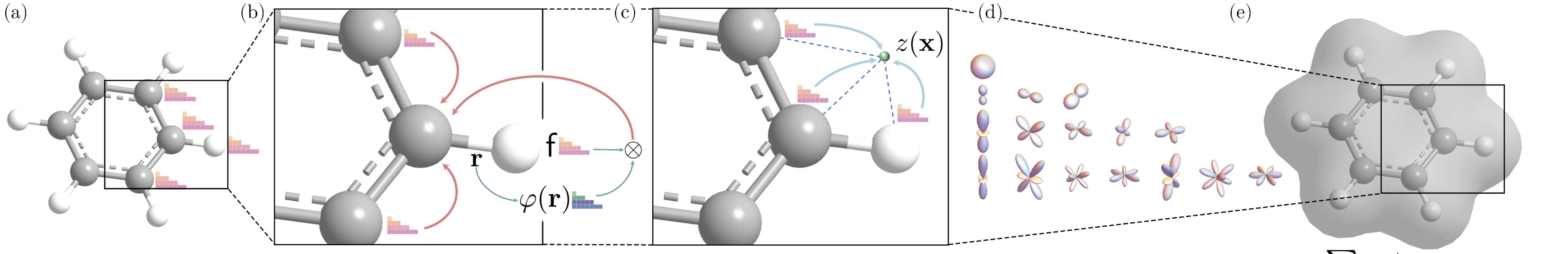
Equivariant Neural Operator Learning with Graphon Convolution

Chaoran Cheng, Jian Peng

NeurIPS23 Spotlight

Department of Computer Science, University of Illinois Urbana-Champaign



$$\hat{\rho}(\mathbf{x}) = \sum_{\ell m} f_{\ell m}^{\text{out}} \cdot \psi_{\ell m}(\mathbf{x}) + z(\mathbf{x})$$

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

NEURAL INFORMATION PROCESSING SYSTEMS

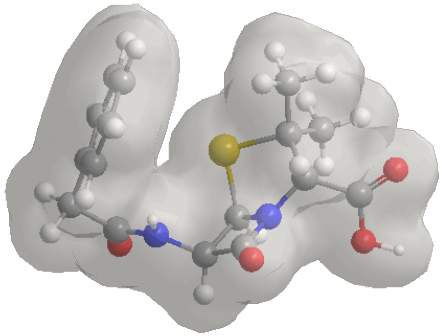
Background

3D Operator Learning

$$\mathcal{T} : f \mapsto g;\, f, g \in L^2(\mathcal{D})$$

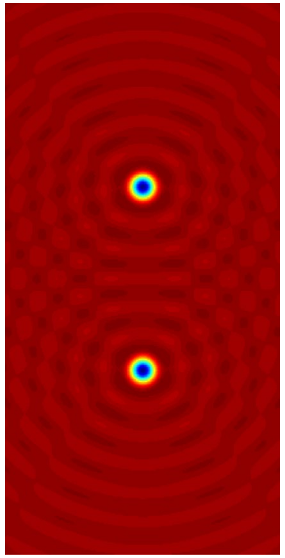
Infinite dimensional
Hilbert space

Mapping between functions
With discrete structure



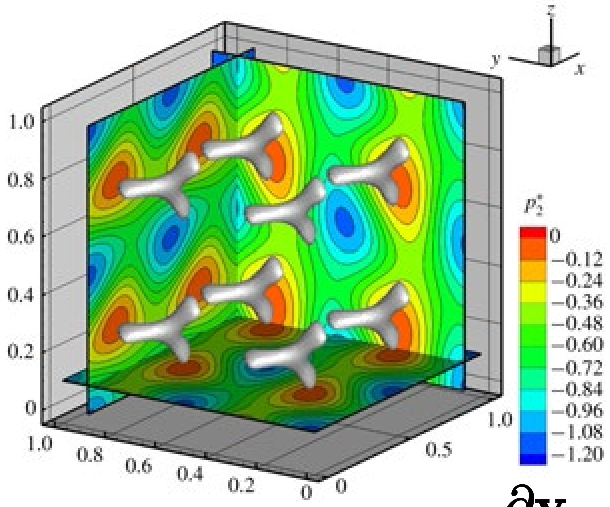
$$\hat{\mathrm{H}}|\Psi\rangle = E|\Psi\rangle$$

Schrödinger equation
Electron density / atom



$$\left(\nabla^2 + \frac{\omega^2}{c^2}\right)\Psi(\mathbf{r}, \omega) = 0$$

Helmholtz equation
Wave / source



$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} = -\frac{1}{\rho}\nabla p + \nu \Delta \mathbf{v} + \mathbf{f}(\boldsymbol{x}, t)$$

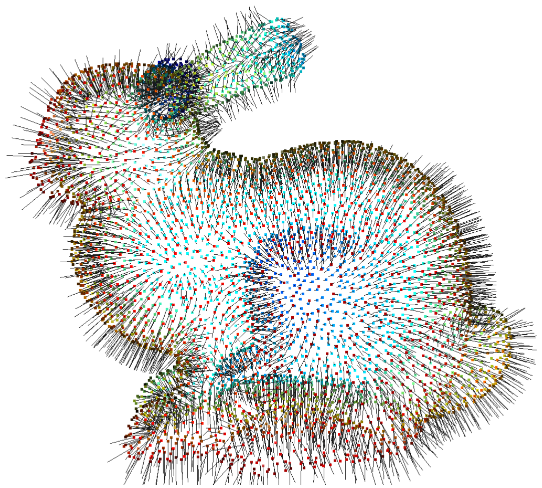
Navier–Stokes equations
Fluid / source, sink

sources: https://en.wikipedia.org/wiki/Helmholtz_equation
Tri-periodic fully three-dimensional analytic solutions for the Navier–Stokes equations. *J. Fluid Mech., 2020*

Background

Equivariance

- Rigid transformation (*rotation and translation*) of the input gives results that transform accordingly.



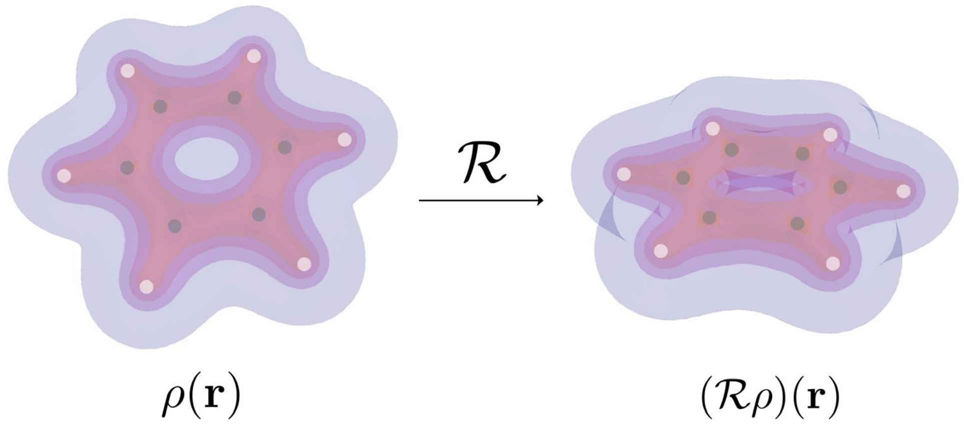
Discrete case
normal estimation

- Rotation of *continuous functions*:

$$(\mathcal{R}f)(\mathbf{x}) := f\left(R^{-1}\mathbf{x}\right)$$

- Equivariant operator:

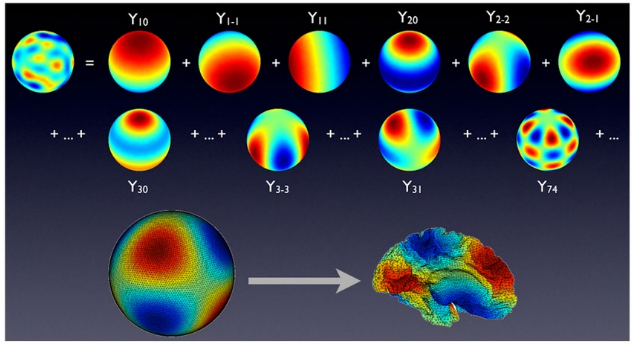
$$\mathcal{T}(\mathcal{R}f) = \mathcal{R}(\mathcal{T}f), \forall \mathcal{R}$$



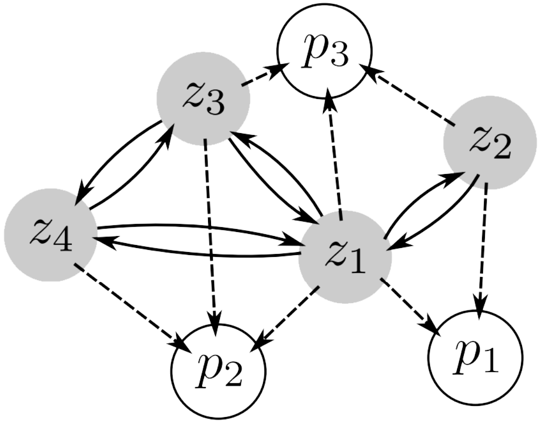
sources: Tensor field networks: Rotation- and translation-equivariant neural networks for 3D point clouds, *Arxiv*, 18

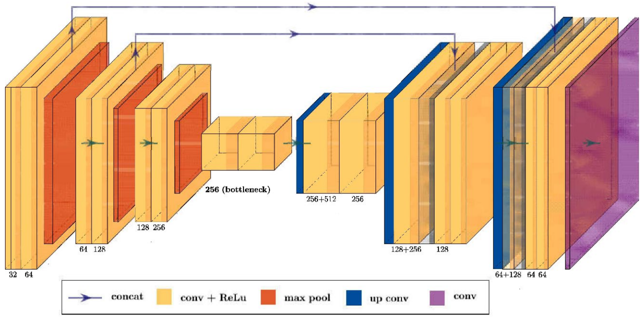
Previous Work

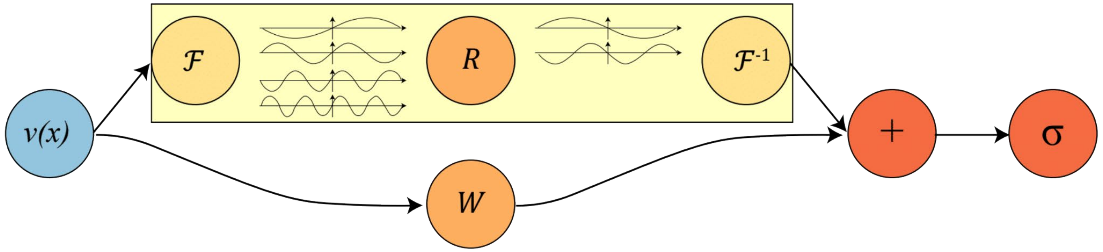
- Traditional quantum chemical calculations are either *slow* (*ab initio* method) or *inaccurate* (KS-DFT).



- *Voxel-based* regression: Memory consuming
- *Coefficient learning*: Finite approximation error
- *Interpolation network*: Lack of long-range interaction
- *Neural Operators*: Hard to scale up







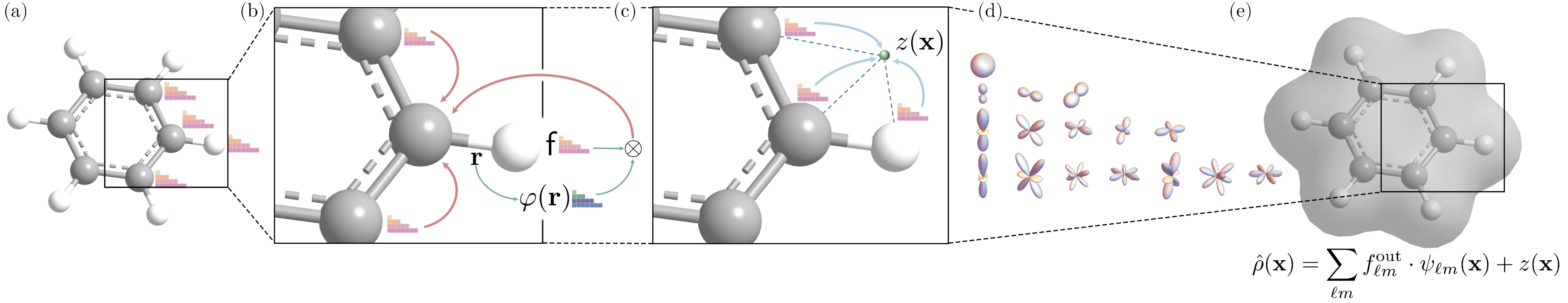
sources: DeepDFT: Neural Message Passing Network for Accurate Charge Density Prediction, *NeurIPS, 2020*
Harmony: EEG/MEG Linear Inverse Source Reconstruction in the Anatomical Basis of Spherical Harmonics, *PLoS ONE, 2012*
Whole Heart Segmentation from CT images Using 3D U-Net architecture, *IWSSIP, 2019*
Fourier neural operator for parametric partial differential equations. *ICLR, 2021*

Our Proposed **InfGCN**

- A combination of coefficient learning nets and interpolation nets
- Equivariance guaranteed *by design*
- Graph spectral theoretical interpretation as *graphon convolution*

InfGCN *layer*

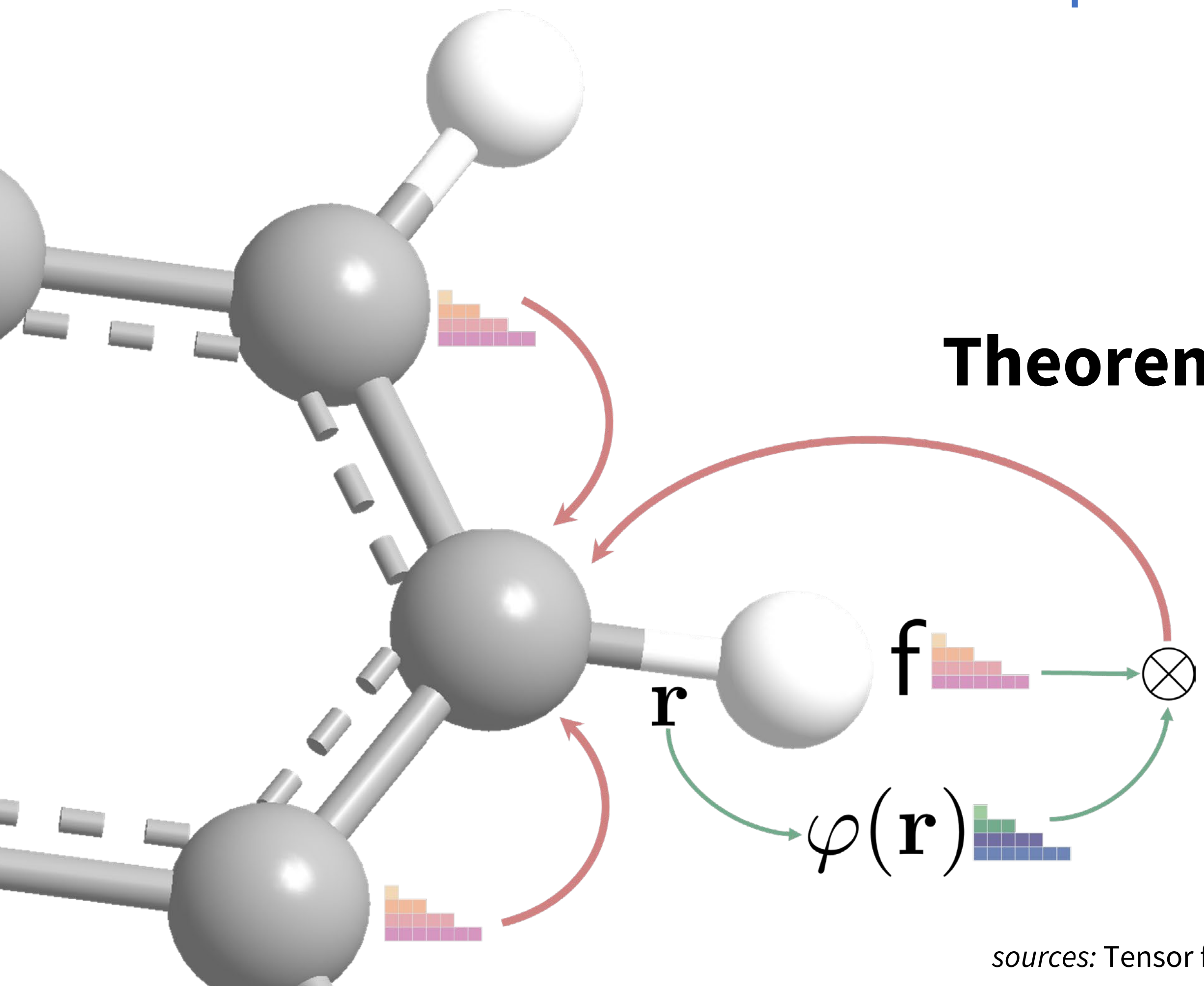
Residual Operator layer



InfGCN Layer

- Use tensor product to achieve equivariance (like in TFN)!

$$\mathbf{x}_u \leftarrow \sum_{v\in\mathcal{N}_u} \sum_{\ell k} w_{\ell k} \sum_{Jm} \mathbf{r}_{uv} \otimes \mathbf{x}_u$$



Theorem. *When interpreted as coefficients for Gaussian-type orbitals (GTOs):*

$$\psi_{n\ell m}(\mathbf{r}) = R_n^\ell(r) Y_\ell^m(\hat{\mathbf{r}}) = c_{n\ell} \exp\left(-a_n r^2\right) r^\ell Y_\ell^m(\hat{\mathbf{r}})$$

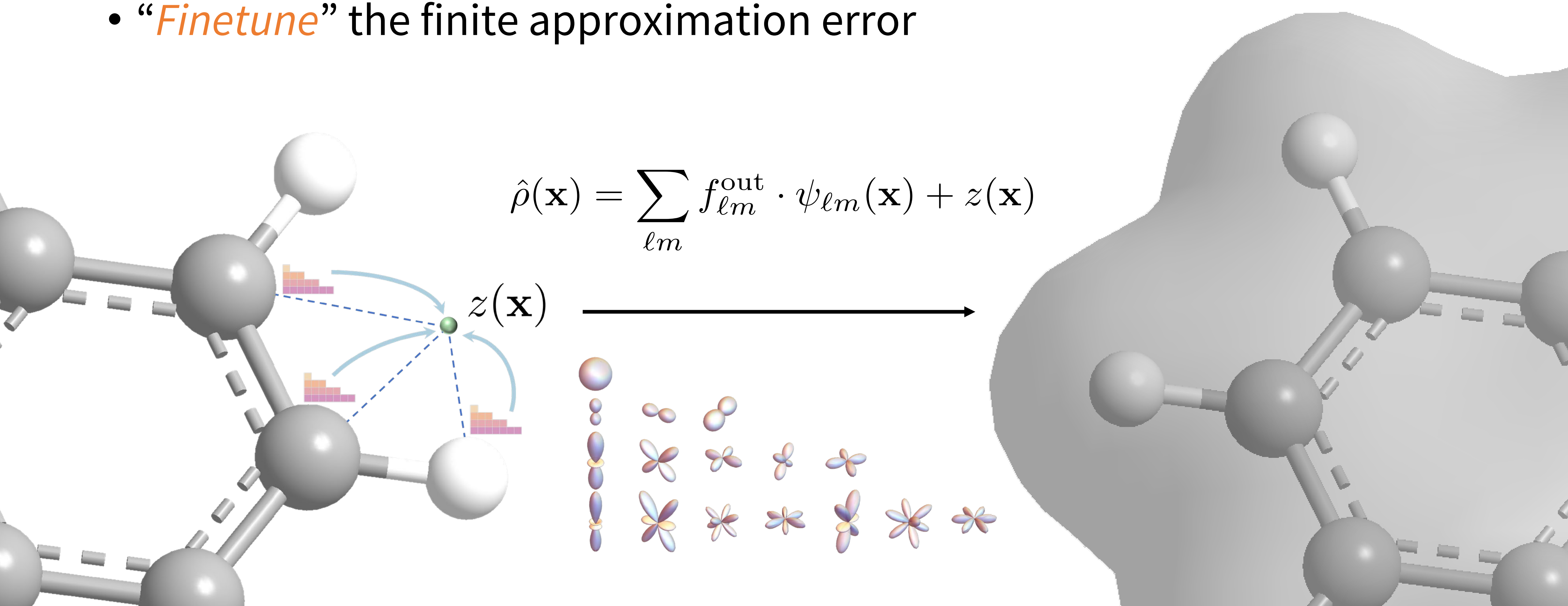
TFN will give an SE(3)-equivariant continuous function.

sources: Tensor field networks: Rotation- and translation-equivariant neural networks for 3D point clouds, *Arxiv, 18*

Residual Operator Layer

- An interpolation-style residual operator layer
- “*Finetune*” the finite approximation error

$$\hat{\rho}(\mathbf{x}) = \sum_{\ell m} f_{\ell m}^{\text{out}} \cdot \psi_{\ell m}(\mathbf{x}) + z(\mathbf{x})$$



Graph Spectral Theory

- Continuous MPNN: *Infinitely-many, continuously-indexable nodes*

$$\mathcal{T}_W f(\mathbf{x}) := \int_{\mathcal{D}} W(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$$

Proposition:

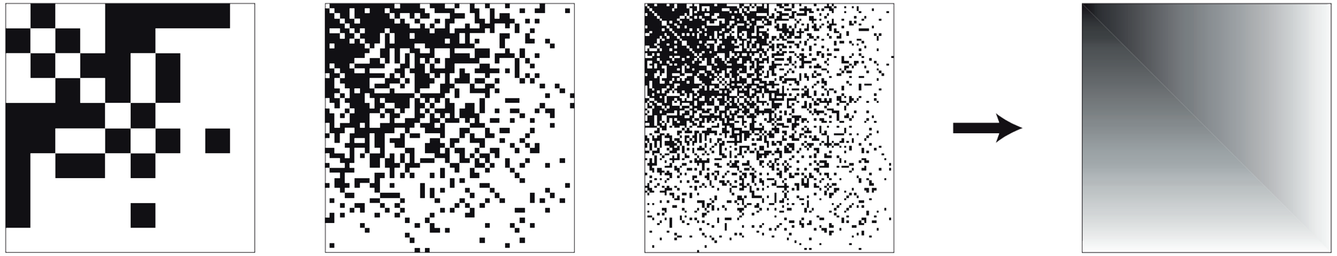
- Basis coefficients are *graphon spectra*.
- Transformation on the coefficients is *graphon convolution*.

- *Graphon* (graph limit, graph function) *slightly generalized*

Symmetric, square-integrable

$$W : \mathcal{D} \times \mathcal{D} \to [0, 1],$$

$$\int_{\mathcal{D}^2} |W(\mathbf{x}, \mathbf{y})|^2 d\mathbf{x} d\mathbf{y} < \infty$$



sources: http://www.ams.org/notices/201501/rnoti-p46.pdf

Graph Spectral Theory

UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

Continuous Version

- *Self-adjoint operator!*

$$\mathcal{T}_W f(\mathbf{x}) := \int_{\mathcal{D}} W(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}$$

$$\mathcal{U}\mathcal{T}_W = \Lambda\mathcal{U}$$

Spectral Theorem for self-adjoint operators

Unitary operator ($\mathcal{U}$)

Multiplicative operator (Λ)

- Power series

$$\mathcal{T}_W^n f(\mathbf{x}) = \mathcal{T}_W \mathcal{T}_W^{n-1} f(\mathbf{x}) = \int_{\mathcal{D}} W(\mathbf{x}, \mathbf{y}) \mathcal{T}_W^{n-1} f(\mathbf{y}) d\mathbf{y}, \mathcal{T}_W^0 = \mathcal{I}$$

- *Graphon Convolution*

$$\mathcal{H}f = \sum_{k=0}^{\infty} w_k \mathcal{T}_W^k f \approx \theta_1 f + \theta_2 \mathcal{T}_W f$$

Results

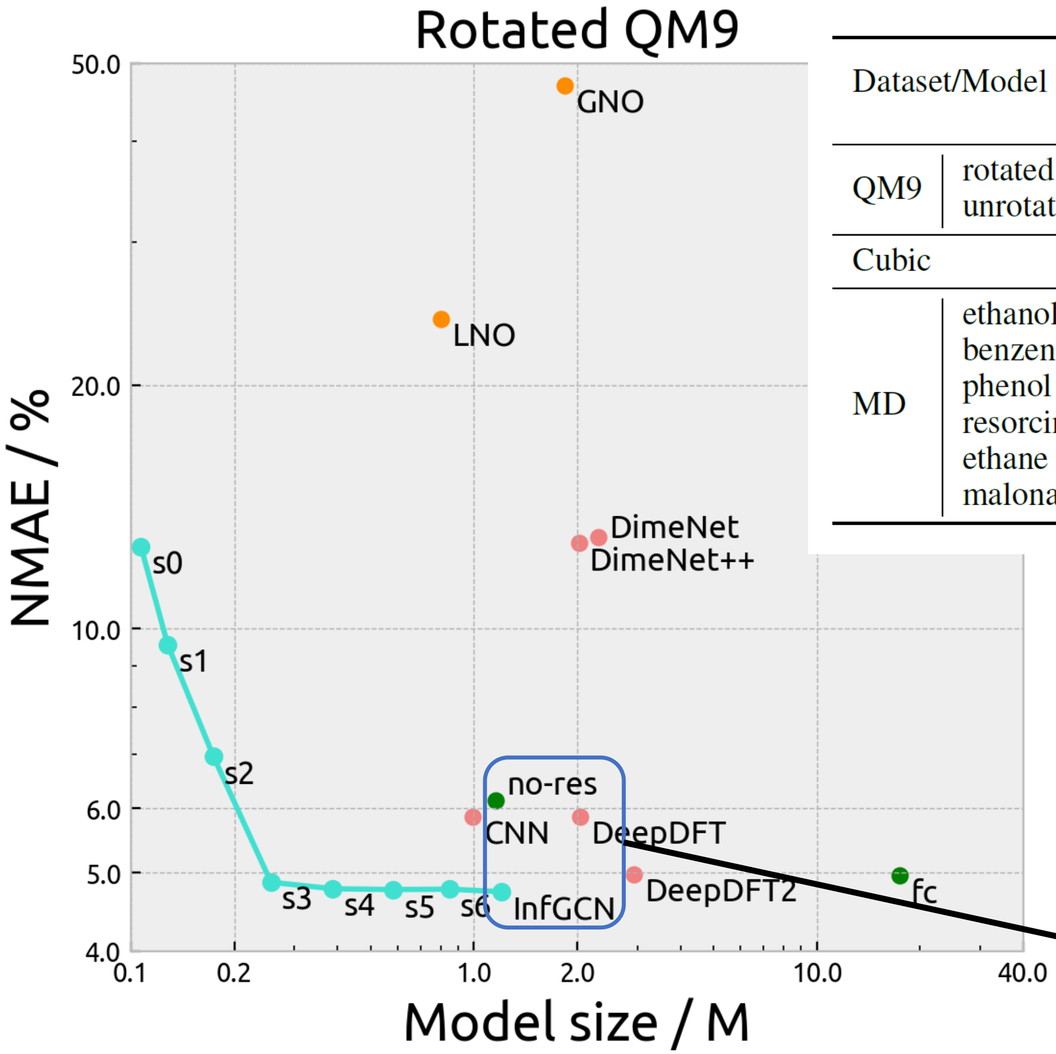


Table 1: NMAE (%) on QM9, Cubic, and MD datasets.

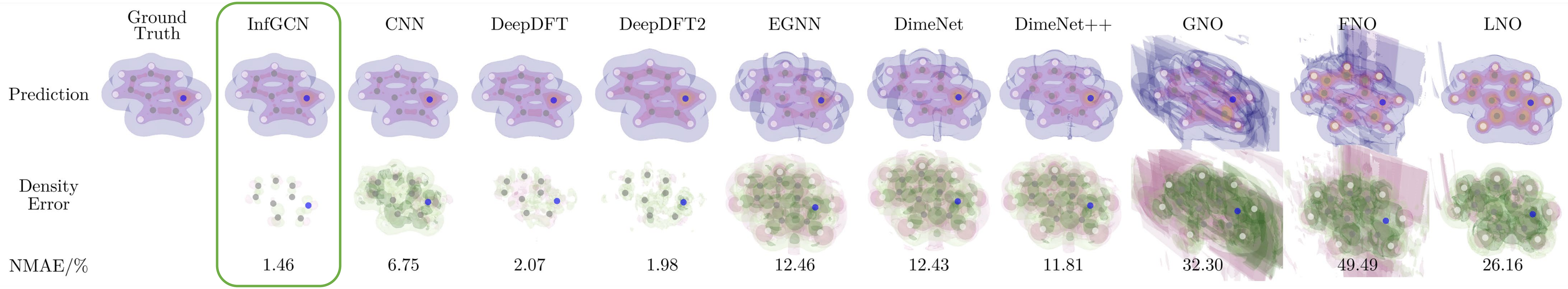
Dataset/Model		**InfGCN**	**CNN**	Interpolation Net				Neural Operator		
				DeepDFT	**DeepDFT2**	**DimeNet**	**DimeNet++**	**GNO**	**FNO**	**LNO**
QM9	rotated	**4.73**	5.89	5.87	4.98	12.98	12.75	46.90	33.25	24.13
	unrotated	**0.93**	2.01	2.95	1.03	11.97	11.69	40.86	28.83	26.14
Cubic		**8.98**	OOM	14.08	10.37	12.51	12.18	53.55	48.08	46.33
MD	ethanol	8.43	13.97	**7.34**	8.83	13.99	14.24	82.35	31.98	43.17
	benzene	**5.11**	11.98	6.61	5.49	14.48	14.34	82.46	20.05	38.82
	phenol	**5.51**	11.52	9.09	7.00	12.93	12.99	66.69	42.98	60.70
	resorcinol	**5.95**	11.07	8.18	6.95	12.04	12.01	58.75	26.06	35.07
	ethane	7.01	14.72	8.31	**6.36**	13.11	12.95	71.12	26.31	77.14
	malonaldehyde	**10.34**	18.52	9.31	10.68	18.71	16.79	84.52	34.58	47.22

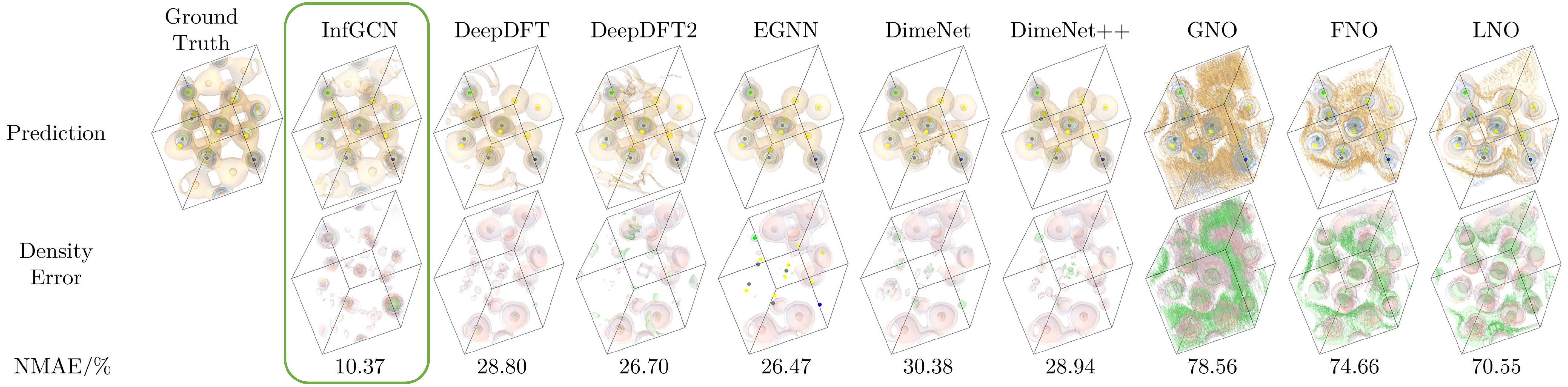
Table 2: NMAE (%) and the parameter count of different model settings on the QM9 dataset.

Model	InfGCN(s_7)	s_6	s_5	s_4	s_3	s_2	s_1	s_0	no-res	fc
QM-rot (%)	**4.73**	4.77	4.76	4.77	4.86	6.95	9.56	12.62	6.14	4.95
QM-unrot (%)	**0.93**	1.01	1.11	1.08	1.46	4.65	8.07	12.05	3.72	1.36
Parameters (M)	1.20	0.85	0.58	0.39	0.26	0.17	0.13	0.11	1.16	17.42

Residual operator layer is important!

Results


Ground Truth
InfGCN
CNN
DeepDFT
DeepDFT2
EGNN
DimeNet
DimeNet++
GNO
FNO
LNO
Prediction
Density Error
NMAE/%
1.46
6.75
2.07
1.98
12.46
12.43
11.81
32.30
49.49
26.16


Ground Truth
InfGCN
DeepDFT
DeepDFT2
EGNN
DimeNet
DimeNet++
GNO
FNO
LNO
Prediction
Density Error
NMAE/%
10.37
28.80
26.70
26.47
30.38
28.94
78.56
74.66
70.55

Thanks

Chaoran Cheng
chaoran7@illinois.edu
https://ccr-cheng.github.io/

UNIVERSITY OF
ILLINOIS
URBANA-CHAMPAIGN