Poisson-Gaussian Holographic Phase Retrieval with Score-based Image Prior

Zongyu Li*, **Jason Hu***, **Xiaojian Xu, Liyue Shen, and Jeffrey A. Fessler**
EECS Department
University of Michigan
zonyul,jashu,xjxu,liyues,fessler@umich.edu *

Abstract

Phase retrieval (PR) is a crucial problem in many imaging applications. This study focuses on resolving the holographic phase retrieval problem in situations where the measurements are affected by a combination of Poisson and Gaussian noise, which commonly occurs in optical imaging systems. To address this problem, we propose a new algorithm called "AWFS" that uses the accelerated Wirtinger flow (AWF) with a score function as generative prior. We calculate the gradient of the log-likelihood function for PR and provide an implementable estimate for it. Additionally, we introduce a generative prior in our regularization framework by using score matching to capture information about the gradient of image prior distributions. The results of our simulation experiments on three different datasets show the following. 1) By using the PG likelihood model, the proposed algorithm improves reconstruction compared to algorithms based solely on Gaussian or Poisson likelihood. 2) The proposed score-based image prior method leads to improved reconstruction quality over the method based on denoising diffusion probabilistic model (DDPM), as well as plug-and-play alternating direction method of multipliers (PnP-ADMM) and regularization by denoising (RED).

1 Introduction

Poisson-Gaussian phase retrieval (PR) is a nonlinear inverse problem, where the goal is to recover a signal from the (square of) magnitude-only measurements that are corrupted by both Poisson and Gaussian noise [1]. The measured pattern is roughly proportional to the square of Fourier transform magnitude of electric field associated with the illuminated objects [2, 3]. Recovering the structure of the sample from its diffraction pattern is a nonlinear inverse problem known as holographic PR. To solve this problem, maximum a posterior (MAP) estimation uses the following form:

$$\hat{\boldsymbol{x}} = \underset{\boldsymbol{x}\in\mathbb{R}^N}{\arg\max}\, p(\boldsymbol{x}|\boldsymbol{y}, \bar{\boldsymbol{b}}, \boldsymbol{A}, \boldsymbol{r}) = \underset{\boldsymbol{x}\in\mathbb{R}^N}{\arg\min}\, g(\boldsymbol{x}; \boldsymbol{A}, \boldsymbol{y}, \bar{\boldsymbol{b}}, \boldsymbol{r}) + h(\boldsymbol{x}),$$

where $\boldsymbol{x}$ denotes the latent image to recover, $\boldsymbol{y}$ is the recorded measurement vector, $\bar{\boldsymbol{b}}$ denotes the mean of background measurements, and $\boldsymbol{A} \in \mathbb{C}^{M\times N}$ denotes the corresponding system matrix in holographic PR, where M denotes the number of measurements and N denotes the dimension of $\boldsymbol{x}$. The known reference image $\boldsymbol{r}$ provides additional information to reduce the ambiguity of $\hat{\boldsymbol{x}}$. Following Bayes' rule, we denote $g(\boldsymbol{x}) = -\log p(\boldsymbol{y}, \boldsymbol{A}, \boldsymbol{r}|\boldsymbol{x})$ and $h(\boldsymbol{x}) = -\log p(\boldsymbol{x})$ as the data fidelity term and the regularization term, respectively. In practical scenarios, the measurements $\boldsymbol{y}$ are contaminated by both Poisson and Gaussian (PG) noise. Because the PG likelihood is complicated, most previous works [4–34] approximate the noise by only a Gaussian or Poisson. The regularizer

*Z. Li and J. Hu contributed equally to this work.

37th Conference on Neural Information Processing Systems (NeurIPS 2023), Vancouver, Canada.

$h(\boldsymbol{x})$ provides prior information about underlying object characteristics. Generative model-based priors are commonly used for this regularizer [35, 36]. Recently, score-based diffusion models have gained significant traction for image generation [37–40]. These models estimate the gradients of data distribution and can be used as plug-and-play priors for inverse problems [41] such as image deblurring and MRI and CT reconstruction [42–47].

2 Methods

Wirtinger Flow. We model the system matrix $\boldsymbol{A}$ by the (oversampled and scaled) discrete Fourier transform applied to a concatenation of the sample $\boldsymbol{x}$, a blank image (representing the holographic separation condition [26]) and a known reference image $\boldsymbol{r}$, so that $\boldsymbol{y}$ follows the Poisson plus Gaussian distribution:

$$\boldsymbol{y} \sim \mathcal{N}(\text{Poisson}\left(|\boldsymbol{A}(\boldsymbol{x})|^2 + \bar{\boldsymbol{b}}\right), \sigma^2 \boldsymbol{I}), \boldsymbol{A}(\boldsymbol{x}) \triangleq \alpha \mathcal{F}\{[\boldsymbol{x}, \boldsymbol{0}, \boldsymbol{r}]\}. \tag{1}$$

Here σ^2 denotes the variance of Gaussian noise, and α denotes a scaling factor (quantum efficiency, conversion gain, etc.) after applying the Fourier transform. Therefore, the negative log-likelihood can be derived through the convolution of the Gaussian and Poisson distributions using (1) to be $g_{\text{PG}}(\boldsymbol{x}) = \sum_{i=1}^{M} g_i(\boldsymbol{x})$, where

$$g_i(\boldsymbol{x}) \triangleq -\log\left(\sum_{n=0}^{\infty} \frac{e^{-\left(|\boldsymbol{a}_i'\boldsymbol{x}|^2 + \bar{b}_i\right)} \cdot \left(|\boldsymbol{a}_i'\boldsymbol{x}|^2 + \bar{b}_i\right)^n}{n!} \cdot \frac{e^{-\left(\frac{(y_i - n)^2}{\sqrt{2}\sigma}\right)}}{\sqrt{2\pi\sigma^2}}\right).$$

Here M denotes the length of $\boldsymbol{y}$ and $\boldsymbol{a}_i'$ denotes the ith row of $\boldsymbol{A}$ (since $\boldsymbol{A}$ is linear). We opt to use WF for estimating $\boldsymbol{x}$ because it is commonly used in practice due to its simplicity and efficiency [19]. The WF algorithm is based on the gradient:

$$\begin{aligned}
&\nabla g_{\text{PG}}(\boldsymbol{x}) = 2\boldsymbol{A}' \operatorname{diag}\{\phi_i(|\boldsymbol{a}_i'\boldsymbol{x}|^2 + b_i; y_i)\}\boldsymbol{A}\boldsymbol{x}, \qquad (2)\\
&\phi(u; v) \triangleq 1 - \frac{s(u, v-1)}{s(u, v)}, \quad s(a, b) \triangleq \sum_{n=0}^{\infty} \frac{a^n}{n!} e^{-\left(\frac{b-n}{\sqrt{2}\sigma}\right)^2}.
\end{aligned}$$

One can show that $\nabla g_{\text{PG}}(\boldsymbol{x})$ is Lipschitz continuous so that a finite sum can be used to approximate it.

Accelerated Wirtinger Flow with Score-based Image Prior. We followed the implementation of [48] for the accelerated WF algorithm. Assuming we have a score function $s_{\boldsymbol{\theta}}(\boldsymbol{x}, \boldsymbol{\sigma})$ that was trained by score matching [37, 49]:

$$\begin{aligned}
&\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{k=1}^{K} \mathbb{E}_{\boldsymbol{x}, \tilde{\boldsymbol{x}}} \left[\left\| s_{\boldsymbol{\theta}}(\boldsymbol{x}, \sigma_k) - \frac{\boldsymbol{x} - \tilde{\boldsymbol{x}}}{\sigma_k^2} \right\|_2^2 \right],\\
&\text{where} \quad \boldsymbol{x} \sim p(\boldsymbol{x}), \quad \tilde{\boldsymbol{x}} \sim \boldsymbol{x} + \mathcal{N}(0, \sigma_k^2 \boldsymbol{I}). \qquad (3)
\end{aligned}$$

The gradient descent algorithm for MAP estimation (1) has the form: $\boldsymbol{x}_{t+1} = \boldsymbol{x}_t - \mu(\nabla g(\boldsymbol{x}_t) + s_{\boldsymbol{\theta}}(\boldsymbol{x}_t, \sigma_k))$. Algorithm 1 summarizes our proposed AWFS algorithm. In a similar fashion as Langevin dynamics, we choose σ_k to be a descending scale of noise levels. The step size factor β in Algorithm 1 can be selected empirically, but one can show that the Lipschitz constant of the gradient $\nabla g_{\text{PG}}(\boldsymbol{x}_t) + s_{\boldsymbol{\theta}}(\boldsymbol{x}_t, \sigma_k)$ exists for any σ_k. Hence with sufficiently small step size β, the inner sequence $\boldsymbol{x}_{t,k}$ generated by Algorithm 1 will converge as $t \to \infty$ to a critical point of the posterior distribution $p_{\sigma_k}(\boldsymbol{x}|\boldsymbol{A}, \boldsymbol{y}, \bar{\boldsymbol{b}}, \boldsymbol{r}) \propto p(\boldsymbol{y}|\boldsymbol{A}, \boldsymbol{x}, \bar{\boldsymbol{b}}, \boldsymbol{r}) p_{\sigma_k}(\boldsymbol{x})$ [48].

3 Experiment

We tested the algorithms on datasets of histopathology images [50], celebrity faces [51], and CT-density dataset. For comparison, we implemented unregularized Gaussian WF, Poisson WF and Poisson-Gaussian WF, smoothed total variation (TV) based on the Huber function [52, p. 184] and

Algorithm 1 Proposed accelerated WF with score-based image prior.

Require: Measurement $\boldsymbol{y}$, system matrix $\boldsymbol{A}$, momentum factor $\eta_0 = 1$, step size factor $\beta > 0$, weighting factor $0 < \gamma < 1$, truncation operator $\mathcal{P}_C(\cdot) \to [0, C]$; initial image $\boldsymbol{x}_0$, initial auxiliary variables $\boldsymbol{z}_0 = \boldsymbol{w}_0 = \boldsymbol{v}_0 = \boldsymbol{x}_0$, initialize $\sigma_1 > \sigma_2 > \cdots > \sigma_K \geq 0$.
  for $k = 1 : K$ **do**
    for $t = 1 : T$ **do**
      Set step size $\mu = \beta \sigma_k^2$.
      Set $\Delta z_{t,k} = \frac{\eta_{t-1,k}}{\eta_{t,k}} (\boldsymbol{z}_{t,k} - \boldsymbol{x}_{t,k})$.
      Set $\Delta x_{t,k} = \frac{\eta_{t-1,k} - 1}{\eta_{t,k}} (\boldsymbol{x}_{t,k} - \boldsymbol{x}_{t-1,k})$.
      Set $\boldsymbol{w}_{t,k} = \mathcal{P}_C\left(\boldsymbol{x}_{t,k} + \Delta z_{t,k} + \Delta x_{t,k}\right)$.
      Compute $s_{\boldsymbol{\theta}}(\boldsymbol{x}_{t,k}, \sigma_k)$ and $s_{\boldsymbol{\theta}}(\boldsymbol{w}_{t,k}, \sigma_k)$.
      Set $\boldsymbol{z}_{t+1,k} = \boldsymbol{w}_{t,k} - \mu\left(\nabla g_{\mathrm{PG}}(\boldsymbol{w}_{t,k}) + s_{\boldsymbol{\theta}}(\boldsymbol{w}_{t,k}, \sigma_k)\right)$.
      Set $\boldsymbol{v}_{t+1,k} = \boldsymbol{x}_{t,k} - \mu\left(\nabla g_{\mathrm{PG}}(\boldsymbol{x}_{t,k}) + s_{\boldsymbol{\theta}}(\boldsymbol{x}_{t,k}, \sigma_k)\right)$.
      Set $\eta_{t+1,k} = \frac{1}{2}\left(1 + \sqrt{1 + 4\eta_{t,k}^2}\right)$.
      Set $\boldsymbol{x}_{t+1,k} = \mathcal{P}_C\left(\gamma_{t,k} \boldsymbol{z}_{t+1,k} + (1 - \gamma_{t,k}) \boldsymbol{v}_{t+1,k}\right)$.
    end for
  end for
  Return $\boldsymbol{x}_{T,K}$.

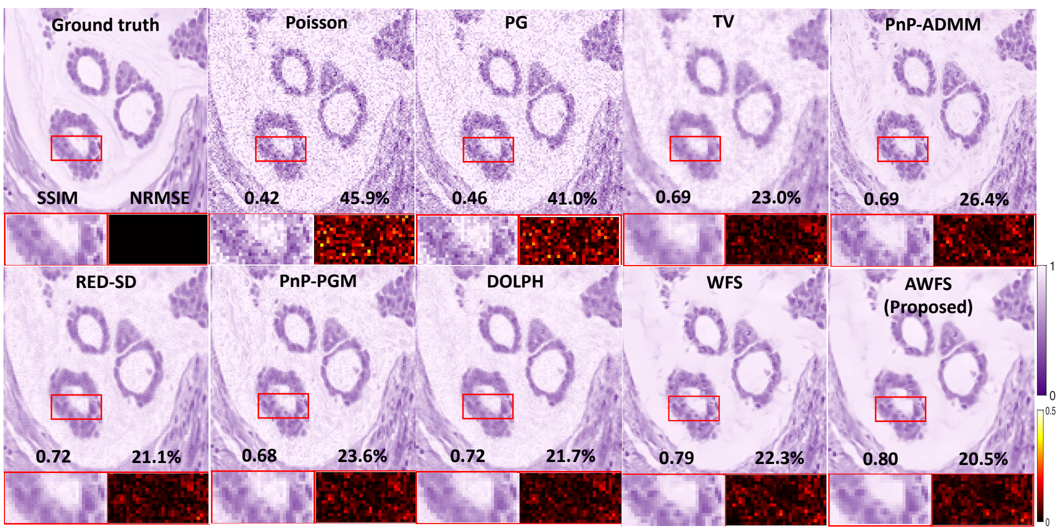

Figure 1: Reconstructed images on histopathology dataset [50]. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image. α and σ were set to 0.035 and 1, respectively.

PnP/RED methods with the DnCNN denoiser [53]: PnP-ADMM [54], PnP-PGM [55], and RED-SD [56]. We also implemented the RED-SD algorithm with "Noise2Self" zero-shot image denoising network [57] (RED-SD-SELF). For diffusion models, we implemented DOLPH [58] and our proposed AWFS.

Results. We compared all implemented algorithms by computing the normalized root mean square error (NRMSE) and structural similarity index measure (SSIM). Due to the global phase ambiguity, i.e., all the algorithms can recover the signal only to within a constant phase shift due to the loss of global phase information, we corrected the phase of $\hat{\boldsymbol{x}}$ by $\hat{\boldsymbol{x}}_{\text{corrected}} \triangleq \operatorname{sign}(\langle \hat{\boldsymbol{x}}, \boldsymbol{x}_{\text{true}} \rangle)\, \hat{\boldsymbol{x}}$.

Fig. 1 visualizes reconstructed images on the histopathology dataset [50]. The WF with PG likelihood outperforms WF with Poisson likelihood with a consistently higher SSIM and lower NRMSE. Of the regularized algorithms with PG likelihood, our proposed AWFS had less visual noise and achieved the best detail recovery. For quantitative evaluations, Table 1 shows that in all cases, usage

Table 1: SSIM and NRMSE for Poisson and Poisson-Gaussian likelihoods. Results were averaged across 7 different noise levels by varying $\alpha \in 0.02 : 0.005 : 0.035$ in (1).

Likelihood	Unregularized (SSIM/NRMSE)		DOLPH (SSIM/NRMSE)		AWFS (SSIM/NRMSE)	
DataSet: Histopathology [50]						
Poisson	0.54 ± 0.18	31.7 ± 10.2	0.72 ± 0.13	19.5 ± 6.1	0.83 ± 0.06	16.2 ± 3.7
Poisson-Gaussian	0.57 ± 0.18	28.9 ± 9.0	0.80 ± 0.06	16.0 ± 2.9	$\mathbf{0.85 \pm 0.05}$	$\mathbf{15.4 \pm 3.7}$
DataSet: CelebA [51]						
Poisson	0.39 ± 0.10	24.5 ± 11.4	0.61 ± 0.12	15.6 ± 10.6	0.72 ± 0.16	15.2 ± 11.8
Poisson-Gaussian	0.42 ± 0.10	21.8 ± 9.1	0.71 ± 0.11	$\mathbf{13.7 \pm 11.1}$	$\mathbf{0.74 \pm 0.15}$	14.8 ± 11.9
DataSet: CT-Density						
Poisson	0.19 ± 0.06	48.9 ± 13.1	0.38 ± 0.11	25.6 ± 7.5	0.84 ± 0.08	17.8 ± 4.3
Poisson-Gaussian	0.24 ± 0.06	40.8 ± 9.5	0.55 ± 0.08	20.0 ± 3.3	$\mathbf{0.88 \pm 0.05}$	$\mathbf{16.4 \pm 3.7}$

Table 2: SSIM and NRMSE using Poisson Gaussian likelihood with different regularization/image prior approaches. Results were averaged across 7 different noise levels by varying $\alpha \in 0.02 : 0.005 : 0.035$ in (1). WFS* runs the same number of iterations as AWFS whereas WFS† runs more iterations until convergence.

Dataset	Histopathology [50]		CelebA [51]		CT-Density	
Methods	SSIM	NRMSE (%)	SSIM	NRMSE (%)	SSIM	NRMSE (%)
Unregularized	0.57 ± 0.18	28.9 ± 9.0	0.42 ± 0.10	21.8 ± 9.1	0.24 ± 0.06	40.8 ± 9.5
RED-SD-SELF [57]	0.66 ± 0.13	21.9 ± 4.5	0.60 ± 0.09	15.9 ± 10.6	0.34 ± 0.04	28.1 ± 4.1
PnP-ADMM [54]	0.71 ± 0.11	20.7 ± 4.2	0.56 ± 0.08	16.7 ± 8.1	0.55 ± 0.03	31.2 ± 2.7
TV regularizer	0.72 ± 0.11	18.2 ± 3.9	0.64 ± 0.07	14.4 ± 8.6	0.41 ± 0.03	23.7 ± 2.8
RED-SD [56]	0.76 ± 0.09	16.8 ± 3.6	0.69 ± 0.11	13.9 ± 10.9	0.38 ± 0.04	25.9 ± 4.0
PnP-PGM [55]	0.78 ± 0.11	16.5 ± 4.5	$\mathbf{0.74 \pm 0.14}$	$\mathbf{13.5 \pm 11.3}$	0.42 ± 0.07	24.6 ± 4.4
DOLPH [58]	0.80 ± 0.06	16.0 ± 2.9	0.71 ± 0.11	13.7 ± 11.1	0.55 ± 0.08	20.0 ± 3.3
WFS*	0.76 ± 0.12	18.2 ± 5.5	0.63 ± 0.16	16.9 ± 11.8	0.53 ± 0.17	21.3 ± 7.6
WFS†	0.83 ± 0.06	16.2 ± 4.0	0.70 ± 0.16	15.7 ± 11.8	0.74 ± 0.13	17.3 ± 4.8
AWFS (Proposed)	$\mathbf{0.85 \pm 0.05}$	$\mathbf{15.4 \pm 3.7}$	$\mathbf{0.74 \pm 0.15}$	14.8 ± 11.9	$\mathbf{0.88 \pm 0.05}$	$\mathbf{16.4 \pm 3.7}$

of the PG likelihood results in improved image quality in terms of both metrics. Table 2 consists of experiments using the PG likelihood and shows that our AWFS had superior quantitative performance over all other compared methods on the histopathology and CT-density datasets.

Conclusion. We proposed a novel algorithm based on **A**ccelerated **W**irtinger **F**low and **S**core-based image prior (AWFS) for Poisson-Gaussian holographic phase retrieval that uses more realistic system and noise models. With evaluation on simulated experiments, we demonstrated that our proposed AWFS algorithm had the best reconstruction quality both qualitatively and quantitatively compared to other state-of-the-art methods. Therefore, our approach has much promise for translation in real-world applications encountering phase retrieval problems.

References

[1] K. Jaganathan, Y. C. Eldar, and B. Hassibi. *Phase retrieval: an overview of recent developments*. `https://arxiv.org/abs/1510.07713`. 2015.

[2] D. A. Barmherzig, J. Sun, E. J. Candes, T. Lane, and P.-N. Li. "Dual-Reference Design for Holographic Phase Retrieval". In: *2019 13th International conference on Sampling Theory and Applications (SampTA)*. 2019, pp. 1–4. DOI: `10.1109/SampTA45681.2019.9030848`.

[3] J. Barmherzig David A. and Sun, P.-N. Li, T. Lane, and E. J. Candès. "Holographic phase retrieval and reference design". In: *Inverse problems* 35.9 (), pp. 94001–. DOI: `10.1088/1361-6420/ab23d1`.

[4] E. J. Candès, T. Strohmer, and V. Voroninski. "PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming". In: *Comm. Pure Appl. Math.* 66.8 (2013), pp. 1241–1274. DOI: `10.1002/cpa.21432`.

[5] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski. "Phase retrieval via matrix completion". In: *SIAM J. Imaging Sci.* 6.1 (2013), 199–225. DOI: `10.1137/110848074`.

[6] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev. "Phase Retrieval with Application to Optical Imaging: A contemporary overview". In: *IEEE Sig. Proc. Mag.* 32.3 (2015), pp. 87–109. DOI: `10.1109/MSP.2014.2352673`.

[7] X. Jiang, S. Rajan, and X. Liu. "Wirtinger Flow Method With Optimal Stepsize for Phase Retrieval". In: *IEEE Sig. Proc. Letters* 23.11 (2016), pp. 1627–1631. DOI: 10.1109/LSP.2016.2611940.

[8] T. T. Cai, X. Li, and Z. Ma. "Optimal Rates of Convergence for Noisy Sparse Phase Retrieval via Thresholded Wirtinger Flow". In: *Annals Stat.* 44.5 (2016), pp. 2221–2251. DOI: 10.1214/16-AOS1443.

[9] M. Soltanolkotabi. "Structured Signal Recovery From Quadratic Measurements: Breaking Sample Complexity Barriers via Nonconvex Optimization". In: *IEEE Trans. Info. Theory* 65.4 (2019), pp. 2374–2400. DOI: 10.1109/TIT.2019.2891653.

[10] T. Qiu, P. Babu, and D. P. Palomar. "PRIME: phase retrieval via majorization-minimization". In: *IEEE Trans. Sig. Proc.* 64.19 (Oct. 2016), 5174–86. DOI: 10.1109/TSP.2016.2585084.

[11] R. W. Gerchberg and W. O. Saxton. "Practical Algorithm for Determination of Phase from Image and Diffraction Plane Pictures". In: *OPTIK* 35.2 (1972), 237–246.

[12] P. Netrapalli, P. Jain, and S. Sanghavi. "Phase Retrieval Using Alternating Minimization". In: *IEEE Trans. Sig. Proc.* 63.18 (2015), pp. 4814–4826. DOI: 10.1109/TSP.2015.2448516.

[13] I. Waldspurger. "Phase Retrieval With Random Gaussian Sensing Vectors by Alternating Projections". In: *IEEE Trans. Info. Theory* 64.5 (2018), pp. 3301–3312. DOI: 10.1109/TIT.2018.2800663.

[14] B. Gao and Z. Xu. "Phaseless Recovery Using the Gauss-Newton Method". In: *IEEE Trans. Sig. Proc.* 65.22 (2017), pp. 5885–5896. DOI: 10.1109/TSP.2017.2742981.

[15] L. Ji and Z. Tie. "On gradient descent algorithm for generalized phase retrieval problem". In: *2016 IEEE 13th International Conference on Signal Processing (ICSP)* (2016), pp. 320–325. DOI: 10.1109/ICSP.2016.7877848.

[16] J. Liang, P. Stoica, Y. Jing, and J. Li. "Phase Retrieval via the Alternating Direction Method of Multipliers". In: *IEEE Sig. Proc. Letters* 25.1 (2018), pp. 5–9. DOI: 10.1109/LSP.2017.2767826.

[17] Y. Yang, M. Pesavento, Y. C. Eldar, and B. Ottersten. "Parallel Coordinate Descent Algorithms for Sparse Phase Retrieval". In: *IEEE Intl. Conf. on Acoustics, Speech and Signal Proc. (ICASSP)*. 2019, pp. 7670–4. DOI: 10.1109/ICASSP.2019.8683363.

[18] J. Sun, Q. Qu, and J. Wright. "A geometric analysis of phase retrieval". In: *2016 IEEE International Symposium on Information Theory (ISIT)*. 2016, pp. 2379–2383. DOI: 10.1109/ISIT.2016.7541725.

[19] E. Candes, X. Li, and M. Soltanolkotabi. "Phase Retrieval via Wirtinger Flow: Theory and Algorithms". In: *IEEE Trans. Info. Theory* 61.4 (Apr. 2015), pp. 1985–2007. DOI: 10.1109/TIT.2015.2399924.

[20] Z. Wu, Y. Sun, J. Liu, and U. Kamilov. "Online Regularization by Denoising with Applications to Phase Retrieval". In: *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. 2019, pp. 3887–3895. DOI: 10.1109/ICCVW.2019.00482.

[21] P. Thibault and M. Guizar-Sicairos. "Maximum-likelihood refinement for coherent diffractive imaging". In: *New J. of Phys.* 14.6 (June 2012), p. 063004. DOI: 10.1088/1367-2630/14/6/063004.

[22] A. Goy, K. Arthur, S. Li, and G. Barbastathis. "Low photon count phase retrieval using deep learning". In: *Phys. Rev. Lett.* 121.24 (Dec. 2018), p. 243902. DOI: 10.1103/PhysRevLett.121.243902.

[23] R. Xu, M. Soltanolkotabi, J. P. Haldar, W. Unglaub, J. Zusman, A. F. J. Levi, and R. M. Leahy. *Accelerated Wirtinger Flow: A fast algorithm for ptychography*. http://arxiv.org/abs/1806.05546. 2018.

[24] D. A. Barmherzig and J. Sun. "Low-Photon Holographie Phase Retrieval". In: *Imaging and Applied Optics Congress*. Optica Publishing Group, 2020. DOI: 10.1364/COSI.2020.JTu4A.6.

[25] I. Vazquez, I. E. Harmon, J. C. R. Luna, and M. Das. "Quantitative phase retrieval with low photon counts using an energy resolving quantum detector". In: *J. Opt. Soc. Am. A* 38.1 (Jan. 2021), 71–9. DOI: 10.1364/JOSAA.396717.

[26] H. Lawrence, D. Barmherzig, H. Li, M. Eickenberg, and M. Gabrie. "Phase Retrieval with Holography and Untrained Priors: Tackling the Challenges of Low-Photon Nanoscale Imaging". In: *Proceedings of the 2nd Mathematical and Scientific Machine Learning Conference*. Vol. 145. Aug. 2022, pp. 516–567.

[27] Z. Li and J. A. Fessler. "Uniform Cramér–Rao Lower Bound for Phase Retrieval: An Empirical Study". In: *2022 56th Asilomar Conference on Signals, Systems, and Computers*. 2022, pp. 111–115. DOI: 10.1109/IEEECONF56349.2022.10052041.

[28] A. Gnanasambandam and S. H. Chan. "Image Classification in the Dark Using Quanta Image Sensors". In: *Computer Vision – ECCV 2020: 16th European Conference*. 2020, pp. 484–501. DOI: 10.1007/978-3-030-58598-3_29.

[29] Z. Li, K. Lange, and J. A. Fessler. "Poisson Phase Retrieval in Very Low-Count Regimes". In: *IEEE Transactions on Computational Imaging* 8 (2022), pp. 838–850. DOI: 10.1109/TCI.2022.3209936.

[30] G. Fatima, Z. Li, A. Arora, and P. Babu. "PDMM: A novel Primal-Dual Majorization-Minimization algorithm for Poisson Phase-Retrieval problem". In: *IEEE Trans. Sig. Proc.* (2022), pp. 1–1. DOI: 10.1109/TSP.2022.3156014.

[31] Y. Chen and E. J. Candes. "Solving random quadratic systems of equations is nearly as easy as solving linear systems". In: *Comm. Pure Appl. Math.* 70.5 (May 2017), 822–83. DOI: 10.1002/cpa.21638.

[32] L. Bian, J. Suo, J. Chung, X. Ou, C. Yang, F. Chen, and Q. Dai. "Fourier ptychographic reconstruction using Poisson maximum likelihood and truncated Wirtinger gradient". In: *Nature Sci. Rep.* 6.1 (2016). DOI: 10.1038/srep27384.

[33] H. Zhang, Y. Liang, and Y. Chi. "A nonconvex approach for phase retrieval: Reshaped Wirtinger flow and incremental algorithms". In: *J. Mach. Learning Res.* 18.141 (2017), 1–35.

[34] H. Chang, Y. Lou, Y. Duan, and S. Marchesini. "Total variation–based phase retrieval for Poisson noise removal". In: *SIAM J. Imag. Sci.* 11.1 (2018), pp. 24–55. DOI: 10.1137/16M1103270.

[35] M. Asim, M. Daniels, O. Leong, A. Ahmed, and P. Hand. "Invertible generative models for inverse problems: mitigating representation error and dataset bias". In: *Proceedings of the 37th International Conference on Machine Learning* 119 (2020).

[36] X. Wei, H. van Gorp, L. Gonzalez-Carabarin, D. Freedman, Y. C. Eldar, and R. J. G. van Sloun. "Deep Unfolding With Normalizing Flow Priors for Inverse Problems". In: *IEEE Transactions on Signal Processing* 70 (2022), pp. 2962–2971. DOI: 10.1109/TSP.2022.3179807.

[37] Y. Song and S. Ermon. "Generative Modeling by Estimating Gradients of the Data Distribution". In: *Advances in Neural Information Processing Systems*. Vol. 32. 2019.

[38] J. Ho, A. Jain, and P. Abbeel. "Denoising Diffusion Probabilistic Models". In: 33 (2020). Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, pp. 6840–6851.

[39] P. Dhariwal and A. Nichol. "Diffusion Models Beat GANs on Image Synthesis". In: *Advances in Neural Information Processing Systems*. Vol. 34. 2021, pp. 8780–8794.

[40] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole. "Score-Based Generative Modeling through Stochastic Differential Equations". In: *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. 2021.

[41] A. Graikos, N. Malkin, N. Jojic, and D. Samaras. "Diffusion Models as Plug-and-Play Priors". In: *Advances in Neural Information Processing Systems*. 2022.

[42] S. Lee, H. Chung, J. Kim, and J. C. Ye. *Progressive Deblurring of Diffusion Models for Coarse-to-Fine Image Synthesis*. 2022.

[43] A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. Tamir. "Robust Compressed Sensing MRI with Deep Generative Priors". In: *Advances in Neural Information Processing Systems*. Vol. 34. Curran Associates, Inc., 2021, pp. 14938–14954.

[44] H. Chung and J. C. Ye. "Score-based diffusion models for accelerated MRI". In: *Medical Image Analysis* 80 (2022), p. 102479. DOI: https://doi.org/10.1016/j.media.2022.102479.

[45] Z.-X. Cui, C. Cao, S. Liu, Q. Zhu, J. Cheng, H. Wang, Y. Zhu, and D. Liang. *Self-Score: Self-Supervised Learning on Score-Based Models for MRI Reconstruction*. 2022.

[46] Y. Long, J. A. Fessler, and J. M. Balter. "A 3D forward and back-projection method for X-ray CT using separable footprint". In: *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.* Winner of poster award. 2009, 146–9.

[47] Y. Song, L. Shen, L. Xing, and S. Ermon. "Solving Inverse Problems in Medical Imaging with Score-Based Generative Models". In: *International Conference on Learning Representations*. 2022.

[48] H. Li and Z. Lin. "Accelerated Proximal Gradient Methods for Nonconvex Programming". In: *Advances in Neural Information Processing Systems*. Vol. 28. 2015.

[49] P. Vincent. "A Connection Between Score Matching and Denoising Autoencoders". In: *Neural Computation* 23.7 (2011), pp. 1661–1674. DOI: 10.1162/NECO_a_00142.

[50] A. Aksac, D. Demetrick, T. Ozyer, et al. "BreCaHAD: a dataset for breast cancer histopathological annotation and diagnosis". In: *MC Res Notes* 12.82 (2019). DOI: https://doi.org/10.1186/s13104-019-4121-7.

[51] Z. Liu, P. Luo, X. Wang, and X. Tang. "Deep Learning Face Attributes in the Wild". In: *Proceedings of International Conference on Computer Vision (ICCV)*. Dec. 2015.

[52] P. J. Huber. *Robust statistics*. New York: Wiley, 1981.

[53] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising". In: *IEEE Transactions on Image Processing* 26.7 (2017), pp. 3142–3155. DOI: 10.1109/TIP.2017.2662206.

[54] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg. "Plug-and-Play Priors for Model Based Reconstruction". In: *Proc. IEEE Global Conf. Signal Process. and Inf. Process.* Austin, TX, USA, Dec. 2013, pp. 945–948.

[55] U. S. Kamilov, H. Mansour, and B. Wohlberg. "A Plug-and-Play Priors Approach for Solving Nonlinear Imaging Inverse Problems". In: *IEEE Signal. Proc. Let.* 24.12 (Dec. 2017), pp. 1872–1876.

[56] Y. Romano, M. Elad, and P. Milanfar. "The Little Engine That Could: Regularization by Denoising (RED)". In: *SIAM Journal on Imaging Sciences* 10.4 (2017), pp. 1804–1844. DOI: 10.1137/16M1102884.

[57] J. Batson and L. Royer. "Noise2Self: Blind Denoising by Self-Supervision". In: *Proceedings of the 36th International Conference on Machine Learning*. Ed. by K. Chaudhuri and R. Salakhutdinov. Vol. 97. Proceedings of Machine Learning Research. PMLR, June 2019, pp. 524–533.

[58] S. Shoushtari, J. Liu, and U. S. Kamilov. *DOLPH: Diffusion Models for Phase Retrieval*. 2022.