ART for Diffusion Sampling: A Reinforcement Learning Approach to Timestep Scheduling

Yilie Huang
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hong Kong
yilie.huang@polyu.edu.hk

Wenpin Tang
Department of Industrial Engineering and Operations Research
Columbia University
New York, NY 10027, USA
wt2319@columbia.edu

Xun Yu Zhou
Department of Industrial Engineering and Operations Research and Data Science Institute
Columbia University
New York, NY 10027, USA
xz2574@columbia.edu

Abstract

We consider time discretization for score-based diffusion models to generate samples from a learned reverse-time dynamic on a finite grid. Uniform and hand-crafted grids can be suboptimal given a budget on the number of time steps. We introduce Adaptive Reparameterized Time (ART), which controls the clock speed of a reparameterized time variable to redistribute computation along the sampling trajectory while preserving the terminal time, with the objective of minimizing the aggregate Euler discretization error. We derive a randomized companion ART-RL that recasts ART as a continuous-time reinforcement learning problem with Gaussian policies, and prove a two-directional bridge between the two: the deterministic ART optimum lifts to an optimal Gaussian policy, and conversely any optimal Gaussian policy must recover the ART control through its mean. This bridge turns continuous-time actor–critic learning into a principled, rather than heuristic, route to the deterministic timestep optimum. Within the official EDM pipeline, ART-RL improves FID on CIFAR–10 across a wide range of budgets; after one-time offline training, the distilled deterministic schedule transfers without retraining to AFHQv2, FFHQ, and ImageNet at no extra inference cost.

1 Introduction

Diffusion models [6, 31, 32] are a family of generative models that create samples from unknown target distributions, and underpin recent advances in text-to-image creation [26, 27], text-to-video generation [25, 29, 5], and diffusion large language models [19, 24].

A diffusion model consists of two steps: learn the target score function by training on given samples (pretraining), and generate new samples from the learned model (inference). The focus of this paper is on diffusion sampling/inference, where a key numerical choice is the timestep grid used to discretize the learned model. Standard samplers often use uniform or hand-crafted timesteps [30, 32, 18, 2], while recent work also studies optimized fixed schedules [28]. Here our purpose is to provide a control-theoretic framework, which allows strategic and systematic time discretization in diffusion sampling in order to minimize the aggregate error when applying the Euler scheme. The main contribution of this paper is as follows:

- *Formulation and methodology*: We propose an optimal control framework – *Adaptive Reparameterized Time* (ART), to schedule timesteps for diffusion sampling. The idea is to treat the speed of the diffusion sampler as control to reparameterize time and adaptively redistribute computation along the sampling trajectory. To solve the ART problem, we

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

propose a continuous-time reinforcement learning (CTRL) approach – ART-RL, premised upon the recent development in Wang et al. [36], Jia and Zhou [15, 16].

- *Theory*: We establish a two-directional bridge between ART and ART-RL: *sufficiency* shows that the deterministic ART optimum lifts to an optimal Gaussian policy whose mean is exactly the ART control, and *necessity* shows that any optimal Gaussian policy must recover the ART control through its mean. This justifies parameterizing the actor as a Gaussian mean and deploying it deterministically with rigorous guarantees, and yields actor–critic updates for learning adaptive time schedules.
- *Experiments*: Our ART-RL schedule consistently outperforms Uniform, the DPM-Solver log-SNR grid [23], and EDM schedules [18] in a matched CIFAR–10 EDM evaluation. It improves FID across a broad range of sampling time budgets, with large gains at low and moderate budgets and a consistent advantage at the largest budget.
- *Generalization and amortization*: ART-RL separates offline schedule learning from inference-time sampling. After training, the stochastic policy is distilled into a deterministic time-only grid that can be reused as a drop-in schedule across timestep counts and transferred across various datasets without retraining.

To the best of our knowledge, this is the first control/RL formulation of diffusion timestep scheduling through adaptive time reparameterization, and the first to provide a two-directional bridge between a deterministic timestep optimum and an actor–critic learning procedure. The proposed method is data-driven and exhibits robust performance with reusable, amortized transfer.

Relevant literature: Diffusion models as generative tools were first proposed by Ho et al. [6] (DDPM) and Song et al. [30] (DDIM) in the discrete setting. The pioneering work of Song et al. [32] introduces a continuous-time formulation of diffusion models, providing a unified treatment that encompasses earlier discrete-time models. Various sampling methods have been proposed for diffusion inference, including predictor-corrector sampler [32], exponential integrator [38], higher-order solvers [39, 37], and optimized fixed timestep schedules [28]. See also Lee et al. [20], Chen et al. [2, 3], Li et al. [22], Benton et al. [1], Li and Yan [21], Huang et al. [7] for convergence analysis of diffusion models (with either uniform or hand-crafted timesteps).

Continuous-time reinforcement learning (CTRL) was first formulated by Wang et al. [36], where exploration is modeled via stochastic relaxed control to capture the trial-and-error nature of reinforcement learning. Subsequent work developed a model-free theoretical foundation via martingale methods [15–17, 35], established performance guarantees [8, 12, 13], and studied policy optimization [40]. The CTRL framework has also been applied to training and fine-tuning diffusion models in generative AI [4, 10, 41, 42].

Organization: Section 2 reviews diffusion models and introduces the ART control formulation. Section 3 describes how CTRL is used to solve ART, leading to the ART-RL approach. Section 4 presents the ART-RL algorithm, followed by empirical results in Section 5. Section 6 concludes. An extended ART–ART-RL analysis, including all proofs, and additional numerical results appear in the extended version [14].

2 Diffusion Models and ART

2.1 Continuous-Time Score-Based Diffusion Models

We briefly review continuous-time score-based diffusion models; see Tang and Zhao [34] for a recent survey. Given samples from an unknown target distribution, a forward process corrupts data over time $\tau \in [0, T]$ toward a tractable reference distribution, and sampling proceeds by integrating a learned reverse-time process.

Diffusion sampling. A forward process obeys the Itô SDE

$$\mathrm{d}\bar{x}(\tau) = -f(\tau)\bar{x}(\tau)\,\mathrm{d}\tau + g(\tau)\,\mathrm{d}w(\tau), \tag{1}$$

where $\tau \in [0, T]$, $\bar{x}(0) \sim p_0$ is the target data distribution, w is a standard Wiener process in $\mathbb{R}^d$, and f and g are coefficient functions. Let p_τ be the law of $\bar{x}(\tau)$ and $S(\tau, x) = \nabla_x \log p_\tau(x)$ be the score.

For sampling, we use the probability flow ODE associated with (1), which shares the same marginals as the reverse-time SDE; see e.g. Tang and Zhao [34, Theorem 5.1]. Denote the backward state by $\tilde{x}(\tau) \coloneqq \bar{x}(T-\tau)$ with initialization $\tilde{x}(0) \sim p_T$. Replacing S with a trained score model $\hat{S}$ yields

$$\frac{\mathrm{d}\tilde{x}(\tau)}{\mathrm{d}\tau} = f(T-\tau)\tilde{x}(\tau) + \tfrac{1}{2}g(T-\tau)^2\hat{S}(T-\tau, \tilde{x}(\tau)), \tag{2}$$

where $\tau \in [0, T]$.

Euler discretization. We integrate (2) on a grid $0 = \tau_0 < \tau_1 < \cdots < \tau_K = T$ with step sizes $h_i = \tau_{i+1} - \tau_i$, and write $\tilde{x}_i \coloneqq \tilde{x}(\tau_i)$. The explicit Euler update is

$$\tilde{x}_{i+1} = \tilde{x}_i + h_i \left[f(T-\tau_i)\tilde{x}_i + \tfrac{1}{2}g(T-\tau_i)^2\hat{S}(T-\tau_i, \tilde{x}_i) \right]. \tag{3}$$

A uniform grid $\tau_i = iT/K$ is simple but allocates evaluations uniformly even when the reverse dynamics vary along the trajectory. A simple intuition is that early stages, where samples resemble noise, may tolerate coarser steps, while later stages will benefit from finer resolutions. Such considerations motivate adaptive, data-driven schedules that redistribute steps under a given, fixed total time budget T.

2.2 ART: Time Reparameterization via Control

Motivated by the drawback of uniform timesteps in backward sampling discussed above, we introduce a reparameterized sampling clock and formalize adaptive timestep selection in the reverse process as a control problem. The key idea is to replace the uniform progression of original time by an adaptive, controlled progression that can accelerate in some segments and decelerate in others, thereby redistributing computational effort along the trajectory.

To this end, let $\psi : [0, T] \to \mathbb{R}$ be a continuous time mapping from the new clock t to the original diffusion time τ (i.e., $\tau = \psi(t)$), with $\psi(0) = 0$ and $\psi(T) = T$. On this new clock we write the state as $x(t) \coloneqq \tilde{x}(\psi(t))$, namely the backward state evaluated at diffusion time $\psi(t)$, with $x(0) \sim p_T$. We define the control as $\theta(t) \coloneqq \dot{\psi}(t)$, which quantifies the instantaneous rate of change in diffusion time with respect to t and satisfies $\int_0^T \theta(t)\,\mathrm{d}t = \psi(T) - \psi(0) = T$. To keep the formulation general, we do not impose monotonicity of ψ, so θ may take either sign. The monotone time-change case (equivalently, $\theta(t) \geq 0$ almost everywhere) is included as a special case of this formulation. In particular, if we discretize the new clock uniformly as $0 = t_0 < t_1 < \cdots < t_K = T$, then the induced original-time grid is $\tau_i \coloneqq \psi(t_i)$ with nonuniform step sizes $\tau_{i+1} - \tau_i$. Intuitively, $x(\cdot)$ tracks the generative trajectory under the reparameterized time, while $\theta(\cdot)$ determines how quickly the trajectory advances in original time per unit step on the new clock (larger $\theta(t)$ corresponds to faster progression at t). We refer to this time reparameterized sampling framework as *Adaptive Reparameterized Time* (ART).

By the chain rule, ART induces the controlled dynamics

$$\begin{cases} \dot{x}(t) = \theta(t)F(x(t), \psi(t)), \quad x(0) \sim p_T, & \text{(4a)} \\ \dot{\psi}(t) = \theta(t), \quad \psi(0) = 0,\ \psi(T) = T, & \text{(4b)} \end{cases}$$

where the backward probability–flow vector evaluated at original time $T - \psi$ is

$$F(x, \psi) \coloneqq f(T-\psi)x + \tfrac{1}{2}g(T-\psi)^2\hat{S}(T-\psi, x). \tag{5}$$

The constraint $\int_0^T \theta(t)\,\mathrm{d}t = T$ is the time budget allocated by the control.

To assess how Euler behaves on the t-clock, we relate its local approximation error over a step to the curvature of the right-hand side in (4a). We proceed analogously to the Euler discretization in (3). On an interval $t \in [t_i, t_{i+1})$ with stepsize $h_i := t_{i+1} - t_i$, the Euler update uses one control value per step; we denote this step value by θ_i.

Define the one-step Euler error proxy on the t-clock by $E_i := x(t_{i+1}) - \{x(t_i) + h_i\theta_i F(x(t_i), \psi(t_i))\}$. A Taylor expansion yields $E_i = \frac{h_i^2}{2}\theta_i^2 Q(x(t_i), \psi(t_i)) + O(h_i^3)$, where

$$Q(x, \psi) = A(s, x)B(s, x) - g(s)g'(s)\hat{S}(s, x) - f'(s)x - \tfrac{1}{2}g(s)^2\,\partial_s\hat{S}(s, x), \tag{6}$$

with $s := T - \psi$, $A(s,x) := f(s)I_d + \frac{1}{2}g(s)^2\nabla_x\hat{S}(s,x)$, and $B(s,x) := f(s)x + \frac{1}{2}g(s)^2\hat{S}(s,x)$. Thus, to leading order, the magnitude of the Euler local error on step i is proportional to $\theta_i^2|Q(x(t_i),\psi(t_i))|$, where $|\cdot|$ denotes the Euclidean norm throughout. In implementation, $\nabla_x\hat{S}(s,x)$ is not formed; automatic differentiation provides the needed Jacobian–vector product.

This motivates using $|Q(x,\psi)|\,\theta(t)^2$ as an error-density surrogate for allocating the time-warping rate. With a fixed time budget, we choose $\theta = \theta(\cdot)$ to minimize the overall residual surrogate subject to the constraint $\int_0^T \theta(t)\,\mathrm{d}t = T$, enforced via a Lagrange multiplier $\gamma \in \mathbb{R}$. Equivalently, we maximize the negative Lagrangian cost

$$J^\theta(s,y,\phi) = \mathbb{E}\Big[\int_s^T \big(-|Q(x(t),\psi(t))|\theta^2(t) - \gamma\theta(t)\big)\,\mathrm{d}t + \gamma T \;\Big|\; x(s) = y,\ \psi(s) = \phi\Big]. \tag{7}$$

Denote the optimal value function to be

$$V(s,y,\phi) = \max_{\theta=\theta(\cdot)} J^\theta(s,y,\phi). \tag{8}$$

In summary, ART recasts timestep allocation as continuous-time control of the time-warping rate θ under (4) and (7). The next section develops an RL-based procedure (ART–RL) to learn θ in a data-driven way.

3 Randomized Control and RL Formulation

The ART problem has no closed-form solution and its HJB is intractable on the high-dimensional state space $x \in \mathbb{R}^d$. We therefore introduce a Gaussian randomization of the control as a technical device that embeds ART into the continuous-time RL framework, enabling actor–critic learning. We call this approach *Adaptive Reparameterized Time via Reinforcement Learning* (ART-RL).

3.1 ART-RL: Auxiliary Problem with Gaussian Policies

We model control randomization with a stochastic policy over the time-warping rate θ. Inspired by entropy-regularized control [43, 11], we consider the Gaussian family indexed by $\lambda \geq 0$:

$$\pi^{(\lambda)}(\cdot \mid t,x,\psi) = \mathcal{N}\left(\mu(t,x,\psi),\ \frac{\lambda}{|Q(x,\psi)|}\right), \tag{9}$$

where μ is a measurable deterministic policy and Q is defined in (6). Let $\Pi^{(\lambda)}$ be the class of such policies with finite second moments. Since Q is proportional to the Euler local-error proxy, the variance is smaller in stiff regions and larger elsewhere; λ controls the overall noise level without changing the mean. For analysis we assume $|Q(x,\psi)| > 0$ almost surely on compact intervals; in practice we use $|Q| \vee \varepsilon$ for small $\varepsilon > 0$.

Given a policy $\pi^{(\lambda)} \in \Pi^{(\lambda)}$, let $\big(x^{\pi^{(\lambda)}}(t), \psi^{\pi^{(\lambda)}}(t)\big)_{t\in[0,T]}$ denote the corresponding state process which, according to Wang et al. [36], satisfy

$$\begin{cases} \dot{x}^{\pi^{(\lambda)}}(t) = \displaystyle\int_{\mathbb{R}} \theta\, F(x^{\pi^{(\lambda)}}(t), \psi^{\pi^{(\lambda)}}(t))\, \pi_t^{(\lambda)}(\theta)\,\mathrm{d}\theta, \quad x^{\pi^{(\lambda)}}(0) = x_0 \sim p_T, & \text{(10a)} \\ \dot{\psi}^{\pi^{(\lambda)}}(t) = \displaystyle\int_{\mathbb{R}} \theta\, \pi_t^{(\lambda)}(\theta)\,\mathrm{d}\theta, \quad \psi^{\pi^{(\lambda)}}(0) = 0,\ \psi^{\pi^{(\lambda)}}(T) = T. & \text{(10b)} \end{cases}$$

The corresponding total-time constraint is $\int_0^T \int_{\mathbb{R}} \theta\pi_t^{(\lambda)}(\theta)\mathrm{d}\theta\mathrm{d}t = T$. The objective function for the auxiliary problem is

$$J^{\pi^{(\lambda)}}(s,y,\phi) = \mathbb{E}\Big[\gamma T + \lambda T + \int_s^T \int_{\mathbb{R}} \pi_t^{(\lambda)}(\theta)\big(-|Q(x^{\pi^{(\lambda)}}(t), \psi^{\pi^{(\lambda)}}(t))|\theta^2 - \gamma\theta\big)\mathrm{d}\theta\,\mathrm{d}t \;\Big|\; x^{\pi^{(\lambda)}}(s) = y,\ \psi^{\pi^{(\lambda)}}(s) = \phi\Big],$$

where the Lagrange multiplier $\gamma \in \mathbb{R}$ enforces the total-time constraint and $\lambda \geq 0$ is the variance parameter from (9). The optimal value function is

$$V^{(\lambda)}(s, y, \phi) \;=\; \max_{\pi^{(\lambda)} \in \Pi^{(\lambda)}} J^{\pi^{(\lambda)}}(s, y, \phi). \tag{11}$$

3.2 Relationship between ART and ART-RL

We now establish the relationship between the ART control problem (8) and the ART-RL auxiliary problem (11). By dynamic programming, the optimal value function V for (8) satisfies

$$V_t + \sup_{\theta} \Big\{ (V_x^\top F(x, \psi) + V_\psi - \gamma)\theta - |Q(x, \psi)|\theta^2 \Big\} = 0, \tag{12}$$

with terminal condition $V(T, x, \psi) = \gamma T$.

Moreover, by Wang et al. [36], the optimal value function $V^{(\lambda)}$ of (11) satisfies the exploratory HJB:

$$V_t^{(\lambda)} + \sup_{\mu} \Big\{ \big(V_x^{(\lambda)\top} F(x, \psi) + V_\psi^{(\lambda)} - \gamma\big)\mu - |Q(x, \psi)| \Big(\mu^2 + \frac{\lambda}{|Q(x, \psi)|} \Big) \Big\} = 0 \tag{13}$$

with terminal condition $V^{(\lambda)}(T, x, \psi) = (\gamma + \lambda)T$.

The following three results characterize the connection: a value-shift theorem aligning the deterministic and auxiliary HJB equations, followed by sufficiency and necessity theorems that together establish a two-directional bridge. All proofs can be found in the extended version [14].

Theorem 3.1 (Value function shift)**.** *If V is a classical solution of* (12)*, then*

$$V^{(\lambda)}(t, x, \psi) = V(t, x, \psi) + \lambda t$$

is a classical solution of (13)*.*

Theorem 3.2 (Sufficiency: deterministic ART optimum lifts to an optimal Gaussian policy)**.** *Under the conditions of Theorem 3.1, define*

$$\mu^*(t, x, \psi) = \frac{V_x^\top F(x, \psi) + V_\psi - \gamma}{2|Q(x, \psi)|}.$$

Then the Gaussian policy

$$\pi^{(\lambda)*}(\cdot \mid t, x, \psi) = \mathcal{N}\Big(\mu^*(t, x, \psi),\; \frac{\lambda}{|Q(x, \psi)|} \Big)$$

is optimal for the auxiliary problem (11) *subject to the dynamics* (10)*. Furthermore, μ^* is an optimal policy for the original problem* (8) *subject to the dynamics* (4)*.*

Theorem 3.3 (Necessity: optimal auxiliary policy recovers the ART control)**.** *Suppose Theorem 3.1 holds, $|Q(x, \psi)| > 0$, and the policy-gap residual $R^{\pi^{(\lambda)}}(t, x, \psi) = |Q(x, \psi)|(\mu^{\pi^{(\lambda)}}(t, x, \psi) - \mu^*(t, x, \psi))^2$ is continuous in (t, x, ψ). Let $\pi^{(\lambda),*} \in \Pi^{(\lambda)}$ be any Gaussian policy that is optimal for the auxiliary problem* (11) *at every initial pair (s, y, ϕ). Then its mean field satisfies*

$$\mu^{\pi^{(\lambda),*}}(t, x, \psi) \;=\; \mu^*(t, x, \psi) \qquad \textit{for every } (t, x, \psi),$$

where μ^* is the optimal ART control of Theorem 3.2. In particular, $\mu^{\pi^{(\lambda),*}}$ is also an optimal control for the original ART problem (8)*.*

Theorems 3.1–3.3 establish a *two-directional* equivalence between ART and ART-RL. Hence the actor in Section 4 can be parameterized as a Gaussian mean and deployed deterministically, with the guarantee that any optimal auxiliary policy recovers the ART control through its mean. This bridge is what makes continuous-time actor–critic learning a principled, rather than heuristic, route to the deterministic timestep problem.

4 ART-RL Actor–Critic Algorithm

Theorems 3.1–3.3 reduce ART to solving the auxiliary ART-RL problem and guarantee that any optimal auxiliary policy mean recovers the deterministic ART optimum.

4.1 Algorithm Design

Our algorithm builds on the continuous-time actor–critic framework of Jia and Zhou [16], adapted to ART-RL with randomized Gaussian control. We use two neural networks NN^{ϑ_c} and NN^{ϑ_a} for the value function and policy mean, and parameterize

$$\hat{V}^{\vartheta_c}(t,x,\psi) = NN^{\vartheta_c}(t,x,\psi) + \lambda t, \qquad \hat{\pi}^{\vartheta_a}(\cdot \mid t,x,\psi) = \mathcal{N}\left(NN^{\vartheta_a}(t,x,\psi), \frac{\lambda}{|Q(x,\psi)|}\right). \tag{14}$$

Let $\theta^{\hat{\pi}^{\vartheta_a}} \sim \hat{\pi}^{\vartheta_a}$ be a Gaussian control process with corresponding (observable) state process $(x^{\theta^{\hat{\pi}^{\vartheta_a}}}(t), \psi^{\theta^{\hat{\pi}^{\vartheta_a}}}(t))$. For brevity, in (15) we write $(x(t),\psi(t),\theta(t)) = (x^{\theta^{\hat{\pi}^{\vartheta_a}}}(t), \psi^{\theta^{\hat{\pi}^{\vartheta_a}}}(t), \theta^{\hat{\pi}^{\vartheta_a}}(t))$. By Jia and Zhou [16], $\hat{V}^{\vartheta_c}$ and $\hat{\pi}^{\vartheta_a}$ satisfy the coupled moment conditions

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \frac{\partial NN^{\vartheta_c}(t,x(t),\psi(t))}{\partial \vartheta_c}\Big(\mathrm{d}\hat{V}^{\vartheta_c}(t,x(t),\psi(t)) - \big(|Q(x(t),\psi(t))|\theta(t)^2 + \gamma\theta(t)\big)\mathrm{d}t\Big)\right] = 0,\\
&\mathbb{E}\left[\int_0^T \frac{\partial \log \hat{\pi}^{\vartheta_a}(\theta(t) \mid t,x,\psi)}{\partial \vartheta_a}\Big(\mathrm{d}\hat{V}^{\vartheta_c}(t,x(t),\psi(t)) - \big(|Q(x(t),\psi(t))|\theta(t)^2 + \gamma\theta(t)\big)\mathrm{d}t\Big)\right] = 0.
\end{aligned} \tag{15}$$

These are standard moment conditions in RL [33, 9]; the parameters ϑ_c and ϑ_a are updated by stochastic approximation.

4.2 Implementable ART-RL Algorithm

At iteration n, a trajectory (x_n,ψ_n,θ_n) is generated under the current policy $\hat{\pi}^{\vartheta_{a,n}}$ on a uniform grid $0 = t_0 < \cdots < t_K = T$ with $\Delta t = T/K$, and parameters are updated with learning rate $a_n > 0$. Writing $\hat{V}_k^{\vartheta_{c,n}} := \hat{V}^{\vartheta_{c,n}}(t_k, x_n(t_k), \psi_n(t_k))$ and applying stochastic approximation with Riemann discretization to the moment conditions (15) gives, with $D_{n,k} := \hat{V}_{k+1}^{\vartheta_{c,n}} - \hat{V}_k^{\vartheta_{c,n}} - \gamma_n\theta_n(t_k)\Delta t - |Q(x_n(t_k),\psi_n(t_k))|\theta_n(t_k)^2\Delta t$, the critic and actor updates

$$\vartheta_{c,n+1} \leftarrow \vartheta_{c,n} + a_n \sum_{k=0}^{K-1} \frac{\partial NN^{\vartheta_{c,n}}}{\partial \vartheta_c}\big(t_k, x_n(t_k), \psi_n(t_k)\big) D_{n,k}, \tag{16a}$$

$$\vartheta_{a,n+1} \leftarrow \vartheta_{a,n} + a_n \sum_{k=0}^{K-1} \frac{\partial \log \hat{\pi}_{t_k}^{\vartheta_{a,n}}(\theta_n(t_k))}{\partial \vartheta_a} D_{n,k}, \tag{16b}$$

where $\hat{\pi}_{t_k}^{\vartheta_{a,n}}(\theta) := \hat{\pi}^{\vartheta_{a,n}}(\theta \mid t_k, x_n(t_k), \psi_n(t_k))$. The Lagrange multiplier is updated by $\gamma_{n+1} \leftarrow \gamma_n + a_n\big(\psi_n(T) - T\big)$, where $\psi_n(T) = \psi_n(t_K)$. The full procedure is summarized in Algorithm 1.

Algorithm 1 Time-discretized ART-RL Actor-Critic

for $n = 1$ to N **do**
  Set $k = 0$, $t = t_k = 0$, initialize $(x_n(t_0), \psi_n(t_0))$
  while $t < T$ **do**
    Sample control $\theta_n(t_k)$ from the Gaussian policy (14)
    Update the states by the ART dynamics (4)
    Increment time: $t_{k+1} = t_k + \Delta t$, $k \leftarrow k+1$
  end while
  Collect trajectory $\{(t_k, x_n(t_k), \psi_n(t_k), \theta_n(t_k))\}_{k=0}^{K-1}$
  Update critic and actor parameters via (16a) and (16b)
  Update the Lagrange multiplier γ_{n+1}
end for

5 Numerical Experiments

We evaluate ART-RL in a one-dimensional known-score setting and in image-generation experiments using the official EDM pipeline [18]. Across CIFAR–10, AFHQv2, FFHQ, and ImageNet, the score

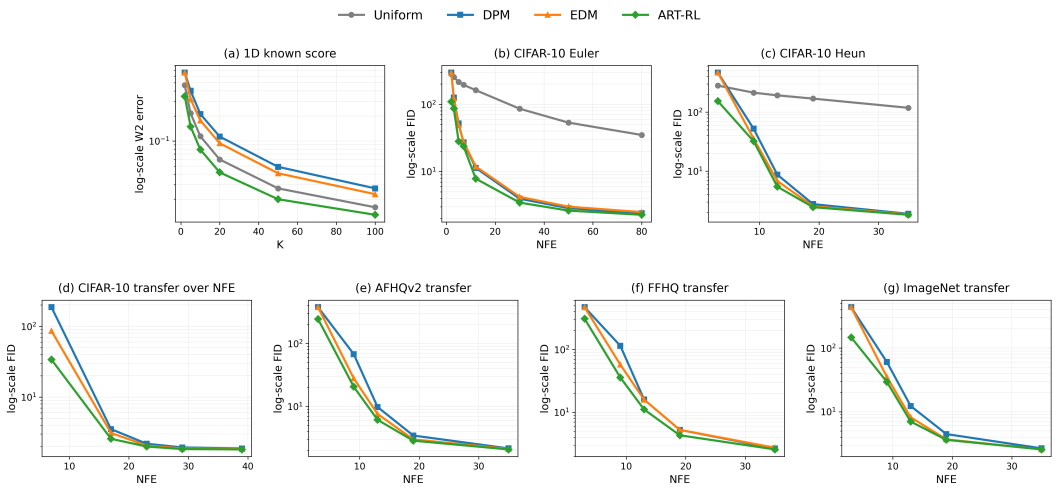


Figure 1: Quantitative overview: ART-RL Pareto-dominates Uniform, DPM-logSNR, and EDM on CIFAR–10 across NFE budgets, and the same distilled grid retains the advantage on AFHQv2, FFHQ, and ImageNet without retraining. DPM-logSNR is the DPM-Solver uniform log-SNR grid.

model, solver, and EDM implementation are kept fixed and only the timesteps are replaced. We compare Uniform, the DPM-Solver log-SNR grid, EDM timesteps, and ART-RL, reporting W_2 on the 1D problem and FID vs. NFE on images; on CIFAR–10 we additionally include AYS [28]. Additional experimental details and results are in the extended version [14].

5.1 One–Dimensional Example with Analytical Score

We first consider a one-dimensional example where the score function is analytical in closed form; hence every performance difference comes from the timestep schedule rather than score-estimation error. The forward process starts from $p_0 = \mathcal{N}(0, 1)$ and follows $\mathrm{d}x(t) = \sqrt{2t}\,\mathrm{d}w(t)$ over $T = 3$, so $x(t) \sim \mathcal{N}(0, 1 + t^2)$, $p_T = \mathcal{N}(0, 10)$, and $S(t, x) = -x/(1 + t^2)$. Substituting this score into (5) and (6) gives $F(x, \psi) = -(T - \psi)x/(1 + (T - \psi)^2)$ and $Q(x, \psi) = x/(1 + (T - \psi)^2)^2$.

This controlled setting separates learned from hand-designed schedules: EDM and DPM-logSNR are strong but fixed analytic prescriptions, whereas ART-RL learns the time allocation from the objective. In Figure 1-(a), ART-RL lowers W_2 across all tested K; for example, at $K = 10$, ART-RL obtains .079 versus .177 (EDM) and .211 (DPM-logSNR), and at $K = 100$, .013 versus .023 and .027.

Distillation diagnostic. Training trajectories concentrate tightly around a smooth mean control, so the stochastic policy can be distilled into a deterministic time-only grid, with two practical benefits. First, distillation removes the per-step cost of actor and Q evaluations, so ART-RL sampling requires no extra computation beyond standard schemes such as Uniform or EDM. Second, it eliminates the residual $\psi(T)$ overshoot/undershoot that occurs when θ is produced by a neural-network actor: normalizing the deterministic increments to sum exactly to T guarantees the induced grid hits the terminal time precisely, improving the numerical fidelity of the discretized probability flow ODE.

5.2 CIFAR–10 under the EDM Pipeline

We first run a solver-aligned Euler ablation to verify the objective–sampler alignment. ART-RL improves FID over Uniform, DPM-logSNR, and EDM across all Euler budgets; for example, at $\mathrm{NFE} = 30$ it attains 3.46 versus 3.95 (DPM-logSNR), 4.21 (EDM), and 85.83 (Uniform).

We report the main image results with the default Heun solver. Heun is the sampler used in the official EDM pipeline and a 2nd-order Runge–Kutta method (one order higher than Euler), which makes it both the standard choice for modern low-NFE deterministic diffusion sampling and the configuration in which EDM reports its strongest FID; running under Heun therefore tests ART-RL against the strongest baseline setup. Although our training surrogate is Euler-based, we assess transfer of the Euler-trained schedule under Heun empirically below. We test $K \in \{2, 5, 7, 10, 18\}$ with Euler only

for the final step update, matching the EDM implementation and giving $\mathrm{NFE} = 2K - 1$. Table 1 shows that ART-RL attains the lowest FID across all baselines. The improvement is most pronounced at small and moderate budgets, and remains present at the largest-budget setting where EDM reported its strongest performance: at $\mathrm{NFE} = 35$, ART-RL obtains 1.82, compared with 1.85 for EDM and 1.89 for DPM-logSNR.

Table 1: CIFAR–10 FID vs. NFE. DPM-logSNR and EDM use fixed analytic schedules, whereas AYS is an optimized fixed reference schedule. At $\mathrm{NFE} = 3$, all three use the same boundary noise levels and therefore yield the same FID; ART-RL instead learns its schedule from data.

NFE	3	9	13	19	35
Uniform	280.29	213.13	191.69	168.87	118.02
AYS	465.83	157.29	86.82	35.30	3.60
DPM-logSNR	465.83	52.29	8.67	2.76	1.89
EDM	465.83	35.54	6.79	2.54	1.85
ART-RL	**152.86**	**32.13**	**5.44**	**2.45**	**1.82**

Visual comparisons corroborate the table: at very small budgets ART-RL already produces recognizable objects while EDM and DPM-logSNR remain close to noise.

Statistical significance. At $\mathrm{NFE} = 35$, three matched 50,000-sample runs give ART-RL FIDs $1.82, 1.79, 1.82$ versus EDM FIDs $1.85, 1.83, 1.85$; ART-RL's worst run beats EDM's best. For context, the ART-RL vs. EDM gaps on CIFAR–10 are 313, 3.4, and 1.35 at $\mathrm{NFE} = 3, 9, 13$, respectively; these selected gaps are larger than the seed-to-seed 50k-sample FID standard deviation reported by EDM (≈ 0.02–0.05).

5.3 Generalization of the ART-RL Time Schedule

The ART-RL schedule is learned once on CIFAR–10 (1–2 hours on a Colab T4, offline), distilled to a fixed grid, and reused across timestep counts and datasets without retraining. After distillation, inference uses only the precomputed grid, just as Uniform, DPM-logSNR, and EDM do. For the transfer studies below we focus on the strongest baselines (DPM-logSNR and EDM); Uniform and AYS are omitted since both are substantially weaker in the matched CIFAR–10 evaluation (Table 1).

5.3.1 Robustness Across Time Grids via Interpolation and Extrapolation

To test reuse across timestep counts, we take the ART-RL schedule learned on CIFAR–10 at $K = 18$ and construct new grids for $K \in \{4, 9, 12, 15, 20\}$ by log-linear interpolation and extrapolation; EDM and DPM-logSNR use their analytic timestep sequences.

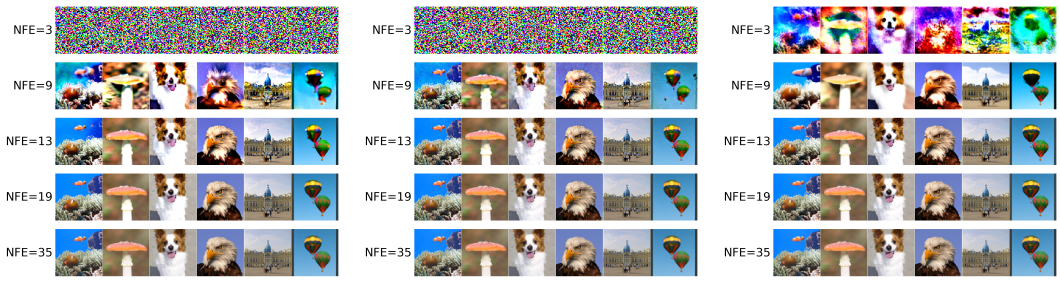
Table 2: CIFAR–10 interpolation/extrapolation FID.

NFE	7	17	23	29	39
DPM-logSNR	185.63	3.52	2.19	1.94	1.88
EDM	85.80	3.11	2.06	1.88	1.85
ART-RL	**33.73**	**2.57**	**2.00**	**1.84**	**1.82**

Table 2 shows ART-RL preserves the Section 5.2 hierarchy across interpolated and extrapolated counts.

5.3.2 Cross-Dataset Transfer: AFHQv2, FFHQ, and ImageNet

We plug the CIFAR–10 ART-RL schedule into the official EDM checkpoints for AFHQv2, FFHQ, and ImageNet, testing a drop-in replacement without dataset-specific retraining.



(a) DPM-logSNR (b) EDM (c) ART-RL

Figure 2: ImageNet samples across NFEs.

Table 3: Cross-dataset FID using the CIFAR–10 ART-RL schedule.

Dataset	Method	3	9	13	19	35
AFHQv2	DPM-logSNR	375.76	67.64	9.77	3.44	2.15
	EDM	375.76	27.88	7.56	2.99	2.11
	ART-RL	**243.48**	**20.48**	**6.12**	**2.85**	**2.07**
FFHQ	DPM-logSNR	466.76	113.87	15.94	5.25	2.66
	EDM	466.76	57.13	15.87	5.26	2.73
	ART-RL	**305.97**	**35.73**	**11.08**	**4.31**	**2.57**
ImageNet	DPM-logSNR	437.42	60.48	12.31	4.46	2.66
	EDM	437.42	35.32	8.18	3.68	2.57
	ART-RL	**147.21**	**29.49**	**7.01**	**3.62**	**2.56**

ART-RL improves over EDM and DPM-logSNR at low and moderate NFEs across all three datasets (Figure 2 and further visual comparisons in Huang et al. [14]), confirming a reusable rather than dataset-specific parameterization.

6 Conclusions

We introduced ART for timestep allocation in score-based diffusion models. Its main modeling step is to replace direct grid selection by a continuous time change whose rate is a learnable control that redistributes computation under a fixed time budget. ART-RL solves this control problem through an auxiliary randomized formulation: sufficiency and necessity theorems show that the deterministic ART optimum and the mean of any optimal Gaussian auxiliary policy coincide, making actor–critic learning a principled route to a deterministic schedule. Empirically, within matched EDM samplers, the distilled ART-RL grid improves FID over Uniform, DPM-logSNR, and EDM on CIFAR–10, generalizes across step counts, and transfers without retraining to AFHQv2, FFHQ, and ImageNet as a drop-in time grid with no inference overhead.

Future directions include extending ART to stochastic samplers, designing higher-order or solver-aware error surrogates, and identifying settings where state-dependent schedules improve over distilled time-only grids. More broadly, adaptive time reparameterization for diffusion sampling remains nascent, and ART with ART-RL offers a control-theoretic starting point for systematic schedule design in generative diffusion models.

Acknowledgments and Disclosure of Funding

Funding. Yilie Huang acknowledges financial support from the Start-up Fund of The Hong Kong Polytechnic University (Project ID: P0063874). Wenpin Tang is supported by NSF CAREER Award DMS-2538791 and NSF MFAI Award DMS-2602038. Xun Yu Zhou is supported by the Nie Center for Intelligent Asset Management and the Yao Artificial Intelligence Research Fund at Columbia University.

Competing Interests. The authors declare no competing interests.

References

[1] Joe Benton, Valentin De Bortoli, Arnaud Doucet, and George Deligiannidis. Nearly d-linear convergence bounds for diffusion models via stochastic localization. In *International Conference on Learning Representations*, 2024.

[2] Hongrui Chen, Holden Lee, and Jianfeng Lu. Improved analysis of score-based generative modeling: User-friendly bounds under minimal smoothness assumptions. In *International Conference on Machine Learning*, pages 4735–4763, 2023.

[3] Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, and Anru R Zhang. Sampling is as easy as learning the score: theory for diffusion models with minimal data assumptions. In *International Conference on Learning Representations*, 2023.

[4] Xuefeng Gao, Jiale Zha, and Xun Yu Zhou. Reward-directed score-based diffusion models via q-learning. *Journal of Machine Learning Research*, 26(302):1–46, 2025.

[5] Google. State-of-the-art video and image generation with Veo 2 and Imagen 3. `https://blog.google/technology/google-labs/video-image-generation-update-december-2024/`, 2024. Accessed: 2025-09-17.

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

[7] Daniel Zhengyu Huang, Jiaoyang Huang, and Zhengjiang Lin. Convergence analysis of probability flow ODE for score-based generative models. *arXiv preprint arXiv:2404.09730*, 2025.

[8] Yilie Huang. Continuous-time reinforcement learning for asset–liability management. In *ACM International Conference on AI in Finance*, pages 360–368, 2025.

[9] Yilie Huang and Xun Yu Zhou. Data-driven exploration for a class of continuous-time indefinite linear–quadratic reinforcement learning problems. *IEEE Transactions on Automatic Control*, 2026. doi: 10.1109/TAC.2026.3734857. Early Access.

[10] Yilie Huang and Xun Yu Zhou. Amortized guidance for image inpainting with pretrained diffusion models. In *Advances in Neural Information Processing Systems*, 2026. To appear; arXiv:2605.13010.

[11] Yilie Huang, Yanwei Jia, and Xun Yu Zhou. Achieving mean–variance efficiency by continuous-time reinforcement learning. In *ACM International Conference on AI in Finance*, pages 377–385, 2022.

[12] Yilie Huang, Yanwei Jia, and Xun Yu Zhou. Sublinear regret for a class of continuous-time linear-quadratic reinforcement learning problems. *SIAM Journal on Control and Optimization*, 63(5):3452–3474, 2025.

[13] Yilie Huang, Yanwei Jia, and Xun Yu Zhou. Mean–variance portfolio selection by continuous-time reinforcement learning: Algorithms, regret analysis, and empirical study. *Management Science*, 2026. To appear; arXiv:2412.16175.

[14] Yilie Huang, Wenpin Tang, and Xun Yu Zhou. Adaptive reparametrized time for score-based diffusion sampling. *Journal of Machine Learning Research*, 2026. To appear; arXiv:2607.02137.

[15] Yanwei Jia and Xun Yu Zhou. Policy evaluation and temporal-difference learning in continuous time and space: A martingale approach. *Journal of Machine Learning Research*, 23(154):1–55, 2022.

[16] Yanwei Jia and Xun Yu Zhou. Policy gradient and actor-critic learning in continuous time and space: Theory and algorithms. *Journal of Machine Learning Research*, 23(275):1–50, 2022.

[17] Yanwei Jia and Xun Yu Zhou. q-learning in continuous time. *Journal of Machine Learning Research*, 24(161):1–61, 2023.

[18] Tero Karras, Miika Aittala, Timo Aila, and Samuli Laine. Elucidating the design space of diffusion-based generative models. In *Advances in Neural Information Processing Systems*, volume 35, pages 26565–26577, 2022.

[19] Samar Khanna, Siddhant Kharbanda, Shufan Li, Harshit Varma, Eric Wang, Sawyer Birnbaum, Ziyang Luo, Yanis Miraoui, Akash Palrecha, and Stefano Ermon. Mercury: Ultra-fast language models based on diffusion. *arXiv preprint arXiv:2506.17298*, 2025.

[20] Holden Lee, Jianfeng Lu, and Yixin Tan. Convergence for score-based generative modeling with polynomial complexity. In *Advances in Neural Information Processing Systems*, volume 35, pages 22870–22882, 2022.

[21] Gen Li and Yuling Yan. Adapting to unknown low-dimensional structures in score-based diffusion models. In *Advances in Neural Information Processing Systems*, volume 37, pages 126297–126331, 2024.

[22] Gen Li, Yuting Wei, Yuxin Chen, and Yuejie Chi. Towards faster non-asymptotic convergence for diffusion-based generative models. In *International Conference on Learning Representations*, 2024.

[23] Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. DPM-Solver: A fast ODE solver for diffusion probabilistic model sampling in around 10 steps. In *Advances in Neural Information Processing Systems*, volume 35, pages 5775–5787, 2022.

[24] Shen Nie, Fengqi Zhu, Zebin You, Xiaolu Zhang, Jingyang Ou, Jun Hu, Jun Zhou, Yankai Lin, Ji-Rong Wen, and Chongxuan Li. Large language diffusion models. *arXiv preprint arXiv:2502.09992*, 2025.

[25] OpenAI. Sora: Creating video from text. `https://openai.com/sora`, 2024. Accessed: 2025-09-17.

[26] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, Casey Chu, and Mark Chen. Hierarchical text-conditional image generation with CLIP latents. *arXiv preprint arXiv:2204.06125*, 2022.

[27] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10684–10695, 2022.

[28] Amirmojtaba Sabour, Sanja Fidler, and Karsten Kreis. Align your steps: Optimizing sampling schedules in diffusion models. In *International Conference on Machine Learning*, pages 42947–42975, 2024.

[29] Uriel Singer, Adam Polyak, Thomas Hayes, Xi Yin, Jie An, Songyang Zhang, Qiyuan Hu, Harry Yang, Oron Ashual, and Oran Gafni. Make-A-Video: Text-to-video generation without text-video data. In *International Conference on Learning Representations*, 2023.

[30] Jiaming Song, Chenlin Meng, and Stefano Ermon. Denoising diffusion implicit models. In *International Conference on Learning Representations*, 2021.

[31] Yang Song and Stefano Ermon. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems*, volume 32, pages 11918–11930, 2019.

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

[33] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT Press, Cambridge, MA, 1998.

[34] Wenpin Tang and Hanyang Zhao. Score-based diffusion models via stochastic differential equations. *Statistics Surveys*, 19:28–64, 2025.

[35] Wenpin Tang and Xun Yu Zhou. Regret of exploratory policy improvement and q-learning. *arXiv preprint arXiv:2411.01302*, 2024.

[36] Haoran Wang, Thaleia Zariphopoulou, and Xun Yu Zhou. Reinforcement learning in continuous time and space: A stochastic control approach. *Journal of Machine Learning Research*, 21 (198):1–34, 2020.

[37] Yuchen Wu, Yuxin Chen, and Yuting Wei. Stochastic Runge-Kutta methods: Provable acceleration of diffusion models. *arXiv preprint arXiv:2410.04760*, 2024.

[38] Qinsheng Zhang and Yongxin Chen. Fast sampling of diffusion models with exponential integrator. In *International Conference on Learning Representations*, 2023.

[39] Qinsheng Zhang, Jiaming Song, and Yongxin Chen. Improved order analysis and design of exponential integrator for diffusion models sampling. *arXiv preprint arXiv:2308.02157*, 2023.

[40] Hanyang Zhao, Wenpin Tang, and David D Yao. Policy optimization for continuous reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

[41] Hanyang Zhao, Haoxian Chen, Ji Zhang, David D Yao, and Wenpin Tang. Scores as Actions: a framework of fine-tuning diffusion models by continuous-time reinforcement learning. *arXiv preprint arXiv:2409.08400*, 2024.

[42] Hanyang Zhao, Haoxian Chen, Ji Zhang, David Yao, and Wenpin Tang. Score as Action: Fine tuning diffusion generative models by continuous-time reinforcement learning. In *International Conference on Machine Learning*, 2025.

[43] Brian D Ziebart, Andrew L Maas, J Andrew Bagnell, and Anind K Dey. Maximum entropy inverse reinforcement learning. In *AAAI Conference on Artificial Intelligence*, volume 8, pages 1433–1438, Chicago, IL, USA, 2008.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes].

   Justification: The abstract and introduction state the method, theory claim, and empirical scope; the experiments and the cited extended version provide the corresponding evidence.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes].

   Justification: The main text specifies the fixed EDM/probability-flow ODE scope and the Euler-proxy/Heun evaluation setting, and the cited extended version explains the distilled time-only grid.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [Yes].

Justification: The theorem statements appear in the main text; the extended version (arXiv:2607.02137) provides the expanded ART–ART-RL equivalence argument.

Guidelines:

- The answer [N/A] means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes].

Justification: The experiments specify the datasets, pipelines, baselines, NFE accounting, and fixed components; additional training and implementation details are in the extended version (arXiv:2607.02137).

Guidelines:

- The answer [N/A] means that the paper does not include experiments.
- If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes].

   Justification: The separately prepared supplementary code archive provides the implementation and instructions.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes].

   Justification: The experimental section and the extended version (arXiv:2607.02137) describe the pipelines, schedules, solvers, training iterations, optimizer, and distillation procedure.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes].

   Justification: The paper reports a three-seed CIFAR–10 evaluation at the largest NFE and confidence/shaded bands for the one-dimensional learned schedule. The main image comparisons are also evaluated across multiple datasets and NFE budgets.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes].

   Justification: The extended version (arXiv:2607.02137) reports the ART-RL schedule-training cost and explains that the distilled schedule has no extra inference-time overhead.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes].

   Justification: The work studies timestep scheduling for existing diffusion samplers using standard public benchmarks and does not involve human-subject experiments or new data collection.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A].

    Justification: The contribution is a foundational sampling-schedule method and does not introduce a new deployed generative model, dataset, or application.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

    Answer: [N/A].

    Justification: The paper does not release a new high-risk pretrained generative model or dataset.

    Guidelines:

    - The answer [N/A] means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes].

    Justification: The datasets, EDM pipeline, and baseline methods are cited to their original sources; the work uses them for standard academic benchmarking.

    Guidelines:

    - The answer [N/A] means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.

- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [Yes].

    Justification: The separately prepared supplementary code archive documents the schedule-learning and evaluation pipeline.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A].

    Justification: The paper does not involve crowdsourcing or human-subject experiments.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A].

    Justification: The paper does not involve human-subject experiments.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.

- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A].

    Justification: LLMs are not an important, original, or non-standard component of the proposed method or experiments.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.