optimization problem becomes more complicated. GSW tends to perform worse because it is more likely to be trapped in local minima and saddle points.

Although there are differences between the objective functions and optimization algorithms of the two methods, the superiority of AdaPS can still be visualized quantitatively. We introduce the "trend error", defined by $\text{err}(L) = \max_l \frac{\widehat{u_l}}{u_l} / \min_l \frac{\widehat{u_l}}{u_l}$, where u_l is the expected two-photon focal intensity and $\widehat{u_l}$ is the actual focal peak intensity after optimization. The minimal value of $\text{err}(L)$ is 1 when $\max_l \frac{\widehat{u_l}}{u_l} = \min_l \frac{\widehat{u_l}}{u_l}$, which further means that the focal intensity is strictly consistent with the target. Otherwise, the larger the error, the greater the deviation is. We initialize two algorithms with the same random phase and plot the decline curves of $\text{err}(L)$, as shown in Fig. 3(a). It is clear that our Adam-based method can robustly converge to the expected intensity distribution [$\text{err}(L) \approx 1$], at least one order of magnitude better than GSW.
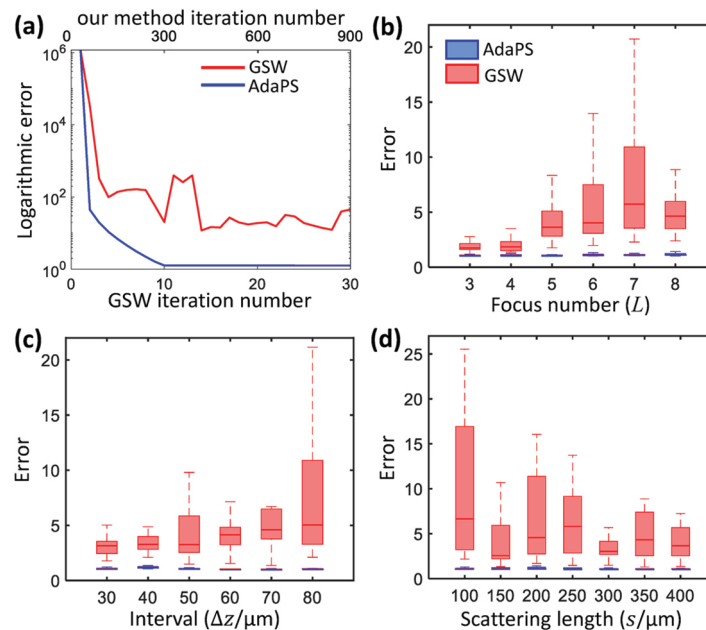
Fig. 3. Quantitative analysis of the algorithm performance. (a) Convergence curves of the trend errors $\text{err}(L)$ of GSW and AdaPS. (b) Boxplot of the errors under different focus numbers (set $\Delta z_l = 50$ μm, $s = 400$ μm). (c) Boxplot of the errors under different focus intervals (set $L=5$, $s = 400$ μm). (d) Boxplot of the errors under different scattering lengths (set $L=5$, $\Delta z_l = 50$ μm). The data of each set of parameters are obtained from 30 independent trials repeatedly.

In order to independently analyze the impact of the three factors (L, Δz_l and s) on GSW and AdaPS, we also conduct independent simulations on each one of them repeatedly, while keeping the other two constant. For each set of parameters, 30 trials are performed and the boxplots are shown in Figs. 3(b), 3(c) and 3(d), respectively. The red boxes are the errors of GSW while the blue boxes at the bottom are the errors of AdaPS. There is a tendency that the performance of GSW becomes worse and more unstable with the number of foci increasing, the focus interval increasing, or the scattering length decreasing. Thus, GSW is likely to fail in generating foci of exponential-increasing intensities to fully compensate the power loss in strong scattering medium.

In contrast, AdaPS maintains strong stability and small errors in these trials. In most cases, our Adam-based method guarantees a better solution to this non-convex problem.

4. Experiment results

All procedures involving mice were approved by the Animal Care and Use Committees of Tsinghua University.

4.1. Experimental validation of flexible intensity control with fluorescent beads

Firstly, we show that the phase searched by our algorithm could be used for generating multi-foci with desired distributions and intensities. We set up a two-photon microscopy with a SLM, as shown in Fig. 4(a). Following the beam expander to adjust the beam size, we simply introduce a SLM (X10468-07, Hamamatsu) to modulate the beam wavefront. The SLM is then conjugated to the dual axis scanner (LSK-GR08, Thorlabs) via L1 (f_1 = 400 mm) and L2 (f_2 = 100 mm), which reduce the beam size to fit the galvo mirror size. The modulated and scanned beam then passes a scan lens (L3, f_3 = 50 mm), a tube lens (L4, f_4 = 200 mm) and an objective (XLPLN25XWMP2, 25x, NA1.05) to form multifocal field near the focal plane. To achieve a relatively large focal range, we underfill the aperture of the objective [14] with effective excitation NA = 0.55. The measured single focus resolution (FWHM) is 0.66 μm laterally and 6 μm axially calibrated by 200 nm fluorescent beads, as shown in Fig. 4(b).
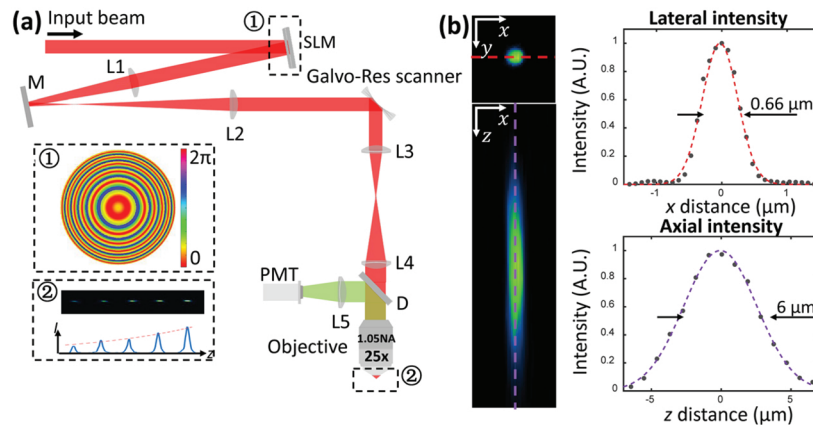
Fig. 4. (a) Experimental setup of a multifocal microscopy. The femtosecond laser beam is modulated by a SLM, and then conjugated to the galvo-resonant scanner via a 4f system (L1 and L2, with f_1 = 400 mm and f_2 = 100 mm). The scanner is then conjugated to the entrance pupil of the objective via scan lens (L3, f_3 = 50 mm) and tube lens (L4, f_4 = 200 mm). The modulated beam finally forms multifocal excitation near the focal plane. Inset ① shows an example SLM phase. Inset ② shows the corresponding multi-foci of exponential-increasing intensities. (b) Lateral (top row) and axial (bottom row) resolution calibration. Scale bar, 1 μm.

To experimentally validate our proposed method, we load the generated phase in Fig. 2(a) onto the SLM, then drive the objective by a piezo scanner to make the fluorescent spheres excited by each focus. To eliminate the influence of unstable factors, such as noise and photobleaching, we extract axial intensities of all 8 beads in the field-of-view (FOV). The statistical results are shown in Fig. 5. We also plot a typical intensity profile of a specified bead in each panel to show the overall intensity distribution (blue lines). All axial intensity profiles are normalized to its corresponding maximum.
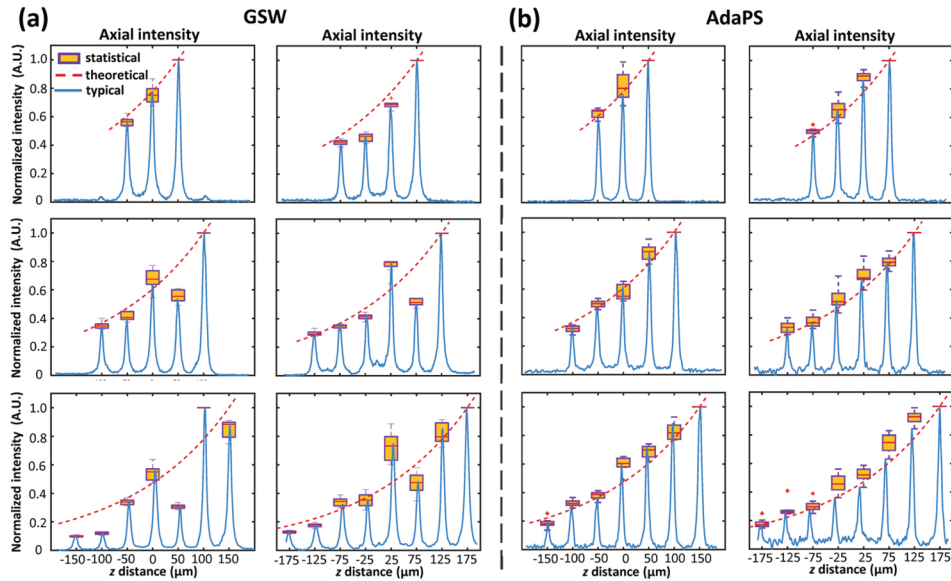
Fig. 5. Performance comparison of GSW and AdaPS on the generation of exponential-increasing multi-foci ($L = 3 \sim 8$, $\Delta z_l = 50$ μm, $s = 400$ μm). (a) Focal intensity with phase generated by GSW. (b) Focal intensity with phase generated by AdaPS. Orange boxes are the statistical results of all 8 beads in the FOV; red dashed lines are exponential trend lines; blue lines are typical axial intensity profiles measured from a specified fluorescent bead.

From Fig. 5, it is clear that the intensity distributions generated by GSW algorithm [Fig. 5(a)] have large deviations from the required intensity distributions (red dashed lines), while the results from AdaPS [Fig. 5(b)] is much closer to the expectation. When the focus number is small ($L = 3$), GSW and AdaPS have comparable performance. But when it is required to increase the focus number ($L > 3$), AdaPS shows its superiority. Benefiting from its adaptive learning rates and momentum term, AdaPS can converge to better solutions and thus keep the deviations at a very low level regardless of the focus number or expected intensity distribution.
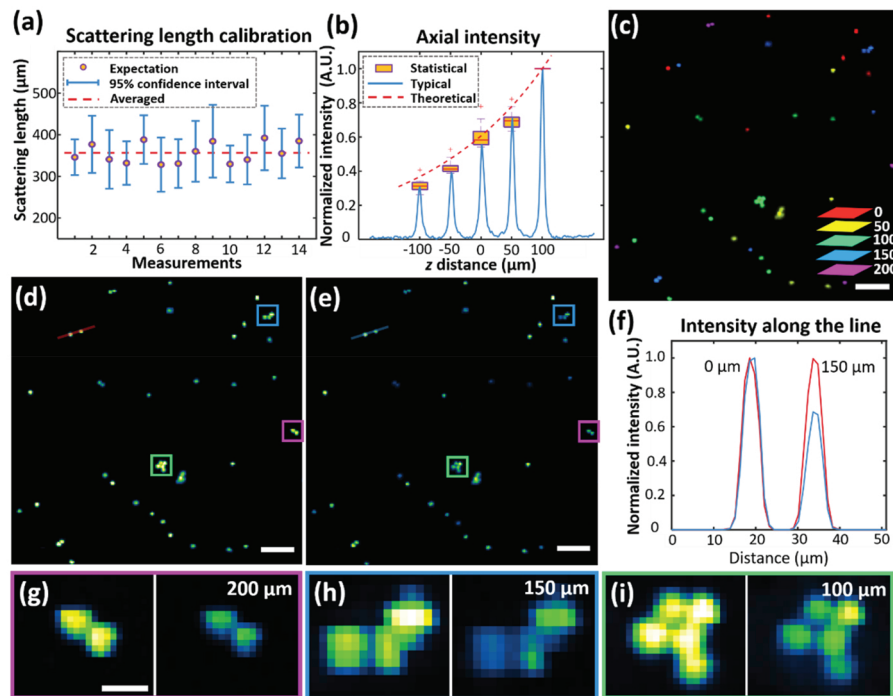
4.2. Uniform multi-depth imaging in scattering phantom

We then show that AdaPS enables uniform axial multi-depth excitation in scattering phantom. We mix 3 μm fluorescent beads and 0.5 μm non-fluorescent beads homogeneously in a ratio of 1:60, and then embed the bead mixture at a concentration of 10% in 2% agarose gel to form a scattering phantom. In order to customize a phase pattern to generate multi-foci of exponential-increasing intensities, we calibrate the scattering length of the phantom by 14 independent trials. In each trial, a single focus scans through the whole volume and the averaged fluorescence signals at each depth are exponentially fitted to estimate the scattering length. The results are shown in Fig. 6(a). The blue line segments are 95% confidence intervals of each estimation and the red dashed line is the averaged value of the scattering length (~350 μm).

We generate 5-foci ($L=5$, $\Delta z_l = 50$ μm, $s = 350$ μm) of exponential-increasing intensities through our proposed method to compensate the scattering. The focal intensity is calibrated as the same procedures in Sec. 4.1, and the corresponding axial two-photon intensity is shown in Fig. 6(b), as well as the target intensity (dashed red line). To locate the depth of each bead, we first scan the 5 planes sequentially using a single focus as in conventional multiphoton microscopy (by loading flat phase onto the SLM). The depth map is shown in Fig. 6(c) that different colors represent different depths. We then drive the scanner to scan the customized 5-foci laterally

for simultaneous 5-depth imaging of the scattering phantom. Fluorescent signals from all the 5 planes are captured in parallel, as shown in Fig. 6(d). To show the necessities of scattering compensation, we also scan the same FOV by 5-foci with the same spatial localizations but equal excitation power (uniform intensities), as shown in Fig. 6(e). It can be seen that the image with scattering compensation [Fig. 6(c)] is much more uniform in brightness than that without [Fig. 6(d)]. The improvement of image contrast is also quantitatively shown in Fig. 6(f), where two beads with 150 μm axial distance have the same signal strength under scattering compensation. Also, some detailed structures are shown in Figs. 6(g)–6(i) to indicate that with the help of our exponential-increasing multi-foci for scattering compensation, structures from different depths can be clearly revealed in the same contrast.

4.3. High-fidelity multi-layer calcium imaging in mouse brain in vivo

Moreover, we demonstrate that AdaPS enables simultaneous multi-depth functional imaging in mouse brain *in vivo* in high fidelity. For mouse preparation, we use wild-type mice at 8-10

Fig. 6. Simultaneous multi-depth imaging in scattering phantom. (a) Scattering length of the phantom (∼350 μm) is calibrated by 14 independent trials. (b) Axial intensity profile of the customized 5-foci of exponential-increasing intensities ($L = 5$, $\Delta z_l = 50$ μm, $s = 350$ μm), measured in the same way as Fig. 5. (c) The depth map of all beads in the FOV, where 5 colors code 5 different depths (0/50/100/150/200 μm). (d) and (e) are the images scanned by 5-foci of exponential-increasing intensities and uniform intensities, respectively. Scale bar, 50 μm. (f) Intensity along the lines in (d) and (e), where the two beads located in plane 1 ($z = 0$ μm) and plane 4 ($z = 150$ μm) show similar contrast when being excited by our exponential-increasing 5-foci. (g-i) Detailed structures in (d) and (e), located in plane 5 ($z = 2000$μm), plane 4 ($z = 150$ μm), and plane 3 ($z = 100$ μm), respectively. The loss of excitation power caused by scattering could be compensated by our exponential-increasing 5-foci. Scale bar, 10 μm.

postnatal weeks and performed craniotomy above the primary somatosensory cortex (S1). Neuron populations at Layer 1 and Layer 2/3 are labeled with fluorescent calcium indicator GCaMP6s [27] through virus infection. After virus injection and optical window implanting, we allow 3-4 weeks for the mice to recover.

For two-photon excitation in mouse brain tissue *in vivo*, the scattering length is about $s = 170$ μm at 920 nm [28]. In order to distribute the excitation power to compensate the power loss, a phase pattern is customized through AdaPS to generate 4-foci of exponential-increasing intensities and the focus interval is set as 60 μm ($L=4$, $\Delta z_l = 60$ μm, $s = 170$ μm). The two-photon intensity is calibrated as methods in Sec. 4.1 and shown in Fig. 7(a). For comparison, we also generate 4-foci of the same spatial localizations but equal intensities. By loading the calculated phase and driving the scanner, 4 planes (80/140/200/260 μm under the dura, respectively) are imaged simultaneously at 30 Hz without moving the sample or objective axially. The locations of neurons are recorded by scanning the four planes sequentially by a single focus, as shown in Fig. 7(b). The temporal standard deviation projections (SDPs) of the recorded image sequences are shown in Fig. 7(c) and Fig. 7(d), respectively. Because of the optimized intensity distribution of our 4-foci, both shallow planes and deeper planes have comparable SNRs, which cannot be achieved by 4-foci excitation of uniform intensities.
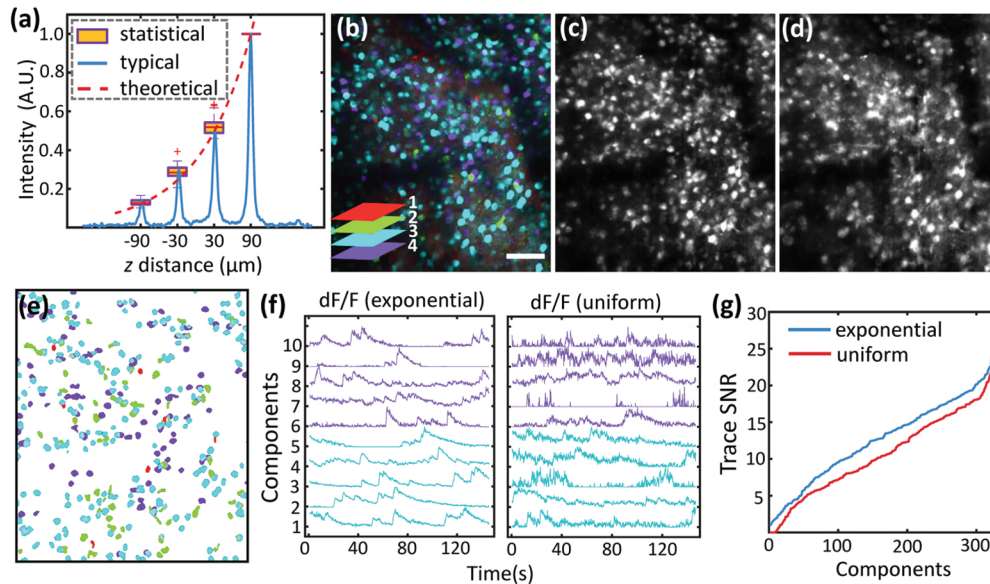
Fig. 7. High-fidelity multi-depth calcium recording *in vivo*. (a) The 4-foci of exponential-increasing intensities ($L = 4$, $\Delta z_l = 60$ μm, $s = 170$ μm) generated by AdaPS. (b) Depth map of neurons in the FOV, composed of the temporal SDPs of each plane (80/140/200/260 μm under the dura) that scanned by a single focus. Scale bar, 100 μm. (c) and (d) are the temporal SDPs of simultaneous 4-plane imaging scanned by 4-foci of exponential-increasing intensities and uniform intensities, respectively. (e) The 517 spatial components (7/87/282/141 in each plane) extracted by CNMF. (f) Representative normalized dF/F traces on plane 3 (cyan) and plane 4 (purple), recorded by 4-foci of exponential-increasing intensities (left panel) and uniform intensities (right panel), respectively. (g) SNRs of the traces on plane 3 and plane 4 (328 out of 423 spatial components, with SNR > 0) are plotted to indicate that our 4-foci of exponential-increasing intensities can record calcium dynamics in high fidelity.

We further extracted corresponding calcium traces to demonstrate that our multi-foci excitation of exponential-increasing intensities designed for scattering compensation is helpful in improving

trace SNR. Our data analysis pipeline includes motion correction using NoRMCorre [29] and source extraction based on constrained non-negative matrix factorization (CNMF) [30], which is generally adopted in demixing spatially overlapping components. Totally 517 source components are extracted, and the spatial contours are shown in Fig. 7(e). We plot the temporal traces of 10 representative spatial components on plane 3 (200 μm depth, in cyan) and plane 4 (260 μm depth, in purple) in Fig. 7(f). The two panels are recorded by exponential-increasing intensities (left) and uniform intensities (right), respectively. Trace SNRs are obviously improved, which means that neural activities can be revealed in high fidelity by AdaPS. For a more general analysis, we calculate the trace SNRs of each spatial components on the two planes (plane 3 and plane 4) according to the formulas as described in [31]. Traces with SNR below zero are discarded because they are severely affected by shot noise. All the non-negative SNRs are sorted and plotted in Fig. 7(g). On average, the SNR of calcium traces recorded by our exponential-increasing 4-foci is improved by 2dB, benefiting from the optimized power distribution of our method. Such improvement shows the capability of AdaPS to enable high-fidelity calcium imaging in scattering tissues under limited illumination power.

5.   Conclusion

In summary, we demonstrate a novel technique to generate axial multi-foci with arbitrary intensity modulations. By introducing Adam as the optimization method, our algorithm could guarantee fast and accurate convergence to the expected phases, regardless of focus number, focus interval, and energy distribution, which is difficult for other algorithms. We demonstrate that AdaPS enables uniform axial multi-depth imaging in scattering phantoms, as well as high-fidelity multi-layer calcium recording in mouse brain *in vivo*. It should be noted that our proposed method could also be adopted in lateral multi-foci generation and 3D computer-generated holography.

References

1. C. Xu, W. Zipfel, J. B. Shear, R. M. Williams, and W. W. Webb, "Multiphoton fluorescence excitation: new spectral windows for biological nonlinear microscopy," Proc. Natl. Acad. Sci. U. S. A. **93**(20), 10763–10768 (1996).
2. M. Deubel, G. von Freymann, M. Wegener, S. Pereira, K. Busch, and C. M. Soukoulis, "Direct laser writing of three-dimensional photonic-crystal templates for telecommunications," Nat. Mater. **3**(7), 444–447 (2004).
3. W. Yang and R. Yuste, "In vivo imaging of neural activity," Nat. Methods **14**(4), 349–359 (2017).
4. W. Yang and R. Yuste, "Holographic imaging and photostimulation of neural activity," Curr. Opin. Neurobiol. **50**, 211–221 (2018).
5. T. Minamikawa, M. Hashimoto, K. Fujita, S. Kawata, and T. Araki, "Multi-focus excitation coherent anti-Stokes Raman scattering (CARS) microscopy and its applications for real-time imaging," Opt. Express **17**(12), 9526 (2009).
6. F. Anselmi, C. Ventalon, A. Begue, D. Ogden, and V. Emiliani, "Three-dimensional imaging and photostimulation by remote-focusing and holographic light patterning," Proc. Natl. Acad. Sci. U. S. A. **108**(49), 19504–19509 (2011).
7. A. G. York, S. H. Parekh, D. Dalle Nogare, R. S. Fischer, K. Temprine, M. Mione, A. B. Chitnis, C. A. Combs, and H. Shroff, "Resolution doubling in live, multicellular organisms via multifocal structured illumination microscopy," Nat. Methods **9**(7), 749–754 (2012).
8. K. Obata, J. Koch, U. Hinze, and B. N. Chichkov, "Multi-focus two-photon polymerization technique based on individually controlled phase modulation," Opt. Express **18**(16), 17193–17200 (2010).
9. S. P. Poland, N. Krstajic, R. D. Knight, R. K. Henderson, and S. M. Ameer-Beg, "Development of a doubly weighted Gerchberg-Saxton algorithm for use in multibeam imaging applications," Opt. Lett. **39**(8), 2431–2434 (2014).
10. E. H. Waller and G. von Freymann, "Multi foci with diffraction limited resolution," Opt. Express **21**(18), 21708–21713 (2013).

11. N. C. Pegard, A. R. Mardinly, I. A. Oldenburg, S. Sridharan, L. Waller, and H. Adesnik, "Three-dimensional scanless holographic optogenetics with temporal focusing (3D-SHOT)," Nat. Commun. **8**(1), 1228 (2017).

12. B. Chen, X. Huang, D. Gou, J. Zeng, G. Chen, M. Pang, Y. Hu, Z. Zhao, Y. Zhang, Z. Zhou, H. Wu, H. Cheng, Z. Zhang, C. Xu, Y. Li, L. Chen, and A. Wang, "Rapid volumetric imaging with Bessel-Beam three-photon microscopy," Biomed. Opt. Express **9**(4), 1992–2000 (2018).

13. R. Lu, W. Sun, Y. Liang, A. Kerlin, J. Bierfeld, J. D. Seelig, D. E. Wilson, B. Scholl, B. Mohar, M. Tanimoto, M. Koyama, D. Fitzpatrick, M. B. Orger, and N. Ji, "Video-rate volumetric functional imaging of the brain at synaptic resolution," Nat. Neurosci. **20**(4), 620–628 (2017).

14. W. Yang, J. E. Miller, L. Carrillo-Reid, E. Pnevmatikakis, L. Paninski, R. Yuste, and D. S. Peterka, "Simultaneous Multi-plane Imaging of Neural Circuits," Neuron **89**(2), 269–284 (2016).

15. M. Dal Maschio, J. C. Donovan, T. O. Helmbrecht, and H. Baier, "Linking Neurons to Network Function and Behavior by Two-Photon Holographic Optogenetics and Volumetric Imaging," Neuron **94**(4), 774–789.e5 (2017).

16. A. Cheng, J. T. Goncalves, P. Golshani, K. Arisaka, and C. Portera-Cailliau, "Simultaneous two-photon calcium imaging at different depths with spatiotemporal multiplexing," Nat. Methods **8**(2), 139–142 (2011).

17. R. D. Leonardo, F. Ianni, and G. Ruocco, "Computer generation of optimal holograms for optical trap arrays," Opt. Express **15**(4), 1913–1922 (2007).

18. T. Haist, M. Schönleber, and H. J. Tiziani, "Computer-generated holograms from 3D-objects written on twisted-nematic liquid crystal displays," Opt. Commun. **140**(4-6), 299–308 (1997).

19. D. F. Swinehart, "The Beer-Lambert Law," J. Chem. Educ. **39**(7), 333–335 (1962).

20. J. A. Davis, D. M. Cottrell, J. Campos, M. J. Yzuel, and I. Moreno, "Encoding amplitude information onto phase-only filters," Appl. Opt. **38**(23), 5004–5013 (1999).

21. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980 [cs] (2014).

22. L. C. Andrews, *Special Functions of Mathematics for Engineers*, 2nd ed. (SPIE Optical Engineering, 1998).

23. U. Chouinard and N. Baddour, "Matlab Code for the Discrete Hankel Transform," J. Open Res. Softw. **5**(1), 4 (2017).

24. J. Zhang, N. Pégard, J. Zhong, H. Adesnik, and L. Waller, "3D computer-generated holography by non-convex optimization," Optica **4**(10), 1306–1313 (2017).

25. X. Chen, S. Liu, R. Sun, and M. Hong, "On the Convergence of A Class of Adam-Type Algorithms for Non-Convex Optimization," arXiv:1808.02941 [cs] (2018).

26. F. Helmchen and W. Denk, "Deep tissue two-photon microscopy," Nat. Methods **2**(12), 932–940 (2005).

27. T. W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda, and D. S. Kim, "Ultrasensitive fluorescent proteins for imaging neuronal activity," Nature **499**(7458), 295–300 (2013).

28. N. G. Horton, K. Wang, D. Kobat, C. G. Clark, F. W. Wise, C. B. Schaffer, and C. Xu, "In vivo three-photon microscopy of subcortical structures within an intact mouse brain," Nat. Photonics **7**(3), 205–209 (2013).

29. E. A. Pnevmatikakis and A. Giovannucci, "NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data," J. Neurosci. Methods **291**, 83–94 (2017).

30. E. A. Pnevmatikakis, D. Soudry, Y. Gao, T. A. Machado, J. Merel, D. Pfau, T. Reardon, Y. Mu, C. Lacefield, W. Yang, M. Ahrens, R. Bruno, T. M. Jessell, D. S. Peterka, R. Yuste, and L. Paninski, "Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data," Neuron **89**(2), 285–299 (2016).

31. A. Giovannucci, J. Friedrich, P. Gunn, J. Kalfon, B. L. Brown, S. A. Koay, J. Taxidis, F. Najafi, J. L. Gauthier, P. Zhou, B. S. Khakh, D. W. Tank, D. B. Chklovskii, and E. A. Pnevmatikakis, "CaImAn an open source tool for scalable calcium imaging data analysis," Elife **8**, e38173 (2019).