

# **An Update Regarding the PhD Dissertation**

by

Zongyu Li

Doctor of Philosophy  
(Electrical and Computer Engineering)  
in the University of Michigan  
2024

Note: This is a revised edition of the doctoral dissertation “Solving Poisson Inverse Problems in Phase Retrieval and Single Photon Emission Computerized Tomography” available at University of Michigan library <https://dx.doi.org/10.7302/23047>. Major changes include additional experiment results, correction of typos in Chapter 3.2 and Chapter 4.4. To enhance readability, the original thesis layout has been retained while many specifics have been excluded.

Zongyu Li

[zonyul@umich.edu](mailto:zonyul@umich.edu)

ORCID iD: [0000-0003-1813-1722](https://orcid.org/0000-0003-1813-1722)

© Zongyu Li 2024

# TABLE OF CONTENTS

<b>List of Figures</b> . . . . .	<b>iv</b>
<b>List of Tables</b> . . . . .	<b>vi</b>
<b>List of Appendices</b> . . . . .	<b>vii</b>
<b>List of Acronyms</b> . . . . .	<b>viii</b>
<b>List of Symbols</b> . . . . .	<b>xii</b>
<b>Abstract</b> . . . . .	<b>xiii</b>
<b>Chapter</b>	
<b>1 Introduction</b> . . . . .	<b>1</b>
1.1 Contributions . . . . .	1
1.2 Outline . . . . .	2
<b>2 Background</b> . . . . .	<b>4</b>
2.1 Inverse Problems . . . . .	4
2.2 Phase Retrieval . . . . .	4
2.3 Quantitative SPECT Imaging . . . . .	4
<b>3 Poisson Inverse Problems in Phase Retrieval</b> . . . . .	<b>5</b>
3.1 Poisson Phase Retrieval in Low-count Regimes . . . . .	5
3.2 Poisson-Gaussian Phase Retrieval with Score-based Image Prior . . . . .	5
3.2.1 Motivation . . . . .	5
3.2.2 Methods . . . . .	7
3.2.3 Experiment . . . . .	12
3.2.4 Discussion . . . . .	18
3.2.5 Conclusion . . . . .	22
<b>4 Poisson Inverse Problems in SPECT Imaging</b> . . . . .	<b>23</b>
4.1 Training End-to-End Unrolled Iterative Neural Networks for SPECT Image Reconstruction . . . . .	23
4.2 DblurDoseNet: A Deep Neural Network for SPECT Dosimetry Estimation and Resolution Recovery . . . . .	23
4.3 Efficient Super Resolution Network (ESR-Net) for SPECT Image Reconstruction . . . . .	23

4.4	Shorter SPECT Scans Using Self-supervised Coordinate Learning to Synthesize Skipped Projection Views . . . . .	23
4.4.1	Motivation . . . . .	23
4.4.2	Materials and Methods . . . . .	25
4.4.3	Results . . . . .	28
4.4.4	Discussion . . . . .	31
4.4.5	Conclusion . . . . .	35
<b>5</b>	<b>Discussion and Future Works . . . . .</b>	<b>36</b>
5.1	Learning on “SmArge” data . . . . .	36
5.2	Generative AI . . . . .	36
5.3	Denoising Projections with Unsupervised Learning . . . . .	36
<b>6</b>	<b>Conclusion . . . . .</b>	<b>37</b>
	<b>Appendices . . . . .</b>	<b>39</b>
	<b>Bibliography . . . . .</b>	<b>42</b>



## LIST OF FIGURES

3.1	Illustration of Poisson and Gaussian noise statistics in holographic phase retrieval. . . . .	6
3.2	Reconstructed images by unregularized methods (Gaussian, Gaussian-Amplitude, Poisson and Poisson-Gaussian) on Histopathology dataset [2], celebA dataset [65] and CT-density dataset. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image. We used $\alpha = 0.035$ and $\sigma = 1$ . . . . .	14
3.3	Reconstructed images on dataset [2]. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image. $\alpha$ and $\sigma$ were set to 0.02 and 1, respectively. . . . .	15
3.4	Reconstructed images on celebA dataset [65]. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image. $\alpha$ and $\sigma$ were set to 0.035 and 1, respectively. . . . .	15
3.5	Reconstructed images on CT-density dataset. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image. $\alpha$ and $\sigma$ were set to 0.035 and 1, respectively. . . . .	16
3.6	Reconstructed images by Gaussian, Poisson and Poisson-Gaussian log-likelihood model with AWFS image prior. Tested on Histopathology dataset [2], celebA dataset [65] and CT-density dataset. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image. $\alpha$ and $\sigma$ were set to 0.025 and 1, respectively. . . . .	17
3.7	Comparison of SSIM and NRMSE varying scaling factor $\alpha \in [0.02, 0.035]$ and STD of Gaussian noise $\sigma \in [0.25, 1.5]$ defined in (3.2). . . . .	19
3.8	Reconstructed images by DOLPH [87] and our proposed AWFS method under different $\sigma$ values. Scaling factor $\alpha$ was set to 0.02 (defined in (3.2)). . . . .	20
3.9	Comparing AWFS vs. WFS with NRMSE vs. number of iterations under different noise levels. The curves and shadows represent the mean and standard deviation, respectively. . . . .	20
3.10	Reconstructed images by the unregularized Poisson method (the second column) as well as with the AWFS method for different scaling factors $\alpha$ (third to fifth columns). The top and bottom rows show reconstructions from different measurement realizations. . . . .	21

4.1	Workflow of the proposed SPECT projection synthesis method. The training process (top) involves inputting 5-dimensional coordinates into the MLP, with a user-defined loss function guiding the network to learn from the patient-specific training targets: measured counts in sparse views. During testing (bottom), the trained network receives the coordinates of missing views and outputs the predicted counts. . . . .	27
4.2	Comparison of measured and synthesized projections for a patient following <sup>177</sup> Lu-PSMA therapy. (a), (b), and (c) show measured projection, linearly interpolated projection, and NERF-synthesized projection, respectively. The images and profile comparison across lacrimal glands show two hot spots/peaks in the NERF synthesized projection (green line) corresponding to left and right lacrimals, closely resembling the profile of the measured projection (red line), whereas the corresponding results for the linear interpolation shows 4 peaks due to distortions. . . . .	29
4.3	Visual comparison of (a) phantom true activity, (b) full recon, and (c) NERF recon, (d) LinInt recon, (e) partial recon for DF=4. All images are in the same color scale. . . . .	30
4.4	AR to noise curves for sphere volumes ranging from 2 to 114 mL for the full recon and across DFs of 2, 4, and 8 (a to c). Distinct markers are consistently used to represent each sphere volume across all subfigures. The comparison illustrates the variations in AR and noise levels across four reconstruction methods: full recon, NERF recon, LinInt recon, and partial recon, for different sphere sizes. . . . .	30
4.5	Coronal MIPs of SPECT reconstructions corresponding to a DOTATATE patient study using four reconstruction methods (columns) and three DFs (rows). Images are displayed with gamma correction with enhanced contrast levels to emphasize the blurring artifacts present in the LinInt recon and the noise present in the partial recon, especially visible at higher DFs. . . . .	32
4.6	Coronal MIPs of SPECT reconstructions corresponding to a PSMA patient study using four reconstruction methods (columns) and three DFs (rows). Images are displayed with gamma correction with enhanced contrast levels to emphasize the blurring artifacts present in the LinInt recon and the noise present in the partial recon, especially visible at higher DFs. . . . .	33

## LIST OF TABLES

3.1	SSIM and NRMSE for Poisson and PG likelihoods. Results were averaged across 7 different noise levels by varying $\alpha \in 0.02 : 0.005 : 0.035$ in (3.2). . . . .	13
3.2	SSIM and NRMSE using Poisson Gaussian likelihood with different regularization/image prior approaches. Results were averaged across 7 different noise levels by varying $\alpha \in 0.02 : 0.005 : 0.035$ in (3.2). WFS* runs the same number of iterations as AWFS whereas WFS <sup>†</sup> runs more iterations until convergence. . . . .	18
4.1	NRMSD (relative to measured projections) comparisons between NERF-synthesized projections and linearly interpolated projections across different DFs for phantom studies and patient studies (average across 11 DOTATATE studies and 6 PSMA studies). . . . .	29
4.2	Comparing synthesized projections using linear interpolation, supervised learning [82] and our proposed method. Results were based on 9 patient scans. . . . .	29
4.3	Average RCNR values of the NERF recon, the LinInt recon, and the partial recon across all eleven DOTATATE patient studies, benchmarked against the full recon, whose RCNR is standardized at 100%. . . . .	31
4.4	Average RCNR values of the NERF recon, the LinInt recon, and the partial recon across all six PSMA patient studies, benchmarked against the full recon, whose RCNR is standardized at 100%. . . . .	32

## LIST OF APPENDICES

<b>A Proof of the Proposed Improved Curvature Formula . . . . .</b>	<b>39</b>
<b>B Uniform Cramér–Rao Lower Bound Analysis for Phase Retrieval Algorithms . . . . .</b>	<b>40</b>
<b>C Proof of the Critical Point Convergence for the “AWFS” Algorithm . . . .</b>	<b>41</b>

## LIST OF ACRONYMS

<b>ADMM</b>	alternating direction method of multipliers
<b>AP</b>	analytical projector
<b>AR</b>	activity recovery
<b>ARNR</b>	activity recovery to noise ratio
<b>AWFS</b>	accelerated Wirtinger flow with score-based image prior
<b>BM3D</b>	block-matching and 3D filtering
<b>CCD</b>	charge-coupled device
<b>CDI</b>	coherent diffractive imaging
<b>CG</b>	conjugate gradient
<b>CNN</b>	convolutional neural network
<b>CNR</b>	contrast-to-noise ratio
<b>CRLB</b>	Cramér-Rao lower bound
<b>CT</b>	computerized tomography
<b>DDPM</b>	denoising diffusion probabilistic model
<b>DF</b>	Downsampling Factor
<b>DFT</b>	discrete Fourier transform
<b>DL</b>	deep learning
<b>DNCNN</b>	Denoising convolutional neural network
<b>DOLPH</b>	diffusion models for phase retrieval
<b>DOTATATE</b>	DOTA-D-Phe (1)-Tyr (3)-octreotate
<b>DPM</b>	dose planning method
<b>DRVH</b>	dose-rate volume histogram

**DVK** dose voxel kernel

**EM** expectation-maximization

**ETM** empirical transmission matrix

**FFT** fast (discrete) Fourier transform

**FISTA** fast iterative shrinkage-thresholding algorithm

**FOV** field of view

**FWHM** full width at half maximum

**GAN** generative adversarial network

**GMM** gaussian mixture models

**GS** Gerchberg Saxton

**IID** independent and identically distributed

**LBFGS** low-memory Broyden–Fletcher–Goldfarb–Shanno

**MAE** mean activity error

**MAP** maximum a posteriori

**MBIR** model-based image reconstruction

**MC** Monte Carlo

**MCRPC** metastatic castration-resistant prostate cancer

**MLE** maximum likelihood estimate

**MLEM** maximum likelihood expectation maximization

**MLP** multi-layer perceptron

**MM** majorize minimize

**MRI** magnetic resonance imaging

**NERF** neural radiance field

**NLM** non-local means

**NRMSD** normalized root mean square difference

**NRMSE** normalized root mean square error

**ODWT** orthogonal discrete wavelet transform

**OSEM** ordered-subset expectation maximization

**PET** positron emission tomography

**PG** poisson-gaussian

**PGM** proximal gradient method

**PNP** plug-and-play

**POGM** proximal optimized gradient method

**PR** phase retrieval

**PSD** positive semi-definite

**PSF** point spread function

**PSMA** prostate-specific membrane antigen

**PSNR** peak signal-to-noise ratio

**PV** partial volume

**RCNR** relative contrast-to-noise ratio

**RED** regularization by denoising

**RED-SD** regularization by denoising with steepest descent

**RELU** Rectified Linear Unit

**RPTS** radiopharmaceutical therapies

**SAM** segment anything model

**SDP** semi-definite programming

**SFBP** SPECT forward-backward projector

**SOTA** state-of-the-art

**SPECT** single photon emission computerized tomography

**SSIM** structural similarity index measure

**STD** standard deviation

**TEW** triple energy windows

**TOF** time-of-flight

**TV** total variation

**UCRLB** uniform Cramér-Rao lower bound

**VAE** variational autoencoder

**VOI** volumes of interest

**VP** virtual patient

**WF** Wirtinger flow

**WFS** Wirtinger flow with score-based image prior

**XCAT** extended cardiac-torso

<sup>68</sup>**GA** Gallium-68

<sup>90</sup>**Y** Yttrium-90

<sup>177</sup>**LU** Lutetium-177



## LIST OF SYMBOLS

$\boldsymbol{x}$  a signal or image

$\hat{\boldsymbol{x}}$  an image estimate

$\boldsymbol{y}$  measurements

$\boldsymbol{A}$  the measurement/system/forward model

$\|\cdot\|_p$   $\ell_p$ -norm

$\nabla\cdot$  the gradient operator

$\nabla^2\cdot$  the Hessian operator

$R(\cdot)$  a regularization function

$\boldsymbol{B}'$  Hermitian transpose of  $\boldsymbol{B}$

$\boldsymbol{B}^T$  transpose of  $\boldsymbol{B}$

$\text{diag}\{\cdot\}$  a diagonal matrix

## ABSTRACT

We live in a world where many objects cannot be imaged directly and hence rely on reconstruction algorithms to solve the corresponding inverse imaging problems. However, lots of information is contaminated or even lost when samples are collected by imaging devices, so that the resulting inverse problem is ill-posed and challenging to solve. As the recorded photon arrivals by the sensor are often assumed to follow Poisson distributions, algorithms for solving Poisson inverse problems are crucial. This thesis tackles two applications where Poisson inverse problems arise: phase retrieval and single photon emission computerized tomography (SPECT).

For phase retrieval, we propose novel optimization algorithms working in low-count regimes, including a novel majorize-minimize (MM) algorithm, a modified Wirtinger flow algorithm using the observed Fisher information for step size and a generative image prior based on score matching. Our proposed algorithms lead to faster convergence rate and improved reconstruction quality evaluated both qualitatively and quantitatively.

For SPECT imaging, we focus on deep learning (DL) solutions including: 1) We propose end-to-end training of unrolled iterative convolutional neural network (CNN) using our memory efficient Julia toolbox for SPECT image reconstruction. 2) We propose a DL algorithm for joint dosimetry estimation and image deblurring for estimating patient's absorbed dose-rate distribution in radionuclide therapy. 3) We propose unsupervised coordinate-based learning for predicting missing SPECT projection views.

# CHAPTER 1

## Introduction

### 1.1 Contributions

The contributions of works in this thesis can be summarized as follows.

1. Phase Retrieval.
  - We develop algorithms to address the challenges of low-count regimes in the Poisson phase retrieval problem. Our proposed approach involves a modified Wirtinger flow ( $w_F$ ) algorithm that uses a step size determined by the observed Fisher information, as well as a novel curvature for majorize-minimize algorithms that incorporates a quadratic majorizer. Through simulated experimentation with different system matrices, we demonstrate the effectiveness and convergence properties of our methods. The simulation results reveal that our approaches not only enable successful recovery in extremely low-count scenarios but also outperform previous methods in terms of speed of convergence. This work is based on published papers [60, 61].
  - In addition, we investigate situations where the measurements are influenced by a mixture of Poisson and Gaussian noise. We introduce a novel method called “AWFS” that employs accelerated  $w_F$  with a score function as a generative prior. We provide theoretical analysis to demonstrate the convergence guarantee of the proposed algorithm at critical points. Results from simulations illustrate that our approach improves reconstruction quality in terms of both visual perception and numerical assessment. This work is based on [39, 59].
2. SPECT Imaging.

- We develop an efficient Julia toolbox for modeling SPECT forward-backward projectors<sup>1</sup>. This toolbox uses multi-threading and in-place operations to enable parallel computing and reduce memory allocations. As a result, our proposed SPECT projector allows for efficient backpropagation during the training of deep learning regularized iterative algorithms in an end-to-end manner. This approach has shown potential for producing higher quality reconstructions compared to methods without end-to-end training. This work is based on published paper [56].
- We propose a deep neural network (DblurDoseNet) for joint dosimetry estimation and image deblurring after SPECT reconstruction that produces accurate dose-rate distribution estimates as well as compensating for SPECT resolution effects. Evaluations both on phantoms and patients demonstrate that the proposed DblurDoseNet can outperform the current gold standard, *i.e.*, Monte Carlo (MC) based methods, and is also fast enough for real-time clinical use in radionuclide therapy dosimetry for treatment planning [58, 54].
- We propose a neural network with unsupervised coordinate-based learning to predict missing SPECT projections before reconstruction. Our method aims to decrease the acquisition time for SPECT by only obtaining a subset (*e.g.*, one fourth) of all projections. Our unsupervised approach achieves improved quality in image reconstruction when compared to linear interpolation methods used for the prediction of absent projection views. This work is based on published abstract [62].

## 1.2 Outline

Chapter 2 provides the necessary background on the mathematical frameworks we use for modelling Poisson inverse problems. It introduces the details of phase retrieval, SPECT image reconstruction and dosimetry. Chapter 3 introduced our proposed algorithms for the phase retrieval problem. Chapter 4 focuses on DL solutions in SPECT imaging, from acquisition, reconstruction, to post processing. Chapter 5 discusses current challenges faced and explores potential avenues for future research. One of the major challenges is addressing the limited amount of training data for large 3D images, prompting the exploration

---

<sup>1</sup><https://github.com/JuliaImageRecon/SPECTrecon.jl>

of techniques like transfer learning and active learning. Additionally, we delve into topics such as generative AI, multi-modality imaging, and optimization approaches from a computational perspective. Chapter 6 concludes this thesis.

The appendix provides more in depth algorithmic detail than is provided in the main chapters as well as other supplementary materials. In particular, Appendix A derives an improved curvature of the quadratic majorizer in the `MM` algorithm for Poisson phase retrieval in Chapter 3. Appendix B derives and analyzes the `UCRLB` for the `WF` algorithm in Chapter 3. Appendix C provided detailed analysis on the critical-point convergence guarantee of the “AWFS” algorithm in Chapter 3.

## CHAPTER 2

# Background

This chapter first introduces inverse problems mathematically and presents a generic framework for addressing such problems. We then present the background for phase retrieval and quantitative SPECT imaging in sufficient detail as two areas where inverse problems arise.

### 2.1 Inverse Problems

### 2.2 Phase Retrieval<sup>1</sup>

### 2.3 Quantitative SPECT Imaging

---

<sup>1</sup>This section is largely taken from [61].

## CHAPTER 3

# Poisson Inverse Problems in Phase Retrieval<sup>1</sup>

## 3.1 Poisson Phase Retrieval in Low-count Regimes<sup>2</sup>

## 3.2 Poisson-Gaussian Phase Retrieval with Score-based Image Prior<sup>3</sup>

### 3.2.1 Motivation

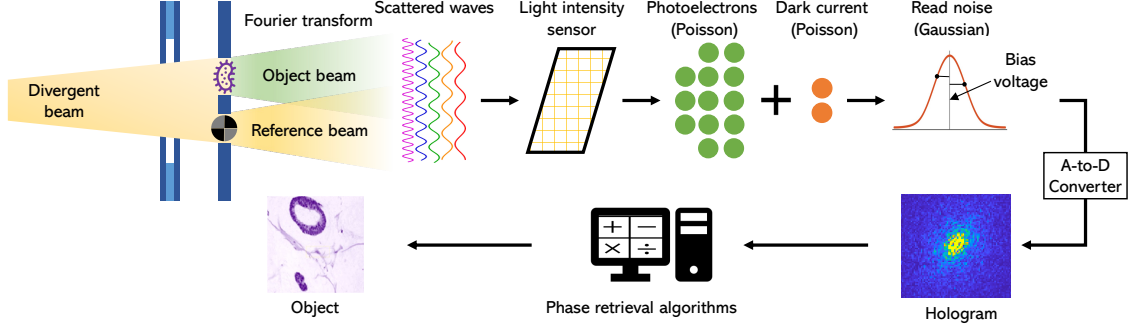
In practical scenarios, the measurements  $\mathbf{y}$  are often contaminated by both PG noise. The Poisson distribution is due to the photon counting and dark current [105]. The Gaussian statistics stem from the readout structures (e.g., analog-to-digital converter (ADC)) of common cameras. Fig. 3.1 illustrates the PG mixed noise statistics in the holographic PR. Because the PG likelihood is complicated, most previous works [11, 10, 86, 43, 9, 90, 77, 9, 33, 74, 104, 32, 42, 63, 113, 96, 12, 110, 98, 35, 111, 4, 100, 49, 57, 34, 61, 30, 16, 6, 116, 14] approximate the Poisson noise statistics by the central limit theorem and work with a substitute Gaussian log-likelihood estimate problem or use the Poisson maximum likelihood model but simply disregard Gaussian readout noise. Other more complicated approximation methods have also been proposed, such as the shifted Poisson model [88], the unbiased inverse transformation of a generalized Anscombe transform [70, 99], and the majorize-minimize algorithm [29]. However, these approximate methods can lead to a suboptimal solution after optimization that results in a lower-quality reconstruction. Apart from the likelihood modeling, the regularizer  $R(\mathbf{x})$  provides prior information about underlying object characteristics that may aid in resolving ill-posed inverse problems. Beyond simple choices of  $R(\mathbf{x})$  such as TV or the L1-norm of coefficients of wavelet transform [23], deep learning

---

<sup>1</sup>This chapter is based on [60, 61, 59].

<sup>2</sup>This section is based on [61].

<sup>3</sup>This section is based on [39, 59].



**FIG 3.1** – Illustration of Poisson and Gaussian noise statistics in holographic phase retrieval.

(DL)-integrated algorithms for solving inverse problems in computational imaging have been reported to be the state-of-the-art [76]. The trained networks can be used as an object prior for regularizing the reconstructed image to remain on a learned manifold [7]. Incorporating a trained denoising network as a regularizer  $R(\cdot)$  led to methods such as plug-and-play (PNP) [13, 117, 46] and regularization by denoising (RED) [80]. In contrast to training a denoiser using clean images, there is growing popularity of self-supervised image denoising approaches that do not require clean data as the training target [51, 5, 107]. In addition to training a denoiser as regularizer, generative model-based priors have also been proposed [3, 109]. Recently, diffusion models have gained significant traction for image generation [91, 37, 28, 93]. These probabilistic image generation models start with a clean image and gradually increase the level of noise added to the image, resulting in white Gaussian noise. Then in the reverse process, a neural network is trained to learn the noise in each step to generate or sample a clean image as in the original data distribution. The score-based diffusion models estimate the gradients of data distribution and can be used as plug-and-play priors for inverse problems [36] such as image deblurring and MRI and CT reconstruction [50, 41, 21, 22, 68, 92]. However, the realm of using score-based models to perform phase retrieval is relatively unexplored; previous relevant works [87, 36] applied DDPM to PR but with less realistic system models and under solely Gaussian or Poisson noise statistics.

In summary, our contribution is three-fold:

- We present a new algorithm known as accelerated WF with a score-based image prior (*i.e.*,  $\nabla R(\mathbf{x})$ ) to address the challenge of holographic PR problem in the presence of PG noise statistics.



- Theoretically, we derive a Lipschitz constant for the holographic PR’s PG log-likelihood and subsequently demonstrate the critical points convergence guarantee of our proposed algorithm.
- Simulation experiments demonstrate that: 1) Algorithms with the PG likelihood model yield superior reconstructions in comparison to those relying solely on either the Poisson or Gaussian likelihood models. 2) With the proposed score-based prior as regularization, the proposed approach generates higher quality reconstructions and is more robust to variation of noise levels without any parameter tuning compared to alternative state-of-the-art methods.

### 3.2.2 Methods

#### 3.2.2.1 Score Function and Diffusion Models

Let  $p_\theta(\mathbf{x})$  denote a model for the prior distribution of the latent image  $\mathbf{x}$ ; the score function is then defined as<sup>4</sup>  $s_\theta(\mathbf{x}) = \nabla_{\mathbf{x}} \log p_\theta(\mathbf{x})$ . Consider a sequence of positive noise scales (for white Gaussian  $\mathcal{N}(0, \sigma_k^2)$ ):  $\sigma_1 > \sigma_2 > \dots > \sigma_K$ , with  $\sigma_K$  being small enough so that noise of this level does not visibly affect the image, and  $\sigma_1$  depending on the application. Score matching can be used to train a noise conditional score network [102, 91] as follows:

$$\hat{\theta} = \arg \min_{\theta} \sum_{k=1}^K \mathbb{E}_{\mathbf{x}, \tilde{\mathbf{x}}} \left[ \left( s_\theta(\mathbf{x}, \sigma_k) - \frac{\mathbf{x} - \tilde{\mathbf{x}}}{\sigma_k^2} \right)^2 \right],$$

where  $\mathbf{x} \sim p(\mathbf{x})$ ,  $\tilde{\mathbf{x}} \sim \mathbf{x} + \mathcal{N}(0, \sigma_k^2 \mathbf{I})$ . (3.1)

With enough data, the neural network  $s_\theta(\mathbf{x}, \sigma)$  is expected to learn the distribution  $p_{\sigma(\mathbf{x})} = \int p(\mathbf{x}) p_\sigma(\mathbf{y}|\mathbf{x}) d\mathbf{x}$  where  $p_\sigma(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{x}, \sigma^2 \mathbf{I})$ . To sample from the prior, the method of Langevin dynamics is frequently used [91]. To leverage diffusion models for solving inverse problems, previous methods generally recast the reconstruction problem as a conditional generation or sampling problem [87, 36, 93, 20, 92, 19]. This involves relying on the capacity of diffusion models to produce high-quality images while complying with data-fidelity constraints. However, in applications where data collection is costly, *i.e.*, with a limited amount of training data, it is often challenging to train a diffusion model that can generate high-quality images even in an unconditional way. Under these conditions, we found that the score function learned during training diffusion models can serve as an effective image prior, which can capture certain data characteristics when trained for the

---

<sup>4</sup>This definition differs from the score function in statistics where the gradient is taken w.r.t.  $\theta$  of  $\log p_\theta(\mathbf{x})$ .

denoising prediction in the reverse process of the diffusion model. Similar to previous works [36] that uses the score function as a PNP prior, here we also incorporate the score function as a regularization in the optimization objective for solving the PR problem. We believe this is a more efficient scheme for incorporating diffusion priors especially for applications with a limited amount of training data, a very common situation in the optical imaging sector.

### 3.2.2.2 Likelihood Modeling and $\mathbf{w_F}$

Based on the physical model as demonstrated in Fig. 3.1, we model the system matrix  $\mathbf{A}$  by the (oversampled and scaled) discrete Fourier transform applied to a concatenation of the sample  $\mathbf{x}$ , a blank image (representing the holographic separation condition [49]) and a known reference image  $\mathcal{R}$ , similar to (??), the measurements  $\mathbf{y}$  follow the Poisson plus Gaussian distribution:

$$\mathbf{y} \sim \mathcal{N}(\text{Poisson}(|\mathbf{A}(\mathbf{x})|^2 + \bar{\mathbf{b}}), \sigma^2 \mathbf{I}), \quad \mathbf{A}(\mathbf{x}) \triangleq \alpha \mathcal{F}\{\mathbf{x}, \mathbf{0}, \mathcal{R}\}. \quad (3.2)$$

Here  $\sigma^2$  denotes the variance of Gaussian noise, and  $\alpha$  denotes a scaling factor (quantum efficiency, conversion gain, etc.) after applying the Fourier transform. So that the negative log-likelihood of (3.2) is

$$g_{\text{PG}}(\mathbf{x}) = \sum_{i=1}^M g_i(\mathbf{x}), \quad g_i(\mathbf{x}) \triangleq -\log \left( \sum_{n=0}^{\infty} \frac{e^{-(|\mathbf{a}'_i \mathbf{x}|^2 + \bar{b}_i)} \cdot (|\mathbf{a}'_i \mathbf{x}|^2 + \bar{b}_i)^n}{n!} \cdot \frac{e^{-\left(\frac{y_i - n}{\sqrt{2}\sigma}\right)^2}}{\sqrt{2\pi}\sigma^2} \right). \quad (3.3)$$

Here  $M$  denotes the length of  $\mathbf{y}$ ,  $\mathbf{a}'_i$  denotes the  $i$ th row of  $\mathbf{A}$  (since  $\mathbf{A}$  is linear).  $\mathbf{w_F}$  can be used for estimating  $\mathbf{x}$ :

$$\nabla g_{\text{PG}}(\mathbf{x}) = 2\mathbf{A}' \text{diag}\{\phi_i(|\mathbf{a}'_i \mathbf{x}|^2 + \bar{b}_i; y_i)\} \mathbf{A} \mathbf{x}, \quad (3.4)$$

$$\phi(u; v) \triangleq 1 - \frac{s(u, v-1)}{s(u, v)}, \quad s(a, b) \triangleq \sum_{n=0}^{\infty} \frac{a^n}{n!} e^{-\left(\frac{b-n}{\sqrt{2}\sigma}\right)^2}.$$

**Lemma.** The function  $\phi(u)$  is Lipschitz differentiable and the Lipschitz constant for  $\dot{\phi}(u)$  is:

$$\max\{|\ddot{\phi}(u)|\} \triangleq \mu = \left(1 - e^{-\frac{1}{\sigma^2}}\right) e^{\frac{2y_{\max}-1}{\sigma^2}}, \text{ where } y_{\max} = \max_{i \in \{1, \dots, M\}} \{y_i\}. \quad (3.5)$$

The proof is given in [18].

**Theorem 1** Assume  $|x_j|$  is bounded above by  $C$  for each  $j$ , a Lipschitz constant of  $\nabla g_{\text{PG}}(\mathbf{x})$  is

$$\begin{aligned}\mathcal{L}(\nabla g_{\text{PG}}) &\triangleq 2\|\mathbf{A}\|_2^2 \left( 2C^2\|\mathbf{A}\|_\infty^2 \tilde{y}_{\max} + \left| 1 - C^2\|\mathbf{A}\|_\infty^2 \tilde{y}_{\max} \right| \right), \\ \tilde{y}_{\max} &\triangleq \left( 1 - e^{-\frac{1}{\sigma^2}} \right) e^{\frac{2y_{\max}-1}{\sigma^2}}.\end{aligned}\quad (3.6)$$

where  $y_{\max}$  is  $\max_i\{|y_i|\}$ ,  $i = 1, \dots, M$ .

**Proof:** Let  $g_{\text{PG}}(\mathbf{x})$  denote a function that maps a vector  $\mathbf{x} \in \mathbb{R}^N$  to a scalar; it is the sum of each  $g_i(\mathbf{x}) \triangleq \phi_i(|\mathbf{a}'_i \mathbf{x}|^2 + b_i; y_i)$  over  $i = 1, \dots, M$ . Let  $\mathbf{g}(\mathbf{x})$  denote a function that maps a vector  $\mathbf{x} \in \mathbb{R}^N$  to the measurement space  $\mathbf{y} \in \mathbb{R}^M$ ; it is the concatenation of each  $g_i(\mathbf{x})$ . So  $\nabla g_{\text{PG}}(\mathbf{x}) \in \mathbb{R}^N$ ,  $\nabla^2 g_{\text{PG}}(\mathbf{x}) \in \mathbb{R}^{N \times N}$ , and  $\nabla \mathbf{g}(\mathbf{x}) \in \mathbb{R}^{M \times N}$ .

By the chain rule, the Hessian of  $g_{\text{PG}}$  is

$$\nabla^2 g_{\text{PG}}(\mathbf{x}) = 2\mathbf{A}' (\text{diag}\{\mathbf{A}\mathbf{x}\} \nabla \mathbf{g}(\mathbf{x}) + \text{diag}\{\mathbf{g}(\mathbf{x})\} \mathbf{A}). \quad (3.7)$$

Assume  $|x_j|$  is bounded above by  $C$  for each  $j$ . Then it follows that  $\|\text{diag}\{\mathbf{A}\mathbf{x}\}\|_2 \leq C\|\mathbf{A}\|_\infty$  by the construction of matrix-vector multiplication, leading to a Lipschitz constant for  $\nabla g_{\text{PG}}(\mathbf{x})$ :

$$\mathcal{L}(\nabla g_{\text{PG}}) = 2C\|\mathbf{A}\|_2\|\mathbf{A}\|_\infty\|\nabla \mathbf{g}(\mathbf{x})\|_2 + 2\|\mathbf{A}\|_2^2\|\text{diag}\{\mathbf{g}(\mathbf{x})\}\|_2. \quad (3.8)$$

Here  $\mathcal{L}(\nabla g_{\text{PG}})$  denotes a Lipschitz constant for  $\nabla g_{\text{PG}}$ , not necessarily the best one. To compute  $\|\nabla \mathbf{g}(\mathbf{x})\|_2$ , we substitute the Lipschitz constant of  $\dot{\phi}(u)$  into (3.4) and apply Lemma 3.2.2.2, leading to

$$\|\nabla \mathbf{g}(\mathbf{x})\|_2 \leq 2C\|\mathbf{A}\|_2\|\mathbf{A}\|_\infty \left( 1 - e^{-\frac{1}{\sigma^2}} \right) e^{\frac{2y_{\max}-1}{\sigma^2}}. \quad (3.9)$$

To compute  $\|\text{diag}\{\mathbf{g}(\mathbf{x})\}\|_2$ , let

$$t \in [b, \max_i\{|\mathbf{a}'_i \mathbf{x}|^2\} + b] \subseteq \mathcal{T} \triangleq [b, C^2\|\mathbf{A}\|_\infty^2 + b]. \quad (3.10)$$

From the fact that  $\dot{\phi}(t) \leq 1$  by its construction, one can derive that

$$\|\text{diag}\{\mathbf{g}(\mathbf{x})\}\|_2 = \|\mathbf{g}(\mathbf{x})\|_\infty \leq \max_{t \in \mathcal{T}}\{|\dot{\phi}(t)|\} \leq \left| 1 - C^2\|\mathbf{A}\|_\infty^2 \max\{|\ddot{\phi}(t)|\} \right|. \quad (3.11)$$

Combining (3.8), (3.9) and (3.11) completes the proof of Theorem 1.

However, due to the infinite sum in Poisson-Gaussian log-likelihood (3.3), we approximate  $s(a, b)$  with a finite sum following [18]:

$$s(a, b) \approx \sum_{n=0}^{n^+} \frac{a^n}{n!} e^{-\left(\frac{b-n}{\sqrt{2}\sigma}\right)^2}, \quad n^+ = \lceil n^* + \delta\sigma \rceil, \quad (3.12)$$

with  $n^*$  given by

$$\begin{aligned} n^* &= \sigma \mathcal{W}\left(\frac{a}{\sigma^2} e^{b/\sigma^2}\right) \\ &\approx \sigma \left( \frac{b}{\sigma^2} \log\left(\frac{a}{\sigma^2}\right) - \log\left(\frac{b}{\sigma^2} \log\left(\frac{a}{\sigma^2}\right)\right) \right) \\ &= \frac{b}{\sigma} \log\left(\frac{a}{\sigma^2}\right) - \sigma \log\left(\frac{b}{\sigma^2} \log\left(\frac{a}{\sigma^2}\right)\right), \end{aligned} \quad (3.13)$$

where  $\mathcal{W}(\cdot)$  denotes the Lambert function. The accuracy of this approximation is controlled by  $\delta$ . Reference [18] provides a comprehensive analysis on the maximum error value of the truncated sum (3.12) and found the bound was very precise.

### 3.2.2.3 Accelerated WF with Score-based Image Prior

---

**Algorithm 1:** Our proposed accelerated WF with score-based image prior.

---

**Input:** Measurement  $\mathbf{y}$ , system matrix  $\mathbf{A}$ , momentum factor  $\eta_0 = 1$ , step size factor  $\beta$ , weighting factor  $\gamma$ , truncation operator  $\mathcal{P}_C(\cdot) \rightarrow [0, C]$ ; initial image  $\mathbf{x}_0$ , initial auxiliary variables  $\mathbf{z}_0 = \mathbf{w}_0 = \mathbf{v}_0 = \mathbf{x}_0$ , initialize  $\sigma_1 > \sigma_2 > \dots > \sigma_K$ .

**for**  $k = 1 : K$  **do**

**for**  $t = 1 : T$  **do**

        Set step size  $\mu = \beta\sigma_k^2$ .

        Set  $\Delta z_{t,k} = \frac{\eta_{t-1,k}}{\eta_{t,k}}(\mathbf{z}_{t,k} - \mathbf{x}_{t,k})$ .

        Set  $\Delta x_{t,k} = \frac{\eta_{t-1,k}-1}{\eta_{t,k}}(\mathbf{x}_{t,k} - \mathbf{x}_{t-1,k})$ .

        Set  $\mathbf{w}_{t,k} = \mathcal{P}_C(\mathbf{x}_{t,k} + \Delta z_{t,k} + \Delta x_{t,k})$ .

        Compute  $s_\theta(\mathbf{x}_{t,k}, \sigma_k)$  and  $s_\theta(\mathbf{w}_{t,k}, \sigma_k)$ .

        Set  $\mathbf{z}_{t+1,k} = \mathbf{w}_{t,k} - \mu(\nabla g_{\text{PG}}(\mathbf{w}_{t,k}) + s_\theta(\mathbf{w}_{t,k}, \sigma_k))$ .

        Set  $\mathbf{v}_{t+1,k} = \mathbf{x}_{t,k} - \mu(\nabla g_{\text{PG}}(\mathbf{x}_{t,k}) + s_\theta(\mathbf{x}_{t,k}, \sigma_k))$ .

        Set  $\eta_{t+1,k} = \frac{1}{2} \left( 1 + \sqrt{1 + 4\eta_{t,k}^2} \right)$ .

        Set  $\mathbf{x}_{t+1,k} = \mathcal{P}_C(\gamma_{t,k}\mathbf{z}_{t+1,k} + (1 - \gamma_{t,k})\mathbf{v}_{t+1,k})$ .

**end**

**end**

**Output:** Return  $\mathbf{x}_{T,K}$ .

---

For accelerating the WF algorithm, we followed the implementation of [52] as its convergence guarantee was proved. Assuming that the true score function can be learned properly, when we have a trained score function  $s_\theta(\mathbf{x}, \sigma)$  by applying (3.1), the gradient descent algorithm for MAP estimation has the form:  $\mathbf{x}_{t+1} = \mathbf{x}_t - \mu(\nabla g(\mathbf{x}_t) + s_\theta(\mathbf{x}_t, \sigma_k))$ . Algorithm 1 summarizes our proposed AWFS algorithm. In a similar fashion as Langevin dynamics, we choose  $\sigma_k$  to be a descending scale of noise levels. In practice, we generally use each noise level a fixed number of times, with geometrically spaced noise levels between some lower and upper bound. The step size factor  $\beta$  in Algorithm 1 can be selected empirically, but we show that the Lipschitz constant of the gradient  $\nabla g_{\text{PG}}(\mathbf{x}_t) + s_\theta(\mathbf{x}_t, \sigma_k)$  exists as demonstrated in Theorem 2 (the proof is given in the Appendix C).

We assume that the data allows the neural network to learn the score function well, i.e.,  $s_\theta(\mathbf{x}, \sigma) \approx \nabla \log(p_\sigma(\mathbf{x}))$ , and  $p_\sigma(\mathbf{x}) = p(\mathbf{x}) \otimes \mathcal{N}(0, \sigma^2)$ , where  $\otimes$  denotes (circular) convolution. One can show that  $\nabla \log(p_\sigma(\mathbf{x}))$  is Lipschitz continuous on  $[-C, C]^N$ . The proof is given in Appendix C. Using  $p_\sigma(\mathbf{x})$ , we define the smoothed posterior as

$$p_\sigma(\mathbf{x}|\mathbf{A}, \mathbf{y}, \bar{\mathbf{b}}, \mathbf{r}) \propto p(\mathbf{y}|\mathbf{A}, \mathbf{x}, \bar{\mathbf{b}}, \mathbf{r})p_\sigma(\mathbf{x}). \quad (3.14)$$

**Theorem 2** *For a smooth density function  $p_{\sigma_k}(\mathbf{x})$  that has finite expectation with  $\sigma_k > 0$ , the Lipschitz constant of  $\nabla g_{\text{PG}}(\mathbf{x}_{t,k}) + s_\theta(\mathbf{x}_{t,k}, \sigma_k)$  exists when each element in  $\mathbf{x}_{t,k}$  satisfies  $0 < |x_j| < C$  for each  $j$ . Furthermore, if the weighting factor  $\gamma \in \{0, 1\}$  is chosen appropriately following [52], i.e., according to the higher posterior probability between  $p_{\sigma_k}(z|\mathbf{y}, \mathbf{A}, \bar{\mathbf{b}}, \mathbf{r})$  and  $p_{\sigma_k}(v|\mathbf{y}, \mathbf{A}, \bar{\mathbf{b}}, \mathbf{r})$ ; then with sufficiently small  $\beta$ , the inner iteration sequence  $\{\mathbf{x}_{t,k}\}$  generated by Algorithm 1 is bounded, and any accumulation point of  $\{\mathbf{x}_{t,k}\}$  is a critical point of the posterior distribution  $p_{\sigma_k}(\mathbf{x}|\mathbf{y}, \mathbf{A}, \bar{\mathbf{b}}, \mathbf{r})$  in (3.14).*

**Proof:** By Lipschitz continuity of  $\log(p_\sigma(\mathbf{x}))$ , and from the design of Algorithm 1,  $\mathbf{x}_{t,k}$  and  $\mathbf{w}_{t,k}$  are both bounded between  $[0, C]$  for all  $t, k$ , so the Lipschitz constant  $\mathcal{L}^*$  of  $\nabla g_{\text{PG}}(\cdot) + s_\theta(\cdot)$  exists. With the step size  $\mu$  satisfying  $0 < \mu < \frac{1}{\mathcal{L}^*}$ , and the weighting factor  $\gamma \in \{0, 1\}$  being chosen according to the higher posterior probability between  $p_{\sigma_k}(z|\mathbf{A}, \mathbf{y}, \bar{\mathbf{b}}, \mathbf{r})$  and  $p_{\sigma_k}(v|\mathbf{A}, \mathbf{y}, \bar{\mathbf{b}}, \mathbf{r})$  (see [52]), we satisfy all conditions in Theorem 1 of [52], which establishes the critical-point convergence of the sequence  $\mathbf{x}_{t,k}$  generated by Algorithm 1 for any  $\sigma_k, k = 1, \dots, K$ . Hence the sequence  $\mathbf{x}_{t,k}$  generated by Algorithm 1 converges as  $t \rightarrow \infty$  to a critical-point of the posterior  $p_{\sigma_k}(\mathbf{x}|\mathbf{A}, \mathbf{y}, \bar{\mathbf{b}}, \mathbf{r})$  for any  $\sigma_k$ .

### 3.2.3 Experiment

#### 3.2.3.1 Experiment Settings

**Dataset.** We tested all algorithms on three datasets: 162 histopathology images related to breast cancer [2] (train/val/test is 122/20/20); 920 images from CelebA dataset [65] (train/val/test is 800/100/20); and 720 images from a homemade CT-density dataset (train/val/test is 600/100/20). The CT-density dataset was generated from SPECT/CT images for Yttrium-90 radionuclide therapy after applying the CT-to-density calibration curve [54]. Although the size of training datasets are relatively small compared to typical datasets such as ImageNet or LSUN [37, 93] that have millions of images, we do not require the score functions to learn image priors strong enough to generate realistic images from white Gaussian noise; rather, it is sufficient for the priors to be able to denoise moderately noisy images.

**System Model.** Similar to (??), we define the system matrix to be discrete Fourier transform of the concatenation of the true image  $\mathbf{x}$ , a blank image  $\mathbf{0}$  and a reference image  $\mathcal{R}$  with scaling and oversampling. We set the scaling factor  $\alpha$  to be in the range  $[0.02, 0.035]$  so that the average counts per pixel range from 6 to 25; the oversampled ratio is set to 2. We set  $\mathcal{R}$  to be a binary random image similar to what was used in [49]. The standard deviation of the Gaussian read noise added to the measurements  $\mathbf{y}$  was set as  $\sigma \in [0.5, 1.5]$ .

**Implemented Algorithms.** For unregularized algorithms, we implemented Gaussian WF, Poisson WF and Poisson-Gaussian WF. For regularized algorithms, we implemented smoothed total variation (TV) based on the Huber function [40, p. 184] and PNP/RED methods with the DL denoiser [118]: PNP-ADMM [101], PNP-PGM [45], and RED-SD [80]. We also implemented the RED-SD algorithm with “Noise2Self” zero-shot image denoising network [5] (RED-SD-SELF). For diffusion models, we implemented DOLPH [87] and our proposed AWFS. The implementation details of each algorithm can be found in the appendix of [59]. We used spectral initialization [69] for the Gaussian PR and Poisson PR methods; we then used the output results from Poisson PR to initialize other algorithms. We ran all algorithms until convergence in normalized root mean squared error (NRMSE) or reached the maximum number of iterations (e.g., 50).

To evaluate the robustness and limitation of these algorithms, we first tuned the parameters for each algorithm at the noise level when  $\alpha = 0.030$  and  $\sigma = 1$ , and then held them fixed throughout all experiments (Table 3.1, Table 3.2, Fig. 3.7 and Fig. 3.8). In practice the ground truths are unknown, so oracle tuning of test datasets is infeasible (though

**TBL 3.1** – SSIM and NRMSE for Poisson and PG likelihoods. Results were averaged across 7 different noise levels by varying  $\alpha \in 0.02 : 0.005 : 0.035$  in (3.2).

Likelihood	Unregularized (SSIM/NRMSE)		DOLPH (SSIM/NRMSE)		AWFS (SSIM/NRMSE)	
DataSet: Histopathology [2]						
Gaussian	0.52 ± 0.18	41.2 ± 25.3	0.76 ± 0.07	18.0 ± 3.0	0.84 ± 0.06	16.2 ± 3.7
Poisson	0.54 ± 0.18	31.7 ± 10.2	0.72 ± 0.13	19.5 ± 6.1	0.83 ± 0.06	16.2 ± 3.7
Poisson-Gaussian	0.57 ± 0.18	28.9 ± 9.0	0.80 ± 0.06	16.0 ± 2.9	<b>0.85 ± 0.05</b>	<b>15.4 ± 3.7</b>
DataSet: CelebA [65]						
Gaussian	0.31 ± 0.09	55.6 ± 13.9	0.70 ± 0.12	14.5 ± 17.4	0.72 ± 0.16	15.3 ± 11.8
Poisson	0.39 ± 0.10	24.5 ± 11.4	0.61 ± 0.12	15.6 ± 10.6	0.72 ± 0.16	15.2 ± 11.8
Poisson-Gaussian	0.42 ± 0.10	21.8 ± 9.1	0.71 ± 0.11	<b>13.7 ± 11.1</b>	<b>0.74 ± 0.15</b>	14.8 ± 11.9
DataSet: CT-Density						
Gaussian	0.29 ± 0.09	50.5 ± 8.0	0.51 ± 0.12	22.4 ± 3.9	0.82 ± 0.11	19.1 ± 4.8
Poisson	0.19 ± 0.06	48.9 ± 13.1	0.38 ± 0.11	25.6 ± 7.5	0.84 ± 0.08	17.8 ± 4.3
Poisson-Gaussian	0.24 ± 0.06	40.8 ± 9.5	0.55 ± 0.08	20.0 ± 3.3	<b>0.88 ± 0.05</b>	<b>16.4 ± 3.7</b>

some form of cross validation may be possible). Though the numbers reported could fluctuate after careful refinement, *e.g.*, by performing grid search on tuning parameters, such techniques would potentially impede the algorithm’s practical use.

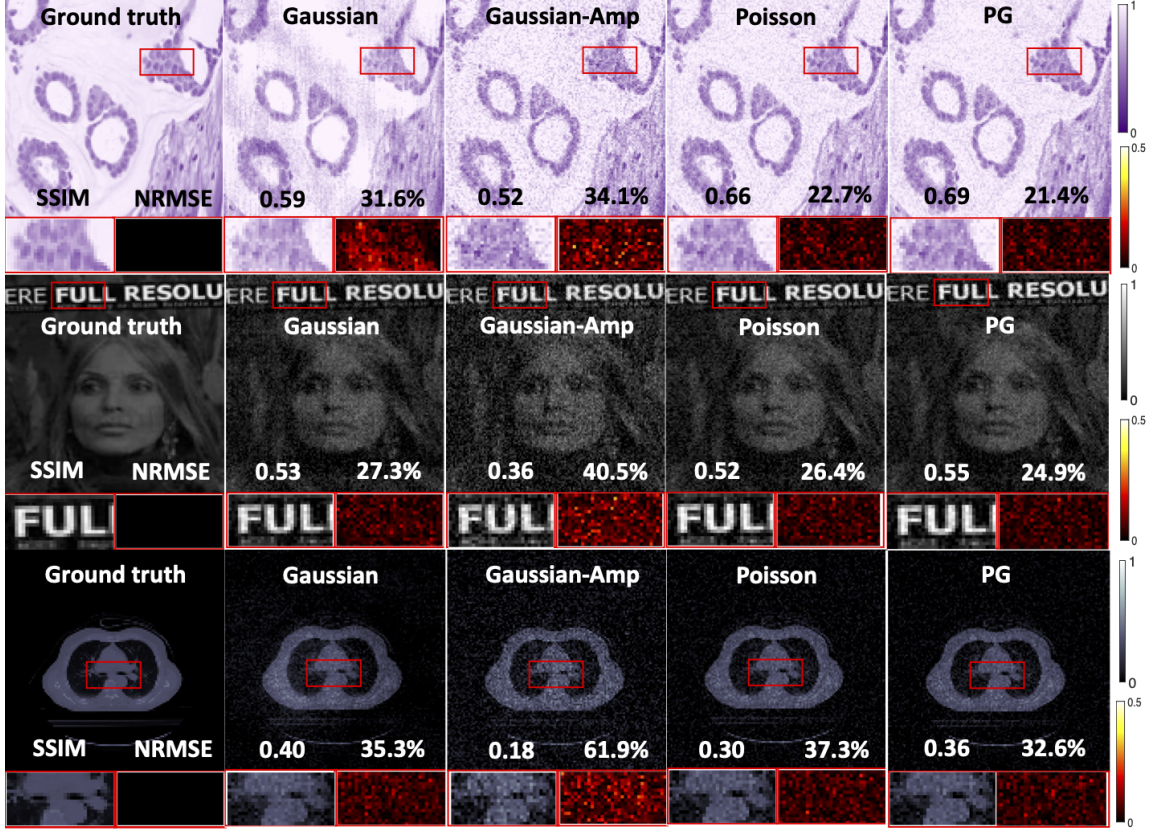
**Network Training.** For PNP denoising networks, we trained all denoisers on different noise levels  $\sigma \in \{9, 11, 13, 15\}$  and found that  $\sigma = 15$  worked the best on our data. We also used the denoiser scaling technique from [112] to dynamically adjust the performance of all PNP methods. To perform score matching, we applied 20 geometrically spaced noise levels between 0.005 and 0.1 on each of the training images. All networks were implemented in PyTorch and trained on an NVIDIA Quadro RTX 5000 GPU using the ADAM optimizer [48] for 1000 epochs with the best one being selected based off the validation error, *i.e.*, the mean squared error (MSE) loss.

### 3.2.3.2 Results

We compared all implemented algorithms both qualitatively, by visualizing the reconstructed images and residual errors, and quantitatively, by computing the NRMSE and structural similarity index measure (SSIM) [108]. Due to the global phase ambiguity, *i.e.*, all the algorithms can recover the signal only to within a constant phase shift due to the loss of global phase information, we corrected the phase of  $\hat{x}$  by  $\hat{x}_{\text{corrected}} \triangleq \text{sign}(\langle \hat{x}, x_{\text{true}} \rangle) \hat{x}$ .

Fig. 3.2 shows experiments of running unregularized methods based on different noise models on the histopathology, CelebA, and CT density datasets. For comparison, we ran



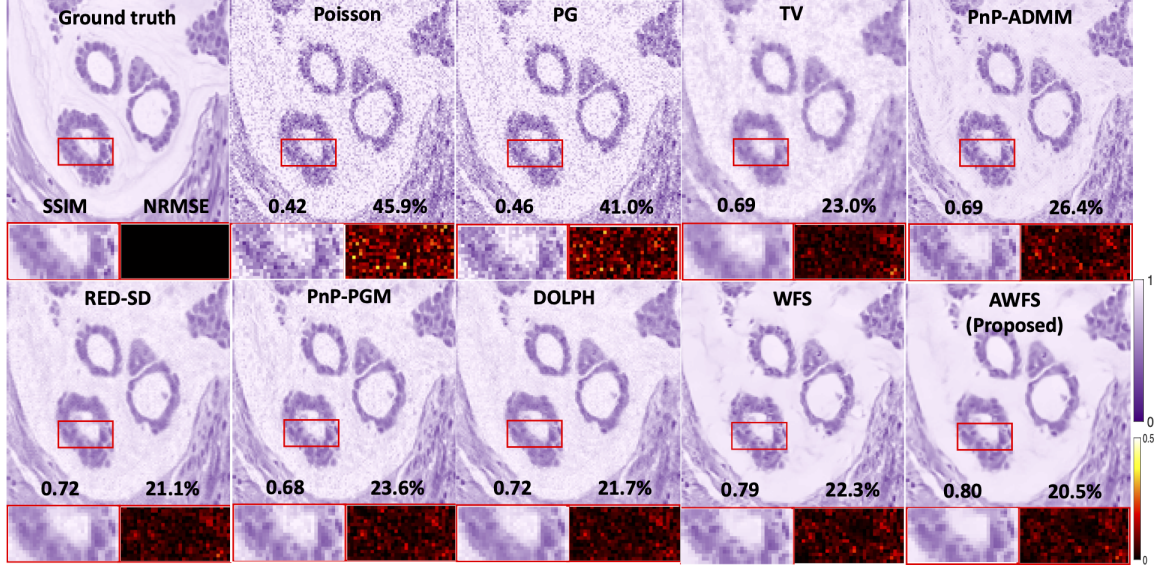


**FIG 3.2** – Reconstructed images by unregularized methods (Gaussian, Gaussian-Amplitude, Poisson and Poisson-Gaussian) on Histopathology dataset [2], celebA dataset [65] and CT-density dataset. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image. We used  $\alpha = 0.035$  and  $\sigma = 1$ .

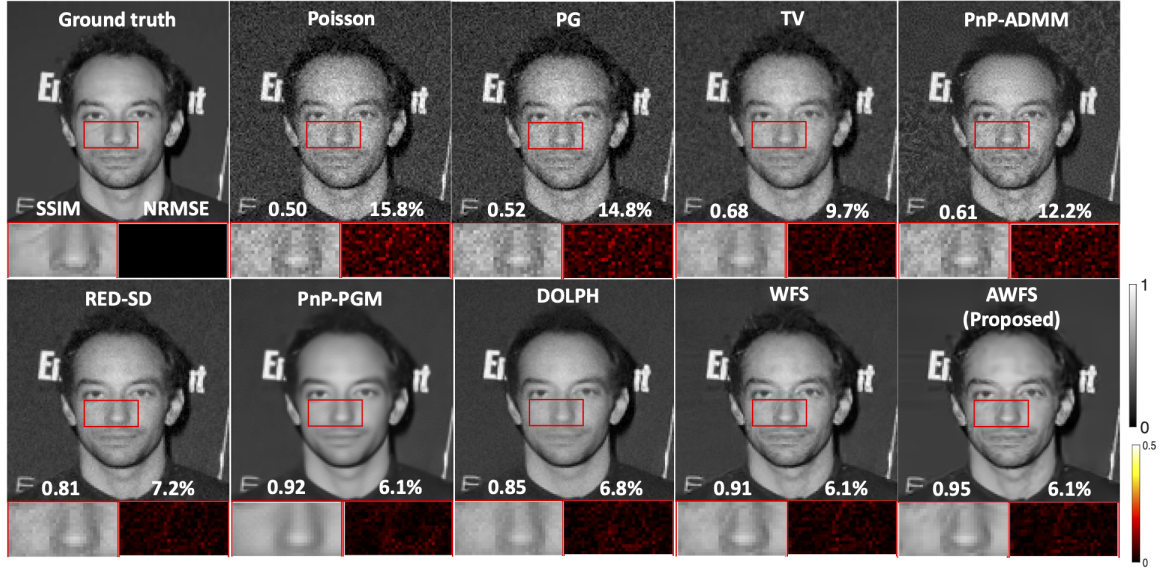
the unregularized methods with a Gaussian only noise model, Poisson only, and PG noise model.

Fig. 3.3, Fig. 3.4 and Fig. 3.5 visualize reconstructed images generated by algorithms mentioned in the previous section. The WF with PG likelihood outperforms WF with Poisson likelihood with a consistently higher SSIM and lower NRMSE. Moreover, we found unregularized Gaussian WF failed to reconstruct images similar to what was reported in [77]. Of the regularized algorithms with PG likelihood, our proposed AWFS had less visual noise and achieved greater detail recovery compared to other methods, as evidenced by the zoomed-in area in these figures. Fig. 3.6 shows that for a variety of datasets, when combined with the AWFS method, while the Poisson only and Gaussian only models lead to reasonable reconstructions, the PG noise model leads to the highest quality image. For all three datasets shown, when used in conjunction with our AWFS method, including both Poisson and Gaussian likelihoods results in the highest quality reconstruction both



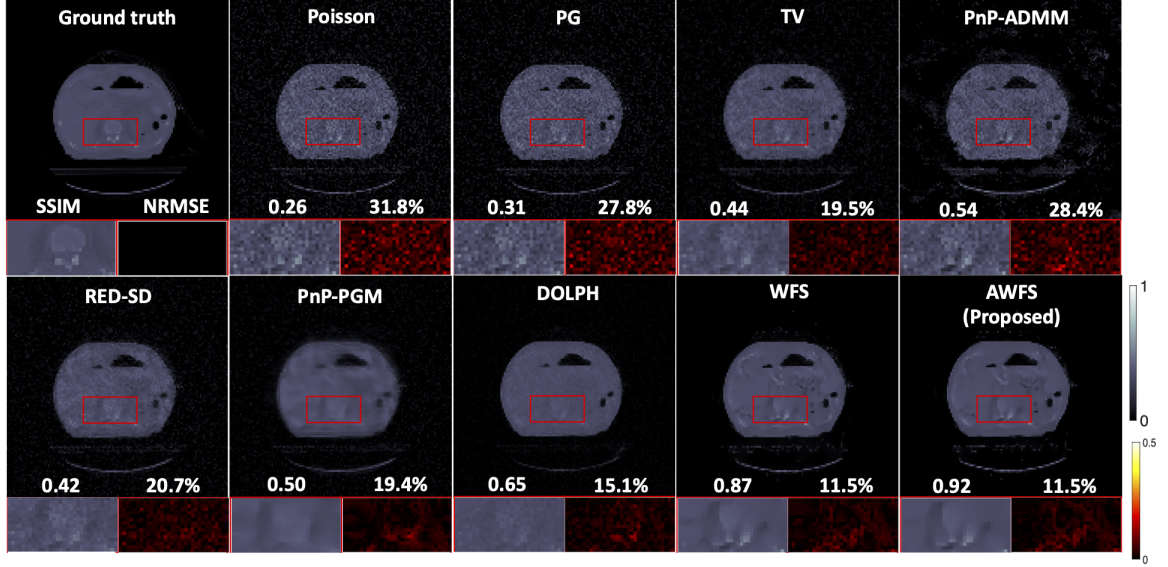


**FIG 3.3** – Reconstructed images on dataset [2]. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image.  $\alpha$  and  $\sigma$  were set to 0.02 and 1, respectively.



**FIG 3.4** – Reconstructed images on celebA dataset [65]. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image.  $\alpha$  and  $\sigma$  were set to 0.035 and 1, respectively.

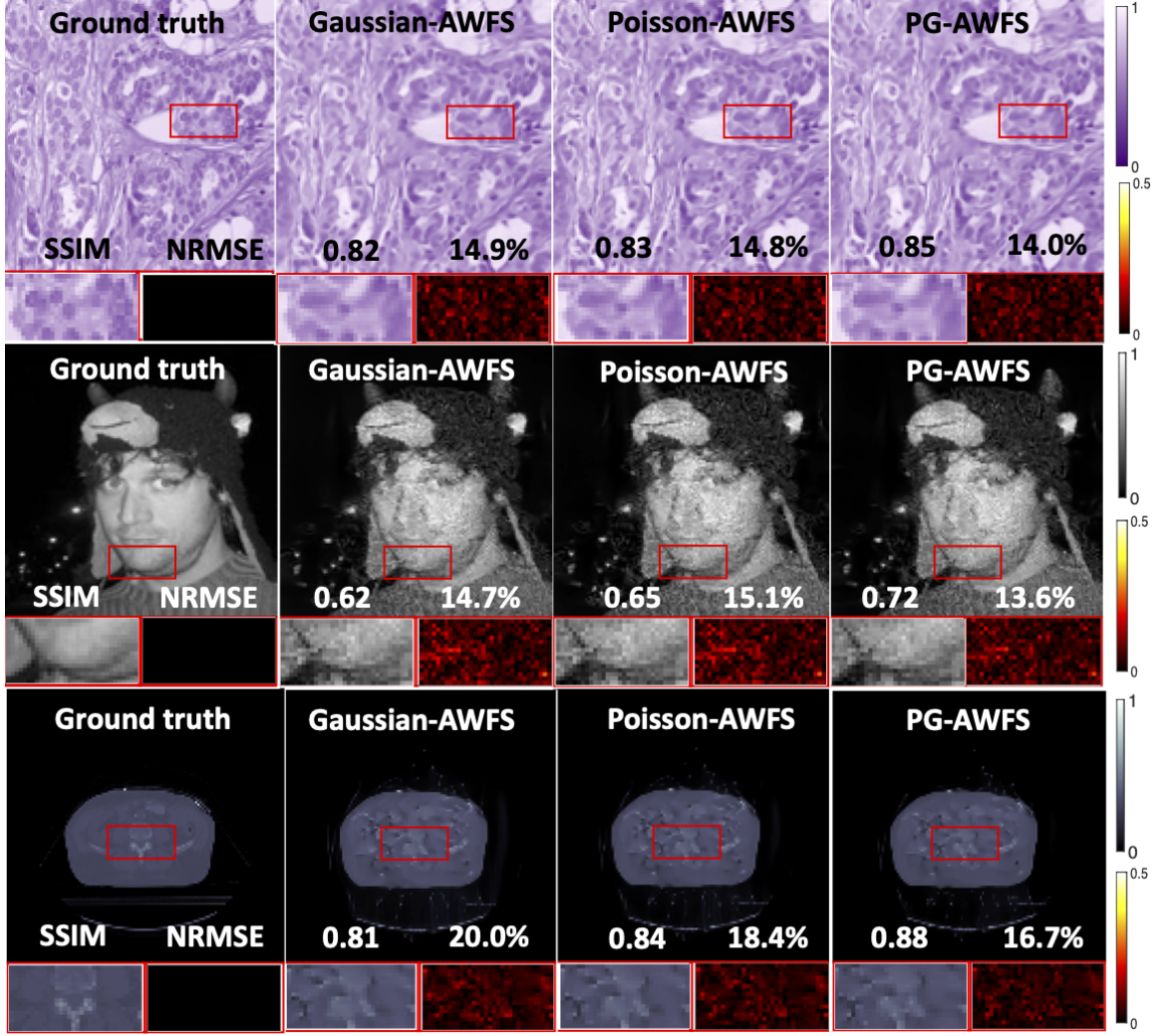
in terms of quantitative metrics as well as visually. Thus, although the score function provides a useful prior for recovering an image when the measurement is very noisy, a proper noise model is also crucial to a high quality reconstruction.



**FIG 3.5** – Reconstructed images on CT-density dataset. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image.  $\alpha$  and  $\sigma$  were set to 0.035 and 1, respectively.

For quantitative evaluations, Table 3.1 exemplifies the effect of using our proposed PG likelihood as compared to the simpler Poisson likelihoods. We did not run the Gaussian likelihood with DOLPH or AWFS due to the abysmal performance with this likelihood. In all cases, usage of the PG likelihood results in improved image quality in terms of both metrics. Table 3.2 consists of experiments using the PG likelihood and shows the efficacy of the proposed AWFS method over other methods. In particular, our AWFS had superior quantitative performance over all other compared methods on the histopathology and CT-density datasets; in contrast, the PnP-PGM showed the lowest NRMSE on celebA dataset. This is likely due to higher randomness in celebrity faces because the effectiveness of generative models can vary depending on the dataset used. Thus, when provided with a small amount of training data with high randomness, image denoising models (DNCNN) may be more effective than generative models.

We also tested the robustness of the leading algorithms in Table 3.2, by varying both scaling factor  $\alpha$  and STD of Gaussian noise  $\sigma$ . Fig. 3.7 and Fig. 3.8 illustrate results, where our AWFS algorithm had the highest SSIM and lowest NRMSE. In Fig. 3.8, AWFS demonstrated minimal variations in SSIM and NRMSE metrics than DOLPH as evidenced by the smaller discrepancies in SSIM (0.17 vs. 0.23) and NRMSE (12.6% vs. 18.2%) when  $\sigma$  varies from 0.75 to 1.5. Fig. 3.9 compares the convergence rate of AWFS vs. WFS for the Poisson and PG likelihood, respectively. Under a variety of noise levels, AWFS consistently converged faster than WFS in terms of number of iterations.



**FIG 3.6** – Reconstructed images by Gaussian, Poisson and Poisson-Gaussian log-likelihood model with AWFS image prior. Tested on Histopathology dataset [2], celebA dataset [65] and CT-density dataset. The bottom left/right subfigures correspond to the zoomed in area and the error map for each image.  $\alpha$  and  $\sigma$  were set to 0.025 and 1, respectively.

It is a known property of diffusion models that they can produce images with hallucinated features if the measurements are insufficiently informative. In the case of low-count phase retrieval with serious corruptions of both Poisson and Gaussian noise, as is investigated here, the measurement is highly corrupted and contains magnitude-only measurements of the original signal. Thus, it may be difficult for the diffusion models to avoid some otherwise realistic hallucinations if the data consistency is not strong enough to guide the model away from such hallucinations. On the other hand, if the measurements



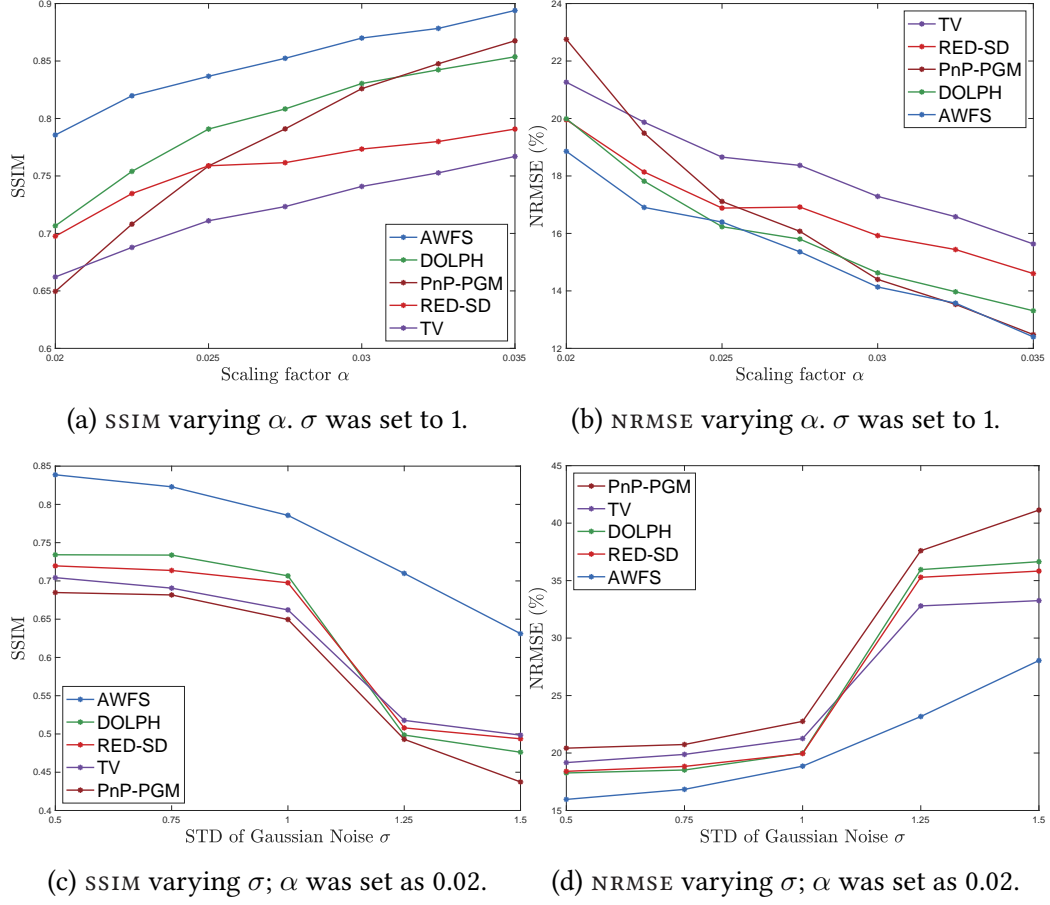
**TBL 3.2** – SSIM and NRMSE using Poisson Gaussian likelihood with different regularization/image prior approaches. Results were averaged across 7 different noise levels by varying  $\alpha \in 0.02 : 0.005 : 0.035$  in (3.2).  $\text{WFS}^*$  runs the same number of iterations as  $\text{AWFS}$  whereas  $\text{WFS}^\dagger$  runs more iterations until convergence.

Dataset	Histopathology [2]		CelebA [65]		CT-Density	
Methods	SSIM	NRMSE (%)	SSIM	NRMSE (%)	SSIM	NRMSE (%)
Unregularized	$0.57 \pm 0.18$	$28.9 \pm 9.0$	$0.42 \pm 0.10$	$21.8 \pm 9.1$	$0.24 \pm 0.06$	$40.8 \pm 9.5$
RED-SD-SELF [5]	$0.66 \pm 0.13$	$21.9 \pm 4.5$	$0.60 \pm 0.09$	$15.9 \pm 10.6$	$0.34 \pm 0.04$	$28.1 \pm 4.1$
PNP-ADMM [101]	$0.71 \pm 0.11$	$20.7 \pm 4.2$	$0.56 \pm 0.08$	$16.7 \pm 8.1$	$0.55 \pm 0.03$	$31.2 \pm 2.7$
TV regularizer	$0.72 \pm 0.11$	$18.2 \pm 3.9$	$0.64 \pm 0.07$	$14.4 \pm 8.6$	$0.41 \pm 0.03$	$23.7 \pm 2.8$
RED-SD [80]	$0.76 \pm 0.09$	$16.8 \pm 3.6$	$0.69 \pm 0.11$	$13.9 \pm 10.9$	$0.38 \pm 0.04$	$25.9 \pm 4.0$
PNP-PGM [45]	$0.78 \pm 0.11$	$16.5 \pm 4.5$	<b><math>0.74 \pm 0.14</math></b>	<b><math>13.5 \pm 11.3</math></b>	$0.42 \pm 0.07$	$24.6 \pm 4.4$
DOLPH [87]	$0.80 \pm 0.06$	$16.0 \pm 2.9$	$0.71 \pm 0.11$	$13.7 \pm 11.1$	$0.55 \pm 0.08$	$20.0 \pm 3.3$
$\text{WFS}^*$	$0.76 \pm 0.12$	$18.2 \pm 5.5$	$0.63 \pm 0.16$	$16.9 \pm 11.8$	$0.53 \pm 0.17$	$21.3 \pm 7.6$
$\text{WFS}^\dagger$	$0.83 \pm 0.06$	$16.2 \pm 4.0$	$0.70 \pm 0.16$	$15.7 \pm 11.8$	$0.74 \pm 0.13$	$17.3 \pm 4.8$
AWFS (Proposed)	<b><math>0.85 \pm 0.05</math></b>	<b><math>15.4 \pm 3.7</math></b>	<b><math>0.74 \pm 0.15</math></b>	$14.8 \pm 11.9$	<b><math>0.88 \pm 0.05</math></b>	<b><math>16.4 \pm 3.7</math></b>

are less corrupted, then the data consistency should be strong enough to avoid such hallucinations. Fig. 3.10 provides examples of this for the CT image dataset via a comparison of the reconstruction quality of the  $\text{AWFS}$  method over a range of count levels. With the lowest scaling factor, *e.g.*,  $\alpha = 0.02$ , the measurements were seriously corrupted with noise, and the method may hallucinate some features. However, at higher count level, *e.g.*,  $\alpha = 0.05$ , there is enough information in the measurement to enforce consistency and avoid noticeable hallucinations. We performed the same experiment twice with different noisy initializations and all other parameters held equal to demonstrate robustness of the method under different initializations.

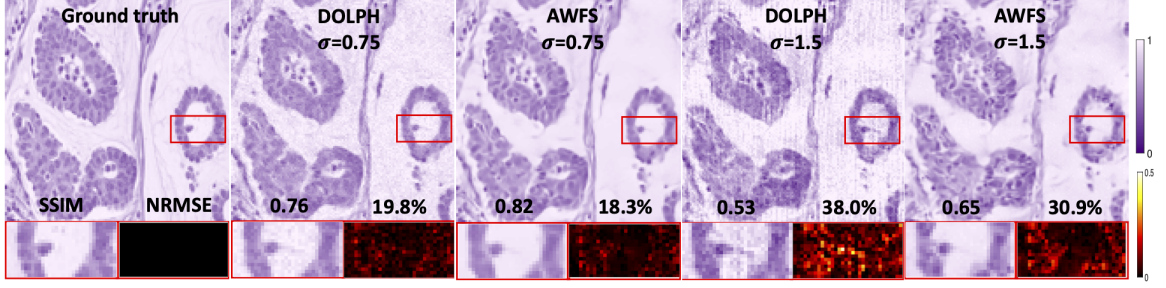
### 3.2.4 Discussion

PR has a long-standing history in the field of signal processing and imaging. Pioneering works such as the error reduction and hybrid input-output algorithms by Gerchberg Saxton [33] and Fienup [31] have been proposed to address this problem. These iterative algorithms involve constraints imposed on evaluations between the image domain and frequency domain. However, these methods have limitations in terms of the quality of reconstructed images and their convergence remains uncertain [115]. Another approach to solving PR problems is through compressed sensing and optimization techniques like  $\text{WF}$  [12], matrix lifting [11, 10, 86],  $\text{MM}$  [77] and  $\text{ADMM}$  [63]. This work focuses on the  $\text{WF}$

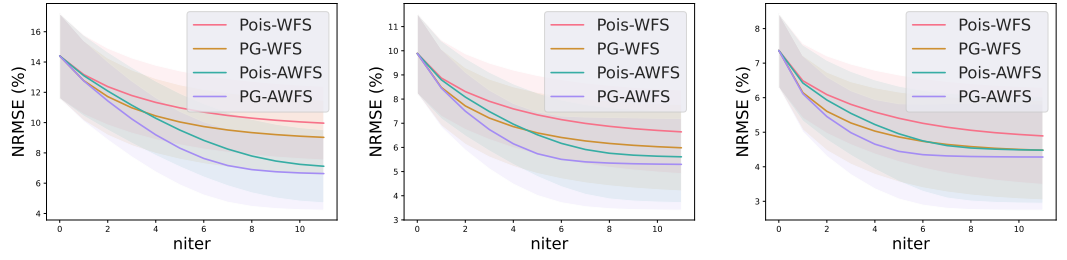


**FIG 3.7** – Comparison of SSIM and NRMSE varying scaling factor  $\alpha \in [0.02, 0.035]$  and STD of Gaussian noise  $\sigma \in [0.25, 1.5]$  defined in (3.2).

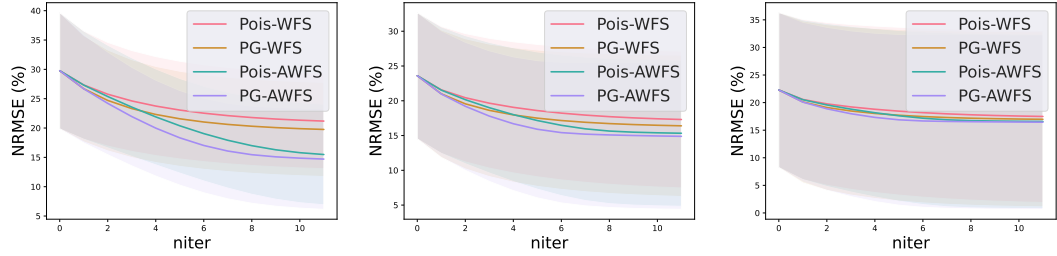
algorithm due to being straightforward to incorporate with the DL regularizer for the image prior. The likelihood modelling of the noise statistics existing in the measurement is also critical. Previous studies have primarily focused on modelling either Gaussian or Poisson likelihood only, but in practical scenarios, both types of noise are often encountered. Therefore, this work contributes to a more practical perspective of addressing the holographic PR problem by using a PG likelihood and incorporating state-of-the-art deep learning image priors. In the case where the measurement is contaminated with Poisson and Gaussian noise, the speedup in reconstruction is crucial, as the bottleneck of our algorithm is in computing the PG likelihood. Additionally, though it is viable to perform a large number of neural network evaluations to perform image reconstruction, it is unrealistic to compute a similarly large number of PG likelihoods. Thus, we perform acceleration in WF algorithm following [52], which guarantees convergence to a critical point for the Holographic PR problems.



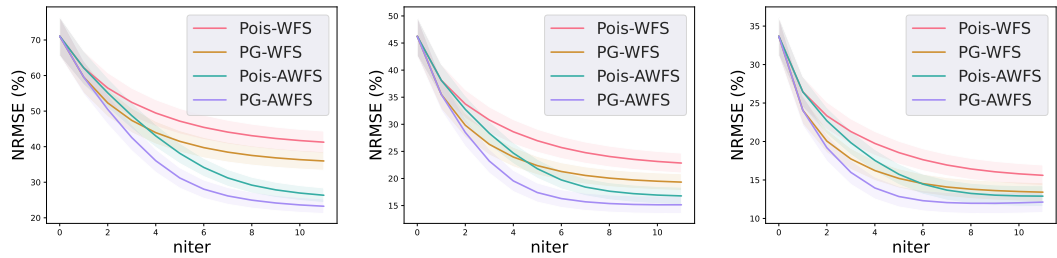
**FIG 3.8** – Reconstructed images by DOLPH [87] and our proposed AWFS method under different  $\sigma$  values. Scaling factor  $\alpha$  was set to 0.02 (defined in (3.2)).



(a) Histopathology dataset:  $\alpha = 0.02$ . (b) Histopathology dataset:  $\alpha = 0.0275$ . (c) Histopathology dataset:  $\alpha = 0.035$ .



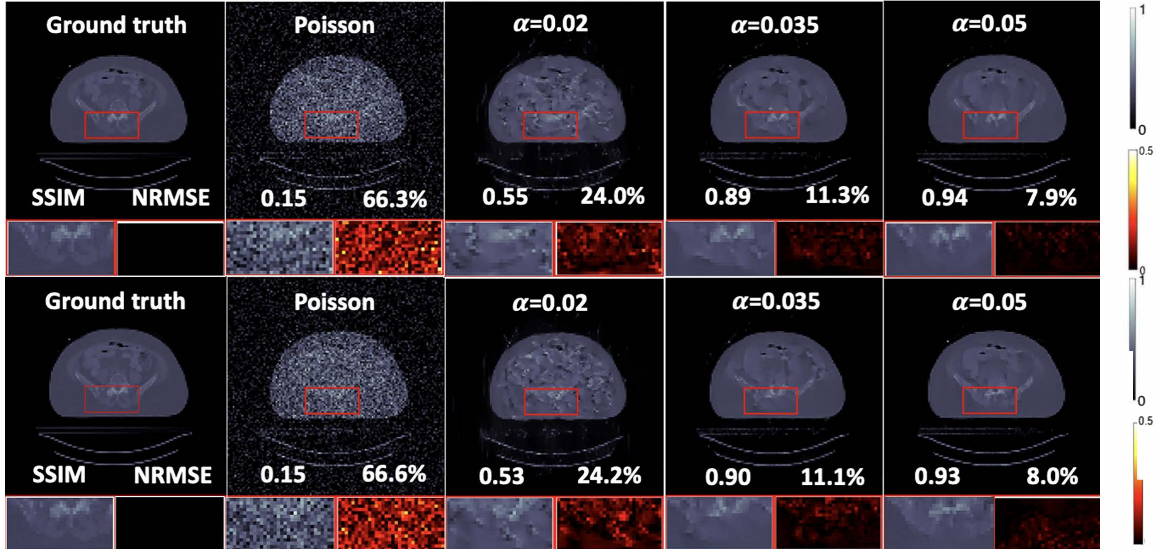
(d) CelebA dataset:  $\alpha = 0.02$ . (e) CelebA dataset:  $\alpha = 0.0275$ . (f) CelebA dataset:  $\alpha = 0.035$ .



(g) CT-density dataset:  $\alpha = 0.02$ . (h) CT-density dataset:  $\alpha = 0.0275$ . (i) CT-density dataset:  $\alpha = 0.035$ .

**FIG 3.9** – Comparing AWFS vs. WFS with NRMSE vs. number of iterations under different noise levels. The curves and shadows represent the mean and standard deviation, respectively.

In our evaluation of three datasets, we consistently observed that the use of PG likelihood yielded superior performance compared to using either Poisson or Gaussian likelihood alone, as expected. Additionally, the results obtained from the CT-density dataset



**FIG 3.10** – Reconstructed images by the unregularized Poisson method (the second column) as well as with the AWFS method for different scaling factors  $\alpha$  (third to fifth columns). The top and bottom rows show reconstructions from different measurement realizations.

were generally of lower quality than those from the other two datasets. This can be attributed to lower average counts per pixel (many zero pixels near the image borders).

Using a DL image prior can be considered from two perspectives: training a denoiser or training to learn the density distribution of images. In our work, we applied both approaches and observed that the effectiveness of these methods differed depending on the dataset tested. Specifically, in the Histopathology dataset [2] and the CT-density dataset, where the images share similar structures, the generative models performs better even when trained with limited data. In the case of the CelebA dataset [65], which includes a wide variety of celebrity faces, generative models did not exhibit as strong performance as denoiser methods when trained on limited data. This is likely due to the fact that generating high-quality images is generally more challenging than removing noise from existing images and may necessitate a larger training dataset. We have also noticed that the PNP-ADMM method provided unsatisfying reconstruction quality, possibly attributable to the non-zero duality gap and slow convergence for non-convex problems [106]. We plan to investigate it further in the future.

The effectiveness of accelerated wF compared to vanilla wF is due to the non-convexity of the PR problem. Although recent advances in geometric landscape analysis of PR can

guarantee that all local minimizers are global even with random initialization [8], in practice the measurements are contaminated by noise so that many more measurements are required for the cost function to have a benign geometric landscape.

Despite the promising results achieved with our proposed **AWFS** approach, there are several limitations of our work. First, the approximate calculation of the infinite sum in (3.3) is accurate but computationally expensive. Future work should seek ways to accelerate this calculation while maintaining accuracy. Second, we did not implement and test the accelerated **WF** applied on the diffusion posterior sampling method [19], for which the network is fine-tuned from a pretrained state-of-the-art diffusion model. This approach has the potential to advance current methods in **PR** problem and we will investigate it in the future. Another limitation of the proposed method is that it has been demonstrated on measurements that are based on simulations. To further demonstrate the efficacy of the method in a real-world setting, future work should consist of evaluating the accuracy of the methods when run on real measurement data. Finally, our experiments are limited to real-valued images, however, our method can be extended to handle complex-values images by splitting real and imaginary components into separate reconstruction routines with different pretrained neural networks [119]. Addressing these limitations will be the future direction of this work.

### 3.2.5 Conclusion

We proposed a novel algorithm based on **Accelerated Wirtinger Flow** and **Score-based image prior (AWFS)** for Poisson-Gaussian holographic phase retrieval. With evaluation on simulated experiments, we demonstrated that our proposed **AWFS** algorithm had the best reconstruction quality both qualitatively and quantitatively and was more robust to various noise levels, compared to other state-of-the-art methods. Furthermore, we proved that our proposed algorithm has a critical-point convergence guarantee. Therefore, our approach has much promise for translation in real-world applications encountering phase retrieval problems.



## CHAPTER 4

# Poisson Inverse Problems in SPECT Imaging<sup>1</sup>

- 4.1 Training End-to-End Unrolled Iterative Neural Networks for SPECT Image Reconstruction<sup>2</sup>
- 4.2 DblurDoseNet: A Deep Neural Network for SPECT Dosimetry Estimation and Resolution Recovery<sup>3</sup>
- 4.3 Efficient Super Resolution Network (ESR-Net) for SPECT Image Reconstruction<sup>4</sup>
- 4.4 Shorter SPECT Scans Using Self-supervised Coordinate Learning to Synthesize Skipped Projection Views

### 4.4.1 Motivation

SPECT imaging has had many advances [78]; however, one continuing limitation is that SPECT acquisition is slow, especially under the low-count conditions encountered when imaging therapy radionuclides, such as  $^{90}\text{Y}$  and  $^{177}\text{Lu}$ . These radionuclides are chosen for the therapeutic properties of their alpha and beta emissions, hence do not have ideal properties for gamma-camera imaging. For example, the photon/gamma-ray yield is relatively

---

<sup>1</sup>This chapter is based on [56, 54, 62].

<sup>2</sup>This section is based on [56].

<sup>3</sup>This section is based on [54].

<sup>4</sup>This section is based on [55].

low, leading to low count conditions. Nevertheless, it is very desirable to perform both therapy and imaging with the same radionuclide, even in very low-count applications.

With  $^{177}\text{Lu}$  where the 208 keV gamma-ray intensity is only 10%, it can take 15-30 mins per bed (40 cm axial) for SPECT on standard gamma-camera systems following radiopharmaceutical therapies (RPTS) such as  $^{177}\text{Lu}$ -DOTATATE and  $^{177}\text{Lu}$ -PSMA [83, 84]. For RPTS involving alpha-emitters, such as Ac-225-PSMA, acquisition times of up to 1 hour have been proposed [26]. This is because both the administered activities and the gamma-ray yields are very low. SPECT under low-count conditions is particularly challenging when multiple beds are needed to encompass metastases and critical organs throughout the body. For example, in PSMA therapy for metastatic castration-resistant prostate cancer (MCRPC), SPECT imaging may require 3-5 bed positions to include all critical organs such as lacrimal glands, salivary gland, bone marrow, and kidneys, as well as lesions that can be throughout the body [44, 103]. Such a procedure demands a significantly greater amount of camera time, which can not only lead to patient discomfort, but can also increase motion artifacts. Additionally, in many facilities, camera availability is limited [103, 85, 71, 25, 17].

To overcome these challenges, a shorter acquisition time is preferable by taking either fewer projection views or shorter acquisition time per view. These strategies pose additional challenges due to either the missing (skipped) view angles or the increased image noise [67]. Numerous algorithms have been proposed with a focus on denoising the reconstructed images from noisy projections to improve image quality [89, 1, 64, 75, 94, 95, 114]. In contrast, the approach of synthesizing the missing projections [82] has been relatively unexplored. Most prior studies have employed deep learning techniques to learn the relationship between one projection and its neighboring views, often relying on ground truth data for training purposes. For instance, Rydén *et al.* [82] used a deep convolutional U-Net trained to generate synthetic intermediate projections. Meanwhile, Li *et al.* [53] introduced a network architecture called LU-Net that integrates Long Short-Term Memory network [38] and U-Net to understand the transformation from sparse-view projection data to full-view data. Chen and Zhou [15] presented a cross-domain method using SPECT images predicted in the image domain as reference for synthesizing full-view projections in the sinogram domain. These approaches are reported to be effective, but they are all based on supervised learning methods that require a sufficient amount of paired data for training. However, in many cases, obtaining enough paired ground truth data for training is challenging or even infeasible. This difficulty is especially true in the case of post-therapy imaging for verifying uptake or dosimetry following RPTS because such imaging is typically not part of routine clinical practice. On the other hand, self-supervised learning, which does not require separate training labels and instead learns from each scan

itself, has the potential to overcome the limitations of supervised learning in such scenarios.

The aim of this research was to reduce SPECT acquisition time by reducing the required number of measured projection views while maintaining image quality by incorporating synthetic projections generated by deep neural networks. We implemented a multi-layer perceptron (MLP) and trained it to generate skipped SPECT projection views through self-supervised coordinate learning [97]. We evaluated the performance of the proposed method both qualitatively and quantitatively in phantom studies and in patients imaged after  $^{177}\text{Lu}$  DOTATATE therapy and  $^{177}\text{Lu}$  PSMA therapy.

## **4.4.2 Materials and Methods**

### **4.4.2.1 Phantom Study**

We used an elliptical phantom with six hot sphere inserts of volumes 2,4,8,16,30,114mL. These “hot” spheres (having the same  $^{177}\text{Lu}$  activity concentration of 0.22 MBq/mL) are placed in a “warm” background (0.035 MBq/mL) to achieve a sphere-to-background ratio of 6.3:1. The sphere volumes of interest (VOIs), corresponding to the physical size, were defined on the CT images.

### **4.4.2.2 Patient Studies**

For the patient studies, we used SPECT/CT scan data from 11 patients imaged after  $^{177}\text{Lu}$ -DOTATATE therapy for neuroendocrine tumor and from 6 patients imaged after  $^{177}\text{Lu}$ -PSMA-617 therapy for MCRPC with the approval of University of Michigan Institutional Review Board (IRB) for retrospective analysis. We defined organs of interest (kidneys for DOTATATE therapy, and kidneys, lacrimal glands, parotid glands, and submandibular glands for PSMA therapy) using deep learning-based methods available within MIM Software. A radiologist manually defined the lesions (78 in total, volume ranging from 2 to 250 mL) as described previously [27].

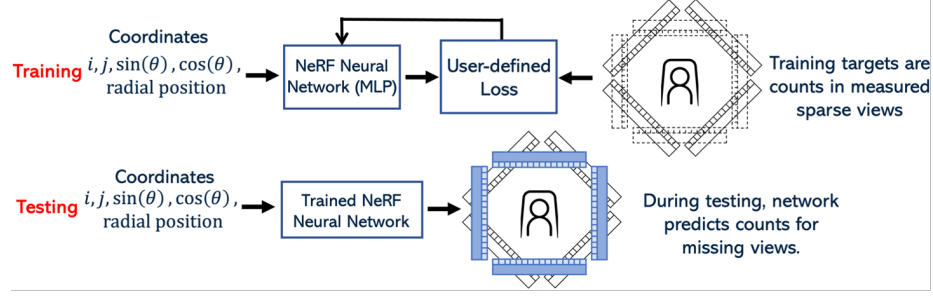
### **4.4.2.3 SPECT/CT Acquisition**

All scans were acquired on a Siemens Intevo Bold SPECT/CT with a 5/8" crystal equipped with medium-energy collimators. Acquisition parameters included 120 views, with 60 views per head, a 20% photopeak window centered at 208 keV, and two adjacent scatter windows of 10% width each. The phantom study used a prolonged acquisition of 196 sec/view to achieve count levels similar to that encountered in patient imaging after  $^{177}\text{Lu}$

therapy. The patient images were acquired under the standard protocols used in our clinic.  $^{177}\text{Lu}$ -DOTATATE SPECT images were acquired for a single bed position at day 2 or day 4 after the cycle 1 administration of 7.4 GBq using an acquisition time of 25 seconds per view (total scan time of 25 min). The  $^{177}\text{Lu}$ -PSMA SPECT images were acquired with two bed positions at day 2 or day 3 after the cycle 1 administration of 7.4 GBq with an acquisition time of about 17 seconds per view per bed (total scan time of about 34 min). The projection view matrix size was  $128 \times 128$ , with a pixel size of  $4.8 \times 4.8\text{mm}$ . The CT images were acquired in low-dose mode (120 kVp; 15 – 80 mAs) under free breathing conditions, with a matrix size of  $512 \times 512$  and pixel size of  $0.98 \times 0.98\text{mm}$ .

#### 4.4.2.4 Self-supervised Coordinate Learning

Given the limited amount of data, we focused on a self-supervised learning approach, rather than supervised methods for this study. Our method draws inspiration from computer vision: the neural radiance field (NERF) approach that models complex 3D scenes through a volumetric scene function [72]. NERF fundamentally uses neural networks to map 3D spatial coordinates to radiance values. In a similar vein, we developed a MLP, comprising 12 hidden layers with 256 neurons each, to synthesize missing projection views in SPECT imaging. Fig. 4.1 illustrates the input to our MLP: 5-dimensional coordinates for each pixel in SPECT projection views. These coordinates consist of pixel position ( $i, j$ ), the sine and cosine of the view angle and radial position (to accommodate noncircular orbits). To enhance the representation of the continuous measurement field, we upscaled the original projection images by a factor of two with the nearest-neighbor resizing method. Consequently, the network input size for each projection view becomes  $(256 \times 256) \times 5$ . The training target consists of measured counts, with a size of  $256 \times 256 \times 1$  for each view. During inference, the model is fed the coordinates of the missing SPECT projections and predicts the corresponding counts both for the main acquisition window and adjacent scatter windows. Our method provides flexible adaption to different numbers of projection views, corresponding to various down-sampling factors (DF). For instance, when trained on 30 measured views and synthesizing 90 views, it achieves a 75% reduction in scan time (DF=4). Additionally, in this study, we also tested our method for DF=2 and DF=8 cases.



**FIG 4.1** – Workflow of the proposed SPECT projection synthesis method. The training process (top) involves inputting 5-dimensional coordinates into the MLP, with a user-defined loss function guiding the network to learn from the patient-specific training targets: measured counts in sparse views. During testing (bottom), the trained network receives the coordinates of missing views and outputs the predicted counts.

#### 4.4.2.5 Training and Optimization

For each scan, we optimized the MLP weights by minimizing the Huber loss function ( $\delta=1$ ), which is given as

$$L_{\delta}(a) = \begin{cases} \frac{1}{2}a^2, & \text{if } |a| < \delta \\ \delta \left( |a| - \frac{1}{2}\delta \right), & \text{otherwise} \end{cases}. \quad (4.1)$$

We employed the ADAM optimizer [47] with an initial learning rate set at 0.001 and a reduce-on-plateau scheduler to minimize the loss function. We used coordinates corresponding to 20% of all pixels from the full projection views as per-patient validation data. The patient-specific model was selected at the lowest validation loss out of 200 training epochs. We used a batch size of 10,000 out of  $256 \times 256 \times n_{\text{bed}} \times n_{\text{view}}$  projection pixel coordinates.

#### 4.4.2.6 SPECT Reconstruction

In this study we performed OSEM SPECT reconstructions (DOTATATE data matrix size:  $128 \times 128 \times 79$  and 2-bed PSMA data matrix size:  $128 \times 128 \times 158$ , both with voxel size in mm:  $4.8 \times 4.8 \times 4.8$ ) with 6 subsets and 16 iterations using in-house software (available at: <https://github.com/JeffFessler/mirt>). No post-processing filter was applied. Scatter correction used a triple energy window method, while the depth-dependent attenuation correction used the standard CT-to-density calibration curve. The point spread function for depth-dependent collimator-detector response modeling was simulated with MC [66] using a point source in air and fitted with Gaussian curves.

#### 4.4.2.7 Evaluation

SPECT image quality was evaluated for four distinct reconstruction methods: 1) Full reconstruction using all 120 measured projections (full recon). 2) Partial reconstruction using a certain DF of the measured projections (partial recon). 3) Linear interpolation reconstruction, where a certain DF of projections were measured, and the remaining projections were generated through linear interpolation (LinInt recon). 4) NERF reconstruction, where a certain DF of projections were measured, and the remaining were MLP-predicted synthetic projections (NERF recon).

We quantified reconstruction performance using multiple evaluation metrics, including NRMSD, AR, ARNR, CNR and RCNR. In the phantom study, the uniform “warm” region served as the background (BKG). For the clinical patient study, we selected a homogeneous region within the lung as the BKG. The noise level was calculated as the standard deviation of voxel counts within this BKG, denoted as  $STD_{BKG}$ . These evaluations provide an assessment of the synthesized projection and reconstructed image compared to a reference image: the true activity map for phantom data and the OSEM reconstruction using all 120 measured projections (*i.e.*, full recon) for patient data. Definitions of the above metrics are given as follows:

$$\begin{aligned} AR &= \frac{\text{mean counts of reconstruction within VOI}}{\text{mean counts of true activity within VOI}}, & ARNR &= \frac{AR}{STD_{BKG}}, \\ CNR &= \frac{\text{mean of VOI} - \text{mean of BKG}}{STD_{BKG}}, & RCNR &= \frac{CNR_{\text{sparse view recon}}}{CNR_{\text{full recon}}} \times 100\%. \end{aligned}$$

### 4.4.3 Results

#### 4.4.3.1 Synthesized Projections

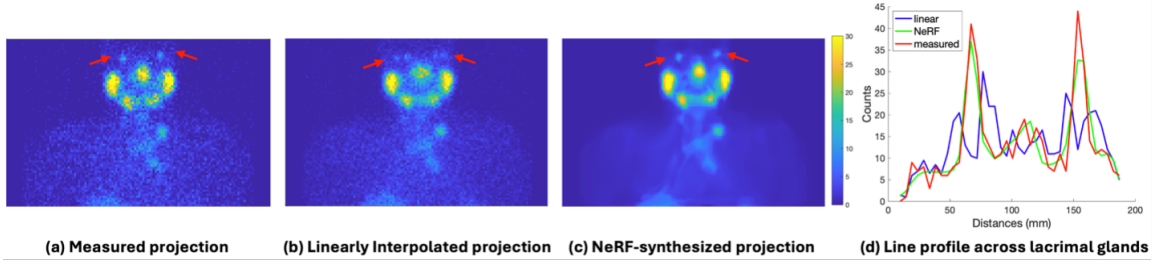
Table 4.1 compares the performance of linearly interpolated projections against NERF-synthesized projections, summarizing the NRMSD values across various DFs for phantom studies and patient studies. The results consistently demonstrate that the NERF-synthesized projections outperform linearly interpolated projections, exhibiting lower NRMSD values in both phantom and patient studies.

Visually, NERF-synthesized projections appear smoother than their measured counterparts. Fig. 4.2 displays the measured (Fig. 4.2 (a)) and synthesized projections for a representative PSMA patient. Close examination of the intensity profiles across the lacrimals reveals notable differences: the NERF-synthesized projection exhibits two peaks (corresponding to high uptake in left and right lacrimals as expected with PSMA), more closely

	Phantoms		DOTATATE Patients		PSMA Patients	
	NERF	Linear	NERF	Linear	NERF	Linear
DF=2	<b>5.9%</b>	9.0%	<b>16.9%</b>	23.4%	<b>17.5%</b>	24.6%
DF=4	<b>6.2%</b>	9.5%	<b>17.5%</b>	25.5%	<b>18.4%</b>	27.4%
DF=8	<b>7.5%</b>	11.1%	<b>18.8%</b>	30.4%	<b>23.7%</b>	34.1%

**TBL 4.1** – NRMSD (relative to measured projections) comparisons between NERF-synthesized projections and linearly interpolated projections across different DFs for phantom studies and patient studies (average across 11 DOTATATE studies and 6 PSMA studies).

aligning with the pattern observed in the measured projection, while the linearly interpolated projection presents four peaks due to angular interpolation.



**FIG 4.2** – Comparison of measured and synthesized projections for a patient following  $^{177}\text{Lu}$ -PSMA therapy. (a), (b), and (c) show measured projection, linearly interpolated projection, and NERF-synthesized projection, respectively. The images and profile comparison across lacrimal glands show two hot spots/peaks in the NERF synthesized projection (green line) corresponding to left and right lacrimals, closely resembling the profile of the measured projection (red line), whereas the corresponding results for the linear interpolation shows 4 peaks due to distortions.

**TBL 4.2** – Comparing synthesized projections using linear interpolation, supervised learning [82] and our proposed method. Results were based on 9 patient scans.

Method	Linear Interpolation	U-Net [82]	Ours
NRMSE (%)	$14.0 \pm 3.0$	$17.3 \pm 1.9$	<b><math>11.1 \pm 2.3</math></b>

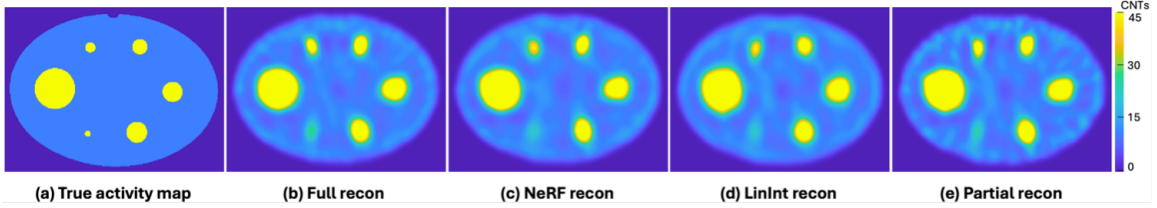
#### 4.4.3.2 Phantom Reconstruction Results

Consider the DF=4 scenario as an illustrative case. Fig. 4.3 compares four reconstructions with the true activity map. Although each reconstruction method exhibits structural similarities with the true activity, the partial recon is noticeably noisier than its counterparts.

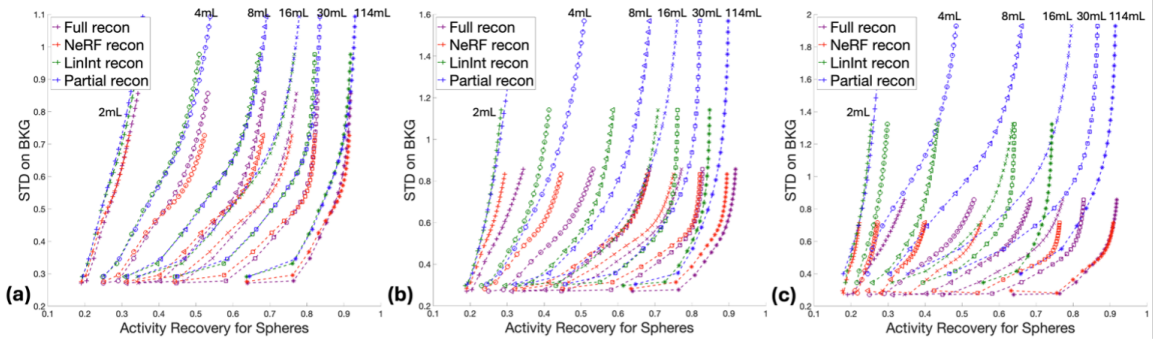


Quantitative comparisons, presented in Fig. 4.4, plot AR to noise curves for various spheres at  $DF=2,4$  and 8. Clearly, the **NeRF** recon outperforms both the partial recon and LinInt recon, delivering results that most closely parallel the full recon through various numbers of iterations of the **OSEM** algorithm. Note that even for the full reconstruction, AR is degraded ( $AR < 1$ ) because of partial volume effects [79].

Moreover, the noise level in all sparse-view reconstructions increases as the  $DF$  becomes larger. But the **NeRF** reconstruction consistently achieved highest activity recoveries for all six lesions at the same noise level. At  $DF=8$ , detailed in Fig. 4.4 (c), the partial reconstruction attained higher activity recovery for small lesions, at the expense of substantially increased noise level, while the **NeRF** reconstruction remains superior for larger lesions. For all sizes of lesions and  $DF$ s, the **NeRF** recon matched the activity recovery of the LinInt recon while maintaining a significantly lower noise level.



**FIG 4.3** – Visual comparison of (a) phantom true activity, (b) full recon, and (c) **NeRF** recon, (d) LinInt recon, (e) partial recon for  $DF=4$ . All images are in the same color scale.



**FIG 4.4** – AR to noise curves for sphere volumes ranging from 2 to 114 mL for the full recon and across  $DF$ s of 2, 4, and 8 (a to c). Distinct markers are consistently used to represent each sphere volume across all subfigures. The comparison illustrates the variations in AR and noise levels across four reconstruction methods: full recon, **NeRF** recon, LinInt recon, and partial recon, for different sphere sizes.



#### 4.4.3.3 Patient Reconstruction Results

Fig. 4.5 and Fig. 4.6 show the coronal Maximum Intensity Projections (MIPs) of an example patient image following DOTATATE and PSMA therapy, respectively, derived from four different reconstruction methods at various DFs. In both studies, the LinInt recons exhibit noticeable artifacts due to angular interpolation, more pronounced at higher DFs. This effect is particularly evident in the PSMA study for organs like the lacrimal, parotid, and submandibular glands at DF=4 and 8, substantially affecting the structural clarity of the SPECT images. Conversely, partial recons became noisier with increasing DFs, making it challenging to discern small hot spots from the background. However, the NERF recons retained a more accurate representation of activity distribution, closely resembling the full reconstructions, while maintaining a balanced noise level.

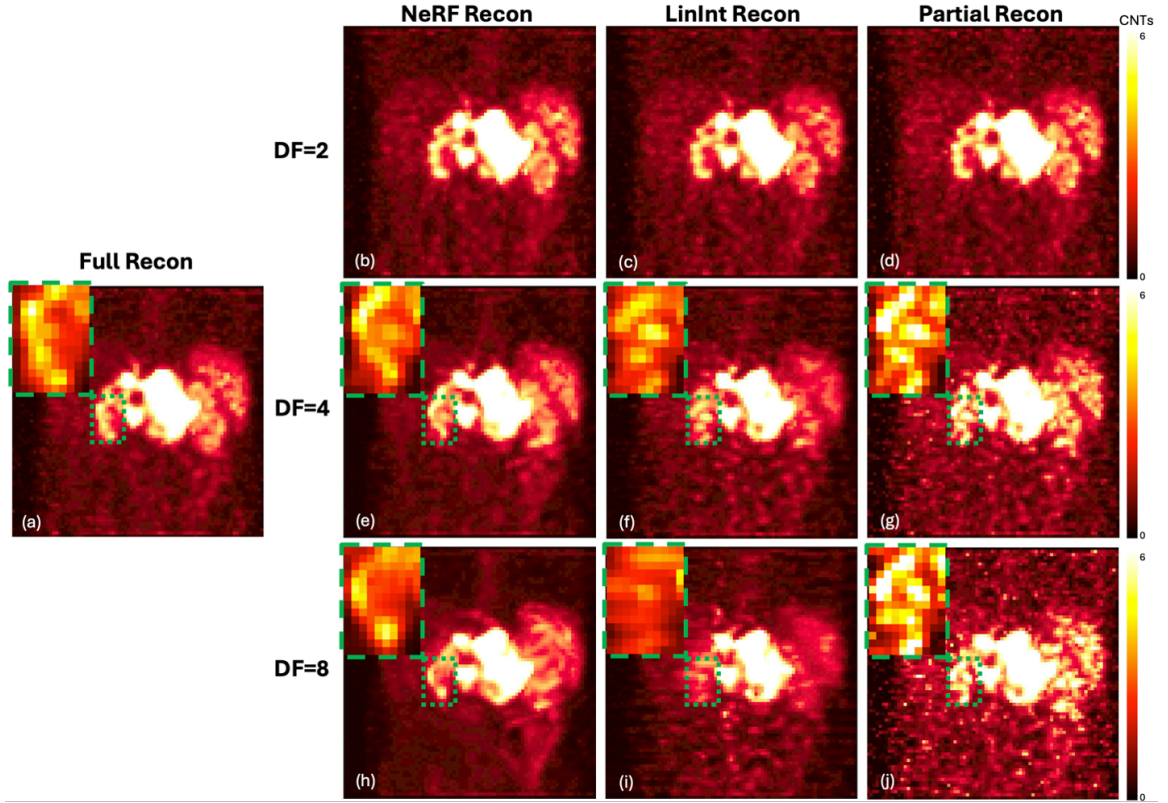
Quantitatively, the NERF recon yielded the highest average RCNR in the DOTATATE study, as shown in Table 4.3, for both lesion and kidney VOIs across all DFs. Similarly, in the PSMA study, the NERF reconstruction had higher average RCNR for all VOIs, as shown in Table 4.4, across all DFs. The limitation of LinInt recon is particularly evident in the lacrimal glands, which are of very small volume (about 0.4 mL) and exhibit exceptionally low RCNR values.

	DF=2			DF=4			DF=8		
	NERF Recon	LinInt Recon	Partial Recon	NERF Recon	LinInt Recon	Partial Recon	NERF Recon	LinInt Recon	Partial Recon
Lesion	<b>88.6%</b>	82.5%	82.7%	<b>87.9%</b>	68.7%	68.7%	<b>73.5%</b>	43.9%	48.2%
Kidney	<b>92.6%</b>	85.8%	84.5%	<b>88.0%</b>	73.1%	67.0%	<b>76.5%</b>	51.3%	48.8%

**TBL 4.3** – Average RCNR values of the NERF recon, the LinInt recon, and the partial recon across all eleven DOTATATE patient studies, benchmarked against the full recon, whose RCNR is standardized at 100%.

#### 4.4.4 Discussion

The field of machine learning, particularly in the domain of DL, is rapidly growing. Compared to other medical imaging modalities, DL applications to SPECT imaging are limited, perhaps due to the challenges of low-count scenarios of gamma-camera imaging. Previous works have demonstrated the effectiveness of using DL to generate missing SPECT projections views with convolutional neural networks, particularly, U-Net [81]. However, the data-intensive nature of supervised DL makes it less feasible for SPECT imaging, where datasets are usually limited, e.g., for our study, only tens of patient data are available, and

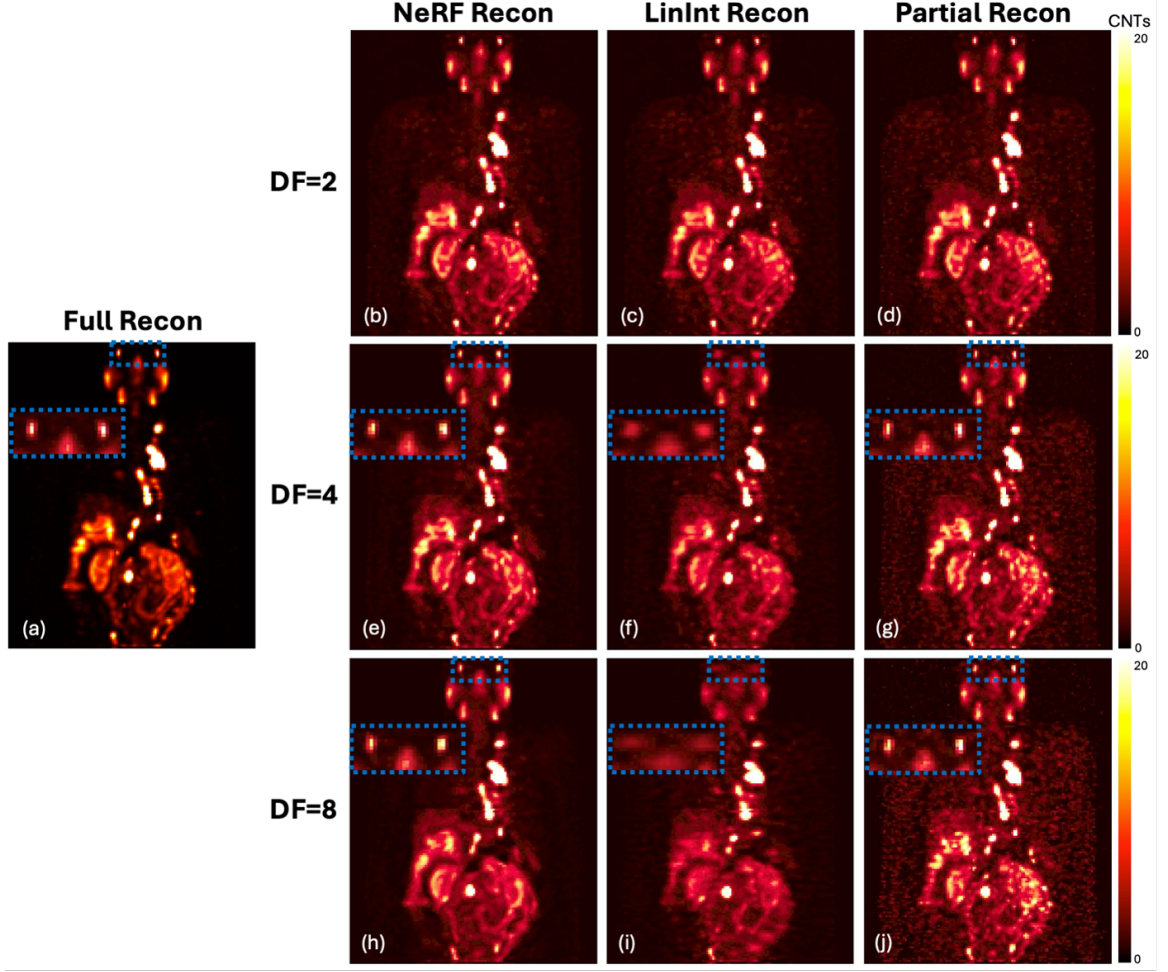


**FIG 4.5** – Coronal MIPs of SPECT reconstructions corresponding to a DOTATATE patient study using four reconstruction methods (columns) and three DFs (rows). Images are displayed with gamma correction with enhanced contrast levels to emphasize the blurring artifacts present in the LinInt recon and the noise present in the partial recon, especially visible at higher DFs.

	DF=2			DF=4			DF=8		
	NeRF Recon	LinInt Recon	Partial Recon	NeRF Recon	LinInt Recon	Partial Recon	NeRF Recon	LinInt Recon	Partial Recon
Lesion	<b>83.8%</b>	79.8%	80.7%	<b>78.4%</b>	70.7%	68.5%	<b>65.7%</b>	55.7%	54.9%
All Organ ROIs	<b>84.7%</b>	75.7%	80.9%	<b>78.4%</b>	56.9%	67.3%	<b>63.2%</b>	31.0%	50.8%
Kidney	<b>84.8%</b>	79.9%	80.3%	<b>80.1%</b>	69.6%	67.6%	<b>65.8%</b>	44.2%	51.3%
Lacrimal	<b>83.6%</b>	63.6%	80.4%	<b>77.5%</b>	29.9%	68.6%	<b>57.2%</b>	10.2%	47.9%
Parotid	<b>84.5%</b>	79.3%	80.9%	<b>79.1%</b>	63.4%	66.0%	<b>67.6%</b>	34.7%	52.0%
Submandibular	<b>85.6%</b>	79.7%	81.8%	<b>77.1%</b>	64.0%	66.9%	<b>62.2%</b>	34.6%	51.7%

**TBL 4.4** – Average RCNR values of the NeRF recon, the LinInt recon, and the partial recon across all six PSMA patient studies, benchmarked against the full recon, whose RCNR is standardized at 100%.

it would be difficult to obtain hundreds or thousands of patient datasets for applying supervised learning methods. Furthermore, the change of camera-specific parameters, for



**FIG 4.6** – Coronal MIPs of SPECT reconstructions corresponding to a PSMA patient study using four reconstruction methods (columns) and three DFs (rows). Images are displayed with gamma correction with enhanced contrast levels to emphasize the blurring artifacts present in the LinInt recon and the noise present in the partial recon, especially visible at higher DFs.

example, the crystal thickness of gamma-cameras and body contour orbits, may also influence the performance of supervised learning approaches. Unlike supervised learning, self-supervised learning methods derive insights directly from the current image itself without the need for labeled datasets, making them inherently adaptable and robust to variations in testing conditions. Thus, this thesis focused on a self-supervised learning method.

With evaluation both on phantoms that covered clinically relevant conditions and patients who underwent  $^{177}\text{Lu}$  DOTATATE and PSMA therapy in our clinic, we have demonstrated that our NERF recon, based on self-supervised coordinate-based learning, effectively compensates for image quality degradation under scenarios of sparse view acquisition. Considering both the reduction in acquisition time and quantitative accuracy/noise, a  $\text{DF}=4$ , appears to be a good compromise. In the patient studies, at a  $\text{DF}$  of 4, the NERF recon achieved CNRS of about 80% and higher for all organs and lesions while the other sparse view methods achieved only about 60 to 70% relative to the full reconstruction (Table 4.4, Table 4.3). Despite these promising outcomes, we observed reduced activity recovery in the NERF recon for smaller spheres ( $\leq 4$  mL) in the phantom reconstruction at higher  $\text{DF}$ s, compared to the other three reconstructions (Fig. 4.4). This limitation could arise from the neural network’s tendency to smooth over areas in low-count SPECT images due to high noise levels, leading to averaged voxel values from high noise variances. Despite the minor loss in recovery (also observed in the LinInt recons), the NERF recons show clearly improved CNR compared with the partial recons. Additionally, at a high  $\text{DF}$  of 8, the MLP faced challenge in accurately learning the representation of the continuous measurement field due to a substantial reduction in training data, particularly impacting finer textures that fluctuate in the measurement projections [97], contributing to reduced activity recovery in small lesions. Future research could explore the integration of variational inference or generative models to diversify the sampling process, potentially mitigating this smoothing effect and enhancing the model’s fidelity in capturing fine details.

In a previous study investigating DL to synthesize missing projection, MC-based reconstruction was used [82]. Although the attenuation, scatter, and collimator-detector response can be included simultaneously and accurately in the MC-based forward projection, this requires the simulation of a large number of photon histories, which is computationally expensive and therefore is less practical for routine clinical application. Instead, we used a reconstruction protocol similar to what is used in the clinic: a publicly available linear forward-backward system model with triple energy window scatter correction [24], which is a widely accepted and practical approach for  $^{177}\text{Lu}$  SPECT reconstruction.

The idea of NERF was to render photorealistic novel views of scenes with complicated geometries and appearances by representing a scene as a continuous function that outputs the radiance emitted in the coordinate space. To learn the continuous representation, a MLP is trained by inputting the coordinate of the scene and the training targets are the three-channel RGB colors. In this work, we conducted a similar training process where the targets were defined as single-channel SPECT projection counts. Moreover, the nature of coordinate-based learning works on the projection domain and hence is not restricted

to a specific image reconstruction method but is compatible with many methods including those based on MBIR or other methods such as plug-and-play (36) approaches. MBIR methods often handle a complete set of projection views but with fewer counts per view. Such methods can improve image quality and reduce noise by incorporating appropriate regularizers and priors; however, choosing the optimal regularizers and regularization parameters remains a challenge. In contrast, our method is tuning-free, as evident from the good performance in two different therapies where the activity distribution in the body is substantially different.

Although our research was initially focused on  $^{177}\text{Lu}$  SPECT imaging, we expect that our coordinate learning-based self-supervised method could be adapted for use in other low-count applications. This includes pure  $\beta^-$ -emitters, like  $^{90}\text{Y}$ , characterized by a low yield of bremsstrahlung photons for SPECT imaging [73], and  $\alpha$ -emitters, like Ac-225 that have low gamma-ray yields [26]. Both present inherent low-count imaging challenges that could potentially benefit from our approach. Furthermore, our method, which allows for skipping projection views could benefit diagnostic SPECT imaging by enabling administration of lower activities, therefore supporting low-dose SPECT protocols that reduce radiation exposure to patients with minimal compromise to image quality.

#### 4.4.5 Conclusion

This study addresses the challenge of extended SPECT imaging durations under low-count conditions, as encountered in  $^{177}\text{Lu}$  SPECT imaging, by developing a self-supervised coordinate learning approach that efficiently synthesizes skipped SPECT projection views without separate training data. The proposed method enables a significant reduction in SPECT acquisition time by allowing for skipping projection views and using a MLP to synthesize skipped projections, while preserving image quality, as indicated by improved NRMSD in projections, and ARNR and RCNR in reconstructions compared with other methods for sparse acquisitions. Unlike deep learning-based approach, this self-supervised method addresses the challenge of limited training data availability commonly encountered in clinical settings. The feasibility for reduction in acquisition time demonstrated in this work is particularly relevant for imaging under low-count conditions and for protocols that require multiple-bed positions.



## CHAPTER 5

### Discussion and Future Works

Poisson inverse problems are an intriguing and flexible category of mathematical and computational difficulties that have a wide range of applications in science and engineering. These problems revolve around the task of reconstructing unknown parameters or functions based on measured data that adhere to Poisson MLE. The allure of Poisson inverse problems lies in their capacity to reveal concealed information from noisy or incomplete observations, thus making them immensely valuable in various fields including medical imaging, environmental science, materials characterization, and astrophysics. This thesis focuses on two specific applications of Poisson inverse problems, namely phase retrieval and SPECT imaging. Thus far, we have presented several effective algorithms for resolving these types of Poisson inverse problems. This chapter discusses the challenges associated with these applications and explores potential future directions that can be investigated based on the research findings presented in this PhD thesis thus far.

#### 5.1 Learning on “SmArge” data

#### 5.2 Generative AI

#### 5.3 Denoising Projections with Unsupervised Learning

## CHAPTER 6

### Conclusion

This thesis demonstrates algorithms that aim to solve Poisson inverse problems in phase retrieval and SPECT imaging. For phase retrieval, our contributions are novel algorithms [60, 61, 59] that have faster convergence speed and lead to improved image reconstruction quality. For example, we propose modifications to the WF algorithm. Our method determines the step size based on observed Fisher information and incorporates a quadratic majorizer into our majorize-minimize approaches. We demonstrate that our methods are effective and exhibit favorable convergence properties [61]. Furthermore, we explore cases involving measurements affected by a combination of Poisson and Gaussian noise. We propose the use of an innovative technique called "AWFS" which uses accelerated WF with a score function as a generative prior. Theoretical analysis is conducted to showcase the critical point convergence guarantee of our algorithm. Simulation results demonstrate that our approach enhances reconstruction quality in terms of both visual perception and numerical assessment.

For SPECT imaging, we focus on DL solutions [56, 54, 62]. We develop a Julia toolbox [56] enables efficient modeling of SPECT forward-backward projectors with parallel computing and minimized memory allocations. This facilitates effective backpropagation during deep learning regularized iterative algorithm training, resulting in higher quality reconstructions compared to non-end-to-end methods. Moreover, we propose DblurDoseNet [54], a deep neural network for joint dosimetry estimation and image deblurring after SPECT reconstruction. It accurately estimates dose-rate distribution and compensates for SPECT resolution effects. Evaluations on phantoms and patients show that DblurDoseNet outperforms conventional dosimetry methods while being fast enough for real-time clinical use in radionuclide therapy dosimetry. Additionally, we propose a neural network with unsupervised learning to predict missing SPECT projections. Our method aims to decrease acquisition time by obtaining only a subset of all projections. Our method outperforms linear interpolation techniques used to predict missing projection views in terms of the achieved image reconstruction quality [62].

As mentioned in Chapter 5, there are several potential avenues for further research. These include investigating transfer learning techniques like finetuned `SAM` for tumor segmentation in `SPECT` images, exploring unsupervised methods for scatter correction in `SPECT` imaging, and incorporating `PET`-guided diffusion into the reconstruction of `SPECT` images. Similar methods can be employed to address 3D phase retrieval problems as well. We will be excited to see explorations on these research directions and believe they have the potential to improve the accuracy and efficiency of algorithms for solving Poisson inverse problems.



## **APPENDIX A**

### **Proof of the Proposed Improved Curvature Formula**

## APPENDIX B

# Uniform Cramér–Rao Lower Bound Analysis for Phase Retrieval Algorithms

This appendix<sup>1</sup> derives and analyzes the UCRLB for the phase retrieval problem, and then compares the bias-variance trade-off between phase retrieval algorithms (*e.g.*, Wirtinger flow, Gerchberg-Saxton, phaselift, MM and ADMM) that were derived from MLE where the measurements follow i.i.d. Gaussian distribution. We also consider regularizers that exploit the assumed properties of the latent signal, *e.g.*,  $\ell_2$  norm and  $\ell_1$  norm (approximated by the Huber function) that corresponds to the sparsity of finite differences (anisotropic TV) or of the detailed coefficients of a discrete wavelet transform. Simulation results show that many phase retrieval algorithms can be biased so that the classical CRLB fails to bound their variance. Regularized algorithms that better approximate the properties of the true signal have better bias-variance trade-offs (when compared to UCRLB) and lower reconstruction error.

---

<sup>1</sup>This work is based on [57].

## **APPENDIX C**

# **Proof of the Critical Point Convergence for the “AWFS” Algorithm**

This appendix proves the “AWFS” algorithm in Chapter 3 has a critical point convergence guarantee.

## BIBLIOGRAPHY

- [1] O. N. Aghakhan, A. Kamali-Asl, *et al.*. “Deep learning-based denoising of low-dose SPECT myocardial perfusion images: quantitative assessment and clinical performance”. In: *Eur. J. Nucl. Med. Mol. Imaging* 5 (2022). DOI: 10.1007/s00259-021-05614-7 (cit. on p. 24).
- [2] A. Aksac, D. Demetrick, T. Ozyer, *et al.*. “BreCaHAD: a dataset for breast cancer histopathological annotation and diagnosis”. In: *MC Res Notes* 12.82 (2019). DOI: <https://doi.org/10.1186/s13104-019-4121-7> (cit. on pp. 12–15, 17, 18, 21).
- [3] M. Asim, M. Daniels, O. Leong, A. Ahmed, and P. Hand. “Invertible generative models for inverse problems: mitigating representation error and dataset bias”. In: *Proceedings of the 37th International Conference on Machine Learning* 119 (2020) (cit. on p. 6).
- [4] D. A. Barmherzig and J. Sun. “Low-photon holographic phase retrieval”. In: *OSA: Computational Optical Sensing and Imaging*. 2020 (cit. on p. 5).
- [5] J. Batson and L. Royer. “Noise2Self: Blind denoising by self-supervision”. In: *Proceedings of the International Conference on International Conference on Machine Learning*. 2019 (cit. on pp. 6, 12, 18).
- [6] L. Bian, J. Suo, J. Chung, X. Ou, C. Yang, F. Chen, and Q. Dai. “Fourier ptychographic reconstruction using Poisson maximum likelihood and truncated Wirtinger gradient”. In: *Nature Sci. Rep.* 6.1 (2016). DOI: 10.1038/srep27384 (cit. on p. 5).
- [7] A. Bora, A. Jalal, E. Price, and A. G. Dimakis. “Compressed Sensing Using Generative Models”. In: *Proceedings of the International Conference on Machine Learning (ICML)*. Ed. by D. Precup and Y. W. Teh. Vol. 70. PMLR, Aug. 2017, pp. 537–546. DOI: 10.5555/3305381.3305437 (cit. on p. 6).
- [8] J.-F. Cai, M. Huang, D. Li, and Y. Wang. “Nearly optimal bounds for the global geometric landscape of phase retrieval”. In: *Inverse Problems* 39.7 (June 2023), p. 075011. DOI: 10.1088/1361-6420/acdab7 (cit. on p. 22).
- [9] T. T. Cai, X. Li, and Z. Ma. “Optimal Rates of Convergence for Noisy Sparse Phase Retrieval via Thresholded Wirtinger Flow”. In: *Annals Stat.* 44.5 (2016), pp. 2221–2251. DOI: 10.1214/16-AOS1443 (cit. on p. 5).
- [10] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski. “Phase retrieval via matrix completion”. In: *SIAM J. Imaging Sci.* 6.1 (2013), 199–225. DOI: 10.1137/110848074 (cit. on pp. 5, 18).
- [11] E. J. Candes, T. Strohmer, and V. Voroninski. “PhaseLift: exact and stable signal recovery from magnitude measurements via convex programming”. In: *Comm. Pure Appl. Math.* 66.8 (Aug. 2013), 1241–74. DOI: 10.1002/cpa.21432 (cit. on pp. 5, 18).

- [12] E. Candes, X. Li, and M. Soltanolkotabi. “Phase Retrieval via Wirtinger Flow: Theory and Algorithms”. In: *IEEE Trans. Info. Theory* 61.4 (Apr. 2015), pp. 1985–2007. DOI: 10.1109/TIT.2015.2399924 (cit. on pp. 5, 18).
- [13] S. H. Chan, X. Wang, and O. A. Elgendy. “Plug-and-Play ADMM for Image Restoration: Fixed-Point Convergence and Applications”. In: *IEEE Transactions on Computational Imaging* 3.1 (2017), pp. 84–98. DOI: 10.1109/TCI.2016.2629286 (cit. on p. 6).
- [14] H. Chang, Y. Lou, Y. Duan, and S. Marchesini. “Total variation-based phase retrieval for Poisson noise removal”. In: *SIAM J. Imag. Sci.* 11.1 (2018), pp. 24–55. DOI: 10.1137/16M1103270 (cit. on p. 5).
- [15] X. Chen and B. Zhou. “DuDoSS: Deep-learning-based dual-domain sinogram synthesis from sparsely sampled projections of cardiac SPECT”. In: *Med. Phys.* 1 (2023). DOI: 10.1002/mp.15958 (cit. on p. 24).
- [16] Y. Chen and E. J. Candes. “Solving random quadratic systems of equations is nearly as easy as solving linear systems”. In: *Comm. Pure Appl. Math.* 70.5 (May 2017), 822–83. DOI: 10.1002/cpa.21638 (cit. on p. 5).
- [17] A. Chicheportiche, S. Grozinsky-Glasberg, *et al.*. “Predictive power of the post-treatment scans after the initial or first two courses of [177Lu]-DOTA-TATE”. In: *EJNMMI Phys.* 1 (2018). DOI: 10.1186/s40658-018-0234-7 (cit. on p. 24).
- [18] E. Chouzenoux, A. Jezierska, J.-C. Pesquet, and H. Talbot. “A Convex Approach for Image Restoration with Exact Poisson–Gaussian Likelihood”. In: *SIAM Journal on Imaging Sciences* 8.4 (2015), pp. 2662–2682. DOI: 10.1137/15M1014395 (cit. on pp. 8, 10).
- [19] H. Chung, J. Kim, M. T. Mccann, M. L. Klasky, and J. C. Ye. “Diffusion Posterior Sampling for General Noisy Inverse Problems”. In: *The Eleventh International Conference on Learning Representations*. 2023 (cit. on pp. 7, 22).
- [20] H. Chung, B. Sim, D. Ryu, and J. C. Ye. “Improving Diffusion Models for Inverse Problems using Manifold Constraints”. In: *Advances in Neural Information Processing Systems*. Vol. 35. 2022, pp. 25683–25696 (cit. on p. 7).
- [21] H. Chung and J. C. Ye. “Score-based diffusion models for accelerated MRI”. In: *Medical Image Analysis* 80 (2022), p. 102479. ISSN: 1361-8415. DOI: <https://doi.org/10.1016/j.media.2022.102479> (cit. on p. 6).
- [22] Z.-X. Cui, C. Cao, S. Liu, Q. Zhu, J. Cheng, H. Wang, Y. Zhu, and D. Liang. *Self-Score: Self-Supervised Learning on Score-Based Models for MRI Reconstruction*. 2022. arXiv: 2209.00835 [eess.IV] (cit. on p. 6).
- [23] I. Daubechies. *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, 1992. DOI: 10.1137/1.9781611970104. eprint: <https://epubs.siam.org/doi/pdf/10.1137/1.9781611970104> (cit. on p. 5).
- [24] N. R. de, V. Lagerburg, T. Klausen, and S. Holm. “Improving quantitative dosimetry in (177)Lu-DOTATATE SPECT by energy window-based scatter corrections”. In: *Nucl. Med. Commun.* 35.5 (2014). DOI: 10.1097/MNM.0000000000000079 (cit. on p. 34).

- [25] A. Delker, W. Fendler, *et al.*. “Dosimetry for (177)Lu-DKFZ-PSMA-617: a new radiopharmaceutical for the treatment of metastatic prostate cancer”. In: *Eur. J. Nucl. Med. Mol. Imaging* 1 (2016). DOI: 10.1007/s00259-015-3174-7 (cit. on p. 24).
- [26] A. Delker, M. Schleske, *et al.*. “Biodistribution and dosimetry for combined [177Lu]Lu-PSMA-I&T/[225Ac]Ac-PSMA-I&T therapy using multi-isotope quantitative SPECT imaging”. In: *Eur. J. Nucl. Med. Mol. Imaging* 2 (2023). DOI: 10.1007/s00259-022-06092-1 (cit. on pp. 24, 35).
- [27] Y. Dewaraja, D. Mirando, A. Peterson, *et al.*. “A pipeline for automated voxel dosimetry: application in patients with multi-SPECT/CT imaging following 177Lu peptide receptor radionuclide therapy”. In: *J. Nucl. Med.* (2022). DOI: 10.2967/jnumed.121.263738 (cit. on p. 25).
- [28] P. Dhariwal and A. Nichol. “Diffusion Models Beat GANs on Image Synthesis”. In: *Advances in Neural Information Processing Systems*. Vol. 34. 2021, pp. 8780–8794 (cit. on p. 6).
- [29] G. Fatima and P. Babu. “PGPAL: A monotonic iterative algorithm for Phase-Retrieval under the presence of Poisson-Gaussian noise”. In: *IEEE Sig. Proc. Letters* (2022), pp. 1–1. DOI: 10.1109/LSP.2022.3143469 (cit. on p. 5).
- [30] G. Fatima, Z. Li, A. Arora, and P. Babu. “PDMM: A Novel Primal-Dual Majorization-Minimization Algorithm for Poisson Phase-Retrieval Problem”. In: *IEEE Transactions on Signal Processing* 70 (2022), pp. 1241–1255. DOI: 10.1109/TSP.2022.3156014 (cit. on p. 5).
- [31] J. Fienup. “Phase Retrieval Algorithms: A Comparison”. In: *Appl. Opt.* 21 (1982) (cit. on p. 18).
- [32] B. Gao and Z. Xu. “Phaseless recovery using the Gauss-Newton method”. In: *IEEE Trans. Sig. Proc.* 65.22 (Nov. 2017), 5885–96. DOI: 10.1109/tsp.2017.2742981 (cit. on p. 5).
- [33] R. W. Gerchberg and W. O. Saxton. “Practical Algorithm for Determination of Phase from Image and Diffraction Plane Pictures”. In: *OPTIK* 35.2 (1972), 237–246. ISSN: 0030-4026 (cit. on pp. 5, 18).
- [34] A. Gnanasambandam and S. H. Chan. “Image Classification in the Dark Using Quanta Image Sensors”. In: *ECCV 2020*. Cham: Springer International Publishing, 2020, pp. 484–501 (cit. on p. 5).
- [35] A. Goy, K. Arthur, S. Li, and G. Barbastathis. “Low photon count phase retrieval using deep learning”. In: *Phys. Rev. Lett.* 121.24 (Dec. 2018), p. 243902. DOI: 10.1103/PhysRevLett.121.243902 (cit. on p. 5).
- [36] A. Graikos, N. Malkin, N. Jojic, and D. Samaras. “Diffusion Models as Plug-and-Play Priors”. In: *Advances in Neural Information Processing Systems*. 2022 (cit. on pp. 6–8).
- [37] J. Ho, A. Jain, and P. Abbeel. “Denoising Diffusion Probabilistic Models”. In: 33 (2020). Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, pp. 6840–6851 (cit. on pp. 6, 12).
- [38] S. Hochreiter and J. Schmidhuber. “Long Short-Term Memory”. In: *Neural Computation* 9.8 (1997), pp. 1735–1780 (cit. on p. 24).
- [39] J. Hu, Z. Li, X. Xu, L. Shen, and J. A. Fessler. “Poisson-Gaussian Holographic Phase Retrieval with Score-based Image Prior”. In: *NeurIPS 2023 Workshop on Deep Learning and Inverse Problems*. 2023 (cit. on pp. 1, 5).
- [40] P. J. Huber. *Robust statistics*. New York: Wiley, 1981 (cit. on p. 12).

- [41] A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. Tamir. “Robust Compressed Sensing MRI with Deep Generative Priors”. In: *Advances in Neural Information Processing Systems*. Vol. 34. Curran Associates, Inc., 2021, pp. 14938–14954 (cit. on p. 6).
- [42] L. Ji and Z. Tie. “On gradient descent algorithm for generalized phase retrieval problem”. In: *2016 IEEE 13th International Conference on Signal Processing (ICSP)* (2016), pp. 320–325. DOI: 10.1109/ICSP.2016.7877848 (cit. on p. 5).
- [43] X. Jiang, S. Rajan, and X. Liu. “Wirtinger flow method with optimal stepsize for phase retrieval”. In: *IEEE Signal Proc. Letters* 23.11 (Nov. 2016), 1627–31. DOI: 10.1109/LSP.2016.2611940 (cit. on p. 5).
- [44] L. Kabasakal, T. Toklu, *et al.*. “Lu-177-PSMA-617 Prostate-Specific Membrane Antigen Inhibitor Therapy in Patients with Castration-Resistant Prostate Cancer: Stability, Bio-distribution and Dosimetry”. In: *Mol. Imaging Radionucl. Ther.* 2 (2017). DOI: 10.4274/mirt.08760 (cit. on p. 24).
- [45] U. S. Kamilov, H. Mansour, and B. Wohlberg. “A Plug-and-Play Priors Approach for Solving Non-linear Imaging Inverse Problems”. In: *IEEE Signal. Proc. Let.* 24.12 (Dec. 2017), pp. 1872–1876 (cit. on pp. 12, 18).
- [46] U. S. Kamilov, C. B. Bouman, G. T. Buzzard, and B. Wohlberg. “Plug-and-Play Methods for Integrating Physical and Learned Models in Computational Imaging”. In: *IEEE Signal Process. Mag.* 40.1 (Jan. 2023), pp. 85–97 (cit. on p. 6).
- [47] D. P. Kingma and J. Ba. “Adam: A Method for Stochastic Optimization”. In: *Proceedings of the International Conference on Learning Representations (ICLR)*. 2015. arXiv: 1412.6980 (cit. on p. 27).
- [48] D. Kingma and J. Ba. “Adam: A method for stochastic optimization”. In: *IEEE International Conference on Learning Representations (ICLR)*. 2015 (cit. on p. 13).
- [49] H. Lawrence, D. A. Barmherzig, H. Li, M. Eickenberg, and M. Gabriele. *Phase retrieval with holography and untrained priors: Tackling the challenges of low-photon nanoscale imaging*. To appear in JMLR. 2020 (cit. on pp. 5, 8, 12).
- [50] S. Lee, H. Chung, J. Kim, and J. C. Ye. *Progressive Deblurring of Diffusion Models for Coarse-to-Fine Image Synthesis*. 2022. arXiv: 2207.11192 [CS.CV] (cit. on p. 6).
- [51] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila. “Noise2Noise: Learning Image Restoration without Clean Data”. In: *Proceedings of the 35th International Conference on Machine Learning*. Vol. 80. Proceedings of Machine Learning Research. 2018, pp. 2965–2974 (cit. on p. 6).
- [52] H. Li and Z. Lin. “Accelerated Proximal Gradient Methods for Nonconvex Programming”. In: *Advances in Neural Information Processing Systems*. Vol. 28. 2015 (cit. on pp. 11, 19).
- [53] S. Li, W. Ye, and F. Li. “LU-Net: combining LSTM and U-Net for sinogram synthesis in sparse-view SPECT reconstruction”. In: *Math. Biosci. Eng.* 4 (2022). DOI: doi : 10.3934/mbe.2022200 (cit. on p. 24).
- [54] Z. Li, J. A. Fessler, J. K. Mikell, S. J. Wilderman, and Y. K. Dewaraja. “DblurDoseNet: A deep residual learning network for voxel radionuclide dosimetry compensating for SPECT imaging resolution”. In: *Med. Phys.* 49.2 (Feb. 2022), 1216–30. DOI: 10.1002/mp.15397 (cit. on pp. 2, 12, 23, 37).

- [55] Z. Li, Y. K. Dewaraja, and J. A. Fessler. “ESR-Net: An Efficient Image Super-resolution Network for SPECT Reconstruction”. In: *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2022, pp. 1–5. DOI: 10.1109/NSS/MIC44845.2022.10399172 (cit. on p. 23).
- [56] Z. Li, Y. K. Dewaraja, and J. A. Fessler. “Training End-to-End Unrolled Iterative Neural Networks for SPECT Image Reconstruction”. In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 7.4 (2023), pp. 410–420. DOI: 10.1109/TRPMS.2023.3240934 (cit. on pp. 2, 23, 37).
- [57] Z. Li and J. A. Fessler. “Uniform Cramér–Rao Lower Bound Analysis for Phase Retrieval Algorithms”. In: *2022 Asilomar Conference on Signals, Systems, and Computers*. 2022 (cit. on pp. 5, 40).
- [58] Z. Li, J. A. Fessler, J. K. Mikell, S. J. Wilderman, and Y. K. Dewaraja. “A Deep Residual Learning Network for Practical Voxel Dosimetry in Radionuclide Therapy”. In: *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2020, pp. 1–4. DOI: 10.1109/NSS/MIC42677.2020.9507764 (cit. on p. 2).
- [59] Z. Li, J. Hu, X. Xu, L. Shen, and J. A. Fessler. *Poisson-Gaussian Holographic Phase Retrieval with Score-based Image Prior*. 2023. arXiv: 2305.07712 (cit. on pp. 1, 5, 12, 37).
- [60] Z. Li, K. Lange, and J. A. Fessler. “Poisson Phase Retrieval With Wirtinger Flow”. In: *2021 IEEE International Conference on Image Processing (ICIP)*. 2021, pp. 2828–2832. DOI: 10.1109/ICIP42928.2021.9506139 (cit. on pp. 1, 5, 37).
- [61] Z. Li, K. Lange, and J. A. Fessler. “Poisson Phase Retrieval in Very Low-Count Regimes”. In: *IEEE Transactions on Computational Imaging* 8 (2022), pp. 838–850. DOI: 10.1109/TCI.2022.3209936 (cit. on pp. 1, 4, 5, 37).
- [62] Z. Li, X. Xu, J. Hu, J. A. Fessler, and Y. K. Dewaraja. “Reducing SPECT acquisition time by predicting missing projections with single-scan self-supervised coordinate-based learning”. In: *Journal of Nuclear Medicine* 64 (supplement 1) (cit. on pp. 2, 23, 37).
- [63] J. Liang, P. Stoica, Y. Jing, and J. Li. “Phase Retrieval via the Alternating Direction Method of Multipliers”. In: *IEEE Sig. Proc. Letters* 25.1 (2018), pp. 5–9. DOI: 10.1109/LSP.2017.2767826 (cit. on pp. 5, 18).
- [64] J. Liu, Y. Yang, M. Wernick, P. Pretorius, and M. King. “Deep learning with noise-to-noise training for denoising in SPECT myocardial perfusion imaging”. In: *Med. Phys.* 1 (2021). DOI: 10.1002/mp.14577 (cit. on p. 24).
- [65] Z. Liu, P. Luo, X. Wang, and X. Tang. “Deep Learning Face Attributes in the Wild”. In: *Proceedings of International Conference on Computer Vision (ICCV)*. Dec. 2015 (cit. on pp. 12–15, 17, 18, 21).
- [66] M. Ljungberg. “The SIMIND Monte Carlo Program”. In: (2012). DOI: 10.1201/b13073-8 (cit. on p. 27).
- [67] M. Ljungberg, A. Celler, *et al.*. “MIRD Pamphlet No. 26: Joint EANM/MIRD Guidelines for Quantitative  $^{177}\text{Lu}$  SPECT Applied for Dosimetry of Radiopharmaceutical Therapy”. In: *J. Nucl. Med.* 1 (2016). DOI: 10.2967/jnumed.115.159012 (cit. on p. 24).
- [68] Y. Long, J. A. Fessler, and J. M. Balter. “A 3D forward and back-projection method for X-ray CT using separable footprint”. In: *Proc. Intl. Mtg. on Fully 3D Image Recon. in Rad. and Nuc. Med.* Winner of poster award. 2009, 146–9 (cit. on p. 6).



- [69] W. Luo, W. Alghamdi, and Y. M. Lu. “Optimal Spectral Initialization for Signal Recovery With Applications to Phase Retrieval”. In: *IEEE Trans. Sig. Proc.* 67.9 (2019), pp. 2347–2356. DOI: 10.1109/TSP.2019.2904918 (cit. on p. 12).
- [70] M. Makitalo and A. Foi. “Optimal Inversion of the Generalized Anscombe Transformation for Poisson-Gaussian Noise”. In: *IEEE Trans. Imag. Proc.* 22.1 (2013), pp. 91–103. DOI: 10.1109/TIP.2012.2202675 (cit. on p. 5).
- [71] G. Marin, B. Vanderlinden, *et al.*. “A dosimetry procedure for organs-at-risk in  $^{177}\text{Lu}$  peptide receptor radionuclide therapy of patients with neuroendocrine tumours”. In: *Phys. Med.* (2018). DOI: 10.1016/j.ejmp.2018.11.001 (cit. on p. 24).
- [72] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng. “NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis”. In: *Commun. ACM* 65.1 (2021), pp. 99–106. ISSN: 0001-0782. DOI: 10.1145/3503250 (cit. on p. 26).
- [73] D. Minarik, G. K. Sjögren, and M. Ljungberg. “Evaluation of quantitative (90)Y SPECT based on experimental phantom studies”. In: *Phys. Med. Biol.* 20 (2008). DOI: 10.1088/0031-9155/53/20/008 (cit. on p. 35).
- [74] P. Netrapalli, P. Jain, and S. Sanghavi. “Phase Retrieval Using Alternating Minimization”. In: *IEEE Trans. Sig. Proc.* 63.18 (2015), pp. 4814–4826. DOI: 10.1109/TSP.2015.2448516 (cit. on p. 5).
- [75] H. Noubari, A. Fayazi, and F. Babapour. “De-noising of SPECT images via optimal thresholding by wavelets”. In: *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* (2009). DOI: 10.1109/IEMBS.2009.5332777 (cit. on p. 24).
- [76] G. Ongie, A. Jalal, C. A. Metzler, R. Baraniuk, A. G. Dimakis, and R. M. Willett. “Deep Learning Techniques for Inverse Problems in Imaging”. In: *IEEE Journal on Selected Areas in Information Theory* 1 (2020), pp. 39–56 (cit. on p. 6).
- [77] T. Qiu, P. Babu, and D. P. Palomar. “PRIME: phase retrieval via majorization-minimization”. In: *IEEE Trans. Sig. Proc.* 64.19 (Oct. 2016), 5174–86. DOI: 10.1109/TSP.2016.2585084 (cit. on pp. 5, 14, 18).
- [78] P. Ritt. “Recent Developments in SPECT/CT”. In: *Semin. Nucl. Med.* 3 (2022). DOI: 10.1053/j.semnuclmed.2022.01.004 (cit. on p. 23).
- [79] P. Ritt, H. Vija, J. Hornegger, and T. Kuwert. “Absolute quantification in SPECT”. In: *Eur. J. Nucl. Med. Mol. Imaging* (2011). DOI: 10.1007/s00259-011-1770-8 (cit. on p. 30).
- [80] Y. Romano, M. Elad, and P. Milanfar. “The Little Engine That Could: Regularization by Denoising (RED)”. In: *SIAM Journal on Imaging Sciences* 10.4 (2017), pp. 1804–44. DOI: 10.1137/16M1102884 (cit. on pp. 6, 12, 18).
- [81] O. Ronneberger, P. Fischer, and T. Brox. “U-net: convolutional networks for biomedical image segmentation”. In: *Medical Image Computing and Computer-Assisted Intervention*. 2015, 234–41. DOI: 10.1007/978-3-319-24574-4\_28 (cit. on p. 31).
- [82] T. Rydén, E. M. Van, I. Marin, J. Svensson, and P. Bernhardt. “Deep-Learning Generation of Synthetic Intermediate Projections Improves  $^{177}\text{Lu}$  SPECT Images Reconstructed with Sparsely Acquired Projections”. In: *J. Nucl. Med.* 62.4 (2021). DOI: 10.2967/jnumed.120.245548 (cit. on pp. 24, 29, 34).

- [83] M. Sandström, U. Garske, D. Granberg, A. Sundin, and H. Lundqvist. “Individualized dosimetry in patients undergoing therapy with  $(^{177}\text{Lu-DOTA-D-Phe (1)-Tyr (3)-octreotate})$ ”. In: *Eur. J. Nucl. Med. Mol. Imaging* 2 (2010). DOI: 10 . 1007 / s00259 - 009 - 1216 - 8 (cit. on p. 24).
- [84] M. Sandström, U. Garske-Román, *et al.*. “Individualized dosimetry of kidney and bone marrow in patients undergoing  $^{177}\text{Lu-DOTA-octreotate}$  treatment”. In: *J. Nucl. Med.* 1 (2013). DOI: 10 . 2967 / jnumed . 112 . 107524 (cit. on p. 24).
- [85] L. Santoro, E. Mora-Ramirez, *et al.*. “Implementation of patient dosimetry in the clinical practice after targeted radiotherapy using  $[^{177}\text{Lu-DOTAo, Tyr3]-octreotate}$ ”. In: *EJNMMI Res.* 1 (2018). DOI: 10 . 1186 / s13550 - 018 - 0459 - 4 (cit. on p. 24).
- [86] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev. “Phase Retrieval with Application to Optical Imaging: A contemporary overview”. In: *IEEE Sig. Proc. Mag.* 32.3 (2015), pp. 87–109. DOI: 10 . 1109 / MSP . 2014 . 2352673 (cit. on pp. 5, 18).
- [87] S. Shoushtari, J. Liu, and U. S. Kamilov. *DOLPH: Diffusion Models for Phase Retrieval*. 2022. arXiv: 2211.00529 [eess . IV] (cit. on pp. 6, 7, 12, 18, 20).
- [88] D. L. Snyder, A. M. Hammoud, and R. L. White. “Image recovery from data acquired with a charge-coupled-device camera”. In: *J. Opt. Soc. Am. A* 10.5 (May 1993), 1014–23. DOI: 10 . 1364 / JOSAA . 10 . 001014 (cit. on p. 5).
- [89] A. Sohlberg, T. Kangasmaa, C. Constable, and A. Tikkakoski. “Comparison of deep learning-based denoising methods in cardiac SPECT”. In: *EJNMMI Phys.* (2023). DOI: 10 . 1186 / s40658 - 023 - 00542 - x (cit. on p. 24).
- [90] M. Soltanolkotabi. “Structured Signal Recovery From Quadratic Measurements: Breaking Sample Complexity Barriers via Nonconvex Optimization”. In: *IEEE Trans. Info. Theory* 65.4 (2019), pp. 2374–2400. DOI: 10 . 1109 / TIT . 2019 . 2891653 (cit. on p. 5).
- [91] Y. Song and S. Ermon. “Generative Modeling by Estimating Gradients of the Data Distribution”. In: *Advances in Neural Information Processing Systems*. Vol. 32. 2019 (cit. on pp. 6, 7).
- [92] Y. Song, L. Shen, L. Xing, and S. Ermon. “Solving Inverse Problems in Medical Imaging with Score-Based Generative Models”. In: *International Conference on Learning Representations*. 2022 (cit. on pp. 6, 7).
- [93] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole. “Score-Based Generative Modeling through Stochastic Differential Equations”. In: *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. 2021 (cit. on pp. 6, 7, 12).
- [94] J. Sun, Y. Du, *et al.*. “Pix2Pix generative adversarial network for low dose myocardial perfusion SPECT denoising”. In: *Quant. Imaging Med. Surg.* 7 (2022). DOI: 10 . 21037 / qims - 21 - 1042 (cit. on p. 24).
- [95] J. Sun, H. Jiang, *et al.*. “Deep learning-based denoising in projection-domain and reconstruction-domain for low-dose myocardial perfusion SPECT”. In: *J. Nucl. Cardiol.* 3 (2023). DOI: doi : 10 . 1007 / s12350 - 022 - 03045 - x (cit. on p. 24).
- [96] J. Sun, Q. Qu, and J. Wright. “A geometric analysis of phase retrieval”. In: *2016 IEEE International Symposium on Information Theory (ISIT)*. 2016, pp. 2379–2383. DOI: 10 . 1109 / ISIT . 2016 . 7541725 (cit. on p. 5).

- [97] Y. Sun, J. Liu, M. Xie, B. Wohlberg, and U. S. Kamilov. “CoLL: Coordinate-Based Internal Learning for Tomographic Imaging”. In: *IEEE Transactions on Computational Imaging* 7 (2021), pp. 1400–1412. DOI: 10.1109/TCI.2021.3125564 (cit. on pp. 25, 34).
- [98] P. Thibault and M. Guizar-Sicairos. “Maximum-likelihood refinement for coherent diffractive imaging”. In: *New J. of Phys.* 14.6 (June 2012), p. 063004. DOI: 10.1088/1367-2630/14/6/063004 (cit. on p. 5).
- [99] X. Tian. “Fourier ptychographic reconstruction using mixed Gaussian-Poisson likelihood with total variation regularisation”. In: *Electronics Letters* 55.19 (Sept. 2019), 1041–3. DOI: 10.1049/el.2019.1141 (cit. on p. 5).
- [100] I. Vazquez, I. E. Harmon, J. C. R. Luna, and M. Das. “Quantitative phase retrieval with low photon counts using an energy resolving quantum detector”. In: *J. Opt. Soc. Am. A* 38.1 (Jan. 2021), 71–9. DOI: 10.1364/JOSAA.396717 (cit. on p. 5).
- [101] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg. “Plug-and-Play Priors for Model Based Reconstruction”. In: *Proc. IEEE Global Conf. Signal Process. and Inf. Process.* Austin, TX, USA, Dec. 2013, pp. 945–948 (cit. on pp. 12, 18).
- [102] P. Vincent. “A Connection Between Score Matching and Denoising Autoencoders”. In: *Neural Computation* 23.7 (2011), pp. 1661–1674. DOI: 10.1162/NECO\_a\_00142 (cit. on p. 7).
- [103] J. Violet, P. Jackson, *et al.*. “Dosimetry of <sup>177</sup>Lu-PSMA-617 in Metastatic Castration-Resistant Prostate Cancer: Correlations Between Pretherapeutic Imaging and Whole-Body Tumor Dosimetry with Treatment Outcomes”. In: *J. Nucl. Med.* 4 (). DOI: 10.2967/jnumed.118.219352 (cit. on p. 24).
- [104] I. Waldspurger. “Phase Retrieval With Random Gaussian Sensing Vectors by Alternating Projections”. In: *IEEE Trans. Info. Theory* 64.5 (2018), pp. 3301–3312. DOI: 10.1109/TIT.2018.2800663 (cit. on p. 5).
- [105] L. Wang. “The Poisson Channel With Varying Dark Current Known to the Transmitter”. In: *IEEE Transactions on Information Theory* 65.8 (2019), pp. 4966–4978. DOI: 10.1109/TIT.2019.2911474 (cit. on p. 5).
- [106] Y. Wang, W. Yin, and J. Zeng. “Global Convergence of ADMM in Nonconvex Nonsmooth Optimization”. In: *J. Sci. Comput.* 78.1 (Jan. 2019), pp. 29–63. ISSN: 0885-7474. DOI: 10.1007/s10915-018-0757-z (cit. on p. 21).
- [107] Z. Wang, J. Liu, G. Li, and H. Han. “Blind2Unblind: Self-Supervised Image Denoising with Visible Blind Spots”. In: *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2022, pp. 2017–2026. DOI: 10.1109/CVPR52688.2022.00207 (cit. on p. 6).
- [108] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. “Image quality assessment: from error visibility to structural similarity”. In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612. DOI: 10.1109/TIP.2003.819861 (cit. on p. 13).
- [109] X. Wei, H. van Gorp, L. Gonzalez-Carabarin, D. Freedman, Y. C. Eldar, and R. J. G. van Sloun. “Deep Unfolding With Normalizing Flow Priors for Inverse Problems”. In: *IEEE Transactions on Signal Processing* 70 (2022), pp. 2962–2971. DOI: 10.1109/TSP.2022.3179807 (cit. on p. 6).

- [110] Z. Wu, Y. Sun, J. Liu, and U. Kamilov. “Online Regularization by Denoising with Applications to Phase Retrieval”. In: *2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*. 2019, pp. 3887–3895. DOI: 10 . 1109 / ICCVW . 2019 . 00482 (cit. on p. 5).
- [111] R. Xu, M. Soltanolkotabi, J. P. Haldar, W. Unglaub, J. Zusman, A. F. J. Levi, and R. M. Leahy. *Accelerated Wirtinger flow: A fast algorithm for ptychography*. 2018 (cit. on p. 5).
- [112] X. Xu, J. Liu, Y. Sun, B. Wohlberg, and U. S. Kamilov. “Boosting the Performance of Plug-and-Play Priors via Denoiser Scaling”. In: *54th Asilomar Conf. on Signals, Systems, and Computers*. 2020, pp. 1305–1312. DOI: 10 . 1109 / IEEECONF51394 . 2020 . 9443410 (cit. on p. 13).
- [113] Y. Yang, M. Pesavento, Y. C. Eldar, and B. Ottersten. “Parallel Coordinate Descent Algorithms for Sparse Phase Retrieval”. In: *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019, pp. 7670–7674. DOI: 10 . 1109 / ICASSP . 2019 . 8683363 (cit. on p. 5).
- [114] Z. Yu, M. Rahman, *et al.*. “Need for objective task-based evaluation of deep learning-based denoising methods: A study in the context of myocardial perfusion SPECT”. In: *Med. Phys.* 7 (2023). DOI: 10 . 1002 / mp . 16407 (cit. on p. 24).
- [115] Z. Yuan and H. Wang. “Phase retrieval via reweighted Wirtinger flow”. In: *Appl. Opt.* 56.9 (Mar. 2017), pp. 2418–2427. DOI: 10 . 1364 / AO . 56 . 002418 (cit. on p. 18).
- [116] H. Zhang, Y. Liang, and Y. Chi. “A nonconvex approach for phase retrieval: Reshaped Wirtinger flow and incremental algorithms”. In: *J. Mach. Learning Res.* 18.141 (2017), 1–35 (cit. on p. 5).
- [117] K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, and R. Timofte. “Plug-and-Play Image Restoration With Deep Denoiser Prior”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44.10 (2022), pp. 6360–6376. DOI: 10 . 1109 / TPAMI . 2021 . 3088914 (cit. on p. 6).
- [118] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang. “Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising”. In: *IEEE Transactions on Image Processing* 26.7 (2017), pp. 3142–3155. DOI: 10 . 1109 / TIP . 2017 . 2662206 (cit. on p. 12).
- [119] Z. Zhuang, D. Yang, D. Barmherzig, and J. Sun. “Phase Retrieval Using Double Deep Image Priors”. In: *NeurIPS 2023 Workshop on Deep Learning and Inverse Problems*. 2023 (cit. on p. 22).