# PHASE RETRIEVAL FROM LOCAL MEASUREMENTS: IMPROVED ROBUSTNESS VIA EIGENVECTOR-BASED ANGULAR SYNCHRONIZATION

MARK A. IWEN, BRIAN PRESKITT, RAYAN SAAB, ADITYA VISWANATHAN

ABSTRACT. We improve a phase retrieval approach that uses correlation-based measurements with compactly supported measurement masks [30]. Our approach admits deterministic measurement constructions together with a robust, fast recovery algorithm that consists of solving a system of linear equations in a lifted space, followed by finding an eigenvector (e.g., via an inverse power iteration). Theoretical reconstruction error guarantees from [30] are improved as a result for the new and more robust reconstruction approach proposed herein. Numerical experiments demonstrate robustness and computational efficiency that compete with other approaches on large problems. Along the way, we show that this approach also trivially extends to phase retrieval problems based on windowed Fourier measurements.

## 1. INTRODUCTION

Consider the problem of recovering a vector $\mathbf{x}_0 \in \mathbb{C}^d$ from measurements $\mathbf{y} \in \mathbb{R}^D$ with entries $y_j$ given by

$$y_j = |\langle \mathbf{a}_j, \mathbf{x}_0 \rangle|^2 + \eta_j, \qquad j = 1, \dots, D. \tag{1}$$

Here the measurement vectors $\mathbf{a}_j \in \mathbb{C}^d$ are known and the scalars $\eta_j \in \mathbb{R}$ denote noise terms. This problem is known as the *phase retrieval problem* (see, e.g., [46, 36]), as we may think of the $|\cdot|^2$ in (1) as erasing the phases of the measurements $\langle \mathbf{a}_j, \mathbf{x}_0 \rangle$ in an otherwise linear system of equations.

The phase retrieval problem arises in many important signal acquisition schemes, including crystallography and ptychography (e.g., [36], see Figure 1), diffraction imaging [23], and optics [36, 46], among many others. Due to the breadth and importance of the applications, there has been significant interest in developing efficient algorithms to solve this problem. Indeed, one of the first algorithms proposed came in the early 1970's with the work of Gerchberg and Saxton [23]. Since then many variations of their method have been proposed (e.g, [21, 41, 43, 42, 8, 7, 20]) and used widely in practice. On the other hand – until recently – there have not been theoretical guarantees concerning the conditions under which these algorithms recover the underlying signal and the extent to which they can tolerate measurement error. Nevertheless, starting in 2006 a growing body of work (e.g., [2, 4, 5, 10, 14, 19, 30, 33]) has emerged, proposing new methods with theoretical performance guarantees under various assumptions on the signal $\mathbf{x}_0$ and the measurement vectors $\mathbf{a}_j$. Unfortunately, the assumptions (especially on the measurement vectors) often do not correspond to the setups used in practice. In particular, the mathematical analysis often requires that the measurement vectors be random or generic (e.g., [4, 5, 14]) while in practice the measurement vectors are a deterministic aspect of the imaging apparatuses employed. A main contribution

Mark A. Iwen: Department of Mathematics, and Department of Computational Mathematics, Science, and Engineering (CMSE), Michigan State University, East Lansing, MI, 48824, USA (markiwen@math.msu.edu).

B. Preskitt: Department of Mathematics, University of California San Diego, La Jolla, CA 92093, USA (bpreskitt@ucsd.edu)

R. Saab: Department of Mathematics, University of California San Diego, La Jolla, CA 92093, USA (rsaab@ucsd.edu).

A. Viswanathan: Department of Mathematics and Statistics, University of Michigan – Dearborn, Dearborn, MI, 48128, USA (adityavv@umich.edu).

of this paper is analyzing a construction that more closely matches practicable and deterministic measurement schemes. We propose a two-stage algorithm for solving the phase retrieval problem in this setting and we analyze our method, providing upper bounds on the associated reconstruction error.

In short, we provide theoretical error guarantees for a numerically efficient reconstruction algorithm in a measurement setting that closely resembles measurements used in practice.

1.1. **Local Correlation Measurements.** Consider the case where the vectors $\mathbf{a}_j$ represent shifts of compactly-supported vectors $\mathbf{m}_j, j = 1, \dots, K$ for some $K \in \mathbb{N}$. Using the notation $[n]_k := \{k, \dots, k+n-1\} \subset \mathbb{N}$, and defining $[n] := [n]_1$ we take $\mathbf{x}_0, \mathbf{m}_j \in \mathbb{C}^d$ with $\operatorname{supp}(\mathbf{m}_j) \subset [\delta] \subset [d]$ for some $\delta \in \mathbb{N}$. We also denote the space of Hermitian matrices in $\mathbb{C}^{k\times k}$ by $\mathcal{H}^k$. Now we have measurements of the form

$$(\mathbf{y}_\ell)_j = |\langle \mathbf{x}_0, S_\ell^* \mathbf{m}_j \rangle|^2, \quad (j,\ell) \in [K] \times P, \tag{2}$$

where $P \subset [d]_0$ is arbitrary and $S_\ell : \mathbb{C}^d \to \mathbb{C}^d$ is the discrete circular shift operator, namely

$$(S_\ell \mathbf{x}_0)_j = (\mathbf{x}_0)_{\ell+j}.$$

One can see that (2) represents the modulus squared of the correlation between $\mathbf{x}_0$ and locally supported measurement vectors. Therefore, we refer to the entries of $\mathbf{y}$ as local correlation measurements. Following [14, 30, 4], the problem may be lifted to a linear system on the space of $\mathbb{C}^{d\times d}$ matrices. In particular, we observe that

$$\begin{aligned}(\mathbf{y}_\ell)_j &= |\langle S_\ell \mathbf{x}_0, \mathbf{m}_j \rangle|^2 = \mathbf{m}_j^* (S_\ell \mathbf{x}_0)(S_\ell \mathbf{x}_0)^* \mathbf{m}_j \\ &= \langle \mathbf{x}_0 \mathbf{x}_0^*, S_\ell^* \mathbf{m}_j \mathbf{m}_j^* S_\ell \rangle,\end{aligned}$$

where the inner product above is the Hilbert-Schmidt inner product. Restricting to the case $P = [d]_0$, for every matrix $A \in \operatorname{span}\{S_\ell^* \mathbf{m}_j \mathbf{m}_j^* S_\ell\}_{\ell,j}$ we have $A_{ij} = 0$ whenever $|i-j| \mod d \geq \delta$. Therefore, we introduce the family of operators $T_k : \mathbb{C}^{d\times d} \to \mathbb{C}^{d\times d}$ given by

$$T_k(A)_{ij} = \begin{cases} A_{ij}, & |i-j| \mod d < k \\ 0, & \text{otherwise.} \end{cases}$$

Note that $T_\delta$ is simply the orthogonal projection operator onto its range $T_\delta(\mathbb{C}^{d\times d}) \supseteq \operatorname{span}\{S_\ell^* \mathbf{m}_j \mathbf{m}_j^* S_\ell\}_{\ell,j}$; therefore,

$$(\mathbf{y}_\ell)_j = \langle \mathbf{x}_0 \mathbf{x}_0^*, S_\ell^* \mathbf{m}_j \mathbf{m}_j^* S_\ell \rangle = \langle T_\delta(\mathbf{x}_0 \mathbf{x}_0^*), S_\ell^* \mathbf{m}_j \mathbf{m}_j^* S_\ell \rangle, \quad (j,\ell) \in [K] \times P. \tag{3}$$

For convenience, we set $D := K|P|$ and define the map $\mathcal{A} : \mathbb{C}^{d\times d} \to \mathbb{C}^D$

$$\mathcal{A}(X) = [\langle X, S_\ell^* \mathbf{m}_j \mathbf{m}_j^* S_\ell \rangle]_{(\ell,j)}. \tag{4}$$

Sometimes, we consider $\mathcal{A}|_{T_\delta(\mathbb{C}^{d\times d})}$, the restriction of $\mathcal{A}$ to the domain $T_\delta(\mathbb{C}^{d\times d})$; indeed, if this linear system is injective on $T_\delta(\mathbb{C}^{d\times d})$, then we can readily solve for

$$T_\delta(\mathbf{x}_0 \mathbf{x}_0^*) =: X_0 \tag{5}$$

using our measurements $(\mathbf{y}_\ell)_j = (\mathcal{A}(\mathbf{x}_0 \mathbf{x}_0^*))_{(\ell,j)}$. In [30], deterministic masks $\mathbf{m}_j$ were constructed for which (3) was indeed invertible for certain choices of $K$ and $P$. An additional construction is given below in §2.

Improving on [30], we can further see that $\mathbf{x}_0$ can be deduced from $X_0$ up to a global phase in the noiseless case as follows: First, $X_0$ immediately gives the magnitudes of the entries of $\mathbf{x}_0$ since $(X_0)_{ii} = |(x_0)_i|^2$. The only challenge remaining, therefore, is to find $\arg((x_0)_i)$ up to a global phase. We proceed by defining $\tilde{\mathbf{x}}_0$ and $\widetilde{X}_0$ by

$$(\tilde{x}_0)_i = \operatorname{sgn}((x_0)_i)$$

$$(\widetilde{X}_0)_{ij} = \begin{cases} \operatorname{sgn}((X_0)_{ij}), & |i-j| \mod d < \delta \\ 0, & \text{otherwise} \end{cases},$$

where $\operatorname{sgn} : \mathbb{C} \to \mathbb{C}$ is the usual normalization mapping

$$\operatorname{sgn}(z) = \begin{cases} \frac{z}{|z|}, & z \neq 0 \\ 1, & \text{otherwise} \end{cases}.$$

We emphasize that

$$\widetilde{X}_0 = \frac{X_0}{|X_0|} = \frac{T_\delta(\mathbf{x}_0\mathbf{x}_0^*)}{|T_\delta(\mathbf{x}_0\mathbf{x}_0^*)|} \text{ and } \widetilde{\mathbf{x}}_0 = \frac{\mathbf{x}_0}{|\mathbf{x}_0|}, \tag{6}$$

where the divisions are taken component-wise. Indeed, in [45], it was shown that the phases of the entries of $\mathbf{x}_0$ (up to a global phase) are given by the leading eigenvector of $\widetilde{X}_0$. Moreover, it was shown that this leading eigenvector is unique. Lemma 2 of this paper improves in these results by giving a lower bound on the gap between the top two eigenvalues of $\widetilde{X}_0$. This better understanding of the spectrum of $\widetilde{X}_0$ is then leveraged to analyze the robustness of this eigenvector-based phase retrieval method to measurement noise.

1.2. **Contributions.** In this paper, we analyze a phase retrieval algorithm (Algorithm 1) for estimating a vector $\mathbf{x}_0$ from noisy localized measurements of the form

$$(\mathbf{y}_\ell)_j = |\langle \mathbf{x}_0, S_\ell^* \mathbf{m}_j \rangle|^2 + n_{j\ell}, \quad (j,\ell) \in [2\delta - 1] \times [d]_0. \tag{7}$$

This algorithm is composed of two main stages. First, we apply the inverse of the linear operator

$$\mathcal{A}|_{T_\delta(\mathbb{C}^{d\times d})} : T_\delta(\mathbb{C}^{d\times d}) \to \mathbb{C}^{(2\delta-1)d}$$

defined immediately after (4), to obtain a Hermitian estimate $X$ of $T_\delta(\mathbf{x}_0\mathbf{x}_0^*)$ given by

$$X = \left((\mathcal{A}|_{T_\delta(\mathbb{C}^{d\times d})})^{-1}\mathbf{y}\right)/2 + \left((\mathcal{A}|_{T_\delta(\mathbb{C}^{d\times d})})^{-1}\mathbf{y}\right)^*/2 \in T_\delta(\mathbb{C}^{d\times d}). \tag{8}$$

In particular, our choice of $\mathbf{m}_j$ as described in Section 2 ensures that $\mathcal{A}|_{T_\delta(\mathbb{C}^{d\times d})}$ is both invertible and well conditioned. Next, once we have an approximation of $T_\delta(\mathbf{x}_0\mathbf{x}_0^*)$, we estimate the magnitudes and phases of the entries of $\mathbf{x}_0$ separately.

For the magnitudes, we simply use the square-roots of the diagonal entries of $X$. For the phases, we use the normalized eigenvector corresponding to the top eigenvalue of

$$\widetilde{X} := \frac{X}{|X|}, \tag{9}$$

where the operations are considered element-wise. The hope is that the leading eigenvector of $\widetilde{X}$ will serve as a good approximation to the leading eigenvector of $\widetilde{X}_0$, which is seen in Section 3 (see also [45]) to indeed be a scaled version of the phase vector $\widetilde{\mathbf{x}}_0$ (up to a global phase ambiguity). The entire method is summarized in Algorithm 1, and its associated recovery guarantees are presented in Theorem 1, while its computational complexity is discussed after the theorem in §1.3. Here and throughout the paper, $\mathbb{e} = 2.71828\ldots$ refers to the base of the natural logarithm and $\mathbb{i}$ refers to the imaginary unit.

**Theorem 1.** *Suppose $\delta > 2$ and $d \geq 4\delta$. Let $(x_0)_{\min} := \min_j |(x_0)_j|$ be the smallest magnitude of any entry in $\mathbf{x}_0 \in \mathbb{C}^d$. Then, the estimate $\mathbf{x}$ produced in Algorithm 1 satisfies*

$$\min_{\theta\in[0,2\pi]} \left\| \mathbf{x}_0 - \mathbb{e}^{\mathbb{i}\theta}\mathbf{x} \right\|_2 \leq C \left( \frac{\|\mathbf{x}_0\|_\infty}{(x_0)_{\min}^2} \right) \left( \frac{d}{\delta} \right)^2 \kappa \|\mathbf{n}\|_2 + C d^{\frac{1}{4}} \sqrt{\kappa \|\mathbf{n}\|_2},$$

*where $\kappa > 0$ is the condition number of the system (8) and $C \in \mathbb{R}^+$ is an absolute universal constant.*

**Algorithm 1** Fast Phase Retrieval from Local Correlation Measurements

**Input:** Measurements $\mathbf{y} \in \mathbb{R}^D$ as per (7)
**Output:** $\mathbf{x} \in \mathbb{C}^d$ with $\mathbf{x} \approx \mathbb{e}^{-\mathbb{i}\theta}\mathbf{x}_0$ for some $\theta \in [0, 2\pi]$

1: Compute the Hermitian matrix $X = \left((\mathcal{A}|_{T_\delta(\mathbb{C}^{d\times d})})^{-1}\mathbf{y}\right)/2 + \left((\mathcal{A}|_{T_\delta(\mathbb{C}^{d\times d})})^{-1}\mathbf{y}\right)^*/2 \in T_\delta(\mathbb{C}^{d\times d})$ as an estimate of $T_\delta(\mathbf{x}_0\mathbf{x}_0^*)$
2: Form the banded matrix of phases, $\tilde{X} \in T_\delta(\mathbb{C}^{d\times d})$, by normalizing the non-zero entries of $X$
3: Compute the top eigenvector $u \in \mathbb{C}^d$ of $\tilde{X}$ and set $\widetilde{\mathbf{x}} := \operatorname{sgn}(u)$.
4: Set $x_j = \sqrt{X_{j,j}} \cdot (\tilde{x})_j$ for all $j \in [d]$ to form $\mathbf{x} \in \mathbb{C}^d$

Theorem 1, which deterministically depends on both the masks and the signal, provides improvements over the first deterministic theoretical robust recovery guarantees proven in [30] for a wide class of non-vanishing signals. Momentarily consider, e.g., the class of "flat" vectors $\mathbf{x}_0 \in \mathbb{C}^d$ for which both $(i)$ $(x_0)_{\min} \geq \frac{\|\mathbf{x}_0\|_2}{2\sqrt{d}}$, and $(ii)$ $\left(\frac{\|\mathbf{x}_0\|_\infty}{(x_0)_{\min}^2}\right) \leq \tilde{C}$ for some absolute constant $\tilde{C} \in \mathbb{R}^+$, hold. The main deterministic result of [30] also applies to this class of vectors and states that an algorithm exists which can achieve the following robust recovery guarantee.

**Theorem 2** (See Theorem 5 in [30])**.** *There exist fixed universal constants $C, C' \in \mathbb{R}^+$ such that the following holds for all $\mathbf{x}_0 \in \mathbb{C}^d$ of the class mentioned above: Let $\|\mathbf{x}_0\|_2 \geq Cd\sqrt{(\delta-1)\|\mathbf{n}\|_2}$. Then, the algorithm in* [30]*, when provided with noisy measurements of $\mathbf{x}_0 \in \mathbb{C}^d$ (7) resulting from the masks discussed in Example 1 of §2, will output a vector $\mathbf{x} \in \mathbb{C}^d$ satisfying*

$$\min_{\theta\in[0,2\pi]} \left\|\mathbf{x}_0 - \mathbb{e}^{\mathbb{i}\theta}\mathbf{x}\right\|_2 \leq C'd\sqrt{(\delta-1)\|\mathbf{n}\|_2}.$$

Comparing the error bounds provided by Theorems 1 and 2 for the class of flat vectors $\mathbf{x}_0$ mentioned above when using measurements resulting from the masks discussed in Example 1 of §2,[1] we can see that Theorem 1 makes the following improvements over Theorem 2:

- Theorem 1 improves on the error bound of Theorem 2 for arbitrary small-norm noise $\mathbf{n}$ having $\|\mathbf{n}\|_2 = \mathcal{O}\left(\delta/d^2\right)$.
- Theorem 2's error bound breaks down entirely for noise $\mathbf{n}$ with $\ell^2$-norm on the order of $\|\mathbf{n}\|_2 = \Theta\left(\frac{\|\mathbf{x}_0\|_2^2}{\delta d^2}\right)$. Theorem 1's error bound, on the other hand, still provides non-trivial error guarantees for such noise levels as long as $\|\mathbf{x}_0\|_2 = \mathcal{O}(\delta)$.[2]

In addition, Theorem 1 also applies to a more general set of masks and a larger class of signals $\mathbf{x}_0$ than Theorem 2 does. And, perhaps most importantly, Algorithm 1 generally outperforms the algorithm referred to by Theorem 2 numerically for all noise levels (see, e.g., Figures 2a and 6a in §6). Theorem 1 provides theoretical error guarantees for this numerically improved method.

We note that the $\mathcal{O}\left(\left(\frac{d}{\delta}\right)^2\right)$-factor in the first term of the error bound provided by Theorem 1 is probably suboptimal, especially in practice. Indeed, Theorem 1 provides a worst-case error guarantee that holds for any arbitrary (including worst-case/adversarial) perturbation $\mathbf{n}$ of the measurements (7). However, while the quadratic dependence on $d/\delta$ is probably suboptimal, *some* dependence on $d/\delta$ likely exists for worst-case additive-noise in our local measurement setting. There is, e.g., numerical evidence that the empirical noise robustness of many phase retrieval

[1] Note that the condition number $\kappa$ mentioned in Theorem 1 for the masks discussed in Example 1 of §2 is $\mathcal{O}(\delta^2)$. See Theorem 3 below for a more exact statement. All asymptotic notation is with respect to $d \to \infty$. Below $\delta$ is always assumed to be independent of (and less than) $d$ unless otherwise noted.

[2] This allows Theorem 1 to cover, e.g., the case of larger $\delta = \Omega\left(\sqrt{d}\right)$.

methods deteriorates as $d$ grows for local measurements whose support size $\delta$ is held fixed. We will leave a rigorous theoretical investigation of the optimal scaling of such noise robustness guarantees with $d/\delta$ to future work. For the time being we will simply note here that the $\mathcal{O}\left(\left(\frac{d}{\delta}\right)^2\right)$-factor is mainly a product of the relatively small eigenvalue gap of the matrix $\widetilde{X}_0$ defined in (6) above. See §3 below for more details.

1.3. **The Runtime Complexity of Algorithm 1.** Consider now the computational complexity of Algorithm 1 (assuming, of course, that $\mathcal{A}|_{T_\delta(\mathbb{C}^{d\times d})}$ is actually invertible). One can see that line 1 can always be done in at most $\mathcal{O}(d\cdot\delta^3+\delta\cdot d\log d)$ flops using a block circulant matrix factorization approach (see Section 3.1 in [30]). In certain cases one can improve on this; for example, the second (new) mask construction of Section 2 allows line 1 to be performed in only $\mathcal{O}(d\cdot\delta)$ flops. Even in the worst case, however, if one precomputes this block circulant matrix factorization in advance given the masks $\mathbf{m}_j$ then line 1 can always be done in $\mathcal{O}(d\cdot\delta^2+\delta\cdot d\log d)$ flops thereafter.

The top eigenvector $\tilde{\mathbf{x}}$ of $\tilde{X}$ is guaranteed to be found in line 3 of Algorithm 1 in the low-noise (e.g., noiseless) setting via the shifted inverse power method with shift $\mu := 2\delta-1$ and initial vector $\mathbf{e}_1$ (the first standard basis vector). More generally, one may utilize the Rayleigh quotient iteration with the initial eigenvalue estimate fixed to $2\delta-1$ for the first few iterations. In either case, each iteration can be accomplished with $\mathcal{O}(d\cdot\delta^2)$ flops due to the banded structure of $\tilde{X}$ (see, e.g., [44]). In the low-noise setting the top eigenvector $\tilde{\mathbf{x}}$ can be computed to machine precision in $\mathcal{O}(\log d)$ such iterations,[3] for a total flop count of $\mathcal{O}(\delta^2\cdot d\log d)$ for line 3 in that case. In total, then, one can see that Algorithm 1 will always require just $\mathcal{O}(\delta^2\cdot d\log d+d\cdot\delta^3)$ total flops in low-noise settings. Furthermore, in all such settings a measurement mask support of size $\delta=\mathcal{O}(\log d)$ appears to suffice.

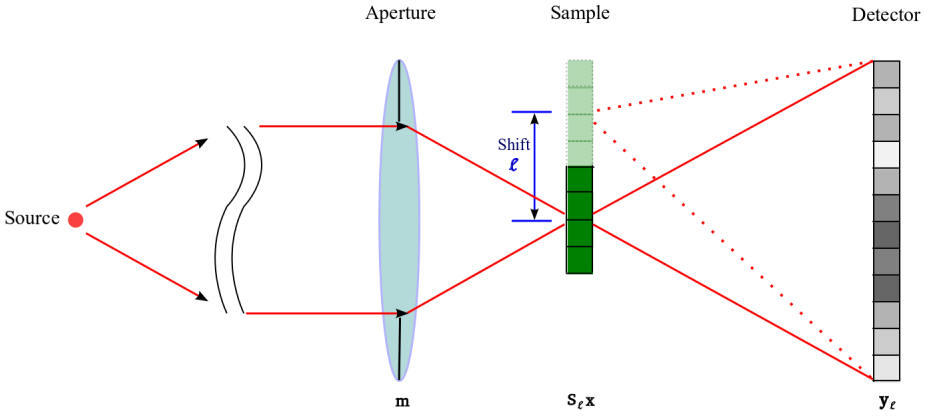



FIGURE 1. Illustration of one-dimensional ptychographic imaging (Adapted from "Fly-scan ptychography", Huang et al., Scientific Reports 5 (9074), 2015.)

[3]To see why $\mathcal{O}(\log d)$ iterations suffice one can appeal to lemmas 1 and 2 below. Let $|\lambda_1|>|\lambda_2|\geq\cdots\geq|\lambda_d|$ be the eigenvalues of $\tilde{X}$ with associated orthonormal eigenvectors $\mathbf{u}_j\in\mathbb{C}^d$. Let $\delta:=|\lambda_1|-|\lambda_2|>0$. When the noise level is sufficiently low (so that $\tilde{X}\approx\tilde{X}_0$) one will have both $(i)$ $|\mathbf{e}_1^*\mathbf{u}_j|=\Theta(1/\sqrt{d})$ $\forall j\in[d]$, and $(ii)$ $\mu\in(\lambda_1-\delta/4,\lambda_1+\delta/4)$ be true. Thus, we will have that there exists some unit norm $\mathbf{r}\in\mathbb{C}^d$ such that

$$\frac{\left(\tilde{X}-\mu I\right)^{-k}\mathbf{e}_1}{\left\|\left(\tilde{X}-\mu I\right)^{-k}\mathbf{e}_1\right\|_2}=\frac{\mathbf{u}_1+\sum_{j=2}^d\mathcal{O}\left(\left|\frac{\lambda_1-\mu}{\lambda_j-\mu}\right|^k\right)\mathbf{u}_j}{1+\mathcal{O}\left(\frac{d}{9^k}\right)}=\mathbf{u}_1+\mathcal{O}\left(\frac{d}{3^k}\right)\mathbf{r}$$

holds for any given integer $k=\Omega\left(\log_3 d\right)$.

1.4. **Connection to Ptychography.** In ptychographic imaging (see Fig. 1), small regions of a specimen are illuminated one at a time and an intensity[4] detector captures each of the resulting diffraction patterns. Thus each of the ptychographic measurements is a local measurement, which under certain assumptions (e.g., appropriate wavelength of incident radiation, far-field Fraunhofer approximation), can be modeled as [24, 17]

$$y(t,\omega) = \left|\mathcal{F}[\widetilde{h}\cdot S_t f](\omega)\right|^2 + \eta(t,\omega). \tag{10}$$

Here, $\mathcal{F}$ denotes the Fourier transform, $f:[0,1]\to\mathbb{C}$ represents the unknown test specimen, $S_t$ is the shift operator defined via

$$(S_t f)(s) := f(s+t),$$

and $\widetilde{h}:[0,1]\to\mathbb{C}$ is the so-called illumination function [47] of the imaging system. To account for the local nature of the measurements in (10), we assume that $\mathrm{supp}(\widetilde{h})\subset\mathrm{supp}(f)$.

As the phase retrieval problem is inherently non-linear and requires sophisticated computer algorithms to solve, consider the discrete version of (10), with $\widetilde{\mathbf{m}},\mathbf{x}_0\in\mathbb{C}^d$ discretizing $\widetilde{h}$ and $f$. Thus (10), in the absence of noise, becomes

$$(\mathbf{y}_\ell)_j = \left|\sum_{n=1}^{d}\widetilde{m}_n\,(x_0)_{n+\ell}\,\mathbb{e}^{-\frac{2\pi\mathbb{i}(j-1)(n-1)}{d}}\right|^2, \quad (j,\ell)\in[d]\times[d]_0, \tag{11}$$

where indexing is considered modulo-$d$, so $(\mathbf{y}_\ell)_j$ is a diffraction measurement corresponding to the $j^{th}$ Fourier mode of a circular $\ell$-shift of the specimen. We use circular shifts for convenience and we remark that this is appropriate as one can zero-pad $\mathbf{x}_0$ and $\widetilde{\mathbf{m}}$ in (11) and obtain the same $(\mathbf{y}_\ell)_j$ as one would with non-circular shifts. In practice, one may not need to use all the shifts $\ell\in[d]_0$ as a subset may suffice. Defining $\mathbf{m}_j\in\mathbb{C}^d$ by

$$(\mathbf{m}_j)_n = \overline{\widetilde{m}_n}\,\mathbb{e}^{\frac{2\pi\mathbb{i}(j-1)(n-1)}{d}} \tag{12}$$

and rearranging (11), we obtain

$$\begin{aligned}(\mathbf{y}_\ell)_j &= \left|\sum_{n=1}^{d}(x_0)_{n+\ell}\,\overline{(\mathbf{m}_j)_n}\right|^2 = \left|\sum_{n=1}^{\delta}(x_0)_{n+\ell}\,\overline{(\mathbf{m}_j)_n}\right|^2 \\ &= |\langle S_\ell\mathbf{x}_0,\mathbf{m}_j\rangle|^2 = \langle S_\ell\mathbf{x}_0\mathbf{x}_0^*S_\ell^*,\mathbf{m}_j\mathbf{m}_j^*\rangle \\ &= \langle T_\delta(\mathbf{x}_0\mathbf{x}_0^*),S_\ell^*\mathbf{m}_j\mathbf{m}_j^*S_\ell\rangle, \quad (j,\ell)\in[d]\times[d]_0\end{aligned} \tag{13}$$

where the second and last equalities follow from the fact that $\widetilde{\mathbf{m}}$ (and hence each $\mathbf{m}_j$) is locally supported. We note that (13) defines a correlation with local masks or window functions $\mathbf{m}_j$. More importantly, (13) shows that ptychography (with $\ell$ ranging over any subset of $[d]_0$) represents a case of the general system seen in (3).

1.5. **Connections to Masked Fourier Measurements.** Often, in imaging applications involving phase retrieval, a mask is placed either between the illumination source and the sample or between the sample and the sensor. Here, we will see that the mathematical setup that we consider is applicable in this scenario, albeit when the masks are band-limited. As before, let $\mathbf{x}_0,\mathbf{m}\in\mathbb{C}^d$ denote the unknown signal of interest, and a known mask (or window), respectively. Moreover, for a vector $\mathbf{x}_0\in\mathbb{C}^d$ we denote its discrete Fourier transform $\widehat{\mathbf{x}_0}\in\mathbb{C}^d$ by

$$(\widehat{x_0})_k := \sum_{n=1}^{d}(x_0)_n e^{-2\pi\mathbb{i}(n-1)(k-1)/d}.$$

[4]By intensity, we mean magnitude squared.

Here, we consider squared magnitude *windowed Fourier transform* measurements of the form

$$(\mathbf{y}_\ell)_k = \left| \sum_{n=1}^{d} (x_0)_n \, m_{n-\ell} \, \mathbb{e}^{-\frac{2\pi \mathbb{i}(k-1)(n-1)}{d}} \right|^2, \ k \in [d], \ \ell \in \{\ell_1, \dots, \ell_L\} \subset [d]_0. \tag{14}$$

As before, $\ell$ denotes a shift or translation of the mask/window, so $(\mathbf{y}_\ell)_k$ corresponds to the (squared magnitude of) the $k^{th}$ Fourier mode associated with an $\ell$-shift[5] of the mask $\mathbf{m}$. Defining the modulation operator, $W_k : \mathbb{C}^d \mapsto \mathbb{C}^d$, by its action $(W_k \mathbf{x}_0)_n = e^{2\pi \mathbb{i}(k-1)(n-1)/d} \ (x_0)_n$ and applying elementary Fourier transform properties[6] one has

$$\begin{aligned}
(\mathbf{y}_\ell)_k &= |\langle \mathbf{x}_0, S_{-\ell}(e^{2\pi \mathbb{i}(k-1)\ell/d} W_k \overline{\mathbf{m}}) \rangle|^2 \\
&= |\langle \mathbf{x}_0, S_{-\ell}(W_k \overline{\mathbf{m}}) \rangle|^2 = |\langle \widehat{\mathbf{x}_0}, \widehat{S_{-\ell}(W_k \overline{\mathbf{m}})} \rangle|^2 \\
&= |\langle \widehat{\mathbf{x}_0}, W_{-\ell+1}(S_{-k+1} \widehat{\overline{\mathbf{m}}}) \rangle|^2 \\
&= |\langle \widehat{\mathbf{x}_0}, S_{-k+1}(W_{-\ell+1} \widehat{\overline{\mathbf{m}}}) \rangle|^2. \qquad (15)
\end{aligned}$$

Defining $\widehat{\mathbf{m}}_\ell := W_{-\ell+1} \widehat{\overline{\mathbf{m}}}$ and assuming that $\mathrm{supp}(\widehat{\overline{\mathbf{m}}}) \subset [\delta]$ (e.g., assuming that $\mathbf{m}$ is real-valued and band-limited), we now have that

$$\begin{aligned}
(\mathbf{y}_\ell)_k &= \langle \widehat{\mathbf{x}}_0 \widehat{\mathbf{x}}_0^*, S_{-k+1} \widehat{\mathbf{m}}_\ell \widehat{\mathbf{m}}_\ell^* S_{-k+1}^* \rangle \\
&= \langle T_\delta(\widehat{\mathbf{x}}_0 \widehat{\mathbf{x}}_0^*), S_{-k+1} \widehat{\mathbf{m}}_\ell \widehat{\mathbf{m}}_\ell^* S_{-k+1}^* \rangle,
\end{aligned}$$

which again represents a case of the general system seen in (3). Moreover, our results all hold for this setting, albeit with the Fourier transforms of signals and conjugated masks.

1.6. **Related Work.** The first approach to the phase retrieval problem was proposed in the 1970's in [23] by Gerchberg and Saxton, where the measurement data corresponded to knowing the magnitude of both the image $\mathbf{x}_0$ and its Fourier transform. This result was famously expanded upon by Fienup [21] later that decade, one significant improvement being that only the magnitude of the Fourier transform of $\mathbf{x}_0$ must be known in the case of a signal $\mathbf{x}_0$ belonging to some fixed convex set $\mathcal{C}$ (typically, $\mathcal{C}$ is the set of non-negative, real-valued signals restricted to a known domain). Though these techniques work well in practice and have been popular for decades, they are notoriously difficult to analyze (see, e.g., [41, 43, 42, 8, 7, 20]). These are iterative methods that work by improving an initial guess until they stagnate. Recently Marchesini et al. proved that alternating projection schemes using generic measurements are guaranteed to converge to the correct solution *if provided with a sufficiently accurate initial guess* and algorithms for ptychography were explored in particular [34]. However, no global recovery guarantees currently exist for alternating projection techniques using local measurements (i.e., finding a sufficiently accurate initial guess is not generally easy).

Other authors have taken to proving probabilistic recovery guarantees when provided with globally supported Gaussian measurements. Methods for which such results exist vary in their approach, and include convex relaxations [11, 14], gradient descent strategies [13], graph-theoretic [1] and frame-based approaches [3, 10], and variants on the alternating minimization (e.g., with resampling) [37].

Several recovery algorithms achieve theoretical recovery guarantees while using at most $D = \mathcal{O}(d \log^4 d)$ masked Fourier coded diffraction pattern measurements, including both *PhaseLift* [12, 27], and *Wirtinger Flow* [13]. However, these measurements are both randomized (which is crucial

[5] As above, all indexing and shifts are considered modulo-$d$.

[6] $\widehat{S_\ell \mathbf{x}_0} = W_{\ell+1} \widehat{\mathbf{x}_0}$, $\widehat{W_k \mathbf{x}_0} = S_{-k+1} \widehat{\mathbf{x}_0}$, and $W_k S_\ell \mathbf{x}_0 = e^{-2\pi \mathbb{i}(k-1)\ell/d} S_\ell W_k \mathbf{x}_0$.

to the probabilistic recovery guarantees developed for both PhaseLift and Wirtinger Flow – deterministic recovery guarantees do not exist for either method in the noisy setting), and provide global information about $\mathbf{x}_0$ from each measurement (i.e., the measurements are not locally supported).

Among the first treatments of local measurements are [18, 9] and [32], in which it is shown that STFT measurements with specific properties can allow (sparse) phase retrieval in the noiseless setting, and several recovery methods are proposed. Similarly, the phase retrieval approach from [1] was extended to STFT measurements in [39] in order to produce recovery guarantees in the noiseless setting. More recently, randomized robustness guarantees were developed for time-frequency measurements in [38]. However, no *deterministic* robust recovery guarantees have been proven in the noisy setting for any of these approaches. Furthermore, none of the algorithms developed in these papers are empirically demonstrated to be competitive numerically with standard alternating projection techniques for large signals when utilizing windowed Fourier and/or correlation-based measurements. In [30], the authors propose the measurement scheme developed in the current paper and prove the first deterministic robustness results for a different greedy recovery algorithm.

1.7. **Organization.** Section 2 discusses two collections of local correlation masks $\mathbf{m}_j$, one of which is novel and the other of which was originally studied in [30]. Most importantly, Section 2 shows that the recovery of $T_\delta(\mathbf{x}_0\mathbf{x}_0^*)$ from measurements associated with the proposed masks can be done stably in the presence of measurement noise. Moreover, since in the noisy regime, the leading eigenvector $\widetilde{\mathbf{x}}$ of $\widetilde{X}$ (associated with line 3 of Algorithm 1) will no longer correspond exactly to the true phases $\widetilde{\mathbf{x}}_0$, we are interested in a perturbation theory for the eigenvectors of $\widetilde{X}_0$. Intuitively, $\widetilde{\mathbf{x}}$ will be most accurate when the eigenvalue of $\widetilde{X}_0$ associated with $\widetilde{\mathbf{x}}_0$ is well separated from the rest of the eigenvalues and so, accordingly, Section 3 studies the spectrum of $\widetilde{X}_0$. Indeed, this eigenvalue is rigorously shown to control the stability of the top eigenvector of $\widetilde{X}_0$ with respect to noise, and Section 4 develops perturbation results concerning their top eigenvectors by adapting the spectral graph techniques used in [1]. Recovery guarantees for the proposed phase retrieval method are then compiled in Section 5. Numerical results demonstrating the accuracy, efficiency, and robustness of the proposed methods are finally provided in Section 6[7], while Section 7 contains some concluding remarks and avenues for further research. In Appendix A, we provide an alternate, weaker but easier to derive eigenvector perturbation result analogous to the one in Section 4 which may be of independent interest.

## 2. Well-conditioned measurement maps

Here, we present two example constructions for which the linear operator $\mathcal{A}|_{T_\delta(\mathbb{C}^{d\times d})}$ used in Step 1 of Algorithm 1 is well conditioned. Such constructions are crucial for the stability of the method to additive noise.

**Example 1:** In [30], a construction was proposed for the masks $\mathbf{m}_\ell$ in (2) that guarantees the stable invertibility of $\mathcal{A}$. This construction comprises windowed Fourier measurements with parameters $\delta \in \mathbb{Z}^+$ and $a \in [4, \infty)$ corresponding to the $2\delta - 1$ masks $\mathbf{m}_j \in \mathbb{C}^d$, $j = 1, ..., 2\delta - 1$ with entries given by

$$(\mathbf{m}_j)_n = \begin{cases} \frac{\mathrm{e}^{-n/a}}{\sqrt[4]{2\delta-1}} \cdot \mathrm{e}^{\frac{2\pi \mathrm{i}\cdot(n-1)\cdot(j-1)}{2\delta-1}} & \text{if } n \le \delta \\ 0 & \text{if } n > \delta \end{cases}. \tag{16}$$

Here, measurements using all shifts $\ell = 1, ..., d$ of each mask are taken. In the notation of (3), this corresponds to $K = 2\delta - 1$ and $P = [d]_0$, which yields $D = (2\delta - 1)d$ total measurements. By

[7] MATLAB code to run the BlockPR algorithm is available online at [31].

considering the basis $\{E_{ij}\}$ for $T_\delta(\mathbb{C}^{d\times d})$ given by

$$E_{i,j}(s,t) = \begin{cases} 1, & (i,j) = (s,t) \\ 0, & \text{otherwise} \end{cases}$$

it was shown in [30] that this system is both well conditioned and rapidly invertible. In particular, if $M'$ is the matrix representing the measurement mapping $\mathcal{A} : T_\delta(\mathbb{C}^{d\times d}) \to T_\delta(\mathbb{C}^{d\times d})$ with respect to the basis $\{E_{ij}\}$, the following estimates of the condition number and cost of inversion hold.

**Theorem 3** ([30])**.** *Consider measurements of the form* (16) *with* $a := \max\left\{4,\ \frac{\delta-1}{2}\right\}$. *Let* $M' \in \mathbb{C}^{D\times D}$ *be the matrix representing the measurement mapping* $\mathcal{A} : T_\delta(\mathbb{C}^{d\times d}) \to T_\delta(\mathbb{C}^{d\times d})$ *with respect to the basis* $\{E_{ij}\}$. *Then, the condition number of* $M'$ *satisfies*

$$\kappa\left(M'\right) \ < \ \max\left\{144\mathrm{e}^2,\ \frac{9\mathrm{e}^2}{4}\cdot(\delta-1)^2\right\},$$

*and the smallest singular value of* $M'$ *satisfies*

$$\sigma_{\min}\left(M'\right) > \frac{7}{20a}\cdot \mathrm{e}^{-(\delta+1)/a} > \frac{C}{\delta}$$

*for an absolute constant* $C \in \mathbb{R}^+$. *Furthermore,* $M'$ *can be inverted in* $\mathcal{O}\left(\delta \cdot d\log d\right)$*-time.*

This theorem indicates that one can both efficiently and stably solve for $\mathbf{x}_0\mathbf{x}_0^*$ using (3) with the measurements given in (16). This measurement scheme is also interesting because it corresponds to a ptychography system if we take the illumination function (i.e. the physical mask) in (11) to be $\widetilde{m}_n = \frac{\mathrm{e}^{-n/a}}{\sqrt[4]{2\delta-1}}$ and assume that $d = k(2\delta-1)$ for some $k \in \mathbb{N}$; in practice, this may be achieved by zero-padding the specimen. Then we may take the subset of the measurements (12) given by $j = (p-1)k+1,\ p \in [2\delta-1]$ to obtain the masks specified in (16). We also remark that in this setup, only one physical mask is required, as the index $j$ in (16) denotes the different frequencies observed in the Fourier domain at the sensor array.

**Example 2:** We provide a second deterministic construction that improves on the condition number of the previous collection of measurement vectors. We merely set $\mathbf{m}_1 = e_1, \mathbf{m}_{2j} = e_1 + e_{j+1}$, and $\mathbf{m}_{2j+1} = e_1 + ie_{j+1}$ for $j = 1,\dots,\delta-1$. A simple induction shows that $\{S_\ell \mathbf{m}_j \mathbf{m}_j^* S_\ell^*\}_{\ell\in[d]_0, j\in[2k-1]}$ is a basis for $T_k(\mathbb{C}^{d\times d})$, so if we take $\mathbf{m}_1,\dots,\mathbf{m}_{2\delta-1}$ for our masks we'll have a basis for $T_\delta(\mathbb{C}^{d\times d})$. Indeed, if we let

$$\mathcal{B} : T_k(\mathbb{C}^{d\times d}) \to \mathbb{C}^{\delta\times d}$$

be the measurement operator defined via

$$\left(\mathcal{B}(X)\right)_{\ell,j} = \langle S_\ell \mathbf{m}_j \mathbf{m}_j^* S_\ell^*, X\rangle, \quad (\ell, j) \in [d]_0 \times [2k-1]$$

we can immediately solve for the entries of $X \in T_k(\mathcal{H}^{d\times d})$ from $\mathcal{B}(X) =: B$ by observing that

$$\begin{aligned} X_{i,i} &= B_{i-1,1} \\ X_{i,i+k} &= \tfrac{1}{2}B_{i-1,2k} + \tfrac{i}{2}B_{i-1,2k+1} - \tfrac{1+i}{2}(B_{i-1,1} + B_{i+k-1,1}), \end{aligned}$$

where we naturally take the indices of $B$ mod $d$. This leads to an upper triangular system if we enumerate $X$ by its diagonals; namely we regard $T_\delta(\mathcal{H}^{d\times d})$ as a $d(2\delta-1)$ dimensional vector space over $\mathbb{R}$ and set, for $i \in [d]$

$$z_{kd+i} = \begin{cases} \mathrm{Re}(X_{i,i+k}), & 0 \le k < \delta \\ \mathrm{Im}(X_{i,i+k-\delta+1}), & \delta \le k < 2\delta-1 \end{cases}, \quad y_{kd+i} = \begin{cases} \mathcal{B}(X)_{i,1}, & k = 0 \\ \mathcal{B}(X)_{i,2k}, & 1 \le k < \delta \\ \mathcal{B}(X)_{i,2(k-\delta+1)+1}, & \delta \le k < 2\delta-1 \end{cases}.$$

Then with $S = S_1 \in \mathbb{R}^{d\times d}$ representing the circular shift operator as before, we have

$$y = \begin{bmatrix} I_d & 0 & 0 \\ D & 2I_{d(\delta-1)} & 0 \\ D & 0 & 2I_{d(\delta-1)} \end{bmatrix} z =: Cz, \text{ where } D = \begin{bmatrix} I_d + S \\ I_d + S^2 \\ \vdots \\ I_d + S^{\delta-1} \end{bmatrix}.$$

Since the matrix $C$ is upper triangular, its inverse is immediate:

$$C^{-1} = \begin{bmatrix} I_d & 0 & 0 \\ -D/2 & I_{d(\delta-1)}/2 & 0 \\ -D/2 & 0 & I_{d(\delta-1)}/2 \end{bmatrix}.$$

To ascertain the condition number of $\mathcal{B}$, then, all we need is the extremal singular values of $C$. We bound the top singular value by considering

$$\begin{aligned} \sigma_{\max}(C) = \max_{||w||^2+||v||^2=1} \left\| C \begin{bmatrix} w \\ v \end{bmatrix} \right\| &= \left\| \begin{bmatrix} w \\ Dw \\ Dw \end{bmatrix} + 2 \begin{bmatrix} 0 \\ v \end{bmatrix} \right\| \\ &\leq \sqrt{||w||^2 + 2||w+Sw||^2 + \cdots + 2||w+S^{\delta-1}w||^2} + ||2v|| \\ &\leq \sqrt{8(\delta-1)+1}||w|| + 2||v|| \leq \sqrt{8(\delta-1)+5} \leq 2\sqrt{2\delta}, \end{aligned}$$

where in the last line we have used $||w||^2 + ||v||^2 = 1$. By a nearly identical argument, we find

$$\frac{1}{\sigma_{\min}(C)} = \sigma_{\max}(C^{-1}) \leq \sqrt{2\delta}$$

so that the condition number is bounded by $\kappa(C) \leq 4\delta$.

## 3. The Spectrum of $\tilde{X}_0$

Consider line 3 of Algorithm 1, which shows that we are trying to recover $\widetilde{\mathbf{x}}_0 := \frac{\mathbf{x}_0}{|\mathbf{x}_0|}$ via an eigenvector method. Here, we show that $\widetilde{X}_0$ has $\widetilde{\mathbf{x}}_0$ as its top eigenvector and we investigate the spectral properties of $\widetilde{X}_0$ in this section, following the intuition that the eigenvalue gap $|\lambda_1 - \lambda_2|$ will affect the robustness of the spectral step in the algorithm.

For the remainder of the paper, we let $\mathbb{1}$ refer to a constant vector of all ones; its size will always be determined by context. To begin, consider $U = T_\delta(\mathbb{1}\mathbb{1}^*)$, i.e.,

$$U_{j,k} = \begin{cases} 1 & \text{if } |j-k| \bmod d < \delta \\ 0 & \text{otherwise} \end{cases}. \tag{17}$$

Observe that $U$ is circulant for all $\delta$, so its eigenvectors are always discrete Fourier vectors. Setting $\omega_j = \mathbb{e}^{2\pi\mathbb{i}\frac{j-1}{d}}$ for $j = 1, 2, \ldots, d$, one can also see that the eigenvalues of $U$ are given by

$$\nu_j = \sum_{k=1}^{d} (U)_{1,k}\omega_j^{k-1} = 1 + \sum_{k=1}^{\delta-1} \omega_j^k + \omega_j^{-k} \;=\; 1 + 2\sum_{k=1}^{\delta-1} \cos\left(\frac{2\pi(j-1)k}{d}\right), \tag{18}$$

for all $j = 1, \ldots, d$. In particular, $\nu_1 = 2\delta - 1$. Set $\Lambda = \operatorname{diag}\{\nu_1, \ldots, \nu_d\}$ and let $F$ denote the unitary $d \times d$ discrete Fourier matrix with entries

$$F_{j,k} := \frac{1}{\sqrt{d}} \mathbb{e}^{2\pi\mathbb{i}\frac{(j-1)(k-1)}{d}},$$

then $U = F\Lambda F^*$.

We consider that $\widetilde{X}_0$ and $U$ are similar; indeed $\widetilde{X}_0 = \widetilde{D}_0 U \widetilde{D}_0^*$, where $\widetilde{D}_0 = \operatorname{diag}\{(\widetilde{x}_0)_1, \ldots, (\widetilde{x}_0)_d\}$. Since $|(\widetilde{\mathbf{x}}_0)_j| = 1$ for each $j$, we have that $\widetilde{D}_0$ is unitary. Thus the eigenvalues of $\widetilde{X}_0$ are given by

(18), and its eigenvectors are simply the discrete Fourier vectors modulated by the entries of $\widetilde{\mathbf{x}}_0$. We now have the following lemma.

**Lemma 1.** *Let $\widetilde{X}_0$ be defined as in* (6)*. Then*

$$\widetilde{X}_0 = \widetilde{D}_0 F \Lambda F^* \widetilde{D}_0^*$$

*where $F$ is the unitary $d \times d$ discrete Fourier transform matrix, $\widetilde{D}_0$ is the $d \times d$ diagonal matrix* $\mathrm{diag}\{(\widetilde{x}_0)_1, \ldots, (\widetilde{x}_0)_d\}$*, and $\Lambda$ is the $d \times d$ diagonal matrix* $\mathrm{diag}\{\nu_1, \ldots, \nu_d\}$ *where*

$$\nu_j := 1 + 2\sum_{k=1}^{\delta-1} \cos\left(\frac{2\pi(j-1)k}{d}\right)$$

*for $j = 1, \ldots, d$.*

We next estimate the principal eigenvalue gap of $\widetilde{X}_0$. This information will be crucial to our understanding of the stability and robustness of Algorithm 1.

3.1. **The Spectral Gap of $\widetilde{X}_0$.** Set $\theta_j = \frac{2\pi j}{d}$ and begin by observing that, for any $\theta \in \mathbb{R}$,

$$\sum_{k=1}^{\delta-1} \cos(\theta k) \;=\; \frac{1}{2}\left(\frac{\sin(\theta(\delta - 1/2))}{\sin(\theta/2)} - 1\right).$$

Accordingly, defining $l_\delta : \mathbb{R} \to \mathbb{R}$ by $l_\delta(\theta) := 1 + 2\sum_{k=1}^{\delta-1} \cos(\theta k)$ we have that

$$\nu_{j+1} \;=\; l_\delta(\theta_j) \;=\; \frac{\sin(\theta_j(\delta - 1/2))}{\sin(\theta_j/2)}. \tag{19}$$

Thus, the eigenvalues of $\widetilde{X}_0$ are sampled from the $(\delta-1)^{\text{st}}$ Dirichlet kernel. Of course, $\nu_1 = 2\delta - 1$ is the largest of these in magnitude, so the eigenvalue gap $\min_j \nu_1 - |\nu_j|$ is at most equal to

$$\begin{aligned}
\nu_1 - \nu_2 &= (2\delta - 1) - \frac{\sin(\pi/d(2\delta-1))}{\sin(\pi/d)} \\
&\leq (2\delta - 1) - \frac{\pi/d(2\delta-1) - \frac{1}{6}(\pi/d(2\delta-1))^3}{\pi/d} \\
&= \frac{1}{6}\left(\frac{\pi}{d}\right)^2 (2\delta-1)^3 \leq \frac{4\pi^2}{3}\frac{\delta^3}{d^2}.
\end{aligned}$$

Thus, $\nu_1 - |\nu_2| \lesssim \frac{\delta^3}{d^2}$. However, a lower bound on the spectral gap is more useful. The following lemma establishes that the spectral gap is indeed $\sim \frac{\delta^3}{d^2}$ for most reasonable choices of $\delta < d$.

**Lemma 2.** *Let $\nu_1 = 2\delta - 1, \nu_2, \ldots, \nu_d$ be the eigenvalues of $\widetilde{X}_0$. Then, there exists a positive absolute constant $C \in \mathbb{R}^+$ such that*

$$\min_{j \in \{2,3,\ldots,d\}} (\nu_1 - |\nu_j|) \geq C\frac{\delta^3}{d^2}$$

*whenever $d \geq 4\delta$ and $\delta \geq 3$.*

*Proof.* Let $\theta_j = \frac{2\pi j}{d}$. We find the lower bound by considering that $\theta_j \in [\pi/d, 2\pi - \pi/d]$ for every $j > 0$, so

$$\nu_1 - \max|\nu_j| \geq \nu_1 - \max_{\theta \in [\pi/d, 2\pi - \pi/d]} |l_\delta(\theta)| = (2\delta - 1) - \max_{\theta \in [\pi/d, \pi]} |l_\delta(\theta)|,$$

where we have used our eigenvalue formula from (19), and the symmetry of $l_\delta$ about $\theta = \pi$.

We now show that $l_\delta$ is decreasing towards its first zero at $\theta = \frac{2\pi}{2\delta-1}$ by considering the derivative

$$l'_\delta(\theta) = \frac{(\delta - 1/2)\cos((\delta-1/2)\theta)\sin(\theta/2) - 1/2\sin((\delta-1/2)\theta)\cos(\theta/2)}{\sin(\theta/2)^2},$$

which is non-positive if and only if

$$(2\delta-1)\sin(\theta/2)\cos((\delta-1/2)\theta) \leq \sin((\delta-1/2)\theta)\cos(\theta/2).$$

Since $\tan(\cdot)$ is convex on $[0,\, \pi/2)$, this last inequality will hold for $\theta \in [0,\, \frac{\pi}{2\delta-1})$. For $\theta \in [\frac{\pi}{2\delta-1}, \frac{2\pi}{2\delta-1})$, $\cos((\delta-1/2)\theta) \leq 0$ while the remainder of the terms are non-negative, so the inequality also holds. Therefore,

$$\nu_1 - \max_{j>1}|\nu_j| \geq (2\delta-1) - \max\left\{\nu_2, \max_{\theta\in[\frac{2\pi}{2\delta-1},\pi]}|l_\delta(\theta)|\right\},$$

which permits us to bound $(2\delta-1)-\nu_2$ and $(2\delta-1) - \max_{\theta\in[\frac{2\pi}{2\delta-1},\pi]}|l_\delta(\theta)|$ separately.

For $\max_{\theta\in[\frac{2\pi}{2\delta-1},\pi]}|l_\delta(\theta)|$, we simply observe that

$$\max_{\theta\in[\frac{2\pi}{2\delta-1},\pi]}|l_\delta(\theta)| \leq \max_{\theta\in[\frac{2\pi}{2\delta-1},\pi]}\frac{1}{\sin(\theta/2)} = \left(\sin\left(\frac{\pi}{2\delta-1}\right)\right)^{-1} \leq \frac{2\delta-1}{2},$$

where the last line uses that $\frac{\pi}{2\delta-1} \leq \pi/2$ (since $\delta \geq 3$). This yields $\nu_1 - \max_{\theta\in[\frac{2\pi}{2\delta-1},\pi]}|l_\delta(\theta)| \geq \frac{1}{2}(2\delta-1)$.

As for $\nu_2$, we have $\theta_1\cdot(\delta-1) \leq \pi/2$ (since $4(\delta-1)\leq d$). Thus, $\cos(\cdot)$ will be concave on $[0,\, \theta_1(\delta-1)]$. Considering (18), this will give $\sum_{k=1}^{\delta-1}\cos(k\theta_1) \leq (\delta-1)\cos\left(\theta_1\frac{\delta}{2}\right)$, so

$$\begin{aligned}
\nu_1 - \nu_2 &\geq 2(\delta-1)\left(1-\cos\left(\pi\tfrac{\delta}{d}\right)\right) \\
&\geq 2(\delta-1)\left(\frac{(\pi\frac{\delta}{d})^2}{4}\right) \\
&\geq \frac{\pi^2}{3}\cdot\frac{\delta^3}{d^2}.
\end{aligned}$$

The stated result follows. □

We are now sufficiently well informed about $\widetilde{X}_0$ to consider perturbation results for its leading eigenvector.

## 4. Perturbation Theory for $\tilde{X}_0$

In this section we will use spectral graph theoretic techniques to obtain a bound on the error associated with recovering phase information using our method. In particular, we will adapt the proof of Theorem 6.3 from [1] to develop a bound for $\min_{\theta\in[0,2\pi]}\|\widetilde{\mathbf{x}}_0 - \mathbb{e}^{\mathbb{i}\theta}\widetilde{\mathbf{x}}\|_2$. This approach involves considering both $\widetilde{X}$ from Algorithm 1 and $\widetilde{X}_0$ from (6) in the context of spectral graph theory, so we begin by defining essential terms. The idea is to consider a graph whose vertices correspond to the entries of $\widetilde{\mathbf{x}}_0$ from (6), and whose edges carry the relative phase data.[8]

[8]The interested reader is also referred to Appendix A where more standard perturbation theoretic techniques are utilized in order to obtain a weaker bound on the error associated with recovering phase information via the proposed approach.

We begin with an undirected graph $G = (V, E)$ with vertex set $V = \{1, 2, \ldots, d\}$ and weight mapping $w : V \times V \to \mathbb{R}^+$, where $w_{ij} = w_{ji}$ and $w_{ij} = 0$ iff $\{i, j\} \notin E$. The **degree** of a vertex $i$ is

$$\deg(i) := \sum_{j \text{ s.t. } (i,j)\in E} w_{ij},$$

and we define the **degree matrix** and **weighted adjacency matrix** of $G$ by

$$D := \operatorname{diag}(\deg(i)) \text{ and } W_{ij} := w_{ij},$$

respectively. The **volume** of $G$ is

$$\operatorname{vol}(G) := \sum_{i\in V} \deg(i).$$

Finally, the **Laplacian** of $G$ is the $d \times d$ real symmetric matrix

$$L := I - D^{-1/2} W D^{-1/2} = D^{-1/2}(D - W)D^{-1/2},$$

where $I \in \{0, 1\}^{d\times d}$ is the identity matrix.

When $G$ is connected, Lemma 1.7 of [15] shows that the nullspace of $(D - W)$ is $\operatorname{span}(\mathbb{1})$, and the nullspace of $L$ is $\operatorname{span}(D^{1/2}\mathbb{1})$. Observing that $D - W$ is diagonally semi-dominant, it follows from Gershgorin's disc theorem that $(D - W)$ and $L$ are both positive semidefinite. Alternatively, one may also note that

$$\mathbf{v}^*(D - W)\mathbf{v} = \sum_{i\in V}\left(v_i^2 \deg(i) - \sum_{j\in V} v_i v_j w_{ij}\right) = \frac{1}{2}\sum_{i,j\in V} w_{ij}(v_i - v_j)^2 \geq 0$$

holds for all $\mathbf{v} \in \mathbb{R}^d$. Thus, we may order the eigenvalues of $L$ in increasing order so that $0 = \lambda'_1 < \lambda'_2 \leq \cdots \leq \lambda'_n$. We then define the **spectral gap** of $G$ to be $\tau = \lambda'_2$.

Herein, though we will state the main theorem of this section more generally, we will only be interested in the case where the graph $G = (V, E)$ is the simple unweighted graph whose adjacency matrix is $U = T_\delta(\mathbb{1}\mathbb{1}^*)$ as in (17). In this case we will have $W = U$ and $D = (2\delta - 1)I$. We also immediately obtain the following corollary of Lemmas 1 and 2.

**Corollary 1.** *Let $G$ be the simple unweighted graph whose adjacency matrix is $U$ from* (17)*. Let $L$ be the Laplacian of $G$. Then, there exists a bijection $\sigma : [d] \to [d]$ such that*

$$\lambda'_{\sigma(j)} = 1 - \frac{1 + 2\sum_{k=1}^{\delta-1}\cos\left(\frac{2\pi(j-1)k}{d}\right)}{2\delta - 1}$$

*for $j = 1, \ldots, d$. In particular, if $d \geq 4(\delta - 1)$ and $\delta \geq 3$ then $\tau = \lambda'_2 > C'''\delta^2/d^2$ for an absolute constant $C''' \in \mathbb{R}^+$.*

Using this graph $G$ as a scaffold we can now represent our computed relative phase matrix $\widetilde{X}$ from Algorithm 1 by noting that for some (Hermitian) perturbations $\eta_{ij}$ we will have

$$\widetilde{X}_{ij} = \frac{(x_0)_i(x_0)_j^* + \eta_{ij}}{|(x_0)_i(x_0)_j^* + \eta_{ij}|} \cdot w_{ij} = \frac{(x_0)_i(x_0)_j^* + \eta_{ij}}{|(x_0)_i(x_0)_j^* + \eta_{ij}|} \cdot \chi_{E(i,j)}. \tag{20}$$

Using this same notation we may also represent our original phase matrix $\widetilde{X}_0$ via $G$ by noting that

$$(\widetilde{X}_0)_{ij} = \frac{(x_0)_i(x_0)_j^*}{|(x_0)_i(x_0)_j^*|} \cdot w_{ij} = \operatorname{sgn}\left((x_0)_i(x_0)_j^*\right) \cdot \chi_{E(i,j)}. \tag{21}$$

We may now define the **connection Laplacian** of the graph $G$ associated with the Hermitian and entrywise normalized data given by $\widetilde{X}$ to be the matrix

$$L_1 = I - D^{-1/2}(\widetilde{X} \circ W)D^{-1/2}, \tag{22}$$

where $\circ$ denotes entrywise (Hadamard) multiplication. Following [6], given $\widetilde{X}$ and a vector $\mathbf{y} \in \mathbb{C}^d$, we define the **frustration of y with respect to** $\widetilde{X}$ by

$$\eta_{\widetilde{X}}(\mathbf{y}) := \frac{1}{2} \frac{\sum_{(i,j)\in E} w_{ij}|y_i - \widetilde{X}_{ij} y_j|^2}{\sum_{i\in V} \deg(i)|y_i|^2} = \frac{\mathbf{y}^*(D - (\widetilde{X} \circ W))\mathbf{y}}{\mathbf{y}^* D \mathbf{y}}. \tag{23}$$

We may consider $\eta_{\widetilde{X}}(\mathbf{y})$ to measure how well $\mathbf{y}$ (viewed as a map from $V$ to $\mathbb{C}$) conforms to the computed relative phase differences $\widetilde{X}$ across the graph $G$.

In addition, we adapt a result from [6]:

**Lemma 3** (Cheeger inequality for the connection Laplacian)**.** *Suppose that $G = (V = [d], E)$ is a connected graph with degree matrix $D \in [0,\infty)^{d\times d}$, weighted adjacency matrix $W \in [0,\infty)^{d\times d}$, and spectral gap $\tau > 0$, and that $\widetilde{X} \in \mathbb{C}^{d\times d}$ is Hermitian and entrywise normalized. Let $\mathbf{u} \in \mathbb{C}^d$ be an eigenvector of $L_1$ from (22) corresponding to its smallest eigenvalue. Then, $\mathbf{w} = \operatorname{sgn}(\mathbf{u}) = \operatorname{sgn}\left(D^{-1/2}\mathbf{u}\right)$ satisfies*

$$\eta_{\widetilde{X}}(\mathbf{w}) \leq \frac{C'}{\tau} \cdot \min_{\mathbf{y}\in\mathbb{C}^d} \eta_{\widetilde{X}}(\operatorname{sgn}(\mathbf{y})),$$

*where $C' \in \mathbb{R}^+$ is a universal constant.*

*Proof.* One can see that

$$\begin{aligned}\inf_{\mathbf{v}\in\mathbb{C}^d\setminus\{\mathbf{0}\}} \frac{\mathbf{v}^* L_1 \mathbf{v}}{\mathbf{v}^*\mathbf{v}} &= \inf_{\mathbf{y}\in\mathbb{C}^d\setminus\{\mathbf{0}\}} \frac{(D^{1/2}\mathbf{y})^* L_1 (D^{1/2}\mathbf{y})}{(D^{1/2}\mathbf{y})^*(D^{1/2}\mathbf{y})} = \inf_{\mathbf{y}\in\mathbb{C}^d\setminus\{\mathbf{0}\}} \frac{\mathbf{y}^*(D - (\widetilde{X}\circ W))\mathbf{y}}{\mathbf{y}^* D \mathbf{y}} \\ &= \inf_{y\in\mathbb{C}^d\setminus\{\mathbf{0}\}} \eta_{\widetilde{X}}(\mathbf{y}) \leq \min_{\mathbf{y}\in\mathbb{C}^d} \eta_{\widetilde{X}}(\operatorname{sgn}(\mathbf{y})).\end{aligned}$$

From here, Lemma 3.6 in [6] gives

$$\eta_{\widetilde{X}}(\mathbf{w}) \leq \frac{44}{\tau}\eta_{\widetilde{X}}\left(D^{-1/2}\mathbf{u}\right) = \frac{44}{\tau} \cdot \inf_{\mathbf{v}\in\mathbb{C}^d\setminus\{\mathbf{0}\}} \frac{\mathbf{v}^* L_1 \mathbf{v}}{\mathbf{v}^*\mathbf{v}} \leq \frac{44}{\tau} \cdot \min_{\mathbf{y}\in\mathbb{C}^d} \eta_{\widetilde{X}}(\operatorname{sgn}(\mathbf{y})).$$

□

We now state the main result of this section:

**Theorem 4.** *Suppose that $G = (V = [d], E)$ is an undirected, connected, and unweighted graph (so that $W_{ij} = \chi_{E(i,j)}$) with spectral gap $\tau > 0$. Let $\mathbf{u} \in \mathbb{C}^d$ be an eigenvector of $L_1$ from (22) corresponding to its smallest eigenvalue, and let*

$$\widetilde{\mathbf{x}} = \operatorname{sgn}(\mathbf{u}) \text{ and } \widetilde{\mathbf{x}}_0 = \operatorname{sgn}(\mathbf{x}_0).$$

*Then for some universal constant $C \in \mathbb{R}^+$,*

$$\min_{\theta\in[0,2\pi]} ||\widetilde{\mathbf{x}} - \mathbb{e}^{\mathbb{i}\theta}\widetilde{\mathbf{x}}_0||_2 \leq C \frac{\|\widetilde{X} - \widetilde{X}_0\|_F}{\tau \cdot \sqrt{\min_{i\in V}(\deg(i))}},$$

*where $\widetilde{X}$ and $\widetilde{X}_0$ are defined as per (20) and (21), respectively.*

The proof follows by combining the two following lemmas, which share the hypotheses of the theorem. Additionally, we introduce the notation $\mathbf{g} \in \mathbb{C}^d$ and $\Lambda \in \mathbb{C}^{d\times d}$, where

$$g_i = (\widetilde{\mathbf{x}}_0)_i^* \widetilde{\mathbf{x}}_i \quad \text{and} \quad \Lambda_{ij} = (\widetilde{X}_0)_{ij}^* \widetilde{X}_{ij},$$

and observe that $|g_i| = |\Lambda_{ij}| = 1$ for each $(i,j) \in E$.

**Lemma 4.** *Under the hypotheses of Theorem 4, there exists an angle $\theta \in [0, 2\pi]$ such that*

$$\tau \sum_{i \in V} \deg(i)|g_i - \mathbb{e}^{\mathbb{i}\theta}|^2 \leq 2 \sum_{(i,j)\in E} |g_i - g_j|^2.$$

**Lemma 5.** *Under the hypotheses of Theorem 4, there exists an absolute constant $C$ such that*

$$2 \sum_{(i,j)\in E} |g_i - g_j|^2 \leq \frac{C}{\tau} \|\widetilde{X} - \widetilde{X}_0\|_F^2.$$

From these lemmas, the theorem follows immediately by observing $\sum_{i\in V} |g_i - \mathbb{e}^{\mathbb{i}\theta}|^2 = |\widetilde{\mathbf{x}} - \mathbb{e}^{\mathbb{i}\theta}\widetilde{\mathbf{x}}_0|_2^2$.

*Proof of Lemma 4.* We set $\alpha = \dfrac{\sum_{i\in V} \deg(i) g_i}{\mathrm{vol}(G)}$ and $w_i = g_i - \alpha$. Then

$$\mathbb{1}^* D\mathbf{w} = \sum_{i \in V} \deg(i)(g_i - \alpha) = 0,$$

so $D^{1/2}\mathbf{w}$ is orthogonal to $D^{1/2}\mathbb{1}$. Noting that the null space of $L$ is spanned by $D^{1/2}\mathbb{1}$ when $\tau > 0$, and recalling that $L \succeq 0$, we have

$$\frac{(D^{1/2}\mathbf{w})^* L (D^{1/2}\mathbf{w})}{\mathbf{w}^* D \mathbf{w}} \geq \min_{\mathbf{y}^* D^{1/2}\mathbb{1} = 0} \frac{\mathbf{y}^* L \mathbf{y}}{\mathbf{y}^* \mathbf{y}} = \tau.$$

Therefore,

$$\begin{aligned} \tau \mathbf{w}^* D \mathbf{w} &\leq \mathbf{w}^*(D - W)\mathbf{w} &&= \mathbf{g}^*(D - W)\mathbf{g} \\ &= \sum_{i\in V} \deg(i)|g_i|^2 - \sum_{i\in V} g_i^* \sum_{(i,j)\in E} g_j &&= \sum_{(i,j)\in E} (1 - g_i^* g_j) \\ &= \frac{1}{2} \sum_{(i,j)\in E} |g_i - g_j|^2. \end{aligned}$$

We note that $\tau \mathbf{w}^* D \mathbf{w} = \tau \sum_{i\in V} \deg(i)|g_i - \alpha|^2$, while we seek a bound on $\sum_{i\in V} \deg(i)|g_i - \mathbb{e}^{\mathbb{i}\theta}|^2$. To that end, we use the fact that $|g_i| = |\mathrm{sgn}(\alpha)| = 1$ to obtain

$$|g_i - \mathrm{sgn}(\alpha)| \leq |g_i - \alpha| + |\alpha - \mathrm{sgn}(\alpha)| \leq 2|g_i - \alpha|.$$

Setting $\theta := \arg \alpha$, we have the stated result. □

*Proof of Lemma 5.* Observe that for any two real numbers $a, b \in \mathbb{R}$, we have $\frac{1}{2}a^2 - b^2 \leq (a - b)^2$. Thus, by the reverse triangle inequality we have

$$\begin{aligned} \sum_{(i,j)\in E} \left( \frac{1}{2}|g_i - g_j|^2 - |\Lambda_{ij} - 1|^2 \right) &\leq \sum_{(i,j)\in E} \left(|g_i - g_j| - |\Lambda_{ij} - 1|\right)^2 \\ &\leq \sum_{(i,j)\in E} |g_i - \Lambda_{ij} g_j|^2 \\ &= \sum_{(i,j)\in E} |\widetilde{\mathbf{x}}_i - \widetilde{X}_{ij}\widetilde{\mathbf{x}}_j|^2 \\ &= 2\mathrm{vol}(G) \cdot \eta_{\widetilde{X}}(\widetilde{\mathbf{x}}), \end{aligned}$$

as the denominator of (23) is $2\mathrm{vol}(G)$ whenever the entries of $\mathbf{y}$ all have unit modulus.

Lemma 3 now tells us that

$$\sum_{(i,j)\in E}\left(\frac{1}{2}|g_i-g_j|^2-|\Lambda_{ij}-1|^2\right)\leq\frac{2C'\mathrm{vol}(G)}{\tau}\min_{\mathbf{y}\in\mathbb{C}^d}\eta_{\widetilde{X}}(\mathrm{sgn}(\mathbf{y}))\leq\frac{2C'\mathrm{vol}(G)}{\tau}\eta_{\widetilde{X}}(\widetilde{\mathbf{x}}_0).$$

Moreover,

$$\begin{aligned}\eta_{\widetilde{X}}(\widetilde{\mathbf{x}}_0) &= \frac{\sum_{(i,j)\in E}|(\widetilde{\mathbf{x}}_0)_i-\widetilde{X}_{ij}(\widetilde{\mathbf{x}}_0)_j|^2}{2\sum_{i\in V}\deg(i)|(\widetilde{\mathbf{x}}_0)_i|^2}\\ &= \frac{\sum_{(i,j)\in E}|(\widetilde{\mathbf{x}}_0)_i(\widetilde{\mathbf{x}}_0)_j^*-\widetilde{X}_{ij}|^2}{2\mathrm{vol}(G)}\\ &= \frac{\|\widetilde{X}_0-\widetilde{X}\|_F^2}{2\mathrm{vol}(G)},\end{aligned}$$

so that $\sum_{(i,j)\in E}\frac{1}{2}|g_i-g_j|^2\leq\frac{C'}{\tau}\|X_0-X\|_F^2+\sum_{(i,j)\in E}|\Lambda_{ij}-1|^2$. Considering also that

$$\sum_{(i,j)\in E}|\Lambda_{ij}-1|^2=\sum_{(i,j)\in E}|\widetilde{X}_{ij}-(\widetilde{X}_0)_{ij}|^2=\|\widetilde{X}-\widetilde{X}_0\|_F^2$$

and $\tau\leq 1$, this completes the proof. □

We may now use Theorem 4 to produce a perturbation bound for our banded matrix of phase differences $\widetilde{X}_0$.

**Corollary 2.** *Let $\widetilde{X}_0$ be the matrix in (6), $\widetilde{\mathbf{x}}_0$ be the vector of true phases (6), and $\widetilde{X}$ be as in line 3 of Algorithm 1 with $\widetilde{\mathbf{x}}=\mathrm{sgn}(\mathbf{u})$ where $\mathbf{u}$ is the top eigenvector of $\widetilde{X}$. Suppose that $\|\widetilde{X}_0-\widetilde{X}\|_F\leq\eta\|\widetilde{X}_0\|_F$ for some $\eta>0$. Then, there exists an absolute constant $C'\in\mathbb{R}^+$ such that*

$$\min_{\theta\in[0,2\pi]}\|\widetilde{\mathbf{x}}_0-\mathbb{e}^{\mathbb{i}\theta}\widetilde{\mathbf{x}}\|_2\leq C'\frac{\eta d^{\frac{5}{2}}}{\delta^2}.$$

*Proof.* We apply Theorem 4 with the unweighted and undirected graph $G=(V,E)$, where $V=[d]$ and $E=\{(i,j):|i-j| \mod d<\delta\}$. Observe that $G$ is also connected and $(2\delta-1)$-regular so that $\min_{i\in V}(\deg(i))=2\delta-1$. The spectral gap of $G$ is $\tau>C'''\delta^2/d^2>0$ by Corollary 1. We know that $\|\widetilde{X}_0\|_F=\sqrt{d(2\delta-1)}$, so that $\|\widetilde{X}_0-\widetilde{X}\|_F\leq C''\eta(d\delta)^{1/2}$. Finally, if $\mathbf{u}$ is the top eigenvector of $\widetilde{X}$ then it will also be an eigenvector of $L_1$ corresponding to its smallest eigenvalue since, here, $L_1=I-\frac{1}{2\delta-1}\widetilde{X}$.

Combining these observations we have

$$\min_{\theta\in[0,2\pi]}\|\widetilde{\mathbf{x}}_0-\mathbb{e}^{\mathbb{i}\theta}\widetilde{\mathbf{x}}\|_2\leq C\frac{C''\eta(d\delta)^{1/2}}{C'''\delta^2/d^2\cdot(2\delta-1)^{1/2}}=C'\frac{\eta d^{5/2}}{\delta^2}.$$

□

We are now properly equipped to analyze the robustness of Algorithm 1 to noise.

## 5. Recovery Guarantees for the Proposed Method

Herein we will assume Algorithm 1 is provided with measurements $\mathbf{y}$ of the form (7) such that the linear operator (4) is invertible on $T_\delta(\mathbb{C}^{d\times d})$ with condition number $\kappa > 0$. Unless otherwise stated, we follow the notation of §1.1- §1.2; therefore, our assumptions imply that $\|X - X_0\|_F \leq \kappa\|\mathbf{n}\|_2$.

We now aim to bound the Frobenius norm of the perturbation error $(\tilde{X} - \tilde{X}_0)$ present in the matrix $\tilde{X}$ formed in line 2 of Algorithm 1. Toward this end we define the set of $\rho$-small indexes of $\mathbf{x}_0$ to be

$$S_\rho := \left\{ j \;\middle|\; |\,(x_0)_j\,| < \left(\frac{\kappa\|\mathbf{n}\|_2}{\rho}\right)^{\frac{1}{4}} \right\} \tag{24}$$

where $\rho \in \mathbb{R}^+$ is a free parameter. With the definition of $S_\rho$ in hand we can bound the perturbation error $(\tilde{X} - \tilde{X}_0)$ using the next lemma.

**Lemma 6.** *Let $\tilde{X}$ be the matrix computed in line 2 of Algorithm 1. We have that*

$$\|\tilde{X} - \tilde{X}_0\|_{\mathrm{F}} \leq C\sqrt{\frac{\rho\frac{\kappa}{\delta}\|\mathbf{n}\|_2 + |S_\rho|}{d}} \cdot \|\tilde{X}_0\|_{\mathrm{F}}$$

*holds for all $\rho \in \mathbb{R}^+$, where $C$ is an absolute constant.*

*Proof.* Set $N_{jk} = X_{jk} - (X_0)_{jk}$ and consider that, for any $j, k \in S_\rho^c$,

$$\begin{aligned} |(\widetilde{X}_0)_{jk} - \widetilde{X}_{jk}| &= \left|(\widetilde{X}_0)_{jk} - \operatorname{sgn}\left(\tfrac{X_{jk}}{|(X_0)_{jk}|}\right)\right| \leq \left|(\widetilde{X}_0)_{jk} - \tfrac{X_{jk}}{|(X_0)_{jk}|}\right| + \left|\tfrac{X_{jk}}{|(X_0)_{jk}|} - \operatorname{sgn}\left(\tfrac{X_{jk}}{|(X_0)_{jk}|}\right)\right| \\ &\leq 2\left|(\widetilde{X}_0)_{jk} - \tfrac{X_{jk}}{|(X_0)_{jk}|}\right| = 2\tfrac{|N_{jk}|}{|(X_0)_{jk}|} \leq 2\rho^{\frac{1}{2}}\tfrac{|N_{jk}|}{(\kappa\|\mathbf{n}\|_2)^{\frac{1}{2}}}. \end{aligned}$$

Thus, there exists an absolute constant $C' \in \mathbb{R}^+$ such that

$$\begin{aligned} \|\tilde{X} - \tilde{X}_0\|_{\mathrm{F}}^2 &\leq \sum_{j,k\in S_\rho^c} 4\rho\frac{|N_{jk}|^2}{\kappa\|\mathbf{n}\|_2} + \sum_{j\in S_\rho,\ \text{or}\ k\in S_\rho} |(\tilde{X}_0)_{jk} - \tilde{X}_{jk}|^2 \\ &\leq 4\rho\frac{\|N\|_{\mathrm{F}}^2}{\kappa\|\mathbf{n}\|_2} + \sum_{j\in S_\rho} 4\cdot(4\delta - 3) = 4\rho\frac{\|N\|_{\mathrm{F}}^2}{\kappa\|\mathbf{n}\|_2} + 4\cdot(4\delta-3)\,|S_\rho| \\ &\leq C'(\rho\kappa\|\mathbf{n}\|_2 + \delta\,|S_\rho|). \end{aligned}$$

The proof is completed by recalling that $\|\widetilde{X}_0\|_F = \sqrt{(2\delta-1)d}$. □

We are finally ready to prove a robustness result for Algorithm 1.

**Theorem 5.** *Suppose that $\tilde{X}$ and $\tilde{X}_0$ satisfy $\|\tilde{X} - \tilde{X}_0\|_{\mathrm{F}} \leq \eta\|\tilde{X}_0\|_F$ for some $\eta > 0$. Then, the estimate $\mathbf{x}$ produced by Algorithm 1 satisfies*

$$\min_{\theta\in[0,2\pi]} \left\|\mathbf{x}_0 - \mathrm{e}^{\mathrm{i}\theta}\mathbf{x}\right\|_2 \leq C\|\mathbf{x}_0\|_\infty \left(\frac{d^{5/2}}{\delta^2}\right)\eta + Cd^{\frac{1}{4}}\sqrt{\kappa\|\mathbf{n}\|_2},$$

*where $C \in \mathbb{R}^+$ is an absolute universal constant. Alternatively, one can bound the error in terms of the size of the index set $S_\rho$ from (24) as*

$$\min_{\theta\in[0,2\pi]} \left\|\mathbf{x}_0 - \mathrm{e}^{\mathrm{i}\theta}\mathbf{x}\right\|_2 \leq C'\|\mathbf{x}_0\|_\infty \left(\frac{d}{\delta}\right)^2 \sqrt{\rho\frac{\kappa}{\delta}\|\mathbf{n}\|_2 + |S_\rho|} + C'd^{\frac{1}{4}}\sqrt{\kappa\|\mathbf{n}\|_2}, \tag{25}$$

*for any desired $\rho \in \mathbb{R}^+$, where $C' \in \mathbb{R}^+$ is another absolute universal constant.*

*Proof.* Let $\phi \in [0, 2\pi)$ be arbitrary; then $\mathbb{e}^{\mathbb{i}\phi}\mathbf{x} = |\mathbf{x}| \circ \mathbb{e}^{\mathbb{i}\phi}\tilde{\mathbf{x}}$ and $\mathbf{x}_0 = |\mathbf{x}_0| \circ \tilde{\mathbf{x}}_0$, where $\circ$ denotes the entrywise (Hadamard) product.

We see that

$$\begin{aligned}\min_{\phi\in[0,2\pi]} \|\mathbf{x}_0 - \mathbb{e}^{\mathbb{i}\phi}\mathbf{x}\|_2 &= \min_{\phi\in[0,2\pi]} \left\| |\mathbf{x}_0| \circ \tilde{\mathbf{x}}_0 - |\mathbf{x}| \circ \mathbb{e}^{\mathbb{i}\phi}\tilde{\mathbf{x}} \right\|_2 \\ &\leq \min_{\phi\in[0,2\pi]} \left\| |\mathbf{x}_0| \circ \tilde{\mathbf{x}}_0 - |\mathbf{x}_0| \circ \mathbb{e}^{\mathbb{i}\phi}\tilde{\mathbf{x}} \right\|_2 + \left\| |\mathbf{x}_0| \circ \mathbb{e}^{\mathbb{i}\phi}\tilde{\mathbf{x}} - |\mathbf{x}| \circ \mathbb{e}^{\mathbb{i}\phi}\tilde{\mathbf{x}} \right\|_2\end{aligned}$$

where the second term is now independent of $\phi$. As a result we have that

$$\min_{\phi\in[0,2\pi]} \|\mathbf{x}_0 - \mathbb{e}^{\mathbb{i}\phi}\mathbf{x}\|_2 \leq \|\mathbf{x}_0\|_\infty \left( \min_{\phi\in[0,2\pi]} \|\tilde{\mathbf{x}}_0 - \mathbb{e}^{\mathbb{i}\phi}\tilde{\mathbf{x}}\|_2 \right) + C''\sqrt{\kappa\sqrt{d} \cdot \|\mathbf{n}\|_2}$$

for some absolute constant $C'' \in \mathbb{R}^+$. Here the bound on the second term follows from Lemma 3 of [30] and the Cauchy-Schwarz inequality. The first inequality of the theorem now results from an application of Corollary 2 to the first term. The second inequality then follows from Lemma 6. □

Looking at the second inequality (25) in Theorem 5 we can see that the error bound there will be vacuous in most settings unless $S_\rho = \emptyset$. Recalling (24), one can see that $S_\rho$ will be empty as soon as $\rho = \kappa\delta\|\mathbf{n}\|_2 / |(x_0)_{\min}|^4$, where $(x_0)_{\min}$ is the smallest magnitude of any entry in $\mathbf{x}_0$. Utilizing this value of $\rho$ in (25) leads to the following corollary of Theorem 5.

**Corollary 3.** *Let $(x_0)_{\min} := \min_j |(x_0)_j|$ be the smallest magnitude of any entry in $\mathbf{x}_0$. Then, the estimate $\mathbf{x}$ produced by Algorithm 1 satisfies*

$$\min_{\theta\in[0,2\pi]} \left\| \mathbf{x}_0 - \mathbb{e}^{\mathbb{i}\theta}\mathbf{x} \right\|_2 \leq C \left( \frac{\|\mathbf{x}_0\|_\infty}{(x_0)_{\min}^2} \right) \left( \frac{d}{\delta} \right)^2 \kappa\|\mathbf{n}\|_2 + Cd^{\frac{1}{4}}\sqrt{\kappa\|\mathbf{n}\|_2},$$

*where $C \in \mathbb{R}^+$ is an absolute universal constant.*

Corollary 3 yields a deterministic recovery result for any signal $\mathbf{x}_0$ which contains no zero entries. If desired, a randomized result can now be derived from Corollary 3 for arbitrary $\mathbf{x}_0$ by right multiplying the signal $\mathbf{x}_0$ with a random "flattening" matrix as done in [30]. Finally, we note that a trivial variant of Corollary 3 can also be combined with the discussion in §1.5 in order to generate recovery guarantees for the windowed Fourier measurements defined by (14). However, we will leave such variants and extensions to the interested reader.

## 6. Numerical Evaluation

We now present numerical simulations supporting the theoretical recovery guarantees in Section 5. In addition to illustrating the performance of our algorithm, we also compare it against other existing phase retrieval methods using *local* measurements. All the results presented here may be recreated using the open source *BlockPR* Matlab software package which is freely available at [31].

In Section 6.2 we utilize the sparse measurement masks of Example 2 (see §2 for details). For each choice of $\delta$ these measurements correspond to $2\delta - 1$ physical masks which are each shifted $d$ times across the sample $\mathbf{x}_0$ by shifts of size 1. In Section 6.3, we then consider the Fourier measurement masks of Example 1 from §2. For each choice of $\delta$ these measurements correspond to $2\delta - 1$ Fourier measurements of *one* physical mask/illumination which is also shifted $d$ times across the sample $\mathbf{x}_0$ by shifts of size 1. These Fourier measurements correspond particularly well to ptychographic measurements of a large sample in the small $\delta$ regime.

For completeness, we also present selected results comparing the proposed formulation against other well established phase retrieval algorithms (such as *Wirtinger Flow*) using *global* measurements such as coded diffraction patterns (CDPs) [12, §1.5]. These measurements correspond to those one might obtain from the diffraction patterns of the sample $\mathbf{x}_0$ after it has been masked by

several different global random windows. Unless otherwise stated, we use i.i.d. zero-mean complex Gaussian random test signals with measurement errors modeled using an additive Gaussian noise model. Applied measurement noise and reconstruction error are both reported in decibels (dB) in terms of signal to noise ratios (SNRs), with

$$\text{SNR (dB)} = 10\log_{10}\left(\frac{\sum_{j=1}^{D}|\langle \mathbf{a}_j, \mathbf{x}_0\rangle|^4}{D\sigma^2}\right), \qquad \text{Error (dB)} = 10\log_{10}\left(\frac{\min_\theta \|\mathrm{e}^{\mathrm{i}\theta}\mathbf{x} - \mathbf{x}_0\|_2^2}{\|\mathbf{x}_0\|_2^2}\right),$$

where $\mathbf{a}_j, \mathbf{x}_0, \mathbf{x}, \sigma^2$ and $D$ denote the measurement vectors, true signal, recovered signal, (Gaussian) noise variance and number of measurements respectively. All simulations were performed on a laptop computer running GNU/Linux (Ubuntu Linux 16.04 `x86_64`) with an Intel® Core™i3-3120M (2.5 GHz) processor, 4GB RAM and Matlab R2015b. Each data point in the timing and robustness plots was obtained as the average of 100 trials.

6.1. **Numerical Improvements to Algorithm 1: Magnitude Estimation.** Looking at the matrix $X$ formed on line 1 of Algorithm 1 one can see that

$$X = X_0 + N',$$

where $X_0$ is the banded Hermitian matrix $T_\delta(\mathbf{x}_0\mathbf{x}_0^*)$ defined in (5), and $N'$ contains arbitrary banded Hermitian noise. As stated and analyzed above, Algorithm 1 estimates the magnitude of each entry of $\mathbf{x}_0$ by observing that

$$X_{jj} = |(x_0)_j|^2 + N'_{jj}, \quad j \in [d].$$

Though this magnitude estimate suffices for our theoretical treatment above, it can be improved in practice by using slightly more general techniques.

Considering the component-wise magnitude of $X$, $|X| \in \mathbb{R}^{d\times d}$, one can see that its entries are

$$|X|_{jk} = \begin{cases} |(x_0)_j||(x_0)_k| + N''_{jk} & \text{if } |j-k| \bmod d < \delta \\ 0 & \text{otherwise} \end{cases},$$

where $N'' = |X| - |X_0|$ represents the changes in magnitude to the entries of $|X_0|$ due to noise. We may then let $D_j \in \mathbb{R}^{\delta\times\delta}$ denote the submatrix of $|X|$ given by

$$(D_j)_{kh} = |X|_{(j+k-1) \bmod d,\ (j+h-1) \bmod d},$$

for all $j \in [d]$; similarly we let $N''_j$ denote the respective submatrices of $N''$. With this notation, it is clear that

$$D_j = |\mathbf{x}_0|^{(j)}(|\mathbf{x}_0|^{(j)})^* + N''_j,$$

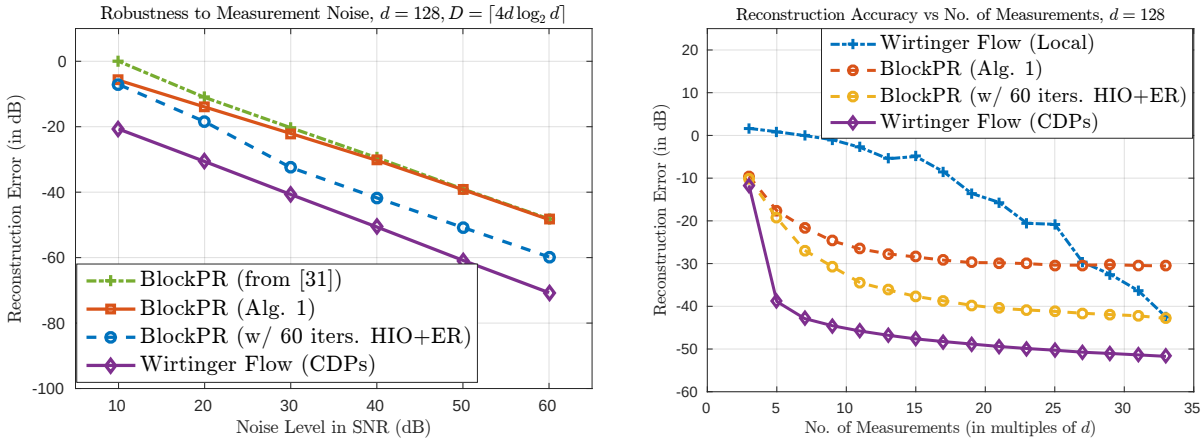
where $|\mathbf{x}_0|_k^{(j)} = |\mathbf{x}_0|_{k+j-1}, k \in [\delta]$. This immediately suggests that we can estimate the magnitudes of the entries of $\mathbf{x}_0$ by calculating the top eigenvectors of these approximately rank one $D_j$ matrices.

Indeed, if we do so for all of $D_1, \ldots, D_d \in \mathbb{R}^{\delta\times\delta}$, we will produce $\delta$ estimates of each $(x_0)_j$ entry's magnitude. A final estimate of each $|(x_0)_j|$ can then be computed by taking the average, median, etc. of the $\delta$ different estimates of $|(x_0)_j|$ provided by each of the leading eigenvectors of $D_{j-\delta+1}, \ldots, D_j$; in our experiments, we used the arithmetic mean. Of course, one need neither use all $d$ possible $D_j$ matrices, nor make them have size $\delta\times\delta$. More generally, to reduce computational complexity, one may instead use $d/s$ matrices, $\tilde{D}_{j'} \in \mathbb{R}^{\gamma\times\gamma}$, of size $1 \le \gamma \le \delta$ and with shifts $s \le \gamma$ (dividing $d$), having entries

$$(\tilde{D}_{j'})_{k,h} = |X|_{(sj'+k-s) \bmod d,\ (sj'+h-s) \bmod d}.$$

Computing the leading eigenvectors of $\tilde{D}_{j'}$ for all $j' \in [d/s]$ will then produce (multiple) estimates of each magnitude $|(x_0)_j|$ which can then be combined as desired to produce our final magnitude estimates. As we shall see below, one can achieve better numerical robustness to noise using this

technique than what can be achieved using the simpler magnitude estimation technique presented in line 4 of Algorithm 1.




(A) Improved Robustness to Measurement Noise – Comparing Variants of the *BlockPR* algorithm

(B) Reconstruction Error vs. No. of Measurements; (Reconstruction at 40dB SNR)

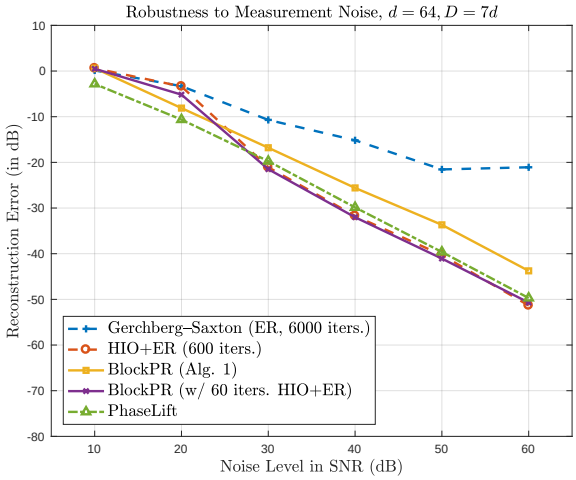
FIGURE 2. Robust Phase Retrieval – Local vs. Global Measurements

6.2. **Experiments with Measurements from Example 2 of §2.** We begin by presenting results in Fig. 2a demonstrating the improved noise robustness of the proposed method over the formulation in [30]. Recall that [30] uses a greedy angular synchronization method instead of the eigenvector-based procedure analyzed in this paper. Fig. 2a plots the reconstruction error when recovering a $d = 128$ length complex Gaussian test signal using $D = \lceil 4d \log_2 d \rceil$ measurements at different added noise levels. As mentioned above, the local correlation measurements described in Example 2 of Section 2 are utilized in this plot and in all the ensuing experiments in this subsection unless otherwise indicated. Three variants of the proposed algorithm are plotted in Fig. 2a:

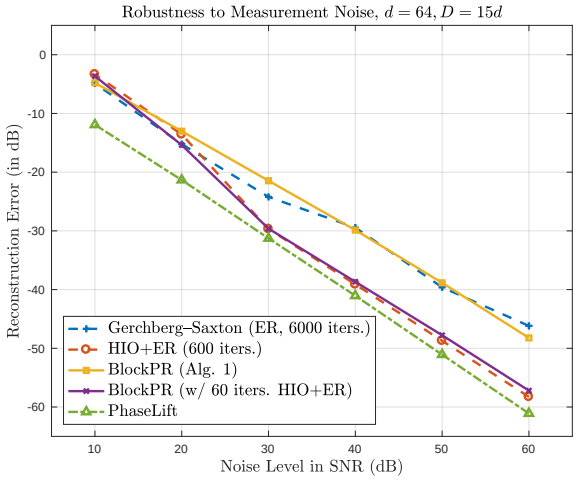
(1) an implementation of Algorithm 1 (denoted by $\square$'s),
(2) an implementation of Algorithm 1 post-processed using 60 iterations of the Hybrid Input–Output (HIO) and Error Reduction (ER) algorithms (implemented in two successive sets, with each set consisting of 20 iterations of HIO followed by 10 ER iterations; denoted by $\circ$'s), and
(3) the algorithmic implementation from [30] (denoted by +'s).

We see that the eigenvector-based angular synchronization method proposed in this paper provides more accurate reconstructions – especially at low SNRs – over the greedy angular synchronization of [30]. Moreover, post-processing using the HIO+ER algorithms as detailed above yields a significant improvement in reconstruction errors over the two other variants. For reference, we also include reconstruction errors with the *Wirtinger Flow* algorithm (denoted by $\diamondsuit$'s) when using (*global*) coded diffraction pattern (CDP) measurements. Specifically, we use $2\delta - 1$ (with $\delta = 2\log_2 d = 14$) octanary modulations/codes as described in [12, §1.5, (1.9)] to construct the CDP measurements. Clearly, using global measurements such as coded diffraction patterns provides superior noise tolerance; however, they are not applicable to the local measurement model considered here. Indeed, when the *Wirtinger Flow* algorithm is used with local measurements such as those described in this paper, the noise tolerance significantly deteriorates. Fig. 2b illustrates this phenomenon by plotting the reconstruction error in recovering a $d = 128$ length complex Gaussian test signal at

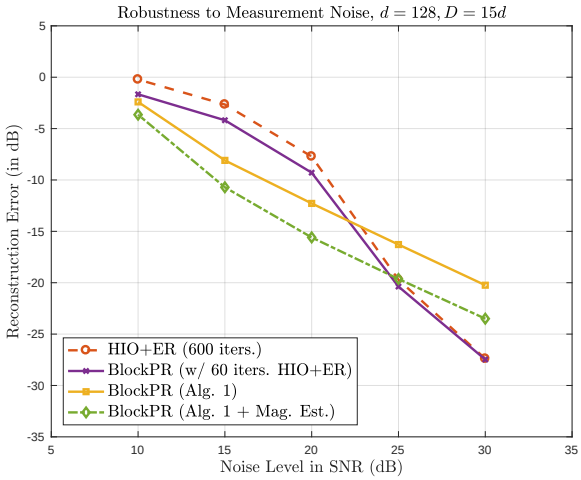
40 dB SNR when using different numbers of measurements, $D$. *Wirtinger flow*, for example, requires a large number of local measurements before returning accurate reconstructions. The wide disparity in reconstruction accuracy between local and global measurements for *Wirtinger Flow* illustrates the significant challenge in phase retrieval from local measurements. Furthermore, we see that the *BlockPR* method proposed in this paper is more noise tolerant than *Wirtinger Flow* for local measurements.



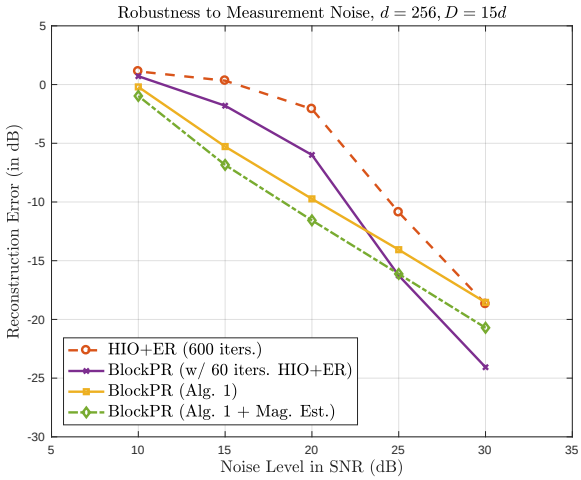

(A) Using $D = 7d$ measurements.




(B) Using $D = 15d$ measurements.




(C) Low SNR Simulations: Using $D = 15d$ measurements, Problem Size $d = 128$




(D) Low SNR Simulations: Using $D = 15d$ measurements, Problem Size $d = 256$

FIGURE 3. Robustness to measurement noise – Phase Retrieval from deterministic local correlation measurements.

Given the weaker performance of *Wirtinger Flow* with local measurements, we now restrict our attention to the empirical evaluation of the proposed method (Alg. 1, as well as the post-processed variant (with HIO+ER iterations)) against *PhaseLift* and alternating projection algorithms. Although numerical simulations suggest that these methods work with local measurements, we note that (to the best of our knowledge) there are no theoretical recovery or robustness guarantees for

these methods and measurements. The *PhaseLift* algorithm was implemented as a trace regularized least-squares problem using CVX [26, 25] – a package for specifying and solving convex programs in Matlab. We consider two variants from the family of alternating projection methods – *Gerchberg–Saxton* (sometimes referred to as the Error Reduction (ER) algorithm) and Hybrid Input-Output (HIO). For both algorithms, the following two projections were utilized: (i) projection onto the measured magnitudes, and (ii) projection onto the span of the measurement vectors $\{\mathbf{a}_j\}_{j=1}^{D}$. This formulation as well as other details and connections to convex optimization theory can be found in [8]. For the ER and HIO implementations, the initial guess was set to be the all-zero vector.[9] For the HIO implementation, as is popular practice (see, for example, [22]) every few (20) HIO iterations were followed by a small number of (10) ER iterations, with the maximum number of HIO+ER iterations limited to 600 – this choice of iteration count ensures convergence of the algorithm (see Fig. 4) while comparing favorably with the computational cost (see Fig. 5b) of the proposed *BlockPR* method. For the ER implementation, 6,000 iterations were necessary to ensure convergence.

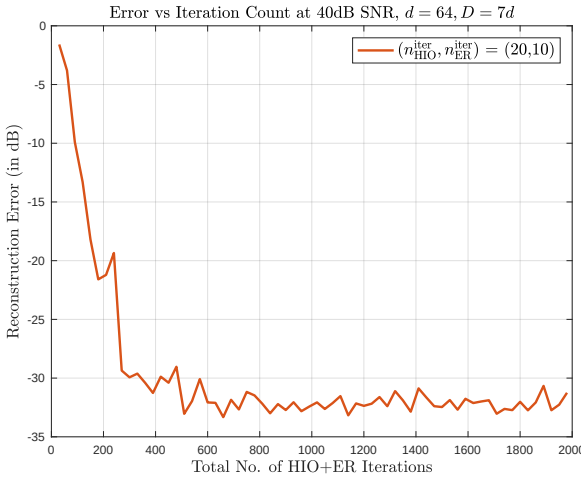


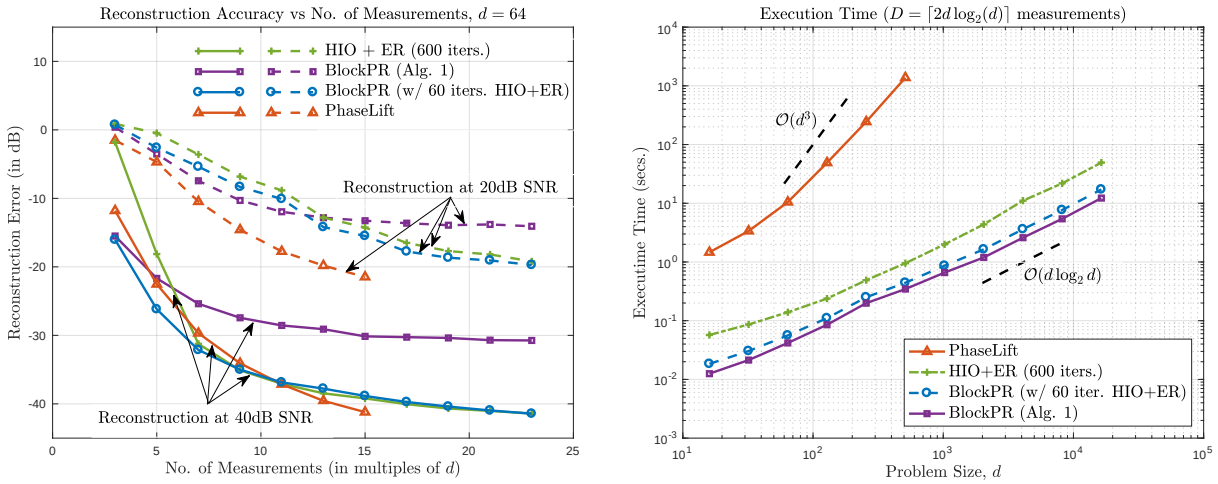

FIGURE 4. Reconstruction Error vs. Iteration Count for HIO+ER Implementation

We begin by presenting numerical results evaluating the robustness to measurement noise. Figs. 3a and 3b plot the error in reconstructing a $d = 64$ length complex vector $\mathbf{x}_0$ using $D = 7d$ and $D = 15d$ local correlation-based phaseless measurements respectively. Note that this corresponds to using $\delta = 4$ (and the associated $2\delta - 1 = 7$ masks) and $\delta = 8$ (and the corresponding $2\delta - 1 = 15$ masks) respectively. Moreover, the well-conditioned *deterministic* (and sparse) measurement construction defined in Example 2 of Section 2 was utilized along with additive Gaussian measurement noise. We see from Fig. 3 that the method proposed in this paper (denoted *BlockPR* in the figure) performs reliably across a wide range of SNRs and compares favorably against existing popular phase retrieval algorithms. When using a small number of measurements (as in Fig. 3a, and modeling real-world situations), both variants of the proposed method – Alg. 1, as well as Alg. 1 post-processed using 60 HIO+ER iterations – outperform the ER algorithm by significant margins and compare well with the popular HIO+ER algorithm. When more measurements are available (as in Fig. 3b), the performance of the ER and HIO+ER algorithms approaches the performance of the *BlockPR* variants proposed in this paper. In addition, the proposed methods also compare well with the significantly more expensive *PhaseLift* reconstructions. We emphasize that the superior performance of the proposed methods demonstrated here comes with rigorous theoretical recovery

[9] We note that using a random starting guess does not change the qualitative nature of the empirical results.

guarantees for local measurements – something that cannot be said of any of the other methods in Fig. 3.

Additionally, Figs. 3c and 3d compare the performance of the HIO+ER algorithm at low SNRs for problems sizes $d = 128$ and $d = 256$ respectively, with the various *BlockPR* implementations – including one with the improved magnitude estimation procedure detailed in §6.1 (with $s = 1$ and using the average of the obtained $\tilde{D}_{j\prime}$ block magnitude estimates). These figures demonstrate the value of the magnitude estimation procedure from §6.1 at low SNRs over the HIO+ER post-processing method utilized in the other figures (and over the HIO+ER algorithm); we defer a more detailed study of this to future work.




(A) Reconstruction Error vs. No. of Measurements

(B) Execution Time vs. Problem Size

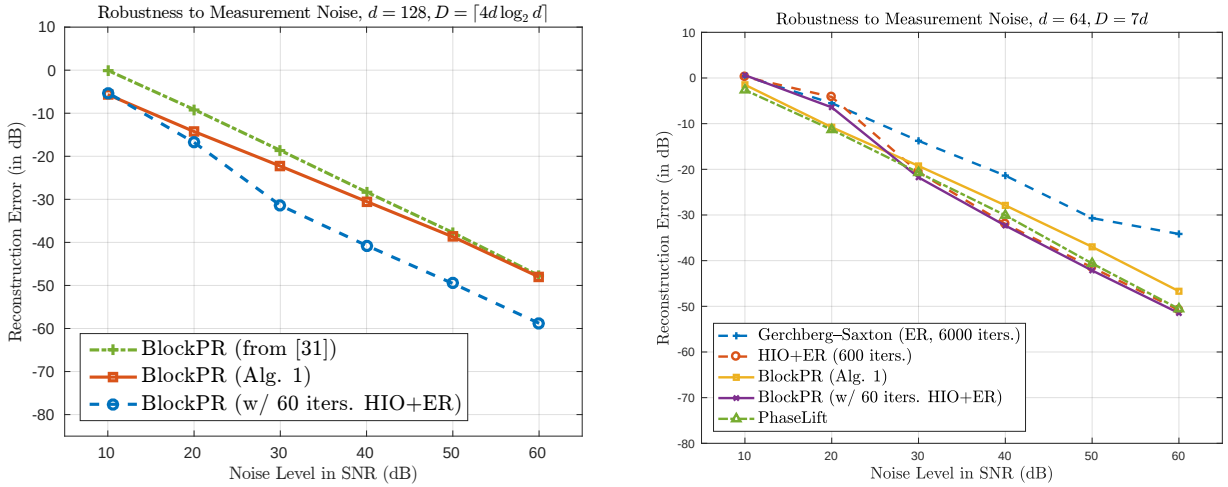
FIGURE 5. Performance Evaluation and Comparison of the Proposed Phase Retrieval Method (with Deterministic Local Correlation Measurements of Example 2, §2 and Additive Gaussian Noise)

Next, Fig. 5a plots the reconstruction error in recovering a $d = 64$-length complex vector as a function of the number of measurements used. This corresponds to using values of $\delta$ ranging from 2 to 12 (and the associated $2\delta - 1$ masks). As with Fig. 3, the deterministic correlation-based measurement constructions of Section 2 (Example construction 2) were utilized along with an additive Gaussian noise model. Plots are provided for simulations at two noise levels – 20 dB and 40 dB. Comparing the performance of the two *BlockPR* variants, we observe that the HIO+ER post-processing procedure provides improved reconstruction errors – with the margin of improvement increasing when more measurements are available. We also notice that both variants of *BlockPR* compare particularly well with the other algorithms (HIO+ER and *PhaseLift*) when small numbers of measurements are available.

Finally, Fig. 5b plots the average execution time (in seconds) required to solve the phase retrieval problem using $D = \lceil 2d \log_2 d \rceil$ measurements. For comparison, execution times for the *PhaseLift* and HIO+ER alternating projection algorithms are provided. We observe that both variants of the proposed method are several orders of magnitude faster than the *PhaseLift* algorithm,[10] and between 2–5 times faster than the HIO+ER implementation. Moreover, the plot illustrates the

[10] For computational efficiency and due to memory constraints, the *PhaseLift* plot in Fig. 5b was generated using the TFOCS software package (http://cvxr.com/tfocs/) instead of the more computationally expensive CVX software package.

essentially FFT-time computational complexity (see §1.3) of the proposed method. Between the two *BlockPR* variants, we see that there is a small trade-off between reduced execution time and improved accuracy; one of the two variants may be more appropriate depending on the application requirements.




(A) Improved Robustness to Measurement Noise – Comparing Variants of the *BlockPR* algorithm

(B) Phase Retrieval from $D = 7d$ Measurements – Comparison with Other Phase Retrieval Algorithms

FIGURE 6. Numerical Evaluation of the Proposed Algorithm with the Ptychographic Measurements from Example 1 of §2

6.3. **Experiments with Ptychographic Measurements from Example 1 of §2.** We now present a selection of empirical results demonstrating the accuracy, robustness, and efficiency of the proposed method when using the ptychographic measurements from Example 1, §2 and observe that, in this case, the comparison between the proposed algorithm and existing methods is similar to that in section §6.2. Recall that (see Example 1 from §2 and §1.4 for details) this measurement construction corresponds to a discretization of the ptychographic measurements $\left|\mathcal{F}[\widetilde{h} \cdot S_l f](\omega)\right|^2$ as per (10), where $f$ is the unknown specimen and $\widetilde{h}(t) = \frac{e^{-t/a}}{\sqrt[4]{2\delta - 1}}$ denotes the *single, deterministic* illumination function (or mask). As before, $\delta$ defines the local support of the illumination function and we consider a discretization involving $2\delta - 1$ Fourier modes and sample/specimen shifts of 1 unit or pixel (yielding a total of $d$ shifts in the discrete problem formulation).

We begin with Fig. 6a, which demonstrates the relative performance of the *BlockPR* variants described in [30] and this paper. More specifically, we plot the error in reconstructing a $d = 128$ length complex Gaussian test signal using $D = \lceil 4d \log_2 d \rceil$ measurements at different added noise levels. As with Fig. 2a, we plot results for three different variants of the *BlockPR* algorithm: (i) Algorithm 1, (ii) Algorithm 1 with the HIO+ER post-processing procedure as described in §6.2, and (iii) the implementation from [30]. We again observe (as in Fig. 2a) that the methods described in this paper (which use eigenvector-based angular synchronization) are more accurate than that detailed in [30] (which uses a greedy angular synchronization method), with the improvement in performance being especially significant at low SNRs. Next, Fig. 6b studies the performance of the proposed method(s) against three other popular phase retrieval algorithms – *PhaseLift*, the *Gerchberg–Saxton* Error Reduction (ER) algorithm, and the Hybrid Input–Output (HIO) algorithm (with implementation parameters identical to those in §6.2). As with Fig. 3a, we consider the

reconstruction of a $d = 64$ length complex vector $x_0$ using $D = 7d$ measurements in the presence of additive Gaussian noise. We observe that the methods proposed in this paper outperform the ER algorithm and compare very well with the HIO+ER and (the significantly more expensive) *PhaseLift* algorithms across a wide range of SNRs. Finally, we note that the execution time plot for this ptychographic measurement construction is qualitatively and quantitatively similar to Fig. 5b – the proposed methods are faster (by a factor of 2–5) than an equivalent HIO+ER implementation and orders of magnitude faster than convex optimization approaches such as *PhaseLift*.

## 7. Concluding Remarks

In this paper new and improved deterministic robust recovery guarantees are proven for the phase retrieval problem using local correlation measurements. In addition, a new practical phase retrieval algorithm is presented which is both faster and more noise robust than previously existing approaches (e.g., alternating projections) for such local measurements.

Future work might include the exploration of more general classes of measurements which are guaranteed to lead to well conditioned linear systems of the type used to reconstruct $X \approx X_0$ in line 1 of Algorithm 1. Currently two deterministic measurement constructions are known (recall, e.g., Section 2) – it should certainly be possible to construct more general families of such measurements.

Other interesting avenues of inquiry include the theoretical analysis of the magnitude estimate approach proposed in Section 6.1 in combination with the rest of Algorithm 1. Alternate phase retrieval approaches might also be developed by using such local block eigenvector-based methods for estimating phases too, instead of just using the single global top eigenvector as currently done in line 3 of Algorithm 1.

Finally, more specific analysis of the performance of the proposed methods using masked/windowed Fourier measurements (recall Section 1.5) would also be interesting. In particular, an analysis of the performance of such approaches as a function of the bandwidth of the measurement mask/window could be particularly enlightening. One might also consider extending the discrete results of this paper to the analytic setting by, e.g., expanding on [35].

## References


[1] B. Alexeev, A. S. Bandeira, M. Fickus, and D. G. Mixon. Phase retrieval with polarization. *SIAM Journal on Imaging Sciences*, 7(1):35–66, 2014.

[2] R. Balan, B. Bodmann, P. Casazza, and D. Edidin. Fast algorithms for signal reconstruction without phase. In *Optical Engineering+ Applications*, pages 67011L–67011L. International Society for Optics and Photonics, 2007.

[3] R. Balan, B. G. Bodmann, P. G. Casazza, and D. Edidin. Painless reconstruction from magnitudes of frame coefficients. *Journal of Fourier Analysis and Applications*, 15(4):488–501, 2009.

[4] R. Balan, P. Casazza, and D. Edidin. On signal reconstruction without phase. *Applied and Computational Harmonic Analysis*, 20(3):345–356, 2006.

[5] A. S. Bandeira and D. G. Mixon. Near-optimal phase retrieval of sparse vectors. In *Wavelets and Sparsity XV*, volume 8858, page 88581O. International Society for Optics and Photonics, 2013.

[6] A. S. Bandeira, A. Singer, and D. A. Spielman. A Cheeger inequality for the graph connection Laplacian. *SIAM Journal on Matrix Analysis and Applications*, 34(4):1611–1630, 2013.

[7] H. Bauschke, P. Combettes, and D. Luke. Hybrid projection-reflection method for phase retrieval. *Journal of the Optical Society of America. A, Optics, Image science, and Vision*, 20(6):1025–1034, 2003.

[8] H. H. Bauschke, P. L. Combettes, and D. R. Luke. Phase retrieval, error reduction algorithm, and Fienup variants: A view from convex optimization. *Journal of the Optical Society of America. A, Optics, Image science, and Vision*, 19(7):1334–1345, 2002.

[9] T. Bendory, Y. C. Eldar, and N. Boumal. Non-convex phase retrieval from STFT measurements. *IEEE Transactions on Information Theory*, 64(1):467–484, 2018.

[10] B. G. Bodmann and N. Hammen. Stable phase retrieval with low-redundancy frames. *Advances in computational mathematics*, 41(2):317–331, 2015.

[11] E. J. Candes and X. Li. Solving quadratic equations via Phaselift when there are about as many equations as unknowns. *Foundations of Computational Mathematics*, 14(5):1017–1026, 2014.

[12] E. J. Candes, X. Li, and M. Soltanolkotabi. Phase retrieval from coded diffraction patterns. *Applied and Computational Harmonic Analysis*, 39(2):277–299, Sept. 2015.
[13] E. J. Candes, X. Li, and M. Soltanolkotabi. Phase retrieval via Wirtinger flow: Theory and algorithms. *Information Theory, IEEE Transactions on*, 61(4):1985–2007, 2015.
[14] E. J. Candes, T. Strohmer, and V. Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013.
[15] F. R. Chung. *Spectral Graph Theory*. Number 92. American Mathematical Society, 1997.
[16] C. Davis and W. M. Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970.
[17] M. Dierolf, O. Bunk, S. Kynde, P. Thibault, I. Johnson, A. Menzel, K. Jefimovs, C. David, O. Marti, and F. Pfeiffer. Ptychography & lensless X-ray imaging. *Europhysics News*, 39(1):22–24, 2008.
[18] Y. Eldar, P. Sidorenko, D. Mixon, S. Barel, and O. Cohen. Sparse phase retrieval from short-time fourier measurements. *IEEE Signal Proc. Letters*, 22(5), 2015.
[19] Y. C. Eldar and S. Mendelson. Phase retrieval: Stability and recovery guarantees. *Applied and Computational Harmonic Analysis*, 36(3):473–494, 2014.
[20] V. Elser. Phase retrieval by iterated projections. *Journal of the Optical Society of America. A, Optics, Image science, and Vision*, 20(1):40–55, 2003.
[21] J. R. Fienup. Reconstruction of an object from the modulus of its Fourier transform. *Optics letters*, 3(1):27–29, 1978.
[22] J. R. Fienup. Phase retrieval algorithms: A comparison. *Applied Optics*, 21(15):2758–2769, 1982.
[23] R. Gerchberg and W. Saxton. A practical algorithm for the determination of the phase from image and diffraction plane pictures. *Optik*, 35:237246, 1972.
[24] J. W. Goodman. *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.
[25] 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, 2008. `http://stanford.edu/~boyd/graph_dcp.html`.
[26] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, Mar. 2014.
[27] D. Gross, F. Krahmer, and R. Kueng. Improved recovery guarantees for phase retrieval from coded diffraction patterns. *Applied and Computational Harmonic Analysis*, 2015.
[28] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2012.
[29] M. Iwen, F. Krahmer, and A. Viswanathan. Technical note: A minor correction of Theorem 1.3 from [1]. *Unpublished note available at `http://users.math.msu.edu/users/markiwen/Papers/PhaseLiftproof.pdf`*, April 2015.
[30] M. Iwen, A. Viswanathan, and Y. Wang. Fast phase retrieval from local correlation measurements. *SIAM Journal on Imaging Sciences*, 9(4):1655–1688, 2016.
[31] M. Iwen, Y. Wang, and A. Viswanathan. BlockPR: Matlab software for phase retrieval using block circulant measurement constructions and angular synchronization, version 2.0. `https://bitbucket.org/charms/blockpr`, Apr. 2016.
[32] K. Jaganathan, Y. C. Eldar, and B. Hassibi. STFT phase retrieval: Uniqueness guarantees and recovery algorithms. *IEEE Journal of selected topics in signal processing*, 10(4):770–781, 2016.
[33] X. Li and V. Voroninski. Sparse signal recovery from quadratic measurements via convex programming. *SIAM Journal on Mathematical Analysis*, 45(5):3019–3033, 2013.
[34] S. Marchesini, Y.-C. Tu, and H.-t. Wu. Alternating projection, ptychographic imaging and phase synchronization. *Applied and Computational Harmonic Analysis*, 41(3):815–851, 2016.
[35] S. Merhi, A. Viswanathan, and M. Iwen. Recovery of compactly supported functions from spectrogram measurements via lifting. In *Sampling Theory and Applications (SampTA), 2017 International Conference on*, pages 538–542. IEEE, 2017.
[36] R. Millane. Phase retrieval in crystallography and optics. *Journal of the Optical Society of America. A, Optics, Image science, and Vision*, 7(3):394–411, 1990.
[37] P. Netrapalli, P. Jain, and S. Sanghavi. Phase retrieval using alternating minimization. In *Advances in Neural Information Processing Systems*, pages 2796–2804, 2013.
[38] G. E. Pfander and P. Salanevich. Robust phase retrieval algorithm for time-frequency structured measurements. *eprint arXiv:1611.02540*, 2016.
[39] P. Salanevich and G. E. Pfander. Polarization based phase retrieval for time-frequency structured measurements. In *Sampling Theory and Applications (SampTA), 2015 International Conference on*, pages 187–191. IEEE, 2015.
[40] G. Stewart and J. Sun. *Matrix Perturbation Theory*. Academic Press, 1990.

[41] H. Takajo, T. Takahashi, H. Kawanami, and R. Ueda. Numerical investigation of the iterative phase-retrieval stagnation problem: Territories of convergence objects and holes in their boundaries. *Journal of the Optical Society of America. A, Optics, Image science, and Vision*, 14(12):3175–3187, 1997.
[42] H. Takajo, T. Takahashi, and T. Shizuma. Further study on the convergence property of the hybrid input–output algorithm used for phase retrieval. *Journal of the Optical Society of America. A, Optics, Image science, and Vision*, 16(9):2163–2168, 1999.
[43] H. Takajo, T. Takahashi, R. Ueda, and M. Taninaka. Study on the convergence property of the hybrid input–output algorithm used for phase retrieval. *Journal of the Optical Society of America. A, Optics, Image science, and Vision*, 15(11):2849–2861, 1998.
[44] L. N. Trefethen and D. Bau III. *Numerical Linear Algebra*, volume 50. SIAM, 1997.
[45] A. Viswanathan and M. Iwen. Fast angular synchronization for phase retrieval via incomplete information. In *Wavelets and Sparsity XVI*, volume 9597, page 959718. International Society for Optics and Photonics, 2015.
[46] A. Walther. The Question of Phase Retrieval in Optics. *Optica Acta*, 10:41–49, 1963.
[47] C. Yang, J. Qian, A. Schirotzek, F. Maia, and S. Marchesini. Iterative Algorithms for Ptychographic Phase Retrieval. *ArXiv e-prints*, May 2011.
[48] Y. Yu, T. Wang, and R. Samworth. A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika*, 102(2):315–323, 2015.

## Appendix A. Alternate Perturbation Bounds

In this section we present a simpler (and easier to derive), albeit weaker, perturbation result in the spirit of Section 4, which is associated with the analysis of line 3 of Algorithm 1. Specifically, we will derive an upper bound on $\min_{\theta\in[0,2\pi]} \|\widetilde{\mathbf{x}}_0 - \mathrm{e}^{\mathrm{i}\theta}\widetilde{\mathbf{x}}\|_2$ (provided by Theorem 6), which scales like $d^3$. While this dependence is strictly worse than the one derived in Section 4, it is easier to obtain and the technique may be of independent interest.

We will begin with a result concerning the top eigenvector of any Hermitian matrix.

**Lemma 7.** *Let $X_0 = \sum_{j=1}^d \nu_j \mathbf{x}_j\mathbf{x}_j^*$ be Hermitian with eigenvalues $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_d$ and orthonormal eigenvectors $\mathbf{x}_1, \ldots, \mathbf{x}_d \in \mathbb{C}^d$. Suppose that $X = \sum_{j=1}^d \lambda_j \mathbf{v}_j\mathbf{v}_j^*$ is Hermitian with eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_d$, orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_d \in \mathbb{C}^d$, and $\|X - X_0\|_{\mathrm{F}} \leq \eta\|X_0\|_F$ for some $\eta \geq 0$. Then,*

$$\left(1 - |\langle \mathbf{x}_1, \mathbf{v}_1\rangle|^2\right) \leq \frac{4\eta^2\|X_0\|_{\mathrm{F}}^2}{(\nu_1 - \nu_2)^2}.$$

*Proof.* An application of the $\sin\theta$ theorem [16, 40] (see, e.g., the proof of Corollary 1 in [48]) tells us that

$$\sin\left(\arccos\left(|\langle \mathbf{x}_1, \mathbf{v}_1\rangle|\right)\right) \leq \frac{2\eta\|X_0\|_F}{|\nu_1 - \nu_2|}.$$

Squaring both sides we then learn that

$$\left(1 - |\langle \mathbf{x}_1, \mathbf{v}_1\rangle|^2\right) = \sin^2\left(\arccos\left(|\langle \mathbf{x}_1, \mathbf{v}_1\rangle|\right)\right) \leq \frac{4\eta^2\|X_0\|_F^2}{(\nu_1 - \nu_2)^2}, \tag{26}$$

giving us the desired inequality. □

The following variant of Lemma 7 concerning rank 1 matrices $X_0$ is of use in the analysis of many other phase retrieval methods, and can be used, e.g., to correct and simplify the proof of equation (1.8) in Theorem 1.3 of [11].

**Lemma 8.** *Let $\mathbf{x}_0 \in \mathbb{C}^d$, set $X_0 = \mathbf{x}_0\mathbf{x}_0^*$, and let $X \in \mathbb{C}^{d\times d}$ be Hermitian with $\|X - X_0\|_F \leq \eta\|X_0\|_F = \eta\|\mathbf{x}_0\|_2^2$ for some $\eta \geq 0$. Furthermore, let $\lambda_i$ be the $i$-th largest magnitude eigenvalue of*

*$X$ and $\mathbf{v}_i \in \mathbb{C}^d$ an associated eigenvector, such that the $\mathbf{v}_i$ form an orthonormal eigenbasis. Then*

$$\min_{\theta\in[0,2\pi]} \|\mathrm{e}^{\mathrm{i}\theta}\mathbf{x}_0 - \sqrt{|\lambda_1|}\mathbf{v}_1\|_2 \leq (1+2\sqrt{2})\eta\|\mathbf{x}_0\|_2.^{11}$$

*Proof.* In this special case of Lemma 7 we have $\nu_1 = \|X_0\|_F = \|\mathbf{x}_0\|_2^2$ and $\mathbf{x}_1 := \mathbf{x}_0/\|\mathbf{x}_0\|$. Choose $\phi \in [0, 2\pi]$ such that $\langle \mathrm{e}^{\mathrm{i}\phi}\mathbf{x}_0, \mathbf{v}_1\rangle = |\langle \mathbf{x}_0, \mathbf{v}_1\rangle|$. Then,

$$\begin{aligned}\|\mathrm{e}^{\mathrm{i}\phi}\mathbf{x}_0 - \sqrt{\nu_1}\mathbf{v}_1\|_2^2 &= 2\nu_1 - 2\nu_1\cdot|\langle \mathbf{x}_0/\|\mathbf{x}_0\|, \mathbf{v}_1\rangle| = 2\nu_1 - 2\nu_1\cdot|\langle \mathbf{x}_1, \mathbf{v}_1\rangle| \\ &\leq 2\nu_1\left(1-|\langle \mathbf{x}_1, \mathbf{v}_1\rangle|\right)\left(1+|\langle \mathbf{x}_1, \mathbf{v}_1\rangle|\right) \\ &= 2\nu_1\left(1-|\langle \mathbf{x}_1, \mathbf{v}_1\rangle|^2\right) \leq 8\eta^2\|X_0\|_{\mathrm{F}}\end{aligned} \tag{27}$$

where the last inequality follows from Lemma 7 with $\nu_1 = \|X_0\|_F = \|\mathbf{x}_0\|_2^2$. Finally, by the triangle inequality, Weyl's inequality (see, e.g., [28]), and (27), we have

$$\begin{aligned}\|\mathrm{e}^{\mathrm{i}\phi}\mathbf{x}_0 - \sqrt{|\lambda_1|}\mathbf{v}_1\|_2 &\leq \|\mathrm{e}^{\mathrm{i}\phi}\mathbf{x}_0 - \sqrt{\nu}_1\mathbf{v}_1\|_2 + \|\sqrt{\nu}_1\mathbf{v}_1 - \sqrt{|\lambda_1|}\mathbf{v}_1\|_2 \\ &\leq 2\sqrt{2}\cdot\eta\sqrt{\nu}_1 + \left|\sqrt{\nu}_1 - \sqrt{|\lambda_1|}\right| \\ &\leq 2\sqrt{2}\cdot\eta\sqrt{\nu}_1 + \frac{|\nu_1-\lambda_1|}{\sqrt{\nu}_1+\sqrt{|\lambda_1|}} \\ &\leq 2\sqrt{2}\cdot\eta\sqrt{\nu}_1 + \frac{\eta\nu_1}{\sqrt{\nu}_1+\sqrt{|\lambda_1|}} \\ &\leq (1+2\sqrt{2})\eta\sqrt{\nu}_1.\end{aligned}$$

The desired result now follows. □

We may now use Lemma 7 to produce a perturbation bound for our banded matrix of phase differences $\tilde{X}_0$ from (6).

**Theorem 6.** *Let $\widetilde{X}_0 = T_\delta(\widetilde{\mathbf{x}}_0\widetilde{\mathbf{x}}_0^*)$ where $|(\widetilde{\mathbf{x}}_0)_i| = 1$ for each $i$. Further suppose $\widetilde{X} \in T_\delta(\mathcal{H}^d)$ has $\widetilde{\mathbf{x}}$ as its top eigenvector, where $||\widetilde{\mathbf{x}}||_2 = \sqrt{d}$. Suppose that $\|\widetilde{X}_0 - \widetilde{X}\|_F \leq \eta\|\widetilde{X}_0\|_F$ for some $\eta > 0$. Then, there exists an absolute constant $C \in \mathbb{R}^+$ such that*

$$\min_{\theta\in[0,2\pi]} \|\widetilde{\mathbf{x}}_0 - \mathrm{e}^{\mathrm{i}\theta}\widetilde{\mathbf{x}}\|_2 \leq C\frac{\eta d^3}{\delta^{\frac{5}{2}}}.$$

*Proof.* Recall that the phase vectors $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{x}}_0$ are normalized so that $\|\tilde{\mathbf{x}}\|_2 = \|\tilde{\mathbf{x}}_0\|_2 = \sqrt{d}$. Combining Lemmas 2 and 7 after noting that $\|\tilde{X}_0\|_{\mathrm{F}}^2 = d(2\delta-1)$ we learn that

$$\left(1 - \frac{1}{d^2}|\langle \tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}\rangle|^2\right) \leq C'\eta^2\left(\frac{d}{\delta}\right)^5 \tag{28}$$

for an absolute constant $C' \in \mathbb{R}^+$. Let $\phi \in [0, 2\pi)$ be such that $\mathrm{Re}\left(\langle \tilde{\mathbf{x}}_0, \mathrm{e}^{\mathrm{i}\phi}\tilde{\mathbf{x}}\rangle\right) = |\langle \tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}\rangle|$. Then,

$$\begin{aligned}\|\tilde{\mathbf{x}}_0 - \mathrm{e}^{\mathrm{i}\phi}\tilde{\mathbf{x}}\|_2^2 &= 2d - 2\,\mathrm{Re}\left(\langle \tilde{\mathbf{x}}_0, \mathrm{e}^{\mathrm{i}\phi}\tilde{\mathbf{x}}\rangle\right) \\ &= 2d\left(1 - \frac{1}{d}\,|\langle \tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}\rangle|\right) \leq 2d\left(1 - \frac{1}{d^2}\,|\langle \tilde{\mathbf{x}}_0, \tilde{\mathbf{x}}\rangle|^2\right).\end{aligned}$$

Combining this last inequality with (28) concludes the proof. □

[11] It is interesting to note that similar bounds can also be obtained using simpler techniques (see, e.g., [29]).

## Acknowledgements

The authors would like to thank Felix Krahmer for helpful discussions regarding Lemma 8. MI and RS would like to thank the Hausdorff Institute of Mathematics, Bonn for its hospitality during its Mathematics of Signal Processing Trimester Program. A portion of this work was completed during that time. In addition, the authors would like to thank and acknowledge the Institute for Mathematics and its Applications (IMA) for the workshop "Phaseless Imaging in Theory and Practice: Realistic Models, Fast Algorithms, and Recovery Guarantees" hosted there in a August of 2017. The revision of this paper benefitted greatly from many discussions with the participants there. In particular, conversations with James Fienup, Andreas Menzel, and Irene Waldspurger were exceedingly helpful. MI was supported in part by NSF DMS-1416752. RS was supported in part by a Hellman Fellowship and NSF DMS-1517204.