

# Dual-Domain Collaborative Diffusion Sampling for Multi-Source Stationary Computed Tomography Reconstruction

Zirong Li<sup>ID</sup>, Dingyue Chang<sup>ID</sup>, Zhenxi Zhang<sup>ID</sup>, Fulin Luo<sup>ID</sup>, Senior Member, IEEE, Qiegen Liu<sup>ID</sup>, Jianjia Zhang<sup>ID</sup>, Guang Yang<sup>ID</sup>, Senior Member, IEEE, and Weiwen Wu<sup>ID</sup>, Member, IEEE

**Abstract**—The multi-source stationary CT, where both the detector and X-ray source are fixed, represents a novel imaging system with high temporal resolution that has garnered significant interest. Limited space within the system restricts the number of X-ray sources, leading to sparse-view CT imaging challenges. Recent diffusion models for reconstructing sparse-view CT have generally focused separately on sinogram or image domains. Sinogram-centric models effectively estimate missing projections but may introduce artifacts, lacking mechanisms to ensure image correctness. Conversely, image-domain models, while capturing detailed image features, often struggle with complex data distribution, leading to inaccuracies in projections. Addressing these issues, the Dual-domain Collaborative Diffusion Sampling (DCDS) model integrates sinogram and image domain diffusion processes for enhanced sparse-view reconstruction. This model combines the strengths of both domains in an optimized mathematical framework. A collaborative diffusion

Manuscript received 1 March 2024; revised 10 May 2024; accepted 22 June 2024. Date of publication 28 June 2024; date of current version 24 October 2024. This work was supported in part by the National Key Research and Development Program of China under Grant 2022YFA1204201; in part by the National Natural Science Foundation of China under Grant 62201628, Grant 62101611, Grant 62131015, and Grant 52394162; in part by Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515011375, Grant 2023A1515012278, and Grant 2023A1515011780; and in part by Shenzhen Science and Technology Program under Grant JCYJ20220530145411027 and Grant JCYJ2022081802414031. The work of Guang Yang was supported by the UK Research and Innovation (UKRI) Future Leaders Fellowship under Grant MR/V023799/1. (Zirong Li, Dingyue Chang, and Zhenxi Zhang contributed equally to this work.) (Corresponding authors: Jianjia Zhang; Weiwen Wu.)

Zirong Li, Zhenxi Zhang, Jianjia Zhang, and Weiwen Wu are with the Department of Biomedical Engineering, Sun-Yat-sen University, Shenzhen Campus, Shenzhen 518107, China (e-mail: 1668005606@qq.com; zhangzhx37@mail2.sysu.edu.cn; seuzjj@gmail.com; 3280722767@qq.com).

Dingyue Chang is with China Academy of Engineering Physics, Institute of Materials, Mianyang 621900, China (e-mail: changdy2014@pku.edu.cn).

Fulin Luo is with the College of Computer Science, Chongqing University, Chongqing 400037, China (e-mail: luoflyn@163.com).

Qiegen Liu is with the School of Information Engineering, Nanchang University, Nanchang 330031, China (e-mail: liuqiegen@ncu.edu.cn).

Guang Yang is with the Department of Bioengineering, Faculty of Engineering, Imperial College London, SW7 2AZ London, U.K., and also with the Imperial-X and the School of Biomedical Engineering and Imaging Sciences, King's College London, WC2R 2LS London, U.K. (e-mail: g.yang@imperial.ac.uk).

Digital Object Identifier 10.1109/TMI.2024.3420411

mechanism underpins this model, improving sinogram recovery and image generative capabilities. This mechanism facilitates feedback-driven image generation from the sinogram domain and uses image domain results to complete missing projections. Optimization of the DCDS model is further achieved through the alternative direction iteration method, focusing on data consistency updates. Extensive testing, including numerical simulations, real phantoms, and clinical cardiac datasets, demonstrates the DCDS model's effectiveness. It consistently outperforms various state-of-the-art benchmarks, delivering exceptional reconstruction quality and precise sinogram.

**Index Terms**—Computed tomography, multiple source, image reconstruction, dual-domain, diffusion model, sparse-view.

## I. INTRODUCTION

COMPUTED Tomography (CT) imaging technology has found widespread use in medical diagnostics [1], [2] [3], [4]. Recently, the multi-source stationary CT which is a system where both the detector and the X-ray source remain stationary, as a novel imaging mode with high temporal resolution, has received extensive attention. However, due to the physical constraints of its structure, the number of X-ray sources within the limited space of a single system is finite, which can lead to sparse-view CT (SVCT) imaging issues. The decrease in the number of scanning views may lead to insufficient measurement data, resulting in a decline in imaging quality [3]. Therefore, the challenge of high-quality reconstruction from sparse-view data has emerged as a pioneering challenge. Deep learning-based approaches have demonstrated remarkable success in the field of sparse-view reconstruction [5]. Among these, the most recently proposed score-based model has garnered remarkable results [6]. The score-based model utilizes a score function (neural network) to learn the score (logarithmic gradient of data distribution) to enable high-quality image generation through sampling from random noise [7], [8].

Most prevailing methodologies appear to overlook the fundamental importance of sinogram's precision of the reconstruction outcomes for addressing SVCT reconstruction [9]. A typical example is the utilization of post-processing-based convolutional neural networks [10], which excels at translating

sparse-view reconstructed images into their full-view counterparts. Furthermore, reconstruction methodologies based on diffusion-based models effectively remove noise from images by incorporating the inherent probability distribution inherent in comprehensive CT images [11]. While standard image-domain diffusion-based reconstruction techniques have made efforts to incorporate constraints tied to data consistency with sparse-view sinogram, these approaches are inadequate in ensuring accurate sinogram fidelity in the presence of an unknown data distribution in full-view projection [12]. Consequently, the sinogram representation of reconstructed images utilizing established reconstruction paradigms remains misaligned with the authentic depiction of a full-view sinogram.

The underlying cause of this issue lies in the inadequate consideration of the distribution of the sinogram [13], [14] [15]. Most existing methods for reconstruction methods primarily focus on the image-domain and perform post-processing on sparse-view reconstruction results [16]. While some methods have utilized the sparse-view sinogram to constrain the reconstruction process, these methods have not leveraged the prior distributional information of the sinogram [17]. For instance, the most effective diffusion model for SVCT learns the prior distribution in the image-domain [18] with utilizing the sparse-view sinogram to constrain the generation. However, it does not factor in the prior distribution of the sinogram itself [19].

In order to overcome the challenge and reconstruct high-quality results with precise sinogram, currently, several sinogram domain diffusion-based models have been proposed [12]. Such methods achieve excellent reconstruction results by progressively sampling the full-view sinogram from noise, guided by the learned prior distribution in the sinogram domain [20]. However, these methods have certain limitations. Firstly, it should be noted that the information contained in the sinogram corresponds to the entire quality in the image-domain [21]. This implies that inaccuracies in the diffusion generation process can have a significant impact on the overall image quality. Secondly, the sinogram-domain diffusion model struggles when addressing reconstruction problems under extremely sparse views ( $\text{views} \leq 30$ ) [22]. When the views are extremely sparse, the initial sparse-view sinogram used for constraints only represents a negligible portion of the full-view sinogram. Therefore, it becomes challenging to generate accurate results for missing views. Consequently, relying solely on the sinogram-based diffusion model is insufficient to achieve optimal outcomes.

To achieve high-quality image reconstruction with accurate sinogram, several dual-domain methods have been proposed, such as DDPTransformer [23] and LEARN++ [24]. However, despite the integration of dual-domain information into these methods, there is a lack of effective mutual enhancement between the two domains. In the case of DDPTransformer, the overall reconstruction process involves both sinogram-domain reconstruction and image-domain reconstruction. However, these two stages are completely independent and sequential. The method iteratively generates full-view measurement data from undersampled measurements and treats them as input for the subsequent image-domain reconstruction stage.

Unfortunately, the iterations in DDPTransformer do not have any mutual feedback between the image-domain and sinogram-domain. On the other hand, the LEARN++ method does have some level of mutual feedback between the image-domain and sinogram-domain. However, it is important to note that the input to the sinogram-domain network in each iteration solely comes from the current image-domain. This means that there is no feedback from the sinogram of the previous iteration. As a result, the LEARN++ method is primarily driven by the image-domain model, with the sinogram-domain model serving as a post-processing correction to the generated image-domain results. This limits the true synergistic generation of dual-domain information in LEARN++.

The stability of these existing dual-domain methods is poor [25]. Firstly, this instability manifests primarily as poor generalizability. Methods like DDPTransformer and LEARN++ require retraining models when faced with reconstructions from different numbers of views or different tasks. In addition, this instability is evident in the difficulty of recovering fine structures in heavily sub-sampled data. Those methods learn the mapping from sub-sampled images to full-view images rather than generating, data that cannot be effectively represented in sub-sampled conditions cannot be accurately generated, leading to distortion of image details.

In order to achieve high-quality image reconstruction with accurate sinogram, the sole application of diffusion models in either the image-domain or the sinogram-domain does not yield optimal results [26]. Hence, a viable solution is to fully explore the advantages of data distribution within both sinogram-domain and image-domain by integrating prior information within dual-domain and enhancing its correlation. However, executing joint diffusion generation in these two domains is a formidable task due to the intricate challenge of aligning noise patterns between the sinogram and image domains. Additionally, as far as we know, an optimized model that effectively merges image-domain and sinogram-domain diffusion models is currently lacking, resulting in a lack of a synergistic modeling approach. Moreover, the presence of two distinct diffusion models within separate domains introduces complexity in achieving a balanced proportion for this dual-domain integration.

Building upon the aforementioned analysis, this study presents a novel approach known as the Dual-Domain Collaborative Diffusion Sampling (DCDS) model for multi-source stationary CT reconstruction. The fundamental concept behind DCDS is to establish a unified optimization framework that seamlessly integrates diffusion models from both the sinogram-domain and image-domain. Our approach brings forth four noteworthy contributions as follows:

- We first propose a unified optimization framework by combining sinogram-domain diffusion model with image-domain diffusion model to simultaneously consider the image prior distribution and sinogram prior distribution for multi-source stationary CT reconstruction.
- We implement a collaborative merging mechanism in order to effectively enhance the reciprocal alignment between the sinogram-domain and image-domain.

By utilizing the results generated through the image-domain diffusion model to facilitate the construction of precise sinogram, and conversely, employing the accurate sinogram to steer the generation of image-domain results, fosters a synergistic interplay between these two domains.

- We employ an innovative optimization approach for solution derivation, where we leverage a conventional iterative technique to enhance data consistency and improve overall stability. Our meticulous analysis of regularization parameters with extensive studies, ensures the practical applicability of our method.
- Our proposed DCDS model demonstrates its effectiveness in achieving high-quality image reconstruction across multi-source stationary CT scanned datasets and simulated AAPM dataset. Through rigorous comparison with state-of-the-art alternatives, including single-domain diffusion models, our DCDS model showcases its remarkable superiority in effectively addressing multi-source stationary CT reconstruction challenges.

## II. RELATED WORK

### A. Sparse-View CT Reconstruction

The core issue to be addressed in multi-source stationary CT reconstruction is the SVCT reconstruction problem. The primary challenge in SVCT reconstruction relates to solving ill-posed inverse problems. The CT imaging process can be expressed as  $\mathbf{Y} = \mathbf{A}(\mathbf{x}) + \epsilon$ , where  $\mathbf{Y}$  represents the measured sinogram,  $\mathbf{x}$  represents the measured object,  $\epsilon$  represents the random noise perturbation and  $\mathbf{A}$  represents the system matrix. [27] In the task of SVCT reconstruction, the matrix  $\mathbf{A}$  is under-determined and the sinogram  $\mathbf{Y}$  is incomplete [28]. Therefore, the goal of the reconstruction task is to solve the inverse problem  $\mathbf{x} = \mathbf{A}^{-1}(\mathbf{Y})$  using the incomplete sinogram  $\mathbf{Y}$ . SVCT reconstruction methods can be classified into two categories. The first category is image-domain reconstruction and the second category is sinogram-domain reconstruction [10]. Typical methods will be introduced as follows. Most of the methods used for SVCT reconstruction operate in the image domain. These image-domain reconstruction methods can be further classified into traditional methods and deep-learning methods. Traditional methods make use of compressed sensing (CS) by incorporating regularization penalties that are based on sparsity and other prior information for the purpose of reconstruction. In recent times, there has been significant progress in SVCT reconstruction achieved through the utilization of deep learning-based methods. These methods involve the use of convolutional neural networks (CNNs) to map sparse-view reconstructed images to high-quality images [29], generative adversarial networks (GANs) [30], and transformers [31], [32] [33].

In recent years, image-domain diffusion models have shown remarkable efficacy in SVCT reconstruction [34]. By learning the prior distribution of high-quality images, the image-domain diffusion model can directly sample full-view reconstructed images from Gaussian noise [35]. It utilizes measurement data from sparse views to constrain the generation process and achieve favorable results. Sinogram-domain methods can be

further classified into two sub-categories. The first category is the sinogram interpolation approach, which involves interpolating sparse-view sinogram to full-view sinogram using methods such as linear interpolation [36]. However, such methods have limitations, as interpolation can introduce interpolation artifacts [37]. The second category involves deep learning-based sinogram reconstruction methods, where neural networks are employed to map sparse-view sinogram to full-view sinogram [38]. The final reconstruction results are obtained through filtering back-projection (FBP). Similarly, the sinogram-domain diffusion models have also achieved remarkable results in sparse-view CT reconstruction [12]. By learning the prior distribution of full-view sinogram and utilizing sparse-view constraints, full-view sinogram can be sampled from noise with high fidelity [20].

Furthermore, there is also a portion of sparse-view imaging and reconstruction implemented based on hardware methods, among which the most representative is many-view undersampling (MVUS) [39], [40], which can provide sparse data sampling without switching the tube power. Specifically, a multi-slit collimator is placed in between a tube and a patient, and it efficiently reduces the radiation dose to the patient. [41]

### B. Diffusion Model

Diffusion models belong to the category of probabilistic generative models, and their distinctive training methodology integrates both forward diffusion and backward generation processes [8], [18]. In the forward progression, the noise of varying magnitudes is progressively introduced to the input image, leading to a gradual degradation of its fidelity until it eventually transforms into pure Gaussian noise [42]. Conversely, the backward generation process follows an iterative approach aimed at noise reduction, effectively generating the image from random Gaussian white noise [43]. During each iteration, a neural network, such as U-Net, is used to estimate the current noise levels. In the context of SVCT reconstruction, addressing the inverse problem involves the intricate task of reconstructing the underlying probability density distribution of the resultant images or sinogram, denoted as  $p(\mathbf{x})$  [43]. Noise Conditioned Score Networks (NCSNs) is a classical diffusion model. The score function associated with the data probability density  $p(\mathbf{x})$  is defined as the logarithmic gradient of this density  $\nabla_{\mathbf{x}} \log p(\mathbf{x})$  [44]. To guide the initial random probability density  $p(\mathbf{x}_0)$  towards regions of increased probability density values, the Langevin annealing algorithm is harnessed [8]. The iterative update equation is given as follows:

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \frac{\rho}{2} \nabla_{\mathbf{x}_{t-1}} \log p(\mathbf{x}_{t-1}) + \sqrt{\rho} \cdot \omega_t, \quad (1)$$

where  $t \in 1, \dots, T$ ,  $\rho$  governs the extent of the update aligned with the score direction,  $\mathbf{x}_0$  is drawn from a prior distribution, and  $\omega_t \sim \mathcal{N}(0, \mathbf{I})$  denotes a stochastic perturbation. As a result, the diffusion model  $S_{\theta}(\mathbf{x}) \approx \nabla_{\mathbf{x}} \log p(\mathbf{x})$  can be subjected to training to forecast the gradients of the probability density gradient, where  $\theta$  is the trained parameter of networks. The formulation of the training loss function is as follows:

$$\mathcal{L}_{\text{sm}} = \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x})} \|S_{\theta}(\mathbf{x}) - \nabla_{\mathbf{x}} \log p(\mathbf{x})\|_2^2. \quad (2)$$

### III. METHODOLOGY

In this section, we commence by introducing our proposed dual-domain collaborative diffusion sampling (DCDS) model that leverages the integration of dual-domain data distribution priors to address the multi-source stationary SVCT reconstruction problem. Subsequently, we progressively describe the implementation of the optimization process implementation.

#### A. DCDS Model

The first step in developing the dual-domain diffusion model to address multi-source stationary SVCT reconstruction is to formulate a collaborative optimization objective function that integrates information from both sinogram and image domains. To enhance clarity, we begin by illustrating the variables that may be subject to confusion during the optimization procedure.  $\mathbf{x}$  represents the image in the optimization process, the transformed sinogram data is  $\mathbf{y}$ , and  $\mathcal{M}$  is the zero-one mask matrix transforming the full-view sinogram to the sparse-view counterpart.  $\mathbf{A}_1$  and  $\mathbf{A}_2$  denote the sparse-view and full-view projection system matrix respectively. Specifically,  $\mathbf{A}_1$  refers to sparse-view projection, where image-domain data is projected onto views corresponding to sparse-view scans, generating a sparse sinogram. On the other hand,  $\mathbf{A}_2$  represents full-view projection, where image-domain data is projected onto 580 views as a full-view sinogram. Next, we present the diffusion model-based reconstruction methods for solving the inverse problem of SVCT using a single image-domain diffusion model or a sinogram-domain diffusion model separately. Subsequently, we establish an optimization collaborative mechanism for simultaneously incorporating dual-domain priors.

*1) Image-Domain Diffusion Model:* Solving the reconstruction of the SVCT using the diffusion model in the image domain can be mathematically represented as  $\mathbf{x} = \mathbf{A}_1^{-1}(\mathbf{Y})$ . Given the ill-posed nature of this problem, it is typically approached through the iterative solution of an optimization equation that incorporates a regularization term:

$$\mathbf{x}^{(*)} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{Y} - \mathbf{A}_1 \mathbf{x}\|_2^2 + \frac{\eta}{2} R_1(\mathbf{x}), \quad (3)$$

where  $R_1(\cdot)$  represents an appropriate regularization term and  $\eta$  is a factor. In the process of utilizing the image-domain diffusion model to solve the inverse problem, by utilizing the Split-Bregman strategy, we employ the model-based iterative result  $\mathbf{u}$  of the image-domain diffusion model to learn the prior information for this term. Specifically, the value of  $\mathbf{u}$  is determined by  $\mathbf{x}$  during the iterative solving process, which includes the values of  $\mathbf{x}$  obtained during the solving process, based on model estimates, and random drift. This prior information guides the optimization process:

$$\mathbf{x}^{(*)} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{Y} - \mathbf{A}_1 \mathbf{x}\|_2^2 + \frac{\eta_1}{2} \|\mathbf{x} - \mathbf{u}\|_2^2. \quad (4)$$

*2) Sinogram-Domain Diffusion Model:* Solving the inverse problem of SVCT reconstruction in the sinogram-domain can be represented as  $\mathbf{y} = \mathcal{M}^{-1}(\mathbf{Y})$ , where  $\mathcal{M}^{-1}$  denotes the operation of recovering a sparse-view sinogram to its full-view counterpart. Similar to the image-domain, the

sinogram-domain reconstruction process can also be optimized using a regularized iterative method:

$$\mathbf{y}^{(*)} = \operatorname{argmin}_{\mathbf{y}} \frac{1}{2} \|\mathbf{Y} - \mathcal{M} \mathbf{y}\|_2^2 + \frac{\psi}{2} R_2(\mathbf{y}), \quad (5)$$

where  $R_2(\cdot)$  represents an appropriate regularization term and  $\psi$  is a factor. In the process of utilizing the sinogram-domain diffusion model to address the inverse problem of SVCT reconstruction, by utilizing the split-Bregman strategy, we employ the model-based iterative result  $\mathbf{v}$  of the sinogram-domain diffusion model to learn prior information about the sinogram. Specifically, the value of  $\mathbf{v}$  is determined by  $\mathbf{y}$  during the iterative solving process, which includes the values of  $\mathbf{y}$  obtained during the solving process, based on model estimates, and random drift. This information is then utilized as a regularization term to restrict the optimization process:

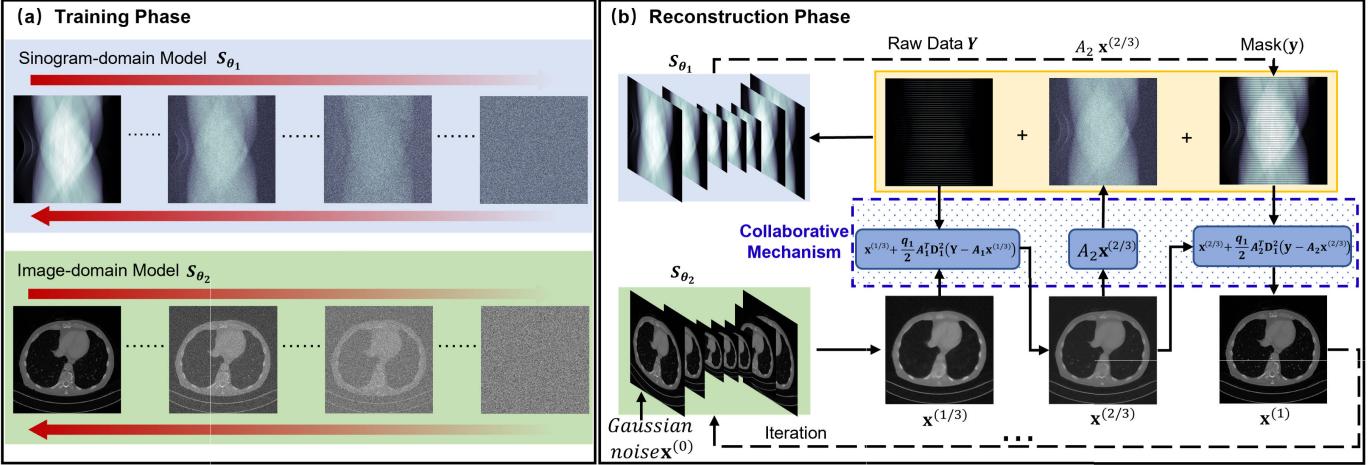
$$\mathbf{y}^{(*)} = \operatorname{argmin}_{\mathbf{y}} \frac{1}{2} \|\mathbf{Y} - \mathcal{M} \mathbf{y}\|_2^2 + \frac{\psi_1}{2} \|\mathbf{y} - \mathbf{v}\|_2^2. \quad (6)$$

*3) Dual-Domain Collaborative Diffusion Sampling Model:* The central aspect of the DCDS model is the creation of a correspondence bridge that facilitates seamless interaction between the reconstructed prior information across the dual domains. In essence, it becomes imperative to leverage the intermediate solutions derived from the image-domain diffusion model to govern the iterative process of generating the sinogram. Conversely, the intermediate solutions obtained from the sinogram impose constraints on the image-domain generation process. Combining the above analyses and Eqs. 3, 5, the optimization function of DCDS can be formulated as follows:

$$\{\mathbf{x}^{(*)}, \mathbf{y}^{(*)}\} = \operatorname{argmin}_{\mathbf{x}, \mathbf{y}} \frac{1}{2} \|\mathbf{Y} - \mathbf{A}_1 \mathbf{x}\|_2^2 + \frac{\lambda_1}{2} \|\mathbf{Y} - \mathcal{M} \mathbf{y}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{y} - \mathbf{A}_2 \mathbf{x}\|_2^2 + \frac{\eta}{2} R_1(\mathbf{x}) + \frac{\psi}{2} R_2(\mathbf{y}), \quad (7)$$

where  $\|\mathbf{y} - \mathbf{A}_2 \mathbf{x}\|_2^2$  plays a crucial role in connecting the image-domain and sinogram-domain throughout the optimization process, with  $\lambda_2 > 0$  as a modulating factor. Considering the practical significance, in the optimization process, the  $\|\mathbf{y} - \mathbf{A}_2 \mathbf{x}\|_2^2$  represents that the generated intermediate result of the reconstruction image  $\mathbf{x}$  should maintain a correspondence with the iteratively generated sinogram  $\mathbf{y}$ , implying that the generated  $\mathbf{y}$  is the full-view projection of  $\mathbf{x}$ .  $\lambda_1$  is a modulating factor to maintain the consistency of generated sinogram  $\mathbf{y}$  with measured sinogram data  $\mathbf{Y}$ . By incorporating this component, we promote the convergence of solutions between the image and sinogram domains.

To solve the optimization Eq. 7 mentioned above, the image-domain diffusion model and the sinogram-domain diffusion model need to be constructed as regularization terms. Furthermore, the outcomes obtained from the image-domain diffusion model should be subjected to mutual constraints imposed by the sinogram-domain diffusion model. The following parts present a comprehensive explanation of the optimization process for DCDS.



**Fig. 1.** The well-defined pipeline of our DCDS method consists of two phases. In the training phase (a), we train the sinogram-domain and image-domain diffusion models. Phase (b) represents the testing phase, featuring a collaborative mechanism.

### Algorithm 1 DCDS for SVCT Reconstruction

**Require:**  $\mathbf{x}, \mathbf{y}, \omega_t, S_{\theta_1}, S_{\theta_2}, \rho, \mu_1, \mu_2, T, \mathbf{A}_1, \mathbf{A}_2$

- 1: **for**  $t = 0$  to  $T$  **do**
- 2:   Draw  $\omega_t \sim N(0, 1)$
- 3:    $\mathbf{x}^{(t+\frac{1}{3})} = \mathbf{x}^{(t)} + \frac{\rho}{2} S_{\theta_1}(\mathbf{x}^{(t)}, t) + \sqrt{\rho} \cdot \omega_t$
- 4:    $\mathbf{x}^{(t+\frac{2}{3})} = \mathbf{x}^{(t+\frac{1}{3})} + \frac{q_1}{2} \mathbf{A}_1^T D_1^2 (\mathbf{Y} - \mathbf{A}_1 \mathbf{x}^{(t+\frac{1}{3})})$
- 5:    $\mathbf{y}^{(t+\frac{1}{2})} = \mathbf{y}^t + \frac{\rho}{2} S_{\theta_2}(\mathbf{y}^{(t)}, t) + \sqrt{\rho} \cdot \omega_t$
- 6:    $\mu_1(t) = 1 - t/T$
- 7:    $\mu_2(t) = t/T$
- 8:    $\mathbf{y}^{(t+1)} = \mathcal{M}\mathbf{Y} + (1 - \mathcal{M})[\mu_1(t)\mathbf{A}_2\mathbf{x}^{(t+\frac{2}{3})} + \mu_2(t)\mathbf{y}^{(t+\frac{1}{2})}]$
- 9:    $\mathbf{x}^{(t+1)} = \mathbf{x}^{(t+\frac{2}{3})} + \frac{q_2}{2} \mathbf{A}_2^T D_2^2 (\mathbf{y}^{(t+1)} - \mathbf{A}_2 \mathbf{x}^{(t+\frac{2}{3})})$
- 10: **end for**

### B. DCDS-Solving

**1) Training Phase:** The essence of developing a sinogram-domain and image-domain collaborative diffusion model for addressing the SVCT reconstruction problem is the reciprocal enhancement of the sinogram-domain and the image-domain reconstruction models. As shown in Fig. 1(a), DCDS accomplishes this through the establishment of both sinogram-domain and image-domain diffusion models. During the training phase, this process is achieved by training two scoring functions that are dedicated to learning the data distributions in both the image-domain and the sinogram-domain. The trained image-domain and sinogram-domain models are represented by  $S_{\theta_1}(\cdot)$  and  $S_{\theta_2}(\cdot)$ , where  $\theta_1$  and  $\theta_2$  are trained parameters of the image-domain model and sinogram-domain model, respectively. Similar to common diffusion models, both our sinogram-domain diffusion model and image-domain diffusion model utilize the U-Net network architecture. Specifically, the network architecture we employed is based on the classic diffusion model network NCSN [18], which is mentioned above in related works.

**2) Reconstruction Phase:** The reconstruction process involves the collaborative sampling of the two constructed diffusion models to solve Eq. 7. The sampling process of the diffusion models in the image domain and sinogram domain

can be described as follows:

$$\mathbf{x}^{(t)} = \mathbf{x}^{(t-1)} + \frac{\rho}{2} S_{\theta_1}(\mathbf{x}^{(t-1)}, t) + \sqrt{\rho} \cdot \omega_t, \quad (8)$$

$$\mathbf{y}^{(t)} = \mathbf{y}^{(t-1)} + \frac{\rho}{2} S_{\theta_2}(\mathbf{y}^{(t-1)}, t) + \sqrt{\rho} \cdot \omega_t, \quad (9)$$

where  $t \in 1, \dots, T$ ,  $\rho$  governs the extent of the update aligned with the score direction,  $S_{\theta_1}(\cdot)$  and  $S_{\theta_2}(\cdot)$  are trained models, and  $\omega_t \sim \mathcal{N}(0, \mathbf{I})$  denotes a stochastic perturbation. To solve Eq. 7, we perform the split-Bregman strategy and embed the Langevin annealed sampling into Eq. 7, resulting in:

$$\begin{aligned} & \{\mathbf{x}^{(t+1)}, \mathbf{y}^{(t+1)}\} \\ &= \operatorname{argmin}_{\mathbf{x}, \mathbf{y}} \frac{1}{2} \|\mathbf{Y} - \mathbf{A}_1 \mathbf{x}\|_2^2 \\ & \quad + \frac{\lambda_1}{2} \|\mathbf{Y} - \mathcal{M}\mathbf{y}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{y} - \mathbf{A}_2 \mathbf{x}\|_2^2 \\ & \quad + \frac{\eta_1}{2} \|\mathbf{x} - \mathbf{u}\|_2^2 + \frac{\psi_1}{2} \|\mathbf{y} - \mathbf{v}\|_2^2 \\ & \text{s.t. } \mathbf{u} = \mathbf{x}^{(t)} + \frac{\rho}{2} S_{\theta_1}(\mathbf{x}^{(t)}, t) + \sqrt{\rho} \cdot \omega_t \\ & \text{s.t. } \mathbf{v} = \mathbf{y}^{(t)} + \frac{\rho}{2} S_{\theta_2}(\mathbf{y}^{(t)}, t) + \sqrt{\rho} \cdot \omega_t. \end{aligned} \quad (10)$$

$\mathbf{u}$  and  $\mathbf{v}$  are iterative sampling results by Langevin annealed sampling. In order to maintain consistency among variables in the subsequent text, we use  $\mathbf{x}^{(t+\frac{1}{3})}$  to replace  $\mathbf{u}$  and  $\mathbf{y}^{(t+\frac{1}{2})}$  to replace  $\mathbf{v}$ . Following the above replacement, the Eq. 10 is transformed as follows:

$$\begin{aligned} & \{\mathbf{x}^{(t+1)}, \mathbf{y}^{(t+1)}\} \\ &= \operatorname{argmin}_{\mathbf{x}, \mathbf{y}} \frac{1}{2} \|\mathbf{Y} - \mathbf{A}_1 \mathbf{x}\|_2^2 \\ & \quad + \frac{\lambda_1}{2} \|\mathbf{Y} - \mathcal{M}\mathbf{y}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{y} - \mathbf{A}_2 \mathbf{x}\|_2^2 \\ & \quad + \frac{\eta_1}{2} \|\mathbf{x} - \mathbf{x}^{(t+\frac{1}{3})}\|_2^2 + \frac{\psi_1}{2} \|\mathbf{y} - \mathbf{y}^{(t+\frac{1}{2})}\|_2^2 \\ & \text{s.t. } \mathbf{x}^{(t+\frac{1}{3})} = \mathbf{x}^{(t)} + \frac{\rho}{2} S_{\theta_1}(\mathbf{x}^{(t)}, t) + \sqrt{\rho} \cdot \omega_t \\ & \text{s.t. } \mathbf{y}^{(t+\frac{1}{2})} = \mathbf{y}^{(t)} + \frac{\rho}{2} S_{\theta_2}(\mathbf{y}^{(t)}, t) + \sqrt{\rho} \cdot \omega_t. \end{aligned} \quad (11)$$

For the conditional optimization Eq. 11, we proceed to solve it step by step with initial Gaussian noise  $\mathbf{x}^{(0)}$  and  $\mathbf{y}^{(0)}$ . First,

we initiate optimization for the image  $\mathbf{x}$ . By decomposing the terms related to  $\mathbf{x}$  in the above equation, we can iteratively proceed to solve for  $\mathbf{x}$  as follows:

$$\mathbf{x}^{(t+\frac{1}{3})} = \mathbf{x}^{(t)} + \frac{\rho}{2} S_{\theta_1}(\mathbf{x}^{(t)}, t) + \sqrt{\rho} \cdot \omega_t, \quad (12)$$

$$\mathbf{x}^{(t+\frac{2}{3})} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{Y} - \mathbf{A}_1 \mathbf{x}\|_2^2 + \frac{\eta_1}{2} \|\mathbf{x} - \mathbf{x}^{(t+\frac{1}{3})}\|_2^2. \quad (13)$$

Eq. 12 illustrates the image update procedure utilizing the image-domain diffusion model. Eq. 13 serves as the data consistency term constraint. We employ the simultaneous iterative reconstruction technique to solve Eq. 13, ensuring that the intermediate result remains consistent with the initial sparse-view projection  $\mathbf{Y}$ . The iteration process is outlined as follows:

$$\mathbf{x}^{(t+\frac{2}{3})} = \mathbf{x}^{(t+\frac{1}{3})} + \frac{q_1}{2} \mathbf{A}_1^T D_1^2 (\mathbf{Y} - \mathbf{A}_1 \mathbf{x}^{(t+\frac{1}{3})}), \quad (14)$$

where  $D_1 = \operatorname{diag}\{1/\|a_1\|, 1/\|a_2\|, \dots, 1/\|a_{q_1}\|\}$  represents a diagonal matrix and  $a_{q_1}$  represents the  $q_1$ -th row vector of matrix  $\mathbf{A}_1$ . It should be noted that the updating process for the image is not complete, as  $\mathbf{x}^{(t+\frac{2}{3})}$  needs to be constrained using obtained sinogram  $\mathbf{y}$  during the iterative process. Therefore, the next step is to optimize for  $\mathbf{y}$ .

Similar to the optimization process for  $\mathbf{x}$ , we focus on the terms related to  $\mathbf{y}$  in the Eq. 10. The optimization process for  $\mathbf{y}$  can be iteratively expressed as follows:

$$\mathbf{y}^{(t+\frac{1}{2})} = \mathbf{y}^{(t)} + \frac{\rho}{2} S_{\theta_2}(\mathbf{y}^{(t)}, t) + \sqrt{\rho} \cdot \omega_t, \quad (15)$$

$$\begin{aligned} \mathbf{y}^{(t+1)} &= \operatorname{argmin}_{\mathbf{y}} \frac{1}{2} \|\mathbf{Y} - \mathcal{M} \mathbf{y}\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{y} - \mathbf{A}_2 \mathbf{x}^{(t+\frac{2}{3})}\|_2^2 \\ &\quad + \frac{\psi_1}{2} \|\mathbf{y} - \mathbf{y}^{(t+\frac{1}{2})}\|_2^2. \end{aligned} \quad (16)$$

Eq. 15 represents the utilization of the sinogram-domain diffusion model for sinogram generation, while Eq. 16 signifies the data consistency constraint. For the data consistency optimization,  $\mathcal{M}$  is a sparse matrix that retains only the original measurement views of the sinogram.

$$\mathbf{y}^{(t+1)} = \mathcal{M} \mathbf{Y} + (1 - \mathcal{M})[(\mu_1(t) \mathbf{A}_2 \mathbf{x}^{(t+\frac{2}{3})} + \mu_2(t) \mathbf{y}^{(t+\frac{1}{2})})]. \quad (17)$$

As represented in Eq. 17, we solve Eq. 16 by performing a weighted combination of the original measurement data  $\mathbf{Y}$ , the projection of the intermediate image  $\mathbf{A}_2 \mathbf{x}^{(t+\frac{2}{3})}$ , and the intermediate sinogram solution  $\mathbf{y}^{(t+\frac{1}{2})}$ . Remarkably, the parameter setting for this weighting process involves certain nuances. During the initial iterations, a significant part of the model-generated sinogram may be inaccurate. At this stage, the constraint from the intermediate image-domain solution  $\mathbf{x}^{(t+\frac{2}{3})}$  should dominate. With  $t$  increases, the generated results from the model become more accurate. Consequently, the weight contribution from the model  $\mu_2(t)$  increases with  $t$  and  $\mu_1(t)$  should decreases with  $t$ . Besides, To maintain the overall numerical range of the sinogram, the sum of  $\mu_1(t)$  and  $\mu_2(t)$  should be equal to 1. Consequently, we have devised the following expressions for  $\mu_1(t)$  and  $\mu_2(t)$ :

$$\mu_1(t) = 1 - t/T, \quad (18)$$

$$\mu_2(t) = t/T, \quad (19)$$

where  $T$  is the total iteration number. Finally, by utilizing the solution  $\mathbf{y}^{(t)}$  obtained from the sinogram-domain diffusion model after the data consistency constraint, we can acquire the ultimate solution within a singular iteration loop of the image-domain diffusion model. The optimization equation can be expressed as follows:

$$\mathbf{x}^{(t+1)} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{y}^{(t+1)} - \mathbf{A}_2 \mathbf{x}\|_2^2 + \frac{\eta_1}{2} \|\mathbf{x} - \mathbf{x}^{(t+\frac{2}{3})}\|_2^2. \quad (20)$$

For Eq. 20, we similarly employ the conventional iterative method for solving, and the iteration process is given as follows:

$$\mathbf{x}^{(t+1)} = \mathbf{x}^{(t+\frac{2}{3})} + \frac{q_2}{2} \mathbf{A}_2^T D_2^2 (\mathbf{y}^{(t+1)} - \mathbf{A}_2 \mathbf{x}^{(t+\frac{2}{3})}), \quad (21)$$

where  $D_2 = \operatorname{diag}\{1/\|b_1\|, 1/\|b_2\|, \dots, 1/\|b_{q_2}\|\}$  represents a diagonal matrix,  $b_{q_2}$  represents the  $q_2$ -th row vector of matrix  $\mathbf{A}_2$ .

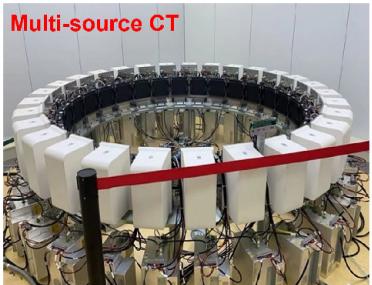
In summary, the reconstruction phase of DCDS is illustrated in Fig. 1(b). The dual-domain collaborative reconstruction process starts with Gaussian noise  $\mathbf{x}^{(0)}$  of size  $512 \times 512$ , which undergoes a pre-trained image-domain diffusion model to obtain an intermediate result  $\mathbf{x}^{(1/3)}$ . Subsequently, the original measurement data  $\mathbf{Y}$  of size  $580 \times 580$  (only the data at the measured viewpoint exists; the rest of the rows are zero) is used to constrain and produce  $\mathbf{x}^{(2/3)}$ . We then perform full-view projection on the intermediate result  $\mathbf{x}^{(2/3)}$  to acquire the mid-sinogram solution. By combining the output  $\mathbf{y}$  of size  $580 \times 580$  of the previous sinogram-domain diffusion model with this intermediate solution as input, the sinogram-domain score-based model updates the sinogram. Finally, the updated sinogram is used to constrain  $\mathbf{x}^{(2/3)}$  for obtaining the ultimate result. The reconstruction algorithm is outlined in Algorithm 1.

#### IV. EXPERIMENTAL SETUP AND RESULTS

In this section, we rigorously validate our approach through a series of experiments and present the findings of our study. We first introduce the utilized datasets. Subsequently, we offer a detailed exposition of the essential parameters and the experimental configuration. Finally, we perform a comparative evaluation of our methodology against the existing state-of-the-art techniques in the field.

##### A. Datasets and Implementation Details

**1) Datasets:** We evaluate the effectiveness of DCDS in multi-source stationary CT reconstruction using three separate datasets. These datasets include numerical simulations of abdominal scenarios, data collected from two physical phantoms using a real multi-source stationary CT system, and real-world clinical cardiac data. The first dataset consists of Mayo Clinic abdomen images from the AAPM Low Dose CT Grand Challenge by McCollough et. al. [45], which encompasses normal-dose CT scans obtained from 10 patients, with 9 patients allocated for training purposes and one patient reserved for evaluation. The training dataset encompasses 4608 slices, each having a thickness of 1.0 mm, and encompassing a pixel grid of  $512 \times 512$ . In the testing phase,



**Fig. 2.** The multi-source CT configuration featuring 29 pairs of X-ray sources and flat panel detectors [3].

we employ fan-beam scanning geometry, with the source-to-rotation center and detector-to-rotation center distances set as  $50\text{cm}$  and  $50\text{cm}$  respectively. The equi-distance detector consists of 580 pixels and each of them is  $0.35\text{mm}$ . For the testing stage, we generate a subset data of 29 views, evenly distributed over the full 360-degree. Using this dataset is to verify the effectiveness of the DCDS algorithm we proposed, similarly adopting 29 simulated imaging views for the feasibility study on future utilization of multi-source stationary CT for human body scanning.

Additionally, to validate the reconstruction ability for real multi-source stationary CT system, we used multi-source stationary CT scanning (as shown in Fig. 2) with 29 views on a bucket filled with  $3\text{mm}$  diameter spheres and applied DCDS to achieve high-quality reconstruction. We scan ten buckets filled with spheres, and each reconstructed slice has a size of  $512 \times 512$ , consisting of 5120 slices. To get full-view training datasets, we supplement the views by rotating the turntable and reconstructing ground truth by utilizing the FDK method with full-view sinogram [46].

Furthermore, To verify the feasibility of using 29 views for high-quality imaging of the cardiac, and to prepare for future real-time imaging of the cardiac using multi-source stationary CT, we employ the actual cardiac clinical datasets as presented in [47]. The curved cylindrical detector array consists of a total of 880 individual units, and the complete scan consists of 2200 views. Specifically, for the purpose of our study, we selected 29 views as the testing dataset. The selection process involved an equal spacing sampling method under an angle greater than  $180^\circ$  from the 2200 views. The field-of-view (FOV) diameter covered a wide expanse measuring  $49.8\text{ cm}^2$ , with image dimensions of  $512 \times 512$  pixels. Notably, the distances from the X-ray source to the center and detector are set at  $53.85\text{ cm}$  and  $103.68\text{ cm}$ , respectively.

Lastly, we must emphasize that the method we designed does not require retraining the model for adjustments to scanning parameters. The reason is that the input size for the sinogram-domain model is fixed at a relatively small size of  $580 \times 580$ . This dimension implies that we are considering a full-view sinogram as the result obtained from imaging systems equipped with 580 detectors capturing 580 views. This number of detectors and views is relatively small compared to typical imaging systems, suggesting that for the 580 dimensions in terms of detector count, downsampling from more complex imaging systems is feasible. For the

580 views, image-domain forward projection is employed during the collaborative process for padding, allowing common imaging systems (with more than 580 detectors) to utilize the sinogram-domain model without the need for retraining.

**2) Model Configurations and Comparisons Selection:** During the process of model training, we utilize the Adam algorithm with a learning rate set to  $10^{-3}$ . The implementation of our method is carried out in Python, leveraging both the ASTRA toolbox and PyTorch [48]. It should be noted that although we trained two models, both models are not simultaneously invoked during the reconstruction process. This design allows all of our experiments to be conducted on a single RTX 3090, ensuring efficient computational performance.

Throughout the training phase, images and sinograms are augmented with random noise corresponding to the twelve levels, they are input into the network. The training process iterates for 100,000 cycles to ensure robust learning. During the reconstruction phase, images and sinograms are generated at all noise levels, with 100-150 iterations applied for each level. Within each iteration, the simultaneous iteration reconstruction technique is employed to enforce data consistency between the sinogram-domain and image-domain models.

For a quantitative assessment of algorithm performance, two established evaluation metrics including the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used. For comparison, several comparative methods for analysis including the FBP algorithm [46], the CNN-based FBPCNN [49], the GAN-based C-GAN [50], GMSD (the sinogram-domain diffusion model) [12], the DPS [43], dual-domain methods DREAM-Net [51], DDPTransformer [23] and LEARN++ [24]. Among these, CGAN and GMSD are methods based on the sinogram domain, which can test the ability of DCDS in full-view sinogram recovery. FBPCNN is a representative method for supervised sparse-view CT reconstruction, and DPS is a state-of-the-art image-domain diffusion model method.

## B. SVCT Reconstruction Results

**1) AAPM Dataset Reconstruction:** As shown in Fig. 4, the FBP results exhibit evident streaky sparse-view artifacts with low-quality reconstruction results. C-GAN aims to recover full-view sinogram from sparse-view sinogram, the reconstruction results suffer from false-positive generation due to the severe instability with 29 views data. Meanwhile, FBPCNN as a classic supervised deep-learning method, effectively suppresses streaking artifacts. However, it fails to reconstruct finer image structures and features, as indicated by the extracted regions of interest (ROIs) in Cases I and II. The representative methods of the dual-domain model, DREAM-Net, DDPTransformer and the LEARN++ method, show a significant improvement in image quality after reconstruction compared to the classic FBPCNN method, with details in the images being observable. However, due to the 29 views reconstruction task we studied being too sparse, the details in the results reconstructed by these two methods are not accurate. The GMSD method is a sinogram-domain diffusion reconstruction approach, which samples the full-view sinogram from Gaussian noise. It can be observed that the GMSD

**TABLE I**  
COMPUTATIONAL COST & QUANTITATIVE EVALUATION RESULTS

Method	Computational Cost		PSNR(dB) and SSIM		
	Model Size	FLOPs	AAPM Dataset	Phantoms Dataset	Clinical Dataset
FBP [46]	-	-	23.04/0.388	16.03/0.362	21.77/0.224
C-GAN [50]	37.3M	274.9G	26.03/0.798	18.32/0.556	23.14/0.613
FBPConvNet [49]	34.5M	262.1G	32.69/0.864	21.02/0.673	25.16/0.689
LEARN++ [24]	1.6M	589.5G	32.97/0.893	21.53/0.703	26.32/0.716
DDPTTransformer [23]	0.42M	8.4G	32.99/0.896	22.13/0.721	26.56/0.720
DREAM-Net [51]	9.9M	479.3G	33.01/0.905	22.83/0.736	26.88/0.723
GMSD [12]	7.4M	629.1G	35/98/0.921	24.54/0.791	28.34/0.741
DPS [43]	9.3M	832.6G	38.17/0.920	27.12/0.822	32.87/0.853
<b>DCDS</b>	15.2M	1263.8G	<b>39.60/0.931</b>	<b>29.82/0.843</b>	<b>34.32/0.883</b>

results exhibit some benefit details, and the overall image quality is relatively good. Because the GMSD does not utilize image domain prior information, the results contain structural blurriness and artifacts in the generated details, marked by red arrows in extracted ROIs. The images reconstructed by DPS exhibit good performance. However, due to it ignores the prior distribution in the sinogram domain, its detail generation ability is still limited, which can be further observed in the red arrows of ROIs. Leveraging the priors distributions within dual-domain, our proposed DCDS exhibits exceptional reconstruction capability, achieving the best performance in reconstructing small image features and details. The extracted ROIs especially the structures indicated by red arrows further confirm the out-performance of our method.

To evaluate the accuracy of sinogram recovery using different methods, as depicted in Fig. 4, we compare the reconstructed results with their corresponding sinogram. This study focuses on examining the disparities between the sinogram generated from the reconstructed results and the full-view sinogram. FBP, C-GAN, and FBPCovNet results exhibit notable inaccuracies in their respective sinogram, which are evident in the difference in sinograms. Even the C-GAN method, which emphasizes sinogram-domain generation, displays noticeable artifacts in its generated sinogram. When observing the sinogram generated by GMSD and DPS, we can observe that the sinogram produced by GMSD is closer to the true full-view sinogram compared to that obtained with DPS. This is attributed to the GMSD method being rooted in sinogram-domain data distribution, while DPS relies on image-domain data distribution. In contrast, the generated sinogram of our method is the closest to the original full-view sinogram.

Furthermore, we conduct a quantitative analysis, and the outcomes are listed in Table I. As seen, our DCDS approach always obtains the best quantitative results than other competitors in terms of PSNR and SSIM. This observation substantiates the superior proficiency of our method in SVCT reconstruction when compared to alternative approaches. In addition, for the purpose of comparing the stability of our method, we conducted a histogram analysis of the results generated by all the comparative methods. As shown in Fig. 3, it can be observed that our proposed DCDS method not only outperforms all current baseline methods in terms of mean values but also exhibits smaller variance compared to advanced diffusion model methods such as DPS. This indicates that our proposed DCDS method is more stable. Besides, to further

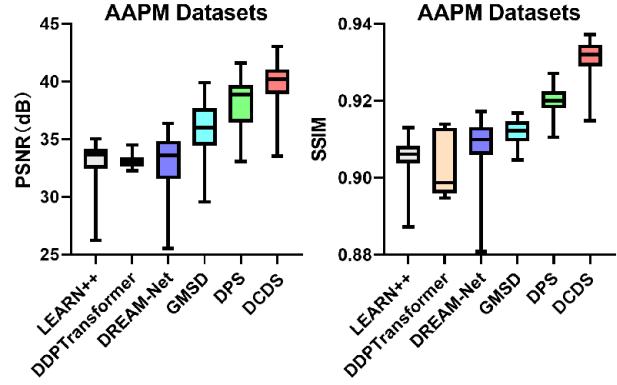
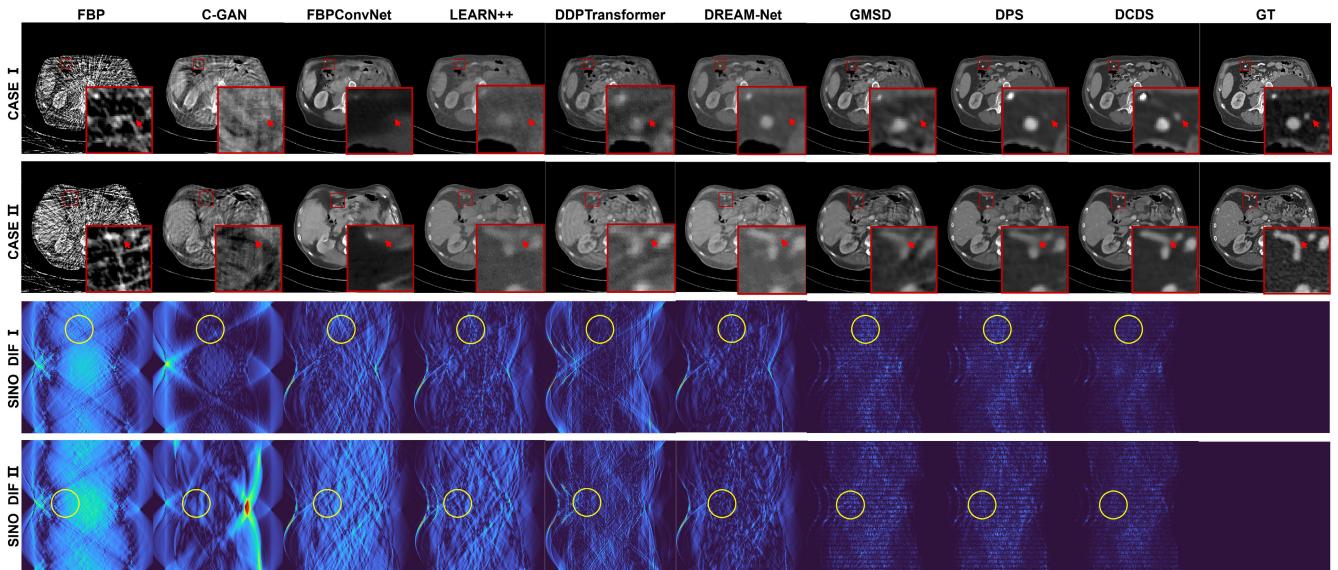


Fig. 3. Statistical analysis in the AAPM dataset.

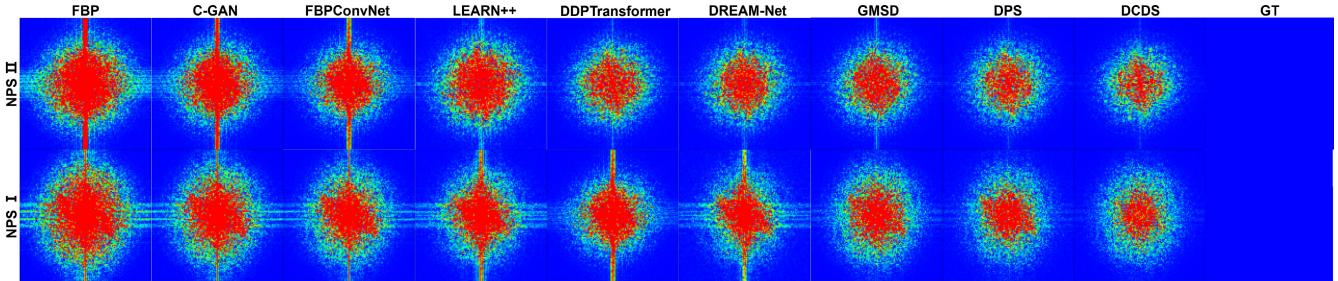
validate the superiority of DCDS, we plot the intensity of the profile in Fig. 6. As observed, the profile reconstructed by our DCDS is closest to the ground truth. Finally, the results of the noise power spectrum (NPS) [52] in Fig. 5. A smaller NPS indicates that the image is closer to the ground truth with the best performance.

Finally, in order to further compare the performance of various models, we conducted comprehensive tests on the computational resource costs of the models. Specifically, as shown in Tab. I, we calculated the model size and Flops for each deep learning method. Methods based on diffusion models, such as GMSD, DPS, and DCDS, did not exceed the model size of previous deep learning methods like FBPCovNet and C-GAN. This is because diffusion models involve the recurrent use of the same network for solving. Flops provide a better representation of the actual computational workload of the network, and our proposed DCDS method and the current state-of-the-art DPS method achieved the highest levels in floating-point operations. It is worth noting that, although DCDS has higher Flops, this does not impact the GPU memory usage, as the two pre-trained diffusion models are based on the same architecture and can share a single model with different weights.

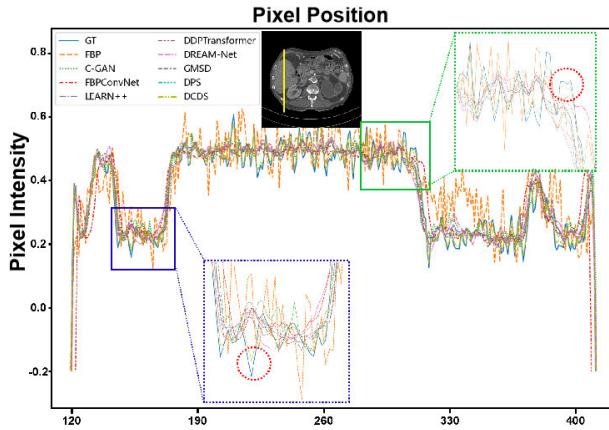
**2) Real-World Multi-Source Stationary CT Reconstruction:** To substantiate the practical applicability of our proposed method, we conducted experiments using real-world data acquired from two physical phantoms. The data were obtained through a multi-source CT scan. As shown in Fig. 2, the multi-source system is equipped with 29 pairs of X-ray sources and detectors, each securely mounted on a circular track. Each



**Fig. 4.** Representative reconstructed results with two cases from 29 views in 1<sup>st</sup> and 2<sup>nd</sup> rows, where the display window is [680,1650] HU. The corresponding sinogram of the reconstructed results from two cases are displayed in 3<sup>rd</sup> and 4<sup>th</sup> rows respectively.



**Fig. 5.** The noise power spectra (NPS) of two cases in Fig. 4. The central area corresponds to low-frequency noise components, whereas the outer region represents high-frequency noise components.

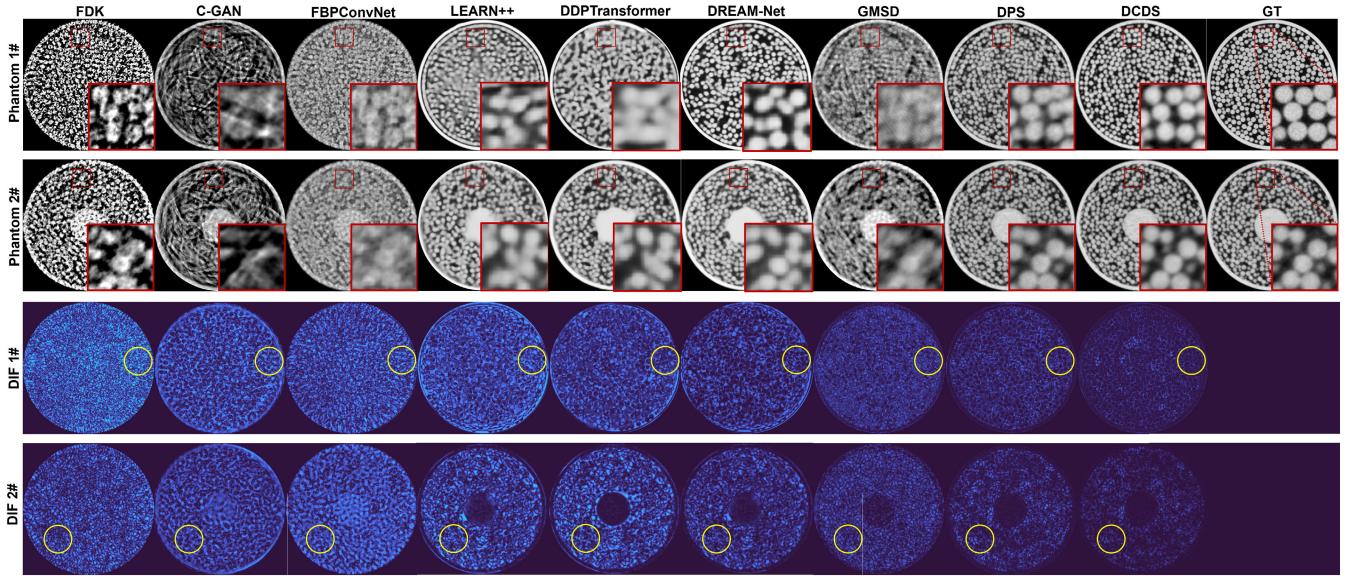


**Fig. 6.** The representative profile of Case II from simulation datasets with different methods. The pixel position is from 120 to 400.

pair features a 5kW monoblock x-ray source coupled with an indium gallium zinc oxide (IGZO) flat panel detector, covering an imaging area of  $153.6 \times 153.6\text{mm}^2$ . Specifically, the multi-source CT system adopts 29 pairs of X-ray source-detector due to limitations in its physical equipment.

Placing the sources too closely together could result in a smaller imaging area while placing them too sparsely could lead to scattering effects from adjacent X-ray beams on the detector. Ultimately, through continuous adjustments, this multi-source cone-beam CT system chose the configuration of 29 pairs of X-ray source-detector. The training datasets are obtained by rotating the gantry to supplement views. Specifically, to acquire needed training datasets, we positioned a turntable beneath the scanned object and exposed the object to 29 continuous X-ray sources, slowly rotating at an extremely slow pace to capture supplemental measurement views of the scanned object. Rearranging these angles enabled us to obtain full-view projections, resulting in the reconstruction of high-quality labels. The test data consists of a sparse-view sinogram acquired from 29 X-ray sources with only one scan. As shown in Fig. 7, we conduct multi-source CT scans on two experimental setups: phantom 1# is a cylinder filled with uniformly sized 3mm spheres, and phantom 2# is a cylinder filled with uniformly sized 3mm spheres and a central cylinder.

As depicted in Fig. 7, the reconstruction quality of the spheres obtained through the FBP is extremely poor, with blurred boundaries and indistinguishable edges. In the case of C-GAN, generation in the sinogram domain results in blurred outcomes due to the presence of numerous image



**Fig. 7.** Reconstructed results of a 29-source CT system with one scan. The 1<sup>st</sup> and 2<sup>nd</sup> rows represent representative results from two physical phantoms. The 3<sup>rd</sup> and 4<sup>th</sup> rows show the difference images between the results of the 1<sup>st</sup> and 2<sup>nd</sup> rows and the corresponding ground truths. The display window is [0,2000] HU.

details in the reconstructed object. The images reconstructed by the LEARN++, DDPTransformer, and DREAM-Net methods exhibit relatively clear spherical structures, but there is adhesion between the spheres, and some details have not been precisely reconstructed. The GMSD in the sinogram domain faces similar issues as C-GAN. Despite generating results with spherical shapes, its outcomes remain blurry and challenging to discern. The image-domain diffusion method (i.e., DPS) achieves favorable results in reconstruction; however, it exhibits unclear boundaries in the vicinity of the spheres. Finally, the reconstruction results of our proposed method show remarkable similarity to the ground truth for both phantom 1# and phantom 2#. The boundaries of the reconstructed spheres are clear, and their intensity distribution is uniform, demonstrating a significant superiority over other methods and highlighting the excellence of our approach.

Similar to the AAPM dataset simulation, we conduct a quantitative analysis of the phantom's reconstruction results. As illustrated in Tab. I, Our method achieved the highest PSNR among all the methods, surpassing the second-best method DPS by 2.7 dB. Furthermore, our method exhibits a 0.021 higher SSIM with DPS.

**3) Clinical Cardiac CT Reconstruction:** To validate the feasibility of using 29-source multi-source static CT for real-time imaging of the human heart in the future by utilizing the DCDS method, we conducted a comprehensive comparison using a clinical cardiac CT dataset as a prospective experiment. As depicted in Fig. 8, the reconstruction results of our DCDS method significantly outperform those of other methods. This superiority is rooted in DCDS's advanced algorithmic framework, which optimizes for both detail preservation and noise reduction, a critical aspect in cardiac CT imaging. The ROI visualizations highlight our method's enhanced clarity and accuracy, notably in reducing streaking artifacts, a common challenge in cardiac CT reconstructions. This is achieved

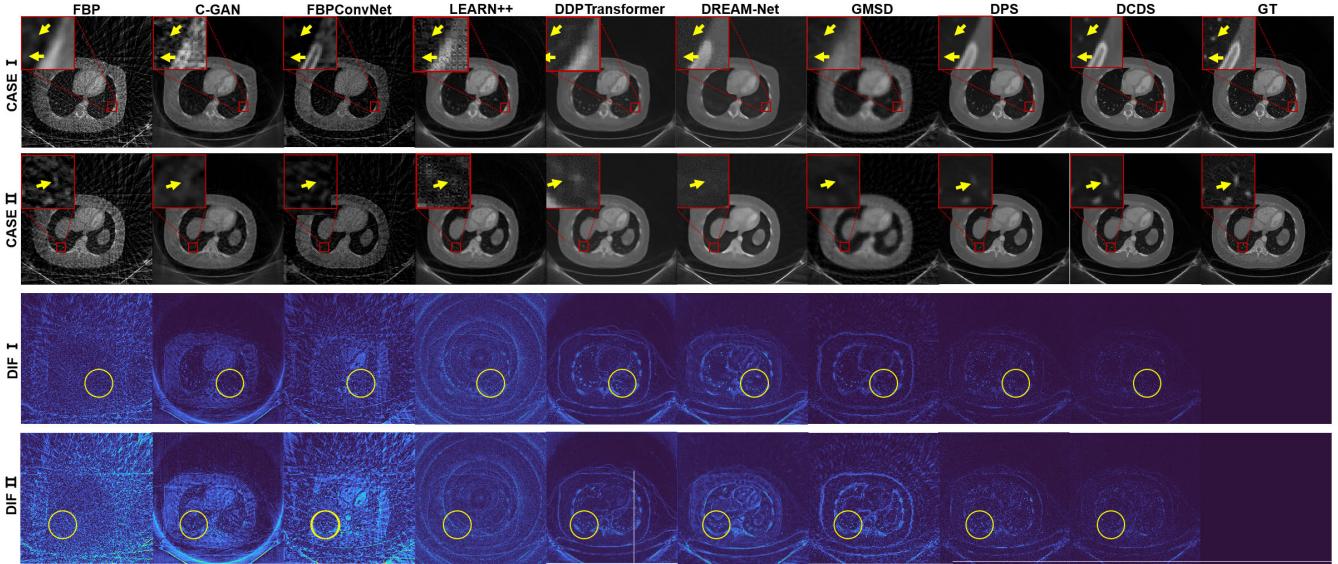
through our novel approach to utilizing the dual-domain information, ensuring high fidelity in structural details.

Furthermore, the difference maps provide empirical evidence of DCDS's closer alignment with the ground truth compared to other methods. This is quantitatively substantiated in Table I, where DCDS demonstrates superior performance in terms of both PSNR and SSIM. These metrics are particularly relevant in our study as they directly correlate with the diagnostic quality of cardiac imaging. PSNR reflects the noise level in reconstructed images, while SSIM indicates the structural similarity to the actual anatomy, both crucial for accurate reconstruction.

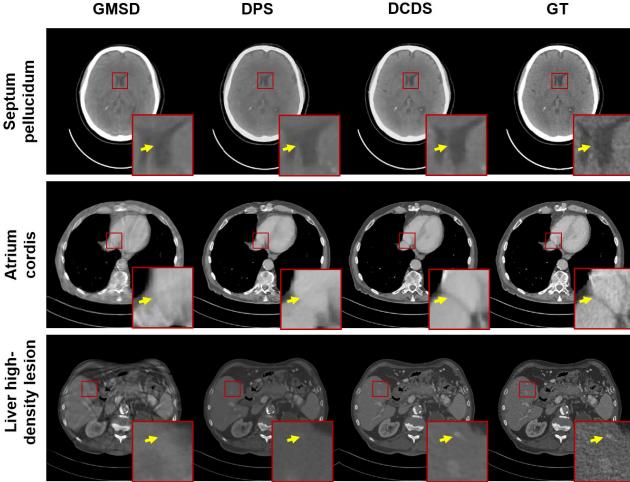
**4) Clinical Structure Reconstruction:** To further validate the effectiveness and accuracy of our proposed DCDS method, we conducted a series of clinical structure reconstruction experiments [53], [54] [55]. These experiments aimed to demonstrate the capability of the DCDS model in reconstructing key anatomical structures, especially in complex medical imaging scenarios [56], [57]. Specifically, as shown in Fig. 9, we conducted experiments about brain CT, cardiac CT, and abdomen CT clinical structure reconstruction.

The brain CT images were sourced from the open-source CQ500 dataset [58]. As shown in the first row of Fig. 9, we focused particularly on the reconstruction of the septum pellucidum. Sandwiched between two high-density regions, the intact structure of the septum pellucidum is crucial for diagnosing certain mental disorders [59]. Our DCDS model successfully reconstructed a complete septum pellucidum extending from the corpus callosum to the fornix, demonstrating its superiority in preserving anatomical integrity compared to competing methods.

The cardiac images were adopted from the AAPM dataset. As displayed in the second row of Fig. 9, we showcased a right atrium reconstructed by DCDS, compared with GMSD and DPS methods. The DCDS model reconstructed a clear



**Fig. 8.** Reconstructed results of the real cardiac dataset using various methods. The 1<sup>st</sup> and 2<sup>nd</sup> rows represent representative results from two cases. The 3<sup>rd</sup> and 4<sup>th</sup> rows show the difference images between the results of the 1<sup>st</sup> and 2<sup>nd</sup> rows and the corresponding ground truths. The display window is [680,1650] HU.



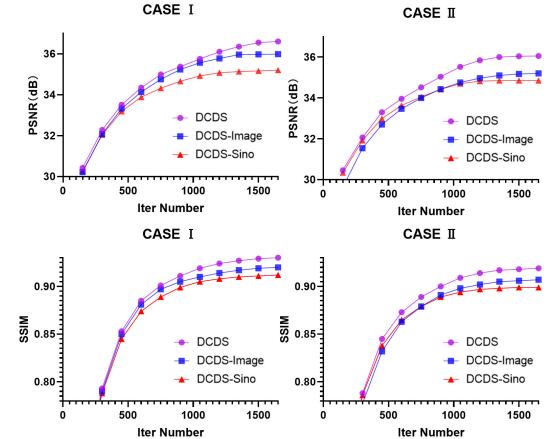
**Fig. 9.** Competing reconstruction methods produced results showcasing vital anatomical structures. In the 1<sup>st</sup> row, the axial section at the corpus callosum exhibits the septum pellucidum. The 2<sup>nd</sup> row highlights the right atrium, while the 3<sup>rd</sup> row displays liver density in abdomen CT.

boundary of the atrium, which can assist in the diagnosis of many heart diseases [60]. This result emphasizes the precision of DCDS in reconstructing cardiac structures.

We also explored the application of the DCDS model in abdomen CT reconstruction. Specifically, accurate reconstruction of liver region densities is vital for screening tumors, cysts, and other lesions. As illustrated in the third row of Fig. 9, compared to other methods, our proposed model exhibits more accurate high-density areas. These accurately reconstructed areas are crucial for clinical diagnosis and treatment planning.

### C. Ablation Study Results

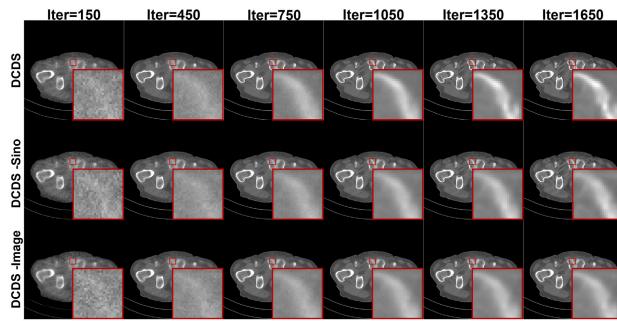
1) **Comparison With Single Domain:** To determine the extent of collaborative improvement achieved by DCDS through the



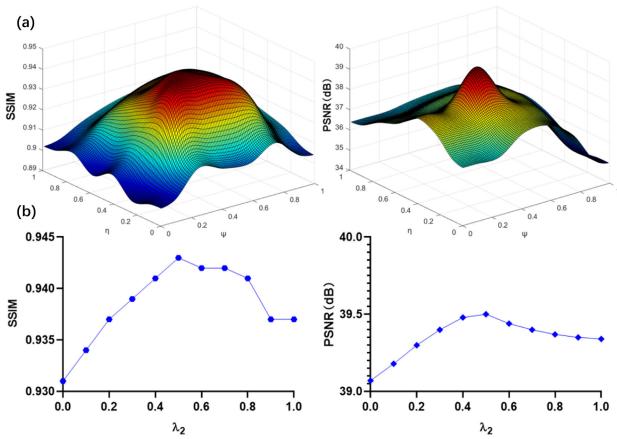
**Fig. 10.** The convergence plots of PSNR and SSIM concerning the number of iterations in the testing phase.

integration of image-domain and sinogram-domain diffusion models, we disabled the sinogram-domain diffusion model in DCDS, resulting in the creation of DCDS-Image. Similarly, deactivating the image-domain diffusion model yields the DCDS-Sino model. As depicted in Fig. 10, it is evident that DCDS outperforms DCDS-Image and DCDS-Sino in terms of both PSNR and SSIM values. Moreover, across the same number of iterations, the metrics of DCDS consistently surpass those of the other two methods. As illustrated in Fig. 11, it is apparent that, at identical iteration counts, the images reconstructed by DCDS exhibit greater clarity compared to those generated by DCDS-Sino and DCDS-Image. This observation underscores the capacity of our dual-domain collaborative approach in accelerating convergence towards a stable solution.

2) **Selection of Regularization Parameters of  $\lambda_1$ ,  $\lambda_2$ ,  $\eta_1$  and  $\psi_1$ :** To maintain the magnification range of the sinogram unaltered, it becomes imperative to ensure the fulfillment of



**Fig. 11.** The reconstructed results of DCDS, DCDS-Image, and DCDS-Sino concerning the number of iterations in the testing phase.



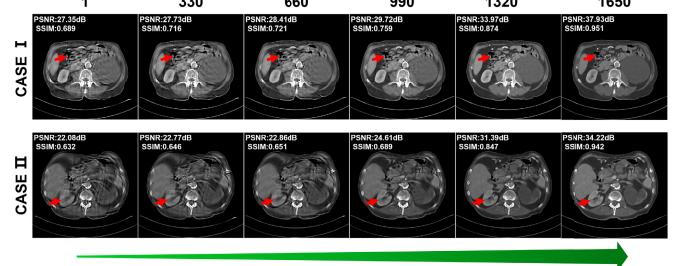
**Fig. 12.** The influence of weight parameter selection between the image-domain diffusion model and the sinogram-domain diffusion model on the PSNR and SSIM metrics.

the equation  $\lambda_1 + \lambda_2 = 1$ . Consequently, our focus narrows down to the parameter selection of  $\lambda_2$ ,  $\eta_1$ , and  $\psi_1$ .

In order to choose the appropriate regularization coefficients for the image-domain diffusion model  $\eta_1$  and the sinogram-domain diffusion model  $\psi_1$ , we conduct parameter selection ablation experiments. Firstly, we find that the values of both  $\eta_1$  and  $\psi_1$  should lie between 0 and 1. We conducted experiments with increments of 0.1 for  $\eta_1$  and  $\psi_1$ , ranging from 0 to 1, as shown in Fig. 12(a). The results indicate that the maximum values of PSNR and SSIM are achieved when  $\eta_1$  is set to 0.6 and  $\psi_1$  is set to 0.4. With the optimal values of  $\eta_1$  and  $\psi_1$  determined, we conduct parameter selection experiments for  $\lambda_2$ . As shown in Fig. 12(b), when  $\lambda_2$  is set to 0.5, both SSIM and PSNR can reach their peak values.

**3) Comparison of Fusion Times:** In order to further verify whether our proposed fusion approach effectively promotes the reconstruction process, we conducted ablation experiments on the fusion times. Specifically, we examined the reconstruction performance under different fusion iteration numbers. We ranged from single-step fusion to every-step fusion (DCDS), linearly increasing the fusion iterations while keeping the total number of iterations constant at 1650. The fusion steps were uniformly sampled from the total number of iterations.

As shown in Fig. 13, we can observe an incremental improvement in image quality from left to right. This indicates



**Fig. 13.** From left to right, it corresponds to single-step fusion to every-step fusion, with the number of fusions increasing linearly.

**TABLE II**  
QUANTITATIVE RESULTS OF DIFFERENT METHODS ON AAPM DATASET.  
THESE METRICS ARE CALCULATED FOR ALL SLICES IN  
THE TEST SET WITH DIFFERENT VIEWS

Method	Time per sample(s)	PSNR(dB)/SSIM			
		23 Views	29 Views	60 Views	120 Views
GMSD	0.63	28.28/0.752	35.98/0.912	36.72/0.919	37.03/0.928
DPS	0.57	31.96/0.868	38.17/0.920	39.78/0.932	40.31/0.936
<b>DCDS</b>	0.97	<b>34.37/0.901</b>	<b>39.60/0.931</b>	<b>41.05/0.943</b>	<b>42.86/0.955</b>

a positive correlation between fusion iterations and the quality of reconstructed images, reflecting that our proposed fusion mechanism significantly enhances the reconstruction effectiveness with each iteration.

## V. DISCUSSION AND CONCLUSION

### A. Discussion

**1) Model Generalization Across Different Views:** In this section, we aim to verify the generalizability of the dual-domain diffusion model method DCDS, in comparison to single-domain diffusion model methods DPS and GMSD, for sparse data from different projection views. Specifically, we compared the quantitative statistical metrics of the three methods when processing the AAPM simulated dataset under three different views: 23, 29, and 60. As shown in Tab. II, the dual-domain diffusion model method DCDS exhibited a smaller decay in PSNR and SSIM metrics when the projection views changed, and it achieved the most outstanding performance overall. This reflects the stronger generalizability of our proposed dual-domain diffusion model method compared to single-domain diffusion models for different projection views. In addition, we also compare the time consumption per sample for these three diffusion modeling methods. The experimental results show that our proposed DCDS sampling method, despite utilizing both image-domain and sinogram-domain diffusion, does not have a huge increase in the sampling time, and can still complete one sample in 0.97 seconds.

**2) Parameter Optimization:** In our study, the parameters  $\lambda_1$ ,  $\lambda_2$ ,  $\eta_1$ , and  $\psi_1$  play pivotal roles in the DCDS method. The equation  $\lambda_1 + \lambda_2 = 1$  is crucial to maintain the magnification range of the sinogram unchanged, which ensures a balanced contribution between different model components.  $\lambda_1$ , specifically, modulates the relative influence of distinct model aspects.

For the selection of regularization coefficients  $\eta_1$  and  $\psi_1$  in the image-domain and sinogram-domain diffusion models,

we performed detailed parameter selection ablation experiments. The choice of the 0 to 1 range for  $\eta_1$  and  $\psi_1$  is grounded in their roles within the diffusion models, where they modulate the diffusion process's strength. Increments of 0.1 were strategically chosen to provide a detailed yet computationally feasible exploration of the parameter space. Our experiments, as illustrated in Fig. 12(a), reveal that the optimal balance of detail preservation and noise reduction is achieved at  $\eta_1 = 0.6$  and  $\psi_1 = 0.4$ . These values reflect an optimal trade-off between image and sinogram domains, as evidenced by the peak values in PSNR and SSIM. Furthermore, the role of  $\lambda_2$  in the reconstruction process is critical. It primarily influences the balance between the fidelity to the data and the regularization strength. Our findings, depicted in Fig. 12(b), indicate that setting  $\lambda_2$  to 0.5 leads to the highest SSIM and PSNR scores.

As the optimal setting for achieving the best reconstruction quality. This setting was determined after observing how variations in  $\lambda_2$  impacted the model's performance, ensuring the reconstructed images' fidelity and detail. To sum up, the careful selection and optimization of parameters  $\lambda_1$ ,  $\lambda_2$ ,  $\eta_1$ , and  $\psi_1$  in our DCDS method directly contribute to its superior performance in clinical cardiac CT reconstruction. The detailed exploration of these parameters not only underscores the method's technical robustness but also its practical applicability in clinical settings, where precision is paramount.

**3) Parameter Transfer:** In our research, we conducted studies on three datasets: the AAPM simulation dataset, a clinical cardiac CT dataset, and a real dataset collected from multi-source CT. Selecting hyper-parameters anew for each dataset is quite complex. Therefore, we discuss the choice of parameters for different datasets here.

Specifically, the selection of hyper-parameters in this study was guided by the performance of parameters on the AAPM simulation data, with further fine-tuning carried out for different datasets as needed. The rationale behind this approach is that although the image characteristics of different datasets vary, the data distributions share certain similarities since all the data are CT images. Furthermore, the physical significance of the hyper-parameter tuning in this paper is the proportion of importance between the sinogram domain and the image domain data. This proportion does not significantly change with the task if the imaging views remain constant. This is also why we adopted this strategy to transfer the hyper-parameter selection.

**4) Future Works:** In our future research endeavors, we plan to conduct simulation studies utilizing the ACR phantom as suggested. By performing scanning experiments on standard phantoms such as the ACR phantom, we aim to further validate our method in a more quantitative manner. In addition, we plan to investigate more imaging scenarios using multi-source stationary CT, such as real-time imaging of the human cardiac, further leveraging the high temporal resolution of multi-source stationary CT.

## B. Conclusion

In this paper, we introduce a groundbreaking approach by combining data distribution priors from both the sinogram

and image domains for multi-source stationary CT reconstruction. The experimental results consistently underscore the effectiveness and adaptability of DCDS across a range of datasets, encompassing sparse-view CT scenarios such as AAPM numerical simulations, multi-source stationary scanned phantoms, and clinical cardiac datasets. Notably, the reconstructed images exhibit enhanced edge sharpness and finer details compared to alternative methods, and all simulation experiments and real-world experiments were conducted using 29 views, consistent with the actual system of the multi-source stationary CT. Additionally, our approach demonstrates superior convergence speed, surpassing that of both image-domain and sinogram-domain methods.

## REFERENCES

- [1] G. N. Hounsfield, "Computed medical imaging," *Med. Phys.*, vol. 7, no. 4, pp. 283–290, Jul. 1980.
- [2] W. Wu, Y. Zhang, Q. Wang, F. Liu, P. Chen, and H. Yu, "Low-dose spectral CT reconstruction using image gradient  $\ell_0$ -norm and tensor dictionary," *Appl. Math. Model.*, vol. 63, pp. 538–557, Nov. 2018.
- [3] W. Wu et al., "Stationary multi-source AI-powered real-time tomography (SMART)," 2021, *arXiv:2108.12076*.
- [4] J. Zhang, H. Mao, X. Wang, Y. Guo, and W. Wu, "Wavelet-inspired multi-channel score-based model for limited-angle CT reconstruction," *IEEE Trans. Med. Imag.*, p. 1, 2024, doi: [10.1109/TMI.2024.3367167](https://doi.org/10.1109/TMI.2024.3367167).
- [5] K. Liang, H. Yang, K. Kang, and Y. Xing, "Improve angular resolution for sparse-view CT with residual convolutional neural network," in *Medical Imaging 2018: Physics of Medical Imaging*, vol. 10573. Bellingham, WA, USA: SPIE, 2018, pp. 382–392.
- [6] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," in *Proc. Int. Conf. Learn. Represent.*, 2021, pp. 1–26.
- [7] Z. Li, Y. Wang, J. Zhang, W. Wu, and H. Yu, "Two-and-a-half order score-based model for solving 3D ill-posed inverse problems," *Comput. Biol. Med.*, vol. 168, Jan. 2024, Art. no. 107819.
- [8] Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–6.
- [9] M. Lee, H. Kim, and H.-J. Kim, "Sparse-view CT reconstruction based on multi-level wavelet convolution neural network," *Phys. Medica*, vol. 80, pp. 352–362, Dec. 2020.
- [10] Z. Zhang, X. Liang, X. Dong, Y. Xie, and G. Cao, "A sparse-view CT reconstruction method based on combination of DenseNet and deconvolution," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1407–1417, Jun. 2018.
- [11] X. Song et al., "Sparse-view reconstruction for photoacoustic tomography combining diffusion model with model-based iteration," *Photoacoustics*, vol. 33, Oct. 2023, Art. no. 100558.
- [12] 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, doi: [10.1109/TRPMS.2023.3309474](https://doi.org/10.1109/TRPMS.2023.3309474).
- [13] C. Shi, Y. Xiao, and Z. Chen, "Dual-domain sparse-view CT reconstruction with transformers," *Phys. Medica*, vol. 101, pp. 1–7, Sep. 2022.
- [14] J. Pan, H. Yu, Z. Gao, S. Wang, H. Zhang, and W. Wu, "Iterative residual optimization network for limited-angle tomographic reconstruction," *IEEE Trans. Image Process.*, vol. 33, pp. 910–925, 2024.
- [15] W. Wu, Y. Wang, Q. Liu, G. Wang, and J. Zhang, "Wavelet-improved score-based generative model for medical imaging," *IEEE Trans. Med. Imag.*, vol. 43, no. 3, pp. 966–979, Mar. 2024.
- [16] X. Sun, X. Li, and P. Chen, "An ultra-sparse view CT imaging method based on X-ray2CTNet," *IEEE Trans. Comput. Imag.*, vol. 8, pp. 733–742, Sep. 2022, doi: [10.1109/TCI.2022.3201390](https://doi.org/10.1109/TCI.2022.3201390).
- [17] H. Chen et al., "LEARN: Learned experts' assessment-based reconstruction network for sparse-data CT," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1333–1347, Jun. 2018.
- [18] Y. Song and S. Ermon, "Improved techniques for training score-based generative models," in *Proc. NIPS*, vol. 33, 2020, pp. 12438–12448.
- [19] H. Chung and J. C. Ye, "Score-based diffusion models for accelerated MRI," *Med. Image Anal.*, vol. 80, Aug. 2022, Art. no. 102479.
- [20] W. Xia, W. Cong, and G. Wang, "Patch-based denoising diffusion probabilistic model for sparse-view CT reconstruction," 2022, *arXiv:2211.10388*.

- [21] W.-A. Lin et al., "DuDoNet: Dual domain network for CT metal artifact reduction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Sep. 2019, pp. 10512–10521.
- [22] B. Huang, L. Zhang, S. Lu, B. Lin, W. Wu, and Q. Liu, "One sample diffusion model in projection domain for low-dose CT imaging," 2022, *arXiv:2212.03630*.
- [23] R. Li et al., "DDPTTransformer: Dual-domain with parallel transformer network for sparse view CT image reconstruction," *IEEE Trans. Comput. Imag.*, vol. 8, pp. 1101–1116, 2022.
- [24] Y. Zhang et al., "LEARN++: Recurrent dual-domain reconstruction network for compressed sensing CT," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 7, no. 2, pp. 132–142, Feb. 2023.
- [25] M. J. Colbrook, V. Antun, and A. C. Hansen, "The difficulty of computing stable and accurate neural networks: On the barriers of deep learning and smale's 18th problem," *Proc. Nat. Acad. Sci. USA*, vol. 119, no. 12, Mar. 2022, Art. no. e2107151119.
- [26] T. Cheslerean-Boghiu, F. C. Hofmann, M. Schultheiß, F. Pfeiffer, D. Pfeiffer, and T. Lasser, "WNet: A data-driven dual-domain denoising model for sparse-view computed tomography with a trainable reconstruction layer," *IEEE Trans. Comput. Imag.*, vol. 9, pp. 120–132, 2023.
- [27] Z. Li et al., "Spectrum learning for super-resolution tomographic reconstruction," *Phys. Med. Biol.*, vol. 69, no. 8, Apr. 2024, Art. no. 085018.
- [28] G. Zang, M. Aly, R. Idoughi, P. Wonka, and W. Heidrich, "Super-resolution and sparse view CT reconstruction," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 137–153.
- [29] D. H. Ye, G. T. Buzzard, M. Ruby, and C. A. Bouman, "Deep back projection for sparse-view CT reconstruction," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Nov. 2018, pp. 1–5.
- [30] Z. Zhao, Y. Sun, and P. Cong, "Sparse-view CT reconstruction via generative adversarial networks," in *IEEE Nucl. Sci. Symp. Med. Imag. Conf. Proc. (NSS/MIC)*, Nov. 2018, pp. 1–5.
- [31] W. Xia, Z. Yang, Q. Zhou, Z. Lu, Z. Wang, and Y. Zhang, "A transformer-based iterative reconstruction model for sparse-view ct reconstruction," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2022, pp. 790–800.
- [32] W. Wu, X. Guo, Y. Chen, S. Wang, and J. Chen, "Deep embedding-attention-refinement for sparse-view CT reconstruction," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–11, 2023.
- [33] W. Wu, D. Hu, C. Niu, H. Yu, V. Vardhanabuti, and G. Wang, "DRONE: Dual-domain residual-based optimization network for sparse-view CT reconstruction," *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 3002–3014, Nov. 2021.
- [34] W. Wu, J. Pan, Y. Wang, S. Wang, and J. Zhang, "Multi-channel optimization generative model for stable ultra-sparse-view CT reconstruction," *IEEE Trans. Med. Imag.*, p. 1, 2024, doi: [10.1109/TMI.2024.3376414](https://doi.org/10.1109/TMI.2024.3376414).
- [35] E. C. Yeats, Y. Chen, and H. Li, "Improving gradient regularization using complex-valued neural networks," in *Proc. Int. Conf. Mach. Learn.*, 2021, pp. 11953–11963.
- [36] M. Kalke and S. Siltanen, "Sinogram interpolation method for sparse-angle tomography," *Appl. Math.*, vol. 5, no. 3, pp. 423–441, 2014.
- [37] H. Kostler, M. Prummer, U. Rude, and J. Hornegger, "Adaptive variational sinogram interpolation of sparsely sampled CT data," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, vol. 3, Aug. 2006, pp. 778–781.
- [38] C. Zhang, Y. Li, and G. Chen, "Accurate and robust sparse-view angle CT image reconstruction using deep learning and prior image constrained compressed sensing (DL-PICCS)," *Med. Phys.*, vol. 48, no. 10, pp. 5765–5781, Oct. 2021.
- [39] D. Lee et al., "A feasibility study of low-dose single-scan dual-energy cone-beam CT in many-view under-sampling framework," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2578–2587, Dec. 2017.
- [40] T. Lee, J. Min, and S. Cho, "Many-view under-sampling (MVUS) technique for low-dose CT: Dose versus image quality," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec. (NSS/MIC)*, Oct. 2012, pp. 3617–3619.
- [41] S. Abbas, T. Lee, S. Shin, R. Lee, and S. Cho, "Effects of sparse sampling schemes on image quality in low-dose CT," *Med. Phys.*, vol. 40, no. 11, Nov. 2013, Art. no. 111915.
- [42] A. Vahdat, K. Kreis, and J. Kautz, "Score-based generative modeling in latent space," in *Proc. NIPS*, 2021, pp. 11287–11302.
- [43] H. Chung, J. Kim, M. T. McCann, M. L. Klasky, and J. C. Ye, "Diffusion posterior sampling for general noisy inverse problems," 2022, *arXiv:2209.14687*.
- [44] H. Ma, L. Zhang, X. Zhu, and J. Feng, "Accelerating score-based generative models with preconditioned diffusion sampling," in *Proc. Eur. Conf. Comput. Vis.*, 2022, pp. 1–16.
- [45] C. McCollough, "TU-FG-207A-04: Overview of the low dose CT grand challenge," *Med. Phys.*, vol. 43, no. 35, pp. 3759–3760, Jun. 2016.
- [46] H. Scherl, M. Koerner, H. Hofmann, W. Eckert, M. Kowarschik, and J. Hornegger, "Implementation of the FDK algorithm for cone-beam CT on the cell broadband engine architecture," in *Medical Imaging 2007: Physics of Medical Imaging*, vol. 6510. Bellingham, WA, USA: SPIE, 2007, pp. 1666–1675.
- [47] 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.
- [48] W. van Aarle et al., "The ASTRA toolbox: A platform for advanced algorithm development in electron tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, Oct. 2015, doi: [10.1016/J.ULTRAMIC.2015.05.002](https://doi.org/10.1016/J.ULTRAMIC.2015.05.002).
- [49] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.
- [50] M. U. Ghani and W. C. Karl, "Deep learning-based sinogram completion for low-dose CT," in *Proc. IEEE 13th Image, Video, Multidimensional Signal Process. Workshop (IVMSP)*, Jun. 2018, pp. 1–5.
- [51] Y. Zhang et al., "DREAM-Net: Deep residual error iterative minimization network for sparse-view CT reconstruction," *IEEE J. Biomed. Health Informat.*, vol. 27, no. 1, pp. 480–491, Jan. 2023.
- [52] M. F. Kijewski and P. F. Judy, "The noise power spectrum of CT images," *Phys. Med. Biol.*, vol. 32, no. 5, pp. 565–575, May 1987.
- [53] S. H. Park, "Regulatory approval versus clinical validation of artificial intelligence diagnostic tools," *Radiology*, vol. 288, no. 3, pp. 910–911, Sep. 2018.
- [54] S. H. Park and K. Han, "Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction," *Radiology*, vol. 286, no. 3, pp. 800–809, Mar. 2018.
- [55] S. H. Park, A.-R. Sul, K. Han, and Y. S. Sung, "How to determine if one diagnostic method, such as an artificial intelligence model, is superior to another: Beyond performance metrics," *Korean J. Radiol.*, vol. 24, no. 7, p. 601, 2023.
- [56] S. H. Park et al., "Methods for clinical evaluation of artificial intelligence algorithms for medical diagnosis," *Radiology*, vol. 306, no. 1, pp. 20–31, Jan. 2023.
- [57] S. H. Park, A.-R. Sul, Y. Ko, H. Y. Jang, and J.-G. Lee, "Radiologist's guide to evaluating publications of clinical research on AI: How we do it," *Radiology*, vol. 308, no. 3, Sep. 2023, Art. no. e230288.
- [58] S. Chilamkurthy et al., "Development and validation of deep learning algorithms for detection of critical findings in head CT scans," *Lancet*, vol. 392, no. 10162, pp. 2388–2396, 2018.
- [59] D. Chen, X. Yan, L. Lu, K. Xue, and X. Dong, "Lymphoma of the central nervous system originating from the septum pellucidum region: Two case reports with literature review," *Medicine*, vol. 102, no. 46, 2023, Art. no. e35954.
- [60] R. M. Lang et al., "Imaging assessment of the right atrium: Anatomy and function," *Eur. Heart J. Cardiovascular Imag.*, vol. 23, no. 7, pp. 867–884, Jun. 2022.