Multiview Scattering Scanning Imaging Confocal Microscopy Through a Multimode Fiber

Sakshi Singh, Simon Labouesse, and Rafael Piestun

Abstract—Confocal and multiphoton microscopy are effective techniques to obtain high-contrast images of 2-D sections within bulk tissue. However, scattering limits their application to depths only up to ~1 millimeter. Multimode fibers make excellent ultrathin endoscopes that can penetrate deep inside the tissue with minimal damage. Here, we present Multiview Scattering Imaging Confocal Microscopy (MUSIC) Microscopy that enables high signal-to-noise ratio (SNR) imaging through a multimode fiber, hence combining the optical sectioning and resolution gain of confocal microscopy with the minimally invasive penetration capability of multimode fibers. The key advance presented here is the high SNR image reconstruction enabled by employing multiple coplanar virtual pinholes to capture multiple perspectives of the object, re-shifting them appropriately and combining them to obtain a high-contrast and high-resolution confocal image. We present the theory for the gain in contrast and resolution in MUSIC microscopy and validate the concept through experimental results.

Index Terms—Confocal microscopy, computational image formation, multiview image reconstruction, multimode fiber imaging.

I. INTRODUCTION

CONFOCAL microscopy [1], [2] is a widely used technique that enables optical sectioning for imaging with high contrast within scattering tissue. It employs a scanning focal spot to sequentially sample small segments of the object followed by filtering of the backscattered light using a small pinhole in the scanning-spot conjugate plane blocking the out-of-focus light. In practice, the pinhole diameter is chosen to be large enough to achieve a desired tradeoff between optical sectioning/resolution and signal integrity. The technique has been widely successful, enabling for instance, clinical studies for imaging of the cornea [3], [4], imaging in body cavities using fiber-optic catheters [5], [6] and skin cancer detection [7], [8]. However, up to date confocal imaging in the deep tissue regime remains infeasible due to the highly scattering nature of tissue and associated insufficient signal-to-noise ratio (SNR) levels.

Multimode fibers make excellent ultrathin endoscopes that can penetrate deep inside the tissue with minimal damage. Multimode fibers (MMFs) make the least invasive and light efficient endoscopes that relay the highest information content in a given cross section. Demonstrations of confocal imaging through multimode fibers have been made by digitally backpropagating from the detector to the object plane and filtering the signal through a virtual pinhole [32], [33], [34] or by means of optical correlation [35]. These demonstrations showed imaging of 2-D samples through MMFs with optical sectioning and improved contrast. However, their application in imaging in thick tissue remains challenging due to SNR limitations.

Confocal microscopy theoretically also has the capability to provide a factor of two in the lateral resolution [35], [40] with respect to the diffraction limited resolution based on the Rayleigh criteria [37], [38]. However, achieving this gain in resolution is impractical as it requires using a collection pinhole smaller than the size of the scanning focal spot, which compromises the signal strength. Improvement in imaging resolution through multimode fibers has been demonstrated using two-photon imaging [39], [40], saturated excitation [41], and by employing a multiple scatterer before the fiber [42], [43]. These approaches however come at the cost of expensive short pulse excitation sources, infeasibly high peak power, loss in transmitted light, or need for using short fibers. Another approach used a parabolic tip design [44] to increase the effective NA however the design reduces the field of view and requires a larger working distance, which makes the endoscope susceptible to tissue induced light distortions due to index mismatch. Recently, resolution beyond the diffraction limit [45] has also been demonstrated using MMFs by assuming sparsity in samples [31] however, it requires SNR levels.
of the sample higher than those feasible with bio-compatible markers.

Here, we present Multiview Scattering Scanning Imaging Confocal (MUSSIC) microscopy through MMFs, an approach to overcome the SNR limitation in confocal microscopy through complex media by employing multiple coplanar virtual pinholes to collect multiple perspectives of the object, followed by proper processing and combining them to retrieve a high SNR confocal image. Our method builds on the principle of image scanning microscopy (ISM) [36], [46], [47], [48], transmission matrices [28], [49], and digital phase conjugation [24]. ISM is used to boost the SNR in traditional confocal microscopy where the system is shift invariant using an algorithm called pixel reassignment. The improvement in SNR can be followed by or associated with a deblurring algorithm. However, in contrast with ISM, MUSSIC microscopy does not require a direct measurement of the images of the scanning focal spots. The so-called pixel reassignment operation typical of ISM is performed in a virtual phase-conjugate plane. Moreover, we demonstrate that given the transmission matrix of the system, MUSSIC microscopy can be employed for a more general, shift-variant system such as a generalized complex medium.

We first present a generalized framework to demonstrate the principle of MUSSIC microscopy through complex media and the theory for SNR and resolution gain. Further, we verify the theory experimentally by performing MUSSIC microscopy through an MMF by measuring its transmission matrix (TM). Using the TM, we generate focal spots on the far (distal) end of the MMF. As the focal spots scan the object, we collect the reflected speckle patterns on the MMF’s near (proximal) end. Using the MMF’s TM, we then back-propagate the collected speckle patterns to the distal object plane [32], [33], [34], [50] by numerically inverting the TM of the MMF. With virtual access to the distal fields reflected from the object, MUSSIC microscopy is implemented using a generalized pixel reassignment method [46], [47].

MUSSIC is ultimately an image reconstruction technique for optical data collected from multiple locations and falls into the broader class of inverse problems. Inverse problems are foundational for applications ranging from medical imaging to remote sensing. In the context of imaging, methods often used include regularization (e.g., Tikhonov, total variation), singular value decompositions, compressive sensing approaches based on sparsity, and learning algorithms. These are used in image restoration tasks like deconvolution, deblurring, inpainting, superresolution and tomographic applications such as magnetic resonance imaging and X-ray computed tomography. In this work, we employ Tikhonov regularization for matrix inversion, which is one of the well-known techniques for stabilization of inverse solutions [51].

Our experimental approach is quite general and applicable to shift invariant and shift variant imaging systems with separate excitation and collection paths [17], [47], [52], [53], [54]. Moreover, it can also be employed using digital phase conjugation [32]. We evaluate the SNR, optical sectioning and resolution of the reconstructed images and compare our approach with the conventional confocal and single pixel imaging [28], [49] approaches.

II. PRINCIPLE OF MUSSIC MICROSCOPY

Imaging through an MMF is performed by calibrating the relationship between the input and output fields through the MMF, described by its TM. The TM can be measured experimentally with both phase and amplitude information by sending an orthogonal set of input fields into the system accompanied with a phase-stepping reference field [28], [49]. A spatial light modulator (SLM) is typically employed to generate different input fields.

Once the TM is captured, it can be used to create focal spots on the distal end of the fiber to scan the object. The light reflected from the object goes back into the fiber and transforms into a speckle pattern, which is detected on the fiber’s proximal side. With knowledge of the TM, these detected proximal speckle fields can be computationally backpropagated to the distal end of the MMF by performing an inversion of the TM. Every pixel in the computed distal field in the neighborhood of the focal spot serves as a virtual pinhole carrying object information. Each of them can therefore be used for obtaining confocal images with space-shifted perspectives. Fig. 1 illustrates this process. The confocal images can subsequently be post-processed and combined to yield a final MUSSIC image reconstruction, as we will describe in Section IV. Before going into the MUSSIC reconstruction however, we describe a generalized matrix formalism to lay the foundation of our technique.
III. GENERALIZED MATRIX FORMALISM FOR MUSSIC MICROSCOPY THROUGH SCATTERING MEDIA

Consider a general imaging system whose input to output transformation is described by a TM, \( T \). Let \( O \) be the backscattering matrix of the object being imaged and let the fields entering and leaving the system be stored in a vectorized form in the columns of 2D matrices. Let these matrices be denoted by the letter \( E \) followed by their corresponding superscripts – “in” for incident fields entering the system, “d” for distal fields illuminating the object in the distal plane or far end of the system and, “p” for proximal fields reflected back from the object to the near or proximal end of the system. The proximal fields are obtained using (1).

\[
E^p = T^b O T E^{in} \tag{1}
\]

The \( T^b \) matrix above represents the transformation from the distal plane to the proximal plane. If the proximal plane coincides with the input plane from which the TM is measured, \( T^b \) is equal to the transpose of the forward transmission matrix. For generalization, we assume that the forward TM, \( T \) is measured between the SLM plane and the distal plane of the imaging system, while the backward TM, \( T^b \), is measured from the distal plane to a proximal plane where a camera detects the proximal speckle fields.

For a shift invariant system, both \( T \) and \( T^b \) matrices are convolution matrices whose columns contain the vectorized and circularly shifted point spread function (PSF) \([55], [56]\) of the system as illustrated in the top row of Fig. 2. For a raster scan approach the incident field matrix, \( E^{in} \) is an identity matrix whose \( k^{th} \) column contains the vectorized focal spot field centered on the \( k^{th} \) input pixel. The \( k^{th} \) column of the \( E^{in} \) matrix and \( k^{th} \) row of the \( E^p \) matrix are expanded to 2D and shown in the top and bottom rows of Fig. 2 using arrows, to illustrate the incident and proximal fields for the \( k^{th} \) scan position. The diagonals and near-diagonal elements of the \( E^p \) matrix comprise the raw data for MUSSIC reconstruction.

On the other hand, for a shift variant system, such as a complex media, the forward transmission matrix, \( T \) is a complex-valued random matrix as shown in the second row of Fig. 2. In this case, we can shape the incident field to create a focal spot at any given pixel in the distal plane using the TM approach for imaging by raster scanning \([24], [25], [28]\). This involves projecting an input field, \( E^{in}_{*k} = T_{*k}^t \), on the SLM to create a diffraction limited focal spot on the \( k^{th} \) pixel on the distal end of the MMF. Here, the subscripts denote the row and column indices of the matrices respectively and the asterisk denotes the full set of indices along the corresponding dimension. The dagger denotes the conjugate transpose operation.

However, since the distal field reflected from the object traverses back through the same medium, which has a complex random TM, the detected proximal fields are random speckle patterns that have no spatial correspondence with the distal fields.

Fig. 2. Generalized matrix formalism for imaging through complex media. Top row: Illustration of matrix structure and fields in a shift invariant system. Forward and backward transmission matrices and a single point example of the input and proximal field vectors reshaped to 2D are displayed using arrows. The transmission matrices in a shift invariant system are convolution matrices (e.g., with an Airy disk kernel). Middle row: Illustration of matrix structure and fields in a shift variant system such as a complex medium. Forward and backward transmission matrices and a single point example of the input and proximal field vectors reshaped to 2D are displayed using arrows. The input field is chosen such that it transforms into a focal spot at a chosen pixel position after propagating through the system. The transmission matrix of the system is a complex random matrix leading to a random speckle field in the proximal plane. Bottom row: Illustration of the virtual backpropagation method. By multiplying the left- and right-hand sides of (1) with the inverse estimate of the \( T^b \) matrix, we can backpropagate to the distal plane using arrows, to illustrate all field vectors denoted using the capital letter ‘E’ have been reshaped to 2D for visualization, whereas the matrices are shown in their original format.
To reverse the effect of the collection path and gain access to the distal fields reflected from the object, we can virtually backpropagate the detected proximal speckle fields to the distal plane [32], [33], [34] using the backward TM. This is achieved by multiplying the left- and right-hand side of (1) with the inverse-estimate of the $T^b$ matrix, which we denote as $T^{b\dagger}$. The resultant virtual distal fields can then be used to calculate the on-axis and off-axis confocal images which are subsequently combined for the MUSSIC image reconstruction. Here we define the on-axis confocal image as the image obtained using the virtual pinhole at the center of the focal spot in the virtual distal field. All surrounding virtual pinholes provide off-axis confocal images, which need to be re-shifted and aligned before they are combined.

If we approximate the inverse estimate of the $T^b$ matrix as its conjugate transpose, $T^{b\dagger} = (T^b)^\dagger$, as we did with the forward TM for creating phase conjugated focal spots on the distal end, and define the field illuminating the object, as $E_{il} = T E_{in}$, then the virtual distal fields, which we denote as $E^d$, can be calculated as follows,

$$E^d = (T^b)^\dagger E^p = (T^b)^\dagger O E_{il}$$  \hspace{1cm} (2)

We define $D_{p^b} = (T^b)^\dagger T^b$ as the virtual collection PSF of the system. Similarly, we define the pre-SLM to distal plane excitation PSF, $D_T = T T^\dagger$, which includes the wavefront projected on the SLM, $T^\dagger$, for generating focal spots. Assuming that a plane wave is incident on the SLM, $D_T$ is also the illumination field matrix, $E_{il}$. The matrices $D_T$ and $D_{p^b}$ have the same strong diagonals as in a convolution matrix for 2D fields used to represent the TMs of shift invariant systems [56], [57]. Their Hadamard product yields the net PSF of the system which is narrower than the individual PSFs as depicted in Fig. 1(c). This narrower net PSF is the source of resolution gain in MUSSIC microscopy, and the resolution enhancement is determined by the size of each virtual pinhole relative to the size of the virtual distal Airy disk. We define 1 Airy unit as the distance from the central peak to the first zero crossing in the virtual distal Airy disks.

IV. MUSSIC RECONSTRUCTION

If $N_{il}$ number of focal spot fields were created for imaging, a $N_{il} \times N_{il}$ matrix, $E^n$, containing the virtual distal fields is obtained. The on-axis confocal image is obtained from the main diagonal of $E^d$, $E_{kl}^d$, where $k \in (1, N_{il})$ is the focal spot position index. This main diagonal comprises the measurements from the central virtual pinhole, $p_2$, indicated in Fig. 1(b). Similar mutually shifted confocal images are also obtained from the diagonals, $E_{kl}^d$, corresponding to the neighboring pixels of $k$ (such as $p_1$ and $p_3$ shown in 1(b)) where $l$ takes $N^2 - 1$ values other than $k$ in the $N \times N$-pixel neighborhood of each scan position $k$. The three confocal images obtained from pinholes $p_1$, $p_2$ and $p_3$ are illustrated in Fig. 3.

To obtain the final MUSSIC image, the confocal images obtained from the chosen pinholes can then be re-shifted to a common perspective with knowledge of their distances from the center of the focal spot or pinhole $p_2$, using the pixel reassignment algorithm [58]. For example, the image obtained from the pinhole $p_1$ at the coordinate $(-d,-d)$ w.r.t. the pinhole $p_2$ at $(0,0)$, will be shifted be a positive distance $d/2$ along both the horizontal and vertical directions before it is combined with the confocal image from $p_2$. We also apply weights to individual images before combining the shifted images. At the boundary of the image the number of pinholes used for reconstruction is smaller since the signal from pinholes outside the boundary are not measured.

We found that a hybrid coherent weighted superposition provides quality reconstruction (see Appendix A). In particular, weighting the images obtained from the pinholes within 1 Airy unit by unity and the images obtained from pinholes outside one Airy unit in proportion to their mean value worked best. We present a comparison of reconstructions from different weighting approaches in the Appendix A. Finally, all the shifted and weighted images are coherently added to yield a high-SNR MUSSIC image.

We provide an algorithmic outline of the entire process below.

1) Use a set of orthogonal inputs to determine the input-output response between the SLM and the associated measured field. Use this to determine $N_{il}$ SLM excitation patterns, each of which produces an approximate single point of illumination at the object, which together form an approximate grid over the object.

Scanning process

2) Repeat the following over each of the generated excitation patterns:

a) Apply an excitation signal via the SLM to obtain an approximate single point of illumination at the object. This illumination appears at the object plane as an excitation PSF denoted by $T E_{il} = T T^\dagger S_{il} = (D_T)^\ast_{il}$.

b) Measure the resulting proximal image $E^p_{il} = T^b O T T^\dagger S_{il}$, where $O T T^\dagger S_{il}$ is the backscattering signal from the object illuminated at the $k^{th}$ focal spot and $T^b$
is the transformation from the distal image/illuminated object back to the proximal/measured image.

2 (a,b) represents the forward model. Therefore, the inverse problem consists in finding an estimate of \( G \) given the measurements \( E_{\tilde{s},k} \). Note that the excitation PSF, \( (D_T)_{s,k} \), as well as the actual measurement PSF, \( T^b(D_T)_{s,k} \), to the proximal plane, varies from point to point. Hence the system is shift variant and cannot be described by a convolution. Therefore, this is a space-variant deblurring problem with an extended complex valued PSF, \( T^b(D_T)_{s,k} \).

Reconstruction process

3) Repeat the following for each of the measured proximal field, \( E_{s,k}^p \):

a) Estimate the distal image corresponding to each scanning location, \( E_{\tilde{s},k}^d = (T^b)^\dagger E_{s,k}^p \approx OT_{s,k}T^b \).

Defining \( D_{T^b} = (T^b)^\dagger T^b \), we have \( E_{\tilde{s},k}^d = D_{T^b}O(D_T)_{s,k} \). This approximates an image obtained by observing directly the object illuminated at the \( k^{th} \) focal spot, where \( (D_T)_{s,k} \) describes the virtual excitation PSF and \( D_{T^b} \) describes the virtual collection PSF.

b) Restrict this estimated distal image to a \( N \times N \) neighborhood of the illuminated point, where \( N \) takes odd number values greater than 1. When the illumination is at the boundary of the imaging field of view, restrict the neighborhood to the subset of points lying within the boundary.

c) Each pixel, \( l \) in the neighborhood of the illumination point, \( k \), including \( l = k \), yields the \( k^{th} \) pixel of a virtual confocal image, \( C_k^l = E_{\tilde{l},k}^l \), where \( C_k^l \) is the virtual confocal image obtained from the \( l^{th} \) neighborhood pixel.

4) Re-shift and take a weighted average of the above confocal images to estimate the final MUSSIC image using (4).

\[
M = \sum l w_l \ast \tilde{C}_k^l
\]  

(3)

\( \tilde{C}_k^l \) is obtained by shifting the confocal image \( C_k^l \) by a distance \( \frac{r_x-r_y}{2} \), where \( r_x \) and \( r_y \) are the 2-D coordinates of the \( i^{th} \) pixel and \( w_l \) are the weights applied to each of the confocal images. Each \( l \) value corresponds to one of the pixels in the chosen neighborhood of the illumination point.

The resolution is determined from the combination of the excitation PSF with the collection PSF. The resulting effective PSF is essentially the Hadamard product of these two PSFs, which has a narrower full-width half-max than any of the two PSF alone.

Note that the reconstruction can be performed in real time (point-by-point) as the scan is performed, or all at once after the scanning process is finished.

V. METHODS

In what follows we present a specific experimental implementation of MUSSIC through an MMF.

A. Calibration of Forward TM, \( T \)

The forward TM \( T \) is measured with both phase and amplitude information, by sending orthogonal input fields into the fiber accompanied by a phase-stepping reference field. We choose the plane-wave basis that transforms to focal spots in the Fourier plane, which are then coupled into the MMF. These patterns are constant in amplitude and their phases are modulated using a spatial light modulator (SLM). The SLM’s active area is divided into two sections each for a changing grating pattern and a phase-stepping reference frame that surrounds it. The intensity measurements at the fiber output for each projected pattern, as the reference field is phase stepped, allow for the recovery of the output fields \([28, 49]\). These output fields are vectorized and used to build all the rows of the matrix \( T \).

B. Calibration of Backward TM, \( T^b \)

The TM of an MMF obeys the reciprocity rule \([59]\). However, in practice, we measure the TM between the SLM and the distal plane and the reciprocity assumption only holds true if the collection plane perfectly matches the SLM plane in scale and orientation. This is challenging in practice, requiring a sensitive and time-consuming alignment procedure \([32]\). Moreover, oftentimes, it is desirable to separate the excitation and collection pathways in endoscopes to improve throughput or to gain some feedback from the distal end \([17, 52, 53, 54]\). In such cases \( T^b \neq T^r \). For other modalities like fluorescence imaging, the excitation and collection PSFs are different by default due to difference in the excitation and fluorescence wavelengths. With these considerations, here we propose a separate calibration of the matrix \( T^b \) from the distal plane to the detector plane.

Towards this end, we place a mirror at the distal end of the fiber and scan focal spots on it, while measuring the reflected fields on the proximal end, denoted as \( E_{p-mirror} \). These fields are measured using phase shifting interferometry, as done earlier for the forward TM calibration \([28, 49]\). The SLM phase masks used for creating focal spots in the distal field are surrounded by a phase-stepping reference frame and the three phase-speckle intensity measurements allow recovery of the proximal speckle fields.

The proximal speckle fields give us an estimate of \( T^b \), which we denote as \( T_{obs}^b \), as follows

\[
T_{obs}^b = E_{p-mirror}^p = T^b I E_{il}^d
\]  

(4)

The matrix \( I \) in the above equation represents the mirror reflection matrix, which we assume to be an identity matrix for a perfectly reflective mirror. The distal fields are then given by

\[
E_{obs}^d = (T_{obs}^b)^\dagger E^p = (E_{il}^d)^\dagger D_{Ts} O E_{il}^d
\]  

(5)

In comparison to (3), the above equation has an additional term \( (E_{il}^d)^\dagger \) on the right-hand side because of our double pass approach for calibration of \( T^b \). Since we use a raster scan approach, both the \( E_{il}^d \) and \( (E_{il}^d)^\dagger \) matrices have the same prominent diagonals as a convolution matrix with a diffraction limited Gaussian kernel and the distal fields obtained from (5) are a good approximation of the distal fields calculated in (3),
hence enabling confocal and MUSSIC image reconstruction. Moreover, the theoretical resolution gain of confocal imaging is also preserved as the bandwidth of the terms on the left and right of the object, \( O \), in the above equation remain unchanged. Appendix B provides a more detailed derivation of (6).

C. Optimal Inversion of Backward TM

As mentioned earlier, we can use the conjugate transpose operator when the inverse of a matrix does not exist. This method works well for generating phase conjugated focal spots, as required when raster scanning on the distal side of the fiber. However, when calculating the backpropagated distal fields, the conjugate transpose is not the best inversion method. We can optimize the inversion of the backward TM using a Tikhonov regularization technique [32], [60]. This involves computing the singular value decomposition of the backward TM, \( T_{\text{obs}}^b = U S V^\dagger \) and finding its inverse as follows

\[
S^{\text{RI}}_{\text{obs}} = V S^{\text{RI}} U^\dagger \tag{6}
\]

\( S^{\text{RI}}_{\text{obs}} \) is the regularized inverse of the diagonal matrix of singular values, \( S \), calculated by replacing the singular values \( \sigma_i \) in the diagonal of \( S \) with \( \frac{\sigma_i}{\sigma_i^2 + \beta^2} \), where \( \beta \) is the regularization parameter. In experiments, we find that by calculating the backpropagated distal fields using Tikhonov regularized inverse of the \( T_{\text{obs}}^b \), instead of \( (T_{\text{obs}}^b)^\dagger \) in (6), yields image reconstructions with improved SNR and contrast. For our results, we chose a \( \beta \) value equal to 10% of the highest singular value of the backward TM. A comparison of reconstructions using the conjugate transpose and Tikhonov regularization is presented in Appendix C and reveals that although the regularization considerably improves the image quality, the faster reconstruction obtained from the conjugate transpose of the TM also provides a good estimate of the object.

D. Lowpass Filtering and Normalization

We perform digital lowpass filtering to bandlimit the spatial frequencies in the acquired data. This eliminates the noise in the high frequencies and ensures that all acquired images have speckles with a minimum grain size limited by diffraction. The frequency cutoff is chosen by computing the average of the Fourier transform of all the measured proximal images and setting the values below a minimum threshold outside the central Gaussian peak to zero.

The MUSSIC reconstruction of a blank object or a perfectly reflective mirror is a speckle pattern. This is because the measured TM does not fully characterize light propagation through the MMF producing a residual background speckle, which appears in the MUSSIC reconstruction. This is also evident in the fact that the excitation and virtual collection PSFs have non-zero values outside of the prominent diagonals present in a convolution matrix. In order to account for this non-uniformity due to the background speckle pattern, we normalize the reconstructed confocal and MUSSIC reconstructions w.r.t to their “blank” counterparts i.e., the reconstruction images obtained when a mirror is placed at the distal end. This normalization significantly improves the image quality. Appendix D shows a comparison of the reconstruction images before and after normalization.

E. Imaging Without Full Field Backpropagation

Calculating the full matrix \( E^d \), involves heavy computation, with a complexity \( O(N^3_{\text{d}} N_{\text{in}}) \), where \( N_{\text{d}} \) is the number of illumination fields scanning the object and \( N_{\text{in}} \) is the number of pixels in the detected proximal speckle fields. However, in fact, access to the full backpropagated distal fields is not necessary to calculate the confocal or MUSSIC images. The only data points required in each distal field are in the neighborhood of the scanning focal spot, for every scan position. This number, which we define as \( N_{\text{pinholes}} \) is chosen to be roughly equal to the number of pixels that sample a focal spot and is much lower than the number of illumination patterns used for imaging. Hence, if we compute only the desired diagonals from the matrix \( E^d \) corresponding to the \( N_{\text{pinholes}} \) neighboring pixels, the complexity of the calculation drops down to only \( O(N_{\text{pinholes}}^3 N_{\text{in}} N_{\text{d}}) \) for the MUSSIC image and to only \( O(N_{\text{in}}^3 N_{\text{d}}) \) for a single confocal image. When using the conjugate transpose of the backward TM to invert it, this method for obtaining the confocal image is similar to the correlation method [32]. Imaging with this reduced computation approach, which we call fast-MUSSIC, enables MUSSIC reconstruction of a 20000-pixel image in 4 minutes on a DELL Desktop computer with a 3.2 GHz Intel Core i5 processor and 64 GB RAM.

VI. EXPERIMENTAL SETUP

The experimental setup used to demonstrate MUSSIC microscopy through an MMF is illustrated in Fig. 4. We use a 785 nm CW Crystal laser and a Meadowlark optics liquid crystal SLM (HSPDM 512) for phase modulation. The phase modulation depth of the SLM is maximized when the incident light is polarized along the vertical axis. We therefore use a half waveplate to rotate the laser polarization to the vertical axis and a linear polarizer that allows only the vertically polarized component of light to reach the SLM. After the SLM, the phase
modulated light is relayed and demagnified using a 4-F system to match the beam diameter to the active area of the SLM. Further, the beam goes through another half-waveplate and a polarizing beam splitter to maximize power along the horizontal polarization component. A quarter-wave plate after the PBS makes the input beam circularly polarized and is used to achieve an optical isolator effect to prevent back-reflections from the near facet of the MMF from reaching the proximal camera, CAM2, when imaging the proximal speckle fields from the MMF.

The SLM plane is then imaged onto the back-aperture of a microscope objective, OBJ 1, which couples the light into the MMF. We used a step-index fiber of diameter 50 μm and 0.22 numerical aperture (NA) for all our experiments. The distal facet of the MMF is imaged onto a camera, CAM 1 using another lens during the forward TM calibration. A linear stage before the camera allows detection of only one polarization component. The sample is mounted on a linear stage (Thorlabs XR50c) which controls its distance from the MMF distal tip.

After the forward TM calibration, a mirror is placed behind the fiber distal tip for calibration of the backward TM. In our experiments, the mirror and the sample were present on the same planar surface (different sub-regions of the USAF target) and hence moving from the mirror to the sample was done by a lateral translation of the target perpendicular to the fiber’s axis.

The backward TM is calibrated using back-reflected fields on the proximal side of the fiber, while focal spots are projected on the distal side. A phase shifting reference frame is simultaneously projected on the SLM along with the phase conjugated patterns for the forward scan, for measuring both the phase and amplitude of the back-reflected fields. The back-reflected light from the mirror couples back into the fiber and is detected on the proximal side after reflecting off the PBS using another camera, CAM 2. This camera images the back aperture of the microscope objective OBJ 1 using another 4-F system and is placed in a plane equivalent to the SLM plane. A linear stage before the camera allows detection of a single polarization component.

After both calibrations, the sample to be imaged replaces the mirror at the distal facet of the MMF, and the back-reflected fields from the object are recorded as it is raster scanned. Different depths can be scanned by moving the sample, moving the fiber, or numerically refocusing the transmission matrix via free space propagation (see Appendix E).

VII. RESULTS

A. SNR and Resolution Analysis

We perform confocal and MUSSIC microscopy in simulation and compare the SNR and resolution of the reconstructed images. We model the MMF TM as a complex random matrix and reconstruct the image of a quadrant of the binary siemens star using the simulated proximal speckle fields, following the backpropagation process described earlier. We added Gaussian noise with 5% variance to the simulated proximal fields before the image reconstruction. Each virtual pinhole in our simulation has a radius of 0.11 Airy unit (a.u.). Hence one Airy disk spans across 9 × 9 individual pinholes.

The ground truth object and its confocal and MUSSIC reconstructions are shown in Figs. 5(a)–(d). Figs. 5(b), (c) show the confocal reconstructions using a 3 × 3 macro-pinhole and a 9 × 9 macro-pinhole respectively. Fig. 5(d) uses the same group of 9 × 9 pinholes as 5(c) but employs the MUSSIC approach. We find that although the SNR improves significantly among the confocal image reconstructions as the size of the macro-pinhole increases, the resolution degrades. On the other hand, the MUSSIC reconstruction, which uses the same group of pinholes as the second confocal reconstruction retains the high-SNR, while also preserving the resolution. The difference in resolution can be more clearly visualized in Fig. 5(e) that shows the normalized cross sections in the image reconstructions corresponding to the green solid lines in Fig. 5(a)–(d). We find that confocal reconstruction with the 1 a.u. pinhole fails to resolve the image features, while the MUSSIC reconstruction using the same raw data resolves them just as well as the confocal reconstruction with the small 0.33 a.u. pinhole.

Next, we analyze in Fig. 5(f) the reconstruction error and correlation as a function of the number of pinholes used, for the green line cross-sections marked in Fig. 5(a)–(d) in the absence of noise. We find that the cross-section error and correlation w.r.t. the ground truth increases and decreases respectively as the number of pinholes constituting the macro-pinhole increases for the confocal reconstruction. On the other hand, both metrics for the MUSSIC reconstructions remain unaffected, indicating that the image quality is preserved.

Furthermore, in Fig. 5(g), we compare the average frequency response for the confocal and MUSSIC methods obtained for a point object using 81 pinholes. These responses were also obtained in the absence of noise. We find that the frequency cutoff of the MUSSIC reconstruction is almost double that of the OTF of the system, which is the theoretically claimed gain in resolution according to Rayleigh’s criterium [37], [38]. On the other hand, the confocal reconstruction obtained from the 1 a.u. pinhole has a frequency cutoff 1.4 times higher than that of the system OTF.

Next, we computed the noiseless PSFs for the confocal and MUSSIC methods and plotted the full width half maxima (FWHM) of the PSFs as a function of the number of pinholes used. We find that the FWHM for the confocal reconstruction increases with the number of pinholes constituting a macro-pinhole, while the FWHM for the MUSSIC reconstruction remains unchanged [Fig. 5(h)].

Finally, in Fig. 5(i), we compare the root mean square error of different normalized reconstructions as a function of the annular radius measured from the center of the Siemens star on the bottom right corner of Fig. 5(a). For this comparison, no noise was added to the reconstructions to analyze the effect of increasing number of pinholes on resolution. We divide the image quadrant into 15 radial zones and plot the error w.r.t. the ground truth image in each zone for the different reconstruction methods. We find that while the error for the confocal reconstruction images increases with the radius of the macro-pinhole, the error in the MUSSIC reconstruction remains almost unchanged as the number of used pinholes increases from 3 × 3 to 13 × 13. The inset images in the figure show the radial
zones 1, 5 and 15 from left to right. Fig. 5(j) shows a schematic of the different pinhole groups used for confocal and MUSSIC reconstructions along with their respective sizes in Airy unit and number of pinholes.

**B. Experimental Results**

We demonstrate MUSSIC microscopy through a multimode fiber and compare the reconstructions in Fig. 6. The field of view (FOV) consists of the fourth and fifth elements of the 7th group in the USAF 1951 resolution target, which have a resolution of 181- and 203-line pairs/mm respectively. Fig. 6(a)–(c) show the MUSSIC reconstructions as the number of used pinholes increases from 1 to 81. We see that the reconstruction SNR improves consistently with the number of pinholes.

We compare our results single pixel imaging (SPI) in which the image is reconstructed pixel-wise by integrating the absolute values of all the pixels in the proximal speckle fields for every focal spot scanning the object. This technique is often used in fluorescence imaging and could be used in reflection [61]. Fig. 6(d) shows this result. Fig. 6(e) shows the normalized average cross sections for all the reconstructions along the horizontal direction for a cropped window within the FOV (grey-dashed lines). The SPI image, obtained without virtual backpropagation and using the signal from all the distal pinholes, has a higher SNR than other reconstruction but lower contrast. For the MUSSIC reconstructions the contrast improves with the number of pinholes.

Next, we demonstrate the optical sectioning capability of the MUSSIC technique. The FOV shows the first element of the 7th group in the resolution target. We move the 2D target in steps of 20 μm in the axial direction and away from the fiber distal facet and capture the back-reflected fields from the object at three z-positions. We compare the SPI reconstructions with the MUSSIC reconstructions at the three positions in Fig. 7. We observe that the object almost disappears in the background
already after a z-shift of 20 μm in the case of the MUSSIC images, while the SPI reconstructions carry a significant amount of light from the sample even after a z-displacement of 40 μm. Hence the MUSSIC approach performs better in rejecting the light out of the image plane.

VIII. DISCUSSION AND CONCLUSION

We have demonstrated MUSSIC microscopy through a multimode fiber to enable imaging with improved optical sectioning, high SNR, and improved resolution. By increasing the number of virtual pinholes used for reconstruction, the image contrast improves with respect to the virtual confocal image. The tradeoff is in the computational complexity which grows linearly with the number of pinholes.

In the experiments, as usual, the quality of the reconstruction and in particular the achieved resolution is affected by overall SNR. While we demonstrate a factor of 2 improvement in the spatial frequency bandwidth of MUSSIC images relative to the optical transfer function of the system in simulation [Fig. 5(g)], a similar improvement is difficult to observe or characterize in the noisy experimental images. This can be attributed to several uncontrolled experimental factors. Firstly, the image quality is dependent on the accuracy of the reconstructed virtual distal fields, which is in turn determined by the quality of the inverse estimate of the backward TM. We used the Tikhonov regularized inversion although alternate techniques could be considered to improve system estimates, including deep learning approaches [60], [61], [62].

Secondly, our technique assumes that the object is illuminated with focal spots with no background speckle. In practice, only about 10% of the total power is in the focal spot, while the rest is distributed in the background speckle and contributes to noise in the reconstruction.

Moreover, in the implemented experimental calibration of the backward TM, we assumed a perfect reflective mirror whose reflection matrix is an identity matrix. In practice, some light is lost at the mirror and does not couple back into the fiber. Moreover, the object must be positioned precisely in the plane of the mirror used during the calibration of the backward TM. Any deviation would cause the signal from the object plane to become out-of-focus which would consequently be rejected by the virtual pinholes. Interestingly, these limitations due to the use of an external mirror for the calibration of the backward TM can be overcome using an alternate calibration method utilizing the reflections from the MMF distal tip. Appendix E presents this approach.

Furthermore, to keep our experimental setup simple and robust to thermal and mechanical fluctuations, we used an internal reference for phase measurements which transforms to a nonuniform speckle in the plane of interest with many intensity nulls, also known as blind spots. The field from these blind spots cannot be recovered, which further degrades the image reconstruction quality. Using complementary reference speckles [65], [66] or an external plane wave reference are possible ways to eliminate the blind spots, although they either require increased measurement time or a more complex setup with phase tracking to account for phase drifts. Bending sensitivity of the fiber is another challenge and any perturbations after calibration lead to systematic errors in the image reconstruction. However, various approaches exist to mitigate fiber perturbations, including feedback mechanisms, fiber selection [70] or illumination pattern optimization [73], [74].

Here we limit our experiments to coherent imaging, but the high SNR capability of MUSSIC microscopy could also be useful for fluorescence imaging. Unlike the presented coherent imaging approach, the monochromatic TM is insufficient to perform virtual backpropagation of the proximal fluorescence patterns due to the partial spatial and temporal coherence of fluorescence. Calibration of the multispectral TM of scattering media has been demonstrated to enable spatio-temporal control of waves through scattering media [75], [76], [77]. With the help of the multispectral TM, one could for instance, scan multi-spectral focal spots on the proximal side while speckle patterns are projected on the object at the distal end. With knowledge of the distal intensity patterns, the object could be
recovered [31], [33], [34]. An advantage of scanning focal spots on the proximal side is that it would eliminate the need for coherent backpropagation and enable imaging by solving an intensity-only inverse problem.

A further generalization of the technique can be made by choosing distal illuminations that are not focal spots, but arbitrary speckle patterns. The mathematical model for this case is described in Appendix F. Speckle illumination is ideal for compressive sampling and can enable imaging with fewer illumination patterns and shorter data acquisition times [31], [78]. Furthermore, it can also eliminate the need for wavefront shaping [34].

In summary, this report demonstrates the capability of MUS-SIC microscopy in enabling high SNR and high-resolution imaging through a MMF endoscope. Given the generalized principle of the technique, its application is not limited to the raster scan approach or to MMF and can easily be adapted to other endoscopic probes that might require different excitation and collection paths such as double-clad fibers.

APPENDIX

A. Different Weighting Methods in the Weighted Pixel Reassignment Algorithm

The standard pixel reassignment method [47] employs a common weighting factor of unity during the summation of the re-shifted images obtained from different pinholes. It has been suggested [58] that a weighting factor dependent on the pinhole location might enable a better noise performance. Here we present a comparison of the reconstruction error for different weighting approaches, namely, unity weighting, mean weighting and hybrid weighting. In the mean weighting approach, the weighting factor is chosen as the mean intensity of the 2D images obtained from different pinholes, whereas the hybrid approach employs unity weighting for pinholes within 1 a.u. and mean weighting for those outside 1 a.u. Optimal weighting approaches, e.g., based on SNR or information theoretic metrics, can be considered beyond these heuristic weighting approaches, and are left for discussion elsewhere. Fig. 8 shows the reconstructions obtained using 81 pinholes (within 1 a.u.) implemented with unity (a) and mean (b) weighting as well as reconstructions using 169 pinholes (extending outside 1 a.u.) using unity (d), mean (e) and hybrid (f) weighting.

We find that when we only employ pinholes within 1 a.u. the image contrast is better using unity weighting. However, as more pinholes outside the 1 a.u. radius are employed, the unity weighted image reconstruction degrades in quality as noise outside the 1 a.u. region dominates over the signal. On the other hand, the image quality is preserved, and contrast slightly enhanced for the mean weighted reconstruction. The hybrid weighting approach leads to a further improved contrast than the mean weighted reconstruction. However, the reconstruction error for the hybrid approach is slightly higher than the mean weighting approach as can be observed from the plot in Fig. 9(c) showing the root mean square error (RMSE) of reconstruction with increasing number of pinholes. The RMSE is calculated for the reconstruction images, $\hat{R}$, optimally rescaled by a factor $\alpha$, with respect to the binary object, $O$:

$$\text{RMSE} = \arg\min_{\alpha} \sqrt{\sum (O - \alpha R)^2 / \|O\|}$$  \hspace{1cm} (7)

Hence, the hybrid approach yields a consistently good reconstruction with increasing number of pinholes, while the mean weighted reconstruction yields the least error when pinholes outside 1.a.u. are used. On the other hand, when only sub-airy unit pinholes are employed, the unity weighting method yields the best reconstruction even in the presence of significant noise.

B. MUSSIC Recovery From a System With Different Excitation and Collection PSFs

In the main text, we explained that when the collection and excitation paths are different, the backpropagated distal fields can be calculated using (6). We derive (6) in this section by expanding each term using the definition of calibrated backward TM, $T^b_\text{obs}$ and the proximal field matrix, $E^p$, as shown in (8).

$$E^d_{\text{obs}} = (T^b I E^d)^\dagger T^b O E^d$$  \hspace{1cm} (8)
Using the definition for the virtual collection PSF, \( D_T = (T^b)^\dagger T^b \), Eq. (8) simplifies to (5) as follows:

\[
E_{\text{obs}}^d = (E_{\text{il}}^d)^\dagger (T^b)^\dagger T^b O E_{\text{il}}^d = (E_{\text{il}}^d)^\dagger D_T O E_{\text{il}}^d
\]  

(9)

It should be noted that the above relationship also holds true when the inverse of the backward TM and the illumination matrix are estimated using the Tikhonov regularization approach instead of the conjugate transpose. To preserve the resolution gain in MUSSIC microscopy, the terms on the left and right side of the object matrix should retain the spatial frequency bandwidths of the excitation and collection systems. The right-hand term is the same as in the earlier case of identical excitation and collection PSFs. The left-hand term is the matrix product of the inverse of the illumination matrix and the inverse of the virtual collection PSF, which also has the bandwidth of the collection or the excitation system, whichever is smaller. When the two systems have the same bandwidth, the resolution gain and sectioning properties of MUSSIC microscopy are preserved.

C. MUSSIC Reconstructions Using Different Inversion Strategies

Fig. 9(a) and (b) show the MUSSIC reconstructions from 25 pinholes when the conjugate transpose of the backward TM and the Tikhonov regularized inverse of the TM are respectively used to backpropagate to the distal plane. The latter requires computation of the backward TM’s singular value decomposition and tuning of the regularization parameter, which requires additional computational effort.

On the other hand, the conjugate transpose is computed faster allowing, for instance, MUSSIC reconstruction (with normalization) within 2.6 minutes using 9 confocal images and within 5 minutes using 25 pinholes on a DELL Desktop computer with a 3.2 GHz Intel Core i5 processor and 64 GB RAM. Although the result from Tikhonov regularization is significantly better, the conjugate transpose also provides a good estimate of the object.

D. Effect of Normalization in MUSSIC Reconstruction

As described in the main text, the MUSSIC reconstruction of a blank object or a mirror is a speckle pattern because the virtual collection PSF, \( D_T \), is similar to but not exactly a convolution matrix. Therefore, we normalize all obtained MUSSIC reconstruction images with respect to a blank image captured when a mirror is placed in the object plane. Fig. 10 shows the experimental MUSSIC reconstruction of the USAF target using 25 pinholes (a) without and (c) with normalization. We can observe that much of the non-uniformity arising from the speckle reference, used for phase measurement, is eliminated because of the normalization, yielding a cleaner reconstruction. Mathematically, the normalized reconstruction in Fig. 10(b) is obtained by the elementwise division of the MUSSIC reconstruction in Fig. 10(a) by the MUSSIC reconstruction of a mirror object.

E. Calibration of the Backward TM Without Using a Mirror

In the main text, we described a procedure for calibrating the observed backward TM of the fiber by scanning focal spots on a mirror placed at the distal end and measuring the reflected speckle fields. The distal tip of the fiber however also acts as a partial reflector and produces speckle patterns with measurable intensity variation on the proximal end. Hence, it is possible to calibrate the observed backward TM without the additional mirror.

Fig. 11 shows the MUSSIC reconstruction obtained using the backward TM calibrated from reflections from (a) the mirror and (b) the MMF distal tip. We observe that the reconstruction in (b) is blurry. This is explained by the fact that the focal spots are scanned in the object plane, which lies at a small distance away from the distal tip. On propagating the MUSSIC reconstruction in (b) along the z-direction, the image quality approaches that in (a) although it still presents lower contrast. On the other hand, one improvement observed in the reconstructions using the distal tip reflections is the lack of background around the edges that is observed in the reconstruction using the external mirror. This background is the result of tilted placement of the mirror with respect to the fiber distal tip, perturbations, and/or multiple reflections. The calibration approach using reflections from the fiber distal tip could therefore be a more robust and easier to implement alternative to the approach using an additional mirror.

F. Generalized MUSSIC Microscopy Using Speckle Illumination

When the fields illuminating the object in the distal plane are not focal spots, but a set of random speckle patterns, the illumination matrices \( E_{\text{il}}^d \) no longer have a structure similar to a convolution matrix. However, MUSSIC microscopy can still be performed by inverting the illumination matrix [31], [34]. (10) shows how we can obtain the raster scan approach equivalent of
the back-propagated distal fields, \( E_{\text{foc}} \), in this case.

\[
E_{\text{foc}} = E^{\text{til}}_m E^{\text{d}}_\text{obs} \left( E^{\text{til}}_m \right)^\dagger = D_D D^\dagger_O D_D
\]  

(10)

where we define \( D_D = E^{\text{til}}_m \left( E^{\text{til}}_m \right)^\dagger \), as the excitation PSF. Again, in this case, the conjugate transpose operation in the above equation can be replaced by the Tikhnov regularized inverse. The added computational complexity of \( O \left( N_{\text{out}}^2 N_{\text{til}} \right) \) above can enable imaging with arbitrary speckle illuminations with reduced number of illuminations, when the chosen patterns have small correlations [31].

Even when focused illumination is employed, there is inevitably some energy in the background due to imperfect control over all the fiber modes due to insufficient overlap between the fiber modes and modes of the spatial light modulator (SLM), and thermal and mechanical perturbations. This background is a source of noise in the reconstruction. By measuring the fields corresponding to the distal illuminations, and using the above generalized approach, we can convert the energy in the background into signal and improve the quality of image reconstruction.

ACKNOWLEDGMENT

We thank Dr. Antonio M. Caravaca for several fruitful discussions that helped improve this manuscript.

REFERENCES

[1] R. H. Webb, “Confocal optical microscopy,” Rep. Prog. Phys., vol. 59, pp. 427–471, 1996.
[2] S. W. Paddock, “Principles and practices of laser scanning confocal microscopy,” Mol. Biotechnol., vol. 16, pp. 127–149, 2000.
[3] I. Jalbert, F. Stapleton, E. Papas, D. F. Sweeney, and M. Conrono, “In vivo confocal microscopy of the human cornea,” Brit. J. Ophthalmol., vol. 87 pp. 225–236, 2003.
[4] W. Wiegand, A. A. Thaer, P. Kroll, O.-C. Geyer, and A. J. Garcia, “Focal plane multiphoton microscopy,” in Biomed. Opt. Exp., vol. 4, pp. 13354–13364, 2009.
[5] R. Kiesslich et al., “Diagnosing Helicobacter pylori in vivo by confocal optical microscopy,” in Biomed. Opt. Exp., vol. 8, pp. 1157–1169, 2014.
[6] J.-H. Han, J. Lee, and J. U. Kang, “Pixelation effect removal from fiber bundle probe based optical coherence tomography imaging,” in Opt. Exp., vol. 18, pp. 7427–7439, 2010.
[7] R. M. Halatari, R. Y. Gu, and J. M. Kahn, “Resolution limits for imaging through multi-mode fiber,” in Opt. Exp., vol. 21, pp. 1656–1668, 2013.
[8] A. M. Caravaca-Aguirre and R. Piestun, “Single multimode fiber endoscope,” in Opt. Exp., vol. 25, pp. 1656–1665, 2017.
[9] M. Rajadhyaksha, M. Grossman, D. Esterowitz, R. H. Webb, and R. R. Anderson, “In vivo confocal scanning laser microscopy of human skin: Melanin provides strong contrast,” in J. Invest. Dermatol., vol. 104, pp. 946–952, 1995.
[10] A. M. Gonzalez Block and A. Mills, “Assessing capacity for health policy and systems research in low and middle income countries,” Health Res. Policy Syst., vol. 1, no. 1, pp. 1–20, 2003.
[11] D. Loterie, S. Farahi, I. Papadopoulos, A. Goy, D. Psaltis, and C. Moser, “Digital confocal microscopy through a multimode fiber,” in Proc. IEEE Photon. Conf., vol. 23, pp. 23845–23858, 2015.
[12] S.-Y. Lee, B. Bouma, and M. Villiger, “Confocal imaging through a multimode fiber without active wave-control,” in Proc. IEEE Photon. Conf., 2019, pp. 1–2.
[13] S.-Y. Lee, Y. J. Parot, B. E. Bouma, and M. Villiger, “Confocal 3D reflectance imaging through multimode fiber without wavefront shaping,” Optica, vol. 9, pp. 112–120, 2022.
[14] D. Loterie, S. A. Goorden, D. Psaltis, and C. Moser, “Confocal microscopy through a multimode fiber using optical correlation,” in Opt. Lett., vol. 40, pp. 5754–5757, 2015.
[15] C. J. R. Sheppard, “Super-resolution in confocal imaging,” Opt. Structit., vol. 80, pp. 53–54, 1998.
[16] M. Bertero, P. Biani, and E. R. Pike, “Super-resolution in confocal scanning microscopy,” in Inverse Problems, vol. 3, pp. 195–212, 1987.
[17] E. E. Morales-Delgado, D. Psaltis, and C. Moser, “Two-photon imaging through a multimode fiber,” in Opt. Exp., vol. 23, pp. 32158–32170, 2015.
G. P. J. Laporte, N. Stasio, C. Moser, and D. Psaltis, “Enhanced resolution in a multimode fiber imaging system,” Opt. Exp., vol. 23, pp. 27484–27493, 2015.

Y. Choi, C. Yoon, M. Kim, J. Yang, and W. Choi, “Disorder-mediated enhancement of fiber numerical aperture,” Opt. Lett., vol. 38, pp. 2253–2255, 2013.

I. N. Papadopoulos, S. Farahi, C. Moser, and D. Psaltis, “Increasing the imaging capabilities of multimode fibers by exploiting the properties of highly scattering media,” Opt. Lett., vol. 38, pp. 2776–2778, 2013.

S. Bianchi, V. P. Rajamaniakam, L. Ferrara, E. D. Fabrizio, C. Liberale, and R. D. Leonardo, “Focusing and imaging with increased numerical apertures through multimode fibers with micro-fabricated optics,” Opt. Lett., vol. 38, pp. 4935–4938, 2013.

L. V. Aminonova and J. F. de Boer, “Endo-microscopy beyond the Abbe and Nyquist limits,” Light: Sci. Appl., vol. 9, 2020, Art. no. 81.

C. B. Müller and J. Enderlein, “Image scanning microscopy,” Phys. Rev. Lett., vol. 104, 2010, Art. no. 198101.

C. J. R. Sheppard, S. B. Mehta, and R. Heintzmann, “Superresolution by image scanning microscopy using pixel reassignment,” Opt. Lett., vol. 38, pp. 2889–2892, 2013.

C. Roider, R. Heintzmann, R. Piestun, and A. Jesacher, “Deconvolution approach for 3D scanning microscopy with helical phase engineering,” Opt. Exp., vol. 24, pp. 15456–15467, 2016.

S. M. Popoff, G. Lerosey, R. Carminati, M. Fink, A. C. Boccara, and S. Gigan, “Measuring the transmission matrix in optics: An approach to the study and control of light propagation in disordered media,” Phys. Rev. Lett., vol. 104, 2010, Art. no. 106601.

S. Singh, S. Labouesse, and R. Piestun, “Multiview virtual confocal microscopy through a multimode fiber,” in Proc. IEEE 19th Int. Symp. Biomed. Imag., 2021, pp. 584–587.

R. A. Willoughby, “Solutions of ill-posed problems (A. N. Tikhonov and V. Y. Arsenin),” SIAM Rev., vol. 21, 1979, Art. no. 266.

D. Yelin, B. E. Bouma, S. H. Yun, and G. J. Tearney, “Double-clad fiber for endoscopy,” Opt. Lett., vol. 29, pp. 2408–2410, 2004.

S. Farahi, D. Ziegler, I. N. Papadopoulos, D. Psaltis, and C. Moser, “Dynamic bending compensation while focusing through a multimode fiber,” Opt. Exp., vol. 21, pp. 23524–23514, 2013.

M. T. Myaing, D. J. MacDonald, and X. Li, “Fiber-optic scanning twophoton fluorescence endoscope,” Opt. Lett., vol. 31, pp. 1076–1078, 2006.

B. Judkewitz, R. Horstmeyer, I. M. Vellekoop, I. N. Papadopoulos, and C. Yang, “Translation correlations in anisotropically scattering media,” Nature Phys., vol. 11, pp. 684–689, 2015.

O. Tzang, E. Niv, S. Singh, S. Labouesse, G. Myatt, and R. Piestun, “Wavefront shaping in complex media with a 350 kHz modulator via a 1D-to-2D transform,” Nature Photon., vol. 13, pp. 788–793, 2019.

P. J. Davis, Circulant Matrices. Hoboken, NJ, USA: Wiley, 1979.

C. J. R. Sheppard et al., “Pixel reassignment in image scanning microscopy: A re-evaluation,” J. Opt. Soc. Amer. A, vol. 37, pp. 154–162, 2020.

S.-Y. Lee, V. J. Parot, B. E. Bouma, and M. Villiger, “Reciprocity-induced symmetry in the round-trip transmission through complex systems,” APL Photon., vol. 5, 2020, Art. no. 106104.

S. M. Popoff, G. Lerosey, M. Fink, A. C. Boccara, and S. Gigan, “Controlling light through optical disordered media; Transmission matrix approach,” New J. Phys., vol. 13, 2011, Art. no. 123021.