HP2024
NAGASAKI

OMNIFOLD-HI

Advanced ML Unfolding for Heavy-Ion Data

Alexandre Falcão*
Adam Takacs

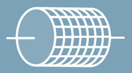
*alexandre.falcao@uib.no

Sep. 24th, 2024

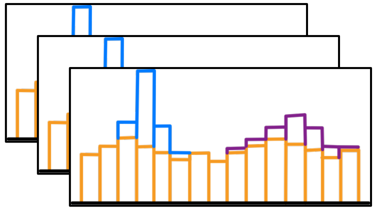
UNIVERSITY OF BERGEN

Recipe to measure a jet observable

detector



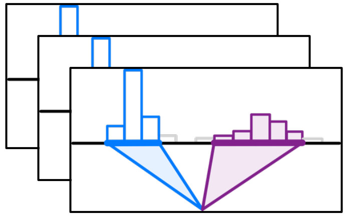
event measurement



event-by-event

backgroud subtraction

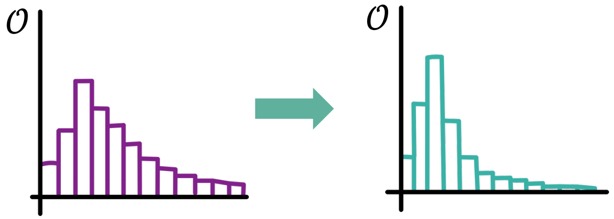
event-by-event background subtraction



statistics

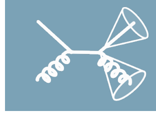
unfolding

unfolding of detector effects

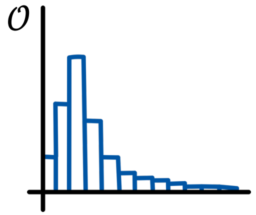


vs.

theory



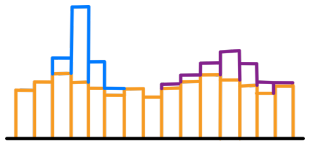
theoretical model



comparison with theoretical models and other experiments

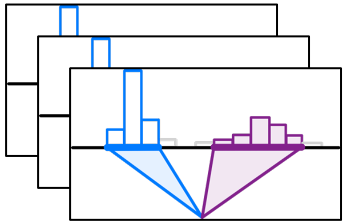
Recipe to measure a jet observable

event measurement



detector

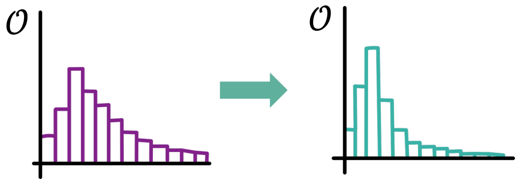
event-by-event
background subtraction



backgroud subtraction

event-by-event

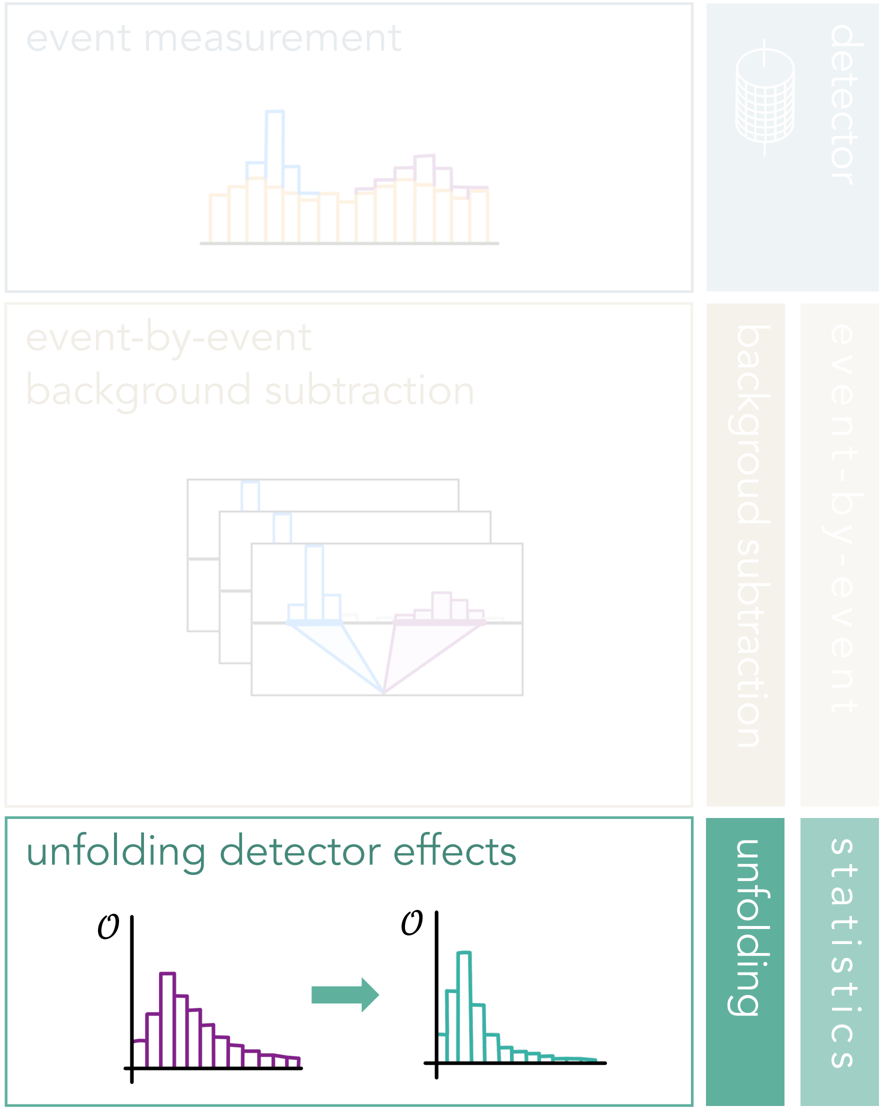
unfolding detector effects



unfolding

statistics

Traditional unfolding

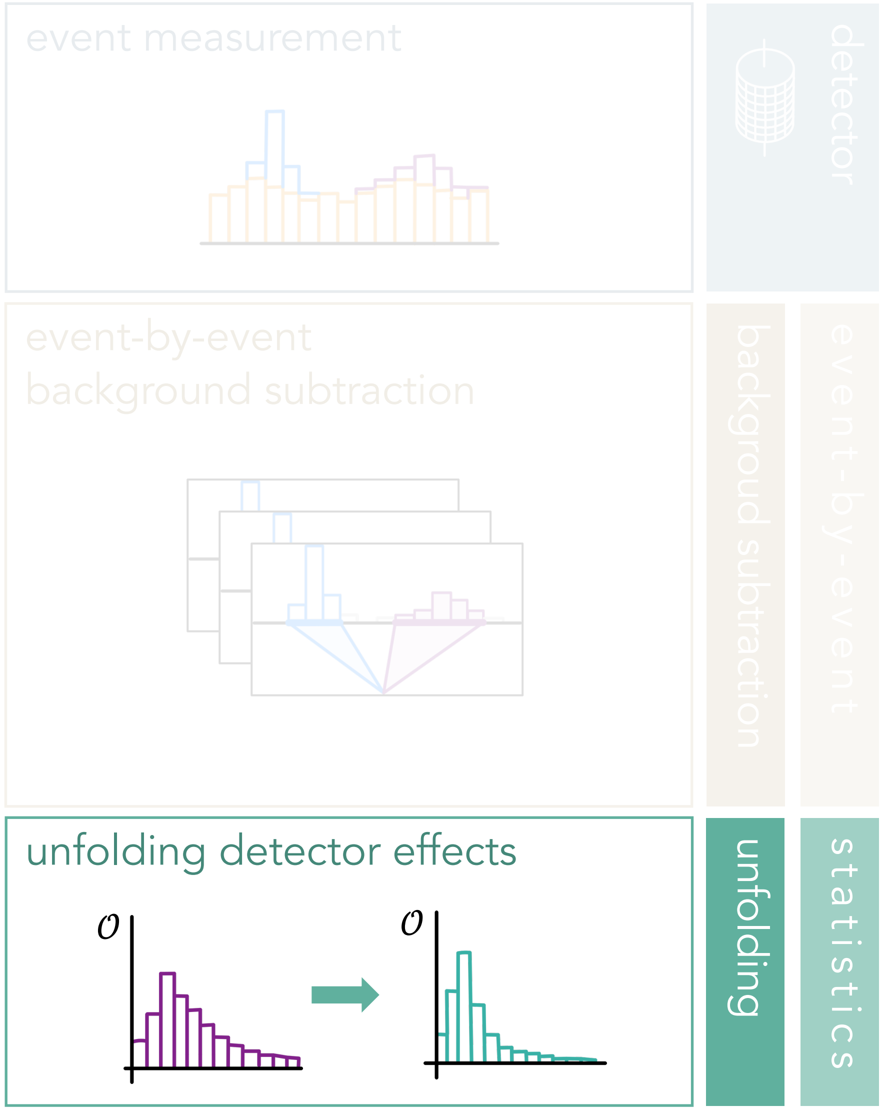


Traditional unfolding

Q: How would my observable look like without detector effects?

- limited resolution
- inefficiency
- distortion or smearing

A: Statistical proceadure to deconvolute detector effects from observables



Traditional unfolding

Q: How would my observable look like without detector effects?

- limited resolution
- inefficiency
- distortion or smearing

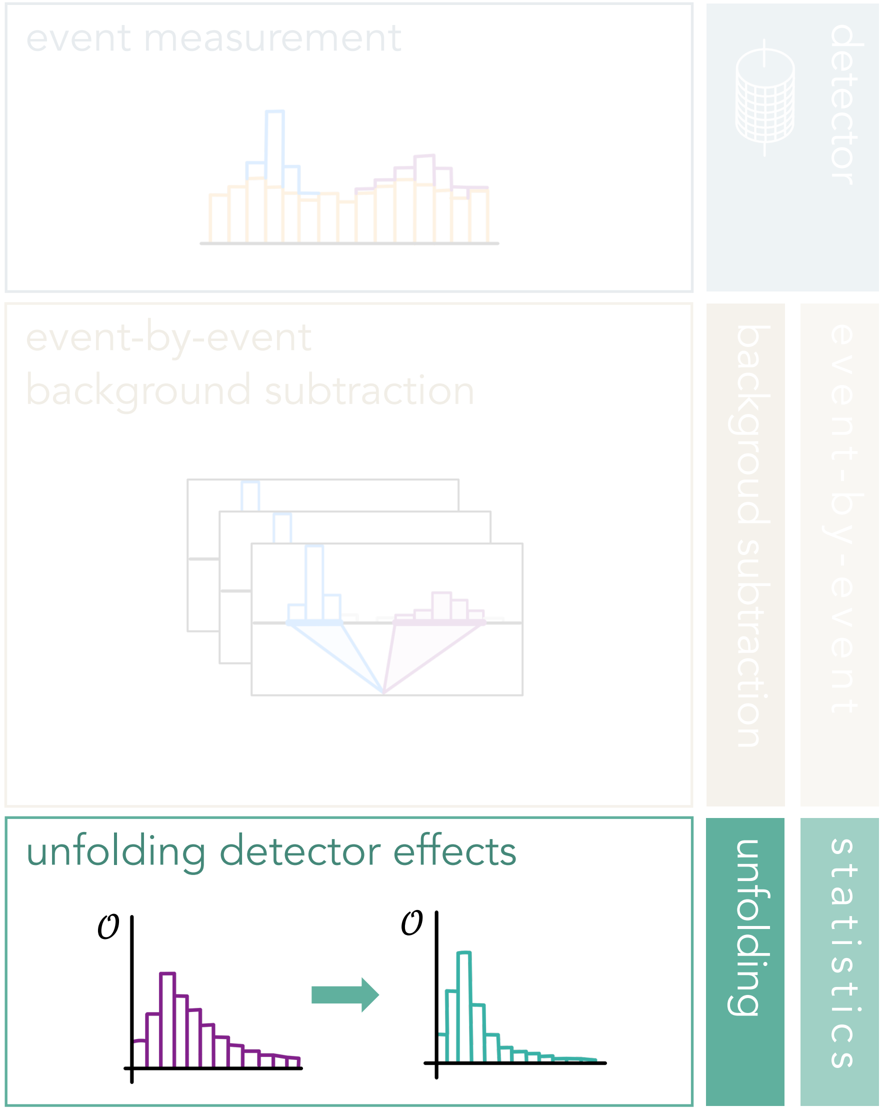
A: Statistical proceadure to deconvolute detector effects from observables

Binned porcedure:

- Iterative Bayesian Unfolding (IBU)

Unbinned procedure (with ML):

- density-based models: GAN's; VAE's; NF's
- classifier-based models: OmniFold



Background subtraction

1. jet clustering
2. background subtraction

jet-by-jet

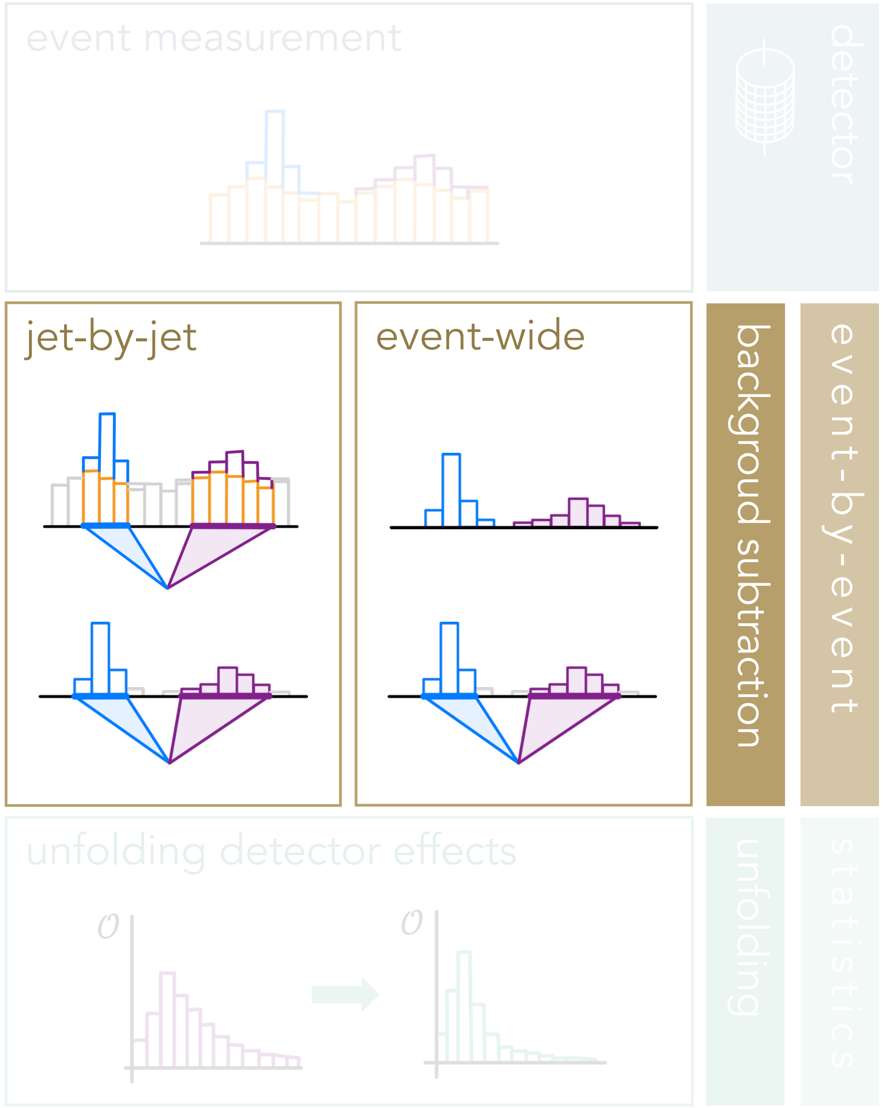
OR

event-wide??

(better for jet substructure)

1. background subtraction
2. jet clustering

e.g.: area subtraction; CS; ICS



Background subtraction

1. jet clustering
2. background subtraction

jet-by-jet

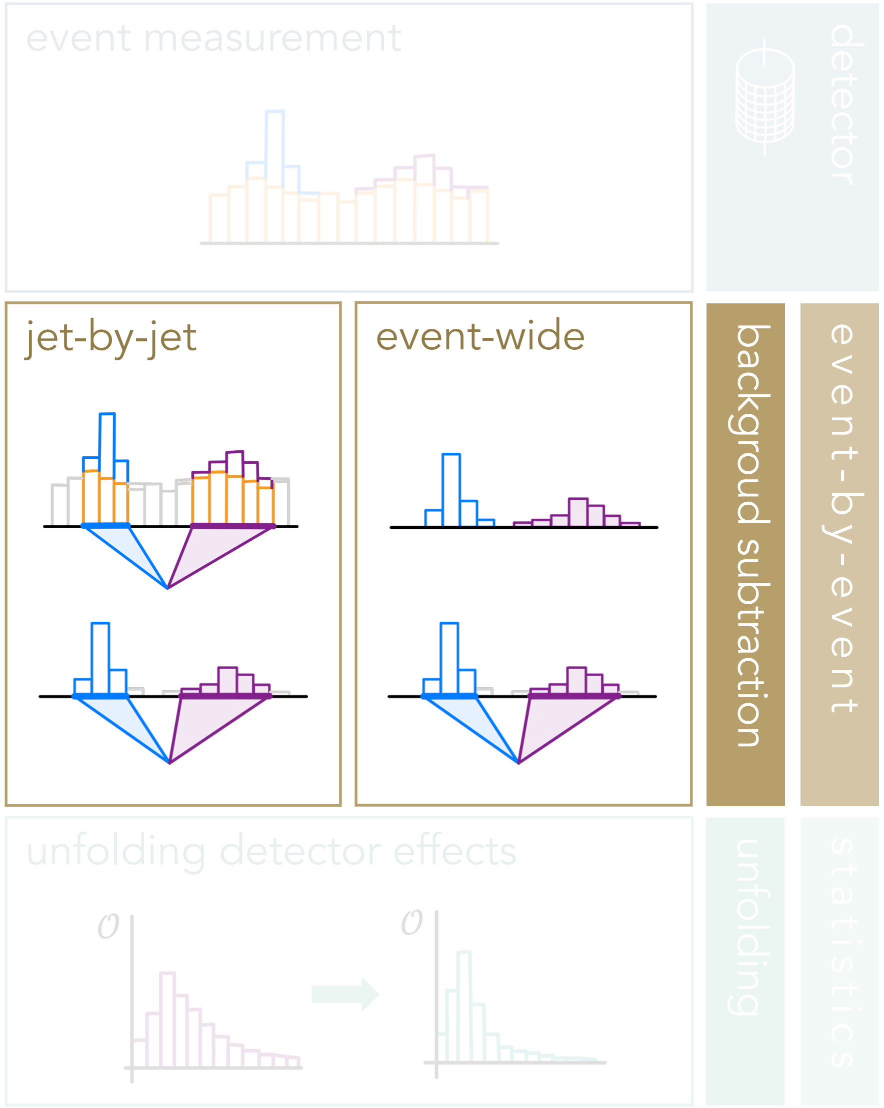
OR

1. background subtraction
2. jet clustering

event-wide??

(better for jet substructure)

e.g.: area subtraction; CS; ICS



Main sources of background:

- **pp:** pileup
- **HI:** underlying event (UE)

The Underlying Event in HI

In HI, complex UE from quark-gluon plasma (QGP):
(formation + hydrodynamical evolution)

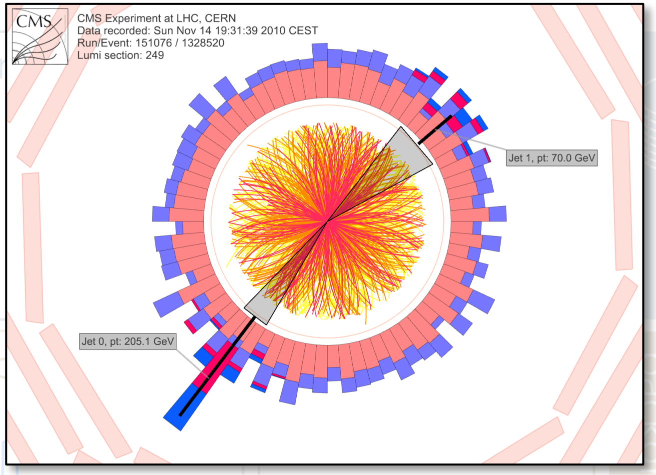
high-multiplicity thermal component

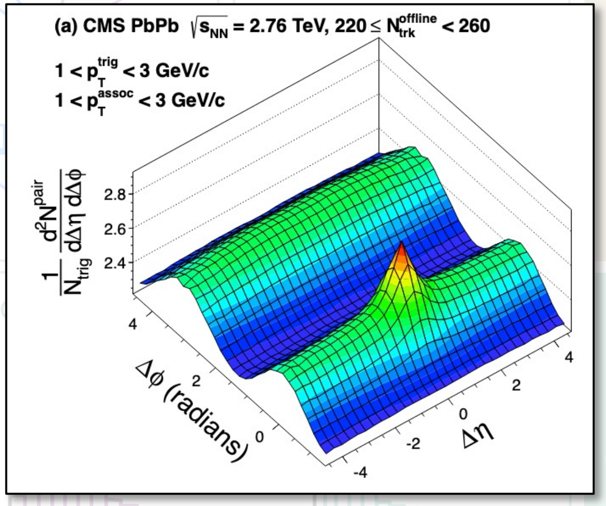
collective flow effects (v_2, v_3, v_4, …)

initial state fluctuations

interaction jet ↔ QGP

- jet quenching ⟺ UE modifies jet p_T and internal structure
- jet medium response modifies UE





The Underlying Event in HI

In HI, complex UE from quark-gluon plasma (QGP):
(formation + hydrodynamical evolution)

high-multiplicity thermal component

collective flow effects (v_2, v_3, v_4, …)

initial state fluctuations

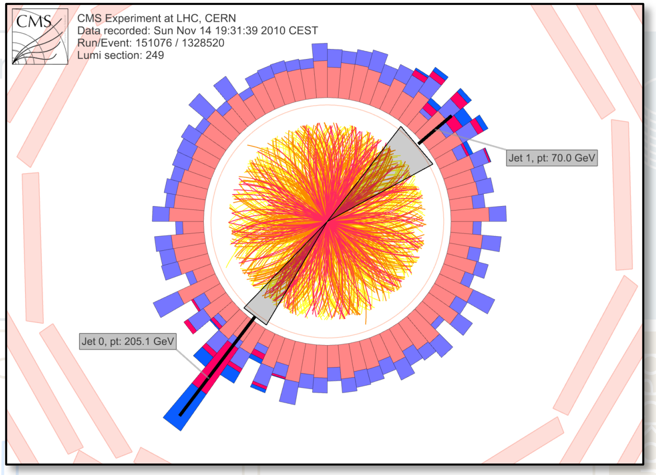
interaction jet ↔ QGP

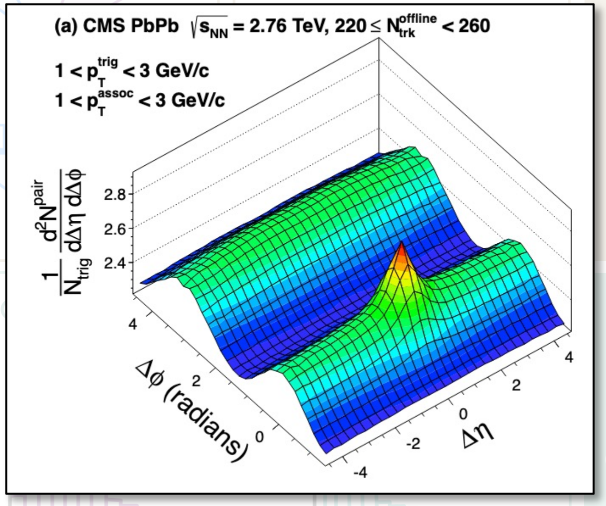
- jet quenching ⟺ UE modifies jet p_T and internal structure
- jet medium response modifies UE

Underlying event *cannot* be separated from the hard event!
(UE is *not* additive)

⇓

HI background is *ill-defined*





The Underlying Event in HI

In HI, complex UE from quark-gluon plasma (QGP):
(formation + hydrodynamical evolution)

high-multiplicity thermal component

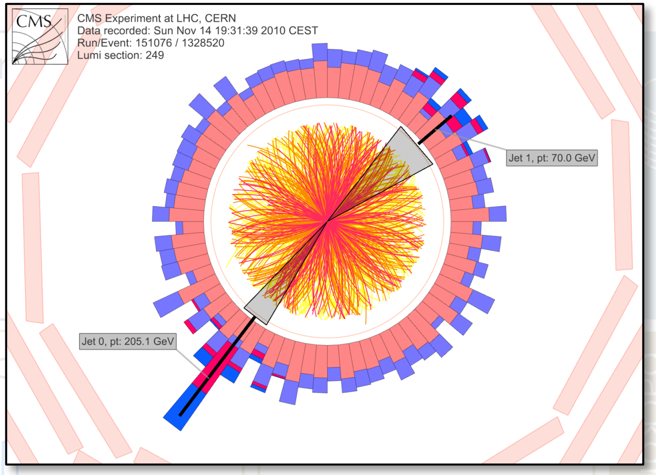
collective flow effects (v_2, v_3, v_4, …)

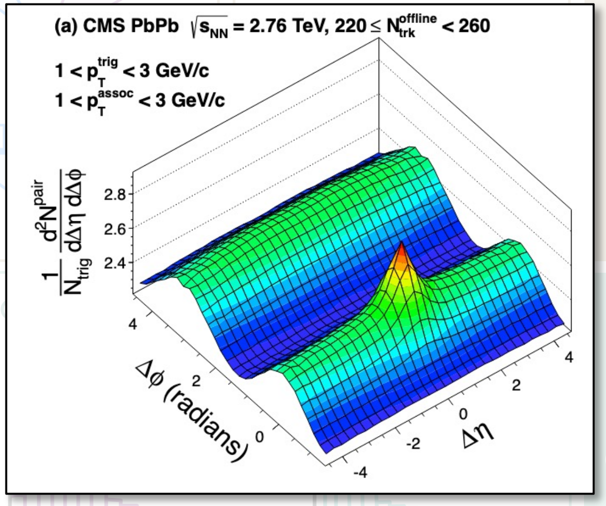
initial state fluctuations

interaction jet ↔ QGP

- jet quenching ⟺ UE modifies jet p_T and internal structure
- jet medium response modifies UE

Underlying event *cannot* be separated from the hard event!
(UE is *not* additive)
⇓
HI background is *ill-defined*





How to deal with it?

consensus: everyone uses the same subtraction method
e.g.: ICS; jet grooming

The Underlying Event in HI

In HI, complex UE from quark-gluon plasma (QGP):
(formation + hydrodynamical evolution)

high-multiplicity thermal component

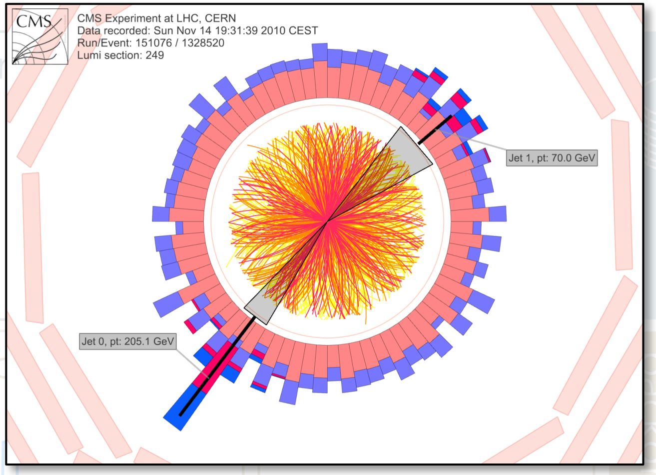
collective flow effects (v_2, v_3, v_4, …)

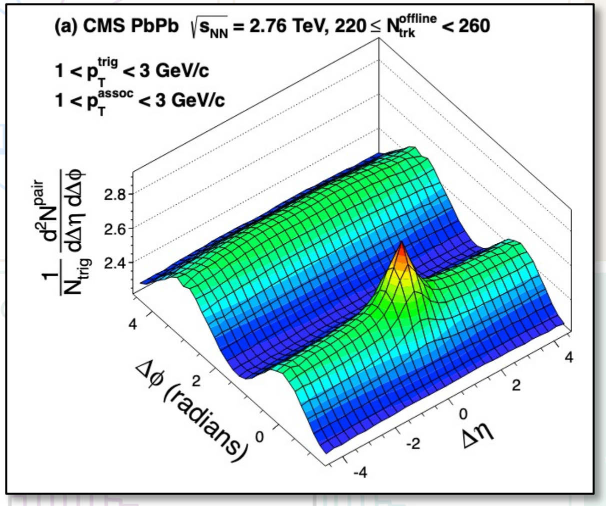
initial state fluctuations

interaction jet ↔ QGP

- jet quenching ⟺ UE modifies jet p_T and internal structure
- jet medium response modifies UE

Underlying event *cannot* be separated from the hard event!
(UE is *not* additive)
⇓
HI background is *ill-defined*





How to deal with it?

consensus: everyone uses the same subtraction method → adds subtraction uncertainty
e.g.: ICS; jet grooming

The Underlying Event in HI

In HI, complex UE from quark-gluon plasma (QGP):
(formation + hydrodynamical evolution)

high-multiplicity thermal component

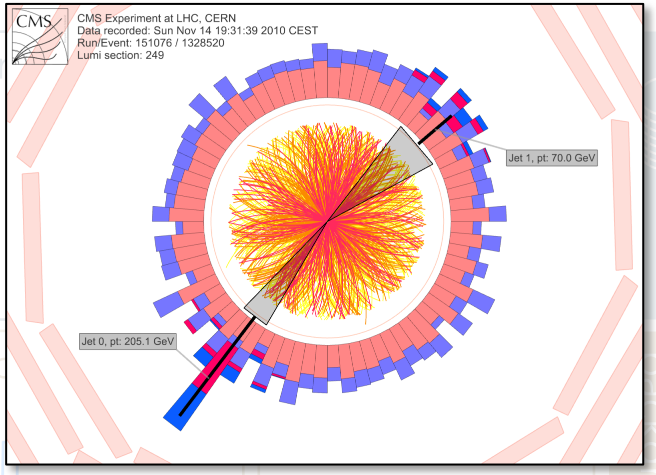
collective flow effects (v_2, v_3, v_4, …)

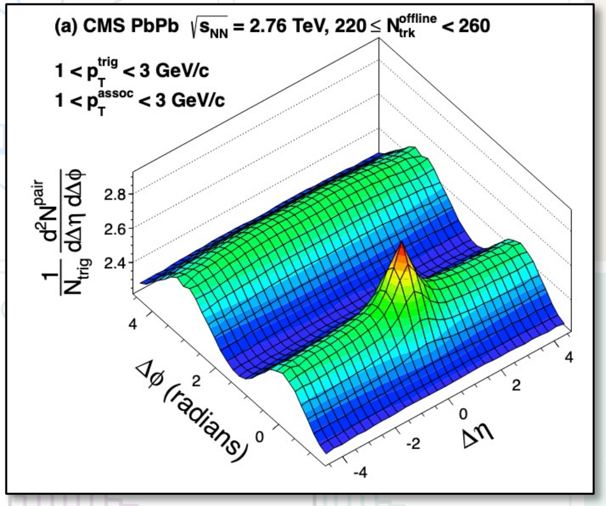
initial state fluctuations

interaction jet ↔ QGP

- jet quenching ⟺ UE modifies jet p_T and internal structure
- jet medium response modifies UE

Underlying event *cannot* be separated from the hard event!
(UE is *not* additive)
⇓
HI background is *ill-defined*





How to deal with it?

consensus: everyone uses the same subtraction method
e.g.: ICS; jet grooming

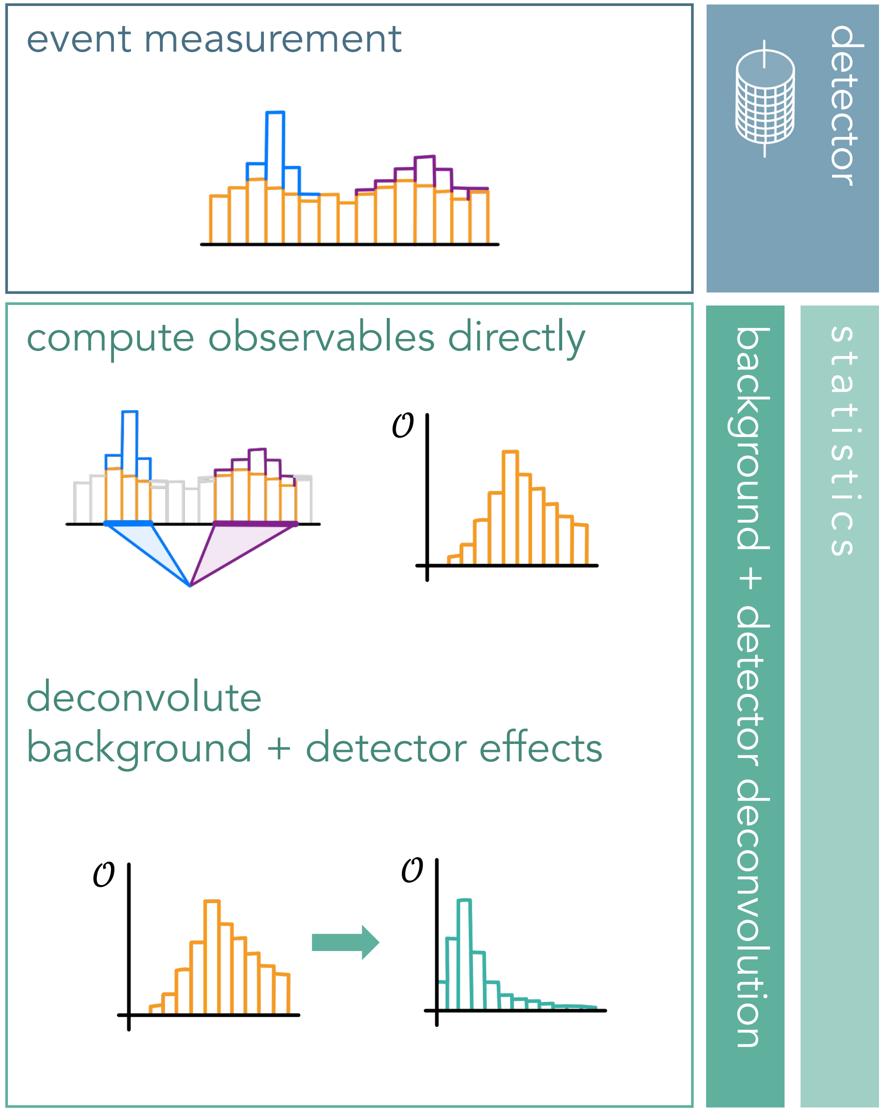
→ adds subtraction uncertainty

Is there as simpler way?

Background + detector deconvolution

We can extend the unfolding of detector effects to include background effects

- combine deconvolution of background and detector effects on observable
- background is defined by simulation needed for deconvolution
- choice of background to deconvolute eliminates subtraction uncertainty



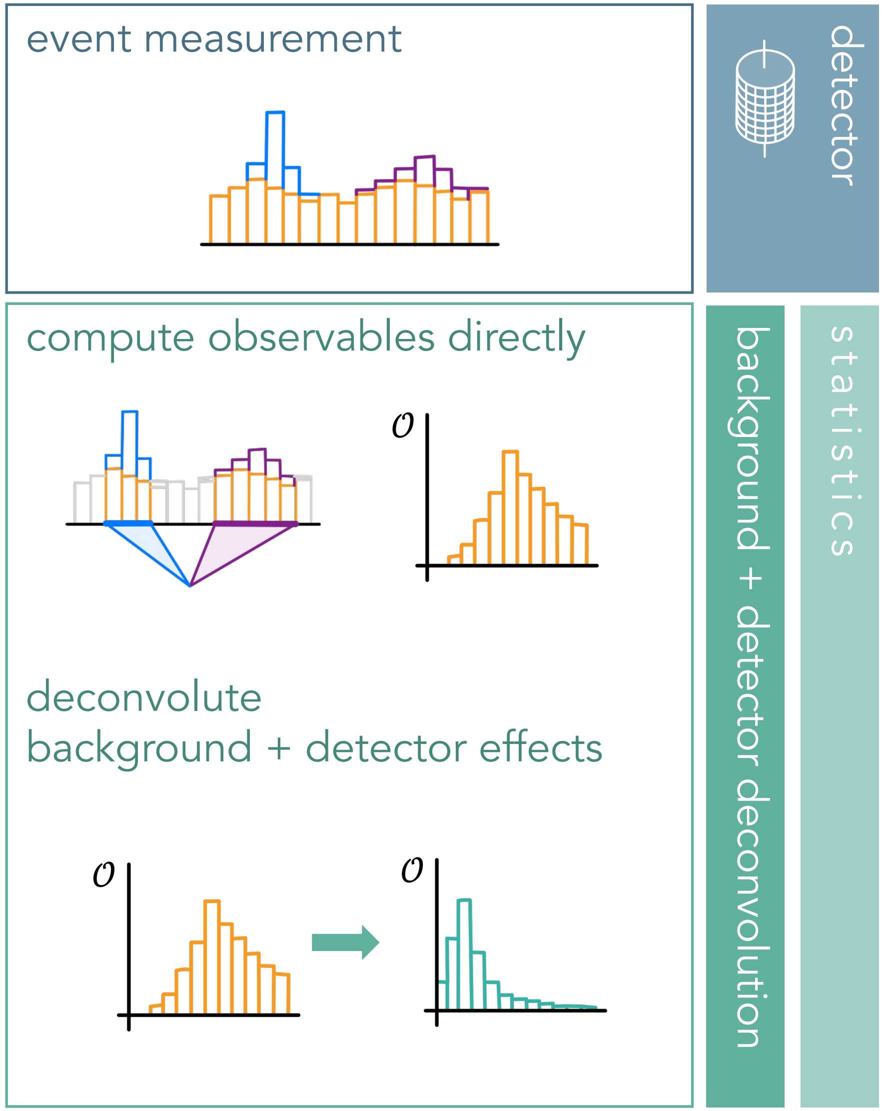
Background + detector deconvolution

We can extend the unfolding of detector effects to include background effects

- combine deconvolution of background and detector effects on observable
- background is defined by simulation needed for deconvolution
- choice of background to deconvolute eliminates subtraction uncertainty
- possible to use unfolding statistical procedures:

  Binned:
  - Iterative Bayesian Unfolding (IBU)

  Unbinned (with ML):
  - OmniFold



Unfolding procedures

IBU:

[arXiv:1010.0632]

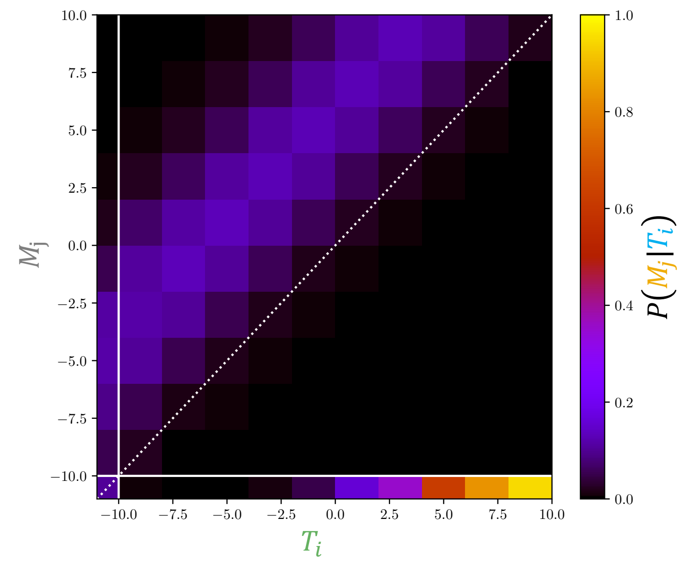
$$x(T_i) = \sum_j P(T_i|M_j) \cdot x(M_j)$$

response matrix

true bin counts

measured bin counts

from Monte Carlo + simualtion



Unfolding procedures

IBU:

[arXiv:1010.0632]

$$x(T_i) = \sum_j P(T_i|M_j) \cdot x(M_j)$$

response matrix

true bin counts

measured bin counts

OmniFold:

[arXiv:1911.09107, arXiv:2105.04448]

$$x(t) = \int \mathrm{d}m \; p(t|m) \cdot x(m)$$

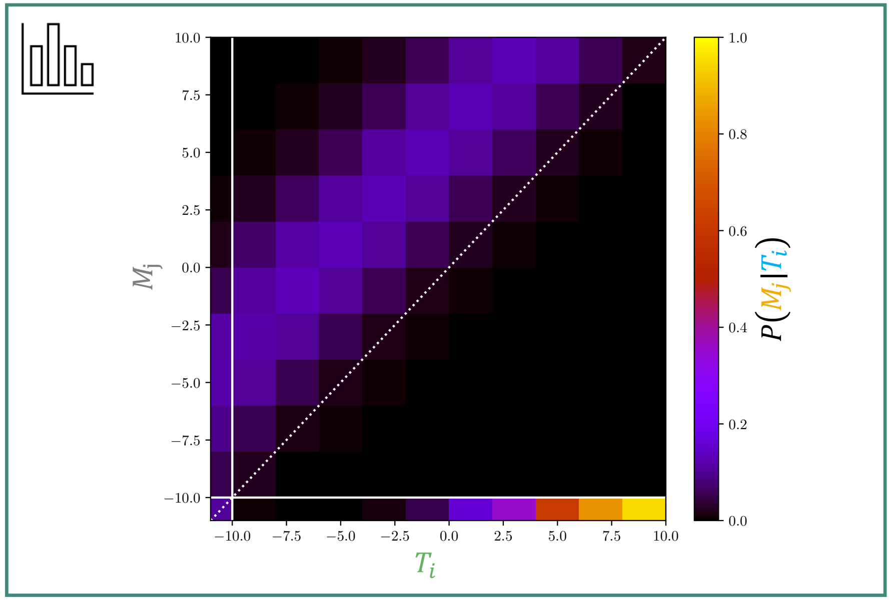
"response matrix" over event space

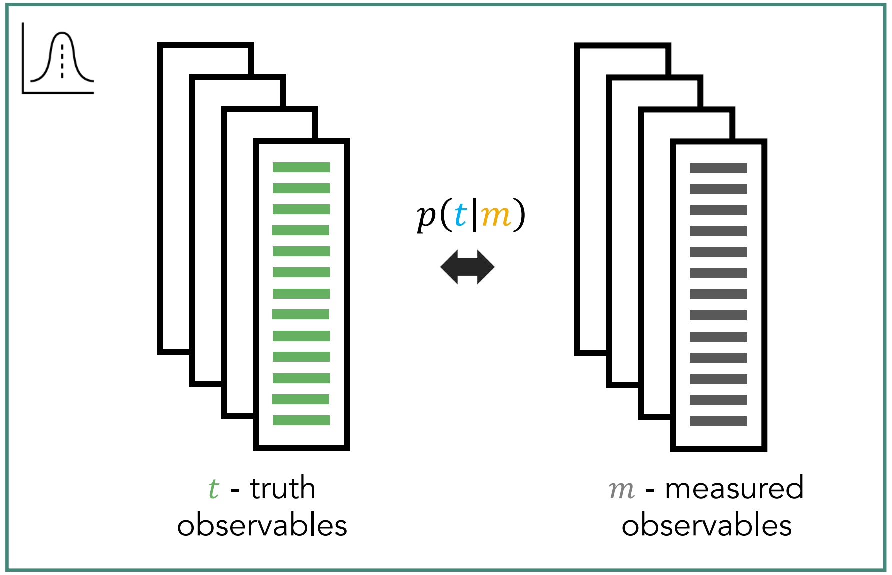
true event distribution

measured event distribution

- easy to extend to higher dimensions (up to whole event unfolding!)

from Monte Carlo + simualtion





OmniFold-HI

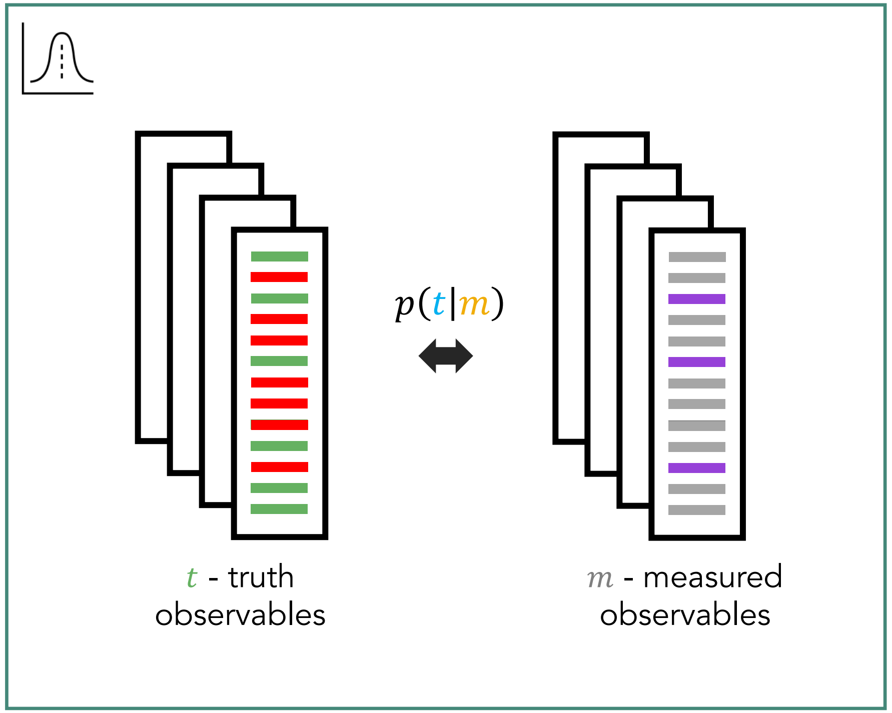
We introduce **OmniFold-HI:**

= Same as OmniFold

\+ Acceptance and efficiency formalized
(rigurous proof of OmniFold from IBU)

\+ Correct estimation of **fake** and **trash** events
(important for fake jets from UE)

\+ Deconvolution of uncertainty



Acceptance: T trash event - true event that are not measured
Efficiency: F fake event - measured event without truth origin

OmniFold-HI

We introduce **OmniFold-HI:**

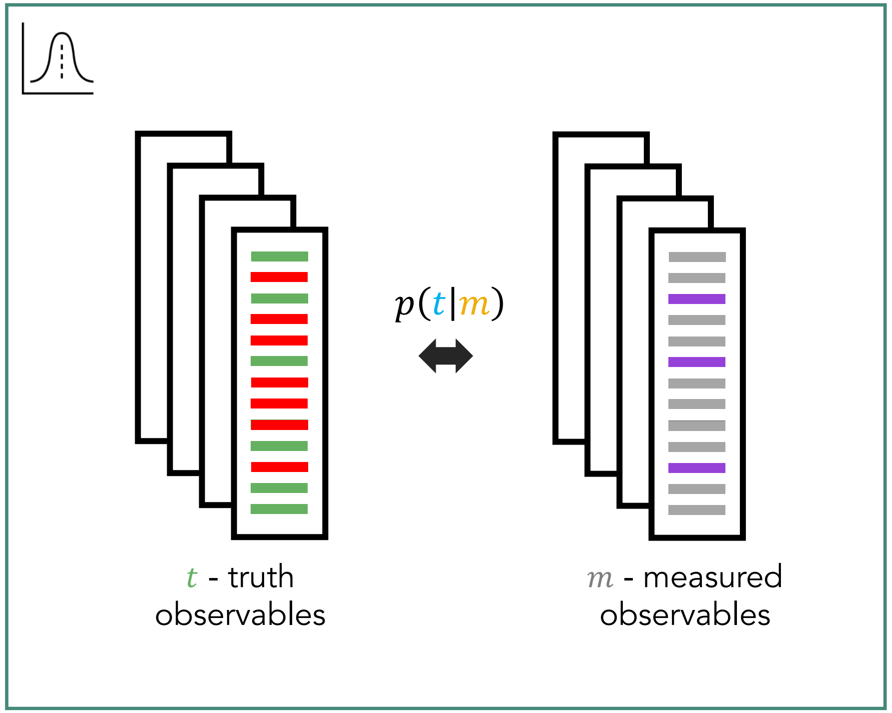
= Same as OmniFold

\+ Acceptance and efficiency formalized
(rigurous proof of OmniFold from IBU)

\+ Correct estimation of fake and trash events
(important for fake jets from UE)

\+ Deconvolution of uncertainty

! algorithm tailored for background deconvolution but useful for traditional detector unfolding !



(**Acceptance:** T trash event - true event that are not measured
Efficiency: F fake event - measured event without truth origin)

Setup

Data: (no real truth data available…)

Truth: Herwig7

Measured: Herwig7 + background + Delphes

detector simulation

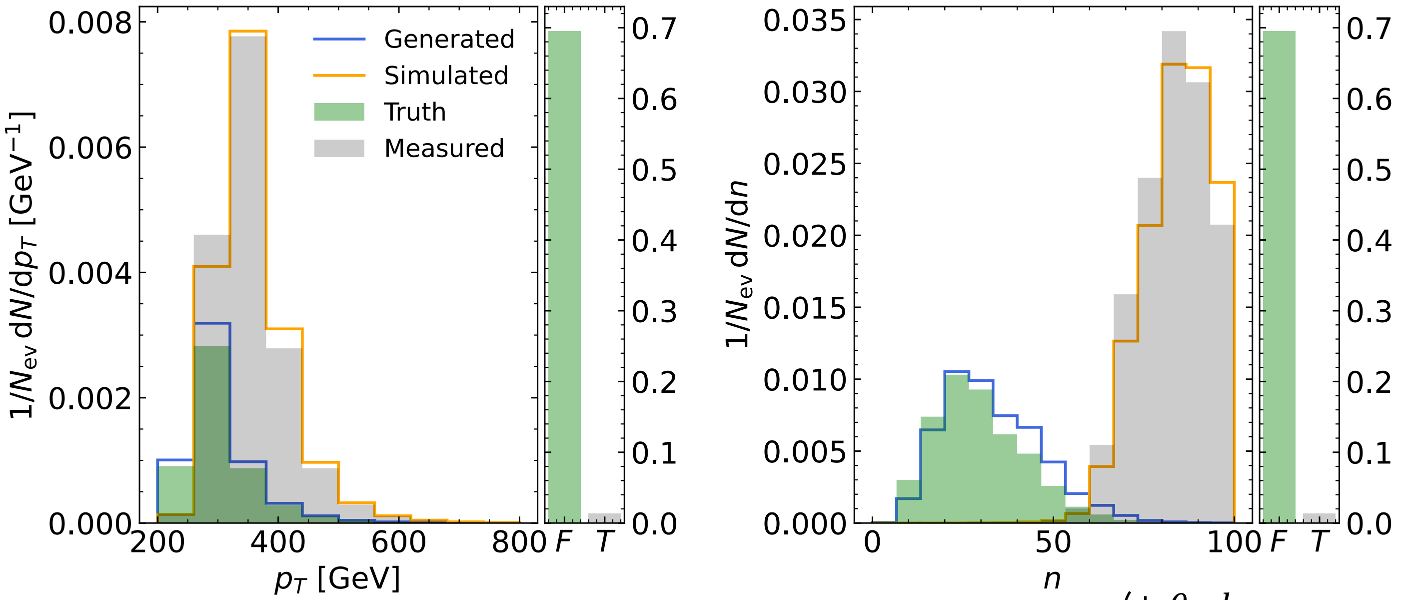
MC for deconvolution procedure:

Generated: Pythia8

Simulated: Pythia8 + background + Delphes

Background:

- Thermal background:
  multiplicity = 7000
  $\langle p_T \rangle$ = 1.2 GeV



$(+\,\theta_g, k_{T,g}, m, m_{sd}, n_{sd})$

Truth/Measured: Herwig7 + thermal bkg. + Delphes
Smeared/Generated: Pythia8 + thermal bkg. + Delphes

Setup

Data: (no real truth data available…)

Truth: Herwig7

Measured: Herwig7 + background + Delphes (detector simulation)

MC for deconvolution procedure:

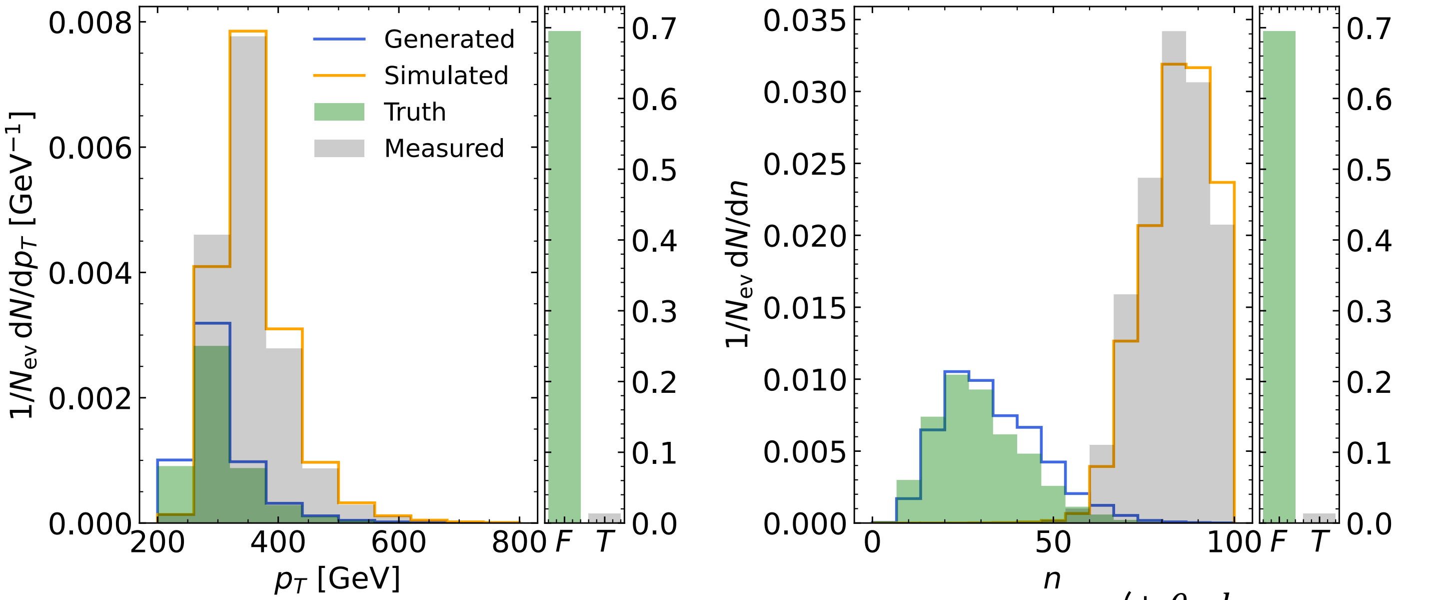
Generated: Pythia8

Simulated: Pythia8 + background + Delphes

Background:

- Thermal background:
  multiplicity = 7000
  $\langle p_T \rangle$ = 1.2 GeV

Goal:
deconvolute *chosen background* effects together with detector effects



$(+\ \theta_g, k_{T,g}, m, m_{sd}, n_{sd})$

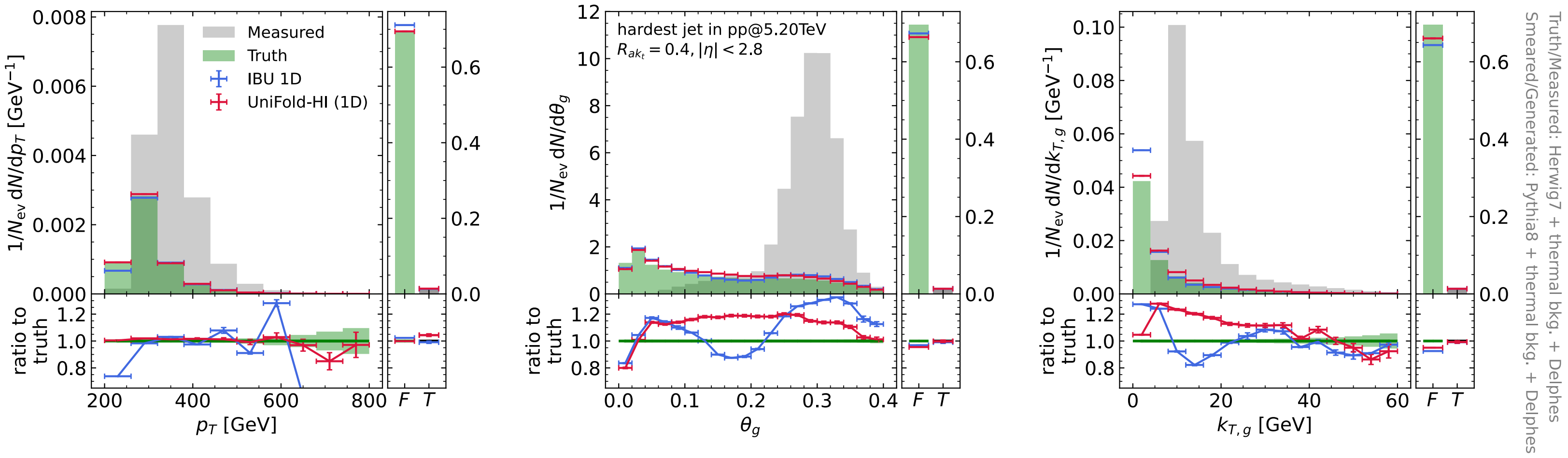
Truth/Measured: Herwig7 + thermal bkg. + Delphes
Smeared/Generated: Pythia8 + thermal bkg. + Delphes

Results: Background + Detector deconvolution

UniFold: one-dimentional OmniFold
MultiFold n**D:** n-dimentional OmniFold

1D deconvolution:

3 observables deconvoluted independently

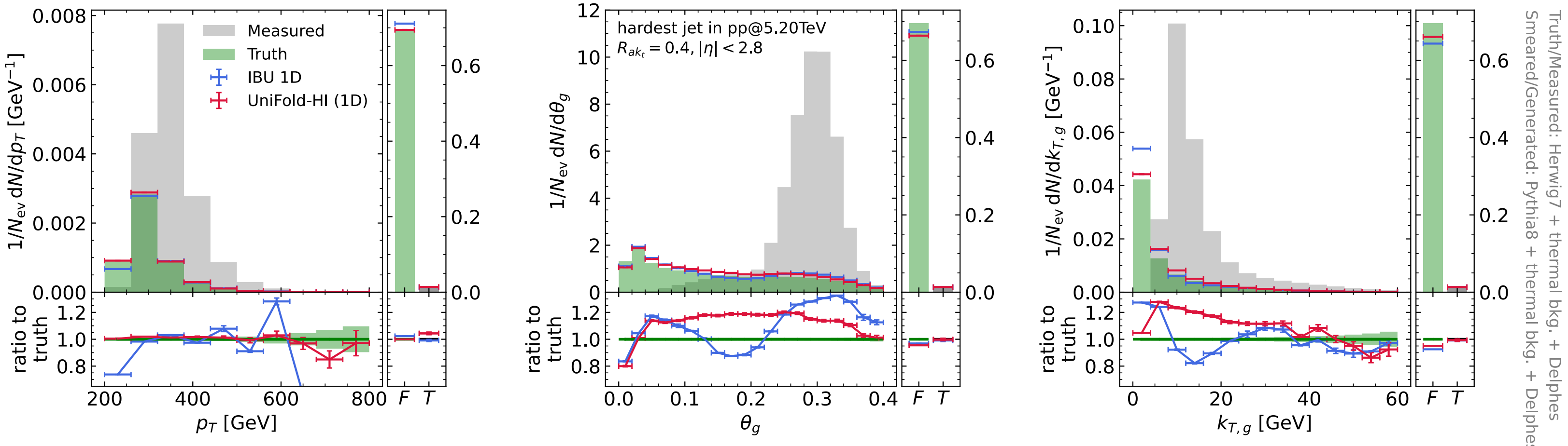


Results: Background + Detector deconvolution

(**UniFold:** one-dimentional OmniFold
MultiFold nD: n-dimentional OmniFold)

1D deconvolution:

3 observables deconvoluted independently

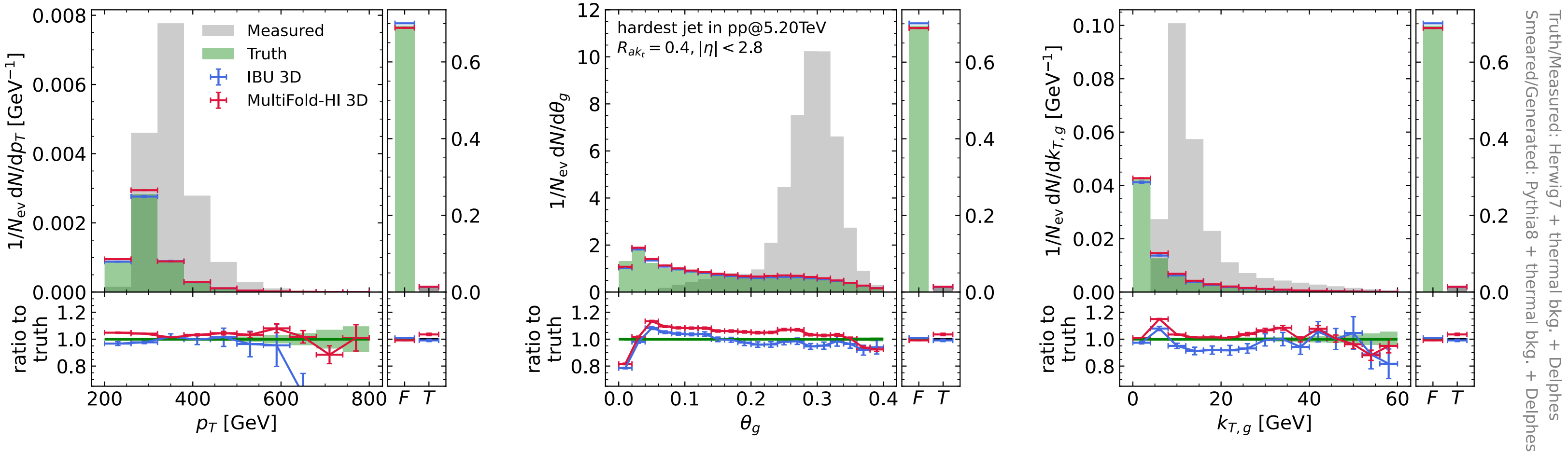


Vs.

3D deconvolution:

3 observables deconvoluted together

- better performance
- agreement between procedures

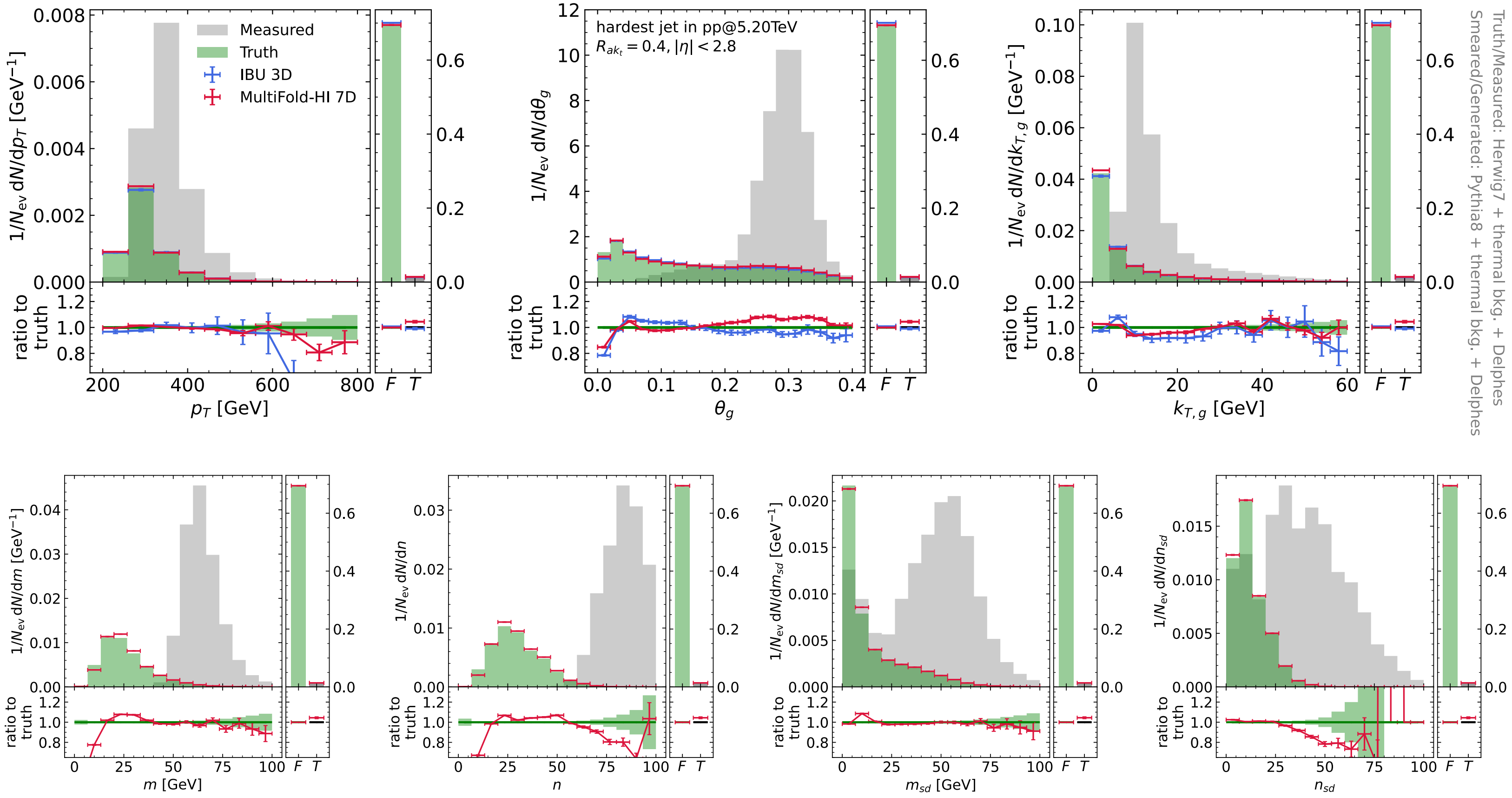


Results: Background + Detector deconvolution

7D deconvolution:

7 observables deconvoluted together

(too computionaly expensive for IBU)

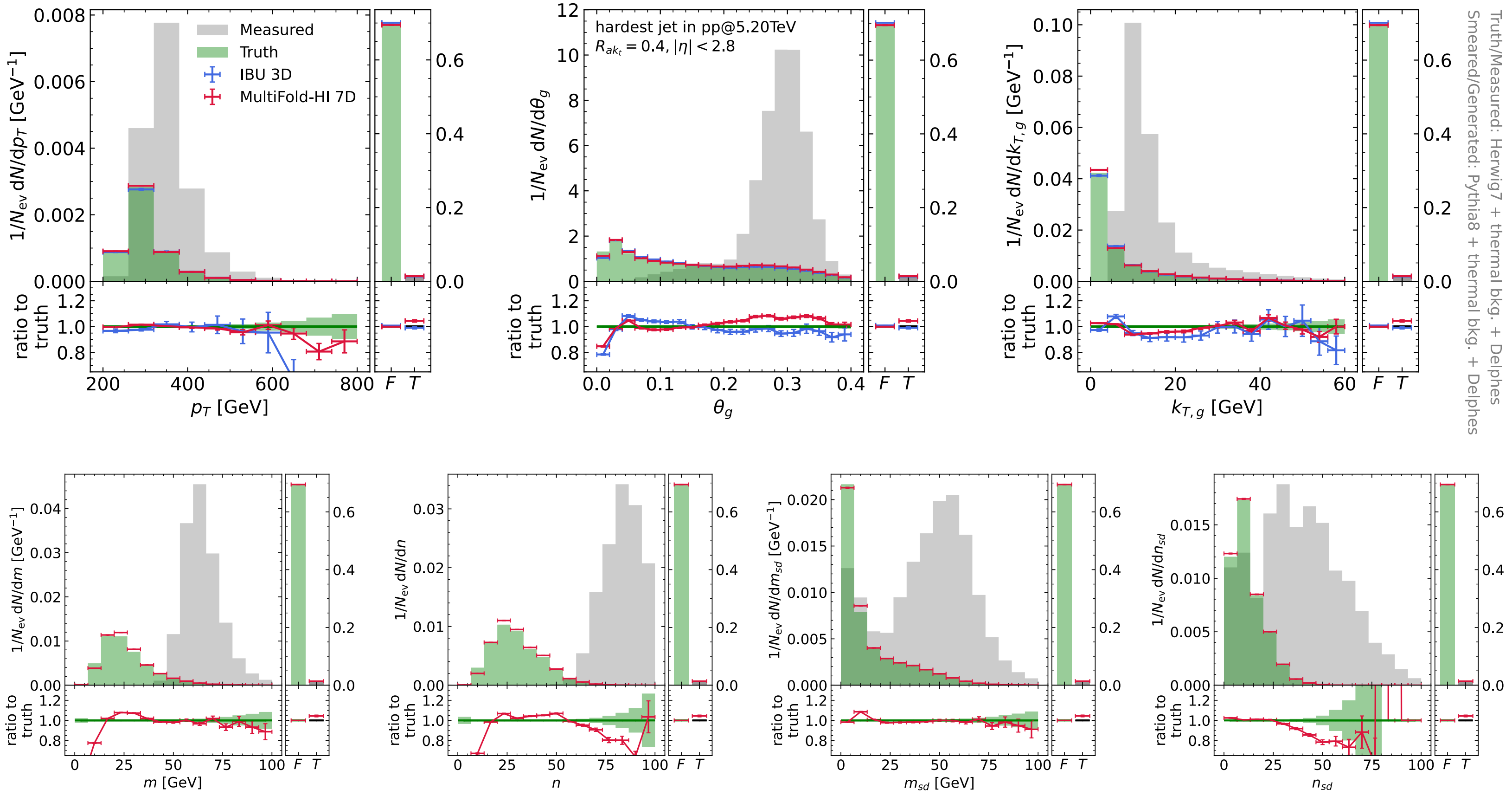


Results: Background + Detector deconvolution

7D deconvolution:
7 observables deconvoluted together

(too computionaly expensive for IBU)

- Both unfolding techniques succeed at deconvoluting background effects
- OmniFold-HI scales easily with deconvolution dimension



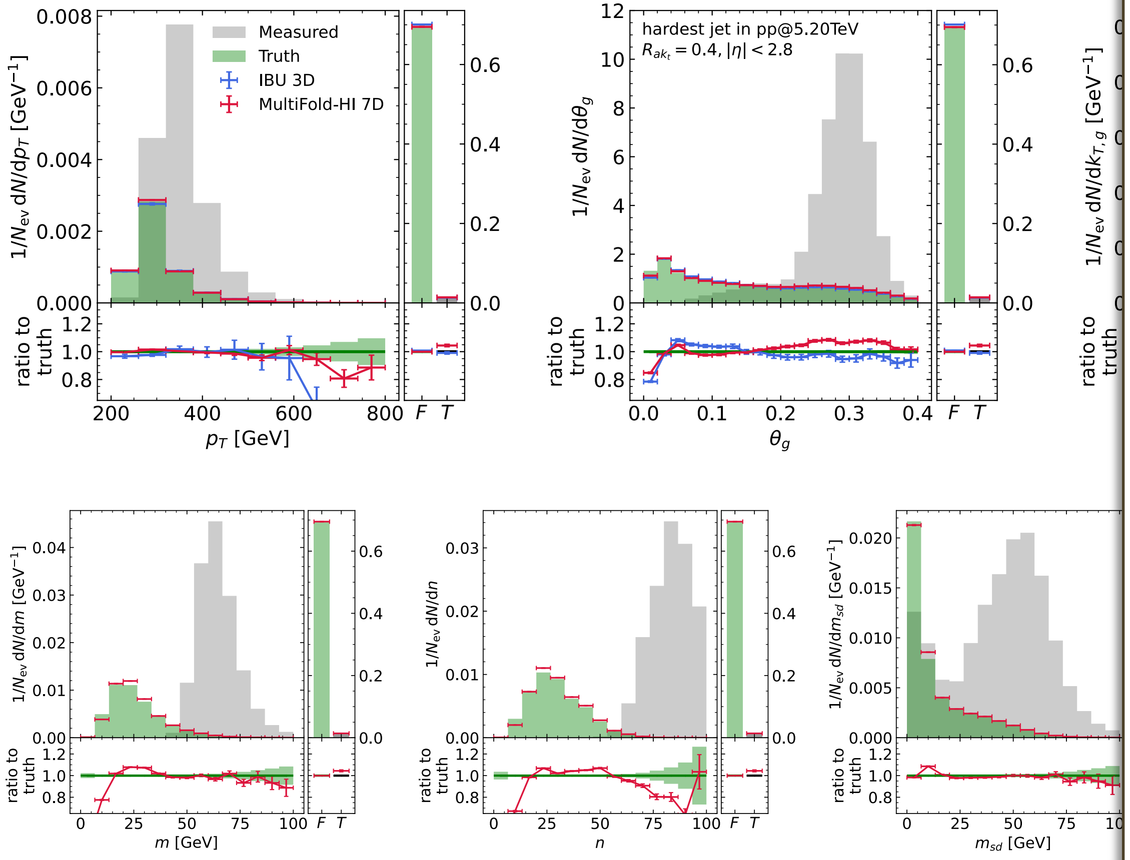
Results: Background + Detector deconvolution

7D deconvolution:

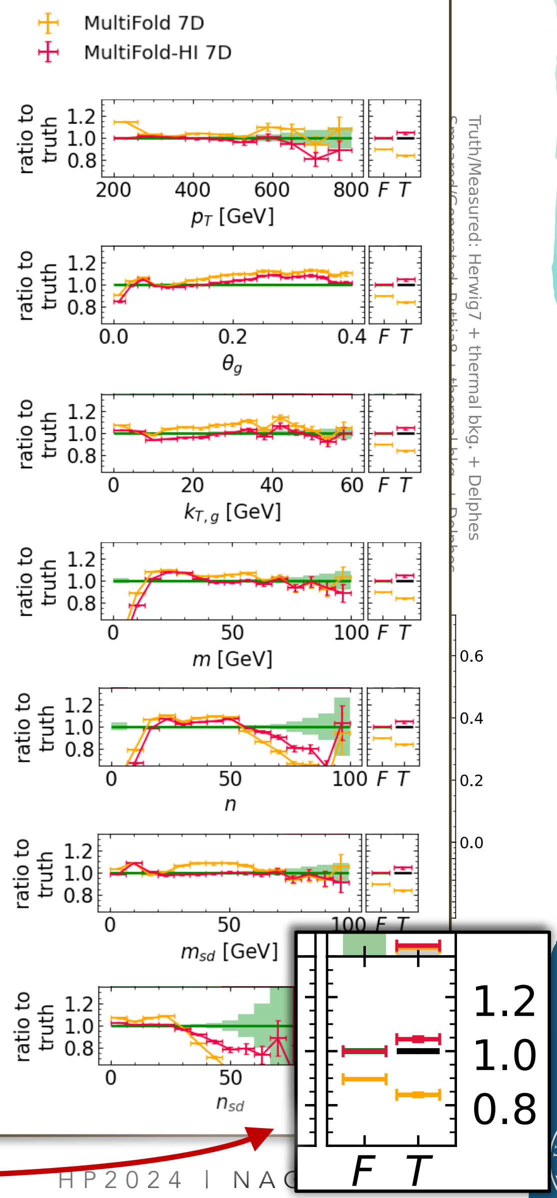
7 observables deconvoluted together

(too computionaly expensive for IBU)

- Both unfolding techniques succeed at deconvoluting background effects
- OmniFold-HI scales easily with deconvolution dimension
- OmniFold-HI outperformes original OmniFold in fake and trash estimation



Comparison with original OmniFold



Summary

- Experimental measurement of jet observables
  - background subtraction + detector unfolding
- Background is ill-defined in HI collisions
- Deconvolution of background effects can be combined with detector unfolding
- Unfolding techniques can be used:
  - IBU
  - OmniFold-HI
    (upgrade to OmniFold + formal approach to acceptance and efficiency)

Next steps

- Better HI simulation to study use in real data
  - hydrodynamic background + jet quenching
- Increase number of observables
- Study use of deconvolution after subtraction

BACKUP

OmniFold vs. OmniFold-HI

(**MultiFold nD:** n-dimentional OmniFold
UniFold: one-dimentional OmniFold)

Unfolding detector effects only