Back-scatter analysis based algorithms for increasing transmission through highly-scattering random media using phase-only modulated wavefronts

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

Dept. of EECS, University of Michigan, Ann Arbor, Michigan 48109-2122, USA

*Corresponding author: jsirius@umich.edu
Recent theoretical and experimental advances have shed light on the existence of so-called ‘perfectly transmitting’ wavefronts with transmission coefficients close to 1 in strongly backscattering random media. These perfectly transmitting eigen-wavefronts can be synthesized by spatial amplitude and phase modulation.

Here, we consider the problem of transmission enhancement using phase-only modulated wavefronts. Motivated by bio-imaging applications in which it is not possible to measure the transmitted fields, we develop physically realizable iterative and non-iterative algorithms for increasing the transmission through such random media using backscatter analysis. We theoretically show that, despite the phase-only modulation constraint, the non-iterative algorithms will achieve at least about $25\pi% \approx 78.5\%$ transmission assuming there is at least one perfectly transmitting eigen-wavefront and that the singular vectors of the transmission matrix obey a maximum entropy principle so that they are isotropically random.

We numerically analyze the limits of phase-only modulated transmission in 2-D with fully spectrally accurate simulators and provide rigorous numerical evidence confirming our theoretical prediction in random media with periodic boundary conditions that is composed of hundreds of thousands of non-absorbing scatterers. We show via numerical simulations that the iterative algorithms we have developed converge rapidly, yielding highly transmitting wavefronts using relatively few measurements of the backscatter field. Specifically, the best performing iterative algorithm yields $\approx 70\%$ transmission using just $15 - 20$ measurements in the regime where the non-iterative algorithms yield $\approx 78.5\%$ transmission but require measuring the entire modal reflection matrix. Our theoretical analysis and rigorous numerical results validate our prediction that phase-only modulation with a given number of spatial modes will yield higher transmission than amplitude and phase modulation with half as many modes.

© 2014 Optical Society of America

OCIS codes: 030.6600

1. Introduction

Multiple scattering by randomly placed particles frustrates the passage of light through ‘opaque’ materials such as turbid water, white paint, and egg shells. Thanks to the theoretical work of Dorokhov [7], Pendry [2, 20], and others [3, 19], as well as the breakthrough experiments of Vellekoop and Mosk [32, 33] and others [1, 5, 6, 15, 16, 21, 23, 27, 29], we now understand that even though a normally incident wavefront will barely propagate through a
thick slab of such media [12], a small number of eigen-wavefronts exist that exhibit a transmission coefficient close to one and hence propagate through the slab transmitted through the slab without significant loss.

In highly scattering random media composed of non-absorbing scatterers, these ‘perfectly’ transmitting eigen-wavefronts are the right singular vectors of the transmission matrix with singular values (or transmission coefficients) close to 1. Thus, if the transmission matrix were measured using the techniques described in [15, 16, 21, 23], one could compute the pertinent singular vector and synthesize a highly transmitting eigen-wavefront via spatial amplitude and phase modulation. The task of amplitude and phase modulating an optical wavefront is not, however, trivial. Calibration and alignment issues prevent its use of two independent spatial light modulators in series that separately modulate the signal amplitude and phase. A viable option is to use the innovative method developed by van Putten et al. in [30] for full spatial phase and amplitude control using a twisted nematic LCD combined with a spatial filter.

In a recent paper [13,14], we assumed that amplitude and phase modulation was feasible, and developed iterative, physically-realizable algorithms for synthesizing highly transmitting eigen-wavefronts using just a few measurements of the backscatter field. We showed that the algorithms converge rapidly and achieve 95% transmission using about 5 − 10 measurements. Our focus on constructing highly transmitting eigen-wavefronts by using the information in the backscatter field was motivated by applications in bio-imaging, where it often is impossible to measure transmitted fields.

Here, we place ourselves in the setting where we seek to increase transmission via backscatter analysis but are restricted to phase-only modulation. The phase-only modulation constraint was initially motivated by the simplicity of the resulting experimental setup (see Fig. 1) and the commercial availability of finely calibrated phase-only spatial light modulators (SLMs) (e.g. the PLUTO series from Holoeye). As we shall shortly see there is another engineering advantage conferred by these methods. We do not, however, expect to achieve perfect transmission using phase-only modulation as is achievable by amplitude and phase modulation. However, we theoretically show that we can expect to get at least (about) 25π% ≈ 78.5% provided that 1) the system modal reflection (or transmission) matrix is known, 2) its right singular vectors obey a maximum entropy principle by being isotropically random, and 3) full amplitude and phase modulation permits at least one perfectly transmitting wavefront. We also develop iterative, physically realizable algorithms for transmission maximization that utilize backscatter analysis to produce a highly transmitting phase-only modulated wavefront in just a few iterations. These rapidly converging algorithms build on the ideas developed in [13,14] by incorporating the phase-only constraint. An additional advantage conferred by these rapidly converging algorithms is that they might facilitate their
use in applications where the duration in which the modal transmission or reflection matrix can be assumed to be constant is relatively small compared to the time it would take to make all measurements needed to estimate the modal transition or reflection matrix or in settings where a near-optimal solution obtained fast is preferable to the optimal solution that takes many more measurements to compute. As in [14], the iterative algorithms we have developed we retain the feature that they allow the number of modes being controlled via an SLM in experiments to be increased without increasing the number of measurements that have to be made.

We numerically analyze the limits of phase-only modulated transmission in 2-D with fully spectrally accurate simulators and provide rigorous numerical evidence confirming our theoretical prediction in random media with periodic boundary conditions that is composed of hundreds of thousands of non-absorbing scatterers. Specifically, we show that the best performing iterative algorithm yields $\approx 70\%$ transmission using just $15 \sim 20$ measurements in the regime where the non-iterative algorithms yield $\approx 78.5\%$ transmission.

This theoretical prediction brings into sharp focus an engineering advantage to phase-only modulation relative to amplitude and phase modulation that we did not anticipate when we started on this line of inquiry. The clever idea in van Putten et al’s work was to use spatial filtering to combine four neighboring pixels into one superpixel and then independently modulate the phase and the amplitude of light at each superpixel. This implies than an SLM with $M$ pixels can control at most $M/4$ spatial modes. For the given aperture, we can expect the undersampling of the $S_{21}$ to reduce transmission. Undersampling the spatial modes by 75% will reduces the amount of transmission by between 65 – 75%. In contrast, controlling all $M$ spatial modes using phase-only modulation will reduce transmission by only 30%. Thus, we can achieve higher transmission with phase-only modulation using all pixels in an SLM than by (integer-valued) undersampling of the pixels to implement amplitude and phase modulation!

The paper is organized as follows. We describe our setup in Section 2. We discuss the problem of transmission maximization using phase-only modulated wavefronts in Section 3. We describe physically realizable, non-iterative and iterative algorithms for transmission maximization in Section 4 and in Section 6, respectively. We identify fundamental limits of phase-only modulated transmission in Section 5, validate the predictions and the rapid convergence behavior of the iterative algorithms in Section 7, and summarize our findings in Section 8.

2. Setup

We study scattering from a two-dimensional (2D) periodic 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. 2) and 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. All fields are TM\(_z\) polarized: electric fields in the \( y < 0 \) (\( i = 1 \)) and \( y > L \) (\( i = 2 \)) halfspaces are denoted \( e_i(\rho) = e_i(\rho) \hat{z} \). These fields (complex) amplitudes \( e_i(\rho) \) can be decomposed in terms of \(+y\) and \(-y\) propagating waves as \( e_i(\rho) = e_i^+(\rho) + e_i^-(\rho) \), where

\[
e_i^{\pm}(\rho) = \sum_{n=-N}^{N} h_n a_i^{\pm,n} e^{-jk_n^{\pm} \rho}.
\]

In the above expression, \( \rho = x \hat{x} + y \hat{y} \equiv (x, y) \), \( k_n^{\pm} = k_{n,x} \hat{x} \pm k_{n,y} \hat{y} \equiv (k_{n,x}, \pm k_{n,y}) \), \( k_{n,x} = 2\pi n/D \), \( k_{n,y} = 2\pi \sqrt{(1/\lambda)^2 - (n/D)^2} \), \( \lambda \) is the wavelength, and \( h_n = \sqrt{\|k_n^{\pm}\|_2 / k_{n,y}} \) 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 \( a_i^{\pm,n}, i = 1, 2; n = -N, \ldots, N \) are related by the scattering matrix

\[
\begin{bmatrix}
a_1^- \\
a_2^+
\end{bmatrix}
= S
\begin{bmatrix}
a_1^+ \\
a_2^-
\end{bmatrix},
\]

where \( a_\pm = [a_{i,-N}^{\pm} \ldots a_{i,0}^{\pm} \ldots a_{i,N}^{\pm}]^T \) and \( T \) denotes transposition. In what follows, we assume that the slab is only excited from the \( y < 0 \) halfspace; hence, \( a_2^- = 0 \). For a given incident field amplitude \( e_1^+(\rho) \), we define transmission and reflection coefficients as

\[
\tau(a_1^+) := \frac{\|S_{21} \cdot a_1^+\|_2^2}{\|a_1^+\|_2^2},
\]

and

\[
\Gamma(a_1^+) := \frac{\|S_{11} \cdot a_1^+\|_2^2}{\|a_1^+\|_2^2},
\]

respectively. We denote the transmission coefficient of a normally incident wavefront by \( \tau_{normal} = \tau([0 \ldots 0 1 0 \ldots 0]^T) \).

3. Problem formulation

We define the phase-vector of the modal coefficient vector \( a_1^+ \), as

\[
\phi(a_1^+) = [\phi(a_{1,-N}^+) \ldots \phi(a_{1,0}^+) \ldots \phi(a_{1,-N}^+)]^T,
\]

where for \( n = -N, \ldots, N, a_{1,n}^+ = |a_{1,n}^+| \exp(j/\phi(a_{1,n}^+)) \) and \( |a_{1,n}^+| \) and \( \phi(a_{1,n}^+) \) denote the magnitude and phase of \( a_{1,n}^+ \), respectively. For a real-valued constant \( c > 0 \), let \( P_c^M \) denote vectors of
the form
\[ p(\theta; c) = \sqrt{\frac{c}{M}} \begin{bmatrix} e^{j\theta_{-N}} & \cdots & e^{j\theta_0} & \cdots & e^{j\theta_N} \end{bmatrix}^T, \]  
(5)
where \( \theta = [\theta_{-N} \cdots \theta_0 \cdots \theta_N]^T \) is a \( 2N + 1 =: M \)-vector of phases. Then, the problem of designing a phase-only modulated incident wavefront that maximizes the transmitted power can be stated as

\[ a_{\text{opt}} = \arg \max_{a^+ \in \mathbb{P}_c} \tau(a^+) = \arg \max_{a^+ \in \mathbb{P}_c} \frac{\| S_{21} \cdot a^+ \|^2}{\| a^+ \|^2} = \arg \max_{a^+ \in \mathbb{P}_c} \frac{\| S_{21} \cdot a^+_1 \|^2}{\| a^+_1 \|^2}. \]  
(6)

Henceforth, let \( p(\theta) := p(\theta; 1) \) denote the setting where \( c = 1 \) in Eq. (5). Consider the optimization problem

\[ \theta_{\text{opt}} = \arg \max_{\theta} \| S_{21} \cdot p(\theta) \|^2. \]  
(7)

Then, from Eq. (6), the optimal wavefront is given by

\[ a_{\text{opt}} = p(\theta_{\text{opt}}). \]  
(8)

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. (7) can be reformulated as

\[ \theta_{\text{opt}} = \arg \max_{\theta} \left( \frac{(p(\theta))^H \cdot S_{11}^H \cdot S_{21} \cdot p(\theta)}{\| p(\theta) \|^2} \right) = \arg \min_{\theta} \| S_{11} \cdot p(\theta) \|^2 = \arg \min_{\theta} \Gamma(p(\theta)). \]  
(9)

Thus the phase-only modulated wavefront that maximizes transmission will also minimize backscatter. The phase-only modulating constraint leads to non-convex cost functions in Eqs. (7) and (9) for which there is no closed-form solution for \( \theta_{\text{opt}} \) or \( a_{\text{opt}} \).

4. Non-iterative, phase-only modulating algorithms for transmission maximization

We first consider algorithms for increasing transmission by backscatter minimization using phase-only modulated wavefronts that utilize measurements of the reflection matrix \( S_{11} \). We assume that this matrix can be measured using the experimental techniques described in [15,16,21,23] by, in essence, transmitting \( K > M \) incident wavefronts \( \{a^+_{i,i}\}_{i=1}^K \), measuring the (modal decomposition of the) backscattered wavefronts \( \{a^-_{i,i}\}_{i=1}^K \), and estimating \( S_{11} \) by solving the system of equations \( \{a^-_{i,i} = S_{11} \cdot a^+_{i,i}\}_{i=1}^K \). We note that, even if the \( S_{11} \) matrix has been measured perfectly, the optimization problem

\[ a_{\text{opt}} = \arg \min_{a^+_i \in \mathbb{P}_1} \| S_{11} \cdot a^+_i \|^2, \]  
(10)
is computationally intractable. In the simplest setting where the elements of $\mathbf{a}^+$ are restricted to be $\pm 1/\sqrt{M}$ instead of continuous values, the optimization problem in (10) is closely related to the binary quadratic programming (BQP) problem which is known to be NP-hard [17].

We can make the problem computationally tractable by relaxing the phase-only constraint in Eq. (10) and allowing the elements of $\mathbf{a}^+$ to take on arbitrary amplitudes and phases while imposing the power constraint $\| \mathbf{a}^+ \|^2 = 1$. This yields the optimization problem

$$
\mathbf{a}_{\text{svd}} = \arg \min_{\| \mathbf{a}^+ \|^2 = 1} \| S_{11} \cdot \mathbf{a}^+ \|^2,
$$

where we have relaxed the difficult constraint $\mathbf{a}^+ \in P_1^M$ into the spherical constraint $\| \mathbf{a}^+ \|^2 = 1$. Although the original unrelaxed backscatter minimization problem in Eq. (9) is hard to solve, the relaxed problem in Eq. (11) is much easier and can be solved exactly.

Let $S_{21} = \sum_{i=1}^{M} \sigma_i \mathbf{u}_i \cdot \mathbf{v}_i^H$ and $S_{11} = \sum_{i=1}^{M} \tilde{\sigma}_i \tilde{\mathbf{u}}_i \cdot \tilde{\mathbf{v}}_i^H$ denote the singular value decompositions (SVD) of $S_{21}$ and $S_{11}$, respectively. Here $\sigma_i$ (resp. $\tilde{\sigma}_i$) is the singular value associated with the left and right singular vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ (resp. $\tilde{\mathbf{u}}_i$ and $\tilde{\mathbf{v}}_i$), respectively. By convention, the singular values are arranged so that $\sigma_1 \geq \ldots \geq \sigma_M$ and $\tilde{\sigma}_1 \geq \ldots \geq \tilde{\sigma}_M$ and $^H$ denotes the complex conjugate transpose. In the lossless setting we have that $S_{11}^H \cdot S_{11} + S_{21}^H \cdot S_{21} = I$ so that $\mathbf{v}_i = \tilde{\mathbf{v}}_{M-i+1}$. Then, a well-known result in matrix analysis [10] states that

$$
\mathbf{a}_{\text{svd}} = \tilde{\mathbf{v}}_1.
$$

(12)

This is an exact solution to the relaxed backscatter minimization problem in Eq. (11).

To get an approximation of the solution to the original unrelaxed problem in Eq. (10) we construct a highly-transmitting wavefront as

$$
\mathbf{a}_{\text{opt,svd}} = p(\mathbf{a}_{\text{svd}}).
$$

(13)

Note that $\mathbf{a}_{\text{opt,svd}}$ given by Eq. (13) is an approximation to the solution of Eq. (10). It is not guaranteed to be the phase-only modulated wavefront that yields the highest transmission. It does, however, provide a lower bound on the amount of transmission that can be achieved using phase-only modulated wavefronts. As we shall see in Section 7, it produces highly transmitting wavefronts for the scattering systems considered here.

The spherical relaxation that yields the optimization problem in Eq. (11) includes all the phase-only wavefronts in the original problem, but also includes many other wavefronts as well. We now consider a ‘tighter’ relaxation that includes all the phase-only wavefronts in the original problem but fewer other wavefronts than the spherical relaxation does.

We begin by examining the objective function on the right hand side of Eq. (11). Note that

$$
\| S_{11} \cdot \mathbf{a}^+ \|^2 = \| (\mathbf{a}^+)^H \cdot S_{11}^H \cdot S_{11} \cdot \mathbf{a}^+ \| = \text{Tr} \left( S_{11}^H \cdot S_{11} \cdot \mathbf{a}^+ \cdot (\mathbf{a}^+)^H \right),
$$

(14)
where $\text{Tr}(\cdot)$ denotes the trace of its matrix argument. Let us define a new matrix-valued variable $A = \mathbf{a}_i^+ \cdot (\mathbf{a}_i^+)^H$. We note that $A$ is a Hermitian, positive semi-definite matrix with rank 1 and $A_{ii} = 1/M$ whenever $\mathbf{a}_i^+ \in P_1$, where $A_{ii}$ denotes the $i$th diagonal element of the matrix $A$. Consequently, from Eq. (14), we can derive the modified optimization problem

$$A_{\text{opt}} = \arg \min_{A \in \mathbb{C}^{M \times M}} \text{Tr} \left( S_{11}^H \cdot S_{11} \cdot A \right)$$

subject to $A = A^H, A \succeq 0, \text{rank}(A) = 1$ and $A_{ii} = 1/M$ for $i = 1, \ldots, M$, 

where the conditions $A = A^H$ and $A \succeq 0$ imply that $A$ is a Hermitian, positive semi-definite matrix. If we can solve Eq. (15) exactly, then by construction, since $A_{\text{opt}}$ is rank 1, we must have that $A_{\text{opt}} = \mathbf{a}_{\text{opt.eig}}^+ \cdot (\mathbf{a}_{\text{opt.eig}}^+)^H$ with $\mathbf{a}_{\text{opt.eig}}^+ \in P_1$ so we would have solved Eq. (10) exactly. Alas, the rank constraint in Eq. (15) makes the problem computationally intractable.

Eliminating the difficult rank constraint yields the semi-definite programming (SDP) problem [31]

$$A_{\text{sdp}} = \arg \min_{A \in \mathbb{C}^{M \times M}} \text{Tr} \left( S_{11}^H \cdot S_{11} \cdot A \right)$$

subject to $A = A^H, A \succeq 0, \text{and } A_{ii} = 1/M$ for $i = 1, \ldots, M$, 

which can be efficiently solved in polynomial-time [17] using off-the shelf solvers such as CVX [9,11] or SDPT3 [28]. See Appendix A for details.

We note that $A_{\text{sdp}}$ is the solution to the relaxed backscatter minimization problem in Eq. (16). If $A_{\text{sdp}}$ thus obtained has rank 1 then we will have solved the original unrelaxed problem in Eq. (10) exactly as well. Typically, however, the matrix $A_{\text{sdp}}$ will not be rank one so we describe a procedure next for obtaining an approximation to the original unrelaxed problem in Eq. (10).

Let $A_{\text{sdp}} = \sum_{i=1}^M \lambda_i \mathbf{u}_{i,\text{sdp}} \cdot (\mathbf{u}_{i,\text{sdp}})^H$ denote the eigenvalue decomposition of $A_{\text{sdp}}$ with the eigenvalues arranged so that $\lambda_1 \geq \ldots \lambda_M \geq 0$. Then we can construct a highly-transmitting phase modulated wavefront as

$$\mathbf{a}_{\text{opt.sdp}} = \mathbf{p} (\| \mathbf{u}_1 \|_{\text{sdp}}).$$

Note that $\mathbf{a}_{\text{opt.sdp}}$ given by Eq. (17) is an approximation to the solution of Eq. (10). It is not guaranteed to be the phase-only modulated wavefront that yields the highest transmission. It does however provide a lower bound on the amount of transmission that can be achieved. Since the SDP relaxation is a tighter relaxation than the spherical relaxation [17], we expect $\mathbf{a}_{\text{opt.sdp}}$ to result in higher transmission than $\mathbf{a}_{\text{opt.svd}}$.

We note that the computational cost of solving Eq. (16) and obtaining $A_{\text{sdp}}$ is $O(M^{4.5})$ [17] while the computational cost for obtaining $\mathbf{a}_{\text{opt.svd}}$ using the Lanczos method for computing only the leading singular vector is $O(M^2)$ [8]. Thus when $M > 1000$, there is a significant extra computational burden in obtaining the SDP solution. Hence, the question of when the extra computational burden of solving the SDP relaxation yields ‘large enough’ gains relative
to the spherical relaxation is of interest. We provide an answer using extensive numerical simulations in Section 7.

We have described two non-iterative techniques for increasing transmission via backscatter analysis that first require the $S_{11}$ to be measured and then compute $a_{\text{opt}, \text{svd}}$ or $a_{\text{opt}, \text{sdp}}$ using Eq. (13) and Eq. (17), respectively. We now provide a theoretical analysis of the transmission power we can expect to achieve using these phase-only modulated wavefronts.

5. Theoretical limit of phase-only modulated light transmission

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

The theoretical distribution [2,3,7,19,20] of the transmission coefficients for lossless random media (referred to as the DMPK distribution) has density given by

$$f(\tau) = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^{M} \delta(\tau - \tau(v_i)) = \frac{l}{2L} \frac{1}{\tau \sqrt{1 - \tau}}, \quad \text{for } 4 \exp(-L/2l) \lesssim \tau \leq 1. \quad (18)$$

In Eq. (18), $l$ is the mean-free path through the medium. This implies that in the regime where the DMPK distribution is valid, we expect $\tau(a_{\text{opt}}) \approx 1$ so that (near) perfect transmission is possible using amplitude and phase modulation. We now analyze the theoretical limit of phase-only modulation in the setting where the $S_{21}$ (or $S_{11}$) matrix has been measured and we have computed $a_{\text{opt}, \text{svd}}$ or $a_{\text{opt}, \text{sdp}}$ as in Eq. (13) and Eq. (17), respectively. In what follows, we prove a lower bound on the transmission we expect to achieve in the regime where the DMPK distribution is valid.

We begin by considering the wavefront $a_{\text{opt}, \text{svd}}$ which yields a transmission power given by

$$\tau(a_{\text{opt}, \text{svd}}) = \tau(p(\tilde{a}_{\text{svd}}) = \|S_{21} \cdot p(\tilde{a}_{\text{svd}})\|_2^2$$

$$= \|U \cdot \Sigma \cdot V^H \cdot p(\tilde{a}_{\text{svd}})\|_2^2 = \|\Sigma \cdot V^H \cdot p(\tilde{a}_{\text{svd}})\|_2^2. \quad (20)$$

Define $\tilde{p}(\tilde{a}_{\text{svd}}) = V^H \cdot p(\tilde{a}_{\text{svd}})$. Then from Eq. (20), we have that

$$\tau(a_{\text{opt}, \text{svd}}) = \|\Sigma \cdot \tilde{p}(\tilde{a}_{\text{svd}})\|_2^2 \quad (21)$$

$$= \sum_{i=1}^{M} \sigma_i^2 |\tilde{p}_i(\tilde{a}_{\text{svd}})|^2 \geq \sigma_1^2 |\tilde{p}_1(\tilde{a}_{\text{svd}})|^2. \quad (22)$$

In the DMPK regime, we have that $\sigma_1^2 \approx 1$ from which we can deduce that

$$\tau(a_{\text{opt}, \text{svd}}) \gtrsim |\tilde{p}_1(\tilde{a}_{\text{svd}})|^2 \quad (23)$$
From Eq. (12), we have that \( a_{svd} = v_1 = \tilde{v}_M \) so that if
\[
\tilde{v}_1^H = \begin{bmatrix} |v_{1,1}| e^{-j/\sqrt{v_{1,1}}} & \ldots & |v_{1,M}| e^{-j/\sqrt{v_{1,M}}} \end{bmatrix},
\]
then
\[
\tilde{p}_1(a_{svd}) = \tilde{v}_1^H \cdot p(v_1) = \frac{1}{\sqrt{M}} \sum_{i=1}^{M} |v_{1,i}|,
\]
and
\[
|\tilde{p}_1(a_{svd})|^2 = \frac{1}{M} \sum_{i=1}^{M} |v_{1,i}|^2 + \frac{2}{M} \sum_{i<j} |v_{1,i}| \cdot |v_{1,j}|.
\]
Substituting Eq. (25) into Eq. (23) gives
\[
\tau(a_{opt,svd}) \gtrsim \frac{1}{M} \sum_{i=1}^{M} |v_{1,i}|^2 + \frac{2}{M} \sum_{i<j} |v_{1,i}| \cdot |v_{1,j}|.
\]
Taking expectations on both sides of Eq. (26) and invoking the linearity of the expectation operator gives us
\[
\mathbb{E} \left[ \tau(a_{opt,svd}) \right] \gtrsim \frac{1}{M} \sum_{i=1}^{M} \mathbb{E} \left[ |v_{1,i}|^2 \right] + \frac{2}{M} \sum_{i<j} \mathbb{E} \left[ |v_{1,i}| \cdot |v_{1,j}| \right].
\]
We now invoke the maximum-entropy principle as in Pendry's derivation [2, 20] and assume that the vector \( v_1 \), is uniformly distributed on the unit hypersphere. Since the uniform distribution is symmetric, for any indices \( i \) and \( j \), we have that \( \mathbb{E} \left[ |v_{1,i}|^2 \right] = \mathbb{E} \left[ |v_{1,1}|^2 \right] \) and \( \mathbb{E} \left[ |v_{1,i}| \cdot |v_{1,j}| \right] = \mathbb{E} \left[ |v_{1,1}| \cdot |v_{1,2}| \right] \). Consequently Eq. (27) simplifies to
\[
\mathbb{E} \left[ \tau(a_{opt,svd}) \right] \gtrsim \frac{1}{M} \sum_{i=1}^{M} \mathbb{E} \left[ |v_{1,i}|^2 \right] + \frac{2M(M - 1)}{2M} \mathbb{E} \left[ |v_{1,1}| \cdot |v_{1,2}| \right],
\]
Since \( \|v_1\|_2^2 = \sum_{i=1}^{M} |v_{1,i}|^2 = 1 \), we have that
\[
\mathbb{E} \left[ |v_{1,1}|^2 \right] = O \left( \frac{1}{M} \right).
\]
Substituting Eq. (29) into Eq. (28) gives
\[
\mathbb{E} \left[ \tau(a_{opt,svd}) \right] \gtrsim (M - 1) \mathbb{E} \left[ |v_{1,1}| \cdot |v_{1,2}| \right] + O \left( \frac{1}{M} \right).
\]
We now note that
\[
\mathbb{E} \left[ |v_{1,1}| \cdot |v_{1,2}| \right] = \mathbb{E} \left[ |v_{1,1}| \right] \cdot \mathbb{E} \left[ |v_{1,2}| \right] + \text{cov} \left( |v_{1,1}|, |v_{1,2}| \right),
\]
\[
= \mathbb{E}^2 \left[ |v_{1,1}| \right] + \text{cov} \left( |v_{1,1}|, |v_{1,2}| \right),
\]
\[
= \mathbb{E}^2 \left[ |v_{1,1}| \right].
\]
where

$$\text{cov} (|v_{1,1}|, |v_{1,2}|) = \mathbb{E} \left[ \{ |v_{1,1}| - \mathbb{E}[|v_{1,1}|] \} \cdot \mathbb{E} \{ |v_{1,2}| - \mathbb{E}[|v_{1,2}|] \} \right],$$  \hspace{1cm} (33)

is the covariance between the random variables $|v_{1,1}|$ and $|v_{1,2}|$. A useful fact that will facilitate analytical progress is that the complex-valued random variable $v_{1,1}$ has the same distribution [22, Chap. 3a] as the vector

$$g_1 \sqrt{|g_1|^2 + \ldots + |g_M|^2}$$

where $g_i = x_i + \sqrt{-1} y_i$ and $x_i$ and $y_i$ are i.i.d. normally distributed variables with mean zero and variance $1/(2M)$. This implies that the variable $|v_{1,1}|^2$ is beta distributed since $|g_1|^2$ and $|g_1|^2 + \ldots + |g_M|^2$ are chi-square distributed. Hence, it can be easily seen that

$$\text{cov} (|v_{1,1}|, |v_{1,2}|) = O \left( \frac{1}{M^2} \right)$$  \hspace{1cm} (34)

and

$$\mathbb{E} [v_{1,1}] = \sqrt{\frac{\pi}{4M}} + O \left( \frac{1}{M} \right),$$  \hspace{1cm} (35)

where the first term on the righthand side of Eq. (35) equals $\mathbb{E}[|g_i|]$. Substituting Eq. (34) and Eq. (35) into Eq. (32) gives us an expression for $\mathbb{E}[|v_{1,1}| \cdot |v_{1,2}|]$, which on substituting into the right-hand side of Eq. (30) yields the inequality

$$\mathbb{E}[\tau(a_{\text{opt, sdp}})] \gtrsim \frac{\pi}{4} + O \left( \frac{1}{\sqrt{M}} \right).$$  \hspace{1cm} (36)

Since $\tau(a_{\text{opt, sdp}}) \geq \tau(a_{\text{opt, svd}})$, Eq. (36) yields the inequality

$$\mathbb{E}[\tau(a_{\text{opt, sdp}})] \geq \mathbb{E}[\tau(a_{\text{opt, svd}})] \gtrsim \frac{\pi}{4} + O \left( \frac{1}{\sqrt{M}} \right).$$  \hspace{1cm} (37)

Letting $M \to \infty$ on both sides on Eq. (37) gives us

$$\lim_{M \to \infty} \mathbb{E}[\tau(a_{\text{opt, sdp}})] \geq \lim_{M \to \infty} \mathbb{E}[\tau(a_{\text{opt, svd}})] \gtrsim \frac{\pi}{4}. \hspace{1cm} (38)$$

From Eq. (38) we expect to achieve at least $25\%$ when the $S_{21}$ (or $S_{11}$) matrix has been measured and we compute the phase-only modulated wavefront using $a_{\text{opt, svd}}$ or $a_{\text{opt, sdp}}$. In contrast, amplitude and phase modulation yields (nearly) $100\%$ transmission; thus the phase-only modulation incurs an (average) loss of at most $22\%$.

We now develop rapidly-converging, physically-realizable, iterative algorithms for increasing transmission by backscatter minimization that utilize significantly fewer measurements than the $O(M)$ measurements it would take to first estimate $S_{11}$ and subsequently construct $a_{\text{opt, svd}}$ or $a_{\text{opt, sdp}}$. 

11
Algorithm 1 Steepest descent algorithm for finding \( \mathbf{a}_{\text{svd}} \)

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

6. Iterative, phase-only modulated algorithms for transmission maximization

6.A. Steepest Descent Method

We first consider an iterative method, based on the method of steepest descent, for finding the wavefront \( \mathbf{a}_{1}^+ \) that minimizes the objective function \( \| S_{11} \cdot \mathbf{a}_{1}^+ \|_2^2 \). At this stage, we consider arbitrary vectors \( \mathbf{a}_{1}^+ \) instead of phase-only modulated vectors \( \mathbf{a}_{1}^+ \in P_1^M \). The algorithm utilizes the negative gradient of the objective function to update the incident wavefront as

\[
\mathbf{a}_{1,(k)}^+ = \mathbf{a}_{1,(k)}^+ - \mu \frac{\partial \| S_{11} \cdot \mathbf{a}_{1}^+ \|_2^2}{\partial \mathbf{a}_{1}^+} \bigg|_{\mathbf{a}_{1}^+=\mathbf{a}_{1,(k)}^+}
\]

\[
= \mathbf{a}_{1,(k)}^+ - 2 \mu S_{11}^H \cdot S_{11} \cdot \mathbf{a}_{1,(k)}^+ ,
\]

where \( \mathbf{a}_{1,(k)}^+ \) represents the modal coefficient vector of the incident wavefront produced at the \( k \)-th iteration of the algorithm and \( \mu \) is a positive stepsize. If we renormalize \( \mathbf{a}_{1,(k)}^+ \) to have \( \| \mathbf{a}_{1,(k)}^+ \|_2 = 1 \), then we obtain Algorithm 1 which iteratively refines the wavefront \( \mathbf{a}_{1,(k+1)}^+ \). In the limit of \( k \rightarrow \infty \), the incident wavefront \( \mathbf{a}_{1,(k+1)}^+ \) will converge to \( \mathbf{a}_{\text{svd}} \).

We now describe how the update equation given by Eq. (40), which requires computation of the gradient \( S_{11}^H \cdot S_{11} \cdot \mathbf{a}_{1,(k)}^+ \), can be physically implemented even though we have not measured \( S_{11} \) apriori.

Let \( \text{flipud}(\cdot) \) 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. Let \( F = \text{flipud}(I) \) where \( I \) is the identity matrix, and let \( ^* \) denote complex conjugation. In recent work [14], we showed that reciprocity of the scattering system implies that

\[
S_{11}^H = F \cdot S_{11}^* \cdot F ,
\]

which can be exploited to make the gradient vector \( S_{11}^H \cdot S_{11} \cdot \mathbf{a}_{1,(k)}^+ \) physically measurable.
To that end, we note that Eq. (41) implies that

$$S_{11}^H \cdot a_{1}^- = F \cdot S_{11}^* \cdot a_{1}^- = F \cdot (S_{11} \cdot (F \cdot (a_{1}^-)^*))^*$,$$

where $a_{1}^- = S_{11} \cdot a_{1,1}^+$. Thus, we can physically measure $S_{11}^H \cdot S_{11} \cdot a_{1,1}^+$, by performing the following sequence of operations and the accompanying measurements:

1. Transmit $a_{1,1}^+$ and measure the backscattered wavefront $a_{1}^- = S_{11} \cdot a_{1,1}^+$.
2. Transmit the wavefront obtained by time-reversing the wavefront whose modal coefficient vector is $a_{1}^-$ or equivalently transmitting the wavefront $F \cdot (a_{1}^-)^*$.
3. Measure the resulting backscattered wavefront corresponding to $S_{11} \cdot (F \cdot (a_{1}^-)^*)$ and time-reverse it to yield the desired gradient vector $S_{11}^H \cdot S_{11} \cdot a_{1,1}^+$ as shown in Eq. (42).

The above represents a physically realizable scheme for measuring the gradient vector, which we proposed in our previous paper [14]. Since time-reversal can be implemented using phase-conjugating mirror, we referred to our algorithm a double phase-conjugating method.

For the setting considered here, we have the additional physically-motivated restriction that all transmitted wavefronts $a_{1}^+ \in P_1^M$. However, the wavefront $a_{1}^-$ can have arbitrary amplitudes and so will the wavefront obtained by time-reversing it (as in Step 2 above) thereby violating the phase-only modulating restriction and making Algorithm 1, physically unrealizable. This is also why algorithms of the sort considered by others in array processing e.g. [26] cannot be directly applied here.

This implies that even though Algorithm 1 probably converges to $a_{svd}$, it cannot be used to compute $a_{opt,svd}$ as in Eq. (13) because it is not physically implementable given the phase-only modulation constraint. To mitigate this problem, we propose modifying the update step in Eq. (40) to

$$\tilde{a}_{1,1}^+ = \mathcal{P} \left( a_{1,1}^+ - 2\mu \tilde{a} \cdot S_{11}^H \cdot \mathcal{P} \left( a_{1,1}^+ \right) \right),$$

where $\tilde{a}$ is chosen such that all magnitudes of modal coefficients of $\mathcal{P} \left( a_{1}^- \right)$ are set to the average magnitude of modal coefficients of $a_{1}^-$. Then, by applying Eq. (41) as before, we can physically measure $\tilde{a} \cdot S_{11}^H \cdot \mathcal{P} \left( a_{1,1}^+ \right)$ by performing the following sequence of operations and the accompanying measurements:

1. Transmit $a_{1,1}^+$ and measure the backscattered wavefront $a_{1}^- = S_{11} \cdot a_{1,1}^+$.
2. Compute the scalar $\tilde{a} = \frac{\sum_{n=-N}^{N} |a_{1,n}|}{\sqrt{M}}$.
3. Transmit the (phase-only modulated) wavefront obtained by time-reversing the wavefront whose modal coefficient vector is $\mathcal{P} \left( a_{1}^- \right)$.
4. Measure the resulting backscattered wavefront, time-reverse it, and scale it with \( \pi \) to yield the desired gradient vector.

This modified iteration in Eq. (43) leads to the algorithm in the left column of Table 1 and its physical counterpart in the right column of Table 1.

| Vector Operation | Physical Operation |
|------------------|--------------------|
| 1: \( a_i^- = S_{11} \cdot a_{i,(k)}^- \) | 1: \( a_i^{\pm,(k)} \rightarrow \text{Backscatter} \rightarrow a_i^- \) |
| 2: \( \bar{a} = \frac{\sum_{n=-N}^{N} |a_{1,n}^-|}{\sqrt{M}} \) | 2: \( \bar{a} = \frac{\sum_{n=-N}^{N} |a_{1,n}^-|}{\sqrt{M}} \) |
| 3: \( a_i^- \leftarrow p/\sqrt{a_i^-} \) | 3: \( a_i^- \leftarrow p/\sqrt{a_i^-} \) |
| 4: \( a_i^+ = F \cdot (a_i^-)^* \) | 4: \( a_i \rightarrow \text{PCM} \rightarrow a_i^+ \) |
| 5: \( a_i^- = S_{11} \cdot a_i^+ \) | 5: \( a_i^+ \rightarrow \text{Backscatter} \rightarrow a_i^- \) |
| 6: \( a_i^+ = F \cdot (a_i^-)^* \) | 6: \( a_i^- \rightarrow \text{PCM} \rightarrow a_i^+ \) |
| 7: \( a_i^+ = a_{1,(k)}^+ - 2\mu a_i^+ \) | 7: \( a_i^+ = a_{1,(k)}^+ - 2\mu a_i^+ \) |
| 8: \( a_{1,(k+1)}^+ = p/a_i^+ \) | 8: \( a_{1,(k+1)}^+ = p/a_i^+ \) |

Table 1. Steepest descent algorithm for refining a highly transmitting phase-only modulated wavefront. The first column represents vector operations. The second column represents the physical (or experimental) counterpart. The operation \( a_i^- \rightarrow F \cdot (a_i^-)^* \) can be realized via the use of a phase-conjugating mirror (PCM). The algorithm terminates when the backscatter intensity falls below a preset threshold \( \epsilon \).

6.B. Gradient Method

The wavefront updating step for the algorithm described in Table 1 first updates both the amplitude and phase of the incident wavefront (in Step 7) and then ‘projects it’ onto the set of phase-only modulated wavefronts (in Step 8). We now develop a gradient-based method that only updates the phase of the incident wavefront. From Eq. (9), the objective function of interest is \( \|S_{11} \cdot p(\theta)\|^2 \) which depends on the phase-only modulated wavefront. The algorithm utilizes the negative gradient of the objective function with respect to the phase vector to...
update the phase vector of the incident wavefront as

$$\theta^+(k+1) = \theta^+(k) - \sqrt{M}\mu \frac{\partial \|S_{11} \cdot p(\theta)\|_2^2}{\partial \theta_{\theta^+(k)}}$$

where $\theta^+(k)$ represents the phase vector of the wavefront produced at the $k$-th iteration of the algorithm and $\mu$ is a positive stepsize. We have separated the $\sqrt{M}$ factor from the stepsize so that $\mu$ can be $O(1)$ and independent of $M$. In Appendix B, we show that

$$\frac{\partial \|S_{11} \cdot p(\theta)\|_2^2}{\partial \theta_{\theta^+(k)}} = 2\text{Im} \left[ \text{diag} \{ p(-\theta^+(k)) \} \cdot S^H_{11} \cdot S_{11} \cdot p(\theta^+(k)) \right],$$

where $\text{diag} \{ p(-\theta^+(k)) \}$ denotes a diagonal matrix with entries $p(-\theta^+(k))$ along its diagonal. Substituting Eq. (45) into the right-hand side of Eq. (44) yields the iteration

$$\theta^+(k+1) = \theta^+(k) - 2\sqrt{M}\mu \text{Im} \left[ \text{diag} \{ p(-\theta^+(k)) \} \cdot S^H_{11} \cdot S_{11} \cdot p(\theta^+(k)) \right].$$

To evaluate the update Eq. (46), it is necessary to measure the gradient vector $S^H_{11} \cdot S_{11} \cdot p(\theta^+(k))$. For the same reason as in the steepest descent scheme, we cannot use double-phase conjugation introduced in our previous paper because of the phase-only modulating restriction. Therefore, we propose modifying the update step in Eq. (46) to

$$\theta^+(k+1) = \theta^+(k) - 2\sqrt{M}\mu \text{Im} \left[ \text{diag} \{ p(-\theta^+(k)) \} \cdot S^H_{11} \cdot p(S_{11} \cdot p(\theta^+(k))) \right],$$

and we use the modified double-phase conjugation as

1. Transmit $p(\theta^+(k))$ and measure the backscattered wavefront $a^- = S_{11} \cdot p(\theta^+(k))$;

2. Compute the scalar $\bar{a} = \frac{\sum_{n=-N}^N |a^-_{1,n}|}{\sqrt{M}}$;

3. Transmit the phase-only modulated wavefront obtained by time-reversing the wavefront whose modal coefficient vector is $p(S_{11} \cdot p(\theta^+(k)))$;

4. Measure the resulting backscattered wavefront, time-reverse it, and scale it with $\bar{a}$ to yield the desired gradient vector.

The phase-updating iteration in Eq. (47) leads to the algorithm in the left column of Table 2 and its physical counterpart in the right column of Table 1.
| Vector Operation | Physical Operation |
|------------------|--------------------|
| 1: \( a_1^- = S_{11} \cdot p(\theta_{1,(k)}^+) \) | 1: \( p(\theta_{1,(k)}^+) \xrightarrow{\text{Backscatter}} a_1^- \) |
| 2: \( \overline{a} = \frac{\sum_{n=-N}^{N} |a_{1,n}^-|}{\sqrt{M}} \) | 2: \( \overline{a} = \frac{\sum_{n=-N}^{N} |a_{1,n}^-|}{\sqrt{M}} \) |
| 3: \( a_1^- \leftarrow p(\overline{a_1^-}) \) | 3: \( a_1^- \leftarrow p(\overline{a_1^-}) \) |
| 4: \( a_1^+ = F \cdot (a_1^-)^* \) | 4: \( a_1^- \xrightarrow{\text{PCM, Backscatter}} a_1^+ \) |
| 5: \( a_1^- = S_{11} \cdot a_1^+ \) | 5: \( a_1^+ \xrightarrow{\text{PCM}} a_1^- \) |
| 6: \( a_1^+ = F \cdot (a_1^-)^* \) | 6: \( a_1^- \xrightarrow{\text{PCM}} a_1^+ \) |
| 7: \( \theta_{1,(k+1)}^+ = \theta_{1,(k)}^+ - 2\sqrt{M \mu \sigma} \text{Im} \left[ \text{diag} \left\{ p(-\theta_{1,(k)}^+) \right\} \cdot a_1^+ \right] \) |

Table 2. Gradient algorithm for transmission maximization. The first column contains the updating iteration in Eq. (47) split into a series of individual updates so that they may be mapped into their physical (or experimental) counterparts in the column to their right. The operation \( a_1^- \xrightarrow{\text{PCM}} F \cdot (a_1^-)^* \) can be realized via the use of a phase-conjugating mirror (PCM). The algorithm terminates when the backscatter intensity falls below a preset threshold \( \epsilon \).

7. Numerical simulations

To validate the proposed algorithms and the theoretical limits of phase-only wavefront optimization, we adopt the numerical simulation protocol described in [14]. Specifically, 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 series of azimuthal harmonics. As in [14], interactions between cylinders are modeled using 2D periodic Green’s functions. The method constitutes a generalization of that in [18], 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. As in [14], all periodic Green’s functions/lattice sums are rapidly evaluated using a recursive Shank’s transform using the methods described in [24, 25]. Our method exhibits exponential convergence in the number of azimuthal harmonics used in the description of the field scattered by each cylinder. As in [14], in the numerical experiments below, care was taken to ensure 11-digit accuracy in the entries of the computed scattering matrices.

First we compare the transmission power achieved by the non-iterative algorithms that
utilize measurements of the $S_{11}$ matrix to compute the wavefronts $a_{\text{opt,svd}}$ and $a_{\text{opt,sdp}}$ given by Eq. (13) and Eq. (17), respectively. Here we have a scattering system with $D = 197\lambda$, $r = 0.11\lambda$, $N_e = 430,000$, $n_d = 1.3$, $M = 395$ and $\bar{l} = 6.69\lambda$, where $\bar{l}$ is the average distance to the nearest scatterer. Fig. 3 plots transmitted power for the SVD and SDP based algorithms as a function of the thickness $L/\lambda$ of the scattering system.

As expected, the wavefront $a_{\text{opt,sdp}}$ realizes increased transmission relative to the wavefront $a_{\text{opt,svd}}$. However, as the thickness of the medium increases, the gain vanishes. Typically $a_{\text{opt,sdp}}$ increases transmission by about $1 - 5\%$ relative to $a_{\text{opt,svd}}$. Fig. 3 also shows the accuracy of our theoretical prediction of $25\pi\% \approx 78.5\%$ transmission using phase-only modulation for highly backscattering (or thick) random media in the same regime where the DMPK theory predicts perfect transmission using amplitude and phase modulated wavefronts.

Fig. 3 also plots the transmitted power achieved by an ‘equal phase’ wavefront with a modal coefficient vector $1/\sqrt{M} \begin{bmatrix} 1 & \ldots & 1 \end{bmatrix}^T$. Both the SVD and the SDP based algorithms realized significant gains relative to this vector.  

Recall that the computational cost of computing $a_{\text{opt,sdp}}$ is $O(M^{1.5})$ while the cost for computing $a_{\text{opt,svd}}$ is $O(M^2)$. Fig. 3 suggests that for large $M$, the significantly extra computational effort for computing $a_{\text{opt,sdp}}$ might not be worth the effort for strongly scattering random media.

We also plot the transmitted power achieved by undersampling the number of control modes by a factor of 4, computing the resulting $S_{21}$ matrix, and constructing the amplitude and phase modulated eigen-wavefront associated with the largest right singular vector. This is what would happen if we were to implement the ‘superpixel’-based amplitude and phase modulation scheme described in [30] in the framework of a system with periodic boundary conditions. As can be seen, phase-only modulation yields higher transmission than amplitude and phase modulation with undersampled modes. We are presently studying whether the same result holds true in systems without periodic boundary conditions as considered in [4].

Fig. 4 compares the rate of convergence of the phase-only modulated steepest descent (with $\mu = 0.6574$) and gradient descent (with $\mu = 1.4149$)-based algorithms and the rate of convergence of the amplitude and phase-only modulated steepest descent (with $\mu = 0.5059$) based algorithm from [14, Algorithm 1]. Here we are in a setting with $D = 197\lambda$, $L = 3.4 \times 10^5\lambda$, $r = 0.11\lambda$, $N_e = 430,000$ dielectric cylinders with $n_d = 1.3$, $M = 395$, $\bar{l} = 6.69\lambda$. In this setting, a normally incident wavefront results in a transmission of $\tau_{\text{normal}} = 0.038$.

The wavefront $a_{\text{svd}}$ yields $\tau_{\text{opt}} = 0.9973$ corresponding to a 26-fold increase in transmission. The amplitude and phase modulated steepest descent algorithm produces a wavefront that converges to 95\% of the near optimum in about $5 - 10$ iterations as shown in Fig. 4.

---

1A normally incident wavefront also yields about the same transmitted power. Note that a normally incident wavefront cannot be synthesized using phase-only modulation using the setup in Fig. 1.
phase-only modulated steepest descent algorithm yields an 19-fold increase in transmission and converges within 5 – 10 iterations. The phase-only modulated gradient descent algorithm yields a 13-fold increase in transmission and converges in 15 – 20 iterations. The fast convergence properties of the steepest descent based method make it suitable for use in an experimental setting where it might be infeasible to measure the $S_{11}$ matrix first.

Fig. 5 compares the maximum transmitted power achieved after 50 iterations as a function of thickness $L/\lambda$ for the iterative, phase-only modulated steepest descent and gradient descent methods and the non-iterative SVD and SDP methods. The non-iterative methods increase transmission by 8.3% relative to the steepest descent method. The gradient descent method performs poorly relative to the steepest descent method but still achieves increased transmission relative to the non-adaptive ‘equal-phase’ wavefront.

We next investigate the choice of stepsize $\mu$ on the performance of the algorithms. Fig. 4 shows the performance with the optimal $\mu$ for the phase-only modulated steepest descent and gradient descent algorithms. The optimal $\mu$ was obtained by a line search, i.e., by running the algorithms over a fixed set of discretized values of $\mu$ between 0 and $\mu_{\text{max}}$, and choosing the $\mu$ that produces the fastest convergence. In an experimental setting, the line search for finding the optimal $\mu$ for the steepest descent algorithm could require additional measurements. Fig. 6 plots the transmitted power as a function of the number of iterations and the stepsize $\mu$ for the phase-only modulated steepest descent algorithm. This plot reveals that there is a broad range of $\mu$ for which the converges in a handful of iterations. We have found that setting $\mu \approx 0.65$ yields fast convergence about 15 – 20 iterations under a broad range of conditions. Fig. 7 shows the transmitted power achieved after 50 iterations of the phase-only modulated steepest descent algorithm as a function of the stepsize $\mu$ and the thickness $L/\lambda$ of the scattering system. There is a wide range of allowed values for $\mu$ where the steepest descent algorithm performs well. Fig. 8 plots the transmitted power after 50 iterations of the phase-only modulated gradient descent algorithm a function of the stepsize $\mu$ and the thickness $L/\lambda$ of the scattering system. In contrast to the steepest descent algorithm, the performance of the gradient descent algorithm is much more erratic. A $\mu$ of about 1.1 is a good choice for the gradient descent based method.

Finally, Fig. 9 plots the average number of iterations required to reach 95% of the respective optimas for the phase-only modulated steepest descent and gradient descent algorithms as a function of the thickness $L/\lambda$ of the scattering system. On average the steepest descent algorithm converges in about in about 15 – 20 iterations while the gradient descent algorithm converges in about 35 – 45 iterations. Here, we selected the optimal $\mu$‘s for the steepest descent algorithm and for the gradient descent algorithm for each depth in the medium.

Since the steepest descent algorithm converges faster and realizes 15 – 20% greater transmitted power, but only loses 10% transmission relative to the non-iterative phase-only
modulated SVD and SDP algorithms, it is the best option for use in an experimental setting.

8. Conclusions

We have shown theoretically and using numerically rigorous simulation that non-iterative, phase-only modulated techniques for transmission maximization using backscatter analysis can expect to achieve about $25\pi\% \approx 78.5\%$ transmission in highly backscattering random media in the DMPK regime where amplitude and phase modulated can yield 100% transmission. We have developed two new, iterative and physically realizable algorithms for constructing highly transmitting phase-only modulated wavefronts using backscatter analysis. We showed using numerical simulations that the steepest descent variant outperforms the gradient descent variant and that the wavefront produced by the steepest descent algorithm achieves about 71% transmission while converging within 15–20 measurements. The development of iterative phase-only modulated algorithms that bridge the 10% transmission gap between the steepest descent algorithm presented here and the non-iterative SVD and SDP algorithms remains an important open problem.

The proposed algorithms are quite general and may be applied to scattering problems beyond the 2D setup described in the simulations. A detailed study, guided by the insights in [4], of the impact of periodic boundary conditions on the results obtained is also underway.

Acknowledgements

This work was partially supported by an ONR Young Investigator Award N000141110660, an NSF grant CCF-1116115, an AFOSR Young Investigator Award FA9550-12-1-0266 and an AFOSR DURIP Award FA9550-12-1-0016. We thank Jeff Fessler for suggesting the gradient descent method so that the advantage of the steepest descent based method could be properly showcased.

A. Solving Eq. (16) in MATLAB

Specifically, the solution to Eq. (16) can be computed in MATLAB using the CVX package by invoking the following sequence of commands:

```matlab
cvx_begin sdp
    variable A(M,M) hermitian
    minimize trace(S11'*S11*A)
    subject to
    A >= 0;
    diag(A) == ones(M,1)/M;

cvx_end
```
Asdp = A; % return optimum in variable Asdp

For settings where $M > 100$, we recommend using the SDPT3 solver. The solution to Eq. (16) can be computed in MATLAB using the SDPT3 package by invoking the following sequence of commands:

cost_function = S11'*S11;
e = ones(M,1); b = e/M;
num_params = M*(M-1)/2;
C{1} = cost_function;
A = cell(1,M); for j = 1:M, A{j} = sparse(j,j,1,M,M); end
blk{1,1} = 's'; blk{1,2} = M; Avec = svec(blk(1,:),A,1);
[obj,X,y,Z] = sqlp(blk,Avec,C,b);
Asdp = cell2mat(X); % return optimum in variable Asdp

B. Derivation of Eq. (45)

Here, we derive Eq. (45). For notational brevity, we replace $S_{11}$ with $B$, and denote $B$’s $m$th row and $n$th column element as $B_{mn}$. We will show that

$$
\frac{\partial \|B \cdot p(\theta)\|_2^2}{\partial \theta_k} = 2 \text{Im} \left[ \text{diag} \{ p(-\theta) \} \cdot B^H \cdot B \cdot p(\theta) \right].
$$

(A1)

To this end, note that the cost function can be expanded as

$$
\|B \cdot p(\theta)\|_2^2 = \sum_{n=1}^{M} |B_{nm} e^{j\theta_m}|^2
= \sum_{n=1}^{M} \sum_{m=1}^{M} |B_{nm}|^2 + 2 \sum_{n=1}^{M} \sum_{p>q} \text{Re} \left( B_{np} B_{nq}^* e^{j(\theta_p - \theta_q)} \right)
= \sum_{n=1}^{M} \sum_{m=1}^{M} |B_{nm}|^2 + 2 \sum_{n=1}^{M} \sum_{p>q} |B_{np}| |B_{nq}| \cos(\theta_p - \theta_q + \angle(B_{np}) - \angle(B_{nq})),
$$

(A2)

where Re($\cdot$) denotes the operator that returns the real part of the argument. Consequently, the derivative of the cost function with respect to the $k$th phase $\theta_k$ can be expressed as

$$
\frac{\partial \|B \cdot p(\theta)\|_2^2}{\partial \theta_k} = -2 \sum_{n=1}^{M} \sum_{q \neq k} \text{Im} \left[ B_{nk} B_{nq}^* e^{j(\theta_k - \theta_q)} \right]
= -2 \text{Im} \left[ e^{j\theta_k} \sum_{n=1}^{M} B_{nk} \sum_{q \neq k} B_{nq}^* e^{-j\theta_q} \right],
$$

(A3)
where $\text{Im}(\cdot)$ denotes the operator that returns the imaginary part of the argument.

Let $\mathbf{e}_k$ be the $k$-th elementary vector. We may rewrite Eq. (A4) as

$$
\frac{\partial \| B \cdot p(\theta) \|_2^2}{\partial \theta_k} = -2 \text{Im} \left[ e^{i\theta_k} \begin{bmatrix} B_{1k} & \cdots & B_{Mk} \end{bmatrix} \cdot B^* \cdot \left\{ I - \mathbf{e}_k \cdot \mathbf{e}_k^H \right\} \cdot p(\theta)^* \right],
$$

or, equivalently, as

$$
\frac{\partial \| B \cdot p(\theta) \|_2^2}{\partial \theta_k} = -2 \text{Im} \left[ e^{i\theta_k} \begin{bmatrix} B_{1k} & \cdots & B_{Mk} \end{bmatrix} \cdot B^* \cdot p(\theta)^* \right] - 2 \text{Im} \left[ \begin{bmatrix} B_{1k} & \cdots & B_{Mk} \end{bmatrix} \cdot B^* \cdot \mathbf{e}_k \right],
$$

(A6)

Stacking the elements into a vector yields the relation

$$
\frac{\partial \| B \cdot p(\theta) \|_2^2}{\partial \theta} = -2 \text{Im} \left[ \text{diag}\{p(\theta)\} \cdot B^T \cdot B^* \cdot p(\theta)^* \right],
$$

(A8)

or, equivalently, Eq. (A1).

**References**

1. 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. *Physical review letters*, 106(10):103901, 2011.

2. C. Barnes and J. B. Pendry. Multiple scattering of waves in random media: a transfer matrix approach. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 435(1893):185, 1991.

3. C. W. J. Beenakker. Applications of random matrix theory to condensed matter and optical physics. *Arxiv preprint arXiv:0904.1432*, 2009.

4. W. Choi, A. P. Mosk, Q.-H. Park, and W. Choi. Transmission eigenchannels in a disordered medium. *Physical Review B*, 83(13):134207, 2011.

5. M. Cui. A high speed wavefront determination method based on spatial frequency modulations for focusing light through random scattering media. *Optics Express*, 19(4):2989–2995, 2011.

6. M. Cui. Parallel wavefront optimization method for focusing light through random scattering media. *Optics letters*, 36(6):870–872, 2011.

7. O. N. Dorokhov. Transmission coefficient and the localization length of an electron in N bound disordered chains. *JETP Lett*, 36(7), 1982.

8. G. H. Golub and C. F. Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.
9. M. Grant and S. Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, Recent Advances in Learning and Control, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. http://stanford.edu/~boyd/graph_dcp.html.
10. R. A. Horn and C. R. Johnson. Matrix analysis. Cambridge university press, 1990.
11. CVX Research Inc. CVX: Matlab software for disciplined convex programming, version 2.0. http://cvxr.com/cvx, August 2012.
12. A. Ishimaru. Wave propagation and scattering in random media, volume 12. Wiley-IEEE Press, 1999.
13. 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 Statistical Signal Processing Workshop (SSP), 2012 IEEE, pages 97–100. IEEE, 2012.
14. C. Jin, R. R. Nadakuditi, E. Michielssen, and S. Rand. Iterative, backscatter-analysis algorithms for increasing transmission and focusing light through highly scattering random media. JOSA A, 30(8):1592–1602, 2013.
15. 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. Nature Photonics, 6(9):583–587, 2012.
16. T. W. Kohlgraf-Owens and A. Dogariu. Transmission matrices of random media: Means for spectral polarimetric measurements. Optics letters, 35(13):2236–2238, 2010.
17. Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang. Semidefinite relaxation of quadratic optimization problems. Signal Processing Magazine, IEEE, 27(3):20–34, 2010.
18. 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. Physical Review E, 60(6):7614, 1999.
19. P. A. Mello, P. Pereyra, and N. Kumar. Macroscopic approach to multichannel disordered conductors. Annals of Physics, 181(2):290–317, 1988.
20. J. B. Pendry, A. MacKinnon, and A. B. Pretre. Maximal fluctuations—a new phenomenon in disordered systems. Physica A: Statistical Mechanics and its Applications, 168(1):400–407, 1990.
21. 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. Physical review letters, 104(10):100601, 2010.
22. C. R. Rao. Linear statistical inference and its applications, volume 22. Wiley.com, 2009.
23. 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,
24. A. Sidi. *Practical extrapolation methods: Theory and applications*, volume 10. Cambridge University Press, 2003.

25. S Singh and R Singh. On the use of Shank’s transform to accelerate the summation of slowly converging series. *Microwave Theory and Techniques, IEEE Transactions on*, 39(3):608–610, 1991.

26. S. T. Smith. Optimum phase-only adaptive nulling. *Signal Processing, IEEE Transactions on*, 47(7):1835–1843, 1999.

27. C. Stockbridge, Y. Lu, J. Moore, S. Hoffman, R. Paxman, K. Toussaint, and T. Bifano. Focusing through dynamic scattering media. *Optics Express*, 20(14):15086–15092, 2012.

28. K. H. Toh, M. J. Todd, and R. H. Tutuncu. SDPT3 version 4.0 – a MATLAB software for semidefinite-quadratic-linear programming.

29. E. G. van Putten, A. Lagendijk, and A. P. Mosk. Optimal concentration of light in turbid materials. *JOSA B*, 28(5):1200–1203, 2011.

30. E. G. van Putten, I. M. Vellekoop, and A. P. Mosk. Spatial amplitude and phase modulation using commercial twisted nematic lcds. *Applied optics*, 47(12):2076–2081, 2008.

31. L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM review*, 38(1):49–95, 1996.

32. I. M. Vellekoop and A. P. Mosk. Phase control algorithms for focusing light through turbid media. *Optics Communications*, 281(11):3071–3080, 2008.

33. I. M. Vellekoop and A. P. Mosk. Universal optimal transmission of light through disordered materials. *Physical review letters*, 101(12):120601, 2008.
Fig. 1. Schematic for the experimental setup considered.

- Periodic repetition
- with period $D$

Fig. 2. Geometry of the scattering system considered.
Fig. 3. Plot of transmitted power obtained by SVD or SDP versus the thickness \( L/\lambda \) in a setting with \( D = 197\lambda, r = 0.11\lambda, N_e = 430,000, n_d = 1.3, M = 395, \bar{l} = 6.69\lambda \). SDP had 2.5\% improvement compared to SVD on average.
Fig. 4. The transmitted power versus the number of iterations is shown for steepest descent algorithm with $\mu = 0.5059$, for phase-only steepest descent algorithm with $\mu = 0.6574$ and for phase-only gradient algorithm with $\mu = 1.4149$ 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 phase-only steepest descent algorithm converged to the optimal transmitted power faster than the phase-only gradient algorithm.
Fig. 5. Maximum transmitted power in 50 iterations of SVD and SDP method, steepest descent, gradient descent and equal-phase input versus the thickness $L/\lambda$ in a setting with $D = 197\lambda, r = 0.11\lambda, N_c = 430,000, n_d = 1.3, M = 395, \bar{t} = 6.69\lambda$. Max of SDP and SVD had 8.3% improvement compared to SD on average.
Fig. 6. Heatmap of the transmitted power on the plane of number of iterations and stepsizes $\mu$ used in steepest descent method for the same setting as in Fig. 4.
Fig. 7. Heatmap of the maximum transmitted power in 50 iterations of steepest descent on the plane of stepsize and the thickness $L/\lambda$ in a setting with $D = 197\lambda, r = 0.11\lambda, N_c = 430,000, n_d = 1.3, M = 395, \tilde{\ell} = 6.69\lambda$. 

Fig. 8. Heatmap of the maximum transmitted power in 50 iterations of gradient descent on the plane of stepsize and the thickness $L/\lambda$ in a setting with $D = 197\lambda$, $r = 0.11\lambda$, $N_c = 430,000$, $n_d = 1.3$, $M = 395\lambda$, $\bar{T} = 6.69\lambda$. 
Fig. 9. Number of iterations to get 95% of the respective maximum transmitted power for steepest descent and gradient descent algorithms versus thickness $L/\lambda$ in a setting with $D = 197\lambda, r = 0.11\lambda, N_c = 430,000, n_d = 1.3, M = 395, \bar{T} = 6.69\lambda$. 