

Diffusion Tree Sampling: Scalable Inference-time alignment of diffusion models

Vineet Jain, Kusha Sareen, Mohammad Pedramfar, Siamak Ravanbakhsh

Problem setting

Problem setting

Goal: Sample from the *reward-tilted* distribution

$$\pi^*(\mathbf{x}) = \frac{1}{Z} p_{\theta}(\mathbf{x}) \exp(\lambda r(\mathbf{x}))$$

Problem setting

Goal: Sample from the *reward-tilted* distribution

$$\pi^*(\mathbf{x}) = \frac{1}{Z} p_{\theta}(\mathbf{x}) \exp(\lambda r(\mathbf{x}))$$

Unseen or varying rewards necessitate alignment at **inference-time**.

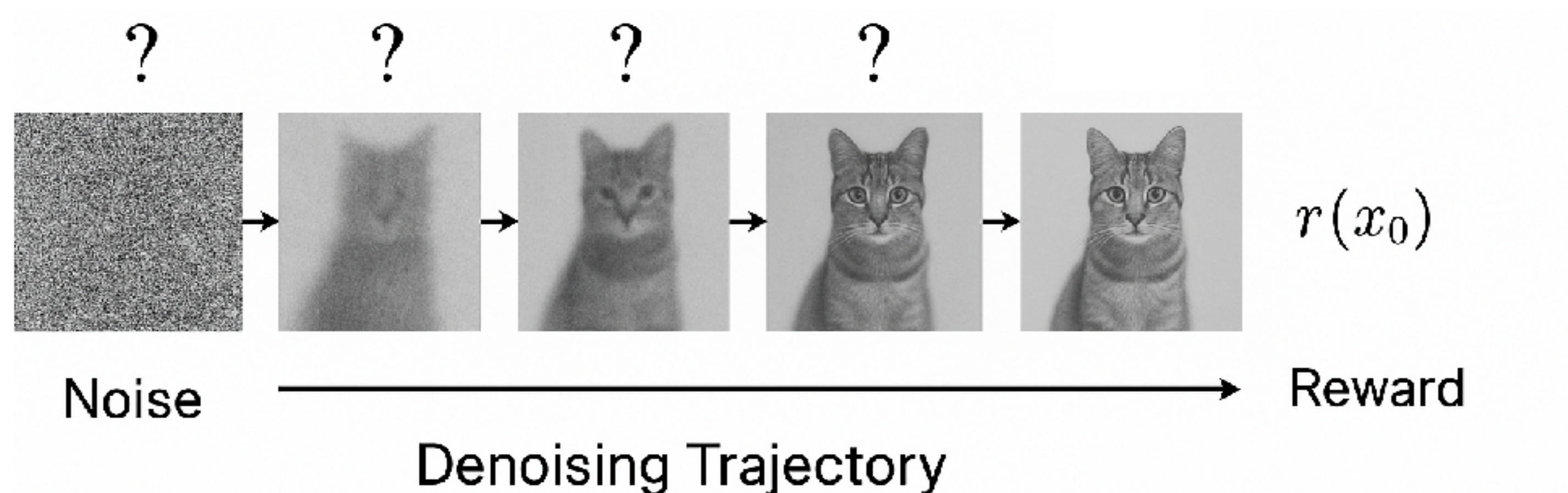
Problem setting

Goal: Sample from the *reward-tilted* distribution

$$\pi^*(\mathbf{x}) = \frac{1}{Z} p_{\theta}(\mathbf{x}) \exp(\lambda r(\mathbf{x}))$$

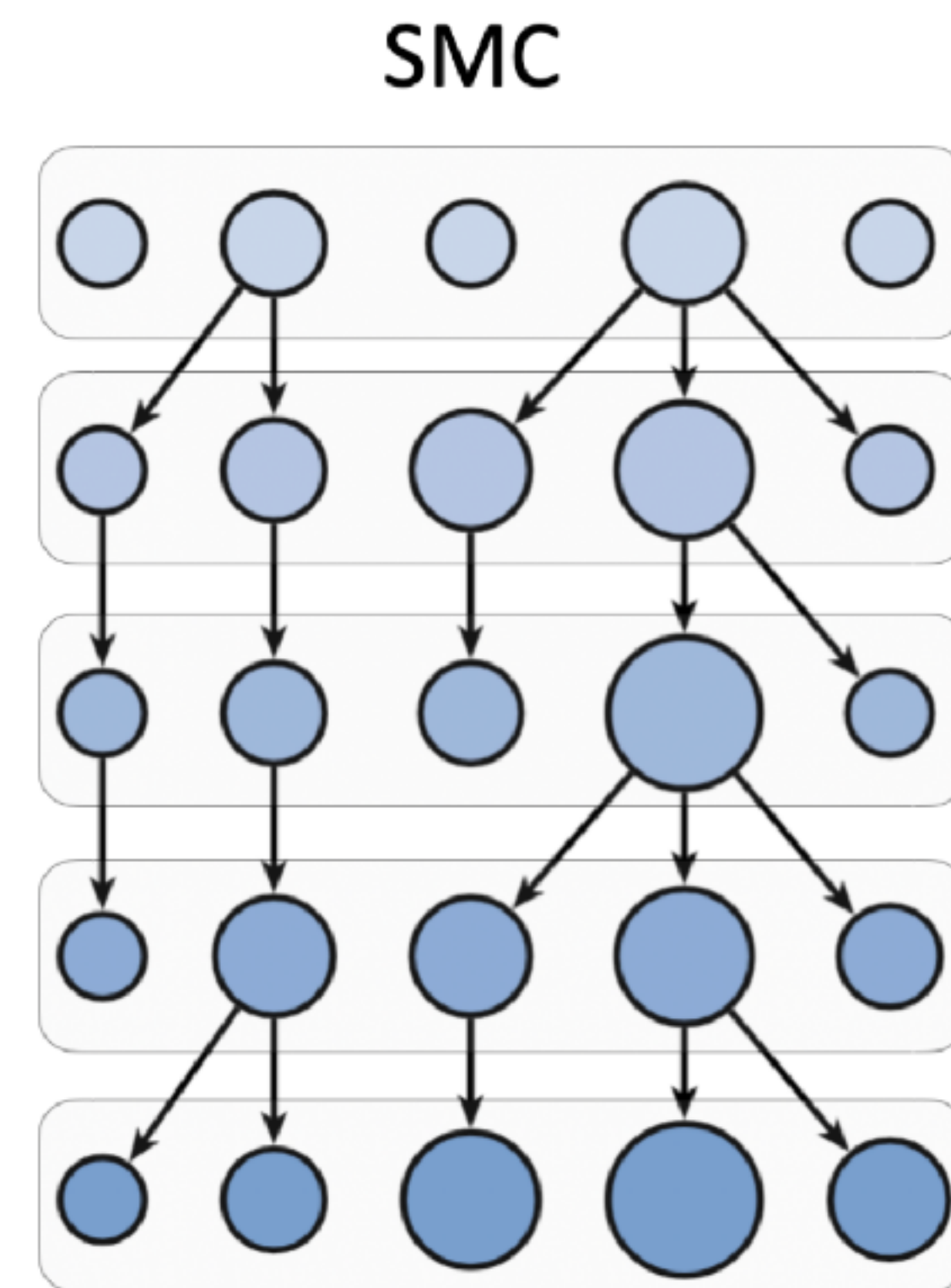
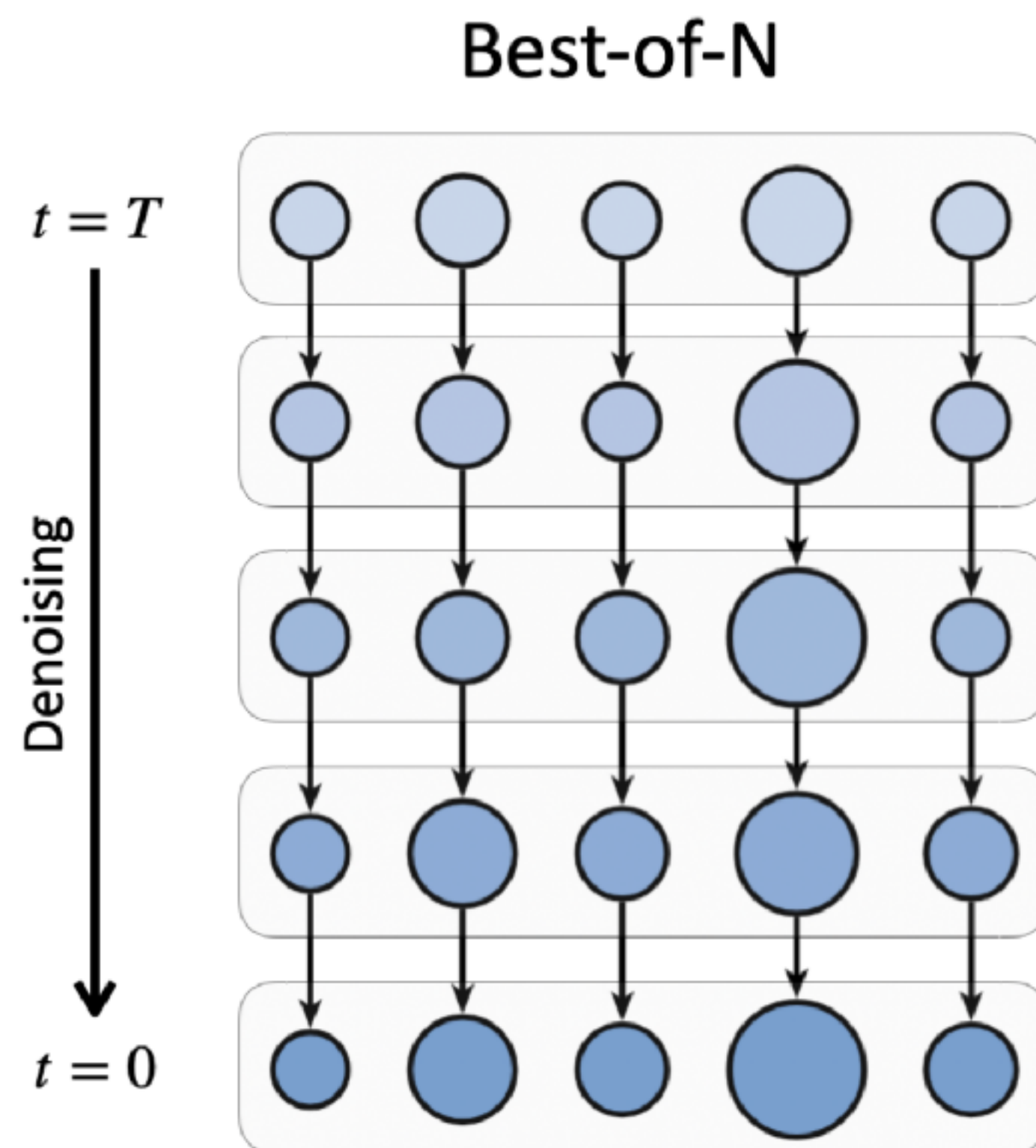
Unseen or varying rewards necessitate alignment at **inference-time**.

But credit assignment is hard!



Existing methods

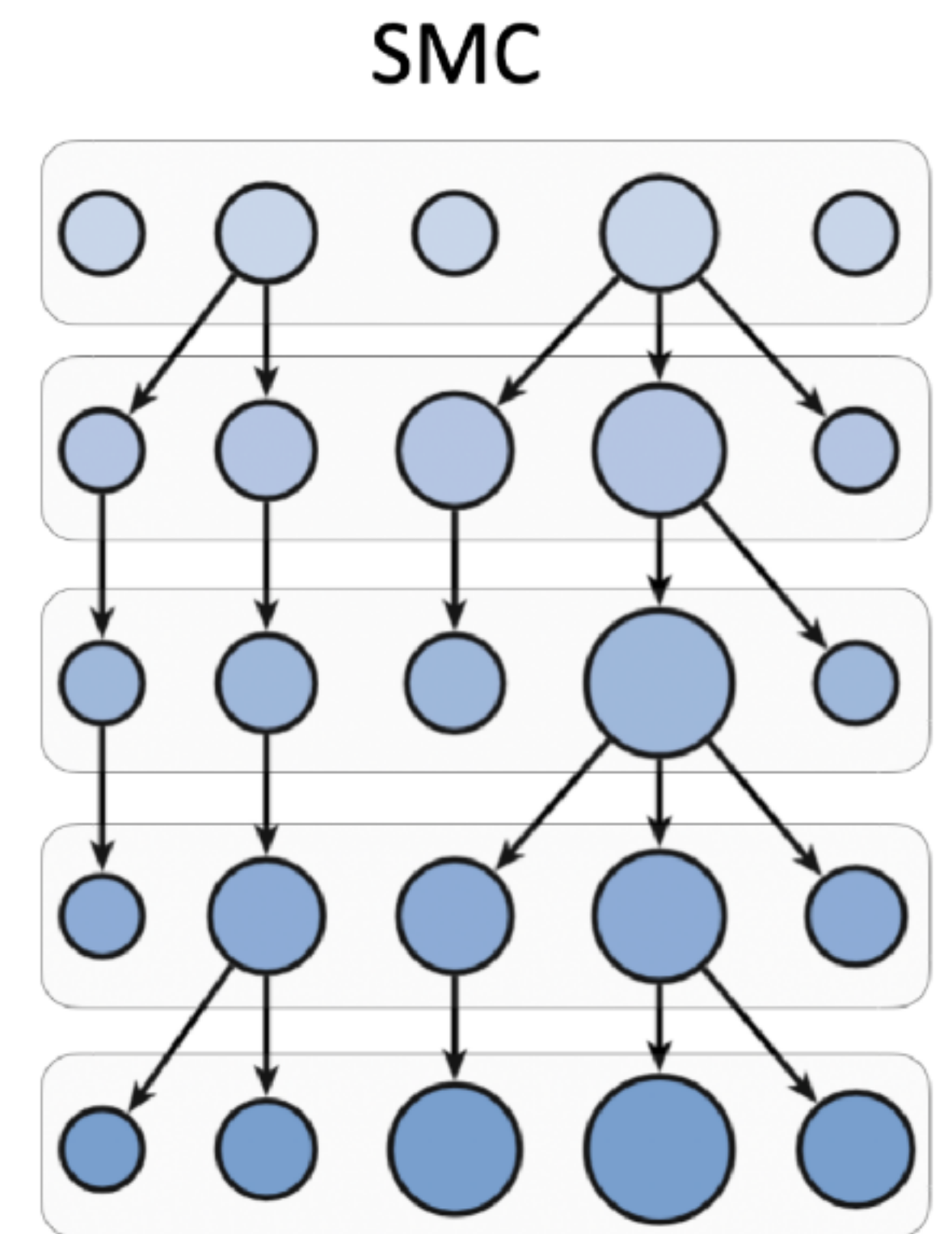
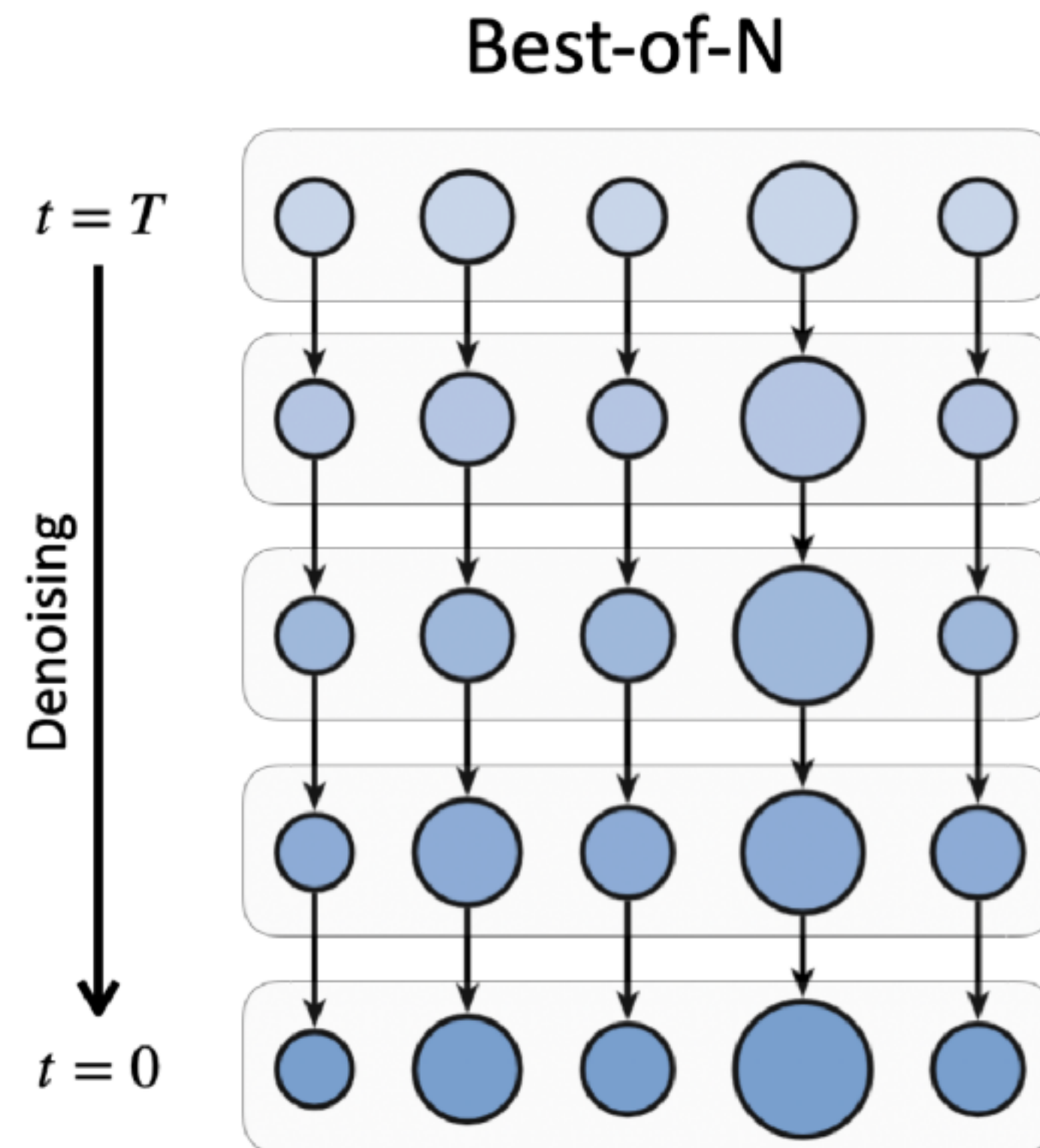
Existing methods



Existing methods

✗ *Biased* approximations of the value function at noisy states

✗ Only scale *in parallel* by increasing particle count



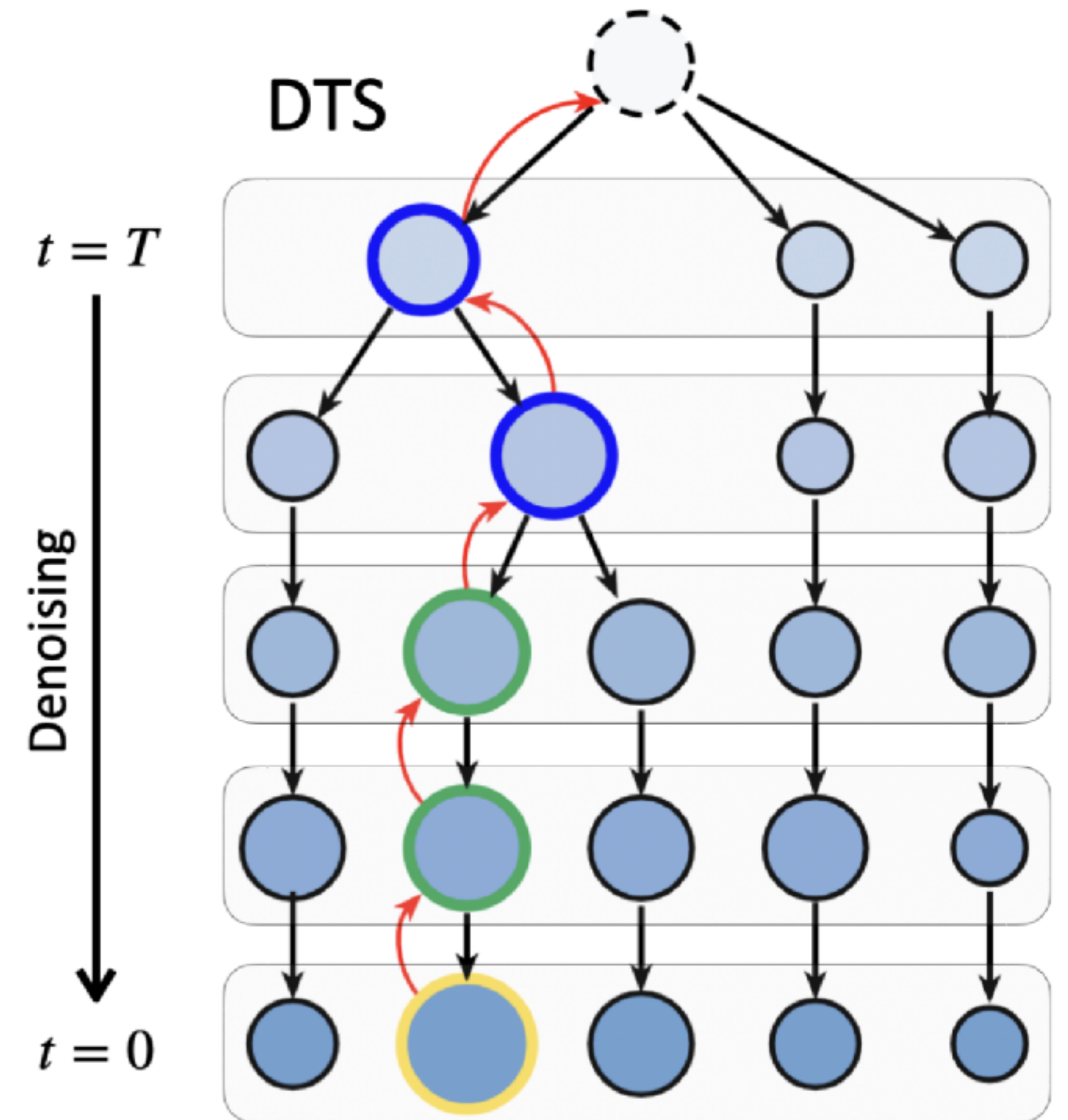
Diffusion Tree Sampling/Search

Diffusion Tree Sampling/Search

Core idea: Denoising as a tree + value backups

Diffusion Tree Sampling/Search

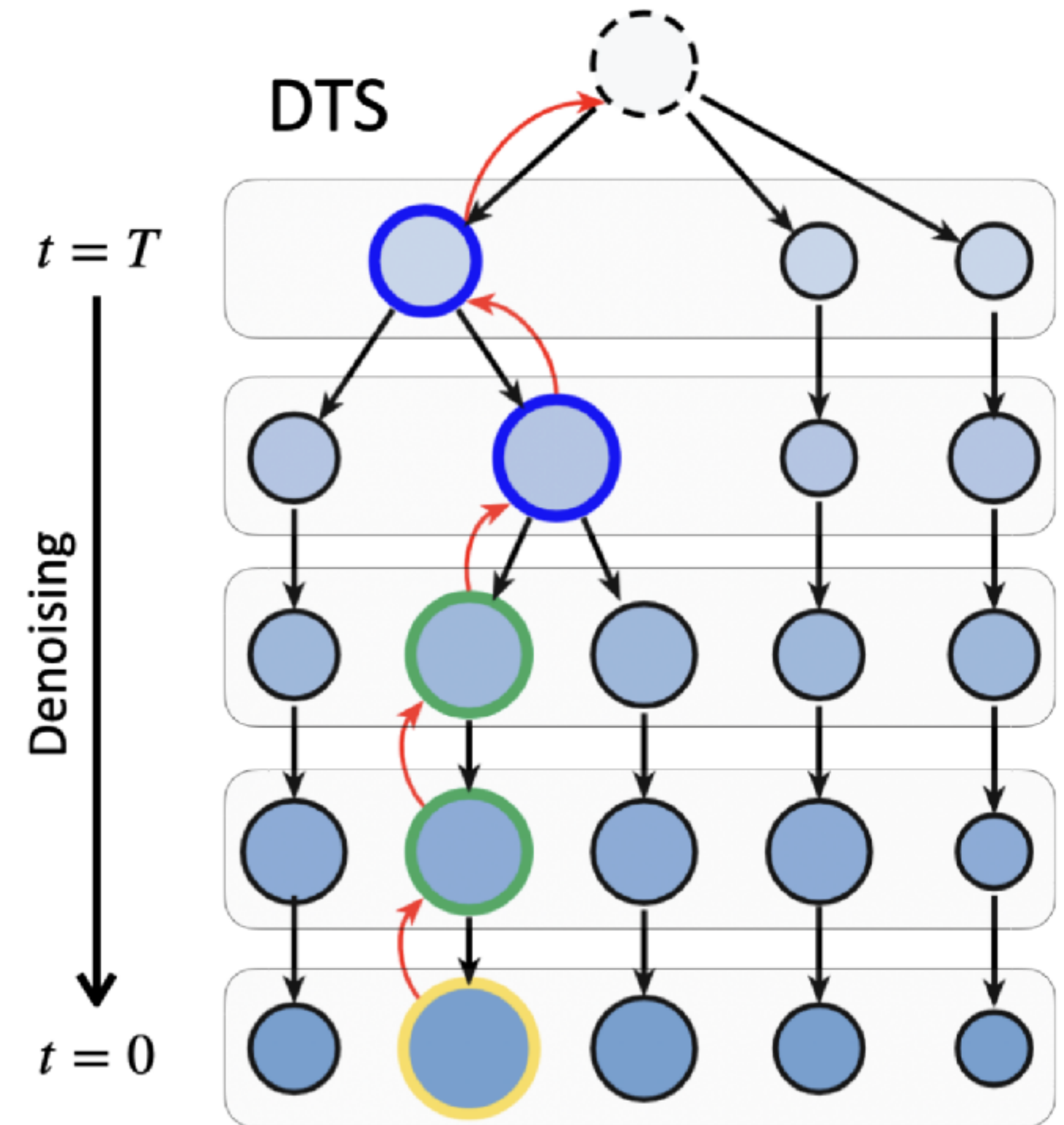
Core idea: Denoising as a tree + value backups



Diffusion Tree Sampling/Search

Core idea: Denoising as a tree + value backups

Treat each noisy x_t as node; denoising as edges
Maintain value estimates $\hat{v}(x_t)$ at each node



Diffusion Tree Sampling/Search

Core idea: Denoising as a tree + value backups

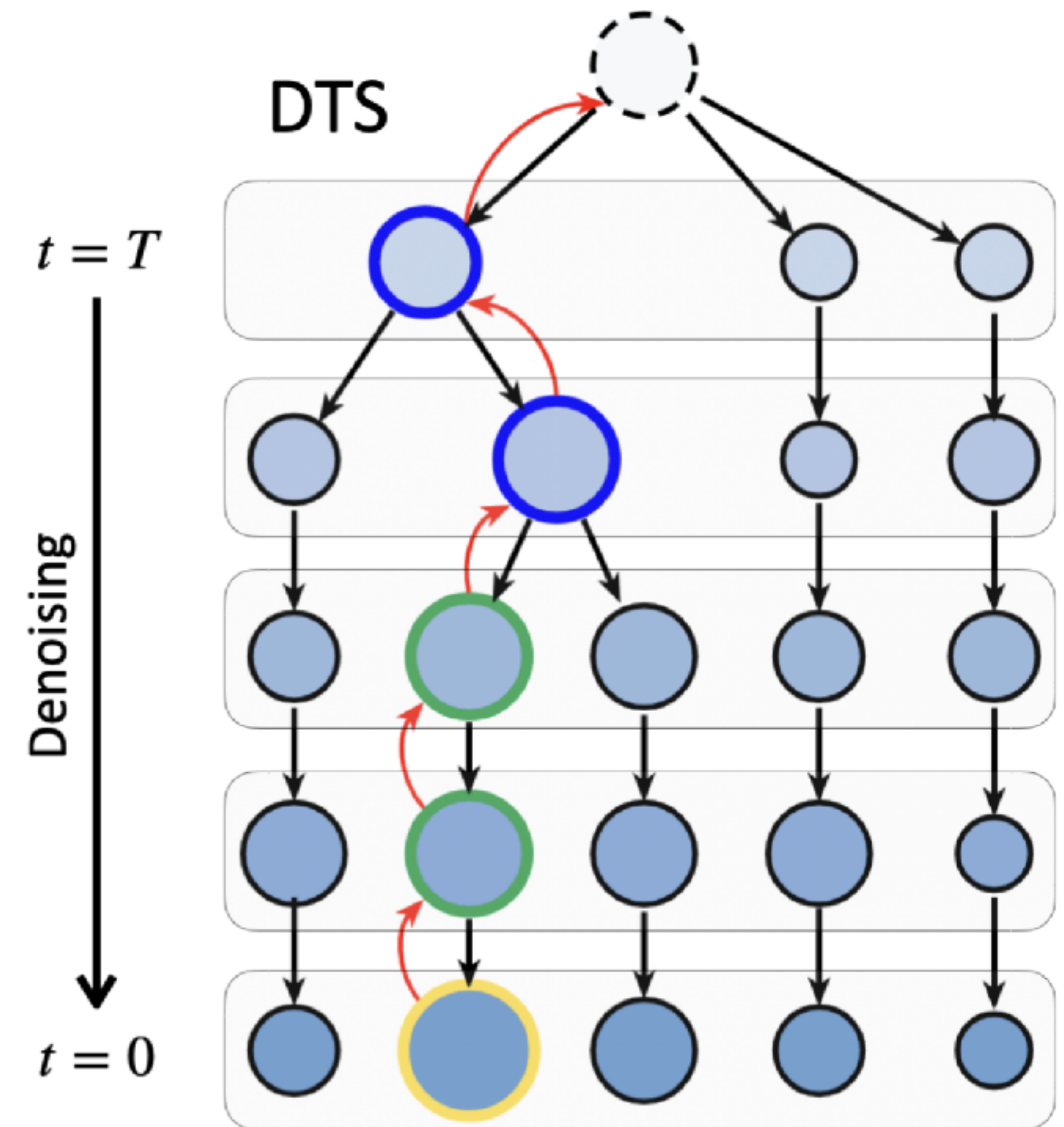
Treat each noisy x_t as node; denoising as edges

Maintain value estimates $\hat{v}(x_t)$ at each node

Selection:

DTS sample $\propto \exp(\hat{v}(x_{t-1})) \rightarrow$ asymptotically exact for π^*

DTS* greedy max $\hat{v}(x_{t-1}) \rightarrow$ efficient high-reward search



Diffusion Tree Sampling/Search

Core idea: Denoising as a tree + value backups

Treat each noisy x_t as node; denoising as edges

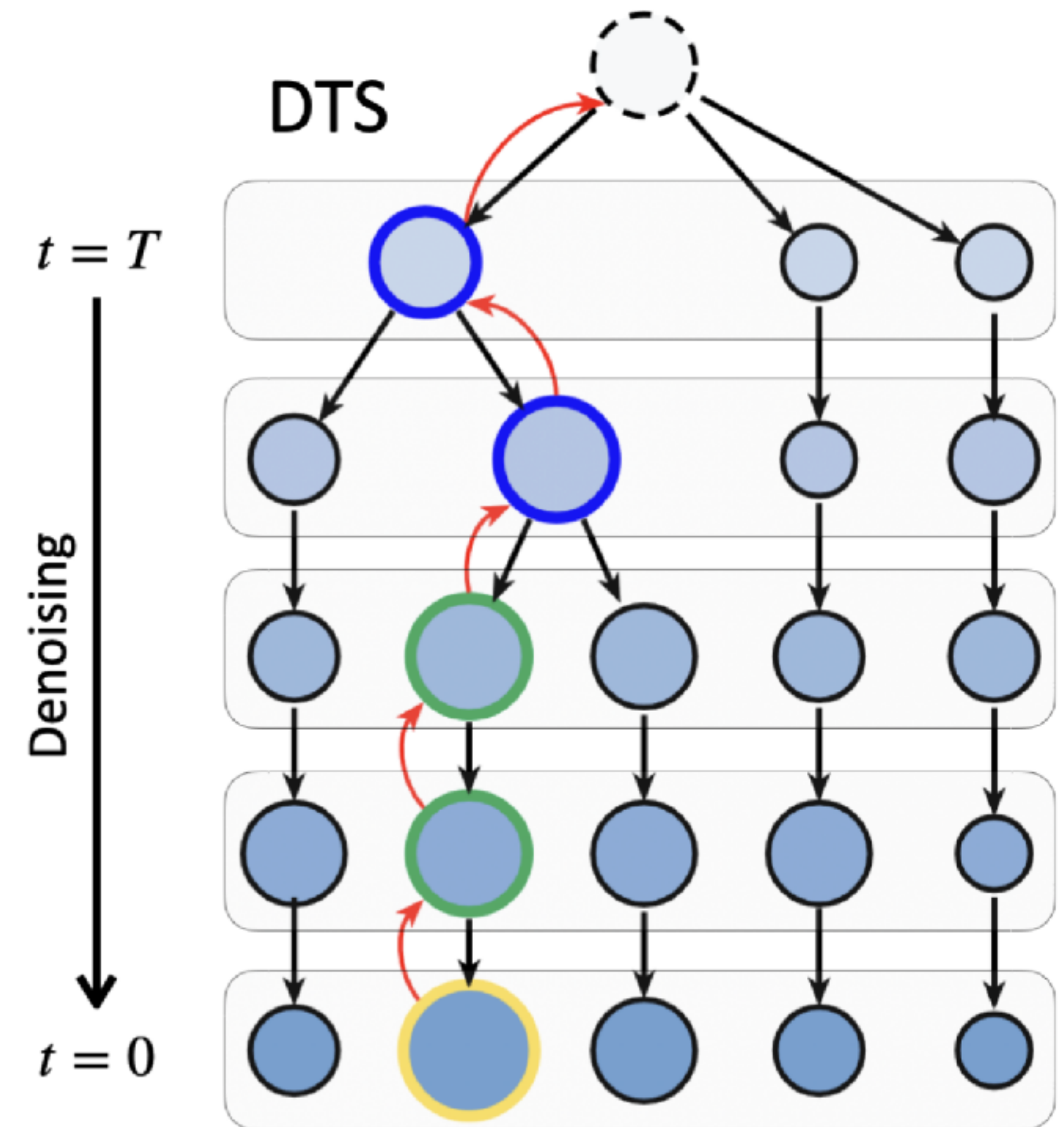
Maintain value estimates $\hat{v}(x_t)$ at each node

Selection:

DTS sample $\propto \exp(\hat{v}(x_{t-1})) \rightarrow$ asymptotically exact for π^*

DTS* greedy max $\hat{v}(x_{t-1}) \rightarrow$ efficient high-reward search

Rollout: Use base model to create child x_{t-1} from node x_t



Diffusion Tree Sampling/Search

Core idea: Denoising as a tree + value backups

Treat each noisy x_t as node; denoising as edges

Maintain value estimates $\hat{v}(x_t)$ at each node

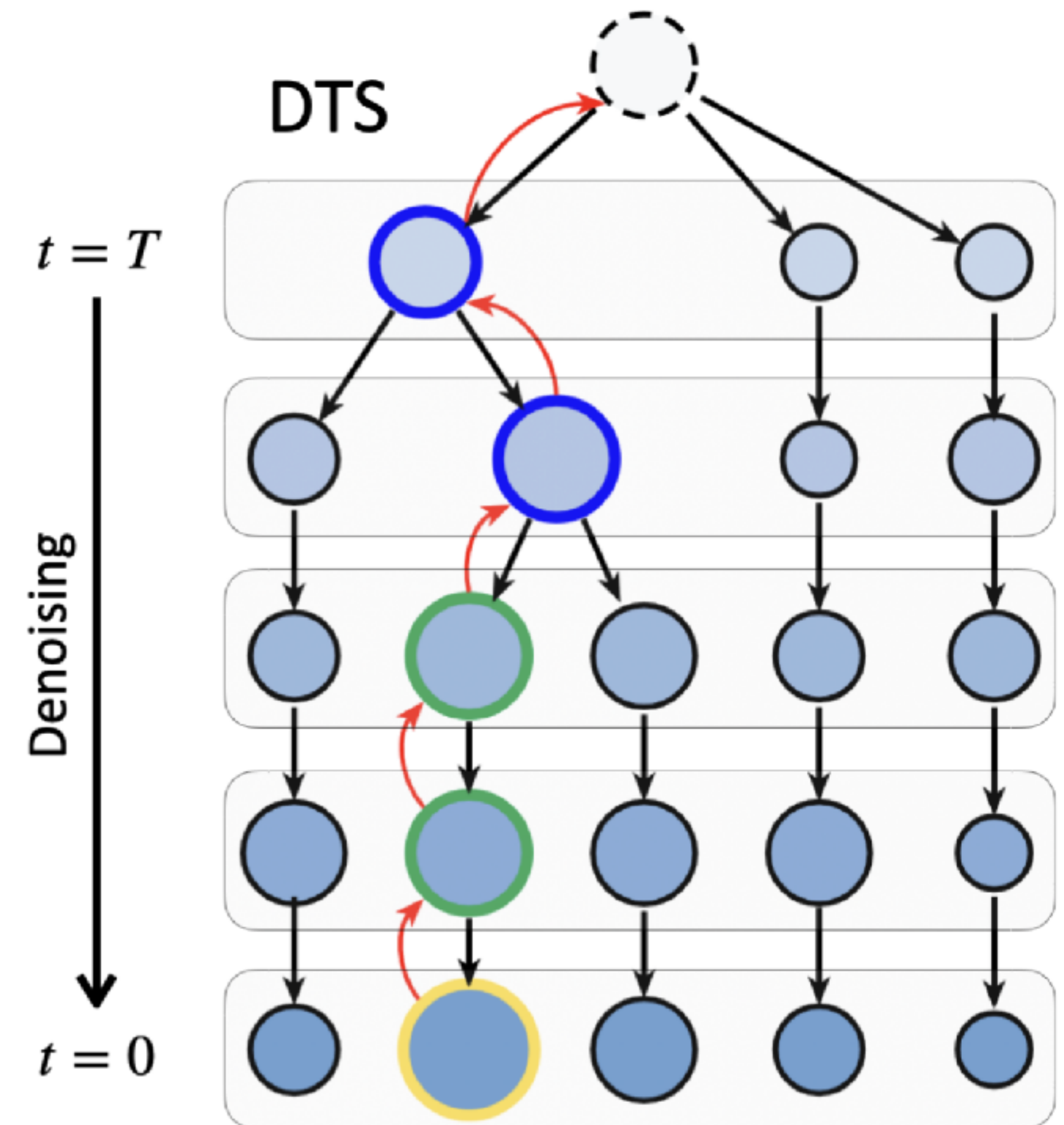
Selection:

DTS sample $\propto \exp(\hat{v}(x_{t-1})) \rightarrow$ asymptotically exact for π^*

DTS* greedy max $\hat{v}(x_{t-1}) \rightarrow$ efficient high-reward search

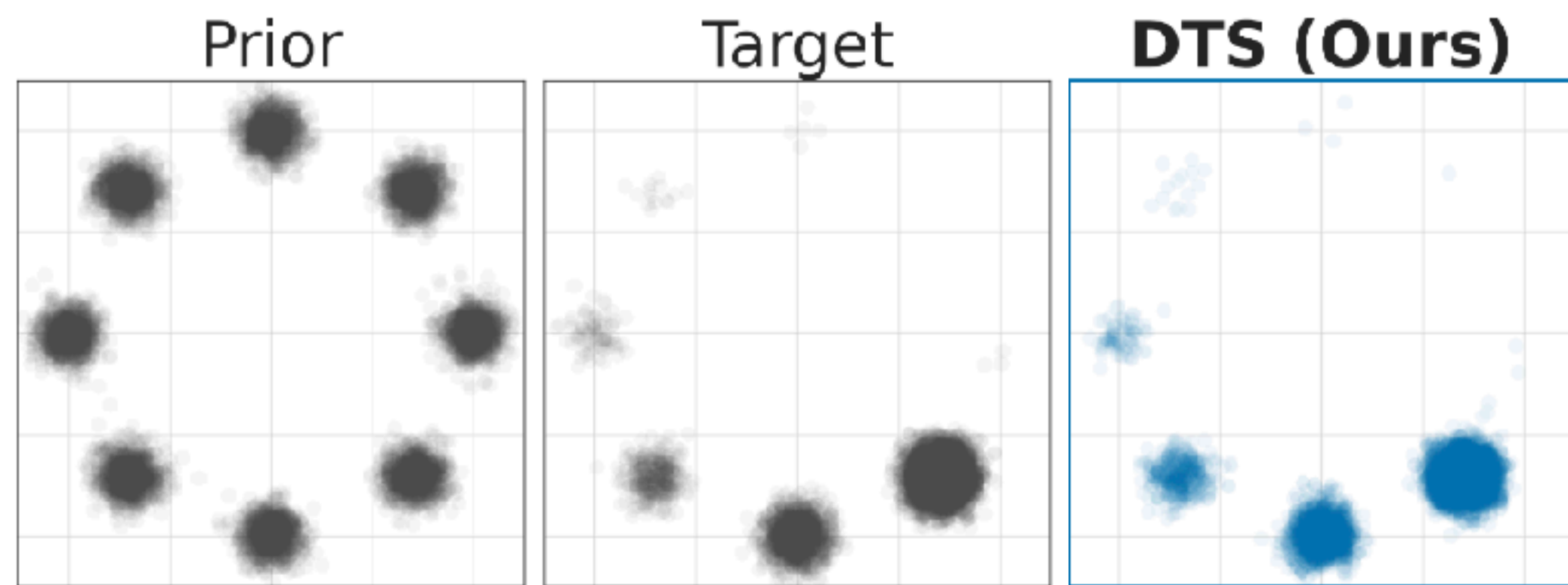
Rollout: Use base model to create child x_{t-1} from node x_t

Backup: Evaluate $r(x_0)$, backup soft (log-sum-exp) values



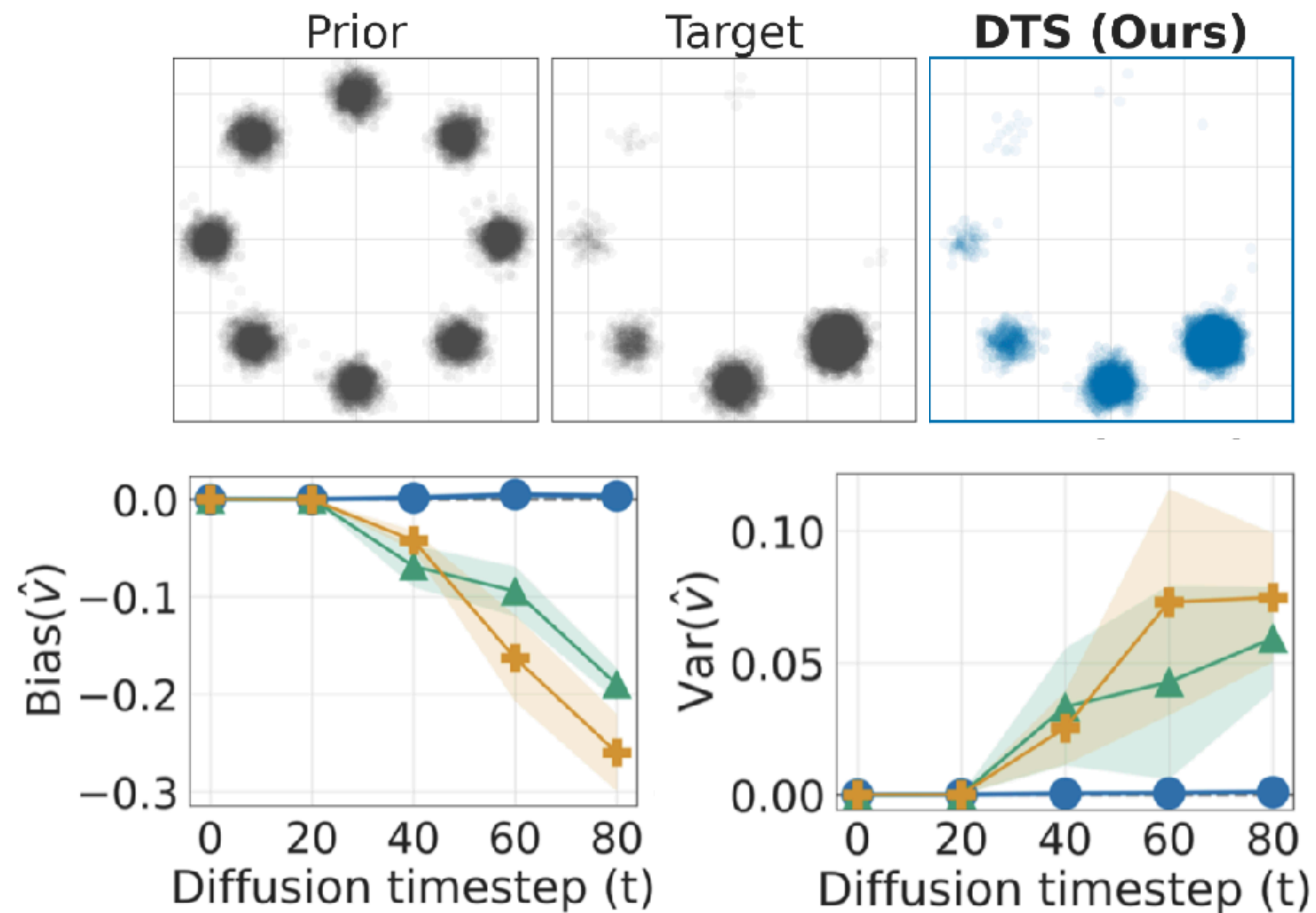
Why does DTS help?

Why does DTS help?



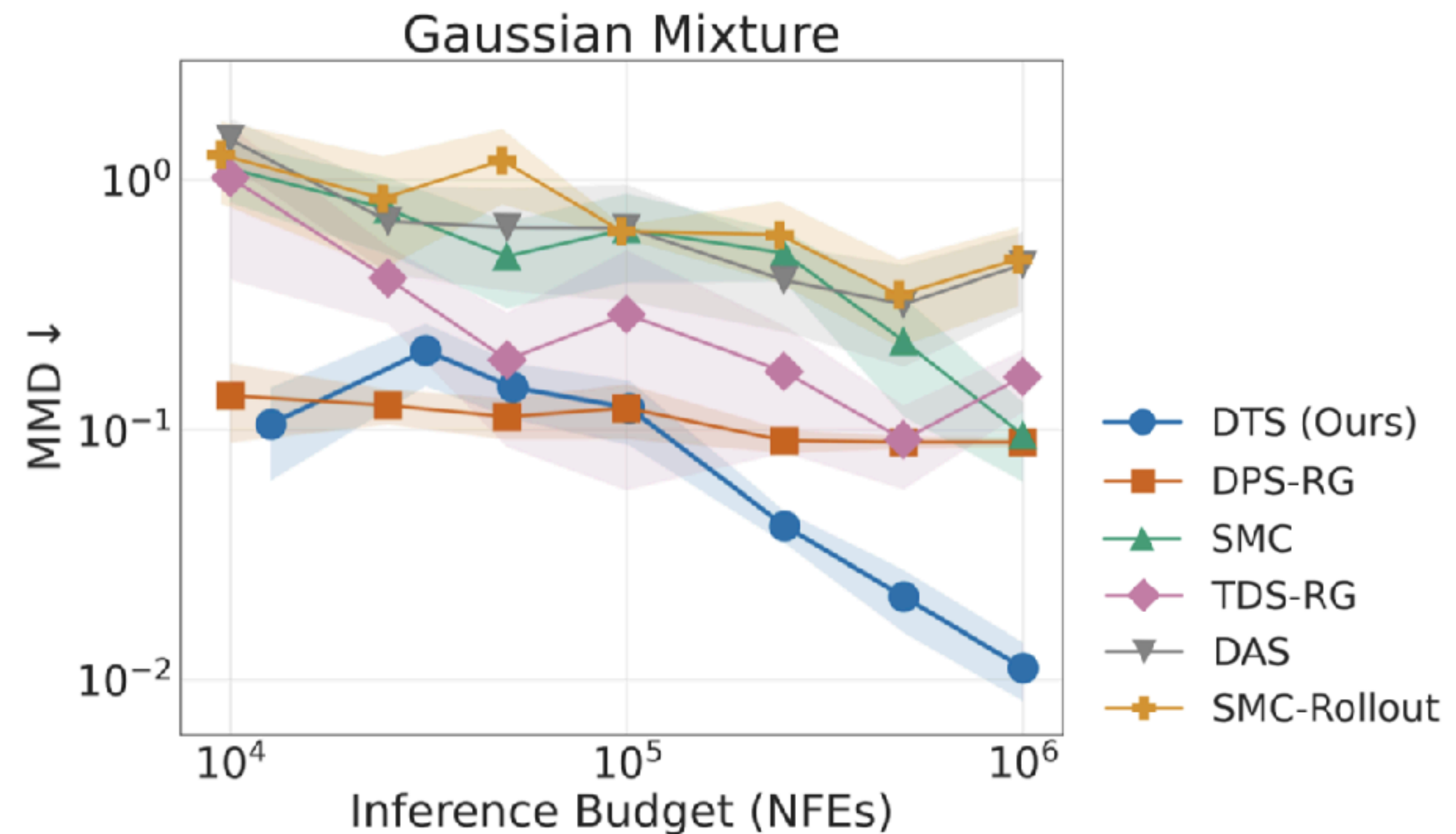
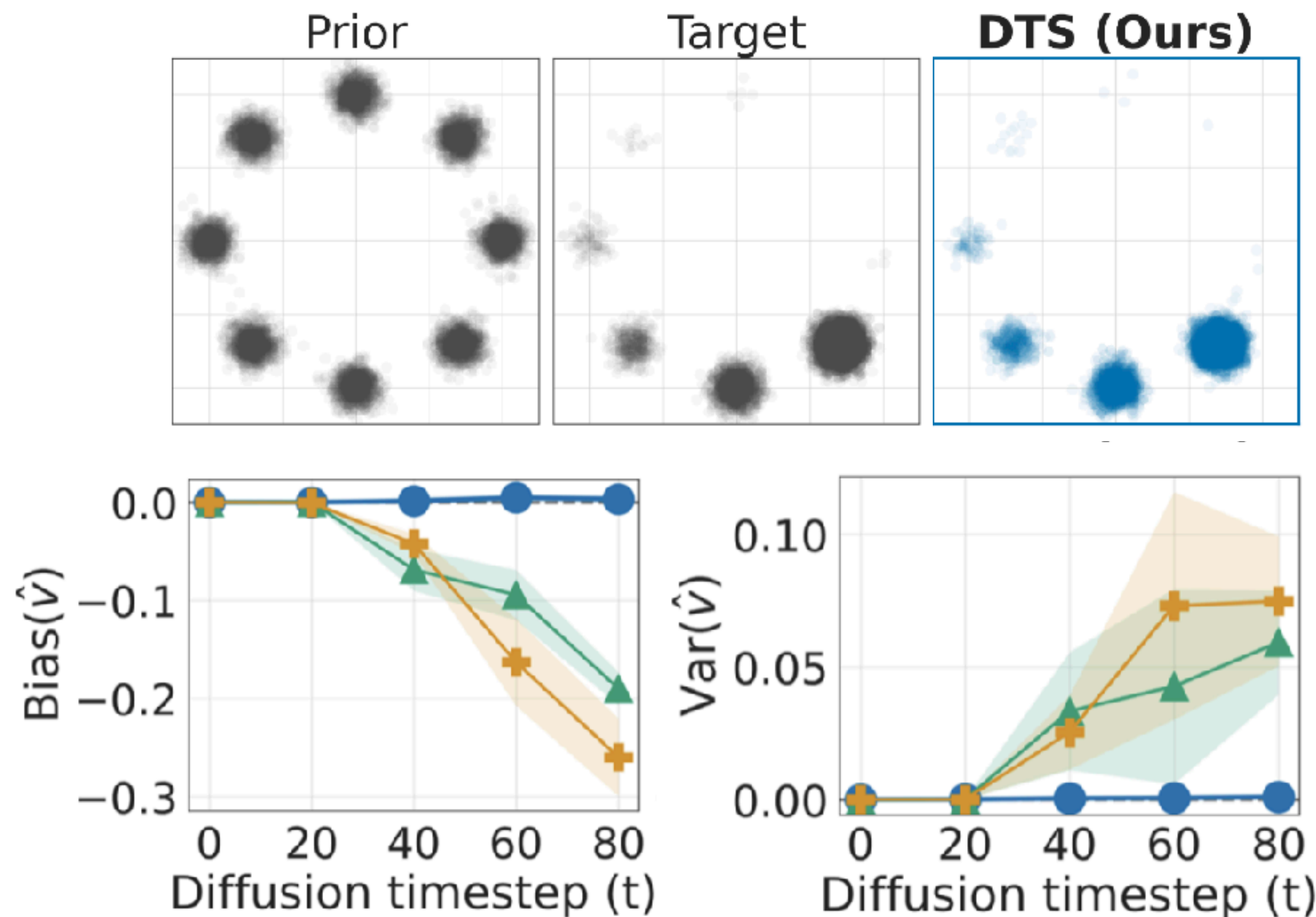
Why does DTS help?

✓ Low-noise rewards back up to high-noise states → **better credit assignment**



Why does DTS help?

- ✓ Low-noise rewards back up to high-noise states → **better credit assignment**
- ✓ Tree reuses estimates → more compute **sequentially** improves values & samples



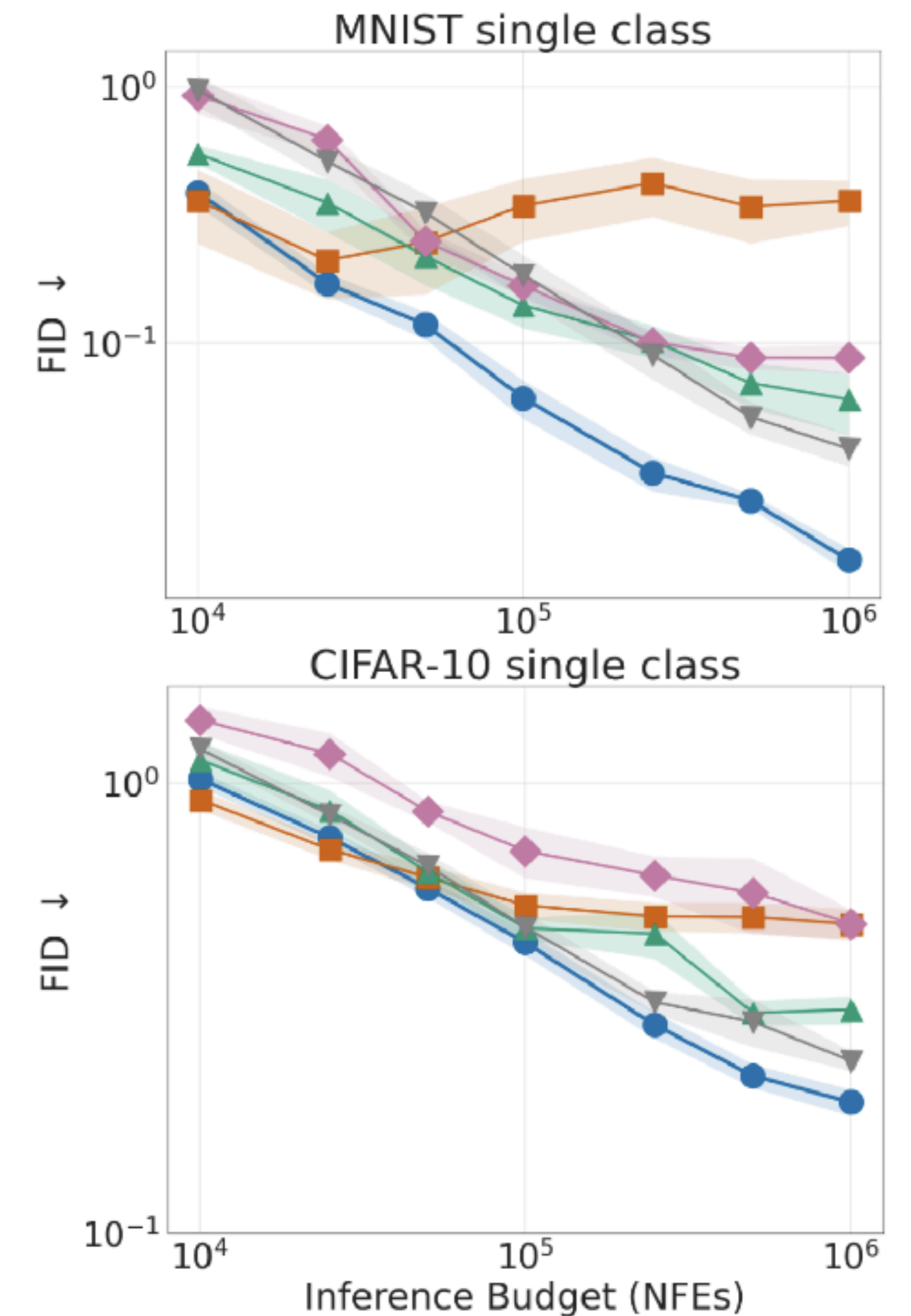
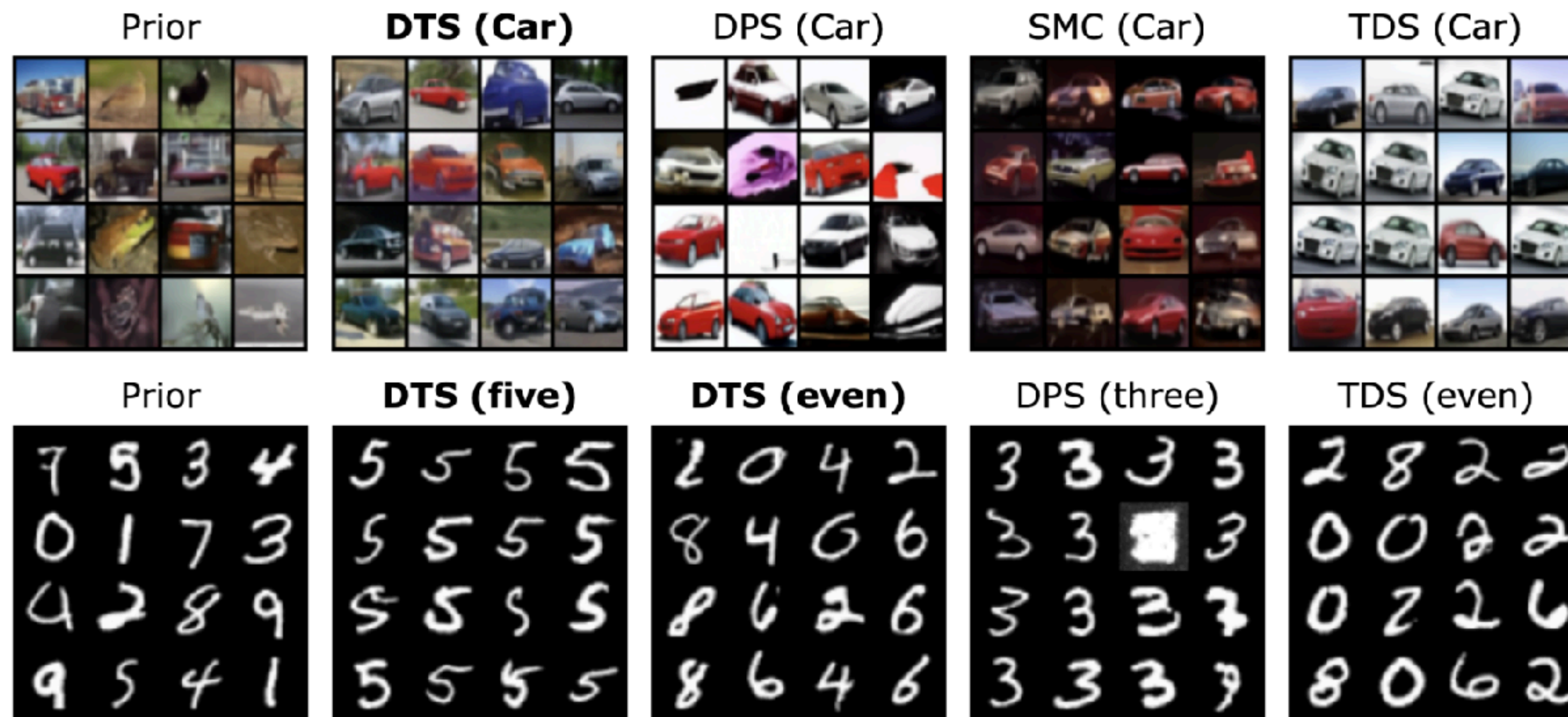
Class-conditional sampling

Class-conditional sampling

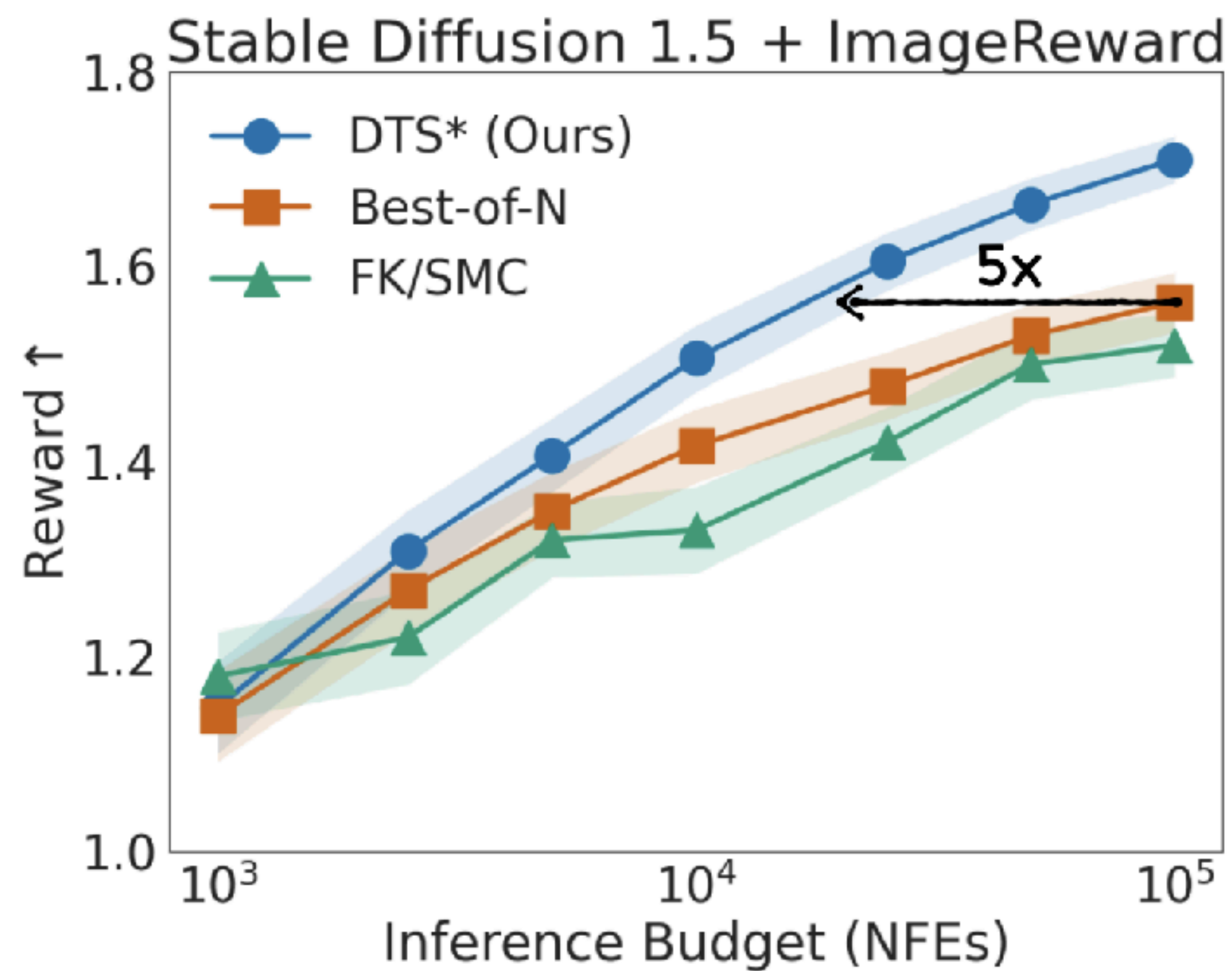
Target distribution becomes $p(x \mid \text{class}) \propto p_{\theta}(x) p(\text{class} \mid x)$

Class-conditional sampling

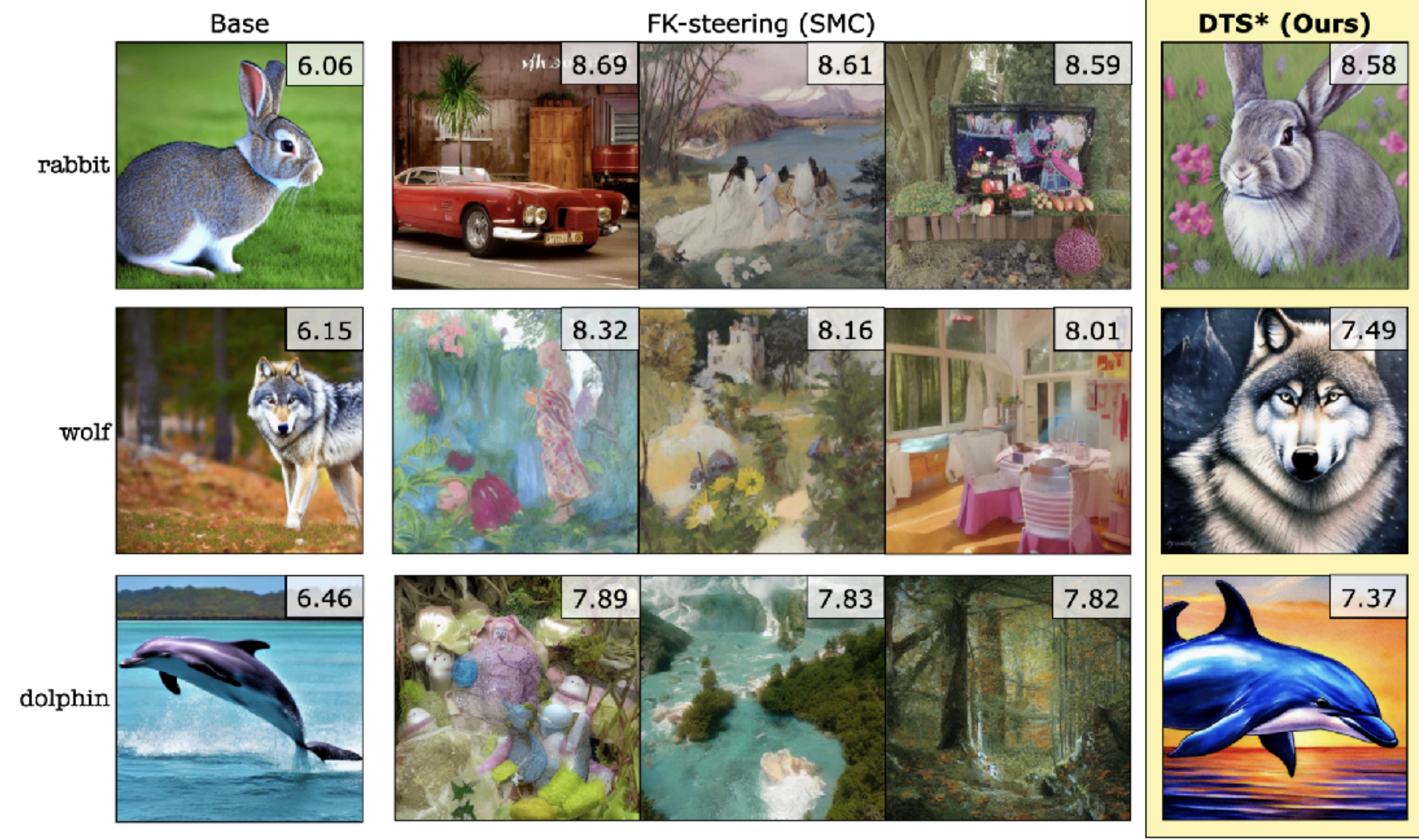
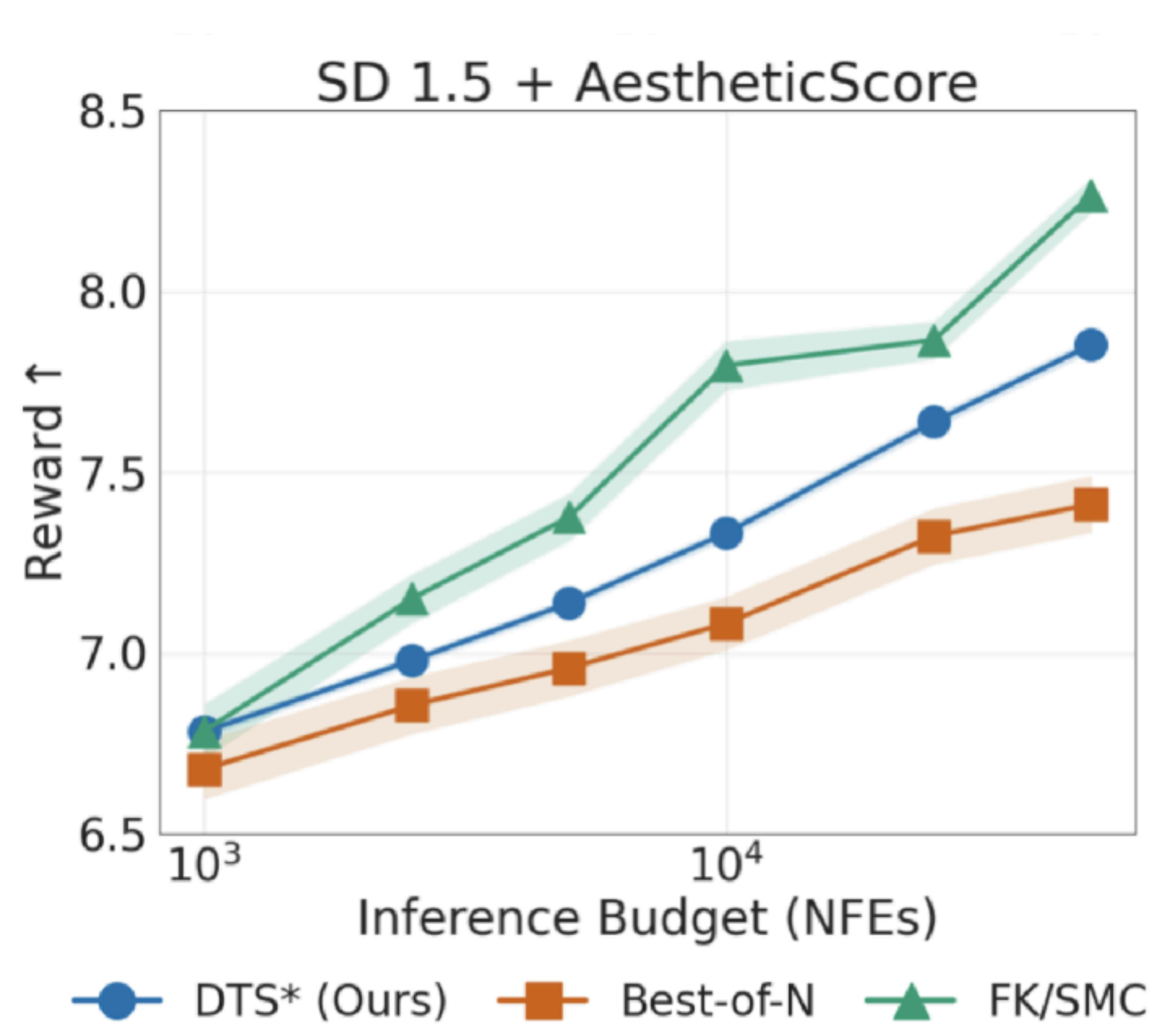
Target distribution becomes $p(x | \text{class}) \propto p_{\theta}(x) p(\text{class} | x)$



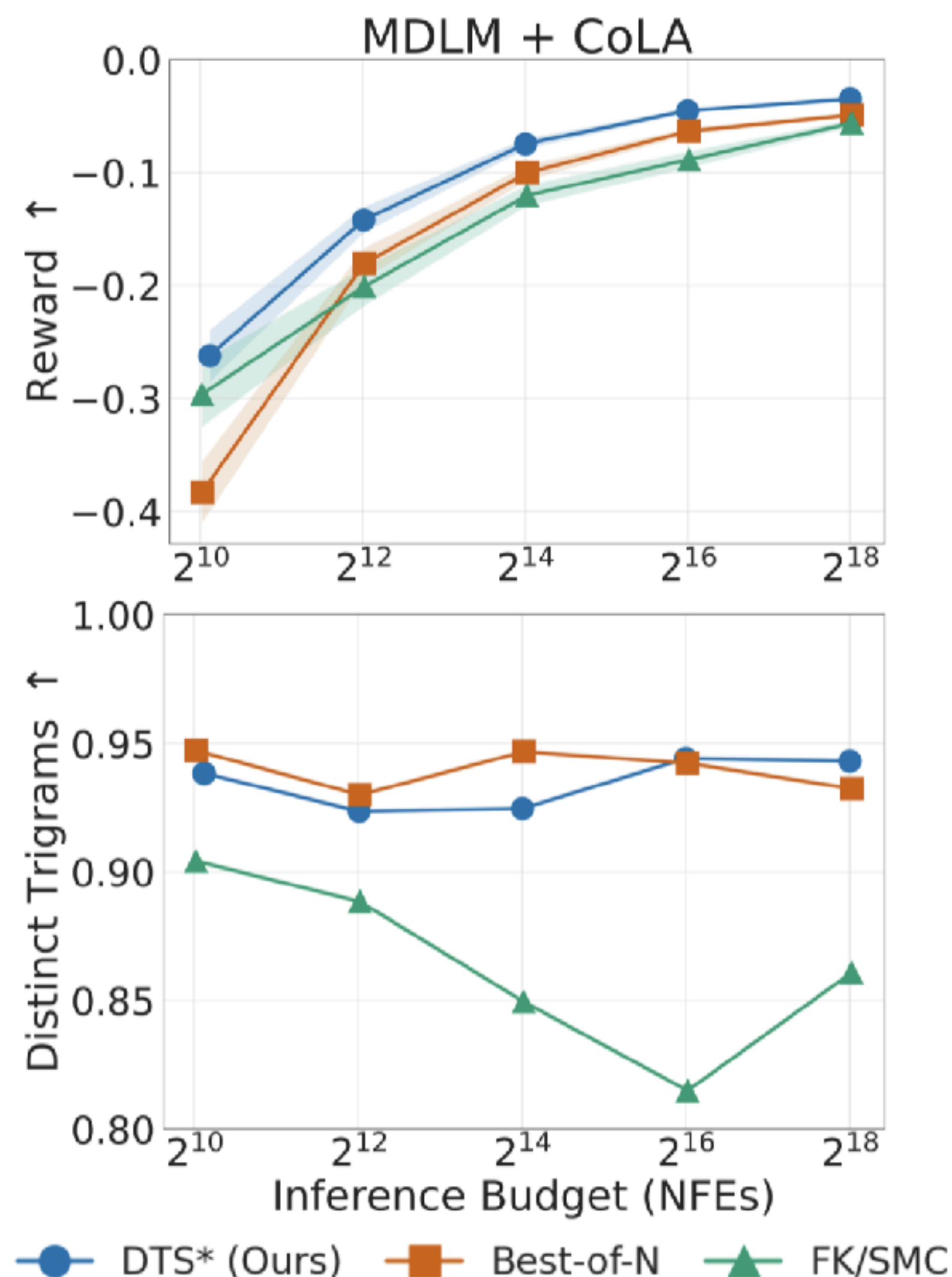
Text-to-image models



Text-to-image models



Diffusion language models



Base

Reward = -1.454

The city centre workers are blamed criticised also for border management in recent years. It has involved taking millions of dollars (nominally Medicare Depenances) out of the US Treasury for months. The package was reduced to \$1.50 million, at the...

FK/SMC

Reward = -0.059

The city is defended by the Royal Jamaican Squadron. The city is defended by the English Defence Squadron. The city is defended by the Continental Guards. The navy is defended by the Catalanian Squadron. The city is defended by the...

DTS* (Ours)

Reward = -0.032

The city's Public Safety Committee has also promised to consider ways to discourage residents from buying private homes of their own. But tax-benefit analysts oppose such a step because they fear it would dramatically reduce the local tax base. Taxpayers...

Thank you!

Website



arXiv

