Low-resolution Prior Equilibrium Network for CT Reconstruction

To cite this article before publication: Yijie Yang *et al* 2024 *Inverse Problems* in press https://doi.org/10.1088/1361-6420/ad5d0d

Low-resolution Prior Equilibrium Network for CT Reconstruction

Yijie Yang[1], **Qifeng Gao**[1] **and Yuping Duan**[2*]

[1] Center for Applied Mathematics, Tianjin University, Tianjin, 300072, China

[2] School of Mathematical Sciences, Beijing Normal University, Beijing, 100875, China

[*] Corresponding author

E-mail: ypduan@bnu.edu.cn

Abstract. The unrolling method has been investigated for learning variational models in X-ray computed tomography. However, for incomplete data reconstruction, such as sparse-view and limited-angle problems, the unrolling method of gradient descent of the energy minimization problem cannot yield satisfactory results. In this paper, we present an effective CT reconstruction model, where the low-resolution image is introduced as a regularization for incomplete data problems. In what follows, we utilize the deep equilibrium approach to unfolding of the gradient descent algorithm, thereby constructing the backbone network architecture for solving the minimization model. We theoretically discuss the convergence of the proposed low-resolution prior equilibrium model and provide the necessary conditions to guarantee its convergence. Experimental results on both sparse-view and limited-angle reconstruction problems are provided, demonstrating that our end-to-end low-resolution prior equilibrium model outperforms other state-of-the-art methods in terms of noise reduction, contrast-to-noise ratio, and preservation of edge details.

Keywords: CT reconstruction, limited-angle, sparse-view, deep equilibrium model, unrolling, low-resolution image prior

1. Introduction

Computed Tomography (CT) is a fundamental imaging tool that finds wide applications in various fields, such as industrial non-destructive testing, medical diagnoses, and security inspections. The CT reconstruction problem aims to reconstruct clean image $\boldsymbol{u} \in \mathcal{X} \subset \mathbb{R}^N$ from the projection data $\boldsymbol{b} \in \mathcal{Y} \subset \mathbb{R}^l$ with unknown noise $\xi \in \Xi \subset \mathbb{R}^l$. The linear inverse problem in imaging is associated with a fixed linear operator $\boldsymbol{A} : \mathcal{X} \to \mathcal{Y}$. In detail, for any given pair $(\boldsymbol{b}, \xi) \in \mathcal{Y} \times \Xi$ the corresponding forward model generates the observations $\boldsymbol{b}$ via

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{u} + \xi. \tag{1}$$

where the reconstruction space $\mathcal{X}$ and the data space $\mathcal{Y}$ are typically Hilbert Spaces, and N represents the number of pixels. Our objective is to reconstruct the clean image

$\boldsymbol{u}$ from the measurement $\boldsymbol{b}$. In general, when the number of pixels in a reconstructed image exceeds the number of projection samples in CT imaging, the inverse problem of (1) becomes ill-posed.

Due to the radiation, various approaches have been explored, which can be primarily categorized into two main categories, either to modify the scanning protocol by reducing the tube voltage and/or tube current [44] or downsample the measured data for CT reconstruction, such as interior CT [36, 51, 60, 70], and sparse-view CT [68, 66, 34, 63, 72, 65]. These approaches aim to achieve dose reduction while maintaining satisfactory image quality. However, employing the first category of approaches for radiation dose reduction may result in the measured CT data being heavily contaminated with noise. The standard filtered back-projection (FBP) method, when applied to reconstruct CT images using noise-contaminated CT data, suffers from a significant degradation in image quality. On the other hand, the second category of approaches, aimed at reducing radiation dose, usually introduces additional ill-posedness into the inverse problem. It is worth noting that the regularization method is an effective technique for tackling ill-posed problems by incorporating prior information. The total variation (TV) regularization, as a prominent example, is considered state-of-the-art in handling low-dose and few-view CT [38]. Moreover, alternative sparsity regularization techniques, e.g., the utilization of wavelet frames [69], have been investigated in CT reconstruction. A noise suppression-guided image filtering reconstruction algorithm has recently been proposed as a solution to address the low signal-to-noise ratio (SNR) problem [31].

With recent advancements in artificial intelligence and hardware performance, deep learning (DL) has emerged as a promising approach for denoising in low-dose computed tomography, which has demonstrated encouraging results. Since then, various deep learning reconstruction methods have been studied including the pre-processing in the projection domain, projection-to-image reconstruction, and post-processing in the image domain [19, 25, 28, 2, 46, 49, 50]. These studies have shown that DL approaches consistently deliver enhanced or comparable noise suppression and maintain structural fidelity [57, 37, 54, 59, 3]. Broadly speaking, we can roughly divide the learning-based methods for CT reconstruction into three categories as the post-processing methods [56, 40], the plug-and-play priors methods [61, 10, 55, 48, 52] and the unrolling methods [4, 22]. There methods have greatly improved the reconstruction qualities compared to conventional methods. The remaining issues mainly include: firstly, the theoretical property guarantees of learning-based reconstruction methods have not been fully addressed, such as the convergence of the learning-based models. Secondly, for severely ill-posed reconstruction problems, such as limited-angle and sparse-view reconstruction issues, it still requires effective learning-based solutions.

1.1. Convergence studies of learning-based methods

Post-processing methods. A deep neural network-based post-processing of a

model-based reconstruction is parameterized as $\mathcal{P}_\theta \circ \mathcal{B}$, where $\mathcal{B}$ denotes a classical reconstruction method (e.g., FBP), $\mathcal{P}_\theta$ represents a deep convolutional network with parameters $\boldsymbol{\theta}$ and $\circ$ represents composite operations. In general, there is no convergence guarantee for the post-processing strategies. That means, a small value of $\|\boldsymbol{A}\boldsymbol{u}-\boldsymbol{b}\|$ does not necessarily imply a small value of $\|\boldsymbol{A}\mathcal{P}_\theta(\boldsymbol{u})-\boldsymbol{b}\|$ for the output of $\mathcal{P}_\theta$. Such issue was addressed in [56] by parametrizing the operator $\mathcal{P}_\theta$ as $\mathcal{P}_\theta = \boldsymbol{I} + \left(\boldsymbol{I} - \boldsymbol{A}^\top \boldsymbol{A}\right) \mathcal{Q}_\theta$, where $\mathcal{Q}_\theta$ is a Lipschitz-continuous deep convolutional neural network. Since $\boldsymbol{I} - \boldsymbol{A}^\top \boldsymbol{A}$ is the projection operator in the null-space of $\boldsymbol{A}$, the operator $\mathcal{P}_\theta$ always satisfies $\boldsymbol{A}\mathcal{P}_\theta(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u}$. Null-space networks were demonstrated to offer convergent regularization schemes. Inspired by the theory of optimal transport, an adversarial framework for learning the regularization term was proposed in [40], which possessed the stability guarantee, subjecting to the condition that the regularization is 1-Lipschitz and coercive.

Plug-and-play (PnP) priors methods. The PnP prior approach leverages the strengths of both realistic modeling through physical knowledge and flexible learning from data patterns. Relying on the *maximum a posteriori probability* (MAP) estimator, the solution of (1) can be formulated as the following minimization problem

$$\boldsymbol{u} \in \arg\min_{\boldsymbol{u}} \mathcal{S}(\boldsymbol{u}, \boldsymbol{b}) + \mathcal{R}(\boldsymbol{u}; \boldsymbol{\theta}), \tag{2}$$

where $\mathcal{S}(\boldsymbol{u}, \boldsymbol{b}) = -\log(p_{\boldsymbol{b}|\boldsymbol{u}}(\boldsymbol{u}))$ is the likelihood relating the solution $\boldsymbol{u}$ to the measurement $\boldsymbol{b}$ and $\mathcal{R}(\boldsymbol{u}; \boldsymbol{\theta}) = -\log(p_{\boldsymbol{u}}(\boldsymbol{u}))$ is the denoiser to be learned by deep neural network models. The PnP priors framework models the likelihood by obeying the noise distributions of the underlying physical process, while it extracts flexible priors from data observation. The learning-based regularization have shown excellent empirical performance, which inspired the theoretical studies on its convergence. The first result demonstrating the global objective convergence of PnP-ADMM was presented in [61], where the learned denoiser $\mathcal{R}(\boldsymbol{u}; \boldsymbol{\theta})$ should satisfy two requirements, i.e., it is continuously differentiable, and its gradient matrix is doubly stochastic. The fixed-point convergence of PnP-ADMM with a continuation scheme and bounded denoiser was established in [10]. Ryu *et al.* [55] demonstrated that the iterations of both PnP-proximal gradient-descent and PnP-ADMM exhibit contraction behavior when the denoiser $\mathcal{R}(\boldsymbol{u}; \boldsymbol{\theta})$ satisfies the Lipschitz condition. The fixed-point convergence of PnP-forward-backward splitting (FBS) and PnP-ADMM were proven in [48]. Specifically, the proof was established for linear denoisers of the form $\mathcal{R}(\boldsymbol{u}; \boldsymbol{\theta}) = \boldsymbol{W}\boldsymbol{u}$, where $\boldsymbol{W}$ is diagonalizable and its eigenvalues are within the range of [0, 1]. The convergence guarantees for the PnP methods were derived in [33] with gradient-step (GS), alleviating the need for such restrictive assumptions. The GS denoiser, denoted as $\mathcal{R}(\boldsymbol{u}; \boldsymbol{\theta}) = \boldsymbol{I} - \nabla g_\sigma$, is formulated with $g_\sigma(\boldsymbol{u}) = |\boldsymbol{u} - P_\sigma(\boldsymbol{u})|$, where P_σ represents a deep neural network devoid of structural constraints. The parametrization was shown to have enough expressive power to achieve state-of-the-art denoising performance in [33]. In [52], a closely related PnP approach was employed with the objective of providing an asymptotic characterization of the iterative PnP solutions. The key concept was to model maximally monotone operators (MMO) using a deep neural network, where the

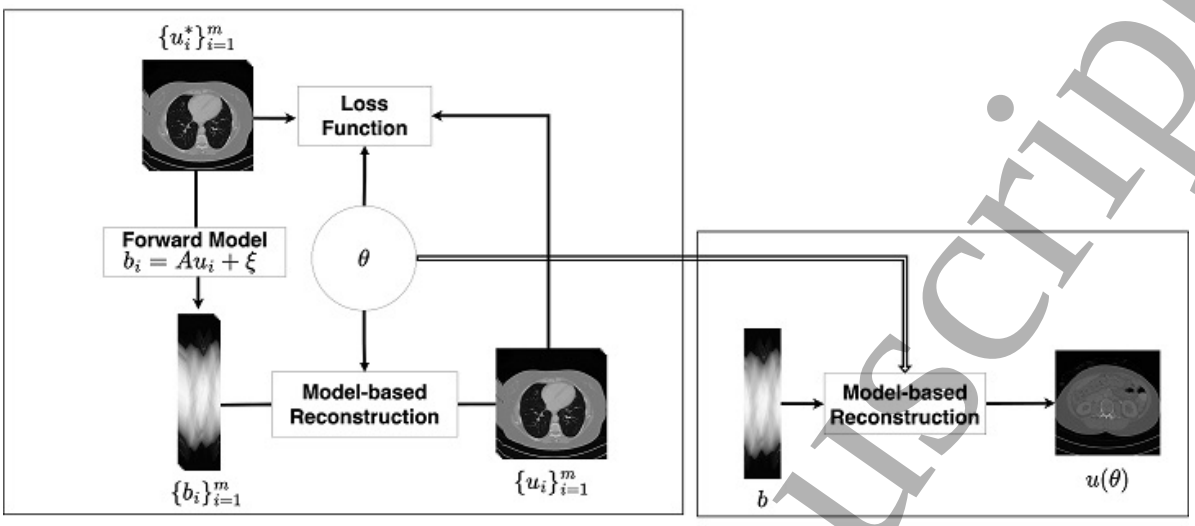


Figure 1. Depiction of a typical bilevel problem for image reconstruction. The left box represents the training process, which includes an upper-level loss and a lower-level cost function. During training, the objective is to minimize the upper-level loss function. Once the parameters $\boldsymbol{\theta}$ are learned, it is employed in the same image reconstruction task, as depicted in the right box.

parameterization of MMOs was achieved by modeling the resolvent through a non-expansive neural network.

Unrolling methods. Algorithm unrolling [47] unfolds the iterative algorithms with loops or recursion into non-recursive neural networks to construct interpretable and effective deep learning models. The origin of unrolling can be traced back to the seminal work by Gregor and LeCun [26] for solving sparse coding via unfolding the iterative soft-thresholding algorithm. We can formally express the unrolling model for (2) as a bilevel minimization problem

$$\begin{aligned} &\min_{\boldsymbol{\theta}} \ \frac{1}{m}\sum_{i=1}^{m} \ell\left(\boldsymbol{u}_i(\boldsymbol{\theta}), \boldsymbol{u}_i^*\right) && \text{(UL)} \\ &\text{s.t.} \ \ \boldsymbol{u}_i(\boldsymbol{\theta}) \in \arg\min_{\boldsymbol{u}} \mathcal{S}\left(\boldsymbol{u}, b_i\right) + \mathcal{R}(\boldsymbol{u}; \boldsymbol{\theta}) && \text{(LL)} \end{aligned} \tag{3}$$

where a loss function $\ell(\cdot,\cdot)$ is employed to enforce the similarity between the reconstructed image $\boldsymbol{u}_i$ and ground truth $\boldsymbol{u}_i{}^*$, $\mathcal{S}$ represents the data term, and $\mathcal{R}$ denotes the learnable regularization term, m indicates the number of training data samples. The upper-level (UL) loss function evaluates the quality of a vector of learnable parameters, while it also depends on the solution to the lower-level (LL) cost function. Figure 1 illustrates a generic bilevel problem for image reconstruction, where the model-based reconstruction represents the process of solving the lower-level minimization problem.

The bilevel problems [5, 7, 13, 16, 15] are known for their inherent difficulties to be solved numerically. To utilize gradient descent methods, it is necessary to compute the derivative of the solution operator for the lower-level problem (2) with respect to the parameters $\boldsymbol{\theta}$. If the objective of the lower-level problem is differentiable and

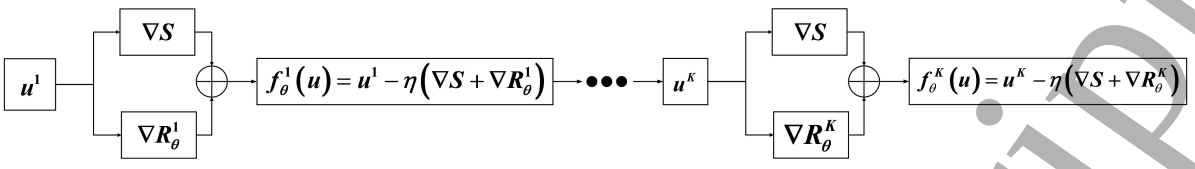


Figure 2. Depiction of a typical unrolling method using the gradient decent.

has a unique minimum, gradients can be computed using implicit differentiation [18]. However, non-smooth objectives, along with potentially multi-valued solution operators, are frequently employed in image and signal processing tasks. The method of unrolling involves utilizing an iterative algorithm that solves the lower-level problem and replaces the optimal solution $\boldsymbol{u}_i(\boldsymbol{\theta})$ with the K stages. We can express the unrolling network as follows

$$\boldsymbol{u}^{k+1} = f_\theta^k\left(\boldsymbol{u}^k\right) \quad \text{for} \quad k = 0, \ldots, K-1, \tag{4}$$

where k is the stage index and $f_\theta^k(\cdot)$ represents a nonlinear transformation, such as convolution followed by the application of a nonlinear activation function.

In fact, different stages can share the same weights, such that $f_\theta^k(\cdot)$ at each stage remains the same, i.e., $f_\theta^k(\cdot) = f_\theta(\cdot)$ for all k. The equilibrium or fixed-point-based networks introduced in [4, 22] used weight-sharing into the unrolling methods, which has demonstrated the capability to attain competitive performance. The main concept was to represent the output of a feed-forward neural network model as a fixed-point of a nonlinear transformation, which enabled the use of implicit differentiation for the back-propagation. Let us consider the K-stage network model with the input $\boldsymbol{b}$ and weights $\boldsymbol{\theta}$. The output of the $(k+1)$-th hidden layer can be denoted as $\boldsymbol{u}^{k+1}$ obtained by the following recursion process

$$\boldsymbol{u}^{k+1} = f_\theta\left(\boldsymbol{u}^k\right). \tag{5}$$

The limit of $\boldsymbol{u}^k$ as $K \to \infty$, provided it exists, is a fixed point of the operator $f_\theta(\cdot)$. The series converges if the spectral norm of the Jacobian $\partial_u f_\theta(\cdot)$ is strictly less than 1 for any $\boldsymbol{u}$, which is true if and only if the iteration map $f_\theta(\cdot)$ is contractive. The deep equilibrium model serves as a bridge between the conventional fixed-point methods in numerical analysis and learning-based techniques for solving the inverse problems. We illustrate the general diagram of the unrolling algorithm based on gradient descent of the low-level minimization problem in Figure 2, while represent the structure of deep equilibrium model in Figure 3. As can be shown, the difference between deep unrolling model and deep equilibrium model lies in the iterative regime, where deep equilibrium model shares the weights for all stage. Indeed, we use the deep neural networks to learn the gradient operator of the data fidelity and regularization used in the gradient descent scheme.

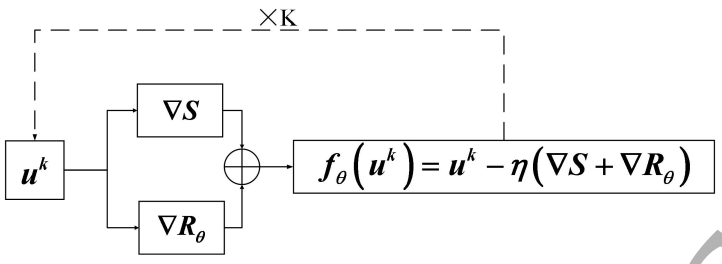


Figure 3. Depiction of a typical deep equilibrium method using the gradient decent.

1.2. Low-resolution Prior for CT reconstruction

For severely degraded CT reconstruction problems such as sparse-view and limited-angle, it is necessary to combine effective prior information to improve the quality of the reconstructed images. Different kinds of prior knowledge have been investigated for incomplete data reconstruction problems. The non-local similarity has been shown highly beneficial for enhancing fine details in reconstructed CT images [32, 35, 21, 45]. Both low-rank prior and dictionary learning have been used to leverage the sparsity of the image in a transform domain [41, 11, 53]. Deep learning methods by integrating geometric priors have also been studied in [58, 39, 64, 71]. Besides, personalized information as a prior has also been demonstrated to significantly enhance the reconstruction results [20, 24].

The ill-posedness of CT problem (1) comes from the finite dimension of measured rays and the infinite-dimensional nature of unknown objects. We can parameterize the object by utilizing a finite-series expansion as follows

$$\boldsymbol{u}(x,y) = \sum_{j=1}^{N} u_j e_j(x,y), \tag{6}$$

where $\{e_j(x,y)\}$ are basis functions that should be linearly independent. The square pixels are the most commonly used basis function defined by

$$e_j(x,y) = rect\Big(\frac{x-\tilde{x}_j}{\Delta x}\Big) rect\Big(\frac{y-\tilde{y}_j}{\Delta y}\Big),$$

where

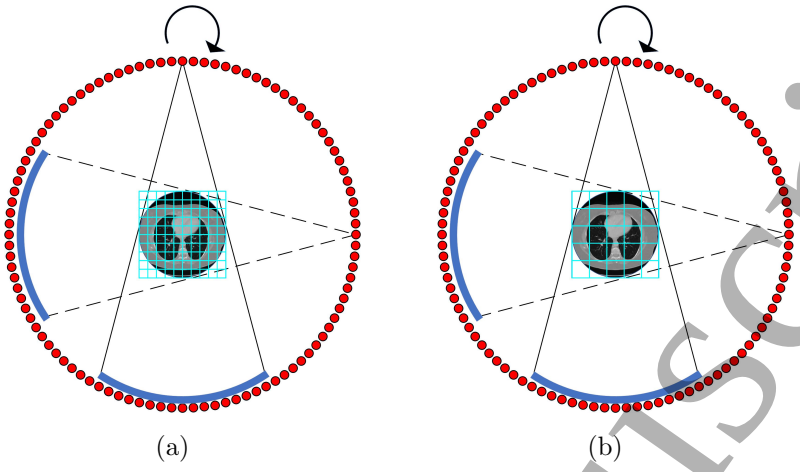
$$rect(t) \triangleq 1_{\{|t|\le 1/2\}} = \begin{cases} 1, & |t| \le 1/2, \\ 0, & \text{otherwise}, \end{cases}$$

and $(\tilde{x}_j,\tilde{y}_j)$ denotes the coordinates of the j-th pixel, $(\Delta x,\ \Delta y)$ denotes the size of each pixel. The length of the ith ray passing through the pixel u_j is denoted by $a_{ij} = \int e_j(x,y)\mathrm{d}\ell$, where all a_{ij} constitute the system matrix $\boldsymbol{A}$ in (1). It is well-known that using a finer grid can improve image resolution, but it also leads to an increase in the number of unknowns, which exacerbates the ill-posedness of the inverse problem and increases sensitivity to disturbances [27]. We can achieve the goal of reducing the number of unknowns by increasing the size of pixels/voxels through modifying the basis functions (6). Actually, there have been studies on the impact of resolutions on CT

reconstruction. In [62], a low-resolution full-view image was initially reconstructed, followed by a fine-resolution reconstruction of a specific region of interest (ROI). In particular, the image attributes of the pixels inside the ROI were set to zero, and a new projection sinogram was calculated using the forward model represented in matrix form by (1). Afterward, the newly calculated sinogram was subtracted from the original sinogram, resulting in a sinogram that contained projections exclusively corresponding to the ROI. Dabravolski *et al.* [14] proposed to obtain an initial reconstruction of a global coarse image, which was subsequently divided into regions comprising both fixed and non-fixed pixels. Specifically, the boundary pixels that shared at least one adjacent pixel with a different attenuation coefficient were assigned to the non-fixed pixels region. The non-fixed pixels region was comprised of boundary pixels that had at least one adjacent pixel with a different attenuation coefficient. Another method in [9] involved dividing the image domain into sub-regions with varying levels of discretization, which were characterized by either a coarse or fine pixel size. The projection data obtained from both resolutions were combined to reconstruct an image with different levels of discretization, effectively incorporating information from both measurements to enhance the reconstructed image quality. Gao *et al.* [17] proposed a low-resolution image prior image reconstruction model, which was realized in an end-to-end regime showing impressive ability in dealing with limited-angle reconstruction problems. He *et al.* [29] introduced a downsampling imaging geometric modeling approach for the data acquisition process, which integrated the geometric modeling knowledge of the CT imaging system with prior knowledge obtained from a data-driven training process to achieve precise CT image reconstruction. Malik *et al.* [42] suggested employing image fusion to generate a composite image that combines the high spatial resolution of a partially ambiguous image obtained from incomplete data with the more dependable quantitative information of a coarser image reconstructed from the same data in an over-complete problem. The aforementioned methods demonstrate that the low-resolution image can be used as an effective prior for ill-posed CT reconstruction problems.

1.3. Our contributions

In this work, we introduce the low-resolution image as the prior and propose a novel image reconstruction model for sparse-view and limited-angle CT reconstruction tasks. More specifically, the low-resolution image prior can introduce effective regularity into the reconstruction model to guarantee the qualities for incomplete data problems. Subsequently, we implement the algorithm unrolling to solve our low-resolution prior image model, where the gradient descent is employed for minimizing the lower-level problem with each Jacobi block being approximated by convolutional neural networks. To balance feature extraction capability and model size, we establish the low-resolution prior equilibrium (LRPE) network, where the weight-sharing strategy is used among all stages. More importantly, our approach has provable convergence guarantees by satisfying certain conditions to guarantee the iterative scheme converge to a fixed-point.



(a) (b)

Figure 4. Illustration of the virtual CT scanning system, where (a) a fine CT scanning system; (b) a coarse CT scanning system.

Through extensive numerical experiments on both sparse-view and limited-angle CT reconstruction problems, our LRPE model demonstrates clear advantages over the state-of-the-art learning methods. To sum up, the major contributions of this paper are as follows

- By leveraging the characteristics of the physical process of CT imaging, we use the low-resolution image as prior and establish an effective regularization model for incomplete data CT reconstruction.
- We employ the deep equilibrium method to solve the low-resolution prior model, where the convergence is established for both explicit data fidelity and learned data fidelity.
- We conduct numerical experiments to demonstrate the advantages of the low-resolution image prior and deep equilibrium strategy for incomplete data reconstruction problems.

2. Our Low-resolution Prior Reconstruction Model

Let $\boldsymbol{A_l} : \mathcal{X}_l \to \mathcal{Y}$ be the system matrix for the low-resolution image $\boldsymbol{u_l} \in \mathcal{X}_l \subset \mathbb{R}^n$. Thus, the low-resolution inverse problem can be formulated as follows

$$\boldsymbol{A_l u_l} = \boldsymbol{b}, \tag{7}$$

where n denotes the number of pixels of the low-resolution image. Without considering the discretization error, the same projection data $\boldsymbol{b}$ is used for the low-resolution image. The illustration of the fine and coarse CT scanning system is provided in Figure 4. As can be observed, the red color dots denote the permissible positions for the emitter, and the blue arc represents the corresponding receiving range of the receiver. It is obviously shown that the coarse scanning system has the same field of view (FOV) but with a

larger step size in spatial discretization. That means the low-resolution image can be obtained from the high-resolution image as follows

$$\boldsymbol{u}_l = \boldsymbol{D}\boldsymbol{u}, \tag{8}$$

where $\boldsymbol{D} : \mathcal{X} \to \mathcal{X}_l$ is the down-sampling operator. In this work, we concern with the following low-resolution image regularized CT reconstruction model

$$\boldsymbol{u} \in \arg\min_{\boldsymbol{u}} \mathcal{S}(\boldsymbol{A}\boldsymbol{u}, \boldsymbol{b}) + \mathcal{R}(\boldsymbol{u}, \boldsymbol{D}^\top \boldsymbol{u}_l), \tag{9}$$

where $\boldsymbol{D}^\top$ represents the transpose of the matrix $\boldsymbol{D}$. Supposing both $\mathcal{S}$ and $\mathcal{R}$ are differentiable, we can use the gradient descent to solve the minimization problem (9). That is, we start with an initial estimate $\boldsymbol{u}^0$ such as $\boldsymbol{u}^0 = 0$ and a step size $\eta > 0$, such that for stage $k = 0, 1, \ldots, K-1$, we estimate $\boldsymbol{u}^{k+1}$ from

$$\boldsymbol{u}^{k+1} = \boldsymbol{u}^k - \eta\Big(\boldsymbol{A}^\top \nabla \mathcal{S}\left(\boldsymbol{A}\boldsymbol{u}^k, \boldsymbol{b}\right) + \nabla \mathcal{R}\left(\boldsymbol{u}^k, \boldsymbol{D}^\top \boldsymbol{u}_l\right)\Big).$$

To leverage the rich information contained in the data, we adopt the deep unrolling strategy to solve the above gradient descent problem. More specifically, we utilize two convolutional neural networks $S_\theta(\cdot,\cdot) : \mathcal{Y} \times \mathcal{Y} \to \mathcal{Y}$ and $R_\theta(\cdot,\cdot) : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ to learn the gradient operator $\nabla \mathcal{S}\left(\boldsymbol{A}\boldsymbol{u}^k, \boldsymbol{b}\right)$ and $\nabla \mathcal{R}\left(\boldsymbol{u}^k, \boldsymbol{D}^\top \boldsymbol{u}_l\right)$, respectively. For $k = 0, \ldots, K-1$, we then have the recursive update as follows

$$\boldsymbol{u}^{k+1} = \boldsymbol{u}^k - \eta(\boldsymbol{A}^\top s^k + t^k), \tag{10}$$

where $s^k := S_\theta(\boldsymbol{A}\boldsymbol{u}^k, \boldsymbol{b})$ and $t^k := R_\theta(\boldsymbol{u}^k, \boldsymbol{D}^\top \boldsymbol{u}_l)$ are the estimations of the two convolutional neural network models, respectively. Moreover, we use the weight-sharing scheme to improve the feature abstraction capabilities by unrolling the network to a sufficient depth. Therefore, the final update scheme is achieved by connecting the unrolled gradient descent stages (10) with the deep equilibrium architecture (5) as follows

$$f_\theta(\boldsymbol{u}^{k+1}) := \boldsymbol{u}^k - \eta\Big(\boldsymbol{A}^\top S_\theta(\boldsymbol{A}\boldsymbol{u}^{\boldsymbol{k}}, \boldsymbol{b}) + R_\theta(\boldsymbol{u}^k, \boldsymbol{D}^\top \boldsymbol{u}_l)\Big). \tag{11}$$

The objective function in the upper-level problem is simply the true risk for the squared loss

$$\ell(\boldsymbol{u}, \boldsymbol{u}^*) = \frac{1}{m}\sum_{i=1}^{m} \|\boldsymbol{u}_i{}^* - \boldsymbol{u}_i{}^K\|_2^2,$$

where $\boldsymbol{u}_i{}^*$ is the ith ground truth, and $\boldsymbol{u}_i^K$ denotes the reconstructed image. Indeed, the use of the mean squared error (MSE) loss is not necessary, as any differentiable loss function can be used.

3. Our Algorithm

The low-resolution prior equilibrium (LRPE) model can be outlined as Algorithm 1. In both training and inference stage, the low-resolution prior image $\boldsymbol{u}_l$ is obtained by established reconstruction methods, where we used the Learned Primal-Dual network [1] to estimate the low-resolution image in our work.

Algorithm 1 Low-Resolution Prior Equilibrium (LRPE) Network

1: **Step 0.** Initialize $\boldsymbol{u}^0$, $\boldsymbol{b}$, $\boldsymbol{u_l}$, step size η;
2: **Step 1.** For $k = 0, 1, 2, \cdots, K-1$ do

$$\begin{cases} s^k \longleftarrow S_\theta(\boldsymbol{A}\boldsymbol{u}^k, \boldsymbol{b}); \\ t^k \longleftarrow R_\theta(\boldsymbol{u}^k, \boldsymbol{D}^\top \boldsymbol{u_l}); \\ \boldsymbol{u}^{k+1} = \boldsymbol{u}^k - \eta(\boldsymbol{A}^\top s^k + t^k); \end{cases}$$

3: **Step 2.** Return $\boldsymbol{u}^K$.

3.1. Network architecture

In Algorithm 1, both $R_\theta(\cdot,\cdot)$ and $S_\theta(\cdot,\cdot)$ are modeled by convolutional neural networks. More specifically, we use a three layer convolutional neural network with residual connections to express the two functionals

$$R_\theta = \boldsymbol{I}_N + H_3\phi_2(H_2\phi_1(H_1)),$$

and

$$S_\theta = \boldsymbol{I}_l + K_3\phi_2(K_2\phi_1(K_1)),$$

where the building blocks are defined as

- $H_i : \mathcal{X} \to \mathcal{X}, i = 1, \ldots, 3$, is the matrix representation of a learned 3×3 convolution kernel.
- $K_i : \mathcal{Y} \to \mathcal{Y}, i = 1, \ldots, 3$, is the matrix representation of a learned 3×3 convolution kernel.
- $\phi_i, i = 1, 2$ is a differentiable element-wise activation function. We adopt the log-student-t-distribution function as the activation function

$$\phi(x) = \frac{1}{2}\log(1 + x^2). \tag{12}$$

- $\boldsymbol{I}_N \in \mathbb{R}^{N\times N}$ and $\boldsymbol{I}_l \in \mathbb{R}^{l\times l}$ stand for the identity transformation.

The inclusion of residual structures [30] is important, which can help prevent the loss of fine details, expedite information flow, and enhance the capcity of the network model. The residual structures can alleviate training challenges and expedite convergence. The overall depth of the network is determined by the number of stages it contains, which is fixed to strike a balance between the receptive fields and the total number of parameters in the network. Therefore, we set the numbers of channels in each stage as $6 \to 32 \to 32 \to 5$ for both R_θ and S_θ. We present the network structure and subblocks in Figure 5, which is an end-to-end convlutional neural network model for CT reconstruction problems.

3.2. Algorithm implementation and network optimization

In implementation, our algorithm was realized using the Operator Discretization Library (ODL), the Adler package, the ASTRA Toolbox, and TensorFlow 1.8.0. In particular,

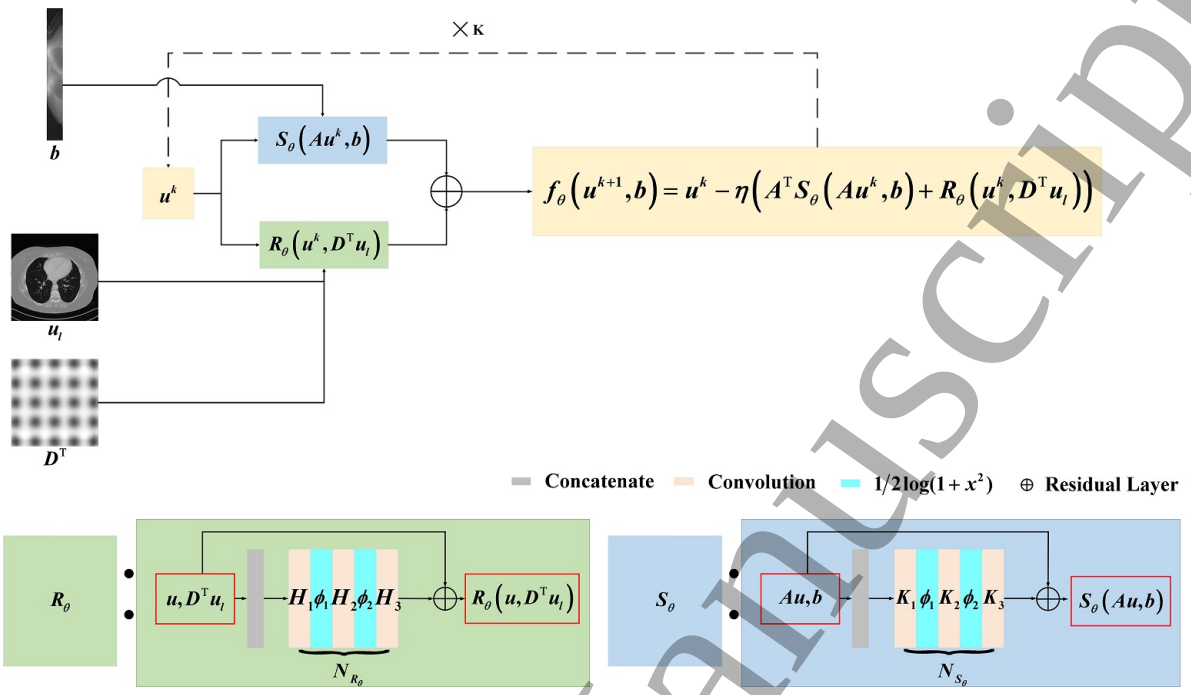


Figure 5. The network structure of our LRPE: Low-Resolution Prior Equilibrium network.

the TensorFlow is a toolkit designed for tackling complex mathematical problems, for which the calculations are represented as graphs, mathematical operations are nodes, and multidimensional data arrays are communicated as edges of the graphs. The ASTRA Toolbox is a MATLAB and Python toolbox that provides high-performance GPU primitives for 2D and 3D tomography. And ODL is a Python library focused on fast prototyping for inverse problems. Adler is a toolkit that facilitates the efficient implementation of neural network architectures.

In our network, the parameters are updated using the backpropagation algorithm within the stochastic gradient descent method implemented in TensorFlow. We employ the Xavier initialization scheme for the convolution parameters and initialize all biases to zero. The adaptive moment estimation (Adam) optimizer is chosen to optimize the learning rate with the parameter β set to 0.99 and the remaining parameters set to their default values. The learning rate follows a cosine annealing strategy, which helps to improve training stability. The initial learning rate is set to 10^{-4} and global gradient norm clipping is performed by limiting the gradient norm to 1. Besides, the batch size is set to 1 for all experiments.

4. Convergence Analysis

Here we study convergence of the proposed LRPE to a fixed-point at inference time. Given the iteration map $f_\theta(\cdot) : \mathcal{X} \to \mathcal{X}$, we discuss the conditions that can guarantee the convergence of the iterates $\boldsymbol{u}^{k+1} = f_\theta(\boldsymbol{u}^k)$ to a fixed-point $\boldsymbol{u}^\infty$ as $k \to \infty$. The

classical fixed-point theorem [23] ensures the convergence of the iterates to a unique fixed point if the iteration map is contractive, i.e. if exists a constant $0 \leq c < 1$ such that $\|f_\theta(\boldsymbol{u}) - f_\theta(\boldsymbol{u}')\| \leq c\,\|\boldsymbol{u} - \boldsymbol{u}'\|$ for all $\boldsymbol{u}$ and $\boldsymbol{u}'$.

Now we establish the conditions on R_θ and S_θ that can ensure the convergence of the iterates $\boldsymbol{u}^{k+1} = f_\theta\left(\boldsymbol{u}^k\right)$ in Algorithm 1 to a fixed-point $\boldsymbol{u}^\infty$ as the stage number k approaches infinity. These conditions guarantee that the resulting mapping satisfies the contractivity condition and, consequently, the fixed-point iteration converges. Relying on the network structure shown in Figure 5, we denote $N_{R_\theta}(\boldsymbol{u}) = H_3\phi_2(H_2\phi_1(H_1(\boldsymbol{u})))$ and $N_{S_\theta}(\boldsymbol{A}\boldsymbol{u}) = K_3\phi_2(K_2\phi_1(K_1(\boldsymbol{A}\boldsymbol{u})))$. Then we have

$$\|\partial_u N_{R_\theta}\| = \left\|H_1^\top \partial_u\phi_1(H_1\boldsymbol{u})H_2^\top \partial_u\phi_2(H_2\phi(H_1\boldsymbol{u}))H_3\right\|,$$
$$\|\partial_u N_{S_\theta}\| = \left\|\boldsymbol{A}^\top K_1^\top \partial_u\phi_1(K_1\boldsymbol{A}\boldsymbol{u})K_2^\top \partial_u\phi_2(K_2\phi(K_1\boldsymbol{A}\boldsymbol{u}))K_3\right\|.$$

Due to $\sup_{\boldsymbol{u}}|\phi'(\boldsymbol{u})| = 0.5$, $\|\partial_u N_{R_\theta}\|$ and $\|\partial_u N_{S_\theta}\|$ can be uniformly bounded independently of $\boldsymbol{u}$. Therefore, we can assume that N_{R_θ} and N_{S_θ} to be ϵ_1-Lipschitz and ϵ_2-Lipschitz. As aforementioned, there is $(R_\theta - \boldsymbol{I_N})(\boldsymbol{u}) = N_{R_\theta}(\boldsymbol{u})$ and $(S_\theta - \boldsymbol{I_l})(\boldsymbol{A}\boldsymbol{u}) = N_{S_\theta}(\boldsymbol{A}\boldsymbol{u})$. That means $(R_\theta - \boldsymbol{I_N})(\boldsymbol{u})$ and $(S_\theta - \boldsymbol{I_l})(\boldsymbol{A}\boldsymbol{u})$ are ϵ_1-Lipschitz and ϵ_2-Lipschitz, i.e.,

$$\|(R_\theta - \boldsymbol{I_N})(\boldsymbol{u}) - (R_\theta - \boldsymbol{I_N})(\boldsymbol{u}')\| \leq \epsilon_1\|\boldsymbol{u} - \boldsymbol{u}'\|,$$
$$\|(S_\theta - \boldsymbol{I_l})(\boldsymbol{A}\boldsymbol{u}) - (S_\theta - \boldsymbol{I_l})(\boldsymbol{A}\boldsymbol{u}')\| \leq \epsilon_2\|\boldsymbol{u} - \boldsymbol{u}'\|.$$

In the following, we discuss convergences for the deep equilibrium unrolling algorithm for the low-resolution prior reconstruction model with empirical data fidelity and learned data fidelity, respectively.

Theorem 1. *(Convergence of LRPE with empirical data fidelity). Assume that the observed data is without noise, there is $\mathcal{S} = \frac{1}{2}||\boldsymbol{A}\boldsymbol{u} - \boldsymbol{b}||_2^2$ in (9). Let $L = \lambda_{\max}\left(\boldsymbol{A}^\top\boldsymbol{A}\right)$ and $\mu = \lambda_{\min}\left(\boldsymbol{A}^\top\boldsymbol{A}\right)$, where $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ denote the maximum and minimum eigenvalue, respectively. If the step-size parameter $\eta > 0$ satisfies $\eta < 1/(L+1)$, we can obtain the following inequality*

$$\|f_\theta(\boldsymbol{u}) - f_\theta(\boldsymbol{u}')\| \leq \underbrace{(1 - \eta(1+\mu) + \eta\epsilon_1)}_{=:\gamma}\|\boldsymbol{u} - \boldsymbol{u}'\|$$

for all $\boldsymbol{u}, \boldsymbol{u}' \in \mathcal{X}$. The coefficient γ is less than 1 if $\epsilon_1 < 1+\mu$, in which case the iterates of the LRPE model converges.

Proof. Let $f_\theta(\boldsymbol{u})$ be the iteration map for the LRPE with empirical data fidelity. The Jacobian of $f_\theta(\boldsymbol{u})$ with respect to $\boldsymbol{u}$ denoted by $\partial_u f_\theta(\boldsymbol{u})$, is given by

$$\partial_u f_\theta(\boldsymbol{u}) = \left(\boldsymbol{I_N} - \eta\boldsymbol{A}^\top\boldsymbol{A}\right) - \eta\partial_u R_\theta(\boldsymbol{u}),$$

where $\partial_u R_\theta(\boldsymbol{u})$ is the Jacobian of $R_\theta(\boldsymbol{u})$ with respect to $\boldsymbol{u}$. To prove $f_\theta(\cdot)$ is contractive, it suffices to show $\|\partial_u f_\theta(\boldsymbol{u})\| < 1$ for all $\boldsymbol{u} \in \mathcal{X}$, where $\|\cdot\|$ denotes the spectral norm.

Towards this end, we have

$$\begin{aligned}\|\partial_u f_\theta(\boldsymbol{u})\| &= \left\|\boldsymbol{I_N} - \eta\boldsymbol{A}^\top\boldsymbol{A} - \eta\partial_u R_\theta(\boldsymbol{u})\right\| \\ &= \left\|\boldsymbol{I_N} - \eta\boldsymbol{A}^\top\boldsymbol{A} - \eta\boldsymbol{I_N} + \eta\boldsymbol{I_N} - \eta\partial_u R_\theta(\boldsymbol{u})\right\| \\ &\leq \left\|(1-\eta)\boldsymbol{I_N} - \eta\boldsymbol{A}^\top\boldsymbol{A}\right\| + \eta\left\|\partial_u R_\theta(\boldsymbol{u}) - \boldsymbol{I_N}\right\| \\ &\leq \max_i |(1-\eta) - \eta\lambda_i| + \eta\epsilon_1,\end{aligned} \tag{13}$$

where λ_i denotes the i th eigenvalue of $\boldsymbol{A}^\top\boldsymbol{A}$. Here, we use the assumption that the map $\left(R_\theta - \boldsymbol{I_N}\right)(\boldsymbol{u}) := R_\theta(\boldsymbol{u}) - \boldsymbol{u}$ is ϵ_1-Lipschitz to guarantee the spectral norm of the Jacobian $\partial_u R_\theta(\boldsymbol{u}) - \boldsymbol{I_N}$ to be bounded by ϵ_1. Based on the assumption $\eta < \frac{1}{1+L}$ with $L := \max_i \lambda_i$, we have $\eta < \frac{1}{1+\lambda_i}$ for all i, which implies $(1-\eta) - \eta\lambda_i > 0$ for all i. Therefore, the maximum in (13) is obtained at $\mu := \min_i \lambda_i$, which gives

$$\|\partial_u f_\theta(\boldsymbol{u})\| \leq 1 - \eta(1+\mu) + \eta\epsilon_1.$$

It shows f_θ is γ-Lipschitz with $\gamma = 1 - \eta(1+\mu) + \eta\epsilon_1$, proving the claim. □

Theorem 2. *(Convergence of LRPE with learned data fidelity). Assume that* $(R_\theta - \boldsymbol{I_N})(\boldsymbol{u})$ *is* ϵ_1*-Lipschitz and* $(S_\theta - \boldsymbol{I_l})(\boldsymbol{A}\boldsymbol{u})$ *is* ϵ_2*-Lipschitz, and let* $L = \lambda_{\max}\left(\boldsymbol{A}^\top\boldsymbol{A}\right)$ *and* $\mu = \lambda_{\min}\left(\boldsymbol{A}^\top\boldsymbol{A}\right)$*, where* $\lambda_{\max}(\cdot)$ *and* $\lambda_{\min}(\cdot)$ *denote the maximum and minimum eigenvalue, respectively. If the step-size parameter* $\eta > 0$ *obeys* $\eta < 1/(L+1)$*, the LRPE iteration map* $f_\theta(\cdot)$ *defined in (11) satisfies*

$$\|f_\theta(\boldsymbol{u}) - f_\theta(\boldsymbol{u}')\| \leq \underbrace{\left(1 - \eta(1+\mu) + \eta(L\epsilon_2 + \epsilon_1)\right)}_{=:\gamma}\|\boldsymbol{u} - \boldsymbol{u}'\|$$

for all $\boldsymbol{u}, \boldsymbol{u}' \in \mathcal{X}$*. The coefficient* γ *is less than 1 if* $\epsilon_1 + L\epsilon_2 < 1 + \mu$*, in which case the iterates of the LRPE-net converges.*

Proof. Let $f_\theta(\boldsymbol{u})$ be the iteration map for LRPE with learned data fidelity. The Jacobian of $f_\theta(\boldsymbol{u})$ with respect to $\boldsymbol{u} \in \mathcal{X}$, denoted by $\partial_u f_\theta(\boldsymbol{u})$, is given by

$$\partial_u f_\theta(\boldsymbol{u}) = \left(\boldsymbol{I_N} - \eta\boldsymbol{A}^\top\partial_u S_\theta(\boldsymbol{A}\boldsymbol{u})\boldsymbol{A}\right) - \eta\partial_u R_\theta(\boldsymbol{u}),$$

where $\partial_u R_\theta(\boldsymbol{u})$ is the Jacobian of $R_\theta(\boldsymbol{u})$ with respect to $\boldsymbol{u} \in \mathcal{X}$ and $\partial_u S_\theta(\boldsymbol{A}\boldsymbol{u})$ is the Jacobian of $S_\theta(\boldsymbol{A}\boldsymbol{u})$. Similarly, we show $\|\partial_u f_\theta(\boldsymbol{u})\| < 1$ for all $\boldsymbol{u} \in \mathcal{X}$. Since $\left(R_\theta - \boldsymbol{I_N}\right)(\boldsymbol{u})$ and $\left(S_\theta - \boldsymbol{I_l}\right)(\boldsymbol{A}\boldsymbol{u})$ are ϵ_1-Lipschitz and ϵ_2-Lipschitz, respectively, we have the spectral norm of its Jacobian $\partial_u R_\theta(\boldsymbol{u}) - \boldsymbol{I_N}$ is bounded by ϵ_1 and $\partial_u S_\theta(\boldsymbol{u}) - \boldsymbol{I_l}$ is bounded by ϵ_2. Then there is

$$\begin{aligned}\|\partial_u f_\theta(\boldsymbol{u})\| &= \left\|\boldsymbol{I_N} - \eta\boldsymbol{A}^\top\partial_u S_\theta(\boldsymbol{A}\boldsymbol{u})\boldsymbol{A} - \eta\partial_u R_\theta(\boldsymbol{u})\right\| \\ &= \left\|\boldsymbol{I_N} - \eta\boldsymbol{I_N} + \eta\boldsymbol{I_N} - \eta\boldsymbol{A}^\top\boldsymbol{A} + \eta\boldsymbol{A}^\top\boldsymbol{A} - \eta\boldsymbol{A}^\top\partial_u S_\theta(\boldsymbol{A}\boldsymbol{u})\boldsymbol{A} - \eta\partial_u R_\theta(\boldsymbol{u})\right\| \\ &= \left\|\boldsymbol{I_N} - \eta\boldsymbol{A}^\top\boldsymbol{A} - \eta\boldsymbol{I_N} - \eta\left(\partial_u R_\theta(\boldsymbol{u}) - \boldsymbol{I_N}\right) - \eta\boldsymbol{A}^\top(\partial_u S_\theta(\boldsymbol{A}\boldsymbol{u}) - \boldsymbol{I_l})\boldsymbol{A}\right\| \\ &\leq \left\|(1-\eta)\boldsymbol{I_N} - \eta\boldsymbol{A}^\top\boldsymbol{A}\right\| + \eta\left\|\partial_u R_\theta(\boldsymbol{u}) - \boldsymbol{I_N}\right\| + \eta\boldsymbol{A}^\top\left\|\partial_u S_\theta(\boldsymbol{A}\boldsymbol{u}) - \boldsymbol{I_l}\right\|\boldsymbol{A} \\ &\leq \max_i \|(1-\eta) - \eta\lambda_i\| + \eta\epsilon_1 + \eta L\epsilon_2,\end{aligned} \tag{14}$$

where λ_i denotes the i the eigenvalue of $\boldsymbol{A}^\top \boldsymbol{A}$. Finally, by our assumption $\eta < \frac{1}{1+L}$ where $L := \max_i \lambda_1$, we have $\eta < \frac{1}{1+\lambda_i}$ for all i, which implies $(1-\eta) - \eta\lambda_i > 0$ for all i. Therefore, the maximum in (14) is obtained at $\mu := \min_i \lambda_i$, which gives

$$\|\partial_u f_\theta(\boldsymbol{u})\| \leq 1 - \eta(1+\mu) + \eta\epsilon_1 + \eta L\epsilon_2.$$

It shows that f_θ is γ-Lipschitz with $\gamma = 1-\eta(1+\mu)+\eta(L\epsilon_2+\epsilon_1)$, proving the claim. □

5. Numerical Results

In this section, we evaluate our LRPE model on both sparse-view and limited-angle reconstruction problems and compare it with several state-of-the-art methods using a dataset of human phantoms. We utilize the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) as evaluation metrics to measure the quality of the reconstructed images produced by different methods.

5.1. Comparison algorithms

We employ several recent CT reconstruction methods for comparison studies, including both traditional and learning-based approaches, as described below:

- TV model: the TV regularized reconstruction model proposed in [8]. We tuned the balance parameter $\lambda \in [1, 3]$, the step size for the primal value τ within the range of [0.5, 0.9], and the step size for the dual value σ within the range of [0.2, 0.5]. These parameter settings were adjusted accordingly for different experiments.
- PD: the Learned Primal-Dual network in [1]. The network is a deep unrolled neural network with 10 stages. The number of initialization channels for both primal and dual values is set to 5. The network parameters are initialized using the Xavier initialization scheme. In all experiments, we employed the mean squared loss as the objective function, which measures the discrepancy between the reconstructed image and the ground truth.
- SIPID: the Sinogram Interpolation and Image Denoising (SIPID) network presented in [67]. The SIPID network utilizes a deep learning framework and achieves accurate reconstructions by iteratively training the sinogram interpolation network and the image denoising network. The network parameters are initialized using the Xavier initialization scheme, and the mean squared loss is employed as the objective function in all experiments.
- FSR: the Learned Full-Sampling Reconstruction From Incomplete Data in [12]. The FSR-Net is an iterative expansion method that uses the corresponding full sampling projection system matrix as prior information. They employed two separate networks, namely IFSR and SFSR. Specifically, the IFSR and SFSR are utilized for reconstructions using IFS and SFS system matrices, respectively. The number of initialization channels for primal values and dual values is set as 6 and

7, respectively. The loss function is the mean square error of the image domain and the Radon domain with the weight α being 1.

- LRIP: the Low-Resolution Image Prior based Network in [17]. It is a low-resolution image prior image reconstruction model for the limited-angle reconstruction problems. The number of initialization channels is set as 5 for both primal and dual values. The loss function used in training is a combination of MSE and SSIM calculated in the image domain. The weight parameter α for balancing the two components of the loss function is set to 1.
- GRAD: gradient descent is used instead of the primal-dual algorithm to solve the low-level problem described in model (9). It does not incorporate low-resolution image prior information. The network is a deep unrolled neural network with 10 stages. The number of initialization channels is set to 5. The Xavier initialization and the mean squared loss of the reconstructed image and the ground truth are used in all experiments.

5.2. Datasets and settings

In the experiments, we utilize the clinical dataset known as "The 2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge" [43]. This dataset comprises 10 full-dose scans of the ACR CT accreditation phantom. To establish the training dataset, we select 9 of these scans, reserving the remaining 1 scan for evaluation purposes. Consequently, the training dataset consists of 2164 images, each with dimensions of 512×512, while the evaluation dataset comprises 214 images. We set the scanning angular interval to be 1 degree. To assess the performance of the reconstruction methods, we introduce various types of noises into the projected data, which allows us to validate and compare the effectiveness of the different reconstruction techniques.

5.3. Test Settings and Parameter Choice

In this subsection, we evaluate the choices of the parameters to the performance of our algorithm. There are several important parameters in Algorithm 1, including the number of stages K, the step size η and the dimension of the variables in the network model.

We first study the influence of the stage number K on the convergence of the algorithm. We use the squared error $\|\boldsymbol{u}^K - \boldsymbol{u}^*\|^2$ to measure the data error during the training process. Both the results of 150° limited-angle and 60 sparse-view reconstruction with and without 5% Gaussian noises are presented in Figure 6. As we can see, our LRPE model demonstrates convergence as the number of epochs increases, regardless of whether there are 5, 10, or 15 stages (iterations). As Theorem 4.8 in [6] indicates, convergence to the optimal value can be achieved when the data itself is noise-free. If the observed data contains noises, the expected objective value of the fixed-point iteration converges to a neighborhood of the optimal value. We can observe from the

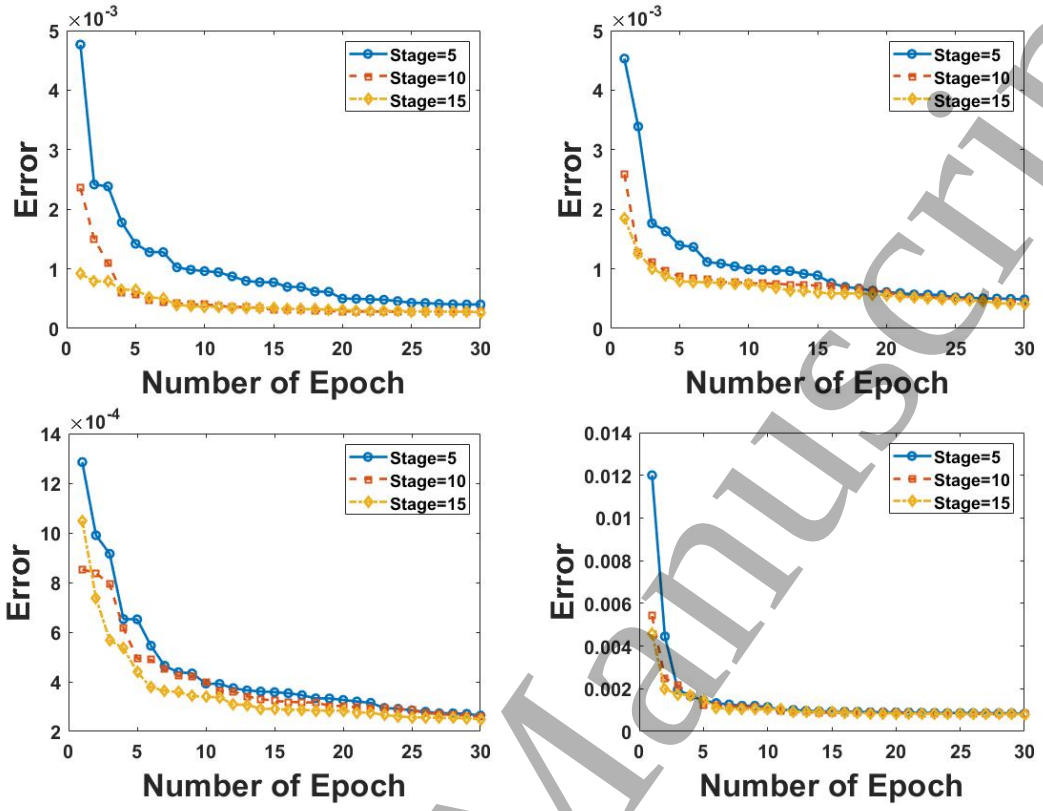


Figure 6. The data error measured by L2 norm with respect to the numbers of epochs in our LRPE model. The curves represent the results for 150° limited-angle (First row)and 60 sparse-view (Second row) with and without 5% Gaussian noises. On each row, the left one is the result without noises, and the right one is the result with 5% Gaussian noises.

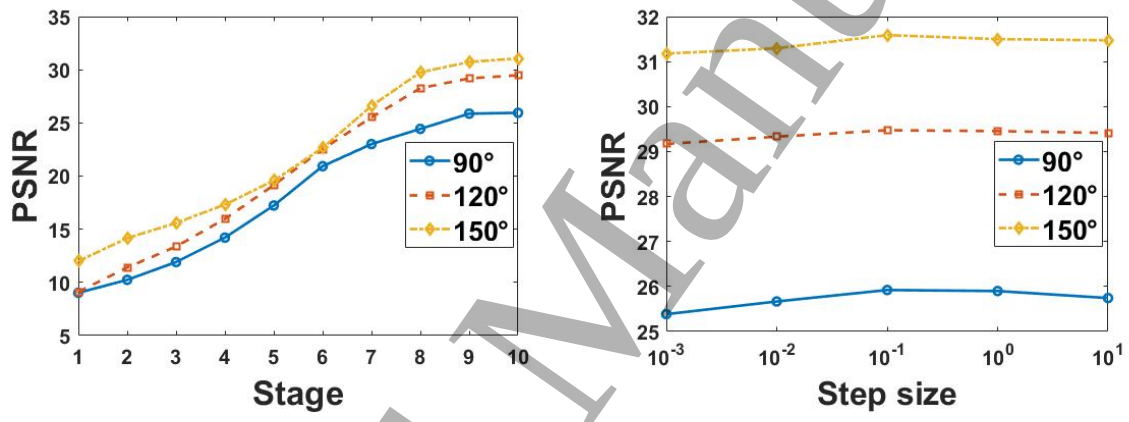
numerical error of our LRPE method converges to lower error for cases without noises. When noises are introduced into observed data, the solution converges to much higher error, which is consistent with the theoretical analysis results in [6]. With the same number of iterations, more stages can optimize to a lower data error faster. Since there is a trade-off between the error reduction and training efficiency, we fix the total number of stages to 10 to achieve a good balance.

On the other hand, we also conducted experiments with different numbers of stages during the inference phase, where the LRPE models trained with 10 stages are used in the experiment. As can be observed in Figure 7(a), as the number of stages increases, the PSNR generally improves for all the limited-angle reconstruction problems. It is worth noting that by using a parameter-sharing strategy in the deep equilibrium model, the number of parameters is significantly reduced compared to the learned PD model. Specifically, the number of parameters in our model is only 1/10th of the number of parameters in the learned PD model. In what follows, we analyze the effect of the hyper-parameter η on the performance of limited-angle reconstruction. We perform a parameter sweep and obtain the results shown in Figure 7(b), which models are used 10 stages. From the results, we observe that the highest PSNR value is achieved when

Table 1. The performance of our LRPE was evaluated in terms of PSNR across different values of N_p.

Limited-angle \ N_p	1	3	4	5	6	7
90°	25.6472	25.8506	25.9506	25.9517	25.9129	**25.9565**
120°	29.1946	29.2797	29.4326	**29.4835**	29.4718	29.4815
150°	31.2969	31.4487	31.5884	**31.636**	31.4477	31.3321

we set η as $\eta = 0.1$. Therefore, in all other experiments, we fix the hyper-parameter η to a value of 0.1.



(a) Results of the different stages (b) Results of the different step size

Figure 7. Evaluation results of our LRPE model with respect to different settings of parameters during inference process.

Expanding the variable space is a common technique used in network optimization to improve stability during the training process. In our case, we expand the variable space by considering multiple variables $\boldsymbol{u} = [\boldsymbol{u}^{(1)}, \boldsymbol{u}^{(2)}, \ldots, \boldsymbol{u}^{(N_p)}]$. In Table 1, we investigate the influence of different choices of N_p on the reconstruction accuracy using limited-angle data. From the perspective of results obtained from various views in Table 1, we fix the value of N_p to 5 for all experiments, which provides a good balance between reconstruction accuracy and computational efficiency.

5.4. Empirical data fidelity or learned data fidelity?

We evaluate the differences between the empirical data fidelity and learned fidelity on 150° limited-angle reconstruction problems, where the empirical one is chosen as the L2 norm following the assumption in Theorem 1 and the learned one following the assumption in Theorem 2. Table 2 presents the PSNR and SSIM values for both cases, for which the raw data are corrupted by 5% Gaussian noises and Poisson noises with up to 100, 1000, and 10000 incident photons per pixel before attenuation, respectively. As

can be seen, when the mixed noises are contained in the projection data, the model with the learned fidelity outperforms the model with the empirical fidelity. It is attributed to the fact that the L2 norm measures the Euclidean distance, which is not ideal for complex noise distributions in mixed noises. When salt-and-pepper noises are introduced into the measured data, we can observe similar results as shown in Table 3. To sum up, the learned data fidelity performs better than the empirical data fidelity term, especially when the noises become significant.

Table 2. Evaluation results on the 150° limited-angle reconstruction problem corrupted by the mixed Gaussian and Poisson noises.

Settings / Method	100 photons		1000 photons		10000 photons	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Empirical	29.6282	0.9172	30.4864	0.9082	31.5067	0.9412
Learned	**29.8253**	**0.9194**	**30.9255**	**0.9193**	**32.2568**	**0.9446**

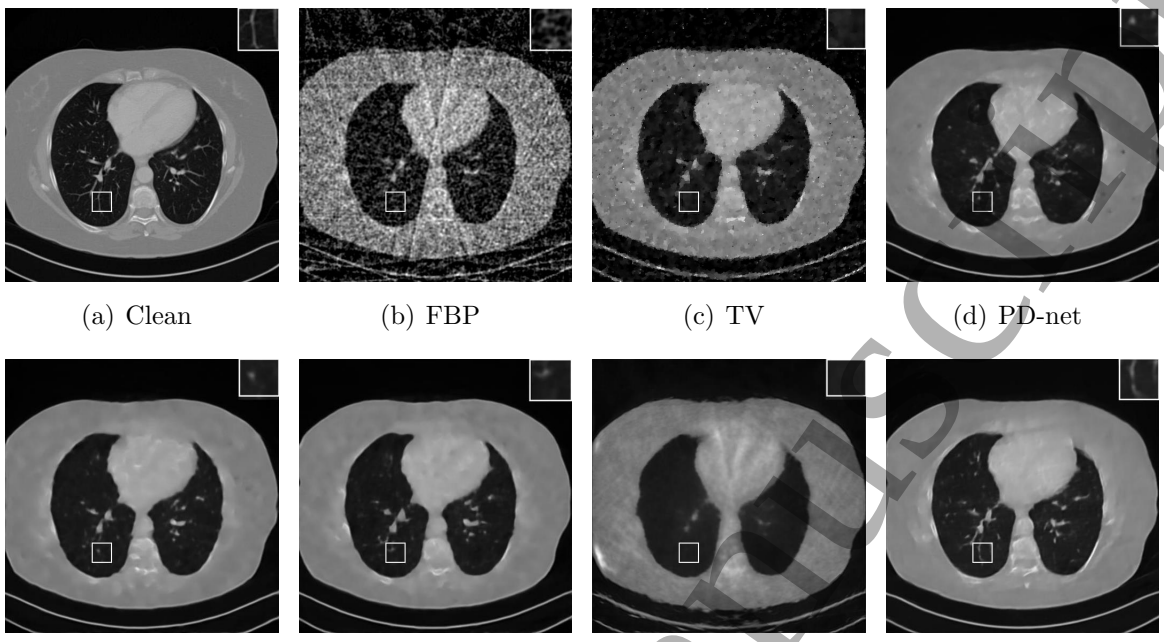
Table 3. Evaluation results on the 150° limited-angle reconstruction problems corrupted by salt-and-pepper noises.

Settings / Method	10% noise		5% noise		1% noise	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Empirical	33.6648	0.9412	33.9457	0.9436	34.5948	0.9493
Learned	**33.9674**	**0.9508**	**34.4504**	**0.9586**	**35.1745**	**0.9589**

5.5. Experiments on the sparse-view reconstruction

We further evaluate the image qualities of our LRPE model under the low-dose sparse-view projection data. We utilize 60 views, 45 views, and 30 views projection data corrupted by 5% white Gaussian noises for evaluation, where the reconstruction results of different reconstruction methods are provided in Table 4. Both PSNR and SSIM demonstrate that our LRPE model outperforms other reconstruction methods. Figure 8 presents the reconstructed images obtained by the comparative methods on the 30-view reconstruction problem. By comparing images (g) and (h), it is evident that the low-resolution image prior can significantly improve the reconstruction quality by preserving fine details and sharp edges. Moreover, the zoomed region exhibits that the deep equilibrium architecture can well maintain structural information.

We further increase the noise level in the raw data to 10% white Gaussian noises and provide the quantitative results in Table 5. It can be observed that the FBP method performs poorly in the presence of high-level noises, with a decrease in PSNR by 4 dB compared to previous experiments. In contrast, the learning-based methods are demonstrated to be less sensitive to noises. Similarly, our LRPE model exhibits the best performance among all the deep learning-based algorithms.



(a) Clean (b) FBP (c) TV (d) PD-net

(e) IFSR (f) SFSR (g) GRAD (h) LRPE

Figure 8. The sparse-view reconstruction experiments are performed on the AAPM phantom dataset using 30 views and 5% Gaussian noise. The display window was set to [0, 1].

Table 4. Comparison of different methods on sparse-view data corrupted by 5% Gaussian noises in terms of PSNR and SSIM.

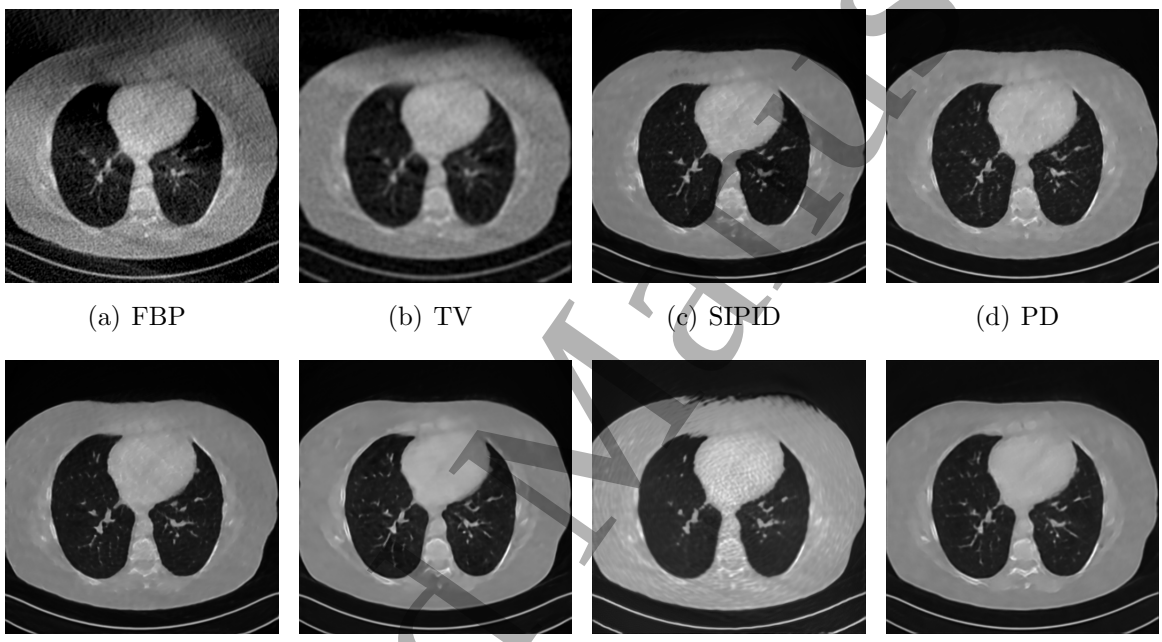
N_{view}	Metrics	FBP	TV	PD	IFSR	SFSR	GRAD	LRPE
60	PSNR	21.4443	27.0301	30.0261	30.4563	30.4681	26.2635	**31.0746**
	SSIM	0.4828	0.8548	0.9237	0.9321	0.9331	0.8647	**0.9337**
45	PSNR	19.7123	25.7201	29.1028	29.6814	29.7915	24.3606	**30.8035**
	SSIM	0.4061	0.8183	0.9135	0.9253	0.9263	0.8031	**0.9342**
30	PSNR	17.8999	23.705	28.9207	29.0632	29.5072	23.713	**29.7237**
	SSIM	0.3257	0.7615	0.9167	0.9198	**0.9244**	0.8355	0.921

Table 5. Comparison of different methods on sparse-view data corrupted by 10% Gaussian noises in terms of PSNR and SSIM.

N_{view}	Metrics	FBP	TV	PD	IFSR	SFSR	GRAD	LRPE
60	PSNR	17.6258	24.6045	28.311	28.3782	28.6113	24.4988	**28.7754**
	SSIM	0.277	0.7826	0.9079	0.9147	0.9161	0.8514	**0.9174**
45	PSNR	16.1392	23.7815	27.5534	28.0502	28.3181	23.6734	**28.6671**
	SSIM	0.2236	0.752	0.9049	0.9116	**0.9141**	0.8318	0.9137
30	PSNR	14.3605	22.9388	26.9146	27.0682	27.653	23.0145	**27.8697**
	SSIM	0.166	0.7416	0.9011	0.9026	**0.9071**	0.8156	0.9043

Table 6. Comparison of different methods on limited-angle data corrupted by 5% Gaussian noises in terms of PSNR and SSIM.

N_{view}	Metrics	FBP	TV	SIPID	PD	SFSR	LRIP	GRAD	LRPE
150°	PSNR	13.5911	25.8815	30.3275	30.3766	30.9411	31.5957	24.3043	**31.636**
	SSIM	0.4854	0.8091	0.9276	0.9301	0.9324	**0.9426**	0.8572	0.9422
120°	PSNR	13.4418	23.5852	27.0428	27.1539	28.3263	29.2763	21.6601	**29.4835**
	SSIM	0.4008	0.7891	0.9024	0.9037	0.9103	0.9361	0.801	**0.9378**
90°	PSNR	13.0314	19.9501	22.7492	22.6047	24.2494	25.1555	19.3469	**25.9517**
	SSIM	0.3881	0.6918	0.8626	0.8612	0.8761	**0.8893**	0.7538	0.8814



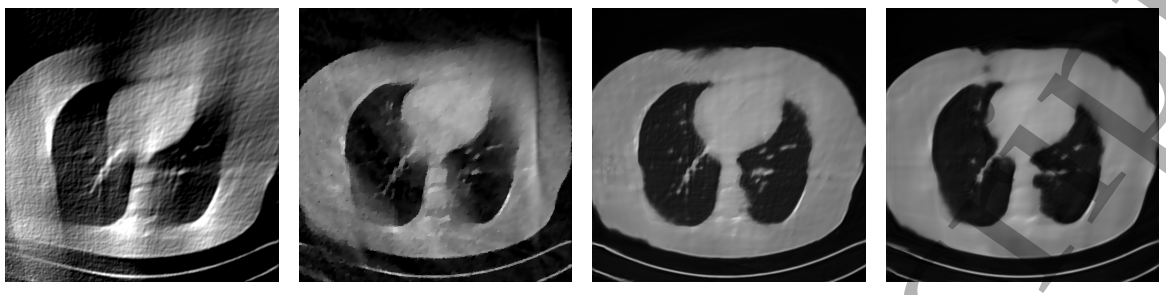
(a) FBP (b) TV (c) SIPID (d) PD

(e) SFSR (f) LRIP (g) GRAD (h) LRPE

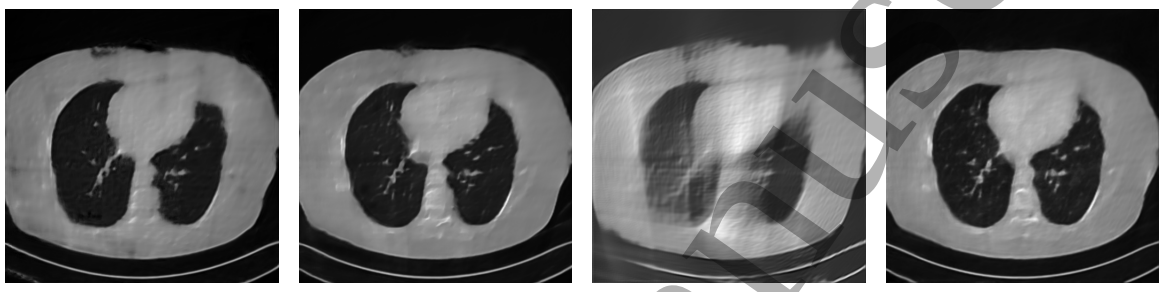
Figure 9. Limited-angle reconstruction experiment with 150° scanning angular range and 5% Gaussian noises.

5.6. Experiments on the limited-angle reconstruction

In this subsection, we evaluate the performance of our LRPE model on the limited-angle reconstruction, where 5% Gaussian noises are corrupted in the projection data. Both PSNR and SSIM of the comparison methods are provided in Table 6. We can observe that the reconstruction qualities of all methods decrease as the scanning angle shrinks. Not surprisingly, our LRPE model has the numerical advantage compared to other comparison algorithms, which provides a 0.8 dB higher PSNR than the LRIP on 90° limited-angle reconstruction task. Particularly, the projection data of the low-resolution prior used by LRIP model is different from ours, which was computed using the down-sampling matrix. Obviously, our setting is more reasonable and in accord with the CT scanner, which also gives better reconstruction results. Figure 9 and Figure 10 present the reconstruction results with a scanning range of 150° limited-angle and 90° limited-angle, respectively. As can be observed, the learning-based methods outperform



(a) FBP (b) TV (c) SIPID (d) PD



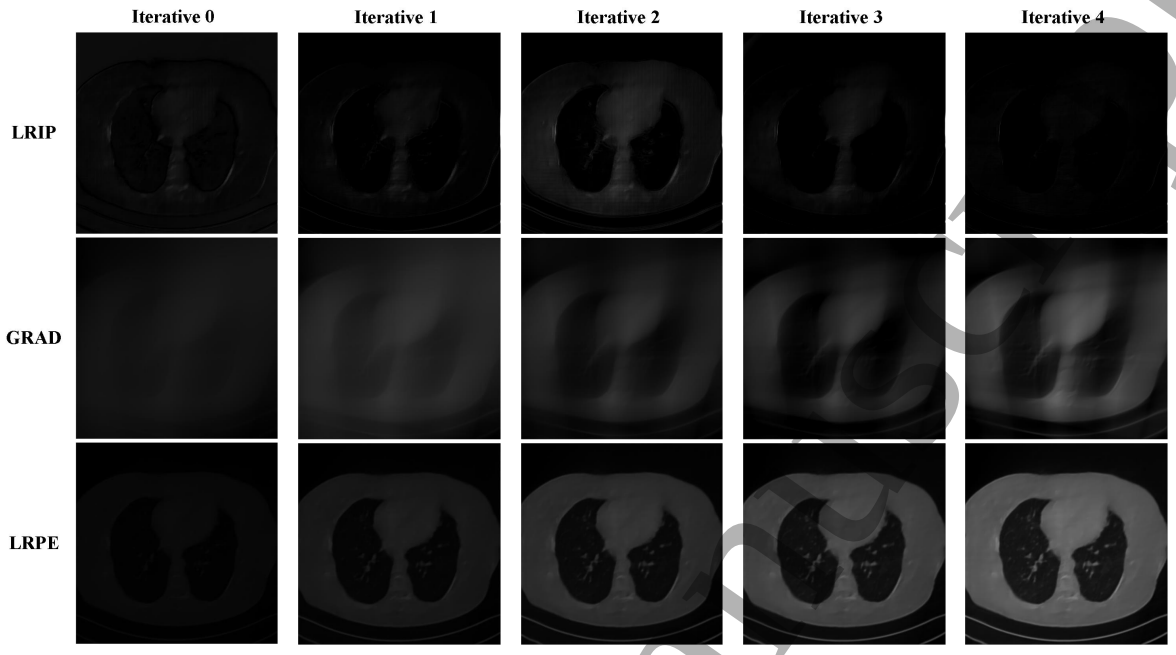
(e) SFSR (f) LRIP (g) GRAD (h) LRPE

Figure 10. Limited-angle reconstruction experiment of the AAPM phantom dataset with 90° scanning angular range and 5% Gaussian noises.

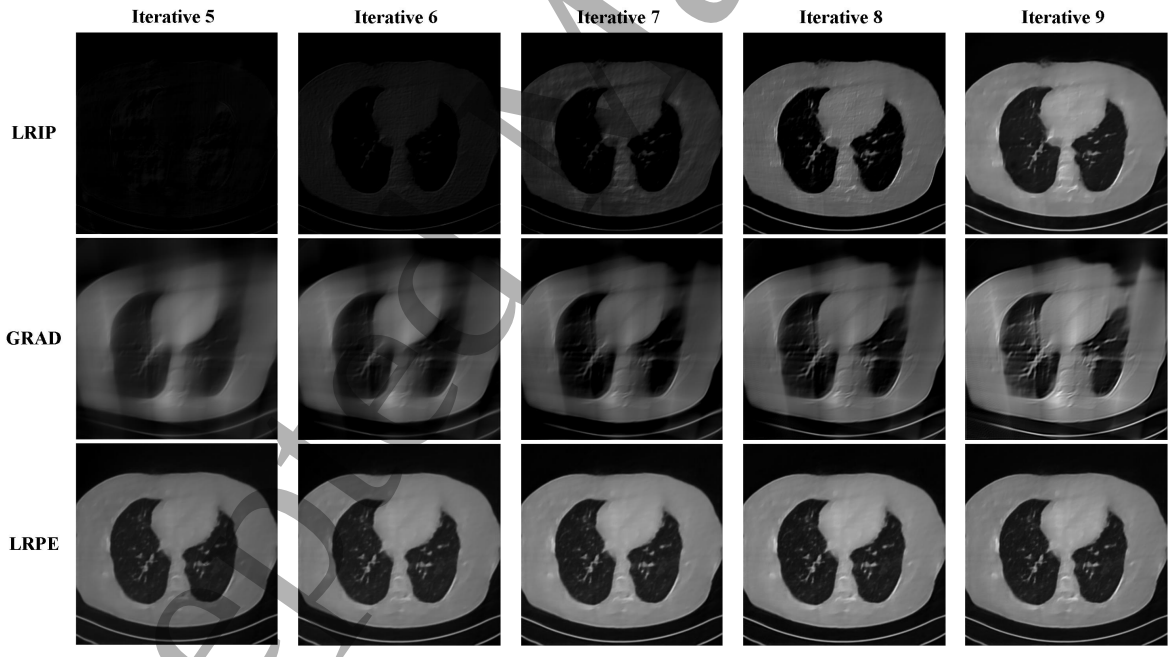
both FBP and TV, which contain noticeable artifacts in regions within missing angles. Moreover, our LRPE model surpasses the SIPID, PD, FSR, and LRIP by providing more image details and continuous contours. Thus, both quantitative and qualitative results confirm that the low-resolution image is a suitable prior for the ill-posed limited-angle reconstruction. Furthermore, by comparing the reconstruction results of 150° limited-angle and 90°, it is not difficult to find that the resolution of the reconstruction results of all methods significantly decreases with the loss of angles. Although our method can provide better structural information by introducing the low-resolution image prior, it still suffers details missing due to angle deficient. It is mainly because our low-resolution image is the direct reconstruction result of 90° limited-angle data. Thus, our future work includes to explore effective reconstruction methods for low-resolution images.

In Figure 11, we present and compare the output results of the LRIP, GRAD, and LRPE models during the iterative process. Although we have shown the numerical convergence of our LRPE model in the inference phase in Figure 7 (a), the visual comparison further demonstrates that our model can provide more meaningful intermediate results in the iterative process. Compared to GRAD model, our method can effectively improve the quality of reconstruction results by introducing a low-resolution image prior. On the other hand, the results of the LRIP model are not stable in the first six stages, and only provide meaningful results in the last two stages. Thus, the deep equilibrium model with convergence guarantee is a better choice for establishing a stable end-to-end reconstruction network.

Last but not least, we conduct the comparison study on the running time of the

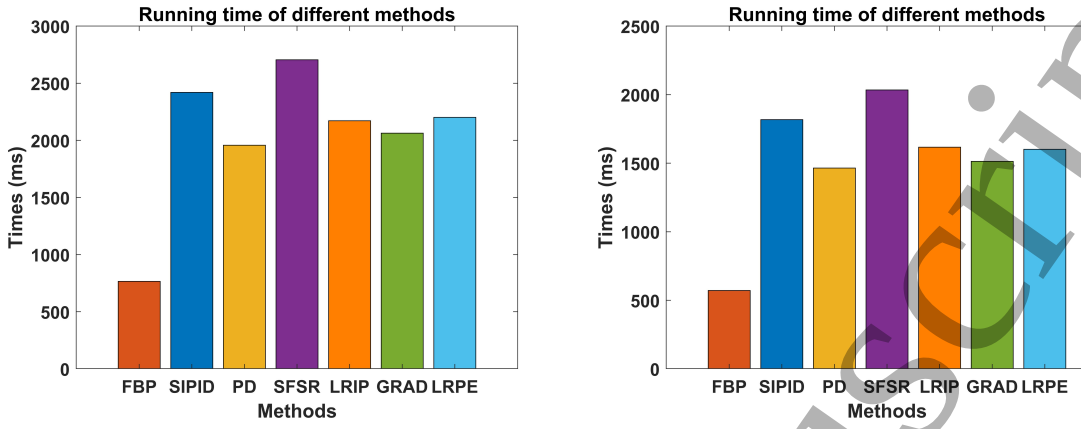


(a) PSNR:9.1227 (b) PSNR:10.2367 (c) PSNR:11.9091 (d) PSNR:14.2079 (e) PSNR:17.2330



(f) PSNR:20.9317 (g) PSNR:23.0121 (h) PSNR:24.4391 (i) PSNR:25.8657 (j) PSNR:25.9517

Figure 11. The comparison results of the LRIP, GRAD, and LRPE model on the human phantom dataset with a scanning angular range of 90° and 5% Gaussian noise.



(a) Sparse-view reconstruction (b) Limited-angle reconstruction

Figure 12. The execution times of various methods were assessed using a sparse angle of 30 degrees and a limited scanning angular range of 90°.

evaluated methods, which are illustrate in Figure 12. Notably, excluding the running time of the TV method, it exhibits an execution time approximately ten times longer than that of the FBP. By analyzing the bar chart results, we can observe that although our method is not as rapid as the PD method, we incur only a 100-millisecond increase in running time to achieve superior reconstruction results.

6. Concluding Remarks

In this paper, we proposed a novel end-to-end equilibrium imaging geometric model for the ill-posed image reconstruction problems. The low-resolution image was used prior to guarantee that the proposed model can achieve good reconstruction results based on incomplete projection data. The advantages of our proposal were validated using a substantial amount of numerical experiments on sparse-view and limited-angle settings. The experimental results demonstrated that our model can enable accurate and efficient CT image reconstruction. Although our downsampling imaging geometric modeling was studied for the fan-beam imaging systems, it can be also used for other modern CT settings. Similar works include [29] and [17], where the low-resolution image prior was incorporated into the reconstruction network. The prior information in [29] was obtained by halving the receiver signal, requiring a second scan of the patient, which is clinically prohibitive. In contrast, our approach eliminates the need of re-scanning patients. On the other hand, Gao *et al.* [17] also requires two distinct sets of projection data. Obviously, the low-resolution image obtained from the same scan can better maintain consistency with the high-resolution image, thereby preserving more structural and fine information. According to the experimental results, although our LRPE model is superior to other algorithms in reconstruction results, there is still a problem of detail loss in the case of incomplete data sampling. Thus, enhancing the quality of low-resolution images through effective sinogram restoration methods is

a possible way to improve the reconstruction results of our LRPE model on severe incomplete data problems. One of the future research directions is to enhance the resolution of reconstructed images under extreme limited-angle and low-dose scanning problems. In theoretical analysis, the noises in data present new challenges in estimating the gradients of network models. We plan to investigate the convergence of the deep equilibrium model based on stochastic gradient similar to [6].

7. Acknowledgments

We would like to thank the anonymous referees for the valuable comments and helpful suggestions to improve this paper. The work was partially supported by the National Natural Science Foundation of China (NSFC 12071345).

References

[1] Jonas Adler and Ozan Öktem. Learned primal-dual reconstruction. *IEEE Transactions on Medical Imaging*, 37(6):1322–1332, 2018.

[2] Clemens Arndt, Felix Güttler, Andreas Heinrich, Florian Bürckenmeyer, Ioannis Diamantis, and Ulf Teichgräber. Deep learning ct image reconstruction in clinical practice. In *RöFo-Fortschritte auf dem Gebiet der Röntgenstrahlen und der bildgebenden Verfahren*, volume 193, pages 252–261. Georg Thieme Verlag KG, 2021.

[3] Daniel Otero Baguer, Johannes Leuschner, and Maximilian Schmidt. Computed tomography reconstruction using deep image prior and learned reconstruction methods. *Inverse Problems*, 36(9):094004, 2020.

[4] Shaojie Bai, J Zico Kolter, and Vladlen Koltun. Deep equilibrium models. *Advances in Neural Information Processing Systems*, 32, 2019.

[5] Jonathan F Bard. *Practical bilevel optimization: algorithms and applications*, volume 30. Springer Science & Business Media, 2013.

[6] Léon Bottou, Frank E Curtis, and Jorge Nocedal. Optimization methods for large-scale machine learning. *SIAM review*, 60(2):223–311, 2018.

[7] Jerome Bracken and James T McGill. Mathematical programs with optimization problems in the constraints. *Operations Research*, 21(1):37–44, 1973.

[8] Ailong Cai, Linyuan Wang, Hanming Zhang, Bin Yan, Lei Li, Xiaoqi Xi, and Jianxin Li. Edge guided image reconstruction in linear scan ct by weighted alternating direction tv minimization. *Journal of X-ray Science and Technology*, 22(3):335–349, 2014.

[9] Qian Cao, Wojciech Zbijewski, Alejandro Sisniega, John Yorkston, Jeffrey H Siewerdsen, and J Webster Stayman. Multiresolution iterative reconstruction in high-resolution extremity cone-beam ct. *Physics in Medicine & Biology*, 61(20):7263, 2016.

[10] Stanley H Chan, Xiran Wang, and Omar A Elgendy. Plug-and-play admm for image restoration: Fixed-point convergence and applications. *IEEE Transactions on Computational Imaging*, 3(1):84–98, 2016.

[11] Guang-Hong Chen, Jie Tang, and Shuai Leng. Prior image constrained compressed sensing (piccs): a method to accurately reconstruct dynamic ct images from highly undersampled projection data sets. *Medical physics*, 35(2):660–663, 2008.

[12] Weilin Cheng, Yu Wang, Ying Chi, Xuansong Xie, and Yuping Duan. Learned full-sampling reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 375–384. Springer, 2019.

[13] Benoît Colson, Patrice Marcotte, and Gilles Savard. An overview of bilevel optimization. *Annals of Operations Research*, 153:235–256, 2007.

[14] Andrei Dabravolski, Kees Joost Batenburg, and Jan Sijbers. A multiresolution approach to discrete tomography using dart. *PloS One*, 9(9):e106090, 2014.

[15] Stephan Dempe and Alain Zemkoho. Bilevel optimization. *Springer optimization and its applications*, 161, 2020.

[16] Heinz Werner Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375. Springer Science & Business Media, 1996.

[17] Qifeng Gao, Rui Ding, Linyuan Wang, Bin Xue, and Yuping Duan. Lrip-net: Low-resolution image prior based network for limited-angle ct reconstruction. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2022.

[18] Saeed Ghadimi and Mengdi Wang. Approximation methods for bilevel programming. *arXiv preprint arXiv:1802.02246*, 2018.

[19] Maryam Gholizadeh-Ansari, Javad Alirezaie, and Paul Babyn. Deep learning for low-dose ct denoising using perceptual loss and edge detection layer. *Journal of Digital Imaging*, 33(2):504–515, 2020.

[20] Suhita Ghosh, Philipp Ernst, Georg Rose, Andreas Nürnberger, and Sebastian Stober. Towards patient specific reconstruction using perception-aware cnn and planning ct as prior. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, pages 1–5. IEEE, 2022.

[21] Guy Gilboa and Stanley Osher. Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, 7(3):1005–1028, 2009.

[22] Davis Gilton, Gregory Ongie, and Rebecca Willett. Deep equilibrium architectures for inverse problems in imaging. *IEEE Transactions on Computational Imaging*, 7:1123–1133, 2021.

[23] Andrzej Granas and James Dugundji. *Fixed point theory*, volume 14. Springer, 2003.

[24] Jace Grandinetti, Yin Gao, Yesenia Gonzalez, Jie Deng, Chenyang Shen, and Xun Jia. Mr image reconstruction from undersampled data for image-guided radiation therapy using a patient-specific deep manifold image prior. *Frontiers in Oncology*, 12:1013783, 2022.

[25] Joël Greffier, Aymeric Hamard, Fabricio Pereira, Corinne Barrau, Hugo Pasquier, Jean Paul Beregi, and Julien Frandon. Image quality and dose reduction opportunity of deep learning image reconstruction algorithm for ct: a phantom study. *European Radiology*, 30(7):3951–3959, 2020.

[26] Karol Gregor and Yann LeCun. Learning fast approximations of sparse coding. In *Proceedings of the 27th international conference on international conference on machine learning*, pages 399–406, 2010.

[27] Per Christian Hansen. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. SIAM, 1998.

[28] Akinori Hata, Masahiro Yanagawa, Yuriko Yoshida, Tomo Miyata, Mitsuko Tsubamoto, Osamu Honda, and Noriyuki Tomiyama. Combination of deep learning–based denoising and iterative reconstruction for ultra-low-dose ct of the chest: image quality and lung-rads evaluation. *American Journal of Roentgenology*, 215:1321–1328, 2020.

[29] Ji He, Shilin Chen, Hua Zhang, Xi Tao, Wuhong Lin, Shanli Zhang, Dong Zeng, and Jianhua Ma. Downsampled imaging geometric modeling for accurate ct reconstruction via deep learning. *IEEE Transactions on Medical Imaging*, 40(11):2976–2985, 2021.

[30] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[31] Yuanwei He, Li Zeng, Wei Yu, and Changcheng Gong. Noise suppression–guided image filtering for low-snr ct reconstruction. *Medical & Biological Engineering & Computing*, 58(11):2621–2629, 2020.

[32] Jing Huang, Jianhua Ma, Nan Liu, Hua Zhang, Zhaoying Bian, Yanqiu Feng, Qianjin Feng, and Wufan Chen. Sparse angular ct reconstruction using non-local means based iterative-correction pocs. *Computers in Biology and Medicine*, 41(4):195–205, 2011.

[33] Samuel Hurault, Arthur Leclaire, and Nicolas Papadakis. Gradient step denoiser for convergent plug-and-play. *arXiv preprint arXiv:2110.03220*, 2021.

[34] Eunhee Kang, Junhong Min, and Jong Chul Ye. A deep convolutional neural network using directional wavelets for low-dose x-ray ct reconstruction. *Medical Physics*, 44(10):e360–e375, 2017.

[35] Hojin Kim, Josephine Chen, Adam Wang, Cynthia Chuang, Mareike Held, and Jean Pouliot. Non-local total-variation (nltv) minimization combined with reweighted l1-norm for compressed sensing ct reconstruction. *Physics in Medicine & Biology*, 61(18):6878, 2016.

[36] Frederick J Larke, Randell L Kruger, Christopher H Cagnon, Michael J Flynn, Michael M McNitt-Gray, Xizeng Wu, Phillip F Judy, and Dianna D

Cody. Estimated radiation dose associated with low-dose chest ct of average-size participants in the national lung screening trial. *American Journal of Roentgenology*, 197(5):1165–1169, 2011.

[37] Marc Lenfant, Olivier Chevallier, Pierre-Olivier Comby, Grégory Secco, Karim Haioun, Frédéric Ricolfi, Brivaël Lemogne, and Romaric Loffroy. Deep learning versus iterative reconstruction for ct pulmonary angiography in the emergency setting: improved image quality and reduced radiation dose. *Diagnostics*, 10(8):558, 2020.

[38] Kuai Li, Ziru Sang, Xuezhu Zhang, Mengxi Zhang, Changhui Jiang, Qiyang Zhang, Yongshuai Ge, Dong Liang, Yongfeng Yang, Xin Liu, et al. Few-view ct image reconstruction using improved total variation regularization. *Journal of X-ray Science and Technology*, 27(4):739–753, 2019.

[39] Wei-An Lin, Haofu Liao, Cheng Peng, Xiaohang Sun, Jingdan Zhang, Jiebo Luo, Rama Chellappa, and Shaohua Kevin Zhou. Dudonet: Dual domain network for ct metal artifact reduction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10512–10521, 2019.

[40] Sebastian Lunz, Ozan Öktem, and Carola-Bibiane Schönlieb. Adversarial regularizers in inverse problems. *Advances in Neural Information Processing Systems*, 31, 2018.

[41] Michael Lustig, David Donoho, and John M Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 58(6):1182–1195, 2007.

[42] Varinder Malik and Esam MA Hussein. A fuzzy inference method for image fusion/refinement of ct images from incomplete data. *Heliyon*, 7(4):e06839, 2021.

[43] C McCollough. Tu-fg-207a-04: overview of the low dose ct grand challenge. *Medical Physics*, 43(6Part35):3759–3760, 2016.

[44] Mingqiang Meng, Sui Li, Lisha Yao, Danyang Li, Manman Zhu, Qi Gao, Qi Xie, Qian Zhao, Zhaoying Bian, Jing Huang, et al. Semi-supervised learned sinogram restoration network for low-dose ct image reconstruction. In *Medical Imaging 2020: Physics of Medical Imaging*, volume 11312, pages 67–73. SPIE, 2020.

[45] Jiang Min, Hongwei Tao, Xinglong Liu, and Kai Cheng. A non-local total generalized variation regularization reconstruction method for sparse-view x-ray ct. *Measurement Science and Technology*, 35(4):045404, 2024.

[46] Payam Mohammadinejad, Achille Mileto, Lifeng Yu, Shuai Leng, Luis S Guimaraes, Andrew D Missert, Corey T Jensen, Hao Gong, Cynthia H McCollough, and Joel G Fletcher. Ct noise-reduction methods for lower-dose scanning: strengths and weaknesses of iterative reconstruction algorithms and new techniques. *RadioGraphics*, 41(5):1493–1508, 2021.

[47] Vishal Monga, Yuelong Li, and Yonina C Eldar. Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Processing Magazine*, 38(2):18–44, 2021.

[48] Pravin Nair, Ruturaj G Gavaskar, and Kunal Narayan Chaudhury. Fixed-point and objective convergence of plug-and-play algorithms. *IEEE Transactions on Computational Imaging*, 7:337–348, 2021.

[49] Ju Gang Nam, Jung Hee Hong, Da Som Kim, Jiseon Oh, and Jin Mo Goo. Deep learning reconstruction for contrast-enhanced ct of the upper abdomen: similar image quality with lower radiation dose in direct comparison with iterative reconstruction. *European Radiology*, 31(8):5533–5543, 2021.

[50] Yoshifumi Noda, Tetsuro Kaga, Nobuyuki Kawai, Toshiharu Miyoshi, Hiroshi Kawada, Fuminori Hyodo, Avinash Kambadakone, and Masayuki Matsuo. Low-dose whole-body ct using deep learning image reconstruction: image quality and lesion detection. *The British Journal of Radiology*, 94(1121):20201329, 2021.

[51] Edward F Patz, Paul Pinsky, Constantine Gatsonis, JoRean D Sicks, Barnett S Kramer, Martin C Tammemägi, Caroline Chiles, William C Black, Denise R Aberle, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. *JAMA Internal Medicine*, 174(2):269–274, 2014.

[52] Jean-Christophe Pesquet, Audrey Repetti, Matthieu Terris, and Yves Wiaux. Learning maximally monotone operators for image recovery. *SIAM Journal on Imaging Sciences*, 14(3):1206–1237, 2021.

[53] Saiprasad Ravishankar and Yoram Bresler. Mr image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE transactions on medical imaging*, 30(5):1028–1041, 2010.

[54] Romke Rozema, Herbert T Kruitbosch, Baucke van Minnen, Bart Dorgelo, Joep Kraeima, and Peter MA van Ooijen. Iterative reconstruction and deep learning algorithms for enabling low-dose computed tomography in midfacial trauma. *Oral surgery, Oral Medicine, Oral Pathology and Oral Radiology*, 132(2):247–254, 2021.

[55] Ernest Ryu, Jialin Liu, Sicheng Wang, Xiaohan Chen, Zhangyang Wang, and Wotao Yin. Plug-and-play methods provably converge with properly trained denoisers. In *International Conference on Machine Learning*, pages 5546–5557. PMLR, 2019.

[56] Johannes Schwab, Stephan Antholzer, and Markus Haltmeier. Deep null space learning for inverse problems: convergence analysis and rates. *Inverse Problems*, 35(2):025008, 2019.

[57] Hongming Shan, Atul Padole, Fatemeh Homayounieh, Uwe Kruger, Ruhani Doda Khera, Chayanin Nitiwarangkul, Mannudeep K Kalra, and Ge Wang. Competitive performance of a modularized deep neural network compared to commercial algorithms for low-dose ct image reconstruction. *Nature Machine Intelligence*, 1(6):269–276, 2019.

[58] Liyue Shen, Wei Zhao, Dante Capaldi, John Pauly, and Lei Xing. A geometry-informed deep learning framework for ultra-sparse 3d tomographic image reconstruction. *Computers in Biology and Medicine*, 148:105710, 2022.

[59] Yoon Joo Shin, Won Chang, Jong Chul Ye, Eunhee Kang, Dong Yul Oh, Yoon Jin Lee, Ji Hoon Park, and Young Hoon Kim. Low-dose abdominal ct using a deep learning-based denoising algorithm: a comparison with ct reconstructed with filtered back projection or iterative reconstruction algorithm. *Korean Journal of Radiology*, 21(3):356–364, 2020.

[60] Byeongsu Sim, Gyutaek Oh, Jeongsol Kim, Chanyong Jung, and Jong Chul Ye. Optimal transport driven cyclegan for unsupervised learning in inverse problems. *SIAM Journal on Imaging Sciences*, 13(4):2281–2306, 2020.

[61] Suhas Sreehari, S Venkat Venkatakrishnan, Brendt Wohlberg, Gregery T Buzzard, Lawrence F Drummy, Jeffrey P Simmons, and Charles A Bouman. Plug-and-play priors for bright field electron tomography and sparse interpolation. *IEEE Transactions on Computational Imaging*, 2(4):408–423, 2016.

[62] Charles W Stearns, Ravindra M Manjeshwar, and Scott D Wollenweber. An efficient algorithm for targeted reconstruction of tomographic data. In *2006 IEEE Nuclear Science Symposium Conference Record*, volume 5, pages 2808–2811. IEEE, 2006.

[63] Jelmer M Wolterink, Tim Leiner, Max A Viergever, and Ivana Išgum. Generative adversarial networks for noise reduction in low-dose ct. *IEEE Transactions on Medical Imaging*, 36(12):2536–2545, 2017.

[64] Dufan Wu, Kyungsang Kim, Georges El Fakhri, and Quanzheng Li. Iterative low-dose ct reconstruction with priors trained by artificial neural network. *IEEE Transactions on Medical Imaging*, 36(12):2479–2486, 2017.

[65] Jinqiu Xu, Yunsong Zhao, Hongwei Li, and Peng Zhang. An image reconstruction model regularized by edge-preserving diffusion and smoothing for limited-angle computed tomography. *Inverse Problems*, 35(8):085004, 2019.

[66] Qingsong Yang, Pingkun Yan, Yanbo Zhang, Hengyong Yu, Yongyi Shi, Xuanqin Mou, Mannudeep K Kalra, Yi Zhang, Ling Sun, and Ge Wang. Low-dose ct image denoising using a generative adversarial network with wasserstein distance and perceptual loss. *IEEE Transactions on Medical Imaging*, 37(6):1348–1357, 2018.

[67] Huizhuo Yuan, Jinzhu Jia, and Zhanxing Zhu. Sipid: A deep learning framework for sinogram interpolation and image denoising in low-dose ct reconstruction. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1521–1524. IEEE, 2018.

[68] Zhicheng Zhang, Xiaokun Liang, Xu Dong, Yaoqin Xie, and Guohua Cao. A sparse-view ct reconstruction method based on combination of densenet and deconvolution. *IEEE Transactions on Medical Imaging*, 37(6):1407–1417, 2018.

[69] Bo Zhao, Hao Gao, Huanjun Ding, and Sabee Molloi. Tight-frame based iterative image reconstruction for spectral breast ct. *Medical Physics*, 40(3):031905, 2013.

[70] Shusen Zhao, Dimeng Xia, and Xing Zhao. A fast image reconstruction method for planar objects ct inspired by differentiation property of fourier transform (dpft). *Inverse Problems*, 37(7):075001, 2021.

[71] Bo Zhu, Jeremiah Z Liu, Stephen F Cauley, Bruce R Rosen, and Matthew S Rosen. Image reconstruction by domain-transform manifold learning. *Nature*, 555(7697):487–492, 2018.

[72] Gustav Zickert, Ozan Öktem, and Can Evren Yarman. Joint gaussian dictionary learning and tomographic reconstruction. *Inverse Problems*, 38(10):105010, 2022.