TRANSPORT ADJOINT MATCHING

Anonymous authors
Paper under double-blind review

ABSTRACT

Sampling from high-dimensional discrete distributions specified only through unnormalized mass functions is a fundamental problem in statistical physics, combinatorial optimization, and probabilistic inference. Adjoint matching provides an effective framework for learning nonequilibrium neural samplers, but the translation-based discrete Schrödinger-bridge construction relies on additive reference dynamics and deterministic cyclic endpoint correspondences, restricting its applicability when such structure is unavailable. We introduce Transport Adjoint Matching (TAM), which represents the optimal-control potential ratio through transport-compatible couplings between reference endpoint kernels, thereby removing the need for deterministic endpoint correspondences. For coordinate-local reference dynamics satisfying a shared-fiber marginal condition, we further derive a fiberwise Rao–Blackwellized target that marginalizes both the auxiliary transport variable and the local endpoint symbol. The resulting target preserves the exact population adjoint identity, requires only known local endpoint weights, local energy differences, and corrector potential ratios, and avoids explicit transport-matrix construction. We evaluate TAM on high-dimensional Ising and four-state Potts models across three temperature regimes, and further test a genuinely non-additive Potts reference for which the deterministic cyclic-shift construction is unavailable. TAM achieves accurate sampling across the standard benchmarks and remains effective under the non-additive reference.

1 INTRODUCTION

Sampling from unnormalized distributions is a fundamental problem in computational statistics (Liu, 2008; Brooks et al., 2011), Bayesian inference (Gelman et al., 2013), and statistical physics (Landau & Binder, 2014). Given an energy function $E : \mathcal{X} \to \mathbb{R}$ on a discrete state space $\mathcal{X}$, the goal is to sample from a probability mass function $\nu(x) \propto \exp(-\beta E(x))$ without evaluating its generally intractable partition function. In high-dimensional discrete systems, the state space grows combinatorially with dimension, while complex energy landscapes can create separated high-probability regions that cause conventional Markov chain Monte Carlo methods to mix slowly.

Neural samplers address this difficulty by learning nonequilibrium stochastic dynamics that transport samples from a tractable source distribution toward the target, thereby shifting part of the sampling cost to training. Stochastic optimal control and Schrödinger bridges provide unified path-space formulations for constructing such dynamics (Léonard, 2014; Chen et al., 2016; Zhang & Chen, 2022; Vargas et al., 2023). Adjoint matching turns their optimality conditions into tractable regression objectives that can be optimized using trajectories generated by the current controlled process, without requiring independent target samples (Domingo i Enrich et al., 2025; Havens et al., 2025). Recent work has extended this principle to discrete continuous-time Markov chains. Discrete Adjoint Matching derives a general discrete adjoint estimator for entropy-regularized fine-tuning (So et al., 2026), while DASBS develops a discrete Schrödinger-bridge sampler whose controller construction exploits cyclic additive reference dynamics (Guo et al., 2026).

In DASBS, the central controller construction uses a deterministic cyclic translation to transform the reference endpoint distribution associated with one current state into that associated with a neighboring state, thereby producing a tractable regression target for the potential ratio. This deterministic correspondence makes the controller target tractable, but ties the construction to endpoint kernels that admit such a map. For reference dynamics without an additive structure, the endpoint distributions may admit a probabilistic transport while lacking any single deterministic map that aligns them

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.

exactly. This raises a more focused question: does the DASBS controller construction intrinsically require a cyclic translation, or only an appropriate transport relation between endpoint distributions?

We introduce **Transport Adjoint Matching** (**TAM**) to answer this question. TAM connects the reference endpoint kernels associated with neighboring current states through a Markov coupling satisfying the required marginal constraints. We show that any such transport-compatible coupling yields an exact population identity for the potential ratio defining the optimal control. This reformulates the endpoint-alignment requirement in the translation-based DASBS controller construction as a marginal compatibility condition between endpoint distributions, rather than a deterministic algebraic transformation of the state space.

Directly sampling from a transport coupling introduces additional randomness and may require explicitly representing a transport kernel. For coordinate-local reference dynamics satisfying a shared-fiber marginal condition, we exploit the common context marginal of the endpoint kernels and partition the state space into low-dimensional fibers. By successively marginalizing the auxiliary transport variable and the local endpoint symbol within each fiber, we obtain a fiberwise Rao–Blackwellized target. This target preserves the exact population adjoint identity, is independent of the particular compatible transport, and can be evaluated using known local endpoint weights together with a finite collection of local energy differences and corrector potential ratios. TAM therefore avoids enumerating the full state space or explicitly storing transport matrices when applied to high-dimensional discrete systems.

Table 1 summarizes the resulting progression: DASBS constructs the adjoint target through a deterministic cyclic translation, whereas TAM generalizes this construction through transport compatibility and reduces the resulting transport target to a fiberwise Rao–Blackwellized regression target. The notation is defined in Sections 3 and 4.

We evaluate TAM on two-dimensional Ising and four-state Potts models across high-, critical-, and low-temperature regimes, and additionally test a genuinely non-additive Potts reference for which the DASBS cyclic-shift construction is unavailable. Across the standard benchmarks, TAM achieves accurate sampling, while the non-additive experiment directly tests the regime targeted by our generalization.

In summary, our contributions are as follows:

- We introduce Transport Adjoint Matching, which represents optimal-control potential ratios through transport-compatible couplings between reference endpoint kernels, removing the dependence on deterministic cyclic translations and additive reference structures.
- We derive, under a shared-fiber marginal condition, a fiberwise Rao–Blackwellized objective that analytically marginalizes the auxiliary transport variable and the local endpoint symbol, eliminating local transport randomness while preserving the exact population adjoint identity and avoiding explicit transport-matrix construction.
- We integrate the resulting objective into a practical controller–corrector training framework and validate TAM on high-dimensional Ising and four-state Potts models across different temperature regimes, including a genuinely non-additive reference for which the DASBS cyclic-shift construction is not applicable.

2 Related Work

Neural samplers for unnormalized distributions. Learning-based samplers amortize the cost of drawing from an unnormalized target by training parameterized nonequilibrium dynamics. In continuous spaces, recent approaches have been developed from diffusion processes, flow-based constructions, and stochastic optimal control (Zhang & Chen, 2022; Vargas et al., 2023; Havens et al., 2026). Discrete neural samplers similarly exploit continuous-time Markov chains (CTMCs), but differ substantially in their training principles. LEAPS learns locally equivariant transition rates using path-space importance weights (Holderrieth et al., 2025), while discrete neural flow samplers learn CTMC rates through the Kolmogorov equation with variance-reduced objectives (Ou et al., 2025). MDNS formulates masked discrete diffusion sampling through stochastic optimal control, with a corresponding uniform-diffusion extension (Zhu et al., 2025). Gradient-informed MCMC provides another direction: Gibbs-with-Gradients (GWG) constructs locally informed Metropolis–Hastings

Table 1: Conceptual comparison between the translation-based DASBS construction and Transport Adjoint Matching. DASBS uses additive endpoint kernels and a deterministic cyclic translation, whereas TAM requires only transport compatibility between endpoint distributions.

	DASBS	**Transport Adjoint Matching**
Endpoint structure	$K_{t,x}(z) = q_t(z - x)$	$K_x C_t^{x\to y} = K_y$
Endpoint transport	$S_{x\to y}(z) = z + y - x, \quad Z' = S_{x\to y}(Z)$	$Z' \sim C_t^{x\to y}(\cdot \mid Z)$
Adjoint Matching	$\frac{\varphi_t(y)}{\varphi_t(x)} = \mathbb{E}_{Z\sim p^\star_{1\mid t}(\cdot\mid x)} \frac{h(S_{x\to y}(Z))}{h(Z)}$	$\frac{\varphi_t(y)}{\varphi_t(x)} = \mathbb{E}_{(Z,Z')\sim \tilde{\Pi}^{x,y}_{t,h}} \frac{h(Z')}{h(Z)} = \mathbb{E}_{A\sim \eta^h_{t,x}} R_t^{x\to y}(A)$
Controller target	$\frac{h(S_{x\to y}(Z))}{h(Z)}$	$R_t^{x\to y}(A) = \frac{H_q(A)}{H_p(A)}$
Practical target	$e^{-\beta(E(S_{x\to y}(Z))-E(Z))}\widehat{c}_\psi(Z, S_{x\to y}(Z))$	$\widehat{R}^{x\to y}_{\psi,t}(A) = \frac{\sum_j q_j(A)\widehat{\gamma}_{\psi,j}(A;j_0)}{\sum_j p_j(A)\widehat{\gamma}_{\psi,j}(A;j_0)}$
Controller loss	$\mathbb{E}\Big[D_f\Big(\mathrm{sg}\Big[e^{-\beta\left(E(S_{X_t\to y}(Z))-E(Z)\right)}\widehat{c}_\psi(Z, S_{X_t\to y}(Z))\Big] \,\|\, g_\theta(t,y,X_t)\Big)\Big]$	$\mathbb{E}\Big[D_f\Big(\mathrm{sg}\Big[\widehat{R}^{X_t\to y}_{\psi,t}(A)\Big] \,\|\, g_\theta(t,y,X_t)\Big)\Big]$

proposals from gradients of a continuous relaxation of the discrete target (Grathwohl et al., 2021). TAM is complementary to these approaches: it focuses on the adjoint objective itself, developing a transport-compatible construction for learning the optimal control.

Adjoint matching and Schrödinger-bridge samplers. Adjoint Matching (AM) casts stochastic optimal control as a regression problem for continuous-state dynamics (Domingo i Enrich et al., 2025). Adjoint Sampling specializes this principle to scalable sampling from unnormalized densities under a memoryless reference structure (Havens et al., 2025), while the Adjoint Schrödinger Bridge Sampler (ASBS) embeds AM in a Schrödinger-bridge formulation and relaxes this restriction through alternating controller–corrector updates (Liu et al., 2025). Related work has also explored informative non-equilibrium reference processes, such as annealed dynamics in NAAS (Choi et al., 2025). These works show that adjoint-based objectives can support scalable diffusion-sampler training without requiring target-sample estimation, but their matching identities are formulated for continuous-state dynamics.

Discrete adjoint matching. Discrete Adjoint Matching (DAM) derives a general CTMC adjoint estimator for entropy-regularized fine-tuning from a statistical perspective and develops importance-weighted estimators for tractable learning in large discrete spaces (So et al., 2026). DASBS provides the closest foundation to our sampling setting by extending Schrödinger-bridge and adjoint-matching principles to discrete CTMCs (Guo et al., 2026). Its controller identity relies on a cyclic group structure and additive endpoint kernels, for which a deterministic translation maps samples from one reference endpoint distribution to another. TAM addresses this endpoint-alignment requirement: it preserves the DASBS optimal-control quantity and controller–corrector framework while generalizing the deterministic endpoint correspondence through transport compatibility. The resulting identity recovers the DASBS construction as a deterministic special case, while fiberwise Rao–Blackwellization removes the auxiliary transport randomness and yields the practical local target summarized in Table 1.

3 Preliminaries

This section reviews the discrete Schrödinger bridge quantities and the DASBS cyclic-translation adjoint identity needed to derive TAM. We work on the finite state space $\mathcal{X} = [N]^D$, with $[N] := \{1, \dots, N\}$ denoting the local alphabet and D the number of coordinates.

3.1 Discrete Schrödinger Bridges

Let $r_t(y, x)$ denote the transition rate from x to y of a reference continuous-time Markov chain, and let p^r be its path measure with initial marginal $p^r_0 = \mu$. The discrete Schrödinger bridge seeks the controlled path measure closest to p^r while matching the target terminal distribution ν:

$$p^\star \in \underset{p^u:\, p^u_0=\mu,\, p^u_1=\nu}{\arg\min} D_{\mathrm{KL}}(p^u \| p^r). \tag{1}$$

Its optimal off-diagonal transition rates take the Doob-transform form (Léonard, 2014; Guo et al., 2026)

$$u_t^\star(y,x) = r_t(y,x)\,\frac{\varphi_t(y)}{\varphi_t(x)}, \qquad y \neq x, \tag{2}$$

where φ_t is a positive Schrödinger potential. Learning the optimal control therefore reduces to estimating $\varphi_t(y)/\varphi_t(x)$ for state pairs connected by the reference dynamics.

For $t \in [0,1]$ and $x \in \mathcal{X}$, define the reference endpoint kernel $K_{t,x}$ and terminal potential h by

$$K_{t,x}(z) := p^r_{1|t}(z \mid x), \qquad h(z) := \varphi_1(z), \qquad \varphi_t(x) = K_{t,x}h := \sum_{z\in\mathcal{X}} K_{t,x}(z)h(z). \tag{3}$$

The terminal conditional distribution under the optimal path measure is

$$p^\star_{1|t}(z \mid x) = \frac{K_{t,x}(z)h(z)}{K_{t,x}h}. \tag{4}$$

Let $\widehat{\varphi}_t$ denote the second Schrödinger potential. The terminal boundary relation $\varphi_1(z)\widehat{\varphi}_1(z) = \nu(z)$ gives

$$\frac{h(z')}{h(z)} = \frac{\nu(z')}{\nu(z)}\,\frac{\widehat{\varphi}_1(z)}{\widehat{\varphi}_1(z')}. \tag{5}$$

The target ratio $\nu(z')/\nu(z)$ is available from the corresponding energy difference, while the ratio of the second potential is estimated by the DASBS corrector. TAM retains this controller–corrector parameterization while generalizing the translation-specific controller construction through transport compatibility, as developed in Section 4.

3.2 Cyclic-Translation Adjoint Identity in DASBS

For its cyclic construction, DASBS identifies the local alphabet with $\mathbb{Z}_N := \{0,\dots,N-1\}$ and uses additive reference endpoint kernels of the form $K_{t,x}(z) = q_t(z-x)$, with all state-space operations taken modulo N. For a state pair satisfying $r_t(y,x) > 0$, define the cyclic translation $S_{x\to y}(z) := z + y - x$. If $Z \sim K_{t,x}$, then $S_{x\to y}(Z) \sim K_{t,y}$. Combining this relation with Eq. (4), for $Z \sim p^\star_{1|t}(\cdot \mid x)$ we obtain

$$\frac{\varphi_t(y)}{\varphi_t(x)} = \mathbb{E}\left[\frac{h(S_{x\to y}(Z))}{h(Z)}\right]. \tag{6}$$

DASBS uses the random potential ratio inside the expectation as the controller regression target. This construction requires a deterministic translation that maps samples from $K_{t,x}$ exactly to $K_{t,y}$. TAM generalizes this deterministic endpoint correspondence to a probabilistic coupling satisfying the same endpoint marginal constraint.

4 Transport Adjoint Matching

4.1 Transport-Compatible Endpoint Couplings

Fix $t \in [0,1]$ and a state pair (x,y) satisfying $r_t(y,x) > 0$, and abbreviate $K_x := K_{t,x}$ and $K_y := K_{t,y}$. A Markov kernel $C_t^{x\to y}(z' \mid z)$ on $\mathcal{X}$ is *transport-compatible* if

$$\sum_{z\in\mathcal{X}} K_x(z)C_t^{x\to y}(z' \mid z) = K_y(z'), \qquad z' \in \mathcal{X}. \tag{7}$$

Equivalently, in row-vector notation, $K_x C_t^{x\to y} = K_y$. The kernel therefore induces a coupling of the two reference endpoint distributions. To match the optimal first marginal in Eq. (4), we further tilt this coupling by the terminal potential:

$$\Pi_t^{x,y}(z,z') := K_x(z)C_t^{x\to y}(z' \mid z), \qquad \widetilde{\Pi}_{t,h}^{x,y}(z,z') := \frac{h(z)}{K_x h}\,\Pi_t^{x,y}(z,z'). \tag{8}$$

The coupling $\Pi_t^{x,y}$ has marginals K_x and K_y. Under the tilted coupling, the first marginal becomes $p^\star_{1|t}(\cdot \mid x)$, while the conditional law of Z' given $Z = z$ remains $C_t^{x\to y}(\cdot \mid z)$.

Theorem 4.1 (Transport adjoint identity). *Suppose that $h > 0$ and that $C_t^{x\to y}$ satisfies Eq. (7). If $(Z, Z') \sim \widetilde{\Pi}_{t,h}^{x,y}$, then*

$$\mathbb{E}\left[\frac{h(Z')}{h(Z)}\right] = \frac{K_y h}{K_x h} = \frac{\varphi_t(y)}{\varphi_t(x)}. \tag{9}$$

Indeed, substituting Eq. (8) into the expectation cancels $h(Z)$, after which transport compatibility converts the remaining marginal from K_x to K_y. Thus, every transport-compatible coupling provides a population-unbiased regression target for the potential ratio in Eq. (2).

The deterministic construction of DASBS is recovered by choosing $C_t^{x\to y}(z' \mid z) = \mathbf{1}_{\{z'=S_{x\to y}(z)\}}$. Thus, as summarized in Table 1, transport compatibility generalizes the translation-based adjoint construction by reformulating deterministic endpoint alignment as a marginal compatibility condition between endpoint distributions. A compatible coupling always exists, for example the independent coupling $C_t^{x\to y}(z' \mid z) = K_y(z')$. Generic choices, however, may be dense and impractical in high-dimensional state spaces. We therefore exploit the local structure of discrete reference dynamics to eliminate the coupling from the practical target.

4.2 Fiberwise Rao–Blackwellization

Consider a coordinate-local transition $x \to y$ for which x and y differ only in coordinate d. Write an endpoint as $z = (a, j)$, where $a := z_{-d} \in [N]^{D-1}$ is the surrounding context and $j := z_d \in [N]$ is the local symbol. We impose the *shared-fiber marginal condition* that the two endpoint kernels have the same marginal over the context. Thus, the endpoint law may differ along the updated coordinate while sharing the same marginal over the remaining coordinates. For every a with $\kappa(a) > 0$, define the corresponding local endpoint distributions by

$$\begin{aligned} \kappa(a) &:= \sum_{j=1}^{N} K_x(a, j) = \sum_{j=1}^{N} K_y(a, j), \\ p_j(a) &:= \frac{K_x(a, j)}{\kappa(a)}, \qquad q_j(a) := \frac{K_y(a, j)}{\kappa(a)}. \end{aligned} \tag{10}$$

The shared-fiber marginal condition holds automatically for factorized endpoint kernels $K_{t,x}(z) = \prod_{\ell=1}^{D} k_{t,\ell}(z_\ell \mid x_\ell)$, since $x_\ell = y_\ell$ for every $\ell \neq d$. In that case, $p_j(a) = k_{t,d}(j \mid x_d)$ and $q_j(a) = k_{t,d}(j \mid y_d)$ are independent of a.

Any local Markov kernel C_a satisfying $p(a)C_a = q(a)$, where $p(a) := (p_j(a))_{j=1}^N$ and $q(a) := (q_j(a))_{j=1}^N$ are row vectors, defines a fiber-preserving transport between the endpoint kernels. Rather than constructing or sampling from C_a, TAM eliminates its explicit dependence through analytic marginalization and obtains a coupling-independent fiberwise target. Define $H_p(a) := \sum_{j=1}^N p_j(a)h(a, j)$ and $H_q(a) := \sum_{j=1}^N q_j(a)h(a, j)$. The resulting fiberwise Rao–Blackwellized target is

$$R_t^{x\to y}(a) := \frac{H_q(a)}{H_p(a)} = \frac{\sum_{j=1}^N q_j(a)h(a, j)}{\sum_{j=1}^N p_j(a)h(a, j)}. \tag{11}$$

Under the optimal endpoint distribution in Eq. (4), the context $A = Z_{-d}$ follows

$$\eta_{t,x}^h(a) := p_{1|t}^\star(A = a \mid x) = \frac{\kappa(a)H_p(a)}{K_x h}. \tag{12}$$

Theorem 4.2 (Fiberwise Rao–Blackwell identity). *Suppose that Eq. (10) holds. If $A \sim \eta_{t,x}^h$, then*

$$\mathbb{E}[R_t^{x\to y}(A)] = \frac{K_y h}{K_x h} = \frac{\varphi_t(y)}{\varphi_t(x)}. \tag{13}$$

The identity follows by substituting Eqs. (11) and (12) and summing $\kappa(a)H_q(a)$ over the contexts. Thus, $R_t^{x\to y}(A)$ provides an exact population-unbiased controller target and does not depend on the choice of the compatible local transport C_a.

This target admits two successive Rao–Blackwellization steps. Let J denote the local symbol of the sampled endpoint, so that $\mathbb{P}(J = j \mid A = a) = p_j(a)h(a, j)/H_p(a)$, and let $J' \sim C_A(\cdot \mid J)$ denote the transported symbol. Then

$$\begin{aligned} \widehat{R}_C(A, J, J') &:= \frac{h(A, J')}{h(A, J)}, \qquad T_C(A, J) := \mathbb{E}\Big[\widehat{R}_C(A, J, J') \mid A, J\Big], \\ R_t^{x\to y}(A) &= \mathbb{E}[T_C(A, J) \mid A] = \mathbb{E}\Big[\widehat{R}_C(A, J, J') \mid A\Big]. \end{aligned} \tag{14}$$

The last line follows from the compatibility of C_A. Consequently, the law of total variance yields

$$\mathrm{Var}[R_t^{x\to y}(A)] \le \mathrm{Var}[T_C(A, J)] \le \mathrm{Var}\Big[\widehat{R}_C(A, J, J')\Big]. \tag{15}$$

The first marginalization removes the auxiliary transport randomness in J', while the second removes the remaining variation in the sampled local symbol J. The residual randomness arises only from the endpoint context A.

4.3 Practical Objective and Training

The fiberwise target in Eq. (11) depends only on potential ratios within the observed context. Choose an anchor symbol $j_0 \in [N]$ and write $\Delta E_j(a; j_0) := E(a, j) - E(a, j_0)$. By Eq. (5),

$$\frac{h(a, j)}{h(a, j_0)} = e^{-\beta \Delta E_j(a;j_0)} \frac{\widehat{\varphi}_1(a, j_0)}{\widehat{\varphi}_1(a, j)}.$$

We estimate the unknown corrector factor with $\widehat{c}_\psi((a, j_0), (a, j)) \approx \widehat{\varphi}_1(a, j_0)/\widehat{\varphi}_1(a, j)$, which yields the practical local ratio and TAM target:

$$\begin{aligned} \widehat{\gamma}_{\psi,j}(a; j_0) &:= e^{-\beta \Delta E_j(a;j_0)}\, \widehat{c}_\psi((a, j_0), (a, j)), \\ \widehat{R}_{\psi,t}^{x\to y}(a) &:= \frac{\sum_{j=1}^N q_j(a)\widehat{\gamma}_{\psi,j}(a; j_0)}{\sum_{j=1}^N p_j(a)\widehat{\gamma}_{\psi,j}(a; j_0)}. \end{aligned} \tag{16}$$

When the corrector ratios are exact, $\widehat{R}_{\psi,t}^{x\to y}(a) = R_t^{x\to y}(a)$ independently of the anchor. Evaluating Eq. (16) requires $O(N)$ local energy and corrector evaluations and does not require constructing an $N \times N$ transport matrix. The numerator and denominator can be evaluated with log-sum-exp for numerical stability.

Let $g_\theta(t, y, x) > 0$ denote the controller prediction of $\varphi_t(y)/\varphi_t(x)$. For a differentiable strictly convex function f, write $D_f(u\|v) := f(u) - f(v) - f'(v)(u - v)$ for the associated Bregman divergence. By Theorem 4.2, for $A \sim \eta_{t,x}^h$ the desired potential ratio is the population minimizer

$$\frac{\varphi_t(y)}{\varphi_t(x)} = \arg\min_{g>0} \mathbb{E}_A[D_f(R_t^{x\to y}(A)\|g)]. \tag{17}$$

Following the alternating controller–corrector procedure of DASBS, we replace the unknown target with the estimate from Eq. (16). Let $\mathrm{sg}[\cdot]$ denote the stop-gradient operator, and define

$$\mathcal{N}_t(x) := \{y \in \mathcal{X} : r_t(y, x) > 0\}.$$

The practical controller loss is

$$\mathcal{L}_{\mathrm{TAM}}(\theta) := \mathbb{E}_{t,X_t,X_1}\left[\sum_{y\in\mathcal{N}_t(X_t)} D_f\Big(\mathrm{sg}\Big[\widehat{R}_{\psi,t}^{X_t\to y}(A_y)\Big] \,\|\, g_\theta(t, y, X_t)\Big)\right], \tag{18}$$

where, if $X_t \to y$ updates coordinate d, we set $A_y := X_{1,-d}$ and $j_{0,y} := X_1^d$. The expectation follows the DASBS sampling procedure: $t \sim \mathrm{Unif}[0, 1]$, an endpoint pair (X_0, X_1) is sampled from the current controlled process, and X_t is drawn from the corresponding reference bridge. The learned transition rate is $u_\theta(t, y, x) := r_t(y, x)g_\theta(t, y, x)$.

Training alternates between TAM controller updates using Eq. (18) and the corresponding DASBS corrector updates. Algorithm 1 summarizes the resulting procedure.

Algorithm 1 Transport Adjoint Matching (TAM)

Require: Energy E, inverse temperature β, reference dynamics r_t, controller g_θ, corrector $\widehat{c}_\psi$, and number of alternating stages K_{stage}.

1: **for** stage $k = 1, \dots, K_{\text{stage}}$ **do**
2: Construct the fiberwise TAM target $\widehat{R}^{x \to y}_{\psi,t}$ according to Eq. (16).
3: Update the controller θ by minimizing Eq. (18).
4: Update the corrector ψ using the DASBS corrector-matching objective (Guo et al., 2026).
5: **end for**

Ensure: Controlled transition rates $u_\theta(t, y, x) = r_t(y, x) g_\theta(t, y, x)$.

5 Experiments

We evaluate TAM on discrete statistical-physics benchmarks, examining its sampling quality on standard Ising and Potts tasks, the variance reduction induced by fiberwise Rao–Blackwellization, and its applicability beyond additive reference dynamics.

Table 2: Learning to sample from lattice Ising models with $L = 24$. **Best** and second best sampling metrics among all *uniform-based* discrete neural samplers are highlighted. * Steps and runtime correspond to the high-temperature setting; full per-task TAM costs are reported in Table 7. † MDNS uses warm-up at the critical- and low-temperature settings; ‡ UDNS and DNFS fail to converge there even with warm-up.

Inv. Temp.				$\beta_{\text{high}} = 0.28$			$\beta_{\text{critical}} = 0.4407$			$\beta_{\text{low}} = 0.6$		
Type	Metrics ↓	Steps ($\times 10^3$)*	Runtime (h)*	ΔMag.	ΔCorr.	EW_2	ΔMag.	ΔCorr.	EW_2	ΔMag.	ΔCorr.	EW_2
Uniform	**TAM (ours)**	3.0	0.2	3.4e − 3	1.0e − 3	**1.9**	**6.8e − 3**	**1.8e − 3**	**2.3**	1.5e − 2	**2.8e − 4**	**1.6**
	DASBS	3.0	0.2	1.5e − 2	2.3e − 3	5.4	5.7e − 1	1.2e − 1	53.7	**9.7e − 3**	1.9e − 3	2.5
	LEAPS	30	8.4	**1.8e − 3**	**9.2e − 4**	3.1	5.9e − 2	2.8e − 1	96.5	3.0e − 2	5.5e − 1	176.6
	UDNS‡	50	11.9	9.0e − 3	8.7e − 3	23.6	–	–	–	–	–	–
	DNFS‡	50	2.1	9.3e − 1	8.0e − 1	661.6	–	–	–	–	–	–
Masked	MDNS†	50	16.8	3.9e − 3	7.4e − 4	0.1	1.1e − 2	5.6e − 3	5.1	9.0e − 3	4.7e − 3	5.3
MCMC	MH	–	–	8.9e − 4	2.9e − 4	1.2	2.5e − 2	3.7e − 3	293.3	4.0e − 2	6.6e − 4	109.9

Table 3: Learning to sample from lattice four-state Potts models with $L = 16$ and $N = 4$. **Best** and second best results among all *uniform-based* discrete neural samplers are highlighted. Full per-task training costs for TAM are reported in Table 7.

Inv. Temp.		$\beta_{\text{high}} = 0.9$			$\beta_{\text{critical}} = 1.0986$			$\beta_{\text{low}} = 1.3$		
Type	Metrics ↓	ΔMag.	ΔCorr.	EW_2	ΔMag.	ΔCorr.	EW_2	ΔMag.	ΔCorr.	EW_2
Uniform	**TAM (ours)**	**2.5e − 3**	**7.5e − 4**	**0.9**	**3.4e − 3**	**1.6e − 3**	**1.0**	**9.9e − 4**	**2.5e − 4**	**0.4**
	DASBS	1.4e − 2	8.2e − 3	6.6	4.2e − 2	3.2e − 2	14.6	1.1e − 3	2.1e − 3	1.4
	LEAPS	5.7e − 3	5.0e − 3	8.2	3.2e − 1	2.6e − 1	79.9	3.6e − 1	3.5e − 1	90.5
Masked	MDNS	6.5e − 3	5.1e − 3	6.4	5.2e − 3	4.6e − 3	1.9	8.4e − 4	6.1e − 4	0.7
MCMC	MH	5.2e − 2	4.7e − 2	98.6	4.9e − 1	4.0e − 1	273.2	6.8e − 1	6.4e − 1	313.1

5.1 Experimental Setup

Target distributions. Following the evaluation protocol of DASBS (Guo et al., 2026), we consider the two-dimensional Ising model and the four-state Potts model. The Ising model is defined on a periodic 24×24 square lattice and evaluated in high-, critical-, and low-temperature regimes with inverse temperatures $\beta \in \{0.28, 0.4407, 0.6\}$. The four-state Potts model is defined on a periodic 16×16 square lattice with inverse temperatures $\beta \in \{0.9, 1.0986, 1.3\}$. The evaluation therefore comprises six sampling tasks.

Baselines. We compare TAM with DASBS (Guo et al., 2026), LEAPS (Holderrieth et al., 2025), UDNS and MDNS (Zhu et al., 2025), DNFS (Ou et al., 2025), and Metropolis–Hastings (MH). For the six standard benchmark settings in Tables 2 and 3, TAM retains the reference process and controller–corrector parameterization of DASBS while training the controller with the proposed fiberwise Rao–Blackwellized target.

Evaluation metrics. We follow the DASBS evaluation protocol (Guo et al., 2026) exactly, using the same definitions and normalization for the magnetization error $\Delta\text{Mag.}$, the two-point correlation error $\Delta\text{Corr.}$, and the energy Wasserstein-2 distance EW_2. All metrics are computed relative

to reference statistics estimated from independent long-run Swendsen–Wang chains (Swendsen & Wang, 1987), and lower values indicate better agreement with the target distribution.

Implementation details. Each TAM experiment uses three independent random seeds. The TAM results in Tables 2 and 3 are averaged over these seeds, with the corresponding standard deviations reported in Table 8. Network architectures, optimization settings, sampling discretization, and additional experimental details are provided in Section B.2.

5.2 Sampling Quality on Ising and Potts Models

Tables 2 and 3 report the sampling errors on the Ising and four-state Potts models across high-, critical-, and low-temperature regimes. Each table includes the magnetization error, two-point correlation error, and energy Wasserstein-2 distance, with lower values indicating better performance.

On the Ising model, TAM achieves the best results on all three metrics at criticality. It also achieves the best $E\mathcal{W}_2$ at high and low temperatures and the best $\Delta\mathrm{Corr.}$ at low temperature among uniform-based neural samplers.

On the four-state Potts model, TAM achieves the best results for all three metrics in the high- and critical-temperature regimes. In the low-temperature regime, it obtains the lowest correlation error and energy Wasserstein-2 distance while remaining close to the best magnetization error.

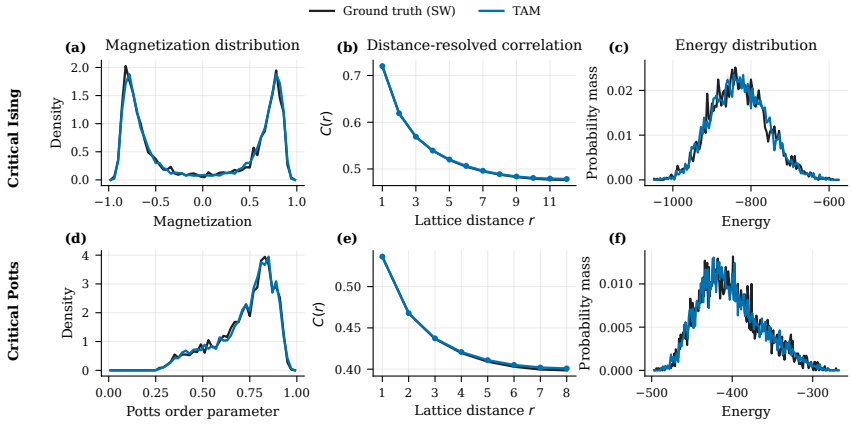


Figure 1: Distributional comparisons at criticality. We compare TAM with Ground Truth (SW) for the critical Ising model at $L = 24$, $\beta = 0.4407$ (top), and the critical four-state Potts model at $L = 16$, $\beta = 1.0986$ (bottom). From left to right, we show the order-parameter distribution, distance-resolved correlation, and energy distribution.

Beyond the aggregate metrics, Fig. 1 further compares the distributional structure at criticality. For both Ising and Potts, TAM closely matches Ground Truth (SW) in the order-parameter distribution, the distance-resolved correlation profile, and the energy distribution. In particular, the agreement of $C(r)$ across lattice distances complements the scalar correlation errors reported in Tables 2 and 3, showing that the observed performance is not confined to a single summary statistic.

5.3 Rao–Blackwell Target-Variance Results

As shown in Fig. 2, marginalizing the transport variable reduces target variance, while the full fiberwise Rao–Blackwellization yields a further reduction at every evaluated bridge time. The empirical ordering is consistent with the theoretical relation $\mathrm{Var}[R_t^{x\to y}(A)] \le \mathrm{Var}[T_C(A, J)] \le \mathrm{Var}[\widehat{R}_C(A, J, J')]$.

5.4 Controller-Target Ablation

For the sampled-transport and transport-marginalized targets, we use the same independent compatible local coupling $C_a(j' \mid j) = q_{j'}(a)$; all remaining training and evaluation settings are held fixed, as detailed in Section B.2.

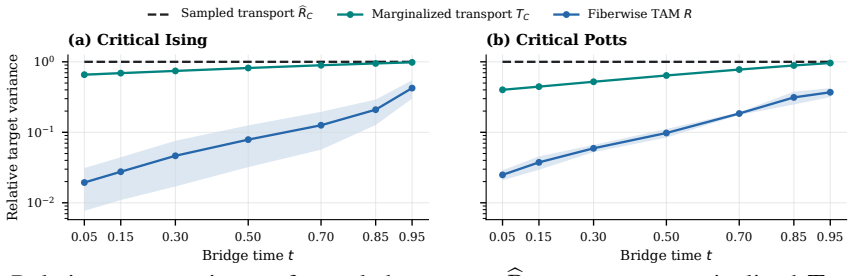


Figure 2: Relative target variance of sampled transport $\widehat{R}_C$, transport-marginalized T_C, and fiberwise TAM $R_t^{x\to y}$ on critical Ising (left) and Potts (right), normalized by $\mathrm{Var}[\widehat{R}_C]$ at each bridge time. Shaded bands show one standard deviation over three random seeds.

Table 4: Controller-target ablation on critical Ising and Potts. All settings are fixed except the regression target. Results are mean $\pm$ standard deviation over three random seeds; lower is better.

	Critical Ising ($\beta = 0.4407$)			Critical Potts ($\beta = 1.0986$)		
Controller target	ΔMag.	ΔCorr.	EW_2	ΔMag.	ΔCorr.	EW_2
Sampled transport $\widehat{R}_C$	$1.80\times10^{-2}\pm4.97\times10^{-4}$	$4.56\times10^{-3}\pm8.01\times10^{-4}$	3.366 ± 0.198	$5.73\times10^{-3}\pm3.76\times10^{-3}$	$4.45\times10^{-3}\pm3.73\times10^{-3}$	1.678 ± 0.676
Transport-marginalized T_C	$1.49\times10^{-2}\pm1.21\times10^{-3}$	$\mathbf{1.41\times10^{-3}\pm2.54\times10^{-4}}$	2.416 ± 0.205	$4.67\times10^{-3}\pm1.72\times10^{-3}$	$3.44\times10^{-3}\pm2.14\times10^{-3}$	1.357 ± 0.288
Fiberwise TAM $R_t^{x\to y}$	$\mathbf{1.47\times10^{-2}\pm1.56\times10^{-3}}$	$1.56\times10^{-3}\pm3.31\times10^{-4}$	$\mathbf{2.364\pm0.121}$	$\mathbf{3.42\times10^{-3}\pm1.70\times10^{-4}}$	$\mathbf{1.65\times10^{-3}\pm3.50\times10^{-4}}$	$\mathbf{1.039\pm0.038}$

Table 4 shows that directly sampling the auxiliary transport produces the largest errors on both critical tasks. Marginalization substantially improves sampling, with fiberwise TAM achieving the best results on two of three Ising metrics and all three Potts metrics. Together with Fig. 2, these results support Rao–Blackwellization of auxiliary transport randomness in the controller target.

5.5 Non-additive Reference Dynamics

Table 5: TAM on critical Potts under additive and non-additive references. Both rows use TAM; the DASBS column marks whether the original cyclic-shift target is applicable. Results are mean $\pm$ standard deviation over three random seeds; lower is better.

TAM reference	DASBS cyclic-shift target	ΔMag.	ΔCorr.	EW_2
Additive, $\rho = (0.25, 0.25, 0.25, 0.25)$	Applicable	$(3.4 \pm 0.2) \times 10^{-3}$	$(1.6 \pm 0.4) \times 10^{-3}$	1.0 ± 0.04
Non-additive, $\rho = (0.1, 0.2, 0.3, 0.4)$	Not applicable	$(4.2 \pm 2.1) \times 10^{-3}$	$(3.1 \pm 2.1) \times 10^{-3}$	1.4 ± 0.2

On critical Potts, we change only the coordinate refresh distribution from $\rho = (0.25, 0.25, 0.25, 0.25)$ to $\rho = (0.1, 0.2, 0.3, 0.4)$ while holding the remaining training and evaluation settings fixed. The non-additive reference breaks translation invariance while preserving the coordinate-factorized endpoint structure and hence the shared-fiber marginal condition in Eq. (10). Although the DASBS cyclic-shift target is no longer applicable, TAM remains well-defined and achieves accurate sampling.

6 Conclusion

We introduced Transport Adjoint Matching (TAM), which generalizes the translation-based DASBS adjoint construction by reformulating deterministic cyclic endpoint alignment as a marginal compatibility condition between endpoint distributions. For coordinate-local reference dynamics satisfying the shared-fiber marginal condition, fiberwise Rao–Blackwellization removes the auxiliary transport randomness and yields a local regression target with $O(N)$ cost in the local alphabet size, requiring only known local endpoint weights, energy differences, and corrector potential ratios. Experiments on Ising and Potts models across high-, critical-, and low-temperature regimes demonstrate strong sampling accuracy, while the non-additive reference experiment confirms that TAM remains applicable when the DASBS cyclic-shift construction does not apply. These results support transport compatibility as a general principle beyond additive references; extending the practical fiberwise construction to correlated or graph-structured references remains future work.

REFERENCES

Steve Brooks, Andrew Gelman, Galin Jones, and Xiao-Li Meng (eds.). *Handbook of markov chain monte carlo*. CRC press, 2011.

Yongxin Chen, Tryphon T Georgiou, and Michele Pavon. On the relation between optimal transport and schrödinger bridges: A stochastic control viewpoint. *Journal of Optimization Theory and Applications*, 169(2):671–691, 2016.

Jaemoo Choi, Yongxin Chen, Molei Tao, and Guan-Horng Liu. Non-equilibrium annealed adjoint sampler. In *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2025, NeurIPS 2025, San Diego, CA, USA, December 2-7, 2025 / Mexico City, Mexico, November 30 - December 5, 2025*, 2025.

Carles Domingo i Enrich, Michal Drozdzal, Brian Karrer, and Ricky T. Q. Chen. Adjoint matching: Fine-tuning flow and diffusion generative models with memoryless stochastic optimal control. In *International Conference on Learning Representations*, volume 2025, pp. 53791–53846, 2025.

Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. *Bayesian Data Analysis*. CRC Press, 3 edition, 2013.

Will Grathwohl, Kevin Swersky, Milad Hashemi, David Duvenaud, and Chris Maddison. Oops i took a gradient: Scalable sampling for discrete distributions. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 3831–3841. PMLR, 18–24 Jul 2021.

Wei Guo, Yuchen Zhu, Xiaochen Du, Juno Nam, Yongxin Chen, Rafael Gómez-Bombarelli, Guan-Horng Liu, Molei Tao, and Jaemoo Choi. Discrete adjoint schrödinger bridge sampler. In *Forty-third International Conference on Machine Learning*, 2026.

Aaron Havens, Brian Karrer, and Neta Shaul. Flow sampling: Learning to sample from unnormalized densities via denoising conditional processes. In *Forty-third International Conference on Machine Learning*, 2026.

Aaron J. Havens, Benjamin Kurt Miller, Bing Yan, Carles Domingo-Enrich, Anuroop Sriram, Daniel S. Levine, Brandon M. Wood, Bin Hu, Brandon Amos, Brian Karrer, Xiang Fu, Guan-Horng Liu, and Ricky T. Q. Chen. Adjoint sampling: Highly scalable diffusion samplers via adjoint matching. In *Forty-second International Conference on Machine Learning, ICML 2025, Vancouver, BC, Canada, July 13-19, 2025*, volume 267 of *Proceedings of Machine Learning Research*. PMLR / OpenReview.net, 2025.

Peter Holderrieth, Michael Samuel Albergo, and Tommi S. Jaakkola. LEAPS: A discrete neural sampler via locally equivariant networks. In *Forty-second International Conference on Machine Learning, ICML 2025, Vancouver, BC, Canada, July 13-19, 2025*, volume 267 of *Proceedings of Machine Learning Research*. PMLR / OpenReview.net, 2025.

David P. Landau and Kurt Binder. *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge University Press, 4 edition, 2014.

Christian Léonard. A survey of the schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems*, 34(4):1533–1574, 2014. doi: 10.3934/dcds.2014.34.1533.

Guan-Horng Liu, Jaemoo Choi, Yongxin Chen, Benjamin Kurt Miller, and Ricky T. Q. Chen. Adjoint schrödinger bridge sampler. In *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2025, NeurIPS 2025, San Diego, CA, USA, December 2-7, 2025 / Mexico City, Mexico, November 30 - December 5, 2025*, 2025.

Jun S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer, 2008. doi: 10.1007/978-0-387-76371-2.

Zijing Ou, Ruixiang Zhang, and Yingzhen Li. Discrete neural flow samplers with locally equivariant transformer. In *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2025, NeurIPS 2025, San Diego, CA, USA, December 2-7, 2025 / Mexico City, Mexico, November 30 - December 5, 2025*, 2025.

Oswin So, Brian Karrer, Chuchu Fan, Ricky T. Q. Chen, and Guan-Horng Liu. Discrete adjoint matching. In *International Conference on Learning Representations*, volume 2026, pp. 32834–32868, 2026.

Robert H Swendsen and Jian-Sheng Wang. Nonuniversal critical dynamics in monte carlo simulations. *Physical review letters*, 58(2):86–88, 1987.

Francisco Vargas, Will Sussman Grathwohl, and Arnaud Doucet. Denoising diffusion samplers. In *The Eleventh International Conference on Learning Representations, ICLR 2023, Kigali, Rwanda, May 1-5, 2023*. OpenReview.net, 2023.

Qinsheng Zhang and Yongxin Chen. Path integral sampler: A stochastic control approach for sampling. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*. OpenReview.net, 2022.

Yuchen Zhu, Wei Guo, Jaemoo Choi, Guan-Horng Liu, Yongxin Chen, and Molei Tao. MDNS: masked diffusion neural sampler via stochastic optimal control. In *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2025, NeurIPS 2025, San Diego, CA, USA, December 2-7, 2025 / Mexico City, Mexico, November 30 - December 5, 2025*, 2025.

AI USE STATEMENT

Generative AI tools were used for assistance with manuscript editing, literature organization, and implementation debugging. All theoretical developments, experiments, and final claims were independently verified by the authors.

A PROOFS

A.1 PROOF OF THE TRANSPORT ADJOINT IDENTITY

Proof of Theorem 4.1. The argument follows by canceling the $h(Z)$ factor introduced by the potential tilt and then applying transport compatibility to the remaining endpoint marginal. By the definition of the tilted coupling in Eq. (8),

$$\begin{aligned}\mathbb{E}_{(Z,Z')\sim\widetilde{\Pi}_{t,h}^{x,y}}\left[\frac{h(Z')}{h(Z)}\right] &= \sum_{z,z'\in\mathcal{X}} \frac{h(z)}{K_x h} K_x(z) C_t^{x\to y}(z' \mid z) \frac{h(z')}{h(z)} \\ &= \frac{1}{K_x h} \sum_{z'\in\mathcal{X}} h(z') \sum_{z\in\mathcal{X}} K_x(z) C_t^{x\to y}(z' \mid z).\end{aligned}$$

Transport compatibility, Eq. (7), gives

$$\sum_{z\in\mathcal{X}} K_x(z) C_t^{x\to y}(z' \mid z) = K_y(z').$$

Therefore,

$$\begin{aligned}\mathbb{E}\left[\frac{h(Z')}{h(Z)}\right] &= \frac{1}{K_x h} \sum_{z'\in\mathcal{X}} K_y(z') h(z') \\ &= \frac{K_y h}{K_x h}.\end{aligned}$$

Finally, since $\varphi_t(x) = K_x h$ and $\varphi_t(y) = K_y h$,

$$\frac{K_y h}{K_x h} = \frac{\varphi_t(y)}{\varphi_t(x)}.$$

This proves the result. □

A.2 PROOF OF THE FIBERWISE RAO–BLACKWELL IDENTITY

Proof of Theorem 4.2. Throughout the proof, all sums over a are restricted to contexts satisfying $\kappa(a) > 0$. The key observation is that the factor $H_p(a)$ in the controlled context distribution cancels the denominator of the fiberwise target, leaving a sum that reconstructs $K_y h$. Using the definitions of $\eta_{t,x}^h$ in Eq. (12) and $R_t^{x\to y}$ in Eq. (11),

$$\begin{aligned}\mathbb{E}_{A\sim\eta_{t,x}^h}[R_t^{x\to y}(A)] &= \sum_a \frac{\kappa(a) H_p(a)}{K_x h} \frac{H_q(a)}{H_p(a)} \\ &= \frac{1}{K_x h} \sum_a \kappa(a) H_q(a).\end{aligned}$$

By the definition of H_q,

$$\begin{aligned}\sum_a \kappa(a) H_q(a) &= \sum_a \kappa(a) \sum_{j=1}^N q_j(a) h(a,j) \\ &= \sum_a \sum_{j=1}^N K_y(a,j) h(a,j) \\ &= K_y h,\end{aligned}$$

where we used $\kappa(a)q_j(a) = K_y(a, j)$. Hence

$$\mathbb{E}_{A\sim\eta^h_{t,x}}\left[R_t^{x\to y}(A)\right] = \frac{K_y h}{K_x h} = \frac{\varphi_t(y)}{\varphi_t(x)}.$$

□

A.3 Variance Reduction by Nested Rao–Blackwellization

From Eq. (14),

$$T_C(A, J) = \mathbb{E}\left[\widehat{R}_C(A, J, J') \mid A, J\right],$$

and

$$R_t^{x\to y}(A) = \mathbb{E}\left[T_C(A, J) \mid A\right].$$

Since $\mathcal{X}$ is finite and $h > 0$, the random variables in Eq. (14) are square-integrable. Therefore, for any square-integrable random variable X and σ-algebra $\mathcal{G}$, the law of total variance gives

$$\mathrm{Var}\left(\mathbb{E}[X \mid \mathcal{G}]\right) \le \mathrm{Var}(X).$$

Applying this first with $X = \widehat{R}_C(A, J, J')$ and $\mathcal{G} = \sigma(A, J)$ gives

$$\mathrm{Var}[T_C(A, J)] \le \mathrm{Var}\left[\widehat{R}_C(A, J, J')\right].$$

Applying it again with $X = T_C(A, J)$ and $\mathcal{G} = \sigma(A)$ gives

$$\mathrm{Var}[R_t^{x\to y}(A)] \le \mathrm{Var}[T_C(A, J)].$$

Combining the two inequalities yields Eq. (15).

B Experimental Details

B.1 Target Energies and Local Ratios

Let $\Lambda = \{1, \dots, L\}^2$ denote the periodic square lattice and let $\mathcal{E}$ be its set of undirected nearest-neighbor edges, with each edge counted once. For the zero-field Ising model, we use

$$E_{\mathrm{Ising}}(x) = -\sum_{\{i,j\}\in\mathcal{E}} x_i x_j, \qquad x_i \in \{-1, +1\}.$$

For the N-state Potts model, we use

$$E_{\mathrm{Potts}}(x) = -\sum_{\{i,j\}\in\mathcal{E}} \mathbf{1}\{x_i = x_j\}, \qquad x_i \in [N].$$

For $i \in \Lambda$, let $x^{i\leftarrow n}$ denote the configuration obtained by replacing only x_i with n, where $n \in \{-1, +1\}$ for Ising and $n \in [N]$ for Potts. Consequently, replacing the symbol at site i by n yields the target ratios

$$\frac{\nu(x^{i\leftarrow n})}{\nu(x)} = \begin{cases} \exp\left(\beta(n - x_i)\sum_{j:j\sim i} x_j\right), & \text{Ising}, \\ \exp\left(\beta\sum_{j:j\sim i}\left(\mathbf{1}\{x_j = n\} - \mathbf{1}\{x_j = x_i\}\right)\right), & \text{Potts}. \end{cases}$$

These local ratios provide the energy factors used in the practical TAM target in Eq. (16).

B.2 Implementation Details

Model backbone. Following DASBS (Guo et al., 2026), we use the same transformer backbone with two-dimensional rotary positional embeddings for all Ising and four-state Potts experiments. Both networks contain six transformer blocks with hidden dimension 32, four attention heads, and an MLP ratio of 4, giving a feed-forward hidden dimension of 128 with GELU activations. The controller additionally receives the continuous time variable through adaptive normalization, whereas the corrector is time-independent. The resulting models contain approximately 144k controller parameters and 95k corrector parameters.

Table 6: Per-temperature training hyperparameters used for TAM. The effective batch size is $b \times$ #GPU, and $B_{\max}$ denotes the maximum replay-buffer capacity on each GPU.

Setting	#GPU	b	K_{time}	$B_{\max}$	T_{ctrl}	T_{corr}	T_r	L_τ	lr
Ising $\beta = 0.28$	2	64	8	512	400	200	20	200	$1\text{e}-3$
Ising $\beta = 0.4407$	2	64	8	512	400	200	20	200	$5\text{e}-4$
Ising $\beta = 0.6$	2	64	8	512	500	250	20	200	$5\text{e}-4$
Potts $\beta = 0.9$	2	128	16	512	500	250	20	100	$1\text{e}-3$
Potts $\beta = 1.0986$	2	128	16	4096	200	100	10	100	$5\text{e}-4$
Potts $\beta = 1.3$	2	128	16	4096	500	250	20	100	$1\text{e}-3$

Table 7: Training cost of TAM across the six benchmark settings. Steps denote the total number of controller and corrector gradient updates over five alternating stages. Runtime is wall-clock training time using two RTX 4080 GPUs.

Setting	Steps ($\times 10^3$)	Runtime (h)
Ising, $\beta = 0.28$	3.0	0.2
Ising, $\beta = 0.4407$	3.0	0.2
Ising, $\beta = 0.6$	3.75	0.3
Potts, $\beta = 0.9$	3.75	0.2
Potts, $\beta = 1.0986$	1.5	0.1
Potts, $\beta = 1.3$	3.75	0.3

Initialization. During the first controller stage, the corrector factor in the regression target is fixed to one, corresponding to a zero log-ratio. The corrector is subsequently trained normally, and its learned EMA parameters are used in later controller stages.

Training. We use AdamW with $\beta_1 = 0.9$, $\beta_2 = 0.999$, and zero weight decay, together with a constant learning rate, generalized KL divergence as the Bregman divergence, and uniform time weighting. Separate exponential moving averages are maintained for the controller and corrector. For the n-th update of either EMA, the decay is

$$\delta_n = \min\left(0.9999, \frac{1+n}{10+n}\right),$$

where n denotes the update count of the corresponding EMA. The EMA states persist across alternating stages. The reference process uses the modified log-linear noise schedule $\gamma_t = \gamma/(t+\alpha)$ with $\alpha = 0.5$. We use $\gamma = 1$ for the high- and critical-temperature settings and $\gamma = 0.25$ for the low-temperature Ising and Potts settings. Training proceeds for five alternating stages, each consisting of T_{ctrl} controller updates followed by T_{corr} corrector updates. At the beginning of each controller or corrector block, an empty per-GPU replay buffer is initialized and populated with b endpoint pairs generated by the controller EMA using τ-leaping. Every T_r local updates within the block, a fresh batch of b endpoint pairs is appended; if the buffer exceeds the maximum capacity $B_{\max}$, the oldest entries are discarded. Each gradient update samples without replacement a mini-batch of size b from the local buffer. For time-dependent objectives, we independently draw K_{time} time points from $\text{Unif}[0,1]$ for every endpoint pair. All TAM models are trained using two RTX 4080 GPUs. The controller is trained with the fiberwise target in Eq. (16), while the corrector follows the corresponding DASBS corrector-matching objective. A detailed implementation-level training procedure for the six standard benchmark experiments in Tables 2 and 3 is given in Algorithm 2. Per-task hyperparameters and training costs are summarized in Tables 6 and 7.

Endpoint-pair generation, bridge-state sampling, and corrector updates follow the DASBS training procedure; TAM modifies the controller update through the fiberwise target in Eq. (16).

Reference bridge sampling. Writing $p^r_{t|s}(x' \mid x)$ for the reference transition kernel from time s to t, the intermediate state conditioned on an endpoint pair (X_0, X_1) is sampled from the reference

Algorithm 2 Detailed training procedure for Transport Adjoint Matching

Require: Energy E, inverse temperature β, source distribution μ, reference dynamics r_t, controller g_θ, corrector $\widehat{c}_\psi$, number of alternating stages K_{stage}, update counts T_{ctrl} and T_{corr}, maximum replay-buffer capacity $B_{\max}$, refresh interval T_r, number of time samples K_{time}, per-GPU mini-batch size b, and τ-leaping steps L_τ.

Initialize controller parameters θ and corrector parameters ψ.
Initialize separate controller and corrector EMA states, which persist across all alternating stages.
for stage $k = 1, \dots, K_{\text{stage}}$ **do**
  Initialize an empty controller replay buffer $\mathcal{B}_{\text{ctrl}}$.
  On each GPU, generate b endpoint pairs (X_0, X_1) by sampling $X_0 \sim \mu$ and simulating the controller-EMA-controlled process using τ-leaping with L_τ steps; append them to $\mathcal{B}_{\text{ctrl}}$.
  for controller update $s = 1, \dots, T_{\text{ctrl}}$ **do**
    if $s > 1$ and $(s-1) \bmod T_r = 0$ **then**
      On each GPU, generate a fresh batch of b endpoint pairs from $X_0 \sim \mu$ using the controller-EMA-controlled process and append them to $\mathcal{B}_{\text{ctrl}}$.
      If $|\mathcal{B}_{\text{ctrl}}| > B_{\max}$, discard the oldest endpoint pairs until $|\mathcal{B}_{\text{ctrl}}| \le B_{\max}$.
    end if
    On each GPU, sample without replacement a mini-batch of b endpoint pairs from the local buffer $\mathcal{B}_{\text{ctrl}}$ and independently draw K_{time} time points $t \sim \text{Unif}[0,1]$ for each endpoint pair.
    Sample the corresponding bridge states X_t from the reference bridge in Eq. (19).
    For each admissible coordinate-local transition $X_t \to y$ with $r_t(y, X_t) > 0$ that updates coordinate d, set the fiber context $A := X_{1,-d}$ and anchor $j_0 := X_1^d$ from the sampled terminal endpoint X_1, and construct the local endpoint weights $p_j(A)$ and $q_j(A)$.
    Evaluate the local energy differences relative to j_0. During the first controller stage ($k = 1$), set the corrector factor to $\widehat{c} = 1$; otherwise evaluate the corrector ratios relative to j_0. Construct $\widehat{R}^{X_t \to y}_{\psi,t}(A)$ using Eq. (16).
    Update θ by minimizing Eq. (18), with the TAM target stop-gradiented.
    Update the controller EMA.
  end for
  Copy the controller EMA parameters to the controller network.
  Initialize an empty corrector replay buffer $\mathcal{B}_{\text{corr}}$.
  On each GPU, generate b endpoint pairs (X_0, X_1) by sampling $X_0 \sim \mu$ and simulating the controller-EMA-controlled process using τ-leaping with L_τ steps; append them to $\mathcal{B}_{\text{corr}}$.
  for corrector update $s = 1, \dots, T_{\text{corr}}$ **do**
    if $s > 1$ and $(s-1) \bmod T_r = 0$ **then**
      On each GPU, generate a fresh batch of b endpoint pairs from $X_0 \sim \mu$ using the controller-EMA-controlled process and append them to $\mathcal{B}_{\text{corr}}$.
      If $|\mathcal{B}_{\text{corr}}| > B_{\max}$, discard the oldest endpoint pairs until $|\mathcal{B}_{\text{corr}}| \le B_{\max}$.
    end if
    On each GPU, sample without replacement a mini-batch of b endpoint pairs from the local buffer $\mathcal{B}_{\text{corr}}$.
    For the denoising-matching corrector objective, additionally draw the required times and reference bridge states following DASBS.
    Update ψ using the task-specific DASBS corrector objective.
    Update the corrector EMA.
  end for
  Copy the corrector EMA parameters to the corrector network.
end for

Ensure: Controlled transition rates $u_\theta(t, y, x) = r_t(y, x) g_\theta(t, y, x)$.

bridge

$$p^r_{t|0,1}(x \mid x_0, x_1) = \frac{p^r_{t|0}(x \mid x_0)\, p^r_{1|t}(x_1 \mid x)}{p^r_{1|0}(x_1 \mid x_0)}. \tag{19}$$

For the coordinate-factorized references used in our experiments, this conditional distribution factorizes across coordinates. The non-additive reference in Section B.3 uses the same bridge construction with the transition kernel in Eq. (23); no cyclic-shift construction is required for bridge sampling.

Source distribution and corrector objective. For the high-temperature tasks, we use the product-uniform source distribution, sampling each coordinate independently and uniformly over its symbols, and train the corrector with the adjoint-matching corrector objective. For the critical- and low-temperature tasks, we use the corresponding zero-temperature source distribution and the denoising-matching corrector objective, following DASBS (Guo et al., 2026). Specifically, for Ising, μ is uniform over the two homogeneous configurations with all spins equal to -1 or $+1$; for the N-state Potts model, μ is uniform over the N homogeneous configurations with all sites taking the same symbol.

Neural sampling. Samples from the learned controlled process are generated using τ-leaping with transition rates $u_\theta(t, y, x) = r_t(y, x)g_\theta(t, y, x)$. We use $L_\tau = 200$ steps for the Ising models and $L_\tau = 100$ steps for the Potts models, as summarized in Table 6.

Ground truth and evaluation. Following the DASBS evaluation protocol (Guo et al., 2026), we evaluate each setting using 4096 samples and compare the resulting statistics with a reference pool of 4096 samples generated by independent long-run Swendsen–Wang chains. For Ising, the reference sampler uses batch size 128, 2^{13} burn-in updates, and collects one batch every 128 updates for 32 collection rounds. For Potts, we use batch size 128, 2^{16} burn-in updates, with the same collection interval and number of rounds. We report magnetization error $\Delta\mathrm{Mag.}$, two-point correlation error $\Delta\mathrm{Corr.}$, and energy Wasserstein-2 distance EW_2, with lower values indicating better agreement with the target distribution.

Rao–Blackwell variance experiment. For Fig. 2, we evaluate the nested targets at bridge times $t \in \{0.05, 0.15, 0.30, 0.50, 0.70, 0.85, 0.95\}$. We use the independent compatible local coupling $C_a(j' \mid j) = q_{j'}(a)$, which satisfies $p(a)C_a = q(a)$. At each bridge time, the variances of $\widehat{R}_C$, T_C, and $R_t^{x\to y}$ are computed under the same endpoint contexts and normalized by $\mathrm{Var}[\widehat{R}_C]$.

Controller-target ablation. For the controller-target ablation in Table 4, the sampled-transport and transport-marginalized variants use the same independent compatible local coupling $C_a(j' \mid j) = q_{j'}(a)$ as in the Rao–Blackwell variance experiment. The sampled variant draws one transported symbol per admissible local transition, whereas the transport-marginalized variant analytically sums over that symbol. The fiberwise TAM variant additionally marginalizes the observed local endpoint symbol. All three variants use the same practical energy-and-corrector approximation of the local terminal-potential ratios; only the degree of marginalization differs. All remaining training and evaluation settings are held fixed across the three regression targets.

B.3 Non-additive Reference Dynamics

For the non-additive reference experiment in Section 5.5, we identify the four Potts states with the cyclic group $\mathbb{Z}_4 = \{0, 1, 2, 3\}$, with addition understood modulo 4. Let $\rho = (\rho_0, \rho_1, \rho_2, \rho_3)$ be a strictly positive probability vector. We define the single-coordinate refresh generator by

$$Q_\rho(i, j) = \begin{cases} \rho_j, & j \neq i, \\ -(1 - \rho_i), & j = i. \end{cases} \tag{20}$$

With the time-dependent reference-noise rate γ_t, an off-diagonal transition from symbol i to j therefore occurs at rate $\gamma_t\rho_j$. Let

$$\Lambda_{s,t} := \int_s^t \gamma_\tau \, d\tau, \qquad \xi_{s,t} := e^{-\Lambda_{s,t}}. \tag{21}$$

Let $P_\rho := \mathbf{1}\rho^\top$. Since $\rho^\top\mathbf{1} = 1$, we have $P_\rho^2 = P_\rho$. Hence $Q_\rho = P_\rho - I$ has eigenvalues 0 and -1, with the latter having multiplicity three, and

$$e^{\Lambda_{s,t}Q_\rho} = P_\rho + e^{-\Lambda_{s,t}}(I - P_\rho). \tag{22}$$

Therefore, the corresponding single-coordinate transition kernel is

$$k^{\rho}_{s,t}(j \mid i) = \xi_{s,t}\mathbf{1}\{j = i\} + (1 - \xi_{s,t})\rho_j. \tag{23}$$

The full reference endpoint kernel factorizes across coordinates as $K^{\rho}_{s,t}(z \mid x) = \prod_{d=1}^{D} k^{\rho}_{s,t}(z_d \mid x_d)$, and therefore satisfies the shared-fiber marginal condition in Eq. (10).

For the standard additive reference, $\rho = (0.25, 0.25, 0.25, 0.25)$, so Eq. (23) depends only on $j - i \pmod 4$. Consequently, the DASBS cyclic translation maps the endpoint kernel associated with i exactly to that associated with any neighboring symbol i'. In the non-additive experiment, we instead fix

$$\rho = (0.1, 0.2, 0.3, 0.4). \tag{24}$$

Both choices of ρ yield the same nonzero spectrum $\{-1, -1, -1\}$ for Q_ρ, so the single-coordinate references share the same spectral relaxation timescale; the experimental change is the coordinate refresh distribution ρ.

The background term in Eq. (23) now depends on the absolute endpoint symbol j. Under a cyclic translation $S_{i\to i'}(j) = j + i' - i \pmod 4$, these background weights are permuted and are no longer equal to ρ. Consequently, if $J \sim k^{\rho}_{s,t}(\cdot \mid i)$, the translated symbol $S_{i\to i'}(J)$ does not in general follow $k^{\rho}_{s,t}(\cdot \mid i')$. Thus, the deterministic cyclic-shift target construction used by DASBS is not applicable.

TAM remains tractable because the endpoint moments required by the fiberwise target can still be evaluated analytically. For a fixed fiber context a, define

$$H^{\rho}_i(a) := \sum_{j=0}^{3} k^{\rho}_{t,1}(j \mid i)h(a, j) = \xi_{t,1}h(a, i) + (1 - \xi_{t,1}) \sum_{j=0}^{3} \rho_j h(a, j). \tag{25}$$

For a coordinate-local transition $x \to y$ differing at coordinate d, the exact fiberwise target is therefore

$$R^{x\to y}_t(a) = \frac{H^{\rho}_{y_d}(a)}{H^{\rho}_{x_d}(a)}. \tag{26}$$

Replacing the local terminal-potential ratios by the practical energy and corrector factors in Eq. (16) gives the implemented TAM target. Its evaluation remains $O(N)$ and requires no explicit $N \times N$ transport matrix. When $\rho_j = 1/4$, the construction reduces exactly to the uniform additive-reference case used in the standard Potts benchmarks.

The experiment uses the critical four-state Potts setting $L = 16$, $\beta = 1.0986$, and the same architecture, optimization, noise schedule, controller–corrector training budget, neural sampling budget, and evaluation protocol as the corresponding standard benchmark. The two conditions differ only in the coordinate refresh distribution, from $\rho = (0.25, 0.25, 0.25, 0.25)$ to $\rho = (0.1, 0.2, 0.3, 0.4)$. Results are reported as the mean and sample standard deviation over three independent random seeds.

B.4 Multi-Seed Results

The main benchmark results in Tables 2 and 3 are reported over three random seeds. Table 8 provides the corresponding mean and standard deviation for the six benchmark settings.

Table 8: Mean ± standard deviation of TAM over three random seeds. Lower is better for all metrics.

Target	β	ΔMag.	ΔCorr.	EW_2
Ising	0.28	$(3.4 \pm 0.1) \times 10^{-3}$	$(1.0 \pm 0.1) \times 10^{-3}$	1.9 ± 0.1
Ising	0.4407	$(6.8 \pm 1.8) \times 10^{-3}$	$(1.8 \pm 0.7) \times 10^{-3}$	2.3 ± 0.1
Ising	0.6	$(1.5 \pm 0.01) \times 10^{-2}$	$(2.8 \pm 0.5) \times 10^{-4}$	1.6 ± 0.1
Potts	0.9	$(2.5 \pm 0.2) \times 10^{-3}$	$(7.5 \pm 0.2) \times 10^{-4}$	0.9 ± 0.01
Potts	1.0986	$(3.4 \pm 0.2) \times 10^{-3}$	$(1.6 \pm 0.4) \times 10^{-3}$	1.0 ± 0.04
Potts	1.3	$(9.9 \pm 0.5) \times 10^{-4}$	$(2.5 \pm 1.4) \times 10^{-4}$	0.4 ± 0.03

C Additional Derivations

C.1 Anchor Invariance of the Practical TAM Target

Assume that the corrector ratio is exact, so that $\widehat{c}_\psi((a, j_0), (a, j)) = \widehat{\varphi}_1(a, j_0)/\widehat{\varphi}_1(a, j)$. By the terminal boundary relation in Eq. (5), the local factor in Eq. (16) satisfies

$$\widehat{\gamma}_{\psi,j}(a; j_0) = e^{-\beta \Delta E_j(a;j_0)} \frac{\widehat{\varphi}_1(a, j_0)}{\widehat{\varphi}_1(a, j)} = \frac{h(a, j)}{h(a, j_0)}.$$

Substituting this identity into the practical target gives

$$\widehat{R}^{x\to y}_{\psi,t}(a) = \frac{\sum_{j=1}^N q_j(a)h(a, j)/h(a, j_0)}{\sum_{j=1}^N p_j(a)h(a, j)/h(a, j_0)} = \frac{\sum_{j=1}^N q_j(a)h(a, j)}{\sum_{j=1}^N p_j(a)h(a, j)} = \frac{H_q(a)}{H_p(a)} = R^{x\to y}_t(a).$$

The common factor $h(a, j_0)^{-1}$ therefore cancels exactly, showing that the TAM target is independent of the anchor j_0 when the corrector ratios are exact.

C.2 Population Minimizer of the TAM Regression Objective

Let $R = R^{x\to y}_t(A)$ and denote $m := \mathbb{E}[R]$. For any $g > 0$, the Bregman divergence satisfies

$$\mathbb{E}[D_f(R\|g)] = \mathbb{E}[D_f(R\|m)] + D_f(m\|g).$$

Since f is strictly convex, $D_f(m\|g) \geq 0$ with equality if and only if $g = m$. Hence the unique population minimizer is

$$g^\star = \mathbb{E}[R].$$

Applying Theorem 4.2 yields

$$g^\star = \mathbb{E}_{A\sim\eta^h_{t,x}}[R^{x\to y}_t(A)] = \frac{\varphi_t(y)}{\varphi_t(x)},$$

which establishes Eq. (17).

C.3 Factorized Endpoint Kernels

Suppose the reference endpoint kernel factorizes across coordinates as $K_{t,x}(z) = \prod_{\ell=1}^D k_{t,\ell}(z_\ell \mid x_\ell)$, and let x and y differ only in coordinate d. Writing $z = (a, j)$ with $a = z_{-d}$, the shared context marginal in Eq. (10) becomes

$$\kappa(a) = \prod_{\ell\neq d} k_{t,\ell}(a_\ell \mid x_\ell) = \prod_{\ell\neq d} k_{t,\ell}(a_\ell \mid y_\ell),$$

because $x_\ell = y_\ell$ for all $\ell \neq d$. Consequently,

$$p_j(a) = k_{t,d}(j \mid x_d), \qquad q_j(a) = k_{t,d}(j \mid y_d),$$

so the local endpoint weights are independent of the surrounding context. The exact fiberwise TAM target therefore reduces to

$$R^{x\to y}_t(a) = \frac{\sum_{j=1}^N k_{t,d}(j \mid y_d)h(a, j)}{\sum_{j=1}^N k_{t,d}(j \mid x_d)h(a, j)}.$$

Using the practical energy-and-corrector approximation in Eq. (16) gives

$$\widehat{R}^{x\to y}_{\psi,t}(a) = \frac{\sum_{j=1}^N k_{t,d}(j \mid y_d)\widehat{\gamma}_{\psi,j}(a; j_0)}{\sum_{j=1}^N k_{t,d}(j \mid x_d)\widehat{\gamma}_{\psi,j}(a; j_0)}.$$

Thus, for coordinate-factorized reference dynamics, evaluating the TAM target requires only the two N-dimensional local endpoint distributions associated with the updated coordinate, together with the corresponding local energy and corrector ratios. No transport kernel or $N \times N$ coupling matrix needs to be constructed.

C.4 Independence from the Compatible Transport

Let C_a be any local Markov kernel satisfying $p(a)C_a = q(a)$. For a fixed context a, define

$$T_{C_a}(a, j) := \sum_{j'=1}^{N} C_a(j' \mid j) \frac{h(a, j')}{h(a, j)}.$$

Under the tilted local distribution proportional to $p_j(a)h(a, j)$, the conditional expectation of $T_{C_a}(a, J)$ given the fixed context $A = a$ is

$$\frac{\sum_{j=1}^{N} p_j(a)h(a, j)T_{C_a}(a, j)}{\sum_{j=1}^{N} p_j(a)h(a, j)} = \frac{\sum_{j,j'} p_j(a)C_a(j' \mid j)h(a, j')}{H_p(a)}.$$

Using the local transport compatibility $\sum_j p_j(a)C_a(j' \mid j) = q_{j'}(a)$ gives

$$\frac{\sum_{j'=1}^{N} q_{j'}(a)h(a, j')}{H_p(a)} = \frac{H_q(a)}{H_p(a)} = R_t^{x \to y}(a).$$

Hence every compatible local transport produces the same Rao–Blackwellized target after marginalization. The explicit transport coupling therefore enters only the derivation of the adjoint identity; after Rao–Blackwellization, the fiberwise TAM objective depends only on the endpoint marginals $p(a)$ and $q(a)$.