



FY27 NPP LDRD Type B Pre-Proposal

Improving LArTPC Clustering through AI-Based Scientific Reasoning and Fine-grained Domain-specific Tools toward Enhanced Discovery Capability in DUNE and SBND

Haiwang Yu (PI, PO, Presenter) , Yi Huang (co-PI, CDS)

Yousen Zhang (PO), Xin Qian (PO), Brett Viren (PO), Yihui Ren (CDS)

Jan 9 2026



FY27 NPP LDRD Type B Pre-Proposal

Proposal title: Improving LArTPC Clustering through AI-Based Scientific Reasoning and Fine-grained Domain-specific Tools toward Enhanced Discovery Capability in DUNE and SBND

Primary Investigator: Haiwang Yu (PI, PO) , Yi Huang (co-PI, CDS)

Other Investigators: Yousen Zhang (PO), Xin Qian (PO), Brett Viren (PO), Yihui Ren (CDS)

Indicate if this is a cross-directorate proposal: Yes

If yes, identify other directorates/organizations: CDS

Proposal Term: From: 2026/10 To: 2028/9

FY27 NPP LDRD Type B Pre-Proposal

Proposal title and brief abstract:

Improving LArTPC Clustering through AI-Based Scientific Reasoning and Fine-grained Domain-specific Tools toward Enhanced Discovery Capability in DUNE and SBND

LArTPC technology is central to the DUNE and SBN programs, which target the next generation of neutrino physics discoveries. Recent MicroBooNE results exclude the single light sterile neutrino explanation of the LSND and MiniBooNE anomalies, highlighting both the strengths of LArTPC detector technology and the limitations of current reconstruction methods, particularly for complex topologies such as photon-related channels. We propose to explore a novel approach that integrates fine-grained domain-specific tools from the Wire-Cell Toolkit with AI-based scientific reasoning to orchestrate reconstruction workflows, with LArTPC clustering as a concrete and representative use case. This approach aims to improve physics performance while establishing LArTPC reconstruction as a high-impact testbed for AI-driven scientific reasoning in experimental physics.

FY27 NPP LDRD Type B Pre-Proposal

Program: HEP

Return on Investment

- The proposed study aligns well with recent U.S. initiatives (e.g., AmSC and Genesis Mission) emphasizing AI-empowered science.
- The application of agentic AI to the HEP reconstruction chain, particularly fine-grained tool orchestration, remains novel and exploratory, strengthening the combined capabilities of BNL HEP and CDS.
- Positions BNL to compete more effectively for future DOE AI-focused funding opportunities.

Broader Laboratory Impact

- Expands and deepens expertise in advanced AI methodologies for HEP reconstruction
- Establishes a pathway for broader AI integration, from coarse- to fine-grained levels, across other scientific domains at BNL

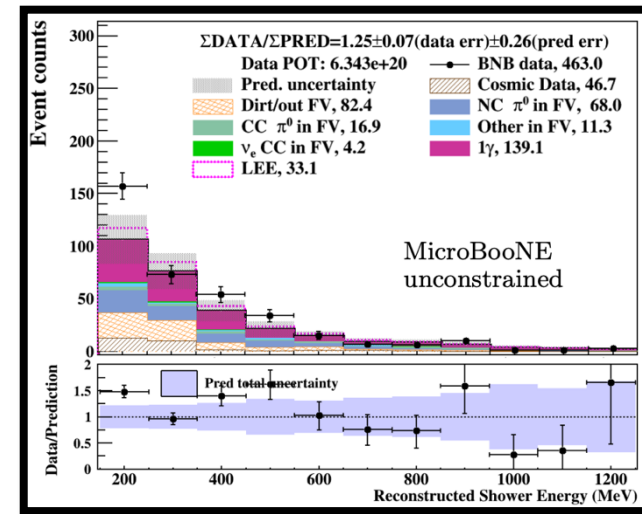
Total planned funding per year in FY27 and FY28: \$250k/year

- Postdoctoral researcher (PO): 0.5 FTE
- Scientific staff (PO + CDS): 0.5 FTE
- Computing resources

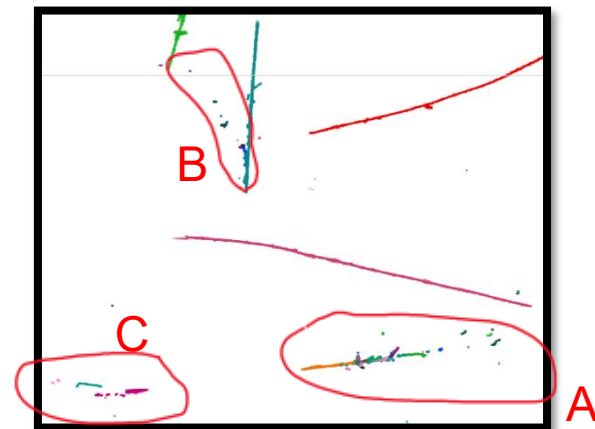
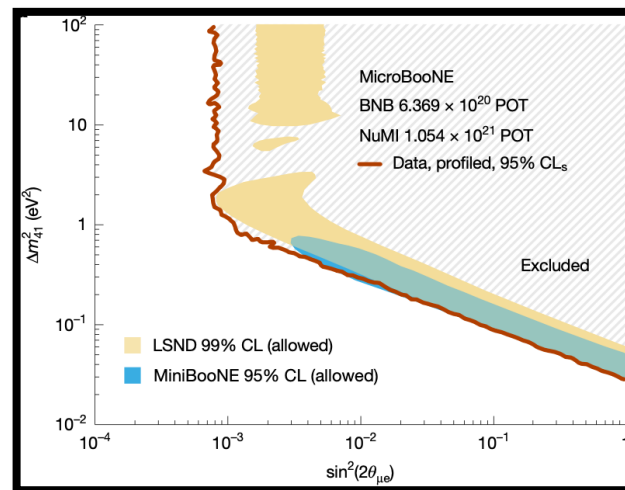
Motivation: Improve LArTPC reconstruction for complex topologies, such as photon-related channels, which are expected to carry rich and potentially new physics messages

- The single light sterile neutrino interpretation has been ruled out
- The first MicroBooNE search for the MiniBooNE LEE in an inclusive photon channel observed a 2.2σ local significance relative to GENIE predictions
- Current LArTPC clustering performance remains limited for complex topologies such as photon-rich final states
- Experience from Wire-Cell development shows that **human reasoning** in selecting and sequencing tools plays a key role in achieving optimized performance.
 - **Human reasoning -> AI reasoning?**

arxiv:2502.06064



Nature volume 648,
pages 64–69 (2025)



[BEE link](#)

Approach: Combine established Wire-Cell tools with agentic AI methodologies

- **Current Limitation:** Existing clustering algorithms are applied in a fixed sequence across all topologies, leading to inefficiencies and suboptimal performance for complex or topology-specific events
- **Proposed Solution**
 - Package existing clustering algorithms (granularity is at C++ function/algorithm level) as MCP tools
 - Develop an event descriptor to summarize key event features in text/JSON
 - Implement an MCP server in C++
 - Build a Python fine-tuning framework leveraging the MCP server and open-weight LLMs
 - Initialize with supervised fine-tuning guided by traditional workflows
 - Apply truth-based reinforcement learning fine-tuning (RLFT)

Deliverables

- A validated workflow to enhance LArTPC reconstruction, with photon-related topologies as a representative use case
- An extensible MCP-based framework integrated into Wire-Cell that is capable of algorithm-level tool orchestration
- A reusable platform for testing and advancing AI methodologies in HEP reconstruction
- *(Optional extension)* Photon-related analyses in SBND using the proposed workflow conducted through collaboration with Universities

Personnel

- **Haiwang Yu (PO):** C++ implementation and system integration
- **Yi Huang (CDS):** Python framework and fine-tuning pipeline
- **Postdoc (PO):** Model training, validation, and performance studies
- **Additional team members:** Advisory and domain-expert support

Summary 1

Proposal Overview

- We propose to improve LArTPC clustering through AI-based scientific reasoning and fine-grained, domain-specific tools.
- **Long-Term Physics Impact:** The final goal is to enhance the discovery capabilities of LArTPC in DUNE and SBND. This study serves as a technology demonstration, delivering **infrastructure** that can be reused in the future.
- **AI Impact:** The proposed work focuses on the **fine granularity** of agentic AI in science, complementing and extending emerging agentic AI applications in scientific research.
- **Near-Term Physics Impact:** The direct outcomes of the proposed work can be utilized by collaborators in physics analyses (**out side scope of this LDRD**) with the potential for high impact.
- Note: Compared to traditional and machine-learning–based reconstruction optimization methods, the proposed agentic AI approach is exploratory in nature and is **not included in the baseline R&D scope**.

Effort and Resources (2 years)

- Postdoctoral researcher (PO): 0.5 FTE
- Scientific staff (PO + CDS): 0.5 FTE
- Computing resources

Summary 2

Return on Investment

- The proposed study aligns well with recent U.S. initiatives (e.g., AmSC and Genesis Mission) emphasizing AI-empowered science.
- The application of agentic AI to the HEP reconstruction chain, particularly fine-grained tool orchestration, remains novel and exploratory, strengthening the combined capabilities of BNL HEP and CDS.
- Positions BNL to compete more effectively for future DOE AI-focused funding opportunities.

Broader Laboratory Impact

- Expands and deepens expertise in advanced AI methodologies for HEP reconstruction
- Establishes a pathway for broader AI integration, from coarse- to fine-grained levels, across other scientific domains at BNL