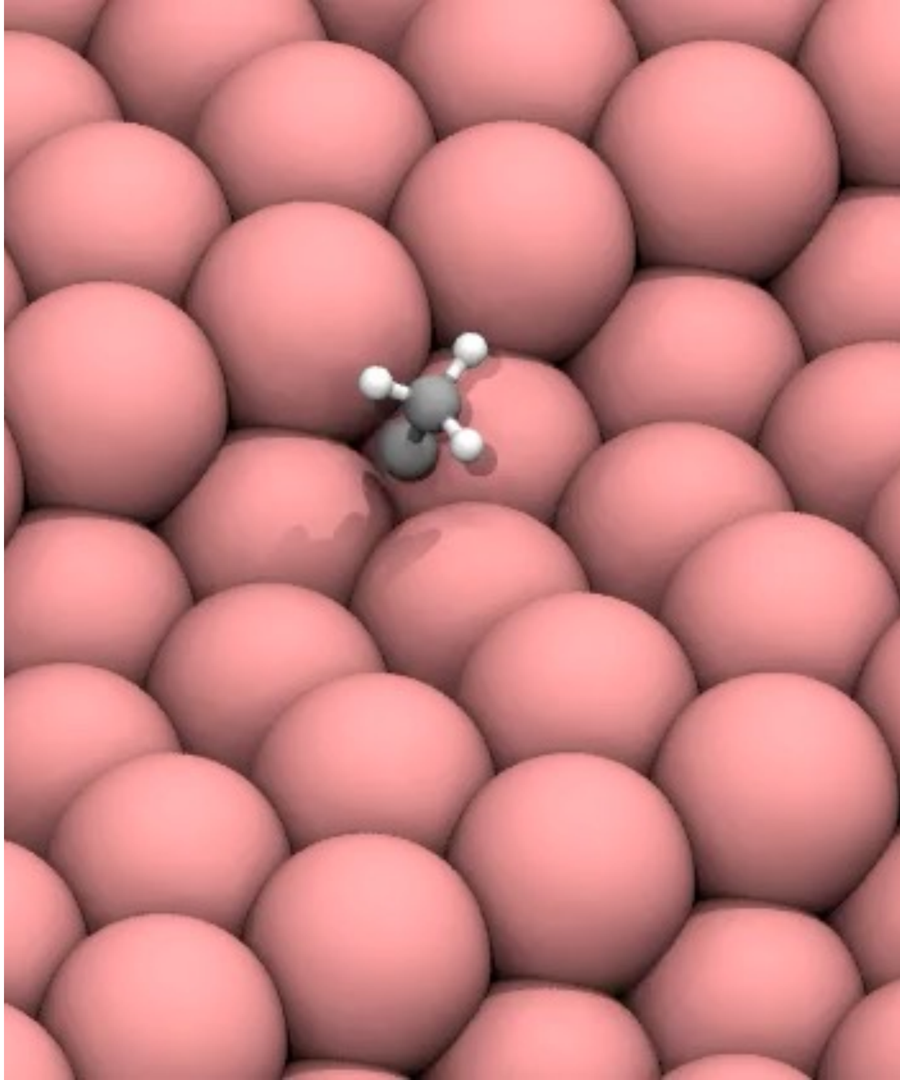
materials design

WEBINAR
Open Catalyst Project

April 26-28, 2022



Materials Design Webinar Series

- Each session runs several times to accommodate schedules
  - Share the webinar series with your colleagues!
  - Registration details http://www.materialsdesign.com/webinars
- We will be recording this webinar
  - Watch any of our earlier webinars anytime
  - We will post upcoming webinars on the webinar page
- Vote for the next webinar topic!
  - Take a 2 minutes brief survey at the end of the webinar!
- Audio issues
  - Log out and log back in again
  - Check your audio output
  - Google Chrome (most recent 2 versions) Mozilla Firefox (most recent 2 versions) Apple Safari (most recent 2 versions) Microsoft Edge (most recent 2 versions)

materials design

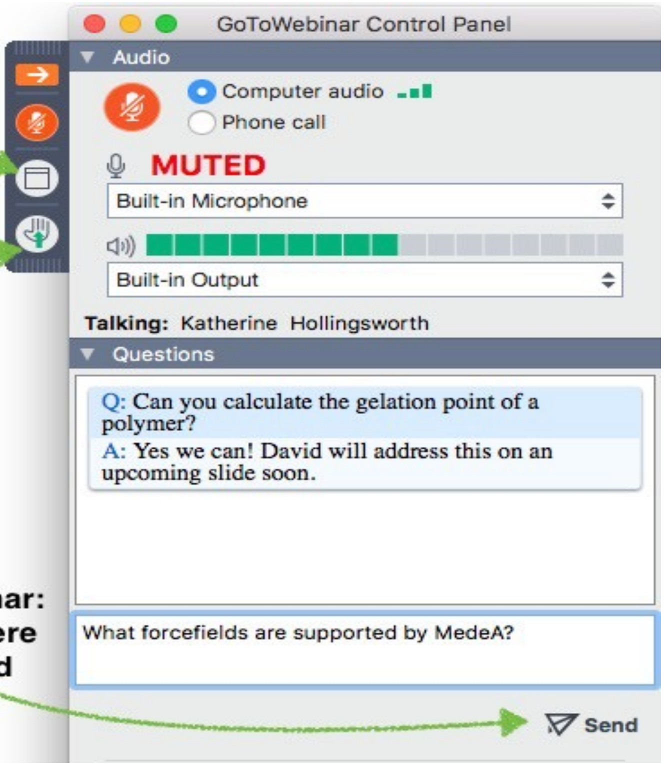
Please Ask Questions!

full screen

during discussion:
raise hand
to speak

Use the raise hand icon to bring attention to your question

any time during webinar:
type your question here
and then press Send



Webinar Speakers

Katherine Hollingsworth

Erich Wimmer

Facebook AI Research

Abhishek Das

Anuroop Sriram

Webinar Presenters

Open Catalyst Project

Abhishek Das

Research Scientist

Facebook AI Research

Anuroop Sriram

Research Engineer

Facebook AI Research

Team

FAIR


Abhishek Das


Brandon Wood


Siddharth Goyal


Anuroop Sriram


Janice Lan


Aditya Grover


Jure Zbontar


Nima Shoghi


Ammar Rizvi


Devi Parikh


Larry Zitnick

CMU Chemical Engineering


Zack Ulissi


Muhammed Shuaibi


Aini Palizhati


Javier Heras-Domingo


Brook Wander

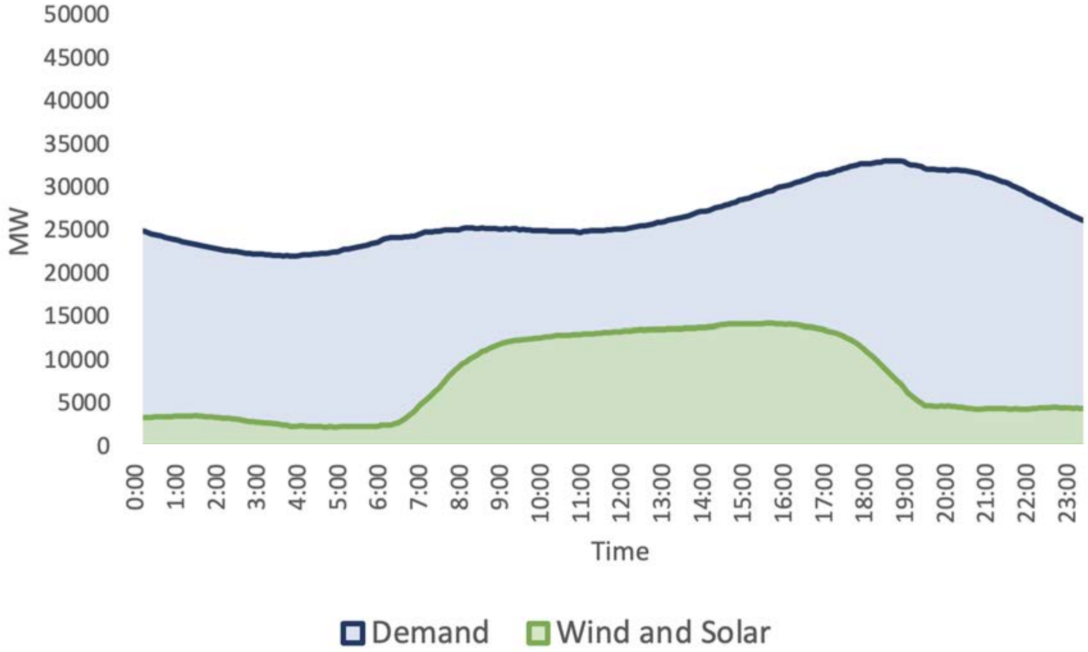

Adeesh Kolluru


Richard Tran

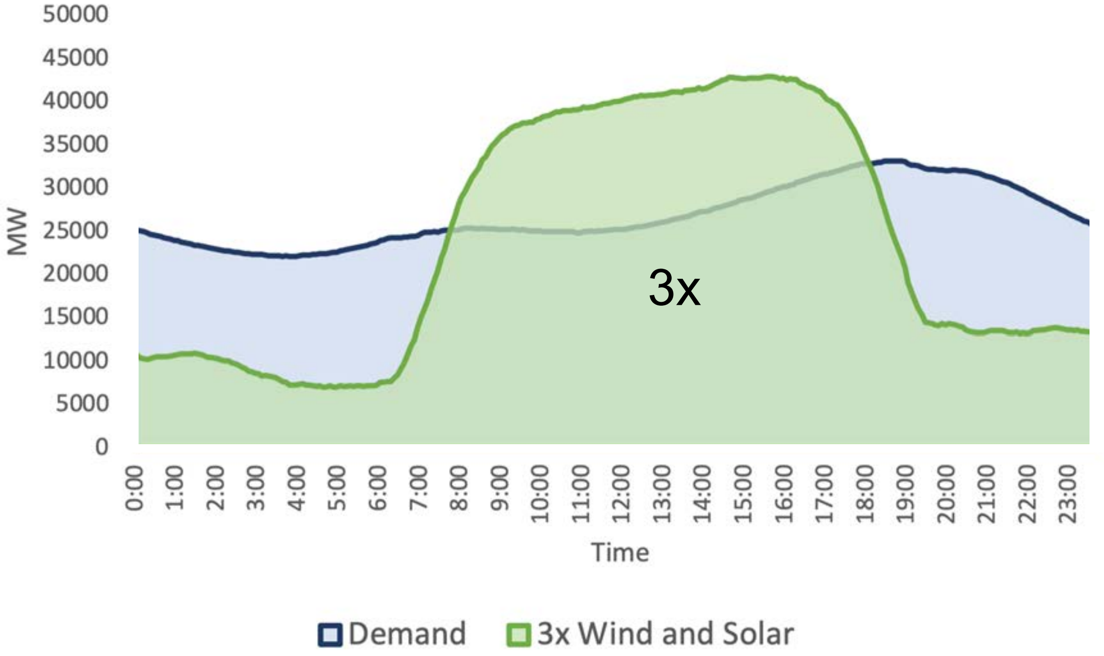
TU Munich


Johannes Klicpera

California Hourly Energy Demand



California Hourly Energy Demand

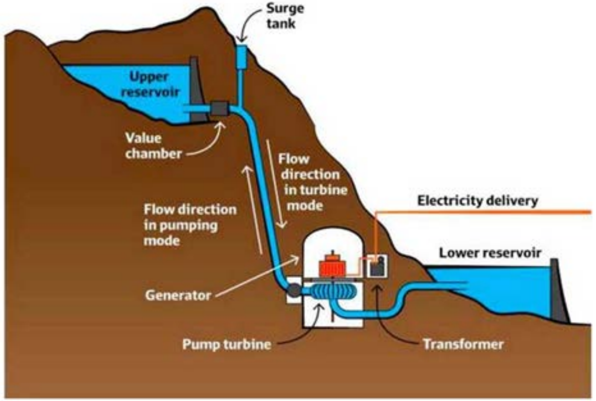


How do we store renewable energy?

Batteries

Tesla's 129 MWh installation ($50M)

Pump water uphill



70-80% efficient

2% of US generating capacity

Excess generation

Storage

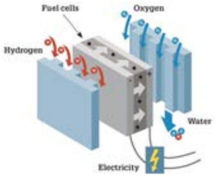
Excess demand

Electricity + $H_2O \rightarrow H_2 + O_2$

Hydrogen

H_2

$H_2 + O_2 \rightarrow$ Electricity + H_2O



Fuel Cells

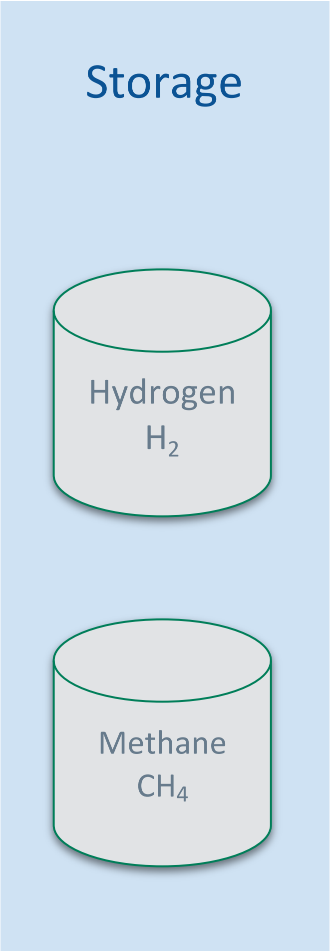
$H_2 + CO_2 \rightarrow CH_4 + H_2O$

Methane

CH_4

CH_4 + Fire $\rightarrow$ Electricity + $H_2O + CO_2$

Excess generation

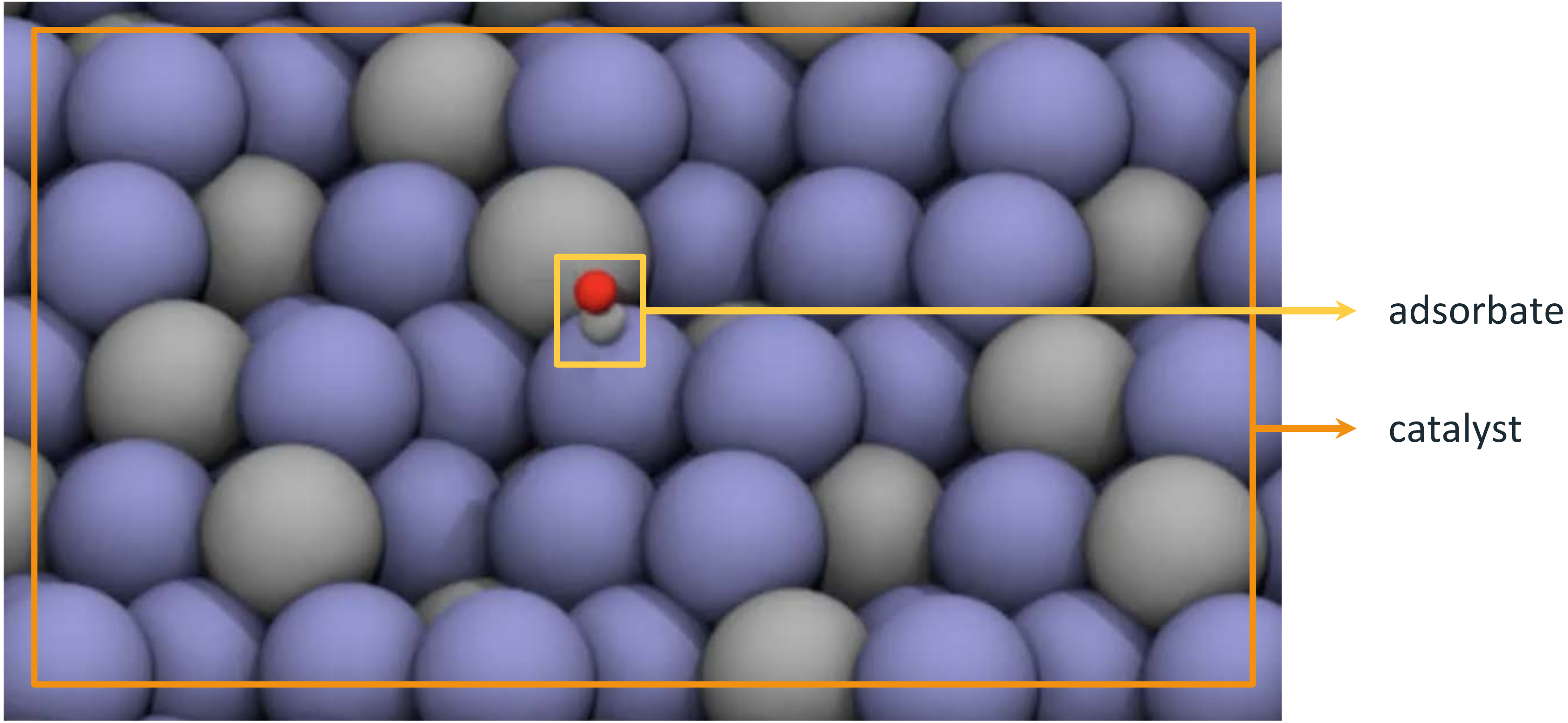


Electricity + $H_2O \rightarrow H_2 + O_2$

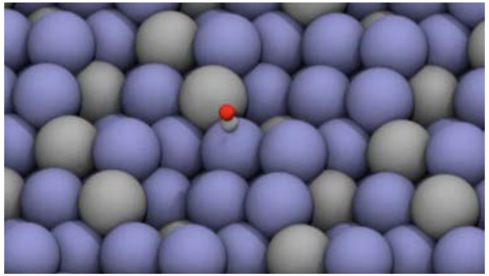
$H_2 + CO_2 \rightarrow CH_4 + H_2O$

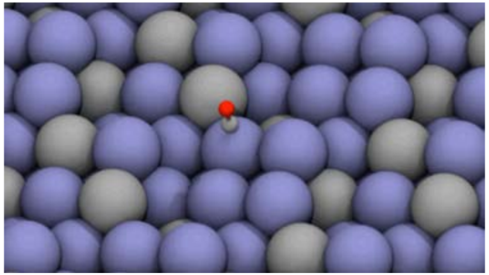
Goal: to find **catalysts** that can drive these chemical reactions at high rates without being consumed in the process.

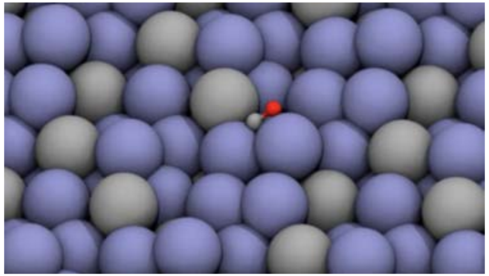
What is a relaxation?



What is a relaxation?







Initial state

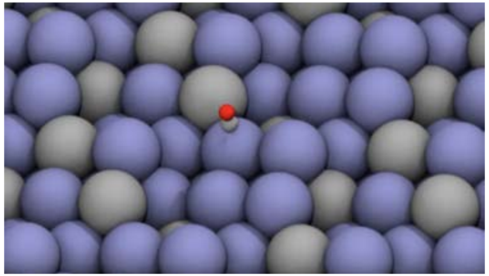
High energy

Adsorbate and catalyst atoms exert force on each other and move around

Relaxed state

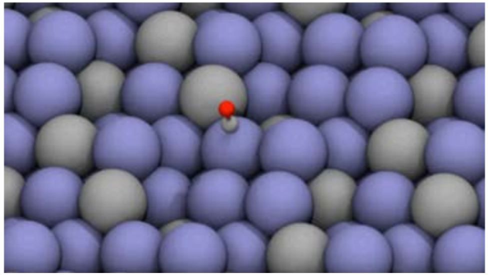
Low energy

What is a relaxation?

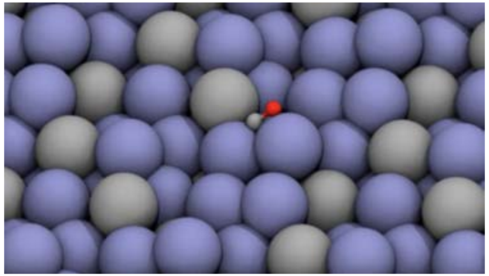


Initial state

High energy



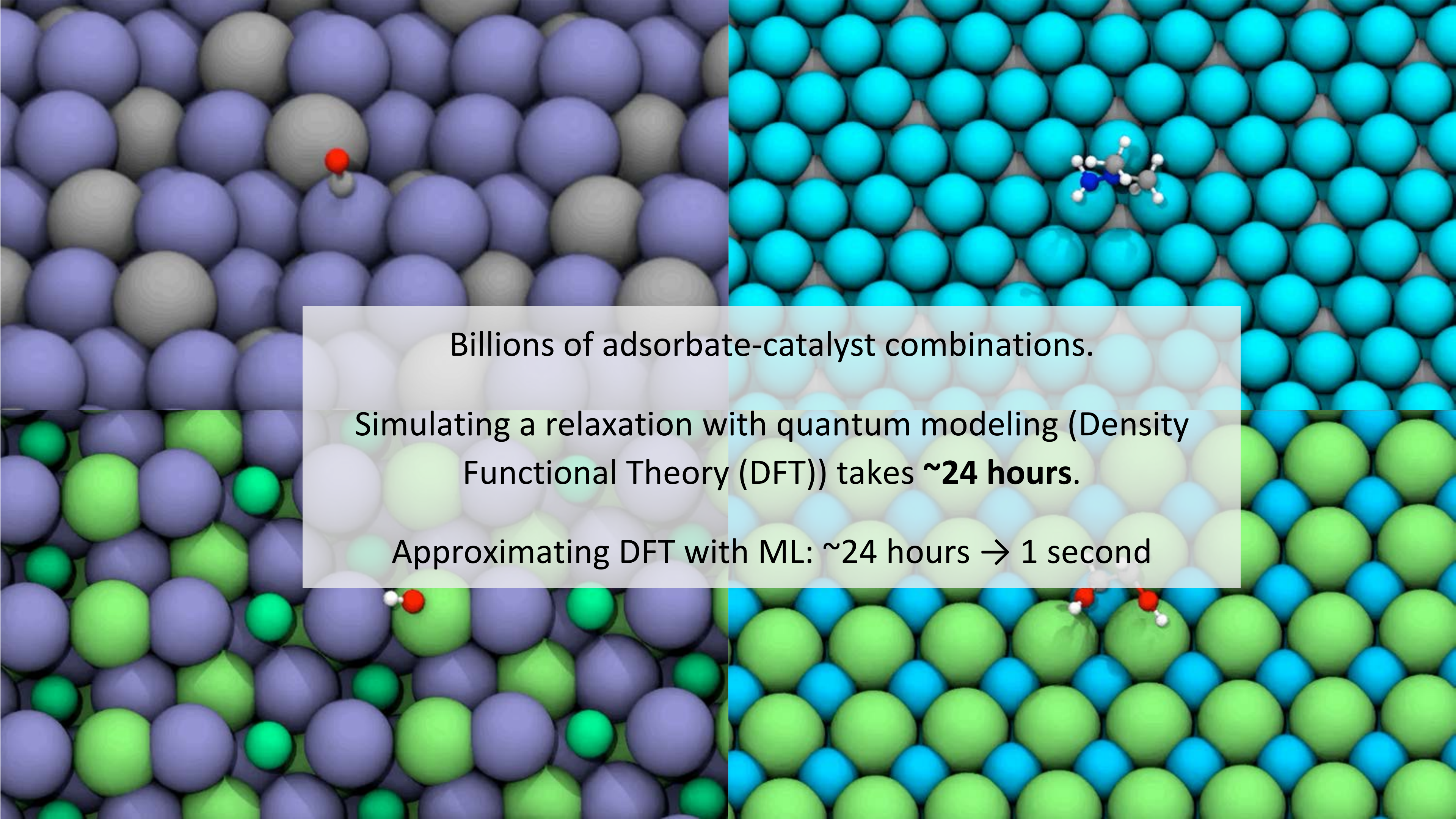
Adsorbate and catalyst atoms exert force on each other and move around



Relaxed state

Low energy

Can be used to determine reaction rate



Billions of adsorbate-catalyst combinations.

Simulating a relaxation with quantum modeling (Density Functional Theory (DFT)) takes **~24 hours**.

Approximating DFT with ML: ~24 hours → 1 second

The Open Catalyst 2020 (OC20) Dataset

- ~1.5M DFT relaxations for training and evaluation
- ~130M+ training examples
- ~200M CPU-hours to compute on 50k servers (3 months on FB Servers*)!

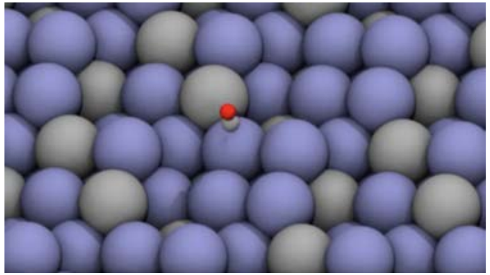
- Publicly available at opencatalystproject.org.

*Facebook servers run on renewable energy

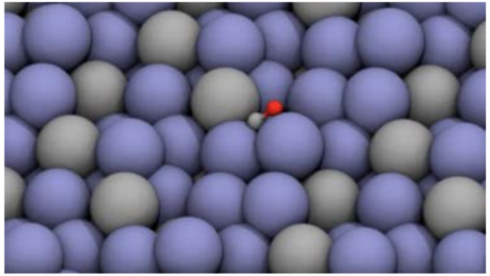
Initial Structure to Relaxed Energy (IS2RE)

Given an initial structure predict the energy of the relaxed structure

Initial

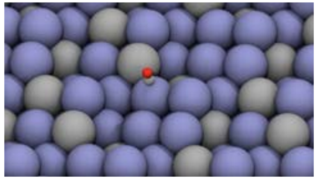


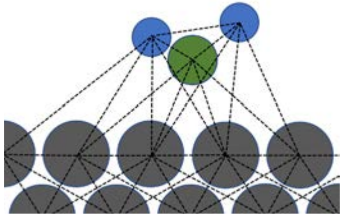
Relaxed



Constructing graphs

Initial Structure

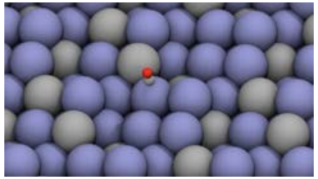


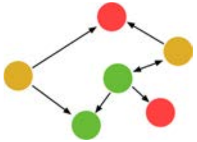


Node: atom
Edge: neighbors

Two ways to do IS2RE

Initial Structure

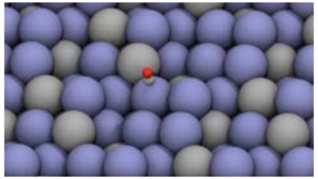


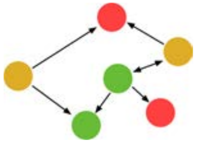


Relaxed energy

Two ways to do IS2RE

Initial Structure





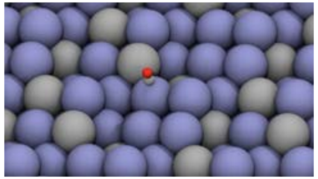
Relaxed energy

“Direct” approach

Can be learned from only IS2RE data

Two ways to do IS2RE

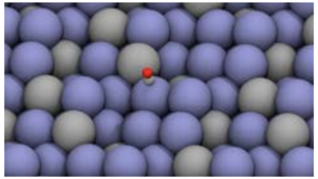
Initial Structure



Atomic forces

Two ways to do IS2RE

Initial Structure

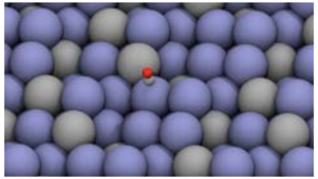


Atomic forces

Iteratively use predicted forces to update atomic positions, till max force per atom is ~0

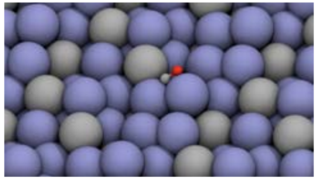
Two ways to do IS2RE

Initial Structure



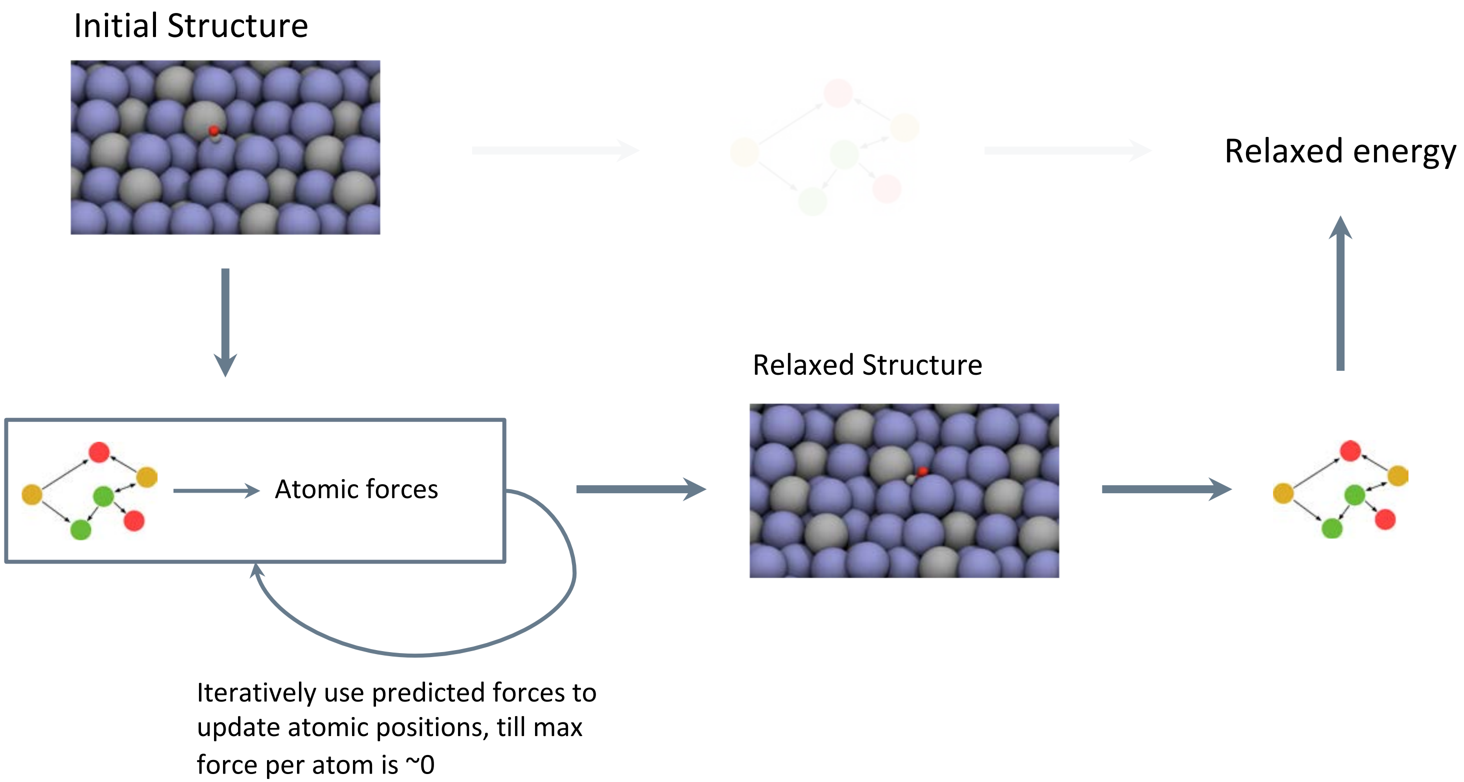
Atomic forces

Relaxed Structure



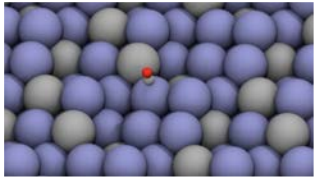
Iteratively use predicted forces to update atomic positions, till max force per atom is ~0

Two ways to do IS2RE



Two ways to do IS2RE

Initial Structure



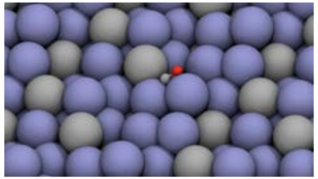
"Relaxation" approach

Typically requires intermediate trajectory data (S2EF)

Relaxed energy

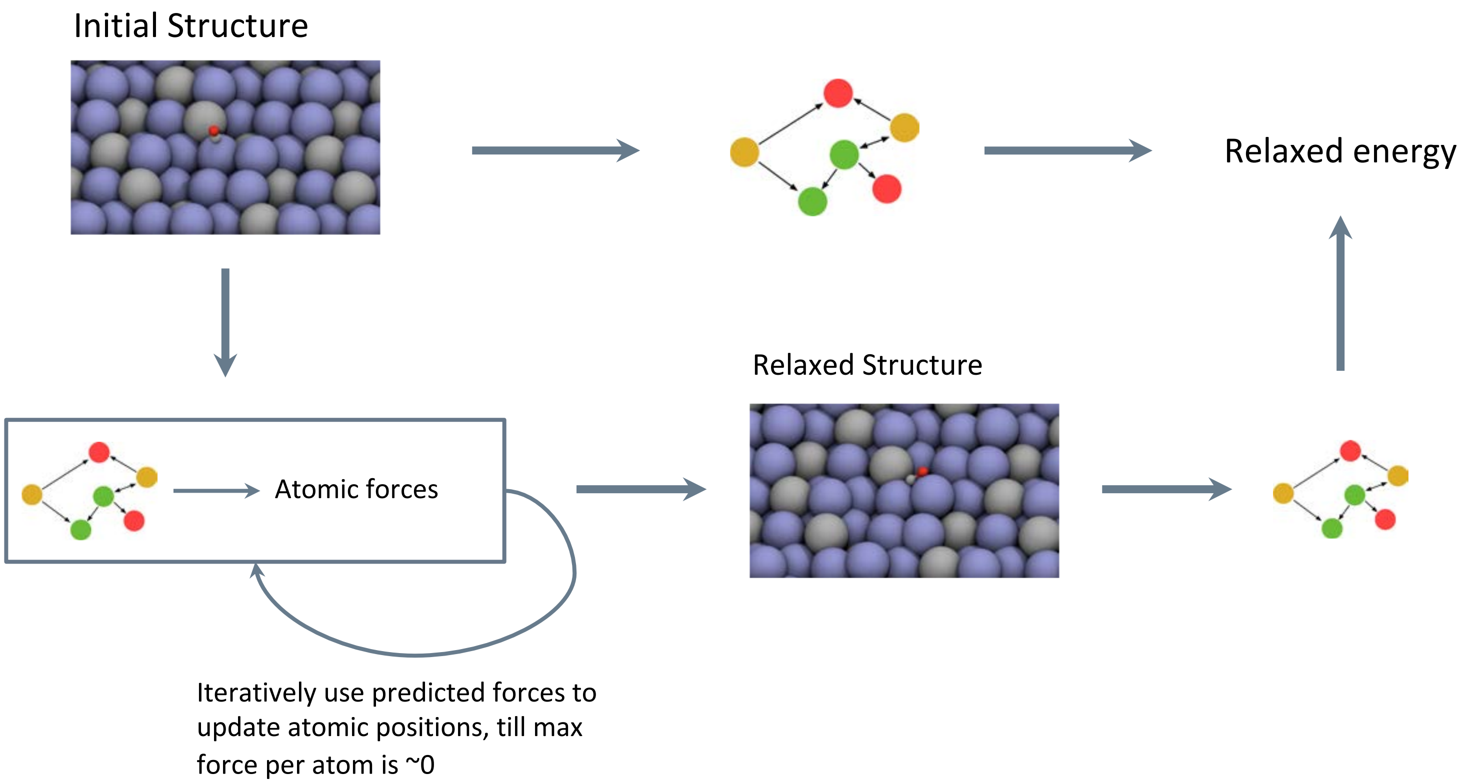
Atomic forces

Relaxed Structure



Iteratively use predicted forces to update atomic positions, till max force per atom is ~0

Two ways to do IS2RE



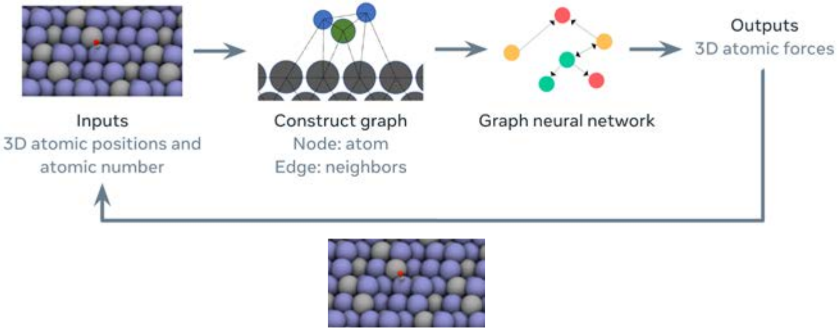
Energy vs force centric models

Force-centric Models

- The network predicts a force vector for each atom (and optionally the energy for the entire structure).

Energy-centric Models

- The network predicts the energy of the structure.
- Forces are computed using the negative gradients of the energy w.r.t atom positions.



vs.

Energy-centric

- Models physical laws by design
  - Ensures forces and energy are consistent
  - Ensures forces are energy conserving
- Slower inference/training

Force-centric

- Less physics built in
  - Forces are not consistent
  - Many times not rotation equivariant
- Faster inference/training

Models

Energy Centric

- CGCNN - GNN using a diverse set of features as input to the node embeddings.
- SchNet - Proposed using continuous edge filters with GNNs to compute per-atom forces through partial derivatives.
- DimeNet - Introduced the use of directional message passing to encode the angular information between triplets of atoms.

Force Centric

- ForceNet - More expressive messages conditioned on both the source and target nodes in GNN. Only computes forces, not energy.
- GemNet-OC - Encodes triplets and quadruplets. Current state-of-the-art

Xie and Grossman, *Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties*, 2018

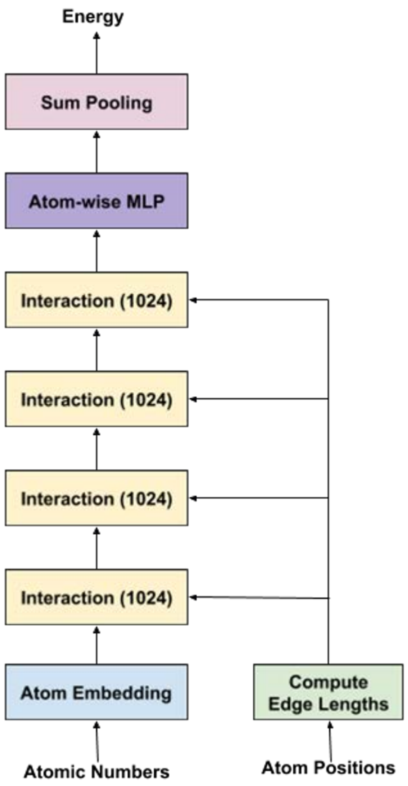
Schutt, et al., Schnet: *A continuous-filter convolutional neural network for modeling quantum interactions*, 2017

Klicpera, Groß, and Gunnemann, *Directional Message Passing for Molecular Graphs*, 2020

Hu et al., *ForceNet: A Graph Neural Network for Large-Scale Quantum Chemistry Simulation*, 2020 (under submission)

SchNet: a deeper dive

- Graph Neural Network mapping atomic graphs to energy
  - Inputs: Nodes (atomic numbers) and Edges (atom positions)
  - Multiple interaction blocks that update node embeddings iteratively using continuous filter convolutions
  - Forces are computed as gradient of predicted energy with respect to atom positions

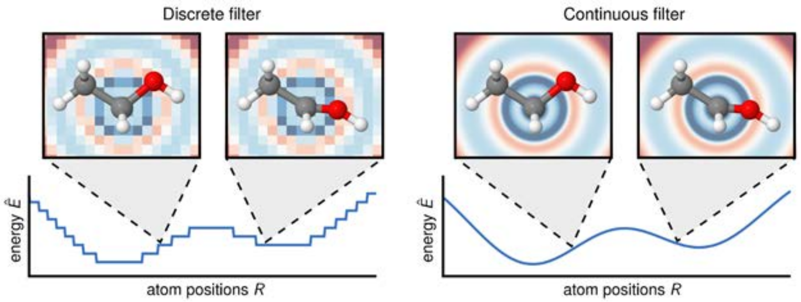


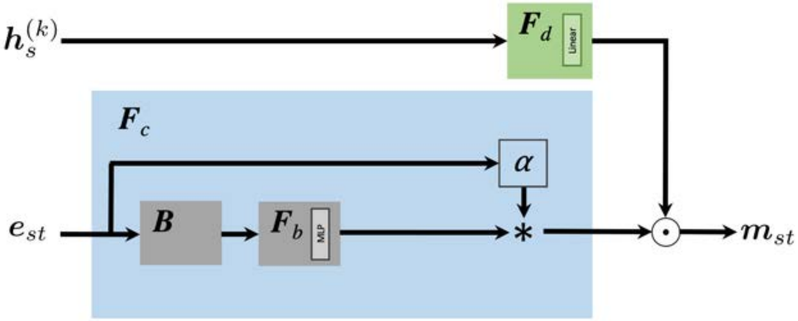
Schutt, et al., Schnet: *A continuous-filter convolutional neural network for modeling quantum interactions*, 2017

SchNet: a deeper dive

- Continuous filter convolutions
  - Graph convolutions without discretizing the atom positions to work with unevenly spaced atoms
  - Allows computing gradient with respect to atom positions
  - Represent edge lengths (atomic distances) using radial basis functions

$$\mathbf{x}_i^{l+1} = (X^l * W^l)_i = \sum_j \mathbf{x}_j^l \circ W^l(\mathbf{r}_i - \mathbf{r}_j),$$

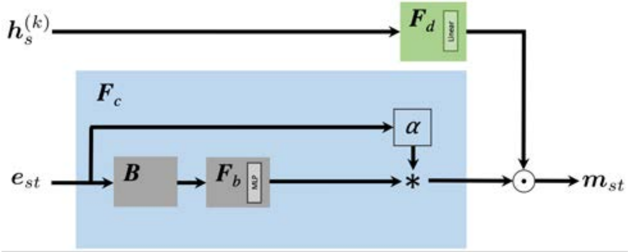




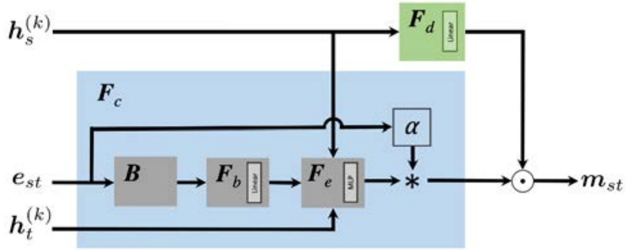
ForceNet

- Force-centric model based on the SchNet architecture
    - Computes more expressive messages that are also conditioned on the source and target node representations
    - More careful choice of activation functions and basis functions
    - Easier to scale up to larger model sizes

SchNet



ForceNet



Hu et al., *ForceNet: A Graph Neural Network for Large-Scale Quantum Chemistry Simulation*, 2020

DimeNet / DimeNet++

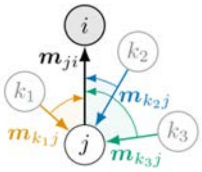
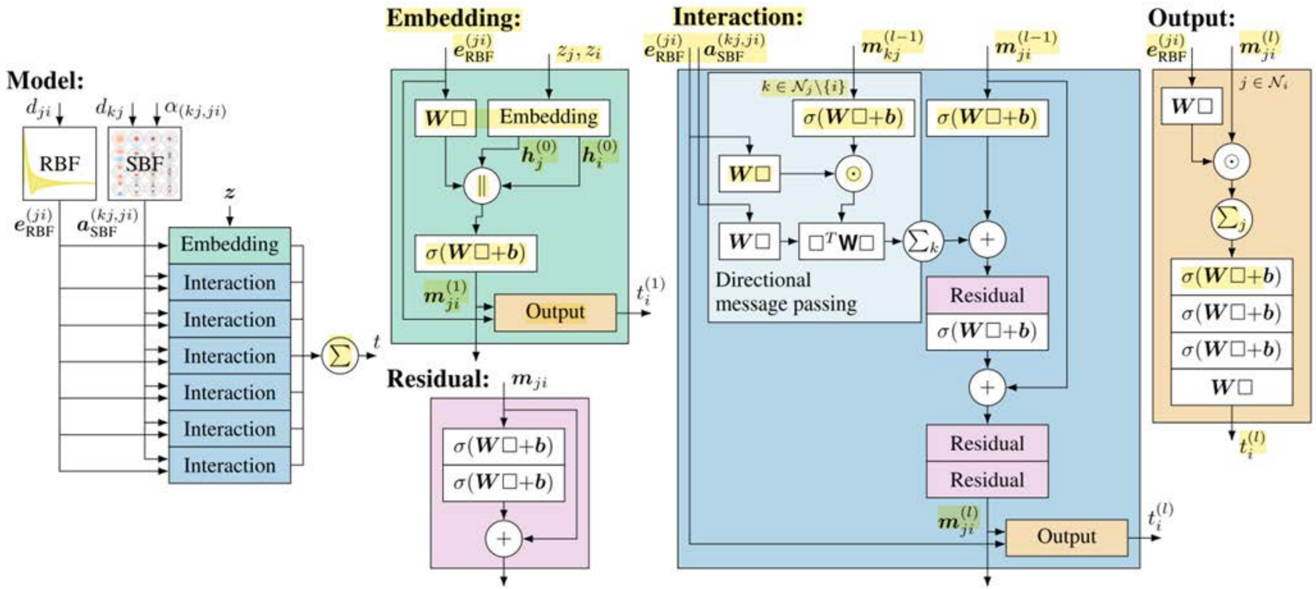


Figure 1: Aggregation scheme for message embeddings.



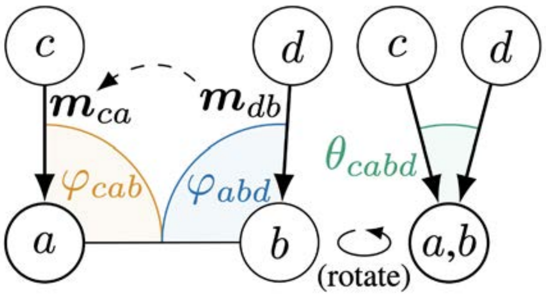
Klicpera et al. 2020

Scaling up model training

Sriram et al., Towards Training Billion Parameter Graph Neural Networks for Atomic Simulations, ICLR 2022

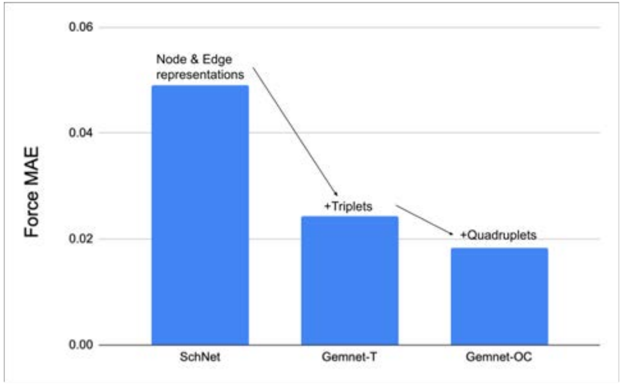
OC20: Modeling Higher Order Interactions

- Simple message passing neural networks do not capture the full 3D geometry of atomic systems
  - Angles can be modeled explicitly using triplets of atoms
  - Dihedral angles can be modeled explicitly using quadruplets of atoms
  - Klicpera et al., 2021 showed that GNNs modeling triplets and quadruplets are universal approximators for predictions that are invariant to translation, and equivariant to permutation and rotation



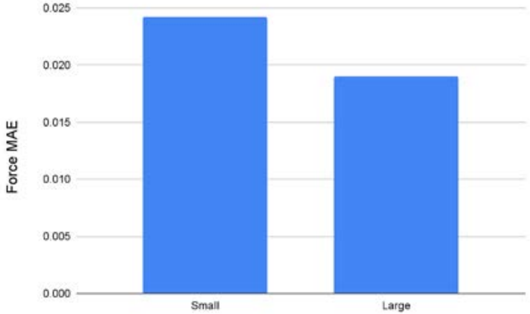
OC20: Modeling Higher Order Interactions

- Modeling higher order interactions improves performance
  - All the top models on OC20 rely on such representations



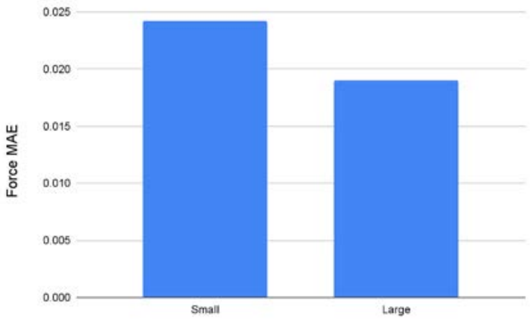
OC20: Model Scaling

- Scaling to large model sizes is critical for good performance
  - Scale is helpful for handling the diversity of OC20



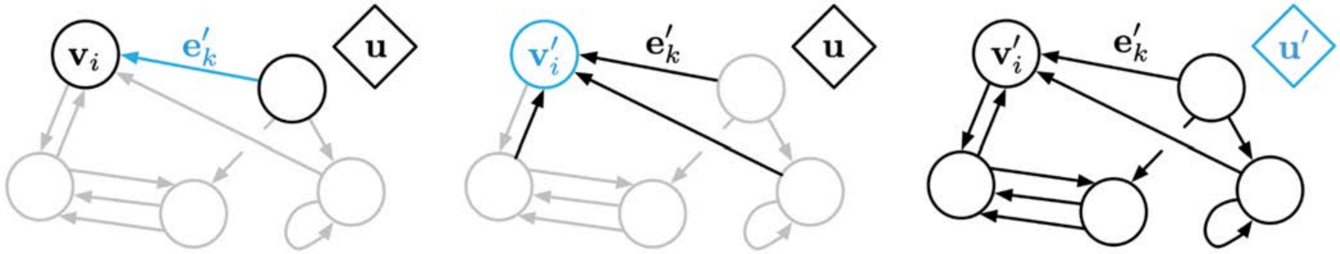
OC20: Model Scaling

- Scaling to large model sizes is critical for good performance
  - Scale is helpful for handling the diversity of OC20
- Higher order interactions increase GPU memory usage significantly, making it challenging to scale up
  - Typical systems have O(100) atoms, O(1000) edges and O(1,000,000) triplets / quadruplets
  - We need a way to parallelize computation across multiple GPUs



Graph Nets (GN)

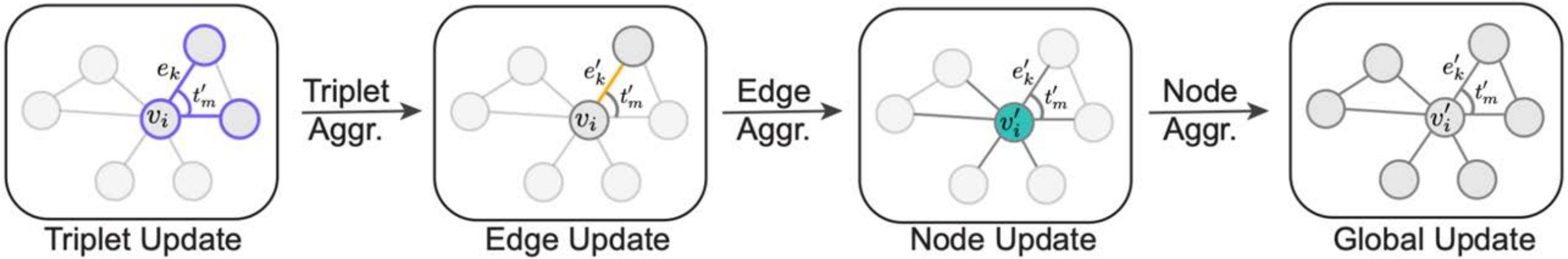
- Battaglia et al, 2018 introduced the Graph Nets framework
  - Provides a general abstraction for many popular GNNs operating on nodes and edges
- Graph Nets consist of a series of GN blocks that iteratively update the edge, node and global representations in that order



(a) Edge update (b) Node update (c) Global update

Extended Graph Nets (EGN)

- A general framework for GNNs that model higher order interactions such as triplets
  - Generalizes the Graph Nets framework from Battaglia et al., 2018
  - Includes multiple message passing steps in each layer
  - Many popular models like GemNet, DimeNet++, GCNs are special cases of EGNs



Dimenet++ as an EGN

- DimeNet++ (Klicpera et al., 2020) uses directional message passing
  - Edge *(i,j)* is represented by a feature representation m_{ji} and node i by h_i
  - Updated using both directional information (via bond angles) & interatomic distances
  - Uses explicit triplet representation for bond angles

$$\mathbf{m}'_{ji} = \mathrm{f}_{\mathrm{update}}(\mathbf{m}_{ji}, \sum_{k \in \mathcal{N}_j \setminus \{i\}} \mathrm{f}_{\mathrm{int}}(\mathbf{m}_{kj}, \mathbf{e}_{RBF}^{(ji)}, \mathbf{a}_{SBF}^{(kj,ji)}))$$

$$\mathbf{h}_i = \textstyle\sum_j \mathbf{m}'_{ji}$$

Dimenet++ as an EGN

- DimeNet++ (Klicpera et al., 2020) uses directional message passing
  - Edge *(i,j)* is represented by a feature representation m_{ji} and node *i* by h_i
  - Updated using both directional information (via bond angles) & interatomic distances
  - Uses explicit triplet representation for bond angles

$$\mathbf{m}'_{ji} = \underbrace{\mathrm{f}_{\mathrm{update}}}_{\text{Edge Update}}\Big(\mathbf{m}_{ji}, \underbrace{\sum_{k\in\mathcal{N}_j\setminus\{i\}}}_{\text{Triplet Aggregation}} \underbrace{\mathrm{f}_{\mathrm{int}}}_{\text{Triplet Update}}\big(\mathbf{m}_{kj}, \mathbf{e}_{RBF}^{(ji)}, \mathbf{a}_{SBF}^{(kj,ji)}\big)\Big)$$

$$\mathbf{h}_i = \underbrace{\textstyle\sum_j}_{\substack{\text{Edge Aggregation}\\ \text{+ Node Update}}} \mathbf{m}'_{ji}$$

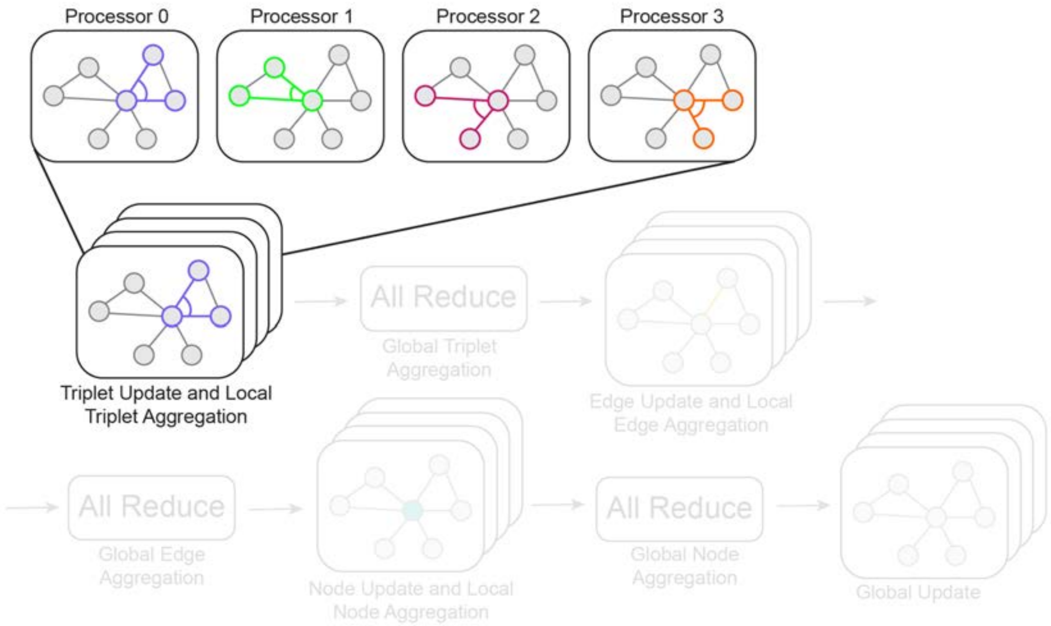
Gemnet-T as an EGN

- Gemnet-T (Klicpera et al., 2021) extends Dimenet++ in a number of ways
  - Uses a more efficient edge update based on bilinear layer instead of the hadamard product
  - Maintains explicit node representations
  - Updated node embeddings are used to update the edge embeddings again (i.e. has an additional edge-update operation)

- Therefore, Gemnet-T can also be represented as an EGN, with an extra edge-update

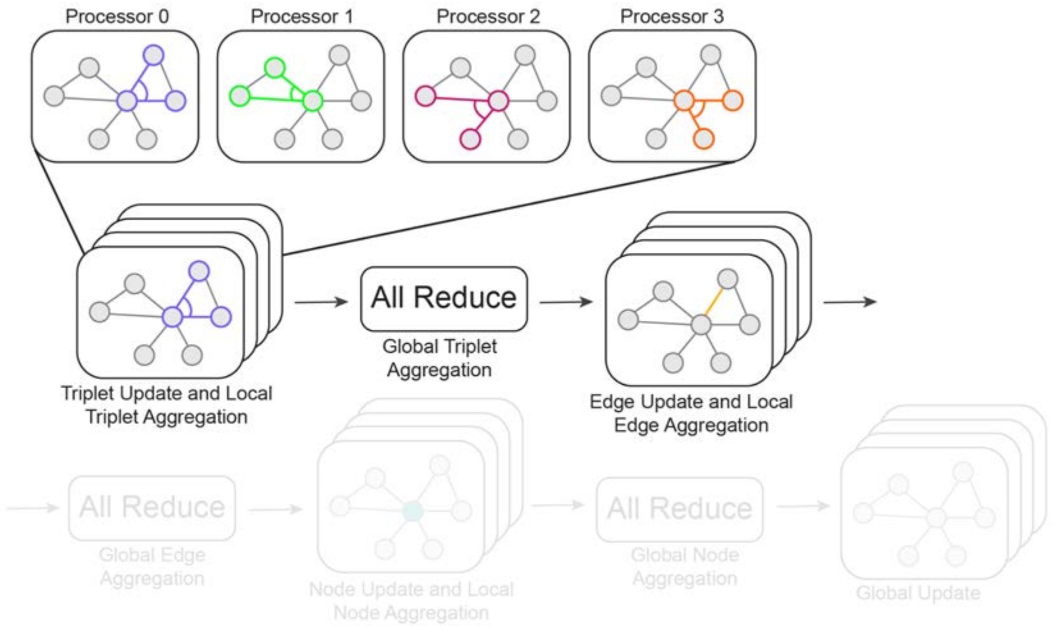
Parallelizing EGNs

- Distribute the triplet, edge and node representations across multiple GPUs
- Update triplet representations and perform local aggregation on each GPU in parallel
  - Ensures that triplets are never communicated across GPUs



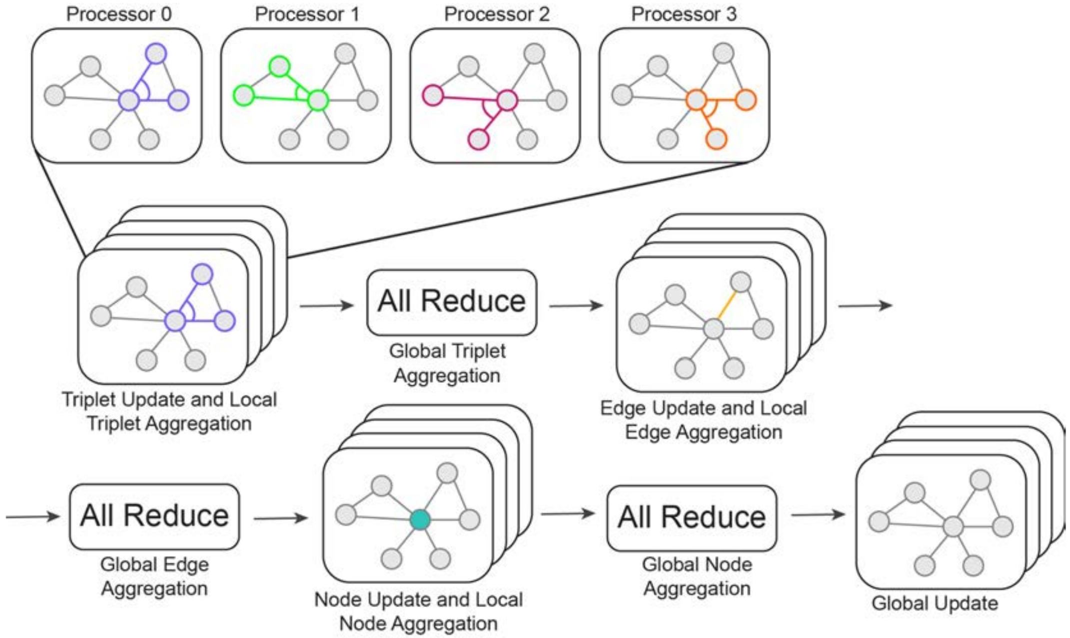
Parallelizing EGNs

- Distribute the triplet, edge and node representations across multiple GPUs
- Update triplet representations and perform local aggregation on each GPU in parallel
    - Ensures that triplets are never communicated across GPUs
- All-reduce the reduced representations to perform a global reduction



Parallelizing EGNs

- Distribute the triplet, edge and node representations across multiple GPUs
- Update triplet representations and perform local aggregation on each GPU in parallel
    - Ensures that triplets are never communicated across GPUs
- All-reduce the reduced representations to perform a global reduction
- Repeat to update edge and node representations



Experiments - Models

- DimeNet++-XL:
  - Graph Parallel implementation of Dimenet++ with 240M parameters trained on 256 Nvidia V100 GPUs.
  - >20x larger than the Dimenet++-Large model reported in Chanussot et al., 2020

- GemNet-XL:
  - Graph Parallel implementation of Gemnet-T with 300M parameters trained on 256 Nvidia V100 GPUs.
  - ~10x larger than Gemnet-T model from Klicpera et al., 2021

Experiments - tasks

- Structure to Energy and Forces (S2EF)
  - Given an atomic system, predict the energy of the system and 3D forces on each atom
- Initial Structure to Relaxed Structure (IS2RS)
  - Given an atomic system, predict the relaxed positions of each atom
  - Involves performing a structure relaxation (iteratively updating positions to get to a local minima in energy.
- Initial Structure to Relaxed Energy (IS2RE)
  - Given an atomic system, predict the relaxed energy of the system
  - Can be predicted directly or by performing a structure relaxation
- Gemnet-XL obtained SOTA on each task (at the time of publication)

Experiments - S2EF

- On S2EF, GemNet-XL obtained a 15% lower MAE than previous SOTA

Force Prediction on OC20 S2EF

Model	#Params	Force MAE
ForceNet-Large	34.8M	0.0312
DimeNet++-Large	10.8M	0.0313
SpinConv	8.9M	0.0297
GemNet-T	31M	0.0242
GemNet-XL (Ours)	**300M**	**0.0204**

Experiments - IS2RS

- Structure Relaxation
  - Use the predicted forces to obtain a low energy configuration
- 21% better AFbT than previous SOTA

Structure Relaxation on OC20 IS2RS

Model	#Params	AFbT
SpinConv	8.9M	16.67%
DimeNet++	10.8M	17.15%
DimeNet++-Large	34.8M	21.82%
GemNet-T	31M	27.60%
GemNet-XL (Ours)	**300M**	**30.82%**
DimeNet++-XL (Ours)	**240M**	**33.44%**

Experiments - IS2RE

- Gemnet-XL obtains a 8% lower Energy MAE than previous SOTA by doing a structure relaxation
- Gemnet-XL model can also be fine-tuned on the IS2RE task for faster inference
  - Fine-tuned Gemnet-XL obtains SOTA for direct prediction

Structure Relaxation on OC20 IS2RS

Model	#Params	Approach	Energy MAE
SpinConv	8.9M	Relaxation	0.4343
Gemnet-T	10.8M	Relaxation	0.3997
GemNet-XL (Ours)	**300M**	Relaxation	**0.3712**
Noisy Nodes	34.8M	Direct	0.4728
3D-Graphormer	31M	Direct	0.4722
GemNet-XL FT (Ours)	**300M**	Direct	**0.4623**

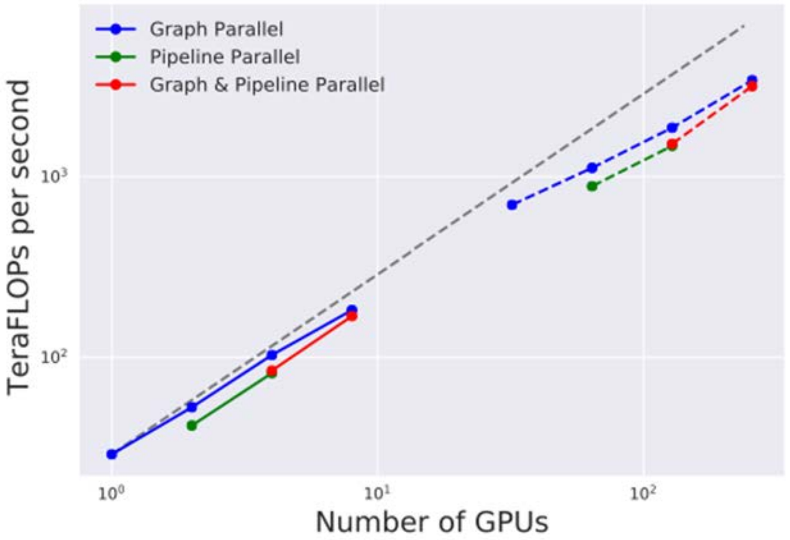
Scaling Analysis

- Weak scaling analysis
  - Scale computation (model size) with number of GPUs and study how runtime scales
  - Trained GemNet-T models with upto 960M parameters with 8-way parallelism

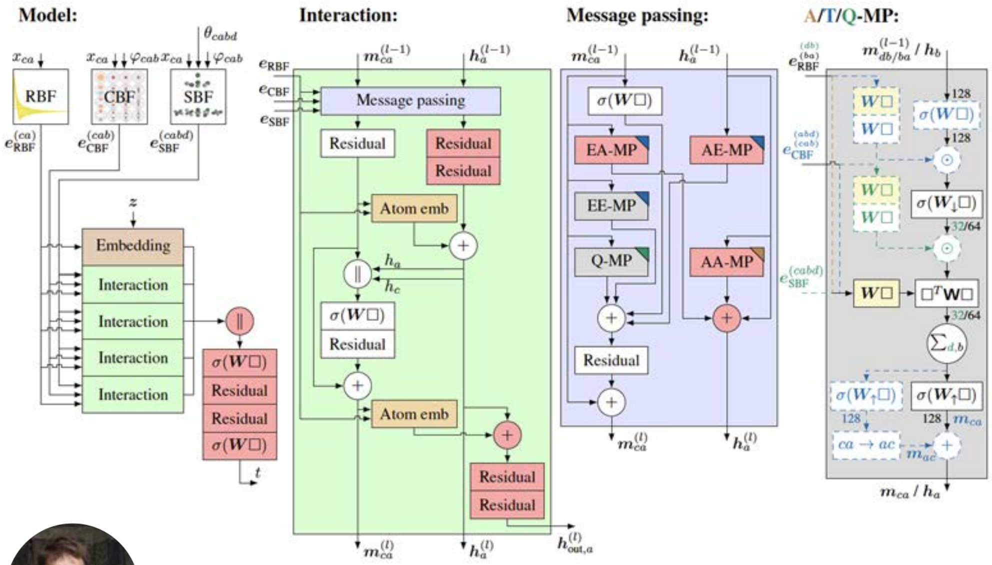
#Blocks	Node Dim	Edge Dim	Trip Dim	Bil Dim	Params	#GP GPUs	#GP+DP GPUs
3	1280	768	128	64	125M	1	32
4	1536	1024	192	96	245M	2	64
6	1792	1184	288	160	480M	4	128
8	2320	1302	512	288	960M	8	256

Scaling Analysis

- We see excellent throughput upto 256 GPUs for the Gemnet-T model.
- Better scaling throughput than pipeline parallel training
- GP & PP can also be combined together



GemNet-OC



Johannes Klicpera

Klicpera et al., How Do Graph Networks Generalize to Large and Diverse Molecular Systems?, 2022

Results

Train set	Model	S2EF validation				S2EF test				IS2RS		IS2RE
		Energy MAE meV ↓	Force MAE meV/Å ↓	Force cos ↑	EFwT % ↑	Energy MAE meV ↓	Force MAE meV/Å ↓	Force cos ↑	EFwT % ↑	AFbT % ↑	ADwT % ↑	Energy MAE meV ↓
OC-2M	SchNet	1400	78.3	0.109	0.00	1370	77.1	0.116	0.00	-	-	-
	DimeNet^{++}	805	65.7	0.217	0.01	761	63.0	0.222	0.01	-	-	-
	SpinConv	406	36.2	0.479	0.13	401	35.5	0.475	0.13	-	-	-
	GemNet-dT	358	29.5	0.557	0.61	323	28.1	0.559	0.69	16.7	54.8	438
	GemNet-OC	**286**	**25.7**	**0.598**	**1.06**	**274**	**24.3**	**0.603**	**1.25**	**19.6**	**56.4**	**407**
OC20	CGCNN	590	74.0	0.142	0.01	608	73.3	0.146	0.01	-	-	-
	SchNet	549	56.8	0.297	0.06	540	54.7	0.302	0.06	-	14.4	764
	ForceNet-large	-	33.5	0.515	-	-	32.0	0.516	0.01	12.7	49.6	-
	DimeNet^{++}-L-F+E	515	32.8	0.541	0.00	480	31.3	0.544	0.00	21.7	51.7	559
	SpinConv	371	41.2	0.473	0.05	336	29.7	0.539	0.45	16.7	53.6	437
	GemNet-dT	315	27.2	0.594	0.54	292	24.2	0.616	1.20	27.6	58.7	400
	GemNet-XL	-	-	-	-	270	**20.5**	0.660	1.81	30.8	**62.7**	371
	GemNet-OC	**244**	**21.7**	**0.662**	**2.07**	**233**	20.7	**0.666**	**2.50**	**35.3**	60.3	**355**
OC20+	GemNet-OC-L-E	**239**	22.1	0.662	2.30	**230**	21.0	0.665	2.80	-	-	-
OC-MD	GemNet-OC-L-F	252	**20.0**	**0.687**	**2.51**	241	**19.0**	**0.691**	**2.97**	**40.6**	60.4	-
	GemNet-OC-L-F+E	-	-	-	-	-	-	-	-	-	-	**348**

Results

Train		S2EF validation				S2EF test				IS2RS		IS2RE
		Energy MAE	Force MAE	Force cos	EFwT	Energy MAE	Force MAE	Force cos	EFwT	AFbT	ADwT	Energy MAE
set	Model	meV ↓	meV/Å ↓	↑	% ↑	meV ↓	meV/Å ↓	↑	% ↑	% ↑	% ↑	meV ↓
OC-2M	SchNet	1400	78.3	0.109	0.00	1370	77.1	0.116	0.00	-	-	-
	DimeNet^{++}	805	65.7	0.217	0.01	761	63.0	0.222	0.01	-	-	-
	SpinConv	406	36.2	0.479	0.13	401	35.5	0.475	0.13	-	-	-
	GemNet-dT	358	29.5	0.557	0.61	323	28.1	0.559	0.69	16.7	54.8	438
	GemNet-OC	**286**	**25.7**	**0.598**	**1.06**	**274**	**24.3**	**0.603**	**1.25**	**19.6**	**56.4**	**407**
OC20	CGCNN	590	74.0	0.142	0.01	608	73.3	0.146	0.01	-	-	-
	SchNet	549	56.8	0.297	0.06	540	54.7	0.302	0.06	-	14.4	764
	ForceNet-large	-	33.5	0.515	-	-	32.0	0.516	0.01	12.7	49.6	-
	DimeNet^{++}-L-F+E	515	32.8	0.541	0.00	480	31.3	0.544	0.00	21.7	51.7	559
	SpinConv	371	41.2	0.473	0.05	336	29.7	0.539	0.45	16.7	53.6	437
	GemNet-dT	315	27.2	0.594	0.54	292	24.2	0.616	1.20	27.6	58.7	400
	GemNet-XL	-	-	-	-	270	**20.5**	0.660	1.81	30.8	**62.7**	371
	GemNet-OC	**244**	**21.7**	**0.662**	**2.07**	**233**	20.7	**0.666**	**2.50**	**35.3**	60.3	**355**
OC20+	GemNet-OC-L-E	**239**	22.1	0.662	2.30	**230**	21.0	0.665	2.80	-	-	-
OC-MD	GemNet-OC-L-F	252	**20.0**	**0.687**	**2.51**	241	**19.0**	**0.691**	**2.97**	**40.6**	60.4	-
	GemNet-OC-L-F+E	-	-	-	-	-	-	-	-	-	-	**348**

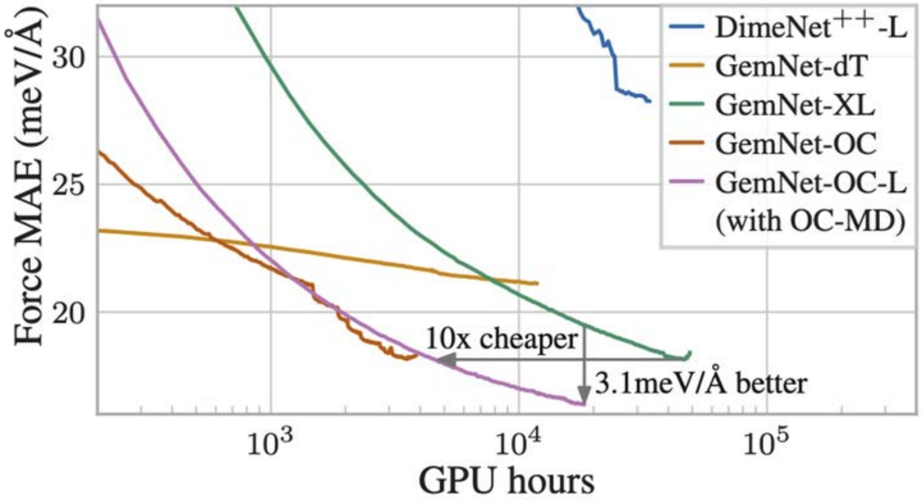
Results

Train set	Model	S2EF validation Energy MAE meV ↓	S2EF validation Force MAE meV/Å ↓	S2EF validation Force cos ↑	S2EF validation EFwT % ↑	S2EF test Energy MAE meV ↓	S2EF test Force MAE meV/Å ↓	S2EF test Force cos ↑	S2EF test EFwT % ↑	IS2RS AFbT % ↑	IS2RS ADwT % ↑	IS2RE Energy MAE meV ↓
OC-2M	SchNet	1400	78.3	0.109	0.00	1370	77.1	0.116	0.00	-	-	-
	DimeNet++	805	65.7	0.217	0.01	761	63.0	0.222	0.01	-	-	-
	SpinConv	406	36.2	0.479	0.13	401	35.5	0.475	0.13	-	-	-
	GemNet-dT	358	29.5	0.557	0.61	323	28.1	0.559	0.69	16.7	54.8	438
	GemNet-OC	**286**	**25.7**	**0.598**	**1.06**	**274**	**24.3**	**0.603**	**1.25**	**19.6**	**56.4**	**407**
OC20	CGCNN	590	74.0	0.142	0.01	608	73.3	0.146	0.01	-	-	-
	SchNet	549	56.8	0.297	0.06	540	54.7	0.302	0.06	-	14.4	764
	ForceNet-large	-	33.5	0.515	-	-	32.0	0.516	0.01	12.7	49.6	-
	DimeNet++-L-F+E	515	32.8	0.541	0.00	480	31.3	0.544	0.00	21.7	51.7	559
	SpinConv	371	41.2	0.473	0.05	336	29.7	0.539	0.45	16.7	53.6	437
	GemNet-dT	315	27.2	0.594	0.54	292	24.2	0.616	1.20	27.6	58.7	400
	GemNet-XL	-	-	-	-	270	**20.5**	0.660	1.81	30.8	**62.7**	371
	GemNet-OC	**244**	**21.7**	**0.662**	**2.07**	**233**	20.7	**0.666**	**2.50**	**35.3**	60.3	**355**
OC20+	GemNet-OC-L-E	**239**	22.1	0.662	2.30	**230**	21.0	0.665	2.80	-	-	-
OC-MD	GemNet-OC-L-F	252	**20.0**	**0.687**	**2.51**	241	**19.0**	**0.691**	**2.97**	**40.6**	60.4	-
	GemNet-OC-L-F+E	-	-	-	-	-	-	-	-	-	-	**348**

Results

Train set	Model	S2EF validation Energy MAE meV ↓	S2EF validation Force MAE meV/Å ↓	S2EF validation Force cos ↑	S2EF validation EFwT % ↑	S2EF test Energy MAE meV ↓	S2EF test Force MAE meV/Å ↓	S2EF test Force cos ↑	S2EF test EFwT % ↑	IS2RS AFbT % ↑	IS2RS ADwT % ↑	IS2RE Energy MAE meV ↓
OC-2M	SchNet	1400	78.3	0.109	0.00	1370	77.1	0.116	0.00	-	-	-
	DimeNet^{++}	805	65.7	0.217	0.01	761	63.0	0.222	0.01	-	-	-
	SpinConv	406	36.2	0.479	0.13	401	35.5	0.475	0.13	-	-	-
	GemNet-dT	358	29.5	0.557	0.61	323	28.1	0.559	0.69	16.7	54.8	438
	GemNet-OC	**286**	**25.7**	**0.598**	**1.06**	**274**	**24.3**	**0.603**	**1.25**	**19.6**	**56.4**	**407**
OC20	CGCNN	590	74.0	0.142	0.01	608	73.3	0.146	0.01	-	-	-
	SchNet	549	56.8	0.297	0.06	540	54.7	0.302	0.06	-	14.4	764
	ForceNet-large	-	33.5	0.515	-	-	32.0	0.516	0.01	12.7	49.6	-
	DimeNet^{++}-L-F+E	515	32.8	0.541	0.00	480	31.3	0.544	0.00	21.7	51.7	559
	SpinConv	371	41.2	0.473	0.05	336	29.7	0.539	0.45	16.7	53.6	437
	GemNet-dT	315	27.2	0.594	0.54	292	24.2	0.616	1.20	27.6	58.7	400
	GemNet-XL	-	-	-	-	270	**20.5**	0.660	1.81	30.8	**62.7**	371
	GemNet-OC	**244**	**21.7**	**0.662**	**2.07**	**233**	20.7	**0.666**	**2.50**	**35.3**	60.3	**355**
OC20+	GemNet-OC-L-E	**239**	22.1	0.662	2.30	**230**	21.0	0.665	2.80	-	-	-
OC-MD	GemNet-OC-L-F	252	**20.0**	**0.687**	**2.51**	241	**19.0**	**0.691**	**2.97**	**40.6**	60.4	-
	GemNet-OC-L-F+E	-	-	-	-	-	-	-	-	-	-	**348**

Results

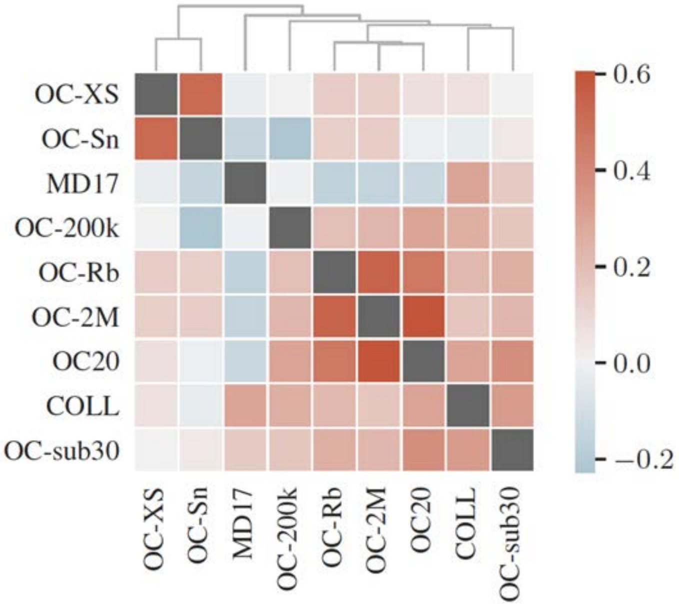


10^4 GPU hours = ~7 days on 64 GPUs

Common molecular simulation datasets

Dataset	Description	Elements	Neighbor pairs	Avg. size	Train set size
MD17	Eight separate molecules	H, C, N, O	3-10	12.5 (9-21)	8 × 1000
ISO17	$C_7O_2H_{10}$ isomers	H, C, O	6	19	404 000
S_N2	Methyl halides, halide ions	H, C, F, Cl, Br, I	20	5.4 (2-6)	400 000
ANI-1x	Selected MD samples	H, C, N, O	10	15.3 (2-63)	4 956 005
QM7-X	Small molecules	H, C, N, O, S, Cl	20	16.7 (4-23)	4 175 037
COLL	Molecule collisions	H, C, O	6	10.2 (2-26)	120 000
OC20	Relaxations of catalysts	56	1454	73.3 (7-225)	133 934 018
OC-Rb	Only H, C, N, O, Rb	H, C, N, O, Rb	15	39.1 (7-220)	524 736
OC-Sn	Only H, C, N, O, Sn	H, C, N, O, Sn	15	59.5 (22-220)	257 757
OC-sub30	At most 30 atoms	55	881	24.6 (7-30)	4 020 568
OC-200k	Random subset	56	1454	73.2 (7-225)	200 000
OC-2M	Random subset	56	1454	73.3 (7-225)	2 000 000
OC-XS	$\leq$ 30 H, C, N, O, Rb atoms	H, C, N, O, Rb	15	19.7 (7-30)	298 797

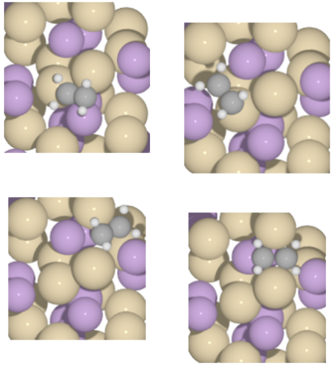
Correlation of model results across datasets



OC-2M correlates well with OC20

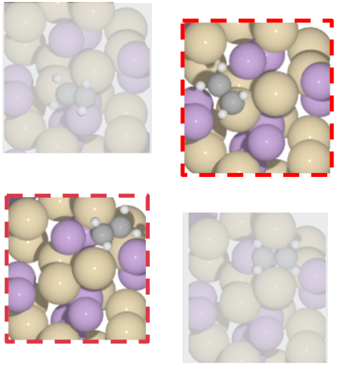
ML + DFT: Retrieving the "best" site

Goal: model's ability to find the lowest energy.



Possible configurations

ML



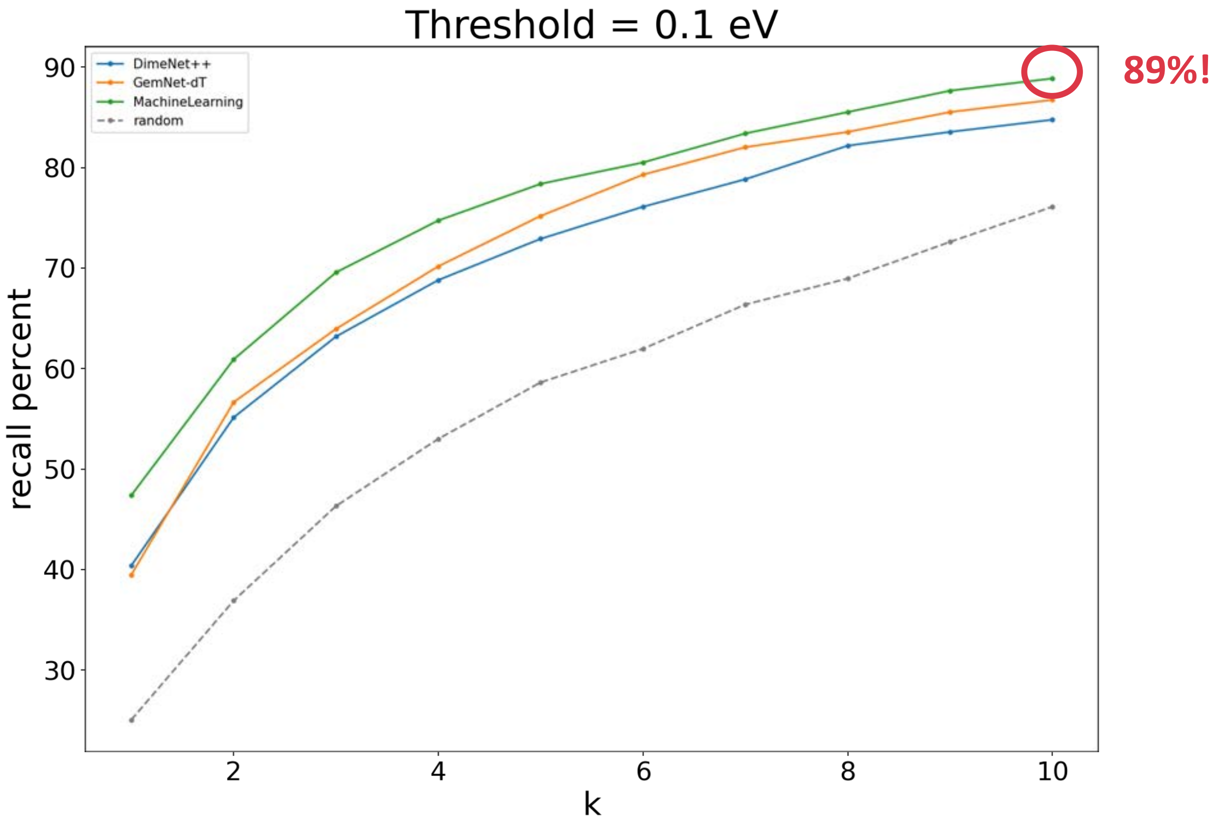
Select top **k** configurations with lowest predicted E_{ads}

DFT

$|\min(E_{ads,\ topk}) - E_{best}| <= \epsilon$?

Y

recalled


Threshold = 0.1 eV
DimeNet++
GemNet-dT
MachineLearning
random
90
80
70
60
50
40
30
recall percent
2
4
6
8
10
k
89%!

Dataset, papers, code:

opencatalystproject.org

github.com/open-catalyst-project

If you have any questions, please reach out to

Abhishek Das (abhshkdz@fb.com)

Anuroop Sriram (anuroops@fb.com)

or on our discussion board

discuss.opencatalystproject.org

MedeA Datasheets Related to This Webinar

MedeA Environment**:** Materials Modeling and Simulation Environment

***MedeA* VASP:** Access the World Leading First-Principles DFT Code

***MedeA* HT:** High Throughput Flowchart Calculations

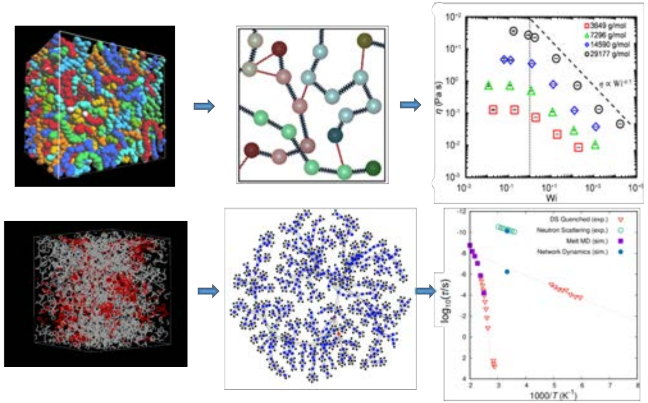
***MedeA* QT:** An Interactive QSPR Toolbox

***MedeA* TSS:** Maps out Reaction and Diffusion Pathways, Locates Transition States, Calculates Activation Barriers

Announcement

Upcoming

- Webinar Next Month: Atomistic and Mesoscopic Modeling of Structure-Property Relations in Polymers
- Professor Doros Theodorou from the National Technical University of Athens
    - May 24-26th
    - www.materialsdesign.com/webinars



materials design

Question and Answer Session

Erich Wimmer

Materials Design

Abhishek Das

Facebook AI Research

Anuroop Sriram

Facebook AI Research

Questions about Materials Design Webinars

Katherine Hollingsworth

khollingsworth@materialsdesign.com

materials design

info@materialsdesign.com

www.materialsdesign.com

MedeA

Innovation by Simulation