Featured in Physics

Learning and Avoiding Disorder in Multimode Fibers

Maxime W. Matthès,[1] Yaron Bromberg,[2] Julien de Rosny,[1] and Sébastien M. Popoff[1,*]

[1]*Institut Langevin, ESPCI Paris, PSL University, CNRS, France*
[2]*Racah Institute of Physics, The Hebrew University of Jerusalem, Israel*

Multimode optical fibers (MMFs) have gained renewed interest in the past decade, emerging as a way to boost optical communication data rates in the context of an expected saturation of current single-mode fiber-based networks. They are also attractive for endoscopic applications, offering the possibility to achieve a similar information content as multicore fibers, but with a much smaller footprint, thus reducing the invasiveness of endoscopic procedures. However, these advances are hindered by the unavoidable presence of disorder that affects the propagation of light in MMFs and limits their practical applications. We introduce here a general framework to study and avoid the effect of disorder in wave-based systems and demonstrate its application for multimode fibers. We experimentally find an almost complete set of optical channels that are resilient to disorder induced by strong deformations. These *deformation principal modes* are obtained by only exploiting measurements for weak perturbations harnessing the generalized Wigner-Smith operator. We explain this effect by demonstrating that, even for a high level of disorder, the propagation of light in MMFs can be characterized by just a few key properties. These results are made possible thanks to a precise and fast estimation of the modal transmission matrix of the fiber which relies on a model-based optimization using deep learning frameworks.

I. INTRODUCTION

The description of light transport in multimode fibers (MMFs) has been widely studied since the 1970s, with a complete analytical understanding available in the case of an ideal straight fiber [1]. However, imperfections of the fabrication, geometrical deformations, or changes of the environmental conditions introduce randomness that drastically modifies their transmission properties. When light injected in one mode statistically explores all the other modes with the same probability, i.e., in the *strong-coupling regime*, some average properties can be predicted [2]. However, from a few centimeters to a few kilometers, typical MMF systems are neither in the no-coupling nor in the strong-coupling regime; disorder strongly influences light propagation, but some aspects of the ordered behavior survive [3–5]. This intermediate regime has been little investigated so far due to the difficulty to experimentally characterize the effect of disorder on the modal content of the fibers. Understanding the transition between these two

*Corresponding author.
sebastien.popoff@espci.psl.eu

regimes remains an important challenge for optical telecommunications, endoscopic imaging, and micromanipulation applications.

It is well known that injecting coherent light into an MMF results in the observation of a random pattern of bright and dark spots at the output, called *speckle pattern*. However, unlike scattering media, the observation of a speckle is not in itself a signature of disorder. Indeed, perfect straight fibers also exhibit this property due to the existence of intermodal dispersion [6]. As long as multiple modes are excited, they quickly accumulate seemingly random relative phases leading to such complex interference patterns. In the past decade, the measurement of the transmission matrix (TM) emerged as a tool of choice to characterize and control the propagation of light in complex but deterministic optical linear systems. Initially introduced in the context of scattering media [7–9], it consists in measuring the field-field linear relation between an input plane and an output plane. This concept was then applied to MMFs, unlocking new applications for endoscopic imaging [10–12], micromanipulation [13], quantum information processing [14], and for the control of optical channels for telecommunications [15]. It allowed one, in particular, to demonstrate the robustness of the propagating modes in the case of short step-index fibers [6] and bent graded-index fibers [16]. However, the observation of the TM does not directly allow assessing the level of disorder, as dispersion and mode interference lead to the observation of a

seemingly random matrix, even without disorder. Only when represented in the basis of the propagating modes does the TM allow us to fully capture the spatial propagation properties of the MMF. This process can be done by directly injecting light and measuring the field in the mode basis [17] or by numerically projecting a TM measured in a basis of diffraction-limited spots [6]. In both cases, a good characterization is achieved only for the low-order modes, as going into higher-order modes places increasingly demanding requirements on the beam quality and on the alignment [18]. A numerical post-treatment is demonstrated [6] to correct the TM measurement, but it still requires a careful and time-consuming procedure. Moreover, such an approach assumes that there is little to no disorder in the fiber, which forbids the study of the transition from the weak to the strong mode-coupling regimes. One of the main challenges of practical applications of multimode fibers is not only to understand the effect of disorder, but to avoid it altogether. In this context, the time-delay operator introduced in quantum mechanics by Wigner and Smith [19,20] has recently attracted renewed interest among the complex media community. For a lossless optical system characterized by its scattering matrix $\mathbf{S}$, which links all input channels to all output ones, the Wigner-Smith operator is constructed using the frequency derivative of $\mathbf{S}$ and defined as $\mathbf{Q} = -i\mathbf{S}^{-1}\partial_\omega\mathbf{S}$. Interestingly, the eigenstates of this operator, also called *principal modes*, are insensitive to small variations of the frequency. The possibility to use wave front shaping techniques to generate those input states opens new applications to improve some properties of light transport, such as to generate particlelike wave packets in chaotic cavities [21] and in scattering media [22,23] or to optimize energy storage in scattering media [24].

For MMFs, the scattering matrix can be approximated by the TM, whose measurement gives access to the principal modes. In the context of telecommunications, they are particularly attractive, as they do not suffer from modal dispersion to the first order [25]. Their ability to be stable over a large bandwidth is observed in the case of weak [15] and strong disorder [26]. The possibility to find channels invariant to small modifications can be extended to other parameters than the frequency using the generalized Wigner-Smith (GWS) operator [27,28]. These studies focus on the interaction between waves and localized targets in scattering environments in the microwave regime.

In the present paper, we first introduce a new approach that relaxes most of the experimental constraints on the procedure to measure quickly and accurately the TM of an MMF in the mode basis. It uses numerical tools based on a modern machine-learning framework. We demonstrate the ability to use the knowledge of the TMs for small deformations to find an almost complete set of channels using the GWS operator, the deformation principal modes, that are insensitive to strong perturbations. To understand this effect, we show that, all across the deformation range, the evolution of the TM can be characterized by only a few parameters that account for the mode coupling between close-by propagating modes.

II. MISALIGNMENT AND ABERRATION ROBUST CALIBRATION

We first define the TM $\mathbf{H}_{\text{pix}}$ measured in the pixel basis of a modulator and a camera, respectively located in planes conjugated with the input and output facets of a fiber. Leveraging a fast digital micromirror modulator and InGaAs camera, we estimate $\mathbf{H}_{\text{pix}}$ with a 1 kHz frame rate in about 10 s. The principle of the experiment is presented in Fig. 1(a) and detailed in the Appendix A. Ideally, the mode basis representation of the TM can simply be recovered using

$$\mathbf{H_{modes}} = \mathbf{M_o}^\dagger . \mathbf{H_{pix}} . \mathbf{M_i},\qquad(1)$$

where $\mathbf{M_i}$ (respectively, $\mathbf{M_o}$) represents the change of basis matrix between the input (respectively, output) pixel basis and the mode basis of the fiber (see Appendix D for the theoretical mode calculation). We use the orbital angular momentum modes basis, in which the modes are defined by a pair of indices l and m, characterizing, respectively, the radial oscillations and the orbital angular momentum. However, in the presence of slight aberrations or misalignments, the change of basis matrices cannot be inferred only from the calculation of the theoretical mode profiles of the fiber. It leads to strong unwanted distortions of the mode basis TM, even with a carefully tuned setup [6]. Such an effect also occurs when working directly in the mode basis, as injected and detected modes, affected by the aberrations and misalignments, are different from the true ones. To overcome this problem, we design a numerical procedure based on the neural network framework PyTorch [29], taking advantage of graphics processing units (GPUs) for optimized computational times. Unlike neural networks, that consists of generalist layers, typically dense or convolutional layers, we use a model-based approach. Our network is composed of custom layers, that each mimic the effect of an aberration by applying a Zernike phase polynomial to the change of basis matrix. The structure of the network is represented in Fig. 1(d). To take advantage of the deep learning optimization procedures, based on gradient descent, each layer is differentiable. Only one parameter per layer, the strength of the corresponding Zernike polynomial, has to be optimized.

We then train the model parameters to maximize $\|\mathbf{H}_{\text{modes}}\|$, where $\|.\|$ represents the L_2 norm (Frobenius norm) of a matrix. Energy conservation imposes that the input and output projections performed in Eq. (1) lead to $\|\mathbf{H}_{\text{modes}}\| \leq \|\mathbf{H}_{\text{pix}}\|$. Since light can be transmitted through the fiber only by the propagating modes, these two quantities are equal when the matrices $\mathbf{M_i}$ and $\mathbf{M_o}$ correctly compensate for the aberrations and misalignments. Unlike neural networks, we do not need a large training set.
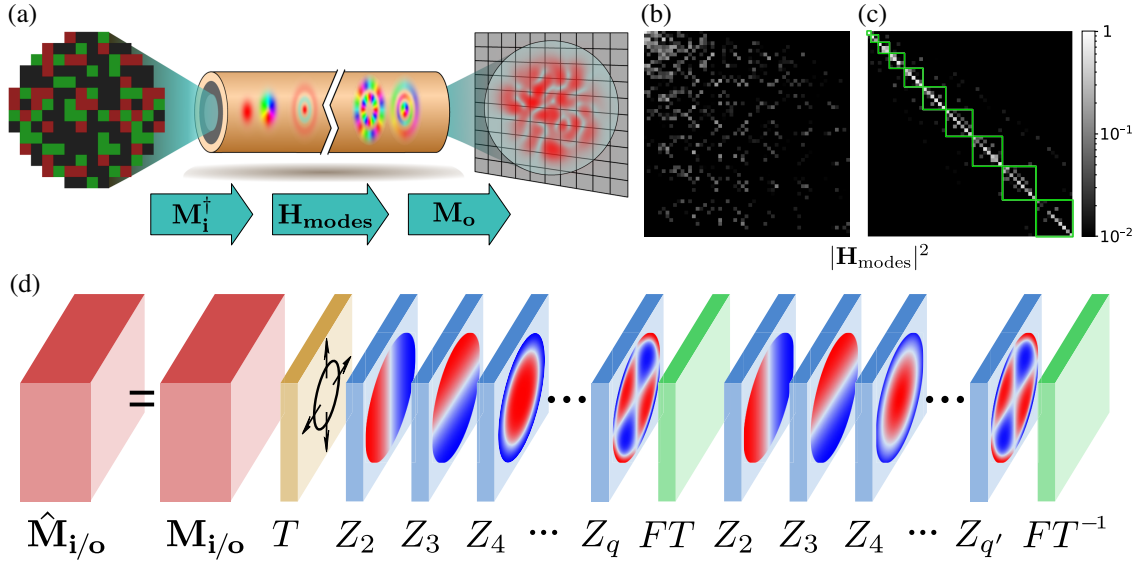
FIG. 1. Principle of the MMF TM reconstruction in the mode basis with the automatic compensation of the aberrations. (a) Simplified sketch of the experiment: The input wave front is modulated using a spatial light modulator and sent on the input facet of an MMF. The light is transmitted through the propagating modes, and the complex output field is imaged onto a camera. The TM is measured in the pixel basis and numerically projected onto the theoretical propagating modes. $\mathbf{M_i}$ and $\mathbf{M_o}$ represent to the input and output change of basis matrices, respectively, and $\mathbf{H}_{\mathrm{modes}}$ is the TM in the mode basis. (b),(c) Intensity of the experimentally measured TMs in the mode basis before and after the numerical compensation of the aberrations for one input and one output polarization. The green squares represent the groups of degenerate modes. The complete TMs are presented in Supplemental Material [30] Sec. S2. (d) Schematic of the model architecture used for the compensation of the aberrations. $\mathbf{M_i}$ and $\mathbf{M_o}$ are modified by differentiable and trainable layers representing a homothety (yellow layer) and phase aberrations characterized by Zernike phase polynomials (blue layers). The Fourier transforms (green layers) allow applying aberrations in the direct and the Fourier planes. The models for the two input and output mode conversions are trained simultaneously against a merit function that maximizes the energy in the projected matrix $\mathbf{H}_{\mathrm{modes}}$ (see Appendix E).

Indeed, we feed to the network only one experimentally measured matrix $\mathbf{H}_{\mathrm{pix}}$. Because of the low number of trainable parameters, 46 corresponding to as many Zernike polynomials plus one for a global scaling on each facet, the optimization takes only a few seconds to converge for a 110-mode fiber. More details about the numerical approach are provided in Appendix E. We effectively shift the complexity of the acquisition from the experimental setup quality to the numerical optimization. It allows changing the fiber in study in a matter of a few seconds, without the need for a precise alignment procedure.

We show in Fig. 1(b) the reconstructed mode basis TM $\mathbf{H}_{\mathrm{modes}}$ assuming no aberration for a 30 cm unperturbed 110-mode graded-index fiber. The TM exhibits little symmetry with high losses for the highest-order modes, reflecting the effect of the aberrations and misalignments. Moreover, only 49% of the energy is conserved in the mode basis. After optimization [Fig. 1(c)], about 94% of the energy is conserved. Moreover, the matrix shows a strong diagonal, which traduces a weak mode-coupling effect. 92% of the energy is in the block diagonal, representing the groups of degenerate modes [in green in Fig. 1(d)]. It is important to stress that the optimization process only maximizes the total energy in the mode basis; the observed strong diagonal appears naturally. The procedure leads to

accurate corrections regardless of the level of disorder (see Supplemental Material [30] S5 for reconstruction comparison for different levels of disorder).

III. PERTURBATION INSENSITIVE CHANNELS

To learn how to be insensitive to disorder, we first characterize the full mode basis TM of an MMF when we introduce and gradually increase a perturbation. We apply a controlled deformation on the fiber along an axis orthogonal to the propagation direction (Fig. 2).

Deformations of the fiber core leads to mode-dependent losses, backreflection, and mode coupling that hinders telecommunication applications. Qualitatively, strong deformations have the effect of progressively populating the off-diagonal elements of the TM while reducing the energy on the diagonal. Thanks to the precise modal projection, we could observe for the first time the crossover from a nearly diagonal TM (weak coupling) [6] to a seemingly random TM (strong coupling) [26].

We first compute the correction for the TM of the unperturbed fiber using our aberration compensation approach. We then apply the same correction parameters for the measurements obtained when the deformation is applied. It ensures that the correction procedure does not
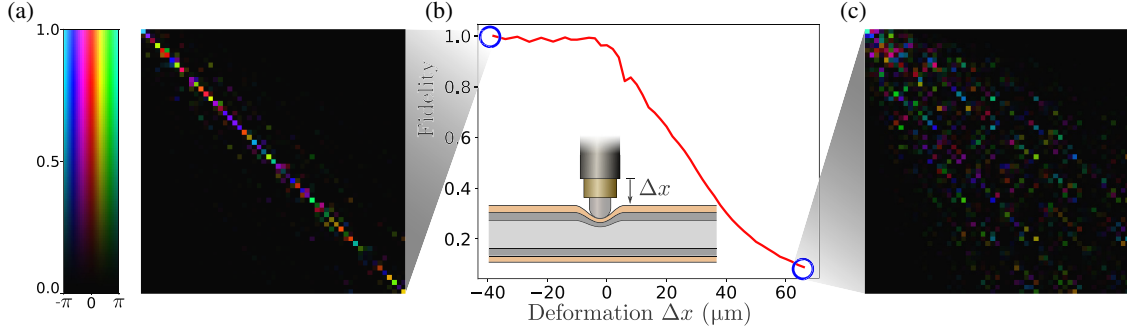
FIG. 2.   Effect of deformations on the mode basis TM. (a),(c) Representation of $\mathbf{H}_{\text{modes}}$ for no deformation and when a transverse deformation $\Delta x = 70~\mu\text{m}$ is applied on the fiber. (b) Fidelity between the TM of the deformed fiber and the unperturbed one as a function of Δx. The fidelity is defined as $F_c = \text{Tr}[|\mathbf{H}_{\text{modes}}(\Delta x).\mathbf{H}_{\text{modes}}(\Delta x = 0)^\dagger|^2]/\sqrt{\text{Tr}[|\mathbf{H}_{\text{modes}}(\Delta x)|^2]\text{Tr}[|\mathbf{H}_{\text{modes}}(\Delta x = 0)|^2]}$. In the inset, we represent a sketch of the deformation procedure.

compensate for some effects of the perturbation. The fidelity between the matrix of the deformed fiber and the reference matrix for the unperturbed configuration (see the caption in Fig. 2) decreases quickly as the displacement Δx increases. While the transmission properties are strongly altered for large deformations, our goal is to find a set of channels that are little affected by them. In the present work, the parameter of interest is the induced displacement Δx; we then study the GWS operator defined as

$$\mathbf{Q}_{\Delta x} = -\frac{i}{2}[\mathbf{H}_{\text{modes}}^{-1}.\partial_{\Delta x}\mathbf{H}_{\text{modes}} - (\mathbf{H}_{\text{modes}}^{-1}.\partial_{\Delta x}\mathbf{H}_{\text{modes}})^\dagger],$$

(2)

where † stands for the conjugate transpose.

The second term appears due to the fact that $\mathbf{H}_{\text{modes}}$ is not unitary [31]. We estimate the GWS operator for a small deformation $\Delta x = 14~\mu\text{m}$. The derivative is numerically estimated using the approximation

$$\partial_{\Delta x}\mathbf{H}_{\text{modes}} \approx \frac{\mathbf{H}_{\text{modes}}(\Delta x_0 + \delta x) - \mathbf{H}_{\text{modes}}(\Delta x_0 - \delta x)}{2\delta x}.$$

(3)

We choose $\delta x = 8~\mu\text{m}$ to mitigate the effect of noise that appears for smaller differences δx.

Its eigenmodes, referred to as the deformation principal modes, are theoretically insensitive to the deformation parameter Δx to the first order. We numerically compute their input profiles and compare the output intensity patterns across the full range of deformations. The stability of the deformation principal modes is shown in Fig. 3 as a function of Δx. For comparison, we also test the injection of the fundamental mode, which is less affected by disorder than the other modes [32], and random wave fronts (the correlation is averaged over 20 realizations). The best deformation principal modes keep a correlation above 95% over the whole range of deformations compared to

the output profile for no deformation. Moreover, all the principal modes but four perform better than the fundamental mode. It is important to note that, while the GWS operator is estimated only for small deformations, it provides an almost complete set of orthogonal channels robust to deformations, even for large values of Δx. The fact that few generalized principal modes do not perform better than the fundamental mode can be attributed to some of the fiber modes close to the cutoff being greatly attenuated (see Supplemental Material [30] S5 for the singular value decomposition of the TM).

IV. DISCUSSION

To further investigate how the deformation principal modes, computed from the TMs for small deformations, can efficiently cancel the effect of large deformations, we study the deformation matrix defined as

$$\mathbf{D}_j = \mathbf{H}_{\text{modes}}^{-1}(\Delta x = 0).\mathbf{H}_{\text{modes}}(\Delta x_j) - \mathbb{I}.$$

(4)

This matrix quantifies how $\mathbf{H}_{\text{modes}}(\Delta x)$ deviates from $\mathbf{H}_{\text{modes}}(\Delta x = 0)$. It is equal to 0 if the TM remains unchanged. We want to determine the main characteristics that best describe how the TM is modified when the perturbation is applied. We then define a *deformation operator* $\bar{\mathbf{D}}$ that links each value of the deformation Δx_j to the corresponding deformation matrix $\mathbf{D}_j$. It characterizes the full evolution of the deformation of the matrix over the range of deformations applied. We compute the singular value decomposition of the operator $\bar{\mathbf{D}}$; it amounts to performing a principal component analysis. The deformation plays the same role as the different realizations in standard principal component analysis implementations (see Appendix C). We show in Fig. 4(a) that the first two singular values amount to more than 96% of the total energy of the operator. The corresponding singular components are represented in Fig. 4(b) for one pair of input
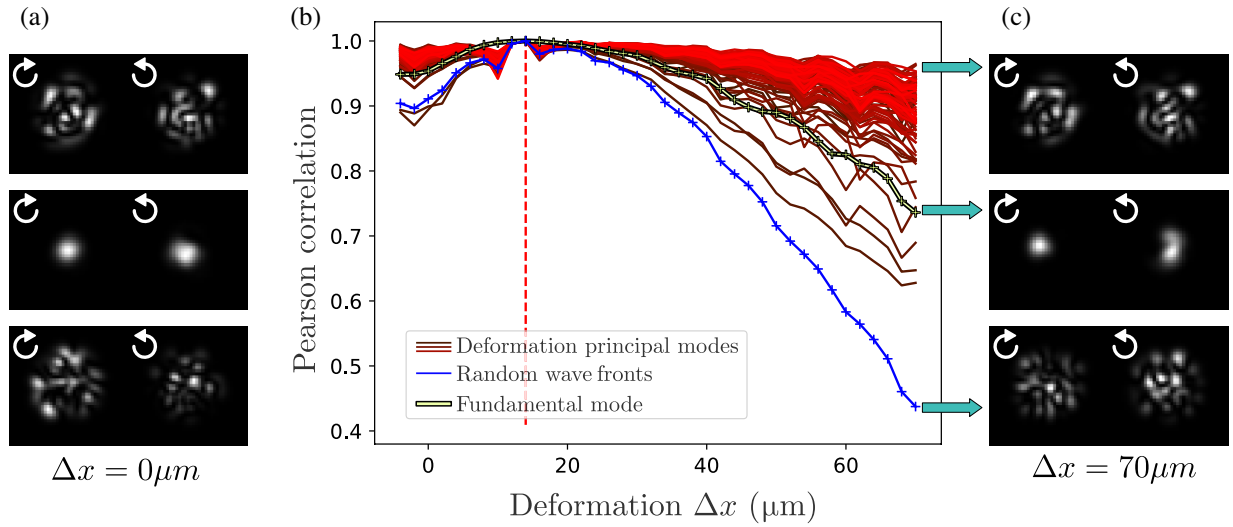
FIG. 3. Examination of the deformation principal modes. (a),(c) Output intensity profiles for the injection of a deformation principal mode, the fundamental mode of the fiber, and a random wave front for the minimal and maximal deformation. Each pair of patterns represents the intensity in the left and right circular polarization states of light. (b) Pearson correlation coefficient between the output intensity pattern at Δx and the one at $\Delta x = 14$ μm for all the deformation principal modes, the fundamental mode, and after averaging over 20 random input wave fronts.

and output polarizations. As the deformation operator is computed for the whole range of deformations, the first principal components characterize the most important modifications applied to the TM during the deformation.

It is shown that, for low perturbations introduced by thermal fluctuations, the distortion of the TM can be parametrized by only one parameter [33]. To test here if the TM of a fiber under strong deformations can be
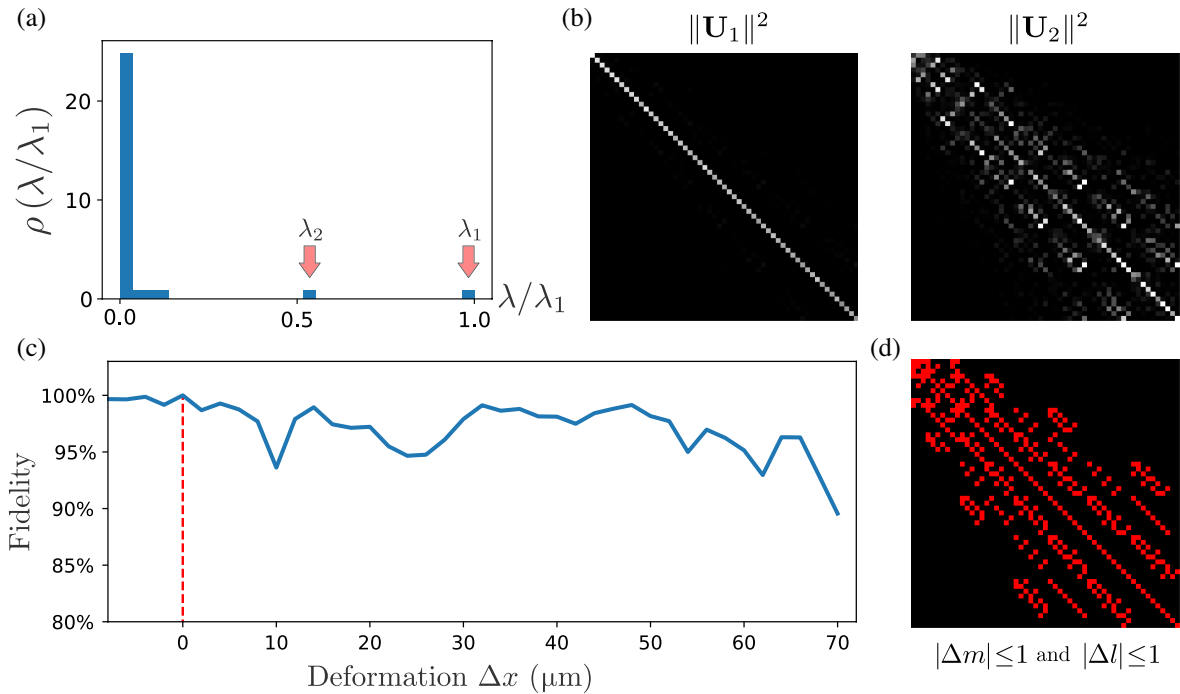

FIG. 4. Analysis of the effect of deformations. (a) Singular value distribution of the deformation operator representing the full range of deformations. (b) Intensity of the first two singular components $\mathbf{U_1}$ and $\mathbf{U_2}$. (d) Matrix representing the coupling between close-by modes, i.e., with a radial order l difference equal or lower than 1 and an orbital momentum m difference equal or lower than 1. (c) Fidelity between the measured TM in the mode basis and the approximated one using Eq. (5) (blue line). The fidelity is defined similarly as in the caption of Fig. 2(b).

parametrized by only a few parameters, we approximate the transmission matrix using just the first two components $\mathbf{U}_1$ and $\mathbf{U}_2$ using

$$\hat{\mathbf{D}}_j = \alpha_j \mathbf{U}_1 + \beta_j \mathbf{U}_2, \qquad (5)$$

where α_j and β_j are directly extracted from the singular value decomposition (see Appendix C). We show in Fig. 4(c) the fidelity between the estimated matrix $\hat{\mathbf{H}}_{\text{modes}}(\Delta x_j) = \mathbf{H}_{\text{modes}}(\Delta x = 0).[\hat{\mathbf{D}}_j + \mathbb{I}]$ and the measured one. Surprisingly, all across the range of deformations, the TM can be estimated using only two parameters with a fidelity above 93%. We can give a qualitative interpretation of the two significant components. $\mathbf{U}_1$ is close to identity, traducing the loss of energy in the diagonal compared to the reference TM at $\Delta x = 0$. It is equivalent to the decay of the ballistic light in the presence of a scattering environment in free space. The second vector $\mathbf{U}_2$ shows a well-defined symmetric pattern that corresponds to an energy conversion between modes with close-by radial and angular momenta l and m [see Figs. 4(b) and 4(d)]. This result is consistent with the previous observations of mode coupling [34] in bent graded-index fibers. It corresponds to photons being injected in one mode and leaving the fiber in a close-by mode, which corresponds to photons whose direction has been modified once by the perturbation. This phenomenon is analogous to the conversion between ballistic and single scattered photons in scattering media. This physical interpretation is made possible thanks to the precise correction of the aberrations that would otherwise destroy the symmetries of $\mathbf{H}_{\text{modes}}$.

The fact that the TM can be estimated precisely using only two terms, only one of them accounting for mode coupling, is counterintuitive considering the fact that $\mathbf{H}_{\text{modes}}$ shows a seemingly random aspect for high-order modes at large deformations [see Fig. 2(c)]. Coupling between modes further away in the l and m space can occur; it is the equivalent of multiple scattered photons in scattering media. However, strong mode coupling also comes with important losses due to coupling to nonguided modes that leak out of the fiber [32], leading to a low-energy contribution of this effect. The fact that the same component $\mathbf{U}_2$ dominates the mode-coupling effect for the whole deformation range explains how the deformation principal modes, estimated for low deformations, are still valid for strong deformations.

V. CONCLUSION

In summary, we present a framework to study the effect of disorder in MMFs, allowing one, in a matter of seconds, to fully characterize light propagation in the mode basis. As precise predictions for the effect of perturbations on multimode fibers in real-life situations are currently lacking, our approach provides a way to quantify such perturbations using measured transmission matrices and could serve as a benchmark for the study of theoretical models. We harness this approach to observe for the first time the existence of deformation principal modes that are robust against strong deformations and that can be found by only using the knowledge of the fiber properties for small deformations. This can be explained by the predominance in the transmission properties of the coupling between nearby modes, even for large deformations. We emphasize that our framework is general and can be used to study any linear propagation system regardless of the presence or the type of perturbations. Moreover, as the complexity of handling the effect of the aberrations and misalignments in the TM estimation is rejected onto a fast automatic postprocessing, our approach is virtually robust to any optical system imperfections, allowing *plug-and-play* operations suitable for real-life applications.

Raw and processed data, custom modules, and sample codes for pre- and postprocessing are available in the dedicated repository [35].

APPENDIX A: EXPERIMENTAL SETUP

The optical setup is represented in Fig. 5. The light source consists of a continuous linearly polarized laser beam at 1550 nm (TeraXion NLL) injected into a 10:90 polarization-maintaining fiber coupler (PNH1550R2F1). The 90% arm is collimated and expanded to illuminate a digital micromirror device (DMD) (Vialux V-650L) composed of 1280 by 800 pixels with a pitch of 10.8 μm working at a maximum frame rate of 10.7 kHz. The light is converted into the left or right circular polarization using a quarter-wave plate and a motorized precision rotation mount (PRM1/MZ8). Two lenses allow the conjugation of the DMD plane with the surface of a standard 30 cm OM2 (50 μm core) graded-index multimode fiber, consisting of a glass core, a glass cladding, and an acrylate coating. The MMF input facet is held by a fiber connector (Thorlabs B30128C3) and a bare fiber terminator (Thorlabs BFT1) and is mounted onto a five-axis translation stage (Thorlabs APY001/M and MAX311D/M). The output
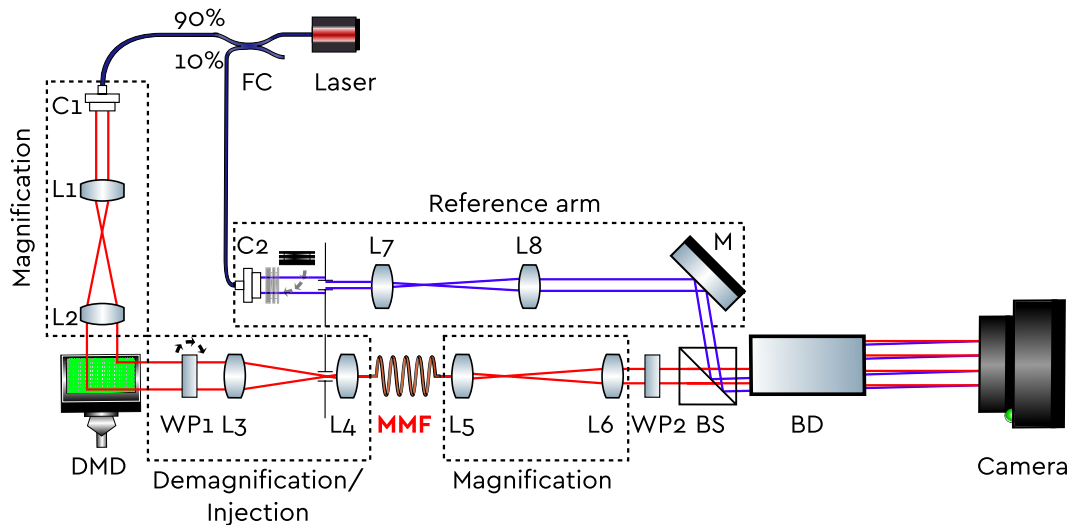
FIG. 5.   Experimental setup. L_i, lenses; C_i, collimators; WP_i, quarter-wave plates; M, mirror; D, diaphragm; BD, beam displacer.

facet is positioned into a V groove (HFV002) and held by magnetic clamps. The fiber is maintained approximately straight. Roughly at half the length, we place a V groove to support the fiber where we introduce a deformation. The perturbation is applied on the fiber by pressing on it using a 50 nm precision dc servo motor actuator (Thorlabs Z812). Magnetic clamps are placed on both sides of the servo motor to prevent the fiber from slipping when the deformation increases. The coating absorbs a significant part of the deformation of the fiber, so that the deformation applied to the fiber core is proportional to but smaller than the translation value Δx. The output facet of the fiber is imaged onto an InGaAs camera (Xenics Cheetah 640-CL 400 Hz) through a beam displacer (Thorlabs BD40) that spatially separates the two polarization contributions in two different areas of the pixel array. A reduced region of interest allows achieving frame rates of about 1 kHz. The 10% arm of the fiber beam splitter is used to illuminate the camera with a tilted reference arm in an off-axis configuration [36].

APPENDIX B: MEASUREMENT OF THE PIXEL BASIS TM

The modulation of the input field is achieved using the Lee hologram method [37]. It allows performing complex amplitude modulation using a binary amplitude DMD [38]. The input modulation patterns consist in square layouts of $N_{\text{pix}}^{\text{in}} = 35 \times 35 = 1225$ square macropixels of size 20×20 pixels. We imprint on each macropixel a periodic pattern of bright (modulation state *on*) and dark (modulation state *off*) stripes. Each macropixel effectively acts as a small grating. The first order of diffraction is selected with a diaphragm represented in Fig. 5. In the plane of the input facet of the MMF, the modulation on the macropixels can be switched on or off by removing the periodic pattern. The phase of the pixels can be modified by offsetting the grating

patterns on each macropixel. We use a grating period of two pixels, giving access to only two levels of phase modulation, 0 and π. We can then create three complex amplitude states: 0, 1, and -1. Sequences of patterns are generated and sent to the control board of the DMD, where they are stored in the onboard memory. The sequence is then displayed at a 1 kHz frame rate on the DMD, which triggers the acquisition of the frames on the camera.

A tutorial on the Lee holograms is made available on our Web site [39]. This modulation procedure is implemented in the PYTHON module SLMlayout [40], and the interface control of the DMD is done using the PYTHON module ALP4lib [41]. We developed, share, and maintain both packages.

The complex output field is measured using an off-axis holographic technique [36]. We share a tutorial on the off-axis holography and some sample codes on our Web site [39]. The complex field is simultaneously measured for the two orthogonal circular polarization states. A quarter-wave plate converts the left and right circular polarizations into two linear orthogonal polarizations. A beam splitter combines the reference arm and the signal arm, and a beam displacer projects the contributions from the two polarization states on two different regions of the camera. For each polarization, the optical field is projected onto a square pattern of 41×41 square macropixels. The field is averaged over each macropixel. The output field for each input wave front is encoded into a vector of size $2N_{\text{pix}}^{\text{out}} = 2 \times 41 \times 41 = 3362$, where $N_{\text{pix}}^{\text{out}}$ is the number of macropixels for each polarization.

The first step of our experiment is to estimate experimentally the TM in the pixel basis $\mathbf{H}_{\text{pix}}$. This matrix describes the linear relationship between the field on one pixel of the modulator to the field on one pixel of the camera. We send a set of input wave fronts described by the vectors X_i, $i \in [1 \ldots N_{\text{masks}}]$ that represent the field on

all the input macropixels. The corresponding output field patterns are represented in the basis of the camera macropixels by the vectors Y_i, $i \in [1\ldots N_{\text{masks}}]$. The relation between the input and output fields reads

$$Y_i = \mathbf{H}_{\text{pix}}.X_i \quad \forall \ i \in [1\ldots N_{\text{masks}}]. \tag{B1}$$

Let us call $\mathbf{X}$ (respectively, $\mathbf{Y}$) the matrix that represents the stack of vectors X_i (respectively, Y_i). Equation (B1) can be rewritten

$$\mathbf{Y} = \mathbf{H}_{\text{pix}}.\mathbf{X}. \tag{B2}$$

An estimation $\hat{\mathbf{H}}_{\text{pix}}$ of the TM can be found by using each vector of the canonical basis for the input excitation patterns, i.e., using $\mathbf{X} = \mathbb{I}$. It gives direct access to the TM using $\hat{\mathbf{H}}_{\text{pix}} = \mathbf{Y}$. One can also use any orthogonal basis, such as the Hadamard basis that is convenient for phase-only modulation [7], so that $\hat{\mathbf{H}}_{\text{pix}} = \mathbf{Y}.\mathbf{X}^{-1}$. However, in the presence of noise, or if one or more measurements fail, the quality of the reconstructed matrix is significantly altered. To mitigate those effects, we choose to use a set of random vectors X_i with $N_{\text{masks}} > N_{\text{pix}}^{\text{in}}$. We can then estimate the TM using

$$\hat{\mathbf{H}}_{\text{pix}} = \mathbf{Y}.\mathbf{X}^+, \tag{B3}$$

where $.^+$ represents the Moore-Penrose pseudoinverse. We choose X_i to be random patterns where the modulation on each pixel can take the value 0, −1, or 1. For each pattern, the percentage of off pixels, i.e., taking the value 0, is drawn from a uniform distribution between 60% and 80%. The percentage of on pixels taking the value 1 is drawn from a uniform distribution between 40% and 60%, the other pixels taking the value −1. The positions of the pixels are random. We choose $N_{\text{masks}} = 6 \times N_{\text{pix}}^{\text{in}} = 7350$ to ensure the existence and the stability of the pseudoinverse of $\mathbf{X}$.

By changing the input polarization state, we measure separately the two corresponding submatrices. They are finally combined into a large matrix of size $2N_{\text{pix}}^{\text{out}} \times 2N_{\text{pix}}^{\text{in}}$.

APPENDIX C: SINGULAR VALUE DECOMPOSITION OF THE DEFORMATION OPERATOR

We first reshape the stack of the matrices $\mathbf{D}_j$ as a two-dimensional matrix $\bar{\mathbf{D}}$ of size $N_{\text{modes}}^2 \times N_{\Delta x}$, where $N_{\Delta x}$ is the number of deformations and $N_{\text{modes}} = 110$ the number of propagating modes. It links each deformation, indexed by j, to all the elements of the matrix $\mathbf{D}_j$, indexed by the composite index $\{kl\} \in [1\ldots N_{\text{modes}}^2]$. The range $[1\ldots N_{\Delta x}]$ of the index j corresponds to deformations between $\Delta x = 0$ μm and the maximal deformation $\Delta x = 70$ μm. Next, we calculate the singular value decomposition of this operator:

$$\bar{\mathbf{D}} = \mathbf{U}.\mathbf{\Lambda}.\mathbf{V}^\dagger. \tag{C1}$$

$\mathbf{\Lambda}$ is a diagonal matrix of size $N_{\Delta x} \times N_{\Delta x}$ containing the singular values, whose distribution is represented in Fig. 4(a). $\mathbf{U}$ is a matrix containing the corresponding output singular vectors U_i, $i \in [1\ldots N_{\Delta x}]$. They can be reshaped as two-dimensional matrices $\mathbf{U}_i$ of size $N_{\text{modes}} \times N_{\text{modes}}$. For any given deformation, we can approximate $\mathbf{D}_j$ using only $\mathbf{U}_1$ and $\mathbf{U}_2$ with Eq. (5). It amounts to replacing Λ in Eq. (C1) by $\tilde{\mathbf{\Lambda}}$ defined by $\tilde{\Lambda}_{11} = \Lambda_{11} = \lambda_1$, $\tilde{\Lambda}_{22} = \Lambda_{22} = \lambda_2$, and $\tilde{\Lambda}_{kl} = 0$ for all other values.

The coefficient α_j and β_j in Eq. (5) are then expressed by

$$\alpha_j = \lambda_1 V_{1j}^*, \tag{C2}$$

$$\beta_j = \lambda_2 V_{2j}^*. \tag{C3}$$

We represent in Fig. 6(a) the evolution of the absolute value of the coefficients of α and β as a function of the deformation. The contribution of α is dominant for small deformations and globally decreases as the deformation increases. Conversely, the contribution of β is small for
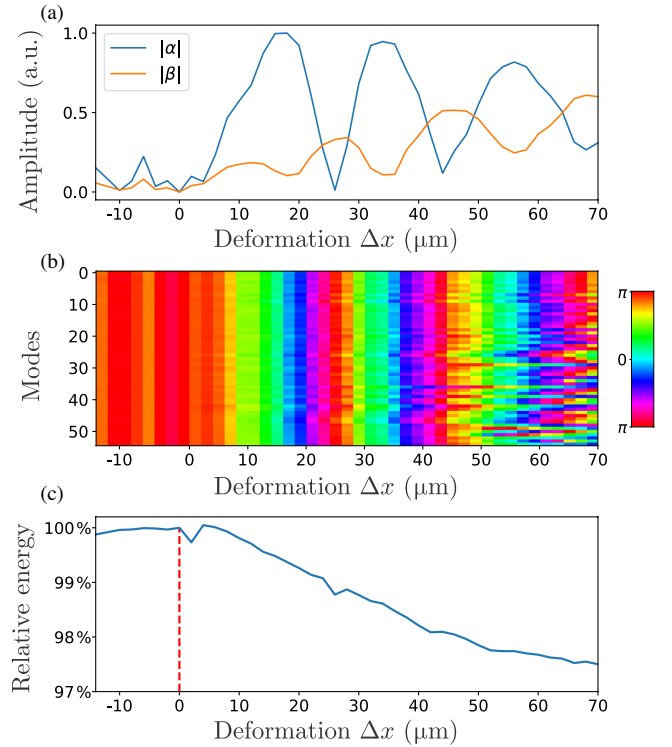

FIG. 6. Analysis of the evolution of the transmission matrix. (a) Coefficients of α and β in Eq. (5) in arbitrary units as a function of the deformation. Coefficients are normalized by the maximal value of $|\alpha|$. (b) Relative phase between the diagonal elements of the mode basis TM as a function of the deformation and the same diagonal elements for $\Delta x = 0$. (c) Energy variation of the mode basis TM as a function of the deformation evaluated as $\|\mathbf{H}_{\text{modes}}(\Delta x)\|^2 / \|\mathbf{H}_{\text{modes}}(\Delta x = 0)\|^2$.

small deformations and globally increases with the deformation. This trend confirms that the first effect to appear is the loss of energy on the diagonal, due to the effect of $\mathbf{U_1}$, and then the coupling to neighboring modes in the momentum space, due to the effect of $\mathbf{U_2}$. We observe that this global trend is modulated by a periodic oscillation. The beating between the two contributions can be attributed to the fact that the two physical effects are not fully decoupled in the two operators, as $\mathbf{U_2}$ also has significant energy on the diagonal that modifies the energy of the ballistic photon similarly to $\mathbf{U_1}$. It is shown that, in addition to mode coupling, deformations are associated with a global rescaling of the fiber which induces phase shifts that dominate for small deformations [6,16,33]. We represent in Fig. 6(b) the evolution of the phase on the diagonal of the mode basis TM. We observe an oscillation with the same periodicity as the beating between α and β. As the perturbation increases, one expects higher-order coupling effects to become significant in Eq. (5), which would be associated with the coupling between modes further away in the momentum space. However, such an effect increases the chance for the photons to couple to nonguided modes, leading to losses [32]. We show in Fig. 6(c) the variation of the energy of the mode basis TM as a function of the deformation. Losses increase with the deformation up to approximately 2.5%, confirming that higher-order coupling effects are still weak in this regime. We restrict ourselves in this study to deformations in the elastic regime of the material, higher deformations leading to nonreversible perturbations and permanent damage of the fiber.

APPENDIX D: CALCULATION OF THE THEORETICAL MODES

The starting point of the mode projection operation is to consider the ideal modes of the fiber. We want to estimate the modes profiles of a perfect straight graded-index fiber under the scalar approximation. Graded-index fiber mode profiles and dispersion relation do not have a closed-form analytical expression. However, approximate analytical expressions can be found, for instance, using perturbation theory or a variational approach [42]. Arguably, the most widely used approximation is the Wentzel-Kramers-Brillouin (WKB) approximation. It leads to an analytical dispersion relation when assuming an infinite quadratic spatial profile of the refractive index. While leading to accurate estimations of the propagation constants, it has a limited accuracy for the expression of the spatial mode profiles [43,44], especially for low radial numbers l. Finite difference methods are easy to implement numerically, but the 2D discretization of the field leads to high memory requirement and computation time and could lead to inaccuracies for high-order modes. Because we consider axiosymmetric index $n(r)$ profiles, we want to simplify the system to solve a 1D problem that depends on only the radial coordinate r, allowing us to increase the accuracy and decrease the computation time.

The 2D scalar Helmholtz equation for a propagating mode can be written in the cylindrical coordinate system as

$$\partial_r^2 \psi(r,\phi) + \frac{1}{r}\partial_r \psi(r,\phi) + \frac{1}{r^2}\partial_\phi^2 \psi(r,\phi)$$
$$+ [k_0^2 n^2(r) - \beta^2] = 0, \tag{D1}$$

where ψ is the optical field, ϕ is the azimuthal coordinate, β is the propagation constant, and $k_0 = 2\pi/\lambda$ with λ the wavelength.

Because the refractive index depends on only the radial coordinate r for a perfect graded-index fiber, we can separate the variables r and ϕ. We are then looking for the orbital angular momentum modes of the form

$$\psi_{ml} = f_l(r)e^{im\phi} \tag{D2}$$

with l the radial order and m the azimuthal order, which also corresponds to the orbital angular momentum. Injecting this expression in Eq. (D1) leads to the 1D equation

$$d_r^2 f_l(r) + \frac{1}{r}d_r f_l(r) + \left[k_0^2 n^2(r) - \beta^2 - \frac{m^2}{r^2}\right]f_l(r) = 0. \tag{D3}$$

The singularity at $r = 0$ arising from the $1/r$ term makes direct finite difference methods unstable. We can use the transformation

$$g_l(r) = \frac{1}{f_l(r)}d_r f_l(r) \tag{D4}$$

and rewrite Eq. (D3) as a quadratic Ricatti equation [45]:

$$d_r g_l(r) + P(r) + Q(r)g_l(r) + g_l^2(r) = 0, \tag{D5}$$

where

$$Q(r) = \frac{1}{r}, \tag{D6}$$

$$P(r) = k_0^2 n^2(r) - \beta^2 - \frac{m^2}{r^2}. \tag{D7}$$

A finite difference approximation of such an equation leads to the recursive expression [45,46]

$$1 + hg_l^{n+1} = \frac{1}{1 + hQ_n/2}\left(h^2 P_n - 2 + \frac{1 - hQ_n/2}{1 + hg_l^n}\right), \tag{D8}$$

where $g_l^n = g_l(r_n)$, $Q_n = Q(r_n)$, $P_n = P(r_n)$, and $h = r_{n+1} - r_n$ is the step size.

The expression (D4) can then be discretized as

$$f_l^{n+1} = f_l^n(1 + hg_l^n). \tag{D9}$$

To find the first steps to initialize the iteration, we need to consider the boundary conditions at the center of the fiber core:

$$\left.\frac{df}{dr}\right|_{r=0} = 0 \quad \text{for } m = 0, \tag{D10}$$

$$f(r = 0) = 0 \quad \text{for } m \neq 0. \tag{D11}$$

For $m = 0$, we discretize the functions at $r_n = nh - 1/2$ and initialize the functions with $f_l^0 = 1$ and $g_l^0 = 0$. For $m \neq 0$, we discretize the functions at $r_n = nh$ and initialize the functions with $f_l^0 = 0$, $f_l^1 = h$, and $g_l^1 = (1 - h^2 P_1)/h$. For a given value of m, the propagation constants β_{ml} that satisfy the Helmholtz equation, corresponding to the propagating modes, are the ones for which the field vanishes at large values of r.

The steps to find the modes of the fiber are the following: We start with $m = 0$ and perform a coarse scan of the propagation constant values between $\beta_{\min} = k_0 n_{\min}$ and $\beta_{\max} = k_0 n_{\max}$. We choose $r_N > a$ large enough to assume that the field at this point, and thus f_N, should be vanishingly small. The number of times $f_N(\beta)$ changes sign gives us the number of propagation modes for the current value of m. It corresponds to the maximal radial number l admissible for the azimuthal number m. We then use a binary search algorithm to find, at a minimum computational cost, the accurate admissible values of β for each l, i.e., the values that minimize f_N under a given tolerance value. We then increment the value of m and repeat the procedure. We stop when no solution is found for the current value of m.

This procedure has been implemented in the PYTHON module pyMMF [47] that we developed and share. Sample codes to compute the ideal modes of the MMF are available at the dedicated repository [35].

APPENDIX E: MODEL-BASED OPTIMIZATION FOR THE COMPENSATION OF THE ABERRATIONS AND MISALIGNMENTS

Recent attempts were made to tackle the problem of modal decomposition using deep learning frameworks. As they use model-free neural network models, using standard convolutional [48] or dense layers [49], these systems require large training sets and significant amounts of memory. Moreover, computational times and limited accuracy forbid their use for more than ten modes (the training for ten modes takes about 43 h with a $> 300\,000$ image training set in Ref. [49]). We develop here a model-based approach that learns only a few relevant parameters, is fast

(a few seconds) to converge, and requires only one TM measurement. The general principle is to apply to the change of basis matrices $\mathbf{M}_i$ and $\mathbf{M}_o$ a set of transformations that mimics the effect of aberrations and misalignments to compensate for the experiment's imperfections. The schematic of the model is presented in Fig. 1(d) in the main text.

In order to implement our model, we first need to use complex-valued matrix operations. However, complex tensors are not natively supported by the PyTorch framework we use. To do so, we add a dimension to our data structure of size 2 to encode the real part and the imaginary part of the complex values. We then create a set of elementary operations: complex conjugation, elementwise, and matrix multiplications. The key parts of our approach are the layers that mimic the effect of aberrations represented by Zernike polynomials. The input of each layer is a batch of complex 2D images of size $N_{\text{modes}} \times N_{\text{pix}} \times N_{\text{pix}} \times 2$. The effect of a layer Z_k, corresponding to the kth Zernike polynomial, is to add, to each 2D image, a phase contribution. It amounts to transforming each input image K_{ij}, $(i, j) \in [1...N_{\text{pix}}] \times [1...N_{\text{pix}}]$ into a modified one K'_{ij} using

$$K'_{ij} = K_{ij} e^{j\alpha_k F_k(r_{ij}, \phi_{ij})}, \tag{E1}$$

where $F_k(r, \phi)$ is the kth Zernike polynomial, r_{ij} and ϕ_{ij} are the polar coordinates corresponding to the pixel indexed by i and j, and α_k is the weight of the aberration. α_k is the only trainable parameter of the layer. The layer automatically calculates and stores the derivative of the output tensor with respect to this parameter, as required for the training process (backpropagation). By adding multiple Zernike layers, we simulate the effect of a high level of aberration. We perform a Fourier transform in the spatial dimensions and add other Zernike layers to simulate aberrations in the Fourier plane [see Fig. 1(d) in the main text]. The first Zernike polynomials correspond to phase slopes in the x and y directions and to a parabolic phase. When applied in the Fourier plane, they introduce spatial shifts in the x and y directions and a defocus. It allows compensating for misalignments in the x, y, and z directions. Finally, we add a transformation T that applies a global scaling transformation in the spatial dimensions. The scaling parameter is the only trainable parameter of this layer.

We treat separately each combination of input and output polarizations. For each optimization, we train simultaneously two models, one for the input and one for the output change of basis matrix. The input data correspond to the matrices $\mathbf{M}_i$ and $\mathbf{M}_o$ of respective size $N^{\text{modes}} \times N_{\text{pix}}^{\text{in}}$ and $N^{\text{modes}} \times N_{\text{pix}}^{\text{out}}$ that we compute using the approach detailed in the previous section. We convert and reshape them as PyTorch tensors of sizes $N^{\text{modes}} \times N_x^{\text{in}} \times N_y^{\text{in}} \times 2$ and

$N^{\mathrm{modes}} \times N_x^{\mathrm{out}} \times N_y^{\mathrm{out}} \times 2$, with $N^{\mathrm{modes}} = 55$ the number of modes per polarization, $N_x^{\mathrm{in}} = N_y^{\mathrm{in}} = \sqrt{N_{\mathrm{pix}}^{\mathrm{in}}} = 35$, and $N_x^{\mathrm{out}} = N_y^{\mathrm{out}} = \sqrt{N_{\mathrm{pix}}^{\mathrm{out}}} = 41$. The first dimension is treated as the batch size in conventional neural networks. The two models return new input and output change of basis matrices $\mathbf{M}'_i$ and $\mathbf{M}'_o$ that are used as input and output projectors, respectively, on the pixel basis TM:

$$\hat{\mathbf{H}}_{\mathrm{modes}} = \mathbf{M}'^{\dagger}_o . \mathbf{H}_{\mathrm{pix}} . \mathbf{M}'_i, \qquad (\text{E2})$$

As explained in the main text, we know that an ideal compensation of the aberrations corresponds to maximizing $\|\mathbf{H}_{\mathrm{modes}}\|$, with $\|.\|$ representing the L_2 norm (Frobenius norm) of a matrix. We choose as the cost function to minimize

$$\mathcal{L} = \frac{\|\mathbf{H}_{\mathrm{pix}}\|}{\|\hat{\mathbf{H}}_{\mathrm{modes}}\|}, \qquad (\text{E3})$$

where $\mathbf{H}_{\mathrm{pix}}$ is the experimentally measured pixel basis TM.

Finally, we run an optimizer based on a stochastic gradient descent approach (Adam optimizer [50]) to find the set of parameters (weights of the Zernike polynomials and the global scaling factors in input and output) that minimizes the cost function $\mathcal{L}$. Once the optimization finished, the obtained change of basis matrices can be used on any newly acquired pixel basis TM as long as the setup stays unchanged. The full optimization takes 18 s on a computer with an Nvidia GeForce 2080 Ti GPU and a Xeon Gold 6142 CPU, 36 s on the same computer with CPU computations only, and 51 s on a regular laptop with an Intel i7-8550U CPU and no GPU. The gain of the GPU computation is expected to increase drastically when the number of modes increases, taking advantage of tensor manipulation optimizations on GPUs.

The full model, the custom layers, and sample codes of aberration correction using experimental data are available at the dedicated repository [35].

[1] D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic, New York, 1974), https://www.elsevier.com/books/theory-of-dielectric-optical-waveguides/marcuse/978-0-12-470950-8.

[2] K.-P. Ho and J. M. Kahn, *Statistics of Group Delays in Multimode Fiber with Strong Mode Coupling*, J. Lightwave Technol. **29**, 3119 (2011).

[3] W. A. Gambling, D. N. Payne, and H. Matsumura, *Mode Conversion Coefficients in Optical Fibers*, Appl. Opt. **14**, 1538 (1975).

[4] R. Ryf, S. Randel, A. H. Gnauck, C. Bolle, A. Sierra, S. Mumtaz, M. Esmaeelpour, E. C. Burrows, R.-J. Essiambre, P. J. Winzer, D. W. Peckham, A. H. McCurdy, and R. Lingle, *Mode-Division Multiplexing over 96 km of Few-Mode Fiber Using Coherent 6×6 MIMO Processing*, J. Lightwave Technol. **30**, 521 (2012).

[5] R. Ryf, N. K. Fontaine, H. Chen, B. Guan, B. Huang, M. Esmaeelpour, A. H. Gnauck, S. Randel, S. Yoo, A. Koonen, R. Shubochkin, Y. Sun, and R. Lingle, *Mode-Multiplexed Transmission over Conventional Graded-Index Multimode Fibers*, Opt. Express **23**, 235 (2015).

[6] M. Plöschner, T. Tyc, and T. Čižmár, *Seeing through Chaos in Multimode Fibres*, Nat. Photonics **9**, 529 (2015).

[7] 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. **104**, 100601 (2010).

[8] Y. Choi, T. D. Yang, C. Fang-Yen, P. Kang, K. J. Lee, R. R. Dasari, M. S. Feld, and W. Choi, *Overcoming the Diffraction Limit Using Multiple Light Scattering in a Highly Disordered Medium*, Phys. Rev. Lett. **107**, 023902 (2011).

[9] S. M. Popoff, G. Lerosey, M. Fink, A. C. Boccara, and S. Gigan, *Image Transmission through an Opaque Material*, Nat. Commun. **1**, 81 (2010).

[10] T. Čižmár and K. Dholakia, *Exploiting Multimode Waveguides for Pure Fibre-Based Imaging*, Nat. Commun. **3**, 1027 (2012).

[11] Y. Choi, C. Yoon, M. Kim, T. D. Yang, C. Fang-Yen, R. R. Dasari, K. J. Lee, and W. Choi, *Scanner-Free and Wide-Field Endoscopic Imaging by Using a Single Multimode Optical Fiber*, Phys. Rev. Lett. **109**, 203901 (2012).

[12] I. N. Papadopoulos, S. Farahi, C. Moser, and D. Psaltis, *Focusing and Scanning Light through a Multimode Optical Fiber Using Digital Phase Conjugation*, Opt. Express **20**, 10583 (2012).

[13] S. Bianchi and R. Di Leonardo, *A Multi-mode Fiber Probe for Holographic Micromanipulation and Microscopy*, Lab Chip **12**, 635 (2012).

[14] S. Leedumrongwatthanakun, L. Innocenti, H. Defienne, T. Juffmann, A. Ferraro, M. Paternostro, and S. Gigan, *Programmable Linear Quantum Networks with a Multimode Fibre*, Nat. Photonics **14**, 139 (2020).

[15] J. Carpenter, B. J. Eggleton, and J. Schröder, *Observation of Eisenbud–Wigner–Smith States as Principal Modes in Multimode Fibre*, Nat. Photonics **9**, 751 (2015).

[16] D. E. B. Flaes, J. Stopka, S. Turtaev, J. F. de Boer, T. Tyc, and T. Čižmár, *Robustness of Light-Transport Processes to Bending Deformations in Graded-Index Multimode Waveguides*, Phys. Rev. Lett. **120**, 233901 (2018).

[17] J. Carpenter, B. J. Eggleton, and J. Schröder, 110×110 *Optical Mode Transfer Matrix Inversion*, Opt. Express **22**, 96 (2014).

[18] F. Stutzki, H.-J. Otto, F. Jansen, C. Gaida, C. Jauregui, J. Limpert, and A. Tünnermann, *High-Speed Modal Decomposition of Mode Instabilities in High-Power Fiber Lasers*, Opt. Lett. **36**, 4572 (2011).

[19] E. P. Wigner, *Lower Limit for the Energy Derivative of the Scattering Phase Shift*, Phys. Rev. **98,** 145 (1955).

[20] F. T. Smith, *Lifetime Matrix in Collision Theory*, Phys. Rev. **118,** 349 (1960).

[21] S. Rotter, P. Ambichl, and F. Libisch, *Generating Particle-like Scattering States in Wave Transport*, Phys. Rev. Lett. **106,** 120602 (2011).

[22] B. Gérardin, J. Laurent, P. Ambichl, C. Prada, S. Rotter, and A. Aubry, *Particlelike Wave Packets in Complex Scattering Systems*, Phys. Rev. B **94,** 014209 (2016).

[23] J. Böhm, A. Brandstötter, P. Ambichl, S. Rotter, and U. Kuhl, *In Situ Realization of Particlelike Scattering States in a Microwave Cavity*, Phys. Rev. A **97,** 021801 (2018).

[24] M. Durand, S. M. Popoff, R. Carminati, and A. Goetschy, *Optimizing Light Storage in Scattering Media with the Dwell-Time Operator*, Phys. Rev. Lett. **123,** 243901 (2019).

[25] S. Fan and J. M. Kahn, *Principal Modes in Multimode Waveguides*, Opt. Lett. **30,** 135 (2005).

[26] W. Xiong, P. Ambichl, Y. Bromberg, B. Redding, S. Rotter, and H. Cao, *Spatiotemporal Control of Light Transmission through a Multimode Fiber with Strong Mode Coupling*, Phys. Rev. Lett. **117,** 053901 (2016).

[27] P. Ambichl, A. Brandstötter, J. Böhm, M. Kühmayer, U. Kuhl, and S. Rotter, *Focusing inside Disordered Media with the Generalized Wigner-Smith Operator*, Phys. Rev. Lett. **119,** 033903 (2017).

[28] M. Horodynski, M. Kühmayer, A. Brandstötter, K. Pichler, Y. V. Fyodorov, U. Kuhl, and S. Rotter, *Optimal Wave Fields for Micromanipulation in Complex Scattering Environments*, Nat. Photonics **14,** 149 (2020).

[29] A. Paszke *et al.*, PyTorch: *An Imperative Style, High-Performance Deep Learning Library*, in *Advances in Neural Information Processing Systems 32*, edited by H. Wallach *et al.* (Curran Associates, Red Hook, NY, 2019), pp. 8026–8037.

[30] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevX.11.021060 for addition data about the measured transmission matrices, the aberration correction and the deformation principal modes.

[31] A. Goetschy and M. Durand (private communication).

[32] D. Marcuse, *Field Deformation and Loss Caused by Curvature of Optical Fibers*, J. Opt. Soc. Am. **66,** 311 (1976).

[33] J. Yammine, A. Tandjè, M. Dossou, L. Bigot, and E. R. Andresen, *Time-Dependence of the Transmission Matrix of a Specialty Few-Mode Fiber*, APL Photonics **4,** 022904 (2019).

[34] S. Li, C. Saunders, D. J. Lum, J. Murray-Bruce, V. K. Goyal, T. Cizmar, and D. B. Phillips, *Compressively Sampling the Optical Transmission Matrix of a Multimode Fibre*, arXiv:2007.15891v1.

[35] http://github.com/wavefrontshaping, http://github.com/wavefrontshaping/article_MMF_disorder [accessed 25 November 2020].

[36] E. Cuche, P. Marquet, and C. Depeursinge, *Spatial Filtering for Zero-Order and Twin-Image Elimination in Digital Off-Axis Holography*, Appl. Opt. **39,** 4070 (2000).

[37] W.-H. Lee, *III Computer-Generated Holograms: Techniques and Applications* (Elsevier, New York, 1978), pp. 119–232.

[38] D. B. Conkey, A. M. Caravaca-Aguirre, and R. Piestun, *High-Speed Scattering Medium Characterization with Application to Focusing Light through Turbid Media*, Opt. Express **20,** 1733 (2012).

[39] Wavefrontshaping.net tutorials, https://www.wavefrontshaping.net/tutorials.

[40] S. M. Popoff and M. W. Matthès, *SLMlayout: A Module to Generate Patterns for Complex Modulation with Liquid Crystal and Micromirror Spatial Light Modulators*, https://doi.org/10.5281/zenodo.4075519.

[41] S. M. Popoff and M. W. Matthès, *ALP4lib: A* PYTHON *Wrapper for the Vialux ALP-4 Controller Suite to Control DMDs*, https://doi.org/10.5281/zenodo.4076193.

[42] A. Sharma and P. Bindal, *International Centre for Theoretical Physics, Solutions of the 2-d Helmholtz Equation for Optical Waveguides: Semi-analytical and Numerical Variational Approaches*, LAMP Series Report No. LAMP/92, 1992.

[43] A. Gedeon, *Comparison between Rigorous Theory and WKB-Analysis of Modes in Graded-Index Waveguides*, Opt. Commun. **12,** 329 (1974).

[44] L. Maksymiuk and G. Stepniak, *On the Precision of Mode Delays Derivation with the Use of the WKB Method*, Opt. Quantum Electron. **48,** 25 (2016).

[45] L. S. Tamil, S. S. Mitra, R. Dutta, and J. M. T. Pereira, *Finite Difference Solution for Graded-Index Cylindrical Dielectric Waveguides: A Scalar Wave Approximation*, Appl. Opt. **30,** 1113 (1991).

[46] G. File and T. Aga, *Numerical Solution of Quadratic Riccati Differential Equations*, Egypt. J. Basic Appl. Sci. **3,** 392 (2016).

[47] S. M. Popoff, *pyMMF: A Simple Module in* PYTHON *for Numerically Finding Multimode Fiber Modes under the Scalar Approximation*, https://doi.org/10.5281/zenodo.4075298.

[48] Y. An, J. Li, L. Huang, L. Li, J. Leng, L. Yang, and P. Zhou, *Numerical Mode Decomposition for Multimode Fiber: From Multi-variable Optimization to Deep Learning*, Opt. Fiber Technol. **52,** 101960 (2019).

[49] S. Rothe, Q. Zhang, N. Koukourakis, and J. W. Czarske, *Intensity-Only Mode Decomposition on Multimode Fibers Using a Densely Connected Convolutional Network*, arXiv:2008.00864v2.

[50] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, arXiv:1412.6980.