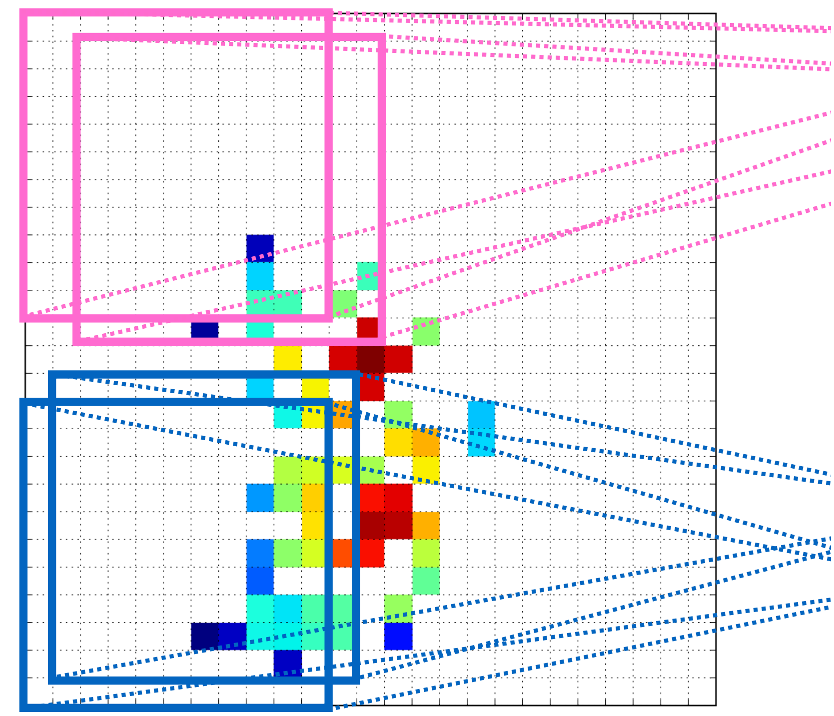
Unbinned measurements for new physics

Benjamin Nachman

Lawrence Berkeley National Laboratory

bpnachman.com
bpnachman@lbl.gov
@bpnachman bnachman



BERKELEY LAB

BIDS
BERKELEY INSTITUTE FOR DATA SCIENCE

LHC Reinterpretation Forum (RIF)
Aug. 29, 2023

Overview of unbinned measurements

2109.13243

Motivation - why unbinned?

2109.13243

Inference-Aware Binning

Derivative Measurements

Extension to Higher Dimensions

Motivation I: Inference-aware binning

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

Current paradigm

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

↓

<u>Do the measurement and publish results</u>

↓

Perform many down-stream analyses using this binning

Unbinned

<u>Do the measurement and publish results</u>

↓

For a given analysis, pick the optimal choice of bins

↓

Repeat for many analyses

Motivation I: Inference-aware binning

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!

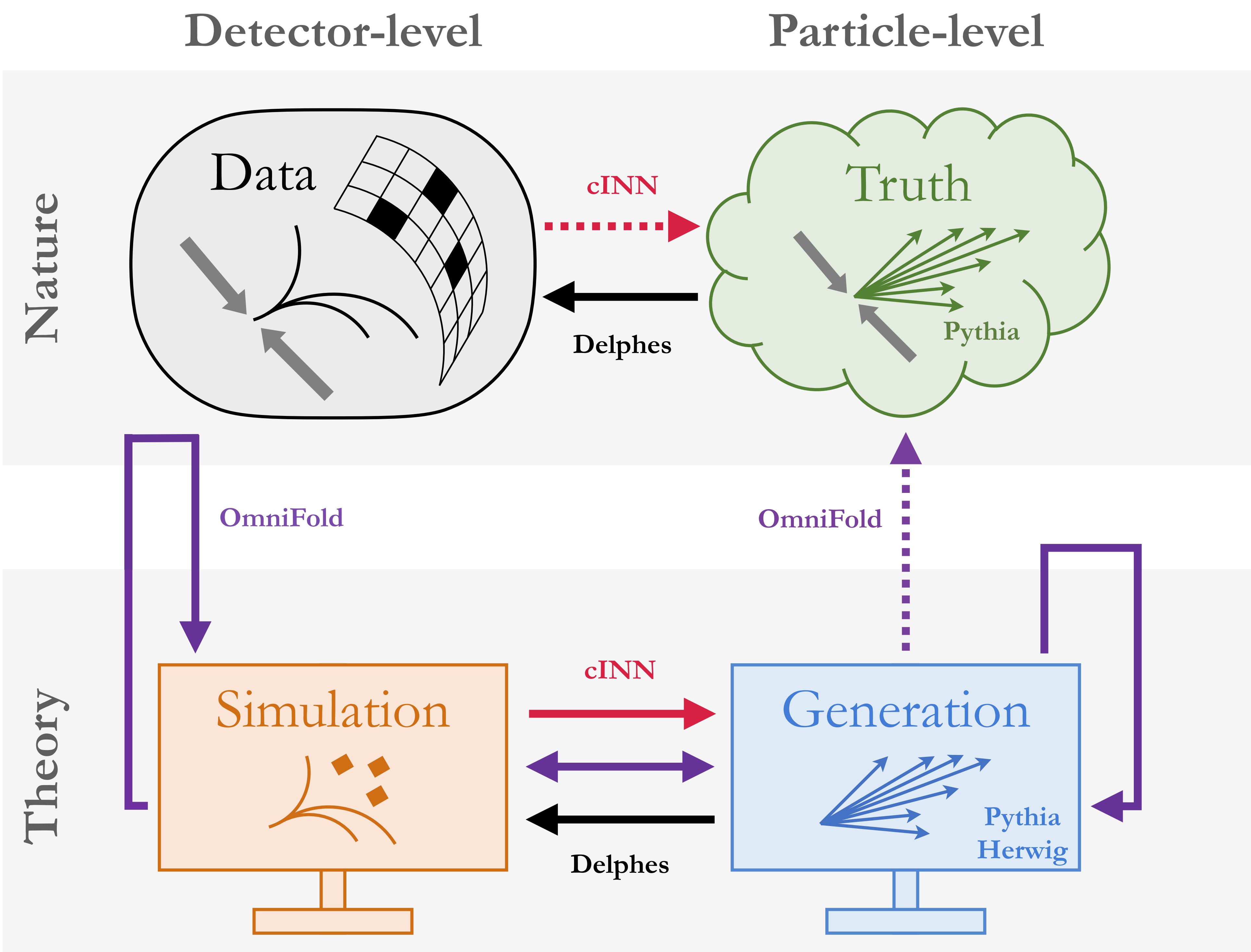
Motivation III: Extension to higher-dim.

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

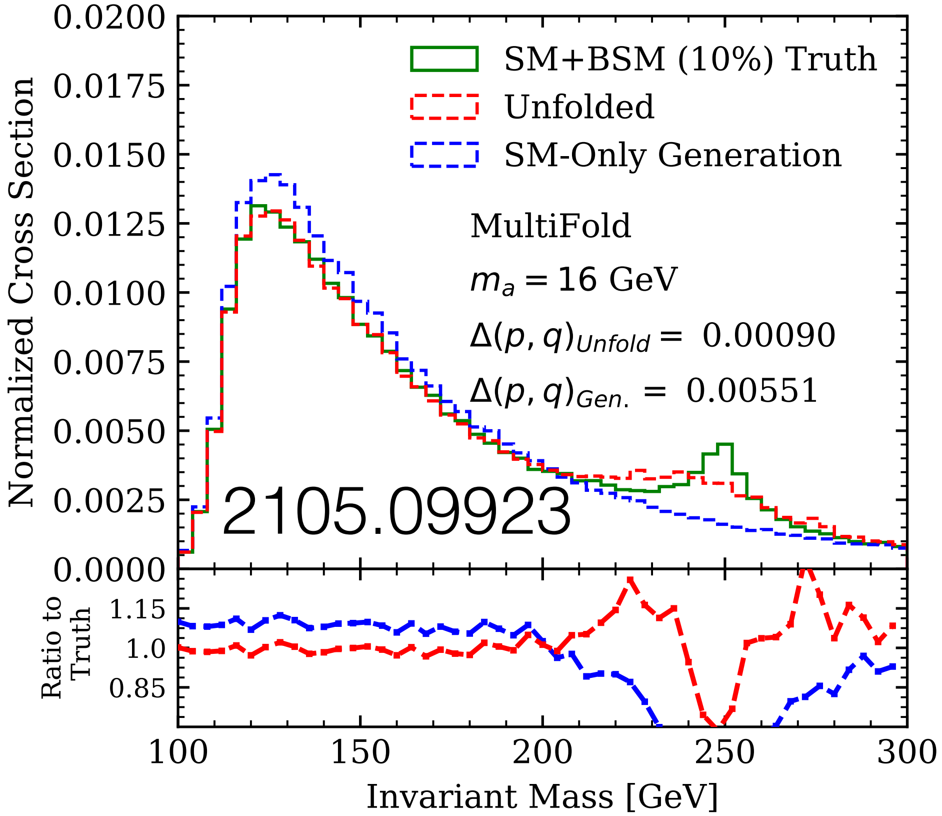
Unbinned Methods

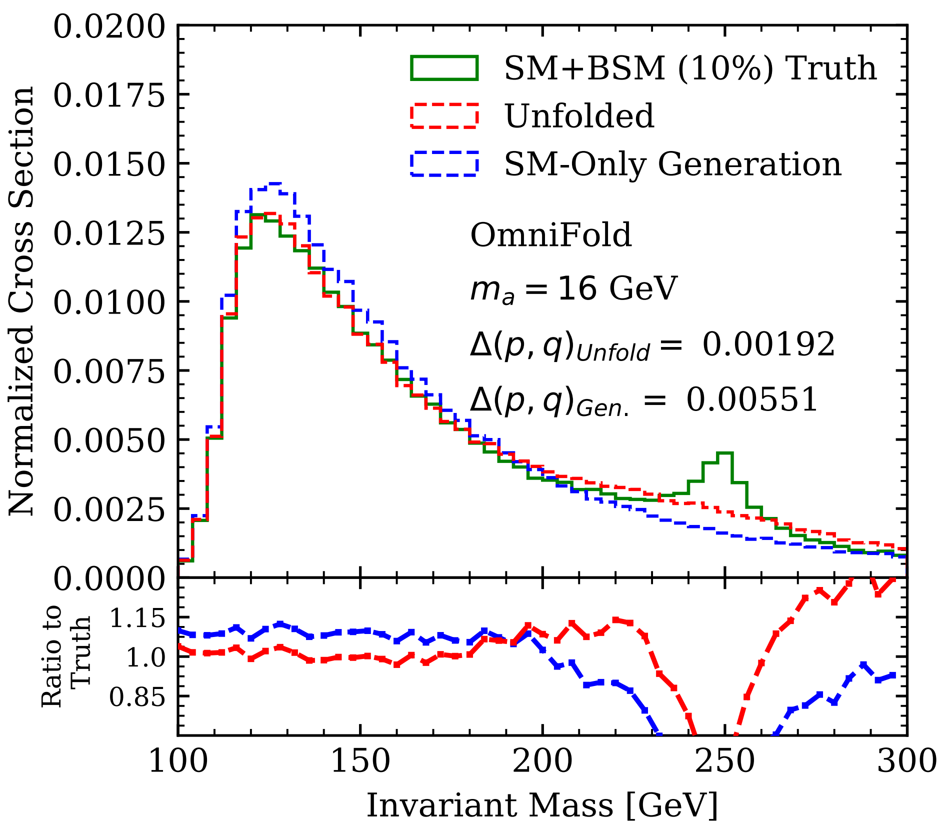


OmniFold: 1911.09107, cINN: 2006.06885

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

Can these approaches preserve BSM?





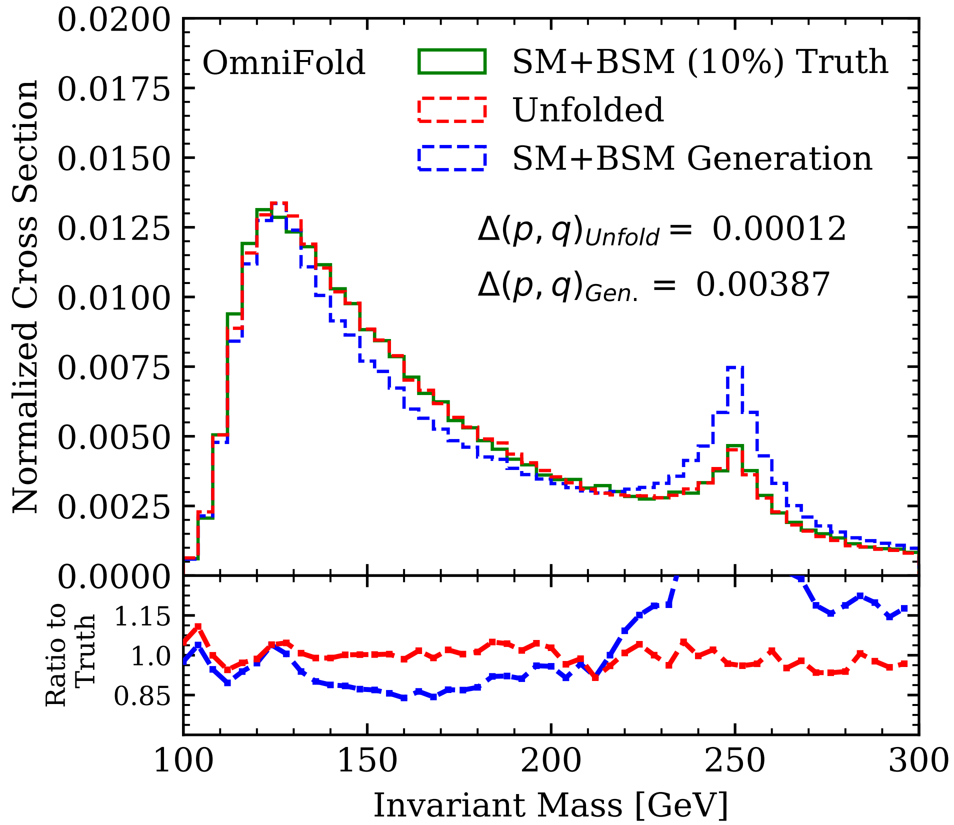
MultiFold: 10d cross section, OmniFold: all particles.

BSM: H->Za

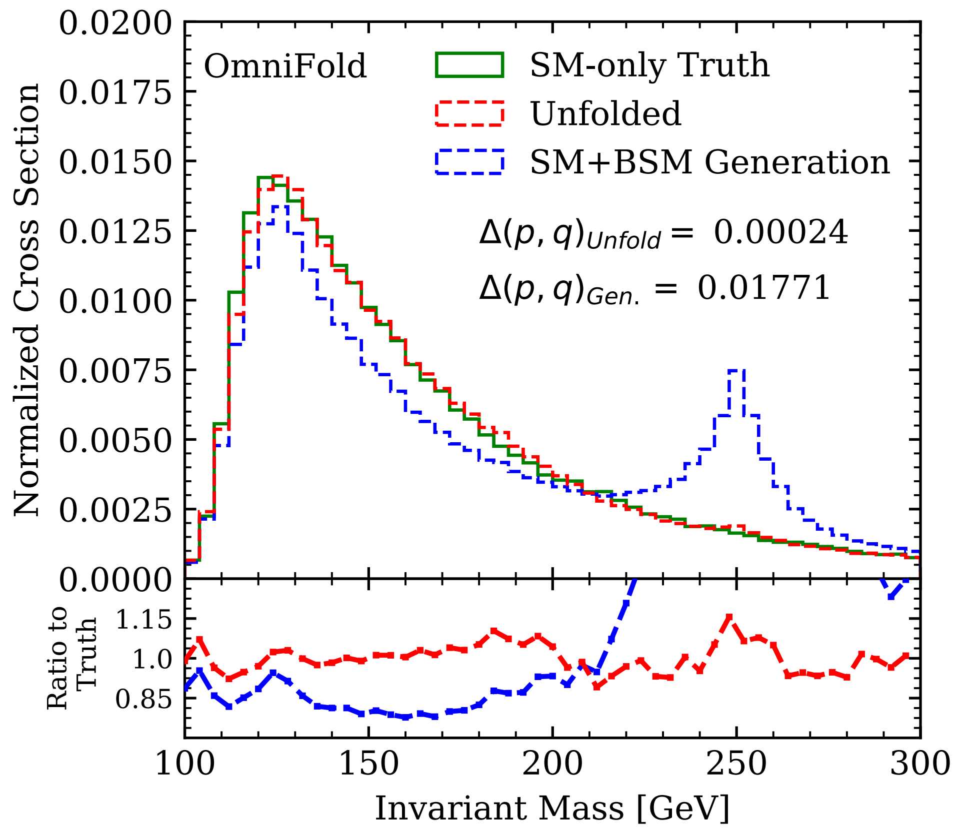
Answer: sort of … preserves anomaly when big enough (>1%). Precision continues to improve!

see also 1912.00477

Can these approaches preserve BSM?



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)

2105.09923

Challenge: how to publish/recast?

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!

+CMS open data study