

Low-Rank Angular Prior Guided Multi-Diffusion Model for Few-Shot Low-Dose CT Reconstruction

Wenhao Zhang, Bin Huang, Shuyue Chen, Xiaoling Xu, Weiwen Wu[✉], Member, IEEE,
and Qiegen Liu[✉], Senior Member, IEEE

Abstract—Low-dose computed tomography (LDCT) is essential in clinical settings to minimize radiation exposure; however, reducing the dose often leads to a significant decline in image quality. Additionally, conventional deep learning approaches typically require large datasets, raising concerns about privacy, costs, and time constraints. To address these challenges, a few-shot low-dose CT reconstruction method is proposed, utilizing low-Rank Angular Prior (RAP) multi-diffusion model. In the prior learning phase, projection data is transformed into multiple consecutive views organized by angular segmentation, allowing for the extraction of rich prior information through low-rank processing. This structured approach enhances the learning capacity of the multi-diffusion model. During the iterative reconstruction phase, a stochastic differential equation solver is employed alongside data consistency constraints to iteratively refine the acquired projection data. Furthermore, penalized weighted least-squares and total variation techniques are integrated to improve image quality. Results demonstrate that the reconstructed images closely resemble those obtained from normal-dose CT, validating the RAP model as an effective and practical solution for artifact and noise reduction while preserving image fidelity in low-dose situation.

Index Terms—Low-dose CT, diffusion model, Hankel matrix, few-shot learning, sinogram domain.

Received 27 May 2024; revised 27 September 2024; accepted 9 November 2024. Date of publication 4 December 2024; date of current version 13 December 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62122033, Grant 62471502, and Grant 62201628, in part by the Guangdong Basic and Applied Basic Research Foundation under Grant 2023A151011780, in part by the Beijing Natural Science Foundation under Grant L232123, and in part by the Nanchang University Interdisciplinary Innovation Fund Project under Grant PYJX20230002. The associate editor coordinating the review of this article and approving it for publication was Prof. Bo Zhao. (W. Zhang and B. Huang are co-first authors.) (Corresponding author: Qiegen Liu.)

Wenhao Zhang and Xiaoling Xu are with the School of Information Engineering, Nanchang University, Nanchang 330031, China (e-mail: zhangwenhao@email.ncu.edu.cn; xuxiaoling@ncu.edu.cn).

Bin Huang and Shuyue Chen are with the School of Mathematics and Computer Sciences, Nanchang University, Nanchang 330031, China (e-mail: huangbin@email.ncu.edu.cn; shuyue.chen@email.ncu.edu.cn).

Weiwen Wu is with the School of Biomedical Engineering, Sun Yat-Sen University, Shenzhen, Guangdong 510275, China (e-mail: wuweiw7@mail.sysu.edu.cn).

Qiegen Liu is with the School of Information Engineering, Nanchang University, Nanchang 330031, China, and also with the Jiangxi Provincial Key Laboratory of Advanced Signal Processing and Intelligent Communications, Nanchang 330031, China (e-mail: liuqiegen@ncu.edu.cn).

Digital Object Identifier 10.1109/TCI.2024.3503366

I. INTRODUCTION

COMPUTED tomography (CT) techniques offer valuable contributions to diagnosis and intervention procedures, including imaging, image-guided needle biopsy, image-guided intervention, and radiotherapy [1]. However, high radiation doses may pose potential cancer risks in clinical tasks [2]. Therefore, obtaining high-quality CT images while reducing radiation dose has become an important area of research.

Currently, the radiation dose received by patients during CT examinations can be reduced by lowering the tube current [3], referring as low-dose CT (LDCT). Thus, an increasing number of researchers are focusing on improving the image quality of LDCT. LDCT algorithms can be roughly divided into three groups: image post-processing methods [4], [5], sinogram domain methods [6], [7], [8], and iterative reconstruction methods [9], [10]. Image post-processing methods enable the direct enhancement of low-quality images without utilizing raw projection data. Chen et al. [4] combined the autoencoder, deconvolution network, and shortcut connections into the residual encoder-decoder convolutional neural network (RED-CNN) to reconstruct CT images with structural preservation and noise suppression. Ding et al. [5] proposed a low-dose CT image reconstruction method based on deep learning regularization. This method unfolds a proximal forward-backward splitting (PFBS) framework with data-driven image regularization via deep neural networks. However, removing severe streak artifacts and accurately recovering image details and features without projection data can be challenging. The sinogram domain-based reconstruction methods are helpful in solving this problem. Yin et al. [6] introduced a domain progressive 3D residual convolution network (DP-ResNet) for the LDCT imaging procedure, while Balda et al. [7] combined the specific properties of the CT measurement process to design a structure adaptive sinogram (SAS) filter. Manduca et al. [8] proposed a bilateral filtering method incorporating the CT noise model, which achieved higher noise resolution compared to commercial reconstruction kernels. Iterative reconstruction focuses on solving the low-dose CT problem iteratively by extracting prior information from target images [9]. Various priors were developed, with total variation (TV)-based reconstructions [10], [11], [12] being the most well-known. Tian et al. [10] developed an iterative CT reconstruction algorithm with edge-preserving TV regularization, which prioritized smoothing in the non-edge regions of the image. Additionally, Zhang et al. [11] proposed the fractional-order

TV model rather than traditional TV model for suppressing the over-smoothing effect. Sagheer et al. [12] considered tensor TV and developed a method based on low-rank approximation to improve global smoothness. However, iterative reconstruction methods seriously affect computational costs, striking a balance between image quality and computational efficiency in practical applications.

Diffusion models have provided a fresh perspective for image processing tasks [13], [14]. Li et al. [15] introduced an unsupervised image domain score generation model for low-dose CT reconstruction. Song et al. [16] incorporated forward and backward diffusion processes into a framework based on stochastic differential equations (SDE) to design a score-based generative model. This model has found extensive application in clinical medical scenarios and has demonstrated remarkable performance [18], [19]. Moreover, diffusion models exhibit higher efficiency when compared to generative adversarial networks (GANs).

Collecting medical data is a complex and expensive task, and the ongoing issues related to privacy and security add to the complexity of sharing clinical data. Consequently, there is a shortage of readily available clinical data, prompting researchers to explore few-shot reconstruction studies such as image denoising [20], image classification [21], and image segmentation [22]. To overcome this challenge, changing the scale of data has become a widely adopted solution. These methods often involve downsizing and sampling larger images to capture a variety of underlying statistical properties, which are subsequently employed in training datasets. Due to its data redundancy and structural reproducibility, Hankel matrix is commonly employed in various fields such as image denoising [23], artifact removal [24], and compressed sensing [25]. Wang et al. [26] designed an encoder-decoder network based on Hankel matrix decomposition, which leverages low-rank Hankel matrices for reconstructing high-dose images from low-dose images using few-shot learning techniques. However, despite research on few-shot natural images is emerging, studies related to medical images are still relatively limited [27], [28].

In this study, we introduce a few-shot learning method called low-Rank Angular Prior (RAP) multi-diffusion model for LDCT reconstruction, leveraging an angular segmentation strategy within the projection domain. This approach combines a low-rank structured Hankel matrix with diffusion modeling to generate high-quality sinograms from low-dose projection data by incorporating prior information from consecutive view segments. Additionally, we employ penalized weighted least-squares (PWLS) and total variation regularization to enhance image quality and accelerate the iterative process. Unlike conventional supervised learning methods, our proposed unsupervised approach eliminates the need for retraining low-dose/normal-dose CT image pairs when projection dosages change. The method requires only a few-shot data samples for training. By structuring projection data into Hankel matrices across angular segments, we reduce the disparity between CT images in the image domain, facilitating better generalization and improving model performance.

The theoretical and practical contributions of this work can be summarized as follows:

- By proposing a few-shot low-dose CT reconstruction method in the projection domain, sufficient prior information is extracted from multiple consecutive views that structured into multiple Hankel blocks. By employing low-rank processing on the multiple consecutive views, each segment of angular information can contribute redundant prior information, thereby enhancing the learning capacity of the few-shot strategy.
- By dividing full-view projections into multiple consecutive views, the complex data distribution can be decomposed into multiple simpler distributions, which can then be learned through multi-diffusion model. This distributed learning approach can greatly enhance the detail reconstruction capability of CT images.

The remaining sections of this study are organized as follows: Section II provides background information on score-based diffusion model and the construction process of the Hankel matrix. In Section III. The experimental results and analysis are shown in Section IV. The discussion is given in Section V. Finally, we draw conclusions in Section VI.

II. PRELIMINARY

A. CT Imaging

Low-dose CT reconstruction attempts to reconstruct the noise parts in low-dose CT data, which is a classic inverse problem. Currently, there are many algorithms proposed for low-dose CT reconstruction, which can be roughly divided into three groups: image post-processing methods (which can directly skip the processing of raw projection data and process the low-quality image), sinogram domain methods, and iterative reconstruction methods. Specifically, assuming that $x \in \mathbb{R}^M$ is the degraded sinogram, the forward formula of the sinogram reconstruction problem can be given by the following equation:

$$y = x + n \quad (1)$$

where $n \in \mathbb{R}^M$ represents additive noise and y denotes a low-dose sinogram. It should be noted that the inverse problem refers to generating x from y .

To avoid ill-posed issues, the CT sinogram reconstruction problem is represented by an optimization equation with the following constraint:

$$\min_x \{ \|x - y\|_2^2 + \mu R(x) \} \quad (2)$$

where $\|x - y\|_2^2$ is the data fidelity term. $R(x)$ represents the adjustment prior degree term, which is selected as the TV semi-norm. $\|\bullet\|_2^2$ represents the ℓ_2 -norm. Besides, μ is a factor to keep a good balance between the data-consistency (DC) term and the regularization term.

Considering the TV aspect, this optimization process distinguishes among the infinite solutions of (2) and selects the best image with desired properties as the reconstructed sinogram.

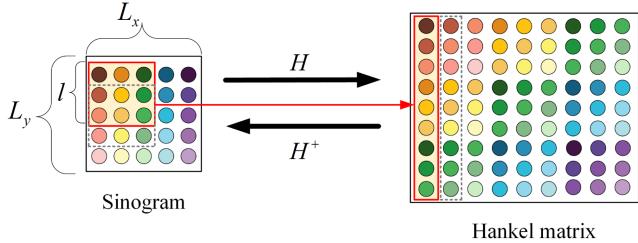


Fig. 1. Demonstration of constructing a new Hankel matrix formulation from a sinogram through H and vice versa H^+ .

Typically, the TV term is defined as:

$$R(x) = \|x\|_{TV}^2 = \int_{\Omega} |\nabla x| dx \quad (3)$$

where Ω is the bounded domain. ∇x represents the gradient of the sinogram x . The TV term is robust in removing noise and artifacts.

B. Construction of Hankel Matrix

Due to the ability of Hankel matrices to fully exploit the redundancy of identical pixels between different positions, Peng et al. [29] constructed a Hankel matrix from the sample data to alleviate the problem of few-shot data samples. Since projection data can be represented by low-rank Hankel matrix, the internal relationship between the sinogram and the Hankel matrix is illustrated in Fig. 1.

As shown in Fig. 1, the sinogram $L_x \times L_y$ undergoes transformation $H(\bullet)$, resulting in the construction of the Hankel matrix of size $l^2 \times (L_x - l + 1)(L_y - l + 1)$. Fig. 1 demonstrates the case of $L_x = 5$, $L_y = 5$, and $l = 3$. Additionally, the step size of the sliding window is set to 1, and the window size is set to $l \times l$.

In the new Hankel matrix formulation, individual block data in the projection domain is vectorized as columns. The linear operator $H(\bullet)$ is defined as the operation to construct the Hankel matrix from the sinogram:

$$H = R^{L_x \times L_y} \rightarrow R^{l^2 \times (L_x - l + 1)(L_y - l + 1)} \quad (4)$$

When performing the reverse operation $H^+(\bullet)$ from (4) to generate the sinogram, multiple anti-diagonal entries are averaged and stored in the projection domain, represented as follows:

$$H^+ = R^{l^2 \times (L_x - l + 1)(L_y - l + 1)} \rightarrow R^{L_x \times L_y} \quad (5)$$

where the superscript $+$ represents the pseudo-inverse operator. It is equivalent to averaging the anti-diagonal elements and placing them in the appropriate locations.

C. Score-Based SDE

The impressive success of diffusion model, specifically score-based stochastic differential equations (SDE), in producing realistic and varied image samples has sparked widespread interest [16]. Score-based SDE encompasses both the forward process and the reverse-time process.

Consider a continuous diffusion process $\{x(t)\}_{t=0}^T$ with $x(t) \in \mathbb{R}^N$, where $t \in [0, T]$ represents the progression time variable. N denotes the dimension of the sinogram. The forward diffusion process can be expressed as the solution to the following SDE:

$$dx = f(x, t)dt + g(t)dw \quad (6)$$

where $f(x, t) \in \mathbb{R}^N$ and $g(t) \in \mathbb{R}$ correspond to the drift coefficient and diffusion coefficient, respectively. $w \in \mathbb{R}^N$ induces the Brownian motion.

Via reversing the above process, samples can be attained. Notably, the reverse-time SDE is also a diffusion process, which could be expressed as follows:

$$dx = [f(x, t) - g(t)^2 \nabla_x \log p_t(x)]dt + g(t)d\bar{w} \quad (7)$$

where dt is the infinitesimal negative time step, \bar{w} is a standard Wiener process when time flows backward from T to 0, and $\nabla_x \log p_t(x)$ is the score of each marginal distribution.

III. PROPOSED METHOD

A. Low-Dose CT Imaging Model

To intuitively represent the simulation process of low-dose CT imaging, n can be defined as the low-dose noise on the sinogram, and I represents the initial CT image. In the presence of n , measuring the sinogram x of size 768×768 , the CT imaging reconstruction problem can be represented by the following equation:

$$y = F^{-1}(I) + n = x + n \quad (8)$$

where $F^{-1}(\bullet)$ corresponds to the filtered back-projection (FBP), y is the low-dose CT sinogram with size 768×768 .

Based on extensive experimental support with a significant amount of real projection data, low-dose sinogram can be approximated as ideal projection data corrupted by additional noise [29]. It is assumed that only a monochromatic light source is used, and the additional noise follows a Poisson distribution. Specifically, the Poisson model for intensity measurements is expressed as follows:

$$L_i \sim \text{Poisson}\{a_i e^{-[x]_i} + r_i\}, \quad i = 1, \dots, M \quad (9)$$

where L_i represents the transmitted number of photons, a_i is the X-ray source intensity of the i -th ray, and r_i signifies the background contribution from electronic noise. x corresponds to the vector representing the attenuation coefficients with inverse length units, M stands for the number of measurements. In (9), the noise level is characterized by a_i , its specific value is provided in Section IV.

The measurement data is transformed into a weighted Gaussian form through a logarithmic operation:

$$y_i \sim N([x]_i, \bar{L}_i / (\bar{L}_i - r_i)^2) \quad (10)$$

where $\bar{L}_i = E[L_i]$. In fact, the low-dose CT problem is formulated as a typical ill-posed inverse problem. To address this issue, the posterior distribution $p(x|y)$ is introduced by the theory of

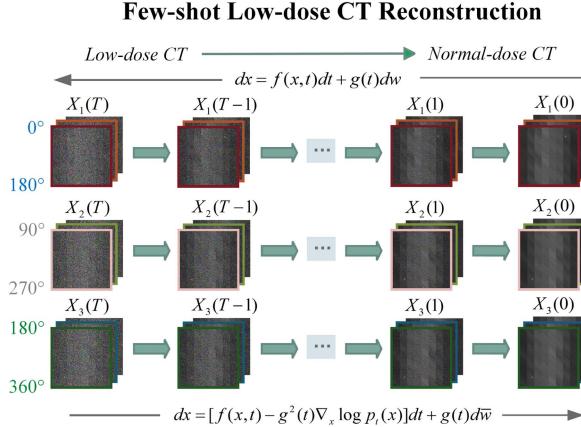


Fig. 2. The perturbed data by noise is smoothed according to the trajectory of an SDE. By estimating the score function $\nabla_x \log p_t(x)$ with SDE, it is possible to approximate the reverse SDE and then solve it to generate sinogram samples from noise.

Bayesian inversion [30]. Consequently, the inverse problem is transformed into a measurement y .

B. Training Stage

During the training stage, the Hankel matrix is preprocessed and input into the network, as shown in Fig. 2. In order to store information within the data, a window of size 8×8 is slid over the initial data, resulting in the construction of a Hankel matrix of size 579121×64 . After applying the Hankel transform, the same information appears at different positions within the matrix, harnessing its redundancy to capture internal statistical information. It is important to note that during the training process, one or few normal-dose sinograms are required. The Hankel matrix is constructed from the projection data as follows:

$$H_s = H(x) \quad (11)$$

where the initial dose projection data is represented as x , $H(\bullet)$ corresponds to the Hankel transform, and H_s represents the constructed Hankel matrix.

During the network training stage, the full-view projection data is divided into three parts: $0^\circ \sim 180^\circ$, $90^\circ \sim 270^\circ$, and $180^\circ \sim 360^\circ$, using a continuous view segmentation strategy. Then, low-rank processing is applied to these three segments of consecutive views, allowing the model to leverage sufficient prior information to support few-shot learning. The first Hankel matrix, with a size of 289560×64 , contains information from the angle range of $0^\circ \sim 180^\circ$. The second Hankel matrix, with a size of 289560×64 , contains information from the angle range of $90^\circ \sim 270^\circ$. The overlapping information can significantly increase data redundancy and enhance data coherence, preventing reconstruction data from being fragmented or cut off due to the segmentation of views. Lastly, the third Hankel matrix is 289561×64 for containing information from the angle range of $180^\circ \sim 360^\circ$. These three Hankel matrices are processed in parallel and cropped into high-dimensional data of

size $64 \times 64 \times 4524$. In addition, multiple small patches are randomly extracted from the matrix as illustrated in Fig. 3. Specifically, the constructed Hankel matrix is partitioned into small blocks using the following process:

$$X = T(H_s) \quad (12)$$

This random partition operation, denoted by the symbol $T(\bullet)$ is applied to the high-dimensional tensor X to extract numerous small blocks. Then these blocks are considered as inputs to the network, resulting in the construction of a large number of tensors. This data augmentation technique enhances the training set, enabling the network to acquire a sufficient amount of prior knowledge.

The score-based model is described in Section II-C learns a prior distribution by leveraging an SDE. This is achieved through a forward SDE, which gradually injects noise, smoothly transforming complex data distributions into the known prior distribution. Fig. 2 illustrates these two processes.

During the training phase, the parameter θ^* of the scoring network is optimized to achieve the peak performance of the network. The objective function can be described as follows:

$$\theta^* = \arg \min_{\theta} \mathbb{E}_t \left\{ \lambda(t) \mathbb{E}_{x(0)} \mathbb{E}_{x(t)|x(0)} \left[\|s_{\theta}(x(t), t) - \nabla_x \log p_t(x(t)|x(0))\|_2^2 \right] \right\} \quad (13)$$

where $\lambda : [0, T] \rightarrow \mathbb{R}^+$ is a positive weighting function and t is uniformly sampled over $[0, T]$. $p_t(x(t)|x(0))$ is the Gaussian perturbation kernel centered at $x(0)$. Once the network satisfies $s_{\theta}(x(t), t) \simeq \nabla_x \log p_t(x)$, $\nabla_x \log p_t(x)$ will be known for all t by solving $s_{\theta}(x(t), t)$.

C. Iterative Reconstruction Stage

This part describes the iterative reconstruction process of the RAP model. First, the forward projection of CT imaging:

$$x = F(I) \quad (14)$$

where I is the original image, $F(\bullet)$ corresponds to the forward projection. The score-based diffusion model is employed to estimate the prior distribution of sinogram data $p_t(x)$, enabling noise suppression and enhanced information. Unlike perturbing data with a finite number of noise distributions, this method takes into account continuous distributions over time during the forward diffusion process. By reversing the SDE, random noise can be transformed into data suitable for sampling. In this paper, as suggested in [31], Predictor-Corrector (PC) sampling is introduced in the sample update step. In PC sampling, the predictor is treated as a numerical solver for the reverse-time SDE. Once the reverse-time SDE process concludes, samples are generated according to the discretized prior distribution, which can be discretized as follows:

$$x^i \leftarrow x^{i+1} + (\sigma_{i+1}^2 - \sigma_i^2)s_{\theta}(x^{i+1}, \sigma_{i+1}) + \sqrt{\sigma_{i+1}^2 - \sigma_i^2} z \quad i = N-1, \dots, 0 \quad (15)$$

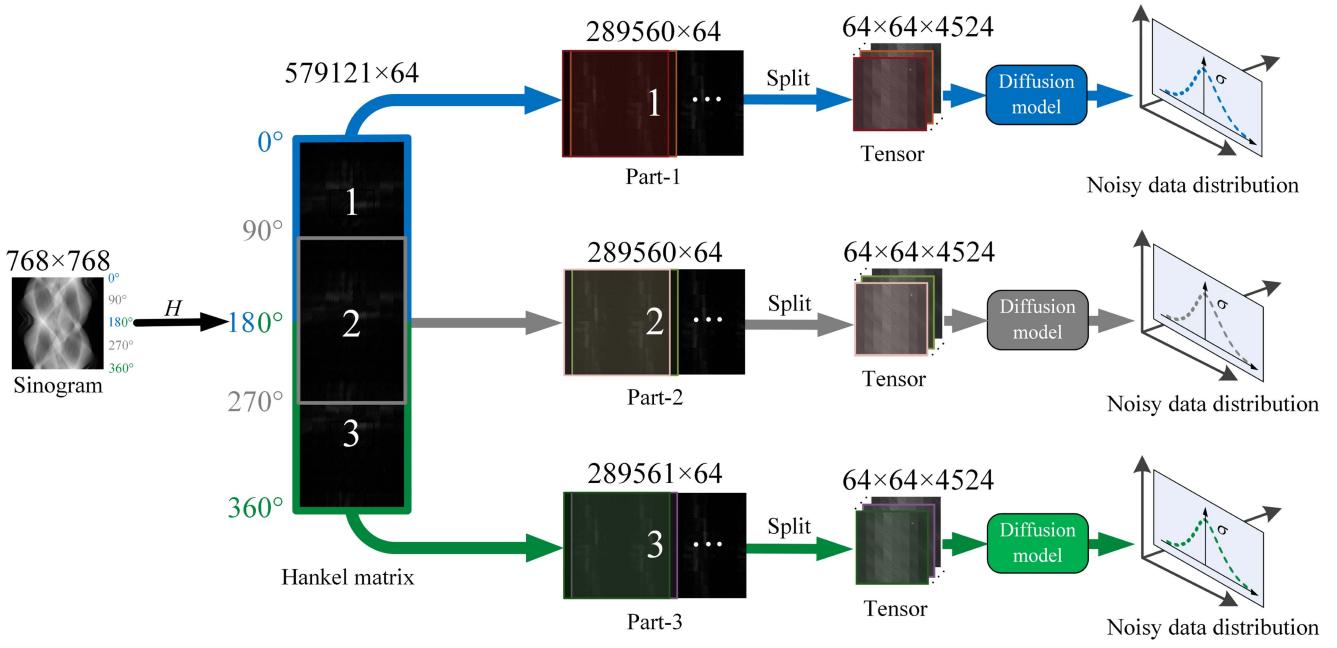


Fig. 3. The training stage of RAP model. First, the projection data is used to construct a low-rank Hankel matrix. Next, three subsets of Hankel matrices are extracted from the Hankel matrix for parallel processing. Finally, each of the three subsets of Hankel matrices is randomly divided into multiple tensors, and denoising score matching is applied to train the network to learn the gradient distribution.

where $z \sim \mathcal{N}(0, 1)$, $x(0) \sim p_0$, and $\sigma_0 = 0$ are chosen to simplify the notation. The above formulation is repeated for $i = N - 1, \dots, 0$. Thus, previous discrete process turns into continuous diffusion process. With adding the conditional constraints to (15), it can be rewritten as follows:

$$\begin{aligned} x^i &= x^{i+1} + (\sigma_{i+1}^2 - \sigma_i^2) \nabla_x [\log p_t(y|x^{i+1}) + \log p_t(x_{LR}^{i+1}) \\ &\quad + \log p_t(x_{TV}^{i+1})] + \sqrt{\sigma_{i+1}^2 - \sigma_i^2} z \end{aligned} \quad (16)$$

In the above equation, $\log p_t(y|x)$ stems from sinogram data knowledge. $\log p_t(x_{LR})$ is derived from low-rank (LR) prior and $\log p_t(x_{TV})$ comes from TV prior.

The flowchart of the iterative reconstruction process is shown in Fig. 4. During the inference stage, iterations are performed between the iterative updates of the numerical SDE solver and the data consistency step to achieve reconstruction.

LR Step: Following the Hankel transform (HT) operation, each updated sample x is further transformed into three distinct new Hankel matrices H_1^i , H_2^i , and H_3^i based on different segments of continuous view information.

$$[H_1^i, H_2^i, H_3^i] = H_p^i \leftarrow H(x^i) \quad (17)$$

The local subset of H_s^i is described as follows:

$$H_s^i = \begin{cases} \{H_p^i(\theta) | \theta \in [0^\circ, 180^\circ]\} = H_s^i(0^\circ, 180^\circ), & s = 1 \\ \{H_p^i(\theta) | \theta \in [90^\circ, 270^\circ]\} = H_s^i(90^\circ, 270^\circ), & s = 2 \\ \{H_p^i(\theta) | \theta \in [180^\circ, 360^\circ]\} = H_s^i(180^\circ, 360^\circ), & s = 3 \end{cases} \quad (18)$$

where θ represents the view of the corresponding projection data. The size of the Hankel matrix H_p^i is 579121×64 .

In order to facilitate processing and analysis, the first step is to use singular value decomposition (SVD) to decompose the Hankel matrices H_1^i , H_2^i , and H_3^i .

$$[U \Delta V^T] = svd[H_1^i, H_2^i, H_3^i]^T \quad (19)$$

where U is an orthogonal matrix, Δ is a diagonal matrix with non-negative diagonal elements, and V is an orthogonal matrix. while $U_{[k]}$, $\Delta_{[k]}$ and $V_{[k]}$ represents the first K columns of U , Δ and V , respectively:

$$\begin{aligned} U_{[k]} &= [u_1, \dots, u_k, \dots, u_K] \\ V_{[k]} &= [v_1, \dots, v_k, \dots, v_K] \\ \Delta_{[k]} &= [\delta_1, \dots, \delta_k, \dots, \delta_K] \end{aligned} \quad (20)$$

The hard-threshold (hard-THR) singular value process can be expressed as:

$$H_{[k]}^i = U_{[k]} \Delta_{[k]} V_{[k]}^T \quad (21)$$

where $H_{[k]}^i$ represents the matrix H reconstructed from the first K eigenvectors. The SVD is particularly useful for ill-posed linear problems with nearly degenerate matrices because it provides the best approximation with lower rank.

After low-rank processing, the Hankel matrix is transformed back to the sinogram through the inverse Hankel transform (IHT) operation denoted by $H^+(\bullet)$, represented as follows:

$$x^i \leftarrow H^+(H_{[k]}^i) \quad (22)$$

TV Step: TV minimization is also employed to remove noise and artifacts. Suppose $\Delta x = \|x - x^i\|$, TV minimization can be

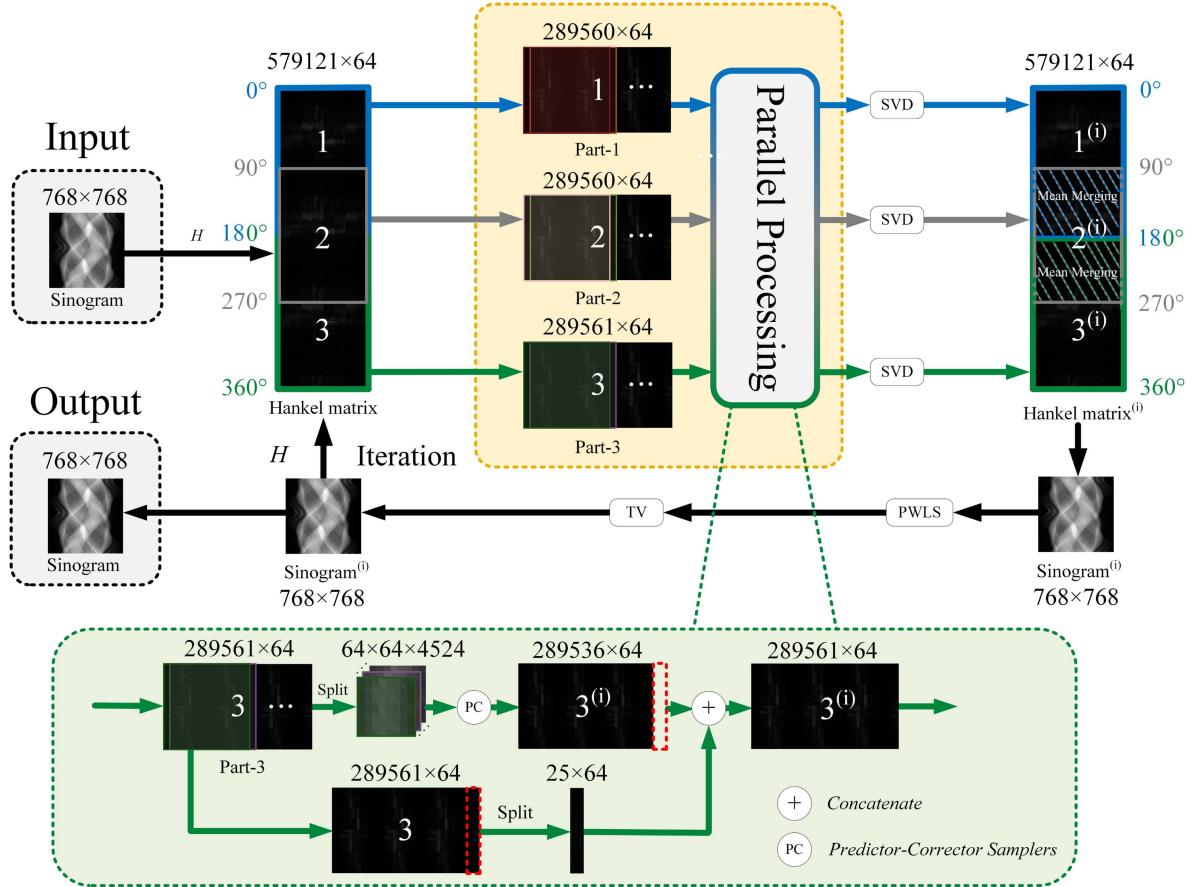


Fig. 4. The pipeline for iterative reconstruction stage of RAP model. During the reconstruction stage, low-dose CT reconstruction is performed by employing parallel processing of multiple partitioned Hankel matrices.

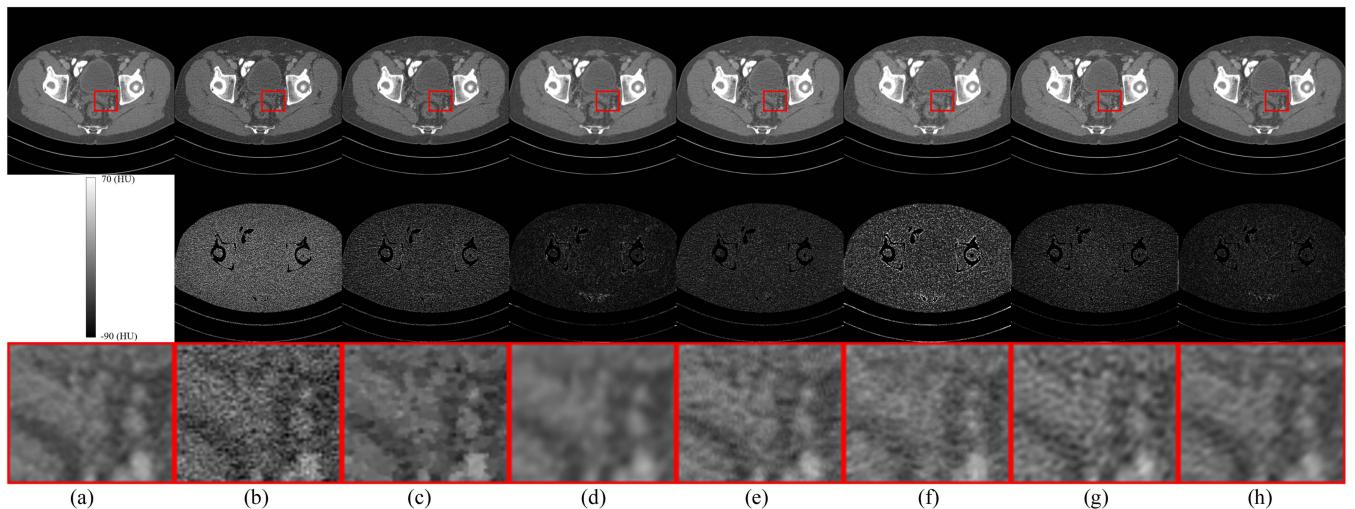


Fig. 5. Reconstruction results from $5e4$ noise level using different methods. (a) The reference image versus the images reconstructed by (b) FBP, (c) SART-TV, (d) CNN, (e) NCSN++, (f) U-ViT, (g) OSDM, (h) RAP (50). The display windows are $[-90, 70]$ HU. The second row shows the error maps of the reconstruction, and the third row shows the enlarged view of the ROI (indicated by the red box in the first row).

stated as follows:

$$TV(x^i) = x^{i+1} - \alpha \Delta x \frac{\nabla \|x^i\|_{TV}}{\|\nabla \|x^i\|_{TV}\|} \quad (23)$$

where α is the length of each gradient-descent step.

DC Step: To improve the noise immunity, the statistical properties of the projection data can be incorporated into the objective function [29], [32]. Additionally, there exists a statistical method for sinogram denoising that utilizes the PWLS method to obtain the best assessment from noisy sinograms. The PWLS prior is integrated into a regularized objective function that is expressed as:

$$\begin{aligned} x^i = \arg \min_x & \left[\|y - x^{i+1}\|_W^2 + \lambda_1 \|x^{i+1} - H^+(H_{[K]}^{i+1})\|_2^2 \right. \\ & \left. + \lambda_2 \|x^{i+1}\|_{TV}^2 \right] \end{aligned} \quad (24)$$

where hyperparameters λ_1 and λ_2 balance the trade-off among the terms of PWLS, LR and TV. $i = N - 1, \dots, 0$ denotes the iteration of outer loop. Specifically, the standard PWLS can be described as follows:

$$x^i = \arg \min_x [(x^{i+1} - y)^T W (x^{i+1} - y) + \mu R(x^{i+1})] \quad (25)$$

where superscript T represents the transposing operation. Equation (23) can be further solved as:

$$x^i = \frac{W(y - x^{i+1}) + \mu R'(x^{i+1})}{W + \mu} \quad (26)$$

In order to decrease the influence of noise, the scale coefficient η for system calibration is set to 22000.

$$W = diag\{w_i\} = diag\{1/\sigma_{x_i}^2\} = diag\{1/l_1 \exp(x_i/\eta)\} \quad (27)$$

Regarding the corrector, it transforms any initial sample $x(t)$ into the final sample $x(0)$ through the following process, which is referred to as Langevin dynamics:

$$\begin{aligned} x^{i,j} & \leftarrow x^{i,j-1} + \varepsilon_i s_\theta(x^{i,j-1}; \sigma_i) + \sqrt{2\varepsilon_i} z \\ i &= N - 1, \dots, 0, \quad j = 1, 2, \dots, M \end{aligned} \quad (28)$$

where $\varepsilon_i > 0$ is the step size, and $z \sim \mathcal{N}(0, 1)$ refers to a standard normal distribution. The above formulation is repeated for $i = N - 1, \dots, 0$, $j = 1, 2, \dots, M$. The theory of Langevin dynamics guarantees that when $M \rightarrow \infty$ and $\varepsilon_i \rightarrow 0$, x^i is a sample from $p_t(x)$ under designated conditions.

Sampling is not performed directly from the distribution $p(x)$, but rather through the posterior distribution $p(x|y)$ as explained in Section II. The DC operation can be viewed as a conditional term, which is incorporated into the sampling process of (26) and results in:

$$\begin{aligned} x^{i,j} &= x^{i,j-1} + \varepsilon_i \nabla_x [\log p_t(y|x^{i,j-1}) + \log p_t(x_{LR}^{i,j-1}) \\ &\quad + \log p_t(x_{TV}^{i,j-1})] + \sqrt{2\varepsilon_i} z \end{aligned} \quad (29)$$

Once the reconstructed projection x is obtained, the final image \tilde{I} can be obtained through $F^{-1}(\bullet)$:

$$\tilde{I} = F^{-1}(x) \quad (30)$$

Algorithm 1: Iterative Reconstruction of RAP.

Training Stage

Dataset: Few projection domain samples x

1: Repeat

2: $x \sim p(x)$, $t \sim \mathcal{U}([0, T])$, $\varepsilon \sim \mathcal{N}(0, I)$

3: $x(t) = x(0) + \varepsilon \sigma(t)$

4: Take a gradient descent step on $\nabla_\theta \|s_\theta(x(t), t) + \varepsilon\|_2^2$

5: Until converged

6: Trained RAP network

Iterative Reconstruction Stage

Setting: $s_\theta, N, M, \sigma, \varepsilon$

1: Initial data $x = F(I)$ (**FP**)

2: $x^N \sim \mathcal{N}(0, \sigma_{\max}^2 I)$

3: For $i = N - 1$ to 0 **do** (**Outer loop**)

4: $x^i \leftarrow Predictor(x^{i+1}, \sigma_i, \sigma_{i+1}, s_\theta)$

5: $[H_1^i, H_2^i, H_3^i] = H_p^i \leftarrow H(x^i)$ (**HT**)

6: H_1^i, H_2^i, H_3^i via Eq. (17)

7: $[U \Delta V^T] = svd[H_1^i, H_2^i, H_3^i]^T$ (**SVD**)

8: $H_{[k]}^i = U_{[k]} \Delta_{[k]} V_{[k]}^T$ (**hard-THR**)

9: $x^i \leftarrow H^+(H_{[k]}^i)$ (**IHT**)

10: $x^i = \frac{W(y - x^{i+1}) + \mu R'(x^{i+1})}{W + \mu}$ (**PWLS**)

11: $x^i = TV(x^i)$ (**TV**)

12: **For** $j = 1$ to M **do** (**Inner loop**)

13: $x^{i,j} \leftarrow Corrector(x^{i,j-1}, \sigma_i, \varepsilon_i, s_\theta)$

14: **Repeat from step 5 to step 11**

15: **End for**

16: **End for**

17: Final image $\tilde{I} = F^{-1}(x)$ (**FBP**)

18: **Return** \tilde{I}

Algorithm 1 describes a detailed description of the training and iterative reconstruction algorithm for the RAP model. The entire process of RAP model reconstruction consists of two loops. In the outer loop, a trained network is used for prediction, followed by correction in the inner loop. The predictor and corrector function together as a whole are utilized to generate the final samples. Additionally, after the operation of both the predictor and corrector, a data fidelity term is applied to ensure the quality of the generated images.

IV. EXPERIMENTS

A. Data Specification

AAPM Challenge Dataset: Mayo Clinic provided simulated human abdominal images for the AAPM Low-Dose CT Grand Challenge to evaluate the performance of different algorithms in low-dose CT imaging [33]. The dataset consists of 2588 normal-dose CT images with a resolution of 512×512 and a slice thickness of 3 mm from 10 patients. After processing, the images were transformed into 7764 normal-dose CT images with a resolution of 512×512 and a slice thickness of 1 mm. Projection data with different noise levels (1e5, 5e4,

and 1e4) were generated by adding Poisson noise to sinograms obtained from the normal-dose CT images. The artifact-free images generated from the normal-dose projection data using the FBP algorithm can be considered as ground truth. For fan-beam CT reconstruction, the Siddon's ray-driven algorithm [34] was utilized to generate the projection data. The distance from the rotation center to the source and detector was set to 40 cm. The detector width was 41.3 cm, composed of 720 detector elements. Over 360 projection views were uniformly distributed throughout the entire imaging process. In this study, a total of 4742 projection data were selected for the training set. All models were validated using 537 sinograms, and 12 sinograms were chosen as the test set.

Somatom Confidence CT Dataset: The generalization study was evaluated on a subset of CT scans from the dataset proposed by Zeng et al. [35]. This dataset comprises data from 307 patients who underwent VMAT treatment, with approximately 100 radiation dose maps and 100 corresponding CT images per patient. The CT images were acquired using Somatom Confidence (Siemens Health-care, Forchheim, Germany) and were in DICOM format with a resolution of $256 \times 256 \times 160$. The CT images were resampled to a resolution of 512×512 , and 12 CT scan images were selected as the test set for the generalization study.

B. Model Training and Parameter Selection

In the experiments, the RAP model is implemented in Python using Operator Discretization Library (ODL) [36] and PyTorch on a personal workstation with 2 GPUs (NVIDIA TITAN Xp-12 GB). To accurately simulate real-world conditions, Poisson noise is added to the testing slices, with X-ray source intensities of $a_i = 1e5$, $a_i = 5e4$, and $a_i = 1e4$. During the training stage, the RAP model utilizes the Adam algorithm with a specific learning rate 0.001 and Kaiming initialization for weight initialization. The reconstruction process involves outer and inner iterations, with $N = 300$ outer iterations and $M = 2$ inner iterations. The correction process in the inner loop is iterated twice using annealing Langevin. η in PWLS scheme is set to a constant 22000. The number of iterations t in TV minimization is set to 2. The singular value thresholding in SVD operation is 38 and the sliding window size is 8×8 . For the convenience of reproducibility, the source code and some representative results are available at: <https://github.com/yqx7150/RAP>.

C. Evaluation Metrics

The evaluation of reconstructed data quality involves the use of quantitative metrics such as peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), and mean squared error (MSE). These metrics are calculated by averaging the results from 12 test data points.

PSNR describes the maximum possible power of the signal in relation to the noise corrupting power. Higher PSNR values correspond to better quality reconstructions. Denoting I and \tilde{I} to be the estimated reconstruction and ground-truth, PSNR is expressed as:

$$PSNR(I, \tilde{I}) = 20\log_{10} \left[MAX(\tilde{I}) / \left\| I - \tilde{I} \right\|_2 \right] \quad (31)$$

The SSIM value is used to measure the similarity between the ground-truth and reconstruction. SSIM is defined as:

$$SSIM(I, \tilde{I}) = \frac{(2\mu_I\mu_{\tilde{I}} + c_1)(2\sigma_{I\tilde{I}} + c_2)}{(\mu_I^2 + \mu_{\tilde{I}}^2 + c_1)(\sigma_I^2 + \sigma_{\tilde{I}}^2 + c_2)} \quad (32)$$

where μ_I and σ_I^2 are the average and variances of I . $\sigma_{I\tilde{I}}$ is the covariance of I and \tilde{I} . c_1 and c_2 are used to maintain a stable constant. MSE is employed to evaluate the errors and defined as:

$$MSE(I, \tilde{I}) = \frac{1}{W} \sum_{i=1}^W \left\| I_i - \tilde{I}_i \right\|_2 \quad (33)$$

where W is the number of pixels within the reconstruction result. If MSE approaches to zero, the reconstructed image is closer to the reference image.

D. Experimental Evaluation

AAPM Challenge Dataset Comparison: We compare the proposed RAP model with six baseline techniques in low-dose CT reconstruction, including FBP [1], SART-TV [37], CNN [38], NCSN++ [16], U-ViT [17] and OSDM [18]. The CNN, NCSN++, and U-ViT models are trained using 4742 projection data. In addition, the OSDM and RAP models use the same amount of projection data for training (1 and 50 projection data). The involved parameters are set following the guidelines in their original papers. This comparative analysis aims to assess the performance and effectiveness of these techniques in low-dose CT reconstruction.

In low-dose CT reconstruction, experiments with different noise levels are conducted by setting 1e5, 5e4, and 1e4 photons along each path of the X-ray. Table I shows the PSNR, SSIM, and MSE values of the reconstructed images. The best PSNR and MSE values are highlighted in bold. The reconstructed images produced by RAP method exhibit fewer artifacts and less noise. It is evident that compared to FBP, RAP exhibits a significant increase in average PSNR in the noise levels of 1e5, 5e4, and 1e4, even when trained with only one sample. The gains are 8.67 dB, 9.1 dB, and 12.17 dB, respectively.

In order to further highlight its advantages, Figs. 6 and 7 display reconstructed images and residual images under different levels of noise. In comparative methods, the FBP algorithm performs the worst due to its high sensitivity to noise, resulting in significant artifacts and incomplete structural organization in the reconstructed images. The performance of the SART-TV algorithm is better than FBP, but it still lacks complex structural details. Meanwhile, CNN sacrifices some edge details and overly smooths the edges. Compared to previous methods, among NCSN++, U-ViT and OSDM exhibit improved reconstruction results, but they do not effectively capture texture details. In contrast, the images reconstructed by RAP exhibit preserved fine details and ensure the integrity of the structure.

After increasing the noise level from 5e4 to 1e4, there is a noticeable differentiation in the quality of the reconstructed images. Images generated by FBP and SART-TV exhibit prominent structural feature omissions and significant noise interference (Fig. 6(b) and (c)), while those generated by CNN tend to overly smooth the edges in the reconstructed images (Fig. 6(d)).

TABLE I
RECONSTRUCTION PSNR/SSIM/MSE OF AAPM CHALLENGE DATA USING DIFFERENT METHODS AT DIFFERENT NOISE LEVEL

Method (Samples) \ Noise level	$a_i=1e5$	$a_i=5e4$	$a_i=1e4$
FBP	34.62/0.9252/3.66e-4	32.43/0.8866/5.81e-4	25.78/0.6897/2.69e-3
SART-TV	41.03/0.9892/8.65e-5	38.72/0.9786/1.39e-4	29.58/0.8710/1.18e-3
CNN	41.35/0.9903/8.45e-5	39.24/0.9859/1.34e-4	37.55/ 0.9793 /1.99e-4
NCSN++	41.52/0.9873/7.19e-5	40.19/0.9810/9.66e-5	37.30/0.9637/1.88e-4
U-ViT	40.16/0.9383/9.76e-5	37.30/0.8920/1.89e-4	34.85/0.8471/3.29e-4
OSDM	42.62/0.9899/5.51e-5	41.20/0.9857/7.74e-5	37.43/0.9683/1.83e-4
RAP (1)	43.29/0.9906/4.71e-5	41.53/0.9861/7.10e-5	37.95/0.9711/1.62e-4
RAP (50)	43.34/0.9907/4.67e-5	41.60/0.9864/6.98e-5	38.01/0.9716/1.60e-4

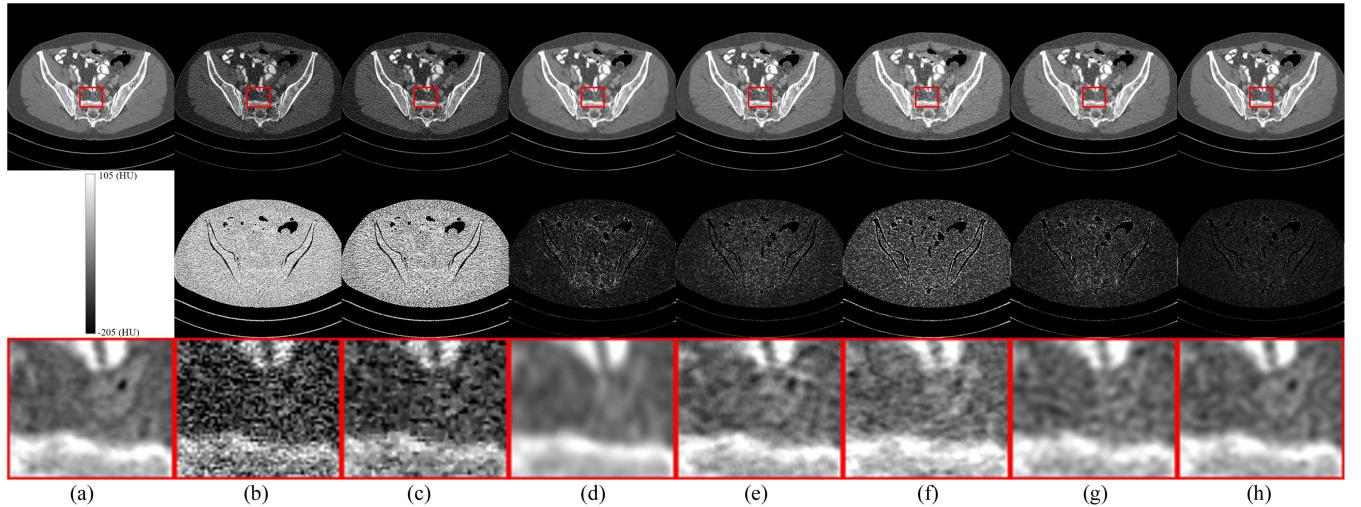


Fig. 6. Reconstruction results from 1e4 noise level using different methods. (a) The reference image versus the images reconstructed by (b) FBP, (c) SART-TV, (d) CNN, (e) NCSN++, (f) U-ViT, (g) OSDM, (h) RAP (1). The display windows are [-205, 105] HU. The second row shows the error maps of the reconstruction, and the third row shows the enlarged view of the ROI (indicated by the red box in the first row).

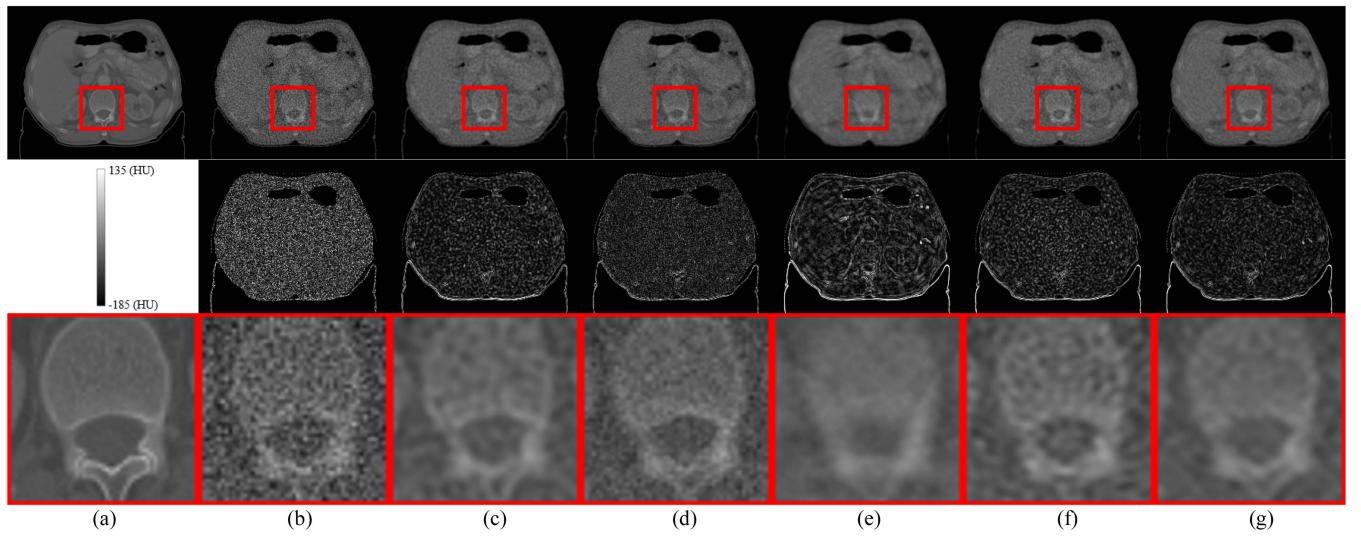


Fig. 7. Reconstruction results from 1e4 noise level using different methods. (a) The reference image versus the images reconstructed by (b) FBP, (c) CNN, (d) NCSN++, (e) U-ViT, (f) OSDM, and (g) RAP (1). The display windows are [-185, 135] HU. The second row shows the error maps of the reconstruction, and the third row shows the enlarged view of the ROI (indicated by the red box in the first row).

TABLE II
RECONSTRUCTION PSNR/SSIM/MSE OF SOMATOM CONFIDENCE CT DATA USING DIFFERENT METHODS AT DIFFERENT NOISE LEVEL

Noise level	FBP	CNN	NCSN++	U-ViT	OSDM	RAP (1)
$a_i=1e5$	40.41/0.9763/9.31e-5	35.65/0.9773/4.45e-4	40.98/0.9855/8.91e-5	43.24/0.9561/4.83e-5	44.10/0.9911/4.09e-5	45.38/0.9938/3.05e-5
$a_i=1e4$	30.53/0.815/9.06e-4	34.63/0.9617/4.74e-4	35.86/0.9377/2.71e-4	37.73/0.9098/1.78e-4	38.09/0.9646/1.68e-4	40.02/0.9779/1.05e-4
$a_i=5e2$	19.42/0.2742/1.17e-2	26.86/0.6382/2.13e-3	28.06/0.7319/1.74e-3	29.25/0.5548/1.39e-3	30.67/0.7987/1.01e-3	32.88/0.8808/6.15e-4

Furthermore, the reconstructed images (Fig. 6(f)) by U-ViT still have a lot of noise interference. Although the reconstruction quality of images by NCSN++ and OSDM are improved, noticeable stripe artifacts still appear due to high noise interference (Fig. 6(e) and (g)). In contrast, under the high noise level of 1e4, RAP is able to preserve complex structural details to the maximum extent while effectively suppressing stripe artifacts (Fig. 6(h)). Furthermore, at this noise level, RAP demonstrates superior reconstruction performance compared to other contrast methods while being trained under the condition of maintaining a single sample, thus greatly reducing the demand for data samples.

Generalization Test: To better explore the generalization of the RAP model, we apply the comparative methods to the Somatom Confidence CT dataset. Furthermore, the pretrained model with a noise level of 1e5 is used to reconstruct the CNN with noise levels of 1e4 and 5e2. The NCSN++, U-ViT, OSDM, and RAP still use the old models with prior knowledge obtained from the AAPM challenge dataset. The quantitative evaluations of PSNR, SSIM, and MSE metrics are presented in Table II. The RAP outperforms other methods in various indicators, demonstrating its superior performance. Additionally, Fig. 7 displays representative reconstruction results and error map of RAP and other methods.

E. Ablation Study

We conduct ablation studies on multiple modules of the RAP and investigate the effectiveness of the proposed partitioned dimension in low-dose CT reconstruction.

Exploring Partitioned Dimension: In this subsection, the impact of different partitioned dimension on the reconstruction performance is primarily investigated, as well as how to effectively improve within the same partitioned dimension. Furthermore, various metrics are employed in the experiments, including partitioned Hankel dimension (i.e., Size Per Hankel, SPH), in conjunction with PSNR, SSIM, and MSE, as shown in Table III.

To ensure a fair comparison, in this experiment, the comparison conditions were standardized by training with the same single sample and testing with the same dataset. We perform low-dose CT reconstruction under a noise level of 1e4 and use dual-partition as the baseline. For the dual-partition (i.e., evenly splitting the Hankel matrix), the two sizes of the Hankel matrices are 289560×64 and 289561×64 , respectively, and are reconstructed separately. Similarly, for the triple-partition, the Hankel matrix is divided into three equal parts, resulting in sizes of 193040×64 , 193040×64 , and 193041×64 for the Hankel matrices. When the triple-partition is extended to

TABLE III
COMPARISON OF PSNR/SSIM/MSE OF FOUR DIFFERENT PARTITIONED DIMENSIONS

Partitions	SPH	PSNR	SSIM	MSE
Dual	(289560,64) (289561,64)	37.60	0.9694	1.79e-4
Triple	(193040,64) (193040,64) (193041,64)	37.71	0.9701	1.72e-4
Triple*	(289560,64) (289560,64) (289561,64)	37.95	0.9711	1.62e-4
Quad	(144780,64) (144780,64) (144780,64) (144781,64)	37.05	0.9665	2.03e-4

quad-partition, the Hankel matrix is divided into four equal parts, and the four resulting Hankel matrices are reconstructed separately.

The triple*-partition (RAP) further improves upon the triple-partition method by obtaining three Hankel matrices (part-1, part-2, part-3) with sizes 289560×64 , 289561×64 , and 289560×64 , respectively. The specific configuration is illustrated in Fig. 4, where part-1 and part-2 are obtained from a Hankel matrix with a size of 579121×64 . Additionally, a Hankel matrix with a size of 289560×64 is extracted from the middle and used as a new part (part-3). After separate reconstructions, part-3 is concatenated back to its original position. For the overlapping Hankel regions of part-1/2, a mean combination is performed to reconstruct a new Hankel matrix with a size of 579121×64 .

As shown in Table III, the triple*-partition approach outperforms other partitioning schemes in terms of PSNR, SSIM and MSE, achieving optimal results. This is also the way adopted by the RAP.

The reconstruction results for different block partitioning dimensions are depicted in Fig. 8. It is evident that, under a high noise level of 1e4, the images generated by Dual/Triple/Quad-partition are all significantly affected by noise, resulting in poor preservation of texture details. Conversely, the triple*-partition demonstrates effective noise reduction, thereby minimizing artifacts and retaining structural details to the maximum extent.

Different Components in RAP Model: In order to better understand the role of the corresponding modules in the RAP model, the ablation study analyzes the role of PWLS and TV modules. Specifically, PWLS serves as a fidelity term module, while TV functions is treated as a denoising module. The quantitative results of the reconstructions at noise levels of 1e5, 5e4, and 1e4 are presented in Table IV. This indicates that the selected module has a positive impact on RAP.

TABLE IV
RECONSTRUCTION PSNR/SSIM/MSE OF AAPM CHALLENGE DATA USING DIFFERENT METHODS AT DIFFERENT NOISE LEVEL

Noise level	(w/o) PWLS&TV	(w/o) TV	(w/o) PWLS	RAP (50)
$a_i = 1e5$	$42.28 \pm 0.49 / 0.9889 \pm 0.0015 / 5.95e-5$	$42.69 \pm 0.55 / 0.9892 \pm 0.0012 / 5.43e-5$	$42.87 \pm 0.68 / 0.9900 \pm 0.0016 / 5.22e-5$	$43.34 \pm 0.49 / 0.9907 \pm 0.0008 / 4.67e-5$
$a_i = 5e4$	$40.71 \pm 0.49 / 0.9841 \pm 0.0015 / 8.56e-5$	$40.89 \pm 0.60 / 0.9846 \pm 0.0014 / 8.21e-5$	$41.29 \pm 0.70 / 0.9854 \pm 0.0028 / 7.51e-5$	$41.60 \pm 0.60 / 0.9864 \pm 0.0019 / 6.98e-5$
$a_i = 1e4$	$35.64 \pm 0.67 / 0.9515 \pm 0.0066 / 2.76e-4$	$36.16 \pm 0.57 / 0.9541 \pm 0.0053 / 2.44e-4$	$37.11 \pm 0.68 / 0.9672 \pm 0.0064 / 1.97e-5$	$38.01 \pm 0.65 / 0.9716 \pm 0.0035 / 1.60e-4$

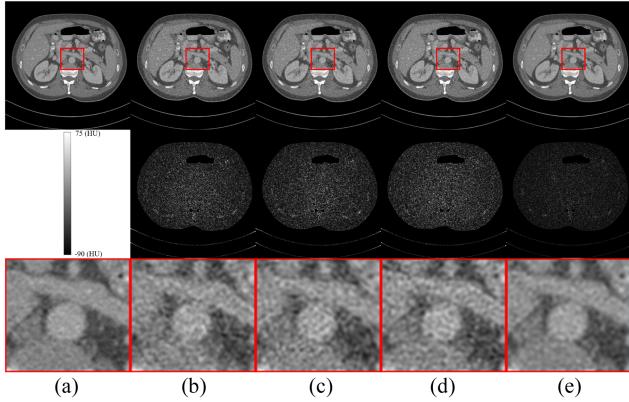


Fig. 8. Reconstruction results from $1e4$ noise level using different partitioned dimensions. (a) The reference image versus the images reconstructed by (b) dual-partition, (c) triple-partition, (d) quad-partition, (e) Triple*-partition. The display windows are $[-90, 75]$ HU. The second row shows the error maps of the reconstruction, and the third row shows the enlarged view of the ROI (indicated by the red box in the first row).

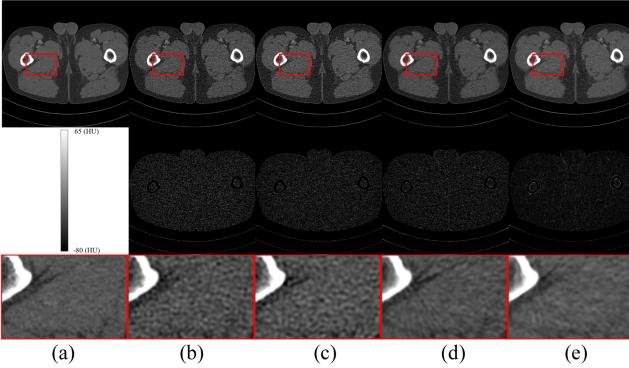


Fig. 9. Reconstruction results from $1e4$ noise level using different methods. (a) The reference image versus the images reconstructed by (b) (w/o) PWLS&TV, (c) (w/o) TV, (d) (w/o) PWLS, (e) RAP (50). The display windows are $[-80, 65]$ HU. The second row shows the error maps of the reconstruction, and the third row shows the enlarged view of the ROI (indicated by the red box in the first row).

Fig. 9 shows the reconstructed images at a noise level of $1e4$. Without the utilization of the PWLS and TV modules, the reconstructed image is significantly impacted by noise, resulting in the degradation of structural details and the prominent existence of artifacts. After the PWLS fidelity module is applied, there is some improvement in the reconstruction values, but the noise removal is still unsatisfactory. Furthermore, with the introduction of the TV module, improvements are noticed in the edge texture of the reconstructed image, along with augmented denoising capabilities. However, it is still accompanied by the

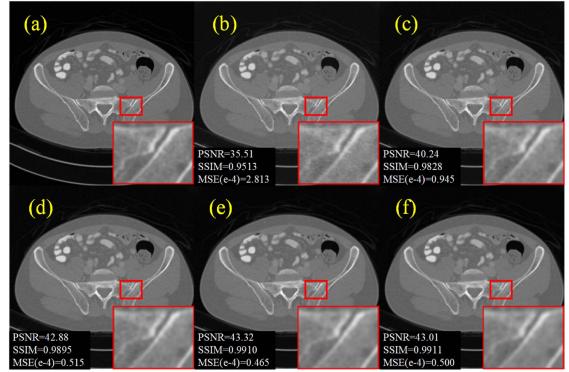


Fig. 10. Reconstruction steps for RAP. (a) Ground truth, (b) LDCT image, (c) step = 1, (d) step = 5, (e) step = 10, (f) step = 15. Quantitative results are provided in the bottom-left corner. The tissue boundaries of the images are enlarged in the bottom-right corner.

generation of undesired artifacts. In contrast, simultaneously applying the PWLS and TV modules can achieve a balance between detail preservation and noise reduction. Even at high noise levels, the results of this combination can still achieve better reconstruction outcomes.

V. DISCUSSION

The input of the RAP model shifts from pure noise to low-dose CT projection data, significantly reducing the number of iterations with the assistance of the TV denoising module, as shown in Fig. 10.

Significantly, in the experimental evaluation with a noise level of $1e5$, the PSNR of the RAP model peaks at the 10th step, whereas most diffusion models typically require 500 to 2000 steps [16], [18], [39], [40]. To achieve faster convergence and minimize the required time, we enforce TV denoising operations during the iterative reconstruction process of the RAP model. While a common practice for many image reconstruction diffusion models involves inputting under sampled images alongside introducing Gaussian noise to aid the reconstruction process, our observations during testing reveal that providing solely low-dose CT projection data without additional Gaussian noise enables the network to reach the desired outcome more rapidly. This approach not only accelerates the entire process but also notably reduces the required number of iterations.

VI. CONCLUSION

In recent years, deep learning-based CT reconstruction methods have rapidly developed, yet challenges persist in enhancing the generalizability and robustness of these models. This paper

introduces a low-rank multi-diffusion model utilizing an angular segmentation strategy for low-dose CT reconstruction. By segmenting consecutive views and structuring them into partitioned Hankel matrices, redundant prior information can be effectively extracted and utilized across multiple angular segments, enhancing the few-shot learning capacity of the model. During the iterative reconstruction stage, a numerical SDE solver is alternated with PWLS fidelity and TV regularization, facilitating efficient and accurate reconstruction. The implementation of a triple-partition parallel processing technique further optimizes both training and testing phases. The method demonstrates robust generalization and effectiveness on the AAPM Challenge dataset and Somatom Confidence CT dataset. Given the limitations in medical data collection, future work will focus on integrating multi-domain and multi-model strategies to fully exploit prior knowledge from limited samples, thereby improving the accuracy and quality of reconstructed images.

ACKNOWLEDGMENT

The authors sincerely thank the anonymous referees for their valuable comments on this work. All authors declare that they have no known conflicts of interest in terms of competing financial interests or personal relationships.

REFERENCES

- [1] D. J. Brenner and E. J. Hall, "Computed tomography—an increasing source of radiation exposure," *New England J. Med.*, vol. 357, no. 22, pp. 2277–2284, 2007.
- [2] D. J. Brenner and H. Hricak, "Radiation exposure from medical imaging: Time to regulate?," *JAMA*, vol. 304, no. 2, pp. 208–209, 2010.
- [3] J. Liu et al., "3D feature constrained reconstruction for low dose CT imaging," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 5, pp. 1232–1247, May 2018.
- [4] H. Chen et al., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017.
- [5] Q. Ding et al., "Low-dose CT with deep learning regularization via proximal forward-backward splitting," *Phys. Med. Biol.*, vol. 65, no. 12, 2020, Art. no. 125009.
- [6] X. Yin et al., "Domain progressive 3D residual convolution network to improve low-dose CT imaging," *IEEE Trans. Med. Imag.*, vol. 38, no. 12, pp. 2903–2913, Dec. 2019.
- [7] M. Balda, J. Hornegger, and B. Heismann, "Ray contribution masks for structure adaptive sinogram filtering," *IEEE Trans. Med. Imag.*, vol. 31, no. 6, pp. 1228–1239, Jun. 2012.
- [8] A. Manduca et al., "Projection space denoising with bilateral filtering and CT noise modeling for dose reduction in CT," *Med. Phys.*, vol. 36, no. 11, pp. 4911–4919, 2009.
- [9] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Trans. Med. Imag.*, vol. 31, no. 9, pp. 1682–1697, Sep. 2012.
- [10] Z. Tian et al., "Low-dose CT reconstruction via edge-preserving total variation regularization," *Phys. Med. Biol.*, vol. 56, no. 18, 2011, Art. no. 5949.
- [11] Y. Zhang et al., "Few-view image reconstruction with fractional-order total variation," *J. Opt. Soc. America A*, vol. 31, no. 5, pp. 981–995, 2014.
- [12] S. V. M. Sagheer and S. N. George, "Denoising of low-dose CT images via low-rank tensor modeling and total variation regularization," *Artif. Intell. Med.*, vol. 94, pp. 1–17, 2019.
- [13] L. Yang et al., "Diffusion models: A comprehensive survey of methods and applications," *Assoc. Comput. Machinery Comput. Surv.*, vol. 56, pp. 1–39, 2023.
- [14] Q. Gao and H. Shan, "CoCoDiff: A contextual conditional diffusion model for low-dose CT image denoising," *Proc. SPIE*, vol. 12242, 2022, Art. no. 122420I.
- [15] Y. Li et al., "Generation model meets swin transformer for unsupervised low-dose CT reconstruction," *Mach. Learn. Sci. Technol.*, vol. 5, 2024, Art. no. 025005.
- [16] Y. Song et al., "Score-based generative modeling through stochastic differential equations," in *Proc. Int. Conf. Learn. Representations*, 2021.
- [17] F. Bao et al., "All are worth words: A ViT backbone for diffusion models," in *Proc. Comput. Vis. Pattern Recognit.*, 2023, pp. 22669–22679.
- [18] B. Huang, S. Lu, L. Zhang, B. Lin, W. Wu, and Q. Liu, "One-sample diffusion modeling in projection domain for low-dose CT imaging," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 8, no. 8, pp. 902–915, Nov. 2024.
- [19] Y. Song et al., "Solving inverse problems in medical imaging with score-based generative models," in *Proc. Int. Conf. Learn. Representations*, 2022.
- [20] B. Jiang, J. Wang, Y. Lu, G. Lu, and D. Zhang, "Multilevel noise contrastive network for few-shot image denoising," *IEEE Trans. Instrum. Meas.*, vol. 71, 2022, Art. no. 5019413.
- [21] Z. Dai et al., "Few-shot medical image classification using prior guided feature enhancement," *Pattern Recognit.*, vol. 134, 2023, Art. no. 109108.
- [22] A. Zhao et al., "Data augmentation using learned transformations for one-shot medical image segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 8543–8553.
- [23] K. H. Jin and J. C. Ye, "Sparse and low-rank decomposition of a Hankel structured matrix for impulse noise removal," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1448–1461, Mar. 2018.
- [24] K. H. Jin et al., "MRI artifact correction using sparse plus low-rank decomposition of annihilating filter-based Hankel matrix," *Magn. Reson. Med.*, vol. 78, no. 1, pp. 327–340, Jul. 2017.
- [25] K. H. Jin, D. Lee, and J. C. Ye, "A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank Hankel matrix," *IEEE Trans. Comput. Imag.*, vol. 2, no. 4, pp. 480–495, Dec. 2016.
- [26] G. Wang, J. C. Ye, K. Mueller, and J. A. Fessler, "Image reconstruction is a new frontier of machine learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1289–1296, Jun. 2018.
- [27] V. Sushko, J. Gall, and A. Khoreva, "One-shot GAN: Learning to generate samples from single images and videos," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 2596–2600.
- [28] Z. Zheng, J. Xie, and P. Li, "Patchwise generative convnet: Training energy-based models from a single natural image for internal learning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 2961–2970.
- [29] J. Wang, T. Li, H. Lu, and Z. Liang, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose X-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 25, no. 10, pp. 1272–1283, Oct. 2006.
- [30] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2015, pp. 234–241.
- [31] Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, vol. 32, pp. 11918–11930.
- [32] I. A. Elbakri and J. A. Fessler, "Statistical image reconstruction for polyenergetic X-ray computed tomography," *IEEE Trans. Med. Imag.*, vol. 21, no. 2, pp. 89–99, Feb. 2002.
- [33] Low Dose CT Grand Challenge, 2017. Accessed: Apr. 6, 2017. [Online]. Available: <http://www.aapm.org/GrandChallenge/LowDoseCT/>
- [34] R. L. Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," *Med. Phys.*, vol. 12, no. 2, pp. 252–255, 1985.
- [35] L. Zeng et al., "TransQA: Deep hybrid transformer network for measurement-guided volumetric dose prediction of pre-treatment patient-specific quality assurance," *Phys. Med. Biol.*, vol. 68, no. 20, 2023, Art. no. 205010.
- [36] J. Adler, H. Kohr, and O. Oktem, "Operator discretization library (ODL)," 2017. [Online]. Available: <https://github.com/odlgroup/odl>
- [37] Y. Censor and T. Elfving, "Block-iterative algorithms with diagonally scaled oblique projections for the linear feasibility problem," *SIAM J. Matrix Anal. Appl.*, vol. 24, no. 1, pp. 40–58, 2002.
- [38] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [39] Z. He et al., "Iterative reconstruction for low-dose CT using deep gradient priors of generative model," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 6, no. 7, pp. 741–754, Sep. 2022.
- [40] B. Guan et al., "Generative modeling in sinogram domain for sparse-view CT reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 8, no. 2, pp. 195–207, Feb. 2024.