

Unbinned measurements for new physics

Benjamin Nachman

Lawrence Berkeley National Laboratory

bpnachman.com

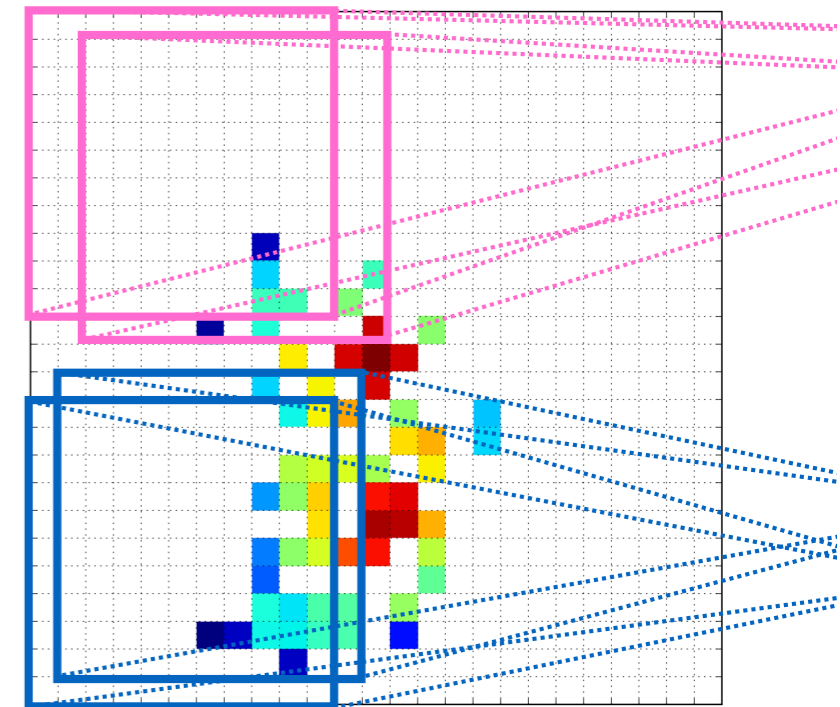
bpnachman@lbl.gov



@bpnachman



bnachman



LHC Reinterpretation
Forum (RIF)
Aug. 29, 2023

Overview of unbinned measurements



2	Motivation	2109.13243	3
2.1	Inference-Aware Binning		3
2.2	Derivative Measurements		4
2.3	Extension to Higher Dimensions		4
3	Overview of Unbinned Methods		5
3.1	Experimental Measurements		5
3.1.1	Density-Based Models		5
3.1.2	Classifier-Based Models		6
3.2	Theoretical Calculations		6
4	Statistics of Unbinned Spectra		7
4.1	Acceptance Effects		7
4.2	Background Subtraction		7
4.3	Local Statistical Uncertainty		7
5	Systematic Uncertainties		9
6	Proposed Format and Storage		9
6.1	Structure		9
6.2	Uncertainties		11
6.3	Storage		11
6.4	Common Tools		11
7	Example		11
8	Conclusions and Outlook		16

Motivation - why unbinned?



2	Motivation	2109.13243	3
2.1	Inference-Aware Binning		3
2.2	Derivative Measurements		4
2.3	Extension to Higher Dimensions		4
3	Overview of Unbinned Methods		5
3.1	Experimental Measurements		5
3.1.1	Decision-Based Models		5
3.1.2	Classification-Based Models		6
3.2	Theoretical Considerations		6
4	Statistics of Unbinned Methods		7
4.1	Acceptance		7
4.2	Background		7
4.3	Local Statistics		7
5	Systematic Uncertainties		7
6	Proposed Framework		7
6.1	Structure		7
6.2	Uncertainty		7
6.3	Storage		7
6.4	Common		7
7	Example		7
8	Conclusions and Outlook		7

Inference-Aware Binning
Derivative Measurements
Extension to Higher Dimensions

Current paradigm

Chose binning based on heuristic (e.g. purity > 50%)



Do the measurement and publish results



Perform many down-stream analyses using this binning

Motivation I: Inference-aware binning

5

Current paradigm

Chose binning based on heuristic (e.g. purity > 50%)



Do the measurement and publish results



Perform many down-stream analyses using this binning

Unbinned

Do the measurement and publish results



For a given analysis, pick the optimal choice of bins



Repeat for many analyses

Motivation I: Inference-aware binning

6

Current paradigm

Chose binning based on heuristic (e.g. purity > 50%)



Do the measurement and publish results



Perform many down-stream analyses using this binning

Unbinned

Do the measurement and publish results



For a given analysis, pick the optimal choice of bins



Repeat for many analyses

Related: makes comparison with other experiments much easier
... do not need to coordinate on binning ahead of time!

Motivation II: Derivative Measurements



Say you measure the observables x ,
but you later want to measure $f(x)$.

If the original measurement is binned,
then you can only make a crude
approximation of f using bin centers.

Motivation II: Derivative Measurements



Say you measure the observables x ,
but you later want to measure $f(x)$.

If the original measurement is binned,
then you can only make a crude
approximation of f using bin centers.

**Optimal f (and the binning) may depend with
time as more data are available for global
fits - this is enabled by unbinned data!**

Motivation III: Extension to higher-dim.

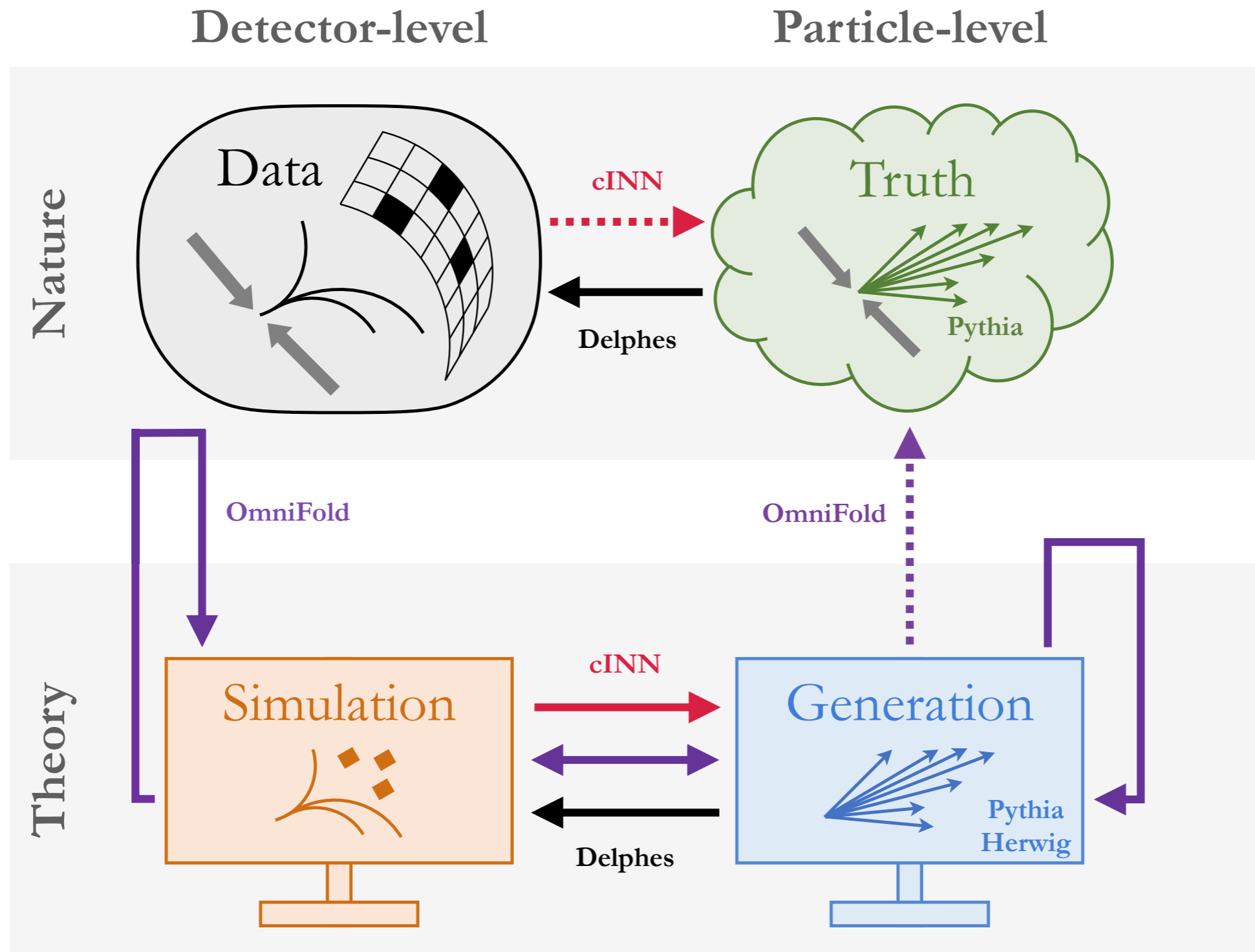


Many of the proposals for unbinned measurements make use of machine learning and readily extend to many (and even variable) dimensions.

While not a direct benefit of unbinned results, this would be a clear game changer for how we do measurements !!

With enough (internal) information, can build correlation matrices between measurements post-hoc, but this comes for free if originally done multidimensional

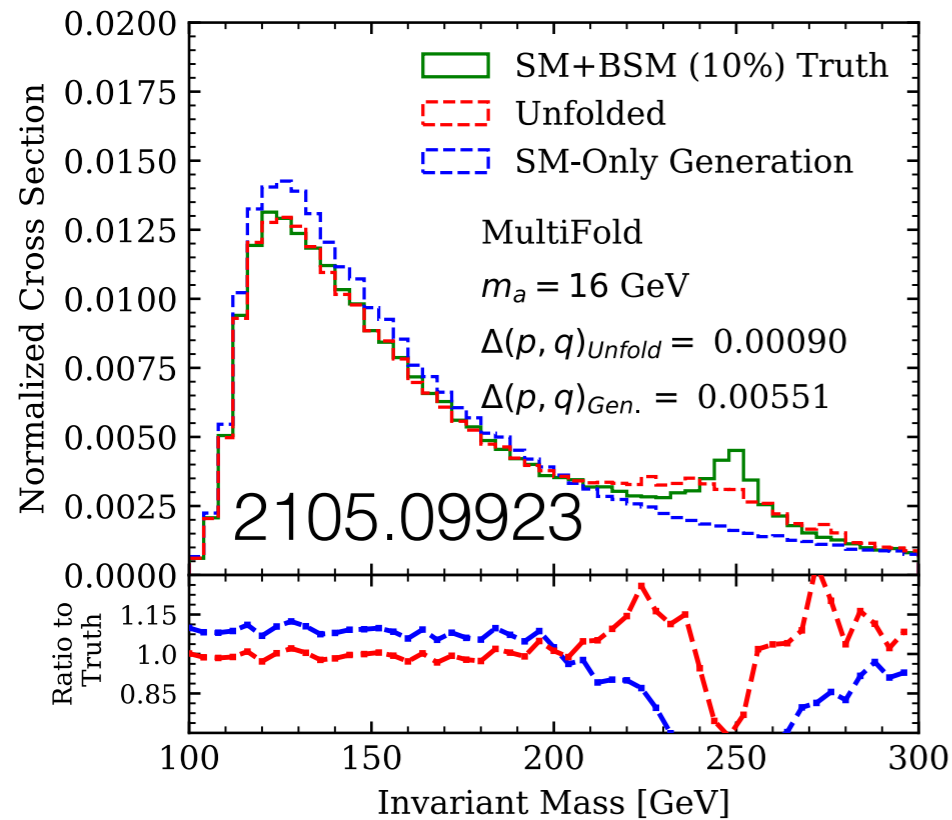
Unbinned Methods



There are other examples, but these ones (and their extensions) are particularly well studied.

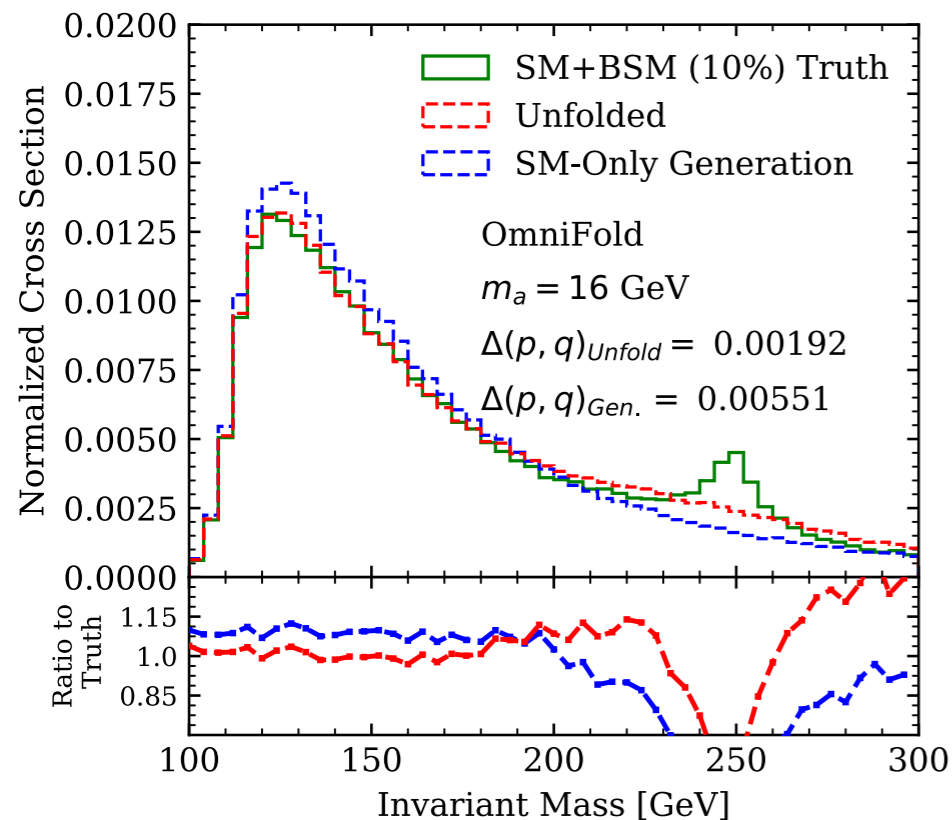
Can these approaches preserve BSM?

11



MultiFold: 10d cross section,
OmniFold: all particles.

BSM: $H \rightarrow Z\alpha$

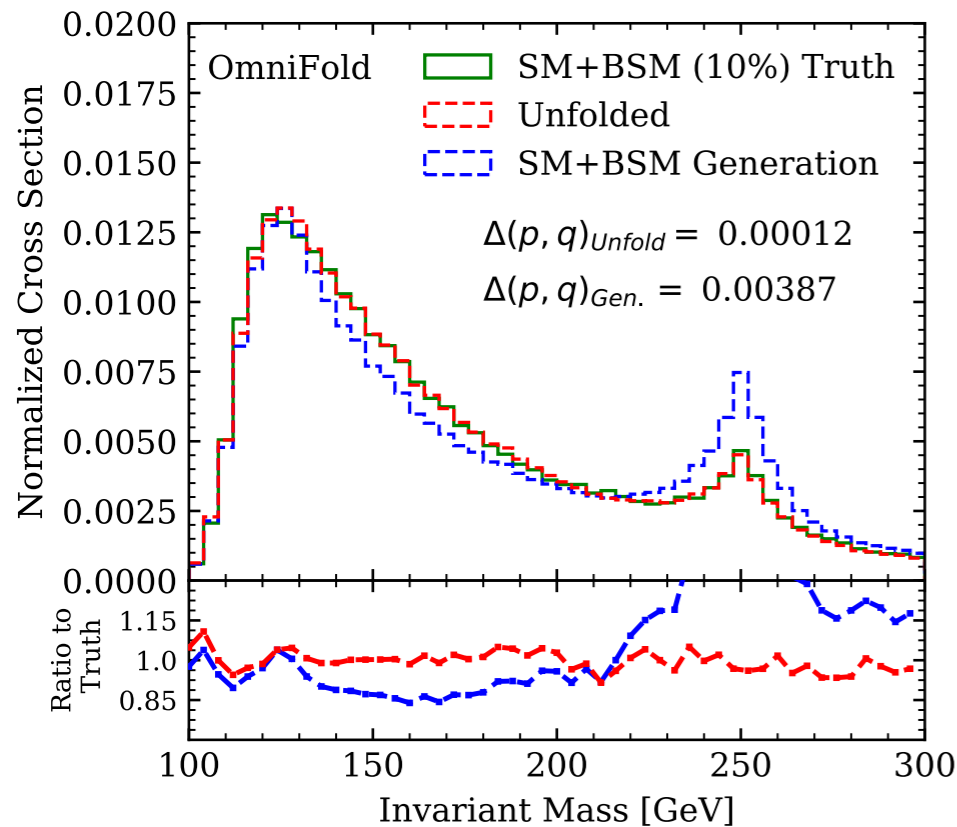


Answer: sort of ... preserves
anomaly when big enough ($> 1\%$).
Precision continues to improve!

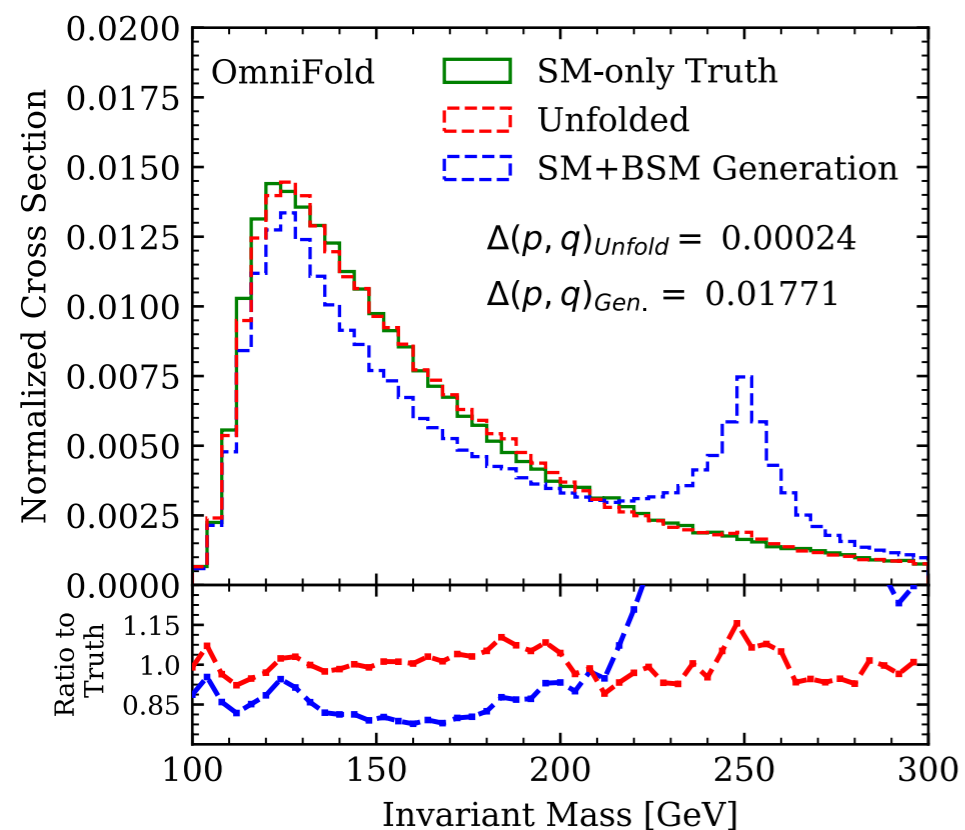
see also 1912.00477

Can these approaches preserve BSM?

12



Step towards improving: add BSM to prior during unfolding.



Has little effect when no signal (bottom) but makes it much easier to preserve signal when present (top)

Challenge: how to publish/recast?

13

This breaks HEPData!

2109.13243 proposed a solution, but it has not been applied yet, despite the fact that OmniFold has been used in a few places now (H1, LHCb, STAR)

...stay tuned!

Conclusions and Outlook



New methods for unbinned unfolding are here! We should be ready to use them also for BSM!

Measurement of lepton-jet correlation in deep-inelastic scattering with the H1 detector using machine learning for unfolding

V. Andreev,²³ M. Arratia,³⁵ A. Bagdasaryan,⁴⁶ A. Baty,¹⁶ K. Begzsuren,³⁹ A. Belousov,²³, V. Boudry,³¹ G. Brandt,¹³ D. Britzger,²⁶ A. Buniatyán,⁶ L. Bystritskaya,²² A.J. Campbell,¹⁴ K.B. Chen,³¹ J.G. Contreras,⁴¹ J. Cvach²⁷, J.B. Dainton¹⁹, K. Daum³⁰, A. Deshpande^{33,36}, C. I. Eckerlin¹⁴, S. Egli³⁷, E. Elsen¹⁴, L. Favart⁴, A. Fedotov⁴⁴, J. Feltesse⁴², M. Fleischer¹⁹, J. Gayler¹⁴, L. Goerlich¹⁷, N. Gogitidze²³, M. Grab¹⁹, T. Greenshaw¹⁹, G. Grindhammer²⁶, D. Haidt¹⁴, R.C.W. Henderson¹⁸, J. Hessler²⁶, D. Hoffmann²¹, R. Horisberger⁴³, T. Hreus⁵⁰, F. Huber¹⁵, P.M. Jacobs⁵, M. Jacquet²⁹, T. Janssen¹⁴, H. Jung¹⁴, M. Kapichine¹⁰, J. Katzy¹⁴, C. Kiesling²⁶, M. Klein¹⁹, C. Kleinwort¹⁴, H.T. Klest³⁸, P. Kostka¹⁹, J. Kretschmar¹⁹, D. Krücker¹⁴, K. Krüger¹⁴, M.P.J. Landon²⁰, W. Lange⁴⁸, P. I. S.H. Lee³, S. Levonian¹⁴, W. Li¹⁶, J. Lin¹⁶, K. Lipka¹⁴, B. List¹⁴, J. List¹⁴, B. Lobodzinski²⁶, E. H.-U. Martyn¹, S.J. Maxfield¹⁹, A. Mehta¹⁹, A.B. Meyer¹⁴, J. Meyer¹⁴, S. Mikocki¹⁷, M.M. Mondal³³, K. Müller⁵⁰, B. Nachman⁵, Th. Naumann⁴⁸, P.R. Newman⁶, C. Niebuhr¹⁴, G. Nowak¹⁷, J.E. D. Ozerov⁴³, S. Park³⁸, C. Pascaud²⁹, G.D. Patel¹⁹, E. Perez¹¹, A. Petrukhin⁴², I. Picuric³², R. Polifka³⁴, S. Preins³⁵, V. Radescu³⁰, N. Raicevic³², T. Ravdandorj³⁹, P. Reimer³³, E. Rizvi²⁰, P. R. Roosen⁴, A. Rostovtsev²⁵, M. Rotaru⁷, D.P.C. Sankey⁹, M. Sauter¹⁵, E. Sauvan^{21,2}, S. Sch. B.A. Schmookler³⁸, L. Schoeffel¹², A. Schöning¹⁵, F. Sefkow¹⁴, S. Shushkevich²⁴, Y. Soloviev²³, D. South¹⁴, V. Spaskov¹⁰, A. Specka³¹, M. Steder¹⁴, B. Stella³⁶, U. Straumann⁵⁰, C. Sun³⁷, T. P.D. Thompson⁶, D. Traynor²⁰, B. Tsepeeldorj^{39,40}, Z. Tu⁴¹, A. Valkárová³⁴, C. Vallée²¹, P. Van D. Wegener⁹, E. Wünsch¹⁴, J. Žáček³⁴, J. Zhang³⁷, Z. Zhang²⁹, R. Zlebčik³⁴, H. Zohrabyan⁴⁶ and (The H1 Collaboration)

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)



CERN-EP-2022-161
LHCb-PAPER-2022-013
August 25, 2022

Multidifferential study of identified charged hadron distributions in Z-tagged jets in proton-proton collisions at $\sqrt{s} = 13$ TeV

Abstract

Jet fragmentation functions are measured for the first time in proton-proton collisions for charged pions, kaons, and protons within jets recoiling against a Z boson. The charged-hadron distributions are studied longitudinally and transversely to the jet direction for jets with transverse momentum $20 < p_T < 100$ GeV and in the pseudorapidity range $2.5 < \eta < 4$. The data sample was collected with the LHCb experiment at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 1.64 fb^{-1} . Triple differential distributions as a function of the hadron longitudinal momentum fraction, hadron transverse momentum, and jet transverse momentum are also measured for the first time. This helps constrain transverse-momentum-dependent fragmentation functions. Differences in the shapes and magnitudes of the measured distributions for the different hadron species provide insights into the hadronization process for jets predominantly initiated by light quarks.

Submitted to Phys. Rev. D Letter

© 2022 CERN for the benefit of the LHCb collaboration. CC BY 4.0 licence.

Unbinned Deep Learning Jet Substructure Measurement in High Q^2 ep collisions at HERA

Andreev⁴⁴, M. Arratia²⁹, A. Bagdasaryan⁴⁰, A. Baty¹⁶, K. Begzsuren³⁴, A. Bolz¹⁴, V. Boudry²⁵, G. Brandt¹³, Britzger²², A. Buniatyán⁷, L. Bystritskaya⁴⁴, A.J. Campbell¹⁴, K.B. Cantun Avila⁴¹, K. C. Chen³¹, J.G. Contreras⁴¹, J. Cvach²⁷, J.B. Dainton¹⁹, K. Daum³⁰, A. Deshpande^{33,36}, C. I. Eckerlin¹⁴, S. Egli³⁷, E. Elsen¹⁴, L. Favart⁴, A. Fedotov⁴⁴, J. Feltesse⁴², M. Fleischer¹⁹, J. Gayler¹⁴, L. Goerlich¹⁷, N. Gogitidze¹⁴, M. Gouzevitch²⁴, C. Grab⁴², T. Greenshaw¹⁹, Haidt¹⁴, R.C.W. Henderson¹⁸, J. Hessler²², J. Hladky²⁷, D. Hoffmann²¹, R. Horisberger²⁷, P.M. Jacobs⁵, M. Jacquet²⁴, T. Janssen⁴, A.W. Jung²⁸, J. Katzy¹⁴, C. Kiesling²⁶, M. Klein¹⁹, K. Klest³³, R. Kogler¹⁴, P. Kostka¹⁹, J. Kretschmar¹⁹, D. Krücker¹⁴, K. Krüger¹⁴, M.P.J. Laycock²⁶, S.H. Lee³, S. Levonian¹⁴, W. Li¹⁶, J. Lin¹⁶, K. Lipka¹⁴, B. List¹⁴, J. List¹⁴, Long²⁹, E. Malinovsky¹⁴, H.-U. Martyn¹, S.J. Maxfield¹⁹, A. Mehta¹⁹, A.B. Meyer¹⁴, J. V.M. Mikuni², M.M. Mondal³³, K. Müller⁴⁹, B. Nachman⁵, Th. Naumann¹⁴, P.R. Newn G. Nowak¹⁷, J.E. Olsson¹⁴, D. Ozerov⁴⁴, S. Park³³, C. Pascaud²⁹, G.D. Patel¹⁹, E. Perez I. Picuric²⁶, D. Pitzl¹⁴, R. Polifka²⁸, S. Preins²⁹, V. Radescu¹⁵, N. Raicevic²⁶, T. Ravdandorj²⁰, P. Robmann⁴³, R. Roosen⁴, A. Rostovtsev⁴⁴, M. Rotaru², D.P.C. Sankey⁹, M. S. S. Schmitt¹⁴, B.A. Schmookler³³, G. Schnell¹², L. Schoeffel¹², A. Schöning¹⁵, F. Sefkow² oloviev¹⁴, P. Sopic¹⁷, D. South¹⁴, A. Specka²⁵, M. Steder¹⁴, B. Stella³⁰, U. Straumann⁵⁰ P.D. Thompson⁷, F. Torales Acosta⁴, D. Traynor²⁰, B. Tsepeeldorj^{34,35}, Z. Tu³⁶, G. Tusti C. Vallée²¹, P. Van Mechelen⁴, D. Wegener¹⁰, E. Wünsch¹⁴, J. Žáček²⁸, J. Zhang³¹, Z. Zhang H. Zohrabyan⁴⁰, F. Zomer²⁴

July 19, 2023

Measurement of CollinearDrop jet mass and its correlation with SoftDrop groomed jet substructure observables in $\sqrt{s} = 200$ GeV pp collisions by STAR

YOUQI SONG (WRIGHT LABORATORY, YALE UNIVERSITY)

on behalf of the STAR Collaboration

Jet substructure variables aim to reveal details of the parton fragmentation and hadronization processes that create a jet. By removing collinear radiation while maintaining the soft radiation components, one can construct CollinearDrop jet observables, which have enhanced sensitivity to the soft phase space within jets. We present a CollinearDrop jet measurement, corrected for detector effects with a machine learning method, Multi-Fold, and its correlation with groomed jet observables, in pp collisions at $\sqrt{s} = 200$ GeV at STAR. We demonstrate that the population of jets with a large non-perturbative contribution can be significantly enhanced by selecting on higher CollinearDrop jet mass fractions. In addition, we observe an anti-correlation between the amount of grooming and the angular scale of the first hard splitting of the jet.

PRESENTED AT

DIS2023: XXX International Workshop on Deep-Inelastic Scattering and Related Subjects, Michigan State University, USA, 27-31 March 2023

arXiv:2108.12376v2 [hep-ex] 1 Apr 2022

arXiv:2208.11691v1 [hep-ex] 24 Aug 2022

arXiv:2307.07718v2 [nucl-ex] 18 Jul 2023

+CMS open data study