PHYSICS-EMBEDDED FLOW MATCHING: GOVERNING THE LOCAL VECTOR FIELD WITH A KNOWN PHYSIOLOGICAL RELATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Physiological signal translation often suffers from limited paired data, motivating the use of governing relationships between source and target measurements as additional supervision. Existing physics-informed models constrain supervised predictions, while their flow matching counterparts constrain reconstructed states, generated endpoints, or generic properties of the transport field. We introduce a physics-informed flow matching framework that instead imposes the physiological relationship directly on the learned local velocity field through a physical constraint. We derive a closed-form velocity-space residual that softly enforces physical consistency at each flow time, and we present a unified error analysis characterizing how physics regularization and multi-step integration can reduce prediction error. Using a learned two-element Windkessel model, we evaluate translation from virtual aortic flow and from measured bioimpedance to pressure against supervised and flow matching baselines with and without physics constraints. Our Physics Embedded Flow Matching (PEFM) method achieves the lowest waveform RMSE on both datasets. On measured data, it achieves 11.8% lower RMSE than the baseline closest in physical consistency and 63.6% lower physics residual than the baseline closest in RMSE. These results support the use of governing relationships to jointly improve predictive accuracy and physical consistency.

1 INTRODUCTION

Physiological signal translation infers a target measurement from a source signal, which is valuable when the target is clinically informative but hard to obtain continuously or non-invasively. Deep learning models can learn accurate source-to-target mappings (Shome et al., 2024; Suzuki et al., 2026; Fang et al., 2025), but depend on the availability and diversity of paired training data (Akbari et al., 2021; Li et al., 2025). Collecting such data often requires synchronized measurements, specialized equipment, or invasive procedures, limiting dataset scale and coverage. Relying solely on these pairs risks overfitting, poor generalization, and predictions that violate known physiological relationships. Physiological signals, however, are observations of shared dynamical systems. Cardiovascular quantities, for example, may be coupled through reduced-order models, circuit analogues, conservation laws, or differential equations that approximately relate a measured source $\mathbf{x}_0$ and target $\mathbf{x}_1$ (Osman et al., 2026; Kissas et al., 2020; Sel et al., 2023; Jara et al., 2025). These relationships offer supervision beyond paired observations.

Physics-informed neural networks (PINNs) augment supervised learning by penalizing violations of governing equations (Kissas et al., 2020; Sel et al., 2023; Jara et al., 2025; Osman et al., 2026; LIU et al., 2026). In signal translation, these constraints typically act on the predicted target waveform, encouraging consistency with a known source-target relationship. For example, a Windkessel model linking flow and pressure can regularize the predicted pressure (Osman et al., 2026). This discourages physically implausible predictions. However, the constraint acts primarily on the predicted endpoint, while the intermediate transformation of the learned map remains implicit. This motivates incorporating the governing relationship directly into the learned transformation, beyond constraining the final prediction.

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

Flow matching (FM) offers a complementary formulation, representing source-to-target prediction as continuous transport governed by a time-dependent vector field (Lipman et al., 2023; Liu et al., 2023b; Albergo et al., 2023). Paired formulations learn transport between observed source and target modalities rather than a direct endpoint predictor (Albergo et al., 2024; Tong et al., 2024; Delbracio & Milanfar, 2023; Liu et al., 2023a). This dynamical representation exposes intermediate states along the translation path and the local velocity governing their transformation. Recent physics-informed generative methods incorporate physics through reconstructed states, generated endpoints, inference-time guidance, or physics-conditioned transport dynamics (Shu et al., 2023; Huang et al., 2024; Jacobsen et al., 2025; Bastek et al., 2025; Baldan et al., 2026; Tauberschmidt et al., 2026; Song & Yu, 2026). However, existing approaches do not generally derive the local transport-velocity constraint implied by a known source-target governing relationship at intermediate flow times (Table 1).

Table 1: **Comparison of method families.** Physics is imposed differently across model families and whether each method uses a paired source-to-target governing law.

Family	Representative	Physics acts on	Paired law?
Supervised	—	Nowhere	—
Flow Matching	PENGUIN	Nowhere	—
PINN	WPINN	Predicted target waveform	Yes
FM+Endpoint PINN	PIDM, PBFM	Reconstructed or generated target state	Yes
Architectural vector-field FM	PIFM-GN	Network class of $\boldsymbol{v}_\theta$	No (generic)
Proposed	**PEFM**	**Induced governing residual on $\boldsymbol{v}_\theta$**	**Yes**

We address the missing source-target constraint with Physics Embedded Flow Matching (PEFM), a physics-informed source-to-target FM framework that converts a governing relationship between paired endpoints into velocity-field supervision. PEFM applies the source-target law directly throughout transport, rather than on a predicted state (Figure 1).These constraints govern both whether a prediction is physically admissible and how one measured physiological state transforms into another. Our contributions are: (I) a theoretical derivation of the local velocity-space constraint induced by a governing relationship between paired source and target endpoints, (II) PEFM, which applies this induced constraint at intermediate flow times, together with an analysis of how physical regularization and multi-step transport affect endpoint consistency and prediction error, and (III) an empirical comparison against direct, FM, and physics-informed baselines on synthetic and measured physiological tasks. By distributing mechanistic supervision across flow time, PEFM extracts additional structure from each paired physiological observation, which is particularly valuable when paired measurements are scarce.

2 Related work

Physics-informed signal translation. Physiological signal translation learns mappings between measurement modalities to recover signals that may be difficult or invasive to acquire (Klein et al., 2025; Shome et al., 2024; Qiu et al., 2026; Wang et al., 2026; Fang et al., 2025; Suzuki et al., 2026). Limited paired data and generalization remain central challenges in these settings (Akbari et al., 2021; Li et al., 2025). Physics-informed physiological modeling complements data-driven translation by incorporating mechanistic relationships into learning. Existing approaches impose vascular conservation laws on estimated states (Kissas et al., 2020; Jara et al., 2025), introduce approximate physiological relationships for cuffless blood pressure estimation (Sel et al., 2023), or regularize predicted pressure waveforms using Windkessel dynamics (Osman et al., 2026). Such constraints provide supervision beyond paired observations and encourage consistency between estimated quantities and an assumed physiological model. However, physics in these approaches primarily constrains estimated physiological states or predicted target waveforms, while the transformation from the source measurement to the target remains implicit. Our work builds on this mechanistic supervision by deriving the constraint that a paired source-to-target relationship induces on the local velocity of a learned translation process.

Physics-informed flow and diffusion models. FM, rectified flows, and stochastic interpolants provide foundations for learning continuous transport between distributions (Lipman et al., 2023;

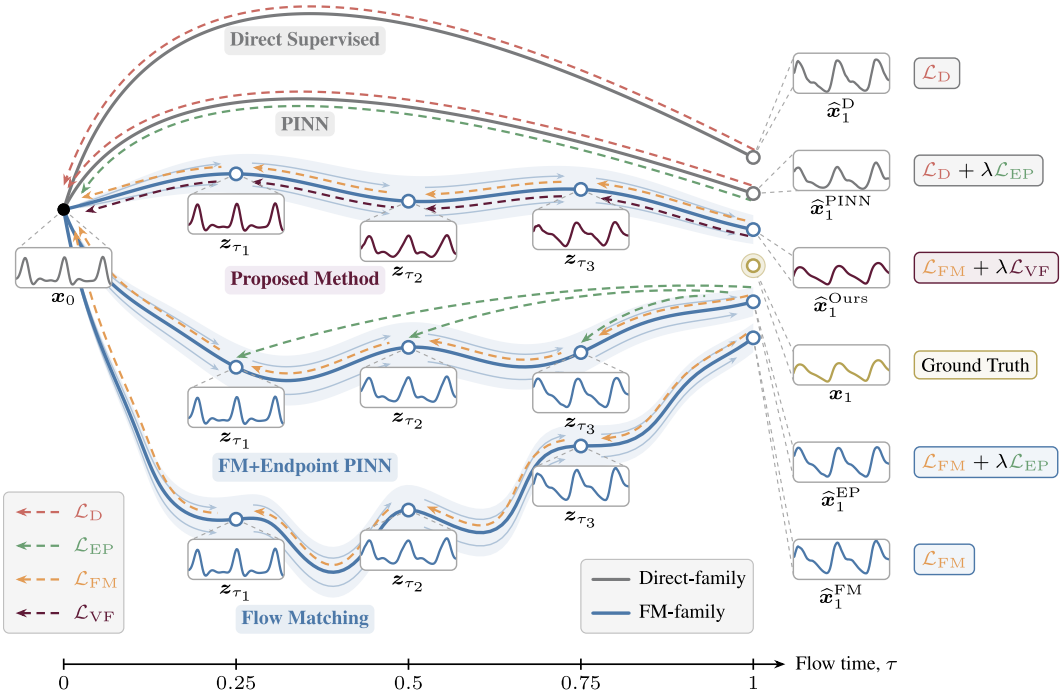


Figure 1: **Overview of translation methods.** PEFM constrains the FM velocity field throughout source-to-target transport. Each method maps source $\boldsymbol{x}_0$ to target $\boldsymbol{x}_1$. Direct Supervised uses $\mathcal{L}_{\mathrm{D}}$, PINN adds the endpoint-physics loss $\lambda\mathcal{L}_{\mathrm{EP}}$, Flow Matching transports intermediate states $\boldsymbol{z}_\tau$ using $\mathcal{L}_{\mathrm{FM}}$, FM+Endpoint PINN adds $\lambda\mathcal{L}_{\mathrm{EP}}$, and PEFM applies the velocity-field physics loss $\lambda\mathcal{L}_{\mathrm{VF}}$ along the path. τ is flow time, λ is the physics-loss weight, and hats denote predicted endpoints.

Liu et al., 2023b; Albergo et al., 2023; Tong et al., 2024). Related developments in data-dependent and multisample couplings, paired restoration, and one-step generative modeling further connect these formulations to translation between corresponding observations (Albergo et al., 2024; Pooladian et al., 2023; Delbracio & Milanfar, 2023; Heitz et al., 2023; Liu et al., 2023a; Zhou et al., 2024; Frans et al., 2025; Song et al., 2023). Physics-informed generative models extend these approaches by incorporating mechanistic constraints into diffusion and flow-based models, including physiological synthesis frameworks such as Cardio-mmFlow (LIU et al., 2026). Existing approaches impose physics on reconstructed or denoised states (Shu et al., 2023; Bastek et al., 2025), through training objectives (Jacobsen et al., 2025), or during sampling (Huang et al., 2024). Among flow-based approaches, physics has been incorporated through integrated flow states (Baldan et al., 2026), post-training fine-tuning (Tauberschmidt et al., 2026), or architectural constraints on the vector field (Song & Yu, 2026). These methods demonstrate several ways to combine physics and generative transport, but do not derive the local velocity constraint induced by a governing relationship between paired source and target measurements. PEFM starts from this relationship and derives a corresponding residual on the transport velocity, directly incorporating source-to-target physics into paired FM training.

3 Physics Embedded Flow Matching

3.1 Problem formulation as paired flow matching

Let $\boldsymbol{x}_0(t)$ and $\boldsymbol{x}_1(t)$ denote paired source and target signals over physical time t, drawn from a coupling π. Flow time is denoted by $\tau \in [0, 1]$. Following conditional FM and related interpolant-based formulations (Liu et al., 2023b; Albergo et al., 2023; Lipman et al., 2023; Tong et al., 2024; Albergo et al., 2024), we consider a differentiable conditional path

$$\boldsymbol{z}_\tau = \Gamma_\tau(\boldsymbol{x}_0, \boldsymbol{x}_1), \qquad \Gamma_0(\boldsymbol{x}_0, \boldsymbol{x}_1) = \boldsymbol{x}_0, \qquad \Gamma_1(\boldsymbol{x}_0, \boldsymbol{x}_1) = \boldsymbol{x}_1, \qquad \boldsymbol{v}_\tau^\star = \partial_\tau \Gamma_\tau(\boldsymbol{x}_0, \boldsymbol{x}_1), \quad (1)$$

where Γ_τ is the interpolant, $\boldsymbol{z}_\tau$ the interpolated state at flow time τ, and $\boldsymbol{v}_\tau^\star$ the oracle velocity. The vector field $\boldsymbol{v}_\theta$ is trained with

$$\mathcal{L}_{\mathrm{FM}} = \mathbb{E}_{(\boldsymbol{x}_0,\boldsymbol{x}_1)\sim\pi,\ \tau}\left[\|\boldsymbol{v}_\theta(\boldsymbol{z}_\tau,\tau) - \boldsymbol{v}_\tau^\star\|_2^2\right]. \tag{2}$$

The paired construction directly transports each observed source $\boldsymbol{x}_0$ toward its corresponding target $\boldsymbol{x}_1$. At inference, the learned vector field defines

$$\frac{\mathrm{d}\widehat{\boldsymbol{z}}_\tau}{\mathrm{d}\tau} = \boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau,\tau), \qquad \widehat{\boldsymbol{z}}_0 = \boldsymbol{x}_0, \tag{3}$$

and the terminal state $\widehat{\boldsymbol{z}}_1$ is the predicted target. We use this paired formulation to place data supervision and the governing physical law on the same learned velocity field.

3.2 Physics constraints in velocity space

Assume that each paired observation satisfies an approximate physical relation

$$\mathcal{F}(\boldsymbol{x}_1,\boldsymbol{x}_0) \approx 0, \tag{4}$$

where $\boldsymbol{x}_0$ is the observed conditioning signal and $\boldsymbol{x}_1$ is the target signal governed by, or coupled to $\boldsymbol{x}_0$ through, $\mathcal{F}$. For example, $\mathcal{F}$ may represent a differential equation relating an input waveform to a target physiological waveform.

Now, consider paths for which the target is locally recoverable from the path state, oracle velocity, and conditioning signal. For these paths, there exists a map H_τ such that

$$\boldsymbol{x}_1 = H_\tau\big(\boldsymbol{z}_\tau, \boldsymbol{v}_\tau^\star, \boldsymbol{x}_0\big). \tag{5}$$

This condition is satisfied by broad classes of endpoint interpolants used in FM, rectified-flow, stochastic-interpolant, and paired bridge constructions, including the affine families considered below (Liu et al., 2023b; Albergo et al., 2023; Lipman et al., 2023; Heitz et al., 2023; Liu et al., 2023a). Substituting Eq. 5 into Eq. 4 yields the corresponding oracle velocity relation

$$\mathcal{F}\left(H_\tau(\boldsymbol{z}_\tau, \boldsymbol{v}_\tau^\star, \boldsymbol{x}_0), \boldsymbol{x}_0\right) \approx 0. \tag{6}$$

This gives a physics constraint directly in the representation learned by FM.

Definition 1 (Velocity-space physics residual)**.** *For a locally recoverable path, define*

$$\boldsymbol{r}_H(\tau) = \mathcal{F}\left(H_\tau\big(\boldsymbol{z}_\tau, \boldsymbol{v}_\theta(\boldsymbol{z}_\tau,\tau), \boldsymbol{x}_0\big), \boldsymbol{x}_0\right). \tag{7}$$

The physics-constrained FM objective is

$$\mathcal{L} = \mathcal{L}_{\mathrm{FM}} + \lambda\mathcal{L}_{\mathrm{VF}}, \qquad \mathcal{L}_{\mathrm{VF}} = \mathbb{E}_{(\boldsymbol{x}_0,\boldsymbol{x}_1)\sim\pi,\ \tau}\left[\|\boldsymbol{r}_H(\tau)\|_2^2\right]. \tag{8}$$

Importantly, the recovery map H_τ serves only as a general derivational link from endpoint physics to a local velocity constraint. In other words, the resulting physics residual is evaluated directly at the same $(\boldsymbol{z}_\tau,\tau)$ samples used by the FM loss and therefore requires neither explicit endpoint reconstruction nor trajectory rollout during training.

3.3 Affine endpoint interpolants under linear physics

We now specialize the physical relation while retaining a general affine family of source-to-target paths. Suppose $\mathcal{A}\boldsymbol{x}_1 = \mathcal{B}\boldsymbol{x}_0$, where $\mathcal{A}$ and $\mathcal{B}$ are linear differential operators acting along physical time t. Consider an affine endpoint interpolant

$$\boldsymbol{z}_\tau = a(\tau)\boldsymbol{x}_0 + b(\tau)\boldsymbol{x}_1, \qquad \boldsymbol{z}_0 = \boldsymbol{x}_0, \quad \boldsymbol{z}_1 = \boldsymbol{x}_1, \tag{9}$$

with $a, b \in C^1([0,1])$. Wherever $b'(\tau) \neq 0$, the endpoint is recoverable from the oracle velocity, so the physical law can be rewritten directly in velocity space. Proposition 3 shows that the resulting velocity residual is

$$\boldsymbol{r}_v(\tau) = \mathcal{A}\boldsymbol{v}_\theta(\boldsymbol{z}_\tau,\tau) - \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0, \tag{10}$$

which vanishes for the oracle field and satisfies $\boldsymbol{r}_v(\tau) = b'(\tau)\,\boldsymbol{r}_H(\tau)$, where $\boldsymbol{r}_H$ is the endpoint residual evaluated at the endpoint implied by the local velocity prediction. The proposed local physics penalty therefore constrains the learned vector field through the physical law satisfied by its implied endpoint, without explicitly reconstructing that endpoint. We use $\mathcal{L}_{\mathrm{VF}} = \mathbb{E}[\|\boldsymbol{r}_v(\tau)\|^2]$, which weights the corresponding squared endpoint residual by $b'(\tau)^2$.

For the straight path $a(\tau) = 1-\tau, b(\tau) = \tau$, the velocity-implied endpoint is $\widehat{\boldsymbol{x}}_1^{(v)} = \boldsymbol{x}_0 + \boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau)$, and the velocity residual becomes

$$\boldsymbol{r}_v(\tau) = \mathcal{A}\widehat{\boldsymbol{x}}_1^{(v)} - \mathcal{B}\boldsymbol{x}_0. \tag{11}$$

These results establish that, for straight interpolation, the original physical law at the velocity-implied endpoint is exactly represented by a local velocity residual, allowing us to impose the physics directly on the learned vector field during training. The results in this subsection follow directly from Proposition 3 and its proof.

3.4 Trajectory-wide control of endpoint physics

The preceding derivation in Eq. 11 connects the physical law to a local velocity prediction. Since inference integrates the learned field to obtain the target, we now examine how these local constraints translate into physical consistency of the final prediction. For the full affine family in Eq. 9, we show that the endpoint physical residual is the integral of the velocity residual along the generated trajectory.

Proposition 2 (Trajectory residual controls endpoint residual)**.** *Let $\widehat{\boldsymbol{z}}_\tau$ satisfy*

$$\frac{\mathrm{d}\widehat{\boldsymbol{z}}_\tau}{\mathrm{d}\tau} = \boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau), \qquad \widehat{\boldsymbol{z}}_0 = \boldsymbol{x}_0, \tag{12}$$

and define the residual along the generated trajectory by

$$\boldsymbol{r}_v^{\mathrm{traj}}(\tau) = \mathcal{A}\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - [a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}]\,\boldsymbol{x}_0. \tag{13}$$

Then

$$\mathcal{A}\widehat{\boldsymbol{z}}_1 - \mathcal{B}\boldsymbol{x}_0 = \int_0^1 \boldsymbol{r}_v^{\mathrm{traj}}(\tau)\,\mathrm{d}\tau, \tag{14}$$

$$\|\mathcal{A}\widehat{\boldsymbol{z}}_1 - \mathcal{B}\boldsymbol{x}_0\|^2 \le \int_0^1 \left\|\boldsymbol{r}_v^{\mathrm{traj}}(\tau)\right\|^2 \mathrm{d}\tau. \tag{15}$$

Proposition 2 shows that the proposed constraint on the velocity imposes a stricter requirement than endpoint regularization. It requires the induced physical law to hold throughout the transport, which in turn controls the physical residual of the generated endpoint through Eq. 15.

For scalar-displacement paths, let $\boldsymbol{g} = (\mathcal{B} - \mathcal{A})\boldsymbol{x}_0$ denote the endpoint-law displacement, and write $\boldsymbol{u}_\theta(\boldsymbol{z}_\tau, \tau) = \boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau)/s'(\tau)$. Then the interpolation schedule determines the relative weighting of the physical constraint across flow time:

$$\mathcal{L}_{\mathrm{VF}} = \mathbb{E}_{(\boldsymbol{x}_0, \boldsymbol{x}_1)\sim\pi,\ \tau}\left[s'(\tau)^2\,\|\mathcal{A}\boldsymbol{u}_\theta(\boldsymbol{z}_\tau, \tau) - \boldsymbol{g}\|^2\right]. \tag{16}$$

Hence $s'(\tau)^2$ is the physics weighting induced by the interpolation schedule and the straight path yields uniform weighting. Proofs of Proposition 2, and the scalar-path specialization are given in Appendix A.

3.5 Error analysis under straight-path integration

The preceding results characterize physical consistency for a general affine interpolation. To analyze endpoint prediction error and compare PEFM with direct supervision, we specialize to the straight path

$$\boldsymbol{z}_\tau = \boldsymbol{x}_0 + \tau(\boldsymbol{x}_1 - \boldsymbol{x}_0), \qquad \boldsymbol{d} := \boldsymbol{x}_1 - \boldsymbol{x}_0, \qquad \mathcal{A}\boldsymbol{d} = \boldsymbol{g}, \tag{17}$$

where $\boldsymbol{d}$ is the oracle velocity. Let $\mathcal{S}(K, \lambda)$ denote the analytical configuration obtained using K uniform Euler evaluations and physics weight λ, with endpoint error $\boldsymbol{e}_{K,\lambda} := \widehat{\boldsymbol{z}}_K - \boldsymbol{x}_1$. Within

this notation, $\mathcal{S}(1,0)$ corresponds to residual direct supervision, $\mathcal{S}(K,0)$ to FM, $\mathcal{S}(1,\lambda)$ to direct supervision with the same physical penalty, and $\mathcal{S}(K,\lambda)$ to the proposed physics-constrained flow.

At flow step $k = 0,\dots,K-1$, with $\tau_k = k/K$, let $\boldsymbol{d} + \boldsymbol{\xi}_k$ denote the data-only FM velocity prediction, where $\boldsymbol{\xi}_k$ is its pre-physics local velocity error relative to $\boldsymbol{d}$. We analyze the effect of the quadratic physics penalty through

$$\boldsymbol{S}_\lambda = (\boldsymbol{I} + \lambda \mathcal{A}^*\mathcal{A})^{-1}, \qquad \widetilde{\boldsymbol{\xi}}_k := \boldsymbol{S}_\lambda \boldsymbol{\xi}_k, \tag{18}$$

where $\widetilde{\boldsymbol{\xi}}_k$ denotes the corresponding analytically regularized local error. If $\mathcal{A} = \boldsymbol{U}\,\mathrm{diag}(\sigma_i)\boldsymbol{V}^*$, then $\boldsymbol{S}_\lambda$ scales the i-th right-singular component of the error by

$$\frac{1}{1+\lambda\sigma_i^2}. \tag{19}$$

Thus, physics suppresses error along directions identifiable by the governing operator, with stronger contraction for larger σ_i, while leaving $\ker \mathcal{A}$ unchanged. Under uniform Euler integration, the endpoint error is

$$\boldsymbol{e}_{K,\lambda} := \widehat{\boldsymbol{z}}_K - \boldsymbol{x}_1 = \frac{1}{K}\sum_{k=0}^{K-1} \widetilde{\boldsymbol{\xi}}_k = \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{S}_\lambda \boldsymbol{\xi}_k. \tag{20}$$

This representation separates two distinct mechanisms in PEFM. Eq. 19 acts across operator directions through $\boldsymbol{S}_\lambda$, whereas Eq. 20 acts across τ by averaging local velocity errors. Eq. 18 suppresses error visible to the governing law and Eq. 20 allows errors that vary along the flow to partially cancel.

To quantify the resulting variance reduction across flow time, suppose the pre-physics local errors have common mean and covariance, and define their post-physics counterpart,

$$\boldsymbol{\mu}_F := \mathbb{E}[\boldsymbol{\xi}_k], \qquad \boldsymbol{\Sigma}_F := \mathrm{Cov}(\boldsymbol{\xi}_k), \qquad \boldsymbol{\Sigma}_{F,\lambda} := \boldsymbol{S}_\lambda \boldsymbol{\Sigma}_F \boldsymbol{S}_\lambda^*. \tag{21}$$

Assume that the post-physics local errors satisfy

$$\mathrm{tr}\,\mathrm{Cov}(\widetilde{\boldsymbol{\xi}}_i, \widetilde{\boldsymbol{\xi}}_j) = \rho_\lambda\,\mathrm{tr}\,\boldsymbol{\Sigma}_{F,\lambda}, \qquad i \neq j. \tag{22}$$

Then Theorem 5 gives

$$\mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] = \|\boldsymbol{S}_\lambda \boldsymbol{\mu}_F\|^2 + \left[\rho_\lambda + \frac{1-\rho_\lambda}{K}\right]\mathrm{tr}\,\boldsymbol{\Sigma}_{F,\lambda}. \tag{23}$$

Eq. 23 decomposes endpoint error into a squared mean term and a variance term. Under fixed marginal local statistics, integration does not reduce the mean error. Its benefit arises from the variance factor

$$\alpha_K(\rho_\lambda) = \rho_\lambda + \frac{1-\rho_\lambda}{K}. \tag{24}$$

When $\rho_\lambda = 1$, local errors are perfectly coherent across flow time and integration provides no variance reduction. When $\rho_\lambda < 1$, the endpoint variance is smaller than the marginal post-physics local variance because nonpersistent components partially cancel during integration. The relevant property of the flow is therefore not only the accuracy of each local velocity prediction, but also the dependence structure of its errors across flow time.

This observation provides a direct comparison with residual direct supervision. Let

$$\boldsymbol{e}_{D,0} := \widehat{\boldsymbol{x}}_{D,0} - \boldsymbol{x}_1, \qquad \boldsymbol{\mu}_D := \mathbb{E}[\boldsymbol{e}_{D,0}], \qquad \boldsymbol{\Sigma}_D := \mathrm{Cov}(\boldsymbol{e}_{D,0}) \tag{25}$$

denote the data-only direct model's endpoint error statistics. Theorem 6 shows that, without physics, K-step FM has lower expected squared endpoint error precisely when

$$\|\boldsymbol{\mu}_F\|^2 + \left[\rho + \frac{1-\rho}{K}\right]\mathrm{tr}\,\boldsymbol{\Sigma}_F < \|\boldsymbol{\mu}_D\|^2 + \mathrm{tr}\,\boldsymbol{\Sigma}_D, \qquad \rho := \rho_0. \tag{26}$$

Under matched marginal statistics, $\boldsymbol{\mu}_F = \boldsymbol{\mu}_D$ and $\boldsymbol{\Sigma}_F = \boldsymbol{\Sigma}_D$, the same result reduces to

$$\mathbb{E}[\|\boldsymbol{e}_{D,0}\|^2] - \mathbb{E}[\|\boldsymbol{e}_{K,0}\|^2] = \frac{K-1}{K}(1-\rho)\,\mathrm{tr}\,\boldsymbol{\Sigma}_F. \tag{27}$$

Hence, for $K > 1$ and nonzero marginal variance, FM has a strict advantage under matched statistics exactly when $\rho < 1$. More generally, when the means are matched,

$$\frac{\operatorname{tr} \boldsymbol{\Sigma}_F}{\operatorname{tr} \boldsymbol{\Sigma}_D} < \frac{1}{\rho + (1-\rho)/K} \tag{28}$$

characterizes when FM remains preferable, provided the denominator is positive. The proof and the corresponding edge cases are given in Theorem 6 of Appendix A.

This result clarifies why a flow formulation can improve upon direct prediction. FM does not require each local velocity estimate to have lower marginal error than a direct endpoint estimate. It can tolerate greater local variance and still produce a more accurate endpoint when those errors are sufficiently weakly correlated across flow time. The potential advantage of FM therefore arises from error cancellation along the integration path rather than from an assumption that every intermediate prediction is individually superior.

Physics regularization contributes a separate guarantee. For the same prescribed pre-physics local error sequence, Theorem 7 establishes

$$\boldsymbol{e}_{K,\lambda} = \boldsymbol{S}_\lambda \boldsymbol{e}_{K,0}, \qquad \mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] \le \mathbb{E}[\|\boldsymbol{e}_{K,0}\|^2]. \tag{29}$$

For $\lambda > 0$, the inequality is strict whenever $\boldsymbol{e}_{K,0}$ has nonzero energy outside $\ker \mathcal{A}$ with positive probability. Under an exact physical law and a fixed pre-physics error sequence, the analytical physics constraint therefore cannot increase endpoint error: it contracts precisely the component of the integrated error that is identifiable through the governing operator. The full argument is given in Appendix A, Theorem 7.

Finally, consider residual direct supervision with the same analytical physics regularization. Define

$$\boldsymbol{e}_{D,\lambda} := \boldsymbol{S}_\lambda \boldsymbol{e}_{D,0}, \qquad \boldsymbol{\Sigma}_{D,\lambda} := \boldsymbol{S}_\lambda \boldsymbol{\Sigma}_D \boldsymbol{S}_\lambda^*. \tag{30}$$

Theorem 8 gives the comparison directly: the proposed K-step physics-constrained flow has lower expected squared endpoint error precisely when

$$\|\boldsymbol{S}_\lambda \boldsymbol{\mu}_F\|^2 + \left[\rho_\lambda + \frac{1-\rho_\lambda}{K}\right] \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda} < \|\boldsymbol{S}_\lambda \boldsymbol{\mu}_D\|^2 + \operatorname{tr} \boldsymbol{\Sigma}_{D,\lambda}. \tag{31}$$

Under matched pre-physics marginal statistics, $\boldsymbol{\mu}_F = \boldsymbol{\mu}_D$ and $\boldsymbol{\Sigma}_F = \boldsymbol{\Sigma}_D$, this comparison simplifies to

$$\mathbb{E}[\|\boldsymbol{e}_{D,\lambda}\|^2] - \mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] = \frac{K-1}{K}(1-\rho_\lambda) \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda}. \tag{32}$$

Thus, once both methods receive the same analytical physics contraction, the additional advantage available to the proposed flow is controlled by the correlation of its post-physics local errors across flow time. Physics reduces error in operator-identifiable directions for both constructions, while the multi-step flow introduces an additional opportunity for nonpersistent errors to cancel.

Taken together, Theorems 5-8 identify the two complementary mechanisms underlying PEFM by showing (1) the contraction across physics-identifiable directions and (2) the variance reduction across flow time. The analysis does not claim that independently trained networks share the same pre-physics errors or that a trained physics-constrained field satisfies $\widetilde{\boldsymbol{\xi}}_k = \boldsymbol{S}_\lambda \boldsymbol{\xi}_k$ exactly. Rather, it provides an analytical fixed-error comparison and identifies measurable conditions under which physics-constrained FM is expected to improve endpoint prediction for which error energy must be present in physics-identifiable directions, and local prediction errors must not remain perfectly coherent along the flow.

4 Results & Experiments

4.1 Experimental setup

Datasets and metrics. We evaluate PEFM on translation from synthetic aortic flow and from real bioimpedance (BioZ) to pressure using the two-element Windkessel relationship (Westerhof et al., 2009). We use simulated waveforms from the Pulse Wave Database (Charlton et al., 2019) and graphene tattoo recordings (Kireev et al., 2022), previously used for PINN-based aortic waveform

reconstruction (Osman et al., 2025) and blood pressure estimation (Sel et al., 2023), respectively. For the measured task, a flow proxy is derived from filtered BioZ. We report root mean squared error (RMSE) and mean absolute error (MAE) for waveform accuracy and endpoint residuals for physical consistency. For graphene recordings, we also evaluate residuals using shared least-squares parameters fitted on training subjects, providing a common reference scale. To assess clinical applicability, we also report systolic (SBP) and diastolic blood pressure (DBP) estimation errors as mean error $\pm$ standard deviation on the graphene tattoo dataset. Together, these tasks assess translation under simulated and measured physiological conditions.

Baselines. We compare PEFM with supervised prediction, standard FM (Lipman et al., 2023), and task-specific adaptations of WPINN (Osman et al., 2026), PBFM (Baldan et al., 2026), and PIFM-GN (Song & Yu, 2026), which constrain predicted waveforms, integrated endpoints, and vector-field architecture, respectively (Table 1). All methods share a one-dimensional U-Net backbone (Ronneberger et al., 2015), with an additional branch for PIFM-GN. Following WPINN, all physics-informed methods learn subject-specific Windkessel parameters, unknown at test time, through an auxiliary backbone head. Each method fits these parameters from its estimated waveform and evaluates its endpoint residual under its own estimates. Appendix B details the Windkessel instantiation, flow-proxy construction, datasets, preprocessing, splits, baseline adaptations, and implementation details.

4.2 Experimental results

Virtual aortic flow-to-pressure. We first evaluate waveform accuracy and physical consistency on the virtual aortic dataset (Table 2). PEFM achieves the lowest MAE and RMSE at 4.399 and 5.335 mmHg, respectively, and the lowest endpoint residual at 214.8. The table additionally reports relative errors in R, C, and P_{out} against the reference parameters.

Table 2: **Virtual aortic flow-to-pressure results.** Waveform errors are in mmHg. Relative errors in R, C, and reconstructed P_{out} use learned parameters when available. Otherwise, they use least-squares Windkessel parameters fitted to the input-output relationship. Lower is better.

Method	RMSE ↓	MAE ↓	Residual ↓	R rel. err.	C rel. err.	P_{out} rel. err.
Naive	15.815	12.863	39987.7	—	—	—
PIFM-GN	5.562	4.588	715.9	0.071	0.022	0.138
WPINN	5.499	4.563	291.2	0.110	0.111	0.097
Flow Matching	5.470	4.559	20761.9	0.090	0.043	0.172
PBFM	5.454	4.572	17586.9	**0.064**	0.078	**0.096**
Supervised	5.434	4.504	970.2	0.180	0.029	0.294
PEFM	**5.335**	**4.399**	**214.8**	0.148	**0.021**	0.321

Measured BioZ-to-pressure. We next evaluate BioZ-to-pressure translation on the graphene tattoo dataset (Table 3), reporting RMSE, MAE, and residuals as mean $\pm$ standard deviation over five seeds. PEFM achieves the lowest RMSE and MAE at 14.2 and 11.5 mmHg, respectively. WPINN records the lowest endpoint residuals under both model-specific and shared least-squares parameters, followed by PEFM. The table also reports SBP and DBP estimation errors, with PEFM obtaining 0.61 ± 14.25 and 1.46 ± 11.22 mmHg, respectively.

Endpoint error analysis. Finally, we measure the held-out error statistics on the graphene tattoo dataset entering the endpoint error decomposition in Eq. 23, using $K = 50$ integration steps and five seeds (Table 4). We report squared mean errors, covariance traces, and cross-flow error correlations for separately trained models with and without physics regularization. The measured correlations are 0.965 without physics and 0.915 with physics. The corresponding direct and flow endpoint mean squared error quantities are 260.8 versus 245.3 and 254.2 versus 203.6, respectively.

4.3 Analysis and discussion

Our theoretical results show that an endpoint governing relationship can be converted into a local constraint on the FM vector field. For affine paths and linear operators, this constraint has a closed

Table 3: **Graphene tattoo BioZ-to-pressure results.** RMSE and MAE are in mmHg. $\|\boldsymbol{r}_v\|^2$ uses each model's learned parameters;$\|\boldsymbol{r}_v\|^2_{\mathrm{LS}}$ uses shared least-squares Windkessel parameters fitted on the training subjects. SBP and DBP estimation errors are reported as mean error (ME) $\pm$ standard deviation (SD), in mmHg. Lower is better.

Method	RMSE ↓	MAE ↓	$\|\boldsymbol{r}_v\|^2$ / $\|\boldsymbol{r}_v\|^2_{\mathrm{LS}}$ ↓	SBP ME $\pm$ SD	DBP ME $\pm$ SD
Naive	22.52	17.38	— / —	-39.47 ± 14.15	7.48 ± 10.57
PIFM-GN	19.0 ± 2.8	16.0 ± 2.5	— / 0.098 ± 0.051	18.06 ± 14.17	5.91 ± 11.99
WPINN	16.1 ± 0.5	13.7 ± 0.7	$\mathbf{0.043 \pm 0.016}$ / $\mathbf{0.043 \pm 0.016}$	-5.65 ± 16.17	-2.57 ± 13.37
Supervised	15.4 ± 1.2	12.9 ± 1.3	— / 0.091 ± 0.043	2.54 ± 14.79	-8.77 ± 10.77
PBFM	15.2 ± 2.5	12.6 ± 2.5	0.085 ± 0.054 / 0.085 ± 0.055	11.22 ± 13.84	$\mathbf{1.31 \pm 10.25}$
Flow Matching	15.1 ± 2.2	12.3 ± 1.8	— / 0.121 ± 0.017	13.07 ± 13.54	1.42 ± 10.10
PEFM	$\mathbf{14.2 \pm 1.4}$	$\mathbf{11.5 \pm 1.2}$	0.045 ± 0.024 / 0.044 ± 0.023	$\mathbf{0.61 \pm 14.25}$	1.46 ± 11.22

Table 4: **Held-out error statistics.** Measured on the graphene tattoo dataset with $K = 50$. $\boldsymbol{\mu}_D, \boldsymbol{\Sigma}_D$, mean and covariance of direct endpoint errors. $\boldsymbol{\mu}_F, \boldsymbol{\Sigma}_F$, of local FM velocity errors (for $\lambda > 0$, measured on a separately trained physics model as the empirical counterpart of $\widetilde{\boldsymbol{\xi}}_k$). ρ_λ, average off-diagonal covariance trace divided by $\operatorname{tr} \boldsymbol{\Sigma}_F$.

Physics	$\|\boldsymbol{\mu}_D\|^2$	$\operatorname{tr} \boldsymbol{\Sigma}_D$	$\|\boldsymbol{\mu}_F\|^2$	$\operatorname{tr} \boldsymbol{\Sigma}_F$	ρ_λ	$\|\boldsymbol{\mu}_D\|^2 + \operatorname{tr} \boldsymbol{\Sigma}_D$	$\|\boldsymbol{\mu}_F\|^2 + \alpha_K(\rho_\lambda) \operatorname{tr} \boldsymbol{\Sigma}_F$
$\lambda = 0$	60.8 ± 37.8	200 ± 3.5	19.3 ± 26.5	234 ± 91.2	0.965 ± 0.006	260.8 ± 38.0	245.3 ± 92.0
$\lambda > 0$	60.2 ± 57.8	194 ± 43.1	6.52 ± 5.91	215 ± 50.2	0.915 ± 0.005	254.2 ± 72.1	203.6 ± 46.4

form, and controlling it along the trajectory bounds the final physical residual. For the straight path, the velocity residual is exactly the physical residual of the endpoint implied by each local prediction, requiring neither endpoint reconstruction nor trajectory rollout during training. The error analysis identifies two complementary effects. Physics contracts errors observable through the governing operator, while multi-step integration reduces errors that do not persist across flow time.

The experiments support these mechanisms. On the virtual dataset in Table 2, PEFM achieves the lowest MAE, RMSE, and residual, reducing the next-best residual by 26.2%. Its weaker recovery of R and P_{out} is expected because nearly equivalent parameter combinations can produce low residuals and closely reproduce the waveform. On measured data in Table 3, PEFM achieves the lowest MAE and RMSE, with 11.8% lower RMSE than the closest baseline in physical consistency and 63.6% lower residual than the closest baseline in RMSE. Moreover, the similar residuals under learned and shared least-squares parameters, 0.045 and 0.044, suggest that this improvement is not solely due to parameter adaptation. PEFM also achieves the lowest absolute SBP ME at 0.61 mmHg. Overall, PEFM provides a better balance between predictive accuracy and physical consistency rather than optimizing either metric alone.

The error statistics in Table 4 provide further support for this interpretation. Without physics, FM slightly lowers the mean estimated endpoint error from 260.8 to 245.3. With physics, the reduction is more pronounced, from 254.2 to 203.6 relative to physics-informed direct prediction. Physics also lowers the squared mean flow error from 19.3 to 6.52 and the cross-flow correlation from 0.965 to 0.915, allowing more local error to cancel during integration. These results support the mechanisms predicted by the analysis. Overall, they support the central conclusion that governing the local transport combines physical supervision and flow-wise error cancellation to produce more accurate and physically consistent predictions.

5 Conclusion

We introduced PEFM, a physics-informed paired flow matching framework that imposes a governing relationship directly on the learned local velocity field. Our limitations include evaluation on only two datasets and a simplified two-element Windkessel relationship, as well as uncertainty in the learned physical parameters due to identifiability and, for BioZ, the use of a proxy for flow. Future work should study more complex and nonlinear governing relationships, improve parameter estimation, and evaluate PEFM across a broader range of physiological translation tasks.

AI USE STATEMENT

We used generative AI tools as general purpose writing and coding assistance. Specifically, they were used to polish the wording of text drafted by the authors and to assist with routine implementation and plotting code. Generative AI tools were not used to generate research ideas, generate figures, derive or verify the theoretical results, or produce experimental results. All AI assisted text was written, checked, and edited by the authors, and all AI assisted code was reviewed and tested against the reported results by the authors. The derivations in Appendix A were produced and verified by the authors. We take full responsibility for the final content of this work, including all text, claims, and artifacts.

ETHICS STATEMENT

This work uses only previously published, publicly available data. The virtual aortic waveforms come from the simulated Pulse Wave Database (Charlton et al., 2019) and contain no human subject data. The BioZ and blood pressure recordings are from the publicly released graphene tattoo dataset (Kireev et al., 2022), collected under the ethical approvals reported in that work. Conflicts of interest will be provided after acceptance.

REPRODUCIBILITY STATEMENT

All theoretical claims are stated with their assumptions in Section 3 and proved in full in Appendix A. The Windkessel instantiation, the flow-proxy construction, dataset sources, preprocessing, and the subject-level splits are described in Appendix B, and both datasets are publicly available (Charlton et al., 2019; Kireev et al., 2022). The backbone architecture, the baseline adaptations, the training hyperparameters, and the evaluation protocol, including the $K = 50$ integration setting and the shared least-squares parameter fit, are given in Section 4 and Appendix B. Source code for training and evaluation, along with trained model checkpoints, will be made publicly available upon acceptance.

REFERENCES

Ali Akbari, Jonathan Martinez, and Roozbeh Jafari. A meta-learning approach for fast personalization of modality translation models in wearable physiological sensing. *IEEE journal of biomedical and health informatics*, 26(4):1516–1527, 2021.

Michael S. Albergo, Nicholas M. Boffi, and Eric Vanden-Eijnden. Stochastic interpolants: A unifying framework for flows and diffusions. *CoRR*, abs/2303.08797, 2023. URL `https://doi.org/10.48550/arXiv.2303.08797`.

Michael S. Albergo, Mark Goldstein, Nicholas M. Boffi, Rajesh Ranganath, and Eric Vanden-Eijnden. Stochastic interpolants with data-dependent couplings. In *International Conference on Machine Learning (ICML)*, 2024.

Giacomo Baldan, Qiang Liu, Alberto Guardone, and Nils Thuerey. Physics vs distributions: Pareto optimal flow matching with physics constraints. In *International Conference on Learning Representations*, volume 2026, pp. 101215–101243, 2026.

Jan-Hendrik Bastek, WaiChing Sun, and Dennis Kochmann. Physics-informed diffusion models. In *International Conference on Learning Representations*, volume 2025, pp. 3360–3385, 2025.

Peter H Charlton, Jorge Mariscal Harana, Samuel Vennin, Ye Li, Phil Chowienczyk, and Jordi Alastruey. Modeling arterial pulse waves in healthy aging: a database for in silico evaluation of hemodynamics and pulse wave indexes. *American Journal of Physiology-Heart and Circulatory Physiology*, 317(5):H1062–H1085, 2019.

Mauricio Delbracio and Peyman Milanfar. Inversion by direct iteration: An alternative to denoising diffusion for image restoration. *Transactions on Machine Learning Research*, 2023.

Xiaocheng Fang, Jiarui Jin, Haoyu Wang, Che Liu, Jieyi Cai, Yujie Xiao, Guangkun Nie, Bo Liu, Shun Huang, Hongyan Li, et al. Ppgflowecg: Latent rectified flow with cross-modal encoding for ppg-guided ecg generation and cardiovascular disease detection. *arXiv preprint arXiv:2509.19774*, 2025.

Kevin Frans, Danijar Hafner, Sergey Levine, and Pieter Abbeel. One step diffusion via shortcut models. In *International Conference on Learning Representations (ICLR)*, 2025.

Eric Heitz, Laurent Belcour, and Thomas Chambon. Iterative α-(de)Blending: a minimalist deterministic diffusion model. In *ACM SIGGRAPH Conference Proceedings*, 2023.

Jiahe Huang, Guandao Yang, Zichen Wang, and Jeong Joon Park. Diffusionpde: Generative pde-solving under partial observation. *Advances in Neural Information Processing Systems*, 37: 130291–130323, 2024.

Christian Jacobsen, Yilin Zhuang, and Karthik Duraisamy. Cocogen: Physically consistent and conditioned score-based generative models for forward and inverse problems. *SIAM Journal on Scientific Computing*, 47(2):C399–C425, 2025.

Sebastián Jara, Julio Sotelo, David Ortiz-Puerta, Pablo A Estévez, Sergio Uribe, Steren Chabert, and Rodrigo Salas. Physics-informed neural network for modeling the pulmonary artery blood pressure from magnetic resonance images: A reduced-order navier–stokes model. *Biomedicines*, 13(9):2058, 2025.

Dmitry Kireev, Kaan Sel, Bassem Ibrahim, Neelotpala Kumar, Ali Akbari, Roozbeh Jafari, and Deji Akinwande. Continuous cuffless monitoring of arterial blood pressure via graphene bioimpedance tattoos. *Nature nanotechnology*, 17(8):864–870, 2022.

Georgios Kissas, Yibo Yang, Eileen Hwuang, Walter R Witschey, John A Detre, and Paris Perdikaris. Machine learning in cardiovascular flows modeling: Predicting arterial blood pressure from non-invasive 4d flow mri data using physics-informed neural networks. *Computer methods in applied mechanics and engineering*, 358:112623, 2020.

Liviu Klein, Marat Fudim, Mozziyar Etemadi, Robert Gordon, Anjan Tibrewala, Jaime Hernandez-Montfort, Patrick McCann, Lucas Zier, Kevin Shah, Allman Rollins, et al. Noninvasive pulmonary capillary wedge pressure estimation in heart failure patients with the use of wearable sensing and ai. *Heart Failure*, 13:102513, 2025.

Xiaoyan Li, Shixin Xu, Faisal Habib, Neda Aminnejad, Arvind Gupta, and Huaxiong Huang. Clepgan: an innovative approach to subject-independent ecg reconstruction from ppg signals. *BMC bioinformatics*, 26(1):306, 2025.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *The Eleventh International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Guan-Horng Liu, Arash Vahdat, De-An Huang, Evangelos A. Theodorou, Weili Nie, and Anima Anandkumar. I^2SB: Image-to-image Schrödinger bridge. In *International Conference on Machine Learning (ICML)*, 2023a.

Xingchao Liu, Chengyue Gong, and qiang liu. Flow straight and fast: Learning to generate and transfer data with rectified flow. In *The Eleventh International Conference on Learning Representations*, 2023b. URL https://openreview.net/forum?id=XVjTT1nw5z.

Ziyang LIU, Ruiqiang Xiao, Chang Huang, KIEREN YU, Siyuan HE, and Kaishun Wu. Cardiommflow: A gaussian-prior-free physics-informed flow matching framework for electrocardiogram to mmwave radar synthesis. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=e43JwTIb2V.

Deen Osman, Kaan Sel, and Roozbeh Jafari. Reconstruction of aortic waveforms from peripheral data using physics informed neural networks. In *2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1–5. IEEE, 2025.

Deen Osman, Kaan Sel, Erica Spatz, and Roozbeh Jafari. Cardiovascular digital twins using a windkessel physics informed neural network. *npj Digital Medicine*, 2026.

Aram-Alexandre Pooladian, Heli Ben-Hamu, Carles Domingo-Enrich, Brandon Amos, Yaron Lipman, and Ricky T. Q. Chen. Multisample flow matching: Straightening flows with minibatch couplings. In *International Conference on Machine Learning (ICML)*, 2023.

Xi Qiu, Hailin Cao, Li Yang, and Hui Wang. A physics-aware diffusion framework for robust ecg synthesis using mesoscopic lattice boltzmann constraints. *Biology*, 15(5):431, 2026.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015. URL `https://arxiv.org/abs/1505.04597`.

Kaan Sel, Amirmohammad Mohammadi, Roderic I Pettigrew, and Roozbeh Jafari. Physics-informed neural networks for modeling physiological time series for cuffless blood pressure estimation. *npj Digital Medicine*, 6(1):110, 2023.

Debaditya Shome, Pritam Sarkar, and Ali Etemad. Region-disentangled diffusion model for high-fidelity ppg-to-ecg translation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 38, pp. 15009–15019, 2024.

Dule Shu, Zijie Li, and Amir Barati Farimani. A physics-informed diffusion model for high-fidelity flow field reconstruction. *Journal of Computational Physics*, 478:111972, 2023.

Xiaozhuang Song and Tianshu Yu. Physically-informed flow matching with graph neural networks for complex fluid dynamics. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 40, pp. 1024–1032, 2026.

Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *International Conference on Machine Learning (ICML)*, 2023.

Shuntaro Suzuki, Shuitsu Koyama, Shinnosuke Hirano, and Shunya Nagashima. Penguin: General vital sign reconstruction from ppg with flow matching state space model. In *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6361–6365. IEEE, 2026.

Jan Tauberschmidt, Sophie Fellenz, Sebastian Vollmer, and Andrew Duncan. Physics-constrained fine-tuning of flow-matching models for generation and inverse problems. In *International Conference on Learning Representations*, volume 2026, pp. 143992–144037, 2026.

Alexander Tong, Kilian FATRAS, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio. Improving and generalizing flow-based generative models with minibatch optimal transport. *Transactions on Machine Learning Research*, 2024. ISSN 2835-8856.

Xiaoda Wang, Minxiao Wang, Kaiqiao Han, Defu Cao, Ching Chang, Yidan Shi, Runze Yan, Xiao Luo, Yan Liu, Xiao Hu, et al. Pg-lrf: Physiology-guided latent rectified flow for electrohemodynamic ppg-to-ecg generation. *arXiv preprint arXiv:2605.12541*, 2026.

Nico Westerhof, Jan-Willem Lankhaar, and Berend E Westerhof. The arterial windkessel. *Medical & biological engineering & computing*, 47(2):131–141, 2009.

Linqi Zhou, Aaron Lou, Samar Khanna, and Stefano Ermon. Denoising diffusion bridge models. In *International Conference on Learning Representations*, volume 2024, pp. 8160–8171, 2024.

A Proofs of Theorems and Propositions

Proposition 3: Affine velocity recovery and induced linear physics

Let $\mathcal{A}$ and $\mathcal{B}$ be linear differential operators acting along physical time t, and suppose the paired endpoints satisfy

$$\mathcal{A}\boldsymbol{x}_1 = \mathcal{B}\boldsymbol{x}_0. \tag{33}$$

Let $a, b \in C^1([0,1])$ satisfy

$$a(0) = 1, \qquad b(0) = 0, \qquad a(1) = 0, \qquad b(1) = 1, \tag{34}$$

and define

$$\boldsymbol{z}_\tau = a(\tau)\boldsymbol{x}_0 + b(\tau)\boldsymbol{x}_1, \qquad \boldsymbol{v}^\star_\tau := \partial_\tau \boldsymbol{z}_\tau = a'(\tau)\boldsymbol{x}_0 + b'(\tau)\boldsymbol{x}_1. \tag{35}$$

Assume $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ lie in the domains required for all operator expressions below. Then the following statements hold.

1. ***Velocity-based local endpoint recovery.*** *At every flow time τ for which $b'(\tau) \neq 0$, the endpoint is recoverable from the oracle velocity:*

$$\boldsymbol{x}_1 = \frac{\boldsymbol{v}^\star_\tau - a'(\tau)\boldsymbol{x}_0}{b'(\tau)}. \tag{36}$$

Replacing $\boldsymbol{v}^\star_\tau$ by a learned vector field therefore defines the velocity-implied endpoint

$$\widehat{\boldsymbol{x}}^{(v)}_1 := \frac{\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau) - a'(\tau)\boldsymbol{x}_0}{b'(\tau)}. \tag{37}$$

Hence, for $b'(\tau) \neq 0$, the endpoint physical law can be evaluated locally as a function of the learned velocity.

2. ***Induced velocity-space physics.*** *The oracle velocity satisfies*

$$\mathcal{A}\boldsymbol{v}^\star_\tau = \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0. \tag{38}$$

Accordingly, define

$$\boldsymbol{r}_v(\tau) := \mathcal{A}\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau) - \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0. \tag{39}$$

Then $\boldsymbol{r}_v(\tau) = 0$ for the oracle field. Moreover, whenever $b'(\tau) \neq 0$,

$$\boldsymbol{r}_v(\tau) = b'(\tau)\left[\mathcal{A}\widehat{\boldsymbol{x}}^{(v)}_1 - \mathcal{B}\boldsymbol{x}_0\right]. \tag{40}$$

Thus the induced velocity residual is exactly a path-speed-scaled endpoint physical residual evaluated on the endpoint implied by the local velocity. In the notation of Definition 1, $\boldsymbol{r}_v = b'(\tau)\boldsymbol{r}_H$; its squared penalty therefore differs from that of $\boldsymbol{r}_H$ by $b'(\tau)^2$.

3. ***Scalar-displacement specialization.*** *If*

$$a(\tau) = 1 - s(\tau), \qquad b(\tau) = s(\tau), \qquad s(0) = 0, \quad s(1) = 1, \tag{41}$$

and

$$\boldsymbol{g} := (\mathcal{B} - \mathcal{A})\boldsymbol{x}_0, \tag{42}$$

then

$$\mathcal{A}\boldsymbol{v}^\star_\tau = s'(\tau)\boldsymbol{g}, \qquad \boldsymbol{r}_v(\tau) = \mathcal{A}\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau) - s'(\tau)\boldsymbol{g}. \tag{43}$$

Wherever $s'(\tau) \neq 0$,

$$\widehat{\boldsymbol{x}}^{(v)}_1 := \boldsymbol{x}_0 + \frac{\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau)}{s'(\tau)}, \tag{44}$$

and

$$\boldsymbol{r}_v(\tau) = s'(\tau)\left[\mathcal{A}\widehat{\boldsymbol{x}}^{(v)}_1 - \mathcal{B}\boldsymbol{x}_0\right]. \tag{45}$$

4. ***Straight-path endpoint equivalence.*** *If $s(\tau) = \tau$, then*

$$\boldsymbol{v}^\star_\tau = \boldsymbol{d} := \boldsymbol{x}_1 - \boldsymbol{x}_0, \qquad \widehat{\boldsymbol{x}}^{(v)}_1 = \boldsymbol{x}_0 + \boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau), \tag{46}$$

and

$$\boldsymbol{r}_v(\tau) = \mathcal{A}\widehat{\boldsymbol{x}}^{(v)}_1 - \mathcal{B}\boldsymbol{x}_0. \tag{47}$$

PROOF OF PROPOSITION 3

From Eq. 35,

$$\boldsymbol{v}_\tau^\star - a'(\tau)\boldsymbol{x}_0 = b'(\tau)\boldsymbol{x}_1. \tag{48}$$

Therefore, whenever $b'(\tau) \neq 0$,

$$\boldsymbol{x}_1 = \frac{\boldsymbol{v}_\tau^\star - a'(\tau)\boldsymbol{x}_0}{b'(\tau)}, \tag{49}$$

which proves Eq. 36. Replacing the oracle velocity by $\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau)$ gives Eq. 37. The condition $b'(\tau) \neq 0$ is essential for this velocity-based reconstruction: if $b'(\tau) = 0$, the local velocity contains no target-endpoint contribution through $\boldsymbol{x}_1$, even though the path state itself may still contain $\boldsymbol{x}_1$.

We next derive the induced velocity law. Applying $\mathcal{A}$ to the oracle velocity and using linearity,

$$\mathcal{A}\boldsymbol{v}_\tau^\star = a'(\tau)\mathcal{A}\boldsymbol{x}_0 + b'(\tau)\mathcal{A}\boldsymbol{x}_1. \tag{50}$$

Because $a'(\tau)$ and $b'(\tau)$ depend only on flow time while $\mathcal{A}$ acts only along physical time, they commute with $\mathcal{A}$. Using the endpoint physical relation $\mathcal{A}\boldsymbol{x}_1 = \mathcal{B}\boldsymbol{x}_0$,

$$\begin{aligned}\mathcal{A}\boldsymbol{v}_\tau^\star &= a'(\tau)\mathcal{A}\boldsymbol{x}_0 + b'(\tau)\mathcal{B}\boldsymbol{x}_0 \\ &= \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0,\end{aligned} \tag{51}$$

which proves Eq. 38. The residual Eq. 39 follows by replacing $\boldsymbol{v}_\tau^\star$ with the learned field and measuring the discrepancy from this identity.

To connect the velocity residual directly to endpoint physics, suppose $b'(\tau) \neq 0$. From Eq. 37,

$$b'(\tau)\widehat{\boldsymbol{x}}_1^{(v)} = \boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau) - a'(\tau)\boldsymbol{x}_0. \tag{52}$$

Applying $\mathcal{A}$ and using again that $a'(\tau)$ and $b'(\tau)$ are scalars with respect to physical time,

$$b'(\tau)\mathcal{A}\widehat{\boldsymbol{x}}_1^{(v)} = \mathcal{A}\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau) - a'(\tau)\mathcal{A}\boldsymbol{x}_0. \tag{53}$$

Subtracting $b'(\tau)\mathcal{B}\boldsymbol{x}_0$ from both sides yields

$$\begin{aligned}b'(\tau)\left[\mathcal{A}\widehat{\boldsymbol{x}}_1^{(v)} - \mathcal{B}\boldsymbol{x}_0\right] &= \mathcal{A}\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau) - \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0 \\ &= \boldsymbol{r}_v(\tau),\end{aligned} \tag{54}$$

which proves Eq. 40. This identity makes explicit that the proposed local physics penalty genuinely constrains the learned vector field through the endpoint it implies.

For the scalar-displacement family,

$$a'(\tau) = -s'(\tau), \qquad b'(\tau) = s'(\tau).$$

Therefore,

$$\begin{aligned}\mathcal{A}\boldsymbol{v}_\tau^\star &= \left[-s'(\tau)\mathcal{A} + s'(\tau)\mathcal{B}\right]\boldsymbol{x}_0 \\ &= s'(\tau)(\mathcal{B} - \mathcal{A})\boldsymbol{x}_0 \\ &= s'(\tau)\boldsymbol{g},\end{aligned} \tag{55}$$

which gives Eq. 43. When $s'(\tau) \neq 0$, Eq. 37 reduces to

$$\begin{aligned}\widehat{\boldsymbol{x}}_1^{(v)} &= \frac{\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau) + s'(\tau)\boldsymbol{x}_0}{s'(\tau)} \\ &= \boldsymbol{x}_0 + \frac{\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau)}{s'(\tau)},\end{aligned} \tag{56}$$

which proves Eq. 44. Substituting $b'(\tau) = s'(\tau)$ into Eq. 40 gives Eq. 45.

Finally, for the straight path $s(\tau) = \tau$, one has $s'(\tau) = 1$, so

$$\boldsymbol{v}_\tau^\star = \boldsymbol{x}_1 - \boldsymbol{x}_0, \qquad \widehat{\boldsymbol{x}}_1^{(v)} = \boldsymbol{x}_0 + \boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau).$$

Eq. 45 therefore reduces to

$$\boldsymbol{r}_v(\tau) = \mathcal{A}\widehat{\boldsymbol{x}}_1^{(v)} - \mathcal{B}\boldsymbol{x}_0,$$

which proves Eq. 47. □

PROPOSITION 2: TRAJECTORY RESIDUAL CONTROLS ENDPOINT RESIDUAL (RESTATED)

Let $\widehat{\boldsymbol{z}}_\tau$ denote the trajectory generated by the learned vector field

$$\frac{\mathrm{d}\widehat{\boldsymbol{z}}_\tau}{\mathrm{d}\tau} = \boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau), \qquad \widehat{\boldsymbol{z}}_0 = \boldsymbol{x}_0, \tag{57}$$

and suppose that the endpoint pair satisfies

$$\mathcal{A}\boldsymbol{x}_1 = \mathcal{B}\boldsymbol{x}_0, \tag{58}$$

with $\mathcal{A}$ and $\mathcal{B}$ linear operators acting only along physical time.

For the affine interpolation

$$\boldsymbol{z}_\tau = a(\tau)\boldsymbol{x}_0 + b(\tau)\boldsymbol{x}_1, \tag{59}$$

assume $a, b \in C^1([0,1])$ and

$$a(0) = 1, \qquad b(0) = 0, \qquad a(1) = 0, \qquad b(1) = 1, \tag{60}$$

and define the trajectory-wise velocity residual

$$\boldsymbol{r}_v^{\mathrm{traj}}(\tau) := \mathcal{A}\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - [a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}]\,\boldsymbol{x}_0. \tag{61}$$

Then the physical residual of the generated endpoint satisfies the exact identity

$$\mathcal{A}\widehat{\boldsymbol{z}}_1 - \mathcal{B}\boldsymbol{x}_0 = \int_0^1 \boldsymbol{r}_v^{\mathrm{traj}}(\tau)\,\mathrm{d}\tau, \tag{62}$$

$$\|\mathcal{A}\widehat{\boldsymbol{z}}_1 - \mathcal{B}\boldsymbol{x}_0\|^2 \le \int_0^1 \left\|\boldsymbol{r}_v^{\mathrm{traj}}(\tau)\right\|^2 \mathrm{d}\tau. \tag{63}$$

For the scalar-displacement family

$$a(\tau) = 1 - s(\tau), \qquad b(\tau) = s(\tau), \qquad s(0) = 0, \quad s(1) = 1, \tag{64}$$

define

$$\boldsymbol{g} := (\mathcal{B} - \mathcal{A})\boldsymbol{x}_0. \tag{65}$$

Then

$$\boldsymbol{r}_v^{\mathrm{traj}}(\tau) = \mathcal{A}\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - s'(\tau)\boldsymbol{g}. \tag{66}$$

To separate the schedule-dependent path speed from the predicted displacement, wherever $s'(\tau) \neq 0$, define the rescaled velocity

$$\boldsymbol{u}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) := \frac{\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau)}{s'(\tau)}. \tag{67}$$

For the oracle velocity, this normalization gives $\boldsymbol{v}_\tau^\star / s'(\tau) = \boldsymbol{x}_1 - \boldsymbol{x}_0$, so $\boldsymbol{u}_\theta$ can be interpreted as a schedule-independent estimate of the endpoint displacement. Then

$$\boldsymbol{r}_v^{\mathrm{traj}}(\tau) = s'(\tau)\left[\mathcal{A}\boldsymbol{u}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - \boldsymbol{g}\right], \tag{68}$$

and hence

$$\left\|\boldsymbol{r}_v^{\mathrm{traj}}(\tau)\right\|^2 = s'(\tau)^2 \left\|\mathcal{A}\boldsymbol{u}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - \boldsymbol{g}\right\|^2. \tag{69}$$

Thus the interpolation schedule induces the multiplicative physics weighting $s'(\tau)^2$ across flow time. In particular, the straight path $s(\tau) = \tau$ yields $s'(\tau)^2 \equiv 1$, and therefore uniform weighting.

PROOF OF PROPOSITION 2

Integrating the trajectory equation Eq. 57 over $\tau \in [0,1]$ gives

$$\widehat{\boldsymbol{z}}_1 - \widehat{\boldsymbol{z}}_0 = \int_0^1 \boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau)\,\mathrm{d}\tau. \tag{70}$$

Since $\widehat{\boldsymbol{z}}_0 = \boldsymbol{x}_0$,

$$\widehat{\boldsymbol{z}}_1 = \boldsymbol{x}_0 + \int_0^1 \boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau)\,\mathrm{d}\tau. \tag{71}$$

Applying $\mathcal{A}$ to Eq. 71 and subtracting $\mathcal{B}\boldsymbol{x}_0$ yields

$$\mathcal{A}\widehat{\boldsymbol{z}}_1 - \mathcal{B}\boldsymbol{x}_0 = \mathcal{A}\boldsymbol{x}_0 - \mathcal{B}\boldsymbol{x}_0 + \mathcal{A}\left(\int_0^1 \boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau)\,\mathrm{d}\tau\right). \tag{72}$$

$\mathcal{A}$ commutes with integration over the independent flow-time variable τ. Therefore,

$$\mathcal{A}\left(\int_0^1 \boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau)\,\mathrm{d}\tau\right) = \int_0^1 \mathcal{A}\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau)\,\mathrm{d}\tau, \tag{73}$$

so

$$\mathcal{A}\widehat{\boldsymbol{z}}_1 - \mathcal{B}\boldsymbol{x}_0 = \mathcal{A}\boldsymbol{x}_0 - \mathcal{B}\boldsymbol{x}_0 + \int_0^1 \mathcal{A}\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau)\,\mathrm{d}\tau. \tag{74}$$

We now integrate the affine forcing term appearing in Eq. 61. Since $\boldsymbol{x}_0$ is independent of τ,

$$\int_0^1 \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0\,\mathrm{d}\tau = \left(\int_0^1 a'(\tau)\,\mathrm{d}\tau\right)\mathcal{A}\boldsymbol{x}_0 + \left(\int_0^1 b'(\tau)\,\mathrm{d}\tau\right)\mathcal{B}\boldsymbol{x}_0. \tag{75}$$

By the fundamental theorem of calculus and the boundary conditions Eq. 60,

$$\int_0^1 a'(\tau)\,\mathrm{d}\tau = a(1) - a(0) = -1, \tag{76}$$

$$\int_0^1 b'(\tau)\,\mathrm{d}\tau = b(1) - b(0) = 1. \tag{77}$$

Hence

$$\int_0^1 \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0\,\mathrm{d}\tau = -\mathcal{A}\boldsymbol{x}_0 + \mathcal{B}\boldsymbol{x}_0. \tag{78}$$

Rearranging Eq. 78,

$$\mathcal{A}\boldsymbol{x}_0 - \mathcal{B}\boldsymbol{x}_0 = -\int_0^1 \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0\,\mathrm{d}\tau. \tag{79}$$

Substituting Eq. 79 into Eq. 74 gives

$$\begin{aligned}
\mathcal{A}\widehat{\boldsymbol{z}}_1 - \mathcal{B}\boldsymbol{x}_0 &= \int_0^1 \mathcal{A}\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau)\,\mathrm{d}\tau \\
&\quad - \int_0^1 \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0\,\mathrm{d}\tau \\
&= \int_0^1 \left\{\mathcal{A}\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - \left[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}\right]\boldsymbol{x}_0\right\}\mathrm{d}\tau \\
&= \int_0^1 \boldsymbol{r}_v^{\mathrm{traj}}(\tau)\,\mathrm{d}\tau,
\end{aligned} \tag{80}$$

which proves the identity Eq. 62.

To obtain the norm bound, apply the norm inequality for the integral and the Cauchy-Schwarz inequality:

$$\begin{aligned}
\left\|\int_0^1 \boldsymbol{r}_v^{\mathrm{traj}}(\tau)\,\mathrm{d}\tau\right\|^2 &\le \left(\int_0^1 1^2\,\mathrm{d}\tau\right)\left(\int_0^1 \left\|\boldsymbol{r}_v^{\mathrm{traj}}(\tau)\right\|^2\,\mathrm{d}\tau\right) \\
&= \int_0^1 \left\|\boldsymbol{r}_v^{\mathrm{traj}}(\tau)\right\|^2\,\mathrm{d}\tau.
\end{aligned} \tag{81}$$

Combining this with Eq. 62 proves Eq. 63. The implication is one-way: controlling the trajectory-wise residual controls the physical residual of the generated endpoint, but a small endpoint residual does not imply that $\boldsymbol{r}_v^{\mathrm{traj}}(\tau)$ is small at every flow time. Indeed, the exact identity Eq. 62 permits nonzero local residuals to cancel under integration.

We now specialize to the scalar-displacement family. From Eq. 64,

$$a'(\tau) = -s'(\tau), \qquad b'(\tau) = s'(\tau). \tag{82}$$

Therefore,

$$\begin{aligned}[a'(\tau)\mathcal{A} + b'(\tau)\mathcal{B}]\,\boldsymbol{x}_0 &= [-s'(\tau)\mathcal{A} + s'(\tau)\mathcal{B}]\,\boldsymbol{x}_0 \\ &= s'(\tau)(\mathcal{B} - \mathcal{A})\boldsymbol{x}_0 \\ &= s'(\tau)\boldsymbol{g}. \end{aligned} \tag{83}$$

Substitution into Eq. 61 yields Eq. 66.

At every τ for which $s'(\tau) \neq 0$, the rescaled field in Eq. 67 is well defined. This normalization removes the path-speed factor because the oracle velocity on a scalar path is $\boldsymbol{v}^\star_\tau = s'(\tau)(\boldsymbol{x}_1 - \boldsymbol{x}_0)$, so $\boldsymbol{v}^\star_\tau / s'(\tau) = \boldsymbol{x}_1 - \boldsymbol{x}_0$. Thus $\boldsymbol{u}_\theta$ represents the schedule-independent endpoint displacement, and its definition gives

$$\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) = s'(\tau)\boldsymbol{u}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau). \tag{84}$$

Because $s'(\tau)$ is a scalar depending only on flow time and $\mathcal{A}$ acts only on physical time,

$$\mathcal{A}\boldsymbol{v}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) = s'(\tau)\mathcal{A}\boldsymbol{u}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau). \tag{85}$$

Thus,

$$\begin{aligned}\boldsymbol{r}_v^{\mathrm{traj}}(\tau) &= s'(\tau)\mathcal{A}\boldsymbol{u}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - s'(\tau)\boldsymbol{g} \\ &= s'(\tau)\left[\mathcal{A}\boldsymbol{u}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - \boldsymbol{g}\right], \end{aligned} \tag{86}$$

which proves Eq. 68. Taking squared norms gives

$$\begin{aligned}\left\|\boldsymbol{r}_v^{\mathrm{traj}}(\tau)\right\|^2 &= |s'(\tau)|^2 \left\|\mathcal{A}\boldsymbol{u}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - \boldsymbol{g}\right\|^2 \\ &= s'(\tau)^2 \left\|\mathcal{A}\boldsymbol{u}_\theta(\widehat{\boldsymbol{z}}_\tau, \tau) - \boldsymbol{g}\right\|^2, \end{aligned} \tag{87}$$

which proves Eq. 69. The scalar interpolation schedule thus weights the squared physical discrepancy by $s'(\tau)^2$. For the straight path $s(\tau) = \tau$, $s'(\tau) = 1$ for all τ, so this weighting is uniform. If $s'(\tau)$ vanishes at isolated flow times, the original residual Eq. 66 remains well defined at those points. Only the rescaled representation $\boldsymbol{u}_\theta = \boldsymbol{v}_\theta / s'(\tau)$ is undefined. Accordingly, neither the trajectory identity nor the endpoint bound requires $s'(\tau) \neq 0$ for every τ. The rescaling is therefore only an analytical device for exposing the schedule-induced weighting. Evaluating the original residual does not require dividing by $s'(\tau)$. □

PROPOSITION 4: PHYSICS SHRINKAGE AND INTEGRATED ENDPOINT ERROR

Consider finite-dimensional real signals with Euclidean norm and a fixed linear operator $\mathcal{A}$, with adjoint $\mathcal{A}^*$. Let $\lambda \geq 0$ and $K \geq 1$. Consider the straight path

$$\boldsymbol{z}_\tau = \boldsymbol{x}_0 + \tau(\boldsymbol{x}_1 - \boldsymbol{x}_0), \tag{88}$$

under the exact linear physical law

$$\mathcal{A}\boldsymbol{x}_1 = \mathcal{B}\boldsymbol{x}_0. \tag{89}$$

Define

$$\boldsymbol{d} := \boldsymbol{x}_1 - \boldsymbol{x}_0, \qquad \boldsymbol{g} := (\mathcal{B} - \mathcal{A})\boldsymbol{x}_0, \tag{90}$$

so that $\mathcal{A}\boldsymbol{d} = \boldsymbol{g}$. At Euler step $k = 0, \ldots, K-1$, suppose the data-only FM velocity prediction is $\boldsymbol{d} + \boldsymbol{\xi}_k$, where $\boldsymbol{\xi}_k$ is its pre-physics local velocity error. For a candidate velocity $\boldsymbol{v}$, define

$$\widehat{\boldsymbol{v}}_k = \arg\min_{\boldsymbol{v}} \left\{ \|\boldsymbol{v} - (\boldsymbol{d} + \boldsymbol{\xi}_k)\|^2 + \lambda \|\mathcal{A}\boldsymbol{v} - \boldsymbol{g}\|^2 \right\}. \tag{91}$$

Let

$$\boldsymbol{S}_\lambda := (\boldsymbol{I} + \lambda \mathcal{A}^*\mathcal{A})^{-1}. \tag{92}$$

Then

$$\widehat{\boldsymbol{v}}_k = \boldsymbol{d} + \boldsymbol{S}_\lambda \boldsymbol{\xi}_k. \tag{93}$$

If $\mathcal{A} = \boldsymbol{U}\,\mathrm{diag}(\sigma_i)\boldsymbol{V}^*$, then an error component in the i-th right-singular direction is multiplied by

$$\frac{1}{1 + \lambda\sigma_i^2}, \tag{94}$$

while components in $\ker \mathcal{A}$ *are unchanged.*

For the uniform Euler updates

$$\widehat{\boldsymbol{z}}_{k+1} = \widehat{\boldsymbol{z}}_k + \frac{1}{K}\widehat{\boldsymbol{v}}_k, \qquad \widehat{\boldsymbol{z}}_0 = \boldsymbol{x}_0, \tag{95}$$

the endpoint error satisfies

$$\boldsymbol{e}_{K,\lambda} := \widehat{\boldsymbol{z}}_K - \boldsymbol{x}_1 = \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{S}_\lambda \boldsymbol{\xi}_k = \boldsymbol{S}_\lambda \boldsymbol{e}_{K,0}, \qquad \boldsymbol{e}_{K,0} := \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{\xi}_k. \tag{96}$$

Moreover,

$$\mathcal{A}\widehat{\boldsymbol{z}}_K - \mathcal{B}\boldsymbol{x}_0 = \mathcal{A}\boldsymbol{e}_{K,\lambda} = \frac{1}{K}\sum_{k=0}^{K-1} \mathcal{A}\boldsymbol{S}_\lambda \boldsymbol{\xi}_k. \tag{97}$$

PROOF OF PROPOSITION 4

Because $\boldsymbol{d} = \boldsymbol{x}_1 - \boldsymbol{x}_0$ and $\mathcal{A}\boldsymbol{x}_1 = \mathcal{B}\boldsymbol{x}_0$,

$$\mathcal{A}\boldsymbol{d} = \mathcal{A}(\boldsymbol{x}_1 - \boldsymbol{x}_0) = (\mathcal{B} - \mathcal{A})\boldsymbol{x}_0 = \boldsymbol{g}. \tag{98}$$

Write a candidate velocity as

$$\boldsymbol{v} = \boldsymbol{d} + \boldsymbol{\delta}, \tag{99}$$

where $\boldsymbol{\delta} = \boldsymbol{v} - \boldsymbol{d}$ is the candidate local velocity error. Then

$$\boldsymbol{v} - (\boldsymbol{d} + \boldsymbol{\xi}_k) = \boldsymbol{\delta} - \boldsymbol{\xi}_k, \qquad \mathcal{A}\boldsymbol{v} - \boldsymbol{g} = \mathcal{A}(\boldsymbol{d} + \boldsymbol{\delta}) - \boldsymbol{g} = \mathcal{A}\boldsymbol{\delta}. \tag{100}$$

Hence Eq. 91 becomes

$$\min_{\boldsymbol{\delta}} \left\{ \|\boldsymbol{\delta} - \boldsymbol{\xi}_k\|^2 + \lambda \|\mathcal{A}\boldsymbol{\delta}\|^2 \right\}. \tag{101}$$

Its first-order optimality condition is

$$\boldsymbol{\delta} - \boldsymbol{\xi}_k + \lambda \mathcal{A}^* \mathcal{A}\boldsymbol{\delta} = 0, \tag{102}$$

or equivalently

$$(\boldsymbol{I} + \lambda \mathcal{A}^* \mathcal{A})\boldsymbol{\delta} = \boldsymbol{\xi}_k. \tag{103}$$

Since $\mathcal{A}^*\mathcal{A}$ is positive semidefinite, $\boldsymbol{I} + \lambda \mathcal{A}^*\mathcal{A}$ is positive definite and invertible. Therefore

$$\boldsymbol{\delta} = \boldsymbol{S}_\lambda \boldsymbol{\xi}_k, \tag{104}$$

which gives Eq. 93.

Now let

$$\mathcal{A} = \boldsymbol{U} \operatorname{diag}(\sigma_i) \boldsymbol{V}^*. \tag{105}$$

Then

$$\mathcal{A}^*\mathcal{A} = \boldsymbol{V} \operatorname{diag}(\sigma_i^2) \boldsymbol{V}^*. \tag{106}$$

For a right-singular vector $\boldsymbol{q}_i$,

$$\boldsymbol{S}_\lambda \boldsymbol{q}_i = \frac{1}{1 + \lambda \sigma_i^2} \boldsymbol{q}_i, \tag{107}$$

proving Eq. 94. The factor is strictly below one when $\lambda > 0$ and $\sigma_i > 0$, and equals one when $\sigma_i = 0$, i.e., along $\ker \mathcal{A}$.

Finally, summing the Euler recursion yields

$$\begin{aligned}
\widehat{\boldsymbol{z}}_K &= \boldsymbol{x}_0 + \frac{1}{K}\sum_{k=0}^{K-1} \widehat{\boldsymbol{v}}_k \\
&= \boldsymbol{x}_0 + \frac{1}{K}\sum_{k=0}^{K-1} (\boldsymbol{d} + \boldsymbol{S}_\lambda \boldsymbol{\xi}_k) \\
&= \boldsymbol{x}_0 + \boldsymbol{d} + \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{S}_\lambda \boldsymbol{\xi}_k \\
&= \boldsymbol{x}_1 + \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{S}_\lambda \boldsymbol{\xi}_k.
\end{aligned} \tag{108}$$

Subtracting $\boldsymbol{x}_1$ gives

$$\boldsymbol{e}_{K,\lambda} = \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{S}_\lambda \boldsymbol{\xi}_k = \boldsymbol{S}_\lambda \left(\frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{\xi}_k\right) = \boldsymbol{S}_\lambda \boldsymbol{e}_{K,0}. \tag{109}$$

Using $\mathcal{A}\boldsymbol{x}_1 = \mathcal{B}\boldsymbol{x}_0$,

$$\mathcal{A}\widehat{\boldsymbol{z}}_K - \mathcal{B}\boldsymbol{x}_0 = \mathcal{A}(\widehat{\boldsymbol{z}}_K - \boldsymbol{x}_1) = \mathcal{A}\boldsymbol{e}_{K,\lambda}, \tag{110}$$

which proves Eq. 97. The physics penalty therefore acts locally through the deterministic shrinkage operator $\boldsymbol{S}_\lambda$, contracting velocity error in directions visible to $\mathcal{A}$ while leaving $\ker \mathcal{A}$ unchanged. The flow solver then forms the endpoint by averaging these locally regularized errors across flow time, so physics acts across operator directions, whereas integration acts across flow-time evaluations. □

THEOREM 5: EXPECTED SQUARED ENDPOINT ERROR OF PHYSICS-CONSTRAINED FLOW MATCHING

Consider finite-dimensional real signals with Euclidean norm, a fixed linear operator $\mathcal{A}$, $\lambda \geq 0$, and $K \geq 1$. Let

$$\boldsymbol{S}_\lambda = (\boldsymbol{I} + \lambda \mathcal{A}^*\mathcal{A})^{-1}. \tag{111}$$

Suppose a straight-path K-step FM construction has pre-physics local velocity errors $\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_{K-1}$ with common mean

$$\boldsymbol{\mu}_F := \mathbb{E}[\boldsymbol{\xi}_k] \tag{112}$$

and common covariance

$$\boldsymbol{\Sigma}_F := \mathrm{Cov}(\boldsymbol{\xi}_k). \tag{113}$$

Define the post-physics local errors

$$\widetilde{\boldsymbol{\xi}}_k := \boldsymbol{S}_\lambda \boldsymbol{\xi}_k, \tag{114}$$

with

$$\boldsymbol{\mu}_{F,\lambda} := \boldsymbol{S}_\lambda \boldsymbol{\mu}_F, \qquad \boldsymbol{\Sigma}_{F,\lambda} := \boldsymbol{S}_\lambda \boldsymbol{\Sigma}_F \boldsymbol{S}_\lambda^*. \tag{115}$$

Assume their cross-flow covariance traces satisfy, for every $i \neq j$,

$$\mathrm{tr}\,\mathrm{Cov}(\widetilde{\boldsymbol{\xi}}_i, \widetilde{\boldsymbol{\xi}}_j) = \rho_\lambda \,\mathrm{tr}\,\boldsymbol{\Sigma}_{F,\lambda}. \tag{116}$$

If the endpoint error is

$$\boldsymbol{e}_{K,\lambda} = \frac{1}{K}\sum_{k=0}^{K-1} \widetilde{\boldsymbol{\xi}}_k, \tag{117}$$

then

$$\mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] = \|\boldsymbol{S}_\lambda \boldsymbol{\mu}_F\|^2 + \left[\rho_\lambda + \frac{1-\rho_\lambda}{K}\right] \mathrm{tr}\,\boldsymbol{\Sigma}_{F,\lambda}. \tag{118}$$

Equivalently, the variance contribution is multiplied by

$$\alpha_K(\rho_\lambda) := \rho_\lambda + \frac{1-\rho_\lambda}{K}. \tag{119}$$

PROOF OF THEOREM 5

By Eq. 117,

$$\mathbb{E}[\boldsymbol{e}_{K,\lambda}] = \frac{1}{K}\sum_{k=0}^{K-1} \mathbb{E}[\widetilde{\boldsymbol{\xi}}_k] = \boldsymbol{S}_\lambda \boldsymbol{\mu}_F. \tag{120}$$

The endpoint covariance is

$$\mathrm{Cov}(\boldsymbol{e}_{K,\lambda}) = \frac{1}{K^2}\sum_{i=0}^{K-1}\sum_{j=0}^{K-1} \mathrm{Cov}(\widetilde{\boldsymbol{\xi}}_i, \widetilde{\boldsymbol{\xi}}_j). \tag{121}$$

Taking traces, the K diagonal terms each contribute $\operatorname{tr}\boldsymbol{\Sigma}_{F,\lambda}$, whereas the $K(K-1)$ off-diagonal terms each contribute $\rho_\lambda \operatorname{tr}\boldsymbol{\Sigma}_{F,\lambda}$. Hence

$$\begin{aligned}\operatorname{tr}\operatorname{Cov}(\boldsymbol{e}_{K,\lambda}) &= \frac{1}{K^2}\left[K + K(K-1)\rho_\lambda\right]\operatorname{tr}\boldsymbol{\Sigma}_{F,\lambda} \\ &= \left[\rho_\lambda + \frac{1-\rho_\lambda}{K}\right]\operatorname{tr}\boldsymbol{\Sigma}_{F,\lambda}. \end{aligned} \tag{122}$$

For every square-integrable random vector $\mathbf{y}$,

$$\mathbb{E}[\|\mathbf{y}\|^2] = \|\mathbb{E}[\mathbf{y}]\|^2 + \operatorname{tr}\operatorname{Cov}(\mathbf{y}). \tag{123}$$

Combining this identity with Eq. 120 and Eq. 122 proves Eq. 118. Therefore, multi-step integration does not reduce the squared mean error under fixed marginal statistics. It reduces only the variance term, and this reduction is governed entirely by the cross-flow correlation factor $\alpha_K(\rho_\lambda)$, so that perfectly coherent errors persist, whereas nonpersistent errors partially cancel during integration. For $K = 1$ there are no cross-flow covariance terms, so $\alpha_1 = 1$ and no such cancellation occurs. □

THEOREM 6: WHEN FLOW MATCHING IMPROVES ON DIRECT SUPERVISION

Consider a K-step straight-path FM model without physics regularization. Let its local velocity errors $\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_{K-1}$ have common mean

$$\boldsymbol{\mu}_F := \mathbb{E}[\boldsymbol{\xi}_k] \tag{124}$$

and common covariance

$$\boldsymbol{\Sigma}_F := \operatorname{Cov}(\boldsymbol{\xi}_k). \tag{125}$$

Assume that for $i \neq j$,

$$\operatorname{tr}\operatorname{Cov}(\boldsymbol{\xi}_i, \boldsymbol{\xi}_j) = \rho \operatorname{tr}\boldsymbol{\Sigma}_F. \tag{126}$$

Let the FM endpoint error be

$$\boldsymbol{e}_{K,0} := \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{\xi}_k. \tag{127}$$

Let a direct supervised model have endpoint error

$$\boldsymbol{e}_{D,0} := \widehat{\boldsymbol{x}}_{D,0} - \boldsymbol{x}_1, \tag{128}$$

with

$$\boldsymbol{\mu}_D := \mathbb{E}[\boldsymbol{e}_{D,0}], \qquad \boldsymbol{\Sigma}_D := \operatorname{Cov}(\boldsymbol{e}_{D,0}). \tag{129}$$

Then FM has lower expected squared endpoint error than direct supervision if and only if

$$\|\boldsymbol{\mu}_F\|^2 + \left[\rho + \frac{1-\rho}{K}\right]\operatorname{tr}\boldsymbol{\Sigma}_F < \|\boldsymbol{\mu}_D\|^2 + \operatorname{tr}\boldsymbol{\Sigma}_D. \tag{130}$$

Under matched marginal statistics,

$$\boldsymbol{\mu}_F = \boldsymbol{\mu}_D, \qquad \boldsymbol{\Sigma}_F = \boldsymbol{\Sigma}_D, \tag{131}$$

the direct-minus-FM gap is

$$\mathbb{E}[\|\boldsymbol{e}_{D,0}\|^2] - \mathbb{E}[\|\boldsymbol{e}_{K,0}\|^2] = \frac{K-1}{K}(1-\rho)\operatorname{tr}\boldsymbol{\Sigma}_F. \tag{132}$$

Hence, for $K > 1$ and $\operatorname{tr}\boldsymbol{\Sigma}_F > 0$, FM has a strict advantage under matched statistics exactly when $\rho < 1$.

More generally, if

$$\boldsymbol{\mu}_F = \boldsymbol{\mu}_D, \qquad \operatorname{tr}\boldsymbol{\Sigma}_D > 0, \qquad \alpha_K(\rho) := \rho + \frac{1-\rho}{K} > 0, \tag{133}$$

then FM has lower expected squared endpoint error if and only if

$$\frac{\operatorname{tr}\boldsymbol{\Sigma}_F}{\operatorname{tr}\boldsymbol{\Sigma}_D} < \frac{1}{\rho + (1-\rho)/K}. \tag{134}$$

PROOF OF THEOREM 6

The FM endpoint mean is

$$\mathbb{E}[\boldsymbol{e}_{K,0}] = \boldsymbol{\mu}_F. \tag{135}$$

Its covariance trace is

$$\begin{aligned}\operatorname{tr}\operatorname{Cov}(\boldsymbol{e}_{K,0}) &= \frac{1}{K^2}\left[K\operatorname{tr}\boldsymbol{\Sigma}_F + K(K-1)\rho\operatorname{tr}\boldsymbol{\Sigma}_F\right] \\ &= \left[\rho + \frac{1-\rho}{K}\right]\operatorname{tr}\boldsymbol{\Sigma}_F. \end{aligned} \tag{136}$$

Therefore

$$\mathbb{E}[\|\boldsymbol{e}_{K,0}\|^2] = \|\boldsymbol{\mu}_F\|^2 + \left[\rho + \frac{1-\rho}{K}\right]\operatorname{tr}\boldsymbol{\Sigma}_F. \tag{137}$$

For the direct model,

$$\mathbb{E}[\|\boldsymbol{e}_{D,0}\|^2] = \|\boldsymbol{\mu}_D\|^2 + \operatorname{tr}\boldsymbol{\Sigma}_D. \tag{138}$$

Comparing Eq. 137 and Eq. 138 gives Eq. 130.

Under $\boldsymbol{\mu}_F = \boldsymbol{\mu}_D$ and $\boldsymbol{\Sigma}_F = \boldsymbol{\Sigma}_D$, the mean terms cancel, and

$$\begin{aligned}\mathbb{E}[\|\boldsymbol{e}_{D,0}\|^2] - \mathbb{E}[\|\boldsymbol{e}_{K,0}\|^2] &= \left[1 - \rho - \frac{1-\rho}{K}\right]\operatorname{tr}\boldsymbol{\Sigma}_F \\ &= \frac{K-1}{K}(1-\rho)\operatorname{tr}\boldsymbol{\Sigma}_F, \end{aligned} \tag{139}$$

proving Eq. 132.

If only the means are matched, the comparison reduces to

$$\left[\rho + \frac{1-\rho}{K}\right]\operatorname{tr}\boldsymbol{\Sigma}_F < \operatorname{tr}\boldsymbol{\Sigma}_D. \tag{140}$$

When $\operatorname{tr}\boldsymbol{\Sigma}_D > 0$ and $\alpha_K(\rho) > 0$, division by these positive quantities gives Eq. 134. The benefit of FM over direct supervision does not require each local FM prediction to be more accurate than the direct endpoint prediction. FM can tolerate larger marginal local variance if the local errors are sufficiently weakly correlated across flow time. Under matched statistics, the entire FM advantage is exactly the variance canceled by multi-step integration. □

THEOREM 7: PHYSICS REGULARIZATION CANNOT INCREASE ANALYTICAL FM ENDPOINT ERROR

Consider finite-dimensional real signals with Euclidean norm and a fixed linear operator $\mathcal{A}$. Let $\lambda \geq 0$, $K \geq 1$, and define

$$\boldsymbol{S}_\lambda := (\boldsymbol{I} + \lambda\mathcal{A}^*\mathcal{A})^{-1}. \tag{141}$$

For any prescribed sequence of pre-physics local FM velocity errors $\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_{K-1}$, define

$$\boldsymbol{e}_{K,0} := \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{\xi}_k \tag{142}$$

and the analytically physics-regularized endpoint error

$$\boldsymbol{e}_{K,\lambda} := \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{S}_\lambda\boldsymbol{\xi}_k. \tag{143}$$

Then

$$\boldsymbol{e}_{K,\lambda} = \boldsymbol{S}_\lambda\boldsymbol{e}_{K,0}, \tag{144}$$

and, pathwise,

$$\|\boldsymbol{e}_{K,\lambda}\| \leq \|\boldsymbol{e}_{K,0}\|. \tag{145}$$

Consequently,

$$\mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] \leq \mathbb{E}[\|\boldsymbol{e}_{K,0}\|^2]. \tag{146}$$

For $\lambda > 0$, the inequality is strict whenever $\boldsymbol{e}_{K,0}$ has nonzero energy outside $\ker\mathcal{A}$ with positive probability.

PROOF OF THEOREM 7

Since $\boldsymbol{S}_\lambda$ is linear and independent of k,

$$\begin{aligned}\boldsymbol{e}_{K,\lambda} &= \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{S}_\lambda \boldsymbol{\xi}_k \\ &= \boldsymbol{S}_\lambda \left(\frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{\xi}_k\right) \\ &= \boldsymbol{S}_\lambda \boldsymbol{e}_{K,0},\end{aligned} \tag{147}$$

which proves Eq. 144.

Let

$$\mathcal{A} = \boldsymbol{U}\,\mathrm{diag}(\sigma_i)\boldsymbol{V}^*. \tag{148}$$

Then

$$\boldsymbol{S}_\lambda = \boldsymbol{V}\,\mathrm{diag}\left(\frac{1}{1+\lambda\sigma_i^2}\right)\boldsymbol{V}^*. \tag{149}$$

Every singular value of $\boldsymbol{S}_\lambda$ therefore lies in $(0, 1]$. Hence

$$\|\boldsymbol{S}_\lambda \boldsymbol{y}\| \le \|\boldsymbol{y}\| \tag{150}$$

for every vector $\boldsymbol{y}$. Taking $\boldsymbol{y} = \boldsymbol{e}_{K,0}$ yields Eq. 145. Squaring and taking expectations gives Eq. 146.

If $\lambda > 0$, the shrinkage factor is strictly smaller than one in every singular direction with $\sigma_i > 0$, so the pathwise inequality is strict whenever $\boldsymbol{e}_{K,0}$ has a nonzero component outside $\ker \mathcal{A}$. When this occurs with positive probability, the expected squared inequality is strict as well. Under exact physics and for the same prescribed pre-physics local errors, the analytical physics penalty can therefore only reduce or preserve the FM endpoint error, since it cannot correct error lying in $\ker \mathcal{A}$, and strict improvement requires endpoint error energy in a physics-identifiable direction. This is an analytical fixed-error comparison, and separately trained networks need not share the same pre-physics error sequence. □

THEOREM 8: WHEN PHYSICS-CONSTRAINED FLOW MATCHING IMPROVES ON PHYSICS-CONSTRAINED DIRECT SUPERVISION

Consider finite-dimensional real signals with Euclidean norm and a fixed linear operator $\mathcal{A}$. Let $\lambda \ge 0$, $K \ge 1$, and

$$\boldsymbol{S}_\lambda := (\boldsymbol{I} + \lambda \mathcal{A}^*\mathcal{A})^{-1}. \tag{151}$$

For the FM model, let the pre-physics local velocity errors $\boldsymbol{\xi}_0, \ldots, \boldsymbol{\xi}_{K-1}$ have common mean

$$\boldsymbol{\mu}_F := \mathbb{E}[\boldsymbol{\xi}_k] \tag{152}$$

and common covariance

$$\boldsymbol{\Sigma}_F := \mathrm{Cov}(\boldsymbol{\xi}_k). \tag{153}$$

Define

$$\widetilde{\boldsymbol{\xi}}_k := \boldsymbol{S}_\lambda \boldsymbol{\xi}_k, \qquad \boldsymbol{\Sigma}_{F,\lambda} := \boldsymbol{S}_\lambda \boldsymbol{\Sigma}_F \boldsymbol{S}_\lambda^*. \tag{154}$$

Assume that, for $i \ne j$,

$$\mathrm{tr}\,\mathrm{Cov}(\widetilde{\boldsymbol{\xi}}_i, \widetilde{\boldsymbol{\xi}}_j) = \rho_\lambda\,\mathrm{tr}\,\boldsymbol{\Sigma}_{F,\lambda}. \tag{155}$$

Let the FM endpoint error be

$$\boldsymbol{e}_{K,\lambda} := \frac{1}{K}\sum_{k=0}^{K-1} \widetilde{\boldsymbol{\xi}}_k. \tag{156}$$

For the direct model, let its pre-physics endpoint error

$$\boldsymbol{e}_{D,0} := \widehat{\boldsymbol{x}}_{D,0} - \boldsymbol{x}_1 \tag{157}$$

have mean and covariance

$$\boldsymbol{\mu}_D := \mathbb{E}[\boldsymbol{e}_{D,0}], \qquad \boldsymbol{\Sigma}_D := \mathrm{Cov}(\boldsymbol{e}_{D,0}). \tag{158}$$

Under the same analytical quadratic physics regularization, define

$$\boldsymbol{e}_{D,\lambda} := \boldsymbol{S}_\lambda \boldsymbol{e}_{D,0}, \qquad \boldsymbol{\Sigma}_{D,\lambda} := \boldsymbol{S}_\lambda \boldsymbol{\Sigma}_D \boldsymbol{S}_\lambda^*. \tag{159}$$

Then the physics-constrained FM model has lower expected squared endpoint error than physics-constrained direct supervision if and only if

$$\|\boldsymbol{S}_\lambda \boldsymbol{\mu}_F\|^2 + \left[\rho_\lambda + \frac{1-\rho_\lambda}{K}\right] \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda} < \|\boldsymbol{S}_\lambda \boldsymbol{\mu}_D\|^2 + \operatorname{tr} \boldsymbol{\Sigma}_{D,\lambda}. \tag{160}$$

Under matched pre-physics statistics,

$$\boldsymbol{\mu}_F = \boldsymbol{\mu}_D, \qquad \boldsymbol{\Sigma}_F = \boldsymbol{\Sigma}_D, \tag{161}$$

the direct-minus-FM expected squared endpoint error gap is

$$\mathbb{E}[\|\boldsymbol{e}_{D,\lambda}\|^2] - \mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] = \frac{K-1}{K}(1-\rho_\lambda) \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda}. \tag{162}$$

Thus, for $K > 1$ and $\operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda} > 0$, the physics-constrained FM model has a strict advantage under matched statistics exactly when $\rho_\lambda < 1$.

PROOF OF THEOREM 8

For the FM model,

$$\mathbb{E}[\boldsymbol{e}_{K,\lambda}] = \boldsymbol{S}_\lambda \boldsymbol{\mu}_F. \tag{163}$$

Using the common marginal covariance and normalized cross-flow covariance trace,

$$\begin{aligned} \operatorname{tr} \operatorname{Cov}(\boldsymbol{e}_{K,\lambda}) &= \frac{1}{K^2}\left[K \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda} + K(K-1)\rho_\lambda \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda}\right] \\ &= \left[\rho_\lambda + \frac{1-\rho_\lambda}{K}\right] \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda}. \end{aligned} \tag{164}$$

Therefore

$$\mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] = \|\boldsymbol{S}_\lambda \boldsymbol{\mu}_F\|^2 + \left[\rho_\lambda + \frac{1-\rho_\lambda}{K}\right] \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda}. \tag{165}$$

For the direct model,

$$\boldsymbol{e}_{D,\lambda} = \boldsymbol{S}_\lambda \boldsymbol{e}_{D,0}, \tag{166}$$

so

$$\mathbb{E}[\boldsymbol{e}_{D,\lambda}] = \boldsymbol{S}_\lambda \boldsymbol{\mu}_D, \qquad \operatorname{Cov}(\boldsymbol{e}_{D,\lambda}) = \boldsymbol{S}_\lambda \boldsymbol{\Sigma}_D \boldsymbol{S}_\lambda^* = \boldsymbol{\Sigma}_{D,\lambda}. \tag{167}$$

Hence

$$\mathbb{E}[\|\boldsymbol{e}_{D,\lambda}\|^2] = \|\boldsymbol{S}_\lambda \boldsymbol{\mu}_D\|^2 + \operatorname{tr} \boldsymbol{\Sigma}_{D,\lambda}. \tag{168}$$

Comparing Eq. 165 with Eq. 168 proves Eq. 160.

Under matched pre-physics statistics,

$$\boldsymbol{S}_\lambda \boldsymbol{\mu}_F = \boldsymbol{S}_\lambda \boldsymbol{\mu}_D, \tag{169}$$

and

$$\boldsymbol{\Sigma}_{F,\lambda} = \boldsymbol{S}_\lambda \boldsymbol{\Sigma}_F \boldsymbol{S}_\lambda^* = \boldsymbol{S}_\lambda \boldsymbol{\Sigma}_D \boldsymbol{S}_\lambda^* = \boldsymbol{\Sigma}_{D,\lambda}. \tag{170}$$

The mean terms therefore cancel, leaving

$$\begin{aligned} \mathbb{E}[\|\boldsymbol{e}_{D,\lambda}\|^2] - \mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] &= \left[1 - \rho_\lambda - \frac{1-\rho_\lambda}{K}\right] \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda} \\ &= \frac{K-1}{K}(1-\rho_\lambda) \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda}, \end{aligned} \tag{171}$$

which proves Eq. 162. Physics-constrained FM combines two distinct mechanisms. $\boldsymbol{S}_\lambda$ suppresses error in physics-identifiable directions, while integration reduces the variance of errors that are not perfectly coherent across flow time. Relative to direct supervision with the same analytical physics penalty, the remaining advantage is therefore controlled by the post-physics local error statistics and, under matched marginals, exactly by ρ_λ. As above, comparisons between separately trained networks require the analytical error model to be checked empirically. □

PROPOSITION 9: GENERAL CROSS-FLOW COVARIANCE AND PERSISTENT-ERROR DECOMPOSITION

Consider the straight-path K-step construction of Proposition 4. Let the square-integrable pre-physics local FM errors $\boldsymbol{\xi}_k$ have common mean $\boldsymbol{\mu}_F := \mathbb{E}[\boldsymbol{\xi}_k]$, and define their post-physics local errors

$$\widetilde{\boldsymbol{\xi}}_k := \boldsymbol{S}_\lambda \boldsymbol{\xi}_k, \qquad k = 0, \dots, K-1. \tag{172}$$

Their common post-physics mean is

$$\mathbb{E}[\widetilde{\boldsymbol{\xi}}_k] = \boldsymbol{S}_\lambda \boldsymbol{\mu}_F, \tag{173}$$

No equicorrelation or common marginal covariance is assumed. Define the scalar post-physics cross-flow covariance traces

$$c_{ij,\lambda} := \operatorname{tr} \operatorname{Cov}\left(\widetilde{\boldsymbol{\xi}}_i, \widetilde{\boldsymbol{\xi}}_j\right). \tag{174}$$

Then

$$\mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] = \|\boldsymbol{S}_\lambda \boldsymbol{\mu}_F\|^2 + \frac{1}{K^2} \sum_{i=0}^{K-1} \sum_{j=0}^{K-1} c_{ij,\lambda}. \tag{175}$$

For $K > 1$, define the average local variance and off-diagonal covariance trace by

$$V_\lambda := \frac{1}{K} \sum_{k=0}^{K-1} c_{kk,\lambda}, \qquad \bar{c}_\lambda := \frac{1}{K(K-1)} \sum_{i \neq j} c_{ij,\lambda}, \tag{176}$$

Then

$$\mathbb{E}[\|\boldsymbol{e}_{K,\lambda}\|^2] = \|\boldsymbol{S}_\lambda \boldsymbol{\mu}_F\|^2 + \frac{1}{K} V_\lambda + \frac{K-1}{K} \bar{c}_\lambda. \tag{177}$$

The equicorrelation formula of Theorem 5 is recovered by setting

$$V_\lambda = \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda}, \qquad \bar{c}_\lambda = \rho_\lambda \operatorname{tr} \boldsymbol{\Sigma}_{F,\lambda}. \tag{178}$$

Furthermore, suppose the local pre-physics errors admit the decomposition

$$\boldsymbol{\xi}_k = \boldsymbol{c} + \boldsymbol{\eta}_k, \tag{179}$$

where $\boldsymbol{c}$ is a random pre-physics local velocity error component shared across flow time and $\boldsymbol{\eta}_k$ is its flow-varying remainder. Assume

$$\mathbb{E}[\boldsymbol{\eta}_k] = 0, \qquad \mathbb{E}[\langle \boldsymbol{c}, \boldsymbol{\eta}_k \rangle] = 0, \tag{180}$$

Define the persistent and flow-varying pre-physics energies

$$\mathcal{E}_c := \mathbb{E}[\|\boldsymbol{c}\|^2], \qquad \mathcal{E}_\eta := \mathbb{E}[\|\boldsymbol{\eta}_k\|^2], \tag{181}$$

where $\mathcal{E}_\eta$ is common to all steps. Let ρ_η denote the normalized cross-flow inner product of the remainders, distinct from $\rho = \rho_0$ for the full errors $\boldsymbol{\xi}_k$, and assume

$$\mathbb{E}\left[\langle \boldsymbol{\eta}_i, \boldsymbol{\eta}_j \rangle\right] = \rho_\eta \mathcal{E}_\eta, \qquad i \neq j. \tag{182}$$

Then, without physics regularization,

$$\mathbb{E}[\|\boldsymbol{e}_{K,0}\|^2] = \mathcal{E}_c + \left[\rho_\eta + \frac{1-\rho_\eta}{K}\right] \mathcal{E}_\eta. \tag{183}$$

Thus integration leaves the persistent component $\boldsymbol{c}$ unchanged and acts only on the flow-time-dependent component through averaging.

After physics regularization,

$$\boldsymbol{e}_{K,\lambda} = \boldsymbol{S}_\lambda \boldsymbol{c} + \frac{1}{K} \sum_{k=0}^{K-1} \boldsymbol{S}_\lambda \boldsymbol{\eta}_k. \tag{184}$$

Therefore the physics term and flow integration act on different structures in the local velocity error: $\boldsymbol{S}_\lambda$ contracts components according to the physical operator, while integration averages the flow-varying components.

PROOF OF PROPOSITION 9

By Proposition 4,

$$\boldsymbol{e}_{K,\lambda} = \frac{1}{K}\sum_{k=0}^{K-1} \widetilde{\boldsymbol{\xi}}_k. \tag{185}$$

Since the local means are common,

$$\mathbb{E}[\boldsymbol{e}_{K,\lambda}] = \boldsymbol{S}_\lambda \boldsymbol{\mu}_F. \tag{186}$$

The covariance of the average is

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{e}_{K,\lambda}) &= \mathrm{Cov}\left(\frac{1}{K}\sum_{k=0}^{K-1} \widetilde{\boldsymbol{\xi}}_k\right) \\ &= \frac{1}{K^2}\sum_{i=0}^{K-1}\sum_{j=0}^{K-1} \mathrm{Cov}\left(\widetilde{\boldsymbol{\xi}}_i, \widetilde{\boldsymbol{\xi}}_j\right).\end{aligned} \tag{187}$$

Taking traces gives

$$\mathrm{tr}\,\mathrm{Cov}(\boldsymbol{e}_{K,\lambda}) = \frac{1}{K^2}\sum_{i=0}^{K-1}\sum_{j=0}^{K-1} c_{ij,\lambda}. \tag{188}$$

Using the bias-variance identity for a square-integrable random vector $\mathbf{y}$,

$$\mathbb{E}[\|\mathbf{y}\|^2] = \|\mathbb{E}[\mathbf{y}]\|^2 + \mathrm{tr}\,\mathrm{Cov}(\mathbf{y}), \tag{189}$$

together with Eq. 186 and Eq. 188, yields Eq. 175.

To obtain the averaged form, separate the diagonal and off-diagonal terms:

$$\begin{aligned}\frac{1}{K^2}\sum_{i,j} c_{ij,\lambda} &= \frac{1}{K^2}\sum_{k=0}^{K-1} c_{kk,\lambda} + \frac{1}{K^2}\sum_{i\neq j} c_{ij,\lambda} \\ &= \frac{1}{K}V_\lambda + \frac{K-1}{K}\overline{c}_\lambda,\end{aligned} \tag{190}$$

which proves Eq. 177. Under the equicorrelation model,

$$c_{kk,\lambda} = \mathrm{tr}\,\boldsymbol{\Sigma}_{F,\lambda}, \qquad c_{ij,\lambda} = \rho_\lambda\,\mathrm{tr}\,\boldsymbol{\Sigma}_{F,\lambda} \quad (i \neq j),$$

and therefore

$$V_\lambda = \mathrm{tr}\,\boldsymbol{\Sigma}_{F,\lambda}, \qquad \overline{c}_\lambda = \rho_\lambda\,\mathrm{tr}\,\boldsymbol{\Sigma}_{F,\lambda}.$$

Substitution recovers the expression in Theorem 5.

We next consider the persistent-error decomposition Eq. 179. Without physics regularization,

$$\begin{aligned}\boldsymbol{e}_{K,0} &= \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{\xi}_k \\ &= \frac{1}{K}\sum_{k=0}^{K-1} (\boldsymbol{c} + \boldsymbol{\eta}_k) \\ &= \boldsymbol{c} + \frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{\eta}_k.\end{aligned} \tag{191}$$

Taking squared norms,

$$\begin{aligned}\|\boldsymbol{e}_{K,0}\|^2 = \|\boldsymbol{c}\|^2 &+ \left\|\frac{1}{K}\sum_{k=0}^{K-1} \boldsymbol{\eta}_k\right\|^2 \\ &+ \frac{2}{K}\sum_{k=0}^{K-1} \langle \boldsymbol{c}, \boldsymbol{\eta}_k\rangle.\end{aligned} \tag{192}$$

Taking expectations and using $\mathbb{E}[\langle \boldsymbol{c}, \boldsymbol{\eta}_k \rangle] = 0$,

$$\mathbb{E}[\|\boldsymbol{e}_{K,0}\|^2] = \mathcal{E}_c + \mathbb{E}\left[\left\|\frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{\eta}_k\right\|^2\right]. \tag{193}$$

The remaining term expands as

$$\begin{aligned}\mathbb{E}\left[\left\|\frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{\eta}_k\right\|^2\right] &= \frac{1}{K^2}\sum_{i=0}^{K-1}\sum_{j=0}^{K-1}\mathbb{E}\left[\langle \boldsymbol{\eta}_i, \boldsymbol{\eta}_j \rangle\right] \\ &= \frac{1}{K^2}\left[K\mathcal{E}_\eta + K(K-1)\rho_\eta \mathcal{E}_\eta\right] \\ &= \left[\rho_\eta + \frac{1-\rho_\eta}{K}\right]\mathcal{E}_\eta. \end{aligned} \tag{194}$$

Substitution into Eq. 193 proves Eq. 183.

The decomposition shows immediately why the persistent component is not reduced by integration: the term $\boldsymbol{c}$ is identical at every flow step, so averaging K copies of it returns $\boldsymbol{c}$ itself.

With physics regularization,

$$\begin{aligned}\boldsymbol{e}_{K,\lambda} &= \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{S}_\lambda \boldsymbol{\xi}_k \\ &= \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{S}_\lambda(\boldsymbol{c} + \boldsymbol{\eta}_k) \\ &= \boldsymbol{S}_\lambda \boldsymbol{c} + \frac{1}{K}\sum_{k=0}^{K-1}\boldsymbol{S}_\lambda \boldsymbol{\eta}_k, \end{aligned} \tag{195}$$

which proves Eq. 184. This identity separates the two mechanisms directly: the physics term acts through $\boldsymbol{S}_\lambda$, whereas integration acts through the average over k. □

B Implementation details

B.1 Datasets

B.1.1 Aortic Flow

We use the two-element Windkessel reference set of the virtual aortic database (Charlton et al., 2019)[1]. The dataset contains 3,125 virtual cases, each comprising one cardiac cycle of aortic inflow Q and arterial pressure P. The source is the inflow waveform and the target is pressure. Each cycle contains 1,136 native samples and has duration $T \in [0.709, 1.135]$ s. We linearly interpolate each waveform onto 256 uniformly spaced points spanning its cycle. The resulting sampling interval, $\Delta t = T/255$, is case-specific and is supplied to the physics residual.

Reference parameters are available for every case, with $R \in [0.470, 0.570]$, $C \in [2.196, 2.340]$, and $P_{\text{out}} \in [27.79, 38.61]$ mmHg. These reference values are used only to evaluate parameter estimation errors in Table 2. The physics-informed models instead estimate their parameters from the source waveform, as described in Appendix B.2.3.

The two-element Windkessel model (Westerhof et al., 2009) relates inflow and pressure through peripheral resistance R, compliance C, and outflow pressure P_{out}:

$$C\frac{\mathrm{d}P}{\mathrm{d}t} + \frac{P - P_{\text{out}}}{R} = Q. \tag{196}$$

[1] http://haemod.uk/aortic

Let $\widehat{P}_\tau^{(v)}$ denote the de-standardized pressure of the velocity-implied endpoint $\widehat{\boldsymbol{x}}_1^{(v)} = \boldsymbol{x}_0 + \boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau)$ entering Eq. 11, given in Eq. 207 of Appendix B.2.2. We write the residual pointwise in physical time, so that $\boldsymbol{r}_v(\tau)$ collects the values $r_v(\tau, t)$ over t:

$$r_v(\tau, t) = C\frac{\partial \widehat{P}_\tau^{(v)}(t)}{\partial t} + \frac{\widehat{P}_\tau^{(v)}(t) - P_{\text{out}}}{R} - Q(t). \tag{197}$$

In standardized coordinates, Eq. 196 reads $\mathcal{A}\boldsymbol{x}_1 = \mathcal{B}\boldsymbol{x}_0 + \boldsymbol{c}$ with

$$\mathcal{A} = \sigma_y\left(C\frac{\mathrm{d}}{\mathrm{d}t} + \frac{1}{R}\right), \qquad \mathcal{B} = \sigma_x \boldsymbol{I}, \qquad \boldsymbol{c} = \left[\mu_x - \frac{\mu_y - P_{\text{out}}}{R}\right]\mathbf{1}, \tag{198}$$

and $\boldsymbol{r}_v(\tau) = \mathcal{A}\widehat{\boldsymbol{x}}_1^{(v)} - \mathcal{B}\boldsymbol{x}_0 - \boldsymbol{c}$. The constant offset is absorbed into the forcing $\boldsymbol{g} = (\mathcal{B} - \mathcal{A})\boldsymbol{x}_0 + \boldsymbol{c}$, which still satisfies $\mathcal{A}\boldsymbol{d} = \boldsymbol{g}$, so Proposition 3 and Eq. 11 apply unchanged. The governing relationship is thus evaluated on the endpoint implied by the sampled local velocity. The operator $\mathcal{A}$ depends on the per-example parameter estimates and is applied in its discretized finite-difference form, which is the operator meant by $\ker \mathcal{A}$ in Section 3.5. We approximate physical-time derivatives with second-order central differences using the case-specific Δt. The residual is evaluated after discarding four samples at each boundary, giving 248 residual samples per waveform. Waveform prediction errors are evaluated over all 256 samples.

B.1.2 BIOZGRAPHENE

The graphene tattoo dataset (Kireev et al., 2022)[2] contains BioZ recordings acquired with graphene electronic tattoos and synchronized continuous reference blood pressure measured with a Finapres device. It comprises seven subjects and eight recording sessions, with one subject recorded on two days. Each session includes ten protocols: a baseline recording, seven hand-grip and cold-pressor trials, a Valsalva maneuver, and a hand-grip trial following exercise.

For each trial, we restrict the paired BioZ and blood pressure streams to their common time support. We verify that the BioZ sampling rate is within 2% of 1250 Hz and apply a zero-phase, second-order Butterworth low-pass filter with a 6 Hz cutoff to BioZ.

Because arterial inflow is not measured, we construct a proxy from the negative time derivative of filtered BioZ,

$$Q_{\text{BioZ}}(t) = -\frac{\mathrm{d}\,\text{BioZ}(t)}{\mathrm{d}t}, \tag{199}$$

using finite differences on the original timestamps. The sign reflects the decrease in impedance associated with increasing local blood volume. We use Q_{BioZ} as the source signal and pressure as the target. For the physiological constraint, we approximate the proxy as $Q_{\text{BioZ}} \approx \kappa Q$, where the unknown scale κ is absorbed into the model parameters. This is a modeling approximation for the measured dataset.

We divide the paired signals into non-overlapping 4 s windows and resample each window to 256 uniform samples. The implementation uses a sampling interval of $\Delta t = 1/64$ s. Windows containing non-finite values are excluded. Of the remaining $5{,}535$ windows assessed for signal quality, $4{,}764$ are labeled acceptable and retained for the experiments.

Dividing Eq. 196 by C and substituting the scaled flow proxy yields the reduced relationship

$$\frac{\mathrm{d}P}{\mathrm{d}t} + k_1(P - P_{\text{out}}) - k_2 Q_{\text{BioZ}} \approx 0, \qquad k_1 = \frac{1}{RC}, \qquad k_2 = \frac{1}{\kappa C}. \tag{200}$$

The effective coefficient k_2 absorbs the unknown proxy scale. The corresponding velocity residual is

$$r_v(\tau, t) = \frac{\partial \widehat{P}_\tau^{(v)}(t)}{\partial t} + k_1\big(\widehat{P}_\tau^{(v)}(t) - P_{\text{out}}\big) - k_2 Q_{\text{BioZ}}(t). \tag{201}$$

As for Aortic Flow, this has the form $\boldsymbol{r}_v(\tau) = \mathcal{A}\widehat{\boldsymbol{x}}_1^{(v)} - \mathcal{B}\boldsymbol{x}_0 - \boldsymbol{c}$, now with $\mathcal{A} = \sigma_y(\mathrm{d}/\mathrm{d}t + k_1)$, $\mathcal{B} = k_2\sigma_x \boldsymbol{I}$, and $\boldsymbol{c} = [k_2\mu_x - k_1(\mu_y - P_{\text{out}})]\mathbf{1}$. The parameter vector $(k_1, k_2, P_{\text{out}})$ is estimated separately for each window and is constant within that window. Derivatives are computed using central differences in the interior and one-sided differences at the boundaries, with the residual evaluated over the full window.

[2] https://physionet.org/content/bp-graphene-bioimpedance/1.0.0/

B.2 TRAINING DETAILS

The experiments use a common training and evaluation protocol, with the model objectives and architectural adaptations described below. Each method minimizes

$$\mathcal{L} = \mathcal{L}_{\mathrm{data}} + \lambda \mathcal{L}_{\mathrm{phys}}, \tag{202}$$

where $\mathcal{L}_{\mathrm{data}}$ is $\mathcal{L}_{\mathrm{FM}}$ in Eq. 2 for flow-based methods and target mean squared error for direct prediction methods. For PEFM, $\mathcal{L}_{\mathrm{phys}}$ is $\mathcal{L}_{\mathrm{VF}}$ in Eq. 8 with the residuals below; for WPINN and PBFM, it is the mean squared endpoint or unrolled-endpoint residual described in Appendix B.2.1. The physics term is included only for methods that use a residual penalty, with $\lambda = 1$.

We use AdamW with learning rate 5×10^{-4}, weight decay 0.01, a cosine learning rate schedule, batch size 32, and gradient clipping at a global norm of 1.0. Models are trained for at most 200 epochs, with early stopping after 50 epochs without improvement in validation RMSE. The checkpoint achieving the lowest validation RMSE is selected for test evaluation. Aortic Flow results are obtained from one training run with a fixed random seed, while BioZGraphene results are aggregated over five seeds.

B.2.1 MODEL ARCHITECTURES AND BASELINE ADAPTATIONS

The common prediction backbone is a one-dimensional U-Net (Ronneberger et al., 2015) with three downsampling blocks, three upsampling blocks, channel widths of 32, 64, and 64, a middle block, Fourier flow-time embeddings, and no dropout. For flow-based models, the single input channel contains the current state $\boldsymbol{z}_\tau$, and the output is a single-channel velocity field $\boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau)$. Direct prediction baselines use the same backbone with fixed flow time and the standardized source as input, and their output represents the target waveform. PIFM-GN introduces an additional U-Net branch, as described below. Methods using learned physical parameters also include the auxiliary estimator in Appendix B.2.3.

All flow-based methods use the straight interpolation path

$$\boldsymbol{z}_\tau = (1-\tau)\boldsymbol{x}_0 + \tau \boldsymbol{x}_1, \qquad \boldsymbol{d} = \boldsymbol{x}_1 - \boldsymbol{x}_0, \tag{203}$$

in standardized coordinates. The observed source waveform provides the initial state. For a batch of size B, stratified flow times are sampled as

$$\tau_i = \frac{i + u_i}{B}, \qquad u_i \sim \mathcal{U}(0,1), \qquad i = 0, \dots, B-1. \tag{204}$$

At inference, the learned field is integrated from $\widehat{\boldsymbol{z}}_0 = \boldsymbol{x}_0$ to $\tau = 1$ using $K = 50$ uniform Euler steps. Direct models produce their prediction in one forward pass. Inference is deterministic, and each method returns one pressure waveform per input.

The supervised and FM baselines optimize their respective data losses without a physics penalty. WPINN (Osman et al., 2026) adds the endpoint Windkessel residual to the supervised prediction objective. PBFM (Baldan et al., 2026) evaluates the physics residual at the endpoint obtained by unrolling the learned field from $(\boldsymbol{z}_\tau, \tau)$ to flow time 1. This unrolling uses n Euler steps of size $(1-\tau)/n$, with gradients propagated through all steps. A curriculum increases n from 1 to 4 in equal training phases, and each sample's residual loss is weighted by τ.

Our PIFM-GN adaptation (Song & Yu, 2026) uses a second U-Net to predict a potential ϕ, combining its physical-time derivative with a directly predicted field:

$$\boldsymbol{v}_\theta = (1-\gamma)\boldsymbol{v}_{\mathrm{direct}} + \gamma\, \partial_t \phi, \qquad \gamma = 0.25. \tag{205}$$

This method uses no explicit physics residual loss. PEFM applies Eq. 197 or Eq. 201 to the local velocity at the sampled state and flow time.

B.2.2 DATA SPLITS AND NORMALIZATION

Aortic Flow cases are randomly divided into 70% training, 15% validation, and 15% test sets. BioZGraphene is partitioned by subject, with three subjects used for training, two for validation, and two for testing. No subject contributes windows to more than one split. The Aortic Flow split uses one

fixed random seed. For BioZGraphene, the seed also determines the subject assignment, so each of the five seeds uses a different subject split.

Source and target waveforms are standardized separately using global scalar means and standard deviations computed exclusively from the training set. Writing $X = Q$ for Aortic Flow and $X = Q_{\mathrm{BioZ}}$ for BioZGraphene, the network coordinates are

$$\boldsymbol{x}_0 = \frac{X - \mu_x}{\sigma_x}, \qquad \boldsymbol{x}_1 = \frac{P - \mu_y}{\sigma_y}. \tag{206}$$

The interpolation path, displacement target, and learned velocity are all defined in these standardized coordinates. Physics residuals are evaluated in physical coordinates by reconstructing

$$\widehat{P}_\tau^{(v)} = \mu_y \mathbf{1} + \sigma_y \widehat{\boldsymbol{x}}_1^{(v)}, \qquad X = \mu_x \mathbf{1} + \sigma_x \boldsymbol{x}_0, \tag{207}$$

with $\widehat{\boldsymbol{x}}_1^{(v)} = \boldsymbol{x}_0 + \boldsymbol{v}_\theta(\boldsymbol{z}_\tau, \tau)$. The final prediction is $\widehat{P} = \mu_y \mathbf{1} + \sigma_y \widehat{\boldsymbol{z}}_K$ for flow-based methods, where $\widehat{\boldsymbol{z}}_K$ is the K-step Euler endpoint of Section 3.5, and the de-standardized network output for direct methods.

Before forming the physics loss, residual samples are divided by σ_x and then squared and averaged. This specifies the numerical scaling of the regularizer for both datasets.

RMSE and MAE are evaluated on de-standardized pressure predictions over the full window and reported in mmHg. SBP and DBP are extracted beat by beat from the peaks and troughs of predicted and reference pressure waveforms, with errors summarized as mean $\pm$ standard deviation over test beats. Physical consistency and parameter estimation metrics are described below.

B.2.3 Physical parameter estimation

Following WPINN (Osman et al., 2026), methods using a physics residual estimate its parameters with an auxiliary network conditioned on the standardized source waveform. The estimator consists of a one-dimensional convolution with 32 channels and kernel size 7, a GELU activation, global average pooling, and a linear output layer. It predicts one parameter vector per example: (R, C, P_{out}) for Aortic Flow and $(k_1, k_2, P_{\mathrm{out}})$ for BioZGraphene. The estimator and prediction model are trained jointly using the data and physics objectives described above. Each method's training residual uses its own parameter estimates.

For evaluation, the endpoint residual is computed by substituting the final predicted pressure into the dataset-specific governing equation. This endpoint quantity is distinct from the local velocity residual used to train PEFM. The reported learned-parameter residual uses the evaluated method's own estimates. On Aortic Flow, we report its mean square in physical residual units.

To compare BioZGraphene predictions under a common physiological model, we additionally fit one shared parameter vector using all training windows. We write

$$-\frac{\mathrm{d}P}{\mathrm{d}t} = k_1 P - k_2 Q_{\mathrm{BioZ}} - k_0, \qquad k_0 = k_1 P_{\mathrm{out}}, \tag{208}$$

and regress $-\mathrm{d}P/\mathrm{d}t$ on $[P, -Q_{\mathrm{BioZ}}, -1]$ in double precision. The fitted coefficients are $(k_{1,\mathrm{LS}}, k_{2,\mathrm{LS}}, k_{0,\mathrm{LS}})$, with $P_{\mathrm{out,LS}} = k_{0,\mathrm{LS}}/k_{1,\mathrm{LS}}$. Thus, $k_{0,\mathrm{LS}}$ is the coefficient of the regressor -1.

These shared parameters are held fixed when evaluating every method's final pressure predictions. The corresponding endpoint residual is

$$r_{\mathrm{end,LS}}(t) = \frac{\mathrm{d}\widehat{P}(t)}{\mathrm{d}t} + k_{1,\mathrm{LS}} \widehat{P}(t) - k_{2,\mathrm{LS}} Q_{\mathrm{BioZ}}(t) - k_{0,\mathrm{LS}}, \tag{209}$$

whose normalized mean square is reported as the shared least-squares residual in Table 3. This evaluation removes variation in the fitted parameters across methods from the residual comparison.

B.3 Ablations

We ablate PEFM on BioZGraphene along two settings: the physics weight $\lambda \in \{0, 0.1, 0.25, 0.5, 0.75, 1\}$ used during training and the number of Euler steps $K \in \{1, 50\}$ used

at inference. Setting $\lambda = 0$ removes the physics term and reduces training to paired flow matching, and $K = 1$ corresponds to a single Euler step from the source waveform. The main experiments use $\lambda = 1$ (Appendix B.2) and $K = 50$ (Appendix B.2.1). All other settings follow Appendix B.2. Tables 5 and 6 report test RMSE for $K = 1$ and $K = 50$, respectively.

Table 5: **Ablation of PEFM on BioZGraphene with $K = 1$.** Test RMSE in mmHg, reported as mean $\pm$ standard deviation over five seeds, for each physics weight λ. Lower is better.

λ	0	0.1	0.25	0.5	0.75	1
RMSE	15.50 ± 0.73	15.78 ± 1.23	15.92 ± 1.24	15.66 ± 1.66	14.69 ± 0.87	14.47 ± 0.98

Table 6: **Ablation of PEFM on BioZGraphene with $K = 50$.** Test RMSE in mmHg, reported as mean $\pm$ standard deviation over five seeds, for each physics weight λ. Lower is better.

λ	0	0.1	0.25	0.5	0.75	1
RMSE	15.10 ± 2.20	15.73 ± 1.42	15.83 ± 1.29	15.53 ± 1.33	14.46 ± 1.59	14.20 ± 1.40

At both inference settings, the two largest physics weights give the lowest mean RMSE. With $K = 50$, $\lambda = 1$ and $\lambda = 0.75$ achieve 14.20 ± 1.40 and 14.46 ± 1.59 mmHg, respectively, compared with 15.10 ± 2.20 mmHg for $\lambda = 0$. With $K = 1$, the corresponding values are 14.47 ± 0.98, 14.69 ± 0.87, and 15.50 ± 0.73 mmHg. Smaller weights, $\lambda \in \{0.1, 0.25, 0.5\}$, do not improve on $\lambda = 0$ at either setting. Increasing K from 1 to 50 reduces mean RMSE for every λ, by 0.05 to 0.40 mmHg, which is smaller than the standard deviation across seeds.