# Iterative, backscatter-analysis algorithms for increasing transmission and focusing light through highly scattering random media

**Curtis Jin,\* Raj Rao Nadakuditi, Eric Michielssen, and Stephen C. Rand**

*Department of EECS, University of Michigan, Ann Arbor, Michigan 48109-2122, USA*
*\*Corresponding author: jsirius@umich.edu*

Scattering hinders the passage of light through random media and consequently limits the usefulness of optical techniques for sensing and imaging. Thus, methods for increasing the transmission of light through such random media are of interest. Against this backdrop, recent theoretical and experimental advances have suggested the existence of a few highly transmitting eigen-wavefronts with transmission coefficients close to 1 in strongly backscattering random media. Here, we numerically analyze this phenomenon in 2D with fully spectrally accurate simulators and provide rigorous numerical evidence confirming the existence of these highly transmitting eigen-wavefronts in random media with periodic boundary conditions that are composed of hundreds of thousands of nonabsorbing scatterers. Motivated by bio-imaging applications in which it is not possible to measure the transmitted fields, we develop physically realizable algorithms for increasing the transmission through such random media using backscatter analysis. We show via numerical simulations that the algorithms converge rapidly, yielding a near-optimum wavefront in just a few iterations. We also develop an algorithm that combines the knowledge of these highly transmitting eigen-wavefronts obtained from backscatter analysis with intensity measurements at a point to produce a near-optimal focus with significantly fewer measurements than a method that does not utilize this information.   © 2013 Optical Society of America

*OCIS codes:*    (030.6600) Statistical optics; (290.5825) Scattering theory.
http://dx.doi.org/10.1364/JOSAA.30.001592

## 1. INTRODUCTION

Media such as glass and air are transparent because light propagates through them without being scattered or absorbed. In contrast, materials such as turbid water, white paint, and egg shells are opaque because the randomly arranged particles cause light to scatter in random directions, thereby hindering its passage. As the thickness of a slab of highly scattering random medium increases, this effect becomes more pronounced, and less and less normally incident light is transmitted through [1].

In this context, the theoretical work of Dorokhov [2], Pendry and co-workers [3,4], and others [5,6] provides unexpected insight into how, and the extent to which, the limitations imposed by random scattering may be overcome. Specifically, these authors predict that in highly scattering random media composed of nonabsorbing scatterers, the eigen-wavefronts associated with the right singular vectors of the $S_{21}$ or transmission matrix will have transmission coefficients whose distribution has a bimodal shape as in Fig. 2. Consequently, while many eigen-wavefronts have a small transmission coefficient, a small number of eigen-wavefronts exist that have a transmission coefficient close to 1; i.e., they propagate with almost no scattering loss.

The breakthrough experiments of Vellekoop and Mosk [7,8] provide evidence of the existence of these highly transmitting eigen-wavefronts in random media. Vellekoop and Mosk showed [7] that intensity measurements on the transmission side of a scattering medium could be used to construct a wavefront that produced about 1000× intensity enhancement at a target point over that due to a normally incident wavefront. Their work set off a flurry of research on methods for measuring the transmission matrix and comparing the transmission coefficient distribution with the theoretical prediction [9–12], faster experimental methods for focusing [13–17], and numerical work on the properties of eigen-wavefronts [18].

Our work is inspired by these three lines of inquiry. We develop iterative, physically realizable algorithms for transmission maximization that utilize backscatter analysis to produce a highly transmitting wavefront in just a few iterations. These algorithms build on the initial work presented in [19].

These algorithms, which utilize the information in the backscatter field, can be useful in applications, such as in bio-imaging, where it might not be possible to measure the transmitted fields. Our algorithms yield a highly transmitting wavefront using significantly fewer measurements than required to measure the whole reflection or $S_{11}$ matrix and then generate the wavefront (associated with the smallest right singular vector of the $S_{11}$ matrix) that produces the smallest backscatter (and hence the highest transmission in a lossless medium).

Since our methods maximize transmission by minimizing backscatter, it is important for most of the backscatter field to be captured to fully realize these advantages. Otherwise, given a limited viewing aperture, the principle of backscatter minimization cannot guarantee increased forward

transmission and might even produce "transmission" into the unobserved portion of the backscatter field.

Furthermore, we develop an iterative, physically realizable algorithm for focusing that utilizes intensity measurements at the desired point and backscatter analysis to produce a near-optimal focusing wavefront with significantly fewer measurements than other approaches. Thus the principal advantage of this approach is that one can get 95% of the optimal intensity using significantly fewer measurements than it would take to get the optimal intensity. In effect, we are increasing the rate of convergence to the optimal focusing wavefront. Changing the focusing point or the number of foci does not affect the convergence behavior. We show that we retain this property even when we control fewer than the total number of propagating modes.

A crucial feature of the algorithms we have developed is that they allow the number of modes being controlled via a spatial light modular in experiments to be increased without increasing the number of measurements that have to be made.

An additional advantage conferred by these rapidly converging algorithms is that they might facilitate their use in applications where the duration in which the $S_{21}$ or $S_{11}$ matrix can be assumed to be quasi-static is relatively small compared to the time it would take to make all measurements needed to estimate the $S_{21}$ or $S_{11}$ matrix or in settings where a near-optimal solution obtained quickly is preferable to the optimal solution that takes many more measurements to compute.

Finally, we numerically analyze the phenomenon using a spectrally accurate simulator for two-dimensional (2D) scattering systems with periodic boundary conditions and provide the first numerically rigorous confirmation of the shape of the transmission coefficient distribution and the existence [8] of an eigen-wavefront with a transmission coefficient approaching 1 for random media with a large number of scatterers.

The paper is organized as follows. We describe our setup in Section 2. We discuss the problem of transmission maximization and focusing in Section 3. To assist in the development of physically realizable algorithms for these applications, we identify physically realizable operations in Section 4, and describe iterative, implementable algorithms for finding transmission-maximizing and focusing inputs in Sections 5 and 6, respectively. We highlight the existence of eigen-wavefronts with transmission coefficients approaching 1, the algorithms' performance, and rapid convergence via numerical simulations in Section 7, and summarize our findings in Section 8.

## 2. SETUP

We study scattering from a 2D random slab of thickness $L$ and periodicity $D$; the slab's unit cell occupies the space $0 \leq x < D$ and $0 \leq y < L$ (Fig. 1). The slab contains $N_c$ infinite and $z$-invariant circular cylinders of radius $r$ that are placed randomly within the cell and assumed either perfect electrically conducting (PEC) or dielectric with refractive index $n_d$; care is taken to ensure the cylinders do not overlap. Fields are TM$_z$ polarized: electric fields in the $y < 0$ ($i = 1$) and $y > L$ ($i = 2$) half-spaces are denoted $\underline{e}_i(\rho) = e_i(\rho)\hat{z}$. The field (complex) amplitude $e_i(\rho)$ can be decomposed in terms of $+y$ and $-y$ propagating waves as $e_i(\rho) = e_i^+(\rho) + e_i^-(\rho)$, where

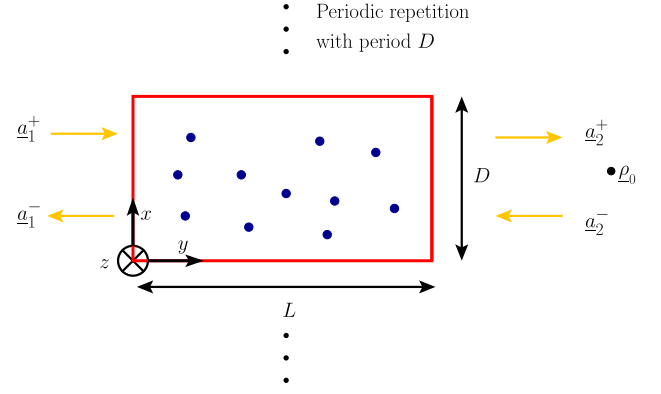$$e_i^{\pm}(\rho) = \sum_{n=-N}^{N} h_n a_{i,n}^{\pm} e^{-j\underline{k}_i^{\pm} \cdot \rho}. \tag{1}$$



Fig. 1.   Geometry of the scattering system considered.

In the above expression, $\rho = x\hat{x} + y\hat{y} \equiv (x, y)$, $\underline{k}_n^{\pm} = k_{n,x}\hat{x} \pm k_{n,y}\hat{y} \equiv (k_{n,x}, \pm k_{n,y})$, $k_{n,x} = 2\pi n/D$, $k_{n,y} = 2\pi\sqrt{(1/\lambda)^2 - (n/D)^2}$, $\lambda$ is the wavelength, and $h_n = \sqrt{\|\underline{k}_n^{\pm}\|_2/k_{n,y}}$ is a power-normalizing coefficient. We assume $N = \lfloor D/\lambda \rfloor$; i.e., we only model propagating waves and denote $M = 2N + 1$. The modal coefficients $a_{i,n}^{\pm}$, $i = 1, 2$; $n = -N, \dots, N$ are related by the scattering matrix

$$\begin{bmatrix} \underline{a}_1^- \\ \underline{a}_2^+ \end{bmatrix} = \underbrace{\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}}_{=:S} \begin{bmatrix} \underline{a}_1^+ \\ \underline{a}_2^- \end{bmatrix}, \tag{2}$$

where $\underline{a}_i^{\pm} = [a_{i,-N}^{\pm} \dots a_{i,0}^{\pm} \dots a_{i,N}^{\pm}]^T$. In what follows, we assume that the slab is only excited from the $y < 0$ half-space; hence, $\underline{a}_2^- = 0$. For a given incident field amplitude $e_1^+(\rho)$, we define transmission and reflection coefficients as

$$\tau(\underline{a}_1^+) := \frac{\|S_{21} \cdot \underline{a}_1^+\|_2^2}{\|\underline{a}_1^+\|_2^2} \tag{3}$$

and

$$\Gamma(\underline{a}_1^+) := \frac{\|S_{11} \cdot \underline{a}_1^+\|_2^2}{\|\underline{a}_1^+\|_2^2}, \tag{4}$$

respectively. We denote the transmission coefficient of a normally incident wavefront by $\tau_{\text{normal}} = \tau([0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0]^T)$; here $T$ denotes transposition.

## 3. PROBLEM FORMULATION

### A. Transmission Maximization

The problem of designing an incident wavefront $\underline{a}_{\text{opt}}$ that maximizes the transmitted power can be stated as

$$\underline{a}_{\text{opt}} = \arg\max_{\underline{a}_1^+} \ \tau(\underline{a}_1^+) = \arg\max_{\underline{a}_1^+} \frac{\|S_{21} \cdot \underline{a}_1^+\|_2^2}{\|\underline{a}_1^+\|_2^2}$$
$$= \arg\max_{\|\underline{a}_1^+\|_2=1} \|S_{21} \cdot \underline{a}_1^+\|_2^2, \tag{5}$$

where $\|\underline{a}_1^+\|_2 = 1$ represents the incident power constraint.

Let $S_{21} = \sum_{i=1}^{M} \sigma_i \underline{u}_i \cdot \underline{v}_i^H$ denote the singular value decomposition (SVD) of $S_{21}$; $\sigma_i$ is the singular value associated with the left and right singular vectors $\underline{u}_i$ and $\underline{v}_i$, respectively. By convention, the singular values are arranged so that $\sigma_1 \geq \ldots \geq \sigma_M$ and $H$ denotes the complex conjugate transpose. A well-known result in matrix analysis [20] states that

$$\underline{a}_{\text{opt}} = \underline{v}_1. \tag{6}$$

When the optimal wavefront $\underline{a}_{\text{opt}}$ is excited, the optimal transmitted power is $\tau_{\text{opt}} := \tau(\underline{a}_{\text{opt}}) = \sigma_1^2$. When the wavefront associated with the $i$th right singular vector $\underline{v}_i$ is transmitted, the transmitted power is $\tau(\underline{v}_i) = \sigma_i^2$, which we refer to as the transmission coefficient of the $i$th eigen-wavefront of $S_{21}$. Analogously, we refer to $\Gamma(\underline{v}_i)$ as the reflection coefficient of the $i$th eigen-wavefront of $S_{21}$.

The theoretical distribution [2–6] of the transmission coefficients for lossless random media has density given by

$$f(\tau) = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^{M} \delta(\tau - \tau(\underline{v}_i))$$
$$= \frac{l}{2L} \frac{1}{\tau \sqrt{1 - \tau}}, \quad \text{for} \quad 4 \exp(-L/2l) \lessapprox \tau \leq 1. \tag{7}$$

In Eq. (7), $l$ is the mean-free path through the medium. Figure 2 shows the theoretical density when $L/l = 3$. From, Eq. (7) we expect $\tau_{\text{opt}} = 1$.

From Eq. (6), it follows that the optimal wavefront can be constructed by measuring the $S_{21}$ matrix and computing its SVD. Techniques for measuring the $S_{21}$ matrix have been developed in recent works by Popoff *et al.* [9] and others [10,11]. Kim *et al.* experimentally measured the $S_{21}$ matrix and demonstrated improved transmission by using the optimal wavefront in Eq. (6) [12].

In the lossless setting, the scattering matrix $S$ in Eq. (2) will be unitary; i.e., $S^H \cdot S = I$, where $I$ is the identity matrix. Consequently, we have that $S_{11}^H \cdot S_{11} + S_{21}^H \cdot S_{21} = I$, and the optimization problem in Eq. (5) can be reformulated as
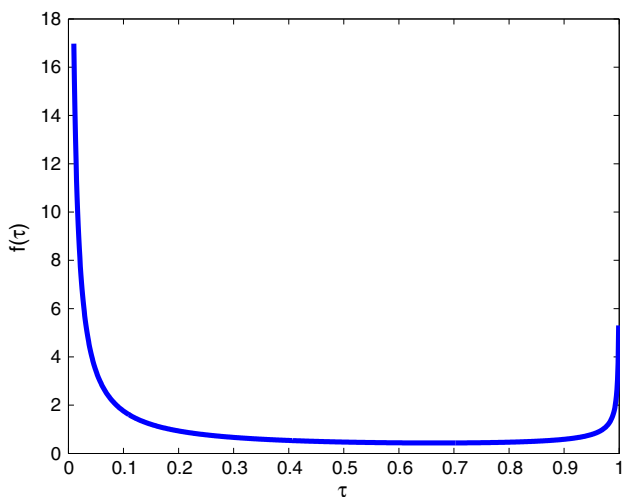


Fig. 2.   Theoretical distribution in Eq. (7) for $L/l = 3$.

$$\underline{a}_{\text{opt}} = \arg\max_{\|\underline{a}_1^+\|_2=1} \underbrace{(\underline{a}_1^+)^H \cdot S_{21}^H \cdot S_{21} \cdot \underline{a}_1^+}_{=(\underline{a}_1^+)^H \cdot (I - S_{11}^H \cdot S_{11}) \cdot \underline{a}_1^+} = \arg\min_{\|\underline{a}_1^+\|_2=1} \|S_{11} \cdot \underline{a}_1^+\|_2^2$$
$$= \arg\min_{\underline{a}_1^+} \Gamma(\underline{a}_1^+). \tag{8}$$

In other words, in a lossless medium the backscatter-minimizing wavefront also maximizes transmission. Let $S_{11} = \sum_{i=1}^{M} \tilde{\sigma}_i \underline{\tilde{u}}_i \cdot \underline{\tilde{v}}_i^H$ denote the SVD of $S_{11}$; $\tilde{\sigma}_i$ is the singular value associated with the left and right singular vectors $\underline{\tilde{u}}_i$ and $\underline{\tilde{v}}_i$, respectively. Then from [20] it follows that

$$\underline{a}_{\text{opt}} = \underline{\tilde{v}}_M. \tag{9}$$

When this optimal wavefront is excited and the medium is lossless, $\tau_{\text{opt}} = 1 - \Gamma(\underline{a}_{\text{opt}}) = 1 - \tilde{\sigma}_M^2 = \sigma_1^2$. When the wavefront associated with the $i$th right singular vector $\underline{\tilde{v}}_i$ is excited, the transmitted power is given by $\tau(\underline{\tilde{v}}_i) = 1 - \Gamma(\underline{\tilde{v}}_i) = 1 - \tilde{\sigma}_i^2$, which we refer to as the transmission coefficient of the $i$th eigen-wavefront of $S_{11}$. Analogously, we refer to $\Gamma(\underline{\tilde{v}}_i)$ as the reflection coefficient of the $i$th eigen-wavefront of $S_{11}$.

A technique for increasing transmission via backscatter analysis would require measurement of the $S_{11}$ matrix and the computation of $\underline{a}_{\text{opt}}$ as in Eq. (9). Our objective is to develop fast, physically realizable, iterative algorithms that converge to $\underline{a}_{\text{opt}}$ by utilizing significantly fewer backscatter field measurements than the $O(M)$ measurements it would take to first estimate $S_{11}$ and then compute its SVD to determine $\underline{\tilde{v}}_M$. Here, we are motivated by applications where it is not possible to measure the transmitted field so that it will not be feasible to measure the $S_{21}$ matrix and compute the optimal wavefront as in Eq. (6).

### B. Focusing
From Eq. (1) and using the fact that $\underline{a}_2^+ = S_{21} \cdot \underline{a}_1^+$ (since $\underline{a}_2^- = 0$), the field at point $\underline{\rho}_0$ is

$$e_2^+(\underline{\rho}_0) = \underbrace{\left[ h_{-N} e^{-j\underline{k}_{-N}^+ \cdot \underline{\rho}_0} \cdots h_N e^{-j\underline{k}_N^+ \cdot \underline{\rho}_0} \right]}_{=:\underline{f}(\underline{\rho}_0)^H} \cdot S_{21} \cdot \underline{a}_1^+. \tag{10}$$

The problem of designing an incident wavefront that maximizes the intensity (or amplitude squared) of the field at $\underline{\rho}_0$ is equivalent to the problem

$$\underline{a}_{\text{foc}} = \arg\max_{\underline{a}_1^+} \frac{\|e_2^+(\underline{\rho}_0)\|_2^2}{\|\underline{a}_1^+\|_2^2} = \arg\max_{\|\underline{a}_1^+\|_2=1} \|\underbrace{\underline{f}^H(\underline{\rho}_0) \cdot S_{21}}_{=:\underline{c}(\underline{\rho}_0)^H} \cdot \underline{a}_1^+\|_2^2, \tag{11}$$

whose solution is

$$\underline{a}_{\text{foc}} = \frac{\underline{c}(\underline{\rho}_0)}{\|\underline{c}(\underline{\rho}_0)\|_2} = \frac{S_{21}^H \cdot \underline{f}(\underline{\rho}_0)}{\|S_{21}^H \cdot \underline{f}(\underline{\rho}_0)\|_2}. \tag{12}$$

Thus the optimal wavefront equals the vector $\underline{c}(\underline{\rho}_0)$ with normalization to satisfy the power constraint. It can be shown that this wavefront may be obtained by time reversing the wavefront received by placing a source at $\underline{\rho}_0$ [21]. This fact was exploited in recent work by Cui and co-workers [22,23].

In Vellekoop and Mosk's breakthrough work [7,8,24], a co-ordinate descent method was employed for constructing the optimal wavefront. The coordinate descent approach finds the amplitude and phase of a single mode that maximize the intensity at $\underline{\rho}_0$ while keeping the amplitudes and phases of the other modes fixed and then repeating this procedure for the remaining modes, one mode at a time. In Vellekoop and Mosk's experiments [7,8,24], they kept the amplitude constant for all modes and considered phase-only modifications of the incident wavefront. While this reduces the complexity of the algorithm, this approach still requires $O(M)$ intensity measurements at $\underline{\rho}_0$ to construct the optimal wavefront. When $M$ is large, the time for convergence will also be large.

This has motivated recent work [15–17] for faster determination of the optimal wavefront. Cui [15,16] considers an approach using multiple frequencies to find the optimal phases of modes simultaneously, while Stockbridge *et al.* [17] have proposed a coordinate descent approach using 2D Walsh functions as a basis set. These methods have accelerated the experimental convergence, but the reported results are still for small $M$ (between 441 and 1024).

Expressing the optimal wavefront in terms of the singular vectors of $S_{21}$ yields the expression

$$\underline{a}_{\text{foc}} \propto S_{21}^H \cdot \underline{f}(\underline{\rho}_0) = \sum_{i=1}^{M} \sigma_i \underbrace{(\underline{u}_i^H \cdot \underline{f}(\underline{\rho}_0))}_{=:w_i} \underline{v}_i = \sum_{i=1}^{M} \sigma_i w_i \underline{v}_i. \quad (13)$$

Recall that $\sigma_i^2 = \tau(\underline{v}_i)$; thus an important insight from Eq. (7) and Fig. 2 is that most of the singular values in Eq. (13) are close to 0. However, there typically are $K \ll M$ singular values close to 1. It is the superposition of these $K$ eigen-wavefronts of $S_{21}$ having transmission coefficients close to 1 whose constructive interference yields the maximal transmission that contributes to maximal intensity.

In the lossless setting, when the scattering matrix $S$ is unitary, we have that $\tau(\underline{v}_i) = 1 - \Gamma(\underline{v}_{M-i+1})$. Hence, the $K$ eigen-wavefronts of $S_{21}$ that have transmission coefficients close to 1 correspond precisely to the $K$ eigen-wavefronts associated with $S_{11}$ that have reflection coefficients close to 0. By using $O(K)$ backscatter field measurements to measure the $K$ eigen-wavefronts of $S_{11}$ with small reflection coefficients and $O(K)$ intensity measurements at $\underline{\rho}_0$, we might expect to approximate $\underline{a}_{\text{foc}}$ in Eq. (13) and yield a near-optimal focus using just $O(K)$ measurements (we expect $K \ll M$).

Our objective is to develop a fast, physically realizable, iterative algorithm that utilizes backscatter field measurements and intensity measurements at $\underline{\rho}_0$ to construct a near-optimal focusing wavefront using significantly fewer measurements than are required by coordinate descent methods that only employ intensity measurements at $\underline{\rho}_0$. The emphasis here is on accelerating the convergence behavior; we do not improve the quality of the focus.

## 4. RECOGNIZING PHYSICALLY REALIZABLE MATRIX–VECTOR OPERATIONS

The iterative algorithms we will develop in Sections 5 and 6 build on the vast literature of iterative methods in numerical linear algebra [25,26]. The algorithms are based on three matrix–vector operations, $S_{11} \cdot \underline{a}_1^+$, $F \cdot (\underline{a}_1^-)^*$, and $S_{11}^H \cdot \underline{a}_1^-$.

These operations can be performed mathematically, but the measurement corresponding to these operations in a physical setting is not obvious. Here, we focus on mapping these matrix–vector operations into their physical counterparts, thus making our algorithms physically realizable.

The first operation, $S_{11} \cdot \underline{a}_1^+$, can be realized by measuring the backscattered wave. In an experimental setting, the modal coefficient vector of the backscattered wave would be extracted from the backscatter intensity measurement by digital holography techniques described in, for example, [27]. We also assume that it is possible to modulate the amplitude and phase of a wavefront, using the methods described in [28]. Thus, the matrix–vector multiplicative operation $S_{11} \cdot \underline{a}_1^+$ corresponds to sending an incident wavefront with modal coefficient vector $\underline{a}_1^+$ and measuring the modal coefficient vector of the backscattered wavefront. Furthermore, we assume that these modal coefficient vectors can be recovered perfectly, and the amplitude and the phase can be perfectly modulated, so that we might investigate the best-case performance of the algorithms.

The second operation, $F \cdot (\underline{a}_1^-)^*$, can be realized by time reversing the wave. Let flipud$(\cdot)$ represent the operation of flipping a vector or a matrix argument upside down so that the first row becomes the last row and so on, and let $*$ denote complex conjugation. We define $F = \text{flipud}(I)$, where $I$ is the identity matrix; then the operation $F \cdot (\underline{a}_1^-)^*$ represents time reversing the wave corresponding to $\underline{a}_1^-$. This can be explained as follows. The expression for a time-reversed wave of $\underline{a}_1^-$ is

$$(e_1^-(\underline{\rho}))^* = \left( \sum_{n=-N}^{N} h_n a_{1,n}^- e^{-j\underline{k}_n^- \cdot \underline{\rho}} \right)^* = \sum_{n=-N}^{N} h_n^* (a_{1,n}^-)^* e^{j\underline{k}_n^- \cdot \underline{\rho}}$$

$$= \sum_{n=-N}^{N} h_n (a_{1,-n}^-)^* e^{-j\underline{k}_n^+ \cdot \underline{\rho}}. \quad (14)$$

Note that we have used the fact that $h_{-n}^* = h_n$ and $\underline{k}_{-n}^- = -\underline{k}_n^+$. From Eq. (14), we see that the modal coefficient vector representation of the time-reversed wave of $\underline{a}_1^-$ is $[(a_N^-)^* \ (a_{N-1}^-)^* \ \dots \ (a_{-N+1}^-)^* \ (a_{-N}^-)^*]^T = F \cdot (\underline{a}_1^-)^*$. Furthermore, we emphasize that the operation $F \cdot (\underline{a}_1^-)^*$ can be physically realized via phase-conjugate mirroring (PCM) [21].

The third operation, $S_{11}^H \cdot \underline{a}_1^-$, can be realized using reciprocity. In a scattering medium that exhibits reciprocity, there are relationships [29–33] between the incident and scattered wavefronts. Consequently, reciprocity requires the reflection matrix $S_{11}$ to satisfy

$$S_{11}^H = F \cdot S_{11}^* \cdot F. \quad (15)$$

This means that if $\underline{a}$ is an input to the system that produces a backscattered wave of $\underline{b}$, then sending $F \cdot (\underline{a})^*$ will produce a backscattered wave of $F \cdot (\underline{b})^*$ in a medium whose reflection matrix corresponds to $S_{11}^H$ (Fig. 3).

An important implication of this equation is that the matrix–vector operation $S_{11}^H \cdot \underline{a}_1^-$ can be cast in terms of physically realizable operations. Note that $S_{11}^H \cdot \underline{a}_1^-$ can be expressed as

$$S_{11}^H \cdot \underline{a}_1^- = F \cdot S_{11}^* \cdot F \cdot \underline{a}_1^- = F \cdot (S_{11} \cdot (F \cdot (\underline{a}_1^-)^*))^*.$$

From the last expression, we see that the operation $S_{11}^H \cdot \underline{a}_1^-$ can be physically realized in a sequence of two steps:
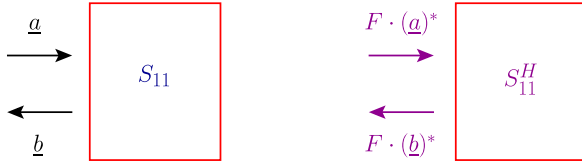
Fig. 3.   Relationship between wavefronts in a medium that exhibits reciprocity. Reciprocity tells us that $S_{11}^H \cdot \underline{a}$ is obtained by time reversing the wave *before and after* sending $\underline{a}$ into the medium, and we call this sequence of operations *double phase conjugation*.

   1.  Time reverse the wavefront whose modal coefficient vector is $\underline{a}_1^-$, and send it to the scattering system.
   2.  Time reverse the resulting backscattered wavefront.

We call this sequence of operations *double phase conjugation*, and we shall leverage it extensively in what follows.

# 5. ITERATIVE, PHYSICALLY REALIZABLE ALGORITHMS FOR TRANSMISSION MAXIMIZATION

We now develop iterative, physically realizable algorithms for transmission maximization that converge to $\underline{a}_{\text{opt}}$ in Eq. (9), by utilizing significantly fewer backscatter field measurements than the $O(M)$ measurements it would take to first estimate $S_{11}$ and then compute its SVD to determine $\tilde{v}_M$.

## A. Steepest Descent Method
The backscatter-minimization problem involves optimization with respect to the objective function $\|S_{11} \cdot \underline{a}_1^+\|_2^2$ that appears on the right-hand side of Eq. (8). The objective function's negative gradient is used as a search direction to correct the previous input as

$$\underline{a}_{1,(k+1)}^+ = \underline{a}_{1,(k)}^+ - \mu \frac{\partial \|S_{11} \cdot \underline{a}_1^+\|_2^2}{\partial \underline{a}_1^+}\bigg|_{\underline{a}_1^+ = \underline{a}_{1,(k)}^+}$$
$$= \underline{a}_{1,(k)}^+ - 2\mu S_{11}^H \cdot S_{11} \cdot \underline{a}_{1,(k)}^+,$$

where $\underline{a}_{1,(k)}^+$ represents the modal coefficient vector of the wavefront produced at the $k$th iteration of the algorithm and $\mu$ is a positive step size. This yields Algorithm 1, which iteratively refines the wavefront $\underline{a}_{1,(k+1)}^+$ until the backscattered intensity $\|S_{11} \cdot \underline{a}_{1,(k)}^+\|_2^2$ drops below a preset threshold $\epsilon$.

---

**Algorithm 1: Steepest Descent Algorithm for Finding $\underline{a}_{\text{opt}}$**

---

1: Input: $\underline{a}_{1,(0)}^+$ = Initial random vector with unit norm
2: Input: $\mu > 0$ = step size
3: Input: $\epsilon$ = Termination condition
4: $k = 0$
5: **while** $\|S_{11} \cdot \underline{a}_{1,(k)}^+\|_2^2 > \epsilon$ **do**
6:     $\tilde{\underline{a}}_{1,(k)}^+ = \underline{a}_{1,(k)}^+ - 2\mu S_{11}^H \cdot S_{11} \cdot \underline{a}_{1,(k)}^+$
7:     $\underline{a}_{1,(k+1)}^+ = \tilde{\underline{a}}_{1,(k)}^+ / \|\tilde{\underline{a}}_{1,(k)}^+\|_2$
8:     $k = k + 1$
9: **end while**

---

Armed with the relationship in Eq. (15), step 6 in Algorithm 1 can be expressed as

$$\tilde{\underline{a}}_{1,(k)}^+ = \underline{a}_{1,(k)}^+ - 2\mu S_{11}^H \cdot S_{11} \cdot \underline{a}_{1,(k)}^+$$
$$= \underline{a}_{1,(k)}^+ - 2\mu F \cdot S_{11}^* \cdot F \cdot S_{11} \cdot \underline{a}_{1,(k)}^+. \tag{16}$$

This allows us to recast each step of Algorithm 1 into the counterparts of the physical operations in the second column of Table 1.

The sequence of steps 1–4 in Table 1, which involves *double phase conjugation*, amplifies the highly backscattering component in the wavefront, analogous to the operations for time-reversal focusing [21,34–36]. In step 5, this component is subtracted, leading to a refined wavefront that will backscatter less. This process is repeated until convergence. A consequence of this technique is that the backscatter field intensity will typically decrease monotonically. This makes measurement of the backscatter modal coefficient vector increasingly difficult as the iteration progresses. An additional disadvantage of this method is the obvious need to carefully set $\mu$ to guarantee convergence, $0 < \mu < (1/\tilde{\sigma}_1^2 + \tilde{\sigma}_M^2) \approx 1$. In an experimental setting, the step size $\mu$ is chosen by a simple line search, i.e., by scanning a set of discretized values and selecting the one that results in the smallest backscatter intensity after a fixed number of iterations.

We describe a method next that maintains high backscatter field intensity throughout the process and does not require selection of any other auxiliary parameters to guarantee convergence.

## B. Conjugate Gradient Method
Consider an iterative solution to Eq. (8) where the iterate (before normalization for power) is formed as

$$\underline{a}_{1,(k+1)}^+ = \underline{a}_{1,(k)}^+ + \mu_{(k+1)}\underline{d}_{(k)}, \tag{17}$$

where $\mu_{(k+1)}$ is a step size and $\underline{d}_{(k)}$ is the search direction. In this framework, Algorithm 1 results from setting $\mu_{(k+1)} = \mu$ and $\underline{d}_{(k)} = -2S_{11}^H \cdot S_{11} \cdot \underline{a}_{1,(k)}^+$.

The conjugate gradient method (see [25, Chap. 5] for a detailed derivation) results from choosing the step size

$$\mu_{(k+1)} = \|\underline{r}_{(k)}\|_2^2 / \|S_{11} \cdot \underline{d}_{(k)}\|_2^2 \tag{18a}$$

with the search direction given by

$$\underline{d}_{(k+1)} = \underline{r}_{(k+1)} + \beta_{(k+1)}\underline{d}_{(k)} \tag{18b}$$

**Table 1.   Steepest Descent Algorithm for Transmission Maximization**[a]

| Vector Operation | Physical Operation |
|---|---|
| 1: $\underline{a}_1^- = S_{11} \cdot \underline{a}_{1,(k)}^+$ | 1: $\underline{a}_{1,(k)}^+ \xrightarrow{\text{Backscatter}} \underline{a}_1^-$ |
| 2: $\underline{a}_1^+ = F \cdot (\underline{a}_1^-)^*$ | 2: $\underline{a}_1^- \xrightarrow{\text{PCM}} \underline{a}_1^+$ |
| 3: $\underline{a}_1^- = S_{11} \cdot \underline{a}_1^+$ | 3: $\underline{a}_1^+ \xrightarrow{\text{Backscatter}} \underline{a}_1^-$ |
| 4: $\underline{a}_1^+ = F \cdot (\underline{a}_1^-)^*$ | 4: $\underline{a}_1^- \xrightarrow{\text{PCM}} \underline{a}_1^+$ |
| 5: $\tilde{\underline{a}}_1^+ = \underline{a}_{1,(k)}^+ - 2\mu\underline{a}_1^+$ | 5: $\tilde{\underline{a}}_1^+ = \underline{a}_{1,(k)}^+ - 2\mu\underline{a}_1^+$ |
| 6: $\underline{a}_{1,(k+1)}^+ = \tilde{\underline{a}}_1^+ / \|\tilde{\underline{a}}_1^+\|_2$ | 6: $\tilde{\underline{a}}_1^+ \xrightarrow{\text{Normalization}} \underline{a}_{1,(k+1)}^+$ |

[a]The first column represents vector operations in Algorithm 1. The second column represents the physical (or experimental) counterpart. The operation $\underline{a}_1^- \mapsto F \cdot (\underline{a}_1^-)^*$ can be realized via the use of a phase-conjugating mirror (PCM). The algorithm terminates when the backscatter intensity falls below a preset threshold $\epsilon$.

and

$$\beta_{(k+1)} = \|\underline{r}_{(k+1)}\|_2^2 / \|\underline{r}_{(k)}\|_2^2. \tag{18c}$$

Here, the residual vector is

$$\underline{r}_{(k+1)} = -S_{11}^H \cdot S_{11} \cdot \underline{a}_{1,(k+1)}^+. \tag{18d}$$

The iteration terminates when $\|\underline{r}_{(k+1)}\|_2 < \epsilon$, a preset threshold.

Plugging Eq. (17) into Eq. (18d) and substituting the expressions in Eqs. (18a)–(18c) gives us an alternate expression for the residual vector

$$\underline{r}_{(k+1)} = \underline{r}_{(k)} - \mu_{(k+1)} S_{11}^H \cdot S_{11} \cdot \underline{d}_{(k)} \tag{19a}$$

or, equivalently,

$$\underline{r}_{(k+1)} = \underline{r}_{(k)} - \frac{\|\underline{r}_{(k)}\|_2^2}{\|S_{11} \cdot \underline{d}_{(k)}\|_2^2} S_{11}^H \cdot S_{11} \cdot \underline{d}_{(k)}. \tag{19b}$$

The utility of Eq. (19b) will become apparent shortly.

To summarize, we described an iterative method for refining the wavefront $\underline{a}_{1,(k)}^+$ via Eq. (17). Inspection of the update Eqs. (18a)–(18c) and (19b) reveals that matrix–vector operations $S_{11} \cdot \underline{d}_{(k)}$ appear in Eq. (18a) while $S_{11}^H \cdot S_{11} \cdot \underline{d}_{(k)}$ appears in Eq. (19b). This means that the vector $\underline{d}_{(k)}$ is transmitted and the associated backscatter is measured. Note that these measurements are used to iteratively refine the vector $\underline{a}_{1,(k)}^+$, but $\underline{a}_{1,(k)}^+$ is *never actually transmitted* until the termination condition $\|\underline{r}_{(k+1)}\|_2 < \epsilon$ is met. This is reflected in the physical description of the proposed algorithm in Table 2. Also, note that we start with a random unit vector $\underline{a}_{1,(0)}^+$, and set $\underline{d}_{(0)}$ and $\underline{r}_{(0)}$ to $-S_{11}^H \cdot S_{11} \cdot \underline{a}_{1,(0)}^+$, since we are using conjugate gradient for finding the input that minimizes reflection, i.e.,

$$-\underline{a}_{1,(0)}^+ \xrightarrow{\text{Backscatter}} \underline{a}_1^- \xrightarrow{\text{PCM}} \underline{a}_1^+ \xrightarrow{\text{Backscatter}} \underline{a}_1^- \xrightarrow{\text{PCM}} \underline{d}_{(0)} = \underline{r}_{(0)}.$$

A feature of the conjugate gradient method is that the intensity of the backscatter measurement $S_{11} \cdot \underline{d}_{(k)}$ is expected

### Table 2. Conjugate Gradient Algorithm for Transmission Maximization[a]

| Vector Operation | Physical Operation |
|---|---|
| 1: $\underline{d}_1^- = S_{11} \cdot \underline{d}_{(k)}$ | 1: $\underline{d}_{(k)} \xrightarrow{\text{Backscatter}} \underline{d}_1^-$ |
| 2: $\underline{d}_1^+ = F \cdot (\underline{d}_1^-)^*$ | 2: $\underline{d}_1^- \xrightarrow{\text{PCM}} \underline{d}_1^+$ |
| 3: $\underline{d}_1^- = S_{11} \cdot \underline{d}_1^+$ | 3: $\underline{d}_1^+ \xrightarrow{\text{Backscatter}} \underline{d}_1^-$ |
| 4: $\underline{d} = F \cdot (\underline{d}_1^-)^*$ | 4: $\underline{d}_1^- \xrightarrow{\text{PCM}} \underline{d}$ |
| 5: $\mu_{(k+1)} = \|\underline{r}_{(k)}\|_2^2 / (\underline{d}_{(k)}^H \cdot \underline{d})$ | 5: $\mu_{(k+1)} = \|\underline{r}_{(k)}\|_2^2 / (\underline{d}_{(k)}^H \cdot \underline{d})$ |
| 6: $\underline{r}_{(k+1)} = \underline{r}_{(k)} - \mu_{(k+1)} \underline{d}$ | 6: $\underline{r}_{(k+1)} = \underline{r}_{(k)} - \mu_{(k+1)} \underline{d}$ |
| 7: $\beta_{(k+1)} = \|\underline{r}_{(k+1)}\|_2^2 / \|\underline{r}_{(k)}\|_2^2$ | 7: $\beta_{(k+1)} = \|\underline{r}_{(k+1)}\|_2^2 / \|\underline{r}_{(k)}\|_2^2$ |
| 8: $\underline{d}_{(k+1)} = \underline{r}_{(k+1)} + \beta_{(k+1)} \underline{d}_{(k)}$ | 8: $\underline{d}_{(k+1)} = \underline{r}_{(k+1)} + \beta_{(k+1)} \underline{d}_{(k)}$ |

　　[a]The first column represents iterates of the conjugate gradient method. The second column represents the physical (or experimental) counterpart. The operation $\underline{d}_1^- \mapsto F \cdot (\underline{d}_1^-)^*$ can be realized via the use of a PCM. The algorithm terminates when the residual vector $\|\underline{r}_{(k+1)}\|_2 < \epsilon$, a preset threshold at which point the optimal backscatter-minimizing wavefront is constructed as $\underline{a}_{1,(k+1)}^+ = \underline{a}_{1,(k)}^+ + \mu_{(k+1)} \underline{d}_{(k+1)}$ followed by a power normalization $\underline{a}_{1,(k+1)}^+ = \underline{a}_{1,(k+1)}^+ / \|\underline{a}_{1,(k+1)}^+\|_2$.

to remain relatively high (for a strongly backscattering medium) throughout the process. It is only when the wavefront corresponding to $\underline{a}_{1,(k+1)}^+$ is excited that a strong transmission (with minimized backscatter) is obtained—this might be a desirable feature for communication or covert sensing applications. Consequently, the algorithm will produce high intensity backscatter measurements, thereby facilitating accurate estimation of the backscatter modal coefficient vectors that are important components of the proposed algorithm. This makes the conjugate gradient method less susceptible to measurement noise than the steepest descent method, where the backscatter intensity decreases with every iteration.

## 6. ITERATIVE, PHYSICALLY REALIZABLE FOCUSING ALGORITHM

We first describe a generalized coordinate descent method for amplitude and phase optimization. Assume we are given a $M \times N_B$ matrix $B = [\underline{b}_1 \ldots \underline{b}_{N_B}]$ whose columns are orthonormal so that $B^H \cdot B = I_{N_B}$. Thus $N_B$ denotes the number of (orthonormal) basis vectors.

The key idea here is to expand $\underline{a}_1^+$ on the right-hand side of Eq. (11) in terms of the basis vectors given by the columns of $B$ as

$$\underline{a}_1^+ = \sum_{l=1}^{N_B} p_l e^{j\phi_l} \underline{b}_l, \tag{20}$$

where $p_l \geq 0$ and $\phi_l \in [-\pi, \pi]$ are the unknown amplitudes and phases, respectively.

The optimal amplitudes can be estimated by transmitting $\underline{a}_1^+ = \underline{b}_l$ for every $l = 1, \ldots N_B$, measuring the corresponding intensity $\mathcal{I}_l$ at the target, and setting $p_l = \sqrt{\mathcal{I}_l}$. This can be accomplished with $O(N_B)$ measurements.

The phases can be estimated by first setting $\phi_1, \ldots \phi_{N_B}$ randomly and then for $l = 1, \ldots, N_B$, sequentially finding the phase that optimizes measured intensity. This can be done via a simple line search, i.e., by scanning the measured intensity over a fixed set of discretized values of the phase or by using more sophisticated algorithms such as the golden section search algorithm with parabolic interpolation [37, Section 10.2]. This too requires $O(N_B)$ measurements.

Setting $N_B = M$ and $B = I$ yields the coordinate descent approach used by Vellekoop and Mosk [7,8,24]. This corresponds to exciting one plane wave mode at a time and inferring the optimal phase and amplitude one mode at time. Such an algorithm requires $O(M)$ iterations to yield the optimal focusing wavefront. Setting $B$ to the 2D Walsh function basis matrix yields the method proposed by Stockbridge *et al.* in [17].

An important insight from Eq. (13) is that if we were to express the optimal focusing wavefront as a superposition of eigen-wavefronts of $S_{21}$, then typically only $K \ll M$ of the combining coefficients will be large. Thus only $K$ of the $p_l$ coefficients in Eq. (20) will be significant if we set $B$ to be the right singular vectors of $S_{21}$. In the lossless setting, the $K$ eigen-wavefronts of $S_{21}$ that have transmission coefficients close to 1 correspond precisely to the $K$ eigen-wavefronts associated with $S_{11}$ that have reflection coefficients close to 0. Hence, we can set $B$ to be the right singular vectors of $S_{11}$ and expect only $K$ of the $p_l$ coefficients in Eq. (20) to be significant as well. Thus, we need to measure the $K$ singular vectors of $S_{11}$ associated with its $K$ smallest singular values.

**Table 3.  Lanzcos Algorithm and Its Physical Counterpart That Computes a Tridiagonal Matrix $H$ Whose Eigenvalues and Eigenvectors Are Closely Related to the Eigenvalues and Eigenvectors of $S_{11}^H \cdot S_{11}$ [a]**

| Vector Operation | Physical Operation |
|---|---|
| 1: $\underline{q}_1^- = S_{11} \cdot \underline{q}_{(k)}$ | 1: $\underline{q}_{(k)} \xrightarrow{\text{Backscatter}} \underline{q}_1^-$ |
| 2: $\underline{q}_1^+ = F \cdot (\underline{q}_1^-)^*$ | 2: $\underline{q}_1^- \xrightarrow{\text{PCM}} \underline{q}_1^+$ |
| 3: $\underline{q}_1^- = S_{11} \cdot \underline{q}_1^+$ | 3: $\underline{q}_1^+ \xrightarrow{\text{Backscatter}} \underline{q}_1^-$ |
| 4: $\underline{v} = F \cdot (\underline{q}_1^-)^*$ | 4: $\underline{q}_1^- \xrightarrow{\text{PCM}} \underline{v}$ |
| 5: $H_{k,k} = \underline{q}_{(k)}^H \cdot \underline{v}$ | 5: $H_{k,k} = \underline{q}_{(k)}^H \cdot \underline{v}$ |
| 6: $\underline{v} = \underline{v} - H_{k,k}\underline{q}_{(k)} - s_{(k-1)}\underline{q}_{(k-1)}$ | 6: $\underline{v} = \underline{v} - H_{k,k}\underline{q}_{(k)} - s_{(k-1)}\underline{q}_{(k-1)}$ |
| 7: $H_{k+1,k} = H_{k,k+1} = s_{(k)} = \|\underline{v}\|_2$ | 7: $H_{k+1,k} = H_{k,k+1} = s_{(k)} = \|\underline{v}\|_2$ |
| 8: $\underline{q}_{(k+1)} = \underline{v}/s_{(k)}$ | 8: $\underline{q}_{(k+1)} = \underline{v}/s_{(k)}$ |

[a]Note that we initialize the algorithm by setting $k = 1$, $\underline{q}_{(1)}$ to a random unit norm vector, and $s_{(0)} = 0$.

The Lanczos algorithm is an iterative algorithm for accomplishing just that [25,26]. The key idea is to create a tridiagonal matrix $H$ whose eigenvalues and eigenvectors (referred to as the Ritz values and vectors) are approximations of the eigenvalues and eigenvectors of $S_{11}^H \cdot S_{11}$. The algorithm is summarized in the first column of Table 3; its physical counterpart is described in the second column. The matrix $B$ in Eq. (20) is obtained as

$$B = Q \cdot U, \tag{21}$$

where $Q = [\underline{q}_{(1)} \ldots \underline{q}_{(N_B)}]$ are the $N_B$ vectors produced by the algorithm (see Table 3) and $U = [\underline{u}_{(1)} \ldots \underline{u}_{(N_B)}]$ are the $N_B$ eigenvectors of $H$ associated with the $N_B$ smallest eigenvalues.

The convergence theory [26] of the Lanczos algorithms predicts that the eigenvector estimates will rapidly converge to the $K$ eigenvectors of $S_{11}^H \cdot S_{11}$ associated with the eigenwavefronts of $S_{11}$ with the smallest reflection coefficients; hence, setting $N_B = O(K)$ will suffice. An estimate of $K$ can be formed from the eigenvalues of $H$ by counting how many of the converged eigenvalues of $H$ are below a preset threshold $\epsilon$.

Estimating these $K$ right singular vectors will require $O(K)$ measurements, and when $K \ll M$, we shall obtain a near-optimal focusing wavefront using significantly fewer measurements than the $O(M)$ measurements required by the coordinate descent when $B = I$. We shall corroborate this convergence behavior using numerical simulations next.

# 7. NUMERICAL SIMULATIONS AND VALIDATION OF THE EXISTENCE OF HIGHLY TRANSMITTING EIGEN-WAVEFRONTS

To validate the proposed algorithms, we compute the scattering matrices in Eq. (2) via a spectrally accurate, T-matrix-inspired integral equation solver that characterizes fields scattered from each cylinder in terms of their traces expanded in a series of azimuthal harmonics. Interactions between cylinders are modeled using 2D periodic Green's functions. The method constitutes a generalization of that in [38], in that

it does not force cylinders in a unit cell to reside on a line but allows them to be freely distributed throughout the cell. All periodic Green's functions/lattice sums are rapidly evaluated using a recursive Shank's transform as in [39,40]. Our method exhibits exponential convergence in the number of azimuthal harmonics used in the description of the field scattered by each cylinder. In the numerical experiments below, care was taken to ensure eleventh digit accuracy in the entries of the computed scattering matrices.

Figure 4 shows the empirical transmission coefficient distribution, i.e., the singular value squared of the $S_{21}$ matrix of a slab with $D = 197\lambda$, $L = 1.2 \times 10^4\lambda$, $r = 0.11\lambda$, $N_c = 14{,}000$ (dielectric), $n_d = 1.3$, $M = 395$, and $\bar{l} = 6.7\lambda$, where $\bar{l}$ is the mean of the minimum-inter-scatterer-distances. The computation validates the bimodal shape of the theoretical distribution in Fig. 2.

Next, we consider scattering system with $D = 14\lambda$, $L = 5.4\lambda$, $r = 0.11\lambda$, $N_c = 50$ (PEC), $M = 27$, and $\bar{l} = 0.8\lambda$. Here $\tau_{\text{normal}} = 0.483$, while $\tau_{\text{opt}} = 0.9997$, so that wavefront optimization produces a twofold increase in transmitted power. Figures 5(a) and 5(b) show the wavefield produced by a normally incident wavefront and the optimal wavefront, respectively. Figure 6 shows the modal coefficients of the optimal wavefront corresponding to Fig. 5(b).

Figure 7 displays the rate of convergence of the algorithms developed for a setting with $D = 197\lambda$, $L = 3.4 \times 10^5\lambda$, $r = 0.11\lambda$, $N_c = 430{,}000$ (dielectric), $n_d = 1.3$, $M = 395$, and $\bar{l} = 6.69\lambda$; this slab has a comparable (slightly lower) packing density than that in Fig. 5(a).

A normally incident wavefront results in a transmission of $\tau_{\text{normal}} = 0.038$. The optimal wavefront yields $\tau_{\text{opt}} = 0.9973$ corresponding to a 26-fold increase in transmission. Steepest descent algorithm and conjugate gradient algorithm produce wavefronts that converge to the near optimum in about 5–10 iterations, as shown in Fig. 7.

Figure 8 plots the transmitted power after the tenth iteration of steepest descent algorithm for different choices of $\mu$. Figure 8 reveals that there is broad range of $\mu$ for which the algorithm converges in a handful of iterations. We have found that setting $\mu \approx 0.5$ yields fast convergence.
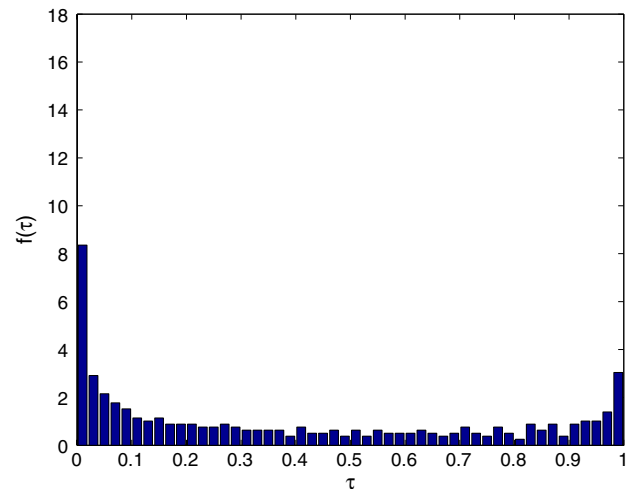


Fig. 4.  Empirical transmission coefficients distribution from a scattering system with $D = 197\lambda$, $L = 1.2 \times 10^4\lambda$, $r = 0.11\lambda$, $N_c = 14{,}000$ (dielectric), $n_d = 1.3$, $M = 395$, $\bar{l} = 6.7\lambda$, where $\bar{l}$ is the mean of the minimum-inter-scatterer-distances.
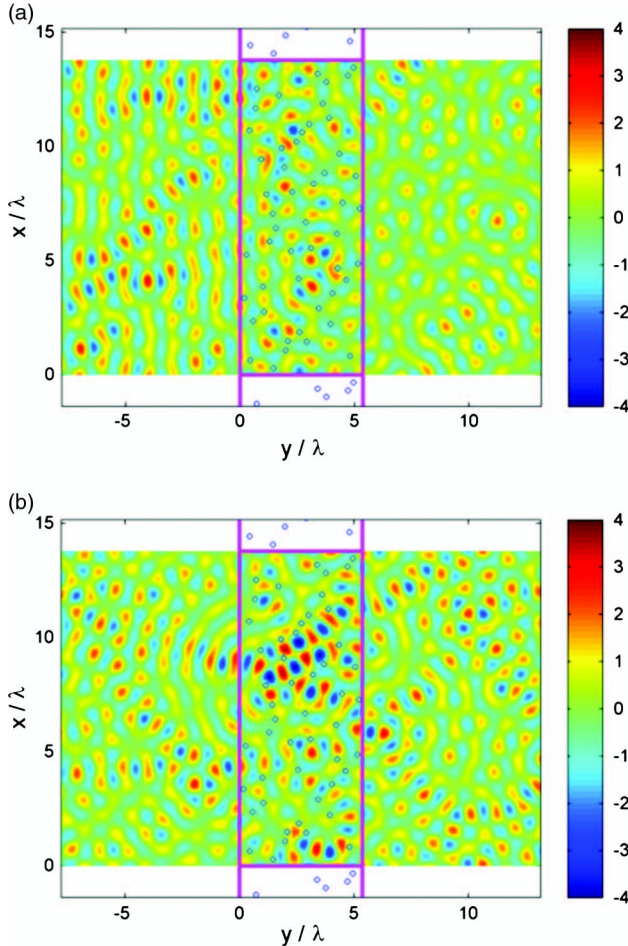
(a)



(b)



Fig. 5. Wavefield plot of the incident-plus-backscatter wave corresponding to (a) normally incident wavefront and (b) optimal wavefront, which were sent to a scattering system with $D = 14\lambda$, $L = 5.4\lambda$, $r = 0.11\lambda$, $N_c = 50$PEC, $M = 27$, $\bar{l} = 0.8\lambda$. The normally incident wavefront has $\tau_{\mathrm{normal}} = 0.483$, while the optimal wavefront yields $\tau_{\mathrm{opt}} = 0.9997$.

The conjugate gradient algorithm converges slightly faster than the steepest descent algorithm in the setting where we chose the optimal $\mu = 0.5037$ for steepest descent algorithm
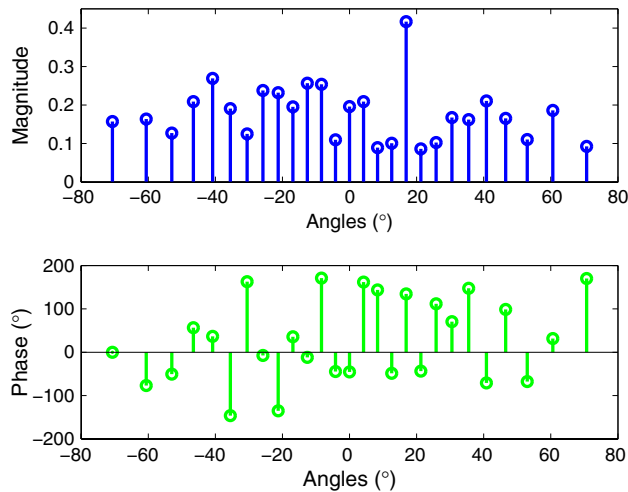


Fig. 6. Modal coefficients of the optimal wavefront corresponding to Fig. 5(b) are shown.
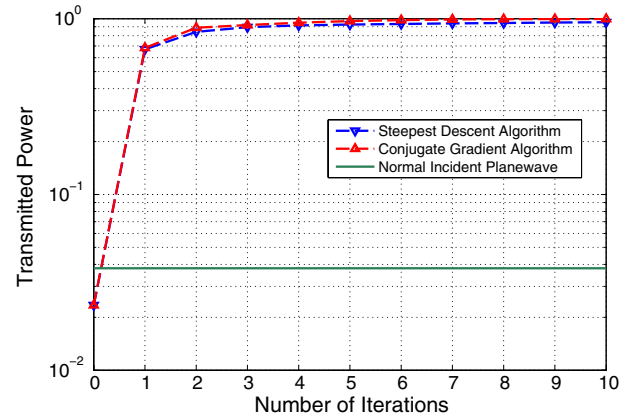


Fig. 7. Transmitted power versus the number of iterations is shown for steepest descent algorithm with $\mu = 0.5037$ and for conjugate gradient in the setting with $D = 197\lambda$, $L = 3.4 \times 10^5\lambda$, $r = 0.11\lambda$, $N_c = 430{,}000$ dielectric cylinders with $n_d = 1.3$, $M = 395$, $\bar{l} = 6.69\lambda$. The conjugate gradient algorithm converged to the optimal transmitted power slightly faster than the steepest descent algorithm. However, since the steepest descent algorithm requires a line search for setting the optimal step size $\mu$, it requires more measurements than the conjugate gradient method, which does not require any parameters to be set.

by a line search; i.e., we ran steepest descent algorithm over a fixed set of discretized values of $\mu$ between 0 and 1, and chose the optimal $\mu$ that gives the fastest convergence result. In an experimental setting, the line search for finding the optimal $\mu$ for the steepest descent algorithm will require additional measurements. Thus, conjugate gradient algorithm will require fewer measurements than steepest descent algorithm with the additional advantage of not requiring any auxiliary parameters to be set.

Next, we consider the setting where a subset of the propagation modes are controlled so that the summation in Eq. (1) is from $-N_{\mathrm{ctrl}}$ to $N_{\mathrm{ctrl}}$. Thus the number of controlled modes is given by $M_{\mathrm{ctrl}} = 2N_{\mathrm{ctrl}} + 1$.

Figure 9 shows the realized gain (relative to a normally incident wavefront) for three different approaches versus the number of control modes in the same setting as in Fig. 7. Here we compute the realized gain for algorithms that control only part of the total number of modes but capture (1) all modes in the backscatter field, (2) only as many modes in the
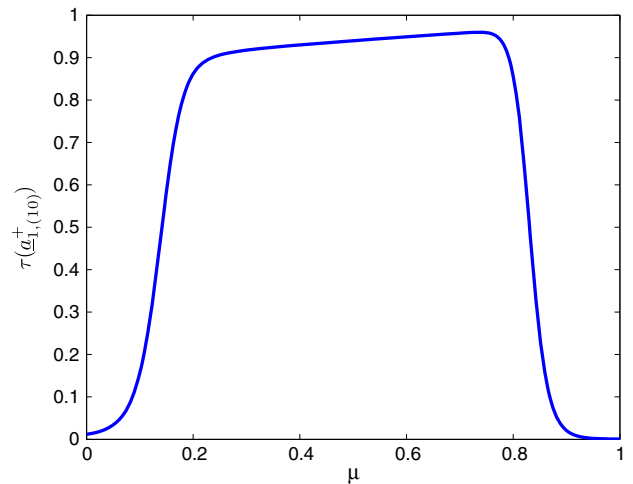


Fig. 8. Transmitted power at the tenth iteration as a function of the step size $\mu$ used in Algorithm 1 for the same setting as in Fig. 7.
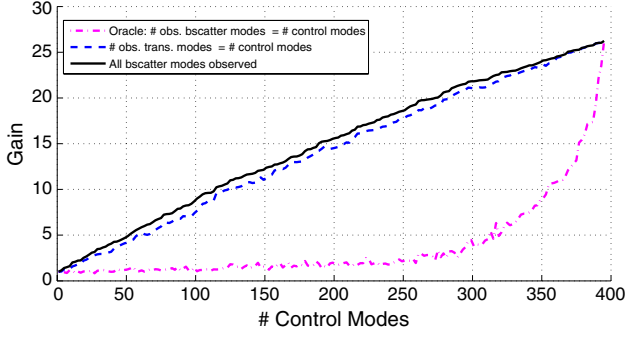
Fig. 9.   Gain ($=:\tau_{\text{opt}}/\tau_{\text{normal}}$) versus the number of control modes for the same setting as in Fig. 7. Here we compute the realized gain for algorithms that control only part of the total number of modes, but capture (1) all modes in the backscatter field, (2) only as many modes in the transmitted field as the number of control modes, and (3) only as many modes in the backscatter field as the number of control modes. For the last algorithm, we transmit the eigen-wavefront of the (portion of the) $S_{11}$ matrix that yields the highest transmission.

transmitted field as the number of control modes, and (3) only as many modes in the backscatter field as the number of control modes. For the last algorithm, we transmit the eigen-wavefront of the (portion of the) $S_{11}$ matrix that yields the highest transmission. Figure 9 shows that if the backscatter field is fully sampled, then it is possible to realize increased transmission with a limited number of control modes. It also emphasizes the important point that when the backscatter field is not fully sampled, the principle of minimizing backscatter might produce "transmission" into the unsampled portion of the backscatter field instead of producing forward transmission.

Figure 10 considers the same setup as in Fig. 7 with a target at $(D/2, L + 5.4\lambda)$ and plots the focus achieved at the target by exciting a focusing wavefront as in Eq. (12). The modal coefficients are plotted in Fig. 11(a). Figure 11(b) shows the sparsity of the modal coefficients of the optimal focusing wavefront when expressed in terms of the basis given by the right singular vectors of the $S_{11}$ matrix or, equivalently, the eigenvectors of $S_{11}^{H} \cdot S_{11}$.

Figure 12 plots the intensity achieved when using $N_B$ basis vectors for the algorithms described in Section 6 in the same setup as in Fig. 10. The new algorithm, which computes the





Fig. 11.   Here, we depict the magnitude of the coefficients of the optimal focusing wavefront, corresponding to the situation in Fig. 10, in terms of two choices of basis vectors. In (a) we decompose the optimal focusing wavefront with respect to the basis vectors corresponding to plane waves; in (b) we decompose the optimal focusing wavefront with respect to the basis vectors associated with the eigen-wavefronts of the $S_{11}$ matrix. A particular important observation is that the eigen-wavefront decomposition yields a sparse representation of the optimal focusing wavefront.
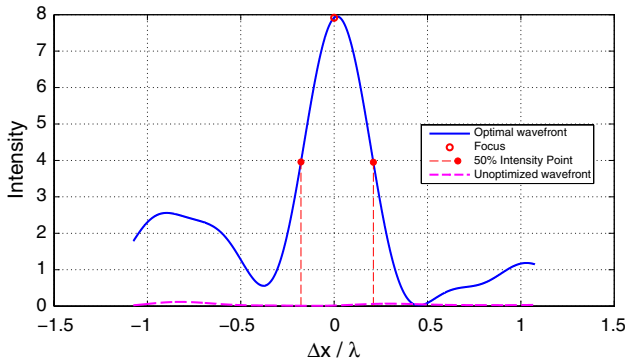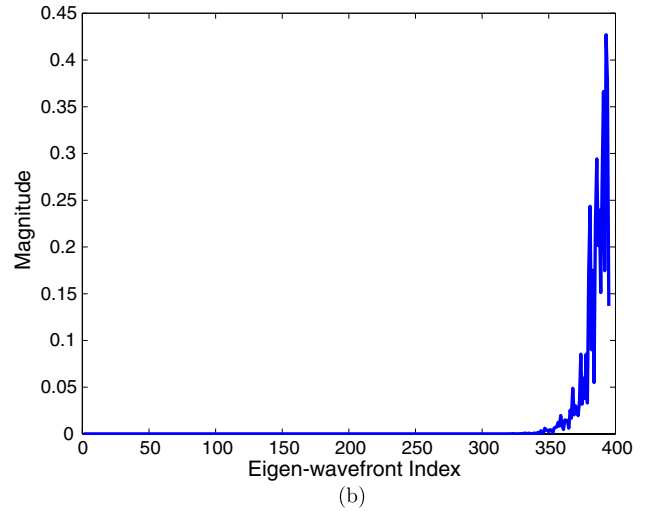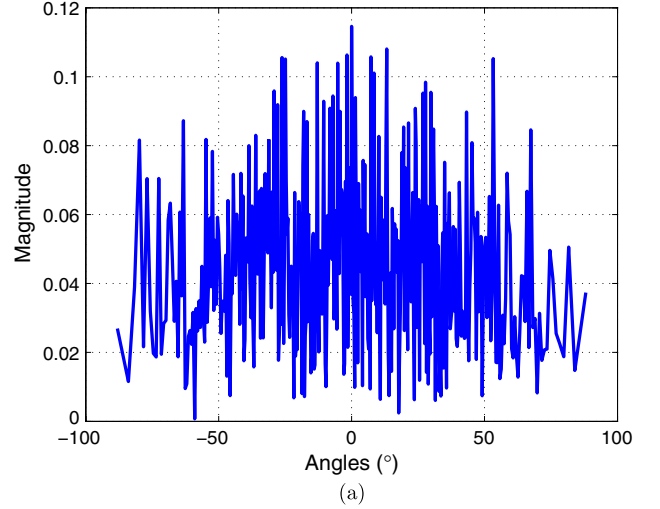


Fig. 10.   Intensity plot around the target at $(D/2, L + 5.4\lambda)$ for the scattering system defined in Fig. 7. The optimal focusing wavefront forms a sharp focus of $1\lambda$ around the target. The unoptimized wavefront solution corresponds to an incident wavefront that would have produced a focus at the target if there were no intervening scattering medium.

basis $B$ from the eigenvectors of $S_{11}^{H} \cdot S_{11}$ associated with its smallest eigenvalues, reaches 95% of the optimal intensity with significantly fewer iterations than the coordinate descent algorithm. This fast convergence to the near-optimum is the principal advantage of the proposed method. Figure 12 shows that this convergence behavior is retained even when the number of control modes is reduced. We obtain similar gains for the setting in which there are multiple focusing points.

Finally, we consider the setting in which the scatterers are absorptive. Here, backscatter minimization as a general principle for increasing transmission is clearly suboptimal since an input with significant absorption can also minimize backscatter. We defined gain as $\tau_{\text{opt}}/\tau_{\text{normal}}$. Here we have $D = 197\lambda$, $L = 3.4 \times 10^5 \lambda$, $r = 0.11\lambda$, $N_c = 4.3 \times 10^5$ (absorbing dielectric), $n_d = 1.3 - j\kappa$, $M = 395$, and $\bar{l} = 6.69\lambda$. In Fig. 13, we compare the gain obtained by using the backscatter-minimizing wavefront to the gain obtained by the optimal
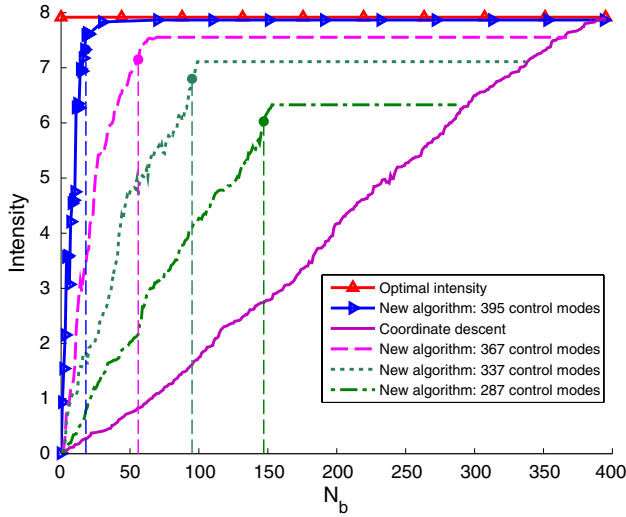
Fig. 12. Intensity at target as a function of the number of basis vectors for the new algorithm [which uses the basis vectors estimated using Eq. (21) and the algorithm described in Table 3] for different number of control modes versus the standard coordinate descent method, which uses the plane wave associated basis vectors (see Section 6) for the same setting as in Fig. 10. The sparsity of the optimal wavefront's modal coefficient vector when expressed using the basis of the eigen-wavefronts [shown in Fig. 11(b)] leads to the rapid convergence observed. The optimal wavefront was constructed as described in Section 3.B using time reversal. The number of basis vectors needed to attain 95% of the optimal focus intensity for a given number of control modes is indicated with a vertical line highlighting the fast convergence of the algorithm and the ability to obtain a near-optimal focus using significantly fewer measurements than the coordinate descent approach.
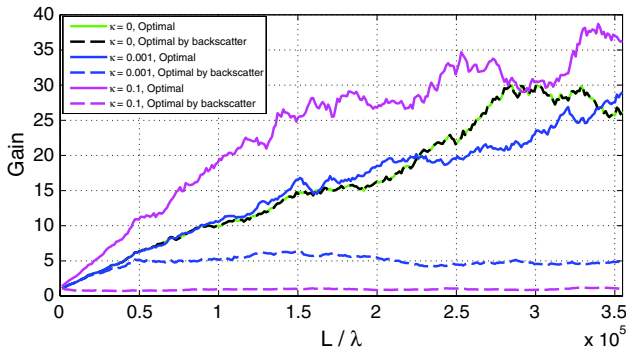


Fig. 13. Gain ($=\tau_{opt}/\tau_{normal}$) versus thickness $L/\lambda$ in a setting with $D = 197\lambda$, $r = 0.11\lambda$, $N_c = 430,000$ Absorbing Dielectric, $n_d = 1.3 - j\kappa$, $M = 395$, $\bar{l} = 6.69\lambda$, for different values of $\kappa$. The solid line represents the maximum possible gain, and the dashed line represents the gain obtained using the backscatter-minimizing algorithm discussed in Section 5.

wavefront (which utilizes information from the $S_{21}$ matrix) for various $\kappa$, as the thickness of the scattering system increases. We obtain an increase in transmission and the methods described again produce dramatic gains whenever the scatterers are weakly absorptive.

## 8. CONCLUSIONS

We have numerically verified the existence of eigen-wavefronts with transmission coefficients approaching 1 in highly scattering systems and developed physically realizable algorithms for finding these highly transmitting

eigen-wavefronts using backscatter analysis. We also developed a physically realizable algorithm for forming a focused input using the highly transmitting eigen-wavefronts identified by the previous algorithm. Via numerical simulations, it was shown that the algorithms converged to a near-optimal wavefront in just a few iterations. The proposed algorithms are quite general and may be applied to scattering problems beyond the 2D setup described in the simulations. We are currently investigating extensions to imaging and sensing applications. A detailed study of the impact of periodic boundary conditions on the results obtained is also underway.

## REFERENCES

1. A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Wiley, 1999), Vol. **12**.
2. O. N. Dorokhov, "Transmission coefficient and the localization length of an electron in $N$ bound disordered chains," JETP Lett. **36**, 458 (1982).
3. J. B. Pendry, A. MacKinnon, and A. B. Pretre, "Maximal fluctuations—a new phenomenon in disordered systems," Phys. A **168**, 400–407 (1990).
4. C. Barnes and J. B. Pendry, "Multiple scattering of waves in random media: a transfer matrix approach," Proc. R. Soc. London Ser. A Math. Phys. Eng. Sci. **435**, 185–196 (1991).
5. P. A. Mello, P. Pereyra, and N. Kumar, "Macroscopic approach to multichannel disordered conductors," Ann. Phys. **181**, 290–317 (1988).
6. C. W. J. Beenakker, "Applications of random matrix theory to condensed matter and optical physics," Arxiv preprint arXiv:0904.1432 (2009).
7. I. M. Vellekoop and A. P. Mosk, "Phase control algorithms for focusing light through turbid media," Opt. Commun. **281**, 3071–3080 (2008).
8. I. M. Vellekoop and A. P. Mosk, "Universal optimal transmission of light through disordered materials," Phys. Rev. Lett. **101**, 120601 (2008).
9. 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).
10. T. W. Kohlgraf-Owens and A. Dogariu, "Transmission matrices of random media: means for spectral polarimetric measurements," Opt. Lett. **35**, 2236–2238 (2010).
11. Z. Shi, J. Wang, and A. Z. Genack, "Measuring transmission eigenchannels of wave propagation through random media," in *Frontiers in Optics* (Optical Society of America, 2010).
12. M. Kim, Y. Choi, C. Yoon, W. Choi, J. Kim, Q.-Han. Park, and W. Choi, "Maximal energy transport through disordered media with the implementation of transmission eigenchannels," Nat. Photonics **6**, 581–585 (2012).
13. E. G. van Putten, A. Lagendijk, and A. P. Mosk, "Optimal concentration of light in turbid materials," J. Opt. Soc. Am. B **28**, 1200–1203 (2011).
14. J. Aulbach, B. Gjonaj, P. M. Johnson, A. P. Mosk, and A. Lagendijk, "Control of light transmission through opaque scattering media in space and time," Phys. Rev. Lett. **106**, 103901 (2011).
15. M. Cui, "A high speed wavefront determination method based on spatial frequency modulations for focusing light through random scattering media," Opt. Express **19**, 2989–2995 (2011).
16. M. Cui, "Parallel wavefront optimization method for focusing light through random scattering media," Opt. Lett. **36**, 870–872 (2011).

17. C. Stockbridge, Y. Lu, J. Moore, S. Hoffman, R. Paxman, K. Toussaint, and T. Bifano, "Focusing through dynamic scattering media," Opt. Express **20**, 15086–15092 (2012).

18. W. Choi, A. P. Mosk, Q.-H. Park, and W. Choi, "Transmission eigenchannels in a disordered medium," Phys. Rev. B **83**, 134207 (2011).

19. C. Jin, R. R. Nadakuditi, E. Michielssen, and S. Rand, "An iterative, backscatter-analysis based algorithm for increasing transmission through a highly-backscattering random medium," in *2012 IEEE Statistical Signal Processing Workshop (SSP)* (IEEE, 2012), pp. 97–100.

20. R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University, 1990).

21. M. Fink, "Time-reversal mirrors," J. Phys. D **26**, 1333–1350 (1993).

22. M. Cui and C. Yang, "Implementation of a digital optical phase conjugation system and its application to study the robustness of turbidity suppression by phase conjugation," Opt. Express **18**, 3444–3455 (2010).

23. M. Cui, E. J. McDowell, and C. Yang, "An in vivo study of turbidity suppression by optical phase conjugation (tsopc) on rabbit ear," Opt. Express **18**, 25–30 (2010).

24. I. M. Vellekoop, A. Lagendijk, and A. P. Mosk, "Exploiting disorder for perfect focusing," Nat. Photonics **4**, 320–322 (2010).

25. H. A. Van der Vorst, *Iterative Krylov Methods for Large Linear Systems* (Cambridge University, 2003), Vol. **13**.

26. L. N. Trefethen and D. Bau III, *Numerical Linear Algebra* (Society for Industrial Mathematics, 1997).

27. S. Grilli, P. Ferraro, S. De Nicola, A. Finizio, G. Pierattini, and R. Meucci, "Whole optical wavefields reconstruction by digital holography," Opt. Express **9**, 294–302 (2001).

28. E. G. van Putten, I. M. Vellekoop, and A. P. Mosk, "Spatial amplitude and phase modulation using commercial twisted nematic LCDs," Appl. Opt. **47**, 2076–2081 (2008).

29. J. A. Kong, *Electromagnetic Wave Theory* (Wiley, 1986).

30. H. A. Haus, *Waves and Fields in Optoelectronics* (Prentice-Hall, 1984), Vol. **1**.

31. R. Carminati, J. J. Saenz, J.-J. Greffet, and M. Nieto-Vesperinas, "Reciprocity, unitarity, and time-reversal symmetry of the s matrix of fields containing evanescent components," Phys. Rev. A **62**, 012712 (2000).

32. R. J. Potton, "Reciprocity in optics," Rep. Prog. Phys. **67**, 717–754 (2004).

33. M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley, 1991).

34. C. Prada and M. Fink, "Eigenmodes of the time reversal operator: a solution to selective focusing in multiple-target media," Wave Motion **20**, 151–163 (1994).

35. M. Fink, C. Prada, F. Wu, and D. Cassereau, "Self focusing in inhomogeneous media with time reversal acoustic mirrors," in *Proceedings of 1989 IEEE Ultrasonics Symposium* (IEEE, 1989), pp. 681–686.

36. M. F. Yanik and S. Fan, "Time reversal of light with linear optics and modulators," Phys. Rev. Lett. **93**, 173903 (2004).

37. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (Cambridge University, 2007).

38. R. C. McPhedran, L. C. Botten, A. A. Asatryan, N. A. Nicorovici, P. A. Robinson, and C. M. De Sterke, "Calculation of electromagnetic properties of regular and random arrays of metallic and dielectric cylinders," Phys. Rev. E **60**, 7614–7617 (1999).

39. S. Singh and R. Singh, "On the use of shank's transform to accelerate the summation of slowly converging series," IEEE Trans. Microwave Theor. Tech. **39**, 608–610 (1991).

40. A. Sidi, *Practical Extrapolation Methods: Theory and Applications* (Cambridge University, 2003), Vol. **10**.