Machine-learned reaction coordinates for energy-entropy disentanglement

Eric Beyerle (Tiwary Group)
Brin Workshop on Rare Events
2/27/23

Institute for Physical Science & Technology IPST

University of Maryland 1856

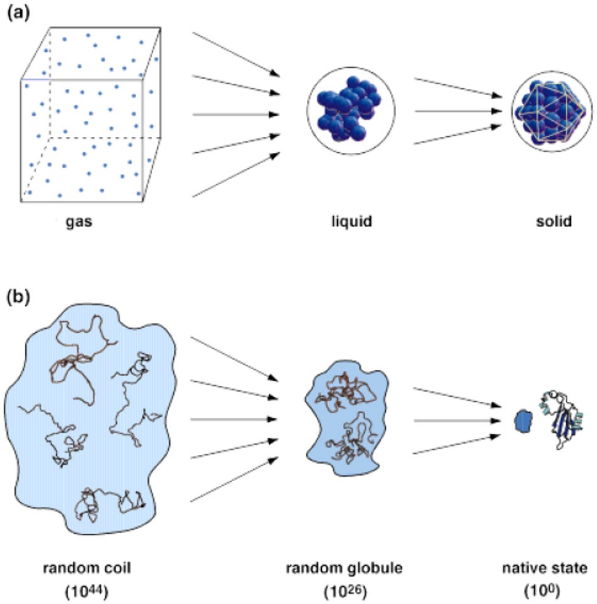
Outline

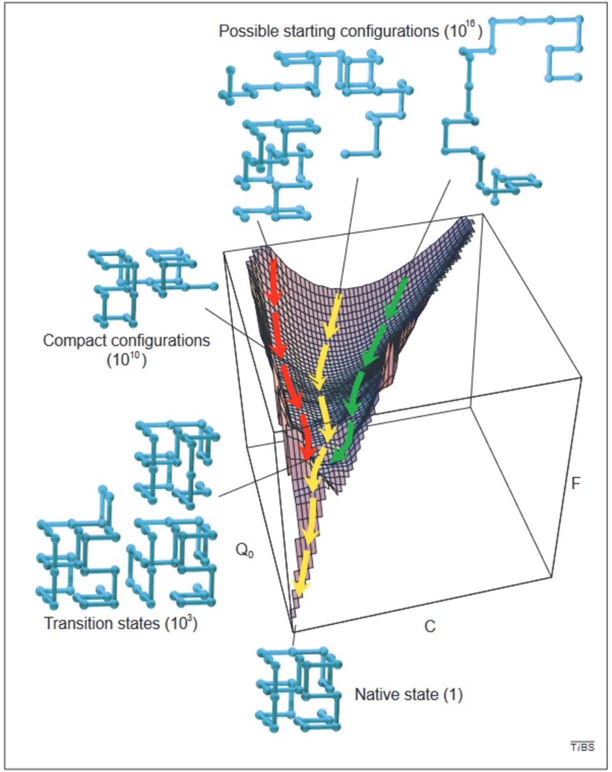
1. Introduction – free energy in physical systems
2. Discovering a relevant latent space – SPIB
3. Energy and entropy in the SPIB latent space
4. Application to simple systems
5. Augmenting the SPIB loss function to learn energetic and entropic pathways
6. Conclusions

Introduction

- All physical processes occurring in a thermostatted ensemble are subject to free-energy barriers, which are a sum of energy and entropy:

$$\Delta A = \Delta U - T\Delta S$$
$$\Delta G = \Delta H - T\Delta S$$
$$= \Delta A + P\Delta V$$





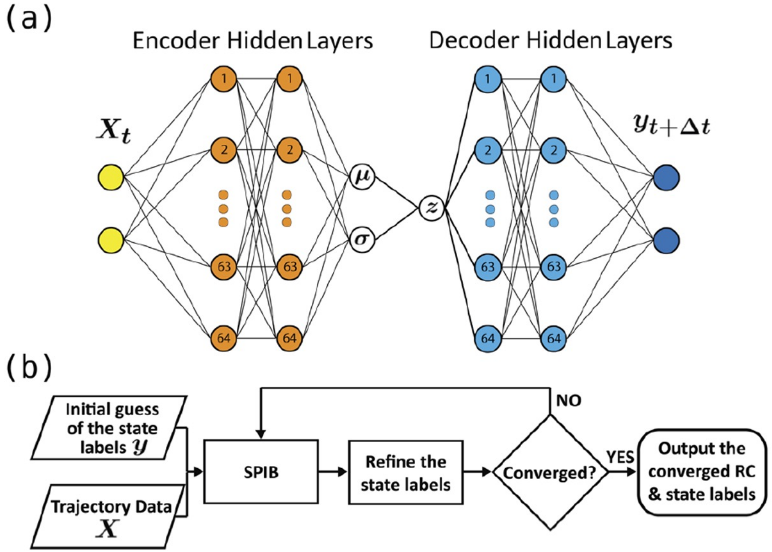
Dinner, Sali, Smith, Dobson, and Karplus, “Understanding protein folding via free-energy surfaces from theory and experiment,” *Trends in Biochemical Sciences*, **25**, 331-339, 2000

Dobson, Sali, and Karplus, “Protein Folding: A Perspective from Theory and Experiment,” *Angew. Chem. Int. Ed.* **37**, 868 - 893, 1998

Introduction

- Generally, there is not a simple, low-dimensional coordinate system describing the entropy- and energy-dominated pathways.
    - Important for determing enhanced sampling method utilized
- Goal of the research described here is to embed the system dynamics in a low-dimensional latent space where energy and entropy pathways are disentangled.
- Our embedding method of choice is to use a modified version of the reweighted autoencoder variational Bayes (RAVE) method called SPIB.

SPIB



$$\mathcal{L}_{\text{SPIB}} = \frac{1}{M \cdot L} \sum_{n=1}^{N} \sum_{l=1}^{L} \left[\underbrace{\log q_\theta \left(\mathbf{y}^{n+s} \middle| \mathbf{z}^{(n,l)} \right)}_{\text{reconstruction error}} - \underbrace{\beta_{\text{SPIB}} \log \frac{p_\theta \left(\mathbf{z}^{(n,l)} \middle| \mathbf{X}^{\mathbf{n}} \right)}{q_\theta \left(\mathbf{z}^{(n,l)} \right)}}_{\text{KL divergence}} \right]$$

$$\mathcal{L}_{\text{SPIB}} \leq \mathcal{L}_{\text{IB}} = \underbrace{I(\mathbf{y}, \mathbf{z})}_{\sim\text{accuracy}} - \beta \underbrace{I(\mathbf{X}, \mathbf{z})}_{\sim\text{compression}}$$

Wang and Tiwary, “State predictive information bottleneck,” *J. Chem. Phys.* **154**, 134111, 2021

Energy and Entropy in the Embedded SPIB Latent Space

- Use the formulation provided by Hartmann et al.[1] to determine the free energy and energy along the discovered RCs:

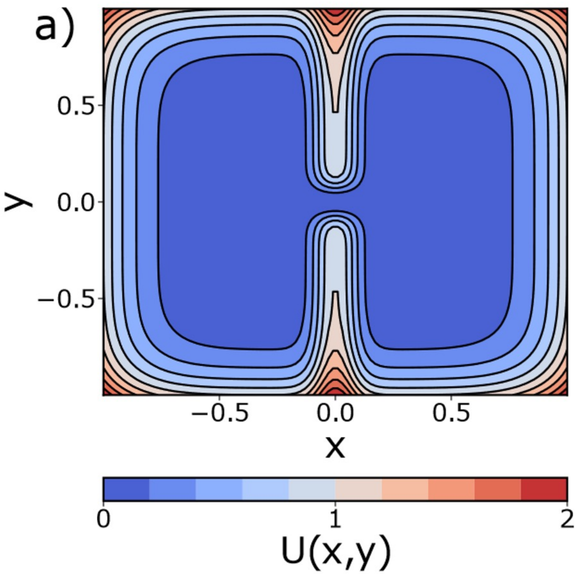
$$G(\mathbf{z}) = -\beta^{-1} \ln \int_{\mathbb{R}^n} \underbrace{d\mathbf{x} \exp(-\beta V(\mathbf{x}))}_{\text{Boltzmann weight of feature } \mathbf{x}} \underbrace{\det\left(\tilde{G}\right)^{\frac{1}{2}}}_{\text{Grammian}} \underbrace{\delta\left(\Phi(\mathbf{x}) - \mathbf{z}\right)}_{\text{Count points on the level set/histogram bin of interest}}$$

$$\begin{aligned} \langle U(z) \rangle_{\Sigma(z)} &:= U(z) \\ &= \frac{1}{N_z} \int U(\mathbf{x}) e^{-U(\mathbf{x})/k_B T} \times \delta(\Phi(\mathbf{x}) - z) \det(\tilde{G})^{1/2} \, d\mathbf{x} \end{aligned}$$

$$\begin{aligned} \Delta G(z) &= \Delta U(z) - T \Delta S(z) \\ &\Rightarrow \Delta S(z) \\ &= \frac{1}{T}(\Delta U(z) - \Delta G(z)) \end{aligned}$$

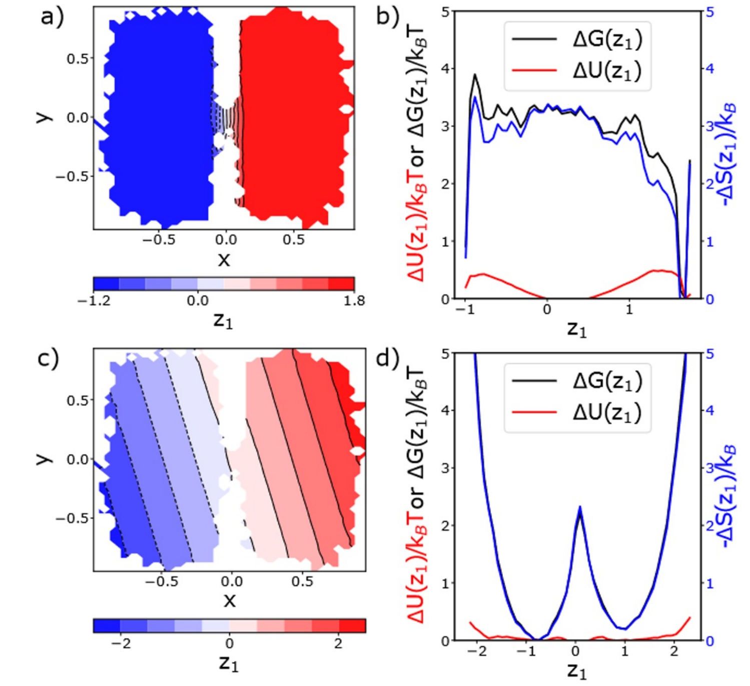
[1]Hartmann, Latorre, and Ciccotti, "On two possible definitions of the free energy for collective variables," *Eur. Phys. J Special Topics*, **200**, 73-89, 2011

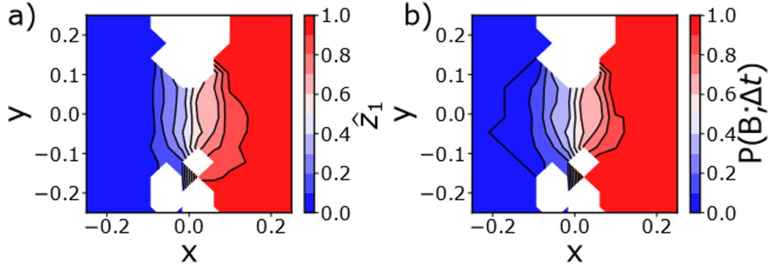
Energy-Entropy for Some Simple Systems



$$U(x,y) = x^6 + y^6 + \exp\left[-y^2/\sigma_y^2\right]\left(1 - \exp\left[x^2/\sigma_x^2\right]\right)$$

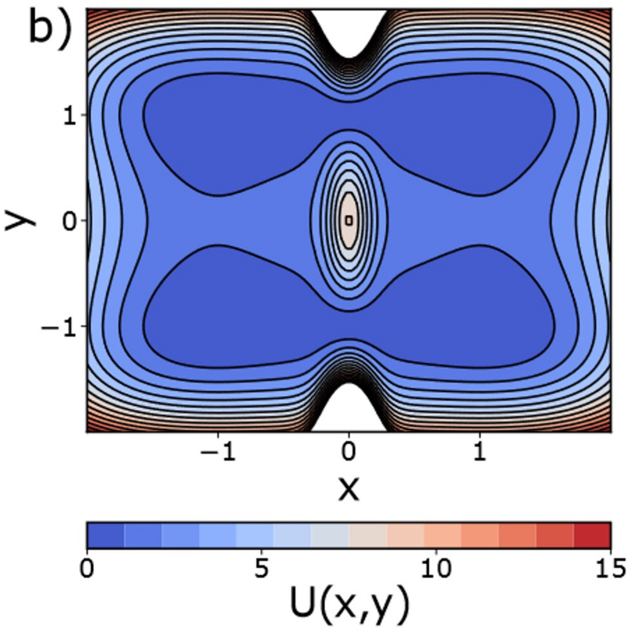
"Elber potential"

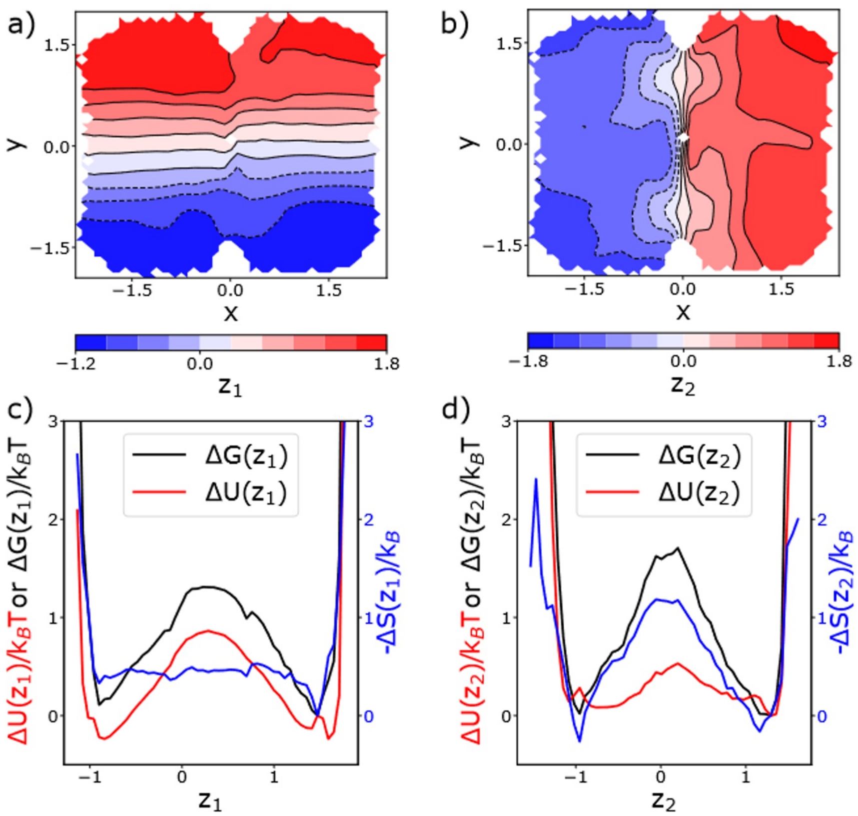




Beyerle, Mehdi, and Tiwary, "Quantifying Energetic and Entropic Pathways in Molecular Systems," *J. Phys. Chem. B*, **126**, 3950-3960, 2022

Energy-Entropy for Some Simple Systems



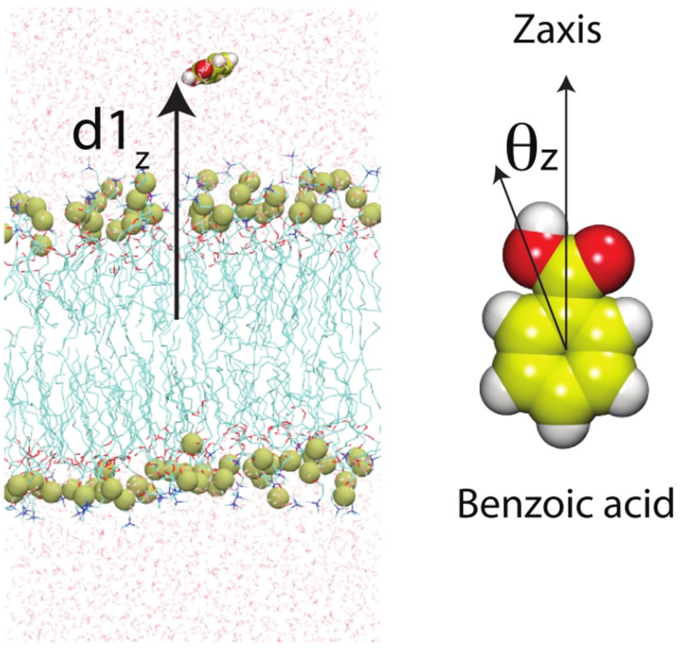


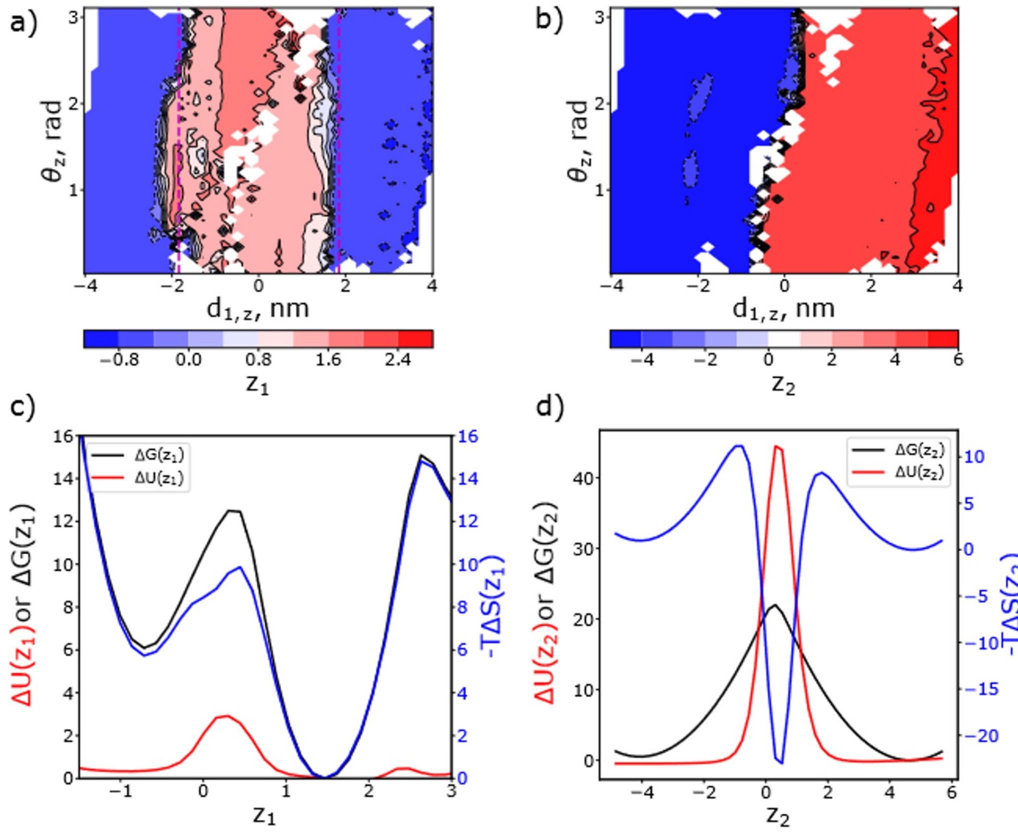
$$U(x, y) = h_x\left(x^2 - 1\right)^2 + (h_y + a((x, \delta))\left(y^2 - 1\right)^2$$

"Temperature switch"

Beyerle, Mehdi, and Tiwary, "Quantifying Energetic and Entropic Pathways in Molecular Systems," *J. Phys. Chem. B*, **126**, 3950-3960, 2022

Energy-Entropy for Some Simple Systems





Beyerle, Mehdi, and Tiwary, "Quantifying Energetic and Entropic Pathways in Molecular Systems," *J. Phys. Chem. B*, **126**, 3950-3960, 2022

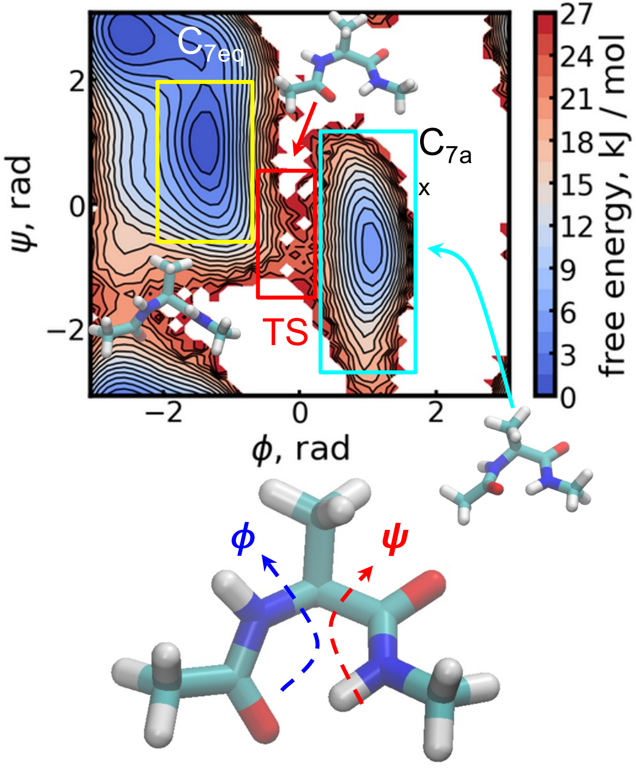
Forcing the Latent Space to Learn Thermodynamics

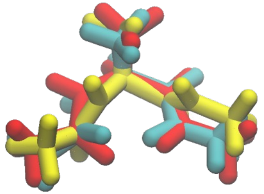
- Add an extra regularization term to the loss function to force the latent space to learn the energy and energy pathways for an arbitrary system.

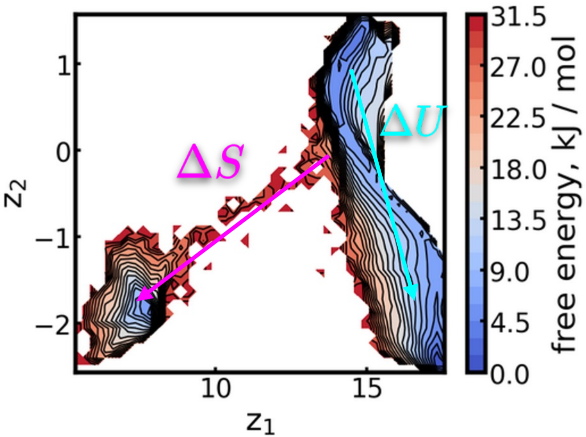
$$\mathcal{L} = \sum_{n=1}^{N} \left[\log(p(\mathbf{y}^{n+s} | \mathbf{z}^{n})) - \beta \log\left(\frac{p(\mathbf{z}^{n} | \mathbf{X}^{n})}{p(\mathbf{z}_{\theta})} \right) \right] + \gamma f(\mathbf{z})$$

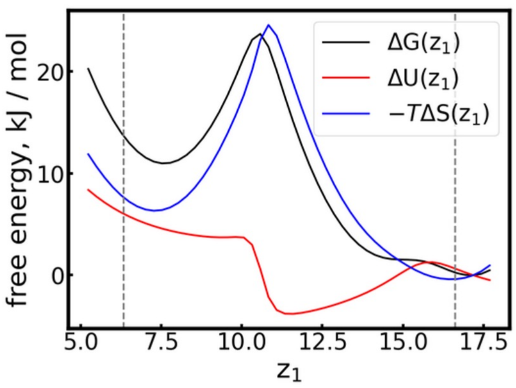
$$f(\mathbf{z}) = \max_{\mathrm{int}(z_1)} \left(-T\Delta S(z_1) \right) + \max_{\mathrm{int}(z_2)} \left(\Delta U(z_2) \right)$$

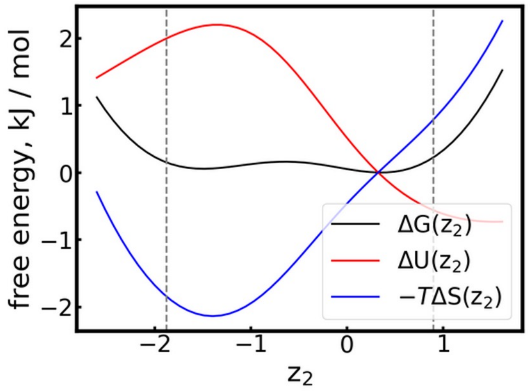
Case Study: Alanine Dipeptide in Vacuum

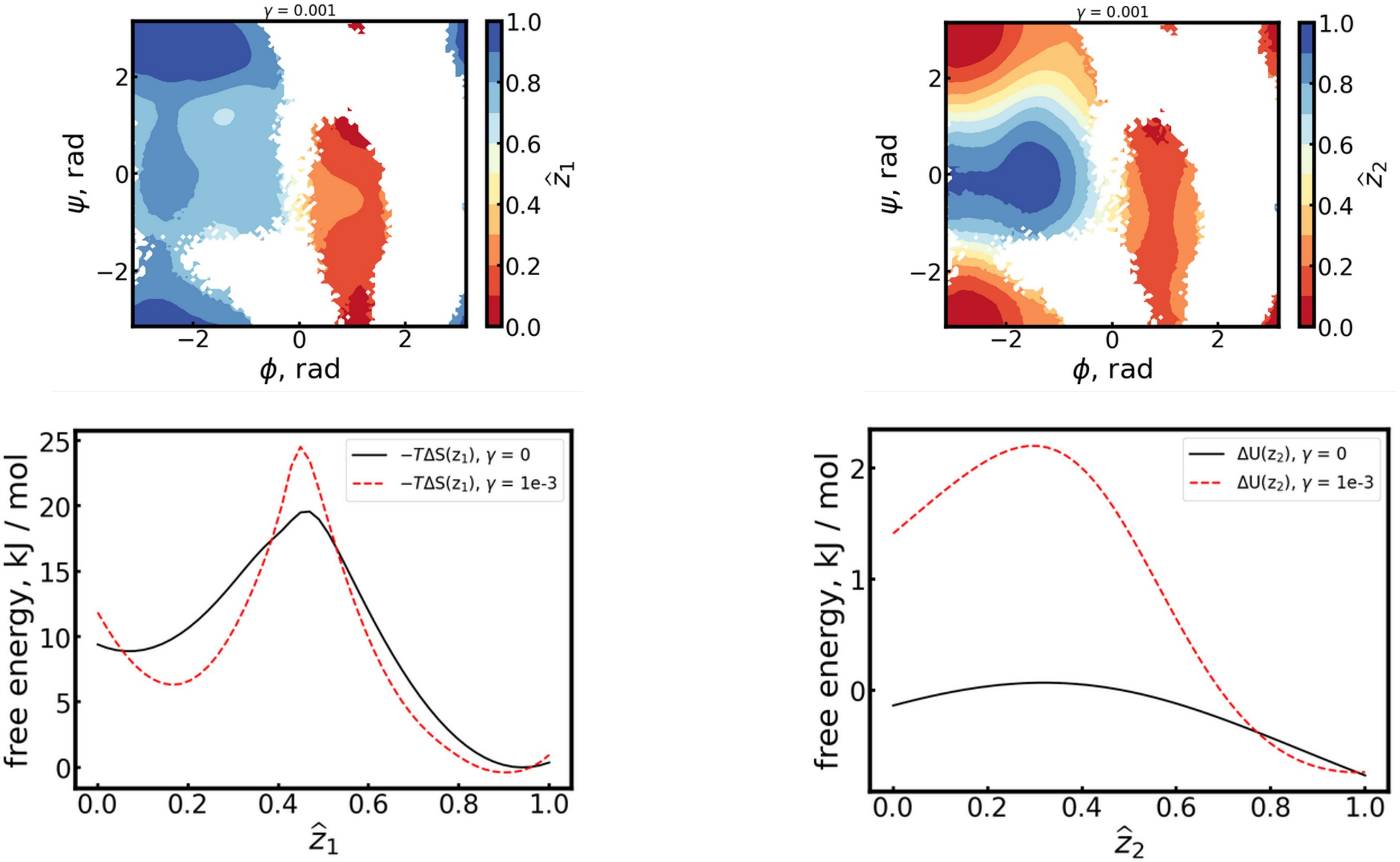












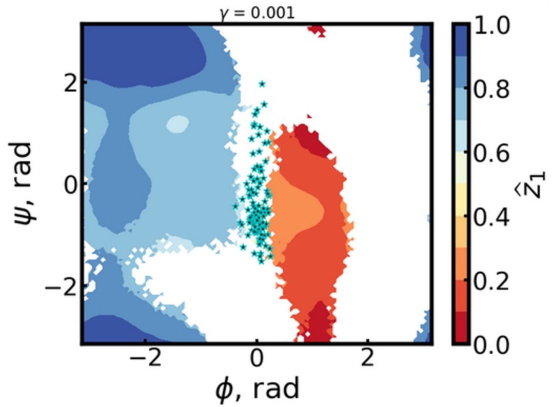
γ = 0.001
ψ, rad
ϕ, rad
$\hat{z}_1$
γ = 0.001
ψ, rad
ϕ, rad
$\hat{z}_2$
free energy, kJ / mol
−TΔS(z1), γ = 0
−TΔS(z1), γ = 1e-3
$\hat{z}_1$
free energy, kJ / mol
ΔU(z2), γ = 0
ΔU(z2), γ = 1e-3
$\hat{z}_2$

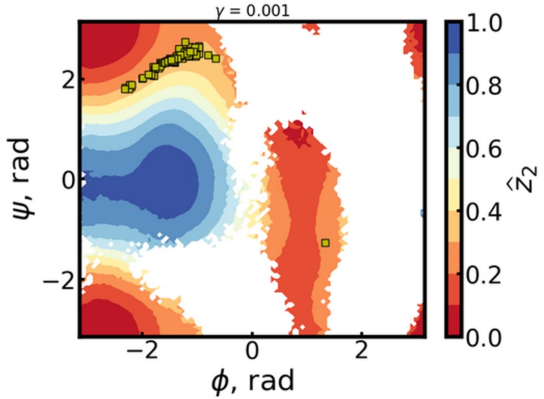
Interpretation

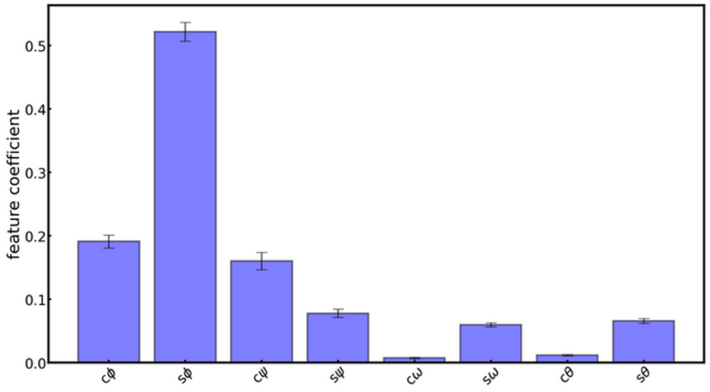
Use an in-house method called TERP[1]:

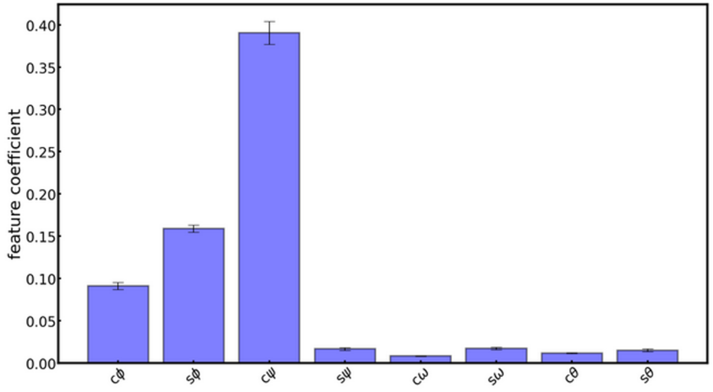
"Unfaithfulness" "Simplicity"

$$\zeta[f, d] = U[f, d] - \theta S[f]$$





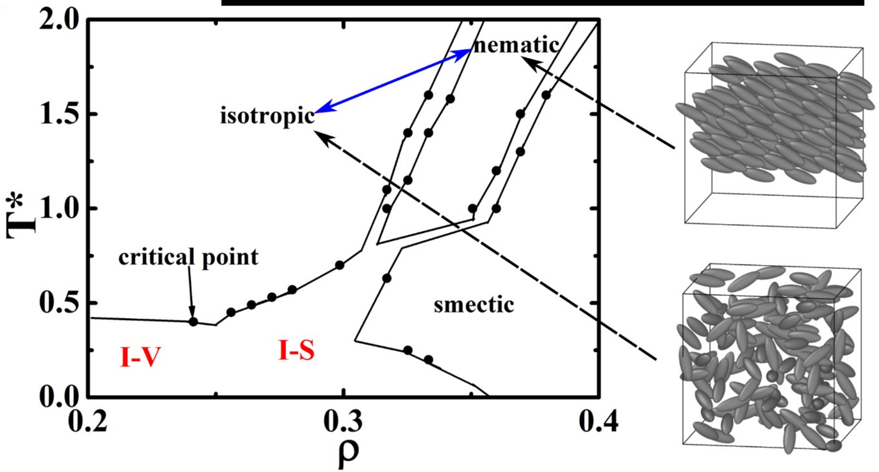




[1]Mehdi and Tiwary, "Thermodynamics of Interpretation," *arXiv,* arXiv:2206.13475v1, 2022

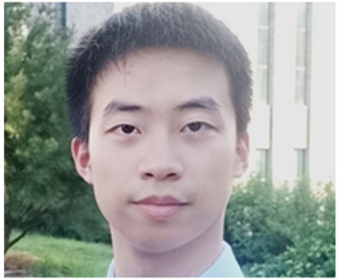
Conclusions and Coming Attractions

- SPIB can discover an effective low-dimensional latent space for crossing energetic and entropic barriers.
- For alanine dipeptide in vacuum, energy and entropy barriers can be enhanced by adding a regularization term to the loss function.
- We look to apply this method to more complicated soft matter systems (hydrophobic ligand binding, liquid crystal models) in the near future.

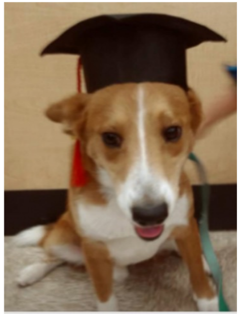


Acknowledgements

Prof. Pratyush Tiwary



Dedi Wang



Pakora

Shams Mehdi

Funding and Resources

Department of Energy · United States of America

XSEDE
Extreme Science and Engineering Discovery Environment

HPC@UMD

ACCESS | Advancing Innovation