Recovery guarantees for low-count Poisson and Bernoulli phase retrieval

Sjoerd Dirksen
Mathematical Institut
Utrecht University
s.dirksen@uu.nl

Felix Krahmer
Department of Mathematics
Technical University of Munich
felix.krahmer@tum.de

Patricia Römer
Department of Mathematics
Technical University of Munich
patricia.roemer@tum.de

Palina Salanevich
Mathematical Institut
Utrecht University
p.salanevich@uu.nl

***Abstract*—In low-dose microscopy, noisy phaseless measurements can be modeled as Poisson or Bernoulli random variables. In this paper, we propose a spectral method for this phase retrieval setup and derive reconstruction guarantees for both noise models with Gaussian measurements. Furthermore, we study how the reconstruction error depends on the radiation dose used for the measurements and on the oversampling ratio.**

***Index Terms*—phase retrieval, spectral method, low-dose imaging, Poisson noise, quantized measurements**

I. Introduction

In X-ray and electron microscopy, the goal is to image small-scale biological objects and study their structure and properties. In such experiments, the object is illuminated with a focused radiation beam and a (pixelated) detector placed behind the object counts the arriving radiation particles scattered from the object. The observed diffraction measurements can be understood as the intensities of the incoming radiation, comprising the information on the measured object.

More formally, we aim to recover the object of interest $x \in \mathbb{R}^n$ from its phaseless measurements of the form

$$y_j \approx |\langle a_j, x\rangle|^2, \quad j \in [m],$$

with measurement vectors $a_j \in \mathbb{R}^n$. This quadratic inverse problem is commonly known as the *phase retrieval problem*.

In many imaging experiments, the objects, such as proteins or various cell materials, may be highly sensitive to the received dose (see, e.g., [6]). In these cases, the measurements are performed in a low-dose regime and each detector pixel essentially measures the received illumination particle count. Due to the discrete nature of this process, the measurements at the detector pixels can be modeled as realizations of independent Poisson random variables

$$Y_j^P \sim \textit{Poisson}\big(|\langle a_j, x\rangle|^2\big). \tag{P}$$

Note that, for notational simplicity, we work with a fixed normalization of the measurement vectors, so the dose is incorporated in x and corresponds to $\|x\|_2^2$.

PS is supported by NWO Talent program Veni ENW grant, file number VI.Veni.212.176. The work for this paper was initiated during a discussion at the Lorentz Center workshop 'Phase Retrieval in Mathematics and Applications' (PRiMA). SD, FK, and PS would like to thank the Lorentz Center for its hospitality and support.

A widely used approach to the phase retrieval problem with Poisson noise is based on likelihood maximization. Common strategies to solve the resulting non-convex optimization problem include ADMM [2] or Wirtinger flow algorithms [3]. While these and similar methods yield a high signal-to-noise ratio in the case when the values $|\langle a_j, x\rangle|^2$ are large, their performance is limited for extremely low doses [4], [8], [9], [12]. In this setting, the values $|\langle a_j, x\rangle|^2$ are very small and the Poisson distributions of Y_j^P can be well approximated by a Bernoulli distribution. That is, one can model the measurements y_j as realizations of Bernoulli random variables

$$Y_j^B \sim \textit{Bernoulli}\left(1 - \exp(-|\langle a_j, x\rangle|^2)\right). \tag{B}$$

The measurement model (B) is further motivated by *4D-scanning transmission electron microscopy* (4D-STEM), in which acquisition is performed with event-driven detectors that have capacity to record only the first particle arrival [7].

In this paper, we propose and analyze a new approach for solving phase retrieval for both Poisson and Bernoulli measurement models and study how the high-count information lost in the Bernoulli model influences the reconstruction accuracy across various dose values.

II. Problem formulation

The Bernoulli measurements can be interpreted as noisy one-bit representations of the squared phaseless measurements. Consequently, the theory for one-bit compressed sensing can serve as an inspiration for reconstruction algorithms. We follow the idea of [13] to seek a vector $z \in \mathbb{R}^n$ whose noiseless measurements $|\langle a_j, z\rangle|^2$ maximize the correlation with the noisy measurements y_j.

In analogy, for both measurement models (P) and (B), we aim to obtain a good approximation for the ground-truth object x by solving the constrained optimization problem

$$\begin{aligned} &\text{maximize} \quad f_x(z) \\ &\text{subject to} \quad \|z\|_2^2 = \alpha, \end{aligned} \tag{1}$$

with the objective function $f_x : \mathbb{R}^n \to \mathbb{R}$ given by

$$f_x(z) = \frac{1}{m}\sum_{j=1}^m y_j |\langle a_j, z\rangle|^2.$$

The constraint $\|z\|_2^2 = \alpha$ here specifies the dose α, which is a part of the experimental setup and is thus known in advance.

Since

$$\frac{1}{m}\sum_{j=1}^{m} y_j|\langle a_j, z\rangle|^2 = z^{\mathrm{T}}\left(\frac{1}{m}\sum_{j=1}^{m} y_j a_j a_j^{\mathrm{T}}\right) z,$$

the solution to (1) is the leading eigenvector of the matrix

$$Y := \frac{1}{m}\sum_{j=1}^{m} y_j a_j a_j^{\mathrm{T}}, \tag{2}$$

scaled according to the norm constraint.

Interestingly, computing the leading eigenvector of this matrix Y is exactly the spectral method proposed in [1]. The spectral method is motivated by the observation proved in [1] that for Gaussian random measurement vectors $a_j \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I_n)$ and the corresponding noise-free phaseless measurements $y_j = |\langle a_j, x\rangle|^2$, the random matrix Y concentrates around its expectation $2xx^{\mathrm{T}} + \|x\|_2^2 \cdot I_n$, whose leading eigendirection is exactly parallel to the underlying signal x.

While we arrive at the same strategy through a different motivation, the resulting algorithm is the same. Consequently, the theoretical analysis can be approached analogously to [1], and the main challenge is to incorporate the effect of the Poisson (P) and the Bernoulli distribution (B), respectively.

For the high-dose scenario, recovery guarantees have been derived in [3] for a truncated version of the spectral method; see [10], [11] for an asymptotic analysis of such a truncated spectral method in the case of more general noise models that also include the Poisson model as a special case.

More precisely, a truncation $\mathcal{T}(y) = y\mathbb{1}_{y\le t}$ is applied to the observations, so that Y is replaced by

$$\frac{1}{m}\sum_{j=1}^{m} \mathcal{T}(y_j) a_j a_j^{\mathrm{T}}.$$

For an appropriate truncation, the resulting method is shown to recover the ground-truth with a sample complexity of $m = \mathcal{O}(n)$. However, the results do not extend to the low-dose setting, but only apply for $\|x\|_2 \ge \log^{1.5}(m)$. In this paper, we show that we can go beyond this limiting minimal dose condition with a sampling complexity only increased by logarithmic factors.

III. Recovery guarantees

In this section, we adopt the following notation and assumptions. Let $x \in \mathbb{R}^n$ be the measured object renormalized by the radiation dose parameter $\alpha := \|x\|_2^2$ and assume that the measurement vectors $a_j,\ j \in [m]$, are drawn independently from the standard normal distribution $\mathcal{N}(0, I_n)$. For the phaseless measurements $\{y_j\}_{j=1}^m$ drawn from the Poisson (P) or Bernoulli (B) observation models, we define the matrix Y as in (2). We denote by x_0 the eigenvector corresponding to the largest eigenvalue of Y, normalized so that $\|x_0\|_2^2 = \alpha$.

A. Poisson measurements

We start by evaluating how well x_0 approximates the ground-truth object x in the case of the Poisson measurement model. For this purpose, we define the phaseless distance between two vectors $u, v \in \mathbb{R}^n$ as

$$\mathit{dist}(u, v) = \min_{\gamma\in\{-1,1\}} \|u - \gamma v\|_2.$$

To derive the recovery guarantees, we need the following technical lemma describing the properties of the matrix Y in the Poisson case. The proof of this lemma can be found in [5].

Lemma 1. *If $y_j,\ j \in [m]$, are distributed according to the Poisson model* (P)*, then*

$$\mathbb{E}[Y] = 2xx^{\mathrm{T}} + \|x\|_2^2 \cdot I_n.$$

Furthermore, for every $\delta > 0$, there exist constants $C, \hat{C}, c > 0$ depending only on δ, such that, if $m \ge n \ge c$, with probability at least $1 - \delta$, it holds that

$$\|Y - \mathbb{E}[Y]\| \le C\left(\|x\|_2^2 + \hat{C}\right)\sqrt{\log(n)}\log(m)\sqrt{\frac{n}{m}}.$$

Using this deviation bound, we can prove the following recovery guarantee for phaseless Poisson measurements.

Theorem 2. *If $y_j,\ j \in [m]$, are distributed according to the Poisson model* (P)*, then, for every $\delta > 0$, there exist constants $C, \hat{C}, c > 0$ depending only on δ, such that, if $m \ge n \ge c$, with probability at least $1 - \delta$, the vector x_0 satisfies*

$$\frac{\mathit{dist}(x_0, x)^2}{\|x\|_2^2} \le 2C\left(1 + \frac{\hat{C}}{\|x\|_2^2}\right)\sqrt{\log(n)}\log(m)\sqrt{\frac{n}{m}}.$$

Proof. We aim to bound the phaseless distance

$$\begin{aligned}\mathit{dist}(x_0, x)^2 &= \min_{\gamma\in\{-1,1\}} \|x_0 - \gamma x\|_2^2 \\ &= \min_{\gamma\in\{-1,1\}} \left(\|x_0\|_2^2 - 2\gamma\langle x_0, x\rangle + \|x\|_2^2\right) \\ &= \|x_0\|_2^2 - 2|\langle x_0, x\rangle| + \|x\|_2^2 \\ &= 2\alpha - 2|\langle x_0, x\rangle|.\end{aligned}$$

Thus, we need to establish a lower bound for $|\langle x_0, x\rangle|$. We proceed similarly to [1], using that x_0 is the eigenvector corresponding to the largest eigenvalue λ_0 of Y normalized so that $\|x_0\|_2^2 = \alpha$. We use that

$$\begin{aligned}&\left|\alpha\lambda_0 - 2|\langle x, x_0\rangle|^2 - \alpha\|x_0\|_2^2\right| \\ &= \left|x_0^{\mathrm{T}} Y x_0 - x_0^{\mathrm{T}}\left(2xx^{\mathrm{T}} + \|x\|_2^2 \cdot I_n\right)x_0\right| \\ &= \left|x_0^{\mathrm{T}}(Y - \mathbb{E}[Y])x_0\right| \\ &\le \|x_0\|_2^2\,\|Y - \mathbb{E}[Y]\|.\end{aligned}$$

Setting $\delta \coloneqq \|Y - \mathbb{E}[Y]\|$, the inequality above yields

$$|\langle x_0, x\rangle|^2 \geq \frac{1}{2}\left(\alpha\lambda_0 - \alpha^2 - \alpha\delta\right).$$

Further, the largest eigenvalue λ_0 of Y satisfies

$$\begin{aligned}\lambda_0 &\geq \frac{1}{\|x\|_2^2} x^{\mathrm{T}} Y x \\ &= \frac{1}{\|x\|_2^2} x^{\mathrm{T}}\left(Y - \mathbb{E}[Y] + \mathbb{E}[Y]\right) x \\ &\geq -\delta + \frac{1}{\|x\|_2^2} x^{\mathrm{T}} \mathbb{E}[Y] x \\ &= -\delta + 3\alpha,\end{aligned}$$

so that

$$|\langle x_0, x\rangle|^2 \geq \frac{1}{2}\left(2\alpha^2 - 2\alpha\delta\right),$$

and we conclude

$$\frac{dist(x_0, x)^2}{\alpha} \leq 2 - 2\sqrt{1 - \frac{1}{\alpha}\cdot\delta} \leq 2\frac{\delta}{\alpha}$$

if $\delta \leq \alpha$. If $\delta > \alpha$, the same upper bound holds for the relative reconstruction error since $|\langle x_0, x\rangle|^2 \geq 0$. Using Lemma 1 to estimate δ yields the result. □

Remark 3. *From the bound in Theorem 2, we conclude that*

$$\frac{m}{\log(m)^2} = \mathcal{O}\left(n\log(n)\cdot\left(1 + \frac{1}{\alpha}\right)^2\right)$$

measurements are required to recover the ground-truth object up to a small constant relative error from Poisson phaseless measurements by solving (1).

B. Bernoulli measurements

We now switch to the study of the case when the phaseless measurements y_j are truncated to $\{0, 1\}$ values and distributed according to the Bernoulli model (B). Similarly to the Poisson case, we start with a technical lemma, the proof of which can be found in [5].

Lemma 4. *If $y_j,\ j \in [m]$, are distributed according to the Bernoulli model* (B)*, it holds that*

$$\mathbb{E}[Y] = \frac{2}{\left(2\|x\|_2^2 + 1\right)^{\frac{3}{2}}}\cdot xx^{\mathrm{T}} + \left(1 - \frac{1}{\left(2\|x\|_2^2 + 1\right)^{\frac{1}{2}}}\right)\cdot I_n.$$

Furthermore, for every $\delta > 0$, there exist constants $C, c > 0$ depending only on δ, such that, if $m \geq n \geq c$, with probability at least $1 - \delta$, it holds that

$$\|Y - \mathbb{E}[Y]\| \leq C\sqrt{\log(n)}\sqrt{\frac{n}{m}}.$$

Using this deviation bound, we prove the following recovery guarantee for phaseless Bernoulli measurements.

Theorem 5. *If $y_j,\ j \in [m]$, are distributed according to the Bernoulli model* (B)*, then, for every $\delta > 0$, there exist constants $C, c > 0$ depending only on δ, such that, if $m \geq n \geq c$, with probability at least $1 - \delta$, the vector x_0 satisfies*

$$\frac{dist(x_0, x)^2}{\|x\|_2^2} \leq 2C\cdot\frac{\left(2\|x\|_2^2 + 1\right)^{\frac{3}{2}}}{\|x\|_2^2}\sqrt{\log(n)}\sqrt{\frac{n}{m}}.$$

Proof. With calculations analogous to the proof of Theorem 2, we obtain that $|\langle x_0, x\rangle|^2$ is bounded from below by

$$\frac{(2\alpha + 1)^{\frac{3}{2}}}{2}\left(\alpha\lambda_0 - \left(1 - \frac{1}{(2\alpha + 1)^{\frac{1}{2}}}\right)\|x_0\|_2^2 - \alpha\delta\right)$$

and

$$\lambda_0 \geq -\delta + \alpha\cdot\frac{2}{(2\alpha + 1)^{\frac{3}{2}}} + \left(1 - \frac{1}{(2\alpha + 1)^{\frac{1}{2}}}\right),$$

so that

$$\begin{aligned}|\langle x_0, x\rangle|^2 &\geq \frac{(2\alpha + 1)^{\frac{3}{2}}}{2}\left(\alpha^2\cdot\frac{2}{(2\alpha + 1)^{\frac{3}{2}}} - 2\alpha\delta\right) \\ &= \alpha^2 - \alpha\left(2\alpha + 1\right)^{\frac{3}{2}}\delta.\end{aligned}$$

Thus, we obtain

$$\frac{dist(x_0, x)^2}{\alpha} \leq 2 - 2\sqrt{1 - \frac{(2\alpha + 1)^{\frac{3}{2}}}{\alpha}\cdot\delta} \leq 2\delta\frac{(2\alpha + 1)^{\frac{3}{2}}}{\alpha}$$

if $\delta \leq \alpha\left(2\alpha + 1\right)^{-\frac{3}{2}}$. If $\delta > \alpha\left(2\alpha + 1\right)^{-\frac{3}{2}}$, the same upper bound holds due to $|\langle x_0, x\rangle|^2 \geq 0$. Applying Lemma 4 completes the proof. □

Remark 6. *From Theorem 5, we conclude that*

$$m = \mathcal{O}\left(n\log(n)\cdot\frac{(2\alpha + 1)^3}{\alpha^2}\right)$$

measurements are required to recover the ground-truth object up to a small constant relative error from Bernoulli phaseless measurements by solving (1).

We note that the factor $\frac{(2\alpha+1)^3}{\alpha^2}$ is minimal for $\alpha = 1$ and its contribution can be interpreted as follows. For very small as well as for rather large values of α, we need a lot of oversampling. In the former case, this is due to the small amount of information captured by the measurements. This phenomenon is naturally the same for Poisson observations. In the latter case, this is due to a loss of information because of the truncation in the Bernoulli model.

IV. Numerical Experiments

To illustrate the findings of Section III, we numerically compute the relative reconstruction error

$$\frac{\min\left\{\|x - \hat{x}\|_2^2, \|x + \hat{x}\|_2^2\right\}}{\|x\|_2^2} \tag{3}$$

for obtained reconstructions $\hat{x} \in \mathbb{R}^n$ across various values of the oversampling ratio $\frac{m}{n}$ and the dose scale $\alpha = \|x\|_2^2$. For our numerical experiments, we set the object dimension to

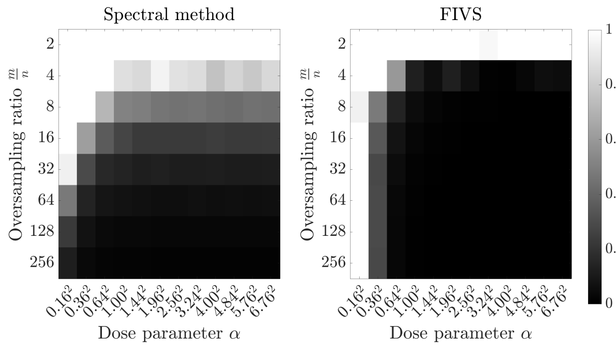

Fig. 1: The dependence of the relative reconstruction error (3) on the dose and the oversampling ratio for the Poisson measurement model (P). Left: the proposed spectral method (1); Right: the proposed method with subsequent FIVS iterations. The plots show averages over 10 random trials.

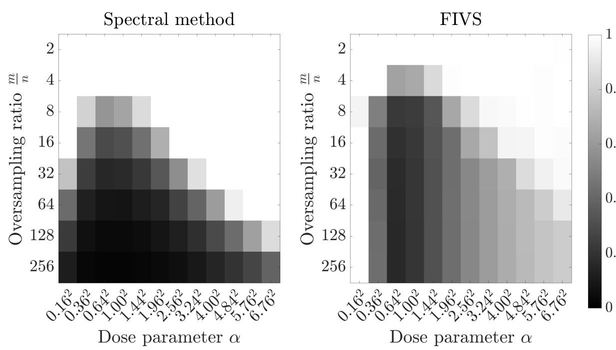

Fig. 2: The dependence of the relative reconstruction error (3) on the dose and the oversampling ratio for the Bernoulli measurement model (B). Left: the proposed spectral method (1); Right: the proposed method with subsequent FIVS iterations. The plots show averages over 10 random trials.

be $n = 256$, and the measurement vectors $a_j \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, I_n)$, $j \in [m]$.

We compare the solution x_0 obtained using the proposed spectral method (1) to the one that uses x_0 as initialization, followed by $k = 100$ iterations of the FIVS method (Wirtinger flow using a Gaussian loss after variance stabilization) that is proposed in [4] for low-dose Poisson phase retrieval. The results of the numerical simulations for the Poisson and Bernoulli measurement models are shown in Figures 1 and 2, respectively. In the case of Poisson phase retrieval, we see that the reconstruction error decreases with higher dose and oversampling ratio values, as predicted by Theorem 2. Furthermore, running iterations of FIVS improves the reconstruction result. In the case of Bernoulli measurements, we see that the reconstruction error deteriorates when the dose α gets large, indicating that too much information is lost in the measurement truncation. FIVS iterations improve the reconstruction here only in the low oversampling ratio regime, otherwise the spectral initialization shows better performance. This can be explained by the fact that FIVS is proven to converge to a stationary point of a non-convex measurement error metric and hence can deteriorate the estimate in terms of the relative reconstruction error (3) if the initialization falls within the basin of attraction of a poor local minimum. We also note that the numerical experiments indicate successful reconstruction for much lower oversampling ratio values than predicted by Theorems 2 and 5.

V. Discussion

We studied the performance of the spectral method for Poisson and Bernoulli measurements, focusing on the influence of the oversampling ratio and the dose value on the reconstruction error. In particular, in the highly truncated case of Bernoulli measurements, we established a trade-off between the oversampling ratio and the achievable dose value. Further research is needed to refine the proposed method. For instance, while the truncation rule in [3] does not apply in the low-count regime, it would be interesting to study reconstruction properties for this type of data preprocessing, expanding upon the results we obtained for the non-truncated spectral method.

References

[1] Emmanuel J Candès, Xiaodong Li, and Mahdi Soltanolkotabi. Phase retrieval via Wirtinger flow: Theory and algorithms. *IEEE Transactions on Information Theory*, 61(4):1985–2007, 2015.

[2] Huibin Chang, Yifei Lou, Yuping Duan, and Stefano Marchesini. Total variation–based phase retrieval for Poisson noise removal. *SIAM journal on imaging sciences*, 11(1):24–55, 2018.

[3] Yuxin Chen and Emmanuel J Candès. Solving random quadratic systems of equations is nearly as easy as solving linear systems. *Communications on pure and applied mathematics*, 70(5):822–883, 2017.

[4] Benedikt Diederichs, Frank Filbir, and Patricia Römer. Wirtinger gradient descent methods for low-dose Poisson phase retrieval. *Inverse Problems*, 40(12):125030, 2024.

[5] Sjoerd Dirksen, Felix Krahmer, Patricia Römer, and Palina Salanevich. Spectral method for low-dose Poisson and Bernoulli phase retrieval. *arXiv preprint arXiv:2502.13263*, 2025.

[6] Robert M Glaeser. Limitations to significant information in biological electron microscopy as a result of radiation damage. *Journal of ultrastructure research*, 36(3-4):466–482, 1971.

[7] Daen Jannis, Christoph Hofer, Chuang Gao, Xiaobin Xie, Armand Béché, Timothy J Pennycook, and Jo Verbeeck. Event driven 4D STEM acquisition with a Timepix3 detector: Microsecond dwell time and faster scans for high precision and low dose applications. *Ultramicroscopy*, 233:113423, 2022.

[8] Felix Krahmer and Patricia Römer. A one-bit quantization approach for low-dose Poisson phase retrieval. In *6th International Workshop on the Theory of Computational Sensing and its applications to Radar, Multimodal Sensing, and Imaging (CoSeRa)*. IEEE, 2024.

[9] Zongyu Li, Kenneth Lange, and Jeffrey A Fessler. Poisson phase retrieval in very low-count regimes. *IEEE Transactions on computational imaging*, 8:838–850, 2022.

[10] Yue M Lu and Gen Li. Phase transitions of spectral initialization for high-dimensional non-convex estimation. *Information and Inference: A Journal of the IMA*, 9(3):507–541, 2020.

[11] Wangyu Luo, Wael Alghamdi, and Yue M Lu. Optimal spectral initialization for signal recovery with applications to phase retrieval. *IEEE Transactions on Signal Processing*, 67(9):2347–2356, 2019.

[12] Philipp Michael Pelz, Wen Xuan Qiu, Robert Bücker, Günther Kassier, and RJ Dwayne Miller. Low-dose cryo electron ptychography via non-convex Bayesian optimization. *Scientific reports*, 7(1):9883, 2017.

[13] Yaniv Plan and Roman Vershynin. Robust 1-bit compressed sensing and sparse logistic regression: A convex programming approach. *IEEE Transactions on Information Theory*, 59(1):482–494, 2012.