PML4HEP

Precision Machine Learning for High Energy Physics

Stefano Carrazza
Consiglio di Sezione - INFN Milano- 9 Luglio 2024

Nodes and coordinators

National Coordinator (Responsabile Nazionale):

Dr. Riccardo Torre

INFN Unit: Genova

e-mail: riccardo.torre@ge.infn.it

Local Coordinators (Responsabili Locali):

Prof. Stefano Forte (Prof. Ordinario - UNIMI)

INFN Unit: Milan

e-mail: stefano.forte@mi.infn.it

Dr. Luca Silvestrini (Dirigente di ricerca - INFN Roma)

INFN Unit: Rome 1

e-mail: luca.silvestrini@roma1.infn.it

Nodes composition

INFN Genova

- Staff members
  - Riccardo Torre (80%)
  - Simone Marzani (10%)
  - Fabrizio Parodi (10%)
  - Carlo Schiavi (10%)
  - Andrea Coccaro (10%)
  - Federica Sforza (10%)
- Students/postdocs/RTD
  - Francesco Armando di Bello (RTD) (10%)
  - Samuele Grossi (PhD student) (50%)

INFN Milan

- Staff members
  - Stefano Carrazza (50%)
  - Stefano Forte (40%)
  - Marco Zaro (10%)
- Students/postdocs/RTD
  - Andrea Barontini (PhD)

INFN Rome

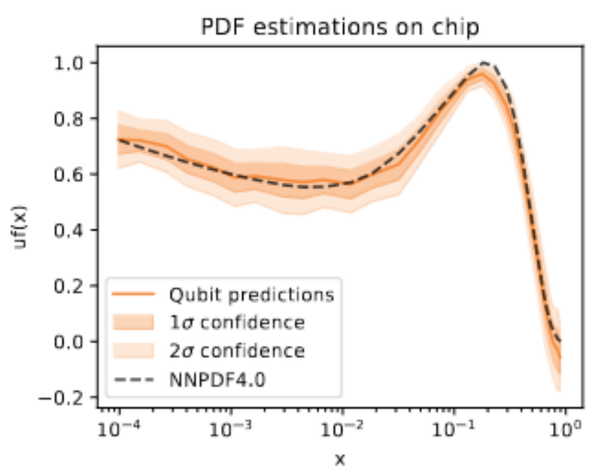
- Staff members
  - Luca Silvestrini (50%)
  - Stefano Giagu (20%)
  - Cecilia Voena (10%)
  - Fracesco Teherani (10%)
  - Andrea Messina (10%)
  - Carlo Terracciano (10%)
  - Valerio Ippolito (10%)
  - Stefano Rosati (10%)
- Students/postdocs/RTD
  - Lorenzo Arsini (PhD student) (20%)

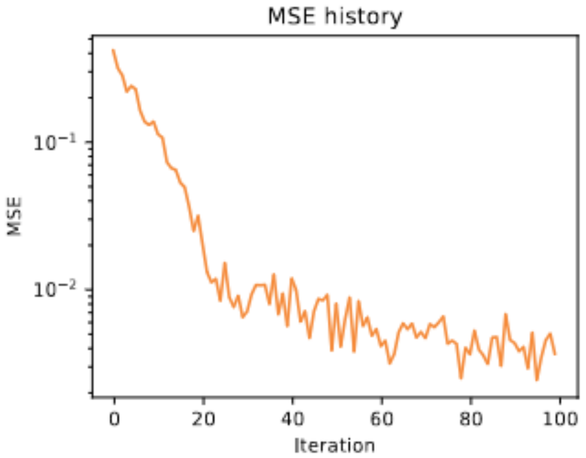
Includes researchers from different communities, both on theory and experimental side.

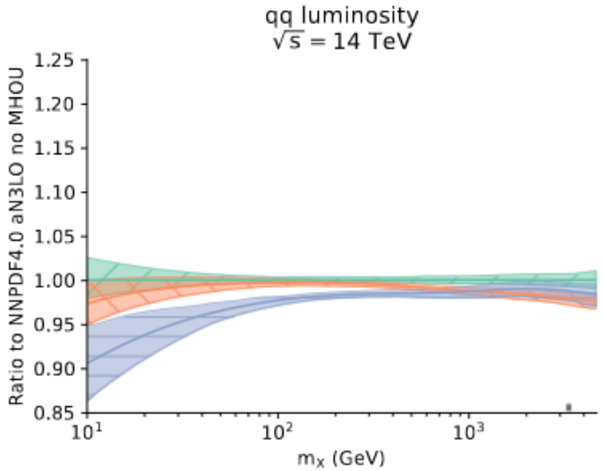
Groups activities that were already ongoing in different projects/ experiments under a unique “scope”.

Scientific goals

- The need of ML techniques in HEP goes together with the need for precision.
- Assess the ability of existing ML techniques to match the required precision, typically much higher than that of industrial and “real-life” applications.
- Develop new techniques tailored to the required precision.
- Very broad scope, ranging from theoretical to experimental applications.







Objectives

WP1: THEORY INPUTS

- Extraction of theory inputs: SM parameters and Parton Distribution Functions
- Simulation of MC events beyond LO
- Inclusion of theory and detector effects: showering, hadronization and detector response

WP2: DATA ACQUISITION

- Implementation of algorithms for online trigger, taggers, anomaly detection and data quality monitoring
- Implementation of algorithms for offline analysis through object reconstruction, jet clustering, boosted object tagging, particle flow, etc.

WP3: DATA ANALYSIS AND STATISTICAL INFERENCE

- Perform statistical inference within and beyond the SM through several techniques with varying level of model dependencies: anomaly detection, statistical learning, parametrized neural networks, model dependent searches, etc.

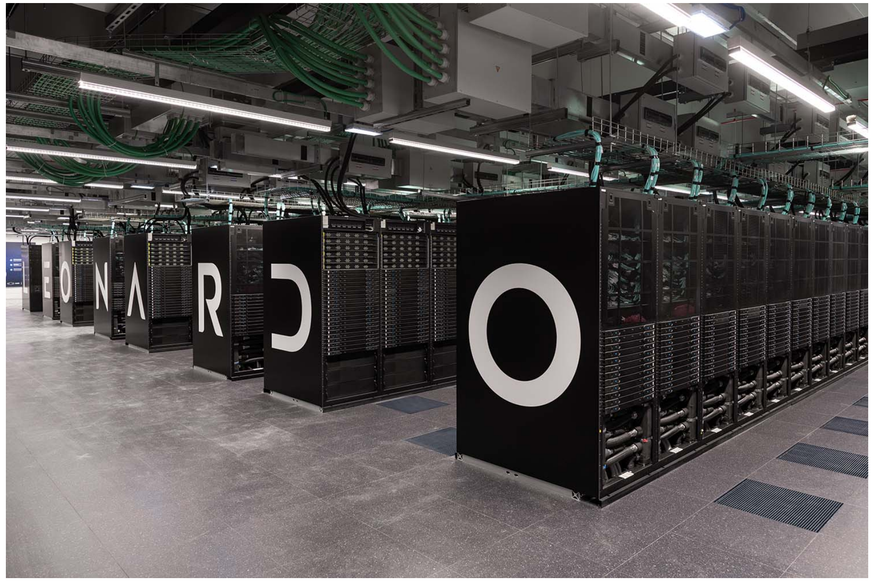
WP4: RESULTS PRESENTATION, DISTRIBUTION AND PRESERVATION

- Present results delivering full likelihood (or even full statistical model) information and by publishing enough information to allow for preservation and reinterpretation.

WP5: GENERAL ML TOOLS AND OTHER STUDIES

- Develop general ML tools such as physics inspired evaluation metrics and new sampling and integration techniques suitable for different tasks.
- Implement physics informed algorithms that exploit symmetries or properties of the data
- Study general properties of density estimation and generative models performances
- Emulation of self-consistent mean field as a function of particle/nucleon positions to accelerate nuclear interaction models of interest for medical applications.

PML4HEP Computational resources



We get a large amount of computing resources on the Leonardo supercomputer (Cineca)

The request for 2025 is:

- 1M core/hours = ~31K node/hours = ~125K A100/hours = ~14 A100/year on booster
- 10TB of storage
- 200K core/hours on DCGP (for NNPDF group)

- Can be used by anybody with a percentage on PML4HEP and collaborators/students (that anyway will have to be allocated a small percentage on PML4HEP for the forthcoming year).
- Usage is much under expectations at the moment, which means we can get much less than requested for next year if we do not increase usage.
- Settings up workflow takes little time: with some help, in a few days one may be ready to submit jobs.

Main external collaborators

Giuliano Panico (Florence U.)
Lukas Heinrich (TUM)
Eilam Gross (Weizmann)
Oliver Mattelaer (Louvain U.)
Luigi Del Debbio (Edinburgh U.)
José Ignacio Latorre (Abu Dhabi and Singapore)
H. Reyes-Gonzalez (postdoc in Genova until 10/2023)
Maurizio Pierini (CERN)
Andrea Wulzer (IFAE Barcelona)
Juan Rojo (Nikhef)
Tobias Golling (Geneva U.)
Michael Kagan (SLAC)
Steve Schramm (Geneva U.)
Daniel Whiteson (UC Irvine)

Thank you