NEEDLE – Workflow Orchestration for Neural Simulation Based Inference Tools

13th KSETA Plenary Workshop, Feb 24-26, 2026, Haus Saron in Wildberg

Kylian Schmidt, the NEEDLE Team

Introduction and FAQ

What is NSBI?

- We train neural networks to perform an *unbinned fit without a known likelihood*
- Learn the underlying distribution directly from simulations
- No data reduction compared to binned fits
- Many flavors of NSBI exist with individual strengths and weaknesses

Should we be hyped?

- Reach higher sensitivities and do more with the same amount of data
- Scales better with more parameters
- Strong improvements in analyses with interference patterns
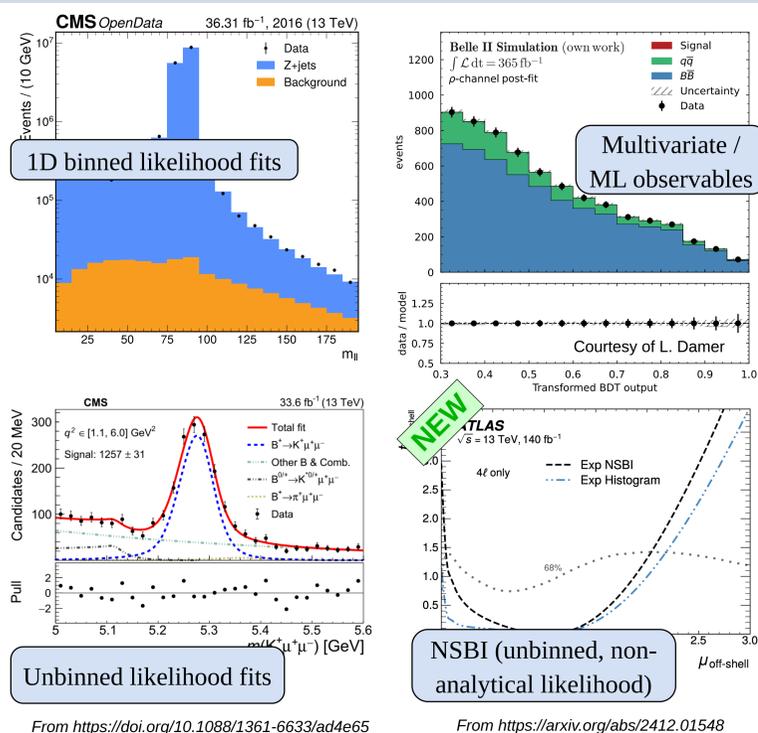- ATLAS demonstrated NSBI in off-shell Higgs analysis

From *https://arxiv.org/abs/2412.01548*

What challenges are there?

- Training and organisation of 100s of neural networks
- Extremely large datasets
- No general-purpose inference tool like CMS Combine yet
- Validation from larger HEP community
- Emerging field with many concurrent efforts

Overview of NSBI in High-Energy Physics

How we usually fit distributions

1D binned likelihood fits

Multivariate / ML observables

Courtesy of L. Damer

Unbinned likelihood fits

NSBI (unbinned, non-analytical likelihood)

From *https://doi.org/10.1088/1361-6633/ad4e65*

From *https://arxiv.org/abs/2412.01548*

Different kinds of NSBI methods

Name	Target	Models	Description
Neural Likelihood Estimation (NLE)	$p(x\|\theta)$	Diffusion / Density Estimation	Estimate the likelihood per event
Neural Likelihood Ratio Estimation (NLRE)	$r(x\|\theta) = \dfrac{p(x\|\theta)}{p(x\|\theta_0)}$	Binary Classifiers	Compare two hypotheses (Neyman-Pearson)
Neural Posterior Estimation (NPE)	$p(\theta\|x)$	Same as NLE	Robust for estimating the posterior per event
Neural Posterior Score Estimation (NPSE)	$\nabla_\theta \log(p(\theta\|x))$	Diffusion / Score-Based	Approximate local features (e.g gradients)

Recipe: NLRE with a classifier (can be done at home)

```
Ingredients:
  1x Classifier
  1x Data
  Nx MC samples
  1x GPU
     (optional)
```

$$D(x|\theta) = \frac{p(x|\theta)}{p(x|\theta) + p(x|\theta_0)}$$

Train your **classifier**, calculate the **density ratio** and compare all hypotheses for your parameter

$$\log \mathcal{L}(\theta) = \sum_N \log r(x, \theta, \theta_0)$$

$$r(x, \theta, \theta_0) = \frac{D(x|\theta)}{1 - D(x|\theta)}$$

Fit by maximizing the **log-likelihood ratio** above

The NEEDLE Project

Orchestration Software

- Framework approach for deploying many large neural network on High Performance Computing clusters
- Take care of the common boilerplate needed for NSBI
  - Data ingestion in the TB scale
  - Configs and experiment tracking
  - Parallelization and workflow management
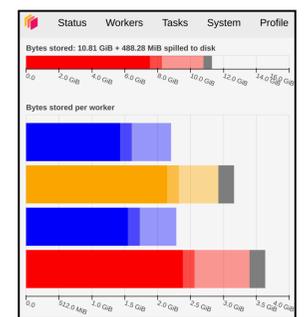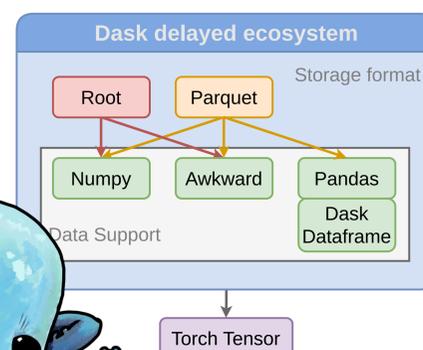- Additional library with out-of-the-box NSBI methods

Data processing · Machine Learning · Orchestration
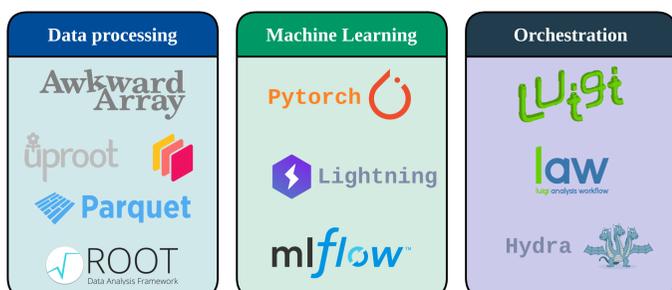
HEP can benefit immensely from using industry ML libraries

Efficient data ingestion with dask

- *Challenge*: Inefficiencies compound over hundreds of trainings
- Dask libraries allow out-of-memory computation
- Delayed arrays are not computed until necessary
- RAM and max CPUs can be configured based on job requirements

Compatible with parquet and root files

Memory and CPUs are capped at runtime

Image from *https://docs.dask.org/en/stable/dashboard.html*

Made using ProCreate
Based on *https://schmidtocean.org/expeditions-chapter_8-2024/*