Stony Brook University

Artificial Intelligence for Science (AI4Sci) in Atomistic and Continuum Systems

Presenter: Wenhan Gao

Advisor: Yi Liu

Department of Applied Mathematics and Statistics

FAR BEYOND

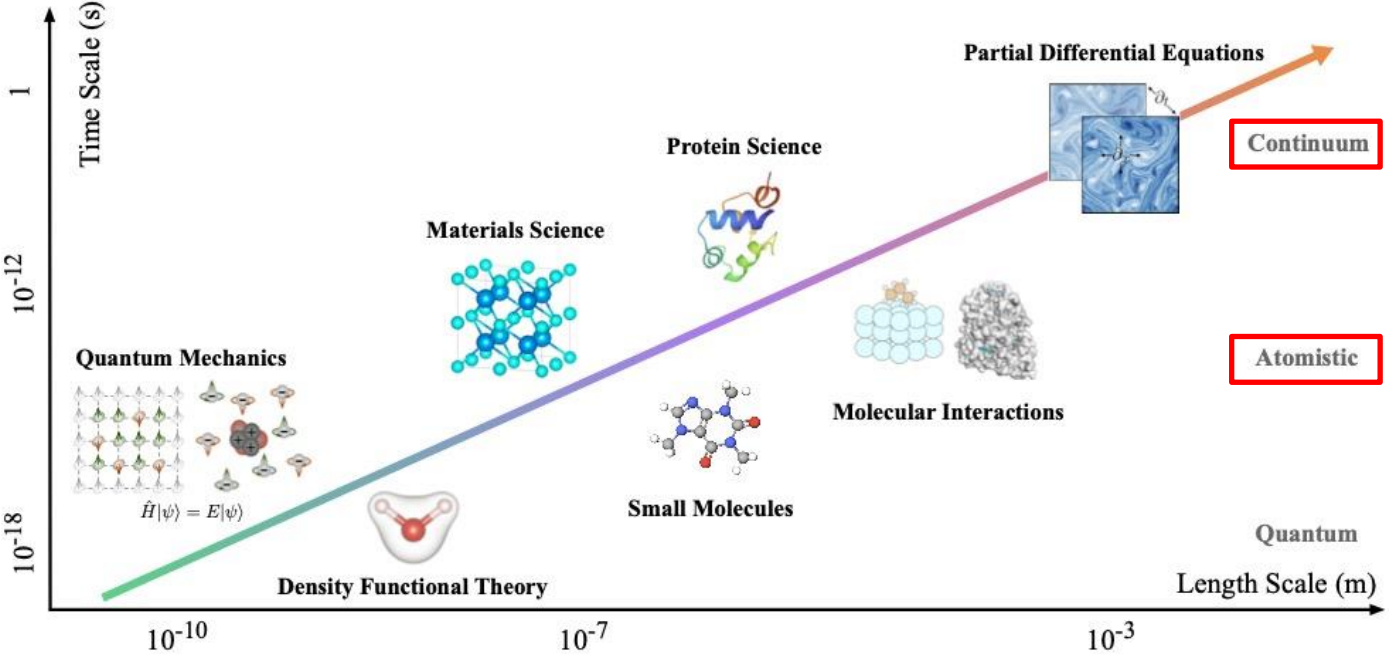
Stony Brook University

About Me

I'm a third year Ph.D. student in Applied Mathematics at Stony Brook University supervised by Professor Yi Liu.
Previously, I obtained a Bachelor of Science degree from Stony Brook University in both applied mathematics 📙 and pure mathematics 📕.
I'm interested in Neural Operators, Equivariant Neural Networks, Generative (Probabilistic) Models, and AI for Science in general; I'm also broadly interested in Machine Learning, discrete math, and many other topics in applied mathematics and computer science.

Homepage: https://wenhangao21.github.io/

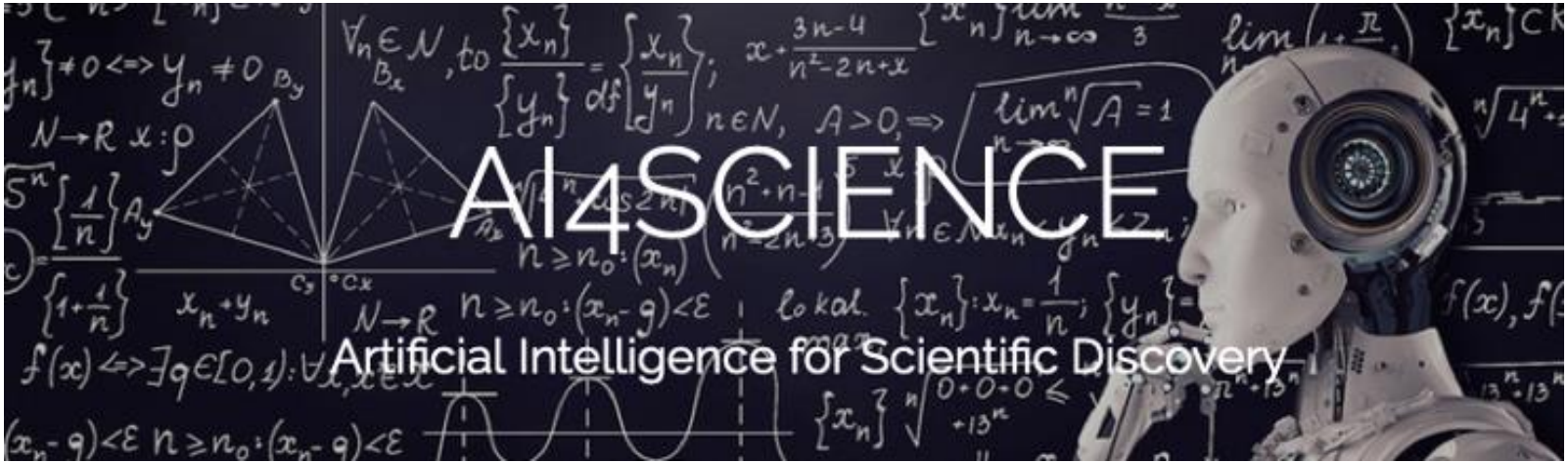
Science at Different Scales



Stony Brook University

AI for Science (AI4Sci)

AI4Sci refers to the use of recent advances in artificial intelligence and deep learning to solve problems in natural sciences: computational chemistry, PDEs, material science, drug design, etc..



FAR BEYOND

Stony Brook University

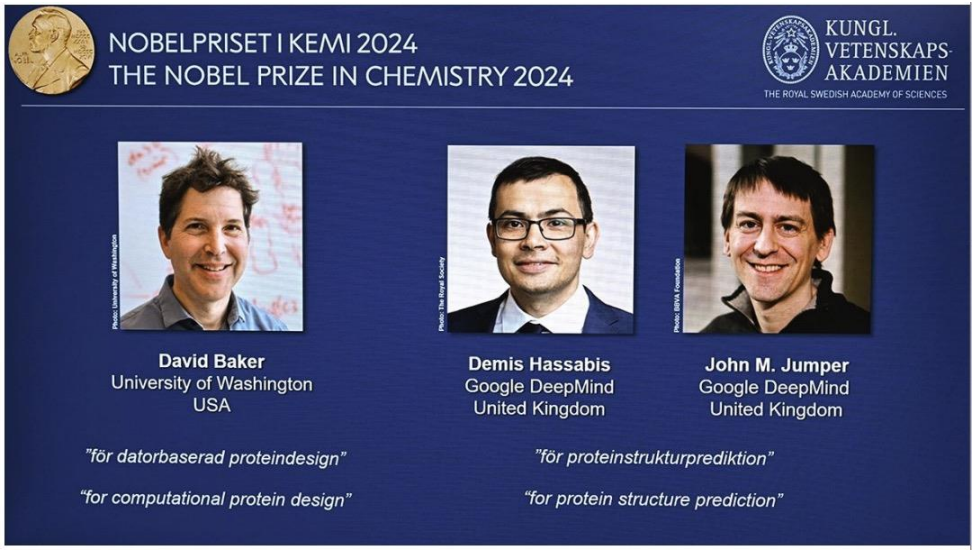
Best Time for AI4Sci

Nobel Prize in Physics → AI

AI pioneers John J. Hopfield and Geoffrey E. Hinton for their contribution to AI and ML

FAR BEYOND

Best Time for AI4Sci

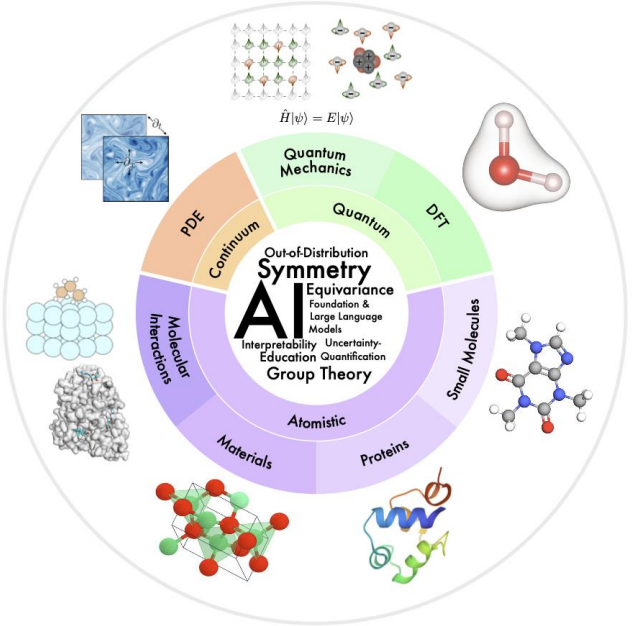


Nobel Prize in Chemistry → AI4Science

Demis Hassabis and John Jumper for their contribution to AlphaFold - protein structure prediction with AI

Stony Brook University

My Work: AI for Science (AI4Sci)



Key mission:

- **Integrate domain knowledge in science (such as symmetry) into AI models.**
- **Innovations in both AI and science.**
- **Analyzing existing AI models for scientific tasks.**

Other perspectives:

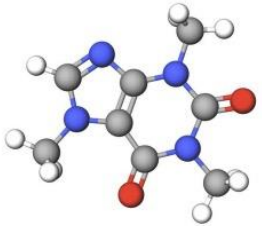
- Explainability
- Large language models

FAR BEYOND

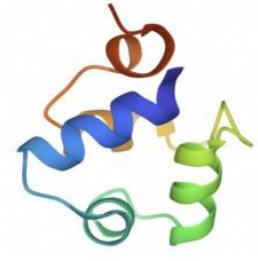
Symmetries in Physics

"It is only slightly overstating the case to say that physics is the study of symmetry."

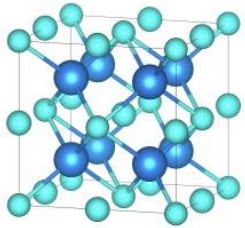
—Philip W. Anderson (1972)



Molecule



Protein



Material

FYI: Dr. Chen Ning Yang received the Nobel Prize in physics (1957) for discoveries about symmetries, and his B.S. thesis is "Group Theory and Molecular Spectra".

Stony Brook University

Equivariance and Invariance

Equivariance is a property of an operator $\Phi : X \to Y$ (such as a neural network layer) by which it commutes with the group action:

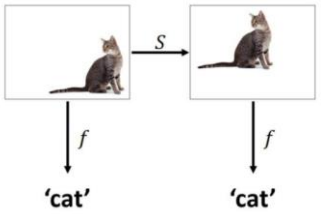
$$\Phi \circ \rho^X(g) = \rho^Y(g) \circ \Phi,$$

Invariance is a property of an operator $\Phi : X \to Y$ (such as a neural network layer) by which it remains unchanged after the group action:

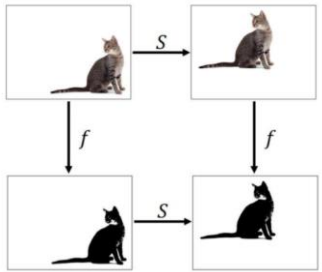
$$\Phi \circ \rho^X(g) = \Phi,$$

- $\rho^X(g)$: group representation action on X
- $\rho^Y(g)$: group representation action on Y
- Invariance is a special case of equivariance when $\rho^Y(g)$ is the identity.

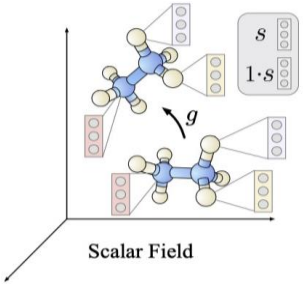
Invariance



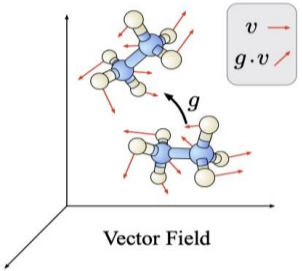
Equivariance



Invariance



Equivariance



FAR BEYOND

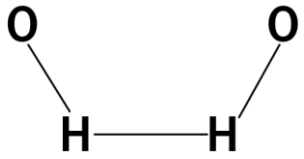
Stony Brook University

Talk Outline

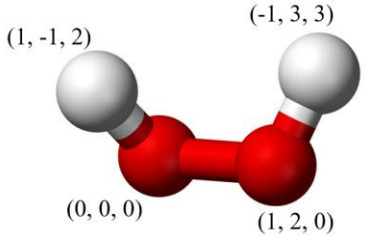
- **Atomistic Systems**
  - Geometric Representation and Learning of Atomistic Systems
  - The Trouble with ML and Symmetries
  - Imperfect Scientific Data
  - Explainability: Does AI Model Learn Science?
- **Continuum Systems**
  - Continuum Systems and PDEs
  - Physics-informed Neural Networks
  - Operator Learning
  - Symmetries in Continuum Systems
  - Paradox in De-biasing from Discretization
- **Future Works**
  - Robustness of Equivariant GNNs
  - Unconstrained Molecular Generation

Geometric Representation of Atomistic Systems

Representation of Atomistic Systems:



Molecules as 2D (planar) graphs



Molecules as 3D (geometric) graphs

- 3D geometric configuration (coordinates) is crucial in determining properties.
- 3D representations outperforms their 2D counterparts by a large margin.

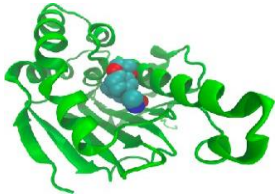
Model	MAE (Lower, better)	
GIN-Virtual	0.2371	→ Best 2D GNN
SchNet	0.1511	3D GNNs outperform 2D GNNs by a large margin
DimeNet++	0.1214	
SphereNet	0.1182	
ComENet	0.1273	

Stony Brook University

Geometric Graph Neural Networks

- A **Fundamental tool** for machine learning on geometric (3D) graphs for atomistic systems.

Geometric Graph → Geometric GNN → Prediction

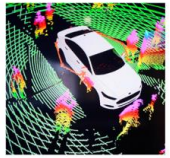


- Functional properties?
- Ligand binding affinity?
- Ligand efficacy?

- There are **many other systems with geometric & relational structures** beyond atomistic systems!

Transportation & Logistics

Robotic Navigation



3D Computer Vision

FAR BEYOND

Image Source: A Hitchhiker's Guide to Geometric GNNs for 3D Atomic Systems

Stony Brook University

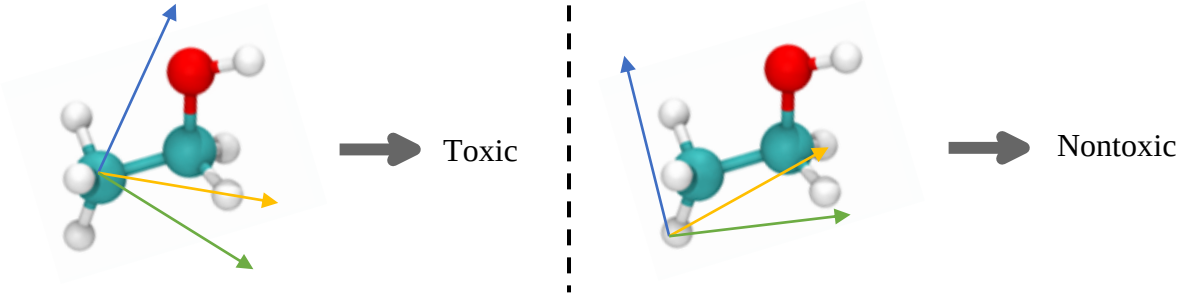
Talk Outline

- **Atomistic Systems**
  - Geometric Representation and Learning of Atomic Systems
  - The Trouble with ML and Symmetries
  - Imperfect Scientific Data
  - Explainability: Does AI Model Learn Science?
- **Continuum Systems**
  - Continuum Systems and PDEs
  - Physics-informed Neural Networks
  - Operator Learning
  - Symmetries in Continuum Systems
  - Paradox in De-biasing from Discretization
- **Future Works**
  - Robustness of Equivariant GNNs
  - Unconstrained Molecular Generation

Stony Brook University

The Trouble with ML and Coordinates

- The nature does not use coordinate systems. Coordinate systems have heavy human bias!
- Us humans can use different coordinates to describe the same object.
- Predictions can be completely different!



FAR BEYOND

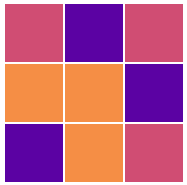
Stony Brook University

The Trouble with ML and Coordinates

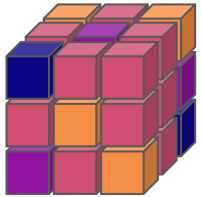
- Many physical properties contain directional information!

3 vector
forces
displacements
positions

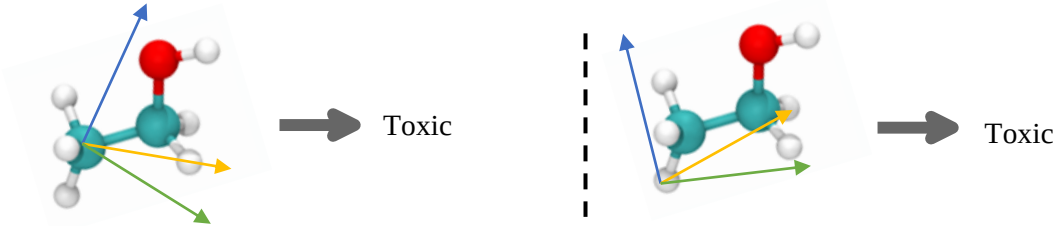
3x3 matrix
moment of inertia
stress
strain



3x3x3 tensor
Levi-Civita tensort



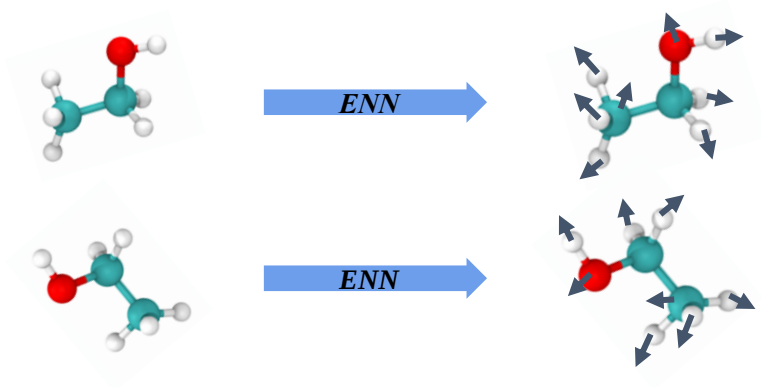
- The model **must spend valuable model capacities to learn to de-bias** (to learn symmetries) for reliable predictions!



FAR BEYOND

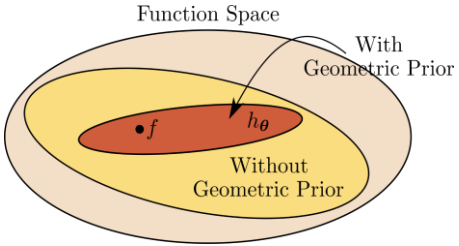
Equivariant Neural Networks

A better solution: **Equivariant Neural Networks (ENNs)!** ENNs guarantee symmetries by design!



Generalization and Efficient Learning:

Geometric priors → smaller search space → less parameters
less data
better generalization

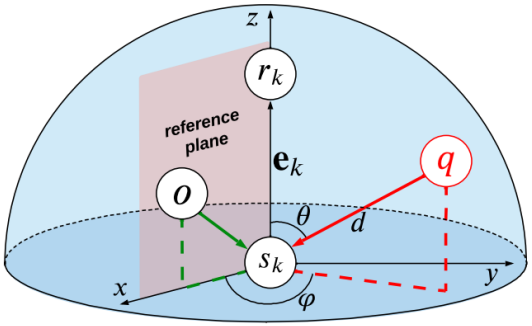


Stony Brook University

Scalarization and Spherical Tensor Networks

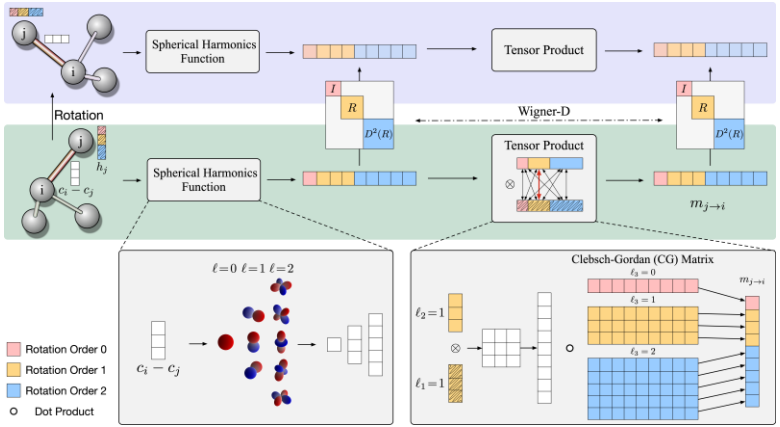
Scalarization Networks

- ❑ Use only invariant quantities as inputs.
- ❑ Can only be used for invariant tasks.



Spherical Tensor Networks

- ❑ Use directional quantities as features.
- ❑ Must maintain directional information throughout.



FAR BEYOND

Stony Brook University

Talk Outline

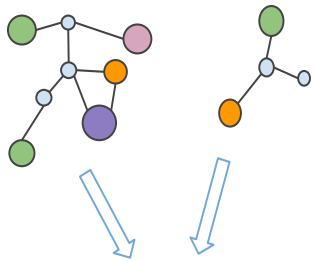
- **Atomistic Systems**
  - Geometric Representation and Learning of Atomic Systems
  - Symmetries in Atomic Systems
  - Imperfect Scientific Data
  - Explainability: Does AI Model Learn Science?
- **Continuum Systems**
  - Continuum Systems and PDEs
  - Physics-informed Neural Networks
  - Operator Learning
  - Symmetries in Continuum Systems
  - Paradox in De-biasing from Discretization
- **Future Works**
  - Robustness of Equivariant GNNs
  - Unconstrained Molecular Generation

Stony Brook University

Imperfect Scientific Data

Motivation

❑ Annotating scientific data is difficult!

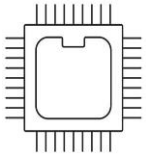


Properties ?

Obtaining Annotations

Experts

High Cost



Extensive Computation

Ronast Subedi*, Lu Wei, **Wenhan Gao***, Shayok Chakraborty, **Yi Liu**, Empowering Active Learning for 3D Molecular Graphs with Geometric Graph Isomorphism, NeurIPS, 2024

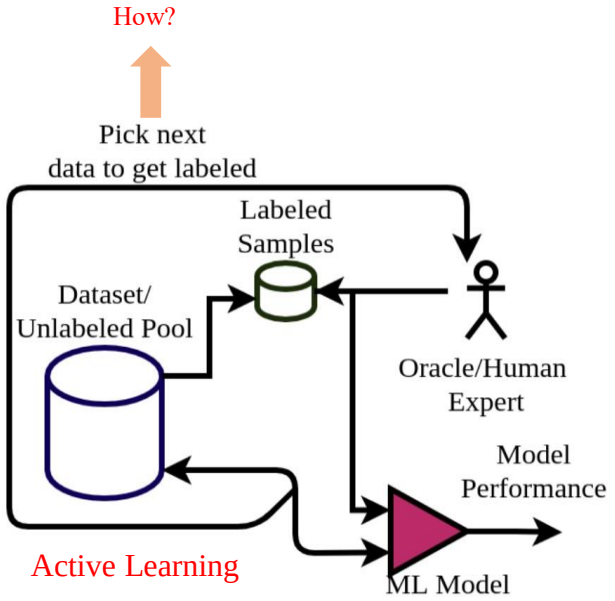
Stony Brook University

Learning with Less Data

What Can We Do?

❑ Select the most informative training samples to get labeled.

- **Diversity:** how a sample is different from other Molecules?
- **Uncertainty:** how the model is confident about a sample Molecules?



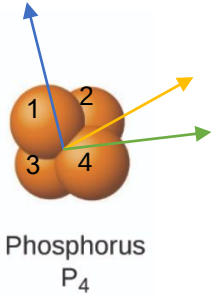
Ronast Subedi*, Lu Wei, **Wenhan Gao***, Shayok Chakraborty, **Yi Liu**, Empowering Active Learning for 3D Molecular Graphs with Geometric Graph Isomorphism, NeurIPS, 2024

Stony Brook University

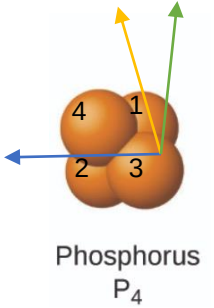
Challenges in Uncertainty Quantification

Challenges

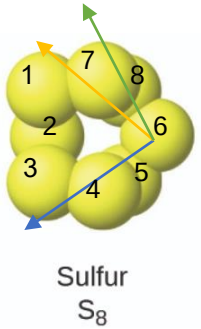
- ❑ Different coordinate systems (human bias).
- ❑ Different Atom ordering (human bias).
- ❑ Different number of atoms.
- ❑ **All existing uncertainty measures fail for geometric graphs!**



Phosphorus
P_4



Phosphorus
P_4



Sulfur
S_8

Ronast Subedi*, Lu Wei, **Wenhan Gao***, Shayok Chakraborty, **Yi Liu**, Empowering Active Learning for 3D Molecular Graphs with Geometric Graph Isomorphism, NeurIPS, 2024

Stony Brook University

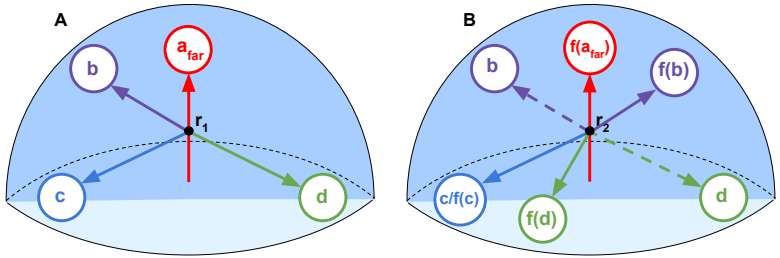
Geometric Isometries

Isometries for Expressive Geometric Representations

- We propose three isometries (< means less expressive) based on geometric priors:

  distance < triangular < cross-angle

- Used as basis for expressive representation of 3D molecular graphs.
- These isometries are E(3)-equivariant (Challenge #1).



A and *B* are triangular isometric but not cross-angular isometric

FAR BEYOND

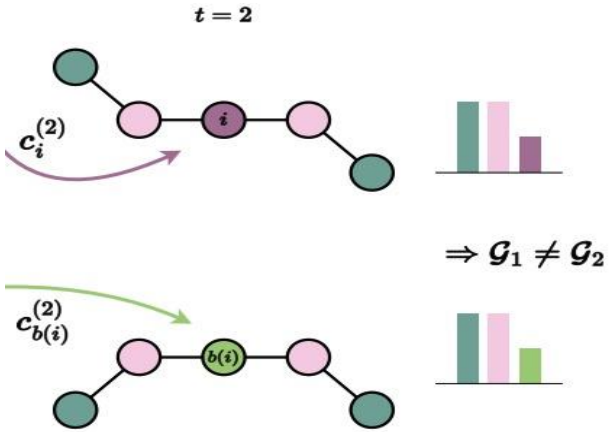
Ronast Subedi*, Lu Wei, **Wenhan Gao***, Shayok Chakraborty, **Yi Liu**, Empowering Active Learning for 3D Molecular Graphs with Geometric Graph Isomorphism, NeurIPS, 2024

Stony Brook University

Geometric Isometries

Expressiveness of Isometries

- Cross-angle isometries are at least as expressive as any geometric GNNs.
- They can distinguish any non-isometric structure a geometric GNN can distinguish.



FAR BEYOND

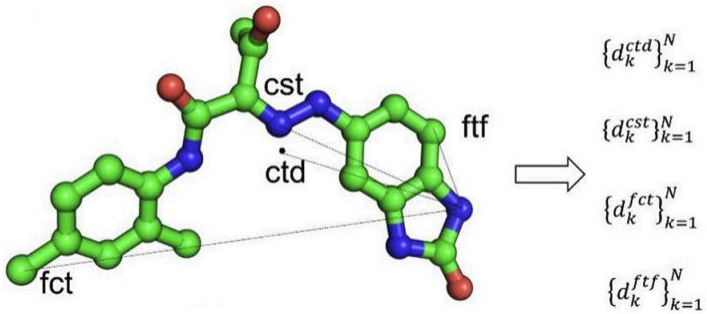
Ronast Subedi*, Lu Wei, **Wenhan Gao***, Shayok Chakraborty, **Yi Liu**, Empowering Active Learning for 3D Molecular Graphs with Geometric Graph Isomorphism, NeurIPS, 2024

Stony Brook University

Geometric Isometries

Diversity Score Computing

- Different orders of statistical moments used to encode the isometries into a geometric descriptor (Challenges #2&3).
- Obtained descriptors are E(3)-equivariant and permutation invariant!



Stony Brook University

Uncertainty Measures

Uncertainty Component

- Will incorporate other chemical properties such as atom type.

The ***variational inference (VI) loss***:

$\mathcal{L}_{\mathrm{VI}}(\theta) = -\int_{R^n} q_\theta(w) \log p_{\mathrm{3DGNN}}(O \mid G, w)\, dw + \mathrm{KL}\big(q_\theta(w) | p_{\mathrm{3DGNN}}(w)\big)$

is used to **quantify molecules that the model has limited knowledge**.

The predictive uncertainty for a new sample o^* given g^* :

$$\widehat{\sigma^2}(o^* \mid g^*) = \frac{1}{N}\sum_{n=1}^{N}\left(\widehat{o_n^*}\right)^2 - \left(\frac{1}{N}\sum_{n=1}^{N}\widehat{o_n^*}\right)^2 + \frac{1}{N}\sum_{n=1}^{N}\widehat{\sigma_n^2}$$

is used to obtain an uncertainty value (variance) for each molecule.

Ronast Subedi*, Lu Wei, **Wenhan Gao***, Shayok Chakraborty, **Yi Liu**, Empowering Active Learning for 3D Molecular Graphs with Geometric Graph Isomorphism, NeurIPS, 2024

Stony Brook University

Results and Broder Impact

Results

- Can learn with much less data: Achieves similar performance with only **5,000 samples** compared to that of **10,000 samples** without our active learning selection.

- Can generalize to any atomistic data: We demonstrate that our method works well on QM9 (stable molecules), MD17 (molecular dynamics trajectories), EC (protein) and can **generalize well to all tasks from molecules to proteins to materials!**

Ronast Subedi*, Lu Wei, **Wenhan Gao***, Shayok Chakraborty, **Yi Liu**, Empowering Active Learning for 3D Molecular Graphs with Geometric Graph Isomorphism, NeurIPS, 2024

Stony Brook University

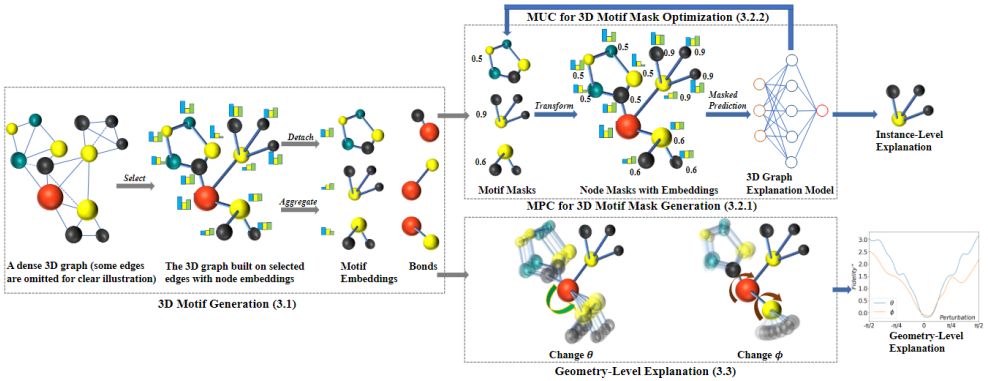
Talk Outline

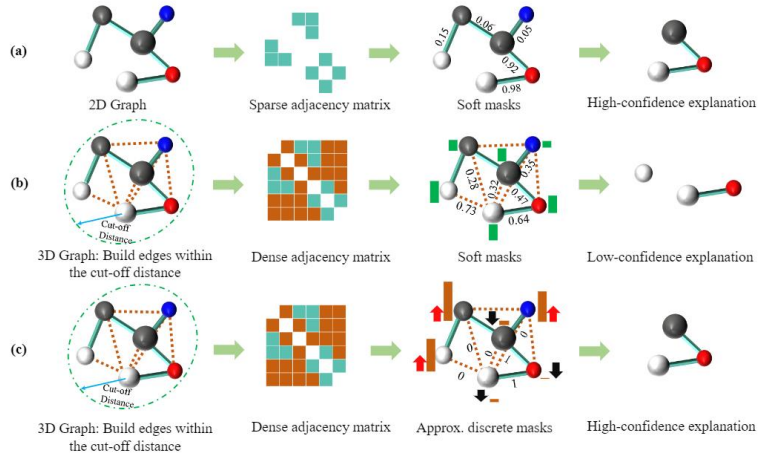
- **Atomistic Systems**
  - Geometric Representation and Learning of Atomic Systems
  - Symmetries in Atomic Systems
  - Imperfect Scientific Data
  - Explainability: Does AI Model Learn Science?
- **Continuum Systems**
  - Continuum Systems and PDEs
  - Physics-informed Neural Networks
  - Operator Learning
  - Symmetries in Continuum Systems
  - Paradox in De-biasing from Discretization
- **Future Works**
  - Robustness of Equivariant GNNs
  - Unconstrained Molecular Generation

Stony Brook University

(Molecular) Graph Explanation

- Identify key molecular substructures that influence the model's predictions.
- Allow domain experts to verify their alignment with chemical knowledge.
- There is a pressing yet often overlooked need for advanced explanation methods for 3D (geometric) GNNs.





Xufeng Liu, Dongsheng Luo, **Wenhan Gao**, **Yi Liu**, 3DGraphX: Explaining 3D Molecular Graph Models via Incorporating Chemical Priors, KDD 2025

Xufeng Liu*, **Wenhan Gao***, **Yi Liu**, Energy-Based Discrete Mask Approximation for 3D Molecular Graph Explanation, Under Review ICLR 2025

FAR BEYOND

Stony Brook University

Talk Outline

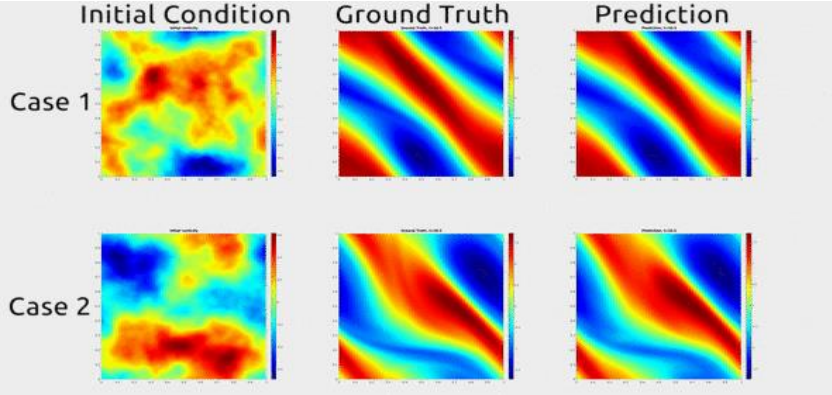
- **Atomistic Systems**
  - Geometric Representation and Learning of Atomic Systems
  - Symmetries in Atomic Systems
  - Imperfect Scientific Data
  - Explainability: Does AI Model Learn Science?
- **Continuum Systems**
  - Continuum Systems and PDEs
  - Physics-informed Neural Networks
  - Operator Learning
  - Symmetries in Continuum Systems
  - Paradox in De-biasing from Discretization
- **Future Works**
  - Robustness of Equivariant GNNs
  - Unconstrained Molecular Generation

Stony Brook University

AI for Continuum Systems: Neural PDE Solvers

- Continuum Systems: We assume a substance can be divided into ever smaller and smaller bits, so it always looks continuous!
- A PDE mathematically describes the behavior of continuum systems by prescribing constraints relating partial derivatives.

Example: The Navier-Stokes Equation



FAR BEYOND

Source for Navier-Stokes Animation: https://zongyi-li.github.io/blog/2020/fourier-pde/

Traditional PDE Solvers

- **Solve one instance**
- **Require the explicit form**
- Speed-accuracy trade-off on resolution
- **Slow** on fine grids; fast on coarse grids
- Suffers from the Curse of Dimensionality (CoD)

PINN

- **Solve one instance**
- Incorporate known physics
- **Can train without data**
- Mesh-free
- Can be **slow to train**
- Lessen the CoD issue

Neural Operators

- **Learn a family of PDEs**
- Compatible with physics-informed ideas
- Black-box, data-driven
- Can be resolution-invariant
- Slow to train; **fast to evaluate** (can be several orders of magnitudes faster)
- Many neural operators suffer from CoD

Stony Brook University

Talk Outline

- **Atomistic Systems**
  - Geometric Representation and Learning of Atomic Systems
  - Symmetries in Atomic Systems
  - Imperfect Scientific Data
  - Explainability: Does AI Model Learn Science?
- **Continuum Systems**
  - Continuum Systems and PDEs
  - Physics-informed Neural Networks
  - Operator Learning
  - Symmetries in Continuum Systems
  - Paradox in De-biasing from Discretization
- **Future Works**
  - Robustness of Equivariant GNNs
  - Unconstrained Molecular Generation

PINNs (Solving One Instance)

Consider the following general form of a PDE for $u(\boldsymbol{x})$:

$$\begin{cases} \mathcal{D}u(\boldsymbol{x}) = f(\boldsymbol{x}), & \text{in } \Omega, \\ \mathcal{B}u(\boldsymbol{x}) = g(\boldsymbol{x}), & \text{on } \partial\Omega, \end{cases}$$

we wish to approximate $u(\boldsymbol{x})$ with a neural network, denoted by $\phi(\boldsymbol{x};\boldsymbol{\theta})$. We can train the neural network with physics-informed loss. That is, we aim to solve the following optimization problem:

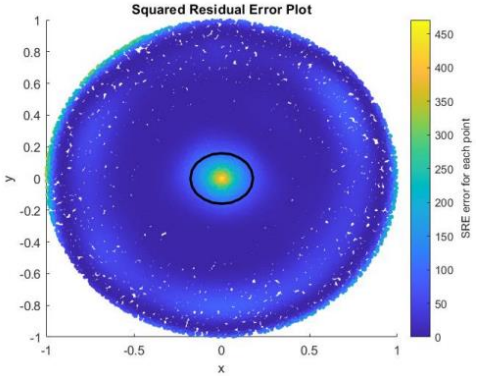
$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}) := \arg\min_{\boldsymbol{\theta}} \underbrace{\|\mathcal{D}\phi(\boldsymbol{x};\boldsymbol{\theta}) - f(\boldsymbol{x})\|_2^2}_{\text{Difference between the L.H.S. and R.H.S. of the differential equation.}} + \lambda\|\mathcal{B}\phi(\boldsymbol{x};\boldsymbol{\theta}) - g(\boldsymbol{x})\|_2^2$$

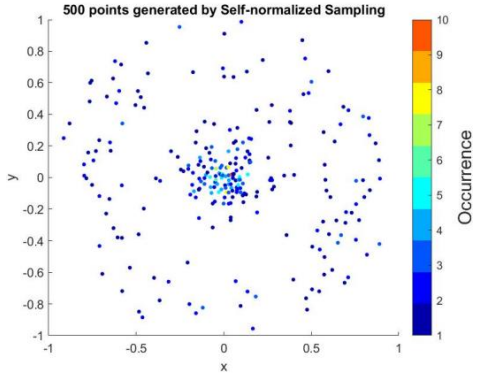
Intuition: We penalize the neural network by the extend to which it violates the PDE/boundary/initial conditions.

Stony Brook University

Active Learning for High-Dimensional PDEs

- PINN still faces significant CoD issues in high-dimensions.
- Proposed an active learning approach based on an unnormalized distribution informed by physics to adaptively sample training collocation points to lessen CoD issues.
- Designed a parallelizable self-normalized algorithm (efficiently run on GPUs) to simulate the proposed distribution.





FAR BEYOND

Wenhan Gao, Chunmei Wang, Active Learning Based Sampling for High-Dimensional Nonlinear Partial Differential Equations, The Journal of Computational Physics (JCP) 2023

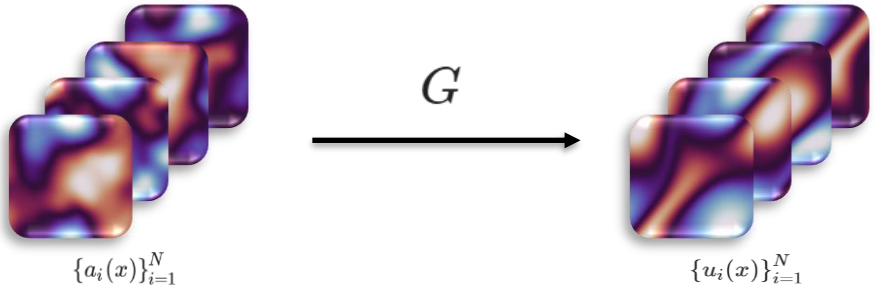
Stony Brook University

Talk Outline

- **Atomistic Systems**
  - Geometric Representation and Learning of Atomic Systems
  - Symmetries in Atomic Systems
  - Imperfect Scientific Data
  - Explainability: Does AI Model Learn Science?
- **Continuum Systems**
  - Continuum Systems and PDEs
  - Physics-informed Neural Networks
  - Operator Learning
  - Symmetries in Continuum Systems
  - Paradox in De-biasing from Discretization
- **Future Works**
  - Robustness of Equivariant GNNs
  - Unconstrained Molecular Generation

Neural Operators (Solving a Family of PDEs)

- In numerous fields, we seek to **study the behavior of physical systems under various parameters.**
- Neural operators approximate the **mapping from parameter function space to solution function space.**
- Once trained, obtaining a solution can be **several orders of magnitude faster than numerical methods.**

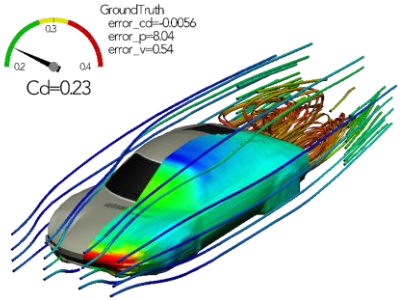


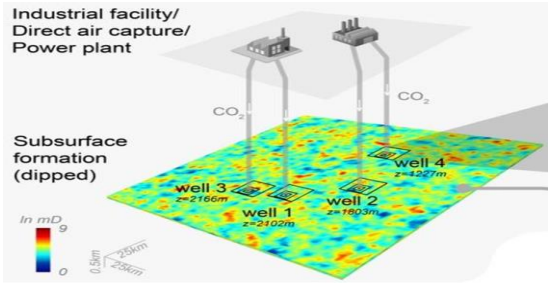
Operator Learning Essentials

- Domain
  - $\Omega \subset \mathbb{R}^d$ be a bounded open set of spatial coordinates
- Input and output function spaces on Ω (e.g., Banach spaces, Hilbert spaces)
  - $\mathcal{A}$ and $\mathcal{U}$
- Ground truth solution operator
  - $G : \mathcal{A} \to \mathcal{U}$ with $G(a) = u$
- Training data
  - Observed (possibly noisy) function pairs $(a_i, u_i) \in \mathcal{A} \times \mathcal{U}, u_i = G(a_i)$ with measures $a_i \sim \nu_a, u_i \sim \nu_u$, where ν_u is the pushforward measure of ν_a by G
- Task: Learn operators from data
  - $G_\theta(a) \approx u$

Stony Brook University

Applications of Operator Learning





Computational Fluid Dynamics [a]

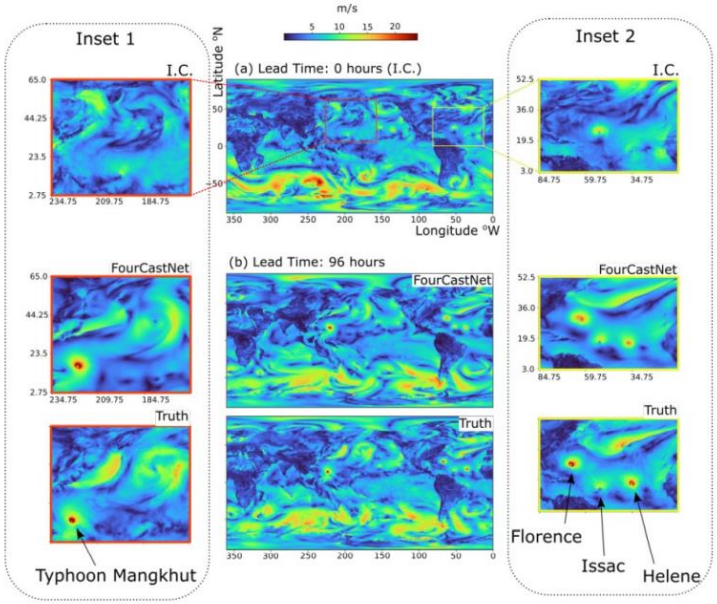
Carbon Storage Modeling [b]

Weather Modeling [c]

[a] Geometry-informed neural operator for large-scale 3D PDEs. arXiv, 2023. Zongyi Li, et al..
[b] Fourier-MIONet: Fourier-enhanced multiple-input neural operators for multiphase modeling of geological carbon sequestration. arXiv, 2023. Zhongyi Jiang, et al..
[c] DeepPhysiNet: Bridging deep learning and atmospheric physics for accurate and continuous weather modeling. arXiv, 2024. Wenyuan Li, et al..

Stony Brook University

Weather Forecasting: FourCastNet



Pathak et al. (2022), FourCastNet: A Global Data-driven High-resolution Weather Model using Adaptive Fourier Neural Operators, arXiv: https://arxiv.org/abs/2202.11214

- Motivation: **Climate change** is making storms both stronger and less predictable, leading to more frequent natural disasters.
- Task: **Emulate the dynamics of global whether patterns** and predict extreme whether events like atmospheric rivers.
- Training Data: 10TB of earth system data from the past.
- Input: The current state of atmospheric fields.
- Result: **Predict the precise path of catastrophic atmospheric rivers a full week in advance**, with only a fraction of a second on powerful GPUs.

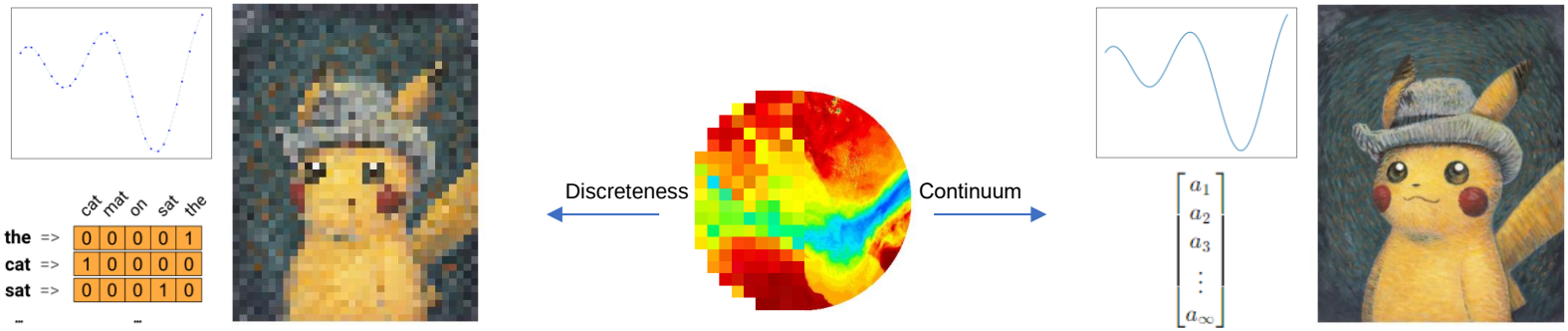
FAR BEYOND

Stony Brook University

A Challenge in Operator Learning

Deep neural networks can only take finite-dimensional inputs and produce finite-dimensional outputs:

$$\phi_{\text{network}} : \mathbb{R}^{d_{in}<\infty} \mapsto \mathbb{R}^{d_{out}<\infty}$$



Finite-dimensional Features

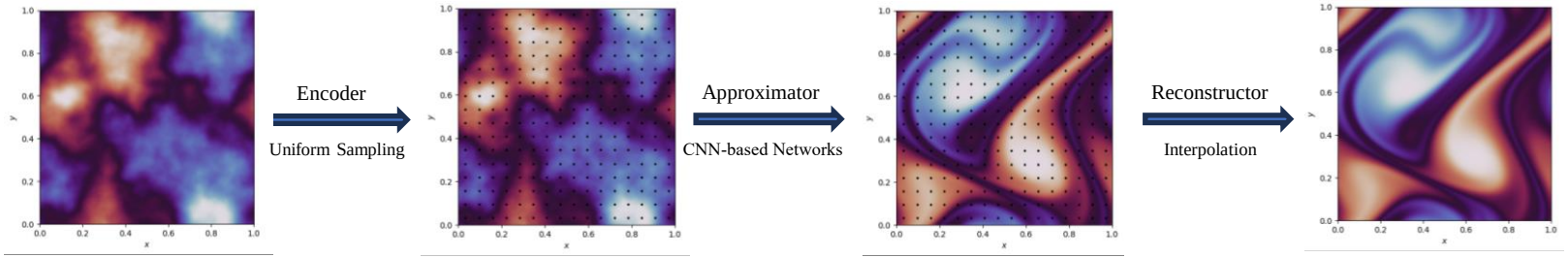
Infinite-dimensional Features

FAR BEYOND

Stony Brook University

Finite-Dimensional Features and Learning

To adapt neural networks to learn operators, a workaround is to use a simplified setting in which functions are characterized by finite-dimensional features.

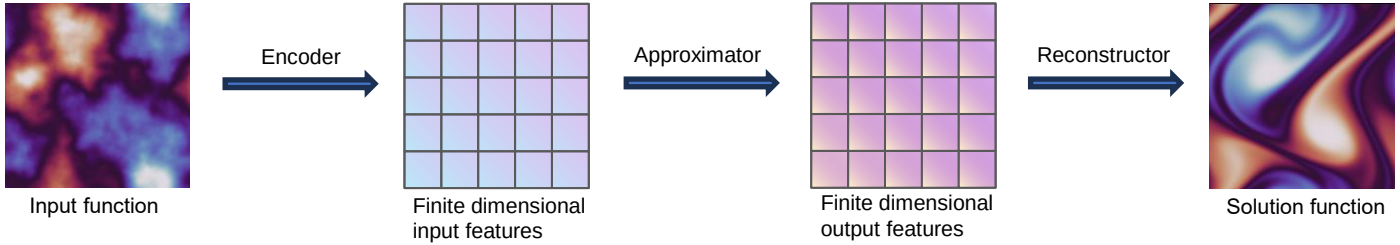


FAR BEYOND

Stony Brook University

Finite-Dimensional Features and Learning

Many numerical schemes can be represented by this diagram as well:



Method	Encoder	Approximator	Example Reconstructor
Finite Difference	Point Values	Numerical Scheme	Polynomial Interpolantion
Finite Element	Node Values	Numerical Scheme	Galerkin Basis
Finite Volume	Cell Averages	Numerical Scheme	Polynomial Interpolantion
Spectral Methods	Fourier Coefs.	Numerical Scheme	Fourier Basis

How we make choices of encoders, reconstructors (decoders), and approximators gives rise to different neural operators.
Moreover, different choices may lead to different properties or various pros and cons for a particular neural operator.

FAR BEYOND

Stony Brook University

Finite-Dimensional Learning: Summary

There have been a lot of work done in this direction, and to name a few:

Method	Encoder	Approximator	Example Reconstructor
Finite Difference	Point Values	Numerical Scheme	Polynomial Interpolantion
Finite Element	Node Values	Numerical Scheme	Galerkin Basis
Finite Volume	Cell Averages	Numerical Scheme	Polynomial Interpolantion
Spectral Methods	Fourier Coefs.	Numerical Scheme	Fourier Basis
CNN-based Networks	Grid Point Values	DNN	Interpolantion
SNO [1]	Fourier/Chebyshev Coefs.	DNN	Fourier/Chebyshev Basis
DeepOnet [2]	Sensor Point Values	Branch Net (DNN)	Trunk Net (DNN)
PCA-Net [4]	PCA	DNN	PCA
IAE-Net [5]	Auto-encoder	DNN	Auto-decoder
CORAL [6]	Implicit Neural Representation (DNN)	DNN	Implicit Neural Representation (DNN)

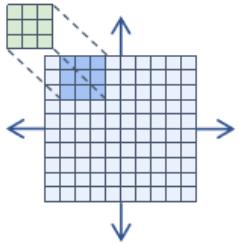
[1] Spectral Neural Operators. V. Fanaskov, I. Oseledets
[2] Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. Lu Lu et al. NMI, 2021.
[4] Model Reduction And Neural Networks For Parametric PDEs. The SMAI Journal of computational mathematics, Volume 7 (2021), pp. 121-157. Kaushik Bhattachary et al..
[5] Integral autoencoder network for discretization-invariant learning. JMLR, 2022. Yong Zheng Ong, Zuowei Shen, and Haizhao Yang
[6] Operator learning with neural fields: Tackling PDEs on general geometries. arXiv, 2023. Louis Serrano, et al..

FAR BEYOND

Stony Brook University

Infinite-Dimensional Learning: Introduction

For finite-dimensional models, the network is **highly dependent on the resolution of the data and/or sensor locations**.



In CNN-based methods, fixed size kernels converge to a point-wise operator as the resolution increases.

- Discretization is also a human-bias!
- The network should be independent of the discretization of the functions.
- Learned parameters should be transferable between discretizations.
- Learns continuous functions instead of discretized vectors.

Kernel Integral Operators

In a standard deep neural network, a layer can be written as:

$$\text{Input: } v_t \longrightarrow \text{Linear Transformation: } W^T v_t + b \longrightarrow \text{ Non-linearity } \longrightarrow \text{ Output: } v_{t+1}$$

Here the input, v_t, and the output, v_{t+1}, are both vectors.

However, we wish to learn continuous functions instead of discretized vectors. We need to adjust the formulation of our linear layers as it must be able to take functions as input:

$$v_t(x) \longrightarrow \text{Integral Linear Operator: } \int \kappa(x,y) v_t(y) dy + b(x) \longrightarrow \text{ Non-linearity } \longrightarrow v_{t+1}(x)$$

Why this is less dependent on the discretization? Numerical integration which will converge, under fairly general conditions, to the true integral.

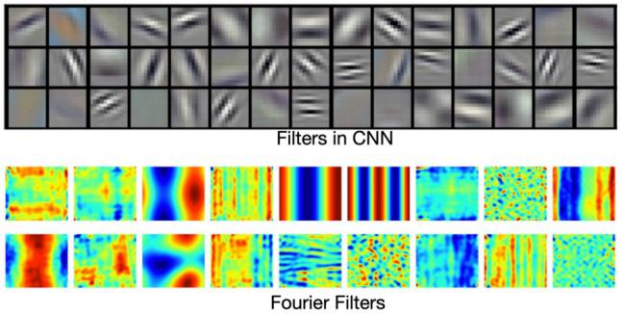
Stony Brook University

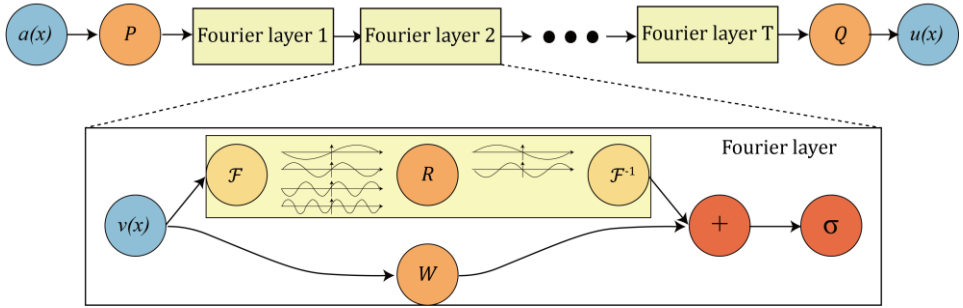
Kernel Integral Operators: FNO

FNO is widely recognized kernel integral operator [8], stands out for its **efficiency and precision**.

Integral Linear Operator	$\int \kappa(x,y)v(y)dy$
Convolution Operator	$\int \kappa(x-y)v(y)dy$
Solving Convolution in Fourier domain	$\mathcal{F}^{-1}(\mathcal{F}(\kappa)\cdot\mathcal{F}(v))$

"Filters in convolution neural networks are usually local. They are good to capture local patterns such as edges and shapes. Fourier filters are global sinusoidal functions. They are better for representing continuous functions."





FAR BEYOND

[8] Fourier neural operator for parametric partial differential equations. ICLR, 2021. Zongyi Li, et al..
Image Source: Stanford CS159: Representation Learning for Science

Stony Brook University

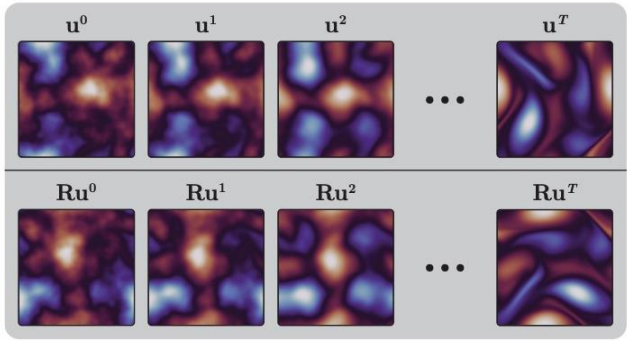
Talk Outline

- **Atomistic Systems**
  - Geometric Representation and Learning of Atomic Systems
  - Symmetries in Atomic Systems
  - Imperfect Scientific Data
  - Explainability: Does AI Model Learn Science?
- **Continuum Systems**
  - Continuum Systems and PDEs
  - Physics-informed Neural Networks
  - Operator Learning
  - Symmetries in Continuum Systems
  - Paradox in De-biasing from Discretization
- **Future Works**
  - Robustness of Equivariant GNNs
  - Unconstrained Molecular Generation

Stony Brook University

Symmetries In Continuum Systems

- Symmetry group of a PDE: Characterizes the transformations (e.g., rotations, reflections, translations, scaling) under which solutions remain solutions.



Example: Solutions are still physically consistent after rotation

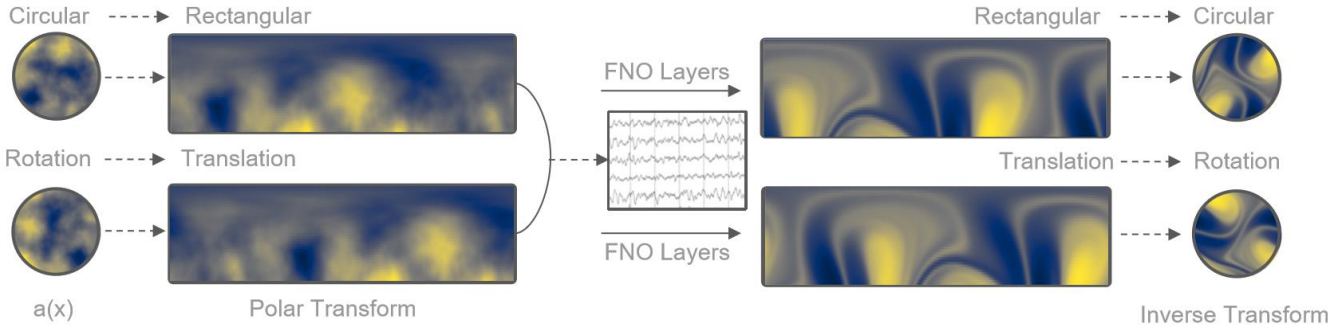
- Symmetry priors: Encode symmetries into architectures to make learning easier and improve generalization

Wenhan Gao, Ruichen Xu, Hong Wang, **Yi Liu**, Coordinate Transform Fourier Neural Operators for Symmetries in Physical Modelings, TMLR 2024

Stony Brook University

Coordinate Transforms and Symmetries

- Coordinate transformations **change the mathematical description but maintain physical properties.**
- Some symmetries can become manifest in a **new coordinate system**.
- Symmetries are **preserved in another coordinate system**.



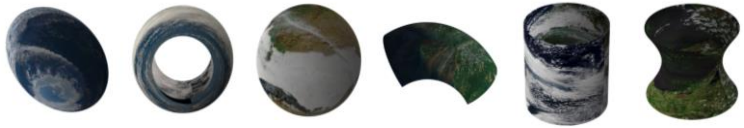
FAR BEYOND

Wenhan Gao, Ruichen Xu, Hong Wang, **Yi Liu**, Coordinate Transform Fourier Neural Operators for Symmetries in Physical Modelings, TMLR 2024

Stony Brook University

Results and Broader Impact

- Generalization: Can adapt to **various domains** commonly encountered in engineering and natural sciences.



- Theoretical Guarantees: Can approximate the behavior of an operator to **any desired precision**.
- Maintaining Symmetries: When symmetries exist, can **reduce the error** from as large as 31.942% to 0.402%

	Common Information			Testing Set A	Testing Set B
	# Par. (M)	Train(%)	TPE[a] (s)	Test (%)	Test (%)
FNO (C_4)	2.37	0.412(0.03)	**0.416**	22.537(1.53)	31.942(1.53)
G-FNO (C_4)	2.24	**0.378(0.04)**	0.709	**0.400**(0.04)	**0.402**(0.04)
Radial-FNO (C_4)	2.63	0.386(0.03)	0.540	0.431(0.04)	0.434(0.04)
CT-FNO (Arbitrary)	2.31	0.399(0.06)	0.459	**0.403**(0.07)	**0.402**(0.06)

[a] TPE: Average Time Per Epoch since G-FNO will increase the channel width and increase FFT computations.

FAR BEYOND

Wenhan Gao, Ruichen Xu, Hong Wang, **Yi Liu**, Coordinate Transform Fourier Neural Operators for Symmetries in Physical Modelings, TMLR 2024

Stony Brook University

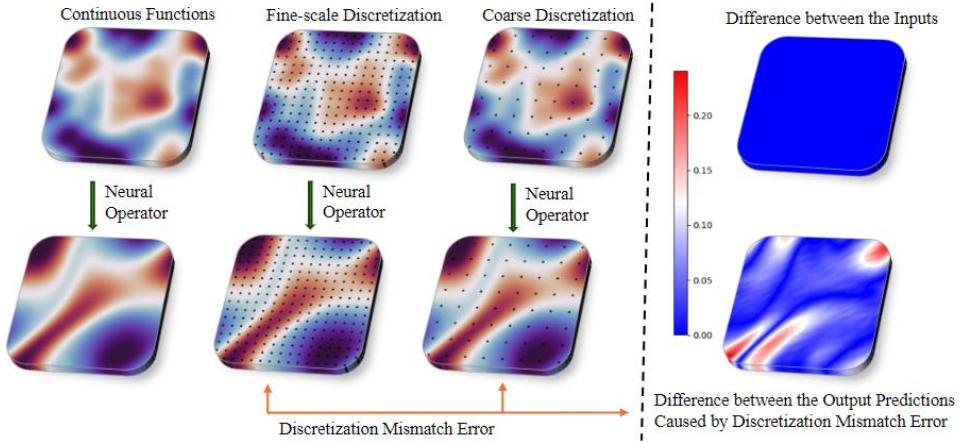
Talk Outline

- **Atomistic Systems**
  - Geometric Representation and Learning of Atomic Systems
  - Symmetries in Atomic Systems
  - Imperfect Scientific Data
  - Explainability: Does AI Model Learn Science?
- **Continuum Systems**
  - Continuum Systems and PDEs
  - Physics-informed Neural Networks
  - Operator Learning
  - Symmetries in Continuum Systems
  - Paradox in De-biasing from Discretization
- **Future Works**
  - Robustness of Equivariant GNNs
  - Unconstrained Molecular Generation

Stony Brook University

Human Bias in Representing Functions

- Discretization is also a **human bias**.
- Neural operators **remove the dependence on the discretization**.
- However, they still suffer from the discretization mismatch errors.



FAR BEYOND

Wenhan Gao, Ruichen Xu, Yuefan Deng, **Yi Liu**, Discretization-invariance? On the Discretization Mismatch Errors in Neural Operators. Under Review ICLR 2025

Discretization Mismatch Errors

- Discretization Mismatch Errors (DMEs) quantifies the **discrepancy between the outputs of the neural operator when using different discretizations**.
- The network introduces DMEs and DMEs propagate through each layer even if the inputs do not observe any DMEs.

Definition 4.3. Let $\Theta \subseteq \mathbb{R}^p$ be a finite dimensional parameter space and $\mathcal{G} : \mathcal{A} \times \Theta \to \mathcal{U}$ a map representing a parametric class of Fourier neural operators with parameters $\theta \in \Theta$. Given a uniform discrete refinement $(\Omega_J)_{J=1}^{\infty}$ of the domain $\Omega \subset \mathbb{R}^d$, the *discretization mismatch error* between the Fourier neural operator taking two different discretization of the function a is defined as

$$E_{MN} := \left\| \mathcal{G}_\theta \left(a|_{\Omega_M} \right) - \mathcal{G}_\theta \left(a|_{\Omega_N} \right) \right\|_{\mathcal{U}}.$$

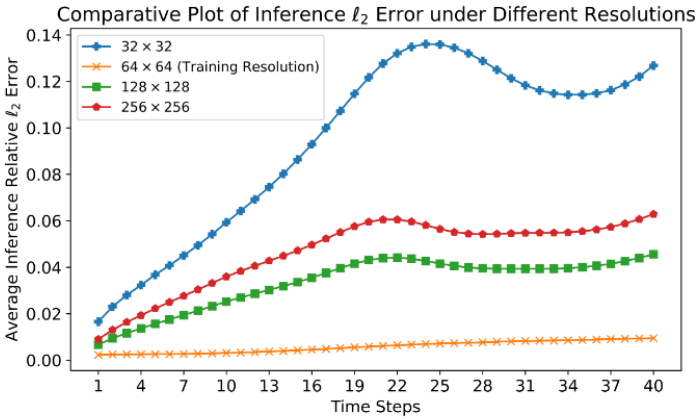
Proposition 4.4. *Let $\Theta \subseteq \mathbb{R}^p$ be a finite dimensional parameter space, $\mathcal{A} = H^s\left(\mathbb{T}^d; \mathbb{R}^{d_a}\right)$, $\mathcal{U} = H^{s'}\left(\mathbb{T}^d; \mathbb{R}^{d_u}\right)$, and $\mathcal{G} : \mathcal{A} \times \Theta \to \mathcal{U}$ a map representing a parametric class of Fourier neural operators with parameters $\theta \in \Theta$ with L Fourier layers. Consider $\sigma(x)$ as the activation function that is (globally) ω-Lipschitz continuous for all layers. Given a discretization Ω_M, then the discretization mismatch error between the Fourier neural operator taking Ω_M and another discretization Ω_N, given by $E_{MN} = \left\| \mathcal{G}_\theta \left(a|_{\Omega_M} \right) - \mathcal{G}_\theta \left(a|_{\Omega_N} \right) \right\|$ with $N > M$, will increase as N increases. Additionally, the discretization mismatch error might increase as L and ω increase.*

Wenhan Gao, Ruichen Xu, Yuefan Deng, **Yi Liu**, Discretization-invariance? On the Discretization Mismatch Errors in Neural Operators. Under Review ICLR 2025

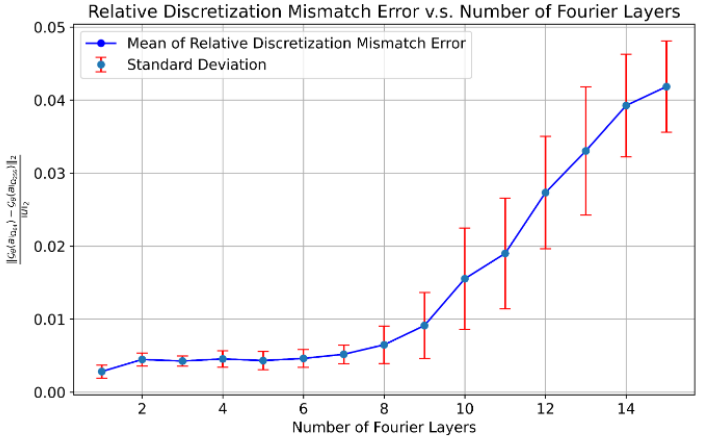
Stony Brook University

Accumulation of DMEs

- Numerical experiments reveals the accumulation of DMEs through network layers and time steps.



(a)



(b)

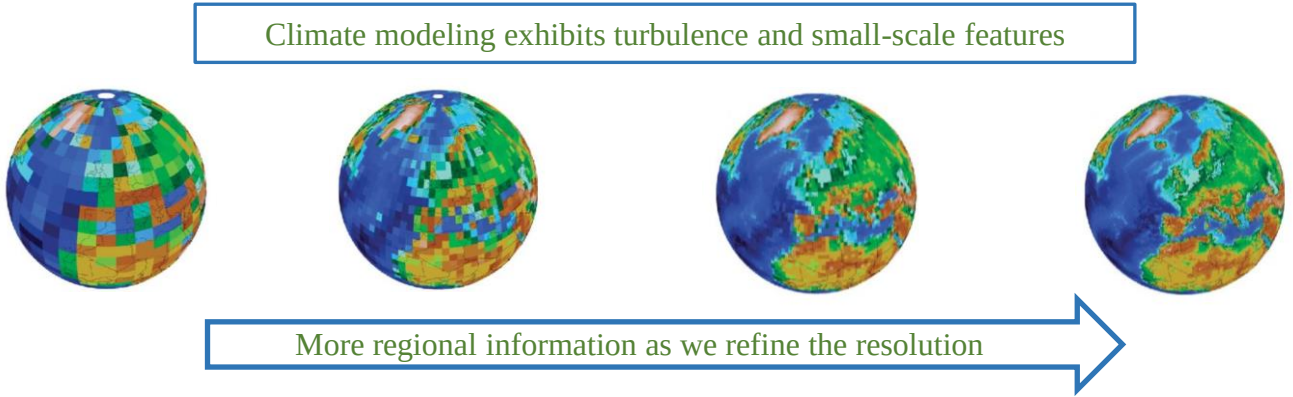
FAR BEYOND

Wenhan Gao, Ruichen Xu, Yuefan Deng, **Yi Liu**, Discretization-invariance? On the Discretization Mismatch Errors in Neural Operators. Under Review ICLR 2025

Stony Brook University

Results and Broder Impact

- CROP reduces the cross-resolution (super-resolution) error of FNO from as high as 9.45% to 0.54%.
- Multi-spatio-learning feature improves performance on difficult learning tasks such as the NS equation under high Reynolds number (turbulent flow with small-scale features).



FAR BEYOND

Wenhan Gao, Ruichen Xu, Yuefan Deng, **Yi Liu**, Discretization-invariance? On the Discretization Mismatch Errors in Neural Operators. Under Review ICLR 2025

Stony Brook University

Publications and Awards

Publications

*Equal Contribution

- Ronast Subedi*, Lu Wei, **Wenhan Gao***, Shayok Chakraborty, **Yi Liu,** Empowering Active Learning for 3D Molecular Graphs with Geometric Graph Isomorphism, NeurIPS, 2024
- **Wenhan Gao**, Ruichen Xu, Hong Wang, **Yi Liu**, Coordinate Transform Fourier Neural Operators for Symmetries in Physical Modelings, TMLR 2024
- **Wenhan Gao**, Chunmei Wang, Active Learning Based Sampling for High-Dimensional Nonlinear Partial Differential Equations**,** The Journal of Computational Physics (JCP) 2023
- Xufeng Liu, Dongsheng Luo, **Wenhan Gao**, **Yi Liu**, 3DGraphX: Explaining 3D Molecular Graph Models via Incorporating Chemical Priors, KDD 2025

Preprints

- **Wenhan Gao**, Ruichen Xu, Yuefan Deng, **Yi Liu**, Discretization-invariance? On the Discretization Mismatch Errors in Neural Operators. Under Review ICLR 2025
- Xufeng Liu*, **Wenhan Gao***, **Yi Liu**, Energy-Based Discrete Mask Approximation for 3D Molecular Graph Explanation, Under Review ICLR 2025
- **Wenhan Gao**, Jian Luo, **Yi Liu**, Dynamic Schwartz-Fourier Neural Operator for Enhanced Expressive Power, Under Review CVPR 2025

Awards

- Excellence in Teaching (Fall 2023; AMS 595 Fundamentals of Computing; Independent Leading Instructor)
- NeurIPS Scholar Award (2024)

Stony Brook University

Thank You!

FAR BEYOND