=======

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:

Professor Benjamin Nachman, Co-chair
Professor Uros Seljak, Co-chair
Professor Joshua Bloom
Professor Saul Perlmutter

Summer 2025

Machine Learning Methods for Cross Section Measurements

To Ben Nachman,

my advisor, mentor, and friend.

Contents

List of Figures

List of Tables

Chapter I

Introduction and physics background.

I.A Units and conventions.

Even though nature does not establish preferred units, the physics that describes it requires establishing a common language and framework. In high energy physics, the choice of units reflects a not simply an arbitrary choice, but a philosophical stance about the nature of reality. While everyday experience suggests that meters, kilograms, and seconds might be the most suitable units to express the physical world, particle physics reveals that speed, action, and energy form more natural rulers for measuring the universe at its smallest scales.

Definition I.1. *Natural Units* are a system where $\hbar = c = 1$, effectively setting the speed of light and quantum of action as fundamental measuring sticks. This choice transforms all physical quantities into powers of energy.

Some common physical quantities, their abbreviated natural units, their full natural units, their approximate SI equivalents and their physical significance are described in Table I.1.

Table I.1: Combined unit conversions in natural units, showing both abbreviated and full natural unit expressions, their SI equivalents and their physical significance.

Quantity	Abbrev.	Full unit	SI (approx.)	Comment
Speed	1	c	$3 \times 10^8\,\mathrm{m\,s^{-1}}$	Speed of light
Action	1	$\hbar$	$10^{-34}\,\mathrm{J\,s}$	Quantum of action
Energy	GeV	GeV	$1.6 \times 10^{-10}\,\mathrm{J}$	Binding energy
Momentum	GeV	GeV/c	$5.3 \times 10^{-19}\,\mathrm{N\,s}$	Typical HEP particle momentum
Mass	GeV	GeV/c^2	$1.8 \times 10^{-27}\,\mathrm{kg}$	Proton mass
Length	GeV^{-1}	$\hbar c/\mathrm{GeV}$	$2 \times 10^{-16}\,\mathrm{m}$	Compton wavelength
Time	GeV^{-1}	$\hbar/\mathrm{GeV}$	$6.6 \times 10^{-25}\,\mathrm{s}$	
Charge	1	$e/\sqrt{4\pi\alpha}$	$5.3 \times 10^{-19}\,\mathrm{C}$	$\mathrm{e} = 1.6 \times 10^{-19}\,\mathrm{C} \approx 0.3$
Cross section	GeV^{-2}	$\hbar^2c^2/\mathrm{GeV}^2$	$4 \times 10^{-32}\,\mathrm{m}^{-2}$	barn$=2.6 \times 10^{-9}\,\mathrm{GeV}^{-2}$
$\boldsymbol{B}$ field	GeV^2	$\mathrm{GeV}^2/\hbar c^2$	$5 \times 10^{16}\,\mathrm{T}$	

I.A.1 Coordinate systems and relativistic geometry.

At a high energy physics (HEP) experiment, the choice of coordinates must respect the underlying symmetries of the physics and the experimental setup. The laboratory frame typically provides a natural Cartesian system. For example, for a circular collider experiment, one can set the z–axis to

be along the beam direction (longitudinal), the x–axis to be horizontal, pointing toward the centre of the accelerator ring and the y–axis to be vertical, completing the right handed system.

However, the partons that participate in an interaction can carry unknown fractions of the beam momentum, making the centre of mass frame of each interaction unknowable. This uncertainty motivates a coordinate system that transforms simply under longitudinal Lorentz boosts.

Definition I.2. The *pseudorapidity* η, is defined as

$$\eta = -\ln\tan\frac{\theta}{2} \tag{I.1}$$

where θ is the polar angle from the beam axis [1]. Under a longitudinal boost with rapidity $\beta = \operatorname{artanh} v/c$, a massless particle's η transforms as $\eta \mapsto \eta + \beta$

Expressing θ in terms of momentum,

$$\eta = -\ln\tan\frac{\theta}{2} = \frac{1}{2}\ln\frac{|p| + p_z}{|p| - p_z} = \operatorname{artanh}\frac{p_z}{|p|} \tag{I.2}$$

Figure I.1 shows the relationship between pseudorapidity η and true rapidity y for particles with different mass to p_T ratios. As expected, the massless

case ($m/p_T = 0$) lies exactly on the diagonal $y = \eta$, while non-zero m/p_T introduces a systematic deviation that grows at large $|\eta|$. Even a modest ratio ($m/p_T = 0.1$) produces a measurable shift, and by $m/p_T = 2.0$ the rapidity is significantly reduced relative to η. This behaviour must be accounted for when inferring kinematic distributions of heavy particles from detector measurements expressed in pseudorapidity.

For massless particles, differences in η remain invariant under longitudinal boosts, motivating the definition of the angular distance as follows.

Definition I.3. The *angular distance metric* is defined as

$$\Delta R^2 = \Delta\eta^2 + \Delta\varphi^2, \tag{I.3}$$

where φ is the azimuthal angle around the beam axis. This metric approximates the geometric angle between particles in the detector while remaining approximately boost invariant.

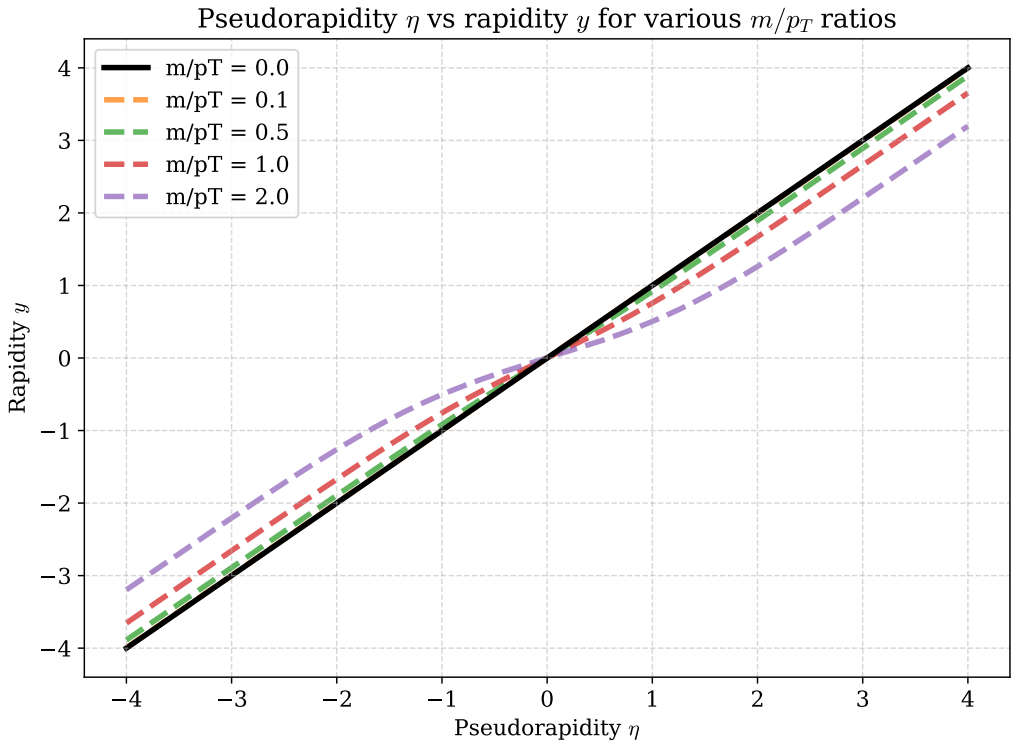

Figure I.1: Comparison of pseudorapidity η and true rapidity $y = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}$ for several mass to transverse momentum ratios. Curves correspond to $m/p_T = 0.0,\ 0.1,\ 0.5,\ 1.0,\ 2.0$, illustrating how finite mass distorts away from the massless limit $y = \eta$ at large $|\eta|$.

I.A.2 Statistical frameworks and uncertainty quantification.

In modern particle physics analyses, every measurement emerges from millions of interaction events, each carrying both statistical and systematic uncertainties.

Definition I.4. The *Poisson distribution* is the fundamental distribution governing counting experiments. For n observed events with expected value λ,

$$\mathbb{P}(n \mid \lambda) = \frac{\lambda^n e^{-\lambda}}{n!} \tag{I.4}$$

In the high statistics limit, $(n \to \infty)$, the Poisson distribution approaches a Gaussian.

Definition I.5. The *profile likelihood ratio*

$$\mathcal{L}(\mu) = \frac{L(\mu, \hat{\theta}_\mu)}{L(\hat{\mu}, \hat{\theta})} \tag{I.5}$$

where $\hat{\theta}_\mu$ maximises L for fixed signal strength μ is the optimal test statistics for hypothesis testing under the Neyman–Pearson lemma.

I.A.3 Machine learning architectures and notation.

Neural networks can be specified by their architecture vectors, $[d_0, d_1, \ldots, d_\ell]$, which denotes a network with input dimension d_0, hidden layers of dimensions d_1 through $d_{\ell-1}$, and output dimension d_ℓ.

For a network $f : \mathbb{R}^{d_0} \to \mathbb{R}^{d_\ell}$ with parameters $\theta = \{W_i, b_i\}$, the forward pass computes

$$h_0 = x \tag{I.6}$$

$$h_i = \sigma(W_i\, h_{i-i} + b_i) \quad i = 1, \ldots, \ell - 1 \tag{I.7}$$

$$f(x; \theta) = W_\ell\, h_{\ell-1} + b_\ell \tag{I.8}$$

where σ denotes the activation function. The universality theorem guarantees that sufficiently wide networks can approximate any continuous function.

Training proceeds via gradient descent on a loss function $L(\theta)$, with the gradient computed through automatic differentiation.

I.A.4 Information theory and optimal observables.

Information theory provides a rigorous framework through the Neyman-Pearson lemma, which implies that the likelihood ratio $\frac{L(x|S)}{L(x|B)}$ provides the most powerful test for distinguishing signal S from background B at any given significance level.

In practice, this optimal observable can be approximated using machine learning classifiers. A well trained classifier computes

$$f(x) \approx \mathbb{P}(S|x) = \frac{L(x \mid S)\,\mathbb{P}(S)}{L(x \mid S)\,\mathbb{P}(S) + L(x \mid B)\,\mathbb{P}(B)]} \tag{I.9}$$

The mutual information $I(Y; f(X))$ between the true labels Y and classifier output $f(X)$ quantifies the information captured

$$I(Y; f(X)) = \iint p(y, f)\, \log \frac{p(y, f)}{p(y)\, p(f)} \mathrm{d}y\, \mathrm{d}f. \tag{I.10}$$

This connects directly to the area under the ROC curve and provides a model-independent measure of classification performance.

I.A.5 Datasets.

Meaningful distinctions in particle physics data arise from the transformation between the underlying physical processes one seeks to understand and the experimental observations one can actually record, as well as from the distinction between Monte Carlo simulation and measurement in nature. Datasets can therefore be classified according to two principal axes that capture the essential distinctions in particle physics data.

The first axis distinguishes between events that occur in nature versus those generated through Monte Carlo simulation. The second axis separates particle level quantities interest from detector level ones.

This classification yields four distinct dataset categories, each serving a unique role in the measurement process. Understanding the relationships

between these categories plays an important role in the development of the unfolding methods discussed in this dissertation.

Figure I.2 illustrates this taxonomy graphically.

I.A.5.i Nature versus Monte Carlo.

The vertical dimension of Figure I.2 distinguishes between natural phenomena and their simulated counterparts. Natural events represent the actual physical processes occurring in particle interactions, while Monte Carlo events are computational reproductions based on theoretical understanding. This distinction is fundamental because although natural events contain true physics, Monte Carlo simulations provide the controlled environment necessary to understand the detector response and develop analysis techniques.

Monte Carlo simulations incorporate theoretical knowledge, including matrix element calculations, parton shower models, and hadronisation prescriptions. While these simulations can achieve remarkable fidelity to natural

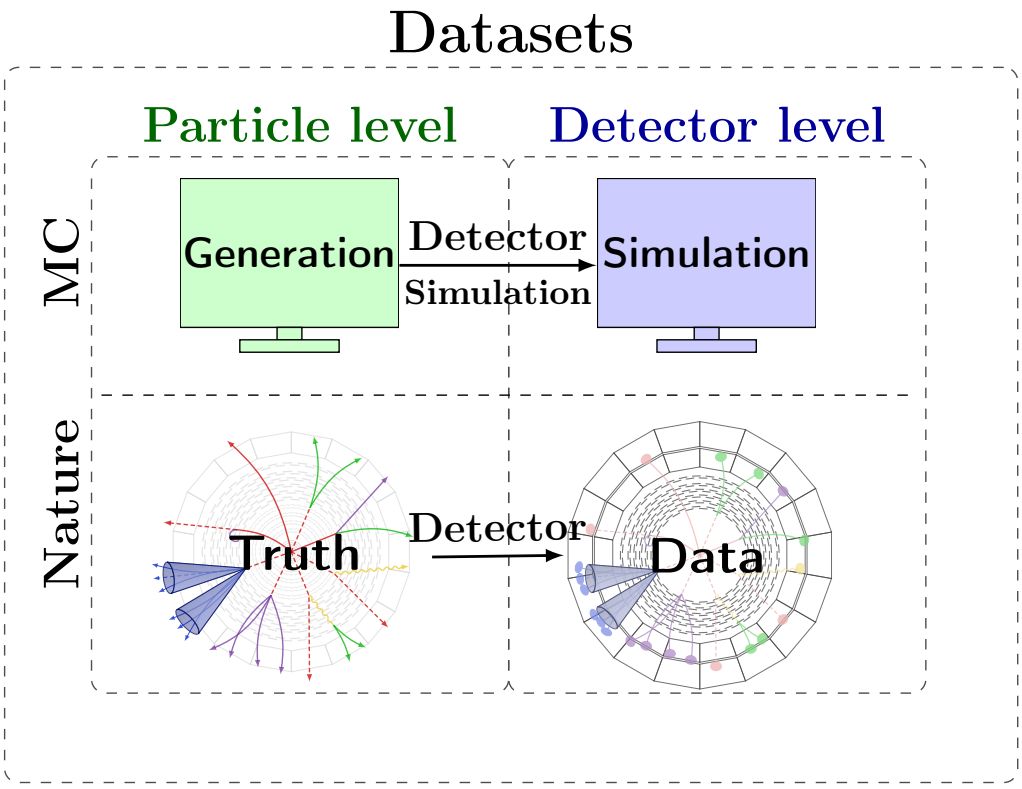

Figure I.2: The four categories of data used in this dissertation organised along two orthogonal dimensions. The vertical axis distinguishes between natural phenomena (Nature) and computational models (MC), while the horizontal axis separates particle level quantities from detector level observations. Arrows indicate the detector response that maps particle level events to their detector level manifestations.

processes, they remain approximations subject to theoretical uncertainties and modelling assumptions.

I.A.5.ii Particle level versus detector level.

The horizontal dimension captures the transformation from particle level information to detector level signals. Particle level quantities represent the properties of particles immediately after the hard scattering process and subsequent decay chains, before interaction with detector. These quantities form the natural language of theoretical predictions and are directly calculable from field theory.

Detector level quantities, in contrast, represent the actual experimental observables after particles have traversed the detector apparatus. This transformation incorporates numerous physical effects. The detector acts as a complex transfer function that irreversibly transforms the particle level distributions into the measured quantities.

I.A.5.iii Four dataset categories.

The intersection of these two dimensions yields four fundamental dataset categories.

Truth represents nature at the particle level—the actual physical distributions HEP experiments ultimately seek to measure. These distributions are never directly observable but represent the ground truth that all analyses attempt to approach. In practice, truth distributions are estimated through unfolding procedures from the information contained in the other three distributions.

Data comprises natural events after detector effects. It is the raw experimental output of all measurements. Data embodies the convolution of the physics of interest with instrumental effects, making it the starting point for any experimental analysis. The challenge lies in disentangling the physics content from the detector effects.

Generation comprises Monte Carlo events at the particle level, providing a controlled means to understanding theoretical predictions. Generation allows the study of particle level distributions with perfect knowledge and unlimited statistics, subject to theoretical modelling uncertainties. Generated events serve as proxies for truth in developing and validating analysis methods.

Simulation represents Monte Carlo events processed through detailed detector modelling, creating artificial detector level datasets with known particle level origins. Simulation enables the study the detector response function by providing matched pairs of particle level and detector level quantities. The fidelity of detector simulation directly impacts the ability to correct for instrumental effects.

I.A.5.iv The detector response.

The arrows in Figure I.2 represent the action of the detector response function, mapping particle level distributions to their detector level counterparts.

This mapping is inherently stochastic, incorporating both resolution effects that blur distributions and acceptance effects that create blind regions in phase space.

For natural events, this transfer function represents the actual physical processes of particle interaction with detector material—ionisation in tracking chambers, showering in calorimeters, and signal processing in readout electronics. For simulated events, the detector simulation models these processes through programs like GEANT4 that track particles through detailed detector geometries, applying the current best understanding of particle matter interactions.

I.A.5.v Operational framework.

In practice, these four dataset categories work in concert to enable precision measurements. Paired Generation–Simulation samples provide the response matrix linking particle and detector levels, while control regions in Data validate the simulation fidelity. Closure tests using independent Monte Carlo

samples verify the unfolding procedure, and the final measurement applies these validated corrections to data to infer the underlying Truth.

This framework implicitly assumes that the detector response is sufficiently similar between data and simulation, an assumption that must be continuously validated through auxiliary measurements and control samples. The robustness of any measurement depends on understanding the limitations of this assumption and properly propagating the associated systematic uncertainties.

The subsequent chapters of this dissertation will repeatedly reference these four dataset categories in the development of sophisticated methods for extracting particle level information from detector level observations. This terminology provides the consistent language necessary for discussing the various approaches to the unfolding problem and their relative merits in different physics contexts.

I.A.6 A note on conventions and clarity.

This work strives to maintain a balance between mathematical rigour and physical insight. Where conventions differ between communities,[1] attempts are made to note both conventions.

[1]For example, particle physicists and machine learning researchers.

I.B The Standard Model: Theoretical framework.

The Standard Model of particle physics represents one of the most significant intellectual achievements in modern science. Developed throughout the latter half of the 20^{th} century, it provides a quantum field theory framework that describes three of the four known fundamental forces—the electromagnetic, weak, and strong interactions—in addition to classifying all known elementary particles. The mathematical formulation of the Standard Model is based on gauge theory, specifically quantum chromodynamics (QCD) and the electroweak theory, underpinned by the gauge symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$.

The predictive power of the Standard Model has been repeatedly validated through precision experiments across multiple energy scales, from low energy nuclear phenomena to the highest energy particle collisions achievable at modern accelerators. Its crowning achievement came with the discovery of the

Higgs boson in 2012 at the Large Hadron Collider (LHC) [2, 3], confirming the mechanism through which elementary particles acquire mass. A schematic illustration of the standard model can be found in Figure I.3.

I.B.1 Fundamental particles and forces.

The Standard Model categorises elementary particles into two main families: fermions, which comprise matter, and bosons, which mediate forces between matter particles.

I.B.1.i Fermions: The building blocks of matter.

Fermions, characterized by half–integer spin, obey the Pauli exclusion principle and satisfy Fermi–Dirac statistics. Fermions are further classified into quarks and leptons, each arranged in three generations of increasing mass.

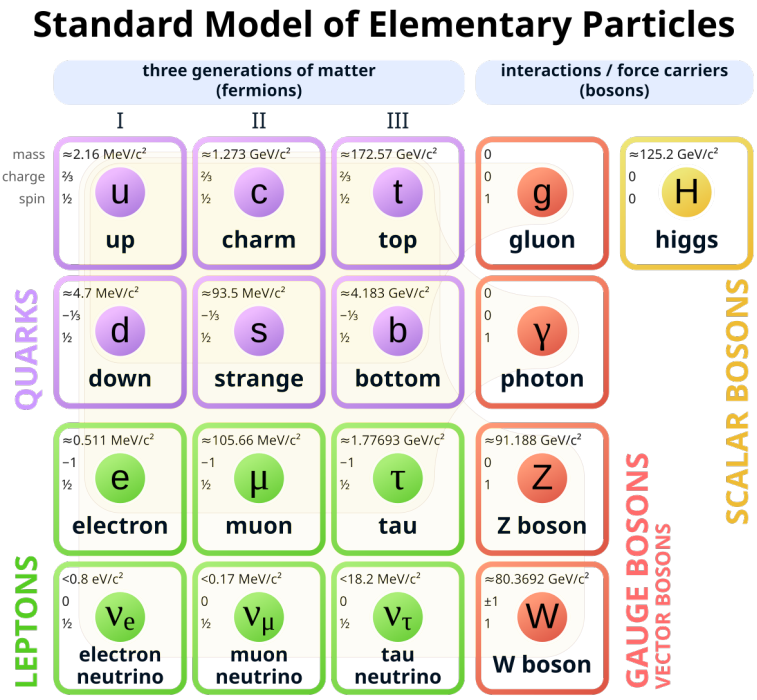

Figure I.3: A schematic illustration of the Standard Model [4].

I.B.1.i.a Quarks.

Quarks are spin = ½ particles that are catagorised into three generations as follows:

- Up (u) and down (d),
- Charm (c) and strange (s),
- Top (t) and bottom (b).

Quarks carry fractional electric charge and colour charge, and experience all fundamental forces. They are confined within hadrons—composite particles categorized as baryons[2] or mesons[3].

I.B.1.i.b Leptons.

Like quarks, leptons too are spin ½ particles that are catagorised into three generations.

[2] three quark states, like protons and neutrons.
[3] quark–antiquark pairs.

- Electron (e) and electron neutrino (ν_e),
- Muon (μ) and muon neutrino (ν_μ),
- Tau (τ) and tau neutrino (ν_τ).

Electrons, muons, and taus carry unit electric charge and interact through the electromagnetic and weak forces, while neutrinos are electrically neutral and interact only through the weak force, making them notoriously difficult to detect.

I.B.1.ii Bosons: Force carriers.

Bosons, with integer spin values, mediate the fundamental interactions. They are not subject to the Pauli exclusion principle and instead satisfy Bose–Einstein statistics. The Standard Model comprises the following bosons:

The photon (γ) is a massless spin–1 boson that mediates the electromagnetic force.

$W^\pm$ **and** Z **bosons** are massive spin–1 bosons that mediate the weak force.

Gluons (g) are a set of eight massless spin–1 bosons that mediate the strong force.

The Higgs boson (H) is a massive spin–0 boson associated with the Higgs field that gives mass to elementary particles.

I.B.2 Theoretical framework and symmetries.

The Standard Model is constructed through principles of quantum field theory where particles are excitations of underlying quantum fields. Its mathematical structure is determined by local gauge invariance under the following specific symmetry transformations:

$\boldsymbol{U(1)_Y}$ is associated with electroweak hypercharge and is the symmetry of electroweak theory,

$\boldsymbol{SU(2)_L}$ describes the weak isospin, and acts on left-handed fermions, and

$\boldsymbol{SU(3)_C}$ governs the strong interactions through colour charge in QCD.

Electroweak unification, demonstrated by Glashow [5], Weinberg [6], and Salam [7], demonstrates how the electromagnetic and weak forces emerge as different aspects of a single electroweak interaction, which undergoes spontaneous symmetry breaking at low energies.

I.B.3 The Higgs mechanism and mass generation.

The Higgs mechanism, proposed by Peter Higgs in the 1960s [8], addresses the theoretical inconsistency of massive gauge bosons in a gauge invariant theory. The mechanism introduces a scalar field—the Higgs field—that permeates space and spontaneously broke the electroweak symmetry when the universe cooled after the Big Bang.

This symmetry breaking generates masses for the W and Z bosons while leaving the photon massless, explaining the significant difference between the electromagnetic and weak forces at ordinary energies. Additionally, the Higgs field couples to fermions through Yukawa interactions [9], generating their masses with coupling strengths proportional to the particle masses.

The discovery of the Higgs boson at the LHC in 2012, with properties consistent with Standard Model predictions, provided crucial experimental validation of this mechanism and completed the Standard Model's particle roster.

I.B.4 Limitations and beyond the Standard Model (BSM) physics.

Despite its remarkable success, the Standard Model has several well recognised limitations, including,

1. It does not incorporate gravity, the fourth fundamental force.
2. It fails to explain the observed matter–antimatter asymmetry in the universe.
3. It does not account for dark matter or dark energy, which together constitute about 95% of the universe's energy content.

4. It requires fine tuning of parameters, raising theoretical concerns like the hierarchy problem.

5. It does not explain neutrino masses, which must exist given observed neutrino oscillations.

These limitations motivate theoretical extensions and experimental searches for physics beyond the Standard Model, including supersymmetry, grand unified theories, and various dark matter candidates. Precision measurements at particle physics experiments provide one of the most powerful approaches to probe these potential extensions, making analysis techniques like those discussed in this thesis essential for advancing our fundamental understanding of nature.

I.C Fundamental role of cross section measurements in particle physics.

Differential cross section measurements are the fundamental currency of scientific exchange in particle physics, serving as the primary bridge between theoretical predictions and experimental observations. These measurements quantify the probability density of specific particle interactions as a function of kinematic variables, providing the essential link between theoretical predictions and experimental observations. A cross section quantifies the probability of a specific particle interaction occurring and is typically expressed in units of area (barns).[4] This seemingly simple concept forms the cornerstone of how we test and validate our understanding of fundamental physics.

The Standard Model makes precise predictions for cross sections that can be directly tested at particle physics experiments. Any statistically significant

[4] 1 barn = $10^{-24}\,\mathrm{cm}^2$.

deviation between measured cross sections and theoretical predictions may signal the presence of new physics beyond the Standard Model [10].[5]

Cross sections are particularly powerful because they encode the underlying quantum field theory structure in a form that can be directly probed by experiment. For instance, measurements of jet production cross sections at different energy scales reveal the running of the strong coupling constant α_S [11], while precision electroweak cross section measurements constrain the properties of the Higgs boson and other fundamental particles [12]. In searches for physics beyond the Standard Model, differential cross section measurements can reveal subtle deviations that point to new particles or interactions, even when direct observation is beyond experimental reach.

These measurements also serve a crucial role in constraining effective field theories (EFTs) that parameterise potential new physics in a model independent way. By measuring differential distributions with high precision,

[5]Such deviations might also signal errors in the theoretical framework used for predictions or in the experimental procedures used to measure the cross section.

experiments can place bounds on EFT coefficients, narrowing the space of viable theoretical extensions to the Standard Model [13].

For example, the ongoing precision program at the Large Hadron Collider (LHC) relies heavily on refined cross section measurements to extract maximum physical insight from collected data. In addition to driving comparisons with theoretical models, cross section measurements are also used at high energy physics experiments for MC tuning [14] and consistency checks [15] among other applications.

I.D Cross section measurements: From theory to experiment.

I.D.1 Theory.

Classically, the cross section (σ) represents the effective area within which two particles must interact for a particular process to occur.

Definition I.6. For collisions between discrete particles, the *cross section* is defined as the area transverse to their relative motion. If the particles were to interact via contact forces (e.g., hard spheres), the cross section corresponds to their geometric size. For long range forces however, the cross section is larger than the physical dimensions of the particles due to action–at–a–distance effects.

Definition I.7. The *differential cross section* ($\frac{d\sigma}{d\Omega}$) provides additional granularity by describing how the probability of scattering depends on specific final state variables, such as scattering angle (θ) or energy transfer. It is

defined as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\text{Number of events scattered into } \mathrm{d}\Omega}{\text{Incident flux} \times \text{Target density}}. \tag{I.11}$$

The total cross section can be recovered by integrating over solid angle:

$$\sigma = \int_{4\pi} \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\, \mathrm{d}\Omega. \tag{I.12}$$

While the classical picture above is intuitive, scattering at HEP experiments is governed by quantum field theory (QFT). In this framework the probability for a process is encoded in a Lorentz invariant matrix element $\mathcal{M}$.

For a $2 \to n$ reaction with incoming $4-$momenta $p_{1,2}$ and final state phase space $\mathrm{d}\Phi_n$, the fully differential cross section is

$$\mathrm{d}\sigma = \frac{(2\pi)^4\, \delta^{(4)}\big(p_1 + p_2 - \sum_{i=1}^{n} p_i\big)}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}}\, |\mathcal{M}|^2\, \mathrm{d}\Phi_n, \tag{I.13}$$

where the denominator is the flux factor and $\mathrm{d}\Phi_n = \prod_{i=1}^{n} \frac{\mathrm{d}^3 p_i}{(2\pi)^3 2E_i}$ is the Lorentz invariant phase space element.[6] Equation (I.13) reduces to the classical area when $|\mathcal{M}|^2$ is replaced by a contact interaction and the final

[6] Standard derivations can be found in [16–19].

state integral collapses to a single kinematic configuration. Integrating Equation (I.13) over final state kinematics yields the total cross section, $\sigma = \int \mathrm{d}\sigma$.

At tree level, $|\mathcal{M}|^2$ is computed from Feynman rules derived from the Lagrangian, while higher order corrections incorporate loops, parton showers, and non-perturbative effects such as hadronisation. For practical experimental predictions predictions one folds $|\mathcal{M}|^2$ with parton distribution functions (PDFs) and convolves the result with detector response—precisely the forward process that the unfolding methods developed in this thesis seek to invert.

Differential cross sections have a long history of providing valuable insights for probing fundamental properties of particles and interactions. Their use dates all the way back to Rutherford's scattering experiments that revealed the existence of atomic nuclei by analysing angular distributions of scattered alpha particles [20].

I.D.2 Experimental Measurement

At modern colliders[7] the two beams themselves act as both "projectile" and "target." The basic experimental quantity is the instantaneous luminosity.

Definition I.8. The *luminosity* $\mathcal{L}(t)$ is defined such that the interaction rate for a process with cross section σ is $\mathrm{d}N/\mathrm{d}t = \mathcal{L}(t)\,\sigma$.

Time integrating over a data taking period $[t_0, t_f]$ yields the integrated luminosity.

Definition I.9. The *integrated luminosity* $\mathcal{L}_{\mathrm{int}}$ is defined as

$$\mathcal{L}_{\mathrm{int}} = \int_{t_0}^{t_f} \mathcal{L}(t)\,\mathrm{d}t, \qquad \sigma = \frac{N_{\mathrm{obs}} - N_{\mathrm{bkg}}}{\mathcal{L}_{\mathrm{int}}\,\epsilon\,A}. \tag{I.14}$$

[7]This section focusses on collider experiments, but similar analyses can be applied to non-collider HEP experiments as well, such as fixed target experiments. [21] and chapters 31 to 38 of [10] provide a comprehensive and detailed exposition of experimental measurement in HEP. [22, 23] focus specifically on fixed target techniques and phenomenology. [24] compares collider and fixed target formalisms, reviews luminosity analogues, and details the role of detector simulations and unfolding in a fixed target context

Here N_{obs} is the number of selected events, N_{bkg} an estimate of background contaminations, ϵ the detector and selection efficiency and A the geometric–plus–kinematic acceptance of the analysis.

For binned measurements one bins events in an observable X[8] and divides by the bin width.

$$\frac{\mathrm{d}\sigma}{\mathrm{d}X}\Big|_{X_i} = \frac{1}{\mathcal{L}_{\text{int}}\,\Delta X_i}\frac{N_i^{\text{obs}} - N_i^{\text{bkg}}}{\epsilon_i\,A_i}, \tag{I.15}$$

with the index i denoting the i^{th} bin.

Luminosity determination is itself a precision measurement, usually performed with dedicated luminometers that exploit van der Meer scans or pileup counting techniques. The efficiency–acceptance term $\epsilon\,A$ is obtained from full detector simulations and corrected in data using control samples and "tag–and–probe" methods.

Equations (I.14) and (I.15) thus link the theoretically calculated parton level cross sections (*vide*Equation (I.13)) to the raw observables recorded by the a detector, completing the chain from theory to experiment.

[8]For example, transverse momentum p_T or rapidity y

I.D.3 Applications in particle physics.

As mentioned above, cross section measurements serve as the fundamental currency of particle physics, translating abstract theoretical predictions into measurable experimental quantities. However, their applications extend far beyond simple theory validation into the operational heart of how experiments function, analyse data, and cross–validate results.

I.D.3.i Theory validation.

As discussed above, comparing measured cross sections with predictions from quantum field theory validates and tests theoretical models like quantum chromodynamics and electroweak theory, by encapsulating interaction probabilities in a measurable form. Deviations from expected cross sections may indicate new phenomena, such as supersymmetric particles or dark matter candidates.

Differential cross sections also provide constraints on effective field theories and parton distribution functions (PDFs), essential for understanding the

internal structure of hadrons. Unfolded cross section measurements allow comparisons with theoretical models years after data collection, even if detector simulations are no longer available, further enhancing their utility, and future proofing the data. Their determination requires careful design and analysis techniques to account for systematic uncertainties introduced by detector effects.

I.D.3.ii Monte Carlo tuning.

Definition I.10. *Monte Carlo tuning* is the iterative process of adjusting simulation parameters to match measured cross sections, ensuring that detector simulations accurately reproduce real experimental data.

Cross section measurements are extensively used in Monte Carlo (MC) tuning.

Every particle physics analysis relies on sophisticated simulations that model everything from the initial parton interactions through hadronisation to the detector response. These simulations contain dozens of phenomenological

parameters, such as the strong coupling constant at various scales and non-perturbative fragmentation functions.

Definition I.11. *Parton distribution functions* (PDFs) are probability distributions that describe the probability of finding a parton (quark or gluon) in a hadron at a given momentum fraction x and scale Q^2. PDFs are determined from global fits to a wide range of hard scattering processes, including deep inelastic scattering, Drell–Yan production, and jet production.

Measured cross sections provide the ground truth that anchors these simulations to reality.

Consider the following example: When CMS measures the inclusive jet cross section at a new centre–of–mass energy, that measurement immediately becomes a crucial input for tuning generators like PYTHIA or HERWIG. The differential distributions—whether in transverse momentum, rapidity, or invariant mass—reveal where the models succeed and where they fail. A discrepancy in the high$-p_T$ tail might indicate a need to adjust the modelling

of initialstate radiation; unexpected structure in angular distributions could point to missing higher order QCD effects.

I.D.3.iii Consistency checks.

Cross–validations between different experimental approaches, detector configurations, or analysis methods ensure measurement reliability and identify systematic biases. Cross sections serve as essential tools for consistency checks across multiple dimensions of experimental physics. Within a single experiment, measuring the same process through different decay channels provides a powerful systematic cross check. For instance, measuring the W boson production cross section through both electronic and muonic decays tests the understanding of lepton universality while simultaneously validating detector calibrations. Any significant deviation signals either new physics, errors in the phenomenological method used, or unaccounted systematic effects.

Between experiments, cross section measurements enable crucial cross–experiment validation. When CMS and ATLAS measure the same process with independent detectors and analysis chains, agreement within uncertainties validates both measurements. Disagreement, conversely, can reveal subtle systematic effects or push calculations to higher precision.

These applications cascade through every level of experimental operations.

I.D.3.iv Luminosity determination.

Luminosity determination becomes possible through processes with well–known theoretical cross sections. Van der Meer scans calibrate the absolute luminosity scale, but elastic scattering and other standard candle processes provide continuous monitoring. The uncertainty on integrated luminosity, typically 2 to 3%, directly impacts every cross section measurement, creating a web of interdependencies.

I.D.3.v Background estimation.

Background estimation in searches for new physics relies on measured cross sections of Standard Model processes. When searching for supersymmetric particles, the irreducible backgrounds from W + jets or $t\bar{t}$ production must be understood at the percent level. Control regions enriched in backgrounds, combined with precise cross section measurements, enable data driven background estimates that would be impossible from simulation alone.

I.D.3.vi Parton luminosity.

Definition I.12. *Parton luminosity* is the effective luminosity for specific parton–parton interactions, calculated by convolving the total luminosity with parton distribution functions.

A particularly elegant application of cross sections emerges in parton luminosity calculations. Since protons are composite objects, the effective luminosity for producing heavy particles depends on the convolution of PDFs with the partonic cross section. Measurements of Drell–Yan production at

different invariant masses directly probe the quark and antiquark distributions, while inclusive jet production constrains the gluon PDF. This creates a self–consistent feedback loop, where better PDFs enable more precise predictions, which enable more sensitive measurements, which further constrain the PDFs.

I.D.3.vii Detector performance validation.

Detector performance validation represents another major application of cross section measurements. Measured cross sections for well understood processes serve as standard candles for monitoring detector stability over time. For example, slow drift in the measured $Z \to \mu\mu$ cross section might indicate degrading muon chamber performance long before it would be noticed in individual event displays. These measurements become part of the experiment's data quality monitoring, flagging problems in real time.

I.D.3.viii Systematic uncertainty evaluation.

The role of cross section measurements in systematic uncertainty evaluation cannot be overstated. Every measurement must account for theoretical uncertainties in signal and background processes. By measuring auxiliary cross sections, for instance, $Z+$ jets production when studying $W+$ jets, experiments can constrain these uncertainties using data rather than relying solely on theoretical estimates. This *in situ* constraint often reduces systematic uncertainties by factors of two or more.

I.D.3.ix Future planning.

Finally, cross sections enable physics program planning for future experiments. The measured production rates at current energies, extrapolated using theoretical calculations, determine required luminosities and detector capabilities for next–generation experiments. For exmaple, the surprisingly large Higgs production cross section at the LHC, for instance, has already influenced design considerations for future electron–positron Higgs factories.

1em In sum, cross sections certainly do function as the Rosetta Stone of particle physics, translating between the languages of theory, simulation, and experimental measurement while maintaining coherence across all three domains. However, they serve not merely as endpoints of analyses, but as the connective tissue that binds together theory, simulation, and experiment into a coherent whole. They simultaneously test theoretical understanding, calibrate experimental tools, and point the way toward new discoveries. This multiplicative utility explains why cross section measurements, even of well studied processes, remain at the heart of every particle physics experiment.

I.E Detector response in precision measurements.

The direct comparison between theoretical predictions and experimental measurements is complicated by detector effects. HEP detectors are technological marvels that capture the trajectories of charged particles, energy deposits in calorimeters, and timing and pattern–recognition information from tracking and particle–identification systems, but they introduce distortions that must be carefully accounted for to extract the true physical distributions of interest. Particle physics detectors represent some of humanity's most sophisticated sensing apparatūs——capturing particle trajectories with silicon sensors operating at liquid helium temperatures, measuring energy deposits in dense calorimeter crystals, and reconstructing vertices with sub–millimetre precision. Yet these technological marvels inevitably introduce systematic distortions that transform the pristine theoretical predictions into the messy reality of experimental data.

Consider the information degradation that occurs in every measurement. Finite resolution creates fundamental blurring, much like how a camera lens distorts an image. The detector's discrete sensing elements can only measure particle energies, momenta, and positions to finite precision, creating an inherent convolution between the true physics distribution and the instrument response function. Geometric acceptance imposes hard boundaries on observable phase space. Particles scattered into the forward beam pipe or extreme backward angles simply vanish from the recorded dataset, creating holes in the measurement that no amount of statistics can fill.

Detection efficiency varies across the detector's active volume, introducing a complex weighting function that depends on particle type, energy, and trajectory. A high–energy muon might traverse the entire detector with near–perfect efficiency, while a low–energy hadron could be absorbed in the first layers of material. Particle misidentification compounds these challenges through cross–contamination between categories, such as when hadronic shower fluctuations cause a pion to masquerade as a kaon [25].

In this way every detector measurement embeds two irreducible probability relationships.

Detection incompleteness $\mathbb{P}(\text{measured} \mid \text{true}) < 1$ True events that fail to be recorded

Measurement impurity $\mathbb{P}(\text{true} \mid \text{neasured}) < 1$ Recorded events that represent contamination

This demonstrates why detector corrections are fundamentally different from simple calibrations. Unlike adjusting a scale that consistently reads, say, 5% high, detector response involves dual information loss that operates asymmetrically.

The first inequality captures the selection bias where certain true configurations have zero probability of detection, creating null spaces in the measurement. The second inequality captures the contamination bias where every reconstruction category contains some fraction of misclassified events.

Background contamination represents a third problem—every measurement category contains some admixture of misclassified events. When hadrons interact in electromagnetic calorimeters, they can mimic electron signatures. When cosmic ray muons traverse the detector during a collision, they contribute to the muon count despite having no connection to the physics of interest. This contamination creates what in signal processing is called the false positive rate.

The mathematical relationship between true particle level distributions and observed detector level measurements follows the convolution integral

$$p(x) = \int r(x \mid z)\, p(z)\, \mathrm{d}z \tag{I.16}$$

Where $p(x)$ is the detector level density, $p(z)$ is the particle level density and $r(x \mid z)$ serves as the response kernel, the conditional probability density that maps each possible true configuration z to the distribution of possible detector measurements x. This kernel encapsulates the entire cascade of finite resolution information degradation.

The response kernel is analogous to the optical transfer function in image processing. It describes how the "lens" of the detector blurs and distorts the perfect theoretical "image". The response function inherently embeds both probabilistic asymmetries. Regions where $\int r(x \mid z)\ \mathrm{d}x < 1$ reveal acceptance holes, i.e. true configurations that produce no detector signal whatsoever. Conversely, the convolution structure itself ensures that multiple truth distributions can yield identical detector observations, creating the degeneracy problem that makes direct inversion impossible.

A central challenge then, for particle physics, is inverting this response kernel to recover $p(z)$ from observed data $p(x)$. This inversion is mathematically ill–posed precisely because of the information loss. Standard matrix inversion fails catastrophically, amplifying statistical noise into wild oscillations that bear no resemblance to the underlying physics.

The resolution requires sophisticated regularisation methods that impose additional constraints such as smoothness assumptions, positivity requirements, and prior knowledge about the expected signal shape. These con-

straints transform the ill–posed inverse problem into a well defined statistical inference challenge, though at the cost of introducing systematic uncertainties that must themselves be carefully validated.

The detector response problem exemplifies a universal pattern in experimental science: the tension between instrumental precision and information preservation. From astronomical imaging through medical diagnosis to HEP measurements, the fundamental trade off between sensitivity and purity governs all attempts to extract signal from noise.

I.F Challenges at modern experiments.

Several challenges at the modern HEP experiments make cross section measurements particularly demanding.

High dimensional phase spaces Modern measurements often involve multiple correlated observables, creating high dimensional distributions that are difficult to analyse with traditional methods.

Limited statistics in extreme regions Rare processes or the tails of distributions often contain valuable physics information but suffer from limited statistics.

Complex detector effects Detectors have non-trivial response functions that can vary significantly across phase space, and are only known implicitly through precision simulations. Their explicit functional form is unknown.

Theoretical uncertainties Precision measurements are increasingly limited by theoretical uncertainties in both signal and background modelling.

Computational constraints Detailed simulation of detector response requires substantial computing resources, limiting the statistical precision of response modelling.

These challenges make the unfolding problem increasingly difficult, particularly as measurements probe more complex final states and differential distributions. For example, measurements of jet substructure, which probe the detailed radiation pattern within collimated sprays of particles, involve observables with complex correlations and detector effects that vary based on jet energy, rapidity, and substructure properties themselves [26–28].

The need for unfolding arises from the fundamental requirement to present results in a detector independent form that can be directly compared with theory predictions or results from different experiments. Without this correction, theoretical interpretations would need to incorporate experiment

specific detector simulations, significantly complicating scientific exchange and theoretical analysis, and inter–experiment comparisons would simply not be possible.

I.G Thesis scope and physics impact.

This dissertation focuses on developing, analysing, and applying novel machine learning methods for cross section measurements in particle physics, with particular emphasis on unbinned approaches that overcome limitations of traditional techniques. The work spans the spectrum from improving binned methods with neural posterior estimation to completely unbinned approaches for both full distributions and statistical moments.

The primary contributions of this thesis include:

1. Development of Neural Posterior Unfolding (NPU), enhancing binned approaches through normalising flows and amortised inference.

2. Introduction of Moment Unfolding, directly deconvolving distribution moments without binning.

3. Creation of Reweighting Adversarial Networks (RAN), a general framework for unbinned spectrum unfolding.

4. Analysis of event correlations in unfolded data and their impact on uncertainty estimation.

5. Investigation of symmetry discovery with SymmetryGAN and its connections to measurement constraints.

These methodological advances address fundamental challenges in experimental particle physics, potentially enhancing the precision and scope of measurements at present and future HEP experiments. These methods can have a wide range of applications in particle physics, including:

- Improved precision in jet substructure measurements, enabling better discrimination between different theoretical models of QCD radiation.

- Enhanced sensitivity to effective field theory parameters by directly deconvolving distribution moments.

- More robust uncertainty quantification in high dimensional measurements.

- Computational efficiency gains allowing for more detailed systematic studies.

- A rigorous statistical framework for incorporating detector response uncertainties in the unfolding process.

By bridging sophisticated machine learning techniques with the specific requirements of particle physics measurements, this work aims to advance the ability to extract fundamental physical insights from complex experimental data. The methods developed here have applications beyond particle physics, potentially benefiting any field where deconvolution of instrumental effects is necessary for scientific inference.

Chapter II

Theoretical foundations.

II.A Statistical formulation of the unfolding problem.

Unfolding, also known as deconvolution, is the process of correcting detector distortions in experimental data to recover the true particle level distributions. This procedure is critical for comparing experimental results with theoretical predictions and for enabling detector independent analyses. The unfolding problem is inherently statistical and presents unique challenges due to its ill posed nature.

II.A.1 The detector response and forward problem.

The relationship between the particle level truth distribution $p(z)$ and the detector level measured distribution $p(x)$ is governed by the detector response function $r(x \mid z)$, which encapsulates the resolution effects.[1] Equation (I.16) describes the **forward problem**, where the true distribution $p(z)$ is mapped

[1] Efficiency and acceptance effects can also be incorporated if "empty" events are allowed.

to the measured distribution $p(x)$. The detector response function $r(x \mid z)$ can often be estimated through detailed simulations.

II.A.2 The inverse problem: Unfolding.

The goal of unfolding is to invert the forward problem and estimate the Truth, $p(z)$ from Data $p(x)$. Mathematically, this requires solving

$$p(z) = \int r^{-1}(z|x)\, p(x)\, \mathrm{d}x, \tag{II.1}$$

where $r^{-1}(z \mid x)$ represents the inverse response kernel. However, this inversion is ill posed because small fluctuations in $p(x)$ can lead to large variations in $p(z)$ [29]. Regularisation techniques are therefore essential to stabilise the solution.

II.A.3 Likelihood based formulation.

In practice, unfolding is performed using statistical inference methods. Given a set of measured data $\mathbf{X}_{i=1}^{N}$, the likelihood function for a proposed

truth distribution $p(z;\theta)$, parametrised by θ, is

$$\mathcal{L}(\theta;\mathbf{X}) = \prod_{i=1}^{N} p(x_i;\ \theta), \tag{II.2}$$

where

$$p(x;\ \theta) = \int r(x \mid z)\ p(x;\ \theta)\,\mathrm{d}z. \tag{II.3}$$

Maximising this likelihood yields an estimate of the parameters θ, which define the unfolded truth distribution. Regularisation can be incorporated into this framework by adding penalty terms to the likelihood or by constraining the parameter space.

II.A.4 Regularisation techniques.

Regularisation mitigates the instability of unfolding by imposing constraints on the solution. Common approaches include Tikhonov Regularization [30], in which one adds a penalty term proportional to the norm of the second derivative of $p(z)$, enforcing smoothness, and iterative methods which gradually refine estimates of $p(z)$, regularising by stopping before convergence.

These techniques balance fidelity to the measured data (prior independence) with stability of the unfolded solution.

II.A.5 Challenges in high dimensional phase spaces.

II.A.5.i Binned methods.

A significant challenge that traditional unfolding methods face is the *curse of dimensionality*. As the number of measured observables increases, the statistical power required to populate discrete bins grows exponentially, quickly overwhelming even the largest datasets collected at modern experiments.

Consider a measurement involving just four kinematic variables, say, transverse momentum, pseudorapidity, azimuthal angle, and invariant mass. With a modest 20 bins per dimension, the resulting joint histogram requires 1.6×10^5 bins. Most of these bins will contain zero events, creating a sparse matrix that renders traditional unfolding techniques numerically unstable.

Theorem II.1 (Scaling of relative uncertainty)**.** Consider a d-dimensional measurement space partitioned into histogram bins of widths $\Delta z_1, \ldots, \Delta z_d$, and let N_{events} denote the total number of events uniformly distributed over the full measurement range. Then the relative uncertainty in the estimated differential event density in any given bin obeys

$$\text{Rel. uncertainty} = \frac{\sigma_Q}{Q} \propto \frac{1}{\sqrt{N_{events}} \prod_{i=1}^{d} \Delta z_i} \tag{II.4}$$

where Q is the differential quantity (events per unit hypervolume).

Proof. A single histogram bin occupies hypervolume

$$V = \prod_{i=1}^{d} \Delta z_i \tag{II.5}$$

in the measurement space. Under the uniform–density assumption, the expected number of events in that bin scales as

$$N_{bin} \propto N_{events}\, V. \tag{II.6}$$

By Poisson statistics the absolute uncertainty on the bin count is

$$\sigma_{N_{bin}} = \sqrt{N_{bin}}, \tag{II.7}$$

and hence the relative uncertainty on N_{bin} is

$$\frac{\sigma_{N_{bin}}}{N_{bin}} = \frac{1}{\sqrt{N_{bin}}} \propto \frac{1}{\sqrt{N_{events}\, V}} \, . \tag{II.8}$$

Since the differential measurement is $Q = N_{bin}/V$, its relative uncertainty coincides with that of the bin count.

$$\frac{\sigma_Q}{Q} = \frac{(\sigma_{N_{bin}}/V)}{(N_{bin}/V)} \tag{II.9}$$

$$= \frac{\sigma_{N_{bin}}}{N_{bin}} \tag{II.10}$$

$$\propto \frac{1}{\sqrt{N_{events}}\, \prod_{i=1}^{d} \Delta z_i} \, , \tag{II.11}$$

which establishes the claimed scaling. □

This relationship reveals why traditional binned methods become increasingly impractical as dimensionality increases: maintaining fixed statistical precision requires exponentially more data or exponentially coarser binning,

both of which severely limit the measurement's resolving power. Beyond a point, the *response matrix*, $\mathbf{R}$, becomes not just ill conditioned but genuinely rank deficient, as sufficiently many bins have not just few events in them, but rather no events in them at all. Hence, entire swaths of phase space remain unmeasurable regardless of accumulated statistics.

Binning artifacts compound these statistical challenges through systematic information loss. The process of discretising continuous distributions into finite bins introduces artificial boundaries onto the underlying physics. This discretisation becomes particularly pernicious when dealing with correlation structures that span multiple dimensions. Real particle interactions generate complex kinematic relationships—the angular distribution of decay products correlates with their energies, jet substructure variables depend on the overall jet momentum, and detector response functions couple seemingly independent observables. Binning destroys these correlations by requiring low dimensional projections.

For instance, in jet substructure measurements, the relationship between different substructure variables contains valuable information about the underlying physics that can be obscured by independent binning of each variable. Moreover, many theoretical predictions in particle physics are at the level of statistical moments or other distribution properties rather than full differential spectra. Traditional unfolding methods require first unfolding the full distribution and then calculating these properties, which can lead to reduced precision in the moment predictions.

The computational burden grows even faster than the statistical requirements. Matrix inversion algorithms scale as $\mathcal{O}(N^3)$ with bin number, meaning that the aforementioned four dimensional histogram with 1.6×10^5 bins requires approximately $\mathcal{O}(10^{16})$ floating point operations to invert. More fundamentally, the *null spaces* and *degeneracies* that plague low dimensional unfolding become dramatically worse in higher dimensions, where multiple truth configurations can project to identical detector signatures along numerous measurement axes simultaneously, because high dimensional mea-

surements create complex geometric relationships between true and observed phase spaces.

In addition to these limitations, traditional binned approaches suffer from an additional systematic weakness; the response matrix $\mathbf{R}$ depends on numerous variables that are typically marginalised over in order to bin the data. While conventional methods construct $\mathbf{R}$ based on a limited set of binned observables, the true detector response depends on a much richer set of event characteristics. These include additional kinematic variables not captured in the chosen binning scheme, event level properties such as particle multiplicity and missing energy, detector conditions like instantaneous luminosity and pileup activity, and correlations with other particles produced in the same collision.

Traditional unfolding methods are effectively forced to assume that $\mathbf{R}$ remains constant when averaged over these marginalised variables, but this assumption is patently incorrect when the detector response varies systematically across different regions of this extended phase space. For example,

the energy resolution for jets may depend not only on the jet's transverse momentum and pseudorapidity, variables often chosen for binning, but also on the jet's substructure, the presence of nearby particles, and the overall event topology. By marginalising over these features, binned methods introduce systematic biases that can propagate through the unfolding procedure and distort the final measurements in ways that are difficult to quantify or correct.

II.A.5.ii Unbinned methods.

Unbinned approaches emerge as a natural response to these challenges, operating directly on individual events rather than aggregated histogram counts. These methods exploit the event level structure that binning destroys, preserving the full kinematic information and correlation patterns within each recorded data point. Instead of discretising phase space into predetermined categories, unbinned techniques allow the data itself to determine the relevant resolution scales and correlation structures.

Rather than explicitly modelling the response function $r(x \mid z)$, modern techniques like OmniFold [31] employ iterative reweighting strategies that learn implicit mappings between particle level and detector level distributions. These methods sidestep the curse of dimensionality by avoiding explicit probability density estimation, instead focusing on weight optimisation that preserves marginal distributions while respecting the detector response.

The computational complexity shifts from matrix algebra to optimisation landscapes, that scale more favourably with dimensionality. While traditional unfolding requires inverting matrices that grow exponentially with dimension, unbinned methods typically employ gradient based optimisation that scales polynomially. This trade off exchanges the well understood numerical properties of linear algebra for the more complex but ultimately more scalable challenges of machine learning optimisation.

It is important to acknowledge that this transition introduces new sensitivities to model and hyperparameter choices that traditional methods avoid. The implicit effect of the learning dynamics make it crucial to understand

how optimisation biases influence the final results. These considerations establish the foundation for exploring how modern machine learning approaches navigate these challenges while preserving the statistical rigour that particle physics demands.

II.B Forward and inverse problems in HEP.

The measurement process in high energy physics experiments inherently involves two complementary mathematical challenges: the forward problem of predicting detector responses from the theoretical understanding of particle level interactions, and the inverse problem of recovering true physics distributions from observed detector measurements. These twin challenges form the conceptual foundation for understanding detector effects and developing unfolding methodologies.

II.B.1 Mathematical formulation.

The relationship between particle level truth distributions and detector level observations is governed by the Fredholm integral equation of the first kind [32],

$$p(x) = \int r(x \mid z)\, p(z)\, \mathrm{d}z + \epsilon(x), \tag{II.12}$$

where $p(z)$ represents the true particle level distribution, $r(x \mid z)$ is the detector response kernel encoding resolution effects and acceptance, $p(x)$ is the observed detector level distribution, and $\epsilon(x)$ accounts for measurement noise [33]. This equation encapsulates the *forward problem* when predicting $p(x)$ given $p(z)$, and the *inverse problem* when estimating $p(z)$ from data $p(x)$.

For discrete histogram representations, this becomes a matrix equation,

$$\boldsymbol{\mu} = \mathbf{R}\boldsymbol{\nu} + \boldsymbol{\epsilon}, \tag{II.13}$$

where $\boldsymbol{\nu}$ and $\boldsymbol{\mu}$ are vectors of true and observed bin counts, respectively, $\mathbf{R}$ is the smearing matrix containing conditional probabilities, $R_{ij} = \mathbb{P}(\text{observed bin } i \mid \text{true bin } j)$ [34], and $\boldsymbol{\epsilon}$ is the vector of measurement noise.

II.B.2 Challenges in inverse problems.

The inverse problem in HEP is fundamentally and intrinsically ill posed. The response kernel is non-injective, i.e. different true distributions can

produce identical observed distributions after detector smearing [33, 35, 36]. and, the distributions are ill conditioned; small measurement errors $\boldsymbol{\epsilon}$ amplify into large fluctuations in unfolded solutions due to small singular values in $\mathbf{R}$ [37–39]. These challenges, intrinsic to inverse problems, are compounded by the fact that modern analyses involve a large number of observables, making brute force phase space discretisation computationally prohibitive [40–42]. The curse of dimensionality is particularly pronounced in high energy physics, where the number of observables can be in the tens or hundreds. Addressing these challenges necessitates regularisation techniques that impose physical constraints on solutions, classic examples of which are Tikhonov regularisation or iteration cut-offs before convergence, described in Section II.C.

II.B.3 HEP specific considerations.

Three aspects particularly complicate unfolding in particle physics compared to other inverse problem domains. First, the response matrix $\mathbf{R}$ is

estimated from from detailed Monte Carlo simulations[2] that encode complex, non-Gaussian systematic uncertainties through their modelling assumptions [45–48].

Second, detector response models contain hundreds to thousands of correlated nuisance parameters—including jet energy scale, $b-$tagging efficiencies, and pile up effects—that must be profiled or marginalised alongside the unfolding procedure [49–52].

Finally, the discrete nature of particle counting combined with highly variable event rates across phase space creates a challenging statistical landscape where Poisson uncertainties dominate in tails and signal regions, while Gaussian approximations may be valid elsewhere [53–56].

A representative unfolding example is differential jet substructure measurements, where detector effects smear the true distribution of observables like jet mass or $N-$subjettiness. The forward problem involves simulating jets through hadronisation models and detector response, producing a migration

[2]E.g., Pythia [43] for hadronisation, Geant4 [44] for detector physics.

matrix that relates true and reconstructed substructure observables. The inverse problem requires unfolding these observables from reconstructed jets while accounting for correlated uncertainties in jet energy scale, angular resolution, and pile up contamination [57].

Similar challenges arise in unfolding differential cross sections as functions of transverse momentum, rapidity, or invariant mass, where detector acceptance and resolution vary significantly across the measurement range [58].

II.C Historical development: From matrix inversion to modern approaches.

The problem of unfolding has a rich history in high energy physics, with methods evolving alongside computational capabilities and statistical sophistication. Early approaches relied primarily on simple correction factors applied to individual bins of histograms, appropriate only when detector effects were minimal.

As measurements became more precise, regularised matrix inversion techniques emerged as the standard approach. These methods discretise both the particle level and detector level distributions into bins, relating them through a response matrix R_{ij} that describes the probability for an event in particle level bin j to be observed in detector level bin i. In component form, Equation (II.13) can be written as

$$\mu_i = \sum_j R_{ij}\nu_j + \epsilon_i, \tag{II.14}$$

where μ_i is the expected number of counts in detector level bin i, ν_j is the expected number of events in particle level bin j, and ϵ_i is the measurement noise in detector level bin i. Naively, one might attempt to solve this system by simply inverting the response matrix:

$$\nu_j = \sum_i R_{ji}^{-1} \mu_i \tag{II.15}$$

However, this direct inversion leads to wildly oscillating solutions with large variances—a manifestation of the ill posed nature of the unfolding problem. Additionally, $\mathbf{R}$ need not, and often is not, a square matrix, and so *the* inverse is not even well defined.

To address this issue, a series of techniques were developed to invert Equation (II.13) by imposing additional constraints.

Iterative Bayesian unfolding [59–61] , also known as Lucy–Richardson deconvolution, uses Bayes' theorem to iteratively update the estimate of the true distribution, with the number of iterations controlling regularisation strength.

SVD unfolding [62] applies singular value decomposition to the response matrix and suppresses contributions from small singular values that amplify statistical fluctuations.

TUnfold [63] formulates unfolding as a least squares problem with Tikhonov regularisation to penalise large second derivatives, preserving smoothness.

These methods, which will be discussed in more detail in Section II.D, have served the field well for decades, particularly for one dimensional measurements where binning is manageable. However, they all share the common limitation of requiring discretisation of the underlying distributions, which becomes increasingly problematic as measurements probe higher dimensional spaces and more complex observables.

II.D Traditional unfolding methods in experimental analyses.

Traditional unfolding methods form the bedrock of detector corrections in high energy physics, balancing statistical rigour with computational practicality. This section provides an overview of established techniques, their mathematical foundations, implementation nuances, and limitations.

II.D.1 Bin by bin correction.

The simplest unfolding approach applies multiplicative correction factors to observed bin counts.

$$\hat{\nu}_j = \frac{\mu_j - b_j}{C_j}, \quad C_j = \frac{\nu_j^{\mathrm{MC}}}{\mu_j^{\mathrm{MC}}}, \tag{II.16}$$

where μ_j is the observed count in bin j, b_j the estimated background, and C_j the correction factor derived from Monte Carlo (MC) simulations relating particle–level generated (ν_j^{MC}) and detector–level simulated (μ_j^{MC}) events [64].

This method has the advantage of being computationally trivial, with no bin–to–bin correlations. However, it fails to account for a non–diagonal response matrix ($\exists i \neq j \; R_{ij} \neq 0$). This is illustrated most dramatically by the observation that biases persist even with $C_j \to 1$ due to ignored cross–bin migrations [65].

Used primarily in early LHC analyses[3], bin–by–bin correction remains viable only for coarse binnings with negligible migration ($\lesssim 5\%$ [34]) between adjacent bins.

II.D.2 Matrix Inversion

When $n_{\text{bins, Gen.}} = n_{\text{bins, Sim.}}$, the response matrix $\mathbf{R}$ is square. Formally one can write an unfolded solution as

$$\hat{\boldsymbol{\nu}} = \mathbf{R}^{-1}\boldsymbol{\mu}, \tag{II.17}$$

and even propagate the covariance as

$$\Sigma_{\hat{\boldsymbol{\nu}}} = \mathbf{R}^{-1}\Sigma_{\boldsymbol{\mu}}(\mathbf{R}^{-1})^T. \tag{II.18}$$

[3] e.g., ATLAS [66, 67] jet cross-sections

However, in practice, direct inversion is highly pathological. This pathology can be quantified by the condition number

$$\kappa(R^{-1}) = \frac{|\lambda_{\max}(R^{-1})|}{|\lambda_{\min}(R^{-1})|} \sim 10^3 - 10^6 \tag{II.19}$$

where $\lambda_{\max}(R^{-1})$ and $\lambda_{\min}(R^{-1})$ are the largest and smallest eigenvalues of R^{-1} respectively. The condition number measures how much a perturbation in the measured counts $\delta\mu$ perturbs the predicted truth counts $\delta\nu$. The large condition number amplifies statistical fluctuations [68, 69]. Further, unphysical solutions such as negative bin count values can arise from noise dominated eigenvectors.

Methods have been suggested to control this variance, such as Truncated SVD [70], involving discard singular values $\sigma_i < \lambda_{\text{cut}}$ [71], and Wiener–SVD [72], a frequency domain filtering method to maximise signal–to–noise ratio [73]. Despite these proposals, matrix inversion's instability limits its utility.

II.D.3 Iterative Bayesian unfolding.

Bayesian methods regularise through prior distributions $p(z)$, yielding posterior estimates.

$$p(z \mid x) \propto \mathcal{L}(x \mid z)\, p(z) \tag{II.20}$$

The initial guess $\boldsymbol{\nu}^{(0)}$ biases the solution. Common priors include Generation, $\boldsymbol{\nu}_{\mathrm{MC}}$, a uniform distribution, and data driven backwards folding $\mathbf{R}^T\boldsymbol{\mu}$. The entropy maximisation prior, $p(z) \propto \exp(-\sum z_j \ln z_j)$ [74] and Gaussian processes for enforcing smoothness [45] have also been proposed as suitable priors for IBU. These provide natural uncertainty quantification but suffer from high computational cost, scaling poorly with dimensionality [75],sensitivity to prior misspecification, especially in low-statistics regions [76], and difficulty interpreting credible intervals as frequentist coverage [77].

IBU, also known as Lucy–Richardson deconvolution or D'Agostini iterative unfolding [78] is an expectation maximization (EM) algorithm iteratively updates truth estimates through the procedure described in Algorithm 1.

Algorithm 1 Iterative Bayesian unfolding (Richardson–Lucy deconvolution/ D'Agostini iteration)

Require: Response matrix R_{ij},
histogram counts of Data, μ_i,
initial estimate $\nu_j^{(0)}$,
iterations K,
bins $N_{\text{Sim.}}$, $N_{\text{Gen.}}$

Ensure: Unfolded (Truth) distribution $\nu_j^{(K)}$

1: **for** $k \leftarrow 0$ **to** $K-1$ **do**
2: **for** $j \leftarrow 1$ **to** $N_{\text{Gen.}}$ **do**
3:
$$\nu_j^{(k+1)} \leftarrow \nu_j^{(k)} \sum_{i=1}^{N_{\text{Sim.}}} \frac{R_{ij}\,\mu_i}{\sum_{\ell=1}^{N_{\text{Gen.}}} R_{i\ell}\,\nu_\ell^{(k)}} \tag{II.21}$$
4: ▷ *Normalisation to enforce* $\sum_j \nu_j = 1$
5: **for** $j \leftarrow 1$ **to** $N_{\text{Gen.}}$ **do**
6: $\nu_j^{(k+1)} \leftarrow \dfrac{\nu_j^{(k+1)}}{\sum_{m=1}^{N_{\text{Gen.}}} \nu_m^{(k+1)}}$
7: **return** $\left\{\nu_j^{(K)}\right\}_{j=1}^{N_{\text{Gen.}}}$

This method regularises via early stopping, by terminating at $K \sim 4-6$ iterations before noise amplification [79, 80].

While computationally efficient, this approach lacks objective stopping criteria, requiring heuristic cross–validation [81] and underestimates uncertainties due to ignored iteration dependent covariance [82]. IBU is dominant in LHC analyses[4] because it balances simplicity, interpretability and computational efficiency with accuracy.

II.D.4 Tikhonov Regularisation

Tikhonov regularisation is a penalised least–squares minimisation method

$$\hat{\boldsymbol{\nu}} = \arg\min_{\boldsymbol{\nu}} \left[||\boldsymbol{\mu} - \mathbf{R}\boldsymbol{\nu}||^2 + \lambda ||\mathbf{L}(\boldsymbol{\nu} - \boldsymbol{\nu}_0)||^2 \right] \tag{II.22}$$

where $\mathbf{L}$ is typically the discrete curvature operator[5] and $\boldsymbol{\mu}_0$ a prior estimate [84] that anchors solutions to MC predictions. L–curve optimisation balances the residual norm against the solution norm to choose λ [85]. The

[4] E.g., differential jet substructure measurements [83]

[5] E.g., discrete second derivatives.

choice of λ sets the bias variance trade off. $\lambda \to 0$ represents the high variance, low bias limit and $\lambda \to \infty (\implies \hat{\boldsymbol{\nu}} \to \boldsymbol{\nu}_0)$ represents the low variance, high bias limit. This method is implemented through the TUnfold package, which also provides automated λ tuning via global correlation minimisation [63]. However Tikhonov regularisation struggles with non-differentiable features like threshold effects due to biased curvature penalties, and requires *ad hoc* λ selection via L–curve curvature maximisation.

Note. The RooUnfold package [86] provides implementations of bin–bin–bin corrections, matrix inversion, IBU, SVD, and TUnfold.

II.D.5 Template fitting.

Template fitting is a method suitable in cases where $N_{\text{Sim.}} \gg N_{\text{Gen.}}$. In this case, one can construct detector level templates for each particle level bin.

$$\mu_i = \sum_{j=1}^{N_{\text{Gen.}}} R_{ij}\nu_j + b_i \tag{II.23}$$

with χ^2 minimisation:

$$\chi^2 = \sum_{i=1}^{N_{\text{Sim.}}} \frac{(\mu_i - \sum_j R_{ij}\nu_j - b_i)^2}{\sigma_i^2} \tag{II.24}$$

The solution then is over-constrained through the use of ${}^{N_{\text{Sim.}}}/_{N_{\text{Gen.}}} \sim 2-3$ for stability [87]. Nuisance parameters are systematically modelled in this approach via template morphing [88].

Template fitting requires dense detector level binning, which inflates statistical uncertainties. This method is commonly used in Higgs coupling measurements where broad mass resolutions necessitate wide truth bins.

II.D.6 Regularised Poisson likelihood.

For low statistics regions, [41] advocates minimising

$$-\log \mathcal{L}(x \mid z) + \lambda S(z). \tag{II.25}$$

$S(z)$ penalises non monotonicity in sharply falling spectra. Using cubic B–splines with entropy regularisation, this method avoids binning artifacts

Method	MC dependence	Uncertainty propagation
Bin–by–bin	Extreme	Underestimated
Matrix inversion	None	Exact but unstable
IBU	Moderate	Partial
Tikhonov	Moderate	Full
Template fit	Low	Full
Regularised Poisson	Moderate	Full

Table II.1: Comparing the MC dependence and uncertainty propagation of traditional unfolding methods.

through continuous representations [41]. However, it requires careful basis function placement to prevent endpoint spikes [89] and demands specialised optimization protocols (e.g., cooling schedules for λ [90]).

II.D.7 Summary.

Table II.1 summarises the properties and limitations of the methods discussed above. These limitations motivate the use of machine learning in unfolding, a transition explored in subsequent sections. However, traditional

methods remain indispensable for validation and low dimensional precision measurements where interpretability is crucial.

II.E Regularisation: Need, approaches, and limitations.

The inherent ill posedness of unfolding necessitates regularisation to stabilise solutions against statistical fluctuations while preserving physical meaning. This section systematically examines the theoretical justification for regularisation, surveys dominant methodologies, and critically evaluates their strengths and limitations in high energy physics applications.

II.E.1 Necessity.

As discussed earlier, inverse problems in HEP exhibit pathological characteristics that demand regularisation. Regularisation counteracts unphysical variance by introducing prior knowledge about $p(z)$, typically favouring smoothness or similarity to Monte Carlo predictions. However, this unavoidably discards information—regularised solutions cannot resolve features finer

than the detector resolution or distinguish theories predicting distributions within the regularisation bias [91, 92].

II.E.2 Limitations and practical challenges.

II.E.2.i Subjectivity–objectivity trade off.

All regularisation methods inject subjective choices—smoothness scales, prior distributions, stopping criteria and so on——that bias results. This trade off reveals a deeper epistemological issue: regularisation transforms the question from "what does the data show?" to "what does the data show given our assumptions about smoothness?"

While [93, 94] argue for publishing unregularised results alongside regularised ones, this approach, though transparent, may be insufficient. The unregularised results often contain artifacts that obscure physical interpretation, and a well chosen regularisation scheme can eliminate solutions that are patently unphysical.

A more nuanced approach would involve explicitly testing regularisation assumptions against physical models where possible, and developing domain specific regularisation schemes that incorporate known physical constraints rather than generic smoothness priors. This would shift the subjectivity from mathematical convenience to physics informed choices, making the trade offs more scientifically meaningful rather than purely computational.

II.E.2.ii High dimensional regimes.

Traditional methods fail catastrophically in $d \gtrsim 4$ phase spaces and struggle even before that for multiple reasons. Binned approaches require n^d histogram bins struggling to effectively sample an increasingly sparse phase space, are shown in Section II.A.5.i. Global smoothness assumptions become untenable for multi–scale features [62, 95, 96] straining regularisation methods that rely on them.

As the number of dimensions increases, the binning also increasingly distorts error propagation. Bayesian credible intervals can exhibit poor

frequentist coverage, as shown Fig. 4 of [97] and as argued in [98, 99], and correlated systematic uncertainties[6] introduce non-convex likelihoods [100, 101] that strain inference methods.

II.E.2.ii.a Spectrum dependent biases.

Sharply falling spectra[7] exacerbate regularisation artifacts [41]. Entropic priors overweight high$-z$ regions, distorting tails [103–106], finite sample sizes truncate the measurable phase space, creating cutoff–induced spikes [107, 108], and curvature penalties conflict with natural spectral shapes, requiring physics informed regularisation strategies [109–112].

Recent advances aim to mitigate these limitations in various ways. For example, adversarial regularisation involves training discriminators to enforce physical consistency rather than explicit smoothness [113]. Differentiable unfolding methods embed the detector response in neural networks enabling gradient based λ optimisation [114].

[6] E.g., jet energy scale.
[7] E.g., proton momentum in [102].

However, no universal solution exists, and the choice of regularisation must align with analysis specific priorities. As detector granularity increases, developing dimension agnostic regularisation schemes remains an open challenge requiring collaboration between statisticians and physicists to design methods that are statistically sound and suitable for the physics being analysed.

II.F Unbinned methods: statistical considerations.

The evolution from binned to unbinned unfolding methodologies represents a paradigm shift in high energy physics, driven by the need to preserve fine grained kinematic information while managing the statistical and computational complexities of high dimensional phase spaces. This section systematically analyses the theoretical foundations, practical challenges, and performance trade-offs that this transition entails.

II.F.1 Principles and implementations.

The unbinned approach to unfolding represents a fundamental shift from traditional histogram based methods. Rather than discretising data into bins, unbinned unfolding preserves the complete kinematic information in the data by operating directly on individual event tuples $\{(z_1, x_1), ..., (z_N, x_N)\}$ where each z_i represents a particle level event and x_i its corresponding detector

level measurement. This approach eliminates information loss from binning and naturally handles high dimensional phase spaces where binning becomes prohibitive.

Modern unbinned methods leverage machine learning techniques, particularly those designed for density (ratio) estimation. These approaches fall into two main categories: reweighting methods and direct generative modelling.

II.F.1.i Reweighting methods.

The central goal of one class of unbinned unfolding methods is to estimate a reweighting function $w(z)$ that transforms Generation to match the underlying Truth that was forward folded into the Data. Formally, this function satisfies

$$p(z) = w(z) \cdot q(z), \tag{II.26}$$

where $q(z)$ is the distribution of Generation.

The most established approach in this category is OmniFold [31], which employs an iterative procedure inspired by the Iterative Bayesian Unfolding

(IBU) algorithm. OmniFold alternates between two steps that estimate likelihood ratios using binary classifiers.

In each iteration k, OmniFold first trains a classifier to distinguish between Data and Simulation yielding the ratio

$$\nu^{(k)}(x) = \frac{p(x)}{q^{(k)}(x)}, \tag{II.27}$$

where $q^{(k)}(x)$ represents the distribution from Simulation weighted by the current iteration's weights.

This detector level ratio is then propagated back to particle level through the Monte Carlo pairing. For each Generation event z_i with corresponding Simulation event x_i, the weight update follows:

$$w^{(k+1)}(z_i) = w^{(k)}(z_i) \cdot \nu^{(k)}(x_i). \tag{II.28}$$

The process, in principle, continues until convergence, at which point the final weights $w(z)$ provide the desired transformation from Generation to Truth.

The connection to likelihood ratio estimation can be explicitly proven—for a binary cross entropy optimised classifier f, the likelihood ratio between the distributions it classifies is

$$LR = \frac{f}{1-f}. \tag{II.29}$$

Algorithm 2 details the OmniFold procedure in pseudocode.

Algorithm 2 OmniFold

Require: Data samples $\{x_j^{\text{Data}}\}$,
Gen.–Sim. pairs $\{(z_i, x_i)\}$,
Initial gen–level weights $w^{(0)}(z_i)$,
iterations K

Ensure: Unfolded weights $w^{(K)}(z_i)$

1: **for** $k \leftarrow 1$ **to** K **do**
2: **Detector-level step:**
3: Train classifier $C_{det}^{(k)}$ on $\{x_j^{\text{Data}}, +1\} \cup \{x_i, -1, w^{(k-1)}(z_i)\}$
4: Compute detector ratio

$$\nu^{(k)}(x) \;=\; \frac{C_{det}^{(k)}(x)}{1 - C_{det}^{(k)}(x)} \quad \forall x \tag{II.30}$$

Algorithm 2 (continued)

Pull-back weights:
for all pairs (z_i, x_i) **do**
    $\omega_i^{(k)} \leftarrow \nu^{(k)}(x_i)$
Particle-level step:
Train classifier $C_{part}^{(k)}$ on $\{z_i, +1, \omega_i^{(k)}\} \cup \{z_i, -1, w^{(k-1)}(z_i)\}$
Compute particle ratio

$$\rho^{(k)}(z) \;=\; \frac{C_{part}^{(k)}(z)}{1 - C_{part}^{(k)}(z)} \quad \forall z \tag{II.31}$$

Update weights:
for all generator events z_i **do**
    $w^{(k)}(z_i) \;\leftarrow\; w^{(k-1)}(z_i)\, \rho^{(k)}(z_i)$
return $\{w^{(K)}(z_i)\}$

II.F.1.ii Generative modelling.

An alternative approach employs generative models to directly learn the conditional distribution $p(z \mid x)$. Among these, Conditional Invertible Neural Networks (cINNs) [115] offer a particularly elegant framework.

cINNs learn a diffeomorphic mapping

$$g_\theta : Z \to X \tag{II.32}$$

between particle and detector levels, parameterised by neural network weights θ. The invertibility constraint ensures that for any detector level observation x, one can directly sample the corresponding particle level distribution through the inverse mapping. The key advantage lies in the tractable Jacobian computation,

$$p(z \mid x) = p(x \mid z)\,\frac{p(z)}{p(x)} = \left|\det \frac{\partial g_\theta}{\partial z}\right|^{-1} p(g_\theta^{-1}(x) \mid x)\ , \tag{II.33}$$

enabling direct sampling from the posterior $p(z \mid x)$ without iterative procedures [116].

Generative models, due to the much greater expressiveness that is necessary for them to model the functional representations needed, are succeptible to training challenges such as mode collapse[8]. Ensuring stable convergence of generative models can get increasingly difficult for high dimensional problems.

More recently, Schrödinger Bridge Unfolding [117] has emerged as a method that frames unfolding as an optimal transport problem. This approach

[8]Normalising flows however, due to the structural constraints that invertibility imposes, are resistant to mode collapse.

minimizes the Kullback–Leibler divergence between joint distributions while constraining the marginal to match observed data.

$$\inf_{p(z,x)} D_{\mathrm{KL}}\Big(p(z,x) \parallel p_{\mathrm{MC}}(z,x)\Big) \text{ subject to } p(x) = p_{\mathrm{Data}}(x). \tag{II.34}$$

This formulation provides theoretical guarantees on the uniqueness of the solution and offers improved stability compared to purely generative approaches.

II.F.2 Statistical considerations in unbinned regimes.

Neural networks implicitly regularise via inductive controls.[9] However, this introduces model dependent smoothing scales requiring careful validation against closure tests [119].

Reweighting based methods can propagate one part of the overall uncertainty through event weights.

$$\Sigma[O] = \sum_{i=1}^{N} w_i^2 O(z_i)^2 - \left(\sum_{i=1}^{N} w_i O(z_i)\right)^2, \tag{II.35}$$

[9] E.g., convolutional layers enforce translational symmetry in jet images [118].

where $\Sigma(O)$ is the covariance of observable $O(z)$ [120]. While avoiding binning induced correlations, unbinned inference requires reconceptualising what the unbinned equivalent of the covariance matrix would be in order to appropriately account for correlations in the unfolded data.

II.F.3 Limitations in complex phase spaces.

II.F.3.i Model misspecification.

Generative models assume that

$$q(z) > 0 \implies p(z) > 0, \tag{II.36}$$

meaning that wherever Generation assigns positive probability density, $q(z) > 0$, the Truth distribution also has positive probability density, $p(z) > 0$. This assumption ensures that the generative model does not learn to generate events in regions of phase space where the true physics has zero probability. It is fundamental to the validity of generative unfolding approaches such as conditional Invertible Neural Networks (cINNs) and Variational Autoen-

coders (VAEs) [121]. Generative models learn to map from detector level measurements to particle level distributions by generating samples from the learned posterior. If the Generation were to sample in kinematically forbidden regions or unphysical phase space, the unfolded results would include spurious events that contaminate the measurement. The assumption requires that the generative model be properly constrained to respect the physical boundaries of the true distribution.

Violations of this assumption can occur when generative models, trained on imperfect or limited data, learn to extrapolate beyond the true physical support and generate samples in unphysical regions. This is particularly problematic given the tendency of neural networks to produce overconfident predictions in regions with sparse training data.

Conversely though, enforcing this support containment requirement has practical implications for unfolding performance. Since the generative model can only learn to assign meaningful probability density to regions of phase space adequately represented in the training data, kinematic regimes that

are poorly sampled or entirely absent from the Monte Carlo generation are rendered effectively invisible to the unfolding procedure regardless of their importance in the truth distribution. This limitation becomes particularly concerning for new physics searches, where signatures may manifest in previously unexplored regions of phase space. Novel phenomena occurring outside the support of the generative model where cannot be properly unfolded and may remain undetected [122].

The assumption thus places stringent requirements on Monte Carlo coverage and highlights the importance of comprehensive phase space sampling in training data preparation for generative unfolding methods. Hybrid approaches combining discriminative and generative components show promise for anomaly detection [113, 123].

II.G Evaluation metrics for unfolding.

The evaluation of unfolding methods presents unique challenges due to the ill posed nature of the inverse problem. While the goal of unfolding is conceptually straightforward, to recover the true particle level distribution from detector level observations, quantifying the success of this recovery requires careful consideration when the downstream use of the data is unknown.

Task specific evaluation becomes more tractable when clear physics objectives are defined. For example, when unfolding is performed to measure a specific parameter such as α_S, evaluation metrics can be designed to directly assess how well the unfolding procedure enables accurate parameter extraction. This section focuses on metrics and approaches for evaluating unfolding performance, considering both traditional binned techniques and modern unbinned methods, for general purpose unfolding applications where the downstream use is not known. Metrics for assessing accuracy, precision, and uncertainty quantification are assessed and practical considerations for their application in high energy physics analyses are discussed.

II.G.1 Statistical metrics for evaluating point estimates.

II.G.1.i Residual based metrics.

The most intuitive approach to evaluating an unfolding method is to compare the unfolded distribution to the true distribution when it is known[10]. Simple residual based metrics quantify the difference between the estimated and true distributions.

Definition II.2. For binned methods, the *bin by bin residual* is defined as

$$\delta_i = \hat{t}_i - t_i \tag{II.37}$$

where $\hat{t}_i$ is the unfolded count in bin i and t_i is the true count.

[10] e.g., in simulation studies

Various summary statistics of these residuals can be computed, including

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{t}_i - t_i)^2 \tag{II.38}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{t}_i - t_i)^2} \tag{II.39}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{t}_i - t_i| \tag{II.40}$$

The MSE might be preferred because it is easier to do calculus with, the RMSE preferred because it is expressed in the units of the original data rather than squared units, but they are manifestly functionally equivalent. The choice between the MSE and MAE depends on the specific requirements of the analysis. MSE more heavily penalizes large errors due to the squaring operation, making it more sensitive to outliers and extreme residuals. MAE treats all errors equally regardless of magnitude and proves more robust to outliers, making it preferable when the unfolding procedure should not be dominated by a few problematic bins or events. The square function though is more amenable to calculus than the absolute value function. In high energy

physics applications, RMSE is the most commonly used metric, while MAE may be provided as valuable complementary information.

While these metrics are straightforward, they have limitations in the context of unfolding. In particular, they treat all bins equally, even though certain regions of phase space may be physically more significant than others. Additionally, these metrics do not account for the correlations between bins introduced by the unfolding process.

For unbinned methods, where the output is a set of weights or a continuous probability density, these metrics must be adapted. One approach is to bin the unbinned unfolded distributions and then apply the above metrics, though this introduces binning artifacts that the unbinned method was designed to avoid.

II.G.1.ii Distributional distance metrics.

Given the limitations of simple residual metrics, distributional distance measures provide a more comprehensive assessment of unfolding perfor-

mance. These metrics compare the entire unfolded distribution to the true distribution.

The *Kullback–Leibler (KL) divergence* [124] measures the information lost when using the unfolded distribution to approximate the true distribution.

$$D_{\mathrm{KL}}(p||q) = \int p(x) \log \frac{p(x)}{q(x)} \, \mathrm{d}x, \tag{II.41}$$

where p is the true distribution and q is the unfolded distribution. While theoretically sound, KL divergence can be numerically unstable when the support of the distributions differs.

The *Vincze–Le Cam (VLC) divergence* [125, 126] is symmetric alternative to KL divergence that is both bounded and highly convex [127].

$$\Delta(p, q) = \frac{1}{2} \int \frac{(p(\lambda) - q(\lambda))^2}{p(\lambda) + q(\lambda)} \, \mathrm{d}\lambda. \tag{II.42}$$

This metric is particularly useful for comparing unfolding methods as it provides a balanced assessment of differences across the entire distribution and has been used in comparative analyses of various unfolding approaches [31, 128].

Table II.2: Distributional distance metrics for unfolding evaluation

	Symmetric	Bounded	Support sensitivity	Compute Cost
D_{KL}	No	No	High	Low
Δ_{VLC}	Yes	Yes	Medium	Low
W_2	Yes	No	Low	High

The *Wasserstein metric* [129], also known as the *Earth mover's distance*," [130] provides a measure of the minimum "work" required to transform one distribution into another.

$$W_p(p,q) = \left(\inf_{\gamma \in \Gamma(p,q)} \int \int |x-y|^p d\gamma(x,y) \right)^{1/p}, \tag{II.43}$$

where $\Gamma(p,q)$ is the set of all joint distributions with marginals p and q. This metric is particularly useful for unfolding evaluations as it accounts for both the magnitude and location of discrepancies between distributions. Table II.2 summarizes the various distributional metrics and their relative strengths for unfolding evaluation.

II.G.2 Uncertainty quantification metrics

Beyond point estimates, properly evaluating unfolding methods requires assessing the accuracy of their uncertainty estimates. This is particularly important in particle physics, where uncertainties propagate to downstream analyses such as parameter fitting.

II.G.2.i Pull distributions

Pull distributions offer a rigorous way to evaluate the calibration of reported uncertainties. For a given unfolded bin or parameter θ, the pull is defined as

$$\text{Pull}\,\theta = \frac{\hat{\theta} - \theta_{\text{true}}}{\sigma_{\hat{\theta}}} \tag{II.44}$$

where $\hat{\theta}$ is the unfolded estimate, θ_{true} is the true value, and $\sigma_{\hat{\theta}}$ is the reported uncertainty. For a well–calibrated method, the pull distribution across many pseudo–experiments should follow a standard normal distribution, $\mathcal{N}(0, 1)$. This follows from the definition of random variables. Since the pull is defined as the ratio of the bias to the estimated uncertainty, for a properly calibrated

unfolding method, two conditions must be satisfied simultaneously. First, the method must be unbiased, meaning that across many pseudo experiments, the estimated values should equal the true values on average, resulting in zero mean for the numerator. Second, the reported uncertainties must accurately reflect the actual variability of the estimates, such that the denominator correctly represents the standard deviation of the numerator.

When both conditions are met, the pull becomes a standardized random variable. For sufficiently large sample sizes, many unfolding estimators approach normality due to the Central Limit Theorem, since they often involve weighted sums or averages of many events. When the underlying estimator is approximately normally distributed, if it satisfies the two conditions above that set its mean and standard deviation to 0 and 1 respectively, the population of sample pulls must be distributed as a standard normal.

A pull distribution with non-zero mean indicates systematic bias in the unfolding procedure, Deviations from unit standard deviation indicate either overestimation or underestimation of uncertainties.

In the context of binned unfolding, pull distributions can be computed for each bin, while for unbinned methods, they can be applied to derived quantities or parameters of interest. For Bayesian methods, pulls can be calculated using the mean and standard deviation of the posterior distribution [131].

II.G.2.ii Coverage properties

Related to pulls but more direct is the evaluation of coverage properties of confidence or credible intervals. For a nominal 68% confidence interval, approximately 68% of intervals computed across many pseudo-experiments should contain the true value. Systematic deviations from nominal coverage indicate issues with the uncertainty estimation. Coverage can be assessed through closure tests. These involve generating multiple datasets from a known truth, applying the unfolding procedure, and checking the fraction of times the true value falls within the reported confidence intervals. Coverage plots plot the actual coverage versus the nominal coverage across different confidence levels. The example in Figure II.1 shows expected coverage

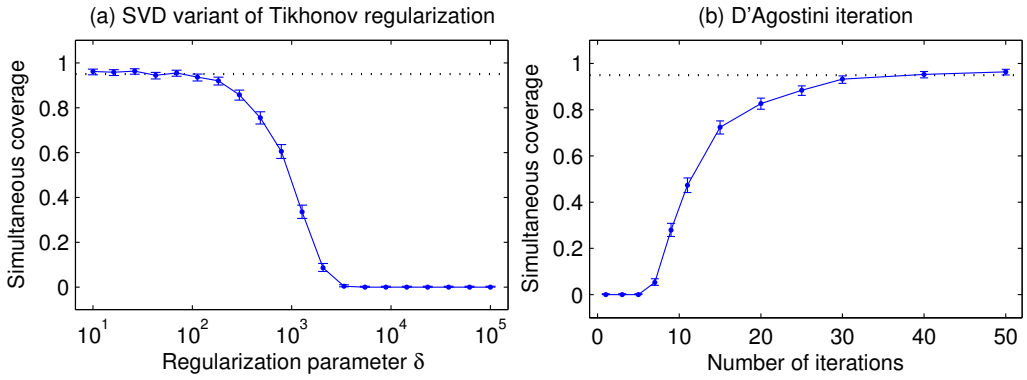

Figure II.1: Coverage of 95% confidence intervals with Tikhonov regularisation (left) and IBU (right). The error bars are given by the 95% Clopper–Pearson intervals and the nominal confidence level is shown by the dotted line. When the regularization is strong (fewer iterations in the case of IBU), both methods undercover substantially. IBU [53]

properties for two different unfolding methods, Tikhonov regularisation and IBU, as a function of the regularisation parameter.

II.G.2.iii Variance and bias decomposition

The total error of an unfolding method can be decomposed into bias and variance components,

$$\mathrm{MSE}(\hat{t}) = \mathrm{Bias}^2(\hat{t}) + \mathrm{Var}(\hat{t}), \tag{II.45}$$

where $\mathrm{Bias}(\hat{t}) = \mathbb{E}[\hat{t}] - t$ and $\mathrm{Var}(\hat{t}) = \mathbb{E}[(\hat{t} - \mathbb{E}[\hat{t}])^2]$.

This decomposition is particularly valuable for understanding the trade–offs inherent in regularized unfolding methods, where stronger regularization typically reduces variance at the expense of increased bias. Different applications might prioritize minimizing one component over the other, making this decomposition essential for method selection.

II.G.3 Evaluation of correlation structure

Traditional evaluation metrics often focus on marginal distributions, overlooking an important aspect of unfolding: the correlation structure between

different bins or events. Properly accounting for these correlations is crucial for downstream analyses.

II.G.3.i Covariance Matrix Assessment

For binned methods, the full covariance matrix of the unfolded distribution provides information about bin–to–bin correlations. A useful measure is the correlation matrix, defined as

$$\mathrm{Corr}_{ij} = \frac{\mathrm{Cov}_{ij}}{\sqrt{\mathrm{Cov}_{ii}\mathrm{Cov}_{jj}}} \tag{II.46}$$

Comparing the correlation structure of the unfolded distribution to that of the true distribution (when known, e.g. in simulation studies) can reveal systematic distortions introduced by the unfolding procedure.

II.G.3.ii Event–to–event correlation metrics

For unbinned methods event–to–event correlations in the unfolded weights can significantly impact downstream inference. These correlations can be quantified by studying the weight correlation as a function of distance. For

any pair of events, one can compute the correlation between their weights as a function of their distance in feature space. One can also estimate the reduction in statistical power due to correlated weights

$$N_{\text{eff}} = \frac{(\sum_i w_i)^2}{\sum_i w_i^2 + 2\sum_{i<j} w_i w_j \text{Corr}(w_i, w_j)} \tag{II.47}$$

In this equation, N_{eff} represents the effective number of independent observations when event weights are correlated, derived from the variance properties of weighted sums. The numerator, $(\sum w_i^2)$ represents the square of the total weighted sample size. The denominator consists of two components, which together account for the increased variance introduced by weight correlations. The first term, $\sum w_i^2$, captures the variance contribution from individual event weights, similar to the standard effective sample size formula for independent weighted observations. The second term, $2\sum_{i<j} w_i w_j \,\text{Corr}(w_i, w_j)$, accounts for the covariance contributions between all pairs of events, where the factor of 2 arises because each pair appears only once in the sum over $i < j$.

When weights are uncorrelated, $\mathrm{Corr}(w_i, w_j) = 0$, and the formula reduces to the familiar $N_{\text{eff}} = {}^{(\sum w_i)^2}/_{\sum w_i^2}$. However, upon unfolding methods, nearby events in phase space often receive similar weight corrections[11], leading to correlations that increase the denominator and reduce the effective sample size. This reduction quantifies the loss of statistical power compared to an ideal scenario with independent weights.

The formula emerges from considering the variance of weighted observables. For a weighted sum, $S = \sum w_i\, x_i$, the variance is

$$\mathrm{Var}(S) = \sum_i \sum_j w_i w_j \,\mathrm{Cov}(x_i, x_j), \tag{II.48}$$

which is the denominator of the Equation (II.47). The effective sample size is then defined as the ratio of the squared expectation to the variance, providing a measure of statistical efficiency that accounts for both weight magnitudes and their correlations.

Figure II.2 illustrates how event correlations typically decay with distance, with the correlation length scale increasing with detector resolution effects.

[11] as will be discusseed in greater detail in **??**

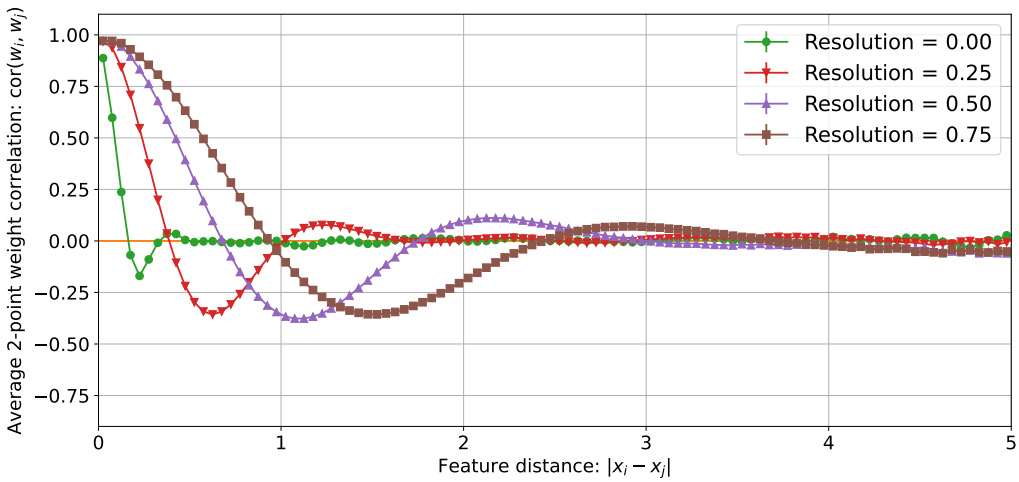

Figure II.2: Average weight correlation between two events as a function of the absolute distance between the events in the observable for Gaussian data unfolded using OmniFold [132].[12]

[12]Figure created by Owen Long

II.G.4 Method specific evaluation metrics

II.G.4.i Iterative methods

For iterative methods like Iterative Bayesian Unfolding (IBU) or OmniFold, convergence behaviour provides important diagnostic information. To study the convergence behaviour, we plot metric values (e.g., χ^2 or NLL) as a function of iteration number. We then compare unfolded distributions at different iterations to assess stability and analyse how the bias–variance trade off evolves with iteration number.

II.G.4.ii Bayesian Methods

For Bayesian unfolding methods such as Fully Bayesian Unfolding (FBU) [133] or Neural Posterior Unfolding (NPU) [134], additional posterior specific metrics are relevant. We can compare detector level data to detector level predictions generated from the posterior. We assess convergence using standard MCMC based diagnostics like Gelman–Rubin statistics or effective

sample size. The width of the posterior allows us to evaluate the posterior uncertainty in relation to the true frequentist variance.

II.G.5 Practical considerations.

In general, when comparing different unfolding methods, a structured evaluation framework ensures fair and comprehensive assessment. Such a framework should consider

- Computational efficiency: Measure training time, inference time, and memory requirements.
- Dimensionality scaling: Assess how performance metrics change as the dimensionality of the problem increases.
- Prior dependence: Evaluate robustness to different initial simulations.
- Regularisation parameter sensitivity: Compare how performance varies with changes in regularization strength.

In real experimental settings where the truth is unknown, evaluation presents additional challenges that require alternative pragmatic approaches. For example, when the true distribution is unavailable, data-splitting techniques can provide useful validation. The two most commonly used techniques are cross-validation, where we split the detector–level data, unfold one portion, then refold it, and compare predictions against the held–out portion; and bootstrapping where we generate multiple resampled datasets to assess the stability of the unfolding procedure.

Closure tests involve applying the full analysis chain (forward model followed by unfolding) to a known input distribution. While not a direct evaluation of performance on real data, closure tests provide confidence in the methodology. The simplest kinds of closure tests involve apply detector simulation to a known particle–level distribution, then unfolding the resulting detector–level distribution and compare with the original input. This procedure can then be modified by using a different particle-level input

than the one used to train the unfolding method, testing robustness to prior misspecification.

Evaluating how unfolding methods propagate systematic uncertainties is crucial for real-world applications. We can test the sensitivity of the method to systematic uncertainties by applying variations to the response matrix based on known systematic uncertainties and assessing the impact on unfolded distributions. For methods that support nuisance parameter profiling, evaluating how effectively nuisance parameters are profiled out is a gold standard test for the effectiveness of the method.

Rigorous evaluation of unfolding methods requires a multi–faceted approach that considers accuracy, uncertainty quantification, and computational performance. The metrics and frameworks presented in this section provide a comprehensive foundation for assessing both traditional and machine learning-based unfolding techniques. For binned methods, established metrics like χ^2 and coverage tests remain valuable, while for unbinned approaches, distributional metrics like Wasserstein distance and VLC divergence offer more

appropriate evaluation. Regardless of the method, uncertainty calibration through pull distributions and correlation structure assessment are important to validate any measurement. As unfolding methods continue to evolve, particularly with the advent of machine learning approaches, evaluation metrics must adapt accordingly. The framework presented here is designed to be extensible, accommodating new methods and application domains while maintaining rigour and comparability

References

[1] A. Rubbia. *Phenomenology of Particle Physics*. Cambridge, UK: Cambridge University Press, June 2022. ISBN: 978-1-316-51934-9 978-1-009-02342-9.

[2] G. Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett. B* 716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[3] T. C. Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1

(Sept. 2012). arXiv:1207.7235 [hep-ex], pp. 30–61. ISSN: 03702693. DOI: 10.1016/j.physletb.2012.08.021. URL: http://arxiv.org/abs/1207.7235 (visited on 07/13/2025).

[4] W. Commons. *File:Standard Model of Elementary Particles.svg — Wikimedia Commons, the free media repository*. [Online; accessed 12-July-2025]. 2025. URL: %5Curl%7Bhttps://commons.wikimedia.org/w/index.php?title=File:Standard%5C%5FModel%5C%5Fof%5C%5FElementary%5C%5FParticles.svg&oldid=1013094830%7D.

[5] S. L. Glashow. "Partial Symmetries of Weak Interactions". In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.

[6] S. Weinberg. "Conceptual Foundations of the Unified Theory of Weak and Electromagnetic Interactions". In: *Rev. Mod. Phys.* 52 (1980). Ed. by S. Lundqvist, pp. 515–523. DOI: 10.1103/RevModPhys.52.515.

[7] A. Salam. "Gauge Unification of Fundamental Forces". In: *Rev. Mod. Phys.* 52 (1980). Ed. by A. Ali, C. Isham, T. Kibble, and Riazuddin, pp. 525–538. DOI: 10.1103/RevModPhys.52.525.

[8] P. W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 Oct. 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[9] E. D'Hoker and E. Farhi. "Decoupling a fermion in the standard electro-weak theory". In: *Nuclear Physics, Section B* 248.1 (1984). ISSN: 05503213. DOI: 10.1016/0550-3213(84)90587-x.

[10] Particle Data Group, R. L. Workman, V. D. Burkert, V. Crede, E. Klempt, U. Thoma, L. Tiator, K. Agashe, G. Aielli, B. C. Allanach, et al. "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2022.8 (Aug. 2022), p. 083c01. ISSN: 2050-3911. DOI: 10.1093/ptep/ptac097. URL: https://doi.org/10.1093/ptep/ptac097 (visited on 07/13/2025).

[11] A. Chiefa, M. N. Costantini, J. Cruz-Martinez, E. R. Nocera, T. R. Rabemananjara, J. Rojo, T. Sharma, R. Stegeman, and M. Ubiali. "Parton distributions confront LHC Run II data: a quantitative ap-

praisal". en. In: *Journal of High Energy Physics* 2025.7 (July 2025), p. 67. ISSN: 1029-8479. DOI: 10.1007/jhep07(2025)067. URL: https://doi.org/10.1007/JHEP07(2025)067 (visited on 07/13/2025).

[12] "Precision electroweak measurements on the Z resonance". In: *Physics Reports* 427.5 (May 2006), pp. 257–454. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2005.12.006. URL: https://www.sciencedirect.com/science/article/pii/S0370157305005119 (visited on 07/13/2025).

[13] R. Contino, A. Falkowski, F. Goertz, C. Grojean, and F. Riva. "On the validity of the effective field theory approach to SM precision tests". en. In: *Journal of High Energy Physics* 2016.7 (July 2016), p. 144. ISSN: 1029-8479. DOI: 10.1007/jhep07(2016)144. URL: https://doi.org/10.1007/JHEP07(2016)144 (visited on 07/13/2025).

[14] A. Albert, S. Alves, M. André, M. Ardid, S. Ardid, J. .-. J. Aubert, J. Aublin, B. Baret, S. Basa, Y. Becherini, et al. "The ANTARES detector: Two decades of neutrino searches in the Mediterranean Sea". In: *Physics Reports* 1121-1124 (June 2025), pp. 1–46. ISSN: 0370-

1573. DOI: 10.1016/j.physrep.2025.04.001. URL: https://www.sciencedirect.com/science/article/pii/S0370157325001450 (visited on 07/13/2025).

[15] A. Buckley, J. Butterworth, J. Egan, C. Gutschow, S. Jeon, M. Habedank, T. Procter, P. Wang, Y. Yeh, and L. Yue. "Constraints On New Theories Using Rivet : CONTUR version 3 release note". en. In: *arXiv e-prints* (May 2025), arXiv:2505.09272. DOI: 10.48550/arXiv.2505.09272. URL: https://ui.adsabs.harvard.edu/abs/2025arXiv250509272B/abstract (visited on 07/13/2025).

[16] M. E. Peskin and D. V. Schroeder. *An Introduction to quantum field theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5 978-0-429-50355-9 978-0-429-49417-8. DOI: 10.1201/9780429503559.

[17] S. Navas, C. Amsler, T. Gutsche, W. Vogelsang, C. Hanhart, U. G. Meißner, J. J. Hernández-Rey, A. Pich, C. Lourenço, A. Ceccucci, et al. "Review of particle physics". In: *Physical Review D* 110.3 (Aug. 2024), p. 51. ISSN: 24700029. DOI: 10.1103/physrevd.110.030001. URL:

https://pdg.lbl.gov/2025/reviews/kinematics%5C%5Fand%5C%5Fcross%5C%5Fsections.html.

[18] W. Hollik. "Quantum Field Theory and the Standard ModelQuantum Field Theory and the Standard Model , Matthew D. Schwartz, Cambridge U. Press, 2014. $90.00 (850 pp.). ISBN 978-1-107-03473-0 ". In: *Physics Today* 67.12 (2014), pp. 57–58. ISSN: 0031-9228.

[19] *Quantum Field Theory and the Standard Model |CambridgeUniversityPress&Assessment*. URL: https://www.cambridge.org/us/universitypress/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-field-theory-and-standard-model.

[20] P. E. R. F.R.S. "LXXIX. The scattering of α and β particles by matter and the structure of the atom". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 21.125 (May 1911), pp. 669–688. ISSN: 1941-5982. DOI: 10.1080/14786440508637080. URL: https://www.tandfonline.com/doi/abs/10.1080/14786440508637080.

[21] W. R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach*. en. Berlin, Heidelberg: Springer, 1994. ISBN: 978-3-540-57280-0 978-3-642-57920-2. DOI: 10.1007/978-3-642-57920-2. URL: https://link.springer.com/10.1007/978-3-642-57920-2 (visited on 07/13/2025).

[22] S. J. Brodsky, F. Fleuret, C. Hadjidakis, and J. P. Lansberg. "Physics opportunities of a fixed-target experiment using LHC beams". In: *Physics Reports* 522.4 (Jan. 2013), pp. 239–255. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2012.10.001.

[23] P. Avery and A. Korytov. "Cross section, Flux, Luminosity, Scattering Rates". In: ().

[24] F. Muheim. "Nuclear and Particle Physics Particle Physics Particle Physics–Measurements and Theory Measurements and Theory Natural Units Relativistic Kinematics Particle Physics Measurements Lifetimes Resonances and Widths Scattering Cross section Collider and Fixed Target Experiments Conservation Laws". In: ().

[25] Belle Collaboration, M. Leitgab, R. Seidl, M. Grosse Perdekamp, A. Vossen, I. Adachi, H. Aihara, D. M. Asner, V. Aulchenko, T. Aushev, et al. "Precision Measurement of Charged Pion and Kaon Differential Cross Sections in ${e}^{+}{e}^{\ensuremath{-}}$ Annihilation at $\sqrt{s}\mathbf{=}10.52\text{ }\text{ }\mathrm{GeV}$". In: *Physical Review Letters* 111.6 (Aug. 2013). Publisher: American Physical Society, p. 062002. DOI: 10.1103/PhysRevLett.111.062002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.111.062002 (visited on 07/13/2025).

[26] A. J. Larkoski, I. Moult, and B. Nachman. "Jet substructure at the Large Hadron Collider: A review of recent advances in theory and machine learning". In: *Physics Reports* 841 (Jan. 2020), pp. 1–63. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2019.11.001.

[27] R. Kogler, B. Nachman, A. Schmidt, L. Asquith, E. Winkels, M. Campanelli, C. Delitzsch, P. Harris, A. Hinzmann, D. Kar, et al. "Jet substructure at the Large Hadron Collider". In: *Reviews of Mod-*

ern Physics 91.4 (Dec. 2019). Publisher: American Physical Society, p. 045003. DOI: 10.1103/RevModPhys.91.045003. URL: https://link.aps.org/doi/10.1103/RevModPhys.91.045003 (visited on 07/13/2025).

[28] M. U. Mozer. "Jet reconstruction and substructure measurements in ATLAS and CMS with first Run-2 data". en. In: *Proceedings of Fourth Annual Large Hadron Collider Physics – PoS(LHCP2016)*. Vol. 276. Conference Name: Fourth Annual Large Hadron Collider Physics. SISSA Medialab, Mar. 2017, p. 090. DOI: 10.22323/1.276.0090. URL: https://pos.sissa.it/276/090 (visited on 07/13/2025).

[29] "Inverse Problems and Interpretation of Measurements". In: *Applied Mathematical Sciences (Switzerland)* 160 (2005), pp. 1–5. ISSN: 2196968x. DOI: 10.1007/0-387-27132-5{_}1.

[30] W. C. Karl. "Regularization in Image Restoration and Reconstruction". In: *Handbook of Image and Video Processing, Second Edition* (Jan. 2005), pp. 183–202. DOI: 10.1016/b978-012119792-6/50075-9.

[31] A. Andreassen, P. T. Komiske, E. M. Metodiev, B. Nachman, and J. Thaler. "OmniFold: A Method to Simultaneously Unfold All Observables". In: *Physical Review Letters* 124.18 (May 2020). Publisher: American Physical Society, p. 182001. DOI: 10.1103/PhysRevLett.124.182001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.124.182001 (visited on 07/13/2025).

[32] I. Fredholm. "Sur une classe d'équations fonctionnelles". In: *Acta Mathematica* 27.none (Jan. 1903). Publisher: Institut Mittag-Leffler, pp. 365–390. ISSN: 0001-5962, 1871-2509. DOI: 10.1007/bf02421317. URL: https://projecteuclid.org/journals/acta-mathematica/volume-27/issue-none/Sur-une-classe-d%c3%a9quations-fonctionnelles/10.1007/BF02421317.full (visited on 07/13/2025).

[33] S. Weinberg. "Elementary Particle Theory of Composite Particles". In: *Physical Review* 130.2 (Apr. 1963). Publisher: American Physical Society, pp. 776–783. DOI: 10.1103/PhysRev.130.776. URL: https:

//link.aps.org/doi/10.1103/PhysRev.130.776 (visited on 07/13/2025).

[34] CMS Collaboration, S. Chatrchyan, V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, C. Fabjan, et al. "Measurement of the differential cross section for isolated prompt photon production in pp collisions at 7 TeV". In: *Physical Review D* 84.5 (Sept. 2011). Publisher: American Physical Society, p. 052011. DOI: 10.1103/PhysRevD.84.052011. URL: https://link.aps.org/doi/10.1103/PhysRevD.84.052011 (visited on 07/13/2025).

[35] B. Palumbo, P. Massa, and F. Benvenuto. *Convergence rates for Tikhonov regularization on compact sets: application to neural networks*. arXiv:2505.19936 [math]. May 2025. DOI: 10.48550/arXiv.2505.19936. URL: http://arxiv.org/abs/2505.19936 (visited on 07/13/2025).

[36] A. Neumaier. "Solving Ill-Conditioned and Singular Linear Systems: A Tutorial on Regularization". In: *SIAM Review* 40.3 (Jan. 1998). Publisher: Society for Industrial and Applied Mathematics, pp. 636–666. ISSN: 0036-1445. DOI: 10.1137/s0036144597321909. URL: https://epubs.siam.org/doi/10.1137/S0036144597321909 (visited on 07/13/2025).

[37] J. Chung and K. Palmer. "A Hybrid Lsmr Algorithm For Large-scale Tikhonov Regularization". In: (). DOI: 10.1137/140975024. URL: http://www.siam.org/journals/sisc/37-5/97502.html.

[38] J. L. Fernández-Martínez, J. L. G. Pallero, Z. Fernández-Muñiz, and L. M. Pedruelo-González. "The effect of noise and Tikhonov's regularization in inverse problems. Part I: The linear case". In: *Journal of Applied Geophysics* 108 (Sept. 2014), pp. 176–185. ISSN: 0926-9851. DOI: 10.1016/j.jappgeo.2014.05.006. URL: https://www.sciencedirect.com/science/article/pii/S0926985114001402 (visited on 07/13/2025).

[39] A. Carpio, O. Dorn, M. Moscoso, F. Natterer, G. C. Papanicolaou, M. L. Rapún, and A. Teta. *Inverse Problems and Imaging*. Ed. by L. L. Bonilla, J. .-. Morel, F. Takens, and B. Teissier. Vol. 1943. Lecture Notes in Mathematics. Berlin, Heidelberg: Springer, 2008. ISBN: 978-3-540-78545-3 978-3-540-78547-7. DOI: 10.1007/978-3-540-78547-7. URL: http://link.springer.com/10.1007/978-3-540-78547-7 (visited on 07/13/2025).

[40] M. Arratia, D. Britzger, O. Long, and B. Nachman. "Optimizing observables with machine learning for better unfolding". en. In: *Journal of Instrumentation* 17.07 (July 2022). Publisher: IOP Publishing, P07009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/17/07/p07009. URL: https://dx.doi.org/10.1088/1748-0221/17/07/P07009 (visited on 07/13/2025).

[41] A. Gaponenko. "A practical way to regularize unfolding of sharply varying spectra with low data statistics". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers,*

Detectors and Associated Equipment 960 (Apr. 2020), p. 163612. ISSN: 0168-9002. DOI: 10.1016/j.nima.2020.163612. URL: https://www.sciencedirect.com/science/article/pii/S0168900220301790 (visited on 07/13/2025).

[42] J. Chan and B. Nachman. "Unbinned profiled unfolding". In: *Physical Review D* 108.1 (July 2023). Publisher: American Physical Society, p. 016002. DOI: 10.1103/PhysRevD.108.016002. URL: https://link.aps.org/doi/10.1103/PhysRevD.108.016002 (visited on 07/13/2025).

[43] C. Bierlich, S. Chakraborty, N. Desai, L. Gellersen, I. Helenius, P. Ilten, L. Lönnblad, S. Mrenna, S. Prestel, C. T. Preuss, et al. "A comprehensive guide to the physics and usage of PYTHIA 8.3". In: *SciPost Physics Codebases* (Nov. 2022), p. 8. DOI: 10.21468/SciPostPhysCodeb.8. URL: https://scipost.org/10.21468/SciPostPhysCodeb.8 (visited on 07/13/2025).

[44] J. Allison, K. Amako, J. Apostolakis, P. Arce, M. Asai, T. Aso, E. Bagli, A. Bagulya, S. Banerjee, G. Barrand, et al. "Recent developments in Geant4". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 835 (Nov. 2016), pp. 186–225. ISSN: 0168-9002. DOI: 10.1016/j.nima.2016.06.125. URL: https://www.sciencedirect.com/science/article/pii/S0168900216306957 (visited on 07/13/2025).

[45] A. Bozson, G. Cowan, and F. Spanò. "Unfolding with Gaussian Processes". In: (Nov. 2018). _eprint: 1811.01242.

[46] S. Schmitt. "Data Unfolding Methods in High Energy Physics". In: *EPJ Web Conf.* 137 (2017). Ed. by Y. Foka, N. Brambilla, and V. Kovalenko. _eprint: 1611.01927, p. 11008. DOI: 10.1051/epjconf/201713711008.

[47] V. Blobel. "Unfolding Methods in Particle Physics". In: *Phystat 2011*. Geneva: Cern, 2011, pp. 240–251. DOI: 10.5170/cern-2011-006.240.

[48] R. G. Huang, A. Cudd, M. Kawaue, T. Kikawa, B. Nachman, V. Mikuni, and C. Wilkinson. "Machine Learning-Assisted Unfolding for Neutrino Cross-section Measurements with the OmniFold Technique". In: (June 2025). DOI: 10.1103/sp1f-n9k2. URL: http://arxiv.org/abs/2504.06857%20http://dx.doi.org/10.1103/sp1f-n9k2.

[49] H. Zhu, K. Desai, M. Kuusela, V. Mikuni, B. Nachman, and L. Wasserman. "Multidimensional Deconvolution with Profiling". In: *NeurIPS. ML4PS*. 150. 2024. DOI: https://doi.org/10.48550/arXiv.2409.10421. arXiv: 0902.0885 [hep-ph].

[50] T. Dorigo and P. de Castro. "Dealing with Nuisance Parameters using Machine Learning in High Energy Physics: a Review". In: (July 2020). URL: https://arxiv.org/abs/2007.09121v2.

[51] K. Cranmer, G. Lewis, L. Moneta, A. Shibata, and W. Verkerke. *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*. New York, 2012. DOI: 10.17181/cern-open-2012-

016. URL: https://cds.cern.ch/record/1456844 (visited on 07/13/2025).

[52] Y. Ke, I. Luise, Q. Buat, and G. Piacquadio. "Recent Developments on the Statistical Treatment of Flavour Tagging Uncertainties in ATLAS". In: *PoS* Lhcp2022 (2023), p. 322. DOI: 10.22323/1.422.0322.

[53] M. Kuusela and P. B. Stark. "Shape-constrained uncertainty quantification in unfolding steeply falling elementary particle spectra". In: *The Annals of Applied Statistics* 11.3 (Sept. 2017). Publisher: Institute of Mathematical Statistics, pp. 1671–1710. ISSN: 1932-6157, 1941-7330. DOI: 10.1214/17-aoas1053. URL: https://projecteuclid.org/journals/annals-of-applied-statistics/volume-11/issue-3/Shape-constrained-uncertainty-quantification-in-unfolding-steeply-falling-elementary-particle/10.1214/17-AOAS1053.full (visited on 07/13/2025).

[54] J. Conrad, O. Botner, A. Hallgren, and C. Pérez de los Heros. "Including systematic uncertainties in confidence interval construction for

Poisson statistics". In: *Physical Review D* 67.1 (Jan. 2003). Publisher: American Physical Society, p. 012002. DOI: 10.1103/PhysRevD.67.012002. URL: https://link.aps.org/doi/10.1103/PhysRevD.67.012002 (visited on 07/13/2025).

[55] P. Giovanni and C. Zheng. "Multi-Dimensional Gaussian Fluctuations on the Poisson Space". In: *Electronic Journal of Probability* 15.none (Jan. 2010). Publisher: Institute of Mathematical Statistics and Bernoulli Society, pp. 1487–1527. ISSN: 1083-6489, 1083-6489. DOI: 10.1214/EJP.v15-813. URL: https://projecteuclid.org/journals/electronic-journal-of-probability/volume-15/issue-none/Multi-Dimensional-Gaussian-Fluctuations-on-the-Poisson-Space/10.1214/EJP.v15-813.full (visited on 07/13/2025).

[56] S. G. From and A. W. Swift. "Some New Bounds And Approximations On Tail Probabilities Of The Poisson And Other Discrete Distributions". en. In: *Probability in the Engineering and Informa-*

tional Sciences 34.1 (Jan. 2020), pp. 53–71. ISSN: 0269-9648, 1469-8951. DOI: 10.1017/s0269964818000347. URL: https://www.cambridge.org/core/journals/probability-in-the-engineering-and-informational-sciences/article/abs/some-new-bounds-and-approximations-on-tail-probabilities-of-the-poisson-and-other-discrete-distributions/5D150AD3C15A0F8219340C7437513C3D?utm%5C%5Fsource=chatgpt.com (visited on 07/13/2025).

[57] N. Zardoshti. "Investigating the Role of Coherence Effects on Jet Quenching in Pb-Pb Collisions at \sqrts_NN =2.76 TeV using Jet Substructure". In: *Nucl. Phys. A* 967 (2017). Ed. by U. Heinz, O. Evdokimov, and P. Jacobs. _eprint: 1705.03383, pp. 560–563. DOI: 10.1016/j.nuclphysa.2017.05.055.

[58] G. Cowan. "Statistics for Searches at the LHC". In: *LHC Phenomenology*. Ed. by E. Gardi, N. Glover, and A. Robson. Cham: Springer International Publishing, 2015, pp. 321–355. ISBN: 978-3-319-05361-5 978-3-319-05362-2. DOI: 10.1007/978-3-319-05362-2_9. URL: http:

//link.springer.com/10.1007/978-3-319-05362-2%5C%5F9 (visited on 07/13/2025).

[59] W. H. Richardson. "Bayesian-Based Iterative Method of Image Restoration*". En. In: *Josa* 62.1 (Jan. 1972). Publisher: Optica Publishing Group, pp. 55–59. DOI: 10.1364/josa.62.000055. URL: https://opg.optica.org/josa/abstract.cfm?uri=josa-62-1-55 (visited on 07/13/2025).

[60] L. B. Lucy. "An iterative technique for the rectification of observed distributions". en. In: *The Astronomical Journal* 79 (June 1974), p. 745. ISSN: 0004-6256. DOI: 10.1086/111605. URL: https://ui.adsabs.harvard.edu/abs/1974AJ.....79..745L/abstract (visited on 07/13/2025).

[61] S. Schmitt. "Data Unfolding Methods in High Energy Physics". In: *EPJ Web of Conferences* 137 (Mar. 2017), p. 11008. ISSN: 2100-014x. DOI: 10.1051/epjconf/201713711008. URL: https://www.epj-conferences.org/articles/epjconf/abs/2017/06/epjconf%5C%

5Fconf2017%5C%5F11008/epjconf%5C%5Fconf2017%5C%5F11008.html.

[62] A. Höcker and V. Kartvelishvili. "SVD approach to data unfolding". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 372.3 (Apr. 1996), pp. 469–481. ISSN: 0168-9002. DOI: 10.1016/0168-9002(95)01478-0. URL: https://www.sciencedirect.com/science/article/pii/0168900295014780 (visited on 07/13/2025).

[63] S. Schmitt. "TUnfold: an algorithm for correcting migration effects in high energy physics". In: *Jinst* 7 (2012). _eprint: 1205.6201, T10003. DOI: 10.1088/1748-0221/7/10/t10003.

[64] G. Cowan. "Statistics". In: *Handbook of Particle Detection and Imaging*. Ed. by C. Grupen and I. Buvat. 2021, pp. 103–124. DOI: 10.1007/978-3-319-93785-4_5.

[65] G. Cowan. *Topics in statistical data analysis for high-energy physics*. en. 2010. DOI: 10.5170/cern-2010-002.197. URL: http://cds.cern.ch/record/1281954 (visited on 07/13/2025).

[66] G. Aad, B. Abbott, J. Abdallah, A. A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, et al. "Measurement of inclusive jet and dijet cross sections in proton-proton collisions at 7 TeV centre-of-mass energy with the ATLAS detector". en. In: *The European Physical Journal C* 71.2 (Feb. 2011), p. 1512. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-010-1512-2. URL: https://doi.org/10.1140/epjc/s10052-010-1512-2 (visited on 07/13/2025).

[67] *Implications of First LHC Data*. en. URL: https://indico.cern.ch/event/94815/contributions/1282695/ (visited on 07/13/2025).

[68] D. A. Belsley, E. Kuh, and R. E. Welsch. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. en. Google-

Books-ID: GECBEUJVNe0C. John Wiley & Sons, Feb. 2005. ISBN: 978-0-471-72514-5.

[69] M. H. Pesaran. *Time Series and Panel Data Econometrics*. en. Google-Books-ID: 7RokCwAAQBAJ. Oxford University Press, Oct. 2015. ISBN: 978-0-19-105847-9.

[70] H. Deng, Y. Yang, J. Li, C. Chen, W. Jiang, and S. Pu. *Fast Updating Truncated SVD for Representation Learning with Sparse Matrices*. arXiv:2401.09703 [math]. Jan. 2024. DOI: 10.48550/arXiv.2401.09703. URL: http://arxiv.org/abs/2401.09703 (visited on 07/13/2025).

[71] *TruncatedSVD – scikit-learn 1.7.0 documentation*. URL: https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.TruncatedSVD.html.

[72] W. Tang, X. Li, X. Qian, H. Wei, and C. Zhang. "Data Unfolding with Wiener-SVD Method". en. In: *Journal of Instrumentation* 12.10 (Oct. 2017), P10002. ISSN: 1748-0221. DOI: 10.1088/1748-0221/12/

10/p10002. URL: https://dx.doi.org/10.1088/1748-0221/12/10/P10002 (visited on 07/13/2025).

[73] S. Zaroubi. "Wiener reconstruction, SVD and 'optimal' functional bases: Application for redshift galaxy catalogs". In: *30th Rencontres de Moriond: Euroconferences: Clustering in the Universe*. _eprint: astro-ph/9505103. 1995, pp. 135–142.

[74] S. Maeda and T. Iguchi. "A new unfolding code combining maximum entropy and maximum likelihood for neutron spectrum measurement". In: *Journal of Nuclear Science and Technology* 50.4 (Apr. 2013). Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00223131.2013.773162, pp. 381–386. ISSN: 0022-3131. DOI: 10.1080/00223131.2013.773162. URL: https://doi.org/10.1080/00223131.2013.773162 (visited on 07/13/2025).

[75] G. Cowan. "Bayesian Statistical Methods In Particle Physics". In: *42nd Rencontres de Moriond on QCD and High Energy Hadronic Interactions*. Hanoi, Vietnam: Gioi Publ., 2007, pp. 7–10.

[76] G. Cowan. "Bayesian statistical methods for parton analyses". In: *14th International Workshop on Deep Inelastic Scattering*. Apr. 2006, pp. 157–160.

[77] F. James, R. Cousins, and G. Cowan. "Statistics". In: *Phys. Lett. B* 592 (2004), pp. 279–288.

[78] G. D'Agostini. *Improved iterative Bayesian unfolding*. arXiv:1010.0632 [physics]. Oct. 2010. DOI: 10.48550/arXiv.1010.0632. URL: http://arxiv.org/abs/1010.0632 (visited on 07/13/2025).

[79] D. A. Fish, A. M. Brinicombe, E. R. Pike, and J. G. Walker. "Blind deconvolution by means of the Richardson–Lucy algorithm". En. In: *Josa A* 12.1 (Jan. 1995). Publisher: Optica Publishing Group, pp. 58–65. ISSN: 1520-8532. DOI: 10.1364/josaa.12.000058. URL: https://opg.optica.org/josaa/abstract.cfm?uri=josaa-12-1-58 (visited on 07/13/2025).

[80] L. A. Shepp and Y. Vardi. "Maximum Likelihood Reconstruction for Emission Tomography". In: *IEEE Transactions on Medical Imaging* 1.2

(Oct. 1982), pp. 113–122. ISSN: 1558-254x. DOI: 10.1109/tmi.1982.4307558. URL: https://ieeexplore.ieee.org/document/4307558 (visited on 07/13/2025).

[81] G. Cowan. "A survey of unfolding methods for particle physics". In: *Conf. Proc. C* 0203181 (2002). Ed. by M. R. Whalley and L. Lyons, pp. 248–257.

[82] G. Cowan. *Statistical data analysis*. 1998. ISBN: 978-0-19-850156-5.

[83] ATLAS Collaboration, G. Aad, B. Abbott, K. Abeling, N. Abicht, S. Abidi, A. Aboulhorma, H. Abramowicz, H. Abreu, Y. Abulaiti, et al. "Measurement of jet substructure in boosted $t\overline{t}$ events with the ATLAS detector using $140\text{ }\text{ }{\mathrm{fb}}^{\ensuremath{-}1}$ of 13 TeV pp collisions". In: *Physical Review D* 109.11 (June 2024). Publisher: American Physical Society, p. 112016. DOI: 10.1103/PhysRevD.109.112016. URL: https://link.aps.org/doi/10.1103/PhysRevD.109.112016 (visited on 07/13/2025).

[84] G. Cowan. "Topics in statistical data analysis for HEP". In: *65th Scottish Universities Summer School in Physics: LHC Physics*. Aug. 2009, pp. 341–369. DOI: 10.1201/b11865-15.

[85] G. Cowan. "Highlights from PHYSTAT 2011". In: *Phystat 2011*. Geneva: Cern, 2011, pp. 215–224. DOI: 10.5170/cern-2011-006.215.

[86] T. Adye. "Unfolding algorithms and tests using RooUnfold". In: *Phystat 2011*. _eprint: 1105.1160. Geneva: Cern, 2011, pp. 313–318. DOI: 10.5170/cern-2011-006.313.

[87] D. Britzger. "The Linear Template Fit". en. In: *The European Physical Journal C* 82.8 (Aug. 2022), p. 731. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-022-10581-w. URL: https://doi.org/10.1140/epjc/s10052-022-10581-w (visited on 07/14/2025).

[88] M. Baak, S. Gadatsch, R. Harrington, and W. Verkerke. "Interpolation between multi-dimensional histograms using a new non-linear moment morphing method". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors*

and Associated Equipment 771 (Jan. 2015), pp. 39–48. ISSN: 0168-9002. DOI: 10.1016/j.nima.2014.10.033. URL: https://www.sciencedirect.com/science/article/pii/S0168900214011814 (visited on 07/14/2025).

[89] Y. Fan, Y. Li, N. Xue, and D. Ding. "Analysis of Regularized Poisson GLM Spike-Train Modeling". en. In: *Journal of Physics: Conference Series* 2173.1 (Jan. 2022). Publisher: IOP Publishing, p. 012019. ISSN: 1742-6596. DOI: 10.1088/1742-6596/2173/1/012019. URL: https://dx.doi.org/10.1088/1742-6596/2173/1/012019 (visited on 07/14/2025).

[90] J. Jia, F. Xie, and L. Xu. "Sparse Poisson regression with penalized weighted score function". In: *Electronic Journal of Statistics* 13.2 (Jan. 2019). Publisher: Institute of Mathematical Statistics and Bernoulli Society, pp. 2898–2920. ISSN: 1935-7524, 1935-7524. DOI: 10.1214/19-ejs1580. URL: https://projecteuclid.org/journals/electronic-journal-of-statistics/volume-13/issue-2/

Sparse-Poisson-regression-with-penalized-weighted-score-function/10.1214/19-EJS1580.full (visited on 07/14/2025).

[91] G. Bohm and G. Zech. "Introduction to Statistics and Data Analysis for Physicists". In: *Introduction to Statistics and Data Analysis for Physicists* (Aug. 2025). DOI: 10.1142/14343.

[92] H. N. Multhei and B. Schorr. "On an Iterative Method for the Unfolding of Spectra". In: *Nucl. Instrum. Meth. A* 257 (1987), p. 371. DOI: 10.1016/0168-9002(87)90759-5.

[93] G. Zech. "Regularization and error assignment to unfolded distributions". In: *Phystat 2011*. Geneva: Cern, 2011, pp. 252–259. DOI: 10.5170/cern-2011-006.252.

[94] M. J. Kuusela. "Uncertainty quantification in unfolding elementary particle spectra at the Large Hadron Collider". PhD Thesis. Ecole Polytechnique, Lausanne, 2016. DOI: 10.5075/epfl-thesis-7118.

[95] J. L. Fernández-Martínez and Z. Fernández-Muñiz. "The curse of dimensionality in inverse problems". In: *Journal of Computational and Applied Mathematics* 369 (May 2020), p. 112571. ISSN: 0377-0427. DOI: 10.1016/j.cam.2019.112571. URL: https://www.sciencedirect.com/science/article/pii/S037704271930576X (visited on 07/14/2025).

[96] Y. Xia and N. Zabaras. "Bayesian multiscale deep generative model for the solution of high-dimensional inverse problems". In: *Journal of Computational Physics* 455 (Apr. 2022), p. 111008. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2022.111008. URL: https://www.sciencedirect.com/science/article/pii/S0021999122000705 (visited on 07/14/2025).

[97] T. Zhang. "A Modification for Bayesian Credible Intervals". In: *Communications in Statistics - Theory and Methods* 35.9 (Sept. 2006), pp. 1703–1711. ISSN: 03610926. DOI: 10.1080/03610920600683838. URL: https://www.tandfonline.com/doi/abs/10.1080/03610920600683838.

[98] L. E. Eberly and G. Casella. "Estimating Bayesian credible intervals". In: *Journal of Statistical Planning and Inference*. Special issue II: Model Selection, Model Diagnostics, Empirical B ayes and Hierarchical Bayes 112.1 (Mar. 2003), pp. 115–132. ISSN: 0378-3758. DOI: 10.1016/s0378-3758(02)00327-0. URL: https://www.sciencedirect.com/science/article/pii/S0378375802003270 (visited on 07/14/2025).

[99] B. Szabó, A. W. v. d. Vaart, and J. H. v. Zanten. "Frequentist coverage of adaptive nonparametric Bayesian credible sets". In: *The Annals of Statistics* 43.4 (Aug. 2015). arXiv:1310.4489 [math]. ISSN: 0090-5364. DOI: 10.1214/14-aos1270. URL: http://arxiv.org/abs/1310.4489 (visited on 07/14/2025).

[100] N. Berger. "Simplified likelihoods using linearized systematic uncertainties". en. In: *Journal of High Energy Physics* 2023.4 (Apr. 2023), p. 84. ISSN: 1029-8479. DOI: 10.1007/jhep04(2023)084. URL: https://doi.org/10.1007/JHEP04(2023)084 (visited on 07/14/2025).

[101] N. Berger. *Lecture on Statistical analysis methods by ATLAS*. Sept. 2017. URL: https://cds.cern.ch/record/2285058.

[102] G. Arnison, A. Astbury, B. Aubert, C. Bacci, R. Bernabei, A. Bézaguet, R. Böck, T. J. V. Bowcock, M. Calvetti, T. Carroll, et al. "Transverse momentum spectra for charged particles at the CERN proton-antiproton collider". In: *Physics Letters B* 118.1 (Dec. 1982), pp. 167–172. ISSN: 0370-2693. DOI: 10.1016/0370-2693(82)90623-2. URL: https://www.sciencedirect.com/science/article/pii/0370269382906232 (visited on 07/14/2025).

[103] A. Caticha and R. Preuss. "Entropic Priors". In: *AIP Conference Proceedings* 707.1 (Apr. 2004), pp. 371–380. ISSN: 0094-243x. DOI: 10.1063/1.1751380. URL: https://doi.org/10.1063/1.1751380 (visited on 07/14/2025).

[104] C. C. Rodriguez. "Entropic priors for discrete probabilistic networks and for mixtures of Gaussians models". In: *AIP Conference Proceedings* 617.1 (May 2002), pp. 410–432. ISSN: 0094-243x. DOI: 10.1063/1.

1477063. URL: https://doi.org/10.1063/1.1477063 (visited on 07/14/2025).

[105] W. Handley and M. Millea. "Maximum-Entropy Priors with Derived Parameters in a Specified Distribution". In: *Entropy 2019, Vol. 21, Page 272* 21.3 (Mar. 2019), p. 272. ISSN: 1099-4300. DOI: 10.3390/e21030272. URL: https://www.mdpi.com/1099-4300/21/3/272/htm%20https://www.mdpi.com/1099-4300/21/3/272.

[106] B. J. Brewer and M. J. Francis. "Entropic Priors and Bayesian Model Selection". In: *AIP Conference Proceedings* 1193.1 (Dec. 2009), pp. 179–186. ISSN: 0094-243x. DOI: 10.1063/1.3275612. URL: https://doi.org/10.1063/1.3275612 (visited on 07/14/2025).

[107] R. Finotello, V. Lahoche, and D. O. Samary. *Functional Renormalization for Signal Detection: Dimensional Analysis and Dimensional Phase Transition for Nearly Continuous Spectra Effective Field Theory*. arXiv:2507.01064 [physics] version: 1. June 2025. DOI: 10.48550/

arXiv.2507.01064. URL: http://arxiv.org/abs/2507.01064 (visited on 07/14/2025).

[108] E. Marchand and W. E. Strawderman. *On Bayesian credible sets in restricted parameter space problems and lower bounds for frequentist coverage*. arXiv:1208.0028 [math]. Dec. 2012. DOI: 10.48550/arXiv.1208.0028. URL: http://arxiv.org/abs/1208.0028 (visited on 07/14/2025).

[109] Y. Lee and F. C. Park. *On Explicit Curvature Regularization in Deep Generative Models*. arXiv:2309.10237 [cs]. Sept. 2023. DOI: 10.48550/arXiv.2309.10237. URL: http://arxiv.org/abs/2309.10237 (visited on 07/14/2025).

[110] S.-M. Moosavi-Dezfooli, A. Fawzi, J. Uesato, and P. Frossard. *Robustness via curvature regularization, and vice versa*. arXiv:1811.09716 [cs]. Nov. 2018. DOI: 10.48550/arXiv.1811.09716. URL: http://arxiv.org/abs/1811.09716 (visited on 07/14/2025).

[111] G. Zech. "Analysis of distorted measurements – parameter estimation and unfolding". In: (July 2016). _eprint: 1607.06910.

[112] P. Baroň and J. Kvita. "Extending the Fully Bayesian Unfolding with Regularization Using a Combined Sampling Method". In: *Symmetry 2020, Vol. 12, Page 2100* 12.12 (Dec. 2020), p. 2100. ISSN: 2073-8994. DOI: 10.3390/sym12122100. URL: https://www.mdpi.com/2073-8994/12/12/2100/htm%20https://www.mdpi.com/2073-8994/12/12/2100.

[113] D. Terjék, Terjék, and Dávid. "Adversarial Lipschitz Regularization". In: *arXiv* (Jan. 2019), arXiv:1907.05681. DOI: 10.48550/arxiv.1907.05681. URL: https://ui.adsabs.harvard.edu/abs/2019arXiv190705681T/abstract.

[114] E. de la Rosa, D. Robben, D. M. Sima, J. S. Kirschke, and B. Menze. "Differentiable Deconvolution for Improved Stroke Perfusion Analysis". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*

12267 Lncs (2020), pp. 593–602. ISSN: 16113349. DOI: 10.1007/978-3-030-59728-3{_}58/figures/3. URL: https://link.springer.com/chapter/10.1007/978-3-030-59728-3%5C%5F58.

[115] G. Anantha Padmanabha and N. Zabaras. "Solving inverse problems using conditional invertible neural networks". In: *Journal of Computational Physics* 433 (May 2021), p. 110194. ISSN: 0021-9991. DOI: 10.1016/j.jcp.2021.110194.

[116] M. Bellagente, A. Butter, G. Kasieczka, T. Plehn, A. Rousselot, R. Winterhalder, L. Ardizzone, and U. Köthe. "Invertible Networks or Partons to Detector and Back Again". In: *SciPost Phys.* 9.5 (Nov. 2020), p. 074. ISSN: 25424653. DOI: 10.21468/scipostphys.9.5.074.

[117] A. Butter, S. Diefenbacher, N. Huetsch, V. Mikuni, B. Nachman, S. P. Schweitzer, and T. Plehn. "Generative unfolding with distribution mapping". In: *SciPost Phys.* 18.6 (June 2025), p. 200. ISSN: 25424653. DOI: 10.21468/scipostphys.18.6.200.

[118] H. Kheddar, Y. Himeur, A. Amira, and R. Soualah. "Image and Point-cloud Classification for Jet Analysis in High-Energy Physics: A survey". In: *Frontiers of Physics* 20.3 (Feb. 2025). DOI: 10.15302/frontphys.2025.035301. URL: http://arxiv.org/abs/2403.11934%20http://dx.doi.org/10.15302/frontphys.2025.035301.

[119] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur.Phys.J.C* 71.2 (Feb. 2011), p. 1554. ISSN: 14346052. DOI: 10.1140/epjc/s10052-011-1554-0.

[120] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, R. Tanaka, S. Alekhin, J. Alwall, E. A. Bagnaschi, A. Banfi, J. Blumlein, et al. "Handbook of LHC Higgs Cross Sections: 2. Differential Distributions". In: (Jan. 2012). DOI: 10.5170/cern-2012-002. URL: http://arxiv.org/abs/1201.3084%20http://dx.doi.org/10.5170/CERN-2012-002.

[121] A. Shmakov, K. Greif, M. J. Fenton, A. Ghosh, P. Baldi, and D. Whiteson. "Full Event Particle-Level Unfolding with Variable-Length Latent Variational Diffusion". In: *SciPost Physics* 18.4 (Jan. 2025). DOI: 10.21468/SciPostPhys.18.4.117. URL: http://arxiv.org/abs/2404.14332%20http://dx.doi.org/10.21468/SciPostPhys.18.4.117.

[122] C. Amsler, M. Doser, M. Antonelli, D. M. Asner, K. S. Babu, H. Baer, H. R. Band, R. M. Barnett, E. Bergren, J. Beringer, et al. "Monte Carlo techniques". In: *Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics* 667.1-5 (Sept. 2008), pp. 1–6. ISSN: 03702693. DOI: 10.1016/j.physletb.2008.07.018.

[123] F. Di Mattia, P. Galeone, M. De Simoni, and E. Ghelfi. "A Survey on GANs for Anomaly Detection". In: (June 2019). URL: https://arxiv.org/abs/1906.11632v2.

[124] S. Kullback and R. A. Leibler. "On Information and Sufficiency". In: *The Annals of Mathematical Statistics* 22.1 (Mar. 1951). Pub-

lisher: Institute of Mathematical Statistics, pp. 79–86. ISSN: 0003-4851, 2168-8990. DOI: 10.1214/aoms/1177729694. URL: https://projecteuclid.org/journals/annals-of-mathematical-statistics/volume-22/issue-1/On-Information-and-Sufficiency/10.1214/aoms/1177729694.full (visited on 07/14/2025).

[125] I. Vincze. "On the Concept and Measure of Information Contained in an Observation*". In: *Contributions to Probability*. Ed. by J. Gani and V. K. Rohatgi. Academic Press, Jan. 1981, pp. 207–214. ISBN: 978-0-12-274460-0. DOI: 10.1016/b978-0-12-274460-0.50023-0. URL: https://www.sciencedirect.com/science/article/pii/B9780122744600500230 (visited on 07/14/2025).

[126] L. L. Cam. "Asymptotic Methods in Statistical Decision Theory". In: Springer Series in Statistics (1986). DOI: 10.1007/978-1-4612-4946-7. URL: http://link.springer.com/10.1007/978-1-4612-4946-7.

[127] J. Melbourne. "Strongly Convex Divergences". In: *Entropy* 22.11 (Nov. 2020). arXiv:2009.10838 [cs], p. 1327. ISSN: 1099-4300. DOI: 10.3390/e22111327. URL: http://arxiv.org/abs/2009.10838 (visited on 07/14/2025).

[128] P. Komiske, W. P. McCormack, and B. Nachman. "Preserving new physics while simultaneously unfolding all observables". In: *Phys. Rev. D* 104.7 (2021). _eprint: 2105.09923, p. 076027. DOI: 10.1103/PhysRevD.104.076027.

[129] *L. N. Vaserstein, "Markov Processes over Denumerable Products of Spaces, Describing Large Systems of Automata", Probl. Peredachi Inf., 1969, Volume 5, Issue 3,Pages <nobr>64–72</nobr>*. URL: https://m.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=ppi&paperid=1811&option%5C%5Flang=eng.

[130] Y. Rubner, C. Tomasi, and L. Guibas. "A metric for distributions with applications to image databases". In: *Sixth International Conference on Computer Vision*. Jan. 1998, pp. 59–66. DOI: 10.1109/iccv.1998.

710701. URL: https://ieeexplore.ieee.org/document/710701 (visited on 07/14/2025).

[131] B. T. Knapik, A. W. van der Vaart, and J. H. van Zanten. "Bayesian inverse problems with Gaussian priors". In: *Annals of Statistics* 39.5 (Feb. 2012), pp. 2626–2657. DOI: 10.1214/11-aos920. URL: http://arxiv.org/abs/1103.2692%20http://dx.doi.org/10.1214/11-AOS920.

[132] K. Desai, O. Long, and B. Nachman. "Unbinned Inference with Correlated Events". In: *E-print* (Apr. 2025). arXiv: 2504.14072 [physics.data-an].

[133] G. Choudalakis. *Fully Bayesian Unfolding*. arXiv:1201.4612 [physics]. May 2012. DOI: 10.48550/arXiv.1201.4612. URL: http://arxiv.org/abs/1201.4612 (visited on 07/14/2025).

[134] F. T. Acosta, K. Desai, V. Mikuni, B. Nachman, and J. Pan. "Multi-dimensional Deconvolution with Profiling". In: *NeurIPS. ML4PS*. 177. 2024.