Inference-Time Reward Alignment of Diffusion Models via Geometric Particle Expansion

Anonymous Authors[1]

Abstract

Diffusion models are powerful generative models that capture the complexity of images and the natural design of proteins, small molecules, and structures. Aligning pre-trained diffusion models with domain-specific objectives without fine-tuning, and using inference only, has recently gained attention as a way to prevent mode-seeking behavior and reward over-optimization. Inspired by recent developments in Sequential Monte Carlo, this paper presents a derivative-free alignment approach for diffusion model inference to provide post-training optimization for non-differentiable rewards. We propose a geometric expansion in Sequential Monte Carlo–based sampling that does not require reward differentiation at any stage of inference. Our approach generates additional particles via geometric noise based on the covariance of elite particles along dominant eigenvalue directions, followed by resampling at the same timestep to prevent weight degeneracy. We formally demonstrate that our method consistently generates designs with improved performance in maximizing non-differentiable reward functions across image and molecule generation tasks, while achieving comparable or superior target rewards for differentiable aesthetic-preference objectives without performing any differentiation.

1. Introduction

Diffusion models are a class of generative models trained via an approximation to the log-likelihood objective and have been applied to images, videos, text, and protein structures (Ho et al., 2020; Song et al., 2020). Most diffusion models are pre-trained on large datasets, and downstream objectives are typically addressed through fine-tuning or inference-time alignment.

While reward maximization in image synthesis and NLP typically relies on differentiable reward functions, many scientific domains, including physics-based discovery and protein design, involve non-differentiable rewards obtained through high-fidelity simulations (Salomon-Ferrer et al., 2013).

Soft-value–based decoding and SMC-based methods are originally aimed to solve conditioning, but are used for reward maximization, especially in non-differentiable reward cases in scientific domains. There are major differences between SMC and soft-value–based models. In SMC, resampling is performed across the full batch with global interactions, whereas in soft-value–based methods, sampling is performed within a single batch. The proposal distribution can be a pre-trained model or an arbitrary proposal distribution, which is common in SMC-based sampling.

While Sequential Monte Carlo (SMC) methods are not originally formulated for reward maximization, they are well suited for preserving fidelity to a pretrained generative model. Even with large batch size and resampling based on effective sample size does not increase reward maximization and protect diversity. A number of variants of Sequential Monte Carlo (SMC) have been proposed, including nested importance sampling SMC, which approximates the locally optimal proposal distribution. Li et al. (Li et al., 2024) leverage this formulation to perform soft-value–based importance-sampling optimization for image and molecular generation.

Our main contributions are:

1. A new sequential monte carlo sampling method based on geometric expansion. We extend SMC with a same-step geometric expansion mechanism to incorporate reward maximization while remaining close to the target distribution.

   Let $q(x)$ be the distribution that generates your candidate pool at some step (parents + `pCN` children, or whatever proposal you have at that moment). Let $r(x)$ be your reward.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

2. Background

2.1. DDPM Formulation

Diffusion models are likelihood-based generative models that transform data into noise via a fixed forward Markov chain and learn the reverse denoising dynamics with a neural network (Sohl-Dickstein et al., 2015; Ho et al., 2020). Given a data sample $x_0 \in \mathbb{R}^d$, DDPMs introduce latent variables $x_{1:T} = \{x_1, \dots, x_T\}$ over T diffusion steps and define the generative model

$$p_\theta(x_{0:T}) = p(x_T) \prod_{t=1}^{T} p_\theta(x_{t-1} \mid x_t), \qquad p(x_T) = \mathcal{N}(0, I), \tag{1}$$

where θ denotes the parameters of the neural network and I is an identity matrix.

Forward (noising) process. A fixed variance schedule $\{\beta_t\}_{t=1}^T$, with $\beta_t \in (0,1)$, defines the forward Markov transitions

$$q(x_t \mid x_{t-1}) = \mathcal{N}\Big(x_t; \sqrt{1-\beta_t}\, x_{t-1}, \beta_t I\Big), \tag{2}$$

where $q(\cdot)$ denotes the forward diffusion distribution.

Defining $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$, the marginal distribution of x_t given x_0 admits the closed form

$$x_t = \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1-\bar{\alpha}_t}\, \epsilon, \quad \epsilon \sim \mathcal{N}(0, I), \tag{3}$$

where $\epsilon \in \mathbb{R}^d$ is standard Gaussian noise.

Reverse (denoising) process. The reverse-time transitions are modeled as Gaussian conditionals

$$p_\theta(x_{t-1} \mid x_t) = \mathcal{N}\big(x_{t-1}; \mu_\theta(x_t, t), \sigma_t^2 I\big), \tag{4}$$

where $\mu_\theta(x_t, t)$ is a learned mean function and σ_t^2 denotes the variance at step t.

Using the noise-parameterization of (Ho et al., 2020), the reverse update can be written as

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}\Big(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\, \epsilon_\theta(x_t, t)\Big) + \sigma_t z, \quad z \sim \mathcal{N}(0, I), \tag{5}$$

where $\epsilon_\theta(x_t, t)$ is a neural network predicting the injected noise and $z \in \mathbb{R}^d$ is standard Gaussian noise.

Training objective. Optimizing the evidence lower bound yields the standard DDPM loss

$$\mathcal{L}_{\text{simple}} = \mathbb{E}_{x_0,\, \epsilon,\, t}\big[\|\epsilon - \epsilon_\theta(x_t, t)\|^2\big], \quad x_t = \sqrt{\bar{\alpha}_t}\, x_0 + \sqrt{1-\bar{\alpha}_t}\, \epsilon, \quad t \sim \mathrm{Unif}\{1, \dots, T\}, \tag{6}$$

where the expectation is taken over the data distribution, Gaussian noise ϵ, and uniformly sampled diffusion steps t. The diffusion formulation described above provides a generative mechanism for producing physically smooth and coherent hull-form variations within the prescribed parametric design space. However, in its standard form, the reverse diffusion process samples designs according to the learned data distribution without preferential emphasis on hydrodynamic performance. To transform the diffusion model from a passive generator into an active design optimization tool, a performance-aware guidance mechanism is required.

2.2. Monte-Carlo Sampling

Consider the optimization problem over distributions $p(x)$:

$$\max_p \;\; \mathbb{E}_p\big[r(x)\big] \;-\; \frac{1}{\lambda}\, \mathrm{KL}(p \,\|\, q)\,. \tag{$\star$}$$

or

$$\{p_t^*\}_t \;=\; \arg\max_{\{p_t\}_{t=T+1}^{1}} \Bigg\{ \mathbb{E}_{\{p_t\}}[r(x_0)] \;-\; \alpha \sum_{t=T+1}^{1} \mathbb{E}_{\{p_t\}}\Big[\mathrm{KL}(p_t(\cdot \mid x_t) \,\|\, p_t^{\text{pre}}(\cdot \mid x_t) \tag{7}$$

This objective says:

- maximize expected reward,
- but don't move "too far" from your proposal q (a trust-region in KL),
- $\lambda > 0$ controls how aggressive you are.

Fact (derivation sketch). The unique maximizer of ($\star$) is:

$$p_\lambda(x) = \frac{q(x)\exp\big(\lambda r(x)\big)}{Z_\lambda}, \qquad Z_\lambda = \int q(x)\exp\big(\lambda r(x)\big)\, dx. \tag{1}$$

That is exactly the exponential-tilt / Boltzmann / softmax distribution.

So when your code does

$$w_i \propto \exp(\lambda r_i)$$

and resamples from the candidate set using those weights, you are implementing a Monte Carlo approximation of the optimizer of ($\star$).

This gives you a strong, defensible claim:

Even though we aren't targeting a Bayesian π_t like SMC samplers do, our selection step is itself an optimal update in distribution space: it is the KL-regularized best-reward distribution closest to the proposal.

This is a standard lens in "policy improvement with KL trust regions" too: "new policy = old policy reweighted by $\exp(\text{advantage}/\alpha)$" (same math).

2. **A simple "effectiveness" proof: exponential tilting always increases expected reward (in expectation)**

   This is a very nice property you can actually prove cleanly.

   Let $X \sim q$. Define the tilted distribution p_λ as in (1). Then:

$$\mathbb{E}_{p_\lambda}\big[r(X)\big] = \frac{\mathbb{E}_q\big[r(X)e^{\lambda r(X)}\big]}{\mathbb{E}_q\big[e^{\lambda r(X)}\big]}. \tag{2}$$

   Now compare $\mathbb{E}_{p_\lambda}[r]$ to $\mathbb{E}_q[r]$. Note:

$$\mathbb{E}_{p_\lambda}[r] - \mathbb{E}_q[r] = \frac{\mathrm{Cov}_q\big(r(X),\, e^{\lambda r(X)}\big)}{\mathbb{E}_q\big[e^{\lambda r(X)}\big]}. \tag{3}$$

   So it's enough to show $\mathrm{Cov}_q\big(r, e^{\lambda r}\big) \geq 0$.

   Lemma 2.1. *For any random variable R and any increasing function g,*

$$\mathrm{Cov}\big(R, g(R)\big) \geq 0.$$

   Proof. Let R' be an independent copy of R. Then

$$\mathrm{Cov}\big(R, g(R)\big) = \frac{1}{2}\,\mathbb{E}\Big[(R - R')\big(g(R) - g(R')\big)\Big].$$

   If g is increasing, then $(R - R')\big(g(R) - g(R')\big) \geq 0$ always, so the expectation is ≥ 0. □

   Here $g(R) = e^{\lambda R}$ is increasing for $\lambda > 0$. Therefore:

$$\mathbb{E}_{p_\lambda}[r] \;\geq\; \mathbb{E}_q[r]. \tag{4}$$

   What this means for your algorithm. At each step, conditional on the candidate pool distribution q, your softmax resampling is guaranteed to bias toward higher-reward regions, and in the idealized infinite-sample limit it increases the expected reward of selected samples relative to the proposal.

   This is already a rigorous "it is an optimizer, not just random sampling" statement.

1. **The cleanest fact: for a Gaussian, covariance = inverse Hessian (exactly).**

   Let $X \sim (\mu, \Sigma)$. Its log-density is

$$\log p(x) = \text{const} - \frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu). \tag{8}$$

   Take the negative log-density:

$$-\log p(x) = \text{const} + \frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu). \tag{9}$$

   Differentiate twice:

$$_x\big(-\log p(x)\big) = \Sigma^{-1} \qquad \text{(constant in } x\text{)}. \tag{10}$$

   So for a Gaussian,

$$(-\log p) = \Sigma^{-1} \iff \Sigma = \Big((-\log p)\Big)^{-1}. \tag{11}$$

   That's the fundamental reason covariance "is" second-order structure: it is literally the inverse curvature (precision) for a Gaussian.

2. **For a general smooth density, it becomes true locally (Laplace / quadratic approximation).**

   Let a target density be

$$\pi(x) \propto \exp\big(\ell(x)\big), \qquad \text{where } \ell(x) = \log \pi(x) \text{ is smooth.} \tag{12}$$

   Assume there is a local mode x (a maximizer of ℓ), and ℓ is twice differentiable. Then $\nabla \ell(x^{)=0}$ and a second-order Taylor expansion around x gives

$$\ell(x) \approx \ell(x^{)+\frac{1}{2}(x-x^{)^\top \ell(x^{)}(x-x^{)}. \tag{13}$$

   At a (strict) local maximum, $\ell(x^{)}$ is negative definite. Define the local precision / curvature matrix

$$A := -\ell(x^{)\succ 0. \tag{14}$$

   Then

$$\ell(x) \approx \ell(x^{)-\frac{1}{2}(x-x^{)^\top A(x-x^{)}, \tag{15}$$

   so

$$\pi(x) \approx \text{const}\cdot\exp\left(-\frac{1}{2}(x-x^{)^\top A(x-x^{)}\right) \;=\; (x^{,A^{-1}). \tag{16}$$

   Hence, near the mode,

$$_\pi(X) \approx A^{-1} = \Big(-\log \pi(x^{)}\Big)^{-1}. \tag{17}$$

   Eigenvalue intuition (very useful). If $Av_i = \kappa_i v_i$, then the local Gaussian has variance along v_i

$$(v_i^\top X) \approx \frac{1}{\kappa_i}. \tag{18}$$

   So:

   large curvature $\kappa_i \;\Rightarrow$ small variance, small curvature $\kappa_i \;\Rightarrow$ large

   That is exactly what you want your noise/proposals to do: explore farther along flatter directions and less along steep ones.

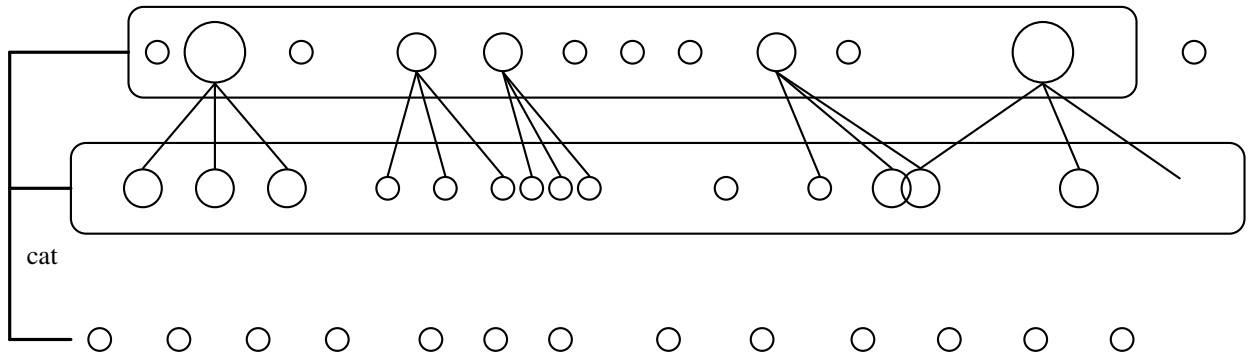

Figure 1. Illustration of particle expansion in the SMC .

3. **Apply it to your diffusion+reward form: the "Hessian" is in the log target, and covariance estimates its inverse.**

   In the DAS / Test-time Alignment paper you uploaded, the intermediate targets are of the form

   $$\pi_t(x_t) \propto p_t(x_t)\exp\big(\lambda_t \alpha\, \hat{r}(x_t)\big), \tag{19}$$

   (see Eq. (11) in `Test_time_AlignmentICLR2025`). Therefore,

   $$\log \pi_t(x) = \log p_t(x) + \lambda_t \alpha\, \hat{r}(x) + \text{const}. \tag{20}$$

   At a high-reward/high-probability point x, the local curvature (negative Hessian) is

   $$A_t := -\log \pi_t(x^{)} \tag{21}$$

   $$= -\log p_t(x^{)} - \lambda_t \alpha\, \hat{r}(x^{)\cdot} \tag{22}$$

   If $\hat{r}$ is concave near x, then $\hat{r}(x^{)\prec 0}$ and thus $-\hat{r}(x^{)\succ 0}$, contributing positive curvature. The Laplace logic says

   $$(X) \approx A_t^{-1}. \tag{23}$$

   So if you can estimate the local covariance of samples concentrated near x, you are implicitly estimating A_t^{-1} and therefore the second-order geometry A_t (up to inversion).

4. **Where your "covariance-from-parents" fits: it's a Hessian-free way to get curvature directions.**

   Your geometric expansion does something like:

   - keep elite parents (high reward),
   - compute empirical covariance $\widehat{\Sigma}$ of those parents,
   - generate anisotropic noise $\eta \sim (0, \widehat{\Sigma})$,
   - inject it via pCN.

   This is exactly the "local Gaussian" idea: elites $\approx$ samples from a locally Gaussian bump, so their empirical covariance $\widehat{\Sigma}$ estimates the bump covariance

   $$\Sigma^{\approx A^{-1}}. \tag{24}$$

   This same core idea appears in classical adaptive proposals: in SMC, "there is a sample from the target readily available" and it "can be used to compute posterior moments and inform the shape of the proposal kernel" (`adaptivesmc`), using proposals like

   $$q(\tilde{\theta} \mid \theta) = \big(\theta,\, h^2\, \widehat{\Sigma}_{\pi_t}\big), \tag{25}$$

   with $\widehat{\Sigma}_{\pi_t}$ an empirical covariance estimate of the target.

5. **Connection to the "locally optimal proposal" derivation in the DAS paper: they do 1st-order, covariance is a route to 2nd-order.**

   In Sec. 3.3.3 of the DAS paper, they derive a locally optimal proposal (for minimizing weight variance) of the form

   $$m_{t-1}^{(x_{t-1}|x_t)\propto\exp\Big(-\frac{1}{2\sigma_t^2}\|x_{t-1}-\mu_\theta(x_t,t)\|^2+\lambda_{t-1}\alpha\,\hat{r}(x_{t-1})\Big)}, \tag{26}$$

   (Eq. (12) in `Test_time_AlignmentICLR2025`). They then approximate this proposal with a Gaussian using a first-order Taylor expansion, yielding a mean shift proportional to $\nabla \hat{r}$ and keeping covariance $\sigma_t^2 I$ (Eq. (13)).

   What's missing from that first-order Gaussian approximation is the Hessian term of $\hat{r}$. If you did a second-order Taylor expansion of $\hat{r}$, you'd get an anisotropic covariance (not just $\sigma_t^2 I$) whose precision includes $-\hat{r}$. That anisotropic covariance is exactly what your "covariance-from-elites" is trying to estimate without explicitly computing Hessians.

6. **Important caveat: covariance is "Hessian-like" only under the right locality.**

   A Hessian is pointwise local: $\hat{r}(x^{)}$. A covariance is a distribution-level second moment.

   They match when your particles are concentrated in a region where the log target is close to quadratic.

   What breaks the link:

- multimodality inside your elite set (covariance becomes a mixture covariance, not a single-mode curvature),
- non-quadratic regions (heavy tails, skew),
- elites that are not actually sampling from π_t locally (e.g., selection too aggressive or too broad).

Your "top-k" elite choice is precisely a practical way to enforce locality so that "local Gaussian bump" is more believable.

Under a Laplace approximation, the local target around a mode x is approximately Gaussian with covariance $\Sigma = \left(-\log \pi(x)\right)^{-1}$. Therefore, the empirical covariance of high-weight or high-reward particles provides a Hessian-free estimate of the inverse curvature, which can be used to precondition proposal noise.

Near x_0, the reverse kernel is almost deterministic $\Rightarrow$ SMC can't explore.

A DDPM-style reverse step can be written (schematically) as

$$x_{t-1} = \mu_\theta(x_t, t, c) + \sigma_t\, \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0, I), \tag{27}$$

with $\sigma_t \downarrow 0$ as $t \to 0$.

So late in the chain, σ_t is tiny: propagation does not create meaningful diversity. If you only do SMC (weight + resample), resampling does not change the support—it just copies existing particles:

$$\{x_{t-1}^{(i)}\}_{i=1}^N \xrightarrow{\text{resample}} \text{a multiset of the same points.} \tag{28}$$

That's why "plain SMC" tends to plateau late: it can select, but it can't generate new high-reward variants once the model's own noise becomes small.

Your covariance-children step is effectively injecting an extra "mutation" kernel

$$x' = \rho x + \beta\, \delta, \tag{29}$$

with β not tied to σ_t. So it restores exploration precisely when the diffusion process stops exploring.

Sequential Monte Carlo (SMC) methods approximate a sequence of probability distributions

$$\pi_t(x) \;=\; \frac{\gamma_t(x)}{Z_t}, \tag{30}$$

where $\gamma_t(x)$ is an unnormalized density and $Z_t = \int \gamma_t(x)\, dx$ is an unknown normalizing constant. Each π_t is represented by a weighted empirical measure

$$\pi_t(x) \;\approx\; \sum_{i=1}^N w_t^{(i)}\, \delta\Big(x - x_t^{(i)}\Big), \qquad \sum_{i=1}^N w_t^{(i)} = 1. \tag{31}$$

Expectations under π_t are approximated as

$$\mathbb{E}_{\pi_t}[f(X)] \;\approx\; \sum_{i=1}^N w_t^{(i)} f\Big(x_t^{(i)}\Big). \tag{32}$$

—

2.3. Generic SMC Recursion

At each step t, SMC performs the following operations:

Propagation. Particles are propagated via a proposal kernel q_t,

$$x_t^{(i)} \sim q_t(\cdot \mid x_{t-1}^{(a^{(i)})}), \tag{33}$$

where $a^{(i)}$ denotes the ancestor index after resampling.

Weight update. Unnormalized importance weights are computed as

$$\tilde{w}_t^{(i)} = w_{t-1}^{(a^{(i)})} \frac{\gamma_t\Big(x_t^{(i)}\Big)}{\gamma_{t-1}\Big(x_{t-1}^{(a^{(i)})}\Big)\; q_t\Big(x_t^{(i)} \mid x_{t-1}^{(a^{(i)})}\Big)}, \tag{34}$$

followed by normalization

$$w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^N \tilde{w}_t^{(j)}}. \tag{35}$$

Resampling. To mitigate weight degeneracy, particles are resampled according to

$$a^{(1)}, \ldots, a^{(N)} \sim \text{Categorical}\Big(w_t^{(1)}, \ldots, w_t^{(N)}\Big). \tag{36}$$

The effective sample size (ESS) is defined as

$$\text{ESS}_t \;=\; \frac{1}{\sum_{i=1}^N \Big(w_t^{(i)}\Big)^2}. \tag{37}$$

—

2.4. SMC Sampler with Reward-Based Potentials

A common specialization of SMC, often referred to as an *SMC sampler*, defines a sequence of targets via a potential (or reward) function $r(x)$:

$$\pi_t(x) \;\propto\; p_\theta(x_t \mid \text{prompt})\, \exp\big(\lambda\, r(x)\big). \tag{38}$$

Particles are propagated through a Markov kernel M_t, and weighted using

$$w_t^{(i)} \;\propto\; \exp\big(\lambda\, r(x_t^{(i)})\big), \tag{39}$$

followed by multinomial resampling.

—

3. Propagation via Reverse Diffusion

In our setting, particles correspond to Stable Diffusion latent variables

$$x_t \in \mathbb{R}^D, \quad D = 4 \times 64 \times 64.$$

Propagation is performed using the guided reverse diffusion kernel

$$x_{t-1} \sim M_t(\cdot \mid x_t), \tag{40}$$

where classifier-free guidance combines conditional and unconditional noise estimates

$$\varepsilon = \varepsilon_{\text{uncond}} + s\,(\varepsilon_{\text{cond}} - \varepsilon_{\text{uncond}})\,, \tag{41}$$

with guidance scale $s > 0$. The diffusion scheduler then induces a Markov transition

$$x_{t-1} = \Phi_t(x_t, \varepsilon).$$

—

4. Reward-Based Selection

Decoded latents are mapped to images and scored via an aesthetic function $r(x)$. Scores are standardized,

$$\hat{r}_i = \frac{r(x_i) - \bar{r}}{\sigma_r + \varepsilon}, \tag{42}$$

and converted to Boltzmann weights

$$w_i = \frac{\exp(\lambda \hat{r}_i)}{\sum_j \exp(\lambda \hat{r}_j)}. \tag{43}$$

Particles are resampled according to these weights, yielding a selection step analogous to fitness-proportional selection.

—

5. Geometric Particle Expansion

To enhance exploration, additional candidate particles are generated via a geometric preconditioned Crank–Nicolson (pCN) proposal. At a given reverse timestep t, let the *parent* particle set be

$$X_{1:N} \subset \mathbb{R}^3, \qquad X_i \in \mathbb{R}^3,$$

with associated scalar rewards $\{r_t(X_i)\}_{i=1}^N$. Let $k \le N$ denote the fixed number of elite parents. Define the elite index set

$$\mathcal{E}_t \;=\; \mathrm{TopK}\big(\{r_t(X_i)\}_{i=1}^N,\, k\big) \;\subset\; \{1, \dots, N\}. \tag{44}$$

Elite geometry. The empirical mean and covariance of the elite parents are given by

$$\begin{aligned} \mu_t &= \frac{1}{k} \sum_{i \in \mathcal{E}_t} X_i, \\ \widehat{\Sigma}_t &= \frac{1}{k-1} \sum_{i \in \mathcal{E}_t} (X_i - \mu_t)(X_i - \mu_t)^\top \;\in\; \mathbb{R}^{3\times 3}. \end{aligned} \tag{45}$$

Let the eigendecomposition be

$$\widehat{\Sigma}_t = U_t \Lambda_t U_t^\top, \qquad \Lambda_t = \mathrm{diag}(\lambda_{t,1}, \lambda_{t,2}, \lambda_{t,3}), \tag{46}$$

and define a numerically stabilized square root

$$\widehat{\Sigma}_t^{1/2} = U_t\, \mathrm{diag}\left(\sqrt{\max(\lambda_{t,j}, \varepsilon)} \right) U_t^\top. \tag{47}$$

Time-dependent pCN parameters. Let T be the total number of reverse steps. We define

$$\beta_t = \beta_0 \sqrt{\frac{t+1}{T}}, \qquad \rho_t = \sqrt{1 - \beta_t^2}, \tag{48}$$

with β_t optionally clipped to ensure $\rho_t \in \mathbb{R}$.

Child generation. Let c be the number of children per elite parent, so that

$$M = k\,c, \qquad Y_{1:M} \subset \mathbb{R}^3. \tag{49}$$

Each child Y_j is associated with an elite parent $X_{p(j)}$, where $p(j) \in \mathcal{E}_t$ repeats each elite index exactly c times. We draw independent noise variables

$$\eta_j^{\text{geom}} \sim \mathcal{N}(0, \widehat{\Sigma}_t), \qquad \eta_j^{\text{iso}} \sim \mathcal{N}(0, I_3), \tag{50}$$

and form a convex combination

$$\eta_j = \eta_j^{\text{geom}} + \eta_j^{\text{iso}}. \tag{51}$$

Here, η_j^{geom} and η_j^{iso} denote the geometric and isotropic perturbation components, respectively. In implementation, their relative contribution is controlled by the hyperparameter $\alpha \in [0, 1]$. Each child particle is then generated via a pCN-style correlated update

$$Y_j = \rho_t\, X_{p(j)} + \beta_t\, \eta_j, \qquad j = 1, \dots, M. \tag{52}$$

Equivalent Gaussian form. Since η_j^{geom} and η_j^{iso} are independent Gaussians,

$$\eta_j \sim \mathcal{N}(0, \Sigma_{t,\text{eff}})\,, \qquad \Sigma_{t,\text{eff}} = (1-\alpha)^2 \widehat{\Sigma}_t + \alpha^2 I_3. \tag{53}$$

Consequently, the conditional proposal distribution of each child satisfies

$$Y_j \mid X_{p(j)} \;\sim\; \mathcal{N}\Big(\rho_t X_{p(j)},\; \beta_t^2 \Sigma_{t,\text{eff}}\Big). \tag{54}$$

5.1. Local Covariance Estimation

Let $\{x^{(j)}\}_{j=1}^{k}$ denote the top-k particles by reward at t. Flattening and centering yields

$$X = \begin{bmatrix} (x^{(1)} - \mu)^\top \\ \vdots \\ (x^{(k)} - \mu)^\top \end{bmatrix} \in \mathbb{R}^{k \times D}, \qquad \mu = \frac{1}{k} \sum_{j=1}^{k} x^{(j)}. \tag{55}$$

The empirical covariance admits the decomposition

$$\hat{\Sigma} = \frac{1}{k-1} X^\top X = V \Lambda V^\top, \tag{56}$$

obtained via the SVD $X = USV^\top$, with $\Lambda = S^2/(k-1)$.

—

5.2. Low-Rank pCN Proposal

A geometric pCN proposal is defined as

$$x' = \rho_t x + \beta_t \eta, \qquad \rho_t = \sqrt{1 - \beta_t^2}, \tag{57}$$

where

$$\eta = (1 - \alpha)\, \eta_{\text{geom}} + \alpha\, \eta_{\text{iso}}, \tag{58}$$

with

$$\eta_{\text{geom}} \sim \mathcal{N}(0, \hat{\Sigma}), \qquad \eta_{\text{iso}} \sim \mathcal{N}(0, I).$$

The step size β_t is annealed over diffusion time, ensuring stability and controlled exploration.

—

6. Covariance and the Inverse Hessian

6.1. Exact Identity for Gaussian Densities

Let $x \in \mathbb{R}^d$. A Gaussian density is

$$p(x) = \mathcal{N}(x; \mu, \Sigma) \;\propto\; \exp\left(-\frac{1}{2} (x - \mu)^\top \Sigma^{-1} (x - \mu) \right). \tag{59}$$

Define the negative log-density (up to an additive constant) as

$$U(x) \;:=\; -\log p(x) = \frac{1}{2} (x - \mu)^\top \Sigma^{-1} (x - \mu) + \text{const}. \tag{60}$$

The gradient of U is

$$\nabla U(x) = \Sigma^{-1} (x - \mu), \tag{61}$$

and the Hessian (matrix of second derivatives) is

$$\nabla^2 U(x) = \Sigma^{-1}. \tag{62}$$

Consequently, for a Gaussian distribution,

$$\Sigma = \big(\nabla^2 (-\log p(x)) \big)^{-1}, \tag{63}$$

and this identity is exact. In particular, it does not depend on x, since the curvature of the Gaussian log-density is constant. Equivalently, the Hessian of the log-density satisfies

$$\nabla^2 \log p(x) = -\Sigma^{-1}. \tag{64}$$

6.2. Laplace Approximation for General Densities

Consider a general density of the form

$$p(x) \propto \exp\big(-U(x)\big), \tag{65}$$

where $U(x)$ denotes the negative log-density up to a constant.

Let $x^\star$ be a mode (MAP point), so that

$$\nabla U(x^\star) = 0. \tag{66}$$

A second-order Taylor expansion of U around $x^\star$ yields

$$U(x) \approx U(x^\star) + \frac{1}{2} (x - x^\star)^\top H (x - x^\star), \qquad H := \nabla^2 U(x^\star). \tag{67}$$

Substituting into the density gives

$$p(x) \propto e^{-U(x)} \approx \exp\big(-U(x^\star)\big) \exp\left(-\frac{1}{2} (x - x^\star)^\top H (x - x^\star) \right). \tag{68}$$

The second factor is a Gaussian kernel with precision matrix H. Hence,

$$p(x) \approx \mathcal{N}(x^\star, H^{-1}) \qquad \Longrightarrow \qquad \text{Cov}[x] \approx H^{-1}. \tag{69}$$

This approximation is local to the mode and becomes accurate when the density is sharply peaked and well approximated by a quadratic potential.

7. Methodology: Setup (one guided step proof).

At one guided step, we pool *parents* $X_{1:N} \subset \mathbb{R}^3$ and *children* $Y_{1:M} \subset \mathbb{R}^3$. For $N \geq k$, we select the top-k parents and generate c children per parent, so that $M = kc$ and M is independent of N. Define the Boltzmann weight

$$w(z) := \exp\big(\beta\, r(z)\big), \qquad \beta > 0, \tag{70}$$

and the total masses

$$S_N := \sum_{i=1}^{N} w(X_i), \qquad C_N := \sum_{j=1}^{M} w(Y_j). \tag{71}$$

Let δ_x [1] denote the Dirac probability measure at x. Define the parent-only weighted empirical measure

$$\eta_N^{\mathrm{par}} := \sum_{i=1}^{N} \frac{w(X_i)}{S_N}\, \delta_{X_i}, \tag{72}$$

and the pooled parent–child weighted empirical measure

$$\eta_N^{\mathrm{all}} := \sum_{i=1}^{N} \frac{w(X_i)}{S_N + C_N}\, \delta_{X_i} \;+\; \sum_{j=1}^{M} \frac{w(Y_j)}{S_N + C_N}\, \delta_{Y_j}. \tag{73}$$

Here, η_N^{all} is a discrete probability measure supported on the finite set $\{X_1, \dots, X_N, Y_1, \dots, Y_M\}$, assigning mass proportional to the particle weights. Conditional on the current particle population, multinomial resampling produces independent draws from the discrete distribution η_N^{all}.

Lemma 1 (deterministic bound). For any bounded measurable φ with $\|\varphi\|_\infty := \sup_z |\varphi(z)| < \infty$,

$$\left|\eta_N^{\mathrm{all}}(\varphi) - \eta_N^{\mathrm{par}}(\varphi)\right| \le 2\|\varphi\|_\infty \frac{C_N}{S_N}. \tag{74}$$

Proof. Write the expectations as fractions:

$$\eta_N^{\mathrm{all}}(\varphi) = \frac{\sum_{i=1}^N w(X_i)\varphi(X_i) + \sum_{j=1}^M w(Y_j)\varphi(Y_j)}{S_N + C_N}, \tag{75}$$

$$\eta_N^{\mathrm{par}}(\varphi) = \frac{\sum_{i=1}^N w(X_i)\varphi(X_i)}{S_N}. \tag{76}$$

Subtract and rearrange:

$$\eta_N^{\mathrm{all}}(\varphi) - \eta_N^{\mathrm{par}}(\varphi) = \frac{\sum_{j=1}^M w(Y_j)\varphi(Y_j)}{S_N + C_N} - \frac{C_N}{S_N + C_N} \cdot \frac{\sum_{i=1}^N w(X_i)\varphi(X_i)}{S_N}. \tag{77}$$

Using $|\varphi| \le \|\varphi\|_\infty$ gives

$$\left|\frac{\sum_{j=1}^M w(Y_j)\varphi(Y_j)}{S_N + C_N}\right| \le \|\varphi\|_\infty \frac{C_N}{S_N + C_N}, \tag{78}$$

and similarly

$$\left|\frac{C_N}{S_N + C_N} \cdot \frac{\sum_{i=1}^N w(X_i)\varphi(X_i)}{S_N}\right| \le \|\varphi\|_\infty \frac{C_N}{S_N + C_N}. \tag{79}$$

Therefore,

$$\left|\eta_N^{\mathrm{all}}(\varphi) - \eta_N^{\mathrm{par}}(\varphi)\right| \le 2\|\varphi\|_\infty \frac{C_N}{S_N + C_N} \le 2\|\varphi\|_\infty \frac{C_N}{S_N}. \tag{80}$$

□

[1] The Dirac measure δ_x satisfies $\int f(z)\, \delta_x(dz) = f(x)$ for any bounded measurable test function f.

Corollary (TV bound). Since $\|\nu\|_{\mathrm{TV}} = \sup_{\|\varphi\|_\infty \le 1} |\nu(\varphi)|$,

$$\|\eta_N^{\mathrm{all}} - \eta_N^{\mathrm{par}}\|_{\mathrm{TV}} \le 2\, \frac{C_N}{S_N + C_N} \le 2\, \frac{C_N}{S_N}. \tag{81}$$

Lemma 2 (probabilistic bound for fixed M). Assume:

$$0 < \mu := \mathbb{E}[w(X_1)] < \infty, \tag{82}$$

$$\mathbb{E}[w(Y_j) \mid X_{1:N}] \le K \quad \text{a.s., for all } j \le M, \tag{83}$$

for some finite K, with M fixed as in (**??**). Then

$$\frac{C_N}{S_N} \xrightarrow[N\to\infty]{\mathbb{P}} 0, \qquad \frac{C_N}{S_N} = O_{\mathbb{P}}(1/N). \tag{84}$$

Proof. By WLLN,

$$\frac{S_N}{N} = \frac{1}{N}\sum_{i=1}^{N} w(X_i) \xrightarrow{\mathbb{P}} \mu, \tag{85}$$

hence for any $\delta \in (0, 1)$,

$$\mathbb{P}(S_N \ge (1-\delta)\mu N) \to 1. \tag{86}$$

Moreover, by conditional expectation and (83),

$$\mathbb{E}[C_N \mid X_{1:N}] = \sum_{j=1}^{M} \mathbb{E}[w(Y_j) \mid X_{1:N}] \le MK, \tag{87}$$

so $C_N = O_{\mathbb{P}}(1)$. On the event in (86),

$$\frac{C_N}{S_N} \le \frac{C_N}{(1-\delta)\mu N} = O_{\mathbb{P}}(1/N), \tag{88}$$

which proves (84). □

Theorem (children are asymptotically negligible). Under (82)–(83) and fixed M,

$$\|\eta_N^{\mathrm{all}} - \eta_N^{\mathrm{par}}\|_{\mathrm{TV}} \xrightarrow[N\to\infty]{\mathbb{P}} 0, \qquad \|\eta_N^{\mathrm{all}} - \eta_N^{\mathrm{par}}\|_{\mathrm{TV}} = O_{\mathbb{P}}(1/N). \tag{89}$$

Corollary (selected children count). Let K_N be the number of children selected after N multinomial draws from η_N^{all}. Conditional on the pool,

$$K_N \mid \{X_{1:N}, Y_{1:M}\} \sim \mathrm{Binomial}(N, p_N), \qquad p_N := \frac{C_N}{S_N + C_N}. \tag{90}$$

Since $p_N \le C_N/S_N = O_{\mathbb{P}}(1/N)$ by (84),

$$K_N = O_{\mathbb{P}}(1), \qquad \frac{K_N}{N} \xrightarrow{\mathbb{P}} 0. \tag{91}$$

7.1. Temperature Scaling

If the density has the form

$$p(x) \propto \exp\bigl(-\beta U(x)\bigr), \tag{92}$$

then the local Hessian is βH, and the corresponding covariance scales as

$$\mathrm{Cov}[x] \approx (\beta H)^{-1}. \tag{93}$$

The identity "covariance = inverse Hessian" is exact only for Gaussian distributions and holds locally under the Laplace approximation. In diffusion models, the intermediate distributions p_t can be highly non-Gaussian and multimodal. As a result, the empirical covariance $\widehat{\Sigma}_t$ should be interpreted as a heuristic proxy for local curvature or spread in the sample population, rather than as a principled or guaranteed substitute for the true Hessian of $-\log p_t$.

If, locally in time, the target distribution at diffusion/SMC step t behaves approximately as a Gaussian,

$$p_t(x) \approx \mathcal{N}(\mu, \Sigma^\star), \tag{94}$$

then the local curvature of the log-density satisfies

$$-\nabla_x^2 \log p_t(x) \approx (\Sigma^\star)^{-1}. \tag{95}$$

The empirical covariance Σ computed from the particle population can be viewed as an estimator of $\Sigma^\star$. Consequently, injecting noise with covariance proportional to Σ is equivalent to using an approximation of the local inverse curvature H^{-1}, without explicitly computing or storing any Hessians.

The isotropic component serves as a robust fallback mechanism when Σ is noisy, low-rank, or misspecified, ensuring stable exploration even when the empirical geometry is unreliable.

8. Overall Algorithm

The resulting method alternates between:

1. diffusion-based propagation,
2. reward evaluation and Boltzmann reweighting,
3. geometric pCN mutation from elite particles,
4. multinomial resampling.

This procedure constitutes an SMC sampler with learned diffusion dynamics and reward-driven selection, closely related to evolutionary strategies while retaining a principled particle-based interpretation.

—

9. Remark on Exactness

Because proposal densities from diffusion and geometric pCN mutations are not explicitly accounted for in the weights, the method should be interpreted as a *reward-tilted SMC sampler* rather than an exact importance sampler for a known posterior. Its primary objective is optimization under a learned generative prior rather than unbiased density estimation.

Local quadratic model and Gaussian approximation.

$$\pi(x) \equiv p_t(x), \qquad \ell(x) := \log \pi(x). \tag{96}$$

$$\ell(x+\delta) \;\approx\; \ell(x) + s^\top \delta - \frac{1}{2}\,\delta^\top H \delta, \tag{97}$$

$$s := \nabla \ell(x), \qquad H := -\nabla^2 \ell(x), \qquad H \succeq 0. \tag{98}$$

$$\frac{\pi(x+\delta)}{\pi(x)} \;\approx\; \exp\!\left(s^\top \delta - \frac{1}{2}\,\delta^\top H \delta\right). \tag{99}$$

$$\exp\!\left(s^\top \delta - \frac{1}{2}\,\delta^\top H \delta\right) \;\propto\; \mathcal{N}\bigl(\delta;\, H^{-1}s,\, H^{-1}\bigr)\,. \tag{100}$$

$$\delta \sim \mathcal{N}\bigl(H^{-1}s,\, H^{-1}\bigr)\,. \tag{101}$$

Curvature–covariance correspondence.

$$H \approx \Sigma_\star^{-1} \quad\Longrightarrow\quad H^{-1} \approx \Sigma_\star. \tag{102}$$

$$\delta \approx \mathcal{N}(\Sigma_\star s,\, \Sigma_\star)\,. \tag{103}$$

Covariance-only geometric move.

$$\delta \sim \mathcal{N}(0,\, \Sigma_\star)\,. \tag{104}$$

9.1. Dimensions

The text of the paper should be formatted in two columns, with an overall width of 6.75 inches, height of 9.0 inches, and 0.25 inches between the columns. The left margin should be 0.75 inches and the top margin 1.0 inch (2.54 cm). The right and bottom margins will depend on whether you print on US letter or A4 paper, but all final versions must be produced for US letter size. Do not write anything on the margins.

The paper body should be set in 10 point type with a vertical spacing of 11 points. Please use Times typeface throughout the text.

9.2. Title

The paper title should be set in 14 point bold type and centered between two horizontal rules that are 1 point thick, with 1.0 inch between the top rule and the top edge of the page. Capitalize the first letter of content words and put the rest of the title in lower case. You can use TeX math in the title (we suggest sparingly), but no custom macros, images, or other TeX commands. Please make sure that accents, special characters, etc., are entered using TeX commands and not using non-English characters.

9.3. Author Information for Submission

ICML uses double-blind review, so author information must not appear. If you are using LaTeX and the `icml2026.sty` file, use `\icmlauthor{...}` to specify authors and `\icmlaffiliation{...}` to specify affiliations. (Read the TeX code used to produce this document for an example usage.) The author information will not be printed unless `accepted` is passed as an argument to the style file. Submissions that include the author information will not be reviewed.

9.3.1. Self-Citations

If you are citing published papers for which you are an author, refer to yourself in the third person. In particular, do not use phrases that reveal your identity (e.g., "in previous work (Langley, 2000), we have shown ...").

Do not anonymize citations in the reference section. The only exception are manuscripts that are not yet published (e.g., under submission). If you choose to refer to such unpublished manuscripts (Author, 2021), anonymized copies have to be submitted as Supplementary Material via OpenReview. However, keep in mind that an ICML paper should be self contained and should contain sufficient detail for the reviewers to evaluate the work. In particular, reviewers are not required to look at the Supplementary Material when writing their review (they are not required to look at more than the first 8 pages of the submitted document).

9.3.2. Camera-Ready Author Information

If a paper is accepted, a final camera-ready copy must be prepared. For camera-ready papers, author information should start 0.3 inches below the bottom rule surrounding the title. The authors' names should appear in 10 point bold type, in a row, separated by white space, and centered. Author names should not be broken across lines. Unbolded superscripted numbers, starting 1, should be used to refer to affiliations.

Affiliations should be numbered in the order of appearance. A single footnote block of text should be used to list all the affiliations. (Academic affiliations should list Department, University, City, State/Region, Country. Similarly for industrial affiliations.)

Each distinct affiliations should be listed once. If an author has multiple affiliations, multiple superscripts should be placed after the name, separated by thin spaces. If the authors would like to highlight equal contribution by multiple first authors, those authors should have an asterisk placed after their name in superscript, and the term "*Equal contribution" should be placed in the footnote block ahead of the list of affiliations. A list of corresponding authors and their emails (in the format Full Name <email@domain.com>) can follow the list of affiliations. Ideally only one or two names should be listed.

A sample file with author names is included in the ICML2026 style file package. Turn on the `[accepted]` option to the stylefile to see the names rendered. All of the guidelines above are implemented by the LaTeX style file.

9.4. Abstract

The paper abstract should begin in the left column, 0.4 inches below the final address. The heading 'Abstract' should be centered, bold, and in 11 point type. The abstract body should use 10 point type, with a vertical spacing of 11 points, and should be indented 0.25 inches more than normal on left-hand and right-hand margins. Insert 0.4 inches of blank space after the body. Keep your abstract brief and self-contained, limiting it to one paragraph and roughly 4–6 sentences. Gross violations will require correction at the camera-ready phase.

9.5. Partitioning the Text

You should organize your paper into sections and paragraphs to help readers place a structure on the material and understand its contributions.

9.5.1. Sections and Subsections

Section headings should be numbered, flush left, and set in 11 pt bold type with the content words capitalized. Leave 0.25 inches of space before the heading and 0.15 inches after the heading.

Similarly, subsection headings should be numbered, flush left, and set in 10 pt bold type with the content words capitalized. Leave 0.2 inches of space before the heading and 0.13 inches afterward.

Finally, subsubsection headings should be numbered, flush left, and set in 10 pt small caps with the content words capitalized. Leave 0.18 inches of space before the heading and 0.1 inches after the heading.

Please use no more than three levels of headings.

9.5.2. Paragraphs and Footnotes

Within each section or subsection, you should further partition the paper into paragraphs. Do not indent the first line of a given paragraph, but insert a blank line between succeeding ones.

You can use footnotes[2] to provide readers with additional information about a topic without interrupting the flow of the paper. Indicate footnotes with a number in the text where the point is most relevant. Place the footnote in 9 point type at the bottom of the column in which it appears. Precede the first footnote in a column with a horizontal rule of 0.8 inches.[3]

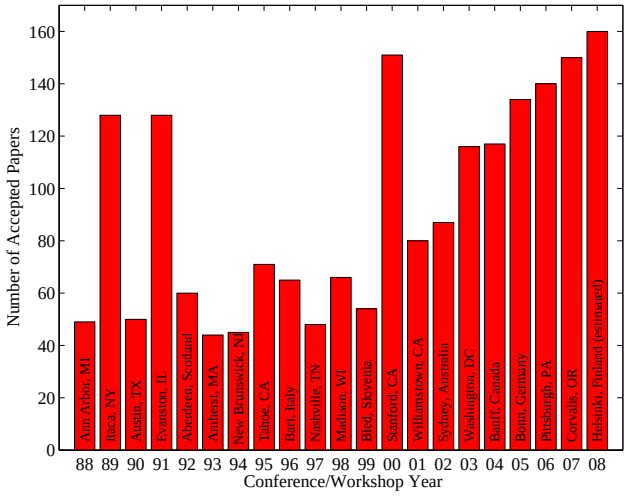

Figure 2. Historical locations and number of accepted papers for International Machine Learning Conferences (ICML 1993 – ICML 2008) and International Workshops on Machine Learning (ML 1988 – ML 1992). At the time this figure was produced, the number of accepted papers for ICML 2008 was unknown and instead estimated.

9.6. Figures

You may want to include figures in the paper to illustrate your approach and results. Such artwork should be centered, legible, and separated from the text. Lines should be dark and at least 0.5 points thick for purposes of reproduction, and text should not appear on a gray background.

Label all distinct components of each figure. If the figure takes the form of a graph, then give a name for each axis and include a legend that briefly describes each curve. Do not include a title inside the figure; instead, the caption should serve this function.

Number figures sequentially, placing the figure number and caption *after* the graphics, with at least 0.1 inches of space before the caption and 0.1 inches after it, as in Figure 2. The figure caption should be set in 9 point type and centered unless it runs two or more lines, in which case it should be flush left. You may float figures to the top or bottom of a column, and you may set wide figures across both columns (use the environment `figure*` in LaTeX). Always place two-column figures at the top or bottom of the page.

[2]Footnotes should be complete sentences.

[3]Multiple footnotes can appear in each column, in the same order as they appear in the text, but spread them across columns and pages if possible.

Algorithm 1 Bubble Sort

Input: data x_i, size m
repeat
  Initialize $noChange = true$.
  for $i = 1$ **to** $m - 1$ **do**
    if $x_i > x_{i+1}$ **then**
      Swap x_i and x_{i+1}
      $noChange = false$
    end if
  end for
until $noChange$ is $true$

Table 1. Classification accuracies for naive Bayes and flexible Bayes on various data sets.

Data set	Naive	Flexible	Better?
Breast	95.9± 0.2	96.7± 0.2	√
Cleveland	83.3± 0.6	80.0± 0.6	×
Glass2	61.9± 1.4	83.8± 0.7	√
Credit	74.8± 0.5	78.3± 0.6	
Horse	73.3± 0.9	69.7± 1.0	×
Meta	67.1± 0.6	76.5± 0.5	√
Pima	75.1± 0.6	73.9± 0.5	
Vehicle	44.9± 0.6	61.5± 0.4	√

9.7. Algorithms

If you are using LaTeX, please use the "algorithm" and "algorithmic" environments to format pseudocode. These require the corresponding stylefiles, algorithm.sty and algorithmic.sty, which are supplied with this package. Algorithm 1 shows an example.

9.8. Tables

You may also want to include tables that summarize material. Like figures, these should be centered, legible, and numbered consecutively. However, place the title *above* the table with at least 0.1 inches of space before the title and the same after it, as in Table 1. The table title should be set in 9 point type and centered unless it runs two or more lines, in which case it should be flush left.

Tables contain textual material, whereas figures contain graphical material. Specify the contents of each row and column in the table's topmost row. Again, you may float tables to a column's top or bottom, and set wide tables across both columns. Place two-column tables at the top or bottom of the page.

9.9. Theorems and Such

The preferred way is to number definitions, propositions, lemmas, etc. consecutively, within sections, as shown below.

Definition 9.1. A function $f : X \to Y$ is injective if for any $x, y \in X$ different, $f(x) \neq f(y)$.

Using Theorem 9.1 we immediate get the following result:

Proposition 9.2. *If f is injective mapping a set X to another set Y, the cardinality of Y is at least as large as that of X*

Proof. Left as an exercise to the reader. □

Theorem 9.3 stated next will prove to be useful.

Lemma 9.3. *For any $f : X \to Y$ and $g : Y \to Z$ injective functions, $f \circ g$ is injective.*

Theorem 9.4. *If $f : X \to Y$ is bijective, the cardinality of X and Y are the same.*

An easy corollary of Theorem 9.4 is the following:

Corollary 9.5. *If $f : X \to Y$ is bijective, the cardinality of X is at least as large as that of Y.*

Assumption 9.6. The set X is finite.

Remark 9.7. According to some, it is only the finite case (cf. Theorem 9.6) that is interesting.

9.10. Citations and References

Please use APA reference format regardless of your formatter or word processor. If you rely on the LaTeX bibliographic facility, use `natbib.sty` and `icml2026.bst` included in the style-file package to obtain this format.

Citations within the text should include the authors' last names and year. If the authors' names are included in the sentence, place only the year in parentheses, for example when referencing Arthur Samuel's pioneering work (1959). Otherwise place the entire reference in parentheses with the authors and year separated by a comma (Samuel, 1959). List multiple references separated by semicolons (Kearns, 1989; Samuel, 1959; Mitchell, 1980). Use the 'et al.' construct only for citations with three or more authors or after listing all authors to a publication in an earlier reference (Michalski et al., 1983).

Authors should cite their own work in the third person in the initial version of their paper submitted for blind review. Please refer to Section 9.3 for detailed instructions on how to cite your own papers.

Use an unnumbered first-level section heading for the references, and use a hanging indent style, with the first line of the reference flush against the left margin and subsequent lines indented by 10 points. The references at the end of this document give examples for journal articles (Samuel, 1959), conference publications (Langley, 2000), book chapters (Newell & Rosenbloom, 1981), books (Duda et al., 2000), edited volumes (Michalski et al., 1983), technical reports (Mitchell, 1980), and dissertations (Kearns, 1989).

Alphabetize references by the surnames of the first authors, with single author entries preceding multiple author entries. Order references for the same authors by year of publication, with the earliest first. Make sure that each reference includes all relevant information (e.g., page numbers).

Please put some effort into making references complete, presentable, and consistent, e.g. use the actual current name of authors. If using bibtex, please protect capital letters of names and abbreviations in titles, for example, use {B}ayesian or {L}ipschitz in your .bib file.

Accessibility

Authors are kindly asked to make their submissions as accessible as possible for everyone including people with disabilities and sensory or neurological differences. Tips of how to achieve this and what to pay attention to will be provided on the conference website `http://icml.cc/`.

Software and Data

If a paper is accepted, we strongly encourage the publication of software and data with the camera-ready version of the paper whenever appropriate. This can be done by including a URL in the camera-ready copy. However, **do not** include URLs that reveal your institution or identity in your submission for review. Instead, provide an anonymous URL or upload the material as "Supplementary Material" into the OpenReview reviewing system. Note that reviewers are not required to look at this material when writing their review.

Acknowledgements

Do not include acknowledgements in the initial version of the paper submitted for blind review.

If a paper is accepted, the final camera-ready version can (and usually should) include acknowledgements. Such acknowledgements should be placed at the end of the section, in an unnumbered section that does not count towards the

paper page limit. Typically, this will include thanks to reviewers who gave useful comments, to colleagues who contributed to the ideas, and to funding agencies and corporate sponsors that provided financial support.

Impact Statement

Authors are **required** to include a statement of the potential broader impact of their work, including its ethical aspects and future societal consequences. This statement should be in an unnumbered section at the end of the paper (co-located with Acknowledgements – the two may appear in either order, but both must be before References), and does not count toward the paper page limit. In many cases, where the ethical impacts and expected societal implications are those that are well established when advancing the field of Machine Learning, substantial discussion is not required, and a simple statement such as the following will suffice:

"This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here."

The above statement can be used verbatim in such cases, but we encourage authors to think about whether there is content which does warrant further discussion, as this statement will be apparent if the paper is later flagged for ethics review.

References

Author, N. N. Suppressed for anonymity, 2021.

Duda, R. O., Hart, P. E., and Stork, D. G. *Pattern Classification*. John Wiley and Sons, 2nd edition, 2000.

Ho, J., Jain, A., and Abbeel, P. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Kearns, M. J. *Computational Complexity of Machine Learning*. PhD thesis, Department of Computer Science, Harvard University, 1989.

Langley, P. Crafting papers on machine learning. In Langley, P. (ed.), *Proceedings of the 17th International Conference on Machine Learning (ICML 2000)*, pp. 1207–1216, Stanford, CA, 2000. Morgan Kaufmann.

Li, X., Zhao, Y., Wang, C., Scalia, G., Eraslan, G., Nair, S., Biancalani, T., Ji, S., Regev, A., Levine, S., et al. Derivative-free guidance in continuous and discrete diffusion models with soft value-based decoding. *arXiv preprint arXiv:2408.08252*, 2024.

Michalski, R. S., Carbonell, J. G., and Mitchell, T. M. (eds.). *Machine Learning: An Artificial Intelligence Approach, Vol. I.* Tioga, Palo Alto, CA, 1983.

Mitchell, T. M. The need for biases in learning generalizations. Technical report, Computer Science Department, Rutgers University, New Brunswick, MA, 1980.

Newell, A. and Rosenbloom, P. S. Mechanisms of skill acquisition and the law of practice. In Anderson, J. R. (ed.), *Cognitive Skills and Their Acquisition*, chapter 1, pp. 1–51. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, 1981.

Salomon-Ferrer, R., Case, D. A., and Walker, R. C. An overview of the amber biomolecular simulation package. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 3(2):198–210, 2013.

Samuel, A. L. Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3):211–229, 1959.

Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., and Ganguli, S. Deep unsupervised learning using nonequilibrium thermodynamics. In *International conference on machine learning*, pp. 2256–2265. PMLR, 2015.

Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., and Poole, B. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

A. You *can* have an appendix here.

You can have as much text here as you want. The main body must be at most 8 pages long. For the final version, one more page can be added. If you want, you can use an appendix like this one.

The `\onecolumn` command above can be kept in place if you prefer a one-column appendix, or can be removed if you prefer a two-column appendix. Apart from this possible change, the style (font size, spacing, margins, page numbering, etc.) should be kept the same as the main body.