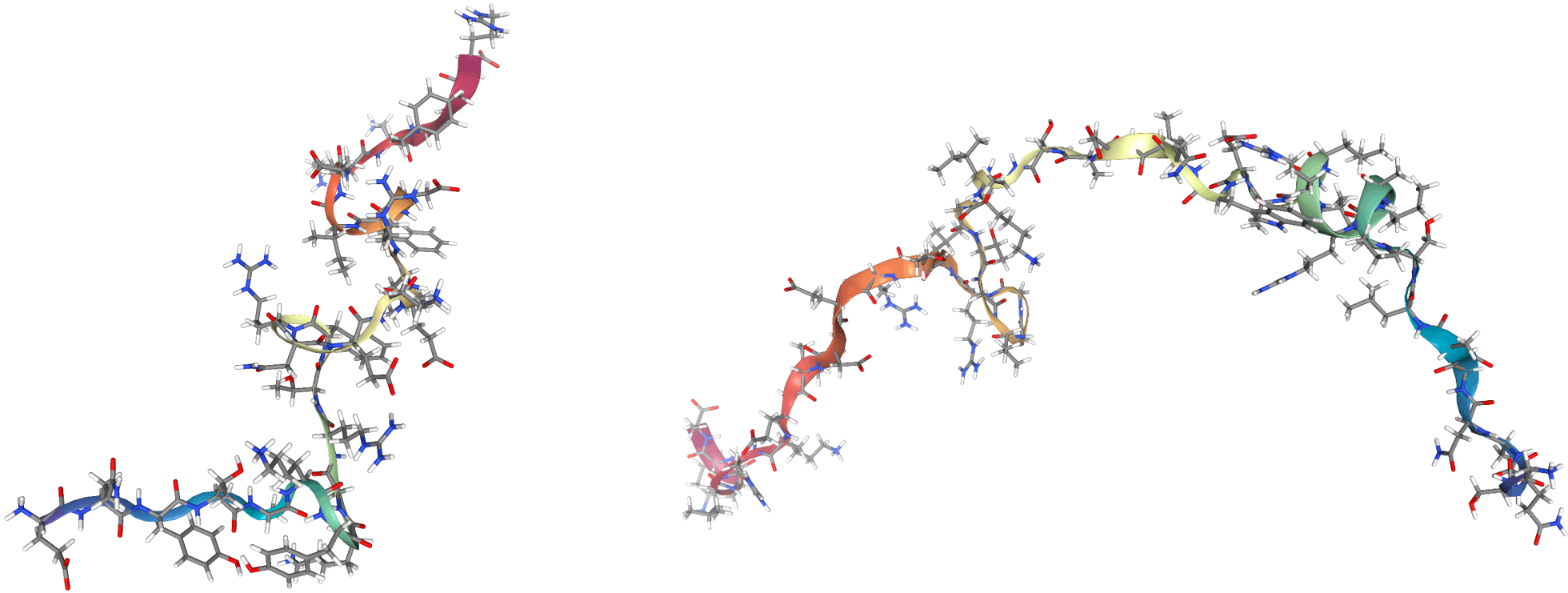
Transition Path Sampling with improved Off-policy training of Diffusion Path Samplers

Kiyoung Seong, Seonghyun Park, Seonghwan Kim
Wooyoun Kim, Sungsoo Ahn

Goal

- Our goal is to sample rare transitions without trapped in meta-stable states.



Fast folding proteins (ours)

Molecular dynamics

- Molecular dynamics (MD) describes the motion of molecules as the following SDE:

$$\mathrm{d}\boldsymbol{R}_t = \boldsymbol{V}_t \mathrm{d}t, \qquad \mathrm{d}\boldsymbol{V}_t = \frac{-\nabla U(\boldsymbol{R}_t)}{\boldsymbol{m}}\mathrm{d}t - \gamma \boldsymbol{V}_t \mathrm{d}t + \sqrt{\frac{2\gamma k_B \lambda}{\boldsymbol{m}}}\mathrm{d}\boldsymbol{W}_t$$

Newton Dynamics friction diffusion

Langevin dynamics

$\boldsymbol{R}_t \in \mathbb{R}^{3N}$	atom-wise positions
$\boldsymbol{V}_t \in \mathbb{R}^{3N}$	atom-wise velocities
$U: \mathbb{R}^{3N} \to \mathbb{R}$	potential energy
$\boldsymbol{m}$	atom-wise mass
γ	friction coefficient
k_B	Boltzmann constant
λ	Temperature
$\mathrm{d}\boldsymbol{w}$	Brownian motion.

Molecular dynamics

- Molecular dynamics (MD) describes the motion of molecules as the following SDE:

$$\mathrm{d}\boldsymbol{R}_t = \boldsymbol{V}_t \mathrm{d}t, \qquad \mathrm{d}\boldsymbol{V}_t = \frac{-\nabla U(\boldsymbol{R}_t)}{\boldsymbol{m}} \mathrm{d}t - \gamma \boldsymbol{V}_t \mathrm{d}t + \sqrt{\frac{2\gamma k_B \lambda}{\boldsymbol{m}}} \mathrm{d}\boldsymbol{W}_t$$

Langevin dynamics

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{u}(\boldsymbol{X}_t)\mathrm{d}t + \Sigma \mathrm{d}\boldsymbol{W}_t$$

$\boldsymbol{R}_t \in \mathbb{R}^{3N}$	atom-wise positions
$\boldsymbol{V}_t \in \mathbb{R}^{3N}$	atom-wise velocities
$U: \mathbb{R}^{3N} \to \mathbb{R}$	potential energy
$\boldsymbol{m}$	atom-wise mass
γ	friction coefficient
k_B	Boltzmann constant
λ	Temperature
$\mathrm{d}\boldsymbol{w}$	Brownian motion.

Amortized transition path sampling

- Transition path $\{X_t\}_{0 \le t \le T}$ is a sample from endpoint conditioned Langevin dynamics

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{u}(\boldsymbol{X}_t)\mathrm{d}t + \Sigma \mathrm{d}\boldsymbol{W}_t, \qquad \boldsymbol{R}_0 \in \mathcal{A}, \boldsymbol{R}_T \in \mathcal{B}$$

- To amortize inference with endpoint condition, we consider controlled dynamics

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{u}(\boldsymbol{X}_t)\mathrm{d}t + \Sigma \boldsymbol{v}_\theta(\boldsymbol{X}_t)\mathrm{d}t + \Sigma \mathrm{d}\boldsymbol{W}_t, \qquad \boldsymbol{R}_0 \in \mathcal{A}$$

$$\boldsymbol{v}_\theta(\boldsymbol{X}_t) = \Sigma^{-1}\left(\mathbf{0}, \frac{\boldsymbol{b}_\theta(\boldsymbol{X}_t)}{\boldsymbol{m}}\right)$$

bias force

policy

[1] Holdijk, Lars, et al. (2023). "Stochastic optimal control for collective variable free sampling of molecular transition paths." Neurips

Amortized transition path sampling

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{u}(\boldsymbol{X}_t)\mathrm{d}t + \Sigma\mathrm{d}\boldsymbol{W}_t, \qquad \boldsymbol{R}_0 \in \mathcal{A}, \boldsymbol{R}_T \in \mathcal{B}$$

Endpoint conditioned MD
Induces target path measure

$$\mathbb{Q}$$

$$\approx$$

$$\mathbb{P}_{v_\theta}$$

$$\mathrm{d}\boldsymbol{X}_t = \boldsymbol{u}(\boldsymbol{X}_t)\mathrm{d}t + \Sigma\boldsymbol{v}_\theta(\boldsymbol{X}_t)\mathrm{d}t + \Sigma\mathrm{d}\boldsymbol{W}_t, \qquad \boldsymbol{R}_0 \in \mathcal{A}$$

Controlled MD
induces path measure

$$\boldsymbol{v}_\theta(\boldsymbol{X}_t) = \Sigma^{-1}\left(\mathbf{0}, \frac{\boldsymbol{b}_\theta(\boldsymbol{X}_t)}{\boldsymbol{m}}\right)$$

bias force

policy

Key ideas

(1) How to design objective function? $\mathbb{Q} \approx \mathbb{P}_{v_\theta}$

(2) How to parameterize bias force? $\boldsymbol{b}_\theta(\boldsymbol{X}_t)$

Key ideas

(1) How to design objective function?

→ log variance divergence for off-policy training

$\mathbb{Q} \approx \mathbb{P}_{v_\theta}$

$$\mathbb{V}_{\mathbb{P}}\left[\log \frac{d\mathbb{P}_{v_\theta}}{d\mathbb{Q}}\right]$$

Reference path measure

Off policy training $(\mathbb{P} \neq \mathbb{P}_{v_\theta})$ allows

1) generating paths at high temperature
2) reusing generated paths with replay buffer

Key ideas

(1) How to design objective function?

→ log variance divergence for off-policy training

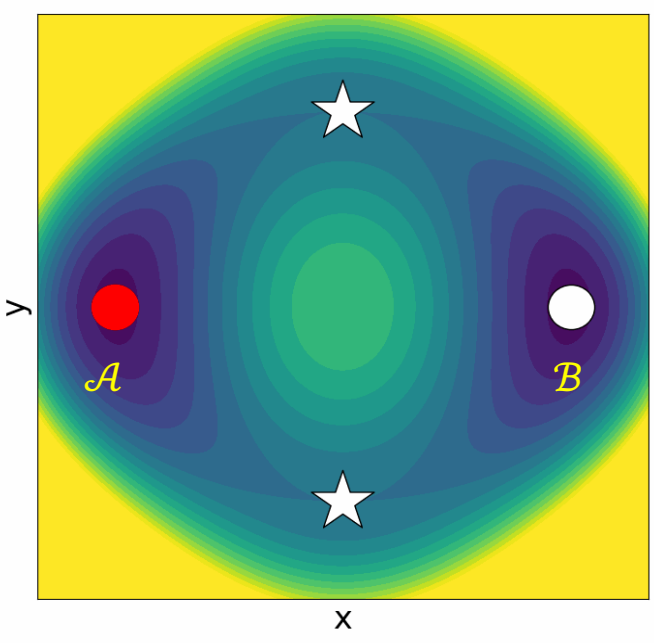
(2) How to parameterize bias force? $\boldsymbol{b}_\theta(\boldsymbol{X}_t)$

→ force $\boldsymbol{b}_\theta(\boldsymbol{X}_t) \in \mathbb{R}^{3N}$ or potential $b_\theta(\boldsymbol{X}_t) \in \mathbb{R}$ for small systems

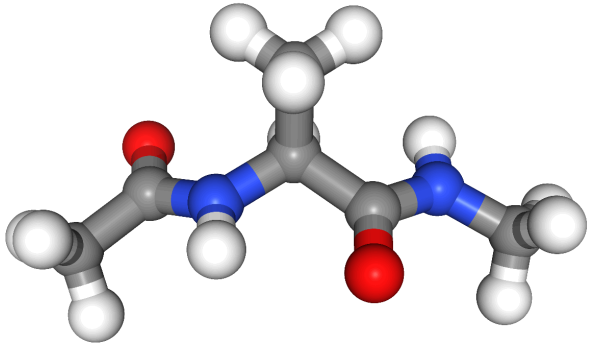
→ positive scaling $\boldsymbol{b}_\theta(\boldsymbol{X}_t) = \boldsymbol{s}_\theta(\boldsymbol{X}_t) * (\boldsymbol{R}_\mathcal{B} - \boldsymbol{R}_t)$ for large systems

Results

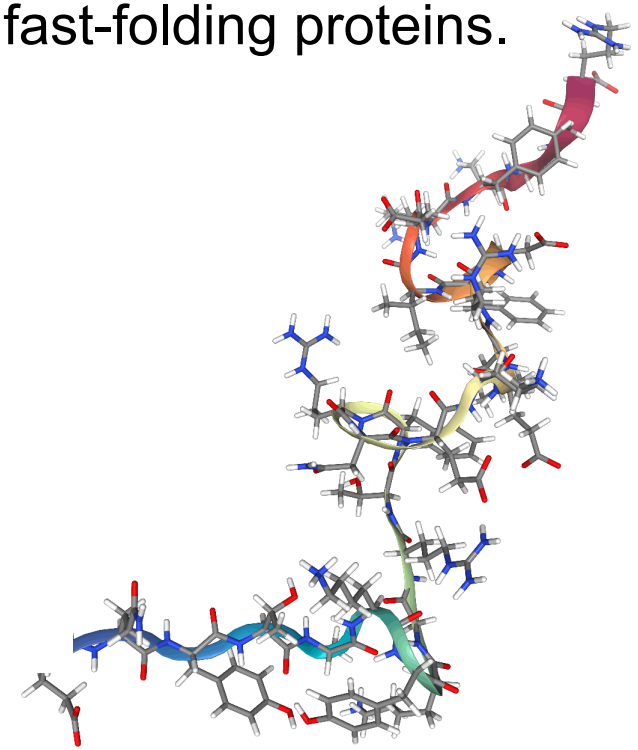
- Our method, named TPS-DPS, samples more realistic and diverse transition paths than baselines on double-well, Alanine Dipeptide, and fast-folding proteins.



Double-well

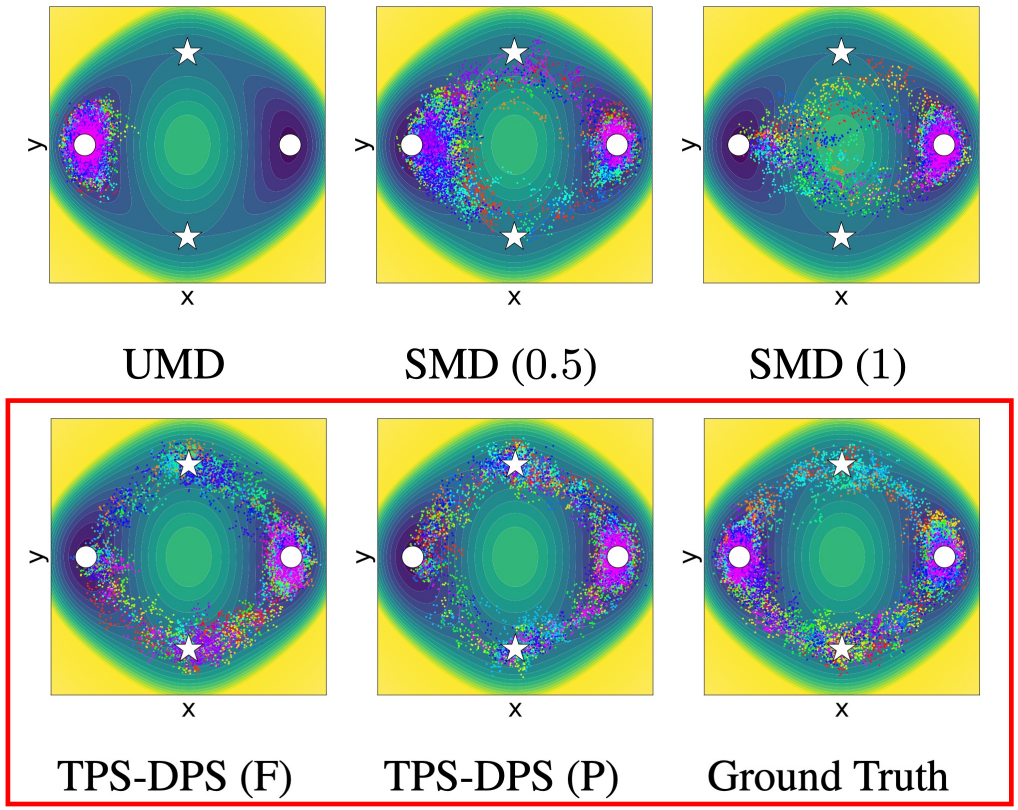


Alanine Dipeptide

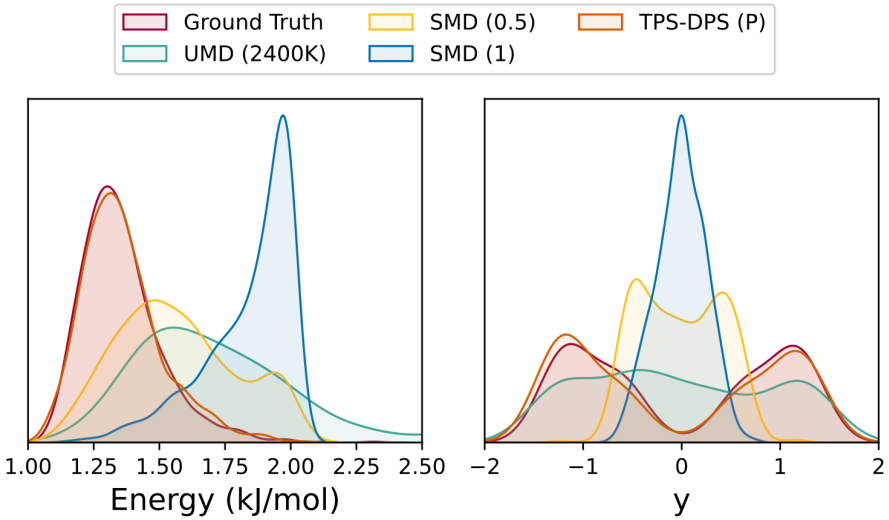


Fast-folding protein

Results – double well



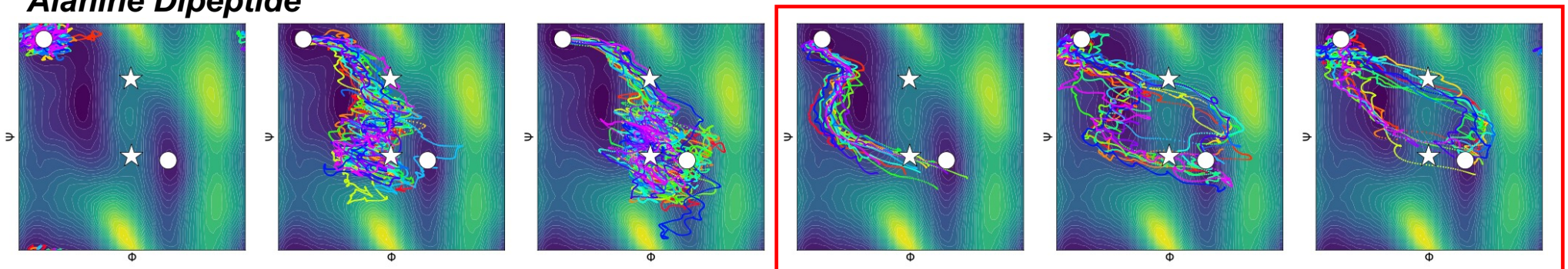
Sampled paths



Distributions of transition state

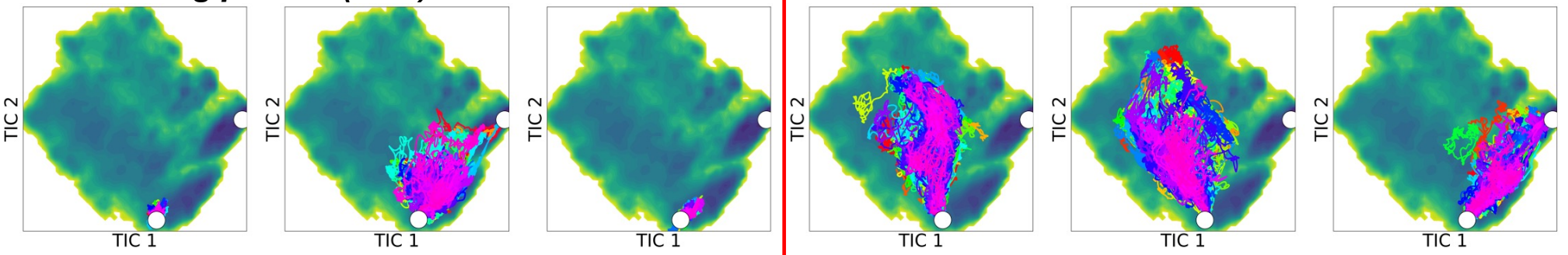
Results – real molecules

Alanine Dipeptide



(a) UMD (b) SMD (10) (c) SMD (20) (d) PIPS (F) (e) TPS-DPS (F) (f) TPS-DPS (P)

Fast folding protein (BBL)



(a) UMD (b) SMD (20k) (c) PIPS (P) (d) TPS-DPS (F) (e) TPS-DPS (P) (f) TPS-DPS (S)

[1] Holdijk, Lars, et al. (2023). "Stochastic optimal control for collective variable free sampling of molecular transition paths." Neurips

Links

- Paper: https://arxiv.org/abs/2405.19961v5
- Project page: https://kiyoung98.github.io/tps-dps/
- Code: https://github.com/kiyoung98/tps-dps