

UC Berkeley

UC Berkeley Previously Published Works

Title

Optimizing observables with machine learning for better unfolding

Permalink

<https://escholarship.org/uc/item/9d89c1z4>

Journal

Journal of Instrumentation, 17(07)

ISSN

1748-0221

Authors

Arratia, Miguel

Britzger, Daniel

Long, Owen

et al.

Publication Date

2022-07-01

DOI

10.1088/1748-0221/17/07/p07009

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

PAPER • OPEN ACCESS

Optimizing observables with machine learning for better unfolding

To cite this article: Miguel Arratia *et al* 2022 *JINST* 17 P07009

View the [article online](#) for updates and enhancements.

You may also like

- [Activity statistics of a forced elastic string in a disordered medium](#)
Marta S de la Lama, Juan M López, José J Ramasco *et al.*
- [Construction of spatial scale-free networks on a lattice by using a democratic rule](#)
A K Nandi and S S Manna
- [Shortest node-disjoint paths on random graphs](#)
C De Bacco, S Franz, D Saad *et al.*



The Electrochemical Society
Advancing solid state & electrochemical science & technology

243rd ECS Meeting with SOFC-XVIII

Boston, MA • May 28 – June 2, 2023

**Abstract Submission Extended
Deadline: December 16**

[Learn more and submit!](#)

Optimizing observables with machine learning for better unfolding

Miguel Arratia,^{a,b} Daniel Britzger,^{c,*} Owen Long^a and Benjamin Nachman^{d,e}

^aDepartment of Physics and Astronomy, University of California, Riverside, CA 92521, U.S.A.

^bThomas Jefferson National Accelerator Facility, 12000 Jefferson Ave., Newport News, VA 23606, U.S.A.

^cMax-Planck-Institut für Physik, Föhringer Ring 6, 80805 München, Germany

^dPhysics Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, U.S.A.

^eBerkeley Institute for Data Science, University of California, Berkeley, CA 94720, U.S.A.

E-mail: britzger@mpp.mpg.de

ABSTRACT: Most measurements in particle and nuclear physics use matrix-based unfolding algorithms to correct for detector effects. In nearly all cases, the observable is defined analogously at the particle and detector level. We point out that while the particle-level observable needs to be physically motivated to link with theory, the detector-level need not be and can be optimized. We show that using deep learning to define detector-level observables has the capability to improve the measurement when combined with standard unfolding methods.

KEYWORDS: Analysis and statistical methods; Large detector-systems performance; Performance of High Energy Physics Detectors; Large detector systems for particle and astroparticle physics

ARXIV EPRINT: [2203.16722](https://arxiv.org/abs/2203.16722)

*Corresponding author.



Contents

1	Introduction	1
2	Machine learning-assisted reconstruction for better unfolding	3
2.1	Technical setup	3
2.2	Simulated datasets	4
3	Numerical results	4
3.1	Neutral-current DIS kinematic observables	5
3.2	Event shapes	8
4	Summary and outlook	11
A	The influence matrix	12
B	Closure tests for rapgap	13

1 Introduction

The measurement of cross sections in particle and nuclear physics experiments usually proceeds as a two step process. First, an observable is constructed and secondly the data are corrected for detector effects (see e.g., [1]). The latter step is the so-called *unfolding* and is based on Monte Carlo event simulations to correct the data for limited resolution, acceptance effects or mis-tagging. The process begins by defining the observable of interest $\mathcal{O}_p : \mathbb{R}^{m_p} \rightarrow \mathbb{R}^{n_p}$, where the particle-level input $x_p \in \mathbb{R}^{m_p}$ can be a list of four-vectors, and usually $n_p = 1$. For example, if \mathcal{O}_p is the p_T of the leading jet, then \mathcal{O}_p acts on all the particles in the event, runs jet clustering, and outputs the p_T of the hardest jet. Then, an observable $\mathcal{O}_d : \mathbb{R}^{m_d} \rightarrow \mathbb{R}^{n_d}$ is defined analogously at detector-level, e.g. using energy-flow objects. The definition of the observable at detector level should yield a reasonable resolution, corresponding to a strong correlation between the detector-level and particle-level observables. Next, a histogram is filled at detector level from the event counts. An event simulation is then used to populate a matrix quantifying detector distortions by linking the event counts between particle-level and detector-level bins. Unfolding is then the art of regularized matrix inversion (see ref. [2–5] for recent reviews).

In most cases, \mathcal{O}_d is defined using detector-level quantities as input to the equivalent definition of the observable at particle-level: $\mathcal{O}_d(x_d) \equiv \mathcal{O}_p(x_d)$. However, although this often yields sufficient resolution, a drawback of this approach is that further quantities are commonly not included in the unfolding, nor in the reconstruction, which then can degrade performance. For example, if the jet energy response worsens at high pseudorapidity, η , then the measurement of jet p_T would not be

optimal, since η is not used in the calculation of O and it is also excluded in the unfolding¹. In this situation, it may be advantageous to consider η when defining O_d or to simultaneously unfold p_T and η (as e.g. done in refs. [15, 16]). The same reasoning would apply to any variable driving jet performance. However, once more variables are involved, this leads to a dimensionality problem, or to limited statistics of data and/or the simulations.

Matrix-based unfolding methods do not scale well to many dimensions. This is because the migration matrices become large when many quantities are considered [15–17] or it is technically challenging to consider migrations in further quantities [18–20]. Thus a variety of multidimensional machine learning (ML) techniques have been proposed [21–24], and first ML-based unfolding techniques were already applied to collider data [25]. However, these pose computational challenges, or are often difficult to implement in existing analysis workflows, and thus low-dimensional unfolding might still continue to be the most prevalent approach in the near future. Additionally, ML approaches still focus on the case $O_d(x_d) = O_p(x_d)$, even if O is multi-dimensional.

We note that there is freedom to define the detector-level quantity O_d , and it need not be a simple analog of O_p . In fact, O_d could even have a different dimensionality than O_p . The only core requirement is that the particle-level observable O_p is linked to theory, and is a useful physics quantity.

We propose to use deep learning to define an improved definition of O_d , given the aim of measuring O_p and using a large number of inputs to better account for detector effects. We use a regression model to learn O_p given x_d , with O_p being the observable for a cross section measurement.² Using standard loss functions, direct regression is prior-dependent so we use standard unfolding techniques to mitigate it.³ Our approach to define O_d is compatible with both binned and unbinned [21–24, 54–64] unfolding methods. The workflow is schematically displayed in figure 1 and compared to other reconstruction and unfolding procedures.

While we are not aware of other proposals to use ML for observable reconstruction, a variety of related ideas have been studied. For example, the Simplified Template Cross Section (STXS) protocol for Higgs studies [65, 66] identifies bins of O_p , but does not specify O_d . Recent ATLAS [67] and CMS [68] measurements have used Boosted Decision Trees to construct O_d as a classification problem with the STXS bins as classes. A related idea was presented in ref. [58], with an iterative, binned unfolding based on ML classifiers. The results ref. [58] are one-dimensional, but the author mentions adding additional features.

The remainder of this paper is organized as follows. Section 2 provides the technical details of our approach. We then provide explicit examples using a full detector simulation in section 3. The paper ends with conclusions and outlook in section 4.

¹These effects are suppressed using dedicated energy calibrations that depend on many features to improve their resolution [6–10]. However, most observables lack dedicated calibrations; for example, ATLAS [11] and CMS [12] do not have a dedicated calibration for the n -subjettiness [13, 14].

²Many proposals do this for observables used in other tasks [26–52] (see also ref. [53]).

³There is a close connection between reconstruction and unfolding. When the reconstruction has no noise, then unfolding is unnecessary. In general, reconstruction moves around the features while unfolding moves around the cross section (e.g. x -axis versus y -axis of a histogram). We thank J. Thaler for many useful discussions about this connection.

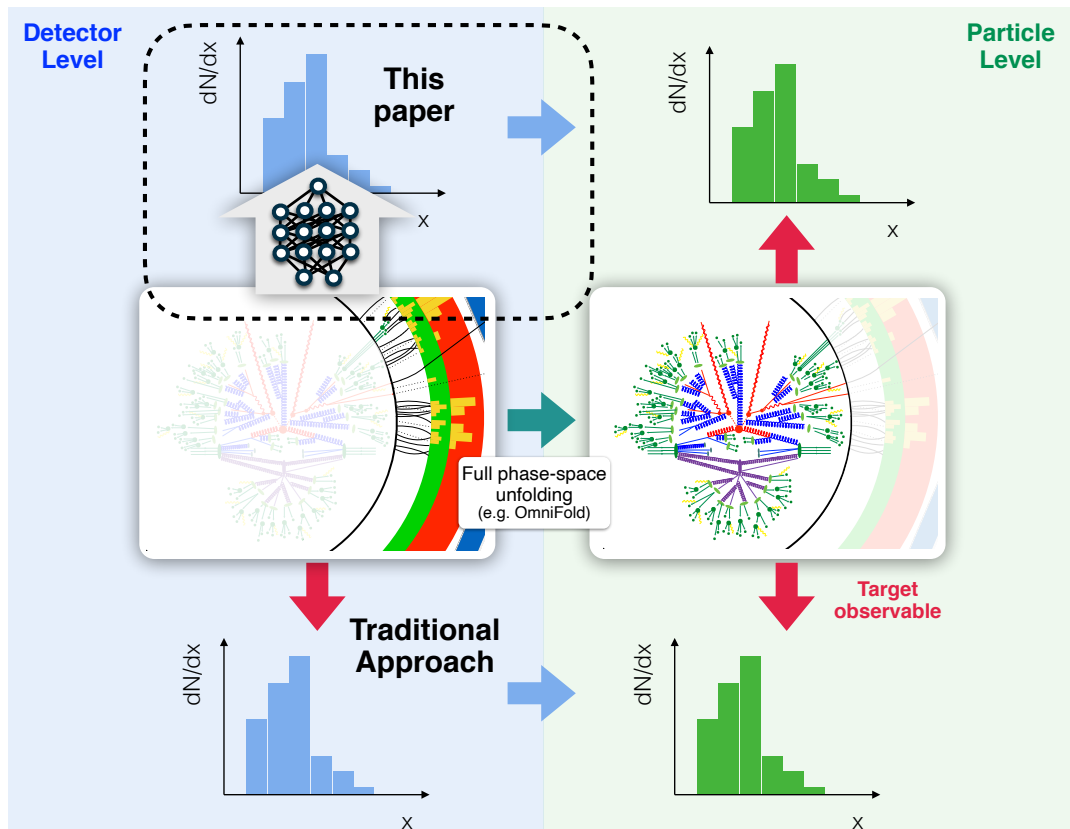


Figure 1. A schematic diagram of unfolding methods, including the one proposed in this paper. The left-hand side images represent detector-level events while the right-hand side images correspond to particle-level events. Horizontal arrows correspond to cross section corrections (e.g., via regularized matrix inversion) while vertical arrows correspond to observable definitions. The traditional path uses the same observable definition at particle-level and detector-level while we propose to use machine learning to define the detector-level observable. The middle horizontal arrow correspond to a full phase space unfolding where observables (and bins) can be constructed after the unfolding (see ref. [21]). Such an approach is complementary to the one proposed in this paper.

2 Machine learning-assisted reconstruction for better unfolding

2.1 Technical setup

The focus will be on the combination of an ML-based reconstruction of an observable and the subsequent application of an unfolding algorithm using summary statistics. It was already demonstrated in ref. [46] that ML-reconstructed observables have the capability to improve the resolution and reduce the bias when compared to classically calculated observables. Hence, we build our study upon this previous ML-setup, and extend it to further observables, and study the unfolding properties of such observables.

The most widely-used regularized unfolding methods are TUNFOLD [69], Iterative Bayesian (also known as Lucy-Richardson) [70–72], and SVD [73]. We use TUNFOLD, which is based on a least-squares fit that includes Tikhonov regularization [74, 75] to control statistical fluctuations. We

use version 17.6 of the TUNFOLD package in the 6.24 distribution of ROOT [76]. The response matrix is constructed from 2D histograms with 10 bins in the particle-level value and 20 bins of the detector-level value. The input to the unfolding is a pseudo-dataset from the MC sample, and the migration matrix is defined by a statistically independent sample, or, alternatively, using an independent MC generator. The scan for the optimal value of the regularization parameter τ is performed with the ScanLcurve method with 30 points [69]. All other settings of TUNFOLD are set to their default values.

All deep neural network (DNN) models are implemented in KERAS [77] and TENSORFLOW [78] and optimized using ADAM [79] with the HUBER [80] loss function as described in our earlier work [46]. The network models use 7 sequential hidden layers with between 64 and 1024 nodes per layer.

The metrics that we will use to compare the standard reconstruction and the ML-based reconstruction are the statistical uncertainties and global correlation coefficients [69] for the unfolded distribution. We also compare the response matrices for the methods and the matrix of correlation coefficients for the unfolded results.

2.2 Simulated datasets

We study simulated data of electron-proton collisions at a center-of-mass energy of $\sqrt{s} = 319$ GeV from the H1 experiment at HERA [81, 82]. These simulated data are well understood and all aspects of the data are in general in good agreement with data from the real experiment, and the H1 detector simulation includes all sub-detectors and run-dependent effects. The simulated data are the same as were used in ref. [46] and are briefly described in the following.

Two simulated samples of deep-inelastic scattering (DIS) were created by the H1 Collaboration, each with a different event generator: RAPGAP 3.1 [83] or DJANGO 1.4 [84], with beam energies $E_e = 27.6$ GeV and $E_p = 920$ GeV, for the lepton and proton, respectively. Both generators use the HERACLES routines [85–87] for Quantum Electrodynamics (QED) radiation. The simulated events are reconstructed in the same way as data. An energy-flow algorithm [88–90] is used to define objects whose sum yields the Hadronic Final State (HFS) four-vector. The scattered electron candidate is defined using the standard H1 approach [15, 91, 92]. Standard selections [15, 91] are applied to suppress backgrounds and mis-measured events. The simulated events are processed by H1’s computing environment [93]. Altogether, $O(10^8)$ events were simulated for each generator. The kinematic region is defined through $Q^2 > 220$ GeV² and for the purpose of a reduction of QED initial-state radiation, the reconstructed events must fulfill 45 GeV $< E - P_z < 62$ GeV, where E and P_z are calculated from the sum of the 4-vectors of the scattered electron and the HFS. This requirement, as well as detector specific acceptance losses, like those from non-instrumented regions (cracks), and cuts to reduce backgrounds, will be taken into account as acceptance corrections.

3 Numerical results

In ref. [46], we introduced a ML approach to reconstruct observables, and applied it to reconstruct the event kinematics in neutral-current DIS events, which are the momentum transfer squared Q^2 , the longitudinal momentum fraction x , and the inelasticity y . Since the benefit was most distinct in x and y , and the ML-reconstructed observables promised a possible extension of the measurement phase space, we will focus on these two DIS kinematic observables in the following. Subsequently we will then study the properties of our method when applied to the 1-jettiness event shape observable τ_1^b .

3.1 Neutral-current DIS kinematic observables

Inputs to the regression DNN are the 4-vectors of the scattered electron, hadronic final state (HFS), and photons identified as QED radiation candidates. The HFS is defined as the sum of all particle candidates that are not associated with the scattered electron. Outputs of the regression DNN are the DIS kinematic variables Q^2 , y , and x . In the presence of QED radiation, the exact definition of the observables at particle level is described and discussed in ref. [46]. Likewise as in ref. [46], the DNN method yields better resolution and smaller biases than any other classical reconstruction methods and the DNN reconstructed observables benefit from the usage of more measured quantities for the calculation of the observables in comparison to classical measurements, and from the inherent classification of QED radiative effects.

Next, we study examples of unfolding in one dimension and quantify the properties of such ML-assisted observables. The 10×20 migration matrices are defined for x in the range $-2.5 < \log_{10}(x) < 1$, and for y in the range $-2.3 < \log_{10}(y) < -0.2$, and the regularization parameter is obtained from the L -curve scan. Figures 2 and 3 show the normalized response matrices and the results of the unfolding for $\log_{10}(x)$ and $\log_{10}(y)$, respectively. Each figure compares the DNN-based reconstruction method with two standard methods: the electron method, and Sigma method [94]. The electron method only uses the scattered electron to compute the DIS kinematic quantities, and the Sigma method is built from HFS inputs. At particle level, all three reconstruction methods use the equivalent particle-level definition for x and y [46]. Hence, only the definition of x_d and $O_d(x_d)$ changes, but not $O_p(x_p)$.

The response matrix for the DNN method is the most diagonal, as expected, since the training of the DNN was performed against the particle-level observables. In contrast, the two classical reconstruction methods have not only larger off-diagonal elements, but also some asymmetries. All three methods have reasonably good resolution at low x , and at high y , while the electron and Sigma methods suffer from poor resolution at high x and low y . The so-called *influence matrix* or *posterior response matrix* can be considered as an alternative representation of the migration matrix and it includes the effect of the regularization [95]. These matrices, and some properties, are discussed in appendix A.

The unfolded distribution for each method agrees well with the generated (Gen) distribution within uncertainties, demonstrating good closure of the unfolding. The classical methods have sizable statistical uncertainties at high x or low y , while those from the DNN method are visibly smaller. When regarding the related correlation matrices, the electron and Sigma methods show significant correlations beyond neighboring bins in areas where the resolution is poor, and even some correlations between distant bins are observed. In contrast, the DNN method has the most diagonal matrix, where the non-zero off-diagonal elements are mostly small correlations between neighboring bins, and no correlations between far apart bins is observed. The average global correlation coefficients ρ_{avg} [69] of these matrices are presented in table 1, and a clear reduction is observed when using the DNN method. In summary, the DNN method directly improves the resolution, which is seen from reduced correlations, and improves mis-reconstruction (e.g. in the presence of QED radiation), which is seen from the absence of distant correlations.

Figure 4 examines the statistical errors and the global correlation coefficients [69] of the unfolding results for the electron, Sigma, and DNN methods in greater detail. At low x and high y , all three methods have good resolution, which results in similar size statistical errors. However,

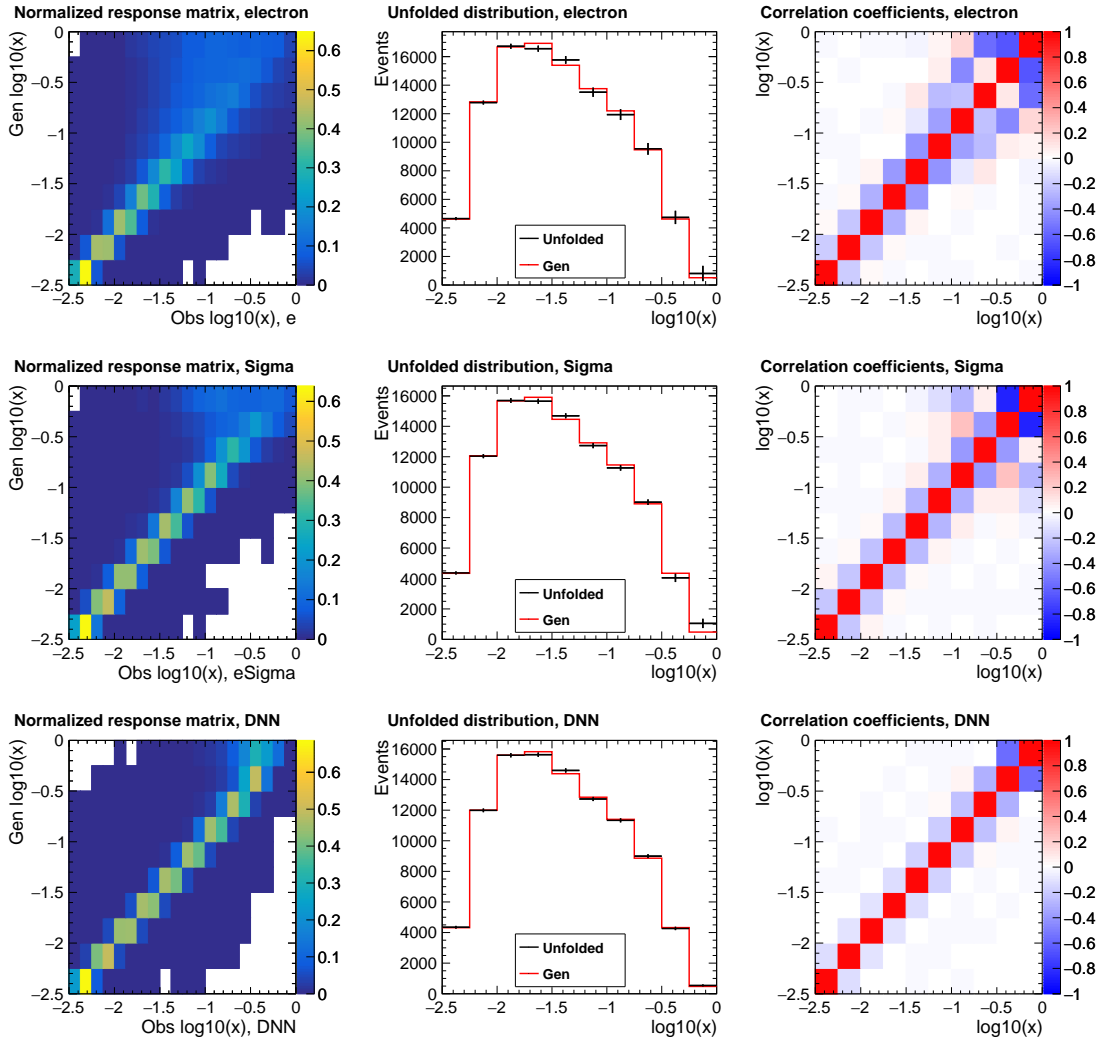


Figure 2. Examples of unfolding $\log_{10}(x)$ for samples of 10^5 events. The response matrix (left), unfolded and gen distributions (middle), and unfolding correlation matrix (right) are shown for the electron (top), Sigma (middle), and DNN (bottom) methods.

Table 1. Value of the average global correlation coefficient ρ_{avg} for the unfolding of x and y when using the electron, Sigma or DNN reconstruction method.

Observable	Reconstruction method		
	electron	Sigma	DNN
x	0.692	0.611	0.400
y	0.837	0.707	0.442

the uncertainties of the DNN methods are still about 10 % smaller at lowest x or highest y , and the global correlation coefficient for the DNN method is found to be smaller by about a factor of 2 or more. At high x and low y , the DNN method shows significantly lower statistical errors and smaller global correlations in all bins. The uncertainties of the electron method become large, due to the

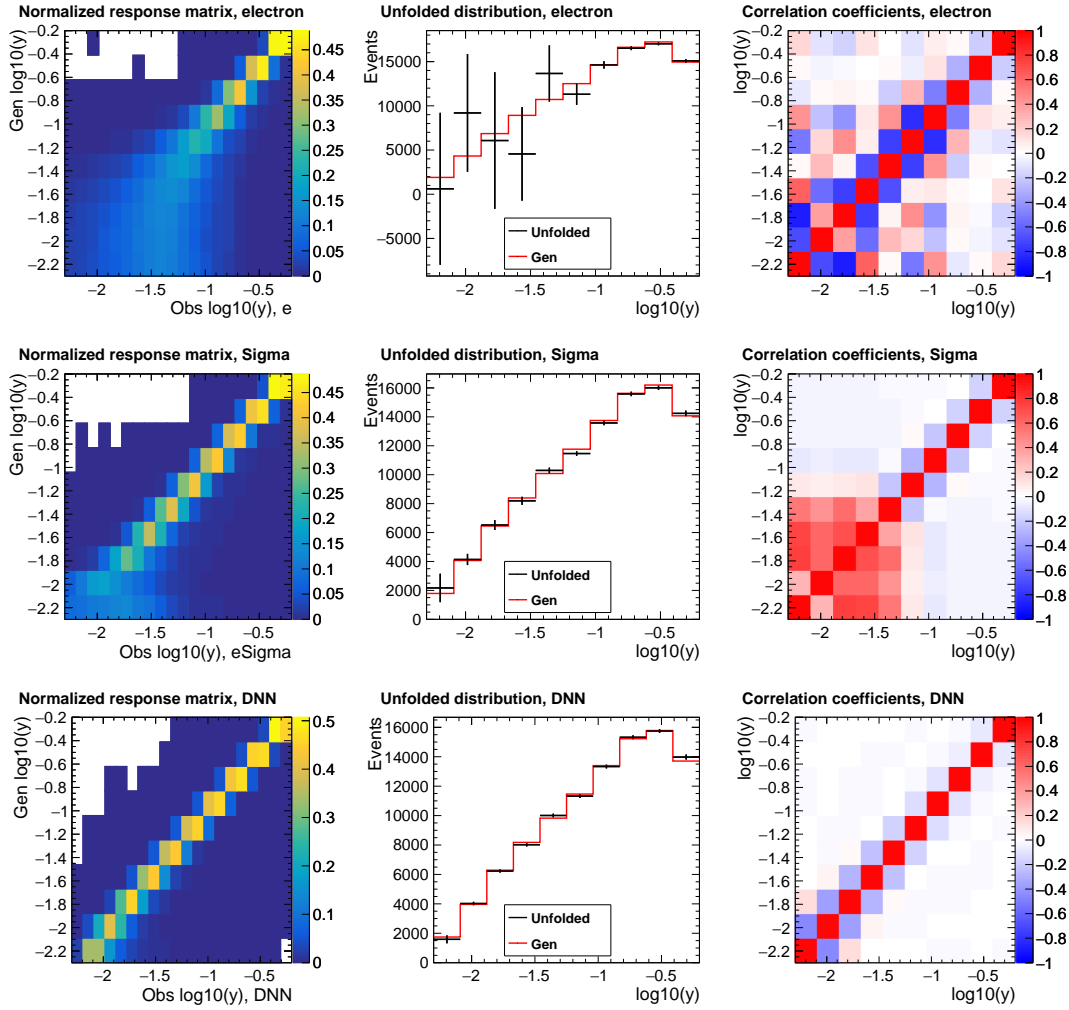


Figure 3. Examples of unfolding $\log_{10}(y)$ for samples of 10^5 events. The response matrix (left), unfolded and gen distributions (middle), and unfolding correlation matrix (right) are shown for the electron (top), Sigma (middle), and DNN (bottom) methods.

particularly large global correlation coefficients at low y . Furthermore, with higher x or lower y both classical methods have a successively increasing uncertainty, whereas the DNN shows continuously reduced uncertainties towards these kinematic regions (except the bin at lowest y).

Lastly, we study the properties of the unfolding when the migration matrix is obtained from a different Monte Carlo event generator as the (pseudo-)data sample. Such tests are of great importance in real data analyses, since a possible bias of the unfolding method due the selection of a certain physics generator is assessed. The differences of the two Monte Carlo models, when unfolded with the simulation from the other generator, is commonly considered as an uncertainty in the data analysis and is denoted as *(generator) model systematic uncertainty*. In order to reveal generator model systematic uncertainties at a statistically significant level, the size of the simulated event sample for these studies is 10^7 , which is 100 times larger than those used in figures 2 through 4. The technical closure plots when using the RAPGAP event generator are displayed in appendix B, and an excellent closure of the unfolded

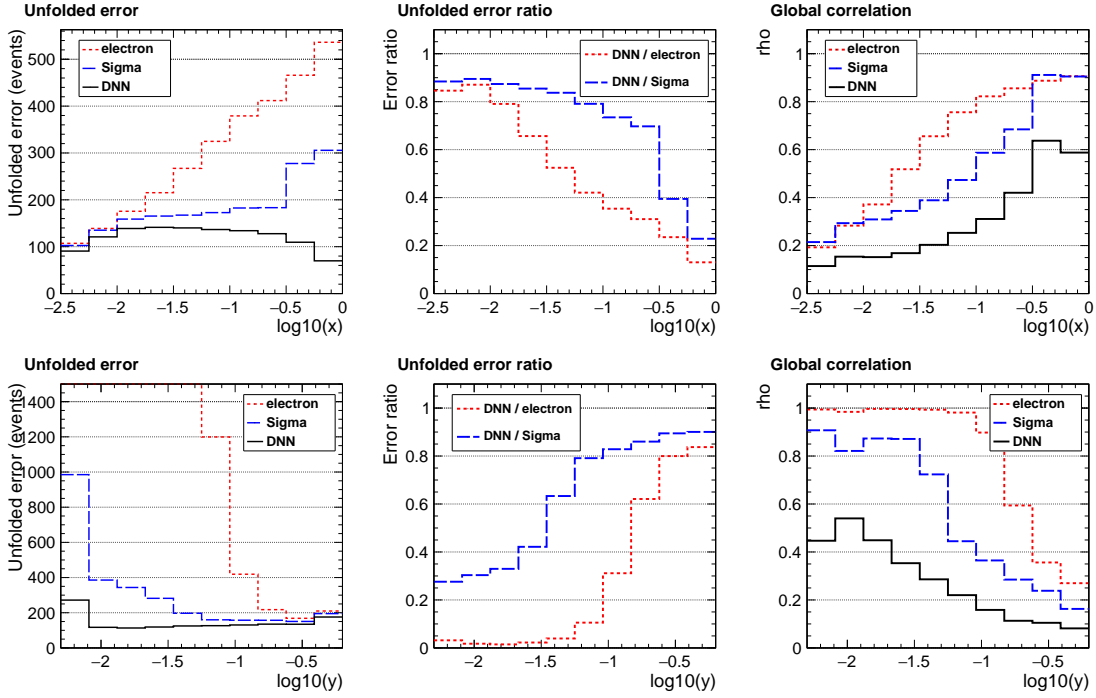


Figure 4. Comparison of unfolding results for $\log_{10}(x)$ (top) and $\log_{10}(y)$ (bottom) using the electron, Sigma, and DNN methods. Distributions for the statistical error on the unfolded result (left), the ratio of errors for the methods (middle), and the global correlation coefficient (right) are shown.

result is observed for all three reconstruction methods. Figure 5 shows the results of unfolding a RAPGAP event sample with a response matrix obtained from the DJANGO event generator, and vice versa, for the electron, Sigma and DNN reconstruction methods. We observe, that the electron method results show very large fluctuations in areas where the resolution is poor, but without any systematic trends. Also the results with the Sigma reconstruction method show large significant deviations. In contrast, the DNN method results in an insignificant model dependence in large parts of the x and y distributions. Only at highest x and lowest y some model dependence is observed, albeit smaller than those of the classical reconstruction methods. Altogether, the DNN reconstruction method results in reduced generator model systematic uncertainty than the classical reconstruction methods, and such would yield a significantly less biased physics result when used in analysis of real data.

3.2 Event shapes

As another example of our method, we explore the global event shape observable 1-jettiness [96, 97], τ_1^b , which is defined as

$$\tau_1^b = \frac{2}{Q^2} \sum_{i \in \text{HFS}} \min\{xP \cdot p_i, (q + xP) \cdot p_i\}, \quad (3.1)$$

where x is the DIS kinematic quantity from the last section, P is the proton beam four-vector, and p_i are the four-vectors of the HFS objects. The measurement of τ_1^b is clearly limited by the acceptance and resolution of the measurements of the HFS, while these measurements are important for the DIS

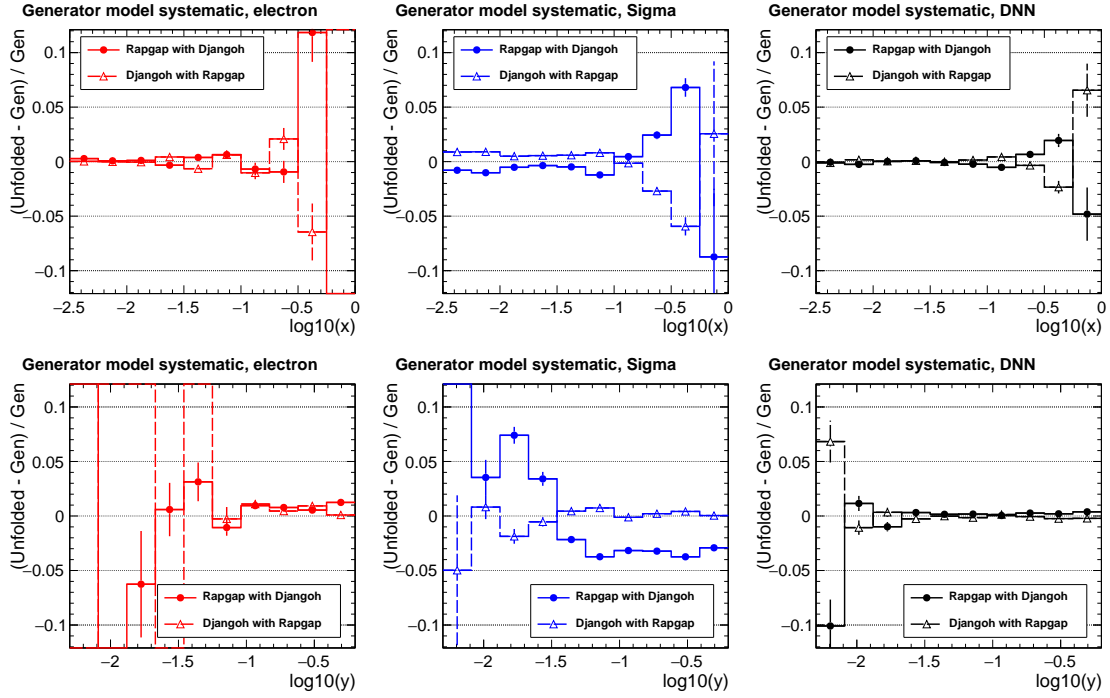


Figure 5. Generator model systematic uncertainties for the unfolding of x (top) and y (bottom) for the electron (left), Sigma (middle) and DNN (right) method.

kinematic observables only at lower y . The DNN-based reconstruction extends the case shown in the previous section, and the regression DNN takes into account further input quantities from HFS 4-vector in the current hemisphere (negative η in the Breit frame). The direct measurement of τ_1^b is not used as an input quantity to the DNN. The outputs of the regression DNN are Q^2 , x , y and τ_1^b .

Figure 6 shows the response matrices and the results of the unfolding for τ_1^b when using the electron, ISigma [94], and DNN methods. The reconstruction method impacts τ_1^b through Q^2 and x in its definition. The influence matrix is again displayed in appendix A. Note, differently than for DIS kinematic observables, we select to study the ISigma method instead of the Sigma method, since it was observed to have superior resolution for τ_1^b [98]. The migration matrices of the ISigma and DNN methods are significantly more diagonal than that of the electron method, whereas differences between ISigma and DNN methods are small. The closure of the three unfoldings is found to be unbiased, but the electron method results in large fluctuations due to its poor resolution for τ_1^b . The correlation matrix has sizable off-diagonal elements for the electron method, while ISigma and DNN are significantly more diagonal. Altogether, the results from the ISigma and DNN reconstruction appear to be quite similar. However, the global average correlation coefficient ρ_{avg} clearly improves for the DNN method, and it is 0.978, 0.869 and 0.842 for the electron, ISigma and DNN reconstruction method respectively. A more detailed study is presented in figure 7, where the ratio of the errors and the correlation coefficients are shown. The DNN method yields significantly smaller errors in all bins, by about 20%, and also reduced correlations coefficients.

While the results are not as dramatic as for inclusive DIS, the DNN still outperforms the traditional methods. In general one can expect that the observables that have a very inhomogeneous

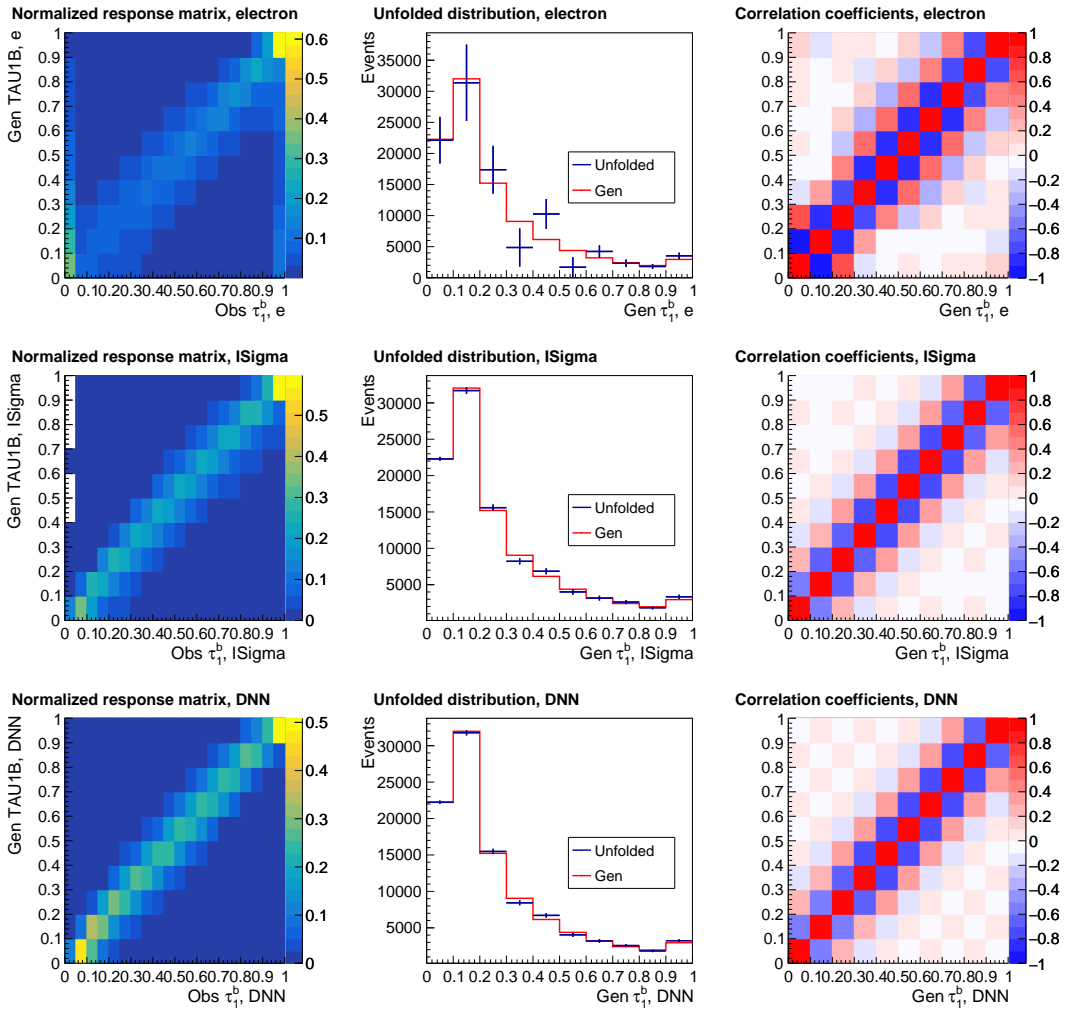


Figure 6. Examples of unfolding τ_1^b in DIS in one dimension for samples of 10^5 events. The response matrix (left), unfolded and gen distributions (middle), and unfolding correlation matrix (right) are shown for the electron (top), ISigma (middle), and DNN (bottom) methods.

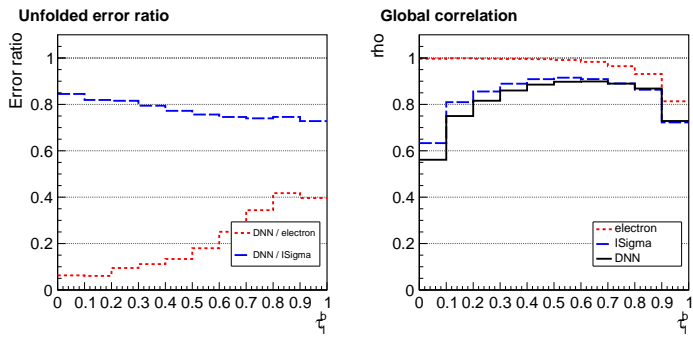


Figure 7. Comparison of unfolding results for τ_1^b using the electron, ISigma and DNN reconstruction method. Distributions for the ratio of the statistical errors of different distributions (left) and the correlation coefficient of the error (right) are shown.

detector response will benefit the most from a DNN-aided unfolding. This is why the benefits for inclusive DIS are more prevalent than for τ_1^b .

4 Summary and outlook

We have argued that the common practice of requiring analogous definition of observables at particle and detector level is not necessary. In fact, abandoning the usual constraint grants us freedom to define the observable at detector level using a large number of inputs to better account for detector effects and nuisance physical phenomena like QED radiation. We have proposed to define observables at detector level with deep learning and we have shown that this improves traditional unfolding methods by providing a more diagonal response matrix. Consequently, this results in smaller correlations between bins and smaller uncertainties of the unfolded results. Furthermore, the DNN reconstructed observables exhibit a smaller model dependence, which would reduce the related systematic uncertainties and provide less biased results. While our examples have focused on inclusive observables in DIS with full detector simulations of the H1 experiment at HERA, this approach can be generalized to any unfolding analysis. The benefits will be largest when there is a highly inhomogeneous detector response and there are many observables that govern this response which are not explicitly part of the unfolding.

Our approach deconstructs unfolding into two steps: the construction of observables and then the correction for detector effects. The first step can be viewed as an event-by-event correction for detector effects, but the second step is still necessary to mitigate prior dependence. Our approach is complementary to proposes for using machine learning to implement detector corrections [21–23, 57–64]. A combination of both methods would result in an even more precise and flexible analysis. The ultimate approach would be to unfold all observables simultaneously using all the available information at detector-level [21]. The approach discussed in this paper is much simpler and will likely yield immediate benefits while many of the statistical and computational challenges are being addressed for the full phase space methods.

Acknowledgments

We thank our colleagues from the H1 Collaboration for allowing us to use the simulated MC event samples. Thanks to DESY-IT and the MPI für Physik for providing some computing infrastructure and supporting the data preservation project of the HERA experiments. Parts of the work were supported by the ORIGINS Data Science Laboratory (ODSL). We thank L. Gouskos, S. Kluth, J. Thaler and Z. Zhang for helpful conversations and feedback on the manuscript. M.A. was supported through DOE Contract No. DE-AC05-06OR23177 under which JSA operates the Thomas Jefferson National Accelerator Facility, and by the University of California, Office of the President award number 00010100. B.N. was supported by the Department of Energy, Office of Science under contract number DE-AC02-05CH11231.

Code availability. The code in this work can be found in: <https://github.com/owen234/MLAssistedUnfolding>.

A The influence matrix

The properties of the *influence matrix* P [4, 99], which is also called *posterior response matrix*, were recently studied as an indicator of an unfolding problem in ref. [95]. The matrix P describes how the unfolding result is distorted with respect to the truth, and it can be regarded as the quadratic response matrix including the effect of the regularization. The influence matrix P for the electron, Sigma and DNN reconstruction methods when unfolding the observables x and y are displayed in figure 8, and for the observable τ_1^b in figure 9.

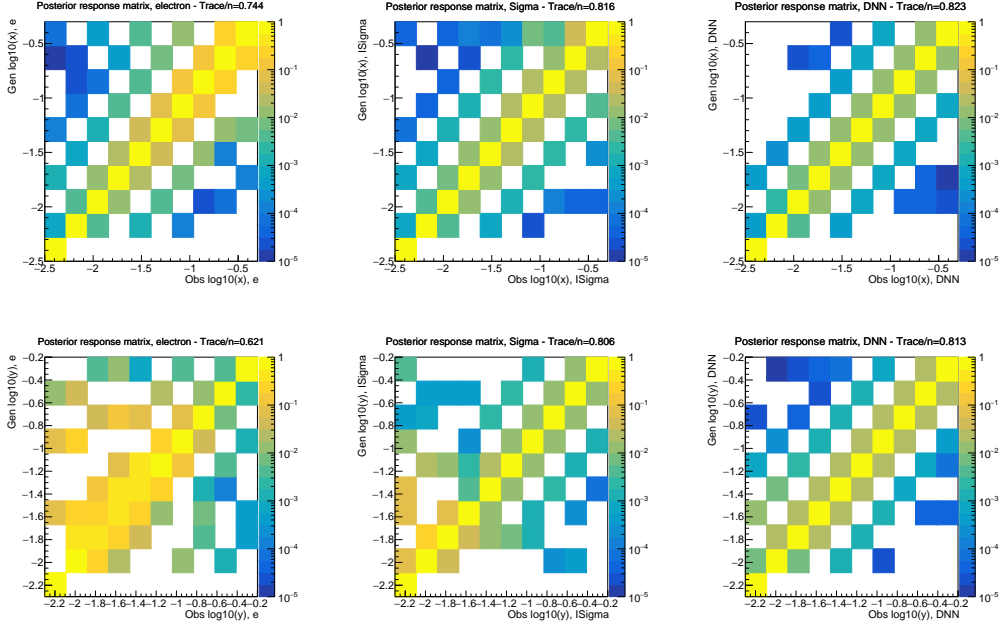


Figure 8. Influence matrices for the unfolding of x (top) and y (bottom) are shown for the electron (left), Sigma (middle) and DNN (right) reconstruction methods.

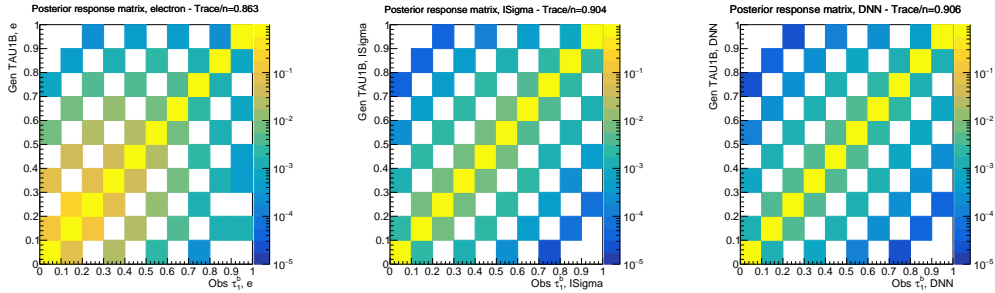


Figure 9. Influence matrices for the unfolding of τ_1^b for the electron (left), ISigma (middle) and DNN (right) reconstruction methods.

The influence matrices of the DNN-based reconstructed observables are more diagonal than those of the classical reconstruction methods. The ‘chessboard’ like structure, which is introduced by the

regularization, is more regular for the DNN observables than for the classical reconstruction methods (note the logarithmic scaling of the color coding). The traces of these matrices normalized to the number of particle-level bins, $\text{tr}(P)/n$, are displayed in table 2, and are seen to increase for the DNN method. The Pearson’s correlation coefficients of these matrices are found to be similar and close to one.

Table 2. Value of the traces of the influence matrices normalized to the number of bins ($n = 10$) for the unfolding of x , y and τ_1^b when using the electron, (I)Sigma or DNN reconstruction method.

Observable	Reconstruction method		
	electron	(I)Sigma	DNN
x	0.744	0.816	0.823
y	0.621	0.806	0.813
τ_1^b	0.863	0.904	0.906

B Closure tests for rapgap

Figure 10 shows the technical closure tests of the unfolding for the RAPGAP sample, where the matrix and pseudo-data are obtained from statistically independent RAPGAP samples. A sample of 10^7 events generated from the RAPGAP sample observed distribution is unfolded with the RAPGAP response matrix. The closure test is successful with the difference between the unfolding result and the generated distribution consistent with zero for all methods and all bins.

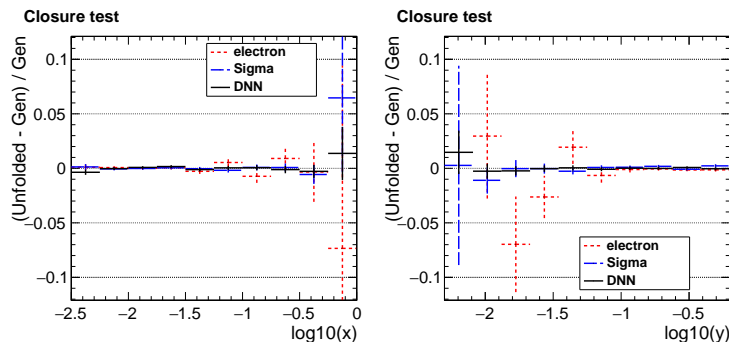


Figure 10. Closure tests for the unfolding of x (left) and y (right) when using the RAPGAP event sample.

References

- [1] O. Behnke, K. Kröniger, G. Schott and T. Schörner-Sadenius, eds., *Data analysis in high energy physics: a practical guide to statistical methods*, Wiley-VCH, Weinheim (2013), DOI: [10.1002/9783527653416](https://doi.org/10.1002/9783527653416)
- [2] G. Cowan, *A survey of unfolding methods for particle physics*, *Conf. Proc. C* **0203181** (2002) 248.
- [3] V. Blobel, *Unfolding Methods in Particle Physics*, in *Proceedings, PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*, CERN, Geneva, Switzerland 17–20 January 2011, CERN Yellow Reports: Conference Proceedings, (Geneva), CERN, 1, 2011, pp. 240–251.

- [4] V. Blobel, *Unfolding*, in *Data Analysis in High Energy Physics*, O. Behnke, K. Kröniger, G. Schott and T. Schörner-Sadenius, eds., Wiley-VCH, Weinheim (2013), pp. 187–225
- [5] L. Brenner, R. Balasubramanian, C. Burgard, W. Verkerke, G. Cowan, P. Verschuuren et al., *Comparison of unfolding methods using RooFitUnfold*, *Int. J. Mod. Phys. A* **35** (2020) 2050145 [[arXiv:1910.14654](#)].
- [6] CMS collaboration, *Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS*, *2011 JINST* **6** P11002 [[arXiv:1107.4277](#)].
- [7] CMS collaboration, *Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV*, *2017 JINST* **12** P02014 [[arXiv:1607.03663](#)].
- [8] ATLAS collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *Phys. Rev. D* **96** (2017) 072002 [[arXiv:1703.09665](#)].
- [9] ATLAS collaboration, *Jet energy measurement and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, *Eur. Phys. J. C* **75** (2015) 17 [[arXiv:1406.0076](#)].
- [10] ATLAS collaboration, *Determination of jet calibration and energy resolution in proton-proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, *Eur. Phys. J. C* **80** (2020) 1104 [[arXiv:1910.04482](#)].
- [11] ATLAS collaboration, *Measurement of jet-substructure observables in top quark, W boson and light jet production in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, *JHEP* **08** (2019) 033 [[arXiv:1903.02942](#)].
- [12] CMS collaboration, *Measurement of jet substructure observables in $t\bar{t}$ events from proton-proton collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. D* **98** (2018) 092014 [[arXiv:1808.07340](#)].
- [13] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N-subjettiness*, *JHEP* **03** (2011) 015 [[arXiv:1011.2268](#)].
- [14] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing N-subjettiness*, *JHEP* **02** (2012) 093 [[arXiv:1108.2701](#)].
- [15] H1 collaboration, *Measurement of multijet production in ep collisions at high Q^2 and determination of the strong coupling α_s* , *Eur. Phys. J. C* **75** (2015) 65 [[arXiv:1406.4709](#)].
- [16] H1 collaboration, *Measurement of Jet Production Cross Sections in Deep-inelastic ep Scattering at HERA*, *Eur. Phys. J. C* **77** (2017) 215 [Erratum *ibid.* **81** (2021) 739] [[arXiv:1611.03421](#)].
- [17] HERMES collaboration, *Azimuthal distributions of charged hadrons, pions, and kaons produced in deep-inelastic scattering off unpolarized protons and deuterons*, *Phys. Rev. D* **87** (2013) 012010 [[arXiv:1204.4161](#)].
- [18] H1 collaboration, *Prompt Photons in Photoproduction at HERA*, *Eur. Phys. J. C* **66** (2010) 17 [[arXiv:0910.5631](#)].
- [19] H1 collaboration, *Elastic and Proton-Dissociative Photoproduction of J/psi Mesons at HERA*, *Eur. Phys. J. C* **73** (2013) 2466 [[arXiv:1304.5162](#)].
- [20] H1 collaboration, *Measurement of Exclusive $\pi^+\pi^-$ and ρ^0 Meson Photoproduction at HERA*, *Eur. Phys. J. C* **80** (2020) 1189 [[arXiv:2005.14471](#)].
- [21] A. Andreassen, P.T. Komiske, E.M. Metodiev, B. Nachman and J. Thaler, *OmniFold: A Method to Simultaneously Unfold All Observables*, *Phys. Rev. Lett.* **124** (2020) 182001 [[arXiv:1911.09107](#)].

- [22] M. Bellagente, A. Butter, G. Kasieczka, T. Plehn, A. Rousselot, R. Winterhalder et al., *Invertible Networks or Partons to Detector and Back Again*, *SciPost Phys.* **9** (2020) 074 [[arXiv:2006.06685](#)].
- [23] M. Vandegar, M. Kagan, A. Wehenkel and G. Louppe, *Neural Empirical Bayes: Source Distribution Estimation and its Applications to Simulation-Based Inference*, [arXiv:2011.05836](#), in *Proceedings of The 24th International Conference on Artificial Intelligence and Statistics*, A. Banerjee and K. Fukumizu, eds., vol. 130 of *Proceedings of Machine Learning Research*, pp. 2107–2115, PMLR, 11, 2021 [<https://proceedings.mlr.press/v130/vandegar21a.html>].
- [24] M. Arratia et al., *Publishing unbinned differential cross section results*, *2022 JINST* **17** P01024 [[arXiv:2109.13243](#)].
- [25] H1 collaboration, *Measurement of Lepton-Jet Correlation in Deep-Inelastic Scattering with the H1 Detector Using Machine Learning for Unfolding*, *Phys. Rev. Lett.* **128** (2022) 132002 [[arXiv:2108.12376](#)].
- [26] CMS collaboration, *Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC*, *2021 JINST* **16** P05014 [[arXiv:2012.06888](#)].
- [27] J. Kieseler, G.C. Strong, F. Chiandotto, T. Dorigo and L. Layer, *Calorimetric Measurement of Multi-TeV Muons via Deep Regression*, *Eur. Phys. J. C* **82** (2022) 79 [[arXiv:2107.02119](#)].
- [28] D. Belayneh et al., *Calorimetry with deep learning: particle simulation and reconstruction for collider physics*, *Eur. Phys. J. C* **80** (2020) 688 [[arXiv:1912.06794](#)].
- [29] ATLAS collaboration, *Deep Learning for Pion Identification and Energy Calibration with the ATLAS Detector*, Tech. Rep., CERN, Geneva (2020), [ATL-PHYS-PUB-2020-018](#).
- [30] N. Akchurin, C. Cowden, J. Damgov, A. Hussain and S. Kunori, *On the use of neural networks for energy reconstruction in high-granularity calorimeters*, *2021 JINST* **16** P12036 [[arXiv:2107.10207](#)].
- [31] N. Akchurin, C. Cowden, J. Damgov, A. Hussain and S. Kunori, *Perspectives on the Calibration of CNN Energy Reconstruction in Highly Granular Calorimeters*, [arXiv:2108.10963](#).
- [32] L. Polson, L. Kurchaninov and M. Lefebvre, *Energy reconstruction in a liquid argon calorimeter cell using convolutional neural networks*, *2022 JINST* **17** P01002 [[arXiv:2109.05124](#)].
- [33] J. Pata, J. Duarte, J.-R. Vlimant, M. Pierini and M. Spiropulu, *MLPF: Efficient machine-learned particle-flow reconstruction using graph neural networks*, *Eur. Phys. J. C* **81** (2021) 381 [[arXiv:2101.08578](#)].
- [34] ATLAS collaboration, *Generalized Numerical Inversion: A Neural Network Approach to Jet Calibration*, Tech. Rep., CERN, Geneva (2018) [ATL-PHYS-PUB-2018-013](#).
- [35] ATLAS collaboration, *Simultaneous Jet Energy and Mass Calibrations with Neural Networks*, Tech. Rep., CERN, Geneva (2020) [ATL-PHYS-PUB-2020-001](#).
- [36] CMS collaboration, *A Deep Neural Network for Simultaneous Estimation of b Jet Energy and Resolution*, *Comput. Softw. Big Sci.* **4** (2020) 10 [[arXiv:1912.06046](#)].
- [37] R. Haake and C. Loizides, *Machine Learning based jet momentum reconstruction in heavy-ion collisions*, *Phys. Rev. C* **99** (2019) 064904 [[arXiv:1810.06324](#)].
- [38] ALICE collaboration, *Machine Learning based jet momentum reconstruction in Pb-Pb collisions measured with the ALICE detector*, *PoS EPS-HEP2019* (2020) 312 [[arXiv:1909.01639](#)].
- [39] P. Baldi, L. Blecher, A. Butter, J. Collado, J.N. Howard, F. Keilbach et al., *How to GAN Higher Jet Resolution*, [arXiv:2012.11944](#).

- [40] P.T. Komiske, E.M. Metodiev, B. Nachman and M.D. Schwartz, *Pileup Mitigation with Machine Learning (PUMML)*, *JHEP* **12** (2017) 051 [[arXiv:1707.08600](#)].
- [41] ATLAS collaboration, *Convolutional Neural Networks with Event Images for Pileup Mitigation with the ATLAS Detector*, Tech. Rep., CERN, Geneva (2019), [ATL-PHYS-PUB-2019-028](#).
- [42] B. Maier, S.M. Narayanan, G. de Castro, M. Goncharov, C. Paus and M. Schott, *Pile-up mitigation using attention*, *Mach. Learn. Sci. Tech.* **3** (2022) 025012 [[arXiv:2107.02779](#)].
- [43] G. Kasieczka, M. Luchmann, F. Otterpohl and T. Plehn, *Per-Object Systematics using Deep-Learned Calibration*, *SciPost Phys.* **9** (2020) 089 [[arXiv:2003.11099](#)].
- [44] J. Arjona Martínez, O. Cerri, M. Pierini, M. Spiropulu and J.-R. Vlimant, *Pileup mitigation at the Large Hadron Collider with graph neural networks*, *Eur. Phys. J. Plus* **134** (2019) 333 [[arXiv:1810.07988](#)].
- [45] M. Diefenthaler, A. Farhat, A. Verbytskyi and Y. Xu, *Deeply Learning Deep Inelastic Scattering Kinematics*, [arXiv:2108.11638](#).
- [46] M. Arratia, D. Britzger, O. Long and B. Nachman, *Reconstructing the kinematics of deep inelastic scattering with deep learning*, *Nucl. Instrum. Meth. A* **1025** (2022) 166164 [[arXiv:2110.05505](#)].
- [47] DUNE collaboration, *Deep-Learning-Based Kinematic Reconstruction for DUNE*, [arXiv:2012.06181](#).
- [48] EXO collaboration, *Deep Neural Networks for Energy and Position Reconstruction in EXO-200*, *2018 JINST* **13** P08023 [[arXiv:1804.09641](#)].
- [49] P. Baldi, J. Bian, L. Hertel and L. Li, *Improved Energy Reconstruction in NOvA with Regression Convolutional Neural Networks*, *Phys. Rev. D* **99** (2019) 012011 [[arXiv:1811.04557](#)].
- [50] R. Abbasi et al., *A Convolutional Neural Network based Cascade Reconstruction for the IceCube Neutrino Observatory*, *2021 JINST* **16** P07041 [[arXiv:2101.11589](#)].
- [51] ICECUBE collaboration, *Cosmic ray spectrum from 250 TeV to 10 PeV using IceTop*, *Phys. Rev. D* **102** (2020) 122001 [[arXiv:2006.05215](#)].
- [52] K. Carloni, N.W. Kamp, A. Schneider and J.M. Conrad, *Convolutional neural networks for shower energy prediction in liquid argon time projection chambers*, *2022 JINST* **17** P02022 [[arXiv:2110.10766](#)].
- [53] M. Feickert and B. Nachman, *A Living Review of Machine Learning for Particle Physics*, [arXiv:2102.02770](#).
- [54] L. Lindemann and G. Zech, *Unfolding by weighting Monte Carlo events*, *Nucl. Instrum. Meth. A* **354** (1995) 516.
- [55] G. Zech and B. Aslan, *Binning-Free Unfolding Based on Monte Carlo Migration*, *eConf C* **030908** (2003) TUGT001.
- [56] H.P. Dembinski and M. Roth, *An algorithm for automatic unfolding of one-dimensional data distributions*, *Nucl. Instrum. Meth. A* **729** (2013) 410.
- [57] N.D. Gagunashvili, *Machine learning approach to inverse problem and unfolding procedure*, [arXiv:1004.2006](#).
- [58] A. Glazov, *Machine learning as an instrument for data unfolding*, [arXiv:1712.01814](#).
- [59] K. Datta, D. Kar and D. Roy, *Unfolding with Generative Adversarial Networks*, [arXiv:1806.00433](#).
- [60] M. Bellagente, A. Butter, G. Kasieczka, T. Plehn and R. Winterhalder, *How to GAN away Detector Effects*, *SciPost Phys.* **8** (2020) 070 [[arXiv:1912.00477](#)].

- [61] A. Andreassen, P.T. Komiske, E.M. Metodiev, B. Nachman, A. Suresh and J. Thaler, *Scaffolding Simulations with Deep Learning for High-dimensional Deconvolution*, in *Proceedings of the 9th International Conference on Learning Representations*, 5, 2021 [[arXiv:2105.04448](#)].
- [62] M. Bunse, N. Piatkowski, K. Morik, T. Ruhe and W. Rhode, *Unification of deconvolution algorithms for cherenkov astronomy*, in *Proceedings of the IEEE 5th International Conference on Data Science and Advanced Analytics*, Turin, Italy, 1–3 October 2018, pp. 21–30.
- [63] T. Ruhe, T. Voigt, M. Wornowizki, M. Börner, W. Rhode and K. Morik, *Mining for spectra — the Dortmund spectrum estimation algorithm*, in *Astronomical Data Analysis Software and Systems XXVI ASP Conference Series* **521** (2019) 394.
- [64] J.N. Howard, S. Mandt, D. Whiteson and Y. Yang, *Foundations of a Fast, Data-Driven, Machine-Learned Simulator*, [arXiv:2101.08944](#).
- [65] LHC HIGGS CROSS SECTION WORKING GROUP collaboration, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, [arXiv:1610.07922](#).
- [66] N. Berger et al., *Simplified Template Cross Sections - Stage 1.1*, [arXiv:1906.02754](#).
- [67] ATLAS collaboration, *Measurement of the properties of Higgs boson production at $\sqrt{s} = 13$ TeV in the $H \rightarrow \gamma\gamma$ channel using 139fb^{-1} of pp collision data with the ATLAS experiment*, Tech. Rep., CERN, Geneva (2020), [ATLAS-CONF-2020-026](#).
- [68] CMS collaboration, *Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{s} = 13$ TeV*, *JHEP* **07** (2021) 027 [[arXiv:2103.06956](#)].
- [69] S. Schmitt, *TUnfold: an algorithm for correcting migration effects in high energy physics*, *2012 JINST* **7** T10003 [[arXiv:1205.6201](#)].
- [70] L.B. Lucy, *An iterative technique for the rectification of observed distributions*, *Astron. J.* **79** (1974) 745.
- [71] W.H. Richardson, *Bayesian-based iterative method of image restoration*, *J. Opt. Soc. Am.* **62** (1972) 55
- [72] G. D’Agostini, *A Multidimensional unfolding method based on Bayes’ theorem*, *Nucl. Instrum. Meth. A* **362** (1995) 487.
- [73] A. Höcker and V. Kartvelishvili, *SVD approach to data unfolding*, *Nucl. Instrum. Meth. A* **372** (1996) 469 [[hep-ph/9509307](#)].
- [74] V. Blobel, *Unfolding Methods in High-energy Physics Experiments*, in *Proceedings of CERN School of Computing*, Aiguablava, Spain, 9–22 September 1984.
- [75] A.N. Tikhonov, *Solution of Incorrectly Formulated Problems and the Regularization Method*, *Soviet Math. Dokl.* **4** (1963) 1035
- [76] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, *Nucl. Instrum. Meth. A* **389** (1997) 81.
- [77] F. Chollet, *Keras*, <https://github.com/fchollet/keras>, 2015
- [78] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean et al., *Tensorflow: A system for large-scale machine learning*, in *Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation*, Savannah, GA, U.S.A., 2–4 November 2016, vol. 16, pp. 265–283.
- [79] D.P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, 12, 2014 [[arXiv:1412.6980](#)].
- [80] P.J. Huber, *Robust Estimation of a Location Parameter*, *Ann. Math. Stat.* **35** (1964) 73.
- [81] H1 collaboration, *The Tracking, calorimeter and muon detectors of the H1 experiment at HERA*, *Nucl. Instrum. Meth. A* **386** (1997) 348.

- [82] H1 collaboration, *The H1 detector at HERA*, *Nucl. Instrum. Meth. A* **386** (1997) 310.
- [83] H. Jung, *Hard diffractive scattering in high-energy $e p$ collisions and the Monte Carlo generator RAPGAP*, *Comput. Phys. Commun.* **86** (1995) 147.
- [84] K. Charchula, G.A. Schuler and H. Spiesberger, *Combined QED and QCD radiative effects in deep inelastic lepton-proton scattering: The Monte Carlo generator DJANGO6*, *Comput. Phys. Commun.* **81** (1994) 381.
- [85] H. Spiesberger et al., *Radiative corrections at HERA*, in *Proceedings of Workshop on Physics at HERA*, DESY, Hamburg, Germany, 29–30 October 1991, pp. 798–839.
- [86] A. Kwiatkowski, H. Spiesberger and H.J. Mohring, *Characteristics of radiative events in deep inelastic $e p$ scattering at HERA*, *Z. Phys. C* **50** (1991) 165.
- [87] A. Kwiatkowski, H. Spiesberger and H.J. Mohring, *Heracles: An Event Generator for ep Interactions at HERA Energies Including Radiative Processes: Version 1.0*, *Comput. Phys. Commun.* **69** (1992) 155.
- [88] M. Peez, *Search for deviations from the standard model in high transverse energy processes at the electron proton collider HERA*, Ph.D. thesis, University of Lyon, France (2003).
- [89] S. Hellwig, *Untersuchung der $D^*-\pi_{\text{slow}}$ Double Tagging Methode in Charmanalysen*, Diploma thesis, University of Hamburg, Germany (2004).
- [90] B. Portheault, *First measurement of charged and neutral current cross sections with the polarized positron beam at HERA II and QCD-electroweak analyses*, Ph.D. thesis, University Paris XI, France (2005).
- [91] H1 collaboration, *Inclusive Deep Inelastic Scattering at High Q^2 with Longitudinally Polarised Lepton Beams at HERA*, *JHEP* **09** (2012) 061 [[arXiv:1206.7007](https://arxiv.org/abs/1206.7007)].
- [92] H1 collaboration, *Measurement of Lepton-Jet Correlation in Deep-Inelastic Scattering with the H1 Detector Using Machine Learning for Unfolding*, *Phys. Rev. Lett.* **128** (2022) 132002 [[arXiv:2108.12376](https://arxiv.org/abs/2108.12376)].
- [93] H1 collaboration, *Preservation through modernisation: The software of the H1 experiment at HERA*, *EPJ Web Conf.* **251** (2021) 03004 [[arXiv:2106.11058](https://arxiv.org/abs/2106.11058)].
- [94] U. Bassler and G. Bernardi, *On the kinematic reconstruction of deep inelastic scattering at HERA: The Sigma method*, *Nucl. Instrum. Meth. A* **361** (1995) 197 [[hep-ex/9412004](https://arxiv.org/abs/hep-ex/9412004)].
- [95] M. Schmelling, *Regularised unfolding with a discrete-valued penalty function*, [arXiv:2203.09579](https://arxiv.org/abs/2203.09579).
- [96] V. Antonelli, M. Dasgupta and G.P. Salam, *Resummation of thrust distributions in DIS*, *JHEP* **02** (2000) 001 [[hep-ph/9912488](https://arxiv.org/abs/hep-ph/9912488)].
- [97] I.W. Stewart, F.J. Tackmann and W.J. Waalewijn, *N -Jettiness: An Inclusive Event Shape to Veto Jets*, *Phys. Rev. Lett.* **105** (2010) 092002 [[arXiv:1004.2489](https://arxiv.org/abs/1004.2489)].
- [98] J. Hessler, *Measurement of the 1-jettiness Event Shape Observable in Deep-inelastic Electron-Proton Scattering*, Ph.D. thesis, Technical University, Munich, Germany (2021).
- [99] G. Wahba, ed., *Spline Models for Observational Data*, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM (1990) [[DOI: 10.1137/1.9781611970128](https://doi.org/10.1137/1.9781611970128)].