

# Phase Retrieval from Coded Diffraction Patterns

Emmanuel J. Candès\* Xiaodong Li† Mahdi Soltanolkotabi‡

November 11, 2013

## Abstract

This paper considers the question of recovering the phase of an object from intensity-only measurements, a problem which naturally appears in X-ray crystallography and related disciplines. We study a physically realistic setup where one can modulate the signal of interest and then collect the intensity of its diffraction pattern, each modulation thereby producing a sort of coded diffraction pattern. We show that *PhaseLift*, a recent convex programming technique, recovers the phase information exactly from a number of random modulations, which is polylogarithmic in the number of unknowns. Numerical experiments with noiseless and noisy data complement our theoretical analysis and illustrate our approach.

## 1 Introduction

### 1.1 The phase retrieval problem

In many areas of science and engineering, we only have access to magnitude measurements; for instance, it is far easier for detectors to record the modulus of the scattered radiation than to measure its phase. Imagine then that we have a discrete object  $\mathbf{x} \in \mathbb{C}^n$ , and that we would like to measure  $\langle \mathbf{a}_k, \mathbf{x} \rangle$  for some sampling vectors  $\mathbf{a}_k \in \mathbb{C}^n$  but only have access to phaseless measurements of the form

$$y_k = |\langle \mathbf{a}_k, \mathbf{x} \rangle|^2, \quad k = 1, \dots, m. \quad (1.1)$$

The phase retrieval problem is that of recovering the missing phase of the data  $\langle \mathbf{a}_k, \mathbf{x} \rangle$ . Once this information is available, one can find the vector  $\mathbf{x}$  by essentially solving a system of linear equations.

The quintessential phase retrieval problem, or phase problem for short, asks to recover a signal from the modulus of its Fourier transform. This comes from the fact that in coherent X-ray imaging, it follows from the Fraunhofer diffraction equation that the optical field at the detector is well approximated by the Fourier transform of the object of interest. Since photographic plates, CCDs and other light detectors can only measure light intensity, the problem is then to recover  $\mathbf{x} = \{x[t]\}_{t=0}^{n-1} \in \mathbb{C}^n$  from measurements of the type

$$y_k = \left| \sum_{t=0}^{n-1} x[t] e^{-i2\pi\omega_k t} \right|^2, \quad \omega_k \in \Omega, \quad (1.2)$$

---

\*Departments of Mathematics and of Statistics, Stanford University, Stanford CA

†Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA

‡Department of Electrical Engineering, Stanford University, Stanford CA

where  $\Omega$  is a sampled set of frequencies in  $[0, 1]$  (we stated the problem in one dimension to simplify matters). We thus recognize an instance of (1.1) in which the vectors  $\mathbf{a}_k$  are sampled values of complex sinusoids. X-ray diffraction images are of this form, and as is well known, permitted the discovery of the double helix [55]. In addition to X-ray crystallography [33, 39], the phase problem has numerous other applications in the imaging sciences such as diffraction and array imaging [18, 24], optics [54], speckle imaging in astronomy [26], and microscopy [38]. Other areas where related problems appear include acoustics [12, 8], blind channel estimation in wireless communications [4, 45], interferometry [28], quantum mechanics [25, 47] and quantum information [34].

## 1.2 Convex relaxation

Previous work [20, 24] suggested to bring convex programming techniques to bear on the phase retrieval problem. Returning to the general formulation (1.1), the phase problem asks to recover  $\mathbf{x} \in \mathbb{C}^n$  subject to data constraints of the form

$$\text{tr}(\mathbf{a}_k \mathbf{a}_k^* \mathbf{x} \mathbf{x}^*) = y_k, \quad k = 1, \dots, m,$$

where  $\text{tr}(\mathbf{X})$  is the trace of the matrix  $\mathbf{X}$ . The idea is then to lift the problem in higher dimensions: introducing the Hermitian matrix variable  $\mathbf{X} \in \mathcal{S}^{n \times n}$ , the phase problem is equivalent to finding  $\mathbf{X}$  obeying

$$\mathbf{X} \geq 0, \quad \text{rank}(\mathbf{X}) = 1, \quad \text{tr}(\mathbf{a}_k \mathbf{a}_k^* \mathbf{X}) = y_k \text{ for } k = 1, \dots, m \quad (1.3)$$

where, here and below,  $\mathbf{X} \geq 0$  means that  $\mathbf{X}$  is positive semidefinite. This problem is not tractable and, by dropping the rank constraint, is relaxed into

$$\begin{aligned} & \text{minimize} && \text{tr}(\mathbf{X}) \\ & \text{subject to} && \mathbf{X} \geq 0 \\ & && \text{tr}(\mathbf{a}_k \mathbf{a}_k^* \mathbf{X}) = y_k, \quad k = 1, \dots, m. \end{aligned} \quad (1.4)$$

PhaseLift (1.4) is a semidefinite program (SDP). If its solution happens to have rank one and is equal to  $\mathbf{x} \mathbf{x}^*$ , then a simple factorization recovers  $\mathbf{x}$  up to a global phase/sign.

We pause to emphasize that in different contexts, similar convex relaxations for optimizing quadratic objectives subject to quadratic constraints are known as Schor’s semidefinite relaxations, see [41, Section 4.3] and [30] on the MAXCUT problem from graph theory for spectacular applications of these ideas. For related convex relaxations of quadratic problems, we refer the interested reader to the wonderful tutorial [52].

## 1.3 This paper

Numerical experiments [20] together with emerging theory suggest that the PhaseLift approach is surprisingly effective. On the theoretical side, starting with [23], a line of work established that if the sampling vectors  $\mathbf{a}_k$  are sufficiently randomized, then the convex relaxation is provably exact. Assuming that the  $\mathbf{a}_k$ ’s are independent random (complex-valued) Gaussian vectors, [23] shows that on the order of  $n \log n$  quadratic measurements are sufficient to guarantee perfect recovery via (1.4) with high probability. A subset of the authors [21] reached the same conclusion from on the order of  $n$  equations only, by solving the SDP feasibility problem; to be sure, [21] establishes that the set of matrices obeying the constraints in (1.4) reduces to a unique point namely,  $\mathbf{x} \mathbf{x}^*$ ,

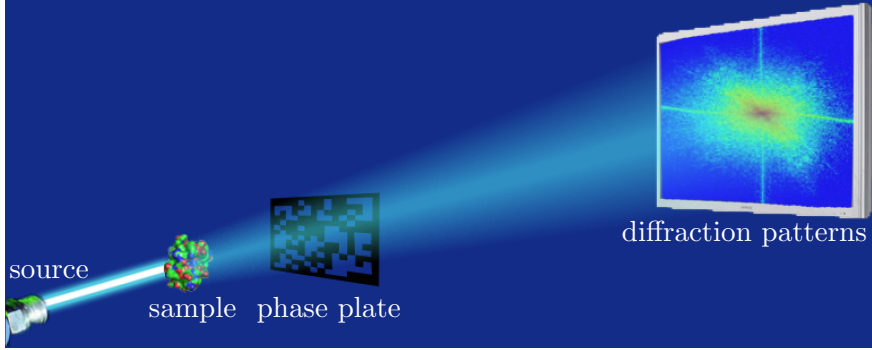


Figure 1: An illustrative setup for acquiring coded diffraction patterns.

see [27] for a similar result.<sup>1</sup> Finally, inspired by PhaseLift and the famous MAXCUT relaxation of Goemans and Williamson, [53] proposed another semidefinite relaxation called PhaseCut whose performance from noiseless data—in terms of the number of samples needed to achieve perfect recovery—turns out to be identical to that of PhaseLift.

While this is all reassuring, the problem is that the Gaussian model, in which each measurement gives us the magnitude of the dot product  $\sum_{t=0}^{n-1} x[t]a_k[t]$  between the signal and (complex-valued) Gaussian white noise, is very far from the kind of data one can collect in an  $X$ -ray imaging and many related experiments. The purpose of this paper is to show that the PhaseLift relaxation is still exact in a physically inspired setup where one can modulate the signal of interest and then let diffraction occur.

#### 1.4 Coded diffraction patterns

Imagine then that we modulate the signal before diffraction. Letting  $d[t]$  be the modulating waveform, we would observe the diffraction pattern

$$y_k = \left| \sum_{t=0}^{n-1} x[t]d[t]e^{-i2\pi\omega_k t} \right|^2, \quad \omega_k \in \Omega. \quad (1.5)$$

We call this a *coded diffraction pattern* (CDP) since it gives us information about the spectrum of  $\{x[t]\}$  modulated by the code  $\{d[t]\}$ . There are several ways of achieving modulations of this type: one can use a phase mask just after the sample, see Figure 1, or use an optical grating to modulate the illumination beam as mentioned in [37], or even use techniques from ptychography which scan an illumination patch on an extended specimen [48, 50]. We refer to [20] for a more thorough discussion of such approaches.

In this paper, we analyze such a data collection scheme in which one uses multiple modulations. Our model for data acquisition is thus as follows:

$$y_{\ell,k} = \left| \sum_{t=0}^{n-1} x[t]d_\ell[t]e^{-i2\pi kt/n} \right|^2, \quad \begin{array}{l} 0 \leq k \leq n-1 \\ 1 \leq \ell \leq L \end{array}. \quad (1.6)$$

<sup>1</sup>[21] also establishes near-optimal estimation bounds from noisy data.

In words, we collect the magnitude of the discrete Fourier transform (DFT) of  $L$  modulations of the signal  $\mathbf{x}$ . In matrix notation, letting  $\mathbf{D}_\ell$  be the diagonal matrix with the modulation pattern  $d_\ell[t]$  on the diagonal and  $\mathbf{f}_k^*$  be the rows of the DFT, we observe

$$y_{\ell,k} = |\mathbf{f}_k^* \mathbf{D}_\ell^* \mathbf{x}|^2.$$

We prove that if we use random modulation patterns (random waveforms  $d[t]$ ), then the solution to (1.4) is exact with high probability provided that we have sufficiently many CDPs. In fact, we will see that the feasible set in (1.4) equal to

$$\{\mathbf{X} : \mathbf{X} \succeq \mathbf{0} \text{ and } \mathcal{A}(\mathbf{X}) = \mathbf{y}\} \quad (1.7)$$

reduces to a single point  $\mathbf{x}\mathbf{x}^*$ . Above  $\mathcal{A} : \mathcal{S}^{n \times n} \rightarrow \mathbb{R}^{m=nL}$  ( $\mathcal{S}^{n \times n}$  is the space of self-adjoint matrices) is the linear mapping giving us the linear equalities in (1.4),

$$\mathcal{A}(\mathbf{X}) = \{\mathbf{f}_k^* \mathbf{D}_\ell^* \mathbf{X} \mathbf{D}_\ell \mathbf{f}_k\}_{\ell,k} = \{\text{tr}(\mathbf{D}_\ell \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}_\ell^* \mathbf{X})\}_{\ell,k}.$$

## 1.5 Main result

In this paper we work with one dimensional signals. We would like to mention that our methods would extend to higher dimensional signals by using unstructured two dimensional (and higher dimensional) codes, although we do not pursue this in this paper. However, one can also see that it is immediate to break down two dimensional (and higher dimensional) phase retrieval problems of the form studied in this paper into one dimensional problems by a “tensorizing” trick which we discuss further in Appendix B. Our model assumes random modulations and we work with diagonal matrices  $\mathbf{D}_\ell$ ,  $1 \leq \ell \leq L$ , which are i.i.d. copies of a matrix  $\mathbf{D}$ , whose entries are themselves i.i.d. copies of a random variable  $d$ . Throughout, we assume that  $d$  is symmetric, obeys  $|d| \leq M$  as well as the moment conditions

$$\mathbb{E} d = 0, \quad \mathbb{E} d^2 = 0, \quad \mathbb{E} |d|^4 = 2 \mathbb{E} |d|^2. \quad (1.8)$$

A random variable obeying these assumptions is said to be *admissible*. The reason why we can have  $\mathbb{E} d^2 = 0$  while  $d \neq 0$  is that  $d$  is complex valued. An example of an admissible random variable is  $d = b_1 b_2$ , where  $b_1$  and  $b_2$  are independent and distributed as

$$b_1 = \begin{cases} 1 & \text{with prob. } \frac{1}{4} \\ -1 & \text{with prob. } \frac{1}{4} \\ -i & \text{with prob. } \frac{1}{4} \\ i & \text{with prob. } \frac{1}{4} \end{cases} \quad \text{and} \quad b_2 = \begin{cases} 1 & \text{with prob. } \frac{4}{5} \\ \sqrt{6} & \text{with prob. } \frac{1}{5} \end{cases}. \quad (1.9)$$

We would like to emphasize that we impose  $\mathbb{E}[d^2] = 0$  mostly to simplify our exposition. In fact, the conclusion of Theorem 1.1 below remains valid if  $\mathbb{E}[d^2] \neq 0$ , although we do not prove this in this paper. In particular, we can also work with  $d$  distributed as

$$d = \begin{cases} 1 & \text{with prob. } \frac{1}{4} \\ 0 & \text{with prob. } \frac{1}{2} \\ -1 & \text{with prob. } \frac{1}{4} \end{cases}. \quad (1.10)$$

**Theorem 1.1** *Suppose that the modulation is admissible and that the number  $L$  of coded diffraction patterns obeys*

$$L \geq c \cdot \log^4 n,$$

*for some fixed numerical constant  $c$ . Then with probability at least  $1 - 1/n$ , the feasibility problem (1.7) reduces to a unique point, namely,  $\mathbf{x}\mathbf{x}^*$ , and thus recovers  $\mathbf{x}$  up to a global phase shift. For  $\gamma \geq 1$ , setting  $L \geq c\gamma \log^4 n$  leads to a probability of success at least  $1 - n^{-\gamma}$ .*

Thus, in a stylized physical setting, it is possible to recover an arbitrary signal from a fairly limited number of coded diffraction patterns by solving an SDP feasibility problem. As mentioned earlier, the equivalence from [53] implies that our theoretical guarantees automatically carry over to the PhaseCut formulation.

Mathematically, the phase recovery problem is different than that in which the sampling vectors are Gaussian as in [20]. The reason is that the measurements in Theorem 1.1 are far more structured and far ‘less random’. Loosely speaking, our random modulation model uses on the order of  $m := nL$  random bits whereas the Gaussian model with the same number of quadratic equations would use on the order of  $mn$  random bits (this can be formalized by using the notion of entropy from information theory). A consequence of this difference is that the proof of the theorem requires new techniques and ideas. Having said this, an open and interesting research direction is to close the gap—remove the log factors—and show whether or not perfect recovery can be achieved from a number of coded diffraction patterns independent of dimension.

The first version of this paper was made publicly available at the same time as [32], which begins to study the performance of PhaseLift from non-Gaussian sampling vectors. There, the authors study sampling schemes from certain finite vector configurations, dubbed t-designs. These models are different from ours and do not include our coded diffraction patterns as a special case. Hence, our results are not comparable. Having said this, there are similarities in the proof techniques, especially in the role played by the robust injectivity property, compare our Lemma 3.7 from Section 3.3 with [32, Section 3.3].

## 1.6 Other approaches to phase retrieval and related works

There are of course other approaches to phase retrieval and we mention some literature for completeness and to inform the interested reader of recent progress in this area. Balan [8] studies a problem where the sampling vectors model a short-time Fourier transform. Balan, Casazza and Edidin [13] formulate the phase retrieval problem as nonconvex optimization problem. In [12], the same authors [12] describe some applications of the phase problem in signal processing and speech analysis and presents some necessary and sufficient conditions which guarantee that the solution to (1.1) is unique. Other articles studying the minimal number of frame coefficient magnitudes for noiseless recovery include [19, 5, 15, 40, 10]. We recommend the two blog posts [2] and [3] by Mixon and the references therein for a comprehensive review and discussion of such results. Lower bounds on the performance of any recovery method from noisy data are studied in [15, 14, 29].

On the algorithmic side, [7] proposes a nonlinear scheme for phase retrieval having exponential time complexity in the dimension of the signal  $\mathbf{x}$  while [11] presents a tractable algorithm requiring a number of measurements at least quadratic in the dimension of the signal; that is to say,  $m \geq c \cdot n^2$  for some constant  $c > 0$ .

We also wish to mention some recent works aiming at presenting efficient reconstruction algorithms for generic frames (for certain types of sampling vectors such as those in [5, 16, 46] fast

implementations already exist) and give two references. The first [9] introduces an iterative regularized least-square algorithm and establishes convergence of the algorithm. It is however not known whether this algorithm enjoys accurate reconstruction guarantees. The second [42] is recent and studies an alternative minimization approach for phase retrieval, which yields very accurate but not exact reconstructions from a number of measurements that needs to be at least on the order of  $n \log n^3$  Gaussian measurements.

There also is a recent body of work studying the phase retrieval under sparsity assumptions about the signal we wish to recover, see [49, 43, 36, 35, 44] as well as the references therein. Finally, a different line of work [5, 16] studies the phase retrieval by polarization, see also [46] for a related approach. This technique comes with an algorithm that can achieve recovery using on the order of  $\log n$  specially constructed masks/codes in the noiseless case. However, to the extent of our knowledge, PhaseLift offers more flexibility in terms of the number and types of masks that can be used since it can be applied regardless of the data acquisition scheme. In addition, when dealing with noisy data PhaseLift behaves very well, see Section 2 below and the experiments in [16].

## 2 Numerical Experiments

In this section we carry out some simple numerical experiments to show how the performance of the algorithm depends on the number of measurements/masks and how the algorithm is affected by noise. To solve the optimization problems below we use Auslender and Teboulle [6] sub-gradient optimization method with a solver written in the framework provided by TFOCS [17] (The code is available online at [1]). The stopping criterion is when the Frobenius norm of the relative error of the objective between two subsequent iterations falls below  $10^{-10}$  or the number of iterations reaches 50,000, whichever occurs first. Before presenting the results we introduce the signal and measurement models we use.

### 2.1 Signal models

We consider two signal models:

- *Random low-pass signals.* Here,  $\mathbf{x}$  is given by

$$x[t] = \sum_{k=-(M/2-1)}^{M/2} (X_k + iY_k) e^{2\pi i(k-1)(t-1)/n},$$

with  $M = n/8$  and  $X_k$  and  $Y_k$  are i.i.d.  $\mathcal{N}(0, 1)$ .

- *Random Gaussian signals.* In this model,  $\mathbf{x} \in \mathbb{C}^n$  is a random complex Gaussian vector with i.i.d. entries of the form  $x[t] = X + iY$  with  $X$  and  $Y$  distributed as  $\mathcal{N}(0, 1)$ ; this can be expressed as

$$x[t] = \sum_{k=-(n/2-1)}^{n/2} (X_k + iY_k) e^{2\pi i(k-1)(t-1)/n},$$

where  $X_k$  and  $Y_k$  are i.i.d.  $\mathcal{N}(0, 1/8)$  so that the low-pass model is a ‘bandlimited’ version of this high-pass random model (variances are adjusted so that the expected power is the same).

## 2.2 Measurement models

We perform simulations based on four different kinds of measurements:

- *Gaussian measurements.* We sample  $m = nL$  random complex Gaussian vectors  $\mathbf{a}_k$  and use measurements of the form  $|\mathbf{a}_k^* \mathbf{x}|^2$ .
- *Binary modulations/codes.* We sample  $(L - 1)$  binary codes distributed as

$$d = \begin{cases} 1 & \text{with prob. } \frac{1}{2} \\ 0 & \text{with prob. } \frac{1}{2} \end{cases}$$

together with a regular diffraction pattern ( $d[t] = 1$  for all  $t$ ).

- *Ternary modulations/codes.* We sample  $(L - 1)$  ternary codes distributed as (1.10) together with a regular diffraction pattern.
- *Octanary modulations/codes.* Here, the codes are distributed as (1.9).

## 2.3 Phase transitions

We carry out some numerical experiments to show how the performance of the algorithm depends on the number of measurements/coded patterns. For this purpose we consider 50 trials. In each trial we generate a random complex vector  $\mathbf{x} \in \mathbb{C}^n$  (with  $n = 128$ ) from both signal models and gather data according to the four different measurement models above. For each trial we solve the following optimization problem

$$\min \frac{1}{2} \|\mathbf{b} - \mathcal{A}(\mathbf{X})\|_{\ell_2}^2 + \lambda \text{tr}(\mathbf{X}) \quad \text{subject to} \quad \mathbf{X} \geq \mathbf{0} \quad (2.1)$$

with  $\lambda = 10^{-3}$  (Note that the solution to (2.1) as  $\lambda$  tends to zero will equal to the optimal solution of (1.4)).

In Figure 2 we report the empirical probability of success for different signal and measurement models with different number of measurements. We declare a trial successful if the relative error of the reconstruction ( $\|\hat{\mathbf{X}} - \mathbf{x}\mathbf{x}^*\|_F / \|\mathbf{x}\mathbf{x}^*\|_F$ ) falls below  $10^{-5}$ ). These plots suggest that for the type of models studied in this paper six coded patterns are sufficient for exact recovery via convex programming.

## 2.4 Noisy measurements

We now study how the performance of the algorithm behaves in the presence of noise. We consider Poisson noise which is the usual noise model in optics. More, specifically we assume that the measurements  $\{y_k\}_{k=1}^m$  is a sequence of independent samples from the Poisson distributions  $\text{Poi}(\mu_k)$ , where  $\mu_k = |\mathbf{a}_k^* \mathbf{x}|^2$  correspond to the noiseless measurements. The Poisson log-likelihood for independent samples has the form  $\sum_k y_k \log \mu_k - \mu_k$  (up to an additive constant factor). Following a classical fitting approach we balance a maximum likelihood term with the trace norm in the relaxation (1.4):

$$\min \sum_k [\mu_k - y_k \log \mu_k] + \lambda \text{tr}(\mathbf{X}) \quad \text{subject to} \quad \boldsymbol{\mu} = \mathcal{A}(\mathbf{X}) \quad \text{and} \quad \mathbf{X} \geq \mathbf{0}.$$

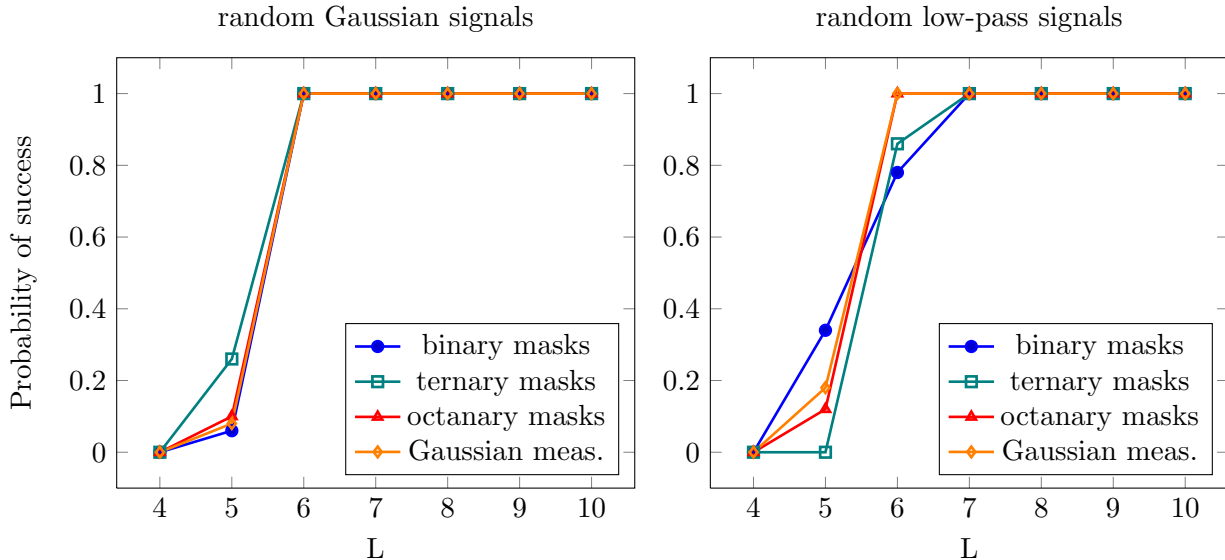


Figure 2: Empirical probability of success based on 50 random trials for different signal/measurement models and a varied number of measurements. A value of  $L$  on the x-axis means that we have a total of  $m = Ln$  samples.

The test signal is again a complex random signal sampled according to the two models described in Section 2.1. We use eight CDP's according to the three models described in Section 2.2. Poisson noise is adjusted so that the SNR levels range from 10 to 50dB. Here,  $\text{SNR} = \|\mathcal{A}(\mathbf{x}\mathbf{x}^*)\|_{\ell_2} / \|\mathbf{b} - \mathcal{A}(\mathbf{x}\mathbf{x}^*)\|_{\ell_2}$  is the signal-to-noise ratio. For the regularization parameter we use  $\lambda = 1/\text{SNR}$ . (In these experiments, the value of SNR is known. The result, however, is rather insensitive to the choice of the parameter  $\lambda$  and a good choice for the regularization parameter  $\lambda$  can be obtained by cross validation.) For each SNR level we repeat the experiment ten times with different random noise and different random CDP's.

Figure 3 shows the average relative MSE (in dB) versus the SNR (also in dB). More precisely, the values of  $10 \log_{10}(\text{rel. MSE})$  are plotted, where  $\text{rel. MSE} = \|\hat{\mathbf{X}} - \mathbf{x}\mathbf{x}^*\|_F^2 / \|\hat{\mathbf{X}}\|_F^2$ . These figures indicate that the performance of the algorithm degrades linearly as the SNR decreases (on a dB/dB scale). Empirically, the slope is close to -1, which means that the MSE scales like the noise. Together with the low offset, these features indicate that all is as in a well-conditioned-least squares problem.

### 3 Proofs

We prove our results in this section. Before we begin, we introduce some notation. We recall that the random variable  $d$  is admissible, i.e. bounded i.e.  $|d| \leq M$ , symmetric, and obeying moment constraints

$$\mathbb{E} d = 0, \quad \mathbb{E} d^2 = 0, \quad \mathbb{E} |d|^4 = 2 \mathbb{E} |d|^2. \quad (3.1)$$

Without loss of generality we also assume that  $\mathbb{E} |d|^2 = 1$ . Throughout  $\mathbf{D}$  is a diagonal matrix with i.i.d. entries distributed as  $d$ . For a vector  $\mathbf{y} \in \mathbb{C}^n$  we use  $\mathbf{y}^T$  and  $\mathbf{y}^*$  to denote the transpose and



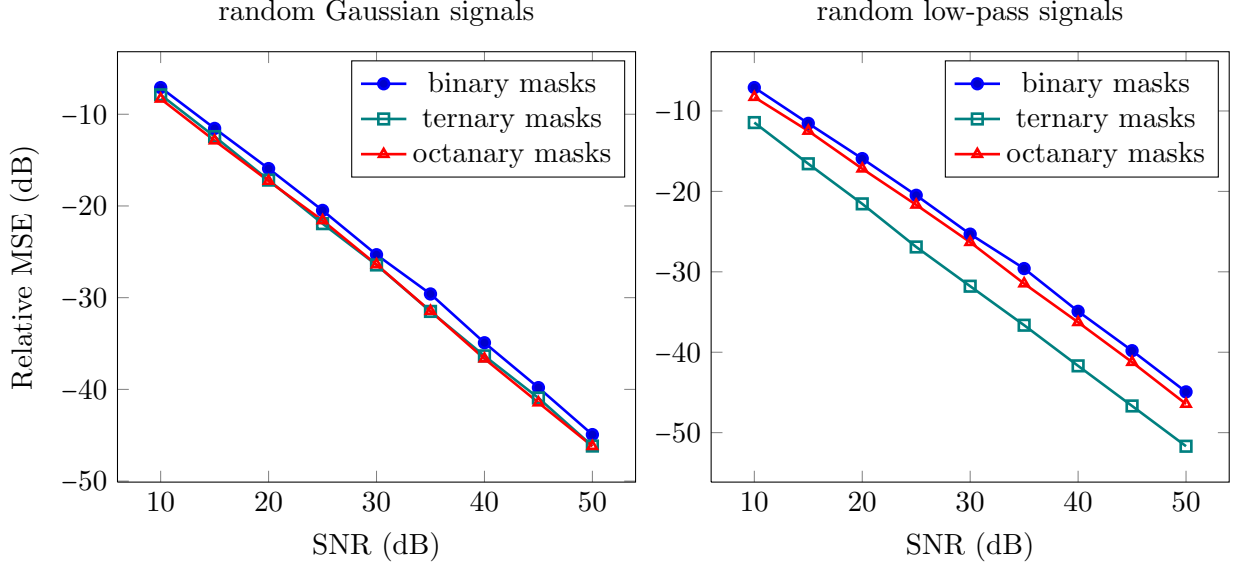


Figure 3: SNR versus relative MSE on a dB-scale for different kinds of signal/measurement models. The linear relationship between SNR and MSE (on the dB scale) is apparent. The MSE behaves as in a well-conditioned least-squares problem.

complex conjugate of the vector  $\mathbf{y}$ . We also use  $\bar{\mathbf{y}}$  to denote elementwise conjugation of the entries of  $\mathbf{y}$ . Since this is less standard, we prefer to be concrete as to avoid ambiguity: for example,

$$\begin{bmatrix} 1+i \\ 1+2i \end{bmatrix}^T = [1+i \quad 1+2i], \quad \begin{bmatrix} 1+i \\ 1+2i \end{bmatrix}^* = [1-i \quad 1-2i], \quad \overline{\begin{bmatrix} 1+i \\ 1+2i \end{bmatrix}} = \begin{bmatrix} 1-i \\ 1-2i \end{bmatrix}.$$

Continuing,  $\|\mathbf{X}\|$  is the spectral or operator norm of a matrix  $\mathbf{X}$ . Finally,  $\mathbf{1}$  is a vector with all entries equal to one.

Throughout, we assume that the fixed vector  $\mathbf{x}$  we seek to recover is unit normed, i.e.  $\|\mathbf{x}\|_{\ell_2} = 1$ . Throughout  $T$  is the linear subspace

$$T = \{\mathbf{X} = \mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^* : \mathbf{y} \in \mathbb{C}^n\}.$$

This subspace may be interpreted as the tangent space at  $\mathbf{x}\mathbf{x}^*$  to the manifold of Hermitian matrices of rank 1. Below  $T^\perp$  is the orthogonal complement to  $T$ . For a linear subspace  $V$  of Hermitian matrices, we use  $\mathbf{Y}_V$  or  $\mathcal{P}_V(\mathbf{Y})$  to denote the orthogonal projection of  $\mathbf{Y}$  onto  $V$ . With this, the reader will check that  $\mathbf{Y}_{T^\perp} = (\mathbf{I} - \mathbf{x}\mathbf{x}^*)\mathbf{Y}(\mathbf{I} - \mathbf{x}\mathbf{x}^*)$ .

### 3.1 Preliminaries

It is useful to record two identities that shall be used multiple times, and defer the proofs to the Appendix A.

**Lemma 3.1** For any fixed vector  $\mathbf{x} \in \mathbb{C}^n$

$$\mathbb{E}\left(\frac{1}{nL}\mathcal{A}^*\mathcal{A}(\mathbf{x}\mathbf{x}^*)\right) = \mathbb{E}\left(\frac{1}{n}\sum_{k=1}^n |f_k^* D^* \mathbf{x}|^2 D f_k f_k^* D^*\right) = \mathbf{x}\mathbf{x}^* + \|\mathbf{x}\|_{\ell_2}^2 \mathbf{I}.$$

**Lemma 3.2** For any fixed  $\mathbf{x} \in \mathbb{C}^n$ ,

$$\mathbb{E} \left( \frac{1}{n} \sum_{k=1}^n (\mathbf{f}_k^* \mathbf{D}^* \mathbf{x})^2 \mathbf{D} \mathbf{f}_k \mathbf{f}_k^T \mathbf{D} \right) = 2\mathbf{x}\mathbf{x}^T.$$

Next, we present two simple intermediate results we shall also use. The proofs are also in the Appendix A.

**Lemma 3.3** Fix  $\delta > 0$  and suppose the number  $L$  of CDP's obeys  $L \geq c \log n$  for some sufficiently large numerical constant  $c$ . Then with probability at least  $1 - 1/n^2$ ,

$$\left\| \frac{1}{nL} \mathcal{A}^*(\mathbf{1}) - \mathbf{I}_n \right\| \leq \delta.$$

**Lemma 3.4** For all positive semidefinite matrices  $\mathbf{X}$ , it holds

$$\frac{1}{nL} \|\mathcal{A}(\mathbf{X})\|_{\ell_1} \leq M^2 \text{tr}(\mathbf{X}).$$

Finally, the last piece of mathematics is the matrix Hoeffding inequality

**Lemma 3.5** [51, Theorem 1.3] Let  $\{\mathbf{S}_\ell\}_{\ell=1}^L$  be a sequence of independent random  $n \times n$  self-adjoint matrices. Assume that each random matrix obeys

$$\mathbb{E} \mathbf{S}_\ell = \mathbf{0} \quad \text{and} \quad \|\mathbf{S}_\ell\| \leq \Delta \quad \text{almost surely.} \quad (3.2)$$

Then for all  $t \geq 0$ ,

$$\mathbb{P} \left( \frac{1}{L} \left\| \sum_{\ell=1}^L \mathbf{S}_\ell \right\| \geq t \right) \leq 2n \exp \left( -\frac{Lt^2}{8\Delta^2} \right). \quad (3.3)$$

## 3.2 Certificates

We now establish sufficient conditions guaranteeing that  $\mathbf{x}^* \mathbf{x}$  is the unique feasible point of (1.7). Variants of the lemma below have appeared before in the literature, see [23, 27, 21].

**Lemma 3.6** Suppose the mapping  $\mathcal{A}$  obeys the following two properties:

1. For all matrices  $\mathbf{X} \in T$

$$\frac{1}{\sqrt{nL}} \|\mathcal{A}(\mathbf{X})\|_{\ell_2} \geq \frac{(1-\delta)}{\sqrt{2}} \|\mathbf{X}\|_F. \quad (3.4)$$

2. There exists a self-adjoint matrix of the form  $\mathbf{Z} = \mathcal{A}^*(\boldsymbol{\lambda})$ , with  $\boldsymbol{\lambda}$  real valued (this makes sure that  $\mathbf{Z}$  is self adjoint), obeying

$$\mathbf{Z}_{T^\perp} \leq -\mathbf{I}_{T^\perp} \quad \text{and} \quad \|\mathbf{Z}_T\|_F \leq \frac{1-\delta}{2M^2\sqrt{nL}}. \quad (3.5)$$

Then  $\mathbf{x}^* \mathbf{x}$  is the unique element in the feasible set (1.7).

**Proof** Suppose  $\mathbf{x}\mathbf{x}^* + \mathbf{H}$  is feasible. Feasibility implies that  $\mathbf{H}$  is a self-adjoint matrix in the null space of  $\mathcal{A}$  and  $\mathbf{H}_{T^\perp} \geq \mathbf{0}$ . This is because for all  $\mathbf{y} \perp \mathbf{x}$ ,

$$\mathbf{y}^*(\mathbf{x}\mathbf{x}^* + \mathbf{H})\mathbf{y} = \mathbf{y}^*\mathbf{H}\mathbf{y} \geq 0,$$

which says that  $\mathbf{H}_{T^\perp}$  is positive semidefinite. This gives

$$\langle \mathbf{H}, \mathbf{Y} \rangle = 0 = \langle \mathbf{H}_T, \mathbf{Z}_T \rangle + \langle \mathbf{H}_{T^\perp}, \mathbf{Z}_{T^\perp} \rangle.$$

On the one hand,

$$\langle \mathbf{H}_T, \mathbf{Z}_T \rangle = -\langle \mathbf{H}_{T^\perp}, \mathbf{Z}_{T^\perp} \rangle \geq \langle \mathbf{H}_{T^\perp}, \mathbf{I}_{T^\perp} \rangle = \text{tr}(\mathbf{H}_{T^\perp}). \quad (3.6)$$

Therefore,

$$\text{tr}(\mathbf{H}_{T^\perp}) \geq \frac{1}{M^2 n L} \|\mathcal{A}(\mathbf{H}_{T^\perp})\|_{\ell_1} \geq \frac{1}{M^2 n L} \|\mathcal{A}(\mathbf{H}_{T^\perp})\|_{\ell_2}.$$

where the first inequality above follows from Lemma 3.4. The injectivity property (3.4) gives

$$\frac{1}{\sqrt{nL}} \|\mathcal{A}(\mathbf{H}_T)\|_{\ell_2} \geq \frac{(1-\delta)}{\sqrt{2}} \|\mathbf{H}_T\|_F$$

and since  $\mathcal{A}(\mathbf{H}_T) = -\mathcal{A}(\mathbf{H}_{T^\perp})$ , we established

$$\text{tr}(\mathbf{H}_{T^\perp}) \geq \frac{1-\delta}{\sqrt{2nLM^2}} \|\mathbf{H}_T\|_F. \quad (3.7)$$

On the other hand,

$$|\langle \mathbf{H}_T, \mathbf{Z}_T \rangle| \leq \|\mathbf{H}_T\|_F \|\mathbf{Z}_T\|_F \leq \frac{1-\delta}{2M^2 \sqrt{nL}} \|\mathbf{H}_T\|_F. \quad (3.8)$$

In summary, (3.6), (3.7) and (3.8) assert that  $\mathbf{H}_T = \mathbf{0}$ . In turn, this gives  $\text{tr}(\mathbf{H}_{T^\perp}) = 0$  by (3.6), which implies that  $\mathbf{H}_{T^\perp} = \mathbf{0}$  since  $\mathbf{H}_{T^\perp} \geq \mathbf{0}$ . This completes the proof.  $\blacksquare$

Property (3.4) can be viewed as a form of robust injectivity of the mapping  $\mathcal{A}$  restricted to elements in  $T$ . It is of course reminiscent of the local restricted isometry property in compressive sensing. Property (3.5) can be interpreted as the existence of an approximate dual certificate. It is well known that injectivity together with an exact dual certificate leads to exact reconstruction. The above lemma essentially asserts that a robust form of injectivity together with an approximate dual certificate leads to exact recovery as in [23, Section 2.1], see also [31]. In the next two sections we show that the two properties stated in Lemma 3.6 above each hold with probability at least  $1 - 1/(2n)$ .

### 3.3 Robust injectivity

**Lemma 3.7** *Fix  $\delta > 0$  and suppose  $L$  obeys  $L \geq c \log^3 n$  for some sufficiently large numerical constant  $c$ . Then with probability at least  $1 - 1/2n$ , for all  $\mathbf{X} \in T$ ,*

$$\frac{1}{\sqrt{nL}} \|\mathcal{A}(\mathbf{X})\|_{\ell_2} \geq \frac{(1-\delta)}{\sqrt{2}} \|\mathbf{X}\|_F.$$

**Proof** First, notice that without loss of generality we can assume that  $\mathbf{x}^* \mathbf{y}$  is real valued in the definition of  $T$ . That is,

$$T = \{\mathbf{X} = \mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^* : \mathbf{y} \in \mathbb{C}^n \text{ and } \mathbf{x}^* \mathbf{y} \in \mathbb{R}\}.$$

The reason why this is true is that for any  $\mathbf{y} \in \mathbb{C}^n$ , we can find  $\lambda \in \mathbb{R}$ , such that  $\mathbf{x}^* \mathbf{y} - i\lambda \mathbf{x}^* \mathbf{x} = \mathbf{x}^* (\mathbf{y} - i\lambda \mathbf{x}) \in \mathbb{R}$  while

$$\mathbf{x}(\mathbf{y} - i\lambda \mathbf{x})^* + (\mathbf{y} - i\lambda \mathbf{x})\mathbf{x}^* = \mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^*,$$

Now for any  $\mathbf{X} = \mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^* \in T$ ,

$$\|\mathbf{X}\|_F = \|\mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^*\|_F \leq \|\mathbf{x}\mathbf{y}^*\|_F + \|\mathbf{y}\mathbf{x}^*\|_F \leq 2\|\mathbf{x}\|_{\ell_2} \|\mathbf{y}\|_{\ell_2} = 2\|\mathbf{y}\|_{\ell_2}, \quad (3.9)$$

where we recall that  $\|\mathbf{x}\|_{\ell_2} = 1$ . Hence, it suffices to show that

$$\frac{1}{\sqrt{nL}} \|\mathcal{A}(\mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^*)\|_{\ell_2} \geq \frac{(1-\delta)}{\sqrt{2}} \|\mathbf{y}\|_{\ell_2}. \quad (3.10)$$

We have

$$\|\mathcal{A}(\mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^*)\|_{\ell_2}^2 = \sum_{\ell=1}^L \sum_{k=1}^n \left( \mathbf{f}_k^* \mathbf{D}_\ell^* (\mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^*) \mathbf{D}_\ell \mathbf{f}_k \right)^2.$$

(The reader might have expected a sum of squared moduli but since  $\mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^*$  is self adjoint,  $\mathbf{f}_k^* \mathbf{D}_\ell^* (\mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^*) \mathbf{D}_\ell \mathbf{f}_k$  is real valued and so we can just as well use squares. For exposition purposes, set

$$\mathbf{A}_k(\mathbf{D}) = |\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}|^2 \mathbf{f}_k \mathbf{f}_k^*, \quad \mathbf{B}_k(\mathbf{D}) = (\mathbf{f}_k^* \mathbf{D}^* \mathbf{x})^2 \mathbf{f}_k \mathbf{f}_k^T.$$

A simple computation we omit yields

$$\begin{aligned} \left( \mathbf{f}_k^* \mathbf{D}^* \mathbf{x} \mathbf{y}^* \mathbf{D} \mathbf{f}_k + \mathbf{f}_k^* \mathbf{D}^* \mathbf{y} \mathbf{x}^* \mathbf{D} \mathbf{f}_k \right)^2 &= \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{y}} \end{bmatrix}^* \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^* \end{bmatrix} \begin{bmatrix} \mathbf{A}_k(\mathbf{D}) & \mathbf{B}_k(\mathbf{D}) \\ \mathbf{B}_k(\mathbf{D}) & \mathbf{A}_k(\mathbf{D}) \end{bmatrix} \begin{bmatrix} \mathbf{D}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{y}} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{y}} \end{bmatrix}^* \mathbf{W}_k(\mathbf{D}) \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{y}} \end{bmatrix}, \end{aligned}$$

where

$$\mathbf{W}_k(\mathbf{D}) := \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}^* \end{bmatrix} \begin{bmatrix} \mathbf{A}_k(\mathbf{D}) & \mathbf{B}_k(\mathbf{D}) \\ \mathbf{B}_k(\mathbf{D}) & \mathbf{A}_k(\mathbf{D}) \end{bmatrix} \begin{bmatrix} \mathbf{D}^* & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{bmatrix}.$$

Fix a positive threshold  $T_n$ . We now claim that (3.10) follows from

$$\frac{1}{nL} \sum_{\ell=1}^L \sum_{k=1}^n \mathbf{W}_k(\mathbf{D}_\ell) \mathbf{1}(|\mathbf{f}_k^* \mathbf{D}_\ell^* \mathbf{x}| \leq T_n) \geq \alpha \begin{bmatrix} \mathbf{x} \\ -\bar{\mathbf{x}} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ -\bar{\mathbf{x}} \end{bmatrix}^* + (1-\delta)^2 \mathbf{I}_{2n} \quad (3.11)$$

in which  $\alpha$  is any real valued number. To see why this is true, observe that

$$\begin{aligned} \frac{1}{nL} \|\mathcal{A}(\mathbf{x}\mathbf{y}^* + \mathbf{y}\mathbf{x}^*)\|_{\ell_2}^2 &\geq \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{y}} \end{bmatrix}^* \frac{1}{nL} \sum_{\ell} \sum_k \mathbf{W}_k(\mathbf{D}_\ell) \mathbf{1}(|\mathbf{f}_k^* \mathbf{D}_\ell^* \mathbf{x}| \leq T_n) \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{y}} \end{bmatrix} \\ &\geq (1-\delta)^2 \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{y}} \end{bmatrix}^* \mathbf{I}_{2n} \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{y}} \end{bmatrix} = 2(1-\delta)^2 \|\mathbf{y}\|_{\ell_2}^2. \end{aligned}$$

The last inequality comes from (3.11) together with

$$\begin{bmatrix} \mathbf{x} \\ -\bar{\mathbf{x}} \end{bmatrix}^* \begin{bmatrix} \mathbf{y} \\ \bar{\mathbf{y}} \end{bmatrix} = 0,$$

which holds since we assumed that  $\mathbf{x}^* \mathbf{y}$  is real valued.

The remainder of the proof justifies (3.11) by means of the matrix Hoeffding inequality. Let  $\langle \mathbf{W} \rangle$  be the left-hand side in (3.11) (we use notation from physics to denote empirical averages since a bar denotes complex conjugation and we would like to avoid overloading symbols). By definition  $\langle \mathbf{W} \rangle$  is the empirical average of  $L$  i.i.d. copies of

$$\mathbf{W}(\mathbf{D}) = \frac{1}{n} \sum_{k=1}^n \mathbf{W}_k(\mathbf{D}) \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| \leq T_n\}}.$$

First,  $\mathbf{W}_k(\mathbf{D}) \geq \mathbf{0}$  since

$$\begin{bmatrix} \mathbf{A}_k(\mathbf{D}) & \mathbf{B}_k(\mathbf{D}) \\ \mathbf{B}_k(\mathbf{D}) & \mathbf{A}_k(\mathbf{D}) \end{bmatrix} = \begin{bmatrix} (\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}) \mathbf{f}_k \\ (\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}) \mathbf{f}_k \end{bmatrix} \begin{bmatrix} (\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}) \mathbf{f}_k \\ (\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}) \mathbf{f}_k \end{bmatrix}^*.$$

Further,

$$\begin{aligned} \begin{bmatrix} \mathbf{A}_k(\mathbf{D}) & \mathbf{B}_k(\mathbf{D}) \\ \mathbf{B}_k(\mathbf{D}) & \mathbf{A}_k(\mathbf{D}) \end{bmatrix} &= \begin{bmatrix} 2\mathbf{A}_k(\mathbf{D}) & \mathbf{0} \\ \mathbf{0} & 2\mathbf{A}_k(\mathbf{D}) \end{bmatrix} - \begin{bmatrix} \mathbf{A}_k(\mathbf{D}) & -\mathbf{B}_k(\mathbf{D}) \\ -\mathbf{B}_k(\mathbf{D}) & \mathbf{A}_k(\mathbf{D}) \end{bmatrix} \\ &\leq \begin{bmatrix} 2\mathbf{A}_k(\mathbf{D}) & \mathbf{0} \\ \mathbf{0} & 2\mathbf{A}_k(\mathbf{D}) \end{bmatrix}. \end{aligned}$$

The inequality comes from

$$\begin{bmatrix} \mathbf{A}_k(\mathbf{D}) & -\mathbf{B}_k(\mathbf{D}) \\ -\mathbf{B}_k(\mathbf{D}) & \mathbf{A}_k(\mathbf{D}) \end{bmatrix} = \begin{bmatrix} (\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}) \mathbf{f}_k \\ -(\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}) \mathbf{f}_k \end{bmatrix} \begin{bmatrix} (\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}) \mathbf{f}_k \\ -(\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}) \mathbf{f}_k \end{bmatrix}^* \geq \mathbf{0}.$$

Hence,

$$\begin{aligned} \sum_{k=1}^n \begin{bmatrix} \mathbf{A}_k(\mathbf{D}) & \mathbf{B}_k(\mathbf{D}) \\ \mathbf{A}_k(\mathbf{D}) & \mathbf{B}_k(\mathbf{D}) \end{bmatrix} \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| \leq T_n\}} &\leq 2 \sum_{k=1}^n \begin{bmatrix} |\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}|^2 \mathbf{f}_k \mathbf{f}_k^* & \mathbf{0} \\ \mathbf{0} & |\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}|^2 \overline{\mathbf{f}_k} \mathbf{f}_k^T \end{bmatrix} \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| \leq T_n\}} \\ &\leq 2T_n^2 \sum_{k=1}^n \begin{bmatrix} \mathbf{f}_k \mathbf{f}_k^* & \mathbf{0} \\ \mathbf{0} & \overline{\mathbf{f}_k} \mathbf{f}_k^T \end{bmatrix} \\ &= 2nT_n^2 \mathbf{I}_{2n}. \end{aligned}$$

In summary,

$$\|\mathbf{W}(\mathbf{D})\| \leq 2T_n^2 \|\mathbf{D}\|^2 \leq 2M^2 T_n^2.$$

We now roughly estimate the mean of  $\mathbf{W}(\mathbf{D})$ . Obviously,

$$\begin{aligned} \mathbf{W}(\mathbf{D}) &= \frac{1}{n} \sum_k \mathbf{W}_k(\mathbf{D}) - \frac{1}{n} \sum_k \mathbf{W}_k(\mathbf{D}) \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| \leq T_n\}} \\ &:= \tilde{\mathbf{W}}(\mathbf{D}) - \frac{1}{n} \sum_k \mathbf{W}_k(\mathbf{D}) \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| \leq T_n\}}. \end{aligned}$$

By Lemmas 3.1 and 3.2, the mean of the first term ( $\tilde{\mathbf{W}}(\mathbf{D})$ ) is equal to

$$\mathbb{E} \tilde{\mathbf{W}}(\mathbf{D}) = \mathbf{I}_{2n} + \begin{bmatrix} \mathbf{x}\mathbf{x}^* & 2\mathbf{x}\mathbf{x}^T \\ 2\bar{\mathbf{x}}\mathbf{x}^* & \bar{\mathbf{x}}\mathbf{x}^T \end{bmatrix}. \quad (3.12)$$

Furthermore, a simple calculation shows that since

$$|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| \leq \|\mathbf{f}_k\|_{\ell_2} \|\mathbf{D}^* \mathbf{x}\|_{\ell_2} \leq \|\mathbf{f}_k\|_{\ell_2} \|\mathbf{D}\| \leq \sqrt{n} \|\mathbf{D}\|,$$

one can verify that

$$\|\mathbf{W}_k(\mathbf{D})\| \leq 4n^2 \|\mathbf{D}\|^4 \leq 4M^4 n^2.$$

Therefore, Jensen's inequality gives

$$\begin{aligned} \|\mathbb{E} \mathbf{W}(\mathbf{D}) - \mathbb{E} \tilde{\mathbf{W}}(\mathbf{D})\| &= \left\| \mathbb{E} \frac{1}{n} \sum_k \mathbf{W}_k(\mathbf{D}) \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| \leq T_n\}} \right\| \\ &\leq \mathbb{E} \left\| \frac{1}{n} \sum_k \mathbf{W}_k(\mathbf{D}) \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| \leq T_n\}} \right\| \\ &\leq 4M^4 n \sum_{k=1}^n \mathbb{P}(|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| > T_n). \end{aligned}$$

Setting  $T_n = \sqrt{2\beta \log n}$ , then a simple application of Hoeffding's inequality gives

$$\mathbb{P}(|\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}| > \sqrt{2\beta \log n}) \leq 2n^{-\beta}$$

(we omit the details). Therefore,

$$\|\mathbb{E} \mathbf{W}(\mathbf{D}) - \mathbb{E} \tilde{\mathbf{W}}(\mathbf{D})\| \leq \frac{8M^4}{n^{\beta-2}}. \quad (3.13)$$

Next,

$$\|\mathbf{W}(\mathbf{D}) - \mathbb{E} \mathbf{W}(\mathbf{D})\| \leq \|\mathbf{W}(\mathbf{D})\| + \|\mathbb{E} \mathbf{W}(\mathbf{D})\| \leq \|\mathbf{W}(\mathbf{D})\| + \|\mathbb{E} \tilde{\mathbf{W}}(\mathbf{D})\| + \|\mathbb{E} \mathbf{W}(\mathbf{D}) - \mathbb{E} \tilde{\mathbf{W}}(\mathbf{D})\|$$

and with  $T_n$  as above, collecting our estimates gives

$$\|\mathbf{W}(\mathbf{D}) - \mathbb{E} \mathbf{W}(\mathbf{D})\| \leq 4M^2 \beta \log n + 4 + \frac{8M^4}{n^{\beta-2}} := \Delta.$$

We have done the groundwork to apply the matrix Hoeffding inequality (3.3), which reads

$$\mathbb{P}(\|\langle \mathbf{W} \rangle - \mathbb{E} \mathbf{W}(\mathbf{D})\| \geq t) \leq 2n \exp\left(-\frac{Lt^2}{8\Delta^2}\right).$$

This implies that when  $\beta$  is sufficiently large and  $L \geq c \log^3 n$  for a sufficiently large constant,

$$\|\langle \mathbf{W} \rangle - \mathbb{E} \mathbf{W}(\mathbf{D})\| \leq \epsilon/2$$

with probability at least  $1 - 1/(2n)$ . Now from (3.12), and (3.13) gives

$$\mathbb{E} \mathbf{W}(\mathbf{D}) = \mathbf{I}_{2n} + \frac{3}{2} \begin{bmatrix} \mathbf{x} \\ \bar{\mathbf{x}} \end{bmatrix} [\mathbf{x}^*, \mathbf{x}^T] - \frac{1}{2} \begin{bmatrix} \mathbf{x} \\ -\bar{\mathbf{x}} \end{bmatrix} [\mathbf{x}^*, -\mathbf{x}^T] + \mathbf{E},$$

where (3.13) gives that  $\|\mathbf{E}\| \leq \epsilon/2$  provided  $\beta \geq 2 + \log(16M^4\epsilon^{-1})/\log n$ . Hence, we have established that

$$\langle \mathbf{W} \rangle \geq (1 - \epsilon)\mathbf{I}_{2n} + \frac{3}{2} \begin{bmatrix} \mathbf{x} \\ \bar{\mathbf{x}} \end{bmatrix} [\mathbf{x}^*, \mathbf{x}^T] - \frac{1}{2} \begin{bmatrix} \mathbf{x} \\ -\bar{\mathbf{x}} \end{bmatrix} [\mathbf{x}^*, -\mathbf{x}^T] \geq (1 - \epsilon)\mathbf{I}_{2n} - \frac{1}{2} \begin{bmatrix} \mathbf{x} \\ -\bar{\mathbf{x}} \end{bmatrix} [\mathbf{x}^*, -\mathbf{x}^T]$$

since  $\begin{bmatrix} \mathbf{x} \\ \bar{\mathbf{x}} \end{bmatrix} [\mathbf{x}^*, \mathbf{x}^T] \geq \mathbf{0}$ . With  $\epsilon = 2\delta - \delta^2$ , this is the desired conclusion (3.11). ■

### 3.4 Dual certificate construction via the golfing scheme

We now construct the approximate dual certificate  $\mathbf{Z}$  obeying the conditions of Lemma 3.6. For this purpose we use the golfing scheme first presented in the work of Gross [31]. Modifications of this technique have subsequently been used in many other papers e.g. [22, 23, 36]. The special form used here is most closely related to the construction in [36]. The mathematical validity of our construction crucially relies on the lemma below, whose proof is the object of the separate Section 3.5.

**Lemma 3.8** *Assume that  $L \geq c \log^3 n$  for a sufficiently large constant  $c$ . Then for any fixed  $\mathbf{X} \in T$ , there exists  $\mathbf{Y}$  of the form  $\mathbf{Y} = \mathcal{A}^*(\boldsymbol{\lambda})$  with  $\boldsymbol{\lambda}$  real valued such that*

$$\|\mathbf{Y} - \mathbf{X}\| \leq \frac{\sqrt{2}}{20} \|\mathbf{X}\|_F$$

holds with probability at least  $1 - 1/n^2$ . This inequality has the immediate consequences

$$\|\mathbf{Y}_T - \mathbf{X}\|_F \leq \frac{1}{5} \|\mathbf{X}\|_F, \quad \|\mathbf{Y}_{T^\perp}\| \leq \frac{\sqrt{2}}{20} \|\mathbf{X}\|_F.$$

To build our approximate dual certificate  $\mathbf{Z}$ , we partition the modulations or CDPs into  $B + 1$  different groups so that, from now on,  $\mathcal{A}_0$  corresponds to those measurements from the first  $L_0$  modulations,  $\mathcal{A}_1$  to those from the next  $L_1$  ones, and so on. Clearly,  $L_0 + L_1 + \dots + L_B = L$ . The random mappings  $\{\mathcal{A}_b\}_{b=0}^B$  correspond to independent modulations and are thus independent. Our golfing scheme starts with  $\mathbf{X}^{(0)} = \frac{2}{nL_0} \mathcal{P}_T(\mathcal{A}_0^*(\mathbf{1}))$  ( $\mathbf{1}$  is the all-one vector) and for  $b = 1, \dots, B$ , inductively defines

- $\mathbf{Y}^{(b)} \in \text{Range}(\mathcal{A}_b^*)$  obeying  $\|\mathbf{Y}^{(b)} - \mathbf{X}^{(b-1)}\| \leq \frac{\sqrt{2}}{20} \|\mathbf{X}^{(b-1)}\|_F$ ,
- and  $\mathbf{X}^{(b)} = \mathbf{X}^{(b-1)} - \mathcal{P}_T(\mathbf{Y}^{(b)})$ .

In the end, we set

$$\mathbf{Z} = \mathbf{Y} - \frac{2}{nL_0} \mathcal{A}_0^*(\mathbf{1}), \quad \mathbf{Y} = \sum_{t=1}^B \mathbf{Y}^{(t)}.$$

Note that Lemma 3.8 asserts that  $\mathbf{Y}^{(b)}$  exists with high probability, and that for each  $b$  both

$$\|\mathbf{X}^{(b)}\|_F \leq \frac{1}{5} \|\mathbf{X}^{(b-1)}\|_F \quad \text{and} \quad \|\mathbf{Y}_{T^\perp}^{(b)}\| \leq \frac{\sqrt{2}}{20} \|\mathbf{X}^{(b-1)}\|_F \quad (3.14)$$

hold on an event of probability at least  $1 - 1/n^2$ .

We now show that our construction  $\mathbf{Z}$  satisfies the required assumptions from Lemma 3.6. First,  $\mathbf{Z}$  is self-adjoint and of the form  $\mathcal{A}^*(\boldsymbol{\lambda})$  with  $\boldsymbol{\lambda} \in \mathbb{R}^{nL}$ . Second,

$$\mathbf{Z}_T = \mathbf{Y}_T - \frac{2}{nL_0} \mathcal{P}_T(\mathcal{A}_0^*(\mathbf{1})) = \sum_{b=1}^B \mathcal{P}_T(\mathbf{Y}^{(b)}) - \mathbf{X}^{(0)} = \sum_{b=1}^B (\mathbf{X}^{(b-1)} - \mathbf{X}^{(b)}) - \mathbf{X}^{(0)} = -\mathbf{X}^{(B)}.$$

Then (3.14) implies that with probability at least  $1 - B/n^2$

$$\|\mathbf{Z}_T\|_F \leq \|\mathbf{X}^{(B)}\|_F \leq \frac{1}{5^B} \|\mathbf{X}^{(0)}\|_F. \quad (3.15)$$

Also, (3.14) gives

$$\|\mathbf{Y}_{T^\perp}\| \leq \sum_{b=1}^B \|\mathbf{Y}_{T^\perp}^{(b)}\| \leq \frac{\sqrt{2}}{20} \sum_{b=1}^B \|\mathbf{X}^{(b-1)}\|_F \leq \frac{\sqrt{2}}{20} \sum_{b=1}^B \frac{1}{5^b} \|\mathbf{X}^{(0)}\|_F < \frac{\sqrt{2}}{16} \|\mathbf{X}^{(0)}\|_F \quad (3.16)$$

with probability at least  $1 - B/n^2$ . If  $L_0 \geq c \log n$  for a sufficiently large constant  $c > 0$ , Lemma 3.3 states that

$$\left\| \frac{2}{nL_0} \mathcal{A}_0^*(\mathbf{1}) - 2\mathbf{I} \right\| \leq \frac{1}{4}$$

with probability at least  $1 - 1/n^2$ . Using the fact that for any matrix  $\mathbf{W}$ , we have  $\|\mathbf{W}_T\| \leq 2\|\mathbf{W}\|$  and  $\|\mathbf{W}_{T^\perp}\| \leq \|\mathbf{W}\|$  we conclude that

$$\|\mathbf{X}^{(0)} - 2\mathbf{I}_T\| \leq 1/2, \quad \|\mathbf{Y}_{T^\perp} - \mathbf{Z}_{T^\perp} - 2\mathbf{I}_{T^\perp}\| \leq 1/4. \quad (3.17)$$

Since  $\mathbf{X}^{(0)}$  has rank at most 2,

$$\|\mathbf{X}^{(0)}\|_F \leq \sqrt{2} \|\mathbf{X}^{(0)}\| \leq \sqrt{2} \|\mathbf{X}^{(0)} - 2\mathbf{I}_T\| + 2\sqrt{2} \|\mathbf{I}_T\|$$

Finally, with (3.17) and  $\|\mathbf{I}_T\| \leq 1$ , we conclude that

$$\|\mathbf{X}^{(0)}\|_F < 4. \quad (3.18)$$

Plugging this into (3.15) we arrive at

$$\|\mathbf{Z}_T\|_F \leq \frac{4}{5^B}. \quad (3.19)$$

Also, (3.16), (3.17) and (3.18) give

$$\|\mathbf{Z}_{T^\perp} + 2\mathbf{I}_{T^\perp}\| \leq \|\mathbf{Y}_{T^\perp}\| + \|\mathbf{Y}_{T^\perp} - \mathbf{Z}_{T^\perp} - 2\mathbf{I}_{T^\perp}\| \leq \frac{\sqrt{2}}{4} + \frac{1}{4} < 1 \quad \Rightarrow \quad \mathbf{Z}_{T^\perp} \leq -\mathbf{I}_{T^\perp}. \quad (3.20)$$

Therefore, the assumptions in Lemma 3.6 hold with probability at least  $1 - 1/2n$  by applying the union bound and using with the proviso that  $B \geq c_1 \log n$  and  $L_b \geq c_2 \log^3 n$  for sufficiently large constants  $c_1$  and  $c_2$  (this is why we require  $L \geq c \log^4 n$  for a sufficiently large constant).



### 3.5 Proof of Lemma 3.8

The immediate consequences hold for the following reasons. First, since any matrix in  $T$  has rank at most 2,

$$\|\mathbf{Y}_T - \mathbf{X}\|_F \leq \sqrt{2}\|\mathbf{Y}_T - \mathbf{X}\| \leq 2\sqrt{2}\|\mathbf{Y} - \mathbf{X}\| \leq \frac{1}{5}\|\mathbf{X}\|_F,$$

where the second inequality follows from  $\|\mathbf{M}_T\| \leq 2\|\mathbf{M}\|$  for any  $\mathbf{M}$ . Second, since  $\|\mathbf{M}_{T^\perp}\| \leq \|\mathbf{M}\|$ ,

$$\|\mathbf{Y}_{T^\perp}\| = \|\mathbf{Y}_{T^\perp} - \mathbf{X}_{T^\perp}\| \leq \|\mathbf{Y} - \mathbf{X}\| \leq \frac{\sqrt{2}}{20}\|\mathbf{X}\|_F.$$

It thus suffices to prove the first property. To this end consider the eigenvalue decomposition of  $\mathbf{X} = \lambda_1 \mathbf{u}_1 \mathbf{u}_1^* + \lambda_2 \mathbf{u}_2 \mathbf{u}_2^*$ . The proof follows from Lemma 3.9 below combined with Lemma 3.3.

**Lemma 3.9** *Assume  $L \geq c \log^3 n$  for a sufficiently large constant  $c$ . Given any fixed self-adjoint matrix  $\mathbf{v} \mathbf{v}^*$ , with probability at least  $1 - 1/(2n^3)$  there exists  $\tilde{\mathbf{Y}} \in \text{Range}(\mathcal{A}^*)$  obeying*

$$\|\tilde{\mathbf{Y}} - (\mathbf{v} \mathbf{v}^* + \|\mathbf{v}\|_{\ell_2}^2 \mathbf{I})\| \leq \epsilon \|\mathbf{v}\|_{\ell_2}^2.$$

**Proof** Without loss of generality, assume  $\|\mathbf{v}\|_{\ell_2} = 1$  and set  $\tilde{\mathbf{Y}} = \langle \mathbf{Y} \rangle$ ,

$$\langle \mathbf{Y} \rangle = \frac{1}{L} \sum_{\ell=1}^L \mathbf{Y}_\ell, \quad \mathbf{Y}_\ell = \frac{1}{n} \sum_{k=1}^n |\mathbf{f}_k^* \mathbf{D}_\ell^* \mathbf{v}|^2 \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}_\ell^* \mathbf{v}| \leq T_n\}} \mathbf{D}_\ell \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}_\ell^*,$$

which is of the form  $\mathcal{A}^*(\lambda)$ . The  $\mathbf{Y}_\ell$ 's are i.i.d. copies of  $\mathbf{Y}$ ,

$$\mathbf{Y} = \frac{1}{n} \sum_{k=1}^n |\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}|^2 \mathbf{D} \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}^* - \frac{1}{n} \sum_{k=1}^n |\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}|^2 \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}| > T_n\}} \mathbf{D} \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}^*.$$

Notice that the random positive semi-definite matrix  $\mathbf{Y}$  obeys

$$\mathbf{Y} \leq \frac{1}{n} \sum_{k=1}^n T_n^2 \mathbf{D} \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}^* = T_n^2 \mathbf{D} \mathbf{D}^*.$$

By Lemma 3.1,

$$\mathbb{E} \left( \frac{1}{n} \sum_{k=1}^n |\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}|^2 \mathbf{D} \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}^* \right) = \mathbf{v} \mathbf{v}^* + \mathbf{I}.$$

Using Jensen's inequality, we have as in the proof of Lemma 3.7

$$\left\| \mathbb{E} \left( \frac{1}{n} \sum_{k=1}^n |\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}|^2 \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}| > T_n\}} \mathbf{D} \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}^* \right) \right\| \leq \frac{1}{n} \sum_{k=1}^n \mathbb{E}(\mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}| > T_n\}}) n^2 M^4. \quad (3.21)$$

Put  $T_n = \sqrt{2\beta \log n}$ . Hoeffding's inequality gives

$$\mathbb{E}(\mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}| > T_n\}}) \leq 2n^{-\beta}.$$

Plugging this into (3.21) we arrive at

$$\left\| \mathbb{E} \left( \frac{1}{n} \sum_{k=1}^n |\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}|^2 \mathbb{1}_{\{|\mathbf{f}_k^* \mathbf{D}^* \mathbf{v}| > T_n\}} \mathbf{D} \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}^* \right) \right\| \leq \frac{2M^4}{n^{\beta-2}}.$$

For sufficiently large  $\beta$ , this implies

$$\|\mathbb{E}(\mathbf{Y}) - (\mathbf{v}\mathbf{v}^* + \mathbf{I})\| \leq \frac{2\|\mathbf{D}\|^4}{n^{\beta-2}} \leq \frac{\epsilon}{2}. \quad (3.22)$$

By using Hoeffding inequality in a similar fashion as in the proof of Lemma 3.7, we obtain (we omit the details)

$$\|\langle \mathbf{Y} \rangle - \mathbb{E}(\mathbf{Y})\| \leq \frac{\epsilon}{2}.$$

Combining the latter with (3.22), we conclude

$$\|\langle \mathbf{Y} \rangle - (\mathbf{v}\mathbf{v}^* + \mathbf{I})\| \leq \epsilon. \quad \blacksquare$$

## 4 Discussion

In this paper, we proved that a signal could be recovered by convex programming techniques from a few diffraction patterns corresponding to generic modulations obeying an admissibility condition. We expect that our results, methods and proofs extend to more general random modulations although we have not pursued such extensions in this paper. Further, we proved that on the order of  $(\log n)^4$  CDPs suffice for perfect recovery and we expect that further refinements would allow to reduce this number, perhaps all the way down to a figure independent of the number  $n$  of unknowns. Such refinements appear quite involved to us and since our intention is to provide a reasonably short and conceptually simple argument, we leave such refinements to future research.

## A Proof of auxiliary lemmas

Set  $\omega = e^{\frac{2\pi i}{n}}$  to be the  $n$ th root of unity so that

$$\mathbf{f}_k^* = [\omega^{-0(k-1)}, \omega^{-1(k-1)}, \dots, \omega^{-(n-1)(k-1)}], \quad \mathbf{f}_k = \begin{bmatrix} \omega^{0(k-1)} \\ \omega^{1(k-1)} \\ \vdots \\ \omega^{(n-1)(k-1)} \end{bmatrix}.$$

For two integers  $a$  and  $b$  we use  $a \stackrel{n}{\equiv} b$  to denote congruence of  $a$  and  $b$  modulo  $n$  ( $n$  divides  $a - b$ ).

### A.1 Proof of Lemma 3.1

Put

$$\mathbf{Y} := \frac{1}{n} \sum_{k=1}^n |\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}|^2 \mathbf{D} \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}^*.$$

By definition,

$$\begin{aligned} |\mathbf{f}_k^* \mathbf{D}^* \mathbf{x}|^2 &= \left( \sum_{a=1}^n \bar{d}_a x_a \omega^{-(a-1)(k-1)} \right) \left( \sum_{b=1}^n d_b \bar{x}_b \omega^{(b-1)(k-1)} \right) \\ &= \sum_{a=1}^n \sum_{b=1}^n \omega^{(b-a)(k-1)} \bar{d}_a d_b x_a \bar{x}_b \end{aligned}$$

Further,

$$\begin{aligned}
\mathbf{Y}_{pq} &= \frac{1}{n} \sum_{k=1}^n \sum_{a=1}^n \sum_{b=1}^n \omega^{(b-a+p-q)(k-1)} \bar{d}_a d_b d_p \bar{d}_q x_a \bar{x}_b \\
&= \sum_{a=1}^n \sum_{b=1}^n \bar{d}_a d_b d_p \bar{d}_q x_a \bar{x}_b \left( \frac{1}{n} \sum_{k=1}^n \omega^{(b-a+p-q)(k-1)} \right) \\
&= \sum_{a=1}^n \sum_{b=1}^n \bar{d}_a d_b d_p \bar{d}_q x_a \bar{x}_b \mathbb{1}_{\{a+q \stackrel{n}{=} b+p\}}.
\end{aligned}$$

Therefore,

$$\mathbb{E}[\mathbf{Y}_{pq}] = \sum_{a=1}^n \sum_{b=1}^n \mathbb{E}[\bar{d}_a d_b d_p \bar{d}_q] x_a \bar{x}_b \mathbb{1}_{\{a+q \stackrel{n}{=} b+p\}}.$$

- Diagonal terms ( $p = q$ ): Here,  $\mathbb{E}[\bar{d}_a d_b |d_p|^2] = 0$  unless  $a = b$ . This gives

$$\begin{aligned}
\mathbb{E}[\mathbf{Y}_{pp}] &= \sum_{a=1}^n \mathbb{E}[|d_a|^2 |d_p|^2] |x_a|^2 \\
&= \mathbb{E}[|d_p|^4] |x_p|^2 + \mathbb{E}[|d_p|^2 \left( \sum_{a \neq p}^n |d_a|^2 |x_a|^2 \right)] \\
&= |x_p|^2 + \|\mathbf{x}\|_{\ell_2}^2.
\end{aligned}$$

- Off-diagonal terms ( $p \neq q$ ): Here  $\mathbb{E}[\bar{d}_a d_b d_p \bar{d}_q] = 0$  unless  $(a = p, b = q)$  so that

$$\mathbb{E}[\mathbf{Y}_{pq}] = (\mathbb{E}[|d|^2])^2 x_p \bar{x}_q = x_p \bar{x}_q.$$

This concludes the proof.

## A.2 Proof of Lemma 3.2

Put

$$\mathbf{R} = \frac{1}{n} \sum_{k=1}^n (\mathbf{f}_k^* \mathbf{D}^* \mathbf{x})^2 \mathbf{D} \mathbf{f}_k \mathbf{f}_k^T \mathbf{D}.$$

By definition,

$$(\mathbf{f}_k^* \mathbf{D}^* \mathbf{x})^2 = \sum_{a=1}^n \sum_{b=1}^n \omega^{-(a+b-2)(k-1)} \bar{d}_a \bar{d}_b x_a x_b$$

and

$$\begin{aligned}
\mathbf{R}_{pq} &= \frac{1}{n} \sum_{k=1}^n \sum_{a=1}^n \sum_{b=1}^n \omega^{(p+q-a-b)(k-1)} \bar{d}_a \bar{d}_b d_p d_q x_a x_b \\
&= \sum_{a=1}^n \sum_{b=1}^n \bar{d}_a \bar{d}_b d_p d_q x_a x_b \left( \frac{1}{n} \sum_{k=1}^n \omega^{(p+q-a-b)(k-1)} \right) \\
&= \sum_{a=1}^n \sum_{b=1}^n \bar{d}_a \bar{d}_b d_p d_q x_a x_b \mathbb{1}_{\{p+q \stackrel{n}{=} a+b\}}.
\end{aligned}$$

Therefore,

$$\mathbb{E}[\mathbf{R}_{pq}] = \sum_{a=1}^n \sum_{b=1}^n \mathbb{E}[\bar{d}_a \bar{d}_b d_p d_q] x_a x_b \mathbb{1}_{\{p+q \equiv a+b\}}.$$

- Diagonal terms ( $p = q$ ): Here,  $\mathbb{E}[\bar{d}_a d_b |d_p|^2] = 0$  unless  $a = b = p$ . This gives

$$\mathbb{E}[\mathbf{R}_{pp}] = \mathbb{E}[|d|^4] x_p^2 = 2x_p^2.$$

- Off-diagonal terms ( $p \neq q$ ): Here,  $\mathbb{E}[\bar{d}_a \bar{d}_b d_p d_q] = 0$  unless  $(a = p, b = q)$  or  $(a = q, b = p)$ . This gives

$$\mathbb{E}[\mathbf{R}_{pq}] = 2\mathbb{E}[|d_p|^2 |d_q|^2] x_p x_q = 2x_p x_q.$$

This concludes the proof.

### A.3 Proof of Lemma 3.3

Note that

$$\mathbf{Z} := \frac{1}{nL} \mathcal{A}^*(\mathbf{1}) = \frac{1}{nL} \sum_{\ell=1}^L \sum_{k=1}^n \mathbf{D}_\ell \mathbf{f}_k \mathbf{f}_k^* \mathbf{D}_\ell^* = \frac{1}{L} \sum_{\ell=1}^L \mathbf{D}_\ell \mathbf{D}_\ell^*.$$

Therefore,  $\mathbf{Z}$  is a diagonal matrix with i.i.d. diagonal entries distributed as  $\frac{1}{L} \sum_{\ell=1}^L X_\ell$ , where the  $X_\ell$  are i.i.d. random variables with  $\mathbb{E}[X_\ell] = \mathbb{E}[|d|^2] = 1$  and  $|X_\ell| = |d|^2 \leq M^2$ . The statement in the lemma then follows from Hoeffding's inequality

$$\mathbb{P}\left\{\left|\frac{1}{L} \sum_{\ell=1}^L X_\ell - 1\right| \geq t\right\} \leq 2e^{-\frac{2L}{M^2} t^2}$$

combined with the union bound.

### A.4 Proof of Lemma 3.4

The proof is straightforward and parallels calculations in [23]. Fix a unit-normed vector  $\mathbf{v}$ , then

$$\|\mathcal{A}(\mathbf{v}\mathbf{v}^*)\|_{\ell_1} = \sum_{\ell=1}^L \sum_{k=1}^n |\mathbf{f}_k^* \mathbf{D}_\ell^* \mathbf{v}|^2 = n \sum_{\ell=1}^L \|\mathbf{D}_\ell^* \mathbf{v}\|_{\ell_2}^2 \leq nM^2 \sum_{\ell=1}^L \|\mathbf{v}\|_{\ell_2}^2 = nLM^2.$$

Consider now the eigenvalue decomposition  $\mathbf{X} = \sum_{j=1}^n \lambda_j \mathbf{v}_j \mathbf{v}_j^*$  where  $\lambda_j$  is nonnegative since  $\mathbf{X} \geq \mathbf{0}$ . Then

$$\|\mathcal{A}(\mathbf{X})\|_{\ell_1} = \sum_{j=1}^n \lambda_j \|\mathcal{A}(\mathbf{v}_j \mathbf{v}_j^*)\|_{\ell_1} \leq nLM^2 \sum_j \lambda_j = nLM^2 \text{tr}(\mathbf{X}).$$

## B Extensions to higher dimensions by tensorization

Consider a two dimensional signal (image)  $\mathbf{x} \in \mathbb{R}^{n_1 \times n_2}$  and assume that we have one dimensional admissible masks of the form  $\mathbf{d} \in \mathbb{R}^{n_1}$  and  $\mathbf{b} \in \mathbb{R}^{n_2}$ .<sup>2</sup> We will use  $\mathbf{B}$  and  $\mathbf{D}$  to denote diagonal matrices with  $\mathbf{b}$  and  $\mathbf{d}$  on the diagonal. Then we will use 2D masks of the form  $\mathbf{d}\mathbf{1}^*$  and  $\mathbf{1}\mathbf{b}^*$ , where  $\mathbf{1}$  denotes the all one vector. Examples of such masks are depicted in Figure 4. Using this form of masks our measurements will be of the form

$$y_{r,s,k}^{(1)} = |\mathbf{f}_{r,n_1}^* \mathbf{x} \mathbf{B}_k \mathbf{f}_{s,n_2}|^2 \quad \text{and} \quad y_{r,s,\ell}^{(2)} = |\mathbf{f}_{r,n_1}^* \mathbf{D}_\ell^* \mathbf{x} \mathbf{f}_{s,n_2}|^2$$

for  $k = 1, 2, \dots, L_1$ ,  $\ell = 1, 2, \dots, L_2$ ,  $r = 1, 2, \dots, n_1$  and  $s = 1, 2, \dots, n_2$  with  $\mathbf{f}_{k,n}^*$  denoting the rows of  $\mathbf{F}_n$  the DFT matrix of  $\mathbb{R}^n$ . Now set  $\mathbf{x}^{(1)} = \mathbf{F}_{n_1} \mathbf{x}$  and  $\mathbf{x}^{(2)} = \mathbf{x} \mathbf{F}_{n_2}^*$ . It is easy to see that we can recover the rows of  $\mathbf{x}^{(1)}$  and the columns of  $\mathbf{x}^{(2)}$  up to a global phase with high probability by applying the SDP feasibility problem (1.7) separately for each row/column (as long as  $L_1 \geq c_1 \log^4 n_1$  and  $L_2 \geq c_2 \log^4 n_2$  of course). Therefore we arrive at matrices of the form

$$\begin{aligned} \tilde{\mathbf{x}}^{(1)} &= \text{diag}(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_{n_1}}) \mathbf{x}^{(1)}, \\ \tilde{\mathbf{x}}^{(2)} &= \mathbf{x}^{(2)} \text{diag}(e^{i\phi_1}, e^{i\phi_2}, \dots, e^{i\phi_{n_2}}). \end{aligned}$$

Using  $\hat{\mathbf{x}} = \mathcal{F}(\mathbf{x})$  to denote the two dimensional Discrete Fourier transform of  $\mathbf{x}$ . Then we have

$$\begin{aligned} \mathbf{z}^{(1)} &:= \tilde{\mathbf{x}}^{(1)} \mathbf{F}_{n_2}^* = \text{diag}(e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_{n_1}}) \hat{\mathbf{x}}, \\ \mathbf{z}^{(2)} &:= \mathbf{F}_{n_1} \tilde{\mathbf{x}}^{(2)} = \hat{\mathbf{x}} \text{diag}(e^{i\phi_1}, e^{i\phi_2}, \dots, e^{i\phi_{n_2}}). \end{aligned} \tag{B.1}$$

Define the bipartite graph of a matrix  $\mathbf{X} = [X_{ab}] \in \mathbb{R}^{n_1 \times n_2}$  as the bipartite graph with vertices  $1, 2, \dots, n_1$  (representing the rows) and  $1, 2, \dots, n_2$  (representing the columns), such that there is an edge joining node  $a$  to node  $b$  if and only if  $X_{ab} \neq 0$ . Let  $a_1, a_2, \dots, a_{n_1}$  denote the nodes corresponding to the rows of  $\hat{\mathbf{x}}$  and  $b_1, b_2, \dots, b_{n_2}$  denote the nodes corresponding to the columns of  $\hat{\mathbf{x}}$ . We assume that the bipartite graph of  $\hat{\mathbf{x}}$  is connected. Then for any node  $b_r$  representing a column of  $\hat{\mathbf{x}}$  we have a path connecting node  $a_1$  to node  $b_j$ . That is, a path of the form  $(a_{i_1}, b_{j_1}), (b_{j_1}, a_{i_2}), \dots, (a_{i_p}, b_{j_p})$  with  $i_1 = 1$  and  $j_p = j$ . Note that when  $\hat{\mathbf{x}}_{ab} \neq 0$  then we have  $\mathbf{z}_{ab}^{(1)} / \mathbf{z}_{ab}^{(2)} = e^{i\theta_a} / e^{i\phi_b}$  so that if we know  $\theta_a$  we can deduce  $\phi_b$ . Applying this argument along the path from  $a_1$  to  $b_j$ , given  $\theta_1$  we can then find  $\phi_j$ . A similar argument holds for any node  $a_i$ . Therefore when the bipartite graph of  $\hat{\mathbf{x}}$  is connected, fixing  $\theta_1$ , we can uniquely determine all parameters  $\theta_2, \theta_3, \dots, \theta_{n_1}$  and  $\phi_1, \phi_2, \dots, \phi_{n_2}$ . As a result it is easy to see that as long as the bipartite graph associated with  $\hat{\mathbf{x}}$  is connected one can recover  $\hat{\mathbf{x}}$  and therefore  $\mathbf{x}$  up to a global phase factor from the two equations in (B.1). We note that in the above  $L = L_1 + L_2 \geq c_1 \log^4 n_1 + c_2 \log^4 n_2$  measurements are sufficient. One can also apply a similar argument to show that  $(\log n_1)^4 \times (\log n_2)^4$  masks of the form  $\mathbf{b}\mathbf{d}^*$  are sufficient without the technical assumption stated above. However, as mentioned before we should emphasize that our methods are directly applicable to the 2D case but we will not pursue this in this paper.

### Acknowledgements

E C. is partially supported by AFOSR under grant FA9550-09-1-0643, by ONR under grant N00014-09-1-0258 and by a gift from the Broadcom Foundation. M. S. is supported by a a Benchmark

<sup>2</sup>The diagonal masks we discussed before would have  $\mathbf{d}$  on the diagonal

$d_1$	$d_1$	$d_1$	$d_1$	$d_1$	$d_1$	$d_1$	$d_1$
$d_2$	$d_2$	$d_2$	$d_2$	$d_2$	$d_2$	$d_2$	$d_2$
$d_3$	$d_3$	$d_3$	$d_3$	$d_3$	$d_3$	$d_3$	$d_3$
$d_4$	$d_4$	$d_4$	$d_4$	$d_4$	$d_4$	$d_4$	$d_4$
$d_5$	$d_5$	$d_5$	$d_5$	$d_5$	$d_5$	$d_5$	$d_5$

$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$b_6$	$b_7$	$b_8$
$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$b_6$	$b_7$	$b_8$
$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$b_6$	$b_7$	$b_8$
$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$b_6$	$b_7$	$b_8$
$b_1$	$b_2$	$b_3$	$b_4$	$b_5$	$b_6$	$b_7$	$b_8$

Figure 4: An example of constructing an admissible two dimensional make of size  $5 \times 8$  of the form  $\mathbf{d}\mathbf{1}^*$  and  $\mathbf{1b}^*$  using one dimensional admissible masks of size 5 and 8.

Stanford Graduate Fellowship. X. L. is supported by the Wharton Dean’s Fund for Post-Doctoral Research and by funding from the National Institutes of Health. We would like to thank V. Voroninski for helpful discussions, especially for bringing to our attention the difficulty of establishing RIP-1 results in the masked Fourier model. M. S. thanks David Brady and Adam Backer for fruitful discussions about the implementation of structured illuminations in X-ray crystallography and microscopy applications.

## References

- [1] [www.stanford.edu/~mahdisol/code](http://www.stanford.edu/~mahdisol/code).
- [2] D. G. Mixon blog: Saving phase: Injectivity and stability for phase retrieval.
- [3] D. G. Mixon blog: AIM workshop: frame theory intersects geometry.
- [4] A. Ahmed, B. Recht, and J. Romberg. Blind deconvolution using convex programming. *arXiv preprint arXiv:1211.5608*, 2012.
- [5] B. Alexeev, A. S. Bandeira, M. Fickus, and D. G. Mixon. Phase retrieval with polarization. *arXiv preprint arXiv:1210.7752*, 2012.
- [6] A. Auslender and M. Teboulle. Interior gradient and proximal methods for convex and conic optimization. *SIAM Journal on Optimization*, 16(3):697–725, 2006.
- [7] R. Balan. A nonlinear reconstruction algorithm from absolute value of frame coefficients for low redundancy frames. In *International Conference on Sampling Theory and Applications (SAMP TA)*, 2009.
- [8] R. Balan. On signal reconstruction from its spectrogram. In *44th Annual Conference on Information Sciences and Systems (CISS)*, pages 1–4, 2010.
- [9] R. Balan. Reconstruction of signals from magnitudes of redundant representations. *arXiv preprint arXiv:1207.1134*, 2012.
- [10] R. Balan. Stability of phase retrievable frames. *arXiv preprint arXiv:1308.5465*, 2013.
- [11] 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.

- [12] R. Balan, P. Casazza, and D. Edidin. On signal reconstruction without phase. *Applied and Computational Harmonic Analysis*, 20(3):345–356, 2006.
- [13] R. Balan, P. Casazza, and D. Edidin. Equivalence of reconstruction from the absolute value of the frame coefficients to a sparse representation problem. *IEEE Signal Processing Letters*, 14(5):341–343, 2007.
- [14] R. Balan and Y. Wang. Invertibility and robustness of phaseless reconstruction. *arXiv preprint arXiv:1308.4718*, 2013.
- [15] A. S. Bandeira, J. Cahill, D. G. Mixon, and A. A. Nelson. Saving phase: Injectivity and stability for phase retrieval. *arXiv preprint arXiv:1302.4618*, 2013.
- [16] A. S. Bandeira, Y. Chen, and D. G. Mixon. Phase retrieval from power spectra of masked signals. *arXiv preprint arXiv:1303.4458*, 2013.
- [17] S. R. Becker, E. J. Candes, and M. C. Grant. Templates for convex cone problems with applications to sparse signal recovery. *Mathematical Programming Computation*, 3(3):165–218, 2011.
- [18] O. Bunk, A. Diaz, F. Pfeiffer, C. David, B. Schmitt, D. K. Satapathy, and J. F. Veen. Diffractive imaging for periodic samples: retrieving one-dimensional concentration profiles across microfluidic channels. *Acta Crystallographica Section A: Foundations of Crystallography*, 63(4):306–314, 2007.
- [19] J. Cahill, P. G. Casazza, J. Peterson, and L. Woodland. *arXiv preprint arXiv:1305.6226*.
- [20] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski. Phase retrieval via matrix completion. *SIAM Journal on Imaging Sciences*, 6(1):199–225, 2013.
- [21] E. J. Candes and X. Li. Solving quadratic equations via phaselift when there are about as many equations as unknowns. *Foundations of Computational Mathematics*, pages 1–10, 2012.
- [22] E. J. Candes, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of the ACM (JACM)*, 58(3):11, 2011.
- [23] 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*, 2012.
- [24] A. Chai, M. Moscoso, and G. Papanicolaou. Array imaging using intensity-only measurements. *Inverse Problems*, 27(1), 2011.
- [25] J. V. Corbett. The pauli problem, state reconstruction and quantum-real numbers. *Reports on Mathematical Physics*, 57(1):53–68, 2006.
- [26] J. C. Dainty and J. R. Fienup. Phase retrieval and image reconstruction for astronomy. *Image Recovery: Theory and Application*, ed. by H. Stark, Academic Press, San Diego, pages 231–275, 1987.
- [27] L. Demanet and P. Hand. Stable optimizationless recovery from phaseless linear measurements. *arXiv preprint arXiv:1208.1803*, 2012.
- [28] L. Demanet and V. Jugnon. Convex recovery from interferometric measurements. *arXiv preprint arXiv:1307.6864*, 2013.
- [29] Y. C. Eldar and S. Mendelson. Phase retrieval: Stability and recovery guarantees. *arXiv preprint arXiv:1211.0872*, 2012.
- [30] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.
- [31] D. Gross. Recovering low-rank matrices from few coefficients in any basis. *IEEE Transactions on Information Theory*, 57(3):1548–1566, 2011.

- [32] D. Gross, F. Kraemer, and R. Kueng. A partial derandomization of phaselift using spherical designs. *arXiv preprint arXiv:1310.2267*, 2013.
- [33] R. W. Harrison. Phase problem in crystallography. *JOSA A*, 10(5):1046–1055, 1993.
- [34] T. Heinosaari, L. Mazzarella, and M. M. Wolf. Quantum tomography under prior information. *Communications in Mathematical Physics*, 318(2):355–374, 2013.
- [35] K. Jaganathan, S. Oymak, and B. Hassibi. On robust phase retrieval for sparse signals. In *50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 794–799, 2012.
- [36] X. Li and V. Voroninski. Sparse signal recovery from quadratic measurements via convex programming. *SIAM J. Math. Anal.*, 45(5):3019–3033, 2013.
- [37] E. G. Loewen and E. Popov. *Diffraction gratings and applications*. CRC Press, 1997.
- [38] J. Miao, T. Ishikawa, Q. Shen, and T. Earnest. Extending x-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes. *Annu. Rev. Phys. Chem.*, 59:387–410, 2008.
- [39] R. P. Millane. Phase retrieval in crystallography and optics. *JOSA A*, 7(3):394–411, 1990.
- [40] D. Mondragon and V. Voroninski. Determination of all pure quantum states from a minimal number of observables. *arXiv preprint arXiv:1306.1214*, 2013.
- [41] A. Nemirovski. Lectures on modern convex optimization. In *Society for Industrial and Applied Mathematics (SIAM)*. Citeseer, 2001.
- [42] P. Netrapalli, P. Jain, and S. Sanghavi. Phase retrieval using alternating minimization. *arXiv preprint arXiv:1306.0160*, 2013.
- [43] H. Ohlsson, A. Y. Yang, R. Dong, and S. S. Sastry. Compressive phase retrieval from squared output measurements via semidefinite programming. *arXiv preprint arXiv:1111.6323*, 2011.
- [44] S. Oymak, A. Jalali, M. Fazel, Y. C. Eldar, and B. Hassibi. Simultaneously structured models with application to sparse and low-rank matrices. *arXiv preprint arXiv:1212.3753*, 2012.
- [45] J. Ranieri, A. Chebira, Y. M. Lu, and M. Vetterli. Phase retrieval for sparse signals: Uniqueness conditions. *arXiv preprint arXiv:1308.3058*, 2013.
- [46] O. Raz, N. Dudovich, and B. Nadler. Vectorial phase retrieval of 1-d signals. *IEEE Transactions on Signal Processing*, 61(7):1632–1643, 2013.
- [47] H. Reichenbach. *Philosophic foundations of quantum mechanics*. University of California Pr, 1965.
- [48] J. M. Rodenburg. Ptychography and related diffractive imaging methods. *Advances in Imaging and Electron Physics*, 150:87–184, 2008.
- [49] Y. Shechtman, Y. C. Eldar, A. Szameit, and M. Segev. Sparsity based sub-wavelength imaging with partially incoherent light via quadratic compressed sensing. *Optics Express*, 19(16):14807–14822, 2011.
- [50] P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer. Probe retrieval in ptychographic coherent diffractive imaging. *Ultramicroscopy*, 109(4):338–343, 2009.
- [51] J. A. Tropp. User-friendly tail bounds for sums of random matrices. *Foundations of Computational Mathematics*, 12(4):389–434, 2012.
- [52] Z. Uo, W. Ma, A. C. So, Y. Ye, and S. Zhang. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Processing Magazine*, 27(3):20–34, 2010.
- [53] I. Waldspurger, A. d’Aspremont, and S. Mallat. Phase recovery, maxcut and complex semidefinite programming. *arXiv preprint arXiv:1206.0102*, 2012.
- [54] A. Walther. The question of phase retrieval in optics. *Journal of Modern Optics*, 10(1):41–49, 1963.
- [55] F. H. C. Watson, J. D. and Crick. A structure for deoxyribose nucleic acid. *Nature*, 171, 1953.