Plan Overview

A Data Management Plan created using DMP Tool

DMP ID: https://doi.org/10.48321/D1XC9J

Title: Hadron spectroscopy from Anisotropic Clover Lattices

Creator: Robert Edwards - **ORCID:** 0000-0002-5667-291X

Affiliation: Thomas Jefferson National Accelerator Facility (jlab.org)

Funder: United States Department of Energy (DOE) (energy.gov)

Funding opportunity number: NP/LQCD Computing Facility

Template: Department of Energy (DOE): Office of Science

Project abstract:

Computation and storage of meson and baryon elementals, perambulators, generalized perambulators, and correlation functions from Lattice QCD by the Hadron Spectrum (HadSpec) collaboration.

Start date: 07-01-2021

End date: 06-30-2030

Last modified: 02-28-2025

Copyright information:

Hadron spectroscopy from Anisotropic Clover Lattices

1. Data sharing and preservation

Data management plans should describe whether and how data generated in the course of the proposed research will be shared and preserved. If the plan is not to share and/or preserve certain data, then the plan must explain the basis of the decision (for example, cost/benefit considerations, other parameters of feasibility, scientific appropriateness, or limitations discussed in #4). At a minimum, DMPs must describe how data sharing and preservation will enable validation of results, or how results could be validated if data are not shared or preserved.

Data types and sources

Data will be generated using lattice QCD techniques. These data objects are to be used in research programs for QCD spectroscopy. The raw and processed data will be archived on local tape resources at Jefferson Lab. It will be made available to other researchers on request.

Content and format

The data are stored in a database format called "filedb" developed at Jefferson Lab and is freely available. The software is available on GitHub in filedb.

Sharing and preservation

The data is a in raw format that is used for other analysis projects. The output from those analysis projects are suitable for publication. The data from the off-line analysis will be preserved in Git repositories on GitHub along with the text and figures for the publications.

Rational

The data generated will be used to elucidate the theory of elementary particles and their interactions. Results of calculations will help interpret experiments at national and international facilities.

2. Data used in publications

Data management plans should provide a plan for making all research data displayed in publications resulting from the proposed research open, machine-readable, and digitally accessible to the public at the time of publication. This includes data that are displayed in charts, figures, images, etc. In addition, the underlying digital research data used to generate the displayed data should be made as accessible as possible to the public in accordance with the Principles published in the DOE Policy for Digital Research Data Management. The published article should indicate how these data can be accessed.

Data for publications

There are two main classes of data. The database objects are in raw format that is used for other analysis projects. This source data has a URL specifying its location.

This raw data is used in analysis projects which produce data and figures suitable for publication. The data from the offline analysis will be preserved in Git repositories on GitHub along with the text and figures for the publications.

Raw data

Below is a table showing the expected production of raw data under the 2025-2026 allocation year.

Production date	Ensemble	Type	Number	Size
Expected 2026	48^3x512, m_pi=170 MeV	Elementals	(light + strange up to mom= (2,0,0))*212 configs	220 TB
Expected 2026	48^3x512, m_pi=170 MeV	Forward propagators (FP) and annihilation lines (AL) for light and strange quarks	FP: 16 time sources, 212 configs AL: 32 time sources, 212 coffins	30 TB
Expected 2025	48^3x512, m_pi=170 MeV	Laplacian solution vectors (N_vecs=640)	Total of 6748	174 TB
Expected 2026	32^3x256, m_pi=239, 283 MeV	Genprop	(light+strange)*500 cfgs	240 TB
Expected 2026	24^3x256, m_pi=330,700 MeV	Genprop	(light+strange)*500 cfgs	70 TB

3. Data management resources

Data management plans should consult and reference available information about data management resources to be used in the course of the proposed research. In particular, DMPs tht explicitly or implicitly commit data management resources at a facility beyond what is conventionally made available to approved users should be accompanied by written approval from that facility. In determining the resources available for data management at DOE Scientific User Facilities, researchers should consult the published description of data management resources and practices at that facility and reference it in the DMP. Information about other Office of Science facilities can be found in the additional guidance from the sponsporing program.

Data management resources at Jefferson Lab

This project will follow the data management policies of Jefferson Lab and USQCD. Files on tape have a unique URL - universal resource locator. A file mapping of the the tape system is available. The tape facilities under the LQCD project are managed as part of the data management of all experimental data at the lab.

4. Confidentiality, security and rights

Data management plans must protect confidentiality, personal privacy, Personally Identifiable Information and U.S. national, homeland, and economic security; recognize proprietary interests, business confidential information, and intellectual property rights; avoid significant negative impact on innovation and U.S. competitiveness; and otherwise be consistent with all applicable laws, regulations, agreement terms and conditions, and DOE orders and policies. There is no requirement to share proprietary data.

Protection

The data generated will not contain PII or compromise U.S. national, homeland, and economic security. It will recognize proprietary interests, business confidential information, and intellectual property rights; avoid significant negative impact on innovation, and U.S. competitiveness; and otherwise be consistent with all applicable laws, regulations, and DOE orders and policies. The data will not involve human or animal subjects.

Planned Research Outputs

Dataset - "Correlation functions"

Two-point and three-point Euclidean correlation functions

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
Correlation functions	Dataset	Unspecified	Restricted	None specified		BSD 3-Clause "New" or "Revised" License	None specified	No	No