

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

Copyright 2025
by
Krish Desai

*To Ben Nachman,
my advisor, mentor, and friend.*

Contents

Contents	ii
List of Figures	xxix
List of Tables	xl
I Introduction and Physics Background	1
I.A Units and Conventions	2
I.B The Standard Model: Theoretical framework	6
I.B.1 Fundamental Particles and Forces	8

I.B.1.i	Fermions: The Building Blocks of Matter	8
I.B.1.i.a	Quarks	8
I.B.1.i.b	Leptons	9
I.B.1.ii	Bosons: Force Carriers	10
I.B.2	Theoretical Framework and Symmetries	10
I.B.3	The Higgs Mechanism and Mass Generation	11
I.B.4	Limitations and Beyond the Standard Model (BSM) Physics	12
I.C	Fundamental role of cross section measurements in particle physics	15
I.D	Cross Section Measurements: From Theory to Experiment	18
I.D.1	Theory	18
I.D.2	Experimental Measurement	21
I.D.3	Applications in Particle Physics	22
I.D.3.i	Theory validation	23
I.D.3.ii	Monte Carlo Tuning	24

I.D.3.iii	Consistency Checks	25
I.D.3.iv	Luminosity determination	26
I.D.3.v	Background estimation	27
I.D.3.vi	Parton Luminosity	27
I.D.3.vii	Detector Performance Validation	28
I.D.3.viii	Systematic uncertainty evaluation	29
I.D.3.ix	Future planning	29
I.E	Detector response in precision measurements	31
I.F	Challenges at modern experiments	37
I.G	Thesis Scope and Physics Impact	40
II	Theoretical Foundations	39
II.A	Statistical Formulation of the Unfolding Problem	40
II.A.1	The Detector Response and Forward Problem	40
II.A.2	The Inverse Problem: Unfolding	41
II.A.3	Likelihood-Based Formulation	42
II.A.4	Regularization Techniques	42

II.A.5	Challenges in High Dimensional Phase Spaces . . .	43
II.A.5.i	Binned methods	43
II.A.5.ii	Unbinned methods	48
II.B	Forward and Inverse Problems in HEP	51
II.B.1	Mathematical Formulation	51
II.B.2	Challenges in Inverse Problems	52
II.B.3	HEP Specific Considerations	53
II.C	Historical development: From matrix inversion to modern approaches	56
II.D	Traditional unfolding methods in experimental analyses. . .	59
II.D.1	Bin by bin correction.	59
II.D.2	Matrix Inversion	60
II.D.3	Iterative Bayesian unfolding	62
II.D.4	Tikhonov Regularization	63
II.D.5	Template Fitting	65
II.D.6	Regularized Poisson Likelihood	66

II.D.7	Summary	66
II.D.8	Regularization: Need, Approaches, and Limitations	67
II.D.8.i	The Necessity of Regularization	68
II.D.8.ii	Limitations and practical challenges . . .	68
II.D.8.ii.a	Subjectivity-objectivity trade-off	68
II.D.8.ii.b	High dimensional regimes	69
II.D.8.ii.c	Spectrum dependent biases	70
II.E	Unbinned Methods: Statistical Considerations	72
II.E.1	Principles and Implementations	72
II.E.1.i	Reweighting Methods	73
II.E.1.ii	Generative Modeling	75
II.E.2	Statistical Considerations in Unbinned Regimes . .	77
II.E.3	Limitations in Complex Phase Spaces	78
II.E.3.i	Model Misspecification	78
II.F	Evaluation metrics for unfolding.	81
II.F.1	Statistical metrics for evaluating point estimates. .	82

II.F.1.i	Residual based metrics.	82
II.F.1.ii	Distributional distance metrics.	85
II.F.2	Uncertainty quantification metrics	87
II.F.2.i	Pull distributions	87
II.F.2.ii	Coverage properties	89
II.F.2.iii	Variance and bias decomposition	90
II.F.3	Evaluation of correlation structure	92
II.F.3.i	Covariance Matrix Assessment	92
II.F.3.ii	Event-to-event correlation metrics	93
II.F.4	Method specific evaluation metrics	96
II.F.4.i	Iterative methods	96
II.F.4.ii	Bayesian Methods	96
II.F.5	Practical considerations.	97
III	Machine Learning for Unfolding	101
III.A	The Emergence of Machine Learning in Particle Physics . .	102
III.A.1	A Paradigm Shift in Data Analysis	102

III.A.2	Evolution of ML Applications in HEP	104
III.A.2.i	Classification Tasks	104
III.A.2.ii	Regression and Anomaly Detection	106
III.A.2.iii	Generative Models	107
III.A.3	Machine Learning Approaches to Unfolding	108
III.A.4	Early successes and current challenges	110
III.B	Introduction to Neural Networks	113
III.B.1	Essential Concepts	113
III.B.1.i	Neural Network Fundamentals	113
III.B.1.ii	Forward propagation	118
III.B.1.iii	Training objectives and loss functions . .	119
III.B.1.iv	Backpropagation and parameter learning .	121
III.B.1.v	Batching and Training Dynamics	122
III.B.2	Neural Networks as Universal Approximators . . .	122
III.B.2.i	Relevance to physics applications	124
III.B.3	Activation Functions, Optimization, and Regularization	126

III.B.3.i	Activation Functions	126
III.B.3.ii	Optimization Techniques	130
III.B.3.ii.a	Gradient Descent	130
III.B.3.ii.b	Stochastic Gradient Descent (SGD)	130
III.B.3.ii.c	Momentum based methods	131
III.B.3.ii.d	Adaptive Methods	131
III.B.3.ii.d.1	AdaGrad	132
III.B.3.ii.d.2	RMSPProp	132
III.B.3.ii.e	Adam	132
III.B.3.ii.f	Learning Rate Scheduling	133
III.B.3.iii	Regularisation techniques.	134
III.B.3.iii.a	L^p Regularisation	134
III.B.3.iii.b	Dropout	134
III.B.3.iii.c	Batch Normalization	135
III.C	Supervised Learning Approaches for Unfolding	136
III.C.1	Mathematical Framework	136

III.C.2	Statistical Interpretation	138
III.C.3	Types of Supervised Learning	139
III.C.4	Learning and Generalization	140
III.C.5	Evaluation	140
III.C.6	Unfolding as a Supervised Learning Problem	141
III.C.6.i	Classification-Based Approaches	142
III.C.6.ii	Regression based calibration	143
III.C.6.ii.a	Deterministic regression	144
III.C.6.ii.b	Probabilistic regression.	144
III.C.7	Regularisation strategies.	145
III.C.8	Advantages and challenges	146
III.D	Deep Learning Architectures	148
III.D.1	Convolutional Neural Networks	148
III.D.2	Recurrent Neural Networks	150
III.D.3	Graph neural networks	151
III.D.4	Transformer based architectures	152

III.D.5	Physics informed neural networks	154
III.D.5.i	Energy Flow Networks and Particle Flow Networks	155
III.D.5.i.a	Relationship to other architectures	158
III.D.5.i.a.1	Deep sets	158
III.D.5.i.a.2	GNNs and transformers	158
III.E	Modern Machine Learning Frameworks	161
III.E.1	Generative Models	161
III.E.1.i	Variational Autoencoders (VAEs)	163
III.E.1.ii	Normalising flows	167
III.E.1.ii.a	Mathematical formulation.	168
III.E.1.ii.b	Variations.	169
III.E.1.ii.b.1	Conditional normalising flows	169
III.E.1.ii.b.2	Other innovations.	170
III.E.1.ii.c	Applications.	171
III.E.1.ii.c.1	Normalising flows for unfolding.	172

III.E.1.ii.d	Limitations.	173
III.E.2	Discriminative Models	175
III.E.2.i	Training dynamics and optimization . . .	177
III.E.2.ii	Applications	178
III.E.2.iii	Traditional discriminative architectures .	180
III.E.2.iii.a	Support vector machines	181
III.E.2.iii.b	Boosted decision trees.	183
III.E.2.iv	Discriminative models and deep learning.	186
III.E.2.v	Graph Neural Networks	186
III.E.2.v.a	Transformer Models	189
III.E.3	Adversarial models.	191
III.E.3.i	Generative Adversarial Networks (GANs)	193
III.E.3.i.a	Applications	196
III.E.3.i.a.1	GANs for unfolding.	197
III.E.3.i.b	Limitations	199
III.E.3.ii	Wasserstein GANs.	200

III.E.3.iii	Conditional Adversarial Networks	202
III.E.3.iv	Other GAN based methods	204
III.E.3.v	Adversarial Autoencoders	205
III.E.3.vi	Cycle-Consistent Adversarial Networks	206
III.F	Comparative analysis of ML-based unfolding methods.	208
III.F.1	Method selection framework.	208
III.F.2	Performance trade offs.	210
III.F.2.i	Synthesis of experimental experience.	212
III.F.3	Open challenges and research frontiers.	214
III.G	Case studies: ML based unfolding in HEP analyses.	218
III.G.1	OMNIFOLD	218
III.G.2	NEURAL POSTERIOR UNFOLDING	220
III.G.3	Invertible neural networks for unfolding	224
III.G.4	MOMENT UNFOLDING	225
III.G.5	Reweighting Adversarial Networks (RAN)	227
III.G.6	Schrödinger Bridge Unfolding	228

III.G.7	Unified Neural Folding GAN	230
III.H	ML upstream and downstream of unfolding.	233
III.H.1	Uncertainty quantification in ML based unfolding. .	233
III.H.2	Hybrid approaches and integration with simulation.	235
III.H.3	Equivariant networks for physics-informed unfolding.	236
III.H.4	Comparison with traditional methods.	238
III.H.5	Implementation Considerations from Real Experiments	240
IV	NEURAL POSTERIOR UNFOLDING	241
IV.A	Motivation for improved binned unfolding	242
IV.A.1	Degeneracy and Null Spaces in Unfolding	242
IV.A.2	Regularization Limitations	243
IV.B	Normalising flows for inverse problems.	246
IV.B.1	A Bayesian perspective on inverse problems.	246
IV.B.2	Conditional normalising flows for inverse problems.	248
IV.C	The NEURAL POSTERIOR UNFOLDING Algorithm	251
IV.C.1	Statistical foundation	251

IV.C.2	Machine Learning Architecture	253
IV.C.3	Addressing degeneracy with NPU.	255
IV.C.4	Implicit regularisation in NPU.	257
IV.D	Numerical results.	260
IV.D.1	2 bin degenerate response example.	260
IV.D.2	Gaussian example.	262
IV.D.2.i	Experimental setup	264
IV.D.2.ii	Results and comparison	266
IV.D.3	Particle physics example.	270
IV.D.3.i	Dataset and Observables	272
IV.D.3.ii	Results	273
IV.D.4	Summary of Numerical Results	275
IV.E	Beyond binning: the path forward	279
V	MOMENT UNFOLDING: direct deconvolution of distribu-	
	tion moments	282
V.A	Why moments: physics context and QCD calculations . . .	283

V.A.1	The Theoretical Significance of Moments in Physics	283
V.A.2	Moments in QCD Calculations	285
V.A.2.i	DGLAP Evolution Equations	285
V.A.2.ii	Operator Product Expansion	286
V.A.2.iii	Event Shape Moments	287
V.A.3	Experimental Significance of Moments in QCD . . .	288
V.A.4	Applications in Jet Physics	291
V.A.4.i	Moments in physics beyond the Standard Model	293
V.A.5	The case for direct MOMENT UNFOLDING	294
V.B	A GAN like method to unfold moments	296
V.B.1	Boltzmann inspired reweighting	297
V.B.2	Adversarial training framework	298
V.B.3	Mathematical formalism.	301
V.B.4	Theoretical properties.	303
V.B.4.i	Connection to the maximum entropy principle	304

V.B.5	Comparison with traditional GAN architectures. . .	304
V.B.6	Practical considerations.	306
V.B.7	Extension to differential measurements.	307
V.B.8	Theoretical foundations in statistical mechanics. . .	308
V.C	Machine learning implementation	310
V.C.1	Training Procedure	313
V.C.2	Gradient Updates	314
V.C.3	Implementation details.	315
V.C.4	Extensions to multiple observables.	316
V.C.5	Uncertainty estimation.	318
V.C.6	Validation procedures.	319
V.C.7	Code availability and reproducibility	320
V.D	Case studies	322
V.D.1	Gaussian experiments	322
V.D.1.i	Experimental Setup	322
V.D.1.ii	Results	323

V.D.2	Jet substructure in collider physics	327
V.D.2.i	Datasets	327
V.D.2.ii	Observables	328
V.D.2.iii	Results	330
V.D.3	Momentum dependent unfolding	335
V.D.3.i	Inclusive distributions.	337
V.D.3.ii	Differential analysis.	338
V.E	Comparison with alternative methods.	343
V.F	Conclusion.	348
V.F.1	Towards unfolding distributions.	350
VI	RAN: Reweighting Adversarial Networks	353
VI.A	The need for full spectral measurements.	354
VI.A.1	Motivation: beyond moments and binned spectra. .	354
VI.A.2	Use cases.	361
VI.A.3	Challenges in unfolding full distributions.	367
VI.A.4	Unbinned unfolding approaches.	373

VI.A.5	Addressing challenges.	379
VI.B	From moments to complete differential cross section spectra.	388
VI.B.1	Moments and the full probability distribution.	389
VI.B.1.i	The maximum entropy principle and exponential families.	391
VI.B.2	Unfolding with moment constraints: classical and modern approaches.	393
VI.B.2.i	Linear regularised unfolding (Tikhonov and SVD).	393
VI.B.2.ii	Iterative Bayesian unfolding.	395
VI.B.2.iii	Entropy-based methods.	396
VI.B.2.iv	Unbinned methods through the lens of moments.	398
VI.B.3	Benefits and challenges of MOMENT UNFOLDING.	400
VI.B.4	Extending MOMENT UNFOLDING.	404
VI.C	Methodology and regularisation.	414

VI.C.1	Historical Approaches to Regularisation	414
VI.C.2	Regularisation in RAN	418
VI.C.2.i	Wasserstein GAN	419
VI.C.2.ii	Regularising the critic: Lipschitz constraints	424
VI.C.2.iii	Regularising the generator: initialisation and activation constraints.	426
VI.C.3	Regularising with MC Prior	434
VI.D	Machine learning implementation.	438
VI.D.1	Neural Network Architecture	438
VI.D.2	Adversarial training procedure.	441
VI.D.3	Datasets and preprocessing.	448
VI.D.3.i	Gaussian example.	449
VI.D.3.ii	Jet substructure dataset	451
VI.D.4	Training, monitoring and validation.	455
VI.E	Results	461
VI.E.1	Gaussian model with smearing.	461

VI.E.2	Jet substructure unfolding results.	465
VI.E.2.i	Overall performance.	467
VI.E.2.ii	Generality of results	477
VII	Unbinned Inference on Correlated Data	480
VII.A	Introduction.	481
VII.B	Statistical independence in HEP.	483
VII.B.1	Poisson point processes for event counts.	483
VII.B.2	Independent and identically distributed (i.i.d.) events.	485
VII.B.3	Likelihood factorisation and conditional independence.	486
VII.C	Violation of statistical independence.	489
VII.C.1	Detector effects.	490
VII.C.2	Pileup	492
VII.C.3	Unfolding and data processing.	495
VII.C.4	Global constraints.	499
VII.D	Consequences for inference.	504
VII.E	Formalism.	509

VII.F Uncertainty quantification.	523
VII.F.1 Analytic approaches: asymptotic theory and its limitations.	523
VII.F.2 Godambe information: the sandwich estimator) . .	527
VII.F.3 Wilks' theorem violations and Bartlett corrections .	531
VII.F.4 Numerical resampling approaches: bootstrapping and toy Monte Carlo.	536
VII.F.4.i Non-parametric bootstrap resampling. . .	537
VII.F.4.ii Parametric bootstrap: toy Monte Carlo simulations	544
VII.F.4.iii Practical considerations and guidelines. .	549
VII.F.4.iii.a Number of replicas.	549
VII.F.4.iii.b Computational cost and parallelisation. . . .	550
VII.F.4.iii.c Diagnostics for coverage and reliability. . . .	551
VII.F.4.iii.d Incorporating systematic uncertainties. . . .	552
VII.F.4.iii.e Reporting and using bootstrap results	553

VII.F.4.iii.f	Common pitfalls.	554
VII.G	Case studies.	557
VII.G.1	Setup.	557
VII.G.2	Fully binned baseline	558
VII.G.2.i	Methodology.	559
VII.G.3	Correlation diagnostics after unbinned unfolding . .	571
VII.G.3.i	Pairwise weight–distance correlations . . .	574
VII.G.3.ii	Histogram covariance matrices.	578
VII.G.3.iii	Implications	579
VII.G.4	Unbinned unfolding: binned and unbinned inference.	581
VII.G.4.i	Parameter bias	583
VII.G.4.ii	Uncertainty estimation and coverage. . . .	586
VII.G.5	Extension to higher dimensions.	594
VII.G.5.i	Evaluation.	602
VII.H	Conclusions and outlook.	607
VIII	Symmetries in data: connections to unfolding challenges.	615

VIII.A Symmetries and unfolding.	616
VIII.A.1 The complementary nature of symmetry discovery and unfolding.	616
VIII.A.2 Symmetry aware cross sections.	620
VIII.B Formalism and Importance	629
VIII.B.1 Fundamental symmetries in HEP.	631
VIII.B.2 Symmetries in detector response functions	645
VIII.B.2.i Spatial uniformity and rotational symmetry.	645
VIII.B.2.ii Polar coverage and boost invariance. . . .	647
VIII.B.2.iii Resolution effects and approximate invari- ance.	650
VIII.B.2.iv Mirror and charge symmetry.	652
VIII.B.2.v Permutation symmetry and identical parti- cles.	654
VIII.B.3 How symmetries manifest in measured cross sections.	662
VIII.B.3.i Exact symmetries and flat distributions. .	663

VIII.B.3.ii Symmetries in kinematic shapes.	665
VIII.B.3.iii Interplay of physical and detector symmetries.	666
VIII.B.4 Challenges in identifying symmetries from noisy data.	668
VIII.B.4.i Dimensionality challenges.	671
VIII.C Statistical definition of dataset symmetries.	675
VIII.C.1 Distinction between point and dataset symmetries.	676
VIII.C.2 Inertial reference densities and their theoretical role	678
VIII.C.3 Inertial densities and relative symmetry.	681
VIII.C.3.i Failure of the naïve PDF preserving criterion.	682
VIII.C.3.ii Relative symmetry.	683
VIII.C.3.iii Choice of the inertial density.	684
VIII.C.4 Applications to particle physics.	685
VIII.D SYMMETRYGAN: Discovering Symmetries with Adversarial	
Learning	689
VIII.D.1 The SYMMETRYGAN Architecture	690
VIII.D.2 Machine Learning with Inertial Restrictions	692

VIII.D.3 Deep learning implementation details.	698
VIII.D.4 Verification.	700
VIII.D.5 Other symmetry discovery methods.	701
VIII.E Empirical experiments.	705
VIII.E.1 Gaussian experiments.	705
VIII.E.1.i One-dimensional Gaussian.	706
VIII.E.1.ii Two dimensional Gaussians.	708
VIII.E.1.iii Gaussian Mixture Models	713
VIII.E.2 Particle physics experiments.	718
VIII.E.3 Interpreting discovered symmetries.	728
VIII.F Towards symmetry inference.	730
VIII.G Symmetry informed unfolding	733
VIII.G.1 Symmetry preserving neural network architectures.	735
VIII.G.2 Reducing dimensionality through symmetry identification.	738
VIII.G.3 Hidden symmetries and emergent simplicity.	739

VIII.G.3.i Formalism	740
VIII.G.4 Applications for unfolding jet substructure variables.	743
VIII.H Improved precision using symmetries	745
VIII.H.1 Data augmentation using discovered symmetries.	745
VIII.H.2 Symmetry constrained unfolding.	749
VIII.I Conclusion	752
VIII.I.1 Beyond linear symmetries.	752
VIII.I.2 Approximate symmetries and symmetry breaking.	755
VIII.I.3 A unified framework.	757
IX Conclusion	759
IX.A Synthesis of Unfolding Methods	760
IX.B A Unified Framework for Modern Unfolding	765
IX.B.1 The Interplay of Physics and Learning	766
IX.B.2 Common Themes Across Methods	768
IX.C Advances in Cross Section Measurement	773
IX.C.1 Summary of Contributions	773

IX.C.2	Impact and Advances	776
IX.D	Applications to Current and Future Experiments	781
IX.E	Open Challenges and Future Directions	790
IX.E.1	Challenges in unfolding methodology	790
IX.E.2	Handling correlations in unbinned inference	791
IX.E.3	Background treatment in complex environments	792
IX.E.4	Systematic uncertainty propagation in ML methods	794
IX.E.5	Future Directions	795
IX.E.5.i	Joint unbinned unfolding and nuisance parameter profiling	795
IX.E.5.ii	Theoretical guarantees for ML-based unfolding	797
IX.E.6	Use on real experimental data	798
IX.F	Closing: The Broader Vision	801
References		808

List of Figures

I.1	A schematic illustration of the Standard Model [3].	7
II.1	Coverage properties of Tikhonov regularisation and IBU . . .	91
II.2	Weight correlation between event pairs as a function of distance between events.	95
III.1	Neural network depth and width illustration	115
III.2	Shallow vs. deep NN	116
III.3	Forward pass	119
IV.1	2 bin example	263

IV.2	Gaussian example setup	265
IV.3	Gaussian unfolding results and posterior analysis	267
IV.4	Statistical calibration analysis for NPU and FBU	269
IV.5	Jet substructure unfolding results using NPU, FBU, and IBU .	271
IV.6	Corner plots for jet substructure observables	276
V.1	A schematic diagram of the training setup for MOMENT UNFOLDING. Like a GAN, g is the generator and d is the discriminator, but unlike a GAN g does not generate samples, it instead generates reweighting factors given by Equation (V.9). The reweighted Simulation dataset inherits its weight from the matching Generation dataset. The detector emulations are only run once, since a new simulated dataset is created via importance weights and not by changing the features themselves.	300
V.2	Effect of detector resolution on Truth and Generation datasets	324

V.3	(a) Distributions from the Gaussian example of particle level truth, generation, and reweighted generation i.e. MOMENT UNFOLDING. The agreement between the truth and reweighted samples demonstrates the qualitative performance of MOMENT UNFOLDING. (b) The weighted MLC loss from Equation (V.19) for fixed g but optimized d , found by scanning over β_1 and β_2 . The correct value is indicated by a red star. Indicated in shaded blue is the 1σ bootstrapped interval for MOMENT UNFOLDING's prediction of β_a . The lower panel in each plot shows the ratio to truth.	326
V.4	Jet substructure distributions comparing truth, generation, and MOMENT UNFOLDING results	333
V.5	Loss landscapes showing MOMENT UNFOLDING parameter estimation accuracy for unfolding jet substructure variable.	334
V.6	Momentum dependent unfolding of inclusive jet substructure distributions	339

V.7	Momentum dependent differential moments of jet substructure observables	340
V.8	Comparison of unfolding methods for momentum-dependent jet moments	344
VI.1	Comparison of activation functions for neural network weight parametrisation	433
VI.2	RAN architecture with generator and critic networks	442
VI.3	Unfolding performance comparison between RAN and OmniFold under varying detector smearing	466
VI.4	Unfolded particle level distributions for six jet substructure observables, comparing RAN, OMNIFOLD, and IBU.	468
VII.1	Detector and particle level distributions for the study on correlations during inference.	560

- VII.2 Characterisation of detector response effects through progressive smearing of Monte Carlo simulation. Panel (a) illustrates the impact of finite detector resolution on the observed distributions, showing how the Generation is broadened by convolution with Gaussian response functions of width σ_{det} . The total observed variance follows the quadrature sum $\sigma_{\text{observed}}^2 = \sigma_{\text{particle}}^2 + \sigma_{\text{det}}^2$. Panel (b) presents the corresponding binned response matrices $R(x_i | z_j)$, which encode the probability of bin migration due to detector effects. For perfect resolution ($\sigma_{\text{det}} = 0$), the response matrix is diagonal, indicating no bin migration. As detector resolution degrades, off diagonal elements become increasingly prominent, quantifying the probability of events generated in bin j being reconstructed in neighbouring bins. 561
- VII.3 Comparison of uncertainty estimation methods for binned unfolded distribution parameters as a function of detector resolution. 566

VII.4	Validation of unbiased parameter extraction from unfolded distributions across varying detector resolutions.	567
VII.5	Event wise weight correlations as a function of separation distance for different unfolding methods.	573
VII.6	Covariance structure of binned unfolded distributions for KDE and NN-based methods.	580
VII.7	Comparison between binned and unbinned parameter inference with unbinned and binned unfolding methods as a function of detector resolution.	584
VII.8	Validation of asymptotic uncertainty estimates across dimensionalities in unbinned unfolding.	603
VII.9	Impact of detector resolution on asymptotic uncertainty reliability in 6-dimensional unfolding.	606
VIII.1	Architectural comparison of standard GAN and SymmetryGAN for automated symmetry discovery.	693

VIII.2 Analytic loss landscape for the one-dimensional uniform distribution $\mathcal{N}((0.5, 1))$ under affine transformations $g(x) = b + cx$. The landscape exhibits two distinct maxima (indicated by stars) corresponding to the identity transformation $(b, c) = (0, 1)$ and reflection $(b, c) = (1, -1)$, separated by a deep valley at $c = 0$. This topological barrier creates a disconnected solution space that deterministically routes optimisation trajectories based on initial parameter values.	707
VIII.3 Empirical validation of SYMMETRYGAN on 1 dimensional Gaussian data showing convergence to discrete symmetries based on initial conditions.	709
VIII.4 Loss landscapes demonstrating SYMMETRYGAN's ability to discover continuous $SO(2)$ and discrete V_4 symmetries in isotropic and anisotropic 2D Gaussians.	711
VIII.5 Controlled discretization of continuous $SO(2)$ symmetry into cyclic subgroups \mathbb{Z}_q via constraint enforcement.	712

VIII.6 Orthogonal slices of $\text{Aff}_2(\mathbb{R})$ symmetries for isotropic Gaussian.	714
VIII.7 Discrete symmetry structure of anisotropic Gaussian in $\text{Aff}_2(\mathbb{R})$ discovered by SYMMETRYGAN.	715
VIII.8 SYMMETRYGAN's predictions for a bimodal Gaussian mixture with \mathbb{Z}_2 reflection symmetry and corresponding loss landscape validation.	716
VIII.9 Discovery of dihedral symmetries D_8 and D_4 in octagonal and square lattice Gaussian mixtures.	717

VIII.10 Symmetry discovery on transverse momenta of dijet events using SYMMETRYGAN with a $SO(2) \times SO(2)$ search space, where each jet undergoes independent azimuthal rotation $g_{\theta_1, \theta_2}(\mathbf{p}) = (R(\theta_1)\mathbf{p}_1, R(\theta_2)\mathbf{p}_2)$. (a) Distribution of converged parameters (θ_1^f, θ_2^f) from independent training runs initialised uniformly in $[0, 2\pi)^2$. The discovered symmetries cluster along the diagonal $\theta_1 = \theta_2$ (blue line), confirming that only simultaneous rotations preserve the dijet distribution, a direct consequence of transverse momentum conservation. (b) Symmetry discovery map $\Omega : (\theta_1^i, \theta_2^i) \mapsto \theta^f$ revealing the loss landscape dynamics. The final rotation angle follows $\theta^f = \frac{\theta_1^i + \theta_2^i}{2}$ when $|\theta_1^i - \theta_2^i| < \pi$ (shown in green), and $\theta^f = \frac{\theta_1^i + \theta_2^i}{2} - \pi$ when $|\theta_1^i - \theta_2^i| > \pi$ (shown in blue). This bisection rule represents the path of steepest ascent in the loss landscape, demonstrating how gradient dynamics naturally discover the physical constraint without prior knowledge. . . . 720

VIII.11	Visual validation of discovered dijet symmetry through transverse momentum preservation.	723
VIII.12	Non-uniform azimuthal distributions from random $SO(4)$ rotations demonstrating symmetry violation, and uniform azimuthal distributions from discovered symmetry confirming invariance of dijet system.	724
VIII.13	Kullback-Leibler divergence analysis showing discovered symmetries approach statistical noise floor.	726

VIII.14	Discriminator loss distributions validating discovered symmetries against theoretical optimum. The histogram shows <i>post hoc</i> discriminator losses when trained to distinguish original from transformed dijet events. Discovered symmetries (blue) achieve losses tightly clustered around $2 \log 2 \approx 1.386$ (dashed line), the theoretical minimum for perfect symmetries where transformed and original distributions are indistinguishable. Random $SO(4)$ rotations (black) yield significantly lower losses, indicating the discriminator easily identifies these as non-symmetries. The discovered transformations achieve losses within 1% of the theoretical optimum, providing independent confirmation that SYMMETRYGAN identifies genuine invariances of the dijet system without prior physics knowledge.	727
---------	--	-----

List of Tables

I.1	Unit conversions in natural units	3
II.1	Comparing the MC dependence and uncertainty propagation of traditional unfolding methods.	67
II.2	Distributional distance metrics for unfolding evaluation	87
IV.1	Computational cost comparison between FBU and NPU . . .	268
V.1	Summary of training hyperparameters used in the model. These values control the architecture, optimization behaviour, and regularization of the training process.	317

V.2	Moments of jet observables at particle level. First and second moments of m, q, w and z_g are shown for truth (HERWIG), generation (PYTHIA) and MOMENT UNFOLDING. Uncertainties in the truth and generation columns are estimated via bootstrap resampling; uncertainties in the unfolding column combine in quadrature the generation bootstrap uncertainty with the empirical 1σ spread from repeated unfolding on the same dataset.	331
VI.1	Comparison of measurement approaches. Reporting only low-order moments loses most distribution information. Binned differential cross sections retain shape information but suffer from discretization and limited dimensionality. Full spectral measurements preserve the complete distribution, enabling maximal reusability and detailed theory comparisons, but require regularizing a much more ill-posed training.	360
VI.2	Comparison of unfolding method characteristics	379

VI.3	Wasserstein distances between unfolded and true distributions	472
VI.4	VLC divergence between unfolded and true distributions . . .	473
VII.1	Sources of correlated uncertainties in cross section measurements	508
VII.2	Gaussian model parameters for multidimensional unfolding validation studies.	596
VIII.1	Comparative analysis of unfolding and symmetry discovery methods in particle physics.	621
VIII.2	Fundamental symmetries and their manifestations in HEP observables.	638
VIII.3	Symmetry breaking induced by detector effects in HEP experiments.	658
IX.1	Applications of methods to experimental scenarios	783

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.

The beauty of natural units emerges when considering quantum field theory. The Klein-Gordon equation for a scalar field : $(\partial^2/\partial t^2 - \nabla^2 + m^2) = 0$ reveals that mass m has dimensions of inverse length - it sets the characteristic

Table I.1: Unit conversions in natural units, expressing common physical quantities in terms of GeV-based natural units alongside their SI equivalents and their physical significance.

Quantity	Natural Units	SI Conversion	Physical Insight
Energy	GeV	$1 \text{ GeV} = 1.602 \times 10^{-10} \text{ J}$	Natural scale of nuclear binding
Mass	GeV/c^2	$1 \text{ GeV}/c^2 = 1.783 \times 10^{-27} \text{ kg}$	Proton mass $\approx 0.938 \text{ GeV}$
Momentum	GeV/c	$1 \text{ GeV}/c = 5.344 \times 10^{-19} \text{ kg m/s}$	Typical LHC particle momentum
Length	GeV^{-1}	$1 \text{ GeV}^{-1} = 1.973 \times 10^{-16} \text{ m}$	Compton wavelength scale
Time	GeV^{-1}	$1 \text{ GeV}^{-1} = 6.582 \times 10^{-25} \text{ s}$	Natural oscillation period
Cross Section	GeV^{-2}	$1 \text{ GeV}^{-2} = 3.894 \times 10^{-32} \text{ m}^2$ $= 1 \text{ barn} = 2.568 \times 10^{-9} \text{ GeV}^{-2}$	Typical nuclear/hadronic cross sections

scale over which quantum fields oscillate. A particle with mass 1 GeV has a Compton wavelength of approximately 0.2 femtometers, roughly the size of a proton. Coordinate Systems and Relativistic Geometry At a hadron collider, the choice of coordinates must respect the underlying symmetries of the physics. The laboratory frame provides a natural Cartesian system:

z-axis: Along the beam direction (longitudinal) x-axis: Horizontal, pointing toward the center of the accelerator ring y-axis: Vertical, completing the right-handed system

However, the partons within colliding hadrons carry unknown fractions of the beam momentum, making the center-of-mass frame of each collision unknowable. This fundamental uncertainty motivates a coordinate system that transforms simply under longitudinal Lorentz boosts.

Pseudorapidity (η): Defined as $\eta = -\ln[\tan(\theta/2)]$, where θ is the polar angle from the beam axis. Under a longitudinal boost with velocity β , a massless particle's η transforms as $\eta \rightarrow \eta + \tanh^{-1}(\beta)$.

The mathematical elegance of pseudorapidity becomes apparent when we expand its definition: $\eta = -\ln[\tan(\theta/2)] = (1/2)\ln[(|p| + p_z)/(|p| - p_z)] = \tanh^{-1}(p_z/|p|)$ For massless particles, differences in η remain invariant under longitudinal boosts, motivating the ubiquitous angular distance metric: $R^2 = \Delta\eta^2 + \Delta\phi^2$ 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.

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 20th century, it provides a quantum field theory framework that describes three of the four known fundamental forces—the electromagnetic, weak, and strong interactions—along with 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 $SU(3) \times SU(2) \times U(1)$.

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

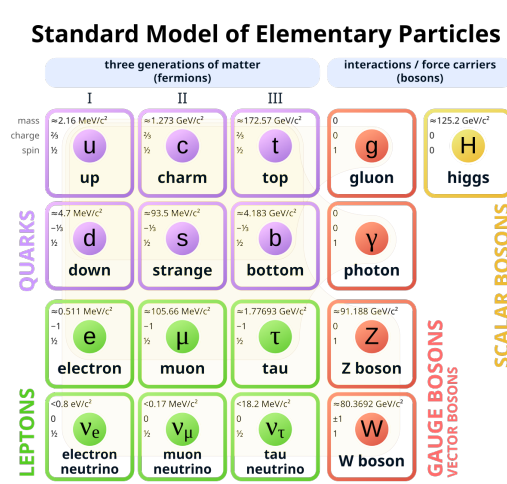


Figure I.1: A schematic illustration of the Standard Model [3].

Higgs boson in 2012 at the Large Hadron Collider (LHC) [1, 2], confirming the mechanism through which elementary particles acquire mass. A schematic illustration of the standard model can be found in Figure I.1.

I.B.1 Fundamental Particles and Forces

The Standard Model categorizes 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 are further classified into quarks and leptons, each arranged in three generations of increasing mass:

I.B.1.i.a Quarks

(spin = $1/2$) Quarks are categorized 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 color charge, and experience all fundamental forces. They are confined within hadrons—composite particles categorized as baryons (three-quark states, like protons and neutrons) or mesons (quark–antiquark pairs).

I.B.1.i.b Leptons

(spin = $1/2$) Like quarks, leptons too are categorized into three generations.

- 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 via 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. The Standard Model comprises the following bosons:

- Photon (γ): The massless spin-1 boson mediating the electromagnetic force,
- W^\pm and Z bosons: Massive spin-1 bosons mediating the weak force
- Gluons (g): Eight massless spin-1 bosons mediating the strong force
- Higgs boson (H): 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

- $U(1)_Y$: Associated with weak hypercharge, the symmetry of electroweak theory
- $SU(2)_L$: Describes the weak isospin, acting on left-handed fermions
- $SU(3)_C$: Governs the strong interactions through color charge in QCD

Electroweak unification, demonstrated by Glashow [4], Weinberg [5], and Salam [6], 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 several physicists including Peter Higgs in the 1960s [7], 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 breaks 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 [8], 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-recognized limitations, including

- It does not incorporate gravity, the fourth fundamental force.
- It fails to explain the observed matter–antimatter asymmetry in the universe.
- It does not account for dark matter or dark energy, which together constitute about 95% of the universe’s energy content.
- It requires fine-tuning of parameters, raising theoretical concerns like the hierarchy problem.
- 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, where $1 \text{ barn} = 10^{-24} \text{cm}^2$). This seemingly simple concept forms the cornerstone of how we test and validate our understanding of fundamental physics.

The Standard Model of particle physics—our most successful theory describing elementary particles and their interactions—makes precise predictions for cross sections that can be directly tested at particle physics experiments.

Any statistically significant deviation between measured cross sections and theoretical predictions may signal the presence of new physics beyond the Standard Model [9].¹

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 [10], while precision electroweak cross section measurements constrain the properties of the Higgs boson and other fundamental particles [11]. 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 parameterize potential new physics in a model-independent way. By measuring differential distributions with high precision,

¹Such 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 [12]. 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 [13] and consistency checks [14] 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. 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, (e.g., electromagnetic or gravitational interactions), the cross section is larger than the physical dimensions of the particles due to action-at-a-distance effects. 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

$$\frac{d\sigma}{d\Omega} = \frac{\text{Number of events scattered into } d\Omega}{\text{Incident flux} \times \text{Target density}}. \quad (\text{I.1})$$

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

$$\sigma = \int_{4\pi} \frac{d\sigma}{d\Omega} d\Omega. \quad (\text{I.2})$$

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 $2 \rightarrow n$ reaction with incoming four--momenta $p_{1,2}$ and final-state phase space $d\Phi_n$, the fully differential cross section is

$$d\sigma = \frac{(2\pi)^4 \delta^{(4)}(p_1 + p_2 - \sum_{i=1}^n p_i)}{4 \sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} |\mathcal{M}|^2 d\Phi_n, \quad (\text{I.3})$$

where the denominator is the flux factor and $d\Phi_n = \prod_{i=1}^n \frac{d^3 p_i}{(2\pi)^3 2E_i}$ is the Lorentz--invariant phase0-space element.² Equation (I.3) reduces to the

²Standard derivations can be found in [15–18].

classical area when $|\mathcal{M}|^2$ is replaced by a contact interaction and the final-state integral collapses to a single kinematic configuration. Integrating Equation (I.3) over final-state kinematics yields the total cross section, $\sigma = \int d\sigma$. 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 one folds $|\mathcal{M}|^2$ with parton distribution functions (PDFs) and convolutes the result with detector response—precisely the forward problem addressed by the unfolding methods developed in this thesis.

Differential cross sections have a long history of providing valuable insights for probing fundamental properties of particles and interactions. For example, Rutherford scattering experiments revealed the existence of atomic nuclei by analysing angular distributions of scattered alpha particles [19].

I.D.2 Experimental Measurement

At modern colliders³ the two beams themselves act as both “projectile” and “target.” The fundamental experimental quantity is the instantaneous *luminosity* $\mathcal{L}(t)$, defined such that the interaction rate for a process with cross section σ is $dN/dt = \mathcal{L}(t) \sigma$.

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

$$\mathcal{L}_{\text{int}} = \int_{t_0}^{t_f} \mathcal{L}(t) dt, \quad \sigma = \frac{N_{\text{obs}} - N_{\text{bkg}}}{\mathcal{L}_{\text{int}} \epsilon A}. \quad (\text{I.4})$$

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.

³This section focusses on collider experiments, but similar analyses can be applied to non-collider HEP experiments as well, such as fixed target experiments. [20] and [9] chapters 31–38 provide a comprehensive and detailed exposition of experimental measurement in HEP. [21, 22] focus specifically on fixed target techniques and phenomenology. [23] compares collider and fixed-target formalisms, reviews luminosity analogues, and details the role of detector simulations and unfolding in a fixed-target context

For binned measurements one bins events in an observable X (e.g. transverse momentum p_T or rapidity y) and divides by the bin width

$$\left. \frac{d\sigma}{dX} \right|_{X_i} = \frac{1}{\mathcal{L}_{\text{int}} \Delta X_i} \frac{N_i^{\text{obs}} - N_i^{\text{bkg}}}{\epsilon_i A_i}, \quad (\text{I.5})$$

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 pile--up counting techniques. The efficiency--acceptance term ϵA is obtained from full detector simulations and corrected in data using control samples and “tag--and--probe” methods.

Equations (I.4)–(I.5) thus link the theoretically calculated parton-level cross sections (cf. Eq. (I.3)) to the raw observables recorded by the a detector, completing the chain from theory to experiment.

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

Comparing measured cross sections with predictions from quantum field theory validates and tests theoretical models like Quantum Chromodynamics (QCD) 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 experimen-

tal design and analysis techniques to account for systematic uncertainties introduced by detector effects.

I.D.3.ii Monte Carlo Tuning

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. Many 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 hadronization to detector response. These simulations contain dozens of phenomenological parameters, from the strong coupling constant at various scales to non-perturbative fragmentation functions. 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 our models succeed and where they fail. A discrepancy in the high- p_T tail might indicate we need to adjust our modelling of initial-state 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 to ensure measurement reliability and identify systematic biases. Cross sections also serve as essential 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 our 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, triggers investigations that can reveal subtle systematic effects or push calculations to higher precision.

The 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 $2 - 3\%$, 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 these 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

Parton luminosity is 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 $Z \rightarrow \mu\mu$ 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 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 example, the surprisingly large Higgs production cross section at the LHC, for instance, has already influenced design considerations for future electron-positron Higgs factories.

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. But 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 our understanding, calibrate our tools, and point the way toward new discoveries. This multiplicative utility explains why cross section measurements, even of well-studied processes, remain priorities for 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 apparatus—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 [24].

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

1. Detection incompleteness: $p(\text{measured}|\text{true}) < 1$ True events that fail to be recorded
2. Measurement impurity: $p(\text{true}|\text{measured}) < 1$ Recorded events that represent background contamination

This reveals 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 signal processing engineers call the false positive rate.

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

$$p(x) = \int r(x|z) p(z) dz \quad (\text{I.6})$$

Where $p(x)$ is the detector level density, $p(z)$ is the particle level density and $r(x|z)$ 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 $\int r(x|z) \, dx < 1$ 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 regularization methods that impose additional constraints such as smoothness assumptions, positivity requirements, and prior knowledge about the expected signal shape. These constraints 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 analyze 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 modeling.
- **Computational constraints:** Detailed simulation of detector response requires substantial computing resources, limiting the statistical precision of response modeling.

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 [25–27].

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, analyzing, 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 binning-free 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 normalizing flows and amortized 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. Specific physics impacts include:

- Improved precision in jet substructure measurements, enabling better discrimination between different theoretical models of QCD radiation
- Enhanced sensitivity to effective field theory parameters through direct moment unfolding
- More robust uncertainty quantification in high-dimensional measurements

- Computational efficiency gains allowing for more detailed systematic studies
- 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 our 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 interpretation.

References

- [1] 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].
- [2] 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).

- [3] 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`.
- [4] S. L. Glashow. “Partial Symmetries of Weak Interactions”. In: *Nucl. Phys.* 22 (1961), pp. 579–588. DOI: 10.1016/0029-5582(61)90469-2.
- [5] 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.
- [6] 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.
- [7] 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>.
- [8] 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.
- [9] 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).
- [10] 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 appraisal”. 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](https://doi.org/10.1007/JHEP07(2025)067) (visited on 07/13/2025).
- [11] “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).
- [12] 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](https://doi.org/10.1007/JHEP07(2016)144) (visited on 07/13/2025).
- [13] 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).

- [14] A. Buckley, J. Butterworth, J. Egan, C. Gutschow, S. Jeon, M. Hagedank, 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).
- [15] 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.
- [16] 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>.

- [17] 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.
- [18] *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>.
- [19] 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>.

- [20] 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).
- [21] 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.
- [22] P. Avery and A. Korytov. “Cross section, Flux, Luminosity, Scattering Rates”. In: ().
- [23] 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: ().

- [24] 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^- Annihilation at $\sqrt{s}=10.52\text{ GeV}$ ”. 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).
- [25] 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.
- [26] 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 Modern 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).
- [27] 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).