Optimization-based wavefront sensorless adaptive optics for multiphoton microscopy

Jacopo Antonello,[1,*] Tim van Werkhoven,[2,3] Michel Verhaegen,[1] Hoa H. Truong,[3] Christoph U. Keller,[2] and Hans C. Gerritsen[3]

[1]*Delft Center for System and Control, Delft University, Mekelweg 2, 2628 CD Delft, The Netherlands*
[2]*Leiden Observatory, Leiden University, P.O. Box 9513, 2300 RA Leiden, The Netherlands*
[3]*Molecular Biophysics, Utrecht University, P.O. Box 80000, 3508 TA Utrecht, The Netherlands*
Corresponding author: j.antonello@tudelft.nl

Optical aberrations have detrimental effects in multiphoton microscopy. These effects can be curtailed by implementing model-based wavefront sensorless adaptive optics, which only requires the addition of a wavefront shaping device, such as a deformable mirror (DM) to an existing microscope. The aberration correction is achieved by maximizing a suitable image quality metric. We implement a model-based aberration correction algorithm in a second-harmonic microscope. The tip, tilt, and defocus aberrations are removed from the basis functions used for the control of the DM, as these aberrations induce distortions in the acquired images. We compute the parameters of a quadratic polynomial that is used to model the image quality metric directly from experimental input–output measurements. Finally, we apply the aberration correction by maximizing the image quality metric using the least-squares estimate of the unknown aberration.   © 2014 Optical Society of America

OCIS codes: (010.1080) Active or adaptive optics; (010.7350) Wave-front sensing; (220.1000) Aberration compensation; (110.0113) Imaging through turbid media; (180.5810) Scanning microscopy; (180.4315) Nonlinear microscopy.

http://dx.doi.org/10.1364/JOSAA.31.001337

1. INTRODUCTION

Multiphoton microscopy techniques, such as two-photon fluorescence microscopy [1] and second-harmonic microscopy [2], are commonly employed to image biological specimens. Exploiting the image sectioning properties of these processes, one can create high-resolution 3D reconstructions that are invaluable for biomedical research. One limiting factor is the presence of specimen-induced aberrations. Because the index of refraction is not homogeneous within the specimen, aberrations affect both the resolution and the maximum depth of penetration [3]. Using adaptive optics [4], these detrimental effects can be minimized by reducing the phase aberrations. A phase aberration can be introduced in the excitation beam by means of a deformable mirror (DM), for example. Chosen correctly, such a phase aberration can suppress some amount of the specimen-induced aberrations.

Direct measurement of the specimen-induced aberrations is challenging. One solution involves measuring the aberrations of the excitation light that is back-scattered from the specimen [5–10]. In this case, the difficulty arises in excluding the light that is reflected from the out-of-focus layers of the specimen [5,6]. Additionally, these measurements are weakly sensitive to odd aberrations [5], due to the double-pass effect [11]. In another solution, instead, the emission from a point source inside the specimen is used to perform Shack–Hartmann wavefront sensing [12–16]. Here, the difficulty stems from the lack of such reference point sources within the specimen and from the limited number of photons available in the emission signal.

An alternative, indirect approach involves deducing the specimen-induced aberrations solely by examining the emission signal. This approach only requires the addition of a DM into the excitation path of an existing microscope. A solution that is based on the segmentation of the pupil has been proposed [17,18]. Other solutions are based on the optimization of an image quality metric, which attains its global maximum when the residual aberration is maximally suppressed. In practice, different trial aberrations are sequentially applied with the DM until the image quality metric attains its maximum.

General optimization algorithms can be used to maximize the image quality metric [19–24]. However, because these algorithms have no prior knowledge about the metric, a large number of trial aberrations must be evaluated before the metric is maximized [3,25,26]. Reducing the number of trial aberrations is critical in achieving short image acquisition times and in limiting side effects, such as photobleaching and phototoxicity. For small aberrations, the response of the image quality metric can be approximated using a simple model, such as a quadratic polynomial [25–35]. Model-based aberration correction algorithms [25–38] exploit the knowledge about the model of the metric to quicken the aberration correction, thus curtailing the aforementioned side effects.

In this paper, we investigate applying a model-based wavefront sensorless aberration correction algorithm to a second-harmonic microscope.

The paper is organized as follows. In Section 2, we discuss the definition of the basis functions for the control of the

DM. In Section [3], we outline our proposed algorithm for the aberration correction. In Section [4], we report the experimental results. The conclusions are drawn in Section [5].

2. DEFINITION OF THE BASIS FUNCTIONS FOR THE CONTROL OF THE DEFORMABLE MIRROR

The basis functions should satisfy two different requirements. In scanning microscopy, the aberration correction should not introduce x-tilt, y-tilt, and defocus Zernike aberrations [39]. These aberrations do not affect the image quality. Instead, they affect the position of the focal point within the specimen and they induce translations or distortions in the acquired images [31,32,40]. For this reason, a first requirement is that the basis functions be orthogonal to the x-tilt, y-tilt, and defocus aberrations.

A second requirement is that the basis functions express the capabilities of the DM in an accurate and concise form, by taking into account the mechanical limitations of the mirror and the misalignment in the optical system as much as possible. This requirement is not satisfied when using Zernike polynomials as the basis functions since a DM with N_a actuations cannot accurately induce a set of N_a Zernike polynomials.

We now discuss a simple procedure to define a new set of basis functions that satisfy the two requirements discussed so far. This procedure is based on the singular value decomposition (SVD) of a matrix H that approximately describes a linear relationship between the control signals of the DM and a set of Zernike coefficients. For completeness, we first report how H can be computed from input–output measurements.

A. Computation of Matrix H from Input–Output Measurements

Let N_a be the number of actuators of the DM. Assuming that the DM is a linear device, the phase aberration $\Phi(\boldsymbol{\xi})$ is given by the superposition of the influence functions [41,42] $\psi_i(\boldsymbol{\xi})$ of each actuator, where $\boldsymbol{\xi}$ is the spatial coordinate in the pupil and u_i is the control signal of the ith actuator:

$$\Phi(\boldsymbol{\xi}) = \sum_{i=1}^{N_a} u_i \psi_i(\boldsymbol{\xi}). \qquad (1)$$

For a suitable number N_z of Zernike polynomials $\mathcal{Z}_j(\boldsymbol{\xi})$, the phase aberration is approximated by

$$\Phi(\boldsymbol{\xi}) \approx \sum_{j=2}^{1+N_z} z_j \mathcal{Z}_j(\boldsymbol{\xi}), \qquad (2)$$

where z_j is the jth Zernike coefficient. We neglect the piston mode $\mathcal{Z}_1$ since this does not affect the image and assume that $\Phi(\boldsymbol{\xi})$ and $\psi_i(\boldsymbol{\xi})$ have zero mean value over the pupil.

The coefficients u_i and z_j are collected, respectively, into vectors $\mathbf{u} \in \mathbb{R}^{N_a}$ and $\mathbf{z} \in \mathbb{R}^{N_z}$. By considering a grid defined in the pupil, N_c samples of $\Phi(\boldsymbol{\xi})$ are collected into a vector $\boldsymbol{\phi} \in \mathbb{R}^{N_c}$. Similarly, we evaluate $\psi_i(\boldsymbol{\xi})$ and $\mathcal{Z}_j(\boldsymbol{\xi})$ over the grid and define two matrices $\Psi \in \mathbb{R}^{N_c \times N_a}$ and $Z \in \mathbb{R}^{N_c \times N_z}$. Using Eqs. ([1]) and ([2]), we find $\boldsymbol{\phi} = \Psi \mathbf{u}$ and $\boldsymbol{\phi} \approx Z \mathbf{z}$.

We would like to recover a matrix H which maps an actuation vector $\mathbf{u}$ into the corresponding vector of Zernike coefficients $\mathbf{z}$, i.e., $\mathbf{z} \approx H\mathbf{u}$. H can be computed using input–output measurements, so that the misalignment in the optical system is accounted for. Using a Shack–Hartmann wavefront sensor or interferometric methods [42–45] one can collect a set of measurements of the phase $\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_D$ corresponding to different settings of the DM $\mathbf{u}_1, \dots, \mathbf{u}_D$.

We compute H by minimizing the following criterion:

$$\min_H \sum_{i=1}^{D} \|\boldsymbol{\phi}_i - ZH\mathbf{u}_i\|^2. \qquad (3)$$

Setting the derivative with respect to H to zero leads to the following normal equation:

$$Z^T Z H \left(\sum_{i=1}^{D} \mathbf{u}_i \mathbf{u}_i^T \right) = Z^T \left(\sum_{i=1}^{D} \boldsymbol{\phi}_i \mathbf{u}_i^T \right), \qquad (4)$$

which can be solved by multiplying from the left and from the right by the inverse matrices of $Z^T Z$ and $\sum_{i=1}^{D} \mathbf{u}_i \mathbf{u}_i^T$. For a properly defined grid, the inverse of $Z^T Z$ exists, since Z is full column rank due to the orthogonality property of the Zernike polynomials. Additionally, vectors $\mathbf{u}_i$ can be selected so that $\sum_{i=1}^{D} \mathbf{u}_i \mathbf{u}_i^T$ is full rank.

In our system we have $N_a = 17$ and $N_c = 75912$. We performed $D = 4N_a$ measurements of the phase $\boldsymbol{\phi}_1, \dots, \boldsymbol{\phi}_{4N_a}$. In each measurement, a single actuator is poked while the other actuators are at rest. We empirically chose $D = 4N_a$, other choices are possible provided $D \geq N_a$. The choice of N_z is more critical. With a poor choice of N_z, the accuracy requirement discussed at the beginning of Section [2] may not be fulfilled and the approximation $\mathbf{z} \approx H\mathbf{u}$ may be too rough. We chose $N_z = 35$ by evaluating the error in approximating the phase measurements $\boldsymbol{\phi}_i$ using an increasing number of Zernike polynomials.

B. SVD-Based Removal of the *x*-Tilt, *y*-Tilt, and Defocus Aberrations

From the previous section, we conclude that, in our system, the $N_a = 17$ influence functions approximately span a subspace of the space spanned by the first $N_z = 35$ Zernike polynomials. Because $\text{rank}(H) < N_z$, there exist nonzero vectors $\mathbf{z}$ that do not belong to the range of H and the Zernike polynomials should not be used as the basis functions for the control of the DM.

We can split vector $\mathbf{z}$ and matrix H so that $\mathbf{z} \approx H\mathbf{u}$ is partitioned as

$$\begin{bmatrix} \mathbf{z}_l \\ \mathbf{z}_h \end{bmatrix} \approx \begin{bmatrix} H_l \\ H_h \end{bmatrix} \mathbf{u}, \qquad (5)$$

where the x-tilt, y-tilt, and defocus coefficients are collected into $\mathbf{z}_l = [z_2, z_3, z_4]^T$. The SVD of H_l is

$$H_l = U_l [\Sigma_l \quad \mathbf{0}] \begin{bmatrix} V_{l1}^T \\ V_{l2}^T \end{bmatrix}, \qquad (6)$$

where Σ_l has dimensions 3×3, V_{l1} has dimensions $N_a \times 3$, and V_{l2} has dimensions $N_a \times N_p$ with $N_p = N_a - 3$. The required constraint that $\mathbf{z}_l \approx \mathbf{0}$ is enforced if we choose $\mathbf{u}$ such that $H_l \mathbf{u} = 0$. This is achieved by parametrizing $\mathbf{u}$ using the

columns of V_{l2}, i.e., letting $\mathbf{u} = V_{l2}\mathbf{p}$, where $\mathbf{p} \in \mathbb{R}^{N_p}$. Therefore, the phase aberration is

$$\Phi(\boldsymbol{\xi}) = \sum_{i=1}^{N_a} u_i \psi_i(\boldsymbol{\xi}), \qquad \text{s.t. } \mathbf{u} = V_{l2}\mathbf{p} \qquad (7)$$

or equivalently

$$\Phi(\boldsymbol{\xi}) = \sum_{i=1}^{N_p} p_i \omega_i(\boldsymbol{\xi}), \qquad (8)$$

where $\omega_i(\boldsymbol{\xi})$ are the new basis functions. These functions are defined by $\omega_i(\boldsymbol{\xi}) = \sum_{j=1}^{N_a} (V_{l2})_{ji}\psi_j(\boldsymbol{\xi})$, where $(V_{l2})_{ji}$ denotes the element of V_{l2} at position (j, i). For a given vector $\mathbf{p}$, we can compute the control signals of the actuators with $\mathbf{u} = V_{l2}\mathbf{p}$. Similarly, for a given $\mathbf{p}$, the Zernike analysis of the induced wavefront aberration is given by $\mathbf{z} \approx HV_{l2}\mathbf{p}$.

In our experiments, we also applied regularization [46] by truncating the SVD of HV_{l2} to $U_1\Sigma_1V_1^T$. Using no more than 80% of the sum of the singular values, Σ_1 was a 7×7 matrix and the DM was controlled with a vector $\mathbf{r} \in \mathbb{R}^N$, where $N = 7$. For a given $\mathbf{r}$, the control signals of the actuators of the DM are computed using $\mathbf{u} = V_{l2}V_1\mathbf{r}$. The Zernike analysis of the induced wavefront aberration is computed using $\mathbf{z}_h \approx H_h V_{l2}V_1\mathbf{r}$ and the rms of the phase profile is given by computing the 2-norm, i.e., $\|\mathbf{z}_h\|$. This is equivalent to applying another parametrization to Eq. (8). We remark that in this way, no pseudoinverse is ever computed or used to control the DM, differently from what is done in [41], for example.

3. LEAST-SQUARES ESTIMATION OF THE UNKNOWN ABERRATION

In this section we discuss the aberration correction algorithm. In [32,34,35], the authors show that, for small aberrations, the image quality metric can be modeled using a quadratic polynomial. We denote a measurement of the image quality metric at time instant k with $\tilde{y}_k$, so that

$$\tilde{y}_k = c_0 - (\mathbf{x} + \mathbf{r}_k)^T Q(\mathbf{x} + \mathbf{r}_k) + \epsilon_k, \qquad (9)$$

where c_0 and Q are the parameters of the quadratic polynomial. Matrix Q is a positive semi-definite matrix, i.e., $Q \succeq 0$ [26]. Vector $\mathbf{x}$ represents the unknown aberration whereas $\mathbf{r}_k$ accounts for the aberration induced by the DM. The term ϵ_k is a placeholder that collects both the uncertainty in modeling the image quality metric and the measurement noise, and as such it cannot be measured by definition. By including this term, a measurement $\tilde{y}_k$ can be set equal to the right-hand side of Eq. (9). Excluding ϵ_k, the right-hand side of Eq. (9) is referred to as the approximate image quality metric in [26]. The parameters c_0 and Q can be computed using the input–output measurements recorded in a calibration experiment described in Subsection 4.C or using the methods described in [31,47].

A. Definition of the Least-Squares Problem

The aberration correction is achieved by maximizing the image quality metric, i.e., by letting $\mathbf{r}_k = -\mathbf{x}$ in Eq. (9). For this reason, we must first estimate the unknown vector $\mathbf{x}$. This can be done by applying $m \geq N + 1$ trial aberrations $\mathbf{r}_1, \ldots, \mathbf{r}_m$

with the DM and by taking the corresponding measurements $\tilde{y}_1, \ldots, \tilde{y}_m$.

Collect $\epsilon_1, \ldots, \epsilon_m$ into a vector $\boldsymbol{\epsilon}$ and $\tilde{y}_1, \ldots, \tilde{y}_m$ into a vector $\tilde{\mathbf{y}}$. By stacking m instances of Eq. (9), we can define a vector-valued function $\mathbf{g}$,

$$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} c_0 - (\mathbf{x} + \mathbf{r}_1)^T Q(\mathbf{x} + \mathbf{r}_1) \\ \vdots \\ c_0 - (\mathbf{x} + \mathbf{r}_m)^T Q(\mathbf{x} + \mathbf{r}_m) \end{bmatrix}, \qquad (10)$$

such that

$$\tilde{\mathbf{y}} = \mathbf{g}(\mathbf{x}) + \boldsymbol{\epsilon}. \qquad (11)$$

The least-squares estimate of $\mathbf{x}$ is obtained by minimizing $\|\boldsymbol{\epsilon}\|^2$, i.e., by solving

$$\min_{\mathbf{x}} f(\mathbf{x}), \qquad (12)$$

where

$$f(\mathbf{x}) = \|\tilde{\mathbf{y}} - \mathbf{g}(\mathbf{x})\|^2. \qquad (13)$$

We note that if vector $\boldsymbol{\epsilon}$ follows a multivariate normal distribution with zero mean and covariance proportional to the identity matrix, solving Eq. (12) provides the maximum likelihood [48] estimate of $\mathbf{x}$.

B. Analysis of the Least-Squares Problem

Finding the global minimum of $f(\mathbf{x})$ appears to be nontrivial as $f(\mathbf{x})$ may be nonconvex. This is illustrated with a two-dimensional example in Fig. 1. Here, the contour plot of Eq. (13) is shown, when $m > N + 1$ measurements of $\tilde{y}$ are taken. The measurement noise is zero, i.e., $\boldsymbol{\epsilon} = \mathbf{0}$. Nevertheless, $f(\mathbf{x})$ is not convex and exhibits two critical points. In addition to the least-squares solution $\mathbf{x}^{ls}$ of Eq. (12), which is the global minimum and for which $f(\mathbf{x}^{ls}) = 0$, a local minimum $\mathbf{x}^{loc}$ is present. In case one uses $\mathbf{r}_k = -\mathbf{x}^{loc}$ to perform the aberration correction, then the residual aberration is not zero and the image quality metric is not maximized.
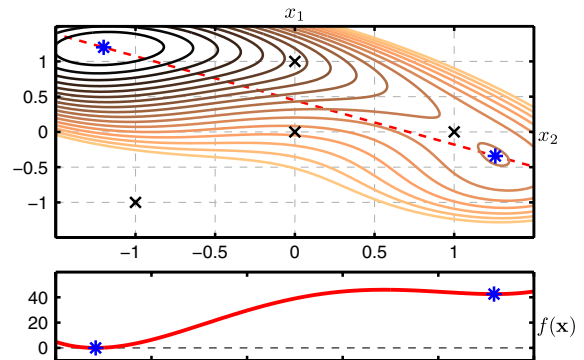

Fig. 1. Contour plot of Eq. (13). In this example, $f(\mathbf{x})$ is not convex and exhibits a local minimum. The parameters are $c_0 = 100$ and $Q = \begin{bmatrix} 1.25 & 0.433 \\ 0.433 & 1.25 \end{bmatrix}$. Four measurements of $\tilde{y}$, taken at $\mathbf{r}_1 = [0, 0]^T$, $\mathbf{r}_2 = [1, 0]^T$, $\mathbf{r}_3 = [0, 1]^T$, and $\mathbf{r}_4 = [0, -1]^T$ are marked with $\times$ symbols. The global minimum $\mathbf{x}^{ls} = [-1.2, 1.2]^T$ and the local minimum $\mathbf{x}^{loc} \approx [1.2582, -0.3421]^T$ are indicated with $*$ symbols. Isolines with an elevation greater than 70 have been removed for clarity. A cross section along the dashed line is reported in the plot in the bottom.

Because the convexity property is not satisfied in general, it is unclear how $\mathbf{x}^{\text{ls}}$ can be computed. For example, a gradient-based method applied to solve Eq. (12) may fail to compute $\mathbf{x}^{\text{ls}}$. Alternatively, more sophisticated algorithms may be unsuitable to meet the requirements of a real-time implementation. Nevertheless, the global solution of Eq. (12) can be computed efficiently even when $f(\mathbf{x})$ is not convex, as is outlined in the following section.

C. Efficient Computation of $\mathbf{x}^{\text{ls}}$

In [49], an efficient algorithm to find the global solution of a possibly nonconvex optimization similar to Eq. (12) was developed, in the context of localization problems. In this section we show how the solution proposed in [49] can be applied to our problem. For the remaining part of the paper we assume that Q is strictly positive definite. This assumption is reasonable since if there are aberrations that do not affect the image quality metric then these cannot be corrected and they should be neglected during the aberration correction [26].

Introducing an additional scalar variable α, we can reformulate Eq. (12) into the following equivalent constrained optimization:

$$\min_{\mathbf{x},\alpha} \sum_{k=1}^{m} (-\alpha - 2\mathbf{r}_k^T Q \mathbf{x} + c_0 - \mathbf{r}_k^T Q \mathbf{r}_k - \tilde{y}_k)^2$$
$$\text{s.t.} \qquad \alpha = \mathbf{x}^T Q \mathbf{x}. \tag{14}$$

Problem (14) is written concisely in matrix form as

$$\min_{\mathbf{w}} \|A\mathbf{w} - \mathbf{b}\|^2 \qquad \text{s.t.} \qquad \mathbf{w}^T D \mathbf{w} + 2\mathbf{f}^T \mathbf{w} = 0, \tag{15}$$

where

$$\mathbf{w}^T = [\,\mathbf{x}^T \quad \alpha\,], \qquad R = [\,\mathbf{r}_1 \quad \dots \quad \mathbf{r}_m\,],$$

$$A = [\,-2R^T Q \quad -\mathbf{1}\,], \qquad \mathbf{b} = \begin{bmatrix} \mathbf{r}_1^T Q \mathbf{r}_1 + \tilde{y}_1 - c_0 \\ \vdots \\ \mathbf{r}_m^T Q \mathbf{r}_m + \tilde{y}_m - c_0 \end{bmatrix},$$

$$D = \begin{bmatrix} Q & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix}, \qquad \mathbf{f}^T = [\,\mathbf{0} \quad -1/2\,]^T \tag{16}$$

and $\mathbf{1}$ and $\mathbf{0}$ denote vectors of appropriate dimensions where all components are, respectively, ones and zeros. The authors in [49] note that Eq. (15) is a generalized trust region subproblem [50]. Such problems, although nonconvex in general, have necessary and sufficient optimality conditions [50]. In particular, from [49,50], we know that $\mathbf{w}$ is a global minimizer of Eq. (15) if and only if there exist a Lagrange multiplier ν such that

$$(A^T A + \nu D)\mathbf{w} = A^T \mathbf{b} - \nu \mathbf{f}$$
$$\mathbf{w}^T D \mathbf{w} + 2\mathbf{f}^T \mathbf{w} = 0$$
$$A^T A + \nu D \succeq \mathbf{0}. \tag{17}$$

We assume that matrix A is full column rank, which in turn implies that $m \geq N + 1$. This assumption on A is by no means restrictive. Because $Q \succ 0$, it can be factored $Q = V\Delta V^T$,

where Δ is diagonal and full rank. Choose $R = [V, \mathbf{0}]$, where $\mathbf{0}$ is a vector of zeros, then A is full column rank. We further assume that the optimal Lagrange multiplier ν^* is such that $A^T A + \nu^* D$ is strictly positive definite. The authors in [49] point out that this more restrictive assumption could be removed with a more refined analysis. However, the case where ν^* is such that $A^T A + \nu^* D$ is not strictly positive definite is unlikely to occur both in theory and in practice [49].

Under these assumptions, one can compute

$$\mathbf{w}(\nu) = (A^T A + \nu D)^{-1}(A^T \mathbf{b} - \nu \mathbf{f}) \tag{18}$$

for a fixed value of ν. By replacing $\mathbf{w}$ in the second equation in Eq. (17) with the right-hand side of Eq. (18), one finds a univariate rational polynomial equation in ν:

$$\mathbf{w}(\nu)^T D \mathbf{w}(\nu) + 2\mathbf{f}^T \mathbf{w}(\nu) = 0. \tag{19}$$

The optimal Lagrange multiplier ν^* can be found examining the solutions of Eq. (19). From the assumption $A^T A + \nu D \succ 0$, it can be derived [49,50] that ν must be in the interval $(I_l, +\infty)$, where $I_l = -1/\lambda_{\max}(\Delta_A^{-1/2} V_A^T D V_A \Delta_A^{-1/2})$, and we used the factorization $A^T A = V_A \Delta_A V_A^T$. In addition, it is known [49,50] that Eq. (19) is strictly decreasing in ν within the considered interval. Therefore, the desired root ν^* of Eq. (19) can be found efficiently, for example via a bisection algorithm [49]. Once ν^* is found, the estimate of $\mathbf{x}$ is extracted from the first N components of $\mathbf{w}(\nu^*)$.

The aberration correction algorithm is therefore applied in the following manner. First, the data collection step takes place, whereby the $m \geq N + 1$ trial aberrations are applied and the corresponding measurements $\tilde{y}_1, \dots, \tilde{y}_m$ are taken. Then, ν^* is computed by finding the root of Eq. (19) within $(I_l, +\infty)$. The estimate $\mathbf{x}^{\text{ls}}$ of the aberration is found in the first N components of $\mathbf{w}(\nu^*)$. The second step involves applying the aberration correction with the DM, by letting $\mathbf{r} = -\mathbf{x}^{\text{ls}}$. These steps can be repeated in the following time instants by including more than m measurements to achieve a refined correction. We note that in [26], the least-squares estimate of $\mathbf{x}$ was not computed since the quadratic constraint in Eq. (14) was neglected to obtain a linear least-squares problem.

4. EXPERIMENTAL RESULTS

We implemented the model-based wavefront sensorless algorithm and report the experimental results in this section. Following [33], we employ the mean image intensity as a metric to correct aberrations in our second-harmonic microscope. Our experiments show successful aberration correction using this metric (see Subsection 4.G).

Our first purpose is to validate a previously proposed method [26] to compute the parameters c_0 and Q of Eq. (9) using input–output measurements. This validation has not been previously done in a realistic setting, since in [26] no microscope and no specimen were involved. Additionally, we intend to validate the aberration correction algorithm described in Subsection 3.C. We report our results in the following sections.

A. Description of the Experimental Setup

A schematic of the experimental setup is shown in Fig. 2. The source is a Coherent Chameleon Ultra II Ti:Sa 140 fs pulsed,
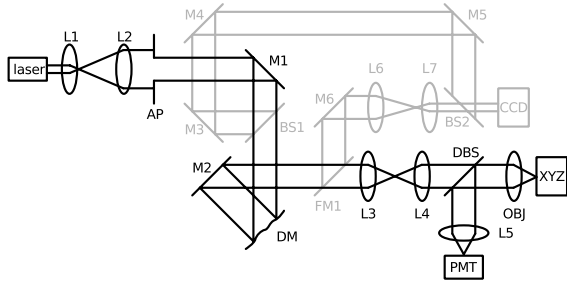
Fig. 2. Illustration of the optical setup. The components in black are used throughout the aberration correction experiments. The components in gray are used only for the initial characterization of the DM. A pulsed laser beam is expanded with lenses L1 and L2, clipped by aperture AP, and reflected by flat mirror M1 onto the DM. The DM is in an image of the back aperture of the microscope objective (OBJ), using lenses L3 and L4. The DM is illuminated under an angle of about 10° using the flat mirrors M1 and M2. The microscope objective (OBJ) focuses the light onto the specimen, which is supported by an xyz stage (XYZ). The second-harmonic signal emitted from the focal point inside the specimen is collected with the objective and separated from the illumination beam using a dichroic beam splitter (DBS). The emitted signal is focused by lens L5 onto a photomultiplier tube (PMT). For characterizing the DM, the surface of the DM is reimaged onto a CCD camera (CCD) using the flip mirror FM1, flat mirror M6, and lenses L6 and L7. A reference arm is created using beam splitter BS1, flat mirrors M3, M4, M5, and beam splitter BS2. A coherence-gated fringe analysis method described elsewhere [6] is applied to the fringe pattern generated onto the CCD.

near-infrared laser, with a beam diameter of 1.2 mm. This beam is expanded to a 14 mm wide beam by lenses L1 and L2.

The beam is stopped down to 9.5 mm (AP) before it is reflected under an angle of about 10° by the DM (Okotech, 17-channel micromachined DM with tip-tilt stage). The DM is reimaged one-to-one onto the objective back-aperture by lenses L3 and L4 (focal lengths, 300 mm). Because the DM can only introduce negative deflection, we bias the mirror so that we can apply positive and negative deflections to correct the wavefront (see [42]). In addition, the relationship between the control signal u_i of each actuator of the DM and the voltage applied to the corresponding electrode is quadratic, so that a linear displacement of the membrane is expected [42]. Due to this bias, the collimated beam coming from L2 is converging after being reflected by the DM. We corrected this by using lenses L4 and L3, so that a collimated beam is fed into the objective.

The sample is mounted on an xyz-piezo stage (XYZ; PI, Nanocube P-611.3S). The second-harmonic signal from the specimen is collected by the objective and split off by a 705 nm cut-off dichroic beam splitter (DBS; Semrock, FF705-Di01-25×36). This light is focused onto a 600 μm multimode fiber that is connected to a photomultiplier tube (PMT; Hamamatsu, GaAsP photocathode H7422P-40). The objective used (OBJ) is a 40 × /0.9 NA Nikon air objective with spherical correction collar. We manually adjusted the collar to correct for the spherical aberration due to the cover glass and the specimen at the selected depth.

For characterization of the DM, we interfere a tilted reference beam with a sample beam deflected off the DM to create fringes that encode the wavefront deformation. To allow this, a 50/50 beam splitter (BS1; Thorlabs, BS016) splits off part of the light into a reference arm beam, which is relayed onto the camera (CCD; AVT, Guppy Pro F-033b) by mirrors M3, M4, and M5. The sample arm beam is deflected by the DM once

before flip mirror FM1 directs the light into the calibration arm. Lenses L6 and L7 reimage the DM onto a camera. For this calibration we use the alignment laser, which is a continuous wave. We used the method described in [6] to decode the wavefront from the fringe patterns.

The piezo stage is controlled with a data acquisition board (National Instruments, PCI-e 6259) on a Windows computer running LabView. The DM is controlled through a PCI DAC card on a Linux computer running MATLAB and custom C code.

In the aberration correction experiments we imaged collagen fiber extracted from rat tail washed four times in distilled water. Following fixation in 4% paraformaldehyde, the fiber was washed in phosphate buffer saline and then embedded in 3% agarose (Sigma-Aldrich chemie GmbH) in a 35 mm glass bottom dish (MatTek Corporation). We used 900 nm excitation light to generate the second-harmonic signal.

B. Preparation of the Experiments

We first imaged a 20 μm × 20 μm region, approximately 33 μm deep into the collagen fiber. The region is labeled with A in Fig. 3. The influence of the size of the region used for the aberration correction has been studied elsewhere [35]. For a certain setting of the DM **r**, the region is scanned using the xyz stage. The corresponding value of the image quality metric $\tilde{y}$ is measured as the mean image intensity [32–35], i.e., the mean pixel value recorded over the region. The pixel dwell time is 0.5 ms and the sampling is 24 pixels × 24 pixels. With these settings, the xyz stage does not reach the full 20 μm distance in the x scanning direction, which is the fast axis. This was not an issue since such a coarse sampling was only used to perform the aberration correction in a short time [34]. The final images taken after the aberration correction were recorded with a higher sampling. The image deformation due to both the nonlinearity and nonuniform speed of the xyz stage were removed from the final images, using interpolation and the signals recorded with the position sensors of the xyz stage.

First, the static aberrations in the system due to misalignment and imperfections in the optical components were corrected. We used the nonzero initial aberration that was found during the calibration of the DM in Section 2 (about
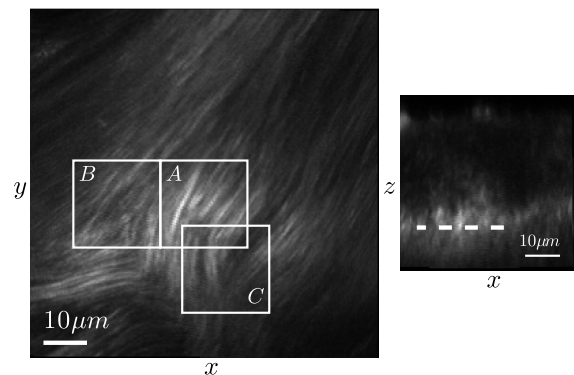

Fig. 3. Cross sections of rat tail collagen fiber used in our experiments. The smaller image on the right-hand side is an xz cross section (50 μm × 50 μm, 128 pixels × 128 pixels). The dashed line denotes an xy cross section (80 μm × 80 μm, 256 pixels × 256 pixels) approximately 33 μm deep, which is shown on the left-hand side. Three different 20 μm × 20 μm regions are marked.

0.79 rms rad at 900 nm, mostly astigmatism). At this point the Nelder–Mead algorithm [51] was executed four times to find a value **r** that maximizes $\tilde{y}$. Unfortunately, this led to the saturation of two actuators, indicating that the stroke of the DM may be insufficient to completely suppress the aberration in this region. We selected a slightly suboptimal vector **r** from the vectors generated by the Nelder–Mead algorithm. For the selected vector, the maximum normalized voltage of the actuators was 0.72, i.e., $\|\mathbf{u}\|_\infty \leq 0.72$, $\tilde{y}$ improved by 3% and a total aberration of about 0.18 rms rad was suppressed. This state was used as the new initial condition for the rest of the experiments, i.e., $\mathbf{r} = \mathbf{0}$ is mapped to this setting of the DM. In the following sections, all the units in rad are referenced to the 900 nm excitation laser light.

C. Computation of the Parameters of the Quadratic Polynomial Using Input–Output Measurements

We executed the computation of the parameters of the quadratic polynomial used for modeling the image quality metric multiple times. Each time, the sequence of input vectors consisted of two subsequences. The first subsequence contained random vectors $\mathbf{r}_1, \ldots, \mathbf{r}_{250}$ and was used for the validation and the cross-validation. The second subsequence contained 70 fixed vectors (each vector having a single nonzero component). The second subsequence was included because the 250 random vectors may be insufficient to uniformly sample the N-dimensional space of the inputs. In [26], 3750 random vectors were used, but this was impractical here, due to the time necessary to move the xyz stage. The maximum rms rad of the input aberrations did not exceed 0.81. This value was empirically tuned by examining the goodness of fit as a function of the maximum rms [26].

For each input vector in the sequence, we measured the corresponding output of $\tilde{y}$. The resulting input–output data, i.e., collections $\tilde{y}_1, \ldots, \tilde{y}_{320}$ and $\mathbf{r}_1, \ldots, \mathbf{r}_{320}$ were used to formulate the following optimization problem [26]:

$$\min_{\tilde{\mathbf{x}}} \|\tilde{A}\,\tilde{\mathbf{x}} - \tilde{b}\| \qquad \text{s.t.}$$

$$\tilde{A} = \begin{bmatrix} 1 & \mathbf{r}_1^T & -\mathbf{r}_1^T \otimes \mathbf{r}_1^T \\ \vdots & \vdots & \vdots \\ 1 & \mathbf{r}_{320}^T & -\mathbf{r}_{320}^T \otimes \mathbf{r}_{320}^T \end{bmatrix},$$

$$\tilde{b} = [\,\tilde{y}_1 \quad \ldots \quad \tilde{y}_{320}\,]^T,$$

$$\tilde{\mathbf{x}} = [\,c_0 \quad \mathbf{c}_1^T \quad \mathrm{vec}(Q)^T\,]^T,$$

$$Q \succeq 0, \qquad\qquad\qquad\qquad (20)$$

where $\mathrm{vec}(\cdot)$ denotes the vectorization operation and $\otimes$ the Kronecker product. This program was solved using CVXOPT [52] (see [53] for further details).

D. Validation and Cross-Validation of the Computed Parameters

The results of applying Eq. (20) in region A marked in Fig. 3 are shown in Fig. 4. The computation of the parameters was repeated six times. Each time, a new input–output dataset was acquired, $\mathcal{D}_1, \ldots, \mathcal{D}_6$. For each input–output dataset, optimization (20) was solved generating six sets of parameters, each set comprising c_0, $\mathbf{c}_1$, and Q. The sets are denoted as $\mathcal{M}_1, \ldots, \mathcal{M}_6$. We quantified the goodness of fit for all combi-
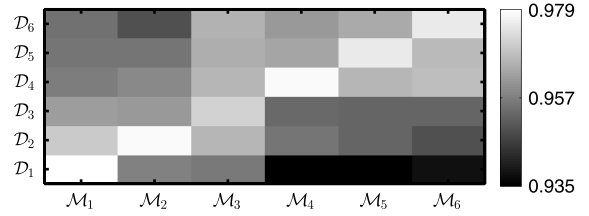

Fig. 4.   Validations and cross-validations of the computation of c_0, $\mathbf{c}_1$, and Q using Eq. (20). The computation has been performed six times in region A in Fig. 3. $\mathcal{D}_i$ denotes the input–output data taken during the ith time. $\mathcal{M}_i$ denotes the set of parameters [c_0, $\mathbf{c}_1$, and Q in Eq. (20)] computed from $\mathcal{D}_i$. For each combination $\mathcal{M}_i$ and $\mathcal{D}_j$, the ith random input subsequence and $\mathcal{M}_j$ are used to compute the predicted output $\hat{\mathbf{o}} \in \mathbb{R}^{250}$. Each rectangle reports the goodness of fit [$\mathcal{R}^2$, see Eq. (21)] computed comparing $\hat{\mathbf{o}}$ with the corresponding measured output $\mathbf{o} \in \mathbb{R}^{250}$ of $\mathcal{D}_j$. A value of one for the goodness of fit indicates that the model fits the data without error. High values of the goodness of fit are reported in all combinations, showing that Eq. (20) is a robust method to compute the parameters.

nations of $\mathcal{D}$ and $\mathcal{M}$ by means of the $\mathcal{R}^2$ indicator. Using the random input subsequence of $\mathcal{D}_i$ and $\mathcal{M}_j$, we computed the predicted output $\hat{\mathbf{o}} \in \mathbb{R}^{250}$ of the image quality metric. The input–output data points obtained from the deterministic input subsequence of 70 vectors were discarded and were not included in the computation of the $\mathcal{R}^2$ indicator, which is obtained using the following equations:

$$\mathcal{R}^2 = 1 - S_r/S_t, \qquad S_r = \|\mathbf{o} - \hat{\mathbf{o}}\|^2,$$

$$S_t = \|\mathbf{o} - \bar{o}\mathbf{1}\|^2, \qquad \bar{o} = (1/250)\mathbf{1}^T\mathbf{o}, \qquad (21)$$

where $\mathbf{o}$ is the measured output of $\mathcal{D}_i$. An $\mathcal{R}^2 = 1$ implies a perfect fit of the experimental data.

Figure 4 reports $\mathcal{R}^2$ indicators that are close to one, implying a good fit of the experimental data. A good fit is also found for the combinations that are off the main diagonal. Here the parameters c_0, $\mathbf{c}_1$, and Q allow to accurately predict cross-validation output data. Similar results were found for the other two regions marked in Fig. 3.

E. Correction of the Residual Aberration

In this section we apply the aberration correction algorithm described in Subsection 3.C. First, we attempt to further reduce the residual aberration in region A, which is marked in Fig. 3. Some aberration may not have been completely suppressed by the Nelder–Mead algorithm, which was applied to region A in Subsection 4.B. We therefore expect no improvement or a small improvement in region A. Second, we apply the aberration correction to regions B and C, where the Nelder–Mead algorithm was not applied. Here, we expect some improvement, provided that the aberrations found in regions B and C are different from the aberration found in region A.

We take the parameters c_0, $\mathbf{c}_1$, and Q that were computed using $\mathcal{D}_6$ in Subsection 4.D. In order to apply the algorithm, the following modified parameters need to be used, i.e., $c_0' = c_0 + (1/4)\mathbf{c}_1^T Q^{-1}\mathbf{c}_1$, $\mathbf{c}_1' = \mathbf{0}$, and $Q' = Q$. This modification is necessary since, for simplicity, in Section 3 we neglected the linear term $\mathbf{c}_1$. This term corresponds to the aberration that is present when computing the parameters of the quadratic polynomial, see [26] for further details. The aberration correction experiment is applied in the three
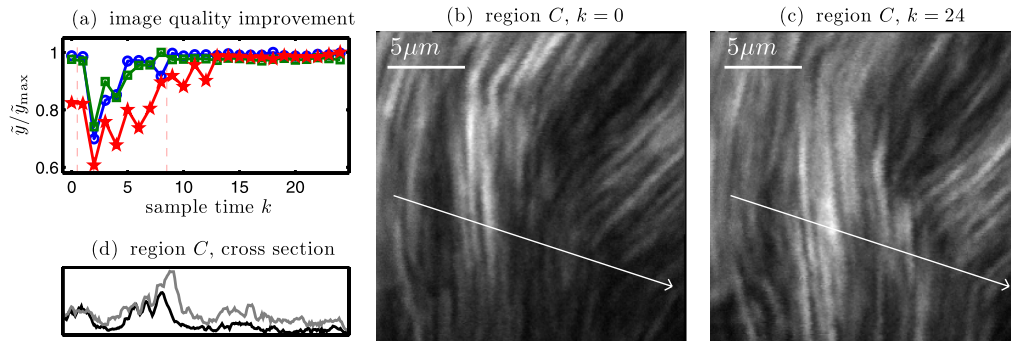
Fig. 5. Summary of three aberration correction experiments. (a) Evolution of the normalized image quality metric. The experiments were performed in region A (curve with $\bigcirc$ markers), B (curve with $\square$ markers), and C (curve with $\star$ markers), which are marked in Fig. 3. For each region, the corresponding parameters computed by solving Eq. (20) were used. $\tilde{y}_{\max}$ is the maximum measurement of $\tilde{y}$ in each region. The estimated rms rad of each aberration is 0.38 for region A, 0.37 for region B, and 1.27 for region C. (b) 256 pixels × 256 pixels image of region C at sample time $k = 0$. (c) 256 pixels × 256 pixels image of region C at sample time $k = 24$. (d) cross sections taken along the arrows marked in (b) and (c), black for (b) and gray for (c).

regions using the corresponding parameters for each region. A summary of the results is given in Fig. 5.

In Fig. 5(a), the normalized measurements of the image quality metric are reported for region A (curve with ∘ markers), region B (curve with $\square$ markers), and region C (curve with $\star$ markers). The measurements are normalized using the corresponding maximum recorded measurement $\tilde{y}_{\max}$ in each region. The initial value of $\tilde{y}$ is reported at sample time $k = 0$. This measurement is not supplied to the aberration correction algorithm. The data collection step is performed between time $k = 1$ and $k = 8$ inclusive, where $N + 1 = 8$ trial aberrations are applied. From time $k = 9$ onward, the aberration correction step is applied.

As expected, a marginal improvement is found in region A (curve with ∘ markers), where an aberration of about 0.38 rms rad is corrected. The rms of each aberration is estimated using $\|H_h V_{l2} V_1 \mathbf{x}^{ls}\|$, adjusted for the 900 nm excitation light. Also in region B (curve with $\square$ markers), a small aberration of about 0.37 rms rad is corrected. In region C (curve with $\star$ markers), an estimated 1.27 rms rad aberration is corrected, leading to an improvement of 20% of the image quality metric. Nevertheless, two actuators of the DM are saturated.

Two 256 pixels × 256 pixels images of region C are reported in Figs. 5(b) and 5(c). These images are recorded before [$k = 0$, Fig. 5(b)] and after [$k = 24$, Fig. 5(c)] the aberration correction. The cross sections marked in the images are reported in a single graph in Fig. 5(d). The image taken at time $k = 24$ is 18% brighter and shows finer detail in the bottom and right parts. Here some structure of the fiber was not visible at time $k = 0$. The improvement is less clear when examining the left and top parts of the region. One possible reason for the variability of the improvement is that the aberration is not spatially invariant over the considered region. We also note that the applied correction was not optimal, due to the saturation of two actuators of the DM. We conclude by observing that this improvement after the aberration correction is compatible with what was achieved by running four iterations of the Nelder–Mead algorithm in Subsection 4.B.

F. Validation of the Aberration Correction Algorithm

To assess whether the aberration correction algorithm is effectively removing aberration we performed a different kind of experiment. First, we introduce a known amount of aberration using the DM. We then apply the aberration correction algorithm to suppress this aberration. The algorithm is not supplied with any information about the known aberration. Finally, we evaluate the residual aberration by comparing the estimate of the aberration provided by the algorithm with the known aberration. This experiment is commonly employed in the literature to assess the effectiveness of the aberration correction [26,29,34,35].

Figure 6 reports a summary of the correction of 20 random aberrations introduced with the DM in region A. The upper plot in Fig. 6 shows some statistical indicators of the normalized measurements of the image quality metric.
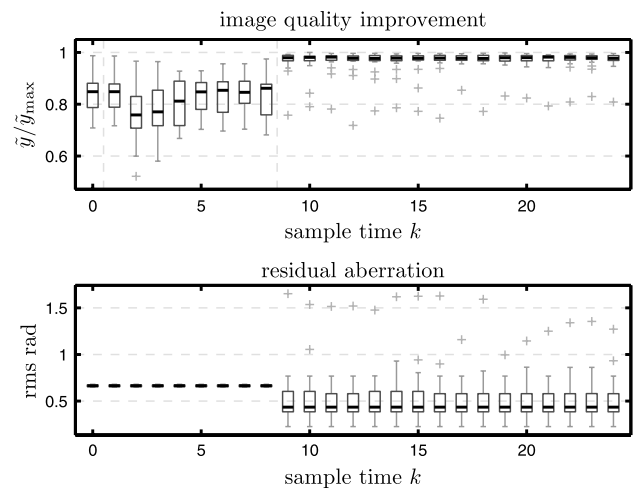

Fig. 6. Summary of the correction of 20 random aberrations induced by the DM in region A, which is marked in Fig. 3. The upper plot reports the normalized measurements of the image quality metric. The measurements are normalized using the maximum measurement $\tilde{y}_{\max}$ that is recorded throughout the 20 experiments. At time $k = 0$ the initial value of $\tilde{y}$ is reported; this data point is not supplied to the aberration correction algorithm. Between time $k = 1$ and $k = 8$, the data collection step is executed. From time $k = 9$ onward, the aberration correction step is applied. A statistical analysis is made at each time instant using the function boxplot from MATLAB. The tops and bottoms of the rectangles denote the 25th and 75th percentiles, the horizontal lines in the middle of the rectangles denote the medians, and the whiskers extend to the furthest measurements not considered as outliers. The + symbols denote single outliers. The same statistical analysis is performed for the residual aberration, and the results are shown in the lower plot.
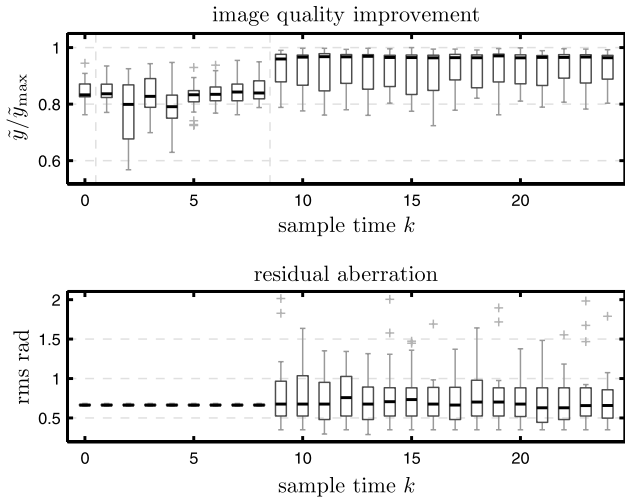
Fig. 7.   Summary of the correction of 20 random aberrations induced by the DM in region B. See the caption of Fig. 6 for a legend of the plots.

The measurements have been normalized using the maximum measurement of the metric $\tilde{y}_{\max}$ that is recorded throughout the 20 experiments. The median, 25th, and 75th percentiles are computed in each time instant, see the caption of Fig. 6 for a detailed legend. The same analysis has been made for the residual aberration and is reported in the lower plot in Fig. 6. The rms of the residual aberration is computed as the rms of the difference between the known aberration introduced by the DM and the respective estimate provided by the algorithm. From this figure, we conclude that the image quality metric is consistently maximized, as the median is close to 1 after the aberration correction is applied from sample time $k = 9$ onward. This is consistent with the reduction in the residual aberration reported in the lower plot.

The same experiments were performed in regions B and C. In both cases we used the corresponding modified parameters, computed using $\mathcal{D}_6$ in Subsection 4.D. The results are reported in Figs. 7 and 8. Whereas, the results for region C are similar to the results obtained in region A, the results
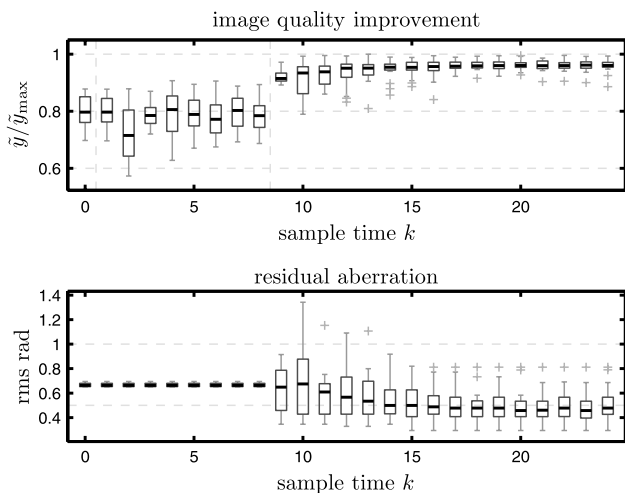

Fig. 8.   Summary of the correction of 20 random aberrations induced by the DM in region C. See the caption of Fig. 6 for a legend of the plots.

in region B do not show a good performance, since the medians of the residual aberration are comparable with the initial aberration before the correction.

Out of the 20 trials in region B, we report, respectively, the ones that resulted in the maximum and in the minimum improvement of $\tilde{y}$ in Fig. 9. In Fig. 9(a), some fine structure of the fiber is more visible after the correction, which is compatible with a successful aberration correction. On the other hand, in Fig. 9(b), the aberration correction failed, as both the image after the correction is visually worse and the intensity is slightly decreased.

The experiments resulting in the maximum and in the minimum improvement of $\tilde{y}$ in region C are also reported in Fig. 10. In Fig. 10(a), a successful aberration correction is shown, with a clear maximization of $\tilde{y}$ and a noticeable improvement in the contrast of the image after the correction. In Fig. 10(b), the improvement is more marginal.

G. Analysis of the Experimental Results

We computed the correlation among different quantities to concisely assess the results of the 60 aberration correction experiments reported in Figs. 6, 7, and 8. Considering the last time instant $k = 24$, we set up a saturation indicator variable $s_1, \ldots, s_{60}$ that is 1 if saturation of some actuators of the DM occurred and -1 otherwise.

We computed a correlation of -0.2477 between the normalized measurement of the image quality metric and the saturation indicator variable. These two quantities are slightly inversely correlated, meaning that saturation of some actuators negatively affected the final value of the image quality metric. The correlation between the final rms of the residual aberration and the saturation indicator variable was 0.3457. This positive correlation shows that a larger amount of residual aberration was found when the stroke of the DM was exhausted. Finally, we computed a correlation of -0.7388 between the normalized measurement of the image quality metric and the rms of the residual aberration, which confirms that some aberration is removed by maximizing the image quality metric. We conclude that saturation of the DM was an issue that hampered the results in our experimental validation.

H. Variations of the Parameters over the Field of View

We report spatial variations in the parameters c_0, $\mathbf{c}_1$, and Q. The parameters differed when computed, respectively, in regions A, B, and C in Fig. 3. Due to the variations, we were not able to apply the aberration correction algorithm using a single set of parameters, e.g., by correcting aberrations in regions B and C using the parameters computed from region A. Variations in the parameters represent a challenge for model-based aberration correction algorithms, since the parameters are computed once only using a calibration experiment [25–37].

Parameter $\mathbf{c}_1$ depends on the nonzero aberration that is present when collecting the input–output measurements used in Eq. (20). This parameter can be removed by applying the aberration correction algorithm as done in Subsection 4.E. Instead, c_0 is dependent on the maximum value of the image quality metric, which differed in the three regions. We found variations in Q. For example, the largest eigenvalue of Q varied by about 30% in region B and by about 22% in region C with
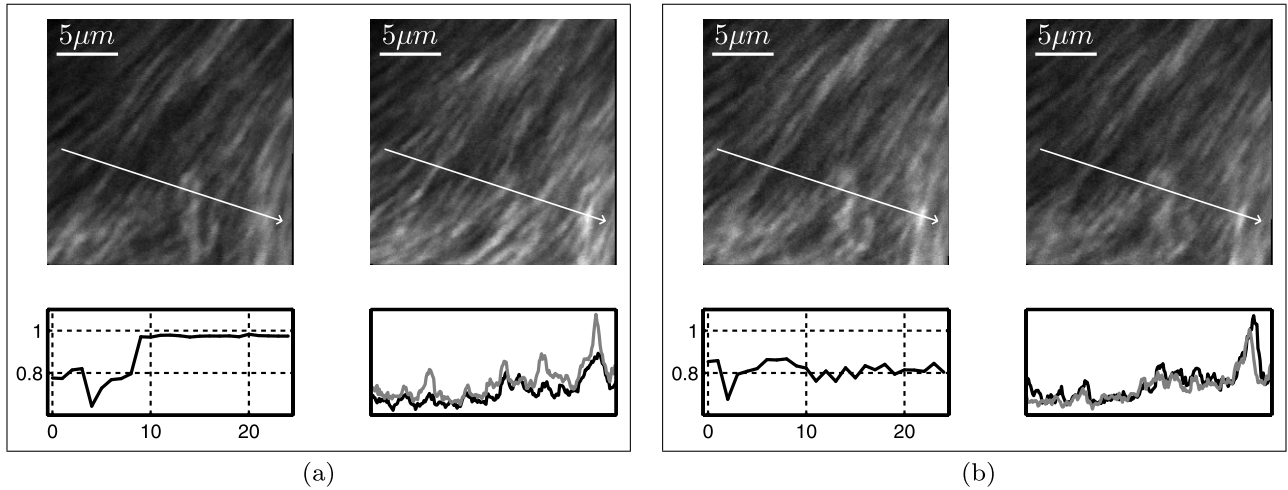
Fig. 9.   Two aberration correction experiments from the set of experiments reported in Fig. 7. These two experiments resulted, respectively, in (a) the maximum and (b) the minimum improvement of $\tilde{y}$. In both (a) and (b), a 256 pixels × 256 pixels image is taken before (on the left, $k = 0$) and after (on the right, $k = 24$) the aberration correction. The graphs in the bottom of (a) and (b) show, respectively, the evolution of the normalized metric (on the left) and the cross sections indicated by the arrows in the images (on the right). In the cross section graphs, the dark and the light lines correspond, respectively, to $k = 0$ and $k = 24$.
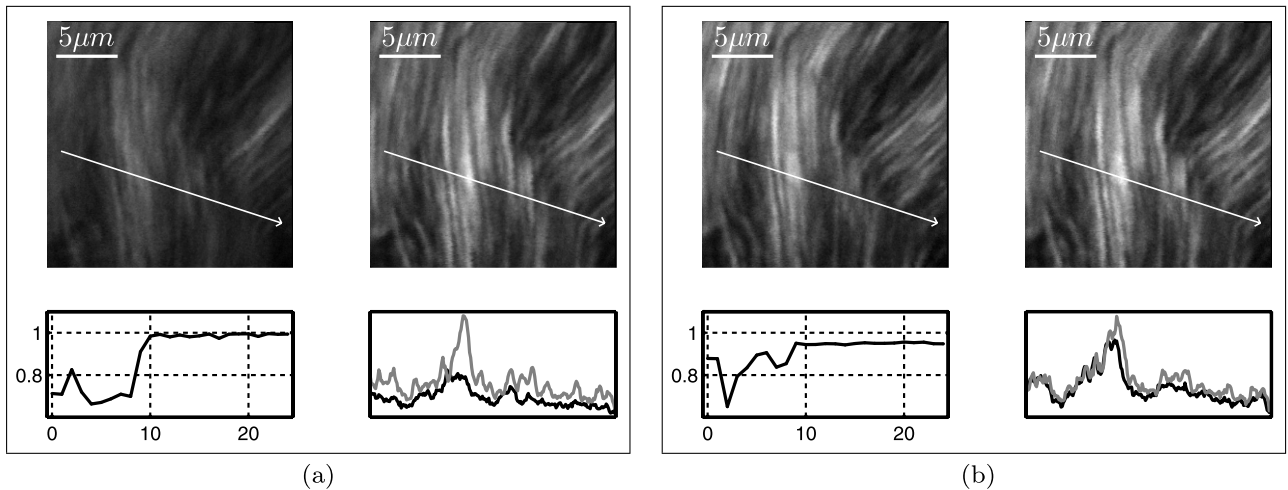

Fig. 10.   Two aberration correction experiments from the set of experiments reported in Fig. 8. These two experiments resulted, respectively, in (a) the maximum and (b) the minimum improvement of $\tilde{y}$. Refer to the caption of Fig. 9 for a detailed legend.

respect to its value in region A. The eigenvectors of Q were also affected. For instance, the eigenvector corresponding to the second largest eigenvalue of Q was rotated by about 7° in region B and by about 21° in region C with respect to its orientation in region A.

From Subsections 4.E and 4.F we conclude that the minimum number of measurements necessary to apply the aberration correction when c_0 and Q are known is $N + 1$, as also found in [25,26,30]. If the parameters c_0 and Q vary during the acquisition of different regions of the specimen, then additional measurements are necessary to update the parameters before the aberration correction can be applied. This is consistent with [29,31–35], where algorithms that use a minimum of $2N + 1$ measurements were employed. By approximating the solution of Eq. (12), these algorithms use the additional N measurements to estimate all the eigenvalues of Q each time the aberration correction is applied (see Section 4 in [26]). Nevertheless, variations in the orientations of the eigenvectors, such as the ones detected during our experi-

ments, are not accounted for and affect the accuracy of the aberration correction (see Section 3 in [34]). For these reasons, detecting variations and updating the parameters of the model as different regions of the specimen are acquired is an open research challenge.

5. CONCLUSIONS

In this paper we present a procedure to define a new set of basis functions for the control of the deformable mirror (DM). The new basis functions can be made approximately orthogonal to a set of Zernike polynomials. This is necessary for applying aberration correction in scanning microscopy applications, where the DM must not induce the x-tilt, y-tilt, and defocus aberrations.

The second contribution concerns the algorithm used for the aberration correction. We consider computing the least-squares estimate of the unknown aberration. Although this problem is nonconvex in general, the least-squares

estimate can be computed efficiently by exploiting results already applied in the solution of localization problems [49]. Once the estimate is computed, the aberration correction is applied by maximizing the image quality metric.

We implement the aberration correction algorithm in a second-harmonic microscope. First, we are able to compute the parameters of the quadratic polynomial used to model the image quality metric directly from input–output measurements, using a previously proposed method [26]. Second, we validate the aberration correction algorithm discussed in this paper. We also report the measurement of variations in the parameters of the quadratic polynomial over the field of view.

ACKNOWLEDGMENTS

REFERENCES

1. W. Denk, J. Strickler, and W. Webb, "Two-photon laser scanning fluorescence microscopy," Science **248**, 73–76 (1990).
2. P. J. Campagnola, A. Lewis, L. M. Loew, H. A. Clark, and W. A. Mohler, "Second-harmonic imaging microscopy of living cells," J. Biomed. Opt. **6**, 277–286 (2001).
3. M. J. Booth, "Adaptive optics in microscopy," Phil. Trans. R. Soc. A **365**, 2829–2843 (2007).
4. R. Tyson, *Principles of Adaptive Optics* (CRC, 2010).
5. S. A. Rahman and M. J. Booth, "Direct wavefront sensing in adaptive optical microscopy using backscattered light," Appl. Opt. **52**, 5523–5532 (2013).
6. T. I. M. van Werkhoven, J. Antonello, H. H. Truong, M. Verhaegen, H. C. Gerritsen, and C. U. Keller, "Snapshot coherence-gated direct wavefront sensing for multi-photon microscopy," Opt. Express **22**, 9715–9733 (2014).
7. T. van Werkhoven, H. Truong, J. Antonello, R. Fraanje, H. Gerritsen, M. Verhaegen, and C. Keller, "Coherence-gated wavefront sensing for microscopy using fringe analysis," Proc. SPIE **8253**, 82530E (2012).
8. M. Rueckel, J. A. Mack-Bucher, and W. Denk, "Adaptive wavefront correction in two-photon microscopy using coherence-gated wavefront sensing," Proc. Natl. Acad. Sci. USA **103**, 17137–17142 (2006).
9. J. W. Cha, J. Ballesta, and P. T. C. So, "Shack–Hartmann wavefront-sensor-based adaptive optics system for multiphoton microscopy," J. Biomed. Opt. **15**, 046022 (2010).
10. J. M. Bueno, E. J. Gualda, and P. Artal, "Adaptive optics multiphoton microscopy to study ex vivo ocular tissues," J. Biomed. Opt. **15**, 066004 (2010).
11. P. Artal, S. Marcos, R. Navarro, and D. R. Williams, "Odd aberrations and double-pass measurements of retinal image quality," J. Opt. Soc. Am. A **12**, 195–201 (1995).
12. O. Azucena, J. Crest, J. Cao, W. Sullivan, P. Kner, D. Gavel, D. Dillon, S. Olivier, and J. Kubby, "Wavefront aberration measurements and corrections through thick tissue using fluorescent microsphere reference beacons," Opt. Express **18**, 17521–17532 (2010).
13. R. Aviles-Espinosa, J. Andilla, R. Porcar-Guezenec, O. E. Olarte, M. Nieto, X. Levecq, D. Artigas, and P. Loza-Alvarez, "Measurement and correction of in vivo sample aberrations employing a nonlinear guide-star in two-photon excited fluorescence microscopy," Biomed. Opt. Express **2**, 3135–3149 (2011).
14. M. Shaw, K. O'Holleran, and C. Paterson, "Investigation of the confocal wavefront sensor and its application to biological microscopy," Opt. Express **21**, 19353–19362 (2013).
15. X. Tao, A. Norton, M. Kissel, O. Azucena, and J. Kubby, "Adaptive optical two-photon microscopy using autofluorescent guide stars," Opt. Lett. **38**, 5075–5078 (2013).
16. X. Tao, Z. Dean, C. Chien, O. Azucena, D. Bodington, and J. Kubby, "Shack–Hartmann wavefront sensing using interferometric focusing of light onto guide-stars," Opt. Express **21**, 31282–31292 (2013).
17. N. Ji, D. E. Milkie, and E. Betzig, "Adaptive optics via pupil segmentation for high-resolution imaging in biological tissues," Nat. Methods **7**, 141–147 (2009).
18. D. E. Milkie, E. Betzig, and N. Ji, "Pupil-segmentation-based adaptive optical microscopy with full-pupil illumination," Opt. Lett. **36**, 4206–4208 (2011).
19. G. Vdovin, "Optimization-based operation of micromachined deformable mirrors," Proc. SPIE **3353**, 902–909 (1998).
20. O. Albert, L. Sherman, G. Mourou, T. B. Norris, and G. Vdovin, "Smart microscope: an adaptive optics learning system for aberration correction in multiphoton confocal microscopy," Opt. Lett. **25**, 52–54 (2000).
21. M. A. Vorontsov, "Decoupled stochastic parallel gradient descent optimization for adaptive optics: integrated approach for wave-front sensor information fusion," J. Opt. Soc. Am. A **19**, 356–368 (2002).
22. P. Marsh, D. Burns, and J. Girkin, "Practical implementation of adaptive optics in multiphoton microscopy," Opt. Express **11**, 1123–1130 (2003).
23. L. Murray, J. C. Dainty, and E. Daly, "Wavefront correction through image sharpness maximization," Proc. SPIE **5823**, 40–47 (2005).
24. A. J. Wright, D. Burns, B. A. Patterson, S. P. Poland, G. J. Valentine, and J. M. Girkin, "Exploration of the optimisation algorithms used in the implementation of adaptive optics in confocal and multiphoton microscopy," Microsc. Res. Tech. **67**, 36–44 (2005).
25. M. Booth, "Wave front sensor-less adaptive optics: a model-based approach using sphere packings," Opt. Express **14**, 1339–1352 (2006).
26. J. Antonello, M. Verhaegen, R. Fraanje, T. van Werkhoven, H. C. Gerritsen, and C. U. Keller, "Semidefinite programming for model-based sensorless adaptive optics," J. Opt. Soc. Am. A **29**, 2428–2438 (2012).
27. M. A. A. Neil, M. J. Booth, and T. Wilson, "New modal wave-front sensor: a theoretical analysis," J. Opt. Soc. Am. A **17**, 1098–1107 (2000).
28. M. J. Booth, M. A. A. Neil, and T. Wilson, "New modal wave-front sensor: application to adaptive confocal fluorescence microscopy and two-photon excitation fluorescence microscopy," J. Opt. Soc. Am. A **19**, 2112–2120 (2002).
29. D. Débarre, M. J. Booth, and T. Wilson, "Image based adaptive optics through optimisation of low spatial frequencies," Opt. Express **15**, 8176–8190 (2007).
30. M. J. Booth, "Wavefront sensorless adaptive optics for large aberrations," Opt. Lett. **32**, 5–7 (2007).
31. D. Débarre, E. J. Botcherby, M. J. Booth, and T. Wilson, "Adaptive optics for structured illumination microscopy," Opt. Express **16**, 9290–9305 (2008).
32. D. Débarre, E. J. Botcherby, T. Watanabe, S. Srinivas, M. J. Booth, and T. Wilson, "Image-based adaptive optics for two-photon microscopy," Opt. Lett. **34**, 2495–2497 (2009).
33. A. Jesacher, A. Thayil, K. Grieve, D. Débarre, T. Watanabe, T. Wilson, S. Srinivas, and M. Booth, "Adaptive harmonic generation microscopy of mammalian embryos," Opt. Lett. **34**, 3154–3156 (2009).
34. A. Facomprez, E. Beaurepaire, and D. Débarre, "Accuracy of correction in modal sensorless adaptive optics," Opt. Express **20**, 2598–2612 (2012).
35. J. Zeng, P. Mahou, M.-C. Schanne-Klein, E. Beaurepaire, and D. Débarre, "3D resolved mapping of optical aberrations in thick tissues," Biomed. Opt. Express **3**, 1898–1913 (2012).
36. H. Song, R. Fraanje, G. Schitter, H. Kroese, G. Vdovin, and M. Verhaegen, "Model-based aberration correction in a closed-loop wavefront-sensor-less adaptive optics system," Opt. Express **18**, 24070–24084 (2010).
37. H. Linhai and C. Rao, "Wavefront sensorless adaptive optics: a general model-based approach," Opt. Express **19**, 371–379 (2011).
38. H. W. Yoo, M. Verhaegen, M. van Royen, and G. Schitter, "Automated adjustment of aberration correction in scanning confocal

microscopy," in *IEEE International Instrumentation and Measurement Technology Conference (I2MTC)* (IEEE, 2012), pp. 1083–1088.

39. R. J. Noll, "Zernike polynomials and atmospheric turbulence," J. Opt. Soc. Am. **66**, 207–211 (1976).

40. A. Thayil and M. Booth, "Self calibration of sensorless adaptive optical microscopes," J. Eur. Opt. Soc. **6**, 11045 (2011).

41. C. Paterson, I. Munro, and J. Dainty, "A low cost adaptive optics system using a membrane mirror," Opt. Express **6**, 175–185 (2000).

42. G. Vdovin, O. Soloviev, M. Loktev, and V. Patlan, *OKO Guide to Adaptive Optics*, 4th ed. (Flexible Optical BV, 2013).

43. B. Wang and M. J. Booth, "Optimum deformable mirror modes for sensorless adaptive optics," Opt. Commun. **282**, 4467–4474 (2009).

44. E. Fernandez and P. Artal, "Membrane deformable mirror for adaptive optics: performance limits in visual optics," Opt. Express **11**, 1056–1069 (2003).

45. M. Booth, T. Wilson, H.-B. Sun, T. Ota, and S. Kawata, "Methods for the characterization of deformable membrane mirrors," Appl. Opt. **44**, 5131–5139 (2005).

46. P. C. Hansen, *Discrete Inverse Problems: Insight and Algorithms* (SIAM, 2010), Vol. **7**.

47. D. Débarre, A. Facomprez, and E. Beaurepaire, "Assessing correction accuracy in image-based adaptive optics," Proc. SPIE **8253**, 82530F (2012).

48. D. Torrieri, "Statistical theory of passive location systems," IEEE Trans. Aeros. Electron. Syst. **AES-20**, 183–198 (1984).

49. A. Beck, P. Stoica, and J. Li, "Exact and approximate solutions of source localization problems," IEEE Trans. Signal Process. **56**, 1770–1778 (2008).

50. J. J. Moré, "Generalizations of the trust region problem," Optim. Methods Softw. **2**, 189–209 (1993).

51. J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder–Mead simplex method in low dimensions," SIAM J. Optim. **9**, 112–147 (1998).

52. M. Andersen, J. Dahl, and L. Vandenberghe, "CVXOPT: A Python package for convex optimization, version 1.1.6," available at http://cvxopt.org (2013).

53. M. Andersen, J. Dahl, Z. Liu, and L. Vandenberghe, "Interior-point methods for large-scale cone programming," in *Optimization for Machine Learning*, S. Sra, S. Nowozin, and S. J. Wright, eds. (MIT, 2011), pp. 55–83.