Near-Optimal Phase Retrieval of Sparse Vectors

Afonso S. Bandeira\textsuperscript{a} and Dustin G. Mixon\textsuperscript{b}

\textsuperscript{a}Program in Applied and Computational Mathematics (PACM), Princeton University, Princeton, NJ 08544, USA; 
\textsuperscript{b}Department of Mathematics and Statistics, Air Force Institute of Technology, Wright-Patterson Air Force Base, OH 45433, USA.

ABSTRACT

In many areas of imaging science, it is difficult to measure the phase of linear measurements. As such, one often wishes to reconstruct a signal from intensity measurements, that is, perform phase retrieval. In several applications the signal in question is believed to be sparse. In this paper, we use ideas from the recently developed polarization method for phase retrieval and provide an algorithm that is guaranteed to recover a sparse signal from a number of phaseless linear measurements that scales linearly with the sparsity of the signal (up to logarithmic factors). This is particularly remarkable since it is known that a certain popular class of convex methods is not able to perform recovery unless the number of measurements scales with the square of the sparsity of the signal. This is a shorter version of a more complete publication that will appear elsewhere.\\

Keywords: Phase Retrieval, Sparse Recovery, Polarization, Angular Synchronization

1. INTRODUCTION

In many areas of imaging science, it is difficult to measure the phase of linear measurements. This motivates the use of absolute values (squared) of linear measurements, called intensity measurements. Formally, given a set of measurement vectors $\Phi = \{\phi_i\}_{i=1}^N \subseteq \mathbb{C}^M$ and a signal $x \in \mathbb{C}^M$ we consider measurements of the form
\[
  z_\ell := |\langle x, \phi_\ell \rangle|^2 + \nu_\ell,
\] (1)

where $\nu_\ell$ is noise; we call these noisy intensity measurements.

Phase retrieval is the problem of recovering a signal from measurements of the form (1) up to a global phase factor, as there is a trivial ambiguity of multiplying $x$ by a unit-modulus complex number. This problem plays an important role in areas such as X-ray crystallography,\textsuperscript{2–4} diffraction imaging,\textsuperscript{5} astronomy\textsuperscript{6} and optics.\textsuperscript{7}

Many recent interesting theoretical work has been trying to understand for which measurement systems $\Phi$ intensity measurements are injective or even stable, potentially allowing stable recovery from noisy intensity measurements of the form (1). In 2006 it was shown\textsuperscript{8} that generic measurement systems with $N \geq 4M - 2$ are injective and it is now conjectured that the same should be true for $N \geq 4M - 4$ and that, moreover, no measurement system with $N < 4M - 4$ is injective.\textsuperscript{9} Recent work has also been done on stability\textsuperscript{9,10} guaranteeing stability, on the real case, for $N = O(M)$.

Classical algorithms\textsuperscript{11} used to tackle this problem are often based on alternating projection ideas, trying to find $y = \Phi^*x$ by alternating between having $y$ satisfy the intensity measurements and belonging to the column space of $\Phi^*$. Most of these methods lack guarantees and often have problems with local minima.

Based on the fact that intensity measurements (1) can be written as linear measurements of a lifted version of the signal $X = xx^*$, Candes, Strohmer and Voroninski\textsuperscript{12} proposed PhaseLift, a method that is able to stably perform phase retrieval, in polynomial time, from only $N = O(M \log M)$ gaussian measurement vectors. This result has later been refined to $N = O(M)$\textsuperscript{13} and a few similar alternatives have been proposed.\textsuperscript{14–16} These

To appear in: Wavelets and Sparsity XV, Proceedings of SPIE Optics+Photonics, 2013.

Further author information: (Send correspondence to ASB) ASB: ajsb@math.princeton.edu DGM: dustin.mixon@afit.edu
methods are however based on Semidefinite Programming, which although solvable in polynomial time, is still rather computationally expensive and not sufficiently efficient for many applications.

With the objective of performing phase retrieval in a more computationally efficient manner, an alternative method was recently proposed, the polarization method,\(^{16}\) which is able to stably perform phase retrieval from \(N = O(M \log M)\) design measurements without the Semidefinite Programming overhead computational cost. In fact, its computational cost is essentially the same as the one of solving the linear system in case one did have access to the phases of the measurements. While the measurement vectors for the polarization method do need to be designed there is significant flexibility on its design, and evidence of this is the application of the polarization method to reconstruct a signal from the power spectrum of masked versions of it.\(^{17}\) As it will become clear below, we will adapt the polarization method to the reconstruction of sparse vectors.

Let us assume that \(x \in \mathbb{C}^M\) is \(k\)-sparse, meaning that it has, at most, \(k\) non-zero entries. The seminal papers of Donoho, Candes, Tao, and Romberg\(^{15,23}\) introduced Compressed Sensing, which essentially is the idea that one can reconstruct \(x\) from \(N \ll M\) linear measurements provided the measurements satisfy certain properties, probably the most popular of which being the Restricted Isometry Property\(^{23}\). Remarkably, efficient and stable recovery, via \(\ell_1\) minimization, of \(k\)-sparse vectors is possible for values of \(N\) as small as \(N = O(k \log \frac{M}{k})\).

This paper is concerned with a hybrid of the Compressed Sensing and Phase Retrieval. We are interested in recovering a \(k\)-sparse signal \(x \in \mathbb{C}^M\) from noisy intensity measurements of the form \((1)\). This problem was introduced at least as early as 2007 by the name of “Compressive Phase Retrieval”\(^{24}\).

**Problem 1.1 (The Sparse Phase Retrieval problem).** Given a set of measurement vectors \(\Phi = \{\varphi_i\}_{i=1}^N \subseteq \mathbb{C}^M\) reconstruct a \(k\)-sparse signal \(x \in \mathbb{C}^M\) from noisy intensity measurements of the form \((1)\).

More recently, the PhaseLift method was adapted to this variant of the Phase Retrieval problem.\(^{25}\) Shortly after is was shown that this method succeeds for \(N = \Omega(k^2 \log(M))\).\(^{25}\) Interestingly, this is rather tight since a certain class of PhaseLift-like problems was shown to fail for \(N = O(k^2 \log^{-2}(M))\).\(^{26}\) This contrasts with the fact that the intensity measurement process is known to, at least in the real case, be injective and stable for \(N\) as small as \(O(k \log \frac{M}{k})\).\(^{16}\)

The main contribution of this paper is to fill this gap and present an efficient algorithm that is able to stably reconstruct any \(k\)-sparse vector from \(N = O(k \log M)\) noisy intensity measurements. Remarkably, the number of measurement is essentially of the same order as the number of ones needed to solve the classical sparse recovery problem where the phases of the measurements are known.

# 2. THE POLARIZATION PHASE-RETRIEVAL PROCEDURE (PPP)

The method we propose is an adaptation of the polarization method.\(^{16}\) We now motivate the main ideas behind the polarization method. Take a finite set \(V\), and suppose we take (noiseless, for the sake of clarity) intensity measurements of \(x \in \mathbb{C}^M\) with a spanning set \(\Phi_V := \{\varphi_i\}_{i \in V} \subseteq \mathbb{C}^M\). Having \(\langle x, \varphi_i \rangle\) for every \(i \in V\), we claim it suffices to determine the relative phase between \(\langle x, \varphi_i \rangle\) and \(\langle x, \varphi_j \rangle\) for sufficiently many pairs of \(i \neq j\). Indeed, if we had this information, we could arbitrarily assign some nonzero coefficient \(c_i = |\langle x, \varphi_i \rangle|\) to have positive phase. If \(\langle x, \varphi_j \rangle\) is also nonzero, then it has well-defined relative phase

\[
\rho_{ij} := \left( \frac{\langle x, \varphi_i \rangle}{|\langle x, \varphi_i \rangle|} \right)^{-1} \left( \frac{\langle x, \varphi_j \rangle}{|\langle x, \varphi_j \rangle|} \right), \tag{2}\]

which determines the phase of the \(|\langle x, \varphi_j \rangle|\) measurement by multiplication: \(c_j = \rho_{ij} |\langle x, \varphi_j \rangle|\). After having the phases of the linear measurements \(x\) can be reconstructed using a standard least-squares approach.

The term “polarization” comes from the fact that we leverage the polarization identity to obtain the relative phase \(\rho_{ij}\) from intensity measurements of other measurement vectors. More precisely, taking \(\zeta := e^{2\pi i/3}\) one has, for any \(\langle x, \varphi_i \rangle\) and \(\langle x, \varphi_j \rangle\),

\[
|\langle x, \varphi_i \rangle \langle x, \varphi_j \rangle| = \frac{1}{3} \sum_{k=0}^{2} \zeta^k |\langle x, \varphi_i \rangle + \zeta^{-k} \langle x, \varphi_j \rangle|^2 = \frac{1}{3} \sum_{k=0}^{2} \zeta^k |\langle x, \varphi_i + \zeta^k \varphi_j \rangle|^2. \tag{3}\]
Thus, if in addition to $\Phi_V$ we measure with $\{\varphi_i + \zeta^k \varphi_j\}_{k=0}^2$, we can use (3) to determine $\langle x, \varphi_i \rangle \langle x, \varphi_j \rangle$ and then normalize to get the relative phase:

$$\rho_{ij} := \left( \frac{\langle x, \varphi_i \rangle}{|\langle x, \varphi_i \rangle|} \right)^{-1} \frac{\langle x, \varphi_j \rangle}{|\langle x, \varphi_j \rangle|} \frac{\langle x, \varphi_i \rangle \langle x, \varphi_j \rangle}{|\langle x, \varphi_i \rangle| |\langle x, \varphi_j \rangle|},$$

provided both $\langle x, \varphi_i \rangle$ and $\langle x, \varphi_j \rangle$ are nonzero.

In practice the measurements are noisy which renders the relative phase calculation noisy as well,

$$\hat{\rho}_{ij} = \frac{\langle x, \varphi_i \rangle \langle x, \varphi_j \rangle + \rho_{ij}}{|\langle x, \varphi_i \rangle \langle x, \varphi_j \rangle + \rho_{ij}|},$$

and a small $|\langle x, \varphi_j \rangle|$ would imply a blowup of the noise because of the normalization (5). There is also the difficulty of how to estimate the phases of the measurements from these noisy relative phases (5).

We can represent the measurement system with a graph, having a vertex for each element in $\Phi_V$ and, for each edge $(i, j)$, include the measurements $\{\varphi_i + \zeta^k \varphi_j\}_{k=0}^2$ denoting this “edge” set of measurements as $\Phi_E$. We will be interested in measurement designs associated with sparse graphs as we want to keep the number of measurements as low as possible. Provided that the vertex measurements are non-zero (and the measurement process is noiseless) a connected graph is necessary and sufficient to ensure that one can obtain the vertex phases from the relative phases. A naive approach is to start at a vertex $i$ and travel through a spanning tree obtaining the new vertex phases by multiplying by the relative phases and the phases of the previous vertex. However, when the measurements are noisy this process will typically suffer from error accumulation (22).

After the measurements are obtained a (small) ratio of the vertex measurements, corresponding to the lowest intensity measurements, is removed to avoid the possible blowup of noise upon the calculation of the relative phase (5). The idea is that, provided that the graph has good enough connectivity properties and the removed ratio is small enough, the surviving graph will contain a large enough subgraph that is still connected and the error on the relative phases for that subgraph is controlled. At this point the phases of the vertex measurements are estimated using a robust and “democratic” method known as Angular Synchronization (21) that takes into account all the relative phases instead of just a spanning tree. Finally, estimates for the surviving vertex measurements are computed by combining the estimated phase with the intensity measurement. In order to control the effect of phase error on the actual linear measurement estimation, a few large vertex associated intensity measurements are discarded (this is a purely technical step). It is then a classical least-squares problem to recover the signal from these linear measurements.

The Polarization Phase-Retrieval Procedure (PPP) (28) is a procedure that uses the ideas briefly described above to, from noisy intensity measurements from a graph associated measurement system $[\Phi_V, \Phi_E]$, recover estimates for a large subset of the vertex associated linear measurements (meaning with phase). In the reminder of this section we describe this procedure.

PPP requires the measurement system to be designed with a graph $G$ having sufficiently good connectivity properties and a few parameters to be set a priori: $r_{SV}$, $r_{LV}$ and $\tau$. They are essentially the ratios of vertex measurements discarded for respectively: the removal procedure that discards low vertex intensity measurement, the removal procedure for the largest vertex measurements, and the total portion of removed vertex measurements (including also the ones removed to insure the connectivity of the graph). In order for the stability guarantees to work, among other things, these parameters need to be smaller than $\frac{1}{3}$ and satisfy

$$\tau > \frac{3}{2} r_{SV} + r_{LV} + \frac{3}{4} (1 - \lambda_2),$$

where $\lambda_2$ is the spectral gap of the graph $G$, a quantitative notion of connectivity (28). Essentially, for fixed $d$, it is possible to build regular graphs with degree $d$, meaning that $|E| = \frac{d}{2} |V|$, and $\lambda_2 = 1 - O\left(\frac{1}{\sqrt{d}}\right)$ (28).

The first step of PPP is to prune the graph by discarding vertices with low intensity measurements that would cause instabilities in the recovery process, this is accomplished by Algorithm 1.

Since Algorithm 1 can potentially destroy the connectivity properties of the graph another pruning step needs to be taken, Algorithm 3 to find a subgraph of the output of Algorithm 1 which has good connectivity.
Algorithm 1: Pruning for reliability

Input: Graph $G = (V, E)$, function $f: E \to \mathbb{R}$ such that $f(i, j) = |\langle x, \varphi_i \rangle \langle x, \varphi_j \rangle + \epsilon_{ij}|$, parameter $\alpha$

Output: Subgraph $H$ with a larger smallest edge weight

Initialize $H \leftarrow G$

for $i = 1$ to $\lfloor (1 - \alpha) |V| \rfloor$ do
  Find the minimizer $(i, j) \in E$ of $f$
  $H \leftarrow H \setminus \{i, j\}$
end

Algorithm 2: Spectral clustering

Input: Graph $G = (V, E)$

Output: Subset of vertices $S$

Take $D$ to be the diagonal matrix of vertex degrees and $A$ to be the adjacency matrix

Compute the Laplacian $L \leftarrow I - D^{-1/2} AD^{-1/2}$

Compute an eigenvector $u$ corresponding to the second smallest eigenvalue of $L$

for $i = 1$ to $|V|$ do
  Let $S_i$ denote the vertices corresponding to the $i$ smallest entries of $D^{-1/2} u$
  Let $E(S_i, S^c_i)$ denote the number of edges between $S_i$ and $S^c_i$
  $h_i \leftarrow E(S_i, S^c_i) / \min\{\sum_{v \in S_i} \deg(v), \sum_{v \in S^c_i} \deg(v)\}$
end

Take $S$ to be the $S_i$ of minimal $h_i$ (or $S^c_i$ if this has smaller size)

properties. In order to find areas of the graph that lack good levels of connectivity, the pruning method uses a Spectral Clustering method$^{28}$ Algorithm 2.

After having the relative phases of a subgraph with sufficiently good connectivity Angular Synchronization$^{27,30}$ (Algorithm 4) outputs estimates for the phases of the vertex measurements whose error is comparable with the noise in the relative phases$^{30}$.

PPP, Algorithm 5, is essentially composed by these subroutines. Section 4 contains a statement about the stability guarantee for PPP that is shown in a previous publication$^{16}$ As we will see in the next Section, PPP is somewhat flexible and can be used as a subroutine to perform sparse phase recovery.

3. SPARSE PHASE RECOVERY ALGORITHM

In this Section we propose an algorithm to solve the Sparse Phase Recovery problem. It will be based on the PPP. For PPP to succeed one needs to design both a favorable vertex measurement set and graph $G$ with sufficient connectivity. Essentially the vertex measurement should guarantee that, for any signal of interest $x \in \mathbb{C}^M$, the number of vertices with either too small or too large of an intensity measurement is small, in order to control the number of vertices that have to be removed. The graph needs to have sufficiently nice connectivity properties so that this vertex removal does not destroy its connectivity. Finally, the vertex measurement system needs to be robust to erasures as one recovers the phase of just a (large) subset of the measurements.

The vertex measurement system is drawn randomly, each measurement vector is i.i.d. with complex gaussian i.i.d. entries with variance $\frac{1}{M}$. To build the graph we leverage the theory of expanders$^{29}$ after picking $d$ sufficiently large, one can build a regular graph with degree $d$ and $\lambda_2 \geq 1 - 2\sqrt{d - 1} \pm \epsilon$, for any a priori chosen $\epsilon$.

After using PPP we get a subset $\hat{V} \subseteq V$ and noisy estimates

$$\hat{y} = \Phi_{\hat{V}}^* x + e.$$

As we will see in the next section, we will have that $\Phi_{\hat{V}}$ satisfies the Restricted Isometry Property which will allow us to reconstruct the $k$-sparse vector $x$ from $\hat{y}$ with Algorithm 6.
Algorithm 3: Pruning for connectivity

**Input:** Graph $G = (V, E)$, pruning parameter $\mu$

**Output:** Subgraph $H$ with spectral gap $\lambda_2(H) \geq \mu$

Initialize $H \leftarrow G$

while $\lambda_2(H) < \mu$ do

Perform spectral clustering (Algorithm 2) to identify a small set of vertices $S$

$H \leftarrow H \setminus S$

end

Algorithm 4: Angular synchronization

**Input:** Graph $G' = (V', E')$, noisy versions of $\{i, j\} \in E'$

**Output:** Vector of phases corresponding to vertex measurements

Let $A_1$ denote the matrix given by the noisy estimations of the relative phase $\rho_{ij}$ whenever $\{i, j\} \in E'$, and otherwise $A_1[i, j] = 0$

Let $D$ denote the diagonal matrix of vertex degrees

Compute the connection Laplacian $L_1 \leftarrow I - D^{-1/2}A_1D^{-1/2}$

Compute an eigenvector $u$ corresponding to the smallest eigenvalue of $L_1$

Output the phases of the coordinates of $u$

The whole procedure reads as follows:

**Measurement Design**

- Fix $d > 2$ even and $\varepsilon$ sufficiently small.
- Given $M$, pick some $d$-regular graph $G = (V, E)$ with spectral gap $\lambda_2 \geq 1 - \frac{\sqrt{d-1} + \varepsilon}{d}$ and $|V| = ck \log M$ for $c$ sufficiently large, and arbitrarily direct the edges.
- Design the measurements $\Phi := \Phi_V \cup \Phi_E$ by taking $\Phi_V := \{\varphi_i\}_{i \in V} \subseteq \mathbb{C}^M$ to have independent entries with distribution $\mathcal{CN}(0, \frac{1}{M})$ and $\Phi_E := \bigcup_{(i, j) \in E} \{\varphi_i + \zeta^k \varphi_j\}_{k=0}^2$.

**Sparse Phase Retrieval Procedure**

- Run PPP (Algorithm 5) picking appropriate parameters $r_{SV}$, $r_{LV}$, and $\tau$ to get a vertex subset $\hat{V}$ and estimates $\hat{y}$ of the linear measurements corresponding to $\hat{V}$.
- Perform $\ell_1$ minimization (Algorithm 6) with $y \leftarrow \hat{y}$ and $A \leftarrow \Phi_{\hat{V}}^*$ to recover $x$.

4. RECOVERABILITY AND STABILITY GUARANTEES

This section deals with the analysis of the Sparse Phase Retrieval Algorithm proposed in the previous section. This section serves only as a brief overview and the proofs are omitted. We direct the reader to a more complete publication for more details.

4.1 Guarantees for PPP

In order to control the size of the smallest and largest vertex intensity measurements after the graph pruning step we define normalized Projective Uniformity for Small Vertices (nPUSV) and normalized Projective Uniformity for Large Vertices (nPULV).
with spectral gap $\lambda$ analysis in the first polarization-based phase retrieval paper.

16 of interest. With these definitions we can present a guarantee for PPP which can be easily obtained from the

$\alpha$ pruning parameter $nPUSV$

Then there exist two constants $c_1$ and $c_2$ depending only on: $\tau - \frac{3}{2}r_{SV} + r_{LV} + \frac{3}{4}(1 - \lambda_2)$, a lower bound on $nPUSV(\Phi_V, x; 1 - \frac{1}{4}r_{SV})$ and an upper-bound on $nPULV(\Phi_V, x; 1 - \frac{1}{2}r_{LV})$ such that, as long as,

$$\text{SNR} := \frac{\|x\|_2}{\|\nu\|_2} \geq c_1 \sqrt{M},$$
it outputs a set $\hat{V} \subseteq V$ such that $|\hat{V}| \geq \tau |V|$ and an estimate $\hat{y}$ indexed by $\hat{V}$ such that

$$\min_{\theta \in [0, 2\pi]} \|\hat{y} - e^{i\theta} \Phi_{\hat{V}}^* x\|^2 \leq c_2 \left( \frac{\sqrt{M}}{\text{SNR}} + \sqrt{\frac{|V|}{M}} \right) \frac{\sqrt{M}}{\text{SNR}} \|x\|^2.$$ 

### 4.2 Main Theorem – Guarantee for Sparse Phase Retrieval

There are essentially two adaptations that need to be done to the original polarization phase retrieval analysis \[^{[3]}\] to adapt it to the sparse case: a new bound on Projective Uniformity constants (Lemma \[^{[4]}\]) and an erasure robust version of the Restricted Isometry Property (Lemma \[^{[4]}\]).

As the original algorithm needs to be able to recover every possible signal $x \in \mathbb{C}^M$ one needs to bound the Projective Uniformity constants (Definition \[^{[4]}\]) over all signals $x \in \mathbb{C}^M$ which required $N \sim M \log M$. Here, however, we only need to control the Projective Uniformity constants for $k$-sparse vectors $x \in \mathbb{C}^M$, and so this vastly decreases the union bounding that is needed and allows for $N \sim k \log M$.

**Lemma 4.3 (Bound on Projective Uniformity constants \[^{[4]}\]).** Let $\frac{2}{3} < \alpha < 1$ be fixed. For sufficiently large $M$, let $\Phi_V := \{\varphi_i\}_{i \in V} \subseteq \mathbb{C}^M$ be drawn randomly to have independent entries with distribution $\mathcal{C} \mathcal{N}(0, \frac{1}{M})$. There exist constants $c_{av}$ and $c_{1v}$, depending only on $\alpha$, such that the following holds with high probability: for all $k$-sparse signals $x \in \mathbb{C}^M$

$$c_{av} \leq n \text{PUSV}(\Phi_V, x; \alpha) \leq n \text{PULV}(\Phi_V, x; \alpha) \leq c_{1v}.$$ 

Lemma \[^{[4]}\] essentially allows us to use the analysis of performance of PPP (Theorem \[^{[4]}\]) to guarantee that it outputs a set $\hat{V} \subseteq V$ such that $|\hat{V}| \geq \tau |V|$ and an estimate $\hat{y}$ indexed by $\hat{V}$ whose distance (up to a global phase factor) to $\Phi_{\hat{V}}^* x$ we can control. This means that after using PPP we have noisy “measurements” $\hat{y} = \Phi_{\hat{V}}^* x$ from which we want to recover the $k$-sparse signal $x$. In the last decade there was significant progress in understanding when such recovery is possible, a very popular sufficient condition on the involved matrix is known as the Restricted Isometry Property.

**Definition 4.4 (Restricted Isometry Property).** We say that a matrix $A$ satisfies the $(2k, \delta)$-Restricted Isometry Property (RIP) if, for every $2k$-sparse vector $x$,

$$(1 - \delta)\|x\|^2 \leq \|Ax\|^2 \leq (1 + \delta)\|x\|^2.$$

It’s popularity is justified by the following theorem \[^{[3]}\].

**Theorem 4.5.** Suppose an $n \times M$ matrix $A$ has the $(2, \delta)$-Restricted Isometry Property (RIP) for some $\delta < \sqrt{2} - 1$. Then, for every $k$-sparse vector $x \in \mathbb{C}^M$, Algorithm \[^{[3]}\] takes as input noisy measurements $y = Ax + \epsilon$ and, assuming $\|\epsilon\| \leq E$, outputs $\hat{x}$ satisfying $\|\hat{x} - x\| \leq C\varepsilon$, where $C$ only depends on $\delta$.

If $\Phi_V^*$ has the $(2K, \frac{1}{K})$-RIP then using Algorithm \[^{[3]}\] with the estimates $\hat{y}$ will provide an estimate $\hat{x}$ for which we can control the error, $\min_{\theta \in [0, 2\pi]} \|\hat{x} - e^{i\theta} x\|$, with respect to the true signal $x$. It turns out $\Phi_V^*$ does have the desired property because one can show that that $\Phi_V$ is RIP in a erasure robust way.

**Definition 4.6 (Erasure Robust RIP).** We say that an $|V| \times M$ matrix $A$ satisfies the erasure robust Restricted Isometry Property for $(k, \delta, \tau)$, $(k, \delta, \tau)$-erRIP, if for every selection $T$ of $|T| = \lfloor (1 - \tau)M \rfloor$ rows, the matrix formed by the rows indexed by $T$ is $(k, \delta)$-RIP.

It turns out that one can show\[^{[3]}\] that if $\tau$ is a suficiently small parameter, the described measurement design renders, with high probability, $\Phi_V$ a $(2k, \frac{1}{k}, \tau)$-erRIP matrix. This is made precise in the lemma below. The proof is omitted and available in a more complete publication\[^{[3]}\] but it essentially consists of adapting a simple
proof\textsuperscript{[11]} that random matrices satisfy RIP by union bounding not only on all possible subset of columns with a given size but also on all possible choices of surviving rows $T$.

**Lemma 4.7 (erRIP Matrices\textsuperscript{[11]}).** Let $\tau$ be sufficiently small. Let $A$ be a $|V| \times M$ matrix with random gaussian iid. entries with distribution $CN(0, \frac{1}{|V|})$. Then, there exist positive constants $c_1$ and $c_2$ such that (for sufficiently large $M$), for all $k < c_1 |V| \log^{-1}(M)$, we have that $\sqrt{\frac{M}{(1-\tau)|V|}}A$ is erasure robust RIP, more precisely, $\sqrt{\frac{M}{(1-\tau)|V|}}A$ is $(k, \frac{1}{10}, \tau)$-erRIP with high probability.

Our main result, whose proof essentially consists of the arguments briefly described above and can be found in a more complete publication\textsuperscript{[11]} reads as follows:

**Theorem 4.8.** Pick $N \sim c_0 k \log M$ with $c_0$ sufficiently large and take $\{\varphi_i\}_{i=1}^N = \Phi_V \cup \Phi_E$ defined in the measurement design. Then there exist constants $c_1, c_2 > 0$ such that the following holds with high probability for all $k$-sparse signals $x \in \mathbb{C}^M$: Consider measurements of the form

$$z_t := |\langle x, \varphi_t \rangle|^2 + \nu_t.$$ 

If the normalized signal-to-noise ratio satisfies $nSNR := \frac{1}{|z_t|} \sum_{\ell=1}^N \min \{\varphi_i \} \geq c_1 \sqrt{\frac{k \log M}{M}}$, then the sparse phase retrieval procedure (described in Section 3) produces an estimate $\hat{x}$ from $\{z_t\}_{t=1}^N$ with squared relative error

$$\frac{||\hat{x} - e^{i\theta} x||^2}{||x||^2} \leq c_2 \left( \frac{1}{nSNR} + 1 \right) \frac{1}{nSNR}$$

for some phase $\theta \in [0, 2\pi)$.

If the intensity measurements $z_t := |\langle x, \varphi_t \rangle|^2$ are noiseless then the typical value of a $z_t$ is on the order of $(\frac{1}{\sqrt{M}} ||x||)^2$ which means that the $\ell_2$ norm of typical $z$ is of the order of $\sqrt{\frac{k \log M}{M}} ||x||^2$. This justifies the choice of normalization for $nSNR$, the normalized signal-to-noise ratio.

5. CONCLUSION AND FUTURE WORK

This paper successfully adapts the polarization method\textsuperscript{[11]} to the problem of recovering a sparse signal from noisy intensity measurements. Our main result (Theorem 4.8) guarantees that the recovery is successful with as few as $N = O(k \log M)$ measurements where $M$ is the ambient dimension of the signal and $k$ its sparsity level. This is particularly remarkable as it has been shown\textsuperscript{[29]} that a certain class of adaptations of PhaseLift to the sparse case fail for $N = O(k^2 \log^{-2} M)$. However, there is still a (tiny) gap with respect to stability of the measurement process which is known\textsuperscript{[10]} to hold, at least in the real case, for $N = O(\frac{k \log M}{T})$.

One advantage of this method is its flexibility to exploit known structure of the signal. After PPP recovers the phases of the linear measurements the linear inverse problem can be solved in any way that might exploit structure in the signal (such as $\ell_1$ exploits sparsity in this application). One immediate observation is that it can be easily adapted to deal with structured types of sparsity, such as partial sparsity\textsuperscript{[32]} or block sparsity\textsuperscript{[33]}.

The measurement design currently exploits randomness in both guaranteeing good Projective Uniformity constants and good sparse recovery properties. In particular, we require $\Phi_V^*$ to be erRIP. Since constructing deterministic RIP matrices seems to be a particularly difficult problem\textsuperscript{[33]} derandomizing the measurement design described above seems to be a very challenging problem. As checking for the RIP on a matrix is known\textsuperscript{[35]} to be an NP-hard problem it also seems difficult to be able to check whether the measurement system has the desired properties.

PPP is tuned to allow for worst-case scenario guarantees, however oftentimes an average-case scenario (like assuming the noise is random) is more representative of the algorithm’s performance on applications. One interesting line for future work is to tune the algorithm to improve its performance in practice and perform some average-case analysis. Another interesting open question is related to Fourier Mask measurements; although the polarization method has been shown to succeed in recovery from noiseless intensity Fourier Mask measurement\textsuperscript{[17]} there is no guarantee of stability as in the original phase retrieval with polarization paper\textsuperscript{[10]} and the present paper.
ACKNOWLEDGMENTS

The authors thank Boris Alexeev for insightful discussions and a suggestion that was the starting point towards the ideas presented in this paper. The authors also thank Vladislav Voroninski for insightful discussions and Nicolas Boumal for reading and commenting on an earlier version of this manuscript. A.S. Bandeira was supported by NSF Grant No. DMS-0914892 and D.G. Mixon was supported by NSF Grant No. DMS-1321779. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

REFERENCES

[1] Bandeira, A. S. and Mixon, D. G., “Compressed sensing with intensity measurements,” to appear (2013).
[2] Harrison, R., “Phase problem in crystallography,” J. Opt. Soc. Am. A 10 (1993) 1046–1055 (1993).
[3] Miao, J., Ishikawa, T., Shen, Q., and Earnest, T., “Extending x-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes,” Annu. Rev. Phys. Chem. 59 (2008) 387–410 (2008).
[4] Millane, R., “Phase retrieval in crystallography and optics,” J. Opt. Soc. Am. A 7 (1990) 394–411. (1990).
[5] Bunk, O. et al., “Diffractive imaging for periodic samples: retrieving one-dimensional concentration profiles across microfluidic channels,” Acta Cryst. A63, 306–314 (2007).
[6] Dainty, J. and Fienup, J., “Phase retrieval and image reconstruction for astronomy,” In: H. Stark, ed., Image Recovery: Theory and Application, Academic Press, New York (1987).
[7] Wolter, A., “The question of phase retrieval in optics,” Opt. Acta 10 (1963) 41–49 (1963).
[8] Balan, R., Casazza, P., and Edidin, D., “On signal reconstruction without phase,” Appl. Comput. Harmon. Anal. 20 (2006) 345–356 (2006).
[9] Bandeira, A. S., Cahili, J., Mixon, D. G., and Nelson, A. A., “Saving phase: Injectivity and stability for phase retrieval,” arxiv (2013).
[10] Eldar, Y. C. and Mendelson, S., “Phase retrieval: Stability and recovery guarantees,” available online (2012).
[11] Fienup, J. R., “Phase retrieval algorithms: a comparison,” Appl. Optics 21, 2758–2769 (1982).
[12] Candès, E. J., Strohmer, T., and Voroninski, V., “Phaselift: exact and stable signal recovery from magnitude measurements via convex programming,” Communications on Pure and Applied Mathematics (2011).
[13] Candès, E. and Li, X., “Solving quadratic equations via phaselift when there are about as many equations as unknowns,” available online (2012).
[14] Candès, E., Eldar, Y., Strohmer, T., and Voroninski, V., “Phase retrieval via matrix completion,” arXiv:1109.0573 (2011).
[15] Waldspurger, I., d’Aspremont, A., and Mallat, S., “Phase recovery, maxcut and complex semidefinite programming,” arXiv:1206.0102 (2012).
[16] Alexeev, B., Bandeira, A. S., Fickus, M., and Mixon, D. G., “Phase retrieval with polarization,” available online (2012).
[17] Bandeira, A. S., Chen, Y., and Mixon, D. G., “Phase retrieval from power spectra of masked signals,” available online (2013).
[18] Donoho, D. L., “Compressed sensing,” IEEE Trans. Inform. Theory 52, 1289–1306 (2006).
[19] Candès, E. J., Tao, T., and Romberg, J., “Robust uncertainty principles: exact signal reconstruction from incompletely sampled Fourier information,” IEEE Trans. Inform. Theory 52, 489–509 (2006).
[20] Candès, E. J., Romberg, J., and Tao, T., “Stable signal recovery from incomplete and inaccurate measurements,” Comm. Pure Appl. Math. 59, 1207–1223 (2006).
[21] Candès, E. J. and Tao, T., “Decoding by linear programming,” IEEE Trans. Inform. Theory 51, 4203–4215 (2005).
[22] Candès, E. J. and Tao, T., “Near optimal signal recovery from random projections: universal encoding strategies?,” IEEE Trans. Inform. Theory 52, 5406–5425 (2006).
[23] Candès, E. J., “The restricted isometry property and its implications for compressed sensing,” C. R. Acad. Sci. Paris, Ser. I 346, 589–592 (2008).
[24] Moravec, M. L., Romberg, J. K., and Baraniuk, R. G., “Compressive phase retrieval,” in [Proc. of SPIE Vol. 6701], (2007).

[25] Ohlsson, H., Yang, A. Y., Dong, R., and Sastry, S. S., “Compressive phase retrieval from squared output measurements via semidefinite programming,” available online (2012).

[26] Li, X. and Voroninski, V., “Sparse signal recovery from quadratic measurements via convex programming,” available online (2012).

[27] Singer, A., “Angular synchronization by eigenvectors and semidefinite programming,” Appl. Comput. Harmon. Anal. 30(1), 20 – 36 (2011).

[28] Chung, F., “Four proofs for the cheeger inequality and graph partition algorithms,” Fourth International Congress of Chinese Mathematicians, pp. 331–349 (2010).

[29] Friedman, J., “A proof of Alon’s second eigenvalue conjecture and related problems,” Mem. Amer. Math. Soc. 195 (2008).

[30] Bandeira, A. S., Singer, A., and Spielman, D., “A Cheeger inequality for the graph connection laplacian,” available online (2012).

[31] Baraniuk, R. et al., “A simple proof of the restricted isometry property for random matrices,” Constr. Approx. 28, 253–263 (2008).

[32] Bandeira, A. S., Vicente, L. N., and Scheinberg, K., “On partial sparse recovery,” tech. rep., CMUC, Department of Mathematics, University of Coimbra, Portugal (2011).

[33] Eldar, Y., Kuppinger, P., and Bolcskei, H., “Block-sparse signals: Uncertainty relations and efficient recovery,” Signal Processing, IEEE Transactions on 58(6), 3042–3054 (2010).

[34] Bandeira, A. S., Fickus, M., Mixon, D. G., and Wong, P., “The road to deterministic matrices with the restricted isometry property,” available online (2012).

[35] Bandeira, A., Dobriban, E., Mixon, D., and Sawin, W., “Certifying the restricted isometry property is hard,” IEEE Trans. Inform. Theory 59(6), 3448–3450 (2013).