From data to noise to data: mixing physics across temperatures with generative artificial intelligence

Yihang Wang, Lukas Herron, Pratyush Tiwary
University of Maryland

Presented by Connor Blake

Outline

- Problem Framing
- Denoising Diffusion Probabilistic Model (DDPM) Review
- Replica Exchange Molecular Dynamics (REMD) Overview
- Methods
- Results

Challenge

- DOFs $> 10{,}000$

Challenge

- DOFs $> 10{,}000$
- $p(q,p) \propto \exp\big(-\beta\, U(q,p)\big) \approx 0$ in almost all of phase space

Challenge

- DOFs $> 10{,}000$
- $p(q,p) \propto \exp\big(-\beta\, U(q,p)\big) \approx 0$ in almost all of phase space
- Transition states and metastable states are rarely observed in MD

Challenge

- DOFs $> 10{,}000$
- $p(q,p) \propto \exp\big(-\beta\, U(q,p)\big) \approx 0$ in almost all of phase space
- Transition states and metastable states are rarely observed in MD
- Same system at different temperature requires new simulation

Challenge

- DOFs $> 10{,}000$
- $p(q,p) \propto \exp\big(-\beta\, U(q,p)\big) \approx 0$ in almost all of phase space
- Transition states and metastable states are rarely observed in MD
- Same system at different temperature requires new simulation

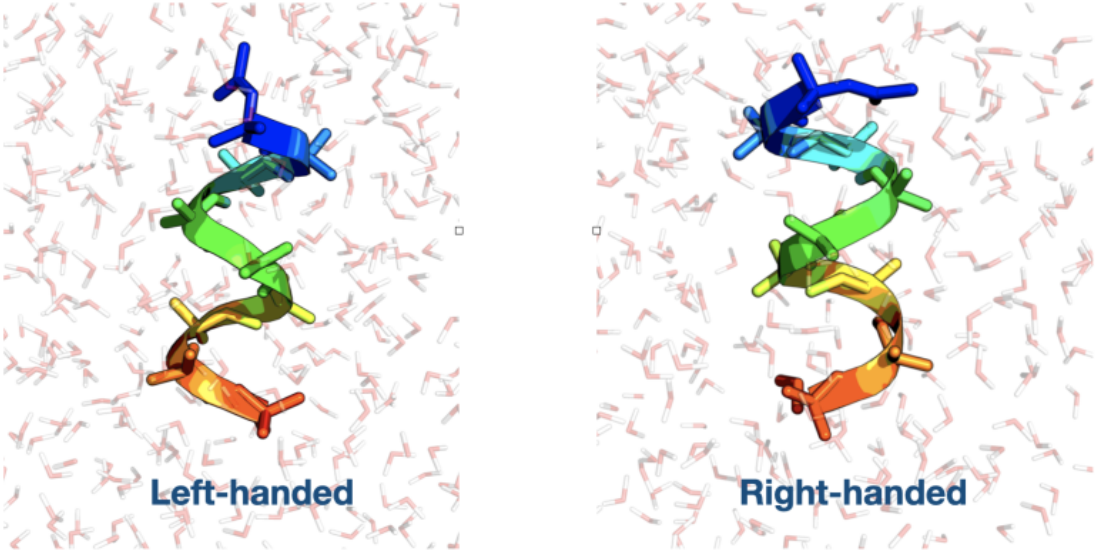


Figure: Aib9 Left and Right Conformations [1]

Notation and Motivation

For a typical sample (numbers here for Aib9)

- $N = 4749$ (all-atoms, explicit TIP3P solvation)
- $x = \{q, p\} \in \mathbb{R}^{N \times 2}$
- $s = \{x, \beta\}$
- $x' = R(x)$ (reduced coordinates)
- $N' = 18$

From equilibrium thermodynamics, $\Delta T \propto \frac{1}{\sqrt{N}}$

Review of Denoising Diffusion Probabilistic Model (DDPM)

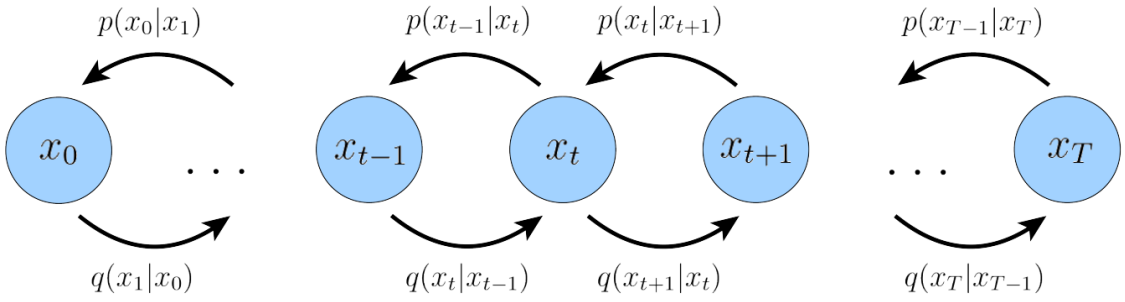


Figure: Markov Chain view of DDPM [2]

1. $$q(s_T \mid s_0) = \prod_{t=1}^{T} q(s_t \mid s_{t-1})$$

Review of Denoising Diffusion Probabilistic Model (DDPM)

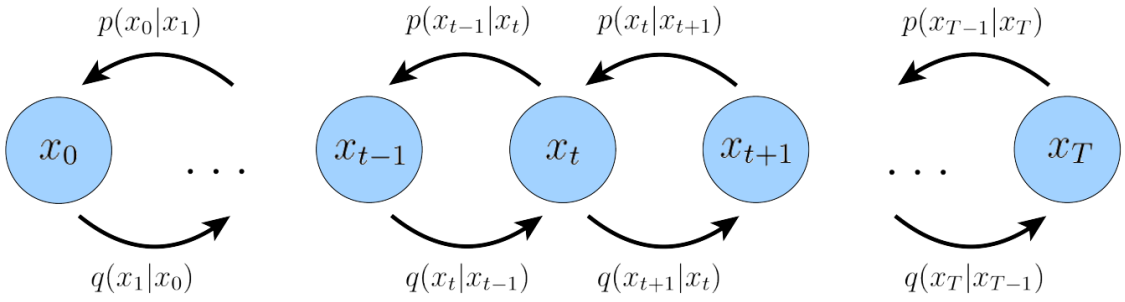


Figure: Markov Chain view of DDPM [2]

1. $$q(s_T \mid s_0) = \prod_{t=1}^{T} q(s_t \mid s_{t-1})$$

2. $$q(s_t \mid s_{t-1}) = \mathcal{N}\left(s_t; \sqrt{\alpha_t}\, s_{t-1},\, (1-\alpha_t)\, I\right)$$

Review of Denoising Diffusion Probabilistic Model (DDPM)

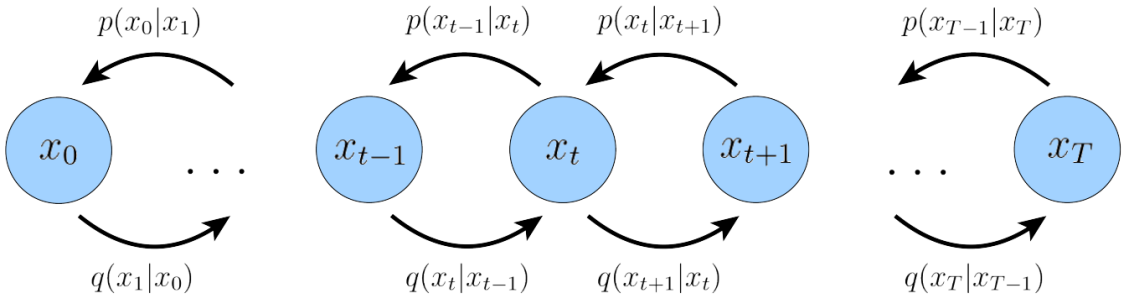


Figure: Markov Chain view of DDPM [2]

1. $$q(s_T \mid s_0) = \prod_{t=1}^{T} q(s_t \mid s_{t-1})$$

2. $$q(s_t \mid s_{t-1}) = \mathcal{N}\big(s_t; \sqrt{\alpha_t}\, s_{t-1},\, (1-\alpha_t)\, I\big)$$

3. $$p(s_0) = p(s_T) \prod_{t=1}^{T} p_\theta(s_{t-1} \mid s_t), \quad p(s_T) = \mathcal{N}(s_T; 0, I)$$

DDPM Architecture Used Here

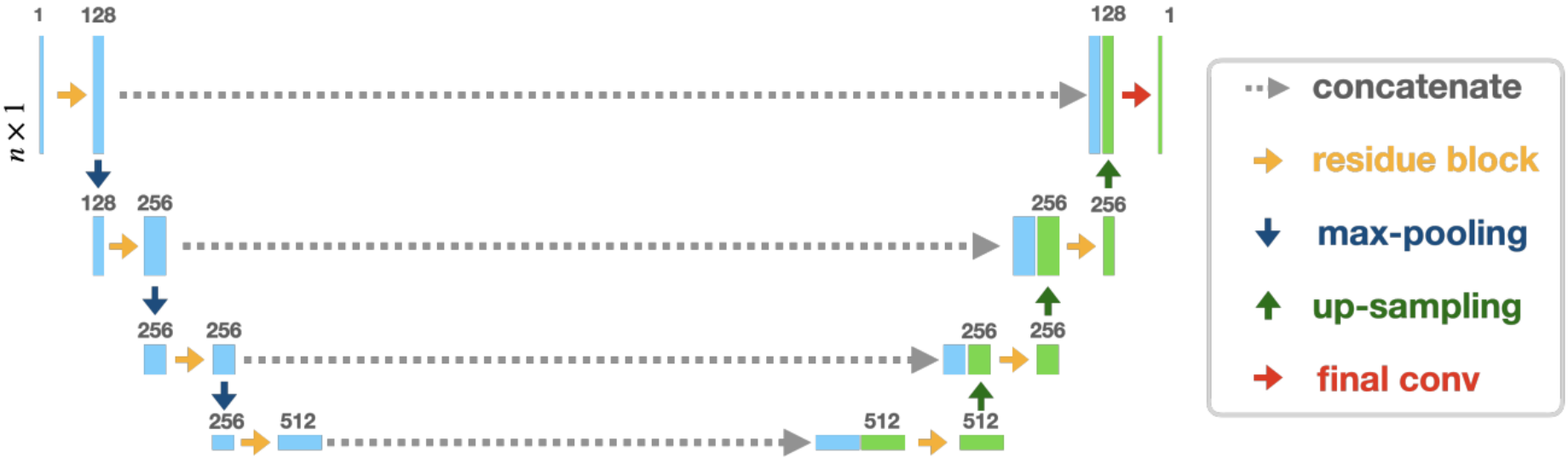


Figure: U-Net Convolutional Architecture for p_θ denoising layers [1]

Replica-Exchange Molecular Dynamics (REMD) [3]
Overview

1. Run M non-interacting MD simulations of system at unique temperatures $\{T_m\}_{m=1}^{M}$

Replica-Exchange Molecular Dynamics (REMD) [3]
Overview

1. Run M non-interacting MD simulations of system at unique temperatures $\{T_m\}_{m=1}^{M}$
2. After K timesteps, randomly select 2 of the systems, i and j at temperatures T_m, T_n respectively

Replica-Exchange Molecular Dynamics (REMD) [3]
Overview

1. Run M non-interacting MD simulations of system at unique temperatures $\{T_m\}_{m=1}^{M}$
2. After K timesteps, randomly select 2 of the systems, i and j at temperatures T_m, T_n respectively
3. Send $(q^i, p^i)_m \rightarrow (q^j, \sqrt{\frac{T_n}{T_m}}\, p^j)_n$ and vice versa

Replica-Exchange Molecular Dynamics (REMD) [3]
Overview

1. Run M non-interacting MD simulations of system at unique temperatures $\{T_m\}_{m=1}^{M}$
2. After K timesteps, randomly select 2 of the systems, i and j at temperatures T_m, T_n respectively
3. Send $(q^i, p^i)_m \rightarrow (q^j, \sqrt{\frac{T_n}{T_m}}\, p^j)_n$ and vice versa
4. Compute

$$\Delta = (\beta_n - \beta_m)\big(U(q^i) - U(q^j)\big)$$

Replica-Exchange Molecular Dynamics (REMD) [3] Overview

1. Run M non-interacting MD simulations of system at unique temperatures $\{T_m\}_{m=1}^{M}$
2. After K timesteps, randomly select 2 of the systems, i and j at temperatures T_m, T_n respectively
3. Send $(q^i, p^i)_m \rightarrow (q^j, \sqrt{\frac{T_n}{T_m}}\, p^j)_n$ and vice versa
4. Compute
$$\Delta = (\beta_n - \beta_m)\big(U(q^i) - U(q^j)\big)$$
5. For detailed balance, compute the probability (Metropolis-style)

$$P(i \leftrightarrow j) = \begin{cases} 1, & \Delta \leq 0, \\ \exp(-\Delta), & \Delta > 0 \end{cases}$$

Replica-Exchange Molecular Dynamics (REMD) [3] Overview

1. Run M non-interacting MD simulations of system at unique temperatures $\{T_m\}_{m=1}^{M}$
2. After K timesteps, randomly select 2 of the systems, i and j at temperatures T_m, T_n respectively
3. Send $(q^i, p^i)_m \rightarrow (q^j, \sqrt{\frac{T_n}{T_m}}\, p^j)_n$ and vice versa
4. Compute
$$\Delta = (\beta_n - \beta_m)\big(U(q^i) - U(q^j)\big)$$
5. For detailed balance, compute the probability (Metropolis-style)
$$P(i \leftrightarrow j) = \begin{cases} 1, & \Delta \leq 0, \\ \exp(-\Delta), & \Delta > 0 \end{cases}$$
6. Accept this swap with probability $P(i \leftrightarrow j)$

Replica-Exchange Molecular Dynamics (REMD) [3]
Overview

1. Run M non-interacting MD simulations of system at unique temperatures $\{T_m\}_{m=1}^{M}$
2. After K timesteps, randomly select 2 of the systems, i and j at temperatures T_m, T_n respectively
3. Send $(q^i, p^i)_m \rightarrow (q^j, \sqrt{\frac{T_n}{T_m}}\, p^j)_n$ and vice versa
4. Compute
$$\Delta = (\beta_n - \beta_m)\big(U(q^i) - U(q^j)\big)$$
5. For detailed balance, compute the probability (Metropolis-style)
$$P(i \leftrightarrow j) = \begin{cases} 1, & \Delta \leq 0, \\ \exp(-\Delta), & \Delta > 0 \end{cases}$$
6. Accept this swap with probability $P(i \leftrightarrow j)$
7. Repeat steps 2–6 until simulation is done

High Level DDPM+REMD Description

- Function to learn: joint PDF $p(s) \equiv p(x, \beta)$ via DDPM diffusion model

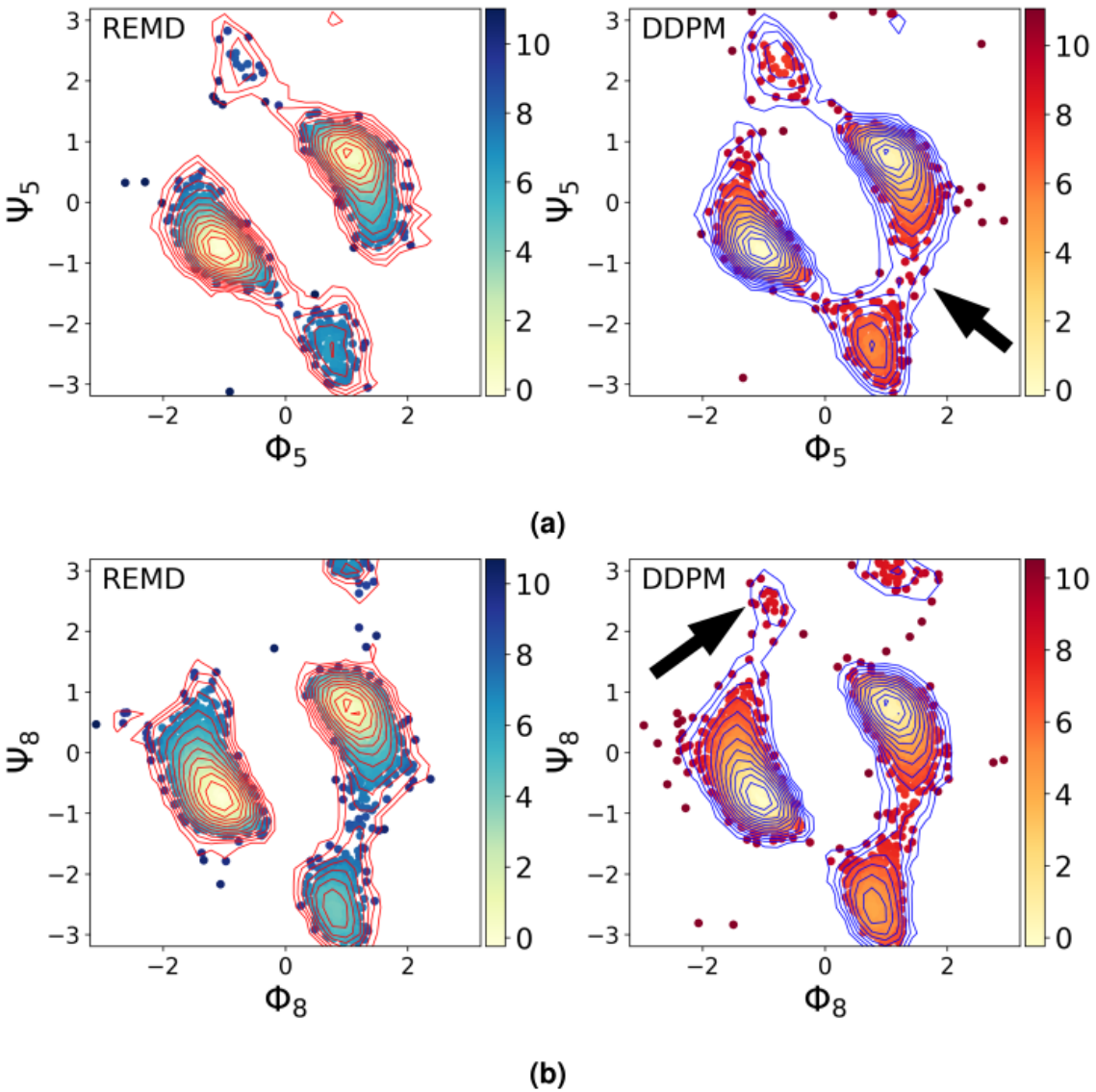
High Level DDPM+REMD Description

- Function to learn: joint PDF $p(s) \equiv p(x, \beta)$ via DDPM diffusion model
- Training Data: REMD trajectories in reduced coordinates
- Benchmarks: Unbiased long-time MD, pure REMD

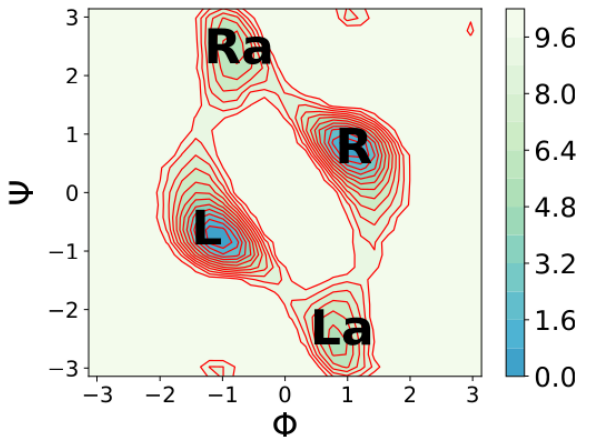
High Level DDPM+REMD Description

- Function to learn: joint PDF $p(s) \equiv p(x, \beta)$ via DDPM diffusion model
- Training Data: REMD trajectories in reduced coordinates
- Benchmarks: Unbiased long-time MD, pure REMD
- Metrics: Can it better-identify transition and metastable states and their free energies? Can it be used to interpolate and extrapolate free energies at temperatures outside $\{T_m\}_{m=1}^{M}$?

Experiment 1: Aib9



Aib9 ϕ, ψ angles on the 5,8 residues from REMD, REMD+DDPM at 400K. [1]



Unbiased 4μs MD, 400K free energy profile for residue 5. All units (both plots) in $k_B T$.

Experiment 1: Aib9 (cont.)

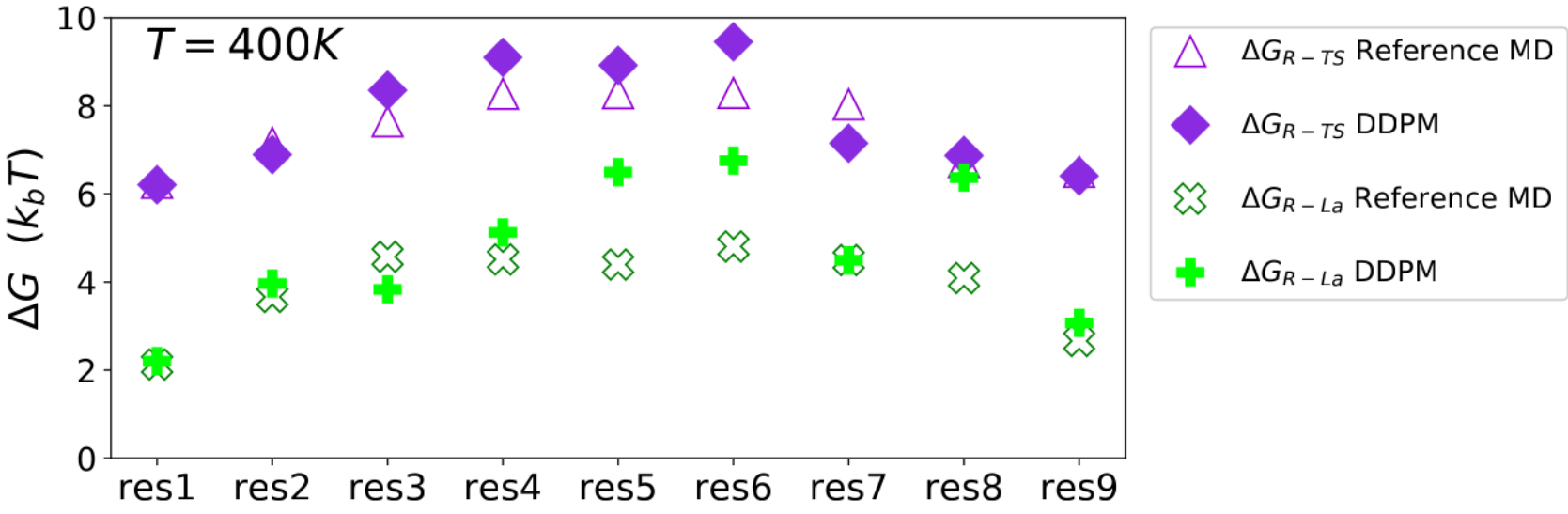


Figure: Extrapolation free energy predictions still work.[1]

Experiment 2: rGACC

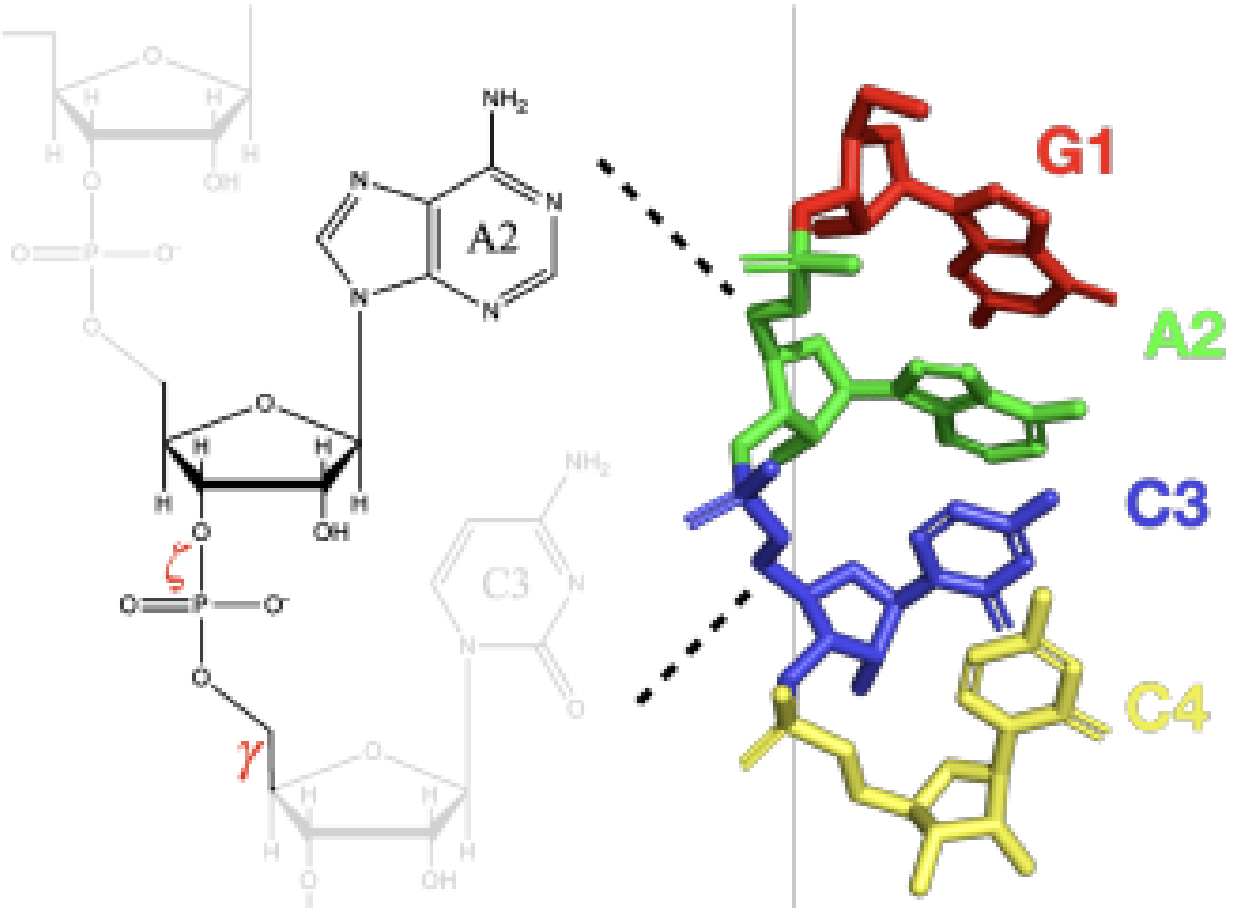
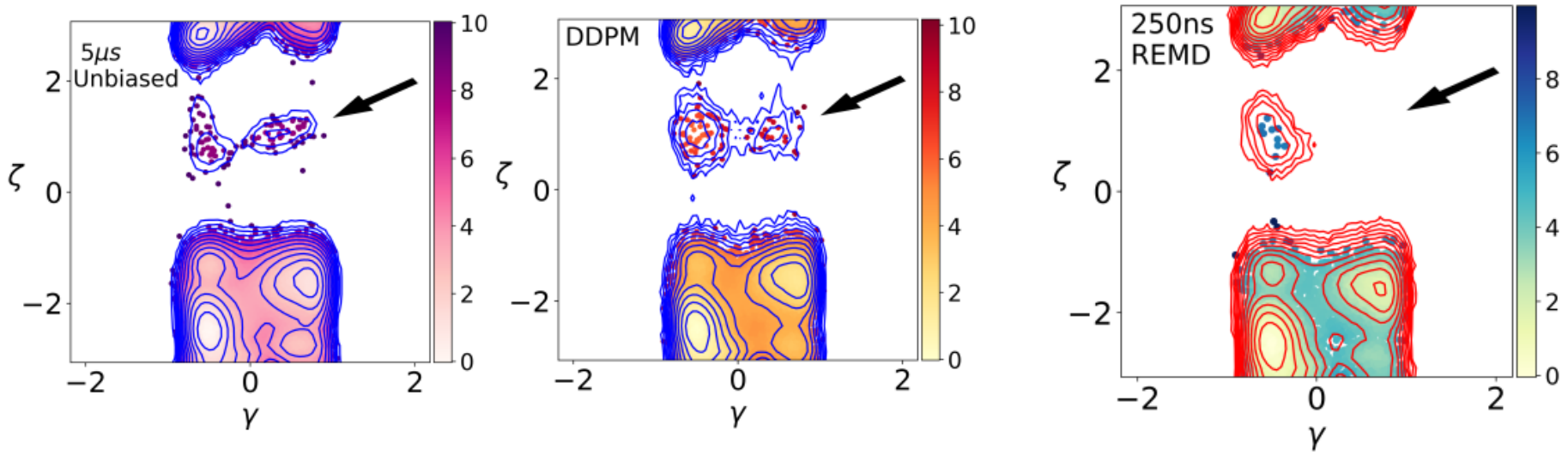


Figure: [1]

Experiment 2: rGACC (cont.)



rGACC ζ, γ angles on the A2 from unbiased MD, REMD+DDPM, REMD at 325K. [1]

Comparison with Boltzmann Generators

This Paper [1]

Boltzmann Generators (Tommy+Joshua) [4]

Similarities

- trying to efficiently construct equilibrium distribution including transition/metastable states
- comparable MD costs in training

Differences

- stochastic map (more expressive) $z \to x$
- generates reduced coordinate structures in explicit solvents

- deterministic map $z \to x$
- generates all-atom structures in implicit solvents
- reaction coordinates easy to generate

Takeaways

- REMD+DDPM outperforms REMD, comparable to full MD in equilibrium prediction of small biomolecules

Takeaways

- REMD+DDPM outperforms REMD, comparable to full MD in equilibrium prediction of small biomolecules
- key insight: treats T as an observation, not control variable (will work for any ensemble)

Takeaways

- REMD+DDPM outperforms REMD, comparable to full MD in equilibrium prediction of small biomolecules
- key insight: treats T as an observation, not control variable (will work for any ensemble)
- allows explicit solvation

Takeaways

- REMD+DDPM outperforms REMD, comparable to full MD in equilibrium prediction of small biomolecules
- key insight: treats T as an observation, not control variable (will work for any ensemble)
- allows explicit solvation
- highly expressive stochastic mappings

Takeaways

- REMD+DDPM outperforms REMD, comparable to full MD in equilibrium prediction of small biomolecules
- key insight: treats T as an observation, not control variable (will work for any ensemble)
- allows explicit solvation
- highly expressive stochastic mappings
- Drawbacks: currently must use reduced coordinates, internal sidechain coordinates not addressed

Questions?

References I


Yihang Wang and Pratyush Tiwary.
Denoising diffusion probabilistic models for replica exchange.
arXiv preprint arXiv:2107.07369, 2021.


Calvin Luo.
Understanding diffusion models: A unified perspective, 2022.


Yuji Sugita and Yuko Okamoto.
Replica-exchange molecular dynamics method for protein folding.
Chemical Physics Letters, 314(1):141–151, 1999.


Frank Noé, Simon Olsson, Jonas Köhler, and Hao Wu.
Boltzmann generators – sampling equilibrium states of many-body systems with deep learning, 2019.