

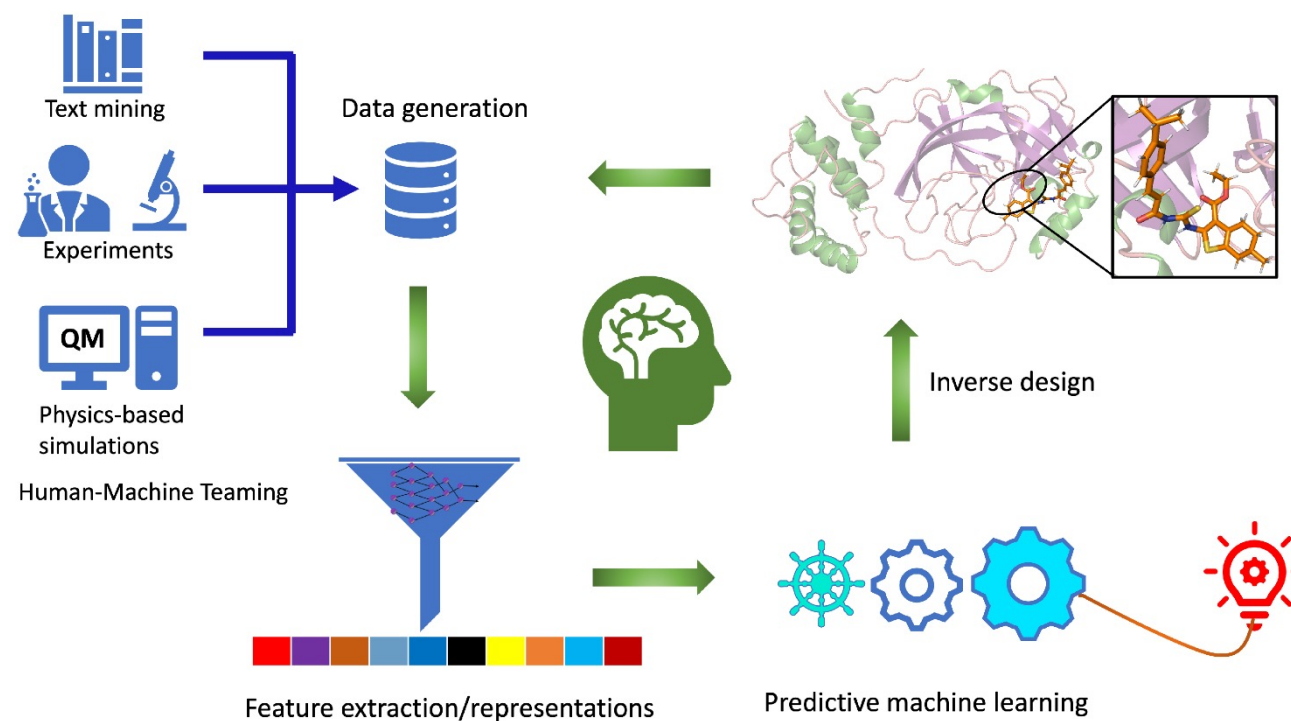
# Equivariant Energy-Guided SDE for Inverse Molecular Design

Fan Bao\*, Min Zhao\*, Zhongkai Hao, Peiyao Li, Chongxuan Li† and Jun Zhu†

# Inverse Molecular Design

Bao & Zhao et al., ICLR 2023

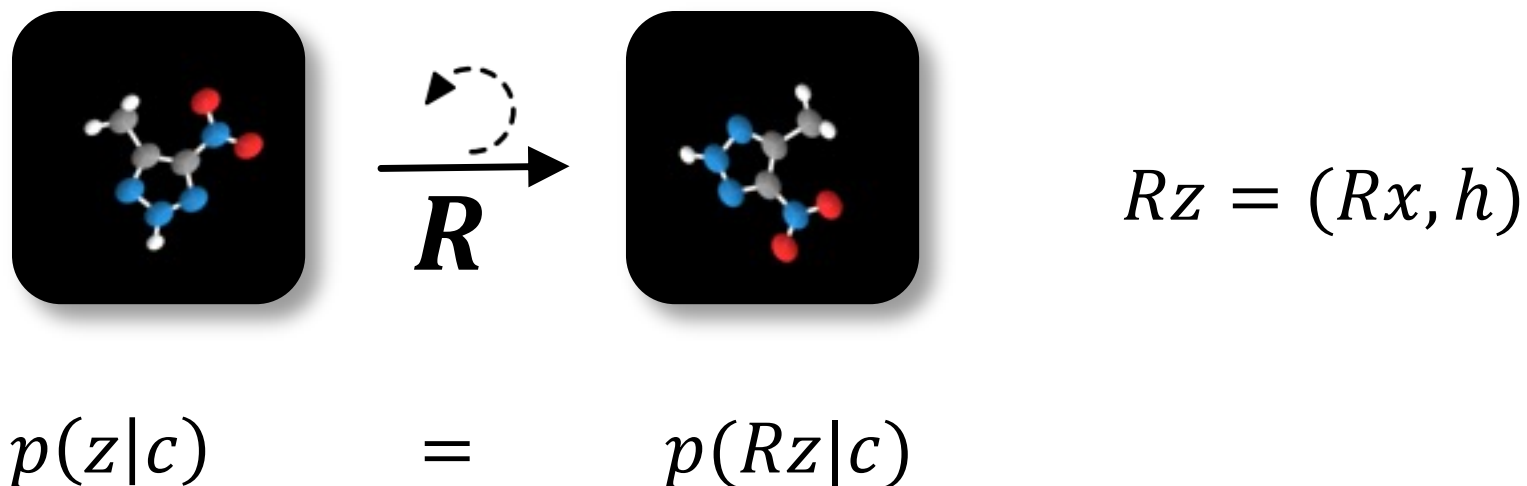
- Generated molecules should **satisfy certain desirable properties**, which involves learning a distribution of molecules conditioned on certain properties from data



# Background

*Bao & Zhao et al., ICLR 2023*

- Let a tuple  $z = (x, h)$  represent a molecule, where  $x = (x^1, \dots, x^M) \in \mathbb{R}^{Mn}$  is the collection of coordinates of the  $M$  atoms and  $h = (h^1, \dots, h^M) \in \mathbb{R}^{Md}$  is the corresponding atom feature.



- The model distribution  $p(z|c)$  should **be invariant to translational and rotational transformations** to leverage the geometric symmetry in 3D molecular conformation

# Equivariant SDE

Bao & Zhao et al., ICLR 2023

- Zero center of mass (CoM) subspace for **translational invariance**:  $X = \left\{x: \frac{1}{M} \sum_{i=1}^M x^i = 0\right\}$
- forward SDE in the product space  $X \times R^{Md}$

$$dz = f(t)zdt + g(t)d(\mathbf{w}_x, \mathbf{w}_h), \quad z_0 \sim q(z_0);$$



independent standard Wiener processes in  $X$  and  $R^{Md}$

# Equivariant SDE

Bao & Zhao et al., ICLR 2023

- reverse SDE in the product space

$$dz = [f(t)z - g(t)^2 \underbrace{(\nabla_{\mathbf{x}} \log q_t(\mathbf{z}) - \overline{\nabla_{\mathbf{x}} \log q_t(\mathbf{z})}, \nabla_{\mathbf{h}} \log q_t(\mathbf{z}))}_{\text{score function form}}]dt + g(t)d(\tilde{\mathbf{w}}_x, \tilde{\mathbf{w}}_h),$$

$$dz = [f(t)z + \frac{g(t)^2}{\sqrt{\beta_{t|0}}} \epsilon_{\theta}(\mathbf{z}, t)]dt + g(t)d(\tilde{\mathbf{w}}_x, \tilde{\mathbf{w}}_h), \quad \mathbf{z}_T \sim p_T(\mathbf{z}_T)$$

in the product space, subtract the CoM for translational invariance

- Rotation invariance


**Theorem 1.** Let  $(\epsilon_{\theta}^x(\mathbf{z}_t, t), \epsilon_{\theta}^h(\mathbf{z}_t, t)) = \epsilon_{\theta}(\mathbf{z}_t, t)$ , where  $\epsilon_{\theta}^x(\mathbf{z}_t, t)$  and  $\epsilon_{\theta}^h(\mathbf{z}_t, t)$  are the predicted noise of  $\mathbf{x}_t$  and  $\mathbf{h}_t$  respectively. If for any orthogonal transformation  $\mathbf{R} \in \mathbb{R}^{n \times n}$ ,  $\epsilon_{\theta}(\mathbf{z}_t, t)$  is equivariant to  $\mathbf{R}$ , i.e.,  $\epsilon_{\theta}(\mathbf{R}\mathbf{x}_t, \mathbf{h}_t, t) = (\mathbf{R}\epsilon_{\theta}^x(\mathbf{x}_t, \mathbf{h}_t, t), \epsilon_{\theta}^h(\mathbf{x}_t, \mathbf{h}_t, t))$ , and  $p_T(\mathbf{z}_T)$  is invariant to  $\mathbf{R}$ , i.e.,  $p_T(\mathbf{R}\mathbf{x}_T, \mathbf{h}_T) = p_T(\mathbf{x}_T, \mathbf{h}_T)$ , then  $p_{\theta}(\mathbf{z}_0)$  is invariant to any rotational transformation.

# Equivariant Energy-guided SDE (EEGSDE)

Bao & Zhao et al., ICLR 2023

- Guide generated molecules towards desired properties  $c$  via a time-dependent energy function  $E(\mathbf{z}, c, t)$ :

$$dz = [f(t)\mathbf{z} + g(t)^2 \left( \frac{1}{\sqrt{\beta_{t|0}}} \epsilon_{\theta}(\mathbf{z}, t) \right. \\ \left. + \underbrace{(\nabla_{\mathbf{x}} E(\mathbf{z}, c, t) - \overline{\nabla_{\mathbf{x}} E(\mathbf{z}, c, t)}, \nabla_{\mathbf{h}} E(\mathbf{z}, c, t))}_{\text{energy gradient taken in the product space}} \right)] dt + g(t) d(\tilde{\mathbf{w}}_x, \tilde{\mathbf{w}}_h), \mathbf{z}_T \sim p_T(\mathbf{z}_T)$$

translational invariance 

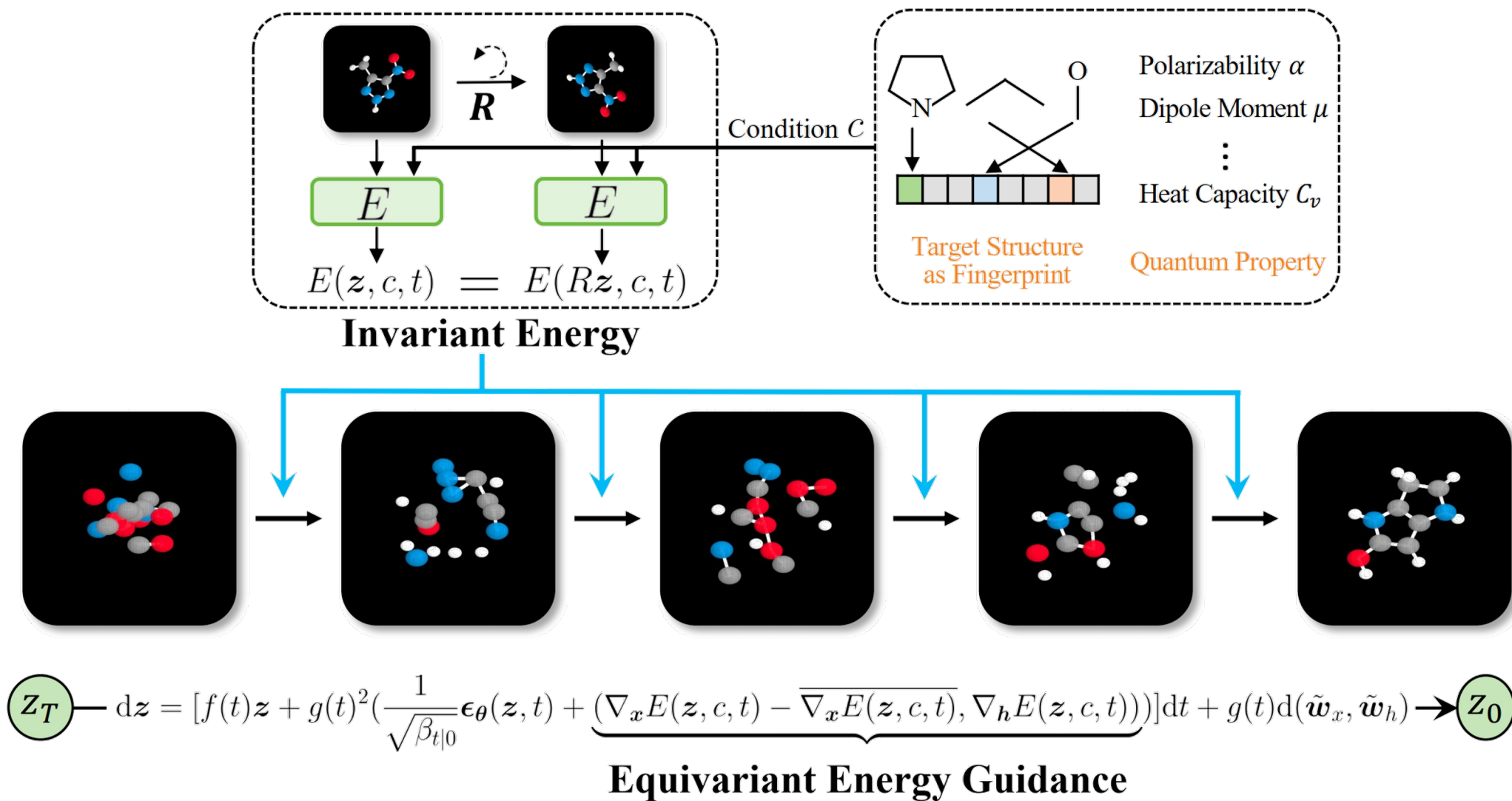
- **Rotation invariance:**

**Theorem 2.** Suppose the assumptions in Theorem 1 hold and  $E(\mathbf{z}, c, t)$  is invariant to any orthogonal transformation  $\mathbf{R}$ , i.e.,  $E(\mathbf{R}\mathbf{x}, \mathbf{h}, c, t) = E(\mathbf{x}, \mathbf{h}, c, t)$ . Then  $p_{\theta}(\mathbf{z}_0|c)$  is invariant to any rotational transformation.

- The EEGSDE defines a distribution  $p(\mathbf{z}_0|c)$  conditioned on the property  $c$ , **which is invariant to translational and rotational transformations**

# Equivariant Energy-guided SDE

Bao & Zhao et al., Arxiv 2022



# Choice of Energy : generating molecules with **desired quantum property**

Bao & Zhao et al., Arxiv 2022

- The energy function is defined as the squared error between the predicted property and the desired property:

$$E(\mathbf{z}_t, c, t) = s |g(\mathbf{z}_t, t) - c|^2$$

↑  
the desired property

↓  
time-dependent property prediction model

- Implementation of property prediction model :

$$g(\mathbf{z}_t, t) = \text{Dec}(\text{EGNN}^h(\mathbf{x}_t, \mathbf{h}'_t)), \quad \mathbf{h}'_t = \text{concatenate}(\mathbf{h}_t, t)$$

↓  
the second component (h feature part) in the output of EGNN  
invariant to orthogonal transformations



# Choice of Energy : generating molecules with **desired quantum property**

Bao & Zhao et al., Arxiv 2022

- EEGSDE is able to generate more accurate molecules than EDM, e.g., **reducing the mean absolute error** by more than 30% on the dipole moment property.

## Evaluation by **test property prediction model**

Method	MAE↓	Method	MAE↓
$C_v$ ( $\frac{\text{cal}}{\text{mol}}\text{K}$ )		$\mu$ (D)	
U-bound	6.879±0.015	U-bound	1.613±0.003
#Atoms	1.971	#Atoms	1.053
Conditional EDM	1.065±0.010	Conditional EDM	1.123±0.013
EEGSDE ( $s=1$ )	1.037±0.010	EEGSDE ( $s=0.5$ )	0.930±0.005
EEGSDE ( $s=5$ )	0.981±0.002	EEGSDE ( $s=1$ )	0.858±0.006
EEGSDE ( $s=10$ )	<b>0.941±0.005</b>	EEGSDE ( $s=2$ )	<b>0.777±0.007</b>
L-bound	0.040	L-bound	0.043
$\Delta\varepsilon$ (meV)		$\varepsilon_{\text{HOMO}}$ (meV)	
U-bound	1464±4	U-bound	645±41
#Atoms	866	#Atoms	426
Conditional EDM	671±5	Conditional EDM	371±2
EEGSDE ( $s=0.5$ )	574±4	EEGSDE ( $s=0.1$ )	357±4
EEGSDE ( $s=1$ )	542	EEGSDE ( $s=0.5$ )	320±1
EEGSDE ( $s=3$ )	<b>487±3</b>	EEGSDE ( $s=1$ )	<b>302±2</b>
L-bound	65	L-bound	39

## Evaluation by **Gaussian software**

Method	MAE↓	Method	MAE↓
$\mu$ (D)		$\alpha$ (Bohr <sup>3</sup> )	
Conditional EDM	1.20	Conditional EDM	2.41
EEGSDE ( $s=0.5$ )	0.96	EEGSDE ( $s=0.5$ )	2.27
EEGSDE ( $s=1$ )	0.78	EEGSDE ( $s=1$ )	2.03
EEGSDE ( $s=2$ )	<b>0.73</b>	EEGSDE ( $s=3$ )	<b>1.85</b>
Method	MAE↓	Method	MAE↓
$\varepsilon_{\text{HOMO}}$ (meV)		$\varepsilon_{\text{LUMO}}$ (meV)	
Conditional EDM	354	Conditional EDM	573
EEGSDE ( $s=0.1$ )	349	EEGSDE ( $s=0.5$ )	495
EEGSDE ( $s=0.5$ )	341	EEGSDE ( $s=1$ )	445
EEGSDE ( $s=1$ )	<b>284</b>	EEGSDE ( $s=3$ )	<b>416</b>

# Choice of Energy : generating molecules with **multiple target properties**

Bao & Zhao et al., Arxiv 2022

- EEGSDE is able to generate molecules targeted to multiple properties by combining the corresponding energy functions linearly:

$$E(\mathbf{z}_t, \mathbf{c}, t) = \sum_{k=1}^K E_k(\mathbf{z}_t, c_k, t)$$

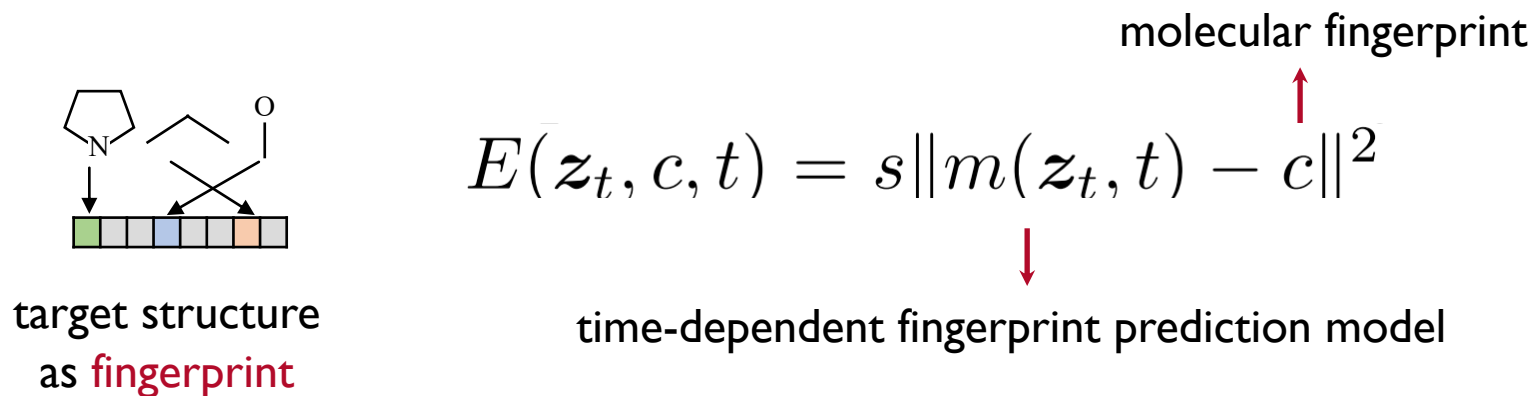
- EEGSDE still has a significantly better MAE than the conditional EDM

Method	MAE1↓	MAE2↓
$C_v$ ( $\frac{\text{cal}}{\text{mol}}\text{K}$ ), $\mu$ (D)		
Conditional EDM	1.079±0.007	1.156±0.011
EEGSDE ( $s_1=10, s_2=1$ )	<b>0.981±0.008</b>	<b>0.912±0.006</b>
$\Delta\varepsilon$ (meV), $\mu$ (D)		
Conditional EDM	683±1	1.130±0.007
EEGSDE ( $s_1=s_2=1$ )	<b>563±3</b>	<b>0.866±0.003</b>
$\alpha$ (Bohr <sup>3</sup> ), $\mu$ (D)		
Conditional EDM	2.76±0.01	1.158±0.002
EEGSDE ( $s_1=s_2=1.5$ )	<b>2.61±0.01</b>	<b>0.855±0.007</b>

# Choice of Energy : generating molecules with **target structure**

Bao & Zhao et al., Arxiv 2022

- The energy function is defined as the squared error between the predicted fingerprint and the molecule fingerprint :



- Implementation of fingerprint prediction model :

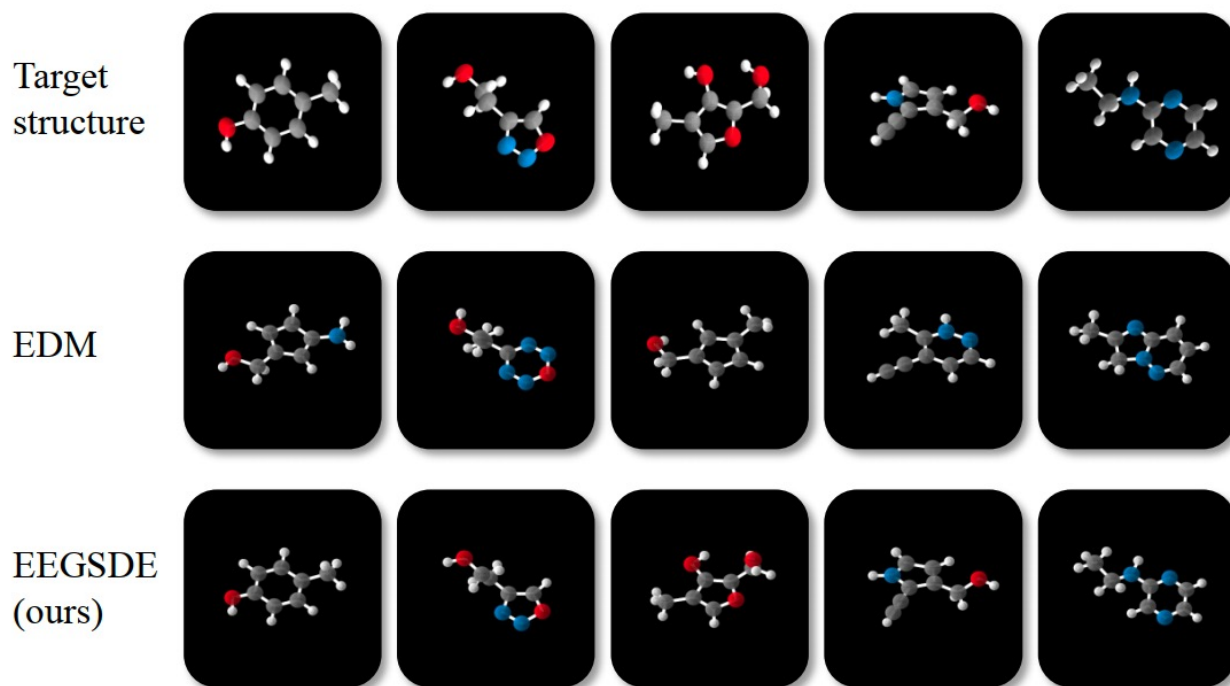
multi-label classifier

$$m(\mathbf{z}_t, t) = \sigma(\text{Dec}(\text{EGNN}^h(\mathbf{x}_t, \mathbf{h}'_t))), \quad \mathbf{h}'_t = \text{concatenate}(\mathbf{h}_t, t)$$

# Choice of Energy : generating molecules with **target structure**

Bao & Zhao et al., Arxiv 2022

- When targeted to specific molecular structures, EEGSDE better capture the structure information in molecules than EDM, e.g, improving the similarity to target structures by more than 10%.



Method	Similarity $\uparrow$
QM9	
cG-SchNet	$0.499 \pm 0.002$
Conditional EDM	$0.671 \pm 0.004$
EEGSDE ( $s=0.1$ )	$0.696 \pm 0.002$
EEGSDE ( $s=0.5$ )	$0.736 \pm 0.002$
EEGSDE ( $s=1.0$ )	<b><math>0.750 \pm 0.003</math></b>
GEOM-Drug	
Conditional EDM	$0.165 \pm 0.001$
EEGSDE ( $s=0.5$ )	$0.185 \pm 0.001$
EEGSDE ( $s=1.0$ )	<b><math>0.193 \pm 0.001</math></b>

Thank you!