

Machine Learning Methods for Cross Section Measurements

by

Krish Desai

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Doctor Benjamin Nachman, Co-chair

Professor Uros Seljak, Co-chair

Professor Joshua Bloom

Professor Saul Perlmutter

Spring 2025

Machine Learning Methods for Cross Section Measurements

Copyright 2025

by

Krish Desai

Abstract

Machine Learning Methods for Cross Section Measurements

by

Krish Desai

Doctor of Philosophy in Physics

University of California, Berkeley

Doctor Benjamin Nachman, Co-chair

Professor Uros Seljak, Co-chair

Precise differential cross section measurements are indispensable for testing Standard Model predictions at the energy frontier and searching for new physics, yet their extraction from collider data is an ill-posed inverse problem. *Unfolding* (also known as deconvolution) is the process of removing detector distortions to reconstruct

particle-level truth from detector-level data. Conventional histogram-based, binned unfolding techniques introduce artifacts, impose arbitrary bin edges, and become computationally prohibitive in high-dimensional phase spaces, potentially obscuring underlying physics.

This dissertation develops a unified framework that leverages modern machine learning techniques to surmount these limitations. Beginning with Neural Posterior Unfolding, I demonstrate how conditional normalizing flows can serve as differentiable surrogates of detector response, enabling likelihood-based unfolding through implicit regularization. Building on this foundation, Moment Unfolding directly extracts distribution moments without binning, providing precise experimental predictions for effective field theories and phenomenological models. The framework is further advanced by Reweighting Adversarial Networks (RANs), which perform full spectral unfolding using adversarial training to implement particle-level reweighting steered by detector-level classifiers, offering theoretical and computational advantages over iterative methods. A critical statistical analysis of event correlations in unfolded data reveals systematic misestimation of uncertainties when these correlations are ignored, leading to methodological recommendations that ensure correct coverage for all derived observables.

The methodology is validated using both idealized Gaussian distributions and realistic

particle physics data, including proton–proton collision simulations of the CMS detector. These methods are applied to Z+jets events, demonstrating significant improvements in precision, accuracy, and computational efficiency, achieving orders of magnitude speed up while maintaining unbiased recovery of sharp spectral features.

By marrying statistical rigour with powerful machine learning methods, this work establishes a scalable blueprint for precision measurements at current and future colliders. The resulting open-source software enables more reliable extraction of fundamental physics parameters from complex detector data, advancing our ability to test theoretical models and potentially discover new phenomena in high energy physics experiments.

To name

text

Contents

List of Figures

List of Tables

Acknowledgments

Acknowledgements