

Joint Reconstruction and Calibration Using Regularization by Denoising with Application to Computed Tomography

Mingyang Xie¹*, Jiaming Liu¹*, Yu Sun¹, Weijie Gan¹, Brendt Wohlberg², Ulugbek S. Kamilov¹

¹Washington University in St. Louis, MO, USA

²†Los Alamos National Laboratory, NM, US

{ericxie, jiaming.liu, sun.yu, weijie.gan, kamilov}@wustl.edu, brendt@ieee.org

Abstract

Regularization by denoising (RED) is a broadly applicable framework for solving inverse problems by using priors specified as denoisers. While RED has been shown to provide state-of-the-art performance in a number of imaging applications, existing RED algorithms require exact knowledge of the measurement operator characterizing the imaging system, limiting their applicability in problems where the measurement operator has parametric uncertainties. We propose a new method, called Calibrated RED (Cal-RED), that enables joint calibration of the measurement operator along with reconstruction of the unknown image. Cal-RED extends the traditional RED methodology to imaging problems that require the calibration of the measurement operator. We validate Cal-RED on the problem of image reconstruction in computerized tomography (CT) under perturbed projection angles. Our results corroborate the effectiveness of Cal-RED for joint calibration and reconstruction using pre-trained deep denoisers as image priors.

1. Introduction

The recovery of an unknown image \mathbf{x} from a set of noisy measurements \mathbf{y} is a central problem in computational imaging. The recovery is commonly formulated as an *inverse problem* that combines a physical-model characterizing the imaging system with a regularizer imposing a prior knowledge on the unknown image. Over the past years, many regularizers have been proposed as imaging priors, including those based on transform-domain sparsity, low-rank penalty, and dictionary learning [1–4].

Recent work has demonstrated the benefit of using deep learning (DL) for solving imaging inverse problems [5–7].

*These authors contributed equally.

†Research presented in this article was supported by the Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20200061DR.

The traditional DL approach is based on training a *convolutional neural network (CNN)* architecture to directly invert the measurement operator by exploiting the natural redundancies in the imaging data [8–10]. Model-based DL is an alternative to the traditional DL that explicitly uses knowledge of the forward model by integrating a CNN into model-based optimization. Two well-known approaches in this context are *Plug-and-play priors (PnP)* [11] and *regularization by denoising (RED)* [12], which have been exploited with pre-trained deep denoisers as priors to achieve excellent performance in a number of imaging tasks [13–19]. *Deep unfolding* is a related approach that interprets the iterations of an image recovery algorithm as layers of a CNN and trains it end-to-end in a supervised fashion [5, 20–24].

There has been a considerable amount of work on model-based DL for imaging inverse problems, the unifying theme being that one assumes accurate knowledge of the measurement operator characterizing the imaging system. In some cases, however, this operator is not known to sufficient accuracy, and can be modelled as depending on some reconstruction parameters θ that have to be calibrated to obtain an accurate characterization of the imaging system. For example, calibration is essential in 2D or 3D *computerized tomography (CT)* when there is uncertainty in the projection angles used in data collection [25]. In such settings, it is common to either manually calibrate the imaging system by using a known phantom [26–29] or embed a calibration step into the reconstruction method [30–32]. Despite the recent popularity of PnP/RED, current PnP/RED algorithms assume exact knowledge of the measurement operator, limiting their ability to solve problems involving parametric uncertainties in the data acquisition.

We address this issue by proposing a new *calibrated RED (Cal-RED)* method that integrates an automatic parameter calibration procedure within the RED reconstruction. Each iteration of Cal-RED performs a θ update step, followed by a traditional \mathbf{x} update step, leading to the joint recovery of the unknown image \mathbf{x} and parameters θ . One of the key benefits of Cal-RED is that it can leverage powerful *deep denoisers*,

such as DnCNN [33], for regularizing the inverse problem. We validate Cal-RED on CT by showing its ability to: (a) reduce the *root mean squared error (RMSE)* of the projection angles from 2° to 0.114° ; and (b) improve the imaging *signal to noise ratio (SNR)* from 22.31 dB (uncalibrated) to 32.39 dB (calibrated). Cal-RED thus addresses an important gap in the current literature on RED by providing a flexible, scalable, and theoretically sound algorithm applicable to a wide variety of imaging problems. Note that while our focus in this paper is on RED, our calibration strategy is also fully compatible with PnP.

The outline for the rest of the paper is as follows. In Section 2, we introduce our notation and review some relevant background material on image reconstruction, discussing both the traditional model-based and the recent learning-based approaches. In Section 3, we present and explain the details of Cal-RED, including our strategy for joint reconstruction and calibration procedure. In Section 4, we provide numerical experiments that corroborate the excellent performance and practical relevance of Cal-RED. Section 5 concludes the paper.

2. Background

2.1. Inverse Problems in Imaging

Consider the imaging problem specified by the linear model

$$\mathbf{y} = \mathbf{A}_\theta \mathbf{x} + \mathbf{e}, \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^n$ denotes the unknown image, $\mathbf{y} \in \mathbb{R}^m$ denotes the measurements, $\mathbf{A}_\theta \in \mathbb{R}^{m \times n}$ is the measurement operator characterizing the response of the imaging system, $\theta \in \mathbb{R}^\ell$ denotes the unknown parameters of the measurement operator, and $\mathbf{e} \in \mathbb{R}^m$ is the noise. When the true value of the parameters θ is known, image reconstruction can be formulated as an optimization problem

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad \text{with} \quad f(\mathbf{x}) = g_\theta(\mathbf{x}) + h(\mathbf{x}), \quad (2)$$

where g_θ is the data-fidelity term that uses the measurement operator to ensure the consistency with the measurements and h is the regularizer that imposes prior knowledge onto \mathbf{x} . For example, consider the smooth ℓ_2 -norm data-fidelity term

$$g_\theta(\mathbf{x}) = \frac{1}{2} \|\mathbf{y} - \mathbf{A}_\theta \mathbf{x}\|_2^2, \quad (3)$$

and the nonsmooth TV regularizer $h(\mathbf{x}) = \tau \|\mathbf{D}\mathbf{x}\|_1$, where $\tau > 0$ is the regularization parameter and \mathbf{D} is the image gradient [1]. *Alternating direction method of multipliers (ADMM)* [34] and *fast iterative shrinkage/thresholding algorithm (FISTA)* [35] are two common optimization algorithms that address the nonsmoothness of h in image reconstruction via the proximal operator

$$\text{prox}_{\tau h}(\mathbf{z}) := \arg \min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \frac{1}{2} \|\mathbf{x} - \mathbf{z}\|_2^2 + \tau h(\mathbf{x}) \right\}, \quad (4)$$

which can be interpreted as a *maximum a prior probability (MAP)* estimator for AWGN denoising.

2.2. Image reconstruction using deep learning

Recently, deep learning has gained popularity due to its effectiveness for solving imaging inverse problems. An extensive review of deep learning in this context can be found in [36, 37]. A widely used approach first brings the measurements to the image domain and then trains a deep network to map the corresponding low-quality images $\{\tilde{\mathbf{x}}_j\}$ to their clean target versions $\{\mathbf{x}_j\}$ by solving an optimization problem [10, 38]

$$\arg \min_{\mathbf{w}} \frac{1}{M} \sum_{j=1}^M \mathcal{L}(\mathcal{T}_\mathbf{w}(\tilde{\mathbf{x}}_j) - \mathbf{x}_j), \quad (5)$$

where $\mathcal{T}_\mathbf{w}$ represents the CNN parametrized by \mathbf{w} trained under the loss function \mathcal{L} . Popular loss functions include the ℓ_2 -norm and ℓ_1 -norm. In practice, (5) can be optimized using the family of stochastic gradient descent (SGD) methods, such as adaptive moment estimation (ADAM). For example, in the context of sparse-view CT, prior methods have trained $\mathcal{T}_\mathbf{w}$ for mapping a filtered backprojected image $\tilde{\mathbf{x}}$ to a reconstruction \mathbf{x} from a fully-sampled groundtruth data [10, 36].

The idea of end-to-end inversion can be refined by including the measurement operator into the CNN architecture. Inspired by LISTA [39], the corresponding *unfolding algorithms* interpret iterations of a regularized inversion as layers of a CNN and train it end-to-end in a supervised fashion. A related class of methods [23, 24], have included a full CNN as a trainable regularizer within an unfolded algorithm. Such unfolding algorithms have been shown to be effective in a number of problems [21, 23, 24] and are closely related to the PnP/RED methods (discussed in Section 2.3) that also combine the measurement operator and the imaging prior. Their main differences are that the former optimize the parameters in an end-to-end manner, and generally produce higher-quality results with fewer iterations. However, the prior in PnP/RED is not trained end-to-end within an iterative algorithm, which significantly reduces the cost and memory requirements of training [40].

2.3. Using denoisers as image priors

Since the proximal operator (4) is mathematically equivalent to regularized image denoising, there has been considerable interest in developing denoiser-based iterative algorithms such as PnP [11] and RED [12]. The key idea in PnP is to replace the proximal operator with an advanced image denoiser D_σ , where $\sigma > 0$ controls the strength of denoising. This simple replacement enables PnP to regularize the problem by using advanced denoisers, such as BM3D [41] or DnCNN [42], that do not correspond to any explicit h . Recent studies have confirmed the effectiveness of PnP in a range of imaging applications [43, 44].

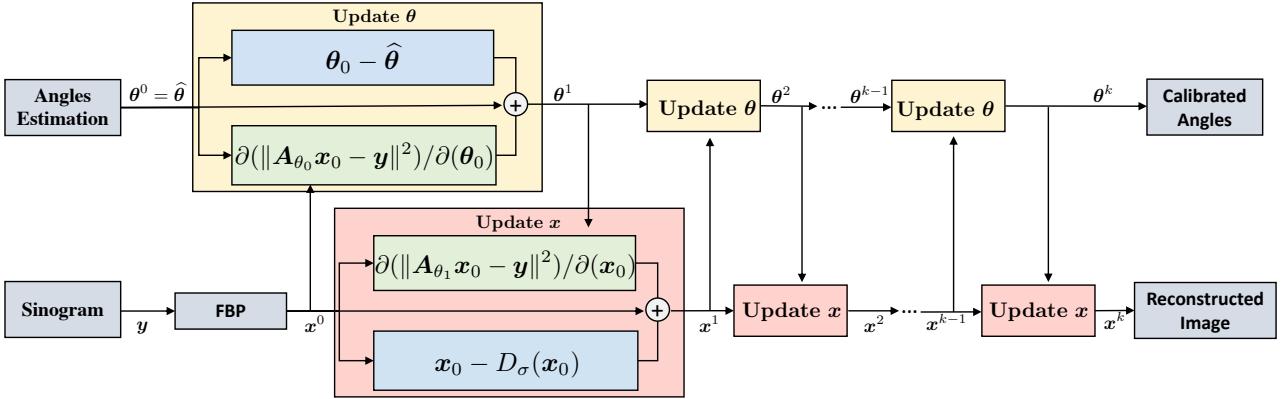


Figure 1. Algorithmic details of Cal-RED. The proposed method embeds automatic angle calibration procedure into the steepest descent variant of RED reconstruction. Each iteration of Cal-RED alternates between the θ update and x update. An accelerated variant of Cal-RED using Nesterov acceleration is presented in Algorithm 1.

RED is an alternative framework to PnP in which the image denoiser can sometimes lead to an explicit regularization function [12]

$$h(\mathbf{x}) = \frac{\tau}{2} \mathbf{x}^\top (\mathbf{x} - D_\sigma(\mathbf{x})) . \quad (6)$$

where τ is the regularization parameter. It has been shown that when the denoiser is locally homogeneous and has a symmetric Jacobian [45], the gradient of the RED regularizer h has a simple form

$$\nabla h(\mathbf{x}) = \tau(\mathbf{x} - D_\sigma(\mathbf{x})) . \quad (7)$$

which enables an efficient implementation of RED for problems of form (2). In the most general setting, RED algorithms seek an fixed point \mathbf{x}^* that satisfies

$$G(\mathbf{x}^*) = \nabla g_\theta(\mathbf{x}^*) + \tau(\mathbf{x}^* - D_\sigma(\mathbf{x}^*)) = \mathbf{0} , \quad (8)$$

where ∇g_θ denotes the gradient of the data-fidelity term using the true calibration parameters θ . Equivalently, \mathbf{x}^* lies in the zero set of $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$

$$\mathbf{x}^* \in \text{Zer}(G) := \{ \mathbf{x} \in \mathbb{C}^n : G(\mathbf{x}) = \mathbf{0} \} . \quad (9)$$

RED thus computes an equilibrium point that balances the data fidelity against the fixed points of the denoiser. The excellent performance of RED under a learned CNN denoiser has been demonstrated in various computational imaging tasks, including in image restoration [18], phase retrieval [17], and tomographic imaging [46]. Later works have developed scalable variants of RED [14, 47] and further replace the AWGN denoising with a general artifact-removal operator [48].

2.4. Calibrating the Measurement Operator

An error in tomographic reconstruction parameters can lead to a large number of artifacts in the reconstructed image.

Hence, target-free calibration of the measurement operator has been an emerging topic in recent years. For CT scanners with slightly misaligned projection angles, existing calibration methods include entropy-based correction [49], frequency-based correction [50], motion-corrected simultaneous algebraic correction [51], and a technique that combined steepest descent and multi-range testing [30]. A popular method to calibrate projection angles of Cryo-EM is called projection-matching [52, 53], which relies on a prior estimate of the density map of the unknown single particle. It seeks the angle of every projection by pairing it with a clean template of the projection of the estimated density map from a known angle. For instance, Malhotra *et al.* [32] propose a three-step pipeline for Cryo-EM to recover the projection angles and the target density map in sequence. Since current PnP/RED algorithms severely rely on accurate knowledge of the measurement operators, these methods can be surprisingly fragile when solving problems related to parameter uncertainty in the data acquisition. The contribution of this paper is the investigation of the calibration problem in the context of PnP/RED algorithms, which has not been done in the prior work.

3. Proposed Method

The proposed Cal-RED algorithm is summarized in Algorithm 1 and Figure 1. It alternates between the updates of θ and x . The parameters θ are initialized with their corresponding nominal values, $\hat{\theta}$, which are assumed to be known from the design of the imaging system. We additionally initialize the image x with a direct inversion from measurements y , which for CT reduces to the traditional FBP [54].

Cal-RED performs an alternating optimization over x and θ . The θ update in lines 5-7 in Algorithm 1 seeks to

Algorithm 1 Cal-RED

Result: the reconstruction image \mathbf{x} and calibrated $\boldsymbol{\theta}$

- 1: **Input:** $\mathbf{y} \in \mathbb{R}^m, \hat{\boldsymbol{\theta}} \in \mathbb{R}^l, \gamma_x > 0, \gamma_\theta > 0, \tau_x > 0, \tau_\theta > 0$, and $\{q_k\}_{k \in \mathbb{N}}$
- 2: $\boldsymbol{\theta} \leftarrow \hat{\boldsymbol{\theta}}$
- 3: $\mathbf{x} \leftarrow \text{DirectInversion}_{\hat{\boldsymbol{\theta}}}(\mathbf{y})$
- 4: **for** $k = 1, 2, \dots$ **do**
- 5: # $\boldsymbol{\theta}$ update step
- 6: $G_{\boldsymbol{\theta}}(\mathbf{s}^{k-1}) \leftarrow \frac{\partial g(\mathbf{s}^{k-1}, \mathbf{u}^{k-1})}{\partial \mathbf{u}^{k-1}} + \tau_\theta(\mathbf{u}^{k-1} - \hat{\boldsymbol{\theta}})$
- 7: $\boldsymbol{\theta}^k \leftarrow \mathbf{u}^{k-1} - \gamma_\theta G_{\boldsymbol{\theta}}(\mathbf{u}^{k-1})$
- 8: $\mathbf{u}^k \leftarrow \boldsymbol{\theta}^k + \frac{q_{k-1}-1}{q_k}(\boldsymbol{\theta}^k - \boldsymbol{\theta}^{k-1})$
- 9: # \mathbf{x} update step
- 10: $G_x(\mathbf{s}^{k-1}) \leftarrow \frac{\partial g(\mathbf{s}^{k-1}, \boldsymbol{\theta}^k)}{\partial \mathbf{s}^{k-1}} + \tau_x(\mathbf{s}^{k-1} - D_\sigma(\mathbf{s}^{k-1}))$
- 11: $\mathbf{x}^k \leftarrow \mathbf{s}^{k-1} - \gamma_x G_x(\mathbf{s}^{k-1})$
- 12: $\mathbf{s}^k \leftarrow \mathbf{x}^k + \frac{q_{k-1}-1}{q_k}(\mathbf{x}^k - \mathbf{x}^{k-1})$
- 13: **end for**

minimize the following objective function for a fixed \mathbf{x}

$$\arg \min_{\boldsymbol{\theta}} \left\{ g_{\boldsymbol{\theta}}(\mathbf{x}) + p(\boldsymbol{\theta}) \right\} \quad \text{with} \quad p(\boldsymbol{\theta}) = \frac{\tau_\theta}{2} \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|_2^2 \quad (10)$$

by performing a step of the steepest descent

$$\boldsymbol{\theta}^k \leftarrow \boldsymbol{\theta}^{k-1} - \gamma_\theta \left(\frac{\partial g_{\boldsymbol{\theta}^{k-1}}(\mathbf{x}^{k-1})}{\partial \boldsymbol{\theta}^{k-1}} + \tau_\theta(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \right). \quad (11)$$

Eq. (10) incorporates a Tikhonov penalty as the prior for $\boldsymbol{\theta}$. Similarly, the simplest algorithm for computing (9) can be obtained by running the following fixed-point algorithm for a fixed $\boldsymbol{\theta}$

$$\mathbf{x}^k \leftarrow \mathbf{x}^{k-1} - \gamma_x G_x(\mathbf{x}^{k-1}), \quad (12)$$

where $\gamma > 0$ is the step-size. Cal-RED relies on Nesterov acceleration [55] for both \mathbf{x} and $\boldsymbol{\theta}$ updates by using the acceleration sequence

$$q_k \leftarrow \frac{1}{2} \left(1 + \sqrt{1 + 4q_{k-1}^2} \right) \quad \text{with} \quad q_1 = 1.. \quad (13)$$

When $q_k = 1$ for all k , the algorithms reverts to the usual gradient method without acceleration. In Algorithm 1, the variables \mathbf{s} and \mathbf{u} store the intermediate values used for acceleration of \mathbf{x} and $\boldsymbol{\theta}$, respectively. Note that Cal-RED does not modify the \mathbf{x} update step of RED, which implies that the $\boldsymbol{\theta}$ update step of Cal-RED could also be integrated within PnP or other variants of RED, such as BC-RED [47].

The convergence of the traditional RED algorithm has been established in prior works [45]. The same analysis applies to Cal-RED if the number of $\boldsymbol{\theta}$ updates is fixed. This can be practically achieved by performing $\boldsymbol{\theta}$ updates only within a fixed number of initial iterations and keeping $\boldsymbol{\theta}$ constant thereafter. For example, in our experiments, $\boldsymbol{\theta}$

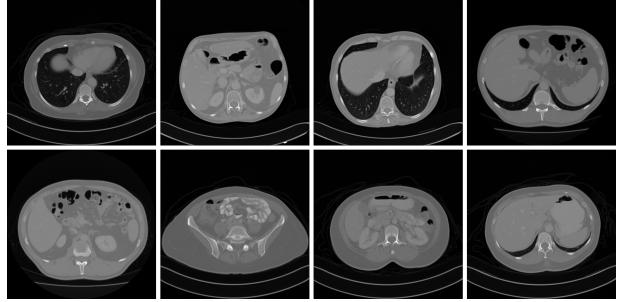


Figure 2. Eight 512×512 images from the scans of three patients in the AAPM human phantom dataset are used for testing.

stabilizes within the first 1000 iterations, which suggest that Cal-RED could stop updating $\boldsymbol{\theta}$ after those initial iterations.

In our implementation of the denoising neural network D_w , we adopt a simple DnCNN network used in [46]. This DnCNN like architecture has been extensively used as imaging priors for various inverse problems [17, 24, 47]. In particular, the DnCNN consists of 12 convolutional layers, including a normal convolutional layer and a rectified linear units (ReLU) layer. All the convolution filters are of size 3×3 and every feature map has 64 channels. We employed the residual learning technique [56] in DnCNN so that the network is forced to learn the noise residual in the noisy input

$$R_w(\mathbf{x}) = (I - D_w)(\mathbf{x}) = \mathbf{x} - D_w(\mathbf{x}). \quad (14)$$

The CNN denoiser was trained to minimize the following loss over a training set

$$\begin{aligned} \arg \min_{\mathbf{w}} \frac{1}{M} \sum_{j=1}^M \mathcal{L}(R_w(\tilde{\mathbf{x}}_j) - \mathbf{x}_j) \quad &\text{with} \\ \mathcal{L} = \alpha \mathcal{L}_{\ell_1} + (1 - \alpha) \mathcal{L}_{\ell_2}, \end{aligned} \quad (15)$$

where $\alpha \in (0, 1)$ adjusts the strength of the ℓ_1/ℓ_2 loss.

4. Numerical Validation

In this section, we empirically validate the performance of Cal-RED on simulated parallel-beam CT reconstruction, where the 2D target \mathbf{x} is rotated around an axis to acquire projections from different angles and form a sinogram. All the experiments were performed on a machine equipped with an Intel Xeon Gold 6130 Processor and four NVIDIA GeForce RTX 1080 Ti GPUs.

4.1. Implementation Details

Conventional CT requires many views for high-quality image reconstruction. In the following experiments, we explore the possibility of high-quality imaging when reducing the number of views in CT imaging. We assume

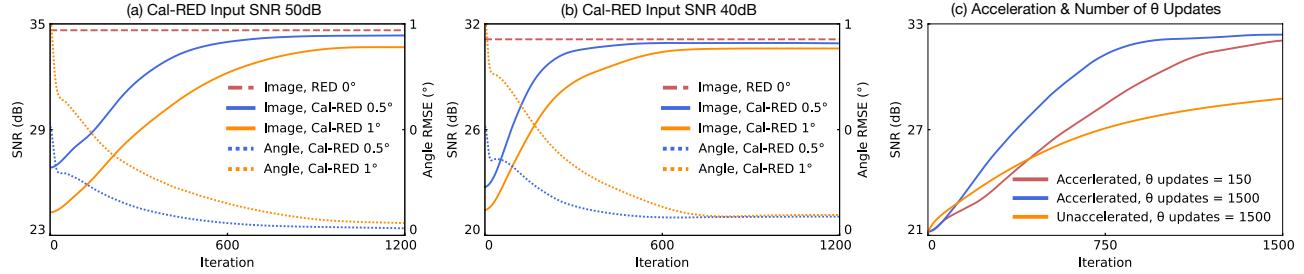


Figure 3. Illustration of the convergence of Cal-RED with projection angles corrupted by AWGN of SD 0.5° and 1° . The left figure corresponds to a sinogram input SNR of 50 dB, while the middle figure corresponds to that of 40 dB. In each figure, the values of SNR and RMSE are plotted against the number of iterations using solid and dotted curves, respectively. The left axis represents the scale of SNR while the right axis represents that of RMSE. The optimal SNR values (obtained by RED using the true angles) are plotted as the horizontal dashed line for reference. Note the significant reduction of angular errors and the nearly optimal SNR results. The rightmost figure shows the influence of using acceleration and the number of θ update on the convergence of Cal-RED with angular error of 1° .

that the CT machine is designed to project from nominal angles $\hat{\theta}$, an array of 90 projection angles that are evenly-distributed on a half circle and 724 detector pixels. However, due to alignment errors, the CT machine actually projects from θ , another array of 90 angles that is assumed to be the result of corrupting $\hat{\theta}$ by an unknown AWGN vector with standard deviation (SD) of $\{0.5^\circ, 1^\circ, 2^\circ\}$, respectively. In our experiments, we synthesize the sinograms by using the ground-truth θ and further corrupt these sinograms by AWGN corresponding to the input SNR of 40 dB and 50 dB, respectively. We consider reconstruction of simulated data obtained from the clinically realistic CT images provided by Mayo Clinic for the *AAPM Low Dose CT Grand Challenge* [57]. In all experiments, we select eight 512×512 test images from the scans of three patients, presented in Figure 2, while the scans of other patients are used for training the deep denoiser in Cal-RED and other learning baseline methods. We implement the measurement operator \mathbf{A}_θ and its adjoint \mathbf{A}_θ^H with PyTorch implementation of Radon and IRadon¹ transform. The partial derivative $\partial g_\theta(\mathbf{x}, \theta)/\partial \theta$ in (10) is thus efficiently calculated using automatic differentiation [58]. While the results in this section are for 2D tomography, Cal-RED can be easily extended to 3D data via minor modifications: replace 2D DnCNN with 3D DnCNN; change \mathbf{A}_θ from 2D Radon transform to 3D Radon transform.

We use the SNR in dB for the quantitatively comparing different algorithms

$$\text{SNR}(\hat{\mathbf{x}}, \mathbf{x}) \triangleq 10 \log_{10} \frac{\|\mathbf{x}\|_2^2}{\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2} \quad (16)$$

where $\hat{\mathbf{x}}$ and \mathbf{x} represents the noisy vector and ground truth that are restricted to the region of interests (ROI) within the body area by removing the background and patient bed [10].

¹The code is publicly available at https://github.com/phernst/pytorch_radon

We also use the angular error by RMSE

$$\text{RMSE}(\hat{\theta}, \theta) \triangleq \sqrt{\frac{\sum_{i=1}^N (\hat{\theta}_i - \theta_i)^2}{N}}. \quad (17)$$

and the *structural similarity index measure (SSIM)* [59] as alternative metrics.

We compare Cal-RED with several baseline methods, including FBP, RED [12], URED [24], U-Net [60], FISTA [35] and the traditional least-square method (LSM). All learning-based methods are trained using Adam optimizer [61] with data parallel and weight decay 3×10^{-8} in Pytorch. In these experiments, RED assimilates the same denoisers with Cal-RED. We train DnCNN denoisers for AWGN removal at four noise levels $\sigma \in \{2, 5, 10, 15\}$ and pick the one that achieves the best performance for evaluation. URED is a recently proposed deep unfolding architecture based on the feed-forward network obtained by unfolding and truncating RED. Since we numerically observed that using the calibration within URED can improve its performance, we opt to equip URED with the same calibration schemes in the proposed Cal-RED, named *angle-calibrating URED (Cal-URED)*. We implement an individual Cal-URED for each combination of $\{40 \text{ dB}, 50 \text{ dB}\} \times \{0.5^\circ, 1^\circ, 2^\circ\}$, while we train additional URED for angular SD of 0° . The implementation of Cal-URED consists of four stages. We first unfold URED for $Q = 8$ iterations and train URED with data generated by filtered-back-projecting the sinogram using the mismatched angles $\hat{\theta}$. Note that this training does not involve θ updates. In the second stage, we continue to train the URED obtained from stage one with additional calibration step in (11) for a few epochs. In the third stage, similar to Cal-RED, we manually optimize the physical parameters of pre-trained Cal-URED using Algorithm 1 in order to further improve the angle calibration accuracy. As can be seen below the θ is well calibrated and closed to the ground truth after stage three. Finally, we train a second URED using

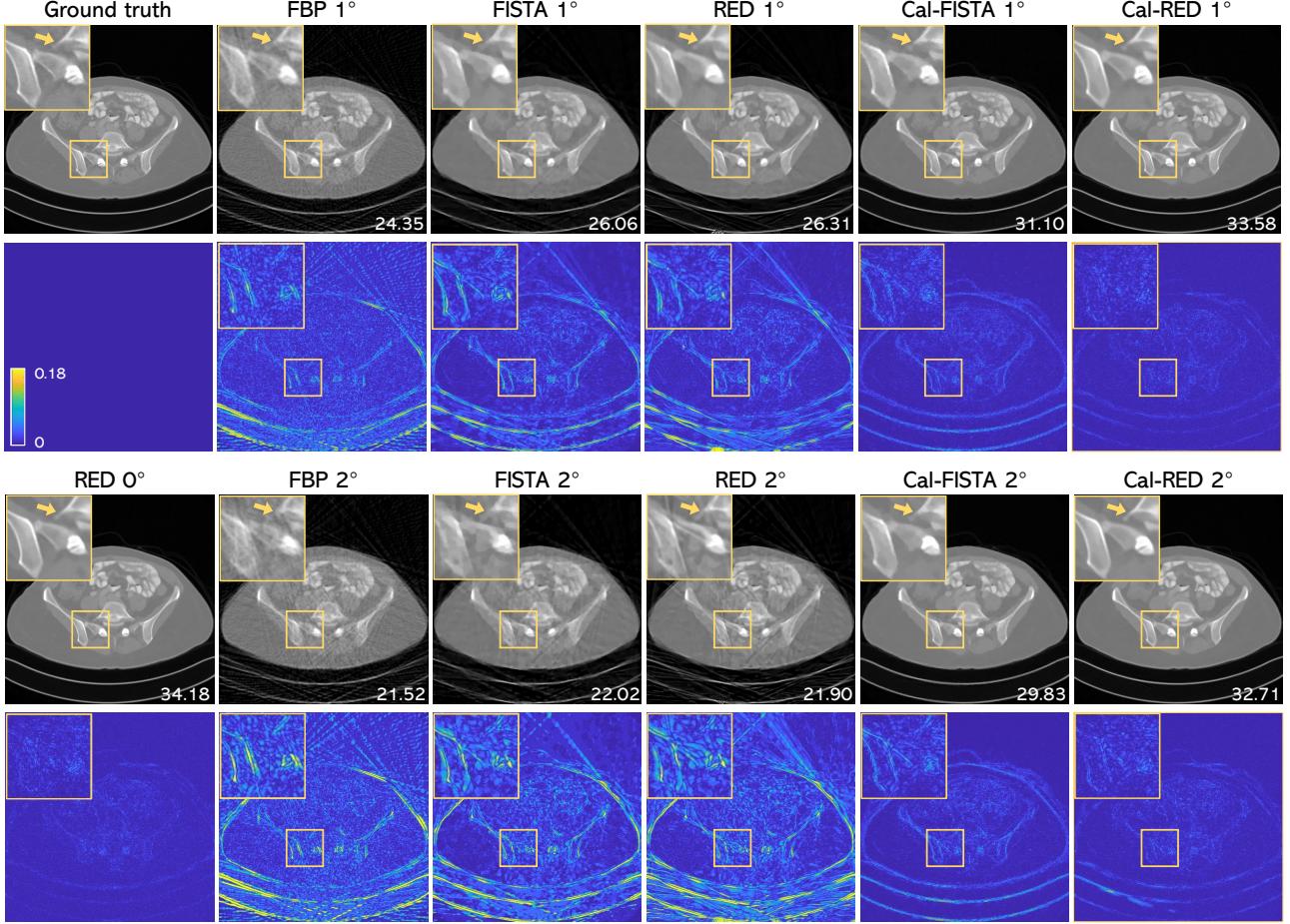


Figure 4. Visual illustration of recovering a 512×512 image from a sinogram with noise corresponding to input SNR of 50 dB and projection angles corrupted by AWGN of SD 1° and 2° . Both the ground truth and the image reconstructed by oracle RED given the true angles are placed at the leftmost column for reference. Each image is labeled with its SNR (dB) values with respect to the original image. The results highlight state-of-the-art performance of Cal-RED due to the calibration of the forward model during reconstruction.

the data generated with true angles based on the assumption that URED can tolerate with minor angle mismatches without significant performance degradation. Hence, the final reconstruction is the output of the second URED taking the calibrated angles from the third stage. Similar to URED, We train a U-Net with corrupted-clean image pairs generated by FBP using the mismatched angles $\hat{\theta}$. We train a separate U-Net for each experiments, respectively. For both Cal-URED and U-Net, the learning rate starts from 5×10^{-4} and is halved at epoch 20, then gradually reduced by a factor of 0.7 every 10 epochs. The number of total training epochs is 120. We adopt gradient clipping [62] in order to accelerate and stabilize the training. In order to better demonstrate the benefit of using deep denoising prior, we run FISTA with the TV regularizer for comparison. We also incorporate the θ update step into both FISTA and LSM to construct the angle-calibrating FISTA (Cal-FISTA) and the angle-calibrating LSM (Cal-LSM), respectively.

4.2. Comparison Studies

Figure 3 (a) and (b) illustrates the convergence of Cal-RED in terms of both image SNR and angular RMSE for input SNR of 50 dB and 40 dB, respectively. In the cases where θ is severely mismatched, Cal-RED is able to reduce the angular RMSE by a factor of more than 13, from 2° to around 0.146° ; moreover, Cal-RED achieves excellent performance compared to RED using the true θ . Figure 3 (c) provides additional evaluations highlighting the influence of different $x - \theta$ update strategies. It can be observed that increasing the number of θ update improves the performance and results in faster convergence.

Figure 4 compares the visual results of FBP, FISTA, RED, Cal-FISTA, and Cal-RED. Both the reconstructions of FBP and uncalibrated RED suffer from obvious line artifacts, which demonstrates their vulnerability to a mismatched measurement operator. While Cal-FISTA reduces those artifacts,

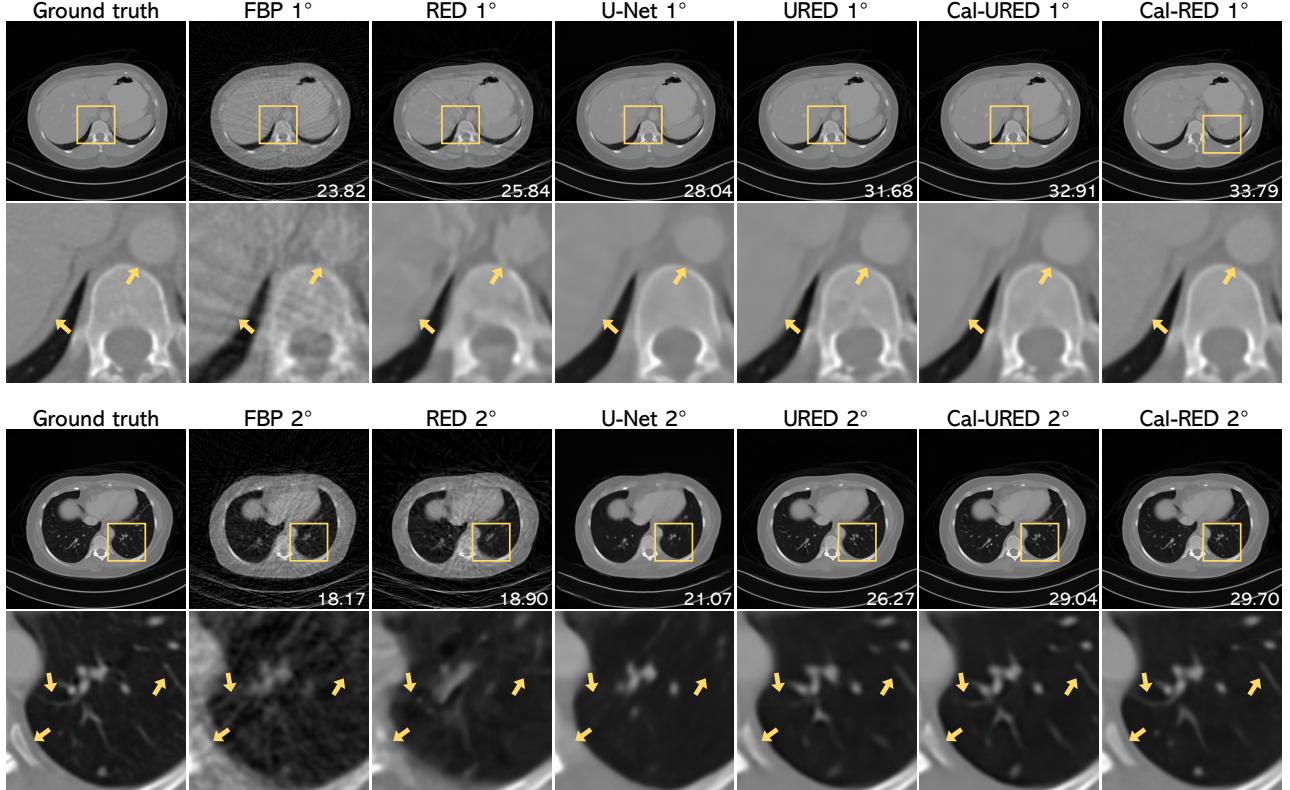


Figure 5. Visual illustration of Cal-RED relative to FBP and several deep learning baseline methods on CT reconstruction with angular SD of 1° and 2° and input SNR of 50 dB. Note that Cal-RED provides substantial improvements over calibrated RED, outperforming Cal-URED that incorporates model calibration into deep unfolding network.

it blurs the fine structures by producing cartoon-like features. On the other hand, Cal-RED performs better than other methods on both angular SD of 1° and 2°, highlighted by the residual error map. The zoomed regions suggest that Cal-RED can accurately reconstruct the fine details as good as RED using truth angles. Figure 5 presents the reconstruction results by all the learning-based methods. The U-Net removes most of the artifacts that are contained in the FBP reconstruction, but it also has an undesirable effect of oversmoothing small features. While the traditional deep learning alone fails to balance the preservation of features and the removal of artifacts, deep denoising priors like DnCNN help Cal-RED to reconstruct high-resolution images with minimal artifacts. While Cal-URED does an excellent job of improving the URED reconstruction by over 1.23 dB, its performance is limited by the poor generalization ability. Since the measurement operator A_θ in Cal-URED changes with different input images, the artifact patterns are different from one to another, which may result in an over-fitting problem. On the other hand, however, since Cal-RED uses denoisers as regularizers, it can easily deal with the artifacts between each iterations. Therefore, one cannot expect Cal-URED to recover significantly better results than Cal-RED

Table 1. Average RMSE of Calibrated CT Angles (°).

Noise level	Angular error	Method			
		Cal-LSM	Cal-TV	Cal-URED	Cal-RED
40 dB	0.5°	0.208	0.106	0.102	0.089
	1.0°	0.492	0.112	0.136	0.096
	2.0°	1.145	0.168	0.207	0.146
50 dB	0.5°	0.081	0.036	0.031	0.030
	1.0°	0.278	0.049	0.062	0.045
	2.0°	0.748	0.087	0.114	0.098

in our experiments.

Table 1, 2, and 3 summarize the numerical results under different combinations of input sinogram SNR and angular error. It is worth noting that Cal-RED and Cal-FISTA significantly outperform RED and FISTA, respectively, which highlights the effectiveness of the θ update step. As expected, URED achieves the best performance when there is no angular error. Despite being designed for the angle calibration purpose, when Cal-RED is applied to a problem

Table 2. Average SNRs of Reconstructed CT Images (dB).

Method	Input SNR 40 dB				Input SNR 50 dB			
	0°	0.5°	1.0°	2.0°	0°	0.5°	1.0°	2.0°
FBP	23.52	22.75	21.43	19.46	28.79	26.84	24.30	21.20
TV	28.90	27.56	25.39	22.29	32.44	29.23	26.14	22.32
U-Net	31.70	29.87	27.46	24.03	34.70	31.56	28.07	24.73
RED (DnCNN)	31.19	29.05	25.74	21.79	34.68	30.24	26.68	22.31
Cal-LSM	22.19	18.31	14.73	12.21	27.16	24.25	20.31	15.96
Cal-TV	28.85	28.64	28.30	27.57	32.24	31.96	31.40	29.85
Cal-URED	32.36	30.99	30.04	29.16	35.41	34.07	32.47	31.36
Cal-RED (DnCNN)	31.10	30.90	30.57	29.98	34.66	34.39	33.80	32.39

Table 3. Average SSIMs of Reconstructed CT Images.

Method	Input SNR 40 dB				Input SNR 50 dB			
	0°	0.5°	1.0°	2.0°	0°	0.5°	1.0°	2.0°
FBP	0.841	0.831	0.813	0.787	0.942	0.925	0.899	0.862
TV	0.967	0.960	0.945	0.921	0.981	0.968	0.952	0.917
U-Net	0.978	0.969	0.961	0.946	0.984	0.977	0.967	0.951
RED (DnCNN)	0.971	0.963	0.939	0.895	0.983	0.969	0.909	0.913
Cal-LSM	0.815	0.730	0.674	0.643	0.932	0.859	0.782	0.713
Cal-TV	0.966	0.965	0.963	0.959	0.980	0.979	0.977	0.971
Cal-URED	0.979	0.971	0.966	0.964	0.987	0.981	0.979	0.976
Cal-RED (DnCNN)	0.971	0.970	0.969	0.967	0.983	0.982	0.981	0.978

without calibration errors, it achieves an SNR only 0.09 dB less than that of the original RED. When the input SNR is 40 dB, the angles recovered by Cal-URED, Cal-RED, and Cal-FISTA are far more accurate than those recovered by Cal-LSM, which demonstrates how advanced priors can also contribute to measurement operator calibration.

5. Conclusion

We propose Cal-RED as a method that takes advantage of advanced denoising priors to jointly recover the measurement operator parameters and the target image. We validate its robustness to severely mismatched measurement operator parameters and noisy measurements via simulated CT reconstruction. Our extensive numerical results indicate that Cal-RED provides competitive imaging quality compared to the traditional and learning-based methods. While our experiments focused on medical CT, the proposed calibration strategy is broadly applicable to many other tomography modalities such as X-ray computed tomography (SXCT) [63] and cryo-electron tomography (cryo-ET) [64], where the evaluation of the measurement operator is not known to sufficient accuracy. Additionally, while we implemented Cal-RED based on the RED framework, the idea can be used for other model-based deep learning architectures.

References

- [1] L. I. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Physica D*, vol. 60, no. 1–4, pp. 259–268, Nov. 1992. [1](#) [2](#)
- [2] M. A. T. Figueiredo and R. D. Nowak, “Wavelet-based image estimation: An empirical bayes approach using Jeffreys’ noninformative prior,” *IEEE Trans. Image Process.*, vol. 10, no. 9, pp. 1322–1331, September 2001. [1](#)
- [3] M. Elad and M. Aharon, “Image denoising via sparse and redundant representations over learned dictionaries,” *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, December 2006. [1](#)
- [4] A. Danielyan, V. Katkovnik, and K. Egiazarian, “BM3D frames and variational image deblurring,” *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1715–1728, April 2012. [1](#)
- [5] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, “Using deep neural networks for inverse problems in imaging: Beyond analytical methods,” *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 20–36, Jan. 2018. [1](#)
- [6] G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, “Deep learning techniques for inverse problems in imaging,” *IEEE Journal on Selected Areas in Information Theory*, vol. 1, no. 1, pp. 39–56, 2020. [1](#)
- [7] W. Gan, Y. Sun, C. Eldeniz, J. Liu, H. An, and U. S. Kamilov, “Deep image reconstruction using unregistered measurements without groundtruth,” in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 2021, pp. 1531–1534. [1](#)
- [8] E. Kang, J. Min, and J. C. Ye, “A deep convolutional neural network using directional wavelets for low-dose x-ray CT reconstruction,” *Medical Physics*, vol. 44, no. 10, pp. e360–e375, 2017. [1](#)
- [9] Y. Chen and T. Pock, “Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restora-

- tion,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1256–1272, 2017. 1
- [10] Y. Han and J. C. Ye, “Framing U-Net via deep convolutional framelets: Application to sparse-view CT,” *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1418–1429, 2018. 1, 2, 5
- [11] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, “Plug-and-play priors for model based reconstruction,” in *Proc. IEEE Global Conf. Signal and Inf. Process. (GlobalSIP)*, 2013. 1, 2
- [12] Y. Romano, M. Elad, and P. Milanfar, “The little engine that could: Regularization by denoising (RED),” *SIAM J. Imaging Sci.*, vol. 10, no. 4, pp. 1804–1844, 2017. 1, 2, 3, 5
- [13] 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, Mar. 2017. 1
- [14] Y. Sun, B. Wohlberg, and U. S. Kamilov, “An online plug-and-play algorithm for regularized image reconstruction,” *IEEE Trans. Comput. Imaging*, vol. 5, no. 3, pp. 395–408, 2019. 1, 3
- [15] A. M. Teodoro, J. M. Bioucas-Dias, and M. Figueiredo, “A convergent image fusion algorithm using scene-adapted Gaussian-mixture-based denoising,” *IEEE Trans. Image Process.*, vol. 28, no. 1, pp. 451–463, Jan. 2019. 1
- [16] T. Meinhardt, M. Moeller, C. Hazirbas, and D. Cremers, “Learning proximal operators: Using denoising networks for regularizing inverse imaging problems,” in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, Venice, Italy, Oct. 2017. 1
- [17] C. A. Metzler, P. Schniter, A. Veeraraghavan, and R. G. Baraniuk, “prDeep: Robust phase retrieval with a flexible deep network,” in *Proc. 35th Int. Conf. Machine Learning (ICML)*, 2018. 1, 3, 4
- [18] G. Mataev, M. Elad, and P. Milanfar, “DeepRED: Deep image prior powered by RED,” in *Proc. IEEE Int. Conf. Computer Vision Workshops (ICCVW)*, Seoul, South Korea, Oct. 2019. 1, 3
- [19] J. Liu, M. S. Asif, B. Wohlberg, and U. S. Kamilov, “Recovery analysis for plug-and-play priors using the restricted eigenvalue condition,” *arXiv preprint arXiv:2106.03668*, 2021. 1
- [20] A. Hauptmann, F. Lucka, M. Betcke, N. Huynh, J. Adler, B. Cox, P. Beard, S. Ourselin, and S. Arridge, “Model-based learning for accelerated, limited-view 3-d photoacoustic tomography,” *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1382–1393, 2018. 1
- [21] H. K. Aggarwal, M. P. Mani, and M. Jacob, “Modl: Model-based deep learning architecture for inverse problems,” *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, Feb. 2019. 1, 2
- [22] J. Adler and O. Öktem, “Learned primal-dual reconstruction,” *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1322–1332, June 2018. 1
- [23] S. A. Hosseini, B. Yaman, S. Moeller, M. Hong, and M. Akcakaya, “Dense recurrent neural networks for accelerated MRI: History-cognizant unrolling of optimization algorithms,” *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 6, pp. 1280–1291, Oct. 2020. 1, 2
- [24] J. Liu, Y. Sun, W. Gan, B. Wohlberg, and U. S. Kamilov, “Sgd-net: Efficient model-based deep learning with theoretical guarantees,” *IEEE Transactions on Computational Imaging*, vol. 7, pp. 598–610, 2021. 1, 2, 4, 5
- [25] S. Basu and Y. Bresler, “Uniqueness of tomography with unknown view angles,” *IEEE Transactions on Image Processing*, vol. 9, no. 6, pp. 1094–1106, 2000. 1
- [26] C. E. Cann, “Quantitative CT for determination of bone mineral density: a review,” *Radiology*, vol. 166, no. 2, pp. 509–522, 1988. 1
- [27] X. Wang, J. G. Mainprize, M. P. Kempston, G. E. Mawdsley, and M. J. Yaffe, “Digital breast tomosynthesis geometry calibration,” in *Medical Imaging 2007: Physics of Medical Imaging*, vol. 6510, 2007. 1
- [28] H. Miao, X. Wu, H. Zhao, and H. Liu, “A phantom-based calibration method for digital X-ray tomosynthesis,” *Journal of X-ray Science and Technology*, vol. 20, pp. 17–29, Jan. 2012. 1
- [29] J. Zou, X. Hu, H. Lv, and X. Hu, “An investigation of calibration phantoms for CT scanners with tube voltage modulation,” *Int. Journal of Biomedical Imaging*, vol. 2013, Dec. 2013. 1
- [30] C. Cheng, Y. Ching, P. Ko, and Y. Hwu, “Correction of center of rotation and projection angle in synchrotron X-ray computed tomography,” *Scientific Reports*, vol. 8, Dec. 2018. 1, 3
- [31] D. Lee, P. Hoffmann, D. Kopperdhal, and T. Keaveny, “Phantomless calibration of CT scans for measurement of bmd and bone strength-inter-operator reanalysis precision,” *Bone*, vol. 103, Aug. 2017. 1
- [32] E. Malhotra and A. Rajwade, “Tomographic reconstruction from projections with unknown view angles exploiting moment-based relationships,” in *2016 IEEE Int. Conf. Image Process. (ICIP)*, 2016. 1, 3
- [33] 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, Feb. 2017. 2
- [34] M. V. Afonso, J. M. Bioucas-Dias, and M. A. T. Figueiredo, “Fast image recovery using variable splitting and constrained optimization,” *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, Sep. 2010. 2
- [35] A. Beck and M. Teboulle, “Fast gradient-based algorithm for constrained total variation image denoising and deblurring problems,” *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, November 2009. 2, 5
- [36] M. T. McCann, K. H. Jin, and M. Unser, “Convolutional neural networks for inverse problems in imaging: A review,” *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 85–95, 2017. 2
- [37] F. Knoll, K. Hammernik, C. Zhang, S. Moeller, T. Pock, D. K. Sodickson, and M. Akcakaya, “Deep-learning methods for

- parallel magnetic resonance imaging reconstruction: A survey of the current approaches, trends, and issues,” *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 128–140, Jan. 2020. 2
- [38] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, “Deep convolutional neural network for inverse problems in imaging,” *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017. 2
- [39] K. Gregor and Y. LeCun, “Learning fast approximation of sparse coding,” in *Proc. 27th Int. Conf. Machine Learning (ICML)*, Haifa, Israel, June 21-24, 2010, pp. 399–406. 2
- [40] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, “A deep cascade of convolutional neural networks for dynamic MR image reconstruction,” *IEEE Trans. Med. Imag.*, vol. 37, no. 2, pp. 491–503, Feb. 2018. 2
- [41] 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. 16, pp. 2080–2095, August 2007. 2
- [42] 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, July 2017. 2
- [43] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman, “Plug-and-play priors for bright field electron tomography and sparse interpolation,” *IEEE Trans. Comput. Imaging*, vol. 2, no. 4, pp. 408–423, Dec. 2016. 2
- [44] A. Brifman, Y. Romano, and M. Elad, “Turning a denoiser into a super-resolver using plug and play priors,” in *Proc. IEEE Int. Conf. Image Proc. (ICIP)*, Phoenix, AZ, USA, Sep. 2016. 2
- [45] E. T. Reehorst and P. Schniter, “Regularization by denoising: Clarifications and new interpretations,” *IEEE Trans. Comput. Imaging*, vol. 5, no. 1, pp. 52–67, Mar. 2019. 3, 4
- [46] Z. Wu, Y. Sun, A. Matlock, J. Liu, L. Tian, and U. S. Kamilov, “SIMBA: Scalable inversion in optical tomography using deep denoising priors,” *IEEE Journal of Selected Topics in Signal Process.*, vol. 14, no. 6, pp. 1163–1175, Mar. 2020. 3, 4
- [47] Y. Sun, J. Liu, and U. S. Kamilov, “Block coordinate regularization by denoising,” in *Advances in Neural Inf. Process. Systems 33*, Vancouver, BC, Canada, Dec. 8-14, 2019. 3, 4
- [48] J. Liu, Y. Sun, C. Eldeniz, W. Gan, H. An, and U. S. Kamilov, “RARE: Image reconstruction using deep priors learned without ground truth,” *IEEE J. Select. Topics Signal Process.*, 2020. 3
- [49] T. Donath, F. Beckmann, and A. Schreyer, “Automated determination of the center of rotation in tomography data,” *Journal of the Optical Society of America. A, Optics, Image science, and Vision*, vol. 23, pp. 1048–57, Jun. 2006. 3
- [50] N. T. Vo, M. Drakopoulos, R. Atwood, and C. Reinhard, “Reliable method for calculating the center of rotation in parallel-beam tomography,” *Optics Express*, vol. 22, pp. 19 078–19 086, Jul. 2014. 3
- [51] J. Wang and X. Gu, “Simultaneous motion estimation and image reconstruction (smeir) for 4d cone-beam ct,” *Medical physics*, vol. 40, no. 10, p. 101912, 2013. 3
- [52] P. A. Penczek, R. A. Grassucci, and J. Frank, “The ribosome at improved resolution: New techniques for merging and orientation refinement in 3D cryo-electron microscopy of biological particles,” *Ultramicroscopy*, vol. 53, no. 3, pp. 251 – 270, 1994. 3
- [53] T. S. Baker and R. H. Cheng, “A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy,” *Journal of Structural Biology*, vol. 116, no. 1, pp. 120 – 130, 1996. 3
- [54] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*. IEEE, 1988. 3
- [55] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*. Kluwer Academic Publishers, 2004. 4
- [56] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. 4
- [57] C. McCollough, “TU-FG-207A-04: Overview of the low dose CT grand challenge,” *Med. Phys.*, vol. 43, no. 6Part35, pp. 3759–3760, 2016. 5
- [58] A. Baydin, B. Pearlmutter, A. Radul, and J. Siskind, “Automatic differentiation in machine learning: A survey,” *Journal of Machine Learning Research*, vol. 18, pp. 1–43, Apr. 2018. 5
- [59] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr 2004. 5
- [60] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2015. 5
- [61] D. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *Int. Conf. Learning Representations (ICLR)*, 2015. 5
- [62] J. Zhang, T. He, S. Sra, and A. Jadbabaie, “Why gradient clipping accelerates training: A theoretical justification for adaptivity,” in *Proc. Int. Conf. on Learn. Represent.*, 2020. 6
- [63] C. Cheng, Y. Ching, P. Ko, and Y. Hwu, “Correction of center of rotation and projection angle in synchrotron x-ray computed tomography,” *Scientific reports*, vol. 8, no. 1, pp. 1–9, 2018. 8
- [64] W. Wan and J. Briggs, “Cryo-electron tomography and subtomogram averaging,” *Methods in enzymology*, vol. 579, pp. 329–367, 2016. 8