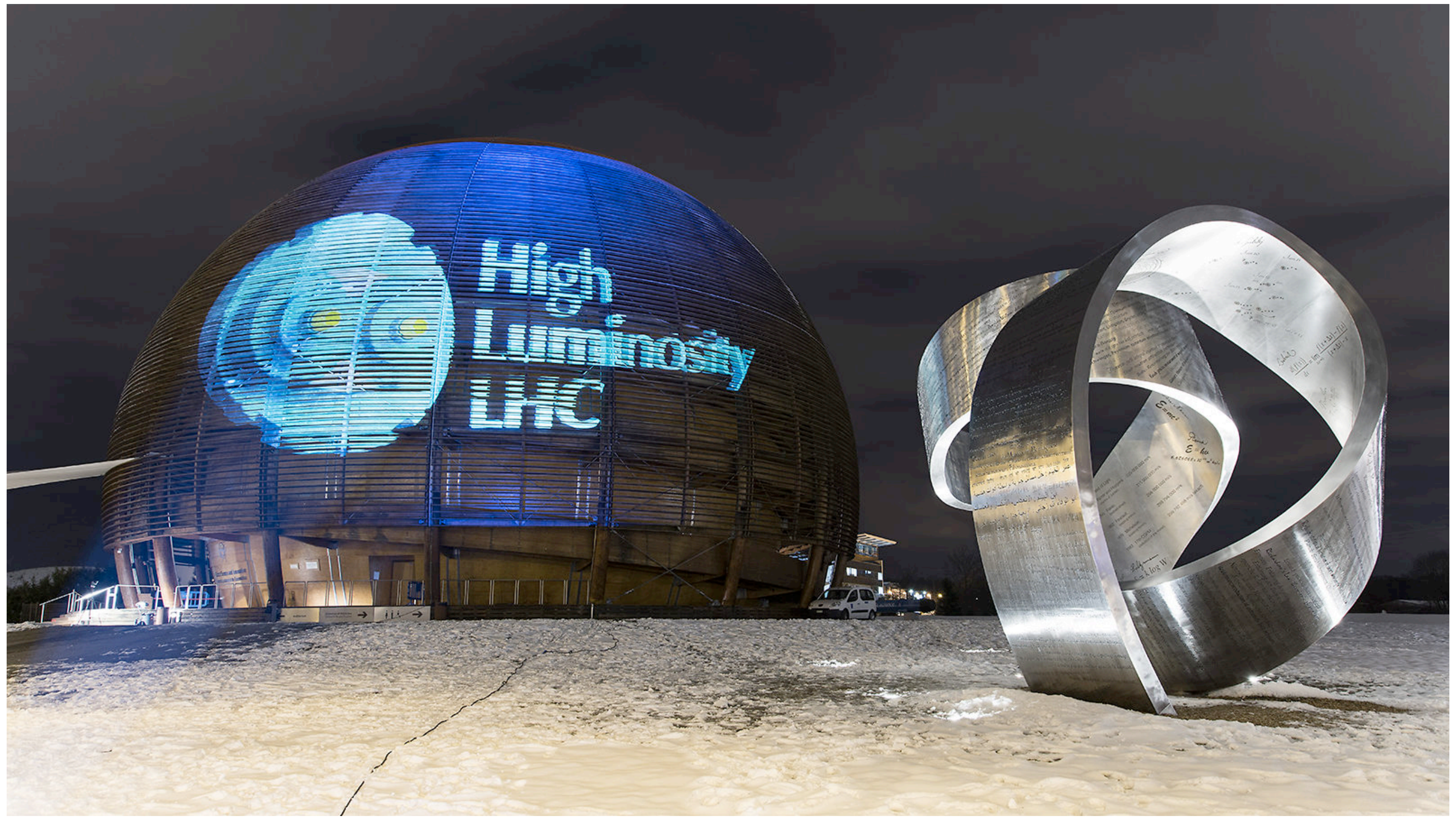
Evolution of analysis techniques and prospects for High-Luminosity LHC in ATLAS and CMS

A biased selection…

16th International Workshop on Top Quark Physics

Traverse City, Michigan (US)
25-29 September 2023

Matteo M. Defranchis (CERN)



What to expect from this talk

-> What this talk is not

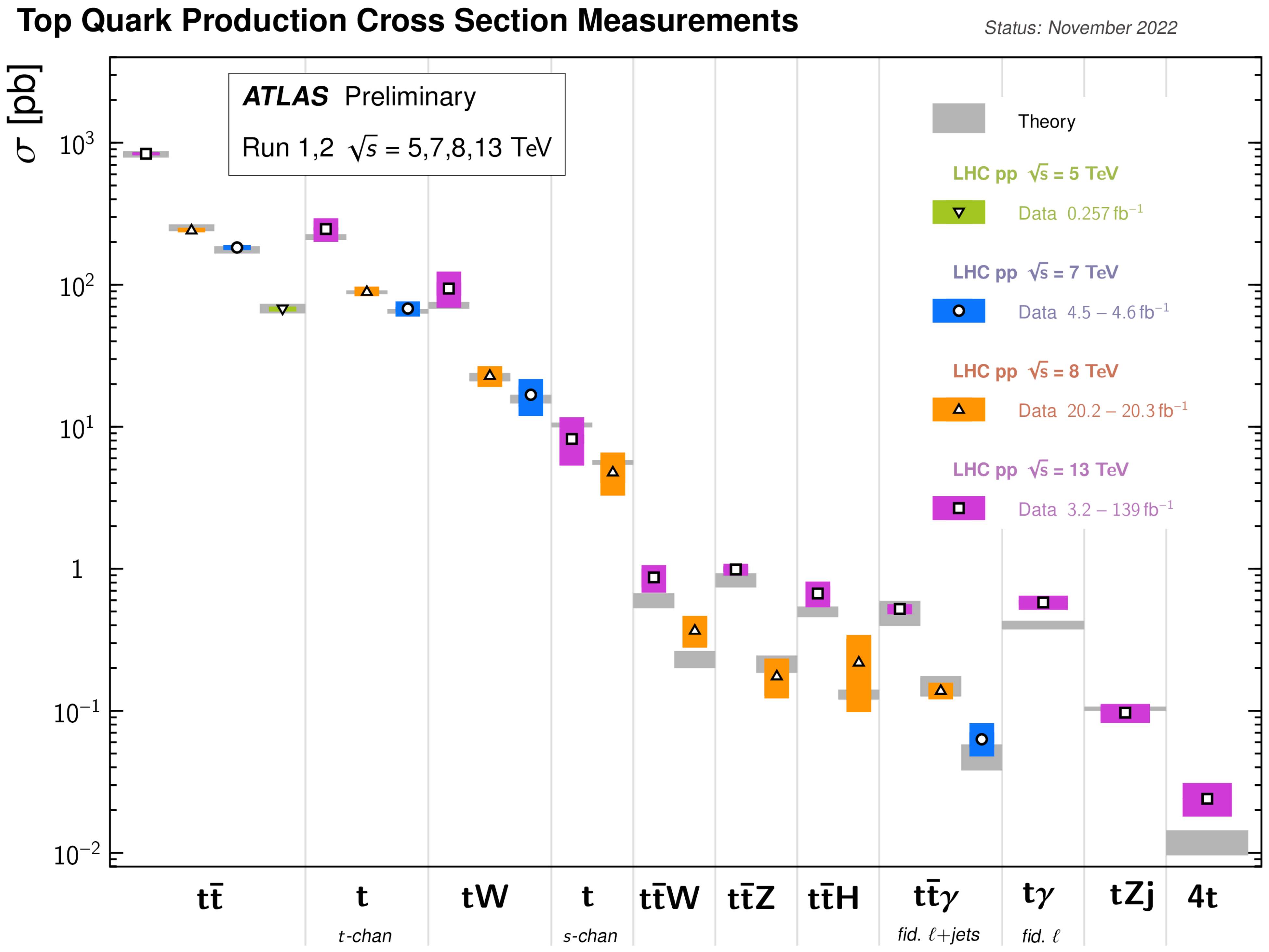
- A collection of HL-LHC projections for various quantities
- An exhaustive discussion of the full Top physics program at HL-LHC

-> What this talk is

- A non-exhaustive personal selection of topics
- An attempt to identify pitfalls and strategies

-> What should be in this talk but it's not

- Searches in the top sector
- A discussion on technical tools, like FastSim or Machine Learning



Would deserve one talk per process

HL-LHC is a long way

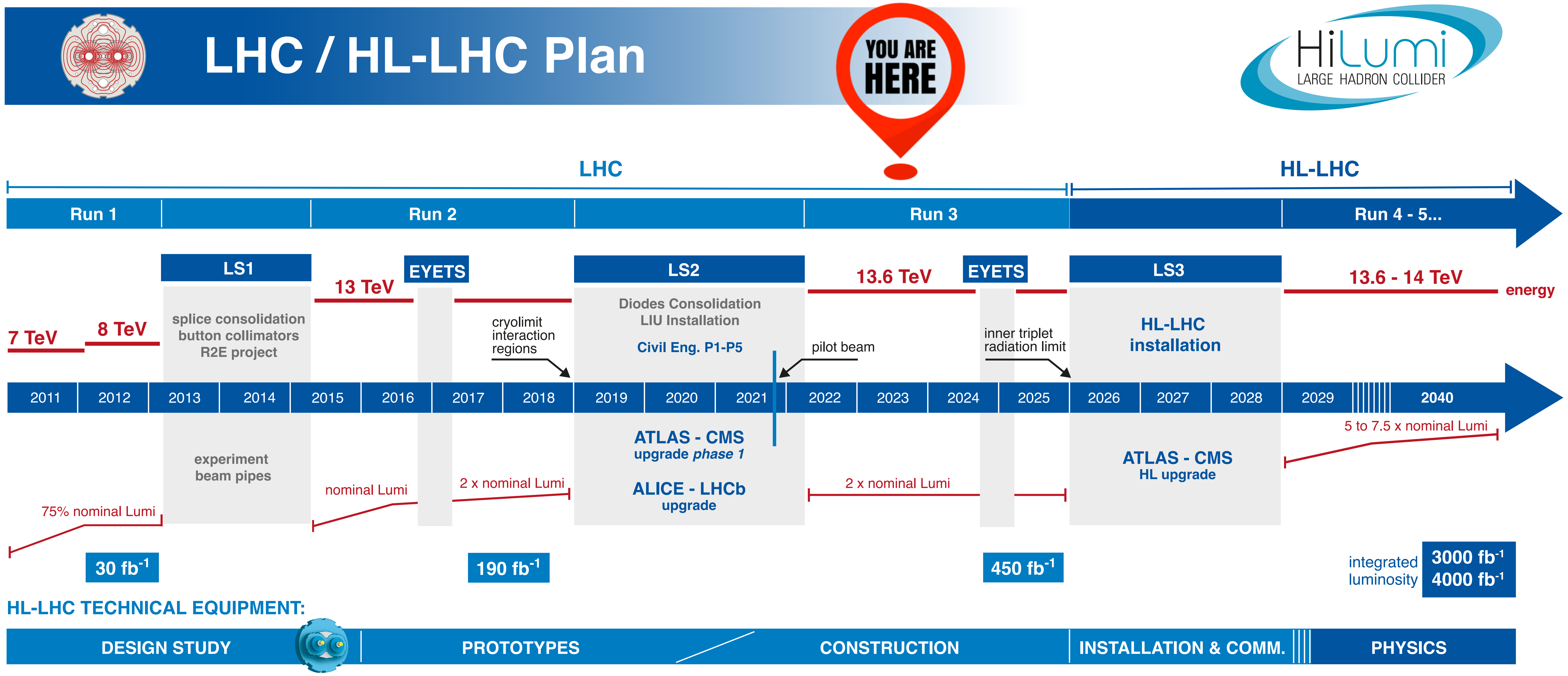
- 20 times more data
- About 20 years time from now
- Major detector upgrades

-> With more data:

1. Rare processes
2. (Multi-) differential measurements
3. Explore corners of phase space

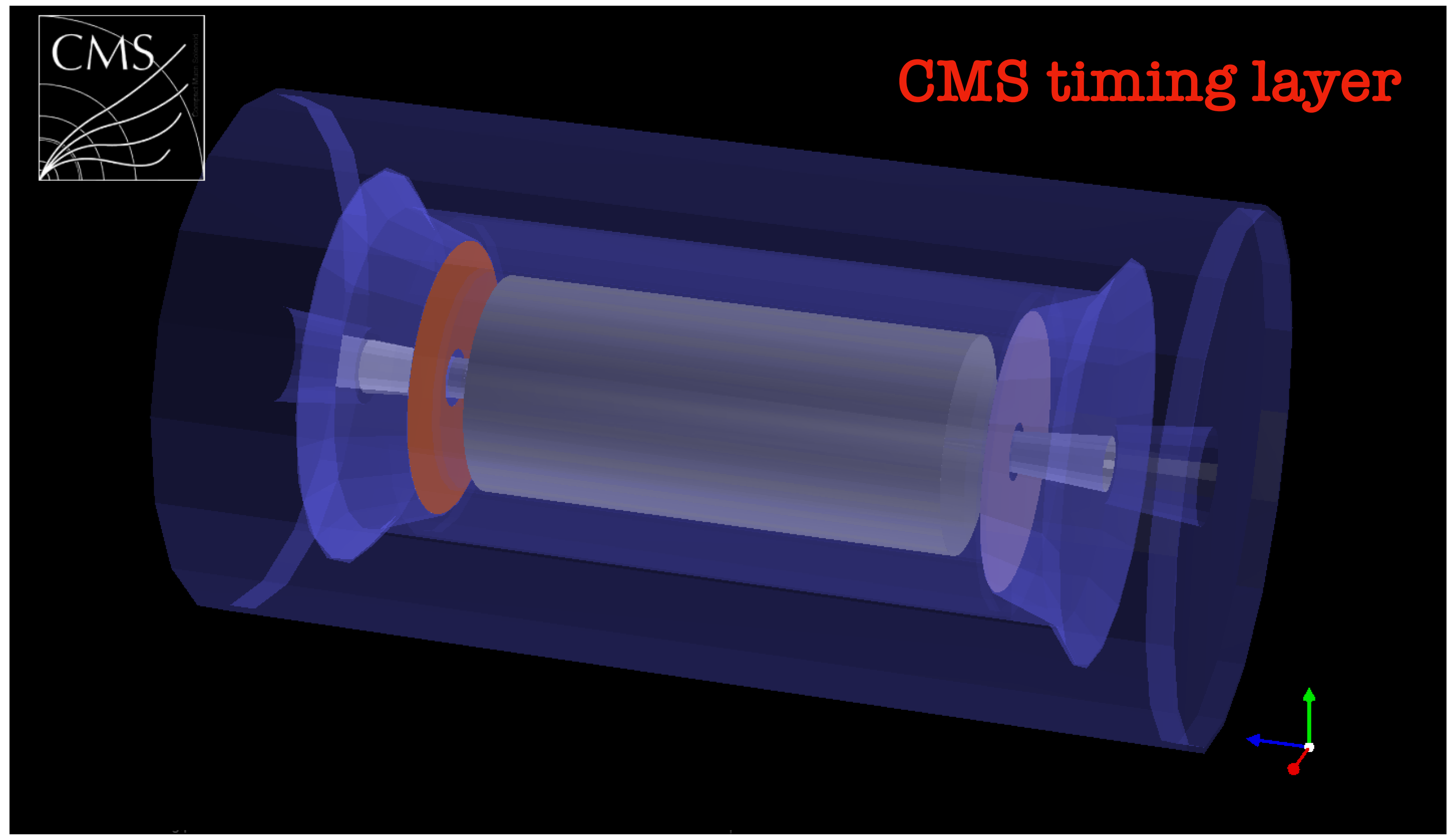
-> With more time:

1. More powerful analysis techniques
2. More accurate theoretical tools
3. Other "technological" breakthroughs (computing, AI, ...)
4. New ideas

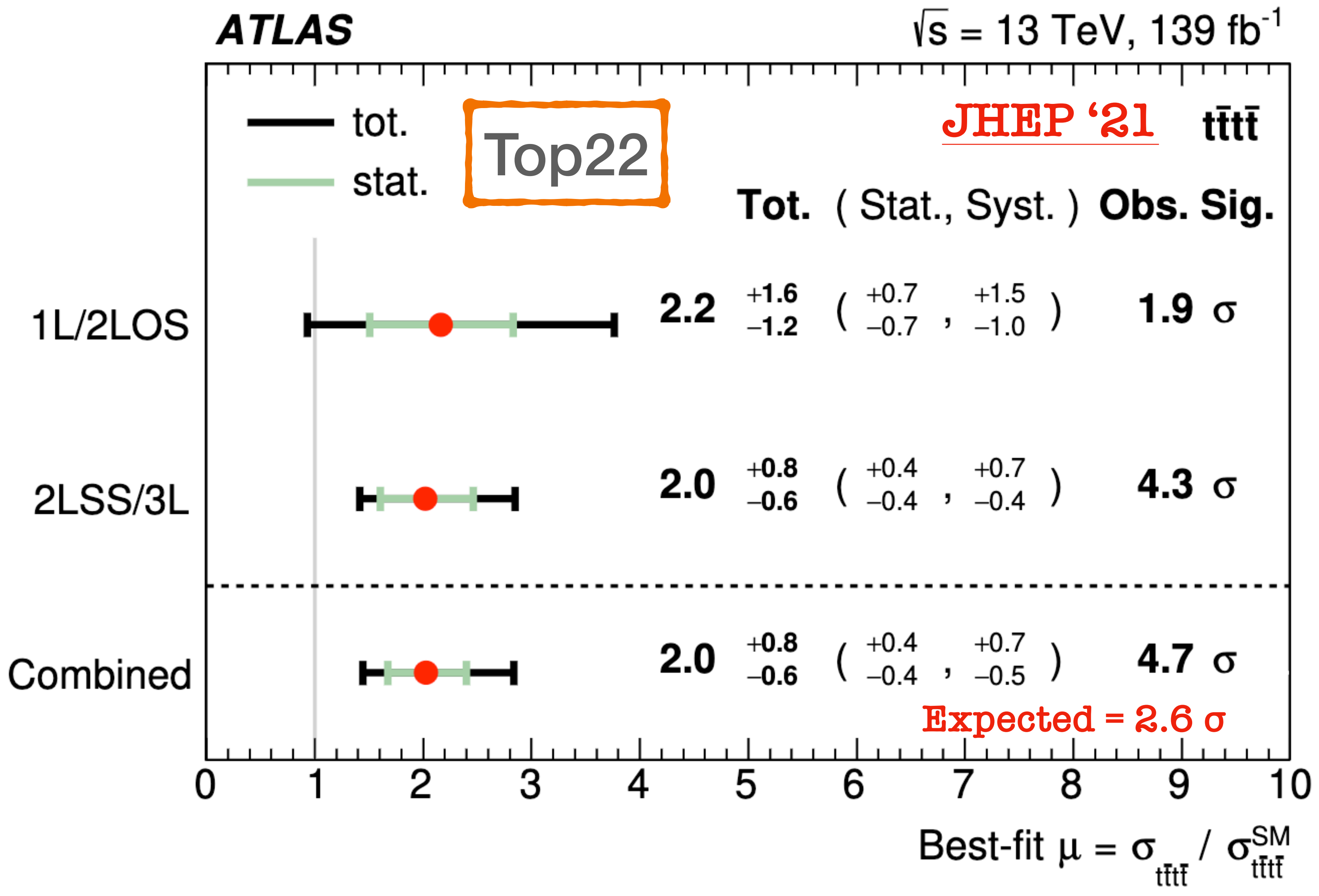


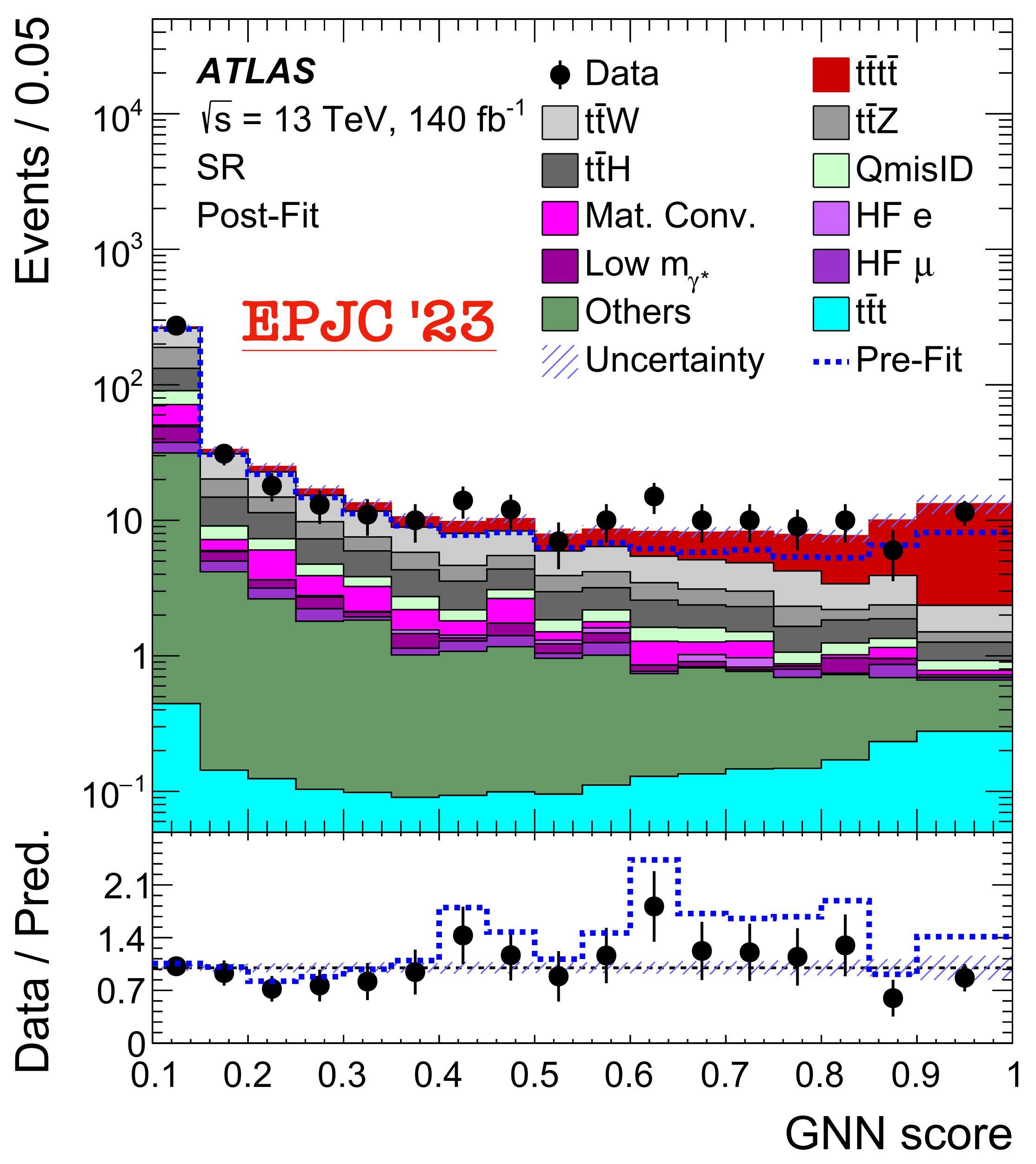
Technical advancements are hard to predict

-> expectations are often exceeded



Observation of 4t production





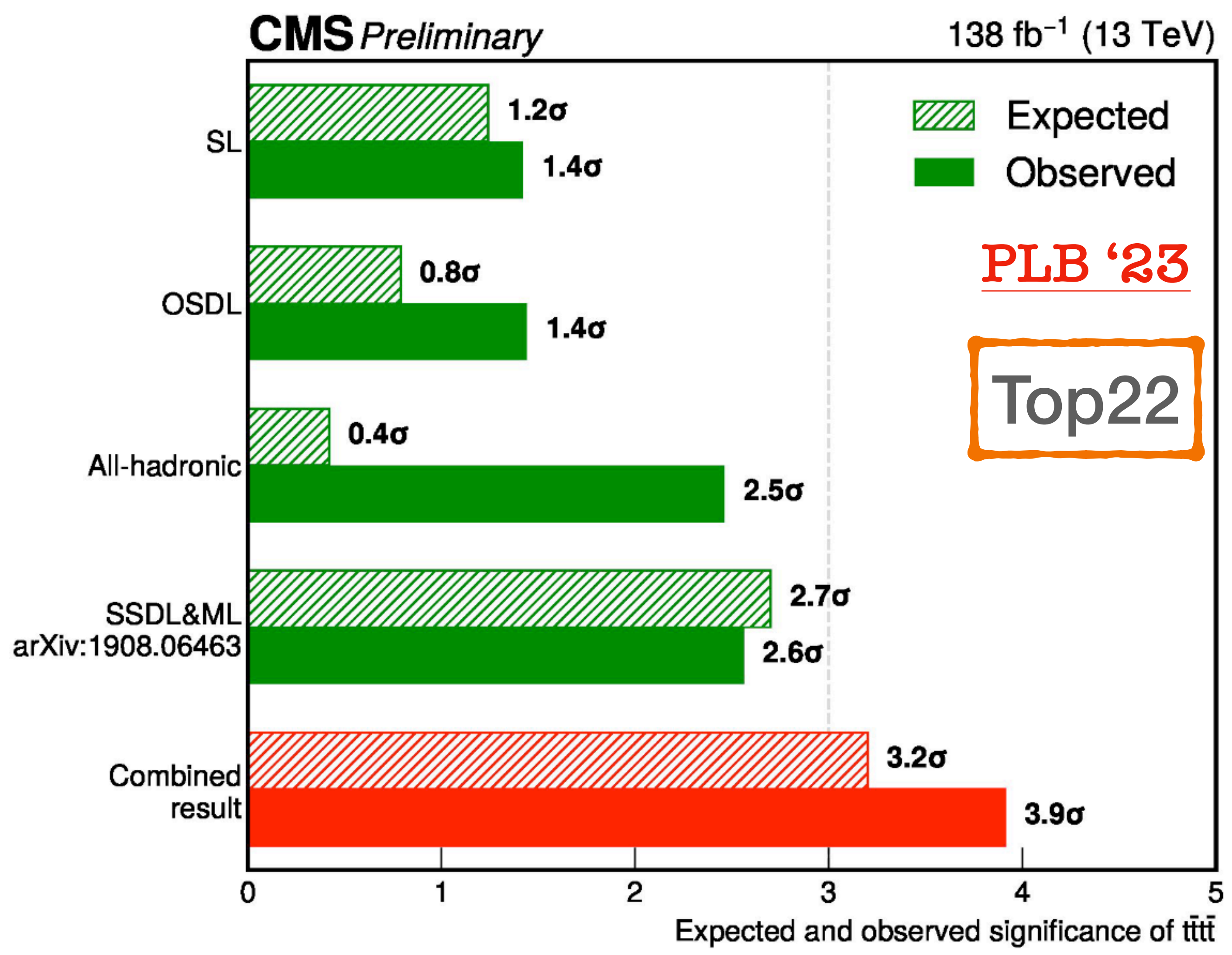
Moriond23: ATLAS & CMS observation on same dataset

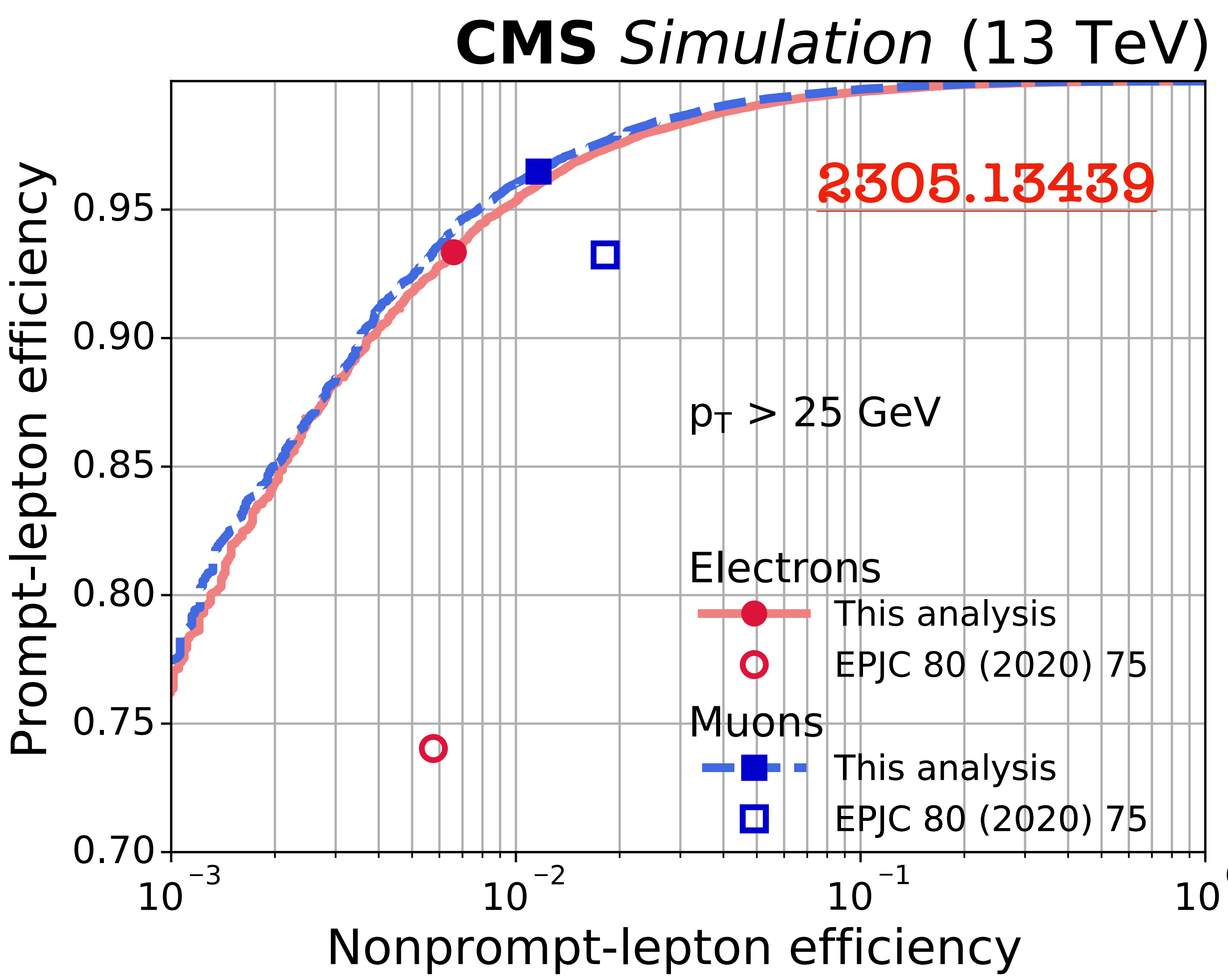
ATLAS

- GNN for signal/bkg separation
- Increased signal acceptance
- In-situ calibration of jet multiplicity in ttW
- Improved MC model for 3t background

CMS

- Lepton and b-jet identification
  - Multiple leptons and b-jets
- Complex multi-category fit to signal and background regions

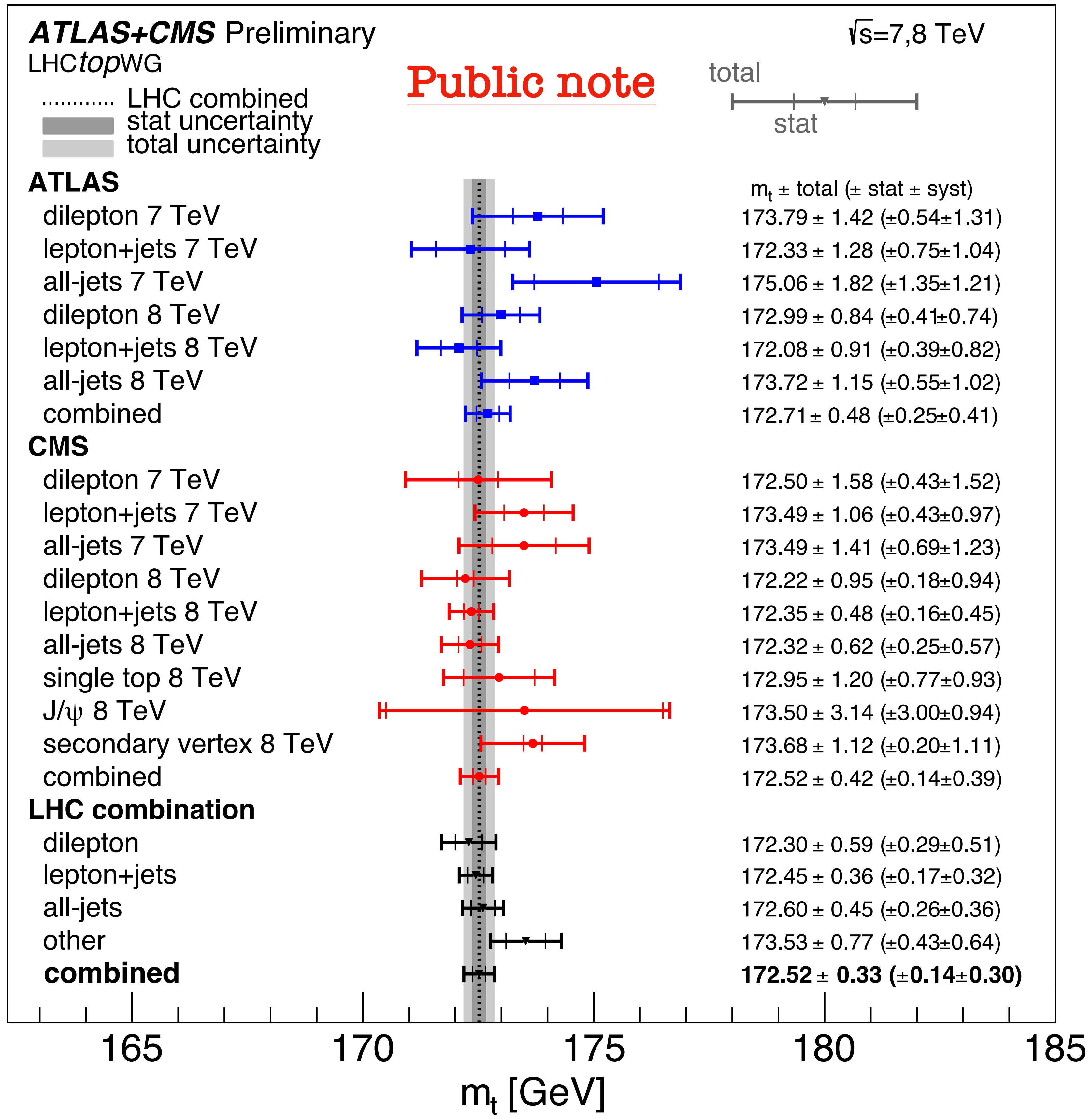


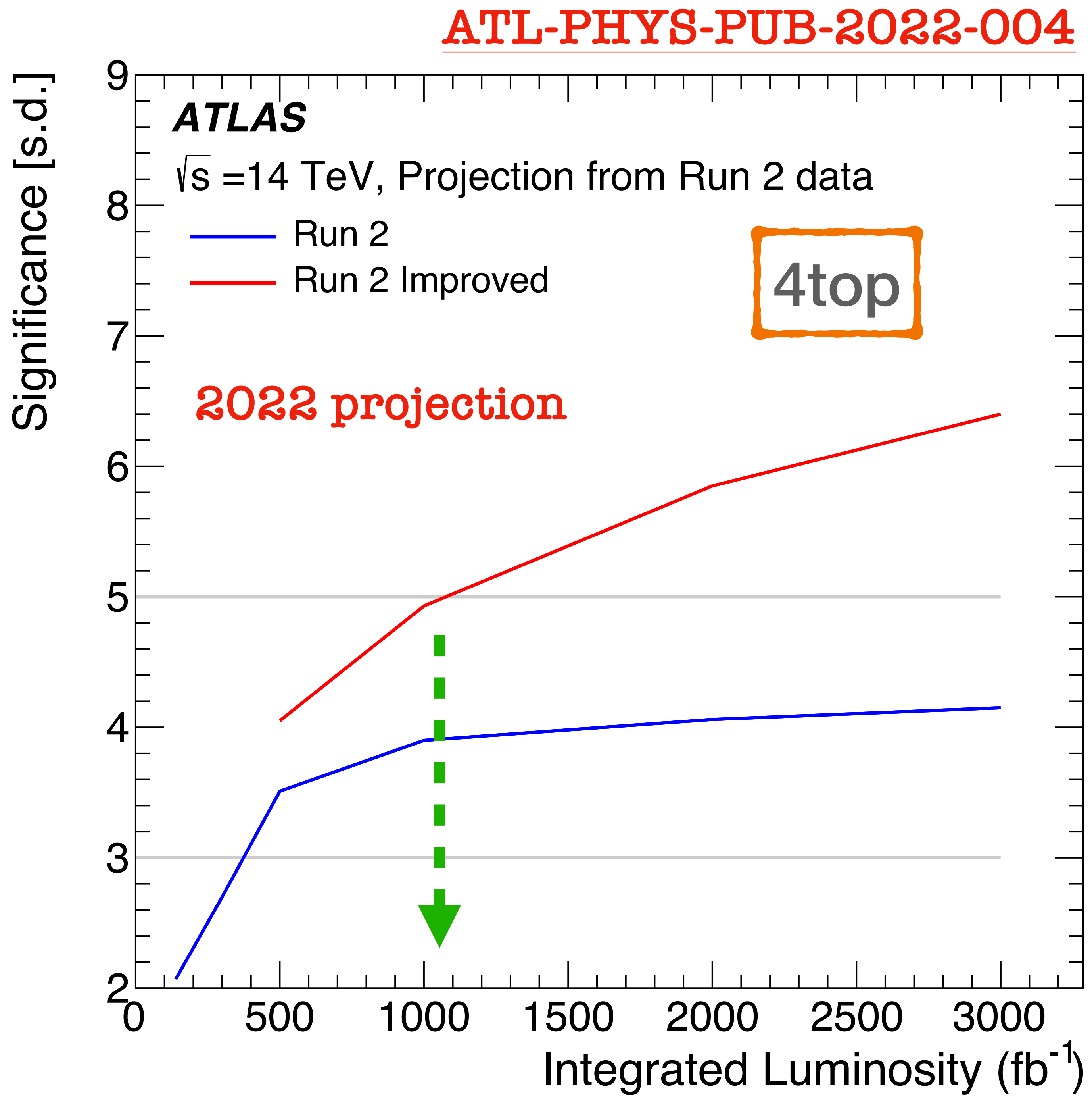


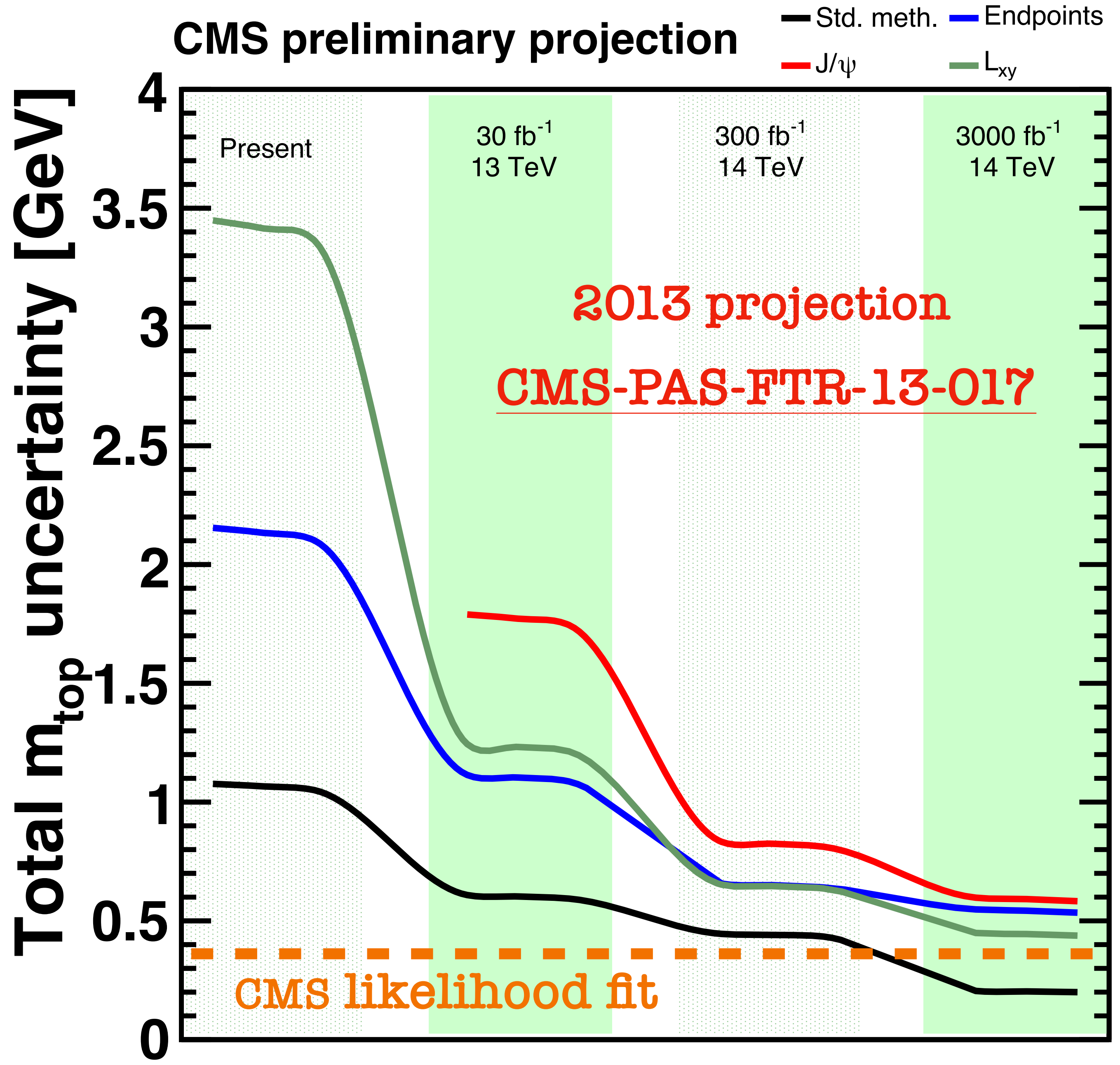
Combination of improvements -> no silver bullet

Projections vs reality

	Expected	Achieved
4t observation	> 1000 fb^{-1}	< 150 fb^{-1}
δm_t < 0.4 GeV	300 fb^{-1}	36 fb^{-1}







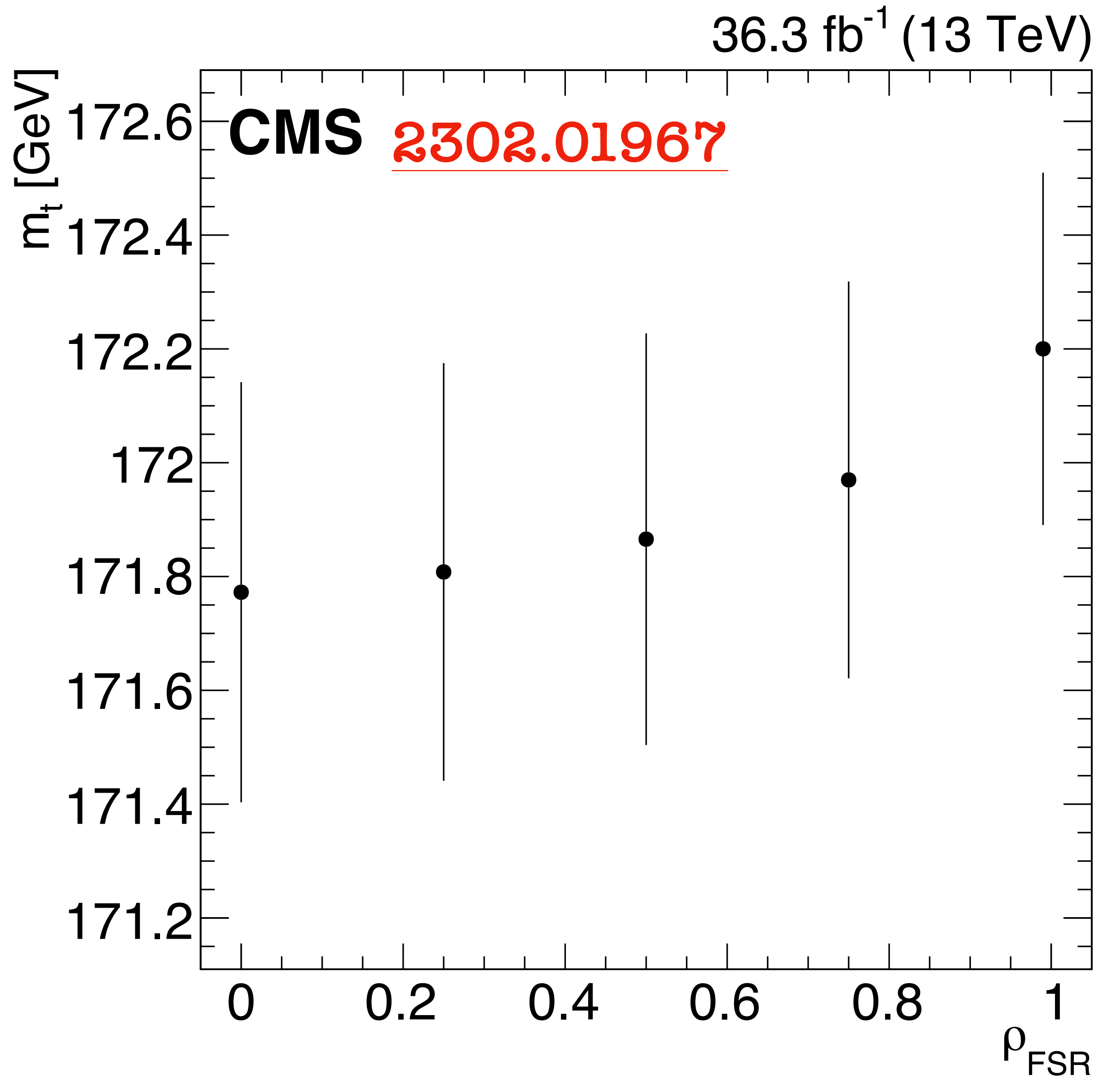
δm_t = **0.33 GeV** -> almost 2x better than best individual input

- Requires good understanding of inter-experiment correlations
- Synchronise on MC model and uncertainties

- **Design analyses with a massive combination effort in mind**

"Free" x2: ATLAS+CMS = 6000 fb^{-1}

Likelihood-based top mass measurements



CMS early 2016 analysis

- Two-observables fit with externalised systematics -> **0.63 GeV**

CMS re-analysis with profiled systematics

- Simultaneous profile-likelihood fit to multiple observables -> **0.37 GeV**

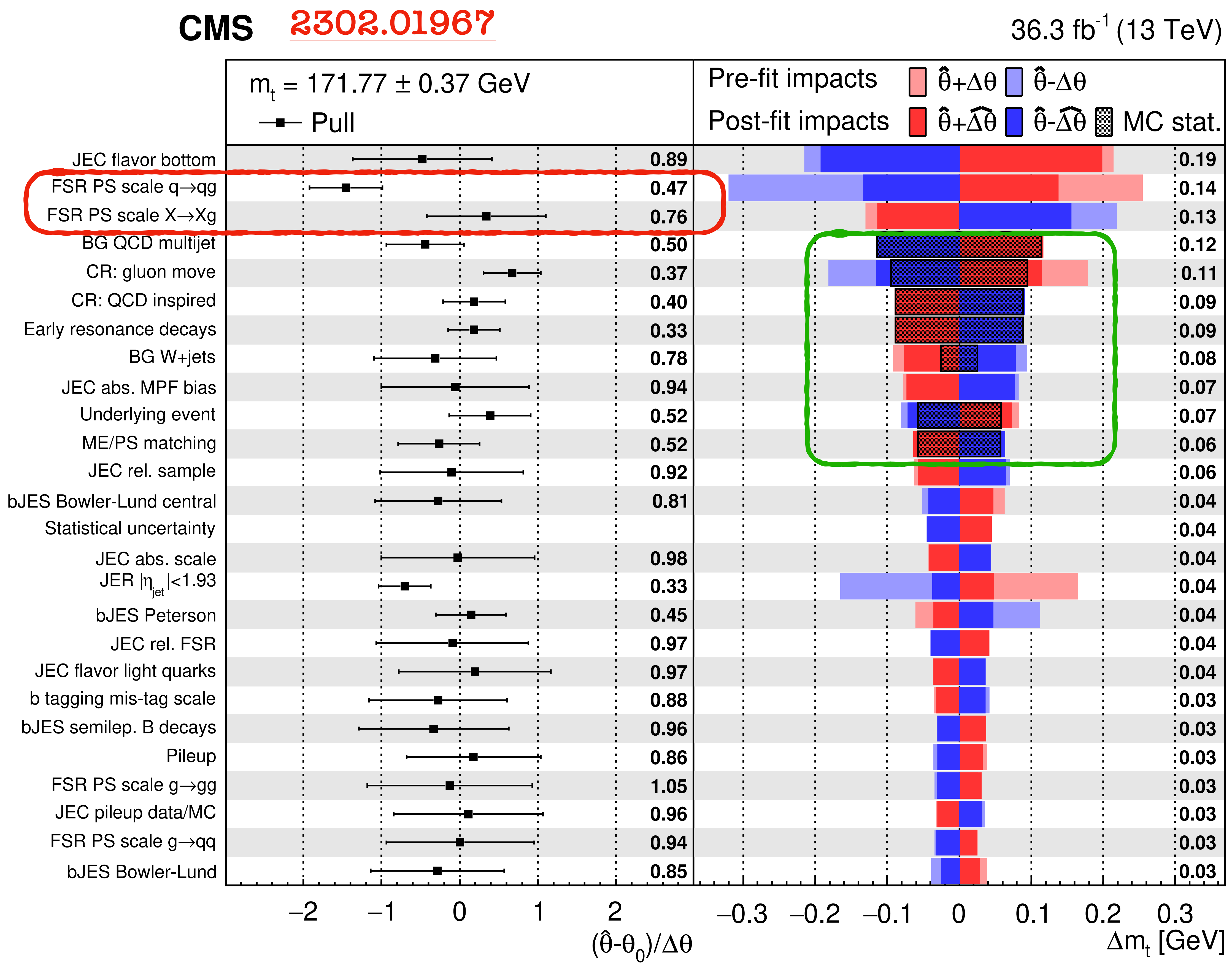
40% improvement with re-analysis of same data set

Parton-shower modelling

- 3-point α_S variation, independently for each splitting kernel
- Studied dependence on correlation assumption

MC statistical uncertainty (independent MC samples)

- Can give rise to unphysical constraints on systematic
- Full treatment of the effect on profiling for the first time



Crucial to move to weight-based variations

-> frontier: DCTR approach with ML-based re-weighting

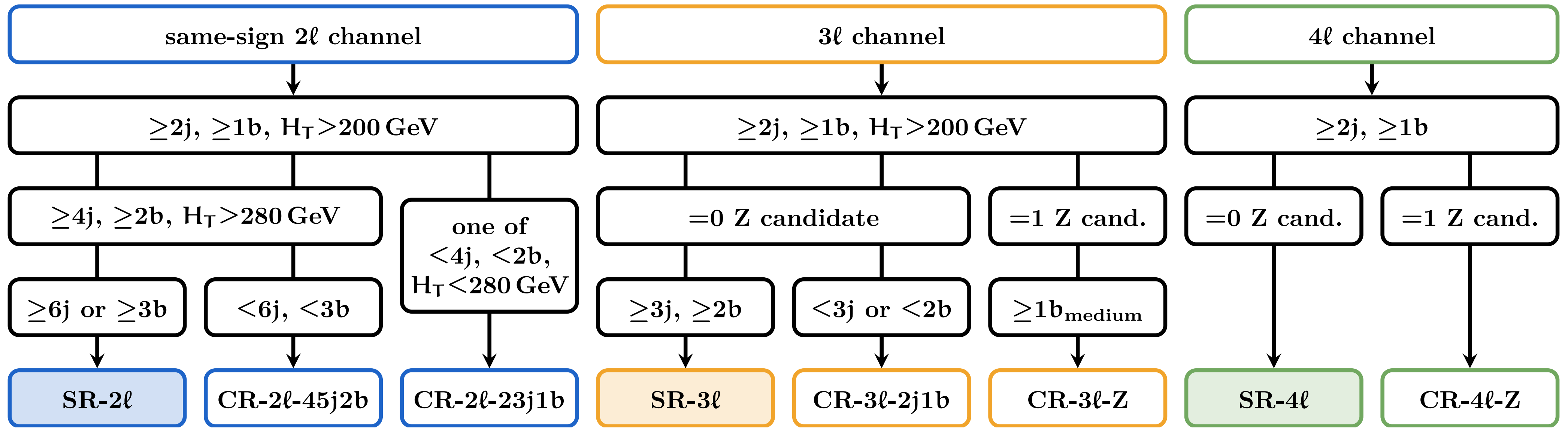
Profiling: advantages, pitfalls, opportunities

Hard to understate the advantages

1. Multi-dimensional, multi-process fits
2. In-situ constraint of systematic uncertainties, backgrounds, ...
3. Reduces bias from choice of nominal model
4. Fully accounts for correlation between processes and systematics

N.B. systematics correlations are fully neglected in standard analyses

2305.13439 CMS 4t



Theory-related uncertainties

a) PDFs, tuned parameters (α_S, h_{damp}) -> statistically meaningful

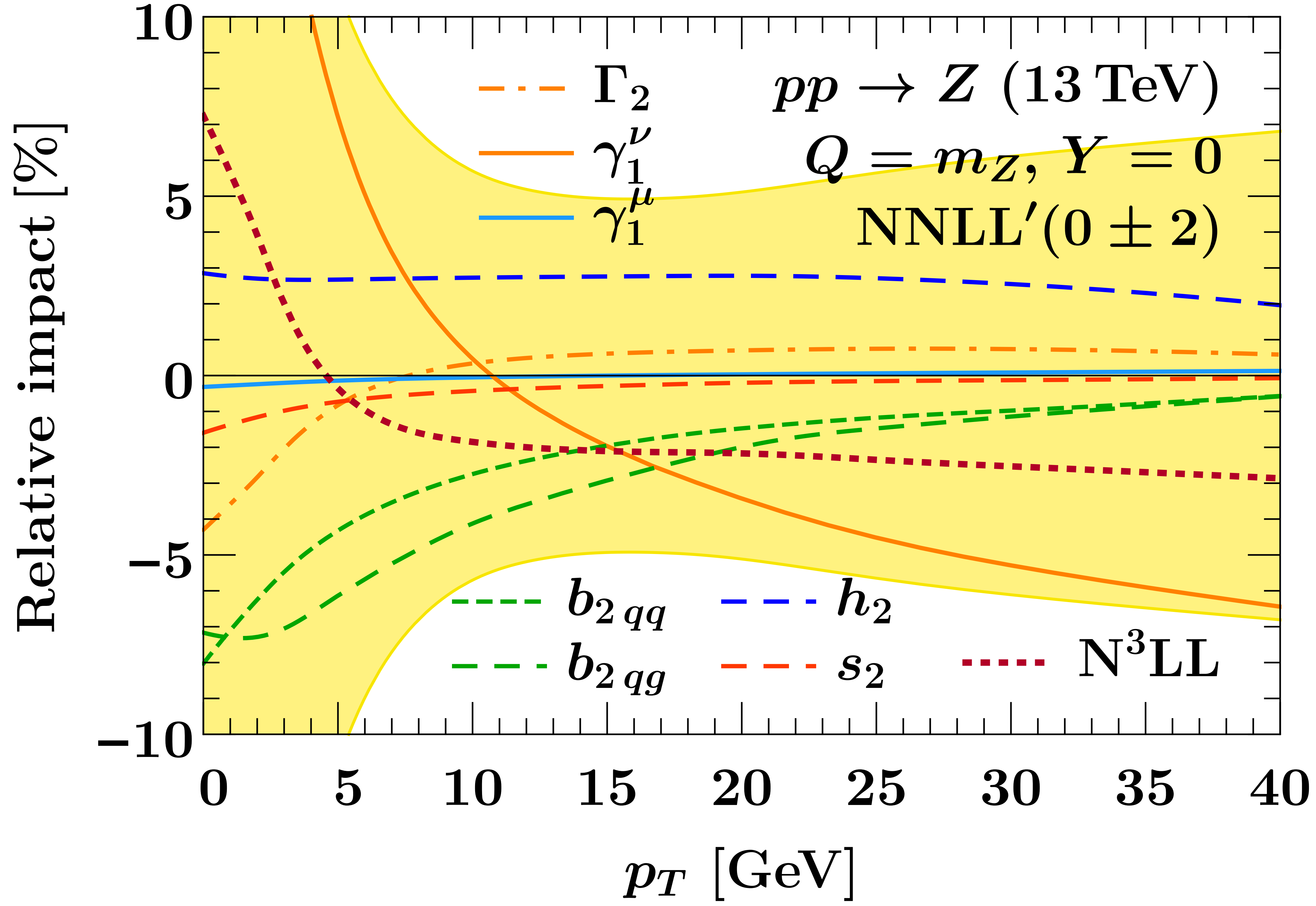
b) 3-point modelling variations (e.g. scales)

- Not very meaningful, but somewhat acceptable

c) 2-point model comparisons

- Not meaningful, but sometimes the only option (CR)

F. Tackmann (BIRS 23)



As the constraining power of data increases, it is a must to move towards statistically meaningful ways of estimating theory uncertainties

$$f(x, p_T) = \exp\left\{\sum_i \left[f_i(0) + f_i'(0)\,x + f_i''(0)\frac{x^2}{2}\right] g_i(p_T)\right\} + \mathcal{O}\left(\frac{p_T^2}{m_Z^2}\right)$$

Novel theory nuisance parameter approach: way forward?

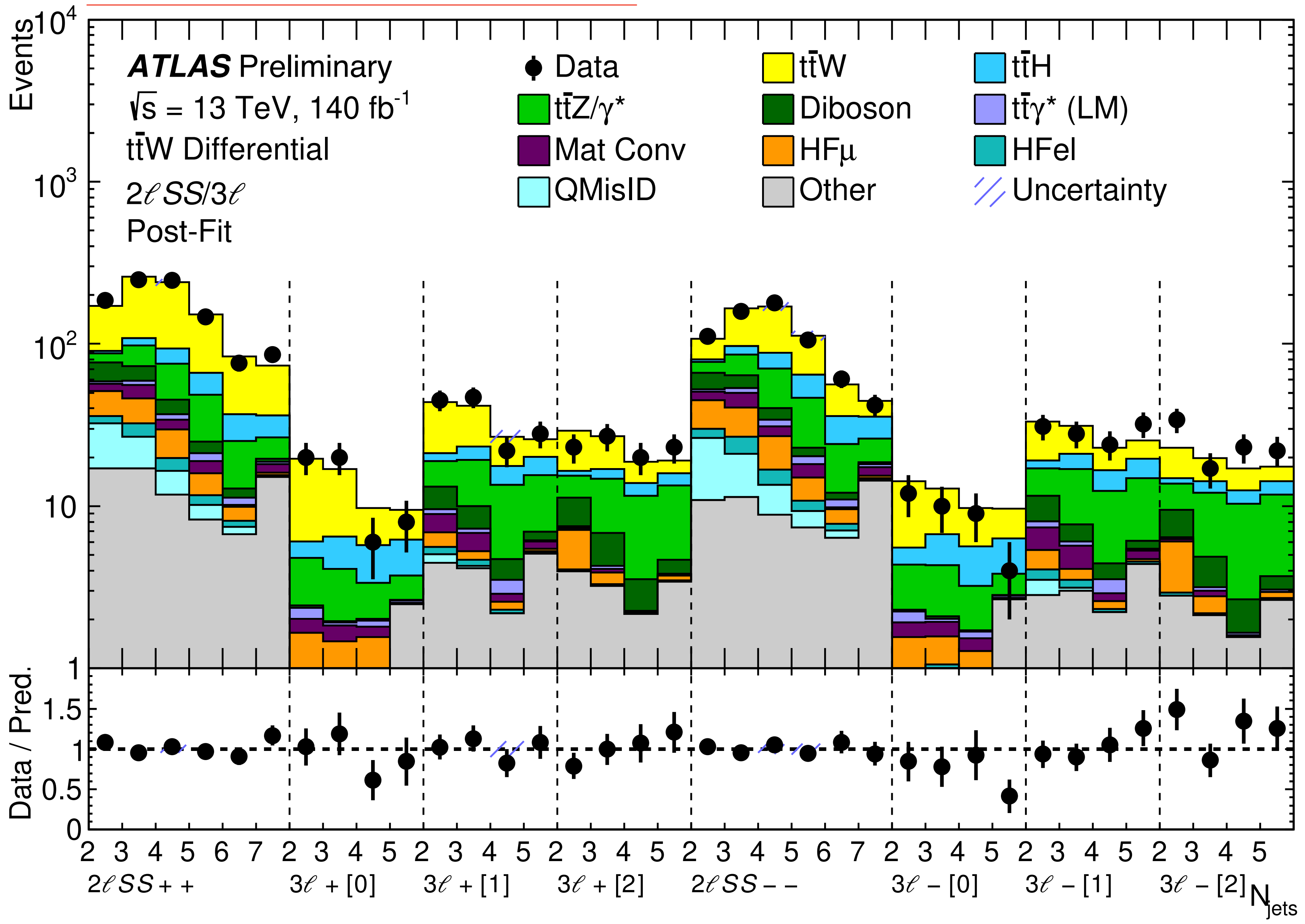
Profile-likelihood unfolding

Used in several differential measurement -> increasingly popular

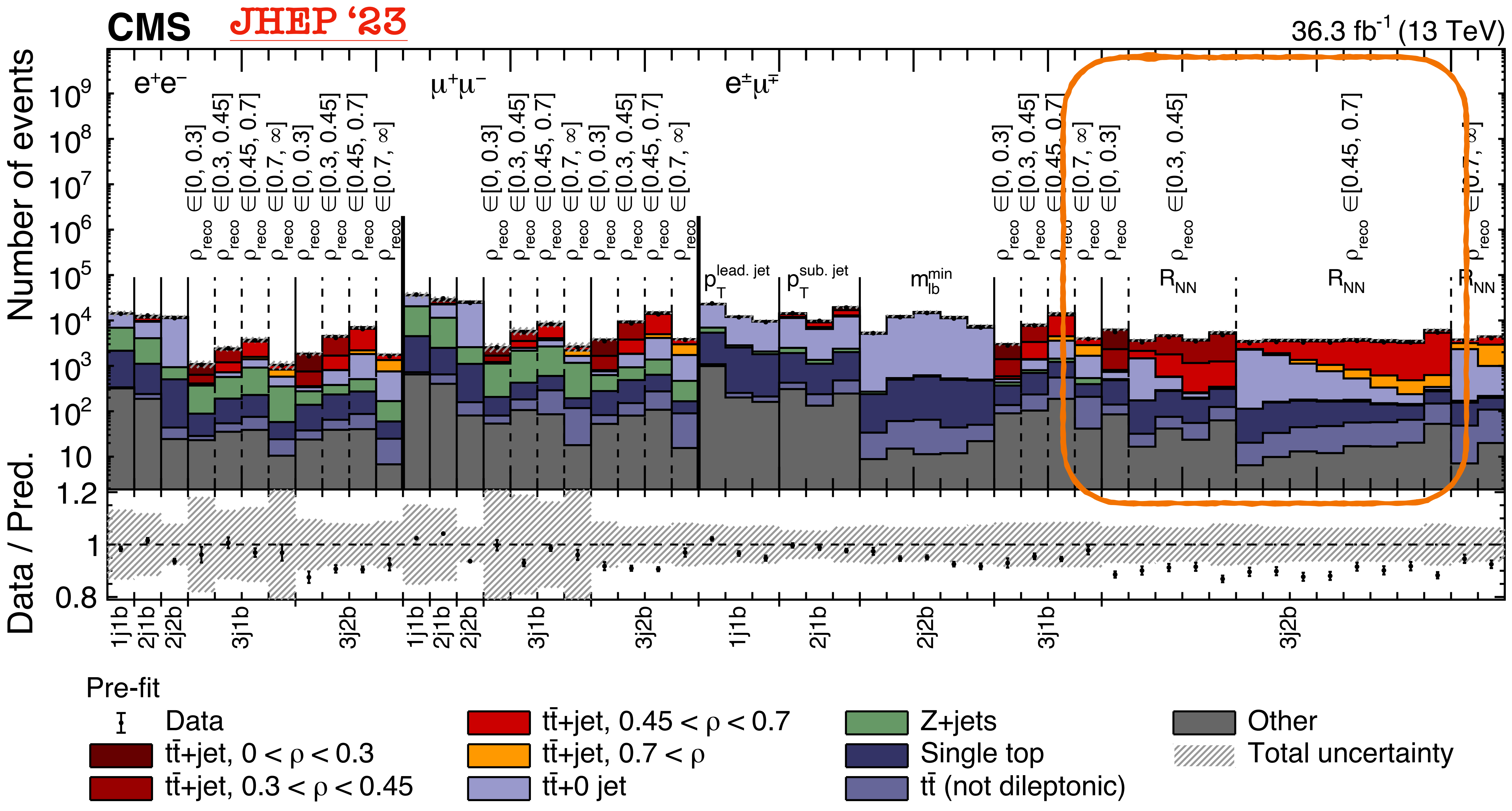
- Multi-category, multi-process fit -> EFT
- Allows to incorporate multi-variate discriminators
- Maximise signal acceptance => minimise extrapolation

Taken one step further, it can be used as a tool to reduce model dependence in parameters extraction in exchange for statistical uncertainty

ATLAS-CONF-2023-019



CMS JHEP '23



Asymptotically theory agnostic m_W (ASIMOW)

$$\frac{\Delta^2\sigma}{\Delta p_T^l\,\Delta\eta^l} = \sum_{\Delta q_T,\Delta|y|} \frac{\Delta^2\sigma_{-1}}{\Delta q_T\,\Delta|y|}\left(T_{-1}(p_T,\eta\mid M_W) + \sum_{i=0\ldots4} A_{i,\Delta q_T,\Delta|y|}\times T_i(p_T,\eta\mid M_W)\right)$$

Measure double-differential cross section and angular coefficients

Parametrise dependence on m_W for each bin in the polarised cross section

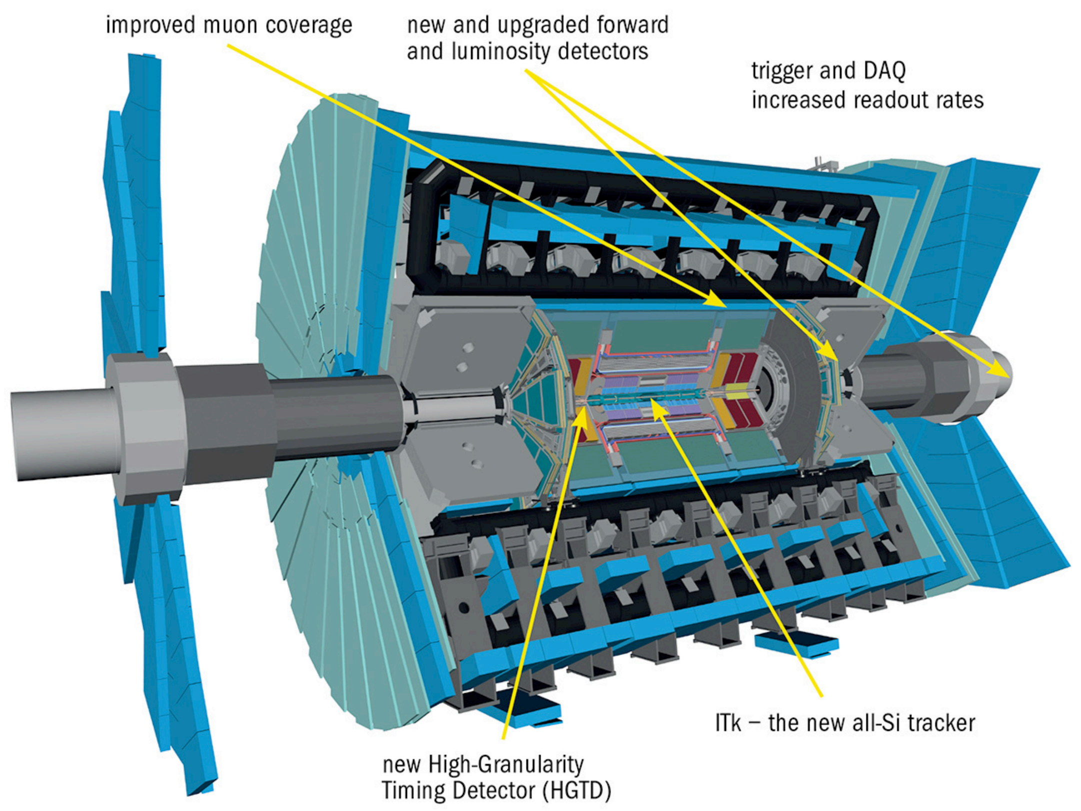
Ultimate precision for inclusive cross sections

Precision tt cross section -> extract α_S and m_t

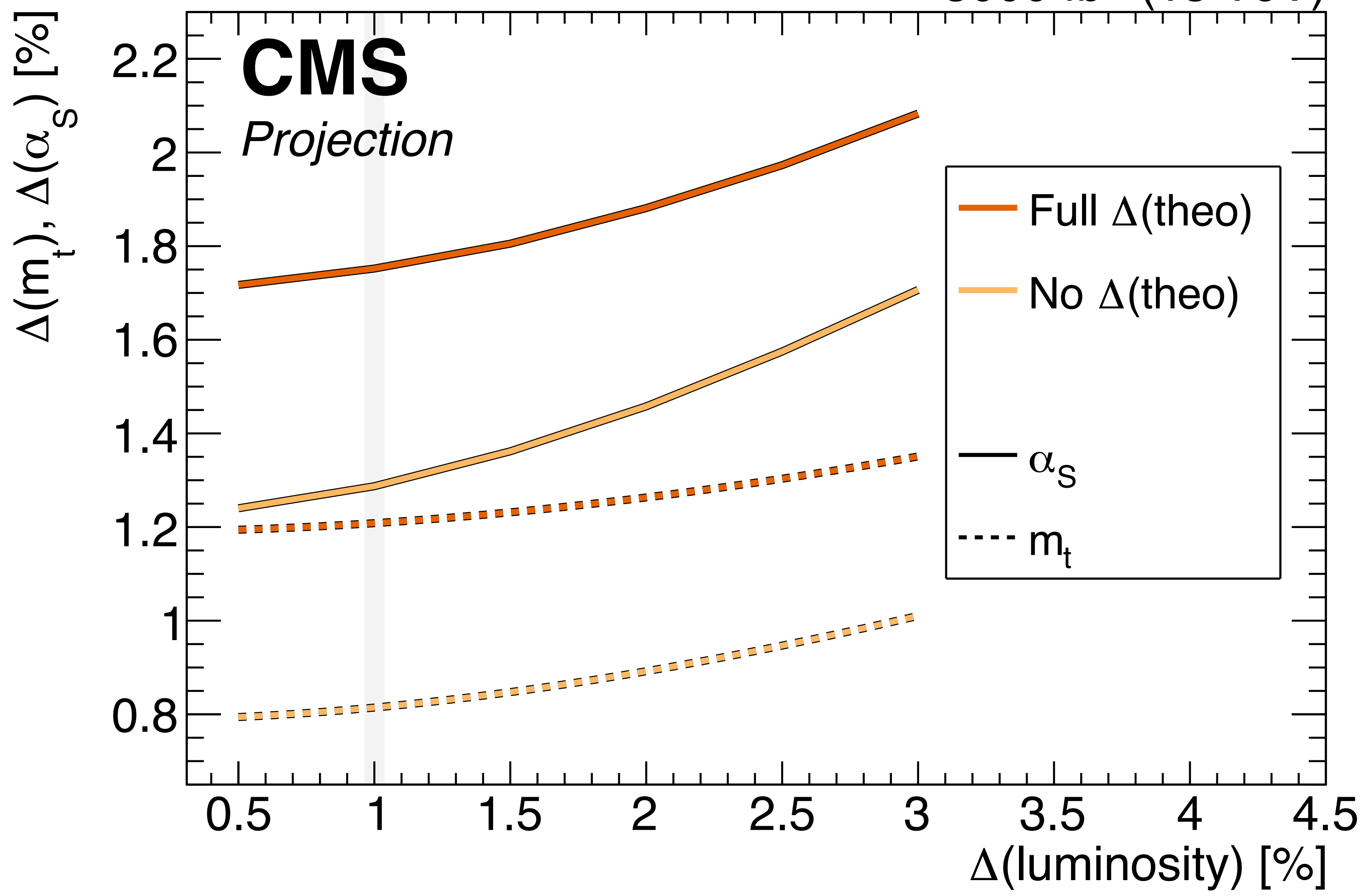
- Critically depends on precision in integrated luminosity (basically irreducible at analysis level)
- Both ATLAS and CMS upgrading luminometers to achieve precision of better than 1%

2212.09379

ATLAS already reached 0.83% precision in Run2



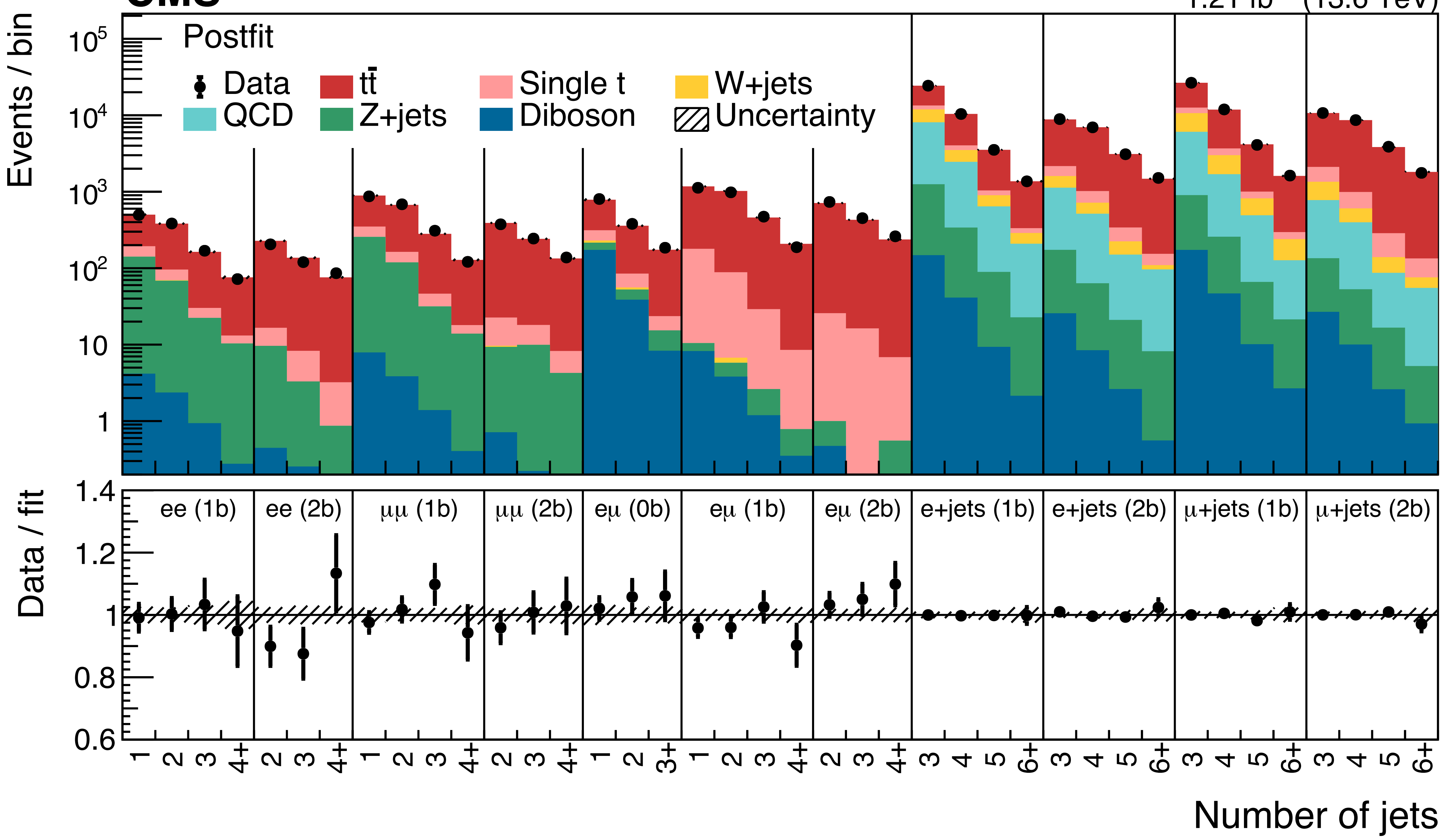
CMS-TDR-023



In-situ calibration of final state objects

- Inspired by the "traditional" ATLAS method
  - In-situ constraint of ε_b from b-jet multiplicity
- CMS: in-situ measurement of ε_e and ε_μ
  - Simultaneous fit to di-lepton and semi-leptonic channels

JHEP '23

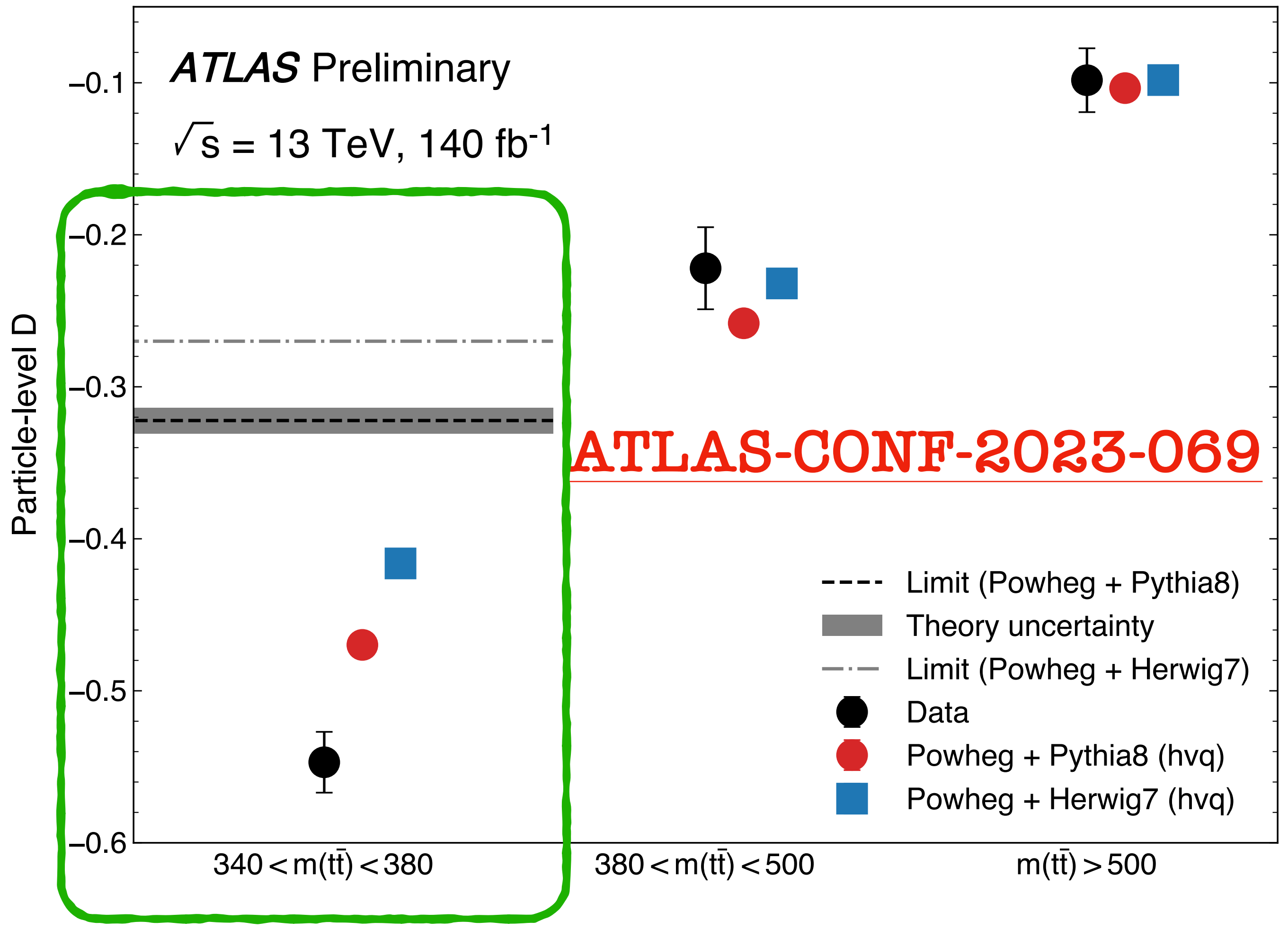


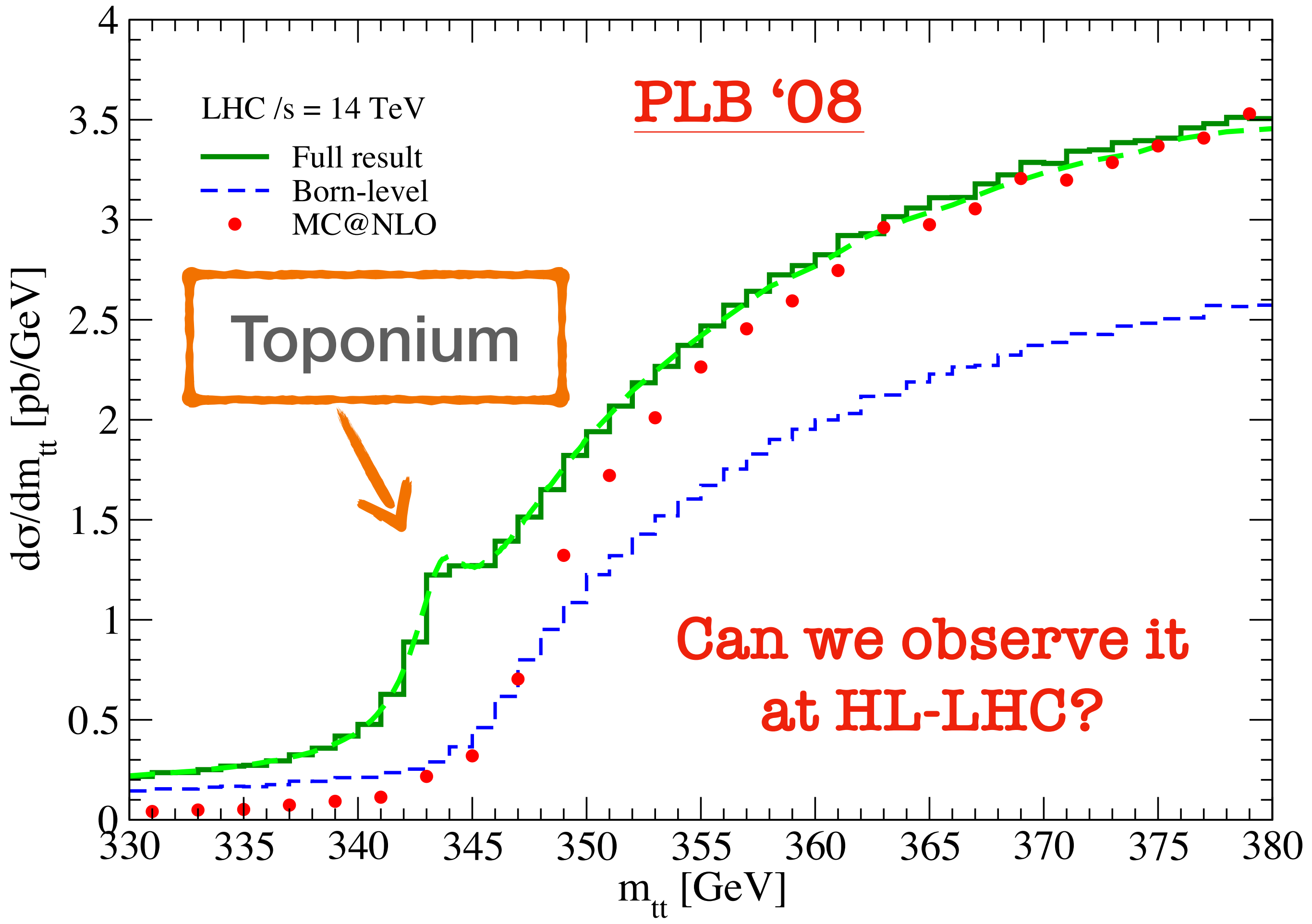
First 13.6 TeV result after just 1 month of data taking

Analysis performed using only inclusive observables

- Can this method be extended to any analysis?

tt in phase-space corners





tt production threshold

- Sensitive to top mass, Yukawa coupling, entanglement
- Contains a bound-state (toponium) contribution
- Threshold resummation approximated by parton shower in MC simulation (LL)

Progress in MC tools is key to fully exploit the information contained in threshold region

Boosted top production

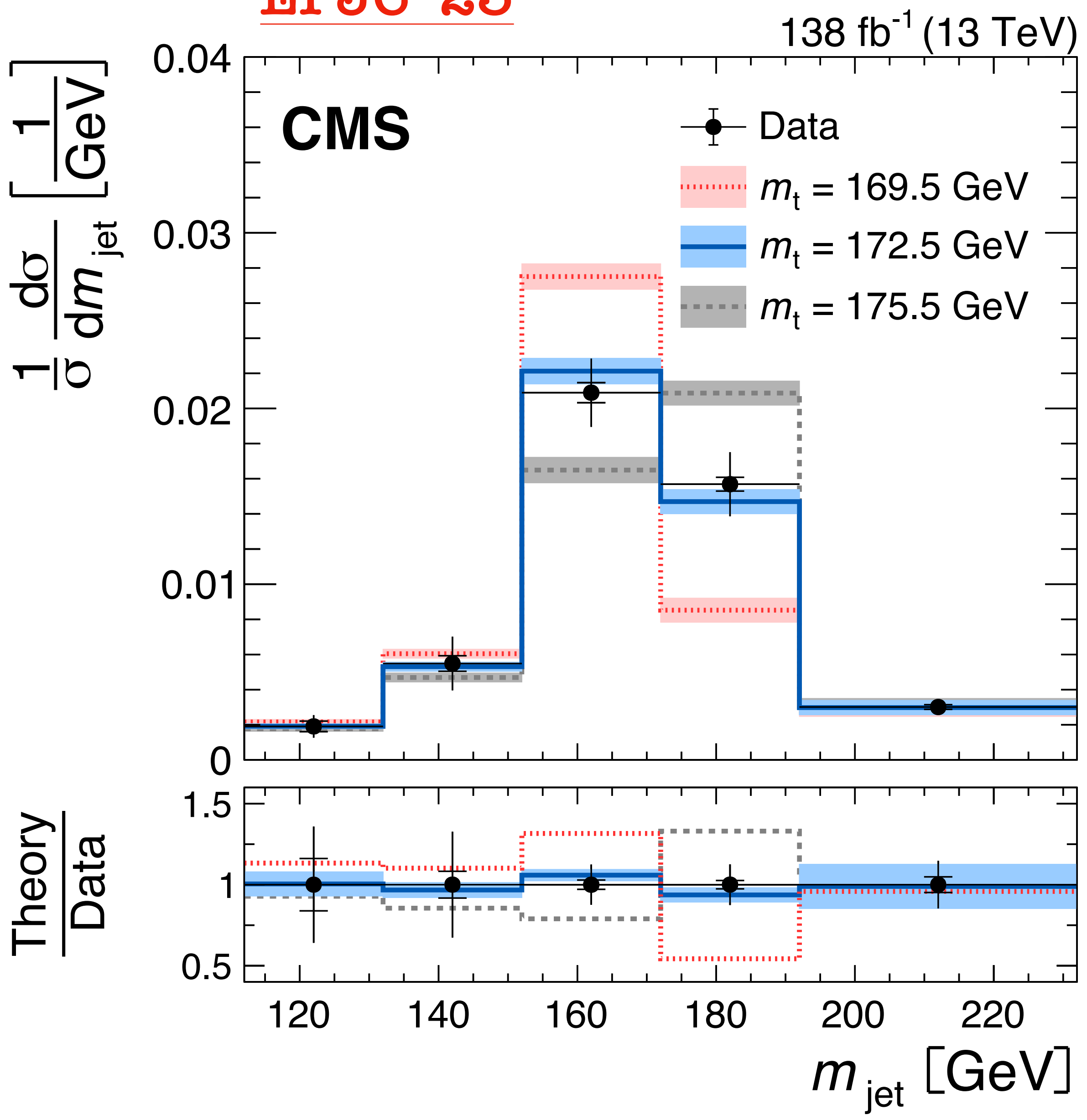
- Fully-merged top decay products -> jet substructure
- Interesting in EFT context + high-x PDFs
- m_t can be extracted from m_{jet}
  - Enormous progress in CMS: 9 GeV (Run1) -> 2.5 GeV (partial Run2) -> **0.84 GeV** (full Run2)
- Interpretation of direct top mass measurements

ATL-PHYS-PUB-2021-034

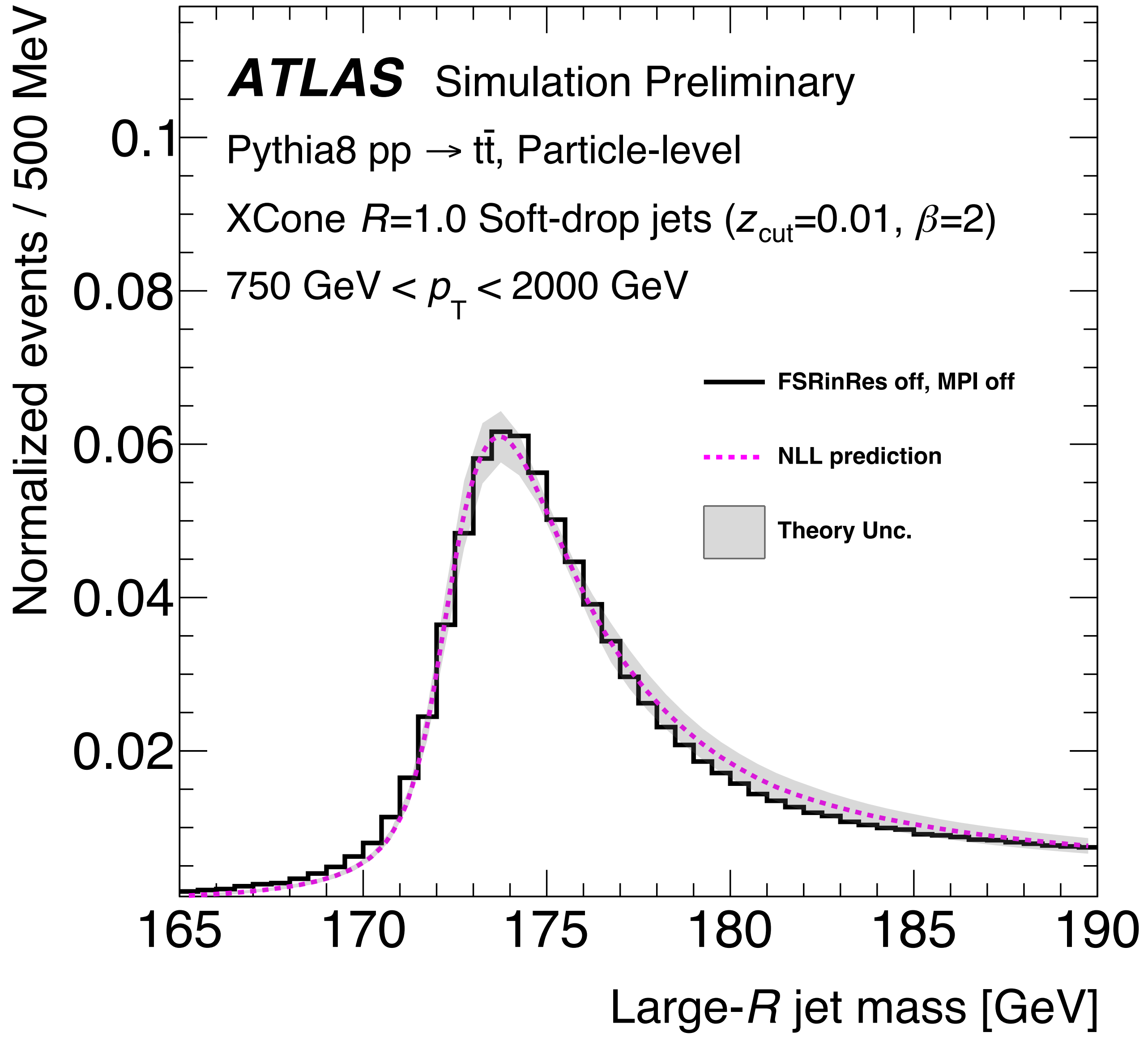
$$\Delta^{MSR} = m_t^{\text{MC}} - m_t^{\text{MSR}}(1\ \text{GeV}) = 80^{+350}_{-400}\ \text{MeV}$$

Large margin for improvement at HL-LHC (+higher top p_T)

EPJC '23



ATL-PHYS-PUB-2021-034



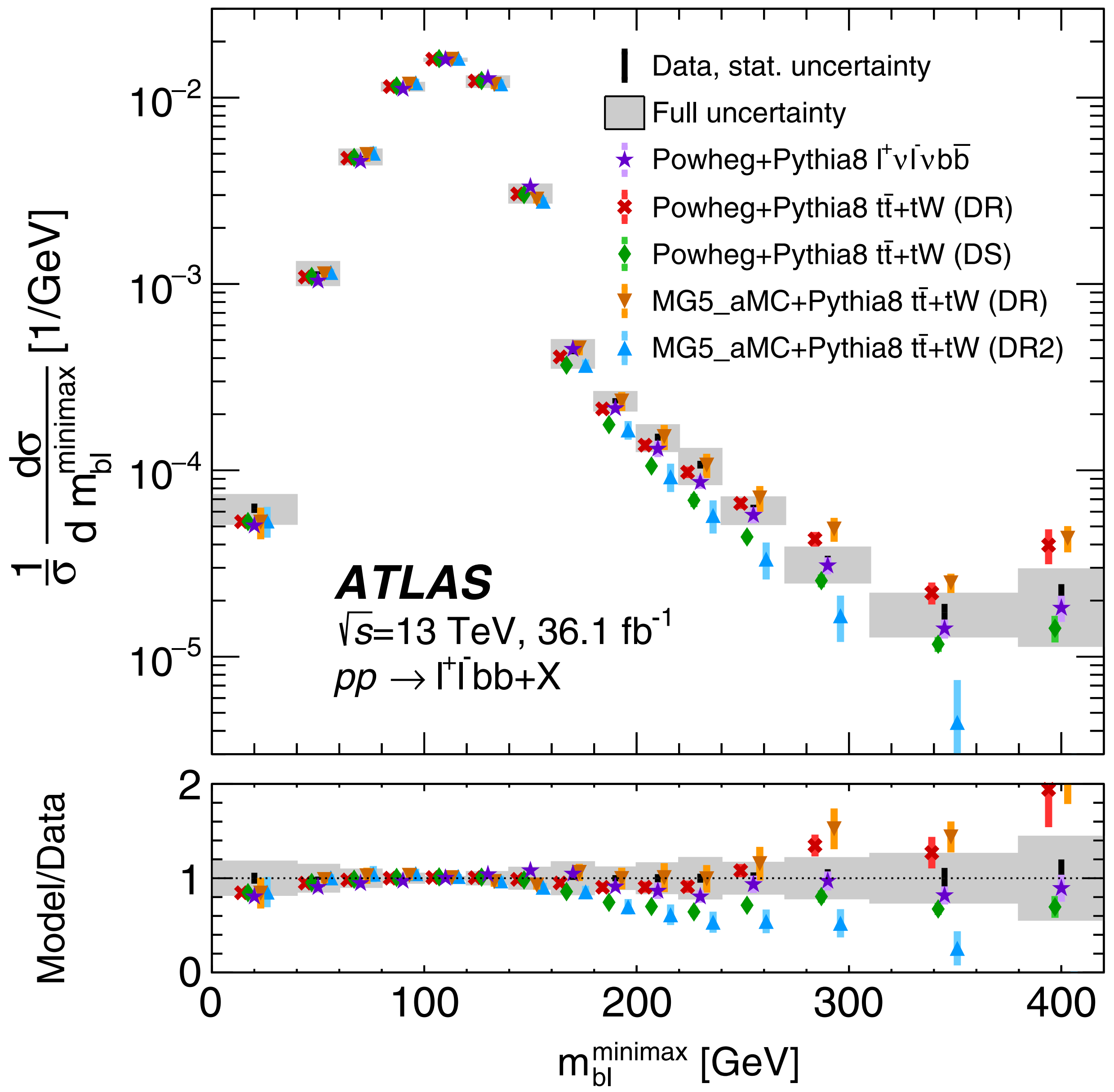
State-of-the-art theory tools

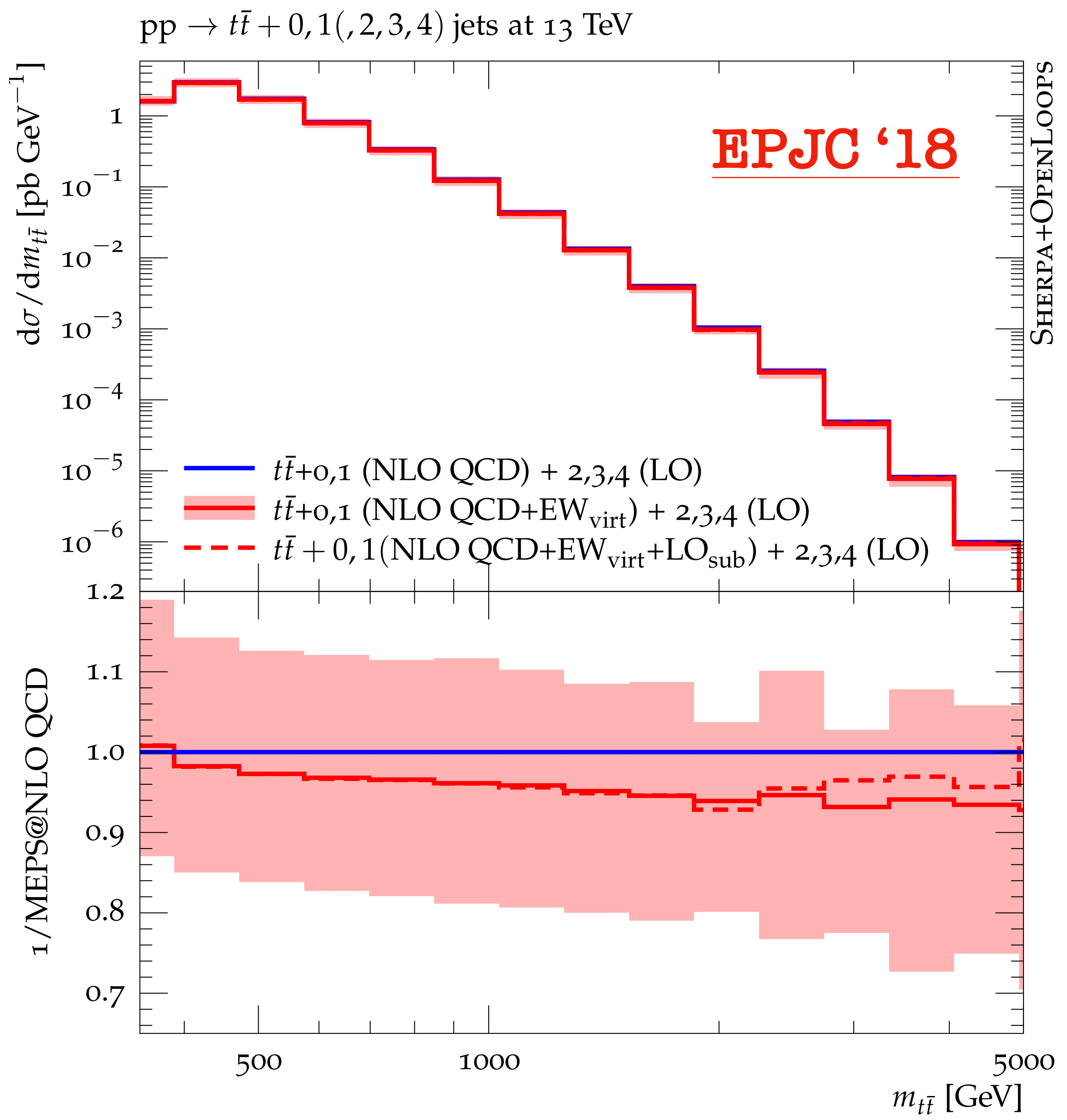
Advancements in ME calculations

- Full tt/tW interference and off-shell effects (bb4l)
- Higher order ME calculations (MiNNLO, Stripper, Matrix)
- EW corrections (especially relevant for m_{tt})

  - ◆ Not available for all processes
  - ◆ Computational costs often limiting factor
  - ◆ Need to put all pieces together

ATL-PHYS-PUB-2021-042





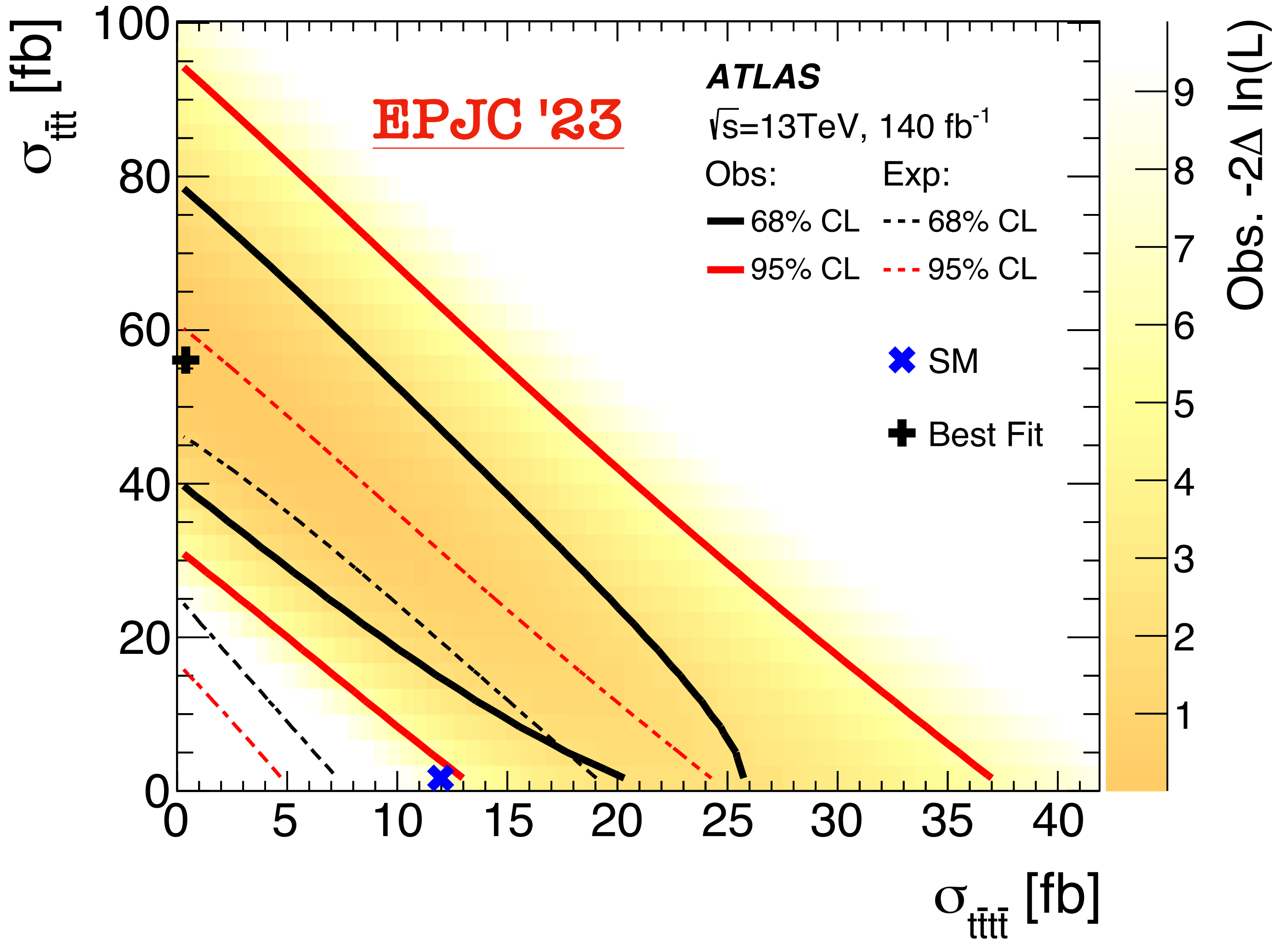
New generation of parton showers

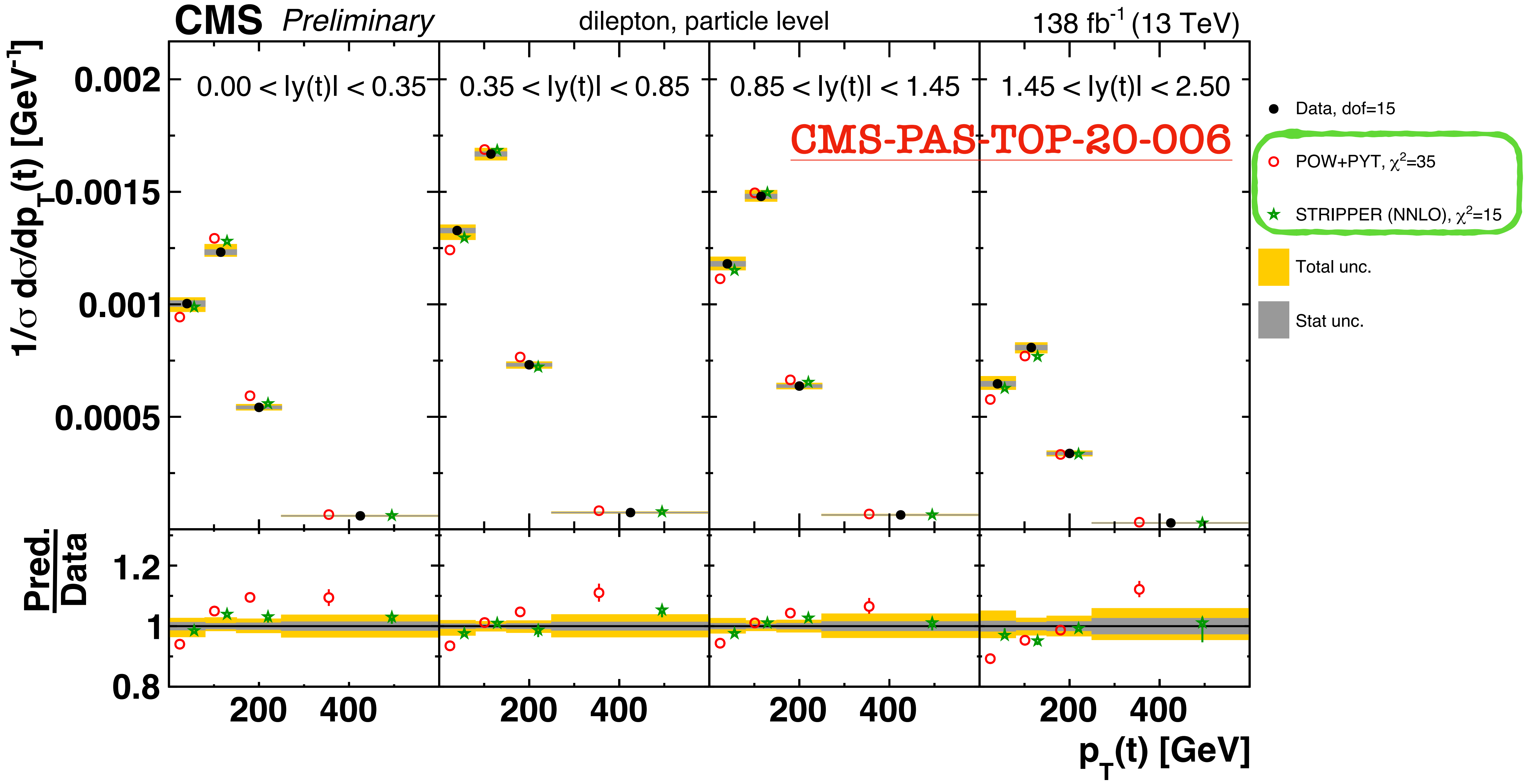
- VINCIA -> automatic weight-based variations, solves recoil ambiguity
- NLL parton showers

Ideally, we should target NNLO QCD + NLO EW for all processes

Re-think process definitions?

- Starting to see the limits of tt/tW separation
- Will soon have same problems in 4t/3tW, ttZ/tWZ, etc
- All definitions bound to a LO picture of processes -> less and less realistic





Matteo M. Defranchis (CERN)

Towards multi-process EFT

1. Different EFT operators affect the same process
2. Difference processes affected by same EFT operators

-> no clear distinction between signal and backgrounds

Simultaneous measurements

- Optimal results obtained by directly targeting EFT at detector level
- However, hard to include into global interpretation

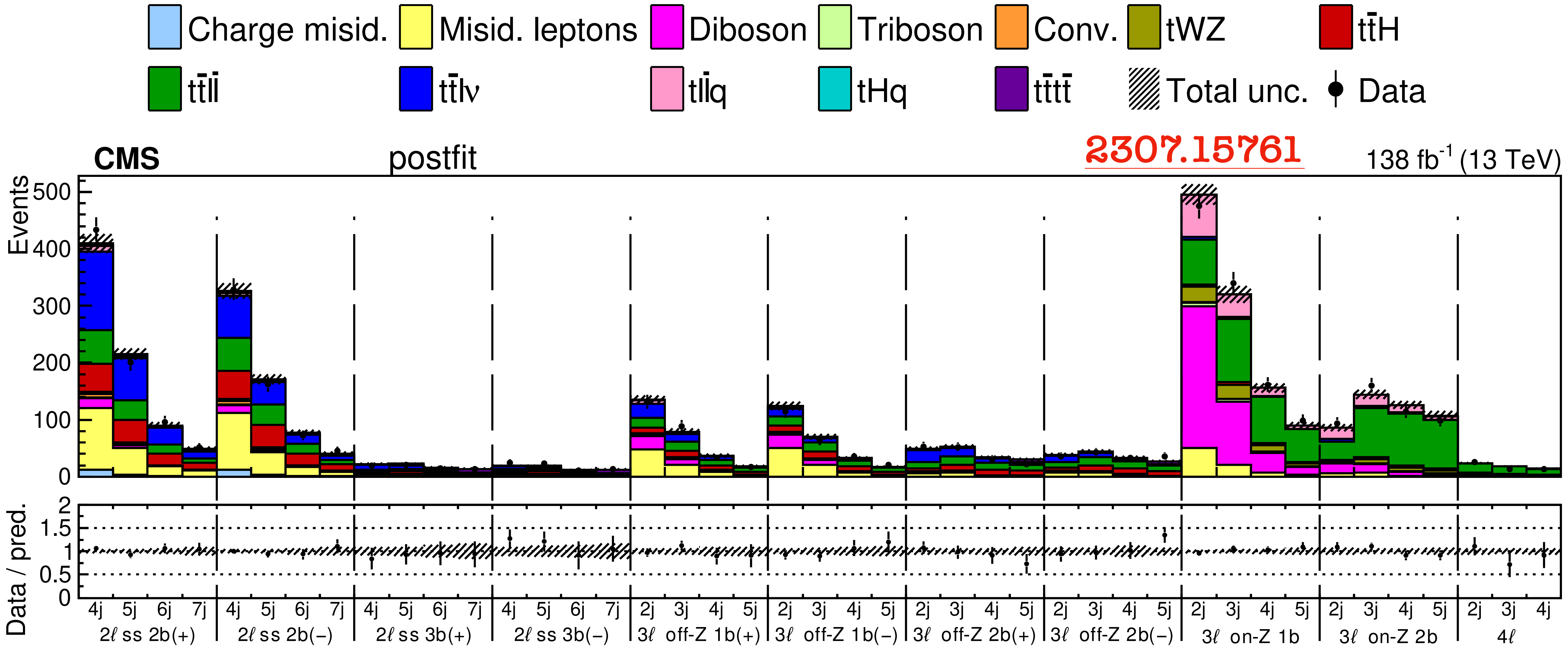
See also: ATL-PHYS-PUB-2023-030

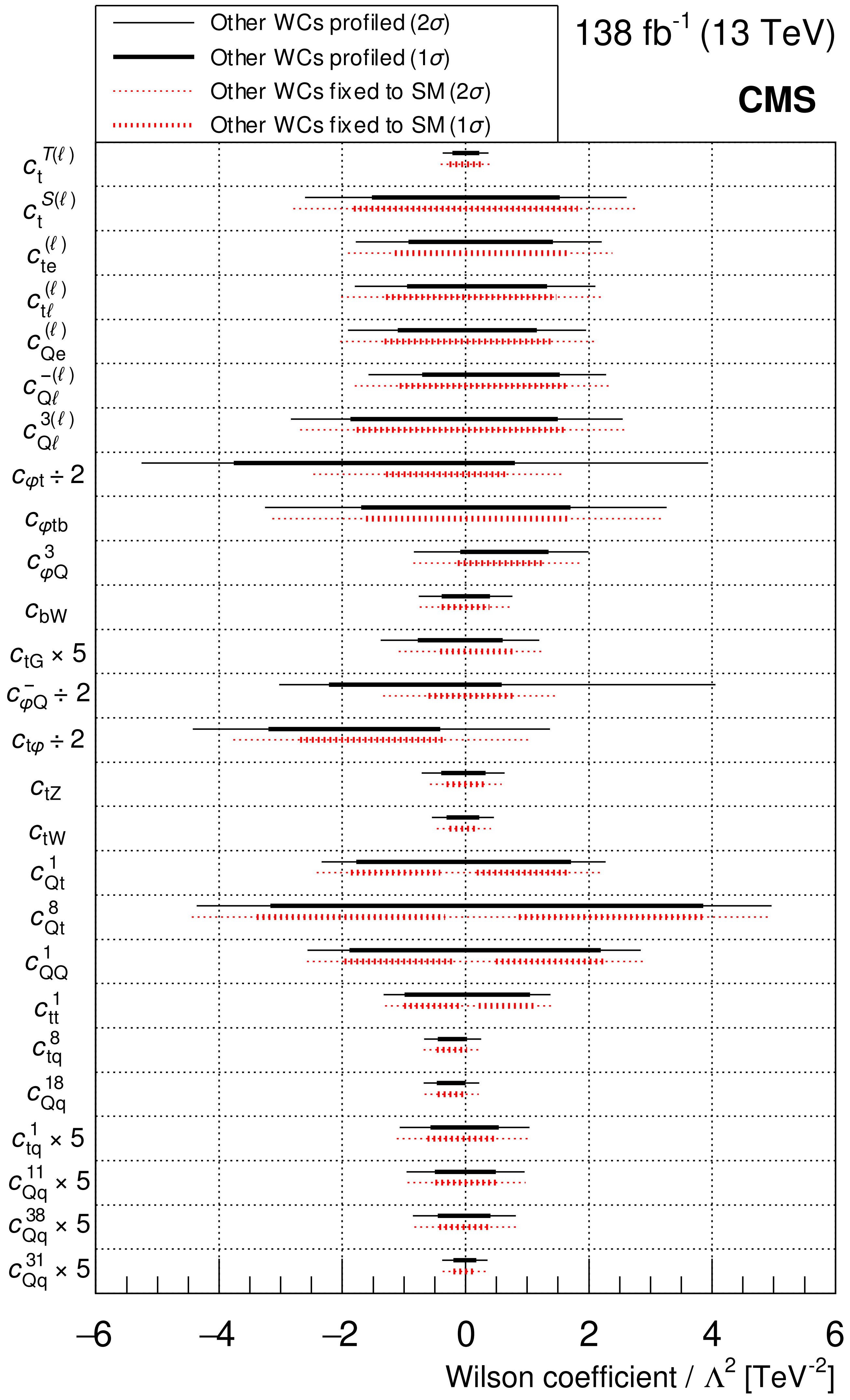
=> **multi-process unfolding of EFT-sensitive processes**
(“easily” combine different experiments, measurements, LHC runs, etc)

Improved separation between processes can be achieved with ML approach (e.g. DNN, GNN, …)

Challenges

- EFT could also affect signal acceptance
- Contribution to PDFs
- How to interface EFT calculations with higher-order QCD MC





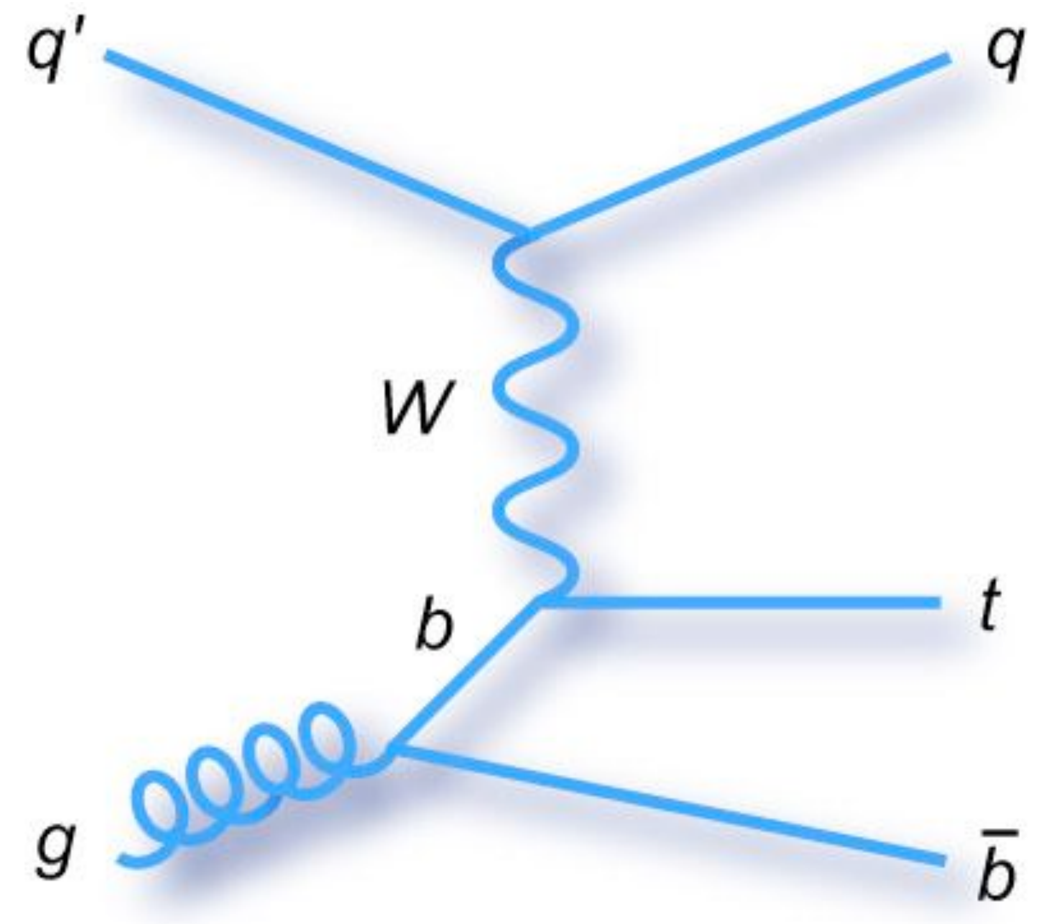
New detectors, new ideas

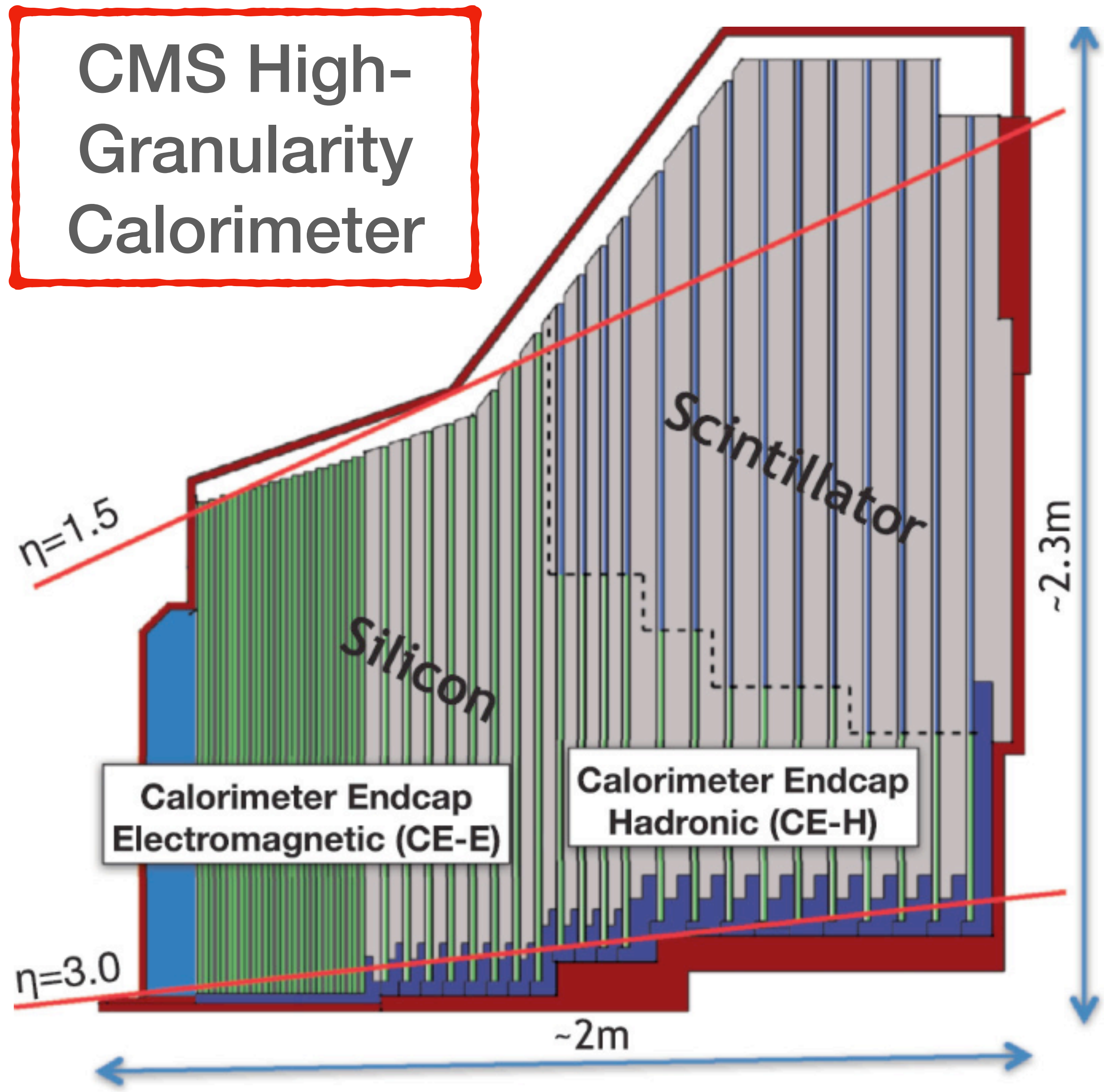
1. Designed to cope with high pile-up and high radiation damage
2. Increased detector granularity, timing capabilities
3. **Extended coverage of key detectors**: trackers, muon systems, calorimeters

Forward object reconstruction

- Forward tt production -> high-x PDFs
- Increase acceptance -> reduce extrapolation uncertainties
- Forward flavour tagging -> **top EW production**

New detector + advanced ML techniques can significantly boost performance





Wb->Wb scattering can be used to measure Γ_t at HL-LHC

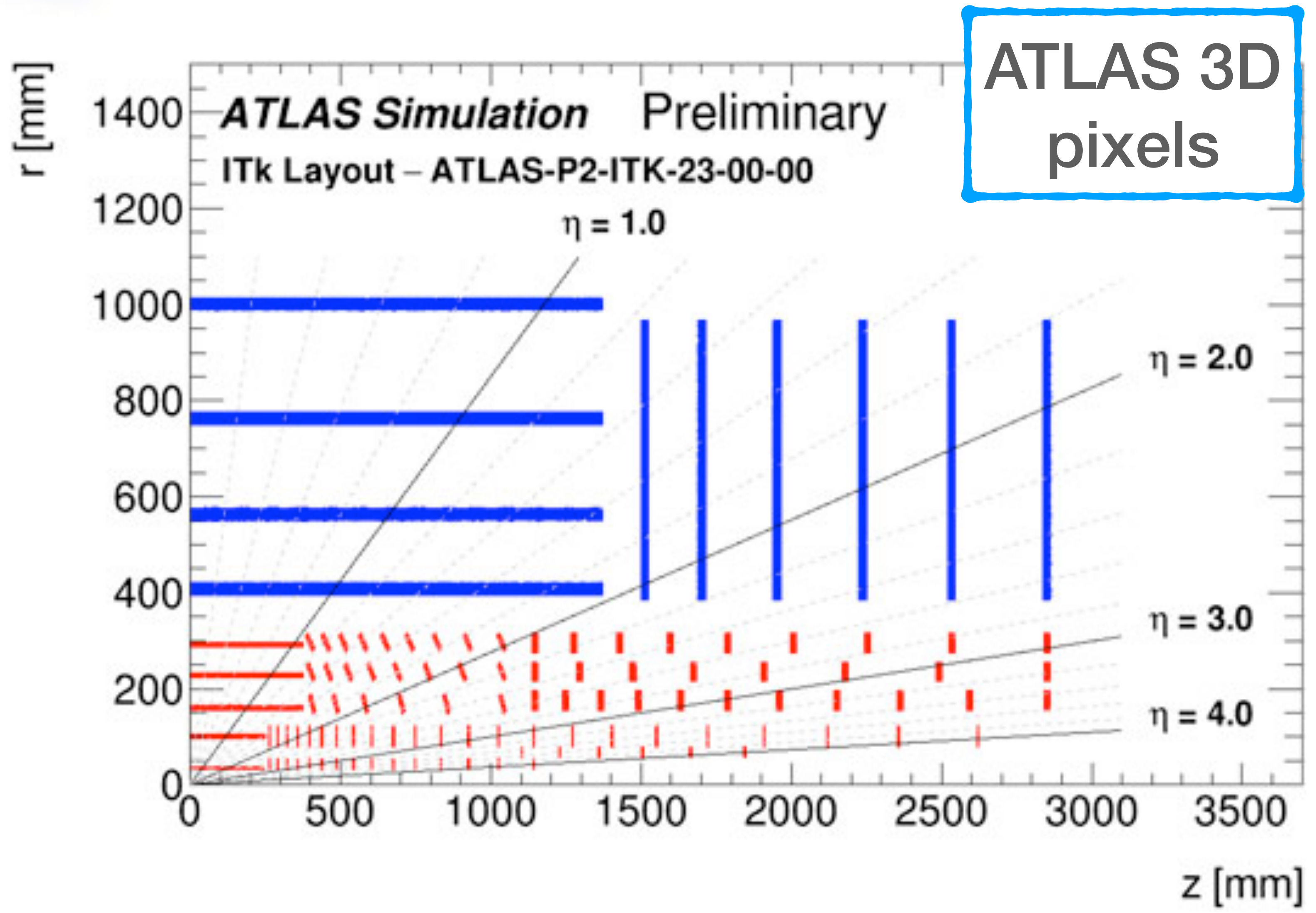
- Higgs-inspired on-shell/off-shell ratio
- Exploit (W-induced) b-charge asymmetry -> **b-charge tagging**
- Different sensitivity to systematics compared to standard approaches

Crucially depends on efficiency of identifying b-quark charge

$\delta\Gamma_t \sim (2\varepsilon-1)/\sqrt{\mathcal{L}}$

PRD '17

Luminosity [fb^{-1}]	30	300	3000
Limits [GeV]	[0.40,2.30]	[1.01,1.73]	[1.14,1.60]
Stat. error	11%	3%	1%



Life after the LHC

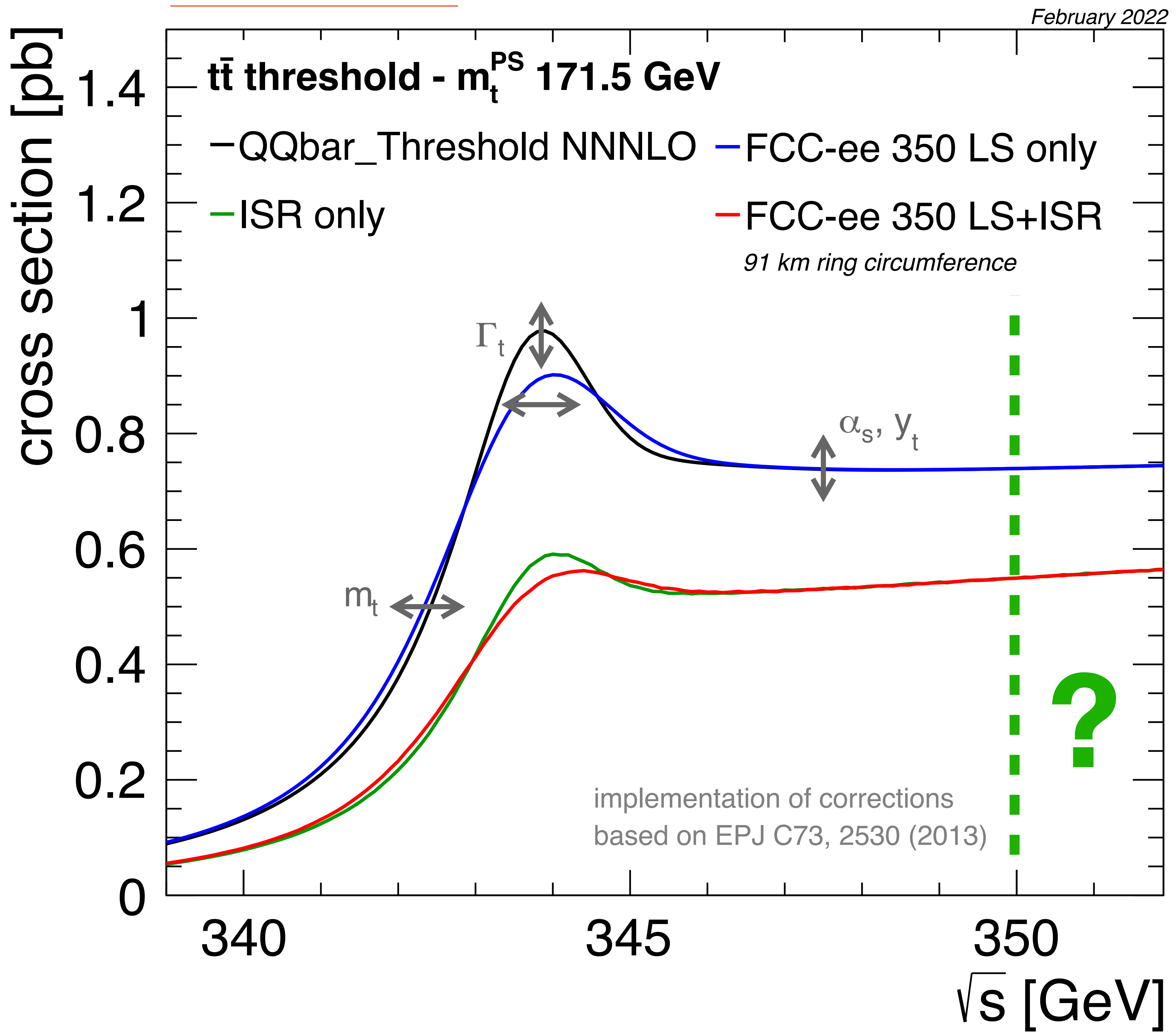
e^+e^- collider at tt threshold

- Unprecedented precision in m_t (few tens of MeV)
- However, unlikely to reach energies much higher than 350 GeV (FCC-ee)

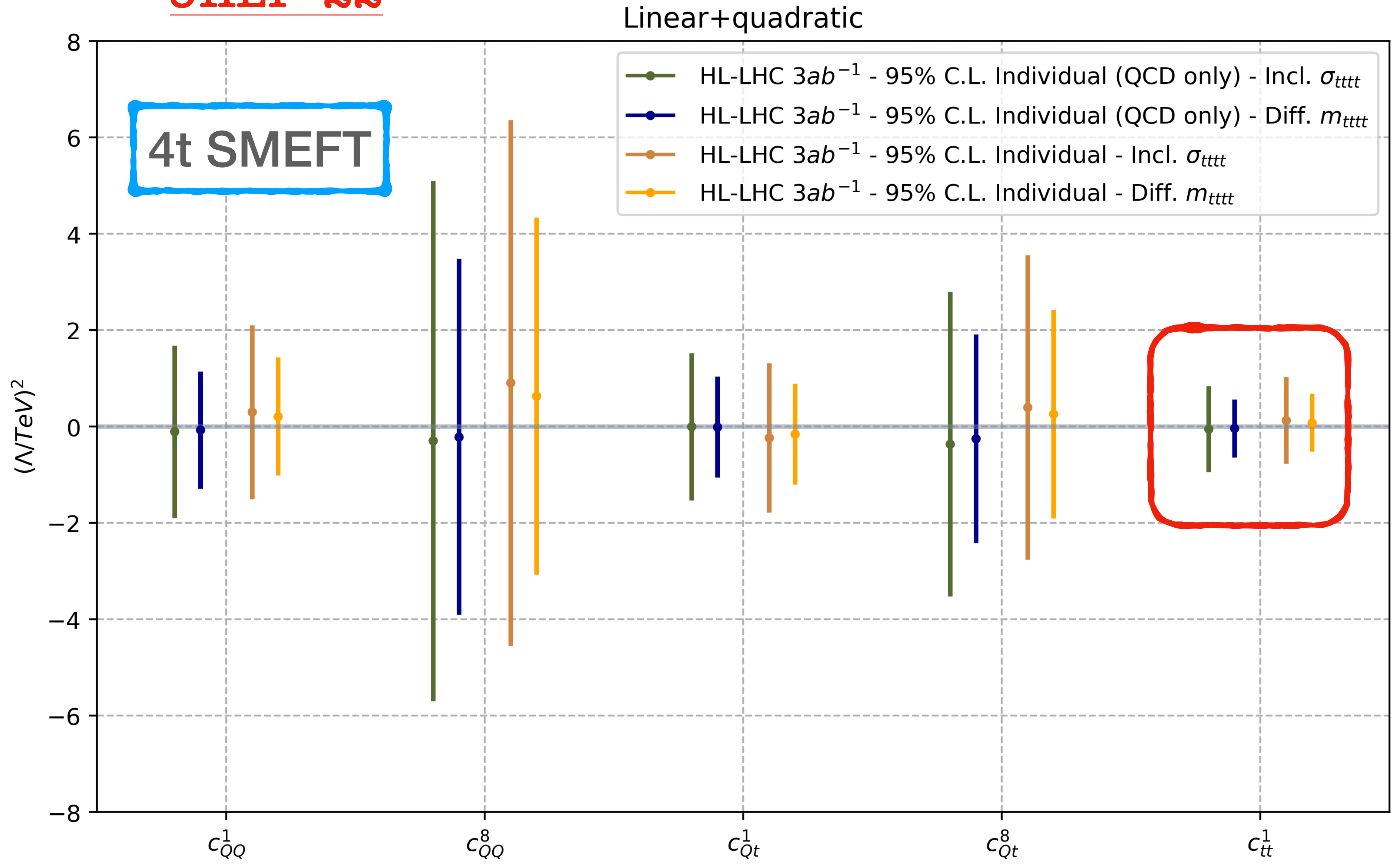
tt+V, tt+H, 3t/4t measurements will be the legacy of HL-LHC

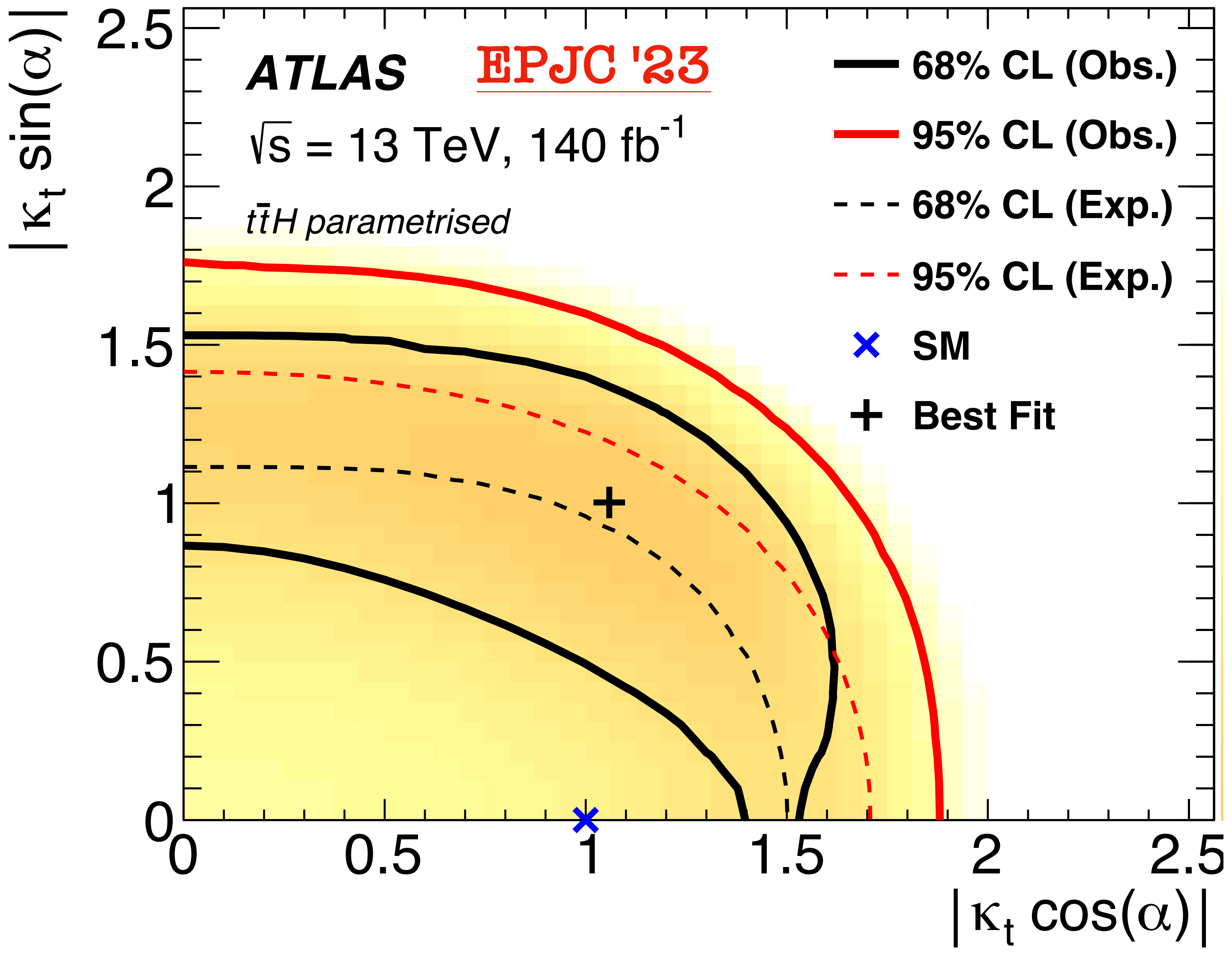
- t-H and t-t (EFT) couplings will still be accessible via loop corrections
- Most sensitive probe of top self-coupling will remain **4t at HL-LHC**

2203.06520



JHEP '22





The "differential" challenge

- Higher sensitivity even with few bins
- Crucial to combine LHC results
- Experiment + theory driven choice of final state observables? (e.g. Higgs STXS)

Take-home messages

- HL-LHC is meant to produce **legacy results** that will last for decades
- It is our mission as a HEP community to make the most out of the HL data

- Community-wide major undertakes to improve tools and tackle pitfalls
- **Combination** of results and **global EFT** interpretations are major goals

- Not everything will be accessible in e^+e^- -> 3t/4t and tt+X will be HL-LHC priorities
- Can profit a lot from **improvements in analysis techniques** in the next 20 years

- **Ambitious** (and expensive) detector and accelerator upgrade plans
- We should be at least as ambitious in planning our analysis efforts

Thank you