

1 We thank the reviewers for taking the time to read and comment on our paper. Below we first group the comments of
2 Reviewers #1 and #2 by topic, and then address the more critical concerns of Reviewer #3.

3 **Scope of applications.** Reviewers #1 and #2 find the work novel and interesting, but comment that it might cater
4 to a small community. While our focus on optical random projections could suggest so, the methodology is widely
5 applicable: our method linearizes any imaging, computation, or communication system in which the phase is important
6 but only the intensity is measured, regardless of whether the transmission matrix (TM) is known or unknown. While
7 mostly optical, these include X-ray and terahertz imaging, electron lasers, etc. The key is to have some control over the
8 input. We give two concrete, relevant examples related to reference [A] suggested by Reviewer #1 (thank you!):

9 1) Once the TM of the medium is calibrated, imaging still requires phase retrieval. We present the method for an
10 unknown \mathbf{A} , but it applies just the same with \mathbf{A} known. If one introduces a set of interferers in the scene whose
11 response is known, then the phase can be obtained rapidly via our proposed method;

12 2) More interestingly, our method reduces TM learning to a *linear* system (vs quadratic). The same Algorithm 1
13 determines the phase of the measurements, yielding a linear system for every row of \mathbf{A} . Finally, unknown TMs remain
14 relevant, as calibration is slow and many media do change over time. We stop here for brevity, but indeed feel that the
15 method can have a broad impact. The above examples can be easily incorporated into the manuscript.

16 **Suggestions for improvement.** The reviewers suggested to improve the clarity of Section 2.3 and experiment and
17 figure descriptions, which we are happy to do. Reviewer #2 also commented as follows:

18 1) *The related work does not clearly show the gap.* We feel that the cited work puts our work in context with the existing
19 literature: Tasks which do not recover the measurement phase [9, 17, 18] have limited performance. Transmission
20 matrix estimation [6, 14] is computationally heavy for large resolution / dynamic media. Our algorithm recovers the
21 phase in real time since it only requires computation of small SVDs, e.g., 6×6 .

22 2) *Difference in performance between MDS and MDS-GD.* Fig. 2b and 3b use different metrics (relative error vs good
23 bits); note that a 1 bit difference which is clearly visible in 3b amounts to only 0.8% relative error. The difference
24 between Fig. 3a and 3b is that in 3a there is only quantization, while in 3b there is thresholding, hence larger bit error.

25 **Reviewer #3.** The reviewer summarizes our work as developing an algorithm for the measurement phase retrieval
26 (MPR) problem in an OPU. Rather, we propose a general algorithm for the MPR, OPU being an example application.

27 **Multiple signals and binary inputs.** The reviewer writes that the paper is confusing, questioning the need for multiple
28 frames. We explain this in l. 134-140: in short, for a single input ξ_1 , since the matrix \mathbf{A} is unknown (and so the global
29 phase of each row), we can choose the phase of $\langle \mathbf{a}_i, \xi_1 \rangle$ arbitrarily and the MPR is vacuous (cf. rotation invariance of
30 complex Gaussians, l. 108-113). We thus use the first frame as a reference and align future frames with it via Procrustes
31 analysis (l. 157-172), yielding correct *relative* phases. The reviewer next states that “*The supp. mat. says that the*
32 *problem in the main text considers only binary signals.*” We respectfully disagree. The main text considers general
33 signals, and simulation experiments use general signals as evident from l. 216-226, 237-249. Only the particular
34 hardware we use for *real-world experiments* has a binary DMD. This is standard and due to the use of cheap, fast
35 modulators developed for movie projectors, as opposed to the slow and costly liquid crystal tech.

36 The reviewer then writes “*The organization is not good, either. Assumptions such as ‘DMD is binary’ seems to be*
37 *popped up all of a sudden from nowhere.*” Again, we find this surprising since it appears only on p. 7 when we show
38 results with real hardware (instead of stopping at numerics): “*Next, we test the linearity error with a real hardware*
39 *OPU. The OPU gives 8-bit unsigned integer measurements. A major challenge is that since the DMD is binary, the*
40 *inputs need to be binary.*” This does not influence the organization of the paper nor the algorithm. It merely affects
41 reference design. The reviewer continues “*The paper also cited some other algorithm without explaining what the*
42 *problems other algorithms solve.*” Unfortunately we are not sure what this refers to. (Perhaps: the pseudocode for
43 RSVD is given in the supplement; the algorithm is standard.)

44 **Randomized SVD.** As RSVD uses multiple random projections, our phase estimates must be consistent over many
45 frames (l. 108-113, Problem 1), hence this is more challenging than a matrix–vector product. This experiment shows
46 that the linearization is good enough for real-life applications. To answer the reviewer’s question “*Perhaps the authors*
47 *just think this is a fun supplementary discovery?*”: This is not the case. Showing that our phases are sufficiently good for
48 RSVD *on real hardware* is a strong demonstration of the method’s practicality in a context relevant for machine learning
49 and scientific computing. Benchmarks for standard and randomized SVD on standard hardware, and optical randomized
50 SVD would be interesting, but those will be very much hardware-dependent and therefore out of the scope of this paper.
51 With large-enough inputs, optical computation inevitably becomes faster than both CPU and GPU, as its timing is
52 *essentially independent* of the input size. To give an order of magnitude, the optical unit we use is faster than GPUs
53 for input/output dimensions above a few thousands—more than the presented MNIST problem, but much below the
54 size of real-life problems. For example, Tropp et al., Streaming Low-Rank Matrix Approximation with
55 an Application to Scientific Simulation. applies a variant of RSVD to climate data of up to 75 GB (!).