COMPRESSIVE PHASE RETRIEVAL UNDER POISSON NOISE

Chinmay Talegaonkar[1], *Parthasarathi Khirwadkar*[2] *and Ajit Rajwade*[3]

[1,2]Department of Electrical Engineering, IIT Bombay
[3]Department of Computer Science and Engineering, IIT Bombay

ABSTRACT

We present a technique for compressive phase retrieval under Poisson noise using the theory of variance stabilization transforms (VSTs). We modify two existing algorithms using VSTs, and derive worst-case performance bounds for both the algorithms. Our proposed modification allows for easy and very principled parameter tuning. Our estimator is tractable and we also show numerical results on phase recovery of sparse signals for Poisson corrupted measurements, and demonstrate the relative advantage of our modification at low intensities. We also present a comparison of the performance and other theoretical aspects of both the algorithms.

***Index Terms*—** Variance Stabilization transforms, Compressive Phase Retrieval, Performance Bounds

1. INTRODUCTION

The problem of phase retrieval aims to estimate an underlying signal from its phaseless (squared) magnitude measurements. It finds applications in various problems like microscopy, ptychography and astrophysical imaging [1]. While acquiring measurements in a physical system, the inability of the sensors to record phase is also accompanied by photon shot noise (Poisson noise), which is a signal dependent noise particularly dominant in low intensity (i.e. low SNR) regimes like astrophysical imaging. Hence, the algorithms for phase retrieval should also take into account the Poisson corrupted nature of the acquired measurements, but not much work has been done in this direction, from a theoretical or algorithmic perspective.

There has been sizeable recent work on both convex and non-convex methods for phase retrieval. Some of the popular non-convex methods for phase retrieval are [2], [3], [4], [5] and [6]. We focus on convex methods, in particular lifting-based approaches in this work. The PhaseLift method [7], which seeks to estimate a rank-one matrix $\boldsymbol{X} \triangleq \boldsymbol{x}\boldsymbol{x}^H$ instead of the signal $\boldsymbol{x}$, is one of the seminal contributions in this area. The lifting-based approaches have performance guarantees but are computationally expensive for high dimensions. [8] provides a convex method based on basis pursuit for phase retrieval that doesn't require the lifting approach. Some recent work also includes projected gradient based iterative algorithms or ADMM for phase retrieval to improve the scalability of the convex algorithms [9], [10]. Phase retrieval is a non-linear ill posed problem, and hence requires oversampling or a large number of measurements for efficient recovery [1]. Incorporating prior knowledge about the signal is one way to improve performance, especially in the case of fewer phaseless measurements. This variant of the phase retrieval problem is known as *compressive phase retrieval*. [10] uses an ℓ_1 sparsity-promoting regularization on $\boldsymbol{X}$ in addition to trace minimization in the phase lift algorithm. [11] proves that for Fourier measurements of a k-sparse signal, $\mathcal{O}(k^2 \log(\frac{n}{k}))$ measurements are sufficient for accurate recovery. In contrast to jointly solving for the low rank and sparse matrix as done in [10], recent works like [12], [13], [14] use a 2-stage iterative method for compressive phase recovery which separately handles the low rank and sparse estimation for compressive phase retrieval and guarantees accurate phase recovery from $\mathcal{O}(k \log(\frac{n}{k}))$ measurements. To our best knowledge, there has been no work so far on compressive phase retrieval under Poisson noise. In [15], expected error bounds for holographic phase retrieval with Poisson noise are derived, but in the non-compressive regime.

Main Contributions: We present the following contributions in this paper:

1. We present a statistically motivated and tractable VST based approach to compressive phase retrieval under Poisson noise.
2. We demonstrate the improvement in performance obtained for two standard algorithms using our approach. We also prove performance bounds for modifications to both these algorithms. The bounds depend on the intensity of the signal observed, but implementing the algorithm does not require any such knowledge.

2. PRELIMINARIES

2.1. Compressive Phase retrieval via Lifting (CPRL)

The standard phase retrieval problem can be formulated as

$$\text{find} \quad \boldsymbol{x} \quad \text{subj. to} \quad \boldsymbol{y} = |\boldsymbol{A}\boldsymbol{x}|^2 + \boldsymbol{e}$$

chinmay0301@iitb.ac.in[1], parthasarathi.k@iitb.ac.in[2]
ajitvr@cse.iitb.ac.in[3], AR thanks IIT-B Seed Grant 14IRCCSG012 and NVIDIA Corporation for the donation of Titan Xp GPUs which were useful for this work

where $|\boldsymbol{Ax}|^2 = \{\boldsymbol{a}_i^H \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{a}_i\}_{1\leq i\leq N}$ ($\boldsymbol{a}_i$ is the i^{th} row of A), $\boldsymbol{y}^T = [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_n] \in \mathbb{R}^{n\times 1}$, $\boldsymbol{x} \in \mathbb{C}^{d\times 1}$, and $\boldsymbol{e}$ is the i.i.d additive Gaussian noise vector. It is well known that this is a general QCQP type of optimization problem, which is combinatorial and NP hard (even for real signals). The PhaseLift method [16] represents this problem as the task of finding a rank 1 positive semi-definite matrix $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^H$ from linear measurements $\boldsymbol{y}$ instead of estimating a d dimensional vector from quadratic measurements $|\boldsymbol{Ax}|^2$, where $\boldsymbol{A} \triangleq \boldsymbol{RF}$, where $\boldsymbol{R} \in \mathbb{C}^{n\times d}$ is a random projection matrix with every entry an i.i.d complex or real Gaussian, and $\boldsymbol{F}$ can be a DFT matrix or an i.i.d random Gaussian matrix. Since finding a rank 1 matrix is a non-convex problem, trace minimization is used as a convex surrogate to rank minimization which makes this problem an SDP, and hence can be solved using convex optimization techniques. Therefore using the lifting idea, the phase retrieval problem can be posed as follows:

$$\min_{\boldsymbol{X}} \quad \mathrm{Tr}(\boldsymbol{X}) \tag{1}$$

$$\text{subj. to} \quad \|B(\boldsymbol{X}) - \boldsymbol{y}\|_2 \leq \varepsilon, \boldsymbol{X} \succeq \boldsymbol{0} \tag{2}$$

where B is a linear operator on $\boldsymbol{X}$ defined as

$$B : \boldsymbol{X} \in \mathbb{C}^{d\times d} \mapsto \{\mathrm{Tr}(\boldsymbol{a}_i^H \boldsymbol{X} \boldsymbol{a}_i)\}_{1\leq i\leq N} \in \mathbb{R}^N, \tag{3}$$

and ε is an upper bound on $\|\boldsymbol{e}\|_2$. The solution to this optimization problem guarantees unique and accurate reconstructions up to global phase ambiguities like time shift, conjugate flip and constant phase [1]. If the signal $\boldsymbol{x}$ is given to be sparse, then the above formulation can be modified to incorporate the sparsity of the signal also using ℓ_1 regularization form of the trace norm, which gives the objective function for compressive phase retrieval via lifting (CPRL) given by Ohlsson et. al [10]:

$$\mathsf{P1} : \min_{\boldsymbol{X}} \quad \mathrm{Tr}(\boldsymbol{X}) + \lambda\|\boldsymbol{X}\|_1 \tag{4}$$

$$\text{subj. to} \quad \|B(\boldsymbol{X}) - \boldsymbol{y}\|_2 \leq \varepsilon, \boldsymbol{X} \succeq \boldsymbol{0}. \tag{5}$$

The estimate of $\boldsymbol{x}$ can be then found using an SVD-based rank-1 decomposition of $\boldsymbol{X}$ up to global ambiguities [10].

2.2. Compressive Phase Retrieval - 2 stage algorithm

Compressive phase retrieval can be interpreted as a problem of recovering a low rank and sparse matrix $\boldsymbol{X}$. [13] provides a 2-stage algorithm for this task. The first stage involves estimating $\widehat{\boldsymbol{A}}$ from the underlying low rank matrix A, followed by the second stage which uses $\widehat{\boldsymbol{A}}$ to obtain a low rank and sparse estimate $\widehat{\boldsymbol{X}}$. For measurements of the form $\boldsymbol{y}_i = |\langle \boldsymbol{a}_i, x\rangle|^2 + e_i$, the sensing vectors $\boldsymbol{a}_i$ are assumed to be of the form $\boldsymbol{a}_i = \boldsymbol{\Psi}^T \boldsymbol{w_i}$, where $\boldsymbol{\Psi} \in R^{m\times d}$ with every element drawn i.i.d from $\mathcal{N}(0, 1/m)$, $\boldsymbol{w_i} \in R^m$ is known, and $\boldsymbol{e_i}$ denotes elements of the additive i.i.d noise vector. Hence, the sensing vectors are assumed to lie in a fixed low-dimensional subspace. Such measurement systems occur in imaging systems where the illumination is controlled. The lifting scheme allows the quadratic measurements to be expressed as -:

$$\boldsymbol{y}_i = \left\langle \boldsymbol{a}_i \boldsymbol{a}_i^T, \boldsymbol{X} \right\rangle + e_i = \left\langle \boldsymbol{\Psi}^T \boldsymbol{w_i} \boldsymbol{w_i}^T \boldsymbol{\Psi}, \boldsymbol{X} \right\rangle + e_i, \tag{6}$$

where $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^H$, and the linear operators W & G defined as

$$W : A \mapsto \left[\left\langle w_i w_i^T, A\right\rangle\right]_{i=1}^n \quad \text{and} \quad G : X \mapsto W\left(\boldsymbol{\Psi} X \boldsymbol{\Psi}^T\right).$$

The measurements can hence be expressed concisely as

$$\boldsymbol{y} = G(\boldsymbol{X}) + \boldsymbol{e}.$$

In this work we assume $\boldsymbol{w_i} \sim \mathcal{N}(0, I)$, $\boldsymbol{\Psi}$ to be RIP-obeying for $2k$-sparse vectors and that $\|e\|_2 \leq \varepsilon$ as assumed in [13]. The algorithm solves the two convex problems of low rank estimation and sparse estimation sequentially. We implemented the low-rank estimation part using CVX, and the sparse estimation stage using iterative hard thresholding for 1000 iterations [17]. The estimated $\widehat{\boldsymbol{X}}$ may not necessarily be a positive semi-definite (PSD) matrix, and hence is then projected on a set of positive semi-definite rank 1 matrices to satisfy the constraint on $\boldsymbol{X}$ [13].

Algorithm 1: 2 Stage Method $\mathsf{P2}$

1 Low-rank estimation stage:

$$\begin{aligned} \widehat{\boldsymbol{A}} \in \underset{\boldsymbol{A}\succcurlyeq \boldsymbol{0}}{\operatorname{argmin}} \quad & \mathrm{Tr}(\boldsymbol{A}) \\ \text{subject to} \quad & \|W(\boldsymbol{A}) - \boldsymbol{y}\|_2 \leq \varepsilon \end{aligned} \tag{7}$$

2 Sparse estimation stage:

$$\begin{aligned} \widehat{\boldsymbol{X}} \in \underset{\boldsymbol{X}}{\operatorname{argmin}} \quad & \|\boldsymbol{X}\|_1 \\ \text{subject to} \quad & \left\|\boldsymbol{\Psi}\boldsymbol{X}\boldsymbol{\Psi}^T - \widehat{\boldsymbol{A}}\right\|_F \leq \frac{C\varepsilon}{\sqrt{n}} \end{aligned} \tag{8}$$

3. COMPRESSIVE PHASE RETRIEVAL FOR POISSON CORRUPTED SIGNALS

The square root is a variance stabilizer for Poisson distribution, i.e. if a measurement $y_i \sim \text{Poisson}(x_i)$ for the i^{th} measurement, then $\sqrt{y_i + c}$ is approximately Gaussian distributed with mean $\sqrt{x_i + c}$ and variance $1/4$ where c is a non-zero constant [18]. When $c = 3/8$, this is called the Anscombe transform. This approximation becomes more accurate as $x_i \to \infty$. Based on our prior work on Poisson denoising using VST [19], we define the residual vector $r(\boldsymbol{y}, \boldsymbol{X}) \triangleq \sqrt{\boldsymbol{y} + c} - \sqrt{B(\boldsymbol{X}) + c}$ and $B(\cdot)$ corresponds to the linear operator mentioned in section 2.1. Motivated by our work in [19],[20], we define an analogous constraint to be incorporated in the SDP for compressive phase retrieval, i.e. $\|r(\boldsymbol{y}, \boldsymbol{X})\|_2 \leq \varepsilon$. $\varepsilon \sim \mathcal{O}(\sqrt{n})$, where n is the number of measurements. This particular choice of ε is motivated by the Theorem 1, proved in our unpublished work [20]. When the measurements $\boldsymbol{y}$ are Poisson corrupted, using the constraints in P1 and P2 is not statistically correct. We extend this VST based modification for compressive phase retrieval from Poisson corrupted measurements. The modifications to P1 and the low rank estimation stage of P2 are as follows:

$$\begin{aligned} \mathrm{Q1} : & \min_{\boldsymbol{X}} \mathrm{Tr}(\boldsymbol{X}) + \lambda \|\boldsymbol{X}\|_1 \\ & \text{subj. to} \left\| \sqrt{\boldsymbol{y} + \frac{3}{8}} - \sqrt{B(\boldsymbol{X}) + \frac{3}{8}} \right\|_2 \leq \varepsilon, \boldsymbol{X} \succeq \boldsymbol{0} \end{aligned} \tag{9}$$

$$\begin{aligned} \mathrm{Q2} : \widehat{\boldsymbol{A}} \in & \operatorname*{argmin}_{\boldsymbol{A} \succcurlyeq \boldsymbol{0}} \quad \mathrm{Tr}\,(\boldsymbol{A}) \\ & \text{subj. to} \quad \left\| \sqrt{\boldsymbol{y} + \frac{3}{8}} - \sqrt{W(\boldsymbol{A}) + \frac{3}{8}} \right\|_2 \leq \varepsilon \end{aligned} \tag{10}$$

The sparse estimation stage in Q2 is the same as that in P2.

3.1. Performance bounds for Q1 and Q2

Theorem 1: *Bounds for Q1*
Abiding by the notations in [21], let $B(\cdot)$ be an (δ, k) RIP obeying linear operator [21], i.e. $\forall \boldsymbol{X} \neq \boldsymbol{0}$ and $\|\boldsymbol{X}\|_0 \leq k$

$$\left| \frac{\|B(\boldsymbol{X})\|_2^2}{\|\boldsymbol{X}\|_2^2} - 1 \right| < \delta. \tag{11}$$

Let $\boldsymbol{X}^*$ be the true rank 1 matrix that satisfies the constraints in Q1, and $\tilde{\boldsymbol{X}}$ be the solution to Q1 for an appropriate choice of parameters. We define $\Delta = \tilde{\boldsymbol{X}} - \boldsymbol{X}^\star$. For a matrix $\boldsymbol{X} \in \mathbb{R}^{d\times d}$, $\boldsymbol{X}_T$ denotes a matrix with all values zero except the indices in T, which are set to the corresponding values of $\boldsymbol{X}$. $\forall \lambda \geq \frac{2d}{1-\rho} + \frac{d}{k^{\frac{1}{2}}}$ and $\delta \leq \sqrt{2} - 1$ we get

$$\|\Delta\|_2 \leq \frac{2\alpha\varepsilon(1-\rho)^{-1} + 2(1+\rho)(1-\rho)^{-1}k^{-\frac{1}{2}}\|\tilde{\boldsymbol{X}} - \boldsymbol{X}_k^\star\|_1}{1 - (\frac{2d}{1-\rho} + \frac{d}{k^{\frac{1}{2}}})\frac{1}{\lambda}} \tag{12}$$

where $\rho \triangleq \frac{\sqrt{2}\delta}{1-\delta} \leq 1$, $\boldsymbol{X}_k^\star$ is the matrix with the k largest elements in $\boldsymbol{X}^\star$ at the corresponding indices and the others elements to be 0, $\varepsilon \sim \mathcal{O}(\sqrt{n})$ and $\alpha \triangleq \frac{4\sqrt{(I+c)(1+\delta)}}{1-\delta}$. Here I is the true intensity of the underlying measurement, i.e. $\|B(\boldsymbol{X}^\star)\|_1$, which is assumed to be known for naturally acquired measurements.
Theorem 2: *Bounds for Q2*
Given that the assumptions in section 2.2 hold, and $\boldsymbol{\Psi}$ has sufficiently small RIP constant δ, then, there exist positive absolute constants C_1, C_2, and C_3 such that if $m \geq c_1 k \log \frac{d}{k}$, and $n \geq C_1 m$ then any estimate $\widehat{\boldsymbol{X}}$ of the algorithm obeys

$$\left\| \widehat{\boldsymbol{X}} - \boldsymbol{X}^\star \right\|_F \leq C_2,$$

for all rank-one and $k \times k$-sparse matrices $\boldsymbol{X}^\star \succcurlyeq \boldsymbol{0}$ with probability exceeding $1 - e^{-C_3 n}$. The constant C_2 depends on $\sqrt{I}$ and since $\varepsilon \sim \mathcal{O}(\sqrt{n})$, we don't get a dependence on $\frac{1}{\sqrt{n}}$ like in [13]. The proof outline for these performance bounds is adapted from [20], [7], [16] and [13]. The complete proof is presented in the supplemental material [22].

4. EXPERIMENTS WITH 1-D SPARSE SIGNALS

4.1. 2 stage recovery method (Q2)

We evaluated the performance of Q2 against P2 for Poisson corrupted measurements $\boldsymbol{y}$ at various Intensity and sparsity levels of the underlying signal $\boldsymbol{x}$. $\boldsymbol{x}$ was taken to be a 64-dimensional k sparse vector, and its intensity is defined as $\|\boldsymbol{x}\|_1$. The support of $\boldsymbol{x}$ was chosen uniformly at random, and the values at those indices were randomly drawn from $\mathcal{N}(0, 1)$. The number of rows in $\boldsymbol{\Psi}$ and size of the measurement vector $\mathbf{y}$, i.e. m and n were taken to be $8k$ and $24k$ respectively in accordance with the experiments in [13]. For every intensity and sparsity level, 20 instances of a signal were generated independently. We fine tuned the ε using cross validation for every signal to get the optimal error for both P2 and Q2. For cross validation, ε was tuned on 80% of the elements of a vector and cross validated over the remaining 20% of the vector's indices. The error was calculated as $\frac{\|\boldsymbol{X}_{true} - \boldsymbol{X}_{recon}\|_F}{\|\boldsymbol{X}_{true}\|_F}$, where $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^H$. Over the 20 runs for a given intensity and sparsity level, we calculated the median errors and plotted them across varying intensity levels for a fixed sparsity level as shown in Fig.1 and across increasing number of measurements at a fixed intensity and sparsity level in Fig.2.

4.2. CPRL - Ohlsson's method (Q1)

To evaluate the performance of Q1 against P1 we performed an experiment similar to that for Q2, with all the signal dimensions same as that in the simulation for Q2. Every row of the measurement matrix $\boldsymbol{A}$ was generated as $\boldsymbol{a}_i = \boldsymbol{\Psi}^T \boldsymbol{w_i}$, where $\boldsymbol{\Psi}$ and $\boldsymbol{w_i}$ are defined as mentioned in the preliminaries. The dimensions of $\boldsymbol{\Psi}$ and the number of measurements were taken to be the same as that in the simulation for Q2. For every intensity and sparsity level, 20 instances of a signal were generated independently. We coarsely tuned λ over an array of random signals and used $\lambda = 100$ to be a constant throughout the experiment for both Q1 and P1 [10]. We fine tuned the ε using cross validation for every signal as done for Q2 to get the optimal error for both P1 and Q1. The median errors across varying intensity for a fixed sparsity level are shown in Fig.3. We also plotted the comparison of Q1 and Q2 across different intensities in Fig. 4.

Observation and Comments We observe that at low intensities, the algorithms with our VST based modification i.e. Q1 and Q2 perform better than P1 and P2, and their perfor-

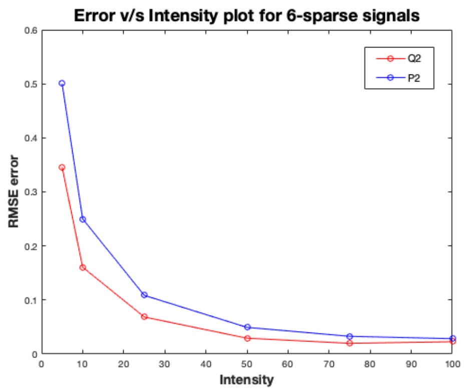

Fig. 1. Median RMSE plots for Q2 and P2 for a 6 sparse signal across increasing intensity levels

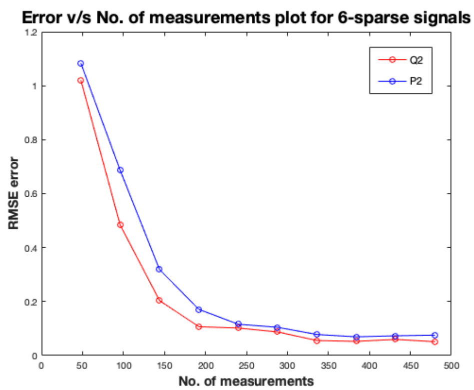

Fig. 2. Median RMSE plots for Q2 and P2 for 6 sparse signals across increasing number of measurements

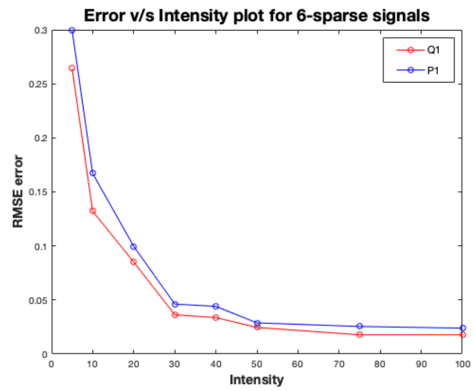

Fig. 3. Median RMSE plots for Q1 and P1 for a 6 sparse signal across increasing intensity levels

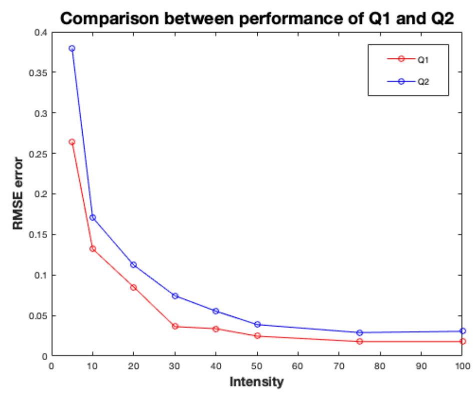

Fig. 4. Median RMSE plots for Q1 and Q2 for 6 sparse signals with increasing intensity

mances converge as the intensity increases. We also observe that Q1 performs slightly better than Q2 but Q1 is almost 7 times slower than Q2 for 64 dimensional 6 sparse signals, and that the speed worsens off with increasing signal dimension and sparsity levels. [23] also presents a similar comparison between a 2 stage recovery algorithm and P1 for signals with additive noise. Since we use the random Gaussian vectors for both Q1 and Q2, as opposed to the sensing matrices for phase retrieval in [23] we don't suffer the loss in performance of Q1. Moreover for 64 dimensional 6 sparse signals, the performance of the 2 stage recovery algorithm and the SDP-l_1 norm method's performance is expected to be similar, as shown in the simulations in [23].

5. DISCUSSION AND FUTURE WORK

We demonstrated a tractable method to achieve compressive phase retrieval from Poisson corrupted measurements at low intensities, and also proved performance bounds for both the algorithms with our proposed algorithm. There are many directions to extending this work. Since we have a square root constraint in our algorithm, obtaining a closed form solution for an ADMM type of implementation for Q1 is not possible, hence speed and scalability of this algorithm is limited. Q2 alleviates this problem to some extent, as the sparse estimation stage can be done iteratively. But Q2 works well with $\mathcal{O}(k \log(\frac{d}{k}))$ measurements only when $\mathbf{\Psi}$ is Gaussian, and under the assumptions that the sensing vectors all lie in the same subspace. The bounds presented in Q2 do not involve the $\operatorname{Tr} \Delta$ term and hence are more practical than the bounds for Q1, which need an additional relaxation to get rid of the trace term. The methods presented in this paper and otherwise work well in general with Gaussian measurement matrices, but phase retrieval from pure Fourier measurements without binary masks or coded diffraction patterns is still a hard problem. Inferring or designing dictionaries specifically for compressive phase retrieval is also an interesting research direction. Sensor saturation is yet another widely encountered phenomenon in sensor acquisitions, and phase retrieval from measurements that are both Poisson corrupted and saturated would be an interesting problem from a both theoretical and practical standpoint.

6. REFERENCES

[1] Y. Eldar and S. Mendelson, "Phase retrieval: Stability and recovery guarantees," *CoRR*, vol. abs/1211.0872, 2012.

[2] J.R. Fienup, "Reconstruction of an object from the modulus of its fourier transform," *Optics Letters*, vol. 3, 1978.

[3] G.Z. Yang, B.Z. Dong, B.Y. Gu, J.Y. Zhuang, and O.K. Ersoy, "Gerchberg–saxton and yang–gu algorithms for phase retrieval in a nonunitary transform system: a comparison," *Appl. Opt.*, , no. 2, pp. 209–218, 1994.

[4] E. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via wirtinger flow: Theory and algorithms," *CoRR*, vol. abs/1407.1065, 2014.

[5] N. Vaswani, S. Nayer, and Y. Eldar, "Low rank phase retrieval," *CoRR*, vol. abs/1608.04141, 2016.

[6] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *NIPS*, 2013.

[7] E. Candès, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *CoRR*, vol. abs/1109.4499, 2011.

[8] T. Goldstein and C. Studer, "Phasemax: Convex phase retrieval via basis pursuit," *CoRR*, vol. abs/1610.07531, 2016.

[9] E. Richard, P. Savalle, and N. Vayatis, "Estimation of simultaneously sparse and low rank matrices," in *ICML*, 2012.

[10] H. Ohlsson, A. Yang, R. Dong, and S. Sastry, "Cprl – an extension of compressive sensing to the phase retrieval problem," pp. 1367–1375, 2012.

[11] M. Moravec, J. Romberg, and R. Baraniuk, "Compressive phase retrieval," *SPIE*, 2007.

[12] L. Zhang, G. Wang, G.B. Giannakis, and J. Chen, "Compressive phase retrieval via reweighted amplitude flow," *CoRR*, vol. abs/1712.02426, 2017.

[13] S. Bahamani and J. Romberg, "Efficient compressive phase retrieval with constrained sensing vectors," in *NIPS*, 2015, pp. 523–531.

[14] M. Iwen, A. Viswanathan, and Y. Wang, "Robust sparse phase retrieval made easy," *Applied and Computational Harmonic Analysis*, vol. 42, no. 1, pp. 135 – 142, 2017.

[15] E. Candes T. Lane P. Li D. Barmherzig, J. Sun, "Holographic phase retrieval and optimal reference design," https://arxiv.org/abs/1901.06453, 2019.

[16] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 910, pp. 589 – 592, 2008.

[17] T. Blumensath and M. Davies, "Iterative hard thresholding for compressed sensing," *CoRR*, vol. abs/0805.0510, 2008.

[18] J. Curtiss, "On transformations used in the analysis of variance," *The Annals of Mathematical Statistics*, vol. 14, no. 2, pp. 107–122, 1943.

[19] C. Talegaonkar and A. Rajwade, "Performance bounds for tractable poisson denoisers with principled parameter tuning," in *IEEE GlobalSIP*, 2018.

[20] D. Garg, P. Bohra, K. S. Gurumoorthy, and A. Rajwade, "Variance stabilization based compressive inversion under poisson or poisson-gaussian noise with analytical bounds," https://www.cse.iitb.ac.in/~ajitvr/Papers/Poisson_PoissonGaussian_VST.pdf, 2018.

[21] H. Ohlsson, A. Yang, R. Dong, and S. Sastry, "Compressive phase retrieval from squared output measurements via semidefinite programming," https://arxiv.org/abs/1111.6323, 2011.

[22] C. Talegaonkar, P. Khirwadkar, and A. Rajwade, "Supplement material," https://www.cse.iitb.ac.in/~ajitvr/Chinmay-supplemental-ICIP.pdf, 2019.

[23] K. Jaganathan, S. Oymak, and B. Hassibi, "Sparse phase retrieval: Convex algorithms and limitations," *CoRR*, vol. abs/1303.4128, 2013.