

PAPER • OPEN ACCESS

Joint ptycho-tomography with deep generative priors

To cite this article: Selin Aslan *et al* 2021 *Mach. Learn.: Sci. Technol.* **2** 045017

View the [article online](#) for updates and enhancements.

You may also like

- [Dynamic Fourier ptychography with deep spatiotemporal priors](#)
Pakshal Bohra, Thanh-an Pham, Yuxuan Long et al.
- [Color Fourier ptychographic microscopy based on symmetrical illumination and wavelength multiplexing](#)
Muyang Zhang, Di Yang and Yanmei Liang
- [Solving ptychography with a convex relaxation](#)
Roarke Horstmeyer, Richard Y Chen, Xiaoze Ou et al.



PAPER

Joint ptycho-tomography with deep generative priors

OPEN ACCESS

RECEIVED
5 March 2021REVISED
26 July 2021ACCEPTED FOR PUBLICATION
12 August 2021PUBLISHED
27 August 2021

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.

Selin Aslan^{1,*} , Zhengchun Liu², Viktor Nikitin¹ , Tekin Bicer^{1,2}, Sven Leyffer³  and Doğa Gürsoy^{1,4} ¹ X-ray Science Division, Argonne National Laboratory, 9700 Cass Avenue, Lemont, IL 60439, United States of America² Data Science and Learning Division, Argonne National Laboratory, Lemont, IL 60439, United States of America³ Mathematics and Computer Science Division, Argonne National Laboratory, 9700 Cass Avenue, Lemont, IL 60439, United States of America⁴ Department of Electrical Engineering and Computer Science, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, United States of America

* Author to whom any correspondence should be addressed.

E-mail: selinaslanphd@gmail.com**Keywords:** ptychography, tomography, imaging, generative models, learned priors**Abstract**

Joint ptycho-tomography is a powerful computational imaging framework to recover the refractive properties of a 3D object while relaxing the requirements for probe overlap that is common in conventional phase retrieval. We use an augmented Lagrangian scheme for formulating the constrained optimization problem and employ an alternating direction method of multipliers (ADMM) for the joint solution. ADMM allows the problem to be split into smaller and computationally more efficient subproblems: ptychographic phase retrieval, tomographic reconstruction, and regularization of the solution. We extend our ADMM framework with plug-and-play (PnP) denoisers by replacing the regularization subproblem with a general denoising operator based on machine learning. While the PnP framework enables integrating such learned priors as denoising operators, tuning of the denoiser prior remains challenging. To overcome this challenge, we propose a denoiser parameter to control the effect of the denoiser and to accelerate the solution. In our simulations, we demonstrate that our proposed framework with parameter tuning and learned priors generates high-quality reconstructions under limited and noisy measurement data.

1. Introduction

Ptychography [23] is a scanning-based coherent diffraction imaging technique that can provide high resolution imaging of thick samples without the need of an optic to form an image. While ptychography can only provide projective imaging of samples in 2D, series of ptychography scans can be acquired in a tomography setting to reconstruct thick volumetric samples in 3D [13]. In ptycho-tomography, a 3D object is scanned with a small coherent beam to collect a series of diffraction patterns through a pixel array detector located in the far-field; see figure 1. The detector records the intensity images of the incident wave on detector plane; therefore, the phase of the wave needs to be recovered through a computational procedure called the phase retrieval. This scanning procedure can be repeated for different view angles of the 3D object around a common rotation axis in order to collect tomographic data and to recover the complex refractive index of the object in 3D. The conventional approach for reconstruction then consists of solving a 2D ptychographic phase retrieval problem independently for each angle, followed by a 3D tomographic reconstruction from the retrieved angular projections of the phase (and amplitude) of the object plane wave. Because phase retrieval algorithms require significant overlap (60% or more) between neighboring illuminations for a successful recovery, the sequential approach is not optimal and limits scanning large volumes within reasonable data collection times.

While the sequential approach, that is, first performing phase retrieval for each angle and then tomographically reconstructing the object, is still the method of choice in practice, recent efforts have focused on relaxing or avoiding the illumination overlap requirement. These methods pose the

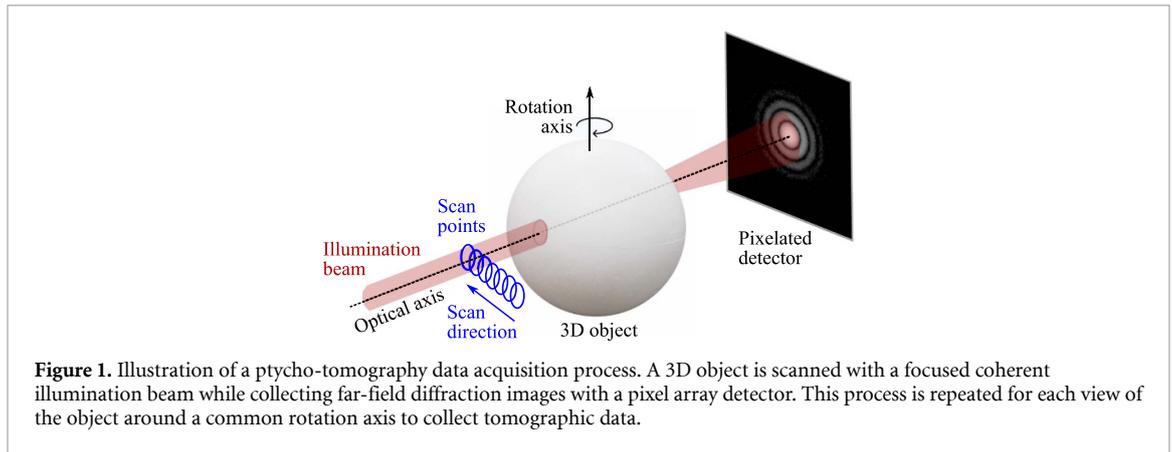
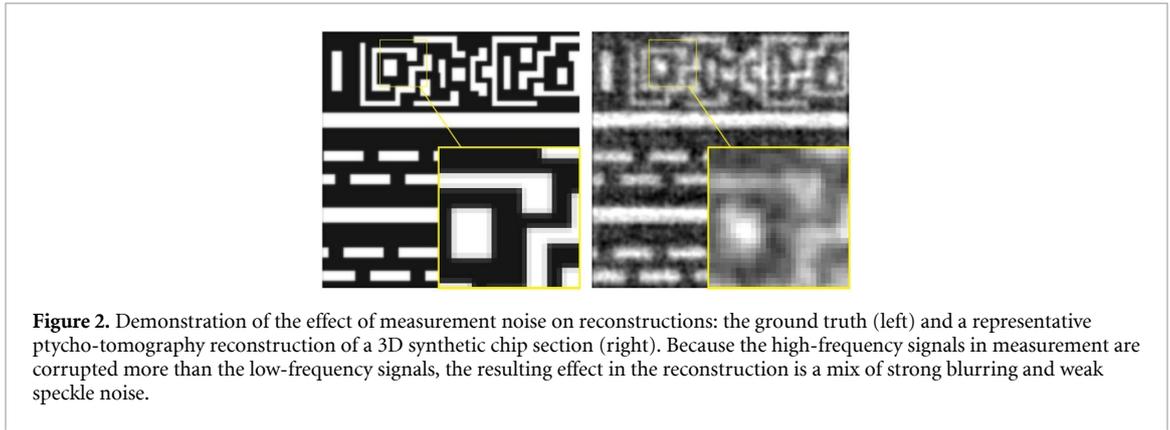


Figure 1. Illustration of a ptycho-tomography data acquisition process. A 3D object is scanned with a focused coherent illumination beam while collecting far-field diffraction images with a pixel array detector. This process is repeated for each view of the object around a common rotation axis to collect tomographic data.

reconstruction problem in a joint fashion. In other words, the phase retrieval problems for each rotation angle are solved simultaneously with the tomographic reconstruction through a joint optimization framework, resulting in a better-posed problem with those extra constraints and allowing for less stringent scanning requirements. Beginning with the first successful demonstration of the joint inversion concept through a numerical simulation [18], and later on experimentally [26], more efforts have focused on further relaxing these overlap constraints and finding new sparse scanning schemes for high-speed or dose-efficient implementations. Different optimizers such as the Levenberg–Marquardt algorithm [38] and Adam algorithm [14] have been used for successfully solving the joint optimization problem. In parallel, an extensible and a generic distributed optimization framework has been proposed in [2] as a solution when additional experimental uncertainties due to noise, motion blur, or other types of model mismatches need to be corrected. The framework is based on the alternating direction method of multiplier (ADMM) [5] and allows splitting the problem into smaller parts where each subproblem can be solved with an independent optimizer. With this modular structure, the whole reconstruction procedure can be expanded by adding new subproblems that often emerge in practical experimental settings. Also, because ADMM sub-problems can be solved independently, we can effectively map those sub-problems onto available computing resources.

Choosing an appropriate prior for the model is a major challenge for many imaging applications. To tackle this challenge, several regularization methods have been introduced. While some methods define priors explicitly in a regularized optimization framework such as total variation (TV) [41], Tikhonov regularization [48], and other types of sparsity-based regularization methods [49]; others do not have explicit formulation as an optimization problem, such as BM3D [10] and WNNM [17]. We also studied incorporating TV as part of our joint optimization scheme to regularize the solution when data points are significantly reduced or when data is heavily corrupted with noise [36]. Also recently, learning-based denoisers have been popularized because of their success in improving the quality of low-dose images [25]. Unlike physics-based optimization methods, learned priors are based on training a mapping between noisy images and a desirable image, and they are often applied after the reconstruction step is completed [30] or, in some cases, before the reconstruction in order to improve the raw data [53]. Furthermore, with the aid of special hardware, the reconstruction times can be improved significantly. One challenge in incorporating learned priors into the ADMM framework is that because the corresponding regularized optimization problem is not explicitly defined, a formal optimization strategy is not applicable. To overcome this challenge, Venkatakrishnan *et al* [50] proposed the plug-and-play (PnP) framework, which enables integrating implicit priors for denoising, to enable use of iterative optimization methods. Although the PnP framework was originally proposed ad hoc, it has been popularized quickly in various inverse problems because of its performance [6, 20, 27, 40, 45, 47, 51, 54]. This success has also led to related studies; for example, convergence of PnP has also been discussed in studies [6, 7, 42, 46]. Another related framework, regularization by denoising, has also been popularized to solve the denoising problem [32, 34, 39].

While the PnP framework provides flexible means to incorporate machine-learning-based denoising models into physics-based models, it has been mainly used for additive white Gaussian noise (AWGN) denoising of linear problems. In ptycho-tomography, or in phase retrieval problems in general, the problem is nonconvex and hard to solve optimally. In addition, the frequency spectrum of reconstruction noise is different from AWGN; see figure 2 for a representative example. This is because the measurements are taken in Fourier space; thus, high-frequency signals dampen quickly, and in turn they are more corrupted than the low-frequency signals because of the Poisson measurement statistics. Therefore, the reconstruction at high



measurement noise is blurry, and the state-of-the-art AWGN denoisers are not effective in addressing these types of noise in ptycho-tomography.

To address the unique challenges of solving ptycho-tomography problem at high noise levels, we propose using pre-trained generative prior models in an ADMM framework. As a generative model, we used conditional coupled generative adversarial network (GAN), however, other types of generative models may also be suitable for the task. We implemented this approach on graphical processing units (GPUs) and validated its effectiveness on realistic data sizes with highly sparse data and noisy measurements. We compare our results with the conventional offline denoising and TV regularization, which are commonly used for denoising in ptycho-tomography applications. Our results show that our optimizations can decrease the total number of required projections (with significantly fewer overlapped regions) by 75% compared with using adequately sampled data (based on Nyquist) while maintaining good image quality.

The remainder of this paper is organized as follows. In section 2, we give an overview of the joint ptycho-tomography problem and its solution using the ADMM method. Section 3 describes the challenges in using the original PnP framework and how we tackle the problem. The training, network design, and other important implementation details of the framework are given in section 4. In section 5, we validate our proposed framework for the joint ptycho-tomography problem via simulated experiments. Discussion and conclusions are given in section 6.

2. Background

In this section, we formulate the ptycho-tomography forward and inverse problems and describe the ADMM scheme for the reconstruction.

2.1. The forward problem

In the ptycho-tomography problem, the model for reconstructing the complex refractive index of a 3D object, $x = \delta + i\beta$, is given by:

$$\text{Poisson}\{|\mathcal{G}\mathcal{H}x|^2\} = d. \quad (1)$$

Here, we use a Poisson-based measurement model, which accurately captures photon-counting statistics in diffraction data in Fourier space. \mathcal{G} is the ptychography operator, \mathcal{H} is the tomography operator, x is the unknown object, and d is the measurement data. \mathcal{G} is defined as $\mathcal{G}\psi = \mathcal{F}\mathcal{Q}\psi$, where $\psi = \mathcal{H}x$ is the object transmission function, \mathcal{F} is the discrete Fourier transform operator, and \mathcal{Q} is the illumination matrix. \mathcal{H} is defined as $\mathcal{H}x = \exp(i\mathcal{R}x)$, where i is $\sqrt{-1}$, c is the wavenumber of the illumination beam, and \mathcal{R} is the Radon transform [21].

2.2. The inverse problem

Let $p(x|d)$ be the posterior conditional probability of having an object x with given measurements d . Then using Bayes's rule, the maximum *a posteriori* probability (MAP) estimate for the solution x_{MAP} is defined as follows:

$$\begin{aligned} x_{\text{MAP}} &= \arg \max_x \frac{p(d|x)p(x)}{p(d)} \\ &= \arg \min_x -\log\{p(d|x)\} - \log\{p(x)\}, \end{aligned} \quad (2)$$

where $\log p(d|x)$ is the log-likelihood of the observation and $\log p(x)$ is the prior of x , also referred to as the regularization term. The MAP estimate in equation (2) for the ptycho-tomography model in equation (1) is given as:

$$x_{\text{MAP}} = \arg \min_x \sum_{j=1}^n \left(|\mathcal{G}\mathcal{H}x|_j^2 - 2d_j \log |\mathcal{G}\mathcal{H}x|_j \right) + \varphi\mathcal{N}(x), \quad (3)$$

where $\varphi\mathcal{N}(x)$ is the regularization term to stabilize or to constrain the solution. For simplicity of notation, j indexes all measurement varieties, namely, detector pixel, rotation angle, and scan position. Next, we rewrite equation (3) into a consensus form by introducing auxiliary variables ψ and η :

$$\begin{aligned} & \min_{\psi, \eta, x} \sum_{j=1}^n \left(|\mathcal{G}\psi|_j^2 - 2d_j \log |\mathcal{G}\psi|_j \right) + \varphi\mathcal{N}(\eta), \\ & \text{subject to} \quad \begin{cases} \mathcal{H}x = \psi, \\ x = \eta. \end{cases} \end{aligned} \quad (4)$$

The objective function is a real-valued function of complex variables, and its augmented Lagrangian is a complex-valued function. We follow [29] and work with the following real-valued augmented Lagrangian:

$$\begin{aligned} \mathcal{L}_{\rho, \tau}^{\lambda, \mu}(\psi, x, \eta) = & \sum_{j=1}^n \left(|\mathcal{G}\psi|_j^2 - 2d_j \log |\mathcal{G}\psi|_j \right) + \varphi\mathcal{N}(\eta) \\ & + 2\text{Re}\{\lambda^H(\mathcal{H}x - \psi)\} + \rho \|\mathcal{H}x - \psi\|_2^2 \\ & + 2\text{Re}\{\mu^H(x - \eta)\} + \tau \|x - \eta\|_2^2, \end{aligned} \quad (5)$$

where $\rho > 0$ and $\tau > 0$ are penalty parameters, λ and μ represent dual variables, and H corresponds to the Hermitian conjugate. This augmented Lagrangian enables us to include the linear terms, $2\text{Re}\{\lambda^H(\mathcal{H}x - \psi)\}$, $\rho \|\mathcal{H}x - \psi\|_2^2$, and $2\text{Re}\{\mu^H(x - \eta)\}$, $\tau \|x - \eta\|_2^2$ in the L2-terms.

2.3. Solution to the inverse problem

Minimization of equation (5) can be achieved by ADMM with iteratively solving the sub-problems followed by dual variable updates:

$$\psi^{k+1} = \arg \min_{\psi} \sum_{j=1}^n \left(|\mathcal{G}\psi|_j^2 - 2d_j \log |\mathcal{G}\psi|_j \right) + \rho \|\mathcal{H}x^k - \psi + \lambda^k / \rho\|_2^2, \quad (6)$$

$$x^{k+1} = \arg \min_x \rho \|\mathcal{H}x - \psi^{k+1} + \lambda^k / \rho\|_2^2 + \tau \|x - \eta^k + \mu^k / \tau\|_2^2, \quad (7)$$

$$\eta^{k+1} = \arg \min_{\eta} \varphi\mathcal{N}(\eta) + \tau \|x^{k+1} - \eta + \mu^k / \tau\|_2^2, \quad (8)$$

$$\lambda^{k+1} = \lambda^k + \rho (\mathcal{H}x^{k+1} - \psi^{k+1}), \quad (9)$$

$$\mu^{k+1} = \mu^k + \tau (x^{k+1} - \eta^{k+1}). \quad (10)$$

Using the ADMM framework, we formulate the joint ptycho-tomography problem in equation (2) in terms of three independently defined subproblems: ptychographic phase retrieval in equation (6), tomographic reconstruction in equation (7), and regularization in equation (8). The dual variable updates promote the satisfaction of the constraints in equations (9) and (10).

2.4. Solutions of the subproblems

For the first subproblem, we minimize the following objection function:

$$F_P(\psi) = \sum_{j=1}^n \left(|\mathcal{G}\psi|_j^2 - 2d_j \log |\mathcal{G}\psi|_j \right) + \rho \|\mathcal{H}x^k - \psi + \lambda^k / \rho\|_2^2. \quad (11)$$

The corresponding gradient is:

$$\nabla_{\psi} F_P(\psi) = \mathcal{G}^H \left(\mathcal{G}\psi - \frac{d}{(\mathcal{G}\psi)^*} \right) - \rho(\mathcal{H}x^k - \psi + \lambda^k/\rho), \quad (12)$$

which is computed by using the Wirtinger calculus [24]. Here * denotes the complex conjugate. For the solution, we use the nonlinear conjugate gradient (CG) method [37]:

$$\psi_{m+1} = \psi_m + \gamma_m \xi_m, \quad (13)$$

where γ_m is a step length computed via a backtracking line search method and ξ_m is the search direction. The first iteration is the steepest descent direction, $\xi_0 = -\nabla_{\psi} F_P(\psi_0)$. For other iterations, ξ_{m+1} is computed recursively by using the Dai–Yuan [11] formula, which gives the fastest convergence in our simulations:

$$\xi_{m+1} = -\nabla_{\psi} F_P(\psi_{m+1}) + \frac{\|\nabla_{\psi} F_P(\psi_{m+1})\|_2^2}{y_m^H \xi_m} \xi_m, \quad (14)$$

where $y_m = (\nabla_{\psi} F_P(\psi_{m+1}) - \nabla_{\psi} F_P(\psi_m))$.

For solving the subproblem with respect to x in equation (7), we transform the nonlinearity introduced by $\mathcal{H}x$ as in [36] and instead minimize the following objection function:

$$F_T(x) = \rho \|\mathcal{K}\mathcal{R}x - \zeta\|_2^2 + \tau \|x^{k+1} - \eta + \mu^k/\tau\|_2^2, \quad (15)$$

where the linear diagonal operator \mathcal{K} is defined as :

$$\mathcal{K}\mathcal{R}x = \frac{2\pi i}{\nu} (\psi^{k+1} - \lambda^k/\rho)\mathcal{R}x, \quad (16)$$

and ζ is given by:

$$\zeta = (\psi^{k+1} - \lambda^k/\rho) \log(\psi^{k+1} - \lambda^k/\rho). \quad (17)$$

Hence, we replace the objective function in equation (7) with equation (15). The gradient is given as follows:

$$\nabla_x F_T(x) = \rho \mathcal{R}^T \mathcal{K}^H (\mathcal{K}\mathcal{R}x - \zeta) + \tau (x - \eta^k + \mu^k/\tau). \quad (18)$$

Similar to the ptychography subproblem, we use the CG method with the Dai–Yuan formula; see equations (13) and (14).

While equations (6) and (7) can be solved via well-known optimization methods, the solution of equation (8) depends on the choice of the image prior. The question of how to choose a prior, $-\log\{p(x)\} = \varphi\mathcal{N}(\eta)$ is a challenging topic in image processing. While one can choose an explicit image prior and measure its distance using the TV norm, we turn our attention to learning-based priors because of their effectiveness.

3. Learned priors for denoising

In this section, we discuss the solution of the denoising problem. We first rewrite equation (8) for some prior $\mathcal{N}(\eta)$ as follows:

$$\eta^{k+1} = \arg \min_{\eta} \mathcal{N}(\eta) + \tau/\varphi \|\tilde{x}^{k+1} - x\|_2^2, \quad (19)$$

where $\tilde{x}^{k+1} = x^{k+1} + \mu^k/\tau$ and x correspond to the noisy and noise-free images, respectively. Several state-of-the art denoisers do not have closed-form expressions for the prior, $\mathcal{N}(\eta)$. Hence, integrating these denoisers into the joint ptycho-tomography problem is challenging. We use the PnP framework [50] to replace equation (19) with a general denoising operator as follows:

$$\eta^{k+1} = \text{Denoiser}(\tilde{x}^{k+1}), \quad (20)$$

where an explicit definition of the image prior, $\mathcal{N}(\eta)$, is not necessarily known. While PnP was originally proposed to remove the AWGN of variance, $\sigma^2 = \tau/2\varphi$, the method has been extended to Poisson inverse problems [40]. In this work, we use a Poisson-based MAP model to accurately capture photon-counting statistics in diffraction data. While we still use the ADMM to solve equation (4) and while the first two

subproblems corresponding to the ptychographic phase retrieval and tomography are the same, the last sub-problem corresponding to the regularization is replaced with a denoising operator in equation (20). The PnP framework allows us to use state-of-the-art denoising algorithms, such as BM3D [10], K-SVD [15], and WNNM [17]. Although the PnP framework does not give a clear definition of the objective function because of the implicit regularization parameter, the method has shown empirical success in various image reconstruction problems [20, 27, 40, 51]. Alternatively, deep-learning-based denoisers have shown great success implementing the PnP framework; see, for example, [8, 33, 56]. In this work, we use our recently developed denoising technique based on GAN, whose implementation details will be discussed in the following section.

We point out that the regularization parameter, φ , that tunes the regularization term in equation (3) is associated with the additive noise in the denoising operator, $\sigma^2 = \tau/2\varphi$. In our application, we have observed that replacing the regularizer problem, equation (19), by the denoising operator using equation (20) can lead to divergence of the overall ADMM scheme. In particular, it appears that the denoiser pushes early iterations to nonphysical solutions from which the ADMM cannot recover. This observation motivates the introduction of a denoising parameter, $\alpha^k \in [0, 1]$, that controls the influence of the denoising operator. Moreover, ADMM can reach a modest accuracy even when the individual subproblems do not converge to optimal values [5], and this fact has been used for accelerating ptycho-tomography reconstruction [2]. When acceleration is used, however, the role of the denoiser for approximate solutions of the subproblems needs to be balanced for stabilizing the solution. To this end, we choose a denoiser parameter that gives weight to the data fidelity term at earlier iterations and gradually increases the weight of the denoiser at later iterations as we get closer to the solution. An alternative approach has also been proposed in [52] when $\mathcal{N}(\eta)$ is a closed, convex, and proper function. In particular, we rewrite equation (20) as

$$\eta_{k+1} = \alpha^k \text{Denoiser}(\tilde{x}^{k+1}) + (1 - \alpha^k) \tilde{x}^{k+1}, \quad (21)$$

which makes η_{k+1} a convex combination of the denoised reconstructions and the noisy reconstructions, \tilde{x}^{k+1} . The extremes $\alpha^k = 0$ and $\alpha^k = 1$ corresponds to the maximum likelihood (ML) estimate (i.e. no regularizer) and full denoising (i.e. PnP denoiser), respectively. In our implementations, we heuristically choose α^k to provide fast convergence to good reconstructions.

One challenge that arises from including equation (21) in the ADMM framework is that it does not directly correspond to an optimization problem (unless the denoiser can be written as a gradient) and therefore cannot directly be included in the augmented Lagrangian in equation (5). This makes it harder to generalize the traditional augmented Lagrangian or ADMM convergence theory.

4. Implementation

In this section, we discuss the implementation aspects of our approach; see algorithm 1. For the ptychography and tomography subproblems, we use the same solvers (CG) as in our previous work; see section 4 in [36]. Hence, we devote this section to the details of the denoising operator used in equation (21).

In our simulations, we use our recently developed denoiser, TomoGAN [30], an image-quality enhancement model based on generative adversarial networks [16], which was originally developed for low-dose x-ray imaging as the learned prior. Figure 3 shows the training pipeline of the model where two neural networks (i.e. generator and discriminator) contend with each other during the training until an equilibrium is reached. Specifically, the generative network generates noise-free images from noisy images while the discriminative network evaluates them; thus both networks are trained from the competition. The VGG [44] is a neural network model with 19 convolutional neural network (CNN) layers followed by three fully connected layers for image classification. Here, the VGG was pretrained with the ImageNet dataset [12], and we only keep the 19 CNN layers to work as a feature extractor for quantifying the difference between denoised image and true image in VGG's feature space. The generator model will work as the learned prior (i.e. for denoising in equation (21)) once trained by using the pipeline. That is, we can input a noisy image to the generator, and it outputs the corresponding enhanced image.

The TomoGAN generator network architecture is a variation of the U-Net architecture proposed for biomedical image segmentation by Shan *et al* [43]. It comprises a down-sampling network followed by an up-sampling network. In the down-sampling process, three sets of two convolution kernels (the three boxes) extract feature maps. Then, followed by a pooling layer, the feature map projections are distilled to the most essential elements by using a signal maximizing process. Ultimately, the feature maps are 1/8 of the original size. Successful training should result in the 128 channels in this feature map, retaining important features. In the up-sampling process, bilinear interpolation is used to expand the feature maps. At each layer, high-resolution features from the down-sampling path are concatenated to the up-sampled output from the

Algorithm 1. Joint Ptycho-tomography reconstruction with learned prior.

Require: Given $0 \leq \alpha^k \leq 1, \rho > 0, \tau > 0$ and initialize: $\psi^0, \eta^0, x^0, \lambda^0, \mu^0$

while not converged **do**

$$\psi^{k+1} \leftarrow \arg \min_{\psi} \sum_{j=1}^n (|\mathcal{G}\psi|_j^2 - 2d_j \log |\mathcal{G}\psi|_j) + \rho \left\| \mathcal{H}x^k - \psi + \lambda^k / \rho \right\|_2^2$$

$$x^{k+1} \leftarrow \arg \min_x \rho \left\| \mathcal{H}x - \psi^{k+1} + \lambda^k / \rho \right\|_2^2 + \tau \left\| x - \eta^k + \mu^k / \tau \right\|_2^2$$

$$\tilde{x}^{k+1} \leftarrow x^{k+1} + \mu^k / \tau$$

for $j = 1 \cdots M$ **do**

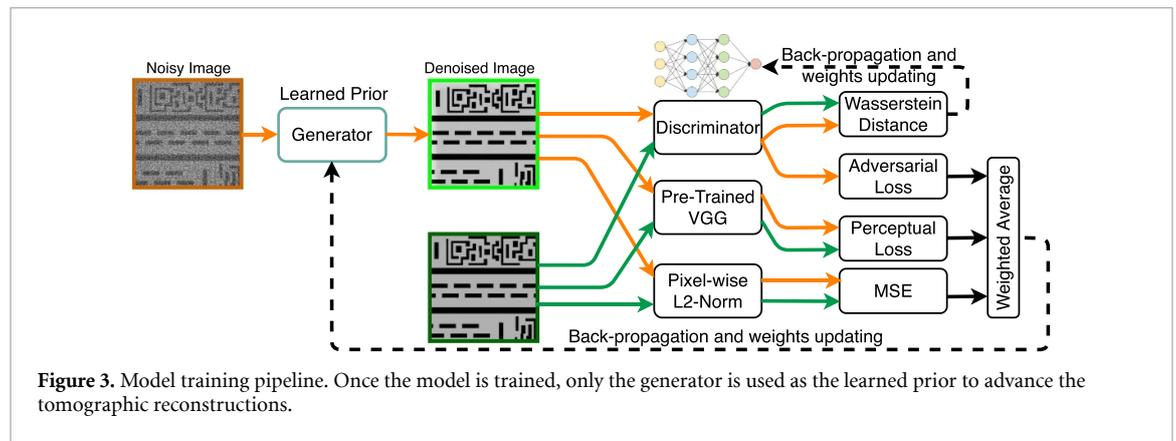
$$\eta_j^{k+1} \leftarrow \alpha^k \text{Denoiser}(\tilde{x}_j^{k+1}) + (1 - \alpha^k) \tilde{x}_j^{k+1}$$

end for

$$\lambda^{k+1} \leftarrow \lambda^k + \rho (\mathcal{H}x^{k+1} - \psi^{k+1})$$

$$\mu^{k+1} \leftarrow \mu^k + \tau (x^{k+1} - \eta^{k+1})$$

end while



layer below to form a large number of feature channels. This structure allows the network to propagate context information to higher-resolution layers, so that the following convolution layer can learn to assemble a more precise output based on this information. The detailed TomoGAN generator architecture can be found in [30].

We implemented TomoGAN with TensorFlow [1] and used one NVIDIA Tesla V100 GPU card for training where the total training time is around 6 h. The Adam algorithm [28] was used to train both the generator and discriminator, with a batch size of 16 samples. In order to train and evaluate the model (discussed in section 5), we synthesized two different 3D samples as shown in figure 4. For each sample, we simulated two different cases with different features (e.g. removed a few features of the chip, and used different random seeds to construct circles of the phantom object) for model training and testing separately. As a data augmentation to avoid overfitting, each image of the batch is a patch (of size 128×128) that was randomly cropped from the original 512×512 image (i.e. slice of the 3D objects).

5. Numerical experiments

In this section, we demonstrate the effectiveness of applying the proposed framework for reconstruction of 3D simulated objects in figure 4.

5.1. Simulation settings

In the first experiment, the object is a simulated chip of size $64 \times 512 \times 512$ and voxel size 5 nm. The 3D simulated chip and its 2D slice are given in figure 4(a). Our interest is to recover the object that is defined by its complex refractive index, $x = \delta + i\beta$. We use a flat-top Gaussian probe function with probe size 16×16 pixels. The far-field diffraction patterns are recorded by a 128×128 pixelated detector. We use 8.8 keV beam energy to simulate the refractive index values for ptychographic data. We emulate a ptychographic experiment by simulating a 3D chip, where δ yields the main imaging contrast. We distort the data with Poisson noise. In figure 5, we demonstrate the effect of Poisson noise on the measured data for three different detector photon counts in the ranges $I = [0, 8644]$, $I = [0, 968]$, and $I = [0, 123]$ on average. As the interval I decreases, the simulations become noisier. Initially, the distance between adjacent probe center positions is

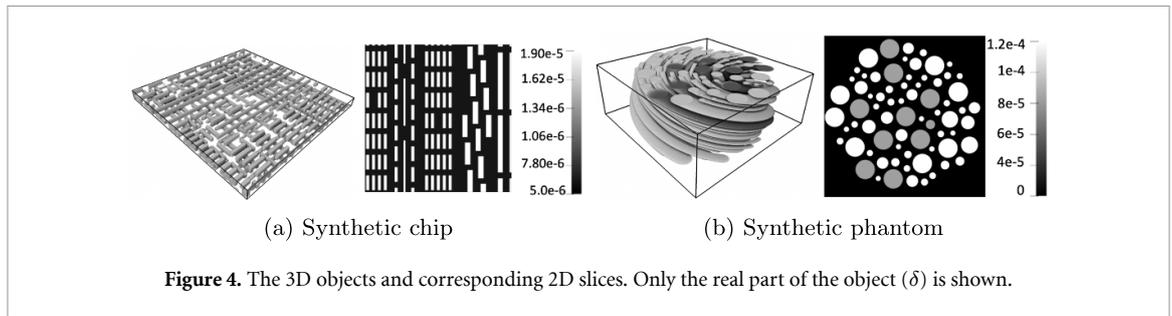


Figure 4. The 3D objects and corresponding 2D slices. Only the real part of the object (δ) is shown.

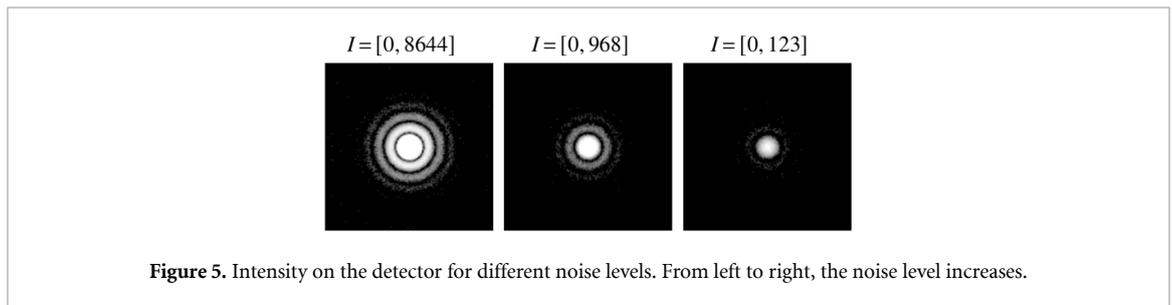


Figure 5. Intensity on the detector for different noise levels. From left to right, the noise level increases.

set to 8 pixels that approximately correspond to 50% overlap. Then, the object is rotated $3N/2$ times at regular intervals from 0 to π , satisfying the Nyquist criterion. We refer to this case as *well-sampled data*. Next, we report reconstructions using *under-sampled data* where we further decrease the probe overlap to 25% and rotate the object $3N/8$ times at regular intervals from 0 to π . It is essential to point out that solving the ptycho-tomography problem jointly enables relaxing high probe overlap restriction where the conventional methods would fail to reconstruct good quality reconstructions. A more detailed study on conventional methods vs jointly solving the ptycho-tomography problem via ADMM can be found in [2].

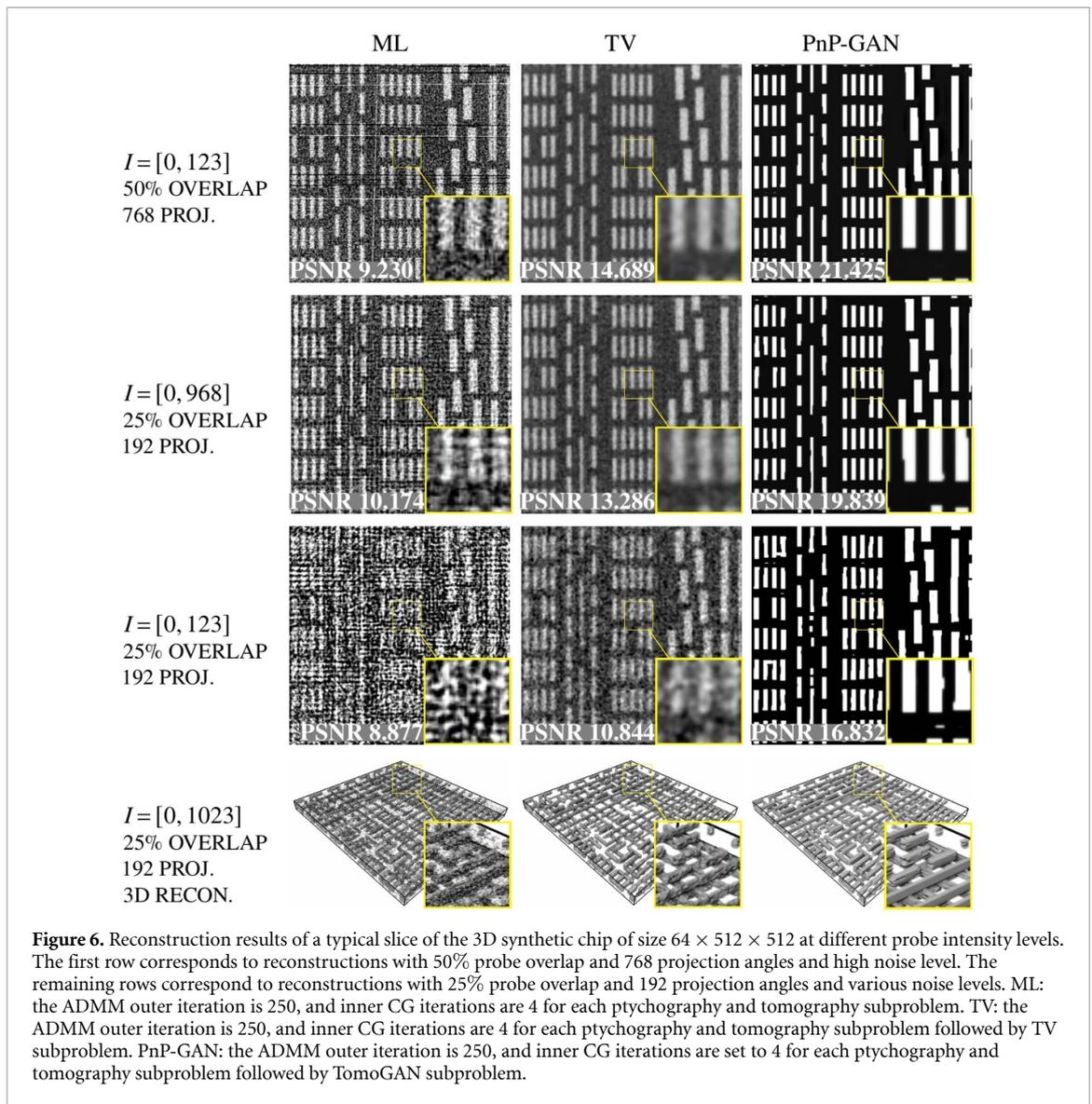
An additional 3D phantom with a shape of $180 \times 512 \times 512$ and a voxel size 5 nm is generated via XDesign [9] to demonstrate the effectiveness of the proposed method. The object consists of two different materials: gold (Au), mercury (Hg) with densities of 19.32 and 13.53 g cm^{-3} , respectively. The 3D object and its 2D slice are given in figure 4(b). While we use 5 keV beam energy to simulate the refractive index values for ptychographic data, the probe and the detector settings are the same as in the first experiment. The detector photon counts are given as $I = [0, 649]$, and $I = [0, 190]$ on average.

For acceleration of the ADMM, we use 4 inner CG iterations for ptychography and tomography problems, and the ADMM outer iteration limit is set to 250. Early termination is a common practice to accelerate the ADMM solution; see the review in [5] and more detailed analysis for our application in [2]. Further accelerations can be possible by varying the penalty parameters ρ and τ dynamically during the ADMM iterations [5, equation (3.13)].

5.2. Simulation results

In this section, we demonstrate the effect of learned priors for the joint ptycho-tomography problem via two 3D simulated objects, see figure 4. To quantify image quality degradation, we use the peak signal-to-noise ratio (PSNR).

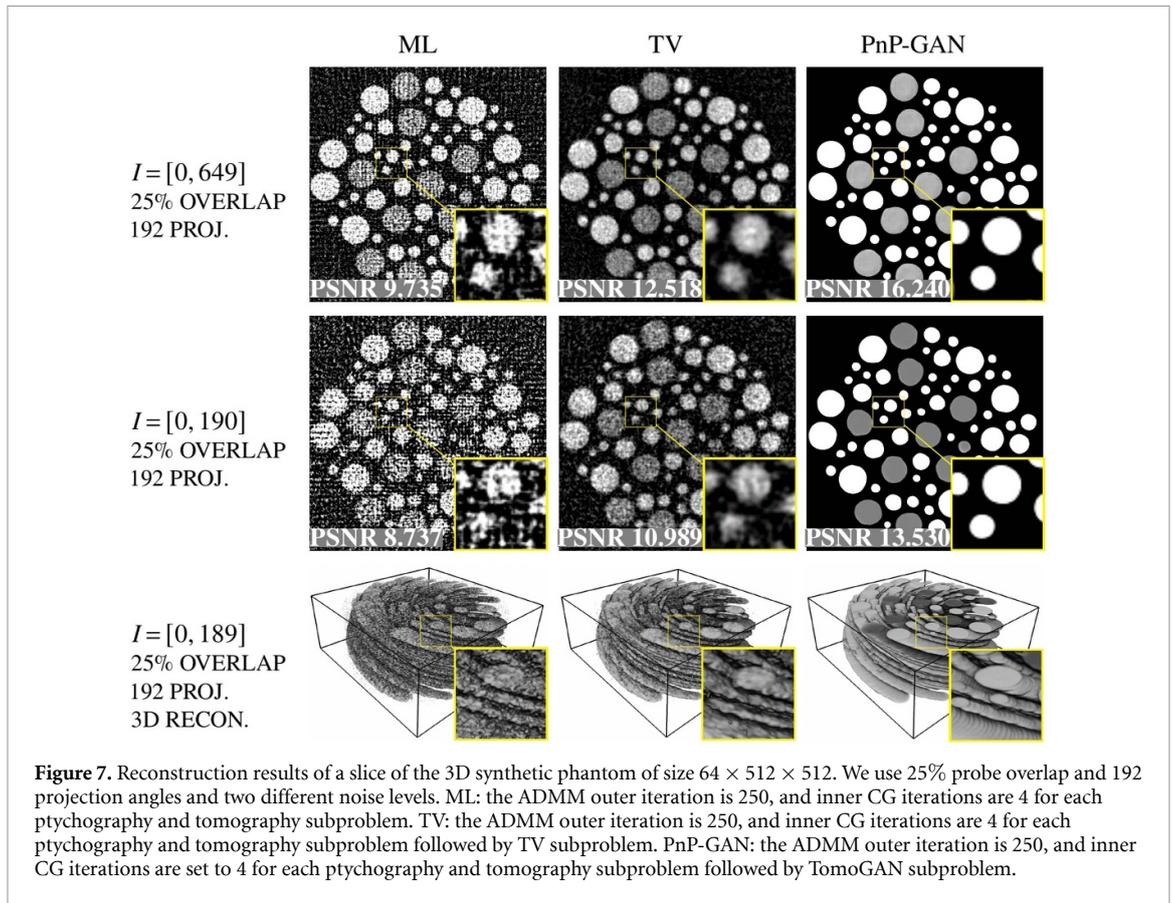
In figure 6, we report reconstruction results for the real part of the object, δ , using three different reconstruction results: (1) the ML estimate (i.e. no regularizer), (2) the MAP estimate with TV prior (i.e. equation (8) is replaced with a TV prior, see [36]), (3) the proposed method denoted by PnP-GAN. In Rows 1–3, we demonstrate 2D slices of the 3D simulated chip to give the details of the image at different noise levels, and in the last row, we show the 3D reconstruction for the high-noise simulation. While the first row of figure 6 corresponds to the well-sampled data at high-noise level, the remaining rows correspond to the under-sampled data at two different noise levels. While most of the features are recovered with well-sampled data using a sparse prior such as TV, the reconstructions are blurred. We observe that PnP-GAN not only removes the artifacts generated by ML, but also denoises images without the blurring effect. Next, we report reconstructions with under-sampled data in figure 6 to highlight the effect of the proposed method, Rows 2–4. While the features are sharper at $I = [0, 968]$, the loss of quality is clear as the noise level increases at $I = [0, 123]$, see Rows 2–4. Without prior knowledge, reconstructions suffer from high noise levels as confirmed by the low PSNR values in the ML reconstructions. While using TV improves the reconstruction



quality compared to the ML reconstructions, the blurring effect is still visible in all reconstructions. On the other hand, PnP-GAN improves reconstruction quality with the help of iterative denoising and generates sharp images with significantly higher PSNR values.

In this paper, our main focus is to generate good-quality reconstructions under limited and noisy measurement data. Therefore, in the next experiment, we only demonstrate reconstructions with under-sampled data where the probe overlap is 25% and the object is rotated $3N/8$ times at regular intervals from 0 to π . Reconstruction results for δ are reported in figure 7 using ML, MAP (with TV), and the proposed PnP-GAN methods. The first two rows of figure 7 demonstrate 2D slices of the 3D simulated chip to provide the details of the image at low- and high-noise levels, and in the bottom row, we show the 3D reconstruction for the high-noise simulation. The imaging artifacts in ML when there is no prior information or regularization is due to the combination of under-sampling artifacts in tomography (also known as the streaking artifacts) and measurement noise in diffracted measurements. TV regularization partly compensates for the high-frequency artifacts in images but a residual artifact pattern is still visible unless the regularization parameter is selected to be too high. PnP-GAN can successfully recover a good-quality image as demonstrated visually and through a high PSNR.

Simulations show that the proposed method can decrease the total number of projections by 75% based on Nyquist sampling with significantly fewer overlapped regions while generating good quality reconstructions. Although small artifacts are introduced in the reconstructions with under-sampled measurements at high noise levels, the proposed method still gives the highest PSNR value. The reconstructions can be further improved by extending the training data or using different



deep-learning-based denoisers. Our goal is not to favor a single deep-learning-based denoiser, but to introduce a generic framework that integrates such learned priors into the ADMM framework to remove the unique type of noise in ptycho-tomography problem.

To give some perspective for the computational performance of the proposed method, consider the second numerical experiment with the object size of $180 \times 512 \times 512$ and detector size of 128×128 . We implemented the main solvers using CUDA and accelerated their computations with NVIDIA RTX 2080 GPUs. The total time for recovering the image is around 10 h when we set the outer ADMM iterations to 250, inner CG iterations to 4 for ptychography and tomography subproblems and an additional denoiser step at each ADMM iteration.

In the remaining of this section, we want to show the advantage of using PnP-GAN as opposed to ML-GAN where TomoGAN is used as a postprocess denoiser. In ML-GAN, we first solve the joint ptycho-tomography problem using the ADMM method as in [2]. Then, TomoGAN is applied to the resulting reconstruction as a postprocess denoiser. To be consistent with the experiments in this paper, we also set the ADMM outer iteration to 250, and inner CG iterations to 4 for each ptychography and tomography subproblem. On the other hand, PnP-GAN splits the joint problem into three parts: ptychography, tomography and denoiser where learned priors are used for iterative denoising at each iteration of the ADMM method. The reconstruction results are shown in the figure 8. While ML-GAN can generate decent reconstructions with well-sampled data, the reconstruction quality highly depends on the noisy input of the image. Therefore, the degradation in image quality is severe when using under-sampled and noisy measurement data. On the other hand, PnP-GAN iteratively denoises the input image, and improves reconstruction quality even at high noise levels. This is potentially due to the nature of the iterative optimization process, where earlier iterations are less noisy and a tunable GAN model is effectively enhancing the image without generating artificial artifacts. In fact, the effect of learned priors is more drastic at high noise levels because the small features in the ML estimate are hardly separable from the background.

5.3. Effect of the denoiser parameter, α^k

In this section, we present an empirical study on the effect of the denoiser parameter, α^k , based on reconstruction quality and residual decay using six representative schemes. The goal of this section is not to provide an optimal denoiser parameter, but to share valuable observations to decide on an effective one.

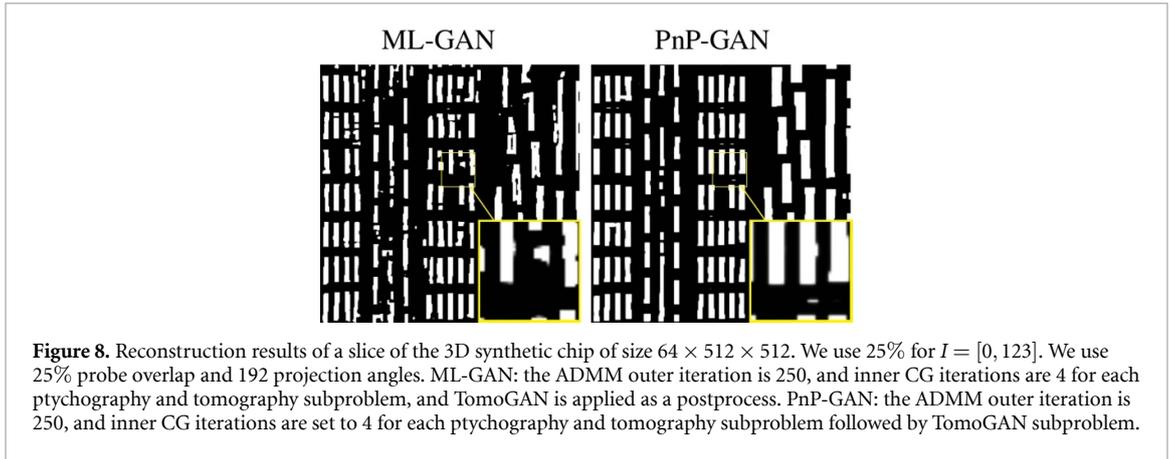


Figure 8. Reconstruction results of a slice of the 3D synthetic chip of size $64 \times 512 \times 512$. We use 25% for $I = [0, 123]$. We use 25% probe overlap and 192 projection angles. ML-GAN: the ADMM outer iteration is 250, and inner CG iterations are 4 for each ptychography and tomography subproblem, and TomoGAN is applied as a postprocess. PnP-GAN: the ADMM outer iteration is 250, and inner CG iterations are set to 4 for each ptychography and tomography subproblem followed by TomoGAN subproblem.



Figure 9. The reconstruction results for the corresponding α^k values.

In figure 9, we give the α -schedules and reconstructions with corresponding α^k values. To demonstrate reconstruction quality, we provide 2D slice of the simulated chip for each denoiser parameter and report the PSNR value on each image. In some cases, we observe that MSE loss in TomoGAN causes some peak amplitude information to be lost since it tries to fit the average. However, this loss does not affect the image quality notably as it is confirmed with relatively high PSNR values.

To highlight the effect of the denoiser parameter on convergence, we also monitor the optimality conditions for the ADMM problem, which are the primal and dual feasibility. For our problem, the primal residuals for the two constraints at iteration $k + 1$ are defined as follows:

$$r_1^{k+1} = \mathcal{H}x^{k+1} - \psi^{k+1} \quad \text{and} \quad r_2^{k+1} = x^{k+1} - \eta^{k+1}, \quad (22)$$

which we call the first and second primal residuals, respectively. In addition, we define the residual for dual feasibility at iteration $k + 1$ as follows:

$$s^{k+1} = \mathcal{H}x^{k+1} - \mathcal{H}x^k; \quad (23)$$

see [2, section (2.3)]. In figure 10, we show the residual decays for each α_k values.

To summarize, we conclude that we obtain poor reconstructions in the early ADMM iterations for the joint ptycho-tomography problem using the general PnP denoising operator in equation (20), α_1 . Hence, reducing the denoiser effect is essential in the first few tens of outer ADMM iterations. Furthermore, we observe that solving ψ and x subproblems higher number of inner iterations does not improve the reconstruction quality in the early ADMM iterations and denoiser parameter is still needed. Next, we implement the denoising operator only incrementally and maximize the denoiser effect as a postprocessing step in the final iteration. This selection not only gives one of the highest PSNR values but also gives the fastest convergence behavior, as can be confirmed in figure 10. While α_5 also generates good-quality reconstructions, we observe that the oscillation in α_k values leads to oscillation in the residual decays. Our

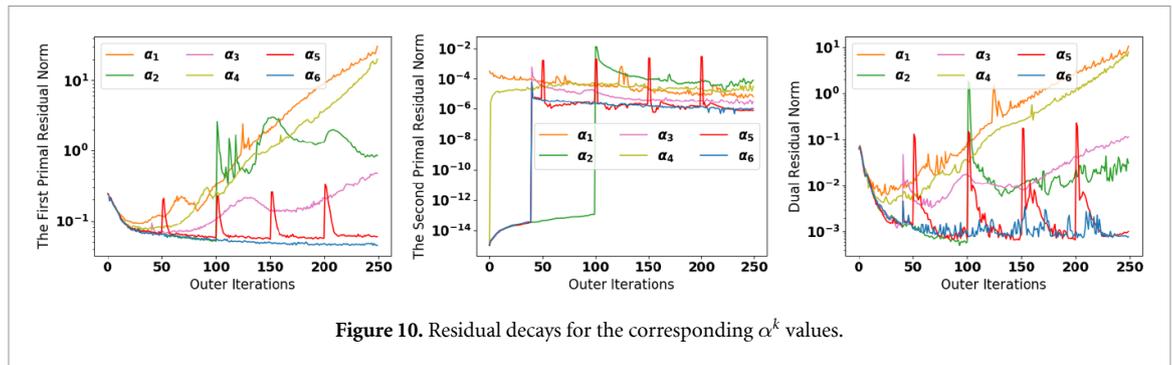


Figure 10. Residual decays for the corresponding α^k values.

observations show that an effective denoiser parameter satisfies the convergence criteria and produces good-quality reconstructions. We obtain reconstructions with high PSNR values in both experiments using the same denoiser parameter, α_6 . While denoiser parameter requires tuning, the observations presented in this paper are applicable to other applications as well.

6. Discussion and conclusions

In this paper, we derive a generic reconstruction framework for solving the joint ptycho-tomography problem with learned priors. The framework splits the joint problem into three parts: ptychography, tomography, and a learned denoiser. The PnP framework is proposed as a flexible way to add state-of-the-art priors to the ADMM. For the joint ptycho-tomography problem, however, these denoisers are not effective because of the different noise characteristics in reconstructions. To this end, we adopted a Poisson process to accurately model our measurements, and further improve reconstruction quality with deep generative models as priors.

A popular way to speed up the ADMM method is through early termination of the subproblems. In our previous work [2], we observed that by solving only a few iterations of ptychography and tomography subproblems, we obtained good-quality reconstructions. In this work, we showed that the general PnP framework leads to poor denoising visible as big white blocks in the reconstructions; see α_1 and α_4 in figure 9. In our simulations, we discuss the importance of the denoiser parameter and introduce an empirical way to control the denoising process. Even though an optimal selection rule is challenging because of the nonconvex nature of the problem, this empirical strategy allows to obtain good results and maintains a robust inversion.

Another way to improve the time-to-solution performance of the ADMM method is to use high-performance many-core architectures, such as GPUs. Depending on the algorithm used, the solution of each subproblems defined in our framework can require significant computational throughput [3, 36]. In our work, we implemented the main solvers using CUDA and accelerated their computations with NVIDIA RTX 2080 GPUs. Similarly, we implemented TomoGAN in TensorFlow, which can be ported to and executed on variety of GPUs, for efficient training and inference operations. While our code is not yet optimized or parallelized, the approach is scalable [35, 36] and there are available software frameworks that we can translate our approach to significantly improve runtimes [55]. We plan to further improve the computational performance of our solvers and intermediate steps using the methods introduced in our previous works [4, 22] and provide a comprehensive evaluation in a future work.

In this work, we focused on under-sampled and highly noisy measurements where we reduced the probe overlap to 25% and projection angles to $3N/8$. Hence, we decrease the total number of projections by 75% compared with well-sampled data. Our simulations show that we can successfully resolve features at high noise level. While a serial approach of using TomoGAN for denoising after reconstruction improves reconstruction quality at lower noise levels, the degradation in image quality is substantial at high noise. In addition, our proposed framework generates a reconstructed object with minimal loss in the quality.

We demonstrated the effectiveness of the framework using synthetic 3D images from under-sampled and noisy measurement data. It should be noted that ptycho-tomography is a relatively new 3D coherent imaging technique and instruments and thus collecting experimental data is not always available for validation studies. For example, at the Advanced Photon Source today, there is no dedicated beamline to ptycho-tomography for general users. While this situation will change soon with the upcoming upgrades of the diffraction limited storage rings at multiple light source facilities worldwide, there are other common challenges in order to work with experimental data that we have not considered in writing this manuscript. For example, because the spatial resolution is on the order of nanometers, it is challenging to precisely know

the position of the measurement geometry. This is a research field by itself, and there are numerous numerical methods to estimate those positions of the data points (also known as the geometrical data alignment or probe position correction; e.g. see [19, 31] and references therein), however those are additional inverse problems and add complexity to the understanding of the learned image priors in the context of ptycho-tomography inverse problem. Naively, we think that those geometrical estimation problems can be solved by introducing an additional set of new auxiliary variables in equation (4). by solving those inverse problems as part of a larger joint solver, however, we leave this work as a future investigation.

Similar to other supervised learning methods, PnP-GAN technique is only applicable when a training dataset is available. This requires either knowing of the expected structure in the images before data is acquired or collecting a representative high-quality dataset (through oversampling) for training the model. While this may not be applicable to all types of samples or specimens, there are key applications that we think will benefit from this supervised approach. For example, the blueprints of integrated circuits (i.e. the GDS file that layouts the design of the chips) could be used to train the model before the experiment, and then the model can be used as part of our framework during image reconstruction. Another potential application could be brain imaging, where the training can be performed from data collected at an electron microscope.

Even there is training data available, it is always finite and often is not completely representative of the whole set of images that are being reconstructed. Because these types of samples have repeating components or structures, our main concern of using GANs as priors is the bias to learned patterns, because sometimes GAN can create imaginary structures (i.e. image artifacts) from arbitrary noise patterns [16]. To evaluate this effect, we performed numerical tests to evaluate this bias. We removed reasonably small wires (rectangles in figure 4(a)) to break the possible learned pattern of the arrangement of wires and check if the network would add those wires back during reconstruction to preserve their arrangement. To our surprise, we have not observed such imaginary artifacts as we may expect at high noise levels. We think this behavior could be explained by the joint (and iterative) solution strategy. Earlier iterations in tomographic reconstruction create blurry but less noisy reconstructions, therefore, the model is not affected by measurement noise as it would be affected in a single image denoising application. We think this may be one of the main advantages of our approach and we believe using GANs as part of an iterative optimization technique may be applicable to other types of imaging modalities as well. As a side note, we also observed through numerous numerical tests that scaling up or down image features does not affect from a potential bias to size of the structures. This is more understandable because we use data augmentation through rotation and scaling for enriching the training data. Ultimately, we conclude that this approach is potentially applicable and can provide improved results if the samples are reasonably sparse such that a well-represented training data can be generated.

Data availability statement

The data generated and/or analysed during the current study are not publicly available for legal/ethical reasons but are available from the corresponding author on reasonable request.

Acknowledgments

Argonne National Laboratory's contribution is based upon work supported by Laboratory Directed Research and Development (LDRD) funding from Argonne National Laboratory, provided by the Director, Office of Science, of the U.S. Department of Energy under Contract No. DE-AC02-06CH11357. Z Liu was supported by the Advanced Scientific Computing Research program of U.S. Department of Energy.

ORCID iDs

Selin Aslan  <https://orcid.org/0000-0002-1077-948X>

Viktor Nikitin  <https://orcid.org/0000-0001-9999-169X>

Sven Leyffer  <https://orcid.org/0000-0002-1077-948X>

Doğa Gürsoy  <https://orcid.org/0000-0002-0080-8101>

References

- [1] Abadi M *et al* 2016 TensorFlow: a system for large-scale machine learning *12th Conf. Operating Systems Design and Implementation, OSDI'16 (Berkeley, CA)* (USENIX Association) pp 265–83
- [2] Aslan S, Nikitin V, Ching D J, Bicer T, Leyffer S and Gürsoy D 2019 Joint ptycho-tomography reconstruction through alternating direction method of multipliers *Opt. Express* **27** 9128–43
- [3] Bicer T, Gürsoy D, Andrade V D, Kettimuthu R, Scullin W, Carlo F D and Foster I T 2017 Trace: a high-throughput tomographic reconstruction engine for large-scale datasets *Adv. Struct. Chem. Imaging* **3** 6

- [4] Bicer T, Gürsoy D, Kettimuthu R, De Carlo F, Agrawal G and Foster I T 2015 Rapid tomographic image reconstruction via large-scale parallelization *Euro-Par 2015: Parallel Processing* ed J L Träff, S Hunold and F Versaci (Berlin: Springer) pp 289–302
- [5] Boyd S, Parikh N, Chu E, Peleato B and Eckstein J 2011 Distributed optimization and statistical learning via the alternating direction method of multipliers *Found. Trends Mach. Learn.* **3** 1–122
- [6] Buzzard G T, Chan S H, Sreehari S and Bouman C A 2018 Plug-and-play unplugged: optimization-free reconstruction using consensus equilibrium *SIAM J. Imaging Sci.* **11** 2001–20
- [7] Chan S H, Wang X and Elgendy O A 2017 Plug-and-play ADMM for image restoration: fixed-point convergence and applications *IEEE Trans. Comput. Imaging* **3** 84–98
- [8] Chang J H R, Li C-L, Poczós B, Kumar B V K V and Sankaranarayanan A C 2017 One network to solve them all—solving linear inverse problems using deep projection models *IEEE Int. Conf. on Computer Vision (ICCV)* pp 5888–97
- [9] Ching D J and Gürsoy D 2017 XDesign: an open-source software package for designing x-ray imaging phantoms and experiments *J. Synchrotron Radiat.* **24** 537–44
- [10] Dabov K, Foi A, Katkovnik V and Egiazarian K 2007 Image denoising by sparse 3-D transform-domain collaborative filtering *IEEE Trans. Image Process.* **16** 2080–95
- [11] Dai Y H and Yuan Y 1999 A nonlinear conjugate gradient method with a strong global convergence property *SIAM J. Optim.* **10** 177–82
- [12] Deng J, Dong W, Socher R, Li L-J, Li K and Fei-Fei L 2009 Imagenet: a large-scale hierarchical image database *2009 Conf. Computer Vision and Pattern Recognition (IEEE)* pp 248–55
- [13] Dierolf M, Menzel A, Thibault P, Schneider P, Kewish C M, Wepf R, Bunk O and Pfeiffer F 2010 Ptychographic x-ray computed tomography at the nanoscale *Nature* **467** 436–9
- [14] Du M, Nashed Y S G, Kandel S, Gürsoy D and Jacobsen C 2020 Three dimensions, two microscopes, one code: automatic differentiation for x-ray nanotomography beyond the depth of focus limit *Sci. Adv.* **6** 1282–9
- [15] Elad M and Aharon M 2006 Image denoising via sparse and redundant representations over learned dictionaries *IEEE Trans. Image Process.* **15** 3736–45
- [16] Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A and Bengio Y 2014 Generative adversarial networks *Advances in Neural Information Processing Systems* pp 2672–80
- [17] Gu S, Zhang L, Zuo W, Zuo W and Feng X 2014 Weighted nuclear norm minimization with application to image denoising *Conf. Computer Vision and Pattern Recognition* pp 2862–9
- [18] Gürsoy D 2017 Direct coupling of tomography and ptychography *Opt. Lett.* **42** 3169–72
- [19] Gürsoy D et al 2017 Rapid alignment of nanotomography data using joint iterative reconstruction and reprojection *Sci. Rep.* **7** 11818
- [20] He J, Yang Y, Wang Y, Zeng D, Bian Z, Zhang H, Sun J, Xu Z and Ma J 2019 Optimizing a parameterized plug-and-play ADMM for iterative low-dose CT reconstruction *IEEE Trans. Med. Imaging* **38** 371–81
- [21] Helgason S 1999 *The Radon Transform* vol 2 (Berlin: Springer)
- [22] Hidayetoğlu M, Biçer T, De Gonzalo S G, Ren B, Gürsoy D, Kettimuthu R, Foster I T and Hwu W-m W 2019 Memxct: memory-centric x-ray CT reconstruction with massive parallelization *Proc. Int. Conf. for High Performance Computing, Networking, Storage and Analysis* pp 1–56
- [23] Hoppe W 1969 Beugung im inhomogenen Primärstrahlwellenfeld, I: Prinzip einer Phasenmessung *Acta Crystallogr. A* **25** 495–501
- [24] Hunger R 2007 An introduction to complex differentials and complex differentiability *Technical Report TUM-LNS-TR-07-06* (Munich University of Technology, Inst. for Circuit Theory and Signal Processing)
- [25] Jin K H, McCann M T, Froustey E and Unser M 2017 Deep convolutional neural network for inverse problems in imaging *IEEE Trans. Image Process.* **26** 4509–22
- [26] Kahnt M et al 2019 Coupled ptychography and tomography algorithm improves reconstruction of experimental data *Optica* **6** 1282–9
- [27] Kamilov U S, Mansour H and Wohlberg B 2017 A plug-and-play priors approach for solving nonlinear imaging inverse problems *IEEE Signal Process. Lett.* **24** 1872–6
- [28] Kingma D P and Ba J 2014 Adam: a method for stochastic optimization *CoRR* (arXiv:1412.6980)
- [29] Li L, Wang X and Wang G 2015 Alternating direction method of multipliers for separable convex optimization of real functions in complex variables *Math. Problems Eng.* **104531** 2015
- [30] Liu Z, Bicer T, Kettimuthu R, Gürsoy D, De Carlo F and Foster I 2020 Tomogan: low-dose synchrotron x-ray tomography with generative adversarial networks *J. Opt. Soc. Am. A* **37** 422–34
- [31] Maiden A, Humphry M, Sarahan M, Kraus B and Rodenburg J 2012 An annealing algorithm to correct positioning errors in ptychography *Ultramicroscopy* **120** 64–72
- [32] Mataev G, Milanfar P and Elad M 2019 Deepred: deep image prior powered by red *Proc. IEEE Int. Conf. Comp. Vis. Workshops (ICCVW)*
- [33] Meinhardt T, Moeller M, Hazirbas C and Cremers D 2017 Learning proximal operators: using denoising networks for regularizing inverse imaging problems *IEEE Int. Conf. on Computer Vision (ICCV)* pp 1781–90
- [34] Metzler C A, Schniter P, Veeraraghavan A and Baraniuk R G 2018 prDeep: robust phase retrieval with a flexible deep network *Proc. 35th Int. Conf. Machine Learning (ICML) (Stockholm, Sweden)*
- [35] Nikitin V, Andrade V D, Slyamov A, Gould B J, Zhang Y, Sampathkumar V, Kasthuri N, Gürsoy D and Carlo F D 2021 Distributed optimization for nonrigid nano-tomography *IEEE Trans. Comput. Imaging* **7** 272–87
- [36] Nikitin V, Aslan S, Yao Y, Bicer T, Leyffer S, Mokso R and Gürsoy D 2019 Photon-limited ptychography of 3D objects via Bayesian reconstruction *OSA Continuum* **2** 2948–68
- [37] Nocedal J and Wright S J 2006 *Numerical Optimization* 2nd edn (New York: Springer)
- [38] Ramos T, Grønager B E, Andersen M S and Andreasen J W 2019 Direct three-dimensional tomographic reconstruction and phase retrieval of far-field coherent diffraction patterns *Phys. Rev. A* **99** 023801
- [39] Romano Y, Elad M and Milanfar P 2017 The little engine that could: regularization by denoising (RED) *SIAM J. Imaging Sci.* **10** 1804–44
- [40] Rond A, Giryes R and Elad M 2016 Poisson inverse problems by the plug-and-play scheme *J. Vis. Commun. Image Represent.* **41** 96–108
- [41] Rudin L, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica D* **60** 259–68
- [42] Ryu E K, Liu J, Wang S, Chen X, Wang Z and Yin W 2019 Plug-and-play methods provably converge with properly trained denoisers *Proc. 36th Int. Conf. Machine Learning (ICML)* pp 5546–57

- [43] Shan H, Zhang Y, Yang Q, Kruger U, Kalra M K, Sun L, Cong W and Wang G 2018 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network *IEEE Trans. Med. Imaging* **37** 1522–34
- [44] Simonyan K and Zisserman A 2014 Very deep convolutional networks for large-scale image recognition *CoRR* (arXiv:1409.1556)
- [45] Sreehari S, Venkatakrishnan S V, Wohlberg B, Buzzard G T, Drummy L F, Simmons J P and Bouman C A 2016 Plug-and-play priors for bright field electron tomography and sparse interpolation *IEEE Trans. Comput. Imaging* **2** 408–23
- [46] Sun Y, Wohlberg B and Kamilov U 2019 An online plug- and-play algorithm for regularized image reconstruction *IEEE Trans. Comput. Imaging* **5** 395–408
- [47] Sun Y, Xu S, Li Y, Tian L, Wohlberg B and Kamilov U S 2019 Regularized fourier ptychography using an online plug-and-play algorithm *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing* pp 7665–9
- [48] Tikhonov A N and Arsenin V Y 1977 *Solutions of Ill-Posed Problems* (New York: Wiley)
- [49] Tropp J A and Wright S J 2010 Computational methods for sparse solution of linear inverse problems *Proc. IEEE* **98** 948–58
- [50] Venkatakrishnan S V, Bouman C A and Wohlberg B 2013 Plug-and-play priors for model based reconstruction *2013 IEEE Conf. Signal and Information Processing* pp 945–8
- [51] Wei K, Aviles-Rivero A, Liang J, Fu Y, Carola-Bibiane S and Huang H 2020 Tuning-free plug-and-play proximal algorithm for inverse imaging problems (arXiv:2002.09611)
- [52] Xu X, Liu J, Sun Y, Wohlberg B and Kamilov U S 2020 Boosting the performance of plug-and-play priors via denoiser scaling (arXiv:2002.11546)
- [53] Yang X, De Andrade V, Scullin W, Dyer E L, Kasthuri N, De Carlo F and Gürsoy D 2018 Low-dose x-ray tomography through a deep convolutional neural network *Sci. Rep.* **8** 1–13
- [54] Ye D H, Srivastava S, Thibault J, Sauer K and Bouman C 2018 Deep residual learning for model-based iterative CT reconstruction using plug-and-play framework *2018 IEEE Int. Conf. Acoustics* 6668–72
- [55] Yu X, Nikitin V, Ching D J, Aslan S, Gursoy D 2021 Scalable and accurate multi-gpu based image reconstruction of large-scale ptychography data (arXiv:2106.07575)
- [56] Zhang K, Zuo W, Chen Y, Meng D and Zhang L 2017 Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising *IEEE Trans. Image Process.* **26** 3142–55