TRACE: RETROSPECTIVE STREAMING GENERATION OF PHYSICAL FIELDS UNDER SPARSE STRUCTURED SENSING

Anonymous authors
Paper under double-blind review

ABSTRACT

Reconstructing continuous physical fields from sparse measurements is central to scientific monitoring, inverse modeling, and digital-twin construction. Generative reconstruction has recently emerged as a promising paradigm for this task by learning data-driven physical priors that complete plausible full fields from limited observations. However, existing methods largely assume fixed, batch conditioning, whereas real sensing systems often produce structured streams: probes scan local regions, instruments observe moving fields of view, and communication constraints may leave entire frames missing. We propose **TRACE**, a retrospective streaming generative reconstruction framework for physical fields under structured sensing. TRACE performs approximate Bayesian inference in a learned continuous-coordinate latent space, converting sparse off-grid measurements into generative latent evidence, fusing it with a state-space temporal prior through Kalman-style filtering, and refining under-observed past frames via retrospective smoothing. Experiments on active matter, ocean sound-speed fields, and supernova simulations show that TRACE remains competitive with or surpasses frame-wise generative reconstructors, offline spatiotemporal methods, and streaming data-assimilation baselines in reconstruction quality under temporally sparse and spatially localized sensing protocols.

1 INTRODUCTION

Reconstructing continuous physical fields from sparse measurements is a foundational problem in scientific machine learning, with applications in physical monitoring, ocean sensing, astronomical simulation, active matter, and digital-twin maintenance. The target quantity is often a field evolving over space and time, while sensors provide only sparse, irregular, and off-grid readings. This sparse-to-full problem is highly ill-posed: many full fields can explain the same limited observations. Generative reconstruction has recently become a promising paradigm for this setting. By learning data-driven priors over plausible physical fields, diffusion posterior sampling (Chung et al., 2023), diffusion-based physical-field solvers (Huang et al., 2024a), and continuous-coordinate generative field models (Du et al., 2024) can recover coherent structures from incomplete measurements, turning reconstruction from interpolation into conditional generation under physical and observational constraints.

However, most generative reconstructors still assume a fixed, batch view of evidence: observations are available before inference either for a single frame or for a complete trajectory. Real sensing systems are often different. A mobile probe scans only a local region at each time, an imaging instrument observes a moving field of view, and power or communication constraints may leave entire frames missing. In such structured sensing streams, sparsity is organized over time. A single frame is usually under-informative, and its meaning depends on evidence observed before and after it. Existing methods do not jointly address this regime. Frame-wise generative reconstructors provide strong field priors but cannot propagate evidence across time. Offline spatiotemporal reconstructors such as SDIFT (Chen et al., 2025b) exploit temporal context but require the full observation horizon before reconstruction. Classical Kalman and ensemble data assimilation (Kalman, 1960; Evensen, 1994), as well as recent neural variants (Rozet & Louppe, 2023; Xiao et al., 2024; Tarumi et al., 2025), support online updates but usually rely on gridded states or explicit dynamics rather

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.

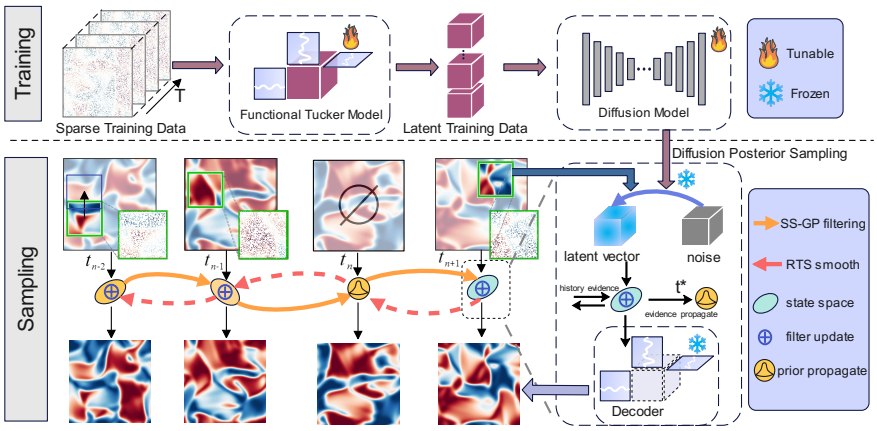


Figure 1: Overview of TRACE. Top: offline pretraining of the continuous-coordinate field representation and latent diffusion prior. Bottom: streaming inference, fusing per-frame generative evidence with a state-space prior via Kalman filtering and refining it via retrospective smoothing.

than continuous-coordinate generative field completion from sparse off-grid measurements. What remains missing is a framework that keeps the learned physical prior of generative reconstruction, accumulates weak evidence as the stream unfolds, and revises earlier ambiguous estimates when later measurements make them identifiable.

We address this problem with **TRACE** (**T**emporal **R**etrospective **A**ccumulation for **C**ontinuous-field **E**stimation), a retrospective streaming generative reconstruction framework for physical fields under structured sensing. TRACE treats each incoming frame as latent evidence rather than as an isolated reconstruction target. It represents fields in a compact continuous-coordinate latent space using a functional Tucker decoder, so sparse off-grid measurements can constrain arbitrary spatial queries. A pretrained generative latent prior supplies plausible single-frame completions, which are summarized as Gaussian latent evidence through generative posterior sampling and moment matching. TRACE then fuses this weak per-frame evidence with a Matérn state-space temporal prior using a Kalman-style update. The forward filter gives causal online estimates from observations seen so far. Once later measurements arrive, a Rauch–Tung–Striebel smoothing pass propagates information backward to refine past frames that were missing, localized, or otherwise under-observed. TRACE thus separates causal filtering from retrospective smoothing, distinguishing the online estimate from the refined offline correction. An overview is shown in Fig. 1.

Our contributions are as follows. **(C1)** We formulate physical-field reconstruction under **structured sensing**, where sparse off-grid measurements arrive as temporally organized streams, including temporally sparse frames and spatially localized moving-window observations. **(C2)** We propose **TRACE**, a retrospective streaming generative reconstruction framework that performs approximate Bayesian evidence fusion in a learned continuous-coordinate latent space. **(C3)** We develop a filtering-and-smoothing inference mechanism that provides causal online estimates and retrospectively refines past under-observed frames as later evidence arrives. **(C4)** We validate TRACE on active matter, ocean sound-speed fields, and supernova simulations, showing that TRACE matches or surpasses frame-wise generative reconstructors, offline spatiotemporal methods, and streaming data-assimilation baselines in reconstruction quality under temporally sparse and spatially localized sensing protocols.

2 Preliminaries

2.1 Continuous Low-Rank Tensor Representations

Many scientific fields are naturally represented as continuous K-variate functions $f : \mathcal{R}_1 \times \cdots \times \mathcal{R}_K \to \mathbb{R}$ with coordinate $\mathbf{r} = (r_1, \ldots, r_K)$. Low-rank tensor representations combined with coordinate-based neural basis functions capture such functions through compact latent coefficients

and continuous coordinate decoders,

$$f(r_1, \ldots, r_K) \approx \left\langle \mathcal{C}, \phi^{(1)}(r_1) \otimes \cdots \otimes \phi^{(K)}(r_K) \right\rangle, \tag{1}$$

where $\phi^{(k)}(\cdot)$ denotes the coordinate-based basis function for the k-th mode, $\mathcal{C}$ is a compact coefficient tensor capturing the multilinear interactions across modes, and $\langle \cdot, \cdot \rangle$ denotes multilinear contraction. This decouples continuous decoding from a compact latent code, enabling off-grid queries and the downstream generative and temporal modeling.

2.2 Diffusion Priors and Posterior Sampling

Diffusion models Song et al. (2021); Karras et al. (2022) learn a generative prior by training a denoising network to recover clean samples from progressively corrupted observations. Once trained, the denoiser implicitly defines a score function

$$s_{\boldsymbol{\theta}}(\mathbf{z}; \sigma) := \nabla_{\mathbf{z}} \log p_\sigma(\mathbf{z}) \approx \frac{D_{\boldsymbol{\theta}}(\mathbf{z}; \sigma) - \mathbf{z}}{\sigma^2}, \tag{2}$$

which estimates the gradient of the log-density of the noisy distribution $p_\sigma(\mathbf{z})$ and defines a learned unconditional prior. For inverse problems, however, generation must additionally satisfy the available observations. Diffusion Posterior Sampling (DPS) Chung et al. (2023) extends the pretrained generative prior to conditional generation by incorporating measurement guidance during reverse diffusion. Given observations $\mathbf{y}$ generated through a forward measurement model $G(\cdot)$, DPS approximates the posterior score as

$$\nabla_{\mathbf{z}} \log p(\mathbf{z} \mid \mathbf{y}; \sigma) \approx s_{\boldsymbol{\theta}}(\mathbf{z}; \sigma) - \zeta \nabla_{\mathbf{z}} \left\| \mathbf{y} - G(D_{\boldsymbol{\theta}}(\mathbf{z}; \sigma)) \right\|_2^2, \tag{3}$$

where $D_{\boldsymbol{\theta}}(\mathbf{z}; \sigma)$ denotes the denoised latent estimate, $G(\cdot)$ is the forward measurement model relating latent representations to observations, and $\zeta > 0$ controls the guidance strength. Reverse diffusion driven by the guided score yields observation-consistent latent samples that jointly satisfy the learned generative prior and the measurement constraints.

2.3 State-Space Gaussian Processes

Gaussian processes (GPs) provide a non-parametric prior over continuous functions $x(t) \sim \mathcal{GP}(0, \kappa(t, t'))$, but exact GP inference scales as $\mathcal{O}(N^3)$ in the trajectory length N. For half-integer Matérn kernels, the GP admits an equivalent finite-dimensional state-space representation Hartikainen & Särkkä (2010); Särkkä & Solin (2019) through an augmented Markov state $\boldsymbol{\xi}(t)$ that stacks $x(t)$ with its first p derivatives, satisfying the linear time-invariant SDE

$$\frac{d\boldsymbol{\xi}(t)}{dt} = \mathbf{F}\boldsymbol{\xi}(t) + \mathbf{L}w(t), \qquad x(t) = \mathbf{H}\boldsymbol{\xi}(t), \tag{4}$$

where $(\mathbf{F}, \mathbf{L}, \mathbf{H})$ are determined by the kernel hyperparameters and $\mathbf{H}$ projects the augmented state onto the function value. Discretising over arbitrary timestamps $t_1 < \cdots < t_N$ with intervals $\Delta_n := t_n - t_{n-1}$ yields an equivalent linear-Gaussian Gauss–Markov transition

$$p(\boldsymbol{\xi}_n \mid \boldsymbol{\xi}_{n-1}) = \mathcal{N}(\mathbf{A}_n \boldsymbol{\xi}_{n-1},\, \mathbf{Q}_n), \tag{5}$$

with $\mathbf{A}_n = \exp(\Delta_n \mathbf{F})$ and process noise $\mathbf{Q}_n$ obtained from the stationary Lyapunov equation (explicit forms in Appendix A.2). This state-space formulation admits $\mathcal{O}(N)$ Bayesian inference via Kalman filtering and Rauch–Tung–Striebel smoothing Rauch et al. (1965), providing an efficient temporal inference interface for streaming trajectories while preserving the underlying Gaussian process prior.

3 Method

Problem Statement: We consider a continuous spatiotemporal physical field $\mathcal{Y} : \Omega \times \mathbb{R}_+ \to \mathbb{R}$ over the spatial domain Ω. Sensing produces an irregular stream of sparse off-grid measurements at timestamps $t_1 < t_2 < \cdots$; at each t_n the active sensors within a frame-dependent region $\Omega_{\text{obs}}^{(n)} \subseteq \Omega$ give the observation set

$$\mathcal{O}_n = \left\{ (\mathbf{r}_{n,m},\, y_{n,m},\, t_n) \right\}_{m=1}^{M_n}, \tag{6}$$

with $\mathbf{r}_{n,m} \in \Omega_{\text{obs}}^{(n)}$ and $y_{n,m} = \mathcal{Y}(\mathbf{r}_{n,m}, t_n) + \epsilon_{n,m}$, where $\epsilon_{n,m}$ denotes observation noise. The per-frame count M_n may vanish ($\mathcal{O}_n = \emptyset$ for entirely unobserved frames), and the streaming history up to t_n is $\mathcal{O}_{1:n} = \{\mathcal{O}_1, \ldots, \mathcal{O}_n\}$. Our objective is to infer $\mathcal{Y}$ at arbitrary continuous queries $(\mathbf{r}^\star, t^\star) \in \Omega \times \mathbb{R}_+$ from $\mathcal{O}_{1:n}$ online, while allowing later observations to retrospectively refine previously under-observed frames.

3.1 Streaming Reconstruction as Latent Inference

Performing sequential inference directly in the original physical field space is challenging under streaming sparse sensing: irregular off-grid observations vary across frames, preventing a common spatial discretisation for temporal aggregation; reconstructing a continuous field from an individual under-observed frame is ill-posed without exploiting intrinsic spatial structure; and recursive inference over the original high-dimensional continuous function space is computationally prohibitive.

To address these challenges, TRACE first reformulates streaming field reconstruction as inference over a compact latent trajectory using the Functional Tucker Model (FTM) Fang et al. (2024b). FTM meets these difficulties in turn: a shared continuous basis serves arbitrary off-grid queries, so differently sampled frames map into one representation; a low-rank core constrains each otherwise ill-posed single-frame reconstruction; and a fixed-dimensional latent keeps the recursion tractable.

Specifically, for a physical field $\mathcal{Y}$ defined over the K-mode spatial domain $\Omega = \mathcal{R}_1 \times \cdots \times \mathcal{R}_K$ with preset multilinear ranks $\{R_k\}_{k=1}^K$, FTM parameterises each frame t_n by a frame-specific Tucker core $\mathcal{W}_n \in \mathbb{R}^{R_1 \times \cdots \times R_K}$ together with shared continuous basis functions $\boldsymbol{\phi}_{\theta_k}^{(k)} : \mathcal{R}_k \to \mathbb{R}^{R_k}$. The field value at an arbitrary continuous coordinate $\mathbf{r} = (r_1, \ldots, r_K)$ is represented as

$$\mathcal{Y}(\mathbf{r}, t_n) \approx \text{vec}(\mathcal{W}_n)^\top \left(\boldsymbol{\phi}_{\theta_1}^{(1)}(r_1) \otimes \cdots \otimes \boldsymbol{\phi}_{\theta_K}^{(K)}(r_K) \right), \tag{7}$$

where $\text{vec}(\cdot)$ denotes vectorisation and $\otimes$ denotes the Kronecker product. Unlike tensor decompositions defined on fixed spatial grids, the continuous basis functions can be evaluated at arbitrary query coordinates, naturally accommodating different off-grid sensing layouts while preserving a unified latent representation.

FTM is pretrained on a training corpus $\mathcal{D} = \{\mathcal{O}^{(b)}\}_{b=1}^B$ of B complete trajectories, where each $\mathcal{O}^{(b)} = \{\mathcal{O}_n^{(b)}\}_{n=1}^{N^{(b)}}$ is a full observation sequence, by jointly learning the basis $\{\boldsymbol{\phi}_{\theta_k}^{(k)}\}$ and per-trajectory cores $\{\mathcal{W}_n^{(b)}\}_{b,n}$; the training procedure and network architectures are deferred to Appendix C.1, and the offline pretraining is summarised in Algorithm 1.

To obtain a unified latent vector for subsequent inference objectives, we vectorise each Tucker core as $z_n = \text{vec}(\mathcal{W}_n) \in \mathbb{R}^d$, where $d = \prod_{k=1}^K R_k$, yielding the latent training dataset $\mathcal{Z} = \{z_n^{(b)}\}_{b,n}$. The physical field is recovered through the decoder

$$\hat{\mathcal{Y}}(\mathbf{r}^\star, t_n) = \mathbf{G}(\mathbf{r}^\star)^\top z_n, \tag{8}$$

where $\mathbf{G}(\mathbf{r}^\star) = \bigotimes_{k=1}^K \boldsymbol{\phi}_{\theta_k}^{(k)}(r_k^\star) \in \mathbb{R}^d$ evaluates the frozen basis functions at the query coordinate $\mathbf{r}^\star = (r_1^\star, \ldots, r_K^\star)$. The same operator $\mathbf{G}(\cdot)$ thus serves both sparse measurements and continuous queries.

Consequently, streaming field reconstruction reduces to inferring the latent trajectory $\{z_n\}_{n=1}^N$ from the incoming observation stream. We formulate this as two complementary Bayesian inference objectives. The *causal filtering* objective estimates the current latent vector using only the observations available up to the current timestamp,

$$q_n^{\text{filt}}(z_n) \;\approx\; p(z_n \mid \mathcal{O}_{1:n}), \tag{9}$$

where $\mathcal{O}_{1:n} = \{\mathcal{O}_1, \ldots, \mathcal{O}_n\}$ denotes the streaming history observed so far. A *retrospective smoothing* objective incorporates future observations,

$$q_k^{\text{smooth}}(z_k) \;\approx\; p(z_k \mid \mathcal{O}_{1:N}), \tag{10}$$

with $k = 1, \ldots, N$.

Given each incoming $\mathcal{O}_n$, TRACE (i) converts it into per-frame latent evidence via a pretrained generative prior, (ii) fuses this evidence with a state-space temporal prior through forward filtering, and (iii) retrospectively refines previously estimated latent states through backward smoothing, decoded to continuous fields via Eq. (8) whenever queried.

3.2 GENERATIVE LATENT EVIDENCE

When a new sparse off-grid observation frame $\mathcal{O}_n$ arrives, our goal is to represent it individually in latent space as probabilistic evidence. Given the pretrained decoder of Eq. (8), each observation frame induces a linear measurement operator connecting the latent vector to the observed values:

$$\mathbf{y}_n = \mathbf{G}_n z_n + \boldsymbol{\epsilon}_n, \qquad \boldsymbol{\epsilon}_n \sim \mathcal{N}(\mathbf{0}, \sigma_{\text{obs}}^2 \mathbf{I}_{M_n}), \tag{11}$$

where $\mathbf{G}_n \in \mathbb{R}^{M_n \times d}$ is the matrix whose m-th row is $\mathbf{G}(\mathbf{r}_{n,m})^\top$, $\mathbf{y}_n = (y_{n,1}, \ldots, y_{n,M_n})^\top$, and σ_{obs} is the observation-noise standard deviation; with $\mathbf{G}_n$ it fixes the Gaussian likelihood $p(\mathbf{y}_n \mid z_n)$ used by the guidance (Eq. (12)) and fusion.

The posterior $p(z_n \mid \mathcal{O}_n)$ is ill-posed without a latent prior. We therefore use a pretrained EDM Karras et al. (2022) denoiser $D_{\boldsymbol{\theta}}$ to provide the score-based prior $s_{\boldsymbol{\theta}}(z;\sigma) = \nabla_z \log p(z;\sigma)$ (Eq. (2)); the denoiser is frozen during inference (training details in Appendix C.2). Given the measurement operator $\mathbf{G}_n$, Diffusion Posterior Sampling (DPS) Chung et al. (2023) steers the reverse diffusion at inference toward $\mathbf{y}_n$-consistent latents via the observation-conditioned score

$$\nabla_{z_n} \log p(z_n \mid \mathbf{y}_n; \sigma) \;\approx\; s_{\boldsymbol{\theta}}(z_n;\sigma) - \zeta \nabla_{z_n} \big\| \mathbf{y}_n - \mathbf{G}_n \hat{z}_0(z_n;\sigma) \big\|_2^2, \tag{12}$$

where $\hat{z}_0(z_n;\sigma) = D_{\boldsymbol{\theta}}(z_n;\sigma)$ is the Tweedie estimate and $\zeta > 0$ sets the guidance strength. Running S independent guided reverse diffusions yields samples $\{z_n^{(s)}\}_{s=1}^S$.

We summarise these samples as a Gaussian by moment matching:

$$q_n^{\text{gen}}(z_n \mid \mathcal{O}_n) = \mathcal{N}\big(\hat{\boldsymbol{\mu}}_n^{\text{gen}}, \hat{\boldsymbol{\Sigma}}_n^{\text{gen}}\big), \tag{13}$$

where $\hat{\boldsymbol{\mu}}_n^{\text{gen}}$ and $\hat{\boldsymbol{\Sigma}}_n^{\text{gen}}$ are the empirical mean and covariance of $\{z_n^{(s)}\}_{s=1}^S$. We call q_n^{gen} the per-frame *generative latent evidence*.

3.3 FORWARD ONLINE FILTERING

The per-frame evidence q_n^{gen} is independent across frames and unavailable when $\mathcal{O}_n = \emptyset$. To achieve the causal filtering objective of Eq. (9), TRACE combines it with a temporal belief propagated from previous frames through a state-space Gaussian process (SS-GP) prior, so that evidence accumulated from earlier frames constrains the current estimate even when the present frame is sparse or entirely missing. The Markovian structure of SS-GP admits an incremental transition kernel, so the temporal belief evolves recursively with each incoming frame without revisiting past observations.

To this end, we place an independent Matérn-$3/2$ SS-GP prior (Eq. (5)) on each latent dimension $j = 1, \ldots, d$ with shared hyperparameters (σ_f^2, ℓ). For Matérn-$3/2$, the augmented Markov state reduces to $\boldsymbol{\xi}_n^{(j)} = (z_n^{(j)}, \dot{z}_n^{(j)})^\top \in \mathbb{R}^2$, with the latent value recovered by $z_n^{(j)} = \mathbf{H}\boldsymbol{\xi}_n^{(j)}$ under $\mathbf{H} = [1, 0]$. The corresponding discrete Gauss–Markov transition (Eq. (5)) has matrices $(\mathbf{A}_n, \mathbf{Q}_n)$ that depend only on $\Delta_n = t_n - t_{n-1}$, so the same temporal prior handles irregular streaming arrivals without retraining.

Stacking the d per-dimension states forms the joint augmented state $\boldsymbol{\xi}_n \in \mathbb{R}^{2d}$ with block-diagonal covariance. Applying Eq. (5) to each dimension propagates $(\mathbf{m}_{n-1}, \mathbf{P}_{n-1})$ to the predicted augmented state $(\mathbf{m}_n^-, \mathbf{P}_n^-)$; projecting the value components via $\mathbf{H}$ yields the latent-space predicted prior

$$p_n^-(z_n) = \mathcal{N}(\boldsymbol{\mu}_n^-, \boldsymbol{\Sigma}_n^-), \tag{14}$$

which encodes the temporal belief before assimilating the current frame.

We fuse p_n^- and q_n^{gen} via a tempered product of experts:

$$q_n^{\text{filt}}(z_n) \propto [p_n^-(z_n)]^\alpha [q_n^{\text{gen}}(z_n \mid \mathcal{O}_n)]^\beta, \tag{15}$$

with temperatures $\alpha, \beta \in (0, 1]$. This fusion is where each new observation enters the running estimate: it combines the belief propagated from earlier frames, p_n^-, with the current frame's evidence, q_n^{gen}, so information accumulates across the stream and past frames inform the present one. We keep it an approximate, tempered fusion rather than an exact Bayesian update because q_n^{gen} is a per-frame posterior, not a likelihood (Eq. (13)): it already embeds the diffusion prior $p(z)$, so a plain product

($\alpha = \beta = 1$) would double-count that prior and over-concentrate the estimate, which the temperatures counteract. As both factors are Gaussian, $q_n^{\text{filt}} = \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$ is closed-form under a diagonal approximation (Appendix A.4).

Since q_n^{filt} is defined over z_n whereas the SS-GP recursion evolves the augmented state $\boldsymbol{\xi}_n$, we perform a moment-consistent lift that replaces the predicted value marginal with q_n^{filt} while retaining the predicted conditional distribution of the derivative components. The resulting $(\mathbf{m}_n, \mathbf{P}_n)$ serves as the filtering state for the next transition; per-dimension expressions are given in Appendix A.3.

When $M_n = 0$ for a streaming frame, no generative evidence is available (q_n^{gen} is undefined). TRACE therefore skips the fusion step and directly sets $q_n^{\text{filt}} = p_n^-$, using the temporal prediction as the best causal estimate. These under-observed frames remain associated with the propagated latent trajectory and can later be refined by the retrospective smoother.

3.4 Retrospective Refinement by Backward Smoothing

The strictly causal filter cannot exploit future measurements that may reveal latent structures hidden in earlier under-observed or missing frames. Given an observation horizon N, TRACE therefore performs retrospective refinement by propagating information from later observations backward through the latent state-space dynamics, yielding the smoothing posterior $q_k^{\text{smooth}}(z_k) \approx p(z_k \mid \mathcal{O}_{1:N})$ for $k = 1, \ldots, N$. Because the temporal prior is linear-Gaussian in the augmented state and the per-frame evidence enters as the Gaussian factor of Eq. (13), this posterior is obtained exactly *for the resulting linear-Gaussian surrogate* by the Rauch–Tung–Striebel (RTS) recursion Rauch et al. (1965) (explicit update formulae in Appendix A.5).

Once the smoothed augmented trajectory is available, the same Markov structure supports prediction at any query time $t^\star \in (t_1, t_N)$ through an SS-GP bridge between the nearest smoothed neighbours; composing its mean with the decoder yields the field estimate

$$\hat{\mathcal{Y}}(\mathbf{r}^\star, t^\star) = \mathbf{G}(\mathbf{r}^\star)^\top \mathbb{E}[z(t^\star)], \tag{16}$$

for any continuous $(\mathbf{r}^\star, t^\star)$ within the observation span (bridge construction in Appendix A.5).

Computational complexity. The state-space formulation reduces temporal inference from the $\mathcal{O}(N^3)$ of batch Gaussian-process regression to linear in the stream length: the SS-GP forward filter and RTS backward smoother each scale as $\mathcal{O}(N)$, while the diagonal Gaussian fusion costs $\mathcal{O}(d)$ per frame. Diffusion posterior sampling dominates the per-frame computation; the temporal inference itself therefore introduces only linear overhead with respect to the streaming horizon. Details of the diagonal approximation are provided in Appendix A.4. Algorithm 2 in Appendix D summarises the full online loop, emitting a causal estimate as each frame arrives and running the backward RTS pass once the horizon is observed.

4 Related Work

Sparse Field Reconstruction. Continuous-coordinate representations reconstruct a field at arbitrary off-grid queries: functional generalisations of Tucker (Fang et al., 2022; 2023) and tensor-train decomposition replace discrete factors with basis functions (Luo et al., 2024b; Fang et al., 2024b; Chen et al., 2025a), while implicit neural representations such as MMGN (Luo et al., 2024a) modulate a coordinate network with per-trajectory latents. Learning-based reconstructors instead map sensor readings directly to the full field, as in Voronoi-tessellation CNNs (Fukami et al., 2021) and the attention-based Senseiver (Santos et al., 2023). All fit each frame or trajectory to its own observations, learning neither a generative prior over the field ensemble nor any accumulation of evidence across time.

Generative Priors for Field Reconstruction. A learned generative prior can complete plausible fields where observations are insufficient. Diffusion posterior sampling (Chung et al., 2023) steers a pretrained score model with likelihood gradients, advancing sparse-to-dense reconstruction (Huang et al., 2024a; Li et al., 2024; Long et al., 2025; Du et al., 2024); most closely, SDIFT (Chen et al., 2025b) pairs a functional Tucker latent with a diffusion prior and couples frames through batch Gaussian-process regression. These methods condition on a fixed observation set and reconstruct offline, presupposing the full horizon at inference rather than updating as observations arrive.

Table 1: Field-domain RMSE under structured sensing. **(a)** Temporally sparse (per-active-frame density $\rho = 1\%$): Miss (gap n=3, one active frame in four) and Blk-L (contiguous blackout of $L \in \{5, 10\}$ frames). **(b)** Spatially localized: moving window at local density $\rho_{\text{loc}} = 15\%$, with S-curve single-pass and circular sweeps of one and three loops. Lower is better; **bold** is best per column; "—" marks cells a per-frame method cannot reconstruct.

	Active Matter			Ocean			Supernova		
(a) Temporally Sparse Observations									
	Miss	Blk-5	Blk-10	Miss	Blk-5	Blk-10	Miss	Blk-5	Blk-10
LRTFR	0.413	0.258	0.461	0.205	0.163	0.177	0.561	0.430	0.503
MMGN	0.313	0.217	0.360	0.097	0.080	**0.092**	**0.318**	0.318	**0.320**
SDIFT	0.490	0.391	0.530	0.158	0.127	0.145	0.436	0.398	0.424
DBF	0.960	0.814	0.873	0.809	0.627	0.709	0.855	0.809	0.845
TRACE-Frame	—	—	—	—	—	—	—	—	—
TRACE-Filter	0.429	0.211	0.385	0.191	0.108	0.257	0.389	0.375	0.539
TRACE-Smoother	**0.290**	**0.186**	**0.332**	**0.095**	**0.059**	0.124	0.334	**0.312**	0.358
(b) Spatially Localized Observations									
	S-curve	1-loop	3-loop	S-curve	1-loop	3-loop	S-curve	1-loop	3-loop
LRTFR	0.769	0.807	0.693	0.245	0.351	0.153	0.917	0.929	0.785
MMGN	0.767	0.782	0.777	0.285	0.277	0.271	**0.321**	**0.321**	0.321
SDIFT	0.851	0.793	0.655	0.281	0.275	0.290	0.430	0.430	0.432
DBF	0.962	0.956	0.950	0.440	0.461	0.439	0.734	0.729	0.724
TRACE-Frame	0.580	0.581	0.577	0.114	0.147	0.141	0.429	0.419	0.414
TRACE-Filter	0.530	0.564	0.510	0.120	0.125	0.112	0.371	0.352	0.342
TRACE-Smoother	**0.509**	**0.537**	**0.476**	**0.097**	**0.106**	**0.088**	0.328	0.334	**0.318**

Streaming and Retrospective Assimilation. Sequential assimilation instead ingests observations over time. Classical filters pair an explicit dynamical model with Kalman recursions (Kalman, 1960; Evensen, 1994); learned variants keep the causal filtering form while embedding data-driven priors, through score-based filters (Bao et al., 2024; Si & Chen, 2025; Xiao et al., 2024), closed-form deep Bayesian filters (Tarumi et al., 2025), and online generative assimilation such as DiffDA (Huang et al., 2024b) and FlowDAS (Chen et al., 2025c). Score-based data assimilation (Rozet & Louppe, 2023), in contrast, infers a whole-trajectory posterior offline, requiring the full horizon. Few methods refine earlier estimates: the recent ForcingDAS (Jia et al., 2026), built on diffusion forcing (Chen et al., 2024), unifies streaming filtering and retrospective smoothing in one model, but is confined to dense fixed-resolution pixel grids, without support for continuous off-grid coordinates or irregular observation times; the state-space GP equivalence (Hartikainen & Särkkä, 2010) yields linear-time forward filtering and backward smoothing, underlying online tensor imputation (Fang et al., 2024a), yet in the raw data domain without a learned generative prior.

TRACE fills the gap these leave open: streaming reconstruction on continuous off-grid fields. It fuses per-frame diffusion posteriors as Gaussian latent evidence within a Matérn state-space model, whose linear dynamics let a single recursion filter causally forward and refine earlier, under-observed frames backward.

5 Experiments

5.1 Experimental Setup

Datasets. We evaluate TRACE on three spatiotemporal physical-field benchmarks spanning distinct physical systems, temporal dynamics, and spatial dimensionalities. Active Matter contains simulated rod-like particle dynamics in a Stokes fluid (256×256, $T = 24$), representing rapidly evolving laboratory-scale phenomena. Ocean consists of Pacific sound-speed reanalysis fields ($5 \times 38 \times 76$, $T = 24$), exhibiting moderately varying large-scale ocean dynamics. Supernova contains astrophysical blast-wave simulations (64^3, $T = 16$), whose temporal evolution is comparatively slow. Together, these datasets cover both 2D and 3D physical fields and range from highly dynamic to

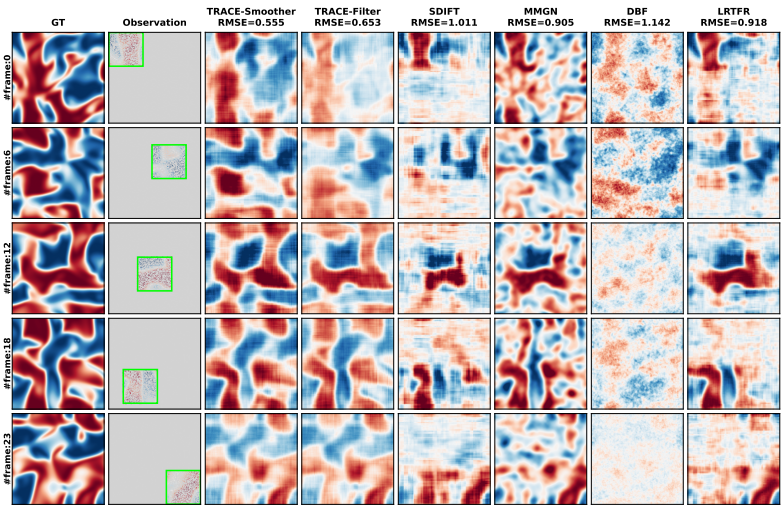
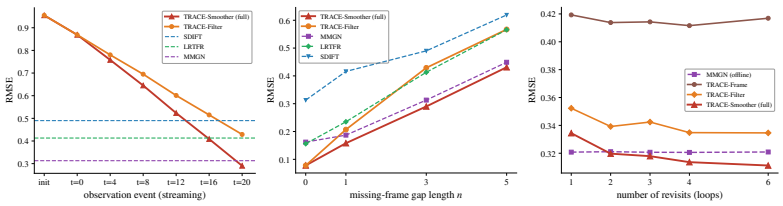


Figure 2: Qualitative reconstruction on Active Matter under a one-lap circular moving window.



(a) Online accumulation (b) Missing-frame gaps (c) Revisit frequency

Figure 3: Regime-structured analysis of TRACE.

nearly stationary regimes. Dataset preprocessing, train/test splits, and latent tensor ranks are provided in Appendix B.

Structured sensing protocols. We evaluate TRACE under two structured sensing regimes: temporal intermittency and spatial partial observability. **Temporally sparse observations** remove measurements from selected frames, requiring temporal propagation and retrospective refinement across missing intervals. **Spatially localized observations** restrict each frame to a moving local region, requiring global reconstruction by accumulating complementary spatial evidence over time.

Baselines. We compare against existing offline reconstruction methods from three paradigms: (i) **LRTFR** Luo et al. (2024b), a low-rank functional Tucker method with temporal TV regularization; (ii) **MMGN** Luo et al. (2024a), an implicit neural field reconstructor; and (iii) **SDIFT** Chen et al. (2025b), a generative spatiotemporal method combining functional Tucker latent with GP-correlated diffusion and posterior sampling. For streaming inference, we include the closest available baseline, **DBF** Tarumi et al. (2025), a deep Bayesian filtering method for state estimation from streaming observations. Additional implementation details are in Appendix E.

TRACE variants and evaluation. We evaluate three variants to isolate the contribution of each temporal inference component. **TRACE-Frame** performs independent per-frame generative reconstruction without temporal inference. **TRACE-Filter** augments it with the strictly causal SS-GP forward filter, enabling online temporal evidence accumulation. **TRACE-Smoother** further applies RTS backward smoothing to retrospectively refine earlier estimates using later observations. Accordingly, the improvement from TRACE-Frame to TRACE-Filter measures the benefit of causal temporal accumulation, while the improvement from TRACE-Filter to TRACE-Smoother quantifies the value of retrospective refinement. We report field-domain RMSE averaged over all frames and held-out test trajectories as the primary evaluation metric.

5.2 Temporally Sparse Observations

We instantiate temporally sparse sensing through two complementary patterns: *Miss*, where a repeating cycle of one observed frame followed by $n{=}3$ missing frames leaves every fourth frame active; and *Blk-*L, a single contiguous blackout of $L \in \{5, 10\}$ frames in an otherwise fully-observed stream. Active frames receive sparse off-grid observations at density $\rho = 1\%$ of the full spatial domain. Panel (a) of Table 1 reports field-domain RMSE.

Since TRACE-Frame is inapplicable on unobserved frames, temporal inference becomes decisive. TRACE-Filter provides causal online estimates via SS-GP forward propagation, but decays over extended gaps and suffers on rapidly evolving dynamics such as Active Matter. TRACE-Smoother improves upon both variants via RTS backward refinement, with gains most pronounced on Active Matter where it matches or surpasses the strongest offline baseline across all settings. On slower-varying Ocean and Supernova, offline methods such as MMGN remain competitive under the longest blackouts, as their per-frame latent codes can be linearly interpolated across the full trajectory; under rapid dynamics, such interpolation fails and TRACE-Smoother achieves the largest gains. Qualitative examples are in Appendix F.9.

5.3 Spatially Localized Observations

The second structured regime keeps every frame active but restricts observations to a moving local window, so that no individual frame provides global coverage. Reconstructing the full field therefore requires accumulating complementary spatial evidence as the window moves through the domain. Each frame observes the window interior with local density $\rho_{\text{loc}} = 15\%$. We consider three window trajectories: S-curve single-pass, circular 1-loop, and circular 3-loop; these trajectories and the dataset-specific window sizes are illustrated and detailed in Appendix C.5 (Fig. 4). Panel (b) of Table 1 reports the resulting RMSE.

TRACE-Frame already reconstructs well from the generative latent prior, and TRACE-Filter and TRACE-Smoother add the benefits of causal accumulation and retrospective refinement. On the rapidly evolving Active Matter and Ocean fields, TRACE-Smoother attains the best accuracy, with its largest margins where dynamics change quickly between observed frames. On the 3D Supernova field, MMGN stays competitive under limited revisits (slow dynamics, full-trajectory access), but as revisits increase TRACE-Smoother surpasses it by integrating spatial evidence through the temporal prior. Figure 2 shows a qualitative example.

5.4 Regime-Structured Analysis

We analyse how the observation structure shapes performance (Fig. 3). **(a) Online accumulation** (Active Matter, Miss $n{=}3$): TRACE-Smoother improves as observations accumulate, outperforming offline batch methods before the full horizon is available. **(b) Missing-frame gaps** (Active Matter): TRACE-Filter degrades with larger gaps, whereas TRACE-Smoother refines earlier frames from later evidence. **(c) Revisit frequency** (Supernova, circular loop): TRACE-Smoother improves with more revisits as the SS-GP prior integrates complementary evidence, while offline baselines are largely insensitive to it.

Additional results. Control results, robustness to sparsity and noise, hyperparameter ablations, runtime, and additional qualitative examples are provided in Appendix F.

6 Conclusion

We introduced **TRACE**, a streaming generative framework combining continuous-coordinate latent decoding, generative posterior sampling, SS-GP temporal fusion, and RTS smoothing for physical fields under structured sensing. Across active matter, ocean, and supernova benchmarks, TRACE attains the best accuracy under rapidly evolving dynamics and repeated spatial coverage, and remains competitive in slowly varying regimes. Future work includes active sensing that steers the window toward informative regions, and replacing the SS-GP prior with a learned physics-informed latent dynamics model.

Reproducibility statement

All datasets (active matter, ocean sound-speed, supernova) are public and their preprocessing, latent ranks, and train/test splits are detailed in Appendix B and Appendix C. Model architectures, training schedules, and inference hyperparameters are given in Appendix C, and full derivations of the state-space prior, generative evidence, and retrospective smoother are in Appendix A. An anonymized reference implementation is provided in the supplementary material to support reproduction of our results.

Ethics statement

This work studies reconstruction of physical fields from simulated and reanalysis scientific data. It does not involve human subjects, personal data, or sensitive attributes, and we foresee no direct ethical concerns beyond the general responsible use of generative models for scientific inference.

AI use statement

Generative AI tools were used only as general-purpose writing and coding assistants: to polish exposition and to help scaffold and refactor experiment and plotting code. They were not used to generate research ideas, derive the method, or produce experimental results. All AI-assisted text and code were reviewed, edited, and verified by the authors, who take full responsibility for the final content of this work.

References

Feng Bao, Zezhong Zhang, and Guannan Zhang. A score-based filter for nonlinear data assimilation. *Journal of Computational Physics*, 514:113207, 2024. doi: 10.1016/j.jcp.2024.113207.

Boyuan Chen, Diego Marti Monso, Yilun Du, Max Simchowitz, Russ Tedrake, and Vincent Sitzmann. Diffusion forcing: Next-token prediction meets full-sequence diffusion. In *Advances in Neural Information Processing Systems*, 2024.

Panqi Chen, Lei Cheng, Jianlong Li, Weichang Li, Weiqing Liu, Jiang Bian, and Shikai Fang. Generalized temporal tensor decomposition with rank-revealing latent-ODE. arXiv preprint arXiv:2502.06164, 2025a.

Panqi Chen, Yifan Sun, Lei Cheng, Yang Yang, Weichang Li, Yang Liu, Weiqing Liu, Jiang Bian, and Shikai Fang. Generating full-field evolution of physical dynamics from irregular sparse observations. In *Advances in Neural Information Processing Systems*, 2025b.

Siyi Chen, Yixuan Jia, Qing Qu, He Sun, and Jeffrey A. Fessler. FlowDAS: A stochastic interpolant-based framework for data assimilation. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, 2025c.

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

Pan Du, Meet Hemant Parikh, Xiantao Fan, Xin-Yang Liu, and Jian-Xun Wang. Conditional neural field latent diffusion model for generating spatiotemporal turbulence. *Nature Communications*, 15(1):10416, 2024. doi: 10.1038/s41467-024-54712-1.

Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5): 10143–10162, 1994. doi: 10.1029/94JC00572.

Shikai Fang, Akil Narayan, Robert M. Kirby, and Shandian Zhe. Bayesian continuous-time Tucker decomposition. In *International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 6235–6245, 2022.

Shikai Fang, Xin Yu, Shibo Li, Zheng Wang, Robert M. Kirby, and Shandian Zhe. Streaming factor trajectory learning for temporal tensor decomposition. In *Advances in Neural Information Processing Systems*, 2023.

Shikai Fang, Qingsong Wen, Yingtao Luo, Shandian Zhe, and Liang Sun. BayOTIDE: Bayesian online multivariate time-series imputation with functional decomposition. In *International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 12993–13009, 2024a.

Shikai Fang, Xin Yu, Zheng Wang, Shibo Li, Robert M. Kirby, and Shandian Zhe. Functional Bayesian Tucker decomposition for continuous-indexed tensor data. In *International Conference on Learning Representations*, 2024b.

Kai Fukami, Romit Maulik, Nesar Ramachandra, Koji Fukagata, and Kunihiko Taira. Global field reconstruction from sparse sensors with Voronoi tessellation-assisted deep learning. *Nature Machine Intelligence*, 3(11):945–951, 2021.

Jouni Hartikainen and Simo Särkkä. Kalman filtering and smoothing solutions to temporal Gaussian process regression models. In *IEEE International Workshop on Machine Learning for Signal Processing*, pp. 379–384, 2010. doi: 10.1109/MLSP.2010.5589113.

Jiahe Huang, Guandao Yang, Zichen Wang, and Jeong Joon Park. DiffusionPDE: Generative PDE-solving under partial observation. In *Advances in Neural Information Processing Systems*, 2024a.

Langwen Huang, Lukas Gianinazzi, Yuejiang Yu, Peter D. Dueben, and Torsten Hoefler. DiffDA: A diffusion model for weather-scale data assimilation. In *International Conference on Machine Learning (ICML), PMLR 235*, pp. 19798–19815, 2024b.

Yixuan Jia, Siyi Chen, Yida Pan, Xiao Li, Lianghe Shi, Chanyong Jung, Haijie Yuan, Ismail Alkhouri, Yue Cynthia Wu, Saiprasad Ravishankar, Jeffrey A. Fessler, and Qing Qu. Forcing-DAS: Unified and robust data assimilation via diffusion forcing. arXiv preprint arXiv:2605.14285, 2026.

Rudolf E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45, 1960. doi: 10.1115/1.3662552.

Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In *Advances in Neural Information Processing Systems*, 2022.

Peter Lancaster and Leiba Rodman. *Algebraic Riccati Equations*. Oxford University Press, 1995.

Zeyu Li, Wang Han, Yue Zhang, Qingfei Fu, Jingxuan Li, Lizi Qin, Ruoyu Dong, Hao Sun, Yue Deng, and Lijun Yang. Learning spatiotemporal dynamics with a pretrained generative model. *Nature Machine Intelligence*, 6(12):1566–1579, 2024. doi: 10.1038/s42256-024-00938-z.

Da Long, Zhitong Xu, Guang Yang, Akil Narayan, and Shandian Zhe. Arbitrarily-conditioned multi-functional diffusion for multi-physics emulation. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, 2025.

Xihaier Luo, Wei Xu, Balasubramanya Nadiga, Yihui Ren, and Shinjae Yoo. Continuous field reconstruction from sparse observations with implicit neural networks. In *International Conference on Learning Representations*, 2024a.

Yisi Luo, Xile Zhao, Zhemin Li, Michael K. Ng, and Deyu Meng. Low-rank tensor function representation for multi-dimensional data recovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(5):3351–3369, 2024b. doi: 10.1109/TPAMI.2023.3341688.

H. E. Rauch, F. Tung, and C. T. Striebel. Maximum likelihood estimates of linear dynamic systems. *AIAA Journal*, 3(8):1445–1450, 1965. doi: 10.2514/3.3166.

François Rozet and Gilles Louppe. Score-based data assimilation. In *Advances in Neural Information Processing Systems*, 2023.

Javier E. Santos, Zachary R. Fox, Arvind Mohan, Daniel O'Malley, Hari Viswanathan, and Nicholas Lubbers. Development of the Senseiver for efficient field reconstruction from sparse observations. *Nature Machine Intelligence*, 5(11):1317–1325, 2023.

Simo Särkkä and Arno Solin. *Applied Stochastic Differential Equations*. Cambridge University Press, 2019.

Phillip Si and Peng Chen. Latent-EnSF: A latent ensemble score filter for high-dimensional data assimilation with sparse observation data. In *International Conference on Learning Representations*, 2025.

Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Advances in Neural Information Processing Systems*, 2020.

Arno Solin and Simo Särkkä. Explicit link between periodic covariance functions and state space models. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, volume 33 of *Proceedings of Machine Learning Research*, pp. 904–912, 2014.

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.

Yuta Tarumi, Keisuke Fukuda, and Shin-ichi Maeda. Deep Bayesian filter for Bayes-faithful data assimilation. In *International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 59182–59209, 2025.

Pengpeng Xiao, Phillip Si, and Peng Chen. LD-EnSF: Synergizing latent dynamics with ensemble score filters for fast data assimilation with sparse observations. arXiv preprint arXiv:2411.19305, 2024.

A Mathematical Derivations

This appendix collects the derivations supporting the state-space inference machinery of Sec. 3.3. We first recall, in standard spectral form, the Hartikainen–Särkkä equivalence that turns a Matérn-ν Gaussian-process prior into an equivalent finite-dimensional LTI-SDE (§A.1). We then specialise to $\nu = 3/2$ and write out the closed-form matrices used throughout this paper (§A.2). The remaining three subsections analyse the tempered fusion introduced in Sec. 3.3: its closed form and its Kalman corollary (§A.3), the rationale for the diagonal-fusion approximation (§A.4), and finally the algorithmic components for backward smoothing and off-grid querying (§A.5). Throughout, we use the same symbols as Sec. 3.3: $\boldsymbol{\xi}(t)$ is the augmented state, $\mathbf{H}$ the projection to the latent value, $(\mathbf{A}_n, \mathbf{Q}_n)$ the discrete Gauss–Markov transition matrices, and (α, β) the tempering coefficients of the fusion.

A.1 Spectral Analysis: From the Matérn Kernel to LTI-SDE and Markov Chain

We take the Matérn kernel as an example to show how to connect a Gaussian process with a linear time-invariant stochastic differential equation (LTI-SDE). The Matérn kernel is defined as

$$\kappa_\nu(t,t') \;=\; \sigma_f^2 \,\frac{\big(\frac{\sqrt{2\nu}\,\Delta}{\ell}\big)^{\nu}}{\Gamma(\nu)\,2^{\nu-1}}\, K_\nu\Big(\tfrac{\sqrt{2\nu}\,\Delta}{\ell}\Big), \tag{17}$$

where $\Delta = |t - t'|$ is the time lag, $\Gamma(\cdot)$ is the Gamma function, $\sigma_f^2 > 0$ is the marginal variance, $\ell > 0$ is the length-scale, K_ν is the modified Bessel function of the second kind, and $\nu > 0$ controls the smoothness of sample paths from the GP prior $f(t) \sim \mathcal{GP}\big(0, \kappa_\nu(t,t')\big)$.

For a stationary Matérn kernel $\kappa_\nu(t,t') = \kappa_\nu(t-t')$, the energy spectral density of $f(t)$ can be obtained via the Wiener–Khinchin theorem by taking the Fourier transform of $\kappa_\nu(\Delta)$:

$$S(\omega) \;=\; \frac{q_s}{(\lambda^2 + \omega^2)^{m+1}}, \tag{18}$$

where ω is the frequency, $q_s > 0$ is a positive constant determined by the prefactor of (17), and we restrict to half-integer smoothness $\nu = m + \frac{1}{2}$ for $m \in \{0, 1, 2, \dots\}$. Throughout this appendix we use the frequency rescaling

$$\lambda \;:=\; \sqrt{2\nu}/\ell, \tag{19}$$

which keeps the matrices below dimensionally homogeneous. The denominator of (18) is a polynomial of degree $2(m+1)$ in ω, which is what eventually permits a finite-dimensional state-space realisation.

Expanding the polynomial $(\lambda + \imath\omega)^{m+1}$ in (18) gives

$$(\lambda + \imath\omega)^{m+1} \;=\; \sum_{k=0}^{m} c_k \, (\imath\omega)^k \;+\; (\imath\omega)^{m+1}, \tag{20}$$

where the real coefficients c_k are given by the binomial formula $c_k = \binom{m+1}{k}\lambda^{m+1-k}$. Since $(\lambda^2 + \omega^2)^{m+1} = \left|(\lambda + \imath\omega)^{m+1}\right|^2$, (20) allows us to construct an equivalent frequency-domain system whose output has PSD (18):

$$\sum_{k=0}^{m} c_k \, (\imath\omega)^k \, \widehat{f}(\omega) \;+\; (\imath\omega)^{m+1} \, \widehat{f}(\omega) \;=\; \widehat{\beta}(\omega), \tag{21}$$

where $\widehat{f}(\omega)$ and $\widehat{\beta}(\omega)$ are the Fourier transforms of $f(t)$ and of a white-noise process $w(t)$ with spectral density q_s, respectively.

Taking the inverse Fourier transform of (21) returns a stochastic differential equation in the time domain,

$$\sum_{k=0}^{m} c_k \, \frac{d^k f}{dt^k} \;+\; \frac{d^{m+1} f}{dt^{m+1}} \;=\; w(t), \tag{22}$$

in which the highest-order derivative of f is driven directly by the white noise. To remove the high-order derivative, we lift the scalar process $f(t)$ to a vector-valued state by collecting f together with its first m time derivatives,

$$\boldsymbol{\xi}(t) \;=\; \big(f(t),\, f^{(1)}(t),\, \dots,\, f^{(m)}(t)\big)^{\top}, \quad f^{(k)} := \tfrac{d^k f}{dt^k}, \tag{23}$$

so that $\boldsymbol{\xi}(t) \in \mathbb{R}^{m_\nu}$ with $m_\nu := m + 1$. Substituting (23) into (22) turns the high-order SDE into a first-order linear time-invariant SDE,

$$\frac{d\boldsymbol{\xi}(t)}{dt} \;=\; \mathbf{F}\,\boldsymbol{\xi}(t) \;+\; \mathbf{L}\,w(t), \qquad f(t) \;=\; \mathbf{H}\,\boldsymbol{\xi}(t), \tag{24}$$

in which the drift matrix $\mathbf{F} \in \mathbb{R}^{m_\nu \times m_\nu}$ and the noise input $\mathbf{L} \in \mathbb{R}^{m_\nu \times 1}$ take the companion form

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & & \\ & \ddots & \ddots & \\ & & 0 & 1 \\ -c_0 & \cdots & -c_{m-1} & -c_m \end{bmatrix}, \quad \mathbf{L} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{25}$$

together with the constant output projection $\mathbf{H} = (1, 0, \dots, 0) \in \mathbb{R}^{1 \times m_\nu}$, which extracts the original process $f(t)$ from the augmented state $\boldsymbol{\xi}(t)$. The companion structure of $\mathbf{F}$ inherits all of its eigenvalues from the roots of $P(s) = (s + \lambda)^{m+1}$ in (18); since these roots all lie at $-\lambda < 0$, $\mathbf{F}$ is Hurwitz and (24) is asymptotically stable.

The LTI-SDE (24) is particularly useful because its finite-dimensional state $\boldsymbol{\xi}(t)$ follows a Gauss–Markov chain at any set of sampled timestamps. Specifically, given arbitrary $t_1 < t_2 < \cdots < t_N$, the joint distribution of $\{\boldsymbol{\xi}(t_n)\}_{n=1}^{N}$ factorises as

$$p\big(\boldsymbol{\xi}(t_1), \dots, \boldsymbol{\xi}(t_N)\big) \;=\; p(\boldsymbol{\xi}(t_1)) \prod_{n=1}^{N-1} p(\boldsymbol{\xi}(t_{n+1}) \mid \boldsymbol{\xi}(t_n)), \tag{26}$$

in which both the initial and the transition distributions are Gaussian,

$$p(\boldsymbol{\xi}(t_1)) \;=\; \mathcal{N}(\mathbf{0},\, \mathbf{P}_\infty), \tag{27}$$

$$p(\boldsymbol{\xi}(t_{n+1}) \,|\, \boldsymbol{\xi}(t_n)) \;=\; \mathcal{N}\big(\mathbf{A}_n \boldsymbol{\xi}(t_n),\, \mathbf{Q}_n\big). \tag{28}$$

The discrete transition matrices in (28) are obtained by solving the LTI-SDE (24) on the inter-frame interval $\Delta_n := t_{n+1} - t_n$, which yields the matrix exponential

$$\mathbf{A}_n \;=\; \exp\bigl(\Delta_n \mathbf{F}\bigr), \tag{29}$$

and the time-integrated process noise

$$\begin{aligned}\mathbf{Q}_n \;&=\; \int_0^{\Delta_n} \exp(s\mathbf{F})\, \mathbf{L}\mathbf{L}^\top \exp(s\mathbf{F})^\top \, q_s \, ds \\ &=\; \mathbf{P}_\infty \;-\; \mathbf{A}_n \, \mathbf{P}_\infty \, \mathbf{A}_n^\top,\end{aligned} \tag{30}$$

in which the second equality follows from the stationary identity $\mathbf{P}_\infty = \mathbf{A}_n \mathbf{P}_\infty \mathbf{A}_n^\top + \mathbf{Q}_n$ valid for any stable LTI-SDE. The stationary covariance $\mathbf{P}_\infty$ is the symmetric positive-definite solution of the algebraic Lyapunov equation Lancaster & Rodman (1995)

$$\mathbf{F}\,\mathbf{P}_\infty \;+\; \mathbf{P}_\infty\,\mathbf{F}^\top \;+\; \mathbf{L}\mathbf{L}^\top \, q_s \;=\; \mathbf{0}, \tag{31}$$

which has a unique solution because $\mathbf{F}$ is Hurwitz, as noted above. The closed form of $\mathbf{P}_\infty$, $\mathbf{A}_n$, and $\mathbf{Q}_n$ for the specific case $\nu = 3/2$ used throughout this paper is worked out in §A.2.

Equations (17)–(31) give a complete spectral construction of the SS-GP backbone used in Sec. 3.3: the Matérn covariance is first Fourier-transformed to the rational PSD (18), which factorises through the polynomial expansion (20) into a Hurwitz frequency-domain system (21); an inverse Fourier transform brings this into the time-domain SDE (22), which is rewritten as the first-order LTI-SDE (24) via state augmentation; and discretising this LTI-SDE on any sequence of timestamps yields the Gauss–Markov chain (26)–(30) that the streaming Kalman recursion of Sec. 3.3 operates on. The construction reduces the cost of joint GP inference from $\mathcal{O}(N^3)$ to $\mathcal{O}(N)$ Hartikainen & Särkkä (2010) while preserving the original Matérn covariance exactly at the sampled times. We close this subsection with a remark on extensibility: for other stationary kernels (e.g. the periodic kernel), the inverse spectral density $1/S(\omega)$ can be approximated by a polynomial of ω^2 with negative roots Solin & Särkkä (2014), after which the same chain of steps applies.

A.2 Closed-Form Matrices for Matérn-3/2

With the canonical form of the LTI-SDE (24) and of the Gauss–Markov chain (26)–(30) derived above, we now work out the closed-form matrices for the Matérn-3/2 kernel that this paper uses throughout. The case corresponds to $\nu = 3/2$, equivalently $m = 1$ and $m_\nu = 2$ in §A.1, so the augmented state of (23) collapses to the two-dimensional vector

$$\boldsymbol{\xi}(t) \;=\; \bigl(f(t),\, \dot{f}(t)\bigr)^\top \;\in\; \mathbb{R}^2, \tag{32}$$

i.e. each scalar latent dimension $f = z^{(j)}(t)$ of the Tucker core is tracked together with its time derivative. This matches the per-dimension definition $\boldsymbol{\xi}^{(j)}(t) = (z^{(j)}(t), \dot{z}^{(j)}(t))^\top$ adopted in Sec. 3.3.

Substituting $\nu = 3/2$ into the spectral derivation (17), the Bessel form simplifies to the familiar Matérn-3/2 covariance function

$$\kappa_{3/2}(\tau) \;=\; \sigma_f^2 \Bigl(1 + \frac{\sqrt{3}\,\tau}{\ell}\Bigr) \exp\Bigl(-\frac{\sqrt{3}\,\tau}{\ell}\Bigr), \tag{33}$$

where $\tau = |t - t'|$ is the lag. Substituting $m = 1$ into the Hurwitz factor of (18) gives $P(s) = (s+\lambda)^2 = s^2 + 2\lambda s + \lambda^2$ with $\lambda = \sqrt{3}/\ell$ from (19), whose companion form (25) specialises to the 2×2 drift matrix and 2×1 noise input

$$\mathbf{F} \;=\; \begin{bmatrix} 0 & 1 \\ -\lambda^2 & -2\lambda \end{bmatrix}, \qquad \mathbf{L} \;=\; \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \tag{34}$$

together with the output projection $\mathbf{H} = (1, 0) \in \mathbb{R}^{1\times2}$ that recovers $f(t)$ from $\boldsymbol{\xi}(t)$. The white-noise spectral density of the equivalent LTI-SDE is fixed by matching the variance to σ_f^2 and evaluates to

$$q_s \;=\; 4\,\sigma_f^2\,\lambda^3. \tag{35}$$

For the stationary covariance we substitute (34)–(35) into the algebraic Lyapunov equation (31) and solve for the symmetric 2×2 matrix $\mathbf{P}_\infty$. The off-diagonal entries cancel by symmetry, and the two diagonal entries decouple into independent scalar equations, yielding the diagonal closed form

$$\mathbf{P}_\infty \;=\; \begin{bmatrix} \sigma_f^2 & 0 \\ 0 & \lambda^2\sigma_f^2 \end{bmatrix}, \tag{36}$$

so that f and $\dot{f}$ are uncorrelated at stationarity with variances σ_f^2 and $3\sigma_f^2/\ell^2$, respectively. The fact that $\mathbf{P}_\infty$ is diagonal is specific to $\nu = 3/2$ and reflects the orthogonality of the process and its derivative at stationarity for this kernel.

The discrete transition matrix $\mathbf{A}_n = \exp(\Delta_n \mathbf{F})$ from (29) admits a known closed form for $\nu = 3/2$. Since the drift matrix $\mathbf{F}$ in (34) has a single eigenvalue $-\lambda$ of multiplicity two, a Jordan decomposition delivers

$$\mathbf{A}_n \;=\; e^{-\lambda\Delta_n} \begin{bmatrix} 1+\lambda\Delta_n & \Delta_n \\ -\lambda^2\Delta_n & 1-\lambda\Delta_n \end{bmatrix}. \tag{37}$$

Substituting $\mathbf{A}_n$ from (37) and $\mathbf{P}_\infty$ from (36) into the process-noise formula (30) yields

$$\mathbf{Q}_n \;=\; \mathbf{P}_\infty \;-\; \mathbf{A}_n\,\mathbf{P}_\infty\,\mathbf{A}_n^\top, \tag{38}$$

which evaluates to a positive-definite 2×2 matrix whose entries depend smoothly on Δ_n, ℓ and σ_f^2; we do not write out the four entries of (38) since they appear only inside matrix-level Kalman operations in Sec. 3.3.

Equations (33)–(38) are the specific closed-form matrices used by every Kalman and RTS recursion in this paper. Sec. 3.3 places an independent Matérn-$3/2$ prior on each of the d scalar latent dimensions $z^{(j)}(t)$ of the Tucker core, so the augmented state of the joint latent process lives in $\mathbb{R}^{2d}$ and inherits a block-diagonal structure: the matrices $(\mathbf{F}, \mathbf{L}, \mathbf{H}, \mathbf{P}_\infty, \mathbf{A}_n, \mathbf{Q}_n)$ of (34)–(38) are replicated along the diagonal once per latent dimension. This block-diagonal structure is what makes the per-dimension Kalman recursion of Sec. 3.3 exact rather than an approximation.

A.3 Tempered PoE Fusion, Augmented-State Update, and the Kalman Corollary

We now complete the Kalman cycle of Sec. 3.3 by deriving the two operations introduced informally in the main text: the value-level tempered fusion of the SS-GP predicted prior with the diffusion posterior, and the injection of the fused belief back into the per-dimension augmented state through a Kalman measurement update. We then show that the composition of these two operations reduces to the classical Kalman measurement update of a state-space GP regression at the canonical setting $\alpha = \beta = 1$.

At each streaming frame n, TRACE combines two Gaussian beliefs over the latent value $z_n \in \mathbb{R}^d$. The first is the SS-GP predicted prior at the value level,

$$p_n^-(z_n) \;=\; \mathcal{N}\big(\boldsymbol{\mu}_n^-,\, \boldsymbol{\Sigma}_n^-\big), \tag{39}$$

obtained from the per-dimension predicted augmented states $(\mathbf{m}_n^{(j),-}, \mathbf{P}_n^{(j),-})$ of the SS-GP forward pass through the output projection $\mathbf{H} = (1, 0)$ derived in (24)–(34): the j-th component of $\boldsymbol{\mu}_n^-$ and the (j, j) entry of $\boldsymbol{\Sigma}_n^-$ are

$$\mu_{n,j}^- \;=\; \mathbf{H}\,\mathbf{m}_n^{(j),-}, \qquad \Sigma_{n,jj}^- \;=\; \mathbf{H}\,\mathbf{P}_n^{(j),-}\mathbf{H}^\top, \tag{40}$$

and the off-diagonal entries of $\boldsymbol{\Sigma}_n^-$ vanish because the per-dimension priors are independent. The second Gaussian belief is the moment-matched diffusion posterior

$$q_n^{\mathrm{gen}}(z_n) \;=\; \mathcal{N}\big(\hat{\boldsymbol{\mu}}_n^{\mathrm{gen}},\, \hat{\boldsymbol{\Sigma}}_n^{\mathrm{gen}}\big) \tag{41}$$

obtained from the S DPS samples via moment matching. Substituting the Gaussian forms (39) and (41) into the tempered PoE rule (15) and computing the log-density gives the two quadratic forms

$$\log p_n^-(z_n) \;=\; -\tfrac{1}{2}\big(z_n - \boldsymbol{\mu}_n^-\big)^\top (\boldsymbol{\Sigma}_n^-)^{-1}\big(z_n - \boldsymbol{\mu}_n^-\big) + c_1, \tag{42}$$

$$\log q_n^{\mathrm{gen}}(z_n) \;=\; -\tfrac{1}{2}\big(z_n - \hat{\boldsymbol{\mu}}_n^{\mathrm{gen}}\big)^\top (\hat{\boldsymbol{\Sigma}}_n^{\mathrm{gen}})^{-1}\big(z_n - \hat{\boldsymbol{\mu}}_n^{\mathrm{gen}}\big) + c_2, \tag{43}$$

where c_1, c_2 collect constants independent of z_n. Taking the tempered linear combination $\alpha\cdot(42) + \beta\cdot(43)$ and gathering the quadratic and linear terms in z_n yields the unnormalised log-density of the fused belief,

$$\begin{aligned}\log q_n(z_n) \;=\; & -\tfrac{1}{2}\, z_n^\top\big[\alpha(\boldsymbol{\Sigma}_n^-)^{-1} + \beta(\hat{\boldsymbol{\Sigma}}_n^{\mathrm{gen}})^{-1}\big]z_n \\ & + z_n^\top\big[\alpha(\boldsymbol{\Sigma}_n^-)^{-1}\boldsymbol{\mu}_n^- + \beta(\hat{\boldsymbol{\Sigma}}_n^{\mathrm{gen}})^{-1}\hat{\boldsymbol{\mu}}_n^{\mathrm{gen}}\big] \\ & + \mathrm{const}. \end{aligned} \tag{44}$$

Reading off the precision matrix and natural parameter from (44) identifies the fused posterior as a Gaussian $q_n(z_n) = \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$ with

$$\boldsymbol{\Sigma}_n^{-1} \;=\; \alpha\,(\boldsymbol{\Sigma}_n^-)^{-1} \;+\; \beta\,(\hat{\boldsymbol{\Sigma}}_n^{\mathrm{gen}})^{-1}, \tag{45}$$

$$\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n \;=\; \alpha\,(\boldsymbol{\Sigma}_n^-)^{-1}\boldsymbol{\mu}_n^- \;+\; \beta\,(\hat{\boldsymbol{\Sigma}}_n^{\mathrm{gen}})^{-1}\hat{\boldsymbol{\mu}}_n^{\mathrm{gen}}. \tag{46}$$

Equations (45)–(46) are the closed-form fusion invoked in (15) of the main text. Under the diagonal fusion default of §A.4, both $\boldsymbol{\Sigma}_n^-$ and $\hat{\boldsymbol{\Sigma}}_n^{\mathrm{gen}}$ are diagonal, and (45)–(46) reduce to d scalar updates of cost $\mathcal{O}(d)$.

The fused belief $q_n(z_n) = \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$ lives at the value level $z_n \in \mathbb{R}^d$, but the SS-GP forward chain operates at the augmented-state level $\boldsymbol{\xi}^{(j)} \in \mathbb{R}^2$. To close the streaming recursion we must inject q_n back into each per-dimension augmented state $(\mathbf{m}_n^{(j),-}, \mathbf{P}_n^{(j),-})$. We do so by treating the j-th component of the fused value-level mean $\boldsymbol{\mu}_n$ as a *virtual measurement* of the scalar latent $z_n^{(j)}$ with measurement noise variance equal to the corresponding diagonal entry of $\boldsymbol{\Sigma}_n$:

$$\mu_{n,j} \;\sim\; \mathcal{N}\big(z_n^{(j)},\, \Sigma_{n,jj}\big), \qquad j = 1, \ldots, d. \tag{47}$$

The d virtual measurements decouple because the diagonal fusion makes $\boldsymbol{\Sigma}_n$ diagonal. The mapping from the augmented state $\boldsymbol{\xi}_n^{(j)} \in \mathbb{R}^2$ to its scalar latent value is the same projection $\mathbf{H} = (1, 0)$ used in (40),

$$z_n^{(j)} \;=\; \mathbf{H}\,\boldsymbol{\xi}_n^{(j)}, \tag{48}$$

so (47) together with (48) defines a linear Gaussian measurement model on each augmented state with measurement matrix $\mathbf{H}$ and noise variance $\Sigma_{n,jj}$.

Applying the standard Kalman measurement update of $(\mathbf{m}_n^{(j),-}, \mathbf{P}_n^{(j),-})$ under this linear Gaussian observation produces the per-dimension augmented-state update of Sec. 3.3. The Kalman gain is

$$\mathbf{K}_n^{(j)} \;=\; \mathbf{P}_n^{(j),-}\,\mathbf{H}^\top\big(\mathbf{H}\,\mathbf{P}_n^{(j),-}\mathbf{H}^\top + \Sigma_{n,jj}\big)^{-1}, \tag{49}$$

the updated state mean is

$$\mathbf{m}_n^{(j)} \;=\; \mathbf{m}_n^{(j),-} \;+\; \mathbf{K}_n^{(j)}\big(\mu_{n,j} - \mathbf{H}\,\mathbf{m}_n^{(j),-}\big), \tag{50}$$

and the updated state covariance is

$$\mathbf{P}_n^{(j)} \;=\; \big(\mathbf{I} - \mathbf{K}_n^{(j)}\mathbf{H}\big)\mathbf{P}_n^{(j),-}. \tag{51}$$

Equations (49)–(51) give the per-dimension form of the augmented-state lift described in Sec. 3.3. The cost of one update is constant in d because all matrices in (49)–(51) are 2×2, and the d updates are run independently for the d latent dimensions, so the total per-frame cost of the augmented-state update is $\mathcal{O}(d)$. We can simplify (49) further by substituting $\mathbf{H} = (1, 0)$: the innovation variance $\mathbf{H}\mathbf{P}_n^{(j),-}\mathbf{H}^\top + \Sigma_{n,jj}$ is the scalar $\big(\mathbf{P}_n^{(j),-}\big)_{1,1} + \Sigma_{n,jj}$, and the Kalman gain $\mathbf{K}_n^{(j)} \in \mathbb{R}^{2\times1}$ is the first column of $\mathbf{P}_n^{(j),-}$ divided by this scalar.

At the canonical setting $\alpha = \beta = 1$, the composition of the value-level fusion (45)–(46) with the augmented-state update (49)–(51) reduces exactly to the classical Kalman measurement update of a state-space GP regression. To see this, set $\alpha = \beta = 1$ in (45)–(46):

$$\boldsymbol{\Sigma}_n^{-1} \;=\; (\boldsymbol{\Sigma}_n^-)^{-1} \;+\; (\hat{\boldsymbol{\Sigma}}_n^{\mathrm{gen}})^{-1}, \tag{52}$$

$$\boldsymbol{\Sigma}_n^{-1}\boldsymbol{\mu}_n \;=\; (\boldsymbol{\Sigma}_n^-)^{-1}\boldsymbol{\mu}_n^- \;+\; (\hat{\boldsymbol{\Sigma}}_n^{\mathrm{gen}})^{-1}\hat{\boldsymbol{\mu}}_n^{\mathrm{gen}}. \tag{53}$$

Equations (52)–(53) are exactly the information-form Kalman measurement update that treats q_n^{gen} as a Gaussian observation $\mathbf{y} = \hat{\boldsymbol{\mu}}_n^{\text{gen}}$ on the value-level prior $p_n^-(z_n) = \mathcal{N}(\boldsymbol{\mu}_n^-, \boldsymbol{\Sigma}_n^-)$ with measurement matrix $\mathbf{I}_d$ and noise covariance $\hat{\boldsymbol{\Sigma}}_n^{\text{gen}}$. Applying the Woodbury identity to (52) returns the equivalent covariance-form update

$$\boldsymbol{\Sigma}_n \;=\; \boldsymbol{\Sigma}_n^- \;-\; \boldsymbol{\Sigma}_n^-\big(\boldsymbol{\Sigma}_n^- + \hat{\boldsymbol{\Sigma}}_n^{\text{gen}}\big)^{-1}\boldsymbol{\Sigma}_n^-. \tag{54}$$

Substituting (52)–(54) into the per-dimension augmented-state update (49)–(51) reproduces the textbook Kalman recursion of an SS-GP regression in which the "observation" at frame n is the moment-matched DPS Gaussian summary q_n^{gen}. For readers familiar with state-space GP regression, this equivalence is visible directly from (52); we record it explicitly only to fix the correspondence between TRACE's tempered fusion plus augmented-state update and the classical Kalman terminology.

For $(\alpha, \beta) \neq (1, 1)$, the composition (45)–(51) is a strict generalisation of the Kalman measurement update. The exponents act as trust weights on each Gaussian source: smaller α down-weights the SS-GP prior under fast dynamics where the $\kappa(\Delta)$ contraction of §A.2 is too tight; smaller β down-weights the diffusion posterior when q_n^{gen} is under-determined by a near-empty observation. The validity of the closed form is unchanged: equations (45)–(51) stay well-defined and positive-definite for any $\alpha, \beta > 0$, so the same per-dimension augmented-state update is reused under any tempering schedule.

A.4 Diagonal Fusion: Rationale and Cost

The diagonal fusion of Sec. 3.3 replaces the moment-matched diffusion covariance $\hat{\boldsymbol{\Sigma}}_n^{\text{gen}}$ entering (15) by its diagonal $\mathbf{D} := \operatorname{diag}(\hat{\boldsymbol{\Sigma}}_n^{\text{gen}})$. The motivation is both statistical and computational. Because the number of DPS samples is far smaller than the latent dimension ($S \ll d$), the empirical full covariance estimated from the S samples is rank-deficient (rank at most $S-1$) and numerically unstable to invert in the high-dimensional latent space. The diagonal moment approximation avoids this inversion, reducing the per-frame fusion cost from $\mathcal{O}(d^3)$ to $\mathcal{O}(d)$ and remaining stable at the Supernova scale ($d = 32{,}768$), where forming and inverting a dense $d \times d$ precision matrix is infeasible. We examine its empirical effect—including the instability of the full-covariance variant—in Appendix F (Table 9).

A.5 RTS Smoother and Off-Grid Bridge Interpolation

After the forward filter terminates with the augmented-state means and covariances $\{(\mathbf{m}_n^{(j)}, \mathbf{P}_n^{(j)})\}_{n=1}^N$ for each latent dimension j, the Rauch–Tung–Striebel smoother Rauch et al. (1965); Särkkä & Solin (2019) delivers the smoothed posterior in a single backward sweep of cost $\mathcal{O}(N)$ per dimension. Initialising at $n = N$ with the terminal filter mean and covariance and recursing backward, the update is

$$\mathbf{G}_n^{(j)} \;=\; \mathbf{P}_n^{(j)}\,\mathbf{A}_{n+1}^{\top}\big(\mathbf{P}_{n+1}^{(j),-}\big)^{-1}, \tag{55}$$

$$\mathbf{m}_n^{(j),s} \;=\; \mathbf{m}_n^{(j)} \;+\; \mathbf{G}_n^{(j)}\big(\mathbf{m}_{n+1}^{(j),s} - \mathbf{m}_{n+1}^{(j),-}\big), \tag{56}$$

$$\mathbf{P}_n^{(j),s} \;=\; \mathbf{P}_n^{(j)} \;+\; \mathbf{G}_n^{(j)}\big(\mathbf{P}_{n+1}^{(j),s} - \mathbf{P}_{n+1}^{(j),-}\big)\mathbf{G}_n^{(j),\top}. \tag{57}$$

The smoothed augmented states are then projected through $\mathbf{H}$ to recover the smoothed latent-value posterior at every grid time t_n, supplying the RTS-corrected estimates used by Sec. 3.3.

For an unseen query time $t^\star \in (t_k, t_{k+1})$, the latent value $z(t^\star)$ is not directly produced by the filter or the smoother; we must instead infer it from the smoothed boundary states $(\boldsymbol{\mu}_k, \boldsymbol{\mu}_{k+1})$ at the neighbouring grid times. The Markov property of the SS-GP gives this inference a closed form, derived below.

Markovianity factorises the joint density of $(\boldsymbol{\xi}_k, \boldsymbol{\xi}^\star, \boldsymbol{\xi}_{k+1})$ given $\mathcal{O}_{1:n}$ into two Gauss–Markov transitions of the form (28):

$$p(\boldsymbol{\xi}^\star \,|\, \boldsymbol{\xi}_k) \;=\; \mathcal{N}\big(\mathbf{A}_1\boldsymbol{\xi}_k,\, \mathbf{Q}_1\big), \tag{58}$$

$$p(\boldsymbol{\xi}_{k+1} \,|\, \boldsymbol{\xi}^\star) \;=\; \mathcal{N}\big(\mathbf{A}_2\boldsymbol{\xi}^\star,\, \mathbf{Q}_2\big), \tag{59}$$

where the matrices $(\mathbf{A}_i, \mathbf{Q}_i)$ are obtained from (37)–(38) at the sub-intervals $\Delta_1 = t^\star - t_k$ and $\Delta_2 = t_{k+1} - t^\star$, respectively.

Conditioning the joint (58)–(59) on the smoothed boundary states $(\boldsymbol{\mu}_k, \boldsymbol{\mu}_{k+1})$ and reading off the quadratic in $\boldsymbol{\xi}^\star$ produces a Gaussian $q(z(t^\star)) = \mathcal{N}(\boldsymbol{\mu}^\star, \mathbf{V}^\star)$. Completing the square gives the precision and natural parameter in closed form,

$$\left(\mathbf{V}^\star\right)^{-1} = \mathbf{Q}_1^{-1} + \mathbf{A}_2^\top \mathbf{Q}_2^{-1} \mathbf{A}_2, \tag{60}$$

$$\left(\mathbf{V}^\star\right)^{-1} \boldsymbol{\mu}^\star = \mathbf{Q}_1^{-1} \mathbf{A}_1 \boldsymbol{\mu}_k + \mathbf{A}_2^\top \mathbf{Q}_2^{-1} \boldsymbol{\mu}_{k+1}, \tag{61}$$

from which the field-level prediction of Eq. (16) is obtained.

Composing the bridge mean $\boldsymbol{\mu}^\star$ from (61) with the FTM decoder $\mathbf{G}(\mathbf{r}^\star)$ delivers a field prediction at the arbitrary off-grid query $(\mathbf{r}^\star, t^\star)$:

$$\hat{\mathcal{Y}}(\mathbf{r}^\star, t^\star) = \mathbf{G}(\mathbf{r}^\star)^\top \boldsymbol{\mu}^\star, \tag{62}$$

in agreement with the field query of Eq. (16) in §3.4.

B Dataset Sources and Physical Context

Table 2: Per-dataset architecture configuration: Tucker multilinear rank R, vectorised-core latent dimension $d = \prod_k R_k$, and SIREN basis frequency ω.

	AM	**Ocean**	**SN**
Tucker rank R	$(1, 48, 48)$	$(5, 20, 20)$	$(32, 32, 32)$
Latent dim. d	2304	2000	32768
SIREN ω	20	10	20

This section provides physical context, data sources, and dataset-level statistics for the three benchmarks.

Active Matter (AM). A two-dimensional active-nematic continuum field describing the collective dynamics of rod-like active particles suspended in a Stokes fluid. Each frame is a scalar field over the unit square (256×256, $T=24$). We use 900 training trajectories and 28 held-out test trajectories, with a fixed 10% training mask. Data are obtained from the *active_matter* collection of The Well.[1]

Ocean. Pacific sound-speed reanalysis fields covering five depth layers, derived from the HYCOM operational reanalysis product.[2] Each frame is a layered scalar field over a fixed latitude–longitude grid ($5\times38\times76$, $T=24$). We use 950 training trajectories and 50 held-out test trajectories, with a fixed 10% training mask.

Supernova (SN). Three-dimensional temperature evolution of a supernova blast wave propagating through a dense monatomic ideal-gas cloud (64^3, $T=16$). We use 370 training trajectories and 26 held-out test trajectories, with a fixed 15% training mask. Data are obtained from the *supernova_explosion_64* collection of The Well.[3]

C Implementation and Hyperparameters

All models are trained and evaluated with PyTorch 2.9 on a single NVIDIA RTX 5090 (32 GB). EDM training uses FP32 throughout; FP16 diverges in the score network at our scale.

[1] https://polymathic-ai.org/the_well/datasets/active_matter/

[2] https://www.hycom.org/

[3] https://polymathic-ai.org/the_well/datasets/supernova_explosion_64/

Table 3: Inference hyperparameters per dataset and observation regime. Shared across all runs: S=20, N=100, σ_{obs}=0.05, β=1, diagonal fusion.

Dataset	Regime	ζ	(α, ℓ, σ_f)
AM	Control	0.001	$(0.5, 1.0, 2.0)$
	Temporally Sparse	0.001	$(0.5, 5.0, 1.0)$
	Spatially Localized	0.01	$(0.5, 5.0, 1.0)$
Ocean	Control	0.06	$(0.5, 2.0, 1.0)$
	Temporally Sparse	0.06	$(0.5, 5.0, 2.0)$
	Spatially Localized	0.2	$(0.5, 5.0, 2.0)$
SN	Control	0.02	$(0.5, 2.0, 1.0)$
	Temporally Sparse	0.02	$(0.3, 5.0, 1.0)$
	Spatially Localized	0.02	$(0.7, 20, 2.0)$

C.1 FTM Pretraining

The shared continuous-coordinate basis $\{\phi_{\theta_k}^{(k)}\}$ is a per-mode SIREN Sitzmann et al. (2020) with frequency ω, jointly trained with the per-trajectory Tucker cores by minimising the masked reconstruction loss on the training split (mask ratio per dataset reported in Appendix B). The full training objective is

$$\mathcal{L}_{\text{FTM}} = \mathbb{E}_{(\mathbf{r},t,y)\sim\mathcal{D}}\big\|y - \hat{\mathcal{Y}}(\mathbf{r}, t)\big\|_2^2 + \beta_{\text{TV}}\,\mathbb{E}_b \sum_{n=2}^{N^{(b)}} \|\mathcal{W}_n - \mathcal{W}_{n-1}\|_F^2, \tag{63}$$

where the first term is evaluated only at observed off-grid coordinates, and the temporal TV regulariser penalises frame-to-frame drift of the Tucker core in the Frobenius norm with weight $\beta_{\text{TV}} = 10^{-7}$. The basis parameters and cores are alternately updated by Adam with learning rate 2×10^{-4}. Per-dataset multilinear ranks R and SIREN frequencies ω are reported in Table 2. The basis is frozen after pretraining and reused by all functional-Tucker methods.

C.2 Latent Diffusion (EDM) Pretraining

We adopt the EDM framework Karras et al. (2022) with an identity reducer (the latent is the vectorised Tucker core). The denoiser $D_{\boldsymbol{\theta}}$ is a Song-UNet Song et al. (2021) internally parameterised over the tensor form $\mathcal{W}_n = \text{vec}^{-1}(z_n)$ to preserve the mode-wise spatial structure of the Tucker core; the training objective is the standard EDM weighted reconstruction loss

$$\mathcal{L}_{\text{EDM}} = \mathbb{E}_{z,\sigma,\boldsymbol{\epsilon}}\Big[\lambda(\sigma)\,\big\|D_{\boldsymbol{\theta}}(z + \sigma\boldsymbol{\epsilon}; \sigma) - z\big\|_2^2\Big], \tag{64}$$

with noise-level weighting $\lambda(\sigma) = (\sigma^2 + \sigma_{\text{data}}^2)/(\sigma \cdot \sigma_{\text{data}})^2$ Karras et al. (2022). The cores are z-scored before training; critically, σ_{data} is read from the post-normalisation training statistics ($\sigma_{\text{data}} \approx 1.0$) rather than from the raw core standard deviation, as the latter produces large reconstruction errors. Training uses Adam (learning rate 2×10^{-4}, batch size 128, EMA half-life 500 kimg) for 500 epochs. All training is in FP32; FP16 diverges in the score network at our scale.

C.3 SS-GP and DPS Inference Setup

The temporal prior is a state-space Gaussian process with a Matérn-$3/2$ kernel, parameterised by a length-scale ℓ and marginal standard deviation σ_f, combined with the diffusion prior through a tempered closed-form Gaussian fusion with coefficients (α, β). Throughout we fix the diffusion-side temper $\beta = 1$ and tune only the prior-side temper α, so α alone controls how strongly the temporal prior is trusted ($\alpha = 0$ recovers per-frame DPS). We use the diagonal fusion variant throughout (§F.3 compares fusion variants). The length-scale ℓ governs how strongly past frames inform the present: a short ℓ suffices in the control regime where every frame is observed, whereas a longer ℓ is essential to bridge the unobserved intervals of the temporally sparse and spatially localized regimes. This dependence is not oracle tuning: ℓ is set from the *observable* missing rate rather than the target, and Fig. 5 shows performance is governed by ℓ with a broad plateau (so a single ℓ=5 is near-optimal across all sparse settings) and is largely insensitive to α.

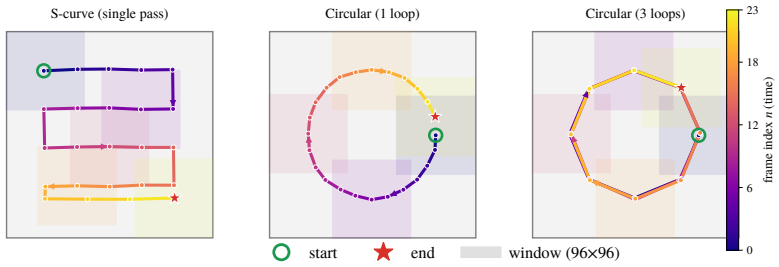


Figure 4: Moving-window trajectories on Active Matter (256×256 domain, 96×96 window). The centre path (colour-coded by time) and shaded squares mark representative window positions; circular revisits accumulate coverage over time.

For the per-frame measurement injection we use the batched diffusion posterior sampling Chung et al. (2023) implementation, which batches all $S = 20$ posterior samples in one network call at each of the $N = 100$ reverse-diffusion steps, measurement-noise scale $\sigma_{\text{obs}} = 0.05$, and a guidance step size ζ set per dataset and observation regime (Table 3). The sample count $S = 20$ is chosen to sit on the accuracy plateau of the Gaussian moment summary. In a controlled ablation on Active Matter (Control regime, $\rho = 1\%$, a 10-trajectory development subset, identical configuration), the field-domain RMSE is 0.083 at $S = 5$, 0.081 at $S = 20$, and 0.080 at $S = 50$ (the S=50 point combines the cached 20 samples with 30 freshly drawn independent ones): increasing S beyond 20 changes RMSE by under 0.0005, i.e. well inside the ≈ 0.03 cross-trajectory standard deviation. (The 0.081 here is on this 10-trajectory subset; the 0.078 reported for TRACE-Frame in the control-regime table (Table 5) is over the full 28-trajectory test set at the same $S = 20$.) We therefore keep $S = 20$, which retains the second-moment information while keeping cost linear in S; the same negligible S=20$\rightarrow$50 gain holds on Supernova and the moving-window setting.

C.4 Per-Dataset Hyperparameter Table

Table 3 summarises the inference hyperparameters. The DPS guidance step ζ is fixed once per dataset and regime. The tempering coefficient α is set to 0.5 as the default, which engages the SS-GP transition prior; the two Supernova settings deviate from this default ($\alpha = 0.3$ under the temporal regime and $\alpha = 0.7$ for the spatial moving-cube) due to the weak frame-to-frame dynamics of the Supernova field. The length-scale ℓ is the principal parameter we increase for missing/blackout and moving-window settings, so that the prior can bridge longer unobserved intervals. Setting $\alpha = 0$ would disable the transition prior and reduce the method to per-frame DPS. In our experiments hyperparameters are fixed per evaluation setting to allow controlled comparison; in deployment one would either choose the conservative regime (longer ℓ) or estimate the missing-frequency regime online — we leave this adaptive tuning to future work.

C.5 Moving-Window Trajectories

Figure 4 illustrates the three trajectory types on the Active Matter domain. Each dataset uses the same three patterns (S-curve, circular 1-loop, circular 3-loop) adapted to its dimensionality, with local observation density $\rho_{\text{loc}} = 15\%$ inside the window on all datasets. Full sweep ranges (up to 6 loops) are documented alongside results in Appendix F.

Active Matter (256×256, T=24; window 96×96). The S-curve follows a serpentine raster: the window sweeps left-to-right along one axis, steps down by half a window width, and reverses direction, covering the domain in a single pass. The circular trajectory moves the window centre along a closed elliptical orbit; each additional lap revisits previously-covered regions. The main text reports 1- and 3-lap results.

Ocean ($5\times38\times76$, T=24; 3D window $5\times24\times24$, i.e. the full 5-layer depth and a 24×24 lateral tile sliding on the 38×76 lateral plane). The window centre follows the same serpentine and elliptical patterns in the lateral dimensions while all five depth layers are observed at every position.

Algorithm 1 Offline Pretraining of TRACE

Input: Training corpus $\mathcal{D} = \{\mathcal{O}^{(b)}\}_{b=1}^{B}$; multilinear ranks $\{R_k\}_{k=1}^{K}$; EDM noise schedule $p(\sigma)$.

Output: Frozen FTM basis $\{\boldsymbol{\phi}_{\theta_k}^{(k)}\}_{k=1}^{K}$ and pretrained latent denoiser $D_{\boldsymbol{\theta}}$.

Stage I: Continuous latent representation learning.
  Initialise basis $\{\theta_k\}$ and cores $\{\mathcal{W}_n^{(b)}\}_{b,n}$.
repeat
  Sample minibatch $(\mathbf{r}, t, y) \sim \mathcal{D}$; reconstruct $\hat{\mathcal{Y}}(\mathbf{r}, t)$ via Eq. (7); update by minimising $\mathcal{L}_{\mathrm{FTM}}$.
until convergence
Freeze basis; construct latent dataset $\mathcal{Z} = \{z_n^{(b)} = \mathrm{vec}(\mathcal{W}_n^{(b)})\}_{b,n}$.
Stage II: Latent diffusion prior learning.
repeat
  Sample $z \sim \mathcal{Z}$, $\sigma \sim p(\sigma)$, $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$; update $D_{\boldsymbol{\theta}}$ by minimising $\mathcal{L}_{\mathrm{EDM}}$.
until convergence
return $\{\boldsymbol{\phi}_{\theta_k}^{(k)}\}$ and $D_{\boldsymbol{\theta}}$.

C.5.1 SUPERNOVA 3D MOVING-WINDOW TRAJECTORIES

For Supernova, the moving window is a 32^3 cube within the 64^3 domain. The cube centre follows a three-dimensional trajectory with two variants. In the 3D S-curve, the cube scans the W–D plane in a serpentine pattern while advancing linearly through the H dimension; the number of rows is determined by the cube size and domain width. In the 3D circular trajectory, the cube centre traces an elliptical orbit in the W–D plane while the H coordinate follows a sinusoidal bounce at half the orbital frequency, producing a helical path. The number of loops n_{loops} controls the angular velocity ($2\pi n_{\text{loops}}/T$) and thus the revisit frequency; the main text reports $n_{\text{loops}} = 1, 3$, with sweeps up to 6 in the appendix. At $T = 16$ frames, one loop yields approximately 6 pixels of lateral displacement per frame.

D ALGORITHMIC PSEUDOCODE

Algorithms 1 and 2 outline the offline pretraining and the online streaming inference of TRACE. All symbols follow the notation of Section 3; implementation details are provided in Appendix C.

E BASELINE REPRODUCTION DETAILS

We detail how each baseline is reproduced, the adaptations required to apply it to off-grid sparse field reconstruction, and the per-method tuning protocol. All methods observe identical scattered masks under each setting and are evaluated on the same held-out trajectories, with each baseline tuned over its principal hyperparameters.

E.1 SHARED-BASIS PROTOCOL

The tensor- and diffusion-based reconstructors are placed on the same footing as our framework by decoding through the *same* pretrained continuous-coordinate basis $\{\boldsymbol{\phi}_{\theta_k}^{(k)}\}$. Concretely, **LRTFR** solves only for the per-frame Tucker cores under a temporal total-variation penalty while keeping our frozen basis fixed (the basis frequency ω matches our FTM pretraining), and **SDIFT** decodes its sampled cores through the identical basis. This isolates the inference algorithm–temporal total variation (LRTFR), batch GP-sequential diffusion (SDIFT), or our streaming SS-GP prior–as the only variable. SDIFT trains its latent diffusion prior on the same shared cores, following its released normalisation and sampling configuration. **MMGN** and **DBF** do not use a Tucker factorisation and retain their native representations.

Algorithm 2 Online Streaming Inference of TRACE

Input: Frozen FTM decoder $\mathbf{G}(\cdot)$ and EDM denoiser $D_{\boldsymbol{\theta}}$; streaming observations $\{\mathcal{O}_n\}_{n=1}^{N}$; SS-GP transitions $\{(\mathbf{A}_n, \mathbf{Q}_n)\}$ from (σ_f, ℓ) and Δ_n; temperatures (α, β); DPS budget (S, ζ).

Output: Causal field estimates $\{\hat{\mathcal{Y}}(\cdot, t_n)\}$; filtering posteriors $\{q_n^{\text{filt}}\}$ and smoothed posteriors $\{q_n^{\text{smooth}}\}$; continuous predictor.

Initialise augmented state $(\mathbf{m}_0, \mathbf{P}_0)$ at the stationary prior.
Forward causal pass.
for $n = 1$ to N **do**
  SS-GP predict: $\mathbf{m}_n^- = \mathbf{A}_n \mathbf{m}_{n-1}, \quad \mathbf{P}_n^- = \mathbf{A}_n \mathbf{P}_{n-1} \mathbf{A}_n^\top + \mathbf{Q}_n$.
  Project to the latent prior $p_n^- = \mathcal{N}(\boldsymbol{\mu}_n^-, \boldsymbol{\Sigma}_n^-)$ via $\mathbf{H}$ (Eq. (14))
  if $\mathcal{O}_n \neq \emptyset$ **then**
    Draw S samples $\{z_n^{(s)}\}_{s=1}^{S}$ by guided reverse diffusion (Eq. (12))
    Moment-match to the evidence $q_n^{\text{gen}} = \mathcal{N}(\hat{\boldsymbol{\mu}}_n^{\text{gen}}, \hat{\boldsymbol{\Sigma}}_n^{\text{gen}})$ (Eq. (13))
    Fuse precisions: $\boldsymbol{\Sigma}_n^{-1} = \alpha(\boldsymbol{\Sigma}_n^-)^{-1} + \beta(\hat{\boldsymbol{\Sigma}}_n^{\text{gen}})^{-1}$
    Fuse means: $\boldsymbol{\mu}_n = \boldsymbol{\Sigma}_n\big(\alpha(\boldsymbol{\Sigma}_n^-)^{-1}\boldsymbol{\mu}_n^- + \beta(\hat{\boldsymbol{\Sigma}}_n^{\text{gen}})^{-1}\hat{\boldsymbol{\mu}}_n^{\text{gen}}\big)$ (Eq. (15))
    Set $q_n^{\text{filt}} = \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)$.
  else
    $q_n^{\text{filt}} \leftarrow p_n^-$ (missing frame: predict-only)
  end if
  Lift q_n^{filt} to the augmented state $(\mathbf{m}_n, \mathbf{P}_n)$ (App. A.3)
  Emit causal field $\hat{\mathcal{Y}}(\mathbf{r}, t_n) = \mathbf{G}(\mathbf{r})^\top \boldsymbol{\mu}_n$.
end for
Backward retrospective pass (once the horizon N is available).
for $n = N - 1$ down to 1 **do**
  Smoother gain: $\mathbf{G}_n = \mathbf{P}_n \mathbf{A}_{n+1}^\top (\mathbf{P}_{n+1}^-)^{-1}$.
  $\mathbf{m}_n^{\text{s}} = \mathbf{m}_n + \mathbf{G}_n(\mathbf{m}_{n+1}^{\text{s}} - \mathbf{m}_{n+1}^-)$.
  $\mathbf{P}_n^{\text{s}} = \mathbf{P}_n + \mathbf{G}_n(\mathbf{P}_{n+1}^{\text{s}} - \mathbf{P}_{n+1}^-)\mathbf{G}_n^\top$ (App. A.5)
  Re-emit refined field $\hat{\mathcal{Y}}(\mathbf{r}, t_n) = \mathbf{G}(\mathbf{r})^\top \boldsymbol{\mu}_n^{\text{s}}$.
end for
Continuous query. For any $(\mathbf{r}^\star, t^\star)$, an SS-GP bridge between the nearest smoothed knots gives $\mathbb{E}[z(t^\star)]$; return $\hat{\mathcal{Y}}(\mathbf{r}^\star, t^\star) = \mathbf{G}(\mathbf{r}^\star)^\top \mathbb{E}[z(t^\star)]$ (App. A.5)

E.2 MMGN and the Interpolation Extension

MMGN fits an implicit neural field independently at each frame and, by construction, produces a meaningful reconstruction only for frames that carry observations; on a frame without measurements its per-frame latent is undetermined. To extend MMGN to unobserved frames—rather than penalise it with an empty output—we exploit the temporal continuity of its latent code and linearly interpolate the codes of the nearest observed frames. This extension *strengthens* the baseline: on Active Matter ($\rho = 1\%$), interpolation reduces the RMSE under missing-frame settings from 0.562 to 0.187 and from 0.762 to 0.313 as the missing interval widens, relative to the native zero-latent variant. We therefore report the interpolated variant as the stronger MMGN baseline throughout.

E.3 SDIFT

SDIFT is trained following its released configuration (effective batch size 16 via gradient accumulation) and its message-passing posterior sampling (MPDPS) is tuned over its guidance strength, step count, and ζ schedule. For the 3D Supernova field we adapt its sampling pipeline to the cubic moving-window observations.

E.4 DBF and the Observation-Density Protocol

The deep Bayesian filter (DBF) is an end-to-end causal filter that does not use a Tucker factorisation. Two adaptations were required for a fair reproduction. First, its decoder applies a terminal ReLU,

Table 4: DBF at its matched-density upper bound ($\rho = 3\%$ for both training and testing, no missing frames) versus our framework evaluated at the same $\rho = 3\%$.

Full-availability RMSE	Ocean	Active Matter	Supernova
DBF (matched $\rho = 3\%$)	0.346	0.473	0.681
TRACE-Smoother (ours)	0.040	0.066	0.298
ratio	8.8×	7.1×	2.3×

which the original DBF paper validates on strictly non-negative fluid simulations and therefore reflects a domain-specific representation choice rather than a method-essential constraint; on our zero-centred fields the same ReLU collapses the output to the mean field, so we remove this activation and z-score the targets, after which DBF produces non-trivial reconstructions (its observation-driven correlation substantially exceeds that of a mean-field template). Second, DBF requires a minimum training density. Trained at our evaluation density $\rho = 1\%$, optimisation *does* proceed but converges to a degenerate solution: a near-constant mean-field predictor (RMSE ≈ 1.1, no better than predicting the dataset mean), so no useful filter is learned. We therefore train it at the lowest density at which it converges to a genuine, observation-dependent reconstruction ($\rho = 3\%$) and evaluate across all test densities. This necessarily introduces a train/test density mismatch (train $\rho = 3\%$, test $\rho = 1\%$). To show the mismatch does not unfairly penalise DBF, Table 4 additionally reports DBF trained *and* tested at $\rho = 3\%$—its strongest attainable setting. Even there it trails our framework by a wide margin with full temporal availability (e.g. 0.473 vs. 0.066 on Active Matter, a $7.1\times$ gap, and a smaller $2.3\times$ gap on the high-rank Supernova), confirming that the shortfall reflects model capacity rather than the density mismatch.

E.5 Streaming Filters That Failed to Reproduce

We also attempted score-based ensemble filters, in both their native and latent forms.

EnSF is training-free but, lacking a learned field prior, collapses under the curse of dimensionality when run directly in the field space ($d = 65{,}536$ on Active Matter, $d = 14{,}440$ on Ocean): the ensemble cannot represent the posterior and the update suffers frequent numerical breakdown (NaN). Enlarging the ensemble from 50 to $2{,}000$ particles does not help—RMSE stays at ≈ 1.15 on Active Matter ($1.147 \to 1.225$) and ≈ 1.03 on Ocean, i.e. no better than the trivial mean-field reconstruction (field std ≈ 1.0). This is consistent with the EnSF design assumption of strong, low-effective-rank PDE structure (e.g. Lorenz-96); we report it as an empirical reproduction failure rather than a definitive diagnosis of its cause.

LD-EnSF instead assimilates in a learned latent space whose dynamics are a neural ODE (an LDNet); this learned forward model failed to fit our data. On Active Matter the LDNet validation relative error stagnated at ≈ 1.0 (no better than the mean field) after training on 855 trajectories, and in a controlled sanity check it could not overfit even two trajectories (relative error plateaued at ≈ 0.71). Without a usable latent forward model its assimilation stage cannot be run. We therefore exclude both filters from the main comparison and report them here for completeness.

F Additional Experimental Results

We use the main-text method names throughout: **TRACE-Frame** (moment-matched $S{=}20$ per-frame posterior, no temporal prior), **TRACE-Filter** (causal SS-GP filter, no smoothing), and **TRACE-Smoother** (full model with RTS smoothing); baselines are **LRTFR**, **MMGN**, **SDIFT**, and **DBF**. A dash (—) marks cells a method cannot produce (TRACE-Frame on unobserved frames) or that were not run. All numbers are field-domain RMSE averaged over test trajectories.

F.1 Control Setting: Dense-Time Full-Domain Sparse Observation

We begin with a *control* regime in which every frame is active and independently observed by an off-grid random sample over the whole domain ($\Omega_{\mathrm{obs}}^{(n)} = \Omega$ for all n, density ρ). This setting serves as a calibration: with each frame already individually informative, it bounds how much of the

Table 5: Control regime (frame-rich: every frame observes the full domain at density ρ), field-domain RMSE. Lower is better; **bold** is best per column.

	Active Matter		Ocean		Supernova	
	1%	3%	1%	3%	1%	3%
LRTFR	0.156	0.090	0.161	0.114	0.406	0.418
MMGN	0.162	0.148	0.070	0.047	0.318	0.313
SDIFT	0.313	0.220	0.112	0.078	0.379	0.342
DBF	0.747	0.473	0.527	0.346	0.771	0.681
TRACE-Frame	0.078	0.066	0.045	0.040	0.322	0.309
TRACE-Filter	0.078	0.066	0.047	0.042	0.321	0.308
TRACE-Smoother	**0.078**	**0.066**	**0.044**	**0.040**	**0.310**	**0.298**

Table 6: Missing pattern with gap $n{=}1$ (every other frame active), field-domain RMSE at $\rho \in \{1\%, 3\%\}$. **Bold** is best per column.

	Active Matter		Ocean		Supernova	
	1%	3%	1%	3%	1%	3%
LRTFR	0.235	0.178	0.175	0.136	0.460	0.424
MMGN	0.187	0.178	0.073	0.055	0.317	0.313
SDIFT	0.416	0.403	0.154	0.147	0.409	0.370
DBF	0.907	0.755	0.713	0.587	0.829	0.771
TRACE-Frame	—	—	—	—	—	—
TRACE-Filter	0.207	0.198	0.082	0.079	0.335	0.322
TRACE-Smoother	**0.158**	**0.146**	**0.055**	**0.051**	**0.312**	**0.301**

reconstruction quality is attributable to the generative field prior versus to cross-time accumulation. Table 5 reports field-domain RMSE at $\rho \in \{1\%, 3\%\}$.

TRACE-Smoother achieves the best RMSE on all three datasets at both observation densities, outperforming every offline reconstructor. Decomposing the contribution of each component, the per-frame prior (TRACE-Frame) alone already surpasses or matches all offline baselines, indicating that the latent diffusion prior drives most of the gain when frames are individually informative. Adding the causal SS-GP filter offers only marginal improvement, as propagating an already-accurate per-frame estimate may accumulate minor error; the RTS smoother consistently corrects this and yields a small but reliable improvement. The overall temporal contribution in this dense-time regime is thus modest—precisely as expected, since cross-time accumulation is most valuable when individual frames are under-informative, as demonstrated in the structured sensing regimes of the main text.

F.2 Temporally Sparse Observations: Additional Results

We supplement the main-text Temporally Sparse experiments (Table 1) along two axes: a *shorter* gap (Table 6) and a *higher* observation density (Table 7). Both confirm that the honest boundary in Sec. 5.4 is confined to the longer unobserved intervals of Table 1.

Shorter gap (Miss-1). Table 6 reports the Missing pattern with gap $n{=}1$ (every other frame active) on all three datasets at $\rho \in \{1\%, 3\%\}$. TRACE-Smoother attains the lowest RMSE in every column.

Higher density ($\rho = 3\%$). Table 7 repeats the main-text Table 1 grid at $\rho = 3\%$. The ranking is preserved across datasets, confirming that the main-text conclusions are stable with respect to observation density.

Table 7: Temporally sparse observations at $\rho = 3\%$, field-domain RMSE. Columns follow main-text Table 1 (*Miss*: gap n=3; *Blk-L*: blackout length L). **Bold** is best per column.

	Active Matter			Ocean			Supernova		
	Miss	Blk-5	Blk-10	Miss	Blk-5	Blk-10	Miss	Blk-5	Blk-10
LRTFR	0.377	0.221	0.442	0.177	0.128	0.142	0.450	0.425	0.445
MMGN	0.307	0.205	0.352	**0.084**	0.060	**0.075**	**0.314**	0.313	**0.316**
SDIFT	0.478	0.366	0.526	0.148	0.110	0.139	0.399	0.356	0.397
DBF	0.901	0.617	0.742	0.747	0.502	0.627	0.829	0.747	0.812
TRACE-Frame	—	—	—	—	—	—	—	—	—
TRACE-Filter	0.426	0.200	0.376	0.190	0.104	0.254	0.377	0.364	0.530
TRACE-Smoother	**0.285**	**0.173**	**0.321**	0.093	**0.056**	0.121	0.323	**0.300**	0.349

Table 8: Effect of the transition prior under full observation (Control regime, $\rho = 1\%$), field-domain RMSE. "DPS (S=1)" is vanilla single-sample DPS.

	Active Matter	Ocean	Supernova
DPS (S=1)	0.091	0.055	0.365
TRACE-Frame	0.078	0.045	0.322
TRACE-Smoother	**0.078**	**0.044**	**0.310**

F.3 Ablation: Prior Contribution and Fusion Variants

Table 8 isolates two contributions under full per-frame observation. The S=1 → 20 gap quantifies the multi-sample moment-estimation gain of the Gaussian moment summary; the further gap to TRACE-Smoother reflects the modest temporal contribution in the Control regime (Sec. 5.3). Temporally-sparse behaviour of the same three variants is in Table 1 (main text) and Table 6.

Table 9 compares diagonal against full-covariance fusion on Active Matter, timing the closed-form step alone. The diagonal variant is $\mathcal{O}(d)$ in compute and memory, costs 1.7 ms per frame, and runs on all 28 trajectories. The full-covariance variant forms and inverts a dense $d \times d$ precision matrix ($\mathcal{O}(d^3)$ compute, $\mathcal{O}(d^2)$ memory), is $\gtrsim$20 s per frame at d=2304, and diverges on 2 of 5 trajectories: the rank-$S-1$=19 empirical covariance from S=20 samples is near-singular in the d=2304 space, so even ridge-regularised inversion is numerically unstable. At the Supernova scale (d=32768) the $\approx$ 4 GB dense covariance exhausts memory outright. This motivates the diagonal default; see Appendix A.4 for the rationale.

Posterior Gaussian summary: sample count S. Table 10 sweeps S on Active Matter (Control, $\rho = 1\%$): RMSE plateaus from $S \geq 20$, justifying the default used throughout (Appendix C).

F.4 Sensitivity: Tempering and Lengthscale

TRACE-Smoother exposes two SS-GP hyperparameters: the Matérn length-scale ℓ and the tempering coefficient α. Figure 5 sweeps each on Active Matter ($\rho = 1\%$) across four Miss levels $n \in \{0, 1, 3, 5\}$ (n=0 is Control, n=1 is Miss-1 of Table 6, n=3 is the main-text Miss) with TRACE-Smoother, holding the other fixed.

Length-scale (left). Each curve is sharply U-shaped, with the optimum shifting with the missing level: full observation (n=0) prefers $\ell \approx 1$, whereas all sparse regimes ($n \geq 1$) flatten into a common plateau at $\ell \approx 5$. A single $\ell = 5$ is therefore near-optimal across sparse regimes, set from the *observable* missing rate rather than from oracle knowledge of the target.

Tempering (right). Fixing $\ell = 5$ and sweeping α leaves every curve nearly flat: RMSE varies by under 3% over $\alpha \in [0.3, 1.0]$ and degrades only mildly at $\alpha = 1.5$. The dotted line at $\alpha = 1$ marks the untempered Kalman update, which differs marginally from the tempered default; tempering is a gentle safeguard rather than the origin of the gains. We fix $\alpha = 0.5$ throughout.

Table 9: Per-frame fusion time (Active Matter, Control regime, ρ=1%, d=2304; CPU, excluding DPS sampling).

Fusion	Time/frame	Outcome
Diagonal (ours)	1.7 ms	RMSE 0.078
Full-covariance	$\gtrsim$20 s	OOM at d>5000

Table 10: Sample-count S ablation for TRACE-Frame on Active Matter (Control regime, $\rho = 1\%$, 10 trajectories), field-domain RMSE.

S	1	5	20	50
RMSE	0.091	0.083	**0.081**	0.080

Table 12 decomposes the total RMSE of Table 1 into averages over *observed* frames ($\mathcal{O}_n \neq \emptyset$) and *missing* frames ($\mathcal{O}_n = \emptyset$) under Miss-3, Blk-5, and Blk-10 at $\rho = 1\%$.

F.5 Robustness: Observation Noise and Sparsity

Table 11 reports robustness to additive measurement noise on Active Matter ($\rho = 1\%$, Control regime). All variants degrade gracefully; TRACE-Smoother is most robust at the highest noise level (0.184 at σ=0.3), where the temporal prior helps most. At low noise the three variants are indistinguishable.

Observation-rate sweep. Table 14 extends ρ to 0.5%–10% on Active Matter (Control). RMSE decreases monotonically with ρ and TRACE-Smoother remains on top of TRACE-Filter at every ρ, with the largest margin at the sparsest setting.

F.6 Streaming Behaviour: Coverage Accumulation and Runtime

Coverage accumulation. Table 15 reports RMSE versus the number of sensor revisits under the circular moving window (matching the Spatially Localized setting of Table 1). TRACE-Smoother improves monotonically as coverage accumulates, while the per-frame INR baseline is essentially flat in the revisit count. On Active Matter and Ocean TRACE-Smoother dominates at every loop; on the high-rank Supernova cube the baseline is competitive at a single pass but is overtaken from the second revisit onward.

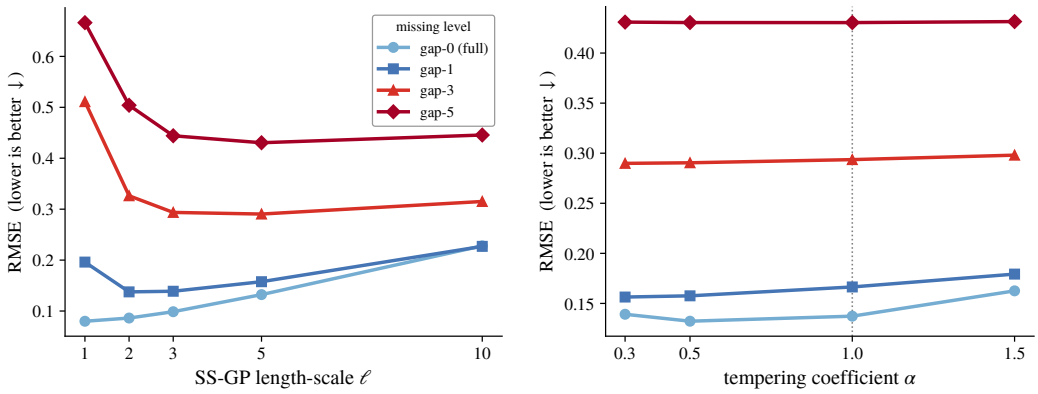
Runtime and complexity. Table 16 reports per-frame cost on Active Matter ($d = 2304$). The DPS reverse ODE ($\approx$3.5 s) dominates; the recursive filter / smoother / fusion / decode components add only a few milliseconds. Diagonal fusion is $\mathcal{O}(d)$ in compute and memory and remains feasible at d=32768 (Supernova), while full-covariance fusion ($\mathcal{O}(d^3)$ compute, $\mathcal{O}(d^2)$ memory) exhausts memory beyond $d \approx 5000$. The RTS backward pass adds one $\mathcal{O}(T)$ recursion at $\approx 3.5\times$ the forward filter cost; the DPS forward cost is unchanged. Per-frame numbers extrapolate linearly in T on our validated $T = 24$ horizon.

F.7 Per-Frame Error Decomposition: Observed vs. Missing Frames

Two patterns emerge. On observed frames, TRACE-Filter and TRACE-Smoother differ negligibly: direct per-frame DPS already constrains them. On *missing* frames, the smoother reduces error substantially over the causal filter, especially under long blackouts—e.g. Active Matter Blk-10 drops from 0.780 to 0.612 (21%). This isolates the role of the RTS backward pass: it improves precisely the frames that the forward filter cannot constrain from later evidence.

F.8 Uncertainty Calibration

TRACE-Smoother produces a Gaussian posterior over the reconstructed field at every frame. Table 13 reports empirical coverage against nominal credible levels on Active Matter (Control regime,



(a) Length-scale ℓ (fixed $\alpha = 0.5$) (b) Tempering α (fixed $\ell = 5$)

Figure 5: Hyperparameter sensitivity on Active Matter ($\rho = 1\%$, full TRACE-Smoother), one curve per Miss level $n \in \{0, 1, 3, 5\}$. **(a)** RMSE is strongly ℓ-sensitive and the optimal ℓ shifts with n, saturating at $\ell \approx 5$ for all sparse regimes. **(b)** RMSE is nearly α-insensitive (note the compressed vertical range); the dotted line is the untempered Kalman update ($\alpha = 1$). Other hyperparameters held fixed ($\sigma_f = 1$, diagonal fusion).

Table 11: Observation-noise robustness (Active Matter, Control regime, $\rho = 1\%$), field-domain RMSE. σ is the additive Gaussian sensor noise. σ=0: 28 trajectories; σ>0: 3 trajectories.

σ	TRACE-Frame	TRACE-Filter	TRACE-Smoother
0.0	0.078	0.078	**0.078**
0.1	0.097	0.097	0.098
0.3	0.195	0.192	**0.184**

28 trajectories). Our posterior is substantially better calibrated than the per-frame ensemble in magnitude (ECE 0.058 vs. 0.220), though it retains some residual overconfidence at the higher credible levels. This is a magnitude-calibration check supporting the probabilistic interpretation, not a claim about spatial informativeness of the variance.

F.9 Qualitative Reconstructions

We provide per-dataset qualitative reconstructions for the structured regimes (Temporally Sparse and Spatially Localized) at $\rho = 1\%$. Unless noted, columns are ground truth, observation, TRACE-Smoother, TRACE-Filter, SDIFT, MMGN, DBF, LRTFR; the Observation panel shows the scattered off-grid samples and is gray on unobserved frames. The Control regime is not shown separately: its per-frame quality is already visible in the observed-frame rows (marked *obs*) of the missing-frame figures.

Temporally sparse: missing frames. The *Miss* pattern (one observed frame in four), matching the main-text Table 1.

Temporally sparse: blackout. A contiguous interval is unobserved (gray in the Observation column); nine frames span the blackout. The causal TRACE-Filter loses structure inside the blackout, while TRACE-Smoother propagates evidence from both ends to bridge it. The blackout (Blk-10) covers frames 7–16 (Active Matter, Ocean) and 3–12 (Supernova).

Spatially localized: moving window. A circular one-lap window (green box) sweeps the domain. TRACE-Smoother carries previously-swept regions forward into a full-field estimate. Columns here are ground truth, observation, TRACE-Smoother, TRACE-Filter, TRACE-Frame, SDIFT, MMGN, LRTFR: the TRACE-Frame column exposes how the memoryless variant collapses outside the current window; DBF is dropped (degenerates to a near-mean field).

Table 12: Per-frame error decomposition at $\rho = 1\%$, stacked by dataset (Active Matter, Ocean, Supernova). Each cell reports total / observed-frame / missing-frame RMSE under Miss-3, Blk-5, and Blk-10 (same symbols as Table 1). **Bold** marks the best total per column within each dataset block.

Method	Miss-3	Blk-5	Blk-10
Active Matter (ρ=1%)			
LRTFR	0.371/0.220/0.422	0.218/0.156/0.451	0.368/0.160/0.660
MMGN	0.313/0.156/0.366	0.217/0.164/0.418	0.360/0.164/0.635
SDIFT	0.458/0.363/0.490	0.352/0.280/0.626	0.472/0.303/0.710
DBF	0.930/0.826/0.965	0.789/0.744/0.957	0.843/0.749/0.975
TRACE-Filter	0.429/0.084/0.544	0.211/0.103/0.619	0.385/0.103/0.780
TRACE-Smoother	**0.290**/0.093/**0.356**	**0.186**/0.132/**0.392**	**0.332**/0.131/**0.612**
Ocean (ρ=1%)			
LRTFR	0.201/0.194/0.203	0.159/0.154/0.180	0.172/0.162/0.187
MMGN	0.097/0.070/0.106	0.080/0.070/0.119	**0.092**/0.070/**0.123**
SDIFT	0.150/0.122/0.159	0.113/0.098/0.168	0.133/0.107/0.170
DBF	0.797/0.628/0.853	0.604/0.536/0.862	0.680/0.542/0.874
TRACE-Filter	0.191/0.045/0.239	0.108/0.044/0.349	0.257/0.045/0.554
TRACE-Smoother	**0.095**/0.045/**0.112**	**0.059**/0.044/**0.116**	0.124/0.045/0.235
Supernova (ρ=1%)			
LRTFR	0.560/0.555/0.562	0.430/0.437/0.414	0.503/0.511/0.498
MMGN	**0.318**/0.315/**0.319**	0.318/0.319/0.314	**0.320**/0.320/**0.319**
SDIFT	0.435/0.420/0.440	0.397/0.392/0.409	0.424/0.407/0.433
DBF	0.853/0.796/0.871	0.806/0.778/0.866	0.841/0.782/0.877
TRACE-Filter	0.389/0.321/0.411	0.375/0.318/0.500	0.539/0.322/0.669
TRACE-Smoother	0.334/0.316/0.339	**0.312**/0.310/**0.314**	0.358/0.312/0.385

Table 13: Reliability on Active Matter (Control regime, 28 trajectories): empirical coverage at two nominal credible levels, and expected calibration error (ECE) over ten levels. Closer to the nominal level is better.

	nominal 0.90	nominal 0.95	ECE
TRACE-Smoother	0.80	0.86	**0.058**
TRACE-Frame ens.	0.56	0.62	0.220

Failure case. When the narrow window leaves the field under-determined, the learned generative prior cannot resolve it: the causal TRACE-Filter propagates the erroneous per-frame estimate forward and even the full smoother cannot recover—a limitation when per-frame evidence is too weak for the prior to be informative.

F.10 Limitations and Future Work

A key limitation of TRACE stems from its temporal prior. Since temporal structure is captured by a stationary state-space Gaussian process rather than a learned dynamical model, the estimate carries no intrinsic account of how the field evolves. When observations are absent over a prolonged interval, the estimate relaxes toward the prior mean and its uncertainty approaches the stationary marginal, so TRACE reconstructs well between observed frames but does not genuinely extrapolate the field forward in time. Retrospective smoothing alleviates this behaviour only when later observations are available to constrain the intervening frames. A promising direction for future work is to replace the stationary Gaussian process with a trajectory-level prior that learns the field dynamics along the temporal axis, enabling the model to propagate physical evolution through unobserved intervals instead of reverting to the prior.

Table 14: Observation-rate sweep (Active Matter, Control regime, $\sigma_{\text{obs}} = 0$, 28 trajectories, DPS reverse steps N=50, guidance $\zeta \propto 1/\rho$). Field-domain RMSE; lower is better. This sweep uses a reduced 50-step sampler and is therefore not directly comparable to the default 100-step results of Table 5.

ρ	TRACE-Filter	TRACE-Smoother
0.5%	0.104	**0.099**
1%	0.093	**0.089**
2%	0.082	**0.079**
10%	0.072	**0.070**

Table 15: Coverage accumulation: RMSE versus number of revisits (loops) under the circular moving window. "TRACE-Smoother" is the full model (RTS smoothing); "MMGN" is the strongest offline per-frame baseline.

Loops	1	2	3	4	6
AM, TRACE-Smoother	0.537	0.498	0.476	0.475	**0.468**
AM, MMGN	0.782	0.770	0.777	0.771	0.768
Ocean, TRACE-Smoother	0.106	0.097	**0.088**	0.083	0.084
Ocean, MMGN	0.277	0.283	0.271	0.284	0.286
SN, TRACE-Smoother	0.334	0.320	0.318	0.314	**0.311**
SN, MMGN	0.321	0.321	0.321	0.321	0.321

Table 16: Per-frame inference cost on Active Matter ($d = 2304$), measured on a single NVIDIA RTX 5090 GPU (the closed-form fusion step of Table 9 is timed on CPU).

Component	Time
DPS reverse ODE (N=100, S=20)	≈ 3.5 s / frame
SS-GP filter (forward)	0.64 ms / frame
RTS smoothing (backward)	2.35 ms / frame
FTM decode	1.26 ms / frame
diagonal fusion ($d = 32768$)	4.1 ms / frame

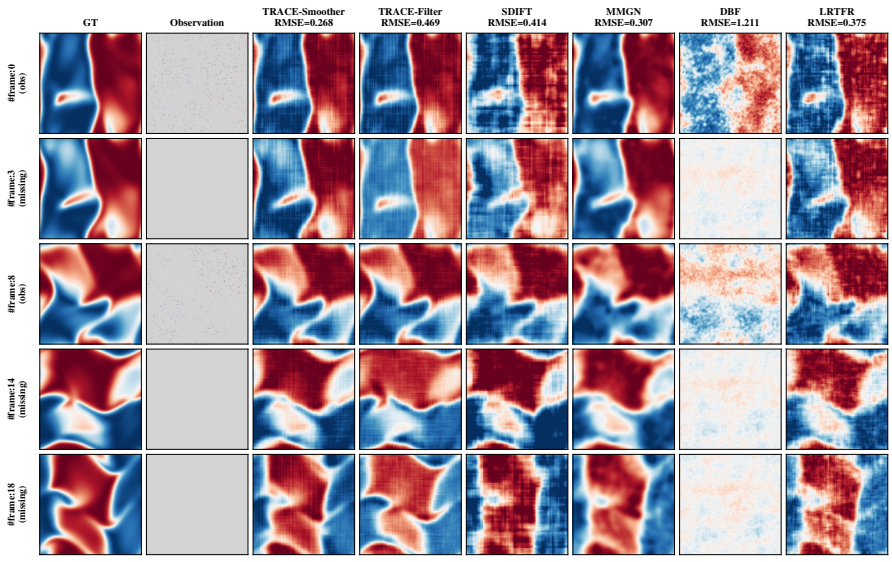


Figure 6: Active Matter, *Miss* pattern ($\rho = 1\%$); five frames mixing observed and missing.

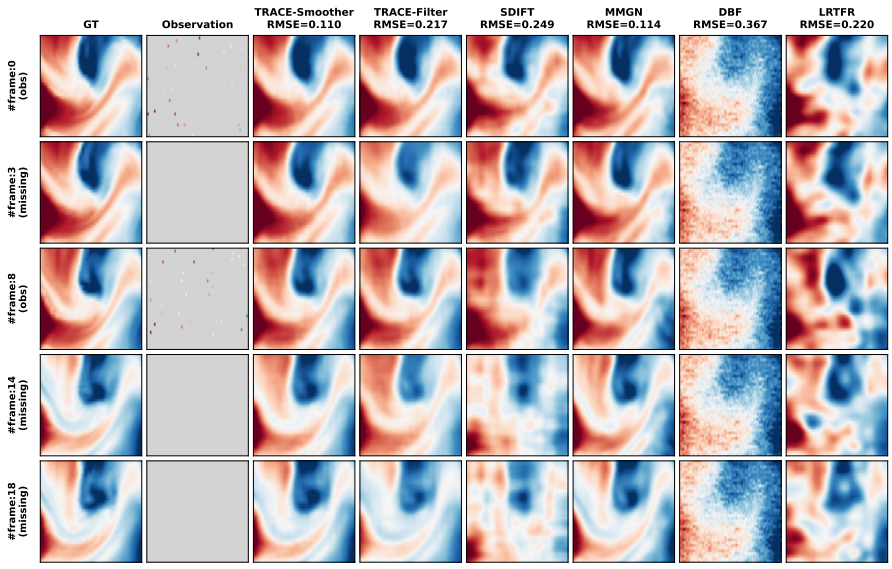


Figure 7: Ocean, *Miss* pattern ($\rho = 1\%$, gap n=3): TRACE-Smoother recovers coherent spatial structures throughout missing intervals; TRACE-Filter drifts on gap frames; offline baselines produce temporally inconsistent transitions. Columns: ground truth, observation, TRACE-Smoother, TRACE-Filter, SDIFT, MMGN, DBF, LRTFR.

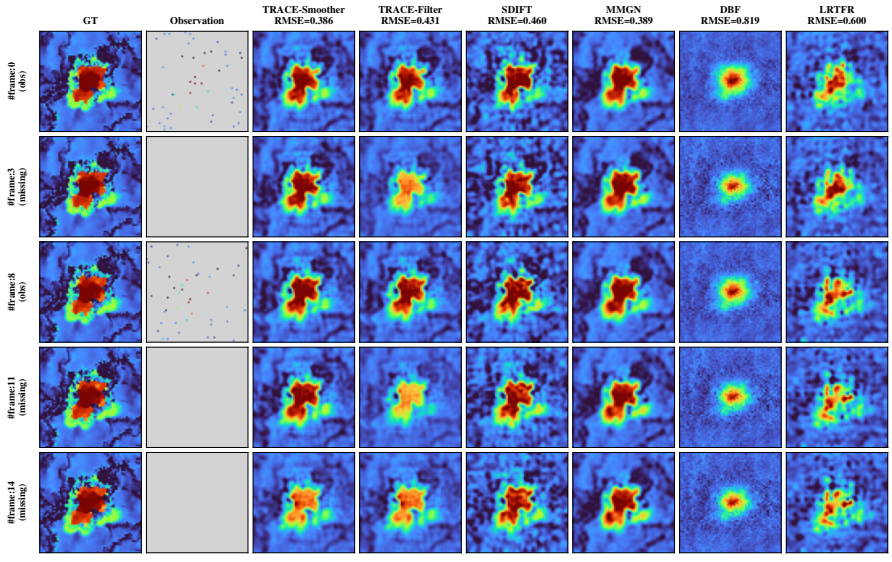


Figure 8: Supernova (z=32 slice), *Miss* pattern ($\rho = 1\%$).

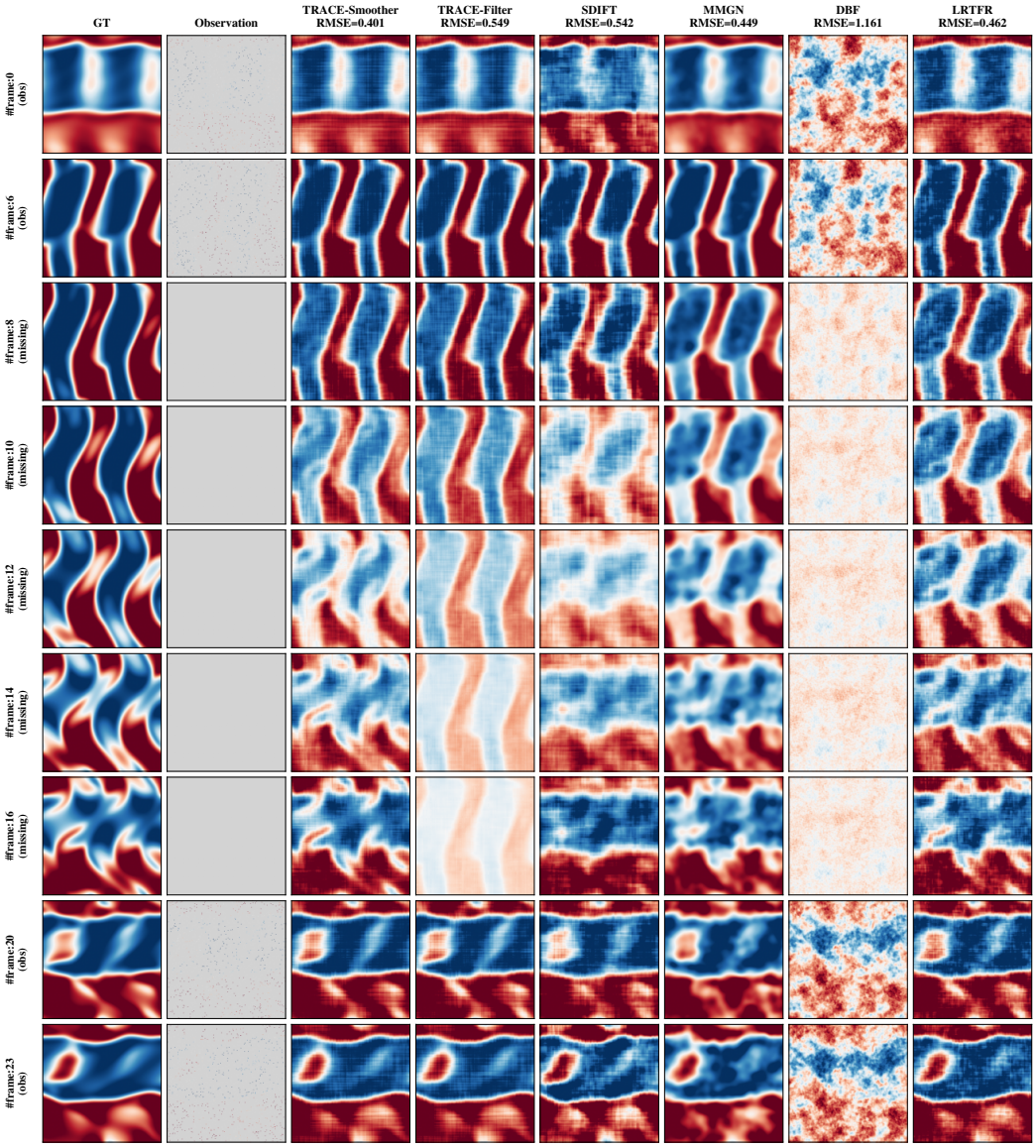


Figure 9: Active Matter, Blk-10 ($\rho = 1\%$): frames 7–16 unobserved.

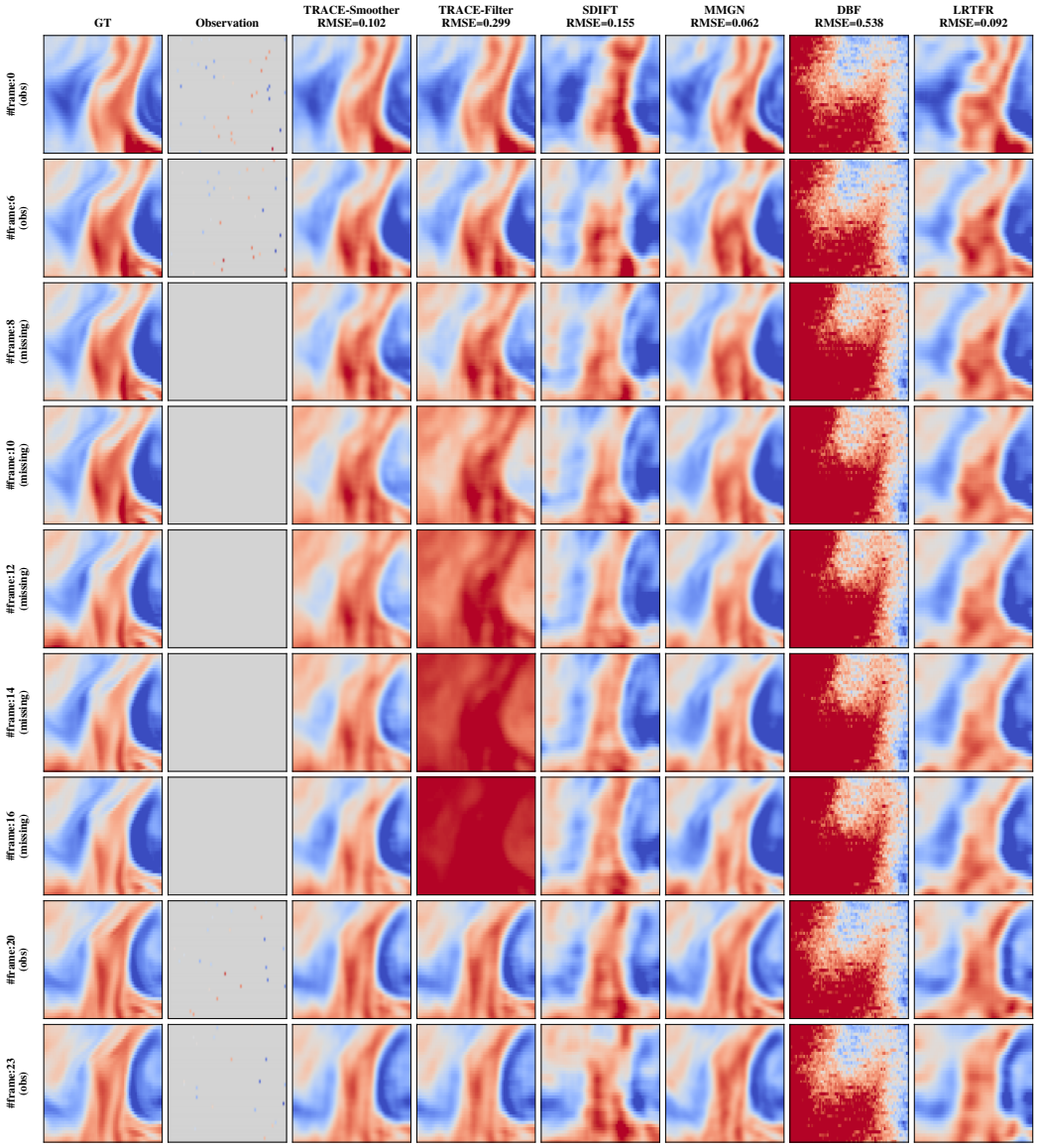


Figure 10: Ocean (layer 2), Blk-10 ($\rho = 1\%$): frames 7–16 unobserved.

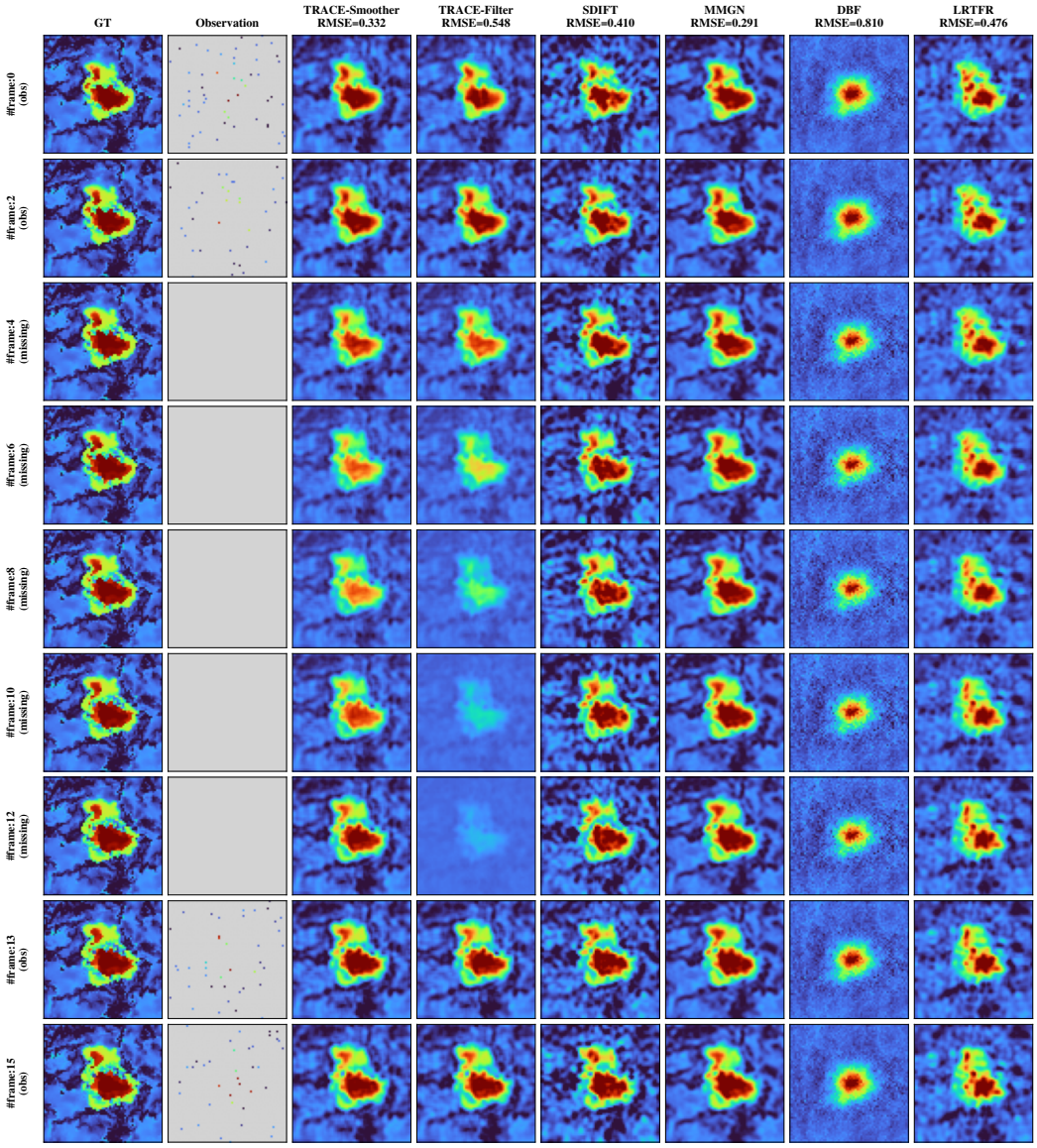


Figure 11: Supernova (z=32), Blk-10 ($\rho = 1\%$): frames 3–12 unobserved.

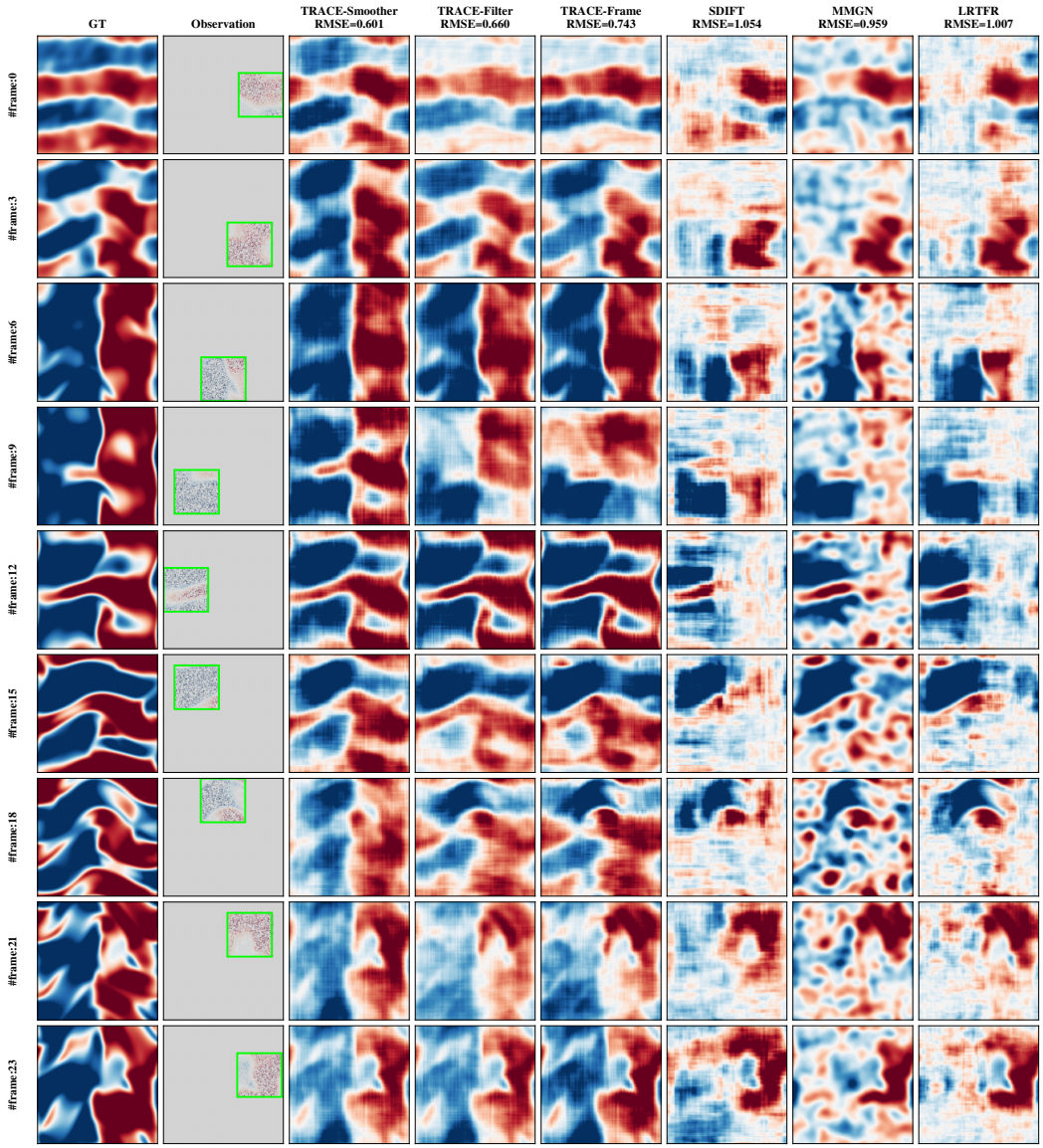


Figure 12: Active Matter, moving window (circular, one lap, $\rho_{\text{loc}} = 15\%$); nine frames span one sweep. TRACE-Smoother accumulates previously-swept regions into a coherent global field, while the per-frame baselines collapse outside the current window.

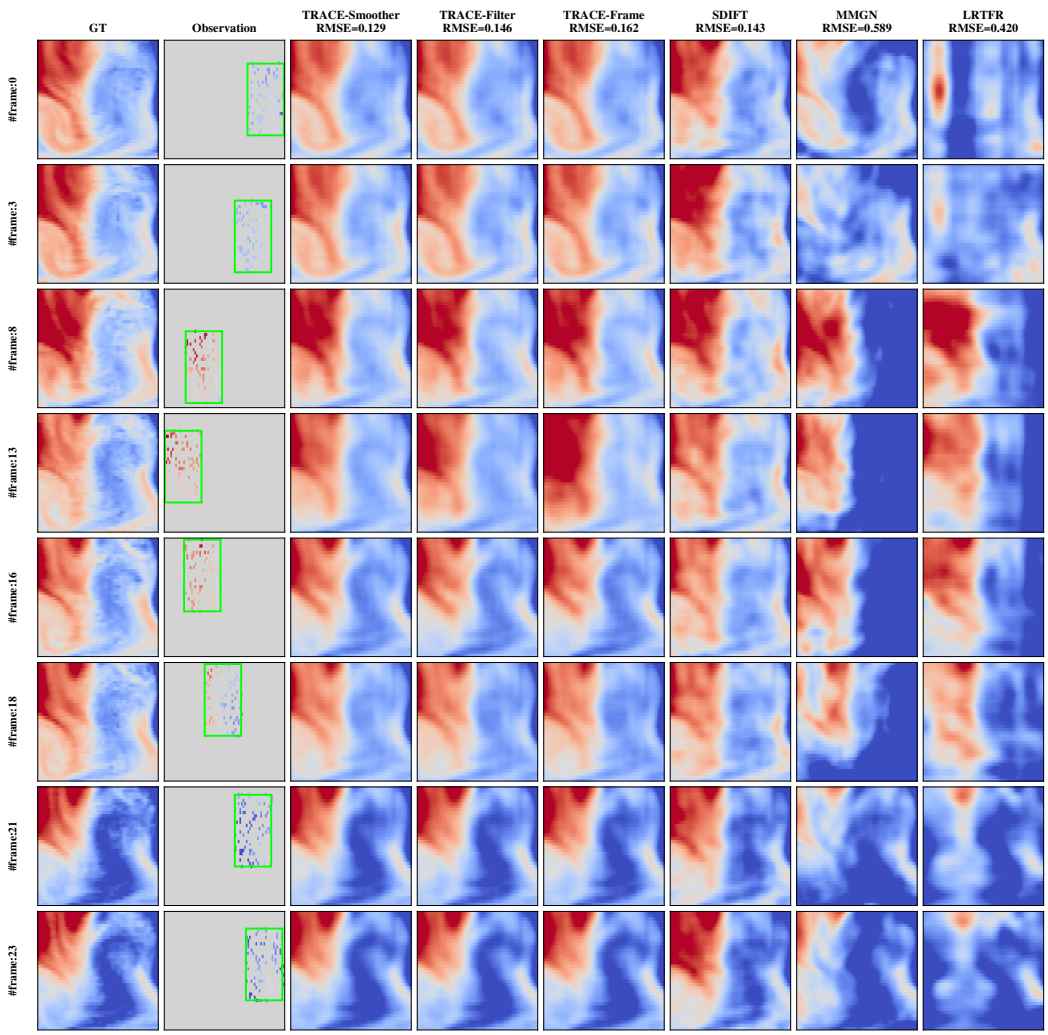


Figure 13: Ocean (layer 2), moving window (circular, one lap, $\rho_{\text{loc}} = 15\%$).

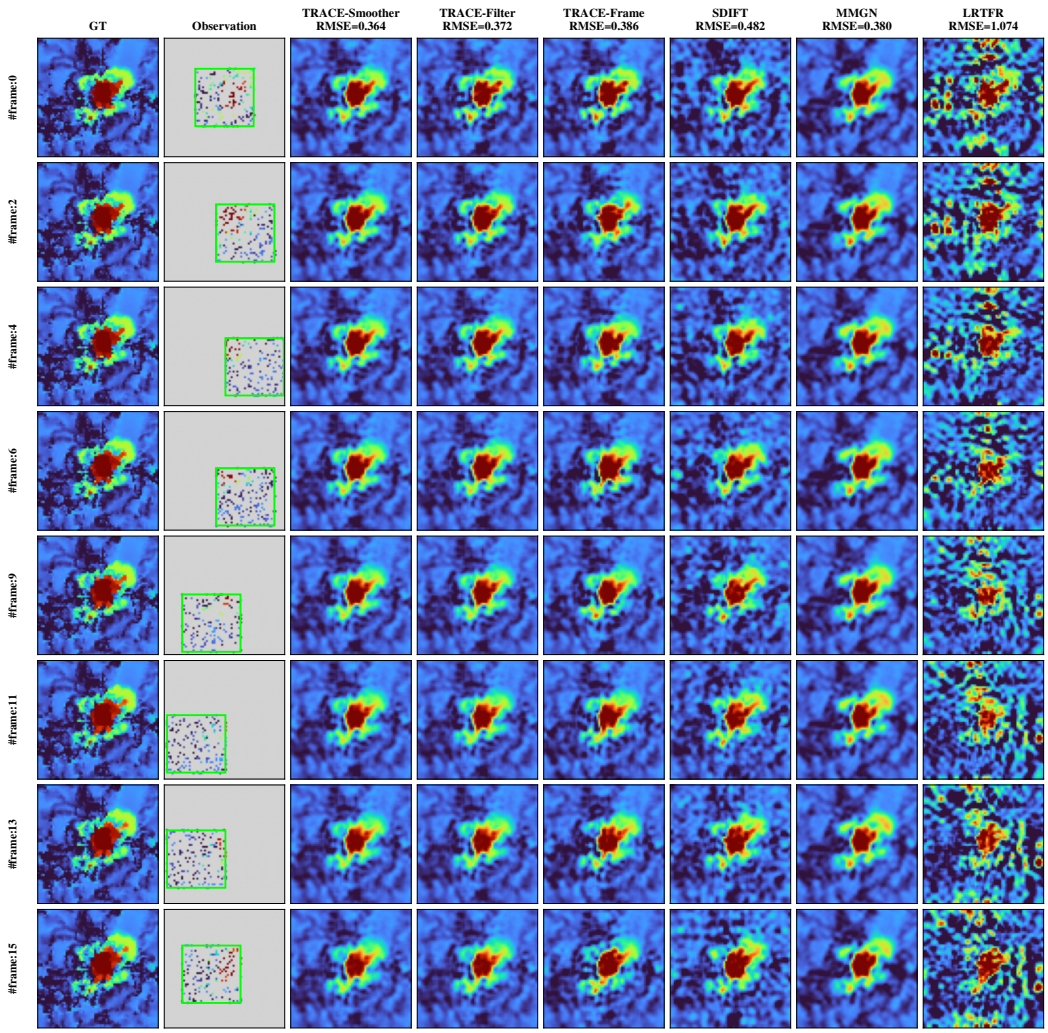


Figure 14: Supernova (z=32), moving cube (circular, one lap).

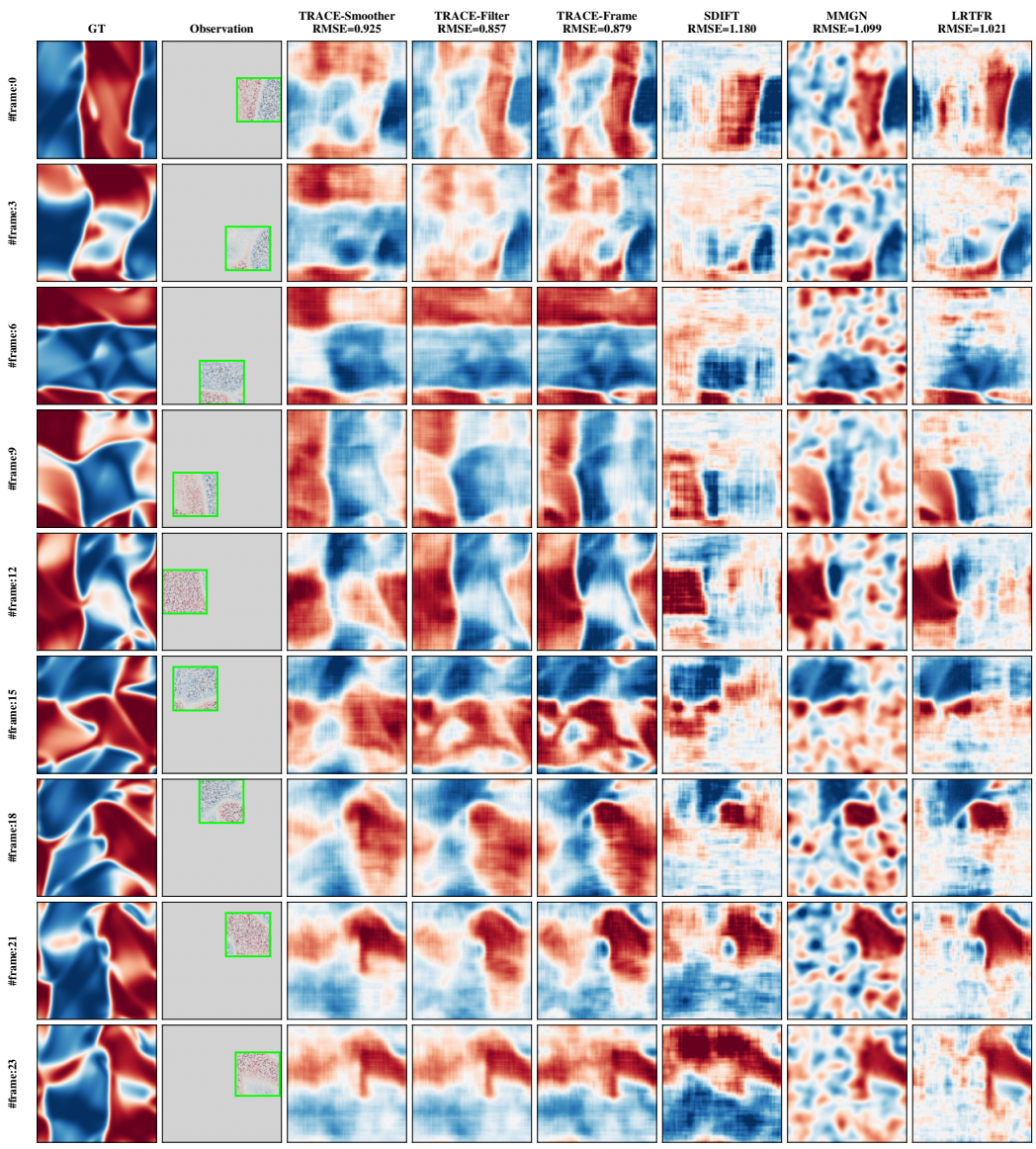


Figure 15: Failure case: Active Matter moving window where per-frame evidence is too weak for the generative prior; TRACE-Filter propagates the error and TRACE-Smoother cannot recover.