SAMSUNG

NEURAL INFORMATION PROCESSING SYSTEMS

Benchmark of Machine Learning Force Fields for Semiconductor Simulations: Datasets, Metrics, and Comparative Analysis

Geonu Kim* Byunggook Na* Gunhee Kim† Hyuntae Cho Seung-Jin Kang
Hee Sun Lee Saerom Choi Heejae Kim Seungwon Lee Yongdeok Kim†

Samsung Advanced Institute of Technology (SAIT)

*Equal contribution †Co-corresponding author

01 | Overview

SAMSUNG

Two
semiconductor datasets
(SiN and HfO)

Six
evaluation metrics

Ten
MLFF model
benchmark results

02 | Machine Learning Force Fields (MLFFs)

SAMSUNG

What is MLFF?

An MLFF model predicts energy $\boldsymbol{E} \in \mathbb{R}$ and forces $\boldsymbol{F} \in \mathbb{R}^{n\times 3}$, given atom positions $\boldsymbol{r} \in \mathbb{R}^{n\times 3}$ and atomic numbers $\mathbf{Z} \in \mathbb{Z}^{n}$, where n indicates the number of atoms in a system.

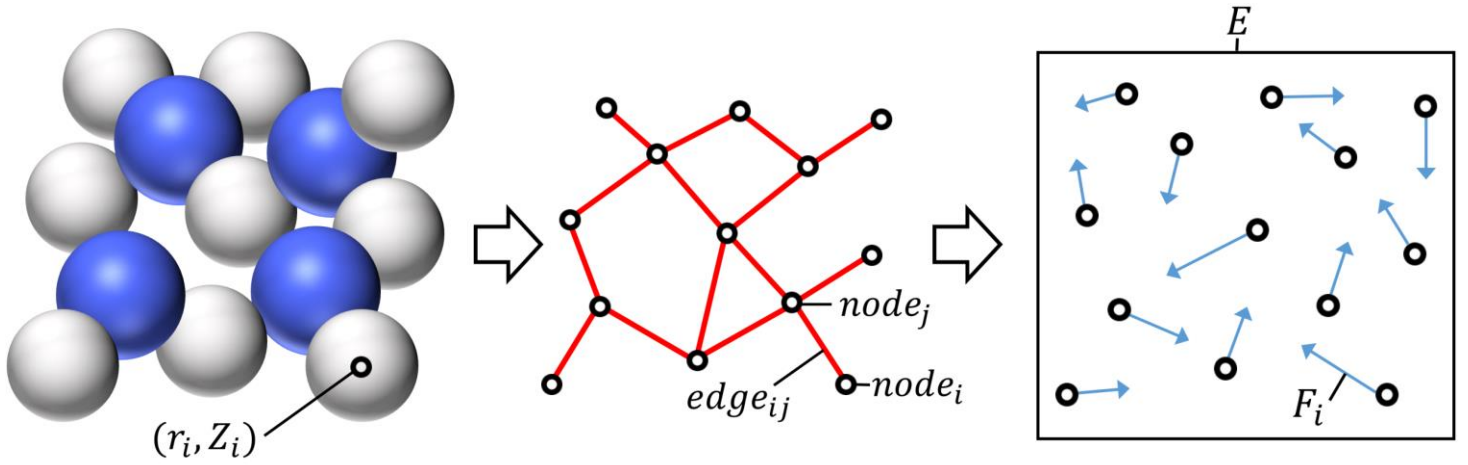
MLFF models are trained using datasets generated by DFT-based simulations that are **accurate but cost-expensive**.

Density functional theory

Why are MLFF models studied?

In classical MD simulations, which are less accurate but faster than DFT-based simulations, MLFF models can play a role in **achieving DFT-level accuracy while maintaining the efficiency and scalability of the MD simulations**.

Molecular dynamics



(a) [Input] An atomic structure (b) [Preprocessed] Graph (Optional) (c) [Output] Energy and Forces

03 | Motivation

SAMSUNG

Insufficient semiconductor datasets

In ML community, molecular datasets (ANI, rMD17, COLL, and 3BPA) and open-catalyst datasets (OC20, and OC22) are mainly used.

Insufficient evaluation metrics

Errors of the prediction of **energy** and **force** are **only used to evaluate MLFF models**.

In previous semiconductor datasets, it is **insufficient** to evaluate various material properties **without domain expertise**.

▶ **Our benchmark suite** includes **two semiconductor datasets** and **six evaluation metrics** considering not only the prediction of energy and force but also the performance in simulations.

04 | Datasets

SAMSUNG

Silicon Nitride (SiN)

Generated by using diverse structures from domain expertise

A varied number of atoms (16 to 510) and a wide range of stoichiometries, including Si:N ratio of 1:1, 3:4, along with pure silicon (1:0) and pure nitrogen (0:1)

Hafnium Oxide (HfO)

Generated by using a modified-MQA method, which requires less domain expertise

A single stoichiometry, Hf:O=1:2 and 96 atoms

	Element	N_{cond}	Snapshots	Datapoints	N_{train}	N_{valid}	N_{test}	Cost (h)
SiN	$Si_m\ N_n$	92	76,213	4,397,744	20,315	2,542	2,585	29,824
	Si	14	6,250	291,600	1,663	212	213	2,963
	N	4	2,000	128,000	532	68	68	3,769
	Total	110	84,463	4,817,344	22,510	2,822	2,866	36,556
SiN^{OOD}	$Si_m\ N_n$	3	3,700	388,500	-	-	1,235	1,166
HfO	$Hf\ O_2$	60	160,000	15,360,000	27,960	3,510	3,510	19,341
HfO^{OOD}	$Hf_m\ O_n$	12	32,000	3,072,000	-	-	6,996	4,182

▶ Total generation cost

: 2600 GPU days (VASP, V100)

05 | Evaluation Metrics

SAMSUNG

Prediction performance for energy and force

EF metric = RMSE of per-atom energy + RMSE of forces

Dynamic indicators required to perform high-temperature MD simulations

Radial distribution function (**RDF**)

Angular distribution function (**ADF**)

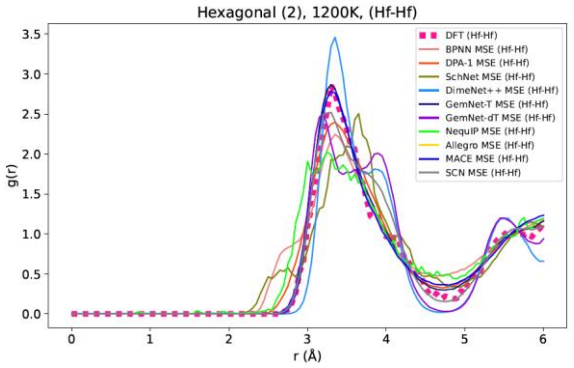
Static indicators required to predict energy estimates for reference structures

Bulk modulus (B_0) and Equilibrium volume (V_0) parameters in the Birch–Murnaghan equation of state (EoS)

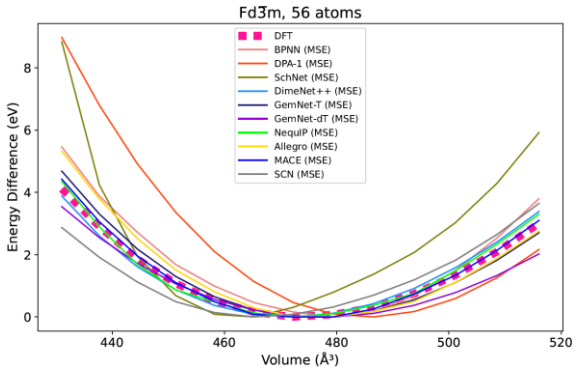
Potential energy curves (**PECs**)

Examples

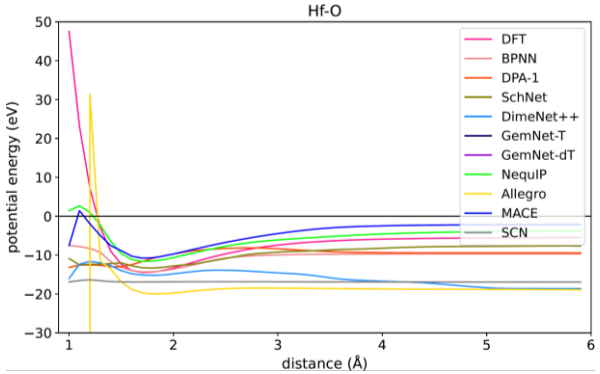
RDF



EoS



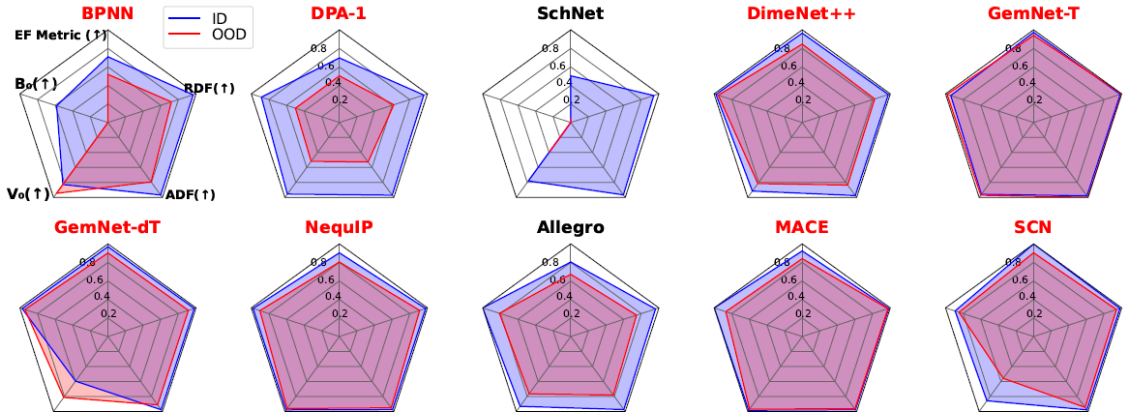
PECs



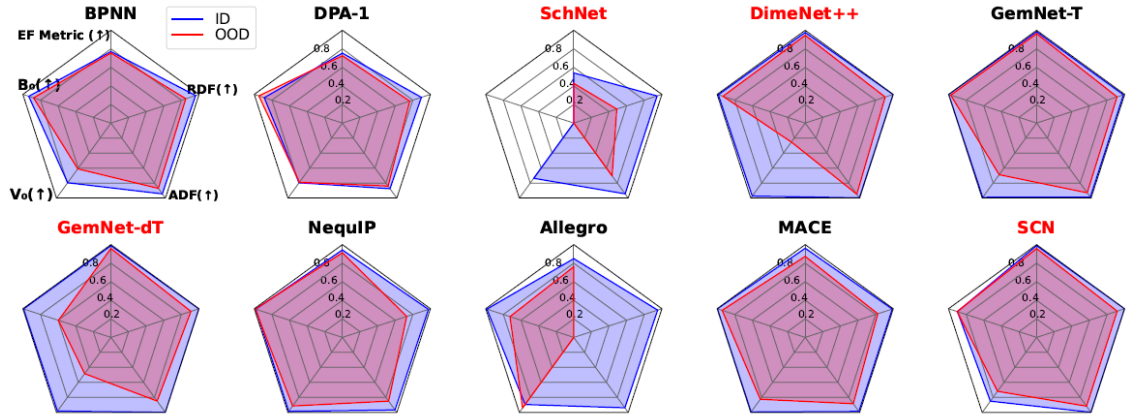
06 | Benchmark Results

SAMSUNG

Evaluation

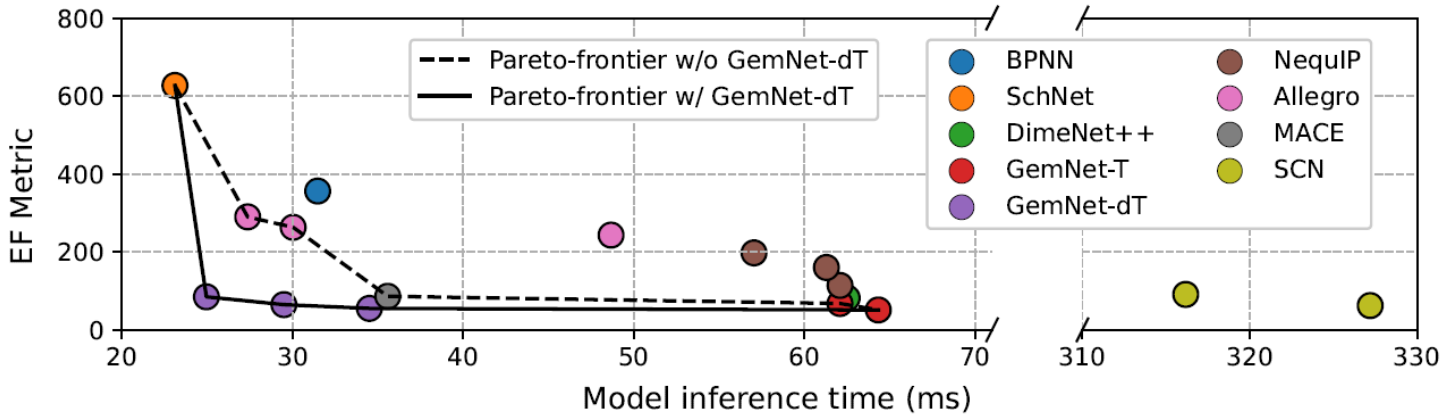


(a) SiN



(b) HfO

Model efficiency



07 | Summary

SAMSUNG

Takeaway messages

To properly evaluate MLFF model performance in simulations, **our simulation indicators are necessary** beyond the prediction of energy and force.

There is **no clear winning model** for large-scale semiconductor MD simulations.

Future work

Extension of diverse semiconductor datasets by including various elements

Efficient training along with the dataset extension

Subjoining additional metrics and loss factors such as stress

Benchmark suite information

https://github.com/SAITPublic/MLFF-Framework

This includes the download links of our datasets and tools to train and evaluate MLFF models