Brookhaven National Laboratory

U.S. DEPARTMENT OF ENERGY

FY27 NPP LDRD Type B Pre-Proposal

Constuction of Foundation Model for Lattice QCD

Chulwoo Jung(PI, Presenter), Peter Boyle(co-PI),
Taku Izubuchi (co-PI), Huan-Hsin Tseng (co-Pi)

January 9, 2026

@BrookhavenLab

FY27 NPP LDRD Type B Pre-Proposal

Proposal title: Construction of Foundation Model for Lattice QCD

Primary Investigator: Chulwoo Jung (PO)

Other Investigators: Peter Boyle, Taku Izubuchi (PO), Huan-Hsin Tseng (CDS)

Indicate if this is a cross-directorate proposal: Yes _X__ No___

If yes, identify other directorates/organizations: CDS

Proposal Term: From: Oct. 2026 To: Sep. 2028

Brookhaven National Laboratory

FY27 NPP LDRD Type B Pre-Proposal

Title: Construction of Foundation Model for Lattice QCD

Abstract: We aim to develop new methods and Foundation Models (FM) for lattice QCD and related high-dimensional scientific problems, leveraging self-supervised learning and modern machine-learning architectures to capture the essential structure of four-dimensional SU(3) gauge field configurations. By constructing and validating a scalable FM for LQCD, we aim to significantly reduce computational cost while enabling new algorithmic capabilities across nuclear and high-energy physics.

Program: HEP & CDS

Return on Investment: This LDRD will positions BNL to lead future DOE initiatives at the intersection of AI and HEP, NP, including SciDAC-6, AI for Science, and EIC computing programs.

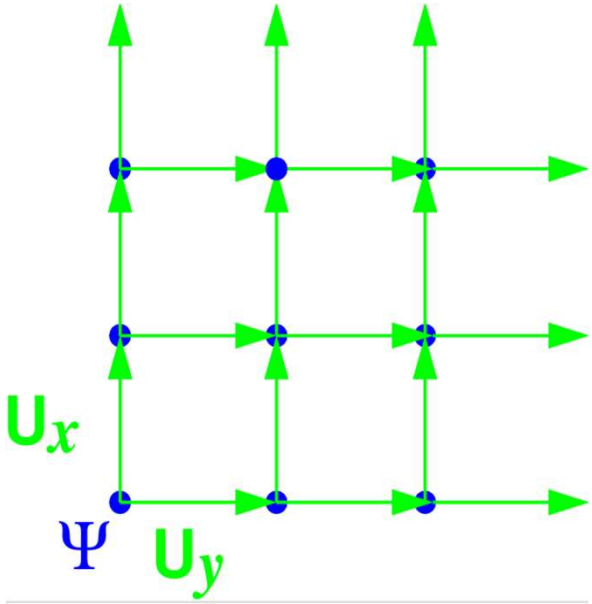
Broader impact on the activities at the laboratory: Reducing cost for LQCD calculation of currently challenging problems, such as Generalized Parton Distibution(GPD) and multi-nucleon systems, which are crucial for physics programs such as EIC & DUNE.

Total planned funding per year in FY27 and FY28 1 postdoc per year (~130K/yr)

Description of the LDRD: Why FM for Lattice QCD (LQCD)?

QCD : Theory of strongly interacting quarks and gluons.

- 4D structured data
- Gauge symmetry SU(3)
- Locality
- Translation & Rotation invariance



LQCD is an ideal scientific testbed for Foundation Models because the data are high-dimensional, structured, symmetry-constrained, and abundant.

Challenges:

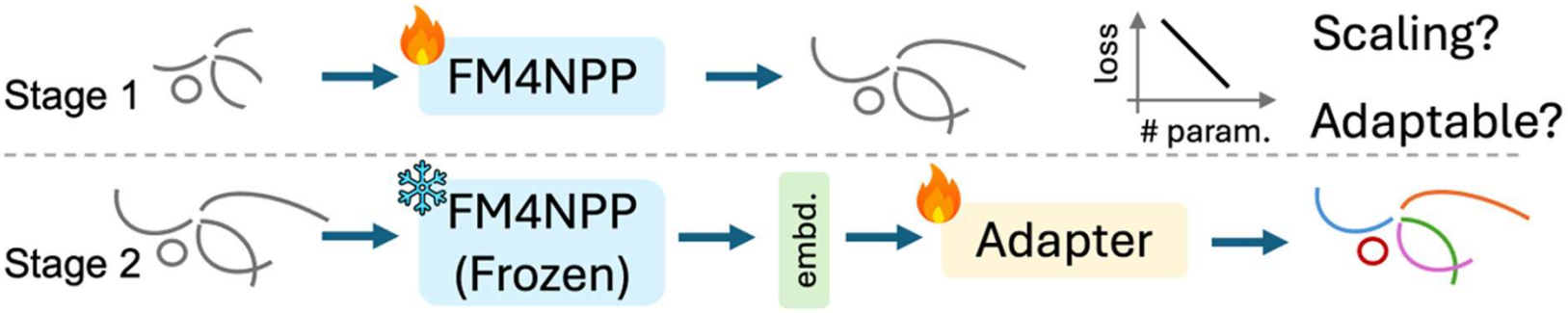
- Big data : V= 96^3x 192, Rank(D) ~ $10^{10-11.}$
- Ensemble generation according to e^{-S} and critical slowing down: rapidly increasing computation needs $(\sim(lattice\ spacing)^{-(10-12)})$
- Physical observable calculation : quite expensive to calculate on each samples and/or requires a lot of samples.

So far, successful application of ML techniques have been limited to simpler systems in lower dimension (2D U(1), etc..)

Foundation Models (Bommasani *et al, arXiv:2108.07258):*

Capture the essential features of data in a data-efficient and scalable manner, then adaptable to various task with minimal learning with labeled data

1. Self supervised learning to construct FM, demonstrate scaling
2. Freeze weights of FM and train adapters for specific observables.

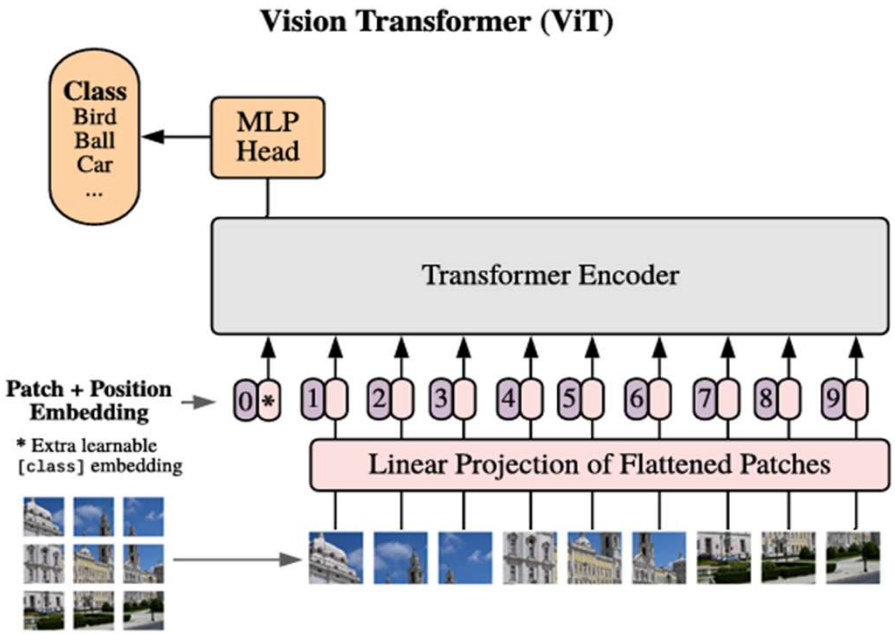


From arXiv: 2508.14087

Strategies for FM training

Convolutional Neural Network(CNN): Highly successful for images. Scaling for higher dimensional data needs to be demonstrated

From arXiv:2010.11929



Vision Transformer (ViT): Use Patching to transform Multidimensional data.

Discretized gluon fields on 2^4 sites = 512 numbers per imbedding.

Explore class-token(CLS) and positional embedding.

Brookhaven National Laboratory

Summary Slide

Despite the spectacular successes in Lattice QCD until now, continuing the current approach to reduce both statistical and systematic errors further is getting increasingly expensive in terms of the needed computing resources.
A successful construction of Foundation Model for Lattice QCD can reduce the cost of both the generation of independent samples and extraction of physics from these samples significantly, leading to significant reductions in errors.

Deliverable: Training algorithm for 4D SU(3) data. Demonstration of scalability for moderate size LQCD ensembles.

Personnel: (co-)PIs, LDRD postdoc, others from HEP and CDS
Required Procurements: 1 Postdoc