
Don't take it lightly: Phasing optical random projections with unknown operators

Sidharth Gupta

University of Illinois at Urbana-Champaign
gupta67@illinois.edu

Rémi Gribonval

Univ Rennes, Inria, CNRS, IRISA
remi.gribonval@inria.fr

Laurent Daudet

LightOn, Paris
laurent@lighton.ai

Ivan Dokmanić

University of Illinois at Urbana-Champaign
dokmanic@illinois.edu

Abstract

In this paper we tackle the problem of recovering the phase of complex linear measurements when only magnitude information is available and we control the input. We are motivated by the recent development of dedicated optics-based hardware for rapid random projections which leverages the propagation of light in random media. A signal of interest $\xi \in \mathbb{R}^N$ is mixed by a random scattering medium to compute the projection $y = A\xi$, with $A \in \mathbb{C}^{M \times N}$ being a realization of a standard complex Gaussian iid random matrix. Such optics-based matrix multiplications can be much faster and energy-efficient than their CPU or GPU counterparts, yet two difficulties must be resolved: only the intensity $|y|^2$ can be recorded by the camera, and the transmission matrix A is unknown. We show that even without knowing A , we can recover the unknown phase of y for some *equivalent* transmission matrix with the same distribution as A . Our method is based on two observations: first, conjugating or changing the phase of any row of A does not change its distribution; and second, since we control the input we can interfere ξ with arbitrary reference signals. We show how to leverage these observations to cast the *measurement phase retrieval problem* as a Euclidean distance geometry problem. We demonstrate appealing properties of the proposed algorithm in both numerical simulations and real hardware experiments. Not only does our algorithm accurately recover the missing phase, but it mitigates the effects of quantization and the sensitivity threshold, thus improving the measured magnitudes.

1 Introduction

Random projections are at the heart of many algorithms in machine learning, signal processing and numerical linear algebra. Recent developments ranging from classification with random features [16], kernel approximation [25] and sketching for matrix optimization [24, 27], to sublinear-complexity transforms [26] and randomized linear algebra are all enabled by random projections. Computing random projections for realistic signals such as images, videos, and modern big data streams is computation- and memory-intensive. Thus, from a practical point of view, any increase in the size and speed at which one can do the required processing is highly desirable.

This fact has motivated work on using dedicated hardware based on physics rather than traditional CPU and GPU computation to obtain random projections. A notable example is scattering of light in random media (Figure 1 (left)) with an optical processing unit (OPU). The OPU enables rapid (20 kHz) projections of high-dimensional data such as images, with input dimension scaling up to

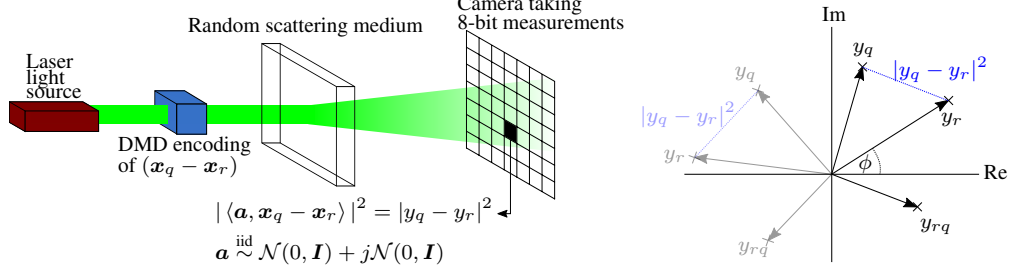


Figure 1: Left: The optical processing unit (OPU) is an example application of where the MPR problem appears. A coherent laser beam spatially encodes a signal $(\mathbf{x}_q - \mathbf{x}_r)$ via a digital micro-mirror device (DMD) which is then shined through a random medium. A camera measures the squared magnitude of the scattered light which is equivalent to the Euclidean distance between complex numbers $y_q \in \mathbb{C}$ and $y_r \in \mathbb{C}$. Furthermore the camera takes quantized measurements; Right: y_q and y_r are points on the two-dimensional complex plane. We can measure the squared Euclidean distance between points and use these distances to localize points on the complex plane and obtain their phase. Note that transformations such as rotations and reflections do not change the distances.

one million and output dimension also in the million range. It works by “imprinting” the input data $\xi \in \mathbb{R}^N$ onto a coherent light beam using a digital micro-mirror device (DMD) and shining the modulated light through a multiple scattering medium such as titanium dioxide white paint. The scattered lightfield in the sensor plane can then be written as

$$\mathbf{y} = \mathbf{A}\xi$$

where $\mathbf{A} \in \mathbb{C}^{M \times N}$ is the transmission matrix of the random medium with desirable properties.

One of the major challenges associated with this approach is that $\mathbf{A}$ is in general *unknown*. Though it could in principle be learned via calibration [6], such a procedure is slow and inconvenient, especially at high resolution. On the other hand, the system can be designed so that the distribution of $\mathbf{A}$ is approximately iid standard complex Gaussian. Luckily, this fact alone is sufficient for many algorithms and the actual values of $\mathbf{A}$ are not required.

Another challenge is that common light sensors are only sensitive to intensity, so we can only measure the intensity of scattered light, $|\mathbf{y}|^2$, where $|\cdot|$ is the elementwise absolute value. The phase information is thus lost. While the use of interferometric measurements with a reference could enable estimating the phase, the practical setup is more complex, sensitive, and it does not share the convenience and simplicity of the one illustrated in Figure 1 (left).

This motivates us to consider the *measurement phase retrieval (MPR) problem*. The MPR sensor data is modeled as

$$\mathbf{b} = |\mathbf{y}|^2 + \boldsymbol{\eta} = |\mathbf{A}\xi|^2 + \boldsymbol{\eta}, \quad (1)$$

where $\mathbf{b} \in \mathbb{R}^M$, $\xi \in \mathbb{R}^N$, $\mathbf{A} \in \mathbb{C}^{M \times N}$, $\mathbf{y} \in \mathbb{C}^M$, and $\boldsymbol{\eta} \in \mathbb{R}^M$ is noise. The goal is to recover the phase of each complex-valued element of $\mathbf{y}$, y_i for $1 \leq i \leq M$, from its magnitude measurements $\mathbf{b}$ when ξ is *known* and the entries of $\mathbf{A}$ are *unknown*. The classical phase retrieval problem which has received much attention over the last decade [15, 4] has the same quadratic form as (1) but with a known $\mathbf{A}$ and the task being to recover ξ instead of $\mathbf{y}$. While at a glance it might seem that not knowing $\mathbf{A}$ precludes computing the phase of $\mathbf{A}\xi$, we show in this paper that it is in fact possible via an exercise in distance geometry.

The noise $\boldsymbol{\eta}$ is primarily due to quantization because standard camera sensors measure low precision values, 8-bit in our case (integers between 0 and 255 inclusive). Furthermore, cameras may perform poorly at low intensities. This is another data-dependent noise source which is modelled in (2) by a binary mask vector $\mathbf{w} \in \mathbb{R}^M$ which is zero when the intensity is below some threshold and one otherwise; $\odot$ denotes the elementwise product.

$$\mathbf{b} = \mathbf{w} \odot (|\mathbf{y}|^2 + \boldsymbol{\eta}) = \mathbf{w} \odot (|\mathbf{A}\xi|^2 + \boldsymbol{\eta}) \quad (2)$$

The distribution of $\mathbf{A}$ follows from the properties of random scattering media [14, 6]. It has iid standard complex Gaussian entries, $a_{mn} \sim \mathcal{N}(0, 1) + j\mathcal{N}(0, 1)$ for all $1 \leq m, n \leq M, N$.

The usefulness of phase is obvious. While in some applications having only the magnitude of the random projection is enough (see [17] for an example related to elliptic kernels), most applications require the phase. For example, with the phase one can implement a more diverse range of kernels as well as randomized linear algebra routines like randomized singular value decomposition (SVD). We report the results of the latter on real hardware in Section 3.1.

Our contributions. We develop an algorithm based on distance geometry to solve the MPR problem (1). We exploit the fact that we control the input to the system, which allows us to mix ξ with arbitrary reference inputs. By interpreting each pixel value as a point in the complex plane, this leads to a formulation of the MPR problem as a pure distance geometry problem (see Section 2.2 and Figure 1 (right)). With enough pairwise distances (corresponding to reference signals) we can localize the points on the complex plane via a variant of multidimensional scaling (MDS) [23, 5], and thus compute the missing phase.

As we demonstrate, the proposed algorithm not only accurately recovers the phase, but also improves the number of useful bits of the magnitude information thanks to the multiple views. Established Euclidean distance geometry bounds imply that even with many distances below the sensitivity threshold and coarse quantization, the proposed algorithm allows for accurate recovery. This fact, which we verify experimentally, could have bearing on the design of future random projectors by navigating the tradeoff between physics and computation.

1.1 Related work

The classical phase retrieval problem looks at the case where A is known and ξ has to be recovered from b in (1) [7, 21, 10]. A modified version of the classical problem known as holographic phase retrieval is related to our approach: a known reference signal is concatenated with ξ to facilitate the phase estimation [1]. Interference with known references for classical phase retrieval has also been studied for known (Fourier) operators [3, 11].

An optical random projection setup similar to the one we consider has been used for kernel-based classification [17], albeit using only magnitudes. A phaseless approach to classification with the measured magnitudes fed into a convolutional neural network was reported by Satat et al. [18].

An alternative to obtaining the measurement phase is to measure, or calibrate, the unknown transmission matrix A . This has been attempted in compressive imaging applications but the process is impractical at even moderate pixel counts [6, 14]. Estimating A can take days and even the latest GPU-accelerated methods take hours for moderately sized A [20]. Other approaches forego calibration and use the measured magnitudes to *learn* an inverse map of $x \mapsto |Ax|^2$ for use with the magnitude measurements [9].

Leaving hardware approaches aside, there have been multiple algorithmic efforts to improve the speed of random projections [12, 25] for machine learning and signal processing tasks. Still, efficiently handling high-dimensional input remains a formidable challenge.

2 The measurement phase retrieval problem

We will denote the signal of interest by $\xi \in \mathbb{R}^N$, and the K reference *anchor* signals by $r_k \in \mathbb{R}^N$ for $1 \leq k \leq K$. To present the full algorithm we will need to use multiple signals of interest which we will then denote $\xi_1, \dots, \xi_S$; each ξ_s is called a frame. We set the last, K th anchor to be the origin, $r_K = 0$. We ascribe ξ and the anchors to the columns of the matrix $X \in \mathbb{R}^{N \times Q}$, so that $X = [\xi, r_1, r_2, \dots, r_K]$ and let $Q = K + 1$. The q th column of X is denoted x_q . For any $1 \leq q, r \leq Q$, we let $y_q = Ax_q$ and $y_{qr} := A(x_q - x_r)$, with $y_{qr,m}$ being its m th entry. Finally, the m th row of A will be denoted by a^m so that $y_{qr,m} = \langle a^m, x_q - x_r \rangle$.

2.1 Problem statement and recovery up to a reference phase and conjugation

Since we do not know A , it is clear that recovering the absolute phase of $A\xi$ is impossible. On the other hand, many algorithms do not require any knowledge of A except that it is iid standard complex Gaussian, and that it does not change throughout the computations.

Let R be an operator which adds a constant phase to each row of its argument (multiplies it by $\text{diag}(e^{j\phi_1}, \dots, e^{j\phi_m})$ for some $\phi_1, \dots, \phi_m$) and conjugates a subset of its rows. Since a standard complex Gaussian is circularly symmetric, $R(\mathbf{A})$ has the same distribution as $\mathbf{A}$. Therefore, since we do not know $\mathbf{A}$, it does not matter whether work with $\mathbf{A}$ itself or with $R(\mathbf{A})$ for some possibly unknown R . As long as the same *effective* R is used for all inputs during algorithm operation, the relative phases between the frames will be the same whether we use $R(\mathbf{A})$ or $\mathbf{A}$.¹

Problem 1. *Given a collection of input frames $\xi_1, \dots, \xi_S$ to be randomly projected and a device illustrated in Figure 1 (left) with an unknown transmission matrix $\mathbf{A} \in \mathbb{C}^{M \times N}$ and a b-bit camera, compute the estimates of projections $\hat{y}_1, \dots, \hat{y}_S$ up to a global row-wise phase and conjugation; that is, so that there exists some R such that $\hat{y}_s \approx R(y_s)$ for all $1 \leq s \leq S$.*

2.2 MPR as a distance geometry problem

Since the rows of $\mathbf{A}$ are statistically independent, we can explain our algorithm for a single row and then repeat the same steps for the remaining rows. We will therefore omit the row subscript/superscript m except where explicitly necessary.

Instead of randomly projecting ξ and measuring the corresponding projection magnitude $|\mathbf{A}\xi|^2$, consider randomly projecting the difference between ξ and some reference vector, or more generally a difference between two columns in $\mathbf{X}$, thus measuring $|\langle \mathbf{a}, \mathbf{x}_q - \mathbf{x}_r \rangle|^2 = |y_q - y_r|^2$. Interpreting y_q and y_r as points in the complex plane, we see that the camera sensor measures exactly the squared Euclidean distance between them. Since we control the input to the OPU, we can indeed set it to $\mathbf{x}_q - \mathbf{x}_r$ and measure $|y_q - y_r|^2$ for all $1 \leq q, r \leq Q$.

This is the key point: as we can measure pairwise distances between a collection of two-dimensional vectors in the two-dimensional complex plane, we can use established distance geometry algorithms such as multidimensional scaling (MDS) to localize points and get their phase. This is illustrated in Figure 1 (right). The same figure also illustrates the well known fact that rigid transformations of a point set cannot be recovered from distance data. We need to worry about three things: translations, reflections and rotations.

The translation ambiguity can be easily dealt with if one notes that for any column $\mathbf{x}_q$ of $\mathbf{X}$, $|y_q| = |\langle \mathbf{a}, \mathbf{x}_q \rangle|$ gives us the distance of y_q to the origin which is a fixed point, ultimately resolving the translation ambiguity. There is, however, no similar simple way to do away with the rotation and reflection ambiguity, so it might seem that there is no way to uniquely determine the phase of $\langle \mathbf{a}, \xi \rangle$. This is where the discussion from the preceding subsection comes to the rescue. Since R is arbitrary, as long as it is kept fixed for all the frames, we can arbitrarily set the orientation of any given frame and use it as a reference, making sure that the relative phases are computed correctly.

2.3 Proposed algorithm

As defined previously, the columns of $\mathbf{X} \in \mathbb{R}^{N \times Q}$ list the signal of interest and the anchors. Recall that all the entries of $\mathbf{X}$ are known. Using the OPU, we can compute a noisy (quantized) version of

$$|y_{qr}|^2 = |\langle \mathbf{a}, \mathbf{x}_q - \mathbf{x}_r \rangle|^2 = |y_q - y_r|^2, \quad (3)$$

for all (q, r) , which gives us $Q(Q-1)/2$ squared Euclidean distances between points $\{y_q \in \mathbb{C}\}_{q=1}^Q$ on the complex plane. These distances can be used to populate a Euclidean (squared) distance matrix $\mathbf{D} \in \mathbb{R}^{Q \times Q}$ as $\mathbf{D} = (d_{qr}^2)_{q,r=1}^Q = (|y_{qr}|^2)_{q,r=1}^Q$, which we will use to localize all complex points y_q .

We start by defining the matrix of all the complex points in $\mathbb{R}^2$ which we want to recover as

$$\mathbf{Y} = \begin{bmatrix} \text{Re}(y_1) & \text{Re}(y_2) & \cdots & \text{Re}(y_Q) \\ \text{Im}(y_1) & \text{Im}(y_2) & \cdots & \text{Im}(y_Q) \end{bmatrix} \in \mathbb{R}^{2 \times Q}.$$

Denoting the q th column of $\mathbf{Y}$ by $\mathbf{v}_q$, we have $d_{qr}^2 = \|\mathbf{v}_q - \mathbf{v}_r\|_2^2 = \mathbf{v}_q^T \mathbf{v}_q - 2\mathbf{v}_q^T \mathbf{v}_r + \mathbf{v}_r^T \mathbf{v}_r$ so that

$$\mathbf{D} = \text{diag}(\mathbf{G}) \mathbf{1}_Q^T - 2\mathbf{G} + \mathbf{1}_Q \text{diag}(\mathbf{G})^T =: \mathcal{K}(\mathbf{G}), \quad (4)$$

¹Up to a sign.

where $\text{diag}(\mathbf{G}) \in \mathbb{R}^Q$ is the column vector of the diagonal entries in the Gram matrix $\mathbf{G} := \mathbf{\Upsilon}^T \mathbf{\Upsilon} \in \mathbb{R}^{Q \times Q}$ and $\mathbf{1}_Q \in \mathbb{R}^Q$ is the column vector of Q ones. This establishes a relationship between the measured distances in $\mathbf{D}$ and the locations of the complex points in $\mathbb{R}^2$ which we seek. We denote by $\mathbf{J}$ the geometric centering matrix, $\mathbf{J} := \mathbf{I} - \frac{1}{Q} \mathbf{1}_Q \mathbf{1}_Q^T$ so that

$$\hat{\mathbf{G}} = -\frac{1}{2} \mathbf{J} \mathbf{D} \mathbf{J} = \mathbf{J} \mathbf{G} \mathbf{J} = (\mathbf{\Upsilon} \mathbf{J})^T (\mathbf{\Upsilon} \mathbf{J}) \quad (5)$$

is the Gram matrix of the centered point set in terms of $\mathbf{\Upsilon}$. $\hat{\mathbf{G}}$ and $\mathbf{J}$ are known as the Gram matrix of the centered point set and the geometric centering matrix because $\mathbf{\Upsilon} \mathbf{J}$ is the points in $\mathbf{\Upsilon}$ with their mean subtracted. An estimate $\hat{\mathbf{\Upsilon}}$ of the centered point set, $\mathbf{\Upsilon} \mathbf{J}$, is then obtained by eigendecomposition as $\hat{\mathbf{G}} = \mathbf{V} \text{diag}(\lambda_1, \dots, \lambda_Q) \mathbf{V}^T$ and taking $\hat{\mathbf{\Upsilon}} = [\sqrt{\lambda_1} \mathbf{v}_1, \sqrt{\lambda_2} \mathbf{v}_2]^T$ where $\mathbf{v}_1$ and $\mathbf{v}_2$ are the first and second column of $\mathbf{V}$ and assuming that the eigenvalue sequence is nonincreasing. This process is the classical MDS algorithm [23, 5]. Finally, the phases can be calculated via a four-quadrant inverse tangent, $\phi(y_q) = \arctan(v_{q2}, v_{q1})$.

Procrustes analysis. As we recovered a centered point set via MDS with a geometric centering matrix $\mathbf{J}$, the point set will have its centroid at the origin. This is a consequence of the used algorithm, and not the “true” origin. As described above, we know that $|y_q|^2$ defines squared distances to the origin and $y_Q = \langle \mathbf{a}, \mathbf{x}_Q \rangle = 0 + 0j$ (as $\mathbf{x}_Q$ was set to the origin), meaning that we can correctly center the recovered points by translating the point set, $\hat{\mathbf{\Upsilon}}$, by $-\mathbf{v}_Q$.

The correct absolute rotation and reflection cannot be recovered. However, since we only care about working with some effective $\mathbf{R}(\mathbf{A})$ with the correct distribution, we only need to ensure that the relative phases between the frames are correct. We can thus designate the first frame as the reference frame and set the rotation (which directly corresponds to the phase) and reflection (corresponding to conjugation) arbitrarily. Once these are chosen, the anchors $\mathbf{r}_1, \dots, \mathbf{r}_K$ are fixed, which in turn fixes the phasing–conjugation operator $\mathbf{R}$.

Since $\mathbf{A}$ is unknown, $\mathbf{R}$ is also unknown, but fixed anchors allow us to compute the correct relative phase with respect to $\mathbf{R}(\mathbf{A})$ for the subsequent frames. Namely, upon receiving a new input ξ_s to be randomly projected, we now localize it with respect to a fixed set of anchors. This is achieved by Procrustes analysis. Denoting by $\tilde{\mathbf{\Upsilon}}_1$ our reference estimate of the anchor positions in frame 1 (columns 2, $\dots$, Q of $\hat{\mathbf{\Upsilon}}$ above which was recovered from $\hat{\mathbf{G}}$ in (5)), and by $\tilde{\mathbf{\Upsilon}}_s$ the MDS estimate of anchor positions in frame s , adequately centered. Let $\tilde{\mathbf{\Upsilon}}_s \tilde{\mathbf{\Upsilon}}_1^T = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T$ be the singular value decomposition of $\tilde{\mathbf{\Upsilon}}_s \tilde{\mathbf{\Upsilon}}_1^T$. The optimal transformation matrix in the least squares sense is then $\mathbf{R} = \mathbf{V} \mathbf{U}^T$ so that $\mathbf{R} \tilde{\mathbf{\Upsilon}}_s \approx \tilde{\mathbf{\Upsilon}}_1$ [19].

Finally, we note that with a good estimate of the anchors, one can imagine not relocalizing them in every frame. The localization problem for ξ then boils down to multilateration, cf. Section C in the supplementary material.

2.4 Sensitivity threshold and missing measurements

As we further elaborate in Section A of the supplementary material, in practice some measurements fall below the sensitivity threshold of the camera and produce spurious values. A nice benefit of multiple “views” of ξ via its interaction with reference signals is that we can ignore those measurements. This introduces missing values in $\mathbf{D}$ which can be modeled via a binary mask matrix $\mathbf{W}$. The recovery problem can be modeled as estimating $\mathbf{\Upsilon}$ from $\mathbf{W} \odot (\mathbf{D} + \mathbf{E})$ where $\mathbf{W} \in \mathbb{R}^{N \times N}$ contains zeros for the entries which fall below some prescribed threshold, and ones otherwise.

We can predict the performance of the proposed method when modeling the entries of $\mathbf{W}$ as iid Bernoulli random variables with parameter p , where $1 - p$ is the probability that an entry falls below the sensitivity threshold and $\mathbf{E}$ as uniform quantization noise distributed as $\mathcal{U}\left(-\frac{\kappa}{2(2^b-1)}, \frac{\kappa}{2(2^b-1)}\right)$, where b is the number of bits, and κ an upper bound on the entries of $\mathbf{D}$ (in our case $2^8 - 1 = 255$).

Adapting existing results on the performance of multidimensional scaling [28] (by noting that $\mathbf{E}$ is sub-Gaussian), we can get the following scaling of the distance recovery error with the number of

Algorithm 1 MPR algorithm for S frames.

Input: Squared distances $[|y_{jQ,m} - y_{lQ,m}|^2]_s$ for all $1 \leq j, l \leq Q$ for frames $1 \leq s \leq S$ and rows $1 \leq m \leq M$; $[\cdot]_s$ denotes frame s

Output: $\mathbf{Y} \in \mathbb{C}^{M \times S}$ containing all localized points such that $\mathbf{y}_s = \mathbf{R}(\mathbf{A})\boldsymbol{\xi}_s$ for some fixed $\mathbf{R}$.

```

1:  $\mathbf{Y} \leftarrow \mathbf{0}_{M \times S}$  ▷ Initialize  $\mathbf{Y}$ 
2:  $m \leftarrow 1$ 
3: while  $m \leq M$  do ▷ Solve each row separately
4:   Populate all frame  $s = 1$  distances into distance matrix  $\mathbf{D}$  ▷  $\mathbf{D} \in \mathbb{R}^{Q \times Q}$ 
5:    $[\boldsymbol{\Upsilon}]_1 \leftarrow \text{MDS}(\mathbf{D})$  ▷  $[\boldsymbol{\Upsilon}]_1 \in \mathbb{R}^{2 \times Q}$ 
6:    $[\boldsymbol{\Upsilon}]_1 \leftarrow \text{GradientDescent}(\mathbf{D}, [\boldsymbol{\Upsilon}]_1)$ 
7:    $[\boldsymbol{\Upsilon}]_1 \leftarrow [\boldsymbol{\Upsilon}]_1 - [\mathbf{v}_Q]_1 \mathbf{1}^T$  ▷ Translate to align with origin
8:    $s \leftarrow 2$ 
9:   while  $s \leq S$  do
10:    Populate all frame  $s$  distances into distance matrix  $\mathbf{D}$ 
11:     $[\boldsymbol{\Upsilon}]_s \leftarrow \text{MDS}(\mathbf{D})$ 
12:     $[\boldsymbol{\Upsilon}]_s \leftarrow \text{GradientDescent}(\mathbf{D}, [\boldsymbol{\Upsilon}]_s)$ 
13:     $[\boldsymbol{\Upsilon}]_s \leftarrow [\boldsymbol{\Upsilon}]_s - [\mathbf{v}_Q]_s \mathbf{1}^T$ 
14:     $\mathbf{R} \leftarrow \text{Procrustes}([\mathbf{v}_2, \dots, \mathbf{v}_Q]_1, [\mathbf{v}_2, \dots, \mathbf{v}_Q]_s)$  ▷  $\mathbf{R}$  aligns frames 1 and  $s$  anchors
15:     $[\boldsymbol{\Upsilon}]_s \leftarrow \text{Align}([\boldsymbol{\Upsilon}]_s, \mathbf{R}, [\mathbf{v}_2, \dots, \mathbf{v}_Q]_1)$  ▷ Align anchors
16:     $s \leftarrow s + 1$ 
17:  end while
18:   $\mathbf{U} \leftarrow [[\mathbf{v}_1]_1, [\mathbf{v}_1]_2, \dots, [\mathbf{v}_1]_S]$  ▷  $\mathbf{U} \in \mathbb{R}^{2 \times S}$ 
19:   $\mathbf{y}^m \leftarrow \mathbf{u}^1 + j\mathbf{u}^2$  ▷ Multiply second row of  $\mathbf{U}$  with  $j$  and add to first row
20:   $m \leftarrow m + 1$ 
21: end while
```

anchors K (for K sufficiently large),

$$\frac{1}{K} \mathbb{E} \left[\left\| \hat{\mathbf{D}} - \mathbf{D} \right\|_F \right] \lesssim \frac{\kappa}{\sqrt{pK}}, \quad (6)$$

where $\lesssim$ denotes inequality up to a constant which depends on the number of bits b , the sub-Gaussian norm of the entries in $\mathbf{E}$, and the dimension of the ambient space (here $\mathbb{R}^2$). An important implication is that even for coarse quantization (small b) and for a large fraction of entries below the sensitivity threshold (small p), we can achieve arbitrarily small amplitude and phase errors *per point* by increasing the number of reference signals K .

Refinement with gradient descent. The output of the classical MDS method described above can be further refined via a local search. A standard differentiable objective called the squared stress is defined as follows,

$$\min_{\mathbf{Z}} f(\boldsymbol{\Upsilon}) = \min_{\mathbf{Z}} \left\| \mathbf{W} \odot \left(\mathbf{D} - \mathcal{K}(\mathbf{Z}^T \mathbf{Z}) \right) \right\|_F^2, \quad (7)$$

where $\mathcal{K}(\cdot)$ is as defined in (4) and $\mathbf{Z} \in \mathbb{R}^{2 \times Q}$ is the point matrix induced by row m of $\mathbf{A}$. In our experiments we report the result of refining the classical MDS results via gradient descent on (7).

Note that the optimization (7) is nonconvex. The complete procedure is thus analogous to the usual approach to nonconvex phase retrieval by spectral initialization followed by gradient descent [15, 4]. Algorithm 1 summarizes our proposed method.

3 Experimental verification and application

We test the proposed MPR algorithm via simulations and experiments on a real OPU. For hardware experiments, we use a scikit-learn interface to a publicly available cloud-based OPU.²

²<https://www.lighton.ai/lighton-cloud/>.
 Reproducible code available at https://github.com/swing-research/opu_phase under the MIT License.

Evaluation metrics. The main challenge is to evaluate the performance without knowing the transmission matrix $\mathbf{A}$. To this end, we propose to use the *linearity error*. The rationale behind this metric is that with the phase correctly recovered, the end-to-end system should be linear. That is, if we recover $\mathbf{y}$ and $\mathbf{z}$ from $|\mathbf{y}|^2 = |\mathbf{A}\boldsymbol{\xi}_1|^2$ and $|\mathbf{z}|^2 = |\mathbf{A}\boldsymbol{\xi}_2|^2$, then we should get $(\mathbf{y} + \mathbf{z})$ when applying the method to $|\mathbf{v}|^2 = |\mathbf{A}(\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2)|^2$. With this notation, the relative linearity error is defined as

$$\text{linearity error} = \frac{1}{M} \sum_{m=1}^M \frac{|(y_m + z_m) - v_m|}{|v_m|}. \quad (8)$$

The second metric we use is the number of “good” or correct bits. This metric can only be evaluated in simulation since it requires the knowledge of the ground truth measurements. Letting $|\mathbf{y}|^2 = |\langle \mathbf{a}, \boldsymbol{\xi} \rangle|^2$ and $\hat{\mathbf{y}}$ be our estimate of $\mathbf{y}$, the number of good bits is defined as

$$\text{good bits} = -\frac{20}{6.02} \log \left(\frac{||\mathbf{y}|^2 - |\hat{\mathbf{y}}|^2|}{|\mathbf{y}|^2} \right).$$

It is proportional to the signal-to-quantization-noise ratio if the distances uniformly cover all quantization levels.³

3.1 Experiments

In all simulations, intensity measurements are quantized to 8 bits and all signals and references are iid standard (complex) Gaussian random vectors.

We first test the phase recovery performance by evaluating the linearity error. In simulation, we draw random frames $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$, and $\mathbf{A} \in \mathbb{C}^{100 \times 64^2}$. We apply Algorithm 1 to $|\mathbf{A}\boldsymbol{\xi}_1|^2$, $|\mathbf{A}\boldsymbol{\xi}_2|^2$ and $|\mathbf{A}(\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2)|^2$ and calculate the linearity error (8). We use classical MDS and MDS with gradient descent (MDS-GD). Figure 2a shows that the system is indeed approximately linear and that the linearity error becomes smaller as the number of reference signals grows. In Figure 2b, we set the sensitivity threshold to $\tau = 6$ and zero the distances below the threshold per (2). Again, the linearity error quickly becomes small as the number of anchors increases showing that the overall system is robust and that it allows recovery of phase for small-intensity signals.

Next, we test the linearity error with a real hardware OPU. The OPU gives 8-bit unsigned integer measurements. A major challenge is that the DMD (see Figure 1) only allows binary input signals. This is a property of the particular OPU we use and while it imposes restrictions on reference design, the method is unchanged as our algorithm does not assume a particular type of signal. Section A in the supplementary material describes how we create binary references and addresses other hardware-related practicalities.

Figure 2c reports the linearity error on the OPU with suitably designed references and the same size $\mathbf{A}$. The empirically determined sensitivity threshold of the camera is $\tau = 6$, and the measurements below the threshold were not used. We ignore rows of $\mathbf{A}$ which give points with small norms (less than two) because they are prone to noise and disproportionately influence the relative error. Once again, we observe that the end-to-end system with Algorithm 1 is approximately linear and that the linearity improves as we increase the number of anchors.

Finally, we demonstrate the magnitude denoising performance. We draw $\mathbf{a} \in \mathbb{C}^{100}$, a random signal $\boldsymbol{\xi} \in \mathbb{R}^{100}$ and a set of random reference anchor signals. We run our algorithm for number of anchors varying between 2 and 15. For each number of anchors, we recover $\hat{\mathbf{y}}$ for $|\mathbf{y}|^2 = |\langle \mathbf{a}, \boldsymbol{\xi} \rangle|^2$ using either classical MDS or MDS-GD. We then measure the number of good bits. The average results over 100 trials are shown in Figure 3a. Figure 3b reports the same experiment with the sensitivity threshold set to $\tau = 6$ (that is, the entries below τ are zeroed in the distance matrix per (2)). Both figures show that the proposed algorithm significantly improves the estimated magnitudes in addition to recovering the phases. The approximately 1 additional good bit with gradient descent in Figure 3b corresponds to the relative value of $2^1/2^8 \approx 0.8\%$ which is consistent with the gradient descent improvement in Figure 2b.

We also test a scenario where the anchor positions on the complex plane are known exactly and we only have to localize a single measurement. We compare this to localizing the anchors and the

³Note that the quantity registered by the camera is actually the squared magnitude, hence the factor 20.

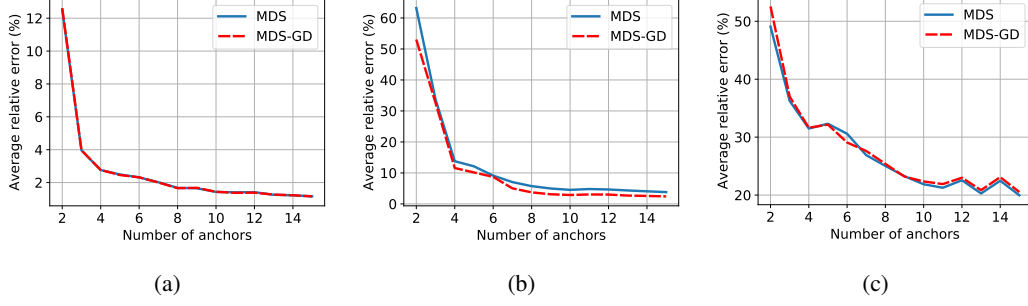


Figure 2: Experiments in simulation and on real hardware to evaluate the linearity error as defined in (8). The input signals are of dimension 64^2 , M in (8) is 100 and the number of anchors signals are increased. The classical MDS and MDS with gradient descent (MDS-GD) are used. In all cases the error decreases as the number of anchors increases. (a) In simulation with Gaussian signals and Gaussian reference signals; (b) In simulation with Gaussian signals and Gaussian reference signals with sensitivity threshold $\tau = 6$; (c) On a real OPU with binary signals and binary references.

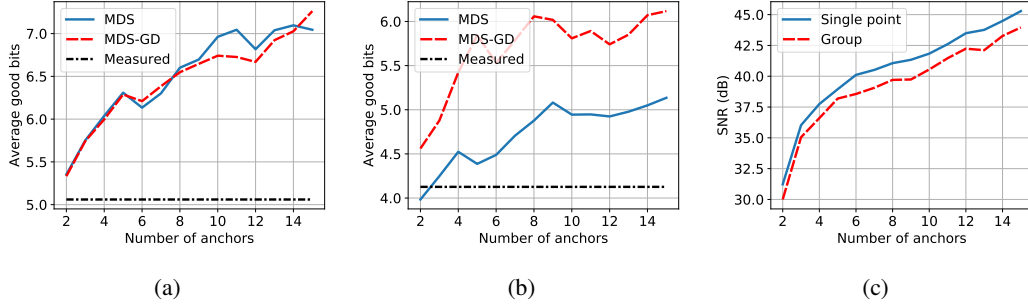


Figure 3: (a) Magnitude denoising performance of MDS and MDS-GD over 100 trials. Input signals are Gaussian and of dimension 100; (b) Magnitude denoising performance of MDS and MDS-GD over 100 trials with 100-dimensional Gaussian signals and sensitivity threshold $\tau = 6$; (c) Comparison between recovering a single point and recovering the point and anchors at the same time. SR-LS is used to locate a single point when anchors are known and MDS is used to locate all points when anchors are unknown.

measurements jointly. Localizing a single point via multilateration is performed by minimizing the SR-LS objective (see (9) in the supplementary material). The input signal dimension is 64^2 and we recover $\hat{y}$ for $|y|^2 = |\langle \mathbf{a}, \xi \rangle|^2$. We perform 100 trials and calculate the SNR of the recovered complex points. Figure 3c shows that although having perfect knowledge of anchor locations helps, classical MDS alone does not perform much worse.

Optical randomized singular value decomposition. We use Algorithm 1 to implement randomized singular value decomposition (RSVD) as described in Halko et al. [8] on the OPU. We use 5 anchors in all RSVD experiments. The original RSVD algorithm and a variant with adaptations for the OPU are described in Algorithms 2 and 3 in the supplementary material.

One of the steps in the RSVD algorithm for an input matrix $\mathbf{B} \in \mathbb{R}^{M \times N}$ requires the computation of $\mathbf{B}\Omega$ where $\Omega \in \mathbb{R}^{N \times 2K}$ is a standard real Gaussian matrix, K is the target number of singular vectors, and $2K$ may be interpreted as the number of random projections for each row of $\mathbf{B}$. We use the OPU to compute this random matrix multiplication. An interesting observation is that since in Algorithm 1 we recover the result of multiplications by a complex matrix with independent real and imaginary parts, we can halve the number of projections when using the OPU with respect to the original algorithm. By treating each row of $\mathbf{B}$ as an input frame, we can obtain $\mathbf{Y} \in \mathbb{C}^{K \times M}$ via Algorithm 1 when $|\mathbf{Y}|^2 = |\mathbf{A}\mathbf{B}^T|^2$ with $\mathbf{A}$ as defined in Problem 1 with K rows. Then, we can construct $\mathbf{P} = [\text{Re}(\mathbf{Y}^*) \quad \text{Im}(\mathbf{Y}^*)] \in \mathbb{R}^{M \times 2K}$ which would be equivalent to computing $\mathbf{B}\Omega$ for real Ω . Section B in the supplementary material describes this in more detail.

Figure 4 shows the results when the OPU is used to perform the random matrix multiplication of the RSVD algorithm on a matrix B . Figure 4 (left) reports experiments with a random binary matrix $B \in \mathbb{R}^{10 \times 10^4}$, different numbers of random projections (number of rows in A), and ten trials per number of projections. We plot the average error per entry when reconstructing B from its RSVD matrices and singular values. Next, we take 500 28×28 samples from the MNIST dataset [13], threshold them to be binary, vectorize them, and stack them into a matrix $B \in \mathbb{R}^{500 \times 28^2}$. Figure 4 (right) shows the seven leading right singular vectors reshaped to 28×28 . The top row shows the singular vectors that are obtained when using the OPU with 500 projections and the bottom row shows the result when using Python. The error is negligible.

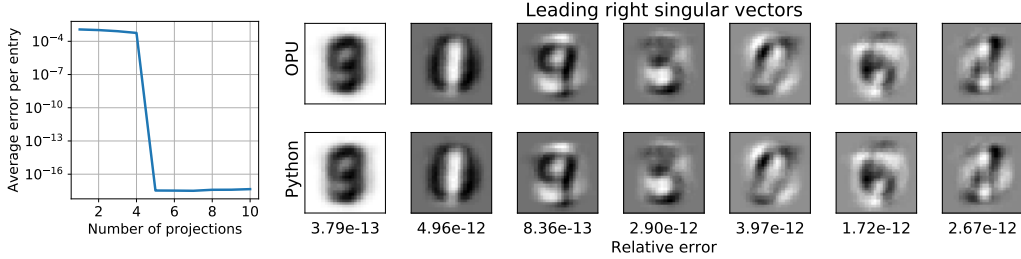


Figure 4: Left: Average RSVD error over 10 trials with varying number of projections on hardware with an input matrix of size 10×1000 ; Right: Reshaped leading right singular vectors of an MNIST matrix of size 500×28^2 . The top rows shows the leading right singular vectors after performing RSVD with the OPU and using our algorithm. The bottom row shows the leading right singular vectors from Python. The relative error is below each singular vector.

4 Conclusion

Traditional computation methods are often too slow for processing tasks which involve large data streams. This motivates alternatives which instead use fast physics to “compute” the desired functions. In this work, we looked at using optics and multiple scattering media to obtain linear random projections. A common difficulty with optical systems is that off-the-shelf camera sensors only register the intensity of the scattered light. Our results show that there is nevertheless no need to reach for more complicated and more expensive coherent setups. We showed that measurement phase retrieval can be cast as a problem in distance geometry, and that the unknown phase of random projections can be recovered even without knowing the transmission matrix of the medium.

Simulations and experiments on real hardware show that the OPU setup combined with our algorithm indeed approximates an end-to-end linear system. What is more, we also improve intensity measurements. The fact that we get full complex measurements allows us to implement a whole new spectrum of randomized algorithms; we demonstrated the potential by the randomized singular value decomposition. These benefits come at the expense of a reduction in data throughput. Future work will have to precisely quantify the smallest achievable data rate reduction due to allocating a part of the duty cycle for reference measurements, though we note that the optical processing data rates are very high to begin with.

Acknowledgement

Sidharth Gupta and Ivan Dokmanić would like to acknowledge support from the National Science Foundation under Grant CIF-1817577.

References

- [1] David A Barmherzig, Ju Sun, Emmanuel J Candes, TJ Lane, and Po-Nan Li. Holographic phase retrieval and optimal reference design. *arXiv preprint arXiv:1901.06453*, 2019.
- [2] Amir Beck, Petre Stoica, and Jian Li. Exact and approximate solutions of source localization problems. *IEEE Transactions on Signal Processing*, 56(5):1770–1778, 2008.
- [3] Robert Beinert. One-dimensional phase retrieval with additional interference intensity measurements. *Results in Mathematics*, 72(1-2):1–24, 2017.
- [4] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval via wirtinger flow: Theory and algorithms. *IEEE Transactions on Information Theory*, 61(4):1985–2007, 2015.
- [5] Ivan Dokmanic, Reza Parhizkar, Juri Ranieri, and Martin Vetterli. Euclidean distance matrices: essential theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 32(6):12–30, 2015.
- [6] Angélique Drémeau, Antoine Liutkus, David Martina, Ori Katz, Christophe Schülke, Florent Krzakala, Sylvain Gigan, and Laurent Daudet. Reference-less measurement of the transmission matrix of a highly scattering material using a dmd and phase retrieval techniques. *Optics express*, 23(9):11898–11911, 2015.
- [7] James R Fienup. Phase retrieval algorithms: a comparison. *Applied optics*, 21(15):2758–2769, 1982.
- [8] Nathan Halko, Per-Gunnar Martinsson, and Joel A Tropp. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM review*, 53(2):217–288, 2011.
- [9] Ryoichi Horisaki, Ryosuke Takagi, and Jun Tanida. Learning-based imaging through scattering media. *Optics express*, 24(13):13738–13743, 2016.
- [10] Kishore Jaganathan, Yonina C Eldar, and Babak Hassibi. Phase retrieval: An overview of recent developments. *arXiv preprint arXiv:1510.07713*, 2015.
- [11] Wooshik Kim and Monson H Hayes. Phase retrieval using two fourier-transform intensities. *JOSA A*, 7(3):441–449, 1990.
- [12] Quoc Le, Tamás Sarló, and Alex Smola. Fastfood-approximating kernel expansions in loglinear time. In *Proceedings of the international conference on machine learning*, volume 85, 2013.
- [13] Yann LeCun, Léon Bottou, Yoshua Bengio, Patrick Haffner, et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [14] Antoine Liutkus, David Martina, Sébastien Popoff, Gilles Chardon, Ori Katz, Geoffroy Lerosey, Sylvain Gigan, Laurent Daudet, and Igor Carron. Imaging with nature: Compressive imaging using a multiply scattering medium. *Scientific reports*, 4:5552, 2014.
- [15] Praneeth Netrapalli, Prateek Jain, and Sujay Sanghavi. Phase retrieval using alternating minimization. In *Advances in Neural Information Processing Systems*, pages 2796–2804, 2013.
- [16] Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems*, pages 1177–1184, 2008.
- [17] Alaa Saade, Francesco Caltagirone, Igor Carron, Laurent Daudet, Angélique Drémeau, Sylvain Gigan, and Florent Krzakala. Random projections through multiple optical scattering: Approximating kernels at the speed of light. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6215–6219. IEEE, 2016.
- [18] Guy Satat, Matthew Tancik, Otkrist Gupta, Barmak Heshmat, and Ramesh Raskar. Object classification through scattering media with deep learning on time resolved measurement. *Optics express*, 25(15):17466–17479, 2017.

- [19] Peter Hans Schoenemann. *A solution of the orthogonal Procrustes problem with applications to orthogonal and oblique rotation*. PhD thesis, University of Illinois at Urbana-Champaign, 1964.
- [20] Manoj Sharma, Christopher A Metzler, Sudarshan Nagesh, Oliver Cossairt, Richard G Baraniuk, and Ashok Veeraraghavan. Inverse scattering via transmission matrices: Broadband illumination and fast phase retrieval algorithms. *IEEE Transactions on Computational Imaging*, 2019.
- [21] Yoav Shechtman, Yonina C Eldar, Oren Cohen, Henry Nicholas Chapman, Jianwei Miao, and Mordechai Segev. Phase retrieval with application to optical imaging: a contemporary overview. *IEEE signal processing magazine*, 32(3):87–109, 2015.
- [22] Petre Stoica and Jian Li. Lecture notes-source localization from range-difference measurements. *IEEE Signal Processing Magazine*, 23(6):63–66, 2006.
- [23] Warren S Torgerson. Multidimensional scaling: I. theory and method. *Psychometrika*, 17(4):401–419, 1952.
- [24] Joel A Tropp, Alp Yurtsever, Madeleine Udell, and Volkan Cevher. Practical sketching algorithms for low-rank matrix approximation. *SIAM Journal on Matrix Analysis and Applications*, 38(4):1454–1485, 2017.
- [25] Yun Yang, Mert Pilanci, Martin J Wainwright, et al. Randomized sketches for kernels: Fast and optimal nonparametric regression. *The Annals of Statistics*, 45(3):991–1023, 2017.
- [26] Felix Xinnan X Yu, Ananda Theertha Suresh, Krzysztof M Choromanski, Daniel N Holtmann-Rice, and Sanjiv Kumar. Orthogonal random features. In *Advances in Neural Information Processing Systems*, pages 1975–1983, 2016.
- [27] Alp Yurtsever, Madeleine Udell, Joel A Tropp, and Volkan Cevher. Sketchy decisions: Convex low-rank matrix optimization with optimal storage. *arXiv preprint arXiv:1702.06838*, 2017.
- [28] Huan Zhang, Yulong Liu, and Hong Lei. Localization from incomplete euclidean distance matrix: Performance analysis for the svd-mds approach. *IEEE Transactions on Signal Processing*, 2019.