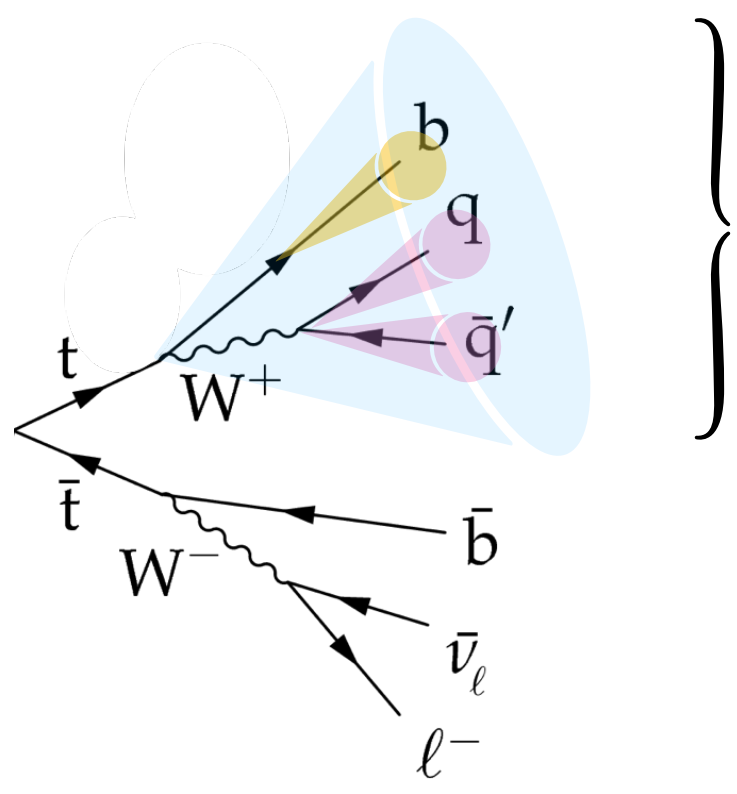


How to Unfold Top Decays

Precisely & Without Bias

Physics problem – boosted top decays

$$p_{T,J} > 400 \text{ GeV}$$



Reconstruct triple jet mass M_{jjj} to measure m_t

Previously done in CMS with TUnfold (classical binned unfolding algorithm, in 5 bins)

CMS 2211.01456

BUT leading uncertainty: choice of m_t in simulation + no access to full phase space

→ **Could generative unfolding help?**

Challenging aspects of top-unfolding & solutions

Multiresonant phase space & combinatorics

- Triple-jet mass (t -decay) & di-jet masses (W -decay)

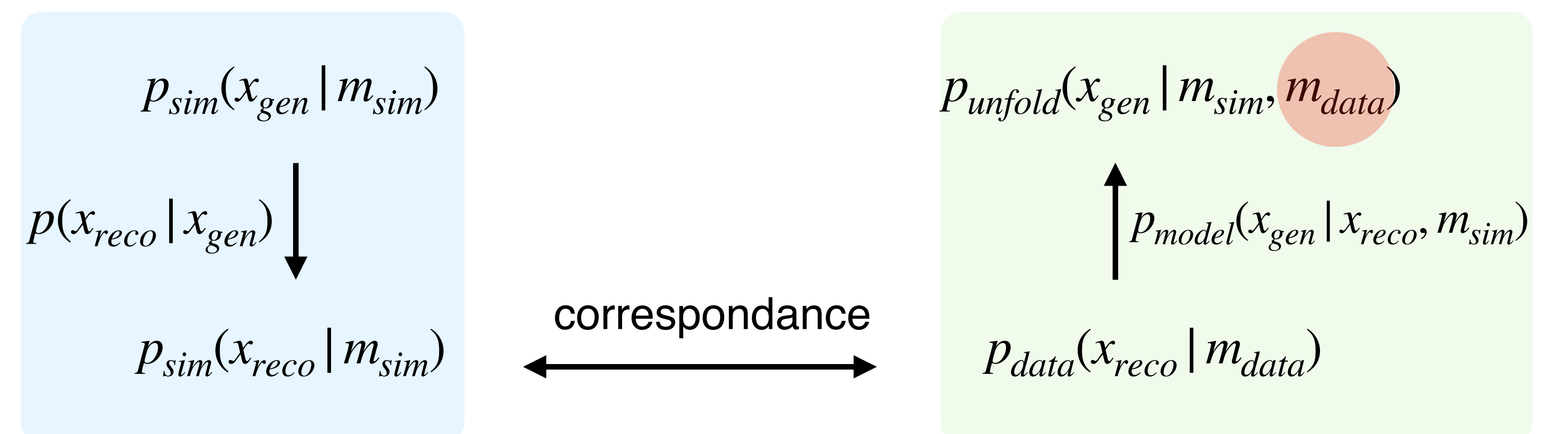
- Jets are sorted according to their p_T

→ any subset of two jet could originate from W -decay

Solution: Choosing a smart parametrisation $p_i = (p_{T,i}, M_{ij}, \eta_i, m_i)$

Detector smearing

- Small local shifts in M_{jjj} (rec) might not result in shifts in M_{jjj} (gen)



Solution: Strengthening m_{data} dependence

Why unfold?

In boosted top regime:

Direct analytical access to m_t after unfolding

For full phase space unfolding:

Direct comparison between theoretical calculations and high-dimensional unfolded data without caring about detector effects

Generative unfolding

- Classical methods are restricted to binned, one-dimensional distributions
- We would like to learn high-dimensional, unbinned unfolding

$$p(x_{gen}) = \int p(x_{reco}) p(x_{gen} | x_{reco}) dx_{reco}$$

target probability

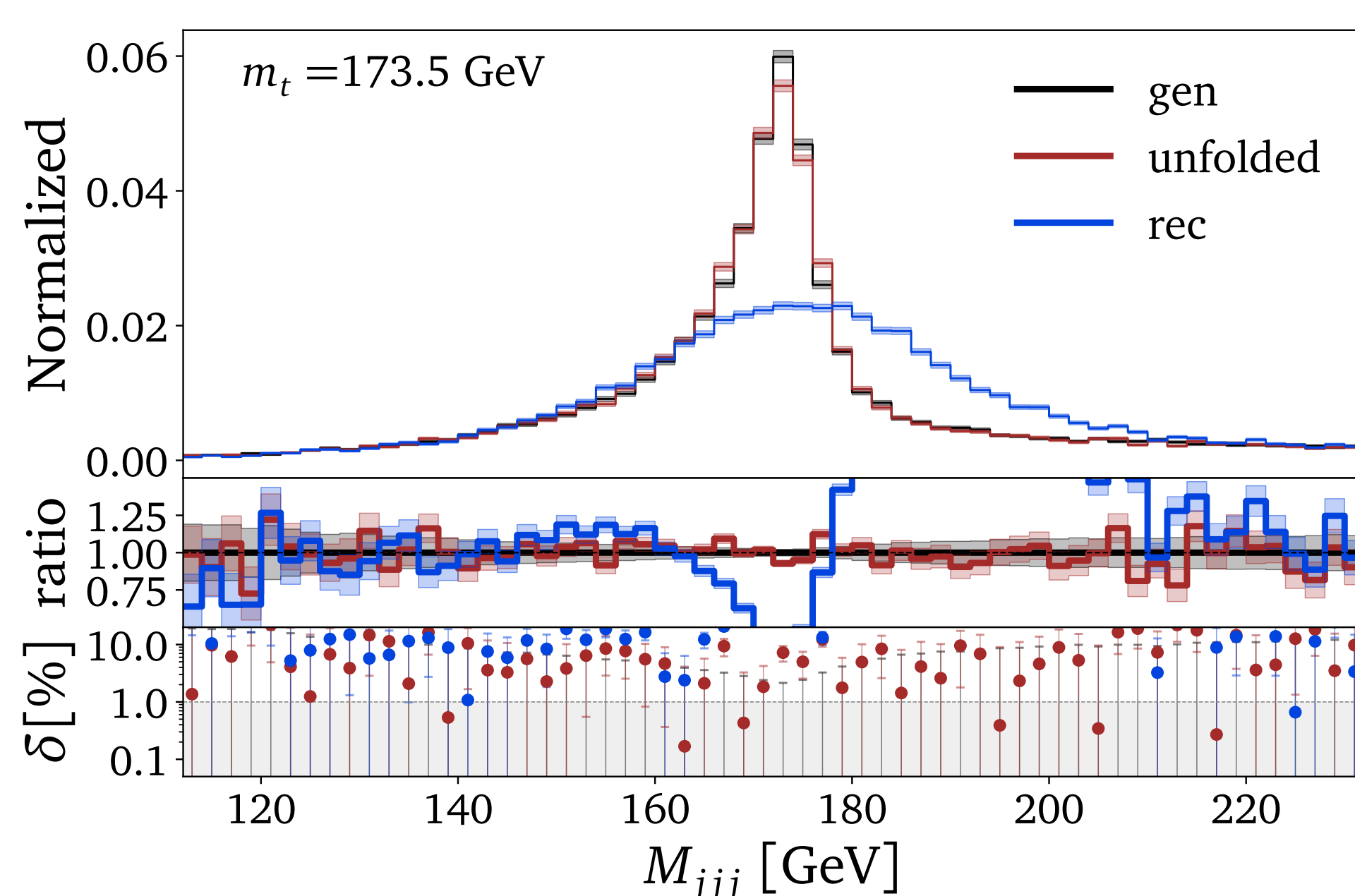
- Use a conditional generative network to sample from $p(x_{gen} | x_{reco})$
- Originally proposed in *Bellagente et al.* 1912.00477, 2006.06685

Results

Train with full CMS simulation with $m_t = [172.5 \text{ GeV}, 169.5 \text{ GeV}, 175.5 \text{ GeV}]$

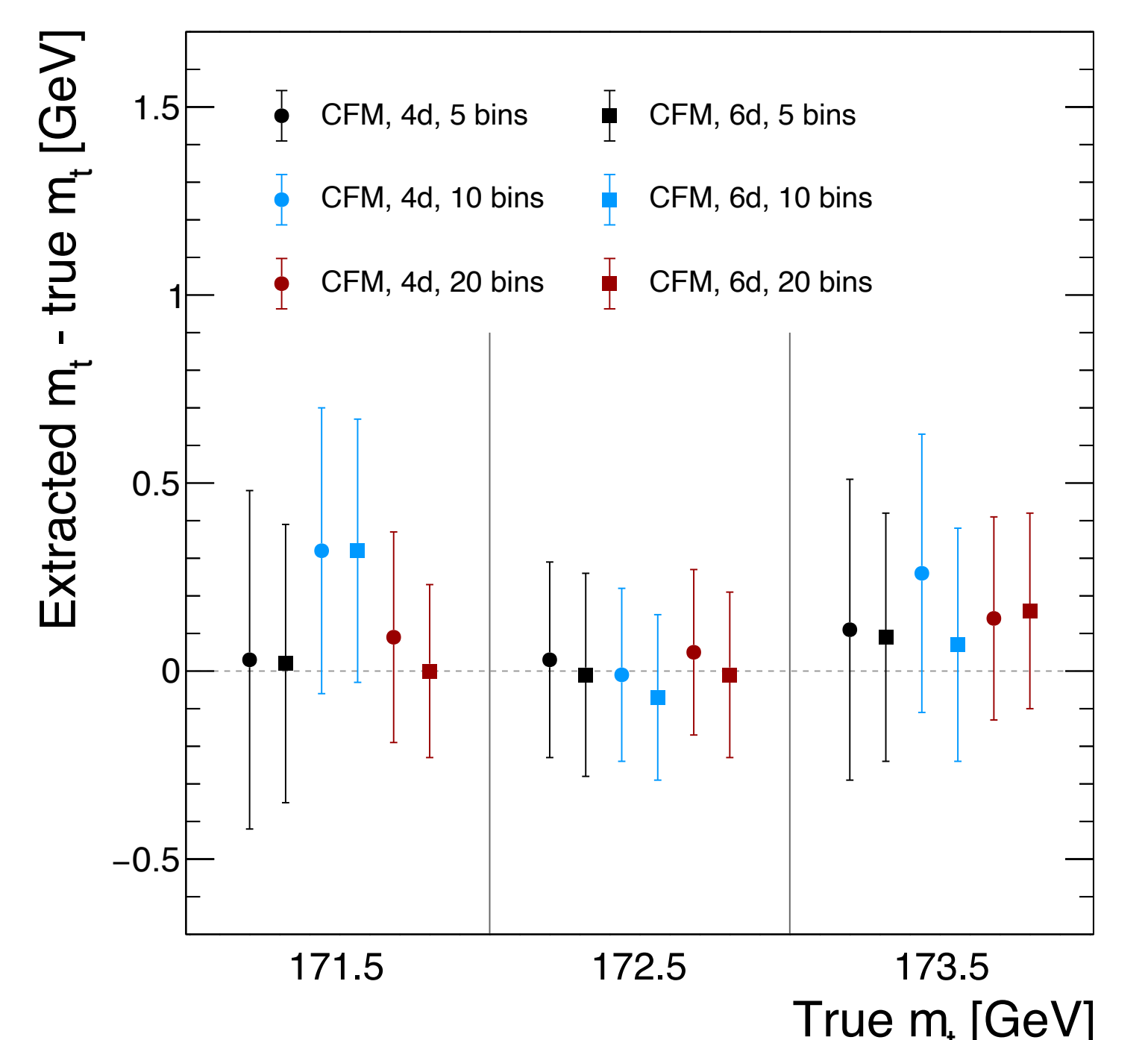
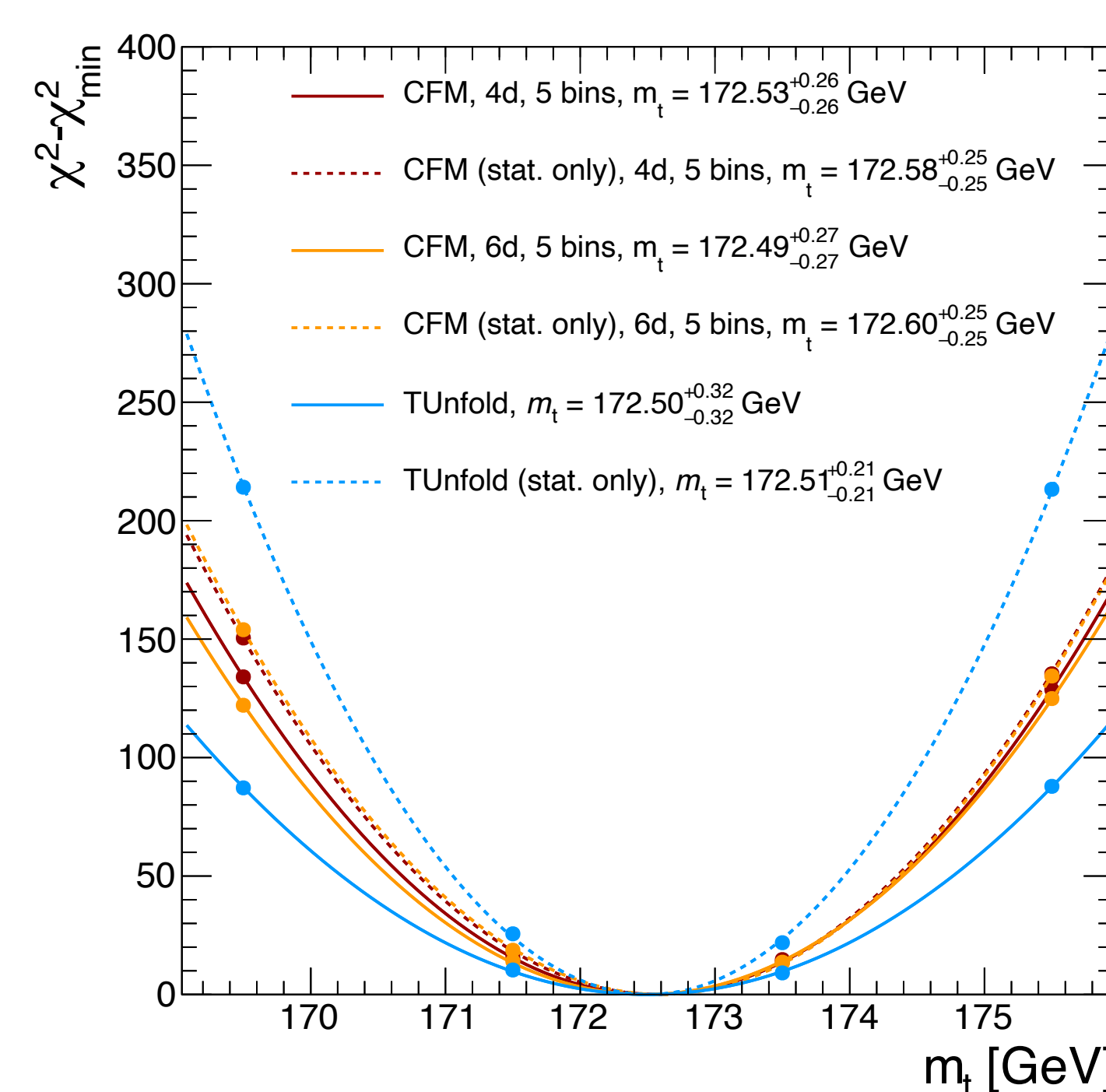
Test by unfolding simulation with $m_t = 171.5 \text{ GeV} \ \& \ 173.5 \text{ GeV}$

Unfolded test distribution of triple jet mass within $\mathcal{O}(1\%)$ of truth gen level **without bias**



For a fixed top mass:

- Choose subset of 41000 reco level events & unfold 1000 bootstrapped replicas
- Estimate covariance matrix and mean from 1000 different unfolded distributions



- Leading systematic uncertainty **reduced by ~80%**
- Statistical uncertainty for 60 bins would be **decreased by 36%**

Proposed analysis pipeline: Observe – Calibrate – Unfold subspace – Measure – Resimulate with measured m_t – Unfold full space – Correct remaining MC/data difference according to *Backes et al.* 2212.08674

SPONSORED BY THE



Federal Ministry of Education and Research

Luigi Favaro¹, Roman Kogler², Alexander Paasch³, Sofia Palacios Schweitzer¹, Tilman Plehn¹, Dennis Schwarz⁴

1 - Institut für theoretische Physik, Universität Heidelberg

2 - Deutsches Elektronen-Synchrotron DESY, Hamburg

3 - Institut für Experimentalphysik, Universität Hamburg

4 - Institut für Hochenergiephysik, Österreichische Akademie der Wissenschaft, Wien



UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386