AGENTIC VOLUMETRIC CT RECONSTRUCTION VIA PHYSICS-GROUNDED HYPOTHESIS SEARCH

Anonymous authors
Paper under double-blind review

ABSTRACT

Diffusion-based 3D CT reconstruction typically follows a single stochastic inference trajectory, limiting its ability to explore alternative volumetric solutions when measurements are incomplete. We introduce an agentic, physics-grounded framework that reformulates reconstruction as adaptive hypothesis search under a fixed test-time computation budget, without retraining the diffusion prior. A vision-language model (VLM) transfers cross-case reconstruction knowledge into a search prior that only prioritises which unexplored volumetric action should be tried first. Each action is realised by Hierarchical Physics-Compatible Evolution (HPCE), which combines shallow diffusion re-entry, full-volume data consistency, and physics-compatible innovation to generate structured 3D hypotheses whose diffusion-proposed structural changes remain constrained by the fit measurements. Realised hypotheses are evaluated using projection views held out from candidate formation, providing a case-specific, ground-truth-free physical state value. Physics-Grounded Monte Carlo Tree Search (PG-MCTS) uses this feedback to adaptively allocate the remaining computation budget across competing reconstruction trajectories. After search, we commit the selected action lineage rather than any search-stage volume, replay it using all acquired measurements, and accept each transition only when the full-measurement projection residual does not increase. Learning therefore prioritises where to search, held-out measurements determine what is worth pursuing, and full measurements govern final acceptance. Experiments on sparse-view and limited-angle CT, including cross-dataset evaluation, demonstrate consistent improvements over strong diffusion reconstruction baselines and matched-budget non-adaptive search, and a cross-task MRI study further shows gains over the diffusion baselines.

1 INTRODUCTION

Computed tomography (CT) reconstructs the three-dimensional (3D) structure of the human body from X-ray projections acquired at multiple angles, and reconstruction quality directly affects subsequent anatomical assessment, lesion identification, and quantitative analysis. An ideal reconstruction should not only accurately recover local anatomical details within individual slices, but also maintain structural continuity and spatial consistency throughout the 3D volume. In recent years, diffusion and score-based generative models have been introduced into CT reconstruction and general inverse problems Song et al. (2021); Chung et al. (2023a;b). By leveraging generative priors learned from data, these methods provide strong structural constraints under incomplete measurements and have shown substantial potential for recovering plausible anatomical structures Chung et al. (2023b; 2024); Du et al. (2026).

However, many existing diffusion-based reconstruction methods still primarily follow a single stochastic inference trajectory Chung et al. (2023a;b; 2024). Once the initialisation, stochastic variables, and sampling process are determined, the algorithm typically continues evolving along the current trajectory and returns a single reconstruction, rather than explicitly maintaining, comparing, and further exploring multiple alternative 3D reconstruction hypotheses. This limitation becomes particularly relevant under severe undersampling, where the ill-posedness of the inverse problem, together with the stochastic nature of diffusion sampling, may admit multiple structurally plausible candidate volumes that remain compatible with the available observations. Recent work on inference-time scaling further suggests that exploring and comparing alternative diffusion trajecto-

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

ries can provide additional gains beyond following a single sampling path Ma et al. (2025); Jain et al. (2025); Ramesh & Mardani (2025). The problem is therefore no longer only how to obtain a better reconstruction along a single trajectory, but also where limited test-time computation should be spent among multiple possible reconstruction directions.

Recent studies on inference-time scaling Ma et al. (2025); Jain et al. (2025); Ramesh & Mardani (2025) have shown that increasing test-time computation and searching over different diffusion generation trajectories can further improve generation quality. However, directly applying this idea to volumetric medical CT reconstruction presents two key challenges. First, effective search requires a reliable evaluation signal. Perceptual, semantic, or task-specific verifiers may contain biases, and stronger search may further exploit such biases Ma et al. (2025). For medical reconstruction, it is not sufficient for a candidate to merely "look plausible" Bhadra et al. (2021); Webber & Reader (2024); it should also be supported by the actual measurements acquired from the current case. Second, the candidates themselves are expensive to generate. A straightforward strategy would be to independently generate a large number of complete 3D reconstructions and then select the best one. However, a single physics-constrained volumetric reconstruction is already computationally expensive, making large-scale random resampling impractical. What is needed is therefore not simply more candidates, but an adaptive mechanism that concentrates the limited computation budget on directions that are more worth exploring. This leads to two questions: without ground truth, what evidence can determine whether a realised 3D candidate deserves further computation? And before an expensive candidate is generated, how can we decide which direction should be tried first?

For the first question, we let the measurements themselves serve as the basis for evaluation. Using measurement splitting to obtain ground-truth-free supervision has precedent in tomography Hendriksen et al. (2020), but it has been used for training. We use the same principle for a different purpose by dividing the acquired measurements into fit views used to construct candidates and held-out review views that do not participate in candidate formation. Each realised 3D hypothesis is evaluated according to how well it explains the projections that were not used to form it, yielding a case-specific, ground-truth-free physical state value. Because all search states are evaluated using the same set of held-out measurements, this value provides a common basis for comparing candidates across branches and search depths.

For the second question, held-out measurement evidence is available only after a candidate has been generated. Before generating a candidate, we therefore need a prior that exploits cross-case experience to decide what is worth trying first. The search state on which this decision is based is inherently multimodal, containing not only the current reconstruction image but also structured context such as the search depth, action history, and physical residuals. Pretrained vision-language models (VLMs) provide rich latent representations that jointly capture visual content and contextual relationships, making them a natural basis for modelling such states. We further adapt the VLM with LoRAHu et al. (2022) so that it learns recurring relationships between reconstruction states and promising actions across cases, thereby converting cross-case experience into an action prior for search. The VLM is used only to prioritise unexplored actions and does not evaluate realised candidates or determine final acceptance.

Building on these ideas, we propose a physics-grounded volumetric hypothesis search framework that treats a frozen diffusion prior as a repeatedly callable hypothesis generator and embeds it within a closed-loop test-time search process. The search starts from a root reconstruction obtained using the fit views. The VLM provides the expansion order over three structured volumetric actions operating at different spatial scales. When a new action is explored, Hierarchical Physics-Compatible Evolution (HPCE) turns it into a new 3D candidate through shallow diffusion re-entry, full-volume data consistency, and physics-compatible innovation, so that structural changes proposed by the frozen diffusion prior remain constrained by the fit measurements. Each newly generated candidate is then physically evaluated using the review views, and Physics-Grounded Monte Carlo Tree Search (PG-MCTS) uses this feedback to adaptively allocate the remaining fixed test-time computation budget. Consequently, our exploration does not rely on broad random sampling over the large space induced by the diffusion prior, but instead performs structured exploration in the neighbourhood of the current reconstruction state. Each subsequent computation is selected according to the current reconstruction state and the physical evidence accumulated so far. This closed-loop state–action–feedback process makes the reconstruction agentic.

After search, we do not use any search-stage volume as the final reconstruction. Instead, we freeze the selected action lineage. We then construct a full-measurement root using all projections and replay the frozen lineage from this root. During replay, each step invokes the frozen diffusion prior again and enforces physical constraints using all measurements. A new state is accepted only if it does not increase the projection residual computed over all measurements; otherwise, replay is terminated and the last accepted state is returned. In this way, the search stage discovers how to reconstruct, while the final stage executes the discovered strategy using all available evidence. In summary, learning prioritises where to search, held-out measurements determine whether computation should continue, and all measurements determine what is ultimately retained.

The main contributions are: 1) We propose an agentic, physics-grounded framework that casts diffusion-based 3D reconstruction as a closed-loop search over reconstruction trajectories under a fixed test-time computation budget. Held-out measurements provide a ground-truth-free state value, while a learned cross-case prior only prioritises the order in which unexplored actions are introduced. 2) We develop HPCE, which realises structured volumetric reconstruction actions through shallow diffusion re-entry, full-volume data consistency, and physics-compatible innovation, producing distinct 3D hypotheses without retraining the diffusion prior. 3) We introduce trajectory replay with physics acceptance, which replays the selected action lineage using all acquired measurements and accepts each step only if the full-measurement projection residual does not increase, ensuring that measurements held out during search still contribute to the final reconstruction. We validate the framework under sparse-view and limited-angle acquisition on AAPM test subjects excluded from the training of all learned components, and on the LIDC dataset without any retraining. Additional cross-task experiments on MRI are provided in the supplementary material.

2 Related Work

Diffusion-Based Volumetric CT Reconstruction. Diffusion and score-based models provide learned priors that can be combined with physical forward models at inference time and have been applied to CT and other inverse problems Song et al. (2022); Chung et al. (2023a); Kawar et al. (2022). Because 3D diffusion models are computationally costly, volumetric reconstruction methods often reuse 2D priors with mechanisms for volumetric coherence. DiffusionMBIR augments a 2D diffusion prior with model-based regularisation along the remaining spatial dimension Chung et al. (2023b); DiffusionBlend learns a 3D prior by blending position-aware scores across neighbouring slices Song et al. (2024); DDS combines diffusion sampling with Krylov-subspace conjugate-gradient updates for large-scale inverse problems, including 3D CT Chung et al. (2024); and ISCS coordinates stochastic noise across slices to improve inter-slice consistency Du et al. (2026). Our method builds on this line but studies how alternative full-volume trajectories can be generated, physically evaluated, and adaptively pursued under a fixed test-time computation budget.

Test-Time Search for Diffusion Models. Additional inference-time computation can improve diffusion generation by explicitly searching over alternative noises or trajectories Ma et al. (2025). Diffusion Tree Sampling constructs a tree over the diffusion process and propagates terminal rewards to reuse computation across trajectories Jain et al. (2025), while Noise Trajectory Search formulates denoising as a sequential decision problem and finds MCTS meaningful but impractical at the level of individual denoising steps Ramesh & Mardani (2025). These methods search over intermediate states along the diffusion chain and rely on external or task-defined rewards to evaluate trajectories. In our volumetric CT reconstruction setting, search states are instead complete 3D reconstruction hypotheses produced under full-volume physics, and the action space is deliberately kept small and coarse to make tree search tractable. We withhold a subset of the acquired measurements from hypothesis construction and use them to evaluate realised hypotheses. Measurement splitting itself is established in computational imaging: Noise2Inverse partitions tomographic measurements to construct ground-truth-free learning signals Hendriksen et al. (2020), while SSDU partitions acquired MRI measurements between data consistency and self-supervised loss construction Yaman et al. (2020). In contrast, we use held-out measurements as an online, case-specific state value for comparing realised hypotheses across branches and search depths. A learned cross-case prior only determines the order in which unexplored actions are instantiated, rather than the state value or final acceptance. Finally, search commits a trajectory rather than a volume and replays it using all acquired measurements, so measurements withheld for search-time evaluation still contribute to the final reconstruction.

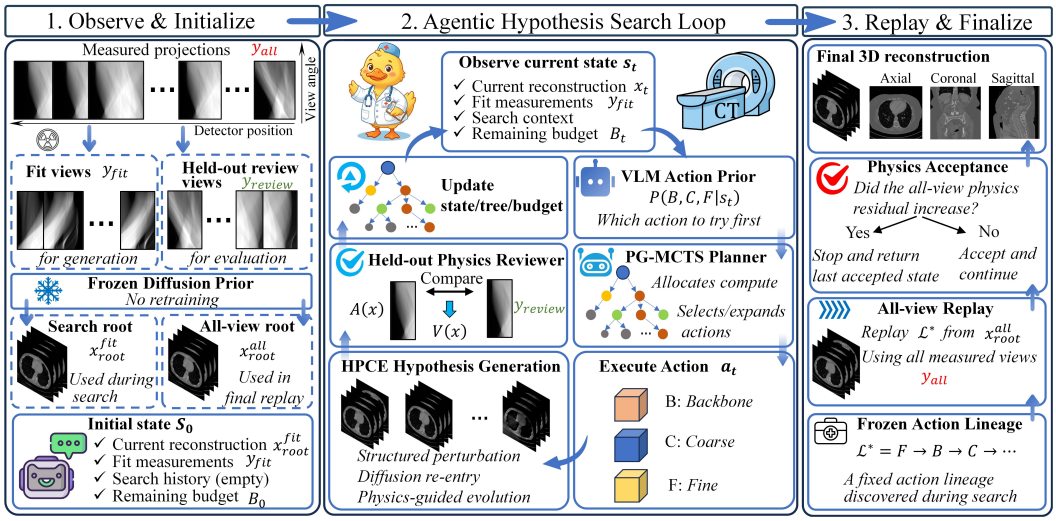

Figure 1: Overview of the proposed agentic, physics-grounded framework for 3D CT reconstruction. Fit views construct the search root, while held-out review views provide physics-based feedback during hypothesis search. A VLM prior only orders which unexplored action is tried first, HPCE generates structured 3D hypotheses, and PG-MCTS allocates test-time computation according to held-out evaluation. After search, the selected action lineage, rather than any search-stage volume, is replayed using all acquired measurements with monotone physics acceptance.

3 METHOD

As illustrated in Fig. 1, we formulate 3D CT reconstruction as a physics-grounded search over volumetric reconstruction hypotheses rather than a single stochastic generation process. Sec. 3.1 introduces structured hypothesis generation from a reconstruction state. Sec. 3.2 presents their exploration under a fixed test-time computation budget, starting from an initial reconstruction obtained from the fit measurements: a cross-case learned prior decides which hypothesis to try first, while held-out measurement evidence determines where subsequent computation is spent. Finally, Sec. 3.3 freezes the selected action lineage and replays it using all acquired measurements, with full-measurement physics determining how far along the lineage the replay is accepted.

3.1 VOLUMETRIC HYPOTHESIS CONSTRUCTION

Reconstructing a 3D volume from sparse-view or limited-angle measurements is inherently ambiguous because incomplete observations may admit multiple plausible solutions. Diffusion and score-based generative priors can constrain such inverse problems by combining learned image statistics with measurement information Song et al. (2021); Chung et al. (2023a). A single stochastic diffusion reconstruction, however, explores only one possible reconstruction trajectory. We therefore use the diffusion prior as a repeatedly callable hypothesis generator and formulate reconstruction as a physics-grounded exploration of alternative volumetric hypotheses. Rather than broadly sampling unrelated volumes from the generative space, which is expensive and ignores information already contained in the current reconstruction, we introduce Hierarchical Physics-Compatible Evolution (HPCE) to perform structured local exploration. HPCE organises possible updates into a small set of structured hypothesis types with different scales and evolution characteristics, each realised as a complete 3D volume. At reconstruction state s_k, let $x_k = x(s_k)$ denote the associated 3D volume. We define three HPCE actions $a_k \in \{B, C, F\}$, corresponding to backbone, coarse, and fine structured full-volume updates, respectively. All three actions share the same HPCE evolution framework but use different predefined hierarchical update configurations. Only one action is applied at each step, producing the next state $s_{k+1} = T_{\text{HPCE}}(s_k, a_k; y_{\text{fit}})$, where y_{fit} denotes the measurements used for candidate formation. A separate subset of measurements is withheld from HPCE candidate formation and used only after a candidate has been generated for held-out evaluation (Sec. 3.2). The exact definitions and implementation settings of B, C, and F are provided in the supplementary material. Because every action operates on the state produced by the preceding step, different action

orderings define different reconstruction trajectories. For example, $F \to B \to B$ represents a single trajectory of three successive HPCE transitions: F maps the initial state to the first successor, after which each B action operates on the state produced by the preceding transition. The search therefore operates over state-dependent 3D reconstruction trajectories rather than isolated candidate volumes. Each selected action is realised as a new full-volume hypothesis by combining shallow diffusion re-entry, full-volume data consistency, and physics-compatible innovation. Shallow diffusion re-entry uses the frozen diffusion prior to propose a structurally meaningful update, while the fit measurements enforce consistency with the observed data. Physics-compatible innovation further constrains the diffusion-proposed change before it is incorporated into the reconstructed volume. The diffusion prior remains frozen throughout, allowing HPCE to explore alternative test-time reconstruction trajectories without retraining the underlying generative model. The mathematical formulation and solution procedure of physics-compatible innovation are provided in the supplementary material.

3.2 VLM-PRIOR PG-MCTS

HPCE turns a single diffusion reconstruction into a branching space of 3D reconstruction trajectories. At each state, the system may choose among the three structured actions B, C, and F for the next evolution step, so the number of possible trajectories can grow exponentially with search depth. Each new direction, however, requires a complete HPCE execution and the generation of a full 3D candidate. Exhaustive enumeration is therefore infeasible, and uniformly distributing a limited budget across directions is inefficient. We formulate this process as a sequential search problem under a fixed test-time budget, with the goal of progressively concentrating computation on more promising reconstruction trajectories according to measurement evidence acquired during search.

Monte Carlo tree search is well suited to this problem because action values are state-dependent and can have long-term consequences. An action not only determines the current 3D candidate but also changes the reconstruction state inherited by all subsequent actions, so its long-term value cannot be judged from a single step alone. Moreover, the consequence of an action becomes observable only after HPCE has been executed and the resulting candidate evaluated against the measurements. Greedy step-by-step selection may therefore overlook trajectories whose immediate gain is limited but whose later potential is greater. PG-MCTS accordingly distributes computation between trajectories with stronger measurement support and directions that remain under-explored, while adapting its search focus as new evaluations become available.

To obtain a case-specific search value, we split the acquired measurements into y_{fit}, used by HPCE for candidate formation, and y_{review}, obtained by withholding one in every five projection views and excluded from candidate formation throughout search. The search root state $s^{\text{fit}}_{\text{root}}$ is constructed using y_{fit} alone, with associated reconstruction volume $x(s^{\text{fit}}_{\text{root}}) = x^{\text{fit}}_{\text{root}}$. The exact view-splitting rule is given in Appendix A.2. Recall that $x(s)$ denotes the reconstruction volume associated with state s. For a realised reconstruction volume x, we define the held-out projection residual

$$r_{\text{review}}(x) = \frac{\|A_{\text{review}}x - y_{\text{review}}\|_2}{\|y_{\text{review}}\|_2}, \tag{1}$$

where A_{review} is the forward projector that simulates the X-ray projections a volume would produce at the held-out views. A smaller residual indicates that the reconstructed volume better explains measurements not used to form it. We convert this residual into a bounded utility, where larger values indicate better held-out consistency, using the root residual $r_{\text{root}} = r_{\text{review}}(x^{\text{fit}}_{\text{root}})$ as a reference:

$$U(x) = \frac{r^2_{\text{root}}}{r^2_{\text{root}} + r_{\text{review}}(x)^2}. \tag{2}$$

A volume with the same held-out residual as the root obtains $U(x) = 0.5$, while better held-out consistency yields a larger utility. This evaluation requires only acquired measurements and no ground-truth 3D reconstruction. Importantly, y_{review} is accessed only after a candidate has been formed; its feedback updates the search values and thereby influences subsequent computation allocation.

When a newly realised state s' receives utility $U(x(s'))$, the utility is backed up along the root-to-s' search path, updating the accumulated utility $W(s, a)$ and visit count $N(s, a)$ of every traversed state–action edge. The empirical value is

$$Q(s, a) = \frac{W(s, a)}{N(s, a)}. \tag{3}$$

Among actions that have already been expanded and evaluated, PG-MCTS allocates subsequent computation using the classical UCT rule Kocsis & Szepesvári (2006):

$$a^* = \arg\max_a \left[Q(s,a) + \sqrt{\frac{2\log N(s)}{N(s,a)}} \right], \tag{4}$$

where $N(s)$ denotes the visit count of state s. PG-MCTS therefore allocates computation according to held-out empirical values while retaining exploration of less-visited trajectories.

UCT can exploit only actions for which held-out feedback has already been obtained. For an unexplored action, no direct measurement evidence from the current case is available before HPCE is executed, while the expansion itself requires generating a complete 3D candidate. We therefore use a cross-case VLM prior to determine the order in which unexplored actions are tried. Given the current search state s, a LoRA-adapted VLM predicts $P_\theta(a \mid s)$ for $a \in \{B, C, F\}$. When progressive widening permits a new action at s, the highest-ranked unexecuted action under $P_\theta(a \mid s)$ is expanded through $T_{\mathrm{HPCE}}(s, a; y_{\mathrm{fit}})$. The exact progressive-widening rule and action-priority procedure are provided in Appendix A.2. Thus, progressive widening determines when a new direction can be introduced, the VLM determines which one is tried next, and UCT determines where subsequent computation is allocated after held-out evaluation. The VLM itself neither evaluates reconstruction quality nor overrides measurement evidence from the current case.

Held-out residual improvements computed on separate training cases supervise this action prior without ground truth. For a training state s and candidate action a, let $s_a = T_{\mathrm{HPCE}}(s, a; y_{\mathrm{fit}})$ denote the resulting state, and define $\Delta r_a = r_{\mathrm{review}}(x(s)) - r_{\mathrm{review}}(x(s_a))$. A larger Δr_a indicates a greater improvement in held-out measurement consistency, providing state–action preferences for LoRA adaptation Hu et al. (2022). The cases used for VLM adaptation are disjoint from both the test cases and those used to train the frozen diffusion prior. The VLM input construction, supervision target, LoRA configuration, and training details are provided in Appendix A.3.

PG-MCTS repeats expansion, held-out evaluation, and backup within a fixed HPCE expansion budget and maximum search depth D, with the experimental settings specified in Sec. 4.1. This closed-loop state–action–feedback process makes the reconstruction agentic: each subsequent search decision is adapted online according to the physical evidence observed for the current case. At termination, $Q(s,a)$ and the visit counts serve only to allocate computation during search. Instead, among all realised states, including the root state $s^{\mathrm{fit}}_{\mathrm{root}}$, we select $s^* = \arg\min_{s \in \mathcal{T}_{\mathrm{realised}}} r_{\mathrm{review}}(x(s))$ and freeze the root-to-s^* action lineage $\mathcal{L}^* = \big(a^*_0, a^*_1, \ldots, a^*_{K-1}\big)$, where $K \le D$. If $s^{\mathrm{fit}}_{\mathrm{root}}$ itself is selected, then $K = 0$ and the lineage is empty. What is committed is this action lineage rather than any search-stage volume. Sec. 3.3 re-executes the frozen lineage using all acquired measurements and applies full-measurement physics acceptance to produce the final reconstruction.

3.3 Final Reconstruction

To ensure that search-time evaluation relies on measurements not used for candidate formation, all search candidates are generated from y_{fit} alone. If a search-stage volume were returned as the final result, however, y_{review} would never contribute to the final reconstruction, and part of the acquired measurement information would remain unused. After search, we therefore do not directly use any search-stage volume and instead freeze the action lineage $\mathcal{L}^*$ determined by the search. Once the lineage is fixed, the search decisions are complete, and all acquired measurements can be used to realise the selected reconstruction path without altering the choices made during search. The held-out measurements y_{review} thus provide an evaluation signal only after candidate formation during search and are reincorporated into the full measurement set for final reconstruction. Together with the frozen lineage $\mathcal{L}^* = (a^*_0, \ldots, a^*_{K-1})$, we store the stochastic configuration and random seed associated with each selected transition as replay metadata. We construct the initial state s^{all}_0 using all acquired measurements, with $x(s^{\mathrm{all}}_0) = x^{\mathrm{all}}_{\mathrm{root}}$, following the same root-construction procedure used during search. We then replay the frozen actions in order, with $s^{\mathrm{all}}_{k+1} = T_{\mathrm{HPCE}}\big(s^{\mathrm{all}}_k, a^*_k; y_{\mathrm{all}}\big)$ for $k = 0, \ldots, K-1$, where y_{all} denotes all acquired measurements and each transition uses its stored stochastic configuration. During replay, B, C, and F are no longer compared, and no VLM prior, UCT selection, or held-out evaluation is used to reselect the path. The action order and stochastic configurations remain fixed, but each step starts from the current full-measurement state and realises

the corresponding HPCE transition using y_{all}. Search thus determines which reconstruction path to attempt, while replay realises that fixed path using all acquired measurements.

Because replay starts from the all-measurement root and every transition is realised using all acquired measurements, the states produced during replay generally differ from their search-stage counterparts. Consequently, later steps of a lineage that performed well under held-out evaluation need not continue improving physical consistency under the full measurements. We therefore apply a monotone full-measurement physics acceptance rule after each replayed step. For a volume x, we define the normalised full-measurement projection residual

$$r_{\mathrm{all}}(x) = \frac{\|A_{\mathrm{all}}x - y_{\mathrm{all}}\|_2}{\|y_{\mathrm{all}}\|_2}, \tag{5}$$

where A_{all} is the forward projector corresponding to all acquired measurements. The k-th replayed transition is accepted only if $r_{\mathrm{all}}\big(x(s^{\mathrm{all}}_{k+1})\big) \le r_{\mathrm{all}}\big(x(s^{\mathrm{all}}_{k})\big)$. If the condition holds, replay continues with the next action; otherwise, replay terminates immediately and the volume of the most recent accepted state is returned as x_{final}. If all K transitions are accepted, then $x_{\mathrm{final}} = x(s^{\mathrm{all}}_K)$; if $K = 0$, then $x_{\mathrm{final}} = x^{\mathrm{all}}_{\mathrm{root}}$. The frozen lineage specifies which evolution steps are attempted, while the full measurements determine how far along that lineage the reconstruction is ultimately accepted for the current case. No stage of the test-time procedure, from candidate formation, held-out evaluation, and lineage selection to full-measurement replay and acceptance, uses the ground-truth volume. Ground truth is used only after x_{final} has been determined, to compute offline evaluation metrics such as PSNR and SSIM.

4 Experiments

We evaluate the proposed framework on sparse-view CT (SVCT) and limited-angle CT (LACT). Our experiments address three questions: (1) whether adaptive volumetric hypothesis search improves reconstruction quality over strong baselines, and whether the gains hold across different underlying diffusion procedures; (2) whether the proposed framework generalises beyond its development setting, through cross-dataset CT evaluation without retraining and its adaptation to cross-task MRI reconstruction; and (3) how its key components contribute to performance.

4.1 Experimental Setup

Datasets. For the main CT experiments, we follow the protocol of DDS+ISCS Du et al. (2026) on the AAPM 2016 Low Dose CT Grand Challenge dataset McCollough et al. (2017), using its released pretrained 2D diffusion checkpoint without retraining and holding out subject L506 ($256 \times 256 \times 300$) for evaluation. We evaluate SVCT with 30 views sampled uniformly from $[0^\circ, 360^\circ)$ and LACT with 100 views sampled from $[0^\circ, 100^\circ]$. The VLM prior is adapted on five additional cases from the Mayo/TCIA LDCT collection Moen et al. (2021), disjoint at the patient level from both the prior-training and test subjects. To assess cross-dataset generalisation, we further evaluate on LIDC-IDRI Armato et al. (2011) without retraining or adapting any component. Pre-processing and acquisition details are given in Appendix B.1, and cross-task MRI experiments in Appendix C.1.

Methods in comparison and metrics. We compare against FDK Feldkamp et al. (1984), ADMM-TV Chung et al. (2023b), DDNM Wang et al. (2023), DDS Chung et al. (2024), DDS+ISCS Du et al. (2026), and the volumetric diffusion method DiffusionBlend++ Song et al. (2024), all run under the protocol of DDS+ISCS. Our framework is instantiated on top of DDS and DDS+ISCS, denoted DDS+Ours and DDS+ISCS+Ours. We report PSNR and SSIM on axial, coronal, and sagittal slices, together with the inter-slice consistency gap $|\Delta| = |\mathrm{SDiff}_{\mathrm{recon}} - \mathrm{SDiff}_{\mathrm{GT}}|$, where SDiff is the mean absolute difference between adjacent slices; a smaller $|\Delta|$ indicates that the reconstruction reproduces the inter-slice variation of the ground truth, rather than being discontinuous or over-smoothed.

Implementation details. PG-MCTS uses a budget of 45 HPCE expansions and a maximum depth of $D = 4$, with one in every five acquired views held out for search-time evaluation. The VLM prior is Qwen3-VL-8B-Instruct Bai et al. (2025), adapted with LoRA Hu et al. (2022). All hyperparameters are fixed across acquisition settings, datasets, and backbones, without per-case tuning, and all experiments are run on a single NVIDIA H100 GPU. HPCE settings and the computational cost are given in Appendices B.1 and C.2.

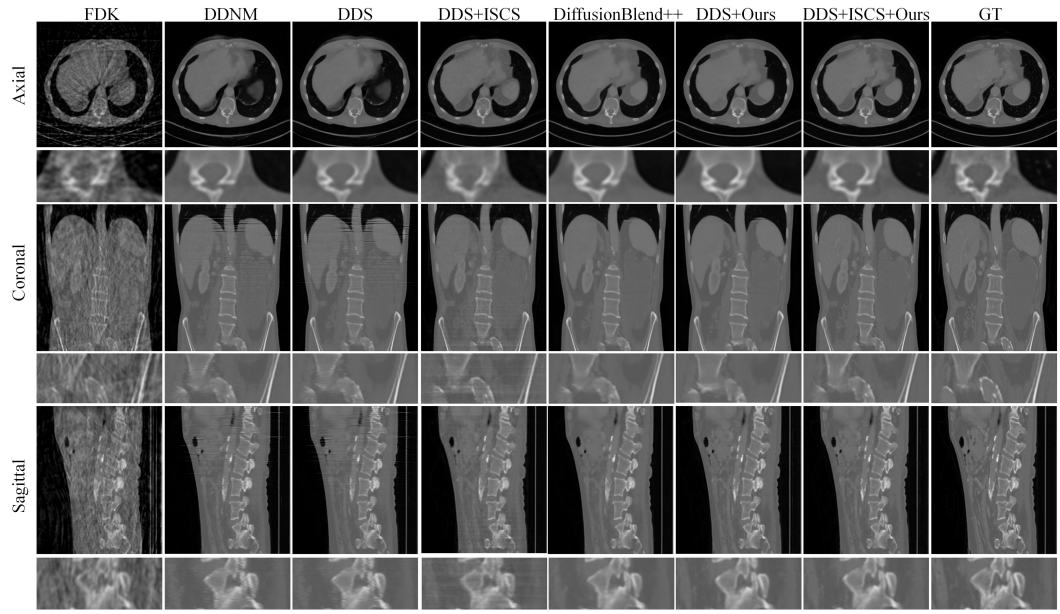

Figure 2: Qualitative comparison for SVCT with 30 views on AAPM, shown in the axial, coronal, and sagittal views.

4.2 Main Results

Sparse-view CT. Table 1 reports results for SVCT with 30 views. On AAPM, DDS+ISCS+Ours achieves the best PSNR and SSIM in all three views. Relative to DDS+ISCS, it improves PSNR by 0.61, 1.53, and 0.79 dB in the axial, coronal, and sagittal views, respectively, while reducing $|\Delta|$ from 0.001835 to 0.000581. It also exceeds DiffusionBlend++ in both PSNR and SSIM across all three views, despite operating with a frozen 2D diffusion prior rather than a dedicated 3D-aware prior. Notably, DDS+Ours improves DDS by 2.71, 2.80, and 3.36 dB and also surpasses DDS+ISCS in PSNR in all three views, even though this instantiation does not use the inter-slice stochasticity mechanism of ISCS. This indicates that the structured slice-axis exploration of HPCE, together with measurement-guided search, can improve volumetric coherence independently of ISCS. Representative AAPM reconstructions are shown in Fig. 2. **Limited-angle CT.** Table 4 reports results for LACT with 100 views over $[0^\circ, 100^\circ]$. On AAPM, DDS+ISCS+Ours improves over DDS+ISCS by 0.68, 1.37, and 0.76 dB PSNR in the axial, coronal, and sagittal views, respectively, exceeds DiffusionBlend++ in PSNR in all three views, and achieves the highest SSIM in the axial and sagittal views. It also reduces $|\Delta|$ from 0.001966 to 0.000570, the lowest value among the compared methods. Representative AAPM LACT results are shown in Fig. 5 in the appendix. **Cross-dataset generalisation.** On LIDC-IDRI, where all components are applied without retraining, the gains persist under both acquisition settings. For SVCT, DDS+ISCS+Ours improves over DDS+ISCS by 1.46, 0.91, and 1.48 dB PSNR in the three views and reduces $|\Delta|$ from 0.001978 to 0.000470. For LACT, the improvements are 1.23, 1.91, and 1.45 dB, and our method achieves the best PSNR, SSIM, and $|\Delta|$ in all views. In both settings, DDS+Ours also outperforms DDS+ISCS in PSNR in all three views. These results indicate that neither the search procedure nor the learned action prior is tied to the AAPM data, despite the change in anatomy from abdomen to chest. Qualitative results are shown in Figures 6 and 4 in the appendix.

4.3 Ablation Studies

Table 2 ablates the three components of our framework on AAPM SVCT with 30 views, using DDS+ISCS as the underlying reconstruction procedure. Search-based variants use the same number of HPCE expansions, and independent resampling is matched in total NFEs. Differences below refer to PSNR averaged over the three views.

Table 1: Quantitative comparison for sparse-view CT reconstruction (SVCT-30) on AAPM and LIDC-IDRI datasets.

Method	AAPM							LIDC-IDRI						
	Axial		Coronal		Sagittal		$\lvert\Delta\rvert$	Axial		Coronal		Sagittal		$\lvert\Delta\rvert$
	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑		PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	
FDK	23.91	0.323	23.92	0.414	23.79	0.348	0.005584	20.05	0.282	20.90	0.258	20.03	0.272	0.011187
ADMM-TV	32.94	0.882	33.67	0.895	33.72	0.893	0.001617	27.21	0.835	34.47	0.831	27.20	0.831	0.000988
DDNM	32.55	0.851	32.51	0.837	32.79	0.834	0.009342	28.61	0.791	32.61	0.785	28.37	0.808	0.018000
DDS	34.76	0.919	35.12	0.906	35.33	0.904	0.005588	27.42	0.761	31.67	0.761	27.05	0.785	0.021643
DDS+ISCS	36.97	0.937	37.75	0.944	38.16	0.942	0.001835	30.96	0.914	36.99	0.904	31.00	0.910	0.001978
DiffusionBlend++	37.54	0.941	38.93	0.943	38.19	0.932	0.001370	32.10	0.917	37.75	0.922	32.32	0.925	0.001024
DDS+Ours	37.47	0.940	37.92	0.941	38.69	0.942	0.001803	31.29	0.915	37.05	0.912	31.75	0.911	0.001363
DDS+ISCS+Ours	37.58	0.947	39.28	0.945	38.95	0.945	0.000581	32.42	0.929	37.90	0.915	32.48	0.928	0.000470

Table 2: Ablation studies on AAPM SVCT (30 views) using DDS+ISCS as the underlying reconstruction procedure.

Variant	Axial		Coronal		Sagittal		$\lvert\Delta\rvert\downarrow$
	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	
Search allocation							
Independent resampling (matched NFEs)	37.05	0.937	38.11	0.942	38.35	0.940	0.00173
Matched-budget non-adaptive search	37.25	0.941	38.73	0.942	38.56	0.944	0.00118
Fit-view evaluation instead of held-out review	37.34	0.943	38.83	0.944	38.62	0.943	0.00109
w/o VLM prior	37.43	0.945	39.02	0.944	38.78	0.944	0.00083
Hypothesis construction							
Best single action type	37.31	0.943	38.91	0.944	38.70	0.943	0.00105
w/o physics-compatible innovation	37.12	0.938	38.30	0.942	38.34	0.941	0.00135
Final reconstruction							
Return search-stage volume	37.24	0.940	38.35	0.942	38.38	0.942	0.00145
Replay w/o physics acceptance	37.53	0.945	39.20	0.945	38.86	0.944	0.00081
Full method (DDS+ISCS+Ours)	**37.58**	**0.947**	**39.28**	**0.945**	**38.95**	**0.945**	**0.00058**

Search allocation. Independent resampling with held-out selection and matched-budget non-adaptive search fall 0.77 and 0.42 dB below the full method, respectively, showing that the gain does not stem from additional computation alone. Evaluating candidates on the fit measurements instead of the held-out ones reduces performance by 0.34 dB, supporting the separation of candidate formation from evaluation, and removing the VLM prior reduces it by 0.19 dB.

Hypothesis construction. Restricting HPCE to the best single action type reduces performance by 0.30 dB, and removing physics-compatible innovation by 0.68 dB, indicating that complementary update scales and measurement-constrained innovation are both needed.

Final reconstruction. Returning the search-stage volume, formed from y_{fit} only, falls 0.61 dB below the full method, confirming the benefit of replaying the frozen lineage with all measurements. Monotone physics acceptance yields a further small gain in PSNR and the lowest $|\Delta|$, preventing replayed transitions from degrading consistency with the full measurements.

5 Conclusion

We presented an agentic, physics-grounded framework that reformulates 3D CT reconstruction as adaptive search over volumetric hypotheses. Learning prioritises where to search, held-out measurements evaluate realised hypotheses, and full measurements govern the final reconstruction through lineage replay and monotone physics acceptance. Across sparse-view and limited-angle CT, the framework consistently improves strong diffusion baselines and generalises to a different CT dataset without retraining, while cross-task MRI experiments further demonstrate its applicability beyond CT. These results support a simple principle for reconstruction with generative priors: learning can guide where to search, while measurements should decide what to trust.

AI USE STATEMENT

In this work, we used generative AI tools for language editing, including improving grammar, clarity, and readability of the manuscript. All AI-assisted edits were reviewed and verified by the authors. The authors take full responsibility for the final content of this work.

ETHICS STATEMENT

This work uses existing publicly available medical imaging datasets and does not involve the collection of new human-subject data. The proposed method is evaluated as a research reconstruction technique and is not intended or validated for clinical diagnosis or decision-making. Generative priors may introduce structures that are insufficiently supported by incomplete measurements; our framework mitigates this risk through measurement-based evaluation and final full-measurement acceptance, but does not eliminate the need for clinical validation. Future evaluation on real clinical acquisitions and broader patient populations is necessary before considering practical deployment.

REPRODUCIBILITY STATEMENT

We provide the information required to reproduce the proposed framework throughout the main paper and appendix. Sections 3.1, 3.2, and 3.3 specify the HPCE transition, VLM-prior PG-MCTS, measurement split, search value, and full-measurement replay procedure. The appendix provides additional details of the HPCE actions and physics-compatible innovation, the progressive-widening and search settings, VLM LoRA adaptation and data splits, reconstruction protocols, and experimental hyperparameters. We also report matched-budget baselines, ablation settings, and cross-task MRI details to clarify the computational and experimental comparisons. All reported quantitative results are obtained from the stated experimental procedures; ground truth is used only for post-hoc evaluation and not for search or reconstruction decisions.

REFERENCES

Samuel G. Armato, III, Geoffrey McLennan, Luc Bidaut, Michael F. McNitt-Gray, Charles R. Meyer, Anthony P. Reeves, Binsheng Zhao, Denise R. Aberle, Claudia I. Henschke, Eric A. Hoffman, et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Medical Physics*, 38(2): 915–931, 2011.

Shuai Bai, Yuxuan Cai, Ruizhe Chen, Keqin Chen, Xionghui Chen, et al. Qwen3-VL technical report. *arXiv preprint arXiv:2511.21631*, 2025.

Sayantan Bhadra, Varun A. Kelkar, Frank J. Brooks, and Mark A. Anastasio. On hallucinations in tomographic image reconstruction. *IEEE Transactions on Medical Imaging*, 40(11):3249–3260, 2021.

Hyungjin Chung, Jeongsol Kim, Michael Thompson McCann, Marc Louis Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In *The Eleventh International Conference on Learning Representations*, 2023a.

Hyungjin Chung, Dohoon Ryu, Michael T. McCann, Marc L. Klasky, and Jong Chul Ye. Solving 3d inverse problems using pre-trained 2d diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 22542–22551, 2023b.

Hyungjin Chung, Suhyeon Lee, and Jong Chul Ye. Decomposed diffusion sampler for accelerating large-scale inverse problems. In *The Twelfth International Conference on Learning Representations*, 2024.

Chenhe Du, Qing Wu, Xuanyu Tian, Jingyi Yu, Hongjiang Wei, and Yuyao Zhang. Improving 2d diffusion models for 3d medical imaging with inter-slice consistent stochasticity. In *The Fourteenth International Conference on Learning Representations*, 2026.

L. A. Feldkamp, L. C. Davis, and J. W. Kress. Practical cone-beam algorithm. *Journal of the Optical Society of America A*, 1(6):612–619, 1984.

Allard Adriaan Hendriksen, Daniël Maria Pelt, and K. Joost Batenburg. Noise2inverse: Self-supervised deep convolutional denoising for tomography. *IEEE Transactions on Computational Imaging*, 6:1320–1335, 2020.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022.

Vineet Jain, Kusha Sareen, Mohammad Pedramfar, and Siamak Ravanbakhsh. Diffusion tree sampling: Scalable inference-time alignment of diffusion models. In *Advances in Neural Information Processing Systems*, 2025.

Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. In *Advances in Neural Information Processing Systems*, volume 35, pp. 23593–23606, 2022.

Levente Kocsis and Csaba Szepesvári. Bandit based monte-carlo planning. In *Machine Learning: ECML 2006*, volume 4212 of *Lecture Notes in Computer Science*, pp. 282–293. Springer, 2006.

Nanye Ma, Shangyuan Tong, Haolin Jia, Hexiang Hu, Yu-Chuan Su, Mingda Zhang, Xuan Yang, Yandong Li, Tommi Jaakkola, Xuhui Jia, and Saining Xie. Scaling inference time compute for diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2523–2534, 2025.

Cynthia H. McCollough, Adam C. Bartley, Rickey E. Carter, Baiyu Chen, Tammy A. Drees, Phillip Edwards, David R. Holmes, III, Alice E. Huang, Farhana Khan, Shuai Leng, Kyle L. McMillan, Gregory J. Michalak, Kristina M. Nunez, Lifeng Yu, and Joel G. Fletcher. Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 low dose CT grand challenge. *Medical Physics*, 44(10):e339–e352, 2017. doi: 10.1002/mp.12345.

Taylor R. Moen, Baiyu Chen, David R. Holmes, III, Xinhui Duan, Zhicong Yu, Lifeng Yu, Shuai Leng, Joel G. Fletcher, and Cynthia H. McCollough. Low-dose CT image and projection dataset. *Medical Physics*, 48(2):902–911, 2021. doi: 10.1002/mp.14594.

Vignav Ramesh and Morteza Mardani. Test-time scaling of diffusion models via noise trajectory search. In *Advances in Neural Information Processing Systems*, 2025.

Bowen Song, Jason Hu, Zhaoxu Luo, Jeffrey A. Fessler, and Liyue Shen. Diffusionblend: Learning 3d image prior through position-aware diffusion score blending for 3d computed tomography reconstruction. In *Advances in Neural Information Processing Systems*, volume 37, pp. 89584–89611, 2024.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

Yang Song, Liyue Shen, Lei Xing, and Stefano Ermon. Solving inverse problems in medical imaging with score-based generative models. In *International Conference on Learning Representations*, 2022.

Yinhuai Wang, Jiwen Yu, and Jian Zhang. Zero-shot image restoration using denoising diffusion null-space model. In *International Conference on Learning Representations*, 2023.

George Webber and Andrew J. Reader. Diffusion models for medical image reconstruction. *BJR—Artificial Intelligence*, 1(1):ubae013, 2024.

Burhaneddin Yaman, Seyed Amir Hossein Hosseini, Steen Moeller, Jutta Ellermann, Kâmil Uğurbil, and Mehmet Akçakaya. Self-supervised learning of physics-guided reconstruction neural networks without fully sampled reference data. *Magnetic Resonance in Medicine*, 84(6):3172–3191, 2020.

A APPENDIX

A.1 HIERARCHICAL PHYSICS-COMPATIBLE EVOLUTION

HPCE procedure. Given the current reconstruction state s_k with volume $x_k = x(s_k)$ and an action $a_k \in \{B, C, F\}$, HPCE realises one complete transition

$$s_{k+1} = T_{\mathrm{HPCE}}(s_k, a_k; y_{\mathrm{fit}}). \tag{6}$$

Rather than restarting from an independent high-noise sample, HPCE re-enters the frozen diffusion process from a shallow noise level, so that alternative continuations can be generated while information already present in the current reconstruction is retained. Starting from the current reconstruction, HPCE first performs shallow diffusion re-entry using the stochastic configuration specified by the selected action. The frozen diffusion prior then evolves the resulting state along a short reverse trajectory, with full-volume data consistency enforced using y_{fit} at each step. The structural change proposed by this evolution is subsequently processed by physics-compatible innovation to produce the next full-volume hypothesis. The re-entry level, the number of reverse steps, and the number of data-consistency iterations are fixed before test-time search. The diffusion model remains frozen throughout HPCE. Measurements in y_{review} are not used during this transition and are accessed only after the resulting candidate has been realised. For transitions on the selected lineage, the stochastic configuration and corresponding random seed are stored and restored during full-measurement replay.

Hierarchical evolution actions. The three HPCE actions B, C, and F perturb the persistent stochastic representation of the underlying reconstruction procedure at global, coarse, and fine scales, respectively. Their scale semantics are shared across instantiations, whereas the stochastic variables they act on follow the underlying procedure. In the DDS+ISCS instantiation, the representation extends the spherically interpolated noise of ISCS Du et al. (2026) with hierarchical multi-scale controls: B modifies the SLERP endpoints (z_0, z_1), C modifies the coarse controls (h_2, h_4), and F modifies the fine controls (h_8, h_{16}), where the subscript denotes the resolution of the control grid along the slice axis. In the DDS instantiation, which does not use interpolated noise, B, C, and F act on the persistent per-step Gaussian stochastic trajectory through global-, coarse-, and fine-scale perturbations, respectively, using the same slice-axis control grids. In both instantiations, each action is realised through a complete HPCE transition and therefore produces a full-volume 3D reconstruction hypothesis rather than a local image update. The action configurations are fixed before test-time search and are not adapted using ground truth..

Physics-compatible innovation. Let X_{DC} denote the volume obtained after the last full-volume data-consistency update of the shallow evolution, and X_{terminal} the volume produced by the subsequent terminal diffusion update. We define the diffusion-proposed innovation as

$$\Delta_p = X_{\mathrm{terminal}} - X_{\mathrm{DC}}. \tag{7}$$

Rather than directly applying this generative update, HPCE computes a physics-compatible innovation Δ^* according to

$$\left(A_{\mathrm{fit}}^{\top} A_{\mathrm{fit}} + \rho I\right) \Delta^* = A_{\mathrm{fit}}^{\top} \left(y_{\mathrm{fit}} - A_{\mathrm{fit}} X_{\mathrm{DC}}\right) + \rho \Delta_p, \tag{8}$$

where A_{fit} denotes the forward projector associated with the fit measurements. Here $A_{\mathrm{fit}}^{\top}$ denotes the backprojection operation used by the reconstruction solver; we do not require the numerical projector–backprojector pair to form an exact adjoint pair. We solve Eq. (8) approximately with a fixed number of conjugate-gradient-type iterations. The coefficient ρ is adapted according to the measurement-domain visibility of the proposed innovation:

$$r_A = \frac{\|A_{\mathrm{fit}} \Delta_p\|_2^2}{\|\Delta_p\|_2^2 + \epsilon}, \qquad \rho = \gamma r_A, \qquad \gamma = 0.1, \tag{9}$$

where ϵ is a small numerical stabiliser, so that the regularisation strength is commensurate with the scale of $A_{\mathrm{fit}}^{\top} A_{\mathrm{fit}}$ along the proposed direction. The resulting child hypothesis is

$$X_{\mathrm{child}} = X_{\mathrm{DC}} + \Delta^*. \tag{10}$$

Thus, physics-compatible innovation constrains the diffusion-proposed change, while full-volume data consistency constrains the reconstructed volume itself. During full-measurement replay (Sec. 3.3), A_{fit} and y_{fit} in Eqs. (8) and (9) are replaced by A_{all} and y_{all}.

A.2 Additional Details of VLM-Prior PG-MCTS

Measurement split. For each acquisition, one in every five projection views is deterministically assigned to the review set, with all remaining views forming the fit set. Using zero-based indexing of the views in acquisition order, the split is

$$I_{\text{review}} = \{\, i \in I_{\text{all}} : i \bmod 5 = 0 \,\}, \qquad I_{\text{fit}} = I_{\text{all}} \setminus I_{\text{review}}. \tag{11}$$

The same rule is applied to every case and acquisition protocol. All HPCE candidate formation during search uses y_{fit} only; y_{review} is accessed only after a candidate has been realised.

Utility, backup, and termination. For the non-zero root residual encountered in our setting, $U(x) \in (0, 1]$: a candidate with the same held-out residual as the root receives $U(x) = 0.5$, improved candidates obtain $U(x) > 0.5$, and degraded ones obtain $U(x) < 0.5$. Because the utility is bounded, the classical UCT exploration term for rewards in $[0, 1]$ is used directly, without an additional tuned exploration constant. When a newly realised state s' receives utility $U(x(s'))$, every traversed edge (s, a) on the root-to-s' path is updated as

$$N(s, a) \leftarrow N(s, a) + 1, \qquad W(s, a) \leftarrow W(s, a) + U(x(s')). \tag{12}$$

Search terminates when the expansion budget is exhausted or when no admissible expansion remains under the depth and action constraints. The total expansion budget and maximum search depth are fixed before inference and reported in Sec. 4.1.

Progressive widening and action priority. All three actions are available at the root. For a non-root state with visit count N, the maximum number of opened actions is

$$m(N) = \min\left(3, \max\left(1, \left\lfloor \sqrt{N} \right\rfloor\right)\right). \tag{13}$$

When the number of opened actions is smaller than $m(N)$, a new action may be introduced; if the currently expanded subtrees are exhausted, an additional unexecuted action may also be introduced, up to the three-action limit.

The VLM determines only which unexecuted action is introduced next. We form the priority ordering

$$\pi_\theta(s) = \operatorname{argsort}^{\downarrow}_{a \in \{B, C, F\}} P_\theta(a \mid s), \tag{14}$$

where $\downarrow$ denotes sorting by decreasing predicted probability. Whenever progressive widening permits a new action, the highest-ranked action in $\pi_\theta(s)$ that has not yet been executed at state s is expanded. After an action has been realised and evaluated, subsequent allocation among expanded actions follows the UCT rule in Sec. 3.2; the VLM prior is not added to the UCT score.

A.3 VLM Prior Training

We use Qwen3-VL-8B-Instruct Bai et al. (2025) as the pretrained VLM and adapt it with LoRA to provide a cross-case prior over HPCE actions. For each state s, the VLM receives the current reconstruction together with serialised search context, including the search depth and action history. The same input construction is used for both LoRA adaptation and test-time action prioritisation.

LoRA adaptation uses five additional cases from the AAPM–Mayo LDCT collection Moen et al. (2021), reserved exclusively for VLM adaptation. These cases are disjoint from both the cases used to train the frozen diffusion prior and all test cases. For each adaptation case, projection measurements are simulated using the same forward model and acquisition geometry as in the reconstruction experiments, and are split into fit and review subsets using the same rule as at test time.

To construct supervision, we realise all three HPCE actions $a \in \{B, C, F\}$ from each training state using the fit measurements,

$$s_a = T_{\text{HPCE}}(s, a; y_{\text{fit}}), \tag{15}$$

and evaluate the corresponding improvement in held-out measurement consistency,

$$\Delta r_a = r_{\text{review}}(x(s)) - r_{\text{review}}(x(s_a)). \tag{16}$$

A larger Δr_a indicates that action a produces a candidate that is more consistent with measurements not used to form that candidate.

Rather than converting these measurements into a hard action label, we retain their relative preference through the soft target

$$q(a \mid s) = \frac{\exp(\Delta r_a/\tau)}{\sum_{b\in\{B,C,F\}} \exp(\Delta r_b/\tau)}, \tag{17}$$

where τ controls the sharpness of the action preference. The VLM is adapted by minimising the cross-entropy between this physics-derived target distribution and its predicted action distribution,

$$\mathcal{L}_{\mathrm{VLM}} = -\sum_{a\in\{B,C,F\}} q(a \mid s) \log P_\theta(a \mid s). \tag{18}$$

This supervision preserves not only which action performs best, but also the relative strength of the physical evidence, so that actions with similar held-out improvements receive similar supervision. Only the LoRA parameters are optimised; the pretrained VLM weights remain frozen.

At test time, the adapted VLM is used only to prioritise unexplored actions, as described in Appendix A.2. It does not evaluate newly realised candidates, enter the UCT value, select the committed lineage, or participate in full-measurement replay or acceptance.

A.4 Data Splits and Matched-Budget Baseline

Data splits. The cases used for diffusion-prior training, VLM adaptation, and final evaluation are disjoint at the patient level. The frozen diffusion prior is trained on the nine AAPM cases used in the original DDS+ISCS protocol, with subject L506 reserved for AAPM evaluation. For VLM adaptation, we use five additional reconstructed CT volumes from the LDCT-and-Projection-data collection Moen et al. (2021). These volumes are reserved exclusively for VLM adaptation and do not overlap with either the diffusion-prior training cases or any test case. Projection measurements for these adaptation volumes are simulated by applying the same CT forward model and acquisition geometry used in the reconstruction experiments.

For cross-dataset evaluation, we use ten volumes from LIDC-IDRI. The frozen diffusion prior, VLM prior, HPCE configuration, and search procedure are applied directly to these cases without retraining or dataset-specific adaptation. The DDS and DDS+ISCS instantiations use the same HPCE procedure and search configuration and differ only in the underlying reconstruction procedure.

Matched-budget non-adaptive search. The non-adaptive search baseline is assigned the same total number of HPCE expansions as PG-MCTS and uses the same HPCE action definitions and maximum search depth. Its computation allocation is fixed in advance and does not adapt subsequent expansions according to held-out measurement feedback. This comparison isolates the effect of adaptive test-time computation allocation under the same candidate-generation budget.

B Additional Experimental Details

B.1 Experimental Protocol

Data and pre-processing. The frozen CT diffusion prior is the released DDS+ISCS checkpoint Du et al. (2026), a variance-exploding diffusion model Song et al. (2021) trained on 5410 slices from nine AAPM subjects. HU values are clipped to $[-1000, 1600]$ and linearly rescaled to $[0, 1]$, and all slices are resized to 256×256. LIDC-IDRI volumes are pre-processed in the same way.

Acquisition geometry. Following DDS+ISCS, projections are simulated with a cone-beam geometry using torch-radon, with a 512×512 detector of 2 mm spacing, a source-to-isocentre distance of 700 mm, and an isocentre-to-detector distance of 500 mm. Under the one-in-five split, SVCT yields 24 fit and 6 review views, and LACT yields 80 fit and 20 review views.

Table 3: Quantitative results for $5\times$ isotropic MRI super-resolution on IXI dataset.

	Axial		Coronal		Sagittal	
Method	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑
Cubic	32.9759	0.9124	34.6093	0.9120	32.4036	0.9073
DDNM	34.5594	0.9249	35.9694	0.9152	33.1219	0.9077
DDS	34.0431	0.9156	35.6468	0.9205	33.2739	0.9162
DDS+ISCS	35.0997	0.9310	36.0328	0.9216	34.0657	**0.9291**
DiffusionBlend++	35.2294	0.9267	36.3692	0.9253	34.4210	0.9220
DDS+Ours	35.2078	**0.9415**	36.3390	0.9208	34.1242	0.9127
DDS+ISCS+Ours	**35.6241**	0.9258	**36.8244**	0.9226	**34.7257**	0.9262

Evaluation metrics. Following DDS+ISCS, the slice difference is defined as

$$\mathrm{SDiff} = \frac{1}{S-1}\sum_{i=1}^{S-1} \mathrm{Mean}\left(|x_{i+1} - x_i|\right), \tag{19}$$

where S is the number of slices and x_i is the i-th slice, and we report $|\Delta| = |\mathrm{SDiff}_{\mathrm{recon}} - \mathrm{SDiff}_{\mathrm{GT}}|$.

HPCE settings. HPCE uses a shallow re-entry level of $\sigma_s = 0.069099$ and 10 reverse diffusion steps, with 10 CG-type iterations for each full-volume data-consistency update and another 10 CG-type iterations for physics-compatible innovation.

C Additional Experimental Results

Unless otherwise stated, all CT experiments use the same frozen diffusion checkpoint, HPCE configuration, and search settings as in the main experiments. Ground truth is used only for post-hoc evaluation.

C.1 Cross-Task MRI Experiments

We further examine whether the proposed framework can be adapted beyond CT. Following the MRI reconstruction protocol of DDS+ISCS Du et al. (2026), we consider isotropic MRI super-resolution on T1-weighted brain scans from the IXI dataset. The evaluation volume has size $256 \times 256 \times 150$ with a voxel spacing of $1.2 \times 0.9375 \times 0.9375\,\mathrm{mm}^3$. It is resampled to an isotropic resolution of $1\,\mathrm{mm}^3$ and padded to $256 \times 256 \times 256$, while anisotropic acquisition is simulated by $5\times$ downsampling along the z-axis. We evaluate on subject 196.

This experiment uses a frozen diffusion prior trained specifically for MRI, rather than transferring the CT diffusion prior across modalities. The VLM prior is adapted with LoRA on 5 IXI cases that are disjoint from the training and test cases. The same overall reconstruction framework is retained, including HPCE, VLM-prior PG-MCTS, held-out measurement evaluation, and full-measurement replay, with the CT projection operator replaced by the MRI z-axis downsampling operator. We compare against cubic interpolation, DDNM, DDS, DDS+ISCS, and DiffusionBlend++, and report PSNR and SSIM on the axial, coronal, and sagittal views.

As shown in Table 3, the proposed framework transfers effectively to the MRI reconstruction setting. Applied to DDS+ISCS, our method improves PSNR by 0.52, 0.79, and 0.66 dB in the axial, coronal, and sagittal views, respectively, achieving the highest PSNR in all three orientations. It also exceeds DiffusionBlend++ by 0.39, 0.46, and 0.30 dB, respectively. Averaged over the three views, DDS+ISCS+Ours improves PSNR from 35.07 to 35.72 dB over DDS+ISCS and from 35.34 to 35.72 dB over DiffusionBlend++.

The improvement is also observed when the framework is instantiated with DDS alone: DDS+Ours improves DDS by 1.16, 0.69, and 0.85 dB across the three views. SSIM results are more mixed across orientations, indicating that the principal gain in this cross-task setting is in reconstruction fidelity rather than uniformly improving every structural metric. Overall, the MRI experiment shows that the proposed hypothesis-search framework can be adapted to a different inverse problem when paired with a task-specific frozen diffusion prior and the corresponding physical forward operator.

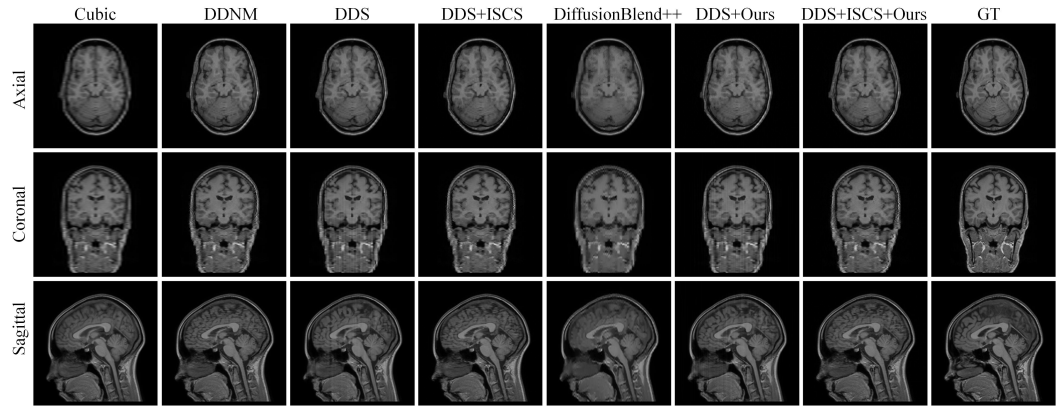

Figure 3: Qualitative comparison for $5\times$ isotropic MRI super-resolution on the IXI dataset, shown in the axial, coronal, and sagittal views.

Table 4: Quantitative comparison for limited-angle CT reconstruction (LACT-100) on AAPM and LIDC-IDRI datasets.

Method	AAPM							LIDC-IDRI						
	Axial		Coronal		Sagittal		$\lvert\Delta\rvert$	Axial		Coronal		Sagittal		$\lvert\Delta\rvert$
	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑		PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	
FDK	15.85	.381	16.30	.453	16.71	.413	.000898	16.62	.295	17.07	.289	16.90	.289	.001854
ADMM-TV	27.27	.795	27.85	.794	27.83	.799	.000877	24.79	.759	29.80	.756	25.60	.760	.001230
DDNM	28.40	.854	28.75	.774	28.22	.775	.016443	23.66	.699	27.69	.657	24.49	.680	.029668
DDS	29.07	.885	29.93	.829	29.20	.830	.011592	23.46	.703	27.60	.665	24.19	.686	.030492
DDS+ISCS	31.65	.911	32.90	.917	32.49	.920	.001966	25.34	.745	31.82	.837	27.25	.821	.001502
DiffusionBlend++	31.96	.918	33.27	.925	32.94	.930	.001619	26.39	.854	33.28	.849	28.41	.840	.001374
DDS+Ours	31.78	0.906	33.15	0.919	32.73	0.921	.001918	26.32	0.842	32.46	0.847	28.31	0.839	.001399
DDS+ISCS+Ours	32.33	0.921	34.27	0.923	33.25	0.934	0.000570	26.57	0.861	33.73	0.850	28.70	0.845	0.000503

C.2 Computational Cost

On a single NVIDIA H100 GPU, DDS and DDS+ISCS each reconstruct a volume in about 5 minutes, whereas DDS+Ours and DDS+ISCS+Ours each take about 10 minutes. Our implementation evaluates independent HPCE candidates in parallel, which reduces the runtime from about 35 minutes with sequential evaluation without changing the search results.

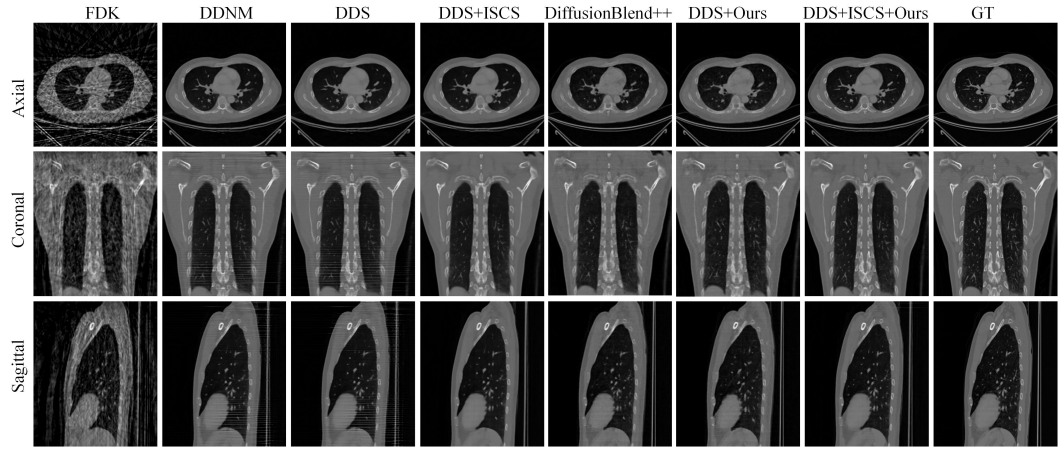

Figure 4: Qualitative comparison for SVCT with 30 views on LIDC-IDRI, shown in the axial, coronal, and sagittal views.

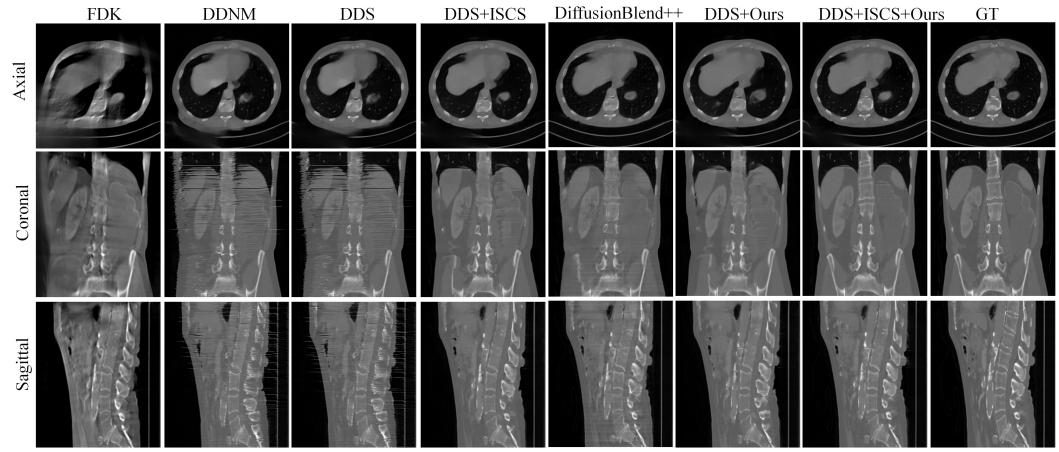

Figure 5: Qualitative comparison for LACT with 100 views over $[0°, 100°]$ on AAPM, shown in the axial, coronal, and sagittal views.

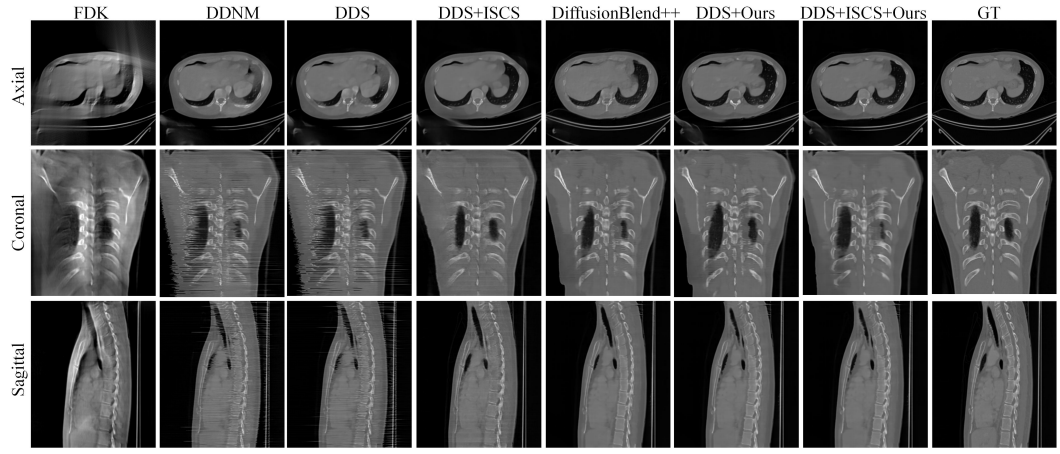

Figure 6: Qualitative comparison for LACT with 100 views over $[0°, 100°]$ on LIDC-IDRI, shown in the axial, coronal, and sagittal views.