MARTINGALE RESIDUALS FOR EVALUATING AND TRAINING DIFFUSION MODELS

Anonymous authors
Paper under double-blind review

ABSTRACT

Diffusion models are typically trained with noise or score targets. We introduce a martingale residual (MR) framework for evaluation and learning from sampled reference paths, without requiring these targets. A martingale residual is the observed change along a reference path minus the change predicted by the model. We prove that, under suitable assumptions, MR orthogonality implies that the learned and reference processes have the same path law. We combine weighted residual expectations into an MR score. When the learned process uses the reference process's known covariance, the MR score formed from all coordinate test functions converges to the squared drift error as the time interval shrinks, provided the instruments form a complete orthonormal basis. Finite instruments give a projection approximation. For DDPMs, suitable residual scaling connects this limit to denoising MSE minus a constant independent of the model parameters. We evaluate the MR score as a diagnostic, a training regularizer, and a sole training objective. On CIFAR-10, MR scores track denoising training, with richer instrument sets capturing a larger fraction of the denoising MSE decrease. MR regularization yields modest improvements in sample quality during continuation of pretrained CIFAR-10 and ImageNet-64 models. Using the MR score alone, we learn stochastic transport from a Gaussian mixture to a two-moons distribution in a Schrödinger bridge task.

1 INTRODUCTION

Diffusion models learn to reverse a process that gradually adds noise to data (Ho et al., 2020; Song et al., 2021). This process supplies noise or score targets for supervised training. For image generation, we construct *reference paths* by adding noise to data images and reversing the time order of the resulting trajectories. More generally, reference paths can come from numerical Schrödinger bridge constructions (Shi et al., 2023), stochastic simulations, or observations of particle dynamics (Frishman & Ronceray, 2020), where noise or score targets may not be readily accessible. We therefore ask whether reference paths themselves can provide a signal for both evaluating and training a diffusion model.

We address this question using *martingale residuals* (MRs), which compare observed changes along reference paths with changes predicted by the learned model. We weight these residuals using instruments, functions of time and the state at the start of each interval. When the learned and reference processes have the same drift and covariance, the weighted residuals have zero expectation. We combine these expectations into a scalar *MR score* for evaluation and training. Figure 1 summarizes the framework.

Theorem 3.1 shows that, under suitable assumptions, MR orthogonality implies that the learned and reference processes have the same path law. Proposition 3.2 shows that coordinate test functions suffice when the learned process uses the reference process's known covariance. For coordinate test functions and a complete orthonormal instrument basis, the MR score converges to the squared drift error as the time interval shrinks. For DDPMs, suitable residual scaling connects this limit to denoising MSE minus a constant independent of the model parameters.

We estimate the MR score from discretely observed reference paths using fixed or learned instruments. We evaluate the MR score as a diagnostic, a training regularizer, and a sole training objective. As a diagnostic on CIFAR-10, the MR score captures an estimated 89.7% of the denoising MSE decrease

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

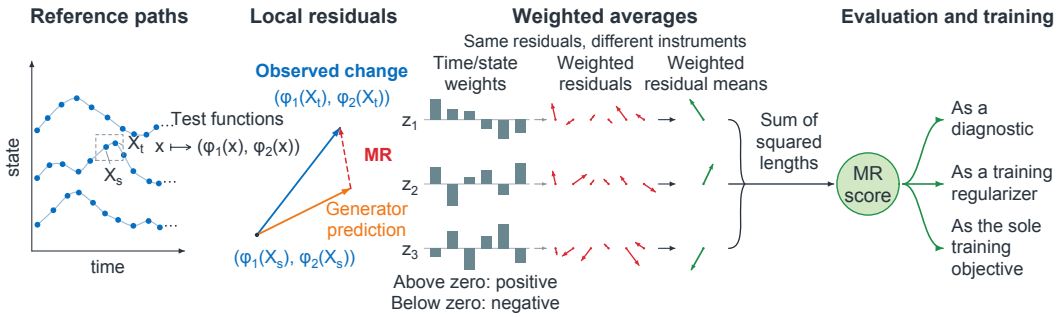

Figure 1: MR score construction from reference paths. Residuals compare observed changes in test functions with model predictions. Their weighted expectations are combined into an MR score for evaluation and training.

between 2k and 40k updates using a rich instrument set. As a regularizer, MR yields modest sample quality improvements during continuation of pretrained CIFAR-10 and ImageNet-64 models. Using MR alone, we improve the pretrained ImageNet-64 model to a FID50k of 19.88. We also learn a Schrödinger bridge from scratch using MR alone, without analytic drift, noise, or score targets. The approach applies beyond constructions based on Brownian bridges.

Related work. Denoising score matching, Flow Matching, and Generator Matching derive training objectives from known conditional constructions (Vincent, 2011; Lipman et al., 2023; Holderrieth et al., 2025). Schrödinger bridge matching learns transport between prescribed marginals (Shi et al., 2023). Stochastic force inference estimates force and diffusion fields from observed trajectories (Frishman & Ronceray, 2020). Diffusion calibration corrects the mean of pretrained score functions (Pang et al., 2023), while martingale guidance enforces constraints through drift corrections (Guo et al., 2026). MR builds on martingale estimating functions, conditional moments, and optimal instruments (Bibby & Sørensen, 1995; Hansen, 1982; Chamberlain, 1987; Newey & Powell, 2003). Learned instruments connect to variational moment estimation (Bennett et al., 2019; Dikkala et al., 2020), and sample splitting separates fitting from evaluation (Chernozhukov et al., 2018).

2 Preliminaries

Diffusion models. A *diffusion process* is an $\mathbb{R}^d$-valued stochastic process $X = (X_t)_{0 \le t \le T}$ satisfying

$$\mathrm{d}X_t = f(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t, \qquad a(t,x) := \sigma(t,x)\sigma(t,x)^\top, \tag{1}$$

where W is a standard Brownian motion. Let p_t denote the marginal density of X_t.

In *diffusion models*, the diffusion process is used as the forward noising process: $X_0 \sim p_{\mathrm{data}}$, and the coefficients f and σ are chosen so that the terminal marginal p_T is close to a tractable prior π, typically Gaussian. The time reversal of equation 1 is itself a diffusion process,

$$\mathrm{d}X_t = b(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}\overline{W}_t, \tag{2}$$

where $\overline{W}$ is a standard Brownian motion in reverse time and the reverse drift is

$$b(t,x) := f(t,x) - (\nabla_x \cdot a)(t,x) - a(t,x)\nabla_x \log p_t(x).$$

The unknown distribution-dependent term is the *marginal score* $\nabla_x \log p_t$ (Song et al., 2021). Replacing it with a learned score model s^θ gives b^θ. To generate data, one samples $X_T \sim \pi$ and simulates the learned reverse-time SDE from T to 0.

The score model s^θ is typically learned by *denoising score matching*, whose population minimizer is the marginal score (Vincent, 2011). For a VP diffusion, the forward diffusion covariance is $a(t,x) = \beta(t)I$ with a known schedule $\beta(t) > 0$, and the perturbation distribution is $X_t \mid X_0 \sim \mathcal{N}(\alpha_t X_0, \sigma_t^2 I)$. The parameterization $s^\theta = -\epsilon_\theta/\sigma_t$ gives the usual denoising loss, the MSE between predicted and sampled noise (Ho et al., 2020). The preceding training objectives operate pointwise

through the known noising mechanism: each corrupted state is paired with either the conditional score or the noise that produced it. We instead study an information regime in which these targets are unavailable. Our path-only analysis evaluates b^θ directly along reference trajectories following the reverse-time process in equation 2.

Martingales. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t\le T}, Q)$ be a filtered probability space. An adapted integrable process $M = (M_t)_{0\le t\le T}$ is a *Q-martingale* if $\mathbb{E}_Q[M_t \mid \mathcal{F}_s] = M_s$ for every $0 \le s < t \le T$. The *martingale orthogonality characterization* states that M is a Q-martingale if and only if $\mathbb{E}_Q[Z_s(M_t - M_s)] = 0$ for every $0 \le s < t \le T$ and every bounded $\mathcal{F}_s$-measurable random variable Z_s.

3 METHOD

We develop a path-only objective that evaluates the learned diffusion process against a fixed reference process using only sampled trajectories and model evaluations, without noise or score targets. Section 3.1 introduces martingale residuals and gives sufficient conditions for equality of the model and reference path laws. Section 3.2 explains what the resulting objective measures under known covariance. Section 3.3 expresses the objective for general instruments and gives an algorithm for estimating an MR score from sampled paths. Section 3.4 describes how to estimate the objective's gradient and update the model. Proofs and computational details are in Appendices A–D.

3.1 MARTINGALE RESIDUALS AND PATH LAW IDENTIFICATION

We first allow the learned and reference processes to have different drifts and covariances. From Section 3.2 onward, we assume that the reference covariance is known and use the same covariance for the learned process.

Martingale residuals and the martingale problem. We reparameterize equation 2 so that t increases in the direction of generation, from the prior toward the data distribution, while retaining the same SDE form and reusing b and a to denote the resulting drift and covariance. Let $Q \in \mathcal{P}(C([0,T], \mathbb{R}^d))$ denote the induced path law and $\mathcal{F}_t := \sigma(X_r : 0 \le r \le t)$ the natural filtration. A trajectory drawn from Q is called a *reference path*. The corresponding *infinitesimal generator* is

$$(\mathcal{L}_t\phi)(t,x) := b(t,x)^\top \nabla_x \phi(t,x) + \frac{1}{2}\operatorname{tr}\big(a(t,x)\nabla_x^2\phi(t,x)\big).$$

For an admissible test function ϕ and $0 \le s \le t \le T$, define the *martingale residual* (MR) over $[s,t]$ by

$$M_{s,t}^\phi := \phi(t, X_t) - \phi(s, X_s) - \int_s^t (\partial_r + \mathcal{L}_r)\phi(r, X_r)\, \mathrm{d}r.$$

Given a generator $\mathcal{L}$ and an initial distribution μ_0, the *martingale problem* $\mathrm{MP}(\mathcal{L}, \mu_0)$ asks for a path law with initial distribution μ_0 under which $(M_{0,t}^\phi)_{0\le t\le T}$ is a martingale for every admissible test function ϕ. The problem is *well posed* if such a law exists and is unique. By Itô's formula, the reference law Q solves $\mathrm{MP}(\mathcal{L}, \mu_0)$ with $\mu_0 = Q_0$.

The learned model specifies the drift b^θ and covariance a^θ and induces the path law P_θ. Its generator $\mathcal{L}_t^\theta$ and residual $M_{s,t}^{\phi,\theta}$ are defined analogously. Comparing $M_{s,t}^{\phi,\theta}$ with $M_{s,t}^\phi$ gives

$$M_{s,t}^{\phi,\theta} = M_{s,t}^\phi + \int_s^t (\mathcal{L}_r - \mathcal{L}_r^\theta)\phi(r, X_r)\, \mathrm{d}r. \tag{3}$$

The first term is a Q-martingale increment, while the second accumulates the generator discrepancy detected by ϕ over the interval. Sufficient admissibility conditions are given in Appendix A.1.

Path law identification. Our goal is to learn a diffusion whose path law matches that of the reference process, $P_\theta = Q$. In particular, the terminal law of the learned diffusion equals Q_T, the target data distribution.

For $0 \le s < t \le T$, we use $Z_s = z(X_s)$, with $z\colon \mathbb{R}^d \to \mathbb{R}$ bounded and measurable, as an *instrument*, a weight determined by the state at time s. If the learned residuals are increments of a Q-martingale, then $\mathbb{E}_Q[Z_s M_{s,t}^{\phi,\theta}] = 0$ by martingale orthogonality. Requiring this orthogonality for every admissible ϕ gives the following identification result.

Theorem 3.1. *Suppose Q and P_θ have the same initial distribution μ_0. If*

$$\mathbb{E}_Q\left[z(X_s)M_{s,t}^{\phi,\theta}\right] = 0 \tag{4}$$

for all admissible test functions ϕ, bounded measurable functions $z\colon \mathbb{R}^d \to \mathbb{R}$, and $0 \le s < t \le T$, then Q solves $\mathrm{MP}(\mathcal{L}^\theta, \mu_0)$. If this martingale problem is well posed, then $Q = P_\theta$.

Appendices A.2 and A.3 state the technical assumptions and sufficient conditions for well-posedness.

A nonzero value of $\mathbb{E}_Q[Z_s M_{s,t}^{\phi,\theta}]$ indicates a departure from the martingale property. Since $\mathbb{E}_Q[M_{s,t}^{\phi} \mid \mathcal{F}_s] = 0$, equation 3 gives

$$\mathbb{E}_Q\left[Z_s M_{s,t}^{\phi,\theta}\right] = \mathbb{E}_Q\left[Z_s \int_s^t (\mathcal{L}_r - \mathcal{L}_r^\theta)\phi(r, X_r)\,\mathrm{d}r\right].$$

The expectation on the left can be estimated from sampled reference paths and evaluations of $\mathcal{L}^\theta$, without knowing $\mathcal{L}$. Expanding the generator difference gives

$$(\mathcal{L}_r - \mathcal{L}_r^\theta)\phi(r,x) = (b - b^\theta)(r,x)^\top \nabla_x \phi(r,x) + \frac{1}{2}\,\mathrm{tr}\big((a - a^\theta)(r,x)\nabla_x^2\phi(r,x)\big).$$

$\nabla_x\phi$ and $\nabla_x^2\phi$ determine which combinations of drift and covariance errors are tested, respectively. The instrument $Z_s = z(X_s)$ weights these errors using the state at time s. When $a = a^\theta$, only the drift term remains. Under this condition, we next show that coordinate functions suffice as test functions.

3.2 Coordinate test functions and the MR score

As in the diffusion models introduced in Section 2, we assume that the covariance of the reference process is known and set $a^\theta(t,x) = a(t,x)$ in the learned model. We use this setting for the methods and experiments that follow.

Sufficiency of coordinate test functions. The following proposition shows that the test functions in Theorem 3.1 can be restricted to the d *coordinate functions* $x \mapsto x_i$, $i = 1, \dots, d$.

Proposition 3.2. *Suppose Q and P_θ have the same initial distribution μ_0 and $a^\theta = a$. If*

$$\mathbb{E}_Q\left[z(X_s)\left(X_t^i - X_s^i - \int_s^t b_i^\theta(r, X_r)\,\mathrm{d}r\right)\right] = 0, \tag{5}$$

for every $i = 1, \dots, d$, $0 \le s < t \le T$, and every bounded measurable function $z\colon \mathbb{R}^d \to \mathbb{R}$, then

$$b^\theta(t, X_t) = b(t, X_t) \qquad \mathrm{d}t \otimes Q\text{-almost everywhere}.$$

If the martingale problem for $(\mathcal{L}^\theta, \mu_0)$ is well posed, then $Q = P_\theta$.

Under the known covariance assumption,

$$(\mathcal{L}_t - \mathcal{L}_t^\theta)\phi(t,x) = \sum_{i=1}^d \partial_i\phi(t,x)\big(b_i(t,x) - b_i^\theta(t,x)\big),$$

so the coordinate discrepancies determine the generator discrepancy. Technical conditions are given in Appendix B.

MR score. We combine the expectations in equation 5 by summing their squared norms. We first consider a complete orthonormal instrument basis and then examine the effect of retaining only finitely many basis functions.

Applying the learned MR formula to the coordinate test functions $x \mapsto x_i$, $i = 1, \dots, d$, and dividing by the interval length h, we define the normalized MR increment

$$r_{t,h}^\theta := \frac{1}{h}\left[X_{t+h} - X_t - \int_t^{t+h} b^\theta(s, X_s)\,\mathrm{d}s\right] \in \mathbb{R}^d,$$

where $h > 0$ and $0 \le t < t + h \le T$.

We sample $t \sim \nu$ independently of $X \sim Q$, and let μ denote the distribution of (t, X_t). Expectations $\mathbb{E}_Q$ below average over both samples. For scalar instruments $z_\ell \in L^2(\mu)$, the following theorem relates the sum of squared norms of $\mathbb{E}_Q[z_\ell(t, X_t) r_{t,h}^\theta] \in \mathbb{R}^d$ to the drift error $b - b^\theta$.

Theorem 3.3. *Suppose $\{z_\ell\}_{\ell\geq 1}$ is a complete orthonormal basis of $L^2(\mu)$. Under the regularity conditions stated in Appendix B.3,*

$$\sum_{\ell=1}^{\infty} \left\|\mathbb{E}_Q[z_\ell(t,X_t)r^\theta_{t,h}]\right\|_2^2 = \left\|\mathbb{E}_Q[r^\theta_{t,h} \mid t, X_t]\right\|^2_{L^2(\mu)}. \tag{6}$$

As $h \downarrow 0$, both sides converge to $\left\|b - b^\theta\right\|^2_{L^2(\mu)}$.

If the drift error $b - b^\theta$ vanishes along the reference paths over the full time horizon, Proposition 3.2 implies $Q = P_\theta$.

For DDPM noise prediction, rescaling the MR increments gives

$$\lim_{h\downarrow 0} \sum_{\ell=1}^{\infty} \left\|\mathbb{E}_Q\left[z_\ell(t,X_t)\frac{\sigma_t}{\beta(t)}r^\theta_{t,h}\right]\right\|_2^2 = \left\|\frac{\sigma_t}{\beta(t)}(b-b^\theta)\right\|^2_{L^2(\mu)} = \mathcal{L}_{\mathrm{DDPM}}(\theta) - C_{\mathrm{irr}}. \tag{7}$$

Here $\mathcal{L}_{\mathrm{DDPM}}$ is the denoising MSE and C_{irr} is independent of θ.

For our population MR objective, we retain the first J terms of the unscaled sum in equation 6. For training, we minimize an empirical approximation of this objective over θ. Write $\mathcal{S} = \mathrm{span}\{z_1, \ldots, z_J\}$ and let $P_\mathcal{S}$ denote the orthogonal $L^2(\mu)$ projection onto this span, applied to each coordinate. By conditional expectation,

$$\begin{aligned}\sum_{\ell=1}^{J} \left\|\mathbb{E}_Q[z_\ell(t,X_t)r^\theta_{t,h}]\right\|_2^2 &= \sum_{\ell=1}^{J} \left\|\mathbb{E}_Q\left[z_\ell(t,X_t)\mathbb{E}_Q[r^\theta_{t,h} \mid t, X_t]\right]\right\|_2^2 \\ &= \left\|P_\mathcal{S}\mathbb{E}_Q[r^\theta_{t,h} \mid t, X_t]\right\|^2_{L^2(\mu)}.\end{aligned} \tag{8}$$

For fixed h, completeness gives

$$\sum_{\ell=J+1}^{\infty} \left\|\mathbb{E}_Q[z_\ell(t,X_t)r^\theta_{t,h}]\right\|_2^2 = \left\|(I-P_\mathcal{S})\mathbb{E}_Q[r^\theta_{t,h} \mid t, X_t]\right\|^2_{L^2(\mu)} \longrightarrow 0, \qquad J \to \infty.$$

We define the *MR score* as the finite sum on the left of equation 8, with coordinate test functions replaced by any finite collection of test functions. Figure 1 illustrates this construction.

3.3 Computing the MR score

For convenience, we retain the notation of Section 3.2 and present the computation of the MR score for coordinate test functions. The same construction extends directly to other finite collections of test functions. Write $Z = (z_1(t,X_t), \ldots, z_J(t,X_t))^\top$. We collect the weighted expectations in equation 8 into the matrix $C_{\theta,h} := \mathbb{E}_Q[Z(r^\theta_{t,h})^\top] \in \mathbb{R}^{J\times d}$, whose ℓth row is

$$(C_{\theta,h})_{\ell,:} = \mathbb{E}_Q[z_\ell(t,X_t)r^\theta_{t,h}]^\top, \qquad \ell = 1, \ldots, J.$$

Hence,

$$\sum_{\ell=1}^{J} \left\|\mathbb{E}_Q[z_\ell(t,X_t)r^\theta_{t,h}]\right\|_2^2 = \mathrm{tr}(C^\top_{\theta,h}C_{\theta,h}).$$

In practice, the instruments are not guaranteed to be orthonormal in $L^2(\mu)$. To express the projection objective without assuming orthonormal instruments, we use the *Gram matrix* $G := \mathbb{E}_Q[ZZ^\top]$. For $\mathcal{S} = \mathrm{span}\{z_1, \ldots, z_J\}$, the MR score can be written as

$$\mathrm{tr}(C^\top_{\theta,h}G^\dagger C_{\theta,h}) = \left\|P_\mathcal{S}\mathbb{E}_Q[r^\theta_{t,h} \mid t, X_t]\right\|^2_{L^2(\mu)}, \tag{9}$$

where $G^\dagger$ is the pseudoinverse of G. This objective depends only on the instrument span. For orthonormal instruments, $G = I$, recovering the sum in equation 8.

To estimate the MR score in equation 9, we estimate G and $C_{\theta,h}$ from sampled paths. The drift integrals in the MR increments cannot be evaluated directly from grid observations, so we use left-endpoint approximations. Algorithm 1 summarizes the computation. Its $\widetilde{r}_k^{\theta,(n)}$ approximates $r^{\theta}_{t_k,\Delta t_k}$ on path $X^{(n)}$. One drift evaluation gives all d coordinate residuals. We analyze the discretization error in Appendix C.

Algorithm 1 MR score estimation with coordinate test functions

Input Independent reference paths $X^{(n)} \sim Q$, $n = 1, \dots, N$, drift b^θ, instruments $z_1, \dots, z_J$, grid $0 = t_0 < \cdots < t_K = T$, time weights $\omega_k \geq 0$ with $\sum_{k=0}^{K-1} \omega_k = 1$, regularization parameter $\lambda > 0$.

$\Delta t_k \leftarrow t_{k+1} - t_k$

Split the reference paths into disjoint sets $\mathcal{D}_G, \mathcal{D}_{C,1}, \mathcal{D}_{C,2}$. Keep all observations from each path in the same set.

for $n = 1, \dots, N$, $k = 0, \dots, K-1$ **do**

  $Z_k^{(n)} \leftarrow \big(z_\ell(t_k, X_{t_k}^{(n)})\big)_{\ell=1}^{J}$

  if $n \in \mathcal{D}_{C,1} \cup \mathcal{D}_{C,2}$ **then**

    $\widetilde{r}_k^{\theta,(n)} \leftarrow \frac{X_{t_{k+1}}^{(n)} - X_{t_k}^{(n)}}{\Delta t_k} - b^\theta(t_k, X_{t_k}^{(n)})$

  end if

end for

$\widehat{G} \leftarrow \frac{1}{|\mathcal{D}_G|} \sum_{n \in \mathcal{D}_G} \sum_{k=0}^{K-1} \omega_k Z_k^{(n)} (Z_k^{(n)})^\top$

for $u = 1, 2$ **do**

  $\widehat{C}_{\theta,h}^{(u)} \leftarrow \frac{1}{|\mathcal{D}_{C,u}|} \sum_{n \in \mathcal{D}_{C,u}} \sum_{k=0}^{K-1} \omega_k Z_k^{(n)} (\widetilde{r}_k^{\theta,(n)})^\top$

end for

return $\operatorname{tr}\Big((\widehat{C}_{\theta,h}^{(1)})^\top (\widehat{G} + \lambda I)^{-1} \widehat{C}_{\theta,h}^{(2)}\Big)$

3.4 Training with the MR score

We construct two training losses from the MR score, one using fixed instruments and the other using learned instruments.

Fixed instruments. We choose the instrument functions $z_1, \dots, z_J$ in advance and keep them fixed during training. To reduce backpropagation cost, we differentiate the estimated MR score in Algorithm 1 only through $\widehat{C}_{\theta,h}^{(1)}$, giving

$$\mathcal{L}_{\mathrm{MR}}(\theta) = \operatorname{tr}\Big((\widehat{C}_{\theta,h}^{(1)})^\top \operatorname{stopgrad}\Big[(\widehat{G} + \lambda I)^{-1} \widehat{C}_{\theta,h}^{(2)}\Big]\Big). \tag{10}$$

We call the paths in $\mathcal{D}_{C,1}$ *update paths* and those in $\mathcal{D}_{C,2}$ *support paths*.

Learned instruments. With fixed instruments, the MR score in equation 9 can be written as

$$\operatorname{tr}(C_{\theta,h}^\top G^\dagger C_{\theta,h}) = \mathbb{E}_Q\big[g(t, X_t)^\top r_{t,h}^\theta\big], \qquad g(t,x) = (G^\dagger C_{\theta,h})^\top Z(t,x).$$

Rather than restricting g to linear combinations of fixed instruments, we learn it with an *auxiliary network* $g_\psi(t,x) \in \mathbb{R}^d$, whose components serve as learned instruments. We use a variational formulation of the MR score. For fixed θ, maximizing the following objective over ψ is equivalent to fitting g_ψ to the numerical residuals $\widetilde{r}^\theta$ from Algorithm 1 by least squares.

$$\min_\theta \max_\psi \; \mathbb{E}_Q\Big[2 g_\psi(t, X_t)^\top \widetilde{r}^\theta - \|g_\psi(t, X_t)\|_2^2\Big]. \tag{11}$$

We alternate between maximizing over ψ and minimizing over θ, as detailed in Appendix D.

4 Experiments

We evaluate MR as a diagnostic, a training regularizer, and the sole training objective. Sole-MR experiments cover both training from scratch and continuation of a pretrained model. Our image experiments use 32×32 RGB images from CIFAR-10 (Krizhevsky, 2009) and 64×64 RGB images from ImageNet (Deng et al., 2009), without class conditioning. All experiments use known covariance and separate sets of reference paths for fitting and evaluation. Protocols and additional results are in Appendix E.

4.1 MR SCORE AS A DIAGNOSTIC

We test whether MR scores capture improvements from denoising training. Using OpenAI's `improved-diffusion` implementation (Nichol & Dhariwal, 2021), we train an unconditional CIFAR-10 model from scratch for 40,000 updates with denoising MSE. MR scores are used only for evaluation. Following Proposition 3.2, we use all coordinate test functions. We choose instruments of the form $q_m(t)c_j(x)$, where q_m is a normalized Legendre time polynomial of degree m and $c_j(x)$ is a discrete cosine transform (DCT) coefficient, linear in the reference state x. Each image has $3 \times 32 \times 32 = 3{,}072$ DCT coefficients. We vary the maximum time degree from zero to three and retain increasing numbers of DCT coefficients from low to high spatial frequency. Using the residual scaling in equation 7, we compute MR scores on the same scale as denoising MSE.

Figure 2A shows that MR captures a larger fraction of the denoising MSE decrease as the DCT coefficient set expands, consistent with the finite projection in equation 8. The largest fraction is 89.7%, using all DCT coefficients and time polynomials through degree three. With the same time polynomials, Figure 2B shows MR scores decreasing alongside denoising MSE. Experimental details and additional results are provided in Appendix E.1.

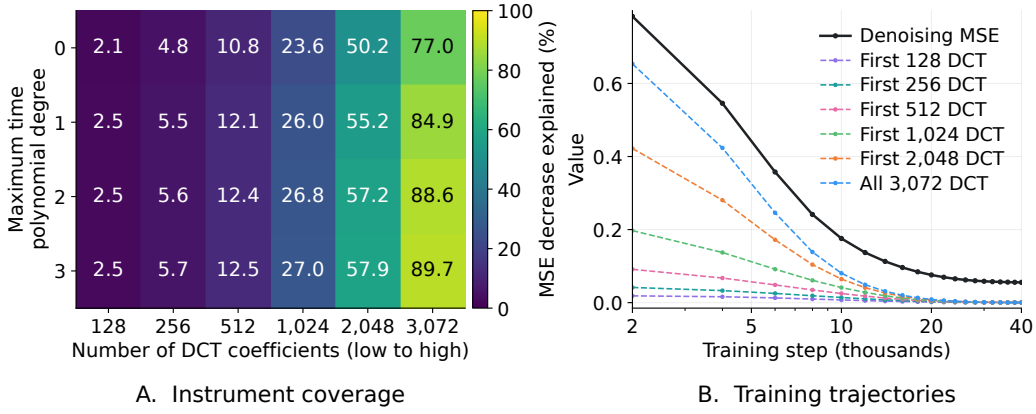

Figure 2: A. Denoising MSE decrease captured by the MR score (%, 2k–40k updates). Rows include time polynomials up to the indicated degree. B. MR scores and denoising MSE during training, using the degree-three configurations in A.

4.2 MR SCORE AS A TRAINING REGULARIZER

We test whether training with $\mathcal{L}_{\text{denoise}} + \gamma\mathcal{L}_{\text{MR}}$ improves sample quality during continuation of pretrained unconditional models. We use an official CIFAR-10 model with EDM preconditioning (Karras et al., 2022) and an OpenAI ImageNet-64 model (Nichol & Dhariwal, 2021). CIFAR-10 and ImageNet-64 are continued for 5,000 and 500 updates, respectively, from the same model weights as their denoising-only baselines.

The Compact configuration uses DCT coefficients to define the test functions $\phi_{j,1}(x) = c_j(x)$ and $\phi_{j,2}(x) = \frac{1}{2}c_j(x)^2$, $j = 1, \ldots, 8$. These 16 tests are paired with instruments that depend only on time. Coordinate-DCT8 and Coordinate-DCT32 instead retain all coordinate test functions and use products of normalized Legendre time polynomials with the first eight or 32 DCT coefficients as instruments, following Section 4.1. The weight γ is calibrated separately for each configuration using gradients at the pretrained model.

We evaluate sample quality using Fréchet Inception Distance (FID) on 50,000 generated images (Heusel et al., 2017). We report means over three paired continuation runs, with ΔFID and MR ratios relative to denoising alone. MR ratios use each configuration's own test functions and instruments. Table 1 shows that Compact improves FID on both datasets. Coordinate-DCT8 performs best on CIFAR-10 but worsens FID on ImageNet-64, and Coordinate-DCT32 worsens FID on both. The benefit therefore depends on the test functions and instruments, and adding DCT coefficients does not guarantee better FID. All three configurations reduce their targeted MR scores on both datasets, but

FID does not always improve, while denoising MSE changes little. Protocols and additional results are in Appendices E.2 and E.3.

Table 1: FID50k and MR score ratios after continued training.

Method	FID50k ↓	ΔFID ↓	MR ratio ↓
CIFAR-10			
Original	1.9916	–	–
Denoising	2.0052	–	1.0000
Compact	1.9818	−0.0234	0.7826
Coordinate-DCT8	**1.9490**	**−0.0562**	0.9771
Coordinate-DCT32	2.0518	+0.0466	0.9853
ImageNet-64			
Original	20.8087	–	–
Denoising	20.7466	–	1.0000
Compact	**20.5793**	**−0.1673**	0.9785
Coordinate-DCT8	20.8716	+0.1250	0.9989
Coordinate-DCT32	20.9606	+0.2140	0.9988

Table 2: ImageNet-64 continuation with MR as the sole training loss.

Method	FID50k ↓	ΔFID ↓	MR ratio ↓
Original	20.8087	–	–
Denoising	20.7466	–	1.0000
Compact MR only (500 updates)	20.2715	−0.4751	0.9560
Compact MR only (2,000 updates)	**19.8759**	**−0.8707**	0.8564

4.3 MR SCORE AS THE SOLE TRAINING OBJECTIVE

ImageNet-64 continuation. Building on Section 4.2, we remove the denoising objective and continue training the pretrained ImageNet-64 model using Compact MR alone. After 500 updates, MR-only continuation achieves a mean FID50k of 20.2715, compared with 20.5793 for Compact MR regularization in Table 1. A separate experiment with 2,000 MR-only updates reaches 19.8759, as shown in Table 2. These results demonstrate that MR can improve the pretrained model without retaining the denoising objective during continuation. Experimental and aggregation details are provided in Appendix E.4.

Table 3: Schrödinger bridge training with fixed and learned instruments.

Instruments	Instrument count	Integrated SW2 ↓	Terminal SW2 ↓	Histogram JS ↓	Separation failure ↓	Thickness W1 ↓
DCT	12	.08171	.13474	.30902	.74573	.18412
Hermite	180	.03710	.04837	.16249	.30548	.03546
RFF	2,068	.03604	.04268	.14910	.24353	.01834
RBF	2,068	.03597	.04267	.13888	.17123	.00501
Learned	N/A	.03495	.03959	.13860	.13962	.00497

Schrödinger bridge learning. We next use the MR score as the sole training loss to learn a Schrödinger bridge from scratch using reference paths, without analytic drift, noise, or score targets. This approach also applies beyond constructions based on Brownian bridges, where such targets can be difficult to obtain (Shi et al., 2023). To simplify reference path sampling, we use Brownian bridges to construct paths from a mixture of eight Gaussians to a thick two-moons distribution. Training does not use the available conditional drift targets. We use coordinate test functions and train from scratch with the fixed and learned instrument losses in Section 3.4. We construct fixed instruments as products of normalized Legendre time polynomials and state functions, which are DCT coefficients, Hermite polynomials, random Fourier features (RFF), and Gaussian radial basis functions (RBF).

Table 3 shows slightly better overall results with learned instruments. Figure 3A shows how instrument choice affects the generated terminal distribution. Figure 3B shows how the model trained with learned instruments transports samples from eight Gaussians to two moons. Details are in Appendix E.5, and additional VP generation experiments with MR alone are in Appendix E.6.

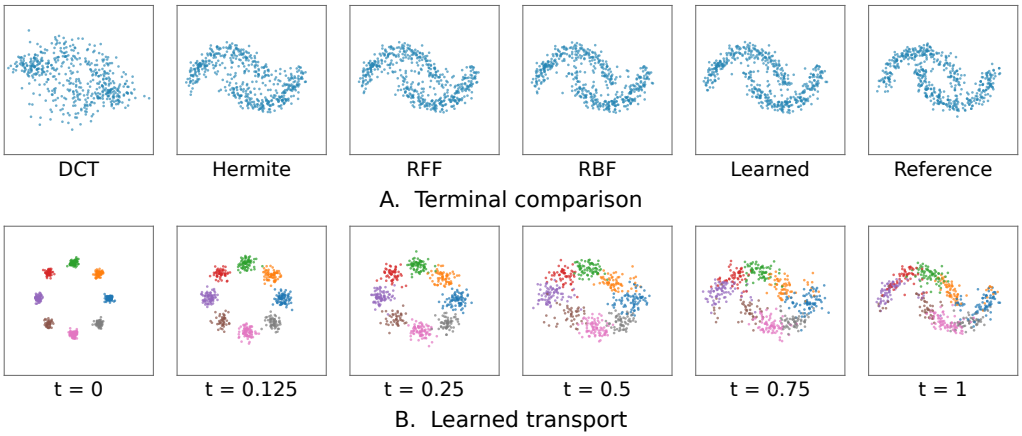

Figure 3: Schrödinger bridge learning with MR score alone. A. Terminal samples from the five models in Table 3 and the reference target. B. Six snapshots of the Learned model in A, colored by initial source component.

5 Discussion and Conclusion

We presented a martingale residual framework for evaluating and training diffusion models from reference paths. Under the stated assumptions, MR orthogonality implies that the learned and reference processes have the same path law. Coordinate test functions suffice when the covariance is known and shared. For VP diffusions, residual scaling connects the complete instrument limit over shrinking time intervals to excess denoising risk. On CIFAR-10, MR scores capture an estimated 89.7% of the denoising MSE decrease from 2k to 40k updates. Compact MR regularization improves FID on both CIFAR-10 and ImageNet-64. Used alone for continuation, MR further improves the pretrained ImageNet-64 model, reaching a mean FID50k of 19.88. We also learn a Schrödinger bridge from scratch using reference paths, without analytic drift, noise, or score targets.

Limitations. These population results do not guarantee stable or efficient optimization with finite samples. Our exploratory CIFAR-10 experiments from random initialization encountered instability or a gap between MR reduction and denoising accuracy, depending on how the residual weighting coefficients in equation 10 were estimated and reused (Appendix E.7). Coordinate test functions suffice for identification with sufficiently rich instruments, but sampling noise, time discretization, instrument truncation, and optimization affect their practical effectiveness. These limitations motivate further investigation of test function design.

The instrument sets considered here may leave substantial components of the conditional residual mean unmeasured, particularly in high dimensions. Increasing their number alone does not ensure better generation, and reducing a finite MR score need not improve sample quality.

Improving the reliability and computational efficiency of MR training remains an open problem. Future work should investigate test function design, adaptive instrument selection, and lower variance estimation, together with broader evaluation across models, datasets, and training settings. These studies may help develop more reliable and scalable MR training methods for a wider range of generative modeling tasks.

AI use statement

In this work, we used generative AI tools for refining and checking mathematical proofs. Additionally, we used generative AI tools for literature searches and manuscript editing. We have reviewed all AI-assisted work. The authors developed the theoretical claims and designed the experiments, checked the mathematical arguments and references, and carefully revised the manuscript. The authors accept full responsibility for the final manuscript, including its text, claims, results, and supporting artifacts.

Reproducibility statement

Appendices A–B provide technical conditions and proofs for the population results. Appendices C–D give discretization and training details. Appendix E documents the experimental protocols and additional results.

References

Andrew Bennett, Nathan Kallus, and Tobias Schnabel. Deep generalized method of moments for instrumental variable analysis. In *Advances in Neural Information Processing Systems*, volume 32, pp. 3564–3574, 2019.

Bo Martin Bibby and Michael Sørensen. Martingale estimation functions for discretely observed diffusion processes. *Bernoulli*, 1(1/2):17–39, 1995.

Gary Chamberlain. Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics*, 34(3):305–334, 1987.

Victor Chernozhukov, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1):C1–C68, 2018.

Jia Deng, Wei Dong, Richard Socher, Li Jia Li, Kai Li, and Li Fei-Fei. ImageNet: A large-scale hierarchical image database. In *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009.

Nishanth Dikkala, Greg Lewis, Lester Mackey, and Vasilis Syrgkanis. Minimax estimation of conditional moment models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 12248–12262, 2020.

Stewart N. Ethier and Thomas G. Kurtz. *Markov Processes: Characterization and Convergence*. Wiley, 1986.

Anna Frishman and Pierre Ronceray. Learning force fields from stochastic trajectories. *Physical Review X*, 10(2):021009, 2020.

Zhengyi Guo, Wenpin Tang, and Renyuan Xu. Conditional diffusion guidance under hard constraint: A stochastic analysis approach. *arXiv preprint arXiv:2602.05533*, 2026.

Lars Peter Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50(4):1029–1054, 1982.

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*, volume 30, 2017.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 33, pp. 6840–6851, 2020.

Peter Holderrieth, Marton Havasi, Jason Yim, Neta Shaul, Itai Gat, Tommi Jaakkola, Brian Karrer, Ricky T. Q. Chen, and Yaron Lipman. Generator matching: Generative modeling with arbitrary Markov processes. In *International Conference on Learning Representations*, pp. 52153–52219, 2025.

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*, volume 35, pp. 26565–26577, 2022.

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

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

Whitney K. Newey and James L. Powell. Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578, 2003.

Alexander Quinn Nichol and Prafulla Dhariwal. Improved denoising diffusion probabilistic models. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8162–8171, 2021.

Tianyu Pang, Cheng Lu, Chao Du, Min Lin, Shuicheng Yan, and Zhijie Deng. On calibrating diffusion probabilistic models. In *Advances in Neural Information Processing Systems*, volume 36, pp. 49234–49249, 2023.

Philip E. Protter. *Stochastic Integration and Differential Equations*. Springer, 2nd edition, 2005.

Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion Schrödinger bridge matching. In *Advances in Neural Information Processing Systems*, volume 36, pp. 62183–62223, 2023.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021.

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

A MARTINGALE RESIDUALS AND PATH LAW IDENTIFICATION

We provide the assumptions and proofs for Section 3.1. We first state the admissibility conditions and derive the residual decomposition, then prove path law identification. The drift and covariance may both differ between the learned and reference processes.

A.1 ADMISSIBILITY AND RESIDUAL DECOMPOSITION

Let $X \sim Q$ and let $(\mathcal{F}_t)_{0\le t\le T}$ denote its natural filtration. We call $\phi \in C^{1,2}([0,T]\times\mathbb{R}^d)$ admissible if the endpoint evaluations $\phi(s,X_s)$, $\phi(t,X_t)$ and the generator integrals defining $M^{\phi}_{s,t}$ and $M^{\phi,\theta}_{s,t}$ are integrable for every $0 \le s \le t \le T$, and $(M^{\phi}_{0,t})_{0\le t\le T}$ is a Q-martingale. A sufficient condition for $(M^{\phi}_{0,t})_{0\le t\le T}$ to be a Q-martingale is

$$\mathbb{E}_Q \int_0^T \left\|\nabla_x \phi(s,X_s)^\top \sigma(s,X_s)\right\|_2^2 \,\mathrm{d}s < \infty. \tag{12}$$

If the test function, its derivatives, and the drift and diffusion coefficients have polynomial growth, suitable moment bounds on X ensure the required integrability (Protter, 2005).

Martingale property and residual decomposition. Applying Itô's formula under Q gives

$$M^{\phi}_{s,t} = \int_s^t \nabla_x \phi(r,X_r)^\top \sigma(r,X_r)\,\mathrm{d}W^Q_r.$$

For each fixed $s \in [0,T]$, condition 12 ensures that $(M^{\phi}_{s,t})_{s\le t\le T}$ is a square-integrable Q-martingale.

Subtracting the definitions of $M^{\phi,\theta}_{s,t}$ and $M^{\phi}_{s,t}$ gives equation 3. □

A.2 PROOF OF THEOREM 3.1

Proof. Fix an admissible test function ϕ and define $D^{\theta}_{\phi}(t,x) := (\mathcal{L}_t - \mathcal{L}^{\theta}_t)\phi(t,x)$. The reference residual is a martingale, so equations 3 and 4 give

$$\mathbb{E}_Q\left[z(X_s)\int_s^t D^{\theta}_{\phi}(r,X_r)\,\mathrm{d}r\right] = 0$$

for any bounded measurable z and $0 \le s < t \le T$. For $\psi \in C_c([0,T]\times\mathbb{R}^d)$ and a partition $0 = t_0 < \cdots < t_K = T$, take $s = t_k$, $t = t_{k+1}$, and $z(x) = \psi(t_k,x)$ in the equation above. Summing over k gives

$$\mathbb{E}_Q\left[\sum_{k=0}^{K-1}\int_{t_k}^{t_{k+1}} \psi(t_k,X_{t_k}) D^{\theta}_{\phi}(r,X_r)\,\mathrm{d}r\right] = 0.$$

As the mesh shrinks, path continuity and integrability give

$$\mathbb{E}_Q \int_0^T \psi(r,X_r) D^{\theta}_{\phi}(r,X_r)\,\mathrm{d}r = 0.$$

Because D^{θ}_{ϕ} depends only on (t,x) and ψ is arbitrary, this implies

$$(\mathcal{L}_t - \mathcal{L}^{\theta}_t)\phi(t,X_t) = 0 \qquad \mathrm{d}t \otimes Q\text{-almost everywhere.}$$

The residual decomposition now gives $M^{\phi,\theta}_{s,t} = M^{\phi}_{s,t}$ almost surely, and hence

$$\mathbb{E}_Q[M^{\phi,\theta}_{s,t} \mid \mathcal{F}_s] = 0. \tag{13}$$

Since this holds for every admissible ϕ, Q solves the learned martingale problem with initial law μ_0. Well-posedness then gives $Q = P_\theta$. □

Converse. If $Q = P_\theta$, the process $(M^{\phi,\theta}_{0,t})_{0\le t\le T}$ is a Q-martingale for every admissible ϕ. Since $z(X_s)$ is bounded and $\mathcal{F}_s$-measurable, the martingale orthogonality characterization in Section 2 gives equation 4.

A.3 Sufficient conditions for well-posedness

For fixed θ, a standard sufficient condition for well-posedness is that the learned drift b^θ and diffusion coefficient σ^θ are locally Lipschitz in the state and have at most linear growth uniformly in time. With an initial law of finite second moment, this gives a unique solution that does not explode on $[0, T]$ and a well-posed martingale problem (Protter, 2005; Ethier & Kurtz, 1986).

Schrödinger bridge models. The learned drift in the Schrödinger bridge experiment of Section 4.3 is parameterized by a finite SiLU MLP with sinusoidal time features and constant diffusion. The bounded derivative of SiLU and bounded time features give a globally Lipschitz drift with at most linear growth, so the learned martingale problem is well posed.

Image diffusion models. Factors from the noise schedule may become unbounded near $t = 0$ or $t = T$. We consider a time interval where these factors remain bounded. If the full drift and diffusion coefficients satisfy the Lipschitz and growth conditions above, the learned martingale problem is well posed on this interval.

B Coordinate Test Functions and Instrument Projections

Following Section 3.2, assume that the learned process uses the reference process's known covariance, $a^\theta = a$. We prove the sufficiency of coordinate test functions, the complete instrument identity and its DDPM specialization, and the finite projection formula.

B.1 Sufficiency of coordinate test functions

Theorem B.1 shows that the learned coordinate residual processes are Q-martingales if and only if the learned and reference drifts agree along reference paths. This equivalence is used to prove Proposition 3.2. We define the learned unnormalized coordinate MR by

$$R^\theta_{s,t} := X_t - X_s - \int_s^t b^\theta(r, X_r)\,\mathrm{d}r.$$

Its components are the learned MRs for the coordinate test functions, and $R^\theta_{t,t+h} = h\, r^\theta_{t,h}$. The reference residual $R_{s,t}$ is defined analogously, with b^θ replaced by b.

Theorem B.1. *Suppose $a^\theta = a$ and the coordinate test functions are admissible. The following are equivalent:*

1. *$(R^\theta_{0,t})_{0\le t\le T}$ is a Q-martingale;*
2. *$(M^{\phi,\theta}_{0,t})_{0\le t\le T}$ is a Q-martingale for every admissible smooth test function ϕ;*
3. *$b^\theta(t, X_t) = b(t, X_t)$ for $\mathrm{d}t \otimes Q$-almost every (t, ω).*

If the initial laws agree and the martingale problem for $(\mathcal{L}^\theta, \mu_0)$ is well posed, these conditions imply $Q = P_\theta$.

Proof. (1) $\Rightarrow$ (3). By admissibility, $(R_{0,t})_{0\le t\le T}$ is a Q-martingale. If $(R^\theta_{0,t})_{0\le t\le T}$ is also a Q-martingale, then

$$R^\theta_{0,t} - R_{0,t} = \int_0^t \big(b - b^\theta\big)(s, X_s)\,\mathrm{d}s$$

is a continuous martingale of finite variation starting at zero, hence is identically zero. This gives condition 3.

(3) $\Rightarrow$ (1). Equal drifts give $R^\theta_{0,t} = R_{0,t}$, which is a Q-martingale.

(1) $\Rightarrow$ (2). Since $a^\theta = a$, applying Itô's formula to $\phi(t, X_t)$ gives

$$\mathrm{d}M^{\phi,\theta}_{0,t} = \nabla_x\phi(t, X_t)^\top\big(\mathrm{d}X_t - b^\theta(t, X_t)\,\mathrm{d}t\big) = \nabla_x\phi(t, X_t)^\top\,\mathrm{d}R^\theta_{0,t}. \tag{14}$$

Equation 14 and admissibility then imply condition 2.

(2) $\Rightarrow$ (1). Choosing $\phi(x) = x_i$ for each coordinate i gives condition 1.

Finally, the common initial law and well-posedness yield $Q = P_\theta$. □

B.2 Proof of Proposition 3.2

Assume that the coordinate test functions are admissible and that

$$\mathbb{E}_Q \int_0^T \left\| b(t, X_t) - b^\theta(t, X_t) \right\|_2 \, \mathrm{d}t < \infty.$$

Proof. Fix a coordinate i and write $e_i(t,x) := b_i(t,x) - b_i^\theta(t,x)$. Since $(R_{0,t})_{0\le t\le T}$ is a Q-martingale, $\mathbb{E}_Q[z(X_s)(R_{s,t})_i] = 0$ for every bounded measurable z. Subtracting this identity from equation 5 yields

$$\mathbb{E}_Q\left[z(X_s) \int_s^t e_i(r, X_r) \, \mathrm{d}r \right] = 0$$

for every bounded measurable z and every $0 \le s < t \le T$. Applying the time partition argument from Appendix A.2, with e_i in place of D_ϕ^θ, gives

$$e_i(t, X_t) = 0 \qquad \mathrm{d}t \otimes Q\text{-almost everywhere.}$$

Repeating the argument for each coordinate yields

$$b^\theta(t, X_t) = b(t, X_t) \qquad \mathrm{d}t \otimes Q\text{-almost everywhere.}$$

By Theorem B.1, this drift equality, the common initial law, and well-posedness imply $Q = P_\theta$.

Conversely, drift equality gives $R^\theta_{s,t} = R_{s,t}$, so $(R^\theta_{0,t})_{0\le t\le T}$ is a Q-martingale. Therefore, for every coordinate i, bounded measurable z, and $0 \le s < t \le T$,

$$\mathbb{E}_Q\big[z(X_s)(R^\theta_{s,t})_i \big] = 0,$$

which is equation 5. □

B.3 Proof of Theorem 3.3

Retain the notation of Section 3.2 and write $e_\theta = b - b^\theta$. Assume that the coordinate test functions are admissible, the residuals and drift error are square integrable, and

$$\frac{1}{h} \int_t^{t+h} e_\theta(s, X_s) \, \mathrm{d}s \longrightarrow e_\theta(t, X_t) \quad \text{in mean square.}$$

Proof. Set $g_h := \mathbb{E}_Q[r^\theta_{t,h} \mid t, X_t]$ and abbreviate $z_\ell = z_\ell(t, X_t)$. The tower property and Parseval's identity applied coordinatewise give

$$\sum_{\ell=1}^{\infty} \left\| \mathbb{E}_Q[z_\ell r^\theta_{t,h}] \right\|_2^2 = \sum_{\ell=1}^{\infty} \left\| \mathbb{E}_Q[z_\ell g_h] \right\|_2^2 = \|g_h\|^2_{L^2(\mu)}.$$

Under Q, the residual decomposes as

$$r^\theta_{t,h} = \frac{1}{h} \int_t^{t+h} e_\theta(s, X_s) \, \mathrm{d}s + \frac{1}{h} \int_t^{t+h} \sigma(s, X_s) \, \mathrm{d}W^Q_s.$$

By the martingale property, the stochastic integral has conditional mean zero given (t, X_t). Hence

$$g_h(t, X_t) = \mathbb{E}_Q\left[\frac{1}{h} \int_t^{t+h} e_\theta(s, X_s) \, \mathrm{d}s \,\middle|\, t, X_t \right].$$

The assumed mean square convergence and the L^2 contraction property of conditional expectation imply

$$\|g_h - e_\theta\|^2_{L^2(\mu)} \le \mathbb{E}_Q \left\| \frac{1}{h} \int_t^{t+h} e_\theta(s, X_s) \, \mathrm{d}s - e_\theta(t, X_t) \right\|_2^2 \longrightarrow 0.$$

The Parseval identity and this convergence prove the theorem. □

B.4 Connection to DDPM

Proof of equation 7. The VP perturbation can be written as $X_t = \alpha_t X_{\text{data}} + \sigma_t \epsilon$, where X_{data} is a data sample and $\epsilon \sim \mathcal{N}(0, I)$ is independent of it. The VP drift difference satisfies

$$b(t,x) - b^\theta(t,x) = \beta(t)\big(s^\star(t,x) - s^\theta(t,x)\big), \qquad s^\star = -\epsilon^\star/\sigma_t, \quad s^\theta = -\epsilon_\theta/\sigma_t,$$

where $\beta(t) > 0$ and $\epsilon^\star(t, X_t) = \mathbb{E}[\epsilon \mid t, X_t]$. For each fixed h, the tower property and Parseval's identity give

$$\sum_{\ell=1}^{\infty} \left\| \mathbb{E}_Q\left[z_\ell(t, X_t) \frac{\sigma_t}{\beta(t)} r^\theta_{t,h} \right] \right\|_2^2 = \left\| \frac{\sigma_t}{\beta(t)} \mathbb{E}_Q[r^\theta_{t,h} \mid t, X_t] \right\|^2_{L^2(\mu)} \xrightarrow{h \downarrow 0} \left\| \frac{\sigma_t}{\beta(t)} e_\theta \right\|^2_{L^2(\mu)}.$$

The limit follows from Appendix B.3 because $\sigma_t/\beta(t)$ is bounded on the time interval. Expressing the drift difference in terms of noise predictions gives

$$\left\| \frac{\sigma_t}{\beta(t)} e_\theta \right\|^2_{L^2(\mu)} = \mathbb{E} \left\| \epsilon_\theta(t, X_t) - \epsilon^\star(t, X_t) \right\|_2^2,$$

Define $\mathcal{L}_{\text{DDPM}}(\theta) := \mathbb{E} \left\| \epsilon_\theta(t, X_t) - \epsilon \right\|_2^2$ and $C_{\text{irr}} := \mathbb{E} \left\| \epsilon - \epsilon^\star(t, X_t) \right\|_2^2$. Conditional orthogonality gives

$$\mathcal{L}_{\text{DDPM}}(\theta) = C_{\text{irr}} + \mathbb{E} \left\| \epsilon_\theta(t, X_t) - \epsilon^\star(t, X_t) \right\|_2^2 = C_{\text{irr}} + \left\| \frac{\sigma_t}{\beta(t)} e_\theta \right\|^2_{L^2(\mu)}.$$

This proves equation 7. □

B.5 Proof of equation 9

Proof. Let $g_h := \mathbb{E}_Q[r^\theta_{t,h} \mid t, X_t] \in L^2(\mu)$. The tower property gives

$$C_{\theta,h} = \mathbb{E}_Q[Z(r^\theta_{t,h})^\top] = \mathbb{E}_Q[Z g_h^\top].$$

Let $c_{h,j}$ denote the jth column of $C_{\theta,h}$. Then $P_{\mathcal{S}} g_{h,j} = (G^\dagger c_{h,j})^\top Z$. Therefore

$$\|P_{\mathcal{S}} g_{h,j}\|^2_{L^2(\mu)} = c_{h,j}^\top G^\dagger c_{h,j}.$$

Summing over j gives

$$\|P_{\mathcal{S}} g_h\|^2_{L^2(\mu)} = \sum_{j=1}^{d} c_{h,j}^\top G^\dagger c_{h,j} = \operatorname{tr}(C_{\theta,h}^\top G^\dagger C_{\theta,h}).$$

This proves equation 9. □

C Discretization Error

Algorithm 1 uses left endpoint approximations to compute MR increments from discrete reference paths. We control the discretization error and show that the numerical conditional residual mean has the same short interval limit as in Theorem 3.3.

Let $0 = t_0 < \cdots < t_K = T$ and $\Delta t_k = t_{k+1} - t_k$. For an admissible test function ϕ, approximate the drift integral in $M^{\phi,\theta}_{t_k,t_{k+1}}$ by the left endpoint rule,

$$\widetilde{M}^{\phi,\theta}_k := \phi(t_{k+1}, X_{t_{k+1}}) - \phi(t_k, X_{t_k}) - (\partial_t + \mathcal{L}^\theta_{t_k})\phi(t_k, X_{t_k}) \Delta t_k. \tag{15}$$

Proposition C.1. *Let $F(t,x) := (\partial_t + \mathcal{L}^\theta_t)\phi(t,x)$ belong to $C^{1,2}$. Suppose $(\partial_t + \mathcal{L}_t)F$ and $\sigma^\top \nabla_x F$ are bounded. There is a constant C independent of the time grid such that*

$$\begin{aligned} \left| \mathbb{E}_Q[\widetilde{M}^{\phi,\theta}_k - M^{\phi,\theta}_{t_k,t_{k+1}} \mid \mathcal{F}_{t_k}] \right| &\le C \Delta t_k^2, \\ \mathbb{E}_Q[|\widetilde{M}^{\phi,\theta}_k - M^{\phi,\theta}_{t_k,t_{k+1}}|^2 \mid \mathcal{F}_{t_k}] &\le C \Delta t_k^3. \end{aligned} \tag{16}$$

Proof. The residual difference satisfies

$$\widetilde{M}_k^{\phi,\theta} - M_{t_k,t_{k+1}}^{\phi,\theta} = \int_{t_k}^{t_{k+1}} \big[F(s, X_s) - F(t_k, X_{t_k})\big]\,\mathrm{d}s.$$

By Itô's formula and the boundedness assumptions, there exists a constant C, independent of the time grid, such that

$$|\mathbb{E}_Q[F(s, X_s) - F(t_k, X_{t_k}) \mid \mathcal{F}_{t_k}]| \le C(s - t_k),$$
$$\mathbb{E}_Q[|F(s, X_s) - F(t_k, X_{t_k})|^2 \mid \mathcal{F}_{t_k}] \le C(s - t_k).$$

Integrating the first inequality over $s \in [t_k, t_{k+1}]$ gives

$$\Big|\mathbb{E}_Q[\widetilde{M}_k^{\phi,\theta} - M_{t_k,t_{k+1}}^{\phi,\theta} \mid \mathcal{F}_{t_k}]\Big| \le C\Delta t_k^2.$$

For the second moment, Cauchy–Schwarz gives

$$\mathbb{E}_Q[|\widetilde{M}_k^{\phi,\theta} - M_{t_k,t_{k+1}}^{\phi,\theta}|^2 \mid \mathcal{F}_{t_k}] \le \Delta t_k \int_{t_k}^{t_{k+1}} C(s - t_k)\,\mathrm{d}s \le C\Delta t_k^3.$$

□

Coordinate test functions. For the coordinate test functions used in Algorithm 1, write $\widetilde{r}_{t,h}^{\theta} = (X_{t+h} - X_t)/h - b^{\theta}(t, X_t)$. Proposition C.1 gives

$$\big\|\mathbb{E}_Q[\widetilde{r}_{t,h}^{\theta} - r_{t,h}^{\theta} \mid t, X_t]\big\|_{L^2(\mu)} = O(h).$$

Combining this bound with Appendix B.3 gives

$$\big\|\mathbb{E}_Q[\widetilde{r}_{t,h}^{\theta} \mid t, X_t] - (b - b^{\theta})\big\|_{L^2(\mu)} \le \big\|\mathbb{E}_Q[r_{t,h}^{\theta} \mid t, X_t] - (b - b^{\theta})\big\|_{L^2(\mu)} + O(h) \longrightarrow 0.$$

Thus $\widetilde{r}_{t,h}^{\theta}$ and $r_{t,h}^{\theta}$ have the same limiting conditional mean $b - b^{\theta}$.

D Training with Learned Instruments

We derive the variational objective in equation 11 and its equivalent residual regression formulation. For fixed θ, we fit the auxiliary network by minimizing the residual regression loss

$$\mathbb{E}_Q\Big[\big\|g_\psi(t, X_t) - \widetilde{r}^{\theta}\big\|_2^2\Big]. \tag{17}$$

Expanding this loss gives

$$\mathbb{E}_Q\,\big\|g_\psi(t, X_t) - \widetilde{r}^{\theta}\big\|_2^2 = \mathbb{E}_Q\,\|g_\psi(t, X_t)\|_2^2 - 2\mathbb{E}_Q\big[g_\psi(t, X_t)^\top \widetilde{r}^{\theta}\big] + \mathbb{E}_Q\,\big\|\widetilde{r}^{\theta}\big\|_2^2.$$

The last term is independent of ψ, so minimizing equation 17 is equivalent to the inner maximization in equation 11.

With ψ fixed, the quadratic term in equation 11 is independent of θ. Absorbing the factor (2) into the step size, the model update uses

$$\mathbb{E}_Q\Big[\mathrm{stopgrad}[g_\psi(t, X_t)]^\top\,\widetilde{r}^{\theta}\Big], \tag{18}$$

which has the same weighted residual form as the expression following equation 9, with g_ψ replacing the fixed instrument weighting function. Finite regression steps approximate the inner optimization within the chosen network class. Algorithm 2 summarizes these alternating updates.

Algorithm 2 Training with learned instruments

Input Disjoint reference path sets $\mathcal{D}_\psi, \mathcal{D}_\theta$; drift b^θ and auxiliary network g_ψ; grid and time weights as in Algorithm 1; model update count M, network update count U, and optimizers.
for $m = 1, \dots, M$ **do**
  Sample a path batch $\mathcal{B}_\psi$ from $\mathcal{D}_\psi$.
  Compute $\widetilde{r}_k^{\theta,(n)}$ on $\mathcal{B}_\psi$ as in Algorithm 1, holding θ fixed.
  for $u = 1, \dots, U$ **do**
    Update ψ by minimizing equation 17 on $\mathcal{B}_\psi$.
  end for
  Sample a path batch $\mathcal{B}_\theta$ from $\mathcal{D}_\theta$.
  Compute $\widetilde{r}_k^{\theta,(n)}$ on $\mathcal{B}_\theta$ as in Algorithm 1.
  Update θ using equation 18 on $\mathcal{B}_\theta$, with g_ψ detached.
end for
return θ

E Experimental Details and Additional Results

The experiments follow the order of Section 4: diagnostics in Appendix E.1, regularization in Appendix E.2, and additional regularizer experiments in Appendix E.3. The sole objective studies then cover ImageNet-64 continuation in Appendix E.4, Schrödinger bridge learning in Appendix E.5, VP generation in Appendix E.6, and exploratory CIFAR-10 training in Appendix E.7.

E.1 MR score as a diagnostic

Figure 2 protocol. Section 4.1 evaluates MR scores during ordinary denoising training, with MR used only for evaluation. Reference paths are sampled at forward noise times in $[0.001, 0.95]$ and reversed into generation time. We compute numerical residuals using Algorithm 1 and apply the scaling in equation 7 at the corresponding forward noise times. MR scores and denoising MSE are averaged over output coordinates.

Table 4: Protocol for Figure 2.

Component	Setting
Model	Unconditional CIFAR-10, `improved-diffusion`, fixed variance
Training	4 seeds, 40k updates, batch 128, FP32, learning rate 10^{-4}
Checkpoints	EMA initialization and every 2k updates
Path law	cosine VP paths reversed into generation time
Time grid	128 intervals
Gram estimation	131,072 reference paths, one uniformly sampled interval per path
Intervals for $C_{\theta,h}$	One uniformly sampled interval from each of eight groups of 16 consecutive intervals
Test functions	all 3,072 coordinate test functions
Instruments	DCT state coefficients times normalized Legendre polynomials
Splitting	Separate path sets for G and $C_{\theta,h}$, with complete paths kept together
Cross products	Distinct pairs over 4 training seeds and 2 path seeds each

Using q_m and c_j from Section 4.1, we take $q_m(t)c_j(x)$ for $m = 0, \dots, q-1$ and $j = 1, \dots, D$. We also include $q_m(t)$, giving $q(D+1)$ instruments. We divide the reference paths used to estimate $C_{\theta,h}$ into $F \geq 3$ independent groups and average all distinct cross products:

$$\widehat{S}_\theta = \kappa(N_G, J)\binom{F}{2}^{-1} \sum_{1 \leq u < v \leq F} \operatorname{tr}\Big((\widehat{C}_{\theta,h}^{(u)})^\top (\widehat{G} + \lambda I)^{-1} \widehat{C}_{\theta,h}^{(v)}\Big).$$

Here N_G is the number of reference paths used to estimate G, J is the instrument count, and $\kappa(N_G, J) = \max\{(N_G - J - 2)/(N_G - 1), 0\}$. The reference paths used to estimate G and $C_{\theta,h}$ share the time design and the DDPM scaling in equation 7. Figure 2A reports the *decrease closure*

$$100\,\frac{\widehat{S}_{2\mathrm{k}} - \widehat{S}_{40\mathrm{k}}}{\widehat{\mathcal{L}}_{2\mathrm{k}} - \widehat{\mathcal{L}}_{40\mathrm{k}}}.$$

Here $\widehat{S}_n$ and $\widehat{\mathcal{L}}_n$ denote the estimated MR score and denoising MSE after n training updates.

Ablation studies. Table 5 examines the effect of increasing the maximum Legendre polynomial degree from three to seven while retaining all 3,072 DCT coefficients. The degree seven setting uses $q_0, \ldots, q_7$, or eight time polynomials. We vary the number N_G of reference paths used to estimate G while keeping those used to estimate $C_{\theta,h}$ fixed. Increasing the maximum time degree from three to seven improves the captured fraction by 2.59 percentage points with $131{,}072$ paths for estimating G, but only 1.12 points with $524{,}288$ paths. Thus, the smaller Gram sample gives a larger apparent benefit from higher time degrees. Inverting an estimated Gram matrix can introduce finite sample bias, which may partly explain this difference.

Table 5: Maximum time polynomial degree, using all DCT coefficients.

Degree	Instruments	$N_G = 131{,}072$	$N_G = 262{,}144$	$N_G = 524{,}288$
3	12,292	89.71%	89.57%	89.59%
4	15,365	90.60%	90.24%	90.20%
5	18,438	90.96%	90.50%	90.49%
6	21,511	91.66%	90.72%	90.64%
7	24,584	92.30%	90.84%	90.71%

Table 6 compares nonlinear extensions of the state features, with the maximum time degree fixed at seven and $524{,}288$ paths used to estimate G. The linear baseline spans all DCT coefficients. The second set adds the standardized coordinate features $H_2(y_i) = (y_i^2 - 1)/\sqrt{2}$, and the third adds $2{,}048$ fixed products of neighboring pixels across spatial positions and channels. Every state feature is multiplied by all eight time polynomials. The tested nonlinear features provide little improvement over the linear baseline. More expressive nonlinear instruments may yield larger gains.

Table 6: Nonlinear state features, with maximum time degree seven.

State feature set	Instruments	MSE decrease captured ($\%$)	95% interval
Linear	24,584	90.701%	$[90.658, 90.744]\%$
Linear + coordinate H_2	49,160	90.563%	$[90.525, 90.601]\%$
Linear + H_2 + neighbor products	65,544	90.753%	$[90.718, 90.788]\%$

E.2 MR SCORE AS A TRAINING REGULARIZER

This section documents the CIFAR-10 and ImageNet-64 rows of Table 1. Both datasets compare Compact with Coordinate-DCT8 and Coordinate-DCT32.

Training protocol. CIFAR-10 uses the official model with EDM preconditioning (Karras et al., 2022), while ImageNet-64 uses the OpenAI 121M model at 1.5M updates (Nichol & Dhariwal, 2021). Within each dataset, all arms start from the same source checkpoint and share the data split and evaluation protocol. All continuations use FP32 AdamW with betas $(0.9, 0.999)$, epsilon 10^{-8}, weight decay 0.01, and unit gradient clipping. The denoising baselines use the same optimizer settings as the corresponding MR runs.

The coefficient γ is calibrated separately for each configuration using gradients at the source checkpoint. The resulting values are reported in Table 7. Support/update gives the independent reference path counts used to estimate the stopped support moment and the update moment in each MR call.

Table 7: Regularizer training protocols.

Configuration	Updates	Learning rate	Batch	Support/update	γ
CIFAR Compact	5,000	10^{-7}	16	1,024/128	.00493394716
CIFAR Coordinate-DCT8	5,000	10^{-7}	16	1,024/128	1.65467460
CIFAR Coordinate-DCT32	5,000	10^{-7}	16	1,024/128	1.43850620
ImageNet Compact	500	2×10^{-8}	4	64/32	.0025312693
ImageNet Coordinate-DCT8	500	2×10^{-8}	4	64/32	7.93694142
ImageNet Coordinate-DCT32	500	2×10^{-8}	4	64/32	5.93759220

MR loss construction. We use the test functions and instruments described in Section 4.2. All configurations use unnormalized MR increments over four intervals of length $h = 0.19$, formed from the same five noise times. For Compact, the known diffusion contribution is included for the quadratic tests. We use the loss in equation 10 with a relative Gram ridge of 10^{-3}.

The coordinate configurations use 3,072 residuals on CIFAR-10 and 12,288 on ImageNet-64. DCT features are standardized using 32,768 paths, and another 32,768 paths estimate the fixed Gram matrix with relative ridge 10^{-3}.

E.3 Additional experiments on MR as a regularizer

These experiments examine different configurations and training settings for MR regularization. We specify the comparison protocol and the numbers of training and evaluation runs for each experiment.

E.3.1 Single batch and cross fitted MR

We compare two ways of estimating the Compact MR score for regularization and evaluate the resulting models using FID. *Single batch MR* reuses one moment estimate on both sides of the quadratic score, whereas *cross fitted MR* uses moment estimates from independent batches of reference paths, as in Algorithm 1. Using the same Gram normalization, the single batch and cross fitted estimates are, respectively,

$$\operatorname{tr}\Big(\widehat{C}_{\theta,h}^{\top}(\widehat{G}+\lambda I)^{-1}\widehat{C}_{\theta,h}\Big) \qquad \text{and} \qquad \operatorname{tr}\Big((\widehat{C}_{\theta,h}^{(1)})^{\top}(\widehat{G}+\lambda I)^{-1}\widehat{C}_{\theta,h}^{(2)}\Big).$$

Both methods supplement denoising training and are compared with denoising alone over eight paired training seeds.

The denoising baseline runs for 2,500 updates, whereas both MR methods run for 5,000 updates. The two MR configurations also differ in coefficients, normalization, and interval phases, so their FID difference cannot be attributed solely to the use of independent path batches.

Table 8: CIFAR-10 continuation with single batch and cross fitted MR regularization. Differences and intervals are relative to denoising alone.

Method	Updates	FID50k ↓	ΔFID ↓	95% interval	Favorable seeds
Denoising	2,500	2.0107	–	–	–
Single batch MR	5,000	1.9952	-0.0155	$[-0.0186, -0.0121]$	8/8
Cross fitted MR	5,000	1.9863	-0.0244	$[-0.0276, -0.0215]$	8/8

Each FID is averaged over two paired sampling ranges within each training seed, and the intervals are paired bootstrap intervals over the eight seed means. Cross fitted MR also improves over single batch MR by 0.0089 FID, with interval $[-0.0113, -0.0061]$, in all eight seeds. However, its independent 9,219 instrument MR ratio is 1.000547, with interval $[1.000518, 1.000576]$. The FID improvement therefore does not imply a decrease in every larger MR diagnostic.

E.3.2 Reference path budget

We reduce the support and update path counts from 1,024 and 128 to 64 and 32, respectively. Table 9 shows that this reduces training time by a factor of about eight, with only a slight increase in mean FID.

Table 9: Effect of the reference path budget.

MR estimator	Support/update paths	FID50k ↓	Δ vs. full	Training time (s)
Full	1,024/128	1.985089	–	24,876
Reduced	64/32	1.985383	+0.000294	3,101

E.3.3 A larger coordinate objective

We test a larger MR regularizer using all 3,072 coordinate test functions and 9,219 instruments formed by combining all 3,072 DCT coefficients with three Legendre time polynomials and adding three time only terms. Table 10 shows that the larger coordinate regularizer worsens FID relative to denoising alone, whereas Compact improves it.

Table 10: FID50k for the larger coordinate objective.

Method	FID50k ↓	Δ vs. denoising ↓
Denoising	2.0128	–
Compact MR	1.9873	−0.0254
Large coordinate MR	3.8723	+1.8595

E.4 ImageNet-64 continuation protocol

For the ImageNet-64 continuation in Section 4.3, we use the pretrained model and Compact MR construction in Appendix E.2, without a denoising loss. Each run starts from the original checkpoint with a fresh FP32 AdamW optimizer. We retain the learning rate of 2×10^{-8}, optimizer betas, epsilon, and gradient clipping from the regularizer protocol, using 64 support paths and 32 update paths per model update. The MR coefficient is 0.0025312692541976135, with a linear warmup over the first 100 updates. Three runs stop after 500 updates with weight decay 0.01. Three further runs use weight decay zero and continue for 2,000 updates without resetting the optimizer or restarting warmup.

We evaluate FID50k using deterministic, unguided DDIM50 sampling. Each training run uses a separate sampling range, shared by the original model and its evaluated checkpoints. The 500 update result averages all six runs, including the intermediate checkpoints of the longer runs. The 2,000 update result averages the three longer runs. Training settings and evaluation horizons are fixed before these runs. Table 2 reuses the Original and Denoising FID means from Table 1.

E.5 Schrödinger bridge protocol and additional results

This section documents Figure 3 and Table 3. Every model uses coordinate test functions and the MR score alone. The known covariance is $\sigma^2 I$ with $\sigma = 0.4$.

Reference paths. The source is an equally weighted mixture of eight isotropic Gaussians, each with standard deviation 0.12 and centered at equally spaced points on a circle of radius two. The target is a two moons distribution with added isotropic Gaussian noise of standard deviation 0.12. Each empirical marginal has 4,096 support points. We construct reference paths over the time interval $[0, T]$, with $T = 1$. The Sinkhorn kernel and coupling are

$$K_{ij} = \exp\left(-\frac{\|x_i - y_j\|_2^2}{2\sigma^2 T}\right), \qquad \pi^\star = \operatorname{diag}(u) K \operatorname{diag}(v),$$

where u and v are obtained by Sinkhorn scaling to match the empirical marginals. Conditional on $(x_i, y_j) \sim \pi^\star$, reference paths follow

$$X_t = (1-t)x_i + ty_j + \sigma(W_t - tW_1),$$

where W is a standard two dimensional Brownian motion. Independent training, calibration, and evaluation archives contain 65,536, 32,768, and 32,768 paths on 128 intervals.

Training protocols. All methods start from the same drift initialization and use the whitened coordinate residual

$$\widetilde{r}_k^{\theta,\mathrm{w}} = \frac{X_{t_{k+1}} - X_{t_k} - b^\theta(t_k, X_{t_k})h}{\sigma\sqrt{h}}, \qquad h = 1/128.$$

Fixed state features are multiplied by four normalized Legendre time polynomials. Their Gram matrices use 131,072 calibration observations, float64 accumulation, and relative ridge 10^{-3}.

The drift model shared by all methods is a SiLU MLP with four hidden layers of width 256 and a 32 dimensional sinusoidal time embedding. The learned configuration uses three independently initialized networks $g_{\psi_k}(t, X_t) \in \mathbb{R}^2$, each taking only the normalized current state and time, without X_0. Each network augments these inputs with Fourier features of both state and time, Legendre time polynomials through degree 15, and a logarithmic terminal time feature, then passes them through four SiLU residual blocks of width 256. Their predictions are averaged for the drift update.

At every update, support paths fit the fixed coefficients or auxiliary networks and independent paths update the drift. The auxiliary networks use learning rate 5×10^{-4} and 200 pretraining rounds with eight updates per network per round. Each model update uses 4,096 gradient paths, 16,384 support paths, and 16 of 128 intervals. Reported training times exclude validation.

Table 11: Schrödinger bridge training configurations.

Method	State representation	Count	Model LR	Step	Auxiliary updates	Time (s)
DCT	constant + 2 DCT features	12	10^{-3}	8,000	–	172.4
Hermite	degree ≤ 8 products	180	10^{-3}	8,000	–	193.8
RFF	base + 512 Fourier features	2,068	10^{-3}	8,000	–	241.4
RBF	base + 512 radial features	2,068	10^{-3}	8,000	–	277.0
Learned	3 auxiliary networks	neural	2.5×10^{-4}	7,000	8 per network	2,758.9

Evaluation protocol. Fixed instruments use the final models. For learned instruments, the lowest mean integrated SW2 on 32,768 independent validation paths selects update 7,000 from checkpoints 500, 1,000, …, 8,000. Evaluation uses four paired rollouts of 8,192 particles with 256 Euler–Maruyama steps. Integrated SW2 averages the distances at $t = 0.25, 0.5, 0.75, 1$, and terminal SW2 uses 512 random directions.

Histogram JS compares smoothed occupancy histograms. If D_1, D_2 are distances to the two target centerlines, separation failure is

$$\Pr\big(\min(D_1, D_2) > q^{\mathrm{ref}}_{.95}\big) + \Pr\big(|D_1 - D_2| < m^{\mathrm{ref}}_{.10}\big)\,.$$

Thickness W1 compares the generated and reference distributions of $\min(D_1, D_2)$.

E.6 Additional study using the sole objective: VP generation

We also train unconditional VP diffusion models from standard Gaussian noise to either eight Gaussians or thick two moons. Each task uses a time conditioned four layer MLP of width 128, coordinate test functions, 40,000 MR updates, and 500 Euler–Maruyama steps for generation. Normalized Legendre polynomials $q_0, \ldots, q_7$ multiply fixed state features: global polar and local cluster features for eight Gaussians, and radial basis plus local tangent and normal features for two moons. Each update selects 16 of 128 intervals, with 16,384 support paths and 8,192 independent gradient paths.

Results average five training runs and three paired evaluations of 10,000 samples per run. All eight modes are recovered in every MR run. Figure 4 shows the generation process from Gaussian noise at $t = 0$ to each target at $t = 1$. Table 12 shows lower mean SW2 for MR on two moons and a similar value on eight Gaussians.

Table 12: SW2 for VP generation.

Target	MR only	Denoising
Eight Gaussians	.04216	.04148
Two moons	.02260	.02415

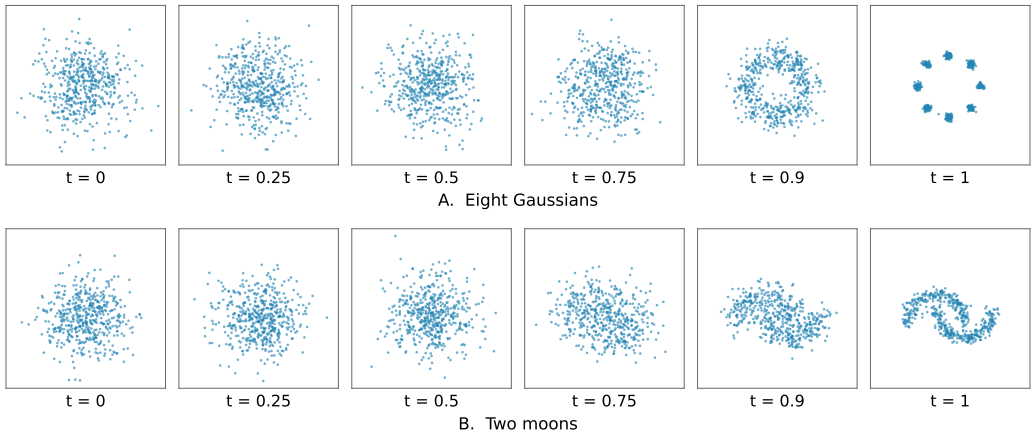

Figure 4: VP generation with MR alone. A. Eight Gaussians. B. Two moons. Each row follows the same 512 samples from $t = 0$ to $t = 1$.

E.7 Exploratory CIFAR-10 training with MR alone

We compare MR alone with denoising training using the DDPM model family in Section 4.1. Three paired random initializations are trained for 1,000 AdamW updates at learning rate 10^{-4}, with batch size 512. MR uses all 3,072 coordinate test functions and 9,219 instruments formed from the full DCT coefficient set, a constant state feature, and three Legendre time polynomials. Reference paths use 128 intervals over forward noise times $[0.001, 0.95]$. A fixed Gram matrix is estimated from 131,072 paths. The fitted moment coefficients are held fixed during model updates, and no denoising targets enter the MR loss.

Estimating the coefficients from 65,536 support paths and reusing them for 100 updates leads to instability in all three runs. Denoising MSE rises from approximately 1 to 6.4×10^3–9.3×10^3, accompanied by large model outputs. Before each model update, we recompute the weighting coefficients in equation 10 using 16,384 support paths. With this procedure, all three runs remain stable, and the independently evaluated MR score decreases by about 99.6%. Final denoising MSE is 0.1333–0.1348, compared with 0.0609–0.0610 for the paired denoising models. Thus, stable MR optimization is possible in this setting, but the large reduction in the finite MR score does not recover the denoising baseline within this training budget.