Gaussian Plane-Wave Neural Operator for Electron Density Estimation

Seongsu Kim and Sungsoo Ahn

Pohang University of Science and Technology (POSTECH)

ICML '24

Presenter: Seongsu Kim

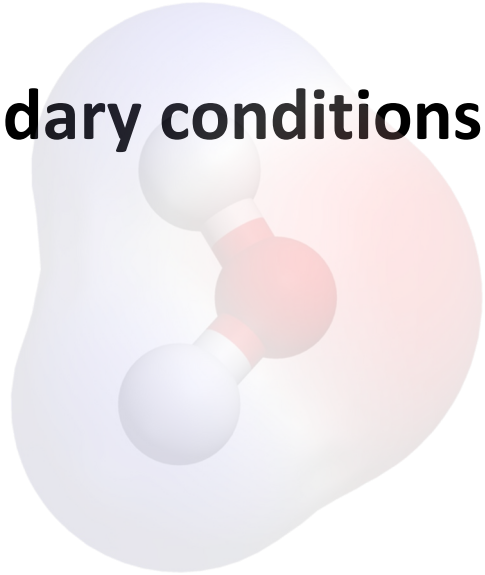
seongsukim@postech.ac.kr
sungsoo.ahn@postech.ac.kr

ICML
International Conference
On Machine Learning

POSTECH
POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

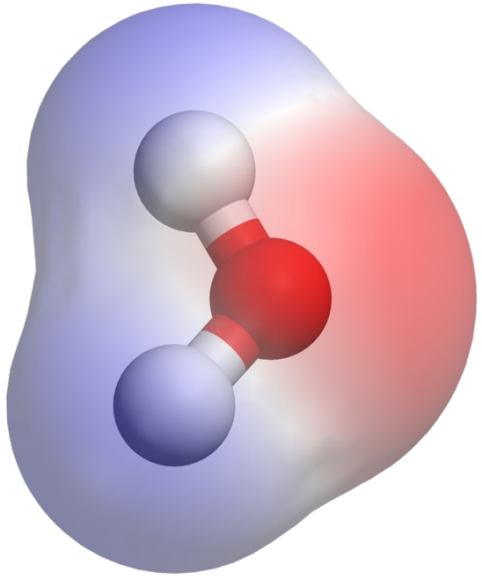
Summary

- Proposed the machine learning model for **estimating electron density**
- Used **Gaussian-type orbital (GTO)** and **Plane Wave (PW)** basis function
- Our framework is a **Neural Operator** since it is the mapping btw. functions
- Our framework satisfy **SE(3)-equivariance** and **periodic boundary conditions**
- We empirically verify **the effectiveness of mixing bases**



Electron density estimation

- **Electron density** is the probability of finding electrons in a specific area
  - According to quantum mechanics, **the exact location of an electron cannot be predicted**, due to uncertainty principle
  - We can only find the probability of electron existence at given point



Electron density plot of water (H_2O)
The surface implies the isosurface where probability is identical

[1] Griffiths, David J., and Darrell F. Schroeter. Introduction to quantum mechanics. Cambridge university press, 2018.

Application of electron density

- Electron density can be used to predict the **material's atomic properties**
  (energy, reactivity, …)
- From **density functional theory (DFT)**, the properties of molecular system can be determined by the functionals depends on the **electron density**

Battery cathodes design

Drug design

Solar cell materials design

[1] Becke, Axel D. "Density-functional exchange-energy approximation with correct asymptotic behavior." *Physical review A* 38.6 (1988): 3098.
[2] Wu, Xi, et al. "Density functional theory calculations: A powerful tool to simulate and design high-performance energy storage and conversion materials." *Progress in Natural Science: Materials International* 29.3 (2019): 247-255.
[3] Rozhenko, Alexander B. "Density functional theory calculations of enzyme–inhibitor interactions in medicinal chemistry and drug design." *Application of computational techniques in pharmacy and medicine* (2014): 207-240.
[4] Chang, Junli, et al. "Lead-free perovskite compounds CsSn 1− x Ge x I 3− y Br y explored for superior visible-light absorption." Physical Chemistry Chemical Physics 23.26 (2021): 14449-14456.

Application of electron density

- Electron density can be used to predict the **material's atomic properties**
  (energy, reactivity, …)
- From **density functional theory (DFT)**, the properties of molecular system can be determined by the functionals depends on the **electron density**

Battery cathodes design

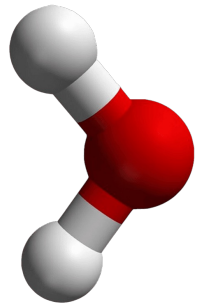
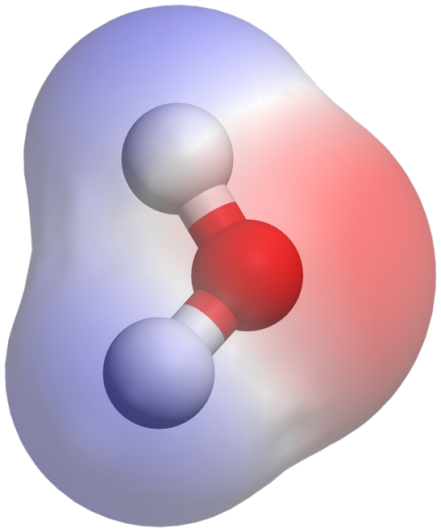
Drug design

Solar cell materials design

[1] Becke, Axel D. "Density-functional exchange-energy approximation with correct asymptotic behavior." *Physical review A* 38.6 (1988): 3098.
[2] Wu, Xi, et al. "Density functional theory calculations: A powerful tool to simulate and design high-performance energy storage and conversion materials." *Progress in Natural Science: Materials International* 29.3 (2019): 247-255.
[3] Rozhenko, Alexander B. "Density functional theory calculations of enzyme–inhibitor interactions in medicinal chemistry and drug design." *Application of computational techniques in pharmacy and medicine* (2014): 207-240.
[4] Chang, Junli, et al. "Lead-free perovskite compounds CsSn 1− x Ge x I 3− y Br y explored for superior visible-light absorption." Physical Chemistry Chemical Physics 23.26 (2021): 14449-14456.

Electron density estimation

- We learn the mapping between **molecules** to the **electron density function**

$f_\theta($  $) \mapsto$ 

Molecular graph
(Sparse scalar field)

Position of atoms $X = [x_1, \cdots, x_n]$
Atomic features $A = [a_1, \cdots, a_n]$

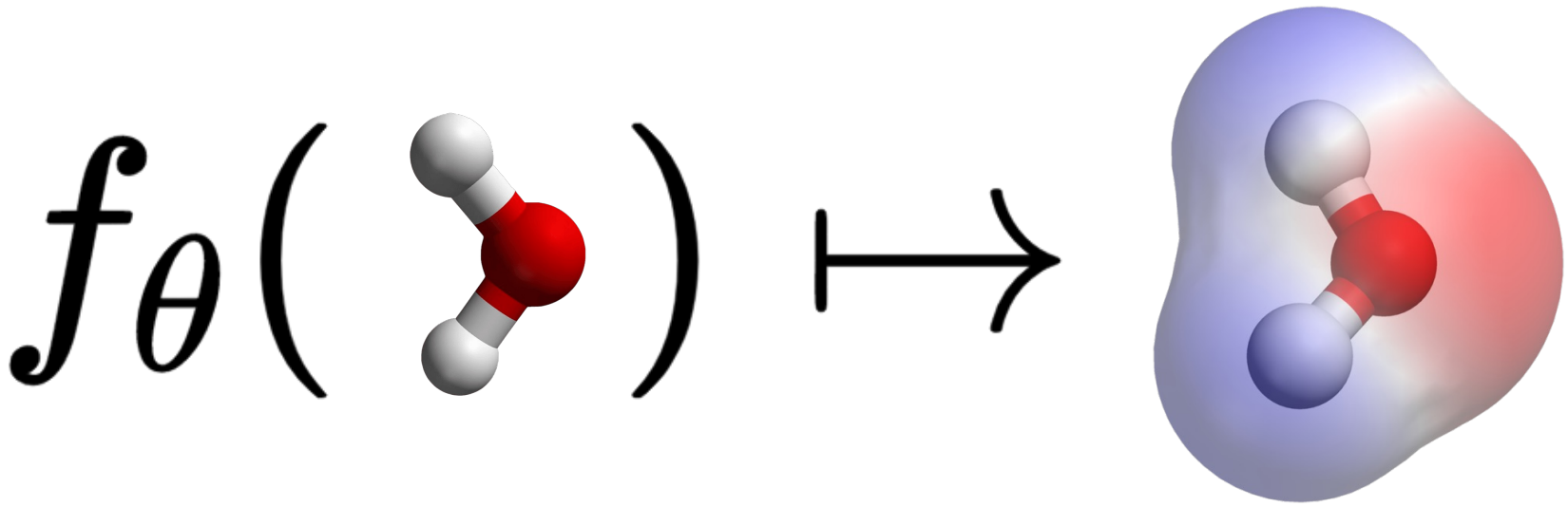
Electron density function
(Continuous scalar field)

Probability (scalar) at query point $\rho(x_q) \in \mathbb{R}$

(Sparse) function to (continuous) function mapping
"Neural Operator"

Electron density estimation

- We learn the mapping between **molecules** to the **electron density function**



Molecular graph
(Sparse scalar field)

Position of atoms $X = [x_1, \cdots, x_n]$
Atomic features $A = [a_1, \cdots, a_n]$

Electron density function
(Continuous scalar field)

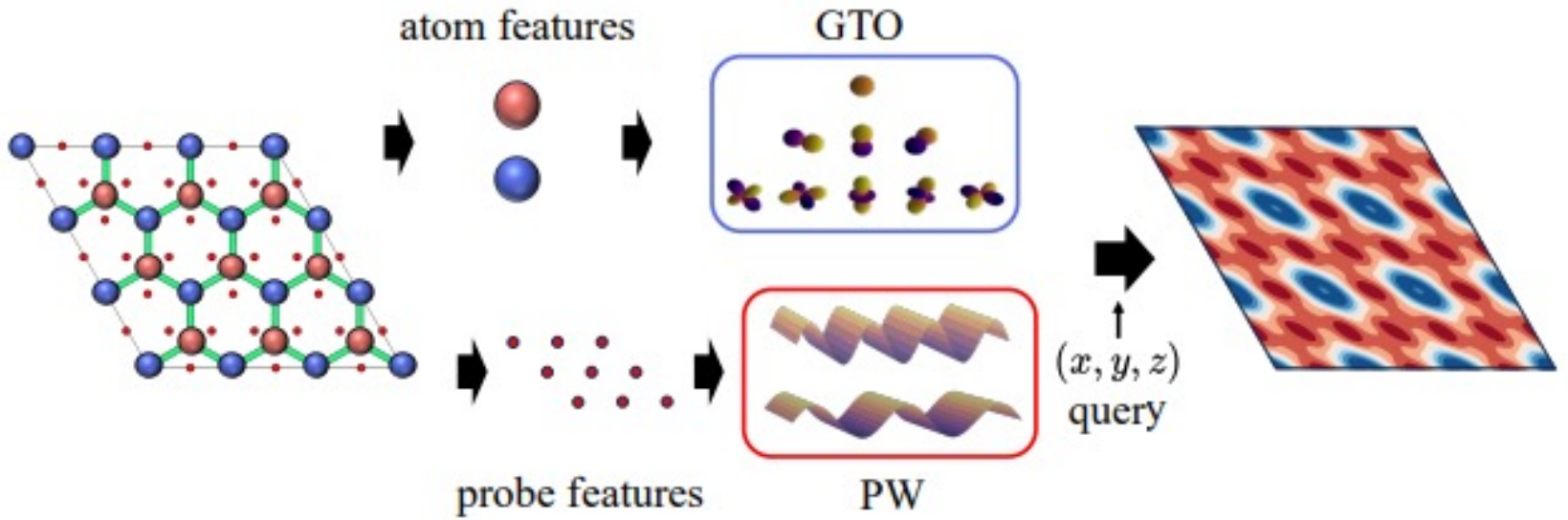
Probability (scalar) at query point $\rho(x_q) \in \mathbb{R}$

(Sparse) function to **(continuous) function** mapping
"Neural Operator"

Overview

- Our GPWNO makes two parallel predictions, GTO- and PW-based predictions.
  - GTO-based prediction utilizes the **coefficient learning** of the GTO basis function
  - PW-based prediction utilizes the **lattice-based discretization** via the **probe nodes**

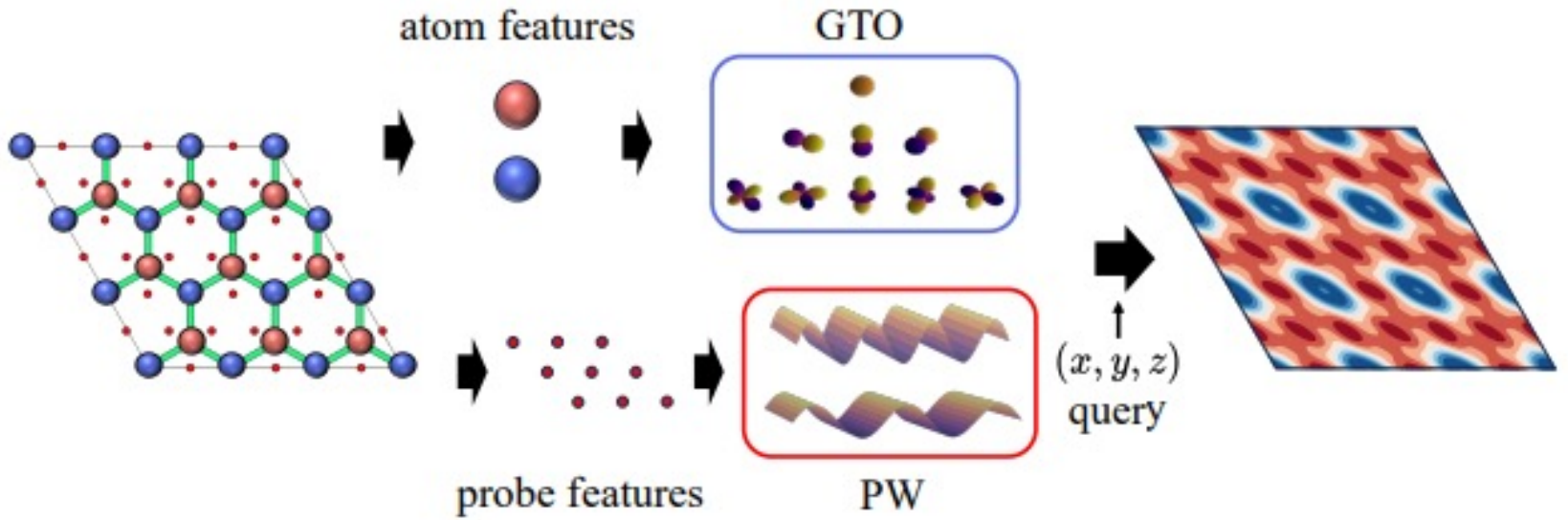
- Output is evaluated at arbitrary **query points** by summing up two prediction layers



Overview

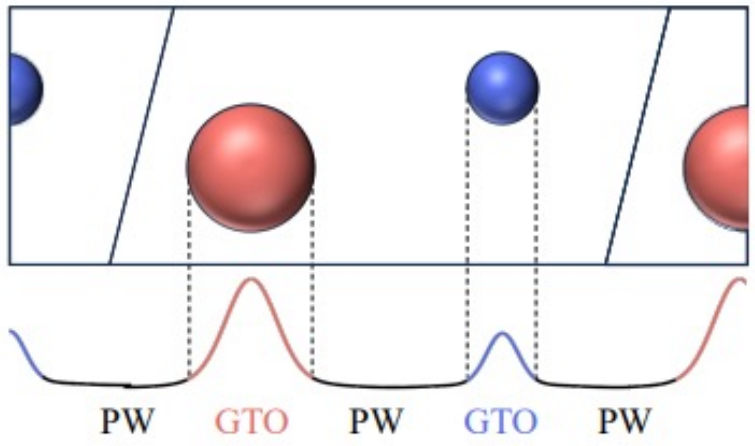
- Our GPWNO makes two parallel predictions, GTO- and PW-based predictions.
  - GTO-based prediction utilizes the **coefficient learning** of the GTO basis function
  - PW-based prediction utilizes the **lattice-based discretization** via the **probe nodes**

- Output is evaluated at arbitrary **query points** by summing up two prediction layers



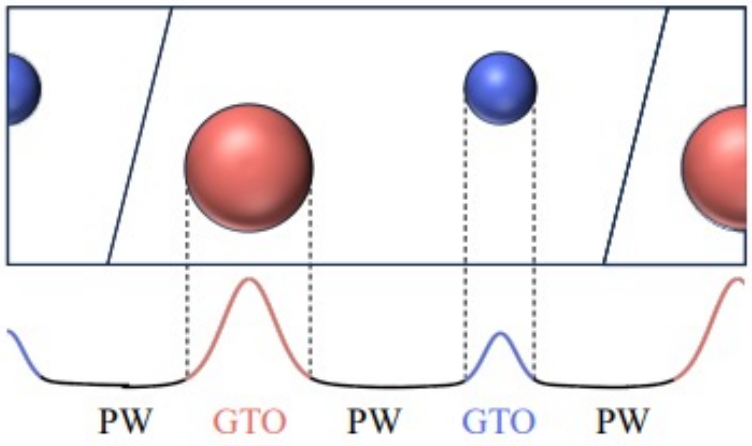
Motivation

- Why do we uses two types of the basis - **GTO** & **PW** ?
- Decomposing the density estimation **into two regions**
  - **Regions near the atoms** / **far from the atoms**
- This makes representation to **focus the region where they are suitable**
  - **Short range – GTO** (Spherical harmonics & radial basis, SO(3) equivariance)
  - **Long range – PW** (Fourier basis, periodicity)



Motivation

- Why do we uses two types of the basis - **GTO** & **PW** ?
- Decomposing the density estimation **into two regions**
  - **Regions near the atoms** / **far from the atoms**
- This makes representation to **focus the region where they are suitable**
  - **Short range – GTO** (Spherical harmonics & radial basis, SO(3) equivariance)
  - **Long range – PW** (Fourier basis, periodicity)

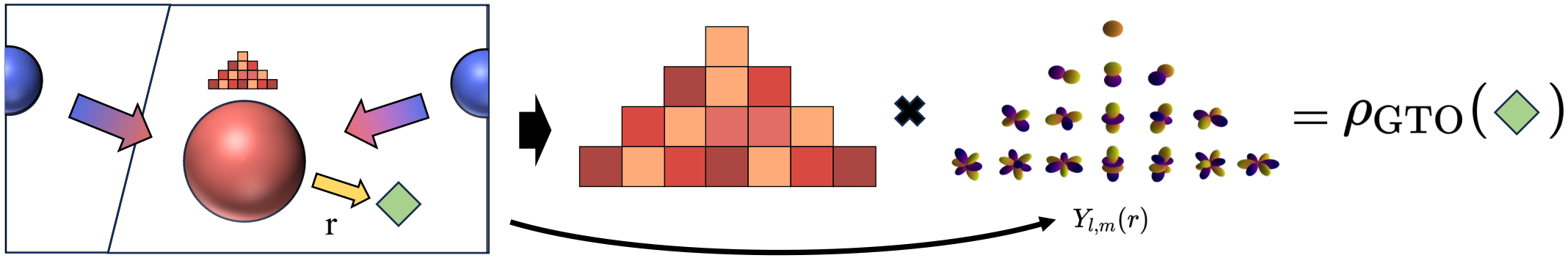


GTO-based prediction

- **Construct prediction from the linear combination of GTO basis**

$$\rho_{\mathrm{GTO}}(\boldsymbol{x}) = \sum_{u \in \text{ atoms}} \sum_{n=1}^{N} \sum_{l=0}^{L} \sum_{m=-l}^{l} \hat{f}_{u,nlm} \phi_{nlm}(\boldsymbol{x}, \boldsymbol{x}_u) \qquad \phi_{nlm} : \text{GTO basis}$$

- Predict the coefficient of spherical harmonics and radial basis by TFN based network and message passing



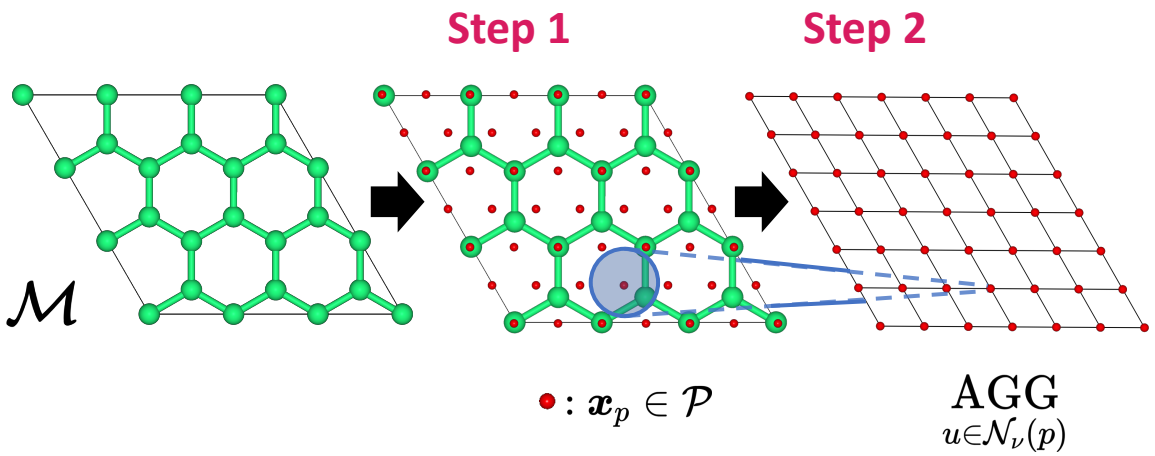
PW-based prediction

Step 1. Placing probe nodes

- Place probe nodes $p \in P$ that discretize the lattice uniformly

Step 2. Aggregating message to probe nodes

- Aggregate the message from the atoms to probe nodes to initialize the probe features



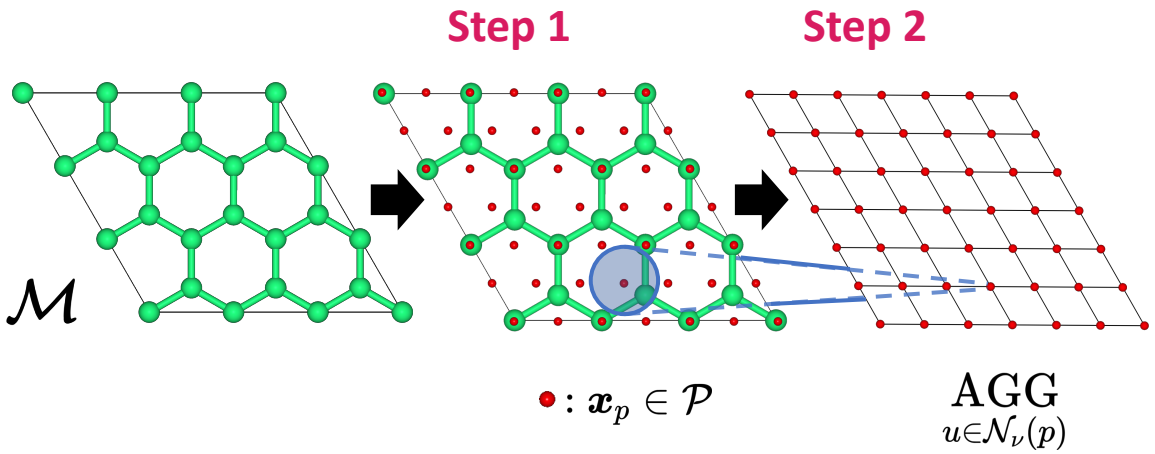
PW-based prediction

Step 1. Placing probe nodes

- Place probe nodes $p \in P$ that discretize the lattice uniformly

Step 2. Aggregating message to probe nodes

- Aggregate the message from the atoms to probe nodes to initialize the probe features



PW-based prediction

Step 3. Iterative convolution on the PW coefficients space

- Iteratively update the probe features by a convolution operator using the PW basis:

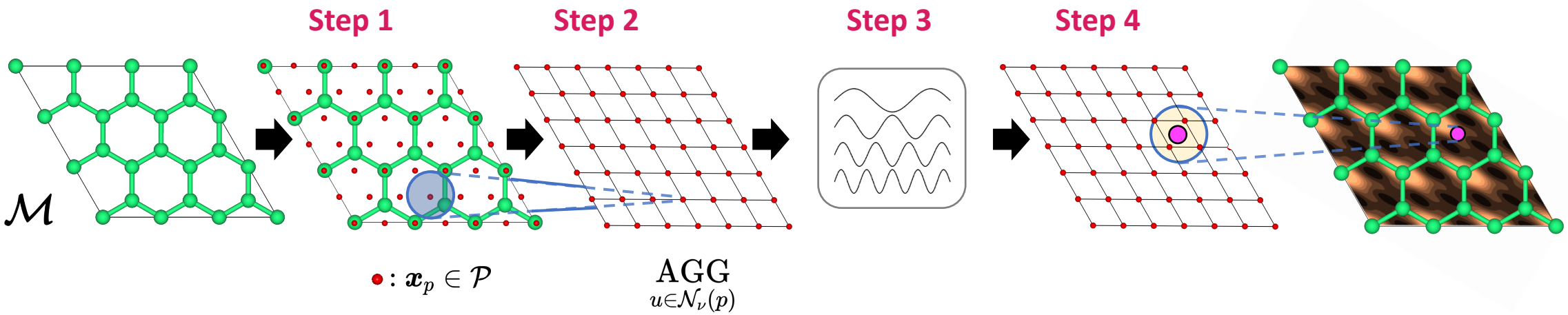
$$f^{(h+1)} = \sigma\left(\mathcal{W} \cdot f^{(h)} + \mathrm{FT}^{-1} \cdot \mathcal{R} \cdot \mathrm{FT}\left(f^{(h)}\right)\right)$$

$$\phi_{\boldsymbol{\lambda}}(\boldsymbol{r}) = \exp\left(j \sum_{d=1}^{3} \lambda_d \boldsymbol{b}_d \cdot \boldsymbol{r}\right)$$

PW basis

Step 4. Prediction network

- At the final layer, given a query point x_q, compute the output by message aggregation from probe nodes to query points



PW-based prediction

Step 3. Iterative convolution on the PW coefficients space

- Iteratively update the probe features by a convolution operator using the PW basis:

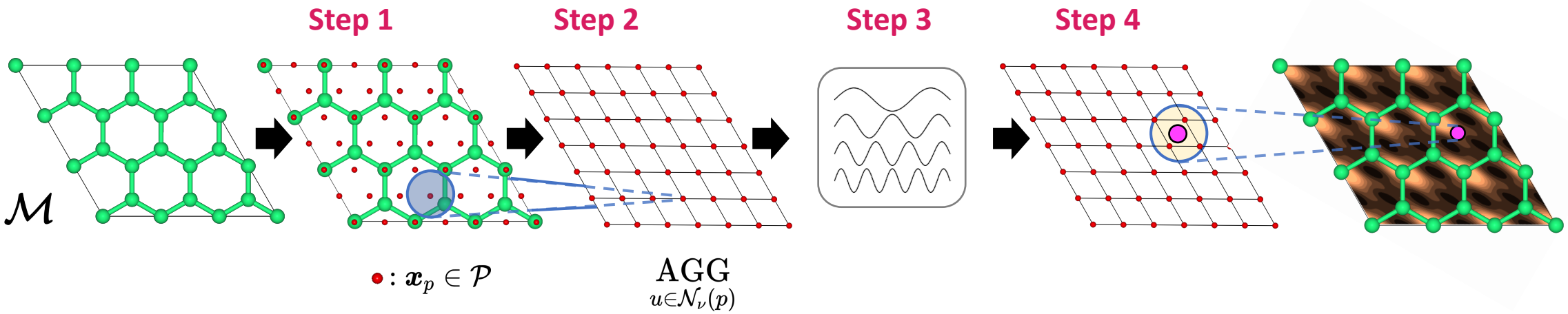
$$f^{(h+1)} = \sigma\left(\mathcal{W} \cdot f^{(h)} + \mathrm{FT}^{-1} \cdot \mathcal{R} \cdot \mathrm{FT}\left(f^{(h)}\right)\right)$$

$$\phi_{\boldsymbol{\lambda}}(\boldsymbol{r}) = \exp\left(j \sum_{d=1}^{3} \lambda_d \boldsymbol{b}_d \cdot \boldsymbol{r}\right)$$

PW basis

Step 4. Prediction network

- At the final layer, given a query point $\boldsymbol{x}_q$, compute the output by message aggregation from probe nodes to query points



Aperiodic materials

Table 1: **Evaluation of GPWNO for aperiodic materials.** We report the performance for QM9 and MD dataset in NMAE (%). The best number is highlighted in **bold**. The baseline results are from Cheng & Peng (2023). Each number is averaged over three runs. For brevity, we denote DimeNet and DeepDFT by DmNet and DDFT, respectively.

Dataset	LNO	FNO	GNO	DmNet++	DmNet	EGNN	DDFT2	DDFT	CNN	InfGCN	GPWNO
QM9	26.14	28.83	40.86	11.69	11.97	11.92	1.03	2.95	2.01	0.93	**0.73**
Ethanol	43.17	31.98	82.35	14.24	13.99	13.90	8.83	7.34	13.97	8.43	**4.00**
Benzene	38.82	20.05	82.46	14.34	14.48	13.49	5.49	6.61	11.98	5.11	**2.45**
Phenol	60.70	42.98	66.69	12.99	12.93	13.59	7.00	9.09	11.52	5.51	**2.68**
Resorcinol	35.07	26.06	58.75	12.01	12.04	12.61	6.95	8.18	11.07	5.95	**2.73**
Ethane	77.14	26.31	71.12	12.95	13.11	15.17	6.36	8.31	14.72	7.01	**3.67**
MDA	47.22	34.58	84.52	16.79	18.71	12.37	10.68	9.31	18.52	10.34	**5.32**

About 30% ↓

About 50% ↓

- Our **GPWNO** outperformed 10 baselines in the aperiodic materials

Aperiodic materials

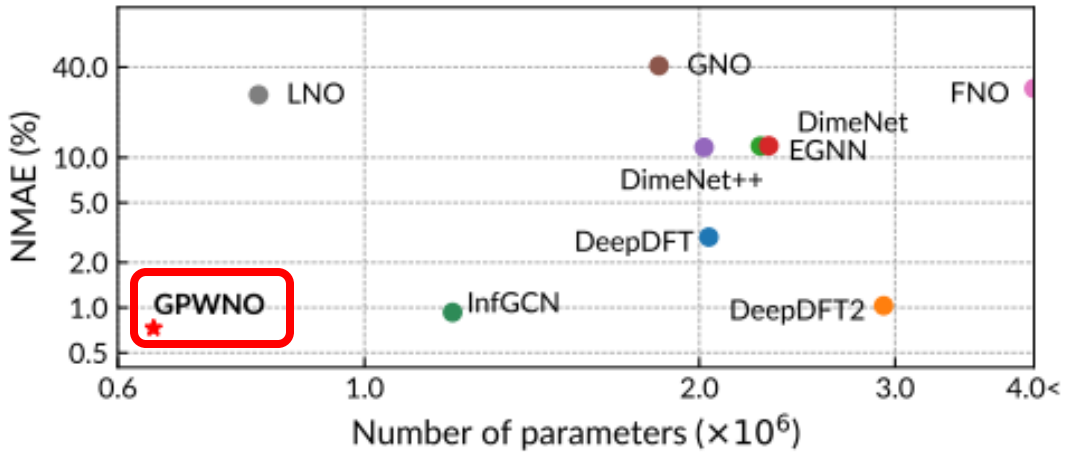


Figure 4: **Comparison between the baselines on QM9 (lower left is better).**

- Also, **GPWNO** excels with the **smallest number of parameters**.

Periodic materials (Crystalline materials)

Table 2: **Evaluation of GPWNO for periodic materials.** NMAE (%) in the MP dataset, categorized by the seven crystal family and their combinations. The best number is highlighted in **bold**. Each number is averaged over three runs.

Model	Mixed	Triclinic	Monoclinic	Orthorhombic	Tetragonal	Trigonal	Hexagonal	Cubic
DeepDFT	11.50	32.33	50.54	22.30	37.68	23.89	18.61	27.30
DeepDFT2	15.11	30.20	49.74	20.50	41.21	22.13	12.87	27.63
InfGCN	5.35	4.36	4.63	5.06	5.02	5.21	4.96	4.93
GPWNO	**4.84**	**3.92**	**4.44**	**4.73**	**4.52**	**4.92**	**4.55**	**4.32**

- Our **GPWNO** outperforms the baselines regardless of the lattice type.
  - We categorized the dataset by the lattice structure

Thank you!

Open to talk! Come say hi in Vienna 👋

Interested in *AI for science* and *geometrical deep learning!*

Poster

Hall C 4-9 #2606
Wednesday 24th July 1:30 PM — 3 PM

PDF

Code