Denoising Particle Beam Micrographs with Plug-and-Play Methods

Minxu Peng, *Student Member, IEEE,* Ruangrawee Kitichotkul, Sheila W. Seidel, *Student Member, IEEE,* Christopher Yu, *Member, IEEE,* and Vivek K Goyal, *Fellow, IEEE*

***Abstract*—In a particle beam microscope, a raster-scanned focused beam of particles interacts with a sample to generate a secondary electron (SE) signal pixel by pixel. Conventionally formed micrographs are noisy because of limitations on acquisition time and dose. Recent work has shown that estimation methods applicable to a time-resolved measurement paradigm can greatly reduce noise, but these methods apply pixel by pixel without exploiting image structure. Raw SE count data can be modeled with a compound Poisson (Neyman Type A) likelihood, which implies data variance that is signal-dependent and greater than the variation in the underlying particle–sample interaction. These statistical properties make methods that assume additive white Gaussian noise ineffective. This paper introduces methods for particle beam micrograph denoising that use the plug-and-play framework to exploit image structure while being applicable to the unusual data likelihoods of this modality. Approximations of the data likelihood that vary in accuracy and computational complexity are combined with denoising by total variation regularization, BM3D, and DnCNN. Methods are provided for both conventional and time-resolved measurements, assuming SE counts are available. In simulations representative of helium ion microscopy and scanning electron microscopy, significant improvements in root mean-squared error (RMSE), structural similarity index measure (SSIM), and qualitative appearance are obtained. Average reductions in RMSE are by factors ranging from 2.24 to 4.11.**

***Index Terms*—electron microscopy, focused ion beam, helium ion microscopy, Neyman Type A distribution, Poisson processes**

I. Introduction

PARTICLE beam microscope (PBM) measurements contain information about topological and chemical composition of a sample. The scanning electron microscope (SEM), used widely in materials and life sciences, is a type of PBM that uses a focused electron beam to form micrographs at micro- and nanometer scales [1]. The less common helium ion microscope (HIM) [2] employs a beam of helium ions and offers larger depth of field and sub-nanometer resolution [3], [4]. The HIM has gained increasing popularity in semiconductor and biological imaging [5]–[8]. Focused beams of heavier ions, such as neon, gallium, and xenon, are also used for imaging.

This work was supported in part by a Draper Fellowship and by the US National Science Foundation under Grant No. 1815896 and Grant No. 2039762.

M. Peng, R. Kitichotkul, S. W. Seidel, and V. K. Goyal are with the Department of Electrical and Computer Engineering, Boston University, Boston, MA 02215 USA (mxpeng@bu.edu; rkitich@bu.edu; sseidel@bu.edu; v.goyal@ieee.org).

C. Yu is with Charles Stark Draper Laboratory, Cambridge, MA 02139 USA (cyu@draper.com).

Conventionally, acquiring accurate micrographs requires a high dose from a long dwell time or high beam current.[1] Desired acquisition speed is one impediment to high dose. Furthermore, high dose is especially problematic for HIM or microscopy with heavier ions, where sputtering and radiation damage are high and increase with dose [9]–[11]. Thus, SEM and HIM micrographs are inherently noisy. Computational noise reduction through post-processing that exploits assumed micrograph structure has received relatively little attention, perhaps because methods that neglect the unusual statistical properties of the measurements perform poorly.

Though not currently common in commercial PBMs, direct detection of secondary electrons (SEs) avoids the noise introduced by scintillators and photomultiplier tubes [12]. Thus, we assume direct SE detection here to characterize favorable performance. Another concept for improving the accuracy–dose trade-off is time-resolved (TR) measurement [13], [14] (see Section II-C); we consider both TR and non-TR measurements here. A concurrent line of research develops models and methods for PBM data without direct SE detection [15], [16]. Advantages from TR measurement persist when direct SE detection is not available [16]. Measurements in [16] also provide preliminary evidence to support the *Poisson–Poisson–Gaussian* model from [13], which builds directly upon the model used in this paper. The amount of data that can be collected with the oscilloscope-based setup described in [16] is limited by the on-board oscilloscope memory; it is not practical for collecting datasets for images with many pixels. The standard interface for a PBM produces 8-bit pixel values on an arbitrary scale that cannot be unambiguously mapped to SE counts. This data is thus not amenable to the goal of developing physics-based, quantitative image formation methods. For these reasons, the results in this paper are limited to an idealization of PBM data that emphasizes the main characteristic that distinguishes a PBM from an optical microscope or digital camera: the compound Poisson nature of the SEs.

In a PBM, the numbers of incident particles and the numbers of detected SEs per incident particle are both random. Each of these random numbers may be modeled as Poisson distributed, resulting in a Neyman Type A distribution for an SE measurement. This distribution results in signal-dependent noise variance that is greater than the variance under a Poisson model for the SE count. Using the distribution explicitly in

[1]Here, *dose* is defined as the mean number of incident particles per pixel, since the absolute spatial scale is not considered. More commonly, dose is the mean number of incident particles per unit area.

a model-based reconstruction method is difficult: to write the measurement likelihood with elementary functions requires an infinite series (see Section II-A), the negative log likelihood is not convex, and minimization problems involving the negative log likelihood do not generally have closed-form solutions.

The central goal of this paper is to provide methods to include PBM measurement distributions in model-based reconstruction. In particular, plug-and-play (PnP) methods have achieved state-of-the-art performance in many applications [17]–[20] and are gaining in popularity [21]. In the PnP alternating direction method of multipliers (ADMM) framework [22], one desires an efficient computation of the proximal operator of the data fidelity term. While no such algorithm is known for PBM data, we introduce approximate data fidelities based on Poisson distributions, which admit simple proximal operators similar to those used in PnP methods for Poisson models [23], [24]. Similarly, for a PnP fast iterative shrinkage-thresholding algorithm (FISTA), one desires an efficient computation of the gradient of the data fidelity term, and we provide this for TR measurements. These proximal operators and gradients may have other uses as well.

TR estimation techniques introduced in [13], [14] operate pixelwise. The only previous regularized reconstructions from TR measurements are restricted to total variation (TV) regularization [25], and they do not emphasize efficient implementations [26], [27]. In this work, we use ADMM or FISTA for different data fidelity terms depending on whether the proximal operator of the data fidelity term can be computed efficiently. In addition to denoising by TV regularization, we plug in the BM3D image denoiser [28] and a pre-trained deep neural network denoiser DnCNN [29].

A. Main Contributions

- *Introduction of data fidelity terms to make PnP methods efficiently applicable to denoising of particle beam micrographs*. Five data fidelity terms are presented: one for conventional (non-TR) measurements, one to study performance assuming an oracle provides the number of incident particles, and three for TR measurements.
- *Experimental evaluations of PnP methods in emulations of HIM and SEM*. We combined the five data fidelity terms with three denoisers: TV, BM3D, and DnCNN. The improvement from regularization is greater for HIM than for SEM and for conventional data than for TR data. RMSE reduction factors range from 2.24 for HIM with TR data to 4.11 for SEM with conventional data.

B. Outline

In Section II, we summarize the abstraction, measurement models and pixelwise estimators in PBM, and in Section III, we review the PnP ADMM and PnP FISTA frameworks. The key novelties of this paper are in Section IV, where we introduce data fidelity terms that allow the application of PnP methods to PBM denoising. The data fidelity terms have varying levels of complexity and accuracy to the negative log-likelihood function of the physical data generation process. Section V describes the collection of PnP algorithms we obtain by combining the data fidelity terms with three different denoisers. We present experimental results in both simulated HIM and SEM settings in Section VI. These show that time-resolved measurements and PnP methods provide significant improvements in estimation accuracy.

II. Measurement Models and Pixelwise Estimators

In this section, we introduce our measurement model for direct SE detection in PBM and several pixelwise estimators; see [13], [14] for additional details. This paper develops regularized estimators analogous to these pixelwise estimators. While the incident particles may be electrons or ions, we refer to them as ions for simplicity.

A. Abstract Model

A sample is raster scanned with a focused beam of ions. During a fixed dwell time t at any one raster scan location, the number of incident ions M is well modeled as a Poisson random variable with mean $\lambda = \Lambda t$, where Λ represents the known rate of incident ions per unit time.[2] Incident ion j interacts with the sample, generating X_j number of detected SEs. Each X_j can be described as a Poisson random variable with mean η, where η is called the SE yield. This is the physical quantity that we wish to estimate at each raster scan location to produce a pixel of the micrograph. Over the duration of the acquisition, the total detected SEs $Y = \sum_{j=1}^{M} X_j$ is a Neyman Type A random variable with probability mass function (PMF)

$$\mathrm{P}_Y(y;\eta,\lambda) = \frac{e^{-\lambda}\eta^y}{y!}\sum_{m=0}^{\infty}\frac{(\lambda e^{-\eta})^m m^y}{m!}, \qquad y = 0,\, 1,\, \ldots, \tag{1}$$

mean

$$\mathbb{E}[Y] = \lambda\eta, \tag{2}$$

and variance

$$\mathrm{Var}(Y) = \lambda\eta(\eta+1). \tag{3}$$

Notice the dependence of both the mean and variance on η and that this differs substantially from Poisson-distributed data, assuming η is not too small.

B. Conventional Measurement

The conventional measurement $\mathbf{y}^{\mathrm{C}} \in \mathbb{R}^d$ gathers measurements across all d pixels, with each entry drawn from the distribution in (1). Because the entries of $\mathbf{y}^{\mathrm{C}}$ are independent, its joint PMF is given by

$$\mathrm{P}_{\mathbf{y}^{\mathrm{C}}}(y^{\mathrm{C}}\,;\boldsymbol{\eta},\lambda) = \prod_{k=1}^{d}\mathrm{P}_Y\left(y_k^{\mathrm{C}}\,;\boldsymbol{\eta}_k,\lambda\right), \tag{4}$$

where $\boldsymbol{\eta}_k$ is the SE yield at the kth pixel and λ is the per-pixel dose.

[2] Although certain operating conditions may cause the beam current to stray from the desired setting due to contamination [30], we have shown that the unknown beam current can be estimated at each pixel using time-resolved measurements [27], [31]. Estimates from TR measurements also have an inherent insensitivity to knowledge of λ [32], [33].

C. Time-Resolved Measurement

With TR measurement, the per-pixel dwell time t is split into n sub-acquisitions, each of length t/n, yielding an n-length vector of measurements at each pixel. The vector $\mathbf{y} \in \mathbb{R}^{dn}$ gathers these TR measurements across all pixels with the measurement vector at the kth pixel given by $\mathbf{y}_k = [\mathbf{y}_k^{(1)}, \mathbf{y}_k^{(2)}, \ldots, \mathbf{y}_k^{(n)}]$. Each observation is sampled from the distribution in (1) with λ being replaced by λ/n. We note that at a given pixel, the conventional measurement may be obtained by summing the vector of time-resolved measurements:

$$\mathbf{y}_k^{\mathrm{C}} = \sum_{i=1}^{n} \mathbf{y}_k^{(i)}. \tag{5}$$

The entries in $\mathbf{y}$ are independent so its joint PMF is given by

$$\mathrm{P}_{\mathbf{y}}(\mathbf{y}\,;\,\boldsymbol{\eta},\lambda) = \prod_{k=1}^{d}\prod_{i=1}^{n} \mathrm{P}_Y\big(y_k^{(i)}\,;\,\boldsymbol{\eta}_k,\lambda/n\big). \tag{6}$$

D. Pixelwise Estimators

We now review estimation methods that operate at a single pixel without regularization. We include an oracle estimator that relies on knowledge of the number of incident ions M (which is *not* available in practice) and estimators that can be applied with conventional or TR data.

1) Conventional Estimator: From (2), scaling the observed SE counts at the kth pixel by λ yields an unbiased estimator:

$$\widehat{\boldsymbol{\eta}}_k^{\mathrm{conv}} = \mathbf{y}_k^{\mathrm{C}}/\lambda. \tag{7}$$

From (3), its mean-squared error (MSE) is

$$\mathrm{MSE}(\widehat{\boldsymbol{\eta}}_k^{\mathrm{conv}}) = \frac{\boldsymbol{\eta}_k(\boldsymbol{\eta}_k+1)}{\lambda}. \tag{8}$$

The $\boldsymbol{\eta}_k + 1$ factor in the MSE stems from the excess variance in (3) compared with the variance of a Poisson random variable with the same mean. This excess variance can be attributed to *source shot noise*—the randomness of the number of incident ions.

2) Oracle Estimator: If one were able to know the number of incident ions $\boldsymbol{M}_k$ at the kth pixel, dividing $\mathbf{y}_k^{\mathrm{C}}$ by $\boldsymbol{M}_k$ will produce a superior estimator

$$\widehat{\boldsymbol{\eta}}_k^{\mathrm{oracle}} = \mathbf{y}_k^{\mathrm{C}}/\boldsymbol{M}_k, \tag{9}$$

which has the MSE

$$\mathrm{MSE}(\widehat{\boldsymbol{\eta}}_k^{\mathrm{oracle}}) = e^{-\lambda}(1-e^{-\lambda})(\boldsymbol{\eta}_k-\eta_0)^2 + \sum_{m=1}^{\infty}\frac{\boldsymbol{\eta}_k}{m}\frac{\lambda^m}{m!}e^{-\lambda}, \tag{10}$$

where η_0 is the estimate assigned when $\boldsymbol{M}_k = 0$ [13]. For large enough λ, the choice of η_0 has little effect on the MSE. One can show [14, App. A] that for large λ, the MSE satisfies

$$\mathrm{MSE}(\widehat{\boldsymbol{\eta}}_k^{\mathrm{oracle}}) \approx \frac{\boldsymbol{\eta}_k}{\lambda}, \tag{11}$$

which eliminates the excess MSE due to the randomness of incident ions. We emphasize that such an estimator is unimplementable since $\boldsymbol{M}_k$ cannot be known exactly from only observing $\mathbf{y}_k^{\mathrm{C}}$.

3) Quotient Mode Estimator: To approach the performance of the oracle estimator, we may naturally seek a proxy for $\boldsymbol{M}_k$ that is computable from observed quantities. When n is large enough, the dose for each subacquisition λ/n becomes so small that the probability that more than one ion will arrive during one subacquisition is negligible. Assuming that η is large enough, most ions will produce at least one SE. In this case, the number of subacquisitions at the kth pixel measuring a positive number of SEs,

$$\boldsymbol{L}_k = \sum_{i=1}^{n} \mathbb{1}_{\{\mathbf{y}_k^{(i)}>0\}}, \tag{12}$$

is a good approximation for the number of incident ions $\boldsymbol{M}_k$, where $\mathbb{1}_{\{\mathbf{y}_k^{(i)}>0\}}$ is equal to 1 when $\{\mathbf{y}_k^{(i)} > 0\}$ and is equal to 0 otherwise. Analogous to the oracle estimator in (9), the *quotient mode* (QM) estimator is defined as

$$\widehat{\boldsymbol{\eta}}_k^{\mathrm{QM}} = \frac{\mathbf{y}_k^{\mathrm{C}}}{\boldsymbol{L}_k} = \frac{\sum_{i=1}^{n}\mathbf{y}_k^{(i)}}{\sum_{i=1}^{n}\mathbb{1}_{\{\mathbf{y}_k^{(i)}>0\}}}. \tag{13}$$

The QM name is taken from a similar concept proposed by John Notte in [34], where counting of the analog-domain pulses produced by SE bursts is adopted as the denominator of an estimator similar to (13). A closed-form expression for the MSE of $\widehat{\boldsymbol{\eta}}_k^{\mathrm{QM}}$ is given in [14]. The MSE of $\widehat{\boldsymbol{\eta}}_k^{\mathrm{QM}}$ is significantly lower than that of $\widehat{\boldsymbol{\eta}}_k^{\mathrm{conv}}$ except when η is small.

4) Lambert Quotient Mode Estimator: When η is small, $\boldsymbol{L}_k = \sum_{i=1}^{n}\mathbb{1}_{\{\mathbf{y}_k^{(i)}>0\}}$ significantly underestimates $\boldsymbol{M}_k$ because the probability of an ion generating zero detected SEs cannot be neglected. In this case, the bias in $\widehat{\boldsymbol{\eta}}^{\mathrm{QM}}$ caused by the underestimation can be reduced by replacing $\boldsymbol{L}_k$ with $(1-e^{-\boldsymbol{\eta}_k})^{-1}\boldsymbol{L}_k$. Since the probability of an incident ion resulting in at least 1 detected SE is $1-e^{-\boldsymbol{\eta}_k}$, the adjusted $(1-e^{-\boldsymbol{\eta}_k})^{-1}\sum_{k=1}^{n}\mathbb{1}_{\{\mathbf{y}_k^{(i)}>0\}}$ is a more accurate estimate of $\boldsymbol{M}_k$. Since $\boldsymbol{\eta}_k$ is unknown, the substitution results in a transcendental equation, which has the solution

$$\widehat{\boldsymbol{\eta}}_k^{\mathrm{LQM}} = W\Big(-\widehat{\boldsymbol{\eta}}_k^{\mathrm{QM}}e^{-\widehat{\boldsymbol{\eta}}_k^{\mathrm{QM}}}\Big) + \widehat{\boldsymbol{\eta}}_k^{\mathrm{QM}}, \tag{14}$$

where $W(\cdot)$ represents the Lambert W function [35]. Hence, we name $\widehat{\boldsymbol{\eta}}_k^{\mathrm{LQM}}$ the *Lambert quotient mode* (LQM) estimator. In MSE, $\widehat{\boldsymbol{\eta}}_k^{\mathrm{LQM}}$ significantly improves upon $\widehat{\boldsymbol{\eta}}^{\mathrm{QM}}$ when η is small and is nearly indistinguishable from $\widehat{\boldsymbol{\eta}}^{\mathrm{QM}}$ otherwise.

5) Time-Resolved Maximum Likelihood Estimator: With TR data at the kth pixel, $[\mathbf{y}_k^{(1)}, \mathbf{y}_k^{(2)}, \ldots, \mathbf{y}_k^{(n)}]$, the value of η that maximizes the joint likelihood is the *time-resolved maximum likelihood* (TRML) estimator:

$$\widehat{\boldsymbol{\eta}}_k^{\mathrm{TRML}} = \operatorname*{arg\,max}_{\eta\in[0,\infty)} \prod_{i=1}^{n} \mathrm{P}_Y(\mathbf{y}_k^{(i)};\eta,\lambda/n), \tag{15}$$

where $\mathrm{P}_Y(\cdot\,;\,\cdot,\cdot)$ is given by (1). According to [14, Fig. 5(a)], $\widehat{\boldsymbol{\eta}}_k^{\mathrm{TRML}}$ has lower MSE than $\widehat{\boldsymbol{\eta}}_k^{\mathrm{conv}}$, $\widehat{\boldsymbol{\eta}}_k^{\mathrm{QM}}$, and $\widehat{\boldsymbol{\eta}}_k^{\mathrm{LQM}}$ across all η values. However, since that domination does not hold for continuous-time observations (see [14, Fig. 3(a)]), in the discrete-time setting of the present work, the relative performances of estimators may depend on the choices of λ and n.

E. Comments on Pixelwise Estimator Performances

The estimators in Section II-D are analyzed and simulated in [14]. The key takeaway from Fisher information analyses is to expect[3]

$$\mathrm{MSE}(\widehat{\boldsymbol{\eta}}_k^{\mathrm{TRML}}) \approx \frac{\boldsymbol{\eta}_k(1-\eta e^{-\boldsymbol{\eta}_k})^{-1}}{\lambda}. \tag{16}$$

This is always better than the conventional estimator MSE (8). It is worse than the oracle estimator MSE (11) by a factor $(1-\boldsymbol{\eta}_k e^{-\boldsymbol{\eta}_k})^{-1}$, which is upper-bounded by $e/(e-1) \approx 1.58$ and approaches 1 for $\boldsymbol{\eta}_k \to 0$ and $\boldsymbol{\eta}_k \to \infty$. Thus, roughly speaking, TR measurement provides an MSE reduction by approximately a factor of $\boldsymbol{\eta}_k + 1$.

The gains from TR measurement in *unregularized* estimation have not previously been combined effectively with regularization.

III. Plug-and-Play Methods

In this section, we review ADMM, FISTA, and their uses in plug-and-play methods for image reconstruction incorporating image priors.

A. Problem Formulation

Given a vector of conventional measurements $\mathbf{y}^{\mathrm{C}} \in \mathbb{R}^d$ or TR measurements $\mathbf{y} \in \mathbb{R}^{dn}$, we seek to reconstruct the underlying SE yield image $\boldsymbol{\eta} \in \mathbb{R}^d$. This image reconstruction task may be written as an optimization problem of the form

$$\widehat{\boldsymbol{\eta}} = \arg\min_{\boldsymbol{\eta}} f(\boldsymbol{\eta}) + \beta g(\boldsymbol{\eta}), \tag{17}$$

where f is a data fidelity term that encourages consistency with $\mathbf{y}^{\mathrm{C}}$ or $\mathbf{y}$, g is a regularizer that promotes solutions with desirable properties, and β is a tuning parameter that controls the regularization strength. The regularizer g is generally non-smooth so solving the optimization problem in (17) is a non-trivial task.

From a Bayesian perspective, (17) can arise as the maximum a posteriori (MAP) estimator when

$$f(\boldsymbol{\eta}) = -\log p(\mathbf{y} \,|\, \boldsymbol{\eta}), \tag{18}$$

the negative log-likelihood of the observation $\mathbf{y}$ (or similarly with $\mathbf{y}^{\mathrm{C}}$); and

$$\beta g(\boldsymbol{\eta}) = -\log p(\boldsymbol{\eta}), \tag{19}$$

the negative log prior of $\boldsymbol{\eta}$.

B. ADMM

ADMM [36] converts the unconstrained problem in (17) into a constrained one:

$$(\widehat{\boldsymbol{\eta}}, \widehat{\boldsymbol{v}}) = \arg\min_{\boldsymbol{\eta}, \boldsymbol{v}} f(\boldsymbol{\eta}) + \beta g(\boldsymbol{v}), \quad \text{subject to } \boldsymbol{\eta} = \boldsymbol{v}. \tag{20}$$

This constrained problem can be solved by minimizing its augmented Lagrangian function

$$\mathcal{L}(\boldsymbol{\eta}, \boldsymbol{v}) = f(\boldsymbol{\eta}) + \beta g(\boldsymbol{v}) + \boldsymbol{u}^T(\boldsymbol{\eta} - \boldsymbol{v}) + \frac{\rho}{2}\|\boldsymbol{\eta} - \boldsymbol{v}\|^2, \tag{21}$$

where ρ is a penalty parameter and $\boldsymbol{u}$ is the Lagrange multiplier. The solution can be obtained by iterating the following steps:

$$\boldsymbol{\eta}^{(t+1)} = \arg\min_{\boldsymbol{\eta} \in \mathbb{R}^d} f(\boldsymbol{\eta}) + \frac{\rho}{2}\|\boldsymbol{\eta} - (\boldsymbol{v}^{(t)} - \boldsymbol{u}^{(t)})\|^2, \tag{22a}$$

$$\boldsymbol{v}^{(t+1)} = \arg\min_{\boldsymbol{v} \in \mathbb{R}^d} g(\boldsymbol{v}) + \frac{1}{2\sigma^2}\|\boldsymbol{v} - (\boldsymbol{\eta}^{(t+1)} + \boldsymbol{u}^{(t)})\|^2, \tag{22b}$$

$$\boldsymbol{u}^{(t+1)} = \boldsymbol{u}^{(t)} + (\boldsymbol{\eta}^{(t+1)} - \boldsymbol{v}^{(t+1)}), \tag{22c}$$

where $\sigma = \sqrt{\beta/\rho}$.

C. FISTA

The fast iterative shrinkage-thresholding algorithm (FISTA) [37] is another popular method to solve the optimization problem (17). FISTA iterates the following steps:

$$\boldsymbol{u}^{(t+1)} = \boldsymbol{v}^{(t)} - \gamma \nabla f(\boldsymbol{v}^{(t)}), \tag{23a}$$

$$\boldsymbol{\eta}^{(t+1)} = \arg\min_{\boldsymbol{\eta} \in \mathbb{R}^d} g(\boldsymbol{\eta}) + \frac{1}{2\sigma^2}\|\boldsymbol{\eta} - \boldsymbol{u}^{(t+1)}\|^2, \tag{23b}$$

$$q_{t+1} = \frac{1}{2}\Big(1 + \sqrt{1 + 4q_t^2}\Big),$$

$$\boldsymbol{v}^{(t+1)} = \boldsymbol{\eta}^{(t+1)} + \frac{q_t - 1}{q_{t+1}}(\boldsymbol{\eta}^{(t+1)} - \boldsymbol{\eta}^{(t)}), \tag{23c}$$

where γ is the step size, $\sigma = \sqrt{\gamma\beta}$, and $q_0 = 1$. In contrast to ADMM, FISTA only needs to evaluate the gradient $\nabla f(\boldsymbol{\eta})$, without needing to solve (22a). This is especially important when (22a) does not have a closed-form solution.

D. Plug-in Denoiser

ADMM and FISTA have similar decouplings into pairs of subproblems. Specifically, (22a) and (23a) can be viewed as *inversion* steps since $f(\boldsymbol{\eta})$ is determined by the forward image measurement model; and (22b) and (23b) can be viewed as *denoising* steps since $g(\boldsymbol{\eta})$ represents an image prior. In particular, (22b) is equivalent to Gaussian denoising on $\boldsymbol{\eta}^{(t+1)} + \boldsymbol{u}^{(t)}$ with noise level $\sigma = \sqrt{\beta/\rho}$. Based on the intuition that any other Gaussian denoiser could be used instead, Venkatakrishnan et al. [22] proposed the *PnP ADMM* algorithm, which does not specify g explicitly. Instead, (22b) is replaced with an off-the-shelf denoiser, denoted as $\mathcal{D}_\sigma$, to yield

$$\boldsymbol{v}^{(t+1)} = \mathcal{D}_\sigma(\boldsymbol{\eta}^{(t+1)} + \boldsymbol{u}^{(t)}). \tag{24}$$

Similarly, replacing (23b) with

$$\boldsymbol{\eta}^{(t+1)} = \mathcal{D}_\sigma(\boldsymbol{u}^{(t+1)}), \tag{25}$$

where $\sigma = \sqrt{\gamma\beta}$, results in a *PnP FISTA* [38].

IV. Data Fidelity Terms

In many applications of PnP methods, the data fidelity term $f(\boldsymbol{\eta})$ is derived from a measurement process that involves linear mixing and signal-independent additive white Gaussian noise (AWGN); (18) then results in $f(\boldsymbol{\eta}) \propto \|\mathbf{y} - A\boldsymbol{\eta}\|_2^2$ for some matrix A, which is quite convenient for computations.

[3] This is the reciprocal of the Fisher information for a *continuous-time* measurement [14, Eq. (26)] so it assumes an efficient estimator and $\lambda/n \ll 1$.

In this work, we wish to apply PnP methods to particle beam micrograph denoising, where the challenge is not rooted in linear mixing. Instead, challenges arise from the measurement likelihood function even though it is separable. To directly apply (18) with the measurement likelihood function (4) or (6) is problematic because of the form of the Neyman Type A PMF (1). This is an obstacle to regularized estimation of any form, whether or not one employs PnP methods.

This section introduces several data fidelity terms $f(\boldsymbol{\eta})$ that vary in their closeness to (18) and their computational complexity. For conventional measurement data, we have a data fidelity term based on a spatially adapted Gaussian approximation. For oracle or TR data, the data fidelity terms have correspondences with the oracle, QM, LQM, and TRML estimators of Section II. In cases in which (22a) has a closed-form solution, the data fidelity term becomes the basis for a PnP ADMM algorithm; for the remaining case, we provide an approximation to $\nabla f(\boldsymbol{\eta})$ so that the data fidelity term becomes the basis for a PnP FISTA method. These PnP methods are detailed in Section V.

A. Gaussian

We may approximate the entries of $\mathbf{y}^{\mathrm{C}}$ as independent Gaussian random variables, each with mean and variance given in (2) and (3):

$$\mathbf{y}_k^{\mathrm{C}} \sim \mathcal{N}(\lambda\boldsymbol{\eta}_k,\, \lambda\boldsymbol{\eta}_k(\boldsymbol{\eta}_k+1)). \tag{26}$$

Simple point evaluation of the Gaussian probability density function gives a reasonable approximation of the Neyman Type A probability mass function provided that $\lambda\boldsymbol{\eta}_k$ is not small, and there is pointwise convergence of the moment generating function for $\lambda \to \infty$ [39]. Omitting a constant term, the corresponding negative log-likelihood function is

$$f_{\mathrm{conv}}(\boldsymbol{\eta}) = \sum_{k=1}^{d} \left(\frac{1}{2}\log\boldsymbol{\eta}_k + \frac{1}{2}\log(\boldsymbol{\eta}_k+1) + \frac{(\mathbf{y}_k^{\mathrm{C}} - \lambda\boldsymbol{\eta}_k)^2}{2\lambda\boldsymbol{\eta}_k(\boldsymbol{\eta}_k+1)} \right). \tag{27}$$

To obtain a closed-form solution when we compute the inverse step (22a), we can further approximate (26) as

$$\mathbf{y}_k^{\mathrm{C}} \sim \mathcal{N}(\lambda\boldsymbol{\eta}_k,\, \lambda\boldsymbol{\eta}_k^{\mathrm{prev}}(\boldsymbol{\eta}_k^{\mathrm{prev}}+1)), \tag{28}$$

where the variance is set to be independent of $\boldsymbol{\eta}_k$ by using $\boldsymbol{\eta}_k^{\mathrm{prev}}$, an estimated value of $\boldsymbol{\eta}_k$ at the previous iteration. In this case, the observation $\boldsymbol{y}^{\mathrm{C}}$ becomes a Gaussian random vector with constant, diagonal covariance matrix, and the negative log-likelihood function is

$$f_{\mathrm{conv}}(\boldsymbol{\eta}) = \sum_{k=1}^{d} \frac{(\mathbf{y}_k^{\mathrm{C}} - \lambda\boldsymbol{\eta}_k)^2}{2\lambda\boldsymbol{\eta}_k^{\mathrm{prev}}(\boldsymbol{\eta}_k^{\mathrm{prev}}+1)} \tag{29}$$

after dropping terms that do not depend on $\boldsymbol{\eta}$. With this f, the ADMM inversion step (22a) is separable over the d components, with closed-form solution

$$\boldsymbol{\eta}_k^{(t+1)} = \left(\frac{\lambda}{\boldsymbol{\eta}_k^{\mathrm{prev}}(\boldsymbol{\eta}_k^{\mathrm{prev}}+1)} + \rho \right)^{-1} \cdot \left(\frac{\mathbf{y}_k^{\mathrm{C}}}{\boldsymbol{\eta}_k^{\mathrm{prev}}(\boldsymbol{\eta}_k^{\mathrm{prev}}+1)} + \rho(\boldsymbol{v}_k^{(t)} - \boldsymbol{u}_k^{(t)}) \right). \tag{30}$$

In the model (28), we are positing a variance for pixel k that gives spatially varying strength to the regularizer g. When the detected $\mathbf{y}_k^{\mathrm{C}}$ is unluckily small relative to a moderate or larger underlying true value η_k, estimation will perform poorly if too much confidence is ascribed to the observed $\mathbf{y}_k^{\mathrm{C}}$. To avoid this, performance is improved by allowing the local variance estimate $\lambda\boldsymbol{\eta}_k^{\mathrm{prev}}(\boldsymbol{\eta}_k^{\mathrm{prev}}+1)$ to depend on a neighborhood of pixel k rather than pixel k alone. Specifically, when computing $\boldsymbol{\eta}_k^{(t+1)}$, we choose $\boldsymbol{\eta}_k^{\mathrm{prev}}$ to be the average of $\boldsymbol{\eta}_k^{(t)}$ across a neighborhood of pixel k:

$$\boldsymbol{\eta}_k^{\mathrm{prev}} = \frac{1}{9}\sum_{\ell\in\mathcal{N}(k)} \boldsymbol{\eta}_\ell^{(t)}, \tag{31}$$

where $\mathcal{N}(k)$ is a 3×3 patch centered at pixel k. We chose the patch size to be 3×3 as using a 5×5 patch was empirically inferior.

B. Oracle

When the number of incident ions is exactly known, the number of SEs observed at the kth pixel may be modeled as a Poisson random variable with parameter $\boldsymbol{M}_k\boldsymbol{\eta}_k$:

$$\mathbf{y}_k^{\mathrm{C}} \sim \mathrm{Poisson}(\boldsymbol{M}_k\boldsymbol{\eta}_k). \tag{32}$$

The negative log-likelihood function is

$$f_{\mathrm{oracle}}(\boldsymbol{\eta}) = \sum_{k=1}^{d} \left(\boldsymbol{M}_k\boldsymbol{\eta}_k - \mathbf{y}_k^{\mathrm{C}}\log\boldsymbol{\eta}_k \right) \tag{33}$$

after dropping terms that do not depend on $\boldsymbol{\eta}$. Since f_{oracle} is convex, the minimum in (22a) is achieved when the gradient of the objective is zero. Hence, we obtain a closed-form update:

$$\boldsymbol{\eta}_k^{(t+1)} = -\frac{1}{2}\varphi + \frac{1}{2}\sqrt{\varphi^2 + \frac{4\mathbf{y}_k^{\mathrm{C}}}{\rho}} \quad \text{where } \varphi = \frac{\boldsymbol{M}_k}{\rho} - \boldsymbol{v}_k + \boldsymbol{u}_k. \tag{34}$$

Although instruments are not capable of measuring $\boldsymbol{M}_k$, the oracle data fidelity term serves as an interesting benchmark. We note that the forward model is reduced to sampling from a Poisson distribution. In fact, similar closed-form updates have been used in PnP ADMM methods for Poisson models, such as equation (16) in [24]. QM and LQM data fidelity terms, which we will discuss next, also follow Poisson distributions and thus have similar closed-form updates.

C. Quotient Mode

As in (13), our QM data fidelity term uses the number of subacquisitions where more that one SE was observed as a proxy for $\boldsymbol{M}_k$. Here we have

$$\mathbf{y}_k^{\mathrm{C}} \sim \mathrm{Poisson}(\boldsymbol{L}_k\boldsymbol{\eta}_k), \tag{35}$$

with $\boldsymbol{L}_k$ defined in (12). The negative log-likelihood function is

$$f_{\mathrm{QM}}(\boldsymbol{\eta}) = \sum_{k=1}^{d} \left(\boldsymbol{L}_k\boldsymbol{\eta}_k - \mathbf{y}_k^{\mathrm{C}}\log\boldsymbol{\eta}_k \right) \tag{36}$$

after dropping terms that do not depend on $\boldsymbol{\eta}$. Similar to the oracle case, since f_{QM} is convex, the ADMM inversion step (22a) again has a closed-form solution:

$$\boldsymbol{\eta}_k^{(t+1)} = -\frac{1}{2}\varphi + \frac{1}{2}\sqrt{\varphi^2 + \frac{4\mathbf{y}_k^{\mathrm{C}}}{\rho}} \quad \text{where } \varphi = \frac{\boldsymbol{L}_k}{\rho} - \boldsymbol{v}_k + \boldsymbol{u}_k. \tag{37}$$

D. Lambert Quotient Mode

Using the same adjustment as in (14) to compensate for the underestimate of $\boldsymbol{M}_k$ in $\boldsymbol{L}_k$, our LQM data fidelity term is based upon the model

$$\mathbf{y}_k^{\mathrm{C}} \sim \mathrm{Poisson}\big((1 - e^{-\boldsymbol{\eta}_k})^{-1}\boldsymbol{L}_k\boldsymbol{\eta}_k\big)\,. \tag{38}$$

The negative log-likelihood function after dropping terms that do not depend on $\boldsymbol{\eta}$ is

$$f_{\mathrm{LQM}}(\boldsymbol{\eta}) = \sum_{k=1}^{d}\left(\frac{\boldsymbol{\eta}_k}{1 - e^{-\boldsymbol{\eta}_k}}\boldsymbol{L}_k + \mathbf{y}_k^{\mathrm{C}}\log(1 - e^{-\boldsymbol{\eta}_k}) - \mathbf{y}_k^{\mathrm{C}}\log\boldsymbol{\eta}_k\right). \tag{39}$$

For faster computation, (38) may be approximated as

$$\mathbf{y}_k^{\mathrm{C}} \sim \mathrm{Poisson}((1 - e^{-\boldsymbol{\eta}_k^{\mathrm{prev}}})^{-1}\boldsymbol{L}_k\boldsymbol{\eta}_k), \tag{40}$$

where $\boldsymbol{\eta}^{\mathrm{prev}}$ is based on the previous iteration as in Section IV-A. Then

$$f_{\mathrm{LQM}}(\boldsymbol{\eta}) \approx \frac{\boldsymbol{L}_k\boldsymbol{\eta}_k}{1 - e^{-\boldsymbol{\eta}_k^{\mathrm{prev}}}} - \mathbf{y}_k^{\mathrm{C}}\log\boldsymbol{\eta}_k. \tag{41}$$

With this approximation, the ADMM inversion step (22a) has closed-form solution

$$\boldsymbol{\eta}_k^{(t+1)} = -\frac{1}{2}\varphi + \frac{1}{2}\sqrt{\varphi^2 + \frac{4\mathbf{y}_k^{\mathrm{C}}}{\rho}} \tag{42a}$$

where

$$\varphi = \frac{\boldsymbol{L}_k}{\rho(1 - e^{-\boldsymbol{\eta}_k^{\mathrm{prev}}})} - \boldsymbol{v}_k + \boldsymbol{u}_k. \tag{42b}$$

E. Time-Resolved Maximum Likelihood

In our TRML data fidelity term, we use the full likelihood in (6). Here we have

$$f_{\mathrm{TRML}}(\boldsymbol{\eta}) = -\log \mathrm{P}_{\mathbf{y}}(\mathbf{y}\,;\,\boldsymbol{\eta}, \lambda). \tag{43}$$

To use derivatives of (43) directly with the substitution of (1) and (6) is computationally expensive and delicate. However, as derived in [27, App. D] with the aid of Touchard polynomials, the derivatives of (43) with respect to the entries in $\boldsymbol{\eta}$ are approximately

$$\frac{\partial f_{\mathrm{TRML}}(\boldsymbol{\eta})}{\partial \boldsymbol{\eta}_k} \approx (n - \boldsymbol{L}_k)\frac{\lambda}{n}e^{-\boldsymbol{\eta}_k} - \frac{\mathbf{y}_k^{\mathrm{C}}}{\boldsymbol{\eta}_k} + \sum_{i\in\mathcal{S}}\frac{n + (2^{\mathbf{y}_k^{(i)}} - 1)\lambda e^{-\boldsymbol{\eta}_k}}{n + (2^{\mathbf{y}_k^{(i)}-1} - 1)\lambda e^{-\boldsymbol{\eta}_k}}, \tag{44}$$

where $\mathcal{S} = \{i \;:\; \mathbf{y}_k^{(i)} > 0\}$. The approximation is accurate when λ/n is small. The experimental results in Section VI are for $\lambda/n = 0.1$, which is small enough for accurate approximation and for most of the gains from TR measurement to be realized [14].

V. Proposed Methods

In this work, we test the five data fidelity terms proposed in Section IV within the PnP framework using three different denoisers: TV-regularized least squares, BM3D [28], and a deep neural network. This section explains the overall algorithm design and different types of denoising.

Algorithm 1 Plug-and-Play ADMM

Input: α, β, ρ, $\mathbf{y}^{\mathrm{C}}$ or $\mathbf{y} = [\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \ldots, \mathbf{y}^{(n)}]$
Initialize: $t = 0$, $\boldsymbol{\eta}^{(0)} = \boldsymbol{v}^{(0)} = \widehat{\boldsymbol{\eta}}^{\mathrm{conv}}$, $\boldsymbol{u}^{(0)} = \mathbf{0}$, $\sigma = \sqrt{\beta/\rho}$
1: **while** not converged **do**
2: $\quad\boldsymbol{\eta}^{(t+1)} = \underset{\boldsymbol{\eta}\in\mathbb{R}^d}{\arg\min}\, f(\boldsymbol{\eta}) + \frac{1}{2}\rho\|\boldsymbol{\eta} - (\boldsymbol{v}^{(t)} - \boldsymbol{u}^{(t)})\|^2$
3: $\quad\boldsymbol{v}^{(t+1)} = \mathcal{D}_\sigma(\boldsymbol{\eta}^{(t+1)} + \boldsymbol{u}^{(t)})$
4: $\quad\boldsymbol{u}^{(t+1)} = \boldsymbol{u}^{(t)} + (\boldsymbol{\eta}^{(t+1)} - \boldsymbol{v}^{(t+1)})$
5: $\quad t = t + 1$
6: **end while**
7: **return** $\boldsymbol{\eta}^{(t+1)}$

Algorithm 2 Plug-and-Play FISTA

Input: α, β, γ, ρ, $\mathbf{y}^{\mathrm{C}}$ or $\mathbf{y} = [\mathbf{y}^{(1)}, \mathbf{y}^{(2)}, \ldots, \mathbf{y}^{(n)}]$
Initialize: $t = 0$, $\boldsymbol{\eta}^{(0)} = \boldsymbol{v}^{(0)} = \widehat{\boldsymbol{\eta}}^{\mathrm{conv}}$, $q_0 = 1$, $\sigma = \sqrt{\gamma\beta}$
1: **while** not converged **do**
2: $\quad\boldsymbol{u}^{(t+1)} = \boldsymbol{v}^{(t)} - \gamma\nabla f(\boldsymbol{v}^{(t)})$
3: $\quad\boldsymbol{\eta}^{(t+1)} = \mathcal{D}_\sigma(\boldsymbol{u}^{(t+1)})$
4: $\quad q_{t+1} = \frac{1}{2}(1 + \sqrt{1 + 4q_t^2})$
5: $\quad\boldsymbol{v}^{(t+1)} = \boldsymbol{\eta}^{(t+1)} + ((q_t - 1)/q_{t+1})(\boldsymbol{\eta}^{(t+1)} - \boldsymbol{\eta}^{(t)})$
6: $\quad t = t + 1$
7: **end while**
8: **return** $\boldsymbol{\eta}^{(t+1)}$

A. Algorithm Overview

Algorithms 1 and 2 outline the key steps of PnP ADMM and PnP FISTA, as detailed in Section III, adapted to our setting. We apply Algorithm 2 with the TRML data fidelity term since (22a) does not have a closed-form solution in that case, and we apply Algorithm 1 to the other cases.

In Algorithm 1, Line 2, the *inversion step*, incorporates the data fidelity term. For the Gaussian, oracle, QM or LQM data fidelity term), it is computed with (30), (34), (37) or (42), respectively. Line 3 is the *denoising* step of (24). Line 4 updates the Lagrange multiplier.

In Algorithm 2, Line 2 makes $\boldsymbol{u}$ a step from $\boldsymbol{v}$ in the direction of the negative gradient of the data fidelity f. This is computed with (44) for the TRML data fidelity term. Line 3 applies a denoiser to remove noise in $\boldsymbol{u}$. Line 4 updates q_t, which controls the convergence rate. Line 5 uses q_t to update $\boldsymbol{v}$.

Both algorithms terminate when convergence is achieved. Here, convergence is declared when $\Delta_{t+1} \le \alpha$, where α is a threshold parameter. For PnP ADMM,

$$\Delta_{t+1} := \frac{1}{\sqrt{d}}\big(\|\boldsymbol{\eta}^{(t+1)} - \boldsymbol{\eta}^{(t)}\|_2 + \|\boldsymbol{v}^{(t+1)} - \boldsymbol{v}^{(t)}\|_2 + \|\boldsymbol{u}^{(t+1)} - \boldsymbol{u}^{(t)}\|_2\big). \tag{45}$$

For PnP FISTA,

$$\Delta_{t+1} := \frac{1}{\sqrt{d}} \|\boldsymbol{\eta}^{(t+1)} - \boldsymbol{\eta}^{(t)}\|_2. \tag{46}$$

If one wishes to constrain $\boldsymbol{\eta}$ to a convex set, it is natural to include a projection to that set within each iteration [23], [40]–[42]. Though the physical meaning of $\boldsymbol{\eta}$ implies nonnegativity of each entry, we have not imposed this explicitly because our experimental conditions keep all entries of the estimates nonnegative at all iterations.

B. Denoisers

We employ three denoising methods.

1) Total Variation: Under this formulation, the denoising step is regularized least-squares estimation (22b) with regularizer $g(\boldsymbol{\eta}) = \|\boldsymbol{\eta}\|_{\mathrm{TV}}$. The isotropic TV cost $\|\boldsymbol{\eta}\|_{\mathrm{TV}}$ is given by

$$\|\boldsymbol{\eta}\|_{\mathrm{TV}} = \|\boldsymbol{D}\boldsymbol{\eta}\|_2, \tag{47}$$

where $\boldsymbol{D}$ denotes the discrete image gradient. This cost term is designed to promote estimates that are piecewise smooth while maintaining edge features. The corresponding denoising step is solved iteratively as in [37].

2) BM3D: One branch of denoising algorithms exploits non-local similarity of image patches to recover a clean image from the noisy observation. BM3D [28] is one of the most widely used among these methods. It groups image patches based on similarity, then applies collaborative filtering and recombines to yield a reconstructed image.

3) Deep Neural Network: In recent years, deep learning-based methods have achieved great success in image denoising tasks. Most existing deep neural networks use a large number of clean–noisy image pairs as training samples, which is one key factor contributing to performance improvements. However, collecting a large dataset of clean–noisy image pairs can be expensive and challenging, especially in PBM. Without access to such a dataset, [26] used synthetic data, generated using accurate knowledge of the PBM forward model, to train a deep neural network. The network's weights were optimized to minimize the $\mathcal{L}_2$ difference between reconstructed and ground truth images. This method did not explicitly leverage the model information during inversion but instead relied on the PBM model to generate a noisy counterpart to each clean training image. In contrast, we use the PnP framework to combine knowledge of the PBM model with the assistance of a deep neural network to extract meaningful features, resulting in improved performance. In the denoising step (Line 3 of Algorithm 1), we adopt a deep neural network called DnCNN [29]. It combines residual learning [43] and batch normalization [44], and it has been demonstrated to be effective in removing AWGN.

In the formulation of ADMM in Section III-B, the Gaussian denoising problem (22b) has a noise standard deviation $\sigma = \sqrt{\beta/\rho}$ derived directly from the regularization parameter β and variable-splitting penalty ρ. Similarly, in PnP FISTA, the noise standard deviation $\sigma = \sqrt{\gamma\beta}$ depends on β and the step size γ. This dependence on σ can be troublesome in PnP methods because the denoiser may require training that depends on σ or it may have no analogous parameter.

Denoiser scaling [45] is a method for introducing tunable regularization to a denoiser trained for a specific noise standard deviation σ. Suppose only that some denoiser $\mathcal{D}$ is given. Then a scaled denoiser is defined as

$$\mathcal{D}_\mu(\boldsymbol{\eta}) \equiv (1/\mu)\mathcal{D}(\mu\boldsymbol{\eta}), \tag{48}$$

where $\mu > 0$ is the denoiser scaling parameter. This approach makes it possible to use a single pre-trained network across a variety of noise levels.

C. Convergence

When f and g are convex functions, both ADMM [36] and FISTA [37] will converge to the global minimum of the convex optimization problem (20) under some additional technical conditions. While a denoiser in PnP methods may not be the proximal operator of any convex function g, they can still have fixed point convergence. Specifically, PnP ADMM and PnP iterative shrinkage-thresholding algorithm (ISTA), i.e., FISTA without Nesterov acceleration, are shown to have fixed point convergence under appropriate choices of the parameters ρ and γ when f is strongly convex and the residue $\mathcal{D}_\sigma - \mathbf{I}$ has a sufficiently small Lipschitz constant, where $\mathbf{I}$ is the identity operator [46].

In our case, not all of the proposed data fidelity terms are convex with respect to $\boldsymbol{\eta}$. While f_{oracle} and f_{QM} are convex, f_{conv}, f_{LQM}, and f_{TRML} are not. We find that f_{LQM} is locally strongly convex for typical values of $\mathbf{y}$ and $\boldsymbol{\eta}$ in our experiments. Consequently, we can establish fixed point convergence of PnP ADMM and PnP ISTA for f_{oracle}, f_{QM}, and f_{LQM} under appropriate choices of ρ, γ, and β, and the denoiser $\mathcal{D}_\sigma$. We note that the TV denoiser is a proximal operator and that DnCNN typically has contractive residue even without further modification [46], so using these denoisers will not hinder fixed point convergence of PnP methods. However, BM3D has been shown to violate the residue contractiveness criterion [46], so the PnP methods that use BM3D may not converge, even if f is convex. We find that the theory outlined here predicts the empirical convergence of PnP ADMM and PnP FISTA well in our experiments.

VI. Experimental Results

In this section, we compare the performances of the proposed methods using simulated HIM and SEM micrograph images. While previous works [13], [14] have characterized the QM, LQM, and TRML estimators of Section II-D, here we aim to compare how these and the benchmark conventional and oracle estimators perform in combination with the three denoisers in Section V-B, within the PnP ADMM and PnP FISTA frameworks. The five estimators are included in the PnP methods through the data fidelity terms detailed in Section IV. Code is available online [47].

A. Datasets and Experiment Details

We use five crops of images from the "porous sponge" NFFA-EUROPE SEM dataset [48] as the ground truth images in our experiments. We scale the images to $\eta \in [2, 8]$ to emulate HIM [49] and to $\eta \in [1, 2]$ to emulate SEM. Given a ground truth image, we apply (6) to generate noisy TR measurements pseudorandomly using total dose $\lambda = 20$ split over $n = 200$ sub-acquisitions for each pixel for the HIM sample and total dose $\lambda = 50$ split over $n = 500$ sub-acquisitions per pixel for the SEM sample. The conventional measurement is obtained by summing over the n subacquisitions at each pixel. All PnP experiments use a stopping threshold of $\alpha = 5 \times 10^{-4}$ as defined in Section V-A. For PnP ADMM methods, we fix $\rho = 2.5$ for HIM and $\rho = 10$ for SEM. For PnP FISTA methods, we fix $\gamma = 0.1$ for HIM and $\gamma = 0.01$ for SEM. We tune σ (for TV and BM3D denoisers) and μ (for the DnCNN denoiser), which implies the β that controls the regularization strength, on a hold-out validation image. For HIM experiments, we use $\sigma \in \{0.24, 0.32, 0.72, 0.96, 1.2\}$ and $\mu \in \{0.5, 0.75, 1, 1.25, 1.5\}$. For SEM experiments, we use $\sigma \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$ and $\mu \in \{0.25, 0.5, 0.75, 1, 1.25, 1.5\}$. We use σ and μ values that yield the least root mean-squared error (RMSE) on the validation image for each combination of f and the denoiser in the corresponding experiments with the remaining 4 test images.

We train the DnCNN denoiser to minimize ℓ_2 loss using 500 natural images from the Berkeley Segmentation Dataset (BSD 500) [50]. The pixel values are scaled to $[0, 1]$, and the standard deviation of the additive white Gaussian noise is $25/255$. When applying DnCNN to an image with a different dynamic range, such as $\eta \in [2, 8]$ in the case of HIM experiments, we shift and scale the image so that the dynamic range of the input to DnCNN is $[0, 1]$. Then, we scale and shift the output back to the original dynamic range. Note that denoiser scaling as in (48) is performed on the image that is already transformed to the $[0, 1]$ dynamic range. In addition, DnCNN-related methods are GPU-accelerated for faster computation.

B. Reconstruction Quality Comparisons

We compare the reconstruction qualities of pixelwise estimations and proposed PnP methods with different data fidelity terms and denoisers. Figures 1 and 2 show reconstructed η micrographs under HIM and SEM settings, respectively. All micrographs are on the same scale shown in the ground truth image. The colorbars are chosen so that no more than 2% of pixels are saturated in any given image. Absolute error images $|\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}|$ for sub-regions marked by green boxes are displayed for each micrograph as well. We report RMSE and also the structural similarity index measure (SSIM) to quantify the perceived image quality.

Pixelwise estimators (top rows): To validate the use of regularization in general, we show the performance of pixelwise application of each of the five estimation methods (right five columns of the top row in Figures 1 and 2). Without regularization, the best TR methods significantly outperform the conventional method, as demonstrated in [14] as well. The RMSE reduction of the best implementable method of these—TRML—is by a factor of 2.1 for HIM and 3.6 for SEM; these factors are consistent with theoretical predictions (see Section II-E and [14]).

Data fidelity terms (comparisons by column): The bottom three rows in Figures 1 and 2 present results obtained by PnP methods. Moving from the second column to the rightmost column, each unregularized estimator is improved through regularization by comparing the top row with the bottom three rows in a single column for both HIM and SEM samples.

In comparing the methods using TR data and including regularization, it is uniformly true that the TRML data fidelity term gives the best performance, LQM second, and QM the worst. For SEM emulation, the QM-related PnP results suffer from large bias due to the large bias of the QM estimator when $\eta < 2$ [14]. Interestingly, the SSIM metric when using the QM data fidelity term is still decent, which is consistent with invariance properties of SSIM [51]. For HIM emulation, the performance of the Gaussian data fidelity term is similar to QM and worse than TRML by a factor of 1.4 in RMSE. For SEM emulation, the performance of the Gaussian data fidelity term is competitive with the best TR method, but still worse than TRML in terms of RMSE. However, in terms of SSIM, both TRML and LQM data fidelity terms outperform the Gaussian data fidelity term by a noticeable margin. These empirical results with regularization are consistent with theoretical results from [13], [14] that show increasing utility of TR measurements over conventional measurements as η increases.

Along with the combinations of data fidelity terms and denoisers yielding fifteen PnP methods (bottom three rows and right five columns in Figures 1 and 2), we compute three additional naïve estimates to demonstrate the virtue of accurate modeling of the acquisition process. The naïve estimators assume that $\mathbf{y}^{\mathrm{C}}/\lambda$ is Gaussian with mean $\boldsymbol{\eta}$ and constant variance (bottom three rows of the left column in Figures 1 and 2). These are computed by solving

$$\widehat{\boldsymbol{\eta}}^{\mathrm{naive}} = \underset{\boldsymbol{\eta} \in \mathbb{R}^d}{\arg\min} \left\| \boldsymbol{\eta} - \mathbf{y}^{\mathrm{C}}/\lambda \right\|_2^2 + \beta g(\boldsymbol{\eta}) \tag{49}$$

for each of the three regularizers g given explicitly or implicitly in Section V-B. The large reconstruction error from assuming a spatially invariant Gaussian distribution reinforces the significance of having accurate modeling of the acquisition process.

Denoisers (comparisons by row): In the HIM emulation, for each of the four PnP ADMM data fidelities (oracle, Gaussian, QM, LQM), the DnCNN denoiser and BM3D give nearly equal performance, better than TV. PnP FISTA with TRML data fidelity exhibits the best performance coupled with BM3D denoiser and nearly equal performance with TV denoiser; these are markedly better in the metrics and recovery of fine details compared to coupling with DnCNN denoiser.

Denoiser comparisons in SEM emulation have different trends. The impact of denoiser choice on RMSE is nearly negligible. We observe that BM3D and DnCNN reconstructions provide better visual reconstruction results than the TV denoiser as the image with TV denoiser displays patchy artifacts. Arguably, BM3D denoiser also results in oversmoothing. This

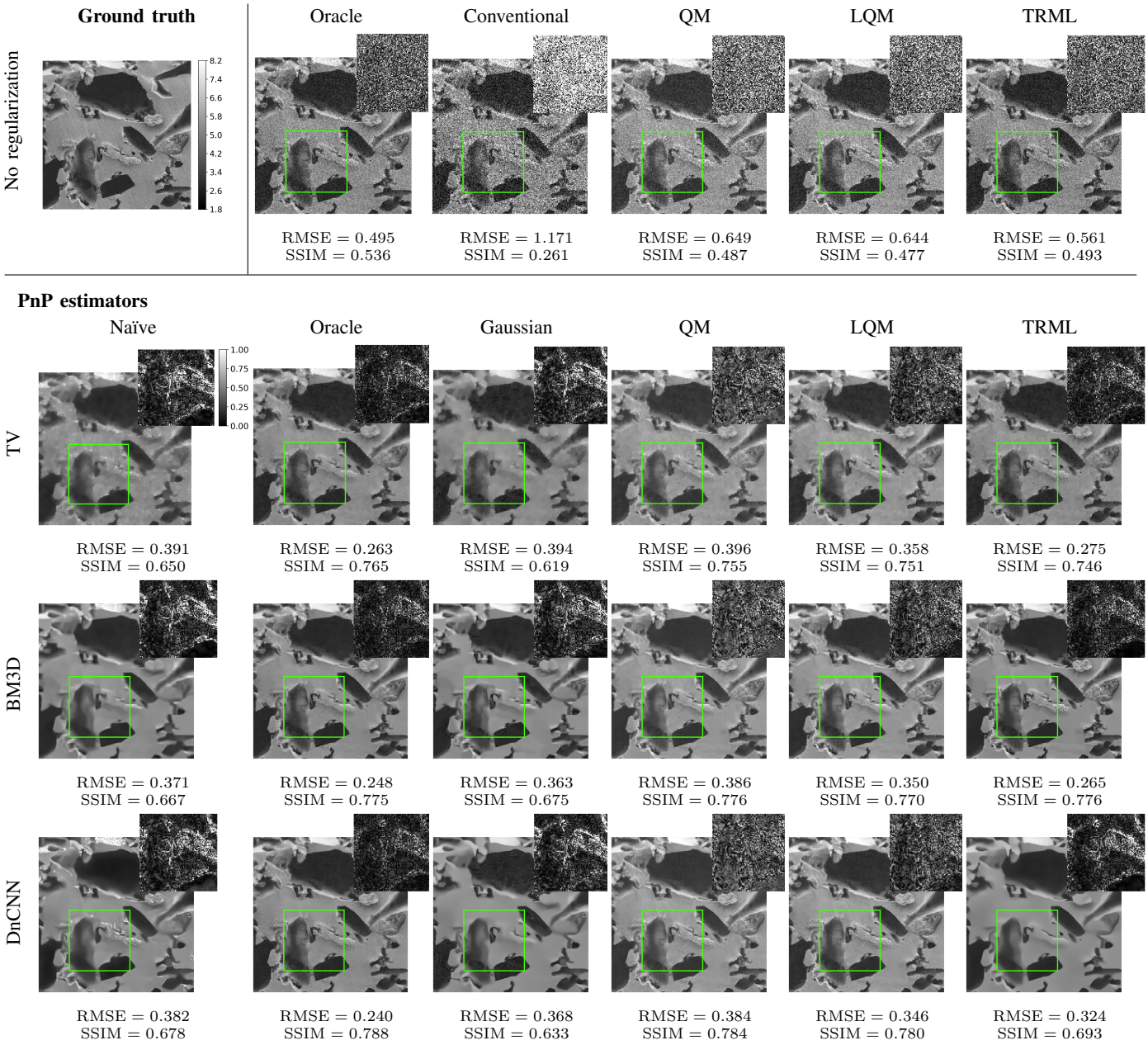


Fig. 1: HIM sample with ground truth $\eta \in [2, 8]$, total dose $\lambda = 20$, and $n = 200$. All images are shown on the same scale as in the ground truth image. Absolute error images for the sub-region in each green square are included for reconstructed micrographs.

may justify employing a deep network, which can represent more complex structural properties of the images.

C. Accuracy and Runtime Trade-Offs

Accuracy summary: Table I summarizes the average quantitative results from the test images. The *Regularized* columns present the best metrics over the choices of denoisers and data fidelity terms, excluding the oracle. The summary highlights the effectiveness of the PnP methods introduced in this paper for both HIM and SEM emulation and both conventional and time-resolved data. Comparing with pixelwise estimators across the four combinations in the table, regularization with PnP methods reduces RMSE by factors ranging from 2.24 to 4.11. Comparing the regularized results using TR data against using conventional data, the RMSE is reduced by a factor of 1.35 for HIM emulation and 1.10 for SEM emulation. This shows that the improvement from regularization is at least partially complementary to the improvement from the use of TR data. Also, the lesser improvement from TR data for SEM emulation is consistent with earlier results [13], [14].

For each of the 4 test images, we repeat Monte Carlo emulations of HIM and SEM with 10 different random seeds. For the HIM emulations, for estimators with regularization, the standard deviation of RMSE of $\widehat{\boldsymbol{\eta}}$ relative to the mean RMSE was 5.1% on average, with the highest not exceeding 10%. The relative standard deviation of RMSE for the conventional estimator without regularization is only slightly lower at 4.1%, suggesting reasonably robust convergence of the PnP methods.

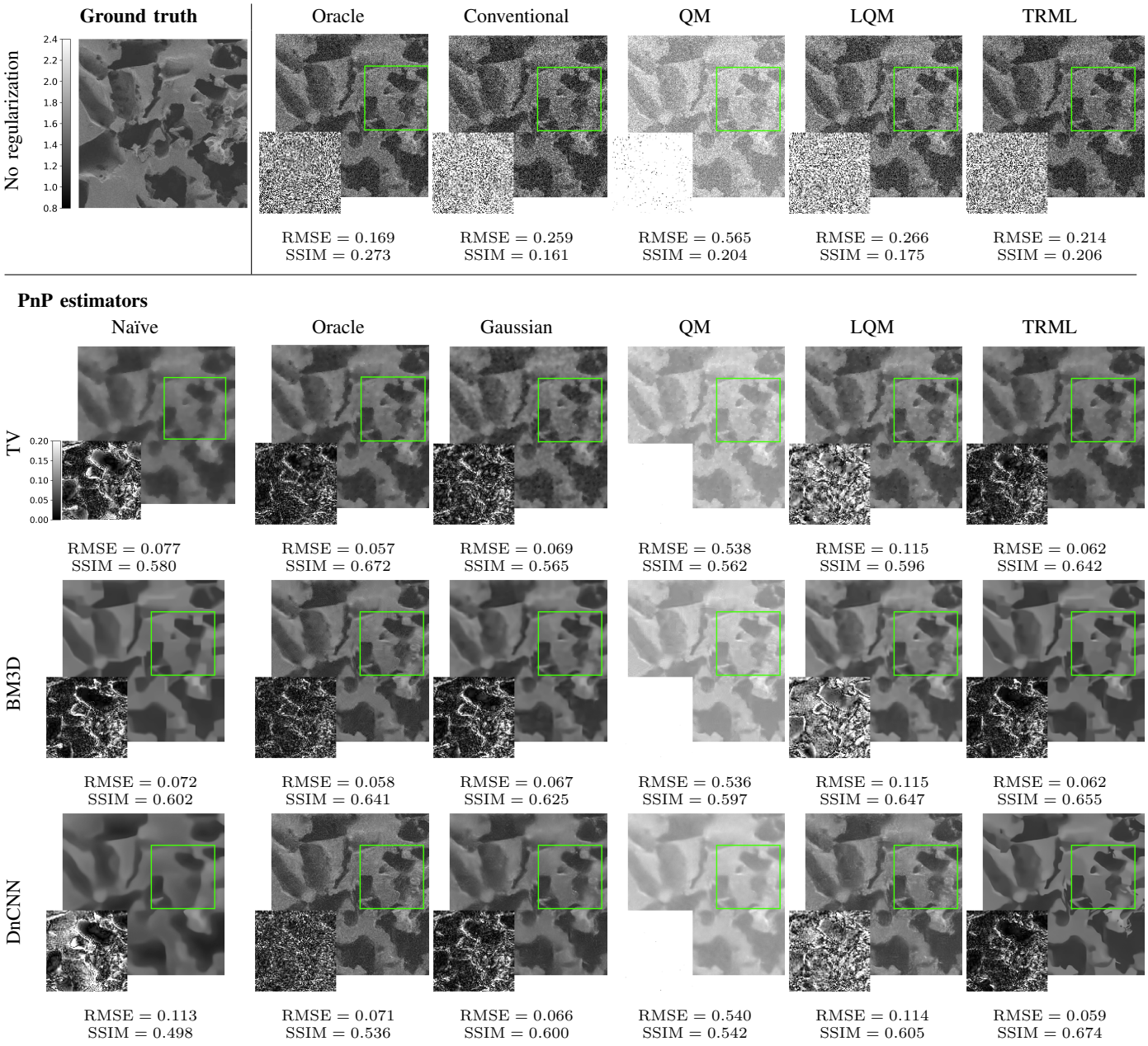


Fig. 2: SEM sample with ground truth $\eta \in [1, 2]$, total dose $\lambda = 50$ and $n = 500$. All images are shown on the same scale as in the ground truth image. Absolute error images for the sub-region in each green square are included for reconstructed micrographs.

TABLE I: Summary of quantitative accuracy metrics.

		Conventional data		Time-resolved data	
		Pixelwise	Regularized	Pixelwise	Regularized
HIM	RMSE	1.173	0.339	0.561	0.251
	SSIM	0.225	0.643	0.448	0.741
SEM	RMSE	0.267	0.065	0.218	0.059
	SSIM	0.147	0.563	0.190	0.626

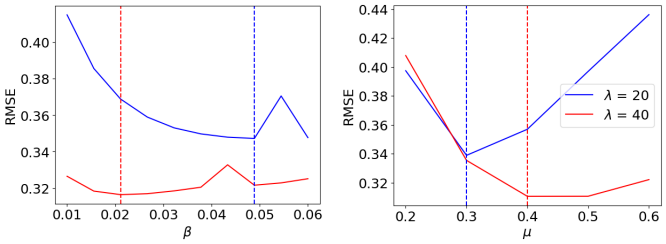
In addition, the performance ranking stays unchanged in each simulation, indicating a fairly robust ranking of methods for the HIM case. The variation for SEM emulations is higher, suggesting that one should consider some of the performances in Figure 2 virtually indistinguishable. The greater variation in SEM emulations may be attributable to higher noise level in the data (the SSIM value for the oracle estimator without regularization is significantly higher in Figure 1 than that in Figure 2) and convergence challenges (see Section VI-E).

Computation times: Table II shows the average runtimes of Algorithms 1 and 2 across the test images for different applicable data fidelity terms and denoisers in HIM emulation.

The runtimes for all PnP ADMM methods are similar across different data fidelity terms while having much more dependence on the choice of denoiser; using DnCNN is the fastest, TV about 10 times slower, and BM3D about another 60 times slower. Considering that none of the PnP ADMM methods with BM3D give appreciably lower RMSE or higher

TABLE II: Runtimes in seconds for all combinations of data fidelity term and denoiser. PnP FISTA is used for the TRML data fidelity term, while PnP ADMM is used for all other data fidelity terms.

		Data fidelity				
		Gaussian	Oracle	QM	LQM	TRML
Denoiser	TV	0.23	2.86	0.68	1.42	46.50
	BM3D	84.82	85.25	85.13	84.88	78.98
	DnCNN	0.14	0.17	0.14	0.13	26.14



(a) TV denoiser (b) DnCNN denoiser

Fig. 3: Relationship between RMSE of the reconstruction and the regularization parameters, β for TV and μ for DnCNN, at $\lambda = 20$ and 40. The vertical lines indicate the regularization parameter value when the RMSE is minimum for the corresponding λ.

SSIM than the corresponding methods with DnCNN, BM3D is generally unattractive in PnP ADMM.

Using PnP FISTA with the TRML data fidelity term, which significantly improves reconstruction quality over the PnP ADMM methods in most cases, the runtimes for TV and DnCNN denoisers increase by approximately 40 and 180 times, respectively, compared to their PnP ADMM counterparts. On the contrary, the runtime for PnP FISTA using BM3D decreases slightly. Comparing the runtimes of PnP FISTA using different denoisers, the runtime when using BM3D is only 1.7 times longer than that of TV and 3 times longer than that of DnCNN. Consequently, BM3D becomes an attractive option in PnP FISTA, especially when it also gives better reconstructions, such as in the case of HIM emulation as shown in Figure 1.

D. Noise Level and Optimal Regularization Strength

Reconstructions from different pixelwise estimators have different expected MSE depending on λ and η as explained in Sections II-D and II-E. However, it is unclear how this effective noise level determines the optimal choice of the regularization strength β in (20). In our experiments, we tune the regularization parameters σ and μ, both of which imply β, on a hold-out validation image. Still, the question of how λ influences the optimal choice of β remains.

To study the relationship between λ and the optimal regularization strength, we simulated two measurements of the hold-out validation image: one with $\lambda = 20$ and $n = 200$ and the other with $\lambda = 40$ and $n = 400$.[4] Then, we compare the RMSE of reconstructions from PnP ADMM with the f_{LQM} data fidelity term using different choices of β. We scale

[4]Scaling n with λ maintains the advantage of TR measurement and the validity of approximations used to justify the LQM estimator [14].

TABLE III: Fractions of 40 independent HIM emulations for which each combination of data fidelity term and denoiser results in convergence to a fixed point.

		Data fidelity				
		Gaussian	Oracle	QM	LQM	TRML
Denoiser	TV	0.70	1	1	1	1
	BM3D	0	0	0	0	0
	DnCNN	0.55	1	1	1	0.025

f_{LQM} by $1/n$ to make the comparison fair, since f_{LQM} scales approximately linearly with n under a fixed expected number of incident ions per subacquisition. The ADMM inversion step then becomes

$$\boldsymbol{\eta}_k^{(t+1)} = -\frac{1}{2}\varphi + \frac{1}{2}\sqrt{\varphi^2 + \frac{4\mathbf{y}_k^{\mathrm{C}}}{\rho n}} \tag{50a}$$

where

$$\varphi = \frac{\boldsymbol{L}_k}{\rho n(1 - e^{-\boldsymbol{\eta}_k^{\mathrm{prev}}})} - \boldsymbol{v}_k + \boldsymbol{u}_k. \tag{50b}$$

We fix $\rho = 0.005$ and $\alpha = 5 \times 10^{-4}$, and we use the TV denoiser and DnCNN with varying β and μ. Note that β cannot be specified exactly when using DnCNN, because the relationship between μ and the Gaussian noise standard deviation σ relies on assumptions about the denoiser that are not true in practice [45].

The RMSE of the reconstructions are shown in Figure 3. As expected, the higher the dose λ, the greater the information in the measurements and thus the smaller the error. Consequently, as λ increases, we observe that the optimal regularization strength decreases, i.e., β decreases and μ increases.

E. Empirical Convergence

The empirical fixed point convergence behaviors of our PnP methods agree with the theory in Section VI-E. As shown in Table III, the PnP methods with a convex data fidelity term—f_{oracle}, f_{QM}, or f_{LQM}—and a denoiser with contractive residue—TV denoiser or DnCNN—always converge to some fixed point. The combination of BM3D with any data fidelity term always fails to converge because BM3D has expansive residues [46]. While there is no convergence guarantee for PnP methods with the nonconvex data fidelity terms f_{conv} and f_{TRML}, they sometimes converge. One explanation is that the iterates converge in a locally convex region of the data fidelity term. We observe similar convergence behaviors in SEM emulations. However, with SEM data, fixed point convergence when using f_{TRML} is more difficult. This may be partially understood through examination of $\partial^2 f_{\mathrm{TRML}}(\boldsymbol{\eta})/\partial \boldsymbol{\eta}_k^2$. The derivative of the leading term of (44) is nonpositive. When $\boldsymbol{\eta}_k$ is small (as in SEM), $n - \boldsymbol{L}_k$ is likely to be relatively large, making the magnitude of the derivative of the leading term of (44) relatively large; empirically, this term is causing a lack of local convexity of $f_{\mathrm{TRML}}(\boldsymbol{\eta})$.

Typical convergence behaviors of some combinations of data fidelity terms and denosiers are shown in Figure 4. For PnP methods that converge to fixed points, the final iterate $\boldsymbol{\eta}^{(\mathrm{final})}$ is approximately the fixed point. For PnP methods

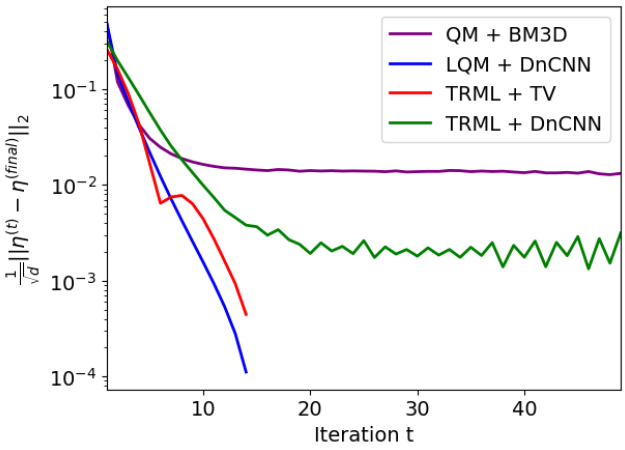


Fig. 4: Root mean-square distance between the reconstruction at each iteration $\boldsymbol{\eta}^{(t)}$ and the final reconstruction $\boldsymbol{\eta}^{(\text{final})}$ for selected combinations of data fidelity terms and denoisers from an HIM emulation.

that do not converge to fixed points, the root mean-square distance between the iterates $\boldsymbol{\eta}^{(t)}$ and $\boldsymbol{\eta}^{(\text{final})}$ either plateaus or fluctuates.

VII. Conclusion

In this paper, we develop data fidelity terms to make model-based reconstruction efficiently applicable to the distinctive measurement model for particle beam microscopy. We use these data fidelity terms in plug-and-play methods for estimation of the mean secondary electron yield $\boldsymbol{\eta}$. Because of the Neyman Type A likelihood for PBM data, using the negative log-likelihood directly would be computationally intractable. We introduce approximations with different computational complexities and accuracies applicable to conventional or time-resolved measurements, and we compare their efficacies when combined with three different denoisers. In synthetic experiments emulating helium ion microscopy and scanning electron microscopy, we demonstrate that our approaches outperform pixelwise (non-regularized) methods substantially in RMSE, SSIM, and qualitative appearance; RMSE reduction is by a factor of 2.24 to 4.11. We provide the first systematic demonstration that improvements due to regularization and to time-resolved measurement can be complementary.

Acknowledgment

The authors thank Akshay Agarwal for insightful discussions. The authors also acknowledge the Boston University Research Computing Services group for providing computational resource support which has contributed to the results reported in this paper.

References

[1] D. McMullan, "Scanning electron microscopy 1928–1965," *Scanning*, vol. 17, no. 3, pp. 175–185, 1995.

[2] B. W. Ward, J. A. Notte, and N. P. Economou, "Helium ion microscope: A new tool for nanoscale microscopy and metrology," *J. Vacuum Sci. Technol. B*, vol. 24, no. 6, pp. 2871–2874, Nov. 2006.

[3] J. Notte and J. Huang, *The Helium Ion Microscope*. Cham: Springer International Publishing, 2016, pp. 3–30.

[4] J. Morgan, J. Notte, R. Hill, and B. Ward, "An introduction to the helium ion microscope," *Microscopy Today*, vol. 14, no. 4, pp. 24–31, 2006.

[5] D. Emmrich, A. Beyer, A. Nadzeyka, S. Bauerdick, J. C. Meyer, J. Kotakoski, and A. Gölzhäuser, "Nanopore fabrication and characterization by helium ion microscopy," *Appl. Phys. Lett.*, vol. 108, no. 16, p. 163103, 2016.

[6] M. S. Joens, C. Huynh, J. M. Kasuboski, D. Ferranti, Y. J. Sigal, F. Zeitvogel, M. Obst, C. J. Burkhardt, K. P. Curran, S. H. Chalasani, L. A. Stern, B. Goetze, and J. A. J. Fitzpatrick, "Helium ion microscopy (HIM) for the imaging of biological samples at sub-nanometer resolution," *Sci. Rep.*, vol. 3, no. 3514, 2013.

[7] S. Bazargan, N. F. Heinig, J. F. Rios, and K. T. Leung, "Electronic transport in tin(IV) oxide nanocrystalline films: Two-medium transport with three-dimensional variable-range hopping mechanism for the ultra-small nanocrystallite size regime," *J. Phys. Chem. C*, vol. 116, no. 8, pp. 4979–4985, Mar. 2012.

[8] T. Wirtz, O. De Castro, J.-N. Audinot, and P. Philipp, "Imaging and analytics on the helium ion microscope," *Ann. Rev. Analytical Chem.*, vol. 12, pp. 523–543, 2019.

[9] J. Orloff, L. W. Swanson, and M. Utlaut, "Fundamental limits to imaging resolution for focused ion beams," *J. Vac. Sci. & Technol. B*, vol. 14, no. 6, pp. 3759–3763, 1996.

[10] R. Livengood, S. Tan, Y. Greenzweig, J. Notte, and S. McVey, "Subsurface damage from helium ions as a function of dose, beam energy, and dose rate," *J. Vac. Sci. & Technol. B*, vol. 27, no. 6, pp. 3244–3249, 2009.

[11] V. Castaldo, C. W. Hagen, P. Kruit, E. van Veldhoven, and D. Maas, "On the influence of the sputtering in determining the resolution of a scanning ion microscope," *J. Vacuum Sci. Technol. B*, vol. 27, no. 6, pp. 982–994, 2009.

[12] Y. Jiang, Z. Chen, Y. Han, P. Deb, H. Gao, S. Xie, P. Purohit, M. W. Tate, J. Park, S. M. Gruner, V. Elser, and D. A. Muller, "Electron ptychography of 2D materials to deep sub-ångström resolution," *Nature*, vol. 559, pp. 343–349, 2018.

[13] M. Peng, J. Murray-Bruce, K. K. Berggren, and V. K. Goyal, "Source shot noise mitigation in focused ion beam microscopy by time-resolved measurement," *Ultramicroscopy*, vol. 211, no. 112948, Apr. 2020.

[14] M. Peng, J. Murray-Bruce, and V. K. Goyal, "Time-resolved focused ion beam microscopy: Modeling, estimation methods, and analyses," *IEEE Trans. Comput. Imaging*, vol. 7, pp. 547–561, 2021.

[15] A. Agarwal, J. Simonaitis, V. K. Goyal, and K. K. Berggren, "Secondary electron count imaging in SEM," *Ultramicroscopy*, vol. 254, no. 113662, 2023.

[16] A. Agarwal, M. Peng, and V. K. Goyal, "Continuous-time modeling and analysis of particle beam metrology," arXiv:2303.04100, Mar. 2023.

[17] R. Ahmad, C. A. Bouman, G. T. Buzzard, S. Chan, S. Liu, E. T. Reehorst, and P. Schniter, "Plug-and-play methods for magnetic resonance imaging: Using denoisers for image recovery," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 105–116, Jan. 2020.

[18] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2017, pp. 3929–3938.

[19] J. He, Y. Yang, Y. Wang, D. Zeng, Z. Bian, H. Zhang, J. Sun, Z. Xu, and J. Ma, "Optimizing a parameterized plug-and-play ADMM for iterative low-dose CT reconstruction," *IEEE Trans. Medical Imaging*, vol. 38, no. 2, pp. 371–382, Feb. 2018.

[20] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play ADMM for image restoration: Fixed-point convergence and applications," *IEEE Trans. Comput. Imaging*, vol. 3, no. 1, pp. 84–98, Jan. 2016.

[21] U. S. Kamilov, C. A. Bouman, G. T. Buzzard, and B. Wohlberg, "Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications," *IEEE Signal Process. Mag.*, vol. 40, no. 1, pp. 85–97, Jan. 2023.

[22] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *Proc. IEEE Global Conf. Signal Inform. Process.*, 2013, pp. 945–948.

[23] M. A. Figueiredo and J. M. Bioucas-Dias, "Restoration of Poissonian images using alternating direction optimization," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3133–3145, 2010.

[24] A. Rond, R. Giryes, and M. Elad, "Poisson inverse problems by the plug-and-play scheme," *J. Visual Commun. Image Representation*, vol. 41, pp. 96–108, 2016.

[25] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlin. Phenom.*, vol. 60, no. 1-4, pp. 259–268, Nov. 1992.

[26] M. Peng, M. Cokbas, U. Dorken Gallastegi, P. Ishwar, J. Konrad, B. Kulis, and V. K. Goyal, "Convolutional neural network denoising of focused ion beam micrographs," in *Proc. IEEE 31st Int. Workshop Mach. Learn. Signal Process.*, Gold Coast, Queensland, Australia, Oct. 2021.
[27] S. W. Seidel, L. Watkins, M. Peng, A. Agarwal, C. Yu, and V. K. Goyal, "Online beam current estimation in particle beam microscopy," *IEEE Trans. Comput. Imaging*, vol. 8, pp. 521–535, 2022.
[28] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
[29] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
[30] F. F. Rahman, J. A. Notte, R. H. Livengood, and S. Tan, "Observation of synchronized atomic motions in the field ion microscope," *Ultramicroscopy*, vol. 126, pp. 10–18, 2013.
[31] S. W. Seidel, L. Watkins, M. Peng, A. Agarwal, C. C. Yu, and V. K. Goyal, "Addressing neon gas field ion source instability through online beam current estimation," *Microsc. Microanal.*, vol. 28, no. S1, pp. 36–39, Aug. 2022.
[32] L. Watkins, S. W. Seidel, M. Peng, A. Agarwal, C. C. Yu, and V. K. Goyal, "Robustness of time-resolved measurement to unknown and variable beam current in particle beam microscopy," in *Proc. IEEE Int. Conf. Image Process.*, Anchorage, AK, Sep. 2021, pp. 3487–3491.
[33] ——, "Prevention beats removal: Avoiding stripe artifacts from current variation in particle beam microscopy through time-resolved sensing," *Microsc. Microanal.*, vol. 27, no. S1, pp. 422–425, Aug. 2021.
[34] J. A. Notte, "Imaging with helium ions – A new detector regime with new challenges and new opportunities," American Vacuum Society 60th Int. Symp. Exhibit., Oct. 31 (presented, no proceedings), 2013.
[35] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Adv. Comput. Math.*, vol. 5, no. 1, pp. 329–359, Dec. 1996.
[36] S. Boyd, N. Parikh, and E. Chu, *Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers*. Now Foundations and Trends, 2011.
[37] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, Nov. 2009.
[38] U. S. Kamilov, H. Mansour, and B. Wohlberg, "A plug-and-play priors approach for solving nonlinear imaging inverse problems," *IEEE Signal Process. Lett.*, vol. 24, no. 12, pp. 1872–1876, 2017.
[39] D. C. Martin and S. K. Katti, "Approximations to the Neyman type A distribution for practical problems," *Biometrika*, vol. 18, no. 3, pp. 354–364, Sep. 1962.
[40] C. A. Bouman, *Foundations of Computational Imaging: A Model-Based Approach*. Philadelphia, PA: SIAM, 2022.
[41] A. Ziabari, J. M. Rickman, L. F. Drummy, J. P. Simmons, and C. A. Bouman, "Physics-based regularizer for joint soft segmentation and reconstruction of electron microscopy images of polycrystalline microstructures," *IEEE Trans. Comput. Imaging*, vol. 5, no. 4, pp. 660–674, 2019.
[42] S. Ono, "Primal-dual plug-and-play image restoration," *IEEE Signal Process. Lett.*, vol. 24, no. 8, pp. 1108–1112, 2017.
[43] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2016, pp. 770–778.
[44] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
[45] X. Xu, J. Liu, Y. Sun, B. Wohlberg, and U. S. Kamilov, "Boosting the performance of plug-and-play priors via denoiser scaling," arXiv:2002.11546, Feb. 2020.
[46] E. Ryu, J. Liu, S. Wang, X. Chen, Z. Wang, and W. Yin, "Plug-and-play methods provably converge with properly trained denoisers," in *International Conference on Machine Learning*. PMLR, 2019, pp. 5546–5557.
[47] R. Kitichotkul and M. Peng, "Goyal-STIR-Group/PnP-Particle-Beam-Micrograph- Denoising: PnP PBM Denoising 1.0.0 Release," May 2023. [Online]. Available: https://doi.org/10.5281/zenodo.7939229
[48] R. Aversa, M. H. Modarres, S. Cozzini, and R. Ciancio, "NFFA-EUROPE - SEM dataset," In Dataset of 18, Vol. 577 SEM images produced at CNR-IOM (Trieste). Not for any other scientific scope related to their content. This work has been done within the NFFA-EUROPE project (www.nffa.eu) and has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 654360 NFFA., 2018. [Online]. Available: https://b2share.eudat.eu/records/19cc2afd23e34b92b36a1dfd0113a89f
[49] J. Notte, B. Ward, N. Economou, R. Hill, R. Percival, L. Farkas, and S. McVey, "An introduction to the helium ion microscope," in *AIP Conference Proceedings*, vol. 931. AIP, 2007, pp. 489–496.
[50] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 33, no. 5, pp. 898–916, May 2011.
[51] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.