

# Unfolding techniques in HEP

Tommaso Cresta  
University of Pavia  
INFN Pavia  
ATLAS Collaboration



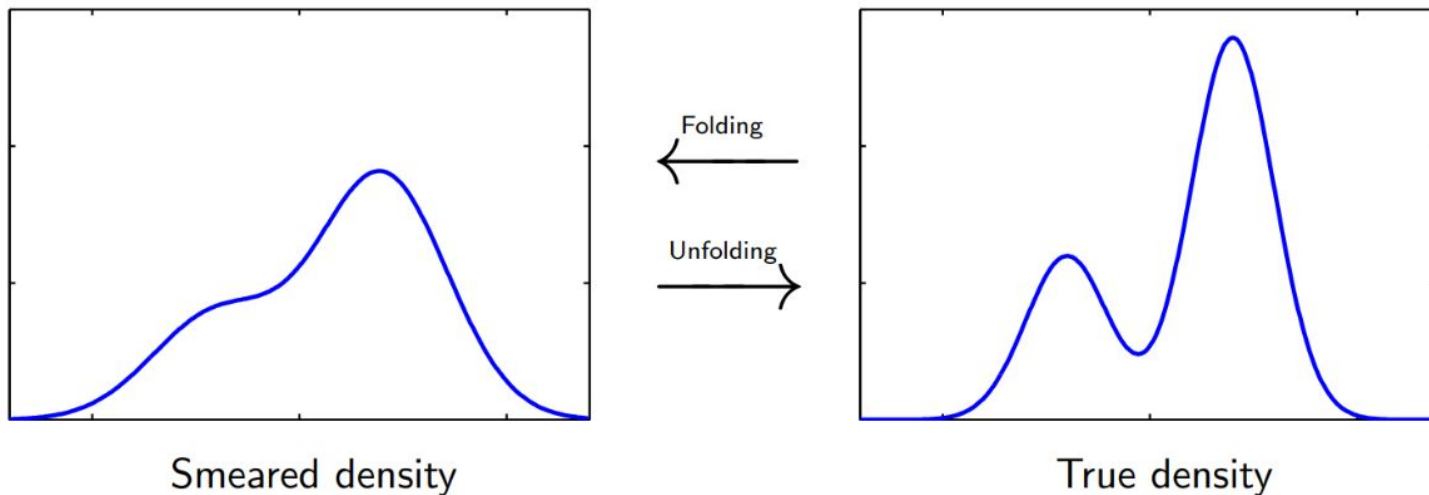
UNIVERSITÀ  
DI PAVIA



# Data Unfolding

The **experimentally measured (reco)** distribution is the convolution between the **true** distribution and **detector response**.

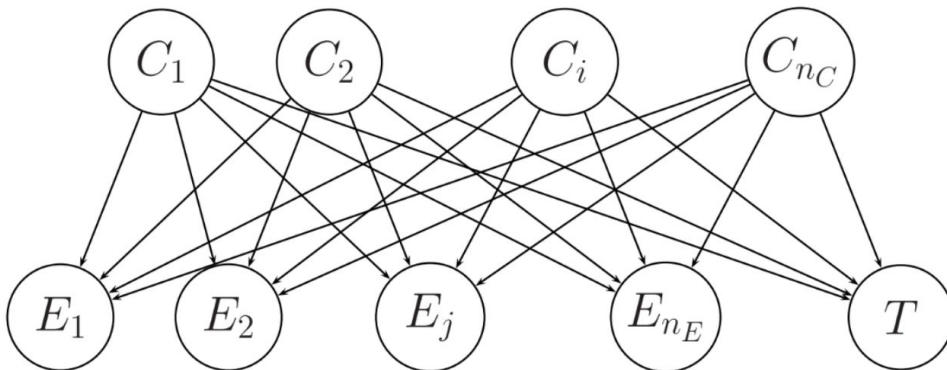
The reco distribution can be deconvoluted with **Unfolding**.



# Iterative Bayesian Unfolding

The most used algorithm is **Iterative Bayesian Unfolding (IBU)** and it is based on a probabilistic approach.

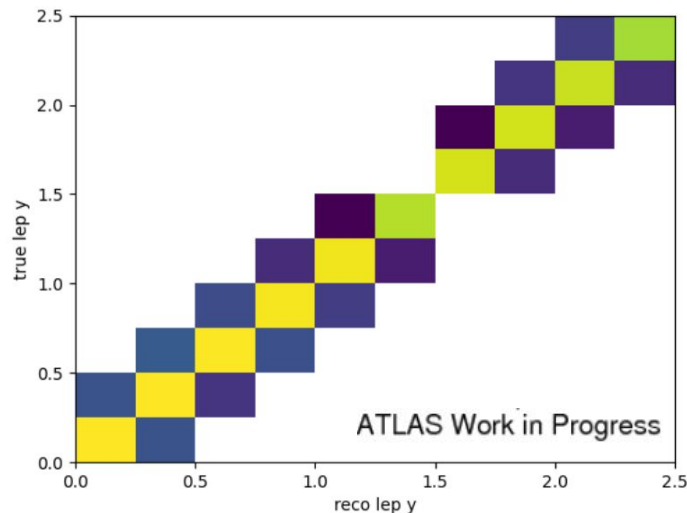
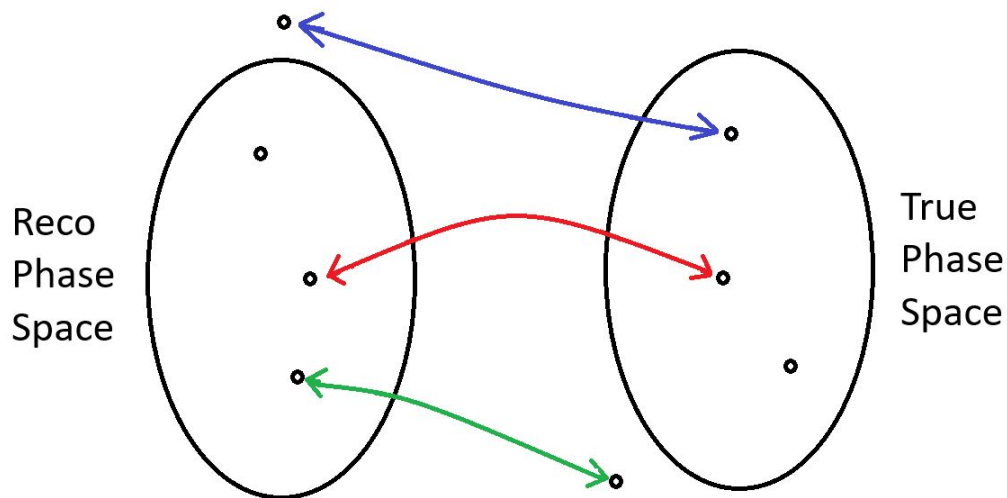
$$P(C_i | E_j) = \frac{P(E_j | C_i) \cdot P(C_i)}{\sum_i P(E_j | C_i) \cdot P(C_i)}$$



# Quantities necessary for unfolding

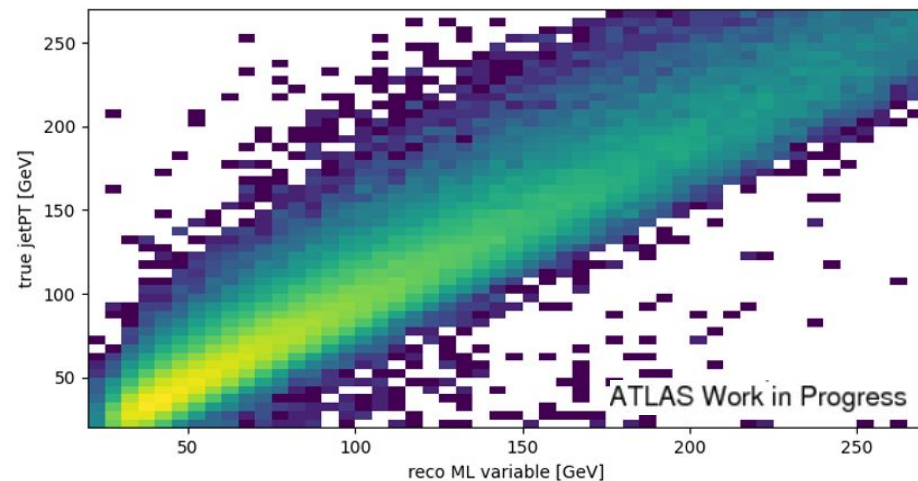
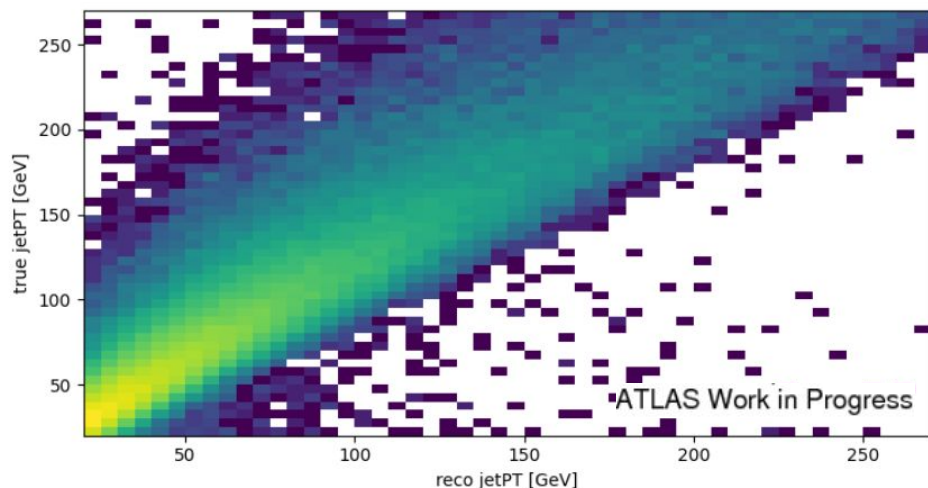
The quantities needed for the unfolding are:

- **Response Matrix:** the ideal case is a diagonal matrix.
- **Inefficiencies** (events reconstructed outside the reco phase space)
- **Fakes** (events generated outside the true phase space)



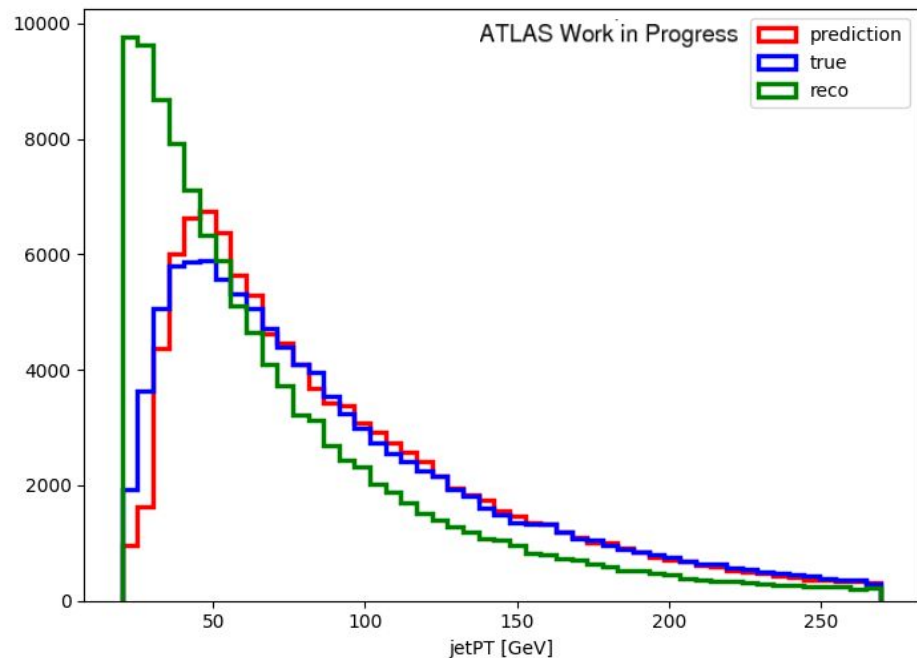
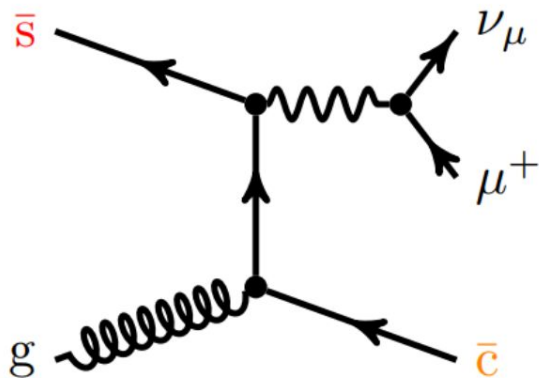
# Observable optimization

It is possible to define a variable at reco level that makes the response matrix more diagonal using Machine Learning (ML) methods.



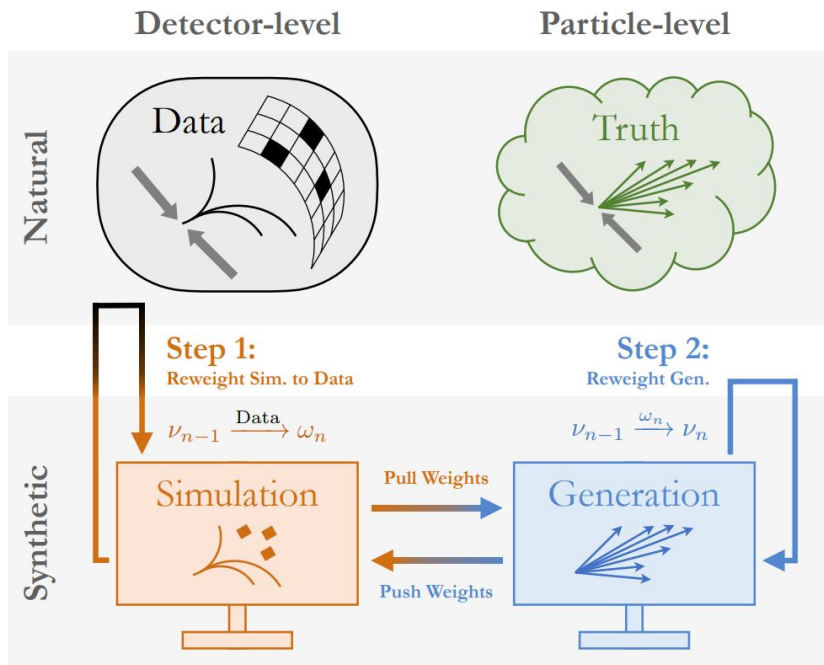
# Application to ATLAS $W+c$ analysis

This method has been tested in order to unfold the transverse momentum of the jet in the ATLAS  $W+c$  analysis.



# Omnifold

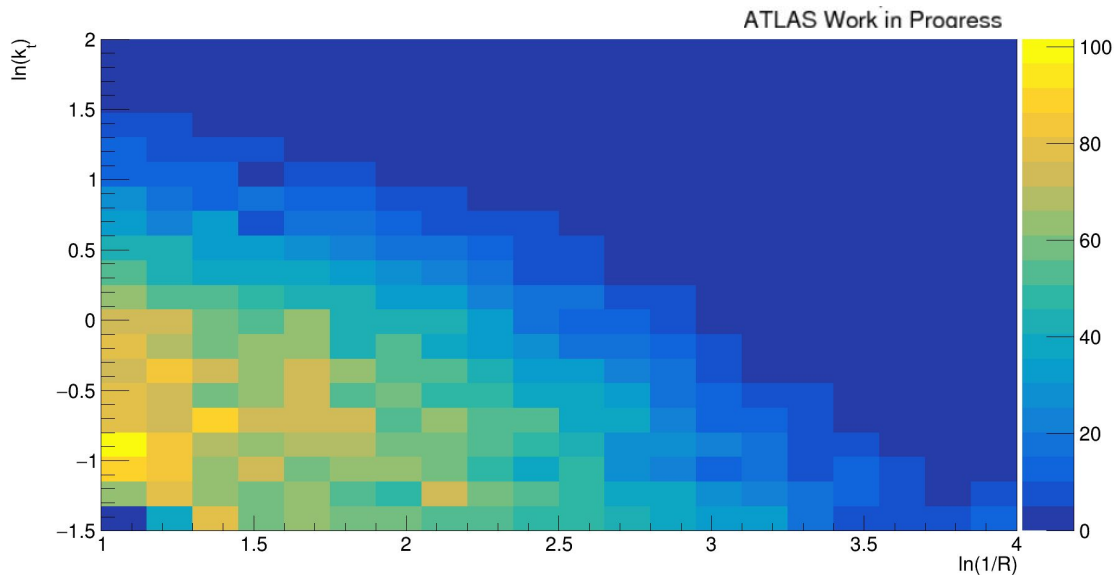
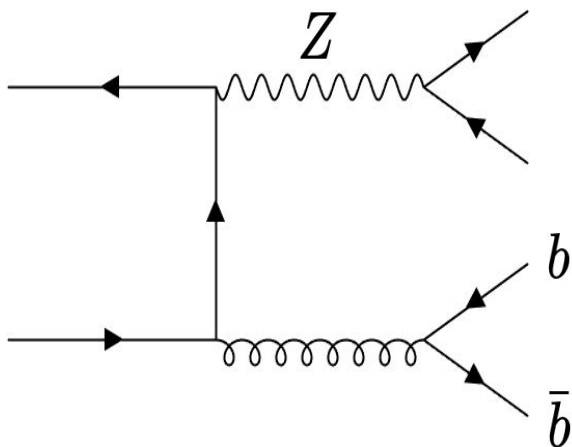
Omnifold is a new ML based unfolding algorithm. Its main advantage is that it unfolds unbinned data.



Omnifold: A Method to Simultaneously Unfold All Observables, A. Andreassen et al., arXiv:1911.09107v2

# Application to ATLAS Z+bb analysis

We are currently comparing Omnifold to IBU in the Z+bb analysis. In this case we are unfolding a **2D variable**: the Lund Jet Plane.



## Summary

New ML based approaches for unfolding have been proposed.  
They need to be tested in a real analysis in order to be correctly evaluated.

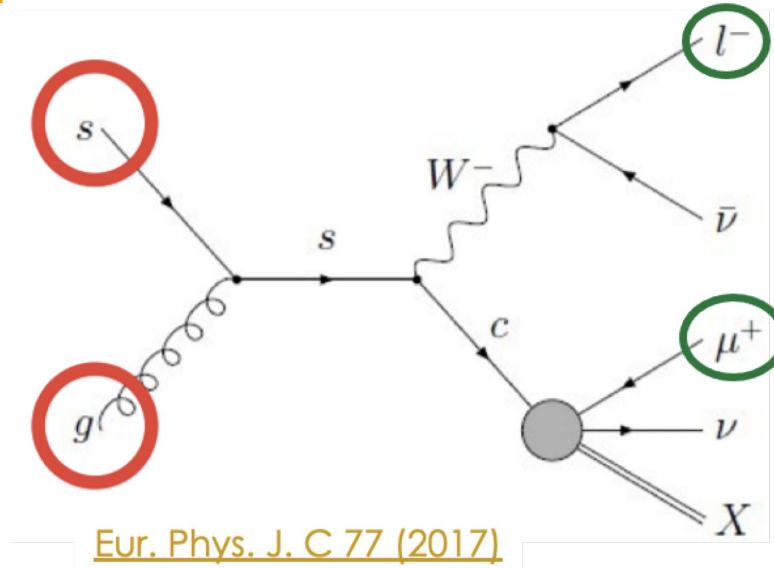
In my research in the ATLAS collaboration:

- Observable optimization has been tested in  $W+c$  analysis.
- Working on a comparison between Omnifold and IBU in  $Z+bb$  analysis.

# Backup

---

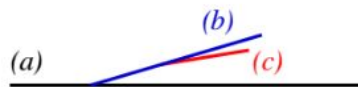
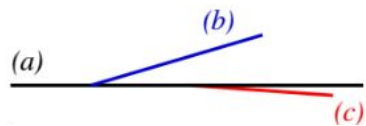
# W+c production



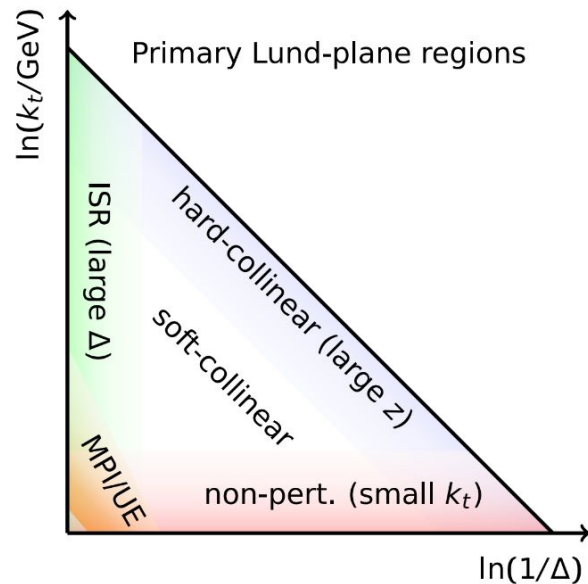
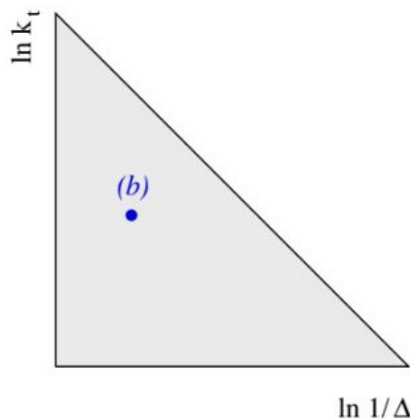
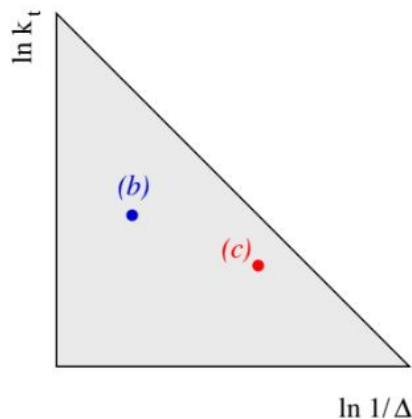
- Main W+c production process sensitive to the s-quark PDF (confirm or not unsuppressed s-quark content at low-x from W,Z 7TeV data)
- Charge correlation between the lepton from W decay and the muon from c-jet exploited
- Soft Muon Tagging(SMT)-algorithm to tag the c-jet, using a muon inside the c-jet

# The Lund Jet Plane

JET



PRIMARY LUND PLANE



# Omnifold

$$\omega_n(m) = \nu_{n-1}^{\text{push}}(m) L[(1, \text{Data}), (\nu_{n-1}^{\text{push}}, \text{Sim.})](m),$$

$$\nu_n(t) = \nu_{n-1}(t) L[(\omega_n^{\text{pull}}, \text{Gen.}), (\nu_{n-1}, \text{Gen.})](t).$$

$$L[(w, X), (w', X')](x) = \frac{p(w, X)(x)}{p(w', X')(x)}$$

# Observable optimization

$$\frac{stat_{ML} - stat_{p_T}}{stat_{p_T}}$$

