

Robust Phase Retrieval Algorithm for Time-Frequency Structured Measurements*

Götz E. Pfander[†] and Palina Salanevich[‡]

Abstract. We address the problem of signal reconstruction from intensity measurements with respect to a measurement frame. This nonconvex inverse problem is known as *phase retrieval*. The case considered in this paper concerns phaseless measurements taken with respect to a Gabor frame. It arises naturally in many practical applications, such as diffraction imaging and speech recognition. We present a reconstruction algorithm that uses a nearly optimal number of phaseless time-frequency structured measurements and discuss its robustness in the case when the measurements are corrupted by noise. We show how geometric properties of the measurement frame are related to the robustness of the phaseless reconstruction.

Key words. phase retrieval, Gabor frames, expander graphs, order statistics of frame coefficients, angular synchronization, spectral clustering

AMS subject classifications. 49N30, 42C15, 05C50, 60B20, 94A12

DOI. 10.1137/18M1205522

1. Introduction. The *phase retrieval* problem arises naturally in many applications within a variety of fields in science and engineering. Among these applications are optics [3], astronomical imaging [4], and microscopy [5].

As an example, let us consider the diffraction imaging problem [6]. To investigate the structure of a small particle, such as a DNA molecule, we illuminate the particle with X-rays and then measure the radiation scattered from it. When X-ray waves pass by an object and are measured in the far field, detectors are not able to capture the phase of the waves reaching them, but only their magnitudes. The measurements obtained in this way are of the form of pointwise squared absolute values of the Fourier transform of the object x , that is, the measurement map \mathcal{A} is given by $\mathcal{A}(x) = \{|\mathcal{F}(x)(n)|^2\}_{n \in \Omega}$, where Ω is the sampling grid. Since \mathcal{A} is not injective, some additional a priori information on the object x is needed for reconstruction. For instance, knowledge of the chemical interactions between parts of a DNA molecule were used for the construction of the DNA double helix model in the Nobel Prize winning work of Watson, Crick, and Wilkins [7].

One way to overcome this noninjectivity when no a priori information is available is *masking*. To modify the phase front, one can insert a known mask after the object, as shown on Figure 1. The measurement map in this case is given by $\mathcal{A}(x) = \{|\mathcal{F}(x \odot f_t)(n)|^2\}_{n \in \Omega, t \in I}$,

*Received by the editors August 6, 2018; accepted for publication (in revised form) February 7, 2019; published electronically April 16, 2019.

<http://www.siam.org/journals/siims/12-2/M120552.html>

Funding: The work of the second author was supported by a Research Grant for Doctoral Candidates and Young Academics and Scientists of German Academic Exchange Service (DAAD).

[†]Department of Scientific Computing, Catholic University of Eichstätt-Ingolstadt, 85072 Germany (pfander@ku.de).

[‡]Department of Mathematics, University of California, Los Angeles, CA 90095-1555 (psalanevich@math.ucla.edu).

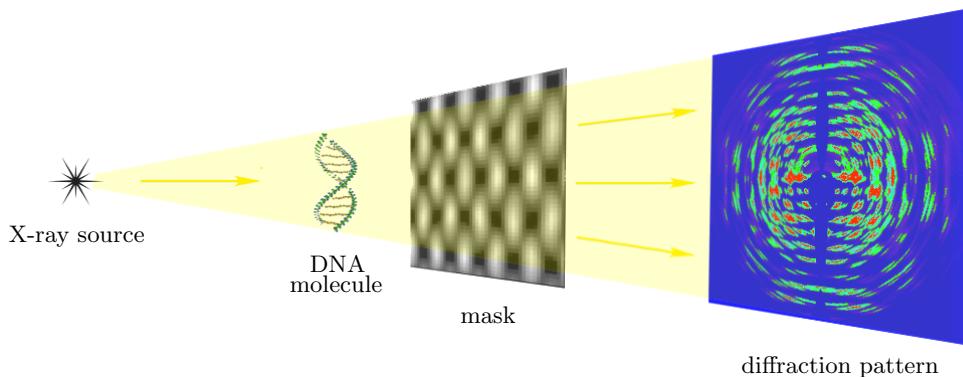


Figure 1. A typical setup for structured illuminations in diffraction imaging using a phase mask.

where f_t , $t \in I$, are the masks used, and \odot denotes pointwise multiplication. By increasing the number of measurements in this way, we reduce the ambiguity in the reconstruction of signal x .

Since the problem of signal reconstruction from magnitudes of Fourier coefficients is particularly hard to handle, a more general frame theoretical setting is frequently considered. Namely, for $\Phi = \{\varphi_j\}_{j=1}^N \subset \mathbb{C}^M$ being a *frame*, that is, a possibly overcomplete spanning set for \mathbb{C}^M , we aim to recover $x \in \mathbb{C}^M$ from its phaseless squared frame coefficients $\mathcal{A}_\Phi(x) = \{|\langle x, \varphi_j \rangle|^2\}_{j=1}^N$.

Note that the masked Fourier coefficients of the signal x with masks $\{f_t\}_{t \in I} \subset \mathbb{C}^M$ can be viewed as the frame coefficients of x with respect to the frame Φ given by $\Phi = \{\varphi_{t,j}\}$, where $\varphi_{t,j}(m) = \frac{e^{2\pi i j m/M}}{\sqrt{M}} \bar{f}_t(m)$.

It is clear that, even in an optimal setting, x can be reconstructed from intensity measurements only up to a global phase. Indeed, for every real θ , the signals x and $e^{i\theta}x$ produce the same intensity measurements. Thus, the goal of phase retrieval is to reconstruct the equivalence class $[x] \in \mathbb{C}/\sim$ of x , where $x \sim y$ if and only if $x = e^{i\theta}y$ for some $\theta \in [0, 2\pi)$. In the following, we are going to identify x with its equivalence class $[x] \in \mathbb{C}/\sim$.

Obviously, not every frame gives rise to an injective measurement map. But even in the case when \mathcal{A}_Φ is known to be injective, the problem of reconstructing $[x]$ from $\mathcal{A}_\Phi(x)$ is NP-hard in general [8]. So, the main goals in this area of applied mathematics is to find conditions on the number of measurements N and vectors φ_j for which there exists an efficient and robust numerical recovery algorithm.

Until recently, very little was known on how to achieve robust and efficient reconstruction given injectivity. Many practical methods used today have their origins in the alternating projection algorithms proposed in the 1970s by Gerchberg and Saxton [9]. Due to their lack of global convergence guarantees, the problem of developing fast phase retrieval algorithms which have provable recovery and robustness guarantees receives significant attention today. Some of the most prominent suggested algorithms are PhaseLift [10, 11, 12, 13, 14], Wirtinger flow algorithms [15, 16], a fast phase retrieval algorithm from local correlation measurements [17], and phase retrieval with polarization [18, 19]. The latter is described in more detail in section 2.1.

While recovery guarantees have been established for all these algorithms, most of them require the measurement frame vectors to be independent random vectors. Since measurements of this type are not implementable in practice, the design of fast and stable recovery algorithms with a small number of application relevant, structured, measurements remains an important problem. We address this problem below.

1.1. Main result. We study the phase retrieval problem for *Gabor frame* measurements, that is, the case when frame vectors are given by time and frequency shifts of a randomly chosen vector, called the *Gabor window* (see section 2.2 for a precise definition). The main motivation for using Gabor frames is that Gabor frame coefficients are masked Fourier coefficients, where the masks are time (or space) shifts of the Gabor window. This makes measurements implementable in applications while preserving the flexibility of the frame-theoretic approach.

Apart from diffraction imaging, phase retrieval with Gabor frames also arises, for example, in speech recognition problems. The use of a noisy phase or its estimation is a critical problem in speech recognition that may prevent the accurate reconstruction of a signal. There is a long-standing belief that speech recognition should be independent of phase. Balan, Casazza, and Edidin addressed this conjecture by using the phase retrieval framework to show that the reconstruction of a signal is possible without using a phase or its estimation for a generic frame [20]. However, construction of Gabor frames with such a property still remains an open problem. Previous work on phase retrieval with Gabor frames concentrated on injectivity conditions for full Gabor frames and shows reconstruction from M^2 time-frequency structured measurements of an M -dimensional signal [21], while no injectivity results for Gabor frames of smaller cardinality are available to this date.

Based on the idea of polarization [18], we propose a reconstruction algorithm for time-frequency structured measurements and investigate its robustness in the case when measurements are corrupted by noise. More precisely, we consider measurements of the form

$$(1) \quad |\langle x, \varphi_j \rangle|^2 + \nu_j, \quad \varphi_j \in \Phi,$$

where ν_j are noise terms and $\Phi = \Phi_V \cup \Phi_E$ is the measurement frame with a Gabor frame Φ_V given by (7) and a set of vectors for additional measurements Φ_E given by (8). We note that, while the frame Φ used for reconstruction is not a Gabor frame, the set of vectors for the additional measurements Φ_E also obeys a time-frequency structure and measurements of a signal x with respect to Φ_E have the form of windowed Fourier transform measurements; see section 3.1 for the details. We prove the following result, a more precise formulation of which we state in section 4.2 as Theorem 4.7.

Theorem 1.1. *Fix $x \in \mathbb{C}^M$ and consider the time-frequency structured phaseless measurements given by (1). If the noise vector ν satisfies $\frac{\|\nu\|_2}{\|x\|_2^2} \leq \frac{c}{M}$ for some c sufficiently small, then there exists a numerical constant C independent of M , so that for the estimate \tilde{x} produced by Algorithm 3 the following holds with overwhelming probability:*

$$(2) \quad \min_{\theta \in [0, 2\pi)} \|\tilde{x} - e^{i\theta} x\|_2^2 \leq \frac{C\sqrt{M}\|\nu\|_2}{\Delta}.$$

Here $\Delta = \min_{\Lambda' \subset \Lambda, |\Lambda'| \geq 2/3|\Lambda|} \sigma_{\min}^2(\Phi_{\Lambda'}^*)$ is the numerically erasure-robust frame bound for the Gabor frame Φ_V .

The reconstruction algorithm we propose in this paper requires a close to optimal number $N = O(M \log M)$ of time-frequency structured measurements. To the best of our knowledge, Theorem 1.1 provides the best existing robustness guarantee for measurements obeying time-frequency structure. Moreover, our result guarantees exact recovery of x (up to a global phase) in the noiseless case, that is, when $\nu = 0$. We note that the reconstruction guarantees given in Theorem 1.1 are nonuniform, in the sense that inequality (2) holds with high probability for each particular signal x , but not for all $x \in \mathbb{C}^M$ simultaneously. Obtaining a uniform version on Theorem 1.1 is one of the main directions for future work. Numerical experiments, presented in section 5, verify the robustness of the proposed phase retrieval algorithm and illustrate dependencies of the error to noise ratio on various parameters.

Remark 1.2. The bound obtained in Theorem 1.1 is similar, up to a log factor, to the recovery guarantees obtained in [18] for phase retrieval with random Gaussian frames with independent frame vectors. Note that, in (2), the norm of the reconstruction error $\tilde{x} - e^{i\theta}x$ is squared, while the norm of the noise vector ν is not. This makes the obtained bound somewhat weaker compared to the recovery guarantees for the PhaseLift reconstruction algorithm (for random frames with independent frame vectors), which ensure that, if $\|\nu\|_2 < \epsilon$, the estimate \tilde{x} obtained using PhaseLift satisfies $\min_{\theta \in [0, 2\pi)} \|\tilde{x} - e^{i\theta}x\|_2 \leq C\epsilon$ for an appropriately chosen numerical constant $C > 0$ [12]. At the same time, while the noise bound ϵ is an input parameter of the PhaseLift algorithm, the method described in this paper is independent of the actual noise size.

The remaining part of this paper is organized as follows. In section 2, we describe the idea of polarization and give some basic definitions and results from Gabor analysis and the theory of expander graphs that are used in the following. In section 3, we describe the reconstruction algorithm for time-frequency structured measurements and discuss the robustness of this algorithm in section 4. The analysis of robustness of the constructed algorithm leads to the investigation of geometric properties of Gabor frames, such as order statistics of frame coefficients. These properties are discussed in section 4.1. Numerical results of the algorithm's robustness are presented in section 5.

2. Notation and setup. Here and in the following, \odot denotes pointwise multiplication of two vectors of the same dimension. We view a vector $x \in \mathbb{C}^M$ as a function $x : \mathbb{Z}_M \rightarrow \mathbb{C}$, that is, all the operations on indices are done modulo M and $x(m - k) = x(M + m - k)$. We denote the complex unit sphere by $\mathbb{S}^{M-1} = \{x \in \mathbb{C}^M, \|x\|_2 = 1\}$.

The adjoint matrix of $A \in \mathbb{C}^{k \times m}$ is denoted by $A^* \in \mathbb{C}^{m \times k}$, and the smallest singular value of A is denoted by $\sigma_{\min}(A)$. Also, by a slight abuse of notation, we identify a frame $\Phi = \{\varphi_j\}_{j=1}^N \subset \mathbb{C}^M$ with its synthesis matrix, having the frame vectors φ_j as columns. For any $V \subset \{1, \dots, N\}$, we set $\Phi_V = \{\varphi_j\}_{j \in V}$.

We denote the Bernoulli distribution with success probability p by $B(1, p)$. Further, $\mathcal{N}(\mu, \sigma)$ denotes the Gaussian distribution with mean μ and variance σ^2 , and $\mathbb{C}\mathcal{N}(\mu, \sigma)$ denotes the complex valued Gaussian distribution.

2.1. Phase retrieval with polarization. The polarization approach to phase retrieval can be described as follows [18]. Suppose $\Phi_V = \{\varphi_j\}_{j \in V} \subset \mathbb{C}^M$ is a measurement frame. We

consider the phase retrieval problem

$$(3) \quad \begin{array}{ll} \text{find} & x \\ \text{subject to} & |\langle x, \varphi_j \rangle|^2 = b_j. \end{array}$$

For any $(i, j) \in V \times V$ with $|\langle x, \varphi_i \rangle| \neq 0$ and $|\langle x, \varphi_j \rangle| \neq 0$, we define the *relative phase* between frame coefficients as

$$(4) \quad \omega_{ij} = \left(\frac{\langle x, \varphi_i \rangle}{|\langle x, \varphi_i \rangle|} \right)^{-1} \frac{\langle x, \varphi_j \rangle}{|\langle x, \varphi_j \rangle|} = \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle}{|\langle x, \varphi_i \rangle| |\langle x, \varphi_j \rangle|}.$$

Note that $\omega_{ij}\omega_{jk} = \omega_{ik}$. Suppose that we are given $\{\omega_{ij}\}_{(i,j) \in E}$ for some set $E \subset V \times V$ in addition to the phaseless measurements with respect to Φ_V . Then we seek to solve the simpler problem

$$(5) \quad \begin{array}{ll} \text{find} & x \\ \text{subject to} & \frac{\overline{\langle x, \varphi_i \rangle} \langle x, \varphi_j \rangle}{|\langle x, \varphi_i \rangle| |\langle x, \varphi_j \rangle|} = \omega_{ij}, \\ & |\langle x, \varphi_i \rangle|^2 = b_i. \end{array}$$

This problem can be solved using *phase propagation*. More precisely, we choose $|\langle x, \varphi_{i_0} \rangle| \neq 0$, set $c_{i_0} = |\langle x, \varphi_{i_0} \rangle|$, and for every $j \in V$ with $(i_0, j) \in E$ define

$$c_j = \begin{cases} \omega_{i_0 j} |\langle x, \varphi_j \rangle| & \text{if } |\langle x, \varphi_j \rangle| \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

In the next step, for each k with c_k not defined yet and $(i_0, j), (j, k) \in E$ for some j with $b_j \neq 0$, we set

$$c_k = \begin{cases} \omega_{j k} \frac{c_j}{|c_j|} |\langle x, \varphi_k \rangle| & \text{if } |\langle x, \varphi_k \rangle| \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

We repeat this step iteratively until values c_i are assigned to all indices $i \in V$ that can be reached from i_0 using edges from E . This process is illustrated in Figure 2 (left).

Assume that we were able to compute c_i for all $i \in V$. Then, using a dual frame $\tilde{\Phi}_V = \{\tilde{\varphi}_i\}_{i \in V}$ and treating c_i 's as frame coefficients, we reconstruct a representative of the “up-to-a-global-phase” equivalence class $[x]$ as

$$\sum_{j \in V} c_j \tilde{\varphi}_j = \sum_{j \in V} \omega_{i_0 j} |\langle x, \varphi_j \rangle| \tilde{\varphi}_j = \left(\frac{\langle x, \varphi_{i_0} \rangle}{|\langle x, \varphi_{i_0} \rangle|} \right)^{-1} \sum_{j \in V} \langle x, \varphi_j \rangle \tilde{\varphi}_j = \left(\frac{\langle x, \varphi_{i_0} \rangle}{|\langle x, \varphi_{i_0} \rangle|} \right)^{-1} x \in [x].$$

Let us consider the graph $G = (V, E)$, later called the *graph of measurements*, with the set of vertices indexed by V and the set of edges E . From the phase propagation procedure, it is apparent that if $\langle x, \varphi_j \rangle = 0$ for some $j \in V$, then the corresponding relative phases ω_{ji} are not defined for all $i \in V$ and the phase cannot be propagated through vertex j . This has the

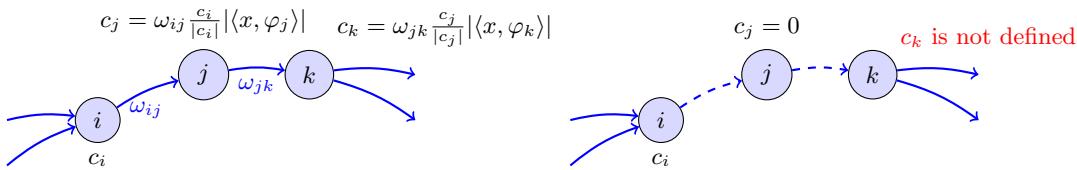


Figure 2. Phase propagation process described above is shown on the left. We iteratively compute phases of the measurements using relative phases (4) and phases computed on previous step. On the right, phase propagation through vertex j fails since the corresponding measurement is zero and relative phases ω_{ij} and ω_{jk} are not defined.

effect of deleting vertex j from G ; see Figure 2 (right). If G remains connected after deleting all “zero” vertices, then, for every vertex i , there exists a path from i_0 to i , and c_i can be computed. This solves problem (5).

Thus, the initial phase retrieval problem (3) is reduced to the problem of finding relative phases between pairs of frame coefficients from a set E , so that the corresponding graph of measurements $G = (V, E)$ satisfies strong connectivity properties. To obtain the relative phase between frame coefficients, the following polarization identity is useful.

Lemma 2.1 (see [18]). *Let $\omega = e^{2\pi i/3}$. If $\langle x, \varphi_i \rangle \neq 0$ and $\langle x, \varphi_j \rangle \neq 0$, then*

$$\omega_{ij} = \frac{1}{3|\langle x, \varphi_i \rangle||\langle x, \varphi_j \rangle|} \sum_{k=0}^2 \omega^k |\langle x, \varphi_i + \omega^k \varphi_j \rangle|^2.$$

In other words, to compute the relative phase ω_{ij} between the nonzero frame coefficients $\langle x, \varphi_i \rangle$ and $\langle x, \varphi_j \rangle$, we may use three additional phaseless measurements of x with respect to $\varphi_i + \varphi_j$, $\varphi_i + \omega \varphi_j$, and $\varphi_i + \omega^2 \varphi_j$. This means that the reconstruction of x using phase propagation involves only phaseless measurements, namely, phaseless measurements with respect to the union $\Phi_V \cup \Phi_E$, where Φ_V is a “vertex” frame and $\Phi_E = \{\varphi_i + \omega^k \varphi_j\}_{(i,j) \in E, k \in \{0,1,2\}}$. Note that $|\Phi_V \cup \Phi_E| = |V| + 3|E|$.

In [18], it is shown that in the noiseless case, one can perform phase retrieval with polarization using only $O(M)$ measurements. This algorithm is robust provided Φ_V consists of independent Gaussian vectors and the number of measurements is $O(M \log M)$. In [19], Bandeira, Chen, and Mixon adapt the polarization method to magnitude measurements of masked Fourier transforms of the signal. Using tools from additive combinatorics, the authors show that the graph of measurements they are using for reconstruction is sufficiently connected provided that the total number of measurements is $O(M \log M)$. However, no stability results are given for the case of structured measurements as considered in [19].

In section 3 we use the idea of polarization to build a recovery algorithm for time-frequency structured measurements and show reconstruction and stability guarantees for the designed algorithm.

2.2. Gabor frames for \mathbb{C}^M . Let us begin by defining two families of unitary operators on \mathbb{C}^M , namely, cyclic shift operators and modulation operators.

Definition 2.2.

1. Translation (or time shift) by $k \in \mathbb{Z}_M$ is given by

$$T_k x = T_k(x(0), x(1), \dots, x(M-1)) = (x(m-k))_{m \in \mathbb{Z}_M}.$$

That is, T_k simply permutes entries of x using k cyclic shifts.

2. Modulation (or frequency shift) by $\ell \in \mathbb{Z}_M$ is given by

$$M_\ell x = M_\ell(x(0), x(1), \dots, x(M-1)) = \left(e^{2\pi i \ell m / M} x(m) \right)_{m \in \mathbb{Z}_M}.$$

That is, M_ℓ multiplies $x = x(\cdot)$ pointwise with the harmonic $e^{2\pi i \ell(\cdot) / M}$.

3. The superposition $\pi(k, \ell) = M_\ell T_k$ of translation by k and modulation by ℓ is a time-frequency shift operator.
4. For $g \in \mathbb{C}^M \setminus \{0\}$ and $\Lambda \subset \mathbb{Z}_M \times \mathbb{Z}_M$, the set of vectors

$$(g, \Lambda) = \{\pi(k, \ell)g\}_{(k, \ell) \in \Lambda}$$

is called the Gabor system generated by the window g and the set Λ . A Gabor system which spans \mathbb{C}^M is a frame and is referred to as a Gabor frame.

The discrete Fourier transform $\mathcal{F} : \mathbb{C}^M \rightarrow \mathbb{C}^M$ plays a fundamental role in Gabor analysis. It is given pointwise by

$$\mathcal{F}x(\ell) = \sum_{m \in \mathbb{Z}_M} x(m) e^{-2\pi i m \ell / M}, \quad \ell \in \mathbb{Z}_M.$$

The short-time Fourier transform (or windowed Fourier transform) $V_g : \mathbb{C}^M \rightarrow \mathbb{C}^{M \times M}$ with respect to the window $g \in \mathbb{C}^M \setminus \{0\}$ is given by

$$(6) \quad (V_g x)(k, \ell) = \langle x, \pi(k, \ell)g \rangle = \mathcal{F}(x \odot T_k \bar{g})(\ell), \quad k, \ell \in \mathbb{Z}_M.$$

Equality (6) indicates that the short-time Fourier transform on \mathbb{C}^M can be efficiently computed using the *fast Fourier transform* (FFT), an efficient algorithm to compute the discrete Fourier transform of a vector. Phase retrieval with time-frequency structured measurements benefits from this, as it reduces the run time of recovery algorithms.

As we shall use polarization for phase retrieval, we would like to choose a window function g so that the frame (g, Λ) is a *full spark frame*, that is, so that for any subset $S \subset (g, \Lambda)$ of frame vectors with $|S| \geq M$, S spans \mathbb{C}^M [22]. Note that if the full Gabor system $(g, \mathbb{Z}_M \times \mathbb{Z}_M)$ is full spark, then so is (g, Λ) , for any $\Lambda \subset \mathbb{Z}_M \times \mathbb{Z}_M$.

The following result on the spark of Gabor frames with random window was shown for M prime in [23] and for M composite in [24].

Theorem 2.3. *Let M be a positive integer and let Λ be a subset of $\mathbb{Z}_M \times \mathbb{Z}_M$ with $|\Lambda| \geq M$. Then, for almost all windows ζ on the complex unit sphere $\mathbb{S}^{M-1} \subset \mathbb{C}^M$, (ζ, Λ) is a full spark frame.*

A more detailed description of Gabor frames in finite dimensions and their properties can be found in [25].

3. Phase retrieval from Gabor measurements. We now describe our design of a measurement frame and a reconstruction process in the noiseless case, also addressed in [1], and then discuss the robustness of our algorithm in the case when measurements are corrupted by noise. The analysis of the algorithm's robustness leads to the investigation of geometric properties of the measurement frame, such as frame bounds and flatness of the vector of frame coefficients. These properties are not only important for the problem at hand, but are of general interest in Gabor analysis.

3.1. Measurement process and frame construction. Consider the phase retrieval problem (3) with measurement frame

$$\Phi = \{\varphi_j\}_{j=1}^N = \Phi_V \cup \Phi_E \subset \mathbb{C}^M,$$

where Φ_V is a Gabor frame and Φ_E is a set of vectors corresponding to the additional edge measurements. As described in section 2.1, phaseless measurements with respect to Φ_E are used to compute relative phases between frame coefficients.

Let us specify Φ_V and Φ_E now. For Φ_V we choose the Gabor frame

$$(7) \quad \begin{aligned} \Phi_V &= (g, \Lambda) \text{ with } \Lambda = F \times \mathbb{Z}_M, \quad F \subset \mathbb{Z}_M, \quad |F| = K, \\ &\text{and } g \in \mathbb{C}^M \text{ uniformly distributed on the unit sphere } \mathbb{S}^{M-1} \subset \mathbb{C}^M. \end{aligned}$$

The integer K is fixed and does not depend on the ambient dimension M , and F is an arbitrary subset of \mathbb{Z}_M of cardinality K . That is, we consider all frequency shifts and only a constant number of time shifts. As (6) indicates, our measurements are magnitudes of masked Fourier transform coefficients with the masks being $T_k \bar{g}$, $k \in F$.

We choose

$$(8) \quad \begin{aligned} \Phi_E &= \{\pi(\lambda_1)g + \omega^t \pi(\lambda_2)g\}_{(\lambda_1, \lambda_2) \in E, t \in \{0, 1, 2\}} \text{ with } \omega = e^{2\pi i/3}, \\ E &= \{((k_1, \ell_1), (k_2, \ell_2)), \text{ s.t. } k_1, k_2 \in F, \ell_2 - \ell_1 \in C\} \subset \Lambda \times \Lambda, \\ &\text{and } C = D \cup (-D) \setminus \{0\} \subset \mathbb{Z}_M \text{ with } \mathbf{1}_D(m) \sim \text{i.i.d. } B\left(1, \frac{d \log M}{M}\right) \end{aligned}$$

(where i.i.d. is independent and identically distributed). In other words, $D \subset \mathbb{Z}_M$ is constructed at random, so that every $m \in \mathbb{Z}_M$ is chosen to be an element of D independently with probability $\frac{d \log M}{M}$ for a parameter $d > 0$ we will specify later, and $C = D \cup (-D) \setminus \{0\}$. Then

$$\pi(\lambda_1)g + \omega^t \pi(\lambda_2)g = \pi(k_1, \ell_1)g + \omega^t \pi(k_2, \ell_2)g = p_{(\ell_2 - \ell_1)k_1 k_2}(t) \odot \pi(\lambda_1)g,$$

where the vector $p_{ck_1 k_2}(t) \in \mathbb{C}^M$ is defined pointwise by

$$p_{c, k_1, k_2}(t)(m) = 1 + e^{2\pi i(\frac{cm}{M} + \frac{t}{3})} \frac{g(m - k_2)}{g(m - k_1)}, \quad m \in \mathbb{Z}_M,$$

with parameters $c \in C$, $k_1, k_2 \in F$, and $t \in \{0, 1, 2\}$. Therefore, for each fixed set of four parameters (c, k_1, k_2, t) , the respective additional measurements are magnitudes of masked

Fourier transform coefficients as well, namely,

$$(9) \quad |\langle x, p_{ck_1k_2}(t) \odot \pi(k_1, \ell)g \rangle| = |\mathcal{F}(x \odot \bar{p}_{ck_1k_2}(t) \odot T_{k_1}\bar{g})(\ell)|, \quad \ell \in \mathbb{Z}_M.$$

Let us note that the frame $\Phi = \Phi_V \cup \Phi_E$ constructed in this way consists of $|\Lambda| + 3|E| = |F|M + 3|F|^2|C|M$ vectors. Since $C = D \cup (-D) \setminus \{0\}$ with $\mathbf{1}_D(m) \sim \text{i.i.d. } B(1, \frac{d \log M}{M})$, we have $|C| = O(\log M)$ with high probability and thus $|\Phi| = O(M \log M)$.

Using the polarization identity in Lemma 2.1 with $\omega = e^{2\pi i/3}$, we compute the relative phases

$$(10) \quad \omega_{\lambda_1 \lambda_2} = \left(\frac{\langle x, \pi(\lambda_1)g \rangle}{|\langle x, \pi(\lambda_1)g \rangle|} \right)^{-1} \frac{\langle x, \pi(\lambda_2)g \rangle}{|\langle x, \pi(\lambda_2)g \rangle|} = \frac{\sum_{t=0}^2 \omega^t |\langle x, \pi(\lambda_1)g + \omega^t \pi(\lambda_2)g \rangle|^2}{3|\langle x, \pi(\lambda_1)g \rangle| |\langle x, \pi(\lambda_2)g \rangle|}$$

for $(\lambda_1, \lambda_2) \in E$, where E is defined by (8). Recall, that $\omega_{\lambda_1 \lambda_2}$ is well defined if and only if $|\langle x, \pi(\lambda_1)g \rangle| \neq 0$ and $|\langle x, \pi(\lambda_2)g \rangle| \neq 0$.

Remark 3.1. As (6) and (9) indicate, all required measurements are magnitudes of masked Fourier transform coefficients. These are relevant for many applications. Moreover, measurements and reconstruction in this case can be implemented using FFT, which allows a noticeable speed up of measurement and reconstruction processes. For comparison, the computational complexity of the measurement process with random Gaussian frame of cardinality $O(M \log M)$ (as considered, for example, in [12] and [18]) is $O(M^2 \log M)$, and the complexity of measurement with the frame Φ constructed above is $O(M \log^2 M)$. Furthermore, in the case of random Gaussian frames, we have to use $O(M^2 \log M)$ memory bits to store the measurement matrix, while for our frame Φ it is enough to store the window g and the set C , and the overall amount of memory used is only $M + O(\log M) = O(M)$. These are some of the advantages of time-frequency structured frames in comparison to randomly generated frames.

3.2. Reconstruction in the noiseless case. We now describe our polarization based reconstruction algorithm for time-frequency structured measurements.

Let us consider graph $G = (\Lambda, E)$, where Λ and E are defined by (7) and (8), respectively. Since $0 \notin C$, the graph G has no loops, and since $C = -C$, it is not directed. Also, each vertex $\lambda = (k, \ell)$ of G is adjacent to any vertex $\lambda' = (k', \ell + c)$ with $c \in C$ and $k' \in F$. Thus, each vertex in G has degree $|F||C|$ and G is regular.

Let $A = A(G)$ be the *adjacency matrix* of G , that is, a $|\Lambda| \times |\Lambda|$ matrix whose (λ_1, λ_2) entry is equal to the number of edges in G connecting vertices λ_1 and λ_2 . Being real and symmetric, $A(G)$ has $|\Lambda|$ real eigenvalues $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_{|\Lambda|}$. We refer to the sequence $\{\alpha_i\}_{i=1}^n$ of eigenvalues of $A(G)$ as the *spectrum* of the graph G . Note that for a d -regular graph G , $\alpha_1 = d$. The spectrum encodes information about the connectivity of the graph. For example, G is connected if and only if $\alpha_1 > \alpha_2$. The value $\text{spg}(G) = \frac{\alpha_1 - \alpha}{\alpha_1}$, where $\alpha = \max\{|\alpha_2|, |\alpha_n|\}$, is known as the *spectral gap* of G . The following version of [26, Lemma 5.2] relates connectivity properties of the graph G and its spectral gap; see also [18].

Lemma 3.2. *Let G be a d -regular graph. For all $\varepsilon \leq \frac{\text{spg}(G)}{6}$, the graph obtained by removing any εn vertices from G has a connected component of size at least $(1 - \frac{2\varepsilon}{\text{spg}(G)})n$.*

To be able to reconstruct a signal x using Algorithm 1, we need to ensure that $|\Lambda''|$ is sufficiently large, so that $\Phi_{\Lambda''} = (g, \Lambda'')$ is a frame (see Figure 3 for an example). Then x

Algorithm 1: Reconstruction in the noiseless case.

Input : phaseless measurements b with respect to $\Phi_\Lambda \cup \Phi_E$, defined by (7) and (8);
 $F \subset \mathbb{Z}_M$, $C \subset \mathbb{Z}_M$, and window $g \in \mathbb{C}^M \setminus \{0\}$

Output: $\tilde{x} \in [x]$, i.e., the signal x up to a global phase.

- 1 construct the graph $G = (\Lambda, E)$ with $\Lambda = F \times \mathbb{Z}_M$ and E as in (8);
- 2 assign to each $\lambda \in \Lambda$ the weight $b_\lambda = |\langle x, \pi(\lambda)g \rangle|^2$;
- 3 assign to each edge $(\lambda_1, \lambda_2) \in E$ the weight $\omega_{\lambda_1 \lambda_2}$ computed using (10);
- 4 delete from G all vertices λ with $b_\lambda = 0$ to obtain $G' = (\Lambda', E') \subset G$;
- 5 choose a connected component $G'' = (\Lambda'', E'') \subset G'$ of the biggest size;
- 6 run the *phase propagation* process (section 2.1) to obtain c_λ , $\lambda \in \Lambda''$;
- 7 reconstruct $\tilde{x} = (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} c$ from $c = \{c_\lambda\}_{\lambda \in \Lambda''}$.

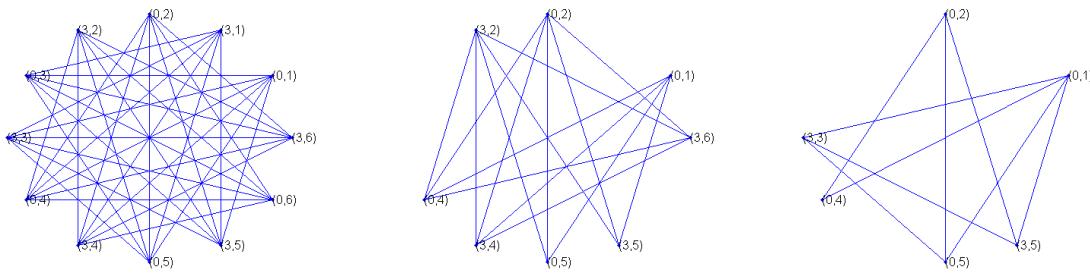


Figure 3. An example of the graph of measurements G with $M = 6$, $F = \{0, 3\}$, and $C = \{2, 3, 4\}$ (left). This graph remains connected after deleting one third of its vertices (middle). After we delete one half of its vertices it has a connected component of size at least 4 (right).

can be recovered from its frame coefficients with respect to $\Phi_{\Lambda''}$. In terms of the graph of measurements G , this means that after we delete all vertices λ with zero weight, the resulting graph has a connected component of sufficiently big size. As Lemma 3.2 shows, this is satisfied provided that G has a sufficiently big spectral gap. The spectral gap of $G = (\Lambda, E)$ can be estimated in terms of the set C . More precisely, for the random set C constructed in (8), the following result is shown in [19].

Lemma 3.3. *Pick $d > 36$ and suppose the entries of the characteristic vector $\mathbf{1}_D$ of a set D are independent with distribution $B(1, \frac{d \log M}{M})$. Take $C = D \cup (-D) \setminus \{0\}$ and construct the graph G as above. Then with overwhelming probability*

$$\text{spg}(G) \geq 1 - \frac{6}{\sqrt{d}}.$$

Now we are ready to establish recovery guarantees for Algorithm 1.

Theorem 3.4. *Let frames Φ_V and Φ_E be constructed as above with $|F| = 12$ and $d = 144$. Then every signal $x \in \mathbb{C}^M$ can be reconstructed from $M + 3|F|^2 M |C| = O(M \log M)$ phaseless measurements with respect to the frame $\Phi_V \cup \Phi_E$ using Algorithm 1.*

Proof. We begin the reconstruction algorithm by assigning to each vertex $\lambda \in \Lambda$ of the constructed graph G the weight $b_\lambda = |\langle x, \pi(\lambda)g \rangle|$ and assigning to each edge $(\lambda_1, \lambda_2) \in E$ of G

the relative phase $\omega_{\lambda_1 \lambda_2}$ which is computed from the additional edge measurements. Theorem 2.3 implies that the frame $\Phi_V = \{\pi(\lambda)g\}_{\lambda \in \Lambda}$ is full spark with probability 1. Thus, for any vector $x \in \mathbb{C}^M$, the number of zero measurements among $\{b_\lambda = |\langle x, \pi(\lambda)g \rangle|\}_{\lambda \in \Lambda}$ is at most $M - 1$. In other words, Algorithm 1 deletes at most $M - 1$ vertices from G to obtain G' .

Next, Φ_V being full spark implies that any its subset $\Phi_{V'} \subset \Phi_V$ of size $|\Phi_{V'}| \geq M$ form a frame. Thus, to recover x , it is enough to know any M of the frame coefficients with respect to $\Phi_{V'}$. To show that G' has a connected component G'' of size at least M , first note that Lemma 3.3 ensures that $\text{spg}(G) \geq 1 - \frac{6}{\sqrt{d}} = \frac{1}{2}$. Then, applying Lemma 3.2 with $n = |\Lambda| = |F|M$ and $\varepsilon = \frac{1}{|F|} = \frac{1}{12} \leq \frac{\text{spg } G}{6}$, we obtain that after deleting any $\varepsilon n = M$ vertices from G , the largest connected component G'' will have at least $(1 - \frac{2\varepsilon}{\text{spg } G})n \geq \frac{2}{3}|F|M = 8M > M$ vertices.

By running the phase propagation algorithm on the connected graph $G'' = (\Lambda'', E'')$, we recover c_λ , $\lambda \in \Lambda''$, which are, up to a global phase $e^{i\theta}$, $\theta \in [0, 2\pi)$, equal to the corresponding frame coefficients of x with respect to $\Phi_{\Lambda''}$, that is,

$$c_\lambda = e^{i\theta} \langle x, \pi(\lambda)g \rangle, \quad \lambda \in \Lambda''.$$

Now, using the canonical dual frame, we obtain

$$\tilde{x} = (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} c = e^{i\theta} (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} \Phi_{\Lambda''}^* x = e^{i\theta} x. \quad \blacksquare$$

4. Robustness of reconstruction in the presence of noise. In many applications, measurements are corrupted by noise. In this section we address the behavior of the presented algorithm in the case when the available measurements are of the form

$$(11) \quad \begin{aligned} b_\lambda &= |\langle x, \pi(\lambda)g \rangle|^2 + \nu_\lambda, \quad \lambda \in \Lambda, \\ b_{\lambda_1 \lambda_2 t} &= |\langle x, \pi(\lambda_1)g + \omega^t \pi(\lambda_2)g \rangle|^2 + \nu_{\lambda_1 \lambda_2 t}, \quad (\lambda_1, \lambda_2) \in E, \quad t \in \{0, 1, 2\}, \end{aligned}$$

where ν_λ , $\nu_{\lambda_1 \lambda_2 t}$ are noise terms. We aim to construct a modification of Algorithm 1 which, in the presence of noise, recovers a close estimate \tilde{x} of the original signal x .

4.1. Order statistics of frame coefficients. To compute a relative phase between two frame coefficients, we rely on formula (10). The calculations include division by $|\langle x, \pi(\lambda_1)g \rangle|$ and $|\langle x, \pi(\lambda_2)g \rangle|$ and are therefore very sensitive to perturbations when $|\langle x, \pi(\lambda_1)g \rangle|$ or $|\langle x, \pi(\lambda_2)g \rangle|$ is small. While vertices with zero corresponding measurements provide no relative phase information, vertices with small vertex measurements lead to unreliable relative phase estimations and should therefore be deleted from the graph. As we require the graph of measurements to have a connected component of size at least M after deleting vertices with small weights, the number of such vertices has to be estimated. To do so, we show that the vector of frame coefficients of a fixed $x \in \mathbb{S}^{M-1}$ with respect to a Gabor frame with random window is “flat” with high probability. That is, most of the frame coefficients are in the range of $\frac{c}{\sqrt{M}}$ to $\frac{K}{\sqrt{M}}$, for some suitably chosen constants $K > c > 0$.

Theorem 4.1. *Fix $x \in \mathbb{S}^{M-1} \subset \mathbb{C}^M$ and consider a Gabor frame (g, Λ) with $\Lambda \subset \mathbb{Z}_M \times \mathbb{Z}_M$ and a random window g uniformly distributed on the unit sphere \mathbb{S}^{M-1} . Then the following hold.*

(a) For any $c > 0$ and $k > 0$ with probability at least $1 - \frac{1}{k^2}$, we have

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| < \frac{c}{\sqrt{M}} \right\} \right| < |\Lambda|(c^2 + kc).$$

(b) For any $K > 0$ and $k > 0$ with probability at least $1 - \frac{1}{k^2}$, we have

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| > \frac{K}{\sqrt{M}} \right\} \right| < |\Lambda| \left(\frac{8}{\pi} e^{-K^2} + k \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} \right).$$

The proof of this result is presented in the appendix.

Remark 4.2. A similar result can be shown for a Gabor frame with a window whose entries are independent Gaussian random variables. The proof in this case involves the same steps as the proof of Theorem 4.1.

Theorem 4.1 is a nonuniform result in the sense that the proven bounds hold with high probability for each *individual* x . Note that this does not imply that the same bounds will hold *simultaneously for all* $x \in \mathbb{S}^{M-1}$ with high probability. We give a uniform bound for the number of large frame coefficients in the following result; see also [2].

Theorem 4.3. Consider a Gabor frame $(g, \mathbb{Z}_M \times \mathbb{Z}_M)$ with window g whose entries are independent Gaussian random variables with zero mean and variance $\frac{1}{\sqrt{M}}$. Then, for some suitably chosen numerical constants $c, c_1 > 0$,

$$\left| \left\{ \lambda \in \mathbb{Z}_M \times \mathbb{Z}_M, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| > \sqrt{\frac{3}{2c} \frac{\log^2 M}{\sqrt{M}}} \right\} \right| < \frac{cM}{\log^4 M}$$

for all $x \in \mathbb{S}^{M-1}$ with probability at least $1 - e^{-c_1 \log^3 M}$.

Proof. Let g be a random vector with $g(m) \sim \text{i.i.d. } \mathcal{CN}\left(0, \frac{1}{\sqrt{M}}\right)$, $m \in \mathbb{Z}_M$, and let $\Phi = \Phi_{\mathbb{Z}_M \times \mathbb{Z}_M}$ be an $M \times M^2$ matrix whose columns are $\pi(\lambda)g$, $\lambda \in \mathbb{Z}_M \times \mathbb{Z}_M$. Fix $s \in \{1, \dots, M^2\}$ and for any $x \in \mathbb{S}^{M-1}$ denote by S_x the set of $\lambda \in \mathbb{Z}_M \times \mathbb{Z}_M$ corresponding to the s biggest in modulus frame coefficients of x with respect to the Gabor frame $(g, \mathbb{Z}_M \times \mathbb{Z}_M)$. Then, for the phase vector $v_x \in \mathbb{C}^{M^2}$ defined by

$$v_x(\lambda) = \begin{cases} \frac{\langle x, \pi(\lambda)g \rangle}{|\langle x, \pi(\lambda)g \rangle|}, & \lambda \in S_x, \\ 0, & \text{otherwise,} \end{cases}$$

we have

$$x^* \Phi v_x = \sum_{\lambda \in S_x} |\langle x, \pi(\lambda)g \rangle|.$$

Applying the Cauchy–Schwarz inequality to $x^* \Phi v_x = \langle \Phi v_x, x \rangle$, we obtain

$$\sum_{\lambda \in S_x} |\langle x, \pi(\lambda)g \rangle| \leq \|x\|_2 \|\Phi v_x\|_2 = \|\Phi v_x\|_2.$$

Note that v_x is an s -sparse vector with $\|v_x\|_2 = \sqrt{s}$. Then, if $s = \frac{cM}{\log^4 M}$ for a suitably chosen numerical constant $c > 0$, we have

$$\frac{1}{2}\|v\|_2^2 \leq \|\Phi v\|_2^2 \leq \frac{3}{2}\|v\|_2^2$$

for any s -sparse vector $v \in \mathbb{C}^{M^2}$ with probability at least $1 - e^{-c_1 \log^3 M}$, where $c_1 > 0$ depends only on c ; see [27, Theorem 5.1]. Thus

$$\sum_{\lambda \in S_x} |\langle x, \pi(\lambda)g \rangle| \leq \|\Phi v_x\|_2 \leq \sqrt{\frac{3s}{2}}$$

with probability at least $1 - e^{-c_1 \log^3 M}$. It follows that with the same probability,

$$\min_{\lambda \in S_x} |\langle x, \pi(\lambda)g \rangle| \leq \sqrt{\frac{3}{2s}} = \sqrt{\frac{3}{2c} \frac{\log^2 M}{\sqrt{M}}}.$$

In other words, with probability at least $1 - e^{-c_1 \log^3 M}$ for any $x \in \mathbb{S}^{M-1}$, all except at most $\frac{cM}{\log^4 M} - 1$ frame coefficients are in modulus bigger than $\sqrt{\frac{3}{2c} \frac{\log^2 M}{\sqrt{M}}}$. \blacksquare

Since Theorem 4.3 holds for the full Gabor frame $(g, \mathbb{Z}_M \times \mathbb{Z}_M)$, it also holds for any of its subframes (g, Λ) , $\Lambda \subset \mathbb{Z}_M$. Also, note that while Theorem 4.3 gives a better bound on the number of large frame coefficients than Theorem 4.1(b), it gives a slightly weaker bound on the modulus of the remaining coefficients, namely, $\frac{C \log^2 M}{\sqrt{M}}$ instead of $\frac{C}{\sqrt{M}}$.

Remark 4.4. Note that [27, Theorem 5.1] is formulated for a window g with independent mean-zero, variance one, L -sub-Gaussian entries. Thus, Theorem 4.3 is also true in this more general case.

4.2. Reconstruction from noisy measurements. To obtain a modification to Algorithm 1 that leads to robust reconstruction, we may assume, without loss of generality, that the signal x lies on the complex unit sphere $\mathbb{S}^{M-1} \subset \mathbb{C}^M$. As mentioned before, instead of deleting vertices with zero weight, a portion of vertices with small weights should be deleted in the first step of the reconstruction algorithm. As shown later, very large measurements can also prevent stable reconstruction, so we delete respective vertices as well. To delete vertices from the graph of measurements, we use Algorithm 2.

Algorithm 2: Deleting “small” and “large” vertices.

Input : graph $G = (\Lambda, E)$ with weighted vertices V , parameters α, β

Output: graph G' with more “flat” vertex weights

```

1 for  $i = 0$  to  $(1 - \alpha)|\Lambda|$  do
2   | find  $\lambda \in \Lambda$  with the smallest value of  $b_\lambda$  and delete it from  $G$ ;
3 end
4 for  $j = 0$  to  $(1 - \beta)|\Lambda|$  do
5   | find  $\lambda \in \Lambda$  with the largest value of  $b_\lambda$  and delete it from  $G$ .
6 end

```

Let $G'' = (\Lambda'', E'')$ be a subgraph of G' and let A be the weighted adjacency matrix of the graph G'' given by

$$(12) \quad A(\lambda_1, \lambda_2) = \begin{cases} \frac{\overline{\langle x, \pi(\lambda_1)g \rangle} \langle x, \pi(\lambda_2)g \rangle + \varepsilon_{\lambda_1 \lambda_2}}{\left| \overline{\langle x, \pi(\lambda_1)g \rangle} \langle x, \pi(\lambda_2)g \rangle + \varepsilon_{\lambda_1 \lambda_2} \right|}, & (\lambda_1, \lambda_2) \in E'' \\ 0, & (\lambda_1, \lambda_2) \notin E'', \end{cases}$$

where $\lambda_1, \lambda_2 \in \Lambda''$ and $\varepsilon_{\lambda_1 \lambda_2} = \frac{1}{3} \sum_{t=0}^2 \omega^t \nu_{\lambda_1 \lambda_2 t}$. Then $A(\lambda_i, \lambda_j)$ can be considered as an approximation of the relative phase between frame coefficients corresponding to Λ'' .

Noise might accumulate while passing from one vertex to another in the phase propagation process. Thus, we seek an efficient method to reconstruct the phases of the vertex frame coefficients using measured relative phases (12). For this purpose, we shall use the *angular synchronization algorithm* [28, 18]. The following result shows robustness of the angular synchronization algorithm in the presence of noise [18, 29].

Theorem 4.5. *Consider a graph $G = (\Lambda'', E'')$ with spectral gap $\tau > 0$, and define $\|\theta\|_{\mathbb{T}} = \min_{k \in \mathbb{Z}} |\theta - 2\pi k|$ for all angles $\theta \in \mathbb{R}/2\pi\mathbb{Z}$. Then, given the weighted adjacency matrix A as in (12), the angular synchronization algorithm outputs $\tilde{v} \in \mathbb{C}^{|V|}$ with unit-modulus entries, such that for some phase $\theta \in \mathbb{R}/2\pi\mathbb{Z}$,*

$$\sum_{\lambda \in \Lambda''} \|\arg(\tilde{v}_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 \leq \frac{C\|\varepsilon\|^2}{\tau^2 P^2},$$

where $P = \min_{(\lambda_1, \lambda_2) \in E''} |\overline{\langle x, \pi(\lambda_1)g \rangle} \langle x, \pi(\lambda_2)g \rangle + \varepsilon_{\lambda_1 \lambda_2}|$ and C is a universal constant.

As Theorem 4.5 shows, the accuracy of the angular synchronization algorithm depends on the spectral gap of the graph G'' . To find a subgraph $G'' \subset G'$ with spectral gap bounded away from zero, we shall use the *spectral clustering* algorithm [30, 18]. To ensure that $|\Lambda''| \geq M$, we rely on the following result. Its proof is based on the Cheeger inequality for the graph connection Laplacian [29] and can be found in [18].

Theorem 4.6. *Take $p \geq q \geq \frac{2}{3}$. Consider a regular graph $G = (V, E)$ with spectral gap $\lambda_2 > g(p, q) = 1 - 2(q(1 - q) - (1 - p))$ and set $\tau = \frac{1}{8}(\lambda_2 - g(p, q))^2$. Then, after Algorithm 2 removes at most $(1 - p)|V|$ vertices from G , the spectral clustering algorithm outputs a subgraph with at least $q|V|$ vertices and spectral gap at least τ .*

We summarize the above discussion in the following reconstruction algorithm.

Now we are ready to prove our main result.

Theorem 4.7. *Fix $x \in \mathbb{C}^M$ and consider the measurement procedure (11) described above, with $|F|$ and d sufficiently large. If the noise vector satisfies $\frac{\|\nu\|_2}{\|x\|_2^2} \leq \frac{C_1}{M}$ for some C_1 small enough, then there exists a constant C'' such that the estimate \tilde{x} produced by Algorithm 3 from the noisy measurements $\{b_j\}_{j=1}^N$ satisfies with overwhelming probability*

$$\min_{\theta \in [0, 2\pi)} \|\tilde{x} - e^{i\theta} x\|_2^2 \leq \frac{C'' \sqrt{M} \|\nu\|_2}{\Delta},$$

where $\Delta = \min_{\Lambda'' \subset \Lambda, |\Lambda''| \geq 2/3|\Lambda|} \sigma_{\min}^2(\Phi_{\Lambda''}^*)$.

Algorithm 3: Reconstruction in the noisy case.

Input : $F \subset \mathbb{Z}_M$, $C \subset \mathbb{Z}_M$, window g ;
noisy measurements b w.r.t. $\Phi_\Lambda \cup \Phi_E$, given by (11);
parameters α, β, τ

Output: approximation \tilde{x} of the signal x (up to a global phase)

- 1 construct graph $G = (\Lambda, E)$ with $\Lambda = F \times \mathbb{Z}_M$ and E as in (8);
- 2 assign to each $\lambda \in \Lambda$ weight b_λ ;
- 3 assign to each edge $(\lambda_1, \lambda_2) \in E$ weight $A_{\lambda_1 \lambda_2}$, given in (12);
- 4 run Algorithm 2 with parameters α, β to obtain $G' = (\Lambda', E')$;
- 5 run *spectral clustering* to find $G'' = (\Lambda'', E'')$ with $\text{spg}(G'') \geq \tau$;
- 6 run *angular synchronization* to obtain approximate phases $\{u_\lambda\}_{\lambda \in \Lambda''}$;
- 7 set $c_\lambda = u_\lambda \sqrt{b_\lambda}$, $\lambda \in \Lambda''$;
- 8 reconstruct $\tilde{x} = (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} c$ from $c = \{c_\lambda\}_{\lambda \in \Lambda''}$.

Proof. Without loss of generality we can assume that $\|x\|_2 = 1$. As follows from Lemma 3.3, with overwhelming probability $\text{spg}(G) \geq 1 - \frac{6}{\sqrt{b}}$. Let us fix parameters $\tau_0 > 0$, $\alpha, \beta \in (0, 1)$ and apply Theorem 4.6 with $g(p, q) = 1 - 2(q(1 - q) - (1 - p)) = 1 - \frac{6}{\sqrt{b}} - \tau_0 < \text{spg}(G)$. Then, after Algorithm 2 deletes $(1 - p)|\Lambda| = (1 - \alpha - \beta)|\Lambda|$ vertices with the smallest and the largest corresponding measurements, we apply the spectral clustering algorithm with parameter $\tau = \frac{1}{8}(\text{spg}(G) - g(p, q))^2 \geq \frac{\tau_0^2}{8}$. We obtain a graph $G'' = (V'', E'')$ with $|V''| \geq q|\Lambda|$ and $\text{spg}(G'') \geq \frac{\tau_0^2}{8}$.

Let us specify q now. Since we set $1 - 2(q(1 - q) - (1 - p)) = 1 - \frac{6}{\sqrt{b}} - \tau_0$, it follows $q(1 - q) = \frac{3}{\sqrt{b}} + (1 - \alpha - \beta) + \frac{\tau_0}{2}$. If τ_0, α, β , and b are chosen appropriately, we can make $\frac{3}{\sqrt{b}} + (1 - \alpha - \beta) + \frac{\tau_0}{2} = A \leq \frac{2}{9}$. Then $1 \geq q = \frac{1 + \sqrt{1 - 4A}}{2} \geq \frac{1 + \sqrt{1/9}}{2} = \frac{2}{3}$. This ensures that $q \in (\frac{2}{3}, 1)$.

Now, after applying the angular synchronization algorithm and using Theorem 4.5, we obtain a universal constant $C > 0$ and phase $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, such that

$$\sum_{\lambda \in V''} \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 \leq \frac{C\|\varepsilon\|_2^2}{\tau^2 P^2},$$

where $\varepsilon_{\lambda_1 \lambda_2} = \frac{1}{3} \sum_{t=0}^2 \omega^t \nu_{\lambda_1 \lambda_2 t}$. Since $\|\nu\|_2 \leq \frac{C_1}{M}$, we also have $|\nu_{\lambda_1 \lambda_2 t}| \leq \frac{C_1}{M}$ for all $\lambda_1, \lambda_2 \in \Lambda''$ and $t \in \{0, 1, 2\}$ and, thus, $|\varepsilon_{\lambda_1 \lambda_2}| \leq \frac{C_1}{M}$. By the Cauchy–Schwarz inequality we have

$$\|\varepsilon\|_2^2 = \sum_{(\lambda_1, \lambda_2) \in E''} |\varepsilon_{\lambda_1 \lambda_2}|^2 \leq \frac{1}{9} \sum_{(\lambda_1, \lambda_2) \in E''} \left(\sum_{i=0}^2 |\omega^t|^2 \right) \left(\sum_{i=0}^2 |\nu_{\lambda_1 \lambda_2 t}|^2 \right) \leq \frac{1}{3} \|\nu\|_2^2.$$

Theorem 4.1 implies that, for any $k, c > 0$ and $\epsilon = c^2$,

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| < \frac{c}{\sqrt{M}} \right\} \right| < |\Lambda|(\epsilon + k\sqrt{2\epsilon})$$

with probability at least $1 - 1/k^2$. Then, since $|\nu_\lambda| \leq \frac{C_1}{M}$, on this event

$$\left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle|^2 + \nu_\lambda < \frac{c^2}{M} - \frac{C_1}{M} \right\} \subset \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle|^2 < \frac{c^2}{M} \right\},$$

that is,

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle|^2 + \nu_\lambda < \frac{c^2 - C_1}{M} \right\} \right| < |\Lambda|(\epsilon + k\sqrt{2\epsilon}).$$

We set $\alpha = 1 - (\epsilon + k\sqrt{2\epsilon})$ and delete $|\Lambda|(\epsilon + k\sqrt{2\epsilon})$ vertices with the smallest corresponding measurements. For the remaining coefficients we have $|\langle x, \pi(\lambda)g \rangle| \geq \frac{\tilde{c}}{\sqrt{M}}$ for some constant $\tilde{c} > 0$, provided c^2 is sufficiently larger than C_1 . Similarly, setting $\beta = 1 - (\eta + k\sqrt{2\eta})$, we delete $|\Lambda|(\eta + k\sqrt{2\eta})$ vertices with the largest corresponding measurements, and for the remaining vertices, with probability at least $1 - \frac{1}{k^2}$, we have $|\langle x, \pi(\lambda)g \rangle| \leq \frac{\tilde{K}}{\sqrt{M}}$ for some constant $\tilde{K} > 0$.

Now, with probability at least $1 - \frac{2}{k^2}$, after applying Algorithms 1 and 2, we have $\frac{\tilde{c}}{\sqrt{M}} \leq |\langle x, \pi(\lambda)g \rangle| \leq \frac{\tilde{K}}{\sqrt{M}}$ for all $\lambda \in V''$. Thus

$$\begin{aligned} P &= \min_{(\lambda_1, \lambda_2) \in E''} |\overline{\langle x, \pi(\lambda_1)g \rangle} \langle x, \pi(\lambda_2)g \rangle + \varepsilon_{\lambda_1 \lambda_2}| \\ &\geq \left| \frac{\tilde{c}^2}{M} - \max_{(\lambda_1, \lambda_2) \in E''} |\varepsilon_{\lambda_1 \lambda_2}| \right| \geq \left| \frac{\tilde{c}^2}{M} - \frac{C_1}{M} \right| \geq \frac{\tilde{C}}{M}. \end{aligned}$$

Summing up, we obtain

$$\sum_{\lambda \in V''} \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 \leq \frac{64C\|\nu\|_2^2 M^2}{3\tau_0^4 \tilde{C}^2}.$$

For every $\lambda \in V''$, we denote the obtained estimate of the corresponding frame coefficient by $c_\lambda = u_\lambda \sqrt{b_\lambda} = u_\lambda \sqrt{|\langle x, \pi(\lambda)g \rangle|^2 + \nu_\lambda}$. Also, set $\delta_\lambda = c_\lambda - e^{i\theta} \langle x, \pi(\lambda)g \rangle$ and $\zeta_\lambda = \sqrt{b_\lambda} - |\langle x, \pi(\lambda)g \rangle|$. Then

$$\begin{aligned} |\delta_\lambda| &= \left| \sqrt{b_\lambda} e^{i\arg(u_\lambda)} - \sqrt{b_\lambda} e^{i(\theta + \arg(\langle x, \pi(\lambda)g \rangle))} + \zeta_\lambda e^{i(\theta + \arg(\langle x, \pi(\lambda)g \rangle))} \right| \\ &\leq \sqrt{b_\lambda} \left| e^{i(\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta)} - 1 \right| + |\zeta_\lambda| \\ &\leq \sqrt{b_\lambda} \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}} + |\zeta_\lambda|. \end{aligned}$$

Further, since $|\delta_\lambda|^2 \leq 2b_\lambda \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 + 2|\zeta_\lambda|^2$, it follows

$$\|\delta\|_2^2 \leq 2 \sum_{\lambda \in V''} b_\lambda \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 + 2 \sum_{\lambda \in V''} |\zeta_\lambda|^2.$$

Using the fact that, for any $a, b \in \mathbb{R}$, $(a - b)^2 \leq |a^2 - b^2|$, we obtain $\zeta_\lambda^2 = (\sqrt{b_\lambda} - |\langle x, \pi(\lambda)g \rangle|)^2 \leq |\nu_\lambda|$. And, since $b_\lambda \leq \frac{\tilde{K}^2}{M}$ and $\|\cdot\|_1 \leq \sqrt{|\Lambda''|} \|\cdot\|_2$,

$$\begin{aligned} \|\delta\|_2^2 &\leq \frac{2\tilde{K}^2}{M} \sum_{\lambda \in V''} \|\arg(u_\lambda) - \arg(\langle x, \pi(\lambda)g \rangle) - \theta\|_{\mathbb{T}}^2 + 2\|\nu_V\|_1^2 \\ &\leq \frac{128C\tilde{K}^2\|\nu\|_2^2 M}{3\tau_0^4 \tilde{C}^2} + 2\sqrt{|F|M}\|\nu\|_2. \end{aligned}$$

Since $\|\nu\|_2 \leq \frac{C_1}{M}$, we have $\|\delta\|_2^2 \leq C''\sqrt{M}\|\nu\|_2$.

For the estimate \tilde{x} of x , constructed by Algorithm 3, we have

$$\tilde{x} = (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} c = e^{i\theta} x + (\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''} \delta.$$

As such, we obtain the desired bound on the reconstruction error,

$$\|\tilde{x} - e^{i\theta} x\|_2^2 \leq \|(\Phi_{\Lambda''} \Phi_{\Lambda''}^*)^{-1} \Phi_{\Lambda''}\|_2^2 \|\delta\|_2^2 = \frac{\|\delta\|_2^2}{\sigma_{\min}^2(\Phi_{\Lambda''}^*)} \leq \frac{C'' \sqrt{M} \|\nu\|_2}{\sigma_{\min}^2(\Phi_{\Lambda''}^*)}. \blacksquare$$

5. Numerical results on the robustness of Algorithm 3. In this section we numerically investigate the behavior of the constructed phase retrieval algorithm for time-frequency structured measurements (Algorithm 3). In particular, we use numerical simulations to demonstrate robustness of Algorithm 3 in the presence of noise and to investigate dependencies of the reconstruction error on various parameters.

Recall that, to construct the measurement frame $\Phi = \Phi_\Lambda \cup \Phi_E \subset \mathbb{C}^M$, we first introduce a random graph of measurements $G = (\Lambda, E)$, where $\Lambda = F \times \mathbb{Z}_M$ with $F \subset \mathbb{Z}_M$ such that $|F|$ is a constant that does not depend on the ambient dimension M , and

$$E = \{((k_1, \ell_1), (k_2, \ell_2)), \text{ s.t. } k_1, k_2 \in F, \ell_2 - \ell_1 \in C\} \subset \Lambda \times \Lambda.$$

Here, set C is a random subset of \mathbb{Z}_M , such that

$$(13) \quad C = D \cup (-D) \setminus \{0\} \subset \mathbb{Z}_M \text{ with } \mathbf{1}_D(m) \sim \text{i.i.d. } B\left(1, \frac{d \log M}{M}\right),$$

that is, $D \subset \mathbb{Z}_M$ is constructed at random, so that every $m \in \mathbb{Z}_M$ is chosen to be an element of D independently with probability $\frac{d \log M}{M}$ for some parameter $d > 0$.

The measurement frame Φ_Λ and the set of vectors for additional measurements Φ_E are then given by

$$\begin{aligned} \Phi_\Lambda &= (g, \Lambda), \quad g \in \mathbb{C}^M \text{ uniformly distributed on } \mathbb{S}^{M-1} \subset \mathbb{C}^M, \\ \Phi_E &= \{\pi(\lambda_1)g + \omega^t \pi(\lambda_2)g\}_{(\lambda_1, \lambda_2) \in E, t \in \{0, 1, 2\}} \text{ with } \omega = e^{2\pi i/3}. \end{aligned}$$

For our numerical simulations, we consider noisy phaseless measurements

$$\begin{aligned} b_\lambda &= |\langle x, \pi(\lambda)g \rangle|^2 + \nu_\lambda, \quad \lambda \in \Lambda, \\ b_{\lambda_1 \lambda_2 t} &= |\langle x, \pi(\lambda_1)g + \omega^t \pi(\lambda_2)g \rangle|^2 + \nu_{\lambda_1 \lambda_2 t}, \quad (\lambda_1, \lambda_2) \in E, t \in \{0, 1, 2\}, \end{aligned}$$

where $\nu_\lambda, \nu_{\lambda_1 \lambda_2 t} \sim \text{i.i.d. } \mathcal{N}(0, \sigma)$ are independent normally distributed additive noise components. Theorem 4.7 then gives the following bound on the reconstruction error of Algorithm 3:

$$(14) \quad \|\tilde{x} - e^{i\theta} x\|_2^2 \leq C \sqrt{M} \|\nu\|_2,$$

where the constant C depends on the spectral gap of the graph of measurements G and on the parameter

$$\Delta = \min_{\substack{\Lambda'' \subset \Lambda, \\ |\Lambda''| \geq 2/3 |\Lambda|}} \sigma_{\min}^2(\Phi_{\Lambda''}^*).$$

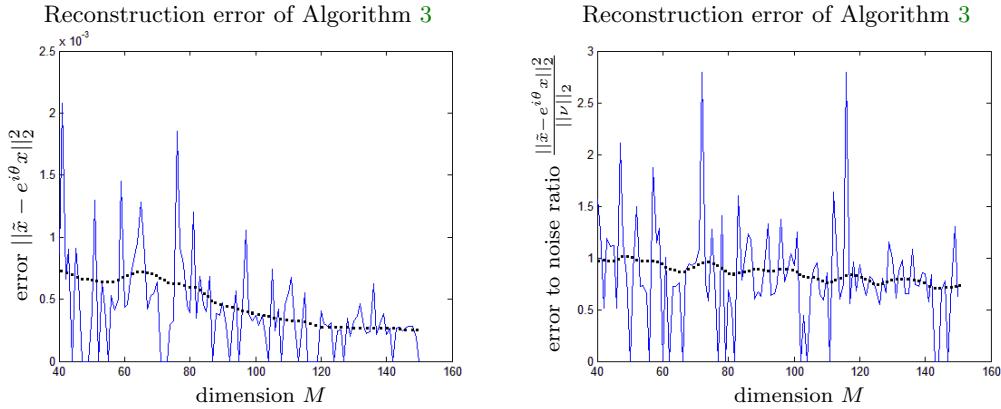


Figure 4. Dependence of the reconstruction error of Algorithm 3 (left) and the error to noise ratio (right) on the ambient dimension M . Here the noise vector is random, such that it has independent normally distributed entries with variance $\sigma = 10^{-3}$. The black dashed lines on the plots show the average (over several simulations with different noise realizations) values of the reconstruction error and the error to noise ratio, respectively. These numerical results suggest that the error to noise ratio does not depend on the signal dimension and is bounded above by a numerical constant.

Singular values of the analysis matrices of Gabor frames with random windows are studied in [31]. In particular, it has been shown that the smallest singular value of the analysis matrix $\sigma_{\min}^2(\Phi_{\Lambda''}^*)$ of a randomly selected subframe (g, Λ'') of a Gabor frame (g, Λ'') is bounded from below by $c \frac{|\Lambda'|}{M}$ for a suitably chosen numerical constant $c > 0$ with high probability. Unfortunately, no uniform bounds, similar to the parameter Δ defined above, are known to the date for Gabor frames. We formulate the following conjecture.

Conjecture 5.1. Consider a Gabor frame (g, Λ) with g uniformly distributed on \mathbb{S}^{M-1} and $\Lambda \subset \mathbb{Z}_M \times \mathbb{Z}_M$, such that $|\Lambda| = O(M \log^\alpha M)$ (where the parameter $\alpha \geq 0$ has to be specified). Then, for any $\varepsilon \in (0, 1)$,

$$\Delta(p) = \min_{\substack{\Lambda'' \subset \Lambda, \\ |\Lambda''| \geq p|\Lambda|}} \sigma_{\min}^2(\Phi_{\Lambda''}^*) \geq c \frac{|\Lambda|}{M}$$

with probability at least $1 - \varepsilon$, where $c > 0$ is some numerical constant that depends only on p and ε .

To illustrate Theorem 4.7, we consider two sets of simulations. For the first one, we let the dimension of the signal vary and explore the reconstruction error of the algorithm for a random normally distributed noise vector with independent entries and fixed variance. On Figure 4 we show the obtained results, which suggest that the error to noise ratio does not depend on the signal dimension, unlike the bound (14) obtained in Theorem 4.7. In fact, the ratio between the reconstruction error and the norm of the noise vector appears to be bounded above by a numerical constant close to 3.

For the second set of simulations, we explore the dependence of the reconstruction error and the error to noise ratio on the noise variance for a fixed signal dimension. The obtained results, shown on Figure 5, illustrate that the reconstruction error grows linearly with the magnitude of noise.

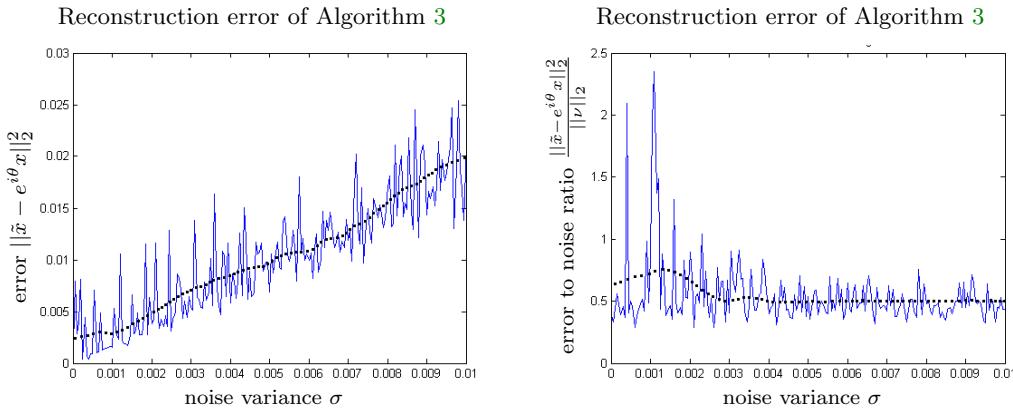


Figure 5. Dependence of the reconstruction error of Algorithm 3 (left) and the error to noise ratio (right) on the variance σ of the entries $\nu_\lambda, \nu_{\lambda_1 \lambda_2 t}$ of the noise vector ν , which are selected independently from the normal distribution $\mathcal{N}(0, \sigma)$. The ambient dimension here is $M = 100$. The black dashed lines on the plots show the average (over several simulations with different noise realizations) values of the reconstruction error and the error to noise ratio, respectively. These numerical results suggest that the reconstruction error grows linearly with the magnitude of noise.

On both Figures 4 and 5, we show the average values of the reconstruction error and the error to noise ratio (over several simulations with different noise realizations) using black dashed lines. In other words, the black dashed lines on the plots show (an approximation of) the expected values of the corresponding quantities. We note that, on both figures, the average error to noise ratio appears to be smaller than 1, which means that noise reduction takes place during signal reconstruction. This can be explained in the following way. Assuming that the graph of measurements G is sufficiently well connected, that is, $\text{spg}(G)$ is sufficiently big, the phase of a frame coefficient can be propagated to the corresponding vertex using various different paths. In Algorithm 3, we use the angular synchronization algorithm, which utilizes relative phase information coming to a vertex $\lambda \in \Lambda$ from all edges $(\lambda, \lambda') \in E$ incident to λ . Since in the simulations we considered noise with independent entries and zero mean, it tends to cancel itself at a vertex.

The reason why the plots on Figures 4 and 5 look quite spiky is that different realizations of the random graph of measurements G are used for the simulations. As we mentioned before, the reconstruction error bound (14) of Algorithm 3 depends on the spectral gap of G , which might differ from one realization to another. In particular, the bigger the cardinality of the random set C is, the better are the connectivity properties of G , and the smaller is the reconstruction error.

The cardinality of the set C and, thus, the reconstruction error of the algorithm depend on the parameter d , as formula (13) shows. More precisely, it follows from Lemma 3.3 that, provided $d > 36$,

$$\text{spg}(G) \geq 1 - \frac{6}{\sqrt{d}}$$

with overwhelming probability. We now investigate the dependence of the reconstruction error of the algorithm on the parameter d numerically.

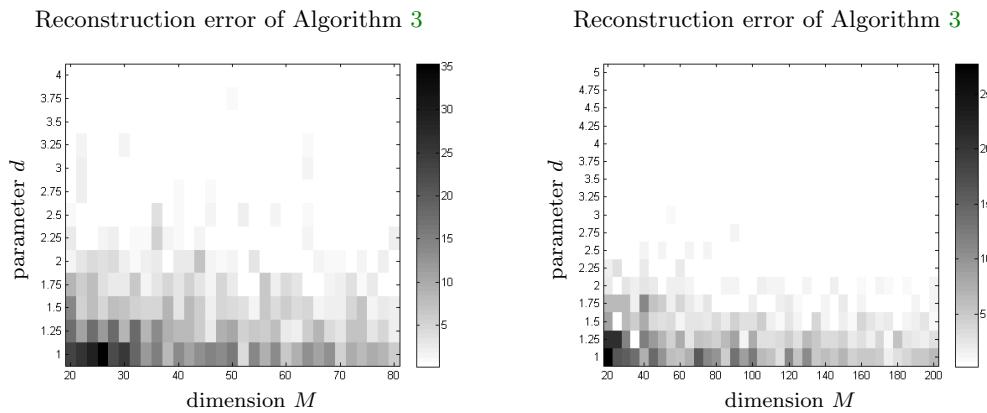


Figure 6. Dependence of the reconstruction error to noise ratio of Algorithm 3 on the parameter d in (13) for various dimensions. Here, values of high error to noise ratio are color coded darker than those with a low ratio. Parameter d controls the connectivity properties of the graph of measurements G . These numerical results suggest that, starting at approximately $d = 3$, the error to noise ratio does not exceed 4, and also does not depend on the ambient dimension M . This observation allows us to reduce the multiplicative constant in the number of measurements required for the reconstruction.

Numerical results presented on Figure 6 show the dependence of the ratio between the reconstruction error of Algorithm 3 and the norm of the noise vector on the parameter d (vertical axis) for the varying ambient dimension (horizontal axis). One can see that starting approximately at $d = 3$, that is, much earlier than the value $d = 144$, predicted by Theorem 4.7, this ratio does not exceed 4.

6. Conclusions. For most of the existing phase retrieval algorithms, including PhaseLift [11, 12] and the Wirtinger flow algorithm [15], recovery and robustness guarantees are proven for the case of a random measurement frame with independent frame vectors. One of the main reasons for this is that properties of such frames are sufficiently well studied. Moreover, such frames appear to have optimal properties in the sense of being “well spread,” which is formalized in different ways for different reconstruction algorithms.

In this paper, we use the idea of polarization [18] to design the first phase retrieval algorithm for time-frequency structured frames Φ with $|\Phi| < M^2$. An investigation of properties of Gabor frames, that is, frames consisting of vectors which are time and frequency shifts of the random window and, thus, are not independent, allow us to conclude recovery and robustness guarantees for the postulated phase retrieval algorithm. In particular, Theorem 4.1, which describes the order statistics of frame coefficients for a Gabor frame with a random window, allows us to obtain robustness guarantees in Theorem 4.7.

The numerical results presented in section 5 suggest that the theoretical bound (14) on the reconstruction error of the proposed phase retrieval algorithm can be further improved by a factor of \sqrt{M} . Thus, one of the important tasks for future research is to understand the gap between theoretically predicted robustness guarantees and results obtained numerically. We hope that a further study of properties of Gabor frames with random windows will allow us to not only remove the factor of \sqrt{M} from the reconstruction error bound (14), but also to prove the following conjecture, which states the uniform robustness guarantees for Algorithm 3.

Conjecture 6.1. Consider the measurement procedure (11) with $|F|$ and d sufficiently large. If the noise vector satisfies $\frac{\|\nu\|_2}{\|x\|_2^2} \leq \frac{C_1}{M}$ for some C_1 small enough, then there exists a numerical constant $C > 0$ such that with overwhelming probability, for every $x \in \mathbb{C}^M$, the estimate \tilde{x} produced by Algorithm 3 satisfies

$$\min_{\theta \in [0, 2\pi)} \|\tilde{x} - e^{i\theta}x\|_2^2 \leq C\|\nu\|_2.$$

We note that one of the main ingredients of the proof of Theorem 4.7 is Theorem 4.1, which gives bounds on the frame order statistics of a Gabor frame with a random window (see section 4.1). Similarly, the main missing ingredient of the proof of Conjecture 6.1 is a uniform version of Theorem 4.1.

Appendix A. Proof of Theorem 4.1. Here we prove Theorem 4.1 which is formulated as follows.

Theorem. Fix $x \in \mathbb{S}^{M-1} \subset \mathbb{C}^M$ and consider a Gabor frame (g, Λ) with $\Lambda \subset \mathbb{Z}_M \times \mathbb{Z}_M$ and a random window g uniformly distributed on the unit sphere \mathbb{S}^{M-1} . Then the following holds.

(a) For any $c > 0$ and $k > 0$, with probability at least $1 - \frac{1}{k^2}$, we have

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| < \frac{c}{\sqrt{M}} \right\} \right| < |\Lambda|(c^2 + kc).$$

(b) For any $K > 0$ and $k > 0$, with probability at least $1 - \frac{1}{k^2}$, we have

$$\left| \left\{ \lambda \in \Lambda, \text{ s.t. } |\langle x, \pi(\lambda)g \rangle| > \frac{K}{\sqrt{M}} \right\} \right| < |\Lambda| \left(\frac{8}{\pi} e^{-K^2} + k \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} \right).$$

Proof. (a) For $x \in \mathbb{S}^{M-1}$ fixed, we set

$$G_\delta(x) = \{ \varphi \in \mathbb{S}^{M-1} \subset \mathbb{C}^M, \text{ s.t. } |\langle x, \varphi \rangle| < \delta \}.$$

We are interested in the distribution of the random variable

$$Z_x = |G_\delta(x) \cap (g, \Lambda)| = \sum_{\lambda \in \Lambda} \mathbf{1}_{\{\pi(\lambda)g \in G_\delta(x)\}},$$

where $\mathbf{1}_{\{\varphi \in G_\delta(x)\}}$ is the characteristic function of the event $\{\varphi \in G_\delta(x)\}$. In words, Z_x is the number of small measurements of the fixed signal x with respect to the Gabor frame (g, Λ) with a random window g .

First, note that for each $\lambda \in \Lambda$, $\pi(\lambda)g$ is also uniformly distributed on \mathbb{S}^{M-1} , that is, the random vectors $\pi(\lambda)g$, $\lambda \in \Lambda$, are equally distributed. Indeed, consider a random vector h , such that $h(m) \sim \text{i.i.d. } \mathcal{CN}(0, \frac{1}{M})$. Then it is well known that $h/\|h\|_2$ is uniformly distributed on \mathbb{S}^{M-1} , since random vector $h/\|h\|_2$ almost surely has unit norm and its distribution is rotation invariant [32]. In other words, we can write $g = h/\|h\|_2$. As such, we have

$$\pi(\lambda)g = \pi(\lambda)h/\|h\|_2 = \pi(\lambda)h/\|\pi(\lambda)h\|_2,$$

since both modulation and time shift are unitary operators. The coordinates of h are i.i.d. random variables, thus translation T_k , which is just a permutation of the vector coordinates, preserves the distribution of h . For the modulation we have $M_\ell h(m) = e^{2\pi im\ell/M} h(m)$ is also normally distributed, with

$$\begin{aligned}\mathbb{E}(e^{2\pi im\ell/M} h(m)) &= e^{2\pi im\ell/M} \mathbb{E}(h(m)) = 0, \\ \text{Var}(e^{2\pi im\ell/M} h(m)) &= \mathbb{E}(e^{2\pi im\ell/M} h(m) e^{-2\pi im\ell/M} \bar{h}(m)) = \text{Var}(h(m)) = \frac{1}{M}.\end{aligned}$$

Thus the distribution of h is preserved by both modulation and time shift, and $\pi(\lambda)g = \pi(\lambda)h/\|\pi(\lambda)h\|$ has the same distribution as $g = h/\|h\|_2$.

Since $\pi(\lambda)g$ has the same distribution as g , we have

$$\mathbb{P}\{|\langle x, \pi(\lambda)g \rangle| < \delta\} = \mathbb{P}\{|\langle x, g \rangle| < \delta\}$$

for all $\lambda \in \Lambda$. Thus

$$(15) \quad \mathbb{E}(Z_x) = \sum_{\lambda \in \Lambda} \mathbb{P}\{\pi(\lambda)g \in G_\delta(x)\} = |\Lambda| \mathbb{P}\{|\langle x, g \rangle| < \delta\}.$$

Let R be an orthogonal matrix, such that $x = Re_1$, where $e_1 = (1, 0, 0, \dots, 0)^T$ is the first vector of the standard basis. Then

$$|\langle x, g \rangle| = |\langle Re_1, g \rangle| = |\langle e_1, R^*g \rangle|.$$

By the rotational symmetry of the distribution of g , we obtain

$$\mathbb{P}\{|\langle x, g \rangle| < \delta\} = \mathbb{P}\{|\langle e_1, g \rangle| < \delta\} = \mathbb{P}\{|g(0)| < \delta\}.$$

Let us identify the complex unit sphere $\mathbb{S}^{M-1} \subset \mathbb{C}^M$ with the real unit sphere $\mathbb{S}_{\mathbb{R}}^{2M-1} \subset \mathbb{R}^{2M}$, using the map $\mathcal{I} : \mathbb{S}^{M-1} \rightarrow \mathbb{S}_{\mathbb{R}}^{2M-1}$ given by

$$(16) \quad \mathcal{I}(z_0, \dots, z_{M-1}) = (\Re(z_0), \Im(z_0), \dots, \Re(z_{M-1}), \Im(z_{M-1})).$$

Since g is uniformly distributed on \mathbb{S}^{M-1} , $\tilde{g} = \mathcal{I}(g)$ is uniformly distributed on $\mathbb{S}_{\mathbb{R}}^{2M-1}$. Thus

$$\mathbb{P}\{|g(0)| < \delta\} = \mathbb{P}\{\tilde{g}(0)^2 + \tilde{g}(1)^2 < \delta^2\} = \frac{S_{<\delta}}{S_1},$$

where $S_1 = \frac{2\pi^M}{(M-1)!}$ is the surface area of $\mathbb{S}_{\mathbb{R}}^{2M-1}$, and $S_{<\delta}$ is the surface area of the set $\{z \in \mathbb{S}_{\mathbb{R}}^{2M-1}, \text{ s.t. } z_0^2 + z_1^2 < \delta^2\}$:

$$S_{<\delta} = \int_{-\delta}^{\delta} \int_{-\sqrt{\delta^2 - z_0^2}}^{\sqrt{\delta^2 - z_0^2}} \frac{2\pi^{M-1}}{(M-2)!} (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 < \frac{2\pi^M \delta^2}{(M-2)!},$$

that is, $\mathbb{P}\{|g(0)| < \delta\} \leq \delta^2(M-1)$. Now, setting $\delta = \frac{c}{\sqrt{M}}$ for c sufficiently small, we obtain

$$(17) \quad \mathbb{P}\left\{|\langle x, g \rangle| < \frac{c}{\sqrt{M}}\right\} \leq c^2.$$

Using (15) and (17), we obtain

$$(18) \quad \mu = \mathbb{E}(Z_x) = |\Lambda| \mathbb{P} \left\{ |\langle x, g \rangle| < \frac{c}{\sqrt{M}} \right\} \leq |\Lambda| c^2.$$

Similarly, using (17), for the variance of Z_x we obtain

$$\begin{aligned} \sigma^2 &= \text{Var}(Z_x) = \mathbb{E}(Z_x^2) - (\mathbb{E}(Z_x))^2 \leq \mathbb{E}(Z_x^2) = \mathbb{E} \left(\left(\sum_{\lambda \in \Lambda} \mathbf{1}_{\{\pi(\lambda)g \in G \frac{c}{\sqrt{M}}(x)\}} \right)^2 \right) \\ &= \mathbb{E} \left(\sum_{\lambda \in \Lambda} \mathbf{1}_{\{\pi(\lambda)g \in G \frac{c}{\sqrt{M}}(x)\}}^2 + \sum_{\substack{(\lambda_1, \lambda_2) \in \Lambda^2, \\ \lambda_1 \neq \lambda_2}} \mathbf{1}_{\{\pi(\lambda_1)g \in G \frac{c}{\sqrt{M}}(x)\}} \mathbf{1}_{\{\pi(\lambda_2)g \in G \frac{c}{\sqrt{M}}(x)\}} \right) \\ &= \sum_{\lambda \in \Lambda} \mathbb{P} \left\{ \pi(\lambda)g \in G \frac{c}{\sqrt{M}}(x) \right\} + \sum_{\substack{(\lambda_1, \lambda_2) \in \Lambda^2, \\ \lambda_1 \neq \lambda_2}} \mathbb{P} \left\{ \pi(\lambda_1)g, \pi(\lambda_2)g \in G \frac{c}{\sqrt{M}}(x) \right\} \\ (19) \quad &\leq (|\Lambda| + (|\Lambda|^2 - |\Lambda|)) \mathbb{P} \left\{ |\langle x, g \rangle| < \frac{c}{\sqrt{M}} \right\} \leq c^2 |\Lambda|^2, \end{aligned}$$

that is, $\sigma \leq c|\Lambda|$. Then, using the Chebyshev inequality and bounds (18) and (19), we have

$$\mathbb{P}\{Z_x \geq |\Lambda|(c^2 + kc)\} \leq \mathbb{P}\{Z_x \geq \mu + k\sigma\} \leq \mathbb{P}\{|Z_x - \mu| \geq k\sigma\} \leq \frac{1}{k^2}$$

for any $k > 0$. In other words, if we delete the $|\Lambda|(c^2 + kc)$ smallest phaseless measurements, for the remaining measurements with probability at least $1 - \frac{1}{k^2}$ we would have $|\langle x, \pi(\lambda)g \rangle| \geq \frac{c}{\sqrt{M}}$. This concludes the proof of (a).

The proof of (b) follows the same steps. Let K be a constant and consider the following random variable

$$U_x = \sum_{\lambda \in \Lambda} \mathbf{1}_{\{|\langle x, \pi(\lambda)g \rangle| > K/\sqrt{M}\}}.$$

Since, for each $\lambda \in \Lambda$, $\pi(\lambda)g$ has the same (uniform on \mathbb{S}^{M-1}) distribution as g , we have $\mathbb{P}\{|\langle x, \pi(\lambda)g \rangle| > K/\sqrt{M}\} = \mathbb{P}\{|\langle x, g \rangle| > K/\sqrt{M}\}$ for all $\lambda \in \Lambda$. As above, we have

$$\mathbb{P}\{|\langle x, g \rangle| > K/\sqrt{M}\} = \mathbb{P}\{|\langle e_1, g \rangle| > K/\sqrt{M}\} = \mathbb{P}\{|g(0)| > K/\sqrt{M}\}.$$

Using the map $\mathcal{I} : \mathbb{S}^{M-1} \rightarrow \mathbb{S}_{\mathbb{R}}^{2M-1}$ defined in (16), for $\tilde{g} = \mathcal{I}(g)$ we obtain

$$\mathbb{P} \left\{ |g(0)| > \frac{K}{\sqrt{M}} \right\} = \mathbb{P} \left\{ \tilde{g}(0)^2 + \tilde{g}(1)^2 > \frac{K^2}{M} \right\} = \frac{S_{>K/\sqrt{M}}}{S_1},$$

where $S_1 = \frac{2\pi^M}{(M-1)!}$ is the surface area of $\mathbb{S}_{\mathbb{R}}^{2M-1}$ and $S_{>K/\sqrt{M}}$ is the surface area of the set

$\{z \in \mathbb{S}_{\mathbb{R}}^{2M-1}, \text{ s.t. } z_0^2 + z_1^2 > \frac{K^2}{M}\}:$

$$\begin{aligned}
S_{>K/\sqrt{M}} &= \int_{|z_0| \leq 1} \int_{\sqrt{\frac{K^2}{M} - z_0^2} < |z_1| < \sqrt{1 - z_0^2}} \frac{2\pi^{M-1}}{(M-2)!} (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 \\
&= 8 \frac{2\pi^{M-1}}{(M-2)!} \int_0^{\frac{K}{\sqrt{2M}}} \int_{\sqrt{\frac{K^2}{M} - z_0^2}}^{\sqrt{1 - z_0^2}} (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 \\
&\quad + 8 \frac{2\pi^{M-1}}{(M-2)!} \int_{\frac{K}{\sqrt{2M}}}^{\frac{1}{\sqrt{2}}} \int_{z_0}^{\sqrt{1 - z_0^2}} (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 \\
&\leq \frac{16\pi^{M-1}}{(M-2)!} \frac{\sqrt{2M}}{K} \int_0^{\frac{K}{\sqrt{2M}}} \int_{\sqrt{\frac{K^2}{M} - z_0^2}}^{\sqrt{1 - z_0^2}} z_1 (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 \\
&\quad + \frac{16\pi^{M-1}}{(M-2)!} \frac{\sqrt{2M}}{K} \int_{\frac{K}{\sqrt{2M}}}^{\frac{1}{\sqrt{2}}} \int_{z_0}^{\sqrt{1 - z_0^2}} z_1 (1 - z_0^2 - z_1^2)^{\frac{2M-1}{2}} dz_1 dz_0 \\
&\leq \frac{8\pi^{M-1}}{(M-1)!} \left(\left(1 - \frac{K^2}{M}\right)^{\frac{2M+1}{2}} + \frac{M}{2K^2} \int_{\frac{K}{\sqrt{2M}}}^{\frac{1}{\sqrt{2}}} 4z_0 (1 - 2z_0^2)^{\frac{2M+1}{2}} dz_0 \right) \\
&\leq \frac{8\pi^{M-1}}{(M-1)!} \left(\left(1 - \frac{K^2}{M}\right)^{\frac{2M+1}{2}} + \frac{1}{2K^2} \left(1 - \frac{K^2}{M}\right)^{\frac{2M+3}{2}} \right) \\
&\leq \frac{8\pi^{M-1}}{(M-1)!} \left(e^{-\frac{K^2}{M} \frac{2M+1}{2}} + \frac{1}{2K^2} e^{-\frac{K^2}{M} \frac{2M+3}{2}} \right) \leq \frac{16\pi^{M-1}}{(M-1)!} e^{-K^2}.
\end{aligned}$$

Here, we used the symmetry of the domain of integration, the fact that $z_0^2 + z_1^2 < \frac{K^2}{M}$ implies $\max\{|z_0|, |z_1|\} > \frac{K}{\sqrt{2M}}$, and the inequality $1 - x \leq e^{-x}$. Using the computed bound for $S_{>K/\sqrt{M}}$, we obtain $\mathbb{P}\left\{ |g(0)| > \frac{K}{\sqrt{M}} \right\} \leq \frac{8}{\pi} e^{-K^2}$. Then

$$(20) \quad \mu = \mathbb{E}(U_x) = \sum_{\lambda \in \Lambda} \mathbb{P}\left\{ |\langle x, \pi(\lambda)g \rangle| > \frac{K}{\sqrt{M}} \right\} \leq \frac{8}{\pi} e^{-K^2} |\Lambda|.$$

Similarly, for the variance of U_x we obtain

$$\begin{aligned}
\sigma^2 &= \text{Var}(U_x) \leq \mathbb{E}(U_x^2) = \mathbb{E}\left(\left(\sum_{\lambda \in \Lambda} \mathbf{1}_{\{|\langle x, \pi(\lambda)g \rangle| > K/\sqrt{M}\}}\right)^2\right) \\
&= \sum_{\lambda \in \Lambda} \mathbb{P}\left\{ |\langle x, \pi(\lambda)g \rangle| > \frac{K}{\sqrt{M}} \right\} + \sum_{\substack{(\lambda_1, \lambda_2) \in \Lambda^2, \\ \lambda_1 \neq \lambda_2}} \mathbb{P}\left\{ |\langle x, \pi(\lambda_1)g \rangle|, |\langle x, \pi(\lambda_2)g \rangle| > \frac{K}{\sqrt{M}} \right\} \\
(21) \quad &\leq |\Lambda| \mathbb{P}\{|\langle x, g \rangle| > K/\sqrt{M}\} + (|\Lambda|^2 - |\Lambda|) \mathbb{P}\{|\langle x, g \rangle| > K/\sqrt{M}\} \leq \frac{8}{\pi} e^{-K^2} |\Lambda|^2,
\end{aligned}$$

that is, $\sigma \leq \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} |\Lambda|$. Then, using the Chebyshev inequality and bounds (20) and (21), we obtain

$$\mathbb{P}\left\{ U_x \geq |\Lambda| \left(\frac{8}{\pi} e^{-K^2} + k \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} \right) \right\} \leq \mathbb{P}\{|U_x - \mu| \geq k\sigma\} \leq \frac{1}{k^2}$$

for any $k > 0$. In other words, if we delete the

$$|\Lambda| \left(\frac{8}{\pi} e^{-K^2} + k \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\frac{K^2}{2}} \right)$$

largest phaseless measurements, with probability at least $1 - \frac{1}{k^2}$ for the remaining measurements we would have $|\langle x, \pi(\lambda)g \rangle| \leq \frac{K}{\sqrt{M}}$. \blacksquare

Acknowledgments. The authors thank the anonymous referees for providing thoughtful suggestions that led to a more complete discussion of the context of our results. P. Salanevich thanks Prof. Dr. Felix Krahmer for insightful discussions during the “Mathematics of Signal Processing” trimester program (Hausdorff Research Institute for Mathematics, Bonn, Germany), and Prof. Dr. Terence Tao for his valuable suggestions during her visit at UCLA in Spring 2015. P. Salanevich also thanks Prof. Dr. Holger Boche for helpful discussions during her visit to Technical University of Munich in December 2014.

REFERENCES

- [1] P. SALANEVICH AND G. E. PFANDER, *Polarization based phase retrieval for time-frequency structured measurements*, in 2015 International Conference on Sampling Theory and Applications (SampTA), IEEE, Piscataway, NJ, 2015, pp. 187–191.
- [2] P. SALANEVICH AND G. E. PFANDER, *Geometric properties of Gabor frames with a random window*, in 2017 International Conference on Sampling Theory and Applications (SampTA), IEEE, Piscataway, NJ, 2017, pp. 183–187.
- [3] R. P. MILLANE, *Phase retrieval in crystallography and optics*, J. Opt. Soc. Amer. A, 7 (1990), pp. 394–411.
- [4] J. R. FIENUP AND J. C. DAINTY, *Phase Retrieval and Image Reconstruction for Astronomy*, Academic, San Diego, CA, 1987, pp. 231–275.
- [5] J. MIAO, T. ISHIKAWA, Q. SHEN, AND T. EARNEST, *Extending X-ray crystallography to allow the imaging of non-crystalline materials, cells and single protein complexes*, Annu. Rev. Phys. Chem., 59 (2008), pp. 387–410.
- [6] O. BUNK, A. DIAZ, F. PFEIFFER, C. DAVID, B. SCHMITT, D. K. SATAPATHY, AND J. F. VAN DER VEEN, *Diffractive imaging for periodic samples: Retrieving one-dimensional concentration profiles across microfluidic channels*, Acta Cryst. A Found. Adv., 63 (2007), pp. 306–314.
- [7] J. D. WATSON AND F. H. C. CRICK, *A structure for deoxyribose nucleic acid*, in *A Century of Nature: Twenty-One Discoveries that Changed Science and the World*, University of Chicago Press, Chicago, 2003, pp. 73–84.
- [8] H. SAHINOGLU AND S. D. CABRERA, *On phase retrieval of finite-length sequences using the initial time sample*, IEEE Trans. Circuits Syst., 38 (1991), pp. 954–958.
- [9] R. W. GERCHBERG AND W. O. SAXTON, *A practical algorithm for the determination of phase from image and diffraction plane pictures*, Optik, 35 (1972), pp. 237–246.
- [10] R. BALAN, B. G. BODMANN, P. G. CASAZZA, AND D. EDIDIN, *Painless reconstruction from magnitudes of frame coefficients*, J. Fourier Anal. Appl., 15 (2009), pp. 488–501.
- [11] E. J. CANDÈS, Y. C. ELDAR, T. STROHMER, AND V. VORONINSKI, *Phase retrieval via matrix completion*, SIAM J. Imaging Sci., 6 (2013), pp. 199–225.
- [12] E. J. CANDÈS, T. STROHMER, AND V. VORONINSKI, *PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming*, Comm. Pure Appl. Math., 66 (2013), pp. 1241–1274.
- [13] E. J. CANDÈS AND X. LI, *Solving quadratic equations via PhaseLift when there are about as many equations as unknowns*, Found. Comput. Math., 14 (2014), pp. 1017–1026.
- [14] E. J. CANDÈS, X. LI, AND M. SOLTANOLKOTABI, *Phase retrieval from coded diffraction patterns*, Appl. Comput. Harmon. Anal., 39 (2015), pp. 277–299.

[15] E. J. CANDÈS, X. LI, AND M. SOLTANOLKOTABI, *Phase retrieval via Wirtinger flow: Theory and algorithms*, IEEE Trans. Inform. Theory, 61 (2015), pp. 1985–2007.

[16] Y. CHEN AND E. J. CANDÈS, *Solving random quadratic systems of equations is nearly as easy as solving linear systems*, in Advances in Neural Information Processing Systems, Neural Information Processing Systems Foundation, La Jolla, CA, 2015, pp. 739–747.

[17] M. A. IWEN, A. VISWANATHAN, AND Y. WANG, *Fast phase retrieval from local correlation measurements*, SIAM J. Imaging Sci., 9 (2016), pp. 1655–1688.

[18] B. ALEXEEV, A. S. BANDEIRA, M. FICKUS, AND D. G. MIXON, *Phase retrieval with polarization*, SIAM J. Imaging Sci., 7 (2014), pp. 35–66.

[19] A. S. BANDEIRA, Y. CHEN, AND D. G. MIXON, *Phase retrieval from power spectra of masked signals*, Inf. Inference, 3 (2014), pp. 83–102.

[20] R. BALANA, P. CASAZZA, AND D. EDIDIN, *On signal reconstruction without phase*, Appl. Comput. Harmon. Anal., 20 (2006), pp. 345–356.

[21] I. BOJAROVSKA AND A. FLINTH, *Phase retrieval from Gabor measurements*, J. Fourier Anal. Appl., 22 (2016), pp. 542–567.

[22] B. ALEXEEV, J. CAHILL, AND D. G. MIXON, *Full spark frames*, J. Fourier Anal. Appl., 18 (2012), pp. 1167–1194.

[23] J. LAWRENCE, G. E. PFANDER, AND D. WALNUT, *Linear independence of Gabor systems in finite dimensional vector spaces*, J. Fourier Anal. Appl., 11 (2006), pp. 715–726.

[24] R.-D. MALIKIOSIS, *A note on Gabor frames in finite dimensions*, Appl. Comput. Harmon. Anal., 38 (2015), pp. 318–330.

[25] G. E. PFANDER, *Gabor frames in finite dimensions*, in Finite Frames: Theory and Applications, P. G. Casazza and G. Kutyniok, eds., Birkhäuser, Boston, 2013, pp. 193–239.

[26] P. HARSHA AND A. BARTH, *Lecture 5: Derandomization (Part II)*, <http://www.tcs.tifr.res.in/~prahladh/teaching/05spring/lectures/lec5.pdf> (2005).

[27] F. KRAHMER, S. MENDELSON, AND H. RAUHUT, *Suprema of chaos processes and the restricted isometry property*, Comm. Pure Appl. Math., 67 (2014), pp. 1877–1904.

[28] A. SINGER, *Angular synchronization by eigenvectors and semidefinite programming*, Appl. Comput. Harmon. Anal., 30 (2011), pp. 20–36.

[29] A. S. BANDEIRA, A. SINGER, AND D. A. SPIELMAN, *A Cheeger inequality for the graph connection Laplacian*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1611–1630.

[30] A. Y. NG, M. I. JORDAN, AND Y. WEISS, *On spectral clustering: Analysis and an algorithm*, Proceedings of the 14th International Conference on Neural Information Processing Systems, MIT Press, Cambridge, MA, 2001 pp. 849–856.

[31] P. SALANEVICH, *Extreme Singular Values of Random Time-Frequency Structured Matrices*, preprint, <https://arxiv.org/abs/1902.01062>.

[32] G. MARSAGLIA, *Choosing a point from the surface of a sphere*, Ann. Math. Statist., 43 (1972), pp. 645–646.