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 [27]. The improved algorithm 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 [27] are improved as a result for the new and more robust reconstruction approach proposed herein. Numerical experiments demonstrate robustness and computational efficiency that outperforms competing approaches on large problems. Finally, 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

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

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., [39, 32]), 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., [32]), diffraction imaging [20], and optics [32, 39], 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 [20]. Since then many variations of their method have been proposed (e.g, [19]) 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, 8, 9, 18, 27, 30]) 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, 9]) while in practice the measurement vectors are a deterministic aspect of the imaging apparatuses employed. A main contribution of this paper is analyzing a

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, Michigan State University, East Lansing, MI, 48824, USA (aditya@math.msu.edu).

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.

1.1. **Local Correlation Measurements.** Consider the case where the vectors $\mathbf{a}_j$ represent shifts of compactly-supported vectors $\mathbf{m}_j, j = 1, \ldots, K$ for some $K \in \mathbb{N}$. Using the notation $[n]_k := \{k, \ldots, k+n-1\} \subset \mathbb{N}$, and defining $[n] := [n]_1$ we take $\mathbf{x}_0, \mathbf{m}_j \in \mathbb{C}^d$ with $\text{supp}(\mathbf{m}_j) = [\delta] \subset [d]$ 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

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

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 [13, 27, 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

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

where the inner product above is the Hilbert-Schmidt inner product. Restricting to the case $P = [d]_0$, for every matrix $A \in \text{span}\{S_\ell^* \mathbf{m}_j \mathbf{m}_j^* S_\ell\}_{\ell,j}$ 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_δ is simply the orthogonal projection operator onto its range $T_\delta(\mathbb{C}^{d \times d}) \supseteq \text{span}\{S_\ell^* \mathbf{m}_j \mathbf{m}_j^* S_\ell\}_{\ell,j}$; therefore,

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

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

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

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

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

using our measurements $(\mathbf{y}_\ell)_j = (\mathcal{A}(\mathbf{x}_0 \mathbf{x}_0^*))_{(\ell,j)}$. In [27], 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 [27], 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

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

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

2

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

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

Indeed, in [38], 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

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

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

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

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

(8) $$\widetilde{X} := \frac{X}{|X|},$$

where the operations are considered elements. The hope is that the leading eigenvector of $\widetilde{X}$ still serves as a good approximation to the leading eigenvector of

(9) $$\widetilde{X}_0 := \frac{X_0}{|X_0|} := \frac{T_\delta(\mathbf{x}_0 \mathbf{x}_0^*)}{|T_\delta(\mathbf{x}_0 \mathbf{x}_0^*)|},$$

which is seen in Section 3 (see also [38]) to indeed be a scaled version of the phase vector

(10) $$\tilde{\mathbf{x}}_0 := \frac{\mathbf{x}_0}{|\mathbf{x}_0|}$$

(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.

Algorithm 1 Fast Phase Retrieval from Local Correlation Measurements

---

Input: Measurements $\mathbf{y} \in \mathbb{R}^D$ as per (6)
Output: $\mathbf{x} \in \mathbb{C}^d$ with $\mathbf{x} \approx \mathrm{e}^{-\mathrm{i}\theta} \mathbf{x}_0$ 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 $\widetilde{\mathbf{x}} := \mathrm{sgn}(u)$.
  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. *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*

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

where $\kappa > 0$ is the condition number of the system (7) and $C \in \mathbb{R}^+$ is an absolute universal constant.

Theorem 1, which deterministically depends on both the masks and the signal, is a strict improvement over the first deterministic theoretical robust recovery guarantees proven in [27] for a wide class of non-vanishing signals.

Consider 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 [27]). 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., [37]). In the low-noise setting the top eigenvector $\tilde{\mathbf{x}}$ can be computed to machine precision in $\mathcal{O}(\log d)$ such iterations,[1] 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.

1.3. **Connection to Ptychography.** In ptychographic imaging (see Fig. 1), small regions of a specimen are illuminated one at a time and an intensity[2] 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 [21, 16]

$$(11) \qquad y(t, \omega) = \left| \mathcal{F}[\widetilde{h} \cdot S_t f](\omega) \right|^2 + \eta(t, \omega).$$

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

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

---

[1]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 $k = \Omega(\log_3 d)$.

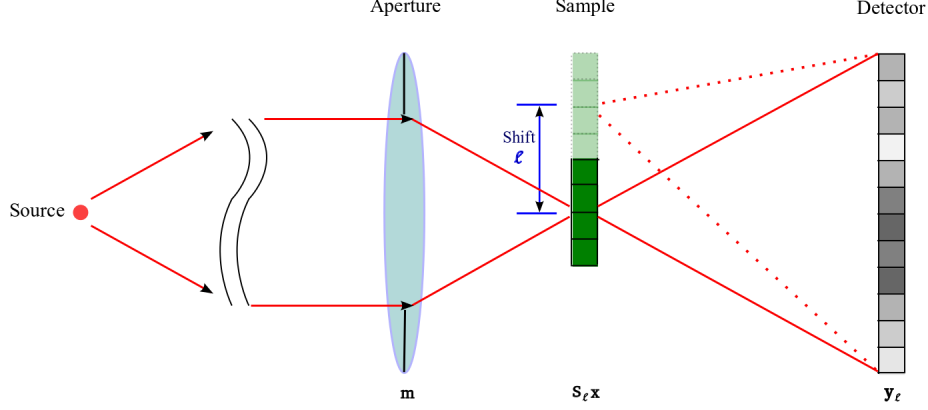[2]By intensity, we mean magnitude squared.

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

and $\widetilde{h} : [0, 1] \to \mathbb{C}$ is the so-called illumination function [40] of the imaging system. To account for the local nature of the measurements in (11), 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 (11), with $\widetilde{\mathbf{m}}, \mathbf{x}_0 \in \mathbb{C}^d$ discretizing $\widetilde{h}$ and f. Thus (11), in the absence of noise, becomes

$$(12) \qquad (\mathbf{y}_\ell)_j = \left| \sum_{n=1}^d \widetilde{m}_n \, (x_0)_{n+\ell} \, \mathrm{e}^{-\frac{2\pi \mathrm{i}(j-1)(n-1)}{d}} \right|^2, \quad (j, \ell) \in [d] \times [d]_0,$$

where indexing is considered modulo-d, so $(\mathbf{y}_\ell)_j$ is a diffraction measurement corresponding to the jth Fourier mode of a circular ℓ-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 (12) 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

$$(13) \qquad (\mathbf{m}_j)_n = \overline{\widetilde{m}_n} \, \mathrm{e}^{\frac{2\pi \mathrm{i}(j-1)(n-1)}{d}}$$

and rearranging (12), we obtain

$$(14) \qquad (\mathbf{y}_\ell)_j = \left| \sum_{n=1}^d (x_0)_{n+\ell} \, \overline{(\mathbf{m}_j)_n} \right|^2 = \left| \sum_{n=1}^\delta (x_0)_{n+\ell} \, \overline{(\mathbf{m}_j)_n} \right|^2$$
$$= |\langle S_\ell \mathbf{x}_0, \mathbf{m}_j \rangle|^2 = \langle S_\ell \mathbf{x}_0 \mathbf{x}_0^* S_\ell^*, \mathbf{m}_j \mathbf{m}_j^* \rangle$$
$$= \langle T_\delta(\mathbf{x}_0 \mathbf{x}_0^*), S_\ell^* \mathbf{m}_j \mathbf{m}_j^* S_\ell \rangle, \quad (j, \ell) \in [d] \times [d]_0$$

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 (14) defines a correlation with local masks or window functions $\mathbf{m}_j$. More importantly, (14) shows that ptychography (with ℓ ranging over any subset of $[d]_0$) represents a case of the general system seen in (3).

1.4. **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

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

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

$$(15) \qquad (\mathbf{y}_\ell)_k = \left| \sum_{n=1}^{d} (x_0)_n\, m_{n-\ell}\, \mathbb{e}^{-\frac{2\pi \mathrm{i}(k-1)(n-1)}{d}} \right|^2 , \;\; k \in [d],\; \ell \in \{\ell_1, \ldots, \ell_L\} \subset [d]_0.$$

As before, ℓ denotes a shift or translation of the mask/window, so $(\mathbf{y}_\ell)_k$ corresponds to the (squared magnitude of) the kth Fourier mode associated with an ℓ-shift[3] of the mask $\mathbf{m}$. Defining the modulation operator, $W_k : \mathbb{C}^d \mapsto \mathbb{C}^d$, by its action $(W_k \mathbf{x}_0)_n = e^{2\pi \mathrm{i}(k-1)(n-1)/d}\,(x_0)_n$ and applying elementary Fourier transform properties[4] one has

$$(\mathbf{y}_\ell)_k = |\langle \mathbf{x}_0, S_{-\ell}(e^{2\pi \mathrm{i}(k-1)\ell/d} W_k \overline{\mathbf{m}}) \rangle|^2$$

$$= |\langle \mathbf{x}_0, S_{-\ell}(W_k \overline{\mathbf{m}}) \rangle|^2 = |\langle \widehat{\mathbf{x}_0}, S_{-\ell}\widehat{(W_k \overline{\mathbf{m}})} \rangle|^2$$

$$= |\langle \widehat{\mathbf{x}_0}, W_{-\ell+1}(S_{-k+1}\widehat{\overline{\mathbf{m}}}) \rangle|^2$$

$$(16) \qquad = |\langle \widehat{\mathbf{x}_0}, S_{-k+1}(W_{-\ell+1}\widehat{\overline{\mathbf{m}}}) \rangle|^2.$$

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

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

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

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.

1.5. **Related Work.** The first approaches to the phase retrieval problem were proposed in the 1970's in [20] by Gerchberg and Saxton, and were famously improved in [19] later that decade. Though these techniques work well in practice and have been popular for decades, they are notoriously difficult to analyze. These iterative methods 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 [31]. 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 [10, 13], gradient descent strategies [12], graph-theoretic [1] and frame-based approaches [3, 8], and variants on the alternating minimization (e.g., with resampling) [33].

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* [11, 24], and *Wirtinger Flow* [12]. However, these measurements are both randomized (which is crucial

---

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

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

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 [17, 7] and [29], 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 [35] in order to produce recovery guarantees in the noiseless setting. More recently, randomized robustness guarantees were developed for time-frequency measurements in [34]. 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 [27], the authors propose the measurement scheme developed in the current paper and prove the first deterministic robustness results for a different greedy recovery algorithm.

1.6. **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 [27]. 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, while Section 7 contains some concluding remarks and avenues for further research. In the appendix, 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 [27], 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

$$(17) \qquad (\mathbf{m}_j)_n = \begin{cases} \frac{e^{-n/a}}{\sqrt[4]{2\delta-1}} \cdot e^{\frac{2\pi i \cdot (n-1) \cdot (j-1)}{2\delta-1}} & \text{if } n \leq \delta \\ 0 & \text{if } n > \delta \end{cases}.$$

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

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 [27] 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 holds.

Theorem 2 ([27]). *Consider measurements of the form* (17) *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 (17). This measurement scheme is also interesting because it corresponds to a ptychography system if we take the illumination function in (12) to be $\widetilde{m}_n = \frac{\mathrm{e}^{-i/a}}{\sqrt[4]{2\delta-1}}$ and assume that $d = k(2\delta-1)$ for some $k \in \mathbb{N}$. Then we may take the subset of the measurements (13) given by $j = (p-1)k+1$, $p \in [2\delta-1]$ to obtain the masks specified in (17). Consider that these assumptions may easily be met by zero-padding $\mathbf{x}_0$ until $d = k(2\delta-1)$ and simply "throwing away" all the measurements but those that correspond to $j = (p-1)k+1$, $p \in [2\delta-1]$.

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 $j = 1,\ldots,\delta-1$. 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 $\mathbf{m}_1,\ldots,\mathbf{m}_{2\delta-1}$ 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 $\tilde{\mathbf{x}}_0 := \frac{\mathbf{x}_0}{|\mathbf{x}_0|}$ via an eigenvector method. Here, we show that $\tilde{X}_0$ has $\tilde{\mathbf{x}}_0$ as its top eigenvector and we investigate the spectral properties of $\tilde{X}_0$ 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.

To begin, consider $U = T_\delta(\mathbb{1}\mathbb{1}^*)$, i.e.,

(18)
$$U_{j,k} = \begin{cases} 1 & \text{if } |j - k| \bmod d < \delta \\ 0 & \text{otherwise} \end{cases}.$$

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

(19)
$$\nu_j = \sum_{k=1}^d (U)_{1,k} \omega_j^{k-1} = 1 + \sum_{k=1}^{\delta-1} \omega_j^k + \omega_j^{-k} = 1 + 2 \sum_{k=1}^{\delta-1} \cos\left(\frac{2\pi(j-1)k}{d}\right),$$

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

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

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

We consider that $\tilde{X}_0$ and U are similar; indeed $\tilde{X}_0 = \tilde{D}_0 U \tilde{D}_0^*$, where $\tilde{D}_0 = \text{diag}\{(\tilde{x}_0)_1, \ldots, (\tilde{x}_0)_d\}$. Since $|(\tilde{\mathbf{x}}_0)_j| = 1$ for each j, we have that $\tilde{D}_0$ is unitary. Thus the eigenvalues of $\tilde{X}_0$ are given by (19), and its eigenvectors are simply the discrete Fourier vectors modulated by the entries of $\tilde{\mathbf{x}}_0$. We now have the following lemma.

9

Lemma 1. *Let $\widetilde{X}_0$ be defined as in (9). 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 Λ 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

$$(20) \qquad \nu_{j+1} = l_\delta(\theta_j) = \frac{\sin(\theta_j(\delta - 1/2))}{\sin(\theta_j/2)}.$$

Thus, the eigenvalues of $\widetilde{X}_0$ are sampled from the $(\delta - 1)^{\mathrm{st}}$ 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 (20), and the symmetry of l_δ about $\theta = \pi$.

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

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

10

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 ν_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 (19), this will give $\sum_{k=1}^{\delta - 1}\cos(k\theta_1) \leq (\delta - 1)\cos\left(\theta_1\frac{\delta}{2}\right)$, so

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

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 $\widetilde{X}_0$

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 $\min_{\theta \in [0, 2\pi]}\|\widetilde{\mathbf{x}}_0 - \mathrm{e}^{\mathrm{i}\theta}\widetilde{\mathbf{x}}\|_2$. This approach involves considering both $\widetilde{X}$ from Algorithm 1 and $\widetilde{X}_0$ from (9) 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 (10), and whose edges carry the relative phase data.[5]

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

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

---

[5]The interested reader is also referred to the appendix 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.

respectively. The **volume** of G is
$$\mathrm{vol}(G) := \sum_{i \in V} \deg(i).$$
Finally, the **Laplacian** of G is the $d \times d$ real symmetric matrix
$$L := I - D^{-1/2}WD^{-1/2} = D^{-1/2}(D - W)D^{-1/2},$$
where $I \in \{0,1\}^{d \times d}$ is the identity matrix.

When G is connected, Lemma 1.7 of [14] shows that the nullspace of $(D - W)$ is $\mathrm{span}(\mathbb{1})$, and the nullspace of L is $\mathrm{span}(D^{1/2}\mathbb{1})$. 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 (18). 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 (18). 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 η_{ij} we will have

$$(21) \qquad \widetilde{X}_{ij} = \frac{(x_0)_i(x_0)_j^* + \eta_{ij}}{|(x_0)_i(x_0)_j^* + \eta_{ij}|} \cdot w_{ij} = \frac{(x_0)_i(x_0)_j^* + \eta_{ij}}{|(x_0)_i(x_0)_j^* + \eta_{ij}|} \cdot \chi_{E(i,j)}.$$

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

$$(22) \qquad (\widetilde{X}_0)_{ij} = \frac{(x_0)_i(x_0)_j^*}{|(x_0)_i(x_0)_j^*|} \cdot w_{ij} = \mathrm{sgn}\left((x_0)_i(x_0)_j^*\right) \cdot \chi_{E(i,j)}.$$

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

$$(23) \qquad L_1 = I - D^{-1/2}(\widetilde{X} \circ W)D^{-1/2},$$

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

$$(24) \qquad \eta_{\widetilde{X}}(\mathbf{y}) := \frac{1}{2} \frac{\sum_{(i,j) \in E} w_{ij}|y_i - \widetilde{X}_{ij}y_j|^2}{\sum_{i \in V} \deg(i)|y_i|^2} = \frac{\mathbf{y}^*(D - (\widetilde{X} \circ W))\mathbf{y}}{\mathbf{y}^* D \mathbf{y}}.$$

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 (23) corresponding to its smallest eigenvalue. Then, $\mathbf{w} = \mathrm{sgn}(\mathbf{u}) = \mathrm{sgn}\left(D^{-1/2}\mathbf{u}\right)$ satisfies*

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

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

Proof. One can see that

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

From here, Lemma 3.6 in [6] gives

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

$\square$

We now state the main result of this section:

Theorem 3. *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 (23) corresponding to its smallest eigenvalue, and let*

$$\widetilde{\mathbf{x}} = \mathrm{sgn}(\mathbf{u}) \; and \; \widetilde{\mathbf{x}}_0 = \mathrm{sgn}(\mathbf{x}_0).$$

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

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

where $\widetilde{X}$ and $\widetilde{X}_0$ are defined as per (21) and (22), 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 3, there exists an angle $\theta \in [0, 2\pi]$ such that*

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

13

Lemma 5. *Under the hypotheses of Theorem 3, 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 $\sum_{i\in V} |g_i - \mathrm{e}^{\mathrm{i}\theta}|^2 = |\widetilde{\mathbf{x}} - \mathrm{e}^{\mathrm{i}\theta}\widetilde{\mathbf{x}}_0|_2^2$.

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 $\sum_{i\in V}\deg(i)|g_i - \mathrm{e}^{\mathrm{i}\theta}|^2$. 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. $\qquad\square$

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

Lemma 3 now tells us that

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

14

Moreover,

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

so that $\sum_{(i,j)\in E} \frac{1}{2}|g_i - g_j|^2 \le \frac{C'}{\tau}\|X_0 - X\|_F^2 + \sum_{(i,j)\in E} |\Lambda_{ij} - 1|^2$. 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 $\tau \le 1$, this completes the proof. $\qquad\square$

We may now use Theorem 3 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 (9), $\widetilde{\mathbf{x}}_0$ be the vector of true phases (10), 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 $\|\widetilde{X}_0 - \widetilde{X}\|_F \le \eta\|\widetilde{X}_0\|_F$ for some $\eta > 0$. Then, there exists an absolute constant $C' \in \mathbb{R}^+$ such that*

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

Proof. We apply Theorem 3 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 $\|\widetilde{X}_0 - \widetilde{X}\|_F \le C''\eta(d\delta)^{1/2}$. 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

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

$\qquad\square$

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 (6) 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 $\|X - X_0\|_F \le \kappa\|\mathbf{n}\|_2$.

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 ρ-small indexes of

$\mathbf{x}_0$ to be

(25)
$$S_\rho := \left\{ j \;\middle|\; |(x_0)_j| < \left(\frac{\kappa \|\mathbf{n}\|_2}{\rho} \right)^{\frac{1}{4}} \right\}$$

where $\rho \in \mathbb{R}^+$ is a free parameter. With the definition of S_ρ 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 $j, k \in S_\rho^c$,

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

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 $\|\tilde{X}_0\|_F = \sqrt{(2\delta - 1)d}$. $\qquad \square$

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

Theorem 4. *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*

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

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_ρ from (25) as

(26)
$$\min_{\theta \in [0,2\pi]} \left\| \mathbf{x}_0 - \mathbb{e}^{\mathbb{i}\theta} \mathbf{x} \right\|_2 \leq C' \|\mathbf{x}_0\|_\infty \left(\frac{d}{\delta} \right)^2 \sqrt{\rho \frac{\kappa}{\delta} \|\mathbf{n}\|_2 + |S_\rho|} + C' d^{\frac{1}{4}} \sqrt{\kappa \|\mathbf{n}\|_2},$$

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 ϕ. As a result we have that

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

for some absolute constant $C'' \in \mathbb{R}^+$. Here the bound on the second term follows from Lemma 3 of [27] 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 (26) in Theorem 4 we can see that the error bound there will be vacuous in most settings unless $S_\rho = \emptyset$. Recalling (25), one can see that S_ρ 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 ρ in (26) leads to the following corollary of Theorem 4.

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*

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

where $C \in \mathbb{R}^+$ 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 [27]. Finally, we note that a trivial variant of Corollary 3 can also be combined with the discussion in §1.4 in order to generate recovery guarantees for the windowed Fourier measurements defined by (15). 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. Our main objective is to evaluate the proposed algorithm against other existing phase retrieval methods using *local* measurements. However, 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). The results presented here may be recreated using the open source *BlockPR* Matlab software package which is freely available at [28]. 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™M-5Y10c processor, 8GB RAM and Matlab R2016a. 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 esetimates 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 δ 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 δ 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 $1 \leq \gamma \leq \delta$ and with shifts $s \leq \gamma$ (dividing d), having entries

$$(\tilde{D}_{j'})_{k,h} = |X|_{(sj'+k-1) \bmod d, \ (sj'+h-1) \bmod d}.$$

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 in order 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.

6.2. Experiments.

We begin by presenting results in Fig. 2a demonstrating the improved noise robustness of the proposed method over the formulation in [27]. Recall that [27] 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. The local correlation measurements described in Example 2 of Section 2 are utilized in this plot in and all the ensuing experiments unless otherwise indicated. Three variants of the proposed algorithm are plotted in Fig. 2a:

(1) an implementation of Algorithm 1 (denoted by $\square$'s),
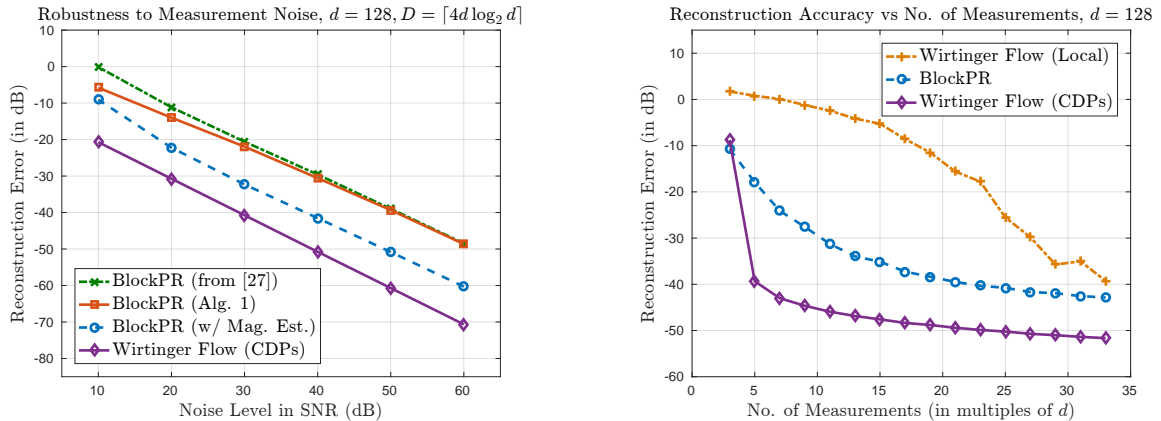
(A) Improved Robustness to Measurement Noise – Comparing Variants of the *BlockPR* algorithm

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

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

(2) an implementation of Algorithm 1 with the improved magnitude estimation procedure detailed above (with $s = 1$ and using the average of the obtained $\tilde{D}_j$, block magnitude estimates) and post-processed using 100 iterations of the *Gerchberg–Saxton* alternating projection algorithm (denoted by ◦'s), and

(3) the algorithmic implementation from [27] (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 [27]. Moreover, the magnitude estimation procedure detailed above yields significant improvement in reconstruction errors over the two other variants; consequently, this implementation is used in all plots henceforth. For reference, we also include reconstruction errors with the *Wirtinger Flow* algorithm (denoted by ◇'s) when using (*global*) coded diffraction pattern (CDP) measurements. Clearly, using global measurements such as coded diffraction patterns provides superior noise tolerance; however, they are not applicable to imaging modalities such as ptychography. 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.

Given the weaker performance of *Wirtinger Flow* with local measurements, we now restrict our attention to the empirical evaluation of the proposed method against the *PhaseLift* and *Gerchberg-Saxton* 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 [23, 22] – a package for specifying and solving convex programs in Matlab; the alternating projection method
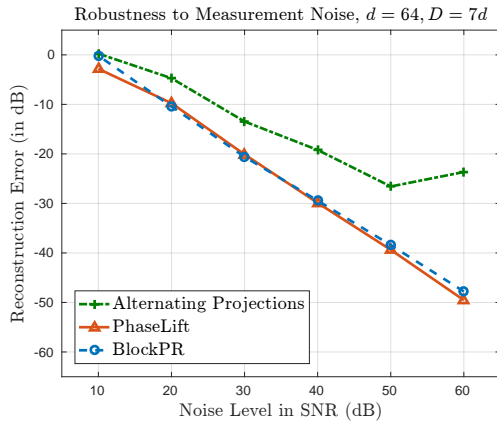
(A) Using $D = 7d$ measurements.
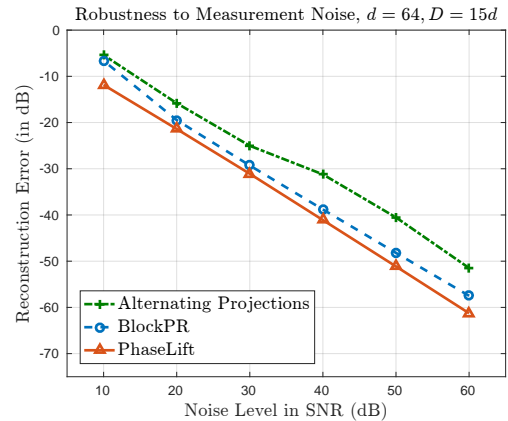
(B) Using $D = 15d$ measurements.

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

was initialized with a random complex Gaussian initial guess and limited to a maximum of $10,000$ iterations. 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. In particular, the well-conditioned *deterministic* 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. In particular, the method performs almost as well as the *PhaseLift* algorithm and returns significantly more accurate reconstructions than the alternating projections algorithm. We remark that the marginally improved noise robustness of *PhaseLift* is at the expense of a significant increase in computational cost, as we will see in Fig. 4b.

Next, Fig. 4a plots the reconstruction error in recovering a $d = 64$-length complex vector as a function of the number of measurements used. 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. We observe that the proposed algorithm outperforms the popular alternating projections method, and is almost as accurate as *PhaseLift*. Moreover, at the 40 dB noise level, the proposed method provides the best reconstruction accuracy when using small numbers of measurements $(D \approx 5d)$ which may be of practical importance.

Finally, Fig. 4b plots the average execution time (in seconds) required to solve the phase retrieval problem using $D = \lceil 2d \log_2 d \rceil$ noiseless measurements. For comparison, execution times for the *PhaseLift* and alternating projection algorithms are provided. We observe that the proposed method is several orders of magnitude faster than the *PhaseLift* and alternating projection algorithms. Moreover, the plot confirms the essentially FFT-time computational complexity (see Section 1) of the proposed method.
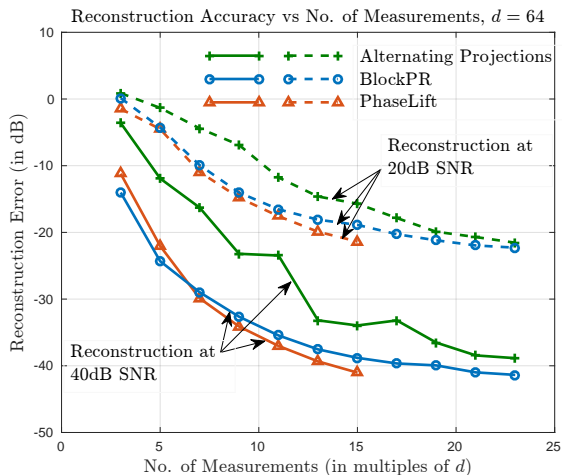
(A) Reconstruction Error vs. No. of Measurements
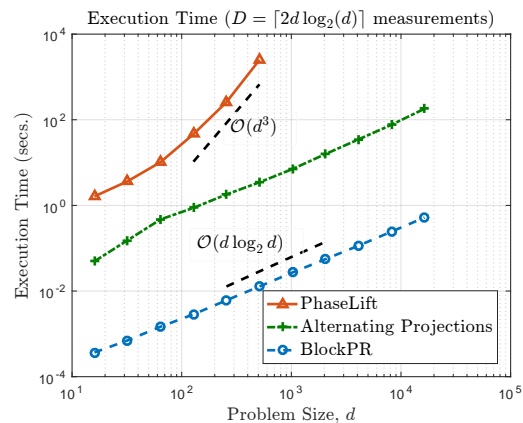
(B) Execution Time vs. Problem Size

FIGURE 4. Performance Evaluation and Comparison of the Proposed Phase Retrieval Method (with Deterministic Local Correlation Measurements and Additive Gaussian Noise)

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 1.4) 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.

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. Bandeira and D. Mixon. Near-optimal phase retrieval of sparse vectors. *arXiv preprint arXiv:1308.0143*, 2013.

[6] A. S. Bandeira, A. Singer, and D. A. Spielman. A Cheeger Inequality for the Graph Connection Laplacian. *ArXiv e-prints*, Apr. 2012.

[7] T. Bendory and Y. C. Eldar. Non-convex phase retrieval from STFT measurements. *CoRR*, abs/1607.08218, 2016.

[8] B. G. Bodmann and N. Hammen. Stable phase retrieval with low-redundancy frames. *ArXiv e-prints*, Feb. 2013.

[9] E. Candès, T. Strohmer, and V. Voroninski. Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 2012.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] F. Chung. *Spectral Graph Theory*. Number no. 92 in CBMS Regional Conference Series. Conference Board of the Mathematical Sciences, 1992.

[15] C. Davis and W. M. Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970.

[16] 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.

[17] 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.

[18] Y. C. Eldar and S. Mendelson. Phase retrieval: Stability and recovery guarantees. *Applied and Computational Harmonic Analysis*, 36(3):473–494, 2014.

[19] J. R. Fienup. Reconstruction of an object from the modulus of its fourier transform. *Optics letters*, 3(1):27–29, 1978.

[20] R. Gerchberg and W. Saxton. A practical algorithm for the determination of the phase from image and diffraction plane pictures. *Optik*, 35:237246, 1972.

[21] J. W. Goodman. *Introduction to Fourier optics*. Roberts and Company Publishers, 2005.

[22] 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`.

[23] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, Mar. 2014.

[24] D. Gross, F. Krahmer, and R. Kueng. Improved recovery guarantees for phase retrieval from coded diffraction patterns. *Applied and Computational Harmonic Analysis*, 2015.

[25] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 2 edition, 2012.

[26] 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.

[27] M. Iwen, A. Viswanathan, and Y. Wang. Fast phase retrieval from local correlation measurements. *SIAM Journal on Imaging Sciences*, 9(4):1655–1688, 2016.

[28] 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.

[29] K. Jaganathan, Y. C. Eldar, and B. Hassibi. Stft phase retrieval: Uniqueness guarantees and recovery algorithms. *arXiv preprint arXiv:1508.02820*, 2015.

[30] X. Li and V. Voroninski. Sparse signal recovery from quadratic measurements via convex programming. *SIAM Journal on Mathematical Analysis*, 45(5):3019–3033, 2013.

[31] S. Marchesini, Y.-C. Tu, and H.-t. Wu. Alternating projection, ptychographic imaging and phase synchronization. *Applied and Computational Harmonic Analysis*, 2015.

[32] R. Millane. Phase retrieval in crystallography and optics. *J. Opt. Soc. Am. A*, 7(3):394–411, 1990.

[33] P. Netrapalli, P. Jain, and S. Sanghavi. Phase retrieval using alternating minimization. In *Advances in Neural Information Processing Systems*, pages 2796–2804, 2013.

[34] G. E. Pfander and P. Salanevich. Robust phase retrieval algorithm for time-frequency structured measurements. *eprint arXiv:1611.02540*, 2016.

[35] 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.

[36] G. Stewart and J. Sun. *Matrix Perturbation Theory*. Academic Press, 1990.

[37] L. N. Trefethen and D. Bau III. *Numerical linear algebra*, volume 50. Siam, 1997.

[38] A. Viswanathan and M. Iwen. Fast angular synchronization for phase retrieval via incomplete information. *Proc. of SPIE Optics + Photonics*, 2015.

[39] A. Walther. The Question of Phase Retrieval in Optics. *Optica Acta*, 10:41–49, 1963.

[40] C. Yang, J. Qian, A. Schirotzek, F. Maia, and S. Marchesini. Iterative Algorithms for Ptychographic Phase Retrieval. *ArXiv e-prints*, May 2011.

[41] Y. Yu, T. Wang, and R. Samworth. A useful variant of the davis–kahan theorem for statisticians. *Biometrika*, 102(2):315–323, 2015.

Appendix

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 5), 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 [15, 36] (see, e.g., the proof of Corollary 1 in [41]) 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

$$(27) \qquad \left(1 - |\langle \mathbf{x}_1, \mathbf{v}_1 \rangle|^2\right) = \sin^2\left(\arccos\left(|\langle \mathbf{x}_1, \mathbf{v}_1 \rangle|\right)\right) \leq \frac{4\eta^2 \|X_0\|_{\mathrm{F}}^2}{(\nu_1 - \nu_2)^2},$$

giving us the desired inequality. $\square$

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 [10].

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 λ_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*

$$\min_{\theta \in [0,2\pi]} \|\mathrm{e}^{\mathrm{i}\theta} \mathbf{x}_0 - \sqrt{|\lambda_1|} \mathbf{v}_1\|_2 \leq (1 + 2\sqrt{2})\eta \|\mathbf{x}_0\|_2.^6$$

---

[6] It is interesting to note that similar bounds can also be obtained using simpler techniques (see, e.g., [26]).

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,

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

(28)

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., [25]), and (28), we have

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

The desired result now follows. $\qquad\square$

We may now use Lemma 7 to produce a perturbation bound for our banded matrix of phase differences $\widetilde{X}_0$ from (9).

Theorem 5. *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 $\|\widetilde{X}_0 - \widetilde{X}\|_F \le \eta\|\widetilde{X}_0\|_F$ for some $\eta > 0$. Then, there exists an absolute constant $C \in \mathbb{R}^+$ such that*

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

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

(29)
$$
\left(1 - \frac{1}{d^2}|\langle \widetilde{\mathbf{x}}_0, \widetilde{\mathbf{x}} \rangle|^2\right) \le C'\eta^2 \left(\frac{d}{\delta}\right)^5
$$

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

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

Combining this last inequality with (29) concludes the proof. $\qquad\square$