香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

ICLR

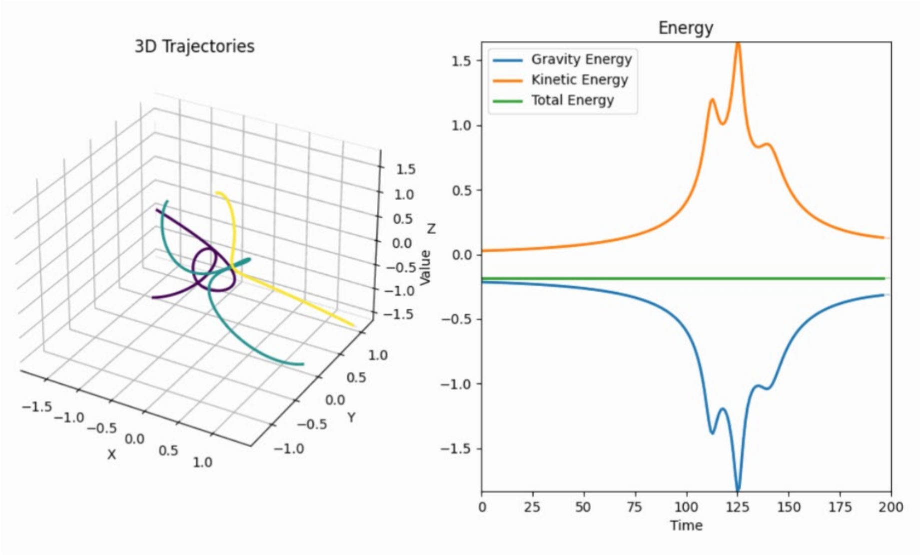
Generating Physical Dynamics under Priors

Zihan Zhou[1], *Xiaoxue Wang*[2], *Tianshu Yu*[1]
[1]School of Data Science, The Chinese University of Hong Kong, Shenzhen
[2]ChemLex Technology Co., Ltd.

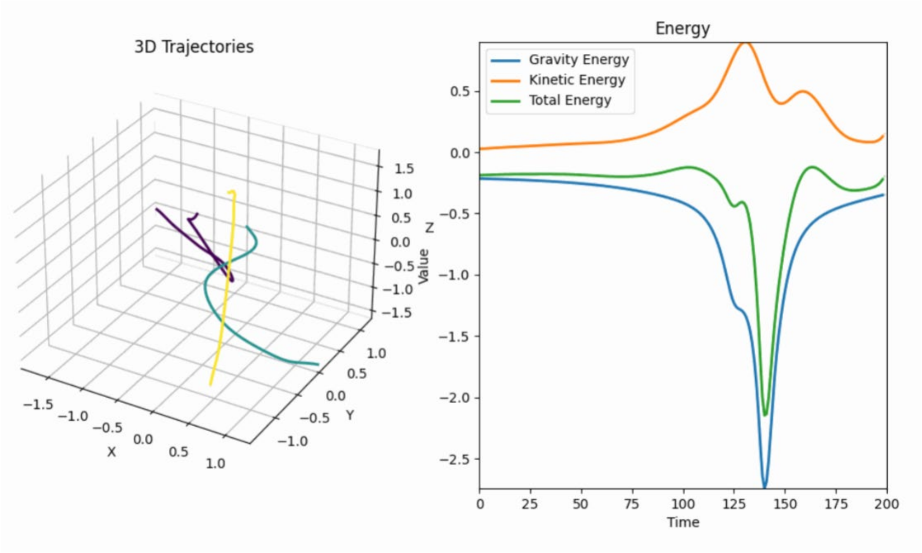
Why Physical Priors Matter?

- Generative models often violate physics laws (energy, momentum, PDEs)
- Consequences: unrealistic simulations, poor generalizability

sample from numerical solver



sample generated from diffusion models



Bridging Physics and Generative Models

Distributional Priors

- **Goal:**
  Respect fundamental symmetries
  (e.g., rotaion/translation invariance)
- **Solution**:
  Equivariant diffusion models
- **Impact**:
  Generated dynamics are invariant to irrelevant transformations
- **Example**:
  3D particle system rotating with fixed pairwise distances

Physical Constraints

- **Goal:**
  Enforce laws (PDEs, energy/momentum conservation)
- **Solution**:
  Constraint decomposition + penalty loss
- **Impact**:
  Avoids unphysical outputs
  (e.g., energy drift)
- **Example**:
  conservation law: $E_{\text{total}} = \text{const}$

Injecting Priors into diffusion

- The integration of **priors** into the generative process is a complex task that necessitates a deep understanding of the relevant mathematical and physical principles.
  - **Distributional Priors**
  - **Physical Feasibility Priors**
- predictive models: predict $\boldsymbol{x}_0$
- generative models: sample $\mathbb{E}[\boldsymbol{x}_0 \mid \boldsymbol{x}_t]$

Injecting priors ensures that generated dynamics are both realistic and physically plausible.

Method: Distributional Priors

- A distribution q is said to be $\mathcal{G}$-invariant under the group $\mathcal{G}$ if for all transformations $G \in \mathcal{G}$, we have $q(G(\boldsymbol{x})) = q(\boldsymbol{x})$

- Sufficient conditions: invariant $q_0 \implies$ invariant q_t

$\mathcal{G}$-invariant distribution
volume-preserving diffeomorphism
isometry, and homogeneity
} q_t is $\mathcal{G}$-invariant

- Property of score functions

$$q(G(\boldsymbol{x})) = q(\boldsymbol{x}) \implies \nabla_{G(\boldsymbol{x})} \log q(G(\boldsymbol{x})) = \left(\frac{\partial G(\boldsymbol{x})}{\partial \boldsymbol{x}}\right)^{-1} \nabla_{\boldsymbol{x}} \log q(\boldsymbol{x})$$

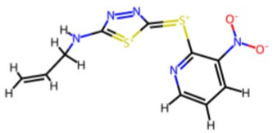
Score function is $(\mathcal{G}, \nabla^{-1})$-equivariant
Models for noise matching should also be $(\mathcal{G}, \nabla^{-1})$-equivariant

$$\mathcal{J}_{\text{noise}}(\boldsymbol{\theta}) = \mathbb{E}_{t, \boldsymbol{x}_0, \boldsymbol{\epsilon}}\left[w(t)\left\|\boldsymbol{\epsilon}_{\boldsymbol{\theta}}\left(\boldsymbol{x}_t, t\right) - \boldsymbol{\epsilon}\right\|^2\right]$$

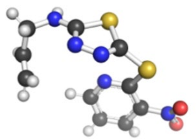
$$\boldsymbol{\epsilon}_{\boldsymbol{\theta}}^*\left(\boldsymbol{x}_t, t\right) = -\sigma_t \nabla_{\boldsymbol{x}} \log q_t\left(\boldsymbol{x}_t\right)$$

Example:

SE(3)- and permutation-invariant



molecular graph



conformations

protein generation

point cloud generation

Distribution	***Backbone***
SE(n)-invariant	SO(n)-equivariant Translational-invariant
Permutation invariant	Permutation invariant

Method: Physical Feasibility Priors

Prior types:

- Conservation laws:
  conservation of momentum, energy, flux, …
- PDE constraints:
  Darcy flow equations, Burger equations, …

Diffusion loss

$$\mathcal{J}_{\text{noise}}(\boldsymbol{\theta}) = \mathbb{E}_{t,\boldsymbol{x}_0,\boldsymbol{\epsilon}}\left[w(t)\left\|\boldsymbol{\epsilon}_{\boldsymbol{\theta}}\left(\boldsymbol{x}_t, t\right)-\boldsymbol{\epsilon}\right\|^2\right], \qquad \boldsymbol{\epsilon}_{\boldsymbol{\theta}}^*\left(\boldsymbol{x}_t, t\right) = -\sigma_t \nabla_{\boldsymbol{x}} \log q_t\left(\boldsymbol{x}_t\right);$$

$$\mathcal{J}_{\text{data}}(\boldsymbol{\theta}) = \mathbb{E}_{t,\boldsymbol{x}_0,\boldsymbol{\epsilon}}\left[w(t)\left\|\boldsymbol{x}_{\boldsymbol{\theta}}\left(\boldsymbol{x}_t, t\right)-\boldsymbol{x}_0\right\|^2\right], \quad \boldsymbol{x}_{\boldsymbol{\theta}}^*\left(\boldsymbol{x}_t, t\right) = \frac{1}{\alpha_t}\boldsymbol{x}_t + \frac{\sigma_t^2}{\alpha_t}\nabla_{\boldsymbol{x}} \log q_t\left(\boldsymbol{x}_t\right).$$

Tweedie's formula

$$\mathbb{E}\left[\boldsymbol{x}_0 \mid \boldsymbol{x}_t\right] = \frac{1}{\alpha_t}\left(\boldsymbol{x}_t - \sigma_t \boldsymbol{\epsilon}_{\boldsymbol{\theta}}^*\left(\boldsymbol{x}_t, t\right)\right)$$

$$\mathbb{E}\left[\boldsymbol{x}_0 \mid \boldsymbol{x}_t\right] = \boldsymbol{x}_{\boldsymbol{\theta}}^*\left(\boldsymbol{x}_t, t\right)$$

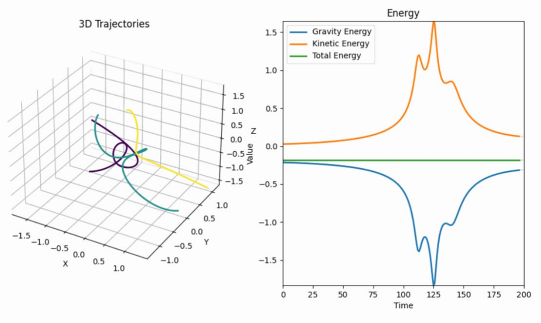
Jensen's gap (main difficulty)

$$\mathcal{R}\left(\mathbb{E}\left[\boldsymbol{x}_0 \mid \boldsymbol{x}_t\right]\right) \neq \mathbb{E}\left[\mathcal{R}\left(\boldsymbol{x}_0\right) \mid \boldsymbol{x}_t\right] = \mathbf{0}$$

Examples:

- conservation of energy:

$$-\sum_{i\neq j}^{K} \frac{Gm^2}{\mathbf{R}_{ij,l}^{(0)}} + \sum_{k=1}^{K}\sum_{d=1}^{D} \frac{1}{2} m \left(\mathbf{V}_{l,k,d}^{(0)}\right)^2 = \text{constant}$$

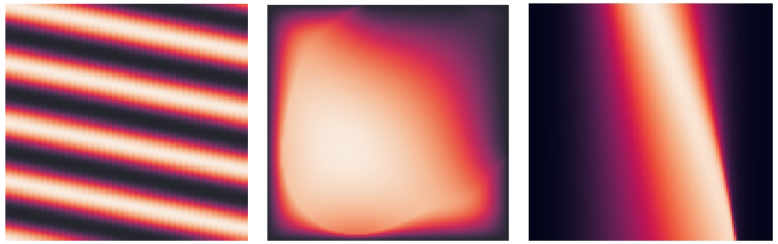


- PDE constraints:

Advection: $\partial_t u(t,x) + \beta \partial_x u(t,x) = 0,$

Darcy flow: $\partial_t u(x,t) - \nabla(a(x)\nabla u(x,t)) = f(x),$

Burger: $\partial_t u(x,t) + u(x,t)\partial_x u(x,t) = 0,$



Method: Physical Feasibility Priors

Apply the penalty loss on $\mathbb{E}\left[\boldsymbol{x}_0 \mid \boldsymbol{x}_t\right]$:

$$\mathcal{J}(\boldsymbol{\theta}) = \mathcal{J}_{\text{score}}(\boldsymbol{\theta}) + \lambda \mathcal{J}_{\mathcal{R}}(\boldsymbol{\theta})$$

Handling Jensen's gap:

$$\mathcal{R}\left(\mathbb{E}\left[\boldsymbol{x}_0 \mid \boldsymbol{x}_t\right]\right) \neq \mathbb{E}\left[\mathcal{R}\left(\boldsymbol{x}_0\right) \mid \boldsymbol{x}_t\right] = \mathbf{0}$$

Case analysis of constraints $\mathcal{R}\left(\boldsymbol{x}_0\right)$

- Linear / convex / reducible nonlinear:
  $$\mathcal{J}_{\mathcal{R}}(\boldsymbol{\theta}) = \mathbb{E}_{t, \boldsymbol{x}_0, \boldsymbol{\epsilon}}\left[w(t)\left\|\mathbf{W}_0 \mathbf{u}_{\boldsymbol{\theta}}\left(\boldsymbol{x}_t, t\right) + \mathbf{b}_0\right\|^2\right]$$
- Multilinear: $\mathcal{R}\left(\boldsymbol{x}_0\right) = \mathbf{W}_0 \mathbf{u}_0 + \mathbf{b}_0 = \mathbf{0}$
  $$\mathcal{J}_{\mathcal{R}}(\boldsymbol{\theta}) = \mathbb{E}_{t, \boldsymbol{x}_0, \boldsymbol{\epsilon}}\left[w(t)\left\|\mathcal{R}\left(\boldsymbol{x}_{\boldsymbol{\theta}}\left(\boldsymbol{x}_t, t\right)\right)\right\|^2\right]$$
- General nonlinear: $\mathcal{R}\left(\boldsymbol{x}_0\right) = \mathbf{g}\left(\mathbf{h}\left(\boldsymbol{x}_0\right)\right) = \mathbf{0}$
  $$\mathcal{J}_{\mathcal{R}}(\boldsymbol{\theta}) = \mathbb{E}_{t, \boldsymbol{x}_0, \boldsymbol{\epsilon}}\left[w(t)\left\|\tilde{\boldsymbol{x}}_{\boldsymbol{\theta}}\left(\boldsymbol{x}_t, t\right) - \mathbf{h}\left(\boldsymbol{x}_0\right)\right\|^2\right]$$

Examples:

conservation of energy:

$$-\sum_{i \neq j}^{K} \frac{G m^2}{\mathbf{R}_{ij,l}^{(0)}} + \sum_{k=1}^{K} \sum_{d=1}^{D} \frac{1}{2} m \left(\mathbf{V}_{l,k,d}^{(0)}\right)^2 = \text{constant}$$

multilinear reduction →

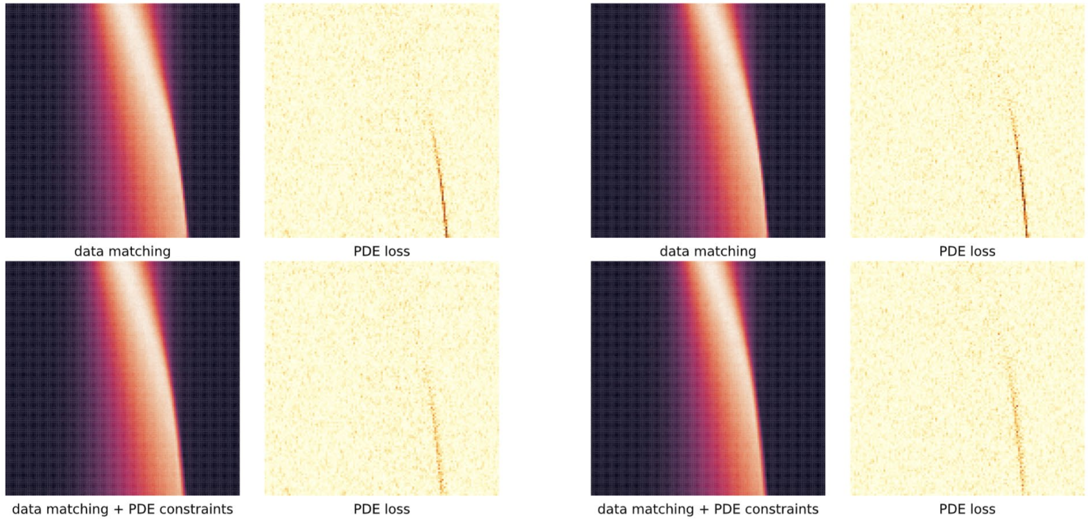
$$\begin{cases} \dfrac{1}{\mathbf{R}_{ij,l}^{(0)}} = \text{potential energy}_l \\ \left(\mathbf{V}_{l,k,d}^{(0)}\right)^2 = \text{kinetic energy}_l \end{cases}$$

convex →

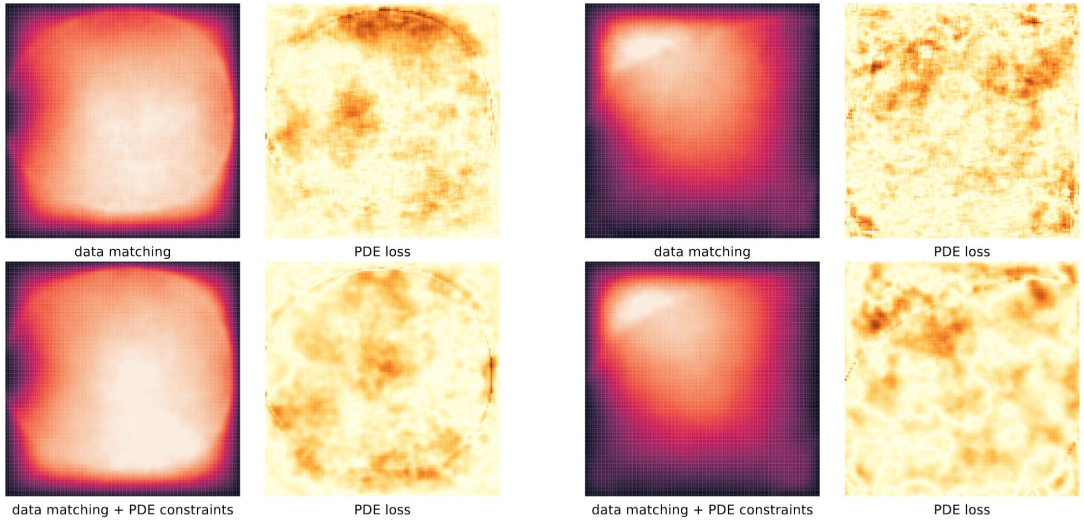
$$\begin{cases} \mathcal{J}_{\text{GPE}}(\boldsymbol{\theta}) = \mathbb{E}_{t, \boldsymbol{x}_0, \boldsymbol{\epsilon}}\left[w_1(t) \sum_{i \neq j, l}\left\|\frac{1}{\left(\mathbf{R}_{\boldsymbol{\theta}}\right)_{ij,l}} - \frac{1}{\mathbf{R}_{ij,l}^{(0)}}\right\|^2\right] \\ \mathcal{J}_{\text{KE}}(\boldsymbol{\theta}) = \mathbb{E}_{t, \boldsymbol{x}_0, \boldsymbol{\epsilon}}\left[w_2(t) \sum_{k,l,d}\left\|\left(\mathbf{V}_{\boldsymbol{\theta}}\right)_{l,k,d}^2 - \left(\mathbf{V}_{l,k,d}^{(0)}\right)^2\right\|^2\right] \end{cases}$$

Experiments: PDE Datasets

Burger: $\partial_t u(x,t) + u(x,t)\partial_x u(x,t) = 0$



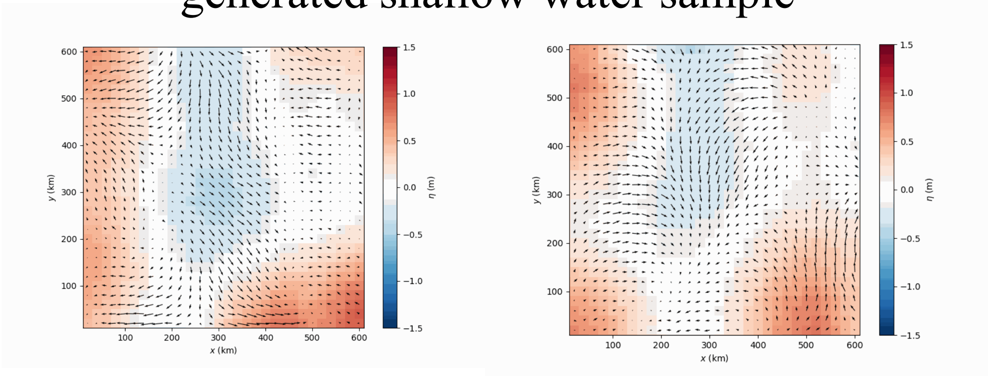
Darcy flow: $\partial_t u(x,t) - \nabla(a(x)\nabla u(x,t)) = f(x)$



Prediction, RMSE compared with ground-truth

Method	Advection ($\times 10^{-2}$)	Darcy flow ($\times 10^{-2}$)
w/o prior	1.7263±0.0491	2.0648±0.0600
w/ prior	**1.6536±0.0677**	**1.9678±0.0651**

generated shallow water sample



Generation, RMSE of PDE constraints

Method	Advection	Burger	Shallow water
w/o prior	2.398±0.024	6.862±0.060	8.0153±0.0960
w/ prior	**2.305±0.001**	**6.610±0.012**	**7.7618±0.0645**

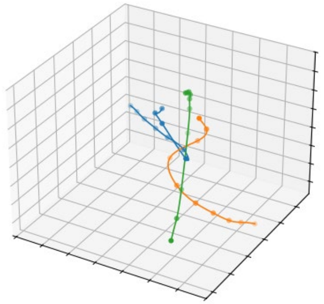
Experiments: Particle Dynamics Datasets

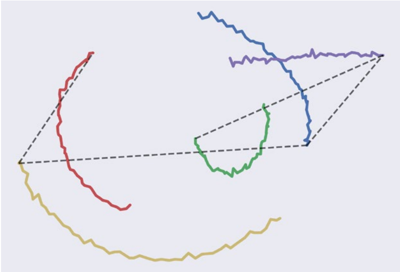
energy conservation

three-body: $-\sum_{i\neq j}^{K} \frac{Gm^2}{\mathbf{R}_{ij,l}^{(0)}} + \sum_{k=1}^{K} \sum_{d=1}^{D} \frac{1}{2} m \left(\mathbf{V}_{l,k,d}^{(0)}\right)^2 = \text{constant}$

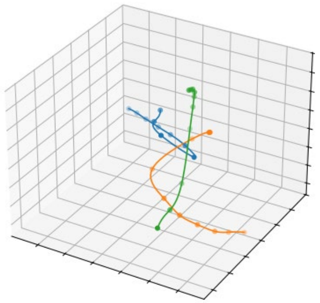
five-spring: $\sum_{(i,j)\in\mathcal{E}} \frac{1}{2}\kappa \left(\mathbf{R}_{ij,l}^{(0)}\right)^2 + \sum_{k=1}^{K} \sum_{d=1}^{D} \frac{1}{2} m \left(\mathbf{V}_{l,k,d}^{(0)}\right)^2 = \text{constant},$

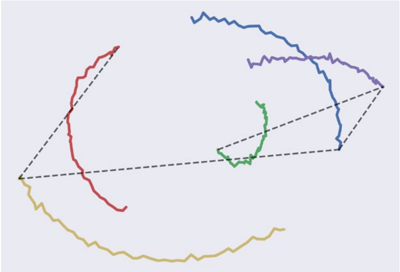
Baseline method



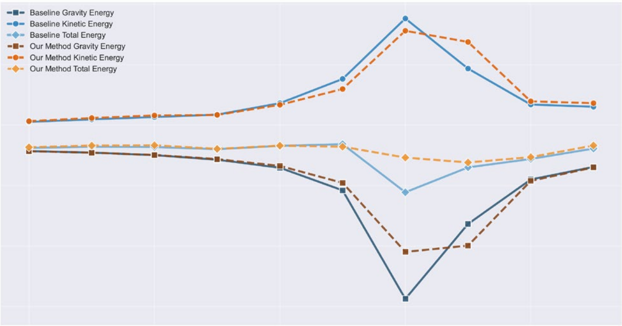


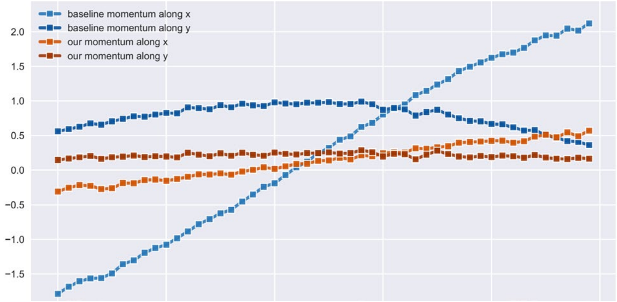
Our method





Conservation over time





Method	Three-body			Five-spring		
	Traj error	Vel error	Energy error	Dynamic error	Momentum error	Energy error
w/o prior	2.4132±0.1208	2.5745±0.0790	4.3292±0.7235	5.1754±0.0286	5.3699±0.0462	1.0618±0.0243
w/ prior	**1.9880±0.3418**	**0.8328±0.1042**	**0.5465±0.0705**	**5.0731±0.0406**	**0.3898±0.0118**	**0.7418±0.0129**

Generation under Physics Feasibility Priors

A framework to generate physics-compliant dynamics by integrating priors into diffusion models.

This advances AI4Physics by:

- Ensuring *feasibility* through principled constraint enforcement.
- Enabling *generalization* via equivariant architectures.

Thanks!