

# Nonconvex optimization for optimum retrieval of the transmission matrix of a multimode fiber

Shengfu Cheng<sup>a,b,†</sup> Xuyu Zhang,<sup>c,d,†</sup> Tianting Zhong,<sup>a,b,†</sup> Huanhao Li,<sup>a,b</sup> Haoran Li,<sup>a,b</sup> Lei Gong,<sup>e</sup> Honglin Liu,<sup>b,c,f,\*</sup> and Puxiang Lai<sup>a,b,g,\*</sup>

<sup>a</sup>The Hong Kong Polytechnic University, Department of Biomedical Engineering, Hong Kong, China

<sup>b</sup>The Hong Kong Polytechnic University, Shenzhen Research Institute, Shenzhen, China

<sup>c</sup>Chinese Academy of Sciences, Shanghai Institute of Optics and Fine Mechanics, Key Laboratory for Quantum Optics, Shanghai, China

<sup>d</sup>University of Shanghai for Science and Technology, School of Optical-Electrical and Computer Engineering, Shanghai, China

<sup>e</sup>University of Science and Technology of China, Department of Optics and Optical Engineering, Hefei, China

<sup>f</sup>University of Chinese Academy of Sciences, Center of Materials Science and Optoelectronics Engineering, Beijing, China

<sup>g</sup>The Hong Kong Polytechnic University, Photonics Research Institute, Hong Kong, China

**Abstract.** Transmission matrix (TM) allows light control through complex media, such as multimode fibers (MMFs), gaining great attention in areas, such as biophotonics, over the past decade. Efforts have been taken to retrieve a complex-valued TM directly from intensity measurements with several representative phase-retrieval algorithms, which still see limitations of slow or suboptimum recovery, especially under noisy environments. Here, we propose a modified nonconvex optimization approach. Through numerical evaluations, it shows that the optimum focusing efficiency is approached with less running time or sampling ratio. The comparative tests under different signal-to-noise levels further indicate its improved robustness. Experimentally, the superior focusing performance of our algorithm is collectively validated by single- and multispot focusing; especially with a sampling ratio of 8, it achieves a 93.6% efficiency of the gold-standard holography method. Based on the recovered TM, image transmission through an MMF is realized with high fidelity. Due to parallel operation and GPU acceleration, our nonconvex approach retrieves a  $8685 \times 1024$  TM (sampling ratio is 8) with 42.3 s on average on a regular computer. The proposed method provides optimum efficiency and fast execution for TM retrieval that avoids the need for an external reference beam, which will facilitate applications of deep-tissue optical imaging, manipulation, and treatment.

Keywords: transmission matrix; phase retrieval; multimode fiber imaging; wavefront shaping.

Received Aug. 1, 2023; revised manuscript received Oct. 18, 2023; accepted for publication Oct. 25, 2023; published online Nov. 17, 2023.

© The Authors. Published by SPIE and CLP under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI.

[DOI: [10.1117/1.APN.2.6.066005](https://doi.org/10.1117/1.APN.2.6.066005)]

## 1 Introduction

Different from ordinary ballistic optics, light propagation in complex media is highly disordered<sup>1,2</sup> due to the multiple scattering occurring in these media, such as biological tissues or mode dispersion in multimode fiber (MMF). Finding an order out of such disorders has been long pursued. Over the past decade,

enormous progress has been made via wavefront shaping<sup>3–8</sup> and especially the transmission-matrix (TM) method<sup>9–13</sup> in controlling light to focus and image through complex media. The TM of a disordered medium describes the complex output responses for an arbitrary point-source input, which is regarded as the transfer function of the medium under the shift-invariance assumption.<sup>11</sup> The measurement of TM offers a versatile tool to control light delivery in spite of scattering,<sup>6,10</sup> as well as recovering object information from acquired speckle patterns.<sup>14,15</sup> The TM method has spurred a wide range of MMF-based applications, such as focusing,<sup>16,17</sup>

\*Address all correspondence to Honglin Liu, [hlliu4@hotmail.com](mailto:hlliu4@hotmail.com); Puxiang Lai, [puxiang.lai@polyu.edu.hk](mailto:puxiang.lai@polyu.edu.hk)

<sup>†</sup>These authors contributed equally to this work.

glare suppression,<sup>18,19</sup> endoscopic imaging,<sup>20–23</sup> manipulation,<sup>24</sup> optogenetics,<sup>25</sup> and communication.<sup>14,26,27</sup>

TM measurement of a scattering medium was first introduced by Popoff et al.<sup>9,10</sup> using coaxial holography with an internal reference. Since then, various forms of TM measurement have been developed. Typically, the TM is measured by recording the complex output fields under a sequence of input modulations. The modulation basis is usually orthogonal, which can be of diverse forms, including the Hadamard matrix,<sup>9,10</sup> the Fourier-transform matrix,<sup>28</sup> point source,<sup>29,30</sup> and random phase.<sup>13</sup> Regardless of the form, the measured TM relates all input modes to each output mode by linear superposition. Depending on the type of input modulation and output measurement, the TM could be complex-valued,<sup>9,30</sup> real-valued,<sup>31,32</sup> or even binary.<sup>33</sup> Among them, complex TM is used most, as it supports both amplitude and phase modulation of light, which, however, usually entails holographic setup. Off-axis holography based on either phase shifting<sup>34</sup> or spatial filtering<sup>35</sup> can acquire the complex TM accurately. Nevertheless, effective off-axis interferometry with an additional reference beam and the high system stability it demands could be impractical in some scenarios. As an example, the coherence length of pulse laser could be too short to use for interferometry-based TM measurement. With coaxial holography, the above issues might be alleviated, but the dark-spot problem with the measured TM caused by speckle reference field<sup>36</sup> is still unsatisfying.

Recent efforts have been made to accurately retrieve the complex TM from intensity-only measurements using advanced phase-retrieval algorithms,<sup>13,28,37–43</sup> which started with a Bayesian inference approach [i.e., phase retrieval Variational Bayes Expectation-Maximization (prVBEM)].<sup>13</sup> This was followed by variants of approximating message passing (AMP) algorithm such as phase retrieval Swept AMP (prSAMP)<sup>37</sup> and phase retrieval Vector AMP (prVAMP).<sup>38</sup> Although robust to noise, a prior knowledge of noise statistics is a must for these approaches. Semidefinite programming (SDP) that uses convex relations has also been introduced for solving the TM retrieval problem,<sup>39</sup> but it usually requires  $N \ln N$  ( $N$  is the input size) measurements and tends to be computationally inefficient. In addition, works based on the extended Kalman filter (EKF)<sup>40</sup> or the generalized Gerchberg–Saxton (GGS) algorithm<sup>41</sup> claim retrieving TM with measurements could be enough. That said, EKF is computationally burdened and hard for parallelization. GGS is efficient in computation, but its performance is still suboptimum in real practice. Most recently, the area also sees the birth of a smoothed Gerchberg–Saxton algorithm<sup>43</sup> and a nonlinear optimization method<sup>28</sup> for TM retrieval.

To overcome the aforementioned limitations, in this study, a state-of-the-art nonconvex optimization approach is adopted and modified for TM retrieval with optimum performance. Compared with existing TM retrieval algorithms, the proposed nonconvex method can achieve optimal efficiency numerically with less running time or sampling ratio. In the experiment, by focus scanning across the field-of-view of an MMF with the acquired TM, the performance of the proposed method is validated to approach the gold standard (i.e., off-axis holography) with a sampling ratio of 8. Moreover, with the assistance of parallel operation and GPU acceleration, multiple rows of TM can be recovered rapidly. Our method for TM retrieval is featured with optimum efficiency and fast implementation in a reference-less and robust setting, which holds potential for many deep-tissue imaging and focusing applications with the usage of MMF.

## 2 Methods

### 2.1 Formulation of the TM Retrieval Problem

The theoretical model of retrieving a TM under a sequence of input modulations is formulated as follows. Suppose the number of discrete modulation units (i.e., input size) and speckle field pixels (i.e., output size) is  $N$  and  $M$ , respectively. Given  $P$  calibration patterns such that the probe matrix  $\mathbf{X} \in \mathbb{C}^{N \times P}$  and the amplitude measurements  $\mathbf{Y} \in \mathbb{R}_+^{M \times P}$ , the TM  $\mathbf{A} \in \mathbb{C}^{M \times N}$  that needs to be estimated as follows:

$$\mathbf{Y} = |\mathbf{A}\mathbf{X}|, \quad (1)$$

where  $|\cdot|$  takes the absolute value for the elements inside. By taking the conjugate transpose of both sides of Eq. (1), we have

$$\mathbf{Y}^H = |\mathbf{X}^H \mathbf{A}^H|, \quad (2)$$

where  $(\cdot)^H$  is the element-wise conjugate transpose operator. Column-wisely,  $\mathbf{Y}^H = [\mathbf{y}_1, \dots, \mathbf{y}_j, \dots, \mathbf{y}_M]$ , where  $\mathbf{y}_i \in \mathbb{R}_+^P$  constitutes the measurements associated with the  $i$ 's output mode, and  $\mathbf{A}^H = [\mathbf{a}_1, \dots, \mathbf{a}_j, \dots, \mathbf{a}_M]$ , where  $\mathbf{a}_i \in \mathbb{C}^N$  denotes the conjugate transpose of the  $i$ 'th row of TM,  $i = 1, \dots, M$ . In this case, the TM retrieval problem is decomposed into  $M$  independent subproblems given as

$$\mathbf{y}_i = |\mathbf{X}^H \mathbf{a}_i|, \quad i = 1, \dots, M. \quad (3)$$

Due to the operation of taking absolute values in Eq. (3), the above problem of estimating one row of TM is nonlinear and falls in the category of phase retrieval.

The phase-retrieval problem has been studied intensively in mathematics, as it is commonly encountered in practice, with representative algorithms including alternating projection<sup>44</sup> (e.g., Gerchberg–Saxton and Fineup), SDP<sup>45</sup> (e.g., PhaseLift, PhaseMax, PhaseCut), approximate message passing (e.g., Generalized AMP<sup>46</sup> and Vector AMP<sup>47</sup>), and nonconvex optimization.<sup>48–52</sup> Among these methods, nonconvex approaches are proven to be superior and have been developed rapidly in recent years. There are mainly two categories of nonconvex approaches: the intensity-flow<sup>49,53</sup> and amplitude-flow models,<sup>50–52</sup> with the latter being better in both empirical success rate and convergence property. In particular, the amplitude-flow models have been proven to converge linearly to the true solution under  $O(n)$  Gaussian measurements for a signal with dimension  $n$ .<sup>52</sup>

### 2.2 Modified Reweighted Amplitude Flow Algorithm

Herein, the cutting-edge reweighted amplitude-flow (RAF) algorithm<sup>52</sup> is adopted for the TM retrieval problem. Solving Eq. (3) can be recast as an optimization problem,

$$\min_{\mathbf{a}_i \in \mathbb{C}^N} L(\mathbf{a}_i) = \|\mathbf{X}^H \mathbf{a}_i - \mathbf{y}_i\|_2^2, \quad (4)$$

where  $\|\cdot\|_2$  denotes the L2 norm of a vector, and  $L(\mathbf{a}_i)$  is an amplitude-based least square error (LSE) loss function. While most nonconvex algorithms contain two stages (i.e., spectral initialization and gradient descent), RAF applies reweighting techniques in both stages that accelerates the signal recovery significantly. Considering Eq. (4), the signal [i.e., one row of TM  $\mathbf{a}$  (the row index  $i$  is omitted for genericity)] is first estimated with the weighted maximum correlation

initialization. A subset of the row vectors in the probe matrix ( $\mathbf{X}^H = [\mathbf{x}_1^H; \dots; \mathbf{x}_p^H]$ ) that correspond to the  $|S|$  (subset  $S \subset \{1, \dots, P\}$ ) largest entries in the measurements  $\mathbf{y} = \{y_j\}_{1 \leq j \leq P}$  are selected, which are called direction vectors, as they are more correlated to the true signal. The signal can be estimated by maximizing its correlation to the direction vectors  $\{\mathbf{x}_j^H | j \in S\}$  such that

$$\max_{\|\mathbf{a}\|=1} \frac{1}{|S|} \sum_{j \in S} |\langle \mathbf{x}_j^H, \mathbf{a} \rangle|^2 = \max_{\|\mathbf{a}\|=1} \mathbf{a}^H \left( \frac{1}{|S|} \sum_{j \in S} \mathbf{x}_j \mathbf{x}_j^H \right) \mathbf{a}. \quad (5)$$

By weighting more to the selected  $\mathbf{x}_j^H$  vectors corresponding to larger  $y_j$  values with the weights  $y_j^\alpha$ ,  $\forall j \in S$  (exponent  $\alpha = 0.5$ , by default), the solution  $\tilde{\mathbf{a}}^0$  of Eq. (5) is the unit-norm principle eigenvector of the Hermitian matrix,

$$\mathbf{H} = \frac{1}{|S|} \sum_{j \in S} y_j^\alpha \mathbf{x}_j \mathbf{x}_j^H = \frac{1}{|S|} \mathbf{X} \cdot \text{diag}(\tilde{y}_1^\alpha, \tilde{y}_2^\alpha, \dots, \tilde{y}_p^\alpha) \cdot \mathbf{X}^H, \quad (6)$$

where  $\tilde{y}_j^\alpha = \begin{cases} y_j^\alpha, & j \in S \\ 0, & \text{otherwise} \end{cases}$ .  $\tilde{\mathbf{a}}^0$  is then scaled to obtain the signal initial guess  $\mathbf{a}^0 = \sqrt{\sum_{j=1}^P y_j^2 / P} \cdot \tilde{\mathbf{a}}^0$ .

The initialized signal  $\mathbf{a}^0$  is further refined by reweighted “gradient-like” iterations. The gradient of the LSE loss in Eq. (4) with respect to  $\mathbf{a}$  is

$$\nabla L(\mathbf{a}) = \frac{1}{P} \cdot \mathbf{X} \left( \mathbf{X}^H \mathbf{a} - \mathbf{y} \circ \frac{\mathbf{X}^H \mathbf{a}}{|\mathbf{X}^H \mathbf{a}|} \right), \quad (7)$$

where  $\circ$  denotes element-wise multiplication. Let  $t$  be the iteration index; then the gradient descent is described as

$$\mathbf{a}^{t+1} = \mathbf{a}^t - \mu \cdot \nabla L(\mathbf{a}^t), \quad (8)$$

where  $\mu$  is the step size, and  $t = 0, 1, \dots$ . One can reweight the gradients in Eq. (7) that have larger values of  $|\mathbf{X}^H \mathbf{a}^t| \oslash \mathbf{y}$  ( $\oslash$  represents element-wise division), which are deemed more reliable in directing to the true signal. The adaptive weight vector is

$$\mathbf{w} = |\mathbf{X}^H \mathbf{a}^t| \oslash (|\mathbf{X}^H \mathbf{a}^t| + \beta \mathbf{y}), \quad (9)$$

where  $\beta$  is a preselected parameter. The above descriptions show the reweighted gradient flow for TM retrieval.

Inspired by Ref. 41, herein we modify the gradient-descent process of the original RAF algorithm, which is divided into two steps. In Step 1, the normalized measurement vector  $\mathbf{y}$  is replaced with its square  $\mathbf{y}^2$  for gradient computation, which enlarges the gradient and functions as the artificial heat data. In Step 2, the amplitude measurement  $\mathbf{y}$  is still used. Step 1 contains the first two-thirds of total iterations, which is set empirically (the rationale is referred to in Appendix A). The modified algorithm is simple, yet surprisingly effective, which is named RAF 2-1 and shown to reduce the measurement error to a much lower level than the original RAF (see Appendix B). The pseudo-code of retrieving one row of TM with RAF 2-1 is summarized in Algorithm 1. The time complexity for spectral initialization and gradient iteration in Algorithm 1 is  $O(N|S|)$  (with the adaptaion of power method or Lanczos algorithm)<sup>52</sup> and  $O(NP)$  (usually  $P \geq N$ ), respectively, so that the total time complexity is (at least)  $O(N^2)$  for retrieving a TM row. Luckily, multiple rows of TM can easily be retrieved in a parallel way.

**Algorithm 1** RAF 2-1 for retrieving a TM row  $\mathbf{a} \in \mathbb{C}^N$ .

- 1: **Input:** Data  $\mathbf{y} \in \mathbb{R}_+^P$  with  $\{y_j\}_{1 \leq j \leq P}$ ,  $\mathbf{X} \in \mathbb{C}^{N \times P}$ ; number of iterations  $T$ ; step size  $\mu = 3$  and weighting parameter  $\beta = 5$ ; subset cardinality  $|S| = \lfloor 3P/13 \rfloor$ , and exponent  $\alpha = 0.5$ .
- 2: **Construct**  $S$  to include indices associated with the  $|S|$  largest entries among  $\{y_j\}_{1 \leq j \leq P}$ .
- 3: **Initialization:** Let  $\mathbf{a}^0 := \sqrt{\sum_{j \in S} y_j^2 / P} \cdot \tilde{\mathbf{a}}^0$  where  $\tilde{\mathbf{a}}^0$  is the unit-norm principle eigenvector of the Hermitian matrix

$$\mathbf{H} := \frac{1}{|S|} \mathbf{X} \cdot \text{diag}(\tilde{y}_1^\alpha, \tilde{y}_2^\alpha, \dots, \tilde{y}_p^\alpha) \cdot \mathbf{X}^H,$$

$$\text{where } \tilde{y}_j^\alpha := \begin{cases} y_j^\alpha, & j \in S \\ 0, & \text{otherwise} \end{cases}.$$

- 4: **Gradient-descent loop**

Step 1: for  $t = 0$  to  $\lfloor \frac{2}{3} T \rfloor - 1$  do

$$\mathbf{a}^{t+1} = \mathbf{a}^t - \frac{\mu}{P} \cdot \mathbf{X} \left[ \mathbf{w} \circ \left( \mathbf{X}^H \mathbf{a}^t - \mathbf{y}^2 \circ \frac{\mathbf{X}^H \mathbf{a}^t}{|\mathbf{X}^H \mathbf{a}^t|} \right) \right]$$

Step 2: for  $t = \lfloor \frac{2}{3} T \rfloor$  to  $T - 1$  do

$$\mathbf{a}^{t+1} = \mathbf{a}^t - \frac{\mu}{P} \cdot \mathbf{X} \left[ \mathbf{w} \circ \left( \mathbf{X}^H \mathbf{a}^t - \mathbf{y} \circ \frac{\mathbf{X}^H \mathbf{a}^t}{|\mathbf{X}^H \mathbf{a}^t|} \right) \right]$$

where  $\mathbf{w} := |\mathbf{X}^H \mathbf{a}^t| \oslash (|\mathbf{X}^H \mathbf{a}^t| + \beta \mathbf{y})$ .

- 5: **Output:**  $\mathbf{a}$ .

## 3 Results

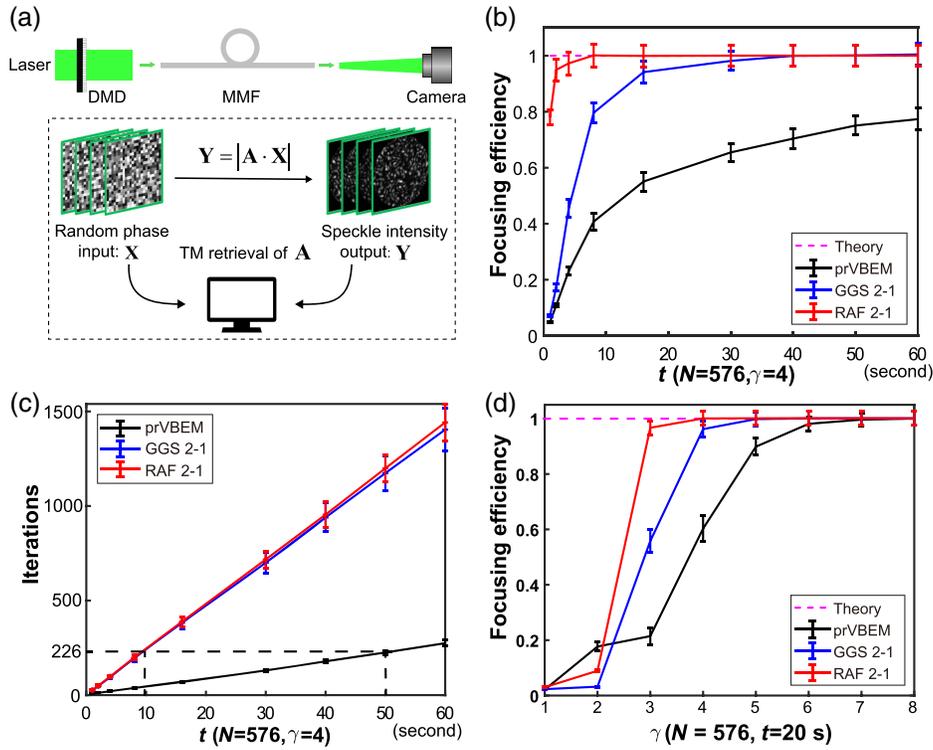
### 3.1 Numerical Evaluation

Numerical evaluations are conducted at first to assess the efficiency of the proposed RAF 2-1, with comparisons with several representative TM retrieval algorithms, including the pioneering prVBEM and the recent GGS 2-1, which outperforms previous ones. Note that all algorithms involved are adjusted to function at their best status, and the comparisons among them are under the same conditions as explained below. Unless otherwise specified, all the following simulations are conducted using a computer with an Intel Xeon CPU (3.50 GHz, 6 cores) and 64 GB RAM in the environment of MATLAB 2022a. For each algorithm, the performance is evaluated by the efficiency of focusing with the retrieved TM, which is indicated by the peak-to-background ratio (PBR). This is performed by taking the conjugate of one TM row to construct the phase mask for focusing light onto the corresponding position. The TM is modeled using the speckle field produced by random phase mask with zero padding in the Fourier domain, whose elements obeyed a circular Gaussian distribution. According to Ref. 3, the theoretical PBR of focusing is linearly proportional to the input size  $N$ , as given by

$$\eta = \frac{\pi}{4} (N - 1) + 1. \quad (10)$$

The focusing efficiency that a certain TM retrieval algorithm can achieve is typically determined by the iterations (or the running time) it has taken and the sampling ratio ( $\gamma = P/N$ ).

Figure 1(a) shows the schematic diagram of TM retrieval for an MMF. We first examined the focusing performance of



**Fig. 1** Theoretical comparisons of TM retrieval performances of prVBEM, GGS 2-1, and our RAF 2-1. (a) Schematic diagram of TM retrieval for an MMF. (b) Focusing efficiency achieved by different algorithms under a range of running times when  $N = 576$  and  $\gamma = 4$ . (c) The iterations taken by different algorithms versus running times for the case of (b). (d) Focusing efficiency achieved by different algorithms under a range of  $\gamma$  when using  $N = 576$  and a running time of 20 s. Note the error bars denote the standard deviations of 30 repeated tests.

different TM retrieval algorithms under a range of running times when the sampling ratio was fixed to be  $\gamma = 4$ . The input phase mask had a size of  $24 \times 24$ . A total of  $20 \times 20$  foci were produced sequentially that corresponded to 400 rows of TM to be retrieved, with the mean PBRs of the foci as the focusing PBR. The focusing PBR is further normalized by the theoretical value to obtain the focusing efficiency. Figure 1(b) shows the average focusing efficiency achieved by different methods with a running time ranging from 1 to 60 s based on 30 repeated tests. It can be seen RAF 2-1 reaches the optimum efficiency after running for  $\sim 8$  s, while GGS 2-1 requires quite longer running time ( $\sim 40$  s) and prVBEM cannot fully approach the optimum within 60 s. This indicates our nonconvex method is superior in algorithm convergence, given the same condition. Figure 1(c) additionally shows the number of iterations versus running times, which reveals the seconds per iteration for prVBEM, GGS 2-1, and RAF 2-1 are roughly 5:1:1 in such a case. Hence, the nonconvex approach is at least as highly efficient as GGS 2-1 in computation time, and both are much better than prVBEM.

The influence of sampling ratio was also explored for TM retrieval algorithms, by fixing the running time to be 20 s when  $N = 576$ . As shown in Fig. 1(d), the focusing efficiency is close to 0 for all algorithms when  $\gamma = 1$ , which increases dramatically starting from  $\gamma = 2$  for RAF 2-1 and GGS 2-1. Higher focusing efficiency can be achieved with a larger  $\gamma$  for a TM retrieval algorithm, since more measurements taken per degree of freedom in the signal allows for more accurate recovery. Obviously, our RAF 2-1 outperforms its competing peers as it averagely

realizes more than 95% efficiency when  $\gamma = 3$  and 100% when  $\gamma = 4$ . By comparison, GGS 2-1 requires a sampling ratio of 5, while prVBEM requires 7 for the optimal performance under the same conditions.

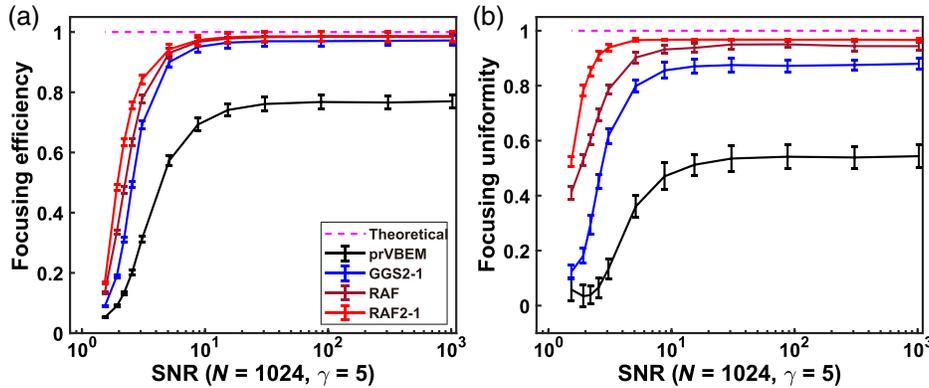
Since data acquisition is inevitably contaminated with noise (mostly signal-dependent) in practice, a good noise robustness is highly preferred for a TM retrieval algorithm. Thus, the algorithm performances were also evaluated under a range of signal-to-noise (SNR) levels using simulated noisy measurements. In the simulation, a multiplicative noise is added to the output field intensity  $\mathbf{I} \in \mathbb{R}_+^P$ . The SNR is defined as

$$\text{SNR} = \langle \mathbf{I} \rangle / \sqrt{\langle (\varepsilon - \langle \varepsilon \rangle)^2 \rangle}, \quad (11)$$

where  $\varepsilon = \mathbf{I}_{\text{noise}} - \mathbf{I}$ , which denotes the noise vector,  $\mathbf{I}_{\text{noise}}$  is the noisy measurement vector, and  $\langle \cdot \rangle$  takes the average for the elements inside. For the focusing results of multiple foci, uniformity is an important metric to measure the focus quality. The focusing uniformity ( $\mu$ ) is given as

$$u = 1 - \sqrt{\langle (I_k - \langle I_k \rangle)^2 \rangle / \langle I_k \rangle}, \quad k \in K, \quad (12)$$

where  $K$  is a set of the indices of foci. With parameter settings that  $N = 1024$ ,  $\gamma = 5$ , and a running time of 50 s, the results of focusing efficiency and uniformity of 400 foci produced using various TM retrieval algorithms are shown in Figs. 2(a) and



**Fig. 2** Simulated (a) focusing efficiency and (b) focusing uniformity achieved by prVBEM, GGS 2-1, RAF, and RAF 2-1 under different SNR levels when using  $N = 1024$ ,  $\gamma = 5$ , and a running time of 50 s. Note the error bars denote the standard deviations of 30 repeated tests.

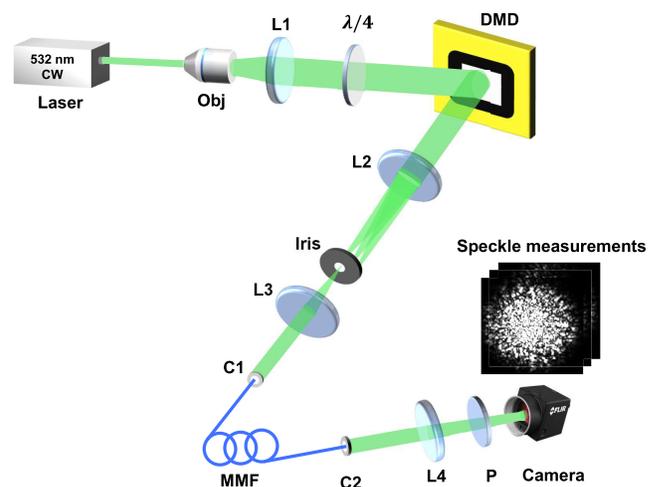
2(b). Note the original RAF was also included to highlight the improved antinoise capability by our modification. It is found that a maximum improvement of 15.5% and 32.4% in terms of focusing efficiency can be realized by RAF 2-1 over RAF and GGS 2-1, respectively, when the SNR is relatively low (e.g., no more than 3.1). Such a performance advantage of RAF 2-1 over other algorithms weakens as the SNR increases, and the focusing efficiency is almost the same when the SNR is about 15. This explains why RAF was excluded in the previous noiseless tests. Besides, an obvious improvement of focusing uniformity is achieved by RAF 2-1, which is at best 25.2% and 60.1% higher than those of RAF and GGS 2-1, respectively, when the SNR is 1.92. The advantage of RAF 2-1 over RAF becomes minor when the SNR is large, while it still sees  $\sim 9\%$  improvement than that of GGS 2-1. Overall, algorithm performance in both focusing efficiency and uniformity follows the order: RAF 2-1 > RAF > GGS 2-1 > prVBEM. The difference between GGS 2-1 and RAF is relatively small, whereas prVBEM falls behind GGS 2-1 considerably.

### 3.2 Experiment

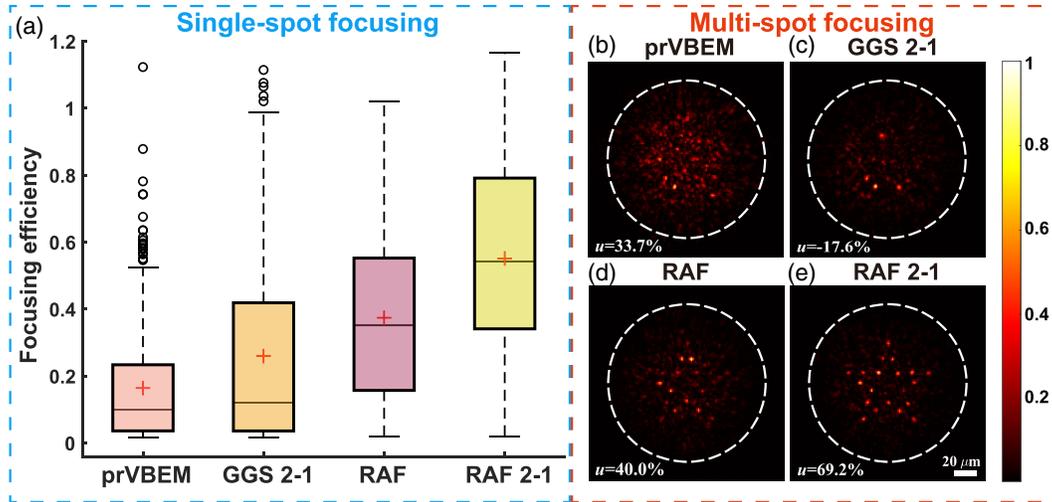
The experimental configuration for retrieving the TM of an MMF is shown in Fig. 3. A beam from a 532 nm continuous-wave laser (EXLSR-532-300-CDRH, Spectra Physics) was expanded by a  $40\times$  objective lens and collimated by a lens (L1). The linearly polarized beam was then modulated into circular polarization by a quarter-wave plate ( $\lambda/4$ ) before impinging onto a digital micromirror device (DMD, DLP9500, Texas Instruments Inc.). Based on the Lee hologram technique, the DMD, combined with a  $4f$  system composed of L2, iris, and L3, allowed for phase modulation at a high-speed pattern rate of up to 23 kHz, rendering fast data acquisition for TM calibration. The phase-encoded and shrunk beam was subsequently coupled into an MMF of 0.22 numerical aperture (NA) and diameter of  $105\ \mu\text{m}$  (SUH105, Xinrui, China) by a collimator (C1). The transmitted light was also imaged with a collimator (C2). The beam then passed through a lens (L4) for convergence and a polarizer before being captured by a CMOS camera (BFS-U3-04S2M, FLIR). For TM retrieval, there was a sequence of speckle intensity measurements under the input modulations of random phase.

In the experiment, the performances of using different TM retrieval algorithms to control light delivery despite scattering

were compared from the aspects of single-spot and multispot focusing through MMF. For single-spot focusing, a total of  $20 \times 20$  foci were generated sequentially on the working plane of the MMF, which meant 400 rows of TM were to be retrieved. The sampling ratio was set to be 5 for all algorithms to ensure the quality of TM retrieval, given that the acquired speckle data suffered from noises such as shot noise, dark current noise, and readout noise. Figure 4(a) presents the histogram distribution of the normalized PBRs of 400 foci achieved with different algorithms. Since the experimentally acquired TM of the MMF also obeyed the circular Gaussian distribution, it could be reasonable to use Eq. (10) for normalizing the focus PBRs generated by the MMF and calculating the focusing efficiency. It can be seen that the average focusing efficiency (denoted by the crosses in the box plots) are 16.45%, 26.01%, 37.46%, and 55.17% for prVBEM, GGS 2-1, RAF, and RAF 2-1, respectively. These results correspond well to the simulated ones with the same parameter settings and under low SNR [see Fig. 2(a)]. Also note for GGS 2-1, the median of the focus PBRs is much lower than



**Fig. 3** Experimental configuration for retrieving the TM of an MMF with speckle-intensity measurements. CW: continuous wave; C, collimator; DMD, digital mirror device; L, lens; P, polarizer; MMF, multimode fiber; Obj, objective lens;  $\lambda/4$ , quarter-wave plate.



**Fig. 4** Comparison of light-focusing results through MMF with the TMs retrieved by different algorithms. (a) The histograms of normalized PBR of  $20 \times 20$  foci and (b) the results of multispot focusing (pentagram) in the output field of MMF, both obtained by prVBEM, GGS 2-1, RAF, and RAF 2-1 with  $N = 1024$  and  $\gamma = 5$ . Note the crosses in (a) represent the mean values, and the white dashed circles in (b) show the fiber region. The scale bar in (b)–(e) is  $20 \mu\text{m}$ .

the mean because of the poor focusing uniformity. According to the box plots of Fig. 4(a), quite a few foci approached or even surpassed the theoretical PBR for GGS 2-1, RAF, and especially RAF 2-1. However, their overall focusing efficiency of 400 different foci on the working plane of MMF still saw a distance from the ideal level, using the retrieved TMs when  $\gamma = 5$ . The fact that RAF 2-1 achieved a considerably higher efficiency than RAF and other algorithms confirms the noisy conditions in a real environment, which may originate from many sources, such as camera noise, imperfect fiber coupling, and other aberrations in the optical path.

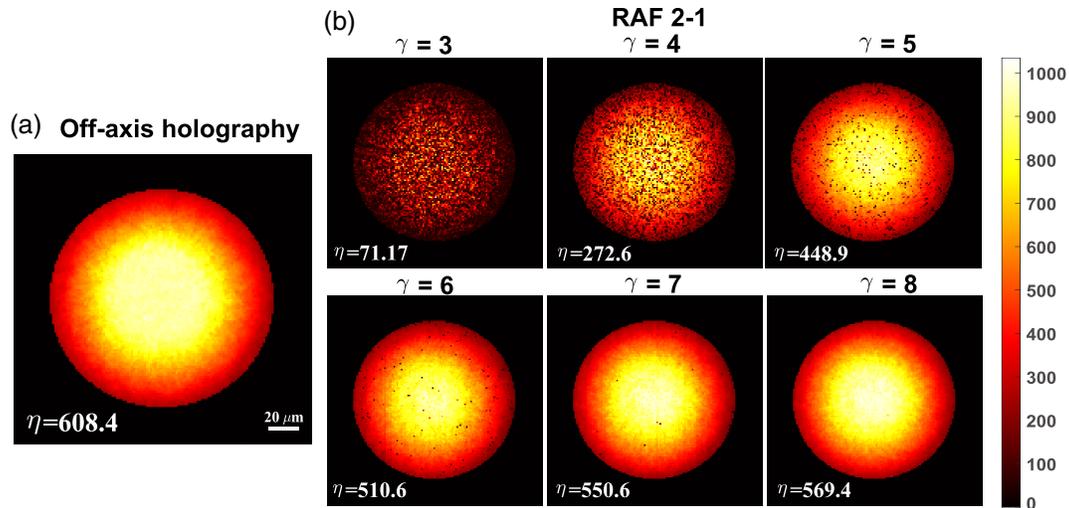
In addition to single-spot focusing, a multispot focusing experiment was also conducted under the same conditions. This was achieved by superposing multiple phase-conjugate rows of the retrieved TM to construct a phase mask. The results of light focusing onto a pentagram pattern composed of 20 spots by different algorithms are shown in Fig. 4(b). The focusing uniformity were 33.7%,  $-17.6\%$ , 40.0%, and 69.2%, respectively, for the four algorithms. It can be observed that only RAF 2-1 produced a high-quality pentagram pattern by focusing light on all the preselected positions, due to its superior performance of TM retrieval.

The accuracy of TM retrieval by our RAF 2-1 was further compared with the off-axis holography, the gold standard for the measurement of TM. To do so, we scanned the whole fiber region on the working plane of the MMF, so that a map of the focusing PBR could be synthesized, which was used to fully evaluate the accuracy of TM. The fiber region was determined by identifying the largest connected region in the binarized image taken when all pixels on the DMD were turned on. Using circular fitting of the fiber region, the center and radius of the fiber region were obtained, which were then used to create a binary mask of the fiber region. In the experiment, there were 8685 pixels inside the circular fiber region of the  $140 \times 140$  output field, which correspond to 8685 rows of TM to be retrieved. Figure 5 summarizes the results of focusing PBR maps with the TM measured by off-axis holography and the TMs retrieved by

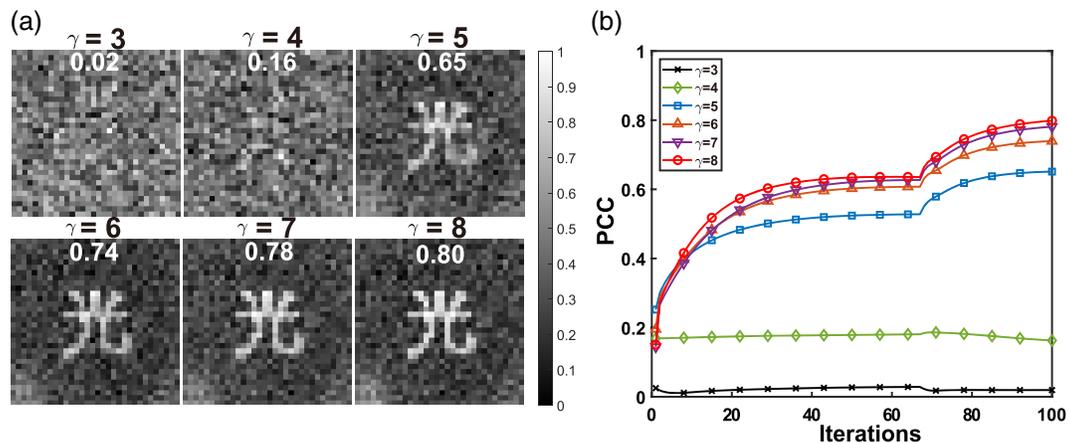
RAF 2-1 under a range of  $\gamma$  from 3 to 8. The mean PBR by the gold standard method is 608.4. Compared with the theoretical PBR (i.e.,  $\eta = 804$ ), it is reasonable, given that the focusing quality degraded in the fiber fringe area due to the inhomogeneous mode excitation inside the MMF. As for RAF 2-1, there are many dark spots in the PBR map synthesized under small  $\gamma$ , indicating poor accuracy of the corresponding rows of TM being retrieved. With a larger  $\gamma$ , the PBR map becomes more homogeneous with fewer dark spots, which means an overall improvement of the TM accuracy. Notably, when  $\gamma = 8$ , the PBR map by RAF 2-1 is comparable to that of the holography, with a mean PBR of 569.4 reaching  $\sim 93.6\%$  efficiency of the gold standard experimentally. In practice, the choice of the sampling ratio should strike a balance between the computing costs and the expected focusing performance with the recovered TM. Compared to off-axis holography, the key advantage of RAF 2-1 is the elimination of reference beam and interferometry, which is compatible with more applications. In addition, with parallel operation and GPU implementation, the TM retrieval process by RAF 2-1 was fast enough. For example, under  $\gamma = 8$ , retrieving an  $8685 \times 1024$  TM by RAF 2-1 took 42.3 s on average when using a computer with an Intel Xeon CPU E5-1650 v3 @3.50 GHz, an NVIDIA RTX2080 Ti GPU, and 128 GB RAM.

Using the retrieved TM by RAF 2-1, one can further reconstruct an input object from intensity-only speckle measurement with one more phase retrieval. The reconstruction result relies heavily on the quality of the recovered TM, which acts as the measurement matrix. Figure 6(a) shows the results of reconstructing a  $32 \times 32$  phase object of the Chinese character “光” (meaning “light”), by taking 100 iterations with the TM of MMF retrieved by RAF 2-1 when  $\gamma$  was increased from 3 to 8. The Pearson correlation coefficient (PCC) is used to quantify the fidelity between the reconstructed image  $\mathbf{I}_R$  and the ground truth  $\mathbf{I}_G$ , which is given as

$$\text{PCC} = \frac{\langle (\mathbf{I}_R - \langle \mathbf{I}_R \rangle) (\mathbf{I}_G - \langle \mathbf{I}_G \rangle) \rangle}{\langle (\mathbf{I}_R - \langle \mathbf{I}_R \rangle)^2 \rangle \langle (\mathbf{I}_G - \langle \mathbf{I}_G \rangle)^2 \rangle}. \quad (13)$$



**Fig. 5** Comparison of the PBR maps of focusing on the working plane of the MMF using the TMs (a) measured by off-axis holography and (b) retrieved by RAF 2-1 under a range of  $\gamma$  with  $N = 1024$ . The scale bar in (a) and (b) is  $20 \mu\text{m}$ .



**Fig. 6** Comparison of image transmission results through MMF using the retrieved TMs by RAF 2-1 under a range of  $\gamma$  with  $N = 1024$ . (a) Normalized reconstructed images for an object of the Chinese character “光” with the values of PCC to the object labeled. (b) The progression curves of PCC during the iterative reconstruction.

The curves of the PCC under different cases of  $\gamma$  are also provided in Fig. 6(b). The upsurges of PCCs at the 67th iteration confirm that the signal recovery is significantly boosted after the gradient heating in the first two-thirds of iterations. The final PCCs are: 0.02, 0.16, 0.65, 0.74, 0.78, and 0.80 for  $\gamma = 3, 4, 5, 6, 7,$  and  $8$ , respectively. As can be observed, the reconstructed image could be recognized starting from  $\gamma = 5$  and attains the best quality when  $\gamma = 8$ . Nonetheless, there is still speckle background noise among the images [Fig. 6(a)], which is common in the reported empirical speckle-imaging results with the TM method,<sup>14,38</sup> mainly because of the imperfect fidelity of the recovered TM. The reconstruction quality can be further improved with a larger  $\gamma$  to overcome the influence of noise. To summarize, image transmission through the MMF is demonstrated with the proposed nonconvex approach, which further validates the effectiveness of the TM retrieved.

## 4 Discussion and Conclusion

There have been various phase-retrieval algorithms used for solving the TM retrieval problem, as introduced earlier. RAF, as one of the best in the nonconvex family, has been reported previously<sup>54</sup> to be highly competitive for image restoration from speckle measurement. To the best of our knowledge, we first adopted it for nonholographic calibration of TM.<sup>42</sup> More importantly, our modified version, RAF 2-1, with an additional step of gradient heating, has shown remarkable improvement in the robustness against noise and the TM retrieval accuracy in both simulations and experiments. The numerical evaluation of RAF and RAF 2-1 can be further seen in Appendix B. Besides the above modification, we resort to speeding up the convergence of RAF for TM retrieval. Efforts include employing adaptive step size in the gradient-descent process or other gradient

acceleration schemes, such as limited memory-BFGS (L-BFGS)<sup>55</sup> and nonlinear conjugate gradient (NCG)<sup>56</sup> methods. However, the improvements are not very impressive, with details referred to in Appendix C.

There are also several limitations to the study. In the experimental setup, the MMF output field was relayed by a collimator instead of an objective lens. Consequently, the working plane of the MMF was immovable, which had a certain distance (about tens of micrometers) away from the fiber end. That said, the setup was sufficient for retrieving a reliable TM and focusing on the working plane for demonstration. An objective lens is needed only for measuring the TMs corresponding to different working planes. In addition, since there is phase ambiguity for the formulated LSE objective function in Eq. (4), a phase offset exists for each row of the retrieved TM. However, it does not affect the intensity of the generated two-dimensional foci. Further phase correction<sup>28</sup> is indispensable when it comes to MMF three-dimensional volumetric focusing and imaging.

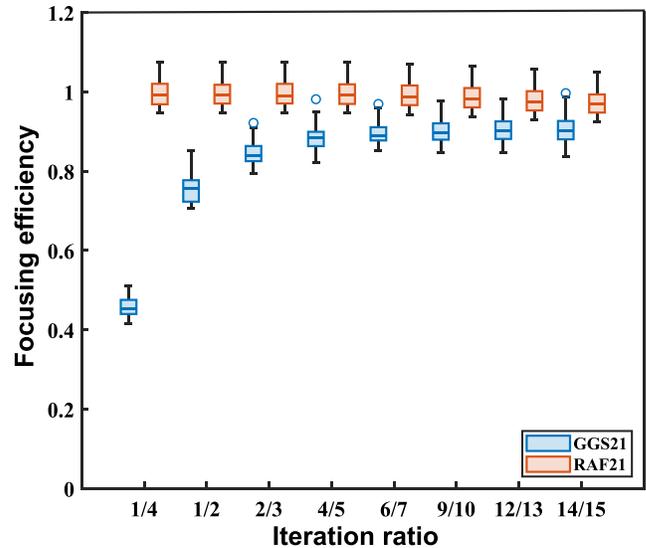
In conclusion, we have proposed a modified nonconvex approach, RAF 2-1, for retrieving the TM of MMF based on speckle-intensity measurements. Theoretically, RAF 2-1 can achieve optimum focusing efficiency with less running time or sampling ratio than the previously reported TM retrieval methods. The experimental results of light control through an MMF confirm a comparable performance of RAF 2-1 to the gold-standard holography method for TM measurement. RAF 2-1 is also computationally efficient, taking 42.3 s on average to recover a  $8685 \times 1024$  TM ( $\gamma = 8$ ) on a regular computer under parallel operation and GPU implementation. Endowed with the advantages of optimum efficiency, fast execution, and a reference-less setup, RAF 2-1 allows for broad applications in MMF-based imaging, manipulation, treatment, etc.

## 5 Appendix A: Best Iteration Ratio for the Two-Step Gradient Iteration Process of RAF 2-1

There are two steps in the gradient iteration process regarding the proposed RAF 2-1. In order to determine the number of iterations in steps 1 and 2 (with the total number fixed) for the best performance, numerical experiments were conducted to compare the focusing efficiency achieved by RAF 2-1 under a series of iteration ratios. Moreover, since GGS 2-1 also involved the two-step gradient descent, it inspires this work and is used for performance comparison. Therefore, the best ratio of iteration of GGS should also be determined. Figure 7 gives the results of both RAF 2-1 and GGS 2-1, with a total iteration of 300. It can be observed for RAF 2-1, there are minor differences of focusing efficiency among different iteration ratios, while the one at two-thirds is chosen as the best iteration ratio due to a slight advantage. As for GGS 2-1, the best focusing efficiency is around an iteration ratio of 9/10, which is also consistent with the original research that adopted 287 and 34 iterations in steps 1 and 2 for GGS 2-1, respectively.

## 6 Appendix B: Numerical Evaluation of RAF and RAF 2-1

As mentioned in Sec. 2, the modified version, RAF 2-1, is more effective for TM retrieval. To evaluate how the performance of

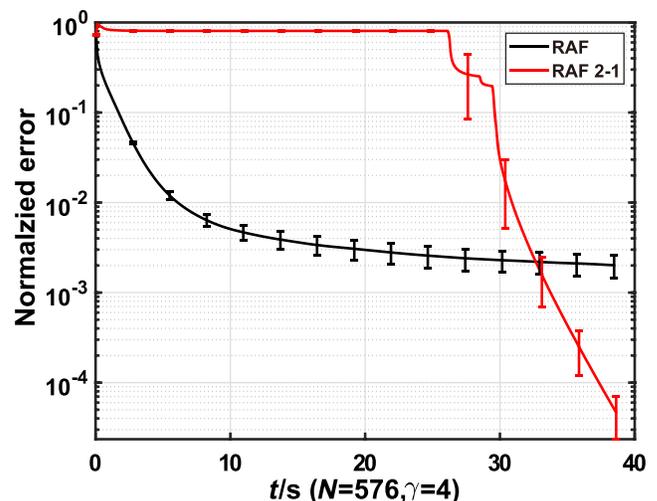


**Fig. 7** Focusing efficiency achieved by GGS 2-1 and RAF 2-1 under a series of iteration ratios during their two-step gradient iterations, with 30 repeated tests.

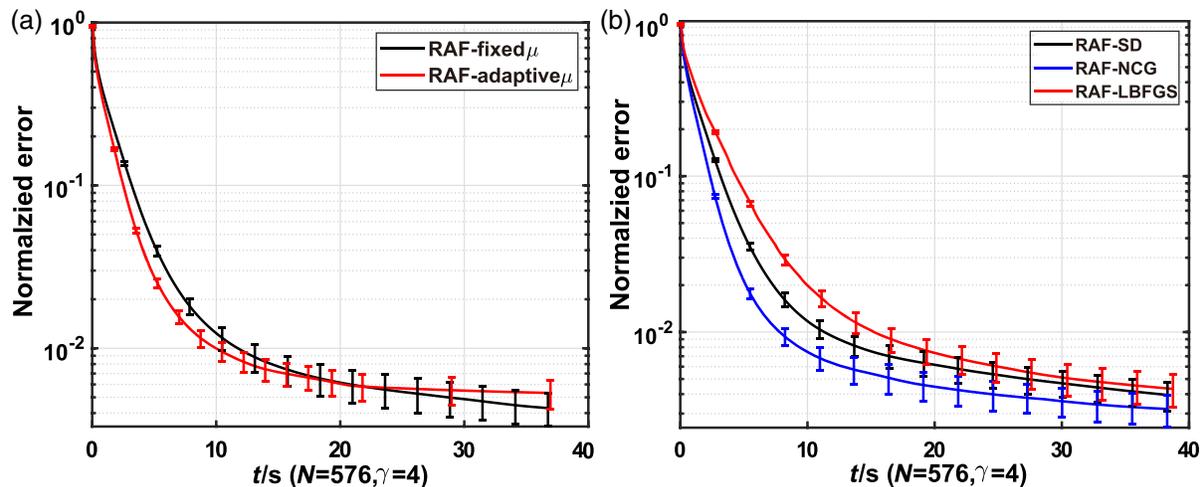
RAF 2-1 is better than the original RAF, a numerical experiment in a noiseless condition was performed for retrieving the TM that corresponds to 400 output modes. The curves of the averaged errors after normalization are presented in Fig. 8. The measurement error for the  $i$ 'th row of TM is defined as

$$\text{error}_i = \|\mathbf{X}^H \hat{\mathbf{a}}_i - \mathbf{y}_i\|_2^2, \quad i = 1, \dots, 400, \quad (14)$$

where  $\hat{\mathbf{a}}_i$  is the estimate of the  $i$ 'th row of TM and other notations are with the same meaning as in Sec. 2. It can be observed that the error of RAF 2-1 can finally decline to a level of as low as  $10^{-4}$ , much lower than that of RAF. This indicates a more accurate result of TM retrieval by RAF 2-1.



**Fig. 8** Normalized curves of error as a function of running time for RAF and RAF 2-1 when  $N = 576$  and  $\gamma = 4$ . Note the error bars denote the standard deviations of 30 repeated tests.



**Fig. 9** (a) Normalized curves of error as a function of running time for RAF with a fixed step size ( $\mu$ ) or an adaptive one. (b) Normalized curves of error as a function of running time for RAF with SD, NCG, or L-BFGS. Note the error bars denote the standard deviations of 30 repeated tests.

## 7 Appendix C: Accelerated Gradient Descent for RAF

As discussed earlier, the ways to accelerate the convergence of RAF were also studied using an adaptive step size and a more advanced gradient-descent solver. First, we adopted the Barzilai–Borwein method for nonmonotonic backtracking line search of step size, which was compared with the fixed one ( $\mu = 3$ ). As shown in Fig. 9(a), the measurement error of using adaptive  $\mu$  drops slightly more rapidly than that of fixed  $\mu$  within the first 20 s of running time, while the latter eventually declines to a lower level. This shows the adaptive step size method is not very effective, although it could be better with parameter fine-tuning. As for the gradient-descent solver, besides the regular steep descent (SD) using the negative first derivative (i.e., the gradient) as the descent direction, acceleration methods, such as NCG and L-BFGS, were employed for comparison. Since NCG and especially L-BFGS entail more computations per iteration than SD, for fair comparison, the curves of error as a function of running time (instead of iterations) for different optimization methods were compared, as shown in Fig. 9(b). We can see that NCG has the fastest convergence with the same running time. The reason that L-BFGS method is even slower in the declining trend of error is attributed to the far more seconds per iteration it requires. In fact, the average number of iterations taken by SD, NCG, and L-BFGS are 671, 656, and 411, respectively. The convergence for L-BFGS could be the fastest if compared from the perspective of error decline versus iterations.

### Disclosures

The authors declare that there are no conflicts of interest for this work.

### Code and Data Availability

The code demo of RAF 2-1 is available in [https://github.com/Ford666/RAF21\\_TM\\_retrieval](https://github.com/Ford666/RAF21_TM_retrieval), and the data that support the findings of this study can be available by reasonable request to the corresponding author.

### Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) (Grant No. 81930048), the Hong Kong Innovation and Technology Commission (Grant Nos. GHP/043/19SZ and GHP/044/19GD), the Hong Kong Research Grant Council (Grant Nos. 15217721, R5029-19 and C7074-21GF), the Guangdong Science and Technology Commission (Grant No. 2019BT02X105), the Shenzhen Science and Technology Innovation Commission (Grant No. JCYJ20220818100202005), and the Hong Kong Polytechnic University (Grant Nos. P0038180, P0039517, P0043485 and P0045762). The authors would also like to thank the Photonics Research Institute of the Hong Kong Polytechnic University for facility support.

### References

1. J.-H. Park et al., “Disordered optics: exploiting multiple light scattering and wavefront shaping for nonconventional optical elements,” *Adv. Mater.* **32**(35), 1903457 (2020).
2. S. Gigan, “Imaging and computing with disorder,” *Nat. Phys.* **18**(9), 980–985 (2022).
3. I. M. Vellekoop and A. Mosk, “Focusing coherent light through opaque strongly scattering media,” *Opt. Lett.* **32**(16), 2309–2311 (2007).
4. S. Rotter and S. Gigan, “Light fields in complex media: mesoscopic scattering meets wave control,” *Rev. Mod. Phys.* **89**(1), 015005 (2017).
5. S. Gigan et al., “Roadmap on wavefront shaping and deep imaging in complex media,” *J. Phys. Photonics* **4**(4), 042501 (2022).
6. Z. Yu et al., “Wavefront shaping: a versatile tool to conquer multiple scattering in multidisciplinary fields,” *Innovation* **3**(5), 100292 (2022).
7. S. Cheng et al., “Long-distance pattern projection through an unfixed multimode fiber with natural evolution strategy-based wavefront shaping,” *Opt. Express* **30**(18), 32565–32576 (2022).
8. C. M. Woo et al., “Optimal efficiency of focusing diffused light through scattering media with iterative wavefront shaping,” *APL Photonics* **7**(4), 046109 (2022).
9. S. M. Popoff et al., “Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media,” *Phys. Rev. Lett.* **104**(10), 100601 (2010).

10. S. Popoff et al., "Controlling light through optical disordered media: transmission matrix approach," *New J. Phys.* **13**(12), 123021 (2011).
11. M. Kim et al., "Transmission matrix of a scattering medium and its applications in biophotonics," *Opt. Express* **23**(10), 12648–12668 (2015).
12. 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**(2), 1733–1740 (2012).
13. A. Drémeau et al., "Reference-less measurement of the transmission matrix of a highly scattering material using a DMD and phase retrieval techniques," *Opt. Express* **23**(9), 11898–11911 (2015).
14. S. Popoff et al., "Image transmission through an opaque material," *Nat. Commun.* **2**(1), 81 (2010).
15. H. Li et al., "Learning-based super-resolution interpolation for sub-Nyquist sampled laser speckles," *Photonics Res.* **11**(4), 631–642 (2023).
16. T. Čížmár and K. Dholakia, "Shaping the light transmission through a multimode optical fibre: complex transformation analysis and applications in biophotonics," *Opt. Express* **19**(20), 18871–18884 (2011).
17. M. Plöschner and T. Čížmár, "Compact multimode fiber beam-shaping system based on gpu accelerated digital holography," *Opt. Lett.* **40**(2), 197–200 (2015).
18. H. Zhang et al., "Large-scale, high-contrast glare suppression with low-transmittance eigenchannels of aperture-target transmission matrices," *Opt. Lett.* **46**(7), 1498–1501 (2021).
19. S. Cheng et al., "Alternating projection-based phase optimization for arbitrary glare suppression through multimode fiber," *Opt. Lasers Eng.* **161**, 107368 (2023).
20. T. Čížmár and K. Dholakia, "Exploiting multimode waveguides for pure fibre-based imaging," *Nat. Commun.* **3**(1), 1027 (2012).
21. S. Turtaev et al., "High-fidelity multimode fibre-based endoscopy for deep brain *in vivo* imaging," *Light Sci. Appl.* **7**(1), 92 (2018).
22. D. Stellinga et al., "Time-of-flight 3D imaging through multimode optical fibers," *Science* **374**(6573), 1395–1399 (2021).
23. Z. Wen et al., "Single multimode fibre for *in vivo* light-field-encoded endoscopic imaging," *Nat. Photonics* **17**, 679–687 (2023).
24. I. T. Leite et al., "Three-dimensional holographic optical manipulation through a high-numerical-aperture soft-glass multimode fibre," *Nat. Photonics* **12**(1), 33–39 (2018).
25. T. Zhong et al., "Optically selective neuron stimulation with a wavefront shaping-empowered multimode fiber," *Adv. Photonics Res.* **3**(3), 2100231 (2022).
26. A. Liutkus et al., "Imaging with nature: compressive imaging using a multiply scattering medium," *Sci. Rep.* **4**(1), 5552 (2014).
27. L. Gong et al., "Optical orbital-angular-momentum-multiplexed data transmission under high scattering," *Light Sci. Appl.* **8**(1), 27 (2019).
28. J. Zhong et al., "Efficient reference-less transmission matrix retrieval for a multimode fiber using fast Fourier transform," *Adv. Photonics Nexus* **2**(5), 056007 (2023).
29. Y. Choi et al., "Overcoming the diffraction limit using multiple light scattering in a highly disordered medium," *Phys. Rev. Lett.* **107**(2), 023902 (2011).
30. Y. Choi et al., "Scanner-free and wide-field endoscopic imaging by using a single multimode optical fiber," *Phys. Rev. Lett.* **109**(20), 203901 (2012).
31. H. Yu, K. Lee, and Y. Park, "Ultrahigh enhancement of light focusing through disordered media controlled by mega-pixel modes," *Opt. Express* **25**(7), 8036–8047 (2017).
32. T. Zhao et al., "Seeing through multimode fibers with real-valued intensity transmission matrices," *Opt. Express* **28**(14), 20978–20991 (2020).
33. X. Tao et al., "High-speed scanning interferometric focusing by fast measurement of binary transmission matrix for channel demixing," *Opt. Express* **23**(11), 14168–14187 (2015).
34. I. Yamaguchi and T. Zhang, "Phase-shifting digital holography," *Opt. Lett.* **22**, 1268–1270 (1997).
35. E. Cuche, P. Marquet, and C. Depeursinge, "Spatial filtering for zero-order and twin-image elimination in digital off-axis holography," *Appl. Opt.* **39**(23), 4070–4075 (2000).
36. P. Jákł et al., "Endoscopic imaging using a multimode optical fibre calibrated with multiple internal references," *Photonics* **9**(1), 37 (2022).
37. B. Rajaei et al., "Intensity-only optical compressive imaging using a multiply scattering material and a double phase retrieval approach," in *IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, IEEE, pp. 4054–4058 (2016).
38. C. A. Metzler et al., "Coherent inverse scattering via transmission matrices: efficient phase retrieval algorithms and a public dataset," in *IEEE Int. Conf. Comput. Photogr. (ICCP)*, IEEE, pp. 1–16 (2017).
39. M. N'Gom et al., "Mode control in a multimode fiber through acquiring its transmission matrix from a reference-less optical system," *Opt. Lett.* **43**(3), 419–422 (2018).
40. G. Huang et al., "Retrieving the optical transmission matrix of a multimode fiber using the extended Kalman filter," *Opt. Express* **28**(7), 9487–9500 (2020).
41. G. Huang et al., "Generalizing the Gerchberg–Saxton algorithm for retrieving complex optical transmission matrices," *Photonics Res.* **9**(1), 34–42 (2021).
42. S. Cheng, T. Zhong, and P. Lai, "Non-convex optimization for retrieving the complex transmission matrix of a multimode fiber," in *TENCON 2022-2022 IEEE Region 10 Conf. (TENCON)*, IEEE, pp. 1–5 (2022).
43. D. Ancora et al., "Speckle spatial correlations aiding optical transmission matrix retrieval: the smoothed Gerchberg–Saxton single-iteration algorithm," *Photonics Res.* **10**(10), 2349–2358 (2022).
44. S. Marchesini, "Invited article: a unified evaluation of iterative projection algorithms for phase retrieval," *Rev. Sci. Instrum.* **78**(1), 011301 (2007).
45. I. Waldspurger, A. d'Aspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Math. Programming* **149**, 47–81 (2015).
46. P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," *IEEE Trans. Signal Process.* **63**(4), 1043–1055 (2014).
47. S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," *IEEE Trans. Inf. Theory* **65**(10), 6664–6684 (2019).
48. P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Proc. 26th Int. Conf. Neural Inf. Process. Syst.- Vol. 2, NIPS'13*, Curran Associates Inc., Red Hook, New York, pp. 2796–2804 (2013).
49. E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: theory and algorithms," *IEEE Trans. Inf. Theory* **61**(4), 1985–2007 (2015).
50. G. Wang, G. B. Giannakis, and Y. C. Eldar, "Solving systems of random quadratic equations via truncated amplitude flow," *IEEE Trans. Inf. Theory* **64**(2), 773–794 (2017).
51. H. Zhang et al., "A nonconvex approach for phase retrieval: reshaped wirtinger flow and incremental algorithms," *J. Mach. Learn. Res.* **18**, 5164–5198 (2017).
52. G. Wang et al., "Phase retrieval via reweighted amplitude flow," *IEEE Trans. Signal Process.* **66**(11), 2818–2833 (2018).
53. Y. Chen and E. Candes, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," in *Adv. in Neural Inf. Process. Syst.*, C. Cortes et al., Eds., Curran Associates, Inc., Vol. **28** (2015).
54. R. Chandra, T. Goldstein, and C. Studer, "PhasePack: a phase retrieval library," in *13th Int. Conf. Sampling Theory and Appl. (SampTA)*, IEEE, pp. 1–5 (2019).
55. L. Ji and Z. Tie, "On gradient descent algorithm for generalized phase retrieval problem," in *IEEE 13th Int. Conf. Signal Process. (ICSP)*, IEEE, pp. 320–325 (2016).

56. T. Goldstein, C. Studer, and R. Baraniuk, "A field guide to forward-backward splitting with a fast implementation," arXiv:1411.3406 (2014).

**Shengfu Cheng** is a PhD student at the Department of Biomedical Engineering in Hong Kong Polytechnic University. He received his bachelor's degree from Sichuan University. His research interests include wavefront shaping, computational optical imaging, and optical-fiber-based endoscopy.

**Xuyu Zhang** is a PhD student at the Optical Engineering, School of Optoelectronics, University of Shanghai for Science and Technology. He received his bachelor's degree from Harbin Institute of Technology at Weihai. His main research interest is the application of deep learning to the study of scattering imaging mechanism.

**Tianting Zhong** received his bachelor's degree from Nanjing Agricultural University and later his PhD from Hong Kong Polytechnic University. His research interests primarily focus on deep-tissue optical focusing, as well as the use of multimode fiber for endoscopic purposes related to imaging, stimulation, and treatment.

**Huanhao Li** is currently a postdoc fellow in the Department of Biomedical Engineering of the Hong Kong Polytechnic University (PolyU). He obtained his PhD from PolyU in 2021. His current research interests include wavefront shaping and its optimization algorithm, speckle imaging, and speckle-based image processing with deep learning. He has published more than 10 papers in journals including *Advanced Science*, *The Innovation*, *Light: Science & Applications*, and *Photonics Research*.

**Haoran Li** is a PhD student in the Department of Biomedical Engineering at PolyU. He received his bachelor's degree and his master's degree in applied physics and optical engineering from Huaqiao University, in 2018 and 2022, respectively. His research interests include fiber imaging, speckle modulation, and endoscopic imaging.

**Lei Gong** is currently an associate professor at the School of Physics of the University of Science and Technology of China (USTC). He received his bachelor's degree from Anhui Normal University in 2011 and his PhD in optics from USTC in 2016. He has published more than 50 scientific articles. His research interests focus on wavefront shaping, computational imaging, and optical trapping.

**Honglin Liu** received her bachelor's degree from the University of Science and Technology of China in 2004 and her PhD from Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences in 2009. Her research interests focus on light propagation in scattering media, techniques and mechanisms of imaging through dense scattering media, and remote sensing in bad weather. She has published more than 40 papers in premium journals, such as *Nature Photonics*, *PhotonIX*, and *Advanced Science*.

**Puxiang Lai** received his bachelor's degree from Tsinghua University in 2002, his master's degree from the Chinese Academy of Sciences in 2005, and his PhD from Boston University in 2011. His research interests focus on deep-tissue optical focusing, imaging, stimulation, and treatment, which have fueled more than 90 journal publications in premium journals, such as *Nature Photonics*, *Nature Communications*, and *The Innovation*.