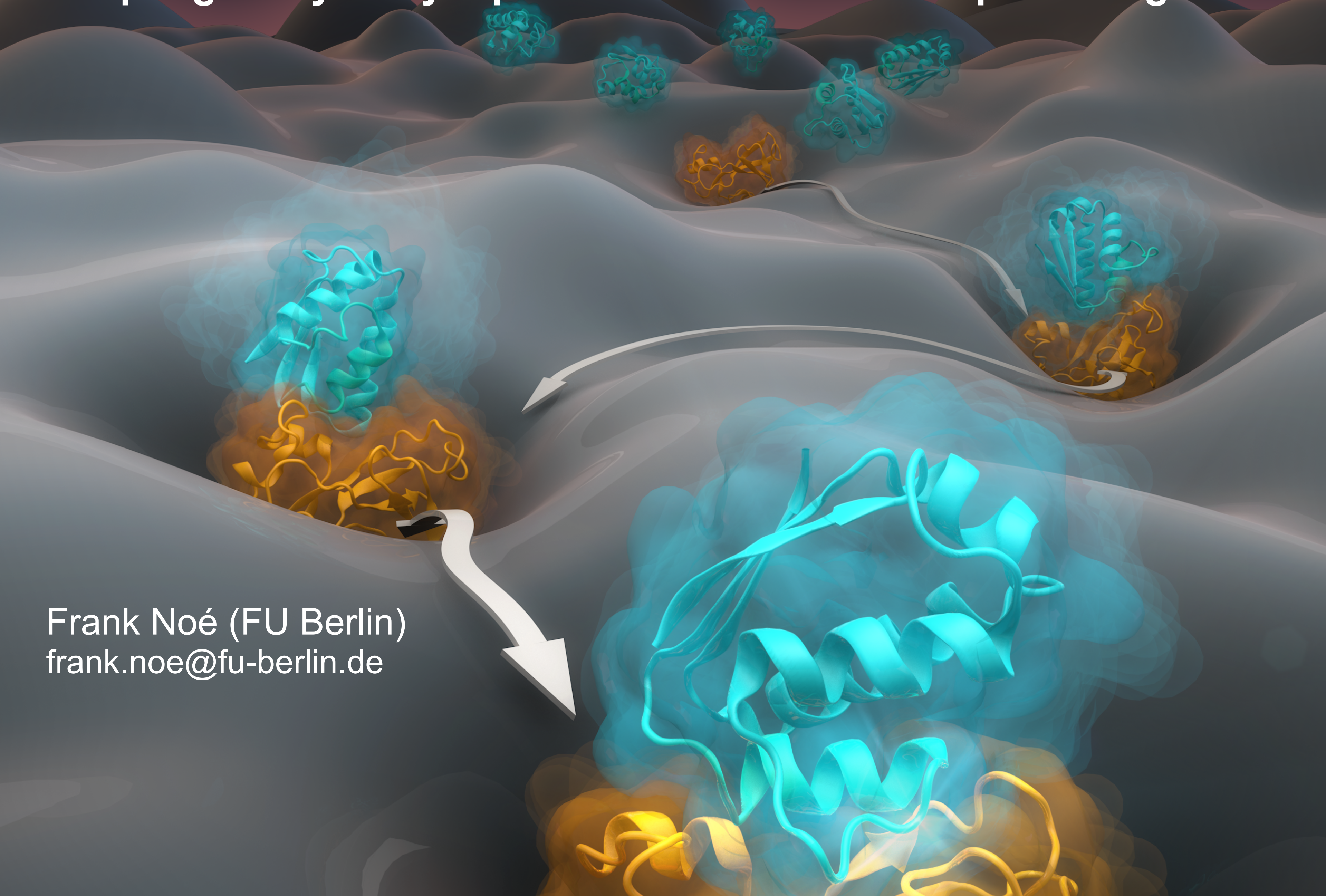
Boltzmann Generators:
sampling many-body equilibrium states with deep learning



Frank Noé (FU Berlin)
frank.noe@fu-berlin.de

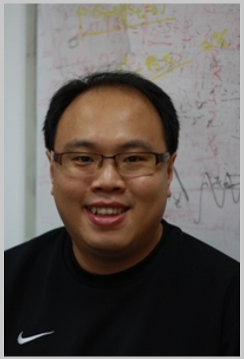
Thermodynamics **Sampling problem**

Boltzmann Generators
sampling equilibrium states of many-body systems with deep learning

Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019)

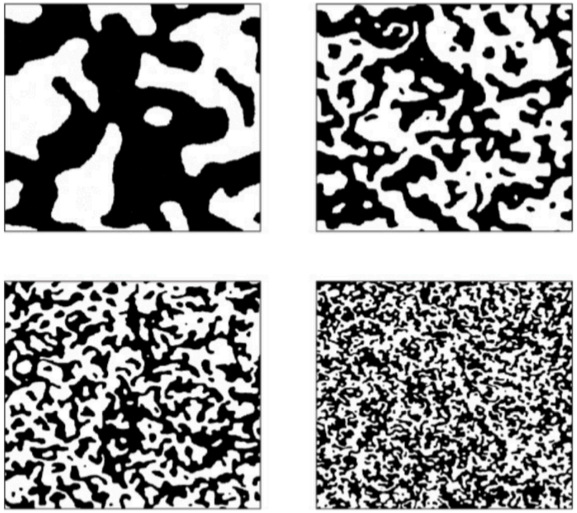
Simon Olsson

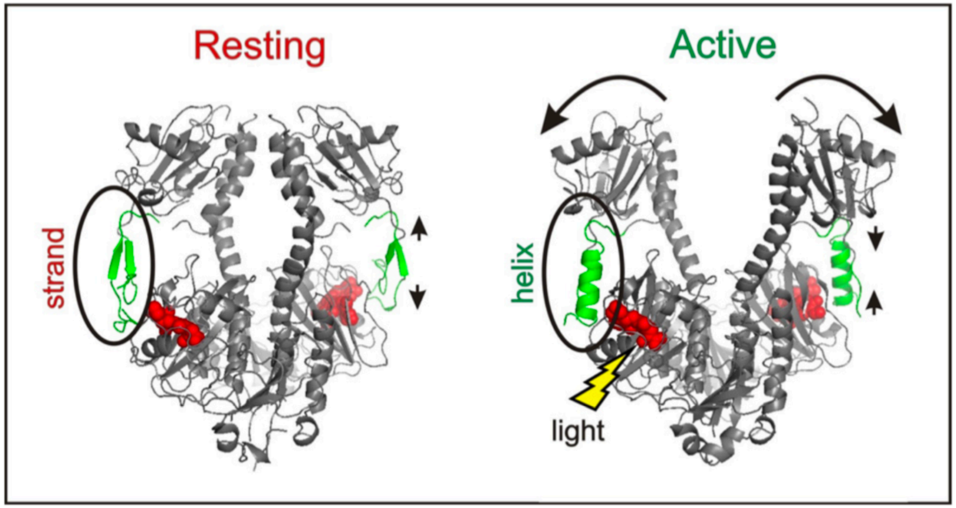
Jonas Köhler



Hao Wu

Boltzmann Generators



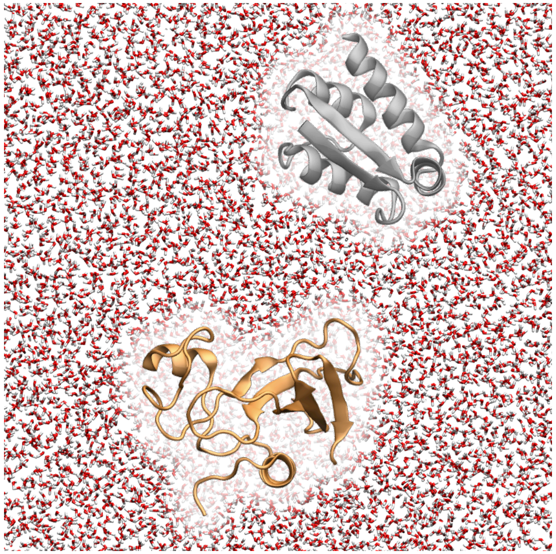


- **Input**: Reduced Potential Energy $u(\mathbf{x})$ in coordinates $\mathbf{x} \in \mathbb{R}^n$, e.g. $u(\mathbf{x}) = U(\mathbf{x})/k_B T$ (canonical ensemble).
- **Aim**: Generate *independent* Samples from Equilibrium Distribution.

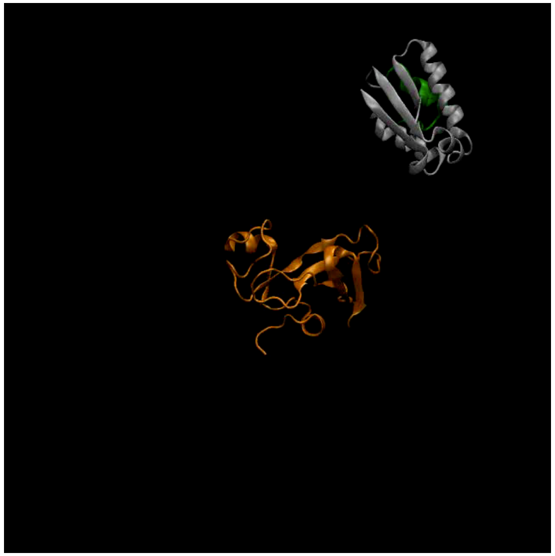
$$\mu(\mathbf{x}) \propto \mathrm{e}^{-u(\mathbf{x})}$$

- **Problem**:
  - Direct MC (proposal for all n degrees of freedom + rejection or reweighting) not available.
  - Standard approach: MD/MCMC with local moves $\rightarrow$ sampling problem.

W1: From passive to active - generative and reinforcement learning with physics



Microsecond
MD Trajectories

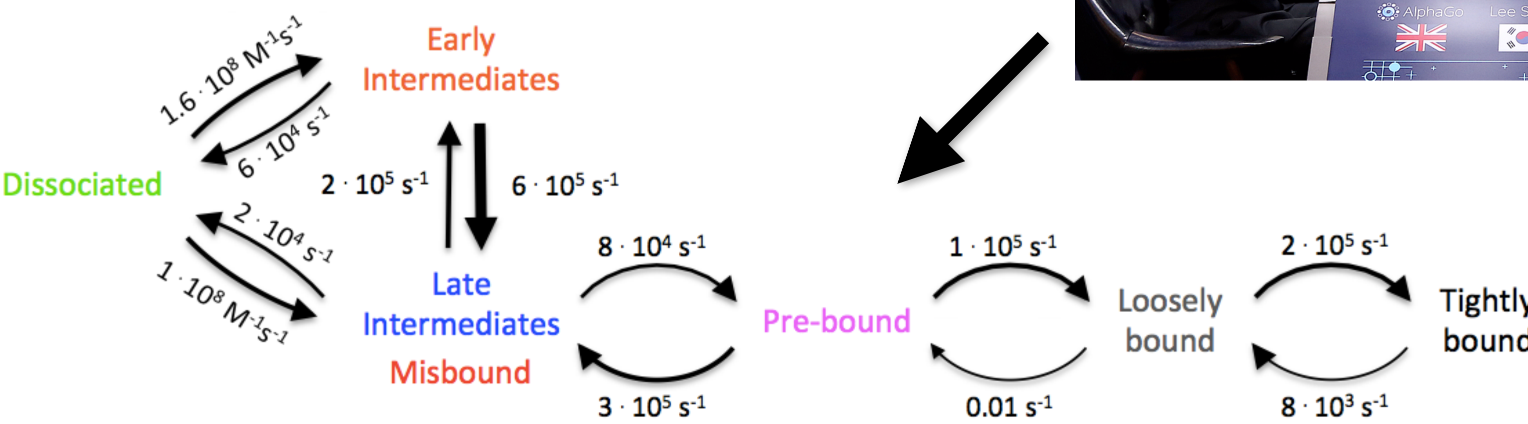


W1: From passive to active - generative and reinforcement learning with physics

Plattner, Doerr, De Fabritiis, Noé
Nature Chemistry 9, 1005 (2017)

Reinforcement learning

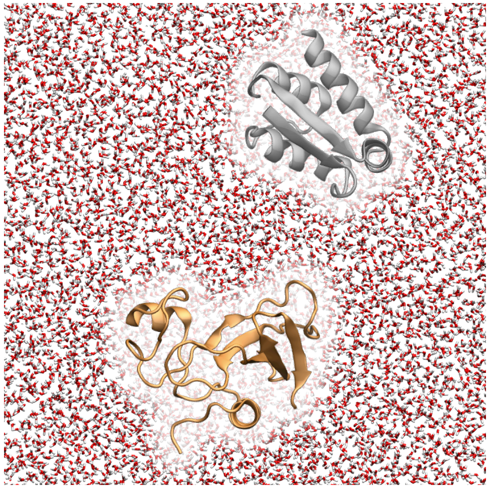
total 2 ms data



Adaptive Markov State Model — **seconds to hours kinetics**

Standard simulation methods in molecular physics are INSANELY expensive

Current approach



very long time

→

a lot of energy

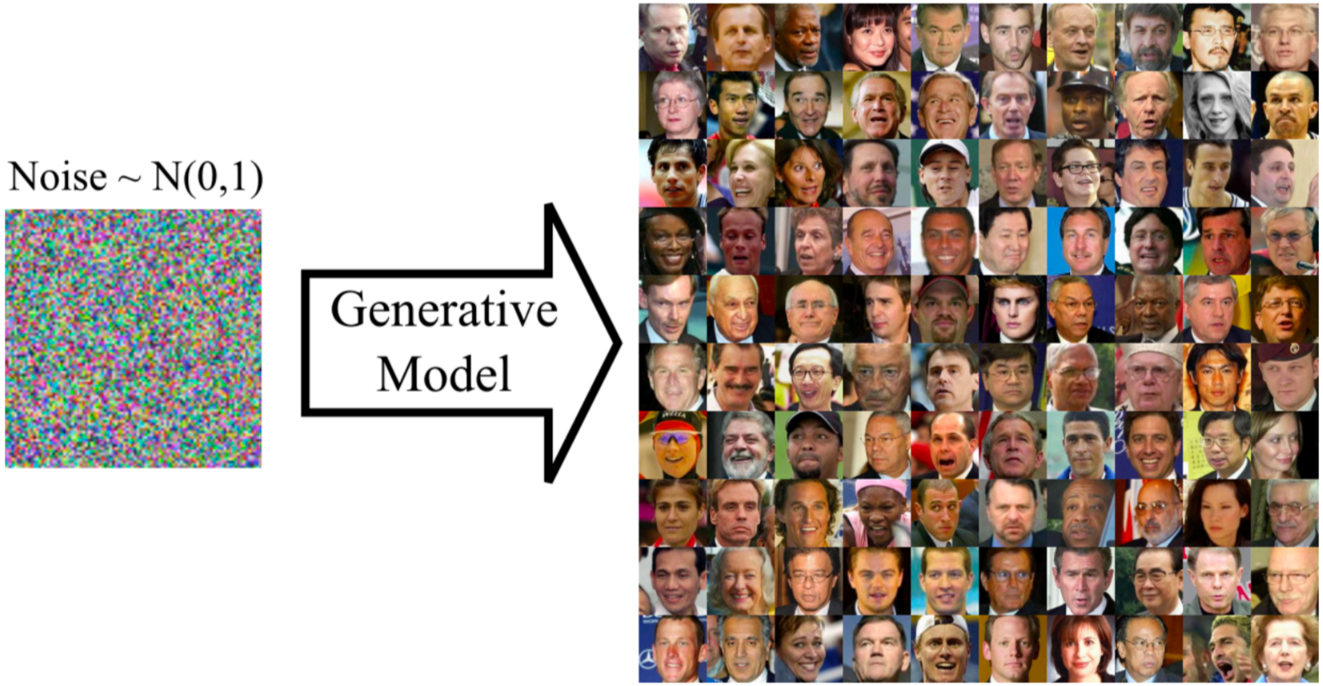
2 ms MD = 20,000 GPU days **500 gigajoule**

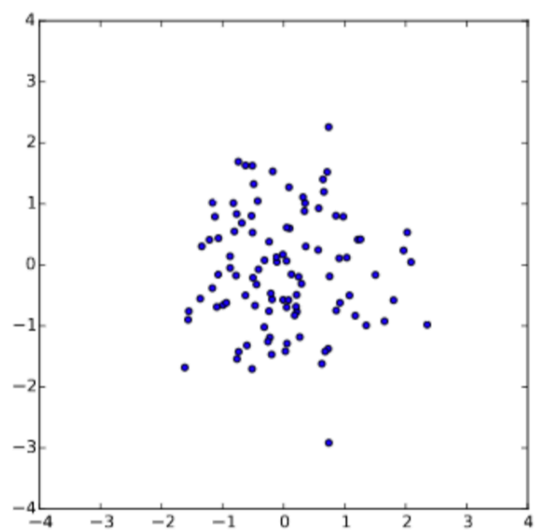
1 non-transferable model

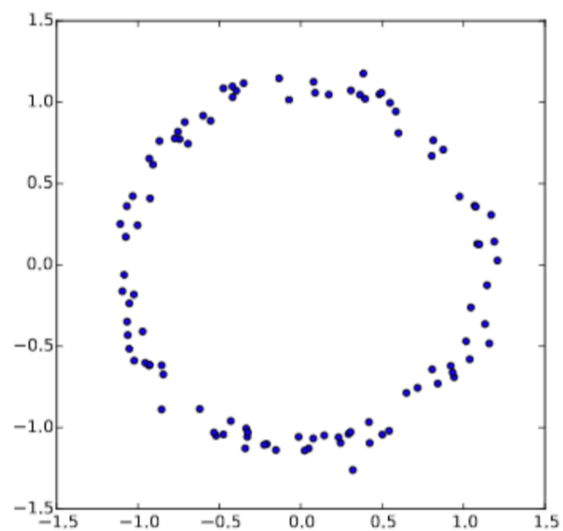
Burn a Saturn V rocket and deliver 50 ton payload to lunar orbit

1500 gigajoule

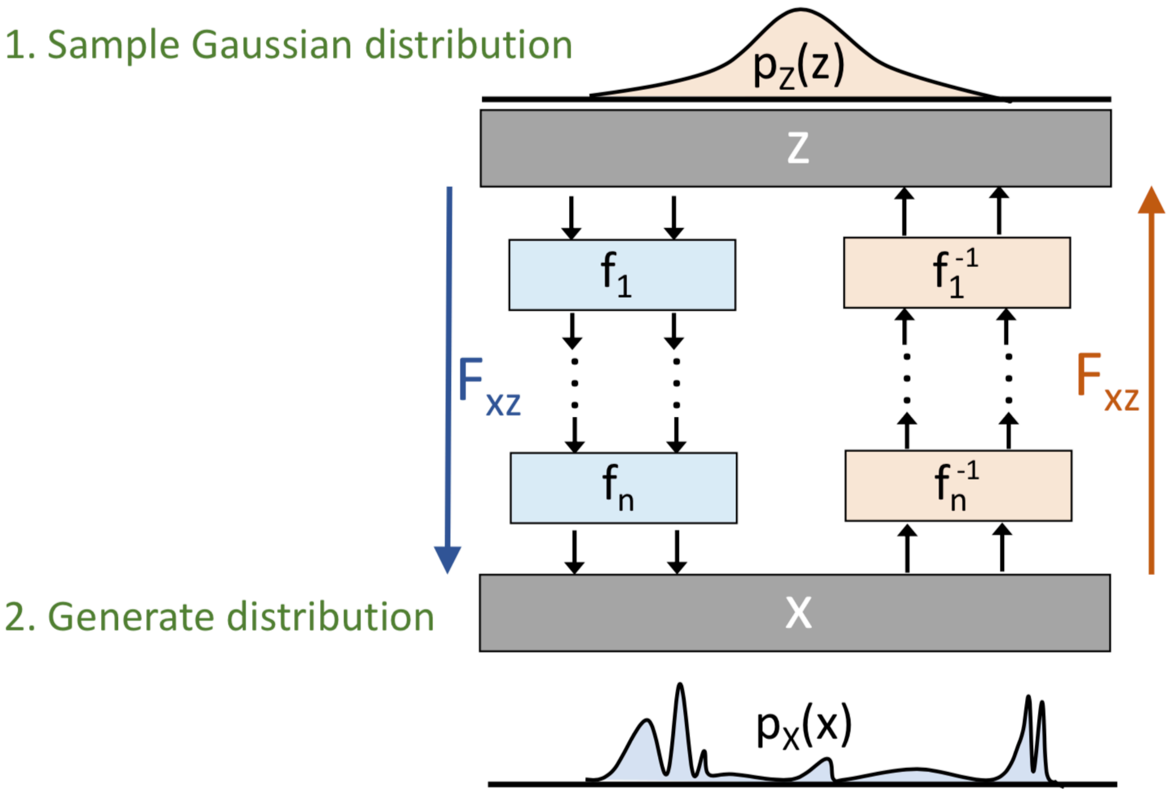
Boltzmann Generators







Flows and Normalizing Flows

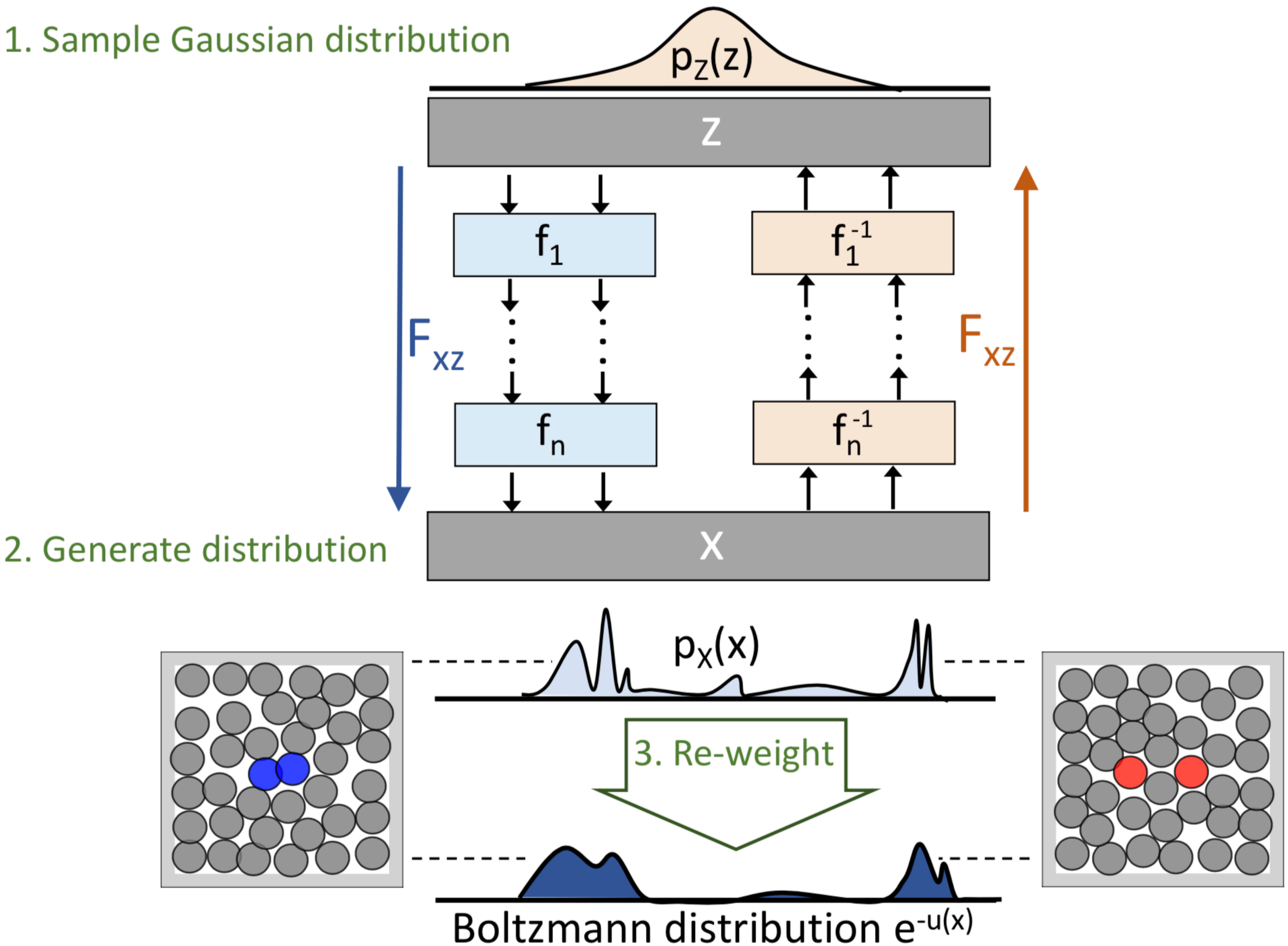


Main idea: Transformation of random variables for bijective transformations:

$$p_X(\mathbf{x}) = p_Z(f(\mathbf{x})) \left| \frac{\partial f(\mathbf{x})}{\partial \mathbf{x}^\top} \right|$$

Parametric PDE Flows: Tabak, Vanden-Eijnden (Commun. Math. Sci. 2010)
NICE: Dinh, Krueger Bengio (ICLR 2015)
RealNVP: Dinh, Sohl-Dickstein, Bengio (ICLR 2016)
Normalizing Flows: Rezende and Mohamed (ICML 2015)

Boltzmann Generators



Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

- Invertible transformation:

$$\mathbf{z} = F_{xz}(\mathbf{x};\theta)$$
$$\mathbf{x} = F_{zx}(\mathbf{z};\theta).$$

- Jacobians:

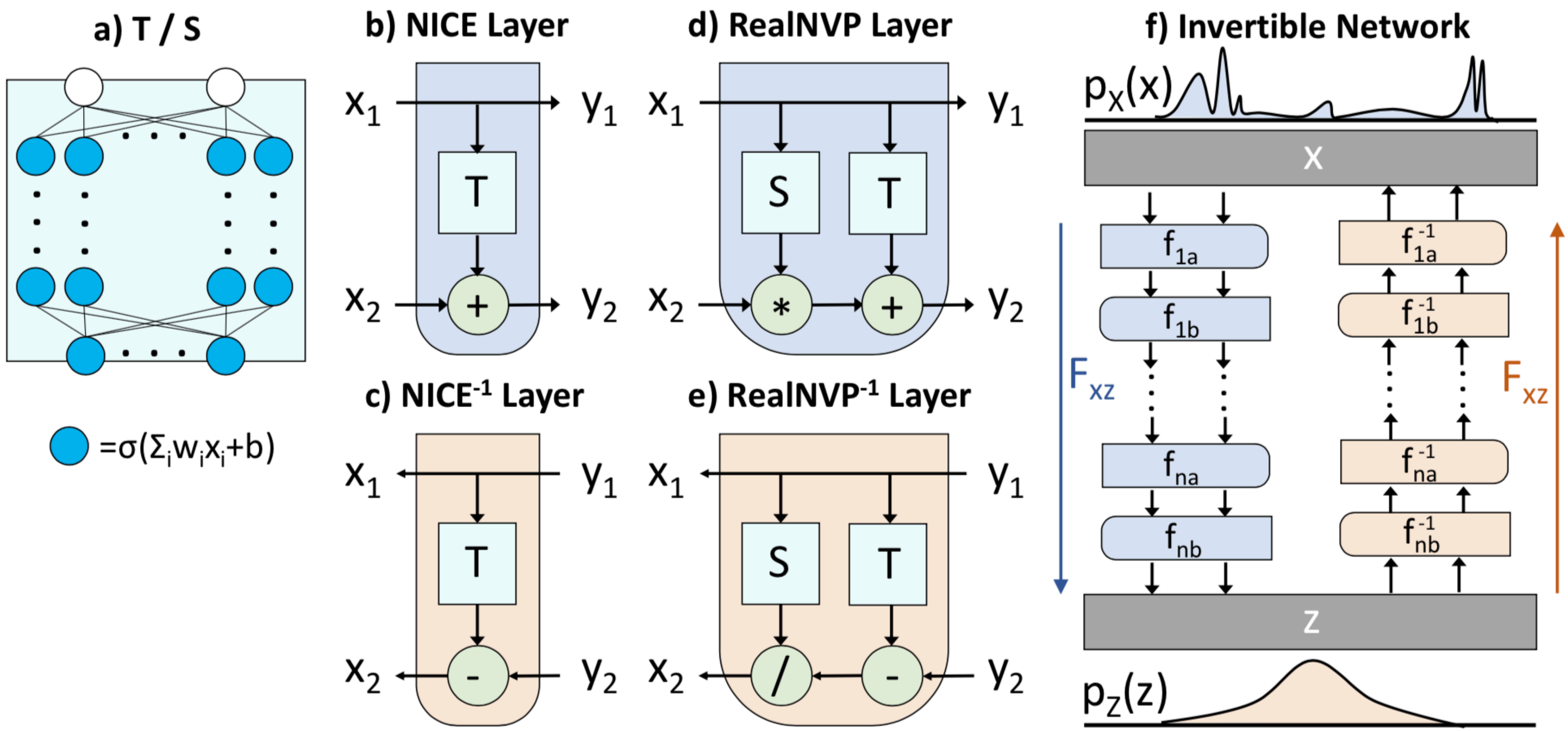
$$\mathbf{J}_{zx}(\mathbf{z};\theta) = \left[\frac{\partial F_{zx}(\mathbf{z};\theta)}{\partial z_1}, \ldots, \frac{\partial F_{zx}(\mathbf{z};\theta)}{\partial z_n}\right]$$
$$\mathbf{J}_{xz}(\mathbf{x};\theta) = \left[\frac{dF_{xz}(\mathbf{x};\theta)}{dx_1}, \ldots, \frac{dF_{xz}(\mathbf{x};\theta)}{dx_n}\right]$$

- Transformation of random variables:

$$p_X(\mathbf{x}) = p_Z(T_{xz}(\mathbf{x}))\left|\det \mathbf{J}_{xz}(\mathbf{x})\right|$$
$$p_Z(\mathbf{z}) = p_X(T_{zx}(\mathbf{z}))\left|\det \mathbf{J}_{zx}(\mathbf{z})\right|$$

Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Boltzmann Generators



- **NICE**: Dinh, Krueger, Y. Bengio, ICLR 2015
- **RealNVP**: Dinh, Sohl-Dickstein, S. Bengio, ICLR 2017
- More advanced flow operations (e.g., normalizing flows)

Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

- **Distributions**:

$$\text{Prior} \quad \mu_Z(\mathbf{z}) \quad \xrightarrow{F_{zx}} \quad p_X(\mathbf{x}) \quad \text{generated}$$
$$\text{Boltzmann} \quad \mu_X(\mathbf{x}) \quad \xrightarrow{F_{xz}} \quad p_Z(\mathbf{z}) \quad \text{generated}$$

- **Aim**: sample configurations $\mathbf{x}$ from **Boltzmann distribution**

$$\mu_X(\mathbf{x}) = Z_X^{-1} \mathrm{e}^{-u(\mathbf{x})}$$

- **Prior distribution**: Sample input in $\mathbf{z}$ from isotropic Gaussian:

$$q_Z(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}) = Z_Z^{-1} \mathrm{e}^{-\frac{1}{2}\|\mathbf{z}\|^2/\sigma^2},$$

Prior energy:

$$u_Z(\mathbf{z}) = -\log q_Z(\mathbf{z}) = \frac{1}{2\sigma^2}\|\mathbf{z}\|^2 + \text{const}$$

Boltzmann Generators

- Loss function:

$$J = \underbrace{w_{ML} J_{ML}}_{\text{max likelihood}} + \underbrace{w_{KL} J_{KL}}_{\text{Kullback–Leibler}} + \underbrace{w_{RC} J_{RC}}_{\text{reaction coordinate}}.$$

- KL divergence between generated $p_X(\mathbf{x})$ and Boltzmann distribution:

$$\begin{aligned} \mathrm{KL}_\theta\left[\mu_Z \,\|\, p_Z\right] &= \int \mu_Z(\mathbf{z}) \left[\log \mu_Z(\mathbf{z}) - \log p_Z(\mathbf{z};\theta)\right] \mathrm{d}\mathbf{z}, \\ &= \underbrace{-H_Z + \log Z_X}_{\text{const}} + \underbrace{\mathbb{E}_{\mathbf{z}\sim\mu_Z(\mathbf{z})}\left[u(F_{zx}(\mathbf{z};\theta)) - \log\left|\mathbf{J}_{zx}(\mathbf{z};\theta)\right|\right]}_{\text{free energy}} \end{aligned}$$

- KL loss:

$$J_{KL} = \mathbb{E}_{\mathbf{z}\sim\mu_Z(\mathbf{z})}\left[u(F_{zx}(\mathbf{z};\theta)) - \log\left|\mathbf{J}_{zx}(\mathbf{z};\theta)\right|\right]. \tag{5}$$

Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Boltzmann Generators

- Loss function:

$$J = \underbrace{w_{ML} J_{ML}}_{\text{max likelihood}} + \underbrace{w_{KL} J_{KL}}_{\text{Kullback-Leibler}} + \underbrace{w_{RC} J_{RC}}_{\text{reaction coordinate}}.$$

- KL divergence between generated distribution $q_Z(\mathbf{z})$ and Gaussian:

$$\mathrm{KL}_\theta(\mu_X \parallel p_X) = -H_X - \int \mu_X(\mathbf{x}) \log p_X(\mathbf{x};\theta) \mathrm{d}\mathbf{x}$$
$$= \underbrace{-H_X + \log Z_Z}_{\text{const}} + \mathbb{E}_{\mathbf{x}\sim\mu(\mathbf{x})}\left[\frac{1}{\sigma^2}\left\|F_{xz}(\mathbf{x};\theta)\right\|^2 - \log\left|\mathbf{J}_{xz}(\mathbf{x};\theta)\right|\right]$$

Problem: sampling $\mathbf{x} \sim \mu(\mathbf{x})$ is difficult and our goal!

- **Maximum Likelihood** loss for initial data distribution $\rho(\mathbf{x})$:

$$J_{ML} = -\mathbb{E}_{\mathbf{x}\sim\rho(\mathbf{x})}\left[\log p_X(\mathbf{x};\theta)\right]$$
$$= \mathbb{E}_{\mathbf{x}\sim\rho(\mathbf{x})}\left[\frac{1}{\sigma^2}\left\|F_{xz}(\mathbf{x};\theta)\right\|^2 - \log\left|\mathbf{J}_{xz}(\mathbf{x};\theta)\right|\right]$$

Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Boltzmann Generators

- Loss function:

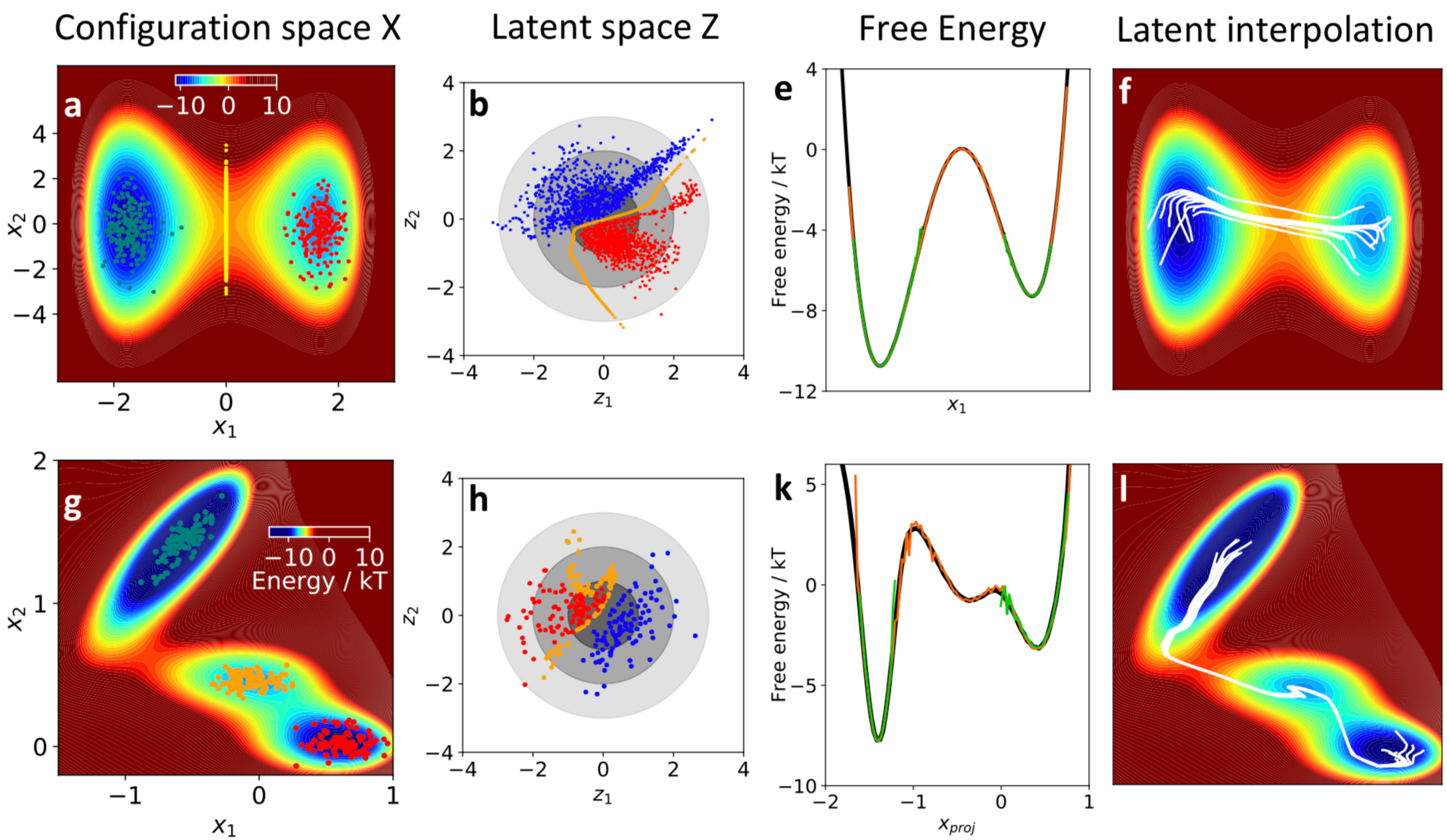
$$J = \underbrace{w_{ML} J_{ML}}_{\text{max likelihood}} + \underbrace{w_{KL} J_{KL}}_{\text{Kullback-Leibler}} + \underbrace{w_{RC} J_{RC}}_{\text{reaction coordinate}}.$$

- Reaction coordinate loss:

$$\begin{aligned} J_{RC} &= \int p(R(\mathbf{x})) \log p(R(\mathbf{x}))\, \mathrm{d}R(\mathbf{x}) \\ &= \mathbb{E}_{\mathbf{x} \sim p_X(\mathbf{x})} \log p(R(\mathbf{x})). \end{aligned}$$

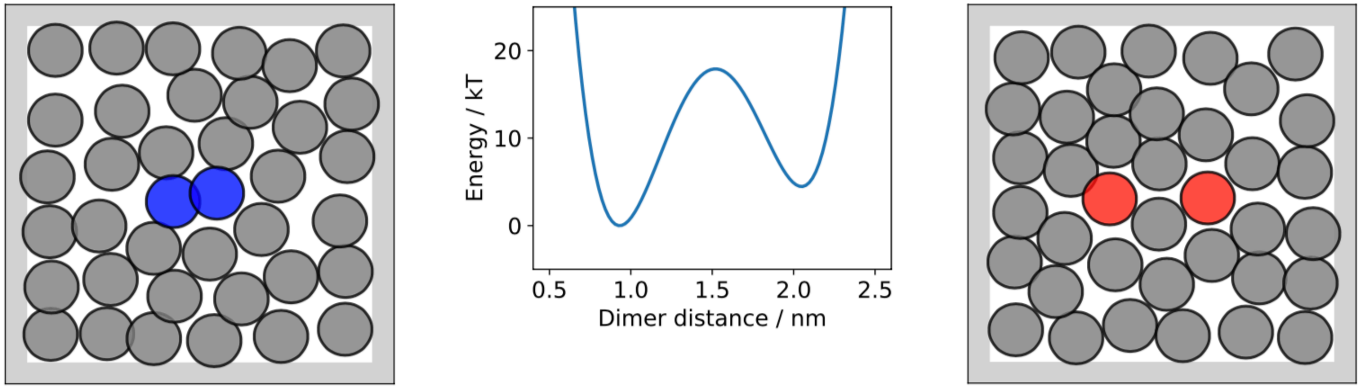
- Implementation:
    - Reaction coordinate function R is user input
    - Min and max bounds are given
    - $p(R(\mathbf{x}))$ is computed as a batchwise kernel density estimate.

Boltzmann Generators: Double well and Mueller potential

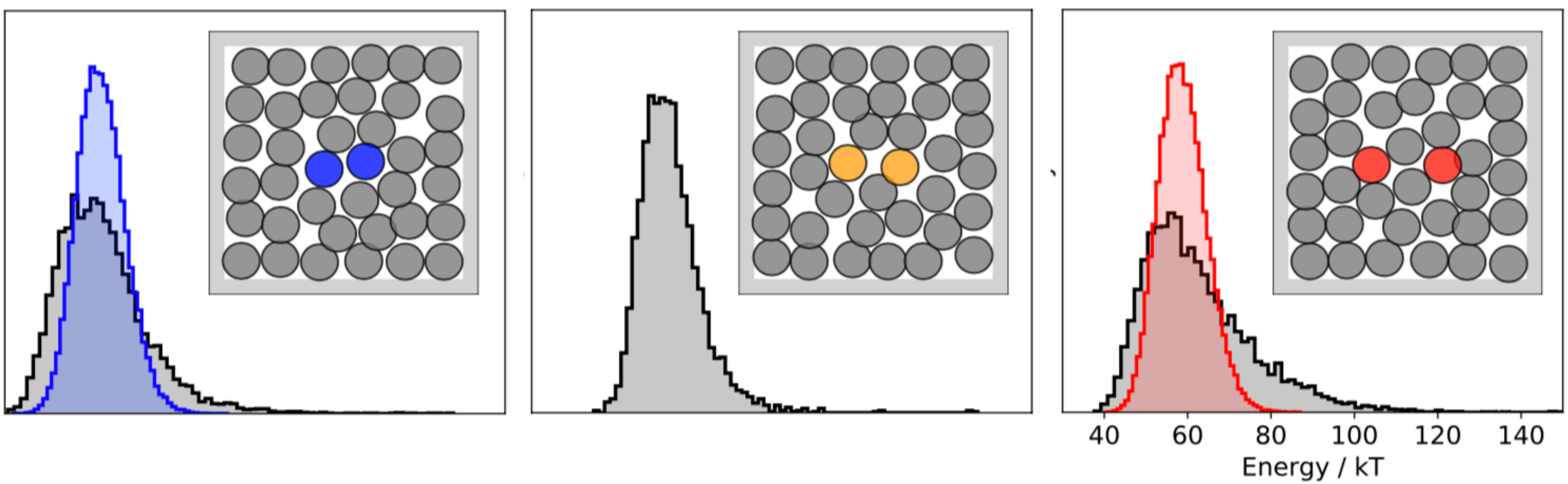


Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Boltzmann Generators: Bistable dimer in dense colloid solvent

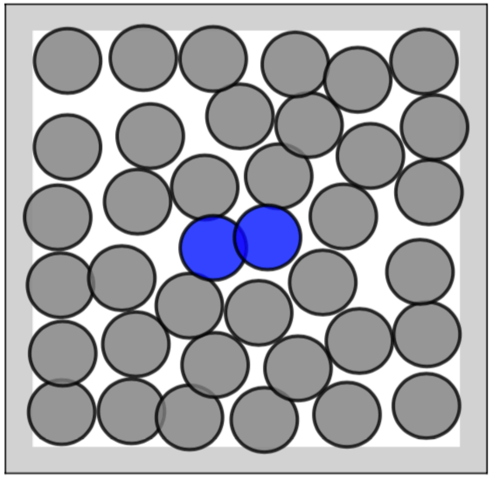


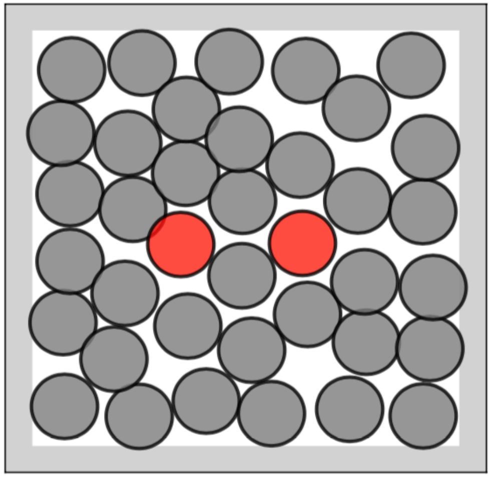
Boltzmann-Generated samples:

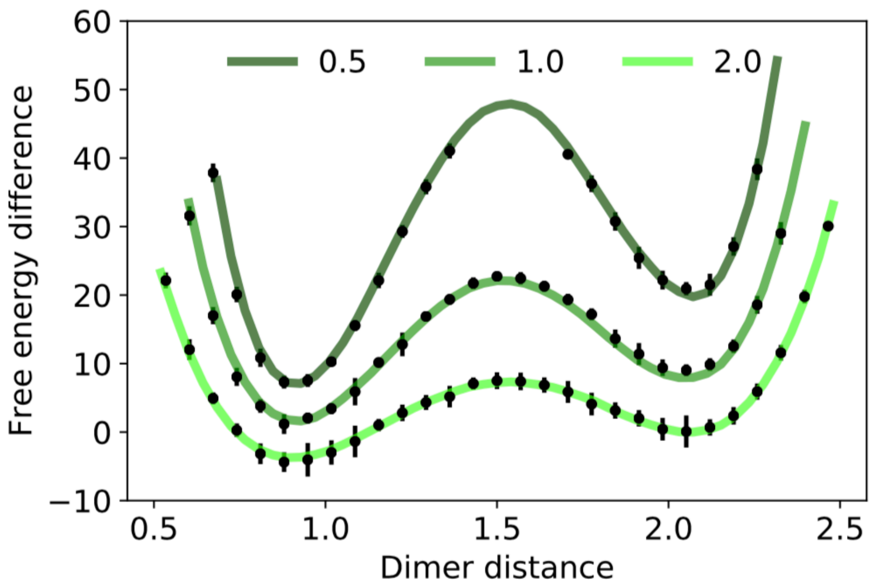


Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Boltzmann Generators: Bistable dimer in dense colloid solvent



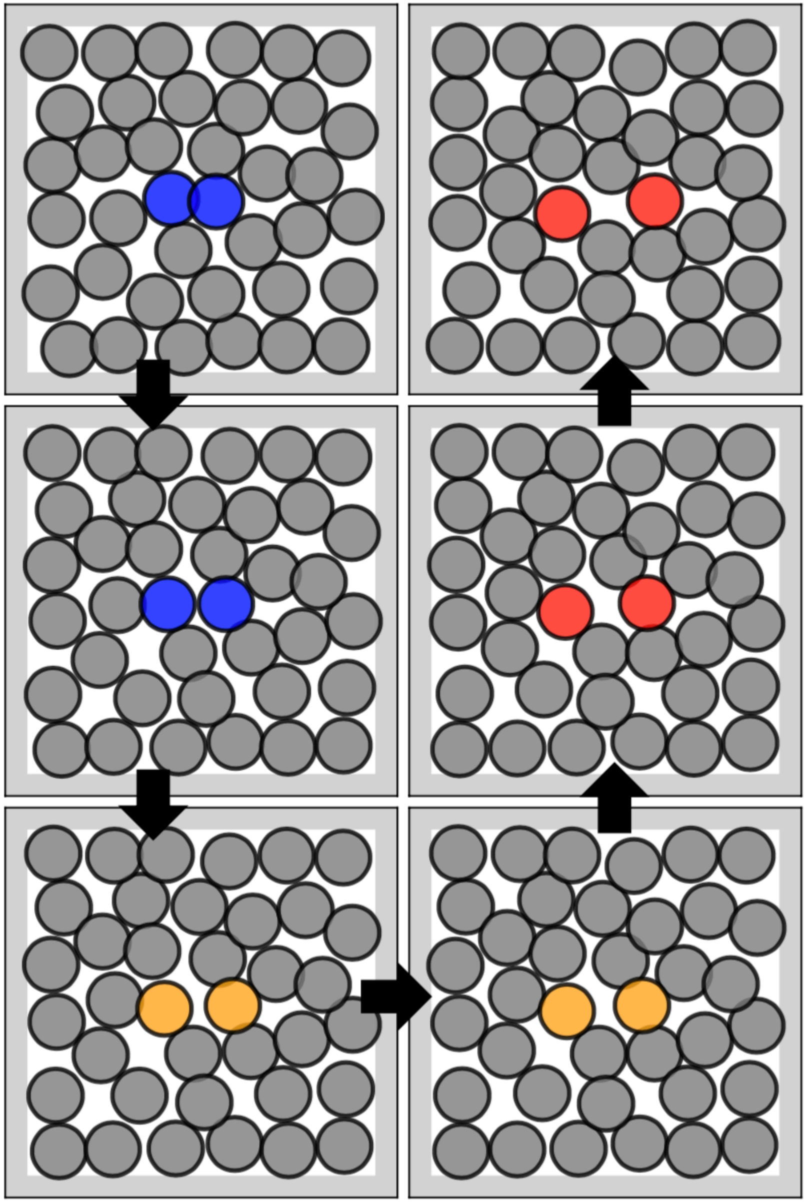




Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Freie Universität Berlin

Boltzmann Generators: Bistable dimer in dense colloid solvent



Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Boltzmann Generators: Exploration

1. Sample batch $\{\mathbf{x}_1, ..., \mathbf{x}_B\}$ from X.

2. Update Boltzmann Generator parameters $\boldsymbol{\theta}$ by training on batch.

3. For each $\mathbf{x}$ in batch, propose a Metropolis Monte Carlo step in latent space with step-size s:

$$\mathbf{z}' = T_{xz}(\mathbf{x}) + s\mathcal{N}(\mathbf{0}, \mathbf{I}).$$

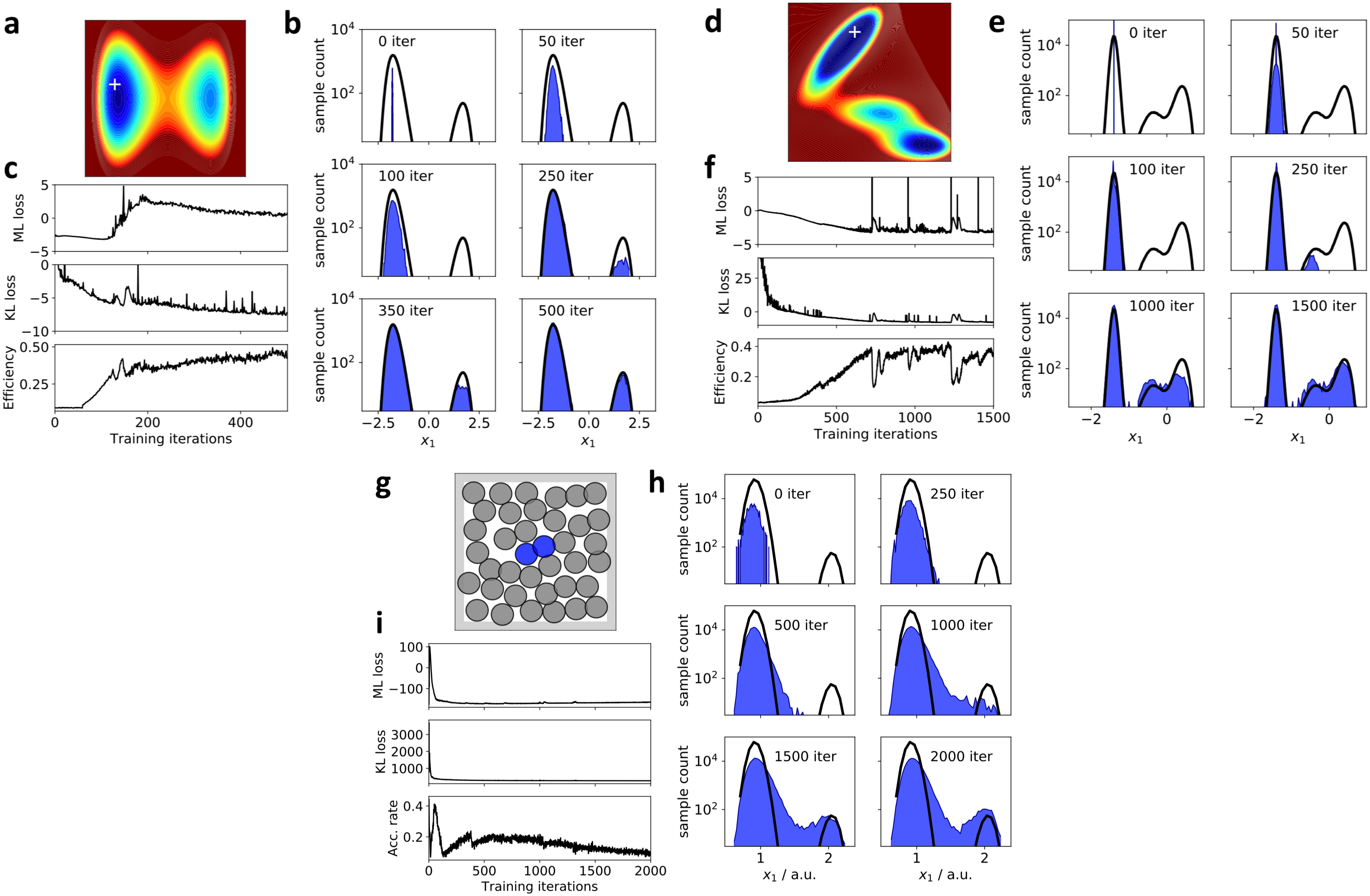
4. Accept or reject proposal with probability $\min\{1, \exp(-\Delta E)\}$ using the energy difference:

$$\Delta E = u(T_{zx}(\mathbf{z}')) - u(\mathbf{x}) - \log R_{zx}(\mathbf{z}'; \boldsymbol{\theta}) + \log R_{xz}(\mathbf{x}; \boldsymbol{\theta})$$

   For the accepted samples, replace $\mathbf{x}$ by $\mathbf{x}' = T_{zx}(\mathbf{z}')$.

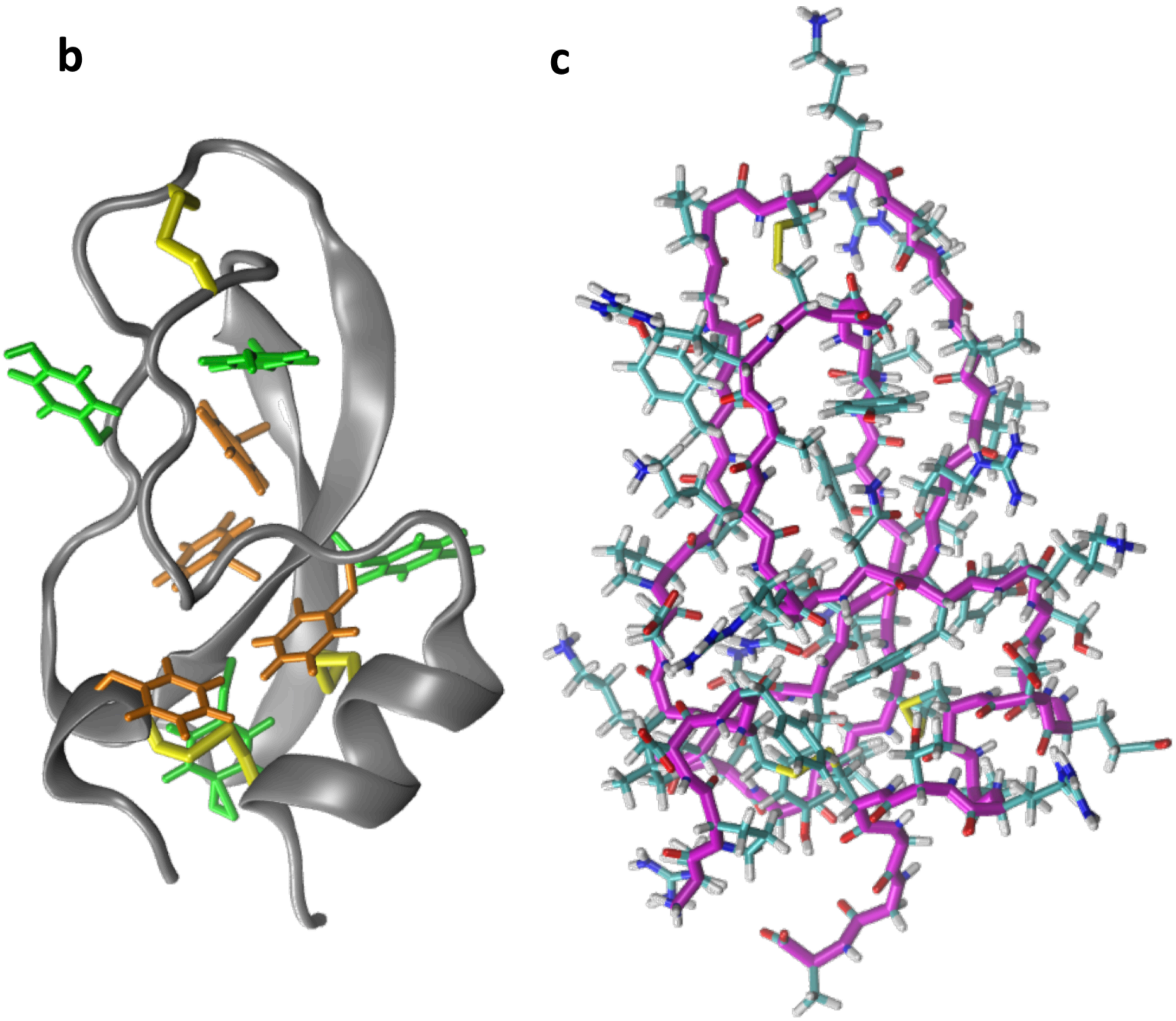
Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Boltzmann Generators: Exploration



Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

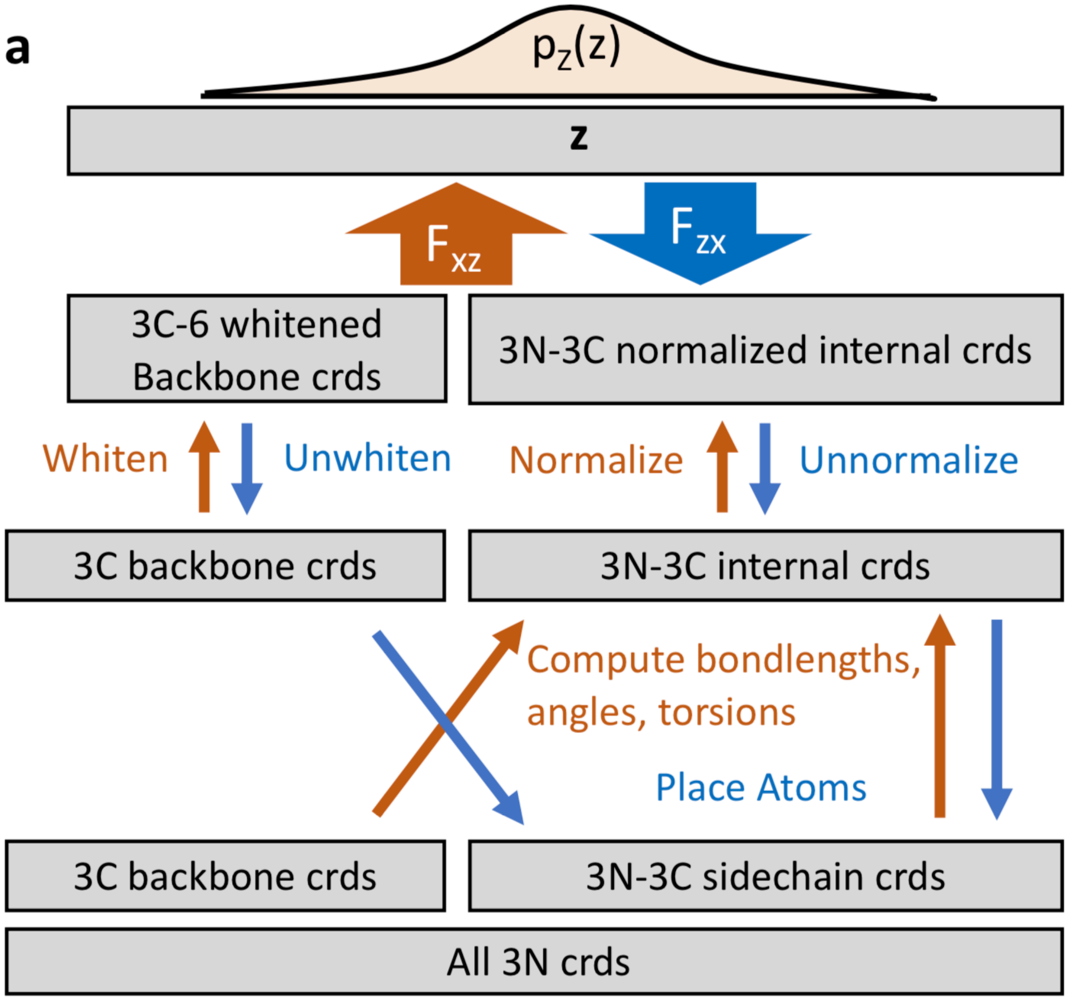
Boltzmann Generators: towards proteins



Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

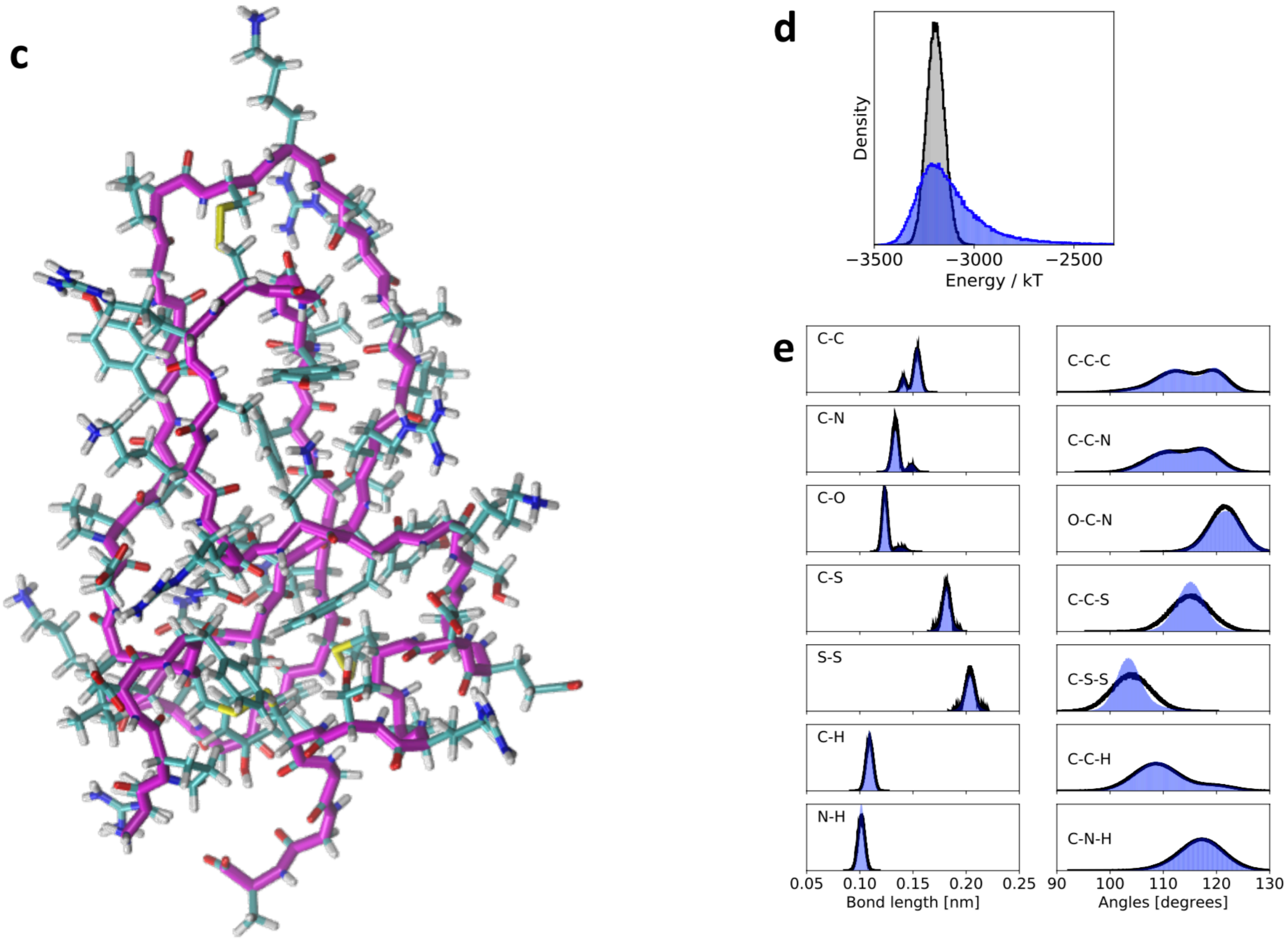
Freie Universität Berlin

Boltzmann Generators: proteins



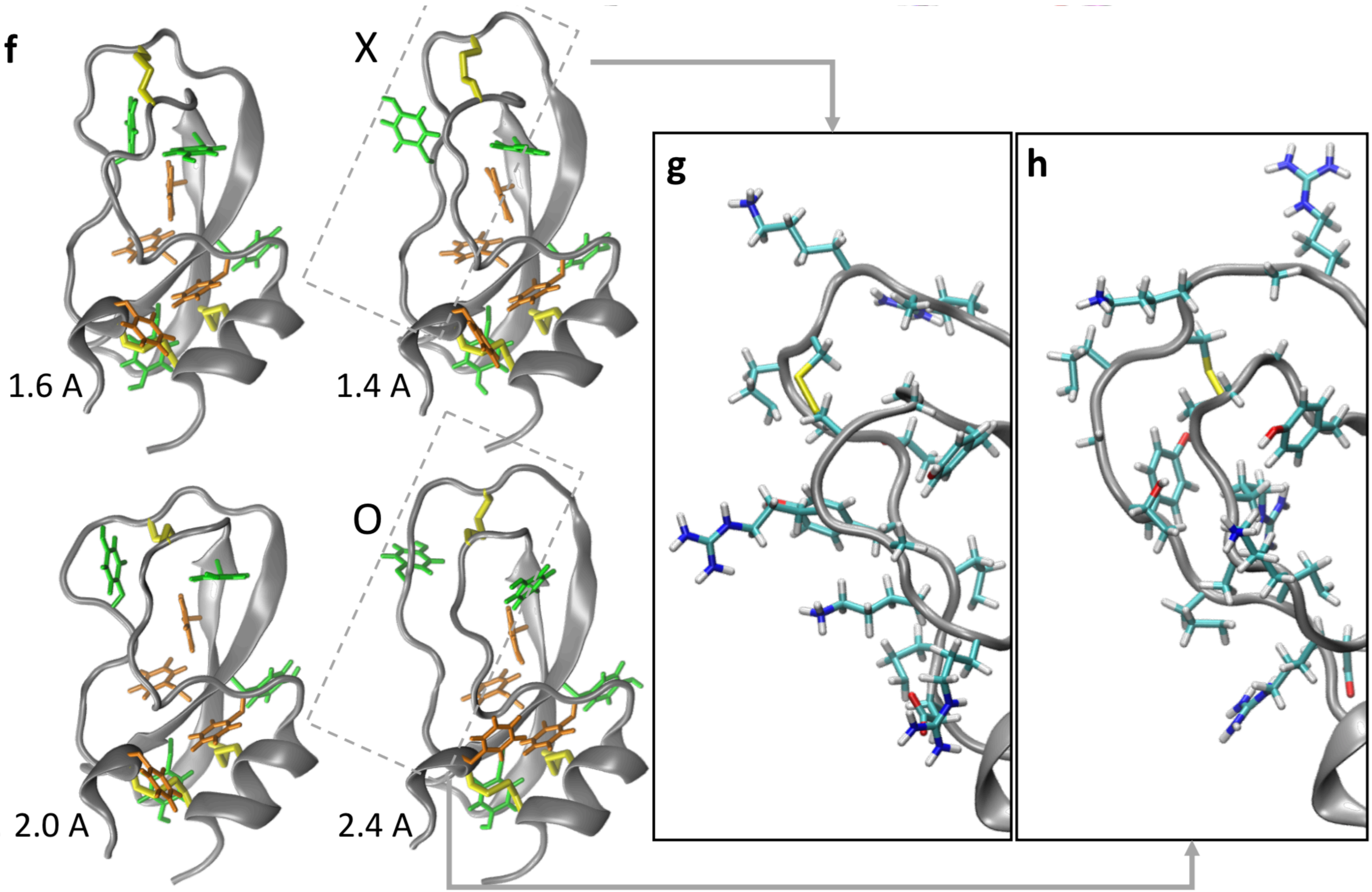
Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Boltzmann Generators: proteins



Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Boltzmann Generators: proteins



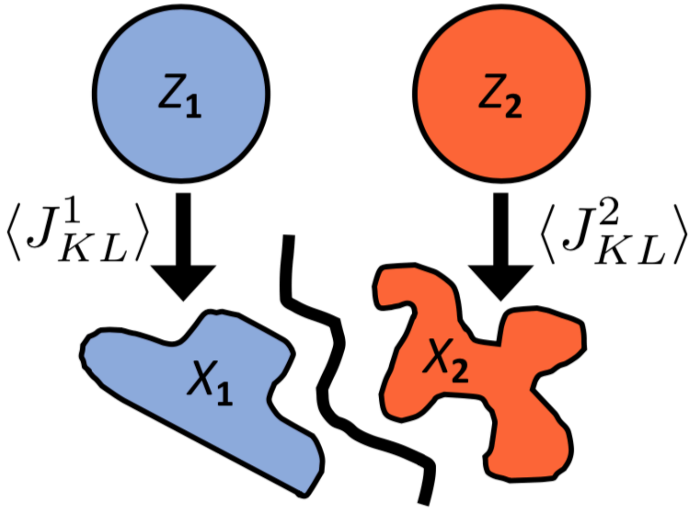
Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

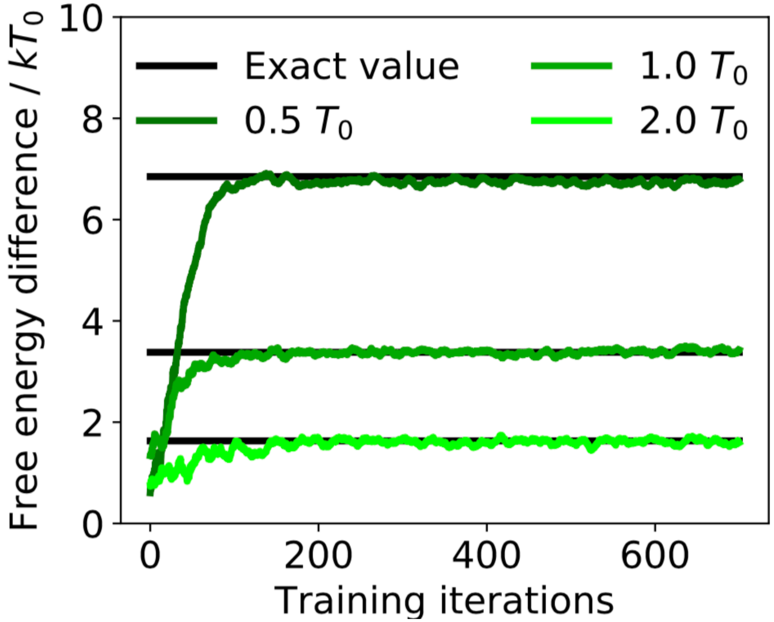
Boltzmann Generators: free energy differences

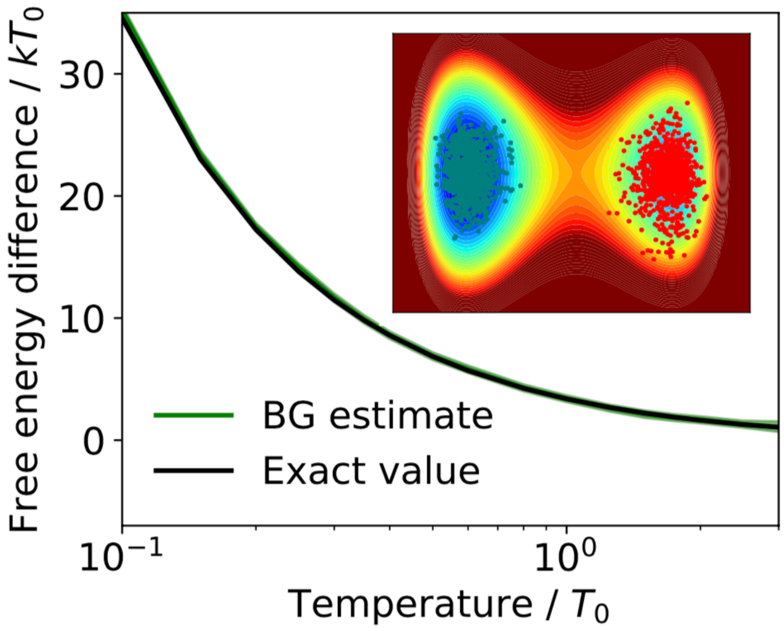
$$J_{KL} = \mathbb{E}_{\mathbf{z}}\left[u(F_{zx}(\mathbf{z})) - \log R_{zx}(\mathbf{z})\right]$$

Free Energy difference from two independent Boltzmann Generators

$$\Delta A_{12} = \langle J_{KL}^2 \rangle - \langle J_{KL}^1 \rangle$$



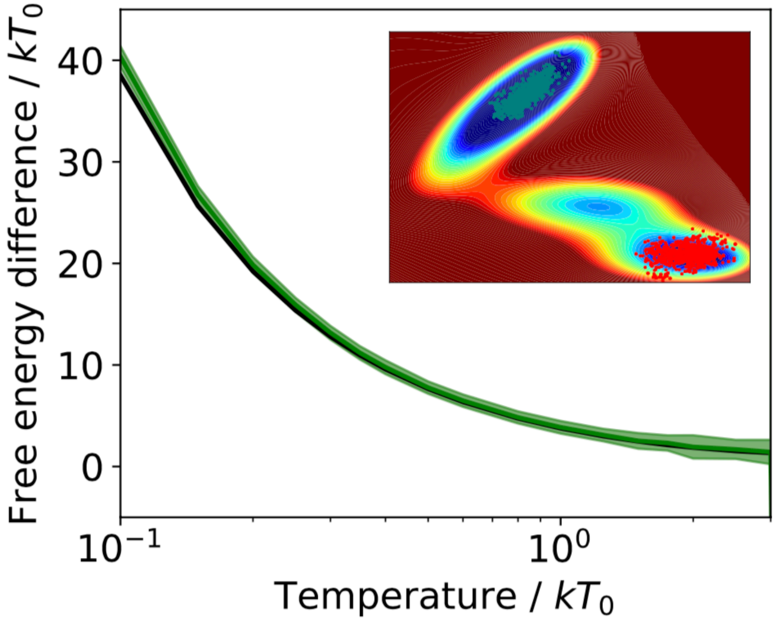


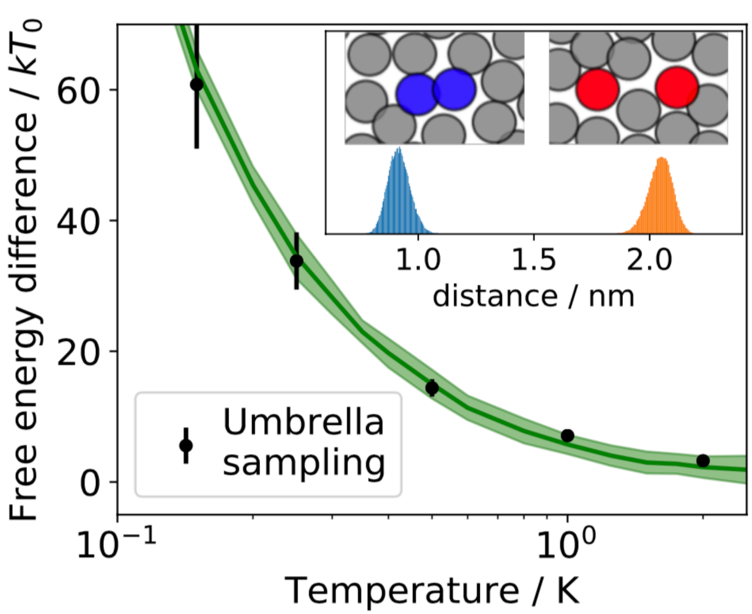


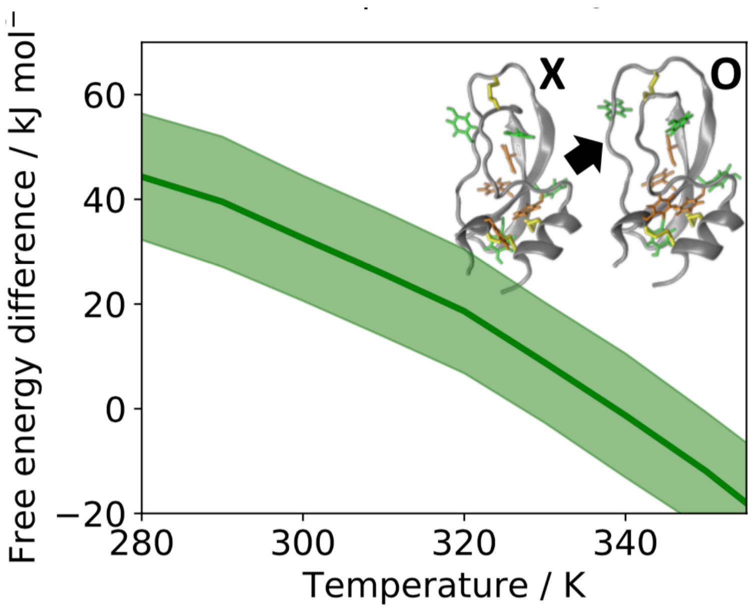
Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Freie Universität Berlin

Boltzmann Generators: free energy differences





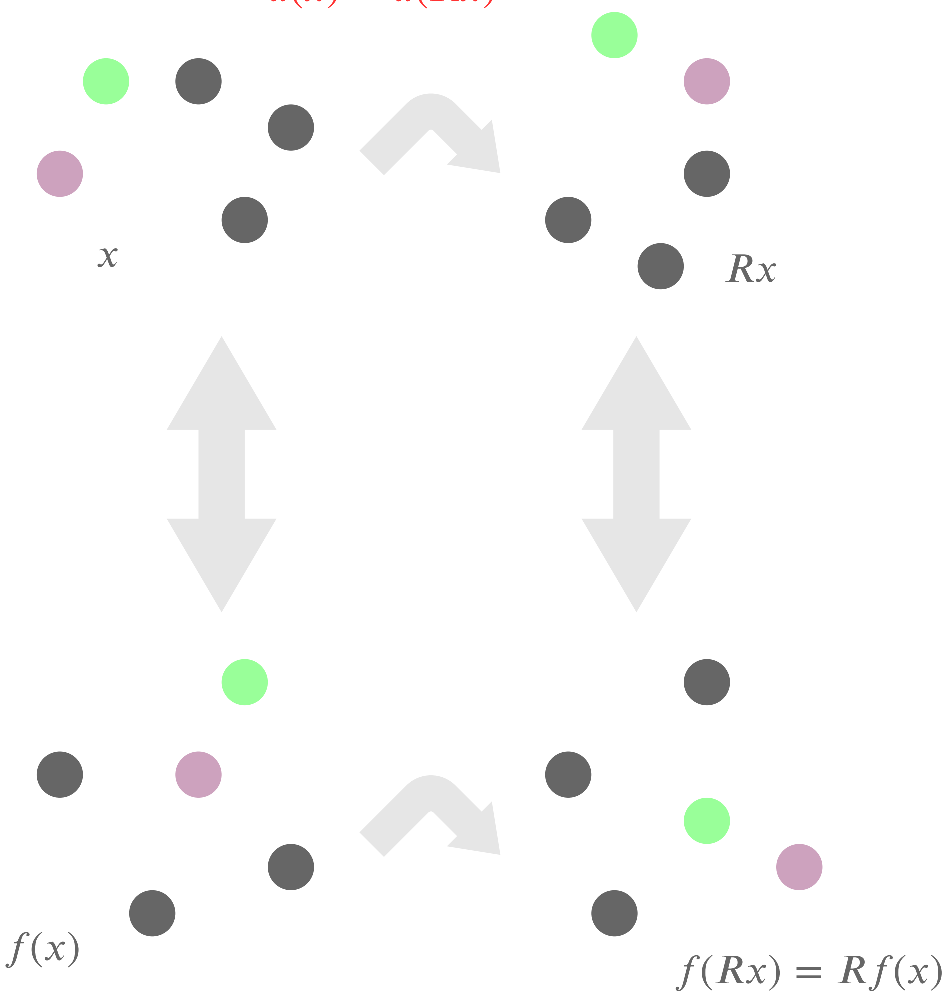


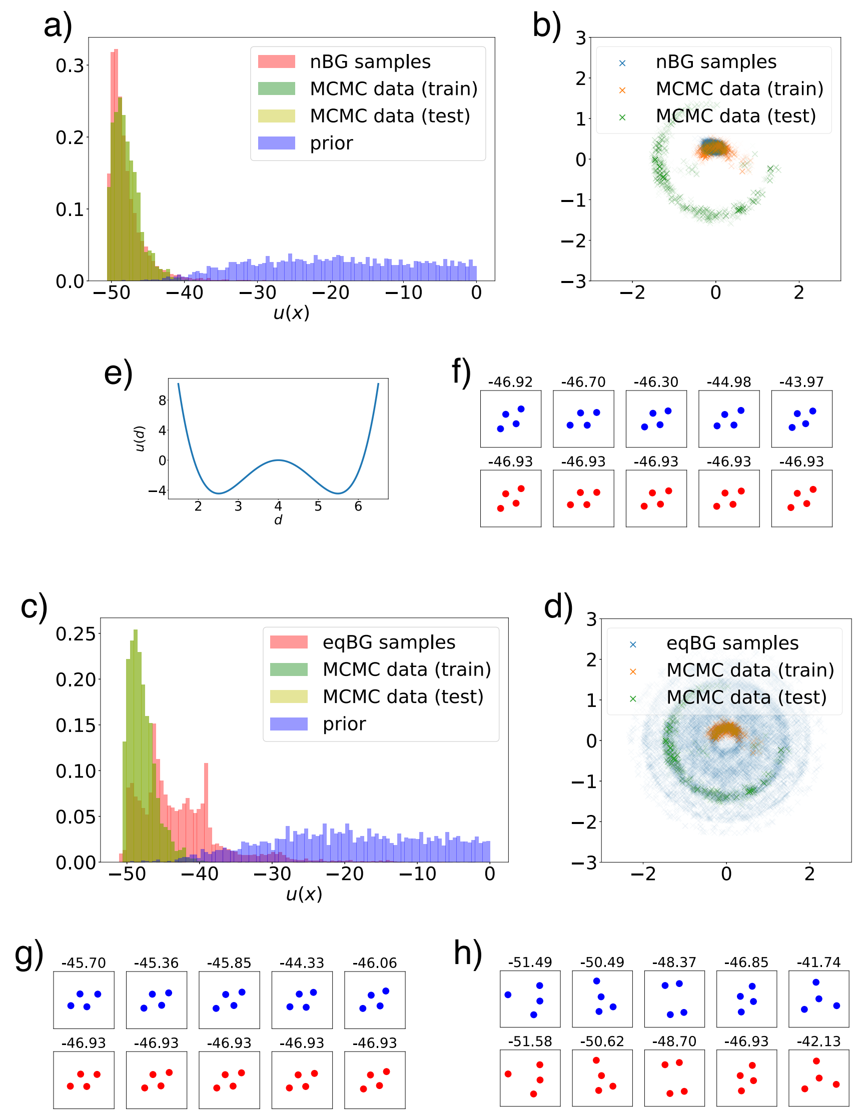
Noé, Olsson, Köhler, Wu, **Science** 365: eaww1147 (2019). arXiv:1812.01729

Freie Universität Berlin

Equivariant Boltzmann Generators

$u(x) = u(Rx)$





$\log p(f(x) = \log p(f(Rx)) = \log p(R(fx))$

Köhler, Klein, Noé, **in prep.**

PauliNet: Deep neural network solution of the electronic Schrödinger equation

Jan Hermann, Zeno Schätzle and Frank Noé
arXiv:1909.08423

September 22, 2019

Freie Universität Berlin

- **Electronic Schrödinger Equation**:

$$\hat{H}\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N) = E\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)$$

  - $\mathbf{r} = (\mathbf{r}_1,\ldots,\mathbf{r}_N)^\top \in \mathbb{R}^{3N}$ spatial electron coordinates
  - E electronic energy, ψ electronic wave function.

- **Hamiltonian** operator:

$$\hat{H} := \sum_i \left(-\tfrac{1}{2}\nabla^2_{\mathbf{r}_i} - \sum_I \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \right) + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

  - Z_I nuclear charges
  - $\mathbf{R}_I$ fixed nuclear coordinates (Born--Oppenheimer approximation).

- **Electronic spins** $s_i \in \{\uparrow,\downarrow\}$.
  ψ must be antisymmetric wrt exchange of equal-spin electrons

$$\psi(\ldots,\mathbf{r}_i,\ldots,\mathbf{r}_j,\ldots) = -\psi(\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_i,\ldots)$$

- **Electronic Schrödinger Equation**:

$$\hat{H}\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N) = E\psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)$$

  - $\mathbf{r} = (\mathbf{r}_1,\ldots,\mathbf{r}_N)^\top \in \mathbb{R}^{3N}$ spatial electron coordinates
  - E electronic energy, ψ electronic wave function.

- **Hamiltonian** operator:

$$\hat{H} := \sum_i \left(-\tfrac{1}{2}\nabla^2_{\mathbf{r}_i} - \sum_I \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \right) + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

  - Z_I nuclear charges
  - $\mathbf{R}_I$ fixed nuclear coordinates (Born--Oppenheimer approximation).

- **Electronic spins** $s_i \in \{\uparrow,\downarrow\}$.
  ψ must be antisymmetric wrt exchange of equal-spin electrons

$$\psi(\ldots,\mathbf{r}_i,\ldots,\mathbf{r}_j,\ldots) = -\psi(\ldots,\mathbf{r}_j,\ldots,\mathbf{r}_i,\ldots)$$

- **Electronic Schrödinger Equation**:

$$\hat{H}\psi(\mathbf{r}_1,\dots,\mathbf{r}_N) = E\psi(\mathbf{r}_1,\dots,\mathbf{r}_N)$$

  - $\mathbf{r} = (\mathbf{r}_1,\dots,\mathbf{r}_N)^\top \in \mathbb{R}^{3N}$ spatial electron coordinates
  - E electronic energy, ψ electronic wave function.

- **Hamiltonian** operator:

$$\hat{H} := \sum_i \left(-\tfrac{1}{2}\nabla^2_{\mathbf{r}_i} - \sum_I \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} \right) + \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

  - Z_I nuclear charges
  - $\mathbf{R}_I$ fixed nuclear coordinates (Born--Oppenheimer approximation).

- **Electronic spins** $s_i \in \{\uparrow,\downarrow\}$.
  ψ must be antisymmetric wrt exchange of equal-spin electrons

$$\psi(\dots,\mathbf{r}_i,\dots,\mathbf{r}_j,\dots) = -\psi(\dots,\mathbf{r}_j,\dots,\mathbf{r}_i,\dots)$$

Variational Monte Carlo

- **Variational Approach** (used in Hartree-Fock, DMRG, QMC)

$$E_0 = \min_{\psi} E[\psi] \leq \min_{\theta} E[\psi(\mathbf{r};\theta)] \qquad E[\psi] = \int \mathrm{d}\mathbf{r}\, \psi(\mathbf{r}) \hat{H} \psi(\mathbf{r})$$

- **VMC**: Variational Quantum Monte Carlo:

$$E[\psi;\theta] = \mathbb{E}_{\mathbf{r}\sim|\psi|^2}\left[E_{\mathrm{loc}}[\psi](\mathbf{r};\theta)\right] \qquad E_{\mathrm{loc}}[\psi](\mathbf{r};\theta) = \hat{H}\psi(\mathbf{r};\theta)/\psi(\mathbf{r};\theta)$$

- **Deep VMC idea**: use deep neural network to represent $\psi(\mathbf{r};\theta)$, minimize variational energy $E[\psi;\theta]$ over θ.
  - G. Carleo, M. Troyer, Science 355:602-606 (2017): Discrete lattice systems, Restricted Boltzmann Machines.
  - **DeepWF**: J. Han et al, arXiv:1807.07014 (2018): Atoms+Molecules, Poor accuracy, moderate computational cost.
  - **FermiNet**: D. Pfau et al, arXiv:1909.02487 (Sept 2019): Atoms+Molecules, highest accuracy, insane computational cost.
  - **PauliNet**: This work, arXiv:1909.08423 (Sept 2019): Atoms+Molecules, high accuracy, moderate computational cost.

Variational Monte Carlo

- **Variational Approach** (used in Hartree-Fock, DMRG, QMC)

$$E_0 = \min_{\psi} E[\psi] \leq \min_{\theta} E[\psi(\mathbf{r};\theta)] \qquad E[\psi] = \int \mathrm{d}\mathbf{r}\, \psi(\mathbf{r}) \hat{H} \psi(\mathbf{r})$$

- **VMC**: Variational Quantum Monte Carlo:

$$E[\psi;\theta] = \mathbb{E}_{\mathbf{r}\sim|\psi|^2}\left[E_{\mathrm{loc}}[\psi](\mathbf{r};\theta)\right] \qquad E_{\mathrm{loc}}[\psi](\mathbf{r};\theta) = \hat{H}\psi(\mathbf{r};\theta)/\psi(\mathbf{r};\theta)$$

- **Deep VMC idea**: use deep neural network to represent $\psi(\mathbf{r};\theta)$, minimize variational energy $E[\psi;\theta]$ over θ.
  - G. Carleo, M. Troyer, Science 355:602-606 (2017): Discrete lattice systems, Restricted Boltzmann Machines.
  - **DeepWF**: J. Han et al, arXiv:1807.07014 (2018): Atoms+Molecules, Poor accuracy, moderate computational cost.
  - **FermiNet**: D. Pfau et al, arXiv:1909.02487 (Sept 2019): Atoms+Molecules, highest accuracy, insane computational cost.
  - **PauliNet**: This work, arXiv:1909.08423 (Sept 2019): Atoms+Molecules, high accuracy, moderate computational cost.

Variational Monte Carlo

- **Variational Approach** (used in Hartree-Fock, DMRG, QMC)

$$E_0 = \min_{\psi} E[\psi] \leq \min_{\theta} E[\psi(\mathbf{r};\theta)] \qquad E[\psi] = \int \mathrm{d}\mathbf{r}\, \psi(\mathbf{r}) \hat{H} \psi(\mathbf{r})$$

- **VMC**: Variational Quantum Monte Carlo:

$$E[\psi;\theta] = \mathbb{E}_{\mathbf{r}\sim|\psi|^2}\left[E_{\mathrm{loc}}[\psi](\mathbf{r};\theta)\right] \qquad E_{\mathrm{loc}}[\psi](\mathbf{r};\theta) = \hat{H}\psi(\mathbf{r};\theta)/\psi(\mathbf{r};\theta)$$

- **Deep VMC idea**: use deep neural network to represent $\psi(\mathbf{r};\theta)$, minimize variational energy $E[\psi;\theta]$ over θ.
  - G. Carleo, M. Troyer, Science 355:602-606 (2017): Discrete lattice systems, Restricted Boltzmann Machines.
  - **DeepWF**: J. Han et al, arXiv:1807.07014 (2018): Atoms+Molecules, Poor accuracy, moderate computational cost.
  - **FermiNet**: D. Pfau et al, arXiv:1909.02487 (Sept 2019): Atoms+Molecules, highest accuracy, insane computational cost.
  - **PauliNet**: This work, arXiv:1909.08423 (Sept 2019): Atoms+Molecules, high accuracy, moderate computational cost.

PauliNet Wavefunction:

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

with:

- $\varphi_\mu(\mathbf{r}_i)$: one-electron molecular orbitals (fixed)
- $J(\mathbf{r};\theta)$: Jastrow factor neural network (trainable)
- $\mathbf{f}(\mathbf{r};\theta)$: Backflow neural network (trainable)
- $\gamma(\mathbf{r})$: Asymptotics (fixed)

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- Assign electrons to $\uparrow$ and $\downarrow$ sets.

$$\psi(\mathbf{r}_1^\uparrow, \ldots, \mathbf{r}_{N_\uparrow}^\uparrow, \mathbf{r}_{N_\uparrow+1}^\downarrow, \ldots, \mathbf{r}_N^\downarrow) \equiv \psi(\mathbf{r}^\uparrow, \mathbf{r}^\downarrow) \equiv \psi(\mathbf{r})$$

- **Antisymmetry** for **single-electron** functions: Slater determinants.

$$\psi_{\mathrm{HF}}(\mathbf{r}) := \det[\varphi_\mu(\mathbf{r}_i^\uparrow)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow)]$$

- **Antisymmetry** for **many-electron** functions $\varphi_\mu(\mathbf{r}_i) \to \varphi_{\mu i}(\mathbf{r})$ holds if $\varphi_{\mu i}(\mathbf{r})$ is equivariant wrt exchange $\mathscr{P}$ of same-spin electrons[1,2]

$$\varphi_\mu(\mathbf{r}_i) \to \varphi_{\mu i}(\mathbf{r}),$$
$$\mathscr{P}\varphi_{\mu i}(\mathbf{r}) = \varphi_{\mu i}(\mathscr{P}\mathbf{r})$$

- **PauliNet**: $\varphi_{\mu i}(\mathbf{r}) = \varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)$, $f_{\mu i}$ equivariant – SchNet[3].

[1] P. López Ríos et al, *Phys. Rev. E*, 74:066701(2006)
[2] D. Luo, B. K. Clark. *Phys. Rev. Lett.* 122:226401 (2019)
[3] K. T. Schütt et al., *J. Chem. Phys.* 148:241722 (2018)

PauliNet – Physics 1: Antisymmetry

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- Assign electrons to ↑ and ↓ sets.

$$\psi(\mathbf{r}_1^\uparrow,\ldots,\mathbf{r}_{N_\uparrow}^\uparrow,\mathbf{r}_{N_\uparrow+1}^\downarrow,\ldots,\mathbf{r}_N^\downarrow) \equiv \psi(\mathbf{r}^\uparrow,\mathbf{r}^\downarrow) \equiv \psi(\mathbf{r})$$

- **Antisymmetry** for **single-electron** functions: Slater determinants.

$$\psi_{\mathrm{HF}}(\mathbf{r}) := \det[\varphi_\mu(\mathbf{r}_i^\uparrow)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow)]$$

- **Antisymmetry** for **many-electron** functions $\varphi_\mu(\mathbf{r}_i) \to \varphi_{\mu i}(\mathbf{r})$ holds if $\varphi_{\mu i}(\mathbf{r})$ is equivariant wrt exchange $\mathscr{P}$ of same-spin electrons[1,2]

$$\varphi_\mu(\mathbf{r}_i) \to \varphi_{\mu i}(\mathbf{r}),$$
$$\mathscr{P}\varphi_{\mu i}(\mathbf{r}) = \varphi_{\mu i}(\mathscr{P}\mathbf{r})$$

- **PauliNet**: $\varphi_{\mu i}(\mathbf{r}) = \varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)$, $f_{\mu i}$ equivariant – SchNet[3].

[1] P. López Ríos et al, *Phys. Rev. E*, 74:066701(2006)
[2] D. Luo, B. K. Clark. *Phys. Rev. Lett.* 122:226401 (2019)
[3] K. T. Schütt et al., *J. Chem. Phys.* 148:241722 (2018)

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)]\det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)]\mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- Assign electrons to $\uparrow$ and $\downarrow$ sets.

$$\psi(\mathbf{r}_1^\uparrow,\ldots,\mathbf{r}_{N_\uparrow}^\uparrow,\mathbf{r}_{N_\uparrow+1}^\downarrow,\ldots,\mathbf{r}_N^\downarrow) \equiv \psi(\mathbf{r}^\uparrow,\mathbf{r}^\downarrow) \equiv \psi(\mathbf{r})$$

- **Antisymmetry** for **single-electron** functions: Slater determinants.

$$\psi_{\mathrm{HF}}(\mathbf{r}) := \det[\varphi_\mu(\mathbf{r}_i^\uparrow)]\det[\varphi_\mu(\mathbf{r}_i^\downarrow)]$$

- **Antisymmetry** for **many-electron** functions $\varphi_\mu(\mathbf{r}_i) \to \varphi_{\mu i}(\mathbf{r})$ holds if $\varphi_{\mu i}(\mathbf{r})$ is equivariant wrt exchange $\mathscr{P}$ of same-spin electrons[1],[2]

$$\varphi_\mu(\mathbf{r}_i) \to \varphi_{\mu i}(\mathbf{r}),$$
$$\mathscr{P}\varphi_{\mu i}(\mathbf{r}) = \varphi_{\mu i}(\mathscr{P}\mathbf{r})$$

- **PauliNet**: $\varphi_{\mu i}(\mathbf{r}) = \varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)$, $f_{\mu i}$ equivariant – SchNet[3].

[1] P. López Ríos et al, *Phys. Rev. E*, 74:066701(2006)
[2] D. Luo, B. K. Clark. *Phys. Rev. Lett.* 122:226401 (2019)
[3] K. T. Schütt et al., *J. Chem. Phys.* 148:241722 (2018)

PauliNet – Physics 1: Antisymmetry

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r}) + J(\mathbf{r};\theta)}$$

- Assign electrons to $\uparrow$ and $\downarrow$ sets.

$$\psi(\mathbf{r}_1^\uparrow, \ldots, \mathbf{r}_{N_\uparrow}^\uparrow, \mathbf{r}_{N_\uparrow+1}^\downarrow, \ldots, \mathbf{r}_N^\downarrow) \equiv \psi(\mathbf{r}^\uparrow, \mathbf{r}^\downarrow) \equiv \psi(\mathbf{r})$$

- **Antisymmetry** for **single-electron** functions: Slater determinants.

$$\psi_{\mathrm{HF}}(\mathbf{r}) := \det[\varphi_\mu(\mathbf{r}_i^\uparrow)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow)]$$

- **Antisymmetry** for **many-electron** functions $\varphi_\mu(\mathbf{r}_i) \to \varphi_{\mu i}(\mathbf{r})$ holds if $\varphi_{\mu i}(\mathbf{r})$ is equivariant wrt exchange $\mathscr{P}$ of same-spin electrons[1,2]

$$\varphi_\mu(\mathbf{r}_i) \to \varphi_{\mu i}(\mathbf{r}),$$
$$\mathscr{P} \varphi_{\mu i}(\mathbf{r}) = \varphi_{\mu i}(\mathscr{P}\mathbf{r})$$

- **PauliNet**: $\varphi_{\mu i}(\mathbf{r}) = \varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)$, $f_{\mu i}$ equivariant – SchNet[3].

---

[1] P. López Ríos et al, *Phys. Rev. E*, 74:066701(2006)
[2] D. Luo, B. K. Clark. *Phys. Rev. Lett.* 122:226401 (2019)
[3] K. T. Schütt et al., *J. Chem. Phys.* 148:241722 (2018)

PauliNet – Physics 2: HF Baseline

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- Use fixed Hartree-Fock (HF) orbitals $\varphi_\mu(\mathbf{r}_i^\uparrow)$
- $\varphi_\mu(\mathbf{r}_i^\uparrow)$: linear combination of basis functions centered on atomic nuclei, coefficients optimized variationally.
- Pro: HF fast, captures physics of atoms and molecules qualitatively
- Con: poor quantitative predictions, no electron-electron correlations.

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- Use fixed Hartree-Fock (HF) orbitals $\varphi_\mu(\mathbf{r}_i^\uparrow)$
- $\varphi_\mu(\mathbf{r}_i^\uparrow)$: linear combination of basis functions centered on atomic nuclei, coefficients optimized variationally.
- Pro: HF fast, captures physics of atoms and molecules qualitatively
- Con: poor quantitative predictions, no electron-electron correlations.

PauliNet – Physics 2: HF Baseline

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- Use fixed Hartree-Fock (HF) orbitals $\varphi_\mu(\mathbf{r}_i^\uparrow)$
- $\varphi_\mu(\mathbf{r}_i^\uparrow)$: linear combination of basis functions centered on atomic nuclei, coefficients optimized variationally.
- Pro: HF fast, captures physics of atoms and molecules qualitatively
- Con: poor quantitative predictions, no electron-electron correlations.

PauliNet – Physics 2: HF Baseline

PauliNet Wavefunction:

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] e^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- Use fixed Hartree-Fock (HF) orbitals $\varphi_\mu(\mathbf{r}_i^\uparrow)$
- $\varphi_\mu(\mathbf{r}_i^\uparrow)$: linear combination of basis functions centered on atomic nuclei, coefficients optimized variationally.
- Pro: HF fast, captures physics of atoms and molecules qualitatively
- Con: poor quantitative predictions, no electron-electron correlations.

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] e^{\gamma(\mathbf{r}) + J(\mathbf{r};\theta)}$$

- **Jastrow factor**[4] $J(\mathbf{r};\theta)$: Antisymmetric part is multiplied by nonnegative symmetric function $e^{J(\mathbf{r};\theta)}$. Can capture complex electron correlations but not change the nodal surface.
- **Backflow**[5,6] $\mathbf{f}(\mathbf{r};\theta)$: Generalize one-electron orbitals $\varphi_\mu(\mathbf{r}_i)$ to many-electron orbitals $\varphi_{\mu i}(\mathbf{r})$. Can change the nodal surface.
- Exploit the flexibility of deep learning to achieve **higher accuracy**!

[4] N. D. Drummond et al. *Phys. Rev. B* 70:235119 (2004)
[5] P. López Ríos et al, *Phys. Rev. E*, 74:066701(2006)
[6] D. Luo, B. K. Clark. *Phys. Rev. Lett.* 122:226401 (2019)

PauliNet – Physics 3: Beyond HF

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- **Jastrow factor**[4] $J(\mathbf{r};\theta)$: Antisymmetric part is multiplied by nonnegative symmetric function $\mathrm{e}^{J(\mathbf{r};\theta)}$. Can capture complex electron correlations but not change the nodal surface.
- **Backflow**[5,6] $\mathbf{f}(\mathbf{r};\theta)$: Generalize one-electron orbitals $\varphi_\mu(\mathbf{r}_i)$ to many-electron orbitals $\varphi_{\mu i}(\mathbf{r})$. Can change the nodal surface.
- Exploit the flexibility of deep learning to achieve **higher accuracy**!

[4] N. D. Drummond et al. *Phys. Rev. B* 70:235119 (2004)
[5] P. López Ríos et al, *Phys. Rev. E*, 74:066701(2006)
[6] D. Luo, B. K. Clark. *Phys. Rev. Lett.* 122:226401 (2019)

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- **Jastrow factor**[4] $J(\mathbf{r};\theta)$: Antisymmetric part is multiplied by nonnegative symmetric function $\mathrm{e}^{J(\mathbf{r};\theta)}$. Can capture complex electron correlations but not change the nodal surface.
- **Backflow**[5],[6] $\mathbf{f}(\mathbf{r};\theta)$: Generalize one-electron orbitals $\varphi_\mu(\mathbf{r}_i)$ to many-electron orbitals $\varphi_{\mu i}(\mathbf{r})$. Can change the nodal surface.
- Exploit the flexibility of deep learning to achieve **higher accuracy**!

[4] N. D. Drummond et al. *Phys. Rev. B* 70:235119 (2004)
[5] P. López Ríos et al, *Phys. Rev. E*, 74:066701(2006)
[6] D. Luo, B. K. Clark. *Phys. Rev. Lett.* 122:226401 (2019)

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r}) + J(\mathbf{r};\theta)}$$

- **Cusp conditions**:

$$\frac{1}{\psi_0} \frac{\partial \psi_0}{\partial |\mathbf{r}_i - \mathbf{R}_I|}\bigg|_{\mathbf{r}_i = \mathbf{R}_I} = -Z_I \qquad \frac{1}{\psi_0} \frac{\partial \psi_0}{\partial |\mathbf{r}_i - \mathbf{r}_j|}\bigg|_{\mathbf{r}_i = \mathbf{r}_j} = \begin{cases} \frac{1}{4} & s_i = s_j \\ \frac{1}{2} & s_i \neq s_j \end{cases}$$

- **PauliNet**:
  - Nuclear cusps are built into $\varphi_\mu(\mathbf{r}_i)$[7].
  - Electronic cusps: use asymptotic function

$$\gamma(\mathbf{r}) := \sum_{i<j} -\frac{c_{ij}}{1 + |\mathbf{r}_i - \mathbf{r}_j|} \qquad c_{ij} = \begin{cases} \frac{1}{4} & s_i = s_j \\ \frac{1}{2} & s_i \neq s_j \end{cases}$$

  - Backflow and Jastrow networks are cuspless by construction:

$$\nabla_{\mathbf{r}_i} J(\mathbf{r})\big|_{\mathbf{r}_i = \{\mathbf{r}_k, \mathbf{R}_a\}} = 0,$$

$$\nabla_{\mathbf{r}_i} f_{\mu i}(\mathbf{r})\big|_{\mathbf{r}_i = \{\mathbf{r}_k, \mathbf{R}_a\}} = 0$$

[7] A. Ma et al: J. Chem. Phys. 122:224322 (2005)

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)]\det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- **Cusp conditions**:

$$\frac{1}{\psi_0}\frac{\partial\psi_0}{\partial|\mathbf{r}_i-\mathbf{R}_I|}\bigg|_{\mathbf{r}_i=\mathbf{R}_I} = -Z_I \qquad \frac{1}{\psi_0}\frac{\partial\psi_0}{\partial|\mathbf{r}_i-\mathbf{r}_j|}\bigg|_{\mathbf{r}_i=\mathbf{r}_j} = \begin{cases}\frac{1}{4} & s_i = s_j\\ \frac{1}{2} & s_i \neq s_j\end{cases}$$

- **PauliNet**:
  - Nuclear cusps are built into $\varphi_\mu(\mathbf{r}_i)$[7].
  - Electronic cusps: use asymptotic function

$$\gamma(\mathbf{r}) := \sum_{i<j} -\frac{c_{ij}}{1+|\mathbf{r}_i-\mathbf{r}_j|} \qquad c_{ij} = \begin{cases}\frac{1}{4} & s_i = s_j\\ \frac{1}{2} & s_i \neq s_j\end{cases}$$

  - Backflow and Jastrow networks are cuspless by construction:

$$\nabla_{\mathbf{r}_i} J(\mathbf{r})\big|_{\mathbf{r}_i=\{\mathbf{r}_k,\mathbf{R}_a\}} = 0,$$
$$\nabla_{\mathbf{r}_i} f_{\mu i}(\mathbf{r})\big|_{\mathbf{r}_i=\{\mathbf{r}_k,\mathbf{R}_a\}} = 0$$

[7] A. Ma et al: J. Chem. Phys. 122:224322 (2005)

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r})+J(\mathbf{r};\theta)}$$

- **Cusp conditions**:

$$\frac{1}{\psi_0}\frac{\partial\psi_0}{\partial|\mathbf{r}_i - \mathbf{R}_I|}\bigg|_{\mathbf{r}_i=\mathbf{R}_I} = -Z_I \qquad \frac{1}{\psi_0}\frac{\partial\psi_0}{\partial|\mathbf{r}_i - \mathbf{r}_j|}\bigg|_{\mathbf{r}_i=\mathbf{r}_j} = \begin{cases} \frac{1}{4} & s_i = s_j \\ \frac{1}{2} & s_i \neq s_j \end{cases}$$

- **PauliNet**:
  - Nuclear cusps are built into $\varphi_\mu(\mathbf{r}_i)$[7].
  - Electronic cusps: use asymptotic function

$$\gamma(\mathbf{r}) := \sum_{i<j} -\frac{c_{ij}}{1+|\mathbf{r}_i - \mathbf{r}_j|} \qquad c_{ij} = \begin{cases} \frac{1}{4} & s_i = s_j \\ \frac{1}{2} & s_i \neq s_j \end{cases}$$

  - Backflow and Jastrow networks are cuspless by construction:

$$\nabla_{\mathbf{r}_i} J(\mathbf{r})\big|_{\mathbf{r}_i=\{\mathbf{r}_k,\mathbf{R}_a\}} = 0,$$

$$\nabla_{\mathbf{r}_i} f_{\mu i}(\mathbf{r})\big|_{\mathbf{r}_i=\{\mathbf{r}_k,\mathbf{R}_a\}} = 0$$

[7] A. Ma et al: J. Chem. Phys. 122:224322 (2005)

Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet Wavefunction:

$$\psi(\mathbf{r};\theta) = \det[\varphi_\mu(\mathbf{r}_i^\uparrow) f_{\mu i}(\mathbf{r};\theta)] \det[\varphi_\mu(\mathbf{r}_i^\downarrow) f_{\mu i}(\mathbf{r};\theta)] \mathrm{e}^{\gamma(\mathbf{r}) + J(\mathbf{r};\theta)}$$

- **Cusp conditions**:

$$\frac{1}{\psi_0} \frac{\partial \psi_0}{\partial |\mathbf{r}_i - \mathbf{R}_I|}\bigg|_{\mathbf{r}_i = \mathbf{R}_I} = -Z_I \qquad \frac{1}{\psi_0} \frac{\partial \psi_0}{\partial |\mathbf{r}_i - \mathbf{r}_j|}\bigg|_{\mathbf{r}_i = \mathbf{r}_j} = \begin{cases} \frac{1}{4} & s_i = s_j \\ \frac{1}{2} & s_i \neq s_j \end{cases}$$

- **PauliNet**:
  - Nuclear cusps are built into $\varphi_\mu(\mathbf{r}_i)$[7].
  - Electronic cusps: use asymptotic function

$$\gamma(\mathbf{r}) := \sum_{i<j} -\frac{c_{ij}}{1 + |\mathbf{r}_i - \mathbf{r}_j|} \qquad c_{ij} = \begin{cases} \frac{1}{4} & s_i = s_j \\ \frac{1}{2} & s_i \neq s_j \end{cases}$$

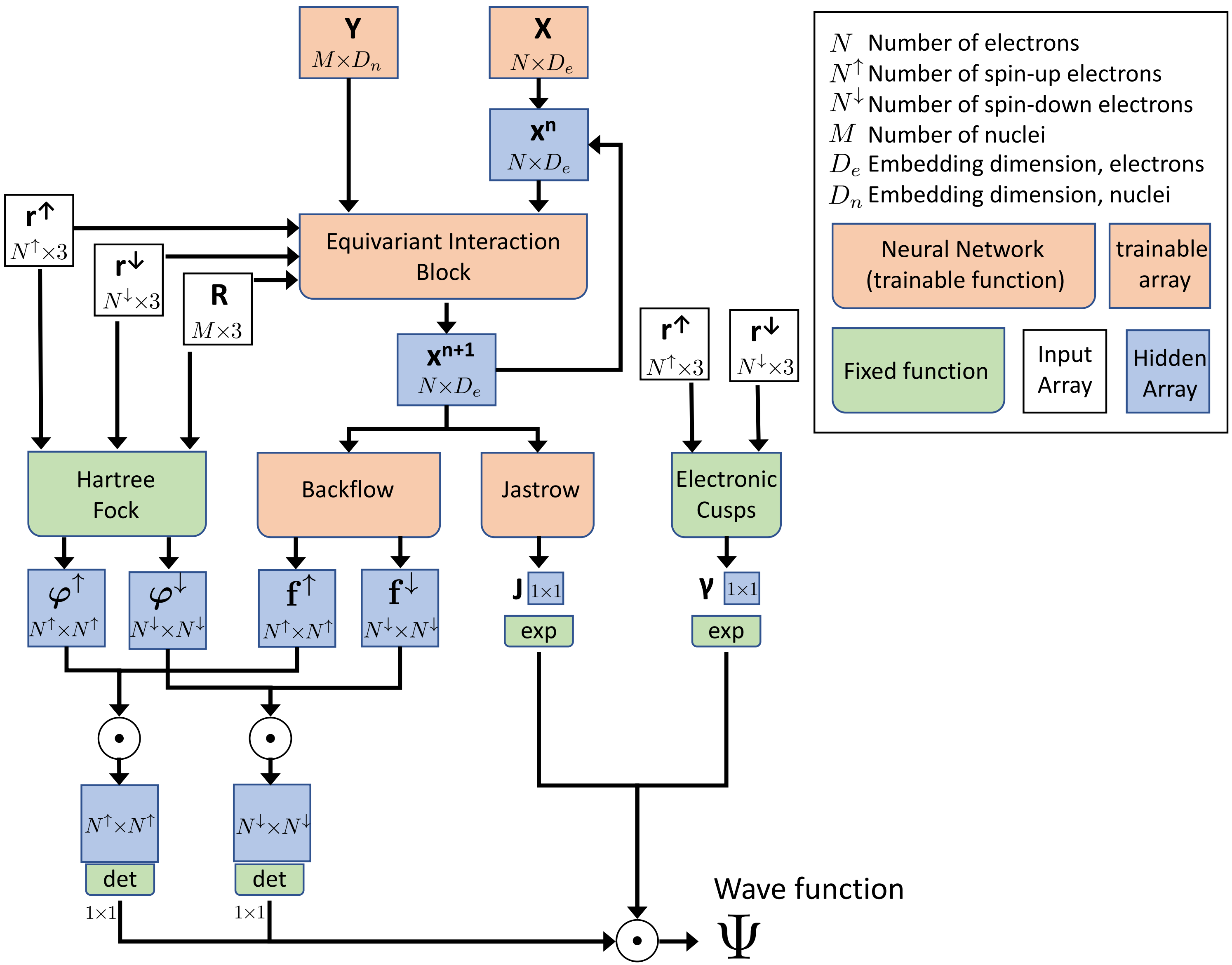
  - Backflow and Jastrow networks are cuspless by construction:

$$\nabla_{\mathbf{r}_i} J(\mathbf{r})\big|_{\mathbf{r}_i = \{\mathbf{r}_k, \mathbf{R}_a\}} = 0,$$

$$\nabla_{\mathbf{r}_i} f_{\mu i}(\mathbf{r})\big|_{\mathbf{r}_i = \{\mathbf{r}_k, \mathbf{R}_a\}} = 0$$

[7] A. Ma et al: J. Chem. Phys. 122:224322 (2005)

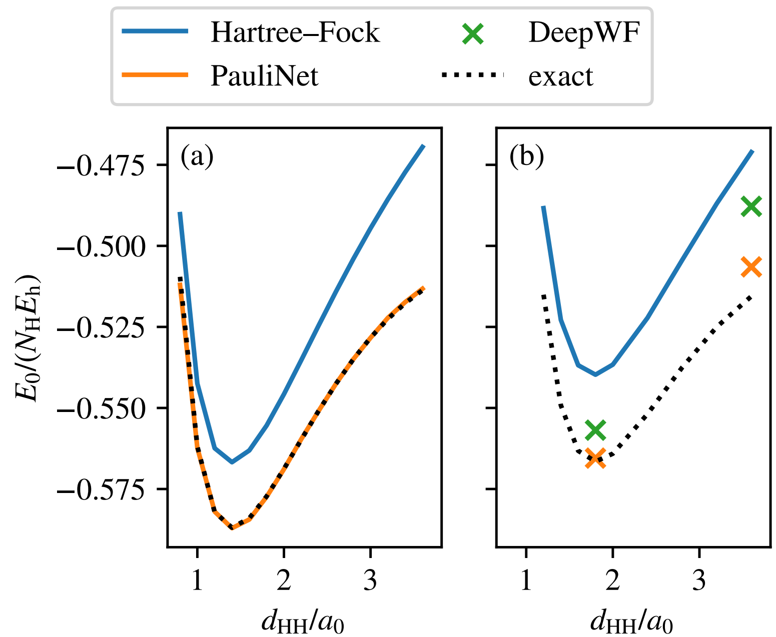
PauliNet



Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

PauliNet: Results

system[a]	VMC		DMC		DeepWF		PauliNet	
	E_0/E_h ($\%\ E_{corr}$)							
H_2[b]					−1.1738	98.4%	−1.174 37(6)	99.7%
LiH[c]	−8.0635[f]	91.5%	−8.0703[f]	99.7%	−7.8732	[g]	−8.0690(3)	98.1%
Be[d]	−14.6311[d]	61.6%	−14.6572[d]	89.2%	−14.6141	43.6%	−14.6546(7)	86.5%
B[d]	−24.6056[d]	60.0%	−24.6398[d]	88.3%	−24.2124	[g]	−24.634(2)	83.5%
H_{10}[e]					−5.5685	63.8%	−5.655(2)	96.0%



Hermann, Schätzle, Noé, **arXiv:1909.08423** (2019).
Deep neural network solution of the electronic Schrödinger equation.

Freie Universität Berlin

Acknowledgements

Collaborations

Cecilia Clementi (Rice University)
Christof Schütte (FU Berlin)
Eric Vanden-Eijnden (Courant Institut NY)
Thomas Weikl (MPI Potsdam)
Edina Rosta (King's College London)

Vijay Pande (Stanford)
Volker Haucke (FMP Berlin)
Stephan Sigrist (FU Berlin)
Katja, Faelber, Oliver Daumke (MDC)
John Chodera (MSKCC NY)
Gianni de Fabritiis (Barcelona)

Funding

erc

EINSTEIN Foundation.de

DFG Deutsche Forschungsgemeinschaft

Freie Universität Berlin