Towards Physically Reliable Molecular Representation Learning

uai2023

Seunghoon Yi, Youngwoo Cho, Jinhwan Sul, Seung Woo Ko,
Soo Kyung Kim, Jaegul Choo, Hongkee Yoon, Joonseok Lee

SEOUL
NATIONAL
UNIVERSITY

Visual Information Processing Lab

Google

Stanford Research Institute

Molecular Energy Prediction Task

Molecular Energy Prediction Task

Estimating the total energy of a given molecule!

Molecular Energy Prediction Task

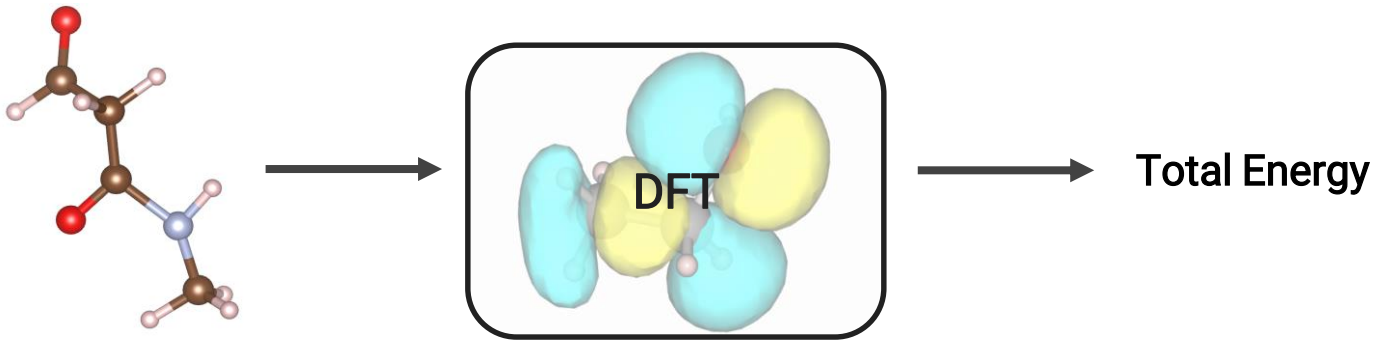
Estimating the total energy of a given molecule!

- To estimate physical / chemical properties of a molecule
- To predict chemical reactions on complex systems

Molecular Energy Prediction Task

Estimating the total energy of a given molecule!

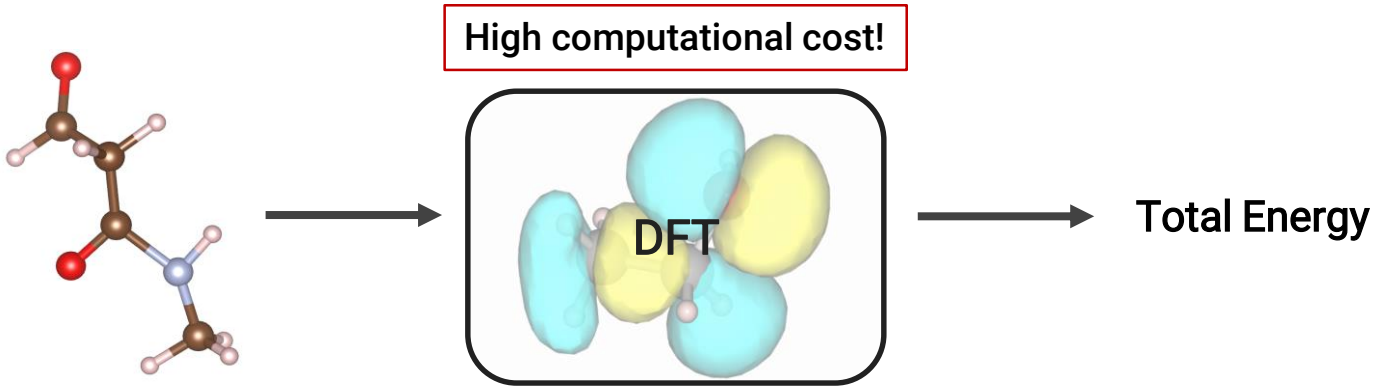
- To estimate physical / chemical properties of a molecule
- To predict chemical reactions on complex systems



Molecular Energy Prediction Task

Estimating the total energy of a given molecule!

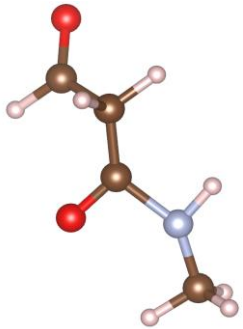
- To estimate physical / chemical properties of a molecule
- To predict chemical reactions on complex systems



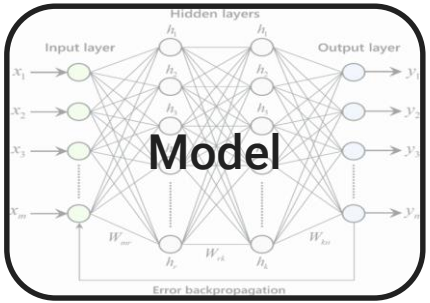
Molecular Energy Prediction Task

Estimating the total energy of a given molecule!

- To estimate physical / chemical properties of a molecule
- To predict chemical reactions on complex systems



High computational cost!



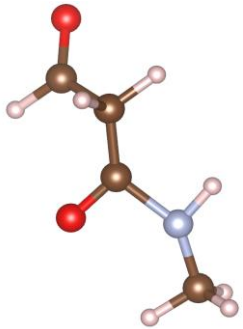
Total Energy

SchNet (Schütt et al., 2018)
DimeNet (Gasteiger et al., 2020)
MXMNet (Zhang et al., 2020)
ForceNet (Hu et al., 2021)
TorchMDNet (Thölke and De Fabritiis, 2022)

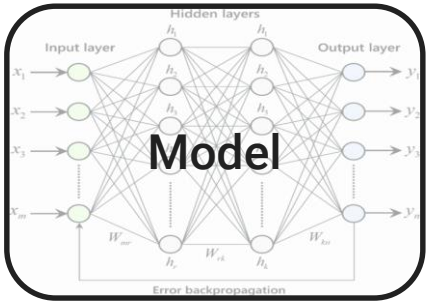
Molecular Energy Prediction Task

Estimating the total energy of a novel molecule!

- To estimate physical / chemical properties of a molecule
- To predict chemical reactions on complex systems



High computational cost!



Total Energy

SchNet (Schütt et al., 2018)
DimeNet (Gasteiger et al., 2020)
MXMNet (Zhang et al., 2020)
ForceNet (Hu et al., 2021)
TorchMDNet (Thölke and De Fabritiis, 2022)

Motivation

Indeed, existing models precisely estimates the energy!

Motivation

Indeed, existing models precisely estimates the energy!

Model	ForceNet	SchNet	DimeNet	TorchMDNet	MXMNet
Energy MAE (meV)	18.6	14.0	7.3	6.2	5.9

- Thermal fluctuation in room temperature(k_bT) ~ 25 meV

Motivation

Indeed, existing models precisely estimates the energy!

Model	ForceNet	SchNet	DimeNet	TorchMDNet	MXMNet
Energy MAE (meV)	18.6	14.0	7.3	6.2	5.9

- Thermal fluctuation in room temperature(k_bT) ~ 25 meV

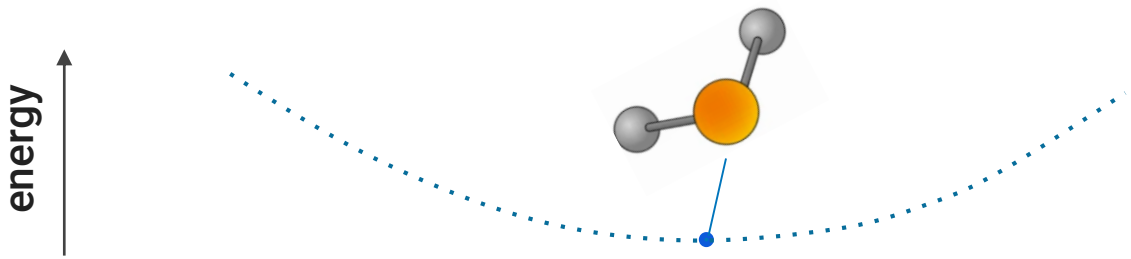
Is small energy error enough for a physically reliable model?

Motivation

- Let’s assess the reliability by structure optimization:

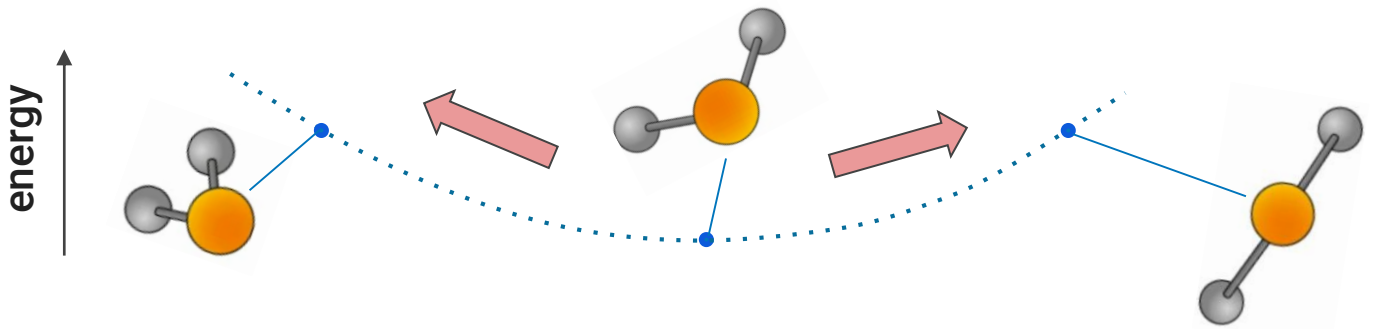
Motivation

- Let’s assess the reliability by structure optimization:
  - A stable molecule has lower energy than unstable ones.



Motivation

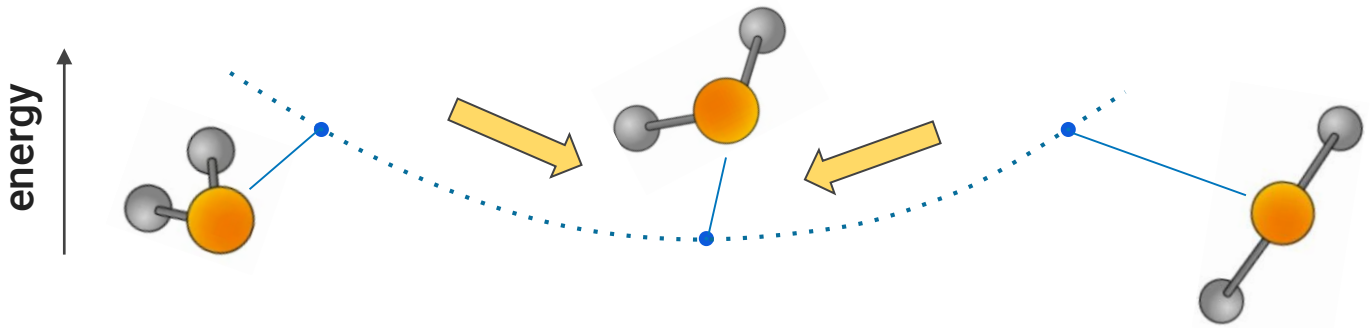
- Let’s assess the reliability by **structure optimization**:
  - A stable molecule has lower energy than unstable ones.
  - If we slightly perturb atoms in a molecule, the energy will get higher.



stable state energy < unstable state energy

Motivation

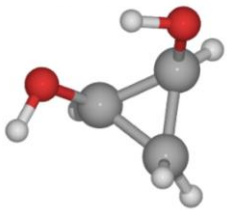
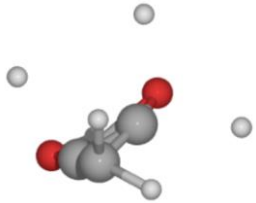
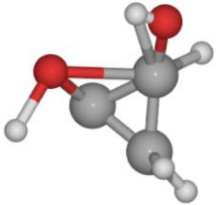
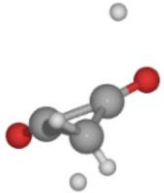
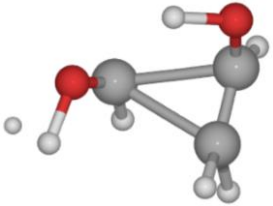
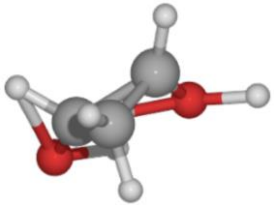
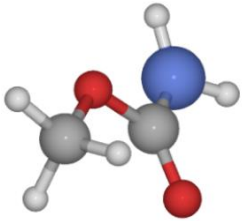
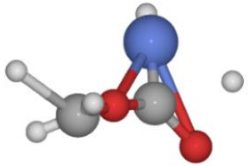
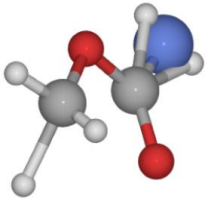
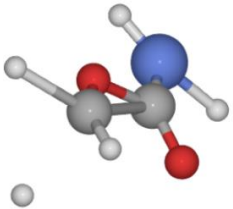
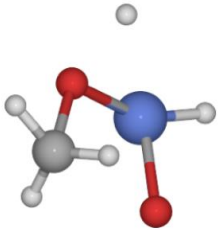
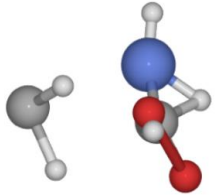
- Let's assess the reliability by structure optimization:
  - A stable molecule has lower energy than unstable ones.
  - If we slightly perturb atoms in a molecule, the energy will get higher.
  - Once we optimize it again, the structure should be recovered if the perturbation is small enough.



stable state energy < unstable state energy

Motivation

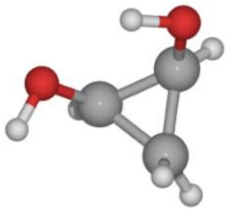
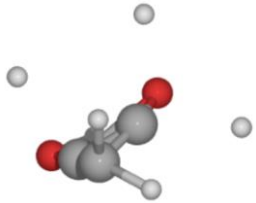
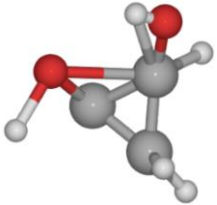
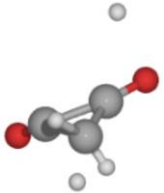
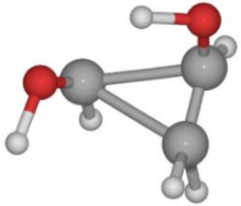
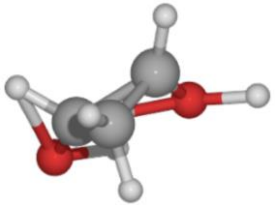
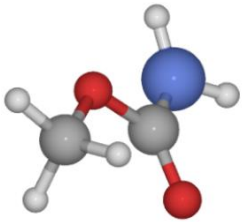
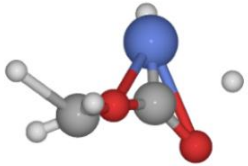
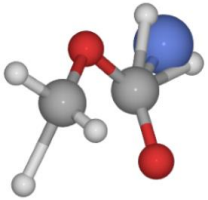
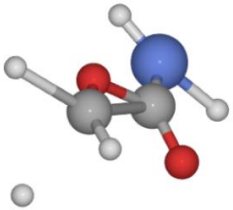
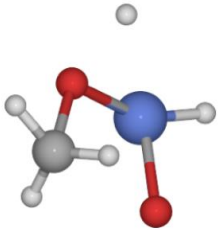
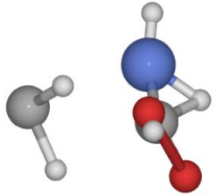
- Surprisingly, previous models fail to recover the stable structure!

Ground Truth	SchNet	ForceNet	DimeNet	MXMNet	TorchMDNet
					
					

Conducted on QM9 dataset.

Motivation

- Surprisingly, previous models fail to recover the stable structure!
  - Indicating they have been over-optimized only to energy estimation.

Ground Truth	SchNet	ForceNet	DimeNet	MXMNet	TorchMDNet
					
					

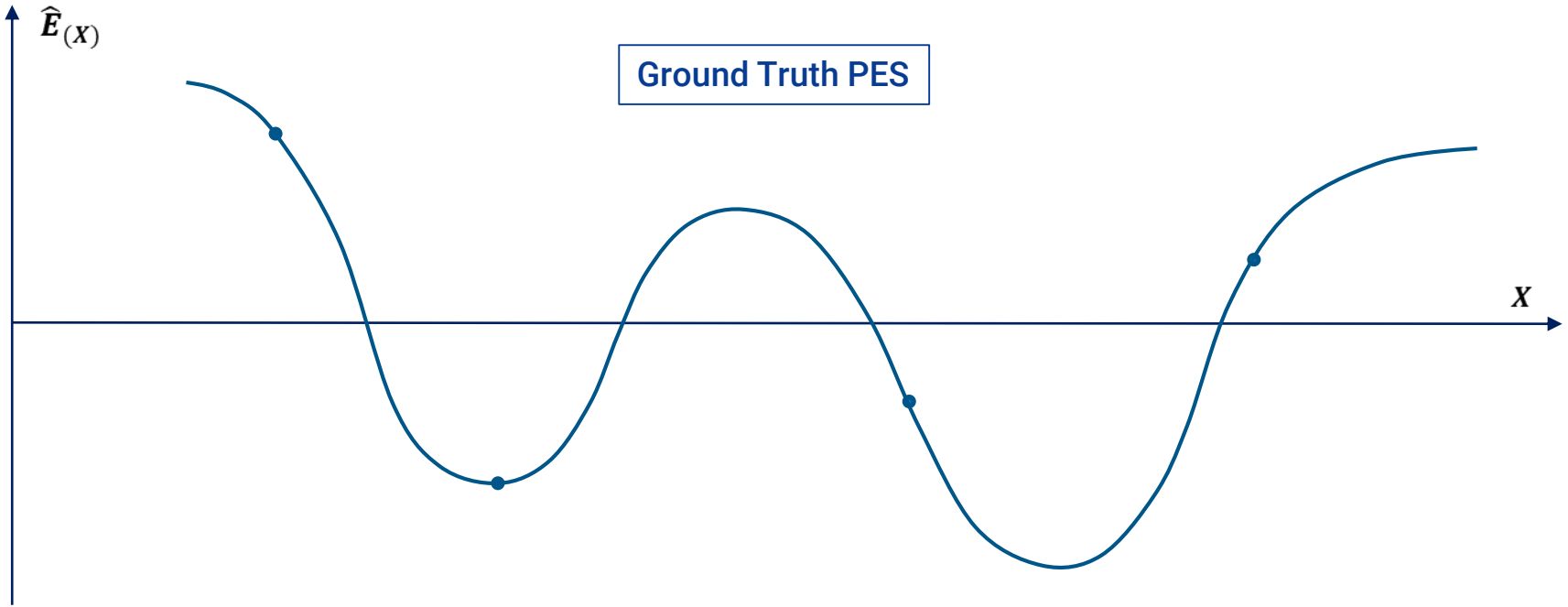
Conducted on QM9 dataset.

Core Issue to the Problem

- **Limited amount of data!**
  - Due to data sparsity, models have trouble learning a reliable potential energy surface (PES).

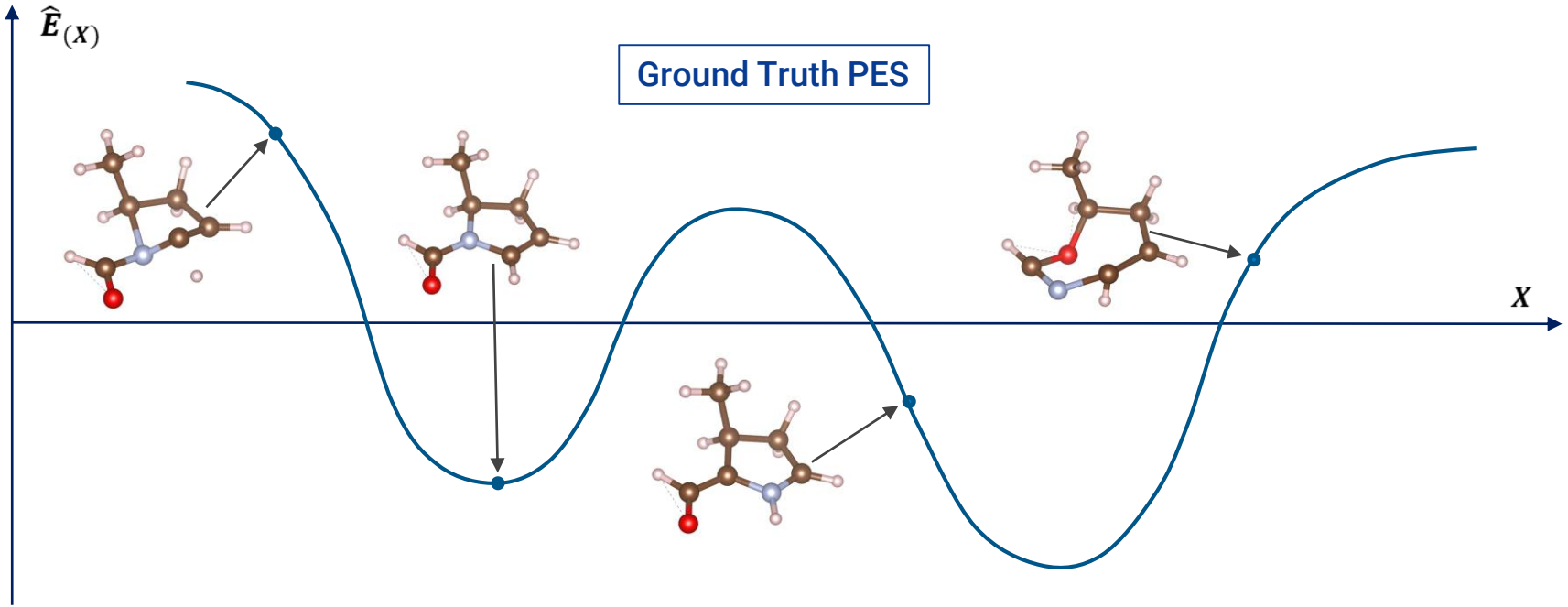
Core Issue to the Problem

- **Limited amount of data!**
  - Due to data sparsity, models have trouble learning a reliable potential energy surface (PES).



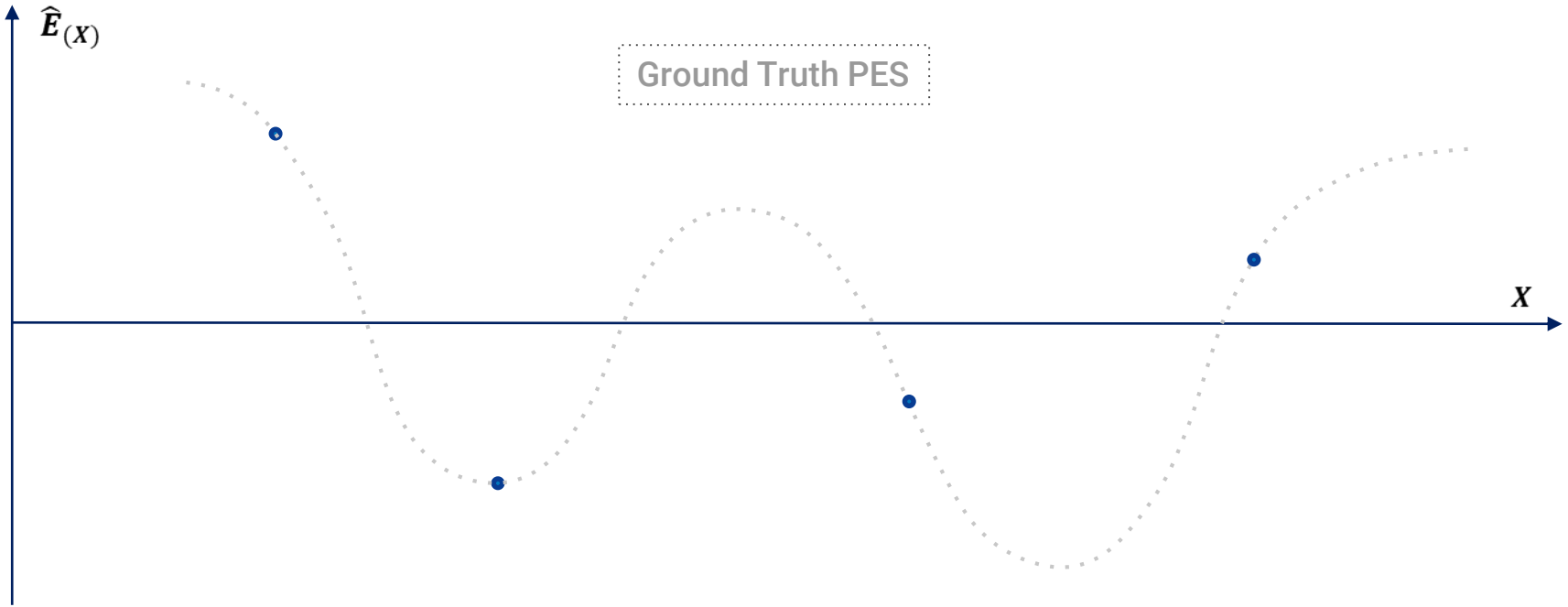
Core Issue to the Problem

- **Limited amount of data!**
  - Due to data sparsity, models have trouble learning a reliable potential energy surface (PES).



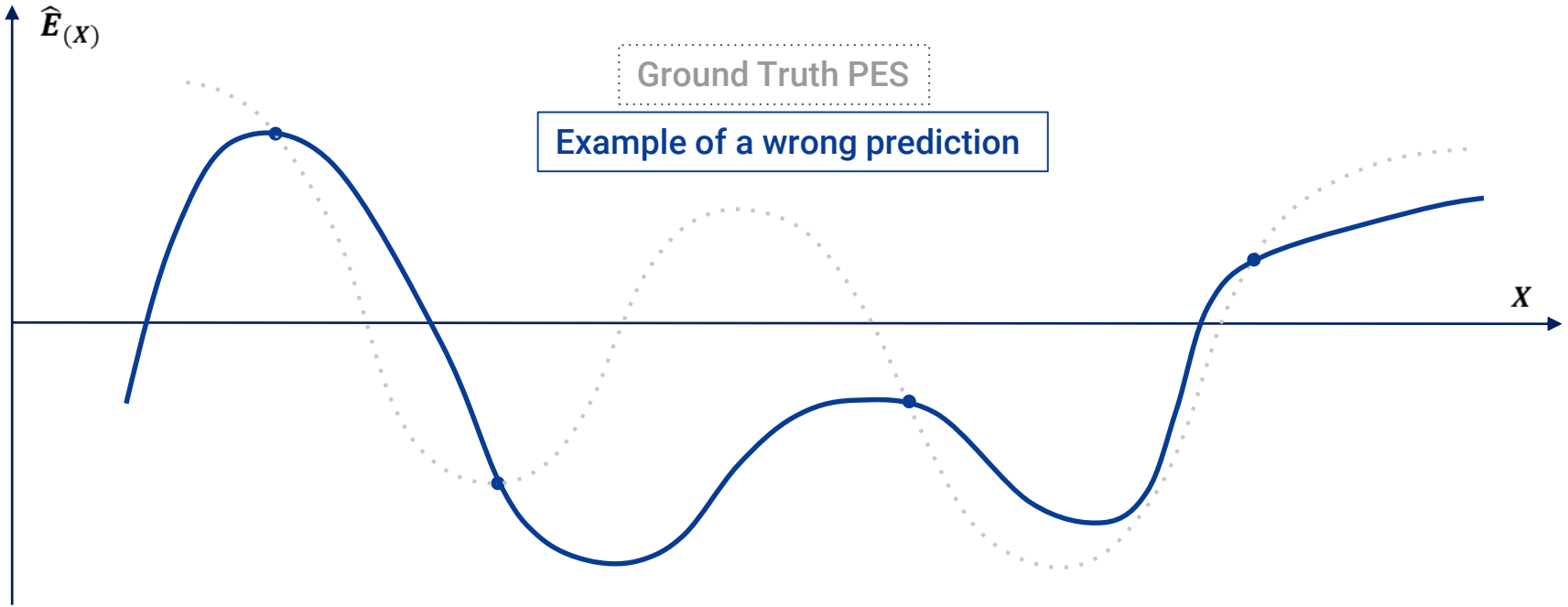
Core Issue to the Problem

- **Limited amount of data!**
  - Due to data sparsity, models have trouble learning a reliable potential energy surface (PES).



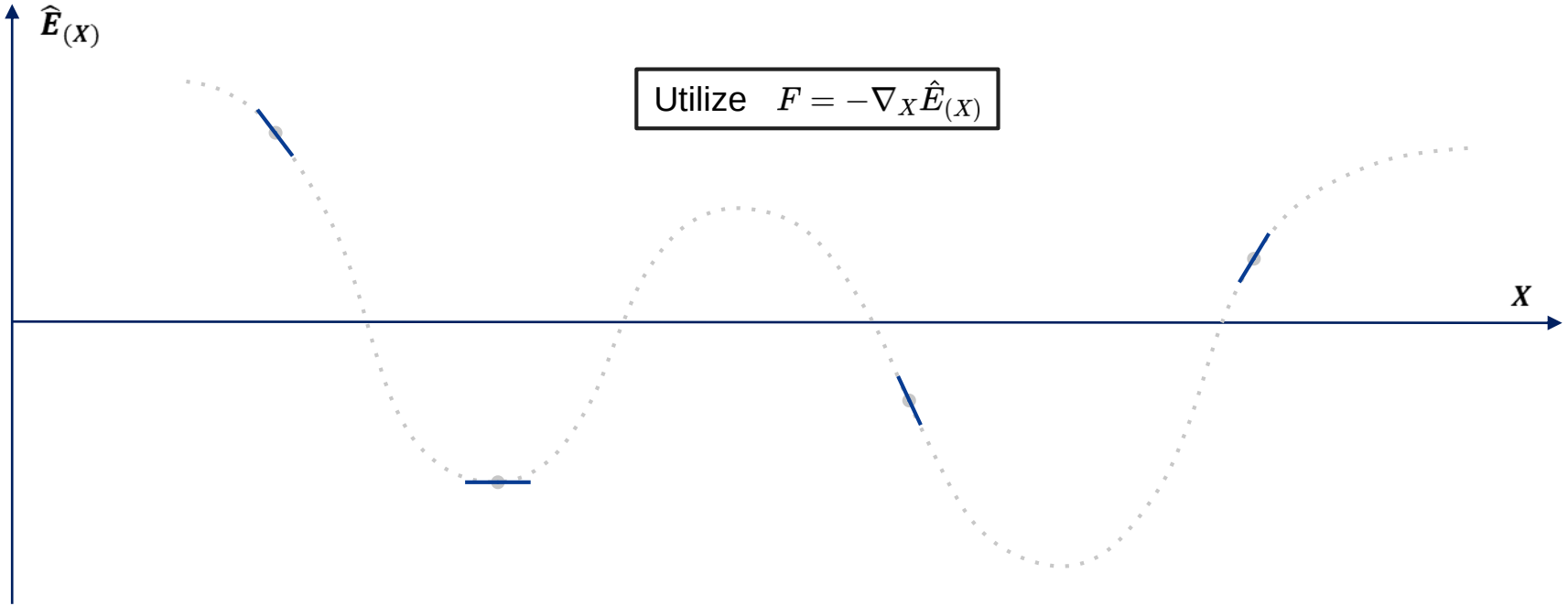
Core Issue to the Problem

- **Limited amount of data!**
  - Due to data sparsity, models have trouble learning a reliable potential energy surface (PES).



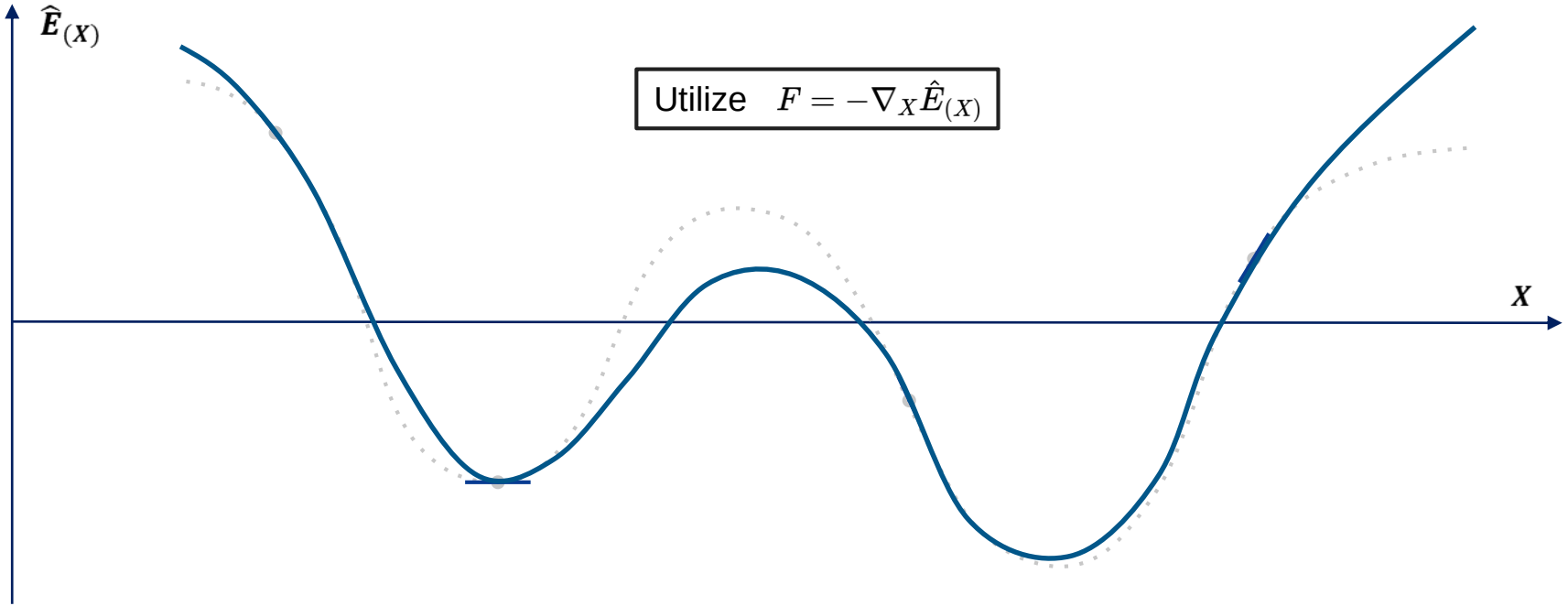
Core Issue to the Problem

- **Limited amount of data!**
  - Due to data sparsity, models have trouble learning a **reliable** potential energy surface (PES).
  - Optimally utilizing physical constraints is essential for training.



Core Issue to the Problem

- **Limited amount of data!**
  - Due to data sparsity, models have trouble learning a **reliable** potential energy surface (PES).
  - Optimally utilizing physical constraints is essential for training.



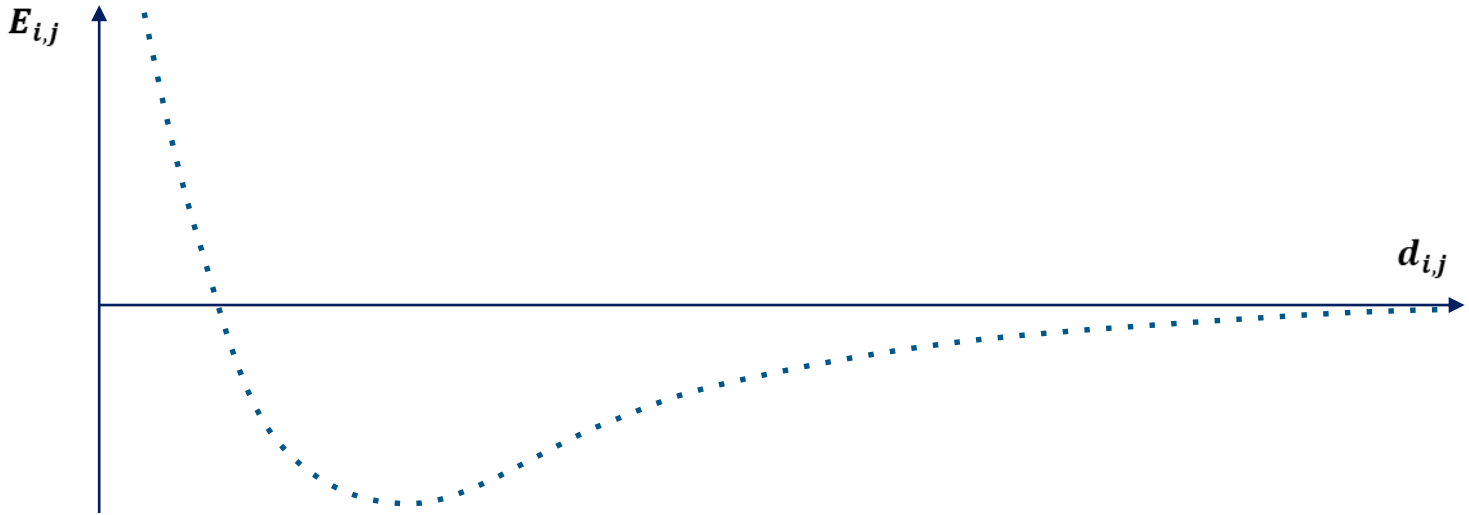
Our Approaches

Our Approach 1: Physics-model Constraints

- Parametrized energy calculation

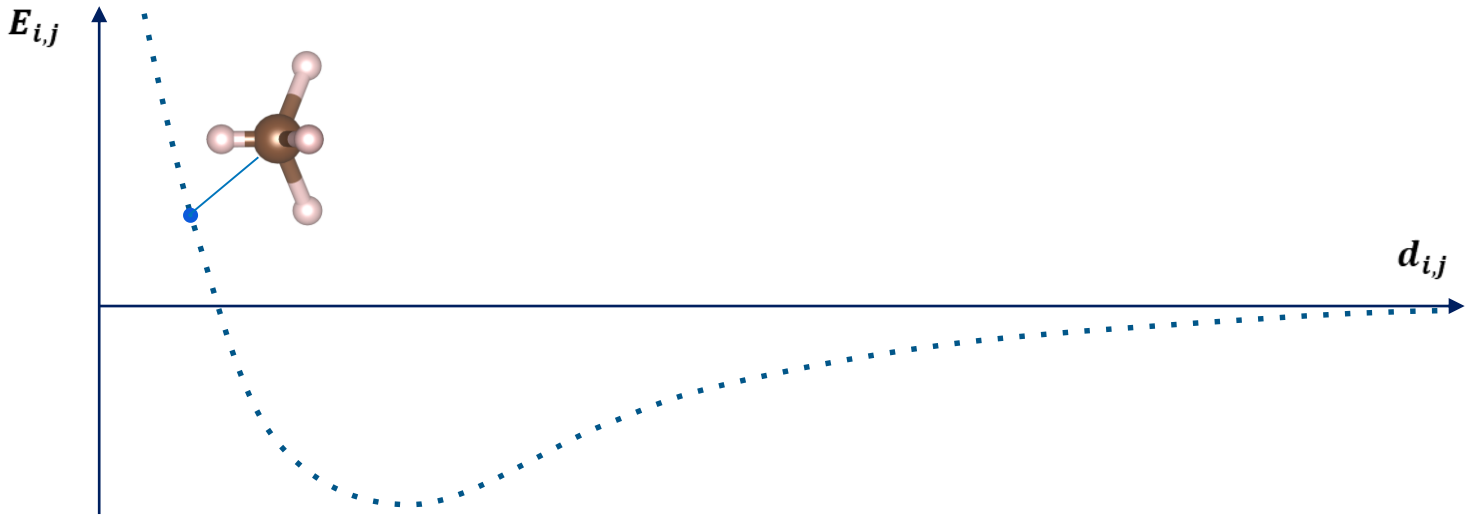
Our Approach 1: Physics-model Constraints

- Parametrized energy calculation
  - We consider self-energy and bond energy:



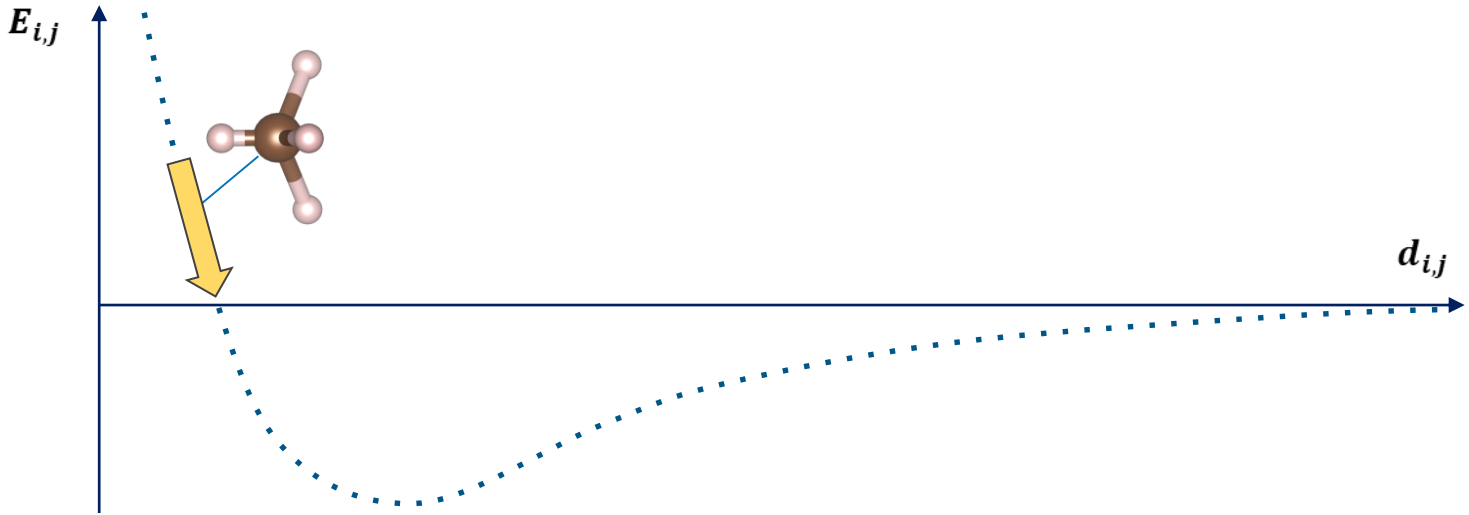
Our Approach 1: Physics-model Constraints

- Parametrized energy calculation
  - We consider self-energy and bond energy:



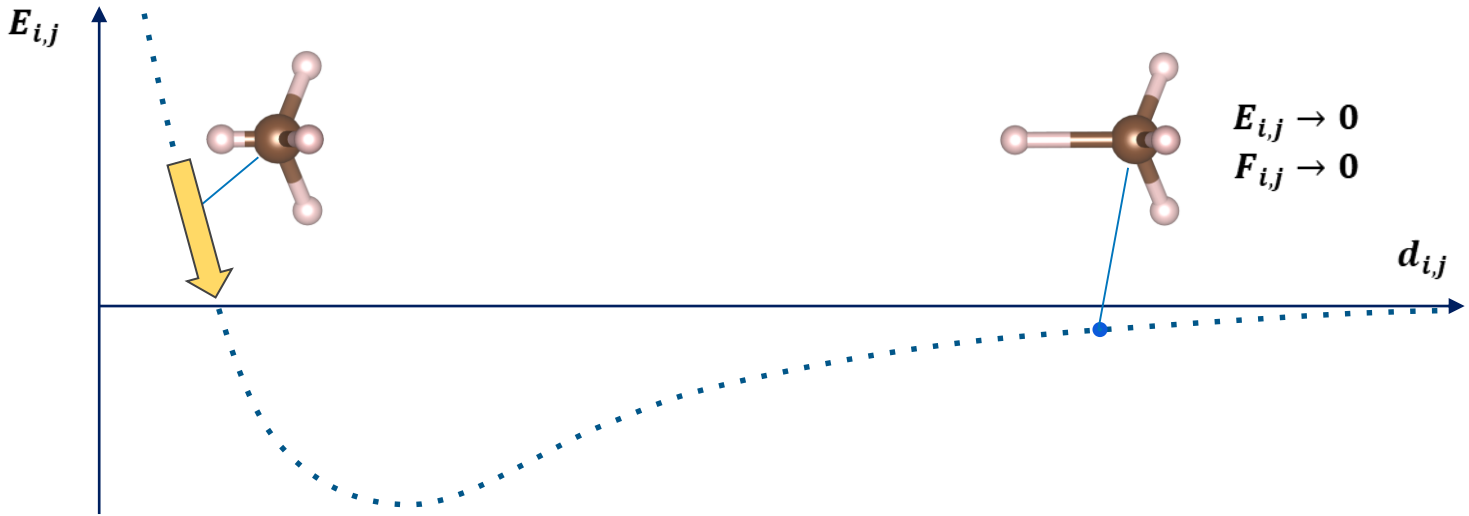
Our Approach 1: Physics-model Constraints

- Parametrized energy calculation
  - We consider self-energy and bond energy:



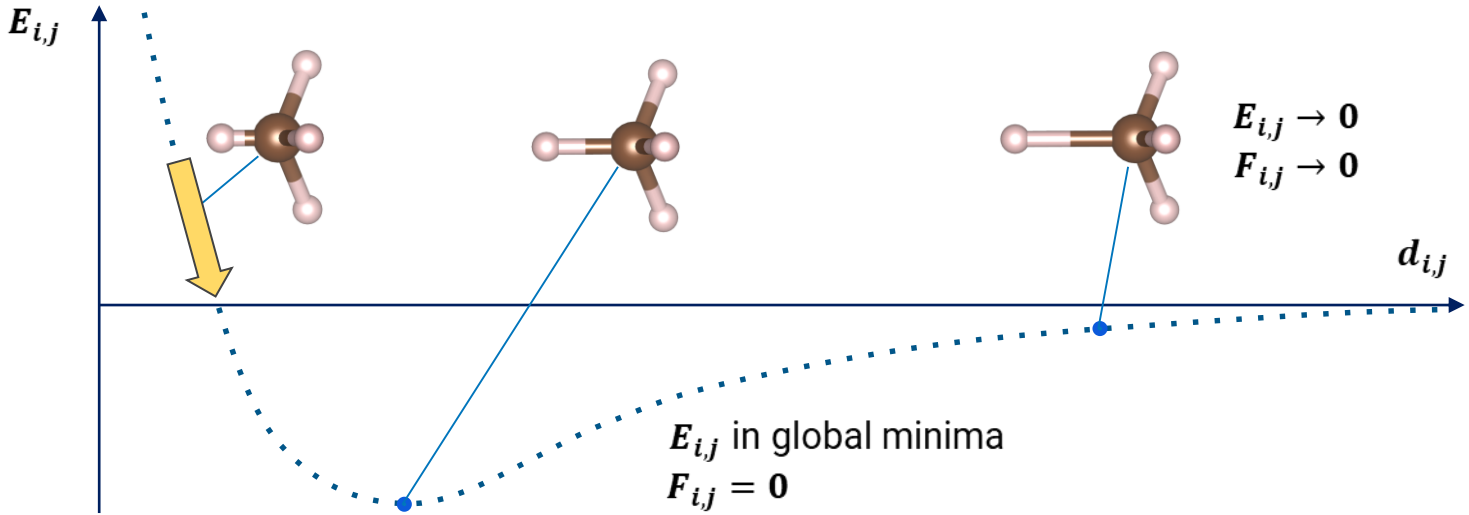
Our Approach 1: Physics-model Constraints

- Parametrized energy calculation
  - We consider self-energy and bond energy:



Our Approach 1: Physics-model Constraints

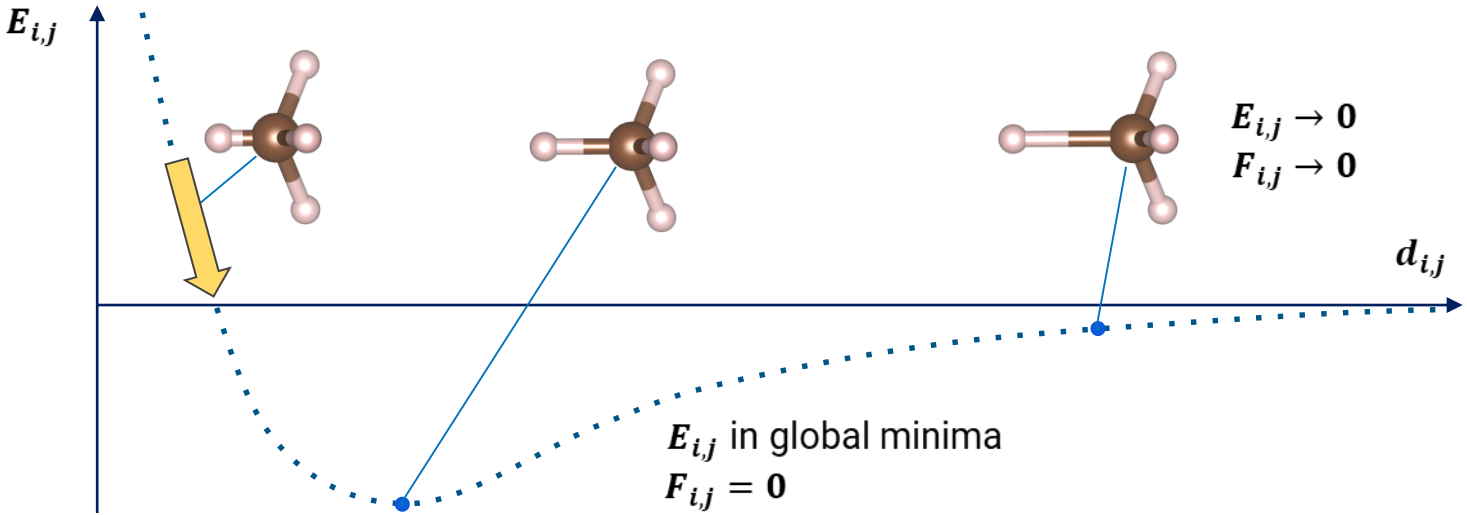
- Parametrized energy calculation
  - We consider self-energy and bond energy:



Our Approach 1: Physics-model Constraints

- Parametrized energy calculation
  - We consider self-energy and bond energy:

$$E_{i,j} = -\beta_1 \frac{\beta_0}{d_{i,j}} + \beta_2 \left[\left(\frac{\beta_4}{d_{i,j}} \right)^{2\beta_3} - 2 \left(\frac{\beta_4}{d_{i,j}} \right)^{\beta_3} \right]$$

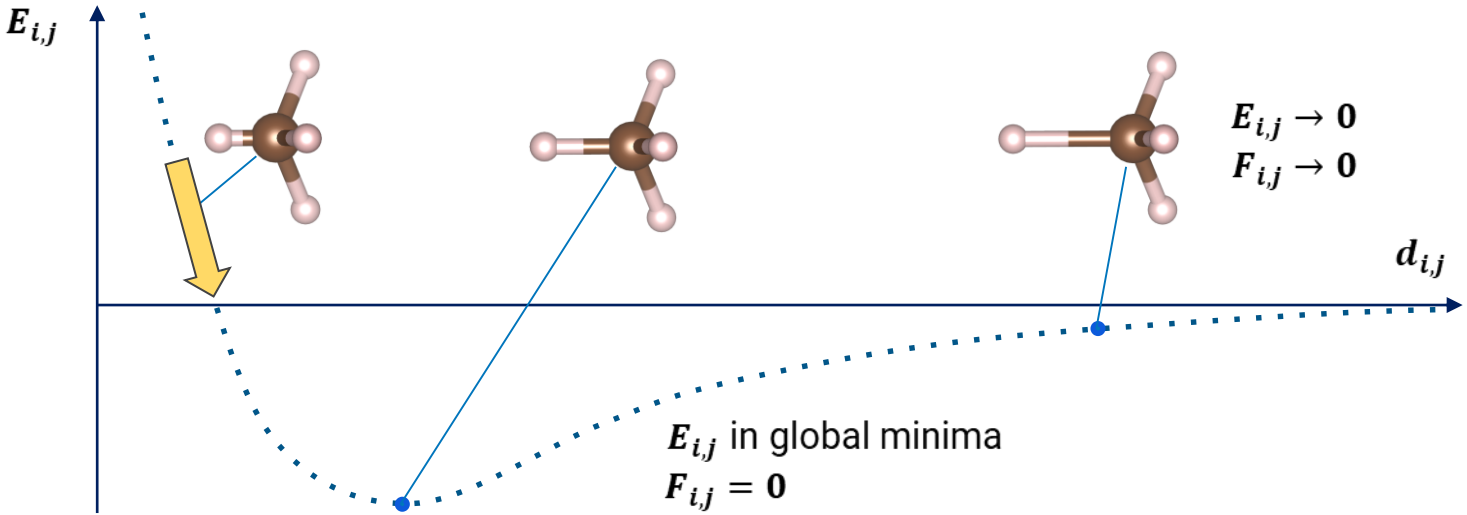


Our Approach 1: Physics-model Constraints

- Parametrized energy calculation
  - We consider self-energy and bond energy:

$$E_{i,j} = -\beta_1 \frac{\beta_0}{d_{i,j}} + \beta_2 \left[\left(\frac{\beta_4}{d_{i,j}} \right)^{2\beta_3} - 2 \left(\frac{\beta_4}{d_{i,j}} \right)^{\beta_3} \right]$$

Van der Waals: $V_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]$



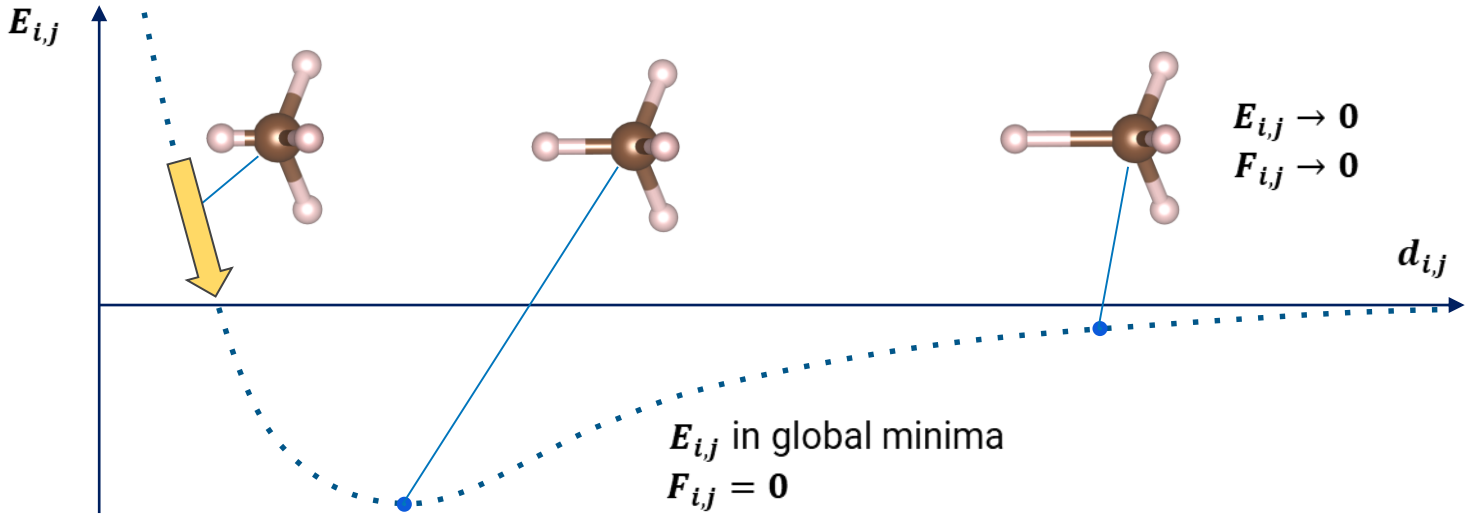
Our Approach 1: Physics-model Constraints

- Parametrized energy calculation
  - We consider self-energy and bond energy:

$$E_{i,j} = -\beta_1 \frac{\beta_0}{d_{i,j}} + \beta_2 \left[\left(\frac{\beta_4}{d_{i,j}} \right)^{2\beta_3} - 2 \left(\frac{\beta_4}{d_{i,j}} \right)^{\beta_3} \right]$$

Electrostatic: $V_E(r) = k \cdot \frac{q_1 q_2}{r}$

Van der Waals: $V_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]$



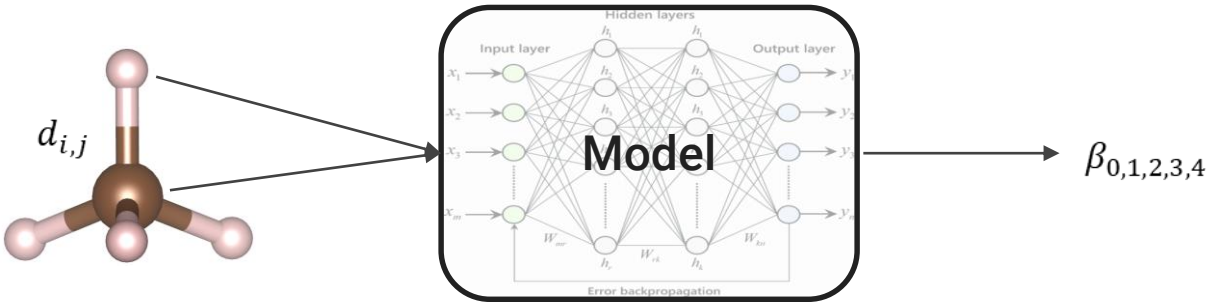
Our Approach 1: Physics-model Constraints

- **Parametrized** energy calculation
  - We consider self-energy and bond energy:

$$E_{i,j} = -\beta_1 \frac{\beta_0}{d_{i,j}} + \beta_2 \left[\left(\frac{\beta_4}{d_{i,j}} \right)^{2\beta_3} - 2 \left(\frac{\beta_4}{d_{i,j}} \right)^{\beta_3} \right]$$

Electrostatic: $V_E(r) = k \cdot \frac{q_1 q_2}{r}$ Van der Waals: $V_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]$

  - Parametrized by β, which depends on the species of pair of atoms



Our Approach 2: Physics-inspired Constraints

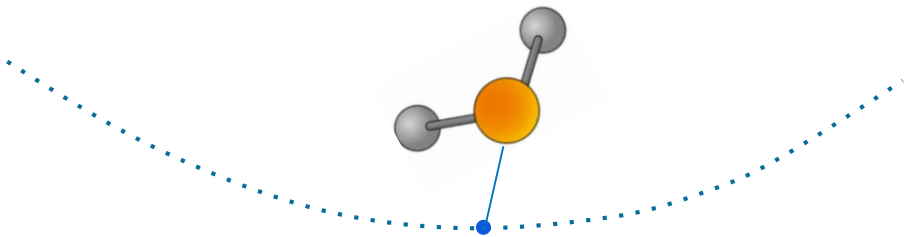
Two constraints when given a stable molecular structure:

Our Approach 2: Physics-inspired Constraints

Two constraints when given a stable molecular structure:

- Inequality bound condition
  - Locally, a perturbed state energy must be greater than the stable state.

$$\mathcal{L}_{\text{bound}} = \begin{cases} \hat{E}_{\text{mol}} - \hat{E}^*_{\text{mol}} & \text{if } \hat{E}^*_{\text{mol}} \leq \hat{E}_{\text{mol}} \\ 0 & \text{otherwise.} \end{cases}$$

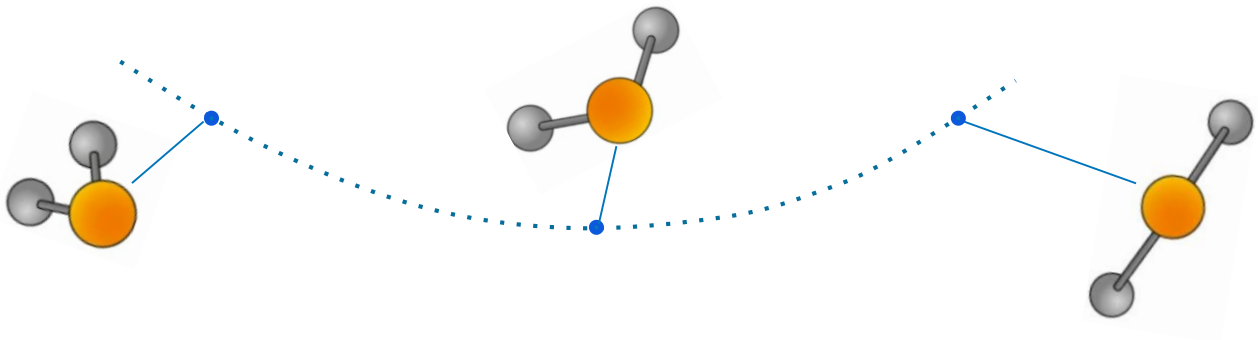


Our Approach 2: Physics-inspired Constraints

Two constraints when given a stable molecular structure:

- Inequality bound condition
  - Locally, a perturbed state energy must be greater than the stable state.

$$\mathcal{L}_{\text{bound}} = \begin{cases} \hat{E}_{\text{mol}} - \hat{E}^*_{\text{mol}} & \text{if } \hat{E}^*_{\text{mol}} \leq \hat{E}_{\text{mol}} \\ 0 & \text{otherwise.} \end{cases}$$

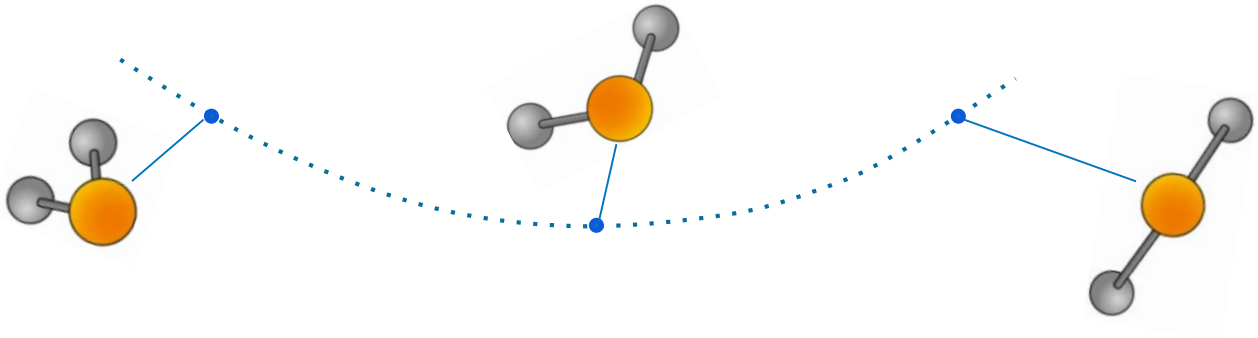


Our Approach 2: Physics-inspired Constraints

Two constraints when given a stable molecular structure:

- Inequality bound condition
  - Locally, a perturbed state energy must be greater than the stable state.

$$\mathcal{L}_{\text{bound}} = \begin{cases} \hat{E}_{\text{mol}} - \hat{E}^{*}_{\text{mol}} & \text{if } \hat{E}^{*}_{\text{mol}} \leq \hat{E}_{\text{mol}} \\ 0 & \text{otherwise.} \end{cases}$$



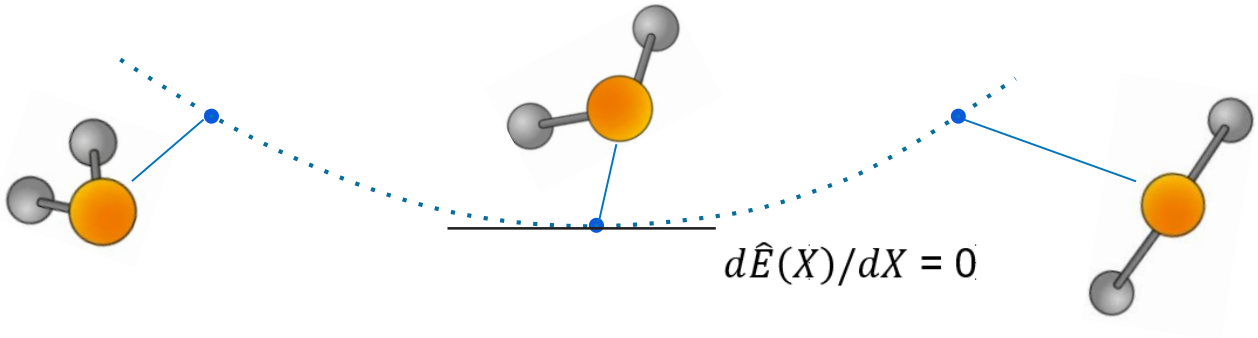
Our Approach 2: Physics-inspired Constraints

Two constraints when given a **stable** molecular structure:

- Inequality bound condition
  - Locally, a perturbed state energy must be greater than the stable state.

$$\mathcal{L}_{\text{bound}} = \begin{cases} \hat{E}_{\text{mol}} - \hat{E}^*_{\text{mol}} & \text{if } \hat{E}^*_{\text{mol}} \leq \hat{E}_{\text{mol}} \\ 0 & \text{otherwise.} \end{cases}$$

- Zero-force condition
  - Net force of the stable structure must be zero.



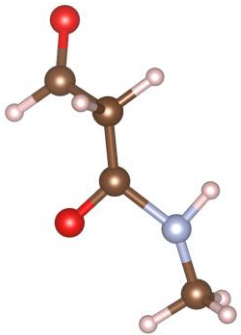
Our Approach 3: Masked Atomic Modeling

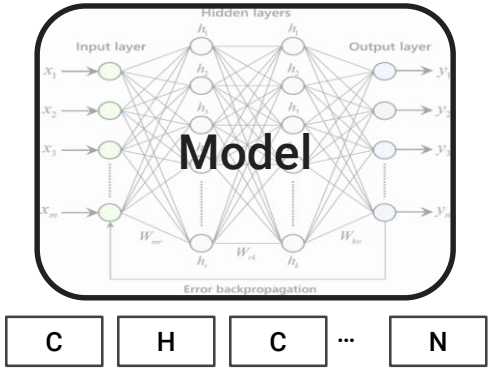
- Masked Atomic Modeling

Concurrently developed in G. Zhou et al., ICLR 2023 (Uni-Mol: A Universal 3D Molecular Representation Learning Framework).

Our Approach 3: Masked Atomic Modeling

- Masked Atomic Modeling
  - Zero-out 30% of the entire atomic embeddings in a given molecule.

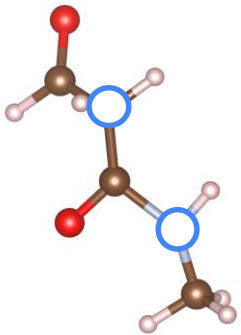


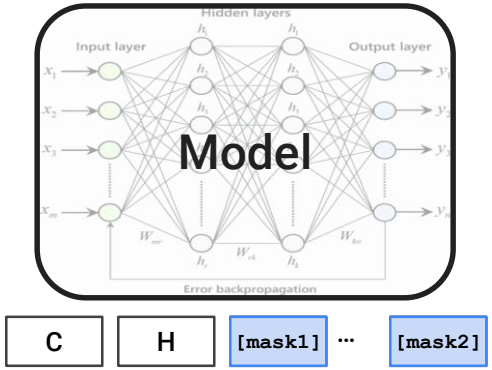


Concurrently developed in G. Zhou et al., ICLR 2023 (Uni-Mol: A Universal 3D Molecular Representation Learning Framework).

Our Approach 3: Masked Atomic Modeling

- Masked Atomic Modeling
  - Zero-out 30% of the entire atomic embeddings in a given molecule.

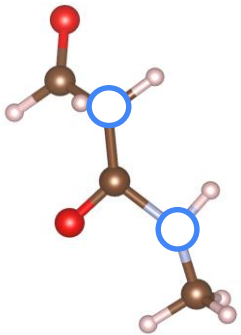


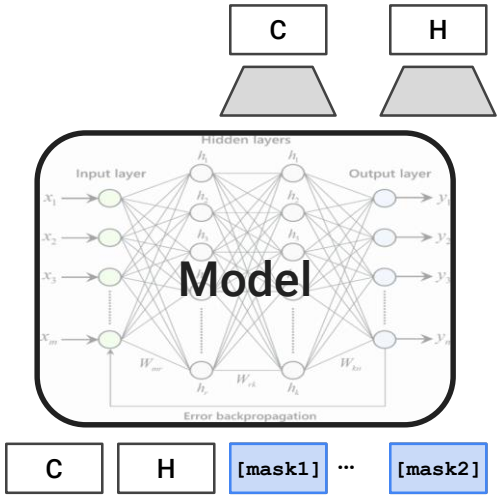


Concurrently developed in G. Zhou et al., ICLR 2023 (Uni-Mol: A Universal 3D Molecular Representation Learning Framework).

Our Approach 3: Masked Atomic Modeling

- Masked Atomic Modeling
  - Zero-out 30% of the entire atomic embeddings in a given molecule.

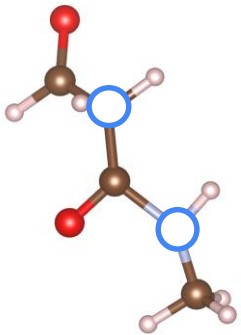


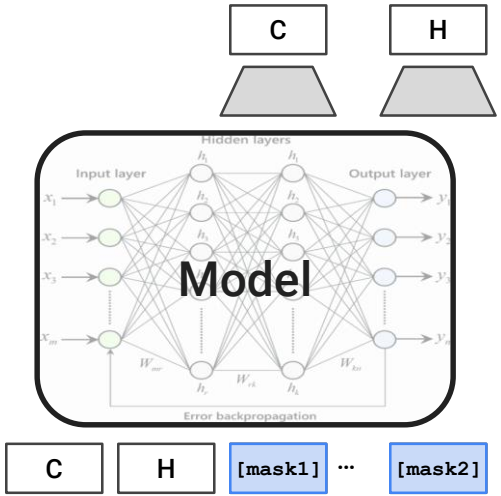


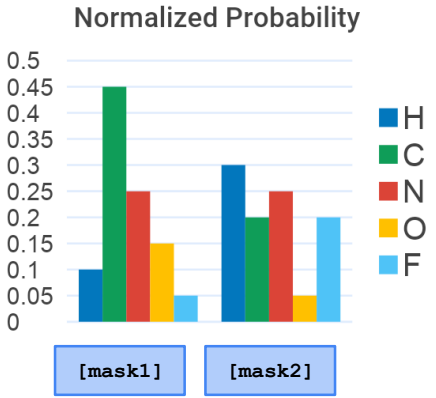
Concurrently developed in G. Zhou et al., ICLR 2023 (Uni-Mol: A Universal 3D Molecular Representation Learning Framework).

Our Approach 3: Masked Atomic Modeling

- Masked Atomic Modeling
  - Zero-out 30% of the entire atomic embeddings in a given molecule.
  - The model must learn fundamental bonding principles to recover them.



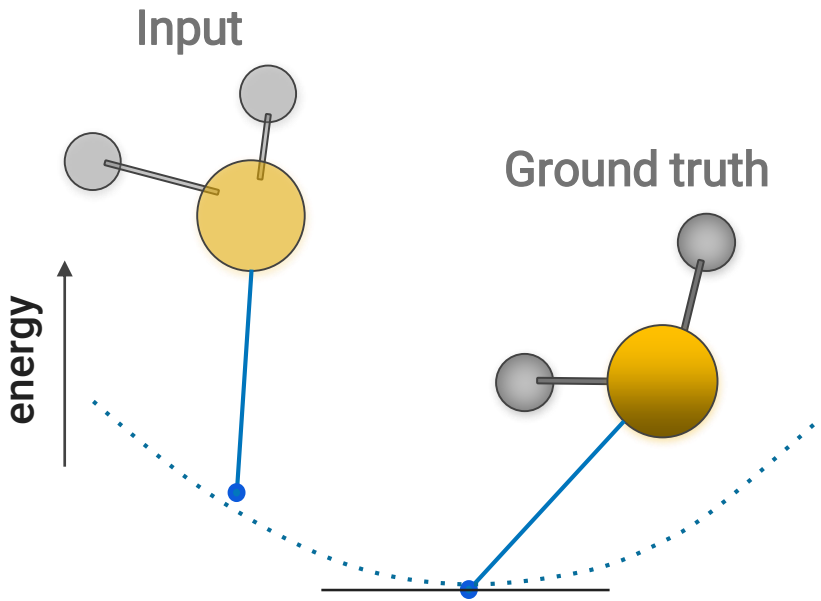




Concurrently developed in G. Zhou et al., ICLR 2023 (Uni-Mol: A Universal 3D Molecular Representation Learning Framework).

Experimental Results

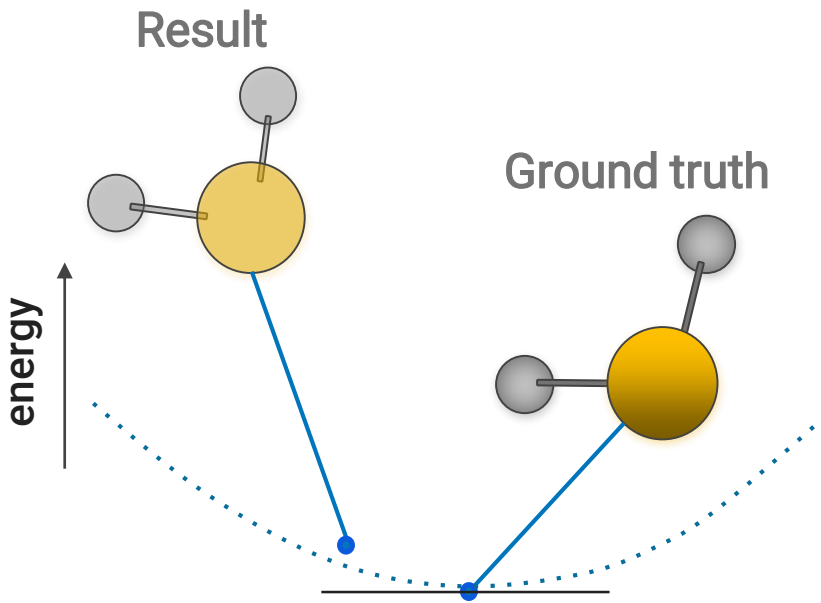
New metrics: ΔE and ΔP



New metrics: ΔE and ΔP

ΔP : Structural distortion

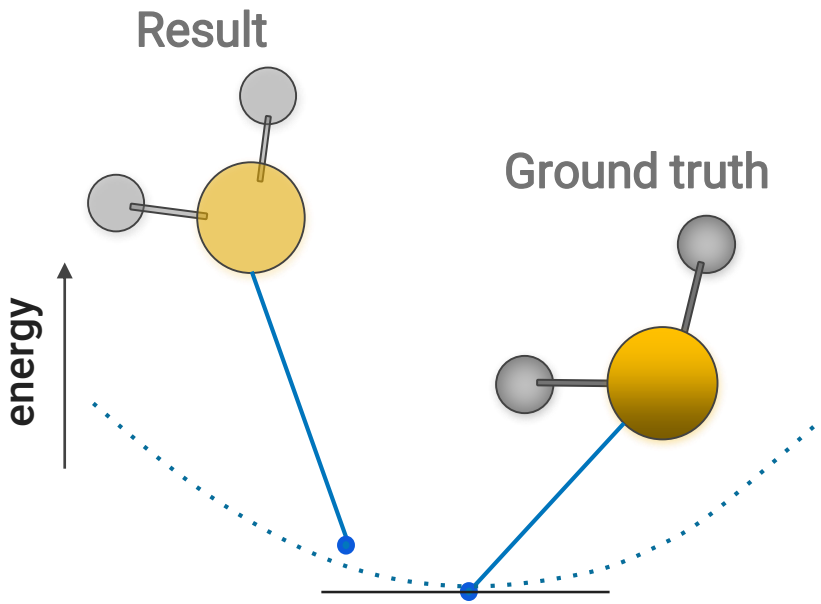
ΔE : Energy difference



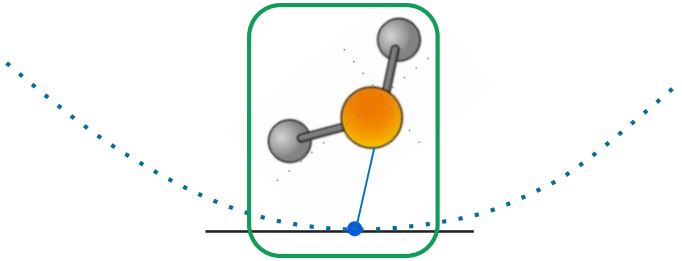
New metrics: ΔE and ΔP

ΔP : Structural distortion

ΔE : Energy difference



Ideal case: ΔE = 0 and ΔP = 0

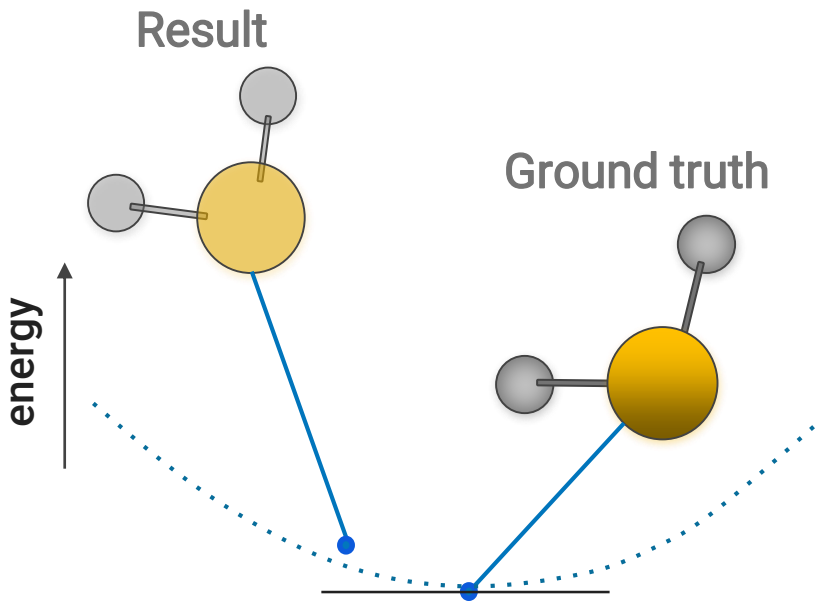


Optimized structure = Stable structure

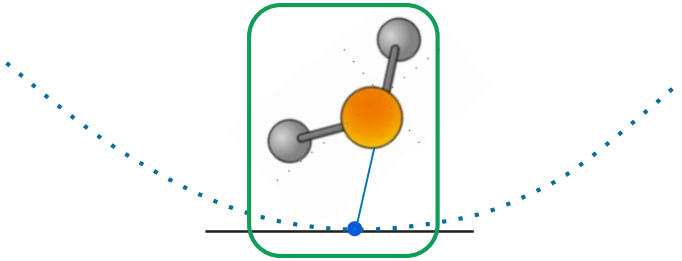
New metrics: ΔE and ΔP

ΔP : Structural distortion

ΔE : Energy difference

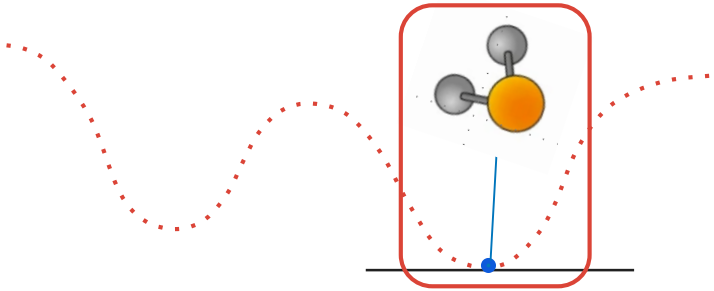


Ideal case: ΔE = 0 and ΔP = 0



Optimized structure = Stable structure

Problematic case: ΔE ≠ 0 and ΔP ≠ 0



Optimized structure ≠ Stable structure

Comparison with Other Models: Analysis

Dataset (Task) / Model	QM9 $\text{MAE}_\text{E}(\downarrow)$	QM9 $\text{MAE}_\text{F}(\downarrow)$	QM9 $\Delta P(\downarrow)$	OC20 (IS2RE) $\text{MAE}_\text{E}(\downarrow)$	OC20 (IS2RE) $\Delta P(\downarrow)$
SchNet [Schütt et al., 2018]	14.00	2.64	0.47	1.059	0.60
CGCNN [Xie and Grossman, 2018]	–	–	–	0.988	0.58
MXMNet [Zhang et al., 2020b]	**5.90**	1.83	1.57	–	–
DimeNet [Gasteiger et al., 2020]	8.02	1.79	0.58	1.012	0.55
ForceNet [Hu et al., 2021]	18.62	0.41	0.21	–	–
TorchMDNet (ET) [Thölke and De Fabritiis, 2022]	6.15	1.15	0.32	–	–
Ours ($\mathcal{L}_{\text{energy}}$ only)	8.35	1.28	1.23	–	–
Ours (full model)	15.16±0.539	**0.0057**±0.001	**0.0251**±0.01	**0.887**±0.024	**0.10**±0.01
p-value	–	0	3.2×10^{-7}	2.6×10^{-4}	7.0×10^{-8}

Comparison with Other Models: Analysis

Energy estimation error

Dataset (Task) / Model	QM9 $\text{MAE}_\text{E}(\downarrow)$	QM9 $\text{MAE}_\text{F}(\downarrow)$	QM9 $\Delta P(\downarrow)$	OC20 (IS2RE) $\text{MAE}_\text{E}(\downarrow)$	OC20 (IS2RE) $\Delta P(\downarrow)$
SchNet [Schütt et al., 2018]	14.00	2.64	0.47	1.059	0.60
CGCNN [Xie and Grossman, 2018]	–	–	–	0.988	0.58
MXMNet [Zhang et al., 2020b]	**5.90**	1.83	1.57	–	–
DimeNet [Gasteiger et al., 2020]	8.02	1.79	0.58	1.012	0.55
ForceNet [Hu et al., 2021]	18.62	0.41	0.21	–	–
TorchMDNet (ET) [Thölke and De Fabritiis, 2022]	6.15	1.15	0.32	–	–
Ours ($\mathcal{L}_{\text{energy}}$ only)	8.35	1.28	1.23	–	–
Ours (full model)	15.16±0.539	**0.0057**±0.001	**0.0251**±0.01	**0.887**±0.024	**0.10**±0.01
p-value	–	0	3.2×10^{-7}	2.6×10^{-4}	7.0×10^{-8}

Comparison with Other Models: Analysis

Dataset (Task) / Model	QM9 Energy estimation error $\mathrm{MAE_E}(\downarrow)$	QM9 Net force (ideally 0) $\mathrm{MAE_F}(\downarrow)$	QM9 $\Delta P(\downarrow)$	OC20 (IS2RE) $\mathrm{MAE_E}(\downarrow)$	OC20 (IS2RE) $\Delta P(\downarrow)$
SchNet [Schütt et al., 2018]	14.00	2.64	0.47	1.059	0.60
CGCNN [Xie and Grossman, 2018]	–	–	–	0.988	0.58
MXMNet [Zhang et al., 2020b]	**5.90**	1.83	1.57	–	–
DimeNet [Gasteiger et al., 2020]	8.02	1.79	0.58	1.012	0.55
ForceNet [Hu et al., 2021]	18.62	0.41	0.21	–	–
TorchMDNet (ET) [Thölke and De Fabritiis, 2022]	6.15	1.15	0.32	–	–
Ours ($\mathcal{L}_{\text{energy}}$ only)	8.35	1.28	1.23	–	–
Ours (full model)	15.16±0.539	**0.0057**±0.001	**0.0251**±0.01	**0.887**±0.024	**0.10**±0.01
p-value	–	0	3.2×10^{-7}	2.6×10^{-4}	7.0×10^{-8}

Comparison with Other Models: Analysis

Dataset (Task) / Model	Energy estimation error QM9 $MAE_E(\downarrow)$	Net force (ideally 0) QM9 $MAE_F(\downarrow)$	Atom position distortion QM9 $\Delta P(\downarrow)$	OC20 (IS2RE) $MAE_E(\downarrow)$	OC20 (IS2RE) $\Delta P(\downarrow)$
SchNet [Schütt et al., 2018]	14.00	2.64	0.47	1.059	0.60
CGCNN [Xie and Grossman, 2018]	–	–	–	0.988	0.58
MXMNet [Zhang et al., 2020b]	**5.90**	1.83	1.57	–	–
DimeNet [Gasteiger et al., 2020]	8.02	1.79	0.58	1.012	0.55
ForceNet [Hu et al., 2021]	18.62	0.41	0.21	–	–
TorchMDNet (ET) [Thölke and De Fabritiis, 2022]	6.15	1.15	0.32	–	–
Ours ($\mathcal{L}_{\text{energy}}$ only)	8.35	1.28	1.23	–	–
Ours (full model)	15.16±0.539	**0.0057**±0.001	**0.0251**±0.01	**0.887**±0.024	**0.10**±0.01
p-value	–	0	3.2×10^{-7}	2.6×10^{-4}	7.0×10^{-8}

Comparison with Other Models: Analysis

Dataset (Task)	Energy estimation error	Net force (ideally 0)	Atom position distortion	OC20 (IS2RE)	
Model	QM9				
	$\text{MAE}_\text{E}(\downarrow)$	$\text{MAE}_\text{F}(\downarrow)$	$\Delta P(\downarrow)$	$\text{MAE}_\text{E}(\downarrow)$	$\Delta P(\downarrow)$
SchNet [Schütt et al., 2018]	14.00	2.64	0.47	1.059	0.60
CGCNN [Xie and Grossman, 2018]	–	–	–	0.988	0.58
MXMNet [Zhang et al., 2020b]	**5.90**	1.83	1.57	–	–
DimeNet [Gasteiger et al., 2020]	8.02	1.79	0.58	1.012	0.55
ForceNet [Hu et al., 2021]	18.62	0.41	0.21	–	–
TorchMDNet (ET) [Thölke and De Fabritiis, 2022]	6.15	1.15	0.32	–	–
Ours ($\mathcal{L}_\text{energy}$ only)	8.35	1.28	1.23	–	–
Ours (full model)	15.16±0.539	**0.0057**±0.001	**0.0251**±0.01	**0.887**±0.024	**0.10**±0.01
p-value	–	0	3.2×10^{-7}	2.6×10^{-4}	7.0×10^{-8}

We achieve comparable energy estimation to the baselines.

Comparison with Other Models: Analysis

Dataset (Task)		Energy estimation error	Net force (ideally 0)	Atom position distortion		
		QM9			OC20 (IS2RE)	
Model		$\text{MAE}_\text{E}(\downarrow)$	$\text{MAE}_\text{F}(\downarrow)$	$\Delta P(\downarrow)$	$\text{MAE}_\text{E}(\downarrow)$	$\Delta P(\downarrow)$
SchNet [Schütt et al., 2018]		14.00	2.64	0.47	1.059	0.60
CGCNN [Xie and Grossman, 2018]		–	–	–	0.988	0.58
MXMNet [Zhang et al., 2020b]		**5.90**	1.83	1.57	–	–
DimeNet [Gasteiger et al., 2020]		8.02	1.79	0.58	1.012	0.55
ForceNet [Hu et al., 2021]		18.62	0.41	0.21	–	–
TorchMDNet (ET) [Thölke and De Fabritiis, 2022]		6.15	1.15	0.32	–	–
Ours ($\mathcal{L}_{\text{energy}}$ only)		8.35	1.28	1.23	–	–
Ours (full model)		15.16±0.539	**0.0057**±0.001	**0.0251**±0.01	**0.887**±0.024	**0.10**±0.01
p-value		–	0	3.2×10^{-7}	2.6×10^{-4}	7.0×10^{-8}

We achieve comparable energy estimation to the baselines.

Force error and ΔP get considerably better than baselines with our full model!

Comparison with Other Models: Analysis

Baseline models fail to learn the actual energy surface, indicated by high force and position error.

Dataset (Task) / Model	Energy estimation error: QM9 $MAE_E(\downarrow)$	Net force (ideally 0): QM9 $MAE_F(\downarrow)$	Atom position distortion: QM9 $\Delta P(\downarrow)$	OC20 (IS2RE) $MAE_E(\downarrow)$	OC20 (IS2RE) $\Delta P(\downarrow)$
SchNet [Schütt et al., 2018]	14.00	2.64	0.47	1.059	0.60
CGCNN [Xie and Grossman, 2018]	–	–	–	0.988	0.58
MXMNet [Zhang et al., 2020b]	**5.90**	1.83	1.57	–	–
DimeNet [Gasteiger et al., 2020]	8.02	1.79	0.58	1.012	0.55
ForceNet [Hu et al., 2021]	18.62	0.41	0.21	–	–
TorchMDNet (ET) [Thölke and De Fabritiis, 2022]	6.15	1.15	0.32	–	–
Ours ($\mathcal{L}_{\text{energy}}$ only)	8.35	1.28	1.23	–	–
Ours (full model)	15.16±0.539	**0.0057**±0.001	**0.0251**±0.01	**0.887**±0.024	**0.10**±0.01
p-value	–	0	3.2×10^{-7}	2.6×10^{-4}	7.0×10^{-8}

We achieve comparable energy estimation to the baselines.

Force error and ΔP get considerably better than baselines with our full model!

Comparison with Other Models: Analysis

Baseline models fail to learn the actual energy surface, **indicated by high force and position error.**

It is same for our model when it is trained only on the energy loss. → **It is indeed important to** consider other physical constraints!

Energy estimation error — Net force (ideally 0) — Atom position distortion

Dataset (Task) / Model	QM9 $\text{MAE}_\text{E}(\downarrow)$	QM9 $\text{MAE}_\text{F}(\downarrow)$	QM9 $\Delta P(\downarrow)$	OC20 (IS2RE) $\text{MAE}_\text{E}(\downarrow)$	OC20 (IS2RE) $\Delta P(\downarrow)$
SchNet [Schütt et al., 2018]	14.00	2.64	0.47	1.059	0.60
CGCNN [Xie and Grossman, 2018]	–	–	–	0.988	0.58
MXMNet [Zhang et al., 2020b]	**5.90**	1.83	1.57	–	–
DimeNet [Gasteiger et al., 2020]	8.02	1.79	0.58	1.012	0.55
ForceNet [Hu et al., 2021]	18.62	0.41	0.21	–	–
TorchMDNet (ET) [Thölke and De Fabritiis, 2022]	6.15	1.15	0.32	–	–
Ours ($\mathcal{L}_{\text{energy}}$ only)	8.35	1.28	1.23	–	–
Ours (full model)	15.16±0.539	**0.0057**±0.001	**0.0251**±0.01	**0.887**±0.024	**0.10**±0.01
p-value	–	0	3.2×10^{-7}	2.6×10^{-4}	7.0×10^{-8}

We achieve comparable energy estimation to the baselines.

Force error and ΔP get considerably better than baselines with our full model!

Comparison with Other Models: Analysis

Baseline models fail to learn the actual energy surface, **indicated by high force and position error.**

It is same for our model when it is trained only on the energy loss. → **It is indeed important to** consider other physical constraints!

Energy estimation error — Net force (ideally 0) — Atom position distortion

Dataset (Task) / Model	QM9 $\text{MAE}_E(\downarrow)$	QM9 $\text{MAE}_F(\downarrow)$	QM9 $\Delta P(\downarrow)$	OC20 (IS2RE) $\text{MAE}_E(\downarrow)$	OC20 (IS2RE) $\Delta P(\downarrow)$
SchNet [Schütt et al., 2018]	14.00	2.64	0.47	1.059	0.60
CGCNN [Xie and Grossman, 2018]	–	–	–	0.988	0.58
MXMNet [Zhang et al., 2020b]	**5.90**	1.83	1.57	–	–
DimeNet [Gasteiger et al., 2020]	8.02	1.79	0.58	1.012	0.55
ForceNet [Hu et al., 2021]	18.62	0.41	0.21	–	–
TorchMDNet (ET) [Thölke and De Fabritiis, 2022]	6.15	1.15	0.32	–	–
Ours ($\mathcal{L}_{\text{energy}}$ only)	8.35	1.28	1.23	–	–
Ours (full model)	15.16±0.539	**0.0057**±0.001	**0.0251**±0.01	**0.887**±0.024	**0.10**±0.01
p-value	–	0	3.2×10^{-7}	2.6×10^{-4}	7.0×10^{-8}

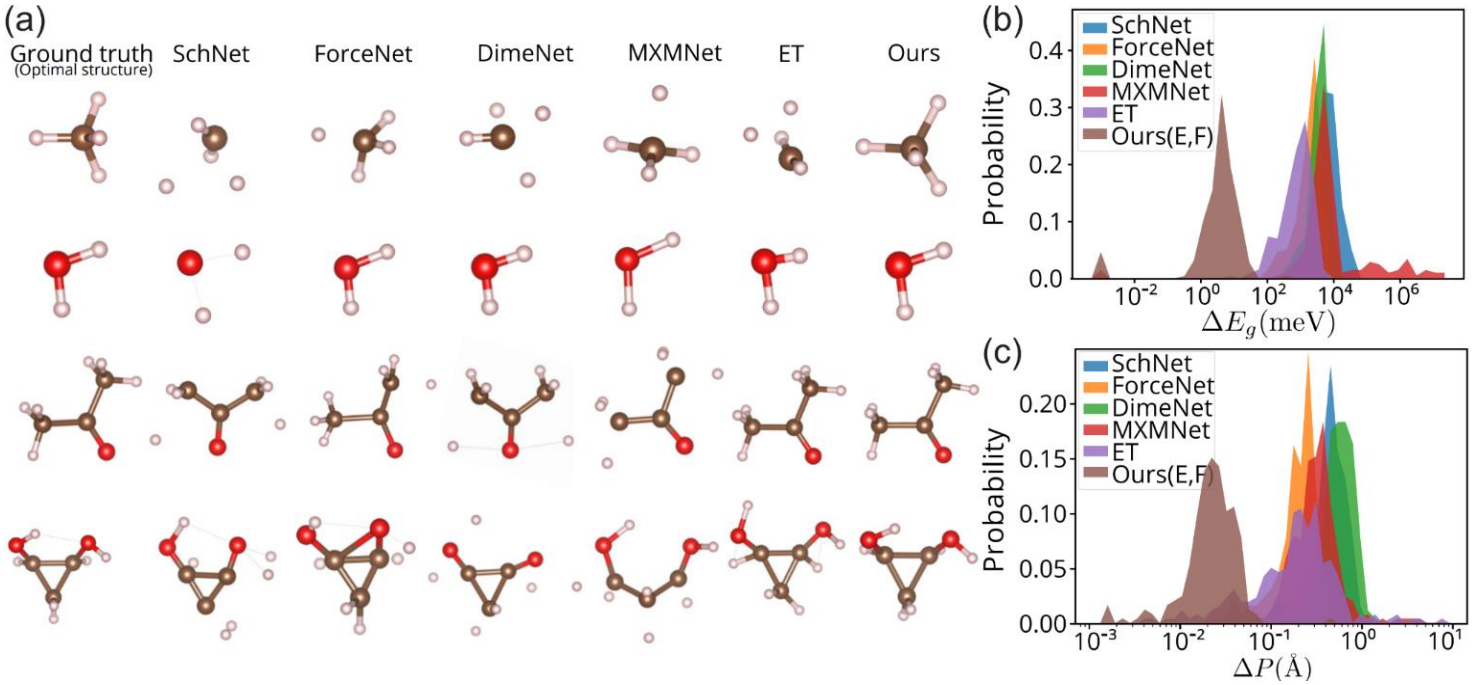
We achieve comparable energy estimation to the baselines.

Force error and ΔP get considerably better than baselines with our full model!

Similarly, our model outperforms baselines on OC20 dataset as well.

Qualitative Results: Structure Optimization

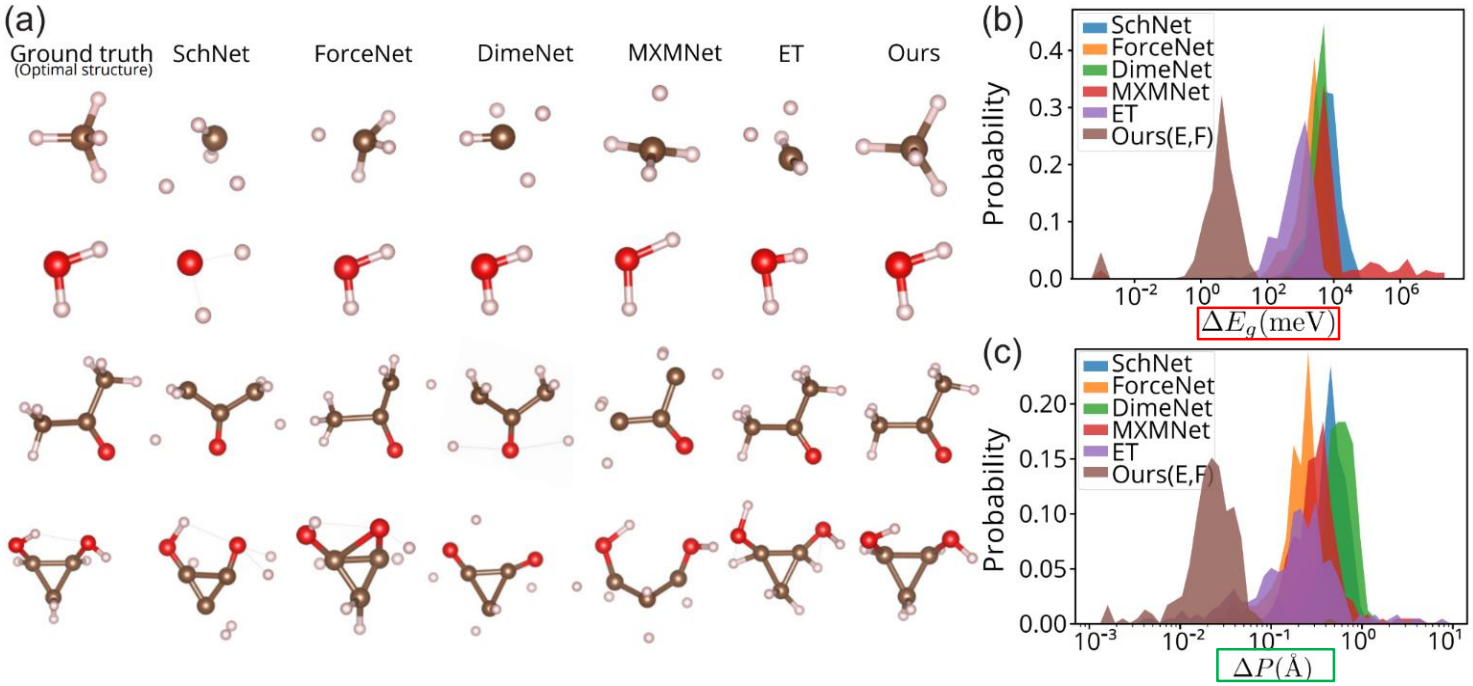
Reduced ΔP and ΔE more than 10x compared to other baselines.



Conducted on QM9 dataset.

Qualitative Results: Structure Optimization

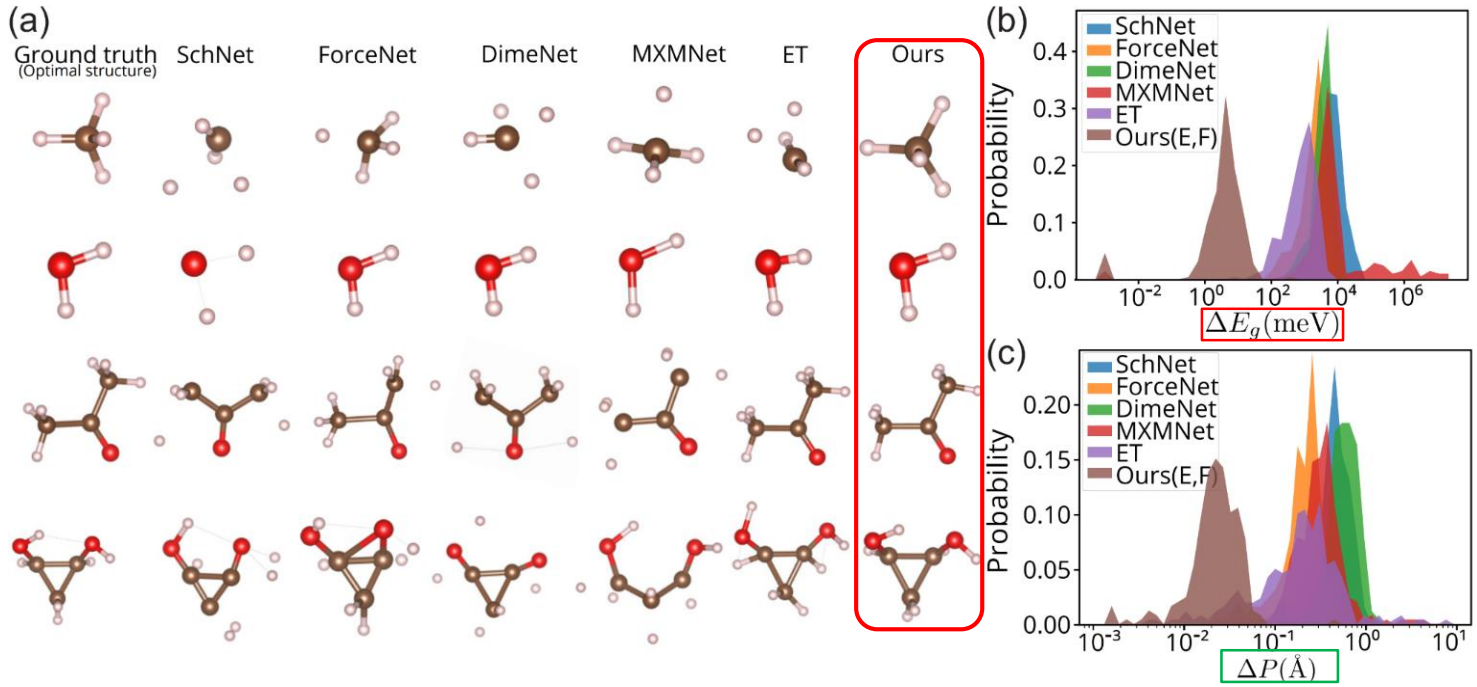
Reduced ΔP and ΔE more than 10x compared to other baselines.



Conducted on QM9 dataset.

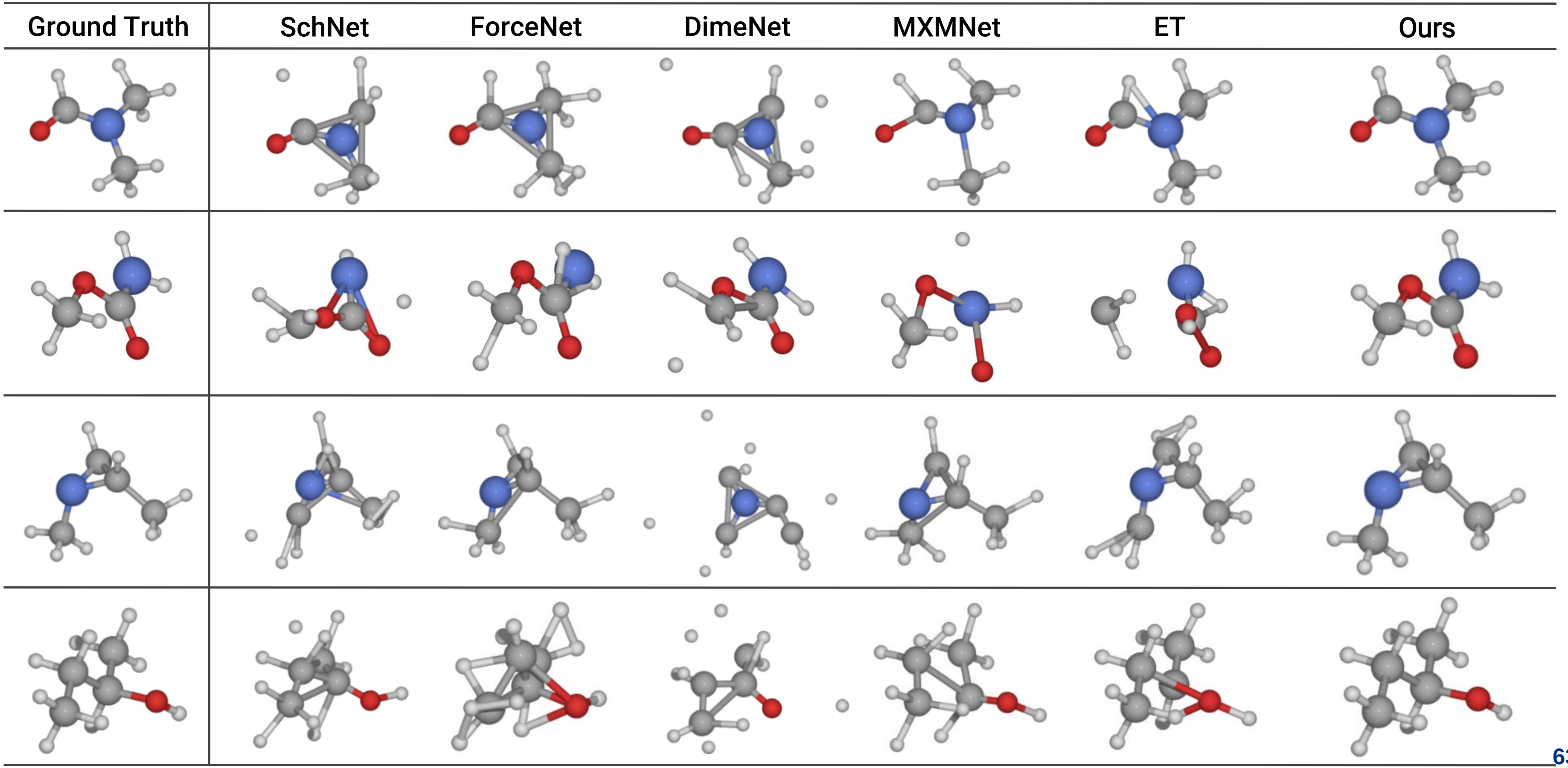
Qualitative Results: Structure Optimization

Reduced ΔP and ΔE more than 10x compared to other baselines.

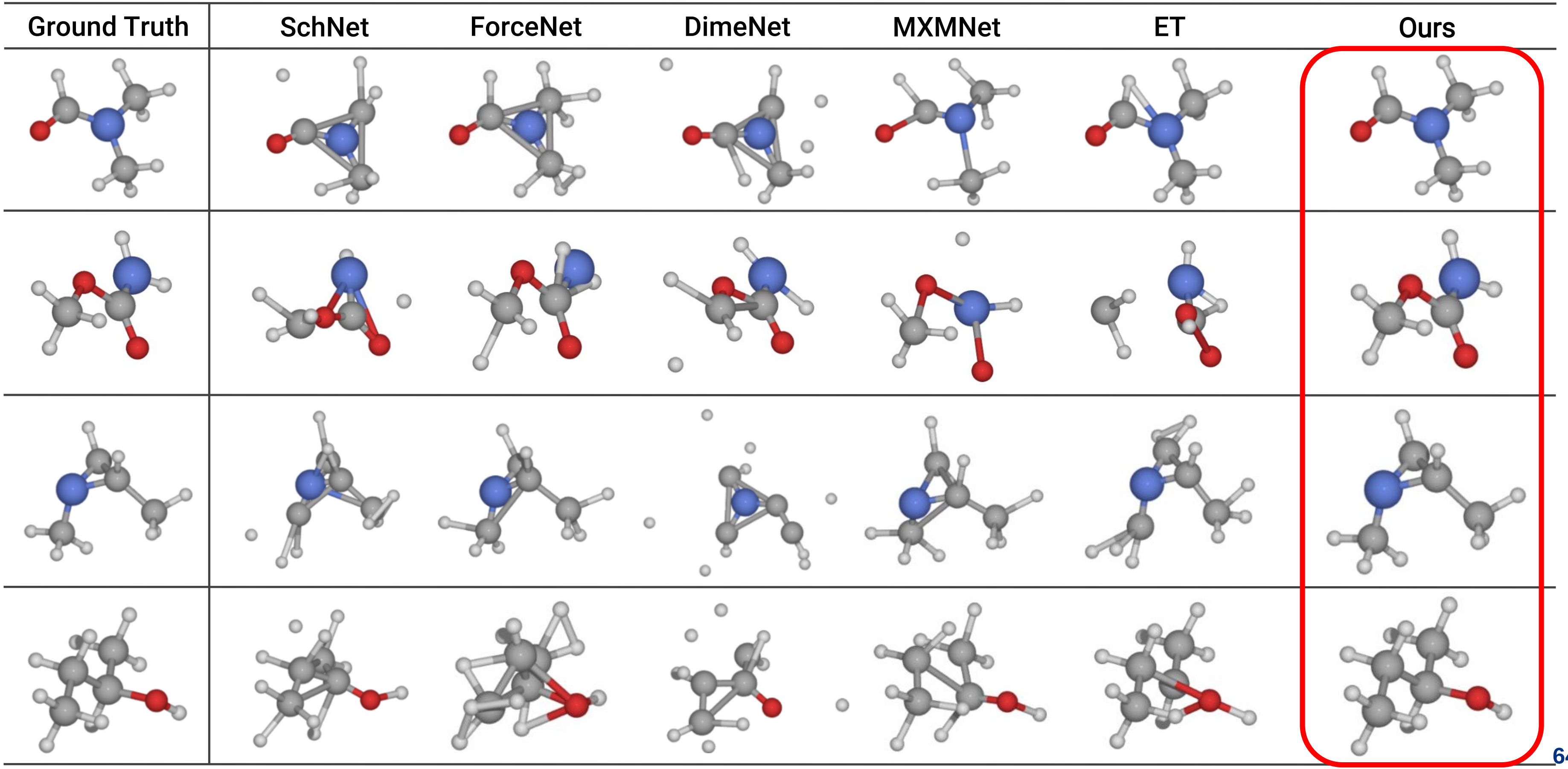


Conducted on QM9 dataset.

Qualitative Results: Structure Optimization

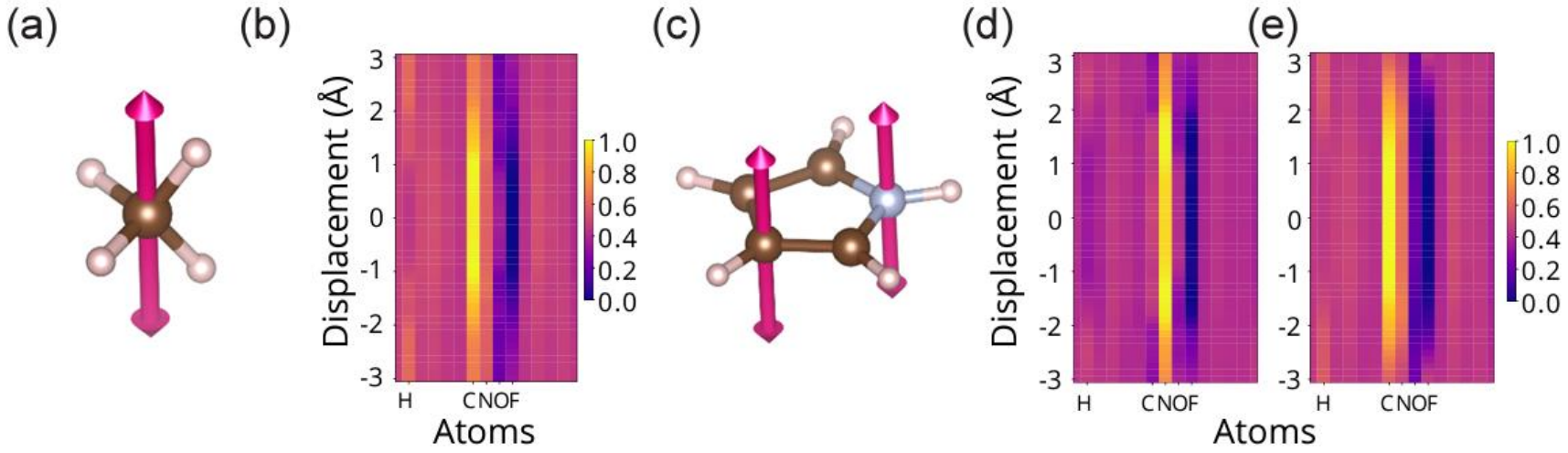


Qualitative Results: Structure Optimization



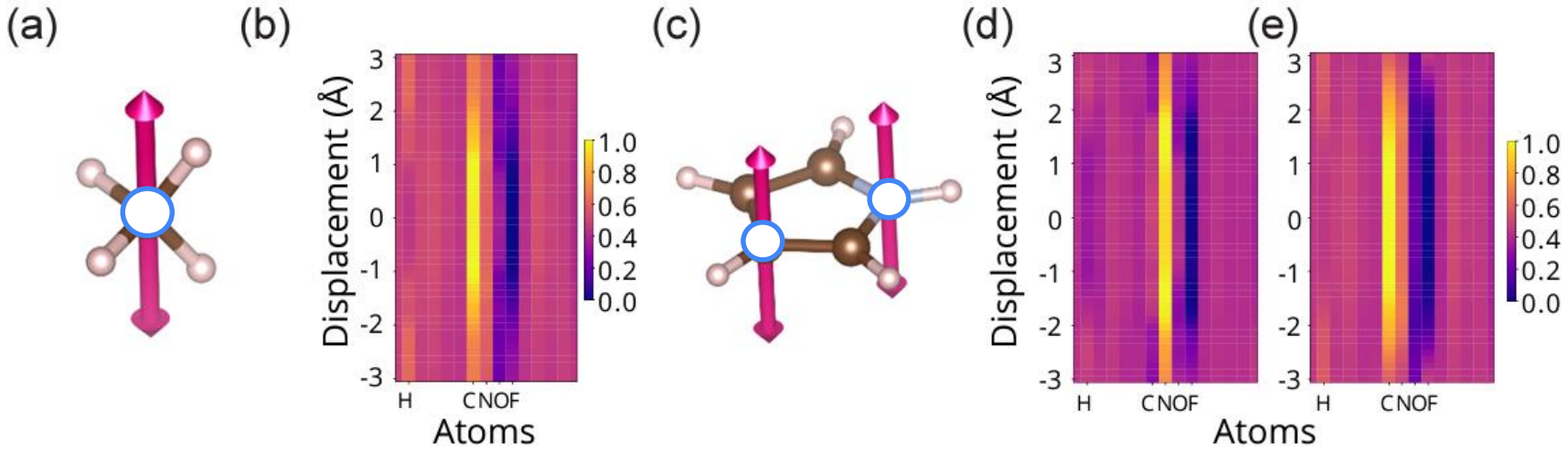
Ablation Study: Effect of MAM

Probability with dragging the atom of interest



Ablation Study: Effect of MAM

Probability with dragging the atom of interest



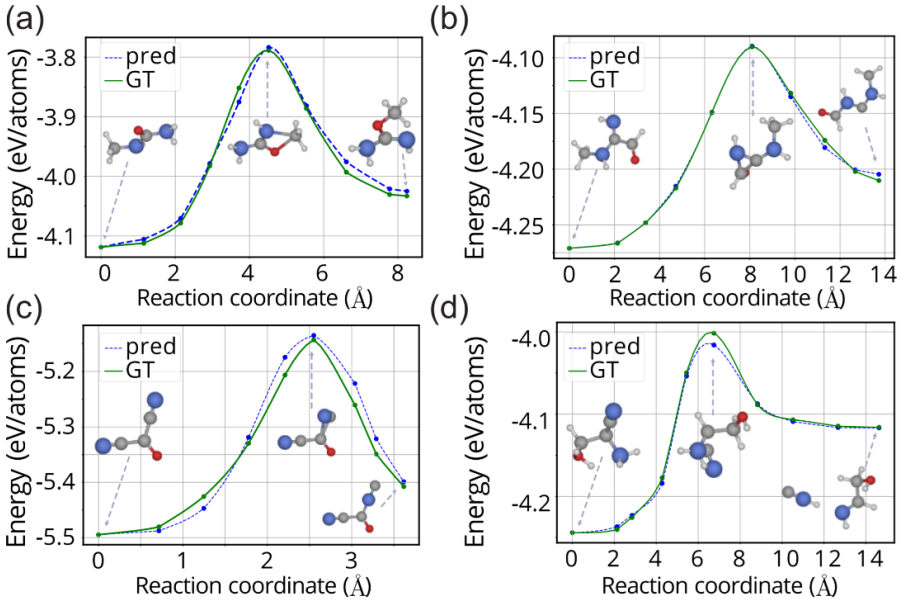
Results: Reaction Barrier Regression

Can our model generalize beyond stable molecules?

Results: Reaction Barrier Regression

Can our model generalize beyond stable molecules?

- Predict total energy along a reaction trajectory!

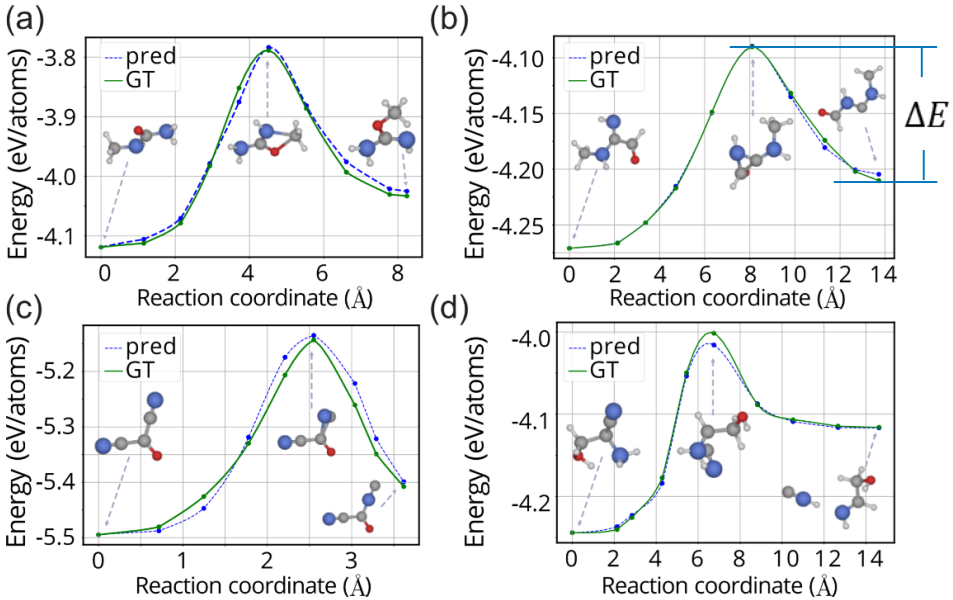


Conducted on Transition-1x dataset.

Results: Reaction Barrier Regression

Can our model generalize beyond stable molecules?

- Predict total energy along a reaction trajectory!



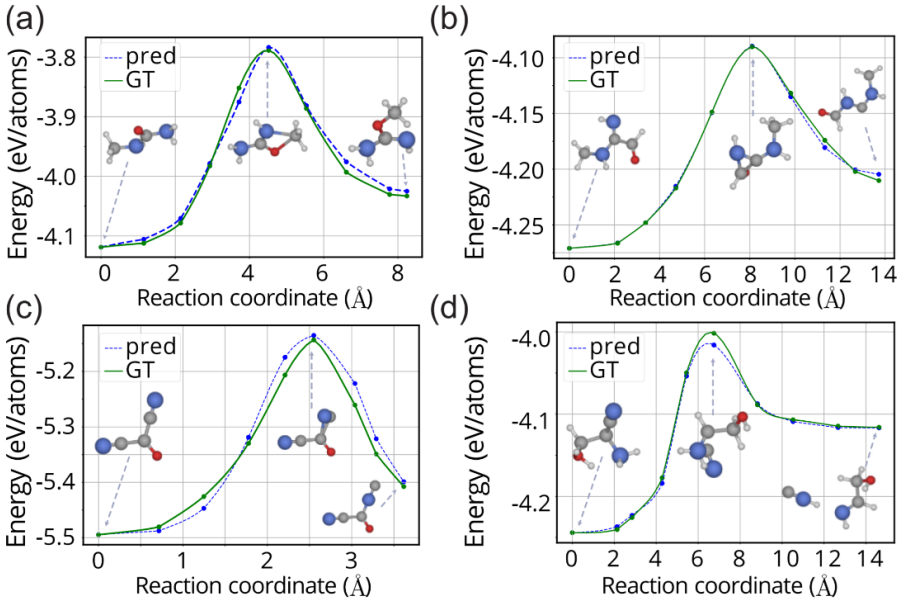
ΔE = Reaction barrier height

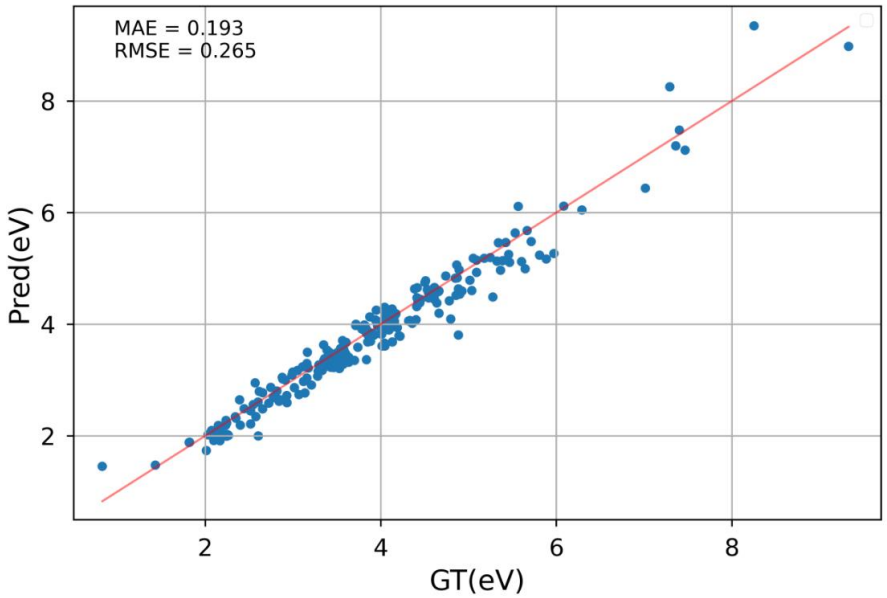
Conducted on Transition-1x dataset.

Results: Reaction Barrier Regression

Can our model generalize beyond stable molecules?

- Predict total energy along a reaction trajectory!



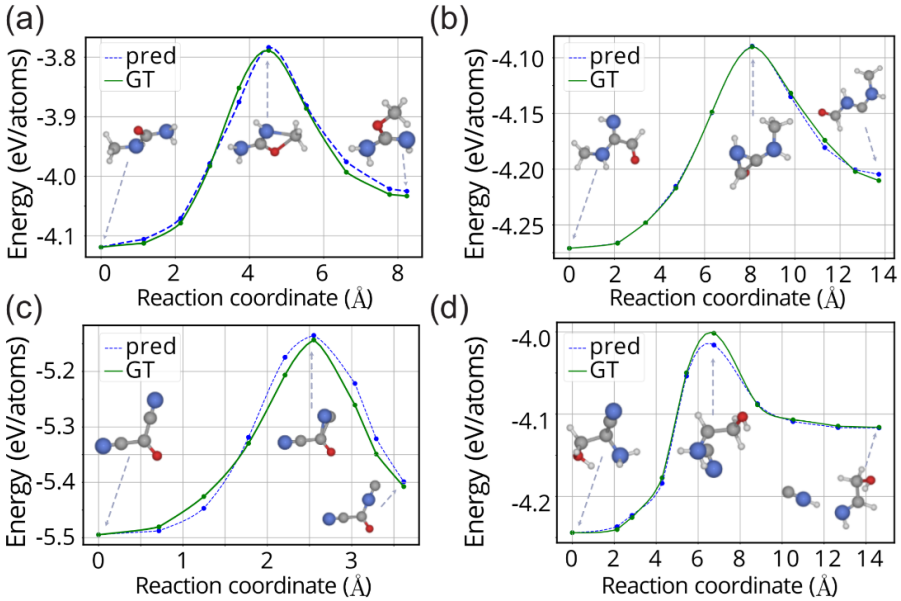


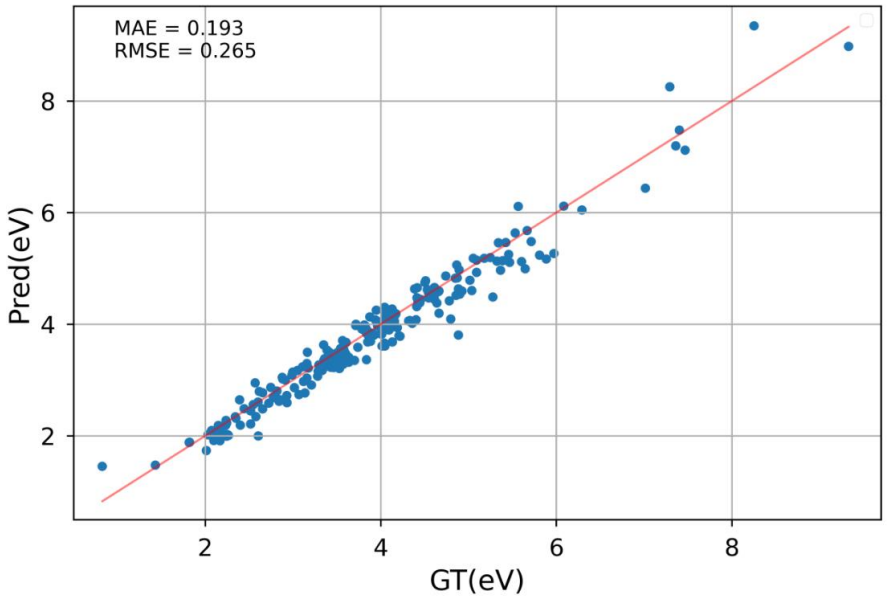
Conducted on Transition-1x dataset.

Results: Reaction Barrier Regression

Can our model generalize beyond stable molecules?

- Predict total energy along a reaction trajectory!

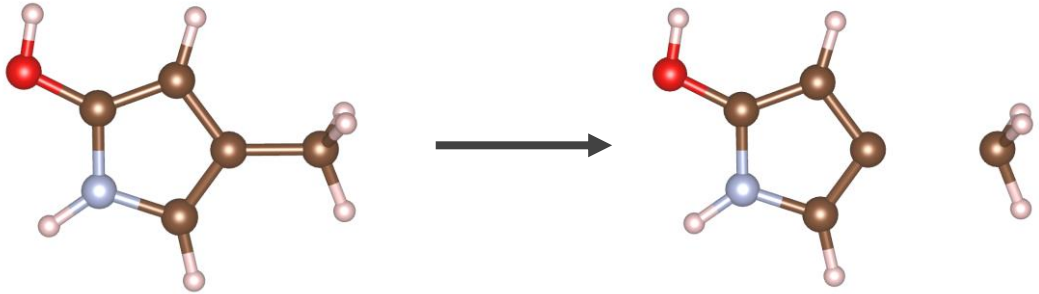




Conducted on Transition-1x dataset.

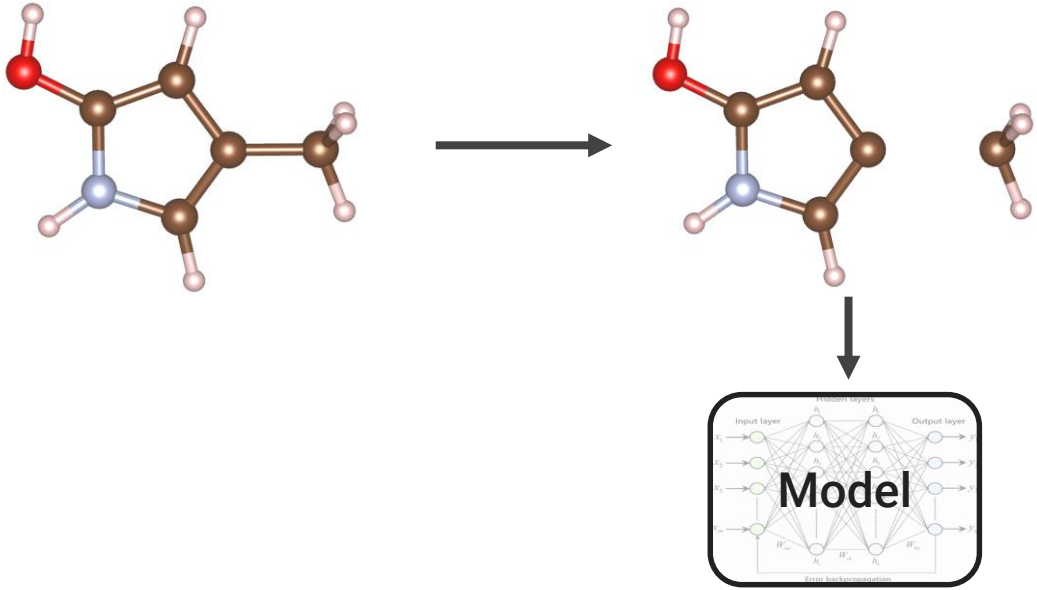
Results: Molecule Fragment Assembly

- A more challenging task to see generalization ability:
  - Split the molecule into functional groups, recover the initial state.



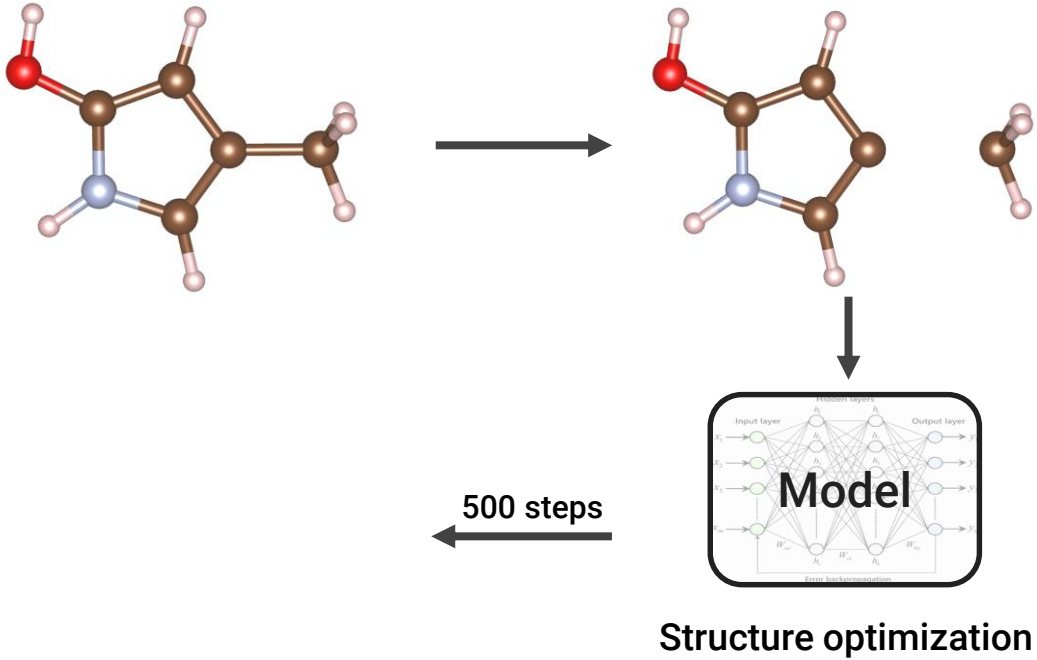
Results: Molecule Fragment Assembly

- A more challenging task to see generalization ability:
  - Split the molecule into functional groups, recover the initial state.



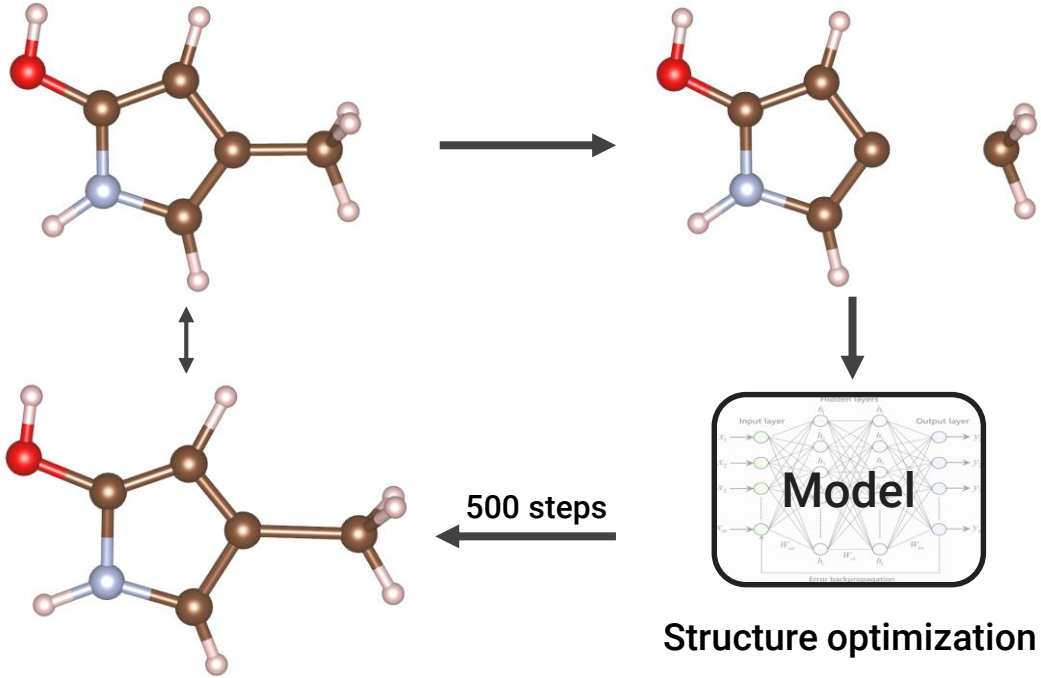
Results: Molecule Fragment Assembly

- A more challenging task to see generalization ability:
  - Split the molecule into functional groups, recover the initial state.



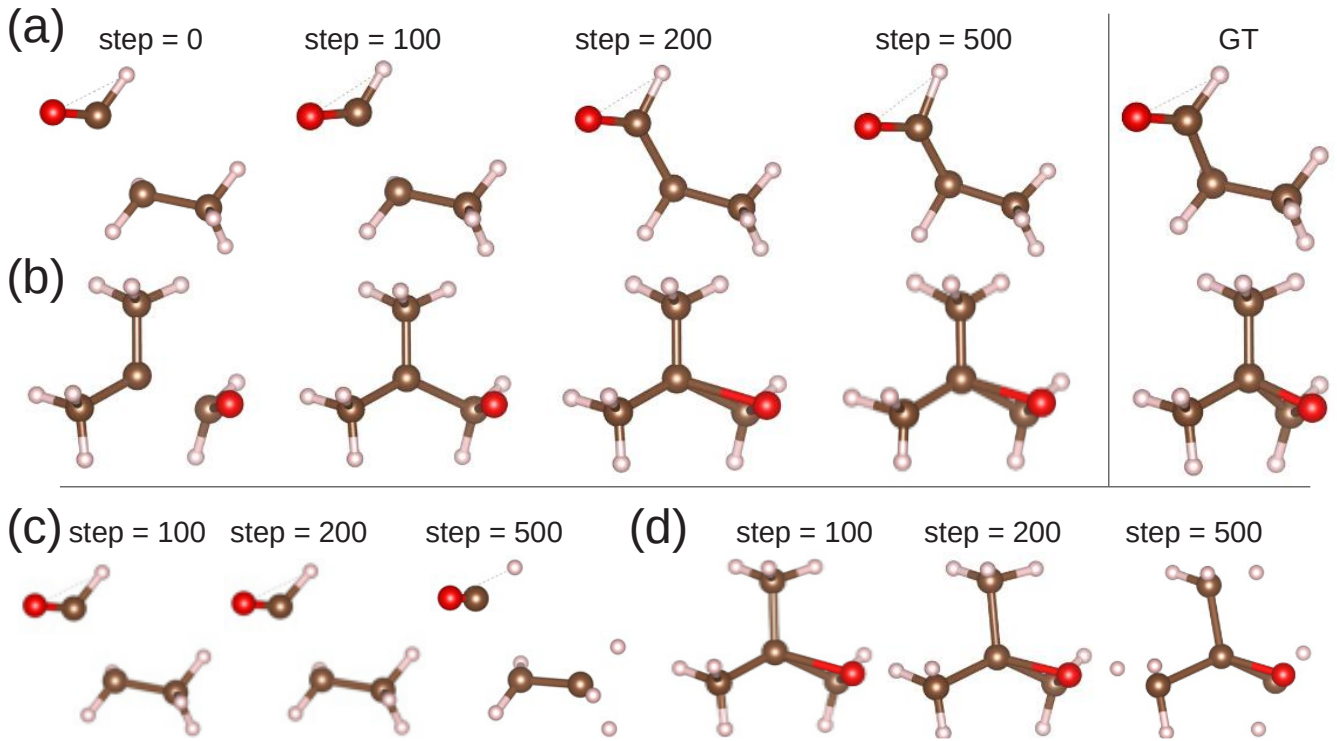
Results: Molecule Fragment Assembly

- A more challenging task to see generalization ability:
  - Split the molecule into functional groups, recover the initial state.



Results: Molecule Fragment Assembly

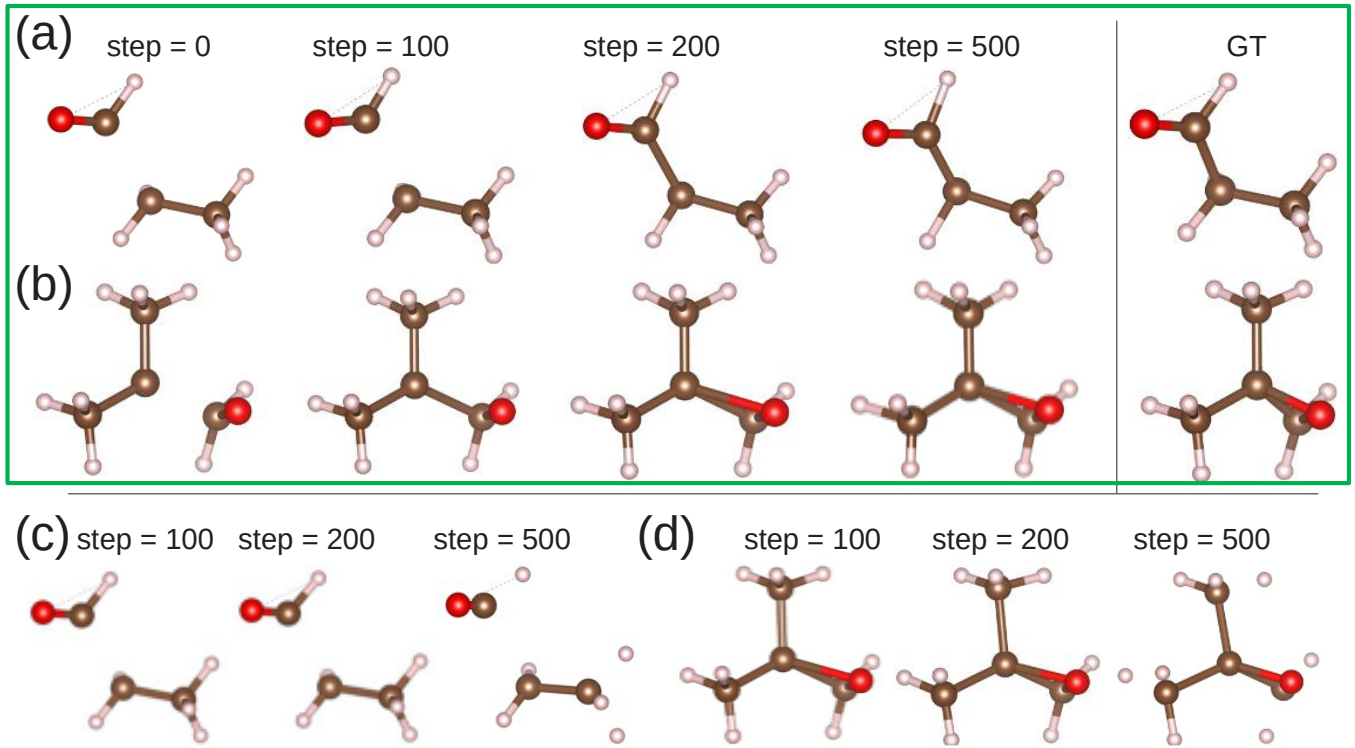
- A more challenging task to see generalization ability:



Results: Molecule Fragment Assembly

- A more challenging task to see generalization ability:
  - Only successful when using the **inequality bound condition**

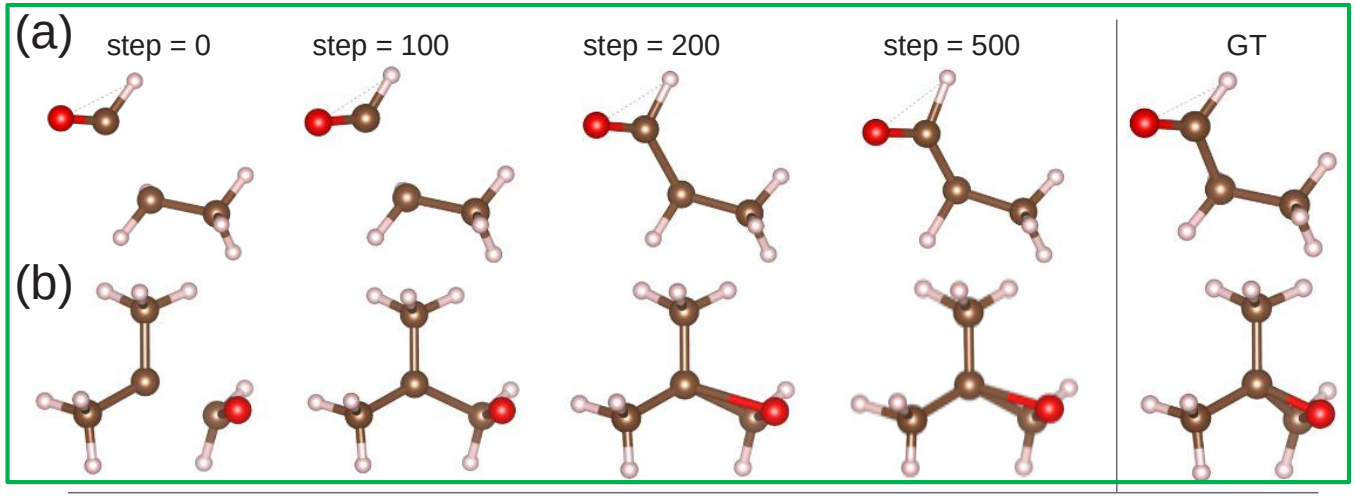
With inequality bound condition



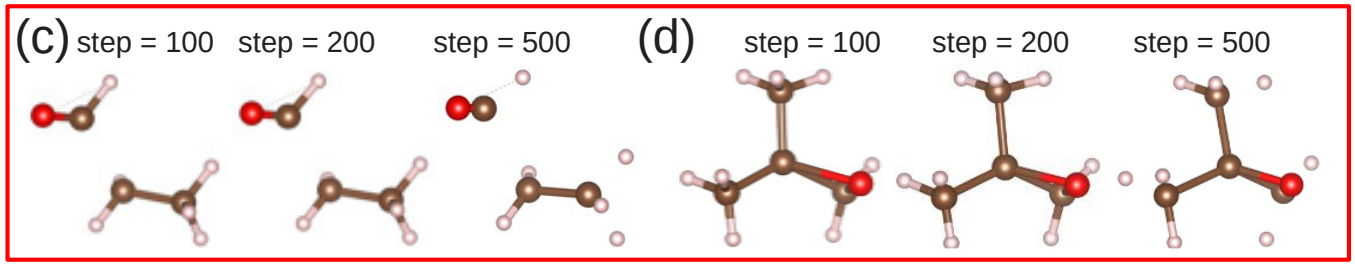
Results: Molecule Fragment Assembly

- A more challenging task to see generalization ability:
  - Only successful when using the **inequality bound condition**

With inequality bound condition



Without inequality bound condition



Conclusion

- We facilitate a more physically reliable model under limited data, by utilizing physical constraints and self-supervised learning method.

Conclusion

- We facilitate a more physically reliable model under limited data, by utilizing physical constraints and self-supervised learning method.
  - The inequality bound condition is crucial for the model to understand the physical properties near the ground state.
  - Zero-force condition is fruitful for optimized structure only datasets (ex: QM9).
  - Masked atomic modeling helps the model to understand basic bonding nature in a molecule.

Conclusion

- We facilitate a more physically reliable model under limited data, by utilizing physical constraints and self-supervised learning method.
  - The inequality bound condition is crucial for the model to understand the physical properties near the ground state.
  - Zero-force condition is fruitful for optimized structure only datasets (ex: QM9).
  - Masked atomic modeling helps the model to understand basic bonding nature in a molecule.

- Our model was able to generalize beyond stable structures, including reaction barrier prediction and molecule assembly task.

Any question?

Questions to: {jaguar6182, joonseok}@snu.ac.kr