FUNCTION-SPACE DECOUPLED DIFFUSION FOR FORWARD AND INVERSE MODELING IN CARBON CAPTURE AND STORAGE

Xin Ju[1] **Jiachen Yao**[2] **Anima Anandkumar**[2] **Sally M. Benson**[1] **Gege Wen**[3]

[1]Stanford University [2]California Institute of Technology
[3]Imperial College London

ABSTRACT

Accurate characterization of subsurface flow is critical for Carbon Capture and Storage (CCS) but remains challenged by the ill-posed nature of inverse problems with sparse observations. We present Function-space *Decoupled* Diffusion Posterior Sampling (Fun-DDPS), a generative framework that combines function-space diffusion models with differentiable neural operator surrogates for both forward and inverse modeling. Our approach learns a prior distribution over geological parameters (geomodel) using a single-channel diffusion model, then leverages a Local Neural Operator (LNO) surrogate to provide physics-consistent guidance for cross-field conditioning on the dynamics field. This decoupling allows the diffusion prior to robustly recover missing information in parameter space, while the surrogate provides efficient gradient-based guidance for data assimilation. We demonstrate Fun-DDPS on synthetic CCS modeling datasets, achieving two key results: (1) For forward modeling with only 25% observations, Fun-DDPS achieves 7.7% relative error compared to 86.9% for standard surrogates (an 11× improvement), proving its capability to handle extreme data sparsity where deterministic methods fail. (2) We provide the first rigorous validation of diffusion-based inverse solvers against asymptotically exact Rejection Sampling (RS) posteriors. Both Fun-DDPS and the joint-state baseline (Fun-DPS) achieve Jensen-Shannon divergence < 0.06 against the ground truth. Notably, Fun-DDPS produces physically consistent realizations free from the high-frequency artifacts observed in joint-state baselines, achieving this with 4× improved sample efficiency compared to rejection sampling. Code is available at `https://github.com/neuraloperator/Fun-DDPS`.

1 INTRODUCTION

Carbon capture and storage (CCS) is a critical technology for mitigating anthropogenic climate change (Pacala & Socolow, 2004; Agency, 2020). The safety and efficacy of gigaton-scale storage depend on two complementary computational tasks: *forward modeling*, to forecast CO_2 plume migration and pressure buildup, and *inverse modeling*, to characterize subsurface geological heterogeneity from sparse monitoring data. However, robust uncertainty quantification in this domain faces a fundamental bottleneck: subsurface parameters are high-dimensional and non-Gaussian, while the governing multiphase flow equations are computationally expensive to solve.

Current data assimilation (DA) methods struggle to effectively tackle these challenges. Traditional ensemble-based methods, such as the Ensemble Kalman Filter (EnKF) and Ensemble Smoother (ES-MDA), remain the industry standard but rely on Gaussian assumptions that fail to capture complex geological features such as discrete facies or channelized reservoirs Nejadi et al. (2012); Emerick & Reynolds (2013). Conversely, rigorous Bayesian sampling methods including Markov Chain Monte Carlo (MCMC) Evensen (2003; 2004); Vrugt et al. (2013) avoid these assumptions but are prohibitively expensive, often requiring thousands of high-fidelity simulations that are infeasible for large-scale 3D models (Oliver et al., 1997). While deep learning surrogates, such as Fourier Neural Operators (FNO), have successfully accelerated forward simulations by orders of magnitude (Li

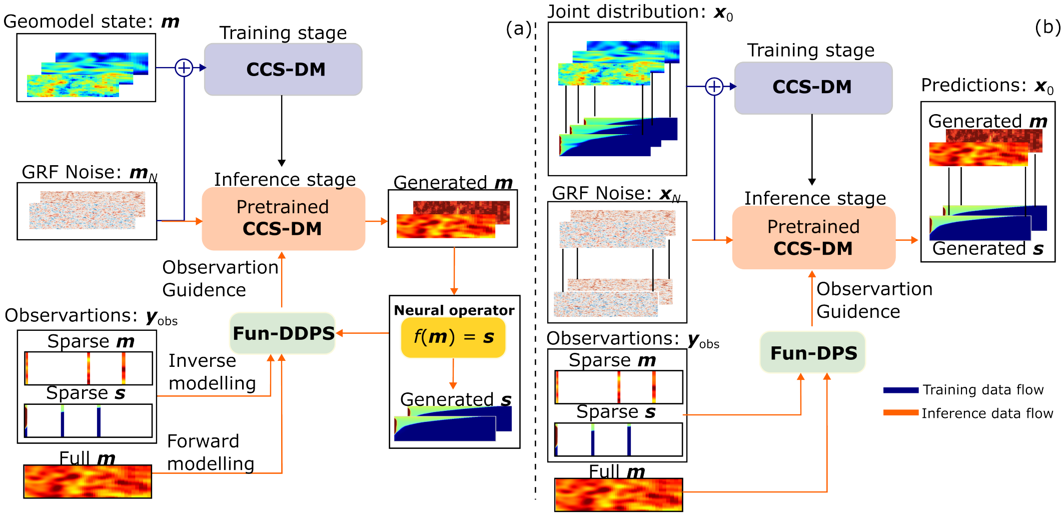

Figure 1: Comparison of decoupled (Fun-DDPS) vs. joint-state (Fun-DPS) architectures. **(a) Fun-DDPS (Decoupled):** Training uses only geomodel samples $\boldsymbol{m}$ to learn the prior $p(\boldsymbol{m})$; a separate neural operator surrogate $\mathcal{L}_\phi$ learns the forward mapping. During inference, the geomodel diffusion model generates $\boldsymbol{m}$, and $\mathcal{L}_\phi$ maps it to dynamics $\boldsymbol{s}$. For inverse problems, dynamics observations guide the diffusion *through* the surrogate gradient. **(b) Fun-DPS (Joint-state):** Training requires paired data $(\boldsymbol{m}, \boldsymbol{s})$ to learn the joint distribution $p(\boldsymbol{m}, \boldsymbol{s})$. During inference, the joint model generates both fields simultaneously, with observations guiding the corresponding channels directly.

et al., 2020; Wen et al., 2022), they are typically deterministic and do not inherently address the ill-posed inverse problem directly.

To bridge this gap, generative diffusion models have emerged as powerful priors for scientific inverse problems (Song et al., 2020; Chung et al., 2022; Yang et al., 2023; Yao et al., 2025). However, applying them to CCS encounters a critical limitation in existing *joint-state* architectures, which learn the joint distribution of geological parameters and dynamic states $p(\boldsymbol{m}, \boldsymbol{s})$ together. We observe that joint training often leads to physical inconsistency, as the model learns statistical correlations rather than explicit physical laws, especially when paired training data is limited (Lin et al., 2026).

In this work, we propose **Fun-DDPS** (Function-space *Decoupled* Diffusion Posterior Sampling), a novel framework that decouples the learning of geological priors from the approximation of flow physics. Our approach learns a prior distribution over geological parameters $p(\boldsymbol{m})$ using a single-channel function-space diffusion model, and independently trains a neural operator surrogate $\mathcal{L}_\phi$ to approximate the forward physics. During inference, we employ the differentiable neural surrogate to backpropagate gradients from sparse observations directly into the diffusion generation process. This decoupled architecture offers three key contributions:

1. **Robust Forward Modeling:** By leveraging the generative prior to reconstruct missing data, Fun-DDPS achieves 7.7% relative error on forward tasks with 25% data coverage, compared to 86.9% error for standard surrogates.
2. **Physics-Consistent Inversion:** We demonstrate that guiding the diffusion process via a physics-based surrogate eliminates the high-frequency artifacts common in joint-state models, producing geologically realistic posterior samples even with limited observations.
3. **Rigorous Validation:** We provide the first rigorous benchmark of diffusion-based inversion against asymptotically exact Rejection Sampling (RS) posteriors, achieving high statistical accuracy (JS divergence < 0.06) with a $4\times$ reduction in computational cost.

2 RELATED WORK

Inverse Problems in Subsurface Systems. The inverse problem in hydrology and reservoir engineering is classically addressed via variational or ensemble methods. Gradient-based variational approaches (Lewis et al., 2006) are efficient but often require adjoint code, which is unavailable for many black-box simulators. Derivative-free methods such as ES-MDA (Emerick & Reynolds, 2013) are widely used but suffer from the "Gaussian limitation," often smoothing out sharp geological interfaces critical for flow connectivity. While recent work has integrated deep generative models (GANs, VAEs) into ensemble smoothers to parameterize non-Gaussian priors (Misra et al., 2023; Forghani et al., 2022; Misra et al., 2024; Teng et al., 2025), these approaches typically rely on low-dimensional latent spaces that may limit the expressivity of the recovered fields.

Deep-Learning-Based Surrogates for Inverse Problems. Recent advances have introduced various deep learning-based (DL) strategies to address the data assimilation challenges. One popular approach involves using DL models as fast surrogates to replace conventional reservoir simulators for forward simulation. These surrogates, including Recurrent Residual U-Nets, Fourier neural operators (FNO), and graph neural networks (GNN), effectively reduce forward simulation costs Tang et al. (2020; 2022); Wen et al. (2023; 2022); Ju et al. (2024). This speedup makes many computationally intensive DA methods feasible. DL surrogates can facilitate Bayesian methods that treat geological hyperparameters (e.g., mean permeability, correlation lengths) as uncertain variables for CCS applications. However, even with fast surrogates, data assimilation remains within conventional frameworks where ESMDA is used to calibrate reservoir parameters, meaning the reliance on Gaussian assumptions in the inversion step is not fully resolved.

Diffusion Models for Inverse Problems. Denoising Diffusion Probabilistic Models (DDPMs) (Ho et al., 2020) currently define the state-of-the-art in generative modeling. For inverse problems, methods such as *Diffusion Posterior Sampling* (DPS) (Chung et al., 2022; Huang et al., 2024) allow for conditional generation by using the gradient of a forward operator to guide the reverse diffusion process. In the scientific domain, this has been extended to infinite-dimensional function spaces (Lim et al., 2023; Kovachki et al., 2021). Most relevant to our work is the *Decoupled Diffusion Inverse Solver* (DDIS) (Lin et al., 2026), which showed the failure modes of joint-state training. We extend this decoupled idea to function-space CCS modeling, specifically addressing the extreme sparsity of dynamic monitoring data ($< 1\%$ coverage) and providing a rigorous quantitative benchmark against the ground truth posterior obtained via Rejection Sampling.

3 METHODOLOGY

We propose **Fun-DDPS** (schematically illustrated in Figure 1), a framework that integrates function-space generative priors with differentiable neural operators to solve subsurface inverse problems. We first formulate the Bayesian inverse problem (Section 3.1), then present the function-space diffusion prior (Section 3.2) and decoupled posterior sampling (Section 3.3), and finally describe the CCS dataset and neural architectures (Section 3.4). Table 1 summarizes the key differences.

3.1 PROBLEM FORMULATION

Let $\boldsymbol{m} \in \mathcal{M}$ denote the static geological parameters (e.g., permeability) and $\boldsymbol{s} \in \mathcal{S}$ denote the dynamic state variables (e.g., saturation), defined on Hilbert spaces of functions. The physics is governed by a forward operator $F : \mathcal{M} \to \mathcal{S}$. We seek to infer the unknown geomodel $\boldsymbol{m}$ given sparse observations $\boldsymbol{y}_{obs}$. The observation process is modeled as:

$$\boldsymbol{y}_{obs} = \mathcal{H}(\boldsymbol{m}, F(\boldsymbol{m})) + \boldsymbol{\eta}, \quad \boldsymbol{\eta} \sim \mathcal{N}(0, \sigma_y^2 \mathbf{I}) \tag{1}$$

where $\mathcal{H}$ is a sparse sampling operator. In the Bayesian setting, we target the posterior distribution:

$$p(\boldsymbol{m}|\boldsymbol{y}_{obs}) \propto p(\boldsymbol{y}_{obs}|\boldsymbol{m}) p(\boldsymbol{m}), \tag{2}$$

	Fun-DDPS	**Fun-DPS**
Components	Diffusion prior $p(\boldsymbol{m})$; LNO surrogate $\mathcal{L}_\phi \approx F$.	Joint diffusion prior $p(\boldsymbol{m}, \boldsymbol{s})$.
Training data		
Paired $(\boldsymbol{m}, \boldsymbol{s})$	For surrogate only; fewer needed.	Required.
Parameter only $(\boldsymbol{m})$	For prior.	Not used.
Output	$\boldsymbol{m}$ from diffusion; $\boldsymbol{s} = \mathcal{L}_\phi(\boldsymbol{m})$.	$(\boldsymbol{m}, \boldsymbol{s})$ from joint model.
Phys. enforcement	Explicit.	Implicit.
Guid. attenuation	Avoided.	Occurs under data scarcity.

Table 1: Comparison of Fun-DDPS (decoupled) and Fun-DPS (joint-state).

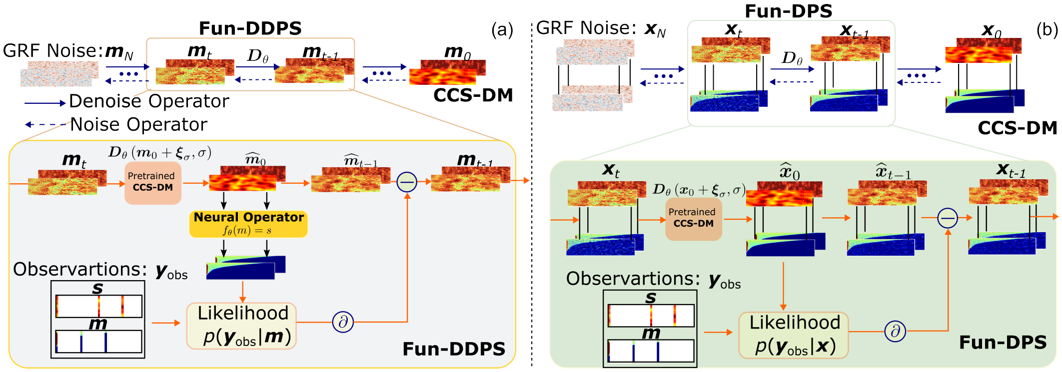

Figure 2: Posterior sampling comparison: Fun-DDPS vs. Fun-DPS. **(a) Fun-DDPS:** The geomodel diffusion model denoises a GRF to produce $\hat{\boldsymbol{m}}_0$. Geomodel observations provide direct guidance; dynamics observations guide via the surrogate gradient $\nabla_{\boldsymbol{m}}\|\mathcal{L}_\phi(\hat{\boldsymbol{m}}_0) - \boldsymbol{y}_{dyn}\|$, translating sparse solution-space constraints into dense parameter-space guidance. **(b) Fun-DPS:** The joint-state model denoises to produce $(\hat{\boldsymbol{m}}_0, \hat{\boldsymbol{s}}_0)$ simultaneously. Observations guide the respective channels directly, but sparse dynamics observations remain localized without the surrogate's global receptive field to propagate information.

where $p(\boldsymbol{m})$ denotes the Radon–Nikodym derivative of the prior measure with respect to a Gaussian reference measure on $\mathcal{M}$. Our decoupled approach addresses this by learning a prior $p(\boldsymbol{m})$ independently and using a pre-trained neural operator surrogate to efficiently approximate the likelihood $p(\boldsymbol{y}_{obs}|\boldsymbol{m})$ during posterior sampling, allowing for robust posterior sampling even under extreme data sparsity.

3.2 FUNCTION-SPACE DIFFUSION PRIOR

We model the geological prior $p(\boldsymbol{m})$ using a Function-Space Diffusion Model (Lim et al., 2023).

Forward Process. Unlike finite-dimensional models, we define the forward process by perturbing $\boldsymbol{m}_0$ with Gaussian Random Fields (GRFs):

$$\boldsymbol{m}_\sigma = \boldsymbol{m}_0 + \boldsymbol{\xi}_\sigma, \quad \boldsymbol{\xi}_\sigma \sim \mathcal{N}(0, \sigma^2 \mathcal{C}_\gamma) \tag{3}$$

where $\mathcal{C}_\gamma$ is a covariance operator defined by a Matérn kernel and σ is the noise level. This ensures that the noise respects the spatial continuity of geological fields, preserving discretization invariance. As σ increases, the perturbed field $\boldsymbol{m}_\sigma$ gradually loses structure and approaches the base measure.

Training Objective. We train a time-dependent denoising neural operator $\boldsymbol{D}_\theta(\boldsymbol{m}_\sigma, \sigma)$ to estimate the clean signal $\boldsymbol{m}_0$ from a noisy input $\boldsymbol{m}_\sigma$. The network minimizes the score matching objective:

$$\mathcal{L}(\theta) = \mathbb{E}_{\boldsymbol{m}_0,\sigma,\boldsymbol{\xi}}\left[\lambda(\sigma)\|\boldsymbol{D}_\theta(\boldsymbol{m}_0 + \boldsymbol{\xi}_\sigma, \sigma) - \boldsymbol{m}_0\|^2_{\mathcal{M}}\right] \tag{4}$$

where we choose $\|\cdot\|_{\mathcal{M}}$ to be L_2 norm, and $\lambda(\sigma)$ is a noise-level dependent weighting function. Once trained, unconditional sampling from the prior $p(\boldsymbol{m})$ is performed via the probability flow ODE:

$$d\boldsymbol{m}_\sigma = -\sigma \nabla_{\boldsymbol{m}} \log p_\sigma(\boldsymbol{m}_\sigma)\, d\sigma \tag{5}$$

where $\nabla_{\boldsymbol{m}} \log p_\sigma(\boldsymbol{m}_\sigma)$ is the score function, approximated via $\boldsymbol{D}_\theta$. We provide additional diagnostics of the learned geomodel prior (e.g., variograms and two-point statistics) in Appendix D.

3.3 FUN-DDPS: DECOUPLED POSTERIOR SAMPLING

To sample from the posterior, we follow Fun-DPS (Yao et al., 2025), which guides the reverse diffusion process in function space using the gradient of the log-likelihood. Appendix A provides unified pseudocode and implementation details for our posterior sampler. The conditional score decomposes into the prior score (provided by $\boldsymbol{D}_\theta$) and the likelihood guidance:

$$\nabla_{\boldsymbol{m}} \log p(\boldsymbol{m}|\boldsymbol{y}_{obs}) = \overbrace{\nabla_{\boldsymbol{m}} \log p(\boldsymbol{m})}^{\text{Prior (via } \boldsymbol{D}_\theta)} + \overbrace{\nabla_{\boldsymbol{m}} \log p(\boldsymbol{y}_{obs}|\boldsymbol{m})}^{\text{Likelihood guidance}}, \tag{6}$$

where $\nabla_{\boldsymbol{m}}$ denotes the Fréchet derivative of the corresponding functional. The form of the likelihood guidance depends on what is observed.

Forward modeling. Observations are of the geomodel itself, e.g. $\boldsymbol{y}_{geo} = M_{geo} \odot \boldsymbol{m} + \boldsymbol{\eta}$, so the likelihood gradient is computed directly and no surrogate is needed:

$$\nabla_{\boldsymbol{m}} \log p(\boldsymbol{y}_{geo}|\hat{\boldsymbol{m}}_0) \approx -\zeta_{geo} \nabla_{\boldsymbol{m}} \|M_{geo} \odot (\hat{\boldsymbol{m}}_0 - \boldsymbol{y}_{geo})\|_2^2 \tag{7}$$

Inverse modeling. Observations are of the dynamic state, $\boldsymbol{y}_{dyn} = M_{dyn} \odot F(\boldsymbol{m}) + \boldsymbol{\eta}$, so computing the likelihood gradient requires differentiating through the forward map $F(\boldsymbol{m})$. Since F is expensive, we use a pre-trained neural operator surrogate $\mathcal{L}_\phi \approx F$ and backpropagate through $\mathcal{L}_\phi$:

$$\nabla_{\boldsymbol{m}} \log p(\boldsymbol{y}_{dyn}|\hat{\boldsymbol{m}}_0) \approx -\zeta_{dyn} \nabla_{\boldsymbol{m}} \|M_{dyn} \odot (\mathcal{L}_\phi(\hat{\boldsymbol{m}}_0) - \boldsymbol{y}_{dyn})\|_2^2 \tag{8}$$

Eq. 8 represents the core innovation of Fun-DDPS. By separating the prior $p(\boldsymbol{m})$ from the physics $\mathcal{L}_\phi$, generated samples remain on the learned geological manifold while data consistency is enforced through surrogate guidance. This mechanism is visualized in Figure 2, which highlights how Fun-DDPS translates sparse constraints in the solution space into dense guidance in the parameter space via the surrogate's Jacobian. This contrasts with some joint-state approaches (e.g. Fun-DAPS in Lin et al. (2026)), where gradients from sparse observations remain localized to the specific channels, often resulting in high-frequency artifacts and physical inconsistencies when training data is limited.

Theoretical Motivation. Recent analysis by Lin et al. (2026) shows that joint-state models suffer from *guidance attenuation* under data scarcity: the gradient signal that updates coefficients vanishes when training data is limited. This occurs because effective guidance requires the diffusion state to lie near multiple training samples simultaneously, a condition that becomes exponentially rare in high dimensions. In contrast, the decoupled design avoids this failure by using the explicit neural operator Jacobian $\nabla_{\boldsymbol{m}}\mathcal{L}_\phi(\boldsymbol{m})$ for guidance, which does not depend on training data density.

3.4 CCS DATASET AND NEURAL ARCHITECTURES

3.4.1 DATASET GENERATION

To validate our framework, we simulate the injection of supercritical CO_2 into a radially symmetrical deep saline aquifer. We generate a dataset of 12,000 training pairs and 1,500 test pairs using the industry-standard simulator ECLIPSE (e300) (Schlumberger, 2014). Each realization consists of a heterogeneous permeability field $\boldsymbol{m}$ generated via SGeMS (Remy et al., 2009b) and the corresponding CO_2 saturation field $\boldsymbol{s}$ after 30 years of injection. The governing multiphase flow equations, numerical grid settings, and prior distributions for geological parameters are detailed in Appendix E.

3.4.2 SURROGATE ARCHITECTURE (LOCALNO)

For the forward surrogate $\mathcal{L}_\phi$, we employ a Local Neural Operator (LNO) (Liu-Schiaffini et al., 2024) to learn the mapping from the geomodel $\boldsymbol{m}$ to the dynamic state $\boldsymbol{s}$. The training set consists of 12,000 pairs $\{(\boldsymbol{m}_i, \boldsymbol{s}_i)\}$. While standard operators such as FNO (Li et al., 2020) excel at capturing global dependencies, they often produce ringing artifacts around sharp discontinuities. The LNO addresses this by combining global Fourier spectral layers with localized kernels implemented via Discrete Continuous Convolutions (DISCO). This dual-path design is critical for our application: the Fourier path resolves global pressure responses, while the DISCO path accurately captures the shock-like saturation fronts characteristic of multiphase flow. We note that the surrogate is relatively robust to train even when paired data are limited, and our sampling pipeline is tolerant to surrogate approximation error.

3.4.3 DIFFUSION BACKBONE (U-NO)

Both the decoupled geomodel prior (Fun-DDPS) and the joint-state baseline (Fun-DPS) utilize a U-shaped Neural Operator (U-NO) architecture (Rahman et al., 2022; Yao et al., 2025). The U-NO integrates neural operator blocks into a multi-scale U-Net structure, ensuring discretization invariance while capturing features at varying spatial resolutions. The same 12,000 training samples are used to train both diffusion models, though the U-NO in Fun-DDPS is trained solely on the geomodel fields $\{\boldsymbol{m}_i\}$, whereas Fun-DPS is trained on the joint input–output pairs. To ensure a rigorous comparison, we fix the architectural hyperparameters to be identical for both methods (See

Appendix E.4). We also examine the quality of the learned geomodel priors both qualitatively and quantitatively in Appendix D.

4 EXPERIMENTS

We evaluate the proposed Fun-DDPS framework on two distinct tasks: (1) forward modeling conditioned on partial static inputs (geomodel), and (2) inverse modeling conditioned on partial dynamic observations (saturation), as schematically shown in Fig. 3. In practical deployment, the observation sparsity ratio is usually unknown a priori and can vary across settings. To reflect this, our training pipeline is agnostic to the observation-mask ratio, and sparsity is imposed only at inference time for evaluation. For the forward problem, we benchmark performance against two baselines: **Fun-DPS:** A joint-state diffusion baseline; and **Surrogate ($\mathcal{L}_\phi$):** A deterministic neural operator trained on complete inputs. For the inverse problem, we provide rigorous validation by comparing our approximate posteriors against a "ground truth" posterior distribution obtained via Rejection Sampling (RS) (Robert et al., 1999).

4.1 FORWARD PROBLEM: PARTIAL GEOMODEL OBSERVATIONS

4.1.1 PROBLEM SETUP

In this task, the goal is to predict the dynamic saturation field $\boldsymbol{s}$ given a partially observed geomodel $\boldsymbol{y}_{geo}$. We evaluate performance across three observation sparsity levels: 100% (full), 50% (random), and 25% (random) coverage. The forward/inverse task setup is shown in Figure 3.

The primary challenge here is that deterministic surrogates ($\mathcal{L}_\phi$) require dense, complete inputs. Standard approaches, such as zero-filling, push the input out of the training distribution, typically leading to severe prediction errors. Fun-DDPS addresses this by utilizing the learned diffusion prior to probabilistically reconstruct missing geomodel features *before* the forward simulation.

For evaluation, we use a test set of 480 samples. We run the diffusion process for 500 iterations with a guidance weight of $\zeta_{geo} = 10{,}000$ to ensure strong consistency with the observed data points.

4.1.2 RESULTS

Table 2 summarizes the relative L_2 errors. At full observation (100%), Fun-DDPS methods perform comparably to the neural surrogate. However, as sparsity increases, the deterministic surrogate degrades catastrophically—rising from 4.4% to 86.9% error—confirming that zero-filling is insufficient for complex physical mappings. In contrast, Fun-DDPS exhibits remarkable robustness, maintaining a low 7.7% error even at 25% coverage. This represents an 11× improvement over the surrogate. By leveraging the generative prior, Fun-DDPS reconstructs a physically plausible full-field geomodel, allowing the underlying operator to function within its valid input domain.

Qualitative results are visualized in Figure 4. While the surrogate produces noisy, incoherent predictions in the presence of gaps, Fun-DDPS maintains coherent spatial structures, particularly capturing the sharp gradients at the CO_2 plume front. Notably, the Fun-DPS baseline shows consistently high error (34–42%) regardless of sparsity. This performance gap highlights a fundamental advantage of the proposed decoupled architecture over joint-state diffusion models. While joint-state models primarily encode geomodel–saturation coupling as statistical correlation in the joint distribution and enforce PDE only at inference time via a physics loss, they can both miss physical laws under limited supervision and have higher sampling cost from physics-loss evaluations.

In contrast, Fun-DDPS decouples the generation and simulation processes: it employs a pre-trained neural operator to strictly enforce physical consistency. Because the operator is trained exclusively on the mapping task, it learns to honor the underlying physics much more effectively than a genera-

Obs. Coverage	Fun-DDPS	Fun-DPS	Surrogate
100%	0.046 ± 0.058	0.418 ± 0.366	**0.044 ± 0.057**
50%	**0.054 ± 0.069**	0.370 ± 0.357	0.850 ± 0.263
25%	**0.077 ± 0.079**	0.336 ± 0.349	0.869 ± 0.258

Table 2: Forward problem: Relative L_2 error (mean ± std). Fun-DDPS maintains robustness under high sparsity, whereas the deterministic surrogate fails due to input artifacts.

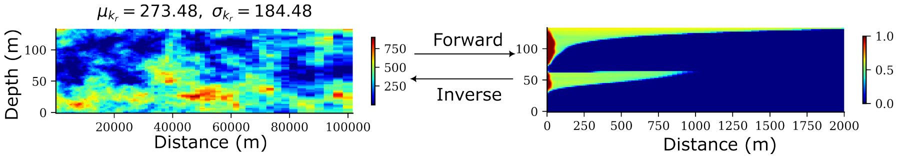

Figure 3: Problem setup. **Forward:** The geomodel is partially observed; the full dynamics are predicted. **Inverse:** The dynamics are partially observed; the full geomodel is inferred.

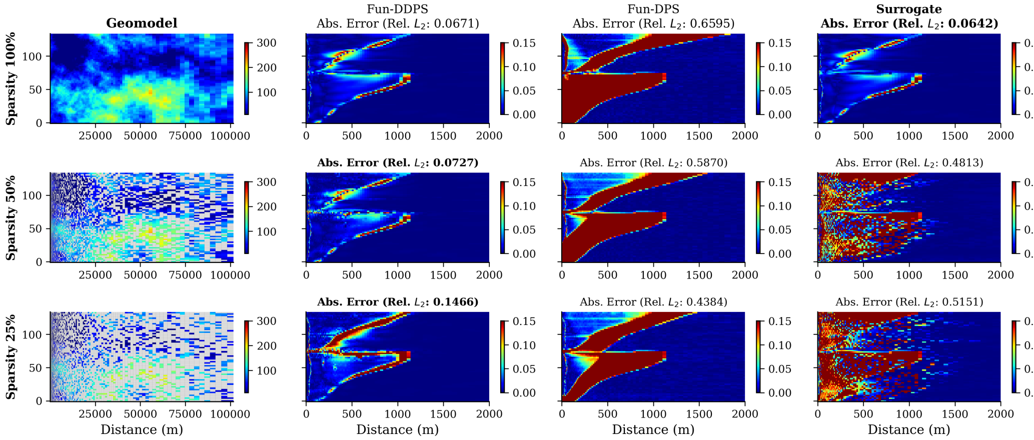

Figure 4: Forward problem under varying observation coverage. **Left:** Observed geomodel at 100%, 50%, and 25% coverage. **Right:** Absolute error in predicted saturation. Fun-DDPS maintains high accuracy (rel. L_2: 0.07–0.15).

tive model attempting to learn both the prior and the forward map simultaneously. This observation aligns with findings in DDIS Lin et al. (2026), which show that decoupled methods often achieve superior sampling efficiency and physical fidelity compared to joint-state approaches when paired training data is scarce.

4.2 INVERSE PROBLEM: PARTIAL DYNAMICS OBSERVATIONS

In this section, we address the inverse problem: conditioning on sparse dynamic observations $\boldsymbol{y}_{dyn}$ to infer the underlying geomodel $\boldsymbol{m}$.

4.2.1 PROBLEM SETUP

Physical Configuration & Ground Truth The ground truth geomodel is generated using SGeMS (Remy et al., 2009a), with permeability hyperparameters μ_{k_r} and σ_{k_r} sampled from uniform priors (See Table 3). The saturation field s_g is simulated over a 30-year injection period using a high-fidelity numerical solver.

We design the experimental setup to reflect the extreme data sparsity encountered in practical Carbon Capture and Storage (CCS) projects. In real-world scenarios, subsurface data is often restricted to well logs, which provide only 1D vertical profiles (column observations) within a vast domain. To mimic this challenge, observations are collected from just two monitoring wells: one at the injector and one 491 meters away (Figure 5). Each well provides 64 dynamic data points, totaling 128 measurements. This configuration corresponds to observing less than 1% of the spatial domain, imposing an extreme challenge for inverse methods to resolve the heterogeneous geomodel from such limited information. We add Gaussian noise with $\sigma_{obs} = 0.04$ to the observations.

Reference Posterior by Rejection Sampling To rigorously validate the diffusion-based posteriors, we compute a "ground truth" reference posterior using Rejection Sampling (RS). We draw from a pool of 2 million prior samples, accepting samples based on the likelihood:

$$L(\boldsymbol{m}) = \exp\left(-\frac{\|\mathbf{M} \odot (\mathcal{L}_\phi(\boldsymbol{m}) - \boldsymbol{y}_{dyn})\|^2}{2\sigma_{obs}^2 \cdot |\mathbf{M}|}\right)$$

	μ_{k_r} (mD)	σ_{k_r} (mD)
Prior	$\mathcal{U}[10, 500]$	$\mathcal{U}[1, 500]$
True	223.0	41.7

Table 3: Geomodel hyperparameters for the ground truth case and their corresponding prior ranges.

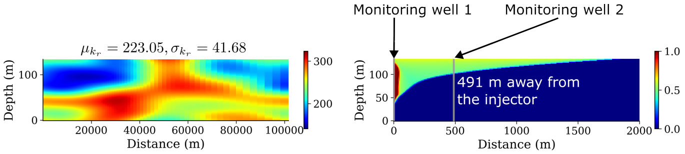

Figure 5: Inverse problem setup. Ground truth permeability (k_r) and gas saturation (s_g) after 30 years. Vertical lines indicate the two monitoring wells where sparse saturation data is assimilated.

Method	JS (Mean)	JS (Std)	N	Filter
RS (Ref.)	–	–	26,082	–
Fun-DPS	**0.047**	**0.037**	1,024	0%
Fun-DDPS	0.051	0.061	998	2.5%

Table 4: Inverse problem validation. JS divergence measures the distance to the RS reference posterior. Both methods achieve high accuracy (JS < 0.06). The Filter column indicates the percentage of diverged samples ($k > 5000\,\text{mD}$). We note that despite slightly worse JS scores, Fun-DDPS produces more physically plausible geomodels as seen in Figure 7.

This yields approximately 26,000 accepted samples, providing a dense representation of the true posterior distribution. The details of the RS algorithm are provided in Appendix C.

Inference Setup We generate 1,024 posterior samples using both Fun-DDPS and Fun-DPS. To ensure a fair comparison, we performed a hyperparameter sweep to determine the optimal guidance strength ζ_{dyn} for each method, effectively weighting the data-consistency term relative to the observation noise. Based on this sweep, the guidance strengths for Fun-DDPS and Fun-DPS are set to $\zeta_{dyn} = 45.0$ and $\zeta_{dyn} = 100.0$, respectively.

4.2.2 RESULTS

Figure 6 compares the marginal posterior distributions of geomodel hyperparameters (μ_{k_r}, σ_{k_r}) obtained from RS, Fun-DPS, and Fun-DDPS, overlaid on the uniform prior distributions (Table 3). Moreover, all three posteriors are sharply concentrated relative to the broad prior, demonstrating significant uncertainty reduction from the measured dynamic data. The RS posterior serves as the ground-truth reference, capturing the true parameter values within its support. Both diffusion methods successfully recover the posterior mode location: Fun-DPS (blue) shows tighter agreement with the RS distribution shape, while Fun-DDPS (orange) exhibits slightly broader tails.

However, a visual inspection of the spatial fields in Figure 7 reveals a critical qualitative difference between the methods. While Fun-DPS is statistically precise, its generated samples tend to exhibit non-physical, high-frequency artifacts. This noise is particularly evident in the Fun-DPS posterior mean (Fig. 7, center column), which appears grainy and textured compared to the smooth reference mean. In contrast, Fun-DDPS produces geologically coherent realizations that maintain physical continuity. By decoupling the generation process, Fun-DDPS effectively leverages the prior to suppress high-frequency noise, resulting in a posterior mean that is cleaner and visually closer to the smoothness of the RS reference.

Finally, in terms of computational efficiency, Fun-DDPS demonstrates a significant advantage. Generating 1,024 posterior samples required approximately 512,000 functional evaluations ($1,024$ samples $\times$ 500 steps), whereas obtaining the reference RS posterior required 2 million evaluations. This represents an approximately $4\times$ reduction in the computational budget.

5 CONCLUSION

We presented Fun-DDPS, a unified framework for geological CO_2 storage that decouples the learning of complex geological priors from the approximation of subsurface flow physics. Our approach addresses the dual challenges of extreme data sparsity and computational expense in subsurface data assimilation.

We demonstrate three critical advantages over existing methods. First, Fun-DDPS exhibits exceptional robustness in forward modeling with partial inputs: while standard deterministic surrogates

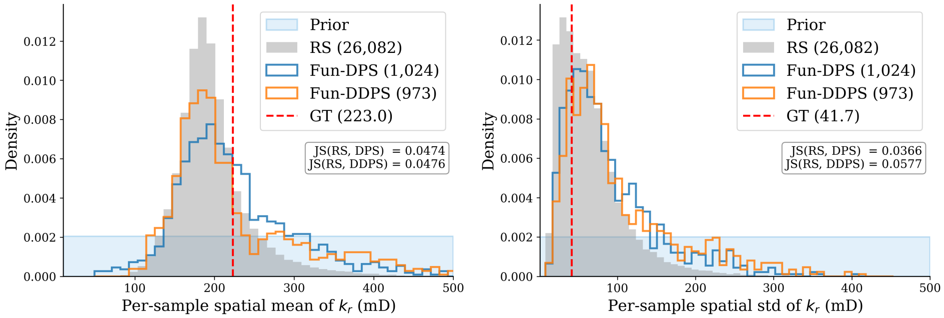

Figure 6: Marginal posterior distributions of geomodel hyperparameters (μ_{k_r}, σ_{k_r}) comparing RS (gray), Fun-DPS (blue), and Fun-DDPS (orange), overlaid on the uniform prior (light blue shaded region). Vertical dashed lines indicate true values.

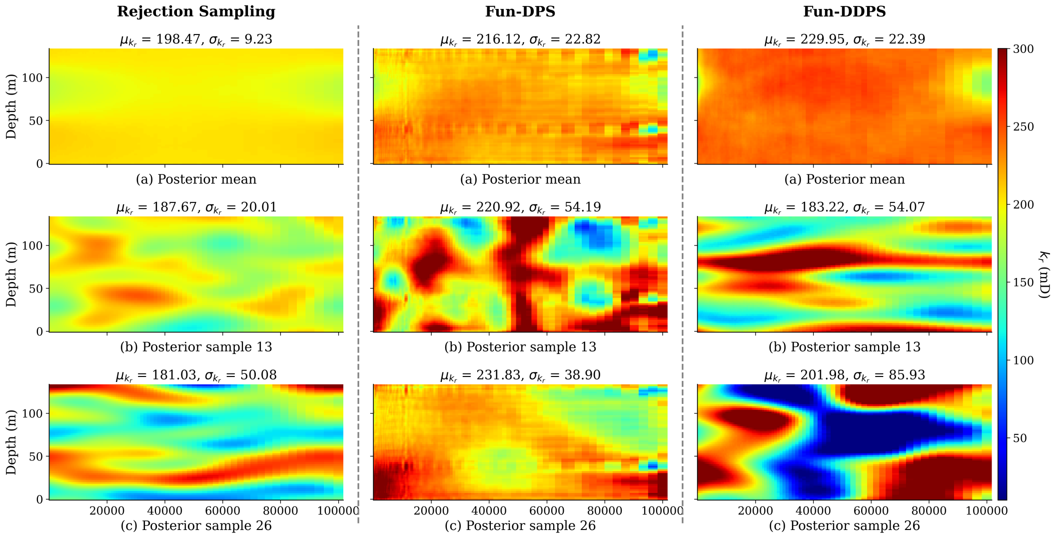

Figure 7: Inverse problem: posterior comparison. Each column shows posterior mean (row a) and two samples (rows b, c). **RS** (left): ground-truth with 26,082 samples. **Fun-DPS** (center): captures posterior structure but suffers from obvious high-frequency artifacts, visible in the grainy posterior mean. **Fun-DDPS** (right): generates more physically plausible geomodels with smoother spatial features, avoiding the artifacts seen in the joint-state baseline.

fail catastrophically (87% error) when facing 25% data coverage, Fun-DDPS leverages its generative prior to achieve a relative error of only 7.7%. Second, we provided the first rigorous validation of diffusion-based inverse solvers against asymptotically exact Rejection Sampling (RS) posteriors. We showed that while both decoupled and joint-state models achieve high statistical accuracy (Jensen-Shannon divergence < 0.06), the decoupled approach produces significantly more physically consistent realizations free from the high-frequency artifacts observed in joint-state baselines. Finally, this performance is achieved with high efficiency, approximating the true posterior with a $4\times$ reduction in the total computational budget compared to Rejection Sampling.

While Fun-DDPS demonstrates strong performance on the presented CCS tasks, our current study simplifies the dynamic state $\boldsymbol{s}$ to a single temporal snapshot (30 years post-injection). In realistic CO_2 monitoring scenarios, data assimilation involves continuous streams of time-series data (e.g., daily pressure gauges, 4D seismic surveys) rather than isolated snapshots. Extending our framework to handle full spatiotemporal trajectories is a natural next step. The decoupled architecture is well-suited for this: the neural operator can map static geomodels to time-varying dynamic states without retraining the expensive diffusion prior.

ACKNOWLEDGMENTS

Anima Anandkumar is supported in part by Bren endowed chair, ONR (MURI grant N00014-23-1-2654), and the AI2050 senior fellow program at Schmidt Sciences. Gege Wen acknowledges the generous support of Schmidt Sciences through the AI2050 fellowship.

REFERENCES

International Energy Agency. *Energy Technology Perspectives 2020-Special Report on Carbon Capture Utilisation and Storage: CCUS in Clean Energy Transitions*. OECD Publishing, 2020.

Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

Hyungjin Chung, Jeongsol Kim, Michael T Mccann, Marc L Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. *arXiv preprint arXiv:2209.14687*, 2022.

Alexandre A. Emerick and Albert C. Reynolds. Ensemble smoother with multiple data assimilation. *Computational Geosciences*, 17(3):503–519, 2013. doi: 10.1007/s10596-013-9347-2.

Geir Evensen. The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, 53(4):343–367, 2003.

Geir Evensen. Sampling strategies and square root analysis schemes for the enkf. *Ocean dynamics*, 54:539–560, 2004.

Mojtaba Forghani, Yizhou Qian, Jonghyun Lee, Matthew Farthing, Tyler Hesser, Peter K Kitanidis, and Eric F Darve. Variational encoder geostatistical analysis (vegas) with an application to large scale riverine bathymetry. *Advances in Water Resources*, 170:104323, 2022.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Jiahe Huang, Guandao Yang, Zichen Wang, and Jeong Joon Park. Diffusionpde: Generative pde-solving under partial observation. *arXiv preprint arXiv:2406.17763*, 2024.

Xin Ju, François P Hamon, Gege Wen, Rayan Kanfar, Mauricio Araya-Polo, and Hamdi A Tchelepi. Learning co2 plume migration in faulted reservoirs with graph neural networks. *Computers & Geosciences*, 193:105711, 2024.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. *Advances in neural information processing systems*, 35:26565–26577, 2022.

Nikola B. Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew M. Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces. *CoRR*, abs/2108.08481, 2021.

John M Lewis, Sivaramakrishnan Lakshmivarahan, and Sudarshan Dhall. *Dynamic data assimilation: a least squares approach*, volume 13. Cambridge University Press, 2006.

Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.

Jae Hyun Lim, Nikola B Kovachki, Ricardo Baptista, Christopher Beckham, Kamyar Azizzadenesheli, Jean Kossaifi, Vikram Voleti, Jiaming Song, Karsten Kreis, Jan Kautz, et al. Score-based diffusion models in function space. *arXiv preprint arXiv:2302.07400*, 2023.

Thomas Y. L. Lin, Jiachen Yao, Lufang Chiang, Julius Berner, and Anima Anandkumar. Decoupled diffusion sampling for inverse problems on function spaces, 2026. URL `https://arxiv.org/abs/2601.23280`.

Miguel Liu-Schiaffini, Julius Berner, Boris Bonev, Thorsten Kurth, Kamyar Azizzadenesheli, and Anima Anandkumar. Neural operators with localized integral and differential kernels. *arXiv preprint arXiv:2402.16845*, 2024.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

Siddharth Misra, Jungang Chen, Yusuf Falola, Polina Churilova, Chung-Kan Huang, and Jose Delgado. Massive geomodel compression and rapid geomodel generation using advanced autoencoders and autoregressive neural networks. In *SPE Europec featured at EAGE Conference and Exhibition?*, pp. D031S010R005. SPE, 2023.

Siddharth Misra, Jungang Chen, Polina Churilova, and Yusuf Falola. Generative artificial intelligence for geomodeling. In *International Petroleum Technology Conference*, pp. D031S128R001. IPTC, 2024.

S. Nejadi, J. Trivedi, and J. Leung. Ensemble kalman filter predictor bias correction method for non-gaussian geological facies detection. In *Offshore Technology Conference*, Houston, TX, 2012. OTC.

Dean S Oliver, Luciane B Cunha, and Albert C Reynolds. Markov chain monte carlo methods for conditioning a permeability field to pressure data. Technical report, 1997.

Stephen Pacala and Robert Socolow. Stabilization wedges: solving the climate problem for the next 50 years with current technologies. *science*, 305(5686):968–972, 2004.

Md Ashiqur Rahman, Zachary E Ross, and Kamyar Azizzadenesheli. U-no: U-shaped neural operators. *arXiv preprint arXiv:2204.11127*, 2022.

Nicolas Remy, Alexandre Boucher, and Jianbing Wu. *Applied geostatistics with SGeMS: A user's guide*. Cambridge University Press, 2009a.

Nicolas Remy, Alexandre Boucher, and Jianbing Wu. *Applied geostatistics with SGeMS: A user's guide*. Cambridge University Press, 2009b.

Christian P Robert, George Casella, and George Casella. *Monte Carlo statistical methods*, volume 2. Springer, 1999.

Schlumberger. Eclipse reservoir simulation software reference manual, 2014.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Meng Tang, Yimin Liu, and Louis J Durlofsky. A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. *Journal of Computational Physics*, 413:109456, 2020.

Meng Tang, Xin Ju, and Louis J Durlofsky. Deep-learning-based coupled flow-geomechanics surrogate model for co2 sequestration. *International Journal of Greenhouse Gas Control*, 118:103692, 2022.

Zheng Teng, Hui Wu, Jize Zhang, Xin Ju, and Shengwen Qi. Generating high-fidelity discrete fracture networks from low-dimensional latent spaces using generative adversarial network. *International Journal of Rock Mechanics and Mining Sciences*, 196:106301, 2025.

Jasper A Vrugt, Cajo JF ter Braak, Cees GH Diks, and Gerrit Schoups. Hydrologic data assimilation using particle markov chain monte carlo simulation: Theory, concepts and applications. *Advances in Water Resources*, 51:457–478, 2013.

Gege Wen, Zongyi Li, Kamyar Azizzadenesheli, Anima Anandkumar, and Sally M Benson. U-fno—an enhanced fourier neural operator-based deep-learning model for multiphase flow. *Advances in Water Resources*, 163:104180, 2022.

Gege Wen, Zongyi Li, Qirui Long, Kamyar Azizzadenesheli, Anima Anandkumar, and Sally M Benson. Real-time high-resolution co 2 geological storage prediction using nested fourier neural operators. *Energy & Environmental Science*, 16(4):1732–1741, 2023.

Ling Yang, Zhilong Zhang, Yang Song, Shenda Hong, Runsheng Xu, Yue Zhao, Wentao Zhang, Bin Cui, and Ming-Hsuan Yang. Diffusion models: A comprehensive survey of methods and applications. *ACM Computing Surveys*, 56(4):1–39, 2023.

Jiachen Yao, Abbas Mammadov, Julius Berner, Gavin Kerrigan, Jong Chul Ye, Kamyar Azizzadenesheli, and Anima Anandkumar. Guided diffusion sampling on function spaces with applications to pdes. *arXiv preprint arXiv:2505.17004*, 2025.

A DIFFUSION POSTERIOR SAMPLING (DPS)

Here we provide the unified algorithm for posterior sampling (Algorithm 1). We formulate the generic algorithm for state variable $\boldsymbol{z}$, which corresponds to $\boldsymbol{m}$ in Fun-DDPS and $\boldsymbol{x} = (\boldsymbol{m}, \boldsymbol{s})$ in Fun-DPS. The following details are presented in a finite-dimensional setting; for the function-space variant, please refer to Fun-DPS (Yao et al., 2025).

To evaluate the conditional score $\nabla_{\boldsymbol{z}} \log p(\boldsymbol{y}_{obs}|\boldsymbol{z}_\sigma)$, we employ the standard DPS approximation (Chung et al., 2022) using the denoised estimate $\hat{\boldsymbol{z}}_0$. By Tweedie's formula, we have:

$$\hat{\boldsymbol{z}}_0(\boldsymbol{z}_\sigma) = \boldsymbol{D}_\theta(\boldsymbol{z}_\sigma, \sigma) \approx \boldsymbol{z}_\sigma + \sigma^2 \nabla_{\boldsymbol{z}} \log p(\boldsymbol{z}_\sigma) \tag{9}$$

The conditional score (a.k.a.the observation guidance gradient) is then approximated via Maximum Likelihood Estimation (MLE):

$$\nabla_{\boldsymbol{z}_\sigma} \log p(\boldsymbol{y}_{obs}|\boldsymbol{z}_\sigma) \approx \nabla_{\boldsymbol{z}_\sigma} \log p(\boldsymbol{y}_{obs}|\hat{\boldsymbol{z}}_0(\boldsymbol{z}_\sigma)) = -\zeta \nabla_{\boldsymbol{z}_\sigma} \|\boldsymbol{y}_{obs} - \mathcal{H}(\hat{\boldsymbol{z}}_0(\boldsymbol{z}_\sigma))\|_2^2 \tag{10}$$

Algorithm 1 Unified Diffusion Posterior Sampler

```
Require: Observations $\boldsymbol{y}_{obs}$, trained denoiser $\boldsymbol{D}_\theta$, noise schedule $\{\sigma_i\}_{i=0}^N$ (descending), guidance
    weight $\zeta$, GRF covariance $\mathbf{C}_\gamma$.
Setup:
if Method is Fun-DDPS then
    State $\boldsymbol{z} \leftarrow \boldsymbol{m}$                                              ▷ Sampling Geomodel only
    Operator $\mathcal{H}_{obs}(\hat{\boldsymbol{z}}_0) \leftarrow M_{dyn} \odot \mathcal{L}_\phi(\hat{\boldsymbol{m}}_0)$        ▷ Backprop through Surrogate $\mathcal{L}_\phi$
else if Method is Fun-DPS then
    State $\boldsymbol{z} \leftarrow (\boldsymbol{m}, \boldsymbol{s})$                                       ▷ Sampling Joint State
    Operator $\mathcal{H}_{obs}(\hat{\boldsymbol{z}}_0) \leftarrow M_{dyn} \odot \hat{\boldsymbol{s}}_0$                  ▷ Direct channel guidance
end if
$\boldsymbol{z}_N \sim \mathcal{N}(0, \sigma_N^2 \mathbf{C}_\gamma)$                                        ▷ Initialize from prior
Loop:
for $i = N$ to $1$ do
    $\hat{\boldsymbol{z}}_0^{(i)} \leftarrow \boldsymbol{D}_\theta(\boldsymbol{z}_i, \sigma_i)$                        ▷ Estimate clean state (Tweedie's formula)
    $\boldsymbol{d}_i \leftarrow (\boldsymbol{z}_i - \hat{\boldsymbol{z}}_0^{(i)})/\sigma_i$                                ▷ Compute score/direction
    // 1. Predictor Step (Euler's method)
    $\boldsymbol{z}_{i-1}^{pred} \leftarrow \boldsymbol{z}_i + (\sigma_{i-1} - \sigma_i)\boldsymbol{d}_i$
    // 2. Corrector Step (2nd-order Heun's method)
    if $i > 1$ and $\sigma_{i-1} > 0$ then
        $\hat{\boldsymbol{z}}_0^{pred} \leftarrow \boldsymbol{D}_\theta(\boldsymbol{z}_{i-1}^{pred}, \sigma_{i-1})$
        $\boldsymbol{d}_{i-1} \leftarrow (\boldsymbol{z}_{i-1}^{pred} - \hat{\boldsymbol{z}}_0^{pred})/\sigma_{i-1}$                       ▷ Direction at next step
        $\boldsymbol{z}_{i-1} \leftarrow \boldsymbol{z}_i + (\sigma_{i-1} - \sigma_i)\left(\frac{1}{2}\boldsymbol{d}_i + \frac{1}{2}\boldsymbol{d}_{i-1}\right)$             ▷ Trapezoidal update
    else
        $\boldsymbol{z}_{i-1} \leftarrow \boldsymbol{z}_{i-1}^{pred}$
    end if
    // 3. Guidance Step
    if $\sigma_{i-1} > 0$ then
        $\hat{\boldsymbol{z}}_0' \leftarrow \boldsymbol{D}_\theta(\boldsymbol{z}_{i-1}, \sigma_{i-1})$                            ▷ Denoise for gradient calculation
        $\boldsymbol{g} \leftarrow \nabla_{\boldsymbol{z}_{i-1}} \|\boldsymbol{y}_{obs} - \mathcal{H}_{obs}(\hat{\boldsymbol{z}}_0')\|_2^2$
        $\boldsymbol{z}_{i-1} \leftarrow \boldsymbol{z}_{i-1} - \zeta \cdot \boldsymbol{g}$                                ▷ Apply gradient guidance
    end if
end for
return $\boldsymbol{D}_\theta(\boldsymbol{z}_0, \sigma_0)$
```

B JOINT-STATE DIFFUSION MODELS

This section details the joint-state diffusion model, with DiffusionPDE (Huang et al., 2024) and Fun-DPS (Yao et al., 2025) as the two representative examples. While the proposed Fun-DDPS (Section 3) decouples the prior $p(\boldsymbol{m})$ from the physics, the joint-state approach learns the joint distribution $p(\boldsymbol{m}, \boldsymbol{s})$ directly.

We first define the joint state $\boldsymbol{x} = (\boldsymbol{m}, \boldsymbol{s}) \in \mathcal{X}$, where $\mathcal{X} = \mathcal{M} \times \mathcal{S}$ is the product Hilbert space of parameters and states. The diffusion model is trained on pairs $\{\boldsymbol{x}^{(i)}\}_{i=1}^{N}$ generated by the simulator, implicitly capturing physical constraints within the joint distribution.

We introduce a centered Gaussian reference measure $\mathcal{N}(0, \mathbf{C}_{\mathcal{X}})$ on the joint space. The forward process perturbs a clean sample $\boldsymbol{x}_0$ via:

$$\boldsymbol{x}_\sigma = \boldsymbol{x}_0 + \boldsymbol{\xi}_\sigma, \quad \boldsymbol{\xi}_\sigma \sim \mathcal{N}(0, \sigma(t)^2 \mathbf{C}_{\mathcal{X}}) \tag{11}$$

where $\sigma(t)$ is the noise schedule.

We train a joint denoising operator $\boldsymbol{D}_\theta(\boldsymbol{x}_\sigma, \sigma)$ to recover $\boldsymbol{x}_0$. The loss function is the function-space analogue of denoising score matching:

$$\mathcal{L}_{joint}(\theta) = \mathbb{E}_{\boldsymbol{x}_0, \sigma, \boldsymbol{\xi}_\sigma} \left[\lambda(\sigma) \| \boldsymbol{D}_\theta(\boldsymbol{x}_0 + \boldsymbol{\xi}_\sigma, \sigma) - \boldsymbol{x}_0 \|_{\mathcal{X}}^2 \right] \tag{12}$$

It is worth noting that during training, the physics is learned implicitly through the cross-state statistics and is not part of the training objective.

C REJECTION SAMPLING BENCHMARK

This section details the Rejection Sampling (RS) procedure used to obtain ground-truth posteriors for validating Fun-DDPS and Fun-DPS in Section 4.2.

C.1 MOTIVATION

RS provides exact posterior samples when proposals are drawn from the prior. Unlike diffusion-based methods that approximate the posterior, RS serves as a rigorous benchmark. However, RS requires many forward evaluations due to low acceptance rates; we use a pre-trained neural operator surrogate $\mathcal{L}_\phi$ to make this computationally tractable.

C.2 PROBLEM SETUP

We use the same setup as the one shown in Fig. 5, where two columns of gas saturation are observed. We use the identical ground truth as the one described in the previous section, where the same observation models and noise variance are used for conducting RS. As the RS method is essentially a rigorous search method that characterizes the exact posterior distribution, we need to generate a large number of prior samples, as the acceptance rate can be very low. Specifically, for generating prior geomodels, we first uniformly sample hyperparameters and later use SGeMS to generate 2 million geomodels that cover the prior space. Then, the pretrained surrogate model is used as the forward model to predict the s_g distribution.

To quantitatively measure the discrepancy between the posterior distributions and the RS ground truth, we use the Jensen-Shannon (JS) divergence. The JS divergence is a symmetric and smoothed version of the Kullback-Leibler (KL) divergence Bishop & Nasrabadi (2006), providing a bounded metric for the similarity between two probability distributions. It is defined as:

$$\text{JS}(P \| Q) = \frac{1}{2} D_{KL}(P \| M) + \frac{1}{2} D_{KL}(Q \| M) \tag{13}$$

where P is the approximate posterior distribution (from machine learning models), Q is the reference posterior (from RS), and $M = \frac{1}{2}(P + Q)$ is the average of the two distributions. A JS divergence of 0 indicates that the distributions are identical, while a larger value signifies a greater dissimilarity.

D PRIOR ASSESSMENT OF FUN-DDPS

A critical step in validating the Fun-DDPS framework is ensuring that the decoupled generation process produces statistically and physically representative samples. Unlike joint-state models, Fun-DDPS generates samples in a two-step pipeline: first, unconditional geomodels are drawn from the learned diffusion prior $\boldsymbol{m} \sim p_\theta(\boldsymbol{m})$; second, these realizations are mapped through the pre-trained LocalNO surrogate to obtain the corresponding dynamic states $\boldsymbol{s} = \mathcal{L}_\phi(\boldsymbol{m})$. To assess this composite pipeline, we generated an ensemble of 1,000 unconditional pairs $\{(\boldsymbol{m}^{(i)}, \boldsymbol{s}^{(i)})\}$ and compared them against a ground-truth test set of equal size.

D.1 QUALITATIVE VISUAL INSPECTION

We first assess the visual fidelity of the generated fields (Fig. 8). The diffusion-generated permeability fields $\boldsymbol{m}$ exhibit geologically realistic features, faithfully reproducing the continuous high-permeability channels and spatial heterogeneity characteristic of the training distribution. Notably, the corresponding saturation plumes $\boldsymbol{s}$—predicted by the surrogate—demonstrate physically consistent migration patterns. The surrogate successfully captures sharp frontal boundaries and complex fluid-rock interactions driven by the underlying geological heterogeneity, with no observable artifacts.

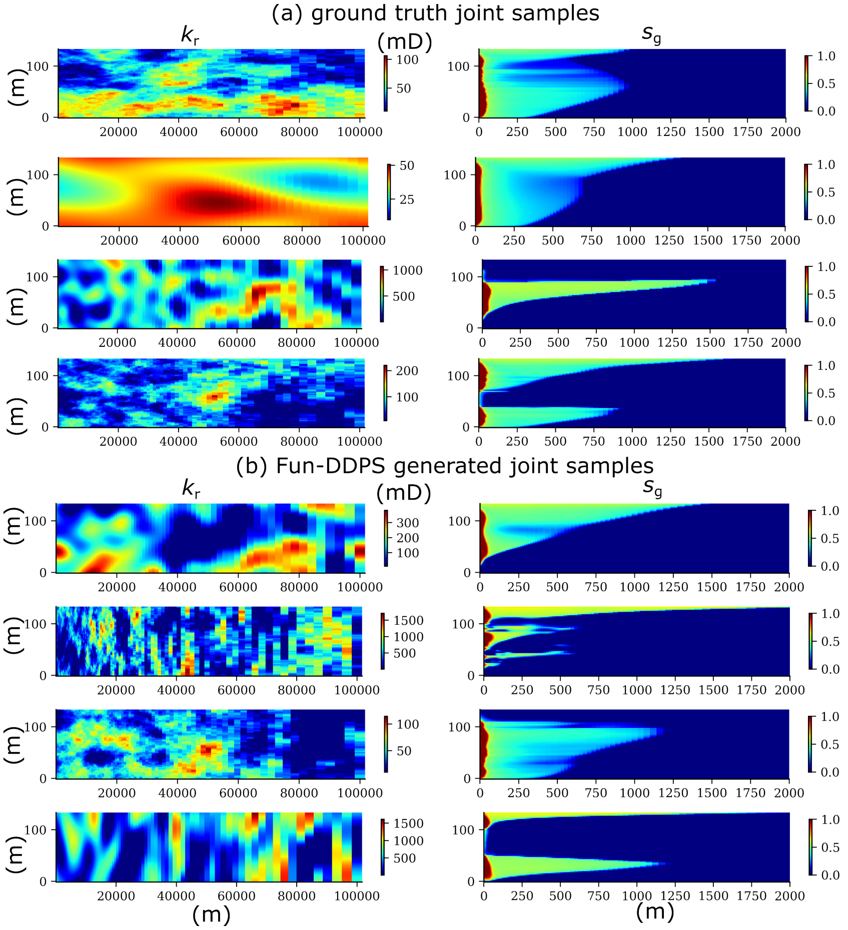

Figure 8: Randomly selected unconditional samples. (a) Ground Truth: Reference pairs from the test set. (b) Fun-DDPS Generated: Geomodels $\boldsymbol{m}$ generated by the diffusion prior and saturation fields $\boldsymbol{s}$ predicted by the LocalNO surrogate. Note that only the s_g field within $r = 2000$ m is shown.

D.2 GEOLOGICAL SPATIAL STATISTICS: VARIOGRAM ANALYSIS

To quantify the spatial structure of the generated geomodels, we employ variogram analysis. The semivariogram $\gamma(\boldsymbol{h})$ measures the spatial dissimilarity of the permeability field as a function of lag distance $\boldsymbol{h}$:

$$\gamma(\boldsymbol{h}) = \frac{1}{2}\mathrm{Var}\left[\boldsymbol{m}(\mathbf{u}) - \boldsymbol{m}(\mathbf{u} + \boldsymbol{h})\right] \tag{14}$$

where Var denotes the variance across the ensemble. Fig. 9(a-b) compares the experimental variograms of the generated ensemble against the ground truth. The Fun-DDPS samples show excellent agreement in both vertical and horizontal directions. Notably, the uncertainty bands (P5–P95) of the generated ensemble overlap consistently with the ground truth, confirming that the diffusion prior $p_\theta(\boldsymbol{m})$ correctly captures the multi-scale spatial variability of the geological distribution.

D.3 DYNAMIC SPATIAL CONNECTIVITY: TWO-POINT PROBABILITY

We evaluate the induced distribution of dynamic states using two-point connectivity functions on the gas saturation plumes. This metric calculates the probability that two points separated by a vector $\boldsymbol{h}$ both lie within the CO_2 plume ($S_g > 0$), serving as a proxy for plume geometry and connectivity. As shown in Fig. 9(c-d), the statistics of the surrogate-predicted saturation fields closely match the ground truth. This result is significant: it confirms that the LocalNO surrogate $\mathcal{L}_\phi$ preserves the complex functional mapping from geological heterogeneity to fluid flow, producing saturation fields with statistically correct spatial correlations.

D.4 ENSEMBLE DISTRIBUTION ASSESSMENT

Finally, we examine the global statistics via Cumulative Distribution Functions (CDFs) in Fig. 9(e-f). The CDFs for both generated permeability and predicted saturation align closely with the reference distributions. This confirms that the Fun-DDPS pipeline is unbiased, effectively capturing the full range of geological and dynamical variability present in the training data without mode collapse.

E DATA GENERATIONS AND TRAINING DETAILS

E.1 GOVERNING EQUATIONS

The underlying physical system is governed by the conservation of mass and momentum for a two-phase (CO_2-water) system in a porous medium. Neglecting chemical reactions and assuming slightly compressible fluids, the mass balance for phase α (where $\alpha \in \{w, g\}$ for water and gas) is given by:

$$\frac{\partial}{\partial t}(\phi \rho_\alpha S_\alpha) + \nabla \cdot (\rho_\alpha \mathbf{u}_\alpha) = q_\alpha \tag{15}$$

where ϕ is porosity, ρ_α is density, S_α is saturation, and $\mathbf{u}_\alpha$ is the Darcy velocity vector defined as:

$$\mathbf{u}_\alpha = -\frac{k_{ri}\mathbf{K}}{\mu_\alpha}(\nabla P_\alpha - \rho_\alpha \mathbf{g}) \tag{16}$$

The system is closed with the saturation constraint $S_w + S_g = 1$ and capillary pressure relationships $P_c = P_g - P_w$.

E.2 SIMULATION SETUP

The domain mimics an infinite-acting aquifer with a radius of 100 km and a thickness of 135 m. We enforce no-flow boundaries at the top and bottom caprock. CO_2 is injected at a constant rate of 0.36 Mt/year for 30 years.

E.3 DATA GENERATION DETAILS

The permeability fields $\boldsymbol{m}$ are generated using Sequential Gaussian Simulation (SGSIM). The geostatistical hyperparameters are sampled from uniform distributions to ensure a wide variety of geological scenarios. The specific ranges are provided in Table 5.

Variable	Distribution	Unit
Radial correlation (x)	$\mathcal{U}[359, 35900]$	m
Vertical correlation (z)	$\mathcal{U}[14, 56]$	m
Permeability Mean (μ_k)	$\mathcal{U}[10, 500]$	mD
Permeability Std (σ_k)	$\mathcal{U}[1, 500]$	mD

Table 5: Prior distributions for geostatistical parameters.

E.4 TRAINING AND ARCHITECTURE DETAILS

We train all diffusion models following the Elucidating Diffusion Models (EDM) framework Karras et al. (2022). The noise level σ is sampled during training from a log-uniform distribution over the range $[\sigma_{\min}, \sigma_{\max}] = [0.002, 80]$, ensuring the model learns to denoise effectively across all scales. Optimization is performed using AdamW Loshchilov & Hutter (2017) with a batch size of 64 for 200 epochs. All experiments were conducted on a single NVIDIA A100 GPU (80GB VRAM).

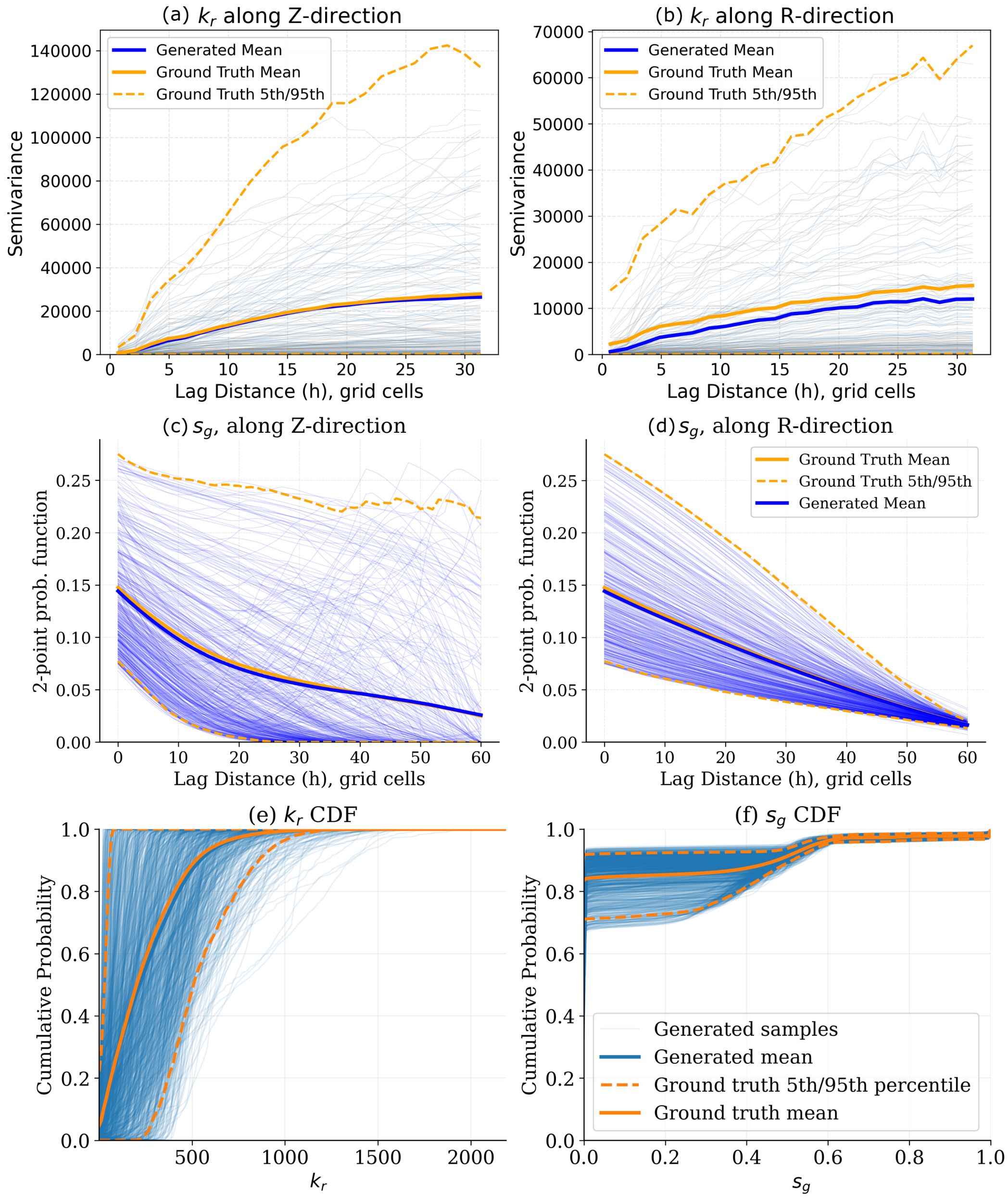

Figure 9: Quantitative statistical comparison of Fun-DDPS generated samples versus ground truth. (a-b) Experimental variograms for permeability $\boldsymbol{m}$ (validating the diffusion prior). (c-d) Two-point connectivity functions for gas saturation $\boldsymbol{s}$ (validating the surrogate mapping). (e-f) Cumulative distribution functions (CDFs) for both fields.

E.4.1 DIFFUSION OPERATOR

The diffusion model backbone is a U-shaped Neural Operator (U-NO) (Rahman et al., 2022). It features a 4-level multi-scale hierarchy (resolutions $64 \times 200 \to 32 \times 100 \to 16 \times 50 \to 8 \times 25$) with a base channel dimension of $d_h = 64$ and multipliers $[1, 2, 4, 4]$. Each resolution level contains residual blocks equipped with spectral convolutions using ComplexTucker factorization (rank 0.1) for parameter efficiency. Fourier modes are truncated at 50% of the spatial resolution at each specific level. A self-attention mechanism is applied at the 8×25 bottleneck to capture global dependencies. Time conditioning σ is injected via sinusoidal positional embeddings, which are mapped through an MLP to adaptive scale and shift parameters (AdaGN) within each residual block. To ensure a fair comparison, both the Fun-DPS joint model (2-channel input $[\boldsymbol{m}, \boldsymbol{s}]$) and the Fun-DDPS geomodel prior (1-channel input $\boldsymbol{m}$) share identical architectural hyperparameters and model size.

E.4.2 NEURAL SURROGATE

We employ a Local Neural Operator (Liu-Schiaffini et al., 2024) that combines global Fourier layers with local Discrete Continuous (DISCO) convolutions. The architecture uses 6 spectral layers with modes truncated at 16 in each spatial direction. The lifting layer projects inputs to a hidden dimension of $d_h = 48$. DISCO kernels of size 4×4 are utilized to capture local spatial features while maintaining resolution invariance.

E.5 EVALUATION METRIC

To quantify performance, we compute the relative L_2 error between the ground truth solution $\boldsymbol{x}_{\text{gt}}$ and the predicted solution $\boldsymbol{x}_{\text{pred}}$:

$$\mathcal{E}_{\text{rel}} = \frac{\|\boldsymbol{x}_{\text{gt}} - \boldsymbol{x}_{\text{pred}}\|_2}{\|\boldsymbol{x}_{\text{gt}}\|_2} \tag{17}$$