



**HAL**  
open science

## Maximum-likelihood estimation in ptychography in the presence of Poisson–Gaussian noise statistics

Jacob Seifert, Yifeng Shao, Rens van Dam, Dorian Bouchet, Tristan van Leeuwen, Allard Mosk

► **To cite this version:**

Jacob Seifert, Yifeng Shao, Rens van Dam, Dorian Bouchet, Tristan van Leeuwen, et al.. Maximum-likelihood estimation in ptychography in the presence of Poisson–Gaussian noise statistics. *Optics Letters*, 2023, 48 (22), pp.6027-6030. 10.1364/OL.502344 . hal-04489456

**HAL Id: hal-04489456**

**<https://hal.science/hal-04489456>**

Submitted on 5 Mar 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Maximum-likelihood estimation in ptychography in the presence of Poisson-Gaussian noise statistics

Jacob Seifert,<sup>1,\*</sup> Yifeng Shao,<sup>2</sup> Rens van Dam,<sup>1</sup> Dorian Bouchet,<sup>3</sup> Tristan van Leeuwen,<sup>4,5</sup> and Allard P. Mosk<sup>1</sup>

<sup>1</sup>*Nanophotonics, Debye Institute for Nanomaterials Science and Centre for Extreme Matter and Emergent Phenomena, Utrecht University, P.O. Box 80000, 3508 TA Utrecht, The Netherlands*

<sup>2</sup>*Imaging Physics Department, Applied Science Faculty, Delft University of Technology, The Netherlands*

<sup>3</sup>*Université Grenoble Alpes, CNRS, LIPhy, 38000 Grenoble, France*

<sup>4</sup>*Mathematical Institute, Utrecht University, Budapestlaan 6, 3584CD, Utrecht, The Netherlands*

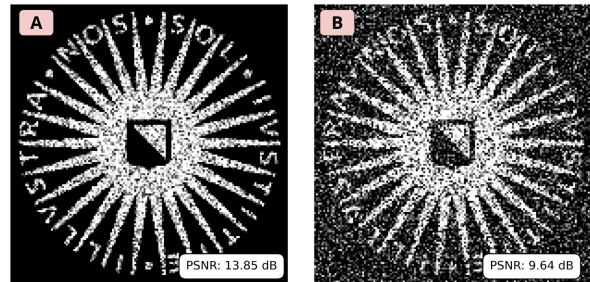
<sup>5</sup>*Centrum Wiskunde & Informatica, Science Park 123, 1098 XG, Amsterdam, The Netherlands*

Optical measurements often exhibit mixed Poisson-Gaussian noise statistics, which hampers image quality, particularly under low signal-to-noise ratio (SNR) conditions. Computational imaging falls short in such situations when solely Poissonian noise statistics are assumed. In response to this challenge, we define a loss function that explicitly incorporates this mixed noise nature. By using maximum-likelihood estimation, we devise a practical method to account for camera readout noise in gradient-based ptychography optimization. Our results, based on both experimental and numerical data, demonstrate that this approach outperforms the conventional one, enabling enhanced image reconstruction quality under challenging noise conditions through a straightforward methodological adjustment.

In the rapidly evolving field of computational imaging, ptychography has emerged as a powerful technique capable of producing high-resolution phase and amplitude images from diffraction patterns. It involves translating a thin object through overlapping illuminations and measuring the resulting diffraction patterns behind the object with a camera sensor [1, 2]. Subsequently, the complex-valued image is constructed through an iterative optimization algorithm, necessitating the formulation and minimization of a loss function, alternatively referred to as the objective, cost, or error function. Ptychography has found applications in a wide range of topics, including label-free biological imaging [3–5], optical metrology [6–10], and atomic-resolution imaging using electron beams [11–13].

Fundamentally, the basis of ptychographic reconstructions is the detection of photon counts on a camera sensor and, therefore, subject to Poissonian noise even under ideal measurement conditions. Given the assumption of an underlying noise model, a powerful and robust optimization strategy is the maximum-likelihood estimation (MLE) principle [14]. By leveraging MLE in ptychography, one seeks to estimate the studied object parameters that render the observed diffraction patterns most probable [15–19]. However, additive camera readout noise is often neglected, thus leaving a gap in the fidelity of the reconstructions. This is a salient concern as the presence of readout noise, typically Gaussian, is an important element of practical ptychographic measurements within the visible spectrum. Ignoring this noise source oversimplifies the underlying statistical model and introduces errors to the reconstructed image, especially when the detected photon counts and signal-to-noise ratio (SNR) are

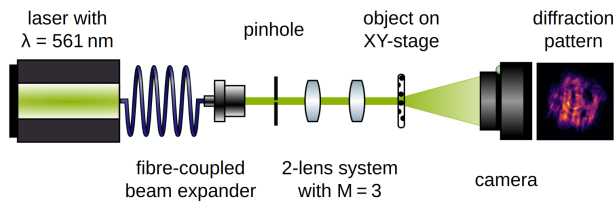
low. The distinction between a Poissonian and a mixed Poisson-Gaussian noise model is depicted in the simulated images presented in Fig. 1.



**Fig. 1.** Comparative visualization of a greyscale image distorted by different noise types. Panel (A) depicts an image with simulated Poissonian noise, while panel (B) illustrates the effect of simulated mixed Poisson-Gaussian noise resulting from additive readout noise. Inlaid values indicate Peak Signal-to-Noise Ratio (PSNR) with respect to the ground truth.

In this letter, we propose a loss function for automatic differentiation ptychography that explicitly incorporates both Poissonian and Gaussian noise sources. This approach brings us closer to the real-world conditions of ptychographic measurements, thereby paving the way for superior performance in image reconstruction under challenging noise conditions. We outline a practical method to incorporate camera readout noise in computational imaging. Furthermore, we provide a comprehensive comparison between the image reconstruction quality using a mixed-statistics loss function and that of a conventional loss function which presumes solely Poissonian noise statistics. For this, we present reconstruction results obtained from both experimental and numerical data.

\* j.seifert@uu.nl



**Fig. 2.** Schematic drawing of our ptychography setup used in the experiment and for numerical simulations. A 500- $\mu\text{m}$  pinhole is illuminated and relayed onto the object using a 2-lens system. The object laterally shifted through the beam using an XY-stage. A CMOS camera records the diffraction intensities at a distance of 38 mm downstream of the object.

In ptychography, the typical reconstruction approach involves minimizing a loss function representing the difference between the intensity values of the observed diffraction pattern  $X_k$  and the anticipated diffraction patterns  $I_k(\boldsymbol{\theta})$  as determined by a parameter set  $\boldsymbol{\theta}$ , which embodies the object under investigation, at all pixel locations indexed by  $k$ . In the presence of measurement noise, it is insightful to tackle the problem of ptychographic reconstruction by maximizing the likelihood of the observed given the object parameters. From this probabilistic perspective, one seeks the object parameters that make the observed data most likely, which renders object retrieval more robust in scenarios of low SNR.

In the supplement (section 1), we elaborate on deriving the maximum-likelihood estimation (MLE) loss functions for two different types of noise statistics. When operating under the assumption of Poissonian counting noise, the loss function  $L_{\text{Poisson}}$  that yields the maximum-likelihood estimate is expressed as:

$$L_{\text{Poisson}}(\boldsymbol{\theta}) = \sum_{k=1}^N \left( \sqrt{X_k} - \sqrt{I_k(\boldsymbol{\theta})} \right)^2, \quad (1)$$

where the sum encompasses  $N$  statistically independent pixels on the camera sensor. To account for Gaussian readout noise on the camera sensor, an additional data acquisition step becomes essential to extend the MLE loss function: the variance  $\sigma_k^2$  of the readout noise of the camera must be determined using multiple full-frame dark images. With this additional information and the assumption that the Poissonian component of the statistics can be approximated by a Gaussian distribution, we formulate the MLE loss function that incorporates mixed Poisson-Gaussian noise statistics:

$$L_{\text{Mixed}}(\boldsymbol{\theta}) = \sum_{k=1}^N \left( \ln[I_k(\boldsymbol{\theta}) + \sigma_k^2] + \frac{[X_k - I_k(\boldsymbol{\theta})]^2}{I_k(\boldsymbol{\theta}) + \sigma_k^2} \right). \quad (2)$$

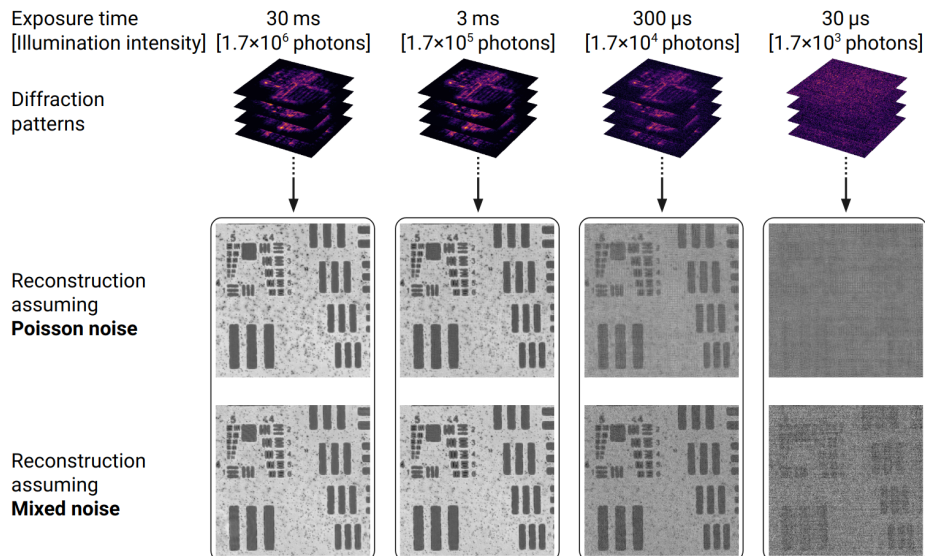
Note that this expression is not only relevant for pixels with high photon counts where the Gaussian approximation of Poisson statistics is most accurate, but also for

pixels with low photon counts. Indeed, for those low-count pixels, the Gaussian readout noise is the dominant source of noise such that the deviation of Poisson statistics from a Gaussian distribution becomes irrelevant.

To validate the beneficial effect of the loss function  $L_{\text{Mixed}}(\boldsymbol{\theta})$  experimentally, we are considering a standard ptychography setup in a transmission geometry (Fig. 2). A circular 500- $\mu\text{m}$  pinhole is illuminated with coherent light with wavelength 561 nm and relayed to the sample plane using two lenses with a magnification  $M = 3$ . There, a binary target sample is illuminated at 80 scanning positions with an overlap of approximately 60% between adjacent positions, which lies within the ideal regime according to [20]. The scattered light is captured by a CMOS camera positioned 38 mm away from the object. To ensure a comprehensive comparison as a function of SNR, we employ four different exposure settings per scanning position, spanning a range from 30  $\mu\text{s}$  to 300 ms, with each subsequent exposure time differing by a factor of 10 (see Fig. 3). For each exposure time, we determine the variances  $\sigma_k^2$  by capturing a stack of 300 dark images. This quantifies the readout noise level associated with each pixel  $k$  of our camera. A comprehensive overview of further details about the experimental implementation and methodology can be found in section 2 of the supplementary information.

By maintaining a constant illumination power, we acquire four distinct ptychographic datasets, each corresponding to a different signal-to-noise ratio (SNR), as demonstrated in the top row of Fig. 3. With these datasets in hand, we proceed to perform image reconstructions utilizing two different loss functions (Eq. 1 and 2) within an automatic differentiation-based ptychography framework, as detailed in [21] and similar to [22]. For this analysis, we precalibrate the illumination field using an additional high-SNR measurement and restrict our optimizations to the complex-valued object transmission functions, particularly under decreasing SNR conditions. This approach enables us to isolate the impact of the loss function choice from the convergence behavior associated with an unknown illumination field. The reconstruction procedure is explained in more detail in section 3 of the supplementary information, and the source code and raw data are available in [23] under open licenses. In scenarios characterized by high SNR, no noticeable disparity in image quality is observed between the two approaches. However, when confronted with low SNR conditions, where the signal becomes immersed within the readout noise, the advantages of the mixed-statistics MLE loss function become clear. The reconstructed images exhibit superior quality and reveal finer details that would otherwise remain obscured without accounting for the readout noise statistics.

For a quantitative analysis and validation of our experimental findings, we generate a simulated object with phase and amplitude contrasts which we treat as the



**Fig. 3.** Comparison of image reconstruction qualities from ptychographic datasets with decreasing signal-to-noise ratio (SNR) from left to right. Top row: Ptychographic datasets at different camera exposures with total illumination intensity, respectively. Second row: Reconstructions using a loss function assuming solely Poissonian noise statistics. Third row: Reconstructions using the loss function defined by Eq. 2 which incorporates mixed Poisson-Gaussian noise statistics. Edge length of every image equals 3.5 mm.

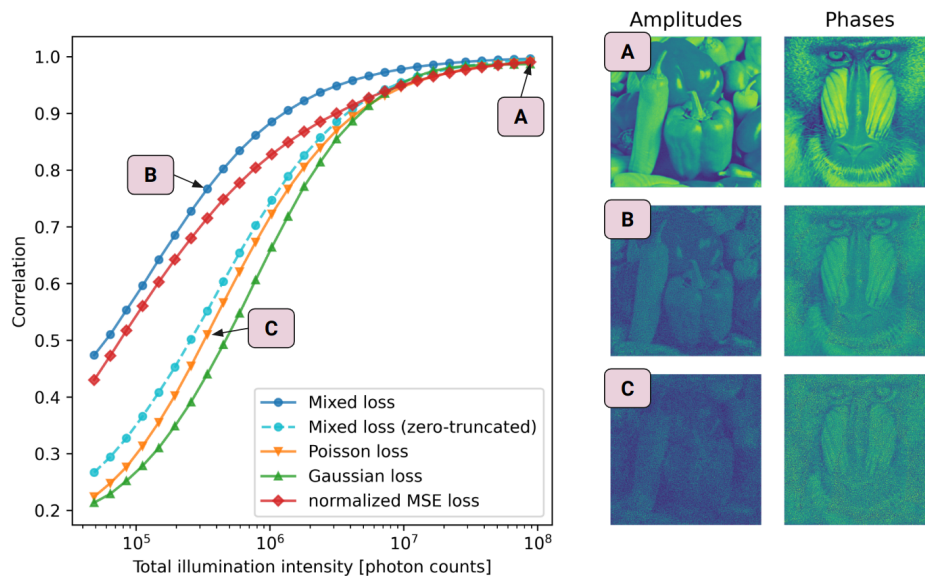
ground truth. Using numerical simulation and the ground truth object, we compute the noisy diffraction patterns assuming Poissonian photon count statistics and an additive Gaussian readout noise with a standard deviation of  $\sigma = 1.5$  counts. The simulation allows for varying the illumination intensity given as the total number of photons in an otherwise fixed illumination field that approximates the experimental conditions shown in Fig. 3, thereby controlling the measurement SNR.

The achieved reconstruction quality of the object  $O$  with respect to the ground truth  $O_{\text{gt}}$  can now be quantified using the correlation coefficient  $C = \frac{|\langle \overline{O}_{\text{gt}}, O \rangle|}{\|O_{\text{gt}}\| \|O\|}$  as defined and motivated in [21] as a function of illumination intensity (Fig. 4). Here,  $\overline{O}_{\text{gt}}$  denotes the complex conjugate of  $O_{\text{gt}}$ ,  $\langle \cdot, \cdot \rangle$  denotes the dot product, and  $\| \cdot \|$  denotes the norm. The simulation confirms the same trend that we observe from the experiments: For low SNR, optimization using a mixed-statistics loss function yields significantly better reconstruction results. In the regime of high illumination intensities, all loss functions converge excellently (up to machine precision) as the readout noise becomes irrelevant, and all assumed underlying probability density functions become valid approximations. Image reconstruction using a Gaussian loss function  $L_{\text{Gaussian}}(\boldsymbol{\theta}) = \sum_{k=1}^N (X_k - I_k(\boldsymbol{\theta}))^2$  performs worst at low illumination intensities. However, introducing a weighting term as motivated in [24] leads to a noteworthy improvement (shown in red) using  $L_{\text{normMSE}}(\boldsymbol{\theta}) = \sum_{k=1}^N \left( \frac{X_k - I_k(\boldsymbol{\theta})}{\text{sg}[I_k(\boldsymbol{\theta})] + \epsilon} \right)^2$ , where  $\epsilon = 10^{-3}$  and  $\text{sg}[\cdot]$  indicates a stop-gradient function. This normalized

MSE loss function can be interesting in cases where  $\sigma_k^2$  is impractical to obtain.

The intensity readout at a given pixel may be negative due to the additive Gaussian component in the noise statistics. This can occur in practice in an experiment via background subtraction, when areas on the camera sensor detect only low intensities. As a practical measure to keep the loss function real-valued when calculating the square root of intensity for the Poissonian loss function, negative intensity values are customarily forced to zero [25, 26]. Hence, in both simulation and experimental scenarios, we assign zero to negative intensity values when optimizing  $L_{\text{Poisson}}$ . However, by zero-cropping the intensity data, we may inadvertently eliminate valuable information, thereby causing a potential bias in our reconstruction results. To examine this bias, and quantify the potential information contained within negative values, we also test  $L_{\text{Mixed}}$  with zero-cropped data, as shown in Fig. 4. The quality of reconstruction in these conditions falls between the results from optimizing  $L_{\text{Poisson}}$  with zero-cropped data and  $L_{\text{Mixed}}$  with unaltered data. This observation suggests that the enhanced reconstruction quality derived from using a mixed-statistics loss function can partially be credited to the statistical information encapsulated in the negative pixel values resulting from background subtraction.

The results presented in this study underscore the importance of considering mixed Poisson-Gaussian noise statistics in ptychographic image reconstruction. We have demonstrated, through both experimental and simulated data, that using an MLE loss function that con-



**Fig. 4.** Correlation between ground truth and reconstructions as a function of total photon count in the illumination field, based on numerical simulations. The dashed blue line represents reconstructions derived from intensity data where negative values have been zero-cropped, while the solid line represents reconstructions that incorporate negative values, which can arise due to the additive Gaussian noise component. (A) Amplitude and phase contrast reconstruction of the simulated object based on high-intensity diffraction patterns. For comparison, with an illumination intensity of  $3.4 \times 10^5$  photons the reconstructions achieved with  $L_{\text{Mixed}}$  (B) and  $L_{\text{Poisson}}$  (C) are shown. The edge length of every image equals 3.5 mm.

siders this mix of noise statistics improves reconstruction quality, particularly in low signal-to-noise ratio conditions. This enhanced performance indicates that the mixed-statistics loss function can extract more information from the measured data by accurately accounting for the underlying noise statistics. An interesting outcome of this study concerns the practice of zero-cropping negative intensity values. We find that this practice introduces a bias into the reconstructions, highlighting the importance of preserving all statistical information in the data.

It is worth noting that some types of detectors bypass the issue of significant Gaussian readout noise, such as high-performance photon-counting hybrid pixel detectors notably used in x-ray ptychography [27, 28]. In such cases, optimizing  $L_{\text{Poisson}}(\theta)$  can yield excellent reconstruction results. Future research could focus on studying the convergence behavior of a mixed-statistics loss function when the illumination field is jointly optimized, as we have observed that optimizing  $L_{\text{Mixed}}(\theta)$  occasionally leads to a less reliable convergence when dealing with a poor initial estimate for the illumination field. To ensure valid comparison and to attain image retrieval under extremely ill-posed conditions, we included the additional step of pre-calibrating the illumination field in this study. Such a step is typically unnecessary in well-posed ptychographic reconstructions [29] or other approaches to noise-robust phase retrieval methods [30, 31].

In summary, the findings presented here could potentially propel significant advancements in the field of

computational imaging, leading to improved image retrieval under challenging noise conditions. By offering a more accurate reflection of real-world ptychographic measurements, a loss function that considers mixed Poisson-Gaussian noise statistics could greatly contribute to various fields, including material science, biology, and nanotechnology, where high-quality image reconstruction under low-SNR conditions is critical. Moreover, the utility of using a mixed-statistics loss function is not just limited to ptychography but extends to many computational and gradient-based imaging methods, broadening its applicability [32–36].

**Funding:** Netherlands Organization for Scientific Research NWO (Perspective P16-08).

**Disclosures:** The authors declare no conflicts of interest.

**Data availability:** Raw data, source code, and reconstruction scripts are available in [23].

**Supplemental Document:** See Supplement 1 for supporting content.

- 
- [1] J. M. Rodenburg and H. M. L. Faulkner, *Appl. Phys. Lett.* **85**, 4795 (2004).
  - [2] J. Rodenburg and A. Maiden, in *Springer Handbook of Microscopy*, edited by P. W. Hawkes and J. C. H. Spence (Springer International Publishing, Cham, 2019) p. 2.
  - [3] K. Giewekemeyer, P. Thibault, S. Kalbfleisch, A. Beerlink, C. M. Kewish, M. Dierolf, F. Pfeiffer, and

- T. Salditt, Proc. Natl. Acad. Sci. U. S. A. **107**, 529 (2010).
- [4] J. Marrison, L. Rätty, P. Marriott, and P. O’Toole, Sci. Rep. **3**, 2369 (2013).
- [5] C. C. Polo, L. Pereira, P. Mazzafera, D. N. A. Flores-Borges, J. L. S. Mayer, M. Guizar-Sicairos, M. Holler, M. Barsi-Andreeta, H. Westfahl, Jr, and F. Meneau, Sci. Rep. **10**, 6023 (2020).
- [6] D. B. Moore and J. R. Fienup, Appl. Opt. **55**, 4596 (2016).
- [7] D. Claus, D. J. Robinson, D. G. Chetwynd, Y. Shuo, W. T. Pike, J. J. De J Toriz Garcia, and J. M. Rodenburg, J. Opt. **15**, 035702 (2013).
- [8] P. Song, S. Jiang, H. Zhang, X. Huang, Y. Zhang, and G. Zheng, APL Photonics **4**, 050802 (2019).
- [9] M. Du, X. Liu, A. Pelekanidis, F. Zhang, L. Loetgering, P. Konold, C. L. Porter, P. Smorenburg, K. S. E. Eikema, and S. Witte, Optica **10**, 255 (2023).
- [10] D. Bouchet, J. Seifert, and A. P. Mosk, Opt. Lett. **46**, 254 (2021).
- [11] H. Yang, R. N. Rutte, L. Jones, M. Simson, R. Sagawa, H. Ryll, M. Huth, T. J. Pennycook, M. L. H. Green, H. Soltau, Y. Kondo, B. G. Davis, and P. D. Nellist, Nat. Commun. **7**, 12532 (2016).
- [12] K. Kharitonov, M. Mehrjoo, M. Ruiz-Lopez, B. Keitel, S. Kreis, M. Seyrich, M. Pop, and E. Plönjes, Opt. Express **29**, 22345 (2021).
- [13] Z. Chen, M. Odstrcil, Y. Jiang, Y. Han, M.-H. Chiu, L.-J. Li, and D. A. Muller, Nat. Commun. **11**, 2994 (2020).
- [14] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory* (Prentice-Hall, Inc., 1993).
- [15] P. Thibault and M. Guizar-Sicairos, New J. Phys. **14**, 063004 (2012).
- [16] P. Godard, M. Allain, V. Chamard, and J. Rodenburg, Opt. Express **20**, 25914 (2012).
- [17] L.-H. Yeh, J. Dong, J. Zhong, L. Tian, M. Chen, G. Tang, M. Soltanolkotabi, and L. Waller, Optics express **23**, 33214 (2015).
- [18] Y. Zhang, P. Song, and Q. Dai, Opt. Express **25**, 168 (2017).
- [19] X. Wei, H. P. Urbach, and W. M. J. Coene, Phys. Rev. A **102**, 043516 (2020).
- [20] O. Bunk, M. Dierolf, S. Kynde, I. Johnson, O. Marti, and F. Pfeiffer, Ultramicroscopy **108**, 481 (2008).
- [21] J. Seifert, D. Bouchet, L. Loetgering, and A. P. Mosk, OSA Continuum **4**, 121 (2021).
- [22] M. Du, Y. S. G. Nashed, S. Kandel, D. Gürsoy, and C. Jacobsen, Sci. Adv. **6**, eaay3700 (2020).
- [23] J. Seifert, Y. Shao, and A. P. Mosk, Data publication platform of Utrecht University (2023), <https://doi.org/10.24416/UU01-4SVE6U>.
- [24] B. Mildenhall, P. Hedman, R. Martin-Brualla, P. Srinivasan, and J. T. Barron, “Nerf in the dark: High dynamic range view synthesis from noisy raw images,” (2021), [arXiv:2111.13679 \[cs.CV\]](https://arxiv.org/abs/2111.13679).
- [25] R. Cao, T. Yang, Y. Fang, C. Kuang, and X. Liu, Appl. Opt. **56**, 6930 (2017).
- [26] S. Kandel, S. Maddali, M. Allain, S. O. Hruszkewycz, C. Jacobsen, and Y. S. G. Nashed, Opt. Express **27**, 18653 (2019).
- [27] M. Guizar-Sicairos, I. Johnson, A. Diaz, M. Holler, P. Karvinen, H.-C. Stadler, R. Dinapoli, O. Bunk, and A. Menzel, Opt. Express **22**, 14859 (2014).
- [28] F. Pfeiffer, Nat. Photonics **12**, 9 (2017).
- [29] A. M. Maiden and J. M. Rodenburg, Ultramicroscopy **109**, 1256 (2009).
- [30] C. Shen, X. Bao, J. Tan, S. Liu, and Z. Liu, Opt. Express **25**, 16235 (2017).
- [31] X. Chang, C. Shen, S. Liu, D. Zheng, S. Wang, C. Yang, N. E. Huang, and L. Bian, Opt. Lett. **48**, 4161 (2023).
- [32] N. Antipa, G. Kuo, R. Heckel, B. Mildenhall, E. Bostan, and others, Optica (2018).
- [33] M. O’Toole, D. B. Lindell, and G. Wetzstein, Nature **555**, 338 (2018).
- [34] R. Raskar, A. Agrawal, C. A. Wilson, and A. Veeraraghavan, ACM Trans. Graph. **27**, 1–10 (2008).
- [35] S.-H. Hong, J.-S. Jang, and B. Javidi, Opt. Express **12**, 483 (2004).
- [36] A. Liutkus, D. Martina, S. Popoff, G. Chardon, O. Katz, G. Lerosey, S. Gigan, L. Daudet, and I. Carron, Sci. Rep. **4**, 5552 (2014).
- [37] X. Huang, H. Yan, R. Harder, Y. Hwu, I. K. Robinson, and Y. S. Chu, Opt. Express **22**, 12634 (2014).
- [38] D. P. Kingma and J. Ba, arXiv preprint [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014)

# Maximum-likelihood estimation in ptychography in the presence of Poisson-Gaussian noise statistics: supplemental document

Jacob Seifert,<sup>1</sup> Yifeng Shao,<sup>1,2</sup> Rens van Dam,<sup>1</sup> Dorian Bouchet,<sup>3</sup> Tristan van Leeuwen,<sup>4,5</sup> and Allard P. Mosk<sup>1</sup>

<sup>1</sup>*Nanophotonics, Debye Institute for Nanomaterials Science and Centre for Extreme Matter and Emergent Phenomena, Utrecht University, P.O. Box 80000, 3508 TA Utrecht, The Netherlands*

<sup>2</sup>*Imaging Physics Department, Applied Science Faculty, Delft University of Technology, The Netherlands*

<sup>3</sup>*Univ. Grenoble Alpes, CNRS, LIPhy, 38000 Grenoble, France*

<sup>4</sup>*Centrum Wiskunde & Informatica, Science Park 123, 1098 XG, Amsterdam, The Netherlands*

<sup>5</sup>*Mathematical Institute, Utrecht University, Budapestlaan 6, 3584CD, Utrecht, The Netherlands*

This document provides supplementary material to *Maximum-likelihood estimation in ptychography in the presence of Poisson-Gaussian noise statistics*.

## DERIVATION OF THE MAXIMUM-LIKELIHOOD ESTIMATION LOSS FUNCTIONS

In this supplementary chapter, we provide a derivation of the maximum-likelihood estimation (MLE) loss functions used in a ptychography framework based on automatic differentiation. MLE operates by selecting the set of parameters that maximize the likelihood function, thus ensuring the best fit to the observed data when the noise follows a known statistical probability distribution.

### Poissonian noise statistics

Let us write the parameters of the ptychography model as a vector  $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_{\hat{N}}]$  with  $\hat{N}$  denoting the total number of free parameters, presented in our case as complex-valued object pixels. Within a physics-based forward model of ptychography (as detailed in [1]), we can denote the expected intensity value at a certain pixel  $k$  as  $I_k(\boldsymbol{\theta})$ . In essence,  $I_k(\boldsymbol{\theta})$  is the noise-free predicted intensity given a specific parameter vector  $\boldsymbol{\theta}$ . Considering a discrete random variable  $Y_k$ , which characterizes the intensity measurement on a camera sensor, we assume a Poisson distribution with an expectation value of  $I_k(\boldsymbol{\theta})$ . The probability mass function is thereby given by

$$p(Y_k|\boldsymbol{\theta}) = \frac{I_k(\boldsymbol{\theta})^{Y_k}}{Y_k!} \exp(-I_k(\boldsymbol{\theta})), \quad Y_k \in \{1, 2, \dots, N\}. \quad (\text{S1})$$

Assuming that the  $N$  measurements are statistically independent, we can express the likelihood function as

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_k^N p(Y_k|\boldsymbol{\theta}). \quad (\text{S2})$$

For the sake of computational convenience, we employ the log-likelihood function  $\ell(\boldsymbol{\theta})$ , as the natural logarithm preserves order while transforming the product into the following sum:

$$\ell(\boldsymbol{\theta}) = \ln \mathcal{L}(\boldsymbol{\theta}) = \sum_k^N (Y_k \ln I_k(\boldsymbol{\theta}) - I_k(\boldsymbol{\theta}) - \ln Y_k!). \quad (\text{S3})$$

As suggested in Chapter 4.1 of [2], we can now find the second-order Taylor expansion in terms of  $\sqrt{I_k(\boldsymbol{\theta})}$  at the point  $\sqrt{I_k(\boldsymbol{\theta})} = \sqrt{Y_k}$  as

$$Y_k \ln I_k(\boldsymbol{\theta}) - I_k(\boldsymbol{\theta}) \approx -Y_k + Y_k \ln Y_k - 2 \left( \sqrt{Y_k} - \sqrt{I_k(\boldsymbol{\theta})} \right)^2. \quad (\text{S4})$$

In practice, we aim to minimize the negative log-likelihood function. This transformation leads to equivalent outcomes and enables us to implement the optimization problem using a reconstruction algorithm based on gradient descent minimization and automatic differentiation libraries such as TensorFlow [3]. Hence, ignoring all constant additive terms and all multiplicative constants, the MLE loss function for Poissonian noise statistics can be written as

$$L_{\text{Poisson}}(\boldsymbol{\theta}) = -\ell(\boldsymbol{\theta}) = \sum_{k=1}^N \left( \sqrt{Y_k} - \sqrt{I_k(\boldsymbol{\theta})} \right)^2. \quad (\text{S5})$$

### Gaussian noise statistics

Minimizing the mean squared error is a common approach to optimization problems and is mathematically closely related to using an MLE loss function with the assumption of Gaussian noise statistics. Even though that is not the ideal assumption for a random variable  $W_k$  representing an intensity measurement, it becomes practicable for a large number of detected photons or for cases where  $W_k$  can be modeled as a sum of a large number of independent, identically distributed variables, regardless of their underlying distributions (central limit theorem). Considering the random variable  $W_k$  following a Gaussian distribution with the mean  $I_k(\boldsymbol{\theta})$  and constant variance  $\sigma^2$ , we can express the probability density function as

$$p(W_k|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(W_k - I_k(\boldsymbol{\theta}))^2}{2\sigma^2}\right). \quad (\text{S6})$$

In analogy to the previous section, we can then express the log-likelihood as

$$\ell(\boldsymbol{\theta}) = \ln \mathcal{L}(\boldsymbol{\theta}) = \ln \prod_k^N p(W_k|\boldsymbol{\theta}) \quad (\text{S7})$$

$$\ell(\boldsymbol{\theta}) = -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{k=1}^N (W_k - I_k(\boldsymbol{\theta}))^2. \quad (\text{S8})$$

Now, by neglecting the constant additive term and multiplicative constants, it becomes evident that  $\ell(\boldsymbol{\theta})$  can be maximized by the least squares method. We can write the MLE loss function for Gaussian noise statistics as

$$L_{\text{Gaussian}}(\boldsymbol{\theta}) = \sum_{k=1}^N (W_k - I_k(\boldsymbol{\theta}))^2. \quad (\text{S9})$$

### Mixed Poisson-Gaussian noise statistics

Continuing from the previous sections, let us consider a random variable  $X_k$  as the intensity measurement on a camera sensor with readout noise. We express  $X_k$  as the sum over two random variables  $X_k = Y_k + Z_k$ , where  $Y_k$  follows a Poisson distribution of expectation value  $I_k(\boldsymbol{\theta})$  (see equation S1) and  $Z_k$  follows a centered Gaussian distribution of variance  $\sigma_k^2$ . The probability density function for  $Z_k$  is given by

$$p_g(Z_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{Z_k^2}{2\sigma_k^2}\right). \quad (\text{S10})$$

The random variable  $Y_k$  can be approximated as a Gaussian distribution with mean  $I_k(\boldsymbol{\theta})$  and variance  $I_k(\boldsymbol{\theta})$ :

$$p_p(Y_k|\boldsymbol{\theta}) \simeq \frac{1}{\sqrt{2\pi I_k(\boldsymbol{\theta})}} \exp\left[-\frac{(Y_k - I_k(\boldsymbol{\theta}))^2}{2I_k(\boldsymbol{\theta})}\right]. \quad (\text{S11})$$

We can express the probability density function of  $X_k$  as

$$p(X_k|\boldsymbol{\theta}) = \int_{-\infty}^{+\infty} p_p(\tau|\boldsymbol{\theta}) p_g(X_k - \tau) d\tau \quad (\text{S12})$$

$$p(X_k|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi(I_k(\boldsymbol{\theta}) + \sigma_k^2)}} \exp\left[-\frac{(X_k - I_k(\boldsymbol{\theta}))^2}{2(I_k(\boldsymbol{\theta}) + \sigma_k^2)}\right]. \quad (\text{S13})$$



In analogy to the case without readout noise above, we can now define the MLE loss function as the negative log-likelihood function:

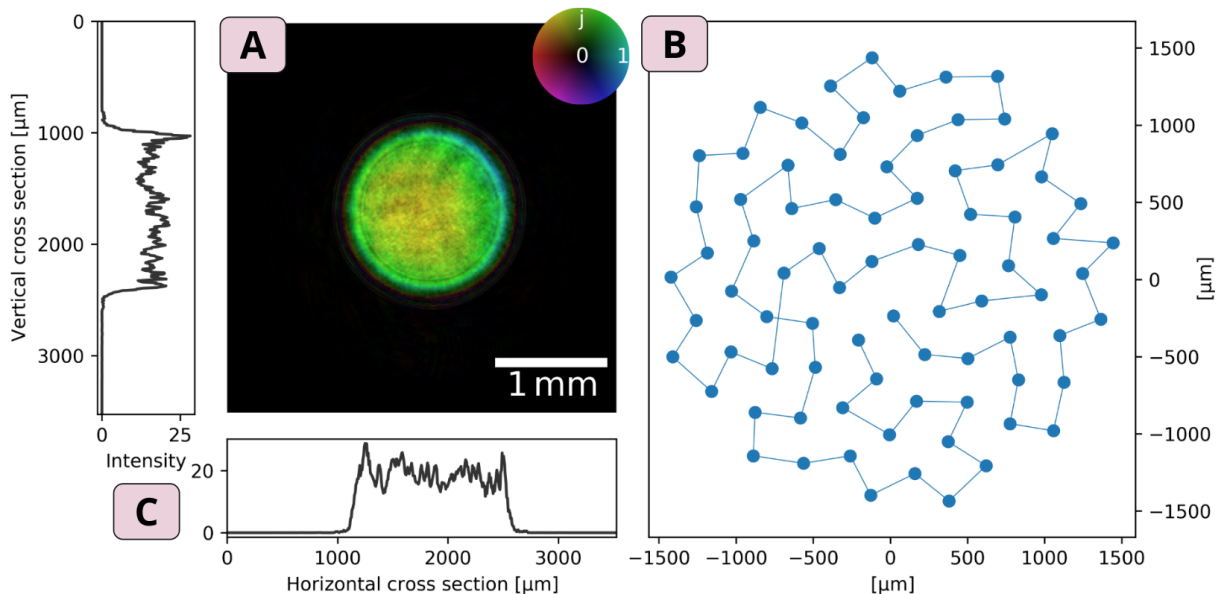
$$L_{\text{Mixed}}(\boldsymbol{\theta}) = -\ell(\boldsymbol{\theta}) = -\ln \mathcal{L}(\boldsymbol{\theta}) = -\ln \prod_k^N p(X_k|\boldsymbol{\theta}) \quad (\text{S14})$$

$$L_{\text{Mixed}}(\boldsymbol{\theta}) = \sum_{k=1}^N \left( \ln[I_k(\boldsymbol{\theta}) + \sigma_k^2] + \frac{[X_k - I_k(\boldsymbol{\theta})]^2}{I_k(\boldsymbol{\theta}) + \sigma_k^2} \right). \quad (\text{S15})$$

Here, we have neglected all constant additive terms and all multiplicative constants. To apply this loss function in an optimization framework, we need to obtain  $\sigma_k^2$ , the variance on the readout noise for each pixel, from dark measurements.

## EXPERIMENTAL SETUP AND METHOD

The experimental setup for this ptychography study, also used for numerical simulations, is depicted in Fig 2 of the main document. A coherent laser beam (Cobolt Jive 100™) with wavelength  $\lambda = 561$  nm is coupled into a single-mode fiber. A fiber collimator (60FC-L-0-M75-26, Schäfter+Kirchoff) expands the beam to around 25 mm in diameter, which then illuminates a 500  $\mu\text{m}$  pinhole. Using a 2-lens system with a magnification of  $M = 3$ , the pinhole is imaged onto the object, resulting in the illumination field shown in Fig. S1, panel A. The two transfer lenses have diameters of 22.9 mm, with focal lengths of 5 mm and 15 mm, respectively. The object ( $\mu\text{Chart}1951$  Test Target, QingYing E&T LLC) is mounted on a motorized XY-stage with stepper motor actuators (ZFS25B, Thorlabs). The scanning trajectory is shown in panel B of Fig. S1. It comprises a total of 80 positions in a Fermat spiral pattern to optimize for overlap uniformness [37]. Using a traveling salesman algorithm, this trajectory is optimized to minimize total travel distance. The linear overlap between adjacent positions is approximately 60% [20]. The diffraction patterns are recorded 37.7 mm behind the object using a CMOS camera (acA2440-35um, Basler) that features a binned pixel size of 6.9  $\mu\text{m}$  and 1024x1024 total pixels.



**Fig. S1.** (A) Visualization of the complex-valued illumination field that is used in the experiment. The image brightness represents the field amplitude, and the color represents the phase (see circular colorbar). (B) The object’s scanning trajectory through the illumination beam in the ptychographic experiment. (C) Horizontal and vertical intensity profiles, centered on the illumination field.

To control the Signal-to-Noise Ratio (SNR) in the measurement, we vary the camera exposure time from 30  $\mu\text{s}$  to 300 ms over 22 steps, covering four orders of magnitude. We derive the spatially varying readout noise variances  $\sigma_k^2$

for each exposure time from 300 dark measurements, during which the laser beam is blocked. To mitigate Johnson-Nyquist noise fluctuations, we operate the camera sensor in a temperature controlled environment at 21 °C. These measurements also provide us with an average background image for each exposure time setting that we subtract from each diffraction pattern. To facilitate exact reproduction of the results presented in this study, the raw background and noise statistics data are included alongside the reconstruction framework in [23].

It is crucial to extract statistical information from potential negative intensity values resulting from Gaussian readout noise. Therefore, we require a black level offset to ensure that no pixel of the sensor ever reads the value of zero in dark measurements. With our Basler camera, we monitor the smallest pixel value for increasing black level settings and observe that an offset of 4 first ensures that all pixel values are larger than zero. To minimize the reduction in dynamic range, we choose this relatively small black level offset for the rest of this work.

For each scanning point, an independent measurement is obtained for each exposure time setting, and an additional high-SNR measurement is taken by averaging 100 images with the highest exposure time. This high-SNR measurement aids the calibration during the reconstruction phase. In Fig. S2 presents an expanded version of Fig. 3 from the main manuscript. In the left column, the noise degradation of a single diffraction pattern is shown for all 22 exposure time settings. The central and right columns (B and C) provide a visual comparison of the reconstruction quality for each of these exposure time settings. Specifically, column B showcases reconstructions obtained by using the Poissonian log-likelihood loss function  $L_{\text{Poisson}}(\theta)$  for optimization, while column C displays reconstructions achieved by employing the mixed Poisson-Gaussian log-likelihood loss function  $L_{\text{Mixed}}(\theta)$ . This comparative illustration provides a clear understanding of the impact of the chosen loss function on the quality of reconstruction across a large range of exposure times.

## RECONSTRUCTION PROCEDURE

The reconstruction procedure begins with diffraction pattern preprocessing. An experimentally acquired mean dark image is subtracted to correct for background noise and account for hot or dead pixels. In cases where  $L_{\text{Poisson}}$  is used for optimization, negative values are set to zero due to the need for a real-valued loss function. Negative values cannot be incorporated into a noise model that solely assumes Poissonian statistics.

Typical CCD and CMOS cameras involve an analog-to-digital converter that converts the number of detected photons into analog-to-digital units (ADU). To rectify the assumption that the intensity measurement is Poisson distributed, we rescale the data by the inverse of the overall system gain. In the case of our CMOS camera, the inverse of the overall system gain is specified by the manufacturer to be  $2.7 \frac{e^-}{\text{ADU}}$ .

Initially, a high-SNR reconstruction is conducted on the calibration dataset discussed in Section . This helps rectify experimental uncertainties such as the object-camera distance and scanning positions, as well as obtaining a high-quality reconstruction of the illumination field (see Fig. S1, panel A). Following this, reconstructions from the lower-SNR diffraction patterns are retrieved using the pre-calibrated illumination field. Each reconstruction is performed in sequence on a commercial GPU (Nvidia RTX A6000) with the same hyperparameter and regularization settings. Over 100 epochs, the learning rate for the ADAM optimizer [38] starts at  $lr = 0.1$  and exponentially decays at a rate  $\lambda = 0.03$ , following the schedule  $lr_{n+1} = lr_n e^{-\lambda}$ .

Three regularization terms are added to the loss function, resulting in a final loss function in the form of  $L = L_{\text{Poisson}/\text{Mixed}} + \sum_{i=1}^3 L_{\text{Reg},i}$ .

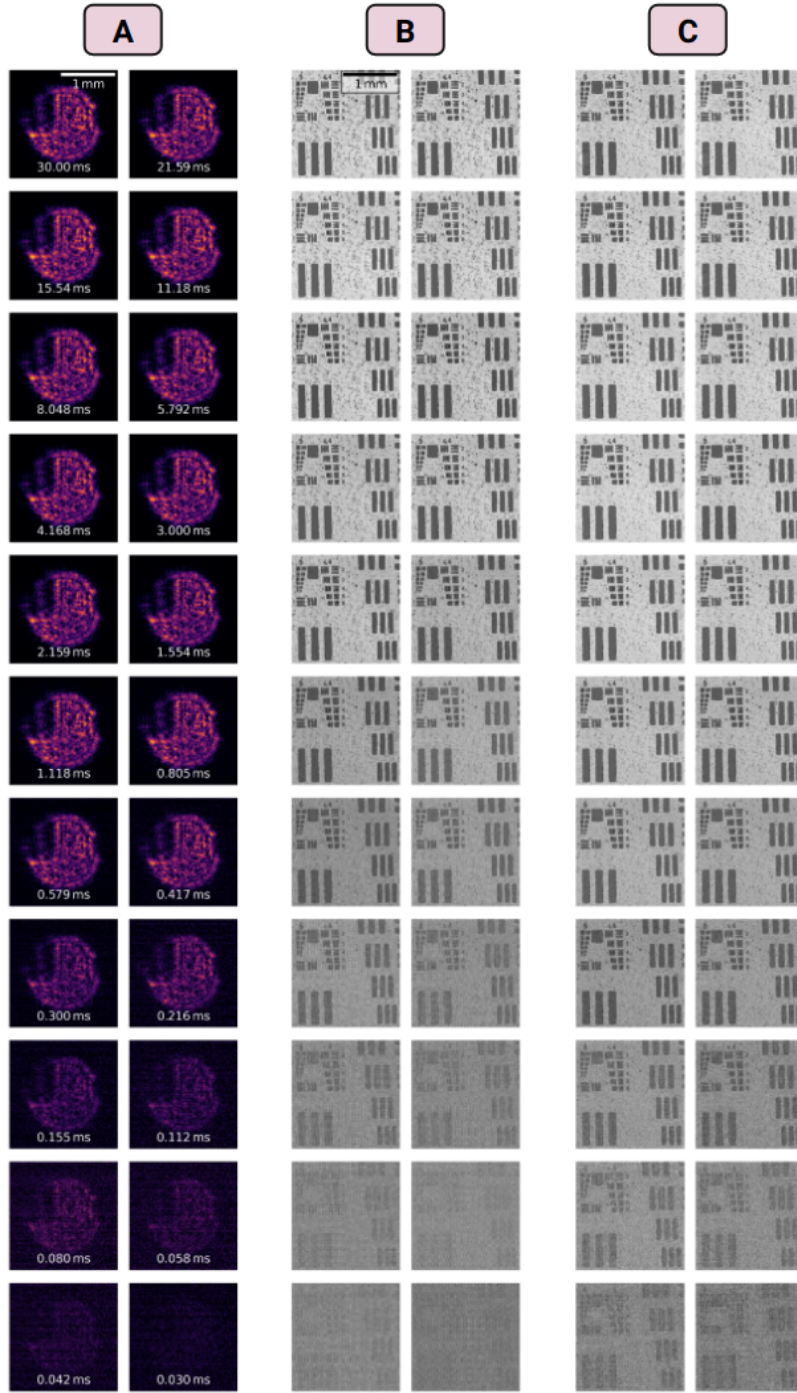
1. An L1 norm on the amplitudes for the illumination field outside a circular support constraint  $\mathcal{S}$  with a radius of 1.5 mm. This regularization term accelerates the convergence of the illumination calibration and is motivated by our experimental setup producing a circular illumination with an approximate radius of 0.75 mm. This is expressed as

$$L_{\text{Reg},1} = \alpha \sum_{(x,y) \in \mathcal{S}} |P(x,y)|, \quad (\text{S16})$$

where  $P(x,y)$  denotes the 2-dimensional illumination field ( $P$  for "probe"). The factor  $\alpha$  regulates the strength of the regularization, typically chosen as  $\alpha \approx 100$  in our calibration procedure.

2. A minor L1 norm on the amplitudes in the object given by

$$L_{\text{Reg},2} = \beta \sum_{(x,y)}^{\tilde{N}} |O(x,y)|, \quad (\text{S17})$$



**Fig. S2.** Column (A): Visualization of one diffraction pattern from the full data set at varying camera exposure time settings. The exposure times range from 30 ms (upper left) to 30  $\mu$ s (lower right). Columns (B and C): Amplitude images reconstructed from the ptychographic data sets with the respective exposure times shown in column A in the same order. Column B is reconstructed with a loss function assuming solely Poissonian noise statistics. Column C is reconstructed with the mixed-statistics loss function assuming Poisson-Gaussian noise.

where  $O(x, y)$  denotes the 2-dimensional complex-valued object with a total number of pixels  $\hat{N}$ . This regularization drives towards finding a compact solution and mitigates high object amplitudes in the object's boundary areas that are insufficiently illuminated. We set  $\beta = 0.0001$ .

3. A minor L1 norm on the summed magnitudes of the object in frequency space is expressed as

$$L_{\text{Reg},3} = \gamma \sum_{(x,y)}^{\hat{N}} |\hat{O}(u, v)|, \quad (\text{S18})$$

where  $\hat{O}(u, v)$  denotes the Fourier transform  $\hat{O}(u, v) = \mathcal{FT}\{O(x, y)\}$ . We observe that this regularization term can stabilize the optimization using  $L_{\text{Mixed}}$ , which sometimes exhibits a poorer convergence behavior than optimizing  $L_{\text{Poisson}}$  or helps prevent numerical divergence phenomena for less-than-optimal learning rates. We set  $\gamma = 0.001$ .

Note that choosing the regularization prefactors  $\alpha, \beta, \gamma$  is arbitrary and depends on non-physical parameters such as the sampling. Therefore, we adopt a heuristical approach to set them small enough to ensure that the data fidelity term strongly dominates the reconstruction process. By doing so, we preserve the valuable comparative basis between the two maximum likelihood estimation (MLE) loss functions while subtly enhancing the reconstructions:

$$\frac{L_{\text{Poisson/Mixed}}}{\sum_{i=1}^3 L_{\text{Reg},i}} \geq 100. \quad (\text{S19})$$