Electrocardiographic Imaging via Posterior Sampling under Operator Discrepancy

Ricarda Buttmann and Andreas M. Lehrmann
Center for Artificial Intelligence and Robotics
Technical University of Applied Sciences
Würzburg-Schweinfurt, Germany
{ricarda.buttmann,andreas.lehrmann}@thws.de

Abstract

Electrocardiographic imaging (ECGI) reconstructs heart-surface potentials from body-surface potentials, an ill-posed inverse problem. Diffusion posterior samplers combine a learned prior with an explicit forward operator that they assume to be accurate. In ECGI this operator is only an estimate, because the body changes with respiration and posture, and body composition such as subcutaneous fat is rarely known for an individual patient. The resulting operator discrepancy is structured and depends on the heart signal. We introduce OpDis-PS, a diffusion posterior sampler that models this discrepancy instead of treating it as signal-independent noise. OpDis-PS learns how the discrepancy depends on the signal, folds this into a corrected forward operator, and reconstructs with a fixed prior across all operator conditions. We study two realistic sources of discrepancy, changing lung conductivity and an added subcutaneous-fat layer. Across both, OpDis-PS reconstructs more accurately than classical regularization and learned predictors, with its largest gains on held-out cardiac pathologies.

1 Introduction

Cardiovascular diseases remain the leading cause of mortality worldwide, accounting for an estimated 19.8 million deaths in 2022 [26]. The diagnosis and treatment of arrhythmias and conduction disorders often rely on identifying where abnormal ventricular activation originates and how it propagates. Electrocardiographic imaging (ECGI) reconstructs heart-surface potentials (HSPs), defined here on the outer surface of the ventricles, from body-surface potentials (BSPs) measured on the torso [38]. These spatio-temporal signals reveal ventricular activation and recovery without requiring electrodes on the cardiac surface and can support the analysis and treatment of arrhythmias [11]. However, the torso attenuates and spatially smooths the cardiac field [35], rendering ECGI an ill-conditioned inverse problem that amplifies measurement and model errors.

Traditional ECGI methods stabilize the inversion using Tikhonov penalties, spatio-temporal constraints, or total variation [7, 27, 37]. While these regularizers improve stability, their generic smoothness assumptions capture only limited physiological structure and can attenuate localized activation patterns. Learning-based methods can capture complex cardiac dynamics directly from data, with approaches ranging from discriminative predictors that estimate HSPs directly from BSPs to conditional generative models that learn the distribution of HSPs given BSP observations [2, 17]. The latter are particularly suited to representing uncertainty and ambiguity in the inverse solution. Despite this advantage, both approaches encode the observation process in their parameterization and do not permit substituting an explicit case-specific forward operator at inference.

An alternative is to learn a generative prior over cardiac signals independently of the observation model and combine the two only during inference. Diffusion posterior samplers such as Denois-

GenAI for Health Workshop @ NeurIPS 2026, Sydney.

ing Diffusion Restoration Models (DDRM) and Diffusion Posterior Sampling (DPS) realize this separation by coupling a learned diffusion prior with an explicit measurement-consistency model [10, 21]. Originally introduced for vision tasks, both methods assume that the supplied forward operator accurately represents the measurement process. In ECGI, however, the nominal operator used for reconstruction is only an estimate derived from an anatomical model and assumed tissue conductivities. The recording operator, which represents the actual volume conductor at measurement time, can differ from this nominal operator because of respiration, posture, electrode placement, or body composition. The resulting operator discrepancy produces a structured error that depends on the underlying cardiac signal. Because ECGI is ill-conditioned, even a modest operator discrepancy can cause substantial reconstruction errors.

We propose OpDis-PS, a diffusion posterior-sampling method for reconstruction under forward-operator discrepancy. Rather than treating operator discrepancy as signal-independent measurement noise, OpDis-PS models its statistical dependence on the nominal forward signal. A signal-conditional Gaussian approximation-error model yields a linear correction to the nominal operator and a covariance for the remaining mismatch. We estimate these quantities offline using HSPs from the training set paired with sampled recording operators. At reconstruction time, OpDis-PS combines the residual discrepancy covariance with the measurement-noise covariance and whitens the corrected observation model before applying a spectral posterior sampler. The unconditional diffusion prior remains fixed across operator conditions, allowing the likelihood to adapt without retraining the generative model.[1] We evaluate OpDis-PS under two physiologically motivated sources of operator discrepancy: lung-conductivity variation, representing different lung inflation states [22, 32], and a superficial subcutaneous-fat layer, representing a body-composition-related torso inhomogeneity known to affect BSPs [6, 23]. Across these settings, OpDis-PS is compared with baselines spanning multiple learning paradigms and demonstrates strong reconstruction performance.

Contributions. In summary, we make the following contributions: (**1**) we present OpDis-PS, a diffusion posterior sampler that reconstructs heart-surface potentials when the reconstruction operator differs from the one that generated the measurement; (**2**) we show how to cast this operator discrepancy as a signal-conditional approximation-error term, absorb its predictable part into a corrected operator and observation, and carry the residual uncertainty into a whitened data-consistency step that leaves the diffusion prior untouched; and (**3**) we build a controlled ECGI evaluation from openCARP [34] simulations that isolates operator discrepancy from measurement noise, and assess OpDis-PS against variational, amortized, and posterior-sampling methods on physiological and unseen pathological activation.

2 Related work

OpDis-PS relates to four lines of prior work. We discuss foundational ECGI studies, classical regularizers, machine-learning models for ECGI, and the diffusion posterior samplers on which OpDis-PS builds.

General ECGI problem. Early studies establish the physical and mathematical setting of ECGI. Rudy and Messinger-Rapport [38] formulate the recovery of HSPs from BSPs as a linear inverse problem and analyze its ill-posed structure. Oster et al. [33] subsequently demonstrate such reconstructions in a torso-tank experiment. Ramanathan and Rudy [35] study the forward problem and examine how volume conduction and torso inhomogeneities affect BSPs.

Traditional methods. Classical ECGI reconstruction relies on variational regularizers that penalize spatial or temporal features of the HSPs. Comparative studies examine Total Variation and first- and second-order Tikhonov regularization among other inverse approaches [20, 28]. Cluitmans et al. [12] replace generic penalties with a truncated basis learned from simulated activation and recovery patterns and reconstruct sparse coefficients in this basis. Moving from deterministic regularization to a probabilistic treatment of forward-operator error, Aldemir and Dogrusoz [1] model the approximation error as an additive Gaussian term with a data-estimated covariance and incorporate it into a Bayesian MAP reconstruction. OpDis-PS carries this approximation-error idea into diffusion posterior sampling and extends it with a signal-conditional operator correction learned from paired training data.

[1] Code, pretrained models, and datasets: projects.cairo.thws.de/asml/papers/ecgi.

Machine learning methods for ECGI. Data-driven ECGI methods differ in their reconstruction targets, treatment of geometry, and use of the forward operator. Amortized end-to-end approaches train networks that map BSPs directly to heart-surface quantities without explicitly using the forward operator at inference [8, 18, 44]. Jara and Meyers [17] use conditional diffusion to map BSPs to HSPs without subject-specific geometry. Our Transformer and Palette [39] baselines represent this amortized-inference line and operate on a common heart-template representation across subjects. Other generative approaches use latent-variable models: Ghimire et al. [13] reconstruct volumetric transmembrane-potential sequences using a subject-specific sequential VAE prior, while Bacoyannis et al. [2] use a conditional VAE to sample volumetric activation-time maps.

Diffusion posterior samplers. Chung et al. [10] propose DPS, which guides diffusion sampling using a gradient of measurement consistency. Kawar et al. [21] formulate DDRM in the singular coordinates of a linear forward operator, with updates determined by the singular values and measurement-noise level. Nezakati et al. [30] extend DDRM to correlated Gaussian measurement noise by whitening the observation before the restoration update. These approaches formulate data consistency using a single observation operator that is assumed to be accurate at inference. OpDis-PS relaxes this assumption by combining a signal-conditional correction of the nominal operator with a residual discrepancy model in the whitened likelihood.

3 Technical prerequisites

The ECGI forward process propagates heart-surface potentials through a torso volume conductor to body-surface potentials. OpDis-PS builds upon previous work on (**1**) approximating this continuous process as a linear model with Gaussian measurement noise (Section 3.1) and (**2**) inverting linear models via posterior sampling using diffusion priors (Section 3.2).

3.1 Linear approximation

Let $\Omega \subset \mathbb{R}^3$ be the torso volume conductor between the heart surface Γ_{H} and the body surface Γ_{B}, with $\partial\Omega = \Gamma_{\mathrm{H}} \cup \Gamma_{\mathrm{B}}$. At every physical time t, the electric potential $\phi(\mathbf{r}, t)$, with $\mathbf{r} \in \Omega$, satisfies a stationary conductivity problem described by a linear second-order elliptic boundary-value problem. Suppressing the time argument t, we have

$$\nabla \cdot (\sigma_{\mathrm{c}} \nabla \phi) = 0 \quad \text{in } \Omega, \qquad \phi = \phi_{\mathrm{H}} \quad \text{on } \Gamma_{\mathrm{H}}, \qquad \mathbf{n}^{\mathsf{T}} \sigma_{\mathrm{c}} \nabla \phi = 0 \quad \text{on } \Gamma_{\mathrm{B}}, \tag{1}$$

where ϕ_{H} denotes the prescribed potential on the heart surface Γ_{H}, $\mathbf{n}$ is the outward unit normal to Ω on Γ_{B}, and σ_{c} is the spatially varying scalar electrical conductivity.

For test functions v that vanish on Γ_{H}, integration by parts and the zero-flux condition on Γ_{B} give the weak form

$$\int_{\Omega} (\nabla v)^{\mathsf{T}} \sigma_{\mathrm{c}} \nabla \phi \, \mathrm{d}\Omega = 0. \tag{2}$$

We approximate Ω by a tetrahedral mesh and associate each node j with a P_1 basis function ψ_j. The resulting global stiffness matrix $\mathbf{K}$, with entries $K_{ij} = \int_{\Omega} (\nabla \psi_i)^{\mathsf{T}} \sigma_{\mathrm{c}} \nabla \psi_j \, \mathrm{d}\Omega$, provides a discrete representation of the conductivity operator [43].

Partitioning $\mathbf{K}$ according to the unknown free nodes F (in Ω and on Γ_{B}) and the prescribed heart nodes H (on Γ_{H}) yields the free–free and free–heart blocks $\mathbf{K}_{\mathrm{FF}}$ and $\mathbf{K}_{\mathrm{FH}}$. With slight abuse of notation, we reuse ϕ_{H} to denote the vector of heart-node potentials and write ϕ_{F} for the vector of free-node potentials. The propagation from the heart nodes to the free nodes at a fixed point in time is then given by

$$\mathbf{A}^*_{\bullet} = -\mathbf{K}_{\mathrm{FF}}^{-1} \mathbf{K}_{\mathrm{FH}}, \qquad \phi_{\mathrm{F}} = \mathbf{A}^*_{\bullet} \phi_{\mathrm{H}}, \tag{3}$$

which we aggregate across T time points using $\Phi_{\mathrm{F}} = [\phi_{\mathrm{F},1}, \ldots, \phi_{\mathrm{F},T}]$ and $\Phi_{\mathrm{H}} = [\phi_{\mathrm{H},1}, \ldots, \phi_{\mathrm{H},T}]$. Letting Φ_{B} denote the restriction of Φ_{F} to the potentials on the body surface Γ_{B} and $\mathbf{A}^*$ the corresponding row restriction of $\mathbf{A}^*_{\bullet}$, we obtain the linear observation model

$$\Phi_{\mathrm{B}} = \mathbf{A}^* \Phi_{\mathrm{H}} + \boldsymbol{\varepsilon}, \tag{4}$$

where $\boldsymbol{\varepsilon}$ is additive measurement noise. The inverse problem is ill-posed because the volume conductor attenuates and spatially smooths the cardiac field, so inversion amplifies measurement noise and model errors in weakly informed directions, requiring prior information or regularization for stable recovery [25].

3.2 Posterior sampling-based inversion

Posterior sampling-based inversion combines a learned generative prior $p_\theta(\Phi_\mathrm{H})$ with a known observation model $\Phi_\mathrm{B} = \mathcal{F}(\Phi_\mathrm{H}) + \boldsymbol{\varepsilon}$ through Bayes' rule, $p(\Phi_\mathrm{H} \,|\, \Phi_\mathrm{B}) \propto p(\Phi_\mathrm{B} \,|\, \Phi_\mathrm{H}) p_\theta(\Phi_\mathrm{H})$, and reconstructs Φ_H by drawing samples from $p(\Phi_\mathrm{H} \mid \Phi_\mathrm{B})$. We recap a common approach with a diffusion prior and a spectral sampler for the linear model $\mathcal{F}(\Phi_\mathrm{H}) = \mathbf{A}^* \Phi_\mathrm{H}$.

Diffusion prior [40]. Starting at $\Phi_\mathrm{H}^{(0)} = \Phi_\mathrm{H}$, a variance-preserving diffusion prior with noise schedule $\{\bar{\alpha}^{(\tau)}\}_{\tau=1}^{S}$ models the reverse of the forward process

$$\Phi_\mathrm{H}^{(\tau)} = \sqrt{\bar{\alpha}^{(\tau)}}\, \Phi_\mathrm{H}^{(0)} + \sqrt{1 - \bar{\alpha}^{(\tau)}}\, \boldsymbol{\xi}, \qquad \xi_{i,t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1) \tag{5}$$

by predicting the velocity

$$\mathbf{v}^{(\tau)} = \sqrt{\bar{\alpha}^{(\tau)}}\, \boldsymbol{\xi} - \sqrt{1 - \bar{\alpha}^{(\tau)}}\, \Phi_\mathrm{H}^{(0)} \tag{6}$$

with a neural network $\mathbf{v}_\theta(\Phi_\mathrm{H}^{(\tau)}, \tau)$, which yields the denoised estimate

$$\widehat{\Phi}_\mathrm{H}^{(0)}(\Phi_\mathrm{H}^{(\tau)}) = \sqrt{\bar{\alpha}^{(\tau)}}\, \Phi_\mathrm{H}^{(\tau)} - \sqrt{1 - \bar{\alpha}^{(\tau)}}\, \mathbf{v}_\theta(\Phi_\mathrm{H}^{(\tau)}, \tau). \tag{7}$$

Unconditional sampling starts from $\Phi_\mathrm{H}^{(S)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and traverses a decreasing sequence of noise levels using DDIM updates [41], ultimately yielding a draw from the prior $p_\theta(\Phi_\mathrm{H})$.

Spectral sampler [21]. Denoising Diffusion Restoration Models (DDRM) turn this prior into a solver for linear inverse problems by combining it with an observation model of the form in Equation (4) that defines a likelihood. Assuming i.i.d. Gaussian noise with variance σ^2, DDRM works in the singular coordinates of $\mathbf{A}^* = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\mathsf{T}$. For each singular value $s_i > 0$, let $\mathbf{u}_i$ and $\mathbf{v}_i$ denote the corresponding left and right singular vectors. Because the same spatial operator $\mathbf{A}^*$ acts on every time point, its singular-coordinate transformation applies to each column of the sequence. For each singular direction i, we define

$$\bar{\Phi}_{\mathrm{H},i} = \mathbf{v}_i^\mathsf{T} \Phi_\mathrm{H}, \qquad \bar{\Phi}_{\mathrm{B},i} = \frac{\mathbf{u}_i^\mathsf{T} \Phi_\mathrm{B}}{s_i}, \tag{8}$$

which yields

$$\bar{\Phi}_{\mathrm{B},i} = \bar{\Phi}_{\mathrm{H},i} + \frac{\sigma}{s_i}\, \boldsymbol{\zeta}_i, \qquad \boldsymbol{\zeta}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_T), \tag{9}$$

i.e., the standard deviation of the effective observation noise in singular direction i is σ / s_i. At every diffusion step, the DDRM reverse update blends the prior prediction $\widehat{\Phi}_\mathrm{H}^{(0)}$ from (7) with $\bar{\Phi}_{\mathrm{B},i}$ according to whether the current diffusion noise exceeds this scale. If it does, the update moves towards $\bar{\Phi}_{\mathrm{B},i}$; otherwise, the prior prediction dominates. Directions with small singular values therefore receive weaker measurement guidance, directions in the null space of $\mathbf{A}^*$ are determined by the prior alone, and for $\sigma = 0$ the sampler enforces exact consistency with the supplied operator $\mathbf{A}^*$ in every observed coordinate.

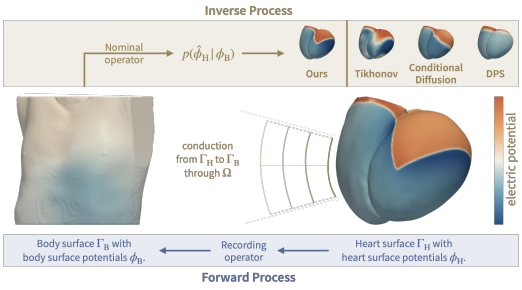
4 Posterior sampling under operator discrepancy

We build upon the formulations in Section 3 by introducing geometry-invariant template representations (Section 4.1). We then present the central concept of operator discrepancy (Section 4.2) and demonstrate its integration into the likelihood while keeping the prior fixed (Section 4.3).

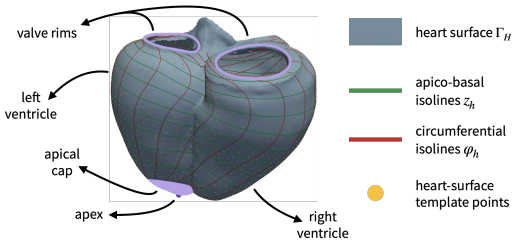
4.1 Template representations

Since the node locations associated with the potentials Φ_H and Φ_B depend on the heart and body geometry, they are subject-specific. To account for inter-subject variation, we introduce shared sets of template points on Γ_H and Γ_B. We denote the potential values at the n_H heart-template points by $\mathbf{h} \in \mathbb{R}^{n_\mathrm{H} \times T}$ and the observed potential values at the n_B body-template points by $\mathbf{b} \in \mathbb{R}^{n_\mathrm{B} \times T}$, with one time point per column. The template points reference equivalent locations across subjects, with two subject-specific linear maps connecting the shared templates to the native meshes: the heart-template map $\mathbf{T}_\mathrm{H}$ yields Φ_H from $\mathbf{h}$ and the body-template map $\mathbf{T}_\mathrm{B}$ maps Φ_B to $\mathbf{b}$. Extending (4), our observation model is thus

$$\mathbf{A} = \mathbf{T}_\mathrm{B} \mathbf{A}^* \mathbf{T}_\mathrm{H}, \quad \mathbf{b} = \mathbf{A}\mathbf{h} + \boldsymbol{\varepsilon}. \tag{10}$$



(a) Inversion under operator discrepancy.



(b) Heart-surface coordinates.

Figure 1: **Overview**. (**a**) We consider the task of recovering latent heart-surface potentials ϕ_{H} from observed body-surface potentials ϕ_{B}. Due to environmental variability, the true ("recording") operator describing the forward process might differ from the theoretical ("nominal") operator employed in the inverse process, giving rise to a signal-dependent *operator discrepancy*. (**b**) For modeling purposes, we represent the heart-surface potentials ϕ_{H} at 1800 discrete template points, shown in orange. Each template point is parameterized by an apico-basal coordinate z_h and a circumferential coordinate φ_h.

Heart template. We describe each point on the heart surface Γ_{H} using an apico-basal coordinate z_h and a periodic circumferential coordinate φ_h, such that equal-sized regions in the (z_h, φ_h) domain correspond to approximately equal heart-surface areas. The $n_{\mathrm{H}} = 1800$ template points form a regular 30×60 grid in this coordinate system, which we construct on one reference heart and then transfer to other subjects through known node correspondences.

The prescribed heart-mesh potentials Φ_{H} are mapped to the heart-template potentials $\mathbf{h} \in \mathbb{R}^{1800\times T}$ by a mass-weighted L^2 fit over the represented surface regions, applied independently at each time point. Conversely, for the forward problem, the subject-specific linear map $\mathbf{T}_{\mathrm{H}}$ yields the prescribed heart-mesh potentials Φ_{H} by interpolating the heart-template potentials $\mathbf{h}$ at the heart-mesh nodes. Both the template and its subject-specific maps remain fixed across operator conditions. Figure 1b illustrates the construction, with mathematical details provided in Appendix A.3.

Body template. The body-template potentials $\mathbf{b}$ are observed at n_{B} electrode locations on the body surface Γ_{B}. Since potentials are only defined up to an additive constant, we introduce the centering matrix $\mathbf{C} = \mathbf{I} - \frac{\mathbf{1}\mathbf{1}^{\mathsf{T}}}{n_{\mathrm{B}}}$ and define the body-template map as $\mathbf{T}_{\mathrm{B}} = \mathbf{C}\mathbf{S}$, where $\mathbf{S}$ selects the body-mesh potentials Φ_{B} at the electrode locations. Consistent with this centering, we draw $\mathbf{z}_t \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_{n_{\mathrm{B}}})$ independently for each time point and set $\boldsymbol{\varepsilon}_t = \mathbf{C}\mathbf{z}_t$, so $\mathrm{Cov}(\boldsymbol{\varepsilon}_t) = \sigma^2\mathbf{C}$.

The covariance $\sigma^2\mathbf{C}$ is rank-deficient, because $\mathbf{C}$ removes one degree of freedom, and this violates the i.i.d. Gaussian noise assumption of the spectral posterior sampler. The centered noise instead lies in the $(n_{\mathrm{B}}-1)$-dimensional subspace of zero-mean measurements. We express $\mathbf{b}$, $\mathbf{A}$, and $\boldsymbol{\varepsilon}$ in an orthonormal basis of this subspace and, with slight abuse of notation, reuse the same symbols for these reduced counterparts, so that the reduced noise becomes isotropic, $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_{n_{\mathrm{B}}-1})$. Appendix A.1 gives the details.

4.2 Operator discrepancy

Section 3.1 defines a generic forward operator $\mathbf{A}^*$ for a given torso geometry and conductivity distribution. In practice, the operator available at reconstruction time is a *nominal* operator $\mathbf{A}$, whereas the operator producing the observation at measurement time is a potentially different *recording* operator $\mathbf{A}_0$. When $\mathbf{A}_0 \neq \mathbf{A}$, the reconstruction method consumes an observation produced by $\mathbf{A}_0$ while only knowing $\mathbf{A}$. Accordingly, the nominal observation model in Equation (10) is formulated in terms of $\mathbf{A}$, despite the observation being generated by $\mathbf{A}_0$. OpDis-PS accounts for the resulting mismatch inside the likelihood.

The reduced observation generated by the recording operator $\mathbf{A}_0$ can be written in terms of the nominal operator $\mathbf{A}$ available during reconstruction as

$$\mathbf{b} = \mathbf{A}_0\mathbf{h} + \boldsymbol{\varepsilon} = \mathbf{A}\mathbf{h} + \underbrace{(\mathbf{A}_0 - \mathbf{A})\mathbf{h}}_{\mathbf{d}} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_{n_{\mathrm{B}}-1}). \tag{11}$$

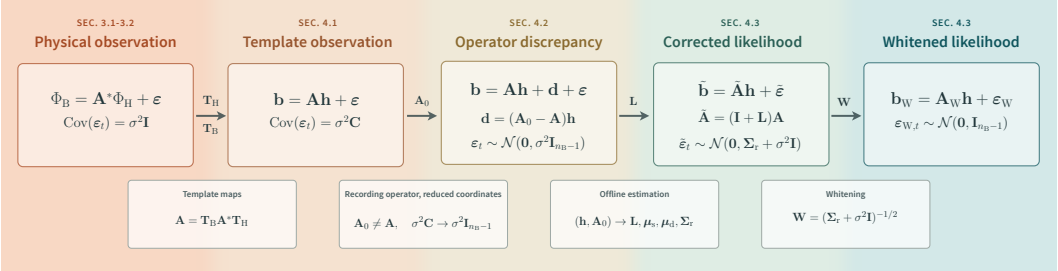


Figure 2: **Model**. The diagram traces the observation model from the physical forward problem to the whitened likelihood consumed by the posterior sampler. The recording operator $\mathbf{A}_0$ generates the observation, while the nominal operator $\mathbf{A}$ is used for reconstruction; from the discrepancy panel onward, all quantities are expressed in reduced coordinates.

The coordinate reduction gives i.i.d. Gaussian measurement noise, as assumed by the spectral posterior sampler. The discrepancy $\mathbf{d}$ vanishes in the nominal condition.

4.3 Approximation-error likelihood

We model the reduced operator discrepancy $\mathbf{d}$ of Equation (11) with a signal-conditional variant of the Bayesian approximation-error framework [19]. Specifically, we approximate its dependence on the heart-template potentials $\mathbf{h}$ through the nominal signal $\mathbf{s} = \mathbf{A}\mathbf{h}$.

For each physical time point t, we model the corresponding columns $\mathbf{s}_t$ and $\mathbf{d}_t$ as jointly Gaussian,

$$\begin{pmatrix} \mathbf{s}_t \\ \mathbf{d}_t \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \boldsymbol{\mu}_\mathrm{s} \\ \boldsymbol{\mu}_\mathrm{d} \end{pmatrix}, \begin{pmatrix} \mathbf{C}_\mathrm{ss} & \mathbf{C}_\mathrm{sd} \\ \mathbf{C}_\mathrm{ds} & \mathbf{C}_\mathrm{dd} \end{pmatrix} \right), \tag{12}$$

with the same means and covariance blocks across time. The joint model implies

$$\mathbf{d}_t \mid \mathbf{s}_t \sim \mathcal{N}(\boldsymbol{\mu}_\mathrm{d} + \mathbf{L}(\mathbf{s}_t - \boldsymbol{\mu}_\mathrm{s}), \boldsymbol{\Sigma}_\mathrm{r}), \tag{13}$$

with linear predictor $\mathbf{L} = \mathbf{C}_\mathrm{ds}\mathbf{C}_\mathrm{ss}^{-1}$ and residual covariance $\boldsymbol{\Sigma}_\mathrm{r} = \mathbf{C}_\mathrm{dd} - \mathbf{L}\mathbf{C}_\mathrm{sd}$. Under the joint-Gaussian model, the residual $\mathbf{r}_t = \mathbf{d}_t - \boldsymbol{\mu}_\mathrm{d} - \mathbf{L}(\mathbf{s}_t - \boldsymbol{\mu}_\mathrm{s})$ follows $\mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathrm{r})$ and is independent of the signal $\mathbf{s}_t$. The residual covariance is shared across time and our approximate likelihood treats the residuals as independent across time; see Appendix A.2 for a full derivation.

Corrected likelihood. We estimate the model parameters separately for each discrepancy source, using training samples of $\mathbf{h}$ paired with recording operators $\mathbf{A}_0$ drawn from the corresponding perturbation distribution. In practice, we compute $\mathbf{L}$ using trace-scaled ridge regularization of $\mathbf{C}_\mathrm{ss}$ and obtain $\boldsymbol{\Sigma}_\mathrm{r}$ from the fitted residuals using Ledoit–Wolf shrinkage [24]; details are provided in Appendix C.4.

We absorb the conditional-mean predictor into a corrected operator $\tilde{\mathbf{A}} = (\mathbf{I} + \mathbf{L})\mathbf{A}$ and subtract the constant bias $\boldsymbol{\mu}_\mathrm{d} - \mathbf{L}\boldsymbol{\mu}_\mathrm{s}$ from each column of $\mathbf{b}$ to obtain the bias-corrected observation $\tilde{\mathbf{b}}$. The corrected observation model reads $\tilde{\mathbf{b}} = \tilde{\mathbf{A}}\mathbf{h} + \tilde{\boldsymbol{\varepsilon}}$, where the combined noise has columns $\tilde{\boldsymbol{\varepsilon}}_t = \mathbf{r}_t + \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathrm{r} + \sigma^2 \mathbf{I}_{n_\mathrm{B}-1})$. Only the residual mismatch enters the noise distribution.

Whitened likelihood. The combined noise is correlated across measurement coordinates. Following [30], we whiten the corrected observation and operator to obtain the isotropic Gaussian noise model required by the spectral posterior sampler. With $\mathbf{W} = (\boldsymbol{\Sigma}_\mathrm{r} + \sigma^2 \mathbf{I}_{n_\mathrm{B}-1})^{-1/2}$, we set $\mathbf{A}_\mathrm{W} = \mathbf{W}\tilde{\mathbf{A}}$ and $\mathbf{b}_\mathrm{W} = \mathbf{W}\tilde{\mathbf{b}}$. The whitened noise $\boldsymbol{\varepsilon}_\mathrm{W} = \mathbf{W}\tilde{\boldsymbol{\varepsilon}}$ has columns $\boldsymbol{\varepsilon}_{\mathrm{W},t} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{n_\mathrm{B}-1})$, and the whitened observation model reads

$$\mathbf{b}_\mathrm{W} = \mathbf{A}_\mathrm{W}\mathbf{h} + \boldsymbol{\varepsilon}_\mathrm{W}. \tag{14}$$

We combine the whitened likelihood with a frozen unconditional prior trained on heart-template potentials. The prior remains fixed across discrepancy conditions. Appendix B.2.1 describes its architecture. For each observation, the spectral posterior sampler consumes $(\mathbf{A}_\mathrm{W}, \mathbf{b}_\mathrm{W})$ and returns K approximate posterior draws. Their Monte Carlo mean defines the point estimate. Figure 2 illustrates the resulting pipeline.

5 Experiments

We first describe the datasets and operator-discrepancy settings (Section 5.1). We then introduce the evaluated baselines (Section 5.2) and define the metrics used for quantitative comparison (Section 5.3). Finally, we report quantitative results (Section 5.4) and qualitative insights (Section 5.5).

5.1 Data

Geometries and physiological simulations. We use 1000 synthetic four-chamber heart meshes from a CT-based statistical shape model [36] and retain the left and right ventricles. Their outer surface defines the heart surface used for reconstruction. Torso anatomies originate from 1228 CT scans in the TotalSegmentator dataset [45, 46]. We retain 167 torso geometries with complete coverage, an extracted body surface, and the lungs included in the volume conductor. We combine each heart with multiple anatomically compatible torso geometries and register it to the ventricular envelope extracted from the corresponding torso CT. A fixed vest layout provides $n_{\mathrm{B}} = 176$ body-surface electrode locations. We split heart and torso identities before pairing, resulting in 1700 training pairs, 100 validation pairs, and 250 physiological test pairs.

We simulate ventricular electrophysiology with openCARP [34] following the MedalCare-XL protocol [14], which couples a pseudo-bidomain formulation with Mitchell–Schaeffer ionic dynamics, transversely isotropic conduction along the myocardial fibers, and a fast endocardial layer (Appendix E); each simulation spans 450 ms sampled every 5 ms. Three physiological rounds resample the action-potential-duration field and five endocardial initiation sites, producing 5100 training sequences over the QRST interval, which covers ventricular depolarization and repolarization. We map the simulated potentials to the heart-template representation $\mathbf{h} \in \mathbb{R}^{1800\times 91}$ defined in Section 4.1. We generate the noise-free training and validation body-template potentials directly as $\mathbf{A}_0\mathbf{h}$, with $\mathbf{A}_0 = \mathbf{A}$ in these nominal splits.

Operator discrepancy. We evaluate the same 250 physiological test pairs in one matched-operator reference and two operator-discrepancy test sets. In the matched-operator reference, the recording and reconstruction operators coincide, $\mathbf{A}_0 = \mathbf{A}$. In each discrepancy test set, the recording operator $\mathbf{A}_0$ produces the observations, while every method with an explicit forward operator receives only the nominal operator $\mathbf{A}$. The heart–torso pairs and electrode locations remain fixed across conditions.

For lung-conductivity discrepancy, we model variability associated with different lung inflation states. The nominal operator $\mathbf{A}$ assigns piecewise-constant conductivities to two tissue compartments, the lungs (0.05 S/m) and the remaining torso (0.22 S/m). For each test case, the recording operator $\mathbf{A}_0$ uses a lung conductivity drawn from $\mathcal{U}[0.03, 0.15]$ S/m while keeping the torso conductivity fixed [22, 32]. The approximation-error model uses operators with lung conductivities drawn from the narrower interval $[0.04, 0.10]$ S/m.

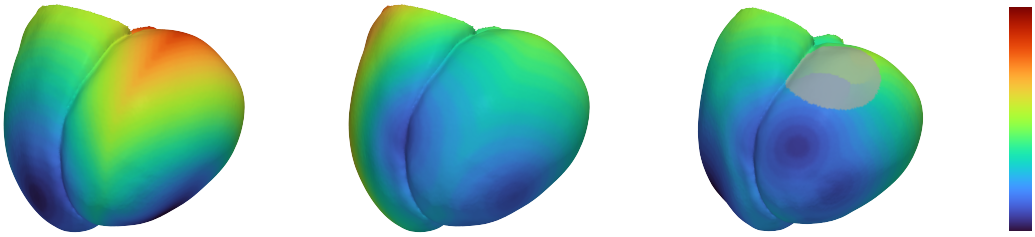
For subcutaneous-fat discrepancy, the nominal operator $\mathbf{A}$ treats the torso outside the lungs as homogeneous, whereas the recording operator $\mathbf{A}_0$ includes a subcutaneous layer with conductivity 0.04 S/m directly beneath the body surface. Its thickness depends on the torso-size category and ranges from 3 to 20 mm across the test cases. The approximation-error model uses operators with thicknesses drawn uniformly from $[5, 15]$ mm.

Pathological data. In addition to physiological data, we evaluate four severity configurations for each of left bundle branch block (LBBB), right bundle branch block (RBBB), and myocardial scar. We simulate every configuration on 60 heart–torso geometries from the test split, yielding 720 pathological sequences per operator-discrepancy condition. No pathological sequence enters training or validation. Figure 3 illustrates the three pathology types, and Table 10 lists all configurations.

Measurement noise. We draw Gaussian noise independently over electrodes and time and apply the centering matrix $\mathbf{C}$ from Section 4.1. We choose its standard deviation to target the sequence-level Frobenius signal-to-noise ratio (SNR)

$$\mathrm{SNR}_{\mathrm{dB}} = 20 \log_{10} \frac{\|\mathbf{A}_0\mathbf{h}\|_F}{\|\boldsymbol{\varepsilon}\|_F}. \tag{15}$$

We use 40 dB for the main comparison to emphasize operator discrepancy and 30 dB for an additional evaluation under stronger measurement noise. Appendix D.1 quantifies the relative magnitudes of measurement noise and operator discrepancy underlying this choice.



(a) Left bundle branch block. (b) Right bundle branch block. (c) Myocardial scar.

Figure 3: **Illustrative examples of pathologies.** Ground-truth local activation-time (LAT) maps for the selected pathologies. The gray overlay marks the modeled scar region. The bundle-branch-block examples share the same heart geometry, whereas the scar example has a different anatomy.

5.2 Baselines

We consider baselines from three categories: variational regularization, amortized inference, and posterior sampling. All baselines are evaluated on the same test cases as OpDis-PS, and every method with an explicit forward operator uses the nominal operator $\mathbf{A}$. Appendix B details the architectures of the data-driven methods and Appendix C reports the full settings for all methods.

Variational regularization. First-order generalized Tikhonov penalizes the template surface gradient and selects λ case-wise by generalized cross-validation (GCV) [15, 27]. Spatio-temporal Tikhonov adds a second-order temporal-difference penalty [7]. Both variants discard reduced-Gram eigenmodes with $w_i \leq 10^{-6} w_{\max}$, making them truncated singular value decomposition (TSVD)–Tikhonov hybrids. Our Total Variation baseline applies the spatial template-gradient penalty [37] to each time point independently and uses the Chambolle–Pock primal–dual algorithm [9].

Amortized inference. We consider a transformer and a conditional diffusion model. Neither baseline explicitly uses the forward operator during inference. Both are trained on nominal pairs with average-referenced Gaussian noise augmentation and use nominal validation pairs for checkpoint selection. Both models map body-template potentials to heart-template potentials. To this end, the transformer baseline implements the architecture of Vaswani et al. [42], whereas the diffusion baseline uses a Palette-style conditional diffusion model [39], following recent geometry-free conditional diffusion in ECGI [17].

Posterior sampling. We consider DDRM [21] and DPS [10]. DDRM combines a frozen unconditional prior over heart-template potentials based on denoising diffusion probabilistic models [16] with the nominal operator $\mathbf{A}$ and an additive noise model. DPS uses the same prior and nominal operator $\mathbf{A}$ for gradient-based guidance, scaled by a weight selected on the validation set.

5.3 Metrics

We evaluate reconstructed heart-template potentials $\widehat{\mathbf{h}}$ against the ground truth $\mathbf{h}$ using five complementary metrics. The relative error (RE) measures global reconstruction error, while the spatial and temporal correlation coefficients (sCC, tCC) measure spatial-pattern and temporal-waveform agreement. The activation time error (ATE) quantifies local activation-time differences in milliseconds, and the localization error (LE) gives the distance between the reconstructed and true earliest activation sites on the heart surface in millimeters. Formal definitions and the statistical testing protocol are provided in Appendix A.4.

5.4 Quantitative results

Tables 1 and 2 report the pathological and physiological evaluations at 40 dB. OpDis-PS achieves the lowest RE and the highest sCC and tCC under both discrepancy sources on both datasets, with the lowest or near-lowest ATE and LE. Variational methods remain clearly less accurate than the leading data-driven methods. Amortized methods are competitive on physiological activation but, together

Table 1: **Quantitative results (pathological data; 40 dB).** We evaluate two sources of operator discrepancy (lung conductivity and subcutaneous fat) on 720 pathological test cases at 40 dB against variational regularization (R), amortized inference (A), and posterior sampling (P) baselines. Data-driven results are mean $\pm$ standard deviation over five training runs; $\star$ denotes $p < 0.05$ for the best versus second-best data-driven method (Appendix A.4).

Discrepancy		Method	Metric				
			RE↓	ATE [ms]↓	LE [mm]↓	sCC↑	tCC↑
lung conductivity	R	Tikhonov (spatial) [27]	0.737	12.058	23.878	0.646	0.742
		Tikhonov (spatio-temporal) [7]	0.720	7.927	9.786	0.667	0.755
		Total Variation [37]	0.618	8.559	11.470	0.733	0.763
	A	Transformer [42]	0.561 ± 0.002	7.143 ± 0.080	7.349 ± 0.251	0.746 ± 0.005	0.792 ± 0.002
		Palette [39]	0.517 ± 0.003	6.384 ± 0.219	7.940 ± 0.513	0.786 ± 0.002	0.822 ± 0.003
	P	DPS [10]	0.617 ± 0.004	11.214 ± 0.180	12.197 ± 0.566	0.724 ± 0.003	0.739 ± 0.003
		DDRM [21]	0.425 ± 0.003	$\mathbf{3.599 \pm 0.077}$	$\mathbf{2.688 \pm 0.070}^\star$	0.849 ± 0.002	0.896 ± 0.002
		OpDis-PS (**ours**)	$\mathbf{0.393 \pm 0.003}^\star$	3.621 ± 0.091	2.927 ± 0.077	$\mathbf{0.853 \pm 0.003}^\star$	$\mathbf{0.903 \pm 0.002}^\star$
subcutaneous fat	R	Tikhonov (spatial) [27]	0.784	11.739	24.112	0.630	0.722
		Tikhonov (spatio-temporal) [7]	0.756	7.640	9.635	0.655	0.738
		Total Variation [37]	0.635	8.488	10.439	0.732	0.764
	A	Transformer [42]	0.565 ± 0.002	7.106 ± 0.085	7.366 ± 0.268	0.746 ± 0.005	0.792 ± 0.002
		Palette [39]	0.520 ± 0.003	6.495 ± 0.246	8.150 ± 0.662	0.785 ± 0.002	0.820 ± 0.003
	P	DPS [10]	0.616 ± 0.004	10.649 ± 0.169	10.515 ± 0.357	0.729 ± 0.003	0.751 ± 0.003
		DDRM [21]	0.460 ± 0.003	3.605 ± 0.085	3.383 ± 0.085	0.852 ± 0.002	0.892 ± 0.002
		OpDis-PS (**ours**)	$\mathbf{0.395 \pm 0.003}^\star$	$\mathbf{3.602 \pm 0.091}$	$\mathbf{3.160 \pm 0.086}^\star$	$\mathbf{0.853 \pm 0.003}$	$\mathbf{0.903 \pm 0.002}^\star$

RE = relative error; ATE = activation time error; LE = localization error; sCC = spatial corr. coefficient; tCC = temporal corr. coefficient

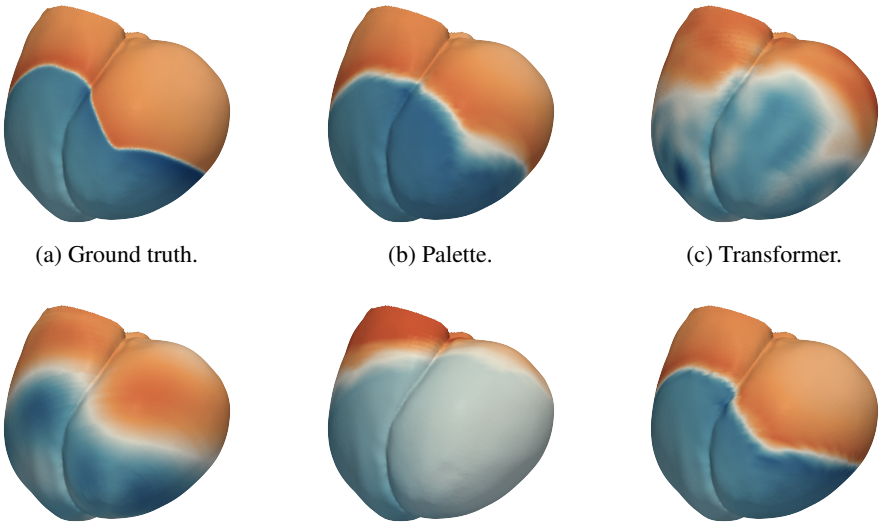
Table 2: **Quantitative results (physiological data; 40 dB).** We evaluate the same two sources of operator discrepancy (lung conductivity and subcutaneous fat) on 250 physiological test cases at 40 dB against variational regularization (R), amortized inference (A), and posterior sampling (P) baselines. Data-driven results are mean $\pm$ standard deviation over five training runs; $\star$ denotes $p < 0.05$ for the best versus second-best data-driven method (Appendix A.4).

Discrepancy		Method	Metric				
			RE↓	ATE [ms]↓	LE [mm]↓	sCC↑	tCC↑
lung conductivity	R	Tikhonov (spatial)	0.765	7.680	15.662	0.587	0.751
		Tikhonov (spatio-temporal)	0.749	5.825	7.408	0.608	0.763
		Total Variation	0.583	6.303	10.199	0.745	0.796
	A	Transformer	0.395 ± 0.002	2.822 ± 0.045	4.379 ± 0.192	0.814 ± 0.009	0.905 ± 0.001
		Palette	0.379 ± 0.003	2.738 ± 0.029	5.380 ± 0.338	0.851 ± 0.002	0.910 ± 0.002
	P	DPS	0.482 ± 0.004	4.466 ± 0.087	8.643 ± 0.378	0.798 ± 0.006	0.853 ± 0.003
		DDRM	0.372 ± 0.001	2.221 ± 0.011	$\mathbf{2.108 \pm 0.101}^\star$	0.837 ± 0.004	0.926 ± 0.000
		OpDis-PS (**ours**)	$\mathbf{0.313 \pm 0.001}^\star$	$\mathbf{2.033 \pm 0.008}^\star$	2.198 ± 0.146	$\mathbf{0.858 \pm 0.003}^\star$	$\mathbf{0.943 \pm 0.000}^\star$
subcutaneous fat	R	Tikhonov (spatial)	0.775	7.351	18.790	0.579	0.741
		Tikhonov (spatio-temporal)	0.750	5.481	7.595	0.606	0.755
		Total Variation	0.601	6.320	9.218	0.740	0.795
	A	Transformer	0.399 ± 0.004	2.883 ± 0.056	4.522 ± 0.228	0.810 ± 0.008	0.903 ± 0.002
		Palette	0.376 ± 0.003	2.698 ± 0.025	4.990 ± 0.223	0.854 ± 0.002	0.912 ± 0.002
	P	DPS	0.501 ± 0.005	4.513 ± 0.094	7.550 ± 0.380	0.793 ± 0.006	0.851 ± 0.003
		DDRM	0.409 ± 0.002	2.365 ± 0.013	2.684 ± 0.125	0.842 ± 0.004	0.920 ± 0.001
		OpDis-PS (**ours**)	$\mathbf{0.321 \pm 0.001}^\star$	$\mathbf{2.109 \pm 0.005}^\star$	$\mathbf{2.470 \pm 0.090}^\star$	$\mathbf{0.855 \pm 0.004}$	$\mathbf{0.940 \pm 0.000}^\star$

RE = relative error; ATE = activation time error; LE = localization error; sCC = spatial corr. coefficient; tCC = temporal corr. coefficient

with DPS, degrade most on pathological activation. Appendix Table 4 provides the matched-operator physiological reference, Appendix D.3 reports the evaluations under stronger measurement noise, and Appendix D.4 ablates the discrepancy model.

Discussion. The results indicate that explicit operator information and a learned prior are complementary. Variational methods use the operator explicitly but rely on generic smoothness assumptions, which limits their accuracy. Amortized methods learn the observation process implicitly and perform well on activation patterns resembling the training data, but generalize less well to pathological activation. Posterior sampling unites both strengths, pairing an explicit operator with a learned prior, which makes the likelihood a direct lever for incorporating operator knowledge. Since OpDis-PS shares the prior, nominal operator, and sampler with DDRM, its consistently lower RE isolates the benefit of the discrepancy-aware likelihood.



(a) Ground truth. (b) Palette. (c) Transformer.

(d) Tikhonov (spatial). (e) DPS. (f) OpDis-PS (**ours**).

Figure 4: **Qualitative insights.** Ground-truth and reconstructed heart-template potentials for an illustrative LBBB case under lung-conductivity discrepancy at $t = 110\,\mathrm{ms}$ and 40 dB. OpDis-PS and Palette reproduce the activation front, which the Transformer blurs and DPS displaces.

5.5 Qualitative insights

Figure 4 compares heart-template potentials reconstructed for an illustrative LBBB case under lung-conductivity discrepancy at $t = 110\,\mathrm{ms}$ and 40 dB. Here, OpDis-PS and Palette recover the sharp activation front most clearly, with OpDis-PS showing the closest visual agreement with the ground truth. Spatial Tikhonov captures the characteristic LBBB-induced displacement of the front, but broadens it through spatial regularization. The Transformer yields a more diffuse front, whereas DPS exhibits a larger spatial displacement and attenuated contrast.

6 Limitations

Our evaluation relies on simulated electrophysiology and two discrepancy sources. Real recordings may contain several sources of mismatch, including variations not represented in the operator distribution used to estimate the approximation-error model. In practice, this distribution would need to be constructed from plausible variation in anatomy, tissue conductivities, electrode placement, and acquisition conditions. In our experiments, it is narrower than the distribution generating the test operators (Section 5.1), but we have not systematically varied the degree of misspecification. The discrepancy model is linear-Gaussian and treats residuals as independent across time, although a recording shares one operator. Its linear predictor explains less of the lung-conductivity discrepancy than of the subcutaneous-fat discrepancy (held-out R^2 of 0.441 versus 0.937). Finally, the improvements are metric-dependent. Future work should assess their clinical relevance and uncertainty calibration on real recordings.

7 Conclusion

We introduced OpDis-PS, a diffusion posterior sampler for reconstruction under forward-operator discrepancy. A signal-conditional approximation-error model corrects the nominal operator and incorporates the remaining mismatch into a whitened likelihood, while leaving the unconditional diffusion prior fixed. Under both discrepancy sources at 40 dB, OpDis-PS achieves the lowest global reconstruction error and the highest spatial and temporal correlation, with its advantage over amortized baselines widening on pathological activation patterns excluded from training. These results show that adapting the likelihood to operator discrepancy can improve diffusion posterior sampling without retraining the generative prior.

Acknowledgments and Disclosure of Funding

The authors acknowledge the HPC resources provided by the Erlangen National High-Performance Computing Center (NHR@FAU) of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU). We thank the developers of openCARP, the authors of the TotalSegmentator dataset, and the authors of the statistical shape model of Rodero et al. [36] for making their resources publicly available.

References

[1] Furkan Aldemir and Yesim Serinagaoglu Dogrusoz. Compensation of model errors in electrocardiographic imaging using Bayesian estimation. In *Computing in Cardiology (CinC)*, 2021.

[2] Tania Bacoyannis, Buntheng Ly, Nicolas Cedilnik, Hubert Cochet, and Maxime Sermesant. Deep learning formulation of electrocardiographic imaging integrating image and signal information with data-driven regularization. *EP Europace*, 2021.

[3] Jason Bayer, Anton J. Prassl, Ali Pashaei, Juan F. Gomez, Antonio Frontera, Aurel Neic, Gernot Plank, and Edward J. Vigmond. Universal ventricular coordinates: A generic framework for describing position within the heart and transferring data. *Medical Image Analysis*, 2018.

[4] Jason D. Bayer, Robert C. Blake, Gernot Plank, and Natalia A. Trayanova. A novel rule-based algorithm for assigning myocardial fiber orientation to computational heart models. *Annals of Biomedical Engineering*, 2012.

[5] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[6] C. P. Bradley, A. J. Pullan, and P. J. Hunter. Effects of material properties and geometry on electrocardiographic forward simulations. *Annals of Biomedical Engineering*, 2000.

[7] Dana H. Brooks, George F. Ahmad, Robert S. MacLeod, and George M. Maratos. Inverse electrocardiography by simultaneous imposition of multiple constraints. *IEEE Transactions on Biomedical Engineering*, 1999.

[8] Kamil Bujnarowski, Pietro Bonizzi, Matthijs Cluitmans, Ralf Peeters, and Joël Karel. CT-scan free neural network-based reconstruction of heart surface potentials from ECG recordings. In *Computing in Cardiology (CinC)*, 2020.

[9] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 2011.

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

[11] Matthijs Cluitmans, Dana H. Brooks, Rob MacLeod, Olaf Dössel, María S. Guillem, Peter M. van Dam, Jana Svehlikova, Bin He, John Sapp, Linwei Wang, and Laura Bear. Validation and opportunities of electrocardiographic imaging: From technical achievements to clinical applications. *Frontiers in Physiology*, 2018.

[12] Matthijs J. M. Cluitmans, Michael Clerx, Nele Vandersickel, Ralf L. M. Peeters, Paul G. A. Volders, and Ronald L. Westra. Physiology-based regularization of the electrocardiographic inverse problem. *Medical & Biological Engineering & Computing*, 2017.

[13] Sandesh Ghimire, Jwala Dhamala, Prashnna Kumar Gyawali, John L. Sapp, B. Milan Horáček, and Linwei Wang. Generative modeling and inverse imaging of cardiac transmembrane potential. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, 2018.

[14] Karli Gillette, Matthias A. F. Gsell, Claudia Nagel, Jule Bender, Benjamin Winkler, Steven E. Williams, Markus Bär, Tobias Schäffter, Olaf Dössel, Gernot Plank, and Axel Loewe. MedalCare-XL: 16,900 healthy and pathological synthetic 12 lead ECGs from electrophysiological simulations. *Scientific Data*, 2023.

[15] Gene H. Golub, Michael Heath, and Grace Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 1979.

[16] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, 2020.

[17] Ramiro Valdes Jara and Adam Meyers. Geometry-free conditional diffusion modeling for solving the inverse electrocardiography problem. *arXiv preprint arXiv:2601.18615*, 2026.

[18] Xiajun Jiang, Maryam Toloubidokhti, Jake Bergquist, Brian Zenger, Wilson W. Good, Rob S. MacLeod, and Linwei Wang. Improving generalization by learning geometry-dependent and physics-based reconstruction of image sequences. *IEEE Transactions on Medical Imaging*, 2023.

[19] Jari Kaipio and Erkki Somersalo. Statistical inverse problems: Discretization, model reduction and inverse crimes. *Journal of Computational and Applied Mathematics*, 2007.

[20] Amel Karoui, Laura Bear, Pauline Migerditichan, and Nejib Zemzemi. Evaluation of fifteen algorithms for the resolution of the electrocardiography imaging inverse problem using ex-vivo and in-silico data. *Frontiers in Physiology*, 2018.

[21] Bahjat Kawar, Michael Elad, Stefano Ermon, and Jiaming Song. Denoising diffusion restoration models. In *Advances in Neural Information Processing Systems*, 2022.

[22] David U. J. Keller, Frank M. Weber, Gunnar Seemann, and Olaf Dössel. Ranking the influence of tissue conductivities on forward-calculated ECGs. *IEEE Transactions on Biomedical Engineering*, 2010.

[23] Ruth N. Klepfer, Christopher R. Johnson, and Robert S. MacLeod. The effects of inhomogeneities and anisotropies on electrocardiographic fields: A 3-D finite-element study. *IEEE Transactions on Biomedical Engineering*, 1997.

[24] Olivier Ledoit and Michael Wolf. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 2004.

[25] Robert S. MacLeod and Dana H. Brooks. Recent progress in inverse problems in electrocardiology. *IEEE Engineering in Medicine and Biology Magazine*, 1998.

[26] George A. Mensah, Valentin Fuster, Christopher J. L. Murray, Gregory A. Roth, and Global Burden of Cardiovascular Diseases and Risks Collaborators. Global burden of cardiovascular diseases and risks, 1990–2022. *Journal of the American College of Cardiology*, 2023.

[27] Barbara J. Messinger-Rapport and Yoram Rudy. Regularization of the inverse problem in electrocardiography: A model study. *Mathematical Biosciences*, 1988.

[28] Matija Milanič, Vojko Jazbinšek, Robert S. MacLeod, Dana H. Brooks, and Rok Hren. Assessment of regularization techniques for electrocardiographic imaging. *Journal of Electrocardiology*, 2014.

[29] Colleen C. Mitchell and David G. Schaeffer. A two-current model for the dynamics of cardiac membrane. *Bulletin of Mathematical Biology*, 2003.

[30] Niki Nezakati, Arnab Ghosh, Amit Roy-Chowdhury, and Vishwanath Saragadam. CARD: Correlation aware restoration with diffusion. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2026.

[31] Alex Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *Proceedings of the 38th International Conference on Machine Learning*, 2021.

[32] P. Nopp, E. Rapp, H. Pfützner, H. Nakesch, and C. Ruhsam. Dielectric properties of lung tissue as a function of air content. *Physics in Medicine and Biology*, 1993.

[33] Hans S. Oster, Bruno Taccardi, Robert L. Lux, Philip R. Ershler, and Yoram Rudy. Noninvasive electrocardiographic imaging: Reconstruction of epicardial potentials, electrograms, and isochrones and localization of single and multiple electrocardiac events. *Circulation*, 1997.

[34] Gernot Plank, Axel Loewe, Aurel Neic, Christoph Augustin, Yung-Lin Huang, Matthias A. F. Gsell, Elias Karabelas, Mark Nothstein, Anton J. Prassl, Jorge Sánchez, Gunnar Seemann, and Edward J. Vigmond. The openCARP simulation environment for cardiac electrophysiology. *Computer Methods and Programs in Biomedicine*, 2021.

[35] Charulatha Ramanathan and Yoram Rudy. Electrocardiographic imaging: I. Effect of torso inhomogeneities on body surface electrocardiographic potentials. *Journal of Cardiovascular Electrophysiology*, 2001.

[36] Cristobal Rodero, Marina Strocchi, Maciej Marciniak, Stefano Longobardi, John Whitaker, Mark D. O'Neill, Karli Gillette, Christoph M. Augustin, Gernot Plank, Edward J. Vigmond, Pablo Lamata, and Steven A. Niederer. Linking statistical shape models and simulated function in the healthy adult human heart. *PLOS Computational Biology*, 2021.

[37] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 1992.

[38] Yoram Rudy and Barbara J. Messinger-Rapport. The inverse problem in electrocardiography: solutions in terms of epicardial potentials. *Critical Reviews in Biomedical Engineering*, 1988.

[39] Chitwan Saharia, William Chan, Huiwen Chang, Chris A. Lee, Jonathan Ho, Tim Salimans, David J. Fleet, and Mohammad Norouzi. Palette: Image-to-image diffusion models. In *ACM SIGGRAPH 2022 Conference Proceedings*, 2022.

[40] Tim Salimans and Jonathan Ho. Progressive distillation for fast sampling of diffusion models. In *International Conference on Learning Representations*, 2022.

[41] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021.

[42] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, 2017.

[43] Dafang Wang, Robert M. Kirby, and Chris R. Johnson. Finite-element-based discretization and regularization strategies for 3-d inverse electrocardiography. *IEEE Transactions on Biomedical Engineering*, 2011.

[44] Tiantian Wang, Joël Karel, Niels Osnabrugge, Kurt Driessens, Job Stoks, Matthijs Cluitmans, Paul Volders, Pietro Bonizzi, and Ralf L. M. Peeters. Deep learning based estimation of heart surface potentials. *Artificial Intelligence in Medicine*, 2025.

[45] Jakob Wasserthal. Dataset with segmentations of 117 important anatomical structures in 1228 CT images. Zenodo, https://doi.org/10.5281/zenodo.10047292, 2023.

[46] Jakob Wasserthal, Hanns-Christian Breit, Manfred T. Meyer, Maurice Pradella, Daniel Hinck, Alexander W. Sauter, Tobias Heye, Daniel T. Boll, Joshy Cyriac, Shan Yang, Michael Bach, and Martin Segeroth. TotalSegmentator: Robust segmentation of 104 anatomic structures in CT images. *Radiology: Artificial Intelligence*, 2023.

A Mathematical details

This appendix provides the mathematical details omitted from the main text. We first derive the reduced average-referenced observation model (Section A.1) and the signal-conditional approximation-error likelihood (Section A.2). We then describe the heart-template construction (Section A.3) and provide precise definitions of the evaluation metrics (Section A.4).

A.1 Coordinate reduction

The centering matrix $\mathbf{C} = \mathbf{I} - \frac{\mathbf{1}\mathbf{1}^\mathsf{T}}{n_\mathrm{B}}$ is symmetric and idempotent with rank $n_\mathrm{B} - 1$, and $\mathbf{C}\mathbf{1} = \mathbf{0}$, so it removes any constant offset across the electrodes. The centered noise $\boldsymbol{\varepsilon}_t = \mathbf{C}\mathbf{z}_t$ then has the rank-deficient covariance $\sigma^2\mathbf{C}$. To recover an isotropic noise model, we choose $\mathbf{R} \in \mathbb{R}^{n_\mathrm{B} \times (n_\mathrm{B}-1)}$ whose columns form an orthonormal basis of the range of $\mathbf{C}$, so that

$$\mathbf{R}^\mathsf{T}\mathbf{R} = \mathbf{I}_{n_\mathrm{B}-1}, \qquad \mathbf{R}\mathbf{R}^\mathsf{T} = \mathbf{C}. \tag{16}$$

Left-multiplying the observation model $\mathbf{b} = \mathbf{A}\mathbf{h} + \boldsymbol{\varepsilon}$ by $\mathbf{R}^\mathsf{T}$ keeps the model linear,

$$\mathbf{R}^\mathsf{T}\mathbf{b} = \mathbf{R}^\mathsf{T}\mathbf{A}\,\mathbf{h} + \mathbf{R}^\mathsf{T}\boldsymbol{\varepsilon}, \tag{17}$$

and by (16) the noise covariance becomes isotropic,

$$\mathrm{Cov}(\mathbf{R}^\mathsf{T}\boldsymbol{\varepsilon}_t) = \sigma^2\mathbf{R}^\mathsf{T}\mathbf{C}\mathbf{R} = \sigma^2\mathbf{R}^\mathsf{T}(\mathbf{R}\mathbf{R}^\mathsf{T})\mathbf{R} = \sigma^2\mathbf{I}_{n_\mathrm{B}-1}. \tag{18}$$

The reduced noise is therefore isotropic, as the spectral posterior sampler assumes. With slight abuse of notation, we reuse $\mathbf{b}$, $\mathbf{A}$, $\mathbf{A}_0$, and $\boldsymbol{\varepsilon}$ throughout the main text for the reduced quantities $\mathbf{R}^\mathsf{T}\mathbf{b}$, $\mathbf{R}^\mathsf{T}\mathbf{A}$, $\mathbf{R}^\mathsf{T}\mathbf{A}_0$, and $\mathbf{R}^\mathsf{T}\boldsymbol{\varepsilon}$.

A.2 Signal-conditional approximation-error derivation

We derive the correction of Section 4.3 for a single time point t. The standard conditional-Gaussian formula [5] applied to the joint model in Equation (12) gives

$$\mathbf{d}_t = \boldsymbol{\mu}_\mathrm{d} + \mathbf{L}(\mathbf{s}_t - \boldsymbol{\mu}_\mathrm{s}) + \mathbf{r}_t, \qquad \mathbf{r}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\mathrm{r}), \tag{19}$$

where

$$\mathbf{L} = \mathbf{C}_\mathrm{ds}\mathbf{C}_\mathrm{ss}^{-1}, \qquad \boldsymbol{\Sigma}_\mathrm{r} = \mathbf{C}_\mathrm{dd} - \mathbf{L}\mathbf{C}_\mathrm{sd}. \tag{20}$$

Under the joint-Gaussian model, the residual distribution does not depend on the nominal signal $\mathbf{s}_t$.

Substituting $\mathbf{s}_t = \mathbf{A}\mathbf{h}_t$ into the reduced observation model $\mathbf{b}_t = \mathbf{A}\mathbf{h}_t + \mathbf{d}_t + \boldsymbol{\varepsilon}_t$ gives

$$\begin{aligned} \mathbf{b}_t &= \mathbf{A}\mathbf{h}_t + \boldsymbol{\mu}_\mathrm{d} + \mathbf{L}(\mathbf{A}\mathbf{h}_t - \boldsymbol{\mu}_\mathrm{s}) + \mathbf{r}_t + \boldsymbol{\varepsilon}_t & (21)\\ &= (\mathbf{I} + \mathbf{L})\mathbf{A}\mathbf{h}_t + (\boldsymbol{\mu}_\mathrm{d} - \mathbf{L}\boldsymbol{\mu}_\mathrm{s}) + \mathbf{r}_t + \boldsymbol{\varepsilon}_t. & (22) \end{aligned}$$

Collecting the signal-dependent part into a corrected operator and the constant part into a corrected observation,

$$\tilde{\mathbf{A}} = (\mathbf{I} + \mathbf{L})\mathbf{A}, \qquad \tilde{\mathbf{b}}_t = \mathbf{b}_t - (\boldsymbol{\mu}_\mathrm{d} - \mathbf{L}\boldsymbol{\mu}_\mathrm{s}), \tag{23}$$

the corrected model reads

$$\tilde{\mathbf{b}}_t = \tilde{\mathbf{A}}\mathbf{h}_t + \tilde{\boldsymbol{\varepsilon}}_t, \qquad \tilde{\boldsymbol{\varepsilon}}_t = \mathbf{r}_t + \boldsymbol{\varepsilon}_t. \tag{24}$$

Since the measurement noise $\boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_{n_\mathrm{B}-1})$ is independent of the residual discrepancy $\mathbf{r}_t$, their covariances add, yielding

$$\tilde{\boldsymbol{\varepsilon}}_t \sim \mathcal{N}\big(\mathbf{0}, \boldsymbol{\Sigma}_\mathrm{r} + \sigma^2\mathbf{I}_{n_\mathrm{B}-1}\big)\,. \tag{25}$$

We therefore define the whitening matrix

$$\mathbf{W} = \big(\boldsymbol{\Sigma}_\mathrm{r} + \sigma^2\mathbf{I}_{n_\mathrm{B}-1}\big)^{-1/2}\,, \tag{26}$$

which maps the combined noise covariance to the identity. We estimate $\mathbf{L}$ and $\boldsymbol{\Sigma}_\mathrm{r}$ from finite training data as described in Appendix C.4. Applying the correction and whitening column-wise yields the sequence-level model in Equation (14).

A.3 Template-coordinate construction and surface mappings

We construct the template coordinates on the heart surface Γ_H of one reference heart. To localize the apex, we solve the surface Poisson problem

$$-\Delta_{\Gamma_\mathrm{H}} w = 1, \qquad w = 0 \quad \text{on the valve rims}, \tag{27}$$

whose solution w is a smooth intrinsic apex indicator. Its maximum defines the apex location, and the 3% of surface area with the largest values of w forms a finite apical cap.

Between this cap and the valve rims, we then solve the Laplace–Beltrami problem

$$-\Delta_{\Gamma_{\mathrm{H}}} u = 0, \qquad u = 0 \quad \text{on the apical cap}, \qquad u = 1 \quad \text{on the valve rims}, \tag{28}$$

whose harmonic solution u varies smoothly from apex to base. We discard the apical cap and the basal collar with $u > u_{\mathrm{cut}}$, where $u_{\mathrm{cut}} = 0.99$, and denote the remaining surface by Γ_{rep}.

Cumulative surface area across the level sets of u defines the apico-basal coordinate z_h and divides Γ_{rep} into 30 approximately equal-area bands. Within each band, cumulative surface area defines the periodic circumferential coordinate φ_h with 60 positions, and we align the circumferential origins of consecutive bands through their nearest surface points. This produces the regular 30×60 grid of heart-template locations, which we transfer to every other heart through the known node correspondence.

Let $\mathbf{M}_{\mathrm{H}}$ be the lumped heart-mesh mass matrix, with weights outside Γ_{rep} set to zero. The mass-weighted L^2 fit of the heart-mesh potentials Φ_{H} is

$$\mathbf{h} = \left(\mathbf{T}_{\mathrm{H}}^{\mathsf{T}} \mathbf{M}_{\mathrm{H}} \mathbf{T}_{\mathrm{H}}\right)^{-1} \mathbf{T}_{\mathrm{H}}^{\mathsf{T}} \mathbf{M}_{\mathrm{H}} \Phi_{\mathrm{H}}. \tag{29}$$

The same map $\mathbf{T}_{\mathrm{H}}$ interpolates heart-template potentials back to the heart-mesh nodes.

A.4 Metrics

Let $\widehat{\mathbf{h}}, \mathbf{h} \in \mathbb{R}^{n_{\mathrm{H}} \times T}$ denote the reconstructed and ground-truth heart-template potentials, with $n_{\mathrm{H}} = 1800$. The entry $h_{i,t}$ is the potential at heart-template location i and time point t, and consecutive time points are separated by 5 ms.

Relative error (RE). The global Frobenius relative error is given by

$$\mathrm{RE} = \frac{\|\widehat{\mathbf{h}} - \mathbf{h}\|_F}{\|\mathbf{h}\|_F}. \tag{30}$$

Spatial and temporal correlation coefficients (sCC, tCC). We compute sCC as the mean spatial Pearson correlation over time,

$$\mathrm{sCC} = \frac{1}{|\mathcal{T}|} \sum_{t \in \mathcal{T}} \frac{\sum_{i=1}^{n_{\mathrm{H}}} (\widehat{h}_{i,t} - \widehat{\mu}_t)(h_{i,t} - \mu_t)}{\sqrt{\sum_{i=1}^{n_{\mathrm{H}}} (\widehat{h}_{i,t} - \widehat{\mu}_t)^2 \sum_{i=1}^{n_{\mathrm{H}}} (h_{i,t} - \mu_t)^2}}. \tag{31}$$

Here $\mu_t = n_{\mathrm{H}}^{-1} \sum_{i=1}^{n_{\mathrm{H}}} h_{i,t}$ is the spatial mean of the ground truth and $\widehat{\mu}_t$ the corresponding reconstruction mean. The set $\mathcal{T}$ collects the time points with nondegenerate spatial variance in both the reconstruction and the ground truth. For tCC, we instead compute the Pearson correlation between reconstructed and true waveforms at each heart-template location and average over locations with nondegenerate temporal variance.

Activation time error (ATE). We estimate the temporal derivative $\dot{h}_{i,t}$ by finite differences and take the local activation time as the time of the steepest downstroke,

$$\mathrm{LAT}_i(\mathbf{h}) = (5\,\mathrm{ms}) \left(\operatorname*{arg\,max}_{t \in \{1,\dots,T\}} (-\dot{h}_{i,t}) - 1 \right), \tag{32}$$

where the trailing -1 places the first time point at 0 ms. ATE is the mean absolute activation-time error in milliseconds,

$$\mathrm{ATE} = \frac{1}{n_{\mathrm{H}}} \sum_{i=1}^{n_{\mathrm{H}}} \left| \mathrm{LAT}_i(\widehat{\mathbf{h}}) - \mathrm{LAT}_i(\mathbf{h}) \right|. \tag{33}$$

Localization error (LE). We call a heart-template location amplitude-gated if its temporal peak-to-peak amplitude reaches 30% of the maximum across locations. An early-breakthrough region is a connected cluster (link radius 12 mm) of gated locations whose LAT lies within 5 ms of the earliest gated LAT, and each region contributes its centroid in millimeters. Because LATs are resolved on the 5 ms sampling grid, several reconstructed regions can share the earliest gated LAT. Let $\{\mathbf{c}_{\widehat{\mathbf{h}}}^{(j)}\}_{j=1}^{J}$

denote the centroids of the reconstructed regions that contain a location at the earliest gated LAT, and $\{\mathbf{c}_{\mathbf{h}}^{(k)}\}$ the set of ground-truth centroids. The localization error is

$$\mathrm{LE} = \frac{1}{J}\sum_{j=1}^{J} \min_{k} \left\| \mathbf{c}_{\widehat{\mathbf{h}}}^{(j)} - \mathbf{c}_{\mathbf{h}}^{(k)} \right\|_2. \tag{34}$$

The one-to-nearest convention accounts for near-simultaneous fronts on the true heart surface, and averaging over the J earliest reconstructed regions makes LE independent of the order of tied regions.

Statistical testing. Data-driven methods are trained five times. DPS, DDRM, and OpDis-PS use the same five trained diffusion models, so their results pair run by run. For each metric and condition, we compare the best and second-best data-driven methods with a two-sided paired t-test when both use the same trained models and with a two-sided Welch t-test otherwise.

B Architectural details

We describe the Transformer and Palette architectures used for amortized inference (Sections B.1.1 and B.1.2), followed by the unconditional diffusion prior shared by all posterior-sampling methods (Section B.2.1). All data-driven architectures are implemented in PyTorch.

B.1 Amortized inference

B.1.1 Transformer

The Transformer baseline is an encoder-only self-attention model [42] applied along the time axis. Each of the 91 time points forms a token containing the 176 centered body-template potentials $\mathbf{b}$ (without the coordinate reduction of Appendix A.1). A linear layer projects each token to a hidden dimension of 384, and fixed sinusoidal positional encodings inject its temporal position.

The encoder uses PyTorch's `TransformerEncoder` with six pre-LayerNorm blocks (`norm_first=True`), each with eight attention heads, a feed-forward width of 1536, GELU activations, and a dropout probability of 0.1. The encoder processes the full sequence without a causal mask. A shared linear head maps each output token to the 1800 heart-template potentials $\mathbf{h}$ at the corresponding time point. The network has $11{,}408{,}520$ trainable parameters.

B.1.2 Palette

We employ a Palette-style conditional VP-DDPM [39] that maps body-template potentials to heart-template potentials, in line with the geometry-free conditional-diffusion approach to ECGI of Jara and Meyers [17]. It supplies a fixed condition alongside the noisy target at every denoising step, with a 3D U-Net backbone and physical boundary conditions.

At each time point, a shared linear layer maps the 176 centered body-template potentials to 1800 conditioning values, reshaped into a 30×60 grid aligned with the heart-template locations. This places the observation in the reconstruction's own coordinate system, and we train the projection jointly with the diffusion model.

The conditional denoiser shares the 3D U-Net architecture of the unconditional prior (Section B.2.1) with separately trained weights. Its six input channels stack the diffusion-corrupted heart-template potentials, the conditioning grid, and the four fixed coordinate channels. The complete model has $15{,}205{,}705$ trainable parameters, including $318{,}600$ in the conditioning projection.

B.2 Posterior sampling

B.2.1 Unconditional diffusion prior

The three posterior samplers, OpDis-PS, DDRM, and DPS, reconstruct with the same frozen unconditional prior and differ only in the likelihood. The prior operates on complete heart-template potentials on the fixed $91 \times 30 \times 60$ grid over time and heart-template locations. Its denoiser has five input channels containing the diffusion-corrupted heart-template potentials and the fixed coordinates

z_h, $\sin(\varphi_h)$, $\cos(\varphi_h)$, and normalized time. Its single output channel predicts the diffusion velocity $\mathbf{v}_\theta$ [40].

The denoiser is a 3D U-Net that jointly processes time and both template axes. The input tensor is padded to $96 \times 32 \times 60$ before the U-Net and cropped back afterward. The $3 \times 3 \times 3$ convolutions use replicated boundaries in time and along z_h and a circular boundary along φ_h.

The encoder levels and bottleneck have widths 32, 64, 128, and 256, with two residual blocks at each level. The three downsampling strides are $(2, 2, 2)$, $(2, 2, 2)$, and $(2, 2, 1)$. The decoder uses corresponding nearest-neighbor upsampling followed by convolutions and concatenates encoder features through skip connections. Each decoder level contains two residual blocks.

The residual blocks use GroupNorm with eight groups, SiLU activations, and a dropout probability of 0.1. A sinusoidal diffusion-step embedding followed by a two-layer MLP enters every residual block. The network contains no attention layers and has $14{,}886{,}241$ trainable parameters.

C Hyperparameters

We summarize the training and inference settings of all methods and describe the validation procedures used for hyperparameter selection. We then detail the estimation of the OpDis-PS discrepancy model (Section C.4).

C.1 Variational regularization

First-order Tikhonov. We compare spatial regularization orders zero, one, and two on nominal validation and first-order regularization achieves the lowest mean global relative error. We use $\mathbf{M}_\mathrm{G} = \mathbf{L}_\mathrm{G} + \gamma\mathbf{I}$, where $\mathbf{L}_\mathrm{G}$ is the unweighted four-neighbor graph Laplacian of the heart-template locations. We set $\gamma = 10^{-3}\lambda^+_{\min}(\mathbf{L}_\mathrm{G})$, with $\lambda^+_{\min}$ denoting the smallest positive eigenvalue. GCV selects one λ per case from 1000 logarithmically spaced values between 10^{-8} and 10^2 using all time points. We retain eigenmodes of $\mathbf{A}\mathbf{M}_\mathrm{G}^{-1}\mathbf{A}^\mathsf{T}$ with $w_i > 10^{-6}w_{\max}$.

Spatio-temporal Tikhonov. Spatio-temporal Tikhonov uses the same spatial penalty and effective-rank cutoff. It adds a second-order temporal-difference penalty and fixes $\rho = \lambda_\mathrm{t}^2/\lambda_\mathrm{s}^2 = 1$. Joint GCV selects the common scale from the same grid for each reconstruction.

Total Variation. We select one SNR-independent strength from $\{0.001, 0.003, 0.01, 0.03, 0.1, 0.3\}$ by mean global relative error across the nominal validation conditions. This yields $\alpha = 0.01$. The solver uses a quadratic gauge-fixing term with $\epsilon = 10^{-6}$, a relative-update tolerance of 10^{-5}, and at most $100{,}000$ iterations.

C.2 Amortized inference

Transformer. We train five independently initialized models on the 5100 nominal training pairs. Body-template potentials and heart-template potentials use separate scalar standard deviations estimated from the training set. During training, we augment each body-template input with probability 0.8 using centered Gaussian noise whose target SNR is drawn uniformly in dB from $[10, 40]$. We use AdamW with batch size 8, learning rate 10^{-4}, weight decay 10^{-4}, and gradient-norm clipping at 1. Training runs for at most 100 epochs and stops after 15 consecutive epochs without lower validation MSE. We select the checkpoint with the lowest MSE on the 100 nominal validation pairs. We fix the architecture of Appendix B.1.1 without a validation sweep.

Palette. We train five independently initialized models on the same training pairs and use the same normalization and augmentation protocol. Each model uses a 1000-step cosine variance-preserving diffusion process with the $\mathbf{v}$-parameterization [31, 40]. We train for 150 epochs with AdamW using batch size 4, learning rate 2×10^{-4}, weight decay 10^{-4}, gradient-norm clipping at 1, and EMA decay 0.9999. We select one checkpoint per training by fixed-seed $\mathbf{v}$-MSE on the same 100 nominal validation pairs and use its EMA weights at inference. Sampling uses 50 DDIM steps with $\eta = 0$, temperature 1, and base inference seed 0. We report the mean of 16 samples as the point estimate.

Table 3: **DDRM sampler selection.** Mean global relative error on 100 nominal validation cases at 40 dB using one trained diffusion model. Bold marks the best parameter combination.

η	Mean RE↓		
	$\eta_b = 0.5$	$\eta_b = 0.75$	$\eta_b = 1$
0.85	0.3719	0.3076	0.2688
1.0	0.3903	0.3088	**0.2609**

C.3 Posterior sampling

Unconditional diffusion prior. We train five independently initialized priors only on the 5100 physiological heart-template potentials, normalized by their global training-set standard deviation. Each prior uses a 1000-step cosine variance-preserving diffusion process with the $\mathbf{v}$-parameterization [31, 40]. We train for 54 epochs with AdamW using batch size 4, learning rate 2×10^{-4}, weight decay 10^{-4}, and EMA decay 0.9999. Unconditional prior evaluation on validation supports the use of $\mathbf{v}$-parameterization and EMA weights. Since the late EMA checkpoints perform similarly, all posterior-sampling methods use the checkpoint after the final training epoch.

DDRM and OpDis-PS. DDRM has two stochasticity hyperparameters: η_b controls how strongly the update moves towards the observation in singular directions where the iterate is noisier than the observation, and η controls the weak measurement guidance in the remaining directions. With $\eta_b = 1$, noiseless observations are enforced exactly in every observed coordinate. DDRM and OpDis-PS use the same sampler settings. On nominal validation at 40 dB with 100 cases and one of the five independently trained priors, we jointly compare $\eta \in \{0.85, 1\}$ and $\eta_b \in \{0.5, 0.75, 1\}$. Table 3 reports the mean global relative errors. The combination $\eta = \eta_b = 1$ achieves the lowest error among these six configurations. We use this setting across the 30- and 40-dB evaluation conditions and all operator conditions. Both methods use 20 reverse steps and report the mean of 16 samples as the point estimate.

DPS. We select the guidance strength ζ on nominal validation at 20 dB using 100 cases and one of the five independently trained priors. We compare $\zeta \in \{0, 10, 40, 60, 100, 150\}$ while fixing 50 DDIM steps, $\eta = 0$, temperature 1, and the point estimate to the mean of 16 samples. The corresponding mean global relative errors are 0.6925, 0.5248, 0.4820, 0.4921, 0.5169, and 0.5417. We therefore select $\zeta = 40$.

C.4 OpDis-PS discrepancy estimation

We fit a separate discrepancy model for each discrepancy source on 64 training heart–torso pairs, each combined with six recording operators $\mathbf{A}_0$ drawn from the corresponding perturbation distribution (Section 5.1), so that a single set of estimates $(\widehat{\boldsymbol{\mu}}_{\mathrm{s}}, \widehat{\boldsymbol{\mu}}_{\mathrm{d}}, \widehat{\mathbf{L}}, \widehat{\boldsymbol{\Sigma}}_{\mathrm{r}})$ per source is shared across all test geometries.

Signal-conditional discrepancy estimator. For each training time point in the $p = n_{\mathrm{B}} - 1 = 175$-dimensional average-referenced measurement space, we form $\mathbf{s}_t = \mathbf{A}\mathbf{h}_t$ and $\mathbf{d}_t = (\mathbf{A}_0 - \mathbf{A})\mathbf{h}_t$, yielding 34,944 pairs per source. We estimate the conditional-mean predictor as

$$\widehat{\mathbf{L}} = \widehat{\mathbf{C}}_{\mathrm{ds}} \left(\widehat{\mathbf{C}}_{\mathrm{ss}} + \lambda \mathbf{I} \right)^{-1}, \qquad \lambda = \kappa \frac{\mathrm{tr}(\widehat{\mathbf{C}}_{\mathrm{ss}})}{p}, \tag{35}$$

where $\widehat{\mathbf{C}}_{\mathrm{ss}}$ and $\widehat{\mathbf{C}}_{\mathrm{ds}}$ are the sample covariance blocks of $(\mathbf{s}_t, \mathbf{d}_t)$. Geometry-level 8-fold cross-validation over $\kappa \in \{0, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 10\}$ selects $\kappa = 10^{-3}$ for both sources by held-out coefficient of determination. The selected predictor achieves held-out R^2 values of 0.441 for lung conductivity and 0.937 for subcutaneous fat. We keep complete geometries within a single fold and never split individual time points across training and validation.

We recompute the residuals $\mathbf{r}_t = \mathbf{d}_t - \widehat{\boldsymbol{\mu}}_{\mathrm{d}} - \widehat{\mathbf{L}}(\mathbf{s}_t - \widehat{\boldsymbol{\mu}}_{\mathrm{s}})$ and estimate their covariance using Ledoit–Wolf shrinkage [24],

$$\widehat{\boldsymbol{\Sigma}}_{\mathrm{r}} = (1 - \rho)\widehat{\mathbf{C}}_{\mathrm{rr}} + \rho \frac{\mathrm{tr}(\widehat{\mathbf{C}}_{\mathrm{rr}})}{p} \mathbf{I}. \tag{36}$$

Table 4: **Nominal quantitative results (physiological data; 40 dB).** We evaluate seven baseline methods on the same 250 physiological test cases. For data-driven methods, values show the mean and standard deviation over five training runs.

	Method	Metric				
		RE ↓	ATE [ms] ↓	LE [mm] ↓	sCC ↑	tCC ↑
R	Tikhonov (spatial)	0.518	5.507	15.574	0.661	0.827
	Tikhonov (spatio-temporal)	0.506	4.715	8.434	0.687	0.835
	Total Variation	0.573	6.058	9.869	0.743	0.805
A	Transformer	0.392 ± 0.002	2.796 ± 0.048	4.561 ± 0.294	0.814 ± 0.009	0.906 ± 0.001
	Palette	0.377 ± 0.002	2.709 ± 0.021	5.035 ± 0.194	0.852 ± 0.003	0.912 ± 0.001
P	DPS	0.472 ± 0.004	4.366 ± 0.091	8.163 ± 0.369	0.801 ± 0.006	0.858 ± 0.003
	DDRM	$\mathbf{0.264 \pm 0.001}$	$\mathbf{1.687 \pm 0.007}$	$\mathbf{1.939 \pm 0.065}$	$\mathbf{0.868 \pm 0.002}$	$\mathbf{0.959 \pm 0.000}$

RE = relative error; ATE = activation time error; LE = localization error; sCC = spatial corr. coefficient; tCC = temporal corr. coefficient

The analytic Ledoit–Wolf rule determines ρ from the training residuals without an additional validation hyperparameter, yielding $\rho \approx 7.2 \times 10^{-4}$ for lung conductivity and 3.8×10^{-4} for subcutaneous fat.

Marginal discrepancy estimator. For the marginal ablations in Appendix D.4, we estimate the discrepancy mean $\widehat{\boldsymbol{\mu}}_{\mathrm{d}}$ and covariance $\widehat{\mathbf{C}}_{\mathrm{dd}}$ directly from the discrepancy samples, applying the same automatic Ledoit–Wolf shrinkage to the sample covariance.

D Additional quantitative results

We first compare the magnitudes of measurement noise and operator discrepancy (Section D.1). We then report a matched-operator reference (Section D.2) and additional 30-dB results (Section D.3), followed by ablations (Section D.4) and an analysis of inference cost (Section D.5).

D.1 Noise and operator-discrepancy magnitude

To compare the magnitude of operator discrepancy with measurement noise, we compute

$$r_{\mathrm{op}} = \frac{\|(\mathbf{A}_0 - \mathbf{A})\mathbf{h}\|_F}{\|\mathbf{A}\mathbf{h}\|_F} \tag{37}$$

before adding measurement noise. Across the 250 physiological and 720 pathological evaluation cases, the median relative discrepancy is 5.4% (interquartile range 2.4–8.1%) for lung conductivity and 15.2% (9.1–21.8%) for subcutaneous fat. Under the same norm convention, these values correspond to effective levels of 25.3 and 16.3 dB, respectively. For comparison, target SNRs of 40 and 30 dB correspond to relative noise norms of 1.0% and 3.2%. Thus, the 40-dB condition emphasizes operator discrepancy without it being masked by measurement noise.

D.2 Nominal reference

Table 4 reports the physiological results at 40 dB when the recording and reconstruction operators coincide, $\mathbf{A}_0 = \mathbf{A}$. We omit OpDis-PS because no operator discrepancy needs to be modeled. Comparing these results with Table 2 shows the degradation caused by operator discrepancy.

D.3 Additional results at 30 dB

Tables 5 and 6 report the corresponding evaluations at 30 dB, where measurement noise is more pronounced relative to the operator discrepancy than in the 40-dB condition.

On physiological activation patterns, OpDis-PS achieves the lowest RE and ATE and the highest tCC under both discrepancy sources. Palette attains the highest sCC. DDRM yields the lowest LE under lung-conductivity discrepancy and OpDis-PS under subcutaneous-fat discrepancy. The differences in LE and sCC between the respective best and second-best methods are not statistically significant.

Table 5: **Quantitative results (physiological data; 30 dB).** We evaluate all methods on 250 physiological test cases under lung-conductivity and subcutaneous-fat discrepancy. Data-driven results are mean $\pm$ standard deviation over five training runs; $\star$ denotes $p < 0.05$ for the best versus second-best data-driven method (Appendix A.4).

Discrepancy		Method	Metric				
			RE↓	ATE [ms]↓	LE [mm]↓	sCC↑	tCC↑
lung conductivity	R	Tikhonov (spatial)	0.701	15.968	39.565	0.569	0.752
		Tikhonov (spatio-temporal)	0.687	6.896	10.128	0.608	0.769
		Total Variation	0.584	6.315	10.373	0.675	0.795
	A	Transformer	0.395 ± 0.002	2.822 ± 0.046	4.362 ± 0.192	0.814 ± 0.009	0.905 ± 0.001
		Palette	0.380 ± 0.003	2.743 ± 0.028	5.360 ± 0.268	$\mathbf{0.851 \pm 0.002}$	0.910 ± 0.002
	P	DPS	0.483 ± 0.004	4.482 ± 0.088	8.695 ± 0.343	0.795 ± 0.005	0.852 ± 0.003
		DDRM	0.363 ± 0.001	2.486 ± 0.006	$\mathbf{2.940 \pm 0.118}$	0.841 ± 0.002	0.923 ± 0.000
		OpDis-PS (**ours**)	$\mathbf{0.347 \pm 0.001}^\star$	$\mathbf{2.426 \pm 0.010}^\star$	2.957 ± 0.170	0.849 ± 0.003	$\mathbf{0.927 \pm 0.000}^\star$
subcutaneous fat	R	Tikhonov (spatial)	0.697	14.876	40.930	0.568	0.756
		Tikhonov (spatio-temporal)	0.679	6.417	12.460	0.608	0.771
		Total Variation	0.602	6.328	9.460	0.668	0.794
	A	Transformer	0.399 ± 0.004	2.882 ± 0.056	4.564 ± 0.257	0.810 ± 0.008	0.903 ± 0.002
		Palette	0.376 ± 0.003	2.701 ± 0.025	4.984 ± 0.173	$\mathbf{0.854 \pm 0.002}$	0.912 ± 0.002
	P	DPS	0.502 ± 0.005	4.533 ± 0.094	7.578 ± 0.414	0.791 ± 0.005	0.850 ± 0.003
		DDRM	0.380 ± 0.002	2.464 ± 0.013	3.067 ± 0.123	0.843 ± 0.004	0.923 ± 0.001
		OpDis-PS (**ours**)	$\mathbf{0.341 \pm 0.001}^\star$	$\mathbf{2.362 \pm 0.011}^\star$	$\mathbf{3.035 \pm 0.188}$	0.850 ± 0.004	$\mathbf{0.931 \pm 0.000}^\star$

RE = relative error; ATE = activation time error; LE = localization error; sCC = spatial corr. coefficient; tCC = temporal corr. coefficient

Table 6: **Quantitative results (pathological data; 30 dB).** We evaluate all methods on 720 pathological test cases comprising LBBB, RBBB, and myocardial scar under lung-conductivity and subcutaneous-fat discrepancy. Data-driven results are mean $\pm$ standard deviation over five training runs; $\star$ denotes $p < 0.05$ for the best versus second-best data-driven method (Appendix A.4).

Discrepancy		Method	Metric				
			RE↓	ATE [ms]↓	LE [mm]↓	sCC↑	tCC↑
lung conductivity	R	Tikhonov (spatial)	0.695	24.121	46.282	0.627	0.738
		Tikhonov (spatio-temporal)	0.677	9.680	15.122	0.664	0.756
		Total Variation	0.618	8.573	11.379	0.694	0.762
	A	Transformer	0.561 ± 0.002	7.144 ± 0.079	7.331 ± 0.223	0.746 ± 0.005	0.792 ± 0.002
		Palette	0.517 ± 0.003	6.404 ± 0.218	7.988 ± 0.471	0.786 ± 0.001	0.821 ± 0.003
	P	DPS	0.618 ± 0.004	11.226 ± 0.179	12.245 ± 0.580	0.722 ± 0.003	0.738 ± 0.003
		DDRM	0.431 ± 0.003	$\mathbf{4.363 \pm 0.116}^\star$	$\mathbf{3.938 \pm 0.142}$	$\mathbf{0.837 \pm 0.002}^\star$	$\mathbf{0.884 \pm 0.003}^\star$
		OpDis-PS (**ours**)	$\mathbf{0.429 \pm 0.004}^\star$	4.506 ± 0.133	4.004 ± 0.107	0.834 ± 0.004	0.882 ± 0.003
subcutaneous fat	R	Tikhonov (spatial)	0.717	24.670	45.018	0.619	0.734
		Tikhonov (spatio-temporal)	0.693	9.421	15.365	0.658	0.753
		Total Variation	0.636	8.510	10.458	0.692	0.763
	A	Transformer	0.565 ± 0.002	7.108 ± 0.085	7.362 ± 0.269	0.746 ± 0.005	0.792 ± 0.002
		Palette	0.521 ± 0.003	6.504 ± 0.239	8.206 ± 0.693	0.784 ± 0.002	0.820 ± 0.003
	P	DPS	0.617 ± 0.004	10.664 ± 0.168	10.552 ± 0.335	0.728 ± 0.003	0.751 ± 0.003
		DDRM	0.444 ± 0.003	$\mathbf{4.083 \pm 0.113}^\star$	$\mathbf{3.791 \pm 0.145}$	$\mathbf{0.840 \pm 0.002}$	$\mathbf{0.889 \pm 0.002}^\star$
		OpDis-PS (**ours**)	$\mathbf{0.421 \pm 0.003}^\star$	4.274 ± 0.121	3.877 ± 0.105	0.837 ± 0.004	0.887 ± 0.003

RE = relative error; ATE = activation time error; LE = localization error; sCC = spatial corr. coefficient; tCC = temporal corr. coefficient

For pathological activation patterns, OpDis-PS retains the lowest RE under both discrepancy sources. DDRM performs best on ATE, LE, sCC, and tCC. The differences in LE under both discrepancy sources and in sCC under subcutaneous-fat discrepancy are not statistically significant.

At 30 dB, the method ordering is largely preserved, but the margins between methods shrink. This narrowing is expected. As measurement noise grows, it increasingly masks the operator discrepancy, so modeling the discrepancy contributes less and the methods converge. The effect is consistent with the noise and discrepancy magnitudes reported in Appendix D.1.

D.4 Ablation of the discrepancy model

We compare the signal-conditional likelihood of OpDis-PS with two marginal approximation-error variants. The first removes signal conditioning and models $\mathbf{d}_t \sim \mathcal{N}(\widehat{\boldsymbol{\mu}}_{\mathrm{d}}, \widehat{\mathbf{C}}_{\mathrm{dd}})$. We handle the resulting correlated-noise likelihood by whitening as in CARD [30]. The second replaces $\widehat{\mathbf{C}}_{\mathrm{dd}}$ with

Table 7: **Ablation (physiological data; 40 dB).** We evaluate the discrepancy-model variants on the 250 physiological test cases under lung-conductivity and subcutaneous-fat discrepancy. Values show the mean and standard deviation over five training runs. $\star$ denotes $p < 0.05$ for the best versus second-best variant (two-sided paired t-test).

Discrepancy	Method	RE↓	ATE [ms]↓	LE [mm]↓	sCC↑	tCC↑
lung conductivity	OpDis-PS (ours)	$\mathbf{0.313 \pm 0.001}^{\star}$	$\mathbf{2.033 \pm 0.008}^{\star}$	$\mathbf{2.198 \pm 0.146}^{\star}$	$\mathbf{0.858 \pm 0.003}^{\star}$	$\mathbf{0.943 \pm 0.000}^{\star}$
	OpDis-PS (marginal)	0.315 ± 0.001	2.074 ± 0.008	2.312 ± 0.094	0.857 ± 0.003	0.941 ± 0.000
	OpDis-PS (marginal, matched isotropic)	0.365 ± 0.001	2.554 ± 0.009	3.110 ± 0.231	0.818 ± 0.002	0.921 ± 0.001
subcutaneous fat	OpDis-PS (ours)	$\mathbf{0.321 \pm 0.001}^{\star}$	$\mathbf{2.109 \pm 0.005}^{\star}$	$\mathbf{2.470 \pm 0.090}^{\star}$	$\mathbf{0.855 \pm 0.004}^{\star}$	$\mathbf{0.940 \pm 0.000}^{\star}$
	OpDis-PS (marginal)	0.353 ± 0.002	2.429 ± 0.016	3.642 ± 0.097	0.852 ± 0.004	0.926 ± 0.001
	OpDis-PS (marginal, matched isotropic)	0.425 ± 0.002	3.270 ± 0.030	4.618 ± 0.115	0.790 ± 0.003	0.893 ± 0.001

Table 8: **Ablation (pathological data; 40 dB).** We evaluate the discrepancy-model variants on the 720 pathological test cases (LBBB, RBBB, and myocardial scar) under lung-conductivity and subcutaneous-fat discrepancy. Values show the mean and standard deviation over five training runs. $\star$ denotes $p < 0.05$ for the best versus second-best variant (two-sided paired t-test).

Discrepancy	Method	RE↓	ATE [ms]↓	LE [mm]↓	sCC↑	tCC↑
lung conductivity	OpDis-PS (ours)	$\mathbf{0.393 \pm 0.003}^{\star}$	$\mathbf{3.621 \pm 0.091}^{\star}$	$\mathbf{2.927 \pm 0.077}$	$\mathbf{0.853 \pm 0.003}^{\star}$	$\mathbf{0.903 \pm 0.002}^{\star}$
	OpDis-PS (marginal)	0.395 ± 0.003	3.716 ± 0.097	3.078 ± 0.088	0.850 ± 0.003	0.901 ± 0.002
	OpDis-PS (marginal, matched isotropic)	0.438 ± 0.004	4.610 ± 0.140	4.328 ± 0.126	0.819 ± 0.002	0.878 ± 0.003
subcutaneous fat	OpDis-PS (ours)	$\mathbf{0.395 \pm 0.003}^{\star}$	$\mathbf{3.602 \pm 0.091}^{\star}$	$\mathbf{3.160 \pm 0.086}^{\star}$	$\mathbf{0.853 \pm 0.003}^{\star}$	$\mathbf{0.903 \pm 0.002}^{\star}$
	OpDis-PS (marginal)	0.430 ± 0.004	4.196 ± 0.123	4.113 ± 0.069	0.833 ± 0.004	0.886 ± 0.003
	OpDis-PS (marginal, matched isotropic)	0.514 ± 0.007	6.750 ± 0.391	6.540 ± 0.223	0.771 ± 0.004	0.833 ± 0.006

the matched-trace isotropic covariance $\widehat{\sigma}^2 \mathbf{I}_{n_{\mathrm{B}}-1}$, where

$$\widehat{\sigma}^2 = \frac{\mathrm{tr}(\widehat{\mathbf{C}}_{\mathrm{dd}})}{n_{\mathrm{B}} - 1}, \tag{38}$$

while retaining $\widehat{\boldsymbol{\mu}}_{\mathrm{d}}$. The first comparison assesses the effect of signal conditioning. The second isolates the covariance structure at fixed total discrepancy variance.

Across both discrepancy sources, the automatic Ledoit–Wolf estimator achieves a mean geometry-held-out Gaussian log score comparable to inner-CV-selected fixed shrinkage and higher than that of the matched-trace isotropic covariance.

Tables 7 and 8 report these three variants under the same evaluation setup as the main comparison.

Signal conditioning yields small improvements for lung-conductivity discrepancy and clearer improvements for subcutaneous-fat discrepancy for both datasets. The structured marginal covariance outperforms its matched-isotropic counterpart across all metrics and both discrepancy types.

D.5 Inference cost

We measure reconstruction time on eight held-out lung-conductivity-discrepancy cases at 20 dB. We use one trained instance per learned model and run all methods sequentially, processing one case at a time. Learned methods use an NVIDIA A100 GPU, while variational methods use 16 CPU threads. Except for Total Variation, whose number of iterations depends on its convergence tolerance, all methods perform a fixed amount of computation per case, so their reconstruction times do not depend on the noise level.

After one excluded warm-up call per method, we reconstruct each case three times. We time complete reconstruction calls with the models already loaded and synchronize CUDA before and after each GPU call. For sampling-based methods, each call includes all 16 draws and their averaging. Table 9 reports the median over the resulting 24 measurements.

E Simulation details

We detail the geometries, electrophysiological simulations, and pathological configurations summarized in Section 5.1.

Table 9: **Inference cost.** Median reconstruction time over three repetitions of eight cases.

Method	Device	Time
Tikhonov (spatial)	CPU	20.12 ms
Tikhonov (spatio-temporal)	CPU	2.14 s
Total Variation	CPU	53.91 s
Transformer	GPU	1.79 ms
Palette	GPU	21.54 s
DPS	GPU	44.62 s
DDRM	GPU	6.03 s
OpDis-PS (ours)	GPU	6.04 s

E.1 Geometries

Hearts. For each ventricular mesh, we compute fiber directions with a rule-based method [4], using helix angles of 80° and -60° at the endocardium and epicardium. Ventricular coordinates [3] provide an apico-basal coordinate $Z \in [0, 1]$, a transmural coordinate $\rho \in [0, 1]$ from endocardium to epicardium, a rotational coordinate φ, and a ventricle indicator V (-1 for the left and $+1$ for the right ventricle). We use them to place stimuli, assign action-potential durations, and define pathological regions.

Torsos and electrodes. We register each heart to the ventricular envelope of the torso CT with a similarity iterative-closest-point fit and accept a pairing only if the epicardium lies inside the torso and does not intersect the lungs. Electrodes lie on a 10×20 grid in normalized torso height and angle around the torso centerline, excluding the arm regions. The resulting 176 positions are the same for every torso, so electrode indices correspond across subjects.

E.2 Electrophysiology

We solve the pseudo-bidomain equations in openCARP [34] with the Mitchell–Schaeffer ionic model [29]. Myocardial conductivities are transversely isotropic with respect to the fiber direction and tuned with openCARP's tuneCV utility on a one-dimensional cable at the mean mesh edge length of 1.13 mm, targeting conduction velocities of 0.6 and 0.3 m/s along and across the fibers [14]. A subendocardial layer ($\rho < 0.1$, $0.1 \leq Z \leq 0.9$) with an isotropic conduction velocity of 2 m/s represents the fast-conducting Purkinje system [14].

Action-potential durations. Following Gillette et al. [14], we map the normalized combination $q_Z Z + q_\rho \rho + q_V V$ affinely to $[\mathrm{APD}_{\min}, \mathrm{APD}_{\max}]$, with $q_Z \sim \mathcal{U}[0.9, 1]$, $q_\rho \sim \mathcal{U}[-0.6, 0]$, $q_V \sim \mathcal{U}[0.1, 0.15]$, $\mathrm{APD}_{\min} \sim \mathcal{U}[150, 175]$ ms, and $\mathrm{APD}_{\max} \sim \mathcal{U}[225, 250]$ ms. This yields apico-basal, transmural, and interventricular repolarization gradients.

Activation. Five endocardial sites represent the fascicles: anterior, posterior, and septal sites in the left ventricle, and septal and moderator-band sites in the right ventricle. Their positions and radii are sampled from fixed ranges in ventricular coordinates.

Physiological rounds. A physiological round is one draw of the action-potential and activation parameters per heart–torso pair. Training pairs are simulated in three rounds, and validation and test pairs in one.

E.3 Pathologies

Table 10 lists the twelve pathological configurations. Within each pathology, all four configurations use the same 60 held-out test pairs, each with its own draw of the physiological parameters.

Bundle branch blocks. Following Gillette et al. [14], LBBB removes the three left-ventricular stimulus sites and RBBB the two right-ventricular sites, so the blocked ventricle is activated through its fast endocardial layer and across the septum. The more severe configurations slow this layer to 1.0 or 0.8 m/s or remove it entirely.

Table 10: **Pathological configurations.** Severity configurations of the held-out pathologies, each simulated on 60 test pairs. For bundle branch blocks, CV denotes the conduction velocity of the fast endocardial layer of the blocked ventricle; in the complete block, this layer conducts like ordinary myocardium. For scar, r and $\rho_{\max}$ denote the radius and transmural depth of the inert region.

Severity	LBBB	RBBB	Scar	
	CV [m/s]	CV [m/s]	r [mm]	$\rho_{\max}$
1	2.0	2.0	10	0.4
2	1.0	1.0	15	0.7
3	0.8	0.8	20	1.0
4	complete	complete	22	1.0

Myocardial scar. The scar configurations keep physiological activation and add an electrically inert region, around which the activation front is routed. The region is anchored in the mid-to-basal anterolateral left-ventricular wall, and its size and transmural extent grow together from a subendocardial to a large transmural scar.