Cross-view Generalized Diffusion Model for Sparse-view CT Reconstruction

Jixiang Chen[1 [0000−0001−9941−8324]], Yiqun Lin[1 [0000−0002−7697−0842]], Yi Qin[1 [0009−0000−2236−652X]], Hualiang Wang[1 [0009−0006−0157−8885]], and Xiaomeng Li[1(✉) [0000−0003−1105−8083]]

Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, China
eexmli@ust.hk

Abstract. Sparse-view computed tomography (CT) reduces radiation exposure by subsampling projection views, but conventional reconstruction methods produce severe streak artifacts with undersampled data. While deep-learning-based methods enable single-step artifact suppression, they often produce over-smoothed results under significant sparsity. Though diffusion models improve reconstruction via iterative refinement and generative priors, they require hundreds of sampling steps and struggle with stability in highly sparse regimes. To tackle these concerns, we present the Cross-view Generalized Diffusion Model (CvG-Diff), which reformulates sparse-view CT reconstruction as a generalized diffusion process. Unlike existing diffusion approaches that rely on stochastic Gaussian degradation, CvG-Diff explicitly models image-domain artifacts caused by angular subsampling as a deterministic degradation operator, leveraging correlations across sparse-view CT at different sample rates. To address the inherent artifact propagation and inefficiency of sequential sampling in generalized diffusion model, we introduce two innovations: Error-Propagating Composite Training (EPCT), which facilitates identifying error-prone regions and suppresses propagated artifacts, and Semantic-Prioritized Dual-Phase Sampling (SPDPS), an adaptive strategy that prioritizes semantic correctness before detail refinement. Together, these innovations enable CvG-Diff to achieve high-quality reconstructions with minimal iterations, achieving **38.34** dB PSNR and **0.9518** SSIM for 18-view CT using only **10** steps on AAPM-LDCT dataset. Extensive experiments demonstrate the superiority of CvG-Diff over state-of-the-art sparse-view CT reconstruction methods. The code is available at https://github.com/xmed-lab/CvG-Diff.

Keywords: Sparse-view CT · CT reconstruction · Diffusion model.

1 Introduction

X-ray computed tomography (CT) is crucial in modern clinical diagnosis. However, exposure to ionizing radiation associated with CT scans poses significant

health risks [2]. Consequently, reducing radiation dose while maintaining diagnostic image quality has become a critical challenge in medical imaging research [21]. Sparse-view CT, which minimizes radiation by subsampling projection views during scanning, offers a promising solution. However, conventional reconstruction methods such as filtered back-projection (FBP) produce severe streak artifacts when applied to undersampled data, limiting their clinical utility.

Recent advances in deep learning have shown great potential for addressing sparse-view CT reconstruction [14,17,15,16,3]. Existing methods mainly focus on developing networks in restoring clean CT image from its sparse-view degraded counterpart using paired training data, operating in either the image domain [29,9,19,12,26], sinogram domain [10,11,7], or both [13,23,24]. To address the challenge of recovering details from heavily undersampled data, researchers have proposed enhancements such as incorporating advanced network building block [23], adding Fourier domain regularization [19], and utilizing feature-level distillation from intermediate-view inputs [12]. Nevertheless, these single-step restoration methods usually produce over-smoothed results. Recent diffusion models leverage generative priors trained on clean CT for iterative posterior sampling, enabling reconstruction across different sparse-view configurations without paired training [4,5,6,27,30,8,25], with [8] achieving best results comparable to state-of-the-art supervised methods. However, their reliance on hundreds of iterative steps and sensitivity to extreme sparsity regimes remains a critical limitation.

Motivated by these challenges, we propose to leverage the inherent correlation between sparse-view CT scans acquired at varying subsampling rates. Our key insight is that sparse-view images from adjacent subsampling rates exhibit shared semantic features and radial artifact patterns, with higher sampling rates preserving more structural details under milder artifacts. Building on generalized diffusion models [1], which enable multi-step restoration by learning inverting arbitrary degradation transformations across severity levels, we introduce the Cross-view Generalized Diffusion Model (CvG-Diff), which reformulates sparse-view CT reconstruction as a generalized diffusion process. Unlike existing diffusion-model-based methods that simulate degradation through Gaussian noise, CvG-Diff explicitly encodes the angular subsampling artifacts as a deterministic degradation operator, enabling unified training across various subsampling rates. In contrast to existing diffusion model [4,5,6,27,8,25] and unified model [28,18] that focus on accommodating arbitrary subsampling rates to minimize the generalization error, CvG-Diff instead aims to exploit the simultaneous reconstruction capability across various subsampling levels, optimizing the reconstruction quality with minimal iterative steps.

Despite these advancements, generalized diffusion models face two critical limitations in sparse-view CT reconstruction. First, intermediate reconstruction errors introduce propagated streak artifacts that models struggle to correct during iterative sampling. Second, conventional sequential sampling often dedicates excessive steps in refining anatomically stable results while under-correcting critical errors. To address these concerns, we propose an Error-Propagating Compos-

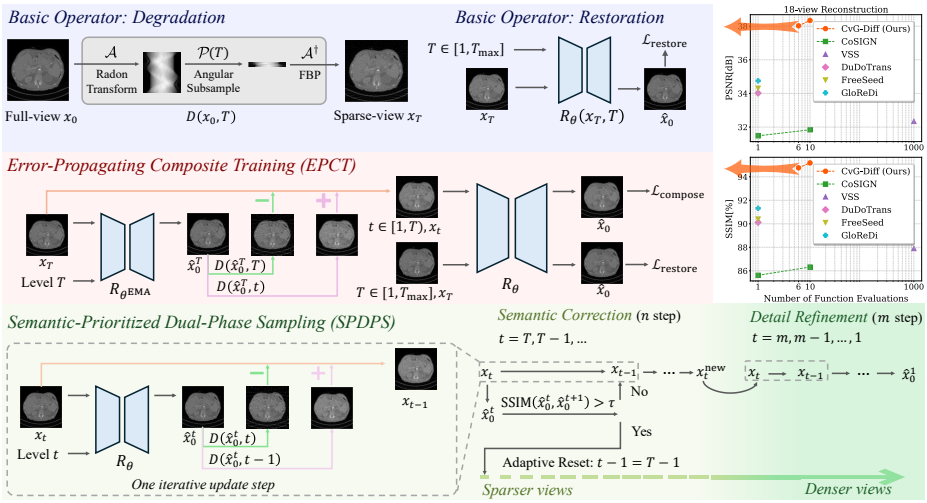

Fig. 1. Overview of the proposed CvG-Diff.

ite Training (EPCT) strategy, which simulates multi-step artifact accumulation during training to tackle propagated errors. Furthermore, we develop Semantic-Prioritized Dual-Phase Sampling (SPDPS), a strategy that prioritizes anatomical correctness before detail refinement. It adaptively resets to input sparse-view level to leverage improved intermediate reconstructions in identifying error-prone regions. As shown in Fig. 1, these innovations enable CvG-Diff to achieve superior results with few sampling steps.

Our contributions are threefold: 1) We propose the CvG-Diff framework that explicitly models angular sparsity artifacts as a deterministic degradation process to harness the simultaneous reconstruction capability across different sparsity regimes, enabling high-quality iterative reconstruction with few sampling steps; 2) We develop the EPCT strategy that suppresses artifact propagation and the SPDPS that further improves reconstruction quality by prioritizing sparse-view level anatomical correctness over detail refinement; 3) Quantitative and qualitative results highlight CvG-Diff's advantages over various state-of-the-art sparse-view CT reconstruction methods.

2 Methodology

2.1 Preliminary: Generalized Diffusion Model

The generalized diffusion process [1] establishes a framework for iterative image restoration through two core components: a degradation operator D that corrupts an input image to various degradation levels, and a restoration operator R trained to invert this degradation. Unlike traditional diffusion models that rely on stochastic Gaussian noise, D can represent arbitrary degradation processes,

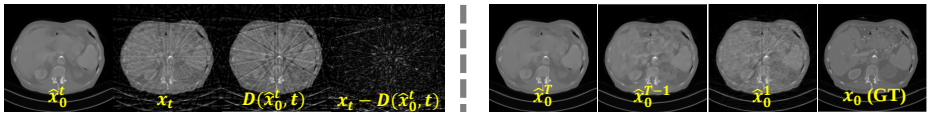

Fig. 2. Error propagation issue in direct extension. **Left:** Incorrect reconstruction introduces streak artifacts. **Right:** Multi-step reconstruction results of performing $I(x_T, T)$. The model struggles to remove accumulated artifacts.

provided it satisfies the boundary condition $D(x_0, 0) = x_0$, where $x_0 \in \mathbb{R}^N$ denotes the clean image.

The forward process generates a degraded image x_t at the severity level t via $x_t = D(x_0, t)$. The restoration operator R_θ, parameterized by a neural network with weights θ, is then optimized to invert this degradation, achieving $R(x_t, t) \approx x_0$ by minimizing the following loss

$$\mathcal{L}_{\text{restore}} = \|R_\theta(D(x_0, t), t) - x_0\|_2. \tag{1}$$

During inference, the reverse sampling process iteratively refines degraded input x_T at known degradation level T by alternating between restoration and degradation operations. Normally, the sampling schedule is the sequential reverse of training levels $t \in \{T, T-1, \ldots, 1\}$, and one updating step is formulated as

$$x_{t-1} = x_t - D(\hat{x}_0^t, t) + D(\hat{x}_0^t, t-1), \tag{2}$$

$$\text{and } \hat{x}_0^t = R_\theta(x_t, t). \tag{3}$$

The final output is $\hat{x}_0 = \hat{x}_0^1 = R_\theta(x_1, 1)$. Formally, we denote the operation of iterative update using Eq. (2) and Eq. (3) following a sequential schedule of $t \in \{T, T-1, \ldots, 1\}$ as $I(x_T, T)$.

2.2 Limitation of Extension to Sparse-view CT Reconstruction

Sparse-view CT reconstruction aims to recover a high-quality image x_0 from subsampled projections of a degraded image x_T. To adapt the generalized diffusion framework to this task, CvG-Diff defines a deterministic degradation operator D that explicitly models artifact patterns induced by angular subsampling:

$$x_T = D(x_0, T) = \mathcal{A}^\dagger \mathcal{P}(T) \mathcal{A} x_0, \tag{4}$$

where $\mathcal{A}$ denotes the Radon Transform, $\mathcal{P}(T)$ applies a subsampling mask to the sinogram at the severity level T, and $\mathcal{A}^\dagger$ represents FBP reconstruction back into image domain. This formulation enables simultaneous training across multiple sparsity levels. In practice, we define a severity level mapping $g(t)$ that assigns each discrete severity level t to a number of views $\mathcal{T}_t$ from a sequence $\mathcal{T}$, spanning from the most-sparse-level $T_{\max}$ to denser ones.

However, direct extension to sparse-view CT suffers from the inherent limitation of artifact propagation during iterative sampling. As shown in Fig 2, in

each step, reconstruction errors within $\hat{x}_0^t$ contribute to additional streak artifacts when calculating $x_t - D(\hat{x}_0^t, t)$. Consequently, networks trained solely with Eq. (1) struggle to correct artifacts that deviate significantly from the current severity level, and the accumulated artifacts lead to subpar reconstructions.

2.3 Cross-view Generalized Diffusion Model

To address error propagation in multi-step sampling while maintaining computational efficiency, CvG-Diff integrates two key innovations: 1) an Error-Propagating Composite Training strategy (EPCT) that mitigates artifact accumulation, and 2) a Semantic-Prioritized Dual-Phase Sampling (SPDPS) mechanism that prioritizes anatomical correctness at sparse-view levels before detail enhancement.

Error-Propagating Composite Training In each training step, we first update R_θ using Eq. (1) with a random target level $T \in [1, T_{\max}]$. Then, we randomly select an intermediate level $t \in [1, T)$ to simulate the propagated artifacts from level T to t. To this end, we maintain an exponential moving averaged (EMA) version of the restoration network $R_{\theta^{\text{EMA}}}$, where $\theta^{\text{EMA}} = \gamma\theta^{\text{EMA}} + (1-\gamma)\theta$ and is updated every p training iterations. We use $R_{\theta^{\text{EMA}}}$ to stably generate reconstruction results at level T, and applies an additional loss as follows

$$\hat{x}_0^T = R_{\theta^{\text{EMA}}}(x_T, T), \tag{5}$$

$$x_t = x_T - D(\hat{x}_0^T, T) + D(\hat{x}_0^T, t), \tag{6}$$

$$\mathcal{L}_{\text{compose}} = \|x_0 - R_\theta(x_t, t)\|_2. \tag{7}$$

We use Eq. (7) to further update R_θ, which is enforced to address potential artifacts introduced in level t and identifies error-prone regions for correction. By this way, CvG-Diff learns to handle artifacts propagated from different levels, enabling reliable reconstruction with larger inter-step subsampling gaps.

Semantic-Prioritized Dual-Phase Sampling During multi-step inference, reconstruction steps at sparse-view levels often introduce over-smoothed errors that blur anatomical boundaries. While reconstruction steps at subsequent denser-view levels refine high-frequency structures, they struggle to fully rectify these significant blurred regions. Therefore, an optimal restoration strategy should prioritize anatomical semantics at sparse-view levels before enhancing finer structures. However, the fixed progression steps in conventional sequential sampling fail to adequately satisfy such requirements (see example in Fig. 4).

To this end, given N sampling steps, we partition $N = n + m$ into first n steps for a semantic correction phase, followed by m steps for a detail refinement phase. For semantic correction, we start with performing $I(x_T, T)$, incorporating an adaptive resetting criterion based on structural similarity (SSIM) comparisons. When updating at t-step, we evaluate whether $\text{SSIM}(\hat{x}_0^t, \hat{x}_0^{t+1}) > \tau$. If the condition is satisfied, we reset to input sparse-view level by

$$x'_{T-1} = x_T - D(\hat{x}_0^t, T) + D(\hat{x}_0^t, T-1). \tag{8}$$

The rationale behind Eq. (8) is that once anatomical convergence is detected, we obtain an improved reconstruction result $\hat{x}_0^t$ over $\hat{x}_0^T$ without wasting excessive

Table 1. Comparison results on AAPM-LDCT. Mean results of RMSE [HU], PSNR [dB], SSIM [%], and reconstruction time [s] are reported. NFE denotes the number of network function evaluation (N) in multi-step reconstruction. The best result is shown in **bold** and the second-best results are underlined.

Method	18-view			36-view			72-view			Time
	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	
One-step feed-forward methods										
DudoTrans [23]	58.83	34.02	90.12	37.78	38.24	94.13	20.56	42.76	97.62	0.13
Freeseed [19]	58.09	34.31	90.40	38.26	37.92	93.93	21.47	42.93	97.53	0.07
Glorei [12]	57.03	34.75	91.32	39.98	37.54	93.52	27.85	41.80	96.30	**0.06**
Multi-step diffusion-based methods										
VSS [8] (NFE=1000)	72.62	32.34	87.90	40.28	37.52	93.99	24.14	41.92	97.07	264.71
CoSIGN [30] (NFE=1)	80.24	31.48	85.62	62.21	33.68	88.71	49.82	35.63	91.40	0.09
CoSIGN [30] (NFE=10)	76.95	31.84	86.31	53.73	34.96	89.67	38.37	37.87	93.20	1.66
CvG-Diff (Ours) (NFE=6)	37.97	38.02	94.76	24.98	41.63	96.92	15.70	45.67	98.54	0.39
CvG-Diff (Ours) (NFE=10)	**36.63**	**38.34**	**95.18**	**24.63**	**41.78**	**97.05**	**15.18**	**45.94**	**98.63**	0.68

steps, and the comparison between $D(\hat{x}_0^t, T)$ and x_T helps to better identify error-prone regions, improving anatomical accuracy at sparser-view levels. Then, we use x'_{T-1} to perform $I(x'_{T-1}, T-1)$. This adaptive mechanism repeats until updating n steps, and we obtain x_t^{new}. Afterwards, for detail refinement, we conduct $\hat{x}_0 = I(x_t^{\text{new}}, m)$ at denser-view levels to obtain the final result.

3 Experiments

3.1 Experiment Settings

We evaluate our method on the AAPM Low-Dose CT (LDCT) dataset [20], which contains 5,936 CT slices from 10 patients. We split the data into 5,410 training/validation slices (9 patients, 9:1 split) and 526 test slices (1 patient). Sparse-view CT projections are simulated using a fan-beam geometry with a source-to-detector distance of 59.5 cm, 672 detector elements, and scan parameters of 120 kVp and 500 mA, implemented via the TorRadon toolbox [22]. To assess robustness across sparsity levels, we reconstruct images from $N_v \in \{18, 36, 72\}$ projection views.

Our reconstruction network adopts a Diffusion UNet architecture with residual blocks, a base feature dimension of 128, and channel multipliers $[1, 2, 2, 2]$ across four resolution scales [1]. The model is trained for 40 epochs with a batch size of 4 using the Adam optimizer ($\beta_1 = 0.9, \beta_2 = 0.999$) with an initial learning rate of 4×10^{-5}, reduced by a factor of 0.8 after epoch 25. The EMA model is updated by setting $\gamma = 0.995$ and $p = 10$. We set $\mathcal{T} = [288, 234, 180, 126, 72, 54, 36, 18]$ that covers all three target numbers of views (*i.e.*, 18, 36, 72), such that only one model needs to be trained to handle all our experimental settings. For SPDPS strategy, we set $\tau = 0.97$ and $m = 4$. Experiments are conducted on a single NVIDIA 3090 GPU. All sparse-view CT reconstruction methods are evaluated quantitatively in terms of root mean squared error (RMSE) in Hounsfield Units (HU), peak signal-to-noise ratio (PSNR), and structural similarity (SSIM).

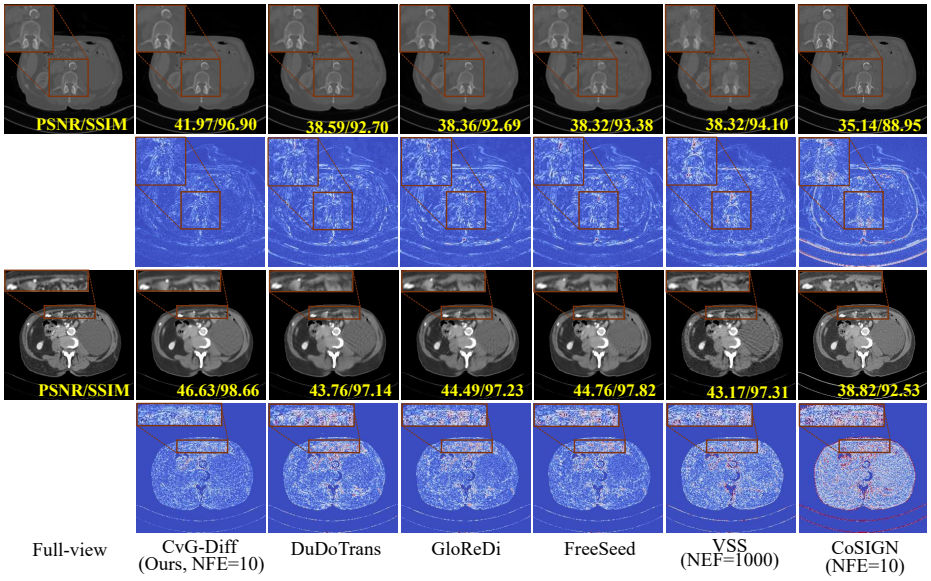

Fig. 3. Visual comparison of different methods. From top to bottom: $N_v = \{36, 72\}$ with display windows $[-1000, 2000], [-200, 300]$ HU, respectively. Red color in error maps indicate a larger error.

3.2 Comparison with State-of-The-Arts

We evaluate CvG-Diff against two categories of state-of-the-art sparse-view CT reconstruction methods: (1) one-step feed-forward methods (FreeSeed [19], GloReDi [12], DuDoTrans [23]), which train networks on paired sparse-view and full-view data for artifact suppression at specific subsampling rates; (2) multi-step diffusion-based methods (VSS [8], CoSIGN [30]), which iteratively refine reconstructions using generative priors on full-view data.

Quantitative results in Tab. 1 demonstrate that one-step methods outperform diffusion baselines due to artifact-specific optimization. On the other hand, diffusion frameworks like VSS achieve comparable results at higher sampling rates (e.g., 72-view) using a unified model that bypasses paired data training, albeit at the cost of prohibitively long sampling steps. CoSIGN mitigates this inefficiency through consistency distillation on paired data, reducing sampling steps while retaining multi-step refinement capabilities. Nevertheless, both diffusion-based approaches struggle under extreme sparsity (e.g., 18-view), where insufficient projection measurements amplify reconstruction errors. In contrast, CvG-Diff consistently outperforms all baselines by leveraging cross-view reconstruction capability across subsampling levels, achieving superior results at all tested sparsity regimes (18/36/72-view) with ≤ 10 sampling steps and comparable inference time. Qualitative results in Fig. 3 highlight the advantages of CvG-Diff. In high-sparsity cases, one-step methods produce over-smoothed outputs, while

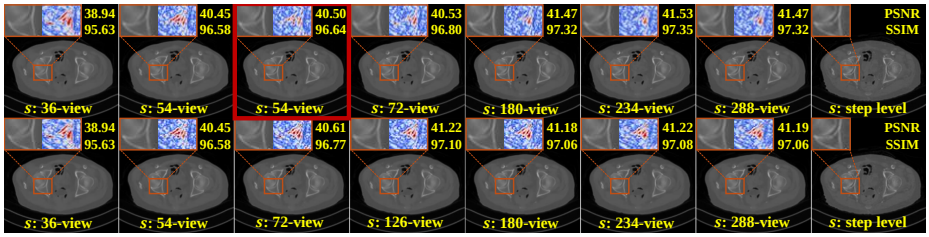

Fig. 4. Visualization of multi-step reconstruction results between SPDPS (top) and sequential sampling (bottom). Error maps show deviations from ground truth (last column) in zoomed region, and red color indicates larger errors. The red box denotes the step when SPDPS resets the degradation level, which rectifies blurred boundaries.

Table 2. Ablation for EPCT and SPDPS.

Settings			18-view	36-view	72-view
ID	EPCT	SPDPS	PSNR↑/SSIM↑	PSNR↑/SSIM↑	PSNR↑/SSIM↑
A)	✗	✗	33.67/82.42	37.02/89.15	43.38/97.30
B)	✗	✓	34.67/85.23	38.23/92.07	43.23/97.54
C)	✓	✗	37.85/94.68	41.48/96.87	45.66/98.54
Ours	✓	✓	**38.34**/**95.18**	**41.78**/**97.05**	**45.94**/**98.63**

Table 3. Influence of τ and m for SPDPS.

Parameter	18-view	Parameter	18-view
τ	PSNR↑/SSIM↑	m	PSNR↑/SSIM↑
0.97	**38.34**/**95.18**	2	38.33/95.13
0.98	38.06/95.03	3	**38.34**/95.15
0.99	37.94/94.96	4	**38.34**/**95.18**

diffusion-based methods reconstruct visually vivid structures with low fidelity. Instead, CvG-Diff successfully recovers precise anatomical structures.

3.3 Ablation Study

We validate our proposed strategies by comparing three CvG-Diff variants: A) The baseline implementation of the generalized diffusion framework using degradation operator defined in Eq. (4); B) Variant A + SPDPS inference strategy; C) Variant A + EPCT strategy during training and uses sequential sampling during inference.

Quantitative results in Tab. 2 demonstrate the incremental benefits of our innovations. While variant B improves over variant A, the limited performance indicates that enhancing inference efficiency alone cannot address the core challenge of artifact propagation. Variant C resolves this limitation through EPCT, which simulates multi-step degradation during training to suppress error accumulation, yielding a significant 3.80 dB averaged PSNR improvement over variant A. The best performance is achieved by combining both EPCT with SPDPS. Fig. 4 compares variant C and D, demonstrating that SPDPS corrects blurred anatomical boundaries by adaptively resetting degradation levels, enhancing subsequent detail refinement. In contrast, sequential sampling propagates uncorrected errors. Further analysis in Tab. 3 validates the robustness of SPDPS across different τ and m configurations, with τ (degradation reset) impacting performance more significantly than m (refinement steps).

4 Conclusion

In this study, we present CvG-Diff, a novel framework for sparse-view CT reconstruction based on the generalized diffusion process. By addressing two critical limitations regarding the artifact propagation across iterative sampling and inefficient correction of over-smoothed artifacts at sparse-view levels, CvG-Diff achieves superior reconstruction results with few (≤ 10) sampling steps across various sparse-view scenarios. Our findings demonstrate the benefits of harnessing the simultaneous reconstruction capability across different sparse-view levels. Future research could explore dual-domain generalized diffusion models that jointly optimize sinogram and image-domain reconstructions and the corresponding optimal sampling strategy with collaborative efforts in both domains. Further investigation of this direction will be a focus of our future work.

Acknowledgments. This work was partially supported by grants from the Hong Kong Innovation and Technology Fund under Project PRP/041/22FX.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bansal, A., Borgnia, E., Chu, H.M., Li, J., Kazemi, H., Huang, F., Goldblum, M., Geiping, J., Goldstein, T.: Cold diffusion: Inverting arbitrary image transforms without noise. Advances in Neural Information Processing Systems **36**, 41259–41282 (2023)
2. Brenner, D.J., Hall, E.J.: Computed tomography—an increasing source of radiation exposure. New England journal of medicine **357**(22), 2277–2284 (2007)
3. Chen, J., Lin, Y., Sun, H., Li, X.: Spatial-division augmented occupancy field for bone shape reconstruction from biplanar x-rays. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 668–678. Springer (2024)
4. Chung, H., Kim, J., Mccann, M.T., Klasky, M.L., Ye, J.C.: Diffusion posterior sampling for general noisy inverse problems. arXiv preprint arXiv:2209.14687 (2022)
5. Chung, H., Kim, J., Ye, J.C.: Direct diffusion bridge using data consistency for inverse problems. Advances in Neural Information Processing Systems **36**, 7158–7169 (2023)
6. Chung, H., Sim, B., Ryu, D., Ye, J.C.: Improving diffusion models for inverse problems using manifold constraints. Advances in Neural Information Processing Systems **35**, 25683–25696 (2022)
7. Dong, X., Vekhande, S., Cao, G.: Sinogram interpolation for sparse-view micro-ct with deep learning neural network. In: Medical Imaging 2019: Physics of Medical Imaging. vol. 10948, pp. 692–698. SPIE (2019)
8. He, L., Du, W., Liao, P., Fan, F., Chen, H., Yang, H., Zhang, Y.: Solving zero-shot sparse-view ct reconstruction with variational score solver. IEEE Transactions on Medical Imaging (2024)

9. Jin, K.H., McCann, M.T., Froustey, E., Unser, M.: Deep convolutional neural network for inverse problems in imaging. IEEE transactions on image processing **26**(9), 4509–4522 (2017)
10. Lee, H., Lee, J., Cho, S.: View-interpolation of sparsely sampled sinogram using convolutional neural network. In: Medical Imaging 2017: Image Processing. vol. 10133, pp. 617–624. SPIE (2017)
11. Lee, H., Lee, J., Kim, H., Cho, B., Cho, S.: Deep-neural-network-based sinogram synthesis for sparse-view ct image reconstruction. IEEE Transactions on Radiation and Plasma Medical Sciences **3**(2), 109–119 (2018)
12. Li, Z., Ma, C., Chen, J., Zhang, J., Shan, H.: Learning to distill global representation for sparse-view ct. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 21196–21207 (2023)
13. Lin, W.A., Liao, H., Peng, C., Sun, X., Zhang, J., Luo, J., Chellappa, R., Zhou, S.K.: Dudonet: Dual domain network for ct metal artifact reduction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10512–10521 (2019)
14. Lin, Y., Luo, Z., Zhao, W., Li, X.: Learning deep intensity field for extremely sparse-view cbct reconstruction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 13–23. Springer (2023)
15. Lin, Y., Wang, H., Chen, J., Li, X.: Learning 3d gaussians for extremely sparse-view cone-beam ct reconstruction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 425–435. Springer (2024)
16. Lin, Y., Wang, H., Chen, J., Yang, J., Guo, J., Li, X.: Deepsparse: A foundation model for sparse-view cbct reconstruction. arXiv preprint arXiv:2505.02628 (2025)
17. Lin, Y., Yang, J., Wang, H., Ding, X., Zhao, W., Li, X.: C^ 2rv: Cross-regional and cross-view learning for sparse-view cbct reconstruction. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11205–11214 (2024)
18. Ma, C., Li, Z., He, J., Zhang, J., Zhang, Y., Shan, H.: Universal incomplete-view ct reconstruction with prompted contextual transformer. arXiv preprint arXiv:2312.07846 (2023)
19. Ma, C., Li, Z., Zhang, J., Zhang, Y., Shan, H.: Freeseed: Frequency-band-aware and self-guided network for sparse-view ct reconstruction. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 250–259. Springer (2023)
20. McCollough, C.: Tu-fg-207a-04: overview of the low dose ct grand challenge. Medical physics **43**(6Part35), 3759–3760 (2016)
21. Miller, D.L., Schauer, D.: The alara principle in medical imaging. philosophy **44**(6), 595–600 (1983)
22. Ronchetti, M.: Torchradon: Fast differentiable routines for computed tomography. arXiv preprint arXiv:2009.14788 (2020)
23. Wang, C., Shang, K., Zhang, H., Li, Q., Zhou, S.K.: Dudotrans: dual-domain transformer for sparse-view ct reconstruction. In: International Workshop on Machine Learning for Medical Image Reconstruction. pp. 84–94. Springer (2022)
24. Wu, W., Hu, D., Niu, C., Yu, H., Vardhanabhuti, V., Wang, G.: Drone: Dual-domain residual-based optimization network for sparse-view ct reconstruction. IEEE Transactions on Medical Imaging **40**(11), 3002–3014 (2021)
25. Wu, W., Pan, J., Wang, Y., Wang, S., Zhang, J.: Multi-channel optimization generative model for stable ultra-sparse-view ct reconstruction. IEEE Transactions on Medical Imaging (2024)

26. Xia, W., Yang, Z., Lu, Z., Wang, Z., Zhang, Y.: Regformer: A local–nonlocal regularization-based model for sparse-view ct reconstruction. IEEE Transactions on Radiation and Plasma Medical Sciences **8**(2), 184–194 (2023)
27. Xu, K., Lu, S., Huang, B., Wu, W., Liu, Q.: Stage-by-stage wavelet optimization refinement diffusion model for sparse-view ct reconstruction. IEEE Transactions on Medical Imaging (2024)
28. Yang, L., Huang, J., Yang, G., Zhang, D.: Ct-sdm: A sampling diffusion model for sparse-view ct reconstruction across various sampling rates. IEEE Transactions on Medical Imaging (2025)
29. Zhang, Z., Liang, X., Dong, X., Xie, Y., Cao, G.: A sparse-view ct reconstruction method based on combination of densenet and deconvolution. IEEE transactions on medical imaging **37**(6), 1407–1417 (2018)
30. Zhao, J., Song, B., Shen, L.: Cosign: Few-step guidance of consistency model to solve general inverse problems. In: European Conference on Computer Vision. pp. 108–126. Springer (2024)