FISHER–RAO RESIDUAL FLOW MATCHING FOR GENERATION FROM CORRUPTED DATA

Anonymous authors
Paper under double-blind review

ABSTRACT

We propose Fisher–Rao Residual Flow Matching (FRRF) for learning unconditional generators of clean signals from corrupted observations and a simulatable measurement model. FRRF alternates between two steps: Fisher–Rao feedback decides what the generated population should change, and residual flow matching learns only that change. This "select, then realize" principle requires only forward simulation, avoids per-observation reconstruction, and separates statistical feedback from generator fitting. We establish geometric reduction of observation mismatch up to statistical and computational error, with clean-law recovery under inverse stability. We demonstrate numerical performance on imaging and scientific problems, and showcase how observation feedback and residual fitting contribute to the resulting updates.

1 INTRODUCTION

In many scientific applications, the population of interest is observed through a measurement process that is better characterized than the population itself: astronomical surveys provide stellar-abundance measurements with observation-specific uncertainties (Chen & Cui, 2025), and camera noise can be modeled through sensor calibration or a fitted conditional noise distribution (Foi et al., 2008; Wei et al., 2020; Abdelhamed et al., 2019). Such knowledge makes it possible to ask a distributional question: can we learn to generate clean signals from a dataset containing only corrupted measurements? The same question arises for image priors from incomplete MRI acquisitions and particle-level distributions from detector measurements (Rozet et al., 2024; Andreassen et al., 2020).

Let X denote a clean signal from an unknown population $\mu^\star$, and let Y be its measurement. The corruption kernel $K(\cdot \mid x)$ specifies the conditional law of a measurement given x:

$$X \sim \mu^\star, \qquad Y \mid X = x \sim K(\cdot \mid x), \qquad p^\star(dy) = \int K(dy \mid x)\, \mu^\star(dx). \tag{1}$$

We observe independent $Y_1, \dots, Y_n \sim p^\star$, without paired clean examples or repeated measurements of one signal, and want a generator of new samples from the clean population.

Separating population variation from measurement error requires additional structure: even an observed Gaussian variance admits many signal–noise decompositions when both components are unknown. The examples above supply a corruption model through acquisition physics, calibration, or a fitted noise model, and we assume only a weak form of this knowledge: the kernel $K(\cdot \mid x)$ can be sampled conditionally on a candidate x, as in the black-box setting of SCSI (Modi et al., 2026). This is less than an evaluable measurement density or an analytical noise identity, neither of which our construction requires. We write $K\mu$ for the observation law induced by a clean law μ; identification further requires that distinct clean laws induce distinct observation laws (Bora et al., 2018).

Fitting a clean law through its observation predictions admits an empirical-Bayes G-modeling view: the clean law is a prior to be estimated from its marginal observations, for instance through the Kullback–Leibler (KL) criterion $F_{\mathrm{KL}}(\mu) = \mathrm{KL}(p^\star \| p_\mu)$ with $p_\mu = K\mu$, whose empirical counterpart is classical nonparametric maximum likelihood (NPMLE) when marginal densities are available (Kiefer & Wolfowitz, 1956). The criterion is convex in the law, but representing and optimizing that law in high dimensions remains challenging (Zhang et al., 2022).

Recent diffusion-based methods bring flexible neural priors to this problem through expectation-maximization (EM) (Dempster et al., 1977; Rozet et al., 2024; Bai et al., 2024; Hosseintabar et al.,

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.

2025). Generative training needs clean endpoints, and EM supplies them by completing each observation with a possible clean signal under the current model. **E-step:** given the current prior μ, draw $x_i \sim \Pi_\mu(\cdot \mid Y_i)$ from the posterior defined by $\Pi_\mu(dx \mid y)p_\mu(dy) = \mu(dx)K(dy \mid x)$; these samples represent the *posterior aggregate*

$$T\mu(dx) = \int \Pi_\mu(dx \mid y)\, p^\star(dy). \tag{2}$$

M-step: train the next generator on the reconstructed endpoints $\{x_i\}$, so that the ideal unrestricted law update is $\mu^+ = T\mu$. The aggregate is also a statistical update: under the usual finite-divergence conditions, $F_{\mathrm{KL}}(T\mu) \le F_{\mathrm{KL}}(\mu)$ (Aubin-Frankowski et al., 2022; Shen et al., 2026). Existing work typically requires training-time *inverse* inference to obtain the endpoints $x_i \sim \Pi_\mu(\cdot \mid Y_i)$ through posterior approximations or an additional conditional model, whereas the learned state we care about is the unconditional generator μ_θ itself. We therefore ask whether the next training population can be built by a route native to that state: *forward* generation from μ_θ and simulation through K.

For unconditional generation, observe that each step needs only the aggregate law $T\mu$, not the individual reconstructions. To construct it from forward measurements, assume $p^\star \ll p_\mu$, that is, p_μ dominates $p^\star$, and let $\rho_\mu = dp^\star/dp_\mu$ be the observation density ratio. Substituting $p^\star = \rho_\mu p_\mu$ into (2) and using the posterior factorization above gives

$$T\mu(dx) = \omega_\mu(x)\mu(dx), \qquad \omega_\mu(x) = \mathbb{E}_{Y\sim K(\cdot|x)}[\rho_\mu(Y)]. \tag{3}$$

The ratio exceeds one where the model underrepresents the observations, so averaging it over a candidate's possible measurements raises the weight of candidates that help explain the deficit: the posterior aggregate can be constructed by reweighting samples from the current generator, using only their forward measurements.

The multiplicative form also explains the population dynamics behind this update. The direction $(\omega_\mu - 1)\mu$ is the Fisher–Rao gradient descent direction of F_{KL} (Fan et al., 2023), and a forward-Euler step of size $\alpha \in (0, 1]$ gives

$$\mu^+ = \{1 + \alpha(\omega_\mu - 1)\}\mu = (1 - \alpha)\mu + \alpha T\mu. \tag{4}$$

The full step $\alpha = 1$ recovers EM; smaller steps interpolate toward its posterior aggregate and inherit the EM descent property by convexity of F_{KL} in the law. This gives a way to guide generator training through prescribed dynamics of its output distribution, a perspective also used by W-Flow and Drifting Models (Han et al., 2026; Deng et al., 2026).

To turn the law update into a training procedure, suppose the current generator μ_θ is a flow with velocity field v_θ (Lipman et al., 2023). We sample it and simulate the kernel to obtain clean candidates $X_i \sim \mu_\theta$ with measurements $Y_{ij} \sim K(\cdot \mid X_i)$; a balanced real/model classifier D on the measurements then estimates the ratio through its odds, $\widehat{\rho}(y) = D(y)/(1 - D(y))$ (Menon & Ong, 2016; Andreassen et al., 2020), so the damped weight of X_i is estimated by $\widehat{\omega}_{\alpha,i} = 1+\alpha\big(m^{-1}\sum_{j\le m}\widehat{\rho}(Y_{ij})-1\big)$, with the averages normalized to mean one. We then use the generated X_i as pseudo-clean endpoints in weighted flow matching (FM), $\mathcal{J}_{\widehat{\omega}_\alpha}(v_\theta) = \frac{1}{2}\mathbb{E}[\widehat{\omega}_\alpha(X)\|v_\theta - V\|^2]$, where V is the path velocity target constructed from an endpoint; reweighting this regression changes the endpoint population the next generator is asked to learn (Section 2). Algorithm 1 summarizes this simple construction.

Algorithm 1 Vanilla ratio-guided flow matching: one round

1: $X_i \sim \mu_\theta, \quad Y_{ij} \sim K(\cdot \mid X_i)$ ▷ endpoints and their measurements
2: $\widehat{\omega}_\alpha \leftarrow 1 + \alpha\{\mathrm{RatioWeights}(\{Y_\ell\}_{\mathrm{real}}, \{X_i, Y_{ij}\}) - 1\}$ ▷ E: one ratio estimate fixes the weights
3: $v^+ \leftarrow \mathrm{Fit}(v_\theta, \mathcal{J}_{\widehat{\omega}_\alpha})$ from v_θ, with $\{X_i, \widehat{\omega}_\alpha\}$ frozen ▷ M: fit the full weighted loss

This construction separates a statistical decision from its neural realization: the generator's own candidates represent possible clean-law corrections, and their measurements come from the forward model without a posterior approximation. The favorable descent statement, however, concerns exact population updates, whereas evidence estimation, finite sampling, and generator fitting all introduce error; damping reduces the weight error passed to fitting by a factor of α, but also the requested change. The challenge is therefore to select a correction whose statistical gain is strong enough to survive these errors, and to transmit it faithfully.

Choosing a strong correction. For KL, the first difficulty is dimension: fitting a high-dimensional marginal likelihood is hard even with explicit component densities (Zhang et al., 2022). The instruction $\rho_\mu = p^\star/p_\mu$ can also become very large where the current model assigns little probability, making a finite-sample correction sensitive to poorly represented observations, and estimating it adds error of its own: $D(y) = 0.9$ and $D(y) = 0.99$ assign the same class label but imply ratios of 9 and 99 (Menon & Ong, 2016). A smaller step limits their effect but does not establish whether the proposed law is a useful explanation of the observations. We therefore revisit the decision itself: assess the observations the reweighted law would produce, revise its weights using the remaining mismatch, and let the observation criterion change when likelihood-based evidence is unsuitable.

Fitting the correction, not refitting. Even a useful target can lose its meaning during generator training. At unit weights the requested population is unchanged, yet ordinary flow matching still regresses the incoming samples, so an imperfect incoming field or a noisy minibatch gradient moves the generator without an evidence-dependent request. Subtracting the unweighted loss removes that force but replaces the positive quadratic weights by the signed values $\widehat{\omega}_\alpha - 1$: wherever $\widehat{\omega}_\alpha < 1$ the objective has negative curvature, which can destabilize the fit. Our proposal therefore removes the baseline refitting force while preserving the positive regression curvature.

We introduce *Fisher–Rao Residual Flow Matching* (FRRF) to realize observation-driven law updates under these two forms of inexactness. First, feedback selects a clean-law correction from the observations it would produce, with a movement cost controlling departure from the incoming population; the backward-Euler interpretation leads to a general evidence-plus-movement decision, and MMD with quadratic movement gives a sample-based, tractable specialization. Second, residual matching trains the selected change while preserving positive curvature and a stationary zero correction when no change is requested. Our analysis follows a complete learning round into the generated law: when the candidate bank supports an accurate observation match at a cost compatible with the quadratic penalty, the generated law's expected observation discrepancy decreases geometrically, up to the resolution of the data and the prescribed computational accuracy.

1.1 Related Work

Fitting a clean generator through its corrupted observations. AmbientGAN, Neural Empirical Bayes, AmbientFlow, Sourcerer, and convMMD deconvolution learn a clean law μ_θ by driving an observation discrepancy $D(p^\star, K\mu_\theta)$ toward zero: adversarially, by marginal likelihood, by a variational bound, by sliced Wasserstein with entropy regularization, or by noise-convolved MMD (Bora et al., 2018; Vandegar et al., 2020; Kelkar et al., 2023; Vetter et al., 2024; Vashistha et al., 2026). All minimize the criterion directly over the generator parameters θ, a nonconvex problem whose gradient passes through the network and then through a differentiable simulator or surrogate, or an evaluable measurement likelihood. FRRF decomposes this burden: with the generated candidates frozen, it minimizes the same kind of criterion over their weights ω, a convex problem that needs only forward draws from K, and generator fitting becomes an ordinary velocity-matching problem toward the selected law $\omega\mu$. SCSI instead learns a reverse transport from observations to clean signals and fits a separate generator on the restored bank (Modi et al., 2026); FRRF keeps the unconditional generator as its learned state and reweights that generator's own samples, reconstructing no observation (Appendix A.5 compares the decisions each state supports). The EM and empirical-Bayes methods of the Introduction, which reconstruct clean samples by posterior inference (Rozet et al., 2024; Hosseintabar et al., 2025; Shen et al., 2026), and training targets built from analytical channel structure (Aali et al., 2023; Chen & Cui, 2025; Daras et al., 2024) are reviewed in Appendix D.

Reweighting, law-driven training, and shared numerical components. Multiplicative population corrections go back to Richardson–Lucy (Richardson, 1972; Lucy, 1974); OmniFold and AUSSIE extend them to simulator-defined channels by pulling classifier ratios back through paired simulation and output a reweighted simulated population (Andreassen et al., 2020; Ore & Plehn, 2026), whereas FRRF revisits the evidence under the proposed weights and realizes the weighted law in a generator. Individual components have close numerical relatives: fixed-support NPMLE optimizes mixture weights against an evaluable likelihood matrix (Zhang et al., 2022), kernel mean matching and weighted MMD quantization solve quadratic feature-mean problems (Huang et al., 2006; Belhadji et al., 2026), and our residual loss recovers implicit Tilt Matching if the regression

reference equals the incoming field (Potaptchik et al., 2026). We analyze and realize the full pipeline: how a controlled clean-law correction, selected from its observation consequences, is faithfully realized by a neural-represented law.

2 PRELIMINARIES

Probability-law descent. Let τ index the evolution of the clean law across learning rounds. For an observation criterion $F_D(\mu) = D(p^\star, K\mu)$, a relative perturbation $(1 + \epsilon h)\mu$ with $\mathbb{E}_\mu h = 0$ preserves mass, and the negative first variation s_μ is defined by

$$\frac{d}{d\epsilon} F_D((1 + \epsilon h)\mu)\Big|_{\epsilon=0} = -\mathbb{E}_\mu[s_\mu h], \tag{5}$$

so $s_\mu(x)$ records how much adding relative mass at x improves observation fit, up to an additive constant that centering removes. Under the Fisher–Rao inner product $\langle h_1, h_2\rangle_\mu = \mathbb{E}_\mu[h_1 h_2]$, with tangent norm $\|h\|_\mu^2 = \mathbb{E}_\mu h^2$, the mass-preserving descent direction is the centered score, $\partial_\tau \mu_\tau = (s_{\mu_\tau} - \mathbb{E}_{\mu_\tau} s_{\mu_\tau})\mu_\tau$, with dissipation $\frac{d}{d\tau} F_D(\mu_\tau) = -\operatorname{Var}_{\mu_\tau}(s_{\mu_\tau})$ under standard regularity. For observation KL, $s_\mu = \omega_\mu$ with $\mathbb{E}_\mu \omega_\mu = 1$, which recovers the EM direction and its forward step (4) (Fan et al., 2023; Aubin-Frankowski et al., 2022).

Flow matching for a weighted endpoint law. Let $t \in [0, 1]$ parametrize one generative path. To fit $\omega\mu$ from samples of μ, draw a clean endpoint $X \sim \mu$, independent source noise ϵ, and auxiliary randomness ξ, and let $Z_t = I_t(X, \epsilon, \xi)$ interpolate from ϵ at $t = 0$ to X at $t = 1$ with path velocity $V = \partial_t I_t$; the straight path $I_t = (1 - t)\epsilon + tX$ is the default, and channel-conditioned intermediate states may be included (Appendix A.5). Write $S = (Z_t, t)$ with t uniform, and call one draw of (X, ϵ, ξ, t) a training tuple. Because ω depends only on X, weighting the endpoint loss $\mathcal{J}_\omega(v) = \frac{1}{2}\mathbb{E}[\omega(X)\|v(S) - V\|^2]$ leaves the source law unchanged and moves the conditional minimizer from v_0 to v_ω:

$$v_0(S) = \mathbb{E}[V \mid S], \qquad m_\omega(S) = \mathbb{E}[\omega(X) \mid S], \qquad v_\omega(S) = \frac{\mathbb{E}[\omega(X)V \mid S]}{m_\omega(S)}. \tag{6}$$

Integrating v_0 or v_ω from source to endpoint generates μ or $\omega\mu$ (Lipman et al., 2023); the field that currently generates μ need not equal v_0 for this interpolation, a distinction Section 3.2 builds on.

3 FISHER–RAO RESIDUAL FLOW MATCHING

Our core idea is to guide generator fitting by a clean-law correction that is itself checked against the observations: first refine what change to request, then train the generator to transmit that change. In this EM-like framework the E-step selects a training law and the M-step realizes its correction. At round k the incoming velocity field v_k generates a clean law μ_k; we describe one round at the population level, dropping the index ($v = v_k$, $\mu = \mu_k$). A nonnegative weight ω with $\mathbb{E}_\mu \omega = 1$ specifies the requested clean law $\omega\mu$, whose observations follow $K(\omega\mu)$. The incoming law stays fixed while the weights are chosen, and the weights stay fixed while the next generator is fitted.

3.1 SELECT THE CORRECTION THROUGH OBSERVATION FEEDBACK

Check the proposed change before fitting, as a backward Fisher–Rao step. The forward step (4) computes its instruction at the incoming law and passes it to fitting; when the instruction is estimated, damping shrinks it but does not check its consequence. A proposed change $\widehat{\omega}$ can instead be examined directly: we simulate forward measurements from $K(\widehat{\omega}\mu)$, measure how much of the discrepancy $D(p^\star, K(\widehat{\omega}\mu))$ remains, and refine $\widehat{\omega}$ *before* any generator is fitted. This check is the backward Euler step of the KL-driven Fisher–Rao flow, which evaluates the instruction at the selected law μ^+,

$$\mu^+ - \mu = \eta\{T\mu^+ - \mu^+\} \quad \Longleftrightarrow \quad \mu^+ = \frac{\mu + \eta T\mu^+}{1 + \eta}, \tag{7}$$

so the new law must agree with evidence computed under its own proposed state, and the step size $\eta > 0$ controls how far it moves.

Choose the weights by a generalized proximal program. An implicit step is also a minimization that balances the objective against a cost for moving away from the current point, the reading that turns implicit Euler into a proximal or minimizing-movement scheme (Jordan et al., 1998; Parikh & Boyd, 2014): with only normalization active, μ^+ minimizes $F_{\mathrm{KL}}(\omega\mu) + \eta^{-1}\,\mathrm{KL}(\mu\|\omega\mu)$ over ω (Proposition 1). Two choices thus separate: how to assess observation fit, and how to price the requested redistribution. We consider the general decision

$$\omega^\star \in \arg\min_{\omega\in\mathcal{A}} J_\eta(\omega), \qquad J_\eta(\omega) = \underbrace{F_D(\omega\mu)}_{\text{observation fit}} + \eta^{-1} \underbrace{R_\mu(\omega)}_{\text{movement cost}}, \tag{8}$$

where $\mathcal{A}$ is a closed convex set of normalized nonnegative weights containing $\mathbf{1}$, and $R_\mu \geq 0$ with $R_\mu(\mathbf{1}) = 0$. Any criterion with a computable first variation may serve as D: observation KL, whose score is the channel-averaged ratio ω_μ and therefore needs ratio estimation; other f-divergences, estimated variationally or through a classifier (Nguyen et al., 2010; Menon & Ong, 2016); kernel scores such as MMD, whose score is a channel-averaged witness; or Fisher-type divergences built from observation scores (Appendix A.5). The first three are convex in the law, so the decision is a convex program; the last is not. The movement cost R_μ controls how strongly that evidence changes the clean population.

To see how these choices determine the update, write $s_\omega = s_{\omega\mu}$ for the negative first variation (5) at the proposed law.

Proposition 1 (Feedback endpoints and selected progress)**.** *Suppose the two terms in* (8) *are convex and a minimizer exists. Then*

$$F_D(\omega^\star\mu) + \eta^{-1} R_\mu(\omega^\star) \leq F_D(\mu). \tag{9}$$

At a differentiable solution where only normalization is active,

$$R_{\mathrm{K}}(\omega) = \mathrm{KL}(\mu\|\omega\mu): \quad \omega - 1 = \eta\omega\{s_\omega - \mathbb{E}_{\omega\mu} s_\omega\}, \tag{10}$$

$$R_{\mathrm{Q}}(\omega) = \tfrac{1}{2}\mathbb{E}_\mu(\omega-1)^2: \quad \omega - 1 = \eta\{s_\omega - \mathbb{E}_\mu s_\omega\}. \tag{11}$$

Movement costs thus yield different discretizations of the same local Fisher–Rao descent geometry. The first endpoint is backward Euler for the Fisher–Rao flow of F_D; observation KL recovers (7). The quadratic cost $R_{\mathrm{Q}}(\omega) = \frac{1}{2}\chi^2(\omega\mu\|\mu)$ is half the squared Fisher–Rao tangent norm of the relative change at the incoming law, the classical proximal-point choice (Parikh & Boyd, 2014) and the weight geometry of weighted MMD quantization and kernel Fisher–Rao flows (Belhadji et al., 2026; Zhu & Mielke, 2024); it agrees with reverse KL to second order around $\omega = 1$, and it makes the decision $1/\eta$-strongly convex, so a smaller η limits sensitivity to errors in the evidence (Appendix B.4.2).

MMD with quadratic movement. We specialize D to MMD (Gretton et al., 2012). Three properties make it a desirable choice: (1) its first variation is a witness function computed from samples, so the feedback loop estimates no density ratio and trains no critic; (2) it is quadratic in the observation law, so (8) becomes a convex quadratic program whose first-order gap certifies the returned weights (Appendix B.4.2); and (3) its resolution is set by the kernel, which allows a domain-aware observation feature map, such as a pretrained perceptual embedding for image tasks (Szegedy et al., 2016; Radford et al., 2021) (Appendix A.5). Observation KL keeps convexity but, as introduced earlier, loses the other two.

Formally, let k be a bounded positive-definite kernel on observations with reproducing kernel Hilbert space $(\mathcal{H}, \|\cdot\|_{\mathcal{H}})$ and feature map $\phi(y) = \mathsf{k}(y,\cdot)$ (Sriperumbudur et al., 2010). The MMD criterion is

$$F_{\mathrm{M}}(\nu) = \tfrac{1}{2}\,\mathrm{MMD}^2_{\mathsf{k}}(p^\star, K\nu) = \tfrac{1}{2}\|m_{p^\star} - m_{K\nu}\|^2_{\mathcal{H}}, \qquad m_Q = \mathbb{E}_{Y\sim Q}\phi(Y),$$

where m_Q is the mean observation feature of a law Q, so that $m_{K\nu} = \mathbb{E}_{X\sim\nu}\mathbb{E}_{Y\sim K(\cdot|X)}\phi(Y)$. Let $\psi(x) = \mathbb{E}_{Y\sim K(\cdot|x)}\phi(Y)$ represent a clean candidate x through its expected observable feature; then $m_{K(\omega\mu)} = \mathbb{E}_\mu[\omega\psi]$ is linear in ω. Writing $u = \omega - 1$ for the relative change, $r = m_{p^\star} - \mathbb{E}_\mu\psi$ for the incoming mismatch, and $Au = \mathbb{E}_\mu[u\psi]$ for the observation change caused by the redistribution, the decision with R_{Q} reads

$$\min_{u:\,1+u\in\mathcal{A}} \tfrac{1}{2}\|r - Au\|^2_{\mathcal{H}} + \tfrac{1}{2\eta}\|u\|^2_\mu. \tag{12}$$

Algorithm 2 Fisher–Rao Residual Flow Matching: one round

1: $X_i \sim \mu_k$, $\lambda_k \leftarrow N^{-1}\sum_i \delta_{X_i}$; simulate $Y_{ij} \sim K(\cdot \mid X_i)$, $j \le m$, and form $\widehat{F}_D$
2: $\widehat{\omega} \leftarrow \arg\min_{\omega \in \mathcal{A}_k} \{\widehat{F}_D(\omega\lambda_k) + (2\eta_k)^{-1}\|\omega - \mathbf{1}\|^2_{\lambda_k}\}$ ▷ E-step: penalized decision
3: $\widetilde{v} \leftarrow v_k$, or fit $\frac{1}{2}\mathbb{E}\|\widetilde{v}(S) - V\|^2$ on λ_k; freeze $\widetilde{v}, \widehat{\omega}$
4: $d \leftarrow \arg\min_d \mathcal{L}_{\widehat{\omega}}(d)$, initialized at $d = 0$ ▷ M-step: residual fit
5: **return** $v_{k+1} = v_k + d$

Differentiating the observation-fit term with respect to u gives the scoring rule $s_\omega(x) = \langle r - Au, \psi(x)\rangle_{\mathcal{H}}$: a clean candidate x is favored when its correspondence $\psi(x)$ aligns with the remaining mismatch $r - Au$. Changing u changes $r - Au$, and with it which candidates the evidence favors. Appendix B.4.2 quantifies the decision's response to observation error; Appendix A gives the scoring-rule interpretation, the conditional-observation variant, and observation guards.

3.2 Fit the correction, not the baseline

With ω fixed by the E-step, we train the generator to track $\mu \mapsto \omega\mu$. Write the student as $v_\theta = v + d$, with v frozen and $d = 0$ initially. Expectations use the current unweighted tuples (X, S, V); we suppress the arguments S and X below.

Weighted FM retains a self-refitting force. Expanding $\mathcal{J}_\omega(f) = \frac{1}{2}\mathbb{E}[\omega\|f - V\|^2]$ gives

$$\mathcal{J}_\omega(v + d) - \mathcal{J}_\omega(v) = \underbrace{\tfrac{1}{2}\mathbb{E}[\omega\|d\|^2]}_{\text{curvature}} - \underbrace{\mathbb{E}[(\omega - 1)(V - v)^\top d]}_{\text{weight-dependent correction}} - \underbrace{\mathbb{E}[(V - v)^\top d]}_{\text{baseline refit}}. \tag{13}$$

The last term is independent of ω: even at unit weights, its population force is $v_0 - v$, where v_0 is the conditional minimizer of ordinary FM as defined in (6). When $v = v_0$, this force vanishes in expectation but generally not on individual minibatches.

Residual FM isolates the requested change. We remove the baseline term and retain the positive quadratic. With a frozen regression reference $\widetilde{v}$ (by default, $\widetilde{v} = v$), we train

$$\mathcal{L}_\omega(d) = \mathbb{E}\left[\tfrac{\omega}{2}\|d\|^2 - (\omega - 1)(V - \widetilde{v})^\top d\right], \qquad v^+ = v + d. \tag{14}$$

We call (14) the *residual flow matching* objective. It gives three properties in field space. *(i) No request, no update:* at $\omega \equiv 1$, each sample loss is $\frac{1}{2}\|d\|^2$, with zero gradient at $d = 0$ for any reference. *(ii) Positive curvature:* for $\omega \ge \ell > 0$ and any field perturbation h, $D^2\mathcal{L}_\omega(d)[h, h] = \mathbb{E}[\omega\|h\|^2] \ge \ell\,\mathbb{E}\|h\|^2$. *(iii) Same ideal target:* with $\widetilde{v} = v$, the loss is equivalent to implicit Tilt Matching with unit control variate, up to scale and a d-independent term (Potaptchik et al., 2026, Eq. 26). If $v = v_0$, its unrestricted minimizer gives $v + d^\star = v_\omega$, the weighted-FM target in (6).

In general $v \ne v_0$: v_0 is the field ordinary FM would fit to the current samples, whereas v is the network trained in the previous round. The gap $v - v_0$ then enters the trained field; we quantify this discrepancy in Appendix B.3. In practice, v_0 can be estimated by ordinary FM against V on the current tuples and supplied as $\widetilde{v}$.

Finite realization of a learning round. In computation the incoming law is replaced by its generated bank $\lambda_k = N^{-1}\sum_i \delta_{X_i}$ in the weight problem and in the residual expectations. The base set $\mathcal{A}_k$ includes a positive lower bound ℓ on the weights (a floor), which the residual fit needs for a bounded objective (Appendix B.3), and may include a convex constraint on a second observation criterion (a guard); Appendix A.3 gives both. Quadratic movement retains the penalty $(2\eta_k)^{-1}\|\omega - \mathbf{1}\|^2_{\lambda_k}$; Section 4 analyzes both fixed and bounded variable step sizes. On a bank the MMD decision is a finite convex quadratic program, which standard solvers handle efficiently (Boyd & Vandenberghe, 2004). Algorithm 2 keeps the generator fixed while selecting weights, then the weights fixed while training, with $d = v_\theta - v_k$ initialized by copying the incoming network. Whether a bank represents a useful correction, how accurately feedback selects it, and how faithfully training realizes it are the three questions addressed next.

A unified generation and denoising framework. FRRF integrates three design considerations: representation, decision, and computation. The path is a representation choice and need not be linear. With an *observation-anchored* path $Z_0 = \epsilon$, $Z_{1/2} = Y$, $Z_1 = X$, straight on $[\frac{1}{2}, 1]$, the learned field is a unified decision system. Writing $\Phi_{s\to t}$ for the flow of v_0 from time s to t, unconditional generation, empirical-Bayes posterior-mean denoising, and deterministic restoration are immediate readouts:

$$\Phi_{0\to 1}(\epsilon), \qquad \mathbb{E}[X \mid Y = y] = y + \tfrac{1}{2}\, v_0(y, \tfrac{1}{2}), \qquad \Phi_{1/2\to 1}(y). \tag{15}$$

Because FRRF is agnostic to the path, the same rounds learn this decision system. Appendix A.5 develops the three design considerations further.

4 From observation-selected corrections to recovery

We follow the selected correction from the candidate bank to the generated law. For Algorithm 2, the empirical quadratic-MMD decision over the bank's feasible set $\mathcal{A}_k$ is

$$\min_{\omega\in\mathcal{A}_k} \widehat{J}_k(\omega), \qquad \widehat{J}_k(\omega) = \tfrac{1}{2}\|\widehat{m} - \widehat{M}_k(\omega)\|^2_{\mathcal{H}} + \frac{1}{2\eta_k}\|\omega - \mathbf{1}\|^2_{\lambda_k}. \tag{16}$$

Every round reuses the empirical observation mean $\widehat{m} = n^{-1}\sum_j \phi(Y_j)$ and draws N fresh candidates with m conditional simulations each. The reweighted bank has observation mean $M_k(\omega) = N^{-1}\sum_i \omega_i \psi(X_{ki})$, which the algorithm cannot evaluate; it uses the simulated estimate $\widehat{M}_k(\omega) = N^{-1}\sum_i \omega_i \widehat{\psi}_{ki}$ with $\widehat{\psi}_{ki} = m^{-1}\sum_j \phi(Y_{kij})$ in place of $\psi(X_{ki})$, and the difference is the simulation error that the penalty controls (Appendix B.1.2). Fix a data-independent kernel with $\mathsf{k}(y, y) \le 1$ and step sizes $0 < \eta_{\min} \le \eta_k \le \eta_{\max} < \infty$; a constant $\eta_k = \eta$ is allowed. Two further errors enter: the solver error $\varepsilon_{\mathrm{opt},k} = \widehat{J}_k(\widehat{\omega}_k) - \min_{\mathcal{A}_k} \widehat{J}_k$ of the returned weights, and the realization error $g_k = \|m_{K_{\mu_{k+1}}} - \widehat{M}_k(\widehat{\omega}_k)\|_{\mathcal{H}}$ of the trained generator, the distance between its observation mean and the selected mean; Appendix B.3 bounds g_k through the residual loss.

The bank must contain a useful correction. The generated candidates need not contain the clean distribution itself; we only require that their relative frequencies can be changed so that their predicted observations match the data. Specifically, suppose there are normalized nonnegative weights c_k, chosen without seeing the simulation draws, such that

$$\tfrac{1}{2}(\mathbf{1} + c_k) \in \mathcal{A}_k \text{ a.s.}, \qquad \mathbb{E}\big[\underbrace{\|M_k(c_k) - m_{p^\star}\|^2_{\mathcal{H}}}_{\text{observation mismatch}} + \underbrace{\eta_{\min}^{-1}\|c_k - \mathbf{1}\|^2_{\lambda_k}}_{\text{cost at the chosen penalty}}\big] \le \varepsilon_{\mathrm{cov}}(\eta_{\min}, N)^2. \tag{17}$$

Thus c_k would match the observation mean up to $\varepsilon_{\mathrm{cov}}$ at a cost the penalty can afford, and its halfway update is feasible in the current round. The algorithm never needs c_k; Appendix B.2 constructs it from a finite importance second moment.

Theorem 1 (Recovery at a prescribed computational accuracy)**.** *Fix a target observation-mean accuracy $\varepsilon > 0$. Assume (17) with $\varepsilon_{\mathrm{cov}}(\eta_{\min}, N) \le \varepsilon$. Suppose $\varepsilon_{\mathrm{opt},k} \le \bar{\varepsilon}_{\mathrm{opt}}$ almost surely in each round, for a deterministic $\bar{\varepsilon}_{\mathrm{opt}}$, and*

$$\varepsilon_{\mathrm{fit}} := \sqrt{2\bar{\varepsilon}_{\mathrm{opt}}} + \sup_{k<T} \|g_k\|_{L^2} \le \varepsilon,$$

where L^2 averages all randomness. With conditionally independent fresh candidates and simulations, and

$$N \ge \varepsilon^{-2}, \qquad m \ge \max\{1, \eta_{\max}\}, \tag{18}$$

the generated law after T rounds satisfies

$$\mathbb{E}\,\mathrm{KL}(\mu^\star\|\mu_T) \lesssim \frac{1}{\alpha_{\mathrm{inv}}}\Big[\underbrace{2^{-T}F_0}_{\textit{initialization}} + \underbrace{\tfrac{1}{n}}_{\textit{finite data}} + \underbrace{\varepsilon^2}_{\textit{model-side accuracy}}\Big], \qquad F_0 = \mathbb{E}F_{\mathrm{M}}(\mu_0). \tag{19}$$

The clean-law conclusion assumes $F_{\mathrm{M}}(\nu) \ge \alpha_{\mathrm{inv}}\,\mathrm{KL}(\mu^\star\|\nu)$ with $\alpha_{\mathrm{inv}} > 0$ on the output class. Without this inverse condition, the same bracket bounds $\mathbb{E}F_{\mathrm{M}}(\mu_T)$. The hidden constant is numerical, and the accuracy is paid for in computation: coverage at accuracy ε requires $\eta_{\min} \gtrsim F_0/\varepsilon^2$, hence $Nm \gtrsim \varepsilon^{-4}$ simulations per round (Appendix B.2). In particular, choose a fixed η for which $\varepsilon_{\mathrm{cov}}(\eta, N) \le n^{-1/2}$ and fit to the same accuracy; then

$$N \gtrsim n, \qquad m \ge \max\{1, \eta\}, \qquad T \gtrsim \log(en) \quad\Longrightarrow\quad \mathbb{E}\,\mathrm{KL}(\mu^\star\|\mu_T) \lesssim \frac{1}{\alpha_{\mathrm{inv}}\, n}. \tag{20}$$



Figure 1: **Recover the population.** Samples after 60 rounds, shown in the first two of 16 coordinates. Likelihood feedback uses a marginal likelihood on coordinates selected from independent noisy observations; the two MMD methods share the resulting observation-derived metric.

The bound geometrically discounts earlier implementation errors, and the n^{-1} term already accounts for reusing the same observations in every round. The proof follows one round along the chain $\mu_k \to \lambda_k \to \widehat{\omega}_k \lambda_k \to \mu_{k+1}$ (sampling, selection, realization): each step adds only its own error, and the comparison with a feasible halfway correction contracts the observation residual (Appendix B.1).

5 Experiments

5.1 Synthetic study

FRRF learns a thin ring with three unequal angular concentrations from noisy observations alone (Fig. 1). The ring is embedded beside 14 independent Gaussian coordinates and observed with Gaussian noise of standard deviation 0.35. All methods start from the same isotropic Gaussian and use 20,480 noisy observations, split between fitting, representation estimation, and validation. After 60 rounds, 90.6% of FRRF samples lie within 0.15 of the unit circle (truth: 99.7%). Two controlled comparisons isolate selection and realization.

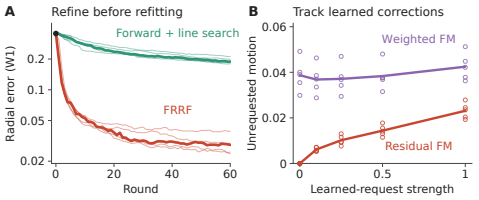

Figure 2: **Select and realize the correction. A:** Radial W_1 error (log scale) from a common Gaussian start; thin curves are five paired runs, bold curves their medians; 2,048 generator updates per round, forward step by exact line search. **B:** Motion orthogonal to the learned correction in a separate two-dimensional test; circles are five incoming-generator means, lines their average.

With the same observation criterion and generator-training budget per round, FRRF reaches the line-searched forward method's round-60 radial error within two rounds on every seed (Fig. 2A). At round 60, the mean radial W_1 errors are 0.029 and 0.192, respectively. Feedback refines the population correction before paying for another generator fit. Given one tenth of a learned correction, residual FM transmits 92% of the requested feature change with one sixth as much motion orthogonal to the request as ordinary weighted FM (Fig. 2B). At unit weights it leaves the generator exactly unchanged. Appendix C.1 gives the protocols and additional population diagnostics.

5.2 Generation from masked CIFAR-10

Setup. We learn an unconditional CIFAR-10 generator (Krizhevsky, 2009) from one 50%-masked observation per image, following Modi et al. (2026); Hosseintabar et al. (2025). We use 45,000 fitting observations and hold out 5,000 images for evaluation. Self-supervised completions initialize a 96-channel U-Net (Dhariwal & Nichol, 2021) through flow matching; clean CIFAR-10 images never enter parameter fitting. FRRF then selects weights by MMD on fixed Inception features, with a raw-pixel MMD guard and a movement radius (Appendix A.3), and realizes the correction by residual FM.

Comparisons and evaluation. We compare with FM continued on the fixed completions, SCSI (Modi et al., 2026), and DiffEM (Hosseintabar et al., 2025) on the same observations. Every row

Table 1: **Generation quality, observation fit, and training cost from the same completions** (seed 1; lower is better, nearer zero for z_{obs} and disp.). Feedback keeps improving where continued FM degrades. In clean space SCSI and DiffEM reach lower FID at comparable cost; on calibration FRRF and the anchored path are the best rows, while DiffEM is 8% under-dispersed and far outside the observation interval.

		clean space		calibration	
	A100-h	FID_{50k}	KID_{10k}	z_{obs}	disp.
Initializer	3.3	—	10.8	2.8	-3.4%
FM (best)	6.3	—	9.0	1.0	-2.3%
FM (matched)	33.3	30.0	21.3	10.7	-7.0%
FRRF	39.2	15.3	9.1	-0.5	$+\mathbf{0.2}\%$
Anchored	33.5	15.2	9.2	$\mathbf{0.3}$	-1.4%
SCSI	31.8	**13.6**	9.5	1.6	$+1.5\%$
DiffEM	39.1	13.7	**8.6**	16.7	-7.9%

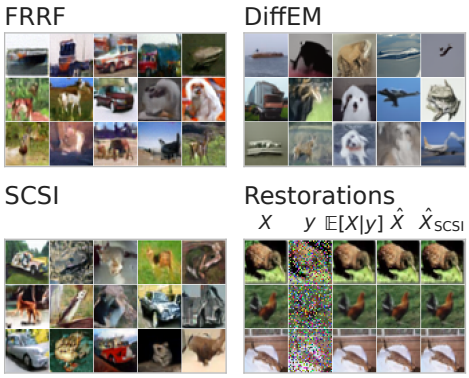

Figure 3: **Samples and readouts.** Samples of the FRRF, DiffEM, and SCSI generators of Table 1; held-out readouts of the anchored field beside SCSI's restoration: clean X, observation y, $\mathbb{E}[X \mid y]$, $\hat{X}$, and $\hat{X}_{\text{SCSI}}$ (more in Fig. 4).

starts from the same self-supervised completions: FM, FRRF, and the anchored path from 16,000-update fits on them, SCSI and DiffEM from the completions as their first restoration bank, each with its native architecture, sampler, and schedule, run to comparable A100-hours (Appendix C.2.6). Table 1 reports training cost in A100-hours including initialization, clean-space quality (FID, KID), and calibration, where z_{obs} tests re-corrupted samples against held-out observations and the dispersion error compares per-channel standard deviations with clean statistics (Appendix C.2.1). On the anchored path, restoration refreshes alone reach a lower FID (Appendix C.2.5). Fig. 3 shows samples and readouts; Appendix C.2 gives the protocols, the full metric set, and the ablations.

6 Conclusion and Limitations

FRRF learns a generator from corrupted data by refining its own population. Each round selects a Fisher–Rao correction of the current law on a bank of generated candidates, through an observation criterion evaluated on simulated corruptions alone, and realizes only that correction by residual flow matching, which is null at unit weights. When the bank covers a useful correction, one round contracts the observation mismatch, and under inverse stability T rounds recover the clean law up to the observation resolution and the fitting error. On masked CIFAR-10, feedback keeps improving the same initialization where fitting to fixed completions degrades, and one observation-anchored field generates, restores, and denoises while matching the observation law at a small cost in FID.

Limitations. As a Fisher–Rao law-update framework, FRRF generally benefits from a good initialization that proposes useful candidate locations, while poor MMD representations can make relevant discrepancies invisible. Our theory assumes sufficient generator-fitting accuracy rather than proving neural optimization or approximation complexity, and residual fitting suppresses unrequested motion without eliminating errors in estimating or realizing the selected correction. Recovery therefore depends jointly on information surviving corruption, observation resolution, candidate-bank coverage, and finite sampling/fitting accuracy; correspondingly, low observation discrepancy need not imply equally accurate fine structure or dependence. Finally, we evaluate only a small set of FRRF instantiations, demonstrating viability rather than optimal achievable performance.

AI USE STATEMENT

In this work, we used generative AI tools to suggest refinement and instantiation of our proposed framework, identify ingredients for proving mathematical claims, suggest strengthened claims and assist rigorous proof writing, implement methods, generate low-dimensional synthetic data and preprocess public datasets, provide feedback on experiments, and summarize quantitative results. We have not used generative AI tools to formulate research questions, develop core conceptual frameworks, formulate the central mathematical claims, and suggest a structure for a research paper; and assisting with translation, cleaning and reformatting dataset, and supporting qualitative and thematic data analysis are not applicable to this work. Additionally, we used generative AI tools to create scientific figures from actual experiment results, populate parts of a research paper from authors' draft writing, and identify a broader scope of relevant literature. We have reviewed all AI-assisted work. We manually drafted the initial version, checked and refined every paragraph above. We manually surveyed closely related literature, credited and cited known ingredients, and did not identify potential plagiarism of the synthesis of our core ideas to the best of our knowledge. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

We provided the link to our anonymized source code and checkpoints in the appendix. The complete proofs of the theoretical results are included in the appendix. We disclose the full data pipeline, model architecture, and implementation details in the appendix as well. Our main paper includes necessary pointers to the proofs and experimental details.

REFERENCES

Asad Aali, Marius Arvinte, Sidharth Kumar, and Jonathan I. Tamir. Solving inverse problems with Score-Based generative priors learned from noisy data. *arXiv preprint arXiv:2305.01166*, 2023. URL https://arxiv.org/abs/2305.01166.

Abdelrahman Abdelhamed, Marcus A. Brubaker, and Michael S. Brown. Noise flow: Noise modeling with conditional normalizing flows. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*, pp. 3165–3173. IEEE, 2019. doi: 10.1109/ICCV.2019.00326.

Sergios Agapiou, Omiros Papaspiliopoulos, Daniel Sanz-Alonso, and Andrew M. Stuart. Importance sampling: Intrinsic dimension and computational cost. *Statistical Science*, 32(3):405–431, 2017. doi: 10.1214/17-STS611.

Anders Andreassen, Patrick T. Komiske, Eric M. Metodiev, Benjamin Nachman, and Jesse Thaler. OmniFold: A method to simultaneously unfold all observables. *Physical Review Letters*, 124(18), 2020. ISSN 1079-7114. doi: 10.1103/physrevlett.124.182001.

Pierre-Cyril Aubin-Frankowski, Anna Korba, and Flavien Léger. Mirror descent with relative smoothness in measure spaces, with application to sinkhorn and EM. In Sanmi Koyejo, S. Mohamed, A. Agarwal, Danielle Belgrave, K. Cho, and A. Oh (eds.), *Advances in Neural Information Processing Systems 35: Annual Conference on Neural Information Processing Systems 2022, NeurIPS 2022, New Orleans, LA, USA, November 28 - December 9, 2022*, 2022. URL http://papers.nips.cc/paper_files/paper/2022/hash/6e3daaeca6be8579573f69082b2dd58b-Abstract-Conference.html.

Weimin Bai, Yifei Wang, Wenzheng Chen, and He Sun. An Expectation-Maximization algorithm for training clean diffusion models from corrupted observations. In Amir Globersons, Lester Mackey, Danielle Belgrave, Angela Fan, Ulrich Paquet, Jakub M. Tomczak, and Cheng Zhang (eds.), *Advances in Neural Information Processing Systems 37: Annual Conference on Neural Information Processing Systems 2024, NeurIPS 2024, Vancouver, BC, Canada, December 10 - 15, 2024*, 2024. URL http://papers.nips.cc/paper_files/paper/2024/hash/22912080c32a244c20bfc101033a6ac7-Abstract-Conference.html.

Joshua Batson and Loic Royer. Noise2Self: blind denoising by self-supervision. In *International Conference on Machine Learning*, 2019.

Heinz H. Bauschke and Patrick L. Combettes. The Baillon-Haddad theorem revisited. *arXiv preprint arXiv:0906.0807*, 2009. URL https://arxiv.org/abs/0906.0807.

Heinz H. Bauschke and Patrick L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer, Cham, 2 edition, 2017. doi: 10.1007/978-3-319-48311-5.

Ayoub Belhadji, Daniel Sharp, and Youssef Marzouk. Weighted quantization using MMD: From mean field to mean shift via gradient flows. In Emtiyaz Khan, Yingzhen Li, Arno Solin, and Aaditya Ramdas (eds.), *Proceedings of The 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 1045–1053. PMLR, 2026. URL https://proceedings.mlr.press/v300/belhadji26a.html.

Joe Benton, George Deligiannidis, and Arnaud Doucet. Error bounds for flow matching methods. *Transactions on Machine Learning Research*, 2024. URL https://arxiv.org/abs/2305.16860.

Mikołaj Bińkowski, Danica J. Sutherland, Michael Arbel, and Arthur Gretton. Demystifying MMD GANs. In *International Conference on Learning Representations*, 2018.

Ashish Bora, Eric Price, and Alexandros G. Dimakis. AmbientGAN: Generative models from lossy measurements. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net, 2018. URL https://openreview.net/forum?id=Hy7fDog0b.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

Gongyu Chen and Ying Cui. Score-Based diffusion modeling for nonparametric empirical Bayes in heteroscedastic Gaussian mixtures. In *Advances in Neural Information Processing Systems 38*, NeurIPS 2025, pp. 124577–124613. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2025. doi: 10.52202/085713-3753.

Badr-Eddine Chérief-Abdellatif and Pierre Alquier. Finite sample properties of parametric MMD estimation: Robustness to misspecification and dependence. *Bernoulli*, 28(1):181–213, 2022. doi: 10.3150/21-BEJ1338.

Giannis Daras, Kulin Shah, Yuval Dagan, Aravind Gollakota, Alex Dimakis, and Adam Klivans. Ambient diffusion: Learning clean distributions from corrupted data. In *Advances in Neural Information Processing Systems 36*, NeurIPS 2023, pp. 288–313. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2023. doi: 10.52202/075280-0015.

Giannis Daras, Alex Dimakis, and Constantinos Daskalakis. Consistent diffusion meets Tweedie: Training exact ambient diffusion models with noisy data. In Ruslan Salakhutdinov, Zico Kolter, Katherine A. Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*, volume 235 of *Proceedings of Machine Learning Research*, pp. 10091–10108. PMLR / OpenReview.net, 2024. URL https://proceedings.mlr.press/v235/daras24a.html.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 39(1): 1–22, 1977. ISSN 1467-9868. doi: 10.1111/j.2517-6161.1977.tb01600.x.

Mingyang Deng, He Li, Tianhong Li, Yilun Du, and Kaiming He. Generative modeling via drifting. *arXiv preprint arXiv:2602.04770*, 2026. URL https://arxiv.org/abs/2602.04770.

Andrew Dennehy, Ramchandran Muthukumar, Rebecca Willett, and Nisha Chandramoorthy. Diffusion models recover accurate mixture weights despite score function insensitivity. *arXiv preprint arXiv:2607.15485*, 2026. URL https://arxiv.org/abs/2607.15485.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat GANs on image synthesis. In *Advances in Neural Information Processing Systems*, 2021.

Zhou Fan, Leying Guan, Yandi Shen, and Yihong Wu. Gradient flows for empirical Bayes in high-dimensional linear models. *arXiv preprint arXiv:2312.12708*, 2023. URL https://arxiv.org/abs/2312.12708.

Alessandro Foi, Mejdi Trimeche, Vladimir Katkovnik, and Karen O. Egiazarian. Practical Poissonian-Gaussian noise modeling and fitting for Single-Image Raw-Data. *IEEE Trans. Image Process.*, 17(10):1737–1754, 2008. doi: 10.1109/TIP.2008.2001399.

Zhengyang Geng, Mingyang Deng, Xingjian Bai, Zico Kolter, and Kaiming He. Mean flows for one-step generative modeling. In *Advances in Neural Information Processing Systems 38*, NeurIPS 2025, pp. 84032–84054. Neural Information Processing Systems Foundation, Inc. (NeurIPS), 2025. doi: 10.52202/085713-2534.

Zhengyang Geng, Yiyang Lu, Zongze Wu, Eli Shechtman, J. Zico Kolter, and Kaiming He. Improved mean flows: On the challenges of fastforward generative models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 30467–30476, 2026. URL https://openaccess.thecvf.com/content/CVPR2026/html/Geng_Improved_Mean_Flows_On_the_Challenges_of_Fastforward_Generative_Models_CVPR_2026_paper.html.

Sulagna Ghosh, Nikolaos Ignatiadis, Frederic Koehler, and Amber Lee. Stein's unbiased risk estimate and Hyvärinen's score matching. *arXiv preprint arXiv:2502.20123*, 2025. URL https://arxiv.org/abs/2502.20123.

Arthur Gretton, Karsten M. Borgwardt, Malte J. Rasch, Bernhard Schölkopf, and Alexander Smola. A kernel Two-Sample test. *Journal of Machine Learning Research*, 13(25):723–773, 2012. URL http://jmlr.org/papers/v13/gretton12a.html.

Jiaqi Han, Puheng Li, Qiushan Guo, Renyuan Xu, Stefano Ermon, and Emmanuel J. Candès. One-Step generative modeling via Wasserstein gradient flows. *arXiv preprint arXiv:2605.11755*, 2026. URL https://arxiv.org/abs/2605.11755.

Martin Heusel, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In *Advances in Neural Information Processing Systems*, 2017.

Trung-Dung Hoang and Lisa M. Koch. Source distribution estimation by posterior averaging. *arXiv preprint arXiv:2609.02622*, 2026. URL https://arxiv.org/abs/2609.02622.

Danial Hosseintabar, Fan Chen, Giannis Daras, Antonio Torralba, and Constantinos Daskalakis. DiffEM: Learning from corrupted data with diffusion models via expectation maximization. *arXiv preprint arXiv:2510.12691*, 2025. URL https://arxiv.org/abs/2510.12691.

Jiayuan Huang, Arthur Gretton, Karsten Borgwardt, Bernhard Schölkopf, and Alex Smola. Correcting sample selection bias by unlabeled data. In B. Schölkopf, J. Platt, and T. Hoffman (eds.), *Advances in Neural Information Processing Systems*, volume 19. MIT Press, 2006. URL https://proceedings.neurips.cc/paper_files/paper/2006/file/a2186aa7c086b46ad4e8bf81e2a3a19b-Paper.pdf.

Aapo Hyvärinen. Estimation of Non-Normalized statistical models by score matching. *J. Mach. Learn. Res.*, 6:695–709, 2005. URL https://jmlr.org/papers/v6/hyvarinen05a.html.

Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker–Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998. doi: 10.1137/S0036141096303359.

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

Bahjat Kawar, Noam Elata, Tomer Michaeli, and Michael Elad. GSURE-Based diffusion model training with corrupted data. *arXiv preprint arXiv:2305.13128*, 2023. URL https://arxiv.org/abs/2305.13128.

Varun A. Kelkar, Rucha Deshpande, Arindam Banerjee, and Mark A. Anastasio. AmbientFlow: Invertible generative models from incomplete, noisy measurements. *arXiv preprint arXiv:2309.04856*, 2023. URL https://arxiv.org/abs/2309.04856.

J. Kiefer and J. Wolfowitz. Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *The Annals of Mathematical Statistics*, 27(4):887–906, 1956. ISSN 0003-4851. doi: 10.1214/aoms/1177728066.

Alex Krizhevsky. Learning multiple layers of features from tiny images. *Technical report, University of Toronto*, 2009.

Tuomas Kynkäänniemi, Tero Karras, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Improved precision and recall metric for assessing generative models. In *Advances in Neural Information Processing Systems*, 2019.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. *arXiv preprint arXiv:2210.02747*, 2023. URL https://arxiv.org/abs/2210.02747.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019.

Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. DPM-Solver++: Fast solver for guided sampling of diffusion probabilistic models. *arXiv preprint arXiv:2211.01095*, 2022. URL https://arxiv.org/abs/2211.01095.

Yiyang Lu, Susie Lu, Qiao Sun, Hanhong Zhao, Zhicheng Jiang, Xianbang Wang, Tianhong Li, Zhengyang Geng, and Kaiming He. One-step latent-free image generation with pixel mean flows. *arXiv preprint arXiv:2601.22158*, 2026. URL https://arxiv.org/abs/2601.22158.

L. B. Lucy. An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79:745, 1974. ISSN 0004-6256. doi: 10.1086/111605.

Aditya Menon and Cheng Soon Ong. Linking losses for density ratio and class-probability estimation. In Maria Florina Balcan and Kilian Q. Weinberger (eds.), *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 304–313, New York, New York, USA, 2016. PMLR. URL https://proceedings.mlr.press/v48/menon16.html.

Chirag Modi, Jiequn Han, Eric Vanden-Eijnden, and Joan Bruna. Generative modeling from Black-Box corruptions via Self-Consistent stochastic interpolants. In C. Vondrick, B. Hariharan, C. Raffel, L. Pinto, D. Yang, and A. Faust (eds.), *International Conference on Learning Representations*, volume 2026, pp. 104078–104107, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/file/aa482e630a0fe3153b7ab3d09950c71f-Paper-Conference.pdf.

Muhammad Ferjad Naeem, Seong Joon Oh, Youngjung Uh, Yunjey Choi, and Jaejun Yoo. Reliable fidelity and diversity metrics for generative models. In *International Conference on Machine Learning*, 2020.

XuanLong Nguyen, Martin J. Wainwright, and Michael I. Jordan. Estimating divergence functionals and the likelihood ratio by convex risk minimization. *IEEE Transactions on Information Theory*, 56(11):5847–5861, 2010. ISSN 1557-9654. doi: 10.1109/tit.2010.2068870.

Ayodele Ore and Tilman Plehn. Unfolding without iterations, adversaries, or surrogates. *arXiv preprint arXiv:2602.24282*, 2026. URL https://arxiv.org/abs/2602.24282.

Lorenzo Pacchiardi, Rilwan A. Adewoyin, Peter Dueben, and Ritabrata Dutta. Probabilistic forecasting with generative networks via scoring rule minimization. *Journal of Machine Learning Research*, 25(45):1–64, 2024. URL http://jmlr.org/papers/v25/23-0038.html.

Neal Parikh and Stephen Boyd. Proximal algorithms. *Foundations and Trends in Optimization*, 1 (3):127–239, 2014. doi: 10.1561/2400000003.

Chicago Y. Park, Shirin Shoushtari, Hongyu An, and Ulugbek S. Kamilov. Measurement Score-Based diffusion model. *arXiv preprint arXiv:2505.11853*, 2025. URL https://arxiv.org/abs/2505.11853.

Peter Potaptchik, Cheuk-Kit Lee, and Michael Samuel Albergo. Tilt matching for scalable sampling and Fine-Tuning. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=dQA4Gjt4KU.

Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning (ICML)*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8748–8763, 2021.

Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems*, 2007.

William Hadley Richardson. Bayesian-Based iterative method of image restoration*. *Journal of the Optical Society of America*, 62(1):55, 1972. ISSN 0030-3941. doi: 10.1364/josa.62.000055.

François Rozet, Gérôme Andry, François Lanusse, and Gilles Louppe. Learning diffusion priors from observations by expectation maximization. In Amir Globersons, Lester Mackey, Danielle Belgrave, Angela Fan, Ulrich Paquet, Jakub M. Tomczak, and Cheng Zhang (eds.), *Advances in Neural Information Processing Systems 37: Annual Conference on Neural Information Processing Systems 2024, NeurIPS 2024, Vancouver, BC, Canada, December 10 - 15, 2024*, 2024. URL http://papers.nips.cc/paper_files/paper/2024/hash/9f94298bac4668db4dc77ddb0a244301-Abstract-Conference.html.

Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training GANs. In *Advances in Neural Information Processing Systems*, 2016.

Mark Schmidt, Nicolas Le Roux, and Francis Bach. Convergence rates of inexact Proximal-Gradient methods for convex optimization. In J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger (eds.), *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc., 2011. URL https://proceedings.neurips.cc/paper_files/paper/2011/file/8f7d807e1f53eff5f9efbe5cb81090fb-Paper.pdf.

Xinwei Shen, Diana Cai, Cheng Zhang, and David M. Blei. Simulation-Based empirical Bayes. *arXiv preprint arXiv:2607.21843*, 2026. URL https://arxiv.org/abs/2607.21843.

Bharath K. Sriperumbudur, Arthur Gretton, Kenji Fukumizu, Bernhard Schölkopf, and Gert R.G. Lanckriet. Hilbert space embeddings and metrics on probability measures. *Journal of Machine Learning Research*, 11(50):1517–1561, 2010. URL http://jmlr.org/papers/v11/sriperumbudur10a.html.

Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, 2016.

Adrian Urbański, Gabriel della Maggiora, and Artur Yakimovich. Noise-Robust conditional flow matching: Generating clean samples from noisy datasets. *arXiv preprint arXiv:2608.00064*, 2026. URL https://arxiv.org/abs/2608.00064.

Maxime Vandegar, Michael Kagan, Antoine Wehenkel, and Gilles Louppe. Neural empirical Bayes: Source distribution estimation and its applications to Simulation-Based inference. *arXiv preprint arXiv:2011.05836*, 2020. URL https://arxiv.org/abs/2011.05836.

Ritwik Vashistha, Abhra Sarkar, and Arya Farahi. Nonparametric deconvolution and denoising using simulation based inference. *arXiv preprint arXiv:2606.21907*, 2026. URL https://arxiv.org/abs/2606.21907.

Julius Vetter, Guy Moss, Cornelius Schröder, Richard Gao, and Jakob H. Macke. Sourcerer: Sample-based maximum entropy source distribution estimation. In Amir Globerson, Lester Mackey, Danielle Belgrave, Angela Fan, Ulrich Paquet, Jakub M. Tomczak, and Cheng Zhang (eds.), *Advances in Neural Information Processing Systems 37: Annual Conference on Neural Information Processing Systems 2024, NeurIPS 2024, Vancouver, BC, Canada, December 10 - 15, 2024*, 2024. URL http://papers.nips.cc/paper_files/paper/2024/hash/a1caa7bf30226ac9370889d00cbd74da-Abstract-Conference.html.

Pascal Vincent. A connection between score matching and denoising autoencoders. *Neural Computation*, 23(7):1661–1674, 2011. doi: 10.1162/NECO_a_00142.

Kaixuan Wei, Ying Fu, Jiaolong Yang, and Hua Huang. A Physics-Based noise formation model for extreme Low-Light raw denoising. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pp. 2755–2764. Computer Vision Foundation / IEEE, 2020. doi: 10.1109/CVPR42600.2020.00283.

Li Kevin Wenliang and Heishiro Kanagawa. Blindness of score-based methods to isolated components and mixing proportions. *arXiv preprint arXiv:2008.10087*, 2020. URL https://arxiv.org/abs/2008.10087.

Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018.

Yangjing Zhang, Ying Cui, Bodhisattva Sen, and Kim-Chuan Toh. On efficient and scalable computation of the nonparametric maximum likelihood estimator in mixture models. *arXiv preprint arXiv:2208.07514*, 2022. URL https://arxiv.org/abs/2208.07514.

Yasi Zhang, Tianyu Chen, Zhendong Wang, Ying Nian Wu, Mingyuan Zhou, and Oscar Leong. Score distillation beyond acceleration: Generative modeling from corrupted data. *arXiv preprint arXiv:2505.13377*, 2025. URL https://arxiv.org/abs/2505.13377.

Yuchen Zhang and Jian Zhou. Inverse flow and consistency models. *arXiv preprint arXiv:2502.11333*, 2025. URL https://arxiv.org/abs/2502.11333.

Jia-Jie Zhu and Alexander Mielke. Kernel approximation of Fisher-Rao gradient flows. *arXiv preprint arXiv:2410.20622*, 2024. URL https://arxiv.org/abs/2410.20622.

A ADDITIONAL METHODOLOGY

This appendix records the observation-score, movement, and regression details underlying the numerical specialization in Section 3 (Sections A.1–A.4), and then the representation choices summarized at the close of Section 3 (Section A.5). As in the theory, λ denotes the fixed reference law, which may be the incoming population or a finite generated bank.

A.1 KERNEL SCORES AND OBSERVATION FEEDBACK

For a bounded positive-definite kernel k with Hilbert space $\mathcal{H}$, let $\phi(y) = \mathsf{k}(y, \cdot)$ and $m_Q = \mathbb{E}_Q \phi(Y)$. The MMD criterion is

$$F_{\mathrm{M}}(\nu) = \tfrac{1}{2}\,\mathrm{MMD}^2_{\mathsf{k}}(p^\star, K\nu) = \tfrac{1}{2}\|m_{p^\star} - m_{K\nu}\|^2_{\mathcal{H}}. \tag{21}$$

It can be evaluated as an expected kernel score

$$S(Q, y) = \tfrac{1}{2}\mathbb{E}_{Z,Z'\sim Q}\mathsf{k}(Z, Z') - \mathbb{E}_{Z\sim Q}\mathsf{k}(Z, y),$$

where Z, Z' are independent. Its expected excess over the true observation law is (21); a characteristic kernel makes the score strictly proper (Gretton et al., 2012; Pacchiardi et al., 2024). This gives a statistical interpretation of the sample averages used by the weight solve.

For a proposed clean law $\omega\lambda$, the witness and its clean-candidate influence are

$$f_\omega(y) = \mathbb{E}_{Y\sim p^\star}\mathsf{k}(y, Y) - \mathbb{E}_{Z\sim K(\omega\lambda)}\mathsf{k}(y, Z), \qquad s_\omega(x) = \mathbb{E}_{K(\cdot|x)} f_\omega(Y). \tag{22}$$

The score is signed: a candidate is favored when its measurements align with the remaining observation deficit. Because a change of weights changes the candidate mean, it also changes the witness. Conditional embeddings reduce this feedback to the quadratic problem in Appendix B.4.2. This has the computational form of regularized kernel mean matching with channel-averaged source representations (Huang et al., 2006; Belhadji et al., 2026). Related kernel approximations of Fisher–Rao flows place kernels in the descent geometry (Zhu & Mielke, 2024); here the kernel specifies the observation criterion.

Local Fisher–Rao geometry. For $\omega = 1 + u$ with $\mathbb{E}_\lambda u = 0$ and sufficiently small $\|u\|_\infty$,

$$\mathrm{KL}(\lambda\|(1+u)\lambda) = \tfrac{1}{2}\mathbb{E}_\lambda u^2 + O(\mathbb{E}_\lambda|u|^3), \qquad R_{\mathrm{Q}}(1+u) = \tfrac{1}{2}\mathbb{E}_\lambda u^2.$$

Both costs therefore have the same Hessian, the Fisher–Rao tangent metric at the incoming law. Their finite endpoint equations differ: reverse KL evaluates the multiplicative geometry at the candidate law, whereas the quadratic penalty retains the incoming metric. This is the distinction between (10) and (11).

A.2 RECORDED ACQUISITION SETTINGS

The observation variable may include recorded acquisition metadata. Alternatively, suppose an exogenous setting $C \sim \pi$ is independent of the clean signal and determines a channel K_C. The conditional criterion is

$$F_{\mathrm{cond}}(\nu) = \tfrac{1}{2}\mathbb{E}_C\,\mathrm{MMD}^2_{\mathsf{k}_C}(p^\star(\cdot \mid C), K_C\nu). \tag{23}$$

Its weight-dependent terms can be estimated by averaging $S_C(K_C\nu, Y)$ over observed pairs (C, Y). Model pairs share the acquisition setting and use independent conditional simulations. Thus the optimization does not require repeated real observations at the same setting (Pacchiardi et al., 2024). With $\psi_C(x) = \mathbb{E}_{K_C(\cdot|x)}\phi_C(Y)$, the model–model term is quadratic in weights and the data–model term is linear. The empirical coefficients have their own statistical errors; Appendix B.4.2 explains their interface with the decision and realization bounds.

A.3 MOVEMENT CONSTRAINTS AND OBSERVATION GUARDS

An alternative implementation constrains quadratic movement instead of penalizing it,

$$\min_{\omega\in\mathcal{W}(R)} \widehat{F}_D(\omega\lambda), \qquad \mathcal{W}(R) = \{\omega \in \mathcal{A} :\ \|\omega - \mathbf{1}\|^2_\lambda \le R\}. \tag{24}$$

For one bank, a constrained solution at which the radius is active with multiplier $\tau > 0$ is the penalized solution with $\eta = (2\tau)^{-1}$, by the Karush–Kuhn–Tucker optimality conditions. This correspondence is per decision: the multiplier depends on the round, so a fixed radius and a fixed penalty define different outer trajectories. Section 4 analyzes the penalized decision directly and requires no multiplier calibration; the fixed-η sensitivity results of Appendix B.4.2 hold that scale and feasible set fixed, and the stability of the constrained alternative is the projection statement of Appendix B.4.1.

A second convex criterion H may be included as an observation guard, $H(K(\omega\lambda)) \le H(K\lambda)$, so improvement in the primary score respects an additional observation comparison. The unit weight remains feasible. MMD guards produce quadratic constraints, giving a convex quadratically constrained problem on a finite bank. Bounds and a positive weight floor may be imposed in the same feasible set. If the guard itself is estimated from data, stability between decisions with different feasible sets needs an additional perturbation argument.

A.4 Residual curvature and the shared-reference identity

Expanding the weighted and unweighted losses around the regression reference $\widetilde{v}$ gives

$$\begin{aligned}\mathcal{L}_\omega(d) = [\mathcal{J}_\omega(\widetilde{v}+d) - \mathcal{J}_\omega(\widetilde{v})] \\ - [\mathcal{J}_1(\widetilde{v}+d) - \mathcal{J}_1(\widetilde{v})] + \tfrac{1}{2}\mathbb{E}\|d\|^2.\end{aligned} \tag{25}$$

The loss difference removes the unweighted force; the final quadratic restores the curvature that would otherwise be removed with it. Its Hessian in a field direction h is $\mathbb{E}[\omega\|h\|^2]$. At unit weights, $\mathcal{L}_1(d) = \frac{1}{2}\mathbb{E}\|d\|^2$, so zero correction is stationary for every training tuple.

For the shared reference $\widetilde{v} = v$ and positive weights, implicit Tilt Matching with unit control variate (Potaptchik et al., 2026, Eq. 26) satisfies

$$\begin{aligned}\mathbb{E}\big[\omega^{-1}\|d + (\omega-1)(v+d-V)\|^2\big] \\ = 2\mathcal{L}_\omega(d) + \mathbb{E}\left[\frac{(\omega-1)^2}{\omega}\|V-v\|^2\right].\end{aligned} \tag{26}$$

The last term is independent of the student correction. Hence the objectives have the same optimization force up to scale, including the positive curvature and stationary unit-weight update. The separate-reference construction and its baseline and calibration errors are analyzed in Appendix B.3.

A.5 Representation, decision, and computation

FRRF fixes three representations, and each decides what the method computes immediately and what it must pay for: the observation criterion the E-step optimizes, the generative path the field represents, and the parameterization of the velocity. This section develops the paragraph that closes Section 3; Section 5 tests the choices.

Observation fit. Besides MMD and KL, Fisher divergence is available without any learned ratio: for Gaussian-smoothed observations the score of $K(\omega\lambda)$ is an explicit responsibility-weighted mixture of conditional scores, so Hyvärinen's identity or its denoising form gives an empirical objective whose gradients in ω are unbiased (Hyvärinen, 2005; Vincent, 2011), the route that estimates empirical-Bayes priors and coincides there with SURE (Ghosh et al., 2025). But the responsibilities make it nonconvex in ω, forfeiting the certificate of (12), and scores are blind to the mixing proportions of well-separated components, exactly the frequencies the E-step adjusts, unless intermediate smoothing scales are added (Wenliang & Kanagawa, 2020; Dennehy et al., 2026). MMD instead delegates sensitivity to the kernel, and that is where representation enters: a kernel on raw observations spends resolution uniformly over coordinates, including filler and mask entries, whereas a kernel on a fixed pretrained embedding (Szegedy et al., 2016; Radford et al., 2021) concentrates it on semantic structure, with a pixel-level MMD guard against implausible redistributions (Appendix A.3). The convex decision is unchanged and the extra cost is one encoder pass over the simulated measurements; identification, and the class on which α_{inv} is positive, move with the encoder.

An anchored path makes restoration and denoising native. The straight interpolation represents generation natively; restoring or denoising an observed y then requires posterior inference under μ_θ and K, which a simulator-only channel does not supply. Nothing in FRRF depends on the straight path. Represent the observation in signal space, with independent noise filling unobserved coordinates, draw $(X, Y) \sim \mu \otimes K$, and let the path visit $Z_0 = \epsilon$, $Z_{1/2} = Y$, $Z_1 = X$: a straight segment $Z_t = (2-2t)Y + (2t-1)X$ on $[\frac{1}{2}, 1]$, so that $V = 2(X - Y)$ at the anchor, joined to a C^1 cubic from ϵ to Y that matches this velocity. The weight depends only on X, so (14) applies unchanged, and the canonical field v_0 supports three readouts: generation integrates from $t = 0$ to 1; restoration integrates from an observed y at $t = \frac{1}{2}$ to 1, a transport that pushes $K\mu$ to μ when the canonical flow from the anchor is well posed, which an atomic observation law such as $Y = \mathbf{1}\{X > 0\}$ does not allow without a randomized augmentation of Y; and denoising is one evaluation at the anchor, since conditioning on $Z_{1/2} = y$ is conditioning on $Y = y$, which gives the denoiser in (15), the posterior mean under the current clean law, marginalized over acquisition variables not supplied to the network. The deployed field inherits the mismatch $v_\theta - v_0$ of Section 3.2 in all three readouts; for the one-evaluation readouts it is the error at the anchor time, $\frac{1}{4}\mathbb{E}\|v_\theta(Y, \frac{1}{2}) - v_0(Y, \frac{1}{2})\|^2$ for denoising, which the time-integrated bound of Appendix B.3 does not control unless training places positive mass at $t = \frac{1}{2}$. Posterior sampling still needs an additional model (Table 2). The restoration readout also invites a collaboration with SCSI (Modi et al., 2026): a restoration pass over real observations proposes new candidate locations while the E-step selects their frequencies, an alternation of locations and weights reminiscent of scalable NPMLE (Zhang et al., 2022); since Section 4 concerns weights on the generator's own samples, restored proposals enter only through the field, and the next decision uses a fresh bank from the refitted field.

	Posterior EM	DiffEM	SCSI	Direct fit	FRRF, straight	FRRF, anchored
State	$\nabla \log \mu_{\theta,t}$	$\Pi_\theta(\cdot \mid y)$	$\Psi_\theta : y \mapsto x$	g_θ	v_θ; $Z_1{=}X$	v_θ; $Z_{1/2}{=}Y$
Access to K	$\nabla_x \log K$	draws	draws	$\nabla_x K$/density	draws	draws
Generation $X \sim \mu_\theta$	✓	$+\{\hat{x}_\ell\}$	$+\{\Psi_\theta(Y_\ell)\}$	✓	✓	✓ $\int_0^1$
Restoration $y \mapsto \hat{x}$	Π	✓	✓	Π	Π	✓ $\int_{1/2}^1$
Denoising $\mathbb{E}[X \mid Y{=}y]$	Π	✓ avg.	$+$	Π	Π	✓ (15)
Posterior $X \sim \Pi(\cdot \mid y)$	Π	✓	$+$	Π	Π	Π

Table 2: Decisions each learned state supports natively, and the access to K its training needs: posterior EM with an unconditional prior (Rozet et al., 2024; Bai et al., 2024), DiffEM (Hosseintabar et al., 2025), SCSI (Modi et al., 2026), direct observation fitting through a generator g_θ (Bora et al., 2018; Vashistha et al., 2026), and FRRF with the straight or anchored path. ✓: direct readout of the learned state. Π: posterior inference combining the learned prior with K, which needs an analytical likelihood or an additional model. $+$: an additional model fitted for this decision, on reconstructions $\{\hat{x}_\ell\}$ or restorations $\{\Psi_\theta(Y_\ell)\}$; SCSI's transport image is a restoration, not $\mathbb{E}[X \mid Y = y]$. The anchored readouts are identities for the canonical field. Cells describe the represented state, not the quality of the decision.

The velocity parameterization is a separate choice. The E-step never touches the network and the M-step is a regression in velocity space, so the backbone can be exchanged. MeanFlow replaces the instantaneous velocity by the average velocity along the trajectory, $\bar{v}(z_{t_0}, t_0, t_1) = (t_1 - t_0)^{-1} \int_{t_0}^{t_1} v(z_s, s)\, ds$, which permits the one-evaluation map $z_{t_1} = z_{t_0} + (t_1 - t_0)\, \bar{v}(z_{t_0}, t_0, t_1)$ and satisfies $\bar{v}(z, t, t) = v(z, t)$ (Geng et al., 2025; 2026; Lu et al., 2026); in this orientation the trainable identity reads $\bar{v} = v(z_{t_0}, t_0) + (t_1 - t_0)\, \frac{d}{dt_0} \bar{v}$, with the total derivative along the trajectory. A MeanFlow realization of FRRF trains $\bar{v}$ on intervals within $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$, keeps the diagonal readout at the anchor so that (15) is unchanged, and centers the residual loss on a frozen average-velocity reference on the same tuples with the same stop-gradient treatment, which preserves the unit-weight null. Its costs are the Jacobian–vector product in the MeanFlow identity and a one-step read that must be validated against multi-step integration of the same field: Appendix B.3 bounds the instantaneous field, and a one-step map can depart from its integration off the diagonal without changing it, a consistency error that belongs in δ_{sam}.

B THEORY

This appendix proves Theorem 1 (Section B.1), constructs the comparator of the coverage condition (Section B.2), bounds the realization error (Section B.3), and collects complementary results on stability, the population penalized step, fixed-penalty progress, and inverse stability (Section B.4).

B.1 PROOF OF THE RECOVERY THEOREM

We analyze the penalized objective directly. Its quadratic term controls the same changes in weights that can amplify conditional-simulation error. A comparison with a partially feasible reweighting then gives a norm-level recurrence for the generated laws. The expanded bound below implies the target-accuracy theorem in the main text.

B.1.1 OBJECTS AND ERROR INTERFACES

Let $Y_1, \dots, Y_n$ be independent observations from $p^\star = K\mu^\star$. Fix a real positive-definite kernel k, independent of these observations, with strongly measurable feature map $\phi(y) = \mathsf{k}(y, \cdot)$ in a separable Hilbert space $\mathcal{H}$ and $\mathsf{k}(y, y) \le 1$. Define

$$m^\star = \mathbb{E}_{p^\star}\phi(Y), \qquad \widehat{m} = \frac{1}{n}\sum_{j=1}^{n}\phi(Y_j), \qquad \psi(x) = \mathbb{E}[\phi(Y) \mid X = x].$$

Then $\|\psi(x)\| \le 1$, $m_{K\nu} = \mathbb{E}_\nu\psi(X)$, and the criterion (21) reads $F_{\mathrm{M}}(\nu) = \frac{1}{2}\|m^\star - m_{K\nu}\|^2$. This is the mean-embedding representation of squared MMD (Gretton et al., 2012, Lemma 4). Unqualified feature norms are in $\mathcal{H}$; L^2 averages the dataset and all algorithmic randomness.

Conditionally on the dataset and past, round k draws independent $X_{ki} \sim \mu_k$, $i \le N$, and sets $\lambda_k = N^{-1}\sum_i \delta_{X_{ki}}$. Given the bank, independently draw $Y_{kij} \sim K(\cdot \mid X_{ki})$, $j \le m$, across all candidate–simulation pairs, and set

$$\widehat{\psi}_{ki} = \frac{1}{m}\sum_j \phi(Y_{kij}), \qquad M_k(\omega) = \frac{1}{N}\sum_i \omega_i\psi(X_{ki}), \qquad \widehat{M}_k(\omega) = \frac{1}{N}\sum_i \omega_i\widehat{\psi}_{ki},$$

and $\|u\|_{\lambda_k}^2 = N^{-1}\sum_i u_i^2$. Weights attach to labeled indices, including when candidate values repeat. Let $\mathcal{A}_k$ be a nonempty closed convex subset of the normalized nonnegative simplex containing $\mathbf{1}$. A measurable feasible correspondence and solver selection are assumed. Compactness gives existence, and the strictly convex quadratic penalty gives uniqueness of the exact minimizer. The step sizes obey deterministic bounds $0 < \eta_{\min} \le \eta_k \le \eta_{\max} < \infty$; their choice may depend on the cache within these bounds.

The computed feasible weight satisfies the certified penalized excess

$$\widehat{J}_k(\widehat{\omega}_k) \le \min_{\omega\in\mathcal{A}_k}\widehat{J}_k(\omega) + \varepsilon_{\mathrm{opt},k}, \qquad 0 \le \varepsilon_{\mathrm{opt},k} \le \bar{\varepsilon}_{\mathrm{opt}} \quad \text{almost surely.} \tag{27}$$

The constant $\bar{\varepsilon}_{\mathrm{opt}}$ is a solver tolerance for the displayed penalized objective. For example, the convex first-order gap at a feasible point bounds this excess. Generator training and sampling produce μ_{k+1}; write

$$\widehat{\nu}_k = \widehat{\omega}_k\lambda_k, \qquad g_k = \|m_{K\mu_{k+1}} - M_k(\widehat{\omega}_k)\|, \qquad \varepsilon_{\mathrm{fit}} = \sqrt{2\bar{\varepsilon}_{\mathrm{opt}}} + \sup_{k<T}\|g_k\|_{L^2}. \tag{28}$$

Section B.3 connects g_k to residual regression and sampling.

A comparator at the chosen penalty. Let $\mathcal{G}_k$ be the σ-field generated by the dataset, the past, and the current bank, but not by its new conditional-simulation draws. Assume that a $\mathcal{G}_k$-measurable normalized nonnegative vector c_k satisfies

$$\bar{\omega}_k = \tfrac{1}{2}(\mathbf{1} + c_k) \in \mathcal{A}_k \quad \text{almost surely}, \qquad a_k = \|M_k(c_k) - m^\star\|, \qquad C_k = \|c_k - \mathbf{1}\|_{\lambda_k}^2,$$
$$\sup_{k<T}\mathbb{E}\left[a_k^2 + \frac{C_k}{\eta_{\min}}\right] \le \varepsilon_{\mathrm{cov}}^2. \tag{29}$$

Thus the comparator depends on the generated candidates, not on the error in their simulated features. It need not be a feasible full update: every floor $\ell \le 1/2$ admits its halfway weights; any further guards must admit them as well. The assumption controls expected correction cost rather than imposing a deterministic cap on the selected weights.

B.1.2 THE PENALTY CONTROLS SELECTION OF SIMULATION NOISE

Suppress k and condition on $\mathcal{G}_k$. Let

$$e_i = \widehat{\psi}_i - \psi(X_i), \qquad \bar{e} = \frac{1}{N}\sum_i e_i, \qquad V_e = \frac{1}{N}\sum_i \|e_i - \bar{e}\|^2.$$

For any normalized weights, writing $u = \omega - \mathbf{1}$ gives

$$\widehat{M}(\omega) - M(\omega) = \bar{e} + \mathcal{E}u, \qquad \mathcal{E}u = \frac{1}{N}\sum_i u_i(e_i - \bar{e}), \qquad \|\mathcal{E}u\| \le \sqrt{V_e}\,\|u\|_\lambda. \tag{30}$$

The last inequality is Cauchy–Schwarz and holds simultaneously for cache-selected weights.

Lemma 1 (Conditional moments of the simulation cache)**.** *Under the stated independent conditional sampling,*

$$\mathbb{E}[\|\bar{e}\|^2 \mid \mathcal{G}_k] \le \frac{1}{Nm}, \quad \mathbb{E}[V_e \mid \mathcal{G}_k] \le \frac{1}{m}, \quad \mathbb{E}[\|\bar{e}\|^4 \mid \mathcal{G}_k] \le \frac{4}{N^2m^2}, \quad \mathbb{E}[V_e^2 \mid \mathcal{G}_k] \le \frac{4}{m^2}. \tag{31}$$

Put $\rho = \eta_{\max}/m$ and $D_k = 1 + \eta_{\max}V_e$. Then

$$\mathbb{E}[D_k \mid \mathcal{G}_k] \le 1 + \rho, \qquad \mathbb{E}[D_k^2 \mid \mathcal{G}_k] \le 1 + 2\rho + 4\rho^2, \qquad \mathbb{E}[D_k\|\bar{e}\|^2 \mid \mathcal{G}_k] \le \frac{1+4\rho}{Nm}. \tag{32}$$

Proof. A feature centered conditionally on $\mathcal{G}_k$, $\xi_{ij} = \phi(Y_{kij}) - \psi(X_{ki})$, satisfies $\mathbb{E}\|\xi\|^2 \le 1$, $\|\xi\| \le 2$, and $\mathbb{E}\|\xi\|^4 \le 4$. For independent centered $\xi_1, \dots, \xi_L$ with these bounds, expanding $\|\sum_j \xi_j\|^4$, dropping the terms that contain an unpaired centered factor, and using $\mathbb{E}\langle\xi_i, \xi_j\rangle^2 \le \mathbb{E}\|\xi_i\|^2\mathbb{E}\|\xi_j\|^2$ for the paired cross terms give

$$\mathbb{E}\Big\|\sum_{j=1}^{L}\xi_j\Big\|^4 \le \sum_j \mathbb{E}\|\xi_j\|^4 + 6\sum_{i<j}\mathbb{E}\|\xi_i\|^2\mathbb{E}\|\xi_j\|^2 \le 4L + 3L(L-1) \le 4L^2.$$

Apply this with $L = m$ for each e_i and $L = Nm$ for $\bar{e}$; the variables need not be identically distributed. Since $V_e \le N^{-1}\sum_i \|e_i\|^2$, Jensen gives $\mathbb{E}V_e^2 \le N^{-1}\sum_i \mathbb{E}\|e_i\|^4 \le 4/m^2$. The second-moment bounds follow from centering and independence. Expansion gives the first two bounds in (32); conditional Cauchy–Schwarz gives $\mathbb{E}[V_e\|\bar{e}\|^2 \mid \mathcal{G}_k] \le 4/(Nm^2)$ and proves the third. □

B.1.3 ONE ROUND AND THE EXPANDED RECOVERY BOUND

Define

$$h_k = \|\widehat{m} - m_{K\mu_k}\|, \qquad r_k = \|m^\star - m_{K\mu_k}\|, \qquad \delta_n = \|\widehat{m} - m^\star\|, \qquad \beta_k = \|M_k(\mathbf{1}) - m_{K\mu_k}\|.$$

The data and bank variance identities give $\|\delta_n\|_{L^2} \le n^{-1/2}$ and $\|\beta_k\|_{L^2} \le N^{-1/2}$, including for data-dependent incoming laws.

Theorem 2 (Expanded recovery bound for penalized updates)**.** *Assume (27)–(29) and $m \ge \max\{1, \eta_{\max}\}$. Set*

$$\rho = \frac{\eta_{\max}}{m} \le 1, \qquad q = \frac{\sqrt{1+\rho}}{2} \le \frac{1}{\sqrt{2}}, \qquad b_\rho = \frac{\sqrt{2+3\rho+4\rho^2}}{2} \le \frac{3}{2}, \qquad F_0 = \mathbb{E}F_{\mathrm{M}}(\mu_0).$$

Then

$$\sqrt{2\mathbb{E}F_{\mathrm{M}}(\mu_T)} \le q^T\sqrt{2F_0} + \frac{1}{(1-q)\sqrt{n}} + \frac{1-q^T}{1-q}\left[\frac{q}{\sqrt{N}} + b_\rho\varepsilon_{\mathrm{cov}} + \frac{2\sqrt{1+4\rho}}{\sqrt{Nm}} + \sqrt{1+\rho}\,\varepsilon_{\mathrm{fit}}\right]. \tag{33}$$

In particular, for a numerical constant C (one valid choice is 720),

$$\mathbb{E}F_{\mathrm{M}}(\mu_T) \le C\left[2^{-T}F_0 + \frac{1}{n} + \varepsilon_{\mathrm{cov}}^2 + \frac{1}{N} + \varepsilon_{\mathrm{fit}}^2\right]. \tag{34}$$

An inverse inequality $F_{\mathrm{M}}(\nu) \ge \alpha_{\mathrm{inv}}\,\mathrm{KL}(\mu^\star\|\nu)$ on the output class converts this to the clean-KL bound by division by α_{inv}.

Proof. Penalized comparison. Work pathwise in one round and put $\eta = \eta_k$, $u = \omega - \mathbf{1}$. Equation (30) and Cauchy–Schwarz in $\mathcal{H} \oplus L^2(\lambda)$ give

$$\begin{aligned}\|\widehat{m} - M(\widehat{\omega})\| &\le \|\bar{e}\| + \sqrt{1+\eta V_e}\,\sqrt{\|\widehat{m} - \widehat{M}(\widehat{\omega})\|^2 + \eta^{-1}\|\widehat{\omega} - \mathbf{1}\|_\lambda^2} \\ &\le \|\bar{e}\| + \sqrt{1+\eta V_e}\left[\sqrt{2\widehat{J}(\bar{\omega})} + \sqrt{2\varepsilon_{\mathrm{opt}}}\right], \end{aligned} \tag{35}$$

where $\bar{\omega} = (\mathbf{1} + c)/2$ is feasible and (27) gives the second line. For $\bar{u} = (c - \mathbf{1})/2$, another triangle inequality in the same product space gives

$$\sqrt{2\widehat{J}(\bar{\omega})} \le \|\widehat{m} - M(\bar{\omega})\| + \|\bar{e}\| + \sqrt{V_e + \eta^{-1}}\,\|\bar{u}\|_\lambda.$$

Using $\sqrt{1+\eta V_e}\sqrt{V_e + \eta^{-1}} = (1 + \eta V_e)/\sqrt{\eta}$, $D_k = 1 + \eta_{\max} V_e$, and $\eta \ge \eta_{\min}$ yields

$$h_{k+1} \le \frac{\sqrt{D_k}}{2}(h_k + \beta_k + \delta_n + a_k) + (1 + \sqrt{D_k})\|\bar{e}\| + \frac{D_k}{2}\sqrt{\frac{C_k}{\eta_{\min}}} + \sqrt{2D_k\bar{\varepsilon}_{\mathrm{opt}}} + g_k. \tag{36}$$

This uses $M(\bar{\omega}) = \frac{1}{2}M(\mathbf{1}) + \frac{1}{2}M(c)$, followed by the bank, comparator, and generator interfaces. No bound on the selected movement is imposed.

Conditional moments. The quantities $h_k, \beta_k, \delta_n, a_k, C_k$ are $\mathcal{G}_k$-measurable. Lemma 1, the inequality $1 + \sqrt{D_k} \le 2\sqrt{D_k}$, and Minkowski therefore give, for $H_k = \|h_k\|_{L^2}$,

$$H_{k+1} \le qH_k + q\|\delta_n\|_{L^2} + \frac{q}{\sqrt{N}} + b_\rho\varepsilon_{\mathrm{cov}} + \frac{2\sqrt{1+4\rho}}{\sqrt{Nm}} + \sqrt{1+\rho}\,\varepsilon_{\mathrm{fit}}. \tag{37}$$

For the two comparator terms, use

$$\frac{\sqrt{1+\rho}}{2}\|a_k\|_{L^2} + \frac{\sqrt{1+2\rho+4\rho^2}}{2}\sqrt{\frac{\mathbb{E}C_k}{\eta_{\min}}} \le b_\rho\sqrt{\mathbb{E}a_k^2 + \frac{\mathbb{E}C_k}{\eta_{\min}}} \le b_\rho\varepsilon_{\mathrm{cov}}.$$

The deterministic solver tolerance ensures that the cache multiplier in its term is controlled by $\mathbb{E}D_k \le 1 + \rho$. No independence of the computed weights from their cache is used.

Iteration and the population comparison. Iterate (37), use $H_0 \le \|r_0\|_{L^2} + \|\delta_n\|_{L^2}$ and $\|r_T\|_{L^2} \le H_T + \|\delta_n\|_{L^2}$, and sum the geometric series. Since $1/2 \le q \le 1/\sqrt{2}$, the coefficient of $\|\delta_n\|_{L^2}$ is

$$1 + q^T + \frac{q(1-q^T)}{1-q} \le \frac{1}{1-q}.$$

This proves (33). Its six terms can be squared using $(\sum_{j=1}^6 z_j)^2 \le 6\sum_j z_j^2$. With $(1-q)^{-1} \le 2 + \sqrt{2}$, $m \ge 1$, and $q^{2T} \le 2^{-T}$, the largest coefficient after division by two and absorption of $1/(Nm)$ into $1/N$ is at most $3(2+\sqrt{2})^2(1/2 + 20) < 720$. This proves (34). Apply inverse stability only at μ_T for clean KL. □

B.1.4 Proof of Theorem 1

The choices $\varepsilon_{\mathrm{cov}}, \varepsilon_{\mathrm{fit}} \le \varepsilon$ and $N^{-1} \le \varepsilon^2$ in (34) give the main rate, for example with hidden constant 2160. For (20), take $\varepsilon = n^{-1/2}$. Since the normalized kernel gives $F_0 \le 2$, $T \ge \lceil \log_2(2n) \rceil$ makes $2^{-T}F_0 \le n^{-1}$.

One dataset, terminal recovery. The same data error is present in every round and has a bounded geometric coefficient; there is no factor T/n. The final comparison uses the same fixed-kernel minimum-distance argument as Chérief-Abdellatif & Alquier (2022, Theorem 3.1 and Remark 3.1); the recurrence supplies its input for the actual generated laws. Data-dependent iterates are allowed; the conditional independence of a bank comparator from that bank's simulation noise is stated in (29). Correlated simulations or a data-fitted kernel require the corresponding moment controls. Inverse stability is required on the output class, not on the atomic selected bank.

The role of the penalty. The sharper squared transient is $[(1+\eta_{\max}/m)/4]^T$; the main theorem uses its uniform bound 2^{-T}. A fixed finite penalty leaves the correction-cost term $\mathbb{E}C_k/\eta$ inside the sufficient accuracy bound. Thus the sufficient bound reaches a smaller target error only with a sufficiently affordable comparator at that penalty, or with a larger η and matching simulation accuracy. The bound does not assert a nonzero limiting bias for exact fixed-penalty dynamics; the complementary fixed-penalty progress and population-KL results analyze those dynamics from other assumptions.

B.2 Coverage at the chosen penalty

The coverage condition pairs observation agreement with its quadratic reweighting cost. Expected cost suffices for Theorem 2, so an unbounded importance ratio with a finite second moment can be used directly.

B.2.1 What the comparator measures

For a fixed bank and a comparison weight c, the two quantities are $\|M(c)-m^\star\|^2$ and $\eta^{-1}\|c-\mathbf{1}\|_\lambda^2$. The first measures the observation discrepancy after changing candidate frequencies; the second prices that change with the same penalty as the algorithm. Only $(\mathbf{1}+c)/2$ must be feasible. The algorithm computes neither the comparator nor the unknown population mean used to assess it.

The condition prices the incoming discrepancy. Since $\|\psi\|\le 1$, $\|M(c)-M(\mathbf{1})\|\le\|c-\mathbf{1}\|_\lambda$, so $\|m^\star-M(\mathbf{1})\|\le a+\sqrt{C}$ and, by Cauchy–Schwarz, $\|m^\star-M(\mathbf{1})\|^2\le(1+\eta)(a^2+C/\eta)$; taking expectations and using the bank variance identity gives $2\mathbb{E}F_{\mathrm{M}}(\mu_k)\le(1+\eta_{\min})\varepsilon_{\mathrm{cov}}^2$. Coverage at accuracy ε therefore requires $\eta_{\min}\ge 2F_0/\varepsilon^2-1$ at initialization, and with $m\ge\eta_{\max}$ and $N\ge\varepsilon^{-2}$ the prescription of Theorem 1 uses $Nm\gtrsim\varepsilon^{-4}$ simulations per round. An affordable accurate correction is thus built into the assumption, which is why no inverse condition number enters the geometric factor; the theorem is a geometric approach to the assumed accuracy, not a fixed-cost contraction for arbitrary incoming laws.

B.2.2 Coverage from bounded or square-integrable ratios

Proposition 2 (A bank comparator from a finite importance second moment)**.** *Fix μ with $\mu^\star\ll\mu$, and let $w=d\mu^\star/d\mu$, taken in a jointly measurable version, satisfy $\mathbb{E}_\mu w^2\le\varsigma<\infty$. For N independent candidates from μ, there is a measurable normalized nonnegative comparator c, depending only on the bank, with*

$$\mathbb{E}\|c-\mathbf{1}\|_\lambda^2\le 4(1-N^{-1})(\varsigma-1),\qquad \mathbb{E}\|M(c)-m^\star\|^2\le\frac{32\varsigma}{N}. \tag{38}$$

Consequently, when its halfway weights satisfy the base constraints,

$$\varepsilon_{\mathrm{cov}}(\eta_{\min},N)^2\le\frac{32\varsigma}{N}+\frac{4(\varsigma-1)}{\eta_{\min}}. \tag{39}$$

If additionally $w\le U$, the same construction has $\|c-\mathbf{1}\|_\lambda^2\le 2U-1$ almost surely and $\mathbb{E}\|M(c)-m^\star\|^2\le 20U/N$.

Proof. Write $w_i=w(X_i)$ and $\bar w=N^{-1}\sum_i w_i$. On $\mathcal{B}=\{\bar w\ge 1/2\}$ set $c_i=w_i/\bar w$; on $\mathcal{B}^c$ set $c_i=1$. On $\mathcal{B}$,

$$\|c-\mathbf{1}\|_\lambda^2=\frac{N^{-1}\sum_i(w_i-\bar w)^2}{\bar w^2}\le 4\,\frac{1}{N}\sum_i(w_i-\bar w)^2.$$

The cost is zero on $\mathcal{B}^c$. Taking expectations and using the mean empirical-variance identity proves the first bound in (38).

Set $Z_i=w_i(\psi(X_i)-m^\star)$. Then $\mathbb{E}Z_i=0$ and $\mathbb{E}\|Z_i\|^2\le 4\varsigma$. On $\mathcal{B}$, the squared observation error is at most $4\|N^{-1}\sum_i Z_i\|^2$, contributing at most $16\varsigma/N$ in expectation. On $\mathcal{B}^c$, the squared error is at most 4, and Chebyshev gives $\Pr(\mathcal{B}^c)\le 4(\varsigma-1)/N$. Its contribution is at most $16(\varsigma-1)/N$, proving the stated bound.

For $w\le U$, on $\mathcal{B}$ one has $N^{-1}\sum_i c_i^2\le U/\bar w\le 2U$. Also $\mathbb{E}\|Z_i\|^2\le U\mathbb{E}_{\mu^\star}\|\psi(X)-m^\star\|^2\le U$. The good and bad events contribute at most $4U/N$ and $16(U-1)/N$, respectively. □

The construction is safeguarded self-normalized importance sampling. Its second-moment dependence is that of Agapiou et al. (2017, Theorem 2.1); the explicit expected quadratic-cost bound and its use in the penalized recurrence are established here. No clipping of the unknown importance ratio is needed, because the analysis controls expected rather than worst-case movement.

Corollary 1 (Recovery with a fixed penalty and a positive floor)**.** *Under the assumptions of Theorem 2 other than (29), suppose $\mu^\star \ll \mu_k$ and $\mathbb{E}_{\mu_k}(d\mu^\star/d\mu_k)^2 \le \varsigma$ almost surely along the actual trajectory. Let the floor satisfy $0 < \ell \le 1/2$, and assume that every other constraint defining $\mathcal{A}_k$ admits the halfway comparator $(\mathbf{1}+c_k)/2$ almost surely. For a fixed $\eta_k = \eta > 0$ and $m \ge \max\{1, \eta\}$,*

$$\mathbb{E}F_{\mathrm{M}}(\mu_T) \lesssim 2^{-T}F_0 + \frac{1}{n} + \frac{\varsigma}{N} + \frac{\varsigma - 1}{\eta} + \varepsilon_{\mathrm{fit}}^2. \tag{40}$$

In particular, $N \gtrsim \varsigma n$, $\eta \gtrsim \varsigma n$, $m \ge \eta$, $\varepsilon_{\mathrm{fit}} \lesssim n^{-1/2}$, and $T \gtrsim \log(en)$ suffice for $\mathbb{E}F_{\mathrm{M}}(\mu_T) \lesssim n^{-1}$. Output inverse stability gives the corresponding $\mathbb{E}\,\mathrm{KL} \lesssim (\alpha_{\mathrm{inv}} n)^{-1}$.

Proof. The comparator is normalized and nonnegative, so $(\mathbf{1} + c)/2 \ge 1/2 \ge \ell$. Apply Proposition 2 conditionally on each incoming law and insert (39) into Theorem 2. Jensen gives $\varsigma \ge 1$, which absorbs the separate $1/N$ term. The stated choices bound the remaining terms by numerical multiples of $1/n$. □

A positive floor thus preserves regression curvature without adding a separate coverage-bias term. The feasibility assumption covers every constraint, not only observation guards: a floor together with the cap $\omega_i \le 1$ leaves only $\mathbf{1}$ feasible, and the corollary then fails; under $w \le U$ the construction has $c_i \le 2U$, so a cap $\omega_i \le U'$ admits the halfway comparator when $U' \ge U + \frac{1}{2}$. The uniform importance-moment bound and fitting accuracy concern the actual trajectory; the corollary does not infer them from initialization alone.

B.2.3 Choosing the penalty and allocating simulations

When comparing step sizes, use a common class of incoming laws and the same base constraints. On a fixed bank with cache-independent base constraints, define

$$\varepsilon_k^\star(\eta)^2 = \min_{\substack{c \ge 0,\ \mathbb{E}_{\lambda_k} c = 1 \\ (\mathbf{1}+c)/2 \in \mathcal{A}_k}} \left\{ \|M_k(c) - m^\star\|^2 + \eta^{-1}\|c - \mathbf{1}\|_{\lambda_k}^2 \right\}. \tag{41}$$

This is nonincreasing in η on that fixed bank. A coverage envelope used to compare trajectories must bound its expectation over banks from one common class. Larger η lowers the cost of the required correction; $m \ge \eta_{\max}$ controls its sensitivity to conditional-simulation error. Candidate count also enters coverage, not only the separate bank approximation term.

For a target observation-mean accuracy ε, Proposition 2 gives the explicit sufficient choices

$$N \ge 64\varsigma\varepsilon^{-2}, \qquad \eta \ge 8\varsigma\varepsilon^{-2}, \qquad m \ge \max\{1, \eta\}, \qquad \varepsilon_{\mathrm{fit}} \le \varepsilon. \tag{42}$$

They make $\varepsilon_{\mathrm{cov}} \le \varepsilon$ and satisfy the main theorem. The constant ς quantifies representation difficulty; it is an analytical quantity, not a statistic supplied to the algorithm. These are sufficient accuracy requirements rather than lower bounds on necessary computation.

For a fixed simulator-call budget $\mathsf{B} = Nm$ and a common envelope $\varepsilon_{\mathrm{cov}}(\eta, N)^2 \le a/N + b/\eta$, the explicit terms to balance are

$$\frac{1+a}{N} + \frac{b}{\eta} + \frac{1}{Nm}, \qquad m \ge \max\{1, \eta\}. \tag{43}$$

In a continuous allocation with $\eta \ge 1$, take $m = \eta$ and $N = \mathsf{B}/\eta$. The interior minimizer is

$$\eta_{\mathrm{opt}} = m_{\mathrm{opt}} = \sqrt{\frac{b\mathsf{B}}{1+a}}, \qquad N_{\mathrm{opt}} = \sqrt{\frac{(1+a)\mathsf{B}}{b}}, \tag{44}$$

clipping η_{opt} to $[1, \mathsf{B}]$ and rounding to feasible sample counts when needed. For $b = 0$, the boundary choice $\eta = m = 1$, $N = \mathsf{B}$ minimizes this bound. Under the finite-second-moment envelope, balancing N, m, η at order $\sqrt{\mathsf{B}}$ suffices to make the model-side squared error $O(\varsigma/\sqrt{\mathsf{B}})$, apart from fitting. This allocation concerns a sufficient bound on simulator calls; candidate generation, solving the weight problem, and residual fitting have additional costs.

B.3 REALIZATION: FROM RESIDUAL TRAINING TO THE GENERATED LAW

The realization error $g_k = \|m_{K\mu_{k+1}} - M_k(\widehat{\omega}_k)\|$ measures preservation of the selected observation mean. We bound it through the residual loss of Section 3.2 and flow stability.

From the residual loss to a field error. Suppress k. Draw a uniform bank index I, the endpoint X_I, and its weight $\Omega = \widehat{\omega}_I$, so that weights attach to labeled candidates even when two share a value; then draw the interpolation state $S = (Z_t, t)$ and target velocity V as in Section 2, conditionally on the round's training artifacts. With

$$v_0 = \mathbb{E}[V \mid S], \qquad m_\omega = \mathbb{E}[\Omega \mid S], \qquad v_\omega = \frac{\mathbb{E}[\Omega V \mid S]}{m_\omega},$$

and finite weighted second moments of V, v, $\widetilde{v}$, and d under a positive floor, quadratic completion of (14) gives its unrestricted optimum and excess,

$$d^\star = v_\omega - v_0 + \frac{m_\omega - 1}{m_\omega}(v_0 - \widetilde{v}), \qquad \delta_{\mathrm{res}} := \mathcal{L}_\omega(d) - \mathcal{L}_\omega(d^\star) = \tfrac{1}{2}\mathbb{E}[m_\omega \|d - d^\star\|^2], \tag{45}$$

where δ_{res} includes neural approximation as well as optimization error; it is the excess over the unrestricted conditional optimum, not the raw training loss, which can be negative. The floor is needed: where $m_\omega(S) = 0$ the conditional loss equals $(v_0 - \widetilde{v})^\top d$, linear in d, so without a floor the objective is bounded below only if $\widetilde{v} = v_0$ there, or under an integrability condition such as $\mathbb{E}[\|b\|^2/m_\omega] < \infty$ for $b = \mathbb{E}[(\Omega - 1)(V - \widetilde{v}) \mid S]$. Measure the two remaining discrepancies by

$$B_\omega = \mathbb{E}[m_\omega \|v - v_0\|^2], \qquad C_\omega^{\mathrm{cal}} = \mathbb{E}\left[\frac{(m_\omega - 1)^2}{m_\omega}\|\widetilde{v} - v_0\|^2\right], \tag{46}$$

the incoming field's mismatch with the present interpolation and the reference-calibration error.

Proposition 3 (Norm-level realization decomposition)**.** *With* $E_v = \mathbb{E}[m_\omega \|v + d - v_\omega\|^2]$,

$$\sqrt{E_v} \le \sqrt{2\delta_{\mathrm{res}}} + \sqrt{B_\omega} + \sqrt{C_\omega^{\mathrm{cal}}}, \tag{47}$$

and for the shared reference $\widetilde{v} = v$,

$$\sqrt{E_v} \le \sqrt{2\delta_{\mathrm{res}}} + \Big(\mathbb{E}\frac{\|v - v_0\|^2}{m_\omega}\Big)^{1/2}. \tag{48}$$

Proof. Decompose $v + d - v_\omega = (d - d^\star) + (v - v_0) + \frac{m_\omega - 1}{m_\omega}(v_0 - \widetilde{v})$ and apply the triangle inequality in $L^2(m_\omega)$; the three norms are $\sqrt{2\delta_{\mathrm{res}}}$, $\sqrt{B_\omega}$, and $\sqrt{C_\omega^{\mathrm{cal}}}$. For $\widetilde{v} = v$ the last two vectors combine exactly into $(v - v_0)/m_\omega$ before any bound is taken. □

Calibration removes C_ω^{cal} when exact; it does not remove the incoming mismatch B_ω, which is not an artifact of the bound. Let the source and the incoming law be $\mathcal{N}(0, 1)$ with the incoming field $v \equiv 0$; on the straight path, $v_0(z, t) = (2t - 1)z/Q_t$ with $Q_t = (1 - t)^2 + t^2$, so $v \neq v_0$ although both generate the incoming law. For the tilt $\omega_b(x) = e^{bx - b^2/2}$, whose requested law is $\mathcal{N}(b, 1)$, one has $v_{\omega_b} = v_0 + b(1 - t)/Q_t$, and exact calibration with exact residual minimization deploys $v + d^\star = b(1 - t)/Q_t$, whose flow ends at $\mathcal{N}(\pi b/4, 1)$. At unit weights the residual gradient vanishes at $d = 0$ for every training tuple, which removes the baseline refitting force; the decomposition nevertheless retains B_ω, so the estimate is a sufficient bound that can be loose exactly there. Specializing the regression c-ITM objective of Potaptchik et al. (2026, Eq. 26) to the unit control variate and rearranging shows that (14), with $\widetilde{v} = v$, is equivalent up to scale and a student-independent term. It therefore has the same samplewise unit-weight null; the separate-reference decomposition and its composition with the outer recurrence are the analysis added here.

From the field to the generated law. The endpoint error of the learned flow is controlled by the following stability theorem, transcribed in the present notation.

Theorem 3 (Flow stability; Benton et al. (2024), Theorem 1). *Let Z_t solve $\dot{Z}_t = v^\star(Z_t, t)$ from $Z_0 = \epsilon \sim \pi_0$, with marginals p_t and endpoint law π_1, and let $\widehat{Z}_t$ solve $\dot{\widehat{Z}}_t = \widehat{v}(\widehat{Z}_t, t)$ from the same ϵ, with endpoint law $\widehat{\pi}_1$. Assume that both flows exist, are unique, and are continuously differentiable in their arguments; that $\widehat{v}(\cdot, t)$ is differentiable and L_t-Lipschitz in space for each t; and that*

$$\int_0^1 \mathbb{E}\|\widehat{v}(Z_t, t) - v^\star(Z_t, t)\|^2 \, dt \le \varepsilon^2.$$

Then $W_2(\widehat{\pi}_1, \pi_1) \le \varepsilon \exp\{\int_0^1 L_t \, dt\}$.

The atomic endpoint $\widehat{\nu} = \omega\lambda$ of a bank is not automatically covered by the smooth-flow hypotheses, so we assume directly what the proof uses: the canonical field v_ω has characteristics with the selected interpolation marginals, the common source, and endpoint $\widehat{\nu}$; the deployed field $\widehat{v} = v + d$ generates a well-posed flow with spatial Lipschitz modulus L_t and $\Lambda = \int_0^1 L_t \, dt < \infty$; and the paths have finite second moments, with endpoint limits and integrals understood in L^2. Under these assumptions the argument of Theorem 3 applies verbatim: coupling the target and learned trajectories at their common source, adding and subtracting $\widehat{v}_t$ along the target trajectory, and applying Grönwall's inequality gives $|\widehat{Z}_1 - Z_1| \le e^\Lambda \int_0^1 \|\widehat{v}_t(Z_t) - v_\omega(Z_t, t)\| \, dt$, and Minkowski with Cauchy–Schwarz gives the bound of the theorem.

Corollary 2 (Realized endpoint of the learned flow). *Let δ_{sam} bound the sampler's W_2 error relative to the endpoint of the exact learned flow. Under the assumptions above,*

$$W_2(\mu_{k+1}, \widehat{\nu}_k) \le e^{\Lambda_k} \sqrt{E_{v,k}} + \delta_{\mathrm{sam},k}. \tag{49}$$

Proof. Apply Theorem 3 with $v^\star = v_\omega$ and $\widehat{v} = v + d$. Its velocity error is E_v: time is uniform, and m_ω is the density of the selected interpolation marginals with respect to the unweighted ones, so $\int_0^1 \mathbb{E}_{p_t}\|\widehat{v} - v_\omega\|^2 dt = \mathbb{E}[m_\omega \|v + d - v_\omega\|^2]$. The triangle inequality for W_2 adds the sampler term. □

A time density bounded below by $\varrho_{\min} > 0$ replaces E_v by $E_v/\varrho_{\min}$.

If the conditional embedding is Lipschitz, $\|\psi(x) - \psi(x')\| \le L_\psi \|x - x'\|$, a coupling of μ_{k+1} and $\widehat{\nu}_k$ bounds the difference of predicted means by $L_\psi W_1 \le L_\psi W_2$, so

$$g_k \le L_\psi \Big[e^{\Lambda_k} \sqrt{E_{v,k}} + \delta_{\mathrm{sam},k} \Big]. \tag{50}$$

The W_2 route is dimension dependent: an absolutely continuous μ_{k+1} on $\mathbb{R}^d$ is at W_2 distance of order at least $N^{-1/d}$ from any N-atom law, whereas g_k compares mean embeddings and can be of order $N^{-1/2}$; (50) is a sufficient certificate, not a dimension-free one. If $\Lambda_k \le \bar{\Lambda}$ almost surely and the expected excess loss, mismatch, calibration, and squared sampler errors are bounded by $\bar{\delta}_{\mathrm{res}}, \bar{B}, \bar{C}^{\mathrm{cal}}, \bar{\delta}^2_{\mathrm{sam}}$, with penalized objective excess $\varepsilon_{\mathrm{opt},k} \le \bar{\varepsilon}_{\mathrm{opt}}$ almost surely, then the fitting accuracy in (28) obeys

$$\varepsilon_{\mathrm{fit}} \le \sqrt{2\bar{\varepsilon}_{\mathrm{opt}}} + L_\psi \Big[e^{\bar{\Lambda}} \big(\sqrt{2\bar{\delta}_{\mathrm{res}}} + \sqrt{\bar{B}} + \sqrt{\bar{C}^{\mathrm{cal}}} \big) + \bar{\delta}_{\mathrm{sam}} \Big]. \tag{51}$$

Equation (51) specifies sufficient solver, regression, and sampling accuracies for the outer theorem. Propagating g_k in the mean norm before squaring preserves their squared-error scales. The sample-wise unit-weight null and the realization decomposition describe what residual matching controls; neither asserts a universal endpoint-error ordering against weighted flow matching.

B.4 Complementary results

This section collects results that support the recovery theorem without being part of its proof: the stability of the constrained alternative to the penalized decision (Section B.4.1), the penalized decision of Proposition 1 with its sensitivity and solver bounds (Section B.4.2), the population clean-KL mechanism that the penalized step admits under two-sided overlap (Section B.4.3), a finite-time progress bound for a fixed penalty (Section B.4.4), and the meaning of the inverse-stability constant (Section B.4.5). Throughout, λ is the fixed reference law of a decision, the incoming population for a population statement and the bank for its finite realization; weight inner products are $\langle a, b \rangle_\lambda = \mathbb{E}_\lambda[ab]$.

B.4.1 THE CONSTRAINED ALTERNATIVE IS A PROJECTION IN OBSERVATION-MEAN SPACE

The constrained alternative (24) to the penalized decision is convex but not strongly convex in the weights, and its minimizer need not be unique when the cached features have a nontrivial null space. Its observation consequence is nevertheless unique and stable. This property is not used in the proof of Theorem 1, which analyzes the penalized decision.

Proposition 4 (Projection contraction and stable predicted means)**.** *Fix a bank, its cached features, and one nonempty closed convex feasible set $\mathcal{W}$ in the finite normalized simplex, and let $\mathcal{C} = \{\widehat{M}(\omega) : \omega \in \mathcal{W}\} \subset \mathcal{H}$. Then $\mathcal{C}$ is compact and convex, the exact selected predicted mean for a target y is the unique Hilbert projection $P_{\mathcal{C}}(y)$, and for any two targets y, y' solved over the same $\mathcal{C}$,*

$$\|P_{\mathcal{C}}(y) - P_{\mathcal{C}}(y')\|^2 \le \langle P_{\mathcal{C}}(y) - P_{\mathcal{C}}(y'),\, y - y'\rangle \le \|P_{\mathcal{C}}(y) - P_{\mathcal{C}}(y')\|\,\|y - y'\|, \tag{52}$$

so $\|P_{\mathcal{C}}(y) - P_{\mathcal{C}}(y')\| \le \|y - y'\|$ even if the minimizing weight vectors are not unique.

Proof. Since $\mathcal{W}$ is a nonempty closed subset of the finite normalized simplex, it is compact. The map $\widehat{M}$ is continuous and linear, so $\mathcal{C} = \widehat{M}(\mathcal{W})$ is nonempty, compact, and convex. Minimizing over $\omega \in \mathcal{W}$ is equivalent to minimizing $\frac{1}{2}\|y - z\|^2_{\mathcal{H}}$ over $z \in \mathcal{C}$. Thus every minimizing weight vector has the same predicted mean $P_{\mathcal{C}}(y)$.

The first inequality in (52) is the classical firm nonexpansiveness of the metric projection onto a nonempty closed convex set in a Hilbert space (Bauschke & Combettes, 2017, Proposition 4.16). The second follows from the Cauchy–Schwarz inequality. □

The stable object of the decision is thus its observation consequence, not an explanation in weight space. If an observation guard changes with the target, the compared feasible sets differ and this fixed-set statement no longer applies. For one bank and one round, a constrained optimum at which the radius is active with multiplier $\tau > 0$ is the penalized optimum with $\eta = (2\tau)^{-1}$, by the Karush–Kuhn–Tucker optimality conditions; the multiplier depends on the round, so a fixed radius and a fixed penalty define different outer trajectories, and the penalized results of this paper are not constrained results through this correspondence.

B.4.2 THE PENALIZED DECISION AND ITS NUMERICAL SOLUTION

Existence and endpoints. For the penalized decision $\min_{\omega\in\mathcal{A}} J_\eta(\omega)$, $J_\eta(\omega) = E_D(\omega) + \eta^{-1}R_\lambda(\omega)$ with $E_D(\omega) = F_D(\omega\lambda)$, linearity of K in its input law makes E_D convex whenever D is convex in the predicted law. On a finite bank the feasible set is compact, and in general, for a closed convex $\mathcal{A}$ and a proper, convex, lower semicontinuous E_D (as for MMD, through the bounded linear mean operator), the quadratic penalty makes the sublevel sets of J_η over $\mathcal{A}$ bounded, closed, and convex, hence weakly compact in $L^2(\lambda)$, so a minimizer exists; strict convexity makes it unique. Comparing it with the feasible unit weight gives the selected-progress inequality

$$F_D(\omega^\star\lambda) + \eta^{-1}R_\lambda(\omega^\star) \le F_D(\lambda). \tag{53}$$

Write $DE_D(\omega)[h] = -\langle s_\omega, h\rangle_\lambda$ on mean-zero directions and $\mathsf{P}_0 h = h - \mathbb{E}_\lambda h$. For $R_{\mathrm{Q}}(\omega) = \frac{1}{2}\|\omega - \mathbf{1}\|^2_\lambda$, first-order optimality reads

$$\left\langle -s_{\omega^\star} + \eta^{-1}(\omega^\star - \mathbf{1}),\, z - \omega^\star\right\rangle_\lambda \ge 0 \qquad (z \in \mathcal{A}), \tag{54}$$

and when normalization is the only active constraint it becomes $\omega^\star - \mathbf{1} = \eta\mathsf{P}_0 s_{\omega^\star}$, the quadratic endpoint of Proposition 1. For $R_{\mathrm{K}}(\omega) = \mathrm{KL}(\lambda\|\omega\lambda) = \mathbb{E}_\lambda[-\log\omega]$, an optimum with $\omega > 0$ λ-a.e. at which only normalization is active satisfies $-s_\omega - (\eta\omega)^{-1} + a = 0$; multiplying by ω and averaging gives $a = \mathbb{E}_{\omega\lambda}s_\omega + \eta^{-1}$ and hence $\omega - \mathbf{1} = \eta\omega(s_\omega - \mathbb{E}_{\omega\lambda}s_\omega)$, the backward-Euler Fisher–Rao endpoint.

Response to observation error. Assume a measurable positive-definite kernel with separable RKHS $\mathcal{H}$, $\mathsf{k}(y, y) \le \kappa^2$, and strongly measurable feature and conditional-mean maps, so the Bochner means ϕ, ψ, m_Q have norm at most κ. On $L^2_0(\lambda) = \{u : \mathbb{E}_\lambda u = 0\}$ set

$$Au = \mathbb{E}_\lambda[u\psi], \qquad A^*h(x) = \langle\psi(x) - \mathbb{E}_\lambda\psi, h\rangle_{\mathcal{H}}, \qquad r = m_{p^\star} - \mathbb{E}_\lambda\psi \tag{55}$$

(with $m_{p^\star} = m^\star$ of Appendix B.1); the adjoint is centered because feasible perturbations preserve mass. For quadratic movement the decision is the regularized least-squares problem $J_\eta(\mathbf{1}+u) = \frac{1}{2}\|r - Au\|^2_{\mathcal{H}} + \frac{1}{2\eta}\|u\|^2_\lambda$ over $\mathbf{1}+u \in \mathcal{A}$. When the unconstrained solution is feasible, its normal equation and the resolvent identity yield

$$u^\star = \eta(I+\eta A^*A)^{-1}A^*r, \qquad r^+ = r - Au^\star = (I+\eta AA^*)^{-1}r : \tag{56}$$

an observable singular mode with squared singular value σ^2 retains the fraction $(1+\eta\sigma^2)^{-1}$ of its discrepancy, and the gain from an observation-mean error into the weights is $\eta\sigma/(1+\eta\sigma^2) \le \sqrt{\eta}/2$. With constraints, let $\widehat{\omega}$ minimize the same quadratic with target $m_{p^\star}+e$, keeping λ, η, and $\mathcal{A}$ fixed, and put $\Delta\omega = \widehat{\omega} - \omega^\star$. Adding the two optimality inequalities gives $\eta^{-1}\|\Delta\omega\|^2_\lambda + \|A\Delta\omega\|^2_{\mathcal{H}} \le \langle e, A\Delta\omega\rangle_{\mathcal{H}}$, whence

$$\|\Delta\omega\|_\lambda \le \tfrac{\sqrt{\eta}}{2}\|e\|_{\mathcal{H}}, \qquad \|A\Delta\omega\|_{\mathcal{H}} \le \|e\|_{\mathcal{H}}, \qquad J_\eta(\widehat{\omega}) - J_\eta(\omega^\star) \le \tfrac{1}{2}\|e\|^2_{\mathcal{H}}. \tag{57}$$

For one dataset of n independent observations, $\mathbb{E}\|\widehat{m} - m_{p^\star}\|^2_{\mathcal{H}} \le \kappa^2/n$, so with exact conditional means $\mathbb{E}\|\widehat{\omega}-\omega^\star\|^2_\lambda \le \eta\kappa^2/(4n)$ and $\mathbb{E}[J_\eta(\widehat{\omega})-J_\eta(\omega^\star)] \le \kappa^2/(2n)$. These inequalities are pathwise in the data error, so the dataset may be reused across rounds and the bank may depend on it, while a kernel fitted to the dataset requires separate uniform control; they concern the penalized problem with fixed λ, η, and $\mathcal{A}$, whereas the stability of the constrained alternative is the projection statement of Section B.4.1.

The finite-bank quadratic form and solver certificates. For a uniform bank, the exact conditional embeddings form the positive-semidefinite Gram matrix $G_{ij} = \langle\psi(X_i), \psi(X_j)\rangle_{\mathcal{H}}$ and the data–model vector $h_i = \langle\psi(X_i), m_{p^\star}\rangle_{\mathcal{H}}$, and up to a weight-independent constant

$$J_\eta(\omega) = \frac{\omega^\top G\omega}{2N^2} - \frac{h^\top\omega}{N} + \frac{\|\omega-\mathbf{1}\|_2^2}{2\eta N}, \qquad N^{-1}\mathbf{1}^\top\omega = 1.$$

This has the form of the kernel mean matching quadratic program (Huang et al., 2006, Eq. (12)), with channel-averaged candidate features in place of covariate features and a movement penalty centered at $\mathbf{1}$ in place of a box constraint; weighted MMD quantization optimizes the same feature-mean agreement over particle weights (Belhadji et al., 2026). A cached plug-in Gram matrix $\widehat{G}_{ij} = \langle\widehat{\psi}_i, \widehat{\psi}_j\rangle$ stays positive semidefinite; its diagonal carries conditional-simulation variance, and (30) with Lemma 1 bounds the resulting mean error through the movement $\|\omega - \mathbf{1}\|_\lambda$ that the penalty controls. For recorded acquisition settings, the conditional score (23) has a normalized quadratic form $E(\omega) = \frac{1}{2}\langle\omega, B\omega\rangle_\lambda - \langle a, \omega\rangle_\lambda + \text{const}$ whose estimated coefficients induce the gradient error $e(\omega) = (\widehat{B}-B)\omega - (\widehat{a}-a)$, a coefficient-error interface distinct from the empirical-mean problem of Appendix B.1.

Suppose R_λ is m_R-strongly convex on the normalization hyperplane and the exact and approximate evidence problems are convex with the same feasible set and movement cost. If $e = \nabla\widehat{E}(\widetilde{\omega}) - \nabla E_D(\widetilde{\omega})$ at the approximate minimizer, monotonicity of the evidence gradient and the constraint normal cone give

$$\|\widetilde{\omega} - \omega^\star\|_\lambda \le \frac{\eta}{m_R}\|\mathsf{P}_0 e\|_\lambda, \tag{58}$$

with $m_R = 1$ for quadratic movement. For a bounded feasible set and a differentiable objective, the first-order gap

$$g(\omega) := \sup_{z\in\mathcal{A}}\langle\nabla J_\eta(\omega), \omega - z\rangle_\lambda, \qquad J_\eta(\omega) - \min_{z\in\mathcal{A}} J_\eta(z) \le g(\omega), \tag{59}$$

is a computable certificate: convexity proves the inequality, and on a finite bank the supremum is a linear program when $\mathcal{A}$ is polyhedral and a second-order cone program when a quadratic guard is present. Formed with $\widehat{J}_k$ over $\mathcal{A}_k$, the gap certifies the penalized objective excess $\varepsilon_{\mathrm{opt},k}$ in (27). For a feasible comparator c the penalized gap also gives $\langle\nabla J_\eta(\omega), c-\omega\rangle_\lambda \ge -g(\omega)$, the stationarity allowance used in (68) below. A standard inexact proximal-gradient argument in the setting studied by Schmidt et al. (2011) gives the following specialization. The evidence gradient is L-Lipschitz for the quadratic MMD objective and hence $1/L$-cocoercive by the Baillon–Haddad theorem (Bauschke & Combettes, 2009), so its gradient step is nonexpansive for $h \le 2/L$. The proximal map of $R_\lambda/\eta + \iota_{\mathcal{A}}$ is $q_h = (1+hm_R/\eta)^{-1}$-Lipschitz; therefore, with gradient errors e_j and proximal-output errors ξ_j, $\|\omega_J - \omega^\star\|_\lambda \le q_h^J\|\omega_0 - \omega^\star\|_\lambda + \sum_{j<J} q_h^{J-1-j}(hq_h\|\mathsf{P}_0 e_j\|_\lambda + \xi_j)$. An exact minimizer of an estimated objective and an inexact numerical solution are different defects: the first is charged through the target error in (57), the second through the gap (59).

B.4.3 POPULATION CLEAN-KL CONTRACTION UNDER THE PENALIZED STEP

The penalized decision admits a direct clean-KL statement when the population update is realized exactly and the truth is reachable by a bounded reweighting. It is a separate mechanism from the recovery theorem: its comparator must represent the truth exactly, so it needs two-sided overlap, whereas Theorem 1 needs only a forward explanation on the bank. It explains why feedback is clean-directed and is the natural point of comparison with the KL guarantees of related methods (Section B.4.5).

Consider the exact penalized population decision of Section 3.1 with quadratic movement and a normalized kernel, realized exactly, $\mu_{k+1} = \omega_k \mu_k$. Write $w_k = d\mu^\star/d\mu_k$ for the truth weight and $D_k = \mathrm{KL}(\mu^\star \| \mu_k)$ for clean error, and assume on a class $\mathcal{C}$ containing the trajectory

$$F_{\mathrm{M}}(\nu) \geq \alpha_{\mathrm{inv}} \,\mathrm{KL}(\mu^\star \| \nu), \qquad \nu \in \mathcal{C}, \quad \alpha_{\mathrm{inv}} > 0. \tag{60}$$

For $0 < \ell \leq 1$ define

$$H(\ell) = \frac{\ell - 1 - \log \ell}{(\ell - 1)^2}, \qquad H(1) = \tfrac{1}{2}, \tag{61}$$

and assume the two-sided overlap

$$w_k \in \mathcal{A}_k, \qquad \omega_k \geq \ell, \qquad w_k \leq U, \qquad UH(\ell) \leq 1 \quad \mu_k\text{-almost everywhere at every round.} \tag{62}$$

Since $H(\ell) \leq 1/(2\ell)$, the condition $U \leq 2\ell$ suffices; the sharper coefficient permits $U \leq 1.294$ at $\ell = \frac{1}{2}$ and $U \leq 1.588$ at $\ell = 0.7$. Because $\mathbb{E}_\mu w = 1$ forces $U \geq 1$, the condition requires $\ell \geq \frac{1}{2}$ under $U \leq 2\ell$, and a floor on every weight additionally requires $w_k \geq \ell$: a quantitative density-ratio condition, not implied by small KL or W_2 in a continuous model.

Proposition 5 (Observation feedback contracts clean-law error)**.** *Fix $\eta > 0$ and an initial law with finite D_0. Under exact decisions and realization and (60)–(62) along the trajectory, with the population observation mean,*

$$D_T \leq (1 + 2\eta\alpha_{\mathrm{inv}})^{-T} D_0. \tag{63}$$

If instead every round uses the same empirical mean from n independent observations, with the kernel and μ_0 fixed independently of that dataset and the trajectory conditions holding almost surely, then

$$\mathbb{E} D_T \leq q_\eta^T D_0 + \frac{1 - q_\eta^T}{2\alpha_{\mathrm{inv}} n}, \qquad q_\eta = (1 + \eta\alpha_{\mathrm{inv}})^{-1}. \tag{64}$$

Proof. Suppress k; let ω be the selected weight, $u = \omega - 1$, and $w = d\mu^\star/d\mu \in \mathcal{A}$ the feasible truth weight; write $E(\omega) = F_{\mathrm{M}}(\omega\mu)$ and $\mu^+ = \omega\mu$. Testing the first-order optimality condition of the penalized convex minimum, in the same form in $L^2(\mu)$, toward w gives

$$\underbrace{\langle u, w - \omega \rangle_\mu}_{\mathbb{E}_{\mu^\star} u - \|u\|_\mu^2} \geq -\eta \, DE(\omega)[w - \omega]. \tag{65}$$

Moving the weights from ω toward w moves the predicted mean from $m_{K\mu^+}$ exactly to $m_{p^\star}$, so with $r^+ = m_{p^\star} - m_{K\mu^+}$ and linearity of the mean in the weights,

$$DE(\omega)[w - \omega] = -\langle r^+, \mathbb{E}_\mu[(w - \omega)\psi] \rangle_{\mathcal{H}} = -\|r^+\|_{\mathcal{H}}^2 = -2F_{\mathrm{M}}(\mu^+). \tag{66}$$

The derivative is evaluated at the proposed weights, so it measures the discrepancy remaining after the decision: this is where feedback enters. Exactly, $\mathrm{KL}(\mu^\star\|\mu) - \mathrm{KL}(\mu^\star\|\mu^+) = \mathbb{E}_{\mu^\star} \log(1 + u)$, and for $u \geq \ell - 1$ the integral identity

$$u - \log(1 + u) = u^2 \int_0^1 \frac{t}{1 + tu} \, dt \leq H(\ell) \, u^2$$

holds because the integrand is decreasing in u and the integral equals $H(\ell)$ at $u = \ell - 1$. With $w \leq U$ this gives $\mathbb{E}_{\mu^\star}[u - \log(1 + u)] \leq UH(\ell)\|u\|_\mu^2$, and combining with (65)–(66),

$$\mathrm{KL}(\mu^\star\|\mu) - \mathrm{KL}(\mu^\star\|\mu^+) \geq 2\eta F_{\mathrm{M}}(\mu^+) + \big[1 - UH(\ell)\big] \|\omega - 1\|_\mu^2. \tag{67}$$

The overlap condition makes the last term nonnegative; applying (60) at μ^+ gives $\mathrm{KL}(\mu^\star\|\mu^+) \leq \mathrm{KL}(\mu^\star\|\mu)/(1 + 2\eta\alpha_{\mathrm{inv}})$, and exact realization identifies μ^+ with the next incoming law, proving (63). The empirical statement is Lemma 2 below with vanishing stationarity and realization allowances, together with $\mathbb{E}\delta_n^2 \leq 1/n$. □

The proof needs only the averaged curvature condition $\mathbb{E}_{\mu^\star}[u - \log(1+u)] \le \|u\|_\mu^2$; uniform ratio and floor bounds are convenient sufficient conditions. For other convex observation criteria, $E(w) = 0$ and convexity give $-DE(\omega)[w-\omega] \ge F_D(\mu^+)$, so the same argument yields the factor η in place of 2η.

One observation dataset and imperfect updates. Freeze one dataset, with $\widehat{m}$ and $\delta_n = \|e_n\|_{\mathcal{H}}$ as in Section B.1.3 and $e_n = \widehat{m} - m^\star$, and let every round depend on it. Let $\widehat{\omega}_k$ be feasible with stationarity error $\zeta_k \ge 0$ in the truth direction,

$$\Big\langle \nabla \widehat{E}_k(\widehat{\omega}_k) + \frac{\widehat{\omega}_k - 1}{\eta},\ w_k - \widehat{\omega}_k \Big\rangle_{\mu_k} \ge -\zeta_k, \tag{68}$$

where $\widehat{E}_k(\omega) = \frac{1}{2}\|\widehat{m} - \mathbb{E}_{\mu_k}[\omega\psi]\|_{\mathcal{H}}^2$; the first-order gap (59) supplies a computable ζ_k. Set $\mu_k^+ = \widehat{\omega}_k \mu_k$ and allow a realization defect in the clean metric,

$$\mathrm{KL}(\mu^\star \| \mu_{k+1}) \le \mathrm{KL}(\mu^\star \| \mu_k^+) + \gamma_k^{\mathrm{KL}}, \qquad \gamma_k^{\mathrm{KL}} \ge 0. \tag{69}$$

Lemma 2 (Perturbed clean contraction)**.** *Under the overlap, finite-divergence, and inverse conditions of Proposition 5 at every incoming and selected law, with $\widehat{\omega}_k$ in place of ω_k, and under (68)–(69), with $q_\eta = (1+\eta\alpha_{\mathrm{inv}})^{-1}$,*

$$D_T \le q_\eta^T D_0 + \frac{\delta_n^2}{2\alpha_{\mathrm{inv}}}(1 - q_\eta^T) + \sum_{k<T} q_\eta^{T-1-k}\big(q_\eta \eta \zeta_k + \gamma_k^{\mathrm{KL}}\big). \tag{70}$$

Proof. With $r_k^+ = m_{p^\star} - m_{K\mu_k^+}$ the derivative toward truth equals $-\|r_k^+\|^2 - \langle e_n, r_k^+\rangle$. Repeating (65) with its stationarity allowance and the logarithm bound gives $D_k - \mathrm{KL}(\mu^\star\|\mu_k^+) \ge \eta\{\|r_k^+\|^2 + \langle e_n, r_k^+\rangle - \zeta_k\}$, and Young's inequality bounds the braces below by $\frac{1}{2}\|r_k^+\|^2 - \frac{1}{2}\delta_n^2 - \zeta_k$. Inverse stability and (69) give $D_{k+1} \le q_\eta D_k + \frac{1-q_\eta}{2\alpha_{\mathrm{inv}}}\delta_n^2 + q_\eta\eta\zeta_k + \gamma_k^{\mathrm{KL}}$, and the geometric sum proves the claim. □

No independence between the dataset and later iterates is used; the factor q_η is slower than $(1 + 2\eta\alpha_{\mathrm{inv}})^{-1}$ because Young's inequality spends half of the exact decrease to absorb the data error. The realization allowance γ_k^{KL} is in clean KL; small W_2 or MMD error of the outgoing generator does not supply it, which is why the recovery theorem applies inverse stability only to the output and never compares the truth with an atomic bank in this metric.

A warm-start basin in which the overlap condition is invariant. On a finite clean state space with truth masses $\pi_i > 0$ and $p_{\min} = \min_i \pi_i$, fix $0 < r_0 < 1/3$ and a floor $\ell \in [\frac{1}{2(1-r_0)}, \frac{1}{1+r_0}]$. If exact penalized population decisions are used over normalized weights $\omega_i \ge \ell$, with any further convex constraints admitting the truth comparator, and $\mathrm{KL}(\pi\|\mu_0) \le 2r_0^2 p_{\min}^2$, then every selected law satisfies $(1-r_0)\pi_i \le \mu_{k,i} \le (1+r_0)\pi_i$ and (62) holds at every round: Pinsker's inequality turns the KL bound into $|\mu_i - \pi_i| \le r_0 p_{\min} \le r_0\pi_i$, so the truth weight lies in $[1/(1+r_0), 1/(1-r_0)]$, the chosen interval makes it feasible with $U = 1/(1-r_0) \le 2\ell$, and (67) keeps clean KL from increasing, so induction proves invariance. If the centered conditional-embedding operator at π has squared smallest singular value $\sigma_{\min}^2 > 0$, Section B.4.5 supplies $\alpha_{\mathrm{inv}} = (1-r_0)\sigma_{\min}^2$ and $\mathrm{KL}(\pi\|\mu_T) \le [1 + 2\eta(1-r_0)\sigma_{\min}^2]^{-T}\,\mathrm{KL}(\pi\|\mu_0)$. The floor can be set without knowing the truth, which enters only through the initialization.

B.4.4 Progress under a genuinely fixed penalty

Theorem 1 prices the comparator's movement into the target accuracy and obtains a geometric approach to that level. The following $O(1/T)$ statement is complementary: it charges a fixed penalty's cost against the objective improvement instead, so the guaranteed fraction of the correction decreases with the residual, and it holds for a general convex observation criterion and finite banks, with allowances measured in the criterion F_D rather than in the observation-mean norm.

At round k keep the incoming law μ_k, its bank λ_k, the selected law $\mu_k^+ = \widehat{\omega}_k\lambda_k$, and the outgoing law μ_{k+1}. On the closed convex feasible set $\mathcal{A}_k$ containing $\mathbf{1}$ write $E_k(\omega) = F_D(\omega\lambda_k)$ and $\widehat{J}_k(\omega) =$

$\widehat{E}_k(\omega) + R_{\lambda_k}(\omega)/\eta_k$, with $R_{\lambda_k} \geq 0$, $R_{\lambda_k}(\mathbf{1}) = 0$, and a minimizer. The empirical score may omit a weight-independent constant C_k; assume the uniform evidence error, the attainable gain and optimization accuracy, and the bank and generator allowances

$$\sup_{\omega \in \mathcal{A}_k} |\widehat{E}_k(\omega) - E_k(\omega) - C_k| \leq \varepsilon_k^{\mathrm{ev}}, \tag{71}$$

$$G_k = \widehat{E}_k(\mathbf{1}) - \min_{\mathcal{A}_k} \widehat{J}_k \geq 0, \qquad \widehat{J}_k(\widehat{\omega}_k) - \min_{\mathcal{A}_k} \widehat{J}_k \leq \varepsilon_k^{\mathrm{opt}}, \tag{72}$$

$$|F_D(\lambda_k) - F_D(\mu_k)| \leq \varepsilon_k^{\mathrm{bank}}, \qquad |F_D(\mu_{k+1}) - F_D(\mu_k^+)| \leq \varepsilon_k^{\mathrm{gen}}. \tag{73}$$

Lemma 3 (Selected gain minus realization cost)**.** *With* $\Xi_k = \varepsilon_k^{\mathrm{opt}} + 2\varepsilon_k^{\mathrm{ev}} + \varepsilon_k^{\mathrm{bank}} + \varepsilon_k^{\mathrm{gen}}$,

$$F_D(\mu_{k+1}) \leq F_D(\mu_k) - G_k - \eta_k^{-1} R_{\lambda_k}(\widehat{\omega}_k) + \Xi_k. \tag{74}$$

Proof. Approximate minimization gives $\widehat{E}_k(\widehat{\omega}_k) \leq \widehat{E}_k(\mathbf{1}) - G_k - R_{\lambda_k}(\widehat{\omega}_k)/\eta_k + \varepsilon_k^{\mathrm{opt}}$. Replacing the two empirical scores by their population values costs $2\varepsilon_k^{\mathrm{ev}}$, the constant C_k cancels, and (73) gives the result. □

Suppose a feasible weight c_k explains the observations to accuracy $\varepsilon_k^{\mathrm{cov}}$ at a controlled movement cost,

$$F_D(c_k \lambda_k) \leq \varepsilon_k^{\mathrm{cov}}, \qquad R_{\lambda_k}(\mathbf{1} + t(c_k - \mathbf{1})) \leq C_{\mathrm{cov}} t^2/2 \quad (0 \leq t \leq 1), \tag{75}$$

which for quadratic movement reads $\|c_k - \mathbf{1}\|_{\lambda_k}^2 \leq C_{\mathrm{cov}}$. A partial move toward c_k improves a convex observation score linearly in t while its cost grows quadratically, which is the source of the rate.

Theorem 4 (Repeated progress with imperfect computation)**.** *Assume* (71)–(75)*, convex* $\widehat{E}_k$*,* $\eta_k \geq \eta_{\min} > 0$*, and* $0 \leq f_k := F_D(\mu_k) \leq M$*. Put* $b_k = \varepsilon_k^{\mathrm{cov}} + 2\varepsilon_k^{\mathrm{ev}} + \varepsilon_k^{\mathrm{bank}}$ *and* $a = \min\{\eta_{\min}/(2C_{\mathrm{cov}}), 1/(2M)\}$*. Then*

$$f_{k+1} \leq f_k - a(f_k - b_k)_+^2 + \Xi_k. \tag{76}$$

Exact coverage and computation give $f_T \leq f_0/(1 + aT f_0)$*. If* $b_k \leq b$ *and* $\Xi_k \leq \Xi$ *for all* k*, then with* $f_\infty = b + \sqrt{\Xi/a}$ *and* $y_0 = (f_0 - f_\infty)_+$*,*

$$f_T \leq f_\infty + \frac{y_0}{1 + aT y_0}, \qquad \textit{and in particular} \qquad \limsup_k f_k \leq \min\{M, f_\infty\}. \tag{77}$$

Proof. Convexity of $\widehat{E}_k$ along the feasible segment toward c_k gives $G_k \geq t[\widehat{E}_k(\mathbf{1}) - \widehat{E}_k(c_k)] - C_{\mathrm{cov}} t^2/(2\eta_k) \geq t(f_k - b_k) - C_{\mathrm{cov}} t^2/(2\eta_k)$; the choice $t = 2a(f_k - b_k)_+$ is feasible since $2aM \leq 1$ and $2aC_{\mathrm{cov}} \leq \eta_{\min}$, and yields $G_k \geq a(f_k - b_k)_+^2$; Lemma 3 proves (76). For zero errors the reciprocal inequality $1/f_{k+1} \geq 1/f_k + a$ gives the ideal rate. For bounded errors, if $f_\infty \geq M$ there is nothing to prove, so let $f_\infty < M$. The scalar map $\Gamma(x) = x - a(x - b)_+^2 + \Xi$ is nondecreasing on $[0, M]$ because $2aM \leq 1$, and for $x \geq f_\infty$, writing $y = x - f_\infty$ and $x - b = y + \sqrt{\Xi/a}$, one has $\Gamma(x) - f_\infty = y - ay^2 - 2y\sqrt{a\Xi} \leq y - ay^2$, while $\Gamma(x) \leq \Gamma(f_\infty) = f_\infty$ for $x < f_\infty$. Hence $y_{k+1} \leq y_k - a y_k^2$ for $y_k = (f_k - f_\infty)_+$, and the reciprocal argument gives (77). □

The threshold b_k records the resolution at which the bank and its evidence reveal an improvement, and Ξ_k the cost of carrying it through one round; the square-root floor is a consequence of the available squared-objective descent estimate, not a lower bound on achievable accuracy. Applying the theorem to the single empirical objective $F_n(\mu) = \frac{1}{2}\|\widehat{m} - m_{K\mu}\|_{\mathcal{H}}^2$ and converting once at the end, as in Appendix B.1.4, permits reuse of one dataset in this branch as well: with a normalized kernel, quadratic movement, exact decisions, embeddings, and realization, the truth comparator $w_k \in \mathcal{A}_k$ with $\|w_k - \mathbf{1}\|_{\mu_k}^2 \leq C_{\mathrm{cov}}$, and $a = \min\{\eta_{\min}/(2C_{\mathrm{cov}}), 1/4\}$, the argument applied to F_n with $F_n(\mu^\star) = \delta_n^2/2$ gives $F_{\mathrm{M}}(\mu_T) \leq 2F_n(\mu_T) + \delta_n^2 \leq 2g_0/(1 + aTg_0) + 2\delta_n^2$ for $g_0 = (F_n(\mu_0) - F_n(\mu^\star))_+$, hence $\mathbb{E}F_{\mathrm{M}}(\mu_T) \leq 2/(aT) + 2/n$ for $T \geq 1$.

B.4.5 When the observations determine the clean law

Both recovery statements convert remaining observation error into clean error through (60). The constant depends on the channel, the kernel, and the class of laws, and a characteristic kernel does not supply it: a characteristic kernel identifies the observation law (Sriperumbudur et al., 2010), and when K is injective the observation law in turn identifies the clean law, as Bora et al. (2018) establish for several measurement models; neither fact gives a uniform inverse rate.

On a finite clean state space with positive truth masses, write $\nu = (1+h)\mu^\star$ with $\mathbb{E}_{\mu^\star} h = 0$ and $\|h\|_\infty \le r_0 < 1$, and let $A_\star h = \mathbb{E}_{\mu^\star}[h\psi]$ be the observation-change operator at the truth. If its smallest squared singular value on mean-zero weights is $\sigma^2_{\min} > 0$, then $F_{\mathrm{M}}(\nu) = \frac{1}{2}\|A_\star h\|^2_{\mathcal{H}} \ge \frac{1}{2}\sigma^2_{\min}\|h\|^2_{\mu^\star}$, while the logarithm bound centered at the truth gives $\mathrm{KL}(\mu^\star\|\nu) \le \|h\|^2_{\mu^\star}/(2(1-r_0))$, so

$$\alpha_{\mathrm{inv}} = (1-r_0)\sigma^2_{\min} \quad \text{is valid on this class.} \tag{78}$$

The Gram matrix of conditional embeddings makes visible which clean redistributions survive corruption. The restriction to a class is essential: with bounded features on an unrestricted infinite-dimensional weight space, $\mathrm{tr}(A_\star^* A_\star) = \mathbb{E}_{\mu^\star}\|\psi - \mathbb{E}_{\mu^\star}\psi\|^2 \le 1$, so $A_\star$ is Hilbert–Schmidt and cannot have a uniformly positive smallest singular value in all directions. The n^{-1} term of Theorem 1 is therefore not a dimension-free minimax statement about image distributions: the inverse constant, the coverage cost, and the realization accuracy determine which clean-law resolution is covered, and a kernel on a pretrained embedding moves this class with the encoder (Appendix A.5). When no quantitative inverse bound is available, a qualitative statement remains: if the generated laws and $\mu^\star$ belong to a compact metric class $(\mathcal{C}, d)$ on which the criterion is lower semicontinuous with $\mu^\star$ as its unique zero, then $F_D(\mu_k) \to 0$ implies $d(\mu_k, \mu^\star) \to 0$, since a subsequence separated from $\mu^\star$ would have a convergent subsequence whose limit is another zero; equivalently $\varpi(t) = \sup\{d(\nu,\mu^\star) : \nu \in \mathcal{C}, F_D(\nu) \le t\} \to 0$ as $t \downarrow 0$, which with Theorem 2 or Theorem 4 gives qualitative clean recovery under vanishing coverage and implementation errors, in the form $\mathbb{E}\, d(\mu_T, \mu^\star) \to 0$ since d is bounded on $\mathcal{C}$. The class must exclude what the kernel cannot see: an additive kernel over coordinate blocks assigns identical observation embeddings to laws that differ only in cross-block dependence, so $\alpha_{\mathrm{inv}} = 0$ on any class containing such pairs.

Comparison with the KL guarantees of SCSI and DiffEM. In the ICLR conference version of SCSI (Modi et al., 2026), Theorem 3 assumes a finite restricted inverse condition number $\chi = \sup_{\rho\in\mathcal{S}} \mathrm{KL}(\pi\|\rho)/\,\mathrm{KL}(K\pi\|K\rho)$ on a class $\mathcal{S}$, a KL-to-drift Lipschitz condition with constant L_ε, a positive diffusion coefficient, and $L_\varepsilon\chi < 4$; its approximate-field error is measured along an oracle interpolation rather than the training path (Remark 4), and Remark 5 excludes the probability-flow ODE limit from that KL proof. In the arXiv version (v2) the same result is numbered Theorem 9 with Remarks 10–11. These assumptions and ours share a purpose, quantitative inverse stability on a specified class, and differ in what is controlled: SCSI controls how an inverse diffusion bridge changes with its current law, whereas Theorem 1 controls whether a bank offers an affordable forward-observation correction and whether the fitted generator realizes it, using the residual-flow bound on the training interpolation and no positive sampling diffusion, while still requiring well-posedness and Lipschitz control of that flow. Neither set of hypotheses implies the other. DiffEM's geometric clean-law result (Hosseintabar et al., 2025, Proposition 3) likewise assumes an inverse comparison, $\mathrm{KL}(P_X^\star\|P) \le \kappa\ \mathrm{KL}(P_Y^\star\|Q_\#P)$ on a KL neighborhood of the truth (Assumption 1), and is stated in terms of the latent-space error $\widetilde{\varepsilon}^{(k)}_{\mathrm{SM}} = \mathbb{E}_{(X,Y)\sim P^\star} \log \frac{P^{(k)}(X|Y)}{q_{\theta^{(k+1)}}(X|Y)}$. Its monotonicity result instead uses $\varepsilon^{(k)}_{\mathrm{SM}} = \mathbb{E}_{Y\sim P_Y^\star}\,\mathrm{KL}(q_{\theta^{(k+1)}}(\cdot\mid Y)\|P^{(k)}(\cdot\mid Y))$ (Lemma 1). Its Section 3 analyzes a variant that draws fresh corrupted observations at every iteration; the fixed-dataset argument of Appendix B.1.4 is a concrete point of contrast. The specific ratio conditions above belong to the quadratic-movement proof and are stated for the classes where they are verified, not asserted for image models.

C Experiments: implementation details and additional results

This appendix gives the protocols and additional results of the synthetic study (Section C.1) and of the masked CIFAR-10 study (Section C.2), and a restoration study on astronomical images (Section C.3). Our code is available at `https://anonymous.4open.science/r/frrf-C05E`.

C.1 SYNTHETIC STUDY

C.1.1 POPULATION LEARNING

Data. The signal is $X_{1:2} = (\cos\Theta, \sin\Theta) + 0.05\epsilon_c$, $\epsilon_c \sim \mathcal{N}(0, I_2)$, with angular density

$$f_\Theta(\theta) = \frac{0.05}{2\pi} + 0.95\sum_{j=1}^{3}\pi_j \frac{\exp\{6\cos(\theta - \vartheta_j)\}}{2\pi I_0(6)}, \qquad \pi = (0.5, 0.3, 0.2), \quad \vartheta = (0, \tfrac{2\pi}{3}, \tfrac{4\pi}{3}),$$

where I_0 is the modified Bessel function of the first kind, order zero. The other 14 coordinates are independent $\mathcal{N}(0, 0.5)$ variables, independent of the signal. Observations follow $Y = X + 0.35Z$, $Z \sim \mathcal{N}(0, I_{16})$. Independent noisy datasets contain 8,192 observations for fitting, 8,192 for estimating the observation metric, and 4,096 for validation; a separate reporting dataset contains 32,768 observations. Clean samples and the signal-coordinate indices are used only for evaluation.

Generator and fitting. All methods start from $\mathcal{N}(0, 0.5I_{16})$ and train all coordinates. The velocity is the analytic initial-Gaussian field plus a zero-output, three-hidden-layer, width-128 SiLU network. Each of 60 rounds draws a fresh bank $\lambda_k = N^{-1}\sum_i \delta_{X_i}$ of $N = 2{,}048$ generated candidates, selects weights, and trains on the straight Gaussian-to-candidate interpolation. We use Adam at 10^{-3}, batch size 512, 2,048 updates, a fresh optimizer each round, and 64-step Heun sampling. All use residual FM with the incoming field as reference, $\widetilde{v} = v_k$.

Observation kernel. A fixed metric estimated from noisy data emphasizes non-Gaussian coordinates without supplying their indices. Let $\widehat{\kappa}_j$ be coordinate j's empirical excess kurtosis on the independent metric dataset, and let $n_{\mathrm{rep}} = 8{,}192$. Set

$$q_j = \big(|\widehat{\kappa}_j| - 3\sqrt{24/n_{\mathrm{rep}}}\big)_+, \qquad \gamma_j = \frac{0.05}{16} + 0.95\frac{q_j}{\sum_r q_r}, \qquad \widehat{\mathcal{I}} = \{j : q_j > 0\}.$$

Here $(a)_+ = \max(a, 0)$; if every q_j is zero, use $\gamma_j = 1/16$ and select all coordinates. The MMD kernel is

$$k(y, y') = \frac{1}{3}\sum_{c\in\{0.5,1,2\}} \exp\left[-\frac{\sum_j \gamma_j (y_j - y'_j)^2}{2c^2\ell_{\mathrm{ker}}^2}\right],$$

where ℓ_{ker} is the median weighted distance between disjoint pairs in that same dataset. The metric, bandwidth, and selected set are frozen and shared. Gaussian kernel expectations are computed analytically with the metric-induced covariance.

Weight decisions. All weights satisfy $N^{-1}\sum_i \omega_i = 1$ and $0.1 \le \omega_i \le 40$. FRRF minimizes the empirical penalized MMD decision $\widehat{F}_M(\omega\lambda_k) + (2\eta)^{-1}\|\omega - 1\|^2_{\lambda_k}$. Forward MMD instead forms the centered score $h = s_{\lambda_k} - \mathbb{E}_{\lambda_k} s_{\lambda_k}$ at unit weights and chooses $\omega = 1 + \alpha h$ by exact line search on $\widehat{F}_M$, over feasible $\alpha \ge 0$. Likelihood feedback replaces $\widehat{F}_M$ by the empirical negative log likelihood of the Gaussian candidate mixture on $\widehat{\mathcal{I}}$, with the same quadratic penalty and residual fitting. Weights apply to full 16-dimensional candidates.

Update strength. Feedback strengths are calibrated on independent pilots to an initial weight change $\|\omega - 1\|_{\lambda_0} \simeq 1.20$, then frozen: $\eta = 3673.73$ for FRRF and 23.300 for likelihood. The smaller likelihood setting uses $\eta = 2.7555$ (initial scale 0.60). Fixed forward coefficients 19.778 and 52.209 target initial scales 0.60 and 1.20. The line-searched forward step is chosen each round, without movement matching.

Checkpoint selection. Students at 128, 512, 1,024, and 2,048 updates are compared with the selected law $\widehat{\omega}\lambda_k$, never the truth. On 8,192 source draws, screen coordinatewise and total second moments at the selected-law value plus the larger of 10% and five standard errors. Choose the eligible student with smallest MMD to the selected law (2,048 draws; coordinates $(\sqrt{\gamma_j}x_j)_j$; bandwidths $0.1, 0.25, 0.5, 1$). Retain the incoming field if none qualifies or the best is worse by more than two jackknife standard errors.

Evaluation. Each of five paired runs supplies 8,192 generated samples. Radial W_1 compares the distributions of $r = \|X_{1:2}\|$; angular errors compare probabilities in three mode-centered 120° sectors or 48 equal bins. Nuisance-variance error is $\max_{j>2}|\widehat{\mathrm{Var}}(X_j) - 0.5|$. These coordinates start correct, so the metric measures preservation. The display uses the first 4,000 draws of the predeclared first run.

C.1.2 Realizing learned corrections

This control gives the *same learned weights* to weighted and residual FM. Five incoming two-dimensional generators are taken after ten rounds on the smooth circle with density $(1+0.45\cos\theta+0.30\cos 3\theta)/(2\pi)$, separately from the study above. For each generator, MMD weights are estimated on 8,192 candidates from three independent 2,048-observation datasets. Set $\omega_\alpha = 1 + \alpha(\widehat{\omega} - 1)$ for $\alpha \in \{0, 0.1, 0.25, 0.5, 1\}$ to vary the magnitude of that learned change. Both objectives receive identical path tuples, two paired training streams, and 2,048 updates; residual FM uses $\widetilde{v} = v_k$.

Let $\chi(x)$ contain the sines and cosines of angular orders 1–6, $r - 1$, and $(r - 1)^2$, with $r = \|x\|$ in this two-dimensional test. Define the requested and realized feature changes by

$$\delta_\alpha = \mathbb{E}_{\omega_\alpha \lambda_k}\chi - \mathbb{E}_{\lambda_k}\chi, \qquad \Delta_\alpha = \mathbb{E}_{\mu_{k+1}}\chi - \mathbb{E}_{\mu_k}\chi.$$

We measure change orthogonal to that request and its useful transmission:

$$e_\perp = \left\| \Delta_\alpha - \frac{\langle \Delta_\alpha, \delta_\alpha \rangle}{\|\delta_\alpha\|^2}\delta_\alpha \right\|, \qquad \text{transmission} = \frac{\sum_s \overline{\langle \Delta_\alpha, \delta_\alpha \rangle}_s}{\sum_s \overline{\|\delta_\alpha\|^2}_s}.$$

Bars average repetitions within incoming generator s; plotted lines average the five generator means. Transmission one denotes full realization along the requested direction. At $\alpha = 0$, transmission is undefined and $e_\perp = \|\Delta_\alpha\|$ measures unrequested change. At $\alpha = 0.1$, residual FM gives $e_\perp = 0.0061$ and transmission 0.923, versus 0.0368 and 0.564 for weighted FM. At full strength, transmission is 0.939 and 0.930, respectively. These metrics assess the stated features, not the full distribution.

C.1.3 What the generated population preserves

FRRF improves radial error over the line-searched forward method on every run (mean difference -0.1630, paired 95% bootstrap interval $[-0.1698, -0.1554]$). Before the first generator fit, feedback already selects a law with radial error 0.158, versus 0.344 for line search (incoming bank: 0.354). Whole-round times average 42.1 and 38.7 seconds, including selection and evaluation.

Method	Initial scale	Radial $W_1\downarrow$	Sector $\ell_1\downarrow$	Angular $\ell_1\downarrow$	Nuisance variance error$\downarrow$
FRRF	1.20	0.0293	0.0247	0.1176	0.0771
Forward MMD + line search	—	0.1923	0.0266	0.1540	0.0271
Likelihood feedback	1.20	0.0192	0.0248	0.1928	0.0794
Likelihood feedback	0.60	0.0373	0.0225	0.1156	0.0484
Forward MMD, fixed step	0.60	0.2554	0.1196	0.2570	0.1086
Forward MMD, fixed step	1.20	0.4372	0.3766	0.5944	0.8969

Table 3: **Recovery and preservation at round 60.** Means over five paired runs at $d = 16$. Initial scale is the pilot target for $\|\omega - 1\|_{\lambda_0}$, not a constraint on later rounds. Angular error uses 48 bins; sector error uses the three mode-centered sectors.

Likelihood feedback at the larger setting yields a thinner ring; FRRF has the more accurate angular profile (Table 3). The smaller likelihood setting trades radial concentration for similar angular accuracy. Both feedback procedures retain signal–nuisance dependence, and FRRF's radius standard deviation is 0.092 versus 0.050 for truth. Visible geometry and mode masses therefore do not establish full-population recovery.

C.2 Masked CIFAR-10: protocol and additional results

C.2.1 Protocol

Data and channel. We split the 50,000 CIFAR-10 training images into 45,000 fitting and 5,000 held-out images; the test set is unused. Each image produces one observation,

$$Y = M \odot X + (1 - M) \odot (c + s \odot E),$$

where M is a Bernoulli$(1/2)$ mask shared across color channels and E is standard Gaussian. The per-channel mean c and standard deviation s set the filler scale, following Modi et al. (2026). The mask is recorded, and kept pixels are exact. Model fitting uses the observations and forward channel simulations; clean images enter only the evaluators.

Observation-only initialization. A blind-prediction U-Net completes the fitting observations (Batson & Royer, 2019). We partition pixels into sixteen interlaced groups (row and column modulo four), hide one group, and score predictions on its originally observed pixels. Completions retain all observed values. The predictor uses base width 32, multipliers $(1, 2, 4)$, 8,000 Adam updates, batch size 128, and learning rate 10^{-3} (0.1 A100-hours). A 96-channel U-Net (Dhariwal & Nichol, 2021) then learns the 45,000 completions by straight-path FM: 16,000 AdamW updates, batch size 256, learning rate 2×10^{-4}, and parameter exponential moving average (EMA) with half-life 1,000 updates (Loshchilov & Hutter, 2019). This 3.2-hour fit initializes the straight-path FRRF and fixed-completion FM variants (held-out FID 23.9 and 25.1). Baseline-specific initialization is in Section C.2.6.

Select, then realize. Each round draws $N = 4{,}096$ candidates and $m = 16$ observations per candidate. An RBF kernel on fixed Inception pool features (Szegedy et al., 2016) is approximated by 1,024 random Fourier features at the median bandwidth (Rahimi & Recht, 2007). Averaging over simulations gives the candidate features; averaging the same map over fitting observations gives the target mean. The constrained decision of Appendix A.3 uses mean-one weights, $0.1 \le \omega_i \le 40$, and $\|\omega - \mathbf{1}\|_\lambda^2 \le 1$. A raw-pixel MMD guard, estimated with eight simulations per candidate, enforces $H(K(\omega\lambda)) \le H(K\lambda)$ alongside the perceptual objective. The convex first-order gap measures solver accuracy.

We take eight SGD steps on (14) with the incoming field as reference, momentum 0.9, batch size 1,024, and step size $\min\{0.05, 0.25/L\}$; L is the largest Hessian eigenvalue estimated on 512 tuples. Candidate sampling mixes equal parts uniform and probabilities proportional to $|\omega_i - 1|$, with importance correction. Candidates use 32 second-order multistep DPM-Solver++ steps (Lu et al., 2022), or 64 Heun steps in the sampler ablation (Karras et al., 2022); evaluation always uses 64 Heun steps. A round costs approximately seven or eleven minutes, respectively, on one A100.

Evaluation. We evaluate after 5, 10, 20, 30, and 36 refinement hours. $\text{FID}_{10\text{k}}$ uses 10,000 samples against the 5,000 held-out images; $\text{FID}_{50\text{k}}$ uses 50,000 samples against the full training split (Heusel et al., 2017). From 20 hours onward, held-out metrics average two latent draws. We also report KID ($\times 10^3$; $\text{KID}_{10\text{k}}$ on the same 10,000 samples), Inception score, and precision, recall, density, and coverage ($k = 5$) (Bińkowski et al., 2018; Salimans et al., 2016; Kynkäänniemi et al., 2019; Naeem et al., 2020). Per-round checks use 2,000 samples for KID, relative per-channel standard-deviation error (*dispersion error*), total variation, and highest-band discrete-cosine-transform (DCT) energy share relative to clean references.

The observation check re-corrupts 10,000 samples and compares them with held-out observations by raw-pixel MMD at three bandwidths. Fitting-versus-held-out observations provide the bootstrap reference; z_{obs} (z in the tables) is the standardized gap at the finest bandwidth. Passing requires agreement with the 95% reference interval at all three scales. Anchored restoration PSNR, LPIPS (Zhang et al., 2018), and posterior-mean-estimate PSNR use 1,024 held-out observations.

Compute and selection. A100-hours include initialization and refinement on one A100-SXM4-40GB, excluding evaluation. Anchored initialization uses a 3.3-hour field fit plus the completion predictor. Our configurations use two training seeds and the declared horizons. A rule fixed before the 30-hour evaluations selects the lowest held-out FID among configurations passing the observation check on at least one draw. Clean held-out images therefore enter configuration selection through FID, not parameter fitting. The rule selects 32-step candidates; selecting by the smallest $|z|$ alone selects 64-step candidates.

C.2.2 Full metric set

Table 4 reports both training seeds and the complementary fidelity and diversity metrics for every configuration. Table 1 shows FRRF with 64-step candidates at 36 hours, the configuration inside the observation interval at all scales; the FID rule of Section C.2.1 selects the 32-step configuration, 0.5 lower in $\text{FID}_{50\text{k}}$, which leaves the interval at the finest scale (z 3.8). Table 1 also shows the anchored alternation at 30 hours, completion-initialized SCSI, and the continued DiffEM run; the

remaining rows are their variants. Feedback maintains recall while continued fitting to fixed completions loses it: recall is 0.66 for FRRF at 36 hours and 0.55 for fixed-completion FM at 30 hours. The early FM checkpoint is competitive, so the contrast concerns sustained refinement rather than every checkpoint.

	A100-h	FID_{50k}	FID_{10k}	KID	IS	Prec.	Rec.	Dens.	Cov.	z
Common start	3.3	—	23.9 / 25.1	10.8	7.67	0.68	0.63	0.65	0.80	2.8
FM on fixed completions, best (3 h)	6.3	—	21.2 / 22.2	9.0	7.96	0.70	0.66	0.68	0.86	1.0
FM on fixed completions, 30 h	33.3	30.0	36.0 / 38.4	21.3	6.94	0.73	0.55	0.71	0.77	10.7
FRRF, 30 h	33.4	—	21.6 / 21.7	8.8	7.85	0.69	0.66	0.66	0.81	3.2
FRRF, 36 h	39.3	14.8 / 14.8	21.5 / 21.5	8.7	7.85	0.68	0.66	0.66	0.82	3.8
FRRF, 64-step candidates, 36 h	39.2	15.3	21.9 / 22.2†	9.1	7.82	0.69	0.65	0.67	0.82	-0.5
Anchored, refresh + feedback, 30 h	33.5	15.2 / 14.5	21.6 / 21.0	9.2	7.96	0.69	0.65	0.67	0.82	0.3
Anchored, refresh only, 36 h	39.3	12.3 / 11.9	18.8 / 18.4	7.3	8.40	0.71	0.67	0.74	0.90	0.7
SCSI generator on completions	4.0	18.0	24.7	12.8	7.62	0.68	0.65	0.64	0.78	0.7
SCSI, completion initialization	31.8	13.6	20.5	9.5	7.88	0.70	0.69	0.69	0.83	1.6
SCSI, 20k restoration steps	30.9	13.3	20.1	9.4	7.90	0.70	0.69	0.71	0.84	0.9
SCSI, full schedule (50k steps)	70.8	12.9	19.5	8.7	7.88	0.71	0.69	0.71	0.85	0.7
DiffEM, continued (laps 4–8 of 21)	39.1	13.7	20.6	8.6	8.60	0.68	0.68	0.66	0.87	16.7
DiffEM, 4.8% schedule	9.9	44.1	50.5	35.7	6.49	0.58	0.48	0.41	0.48	15.3

Table 4: **Generation metrics.** FID values are for training seeds 1 / 2 where available (†: seed 2 at 30 hours). Other metrics use seed 1. KID is multiplied by 10^3; IS denotes Inception score, and Prec., Rec., Dens., and Cov. denote precision, recall, density, and coverage ($k = 5$). Anchored models use exponential moving averages of the parameters. The partial DiffEM run precedes the continued run; the continued run fails the observation check on either weights (raw: FID 23.5 / 17.1, dispersion $+5.4\%$, z 22.7).

C.2.3 Ablation

We vary the candidate sampler, residual-training batch size, and movement control one at a time (Table 5). The controls continue ordinary FM on fixed completions, either from the common initialization or without interrupting the initial fit.

	Hours	FID seed 1	FID seed 2	KID	IS	Cov.	Disp. (%)	z
Common start	0	23.9	25.1	10.8	7.67	0.80	-3.4	2.8
FRRF (32-step candidates)	30	21.6	21.7	8.8	7.85	0.81	$+5.2$	3.2
FRRF (32-step candidates)	36	21.5	21.5	8.7	7.85	0.82	$+5.7$	3.8
64-step candidates	30	22.0	22.2	9.3	7.78	0.81	$+0.2$	-0.5
Residual-step batch 4,096	30	22.1	22.5	9.4	7.78	0.81	-0.4	-0.4
Penalized decision, fixed η	20	23.2	24.0	10.5	7.70	0.80	-0.8	-0.1
FM on fixed completions, best	3	21.2	22.2	9.0	7.96	0.86	-2.3	1.0
FM on fixed completions	20	29.6	31.1	15.9	7.37	0.82	-5.2	6.0
FM on fixed completions	30	36.0	38.4	21.3	6.94	0.77	-7.0	10.7

Table 5: **Effect of individual training choices.** Hours count refinement only; FID uses the held-out images and reports both training seeds. KID is multiplied by 10^3; other metrics use seed 1. The fixed-penalty result is reported at 20 hours.

Renewing the population sustains progress. Feedback improves at the recorded checkpoints on both seeds. Fixed-completion FM peaks after three hours; uninterrupted FM shows the same turnover at 32,000 updates. The contrast supports renewing the training population from observation evidence.

Sampling accuracy is part of the update. With 64-step candidates, dispersion stays within 0.4% of the clean reference and observation checks pass at every recorded checkpoint. The 32-step sampler gives slightly lower FID, but its candidates are about 5% less dispersed than 64-step samples of the same field. Correcting this bank leaves final samples 5–6% over-dispersed. The comparison motivates evaluating observation calibration alongside generation quality.

C.2.4 SELECTION AND GENERALIZATION OF CORRECTIONS

The weights select meaningful changes. Across 649 decisions from ten configurations and two seeds, we compare the weighted bank with its uniform counterpart and with 200 random permutations of the weights. Early selected changes reach 4–10 permutation standard deviations, improving calibration and KID (Table 6). With 64-step candidates, the generator's dispersion error moves from -3.5% to 0.0% by ten hours, and its total-variation ratio from 0.92 to 0.99. The dispersion request then vanishes, while a persistent high-frequency request is only partly realized: the generated high-frequency share levels off near 0.6 of the clean value.

Decision (seed 1, rounds 1–8)	Movement	Δdisp.	ΔTV	ΔHF ($\times 10^5$)	ΔKID ($\times 10^3$)
Constrained, 64-step candidates	0.89	$+0.0109$	$+0.0058$	$+3.1$	-1.02
Constrained, batch 4,096	0.87	$+0.0087$	$+0.0052$	$+3.4$	-0.95
Penalized, fixed η, on its bank	0.024	$+0.0069$	$+0.0021$	$+0.05$	-0.67
Penalized, fixed η, fresh candidates	0.026	$+0.0070$	$+0.0022$	-0.09	-0.68

Table 6: **Selected changes on the bank and on fresh candidates.** Weighted-minus-uniform statistics, averaged over rounds 1–8 of training seed 1. Movement is $\|\omega - \mathbf{1}\|_\lambda^2$. Dispersion and total variation (TV) use standardized pixel units; HF is the highest-frequency energy share. Clean references are approximately 0.50, 0.230, and 1.65×10^{-3}, respectively. KID uses held-out images.

Smaller corrections transfer beyond the bank. We solve on three quarters of each bank and evaluate the scoring rule on the withheld quarter (Table 7). Moderate squared movement removes up to 48% of the round-one observation gap; nearly interpolating decisions worsen it. The constrained weights closely follow minimum-norm feature interpolation (correlation 0.99–0.998; effective feature rank 18). The fixed-penalty decisions instead transfer to 8,192 fresh candidates and attain near-zero observation z, with FID 0.8–1.5 higher at 20 hours. Thus movement control trades immediate bank fit for an instruction that generalizes to new samples.

Movement $\|\omega - \mathbf{1}\|_\lambda^2$	0.004	0.018	0.030	0.047	0.15	0.26	0.9 (constrained)
Gap removed, round 1	0.31	0.44	0.47	0.48	0.32	-0.04	-3.4
Gap removed, round 10	—	—	0.44	0.44	—	0.21	-1.8
Gap removed, round 20	—	0.33	0.34	—	—	-0.21	-4.6
Gap removed, round 28	0.21	0.21	0.19	—	-0.57	—	-5.4

Table 7: **Smaller decisions generalize better to new candidates.** Four-fold bank splits at four rounds of training seed 1. Each entry is the fraction of the observation gap removed on the withheld quarter: one is a perfect match, zero makes no improvement, and negative values worsen the gap. Columns give squared relative movement. With 512, 1,024, 2,048, and 3,072 fitting candidates in round one, the best η is 12, 48, 192, and 192, respectively.

Separate selection from realization. A channel-specific control fits 39 pixel moments estimated from kept pixels by inverse-probability weighting: channel means and second moments, neighbor differences, and three DCT-band energies. It uses perceptual MMD as the guard and a fixed movement penalty. All targets agree with clean references within their standard errors, with high-frequency energy the least precisely estimated. The decision requests about 15% more high-band energy and transfers the request to fresh candidates. Eight residual updates instead raise broadband contrast; the high-frequency share rises only from 0.50 to 0.62. FID reaches 26.1 and 27.2 after five hours, versus 23.1 and 24.1 for the perceptual-kernel control, triggering the declared stop. The experiment isolates generator fitting as a further limit on fine-detail recovery, even when the selected statistic is informative.

C.2.5 ANCHORED PATH

One field supports three tasks. The anchored path of Appendix A.5 joins source noise at $t = 0$, the observation at $t = 1/2$, and the image at $t = 1$. The network takes the state and mask as four channels; a gate fixes observed pixels on the second segment. A full trajectory generates, the second segment restores, and (15) gives a one-evaluation posterior-mean estimate. Training puts 45% of times in each segment and 10% at the anchor.

After 16,000 FM updates on completions, a *restoration refresh* transports 4,096 fitting observations to image space and trains for 256 updates on these endpoints with freshly simulated observations. We compare refreshes alone with alternation between a refresh and one feedback round on fresh generated candidates (Algorithm 3). Both use parameter EMA with half-life 1,000 updates; unaveraged checkpoints vary by 5–17 FID.

Algorithm 3 Anchored training: restoration refresh and observation feedback

1: Restore 4,096 fitting observations: integrate v_k from $(Y_\ell, \frac{1}{2})$ to $t = 1$ to obtain $\hat{X}_\ell$.
2: Simulate $Y'_\ell \sim K(\cdot \mid \hat{X}_\ell)$ and take 256 FM updates on the anchored path; update v_k.
3: Apply Algorithm 2 to a fresh bank of 4,096 samples from the updated field.
4: **return** the updated field. The refresh-only control omits step 3.

	5 h	10 h	20 h	30 h	36 h	KID	IS	Cov.	PSNR restore / mean (dB)
Refresh + feedback, seed 1	28.3	26.0	23.4	21.6	—	9.2	7.96	0.82	28.6 / 31.6
Refresh + feedback, seed 2	28.0	25.6	23.1	21.0	—	8.7	7.99	0.84	28.6 / 31.6
Refresh only, seed 1	25.3	22.6	19.6	18.8	18.8	7.3	8.40	0.90	28.6 / 31.4
Refresh only, seed 2	24.9	21.4	19.1	18.6	18.4	7.0	8.34	0.90	28.6 / 31.4

Table 8: **An anchored field supports both generation and conditional readouts.** The time columns give held-out FID; 30- and 36-hour values average two evaluation draws. KID ($\times 10^3$), IS, and coverage are measured at the final reported checkpoint. Restoration and posterior-mean PSNR use 1,024 held-out observations. All results use averaged parameters.

Path choice expands the available readouts. Refresh-only training reaches held-out FID 18.8 and 18.4 while retaining restoration and one-evaluation denoising. Restoration stays at 28.4–28.8 dB throughout refinement; the posterior-mean estimate reaches 31.4–31.6 dB. SCSI's separate restoration model reaches 29.0–29.1 dB on the same observations. Added feedback does not improve FID under the tested schedules, including a lower refresh learning rate and the fixed-penalty decision. This motivates coordinating restoration-driven changes in candidate locations with feedback-driven changes in their frequencies, rather than expecting their gains to add. Refresh-only samples are closer to training images in feature distance, with no duplicates at the evaluation threshold.

C.2.6 Baselines

SCSI. The released implementation (Modi et al., 2026) fits a 64-channel restoration interpolant for 50,000 cosine-scheduled updates (64-step Euler transport, native EMA), then a 96-channel EDM generator (Karras et al., 2022) for 20,001 updates (3.9 A100-hours). Generators fitted from restoration checkpoints at 5,000, 10,000, 20,000, and 35,000 updates are charged separately; these checkpoints precede completion of the cosine schedule. We also test completion-initialized restoration (22,500 updates) and the generator stage alone on completions.

DiffEM. The released code (Hosseintabar et al., 2025) initializes its first conditional fit from completions. The partial run uses four 64-epoch conditional fits and 64 unconditional epochs (4.8% of the published schedule; 9.9 A100-hours). Continuation uses 256 epochs per conditional fit through fit 8 of 21 (28.6% of the native conditional updates), then the native unconditional stage (39.1 A100-hours of training); its reads use parameter EMA, and its Inception score averages the 50,000-sample draw and the second 10,000-sample draw. Validation uses observations. The partial and continued results are reported separately.

Published values use different data protocols and serve as context, not matched rows: SCSI reports FID 6.74 at 50% masking (86 GPU-hours), EM-Posterior 6.76 (512 GPU-hours), DiffEM 10.24 at 75% masking, and Ambient Diffusion 18.85 at 40% (Modi et al., 2026; Rozet et al., 2024; Hosseintabar et al., 2025; Daras et al., 2023).

	A100-h	FID_{50k}	FID_{10k}	KID	IS	Prec.	Rec.	Dens.	Cov.	Disp.	z
SCSI, 5k restoration steps	10.9	—	44.0	31.5	7.34	0.59	0.54	0.46	0.60	+2.3	3.4
SCSI, 10k	17.6	—	21.8	10.1	7.77	0.71	0.68	0.71	0.82	+1.1	1.4
SCSI, 20k	30.9	13.3	20.1	9.4	7.90	0.70	0.69	0.71	0.84	+0.8	0.9
SCSI, 35k	50.9	—	19.4	8.9	7.84	0.72	0.69	0.72	0.85	+1.3	1.0
SCSI, 50k (annealed)	70.8	12.9	19.5	8.7	7.88	0.71	0.69	0.71	0.85	+1.0	0.7
SCSI, completion initialization	31.8	13.6	20.5	9.5	7.88	0.70	0.69	0.69	0.83	+1.5	1.6
SCSI generator on completions	4.0	18.0	24.7	12.8	7.62	0.68	0.65	0.64	0.78	+0.2	0.7
DiffEM, 4.8% of its schedule	9.9	44.1	50.5	35.7	6.49	0.58	0.48	0.41	0.48	+8.3	15.3
DiffEM, conditional fits 4–8 of 21	39.1	13.7	20.6	8.6	8.60	0.68	0.68	0.66	0.87	−7.9	16.7
Completed dataset (not generated)	0.1	—	11.3	2.0	9.96	0.75	0.76	0.82	0.98	−1.1	—

Table 9: **Baseline learning trajectories.** Metrics follow Table 4; Disp. is dispersion error in percent. The final row evaluates the 45,000 completed training images themselves, with observed pixels retained; it is not a generator result. All other rows evaluate independently generated samples.

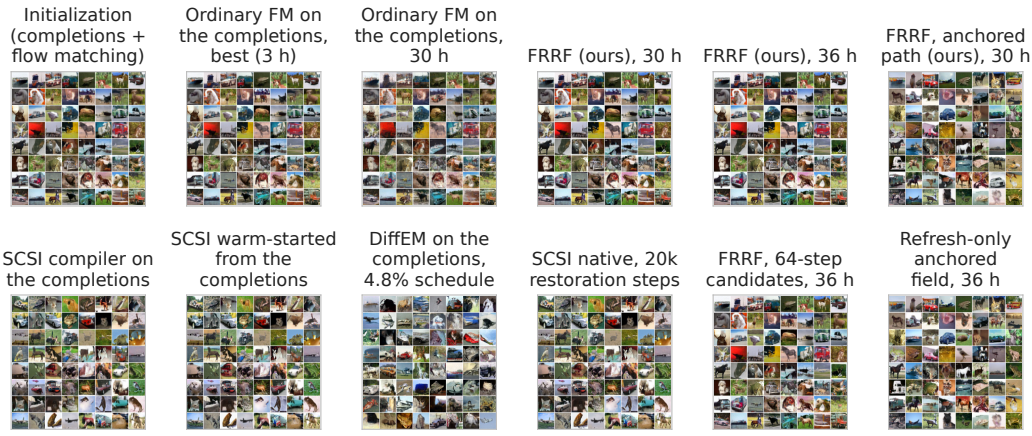

Figure 4: **Unconditional samples from the compared generators.** Each grid uses fixed latent indices. Model labels identify the training procedure and checkpoint; the DiffEM grid is from its partial schedule.

C.3 Astronomical Restoration from the Generative Path

Our primary objective is to learn an unconditional generative model of clean data from corrupted observations. The construction also provides a restoration map without introducing a separate inverse model. In our astronomy specialization, the clean image is at $t = 0$ and the physical observation $A = KX + \sigma E$ is the midpoint state at $t = 1/2$. A measured image Y can therefore initialize the learned path at this midpoint, and integration to the clean endpoint gives

$$\widehat{X}(Y) = F^{\theta}_{0\leftarrow 1/2}(Y). \tag{79}$$

Unlike unconditional sampling, this map produces a reconstruction paired with the particular input observation. The time direction differs from the convention used in the main derivation, but the principle is the same: placing the known corruption at an interior landmark makes part of the generative path an observation-to-clean transport.

We study this secondary capability on the COSMOS 23.5 RealGalaxy data. The evaluation uses a fixed held-out set of 2,000 identity-disjoint 48×48 galaxies, observed through a circular Gaussian point-spread function of width 1.2 pixels with additive Gaussian noise ($\sigma = 0.0496$). We restore each observation with 64 Heun steps; its paired clean image is used only for evaluation. The resulting restorations reduce the log-flux distribution distance from 0.120 for the squared-loss inversion baseline to 0.032, a $3.7\times$ improvement. The examples in Figure 5 demonstrate that the same path learned for unconditional generation can also recover individual astronomical sources.

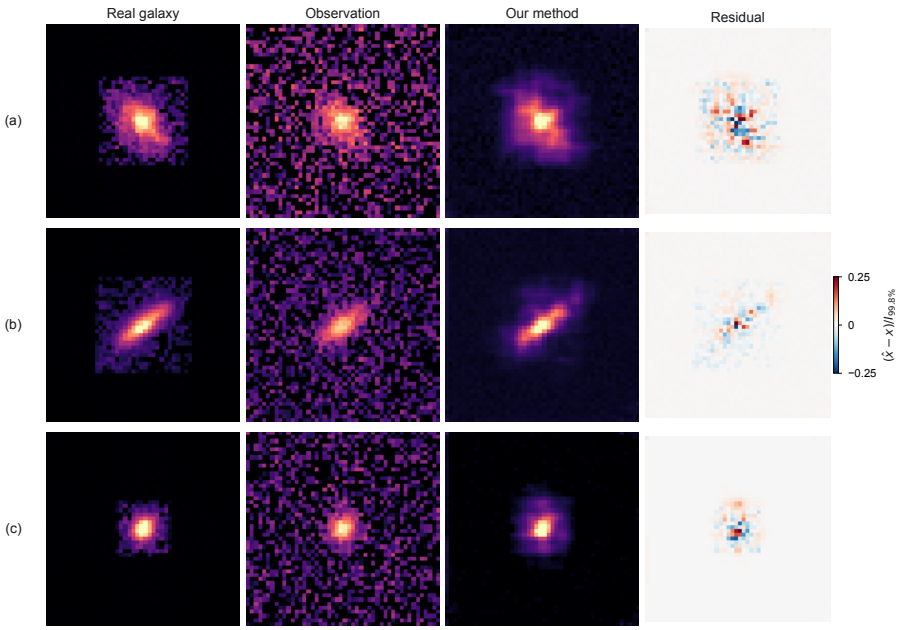

Figure 5: **Restoration of COSMOS galaxy observations.** Columns show the real galaxy used only for evaluation, the blurred and noisy observation, our restoration $F^{\theta}_{0\leftarrow 1/2}(Y)$, and the signed residual. The displayed examples are selected for visualization; the reported log-flux distance is computed on all 2,000 galaxies. Residuals are normalized by the row's 99.8th-percentile clean intensity and shown on the common range $[-0.25, 0.25]$.

D Additional related work

This appendix reviews two lines of work that Section 1.1 summarizes in one sentence: methods that reconstruct or aggregate clean samples through posterior inference, and methods that build training targets from analytical channel structure.

Posterior reconstruction and restoration. DiEM, EM-Diffusion, and DiffEM realize the EM update by reconstructing a clean sample for each observation, through posterior sampling under an unconditional prior or a learned conditional model, and refitting on the reconstructions (Rozet et al., 2024; Bai et al., 2024; Hosseintabar et al., 2025); simulation-based empirical Bayes and posterior averaging aggregate amortized posteriors (Shen et al., 2026; Hoang & Koch, 2026).

Supervision from analytical channel structure. When the corruption has usable analytical form, clean-directed training targets can be built directly: Stein identities for Gaussian and noisy linear measurements (Aali et al., 2023; Kawar et al., 2023), diffusion G-modeling for observation-specific Gaussian covariances (Chen & Cui, 2025), re-corruption and Tweedie consistency (Daras et al., 2023; 2024), and score-distillation, measurement-score, inverse-flow, or Gaussian-bridge constructions (Zhang et al., 2025; Park et al., 2025; Zhang & Zhou, 2025; Urbański et al., 2026). Each identity holds only for its channel family; FRRF instead assumes only a simulatable channel and relies on the forward average $s_\omega(x) = \mathbb{E}_{K(\cdot|x)} f_\omega(Y)$ to transfer an observation-fit score f_ω to the clean candidates that guide generator training.