Phase Retrieval with Holography and Untrained Priors

Tackling the Challenges of Low-Photon Nanoscale Imaging

Hannah Lawrence[*1], David A. Barmherzig [*2], Henry Li [3], Michael Eickenberg [2], Marylou Gabrié[†2, 4]

[1]Massachusetts Institute of Technology

[2]Flatiron Institute

[3]Yale University

[4]New York University

MSML 2021

Motivation

Guiding question

How can we most practically apply modern machine learning techniques to *experimentally realistic* phase retrieval settings?

Motivation

Guiding question

How can we most practically apply modern machine learning techniques to *experimentally realistic* phase retrieval settings?

Examples of practical considerations:

1. No dataset required
2. Can handle noisy data
3. Can handle lost data
4. Faithful to modern experimental setups

How do you image nanoparticles?

How do you image nanoparticles?

Coherent diffraction imaging: diffract X-rays from a specimen (such as a virus, protein, or crystal) onto a beamstopped CCD detector

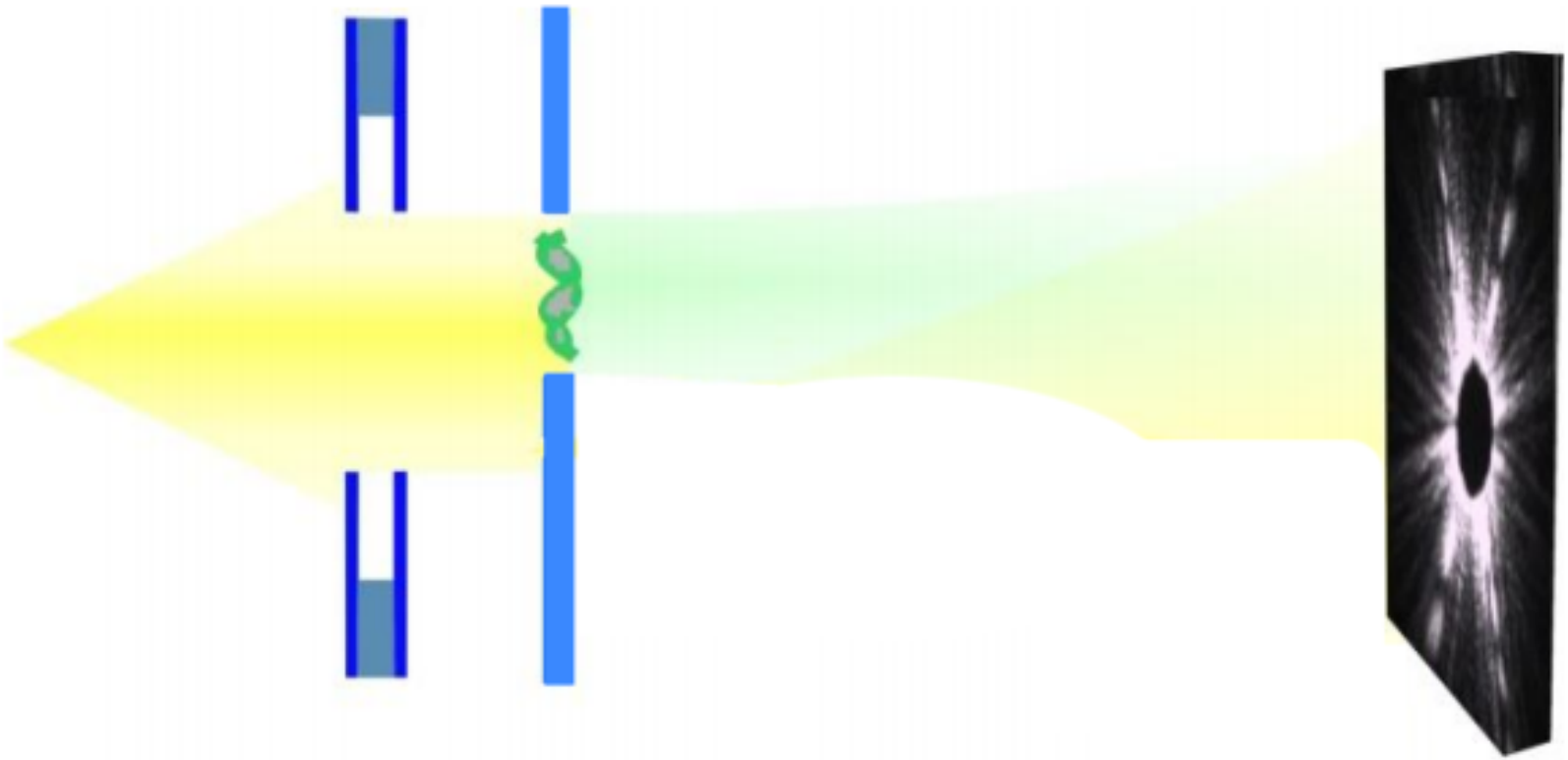
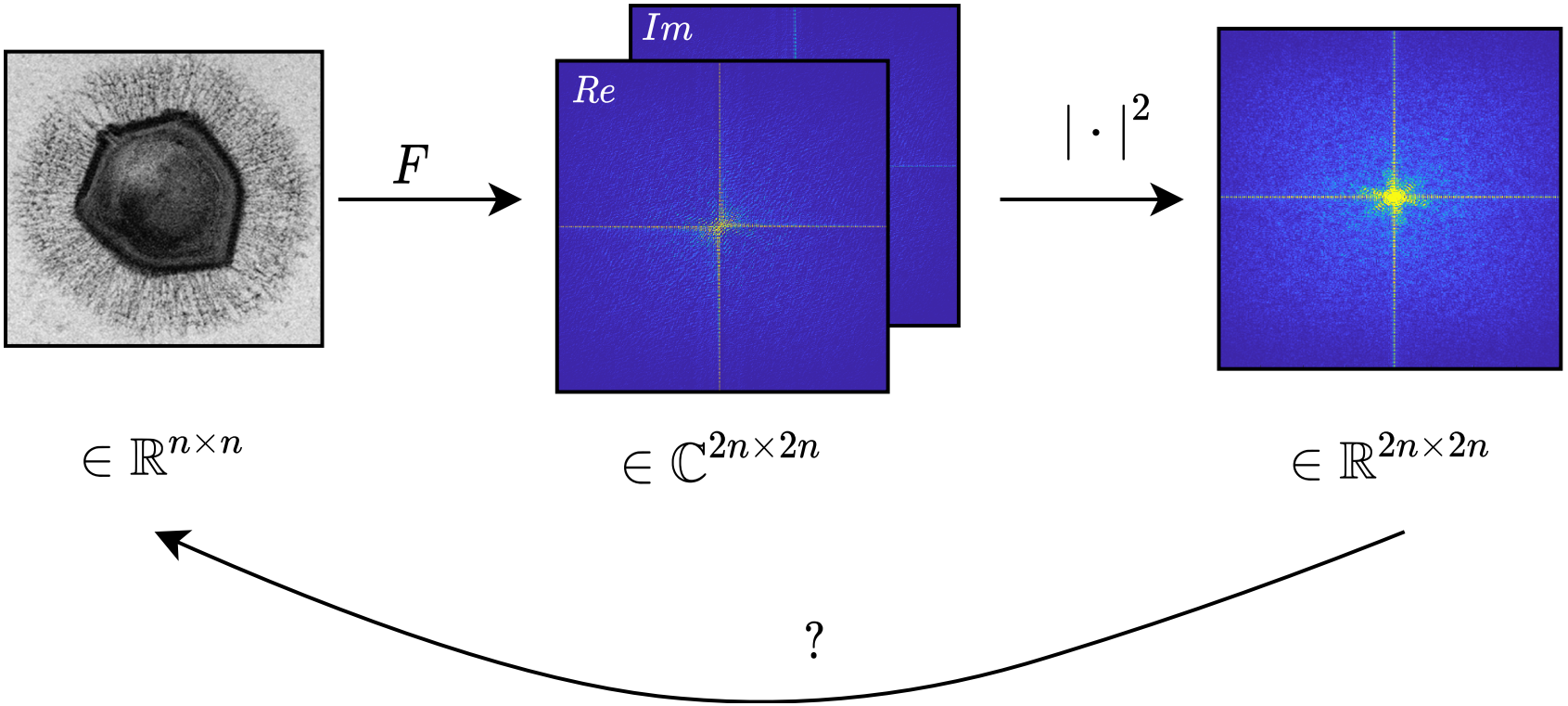


Figure: CDI schematic. The recorded data $\boldsymbol{Y}$ has its low frequencies occluded by a beamstop. (Original image courtesy of Saliba et al. [2012].)

Phase retrieval, mathematically



Here, F is a $2x$ oversampled Fourier operator:

$$(Fx)_{\ell k} = \sum_{a,b=1}^{n} x_{ab} e^{-2\pi i(\frac{a\ell+bk}{2n})}$$

Trivial symmetries cannot be avoided

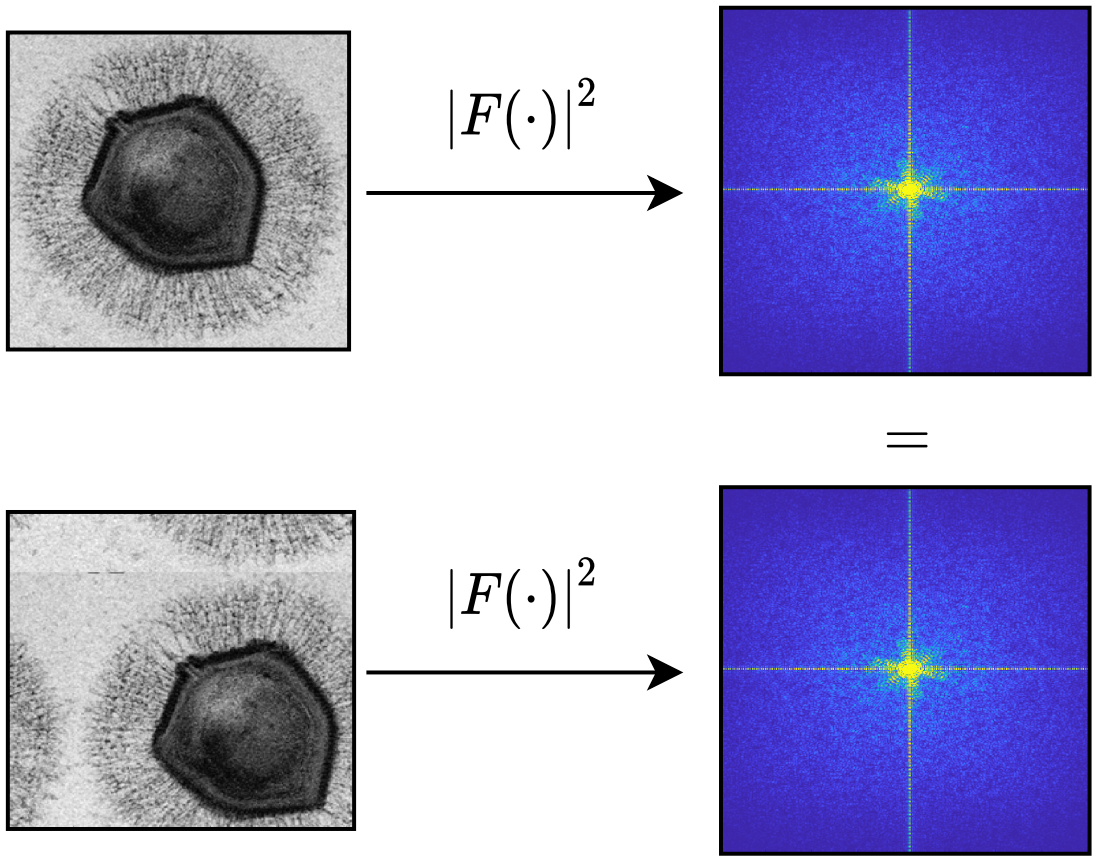


Figure: The Fourier magnitudes of x and those of any translation of x are identical.

Caution: practical $\neq$ theoretical phase retrieval

	Practice	Theory
Measurements	Oversampled Fourier	Gaussian or random Fourier
Specimen	$x \in \mathbb{R}^{n \times n}$	$x \in \mathbb{R}^n$
Forward model	$y = \lvert Fx \rvert \in \mathbb{R}^{2n \times 2n}$	$y = \lvert Ax \rvert \in \mathbb{R}^{n \log n}$
Ambiguity	Complex phase	Real sign[1] ± 1
Recovery algorithms	HIO	PhaseLift; Wirtinger flow

[1] Some theoretical works do consider a complex phase ambiguity, or even a complex-valued signal x, but typically not with Fourier measurements.

Recovery is possible, but ill-posed

Theorem (Hayes' theorem, informal [Hayes, 1982])

2x oversampling is sufficient to recover a two-dimensional image from its Fourier magnitudes, up to trivial symmetries (translation, conjugate reflection, or scaling by a unit vector).

Hybrid input-output (HIO) algorithm is a popular iterative approach which alternates between updates in real and Fourier space, but its convergence to the correct solution is not guaranteed.

Ill-posedness: there exist non-trivially distinct signals with near-indistinguishable Fourier magnitudes [Barnett et al., 2020].

Holographic phase retrieval uses a known reference to simplify reconstruction

Holographic phase retrieval uses a known reference to simplify reconstruction

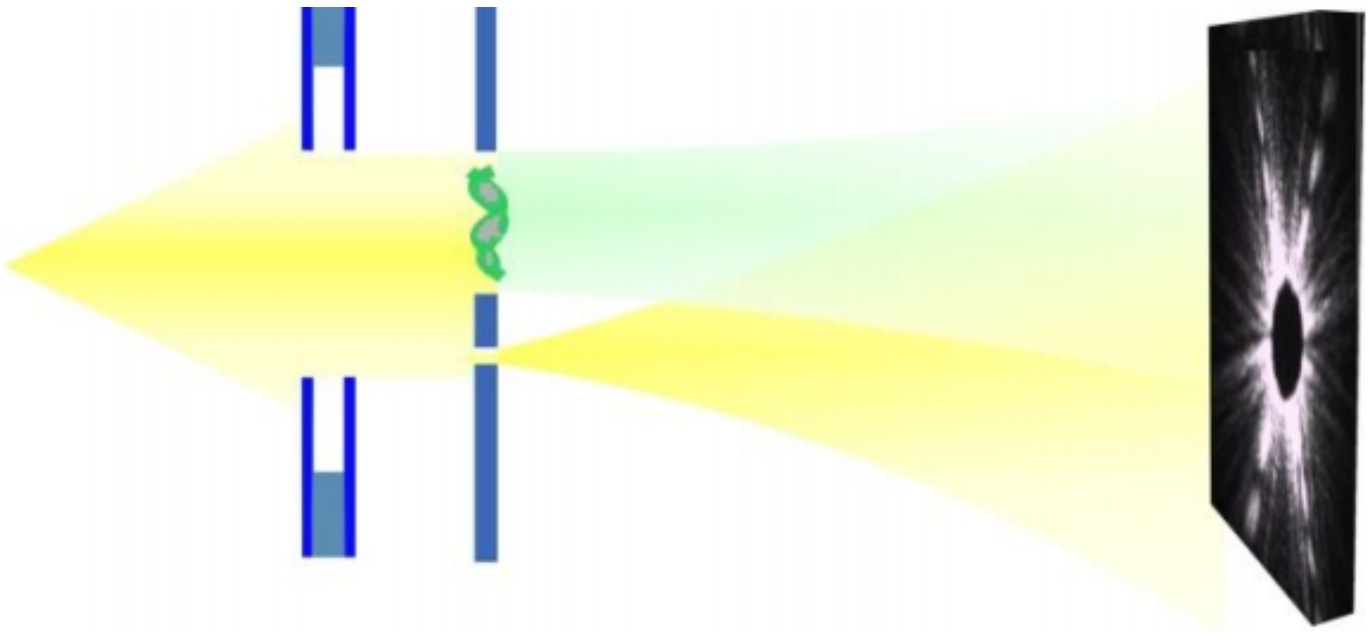


Figure: Holographic CDI schematic. The upper portion of the diffraction area contains the specimen of interest $\boldsymbol{X}_0$, and the adjacent portion consists of a known "reference" $\boldsymbol{R}_0$. The recorded data $\boldsymbol{Y}$ has its low frequencies occluded by a beamstop. (Image courtesy of Saliba et al. [2012].)

Holographic phase retrieval uses a known reference to simplify reconstruction

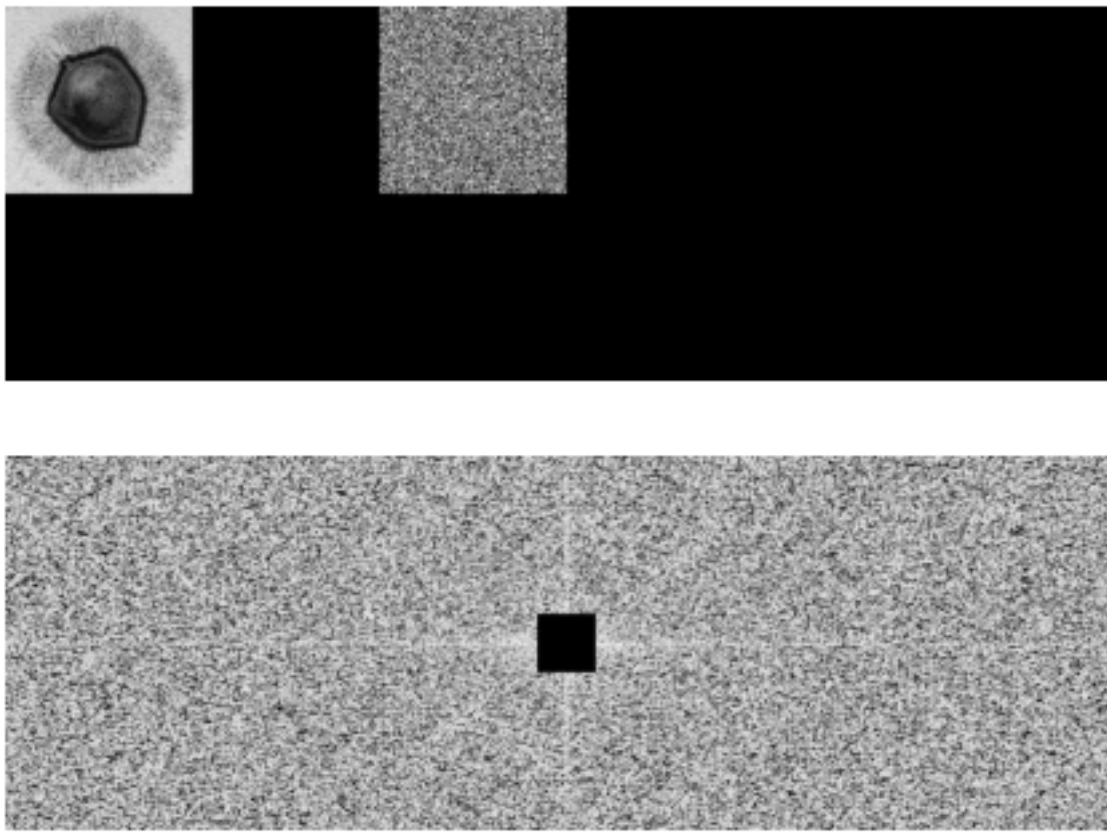
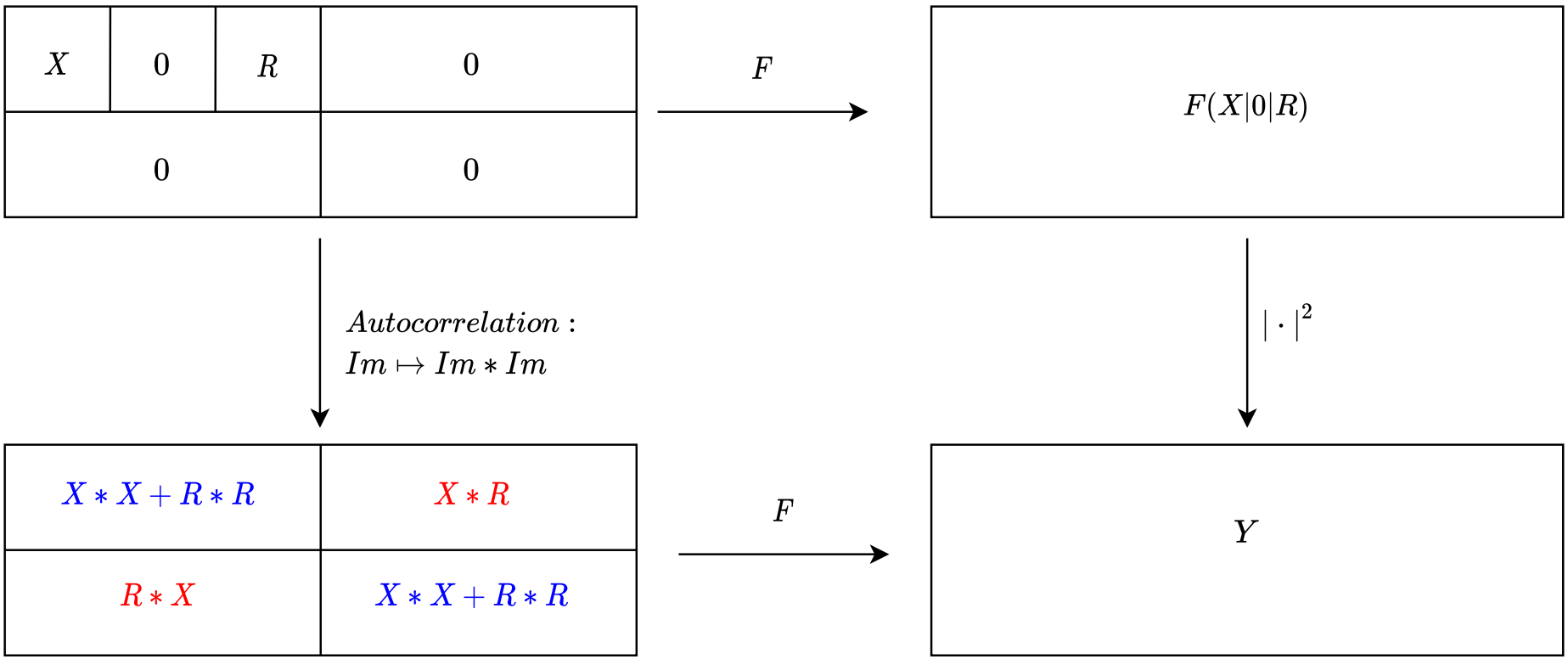


Figure: Holographic CDI schematic. The upper portion of the diffraction area contains the specimen of interest $\boldsymbol{X}_0$, and the adjacent portion consists of a known "reference" $\boldsymbol{R}_0$. The recorded data $\boldsymbol{Y}$ has its low frequencies occluded by a beamstop.

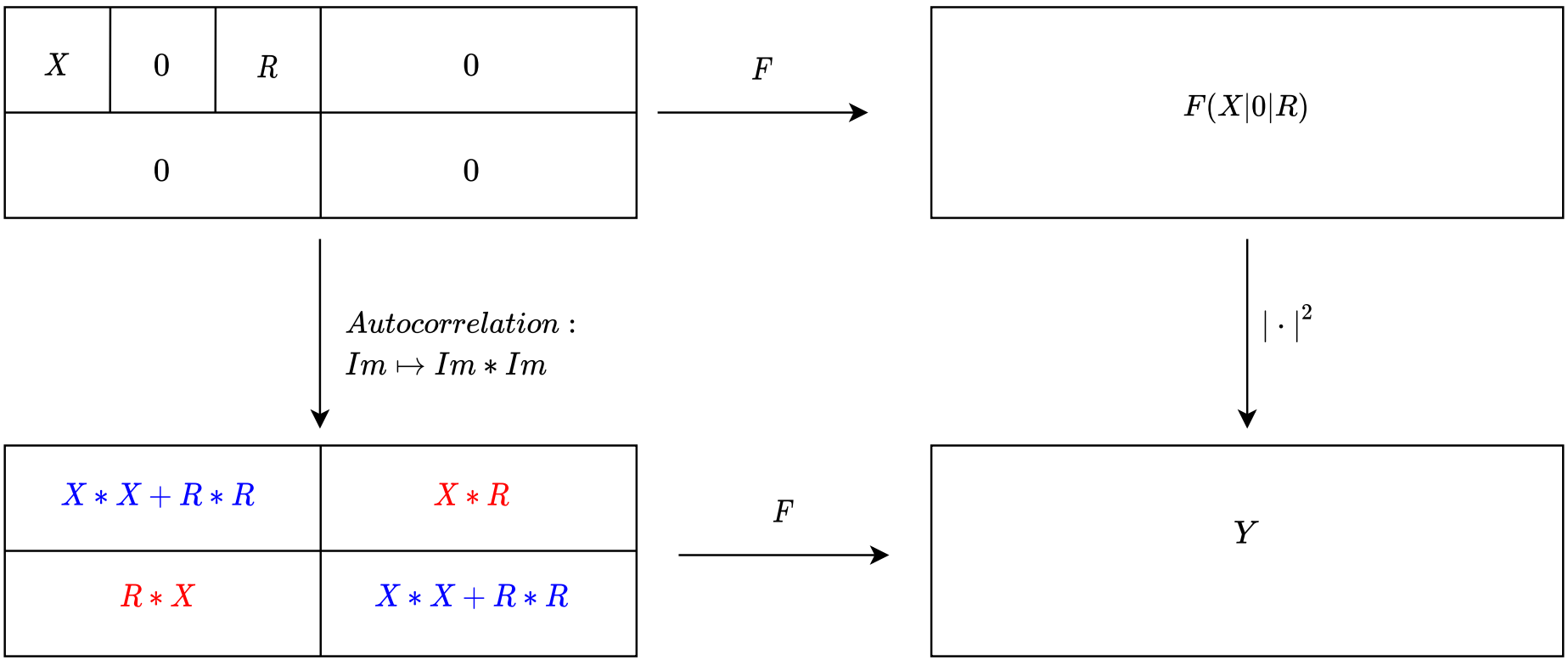
Holographic phase retrieval is well-posed



The autocorrelation of the complete specimen-reference image contains both linear and nonlinear information about X. Since R is known, one can recover X from $X * R$ via deconvolution: this is a *linear* inverse problem.[2]

[2] We use $*$ to denote cross-correlation throughout, such that $F(X * X) = |F(X)|^2$

Holographic separation condition



In order for $X * R$ to appear unobstructed in the autocorrelation, it is important that X and R be separated by a zero block of equal size. If there is no separating zero block, the bottom right pixel of R must be nonzero to extract a solvable linear system Barmherzig et al. [2019].

Nonetheless, practical challenges remain

1. **Noise:** Let $N_p > 0$ represent the expected number of photons incident per detector pixel, and Y the true Fourier magnitudes. Then the measurements are distributed as:

$$\boldsymbol{Y} \sim \frac{||Y||_{Fro}^2}{N_p} \text{Poisson}\left(\frac{N_p}{||Y||_{Fro}^2} Y\right). \tag{1}$$

   For large N_p, additive Gaussian noise is a good approximation to Poisson shot noise. However, for small N_p (the "low-photon" or high noise regime), this approximation suffers.

2. **Beamstop:** low frequencies are occluded.
3. **Reference size and separation:** physical constraints can limit reference manufacturing and separation from the specimen.

Machine learning for inverse problems (IPs)

		General IPs	Fourier phase retrieval
Trained	End-to-end	5	6, 7*
	Unrolling	8	
	Pretrained denoisers	9	10, 11, 12
	Generative models	13	
Untrained	Generative models	14, 15	16, 17

* denotes the holographic setting

The deep phase decoder [Bostan et al., 2020] uses as its untrained prior the **deep decoder** architecture, which we will also apply in the holographic setting.

Deep decoder

The **deep decoder** is an untrained image prior, where its architecture imposes an inductive bias favoring image statistics via a sequence of upsampling and 1×1 convolution layers.

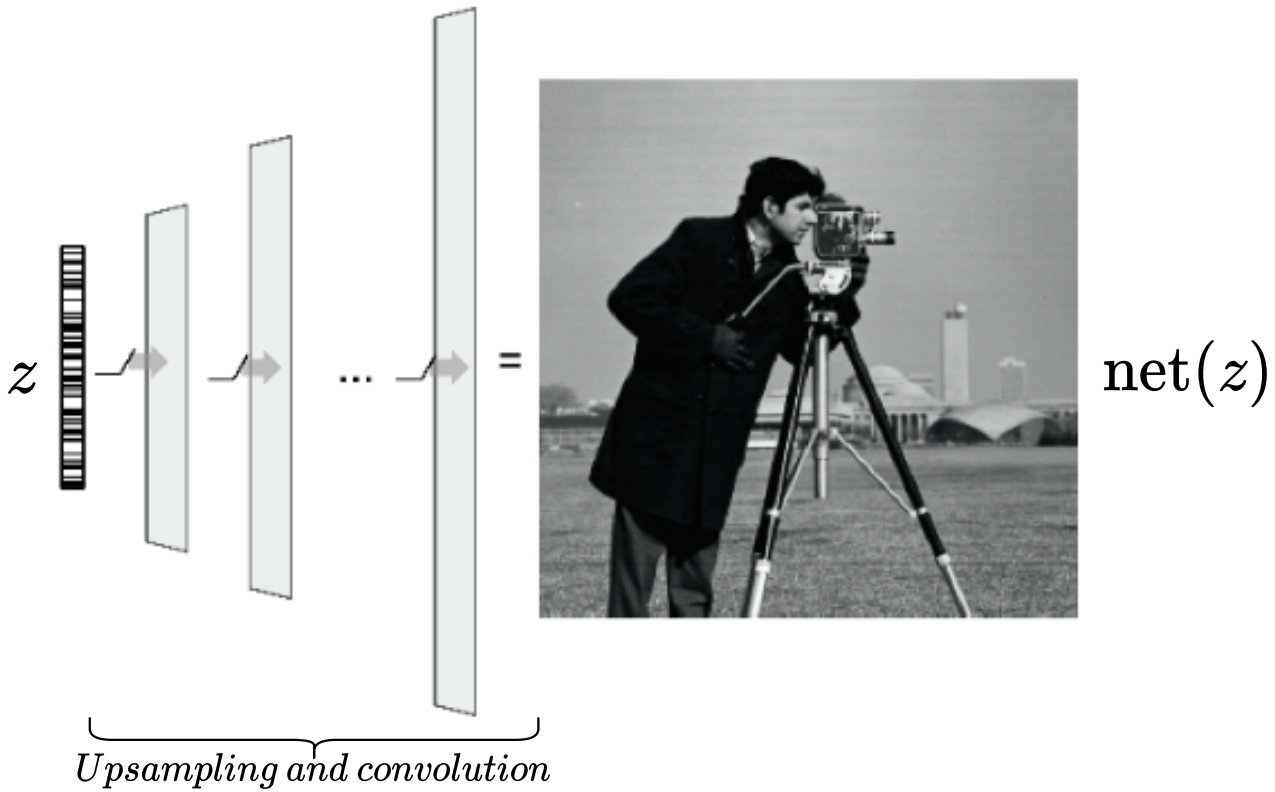


Figure: For a fixed, random input $z \in \mathbb{R}^{1 \times c_1 \times \kappa \times \lambda}$, the deep decoder outputs $\sigma(\vartheta, z) = \mathrm{net}(z)$ where ϑ denotes all convolution parameters, $\mathrm{block}_i := \mathrm{up}_2 \circ \mathrm{relu} \circ \mathrm{conv}_{\vartheta_i}$ and $\mathrm{net} := \mathrm{block}_d \circ \cdots \circ \mathrm{block}_1$.

Our method: Poisson maximum likelihood and the deep decoder

We present the first untrained machine learning method for holographic phase retrieval.

$$\begin{aligned}\hat{\boldsymbol{X}} &= \arg\max_{\boldsymbol{X}} \log p(\boldsymbol{Y}|\boldsymbol{X}, N_p) \\ &= \arg\max_{\boldsymbol{X}} \sum_{ij|(\boldsymbol{B})_{ij}=1} Y_{ij} \log |F(\boldsymbol{X}|0_{m\times m}|\boldsymbol{R})_{ij}|^2 - |F(\boldsymbol{X}|0_{m\times m}|\boldsymbol{R})_{ij}|^2\end{aligned}$$

Deep decoder: restrict maximization to $\boldsymbol{X}$ in the range of the deep decoder, i.e. such that $\boldsymbol{X} = \sigma(\vartheta, z)$ for some parameters ϑ

We optimize this objective using automatic differentiation in PyTorch.

Some intuition

$X * X + R * R$	$X * R$
$R * X$	$X * X + R * R$

Unlike linear approaches, based on $X * R$, we also implicitly leverage the nonlinear part of the data, $X * X + R * R$.

The deep decoder provides further information for the method to leverage.

Hopefully, the combination of these may compensate for missing or "challenging" data, and enable reconstruction that competing methods cannot.

Experimental setup: datasets

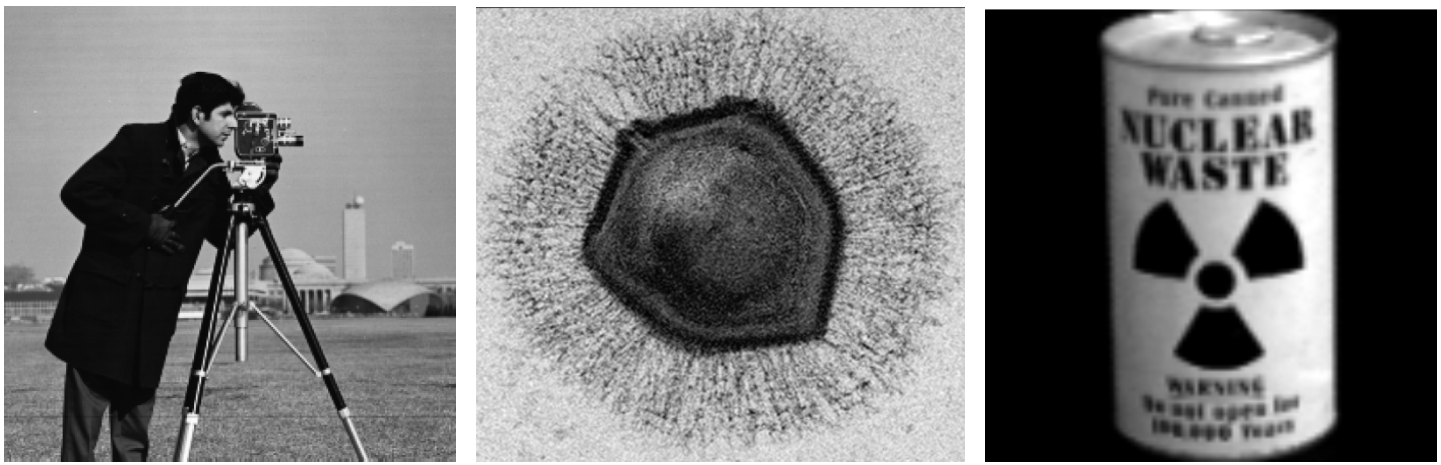


Figure: (a) Example from SET12: 12 images, 128×128 (b) Example from BIO10: 10 images, 256×256 (c) Example from COIL100: 100 images, 128×128

Experimental setup: metric and baselines

Metric: Structural Similarity Index (SSIM), used for measuring image similarity [Wang et al., 2004]

Experimental setup: metric and baselines

Metric: Structural Similarity Index (SSIM), used for measuring image similarity [Wang et al., 2004]

Baselines:

- HIO-Holo: custom variant of HIO; enforces R in real-space iterations
- Inverse filtering: deconvolution of the linear $X * R$ term
- Wiener filtering: deconvolution, under a Gaussian noise assumption, of the linear $X * R$ term

Experimental setup: metric and baselines

Metric: Structural Similarity Index (SSIM), used for measuring image similarity [Wang et al., 2004]

Baselines:

- HIO-Holo: custom variant of HIO; enforces R in real-space iterations
- Inverse filtering: deconvolution of the linear $X * R$ term
- Wiener filtering: deconvolution, under a Gaussian noise assumption, of the linear $X * R$ term

Variants of our method, HolOpt:

- "-S": **squared** loss; included as ablation study only in basic experiment
- "-P": **Poisson** loss
- "-TV": **total variation** (TV) norm regularization
- "-DD": the **deep decoder** prior (if omitted, optimization is performed in the full space of pixels)

Reconstruction at different noise levels

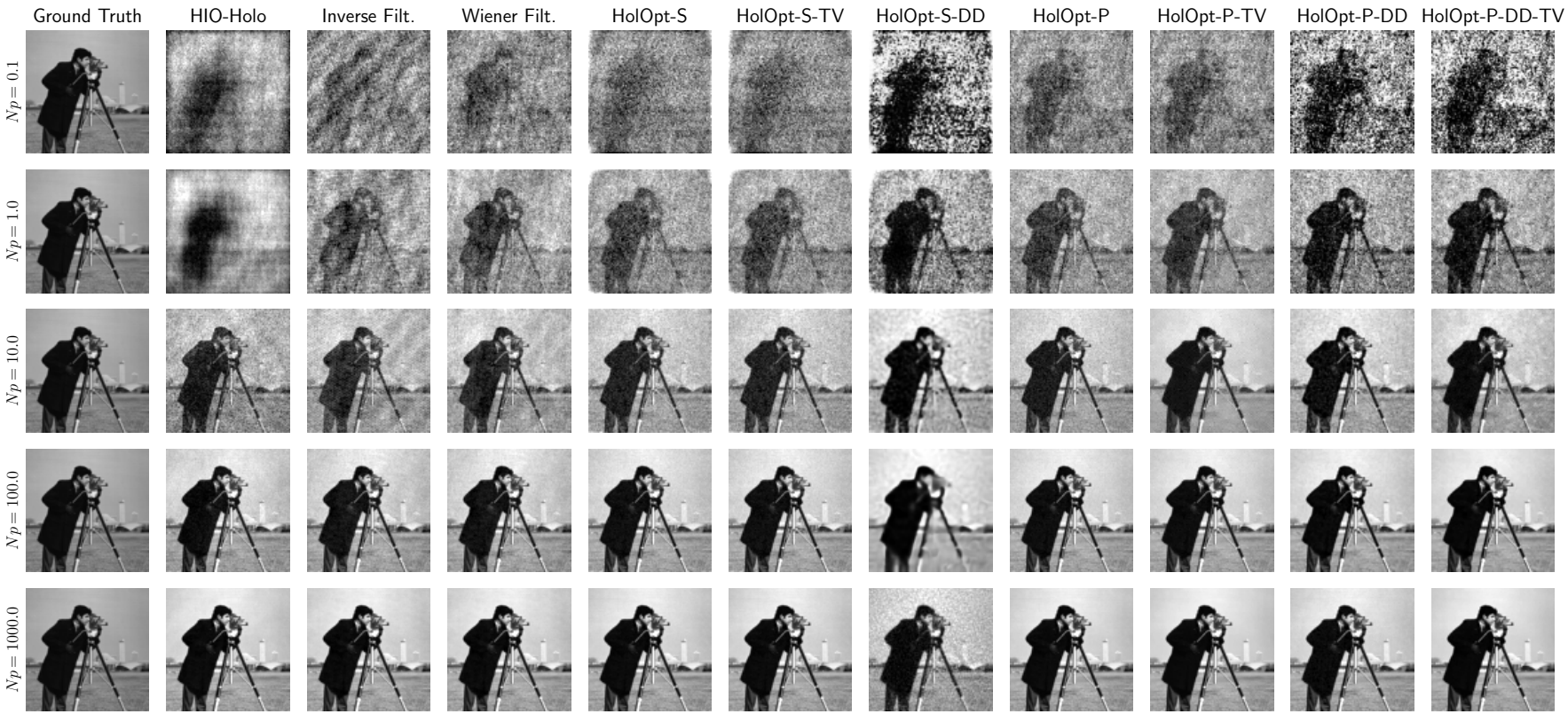


Figure: Comparing reconstructions across algorithms and noise levels on a sample image from SET12 (CAMERA) with a binary random reference and without beamstop. To improve contrast black and white are set respectively to first and last (99th) percentile of all pixels within each image. This convention is adopted for all subsequent visuals.

Reconstruction at different noise levels

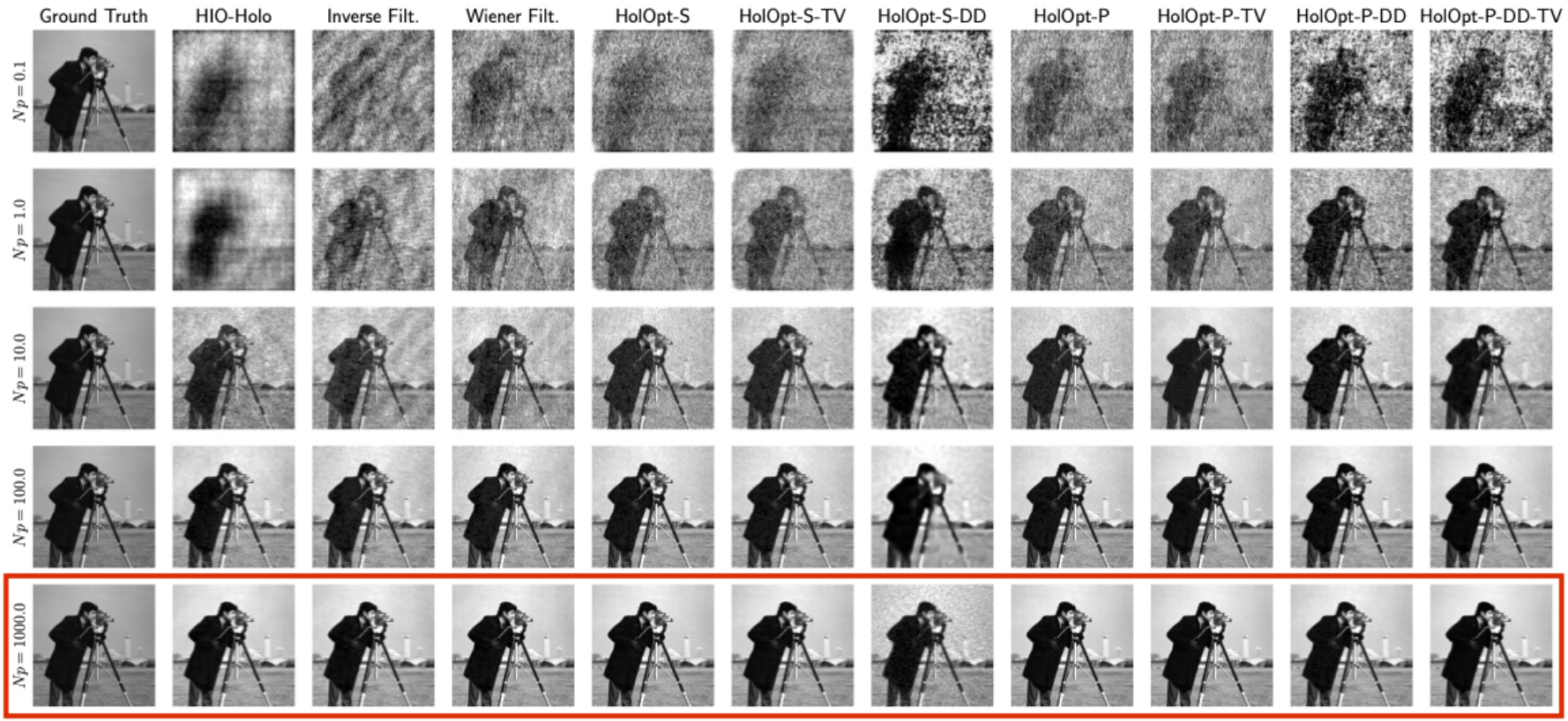


Figure: At low noise (high N_p), all methods do relatively well.

Reconstruction at different noise levels

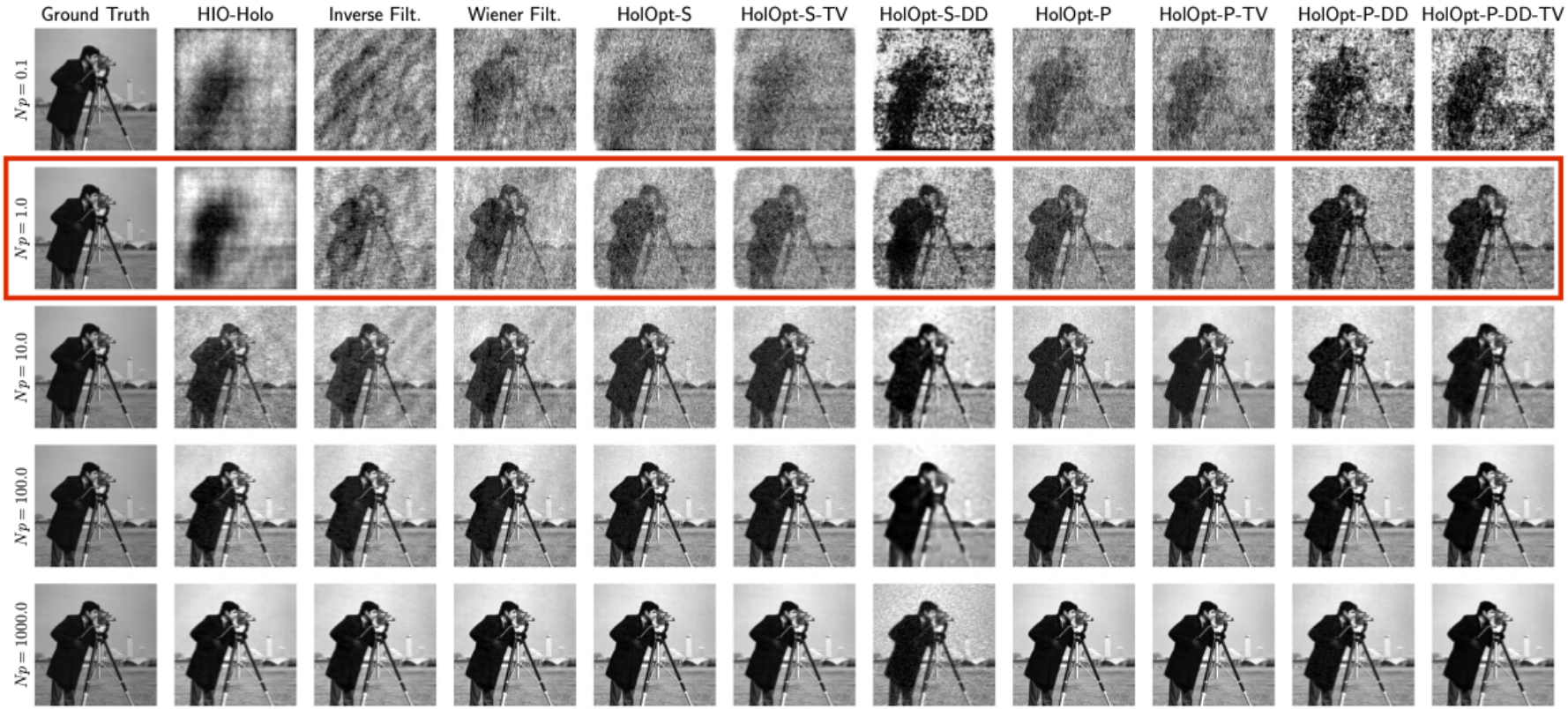


Figure: At high noise (low N_p), our methods (right-most columns) obtain cleaner reconstructions.

Reconstruction at different noise levels

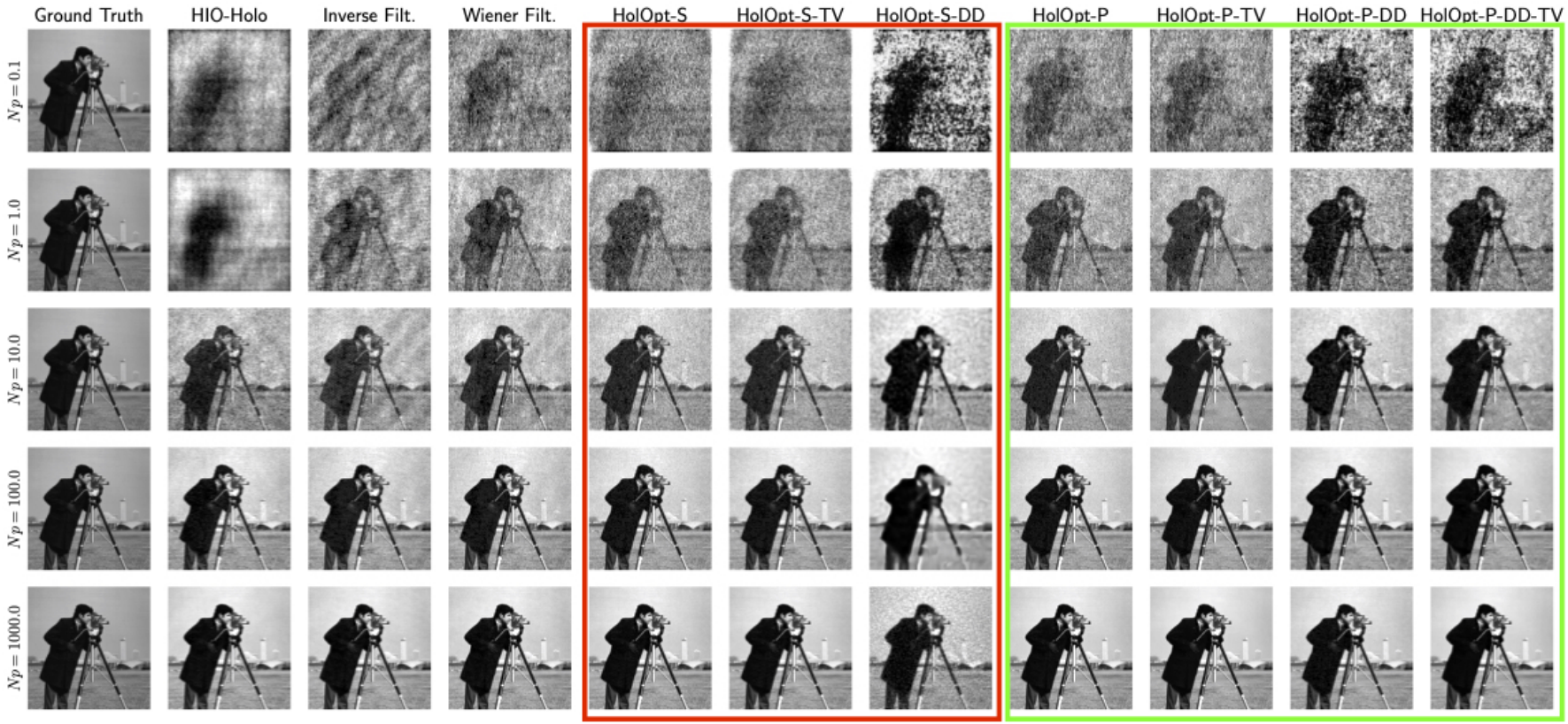


Figure: Poisson loss always performs at least as well as squared (MSE) loss. The squared loss sometimes leads to artifacts in the images reconstructed by HolOpt-S-DD.

Reconstruction at different noise levels

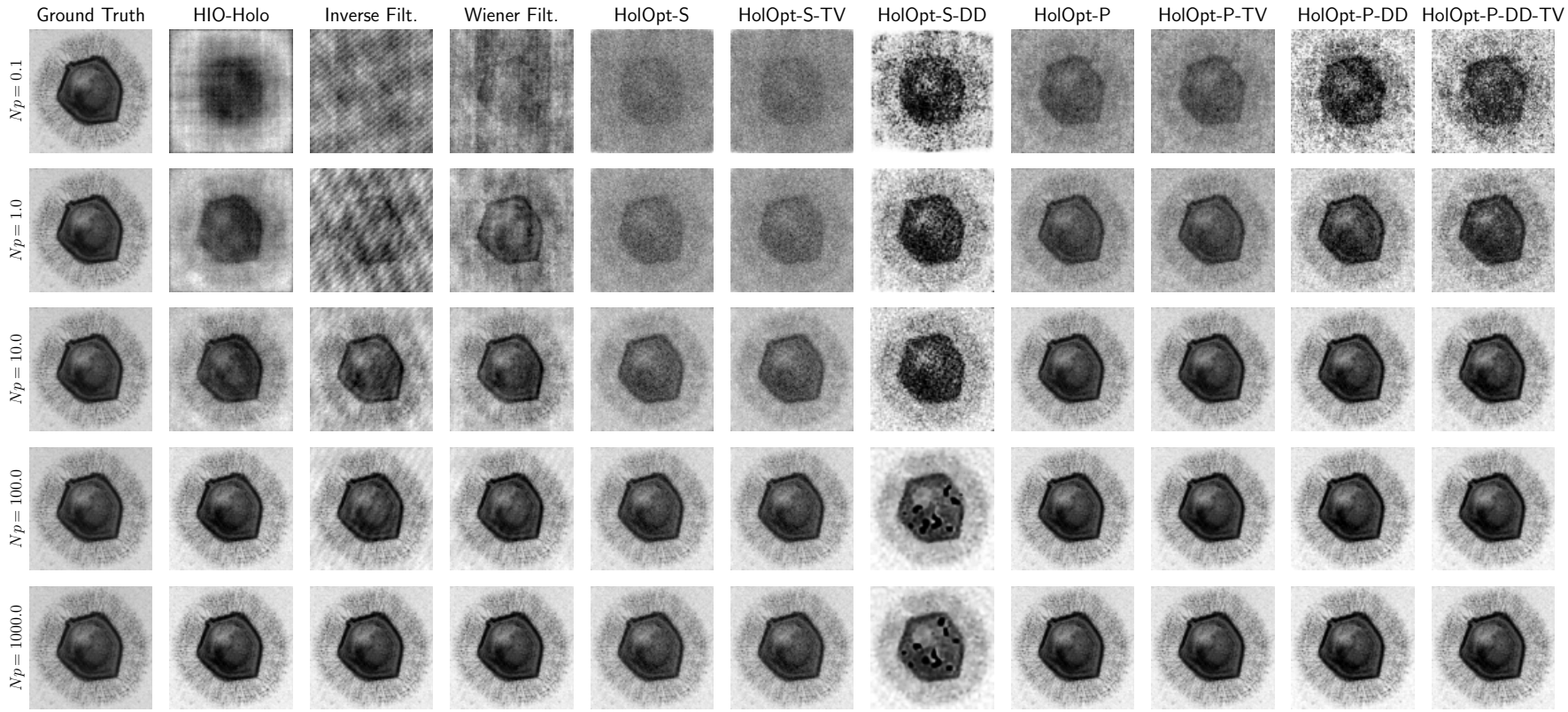


Figure: Comparing reconstructions across algorithms and noise levels on a sample image from from BIO10 (VIRUS) with a binary random reference and without beamstop.

Reconstruction at different noise levels

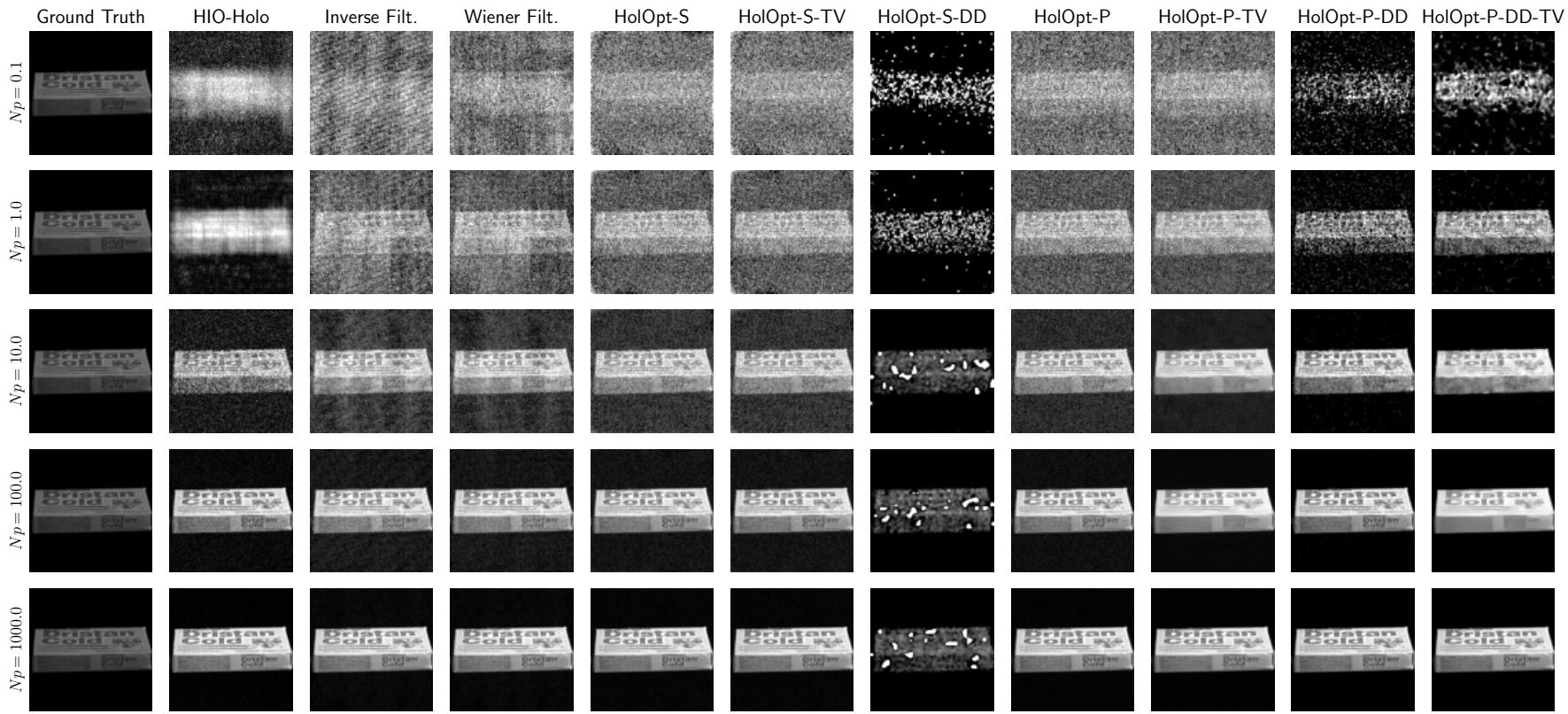


Figure: Comparing reconstructions across algorithms and noise levels on a sample image from COIL100 with a binary random reference and without beamstop.

Reconstruction at different noise levels

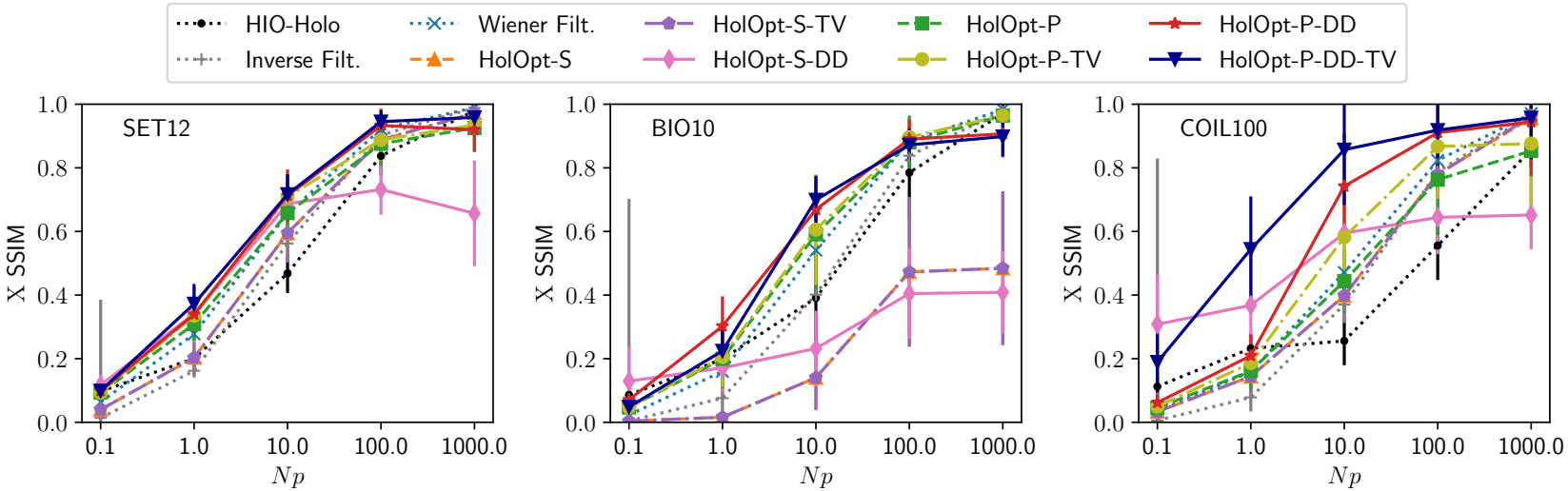


Figure: Reconstruction scores (SSIM) as a function of the photon count N_p with a binary random reference and without beamstop.

- HolOpt-S $<$ HolOpt-P and HolOpt-S-DD $<$ Holopt-P-DD
- Deep decoder is most helpful at high noise regime
- The relative performances of HolOpt-P-TV, Holopt-P-DD and Holopt-P-DD-TV are image-dependent: the deep decoder usually allows better contrast, while TV regularization smooths constant backgrounds (but sometimes also erases details)

Missing low-frequency data

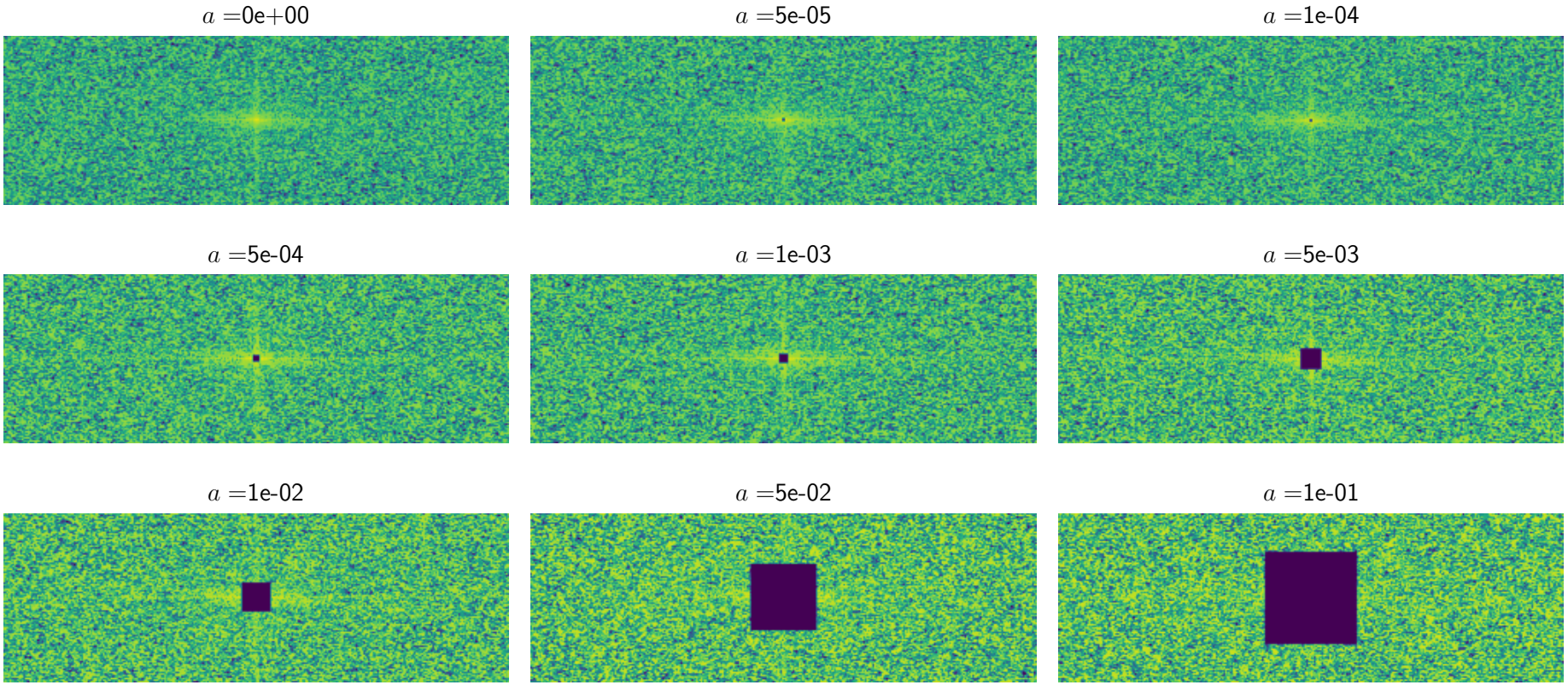


Figure: Examples of the measured magnitudes corresponding to varying beamstop area fractions for the SET12 images, with $N_p = 10$.

Missing low-frequency data

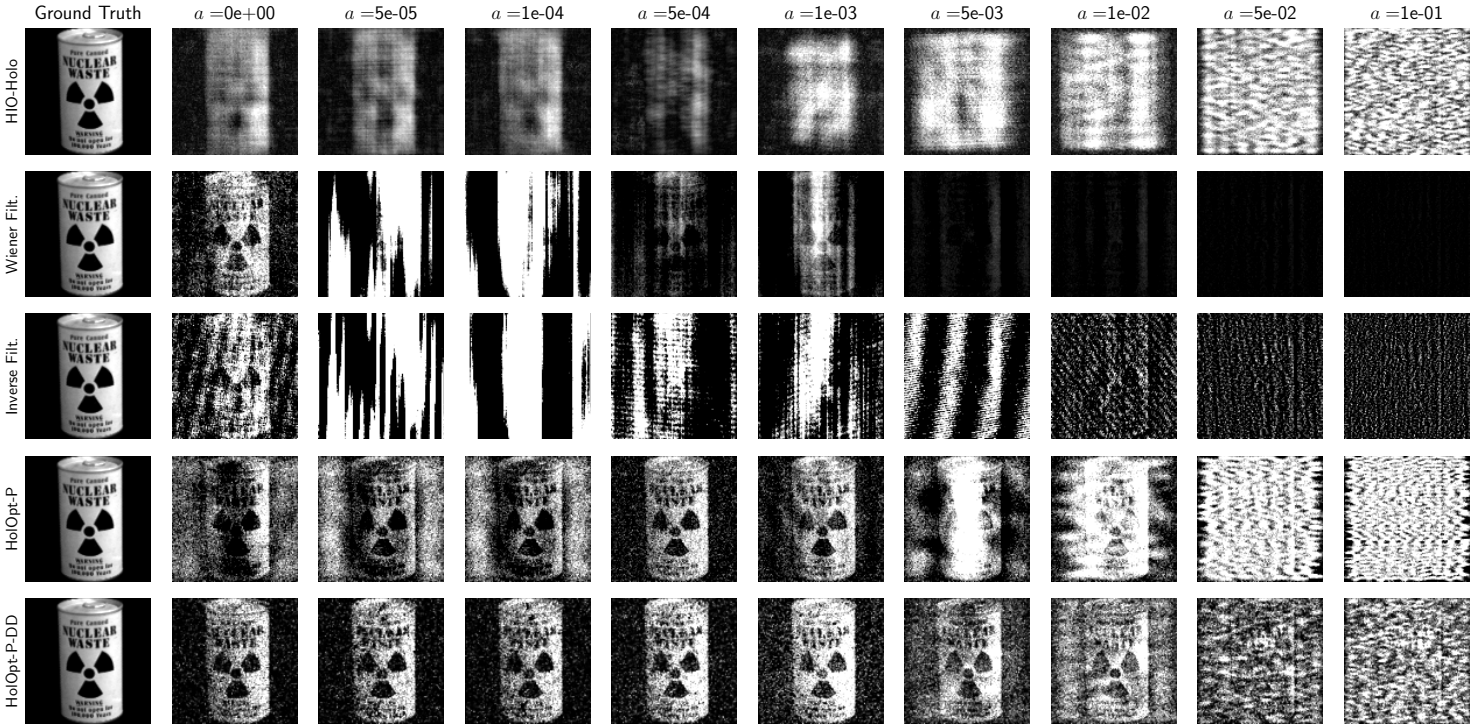


Figure: An example reconstructed image from COIL dataset as a function of beamstop area fraction a for fixed photon count $N_p = 1$. Our methods are shown on the bottom two rows.

Missing low-frequency data

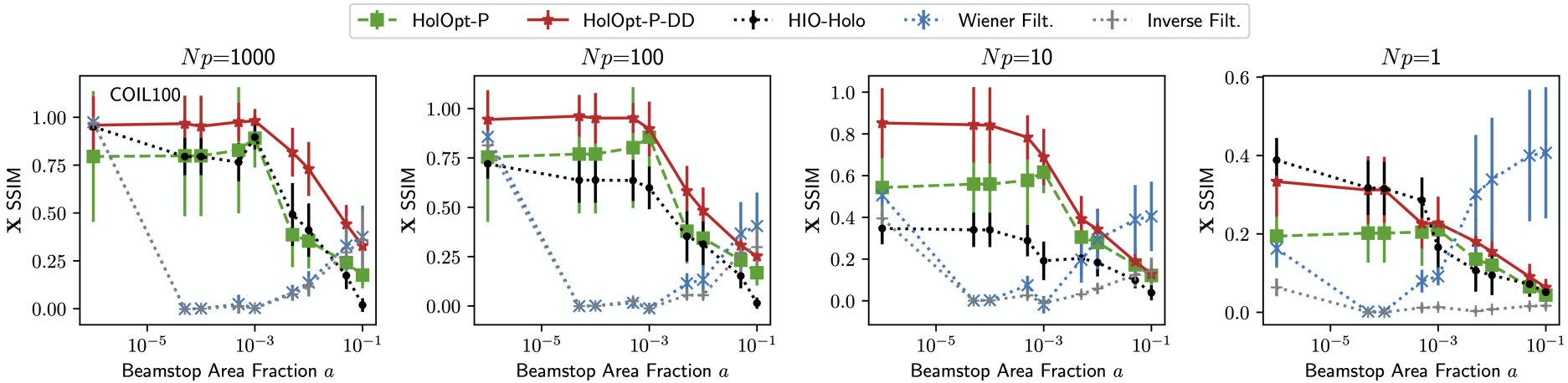


Figure: Reconstruction SSIM as a function of beamstop area fraction. Baseline methods are run for 5 trials per image while for tractability our methods are run 1 trial each on COIL100. Average SSIM and one standard deviation error bars are shown. The leftmost datapoints correspond to no missing data.

The performances of our methods smoothly degrade with increasing fraction of missing magnitudes a, and enable reconstructions with generally lower error and visually improved features relative to baselines. [3]

[3] At $N_p = 1$ and large a regime: visually, all methods fail; Wiener filtering advantage seems to be an idiosyncrasy of SSIM.

Breaking the holographic separation condition

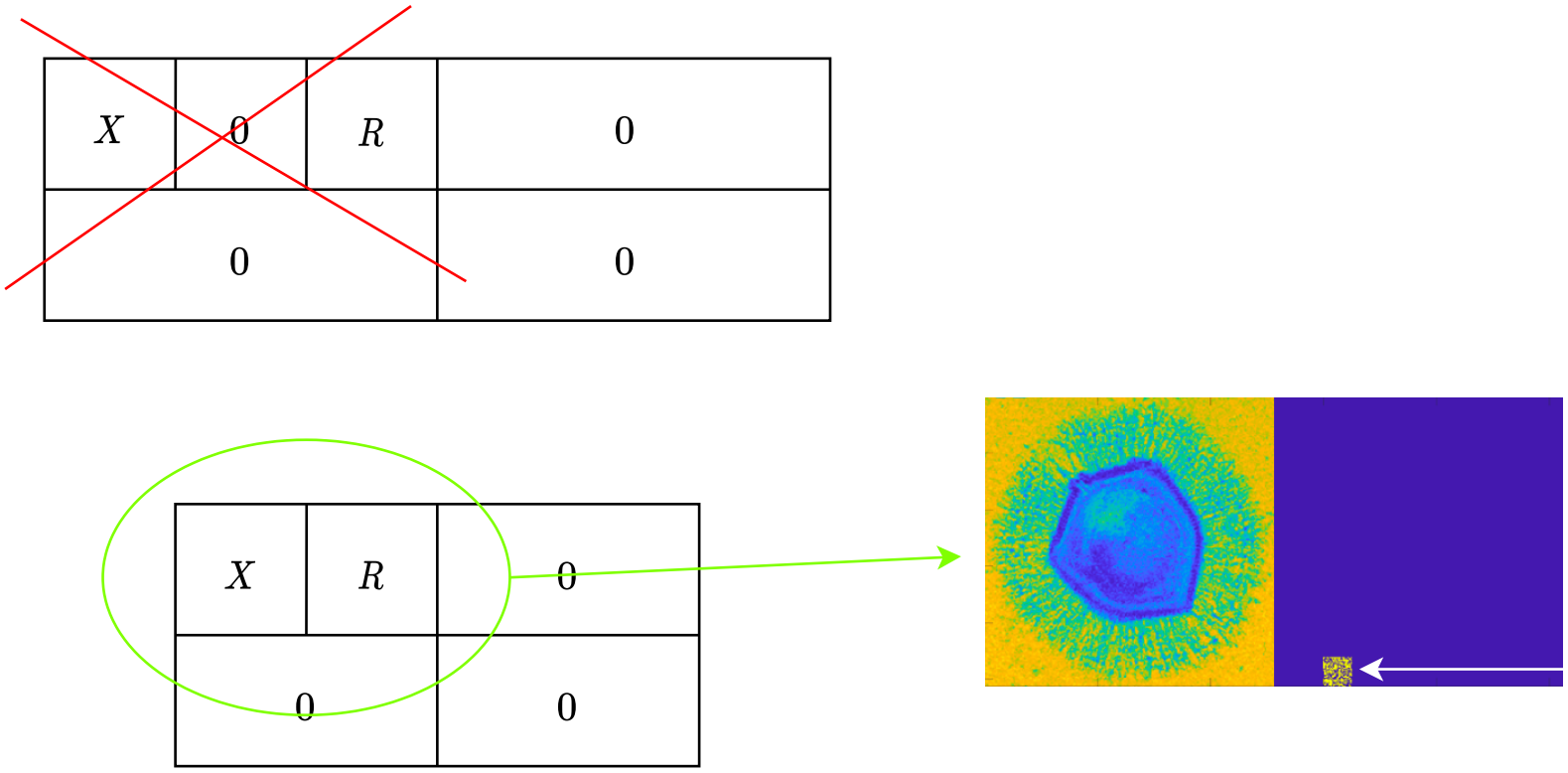


Figure: In this setting, we vary the position of a small, binary random reference.

Note: Wiener and inverse filtering do not apply without sufficient separation, conveying how challenging the setting is. As such, we can only compare to HIO-Holo.

Breaking the holographic separation condition

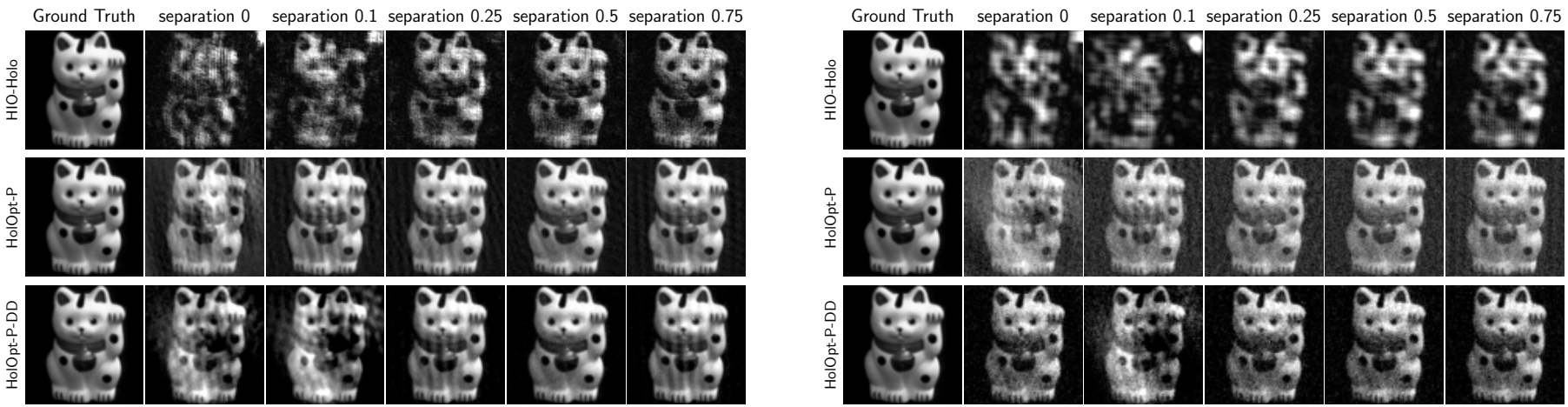


Figure: Reconstructions for photon counts $N_p = 10$ (left) and $N_p = 1$ (right) with a $0.1m \times 0.1m$ binary random reference as a function of the relative separation. A separation of 0.5 implies that the left-most non-zero pixel of the reference is $0.5n$ pixels away from the image.

Breaking the holographic separation condition

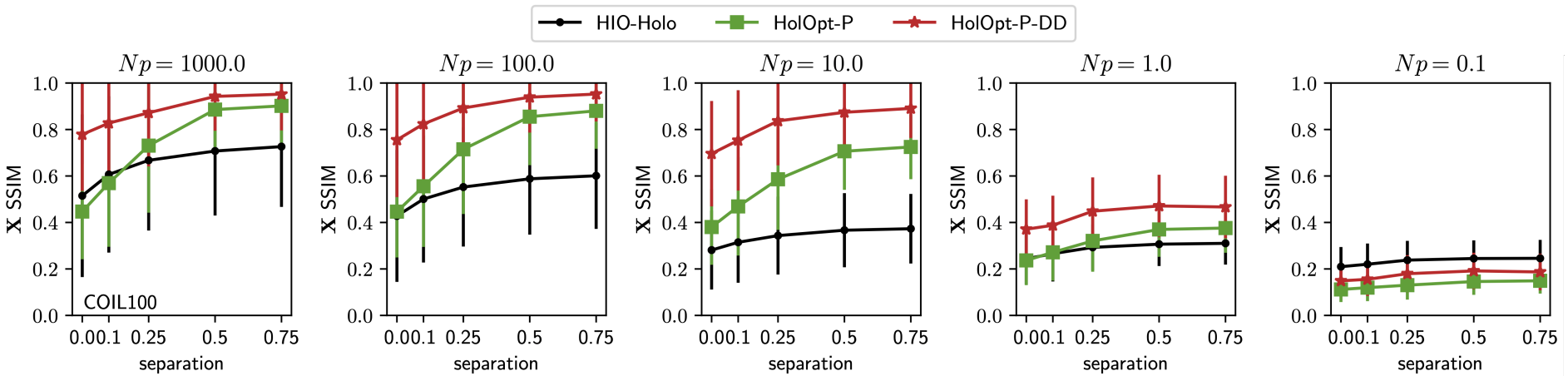
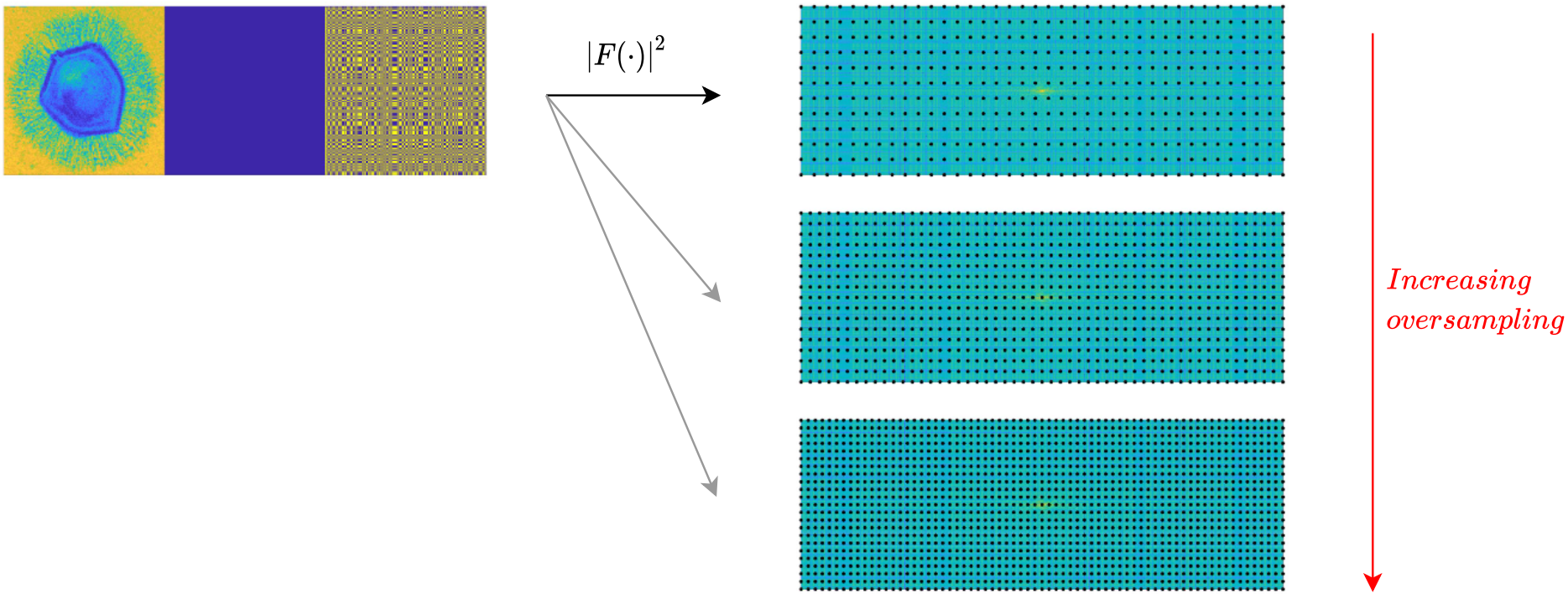


Figure: Reconstruction SSIM for decreasing photon counts N_p as a function of the relative separation (see caption of Figure 18).

Reduced oversampling rate



We consider oversampling ratios less than two (the standard), which correspond to less fine sampling of the same region of Fourier space. An oversampling ratio of r applied to an input of size $n \times 3n$ corresponds to measurements of size $rn \times 3rn$.

Reduced oversampling rate

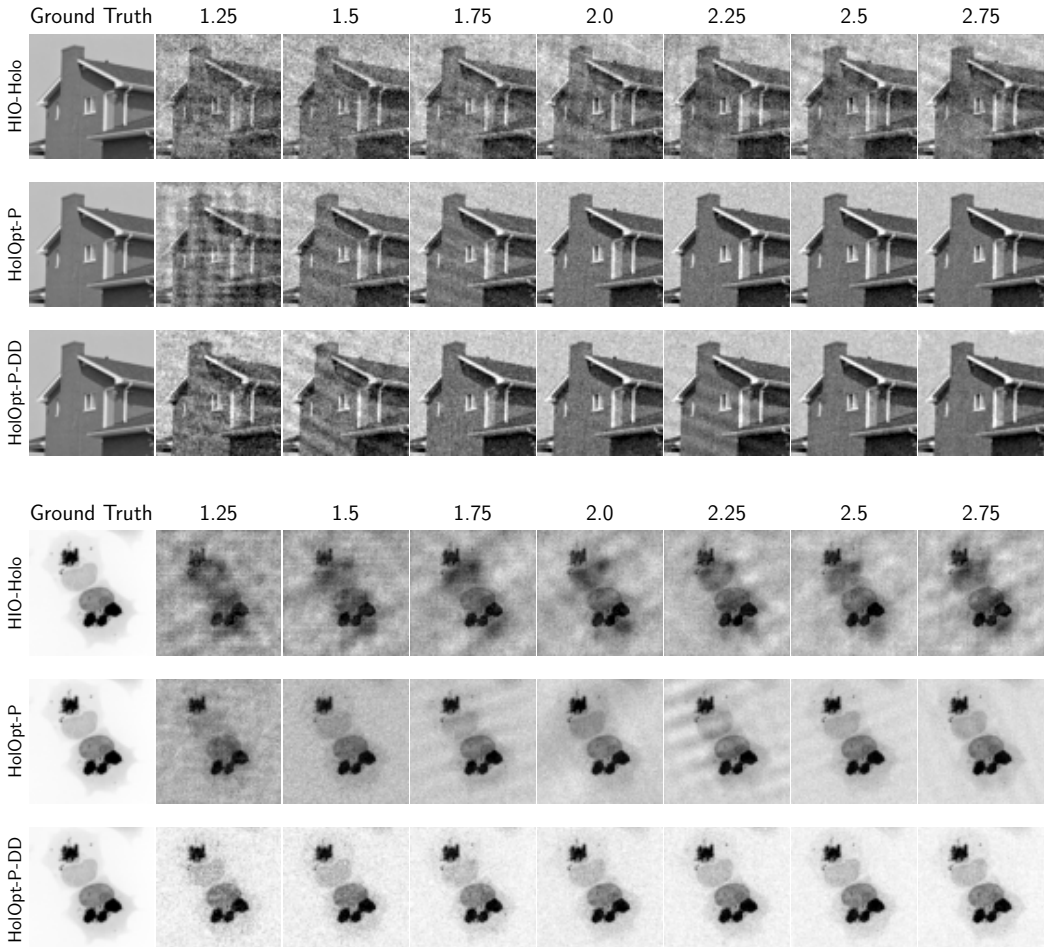


Figure: Reconstructed images for samples from SET12, BIO10 datasets with varying oversampling factors (numbers above each column) at $N_p = 10$ photon/pixel.

Reduced oversampling rate

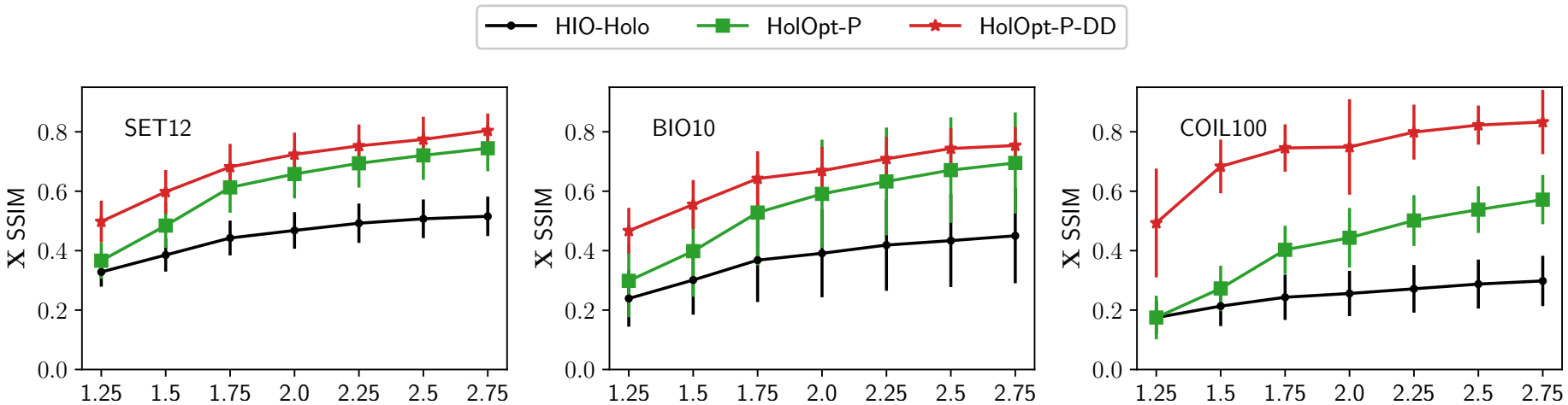
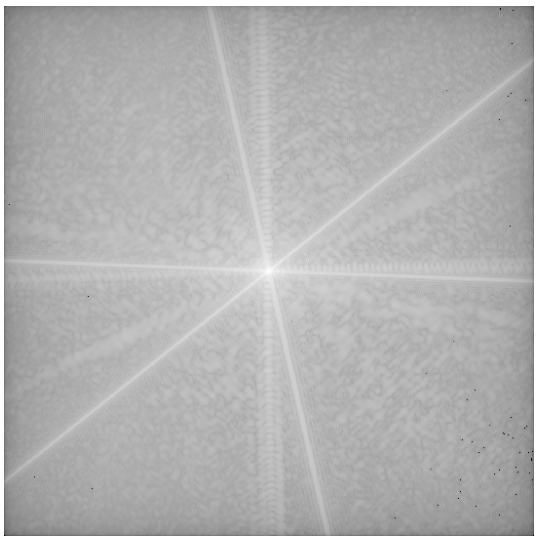
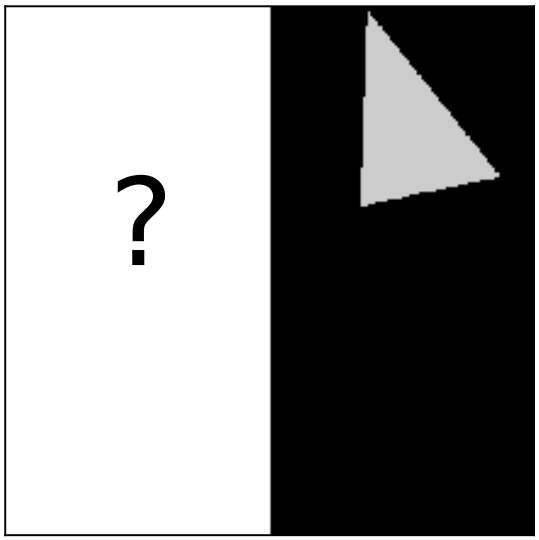


Figure: Average SSIM scores for SET12, BIO10 and COIL100 with varying oversampling factor at $N_p = 10$ photon/pixel. Errors bars represent standard deviations.

Real experimental data





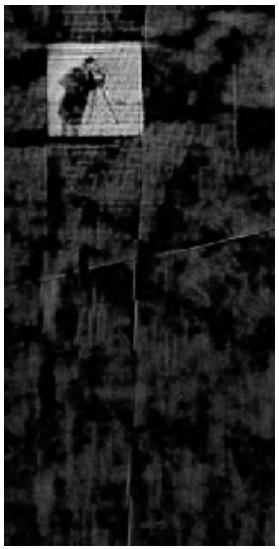


Figure: Reconstruction from the experimental data of Guizar-Sicairos and Fienup [2008], which used a He-Ne laser of 632.8 nm wavelength and a triangular reference. (a) Measurements (1024×1024 pixels), (b) Reference positioning (512×512 pixels), (c) Full reconstruction for HolOpt-P-DD-TV (512×256 pixels)

Real experimental data

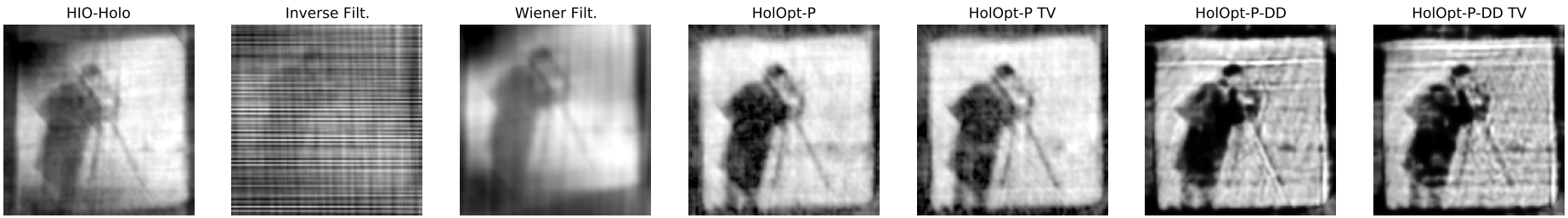


Figure: Reconstruction comparison from experimental data (He-Ne laser, 632.8 nm wavelength, from Guizar-Sicairos and Fienup [2008]) with a triangular reference.

We observe that TV-regularization alone noticeably smooths out the horizon line, effectively removing it, while the addition of a deep decoder prevents this phenomenon.

Conclusions

- We can obtain state-of-the-art reconstruction in holographic phase retrieval by combining a deep decoder prior, Poisson maximum likelihood, and automatic differentiation for easy optimization
- It is advantageous to harness the complete measurement and forward model information in holography, even if the resultant reconstruction problem is nonlinear
- Untrained image priors are confirmed to be powerful tools, especially when a significant amount of information is missing from the measured data due to low photon counts, beamstop-obscured frequencies and small oversampling

Future directions:
(1) evaluation on real data from other physical experiments, and
(2) theory for *Fourier* phase retrieval under a generative prior.

References + acknowledgements I

Acknowledgements: We are grateful to Charles Epstein, Leslie Greengard and Alex Barnett for insightful discussions. We also thank Manuel Guizar-Sicairos and Jim Fienup for sharing Heraldo experimental data, and Ju Sun for suggesting SSIM as an image metric and other helpful comments.

M Saliba, T Latychevskaia, J Longchamp, and H Fink. Fourier Transform Holography: A Lensless Non-Destructive Imaging Technique. *Microscopy and Microanalysis*, 18(S2):564–565, 2012. doi: 10.1017/S1431927612004679.

M. Hayes. The reconstruction of a multidimensional sequence from the phase or magnitude of its fourier transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 30(2):140–154, 1982.

References + acknowledgements II

Alexander H Barnett, Charles L Epstein, Leslie F Greengard, and Jeremy F Magland. Geometry of the phase retrieval problem. *Inverse Problems*, 36(9):094003, sep 2020. doi: 10.1088/1361-6420/aba5ed. URL `https://doi.org/10.1088/1361-6420/aba5ed`.

David A Barmherzig, Ju Sun, Po-Nan Li, T J Lane, and Emmanuel J Candès. Holographic phase retrieval and reference design. *Inverse Problems*, 35(9):094001, Aug 2019. doi: 10.1088/1361-6420/ab23d1. URL `https://doi.org/10.1088%2F1361-6420%2Fab23d1`.

Michael T. McCann, Kyong Hwan Jin, and Michael Unser. Convolutional neural networks for inverse problems in imaging: A review. *IEEE Signal Processing Magazine*, 34(6):85–95, 2017. ISSN 10535888. doi: 10.1109/MSP.2017.2739299.

References + acknowledgements III

Tobias Uelwer, Alexander Oberstraß, and Stefan Harmeling. Phase Retrieval using Conditional Generative Adversarial Networks. *arXiv preprint*, 1912.04981, 2019. URL `http://arxiv.org/abs/1912.04981`.

Yair Rivenson, Yibo Zhang, Harun Günaydın, Da Teng, and Aydogan Ozcan. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light: Science and Applications*, 7(2): 17141, 2018. ISSN 20477538. doi: 10.1038/lsa.2017.141.

Tim Meinhardt, Michael Moeller, Caner Hazirbas, and Daniel Cremers. Learning Proximal Operators: Using Denoising Networks for Regularizing Inverse Imaging Problems. In *2017 IEEE International Conference on Computer Vision (ICCV)*, number 12, pages 1799–1808. IEEE, oct 2017. ISBN 978-1-5386-1032-9. doi: 10.1109/ICCV.2017.198. URL `http://ieeexplore.ieee.org/document/8237460/`.

References + acknowledgements IV

Yaniv Romano, Michael Elad, and Peyman Milanfar. The little engine that could: Regularization by Denoising (RED). *SIAM Journal on Imaging Sciences*, 10(4):1804–1844, 2017. ISSN 19364954. doi: 10.1137/16M1102884.

Christopher A. Metzler, Philip Schniter, Ashok Veeraraghavan, and Richard G. Baraniuk. prDeep: Robust phase retrieval with a flexible deep network. *35th International Conference on Machine Learning, ICML 2018*, 8:5654–5663, 2018.

Yaotian Wang, Xiaohang Sun, and Jason W. Fleischer. When deep denoising meets iterative phase retrieval. In *International Conference on Machine Learning*, 2020a. URL `http://arxiv.org/abs/2003.01792`.

References + acknowledgements V

Baoshun Shi, Qiusheng Lian, and Huibin Chang. Deep prior-based sparse representation model for diffraction imaging: A plug-and-play method. *Signal Processing*, 168:107350, 2020. ISSN 01651684. doi: 10.1016/j.sigpro.2019.107350. URL `https://doi.org/10.1016/j.sigpro.2019.107350`.

Eric W. Tramel, Andre Manoel, Francesco Caltagirone, Marylou Gabrié, and Florent Krzakala. Inferring sparsity: Compressed sensing using generalized restricted Boltzmann machines. In *2016 IEEE Information Theory Workshop (ITW)*, pages 265–269. IEEE, sep 2016. ISBN 978-1-5090-1090-5. doi: 10.1109/ITW.2016.7606837. URL `http://ieeexplore.ieee.org/document/7606837/`.

References + acknowledgements VI

Victor Lempitsky, Andrea Vedaldi, and Dmitry Ulyanov. Deep Image Prior. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9446–9454. IEEE, jun 2018. ISBN 978-1-5386-6420-9. doi: 10.1109/CVPR.2018.00984. URL `https://box.skoltech.ru/index.php/s/ib52BOoV58ztuPM{#}pdfviewerhttps://ieeexplore.ieee.org/document/8579082/`.

Reinhard Heckel and Paul Hand. Deep Decoder: Concise Image Representations from Untrained Non-convolutional Networks. *International Conference on Learning Representations*, 2019. URL `http://arxiv.org/abs/1810.03982`.

References + acknowledgements VII

Fei Wang, Yaoming Bian, Haichao Wang, Meng Lyu, Giancarlo Pedrini, Wolfgang Osten, George Barbastathis, and Guohai Situ. Phase imaging with an untrained neural network. *Light: Science and Applications*, 9 (1), 2020b. ISSN 20477538. doi: 10.1038/s41377-020-0302-3. URL http://dx.doi.org/10.1038/s41377-020-0302-3.

Emrah Bostan, Reinhard Heckel, Michael Chen, Michael Kellman, and Laura Waller. Deep phase decoder: self-calibrating phase microscopy with an untrained deep neural network. *Optica*, 7(6):559, jun 2020. ISSN 2334-2536. doi: 10.1364/OPTICA.389314. URL https://www.osapublishing.org/abstract.cfm?URI=optica-7-6-559.

Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, apr 2004. ISSN 1057-7149. doi: 10.1109/TIP.2003.819861. URL http://ieeexplore.ieee.org/document/1284395/.

References + acknowledgements VIII

Manuel Guizar-Sicairos and James R. Fienup. Direct image reconstruction from a fourier intensity pattern using heraldo. *Opt. Lett.*, 15(33(22)): 2668–70, Nov 2008. doi: 10.1364/ol.33.002668. URL `https://pubmed.ncbi.nlm.nih.gov/19015703/`.