Simulation-Based Inference Blueprint Workshop

Welcome, Overview and Introduction

Jay Sandesara, Nick Smith

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

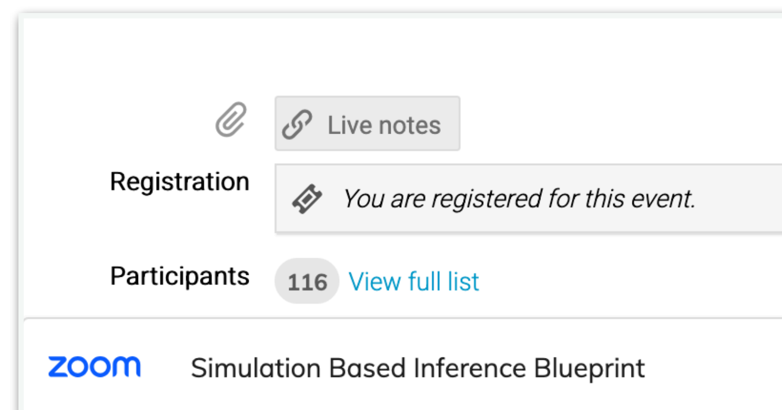
Fermilab

iris hep
Institute for Research & Innovation
in Software for High Energy Physics

Introduction

- Welcome everyone to the IRIS-HEP Blueprint Worksop on Simulation-Based Inference!

- Organizational points:
  - This is a hybrid workshop
  - Please abide by CERN Code of Conduct
  - Workshop live notes: google docs & —>



- Please help us to capture the discussion!

- All the talks and discussions will also be recorded.

Introduction

- Overall vision of Blueprint workshops: Inform the development and evolution of the (IRIS-HEP) software ecosystem strategic vision.

- 2026 Q1 Blueprints:
  - Statistical Ecosystem (24-25 February) indico
  - Simulation-Based Inference (26-27 Feb) indico
  - Differentiable Analysis (5-6 March) indico

All links here: https://indico.cern.ch/category/11329/

Introduction

- Simulation-Based Inference within HEP has a long history, starting with the first CARL paper by Cranmer, et al.

- Core idea: use deep learning to model high-dimensional data for statistical inference.

10 year anniversary of the "CARL" paper

https://arxiv.org/abs/1506.02169

Approximating Likelihood Ratios with Calibrated Discriminative Classifiers

Kyle Cranmer[1], Juan Pavez[2], and Gilles Louppe[1]
[1]New York University
[2]Federico Santa María University

March 21, 2016

https://arxiv.org/abs/1805.00020

A Guide to Constraining Effective Field Theories with Machine Learning

Johann Brehmer,[1] Kyle Cranmer,[1] Gilles Louppe,[2] and Juan Pavez[3]
[1]*New York University, USA*
[2]*University of Liège, Belgium*
[3]*Federico Santa María Technical University, Chile*
(Dated: 30th July 2018)

.....

Introduction

- It is now almost one year since the first paper that used fully differential, high-dimensional data directly to perform hypothesis test using real data from the ATLAS experiment.

- Pheno -> Experiment took almost 10 years!

Rep. Prog. Phys. 88 067801

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Rep. Prog. Phys. 88 (2025) 067801
DOI: 10.1088/1361-6633/add370

CERN-EP-2024-305
June 16, 2025

An implementation of neural simulation-based inference for parameter estimation in ATLAS

The ATLAS Collaboration

Rep. Prog. Phys. 88 057803

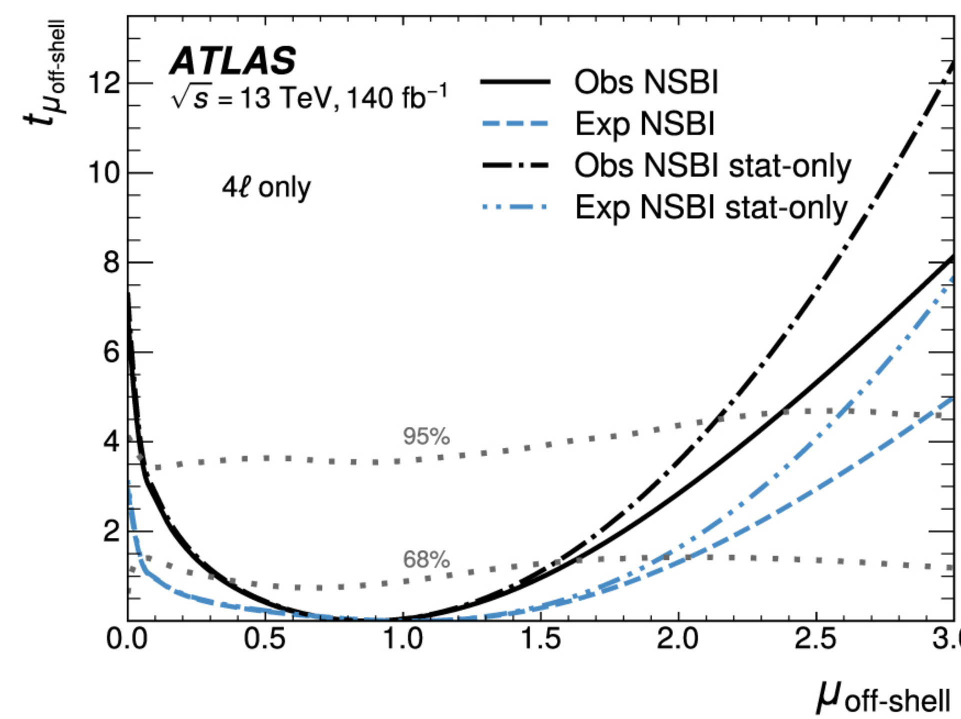
EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Rep. Prog. Phys. 88 (2025) 057803
DOI: DOI:10.1088/1361-6633/adcd9a

CERN-EP-2024-298
May 23, 2025

Measurement of off-shell Higgs boson production in the $H^* \to ZZ \to 4\ell$ decay channel using a neural simulation-based inference technique in 13 TeV pp collisions with the ATLAS detector

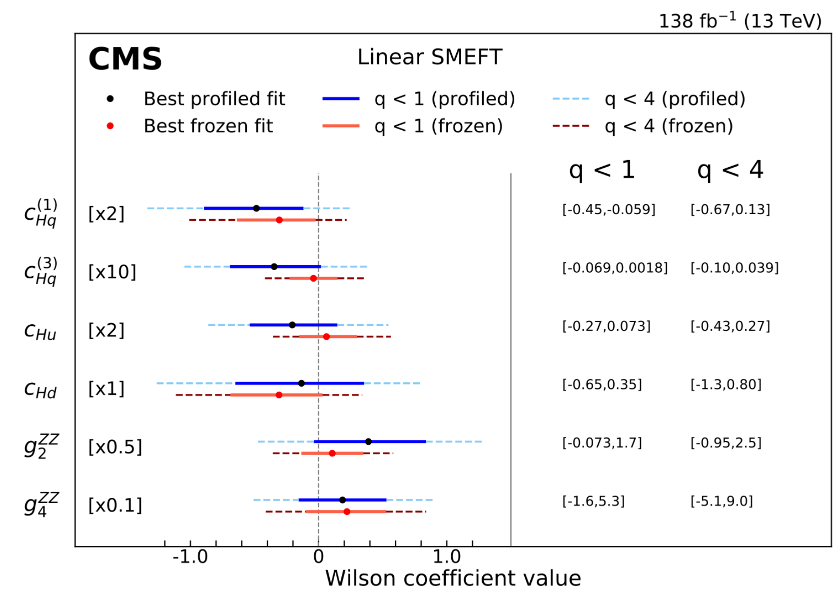
The ATLAS Collaboration

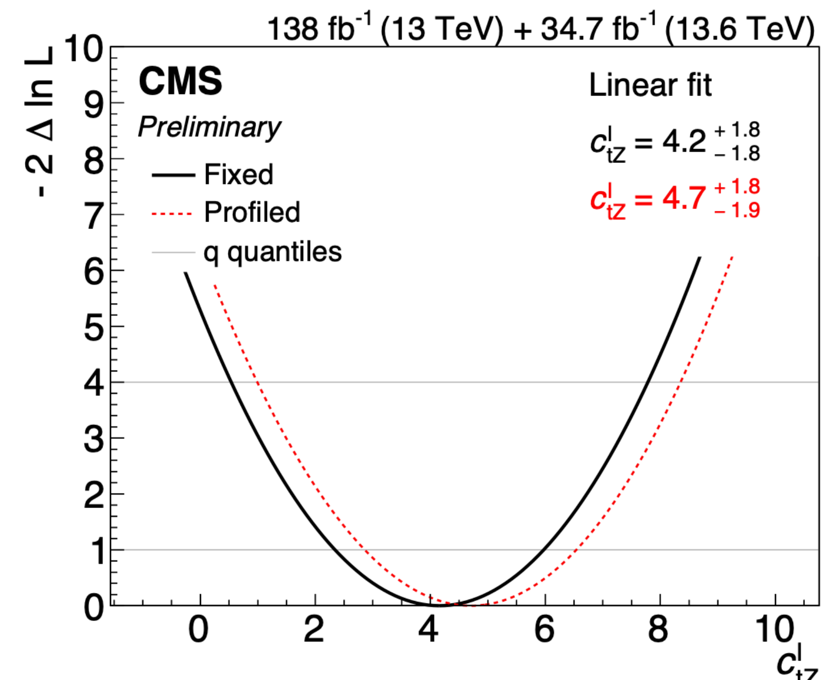


First ever measurement with real data using fully-differential high-dimensional information

Introduction

- Other papers have delved into Simulation-Based Inference, using approaches that convert high-dimensional data into efficient low-dimensional summaries, retaining a lot of the information useful for inference.





Introduction

- Several novel developments go into making Simulation-Based Inference feasible for a full analysis at ATLAS and CMS experiments.

- It also involve learning and adapting from the traditional approaches - especially when tackling the particularly challenging domain problem of huge systematic uncertainty models

13:30 → 15:00 **Statistical Modeling (Systematic Uncertainties)**

13:30 **HistFactory Introduction**
Speaker: Alexander Held (University of Wisconsin Madison (US))

13:50 **HistFactory v2: Systematics Modeling with Gaussian Processes**
Speaker: Jay Ajitbhai Sandesara (University of Wisconsin Madison (US))

14:10 **Unbinned measurements with machine-learned systematic uncertainties**
Speaker: Ricardo Barrué (Marietta Blau Institute for Particle Physics)

14:30 **Discussion**

Introduction

- We will also hear from the community on how the big deep-learning models essential to SBI are trained and can be used for improved techniques.

15:30 → 16:30 **Neural Networks For Likelihood (Ratio) Estimation**

15:30 **Direct Likelihood Learning and Uncertainty Profiling with Normalizing Flows** 20m

Recent advances in Simulation-Based Inference (SBI) often rely on training classifiers to approximate likelihood ratios. However, direct density estimation using Normalizing Flows offers distinct advantages, particularly in the flexibility of the learned statistical model. In this presentation, we explore the use of Normalizing Flows to learn the likelihood function directly to infer physics parameters of interest. Crucially, we address the integration of systematic uncertainties, which is often the bottleneck in precision analyses. We introduce a Factorizable Normalizing Flow architecture that directly learns the conditional dependence of the data on nuisance parameters. Crucially, this allows us to model multiple systematic effects simultaneously without the combinatorial explosion in training cost, enabling efficient profiling without the need for interpolation between alternative likelihood-ratio models.

Speaker: Davide Valsecchi (ETH Zurich (CH))

15:50 **Neural (Quasi-)Probabilistic Likelihood Ratio Estimation** 20m

Speakers: Matthew Drnevich (New York University (US)), Stephen Jiggins (Deutsches Elektronen-Synchrotron (DE))

16:10 **Discussion** 20m

Introduction

- Tomorrow will be dedicated primarily to building toolkits for facilitate making Simulation-Based Inference into the mainstream. Morning will be focused on high-dimensional inference tools

10:00 → 12:00 **Toolkits for Simulation-Based Inference**

10:00 **Toolkit for Simulation-Based Inference (IRIS-HEP)** 30m

Speaker: Jay Ajitbhai Sandesara (University of Wisconsin Madison (US))

10:30 **Q&A** 10m

10:40 **NEEDLE: Workflow Orchestration for Large-Scale NSBI Deployment** 25m

Neural Simulation Based Inference is a fast-moving field with many ongoing efforts to share these promising methods with the wider HEP community. Whilst the specific NSBI methods that will eventually find adoption in actual analyses are not yet known, it is clear that most approaches face a set of common challenges. Foremost, the reliance on many large neural networks that train on large datasets to provide bias-free estimations leads to a considerable computational demand compared to traditional binned approaches.
Firstly, the NEEDLE project aims to provide a flexible framework that takes care of the common boilerplate shared by NSBI tools. This includes standardized model management with PyTorch and Lightning, flexible task-based orchestration with law and experiment tracking. In addition, ready-to-use data ingestion modules written with dask allow for out-of-memory computations for both root and parquet data formats. These modules can be used together in a finished workflow or used individually to accommodate existing code. The goal is to facilitate the deployment of NSBI methods without imposing strong constraints on the specific implementations.
Second, alongside the orchestration infrastructure, a companion benchmarking library to systematically evaluate generative models on toy datasets is being developed. Its goal is to help establish best-practices for the selection and configuration of density estimators. The investigation of powerful and expressive generative models is motivated by their increasing adoption for NSBI in fields such as cosmology and astro-particle physics.
In this contribution, we present an overview of how NEEDLE will provide a standard and flexible framework for NSBI deployment on HPC with minimal constraints, together with a toolbox of new generative models.

Speakers: Kylian Schmidt (KIT - Karlsruhe Institute of Technology (DE)), Levi Evans (Deutsches Elektronen-Synchrotron (DE))

11:05 **Q&A** 10m

11:15 **SBI Toolkit Modification and robust Uncertainty Quantification** 20m

We are exploring the application of the IRIS-HEP Simulation-Based Inference (SBI) toolkit to precision Higgs measurements. In this talk, we discuss methodological developments aimed at improving robustness and of SBI workflows in realistic LHC settings.

On the tooling side, we explore physics-informed inductive biases in neural architectures, energy-conserving optimization schemes as alternatives to Adam, pre-training strategies for improving computing efficiency, and application of SBI to precision cross-section measurements (e.g. OmniFold). On the modeling side, we discuss strategies to mitigate Monte Carlo statistical uncertainties through the wifi ensembling approach.

The goal is to identify practical improvements that strengthen SBI applications in precision LHC physics and to foster collaboration between methodological and experimental communities.

Speaker: Jingjing Pan (KIT - Karlsruhe Institute of Technology (DE))

11:35 **Broader Discussion** 25m

Introduction

- Afternoon will focus on tooling for SBI that involve building efficient summaries of high-dimensional data.

13:30 → 15:00 **SBI with semi-parameterized Density Ratios**

13:30 **Building summaries with event2vec** 20m

Speakers: Nick Smith (Fermi National Accelerator Lab. (US)), Prasanth Shyamsundar (Fermi National Accelerator Laboratory)

13:50 **Parametrized Optimal Observable Approach to NSBI** 20m

The parametrized optimal observable approach is a binned approximation to the full NSBI formalism introduced in [Rep. Prog. Phys. 88 (2025) 067801]. We present the method highlighting it's advantages and limitations. We will show a practical implementation of the parametrized optional observable formalism in RooFit and show how it can be used to construct Asimov datasets and to perform Neyman Construction. We will discuss the challenges faced when trying to implement these ideas with the analysis presented in [Rep. Prog. Phys. 88 (2025) 057803] and ways we found to circumvent them.

Speaker: Matthew Kenneth Maroun (University of Massachusetts (US))

14:10 **Tooling issues for binned nSBI analyses with pyhf and how to overcome them** 20m

Neural estimates of likelihood ratios provide a powerful approach to extending sensitivity across wide regions of phase space, but their integration into full HEP analyses presents significant technical challenges. The computational cost of unbinned neural simulation-based inference (nSBI) can be reduced by performing binned fits using optimal observables - whilst still retaining the benefits from a parameterised observable. However, even in for this approach, commonly used statistical tools such as QuickStats and pyhf introduce practical limitations. In this work, we identify the key technical hurdles encountered in binned nSBI analyses and demonstrate solutions that enable robust and fast-turnaround statistical inference.

Speaker: Malin Elisabeth Horstmann (Technische Universitat Munchen (DE))

14:30 **Discussion** 30m

Introduction

- We end the workshop with more tooling related to large-scale training workflows and then hear about EFT applications before closeout.

15:30 → 16:00 **Scaling Neural Simulation-Based Inference at High Performance Computing for LHC analysis** 30m

This talk will present the preliminary work on scaling the Machine Learning training and Hyperparameter optimization (HPO) at High Performance Computers. We leveraged the Pytorch Lightning framework for distributed training and Ray Tune for HPO. In addition, the ML training framework automatically monitors the model's physics performance by evaluating Neural Simulation Based Inference (NSBI) calibration and closure metrics using goodness-of-fit measures such as χ^2 and Wasserstain distances.

We plan to add NSBI-oriented functionality to the framework including wifi ensembling, ensembling and inference within the HPO loop. The framework currently focuses on ML training but we plan to add downstream tasks like unbinned fit.

Speakers: Walter Hopkins (Argonne National Laboratory (US)), Xiangyang Ju (Lawrence Berkeley National Lab. (US))

16:00 → 16:30 **SBI for Effective Field Theories** 30m

Speaker: Nick Smith (Fermi National Accelerator Lab. (US))

16:30 → 16:45 **Closeout** 15m

Speakers: Jay Ajitbhai Sandesara (University of Wisconsin Madison (US)), Nick Smith (Fermi National Accelerator Lab. (US))

Concluding Remarks

- A hybrid discussion-oriented workshop with 115 participants will be challenging to run successfully. Let's work together:
  - Speakers: please prime discussion on particular elements
    - Consider pausing at appropriate points rather than queueing up discussion for the end?
  - Participants: raise hands, write comments in our live notes, etc.
- The outcome of this workshop will be a community whitepaper (details in a separate mail). What do we want to know?
  - What kind of tools or new developments do you need to start doing SBI? What would a successful SBI ecosystem look like?
  - How can we help in making the ideas presented in the slides to be integrated into tools for the whole community to use? (short-term and long-term)

Have Fun Inference-ing

(Simulation-Based or otherwise)