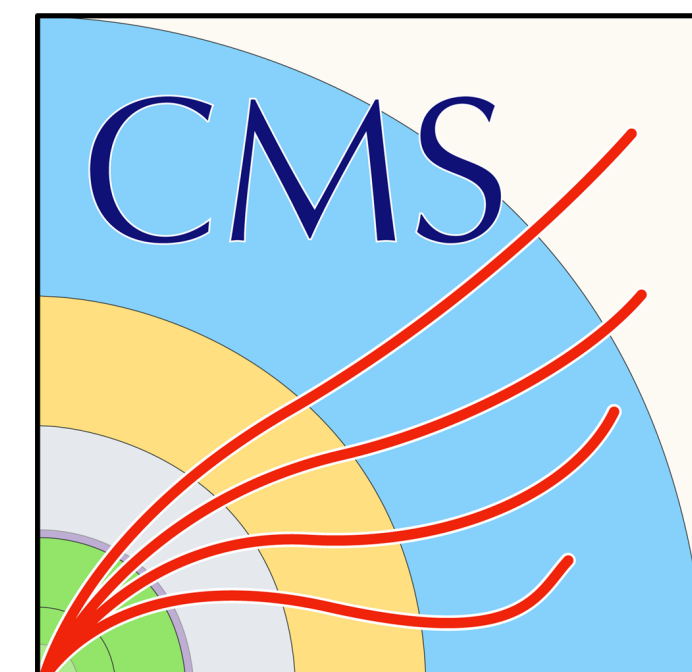
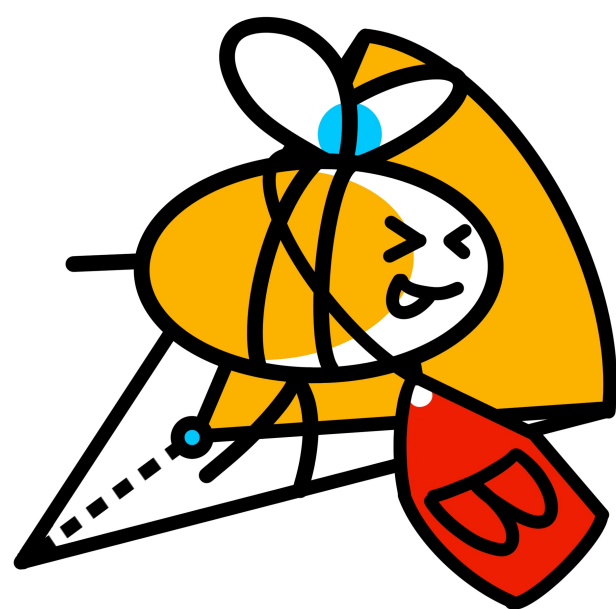
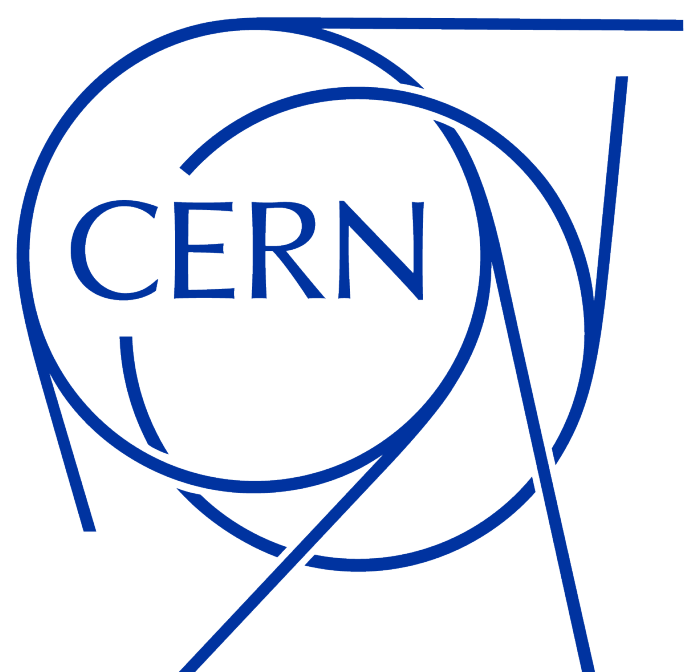


From jet flavour tagging to $H \rightarrow c\bar{c}$: probing Higgs couplings at CMS

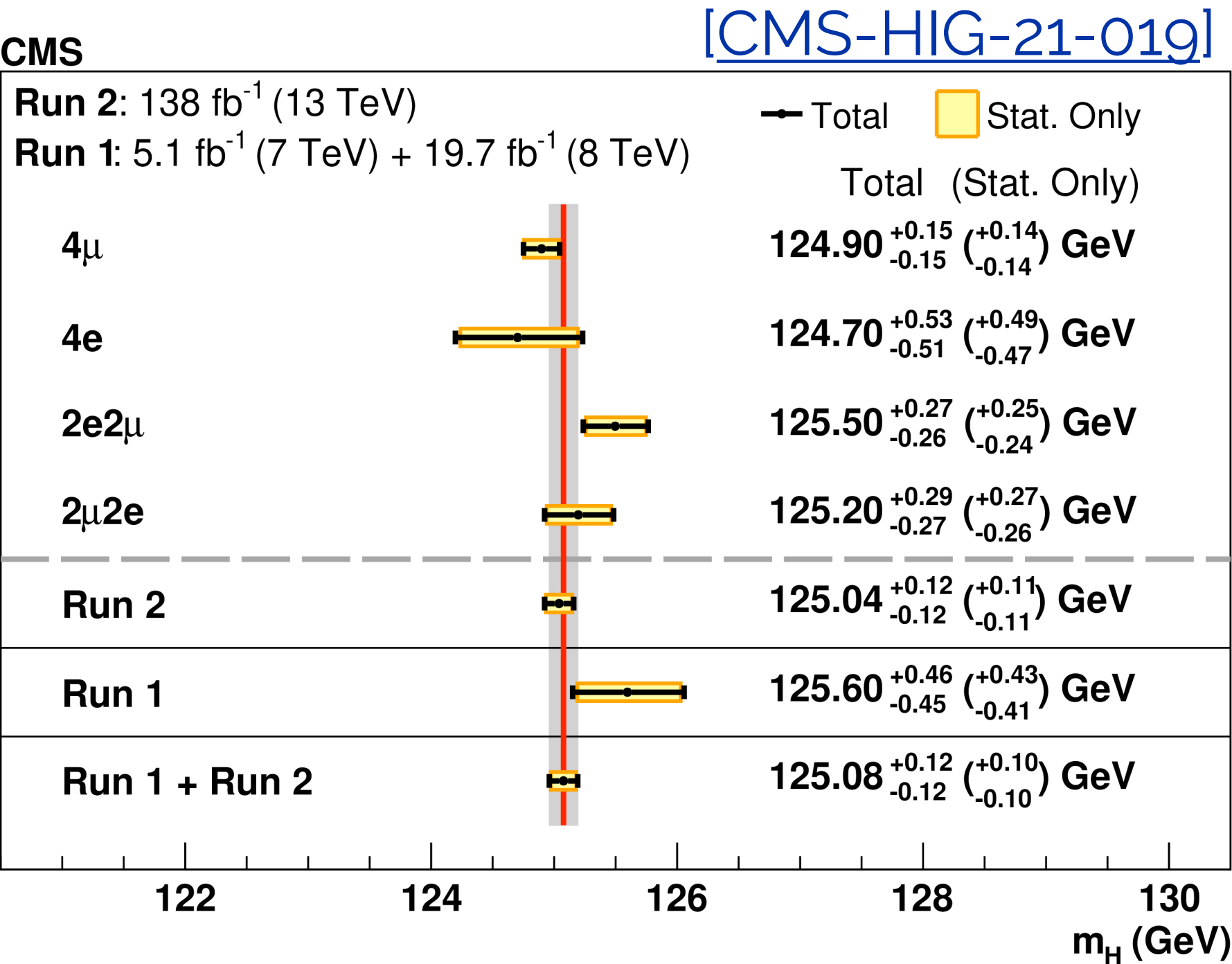
Sebastian Wuchterl (CERN)



Introduction and motivation



- 10+ years ago: A new era of particle physics
 - Confirmation of the Higgs-mechanism
 - Masses of W/Z bosons through EW symmetry breaking
 - Fermion masses added ad-hoc via Yukawa couplings



- Tremendous progress in understanding the Higgs boson
- Mass measured with per-mil precision
- From discovery to precision in one decade!



Introduction and motivation

- 10+ years ago: A new era of particle physics
 - Confirmation of the Higgs-mechanism
 - Masses of W/Z bosons through EW symmetry breaking
 - Fermion masses added ad-hoc via Yukawa couplings



Mass hierarchy?

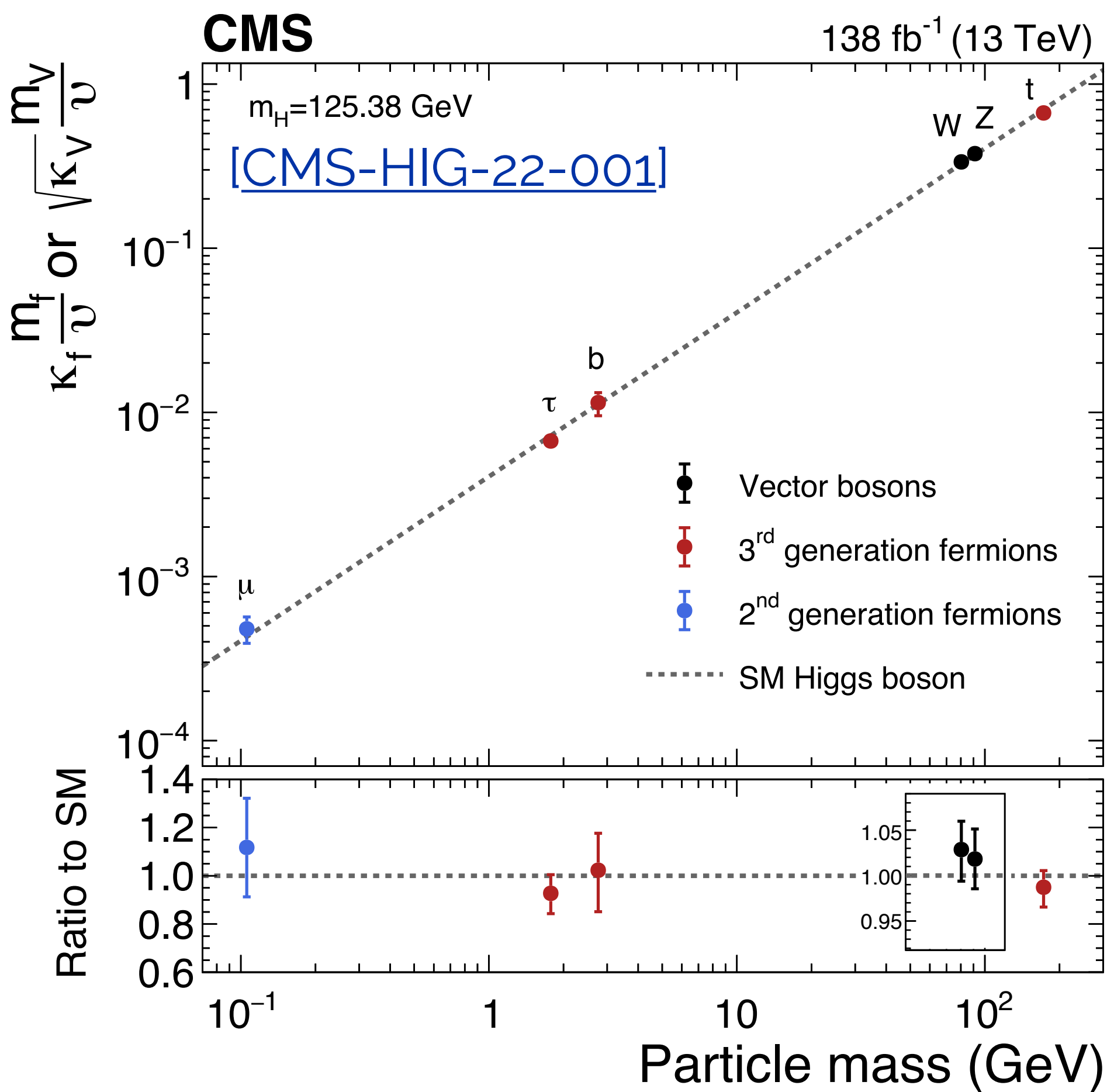
**BSM
modifications?**

See e.g. [arXiv:2410.05236](https://arxiv.org/abs/2410.05236)
[Phys. Rev. D 100, 115041](https://arxiv.org/abs/2410.05236)
[Phys. Rev. D 94, 115031](https://arxiv.org/abs/2410.05236)

- But how much closer are we to fundamentally understanding fermion mass generation?
- With known m_H , fixed predictions for decay and production rates

Study of cross section and decay rates of utmost importance!

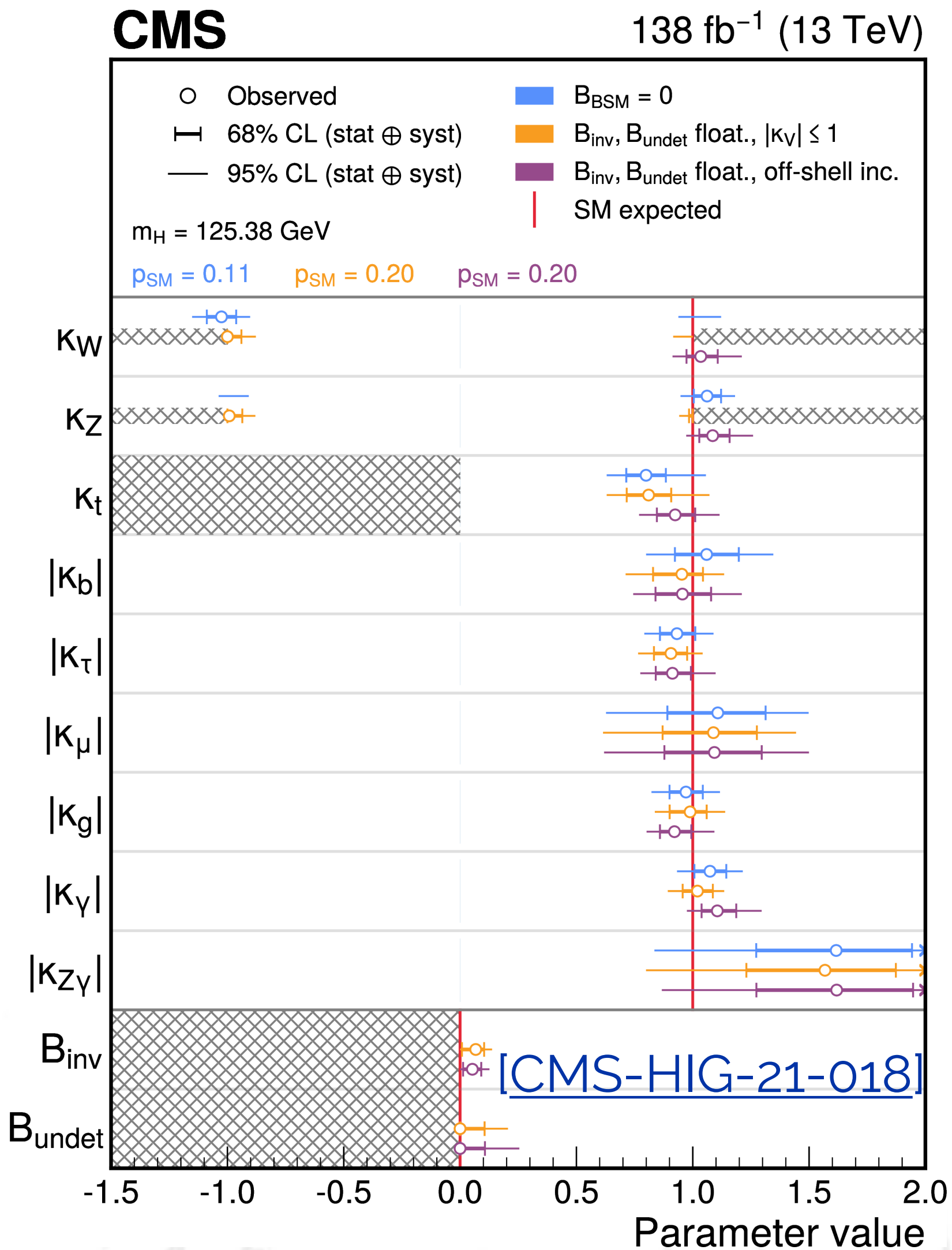
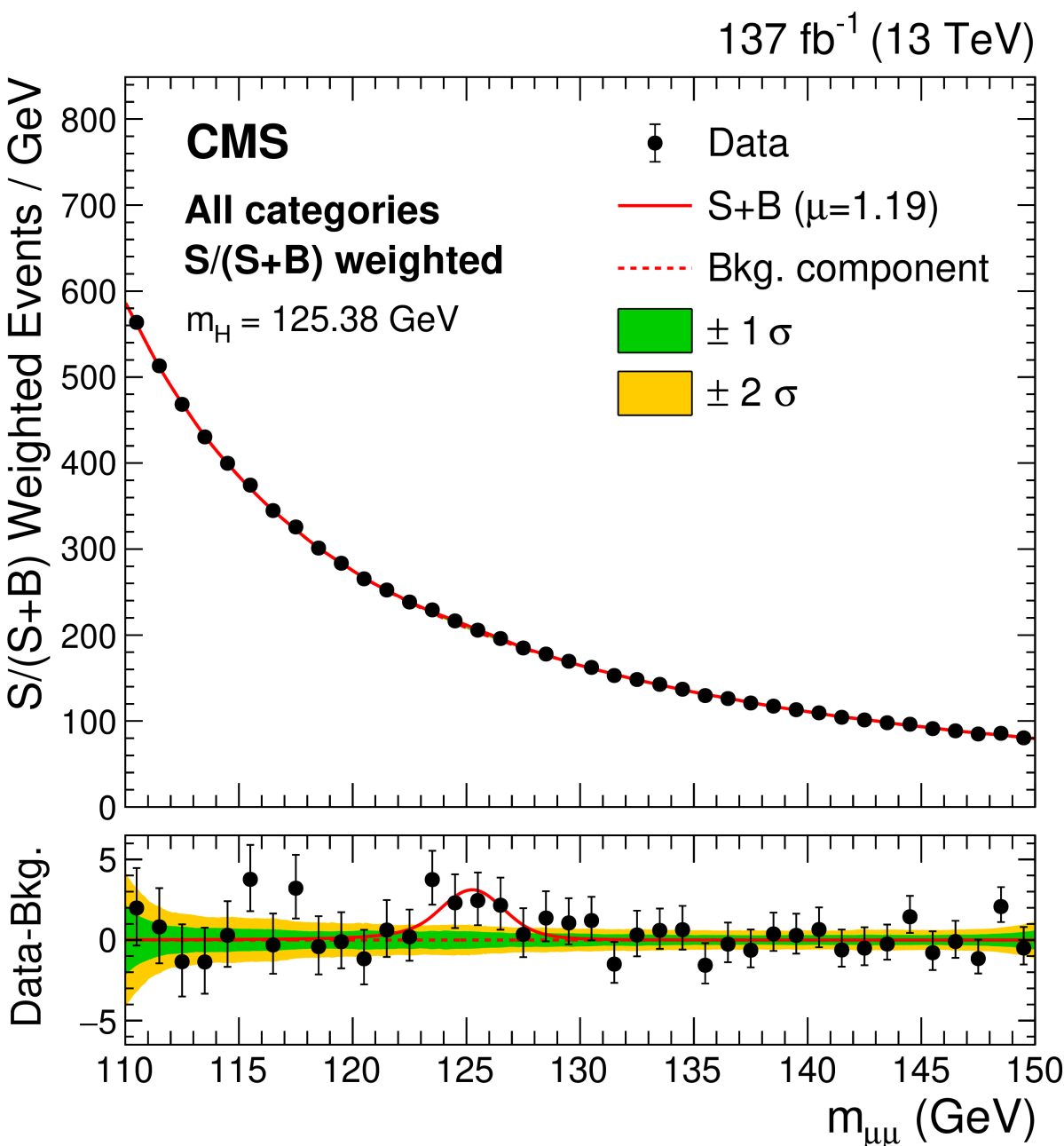
Confirmation of the Yukawa picture



*Evidence for $H \rightarrow \mu\mu$ in 2020!
3.0 σ obs. (2.5 σ exp.)*

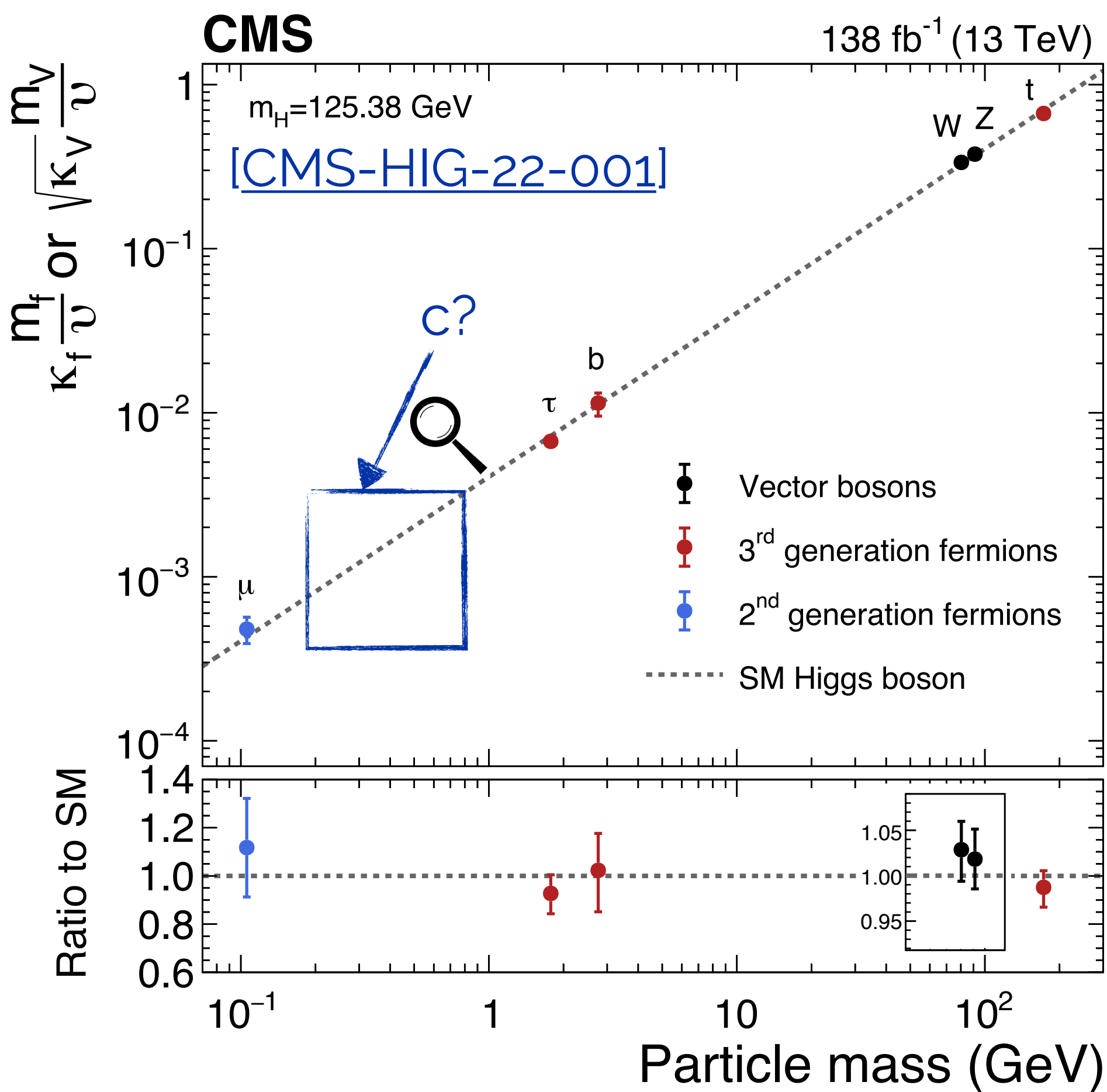
[CMS-HIG-19-006]

- Couplings to bosons and third generation fermions established (O(10%))
- Evidence for coupling to muons
- Any deviation could point to BSM physics



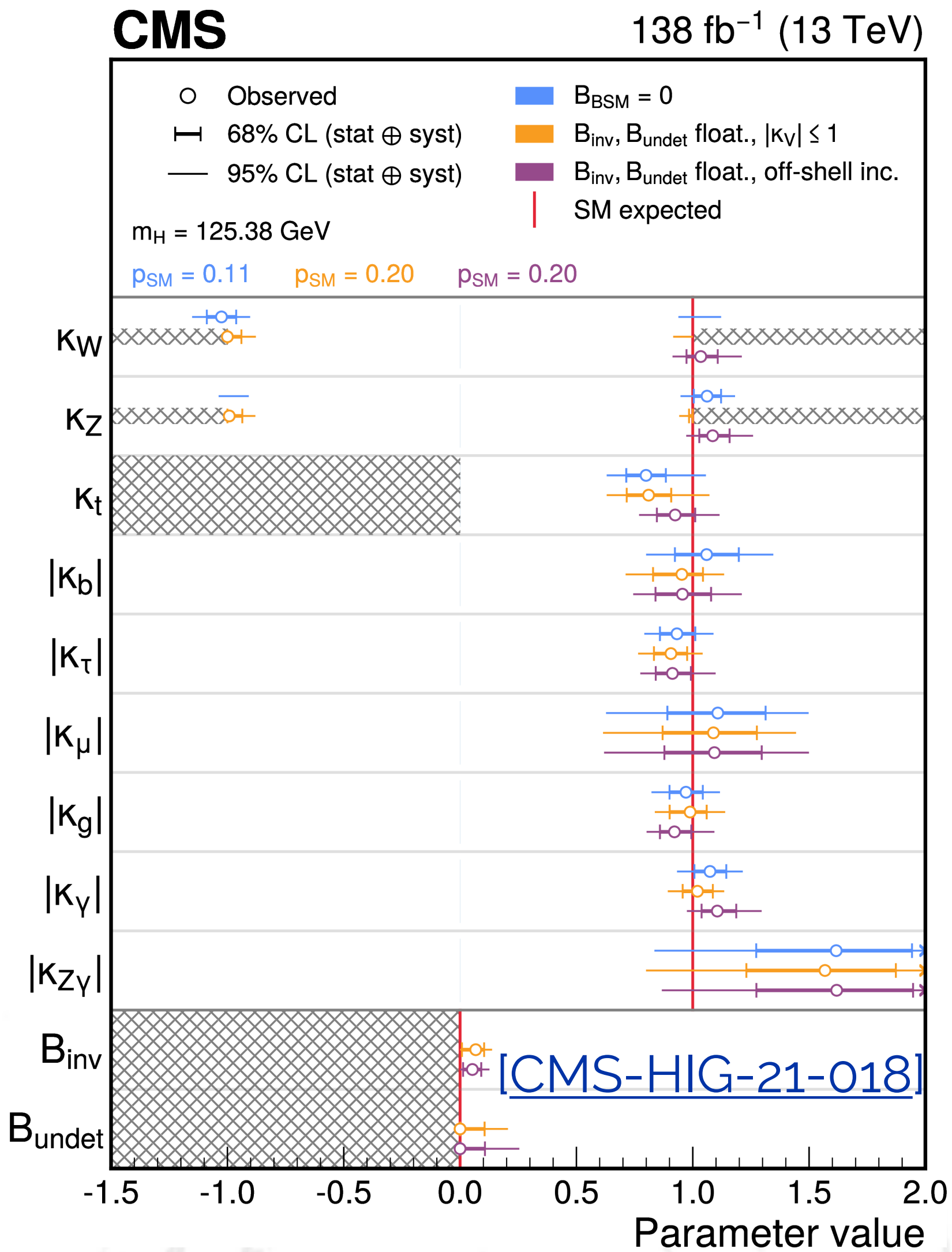
CMS combined coupling interpretation

Confirmation of the Yukawa picture



- Couplings to bosons and third generation fermions established (O(10%))
- Evidence for coupling to muons
- Any deviation could point to BSM physics

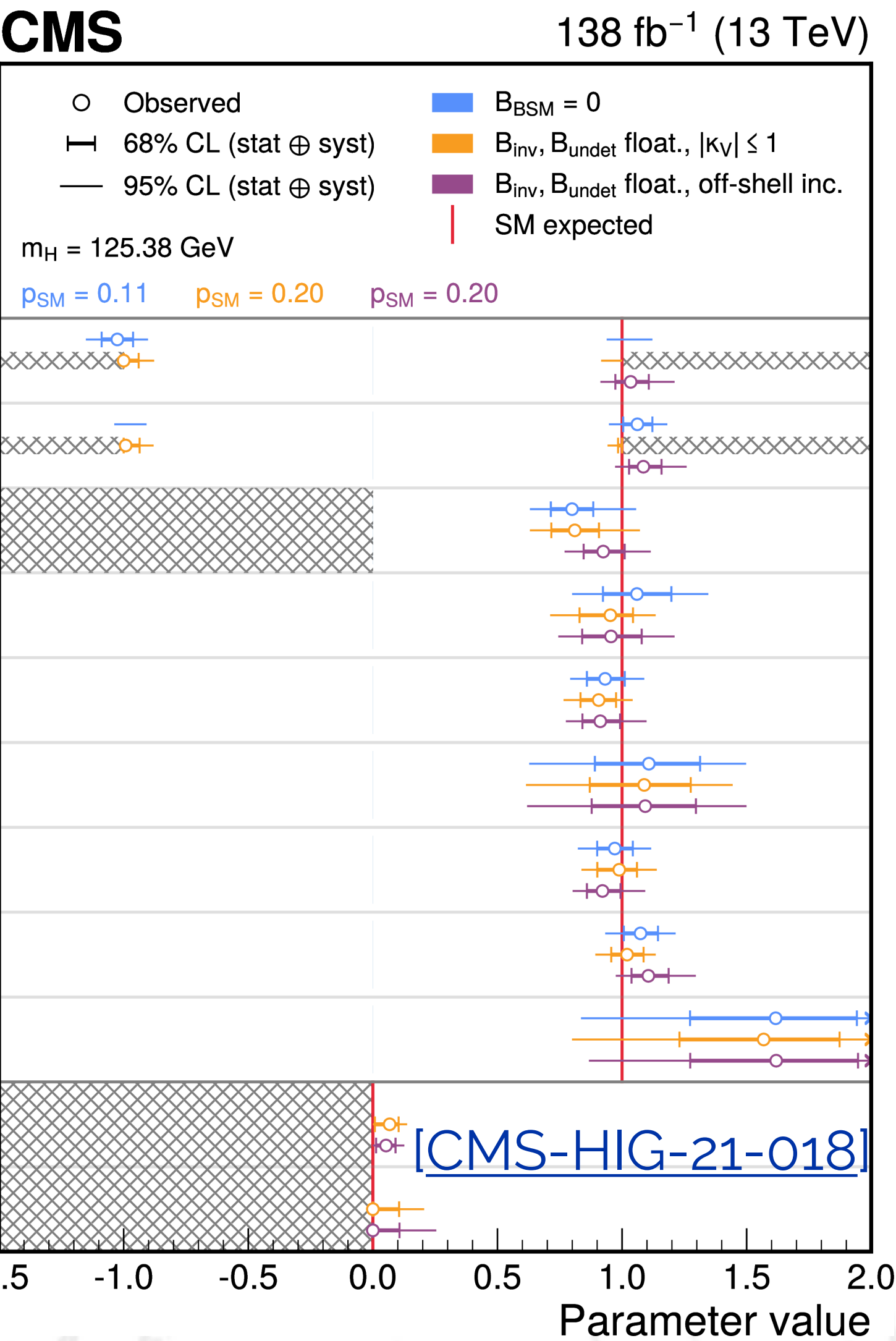
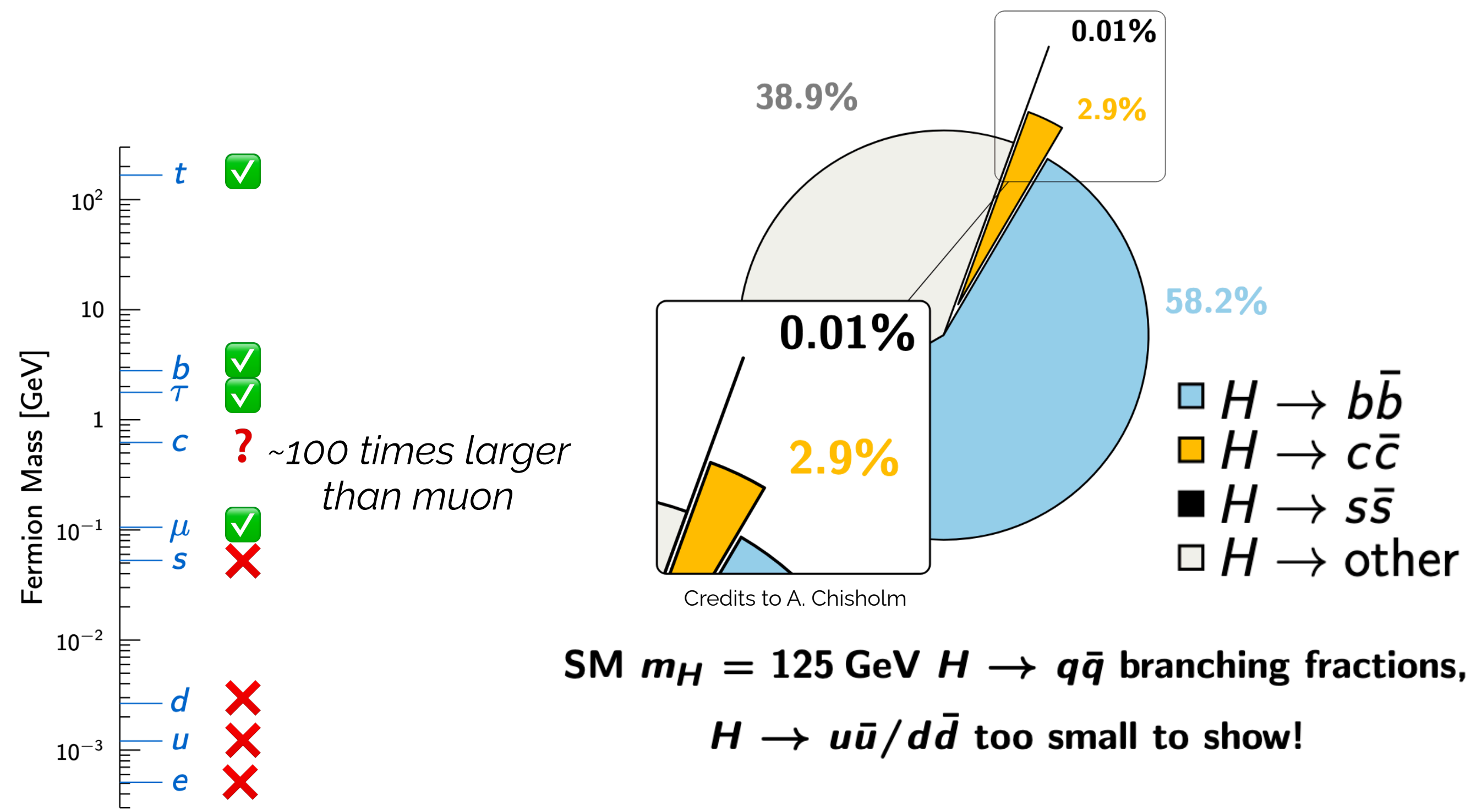
The measurement of the Higgs-charm coupling is at the frontier of Higgs physics!



CMS combined coupling interpretation

The Higgs-charm Yukawa coupling

- Smallness of the charm-quark Yukawa coupling
 - Any deviation could be interpreted as BSM physics
- Largest contribution to the undetected width
 - Remaining quarks and electron impossible at the LHC



Charm is ~what is left
(not considering first generation)

But... the quark sector is much harder!

So far: probing the Higgs-charm coupling



- Several methods explored by ATLAS and CMS to probe the Higgs-charm Yukawa coupling (y_c)

Indirect constraint from Higgs kinematics

Model dependence?

Direct searches for $H \rightarrow c\bar{c}$



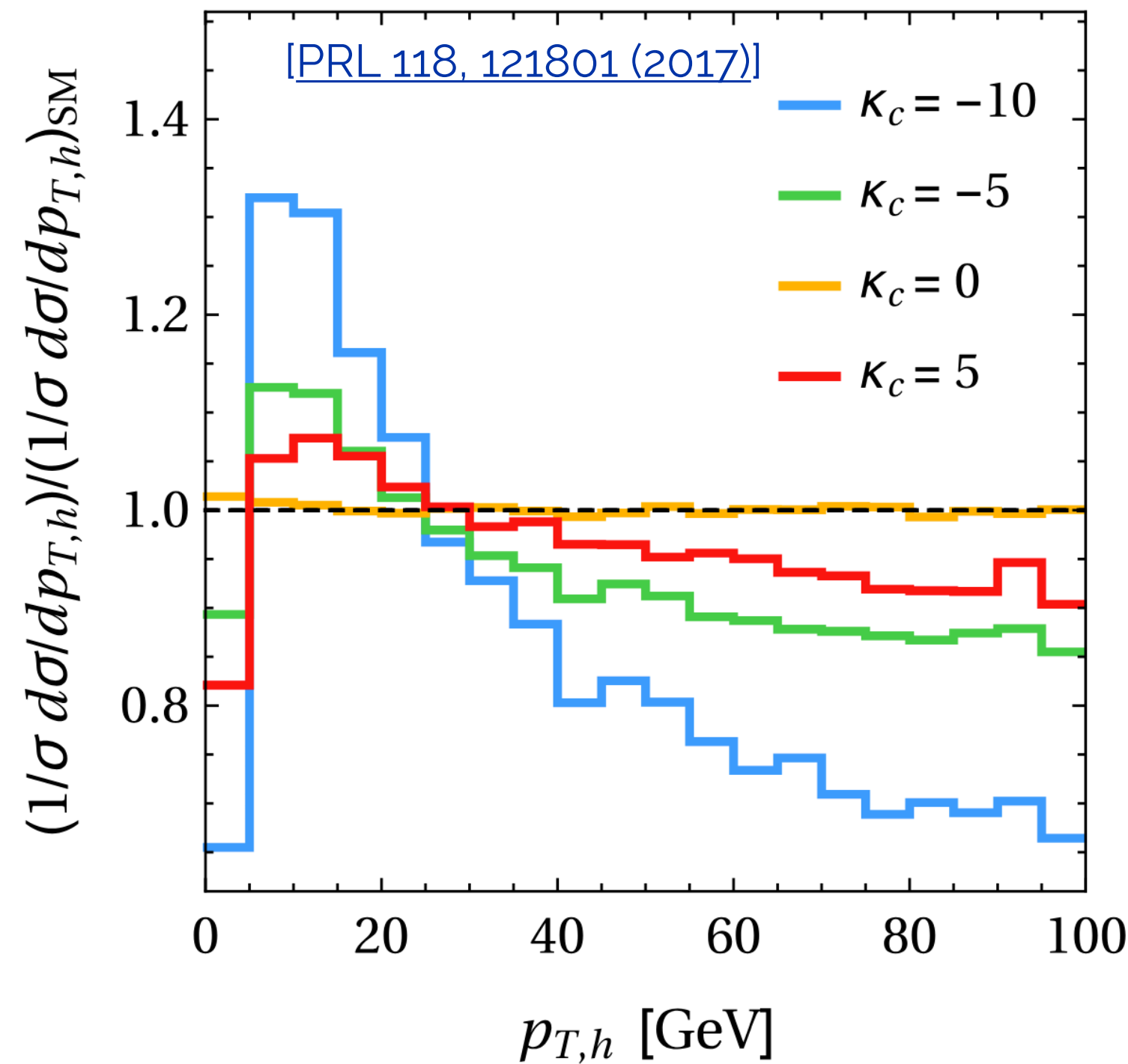
So far: probing the Higgs-charm coupling

- Several methods explored by ATLAS and CMS to probe the Higgs-charm Yukawa coupling (y_c)

Indirect constraint from Higgs kinematics

Model dependence?

Direct searches for $H \rightarrow c\bar{c}$

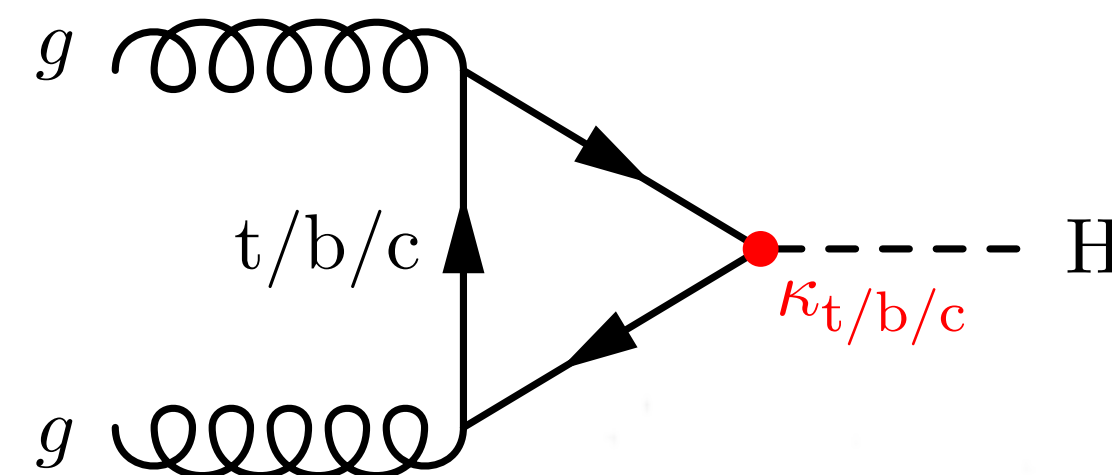


- Differential distributions are sensitive to y_c

Full Run 2 Higgs differential cross-section combination

[CMS-HIG-23-013]

$|\kappa_c| < 2$



Kappa framework

$$(\sigma \cdot \mathcal{B})(ii \rightarrow H \rightarrow ff) = \frac{\sigma_{ii} \cdot \Gamma_{ff}}{\Gamma_H} = \sigma_{ii}^{SM} \cdot \mathcal{B}_{ff}^{SM} \cdot \frac{\kappa_i^2 \kappa_f^2}{\kappa_H^2}$$

κ_i : coupling wrt. SM expectation of 1

So far: probing the Higgs-charm coupling

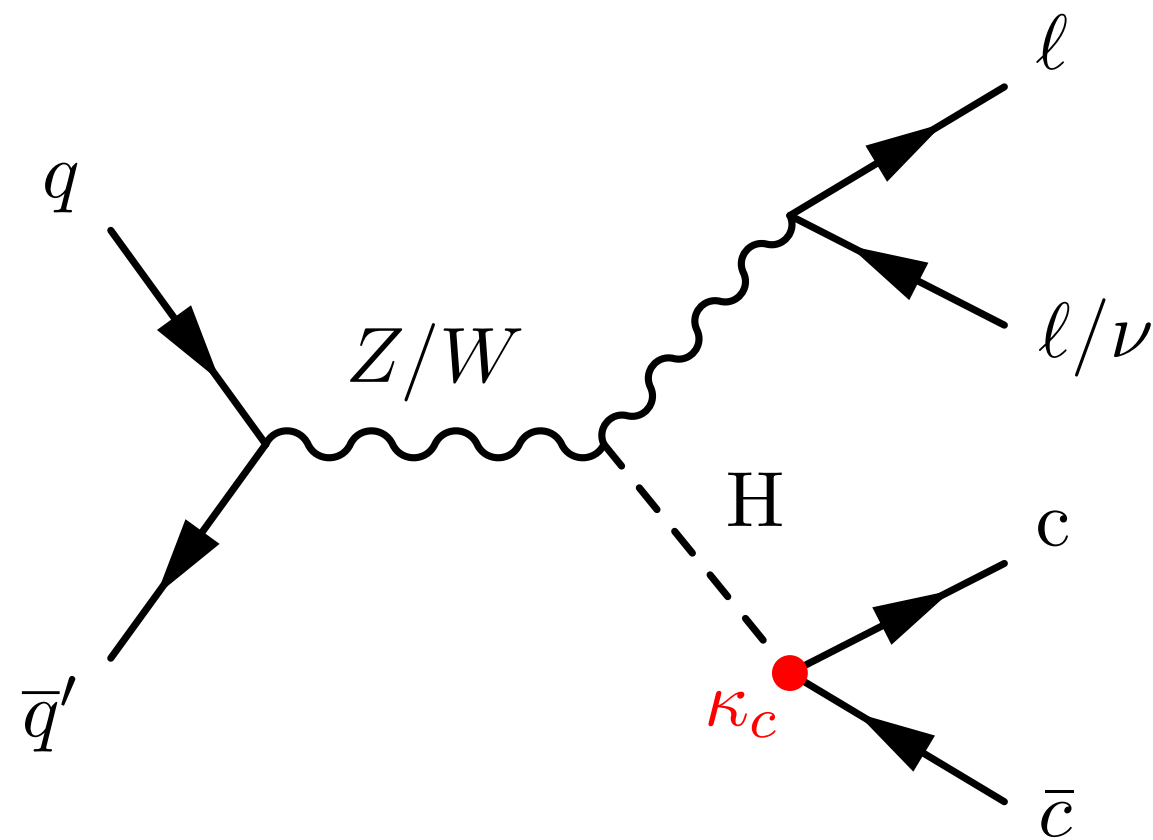


- Several methods explored by ATLAS and CMS to probe the Higgs-charm Yukawa coupling (y_c)

Indirect constraint from Higgs kinematics

Model dependence?

Direct searches for $H \rightarrow c\bar{c}$

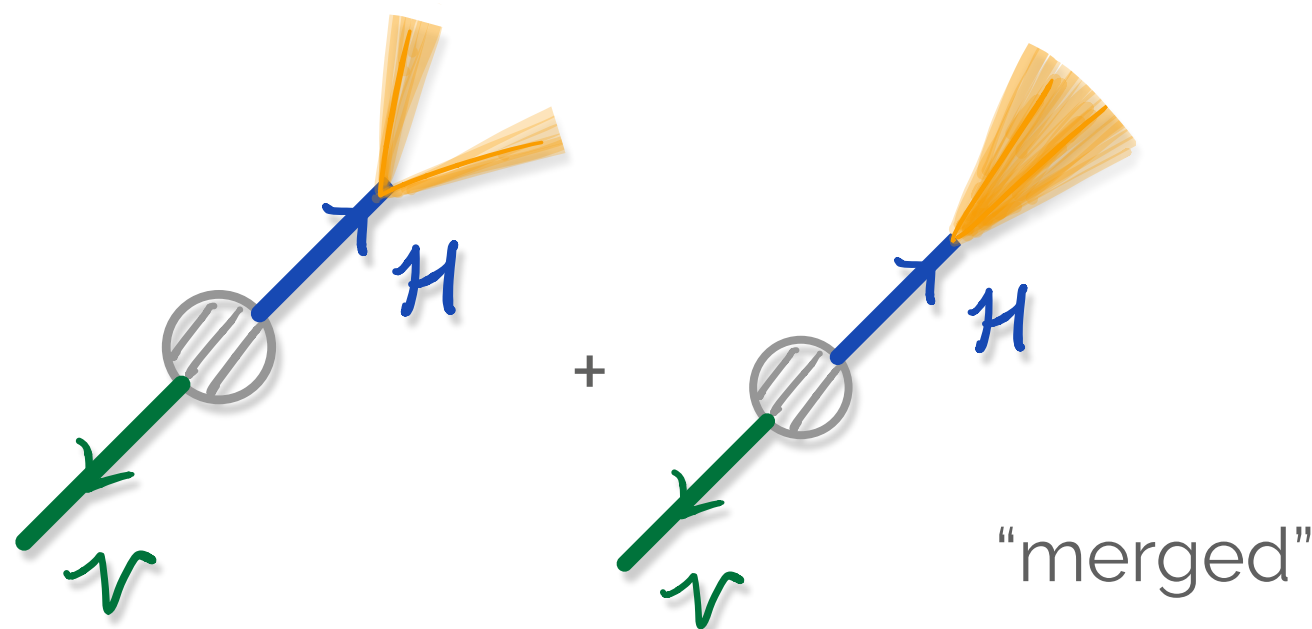


VH as the “golden channel”

$$1.1 < |\kappa_c| < 5.5 (< 3.4)$$

$$\mu < 14 (7.6)$$

“resolved”



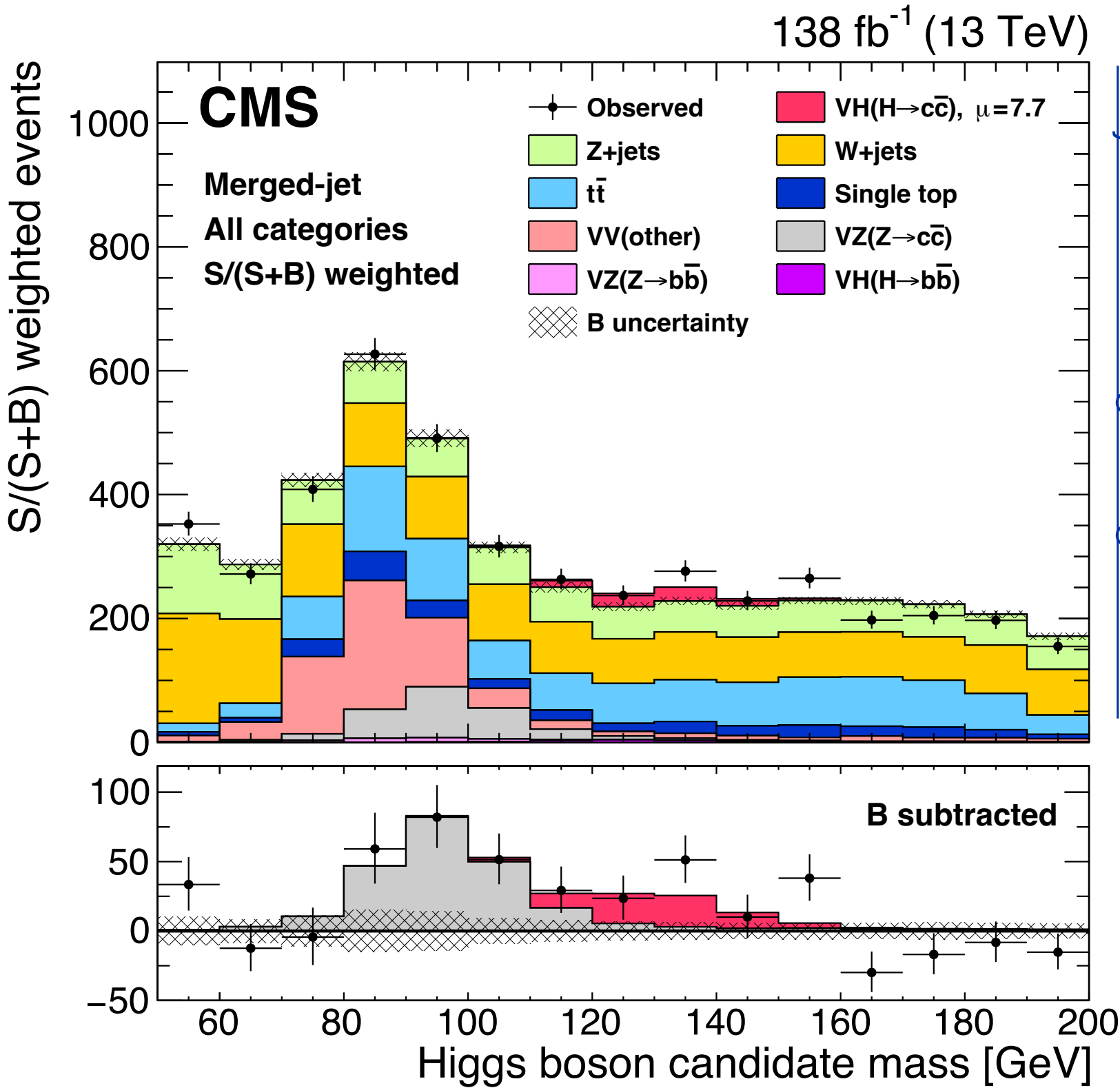
From ATLAS:

$$|\kappa_c| < 4.2 (< 4.1)$$

$$\mu < 12 (11)$$

[JHEP 04(2025) 075]

We will come back to this..



[Phys. Rev. Lett. 131 (2023) 061801]

Which tools do we have?

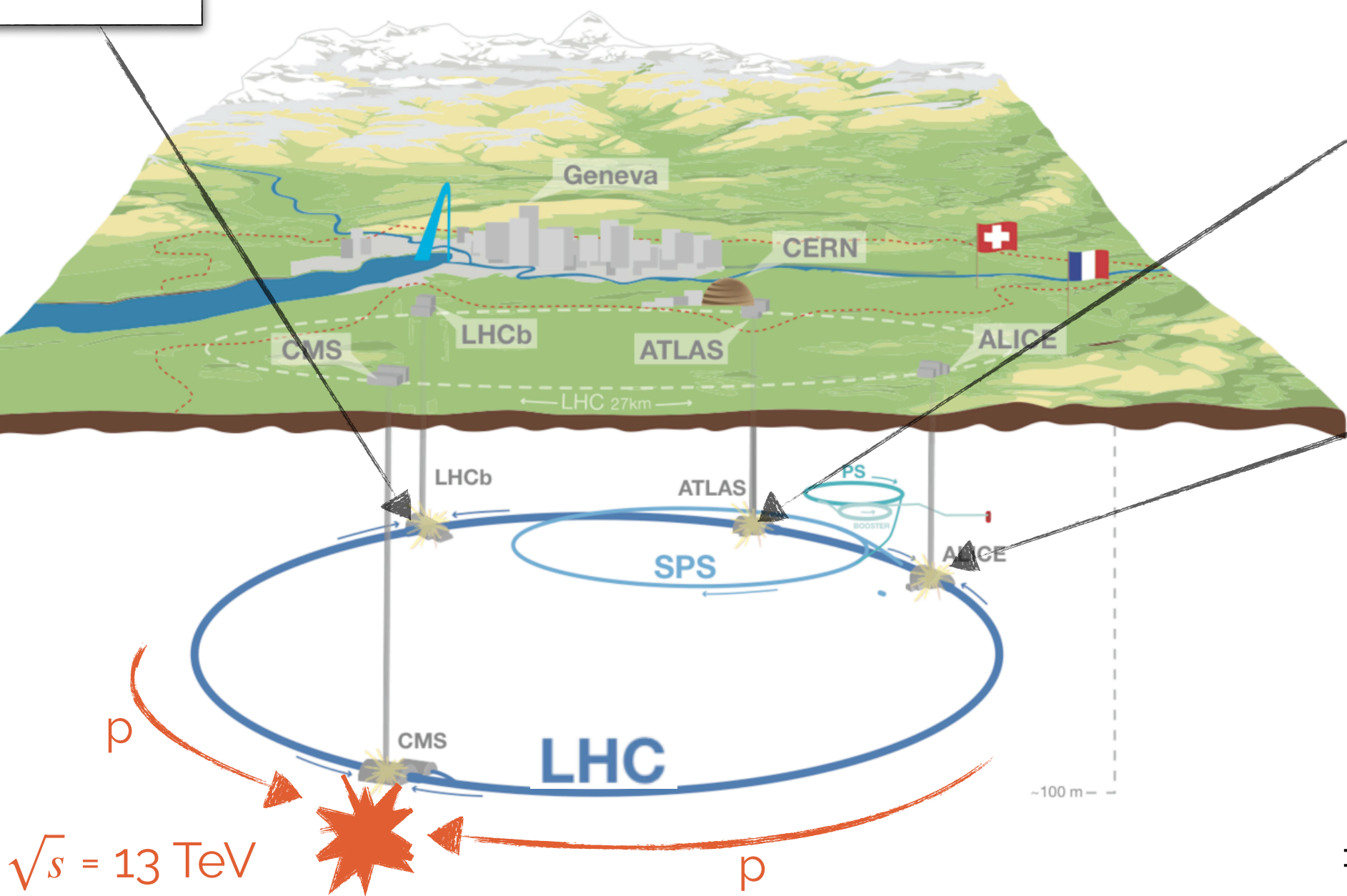
The Large Hadron Collider (LHC)

LHCb: b quark physics

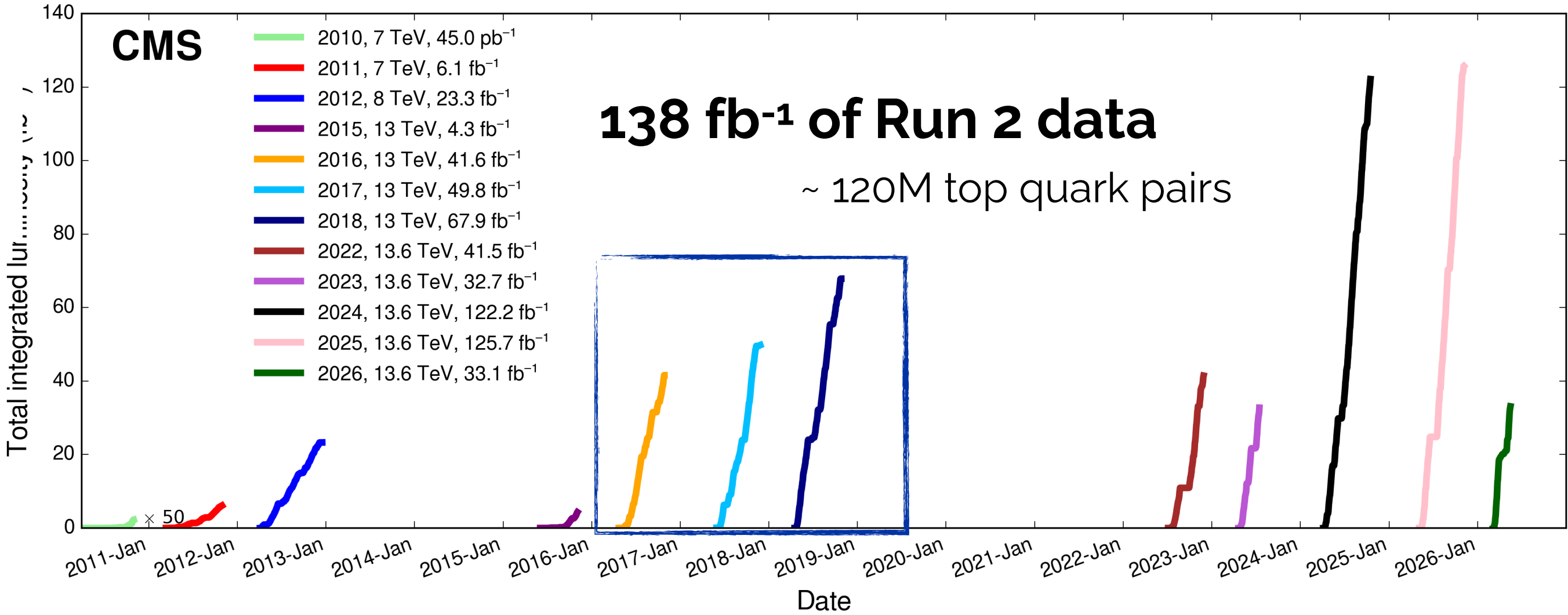
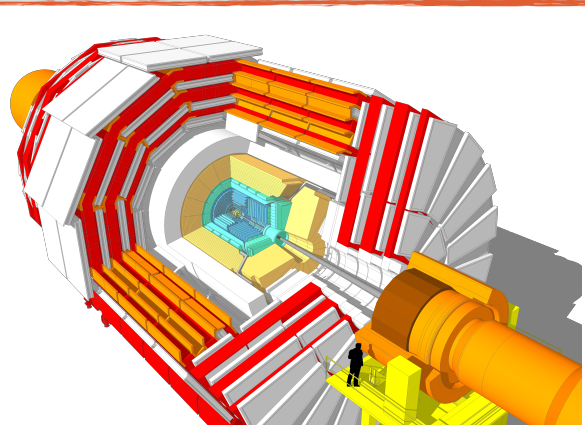
ATLAS: general purpose experiment

ALICE: heavy-ion physics

... just 1/3 of the Run 2 + 3 data

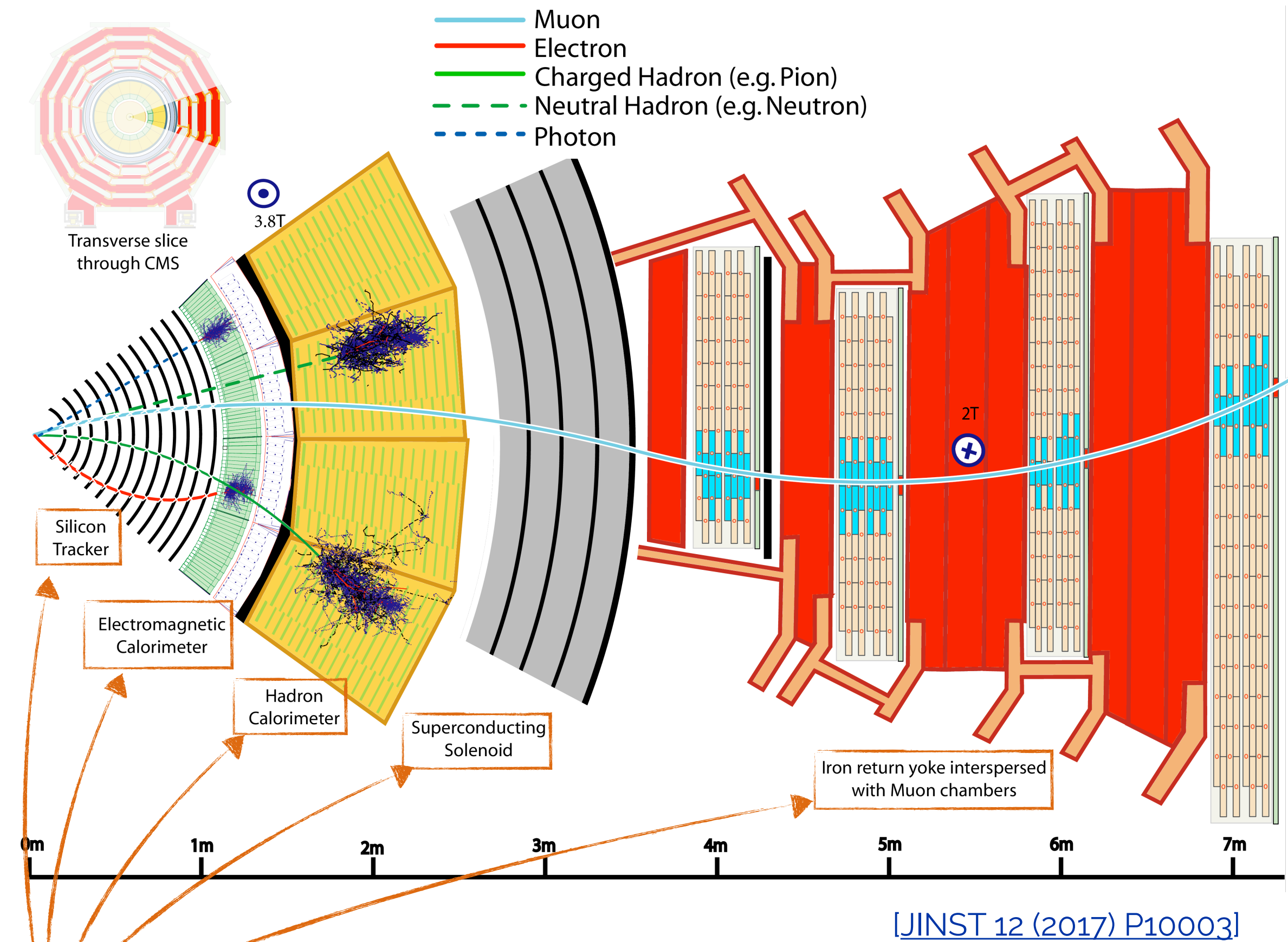


CMS: general purpose experiment



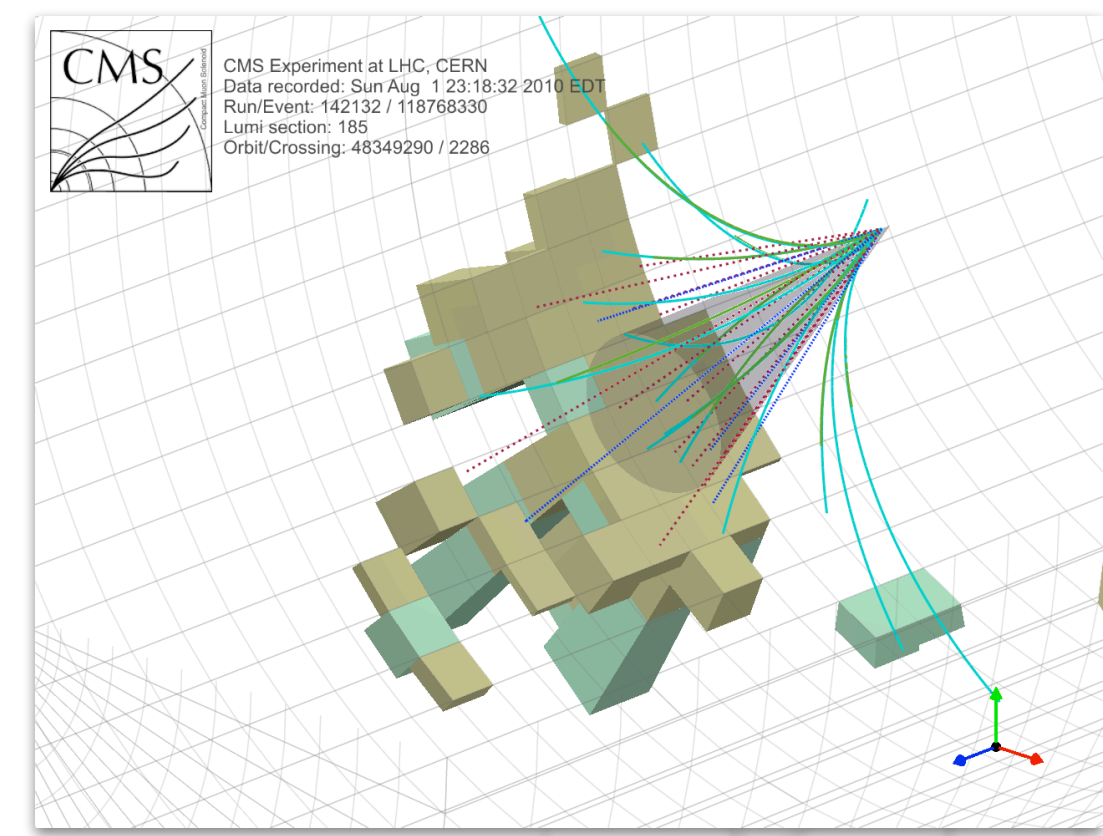
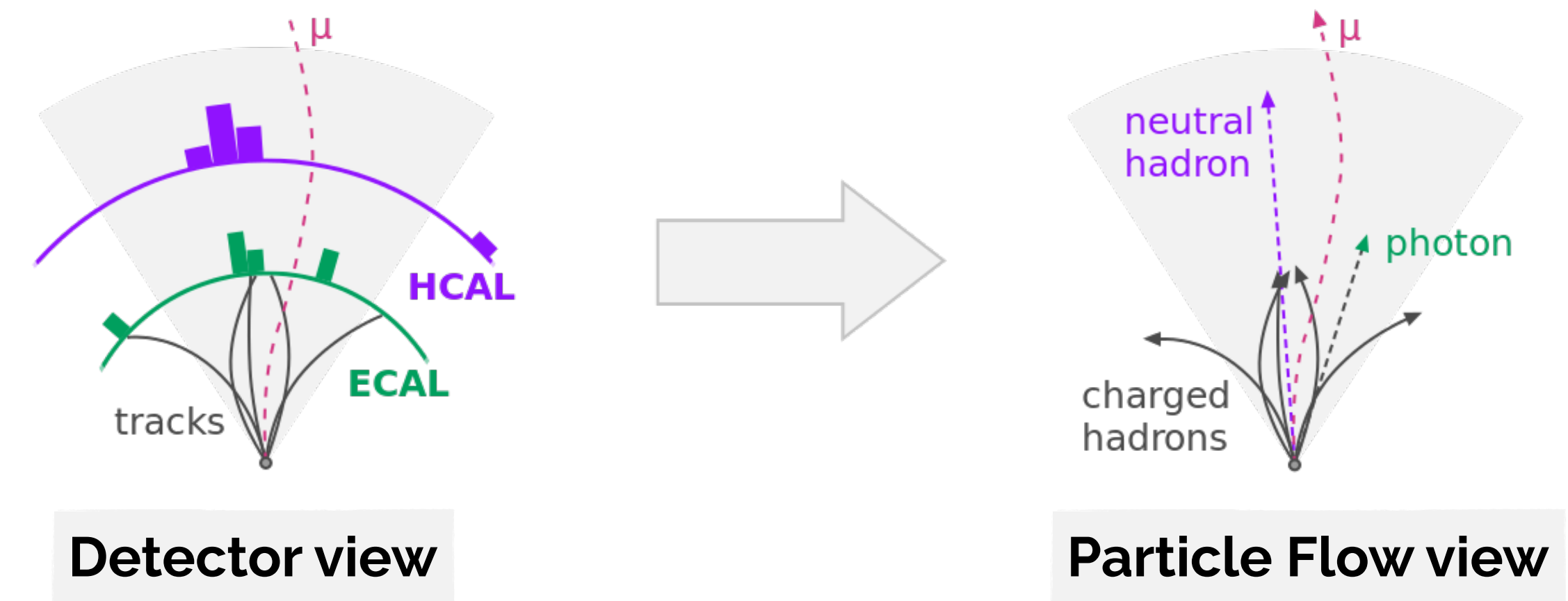
- Underground superconducting accelerator in Geneva at CERN
- Hadron-hadron collisions at the high-energy frontier, up to 14 TeV

Which tools do we have?



- **Particle flow reconstruction:**
 - Combined particle reconstruction with detailed properties
 - Excellent energy and momentum resolution

The CMS experiment



Jet reconstruction

- Clustering collimated sprays of particles as jets, initiated by partons
- Typical $R=0.4$, Anti k_T

Now let's do some analysis!

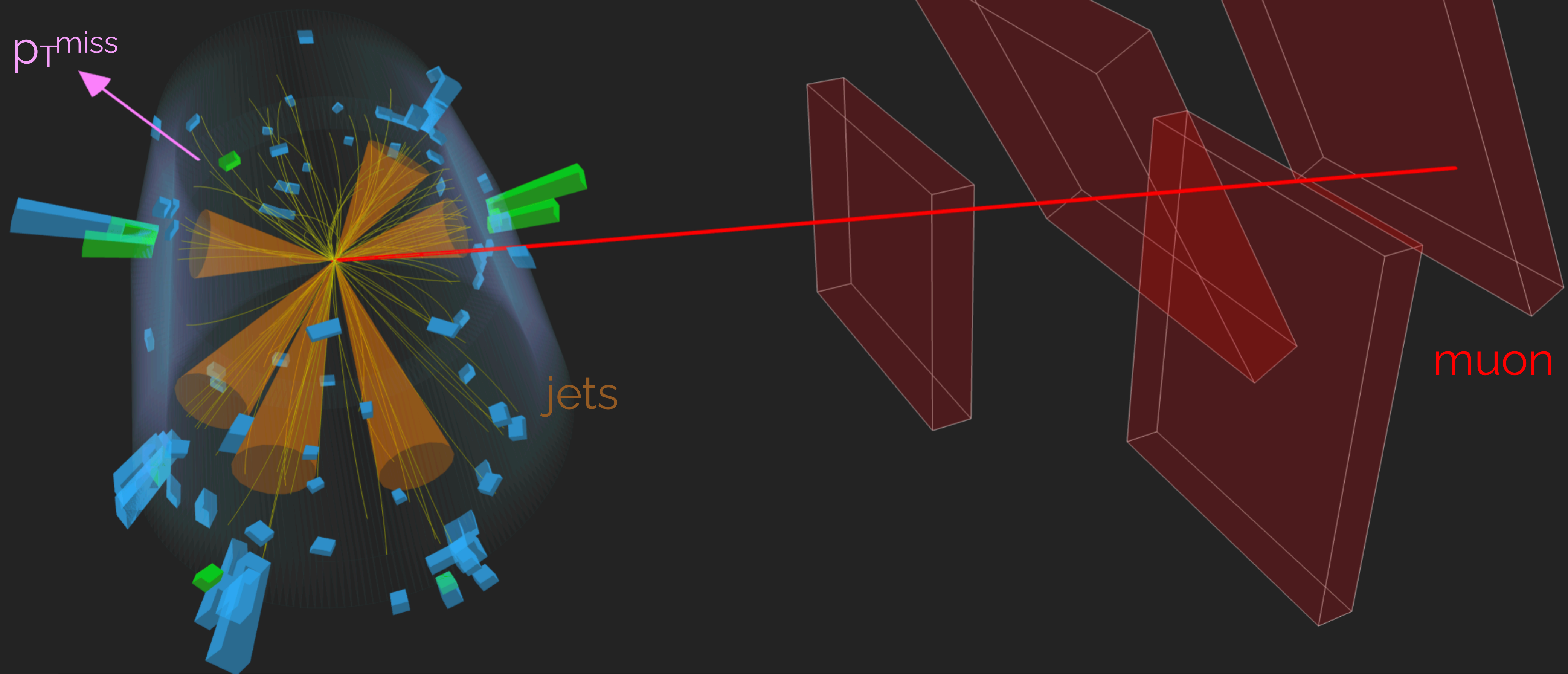


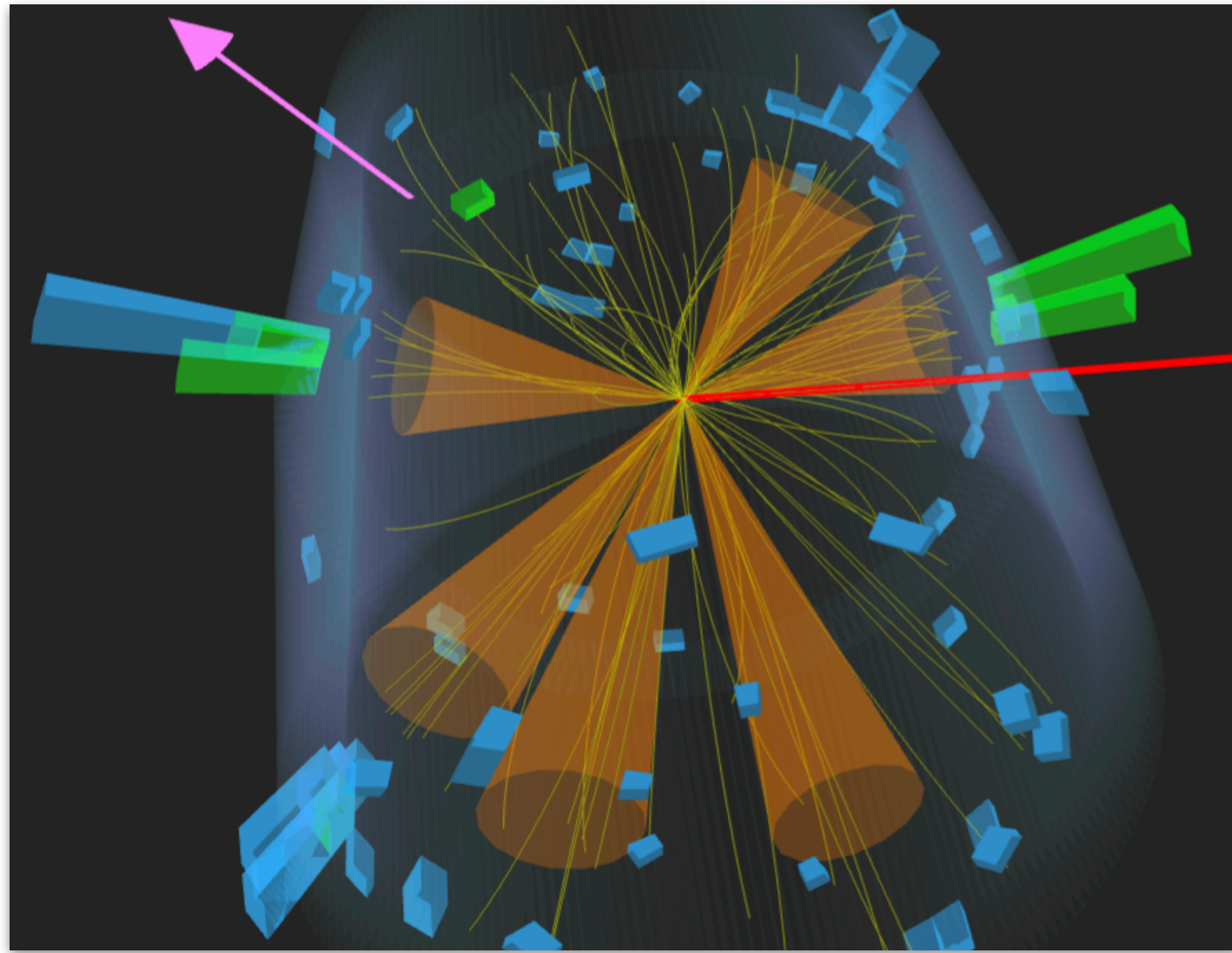
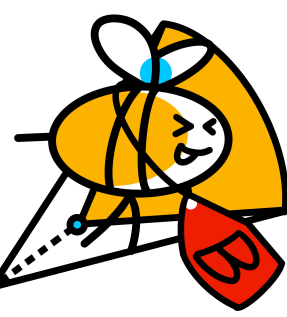
CMS Experiment at the LHC, CERN

Data recorded: 2018-Oct-01 01:05:14.054528 GMT

Run / Event / LS: 323778 / 1280408227 / 680

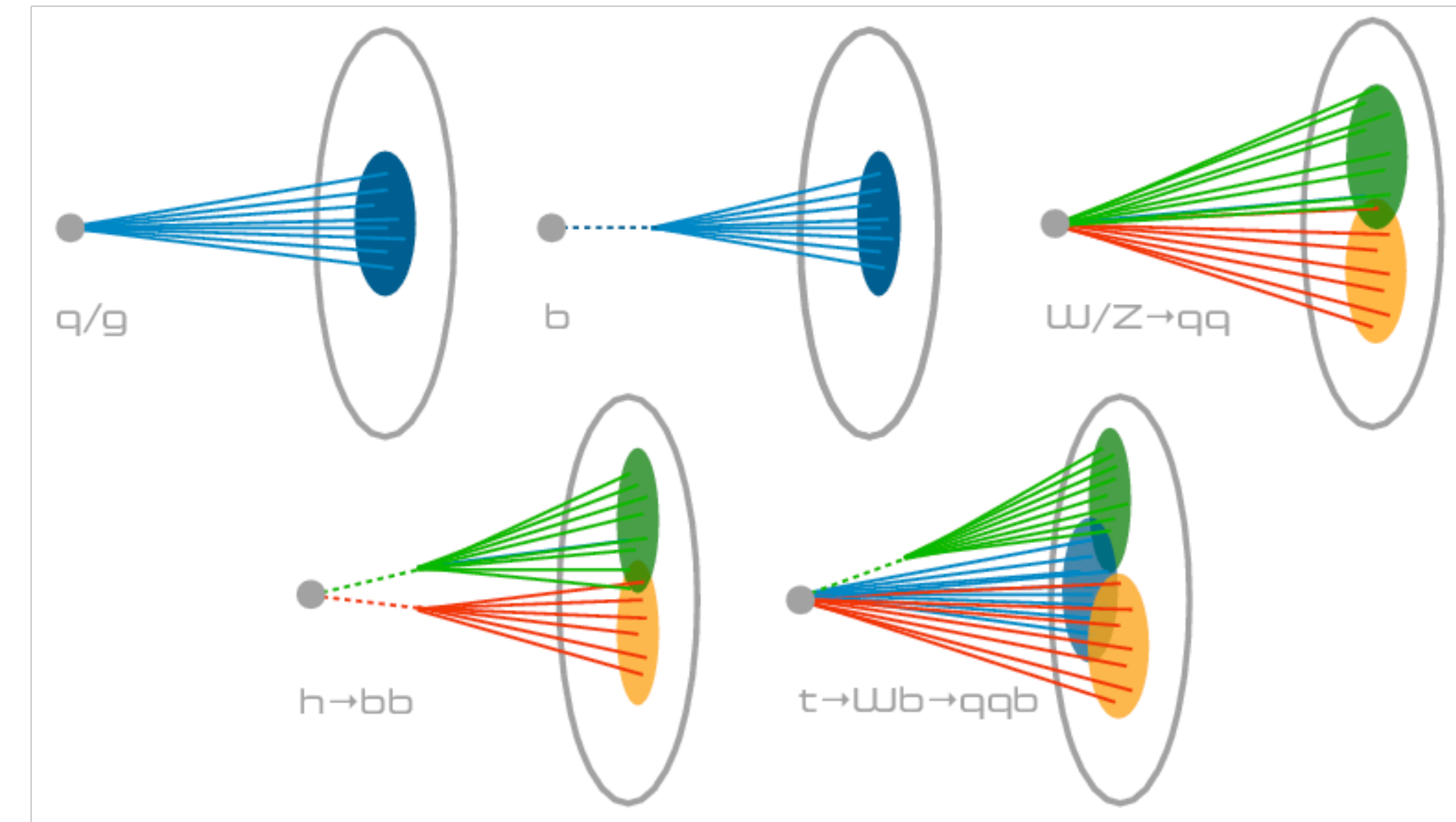
$H \rightarrow c\bar{c}$ candidate event





What initiated the jets?

Resolved jet tagging



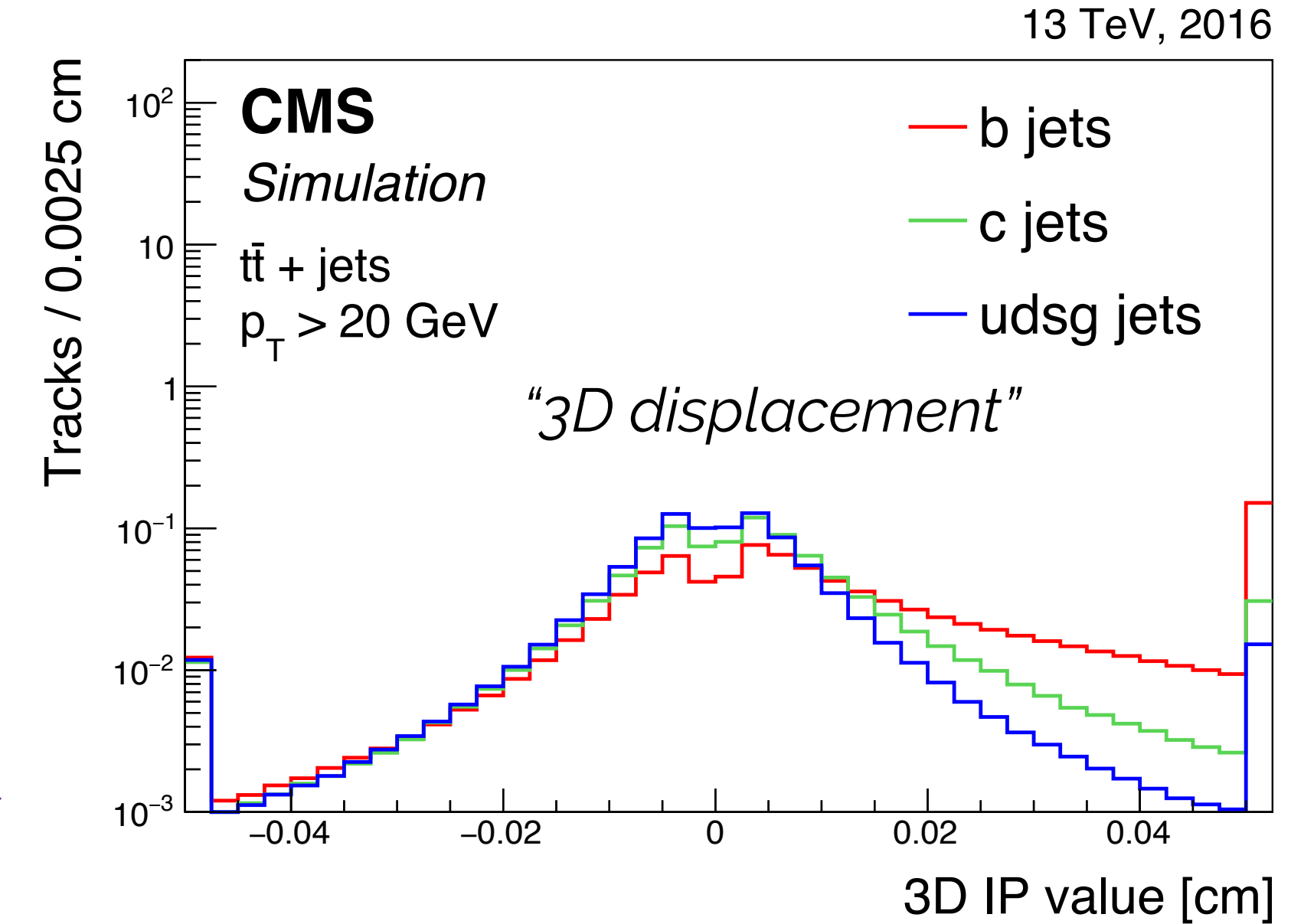
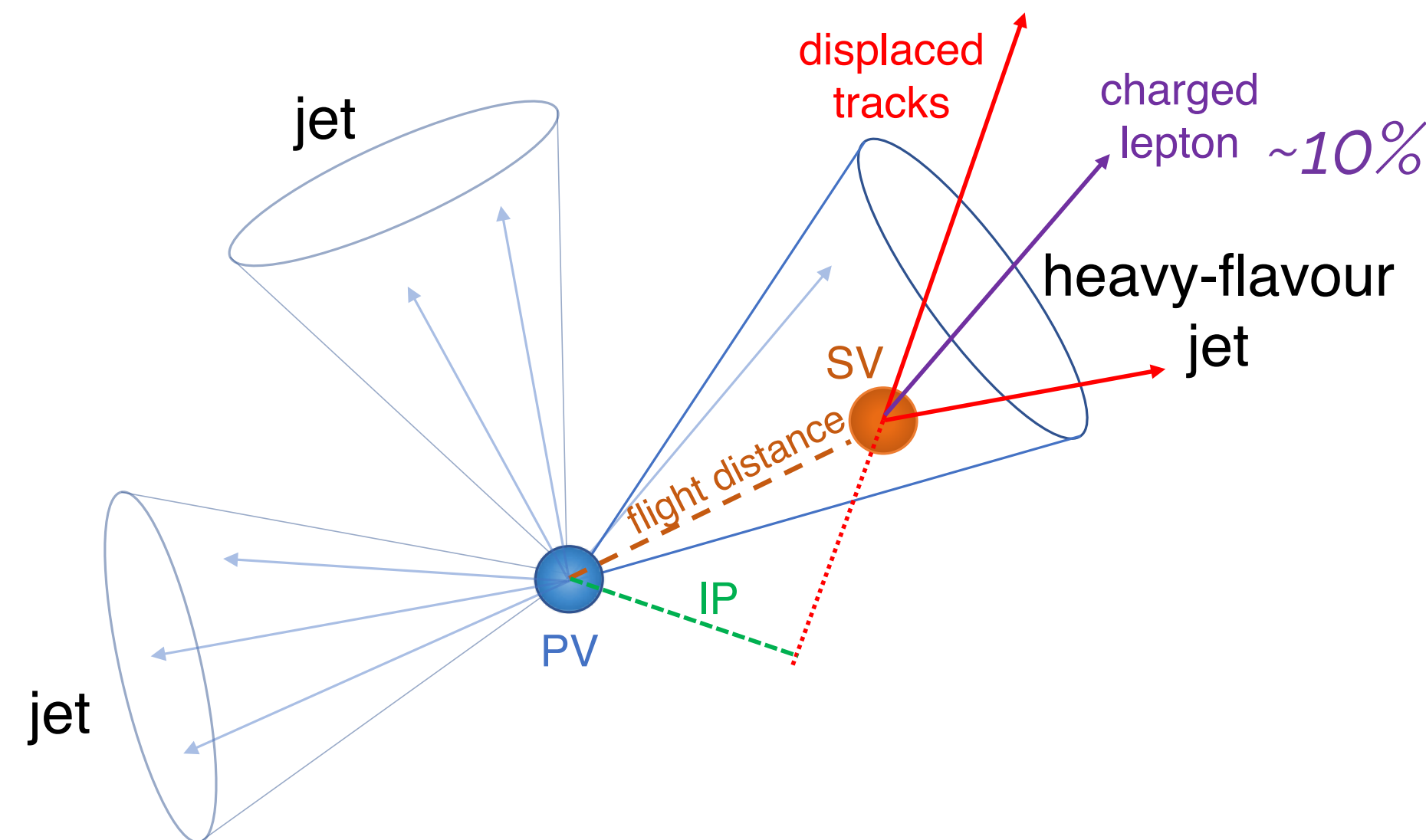
Boosted jet tagging

- Fragmentation structure depends on incoming parton type
- B/C hadrons have a larger lifetime: displaced tracks / secondary vertices (SV)
- Presence of non-isolated leptons in jets
- Jet-internal substructure: prongs

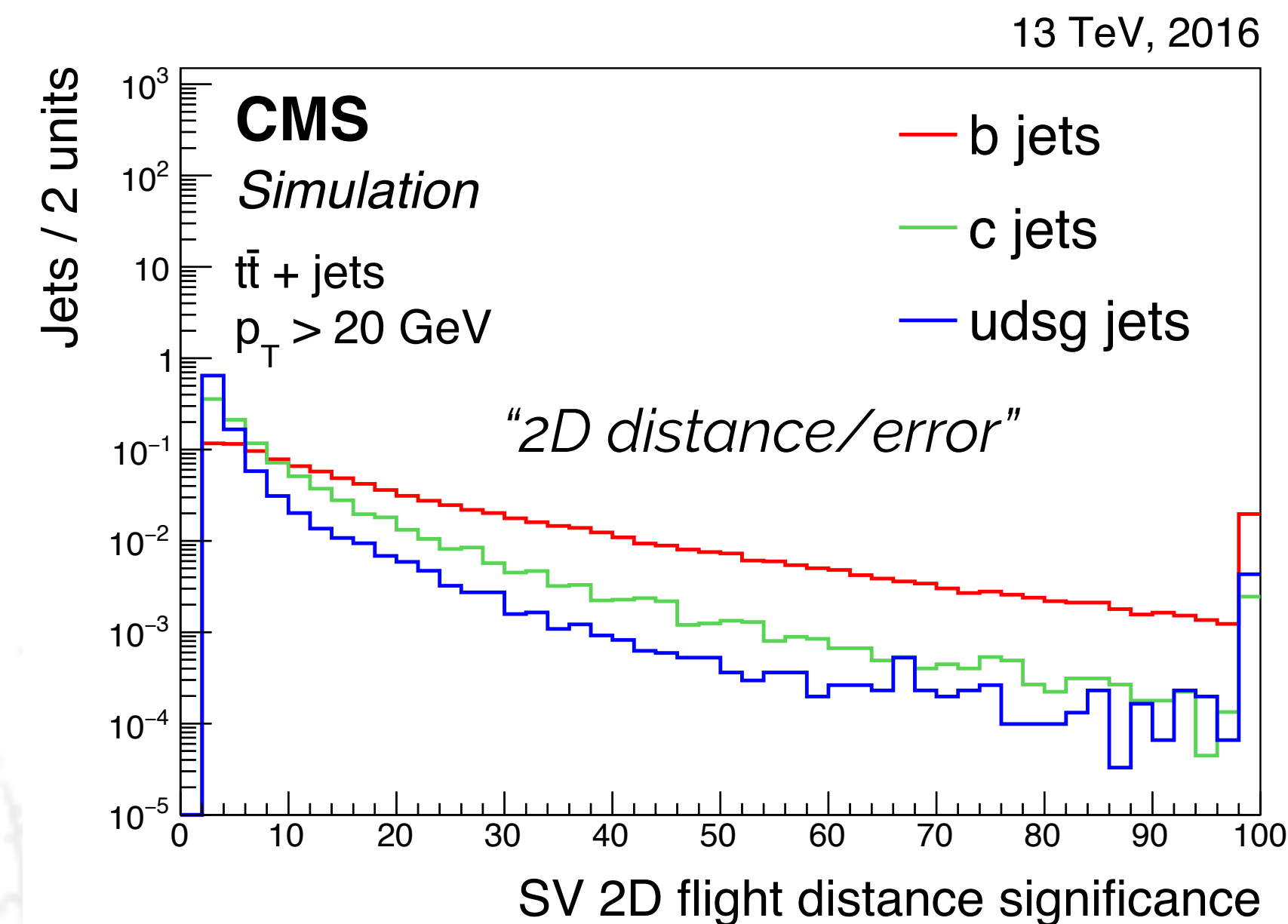
→ **CMS Particle Flow provides perfect input for ML-based jet tasks!**

Jet flavour tagging

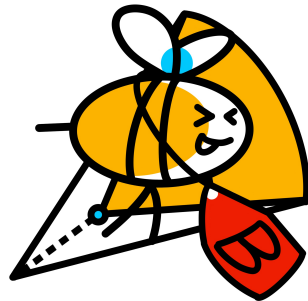
- Direct $H \rightarrow c\bar{c}$ searches only as good as the identification of real **c-jets**!
 - c-jets have intermediate properties to b- and light jets
→ A simultaneous 2D discrimination is needed
- light** ↔ **c** ↔ **b**
- Exploit low-level track and secondary vertex features in ML-algorithms
 - Significant improvement hand-in-hand with ML developments



[\[JINST 13 \(2018\) P05011\]](#)



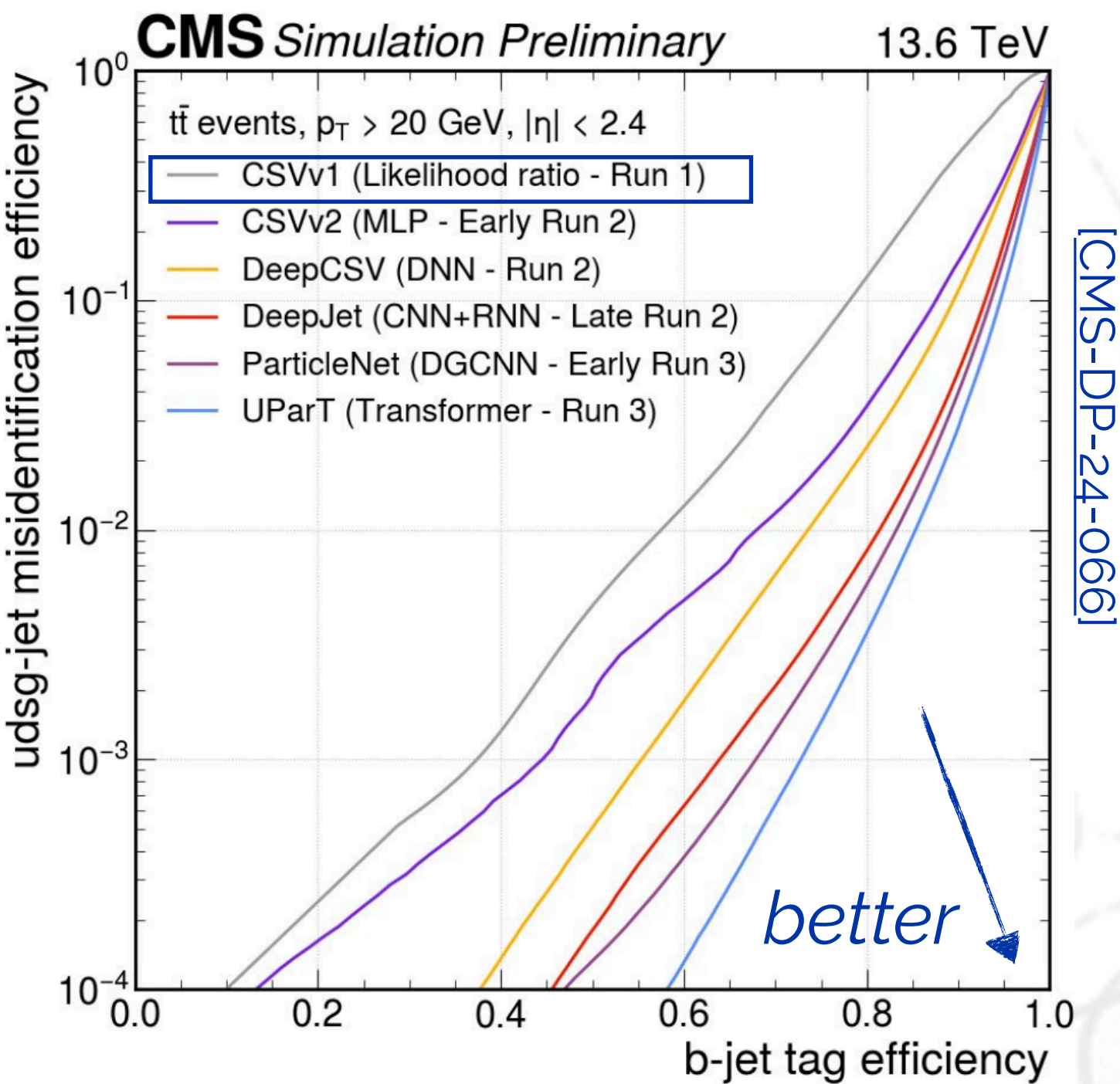
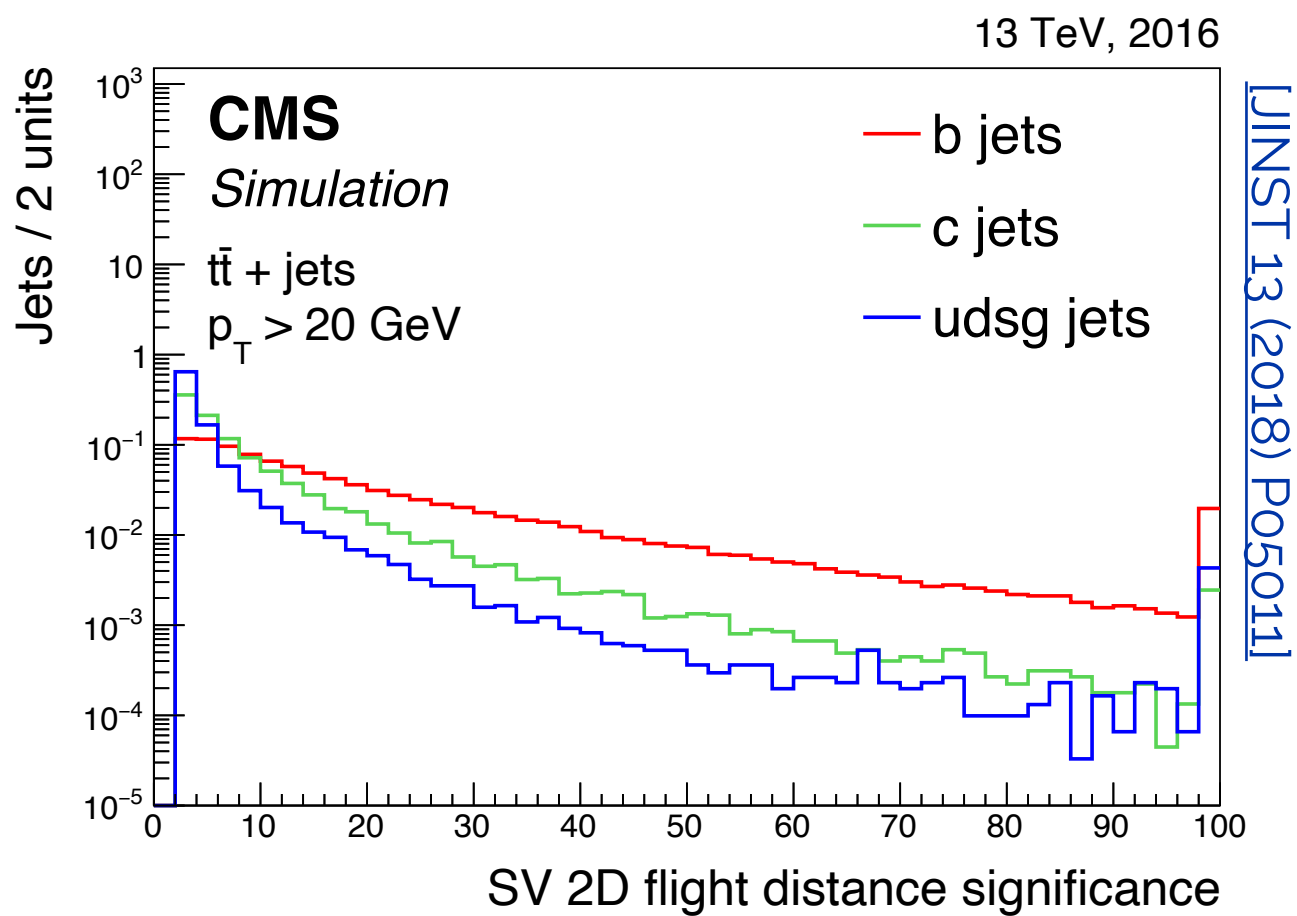
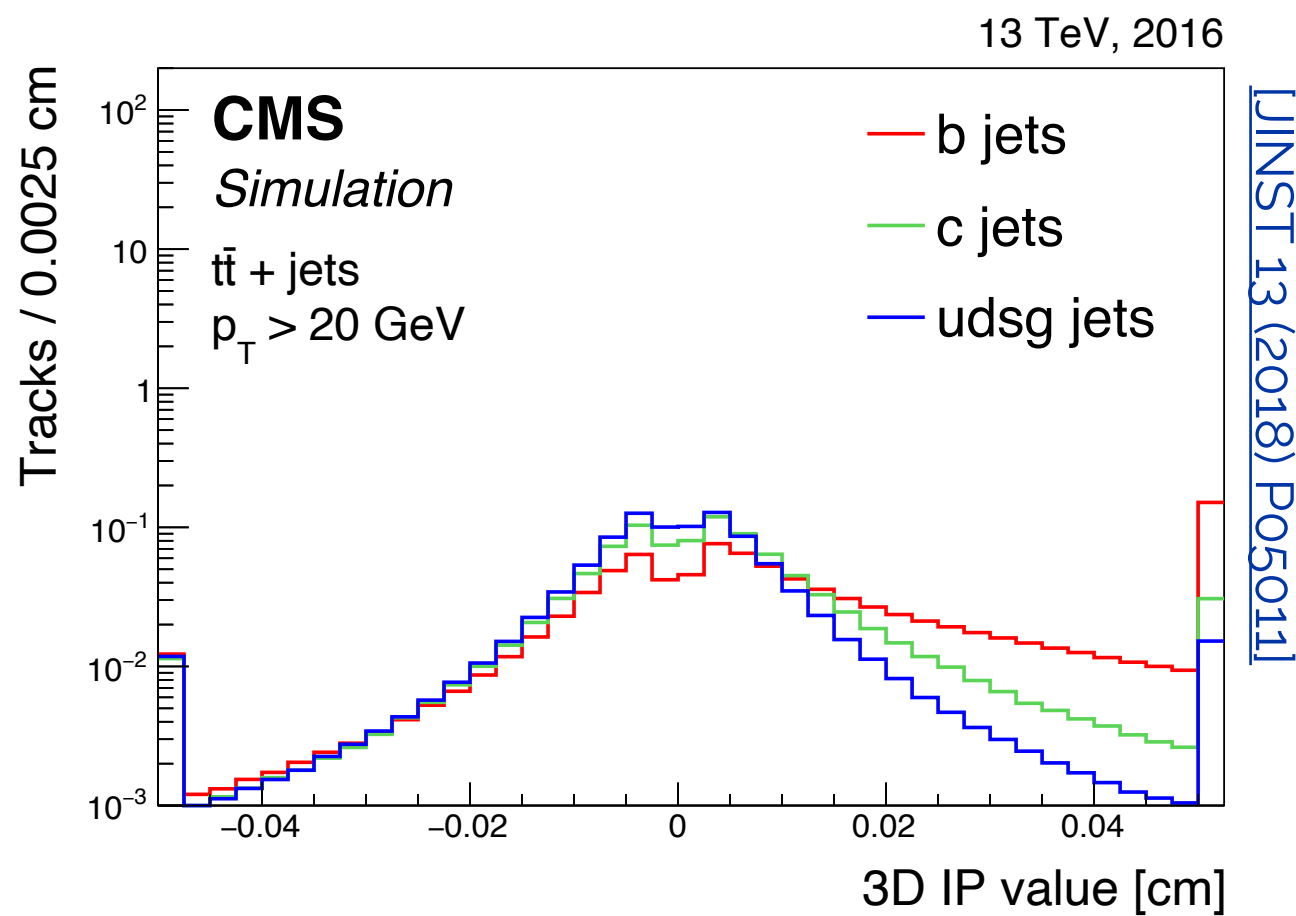
A bit of CMS history...



Run 1 (2011 - 2016)

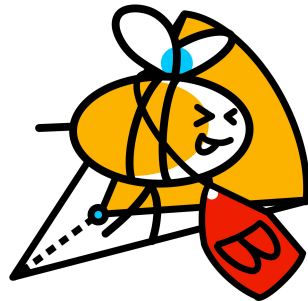
Jet probabilities

- Likelihood based sum over track-from-PV probabilities
- Combination of track/SV properties “by hand”



A bit of CMS history...

Run 2 (2016 - 2018)



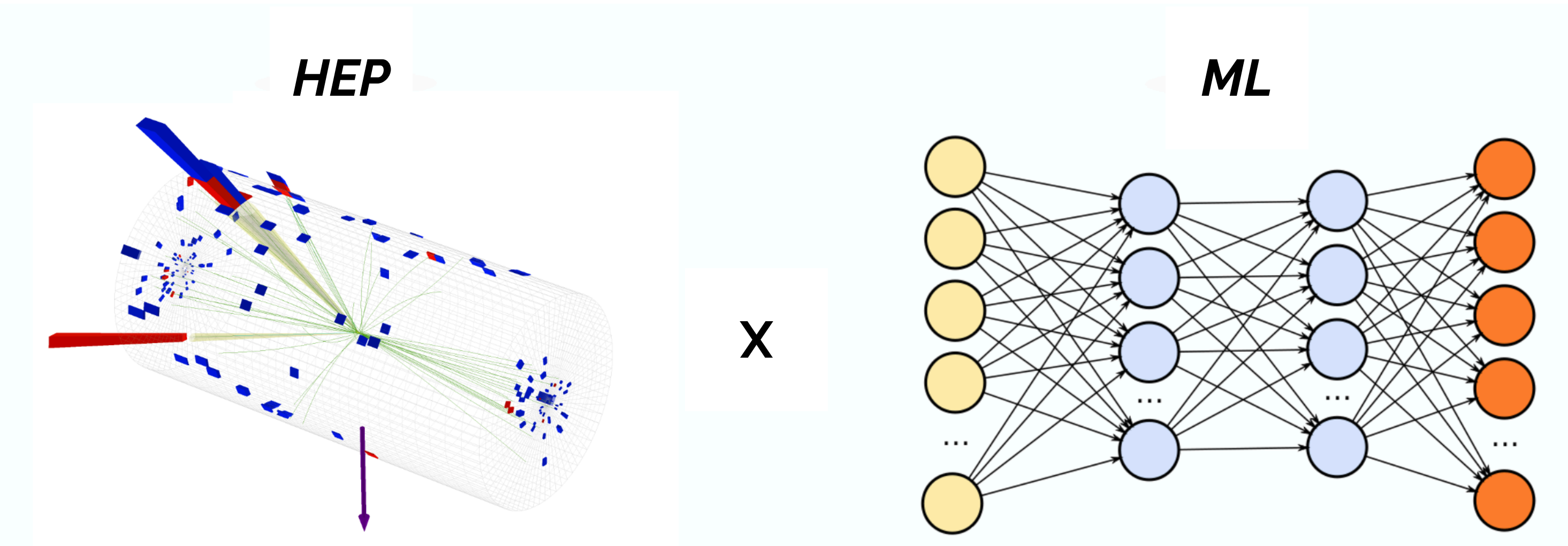
Run 1 (2011 - 2016)

Machine Learning

Jet probabilities

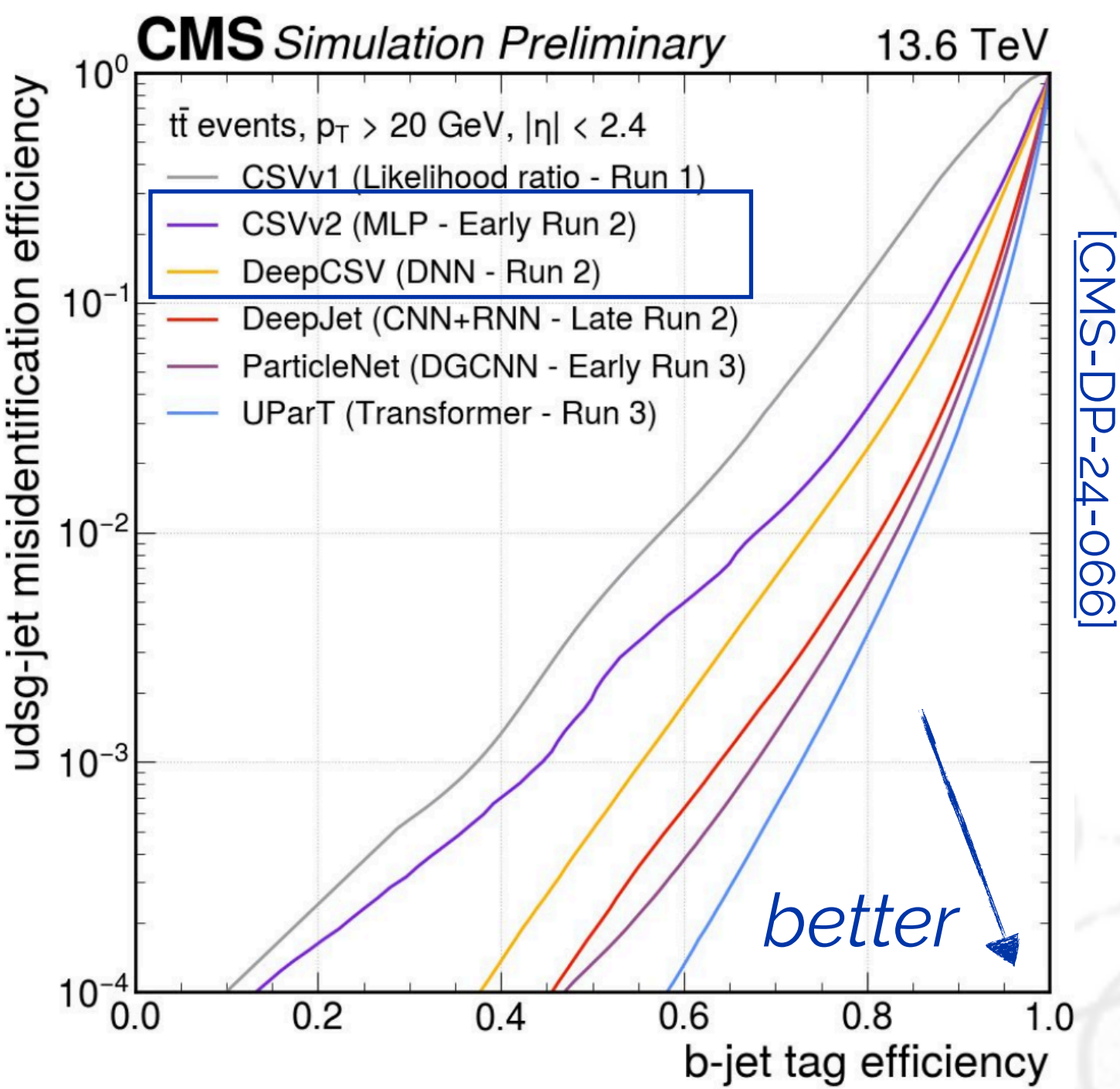
CSV/DeepCSV
Sub-jet tagging

- How do we represent the data?



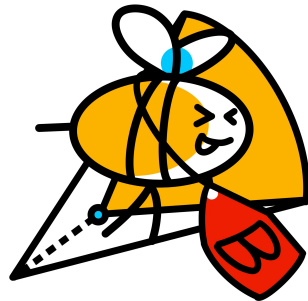
Collisions, detector events, sensor arrays, ...

- The traditional way: summarizing jets into a few high-level features
 - p_T , η , ϕ , mass, ...
- Then use boosted decision trees or fully-connected neural networks



A bit of CMS history...

Run 2 (2016 - 2018)



Run 1 (2011 - 2016)

Jet probabilities

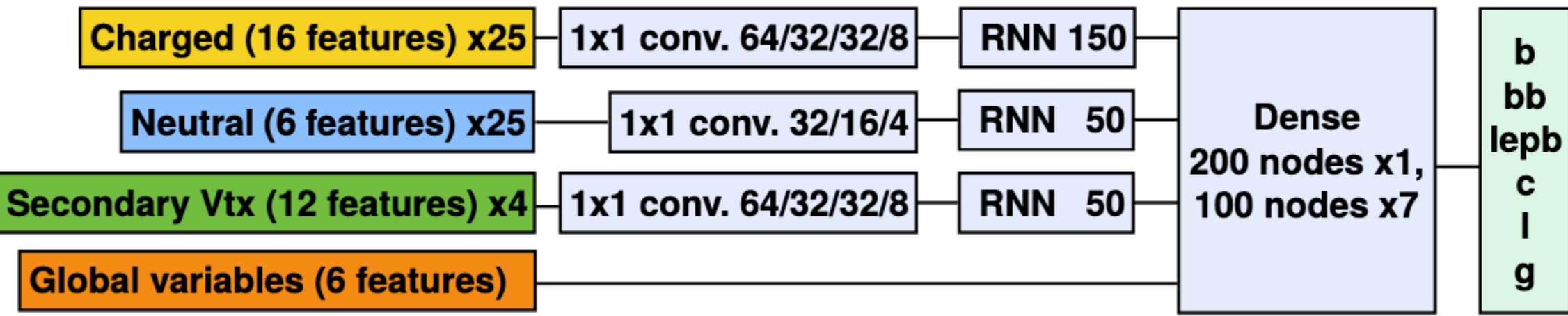
Machine Learning

CSV/DeepCSV
Sub-jet tagging

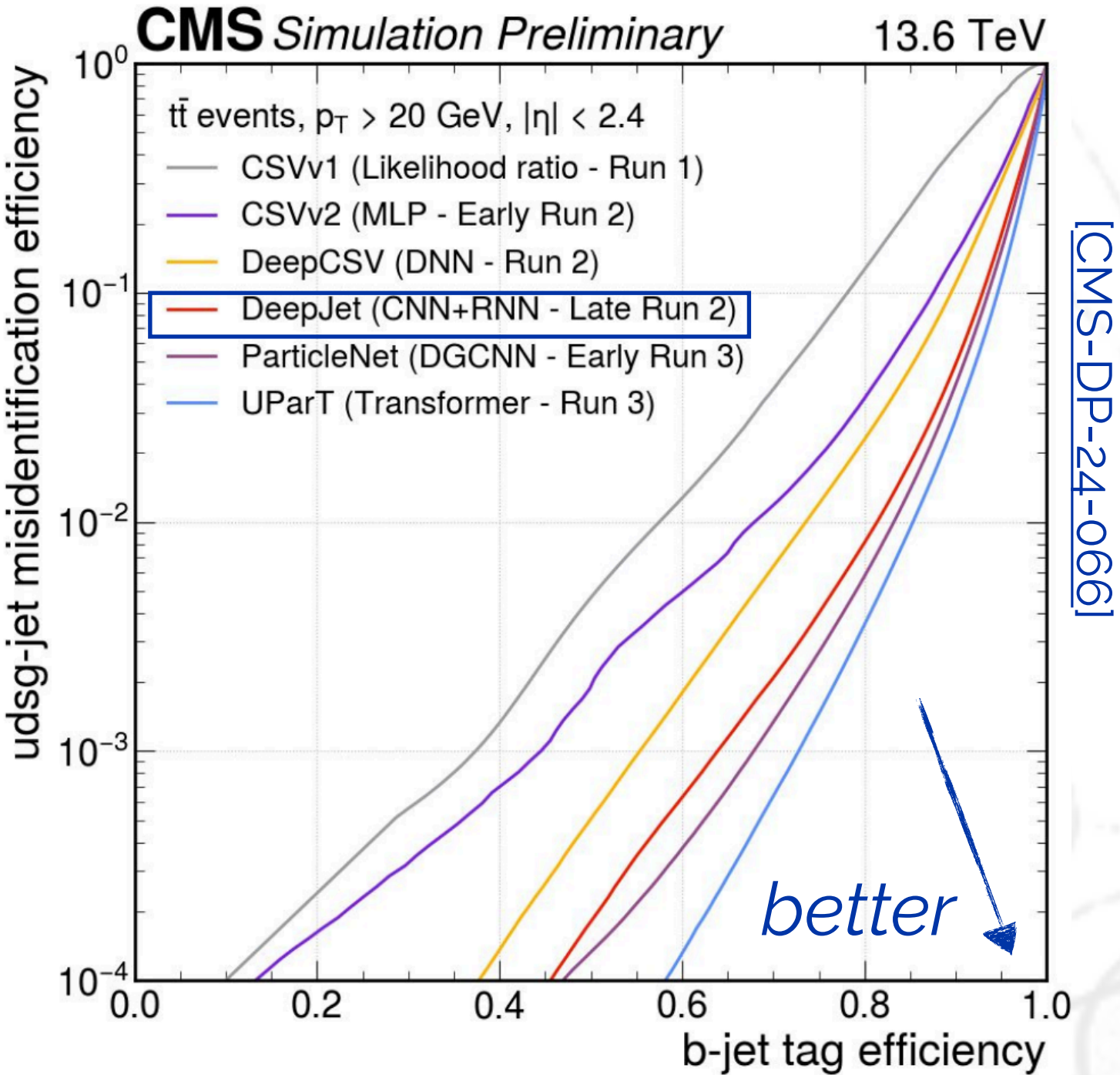
Jets as a sequence

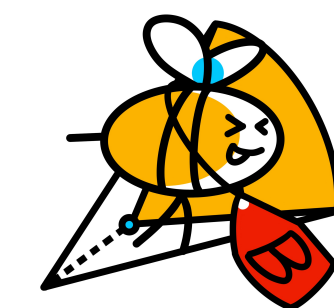
DeepJet/
DeepAK8

- A jet is a set of O(50-100) particles
→ Sequence as in language processing
- First models to focus on jet constituents: RNN+LSTM



[Bols et al (2020)]





A bit of CMS history...

Run 2 (2016 - 2018)

Run 1 (2011 - 2016)

Machine Learning

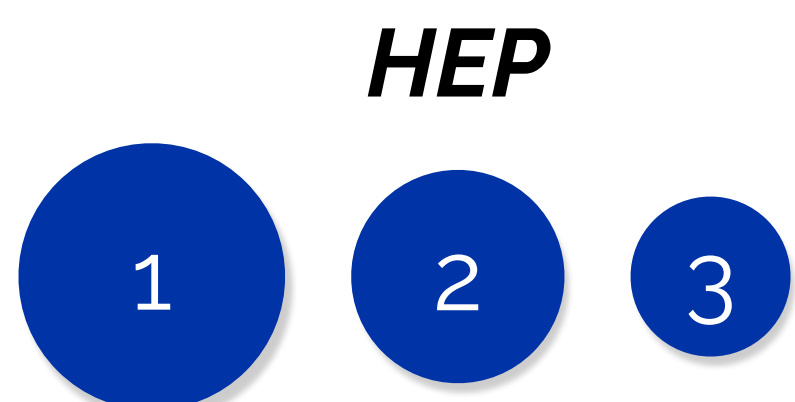
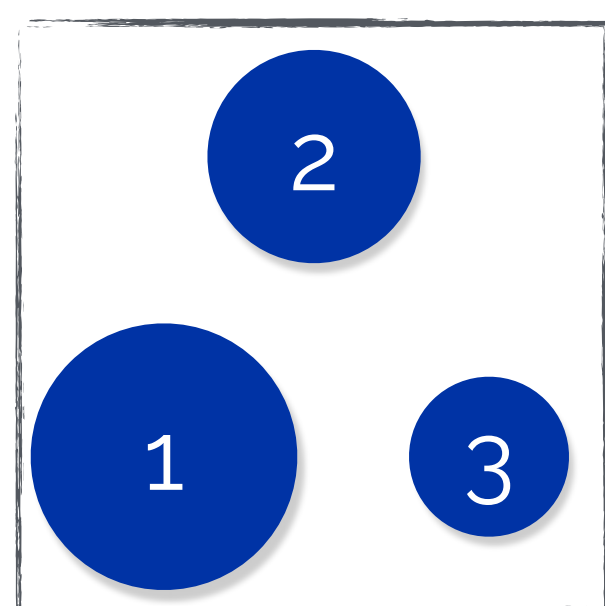
Jets as a sequence

Jet probabilities

CSV/DeepCSV
Sub-jet tagging

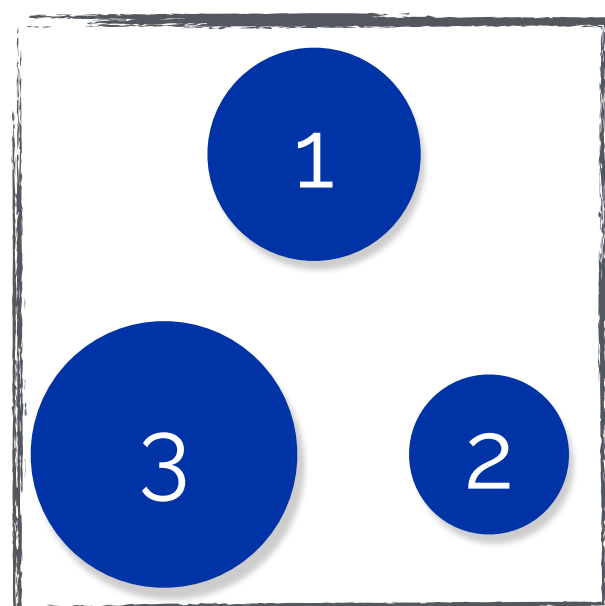
DeepJet/
DeepAK8

- Ordering of the sequence? We must impose a human inspired ordering (energy, displacement)...
- Jet is intrinsically an un-ordered set of particles!**



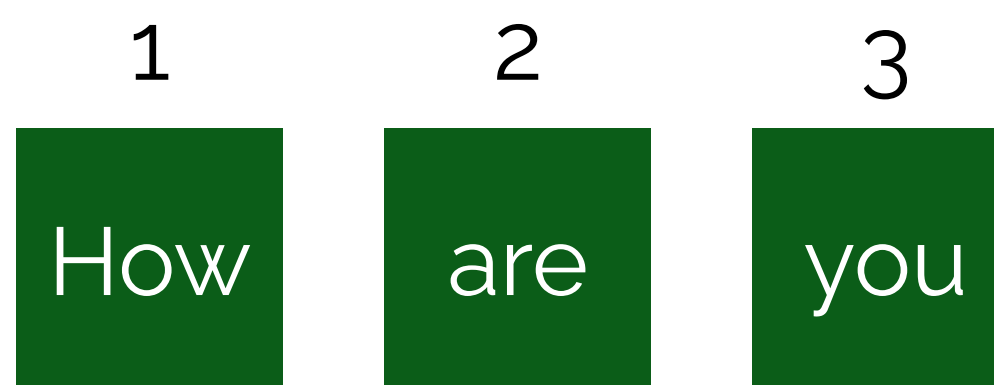
HEP

=

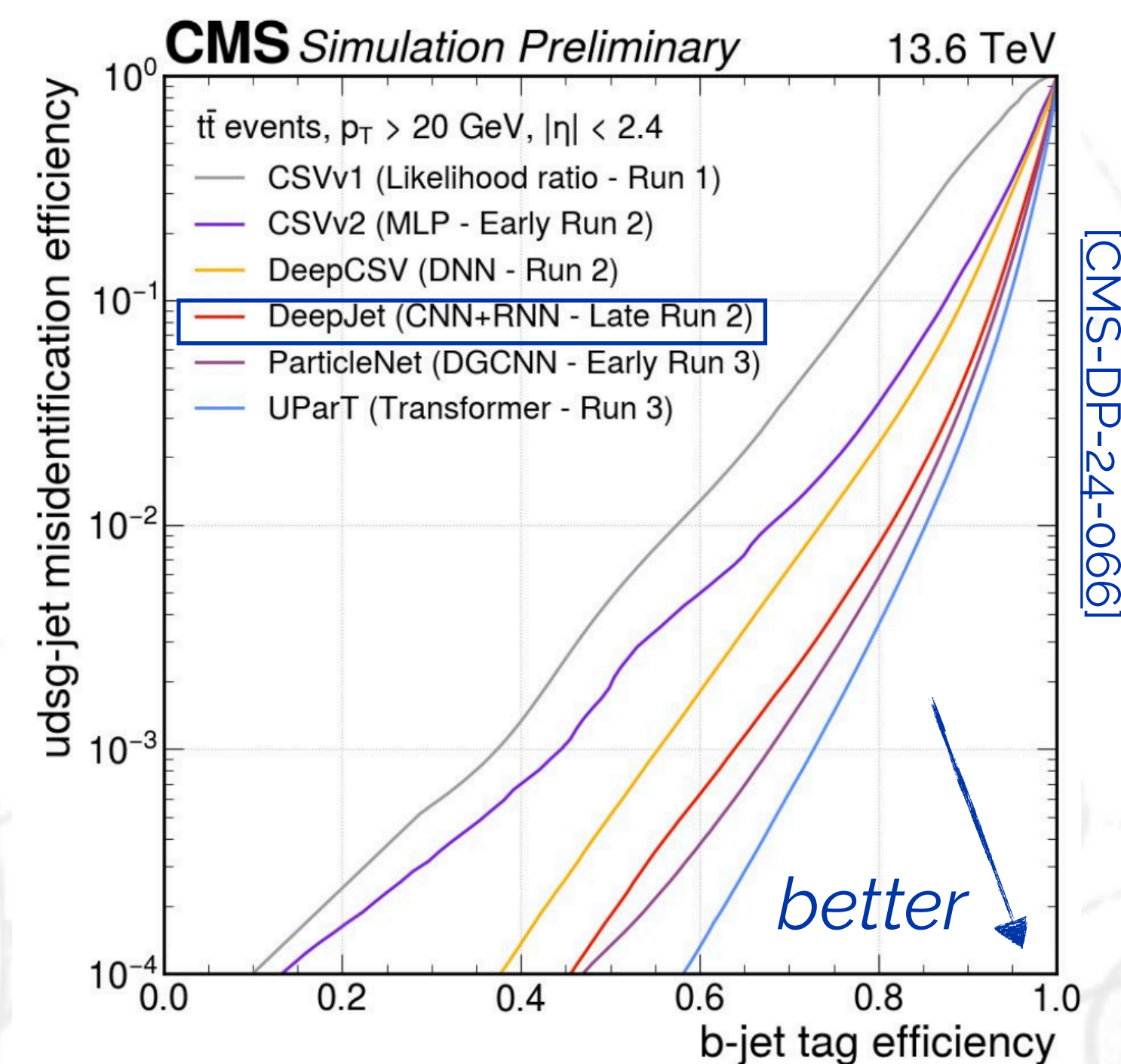
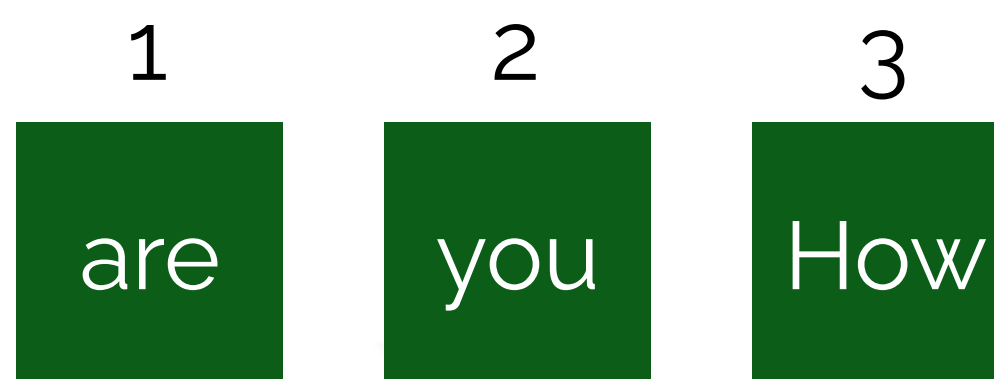


Permutation symmetry

Sequence

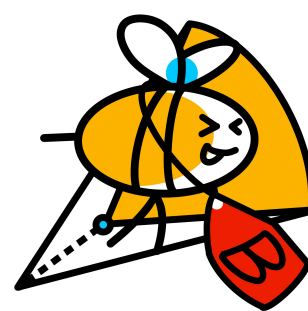


≠



A bit of CMS history...

Run 2 (2016 - 2018)



Run 1 (2011 - 2016)

Jet probabilities

Machine Learning

CSV/DeepCSV
Sub-jet tagging

Jets as a sequence

DeepJet/
DeepAK8

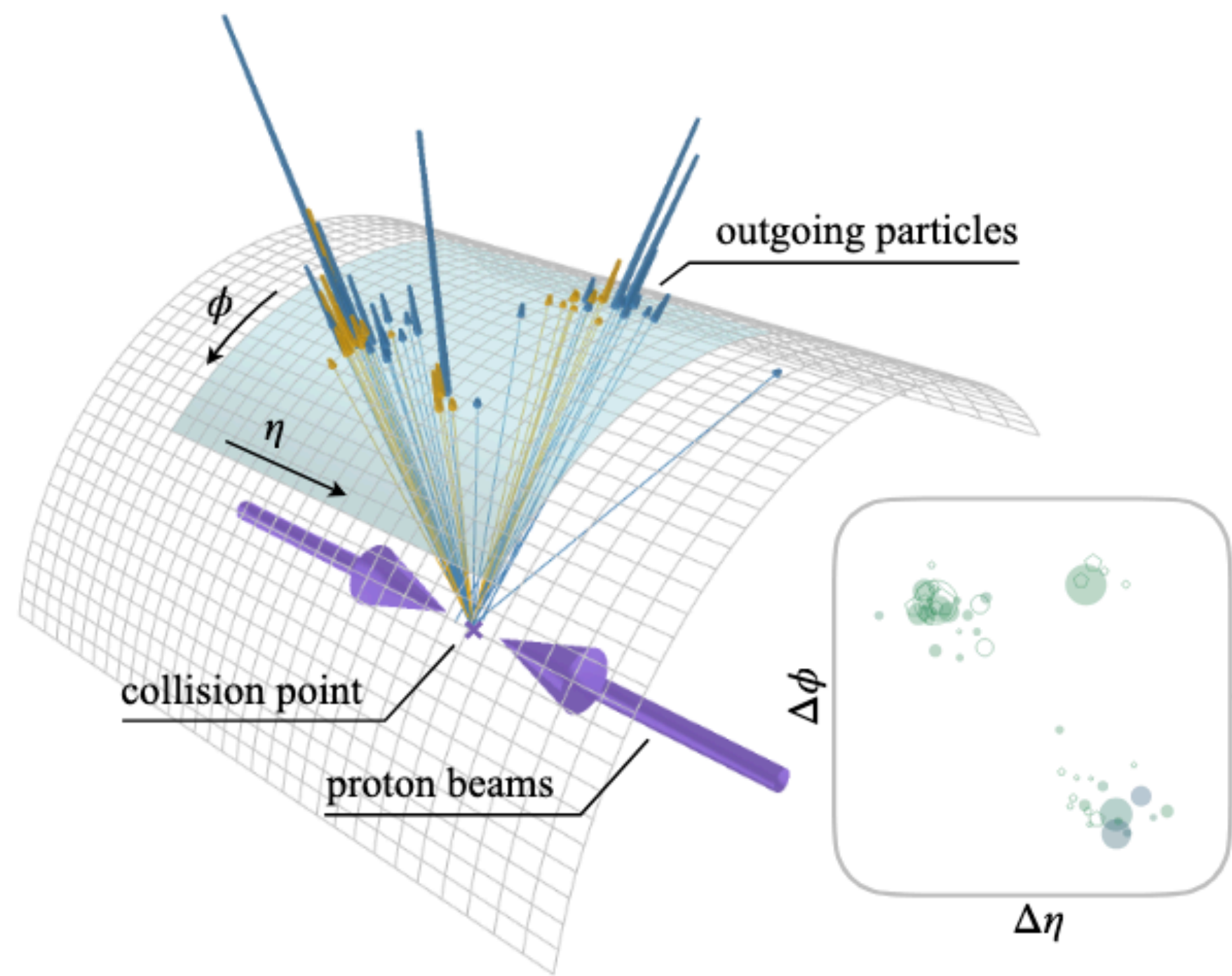
Jets as a point particle cloud

ParticleNet

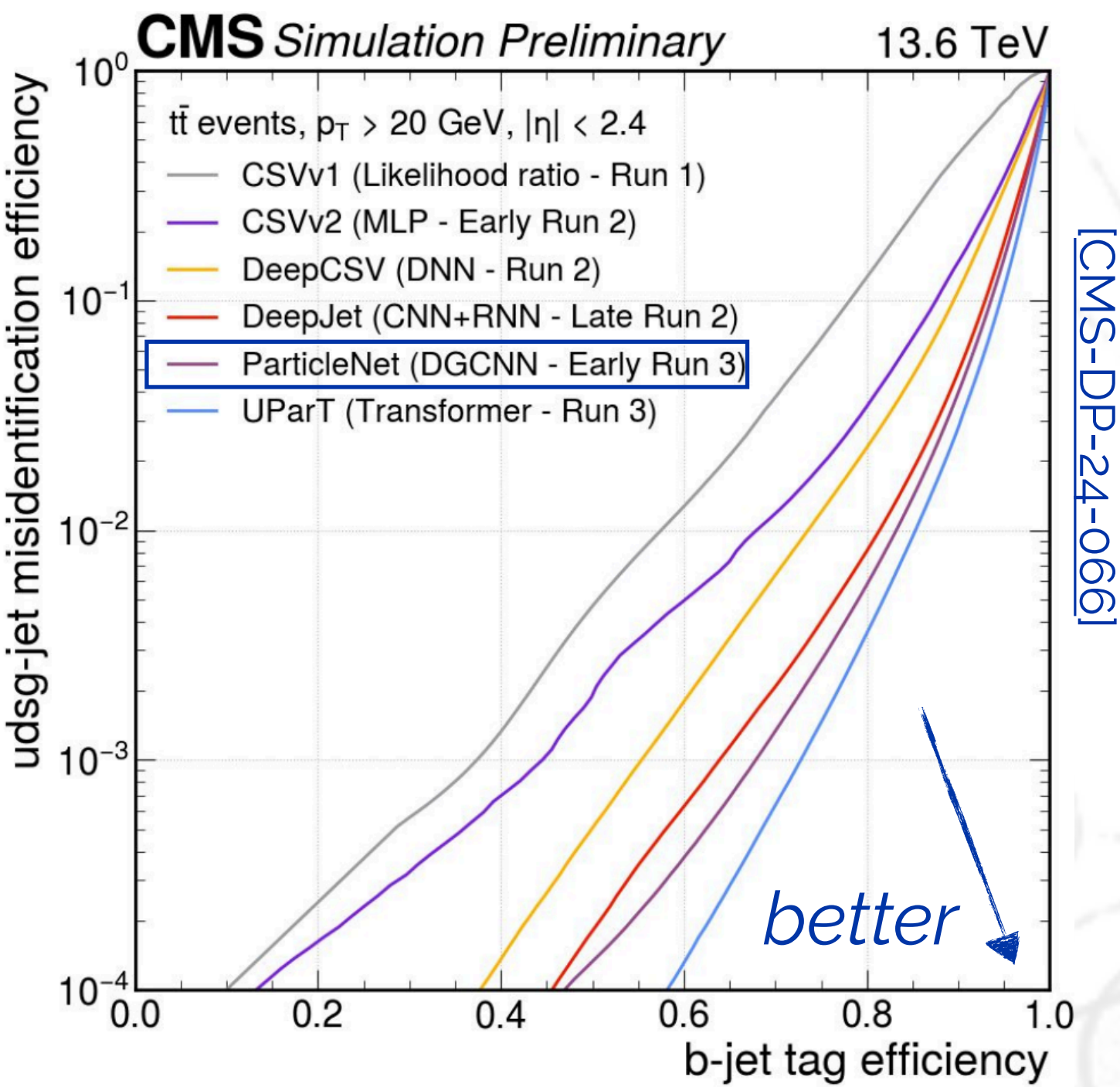
- The revolution: unordered set of particles in graph / **point cloud**
- For each point: (spatial) coordinates + properties (energy/momentum, detector response, ...)
- Key feature: **permutation symmetry** via e.g. locally connected graph

[Qu, H., Gouskos, L. (2019)]

**Main tagger for CMS Run 2
H(cc) results!**



From [PMLR162:18281-18292,2022]



[CMS-DP-24-066]

A bit of CMS history...

Run 1 (2011 - 2016)

Jet probabilities

Machine Learning

CSV/DeepCSV
Sub-jet tagging

Run 2 (2016 - 2018)

Jets as a sequence

DeepJet/
DeepAK8

Run 3 (2022 - 2026)

Jets as a point particle cloud

ParticleNet

ParticleTransformer

- Transformers are taking on every task!
- Core concept: Self-attention mechanism

Attention Is All You Need

[\[Vasvani et al. \(2017\)\]](#)

Ashish Vaswani*
Google Brain
avaswani@google.com

Llion Jones*
Google Research
llion@google.com

Noam Shazeer*
Google Brain
noam@google.com

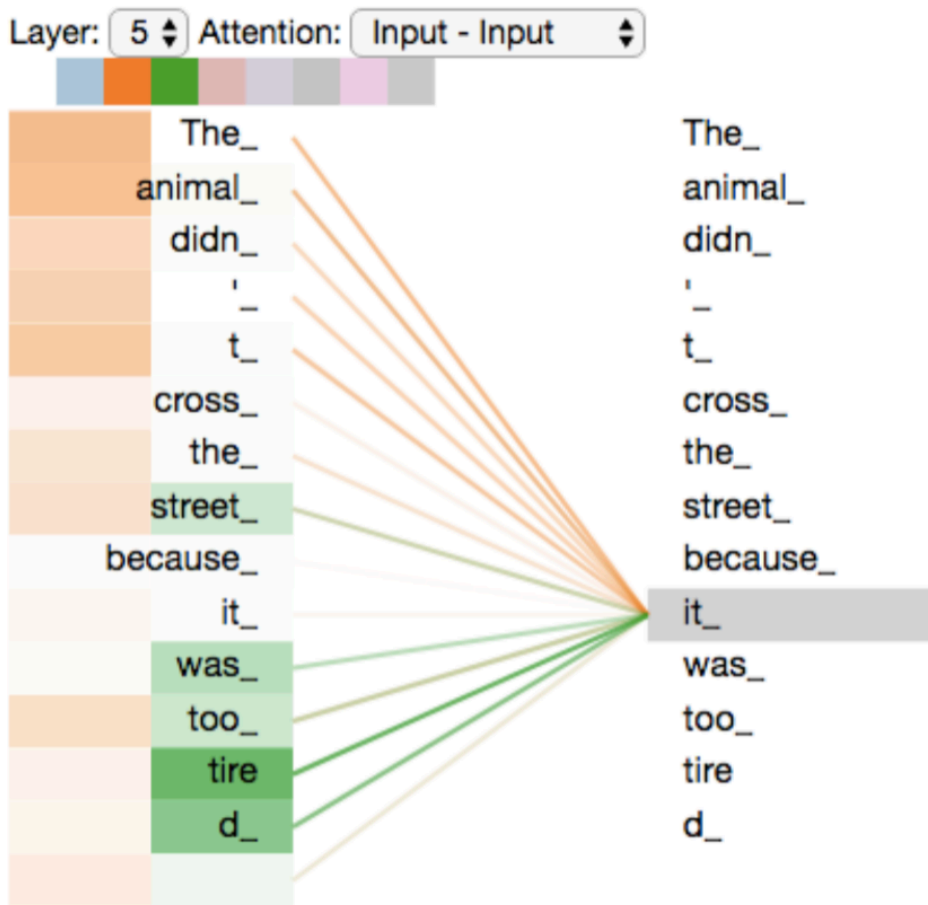
Aidan N. Gomez*[†]
University of Toronto
aidan@cs.toronto.edu

Niki Parmar*
Google Research
nikip@google.com

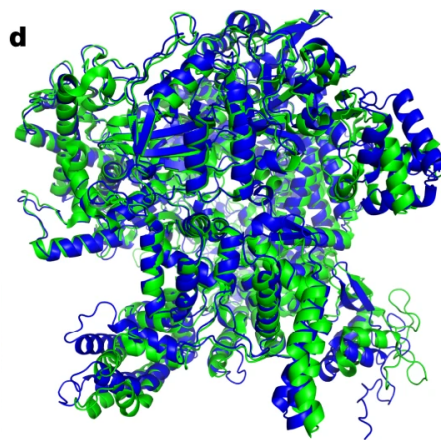
Lukasz Kaiser*
Google Brain
lukaszkaizer@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Illia Polosukhin*[‡]
illia.polosukhin@gmail.com

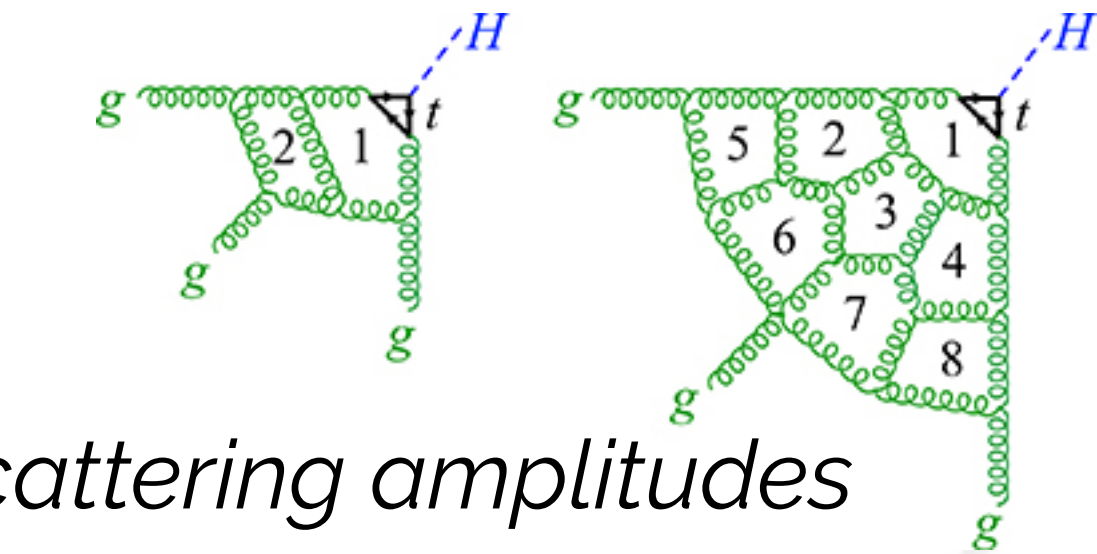


Protein structure

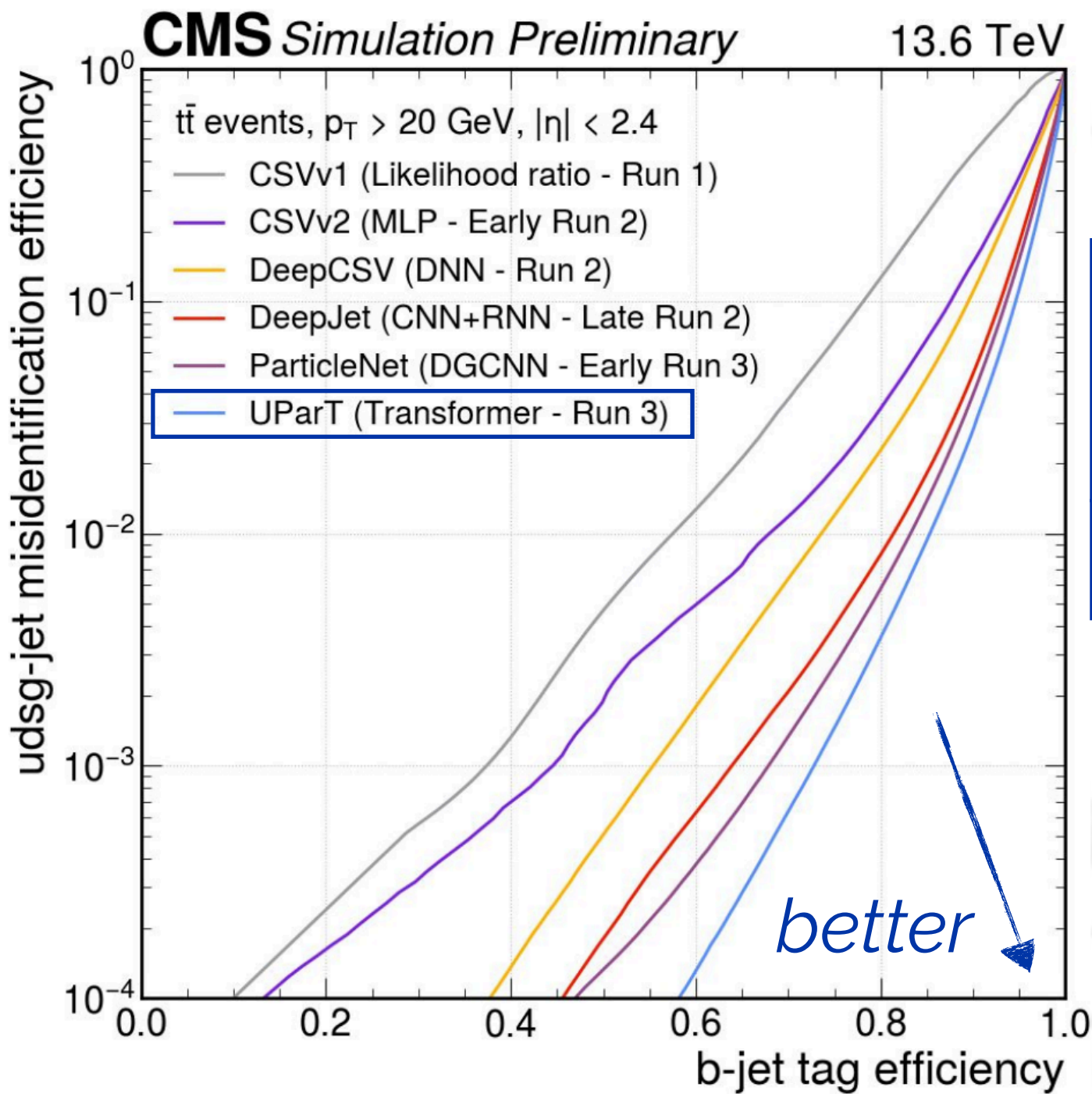


[\[Nature 596 \(2021\) 583\]](#)

Scattering amplitudes



[\[MLST 5 \(2024\) 035073\]](#)



[CMS-DP-24-066]

A bit of CMS history...

Run 1 (2011 - 2016)

Jet probabilities

Machine Learning

CSV/DeepCSV
Sub-jet tagging

Run 2 (2016 - 2018)

Jets as a sequence

DeepJet/
DeepAK8

Run 3 (2022 - 2026)

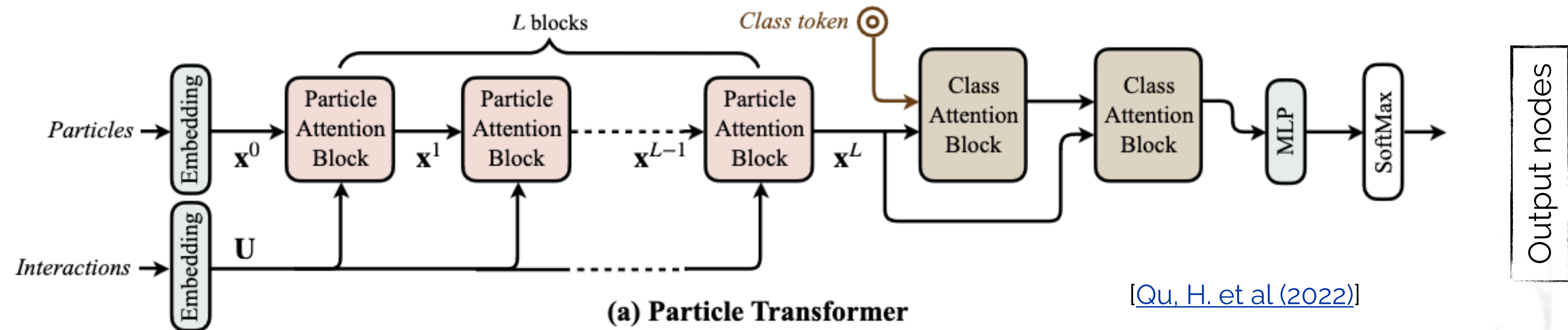
Jets as a point particle cloud

ParticleNet

ParticleTransformer

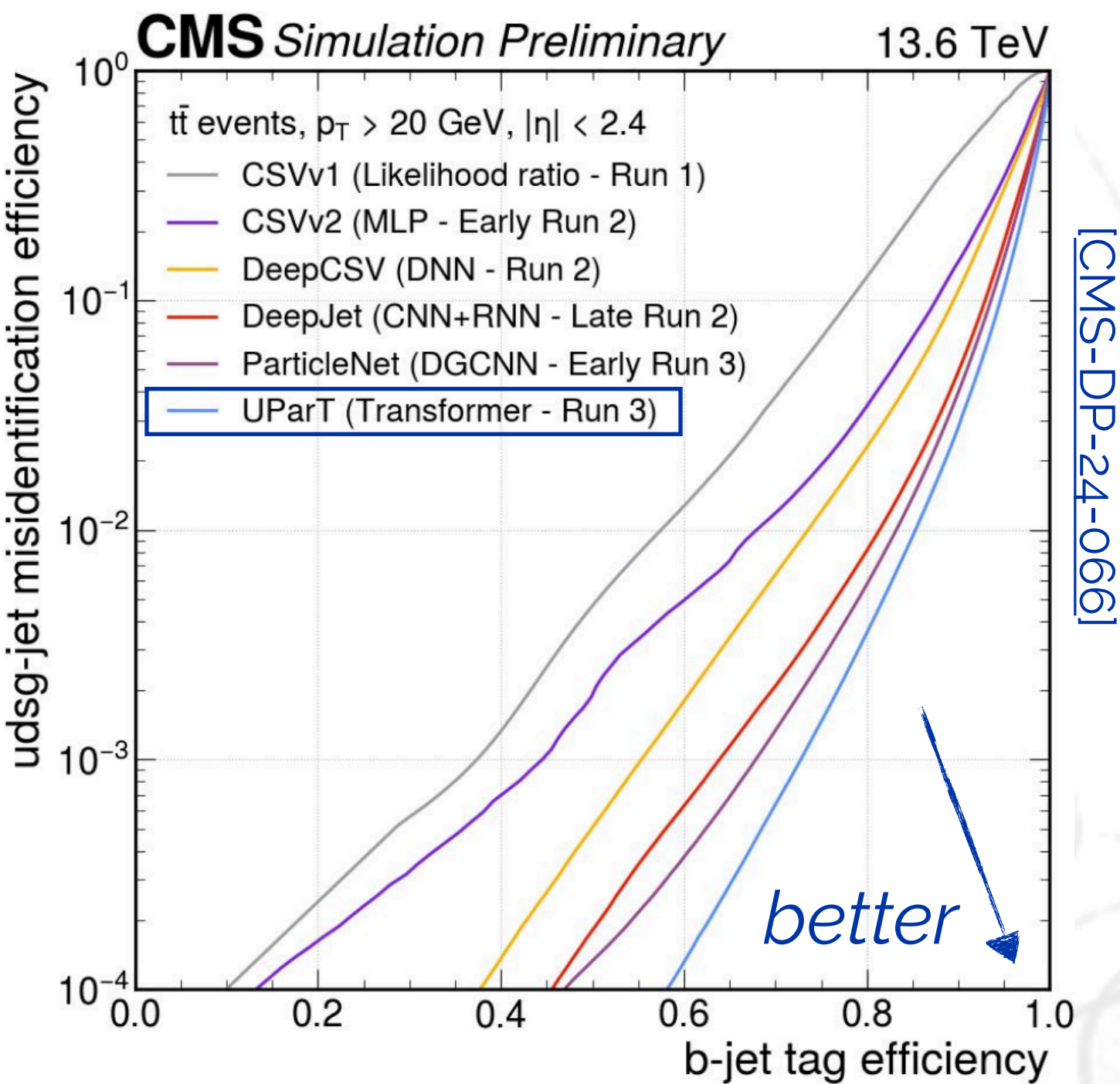
- Transformers are taking on every task!
- Core concept: Self-attention mechanism
- Physics-inspired pairwise features

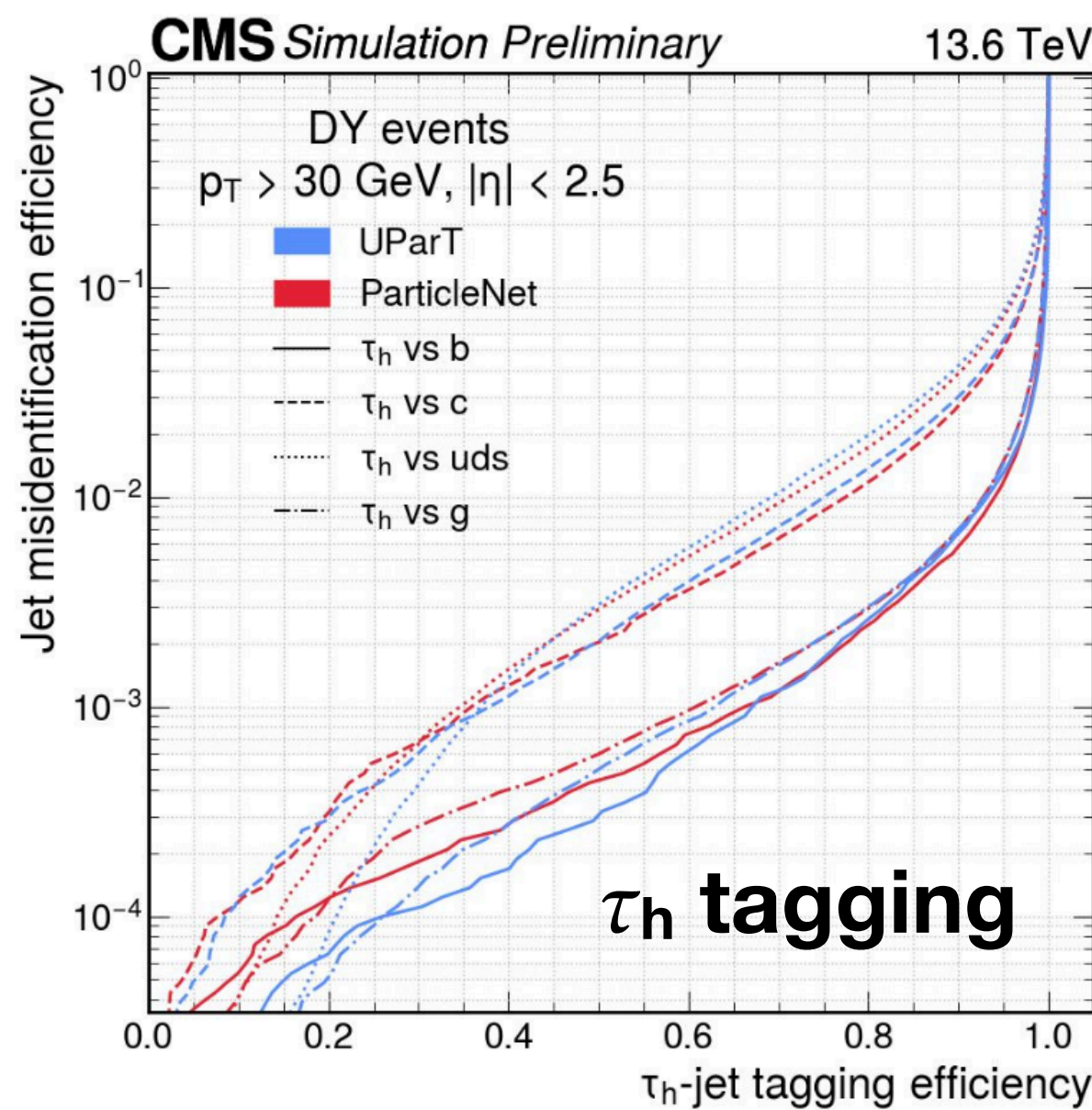
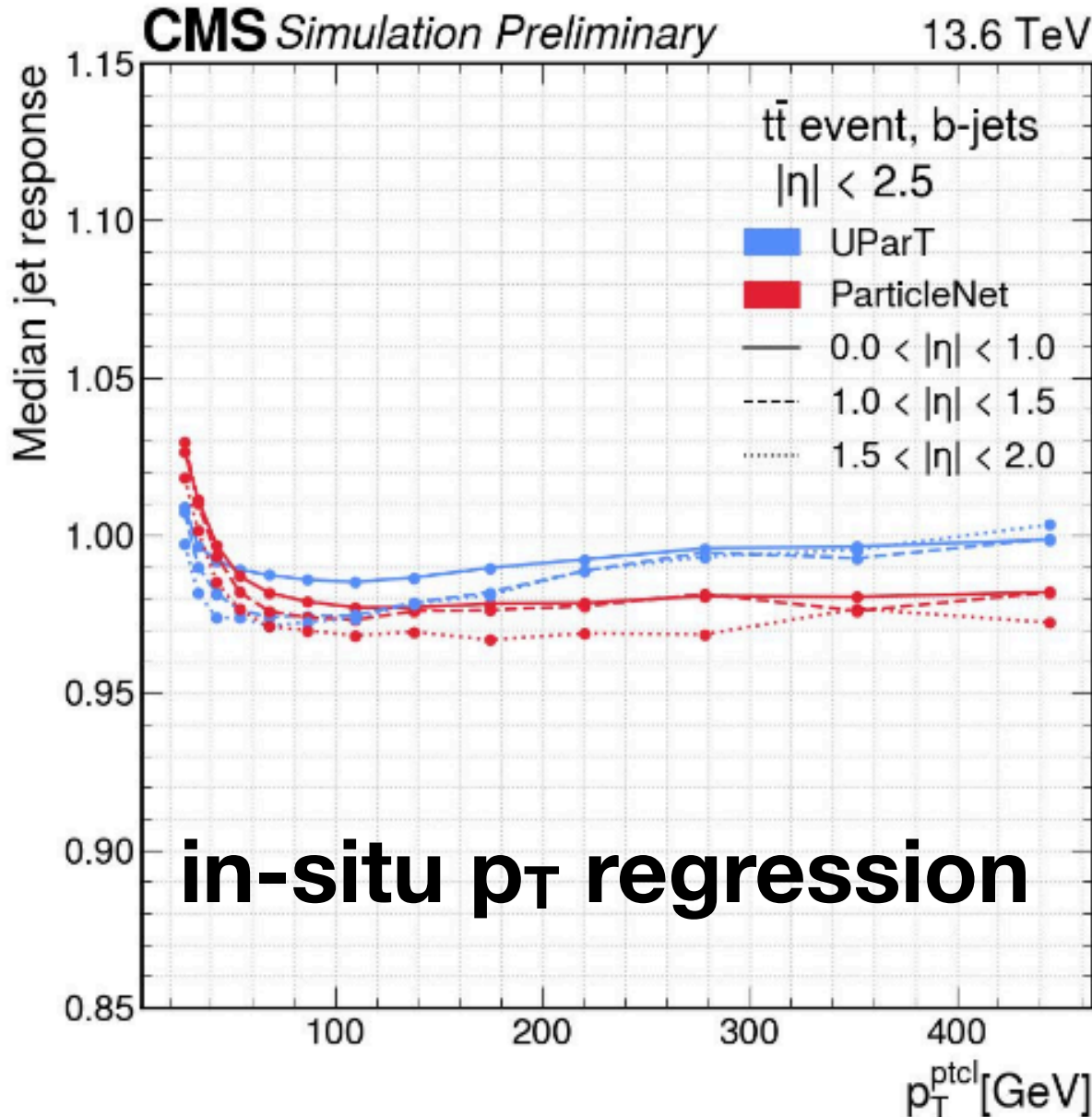
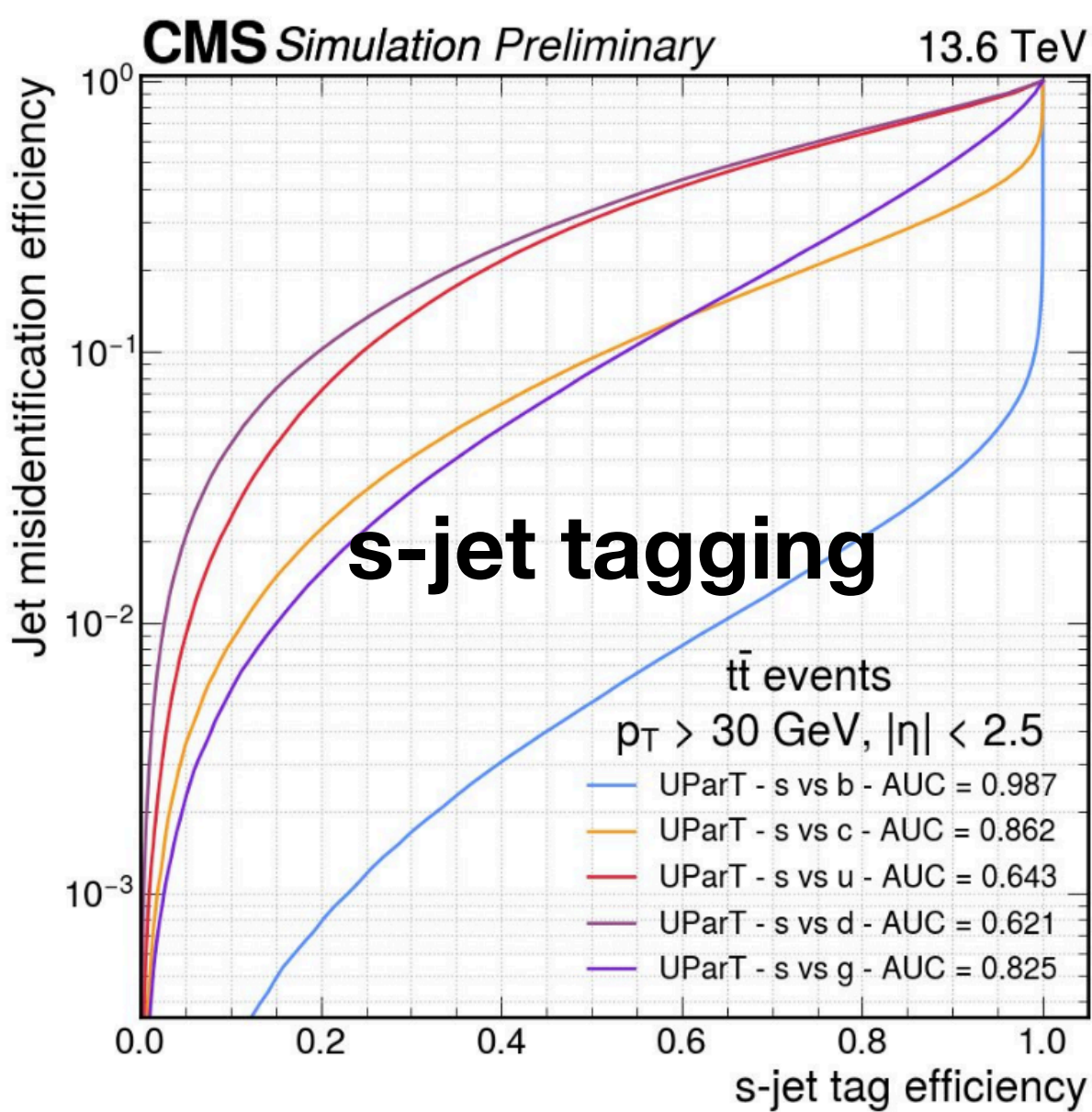
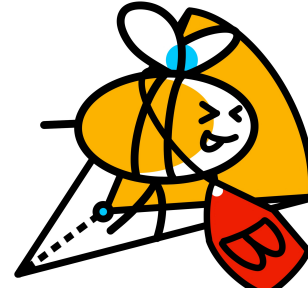
Our tagger for CMS Run 3!



$$\Delta = \sqrt{(y_a - y_b)^2 + (\phi_a - \phi_b)^2},$$
$$k_T = \min(p_{T,a}, p_{T,b})\Delta,$$
$$z = \min(p_{T,a}, p_{T,b})/(p_{T,a} + p_{T,b}),$$
$$m^2 = (E_a + E_b)^2 - \|\mathbf{p}_a + \mathbf{p}_b\|^2,$$

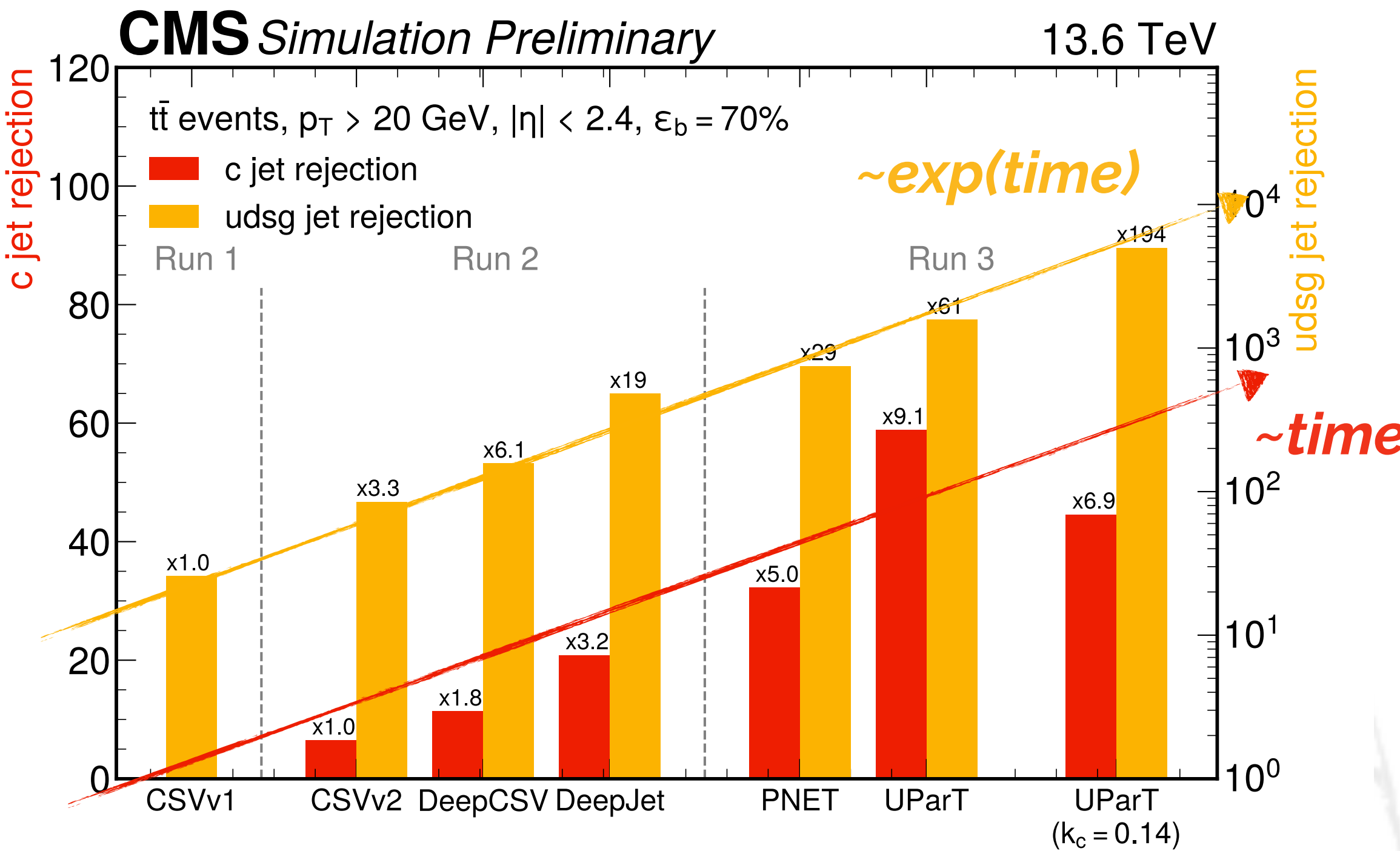
Inject physics to “steer” model





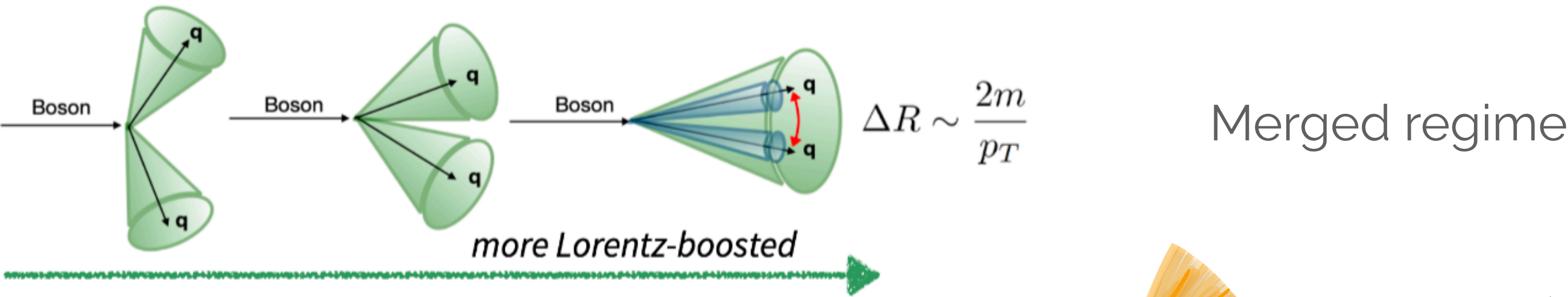
Towards a foundation model of jets?

N.B. ATLAS has similar results, although very different development over time..

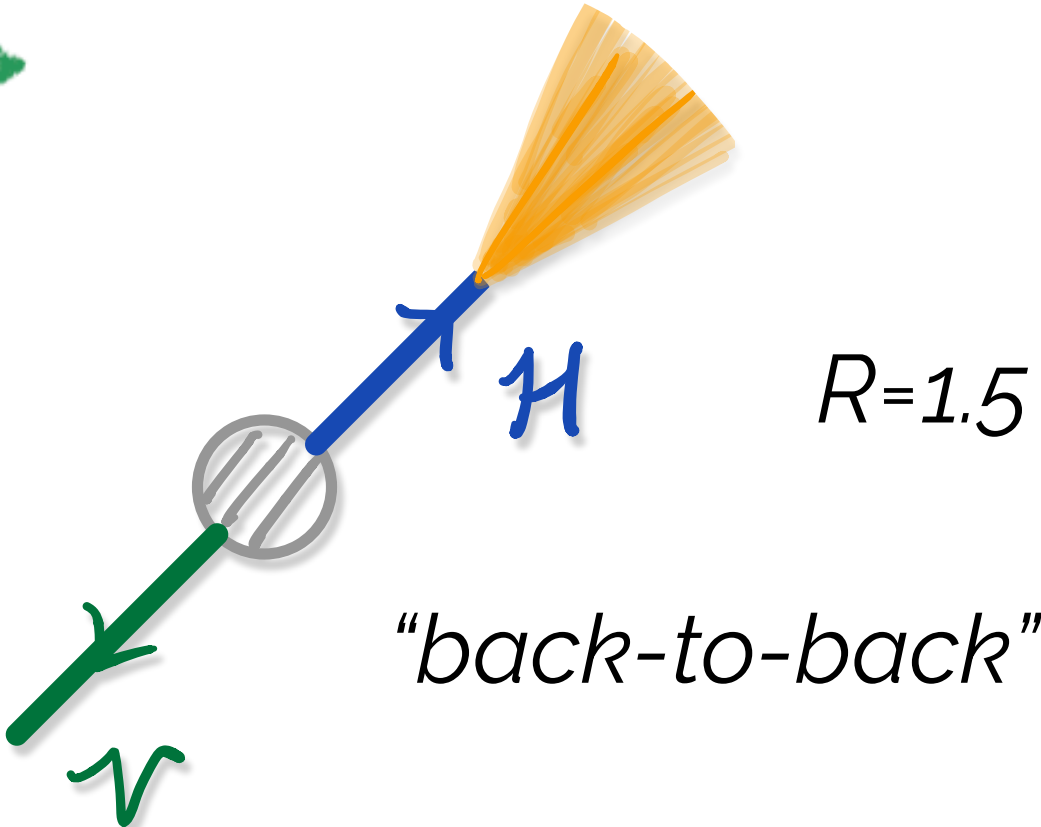
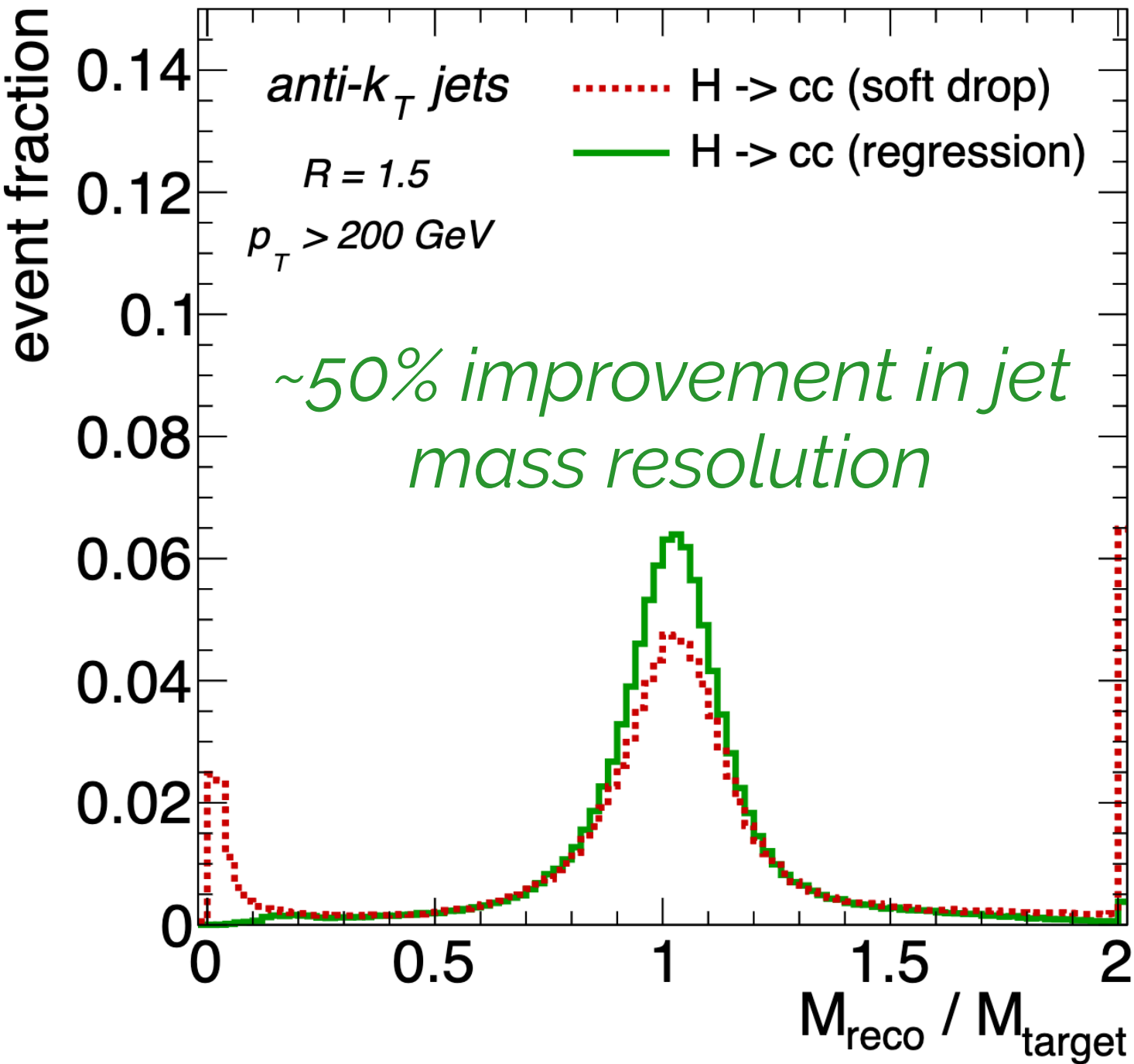


Impact of ParticleNet on the example of VH(cc)

- Similar improvement in boosted jet tagging!

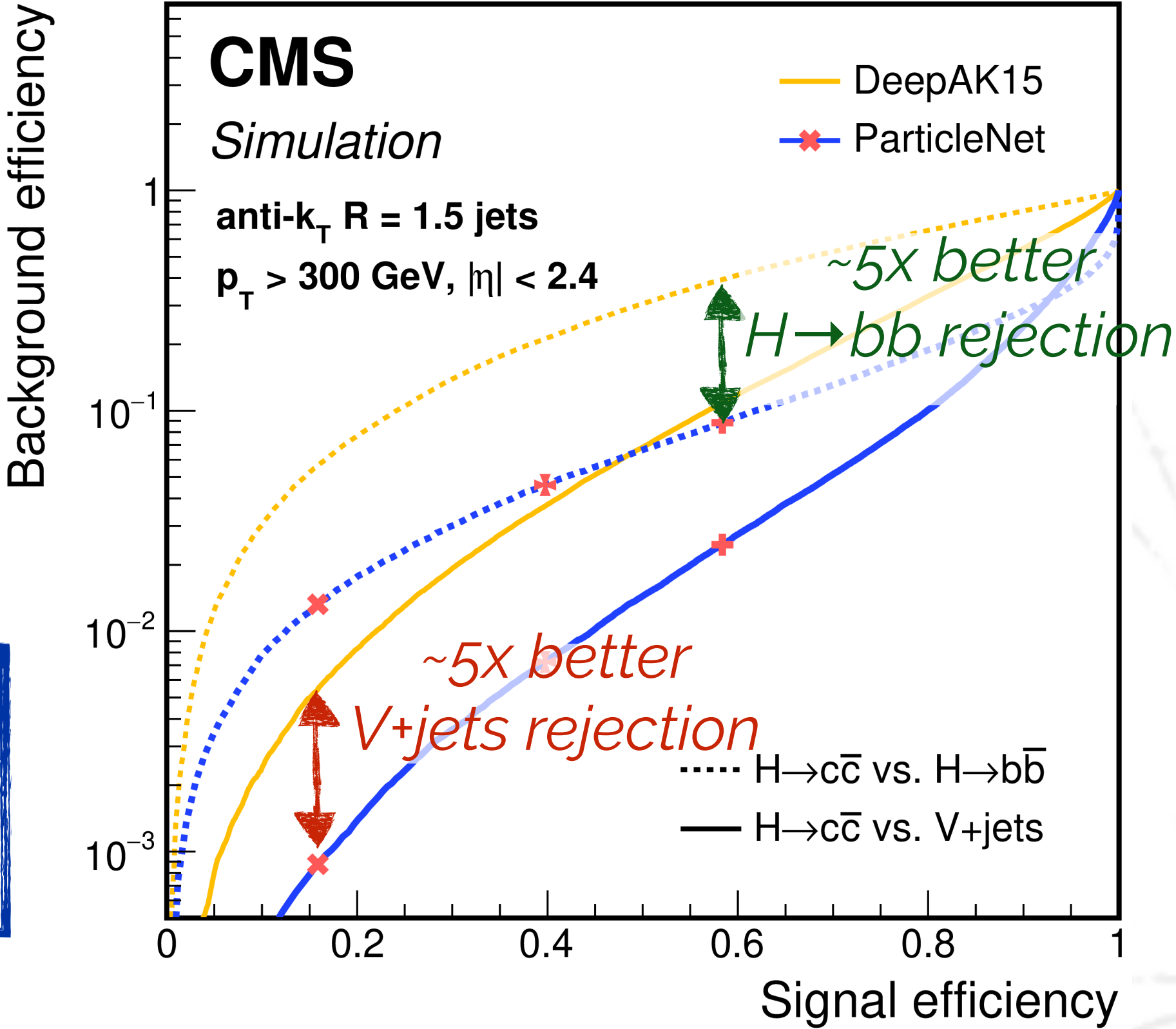


[CMS-DP-2021-017] CMS Simulation Preliminary



25% and 200% improvement in final sensitivity!

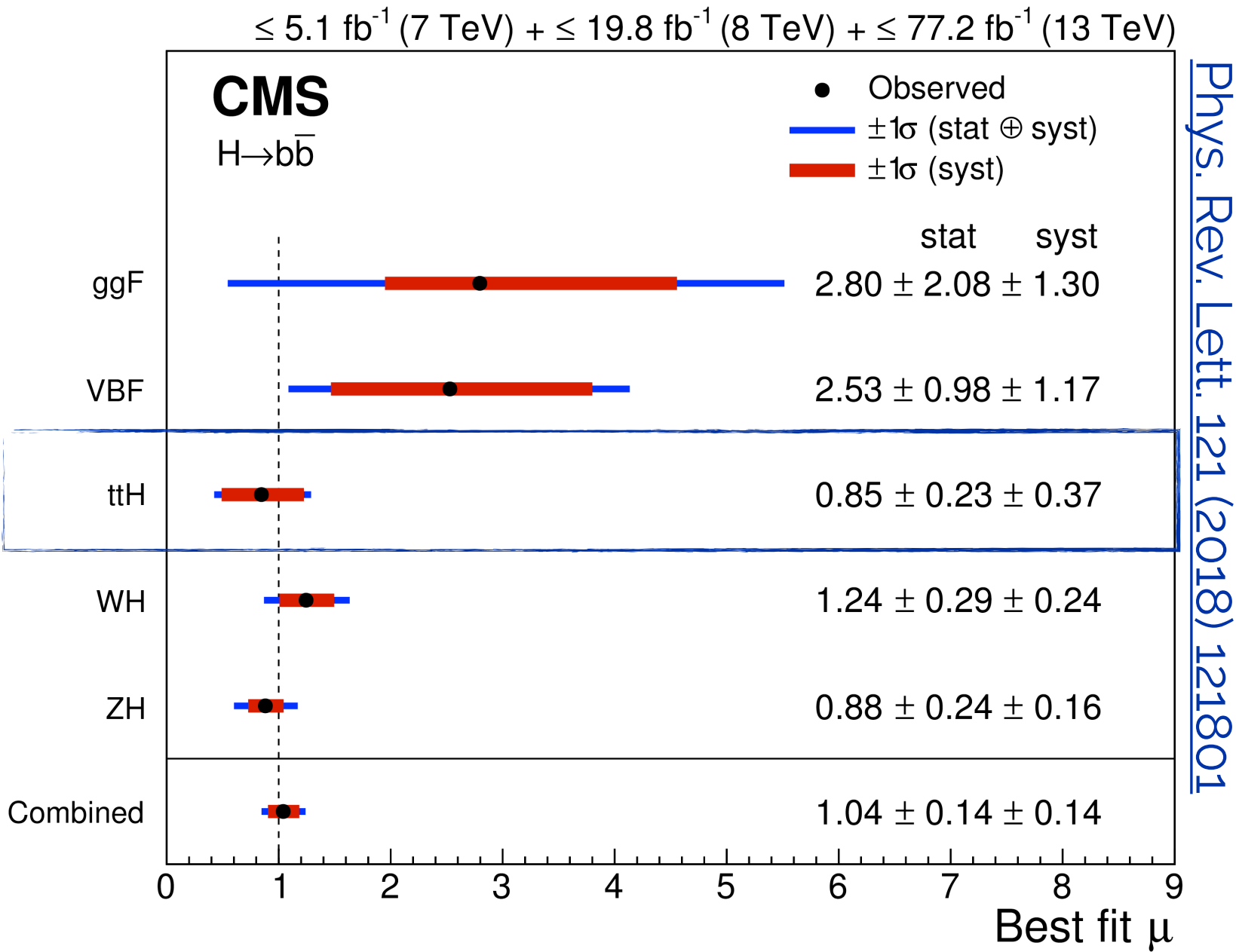
[PRL 131 (2023) 061801] (13 TeV)



What can we add to VH?



Let's check out the previous generation!



g g t $\bar{t}$ H c $\bar{c}$ κ_c

- **First search for $t\bar{t}H(H \rightarrow c\bar{c})$**
 - [PRL 36 \(2026\) 011801](#)
 - LHC Run 2 (138 fb^{-1})
 - Carried out as combination of 0L/1L/2L channels
 - Simultaneously measure $t\bar{t}H(H \rightarrow b\bar{b})$

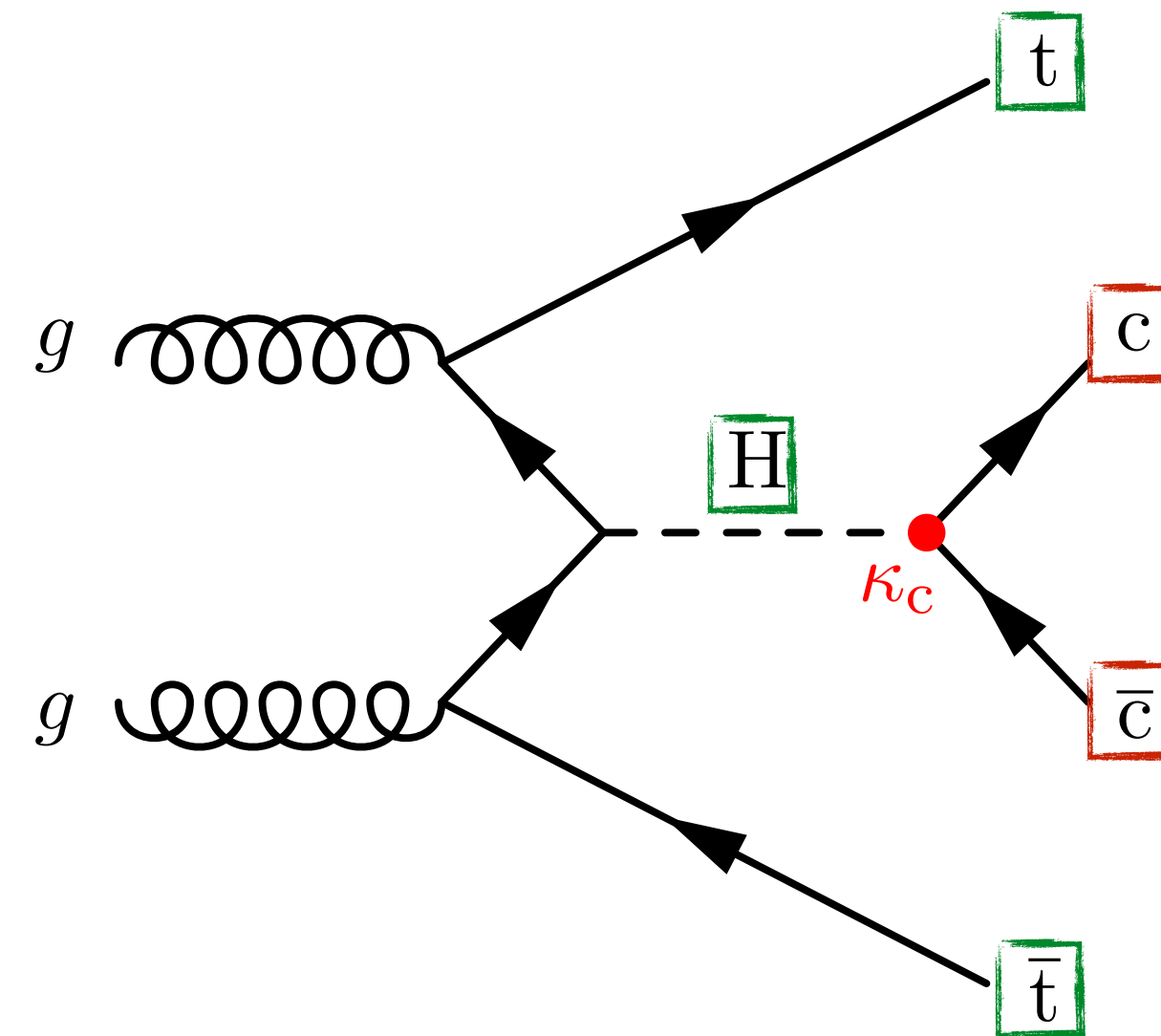
- $t\bar{t}H(H \rightarrow b\bar{b})$ important piece in the observation of $H \rightarrow b\bar{b}$
- Adding $t\bar{t}H(H \rightarrow c\bar{c})$ as a new puzzle piece
 - 0.5 pb vs. 2.5 pb of VH / large backgrounds
 - Challenging measurement, only **2000 $t\bar{t}H(H \rightarrow c\bar{c})$** events produced in CMS in Run 2



Why is $t\bar{t}H(H \rightarrow c\bar{c})$ challenging?

$t\bar{t}H(H \rightarrow c\bar{c})$ as an example of a modern LHC analysis

→ Smart usage of **physics-oriented machine learning**





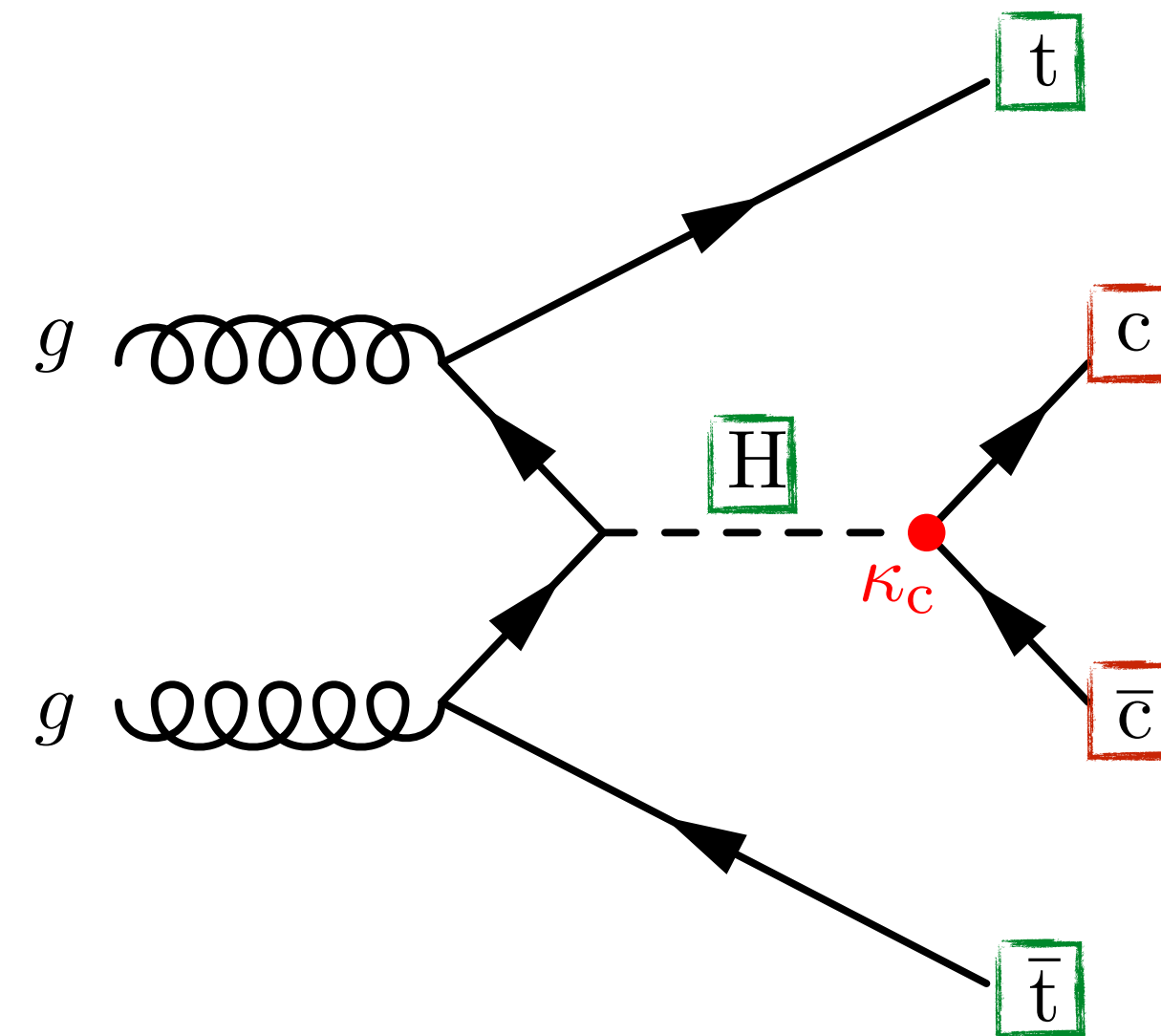
Why is $t\bar{t}H(H \rightarrow c\bar{c})$ challenging?

- **1) Jet flavour identification (tagging):**
 - Multiple **heavy-flavour (HF) jets** in the final state
 - Crucial to identify HF jets efficiently

Exploit modern graph-based tagger (ParticleNet)

$t\bar{t}H(H \rightarrow c\bar{c})$ as an example of a modern LHC analysis

→ Smart usage of **physics-oriented machine learning**





Why is $t\bar{t}H(H \rightarrow c\bar{c})$ challenging?

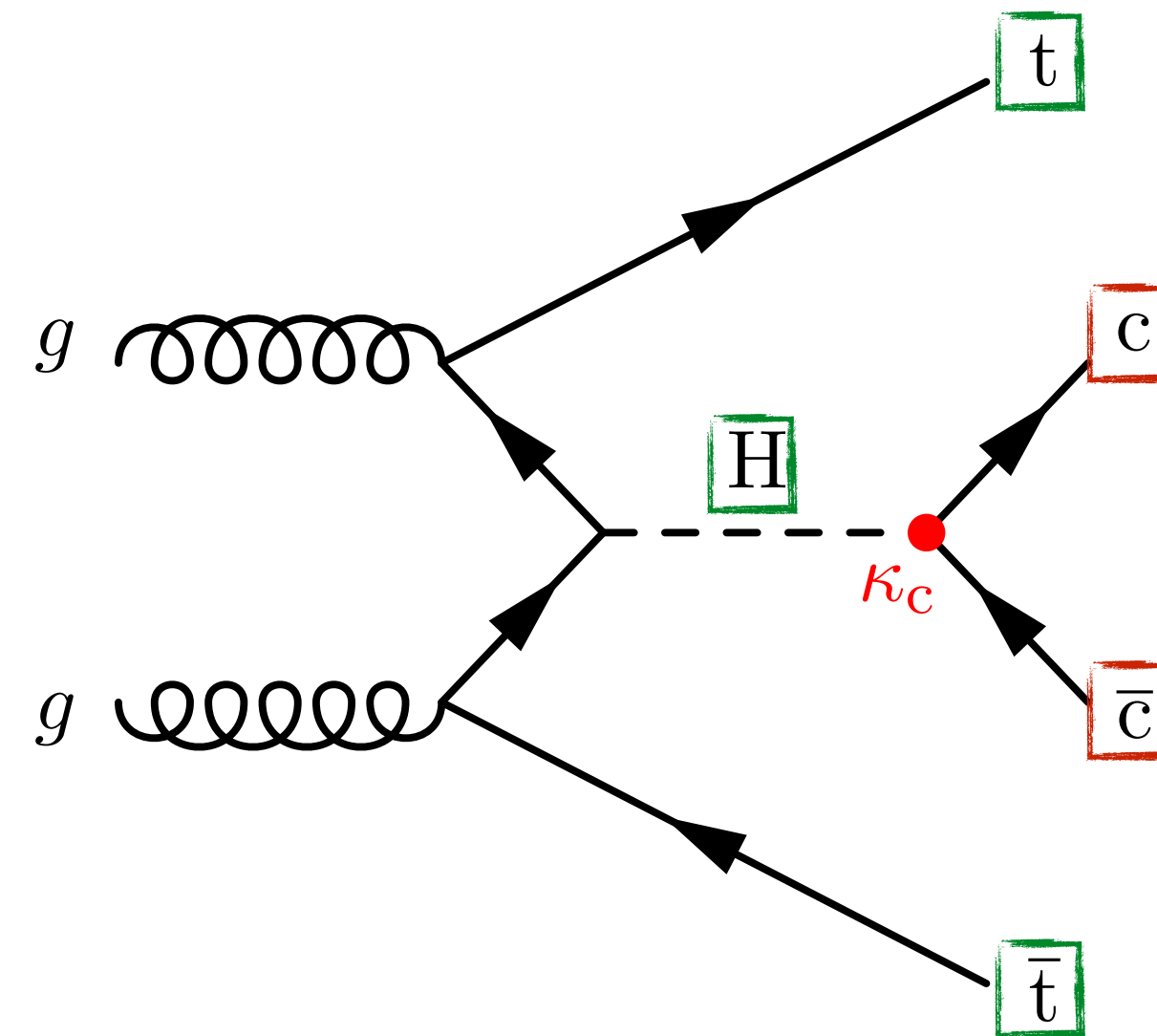
- **1) Jet flavour identification (tagging):**

- Multiple **heavy-flavour (HF) jets** in the final state
- Crucial to identify HF jets efficiently

Exploit modern graph-based tagger (ParticleNet)

$t\bar{t}H(H \rightarrow c\bar{c})$ as an example of a modern LHC analysis

→ Smart usage of **physics-oriented machine learning**



- **2) Event reconstruction:**

- Complicated topology: High jet multiplicity & leptons (only $e\backslash\mu$)
 - e.g. 8 jets in 0L channel
- Ambiguous combinatorics

→ **Higgs reconstruction challenge**

Neural network classifier learning correlations (ParticleTransformer)



Why is $t\bar{t}H(H \rightarrow c\bar{c})$ challenging?

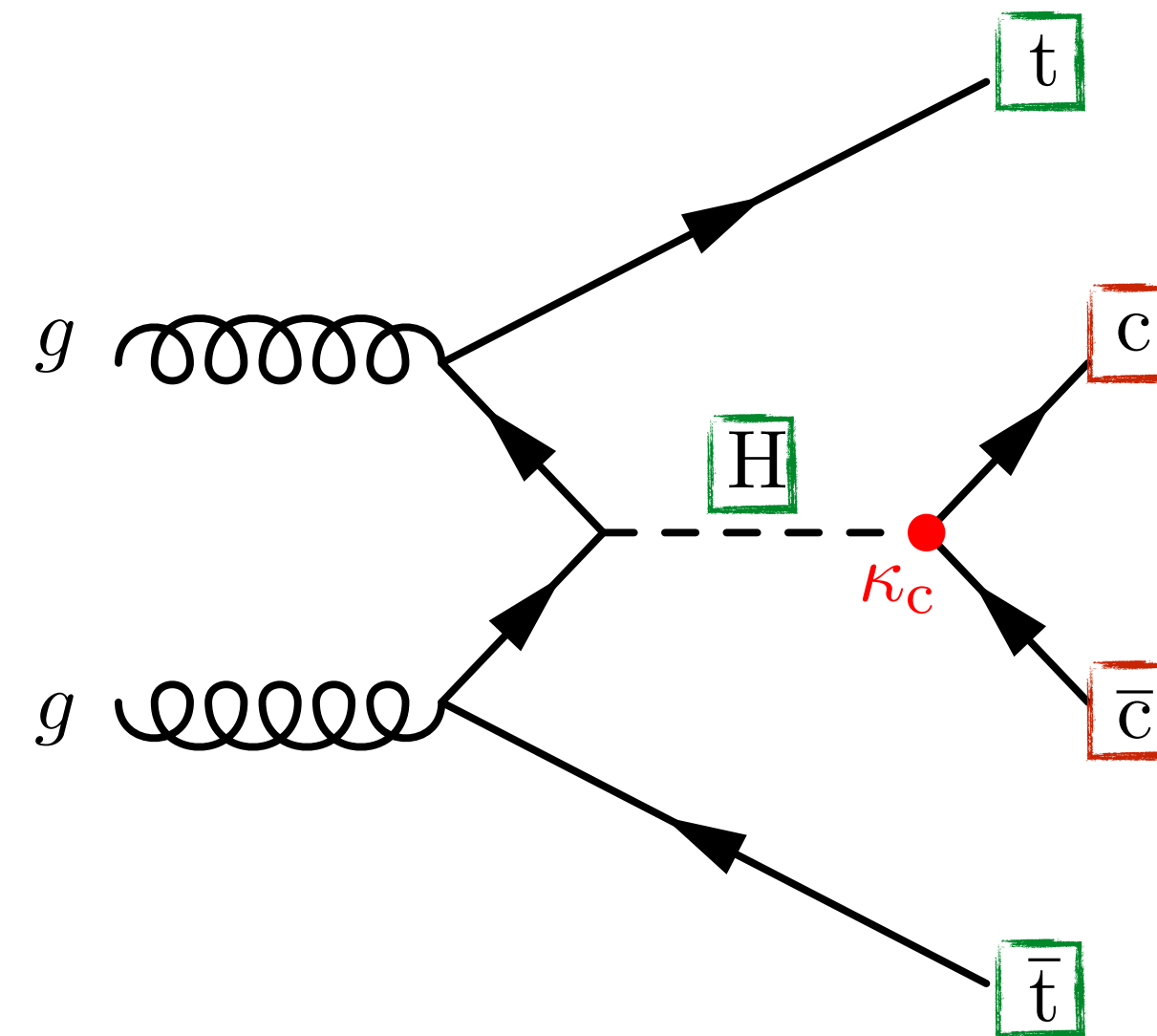
- **1) Jet flavour identification (tagging):**

- Multiple **heavy-flavour (HF) jets** in the final state
- Crucial to identify HF jets efficiently

Exploit modern graph-based tagger (ParticleNet)

$t\bar{t}H(H \rightarrow c\bar{c})$ as an example of a modern LHC analysis

→ Smart usage of **physics-oriented machine learning**



- **2) Event reconstruction:**

- Complicated topology: High jet multiplicity & leptons (only $e\backslash\mu$)
 - e.g. 8 jets in 0L channel
- Ambiguous combinatorics

→ **Higgs reconstruction challenge**

Neural network classifier learning correlations (ParticleTransformer)

- **3) Background estimation:**

- Dominant $t\bar{t}$ +heavy-flavor (HF) jets (0L/1L/2L) and QCD multijet (0L)
- Difficult to model in simulation

Data-driven background estimation

Heavy flavour jet identification for $t\bar{t}H(H \rightarrow c\bar{c})$



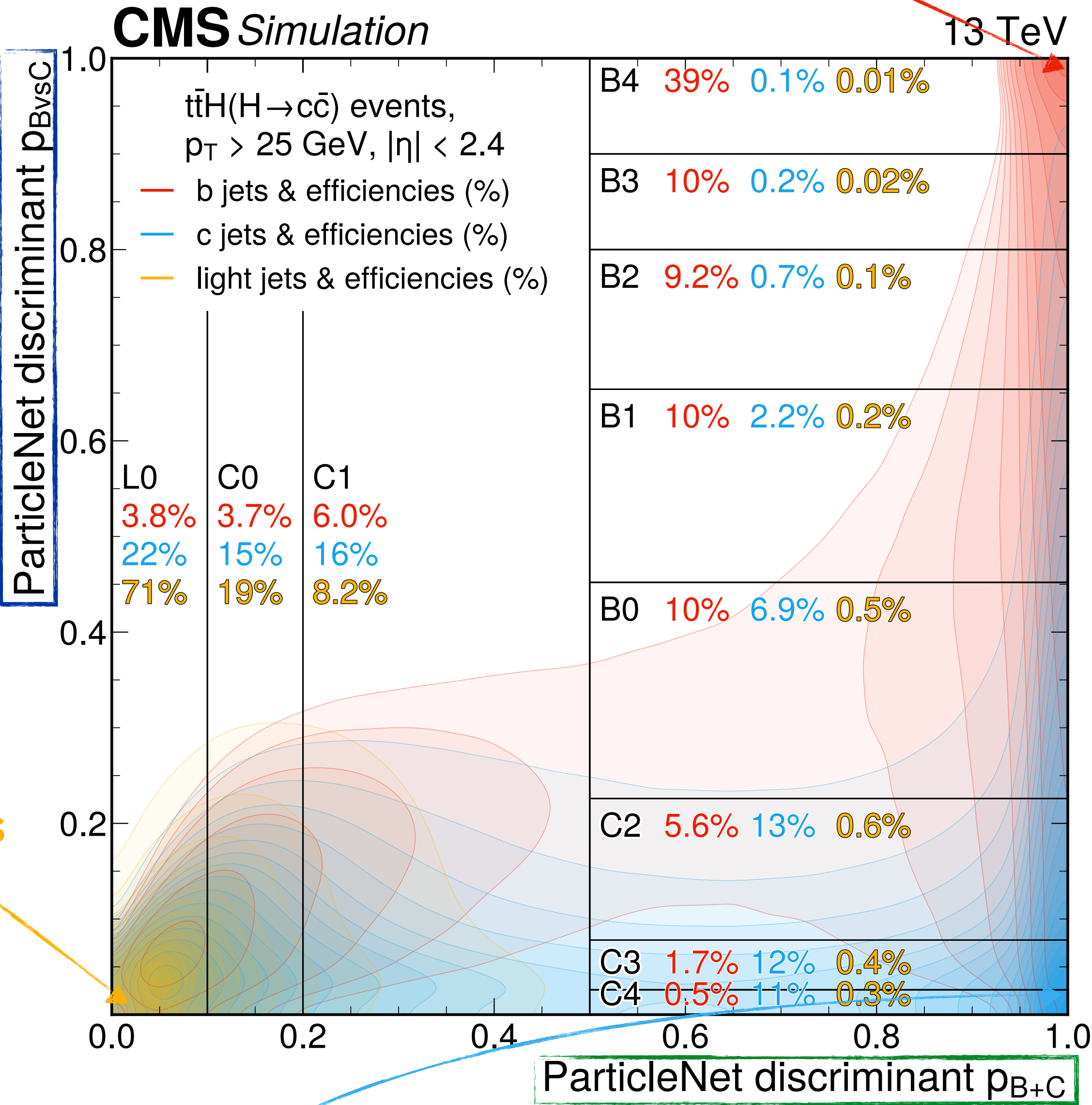
- This analysis exploits for the first time ParticleNet for small-radius jets in a Run 2 analysis in CMS
- For simultaneous b- and c-jet identification:

$$p_{B+C} = \frac{p_b + p_{bb} + p_c + p_{cc}}{p_b + p_{bb} + p_c + p_{cc} + p_{uds} + p_g}$$
$$p_{BvsC} = \frac{p_b + p_{bb}}{p_b + p_{bb} + p_c + p_{cc}}$$

- Custom tagging categories:
 - Granular score information “pseudo continuous”
 - More robust calibration than full score
 - 5 b-jet enriched bins (B0 - B4)
 - 5 c-jet enriched bins (C0 - C4)
 - 1 light-jet (udsg) enriched bin (L0)

light jets

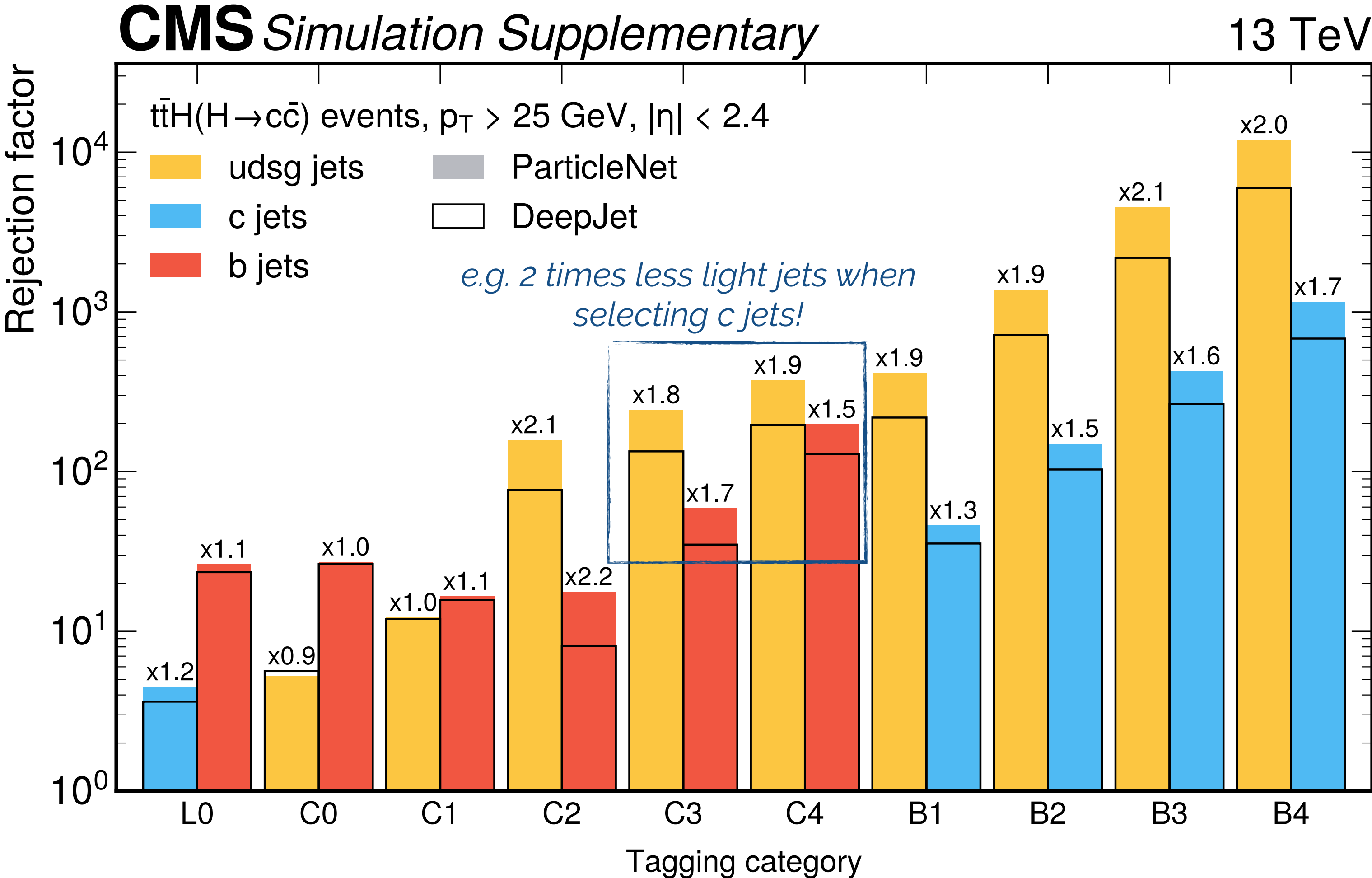
c jets



Heavy flavour jet identification for $t\bar{t}H(H \rightarrow c\bar{c})$

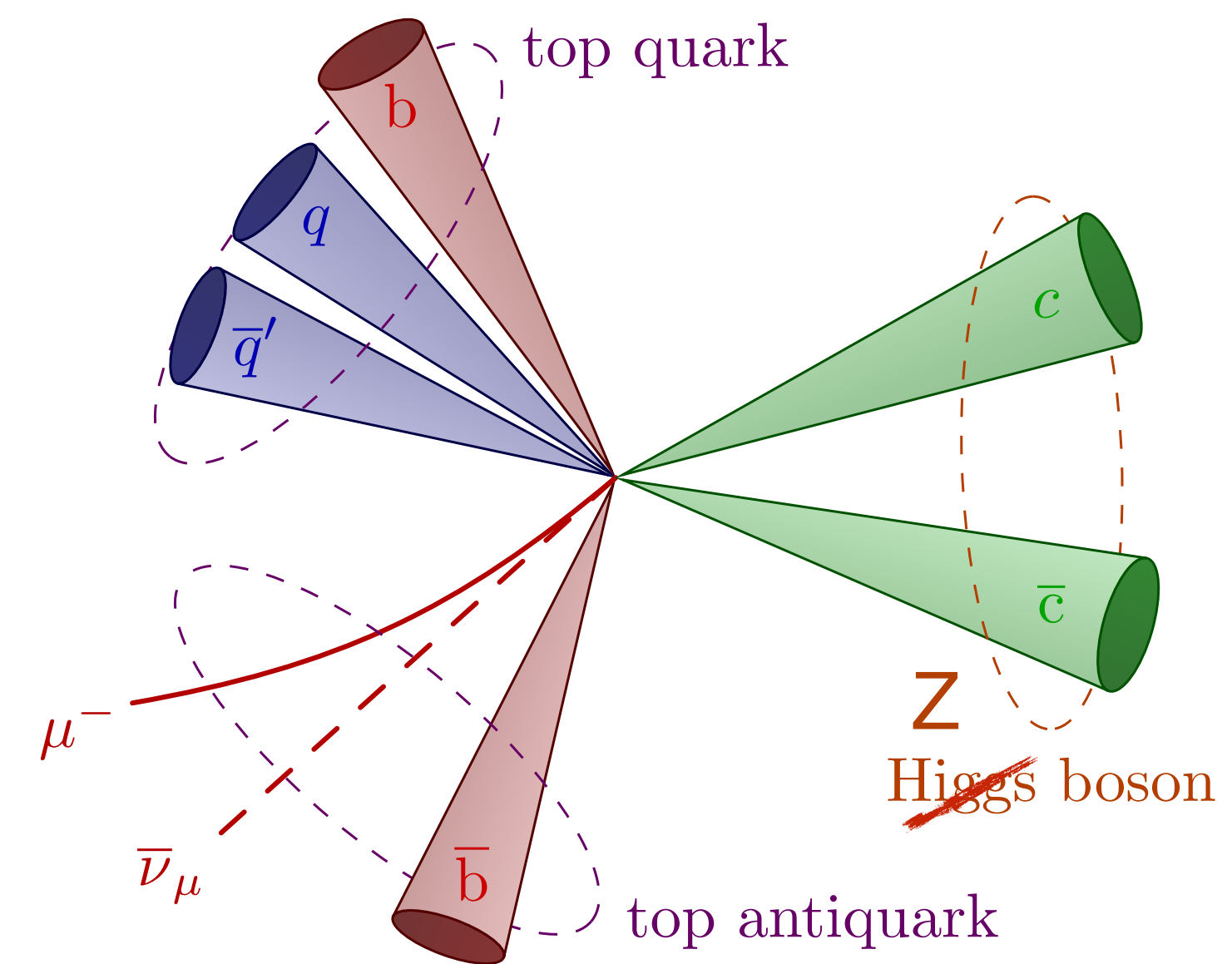
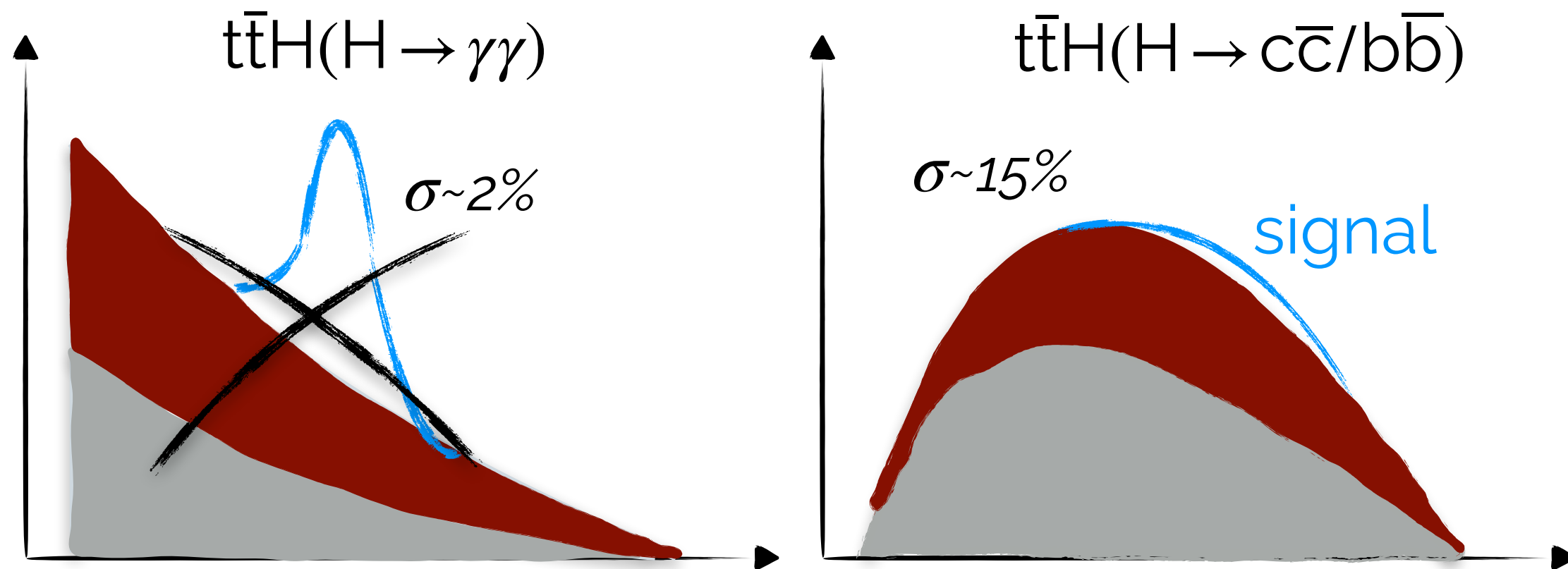


- How much do we gain by using ParticleNet in the $t\bar{t}H(H \rightarrow c\bar{c})$ analysis?
- Study **per-jet rejection for fixed signal-jet efficiencies** comparing DeepJet (CNN+RNN) vs. ParticleNet



Factor ~2 better than DeepJet (default for Run 2)!

Let's define the analysis strategy



- Combinatorics + jet resolution + complicated backgrounds
→ Mass fits not optimal
- Fitting neural network scores D_X
 - Need to validate background model, in e.g., CRs
 - But how well does it work for our signal?
- Replace the Higgs boson with a **Z boson**
 - Very similar kinematic properties (except lower mass)
 - Known well enough from dedicated measurements
 - Large enough cross section (0.86 pb for $t\bar{t}Z$ vs. 0.51 pb for $t\bar{t}H$)

- Find a good **proxy for validation!**

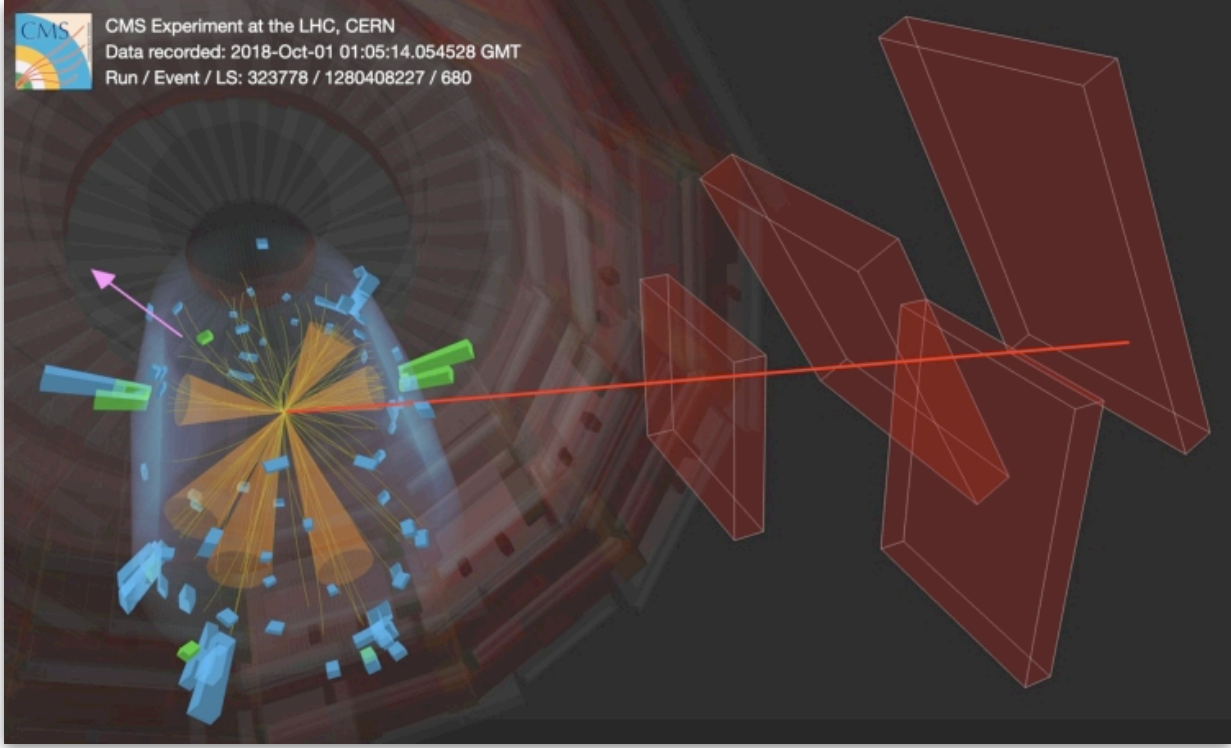
Analysis symmetric in $t\bar{t}Z$ and $t\bar{t}H$

$b\bar{b}$: 12 % vs. 58 % BR
 $\rightarrow t\bar{t}Z(Z \rightarrow b\bar{b}) \sim 1/3$ of $t\bar{t}H(H \rightarrow b\bar{b})$
 $c\bar{c}$: 15 % vs. 3 % BR
 $\rightarrow t\bar{t}Z(Z \rightarrow c\bar{c}) \sim \mathbf{13 \text{ times}}$ $t\bar{t}H(H \rightarrow c\bar{c})$

Multiclass event classification



- Complicated hadronic event topology can be a benefit:
 - 8 jets (0L), 1 lepton + 6 jets (1L), 2 leptons + 4 jets (2L)
 - 2 b jets, 2 c jets
 - Rich correlations and resonance structures (1 Higgs, 2 W's, 2 top quarks)
- Event classification using state-of-the-art transformer network
 - Let's bring the tagging improvements to the analysis level!



Vasvani et al. 2017

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

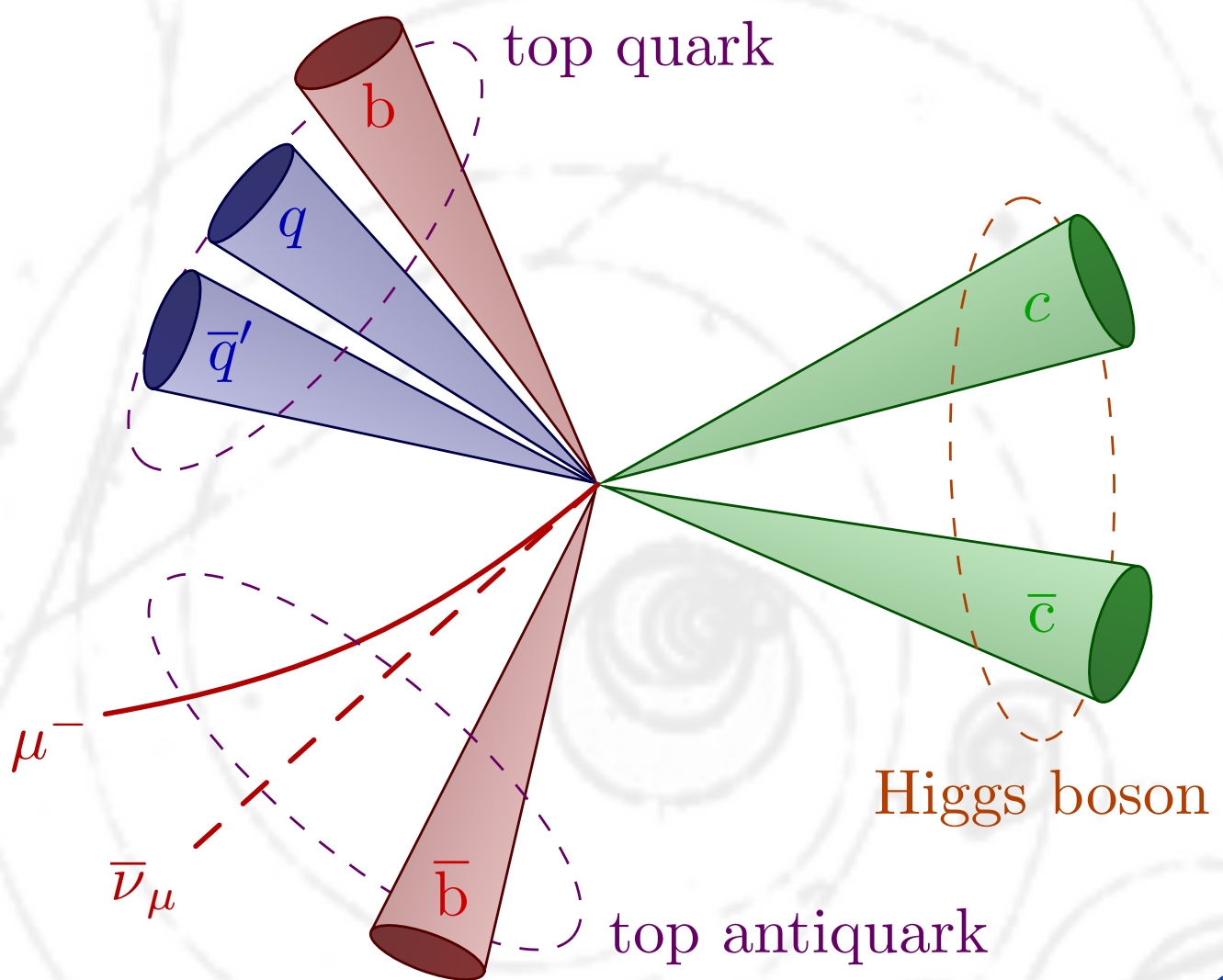
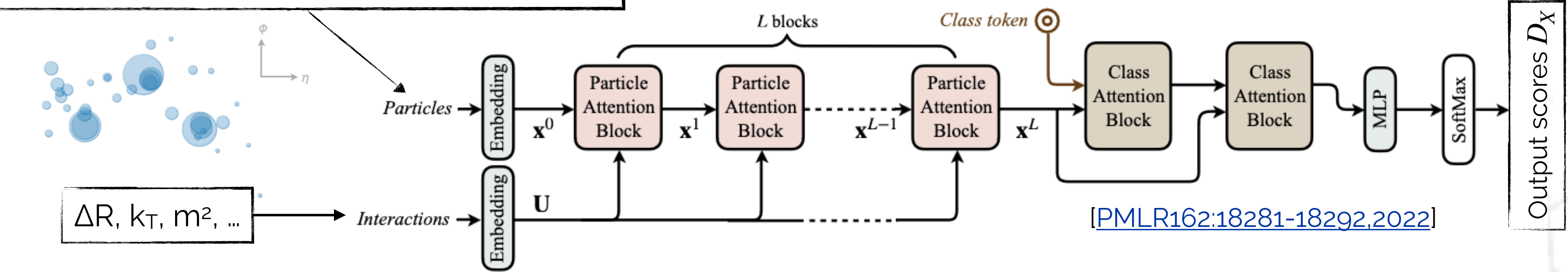
Aidan N. Gomez*[†]
University of Toronto
aidan@cs.toronto.edu

Lukasz Kaiser*
Google Brain
lukasz.kaiser@google.com

Illia Polosukhin*[‡]
illia.polosukhin@gmail.com

ML

Jets (up to 10/8/6 in FH/SL/DL), leptons, MET:
four-momenta, PID, tagging information



Background modeling & estimation

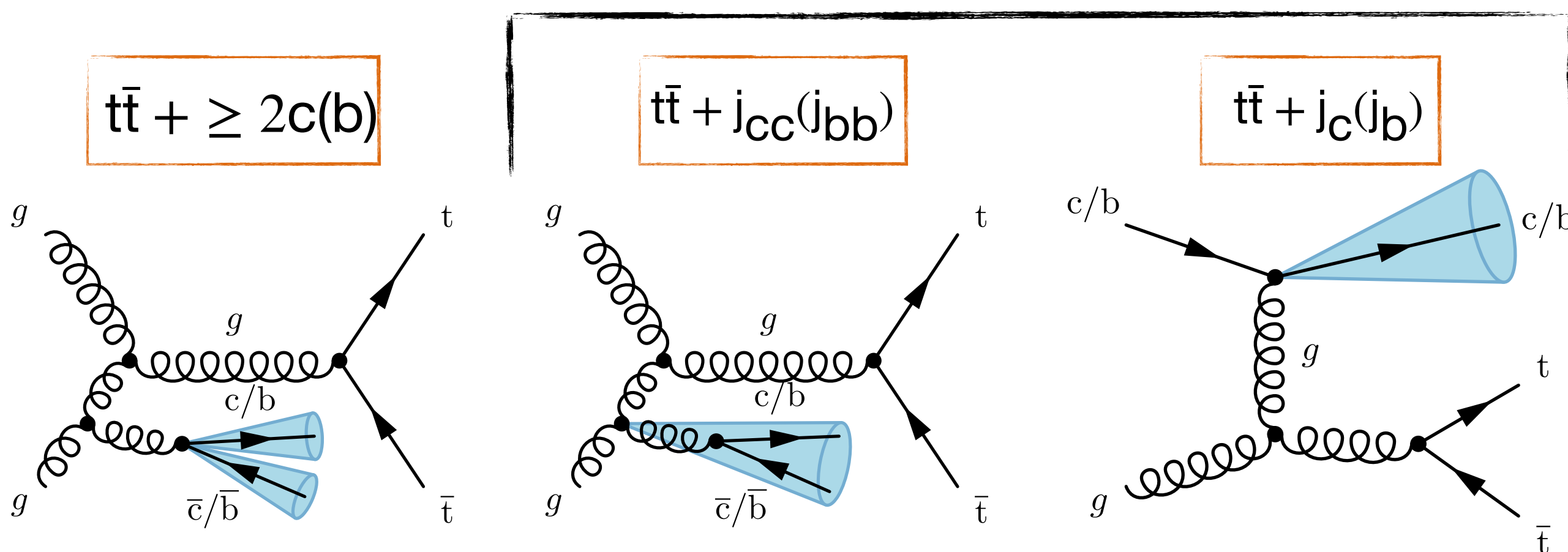


$t\bar{t}$ +jets background model:

- **State-of-the-art merging of 5FS and 4FS predictions**

- $t\bar{t} + \geq 2b, t\bar{t} + j_b, t\bar{t} + j_{bb}$ from 4FS $t\bar{t}b\bar{b}$
- $t\bar{t} + \geq 2c, t\bar{t} + j_c, t\bar{t} + j_{cc}, t\bar{t} + \text{light}$ from 5FS $t\bar{t}$ +jets

$t\bar{t} + c(b)$



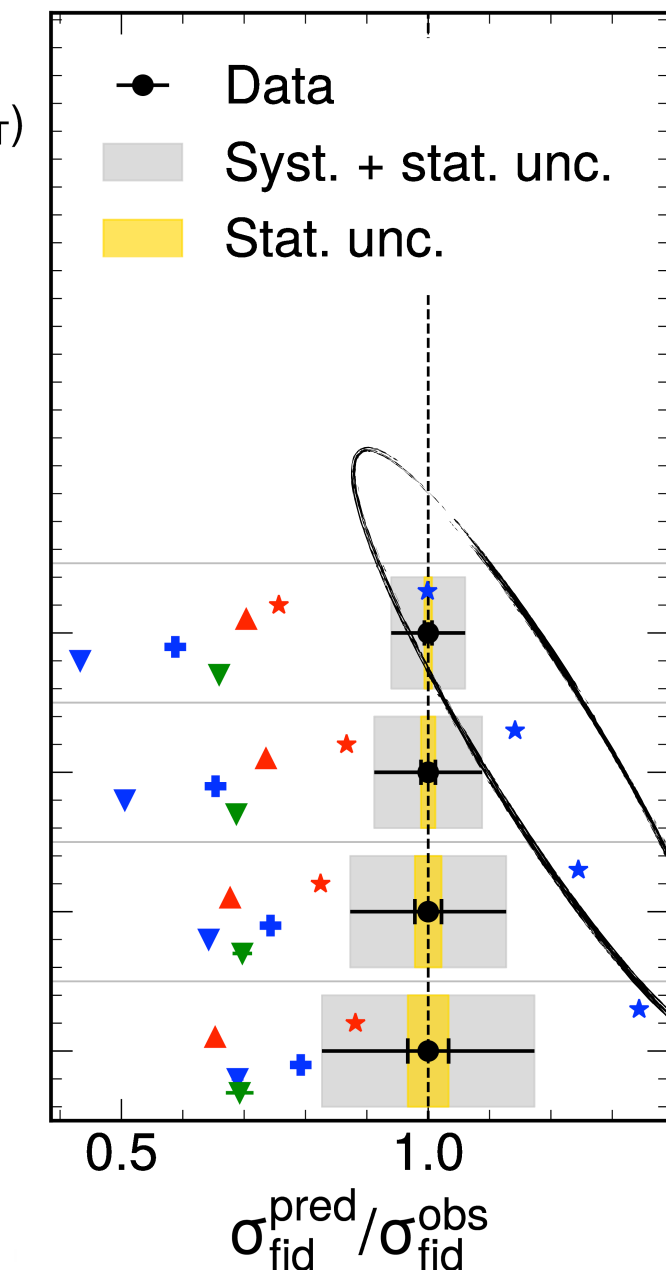
Yielding a fit model with sufficient flexibility

CMS 138 fb⁻¹ (13 TeV)

- ★ Powheg+OL+P8 $t\bar{t}b\bar{b}$ 4FS ($\mu_R = \frac{1}{2}\prod m_T^{1/4}, \mu_F = \frac{1}{4}\sum m_T$)
- ★ Powheg+P8 $t\bar{t}$ 5FS ($\mu_{R/F} = m_T$)
- ★ Powheg+H7 $t\bar{t}$ 5FS ($\mu_{R/F} = m_T$)
- ★ Sherpa+OL $t\bar{t}b\bar{b}$ 4FS ($\mu_R = \prod m_T^{1/4}, \mu_F = \frac{1}{2}\sum m_T$)
- ★ MG5_aMC+P8 $t\bar{t}b\bar{b}$ 4FS ($\mu_{R/F} = \sum m_T$)
- ★ MG5_aMC+P8 $t\bar{t}$ +jets FxFx 5FS ($\mu_{R/F} = \sum m_T$)

[JHEP 05 \(2024\) 042](#)

≥ 5 jets: $\geq 3b$
 ≥ 6 jets: $\geq 3b, \geq 3$ light
 ≥ 6 jets: $\geq 4b$
 ≥ 7 jets: $\geq 4b, \geq 3$ light



[JHEP 05 \(2024\) 042](#)

Data-driven estimation:

- One control region (CR) for each contribution

$t\bar{t} + \geq 2c$

$t\bar{t} + c$

$t\bar{t} + \geq 2b$

$t\bar{t} + b$

$t\bar{t} + \text{light}$

*For $t\bar{t}$ +jets: 184 nuisance parameters in fit
+ 10 free-floating parameters*

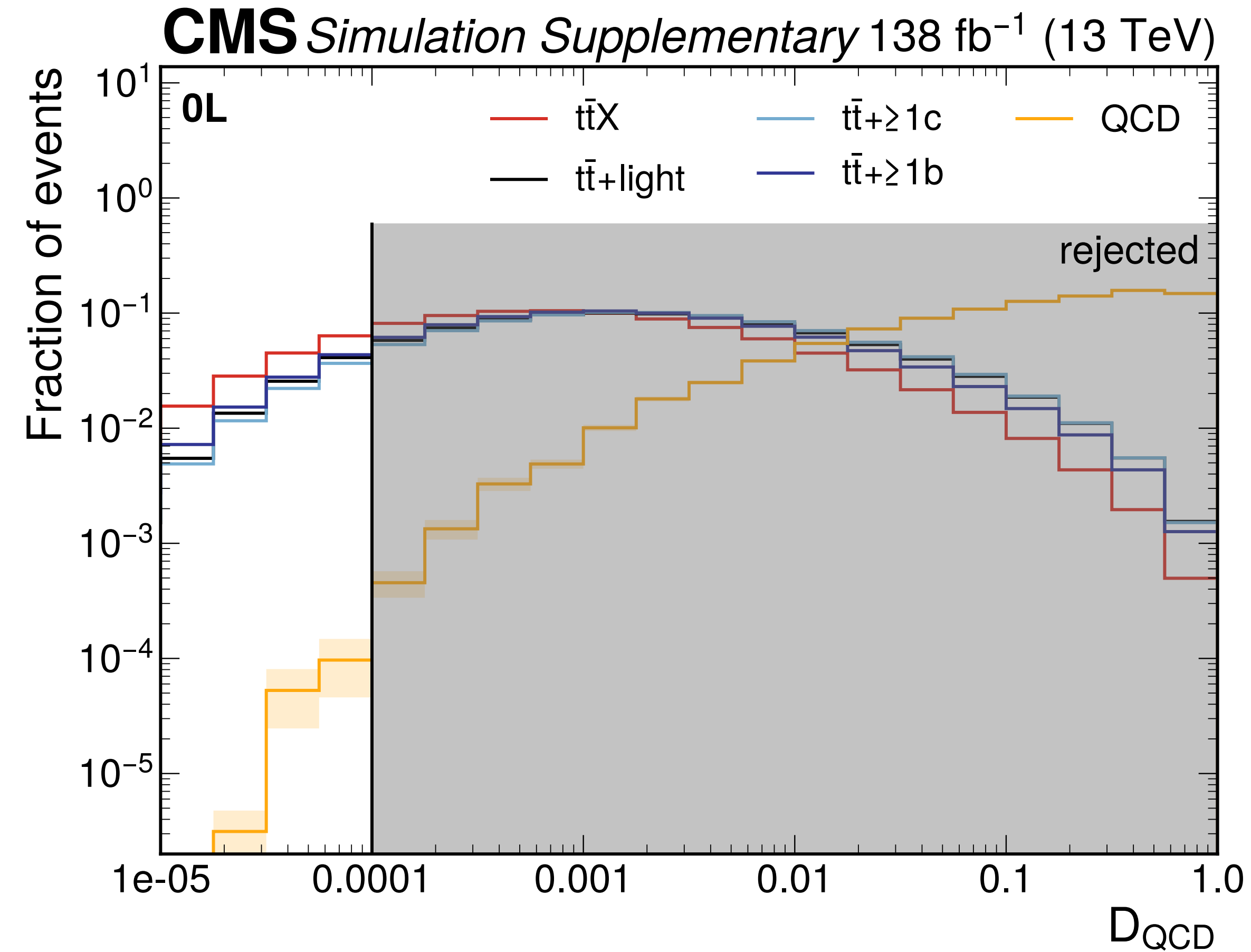


Event selection & categorisation

- Use scores (D_X) to select & categorise events

QCD suppression in FH channel:

- Event classifier very good at separating QCD $\leftrightarrow$ $t\bar{t}$
 - Only use events with QCD score $< 10^{-4}$
 - **QCD contribution becomes negligible ($< 1\%$)**
 - At the cost of signal efficiency ($\sim 15\%$)
- Same strategy & background estimation in 0L, 1L, 2L





Event selection & categorisation

- Use scores (D_X) to select & categorise events

Sideband for validation (VR)

- Independently test the background model

Sideband for background estimation (CR)

- **5 control regions:**

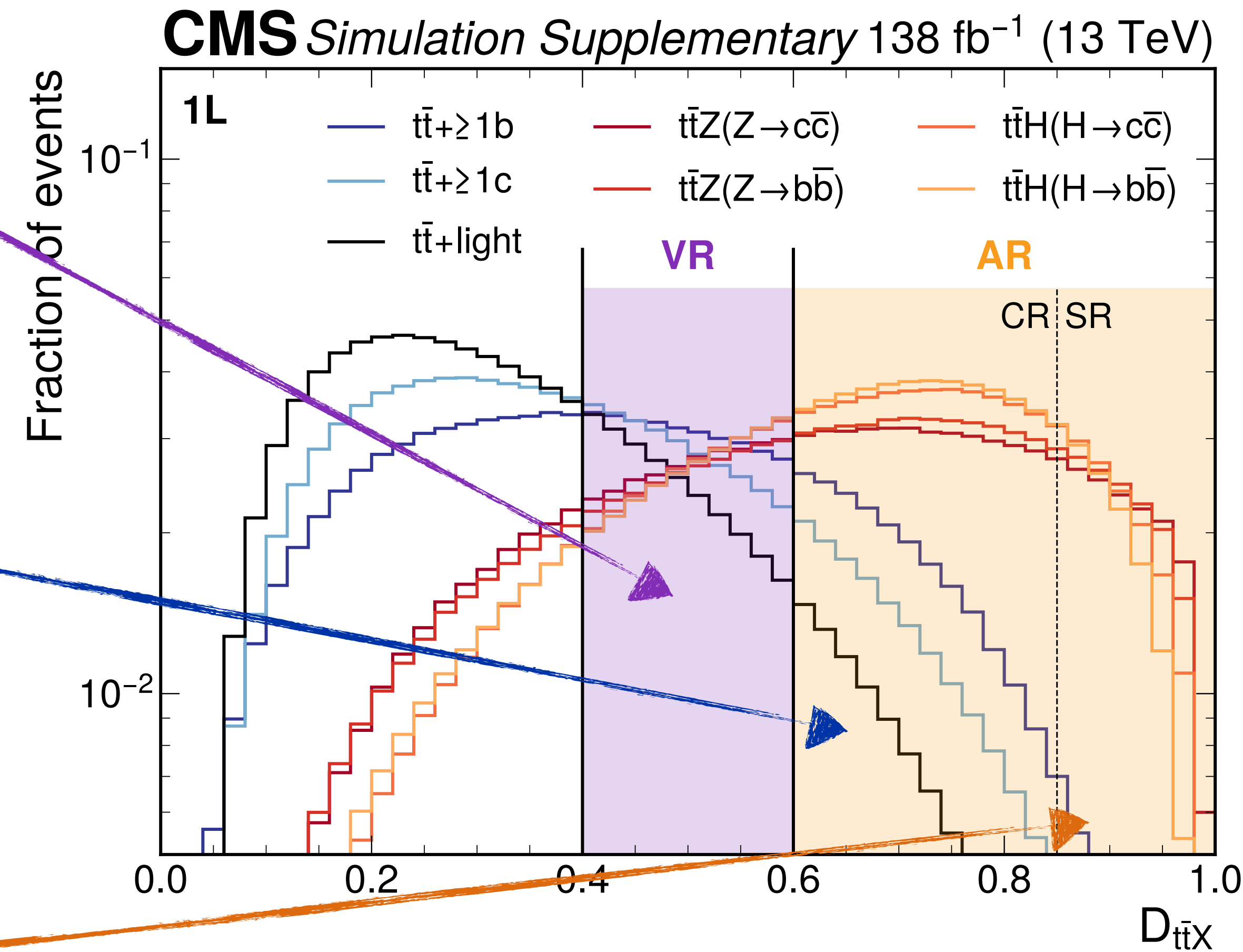
- $t\bar{t} + \text{light}$, $t\bar{t} + c$, $t\bar{t} + \geq 2c$, $t\bar{t} + b$, $t\bar{t} + \geq 2b$

Signal sensitive region (SR)

→ Reduced background extrapolation

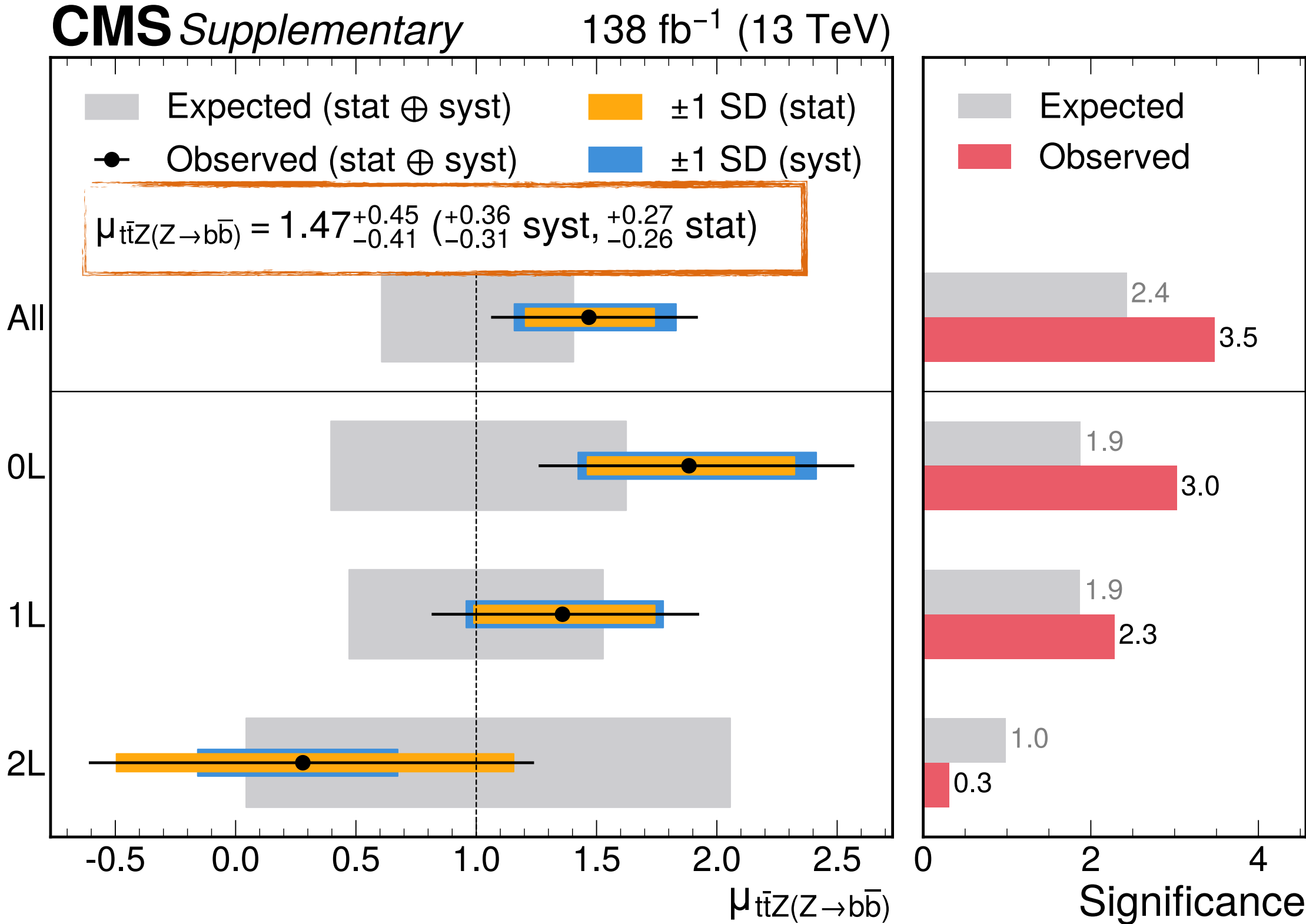
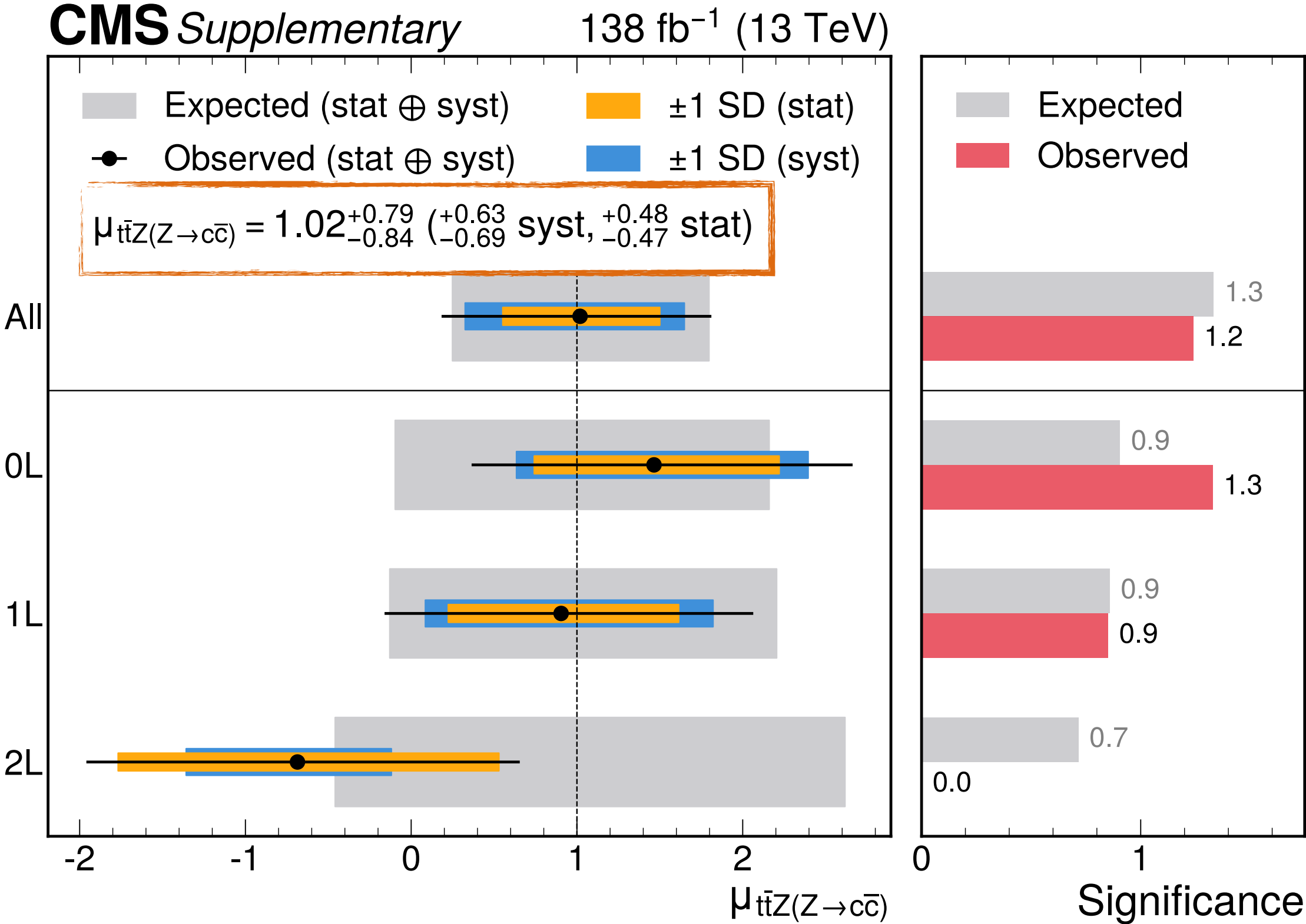
- **2+2 dedicated signal regions:**

- $t\bar{t}H(H \rightarrow c\bar{c})$, $t\bar{t}H(H \rightarrow b\bar{b})$, $t\bar{t}Z(Z \rightarrow c\bar{c})$, $t\bar{t}Z(Z \rightarrow b\bar{b})$



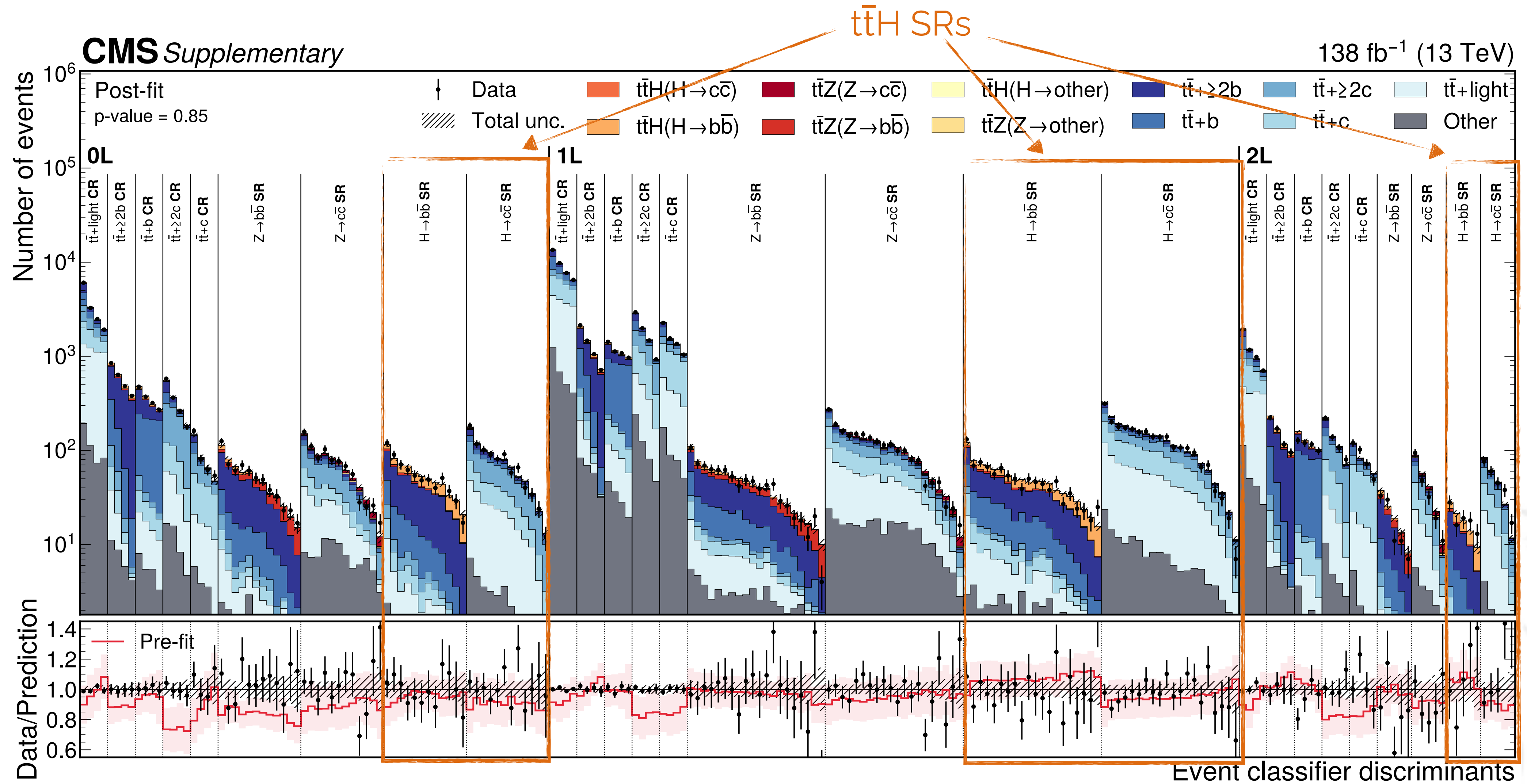
- Split **AR** into **signal region part (SR)** and **control region part (CR)**

Validation with $t\bar{t}Z(Z \rightarrow c\bar{c}/b\bar{b})$ production



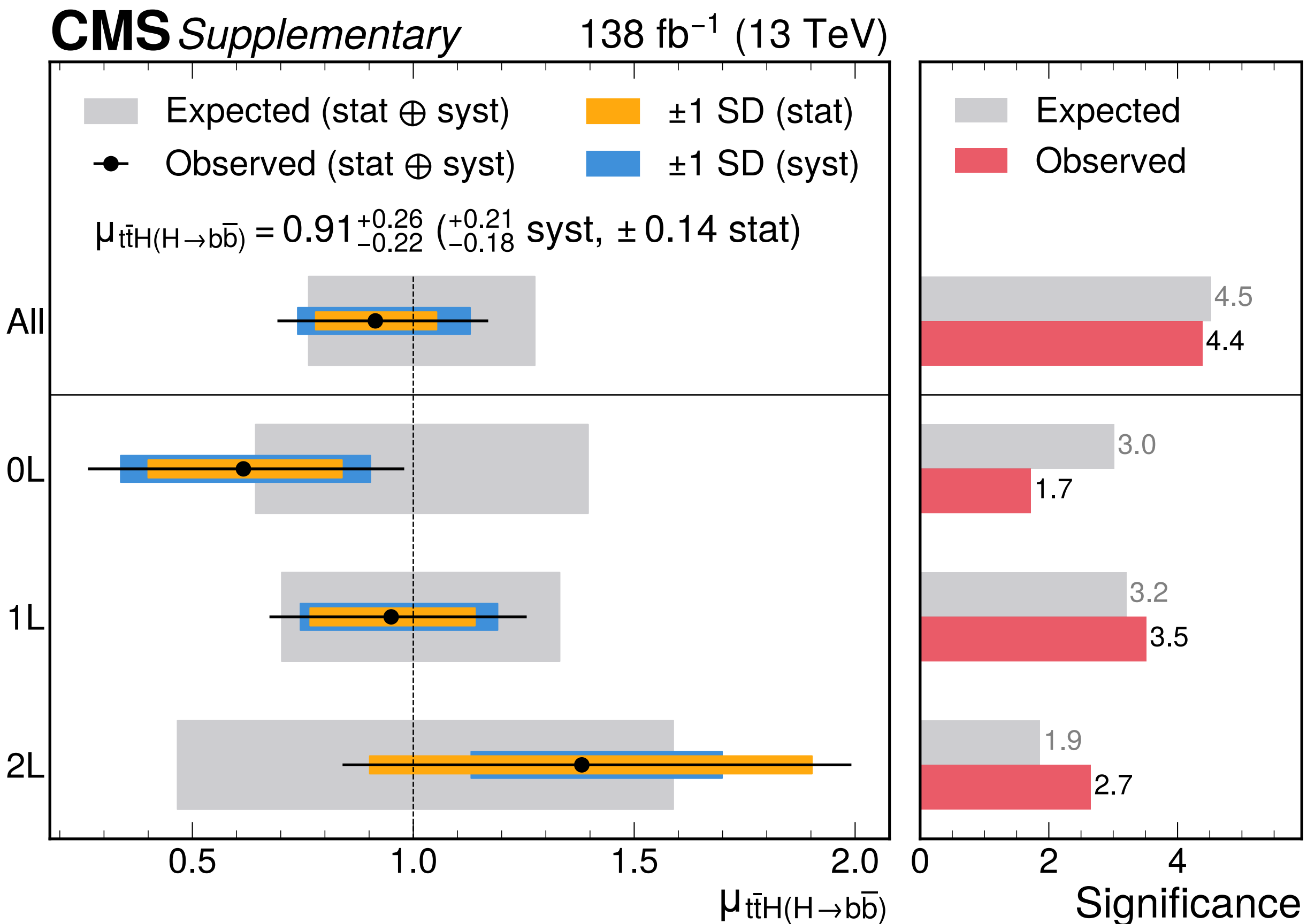
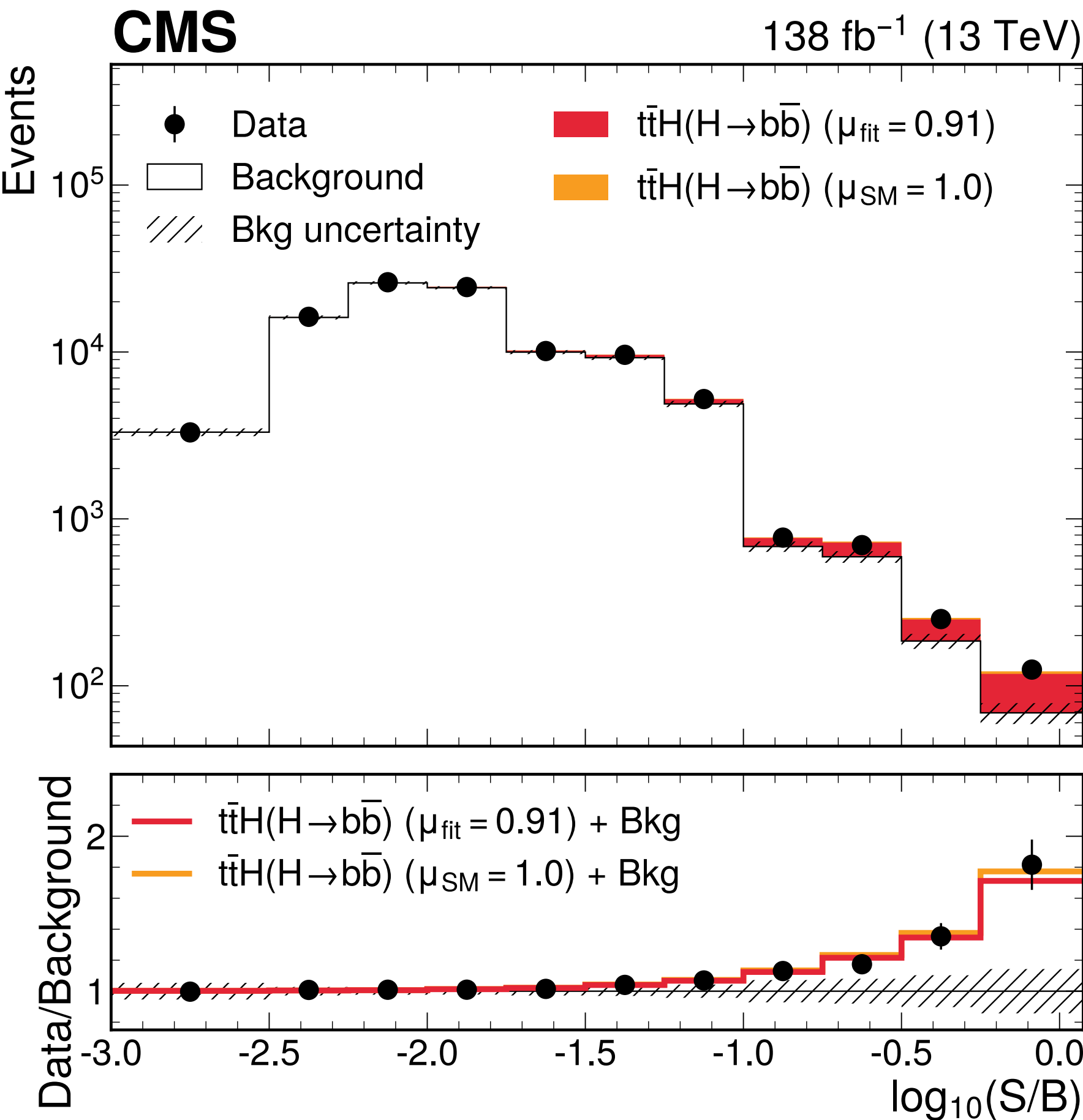
- **$t\bar{t}Z$ signal strengths** compatible with SM and LHC results
 - From leptonic channels: $\mu_{t\bar{t}Z} \sim 1.1$ with 8% precision
- **Good validation for $t\bar{t}H(H \rightarrow c\bar{c}/b\bar{b})$ analysis!**

Comparing data to prediction in the full fit distribution



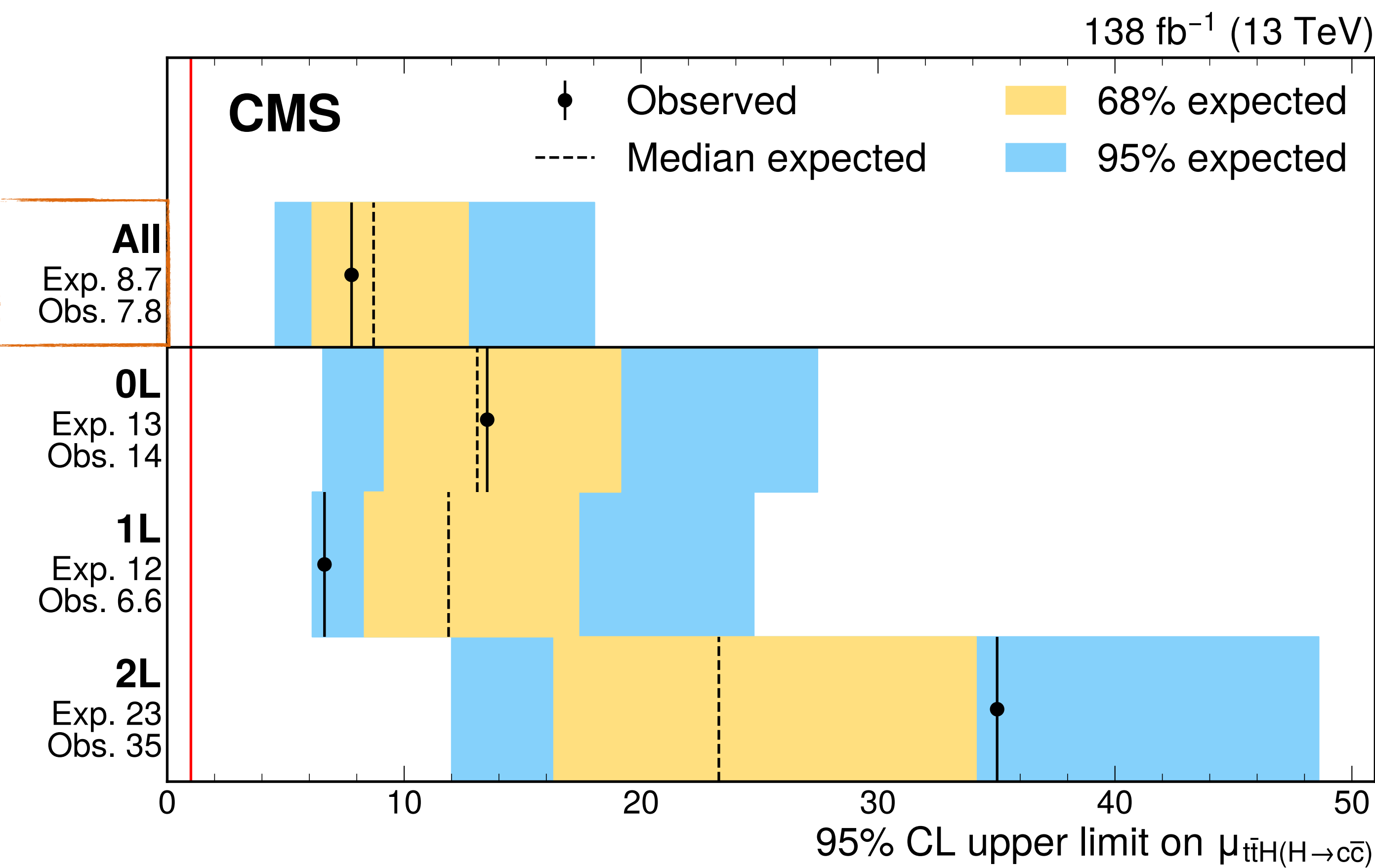
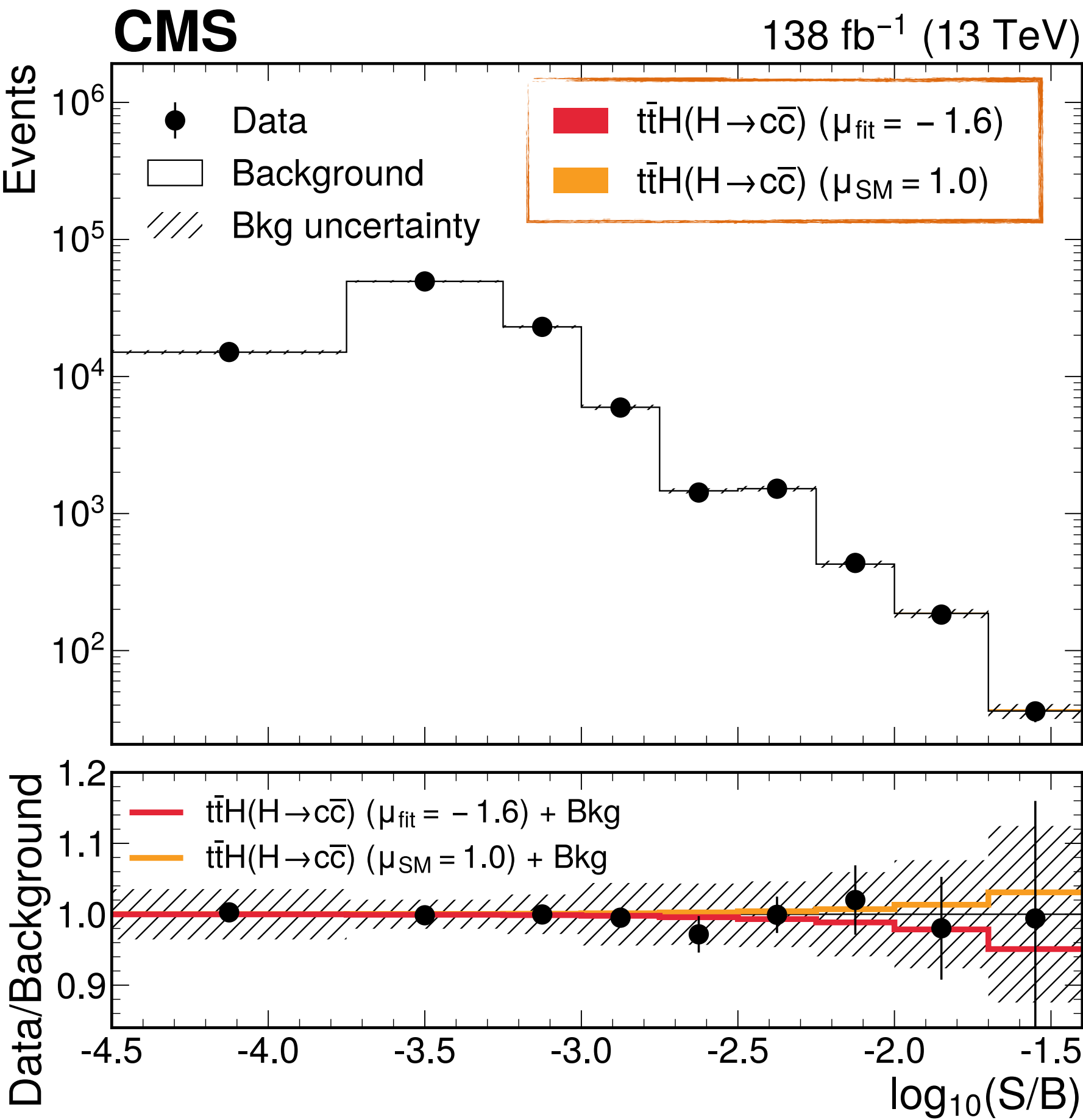
- Four POIs simultaneously extracted – $t\bar{t}H(H \rightarrow c\bar{c})$, $t\bar{t}H(H \rightarrow b\bar{b})$, $t\bar{t}Z(Z \rightarrow c\bar{c})$, $t\bar{t}Z(Z \rightarrow b\bar{b})$
→ Very good pre- and post-fit agreement!

Results: $t\bar{t}H(H \rightarrow b\bar{b})$ production



- $\mu_{t\bar{t}H(H \rightarrow b\bar{b})} = 0.91^{+0.26}_{-0.22} - 4.4\sigma$ observed
 - Good agreement with SM expectation

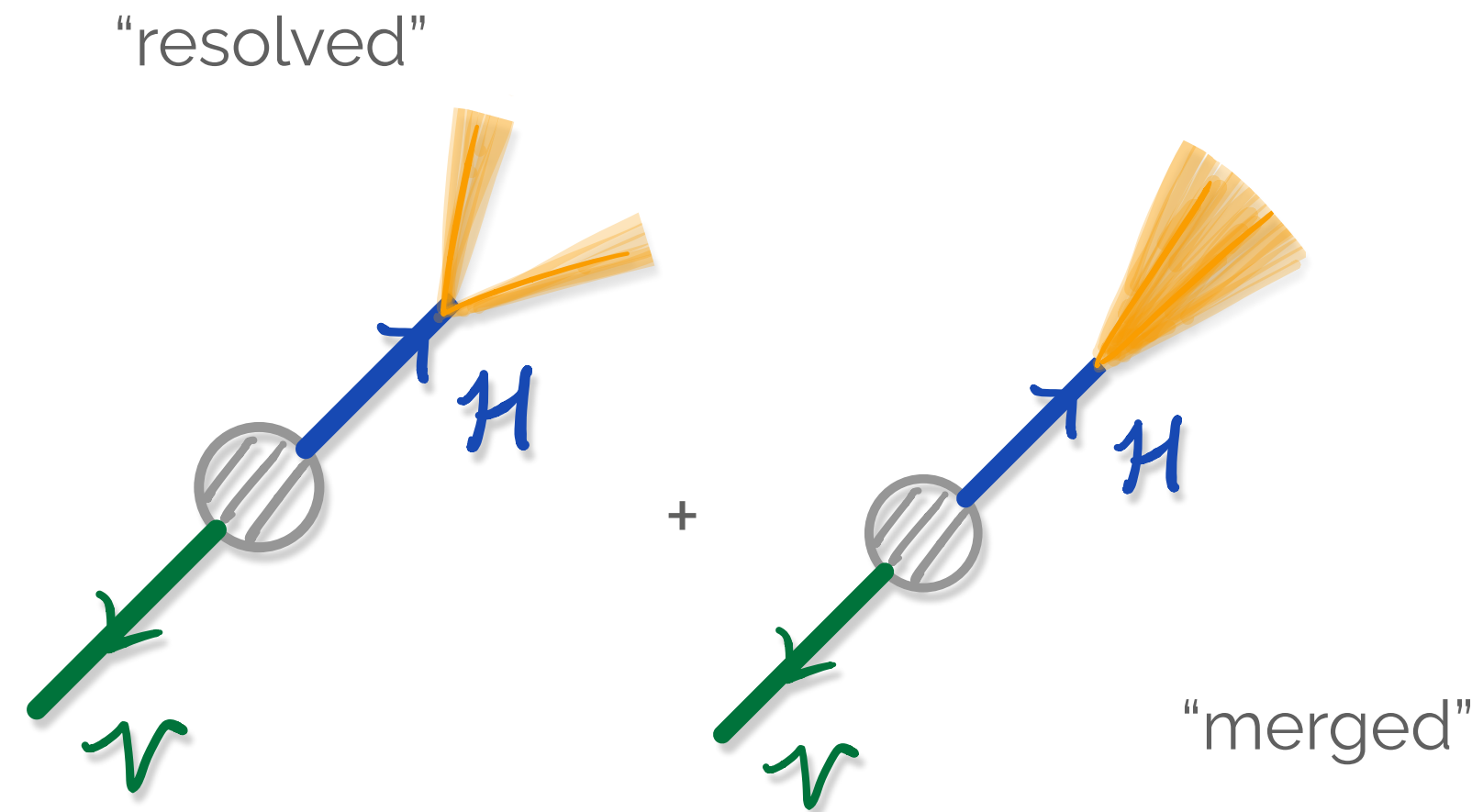
Results: $t\bar{t}H(H \rightarrow c\bar{c})$ production



- **Limits on $\mu_{t\bar{t}H(H \rightarrow c\bar{c})}$: $\lesssim 7.8$ wrt. SM observed ($\lesssim 8.7$ expected)**
 - Similar sensitivity as compared to $VH(cc)$ measurements!
From $VH(cc)$: $\mu \lesssim 12$ ($\lesssim 7.6$ expected)
- **1D constraints: $|\kappa_c| \lesssim 3.0$ obs. (3.3 exp.)**
→ **Most stringent constraints to date!**



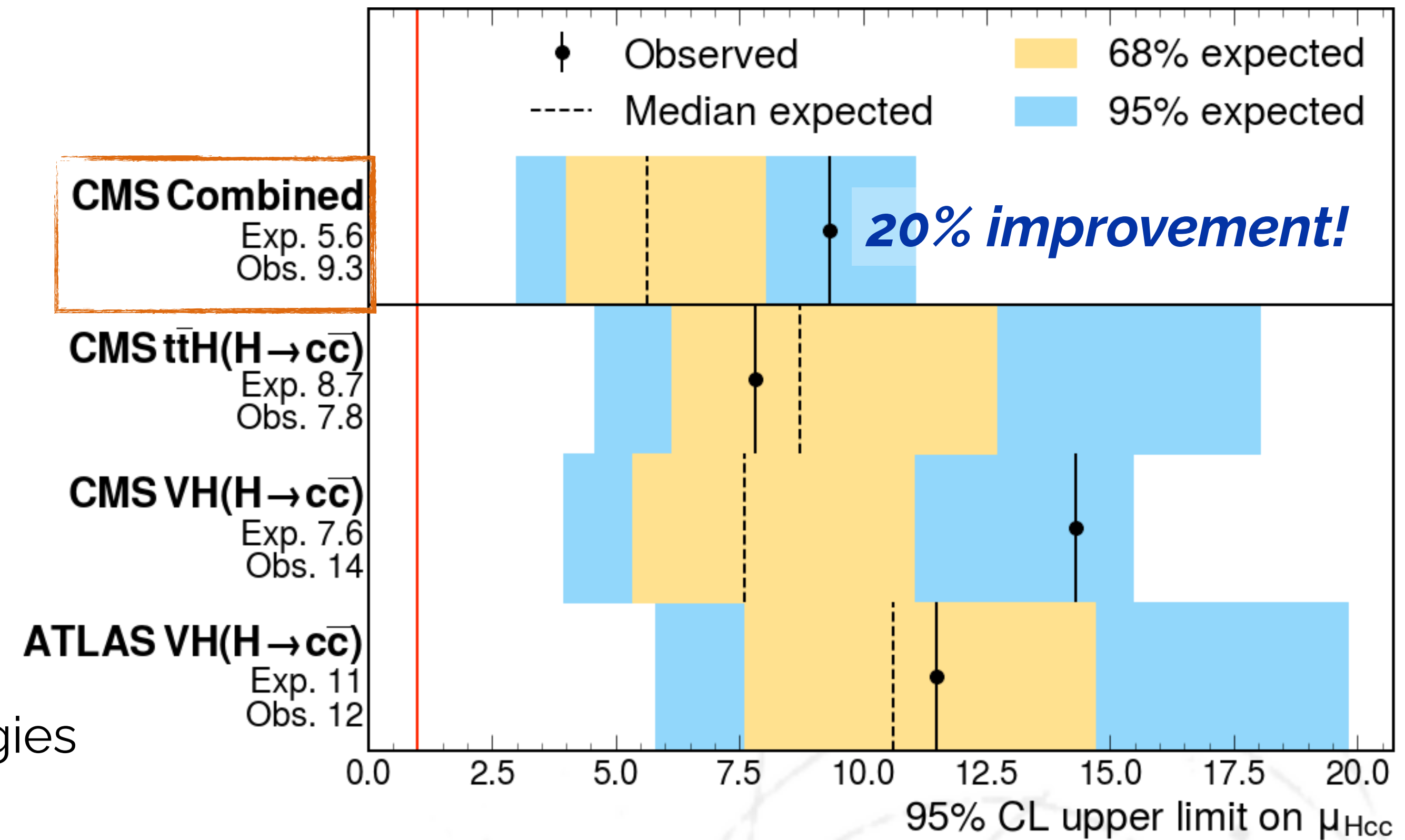
Combination with $VH(H \rightarrow c\bar{c})$



- Combine with $VH(H \rightarrow c\bar{c})$:

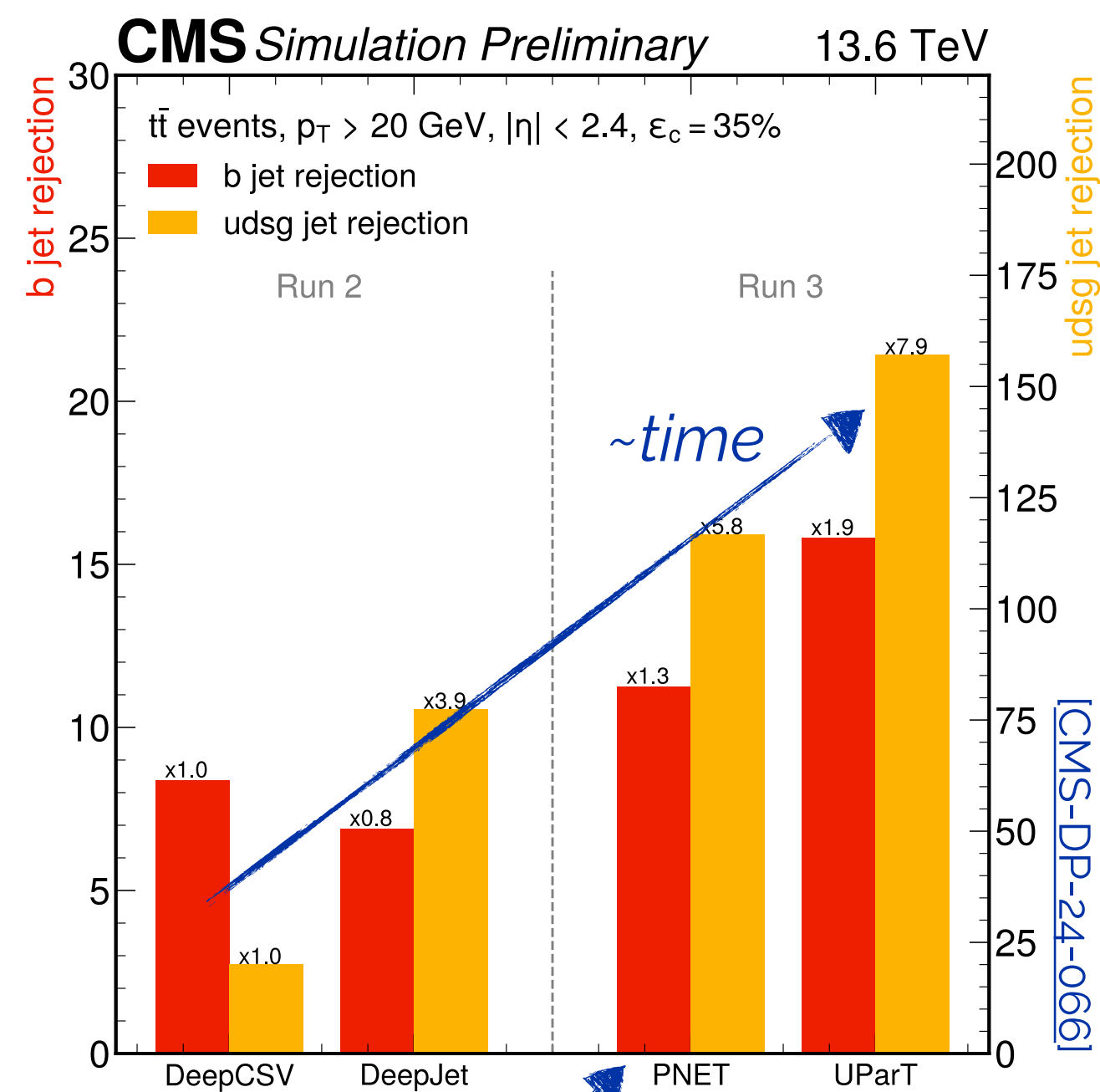
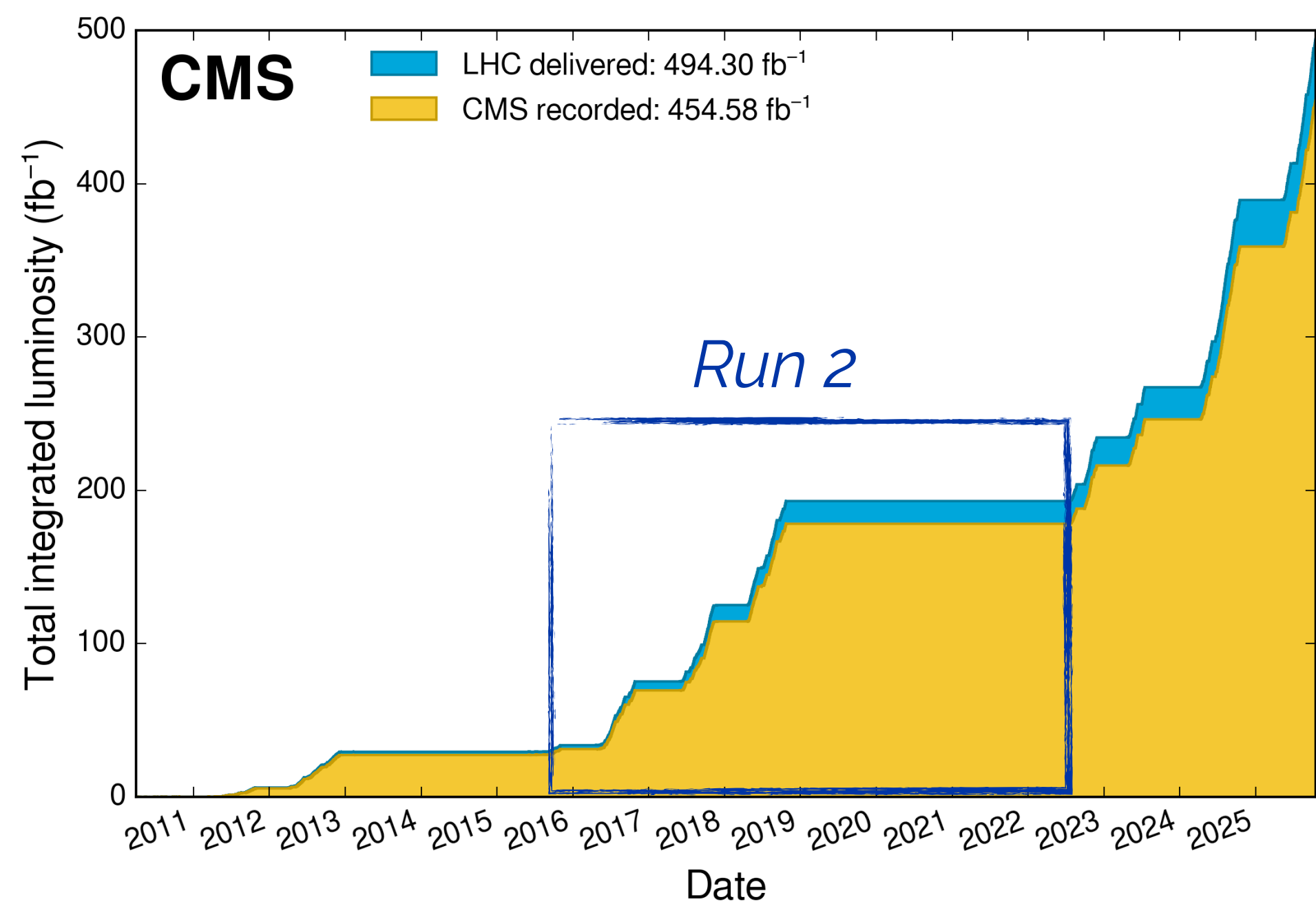
- Full Run 2 result from resolved+merged topologies
- Correlated experimental uncertainties
- Uncorrelated tagging uncertainties
- Correlated theor. uncertainties (cross section / BR)

- $VH(H \rightarrow c\bar{c})$: $1.1 < |\kappa_c| < 5.5$ (< 3.4 exp.)
- $t\bar{t}H$: $|\kappa_c| < 3.0$ (< 3.3 exp.)
- Combined: $|\kappa_c| < 3.5$ (< 2.7 exp.)

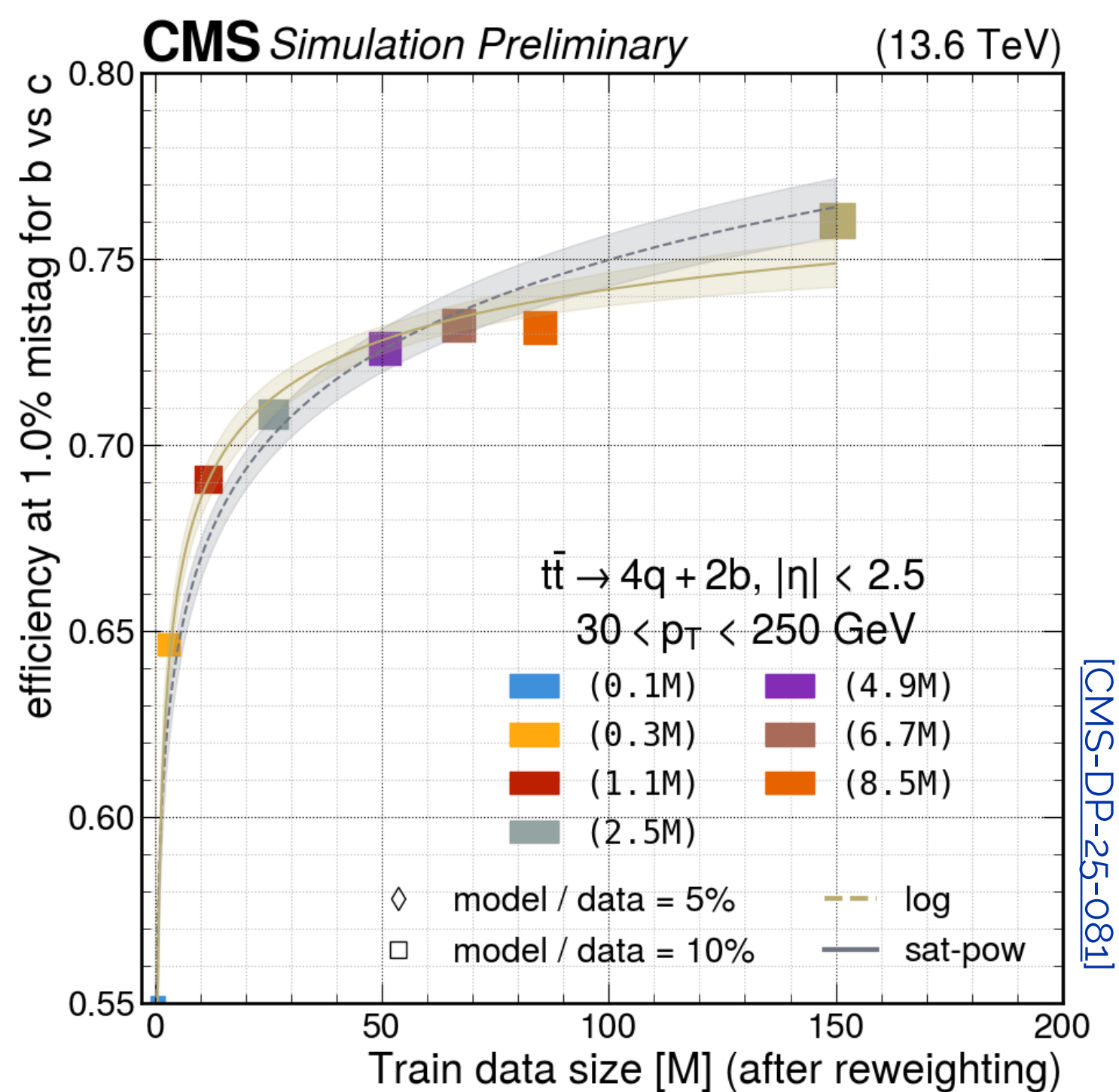


Combined signal strength interpretation:
 $\mu_{(H \rightarrow c\bar{c})} < 9.3$ (5.6 expected)

Outlook: What's next?



We are not saturating..

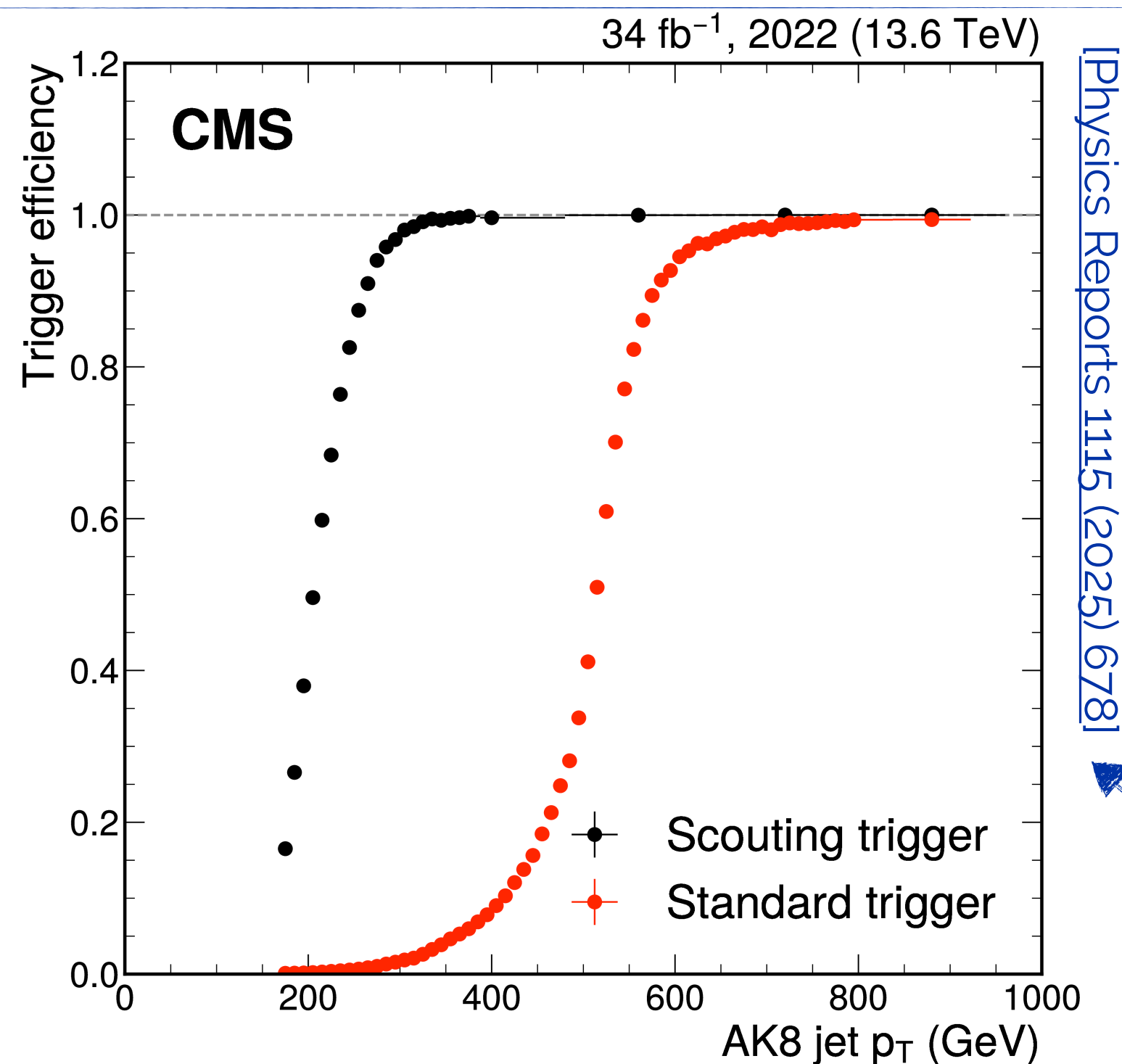


- Run 3 data taking:
 - With 2025+26 runs ~ 500/fb
- Will allow us to extend the analyses:
 - Explore merged regime in $t\bar{t}H$?
 - Additional channels such as $H+c$ & VBF

Used in results shown today

- Improvements in tagging are impressive, but can only help us that far...
 - Backgrounds are likely also real charm quarks
- Improvement in modeling & calibration needed to reduce uncertainties

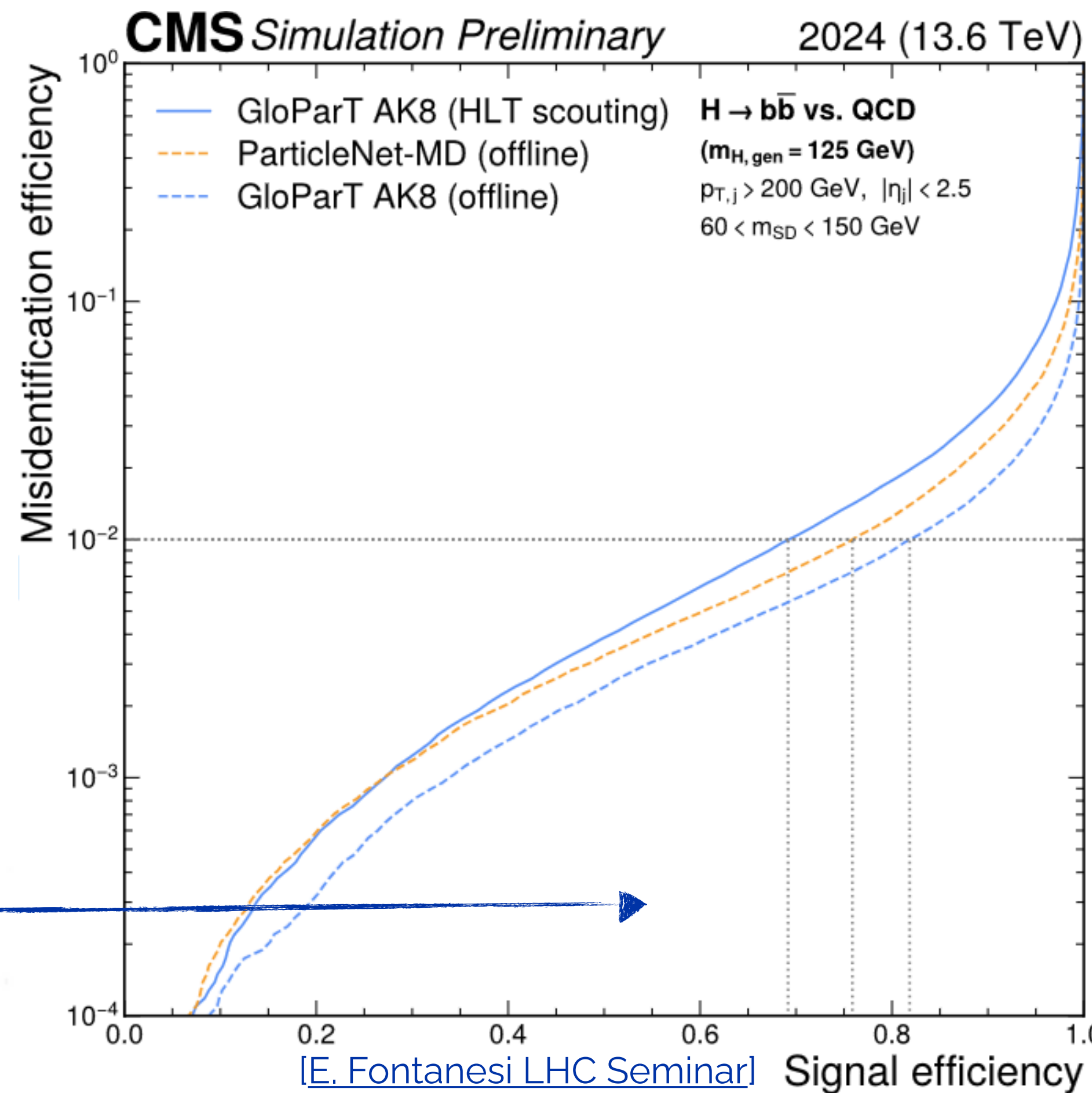
Outlook: What's next?



[Physics Reports 1115 (2025) 678]

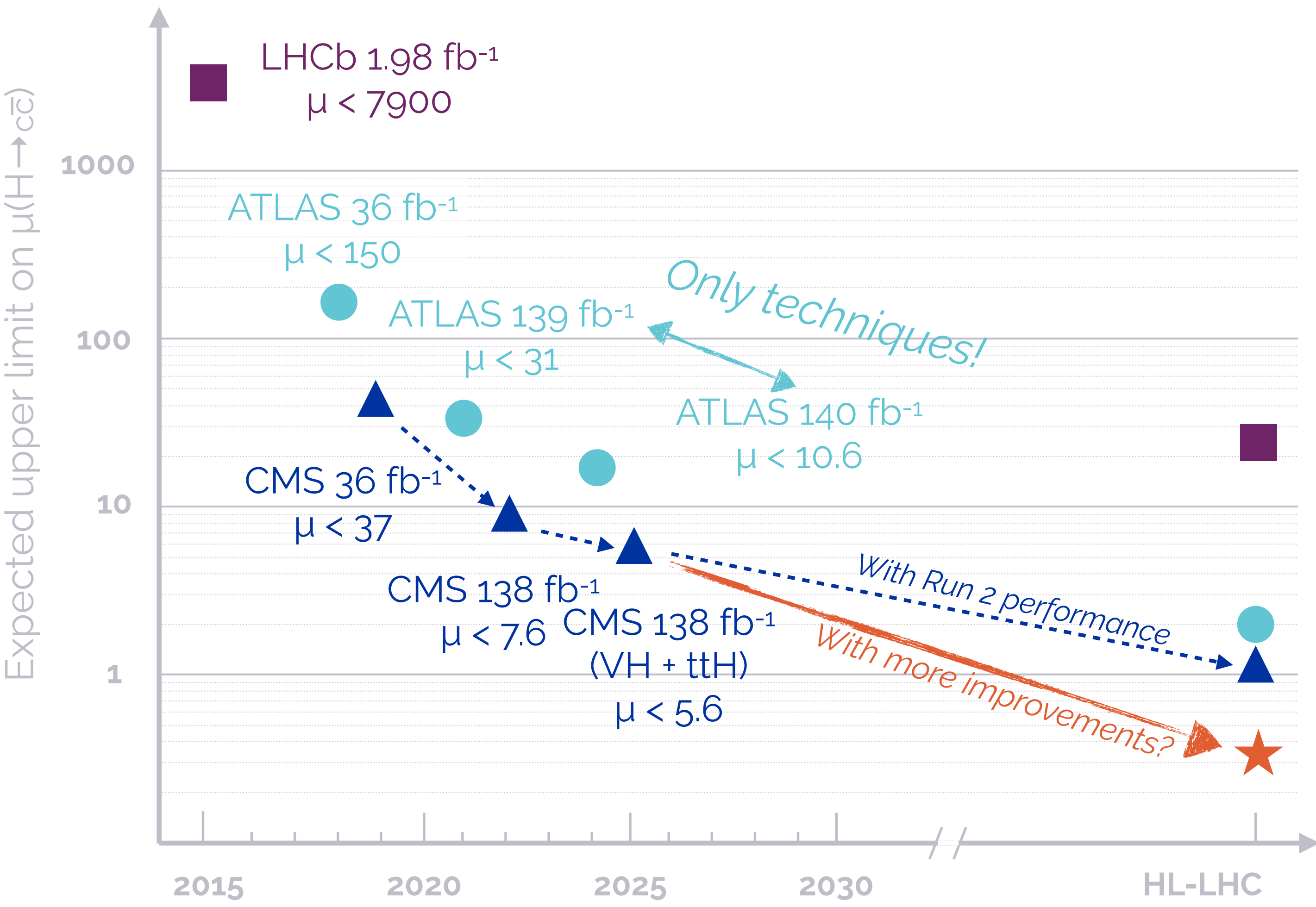
- Run 3 data taking:
 - With 2025+26 runs ~ 500/fb
- Will allow us to extend the analyses:
 - Explore merged regime in ttH?
 - Additional channels such as H+c & VBF

Usage of advanced taggers in software trigger and HLT scouting



[E. Fontanesi LHC Seminar]

Conclusions & some outlook...?



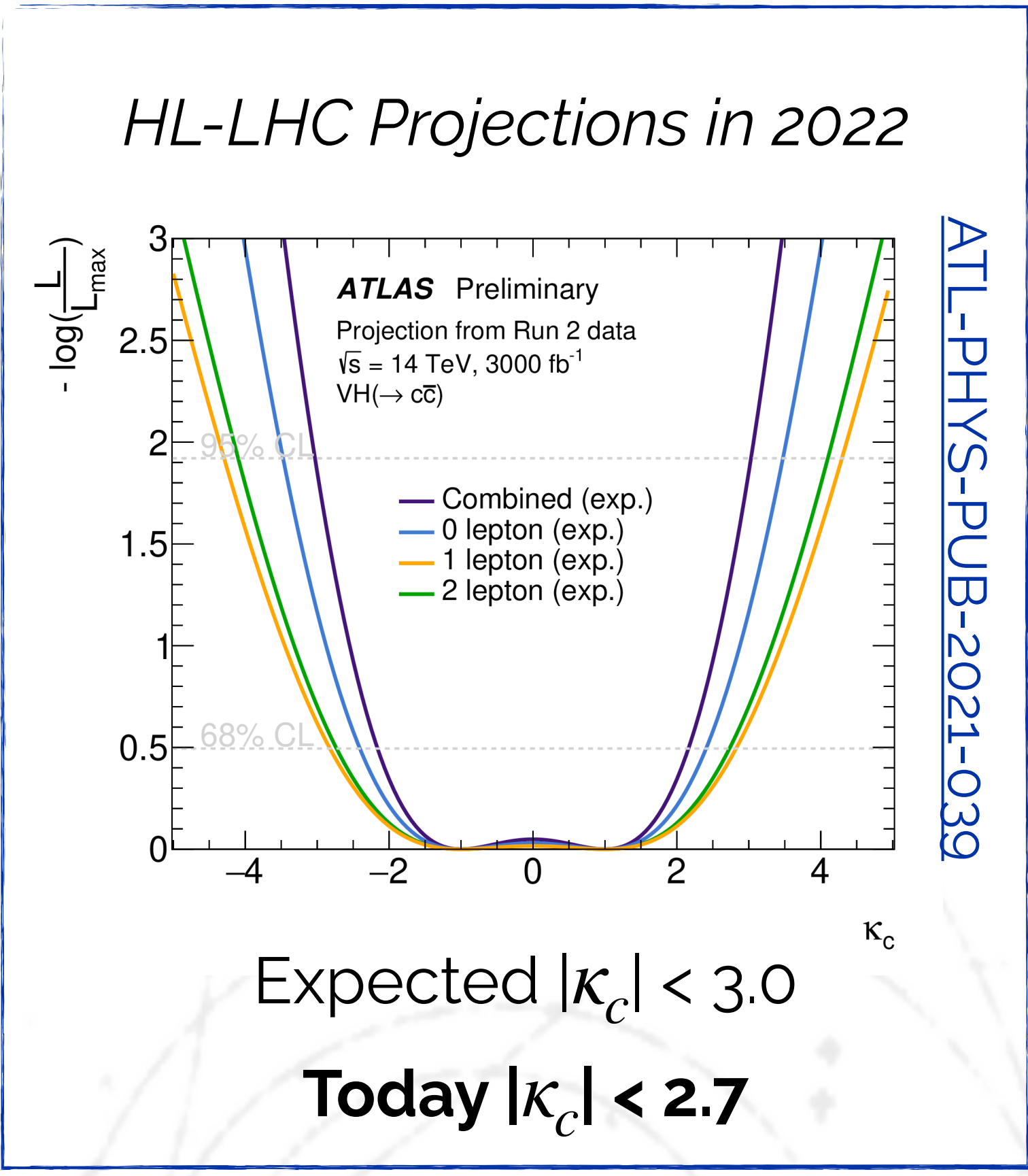
[S. Wuchterl LHC Seminar]

LHCb 300 fb⁻¹
 $\mu < O(10)$

ATLAS 3000 fb⁻¹
 $\mu < O(2)$

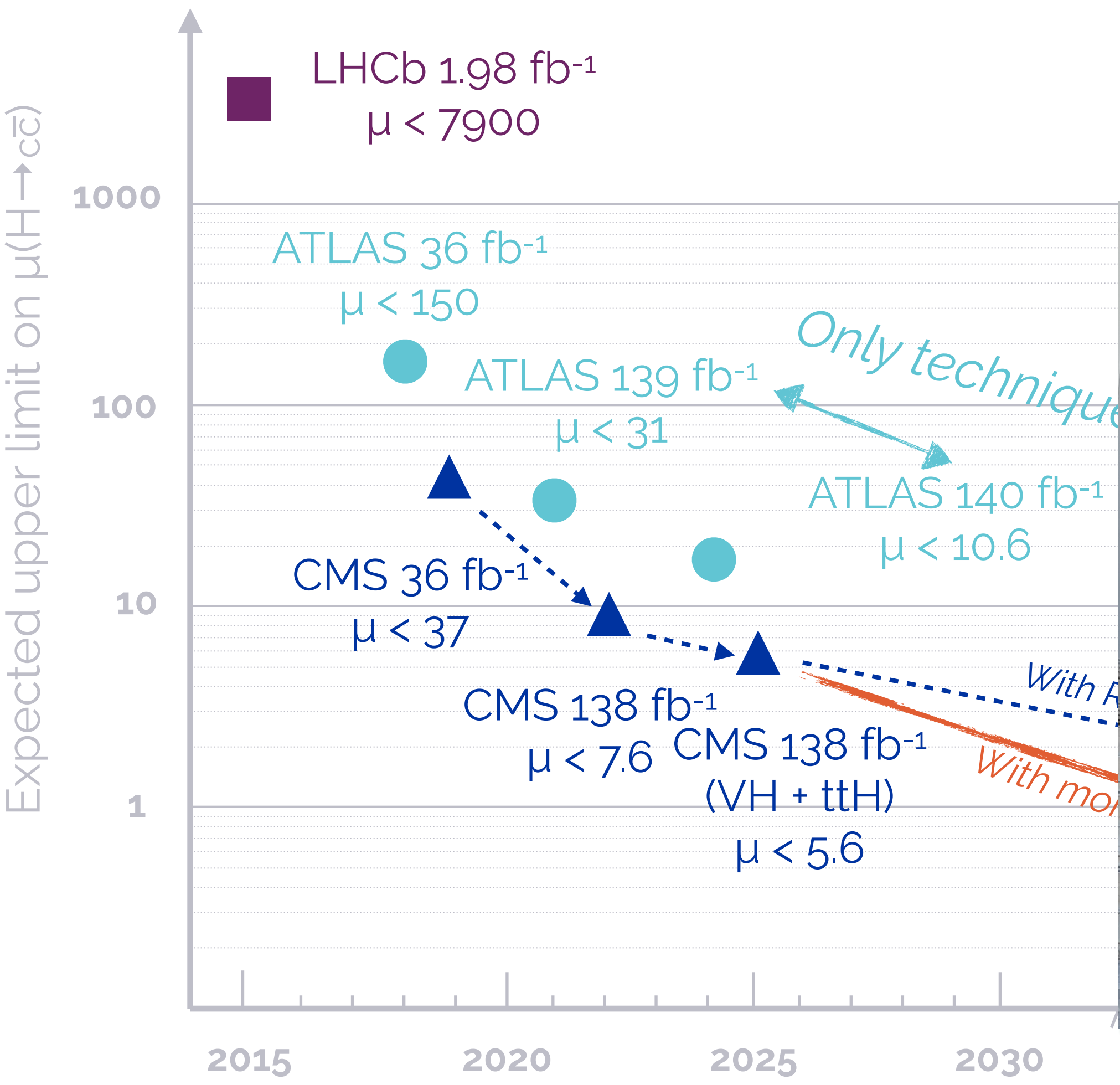
CMS 3000 fb⁻¹
 $\mu < O(1)$

A first evidence at HL-LHC?



- Compared to old ATLAS VH(cc) projections
 - 18 years of progress in 3 years!
- **We have a charming future ahead of us!**

Conclusions & some outlook...?



Is. Wuch

HL-LHC Projections in 2022



2nd CMS Higgs-charm workshop

HL-LHC

Thank you very much for your attention!

- Compared to old ATLAS VH(cc) projections
 - 18 years of progress in 3 years!
- **We have a charming future ahead of us!**

Additional slides





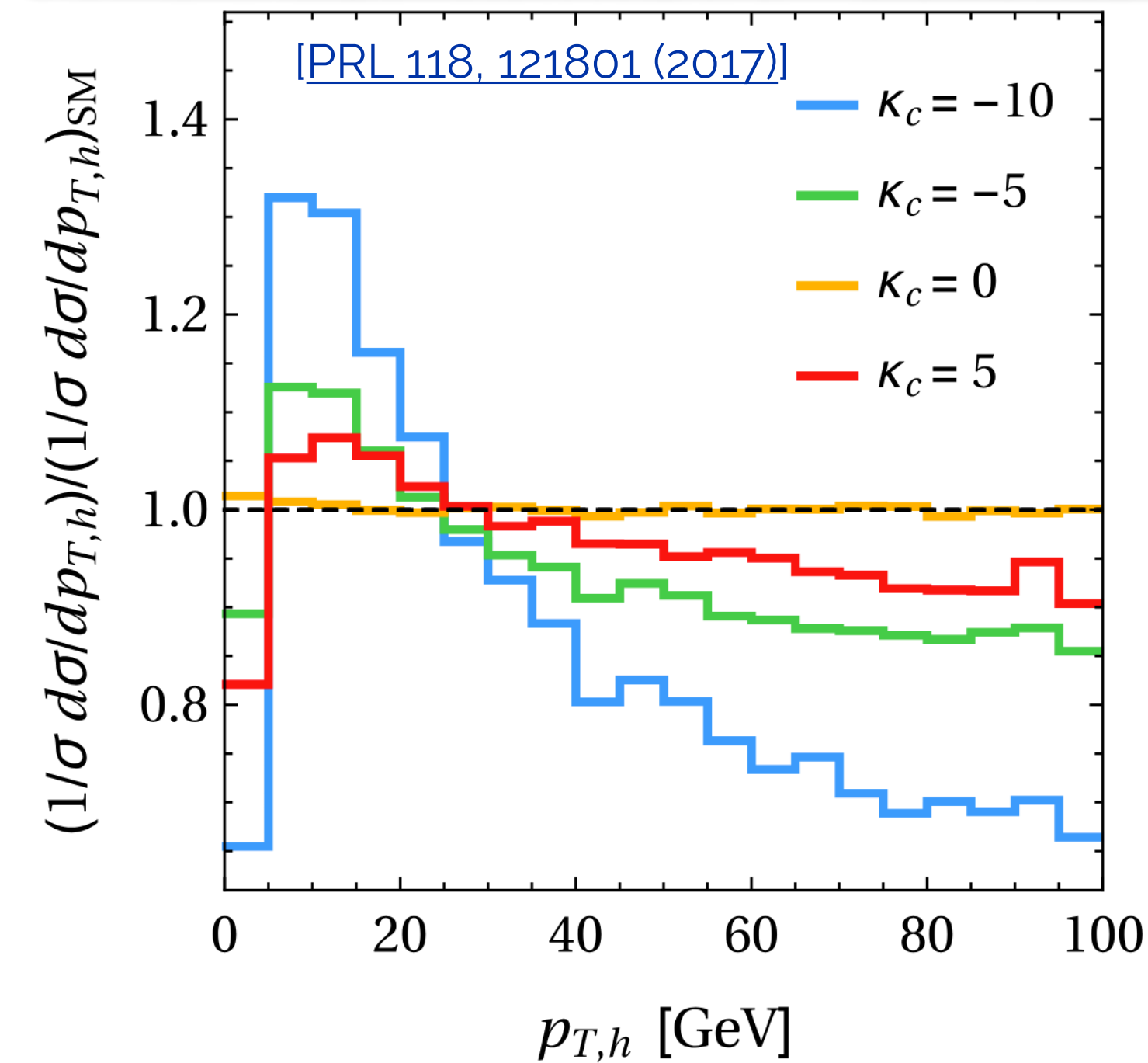
So far: probing the Higgs-charm coupling

- Several methods explored by ATLAS and CMS to probe the Higgs-charm Yukawa coupling (y_c)

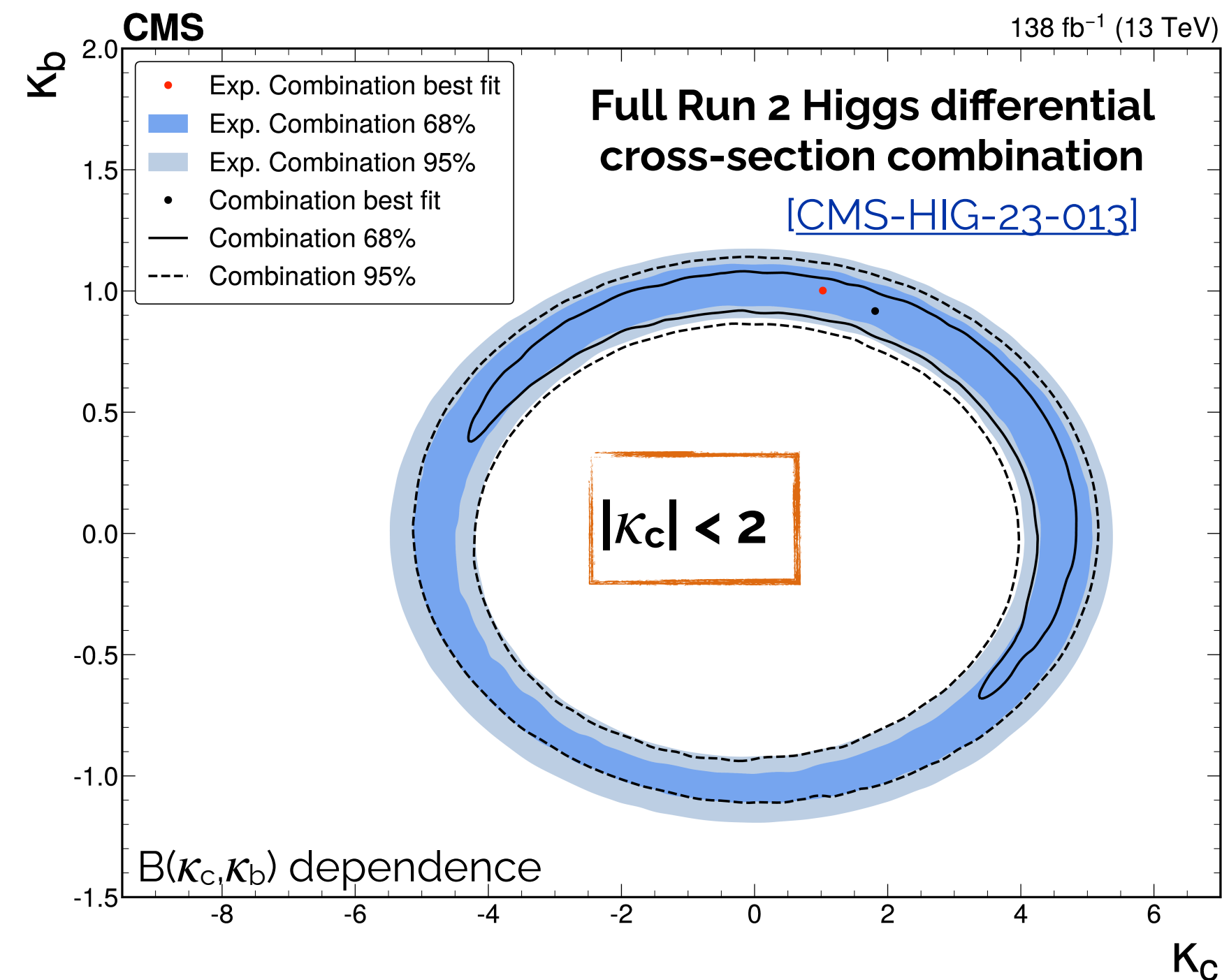
Indirect constraint from Higgs kinematics

Model dependence?

Direct searches for $H \rightarrow c\bar{c}$



- Differential distributions are sensitive to y_c



Kappa framework

$$(\sigma \cdot \mathcal{B})(ii \rightarrow H \rightarrow ff) = \frac{\sigma_{ii} \cdot \Gamma_{ff}}{\Gamma_H} = \sigma_{ii}^{SM} \cdot \mathcal{B}_{ff}^{SM} \cdot \frac{\kappa_i^2 \kappa_f^2}{\kappa_H^2}$$

κ_i : coupling wrt. SM expectation of 1

So far: probing the Higgs-charm coupling

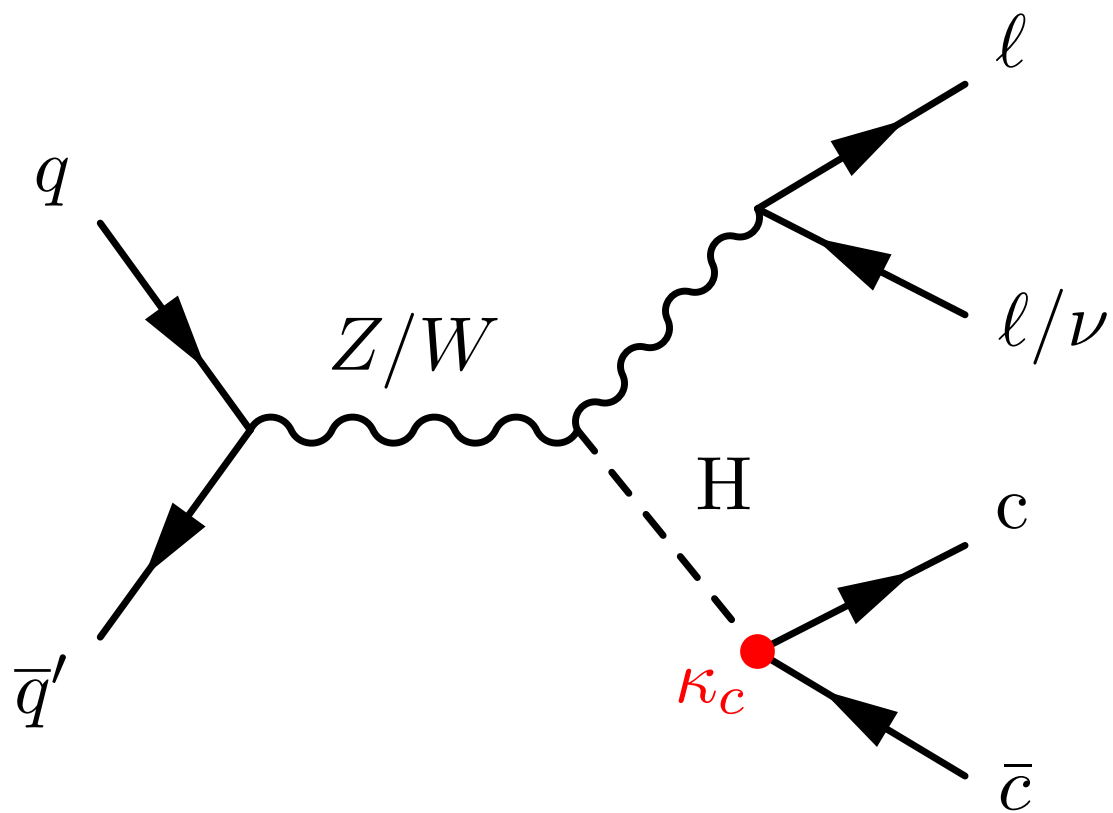


- Several methods explored by ATLAS and CMS to probe the Higgs-charm Yukawa coupling (y_c)

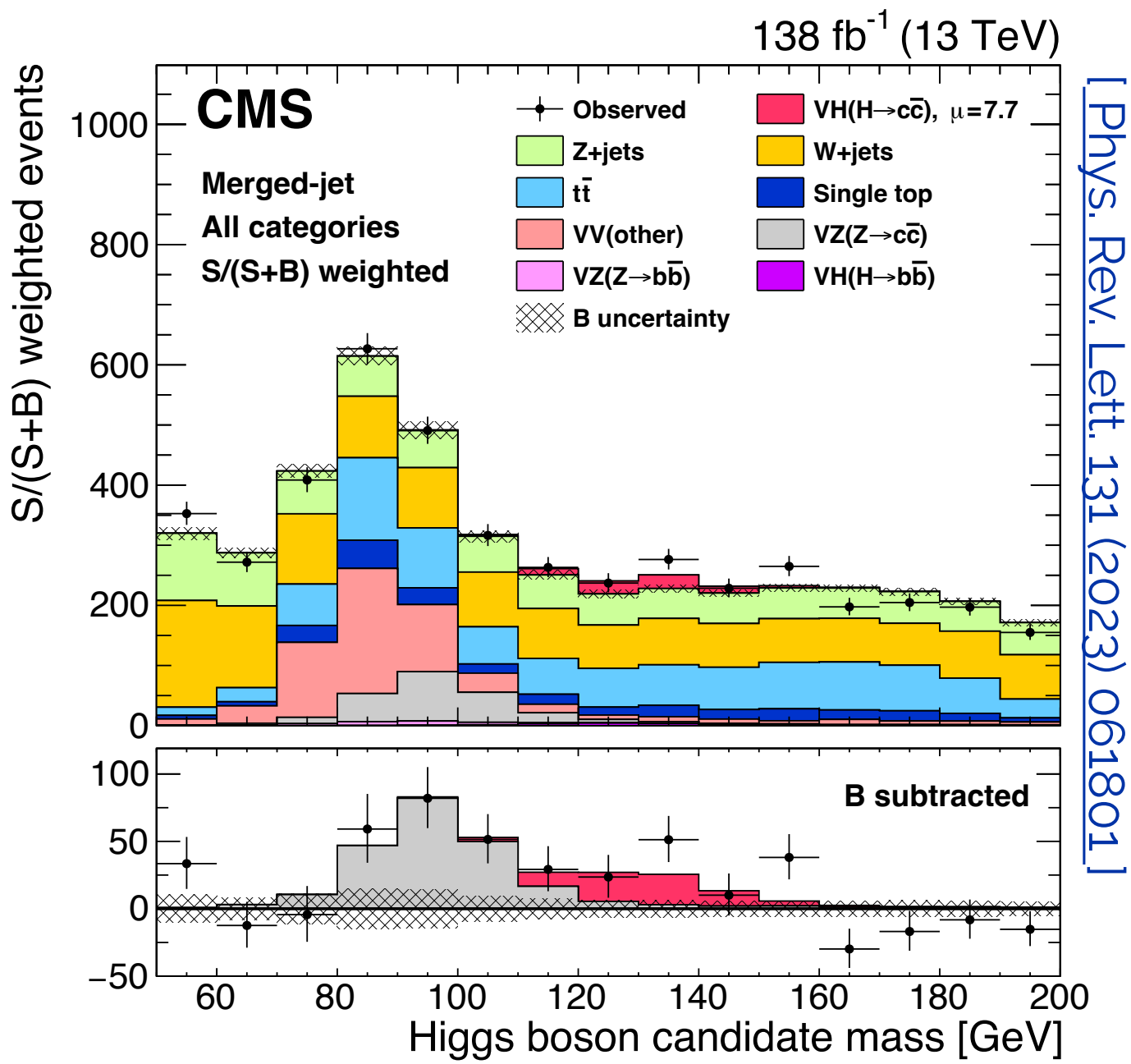
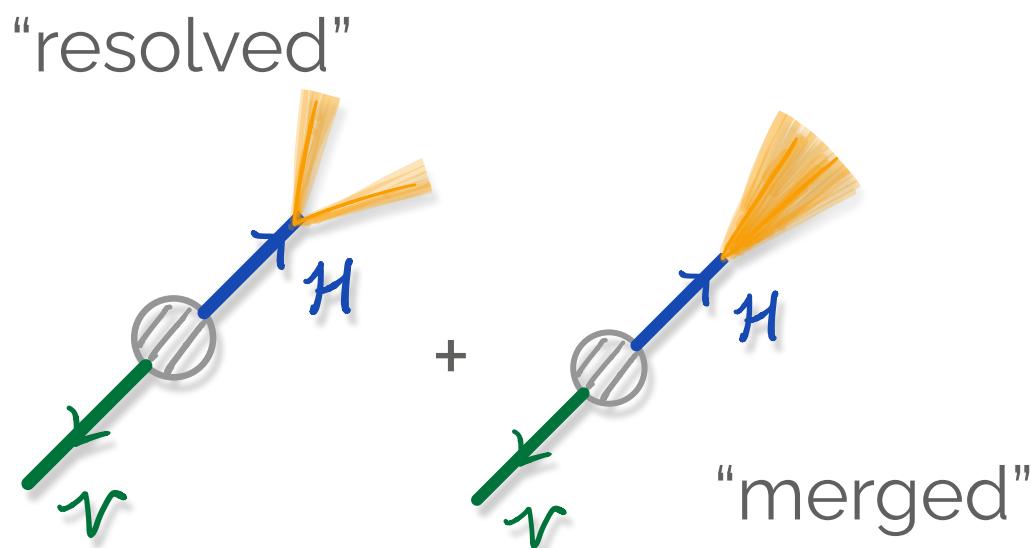
Indirect constraint from Higgs kinematics

Model dependence?

Direct searches for $H \rightarrow c\bar{c}$



- VH as the “golden channel”



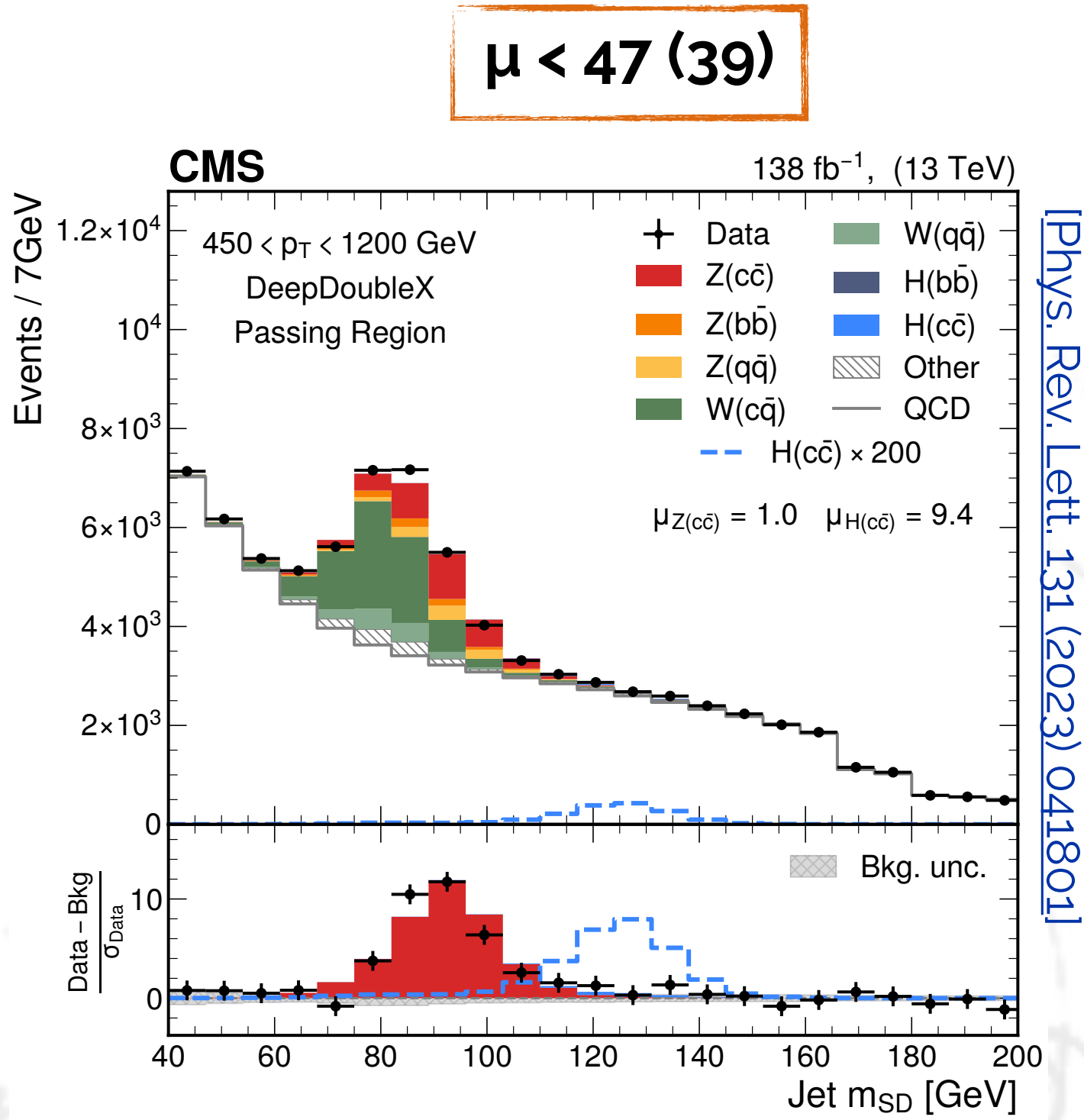
[Phys. Rev. Lett. 131 (2023) 061801]

$$1.1 < |\kappa_c| < 5.5 (< 3.4)$$

$$\mu < 14 (7.6)$$

From ATLAS:
 $|\kappa_c| < 4.2 (< 4.1)$
 $\mu < 12 (11)$
[JHEP 04(2025) 075]

- Inclusive boosted $H \rightarrow c\bar{c}$



[Phys. Rev. Lett. 131 (2023) 041801]

So far: probing the Higgs-charm coupling



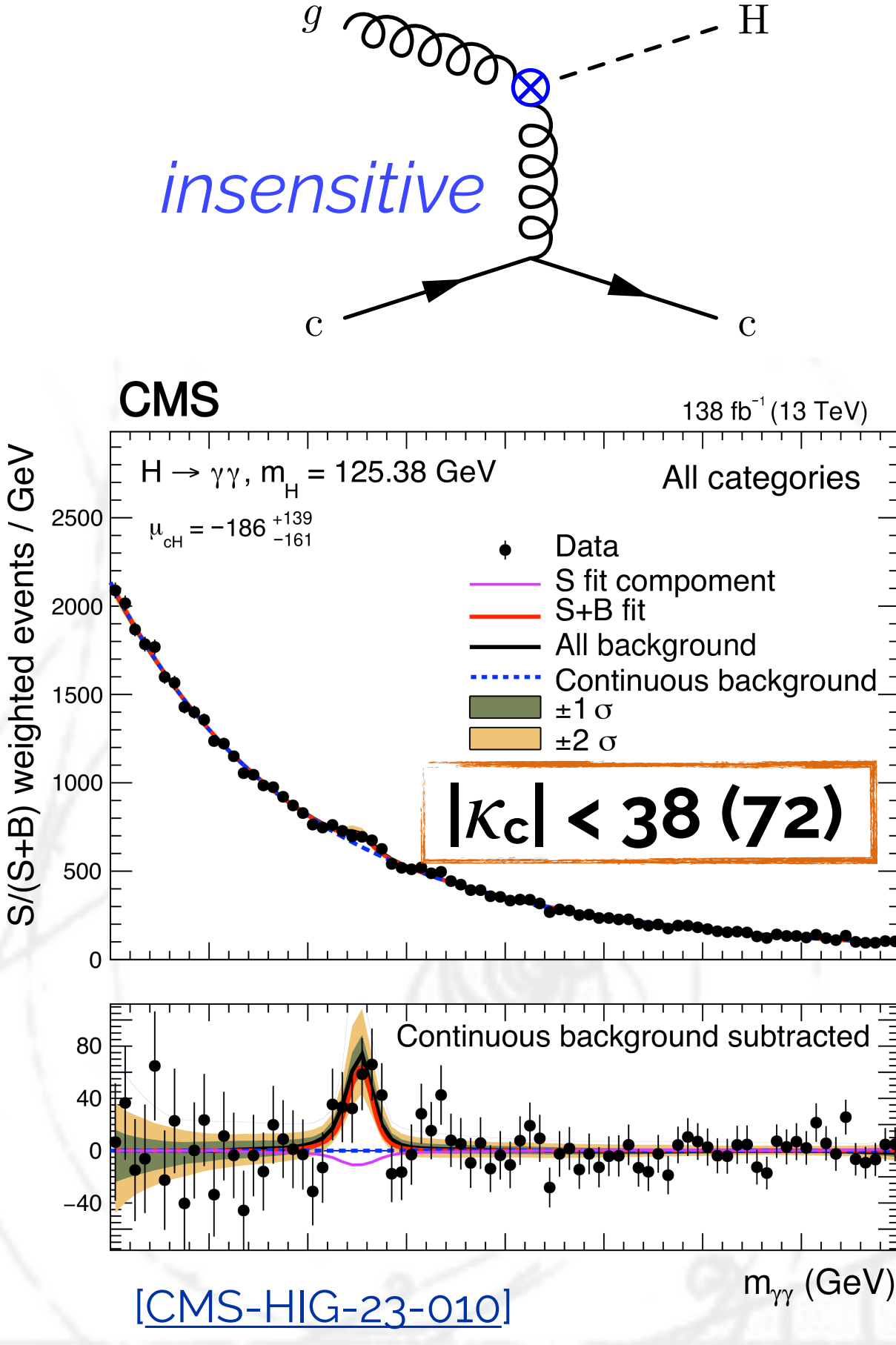
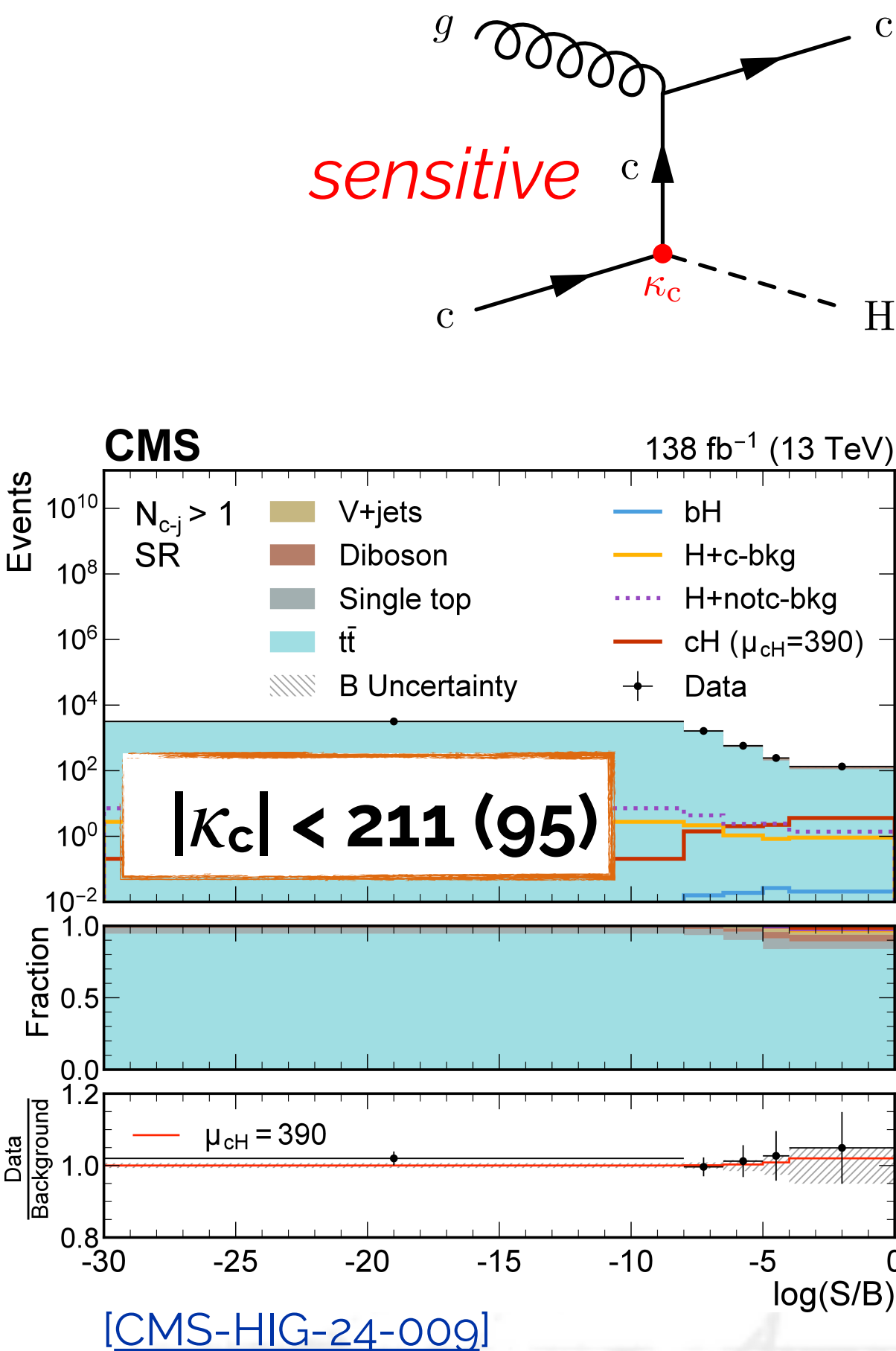
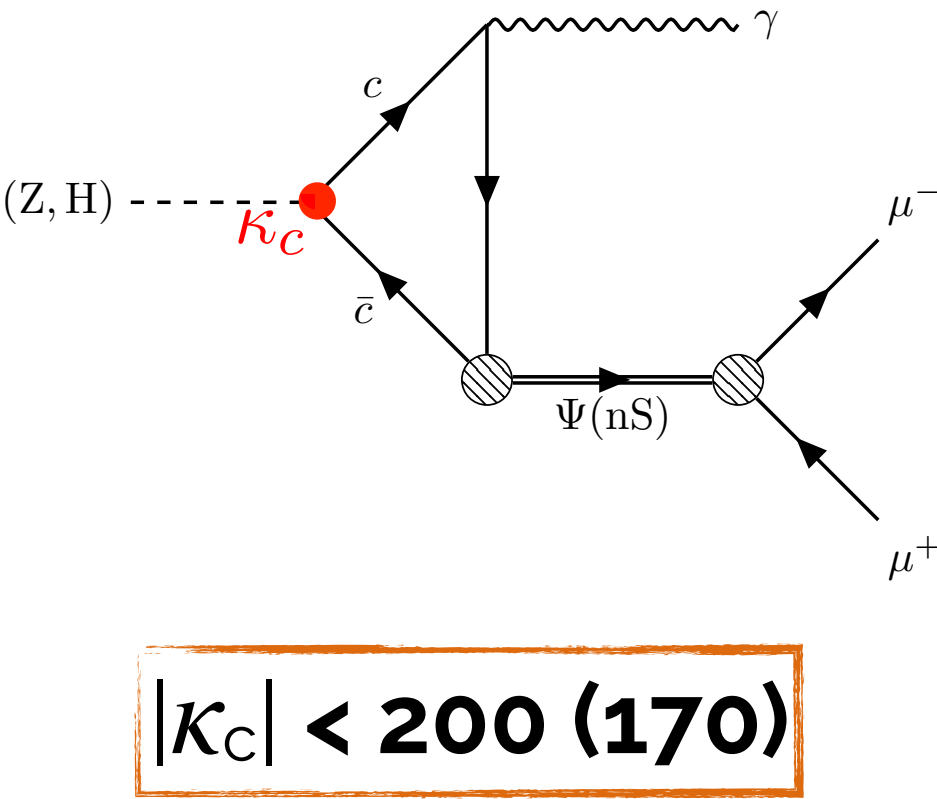
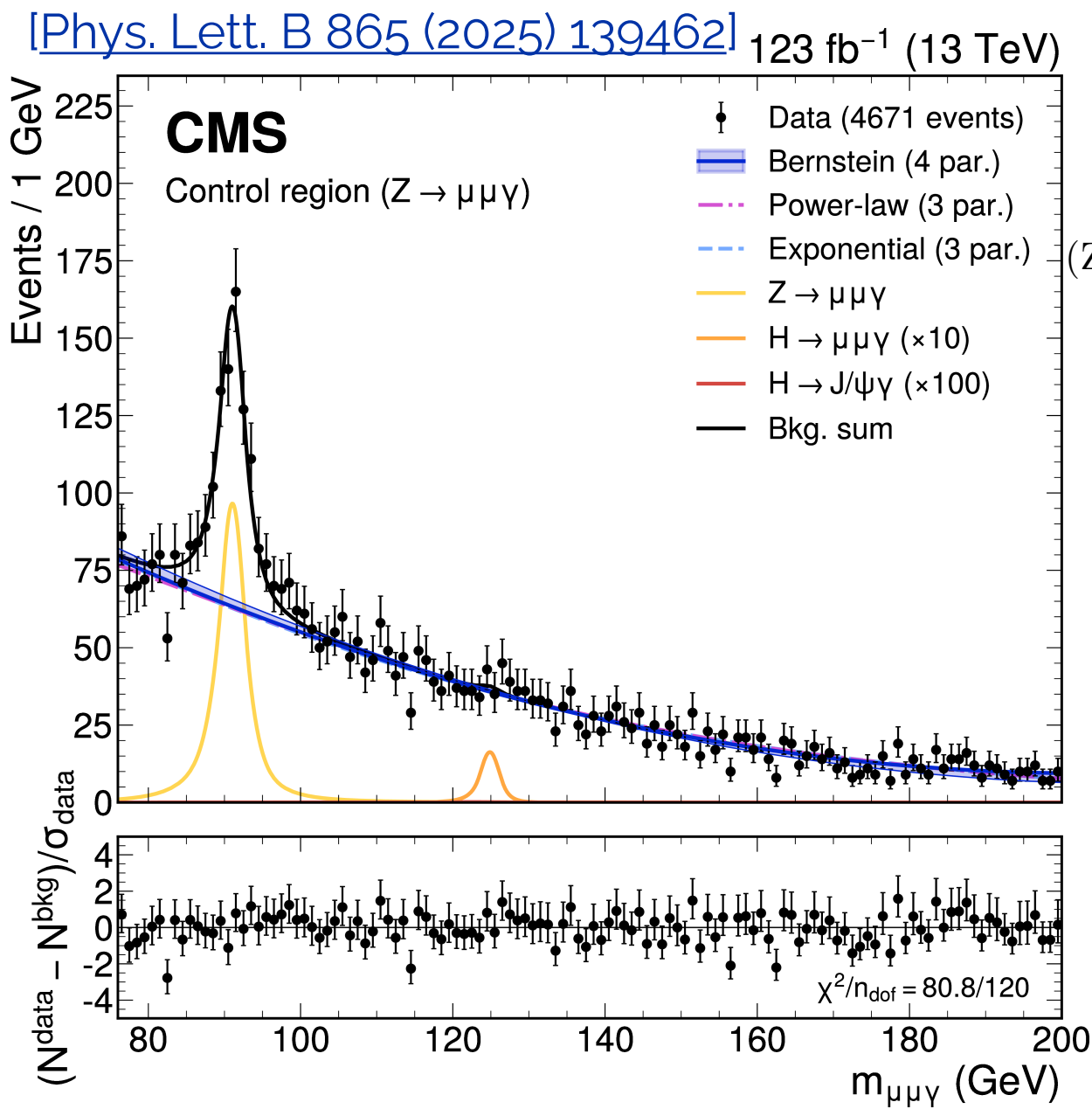
- Several methods explored by ATLAS and CMS to probe the Higgs-charm Yukawa coupling (y_c)

Indirect constraint from Higgs kinematics

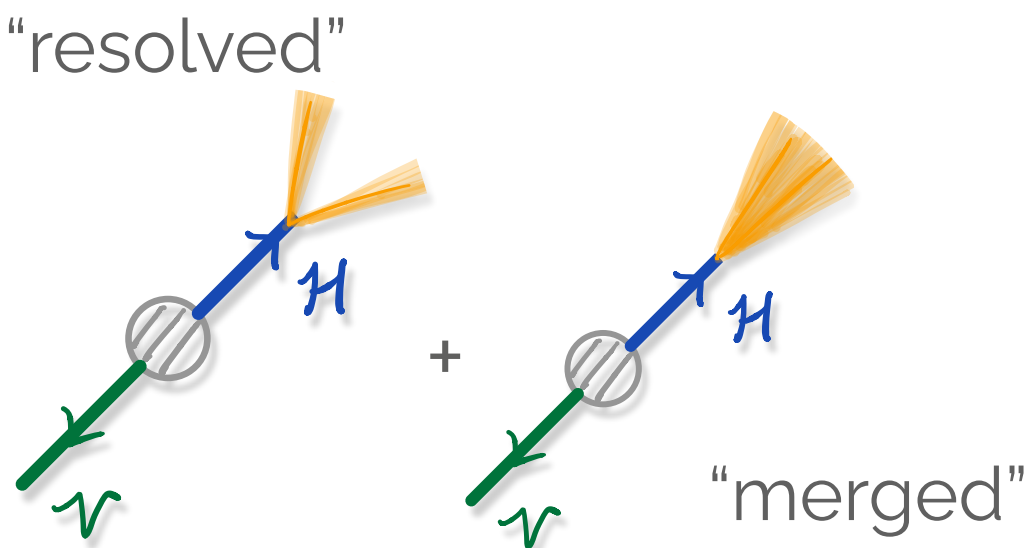
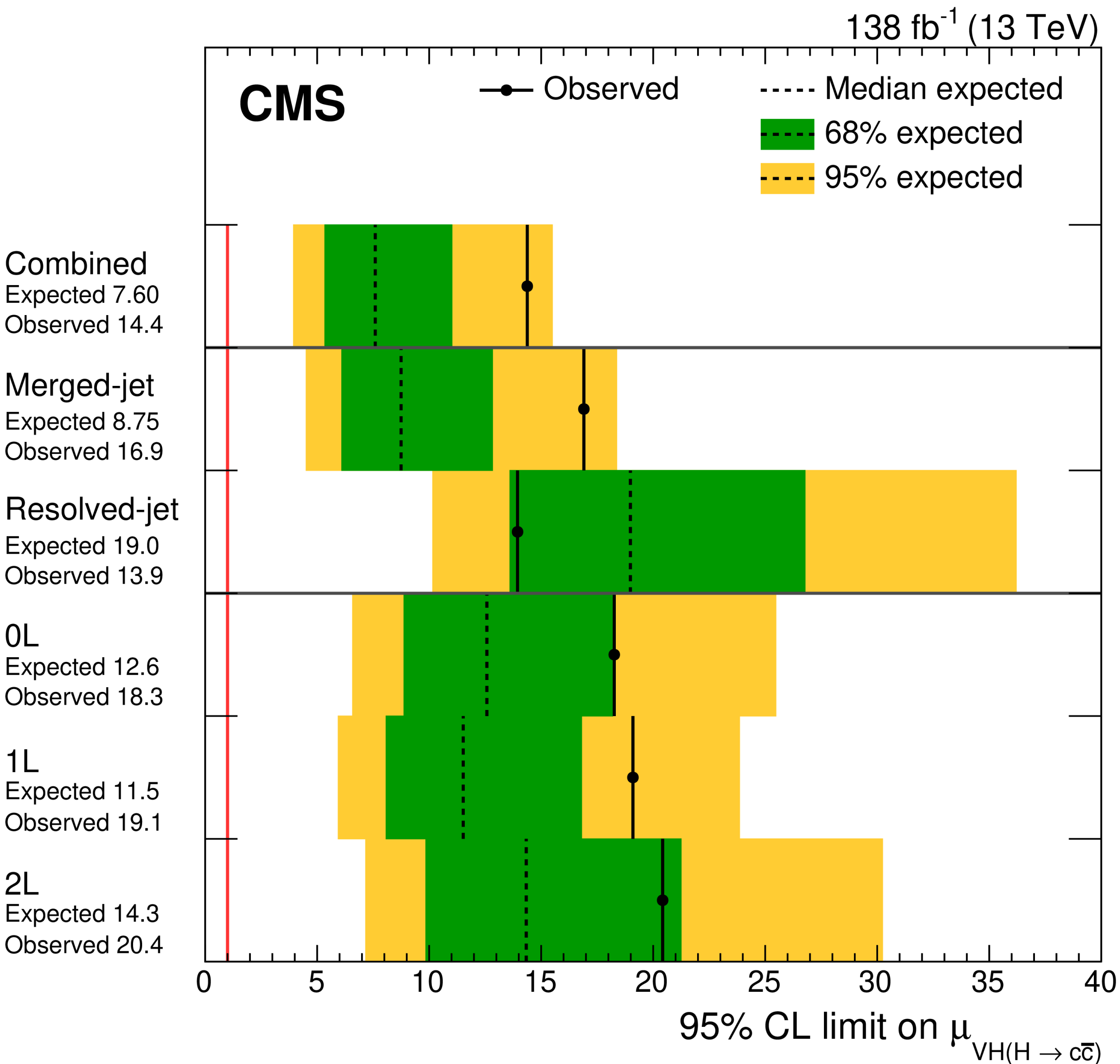
Model dependence?

Direct searches

- Rare Higgs decays
- Associated production (H+c)



So far: probing the Higgs-charm coupling

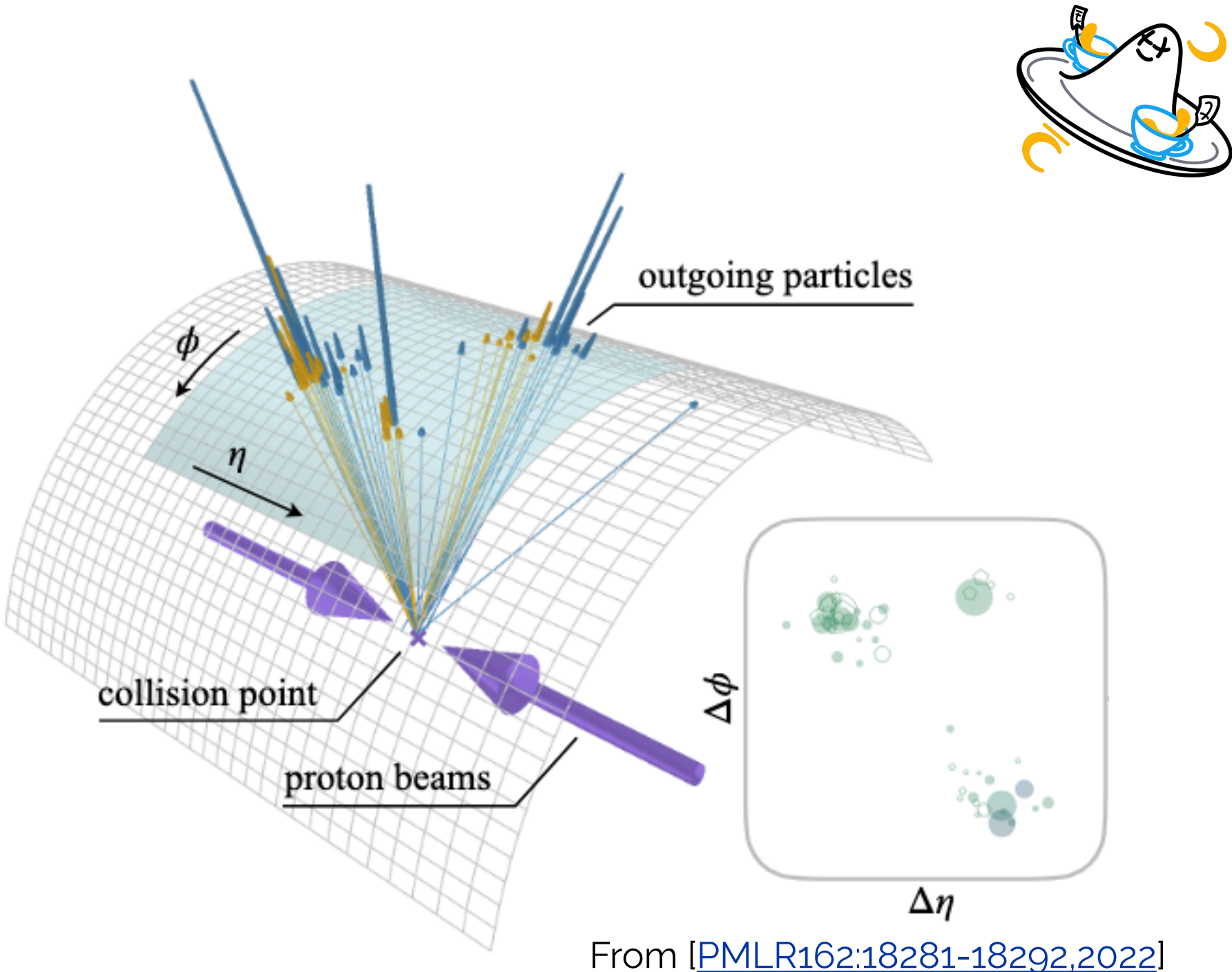


Uncertainty source	$\Delta\mu / (\Delta\mu)_{\text{tot}}$
Statistical	85%
Background normalizations	37%
Experimental	48%
Sizes of the simulated samples	37%
c jet identification efficiencies	23%
Jet energy scale and resolution	15%
Simulation modeling	11%
Integrated luminosity	6%
Lepton identification efficiencies	4%
Theory	22%
Backgrounds	17%
Signal	15%

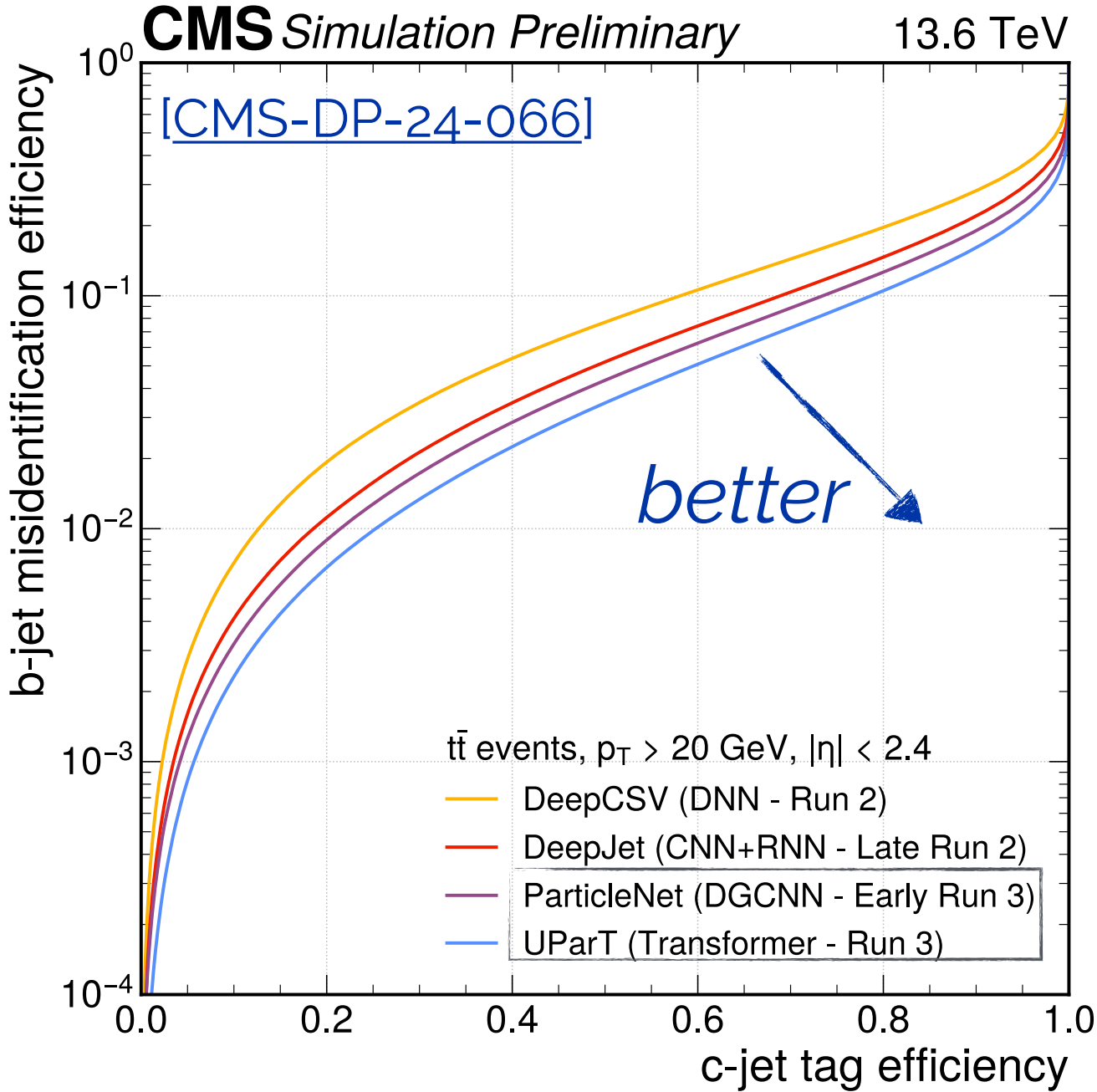
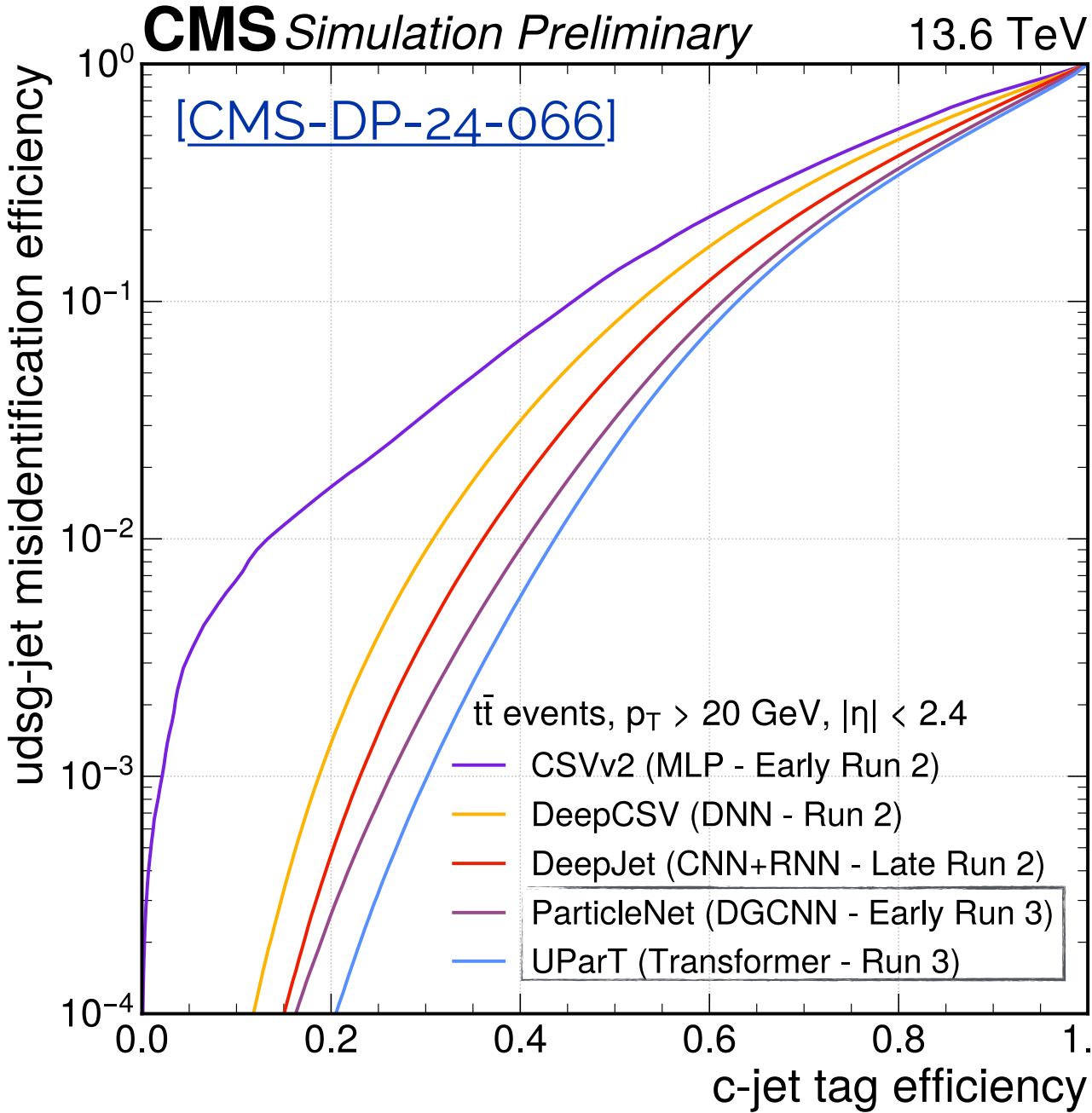
[Phys. Rev. Lett. 131 (2023) 061801]

Jet flavour tagging

- c-jet identification much harder task than b vs. light separation
- We need modern algorithms that:
 - Represent the “physics picture”
 - Can learn particle-particle correlations→ Treat jet as unordered set of inputs (“particle cloud”) in *graph-structure*



Order of
magnitude
improvement!



ParticleNet [2] algorithm originally designed for large-radius tagging successfully used here for c-jet identification

[2] [Phys.Rev. D 101, 056019 (2020)]

Tagging calibration



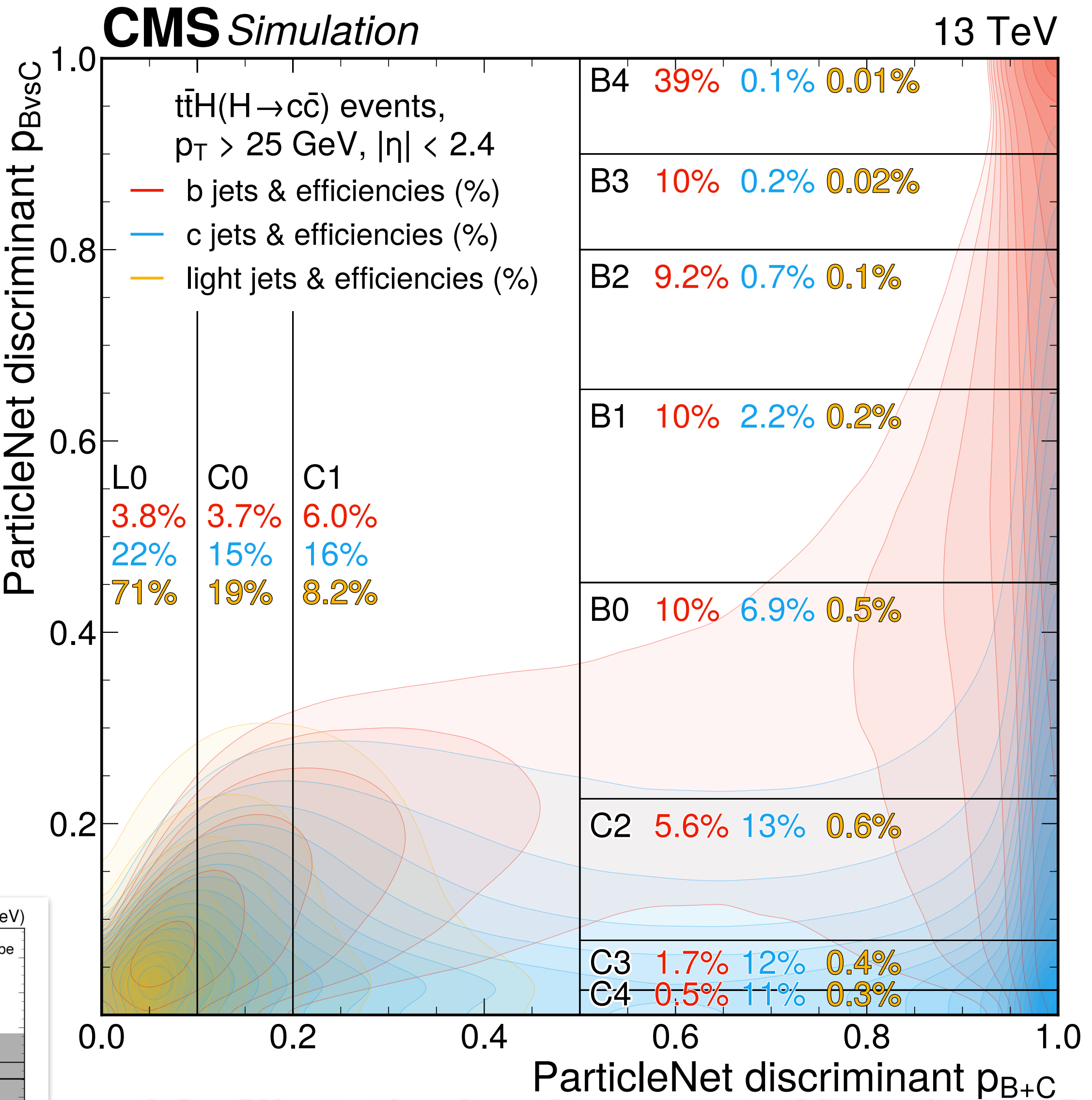
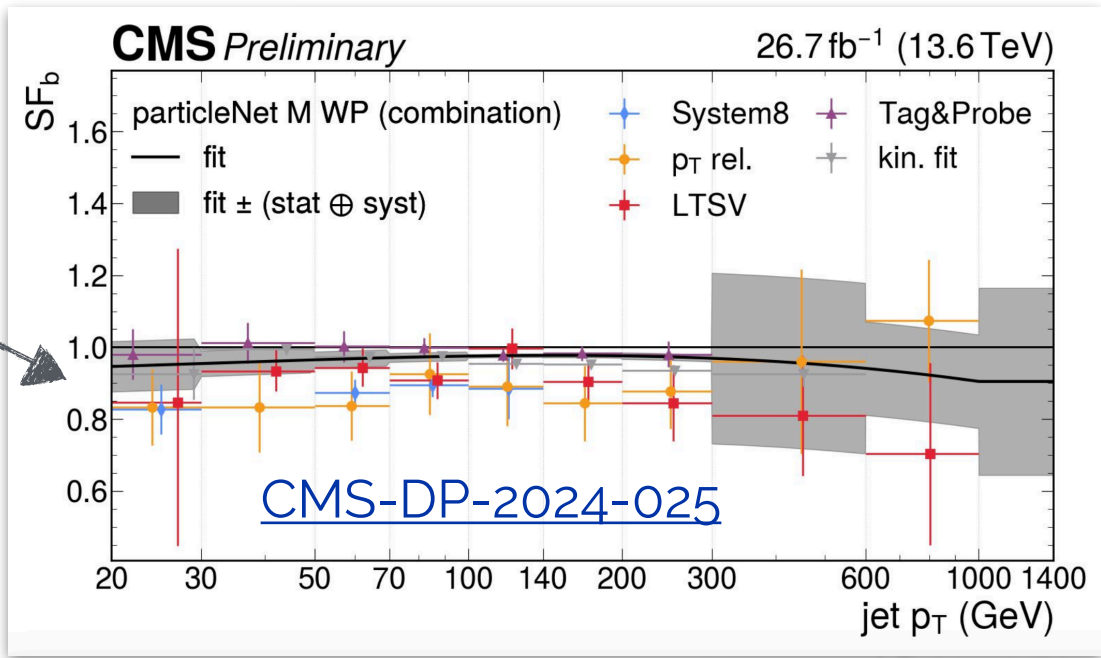
- Leverage established methods:
 - Simultaneous fit to determine all scale factors (SF)

$$\sum_{WP \in [L0, C0-C4, B0-B4]} \epsilon_{MC}(WP|flav) = \sum_{WP} \epsilon_{Data}(WP|flav) \equiv 1$$

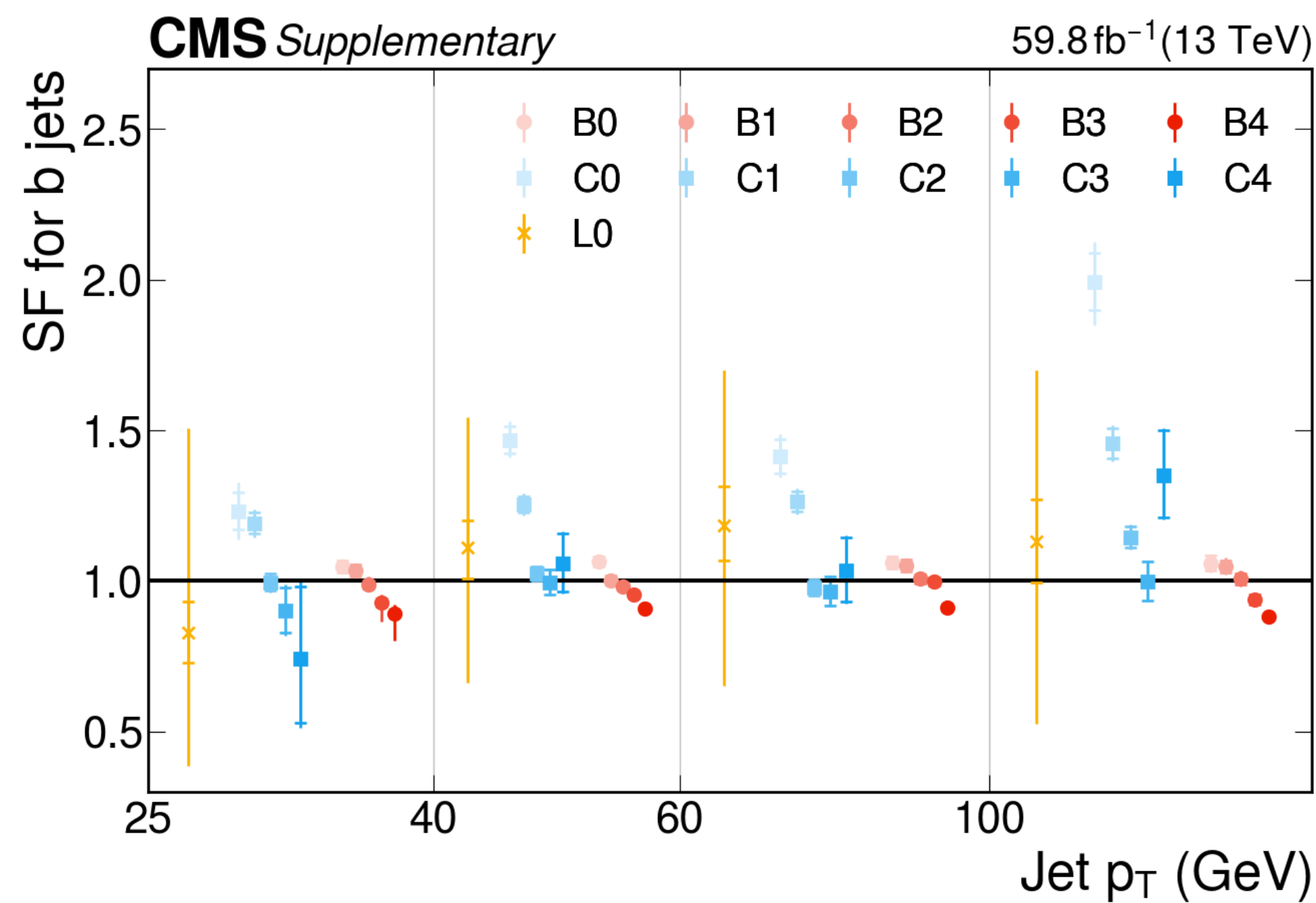
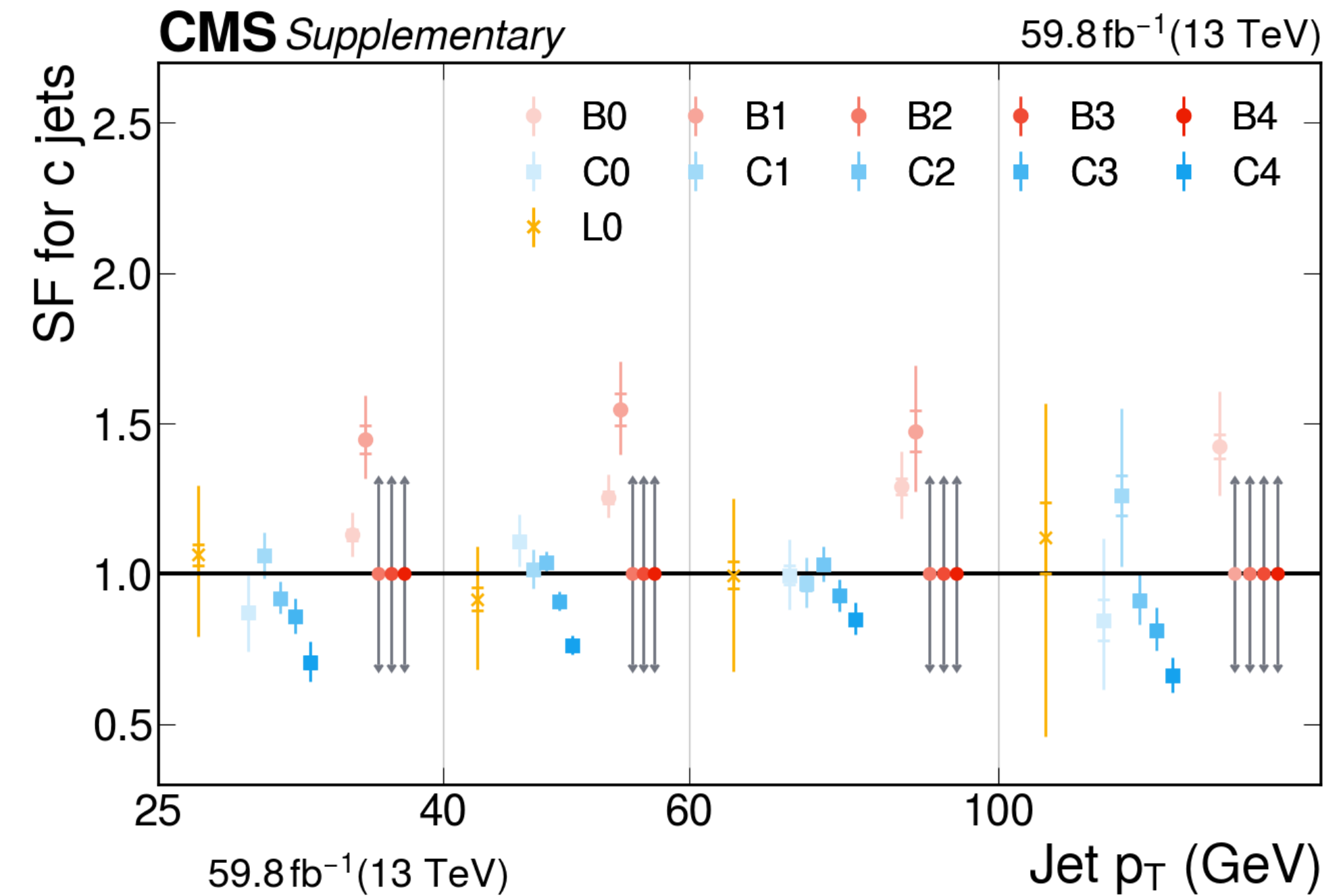
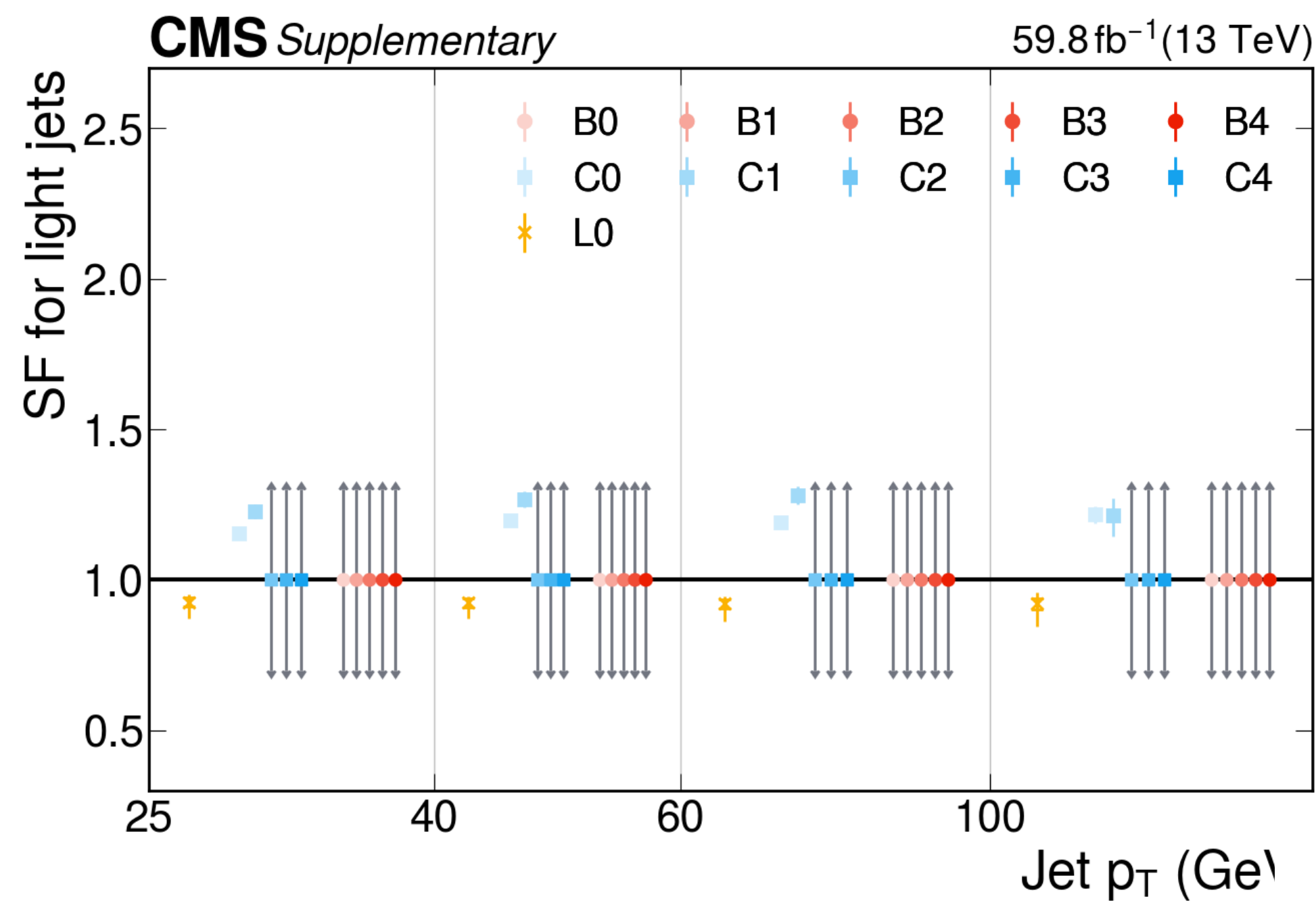
- Fit 3 x 11 - 3 SFs in three selections:
 - W+jets: $l\nu$ + soft muon in jet $\rightarrow$ **c jets**
 - tt: dilepton $e\mu \rightarrow$ **b jets**
 - Z+jets: dilepton $ee/\mu\mu \rightarrow$ **light jets**
- Fit in different bins of jet p_T :

$$p_T(\text{jet}) \in [25, 35, 50, 70, 90, 120, \infty] \text{ GeV}$$

- SFs typically in range of [0.75, 1.5]
 - Also already calibrated in Run 3

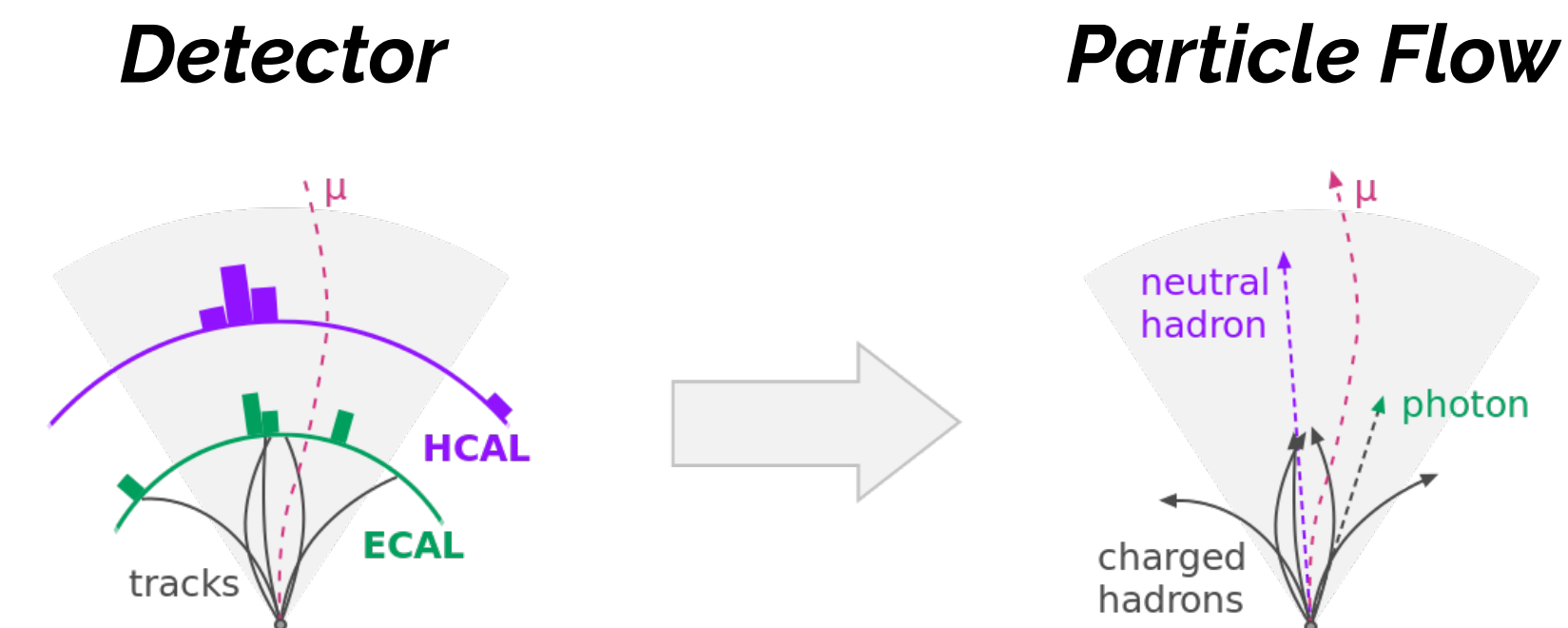


Tagging calibration



The basics: ingredients and CMS strategy

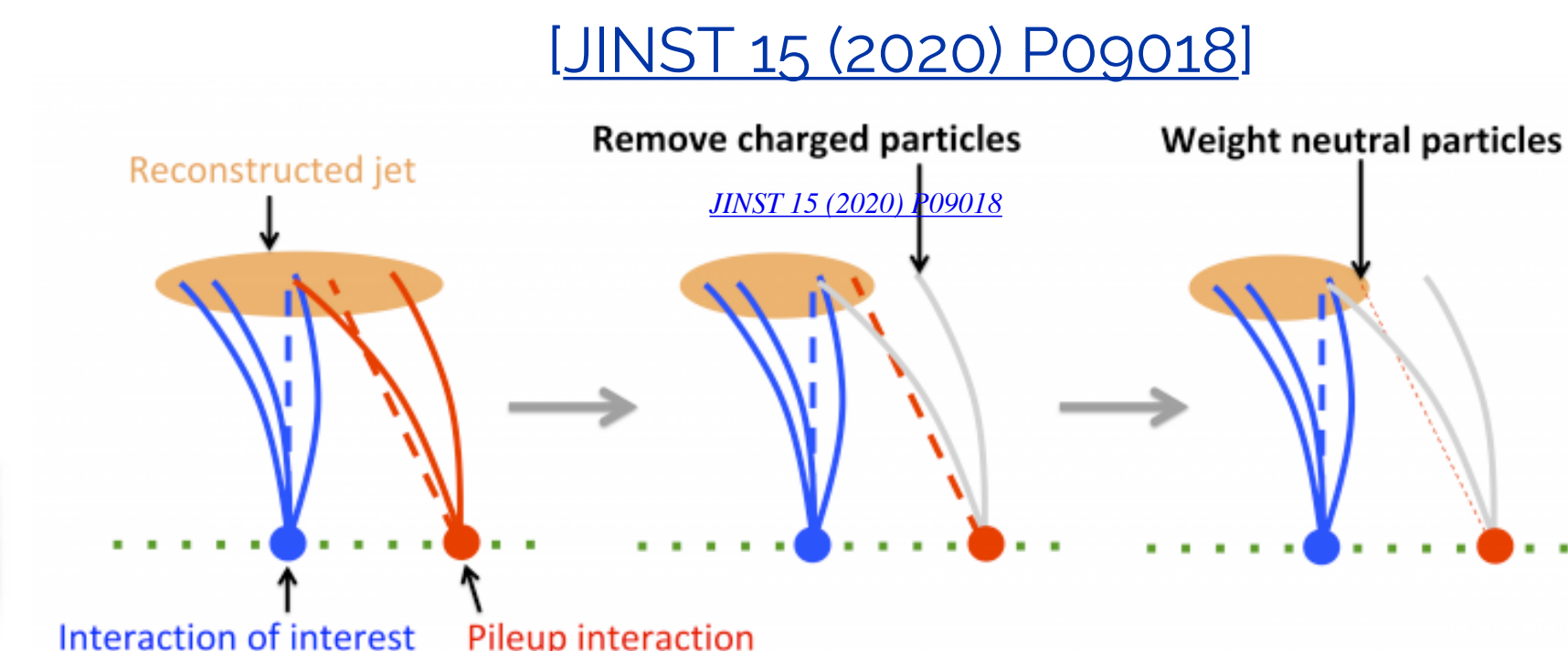
- Particle flow (PF) algorithm:
 - PF provides an improved event description by correlating sub-detector measurements
 - PF allows to reconstruct and identify each single particles
 - Outputs: energy/momentum, position, particle-type, track info and displacement, calorimeter deposit and cluster shape, etc.



[JINST 12 (2017) P10003]

- Pileup mitigation via PileUp Per Particle Identification (PUPPI):
 - Remove particles that are produced by in-time pileup interactions and alter the jet radiation pattern
 - PUPPI assigns to each PF-candidate a probability (α_i) of coming from a pileup interaction
 - Charged particles have weight 1 or 0 depending on their association to PV, neutral particles have a weight assigned by

$$\alpha_i = \log \sum_{j \neq i, \Delta R_{ij} < R_0} \left(\frac{p_{T,j}}{\Delta R_{ij}} \right)^2$$



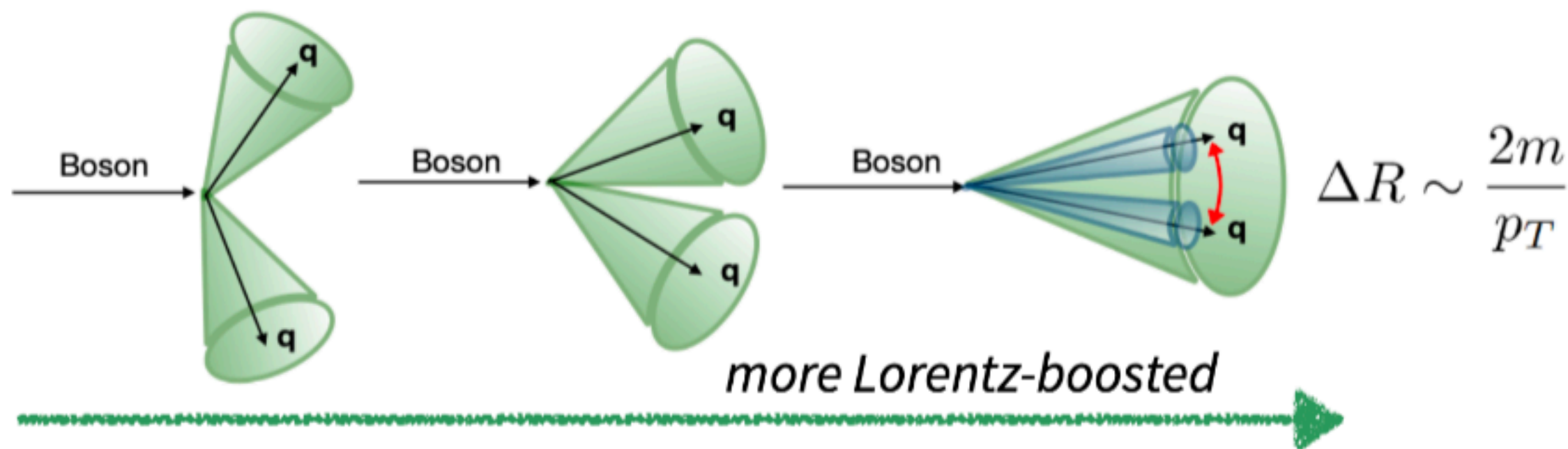
Recent important updates for tau-ID

[CMS-DP-2024-043]

→ Perfect inputs for ML-based jet tasks! (Not only classification 😊)

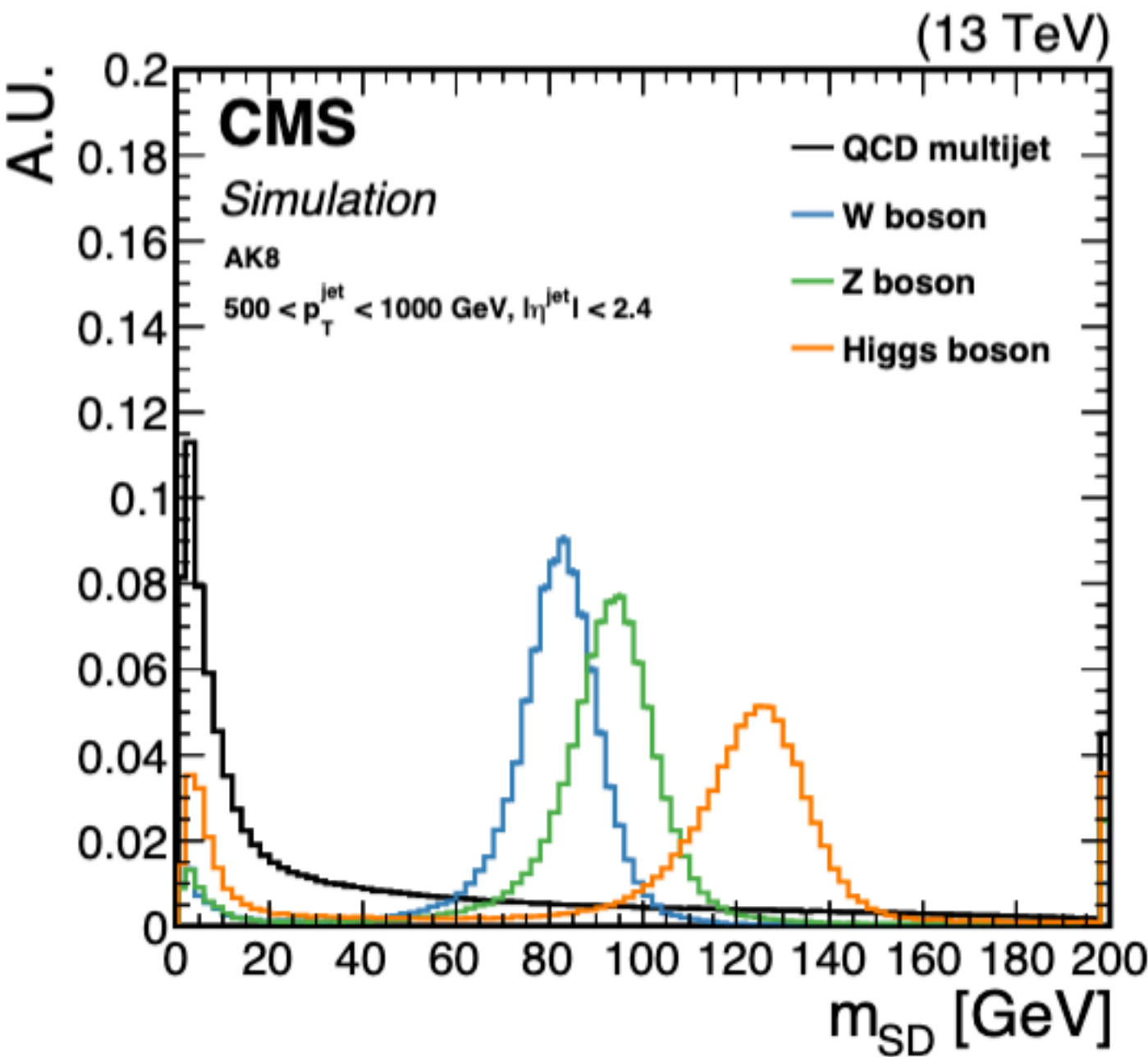
Boosted jet tagging

- Resonance decay with high Lorentz-boost:
 - Decay products clustered in single large-radius jet (AK8 or AK15)

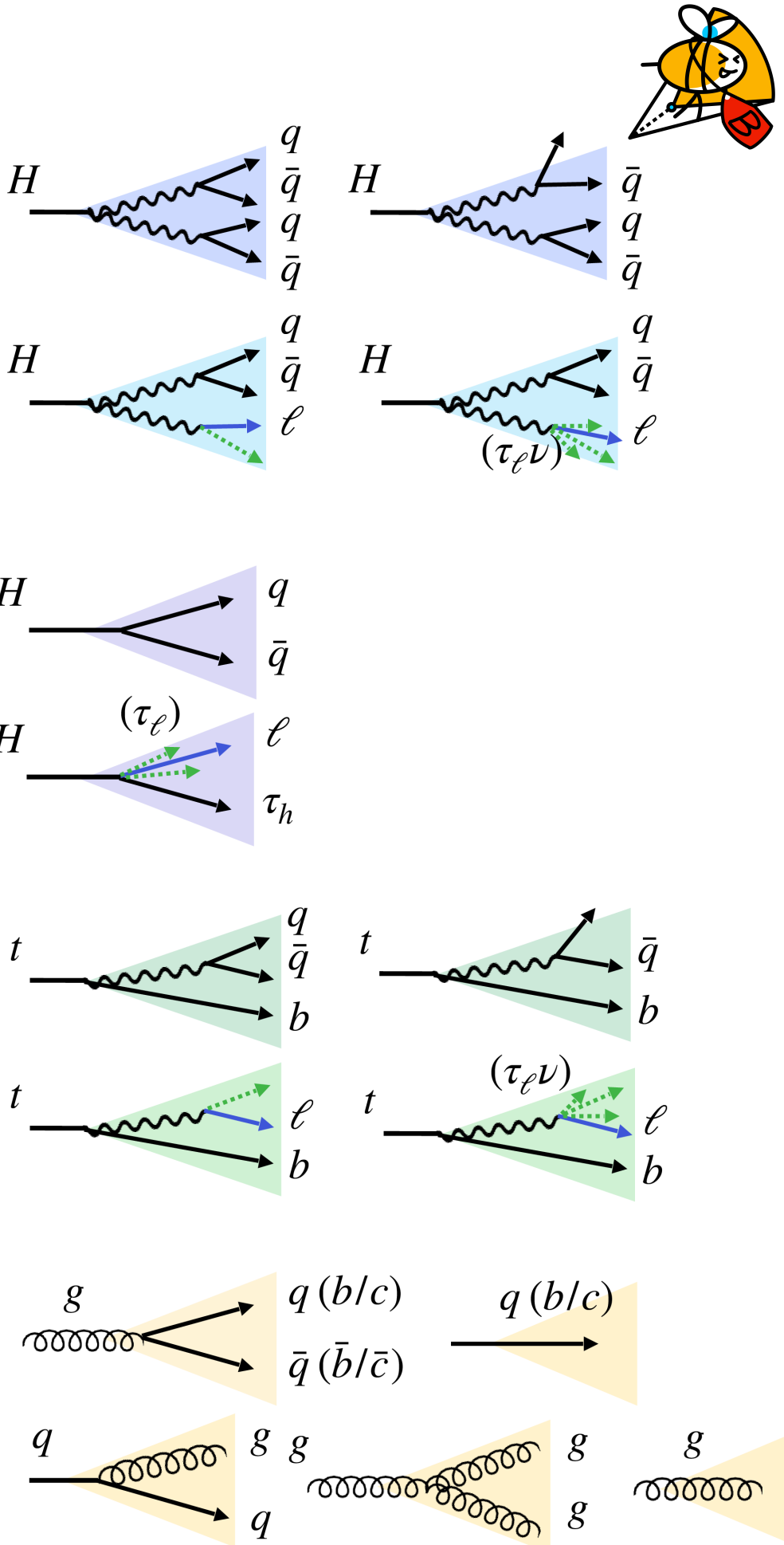


- **ParticleNet** and **GloParT** established as strong baseline for multitude of developments in CMS

But what do we do with the jet mass?



Process	Final state/ prongness	heavy flavour	# of classes
$H \rightarrow VV$ (full-hadronic)	qqqq	0c/1c/2c	3
	qqq		3
$H \rightarrow WW$ (semi-leptonic)	$e\nu qq$	0c/1c	2
	$\mu\nu qq$		2
	$\tau_e\nu qq$		2
	$\tau_\mu\nu qq$		2
	$\tau_h\nu qq$		2
$H \rightarrow qq$		bb	1
		cc	1
		ss	1
		qq (q=u/d)	1
$H \rightarrow \tau\tau$	$\tau_e\tau_h$		1
	$\tau_\mu\tau_h$		1
	$\tau_h\tau_h$		1
$t \rightarrow bW$ (hadronic)	bqq	1b + 0c/1c	2
	bq		2
$t \rightarrow bW$ (leptonic)	$b e \nu$	1b	1
	$b \mu \nu$		1
	$b \tau_e \nu$		1
	$b \tau_\mu \nu$		1
	$b \tau_h \nu$		1
QCD		b	1
		bb	1
		c	1
		cc	1
		others (light)	1

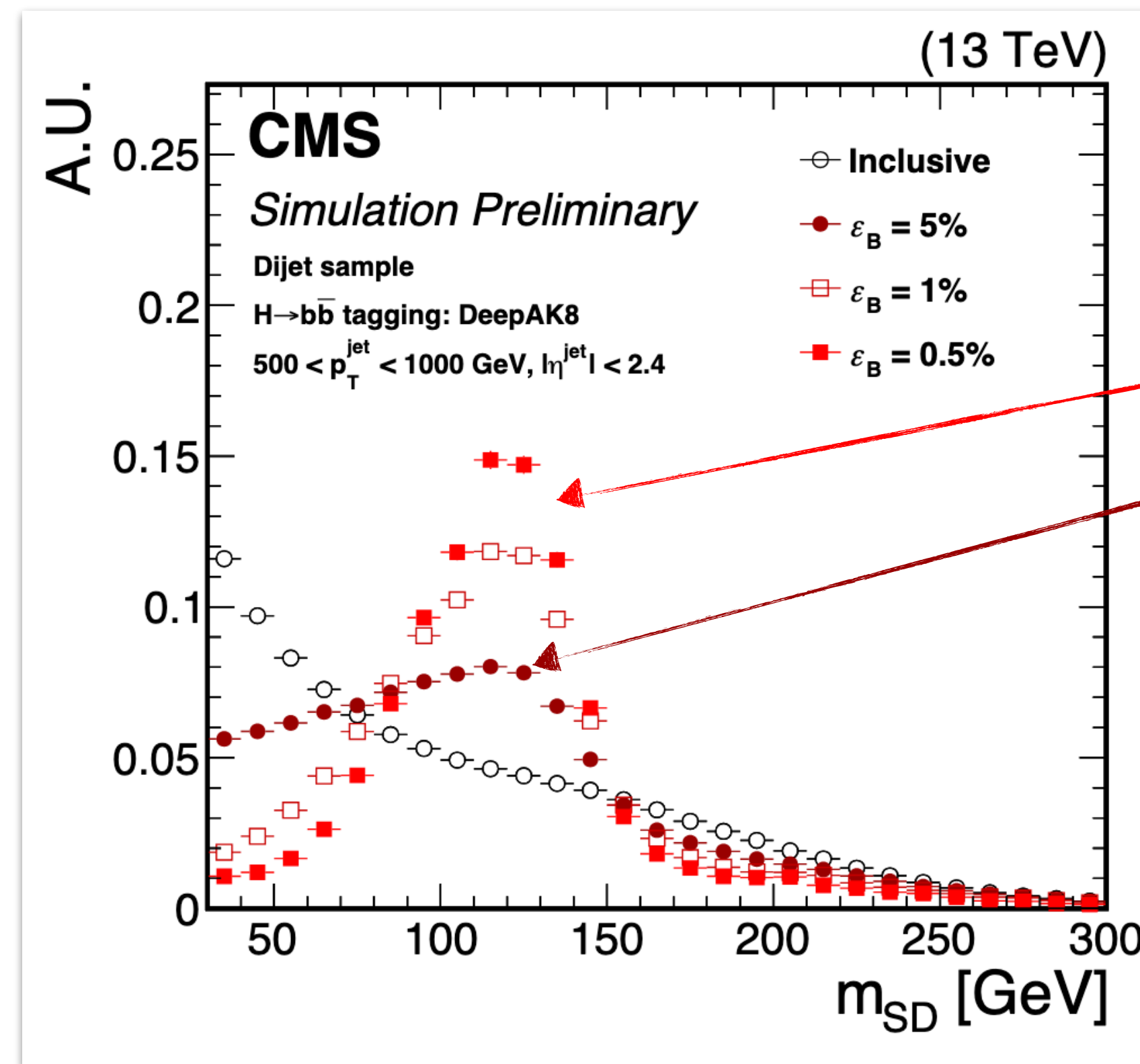


See, e.g., recent $H \rightarrow 4b$ results for Run 3
[\[HIG-24-010\]](#)

Mass dependence

- In several searches with boosted $X(\rightarrow YY)$ resonances → **signal is extracted from a fit to the jet mass**
- If the $X(\rightarrow YY)$ tagger is correlated with the jet mass → jet mass in background events becomes signal-like (**sculpting**) → extracting signal becomes more difficult

*N.B.: It can also be desired to “learn” the mass and use it
→ both versions of the tagger exist!*

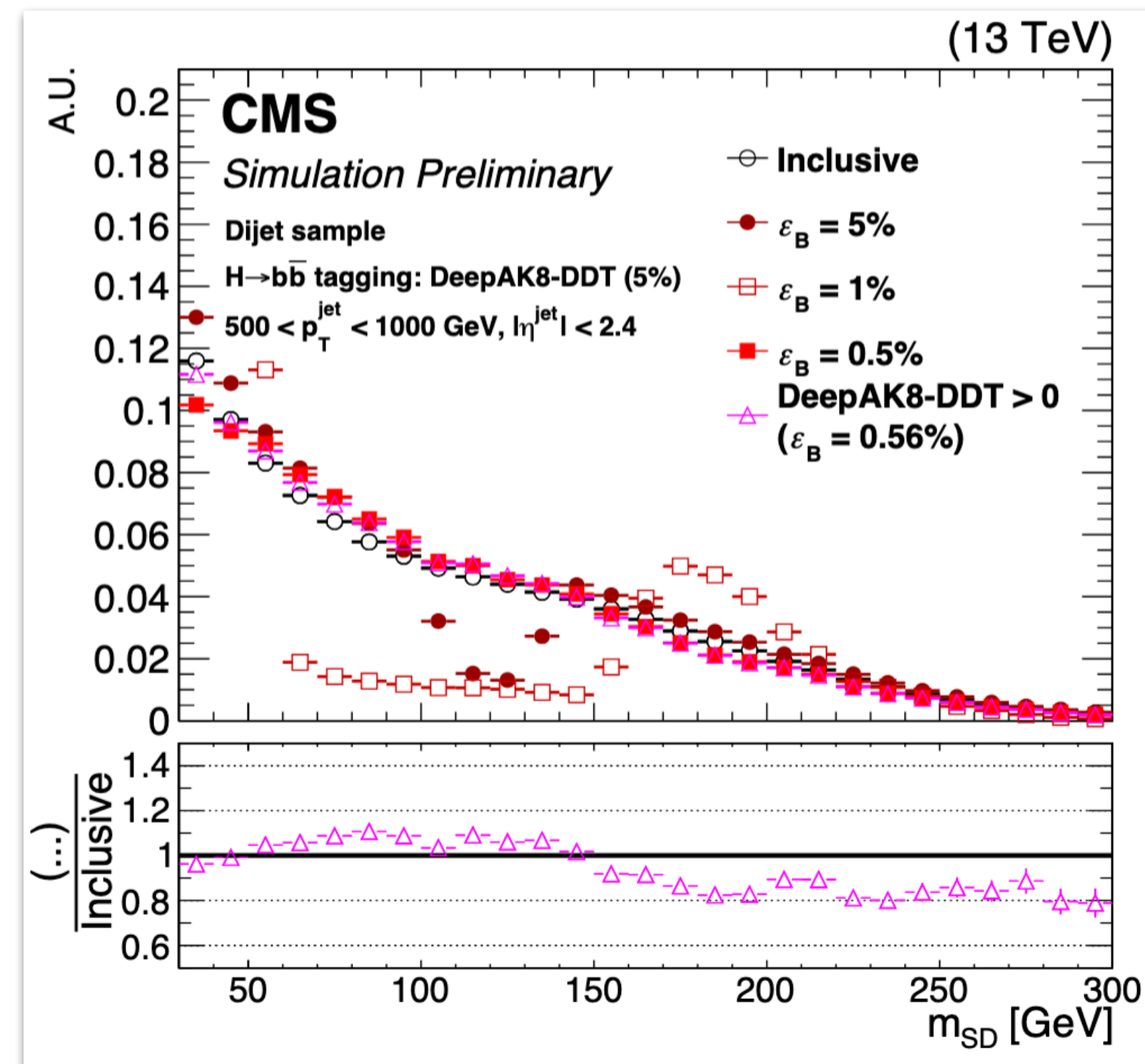


Mass decorrelation techniques

- In several searches with boosted $X(\rightarrow YY)$ resonances → **signal is extracted from a fit to the jet mass**
- If the $X(\rightarrow YY)$ tagger is correlated with the jet mass → jet mass in background events becomes signal-like (**sculpting**) → extracting signal becomes more difficult

A-posteriori de-correlation (DDT)

- Define a metric that is sensitive to jet-mass through correlations: $\ln(m_{SD}^2/p_T^2)$ and p_T
- Transform tagger score to preserve constant background rejection vs. m_{jet} for a given efficiency



Mass decorrelation breaks down as soon as the selection is changed

Mass decorrelation techniques

- In several searches with boosted $X(\rightarrow YY)$ resonances → **signal is extracted from a fit to the jet mass**
- If the $X(\rightarrow YY)$ tagger is correlated with the jet mass → jet mass in background events becomes signal-like (**sculpting**) → extracting signal becomes more difficult

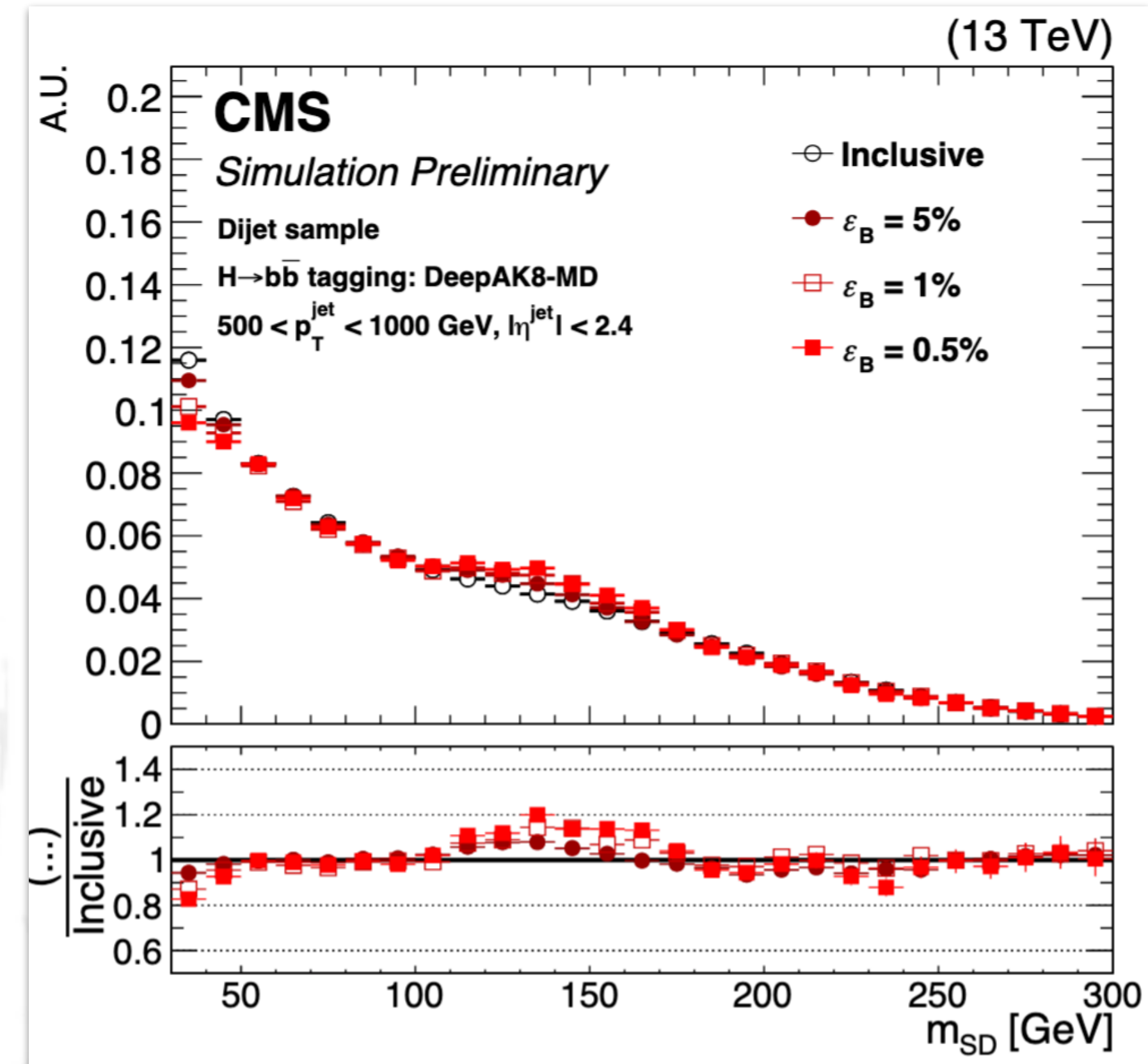
A-posteriori de-correlation (DDT)

- Define a metric that is sensitive to jet-mass through correlations:
 $\ln(m_{SD}^2/p_T^2)$
- Transform tagger score to preserve constant background rejection vs. m_{jet} for a given efficiency

Adversarial networks

- Add a NN that predicts m_{jet} from the feature embeddings of the nominal network
- Add penalty term in the loss function related to the prediction of the correct m_{jet}

Painful to train, significant loss in performance and not perfect mass de-correlation

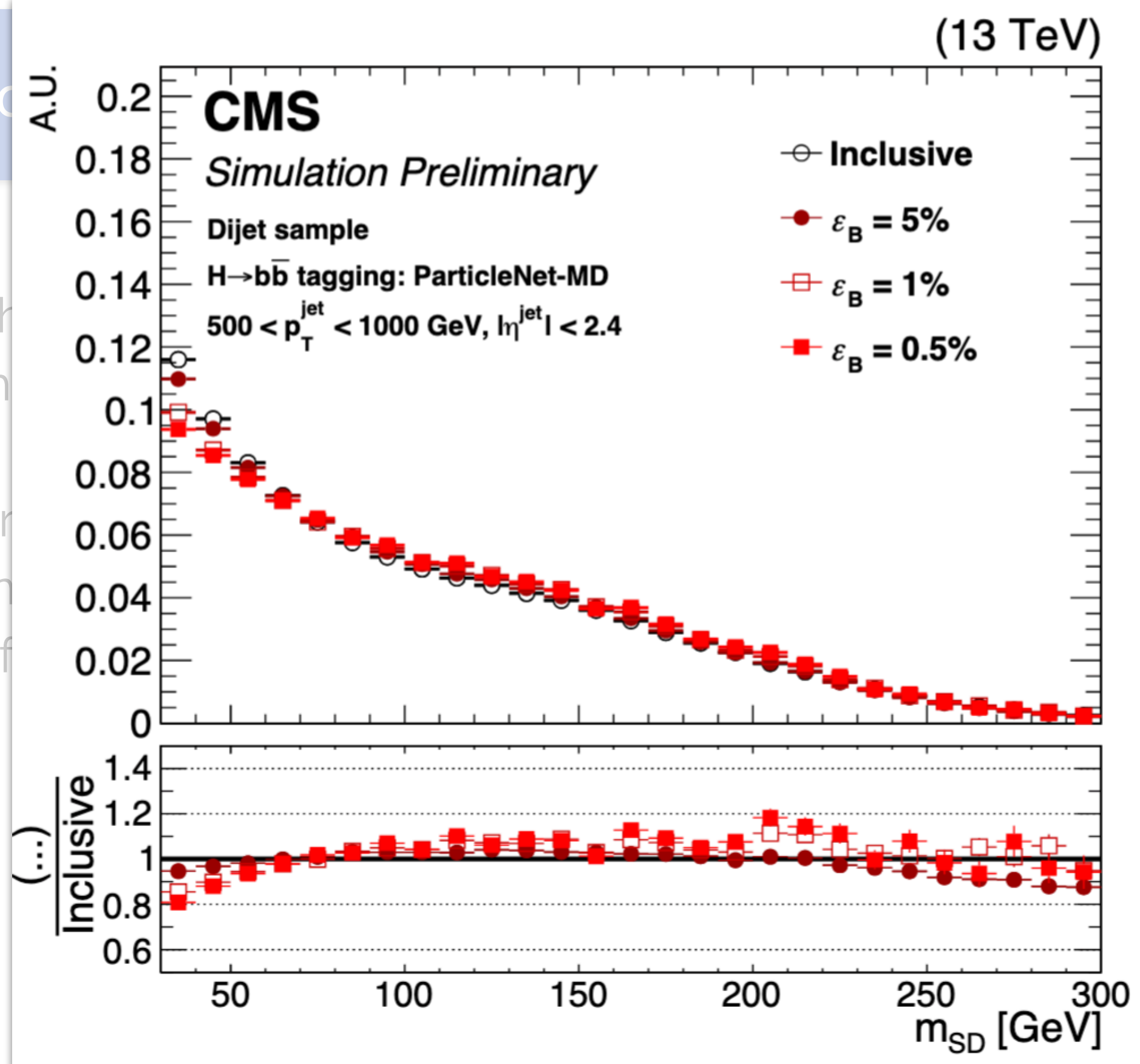


Mass decorrelation techniques

- In several searches with boosted $X(\rightarrow YY)$ resonances → **signal is extracted from a fit to the jet mass**
- If the $X(\rightarrow YY)$ tagger is correlated with the jet mass → jet mass in background events becomes signal-like (**sculpting**) → extracting signal becomes more difficult

A-posteriori de-correlation

- Define a metric that decorrelates jet-mass through $\ln(m_{SD}^2/p_T^2)$
- Transform tagger to preserve constant rejection vs. m_{jet} for efficiency



works

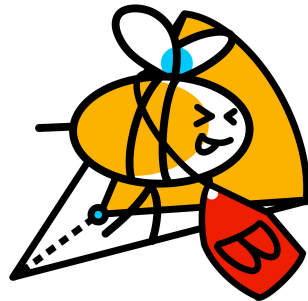
ts m_{jet} from
ngs of the
he loss
e prediction

Mass uniform sampling in training

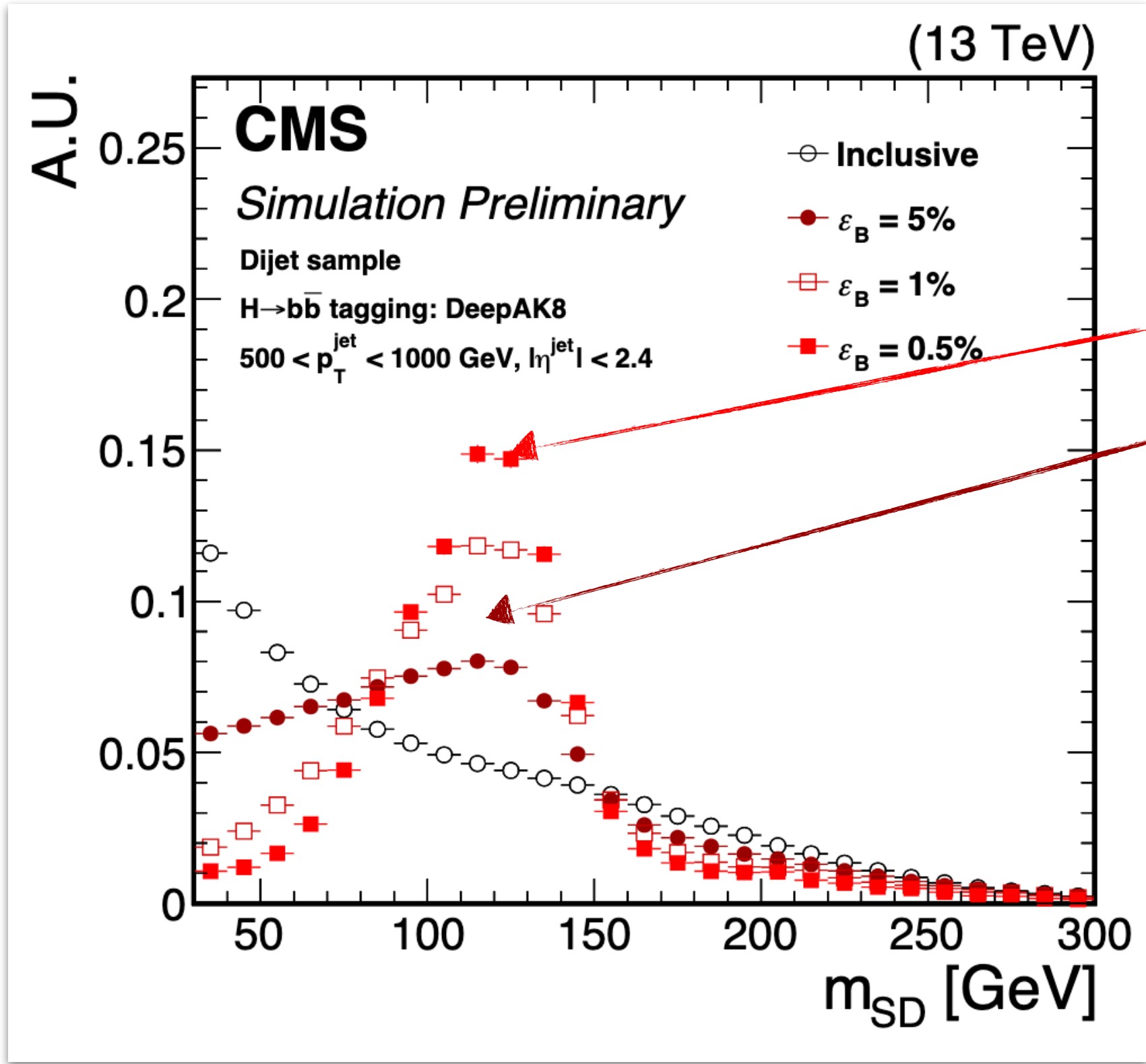
- Dedicated signal samples with flat resonance mass (m_X)
- All jet classes included in the training reweighted to a flat jet mass (m_{SD}) and p_T

**Best method developed so far:
minimal loss in performance and
good mass decorrelation**

Mass decorrelation techniques



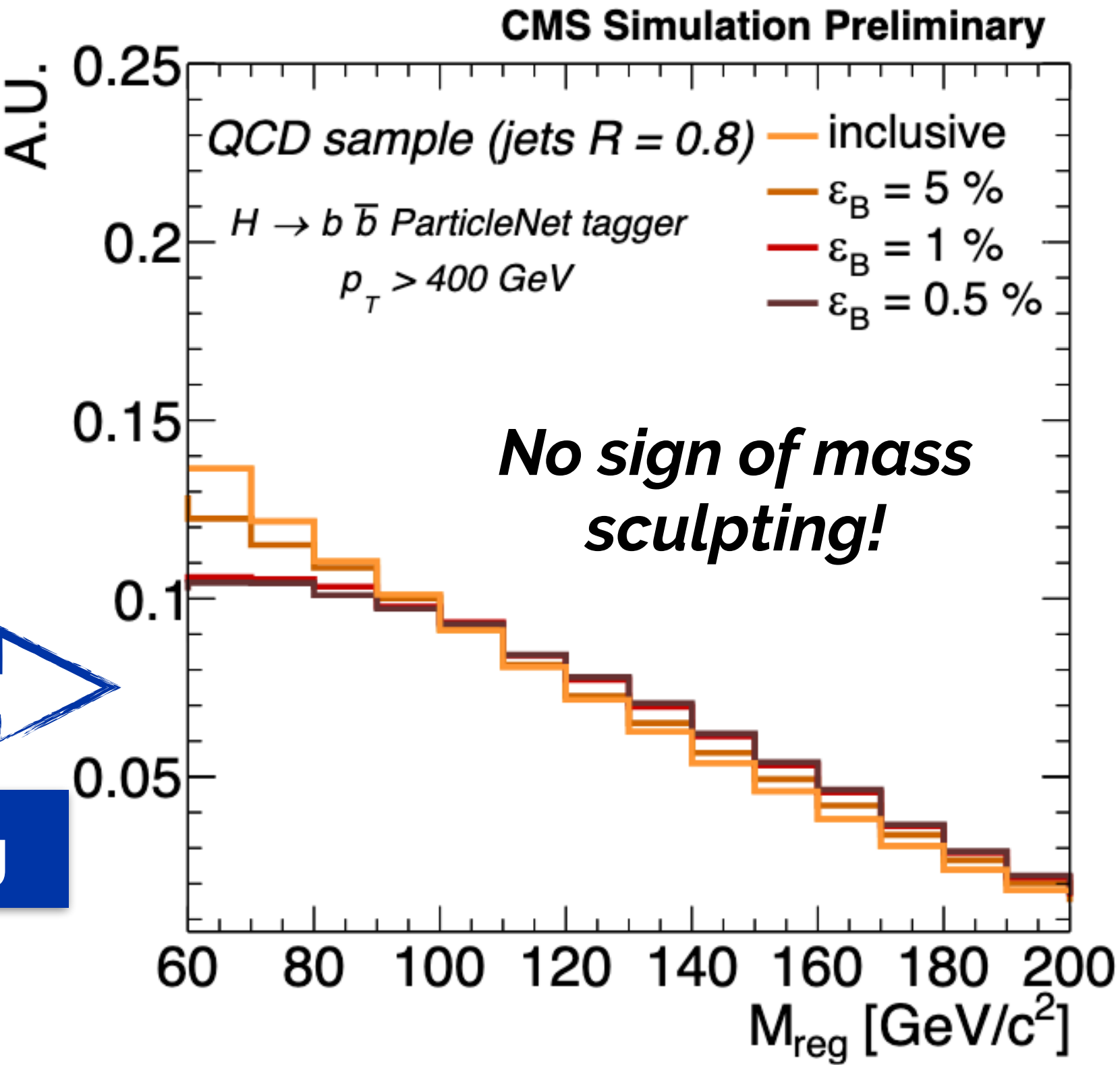
- In several searches with boosted $X(\rightarrow YY)$ resonances → **signal is extracted from a fit to the jet mass**
- If the $X(\rightarrow YY)$ tagger is correlated with the jet mass → jet mass in background events becomes signal-like (**sculpting**) → extracting signal becomes more difficult



Gets worse with tighter selection!

Mass uniform sampling in training

- Dedicated signal samples with flat mass (m_X)
- All jet classes included in the training reweighted to a flat jet mass (m_{SD}) and p_T

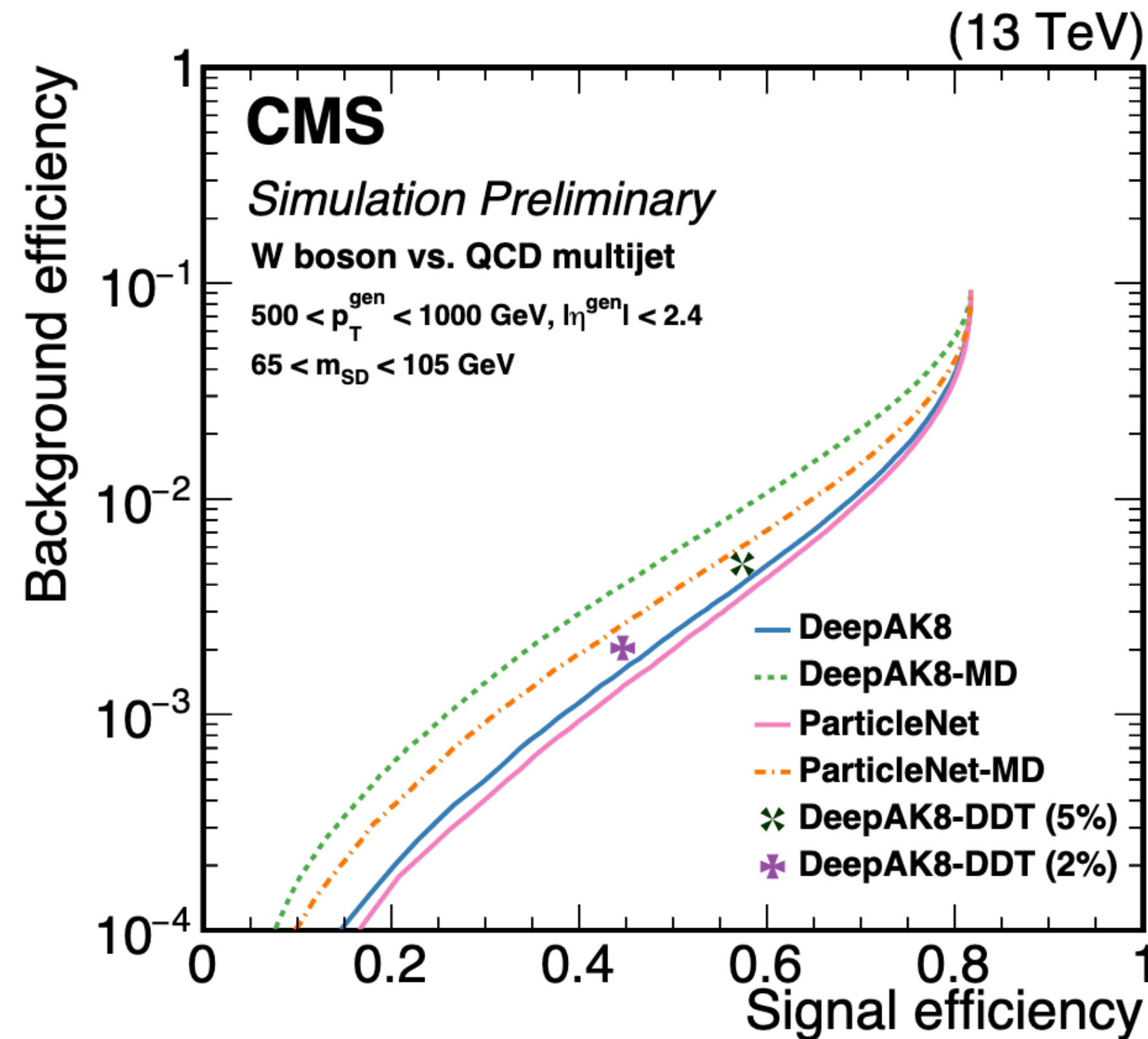
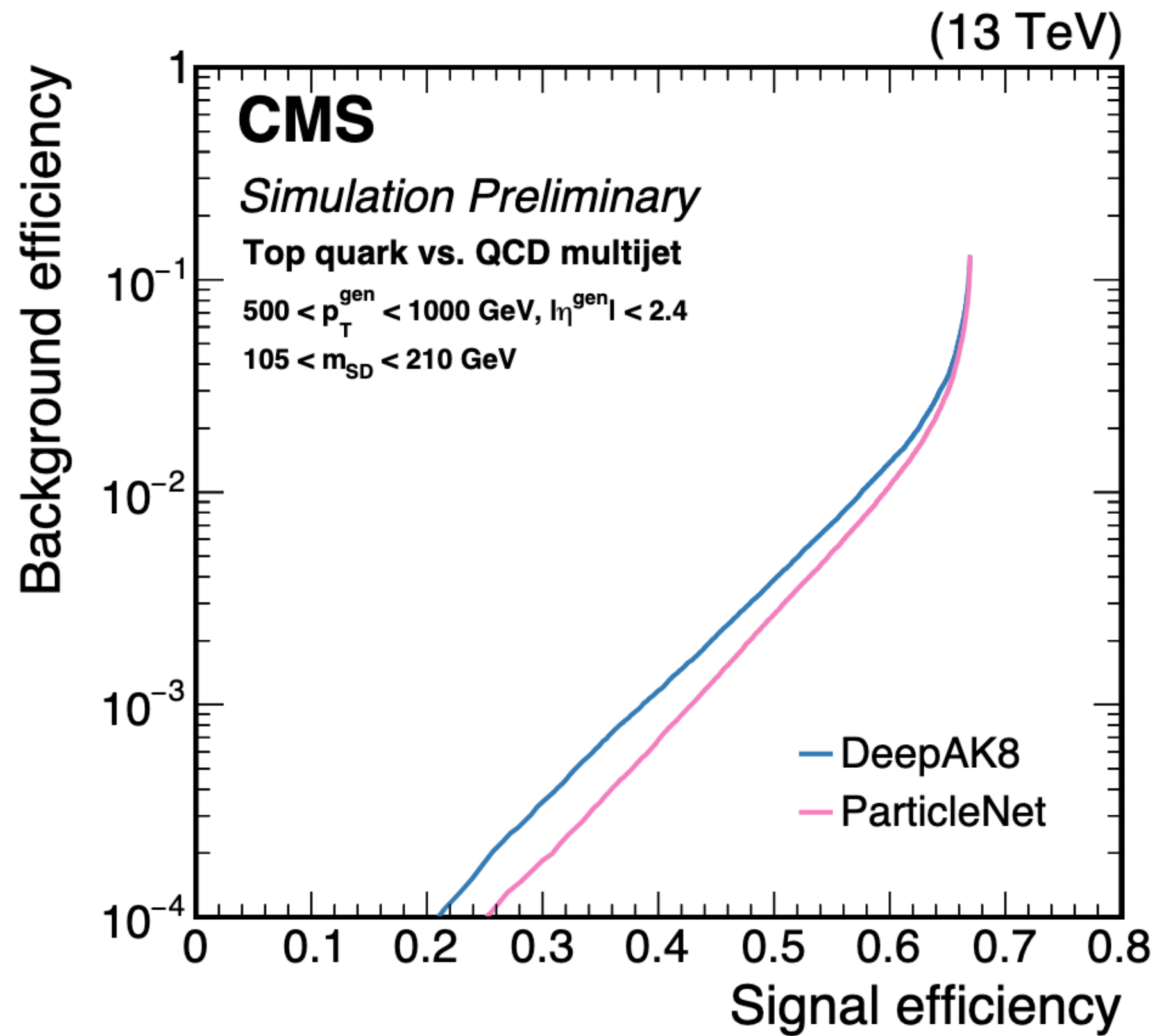


Minimal loss in performance and good mass decorrelation

N.B.: It can also be desired to “learn” the mass → both versions of the tagger exist!

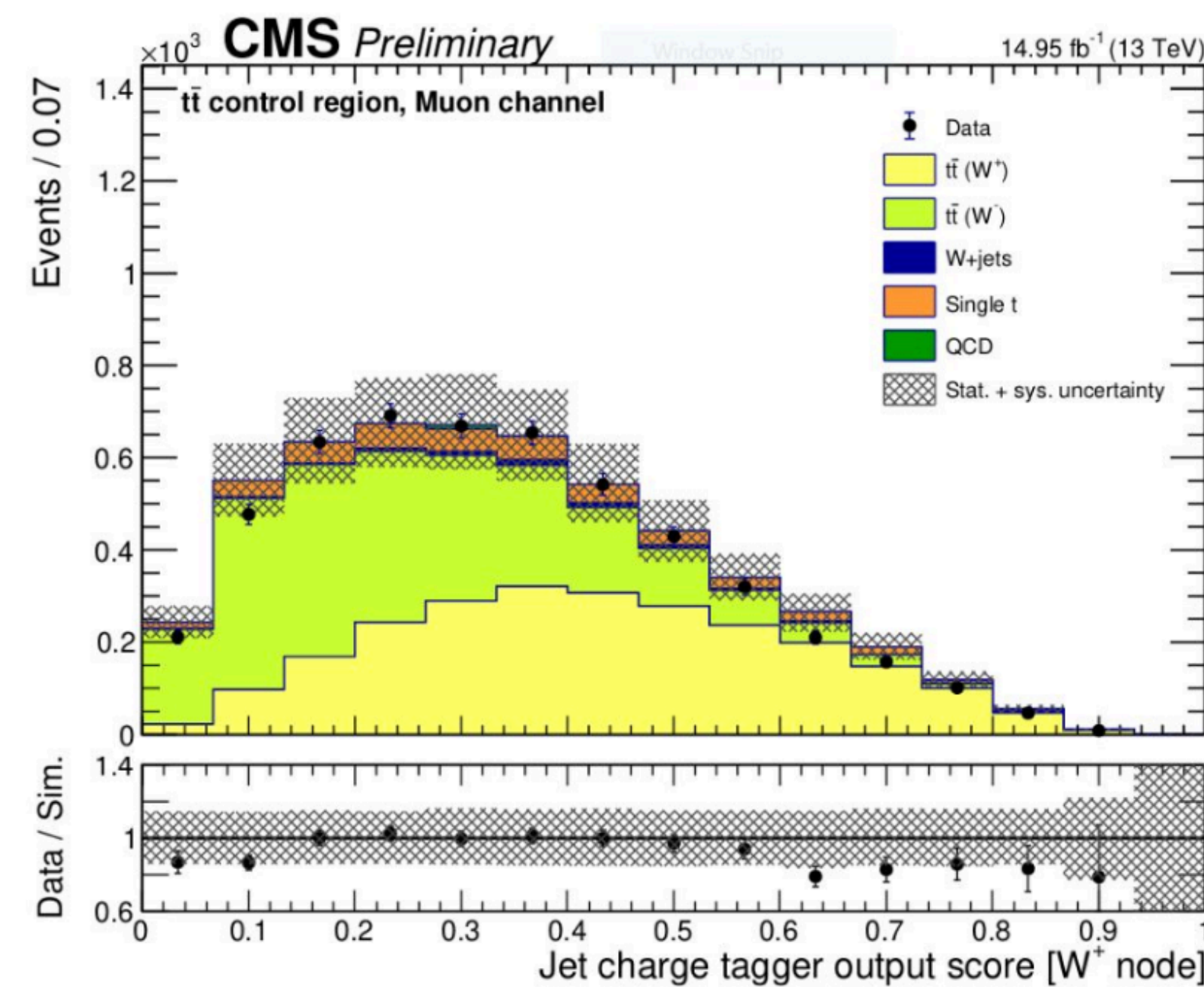
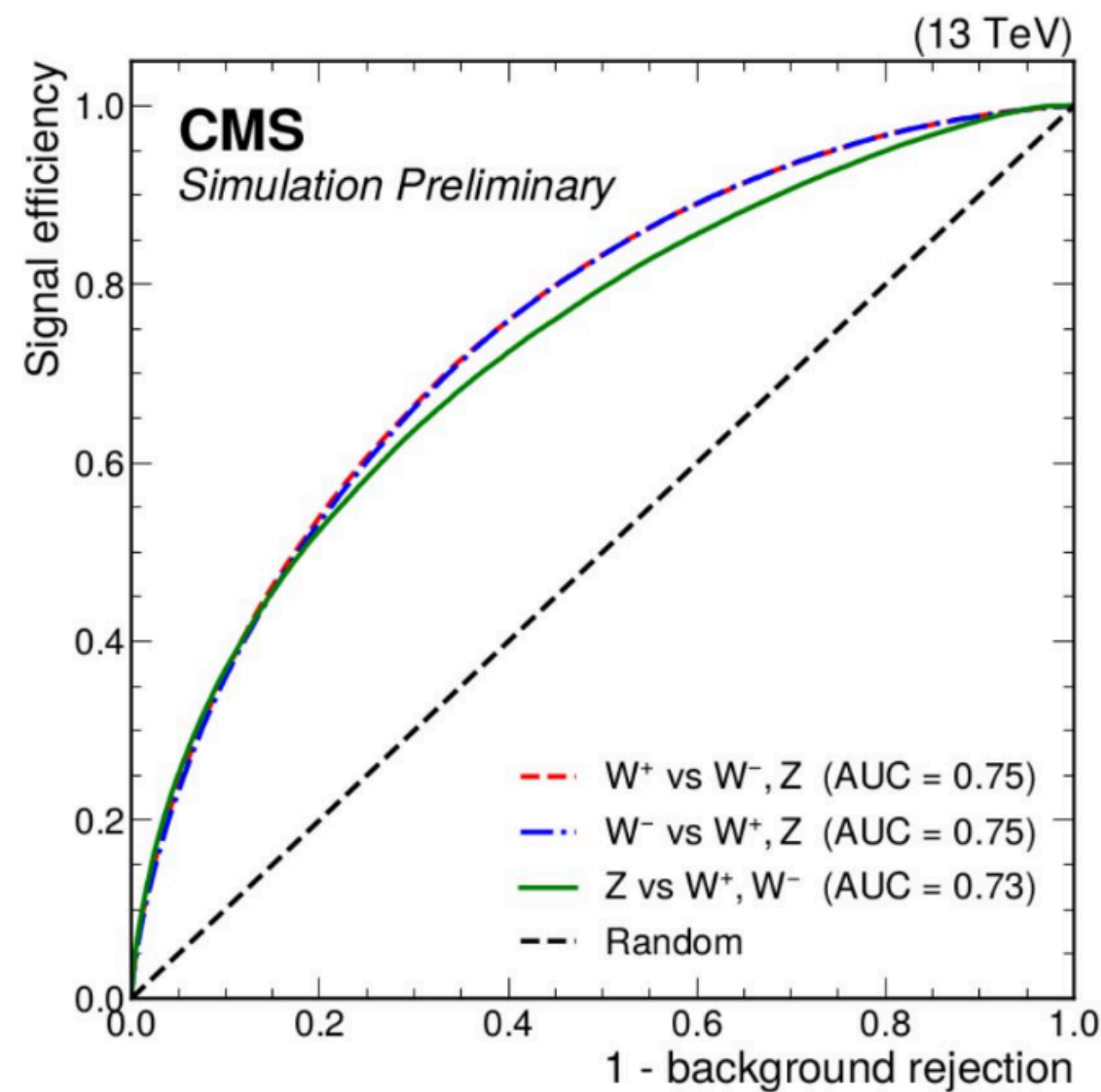
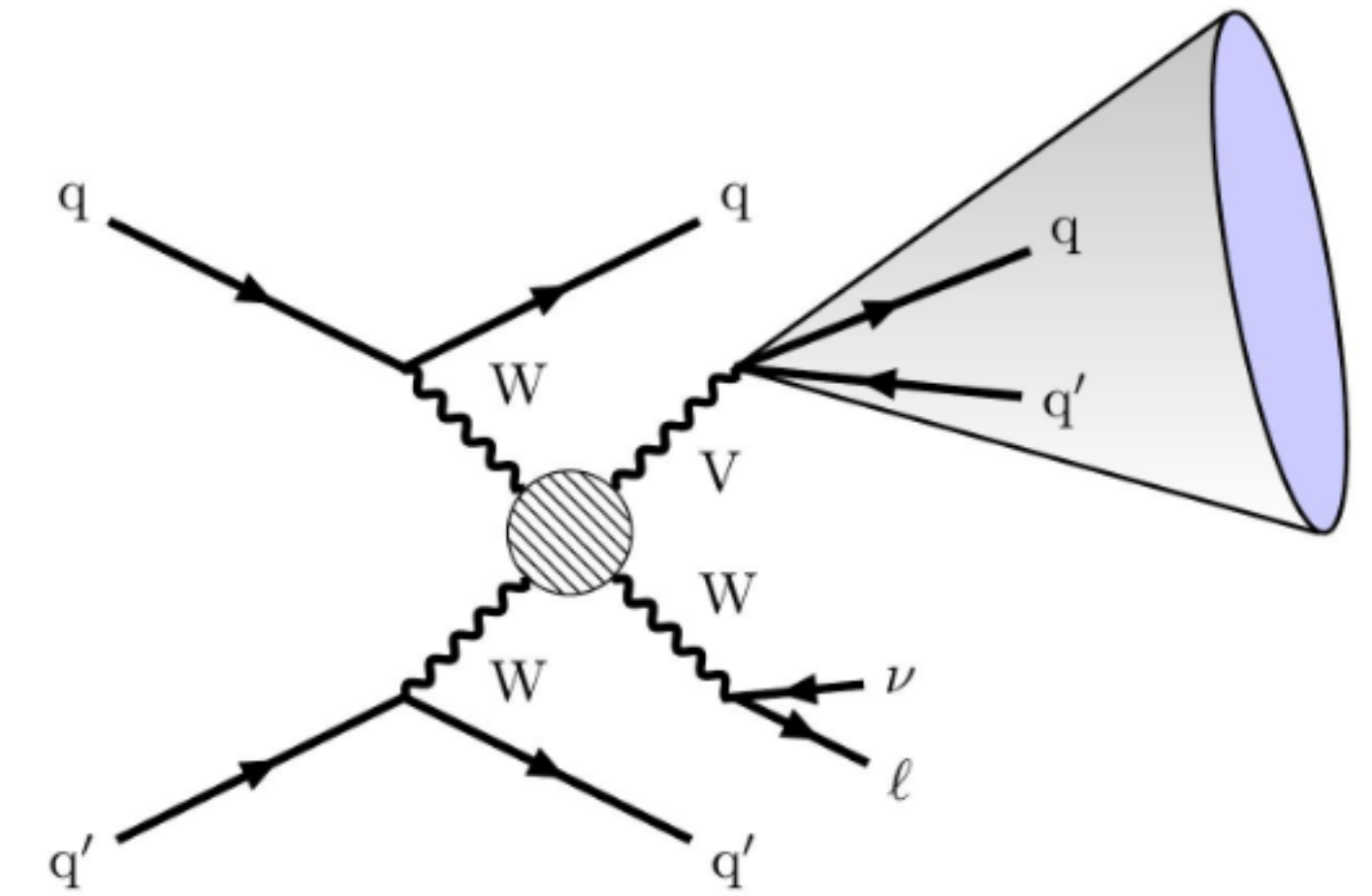
What about t and W?

- Also here, ParticleNet is the state-of-the-art
 - Improvement with and without mass-decorrelation



Extensions: Charge tagging

- When studying processes like same-sign VBS, signal can be isolated using jet charge
- Train ParticleNet to separate W^+ , W^- , Z^0



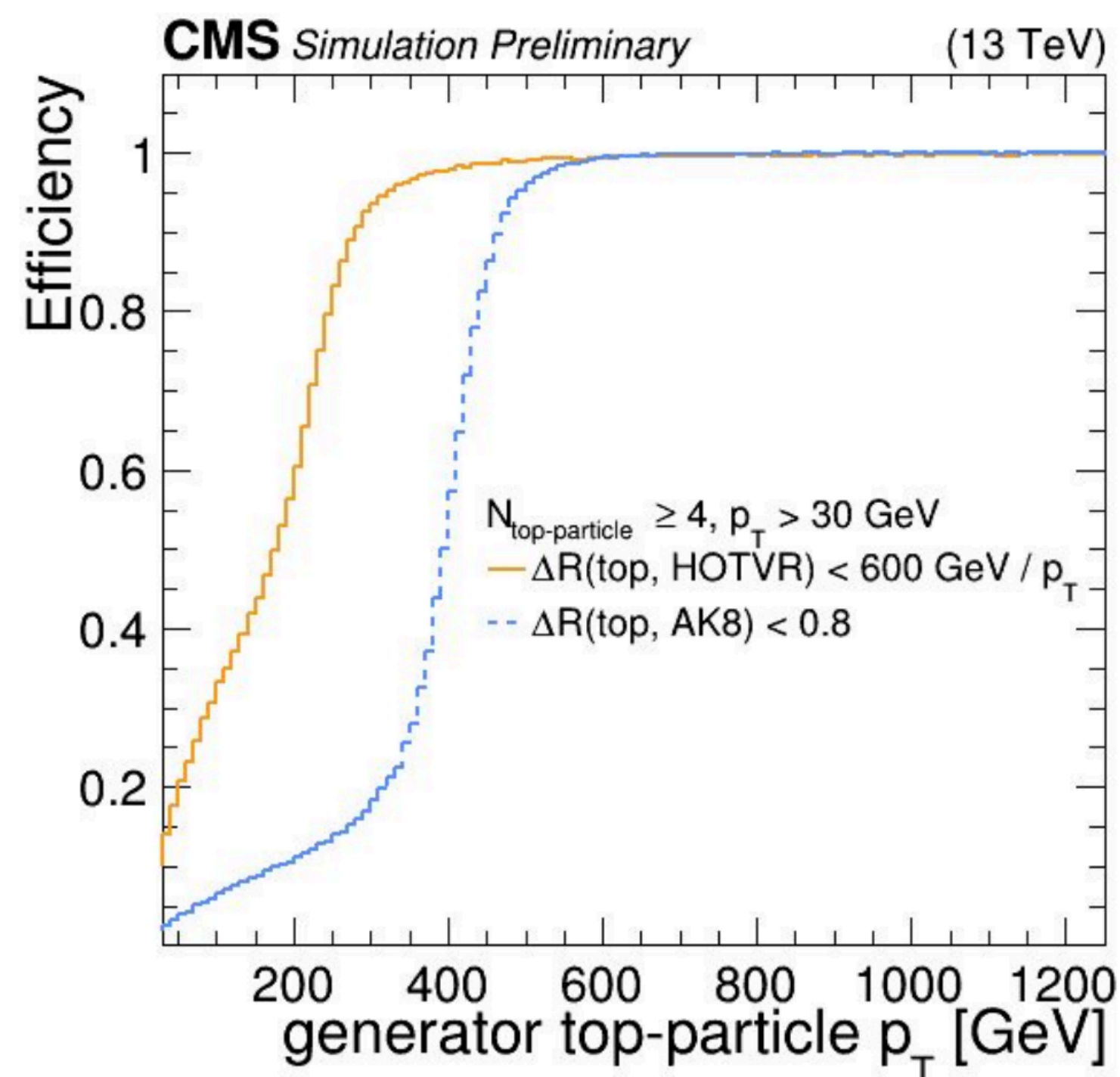
wrt. classical approach:

$$Q_k = \frac{\sum_i q_i (p_{T,i})^k}{(p_{T,jet})^k}$$

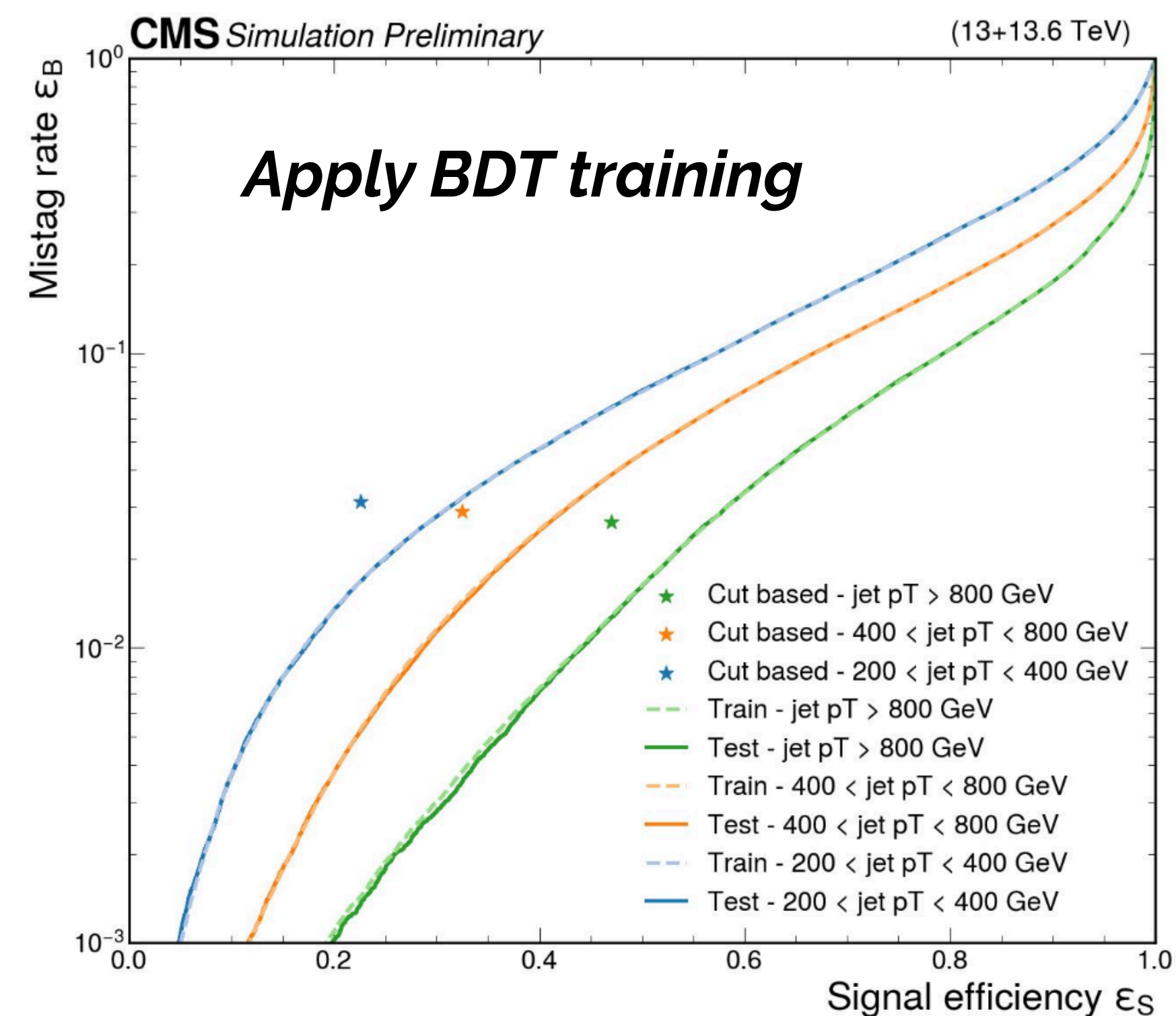
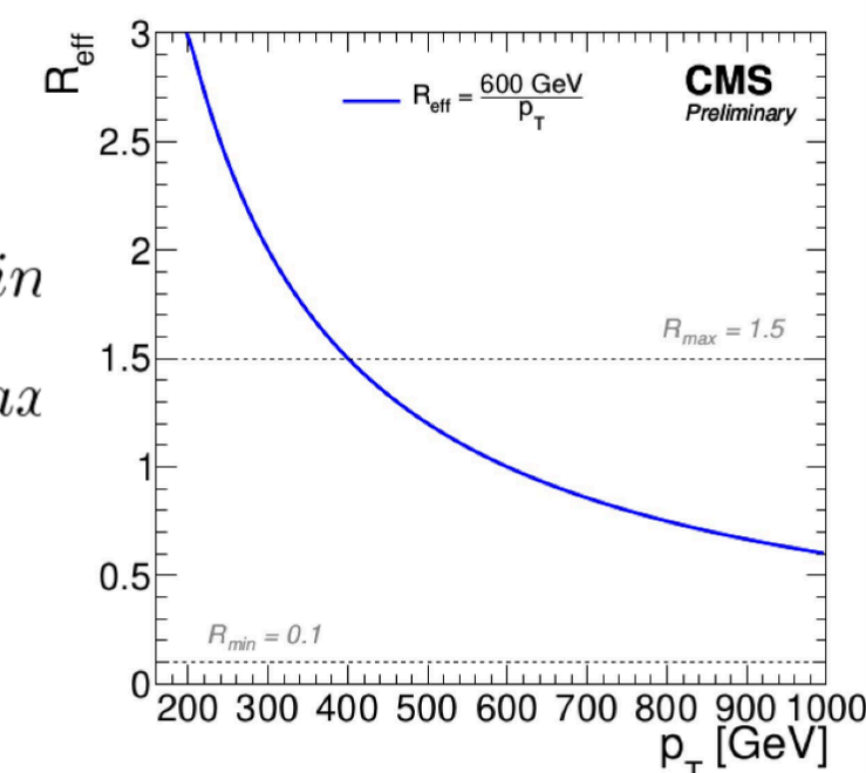
Extensions: HOTVR

- Above methods work well for “standard” boosted regime
 - Suboptimal for highly boosted (> 800 GeV) and intermediate (~ 200 GeV)
 - Use HOTVR (Heavy Object Tagger with Variable Radius) to cluster jets
 - Apply PUPPI

Extend reach to lower top quark p_T



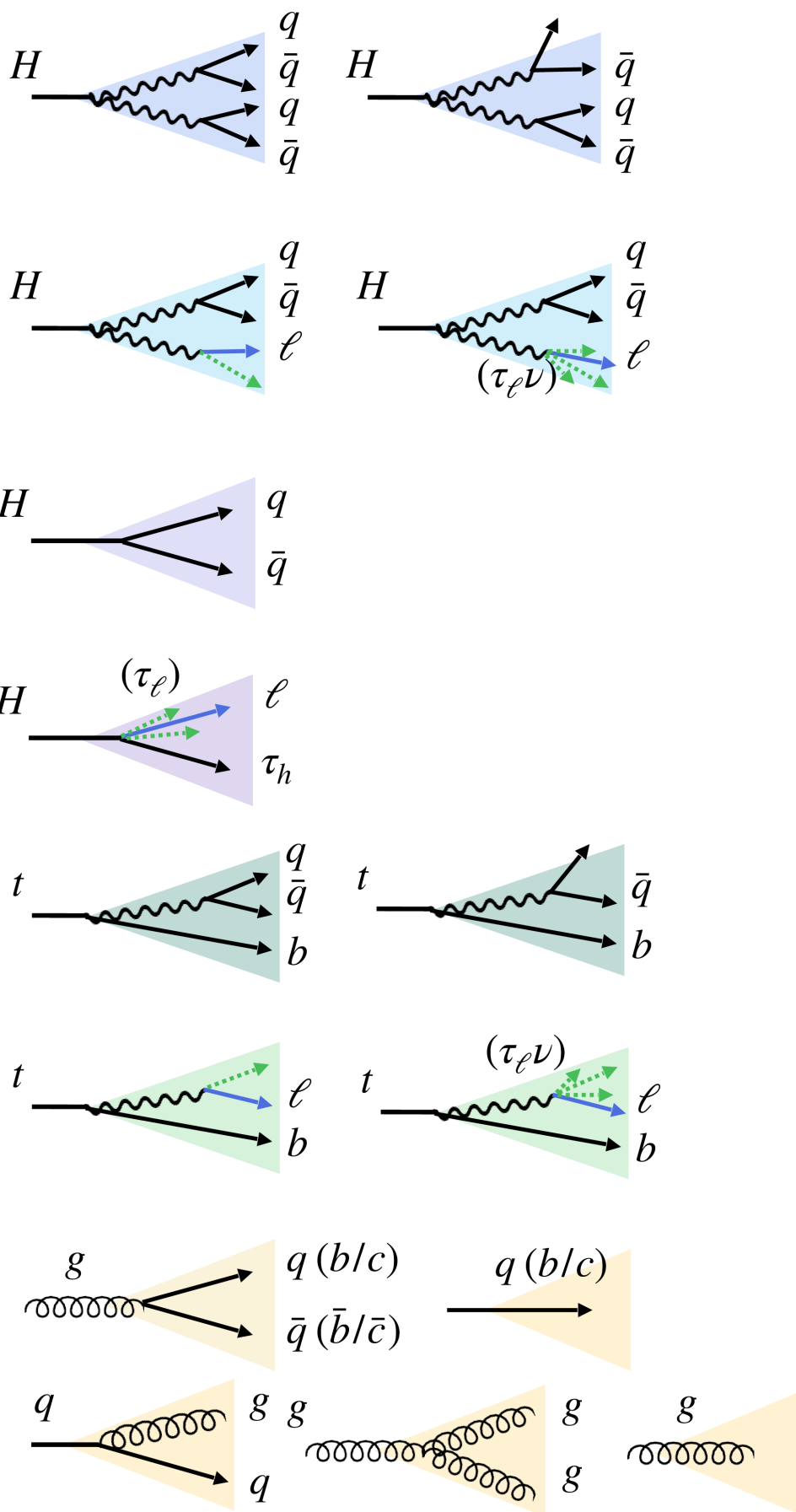
$$R_{eff} = \begin{cases} R_{min} & \rho/p_T < R_{min} \\ R_{max} & \rho/p_T > R_{max} \\ \rho/p_T & \text{else.} \end{cases}$$



Towards a global tagger?

- Leverage ParT architecture:
-
- Simultaneous classification + regression
 - Extension to 37 (!) classes
 - Allowing H->WW tagging for the first time!

Process	Final state		Flavor	# of classes
H→WW (full-hadronic)	qqqq	⊗	0c / 1c / 2c	3
	qqq			3
H→WW (semi-leptonic)	eνqq	⊗	0c / 1c	2
	μνqq			2
	τ _e νqq			2
	τ _μ νqq			2
	τ _h νqq			2
H→qq		⊗	bb	1
			cc	1
			ss	1
			qq (q=u/d)	1
H→ττ	τ _e τ _h			1
	τ _μ τ _h			1
	τ _h τ _h			1
t→bW (hadronic)	bqq	⊗	1b + 0c / 1c	2
	bq			2
t→bW (leptonic)	b e ν	⊗	1b	1
	b μ ν			1
	b τ _e ν			1
	b τ _μ ν			1
	b τ _h ν			1
QCD			b	1
			bb	1
			c	1
			cc	1
			others (light)	1



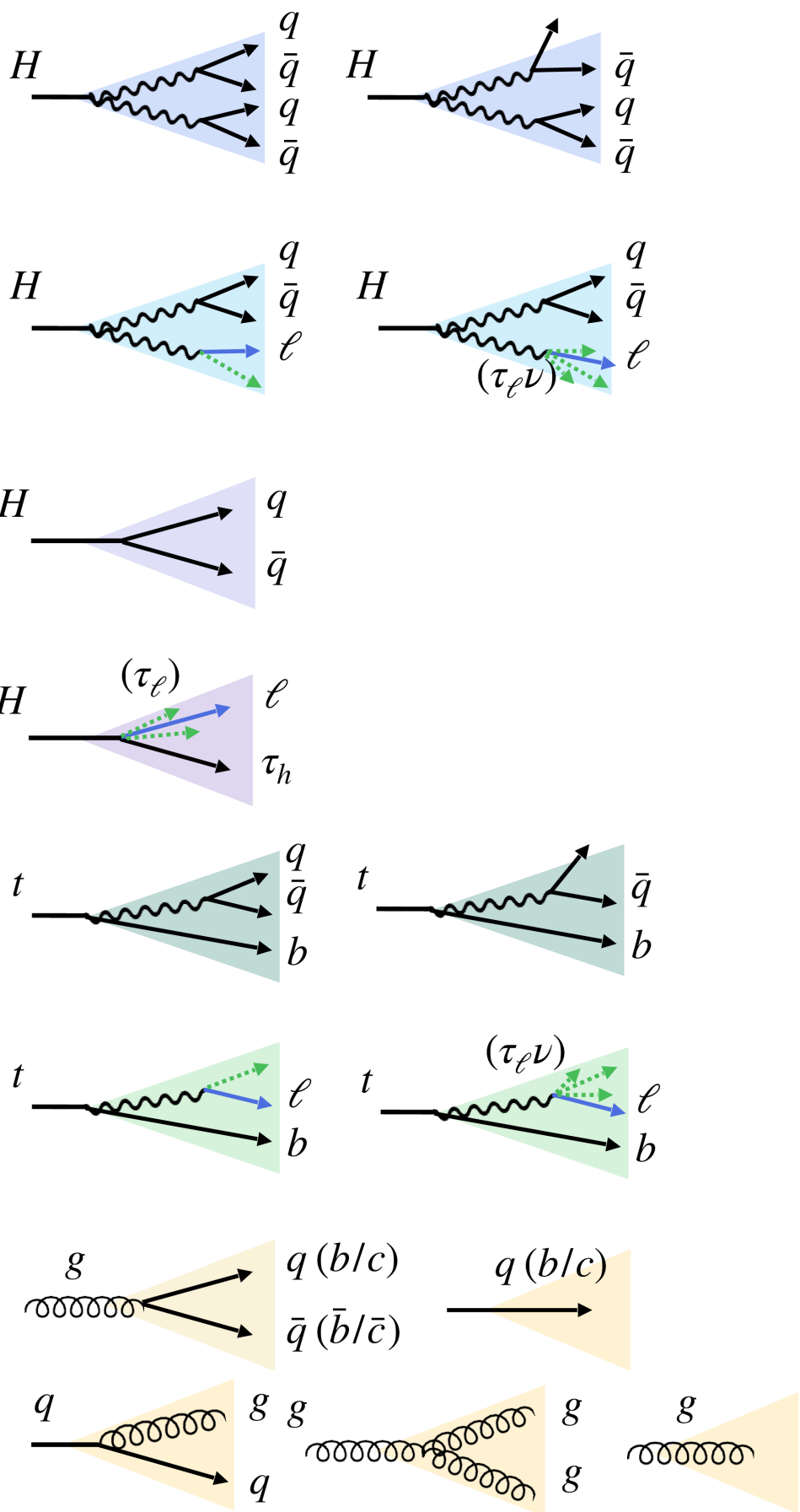
Towards a global tagger?

- Leverage ParT architecture:
 - Charged PF cand.
Neutral PF cand.
SVs
Pairwise features
 - Embedding
 - PAB (x8)
 - Class token
 - CAB
 - MLP
 - 37 class outputs
 - SoftMax
 - Mass output
- Simultaneous classification + regression
- Extension to 37 (!) classes
 - Allowing H->WW tagging for the first time!

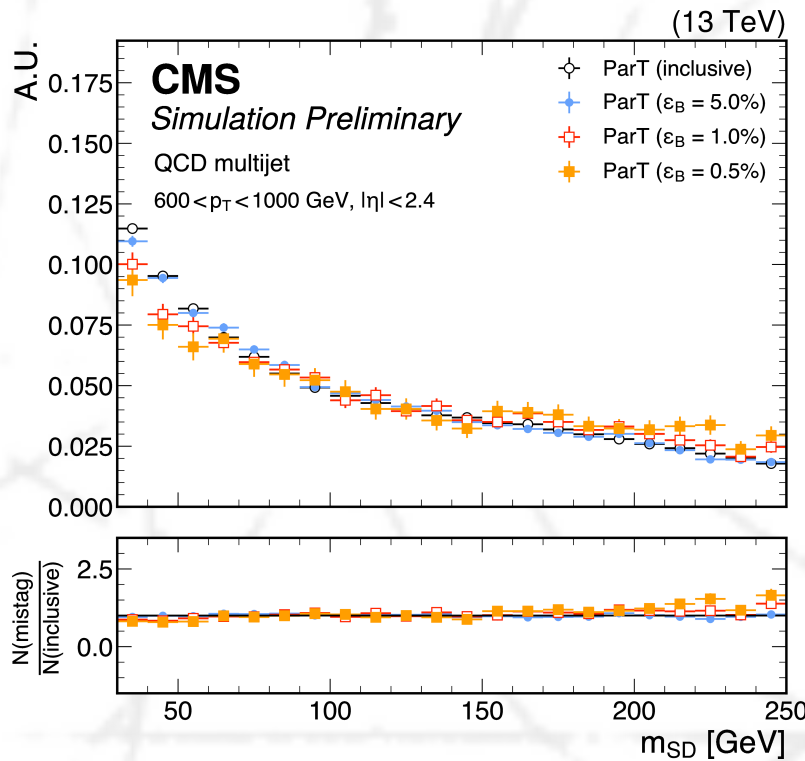
Mass decor relation via flat sampling

H-like ($G \rightarrow HH$)		t-like ($Z' \rightarrow t\bar{t}$)	
m_G	600–6000 GeV	$m_{Z'}$	600–6000 GeV
m_H	15–250 GeV	m_t	15–250 GeV
m_W	Fixed to SM ratio: $\frac{80.4}{125} m_H$	m_W	Fixed to SM ratio: $\frac{80.4}{172.5} m_t$
All other particles: SM values			

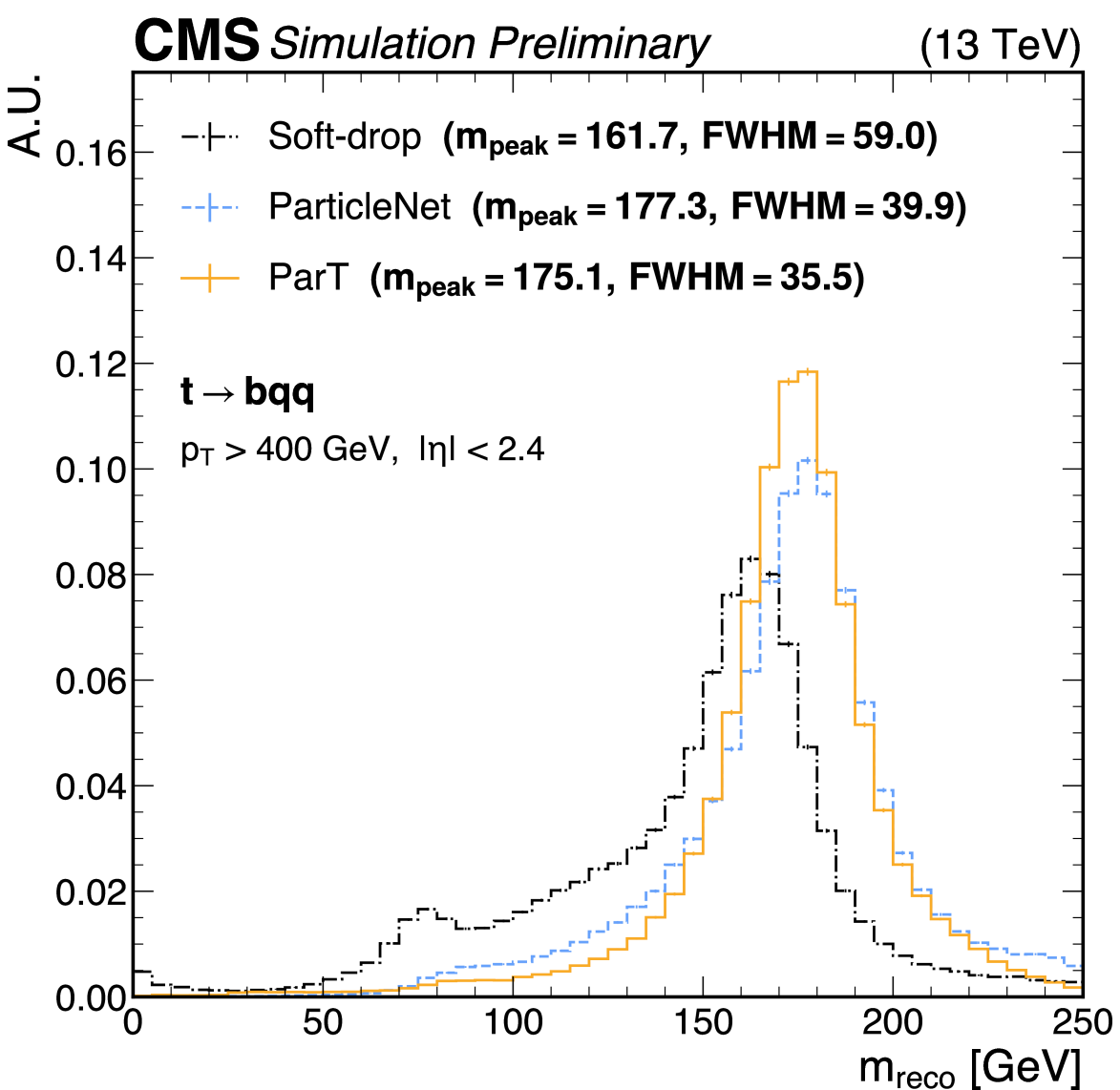
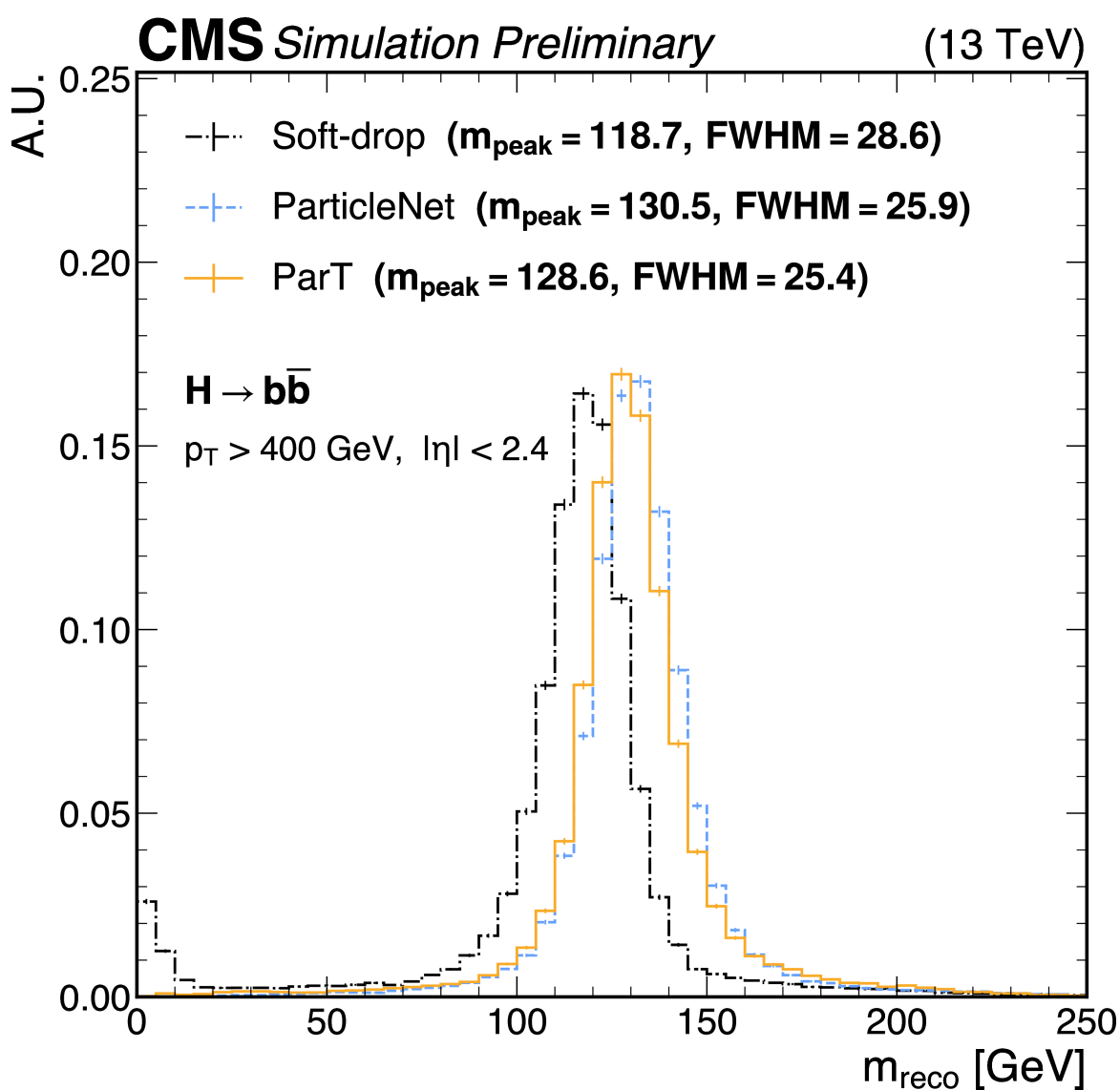
Process	Final state	Flavor	# of classes
H→WW (full-hadronic)	qqqq	$\otimes$ 0c / 1c / 2c	3
	qqq		3
H→WW (semi-leptonic)	eνqq	$\otimes$ 0c / 1c	2
	μνqq		2
	τeνqq		2
	τμνqq		2
	τhνqq		2
H→qq	$\otimes$	bb	1
		cc	1
		ss	1
		qq (q=u/d)	1
H→ττ	τeτh	$\otimes$	1
	τμτh		1
	τhτh		1
t→bW (hadronic)	bqq	$\otimes$ 1b + 0c / 1c	2
	bq		2
t→bW (leptonic)	bев	$\otimes$ 1b	1
	bμν		1
	bτeν		1
	bτμν		1
	bτhν		1
QCD	b	$\otimes$	1
	bb		1
	c		1
	cc		1
	others (light)		1



No sign of mass sculpting!

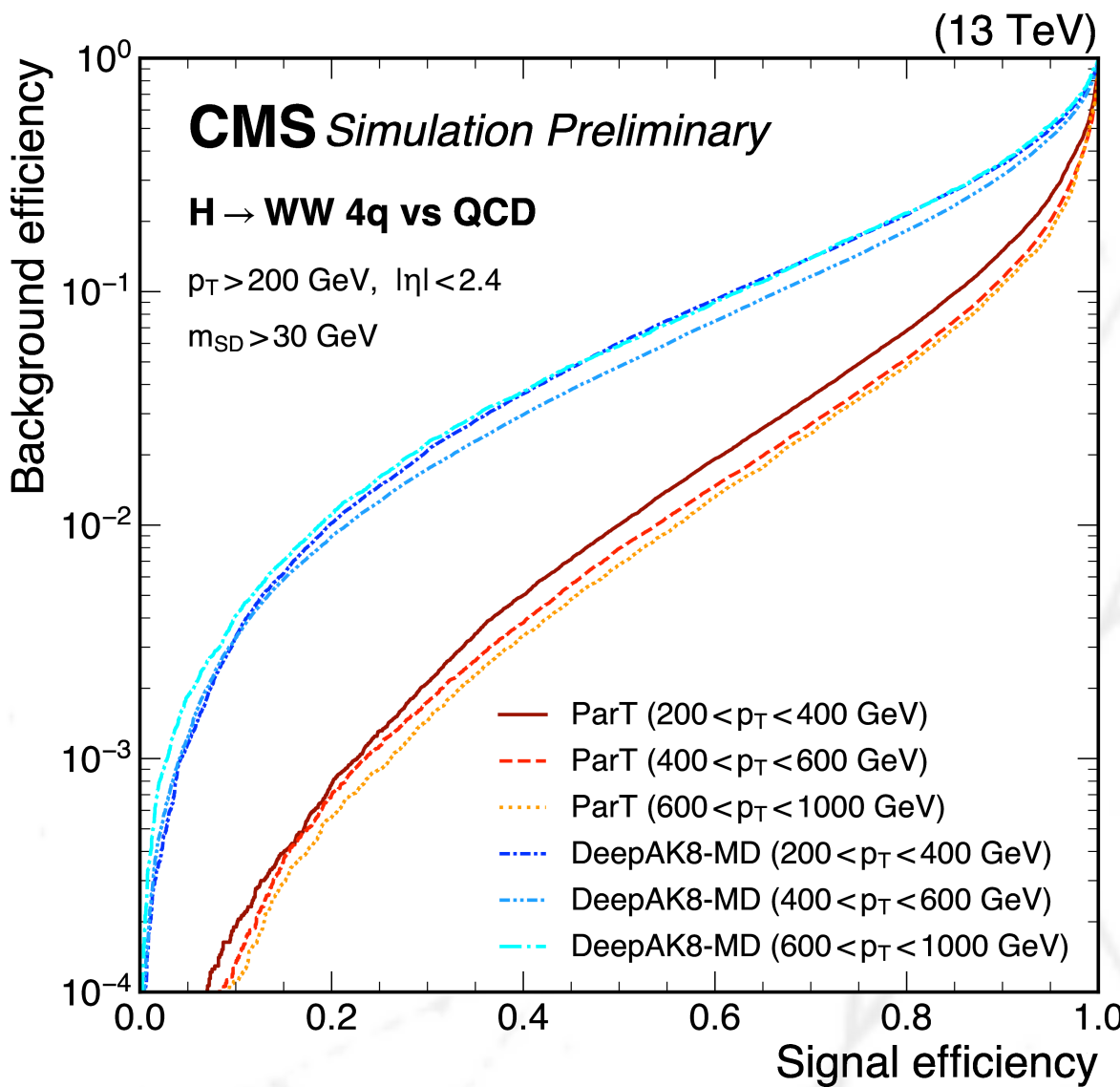
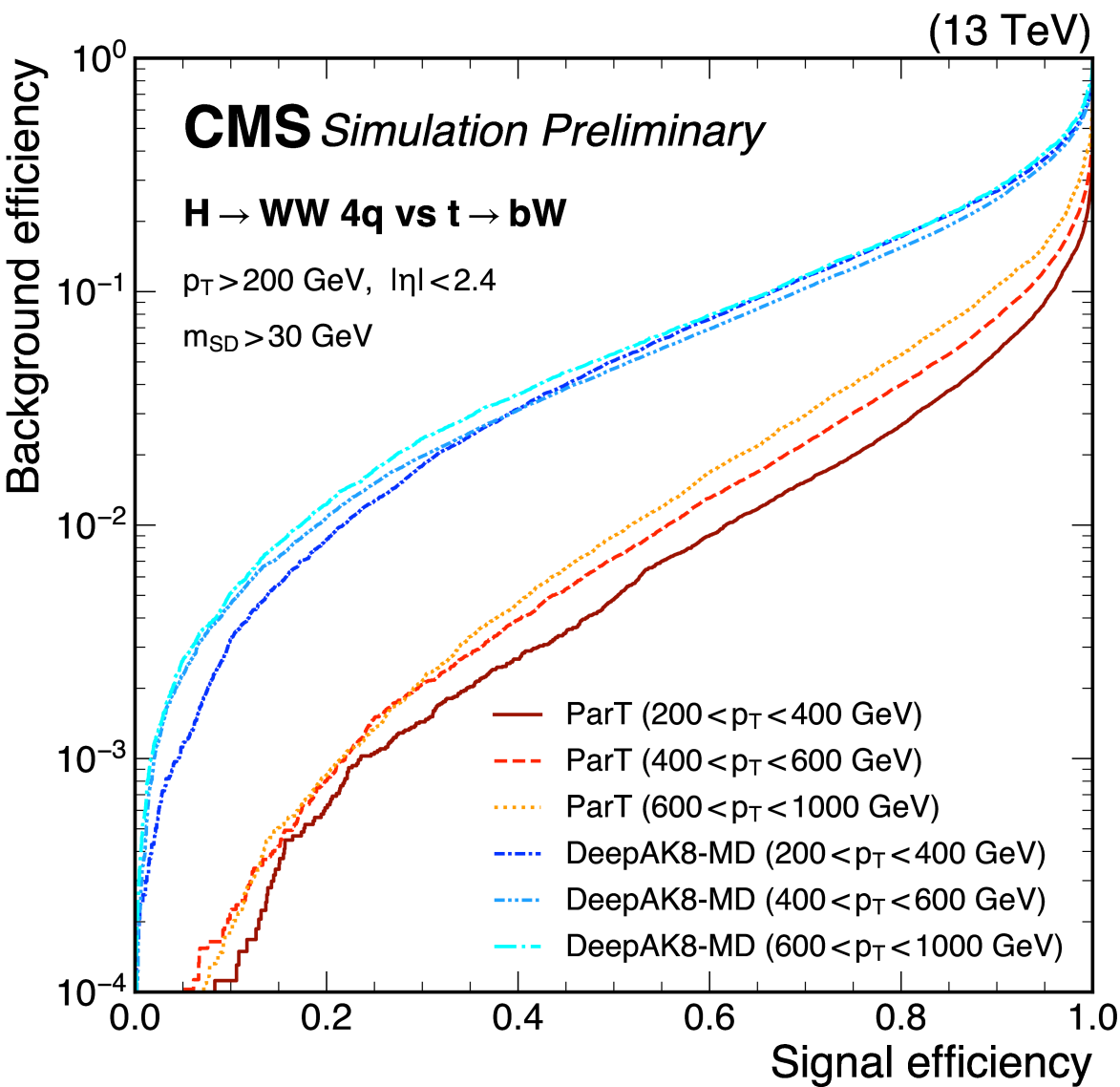
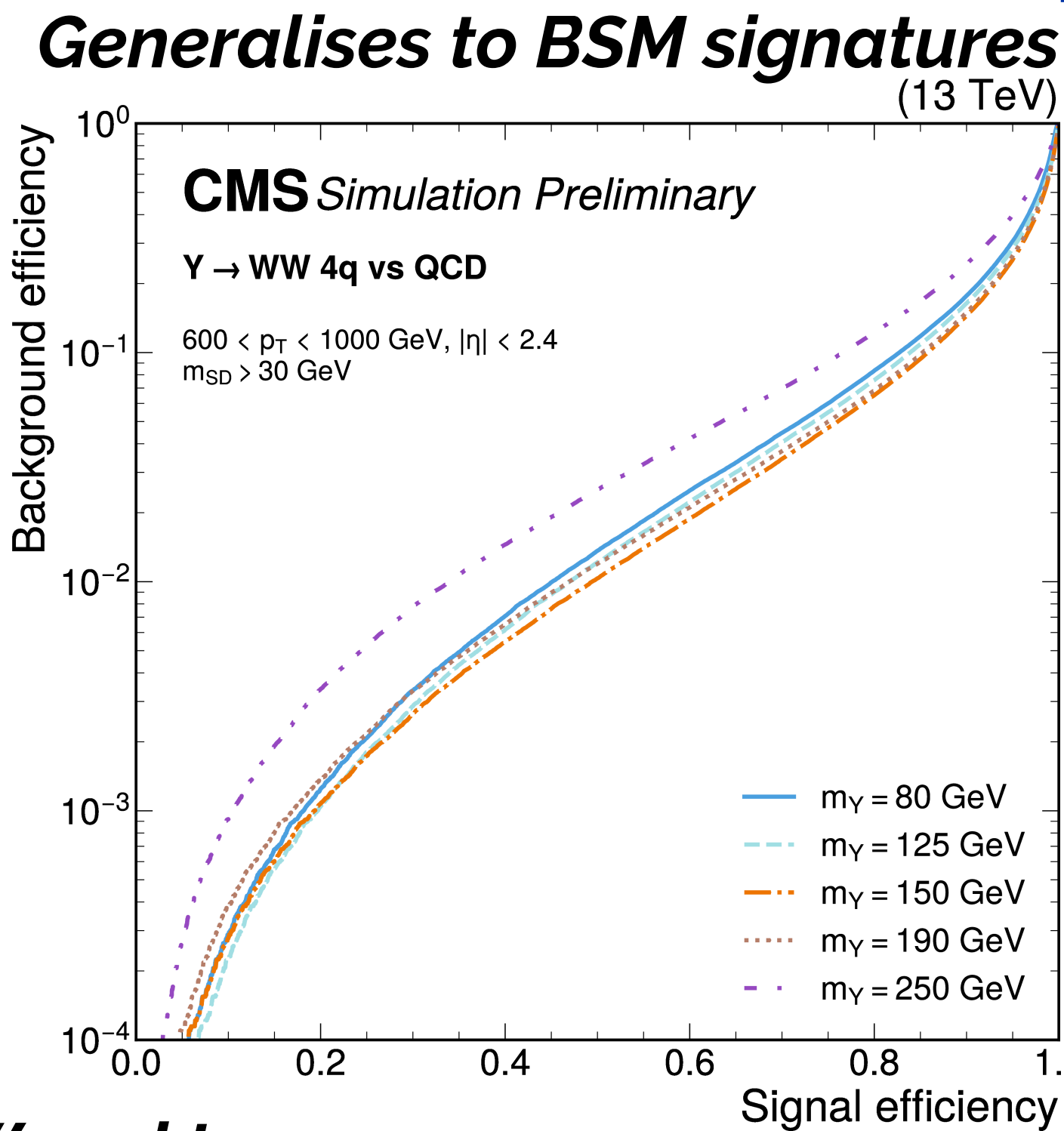


Towards a global tagger?

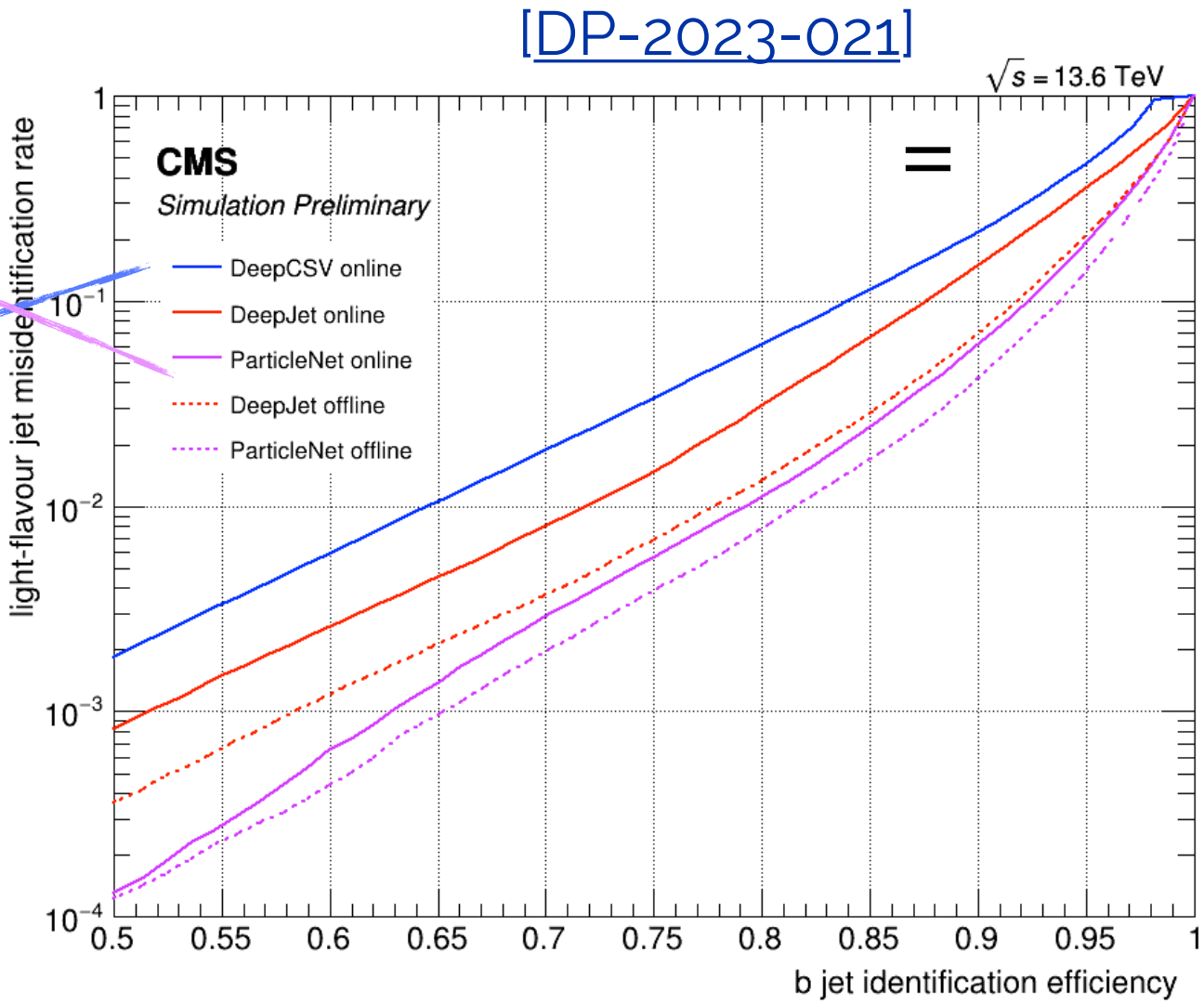
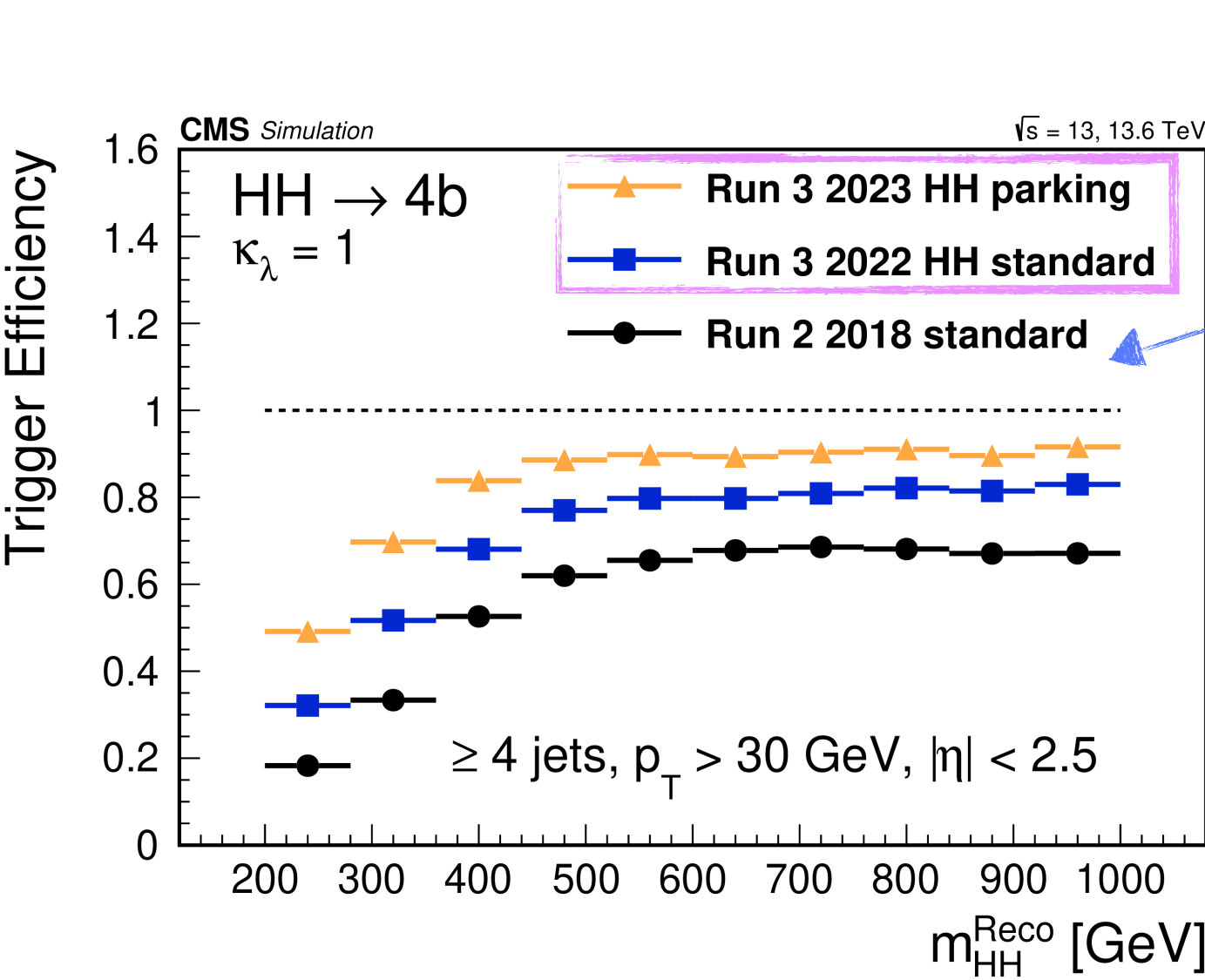
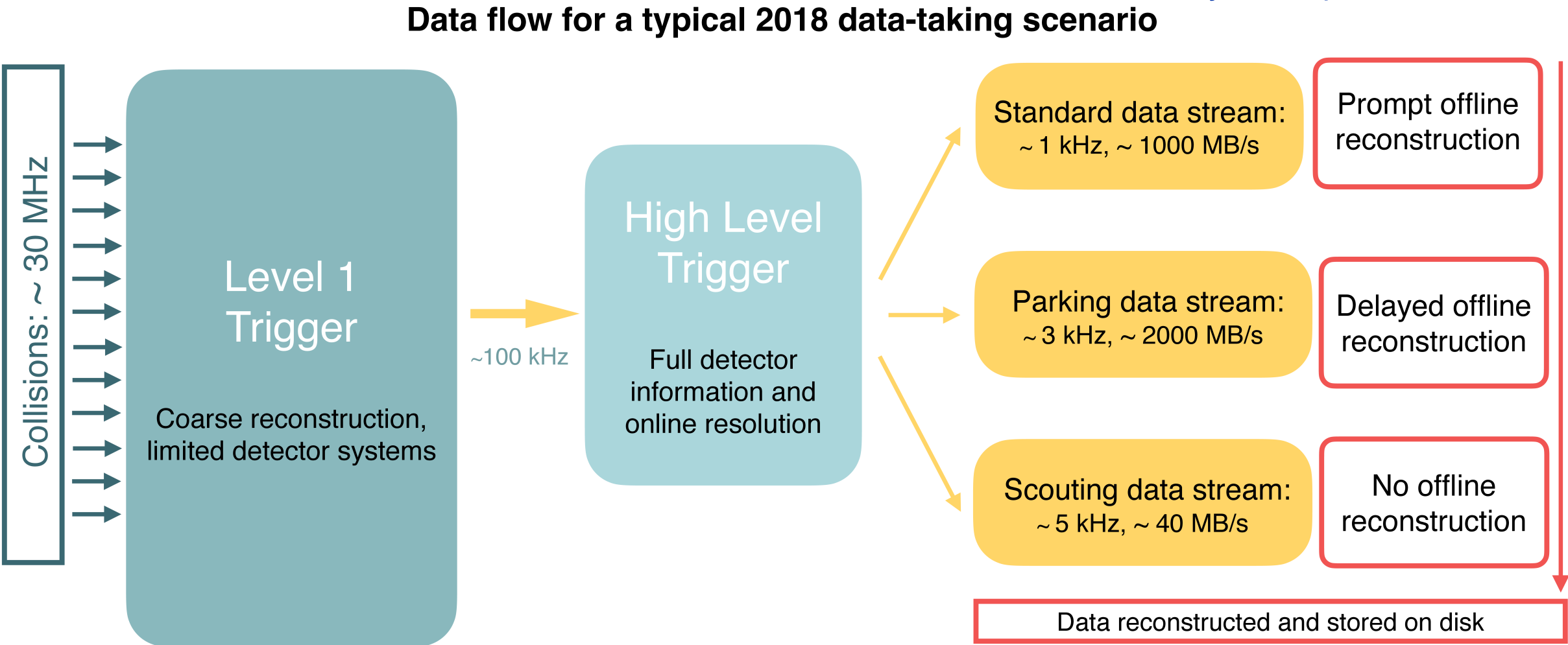
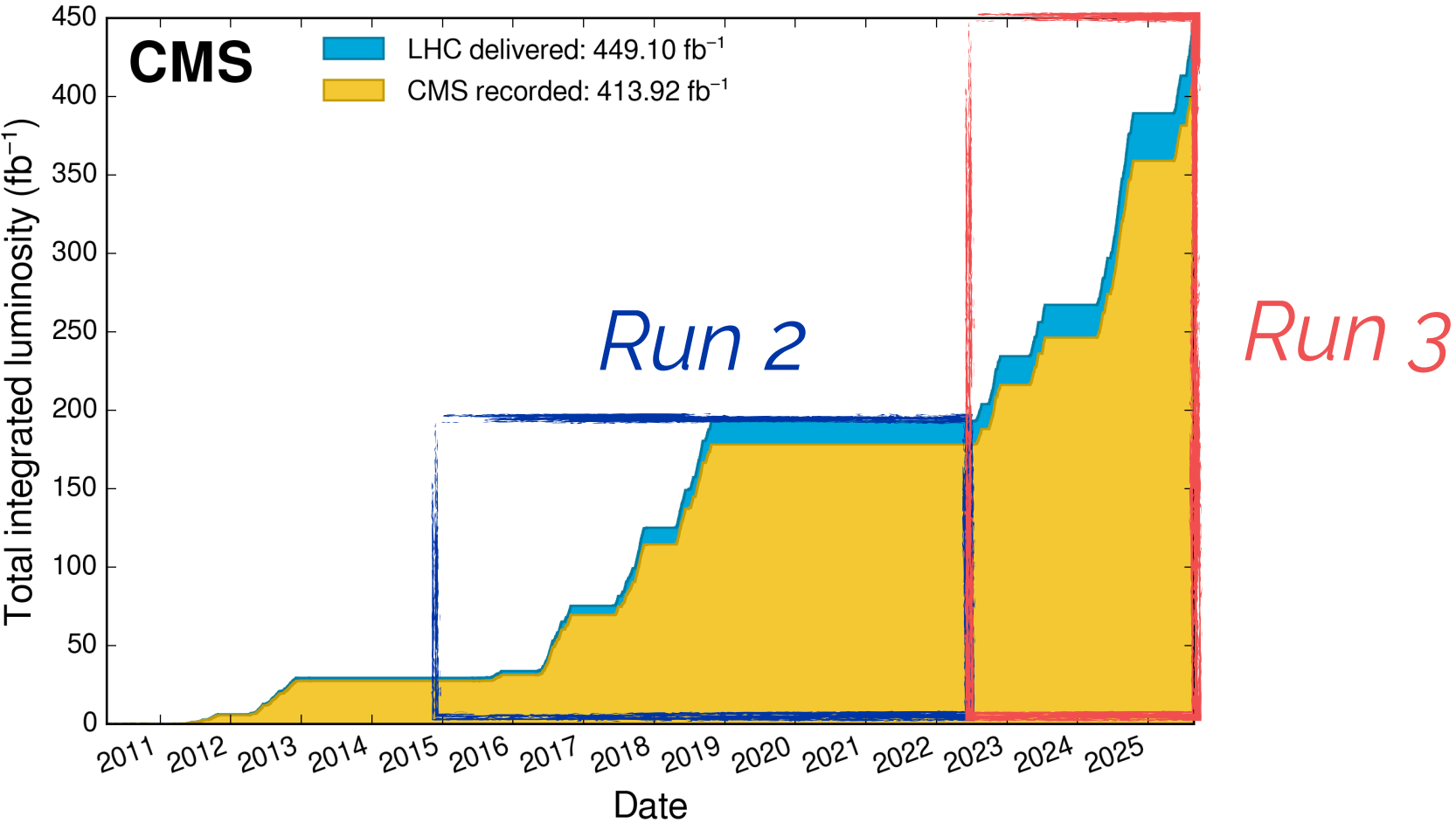


**Powerful multi-prong
mass regression**

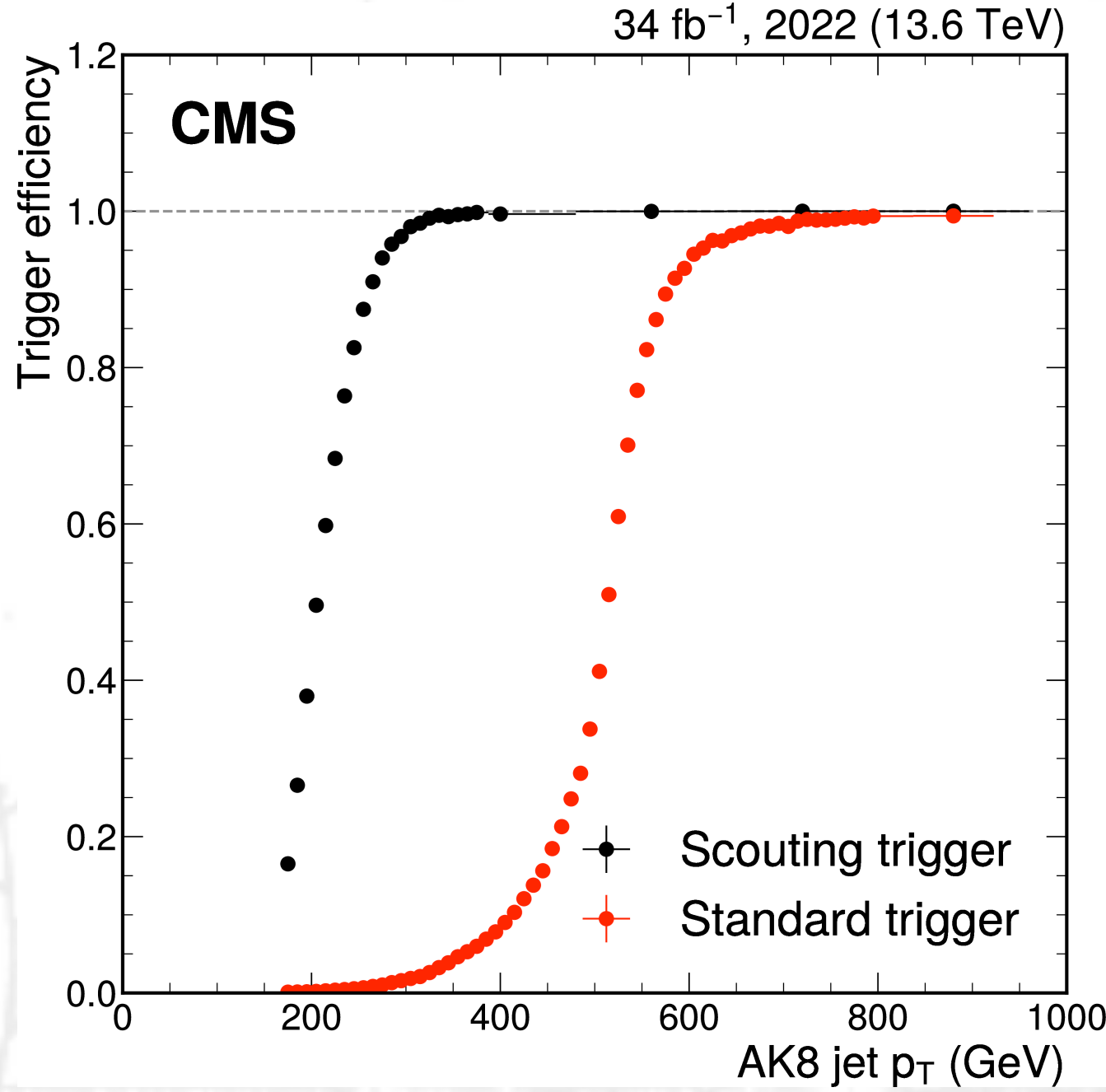
10x better background rejection for WW and top



Tagging at the trigger level



Extended acceptance + ParticleNetAK8 in scouting and trigger



Usage of ParticleNetAK4 in software trigger and scouting, up to 60% improvement

See Jieun' and Marko's talks

Multiclass event classification



- Simultaneous $t\bar{t}H(H \rightarrow c\bar{c})$ and $t\bar{t}H(H \rightarrow b\bar{b})$ measurement to handle $t\bar{t}H(H \rightarrow b\bar{b})$ background
- **Introduced dedicated categories for $t\bar{t}Z(Z \rightarrow c\bar{c})$ and $t\bar{t}Z(Z \rightarrow b\bar{b})$**
 - Validation for $t\bar{t}H$ measurement, similar in the $VH(H \rightarrow c\bar{c})$ / $ggH(H \rightarrow c\bar{c})$ analyses
 - Employs a symmetric approach between $t\bar{t}H$ and $t\bar{t}Z$

- The target are four signal classes:

$t\bar{t}H(H \rightarrow c\bar{c})$

$t\bar{t}H(H \rightarrow b\bar{b})$

$t\bar{t}Z(Z \rightarrow c\bar{c})$

$t\bar{t}Z(Z \rightarrow b\bar{b})$

- Five background classes:

$t\bar{t} + \geq 2c$

$t\bar{t} + c$

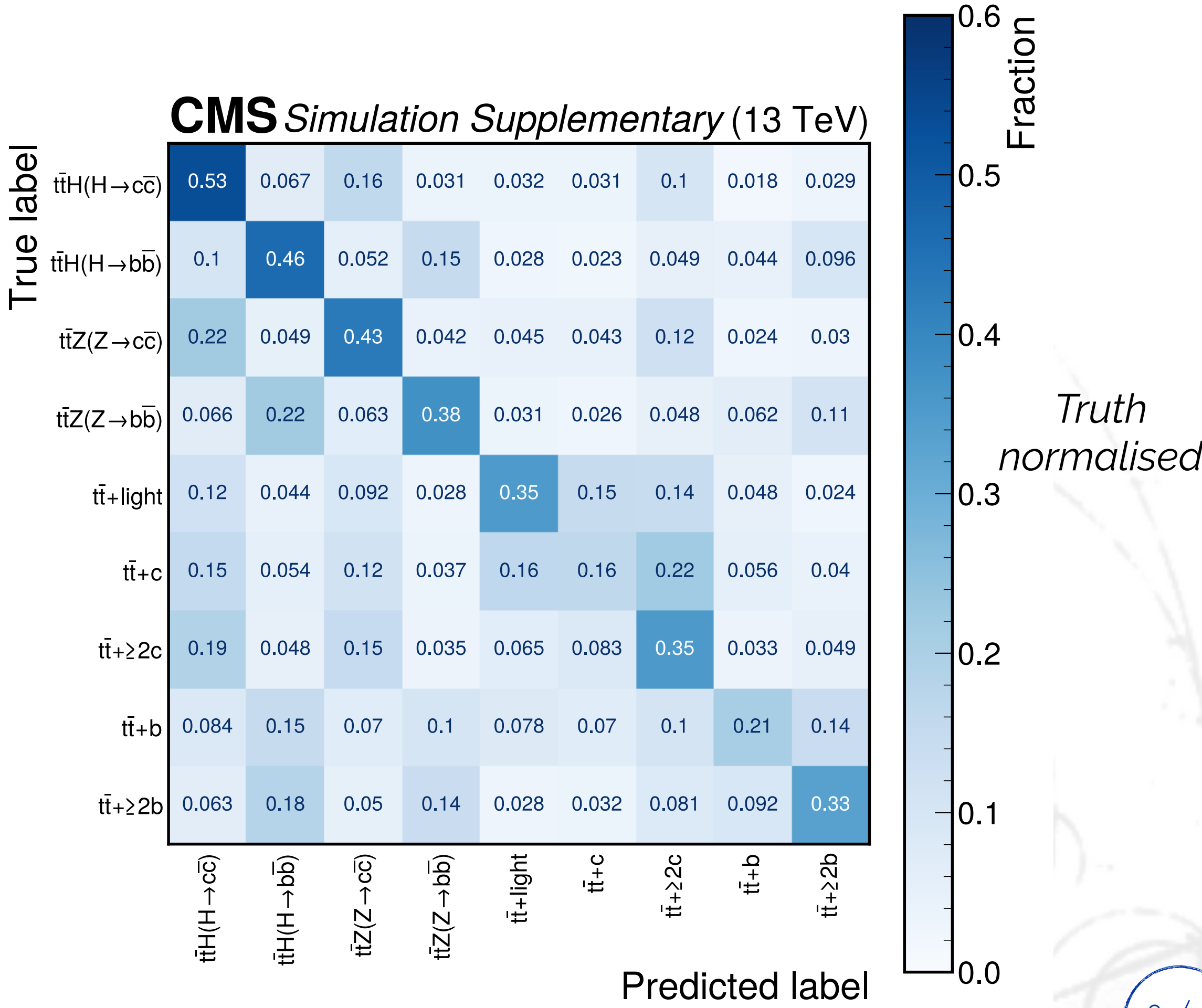
$t\bar{t} + \geq 2b$

$t\bar{t} + b$

$t\bar{t} + \text{light}$

- And in the oL channel one additional background class:

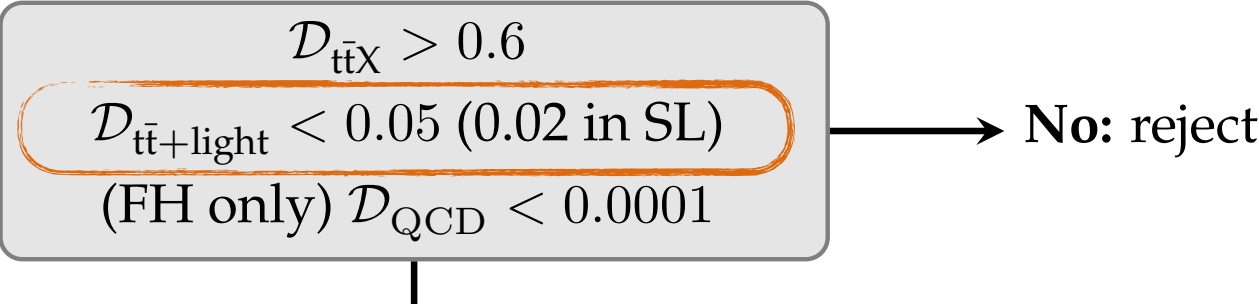
QCD



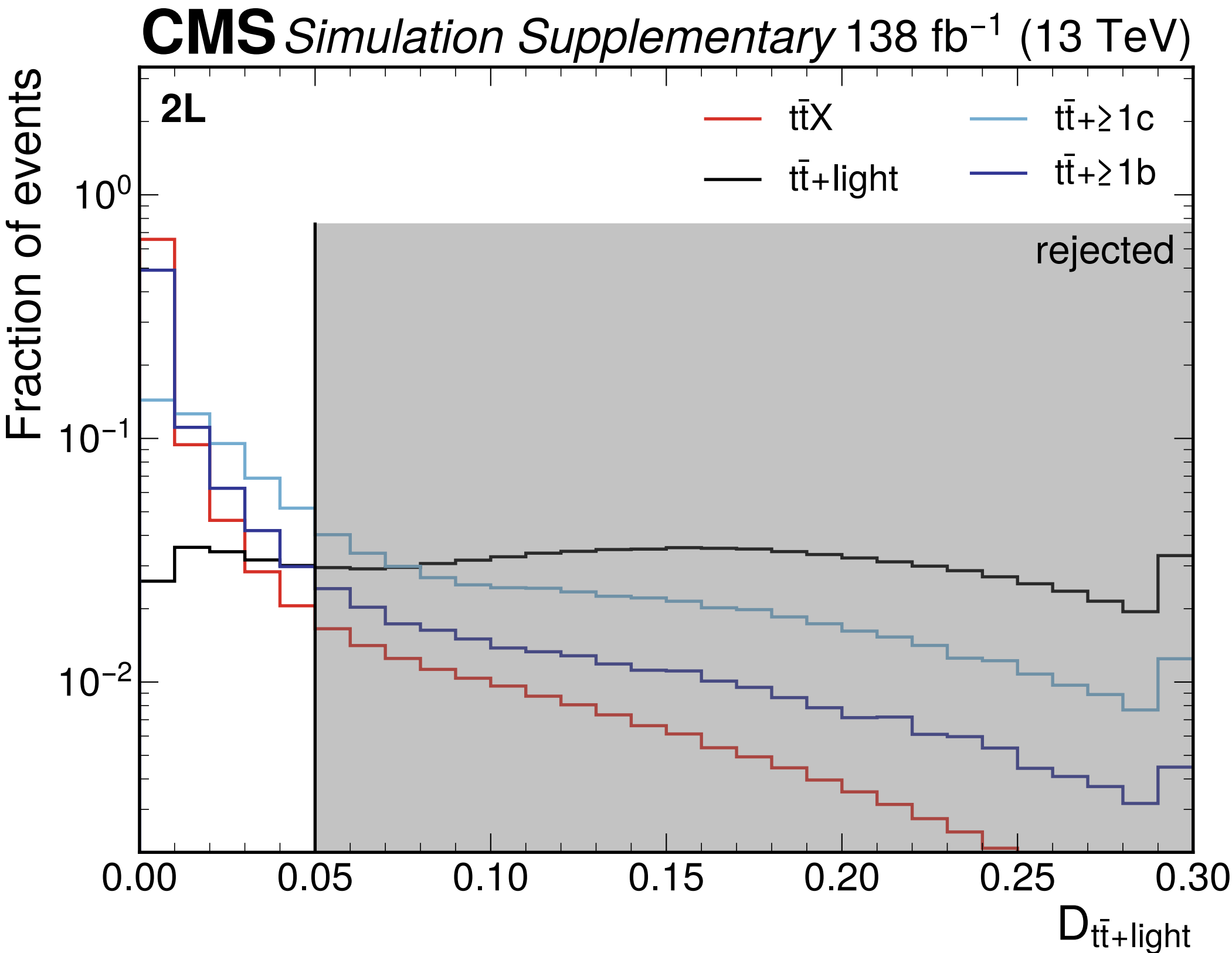
Event selection & categorisation



- Use scores (D_X) to select & categorise events



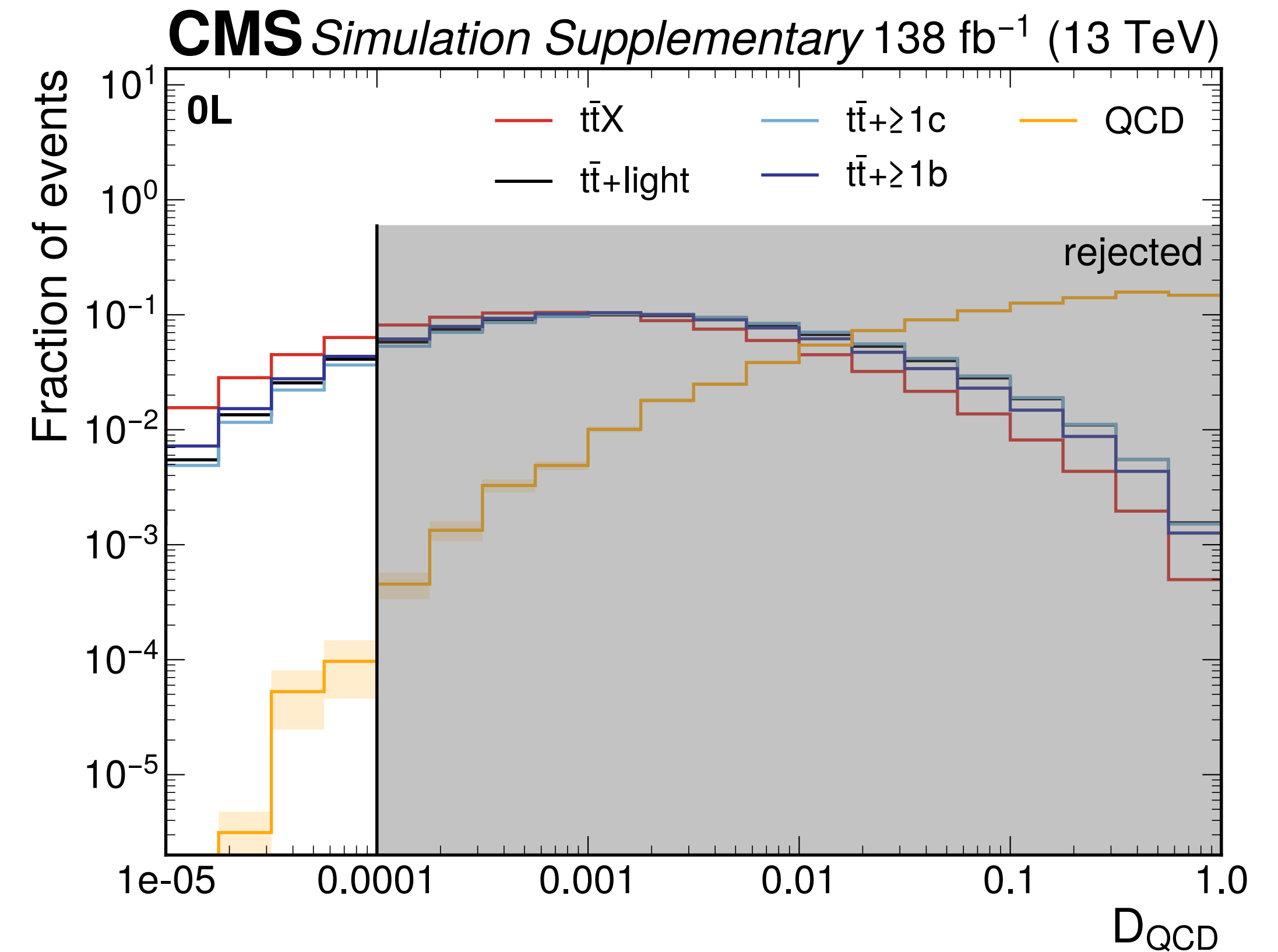
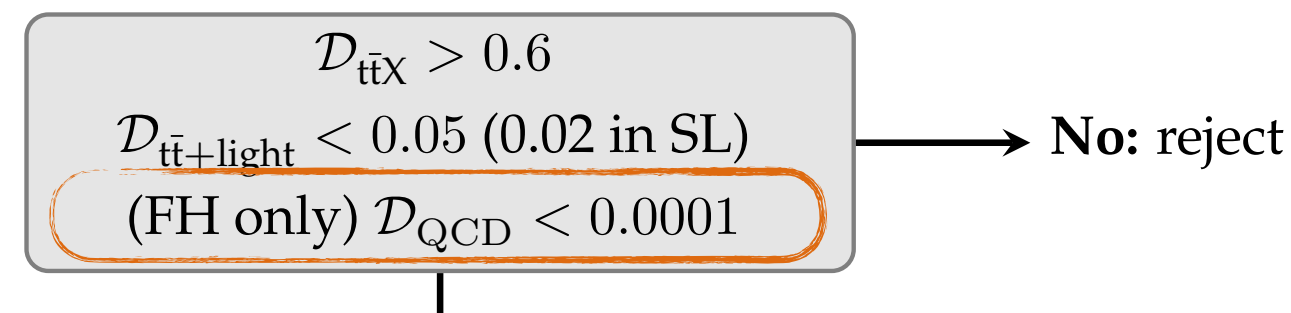
- Remove QCD and $t\bar{t}$ +light contributions via selections on background scores





Event selection & categorisation

- Use scores (D_X) to select & categorise events



- Remove QCD and $t\bar{t}+light$ contributions via selections on background scores

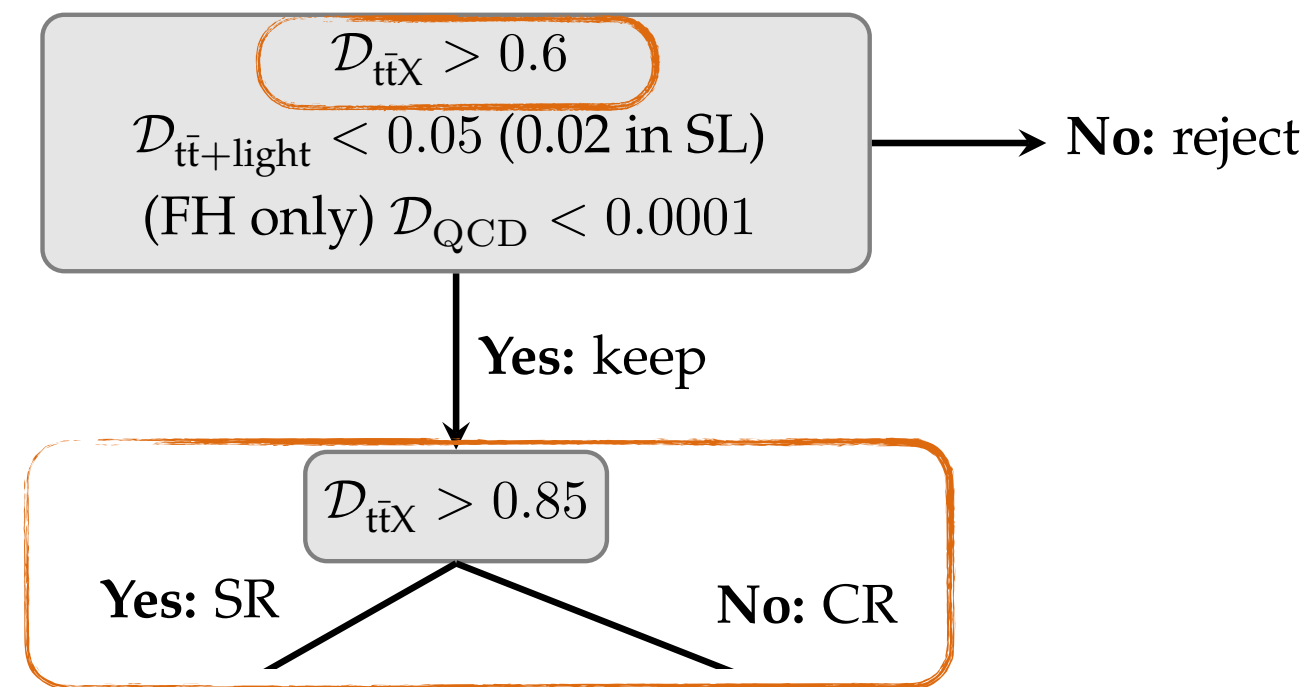
QCD suppression in FH channel:

- Event classifier very good at separating QCD $\leftrightarrow$ $t\bar{t}$
- Only use events with QCD score $< 10^{-4}$
 - QCD contribution becomes negligible ($< 1\%$)**
 - At the cost of signal efficiency ($\sim 15\%$)

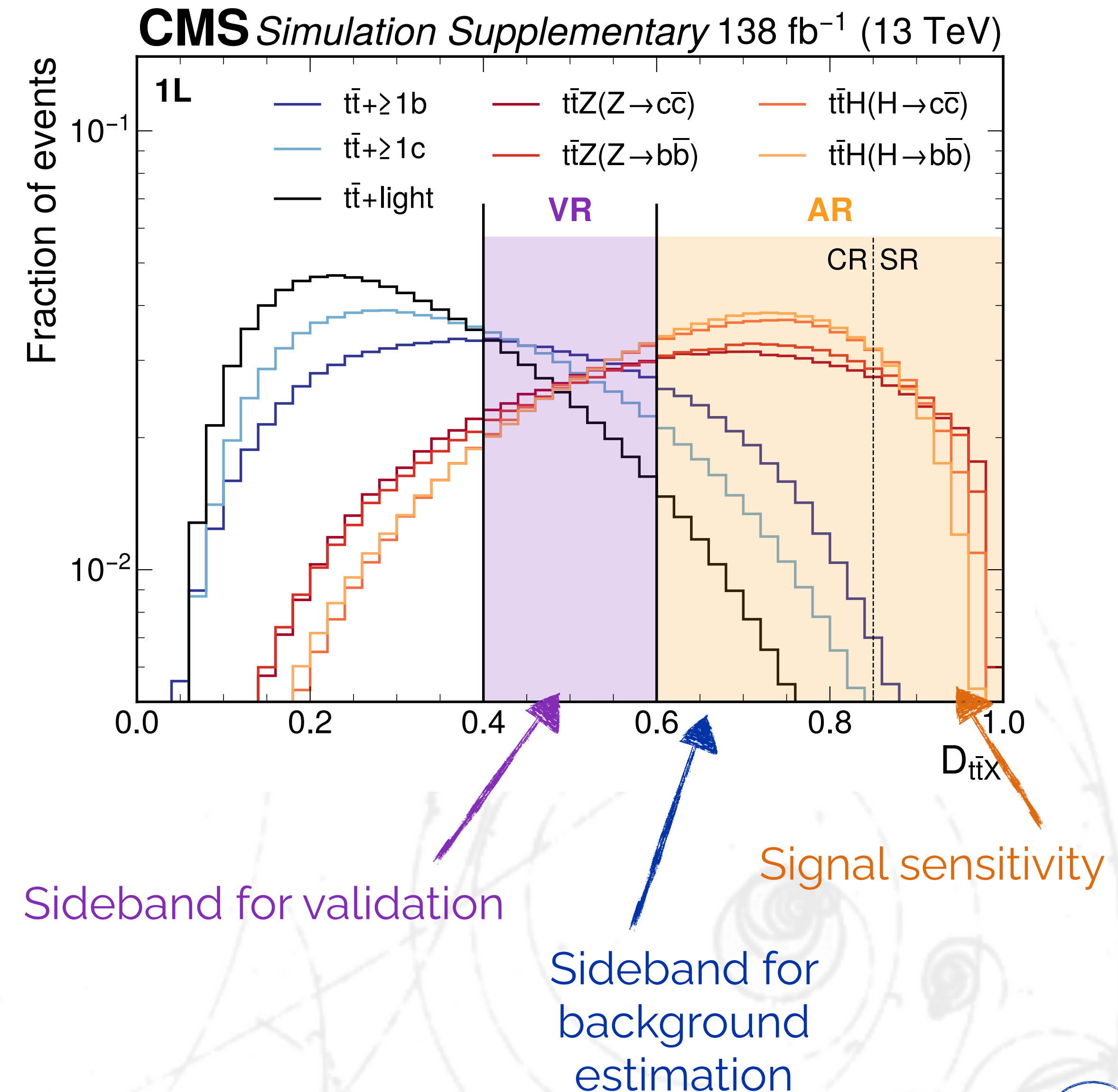


Event selection & categorisation

- Use scores (D_X) to select & categorise events



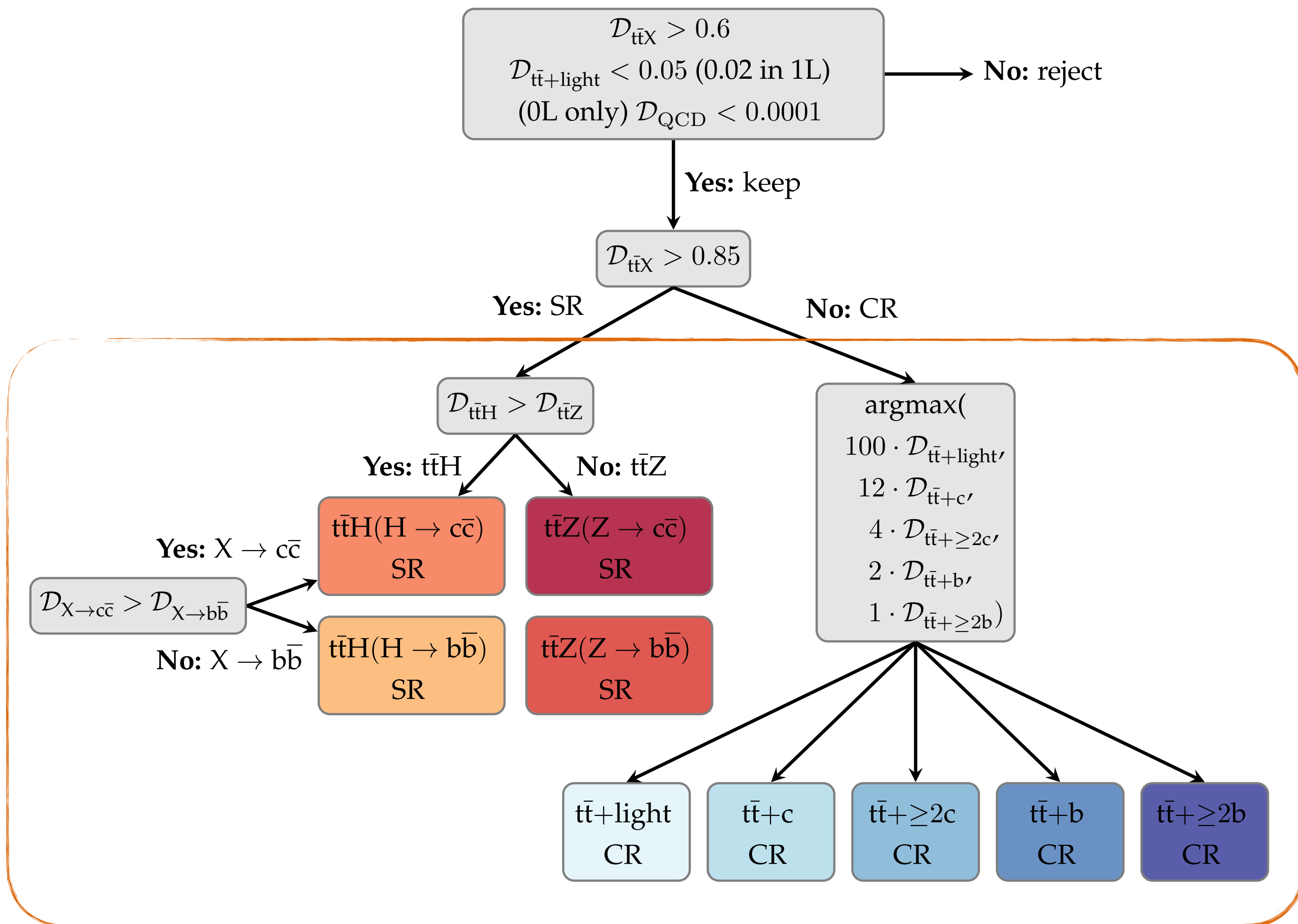
- Analysis region (**AR**) with signal-like events
→ Reduce background extrapolation
- Split **AR** into **signal region part (SR)** and **control region part (CR)**
- Introduce validation region (**VR**) to test the background model





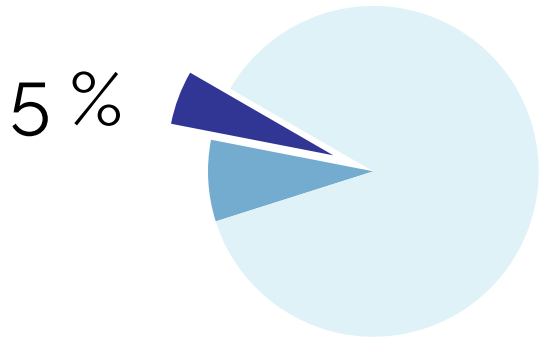
Event selection & categorisation

- Use scores (D_X) to select & categorise events



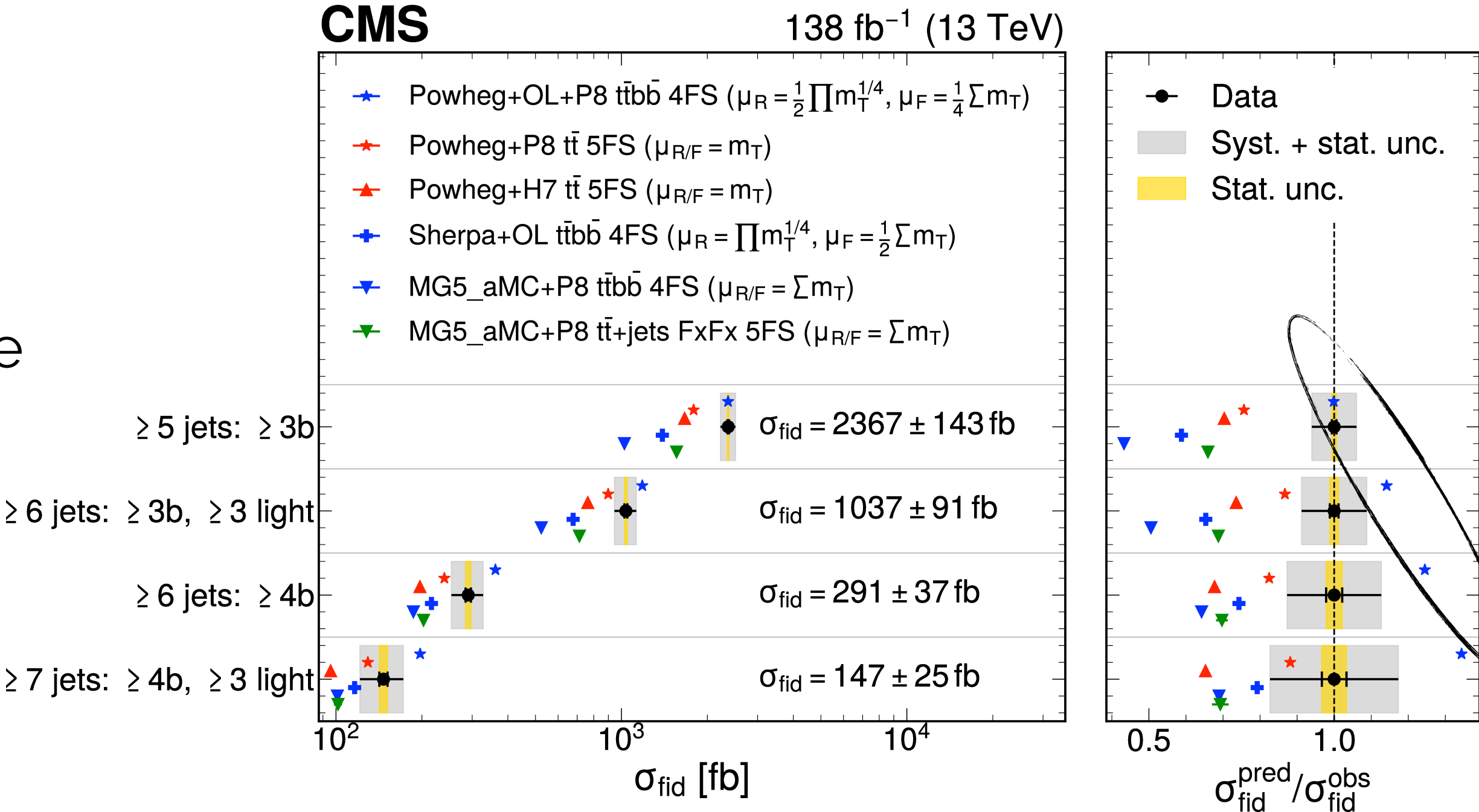
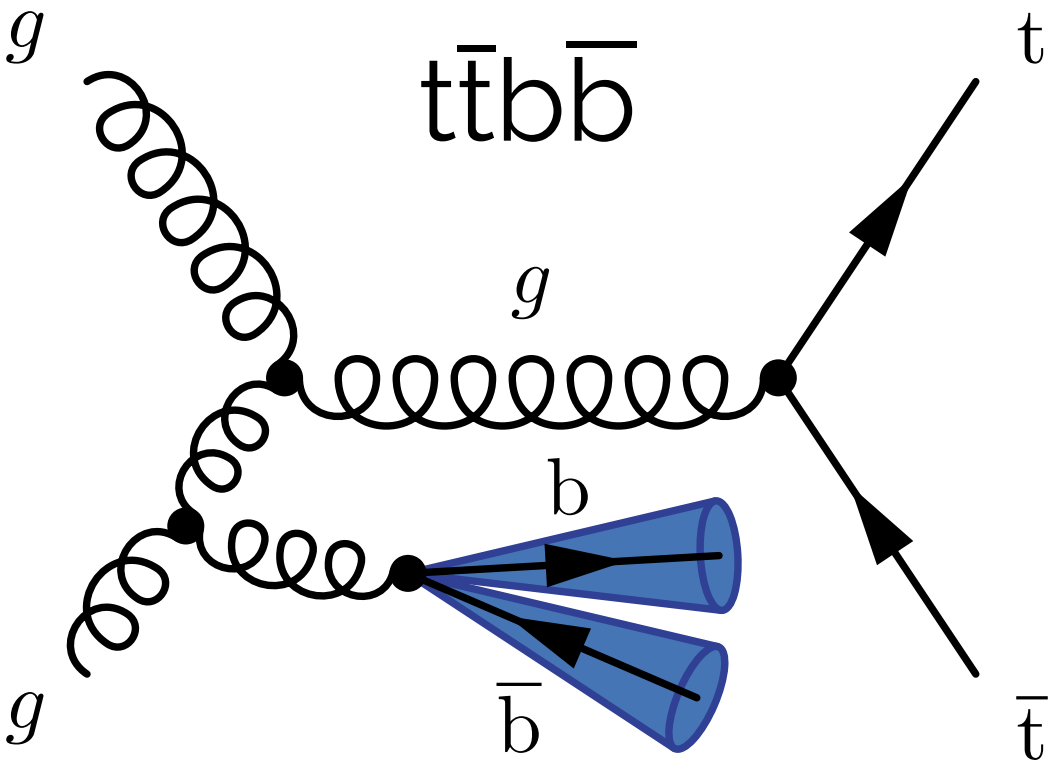
- Categorise signal-like events within **SR** into **2+2** dedicated **signal regions (SRs)**:
 - $t\bar{t}H(H \rightarrow c\bar{c})$, $t\bar{t}H(H \rightarrow b\bar{b})$, $t\bar{t}Z(Z \rightarrow c\bar{c})$, $t\bar{t}Z(Z \rightarrow b\bar{b})$
- And **5 control regions (CRs)**:
 - $t\bar{t}+light$, $t\bar{t}+c$, $t\bar{t}+ \geq 2c$, $t\bar{t}+b$, $t\bar{t}+ \geq 2b$
- 9 regions in total for **AR**
- Employ importance weights to increase purities
- VR**: ensure same background purities as **AR** to validate the final fit strategy

How well do we know the $t\bar{t}$ +jets background?



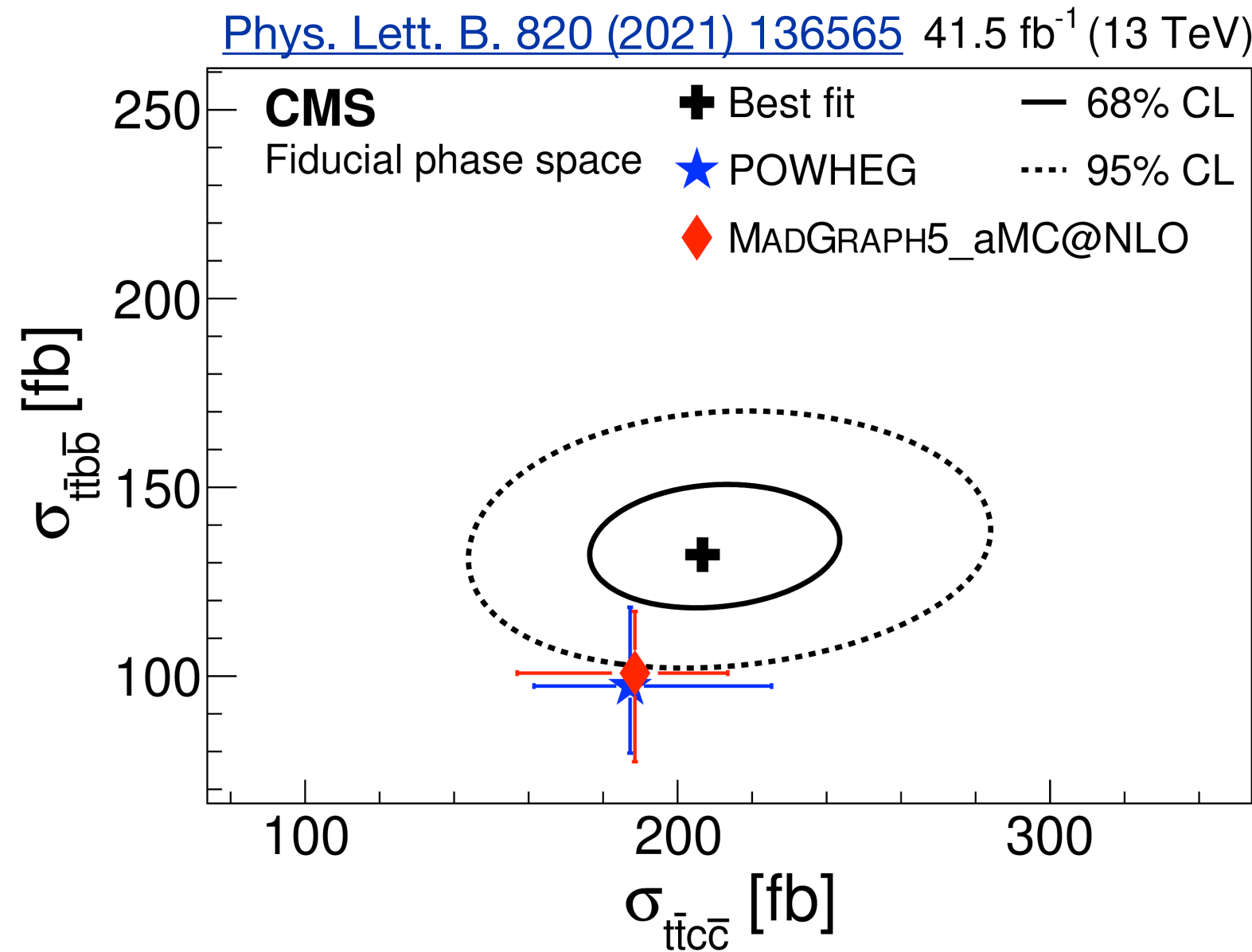
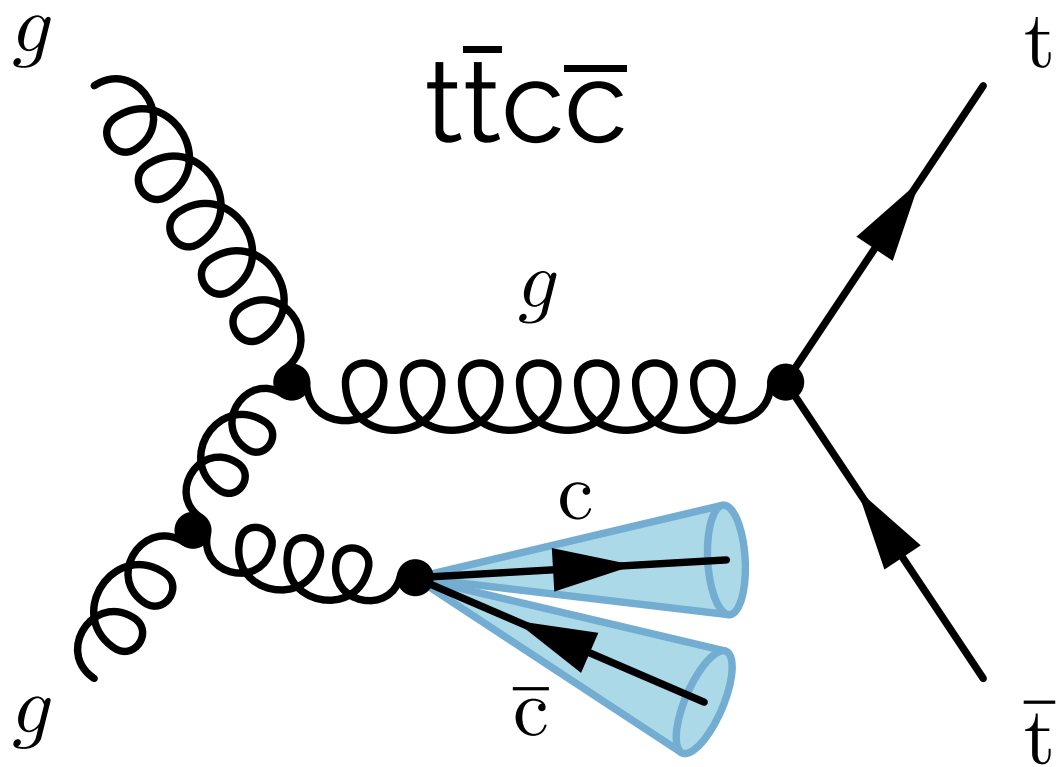
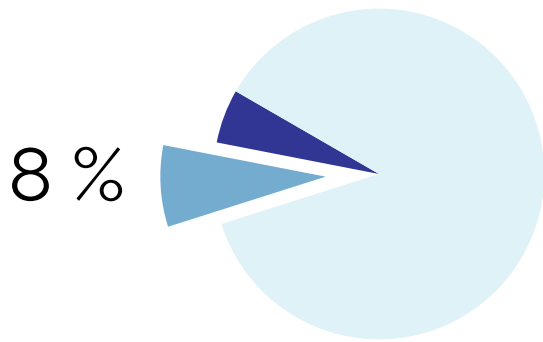
- Inclusive $t\bar{t}$ well described by **NLO Powheg+Pythia8 5FS** prediction
 - See e.g. multi-differential cross section measurements
- For $t\bar{t}b\bar{b}$:
 - Theoretical difficulties:
 - Massless or massive b-quarks?
e.g. **5FS inclusive NLO** prediction vs. **4FS dedicated $t\bar{t}b\bar{b}$** simulation
 - Experimental difficulties:
 - What do we consider a b-quark jet?
- Rich set of $t\bar{t}b\bar{b}$ measurements:
 - No generator has a good description everywhere
 - **4FS $t\bar{t}b\bar{b}$ simulation** close to data inclusively
 - For differential distributions, agreement varies...

A background model with enough flexibility needed, based on 4FS $t\bar{t}b\bar{b}$

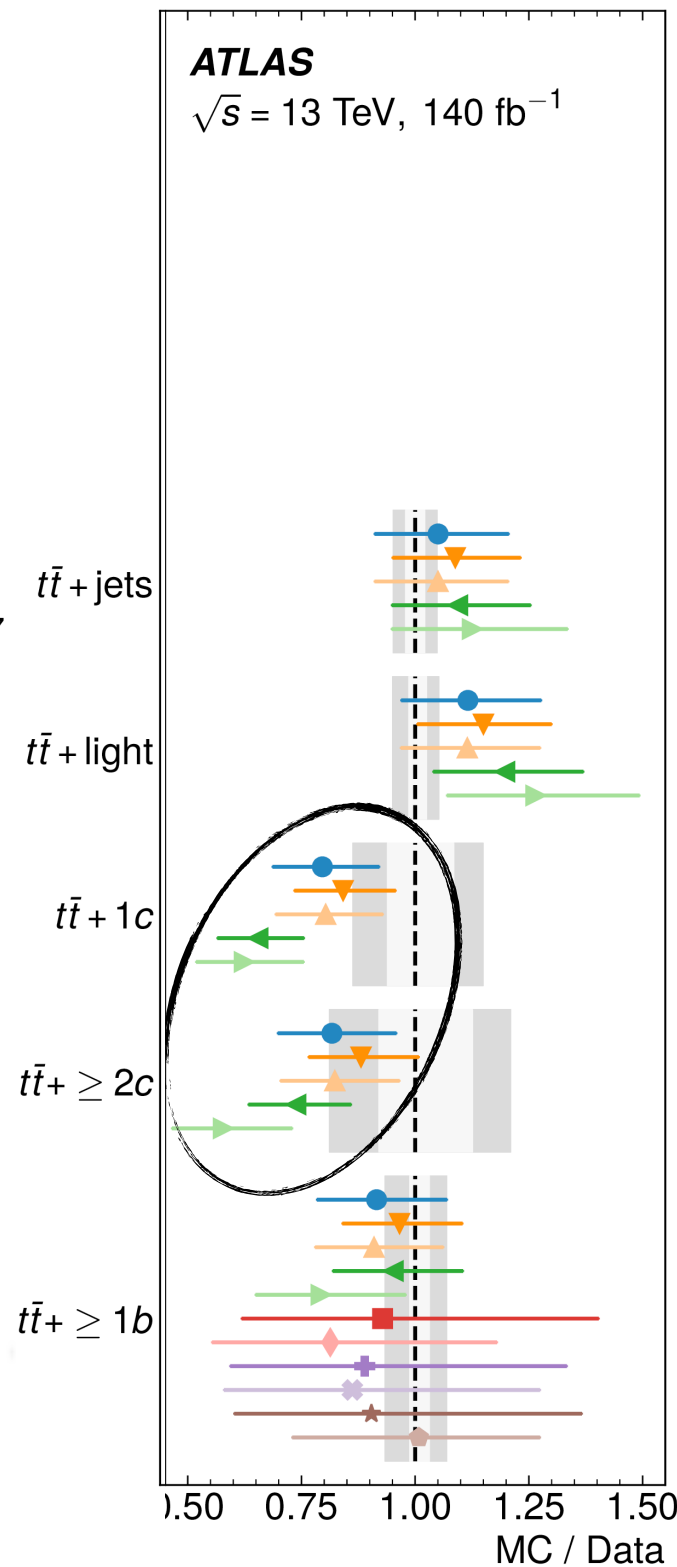


How well do we know the $t\bar{t}$ +jets background?

- For $t\bar{t}c\bar{c}$:
 - Only inclusive measurements available
 - General trend consistent
 - Also as seen in $t\bar{t}H(H \rightarrow b\bar{b})$ measurements
 - $\sim \mu = 1.1 - 1.5$, varying with selections



- Data total/stat. uncertainty
- $t\bar{t}$ Powheg+Pythia8
- $t\bar{t}$ Powheg+Pythia8 $h_{\text{damp}} = 3m_t$
- $t\bar{t}$ Powheg+Pythia8 $p_T^{\text{hard}} = 1$
- $t\bar{t}$ Powheg+Herwig7
- $t\bar{t}$ MadGraph5_aMC@NLO+Herwig7
- $t\bar{t} + b\bar{b}$ Powheg+Pythia8
- $t\bar{t} + b\bar{b}$ Powheg+Pythia8 $p_T^{\text{hard}} = 1$
- $t\bar{t} + b\bar{b}$ Powheg+Pythia8 $h_{\text{bzd}} = 2$
- $t\bar{t} + b\bar{b}$ Powheg+Pythia8 dipole recoil
- $t\bar{t} + b\bar{b}$ Powheg+Herwig7
- $t\bar{t} + b\bar{b}$ Sherpa 2.2.10



[Phys. Lett. B 860 \(2025\) 139177](#)

[Eur. Phys. J. C 85 \(2025\) 210](#)

Normalisation factor	$t\bar{t} + \text{light}$	$t\bar{t} + \geq 1c$	$t\bar{t} + 1b$	$t\bar{t} + 1B$	$t\bar{t} + \geq 2b$
Single-lepton	$0.78^{+0.08}_{-0.08}$	$1.51^{+0.19}_{-0.18}$	$1.06^{+0.10}_{-0.10}$	$1.15^{+0.15}_{-0.14}$	$0.94^{+0.08}_{-0.08}$
Dilepton	$0.88^{+0.11}_{-0.10}$	$1.36^{+0.10}_{-0.10}$	$1.24^{+0.09}_{-0.09}$		

Beware the different nomenclature...

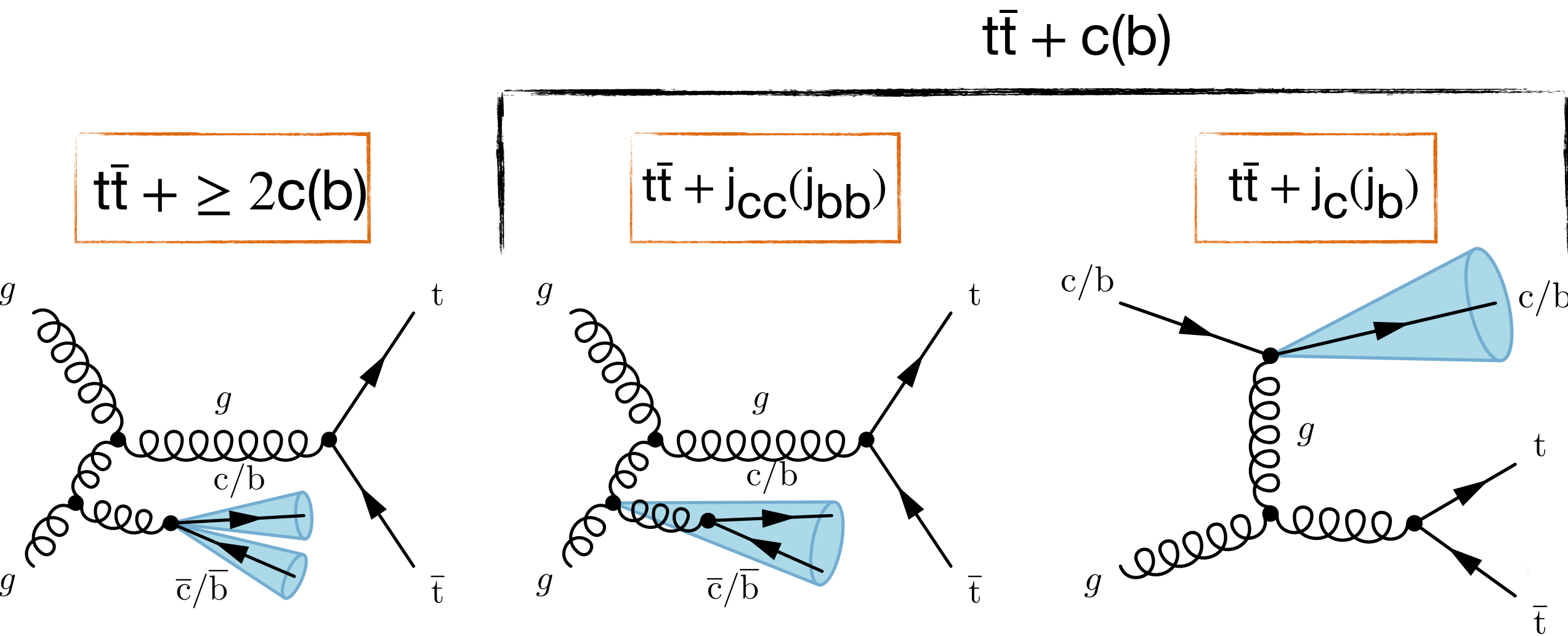
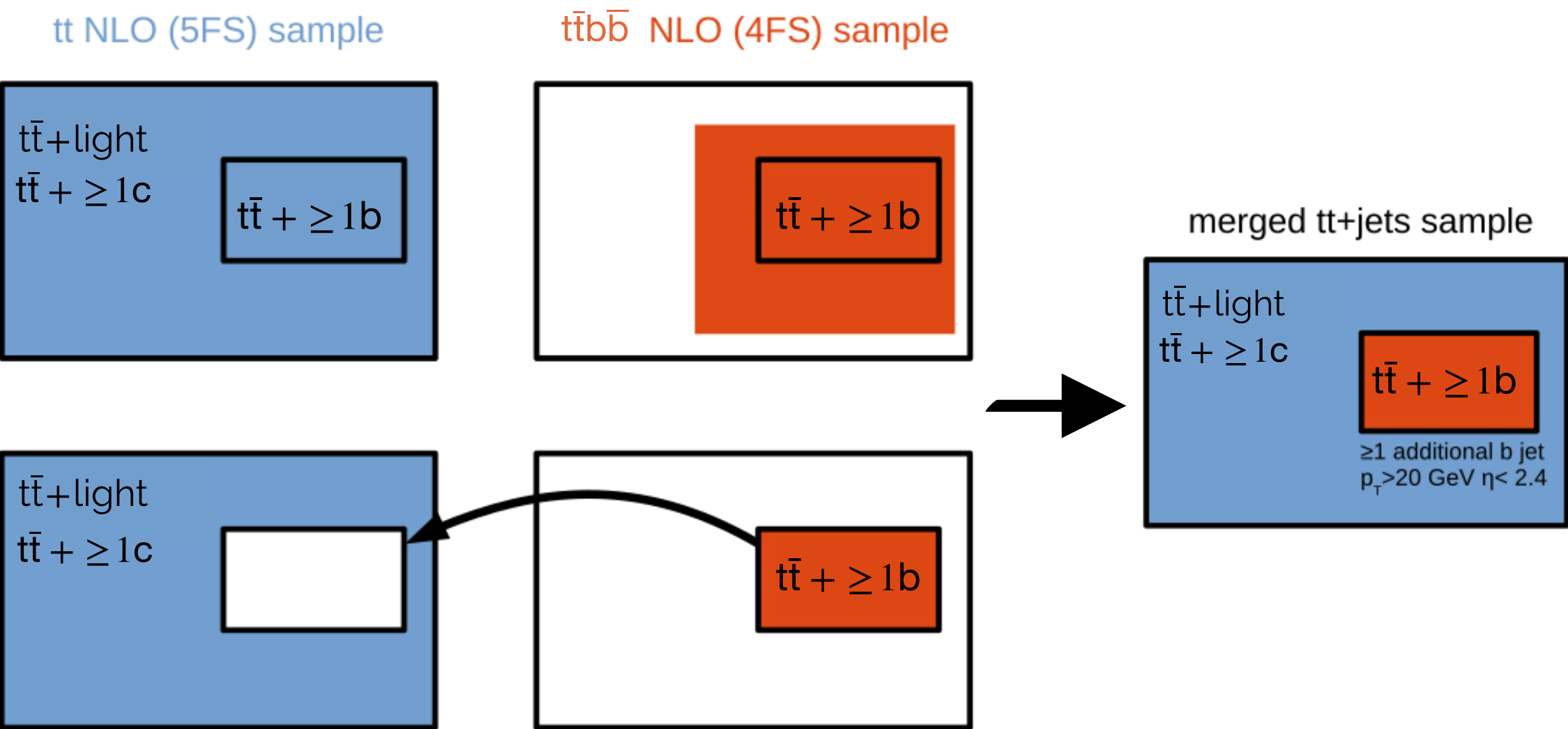
Future differential measurements particularly interesting!

Background modeling & estimation



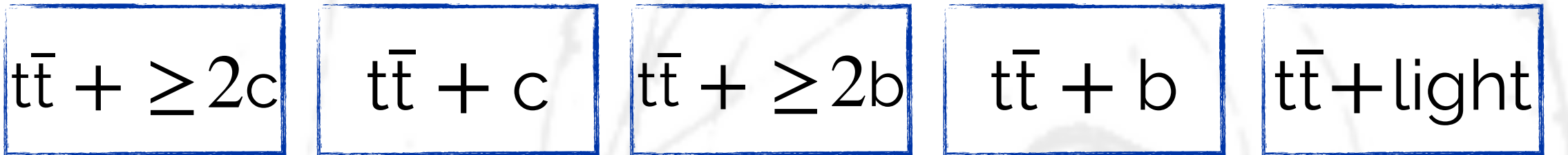
$t\bar{t}$ +jets background model:

- **State-of-the-art merging of 5FS and 4FS predictions**
 - $t\bar{t} + \geq 2b, t\bar{t} + j_b, t\bar{t} + j_{bb}$ from 4FS $t\bar{t}b\bar{b}$
 - Additional double-parton-scattering $t\bar{t}b\bar{b}$ component, missing in merging, considered for the first time
 - $t\bar{t} + \geq 2c, t\bar{t} + j_c, t\bar{t} + j_{cc}, t\bar{t} + \text{light}$ from 5FS $t\bar{t}$ +jets



Data-driven estimation:

- One control region (CR) for each contribution



Yielding a fit model with sufficient flexibility

*For $t\bar{t} + \text{jets}$: 184 nuisance parameters in fit
+ 10 free-floating parameters*

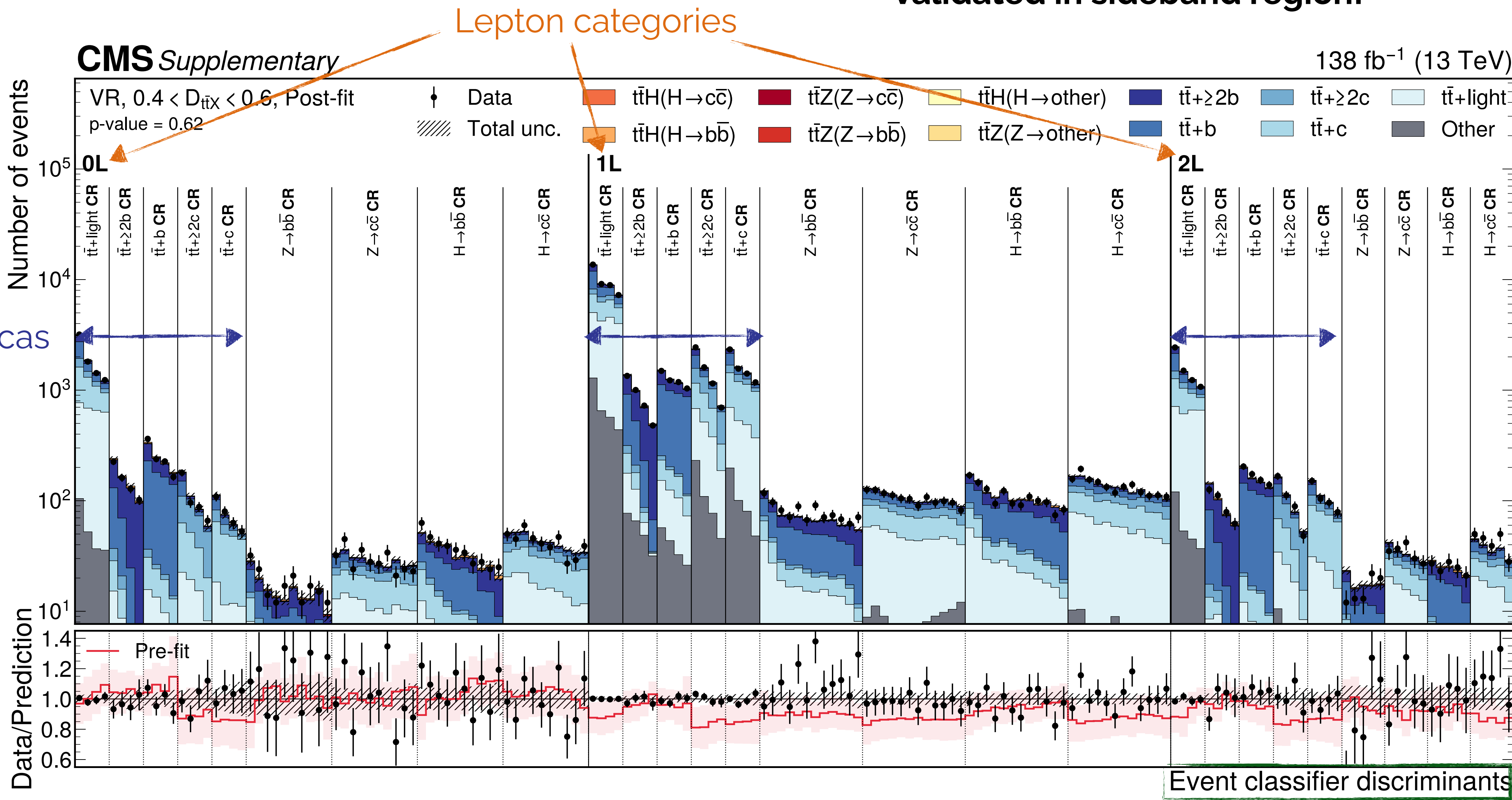
Validating the background model



→ Background model successfully validated in sideband region!

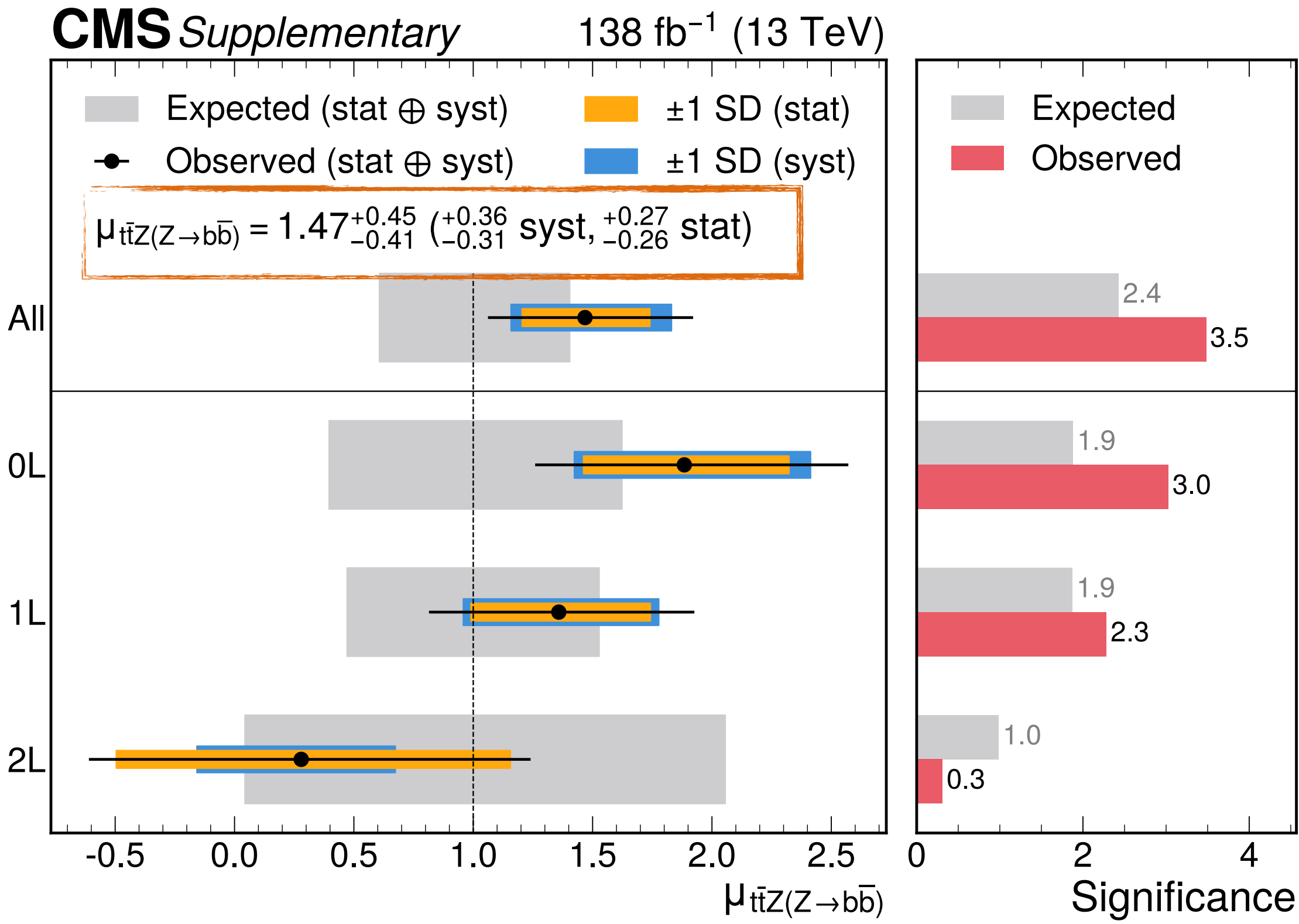
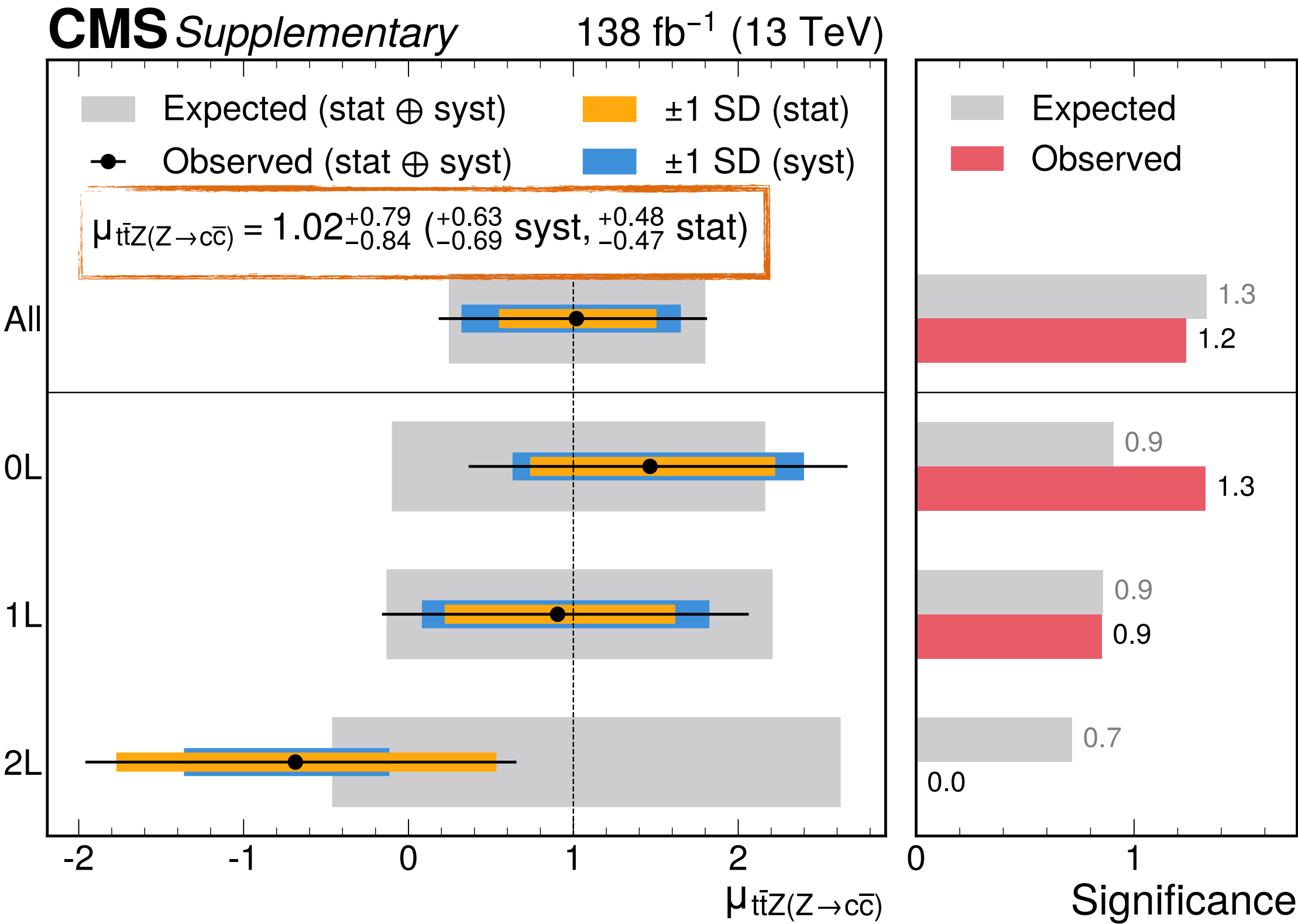


CR replicas



- Note: For robustness and more granular binning: merge e/ μ events in 1L/2L channels

Validation with $t\bar{t}Z(Z \rightarrow c\bar{c}/b\bar{b})$ production



- $t\bar{t}Z$ signal strengths SM & measurement compatible
- Good validation for $t\bar{t}H(H \rightarrow c\bar{c}/b\bar{b})$ analysis!

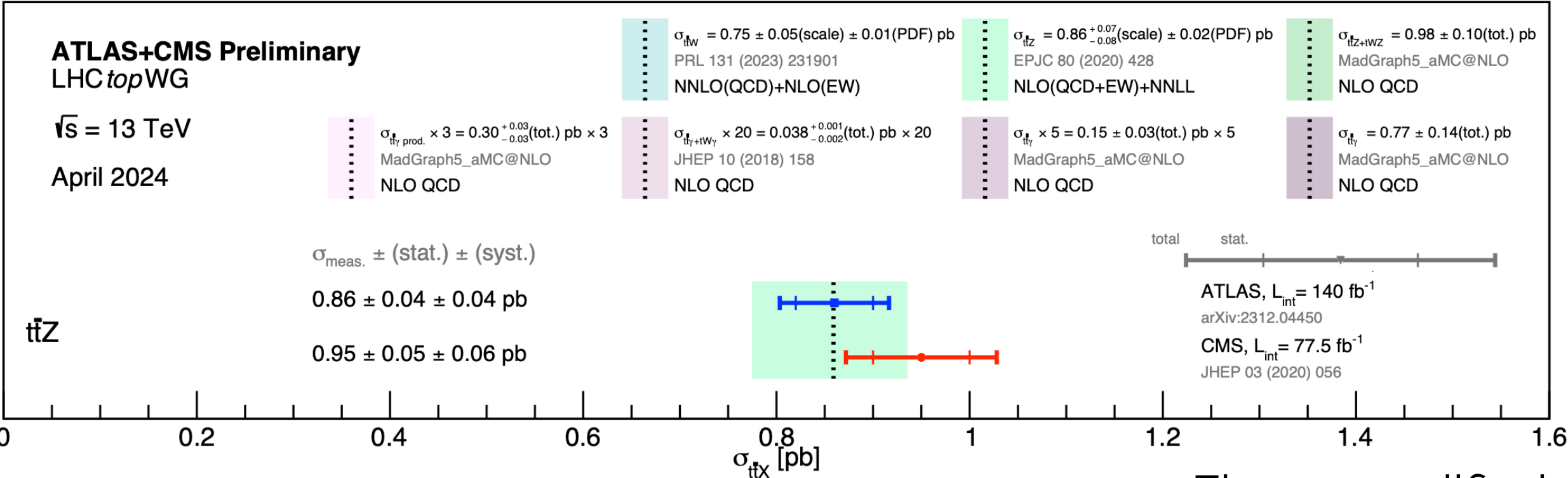
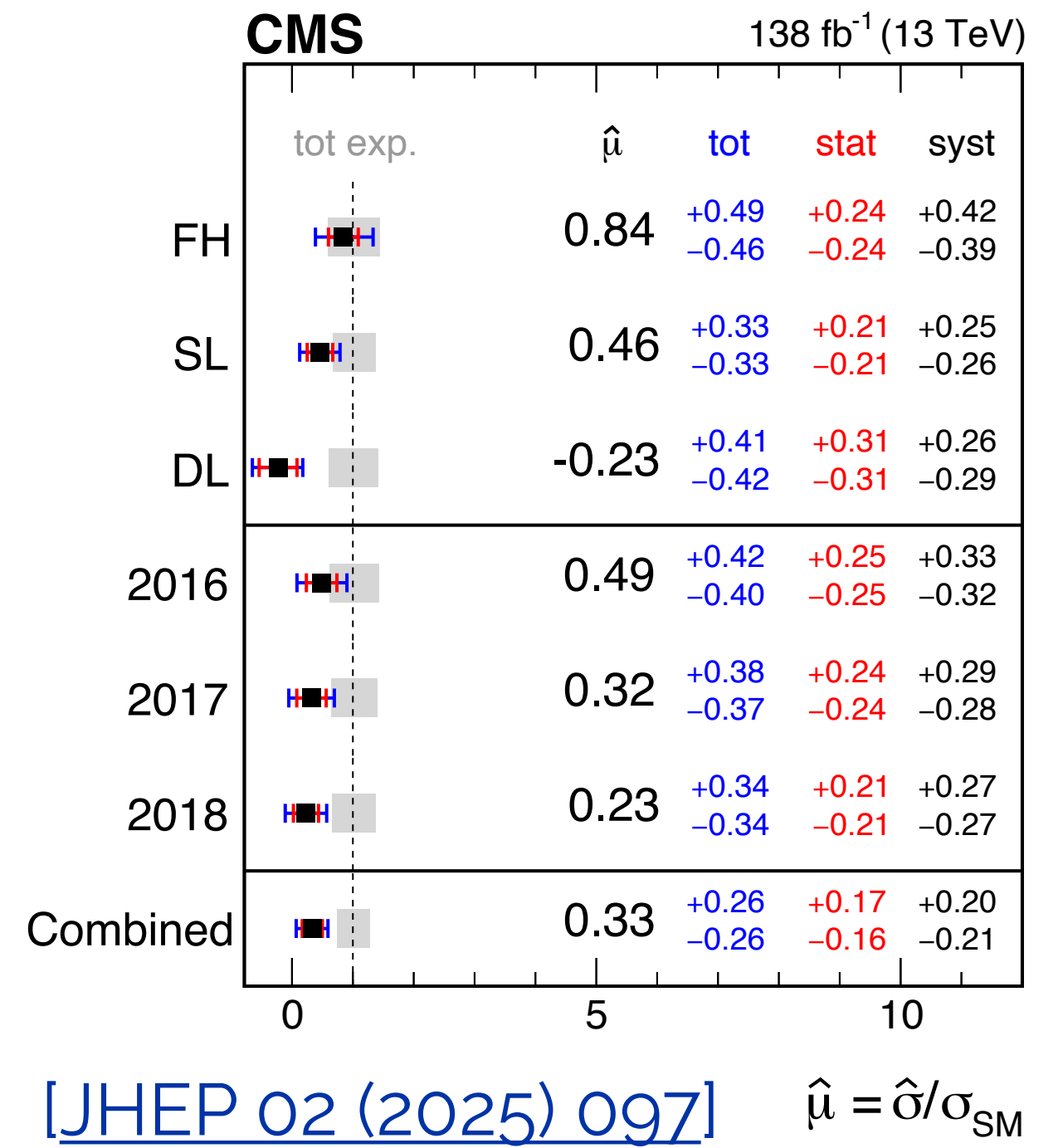
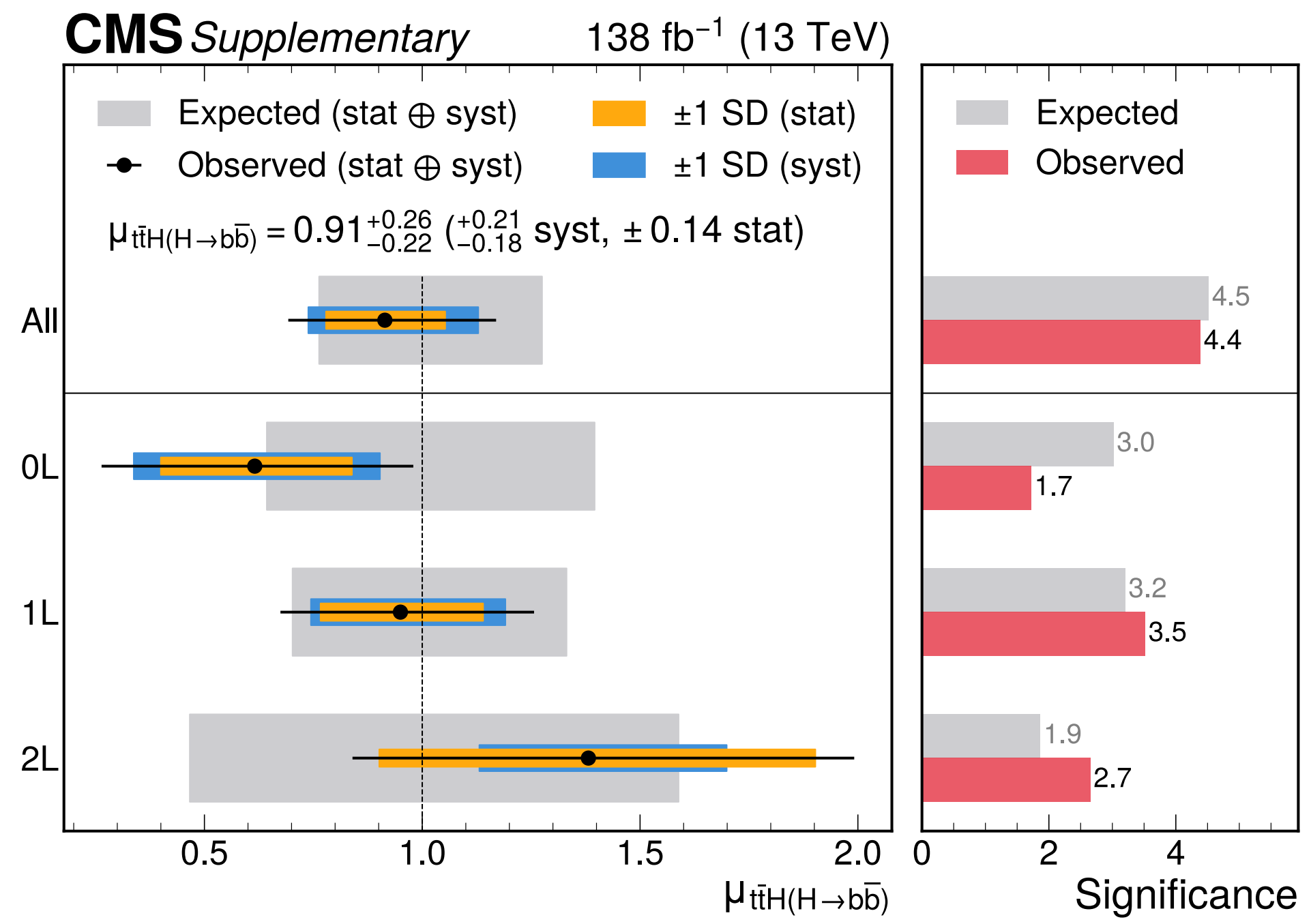
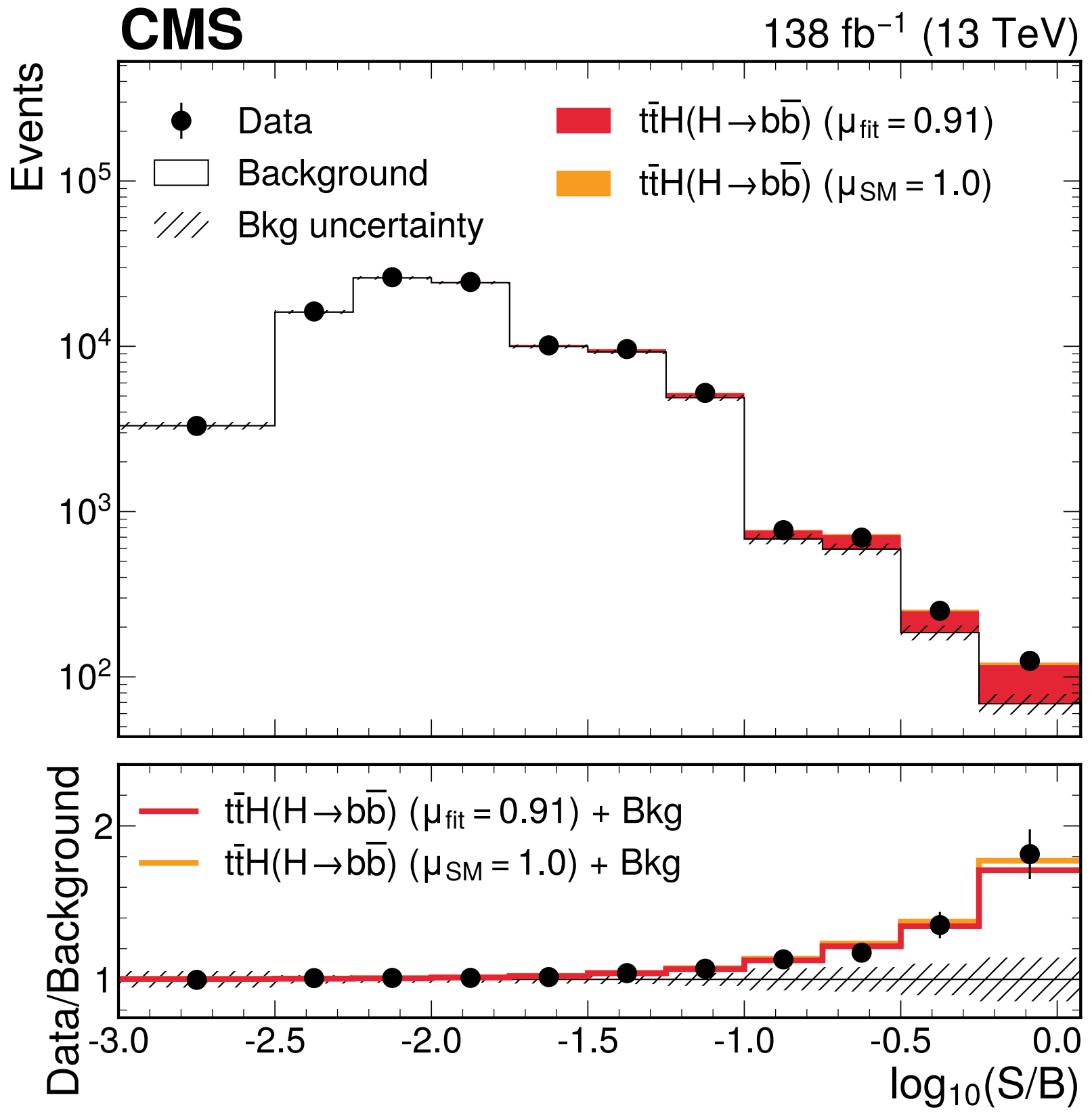


Figure modified

Results: $t\bar{t}H(H \rightarrow b\bar{b})$ production



- $\mu_{t\bar{t}H(H \rightarrow b\bar{b})} = 0.91^{+0.26}_{-0.22} - 4.4\sigma$ observed
 - Good agreement with SM expectation

- Compared to previous results:
 - More sophisticated & granular background model
 - More inclusive selection
 - Better MVA

Uncertainty breakdown



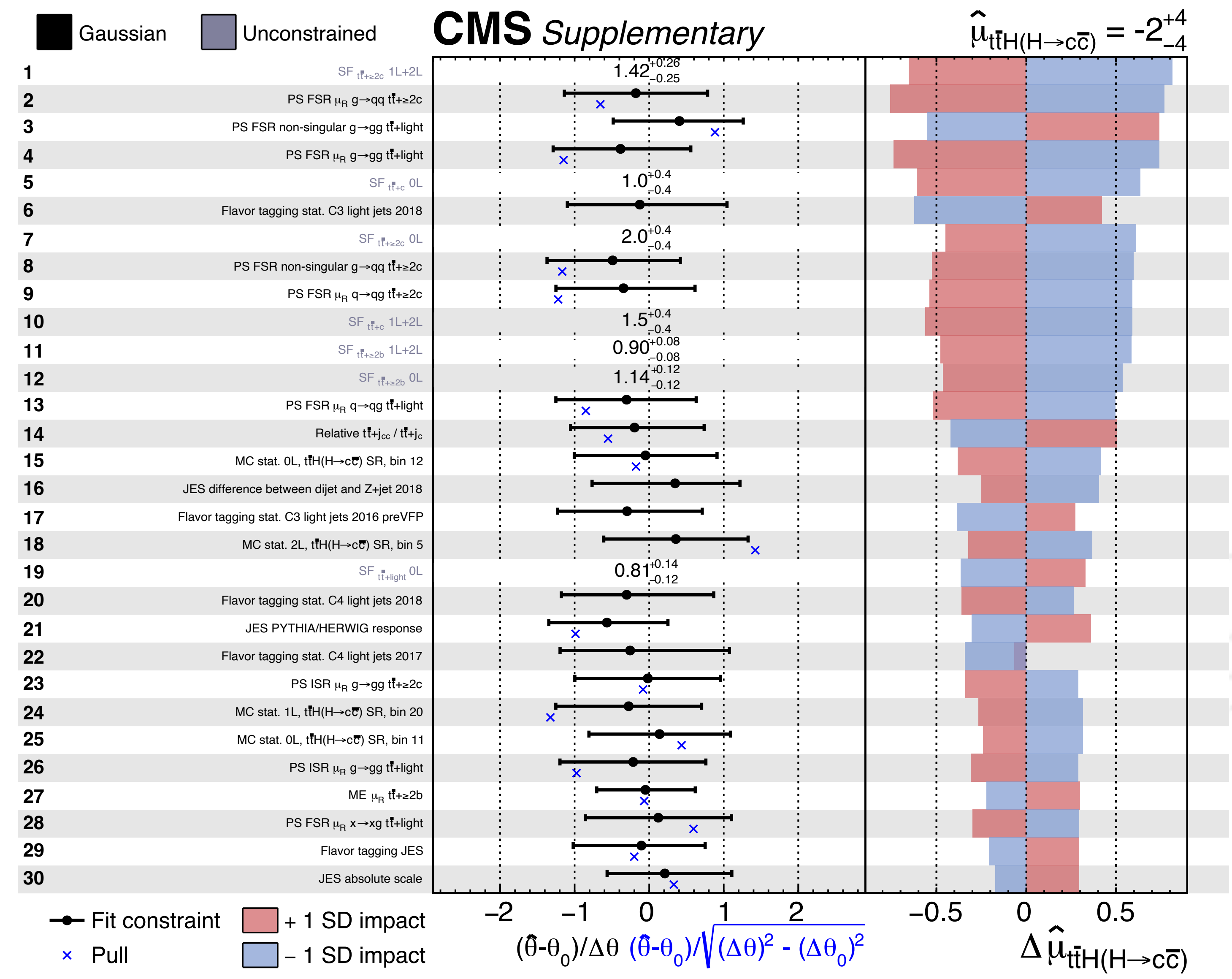
- Large statistical contribution:
 - Dominant for $ttH(cc)$ and significant for $ttH(bb)$
- Tagging efficiency as main experimental uncertainty
 - Largest single contribution for $ttH(cc)$
- Modeling uncertainties:
 - Control over tt +jets backgrounds
 - $tt+bb$ and $tt+cc$
 - Modeling of the ttH signal itself
- $ttH(bb)$ and $ttH(cc)$ correlation is only 1%

Uncertainty source	$\Delta\mu$ ($\Delta\mu / \Delta\mu_{\text{tot}}$)			
	$\mu_{t\bar{t}H(H\rightarrow c\bar{c})}$		$\mu_{t\bar{t}H(H\rightarrow b\bar{b})}$	
Statistical	3.3	(74%)	0.14	(57%)
$t\bar{t}$ +jets normalizations	1.4	(32%)	0.06	(26%)
$t\bar{t}Z$ normalizations	0.4	(8.4%)	0.06	(30%)
Theory	2.1	(47%)	0.18	(75%)
Signal	0.7	(15%)	0.11	(47%)
$t\bar{t}+\geq 1b$	0.7	(15%)	0.14	(60%)
$t\bar{t}+\geq 1c$	1.4	(32%)	0.01	(5.8%)
$t\bar{t}$ +light	1.3	(29%)	0.01	(5.2%)
Minor backgrounds	0.2	(4.6%)	0.01	(4.6%)
Experimental	2.0	(47%)	0.07	(31%)
Jet flavor tagging	1.7	(39%)	0.07	(28%)
Size of the simulated samples	1.1	(24%)	0.05	(21%)
Jet energy scale and resolution	0.8	(18%)	0.02	(8.6%)
Lepton identification	0.3	(6.0%)	0.02	(6.3%)
Integrated luminosity	0.1	(2.0%)	0.02	(6.2%)
Total	4.5	(100%)	0.24	(100%)

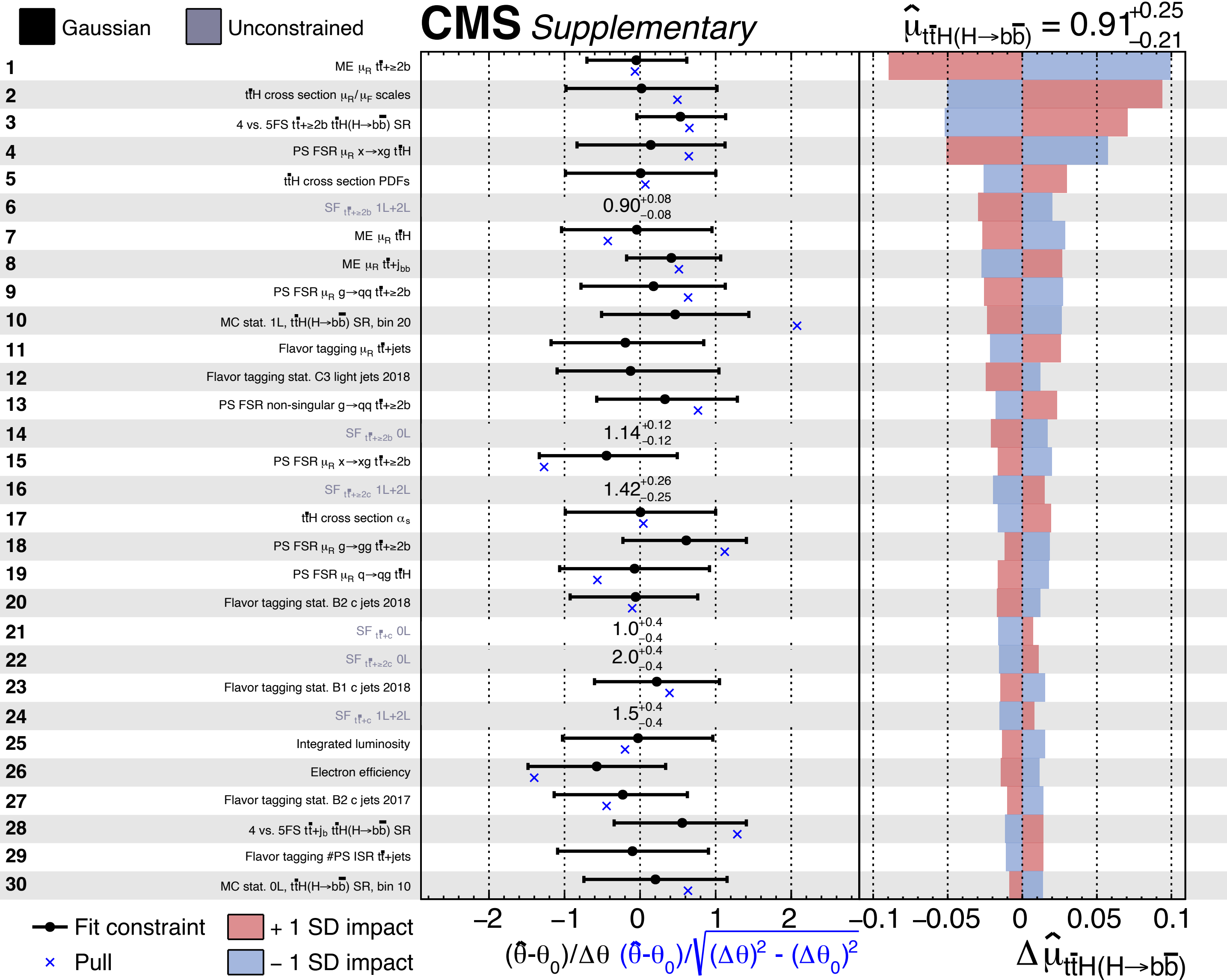
Considered uncertainties and correlations

	Source	Corr. (period)	Corr. (process)	Processes
Experimental	Integrated luminosity	~	✓	all
	Electron reconstruction and identification	✓	✓	all
	Muon reconstruction and identification	✓	✓	all
	Trigger efficiencies	×	✓	all
	Pileup reweighting	✓	✓	all
	Jet energy scale (11 sources)	~	✓	all
	Jet energy resolution	×	✓	all
	Unclustered p_T^{miss}	×	✓	all
	Jet flavor tagging	~	✓	all
Modeling	Incl. cross sections & normalization	✓	~	all
	μ_R scale	✓	×	$t\bar{t}$ +jets, $t\bar{t}X$, single t
	μ_F scale	✓	×	$t\bar{t}$ +jets, $t\bar{t}X$, single t
	PDF shape	✓	×	$t\bar{t}$ +jets, $t\bar{t}X$
	PS scales: ISR	✓	×	$t\bar{t}$ +jets, $t\bar{t}X$, single t
	PS scales: FSR	✓	×	$t\bar{t}$ +jets, $t\bar{t}X$, single t
	ME-PS matching (h_{damp})	✓	×	$t\bar{t}$ +jets
	ME (4 vs. 5 flavor scheme)	✓	×	$t\bar{t} + \geq 1b$

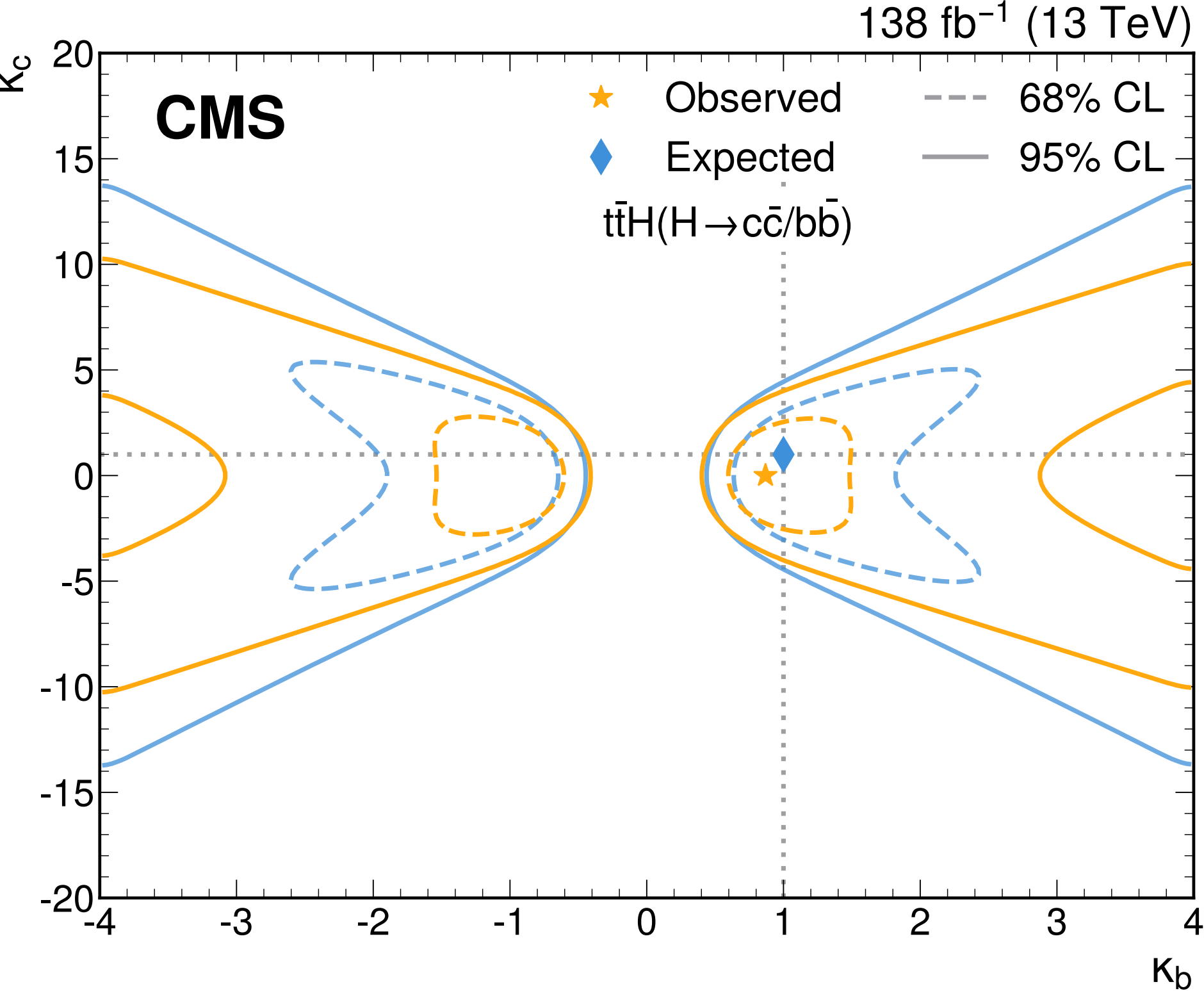
Nuisance parameter impacts



Nuisance parameter impacts



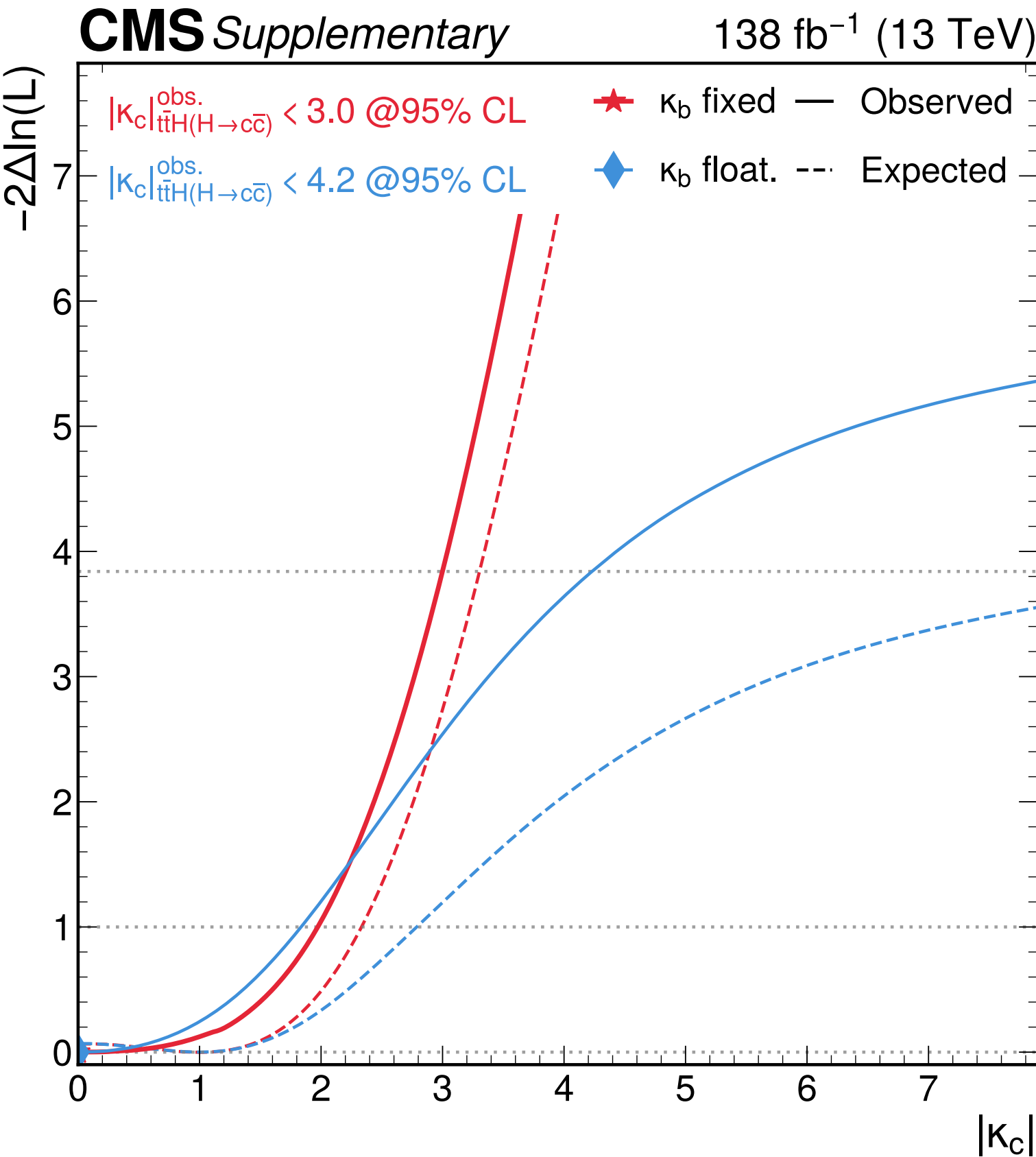
Results: coupling interpretation



All other kappas/couplings fixed to SM ($k=1$)

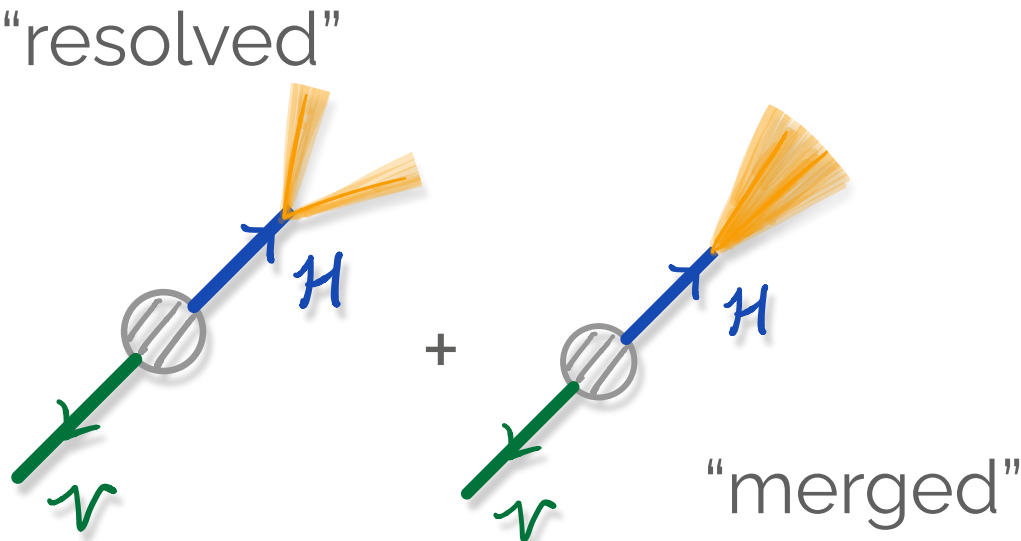
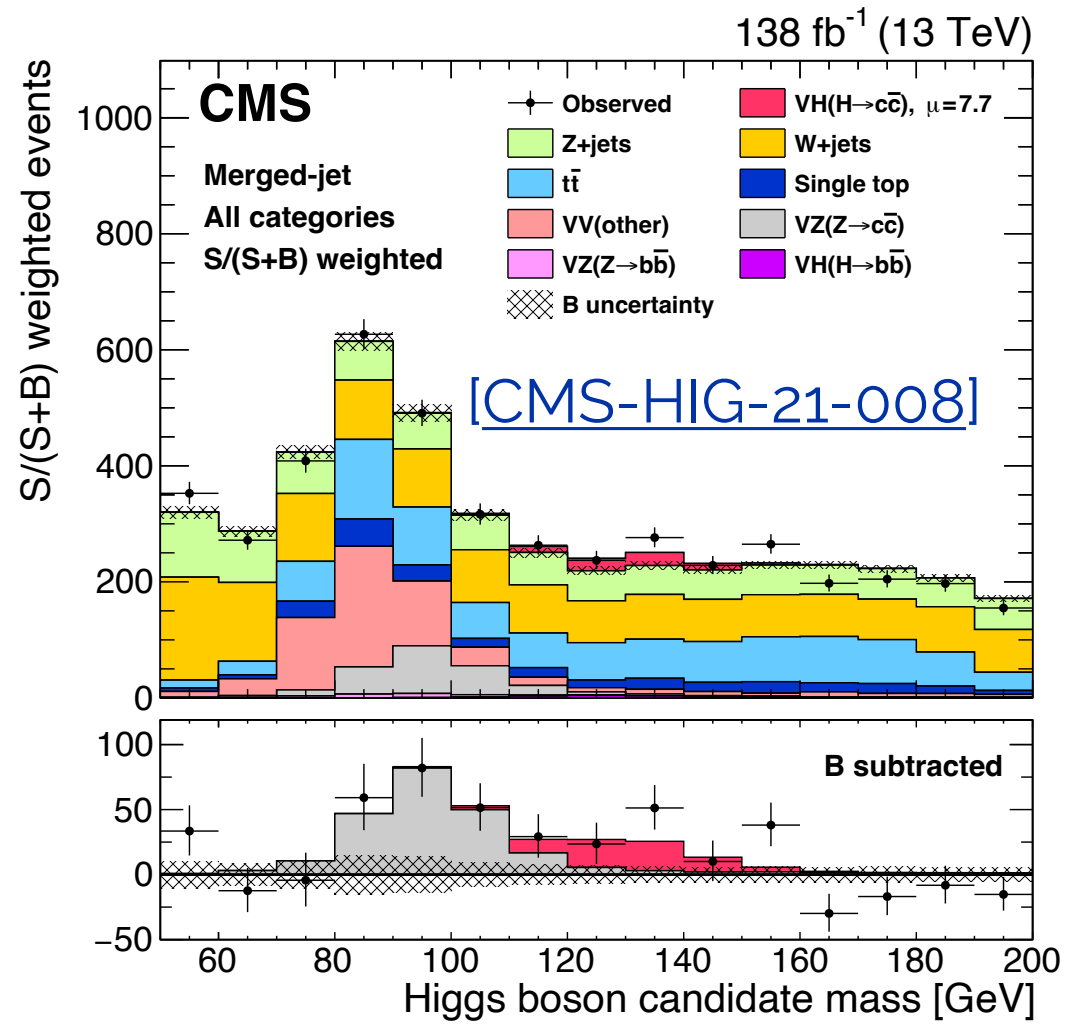
$$\mathcal{B}(H \rightarrow c\bar{c}) = \frac{\kappa_c^2 \cdot \mathcal{B}_{SM}^{H \rightarrow c\bar{c}}}{1 + (\kappa_c^2 - 1) \cdot \mathcal{B}_{SM}^{H \rightarrow c\bar{c}} + (\kappa_b^2 - 1) \cdot \mathcal{B}_{SM}^{H \rightarrow b\bar{b}}},$$

$$\mathcal{B}(H \rightarrow b\bar{b}) = \frac{\kappa_b^2 \cdot \mathcal{B}_{SM}^{H \rightarrow b\bar{b}}}{1 + (\kappa_c^2 - 1) \cdot \mathcal{B}_{SM}^{H \rightarrow c\bar{c}} + (\kappa_b^2 - 1) \cdot \mathcal{B}_{SM}^{H \rightarrow b\bar{b}}}.$$



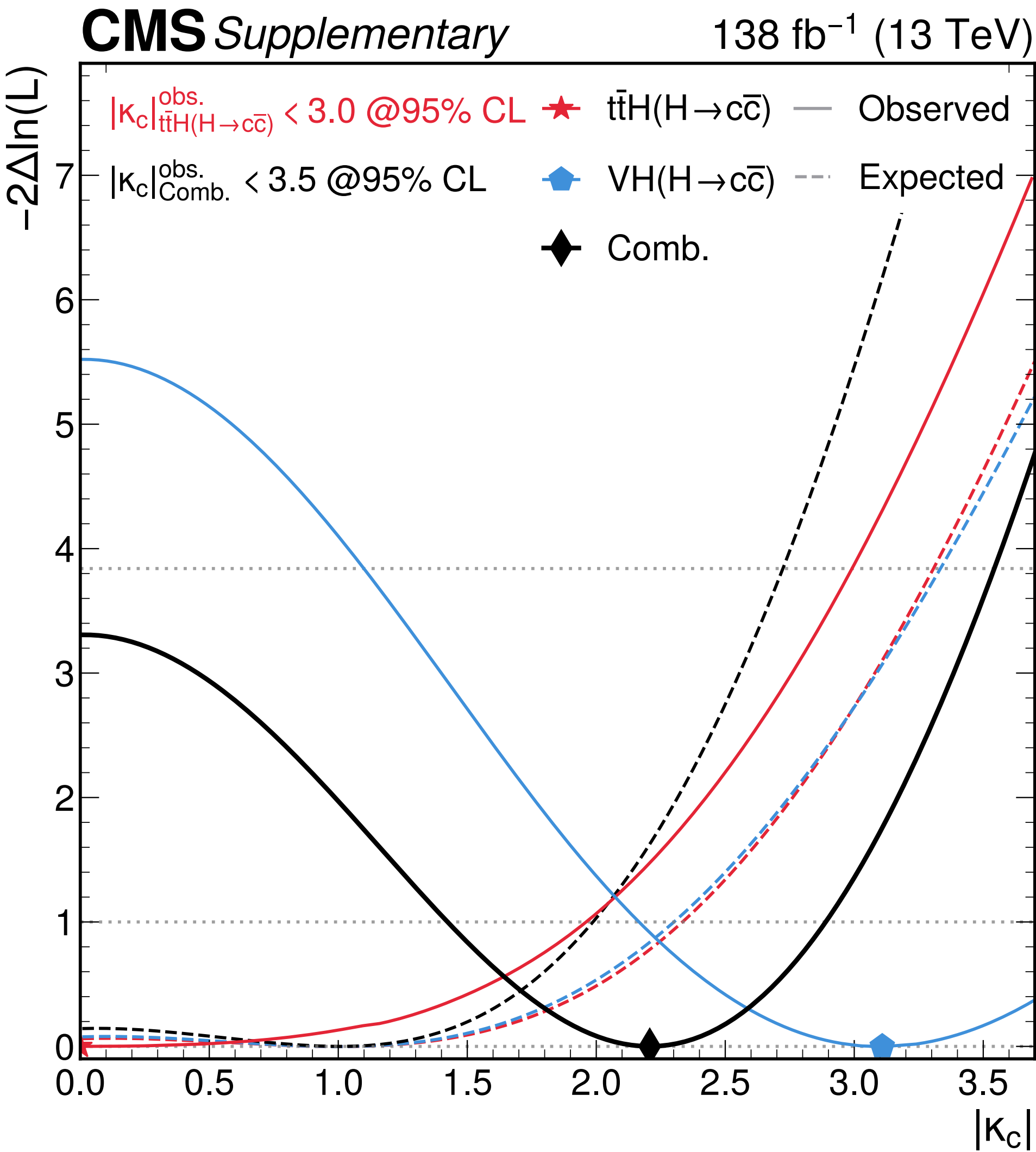
- **1D constraints on κ_c : $|\kappa_c| \approx 3.0$ obs. (3.3 exp.)**
- **Most stringent constraints to date!** κ_b fixed to 1
 - Even better than VH(cc) results (4.1 (ATLAS) and 3.4 (CMS) exp.)

Combination with $VH(H \rightarrow c\bar{c})$



- Combine with $VH(H \rightarrow c\bar{c})$:
 - Full Run 2 result from resolved+merged topologies
 - Correlated experimental uncertainties
 - Uncorrelated tagging uncertainties
 - Correlated theor. uncertainties (cross section / BR)
- $VH(H \rightarrow c\bar{c})$ results: $1.1 < |\kappa_c| < 5.5$ (< 3.4 exp.)
- $t\bar{t}H$ results: $|\kappa_c| < 3.0$ (< 3.3 exp.)
- Combined: $|\kappa_c| < 3.5$ (< 2.7 exp.)

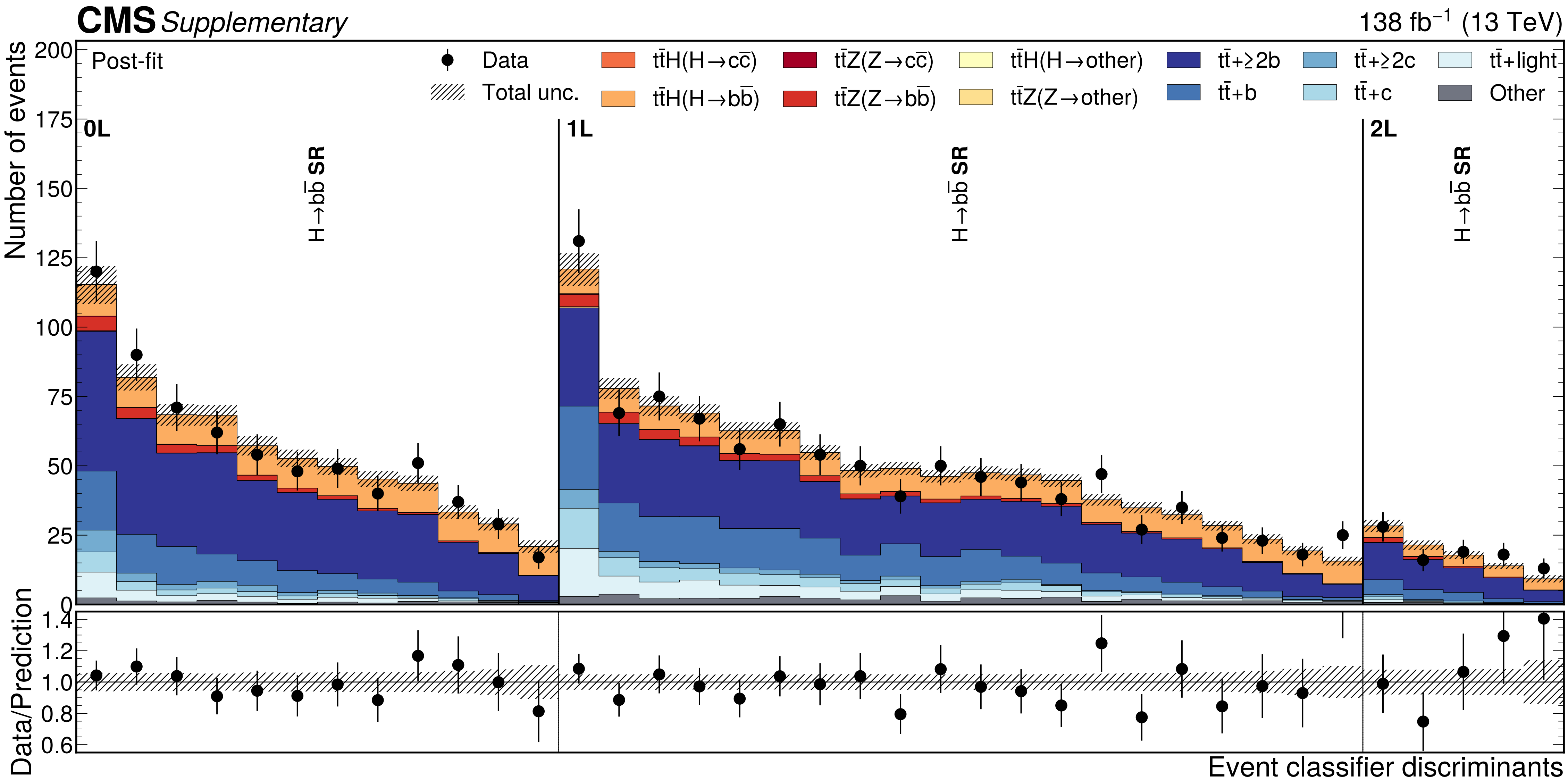
20% improvement!



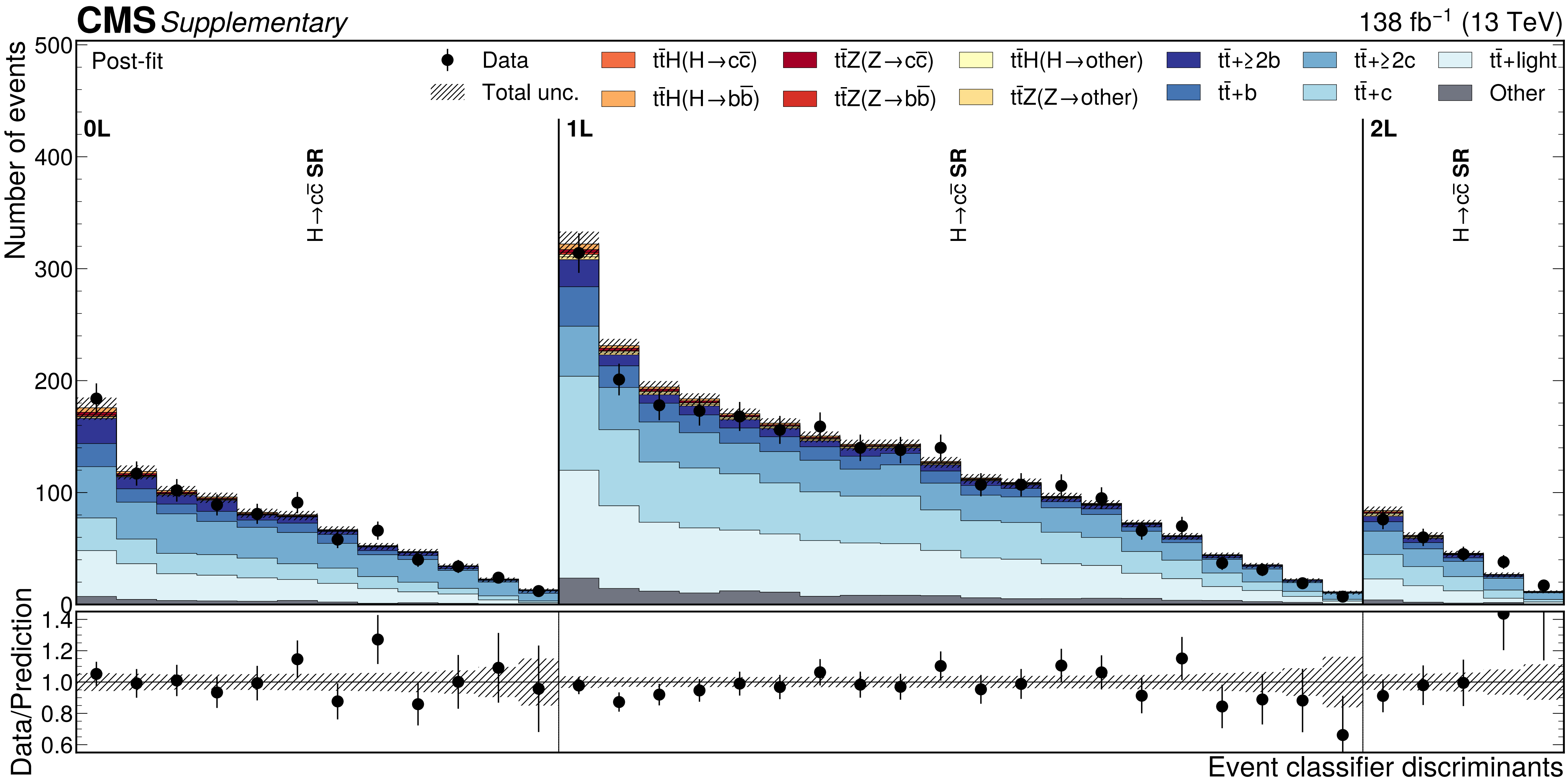
Objects and event selection

Selection criteria	0L channel	1L channel	2L channel
Number of leptons	0	1	2
Sign and flavor of leptons	—	$e^{\pm}, \mu^{\pm}$	$e^+e^-, \mu^{\pm}e^{\mp}, \mu^+\mu^-$
Min. p_T of p_T -leading electron [GeV]	—	29/30/30	25
Electron identification	—	Tight (80% efficiency)	Loose (90% efficiency)
Electron isolation	—	Tight (80% efficiency)	Loose (90% efficiency)
Min. p_T of p_T -leading muon [GeV]	—	26/29/26	25
Max. muon relative isolation	—	0.15	0.25
Min. $m_{ee/\mu\mu}$ [GeV]	—	—	20
Range of $m_{ee/\mu\mu}$ [GeV]	—	—	$<76, >106$
Min. p_T of jets [GeV]	25	25	25
Min. p_T of 6 th jet [GeV]	40	—	—
Max. $ \eta $ of jets	2.4	2.4	2.4
Min. number of jets	7	5	4
Min. number of b-tagged jets	1	1	1
Min. number of HF (b- or c-tagged) jets	3	3	3
Min. H_T [GeV]	500	—	—
Min. p_T^{miss} [GeV]	—	20	20

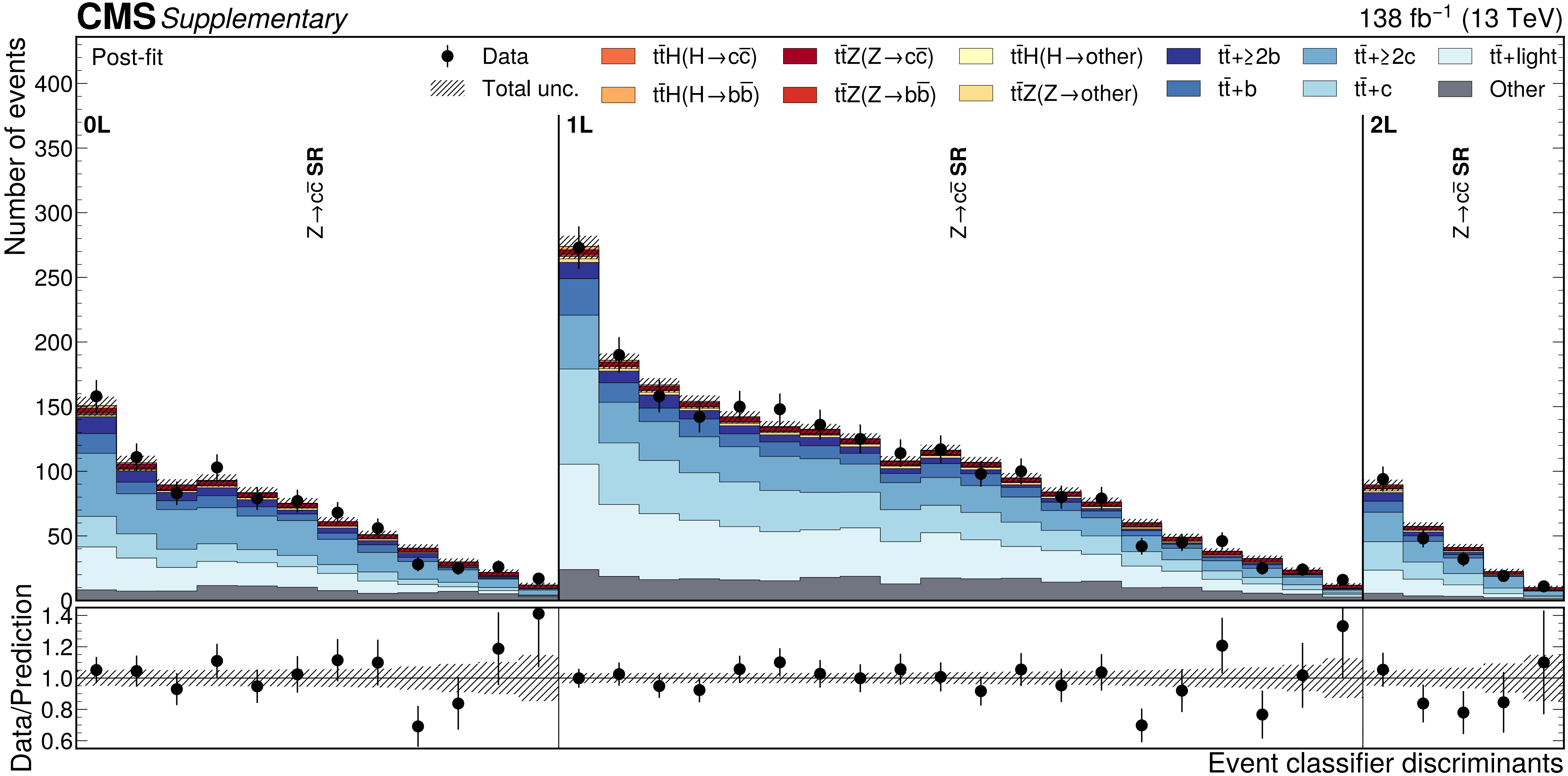
More post-fit plots



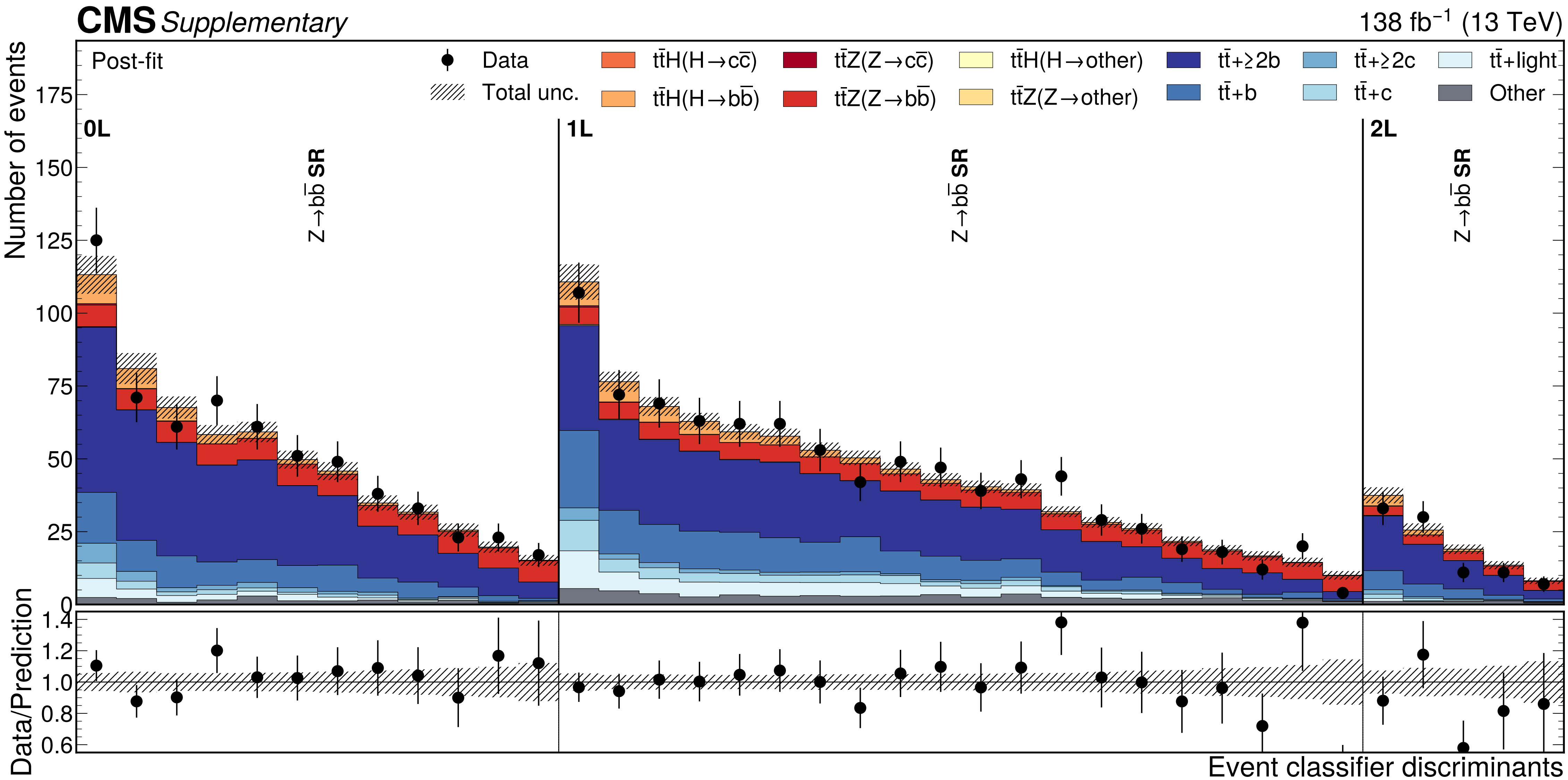
More post-fit plots



More post-fit plots



More post-fit plots





- Background normalisation SFs compatible with:
 - Dedicated measurements
 - Previous CMS and ATLAS ttH analyses

Cat.	$t\bar{t} + \geq 2b$	$t\bar{t} + b$	$t\bar{t} + \geq 2c$	$t\bar{t} + c$	$t\bar{t} + \text{light}$
2L/1L	0.88 ± 0.07	0.92 ± 0.16	1.39 ± 0.23	1.41 ± 0.36	0.89 ± 0.11
0L	1.12 ± 0.11	0.89 ± 0.18	2.03 ± 0.41	0.95 ± 0.34	0.89 ± 0.17

$t\bar{t} + \geq 1b$ bkg. model	$\mu_{t\bar{t}H(H \rightarrow c\bar{c})}$	(95% CL)	$\mu_{t\bar{t}H(H \rightarrow b\bar{b})}$	(Sign. [s.d.])
Default	-1.6 ± 4.5	(<7.8)	$0.91^{+0.26}_{-0.22}$	(4.4)
Simple PS	-0.4 ± 4.4	(<8.8)	$0.91^{+0.27}_{-0.22}$	(4.4)
$\mu_F \times 2$	$-1.7^{+4.2}_{-4.3}$	(<8.0)	$0.94^{+0.26}_{-0.23}$	(4.4)
$\mu_F \times 2$ & Simple PS	-1.0 ± 4.4	(<8.5)	$0.97^{+0.28}_{-0.23}$	(4.5)
5FS	-1.8 ± 4.3	(<7.7)	$0.97^{+0.24}_{-0.20}$	(5.7)
5FS & Simple PS	-0.9 ± 4.5	(<7.6)	$1.02^{+0.26}_{-0.21}$	(6.0)

- Vary the background model and repeat the fit:
 - Default model: 8 independent variations per process:
 - Scale and splitting-kernel variations (x2) [Phys. Rev. D 94, 074005 \(2016\)](#)
 - Each for $g \rightarrow gg$, $g \rightarrow qq$, $q \rightarrow qg$, $x \rightarrow xg$ (where $x = t, b$) splittings (x4)
 - Varied factorization scale for $tt+bb$ 4FS (as used in recent ATLAS $ttH(bb)$ measurement)
 - Using only the 5FS tt +jets simulation

→ Results robust wrt. all variations!

t $\bar{t}$ bb in t $\bar{t}$ H(H $\rightarrow$ b $\bar{b}$) over time

	ATLAS 36 fb ⁻¹	ATLAS 139 fb ⁻¹	ATLAS 140 fb ⁻¹	CMS 36 ⁻¹ +42 ⁻¹	CMS 138 fb ⁻¹	This result
Generator	P+P8 tt 5FS scaled to Sherpa ttbb 4FS	P+P8 ttbb 4FS	P+P8 ttbb 4FS	P+P8 tt 5FS	P+P8 ttbb 4FS	P+P8 ttbb 4FS
μ_F	-	$\frac{1}{2} \cdot \sum m_T$	$\frac{1}{2} \cdot \sum m_T$	-	$\frac{1}{4} \cdot \sum m_T$	$\frac{1}{4} \cdot \sum m_T$
μ_R	-	$\prod m_T$	$\frac{1}{2} \cdot \prod m_T$	-	$\frac{1}{2} \cdot \prod m_T$	$\frac{1}{2} \cdot \prod m_T$
$\mu_{t\bar{t}H(bb)}$	0.84 ^{+0.64} _{-0.61}	0.35 ^{+0.36} _{-0.34}	0.81 ^{+0.22} _{-0.19}	1.15 ^{+0.44} _{-0.40}	0.33 ^{+0.26} _{-0.26}	0.91 ^{+0.26} _{-0.22}

$$\sum m_T = \sum_{i=t,\bar{t},b\bar{b}} m_{T,i}$$

$$\prod m_T = \sqrt[4]{\prod_{i=t,\bar{t},b\bar{b}} m_{T,i}}$$

Comparison to CMS' 138 fb⁻¹ t $\bar{t}$ H(H $\rightarrow$ b $\bar{b}$) analysis

- CMS Full Run 2 ttH(bb) analysis: [\[JHEP 02 \(2025\) 097\]](#)

- Split t $\bar{t}$ bb into:

- t $\bar{t}$ + j_{bb} / t $\bar{t}$ + j_b + t $\bar{t}$ + $\geq 2b$ / t $\bar{t}$ cc / t $\bar{t}$ + light

- Categories / network scores for:

- In 1L: t $\bar{t}$ + j_{bb} / t $\bar{t}$ + j_b + t $\bar{t}$ + $\geq 2b$ / t $\bar{t}$ cc / t $\bar{t}$ + light

- In 2L: t $\bar{t}$ bb / t $\bar{t}$ cc / t $\bar{t}$ + light

- Free normalisation for:

- t $\bar{t}$ cc / t $\bar{t}$ bb

- This analysis:

- Split t $\bar{t}$ bb into:

- t $\bar{t}$ + j_{bb} + t $\bar{t}$ + j_b / t $\bar{t}$ + $\geq 2b$ / t $\bar{t}$ + j_{cc} + t $\bar{t}$ + j_c / t $\bar{t}$ + $\geq 2c$ / t $\bar{t}$ + light

- Categories / network scores for:

- t $\bar{t}$ + j_{bb} + t $\bar{t}$ + j_b / t $\bar{t}$ + $\geq 2b$ / t $\bar{t}$ + j_{cc} + t $\bar{t}$ + j_c / t $\bar{t}$ + $\geq 2c$ / t $\bar{t}$ + light

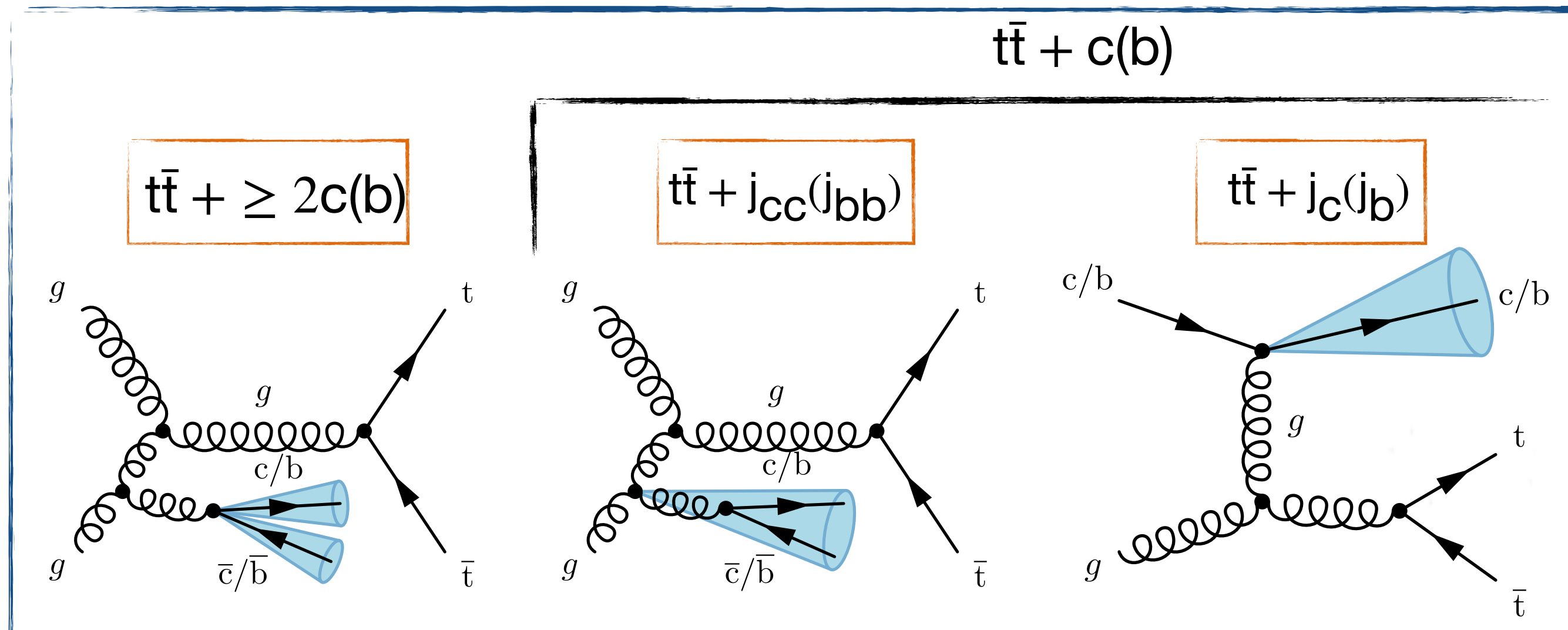
- Free normalisation for:

- t $\bar{t}$ + j_{bb} + t $\bar{t}$ + j_b / t $\bar{t}$ + $\geq 2b$ / t $\bar{t}$ + j_{cc} + t $\bar{t}$ + j_c / t $\bar{t}$ + $\geq 2c$ / t $\bar{t}$ + light

- Granular Parton-Shower uncertainty treatment

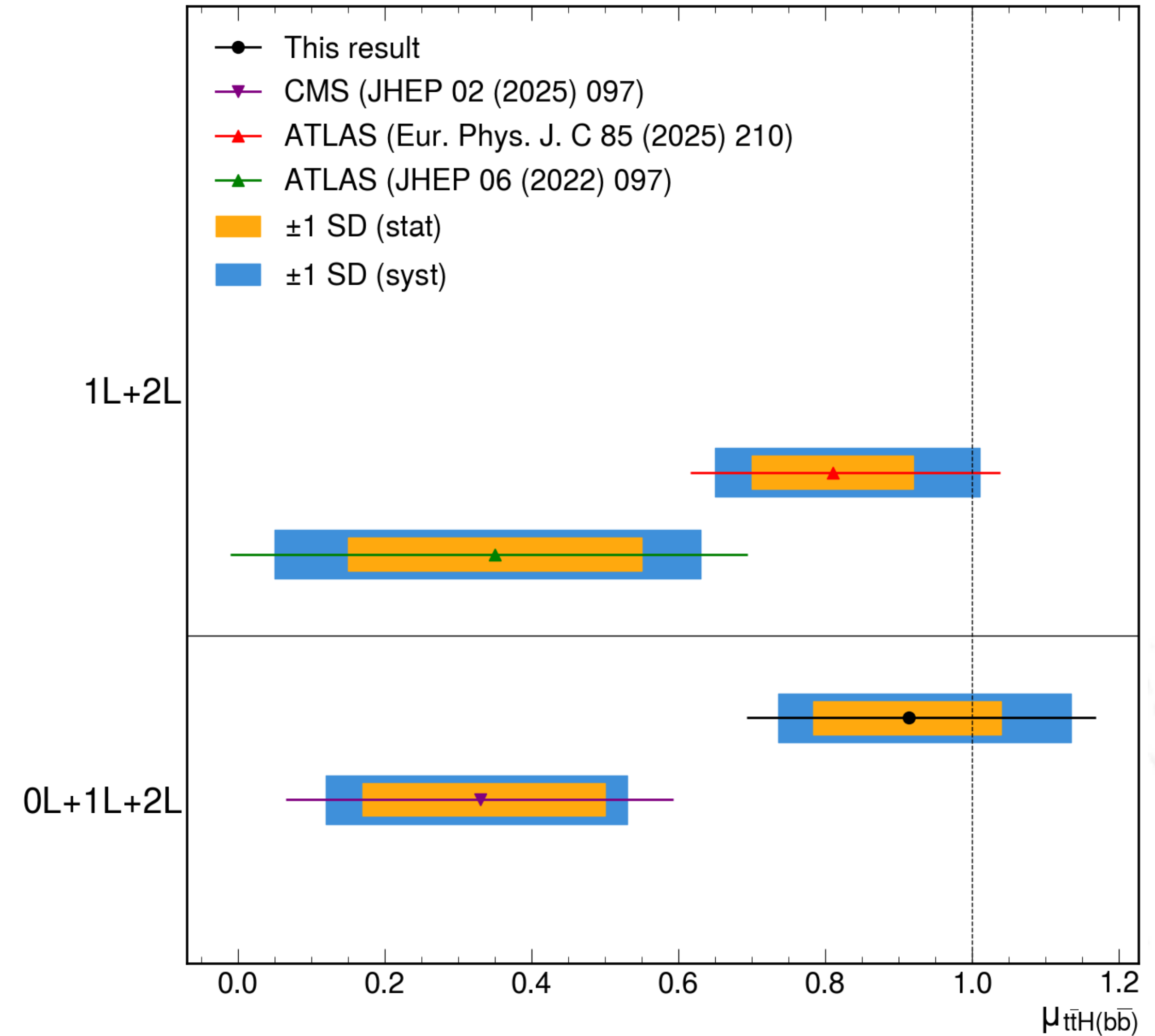
- 4FS vs. 5FS uncertainties

- Better separation (48% vs. 10% correlation with ttbb)



Comparison to CMS' 138 fb⁻¹ t $\bar{t}$ H(H $\rightarrow$ b $\bar{b}$) analysis

- Using *bootstrapping* method:
resample data events and analyze simultaneously
in both analyses to estimate their statistical
correlation
- Statistical correlation of the most sensitive
channel is found to be 13%
 - Assuming independent systematic
uncertainties, p- value that results are
compatible is 10%
 - Very different background model
 - More inclusive event selection
 - Low correlation of the neural network scores
 - For an extreme systematic correlation of 50%,
the p-value is found to be 3%

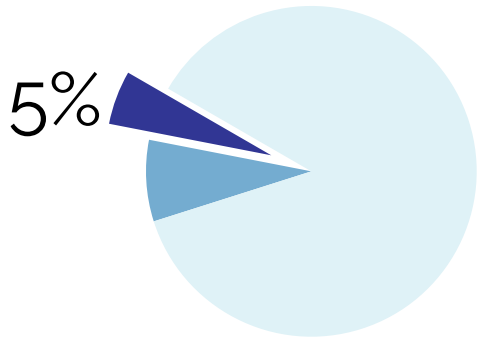


Generator choices for signal and background

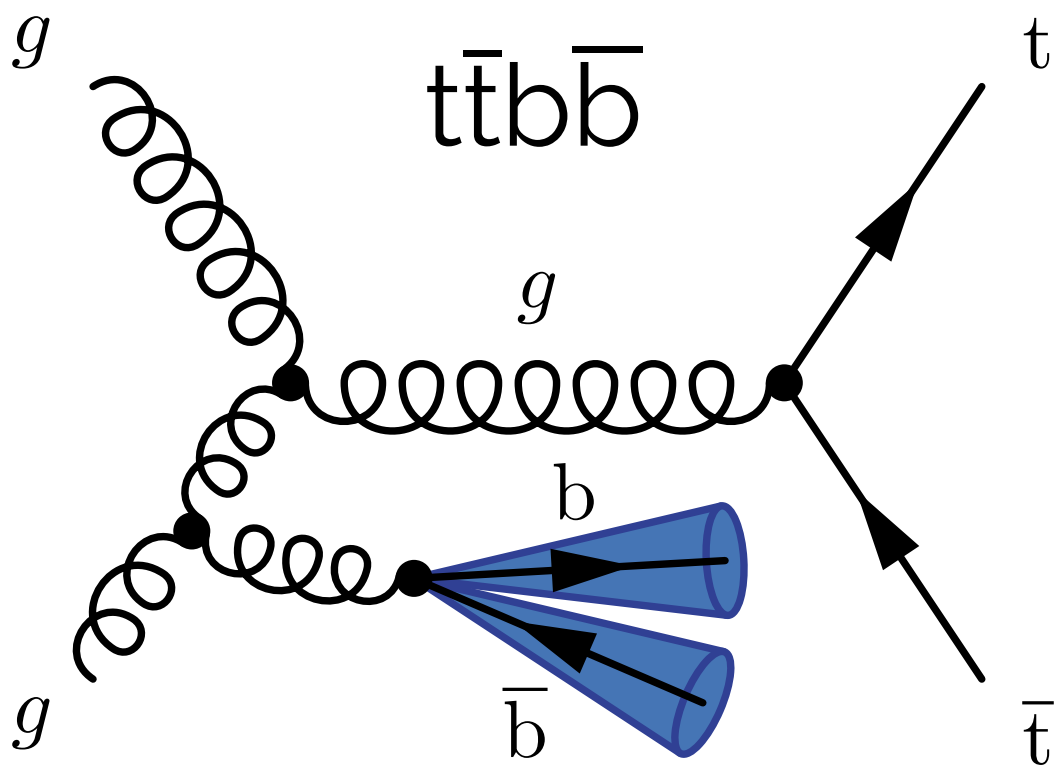
Process	Groups	ME order	Generator	Notes
$t\bar{t}H$	$t\bar{t}H(H\rightarrow c\bar{c}), t\bar{t}H(H\rightarrow b\bar{b}),$ $t\bar{t}H(H\rightarrow \text{Other})$	NLO	POWHEG v2	
$t\bar{t}Z$	$t\bar{t}Z(Z\rightarrow c\bar{c}), t\bar{t}Z(Z\rightarrow b\bar{b}),$ $t\bar{t}Z(Z\rightarrow \text{Other})$	NLO	MADGRAPH5_aMC@NLO v2.6.1	MADSPIN for heavy particle decays
$t\bar{t}$	$t\bar{t}+\text{light}, t\bar{t}+\geq 2c, t\bar{t}+c$	NLO	POWHEG v2	
$t\bar{t}b\bar{b}$	$t\bar{t}+\geq 2b, t\bar{t}+b$	NLO	POWHEG-BOX-RES	
$t\bar{t}b\bar{b}$ DPS	$t\bar{t}+\geq 2b, t\bar{t}+b$	NLO / LO	POWHEG v2 / PYTHIA v8.240	$t\bar{t}$ NLO ME with POWHEG / LO $b\bar{b}$ DPS with PYTHIA

Process	Group	ME order	Generator	Notes
tW			POWHEG v2	
t channel	single t	NLO	POWHEG v2	MADSPIN for heavy particle decays
s channel			MADGRAPH5_aMC@NLO v2.6.1	
$t\bar{t}W$	$t\bar{t}W$	NLO	MADGRAPH5_aMC@NLO v2.6.1	FxFx merging up to 1 additional jet MADSPIN for heavy particle decays
tWZ	tWZ	LO	MADGRAPH5_aMC@NLO v2.6.5	MADSPIN for heavy particle decays
QCD	QCD	LO	MADGRAPH5_aMC@NLO v2.5.1	MLM merging for ME-PS matching

How well do we know the $t\bar{t}$ +jets background?

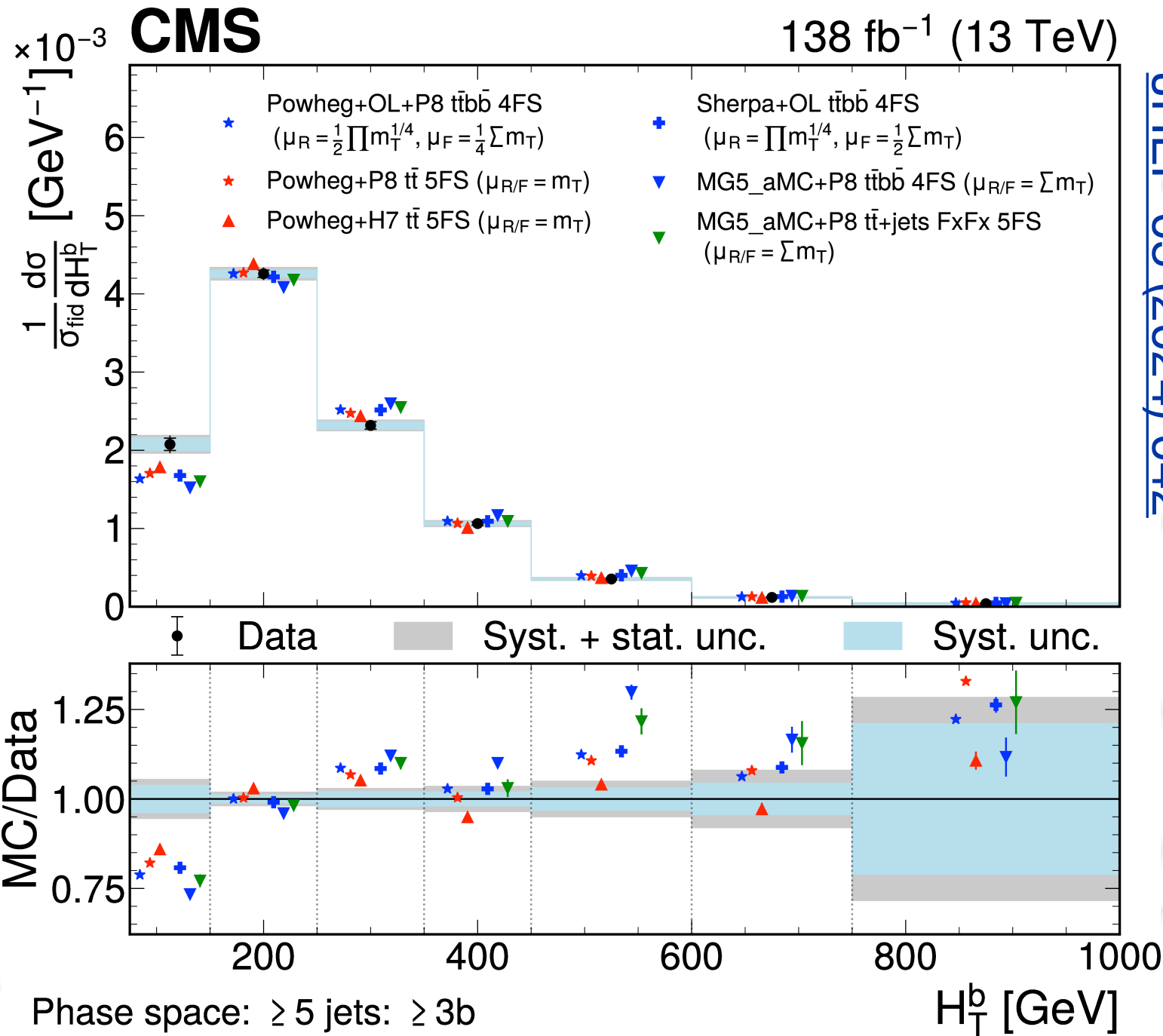


- Inclusive $t\bar{t}$ well described by NLO Powheg+Pythia8 prediction
 - See e.g. multi-differential cross section measurements
- Associated production with heavy-flavour quarks difficult:
 - Theoretical difficulties:
 - Massless or massive b-quarks?
 - E.g. 5FS inclusive NLO prediction vs. 4FS dedicated $t\bar{t}b\bar{b}$ simulation
- Experimental difficulties:
 - What do we consider a b-quark jet?



- Rich set of $t\bar{t}b\bar{b}$ measurements:
 - No generator has a good description everywhere
 - 4FS $t\bar{t}b\bar{b}$ simulation close to data inclusively
 - For differential distributions, agreement varies...

Model with enough flexibility needed

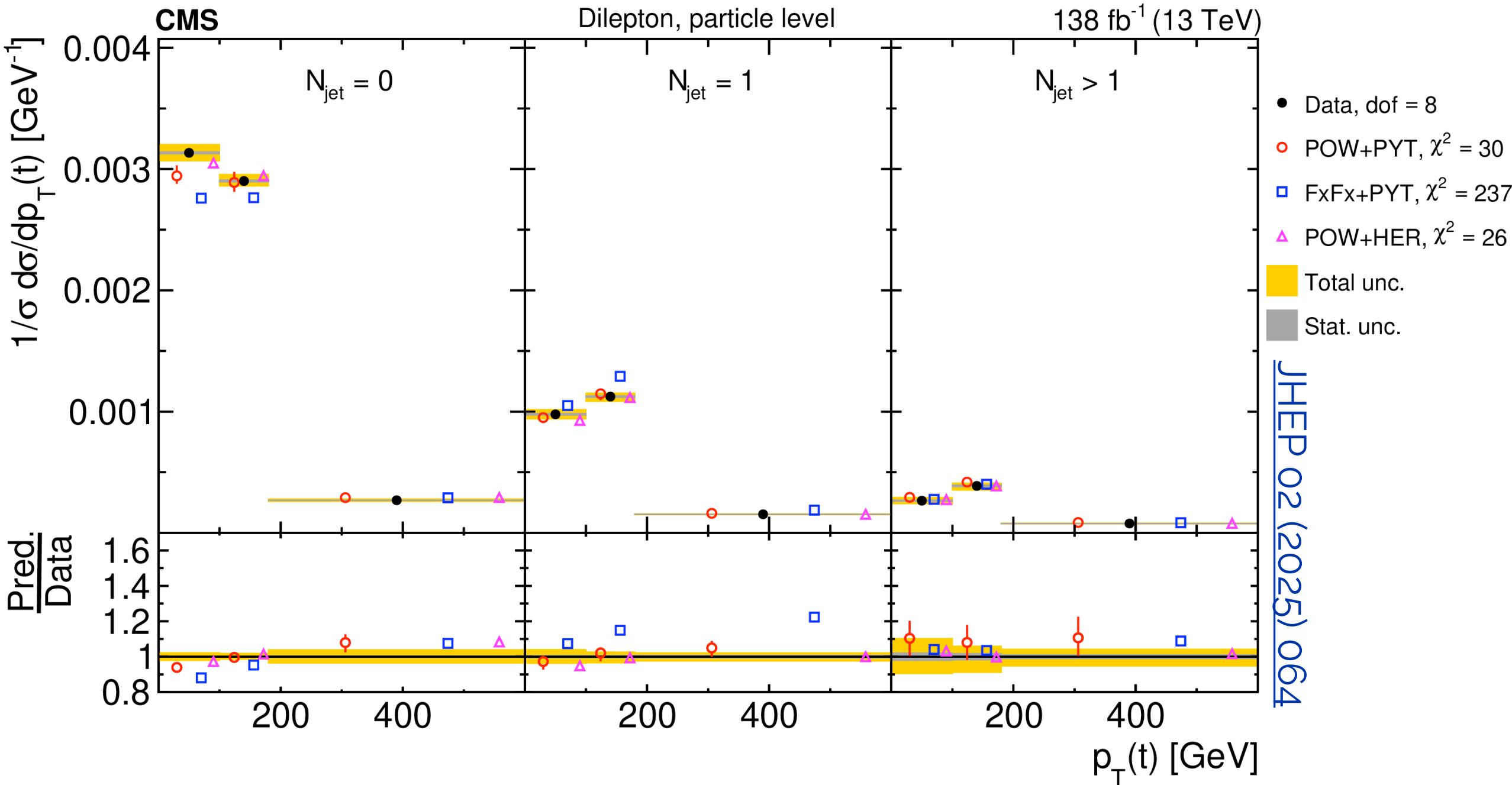
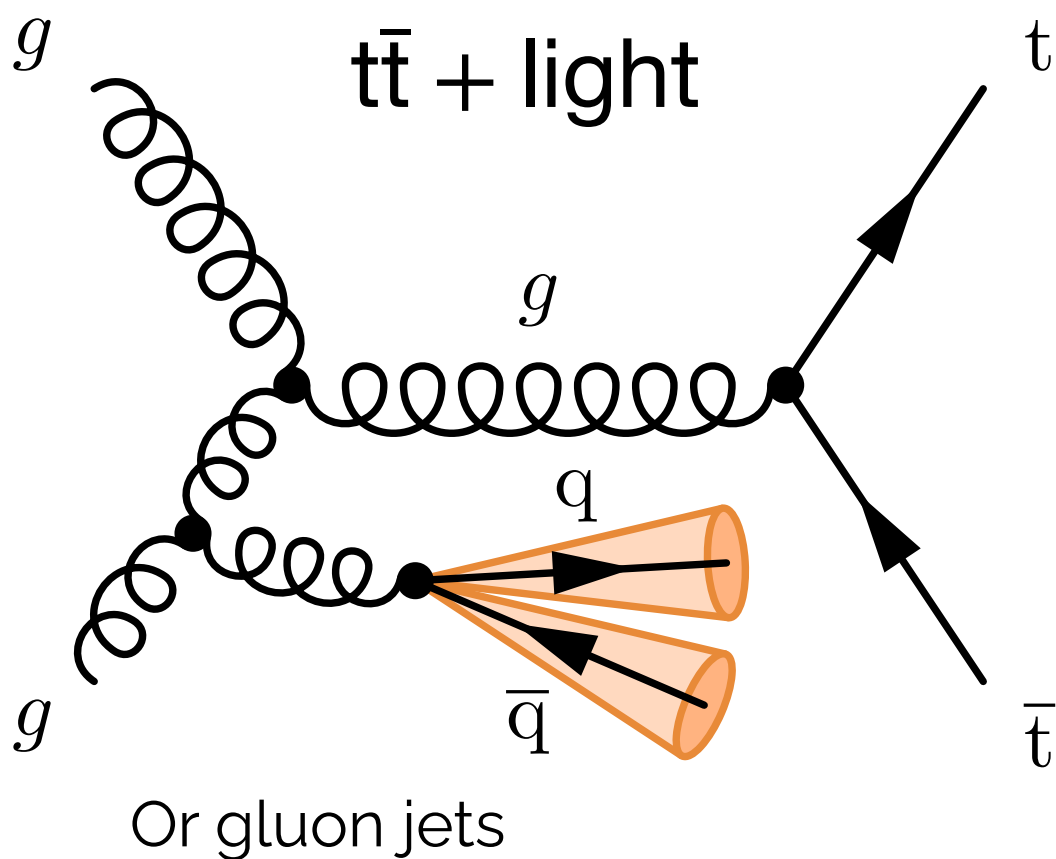


JHEP 05 (2024) 042

How well do we know the tt+jets background?



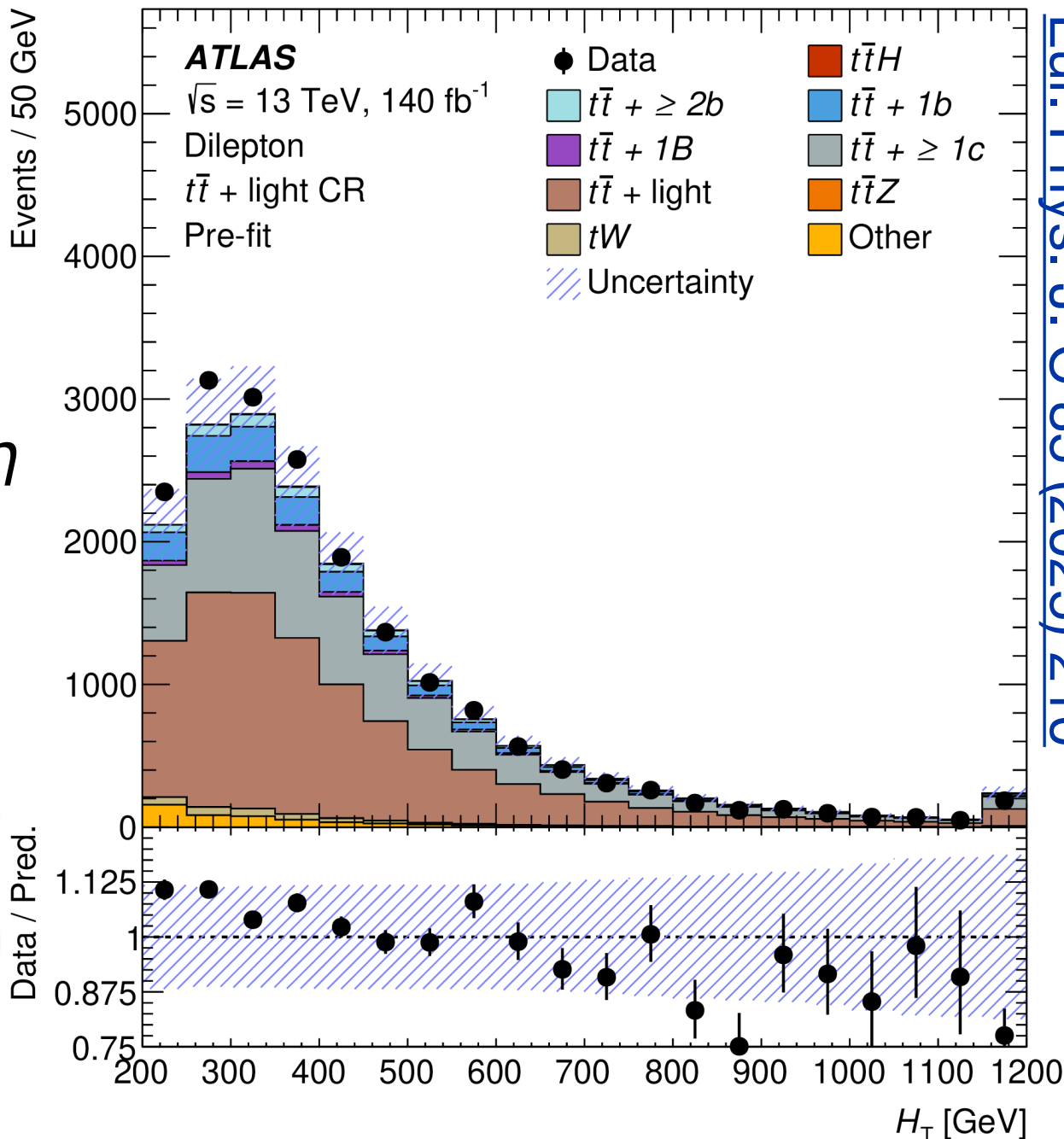
- For $t\bar{t}$ + light:
 - Extensive set of inclusive and differential measurements
 - Up to 3D in parton & particle level
 - Very good agreement by [Powheg+Pythia](#) simulation



Dedicated $t\bar{t}$ measurements

JHEP 02 (2025) 064

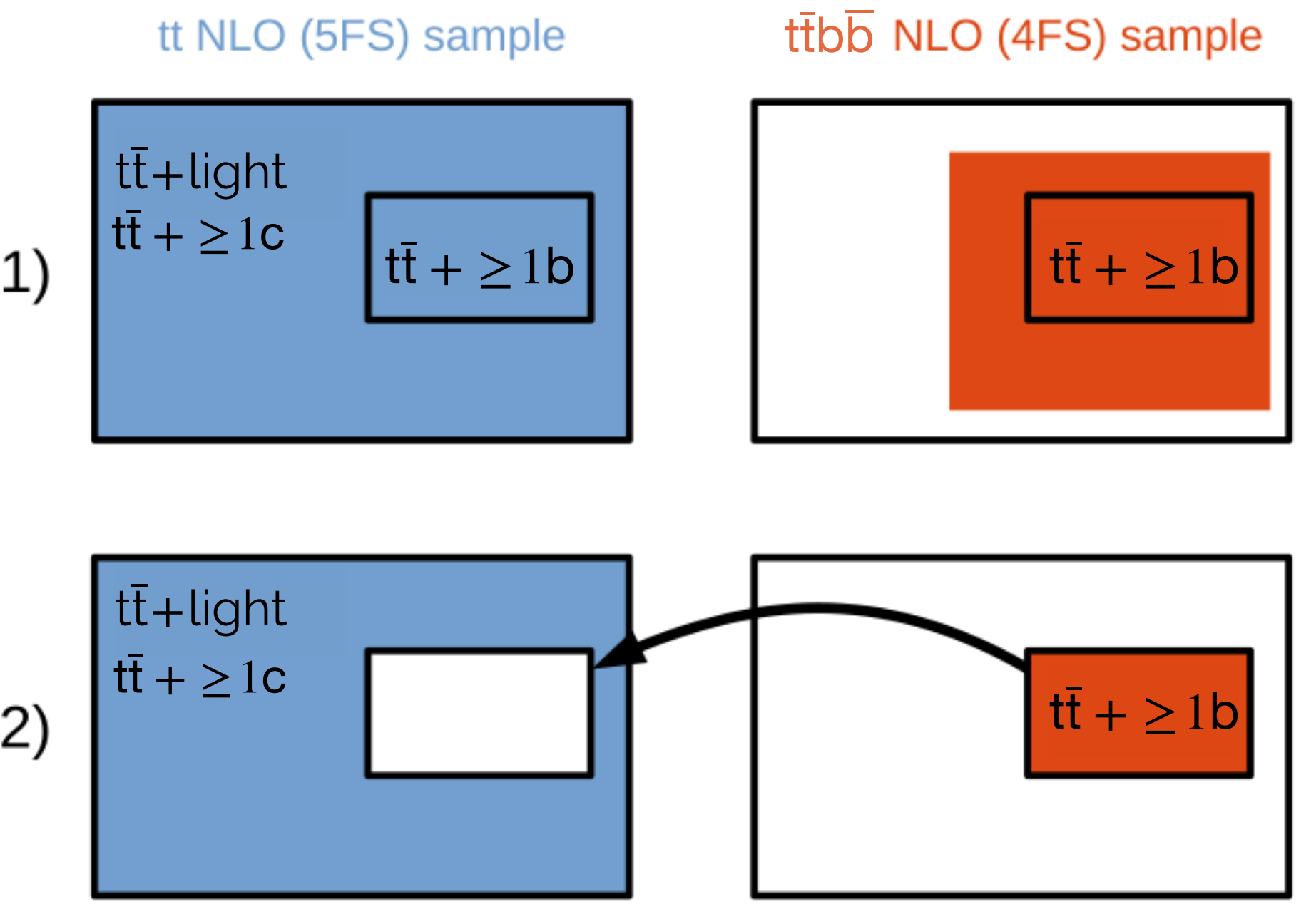
As background in $t\bar{t}H(H \rightarrow b\bar{b})$ searches



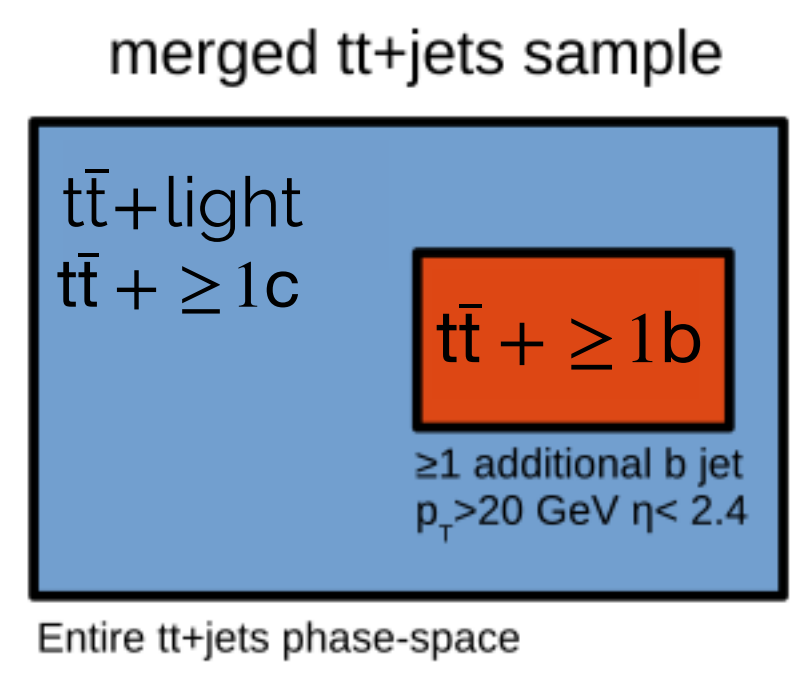
Eur. Phys. J. C 85 (2025) 210



	t $\bar{t}$ NLO (5FS)	t $\bar{t}b\bar{b}$ NLO (4FS)
ME generator	POWHEG v2	POWHEG-BOX-RES
Parton shower	PYTHIA v8.230	PYTHIA v8.230
Flavor scheme	5FS	4FS
PDF set	NNPDF31_nnlo_hessian_pdfas	NNPDF31_nnlo_as_0118_nf_4
m_t	172.5 GeV	172.5 GeV
m_b	0	4.75 GeV
μ_R	$\sqrt{\frac{1}{2}(m_{rmT,t}^2 + m_{T,\bar{t}}^2)}$	$\frac{1}{2}\sqrt[4]{m_{t,t} m_{t,\bar{t}} m_{t,b} m_{t,\bar{b}}}$
μ_F	μ_R	$\frac{1}{4}[m_{t,t} + m_{t,\bar{t}} + m_{t,b} + m_{t,\bar{b}} + m_{t,g}]$
h_{damp}	$1.379m_t$	$1.379m_t$
Tune	CP5	CP5



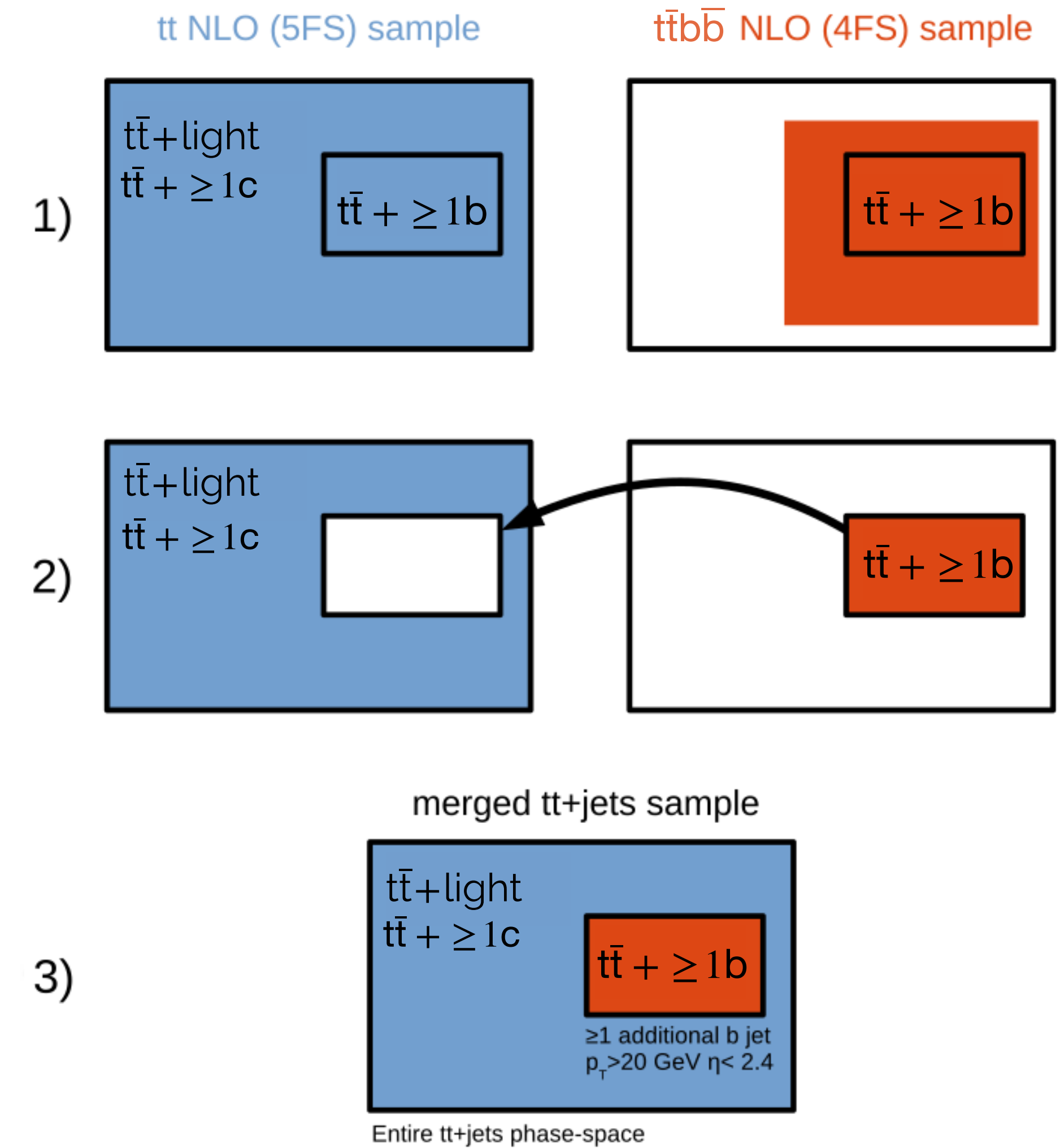
- State-of-the-art $t\bar{t}b\bar{b}$ background model:
 - Same procedure as in ATLAS measurement [1] (except for $\mu_F/2$)
- Replace PS-based emissions in inclusive 5FS $t\bar{t}$ +jets with dedicated ME $t\bar{t}b\bar{b}$ 4FS simulation
- Consider 4FS cross section prediction (43.76 pb)





tt+jets modeling (DPS contribution)

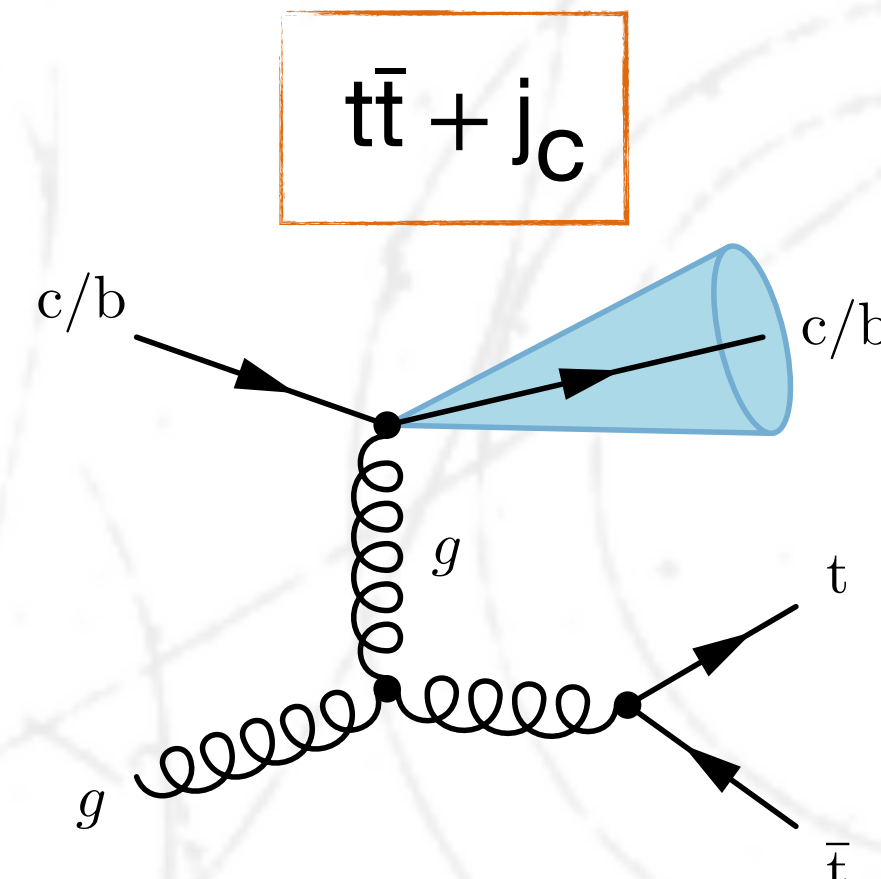
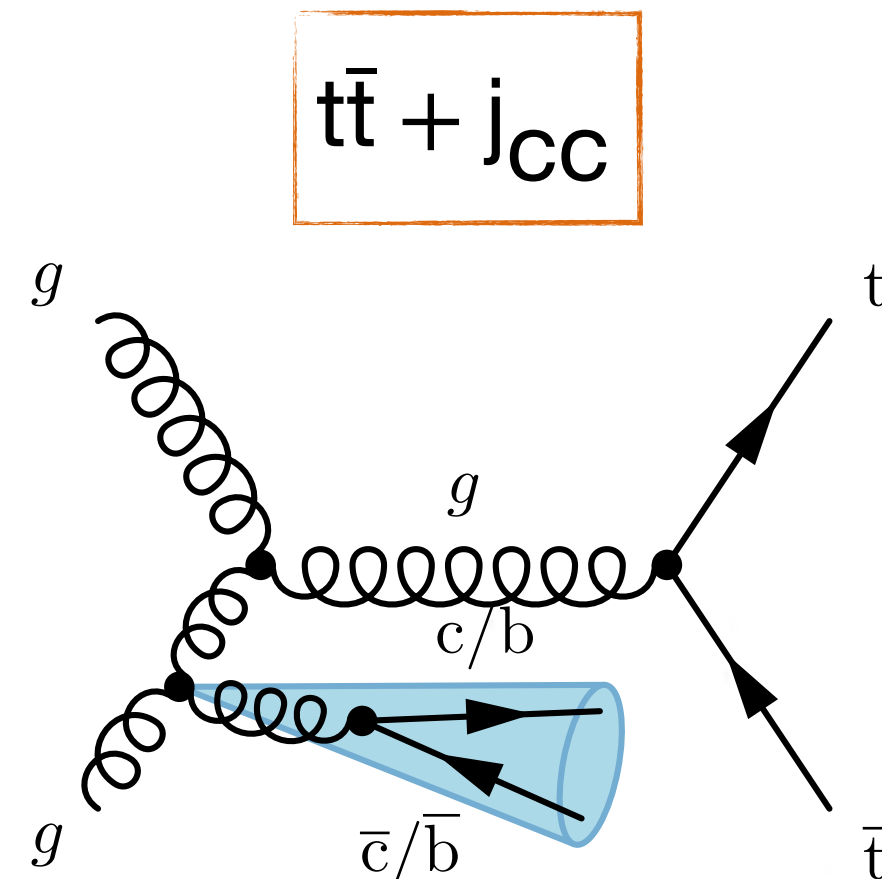
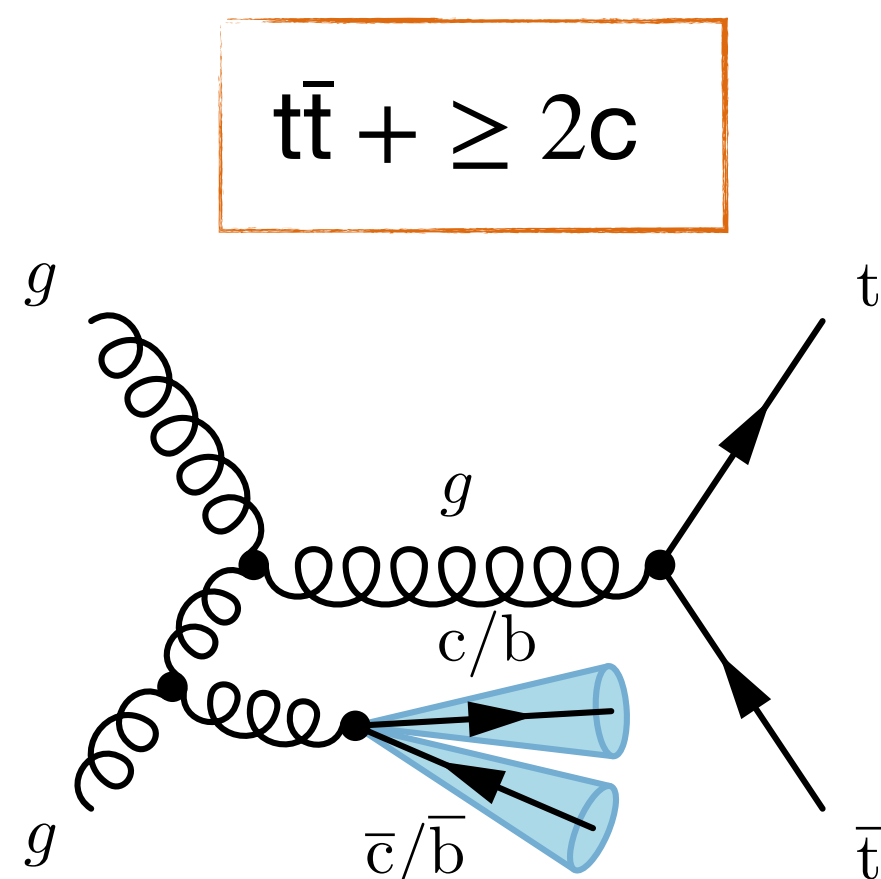
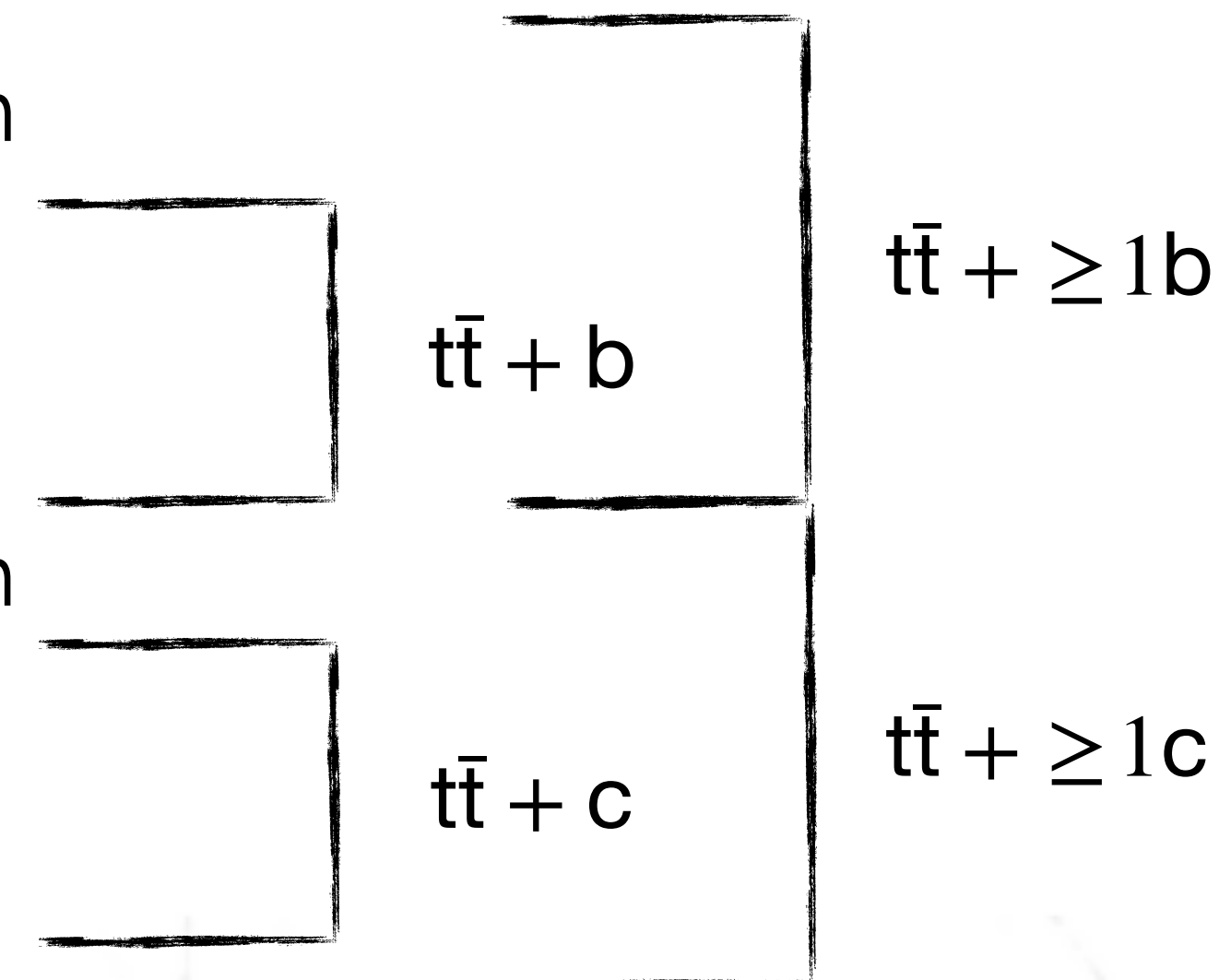
- With stitching procedure, we remove also contributions from MPI
 - $t\bar{t}$ hard scatter + $b\bar{b}$ add. scatter removed from $t\bar{t}$ 5FS sample
 - $t\bar{t}b\bar{b}$ 4FS does not include such events
 - Generate additional $t\bar{t}$ and $b\bar{b}$ DPS samples
 - At LO QCD using Powheg+Pythia
 - Estimated to be ~ 8 pb $\rightarrow \sim 15\%$ of total $t\bar{t}b\bar{b}$
 - Small compared to $t\bar{t}H$
- Discussed in CMS $t\bar{t}b\bar{b}$ measurement [JHEP 05 \(2024\) 042](#)



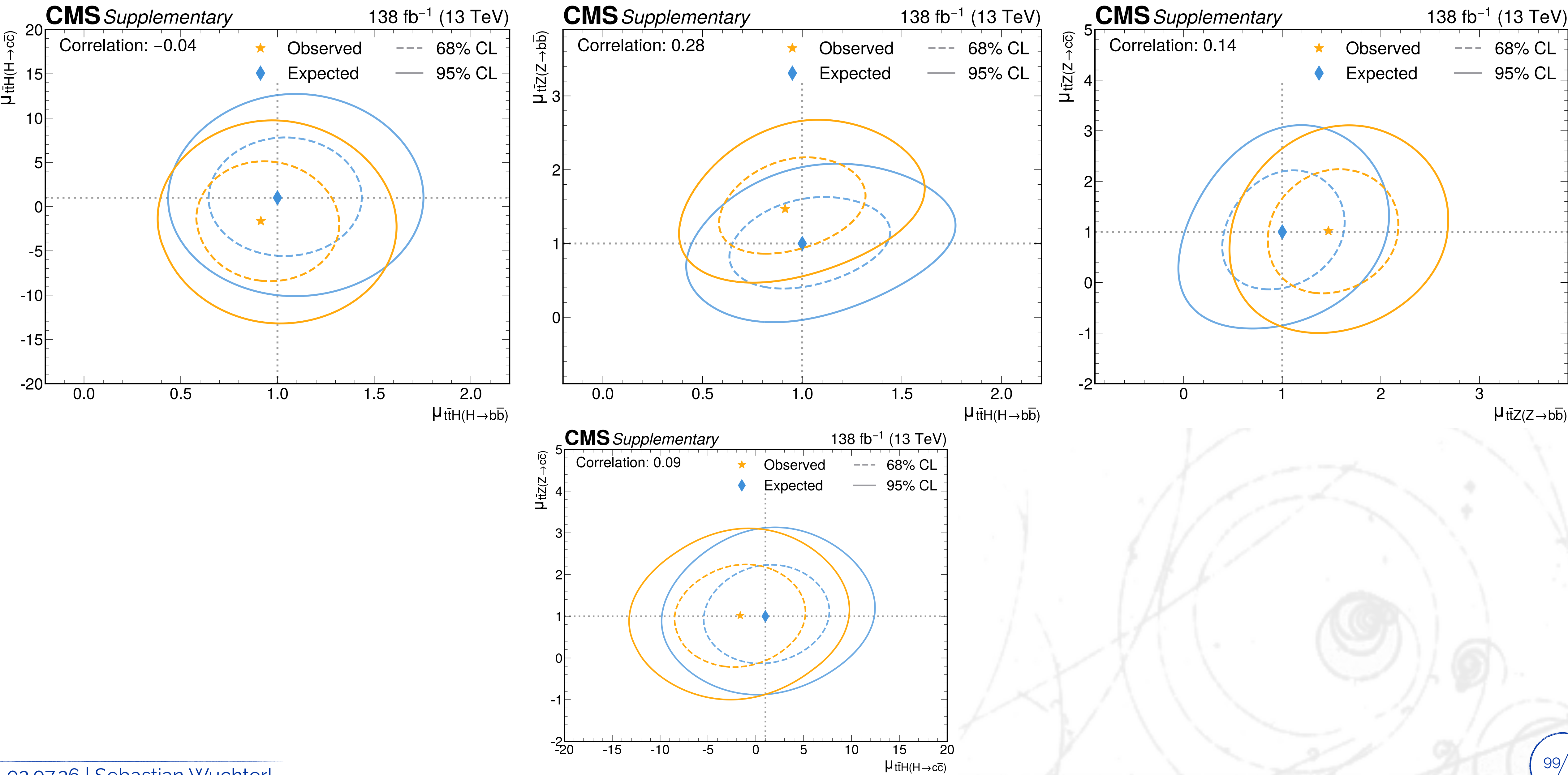


- Processes distinguished at generator level:

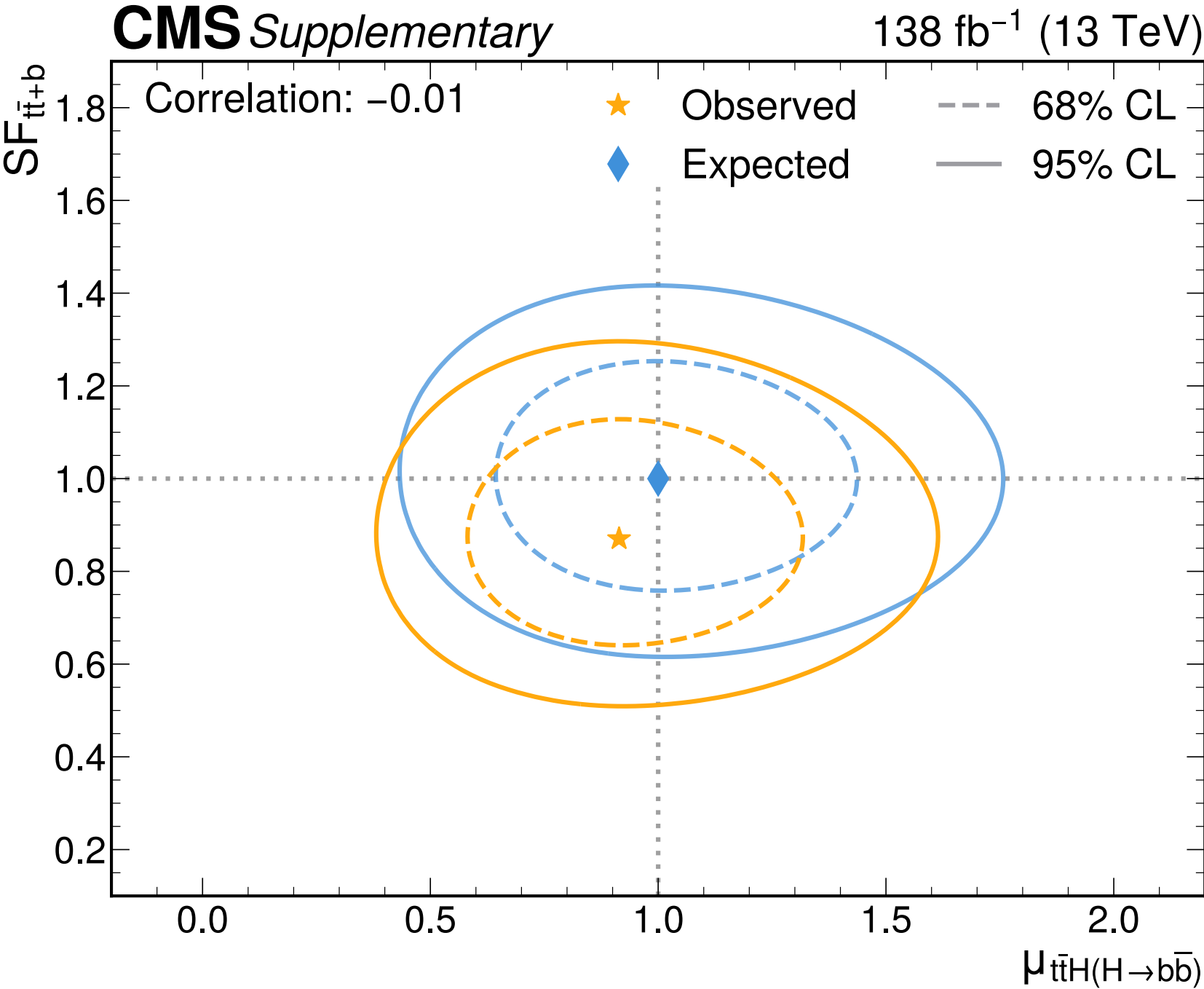
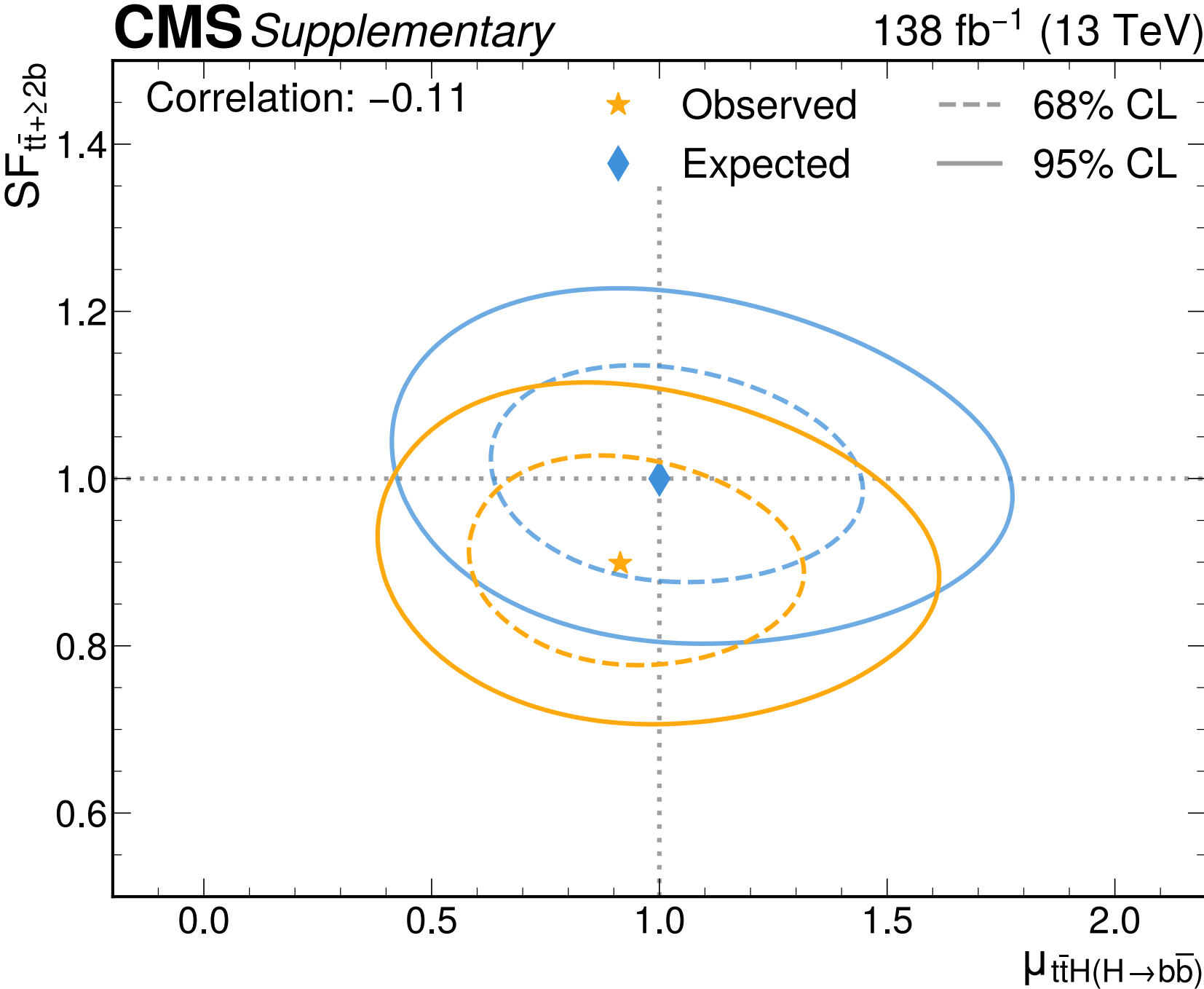
- $t\bar{t} + \geq 2b$: at least **two** additional b jets, from **one or more b hadrons** each
- $t\bar{t} + j_b$: exactly **one** additional b jet, from **one b hadron**
- $t\bar{t} + j_{bb}$: exactly **one** additional b jet, from **two b hadrons**
- $t\bar{t} + \geq 2c$: at least **two** additional c jets, from **one or more c hadrons** each
- $t\bar{t} + j_c$: exactly **one** additional c jet, from **one c hadron**
- $t\bar{t} + j_{cc}$: exactly **one** additional c jet, from **two c hadrons**
- $t\bar{t} + \text{light}$: events **without** any additional c or b jet



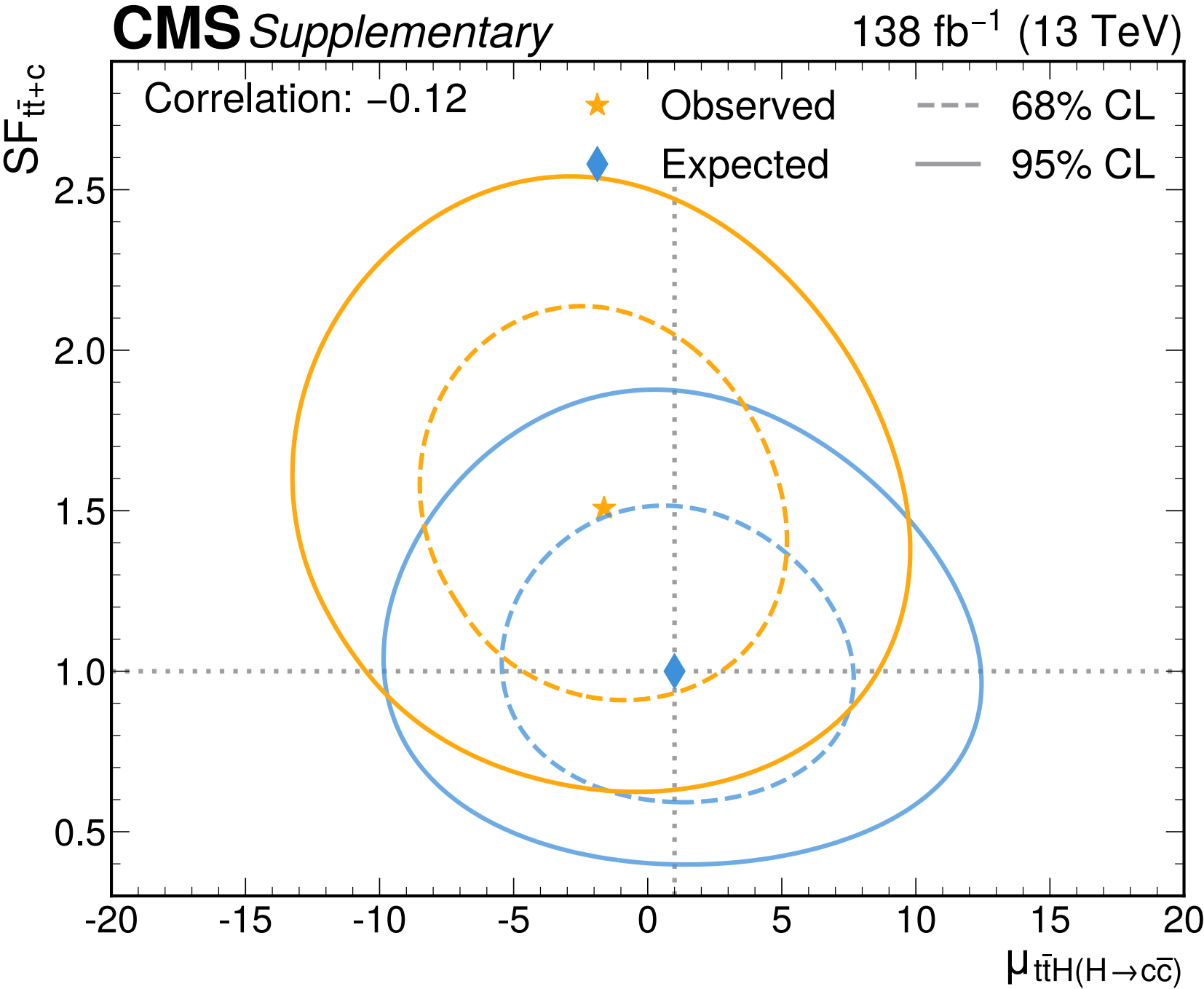
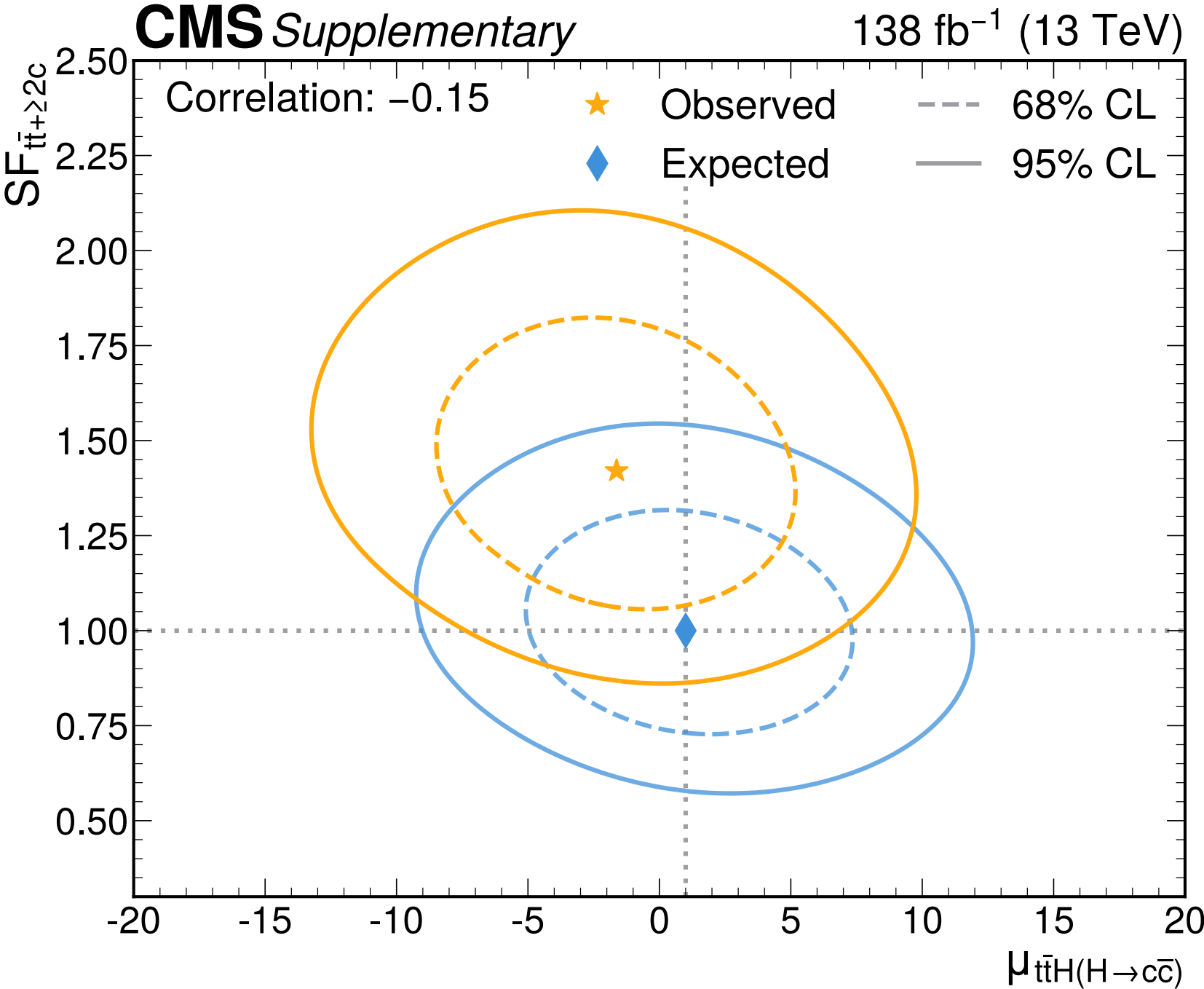
2D correlations - among signal strengths



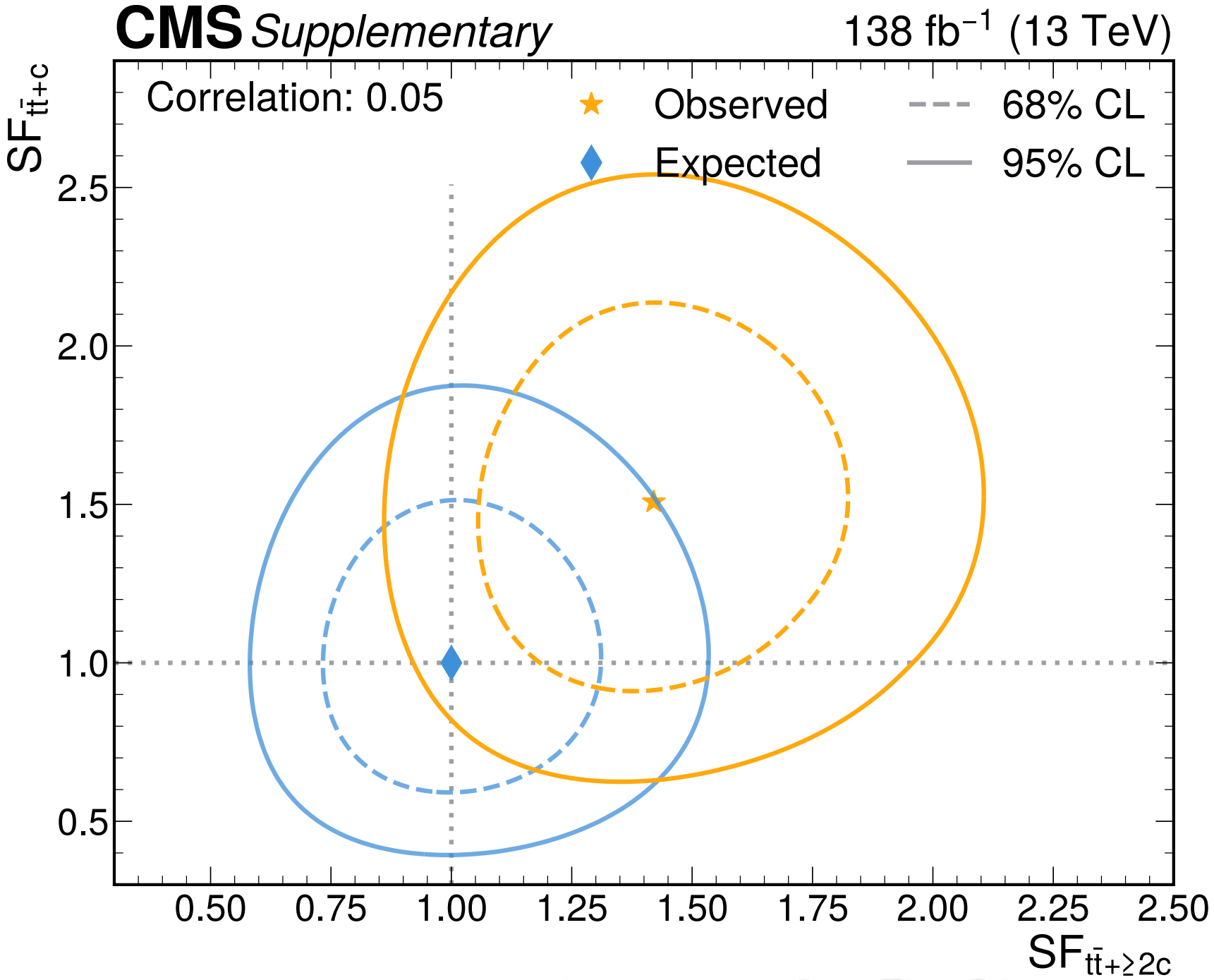
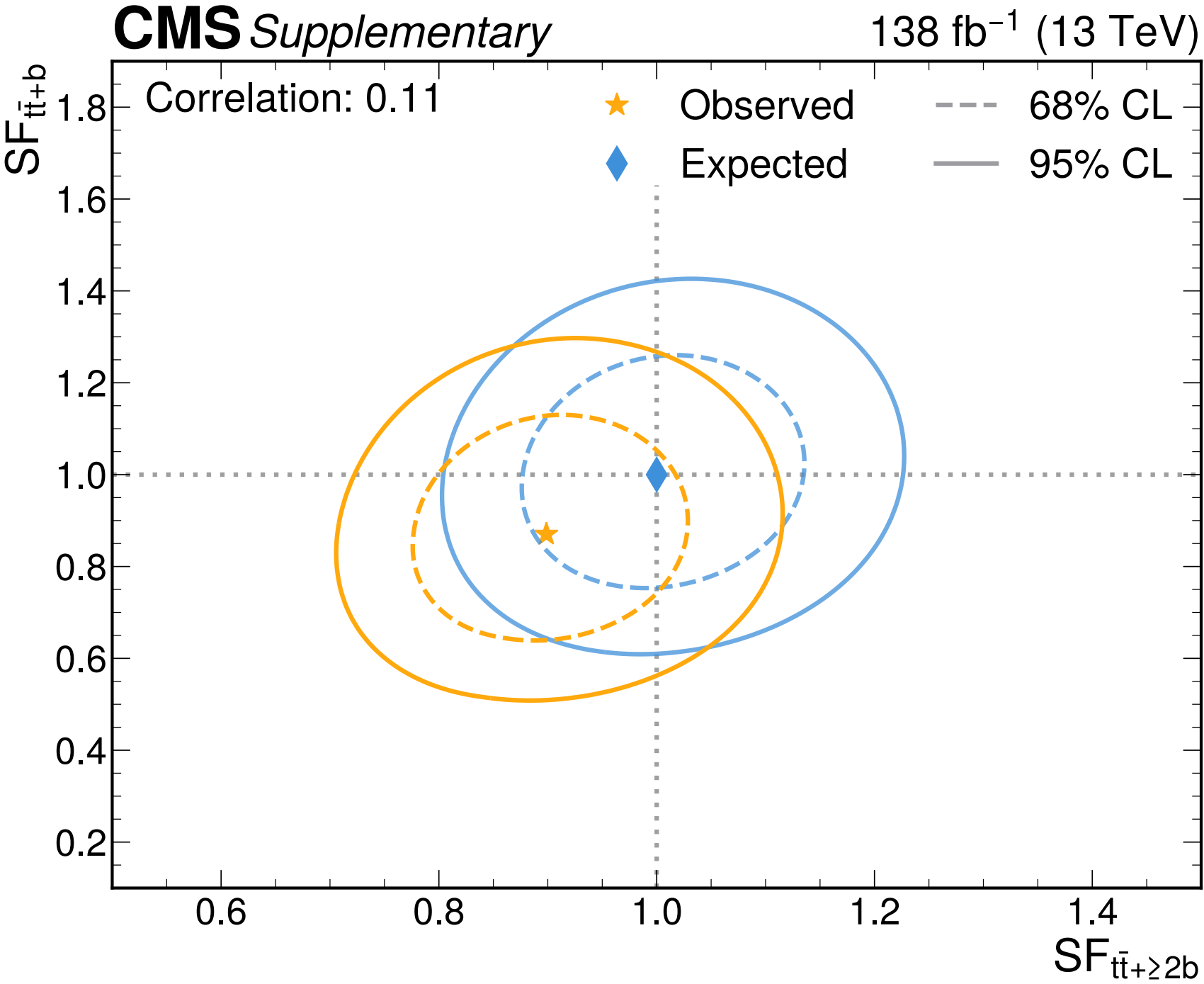
2D correlations - with background SFs



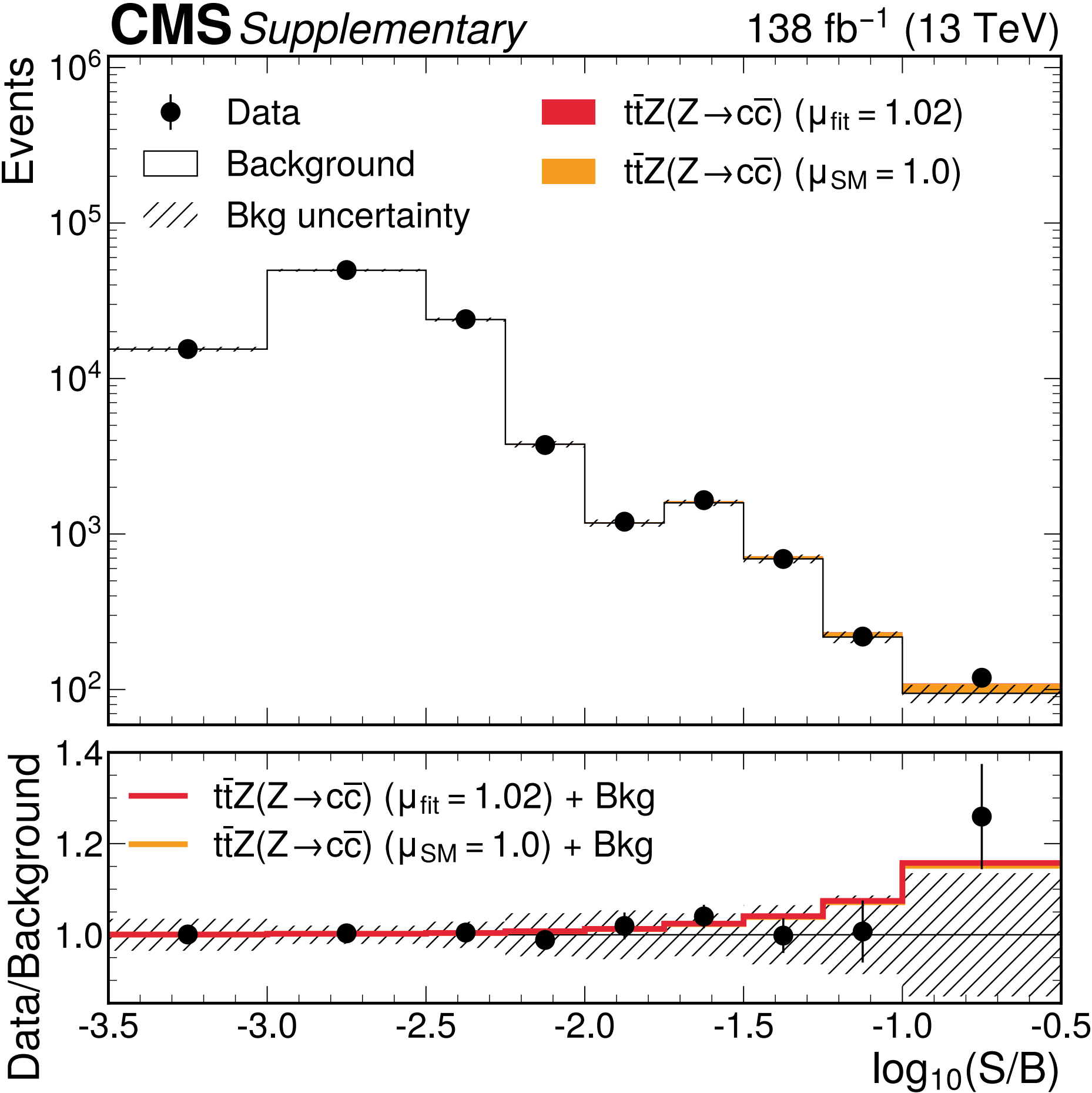
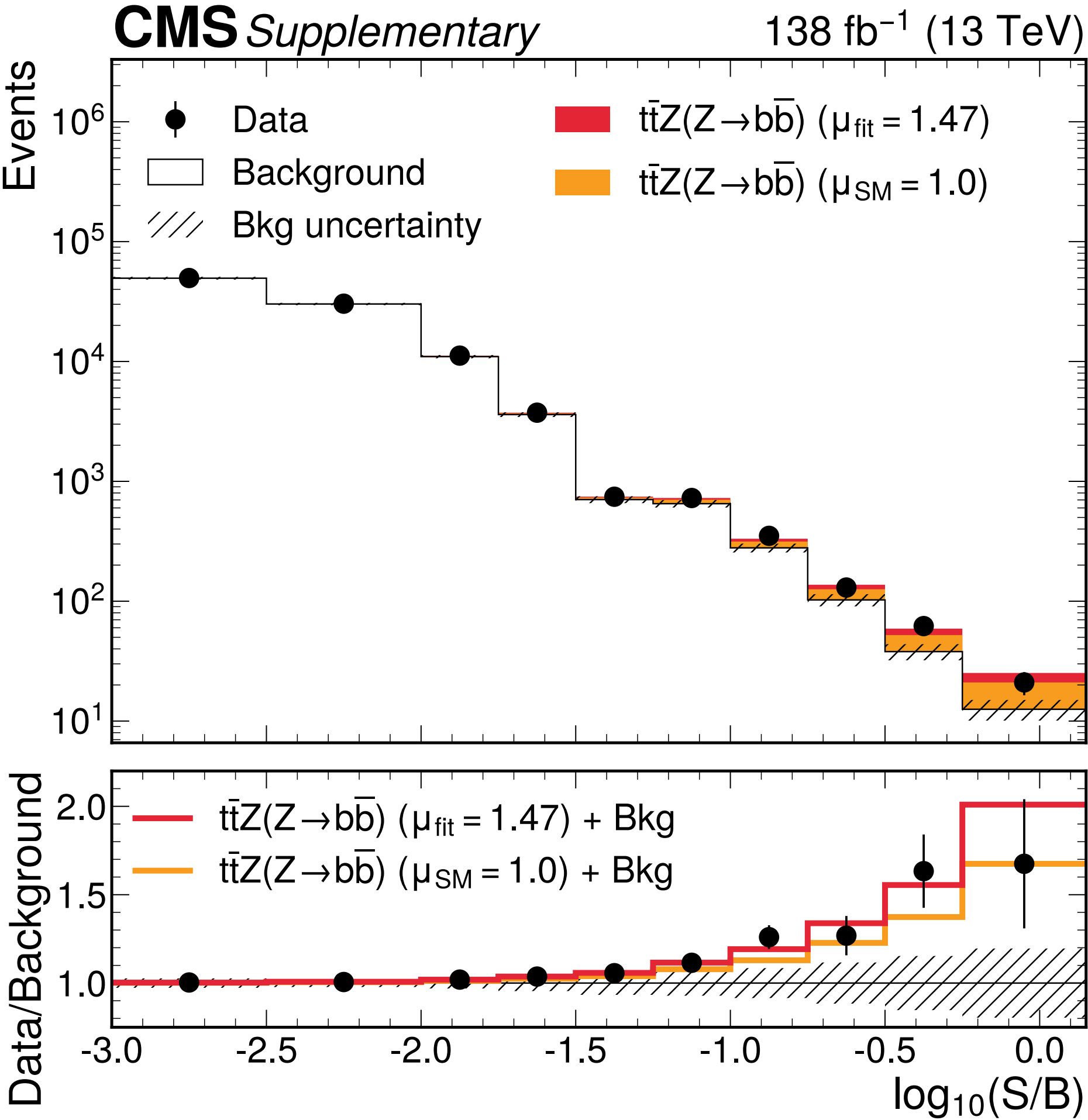
2D correlations - with background SFs



2D correlations - among background SFs

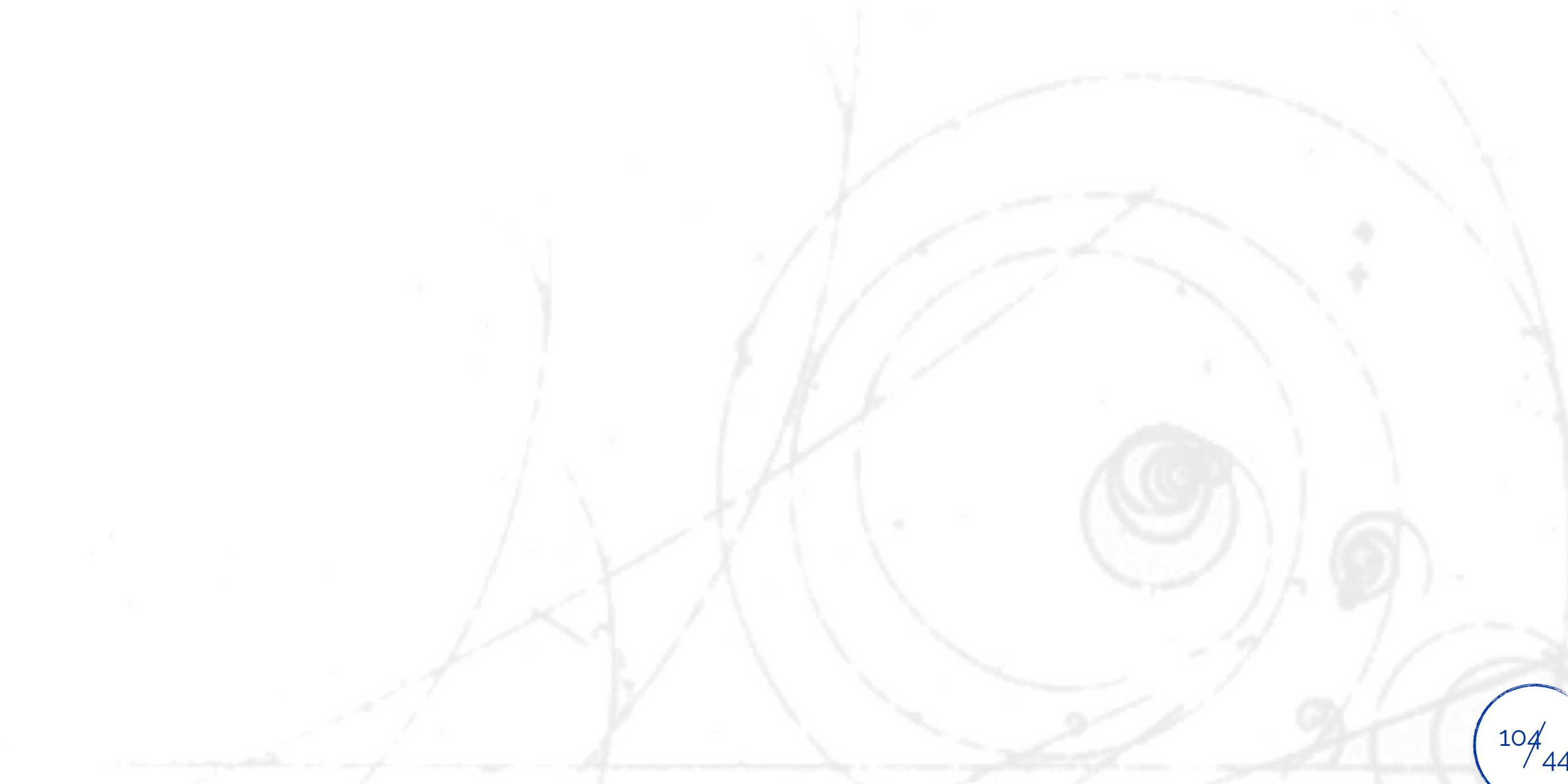


Accumulated plots for ttZ



Definition of scores and categories

Category	Discriminant
$H \rightarrow c\bar{c}$ SR	$\mathcal{D}_{t\bar{t}H(H \rightarrow c\bar{c})}$
$H \rightarrow b\bar{b}$ SR	$\mathcal{D}_{t\bar{t}H(H \rightarrow b\bar{b})}$
$Z \rightarrow c\bar{c}$ SR	$\mathcal{D}_{t\bar{t}Z(Z \rightarrow c\bar{c})}$
$Z \rightarrow b\bar{b}$ SR	$\mathcal{D}_{t\bar{t}Z(Z \rightarrow b\bar{b})}$
$t\bar{t} + \text{light}$ CR	$\mathcal{D}_{t\bar{t} + \text{light}} / \mathcal{D}_{t\bar{t} + \text{jets}}$
$t\bar{t} + c$ CR	$\mathcal{D}_{t\bar{t} + c} / \mathcal{D}_{t\bar{t} + \text{jets}}$
$t\bar{t} + \geq 2c$ CR	$\mathcal{D}_{t\bar{t} + \geq 2c} / \mathcal{D}_{t\bar{t} + \text{jets}}$
$t\bar{t} + b$ CR	$\mathcal{D}_{t\bar{t} + b} / \mathcal{D}_{t\bar{t} + \text{jets}}$
$t\bar{t} + \geq 2b$ CR	$\mathcal{D}_{t\bar{t} + \geq 2b} / \mathcal{D}_{t\bar{t} + \text{jets}}$



Event yields - integrated

- Acceptance:
 - ttH(cc): ~ 3.2%
 - In SR: ~ 0.8%
 - ttH(bb): ~ 4.3%
 - In SR: ~0.8%

Process	2L	1L	0L	Total
$t\bar{t}H(H\rightarrow c\bar{c})$	-9 ± 18 (6 ± 11)	-62 ± 69 (38 ± 42)	-33 ± 42 (20 ± 25)	-100 ± 130 (64 ± 78)
$t\bar{t}H(H\rightarrow b\bar{b})$	137 ± 56 (151 ± 62)	1060 ± 220 (1150 ± 240)	580 ± 140 (630 ± 150)	1770 ± 410 (1930 ± 450)
$t\bar{t}Z(Z\rightarrow c\bar{c})$	32 ± 27	210 ± 110	112 ± 65	360 ± 200
$t\bar{t}Z(Z\rightarrow b\bar{b})$	75 ± 53	580 ± 210	290 ± 130	940 ± 400
$t\bar{t}+\geq 2b$	1060 ± 180	8420 ± 650	4670 ± 480	14100 ± 1300
$t\bar{t}+b$	1310 ± 260	12100 ± 1000	3730 ± 530	17100 ± 1800
$t\bar{t}+\geq 2c$	940 ± 160	7320 ± 610	3850 ± 460	12100 ± 1200
$t\bar{t}+c$	1730 ± 290	15600 ± 1200	3220 ± 390	20600 ± 1900
$t\bar{t}+\text{light}$	1700 ± 140	17080 ± 670	4990 ± 280	23800 ± 1100
Other	430 ± 200	5040 ± 810	870 ± 320	6300 ± 1300
Total	7410 ± 240	67300 ± 1400	22270 ± 720	97000 ± 2400
Data	7366	67370	22271	97007

Event yields - 2L channel

Process	SR $t\bar{t}H(H\rightarrow c\bar{c})$	SR $t\bar{t}H(H\rightarrow b\bar{b})$	SR $t\bar{t}Z(Z\rightarrow c\bar{c})$	SR $t\bar{t}Z(Z\rightarrow b\bar{b})$	CR $t\bar{t}+\geq 2b$	CR $t\bar{t}+b$	CR $t\bar{t}+\geq 2c$	CR $t\bar{t}+c$	CR $t\bar{t}+light$
$t\bar{t}H(H\rightarrow c\bar{c})$	-2.5 ± 4.4 (1.5 ± 2.8)	-0.06 ± 0.62 (0.03 ± 0.37)	-0.5 ± 1.8 (0.3 ± 1.1)	-0.02 ± 0.35 (0.01 ± 0.21)	-0.13 ± 0.83 (0.08 ± 0.50)	-0.04 ± 0.51 (0.02 ± 0.31)	-1.2 ± 2.8 (0.7 ± 1.7)	-0.4 ± 1.5 (0.24 ± 0.97)	-4.4 ± 5.2 (2.7 ± 3.2)
$t\bar{t}H(H\rightarrow b\bar{b})$	2.4 ± 3.1 (2.7 ± 3.4)	20.2 ± 9.6 (22 ± 11)	1.3 ± 2.1 (1.4 ± 2.2)	6.4 ± 4.7 (7.1 ± 5.2)	37 ± 12 (40 ± 13)	18.6 ± 8.1 (20.8 ± 9.0)	4.1 ± 3.4 (4.4 ± 3.6)	1.8 ± 2.4 (2.1 ± 2.6)	45 ± 11 (50 ± 13)
$t\bar{t}Z(Z\rightarrow c\bar{c})$	1.9 ± 2.7	0.11 ± 0.64	9.9 ± 7.1	0.21 ± 0.90	0.4 ± 1.1	0.12 ± 0.71	4.3 ± 4.2	1.2 ± 2.2	14.2 ± 7.4
$t\bar{t}Z(Z\rightarrow b\bar{b})$	0.7 ± 2.1	3.8 ± 4.7	1.1 ± 2.7	16 ± 11	19 ± 11	9.1 ± 7.2	2.1 ± 3.1	0.9 ± 2.2	23 ± 10
$t\bar{t}+\geq 2b$	12.7 ± 7.9	45 ± 16	11.8 ± 7.4	52 ± 17	370 ± 42	121 ± 24	52 ± 15	14.3 ± 8.0	385 ± 40
$t\bar{t}+b$	19 ± 14	14 ± 13	16 ± 12	17 ± 14	120 ± 32	303 ± 56	54 ± 22	37 ± 19	727 ± 81
$t\bar{t}+\geq 2c$	67 ± 21	1.9 ± 3.7	63 ± 20	2.4 ± 3.3	16.0 ± 7.7	5.5 ± 5.6	214 ± 35	44 ± 15	524 ± 52
$t\bar{t}+c$	61 ± 28	1.7 ± 4.0	49 ± 24	3.4 ± 6.8	16 ± 11	23 ± 16	121 ± 37	166 ± 43	1290 ± 120
$t\bar{t}+light$	50 ± 14	1.7 ± 2.7	44 ± 12	1.8 ± 2.7	8.1 ± 4.7	11.2 ± 6.1	60 ± 14	54 ± 13	1470 ± 74
Other	15 ± 18	2.8 ± 7.3	24 ± 24	5.4 ± 9.7	22 ± 17	8.5 ± 8.1	32 ± 25	13 ± 14	313 ± 75
Total	228 ± 14	90.7 ± 7.9	220 ± 14	103.6 ± 9.1	608 ± 22	499 ± 19	541 ± 27	332 ± 18	4790 ± 110
Data	236	93	204	92	605	463	546	343	4784

Event yields - 1L channel

Process	SR $t\bar{t}H(H\rightarrow c\bar{c})$	SR $t\bar{t}H(H\rightarrow b\bar{b})$	SR $t\bar{t}Z(Z\rightarrow c\bar{c})$	SR $t\bar{t}Z(Z\rightarrow b\bar{b})$	CR $t\bar{t}+\geq 2b$	CR $t\bar{t}+b$	CR $t\bar{t}+\geq 2c$	CR $t\bar{t}+c$	CR $t\bar{t}+light$
$t\bar{t}H(H\rightarrow c\bar{c})$	-16 ± 22 (10 ± 14)	-0.5 ± 3.6 (0.3 ± 2.1)	-3.1 ± 8.9 (1.9 ± 5.4)	-0.2 ± 1.9 (0.1 ± 1.1)	-1.2 ± 2.6 (0.7 ± 1.5)	-0.5 ± 1.8 (0.3 ± 1.0)	-11.7 ± 8.8 (7.0 ± 5.3)	-5.4 ± 5.9 (3.3 ± 3.6)	-24 ± 13 (14.8 ± 7.7)
$t\bar{t}H(H\rightarrow b\bar{b})$	21 ± 18 (23 ± 19)	166 ± 55 (181 ± 60)	9 ± 11 (10 ± 12)	43 ± 25 (47 ± 27)	264 ± 31 (284 ± 34)	139 ± 22 (153 ± 25)	52 ± 12 (55 ± 13)	35 ± 10 (38 ± 11)	327 ± 32 (356 ± 35)
$t\bar{t}Z(Z\rightarrow c\bar{c})$	16 ± 16	1.1 ± 4.1	59 ± 35	1.5 ± 4.9	3.5 ± 3.6	1.6 ± 2.5	37 ± 12	18.1 ± 8.6	78 ± 18
$t\bar{t}Z(Z\rightarrow b\bar{b})$	6 ± 12	32 ± 28	8 ± 14	114 ± 58	136 ± 28	70 ± 20	25 ± 11	19.2 ± 9.7	166 ± 29
$t\bar{t}+\geq 2b$	109 ± 48	387 ± 91	85 ± 41	360 ± 88	2760 ± 120	957 ± 67	633 ± 51	293 ± 34	2830 ± 110
$t\bar{t}+b$	198 ± 96	205 ± 97	168 ± 88	165 ± 87	1300 ± 110	2700 ± 170	841 ± 88	719 ± 81	5760 ± 240
$t\bar{t}+\geq 2c$	470 ± 110	26 ± 24	389 ± 99	19 ± 21	153 ± 26	60 ± 18	2080 ± 110	718 ± 63	3400 ± 140
$t\bar{t}+c$	700 ± 190	61 ± 54	470 ± 150	45 ± 46	256 ± 48	291 ± 59	1910 ± 150	2680 ± 180	9220 ± 330
$t\bar{t}+light$	780 ± 110	73 ± 33	605 ± 94	69 ± 33	205 ± 24	226 ± 29	1180 ± 65	1326 ± 70	12620 ± 220
Other	190 ± 120	43 ± 48	320 ± 160	58 ± 55	249 ± 50	136 ± 33	550 ± 86	423 ± 70	3080 ± 190
Total	2470 ± 92	993 ± 52	2115 ± 80	875 ± 48	5330 ± 110	4580 ± 110	7300 ± 200	6230 ± 190	37450 ± 530
Data	2411	983	2108	879	5372	4594	7317	6209	37497

Event yields - oL channel

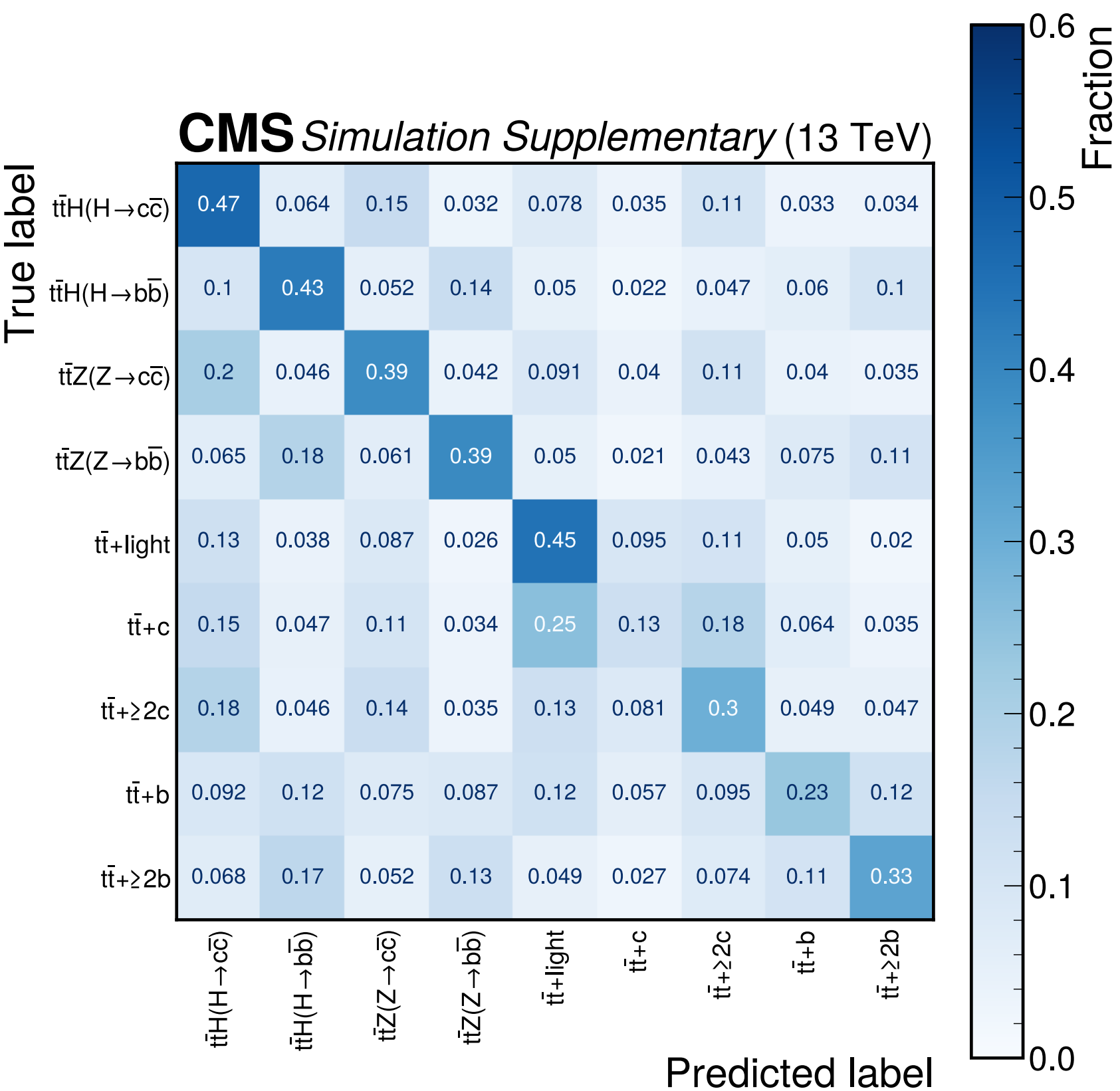
Process	SR $t\bar{t}H(H\rightarrow c\bar{c})$	SR $t\bar{t}H(H\rightarrow b\bar{b})$	SR $t\bar{t}Z(Z\rightarrow c\bar{c})$	SR $t\bar{t}Z(Z\rightarrow b\bar{b})$	CR $t\bar{t}+\geq 2b$	CR $t\bar{t}+b$	CR $t\bar{t}+\geq 2c$	CR $t\bar{t}+c$	CR $t\bar{t}+light$
$t\bar{t}H(H\rightarrow c\bar{c})$	-9 ± 13 (5.5 ± 8.1)	-0.4 ± 2.5 (0.2 ± 1.4)	-1.6 ± 5.0 (1.0 ± 3.0)	-0.1 ± 1.4 (0.08 ± 0.79)	-0.7 ± 1.9 (0.4 ± 1.1)	-0.3 ± 1.4 (0.17 ± 0.79)	-3.4 ± 4.7 (2.0 ± 2.8)	-0.5 ± 1.9 (0.3 ± 1.1)	-17 ± 10 (10.2 ± 6.2)
$t\bar{t}H(H\rightarrow b\bar{b})$	12 ± 10 (13 ± 11)	128 ± 38 (137 ± 41)	5.4 ± 6.5 (5.9 ± 7.1)	31 ± 16 (34 ± 17)	140 ± 23 (151 ± 24)	62 ± 15 (68 ± 16)	14.9 ± 6.5 (15.8 ± 6.9)	2.9 ± 2.8 (3.1 ± 3.0)	185 ± 23 (201 ± 25)
$t\bar{t}Z(Z\rightarrow c\bar{c})$	8.0 ± 9.0	0.8 ± 2.9	36 ± 21	1.3 ± 3.6	1.9 ± 2.6	1.0 ± 2.0	10.4 ± 6.7	2.0 ± 2.9	51 ± 14
$t\bar{t}Z(Z\rightarrow b\bar{b})$	3.1 ± 6.5	22 ± 18	4.7 ± 8.3	87 ± 40	62 ± 19	28 ± 13	6.2 ± 5.4	1.5 ± 2.7	75 ± 19
$t\bar{t}+\geq 2b$	73 ± 32	336 ± 73	55 ± 28	320 ± 70	1548 ± 97	452 ± 50	178 ± 30	28 ± 11	1677 ± 89
$t\bar{t}+b$	86 ± 47	96 ± 49	70 ± 43	88 ± 50	405 ± 60	742 ± 86	159 ± 37	45 ± 20	2040 ± 130
$t\bar{t}+\geq 2c$	296 ± 81	22 ± 21	260 ± 75	20 ± 19	74 ± 22	30 ± 15	653 ± 72	85 ± 25	2410 ± 130
$t\bar{t}+c$	163 ± 59	21 ± 20	113 ± 48	16 ± 17	49 ± 18	44 ± 19	180 ± 39	113 ± 30	2520 ± 140
$t\bar{t}+light$	213 ± 42	27 ± 15	165 ± 36	21 ± 13	44 ± 11	43 ± 12	122 ± 20	50 ± 13	4310 ± 120
Other	39 ± 40	13 ± 20	106 ± 67	18 ± 23	32 ± 19	24 ± 15	52 ± 29	17 ± 16	565 ± 88
Total	883 ± 50	665 ± 42	814 ± 47	602 ± 40	2357 ± 74	1426 ± 63	1372 ± 64	343 ± 26	13800 ± 310
Data	898	668	831	622	2345	1435	1382	357	13733

Event yields - split by categories

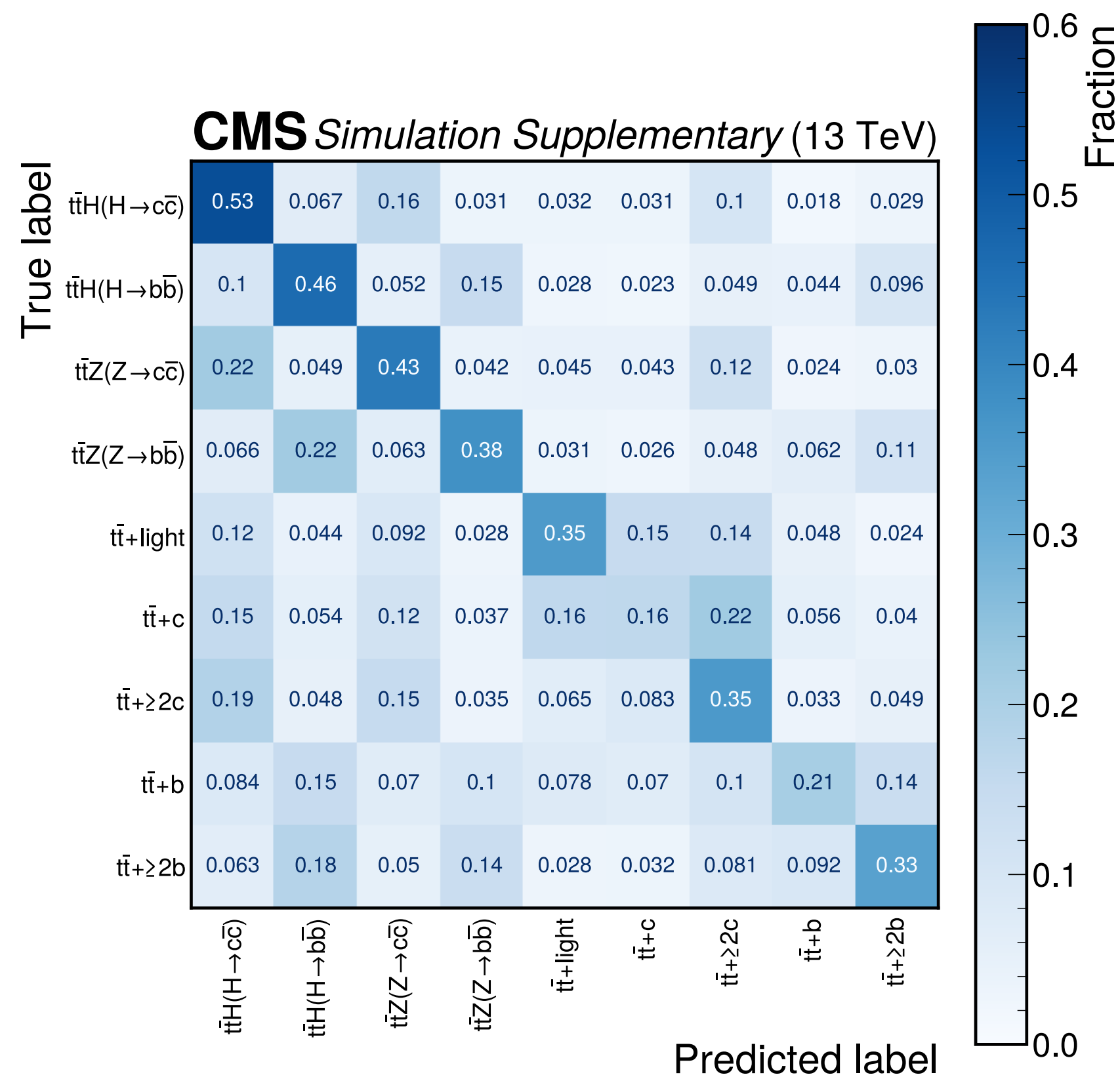
Process	SR	SR	SR	SR	CR	CR	CR	CR	CR
	$t\bar{t}H(H\rightarrow c\bar{c})$	$t\bar{t}H(H\rightarrow b\bar{b})$	$t\bar{t}Z(Z\rightarrow c\bar{c})$	$t\bar{t}Z(Z\rightarrow b\bar{b})$	$t\bar{t}+\geq 2b$	$t\bar{t}+b$	$t\bar{t}+\geq 2c$	$t\bar{t}+c$	$t\bar{t}+light$
$t\bar{t}H(H\rightarrow c\bar{c})$	-27 ± 40 (17 ± 25)	-0.9 ± 6.7 (0.5 ± 3.8)	-5 ± 16 (3.2 ± 9.5)	-0.3 ± 3.7 (0.2 ± 2.1)	-2.0 ± 5.4 (1.2 ± 3.1)	-0.8 ± 3.6 (0.5 ± 2.1)	-16 ± 16 (9.8 ± 9.8)	-6.3 ± 9.3 (3.9 ± 5.7)	-46 ± 28 (28 ± 17)
$t\bar{t}H(H\rightarrow b\bar{b})$	35 ± 31 (38 ± 34)	310 ± 100 (340 ± 110)	16 ± 20 (17 ± 21)	81 ± 45 (88 ± 49)	441 ± 65 (475 ± 71)	220 ± 45 (242 ± 50)	71 ± 22 (75 ± 24)	40 ± 15 (43 ± 17)	557 ± 67 (607 ± 73)
$t\bar{t}Z(Z\rightarrow c\bar{c})$	25 ± 28	2.0 ± 7.6	105 ± 63	2.9 ± 9.4	5.8 ± 7.3	2.7 ± 5.3	51 ± 23	21 ± 14	143 ± 40
$t\bar{t}Z(Z\rightarrow b\bar{b})$	10 ± 21	57 ± 51	14 ± 25	220 ± 110	218 ± 58	108 ± 40	34 ± 20	22 ± 15	263 ± 58
$t\bar{t}+\geq 2b$	194 ± 88	770 ± 180	152 ± 76	730 ± 170	4680 ± 260	1530 ± 140	862 ± 95	335 ± 53	4890 ± 240
$t\bar{t}+b$	300 ± 160	320 ± 160	250 ± 140	270 ± 150	1830 ± 200	3740 ± 310	1050 ± 150	800 ± 120	8520 ± 450
$t\bar{t}+\geq 2c$	830 ± 210	50 ± 49	710 ± 190	41 ± 43	244 ± 56	96 ± 39	2950 ± 220	850 ± 100	6330 ± 320
$t\bar{t}+c$	920 ± 270	84 ± 78	640 ± 230	64 ± 70	321 ± 77	358 ± 94	2210 ± 230	2960 ± 250	13030 ± 590
$t\bar{t}+light$	1050 ± 160	102 ± 51	810 ± 140	92 ± 49	257 ± 39	280 ± 47	1362 ± 99	1430 ± 96	18390 ± 410
Other	240 ± 180	58 ± 76	450 ± 250	82 ± 88	303 ± 87	169 ± 56	630 ± 140	453 ± 100	3950 ± 350
Total	3580 ± 160	1750 ± 100	3150 ± 140	1580 ± 97	8300 ± 210	6500 ± 200	9210 ± 290	6910 ± 240	56040 ± 950
Data	3545	1744	3143	1593	8322	6492	9245	6909	56014

Classifier performance

0L



1L



2L

