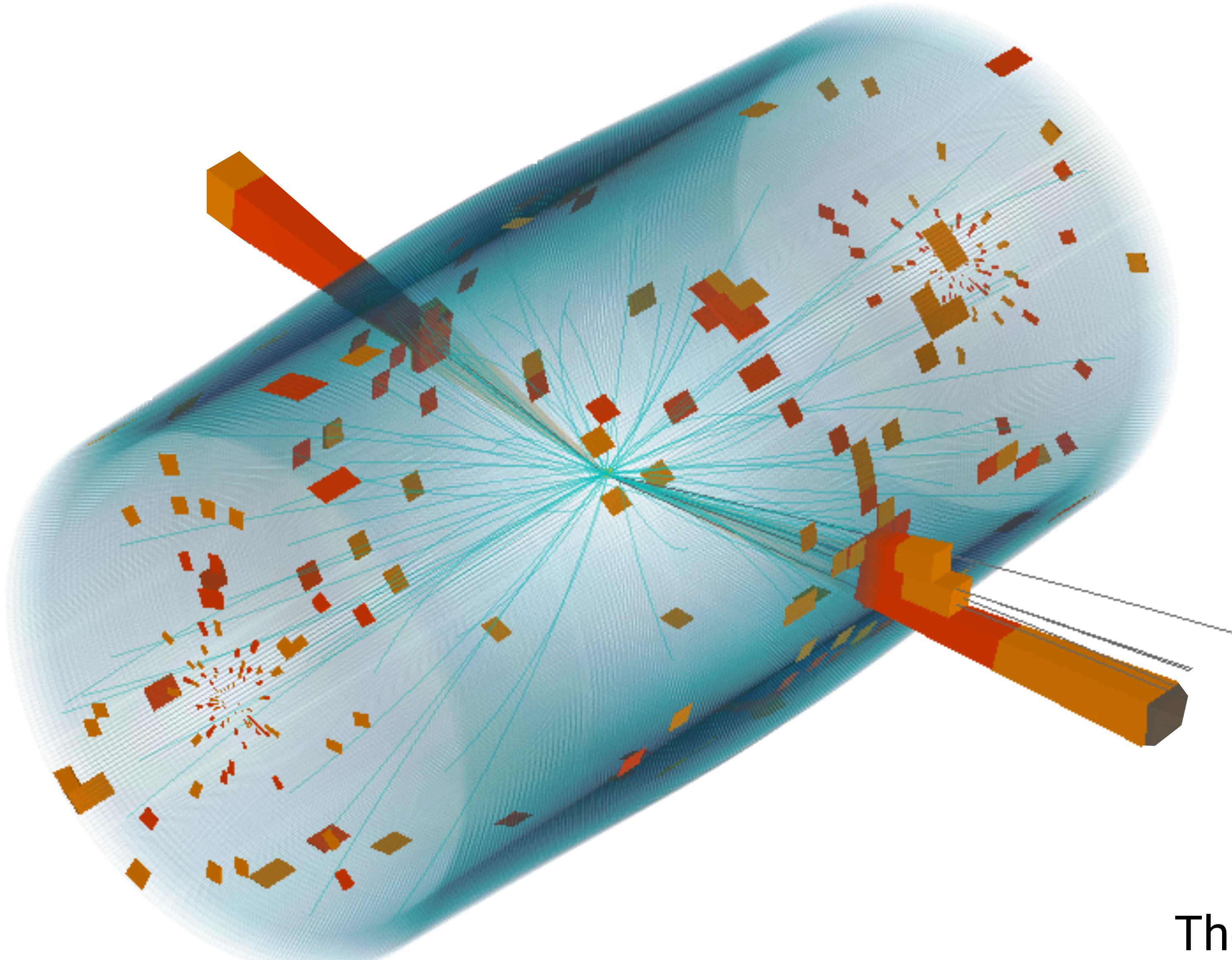
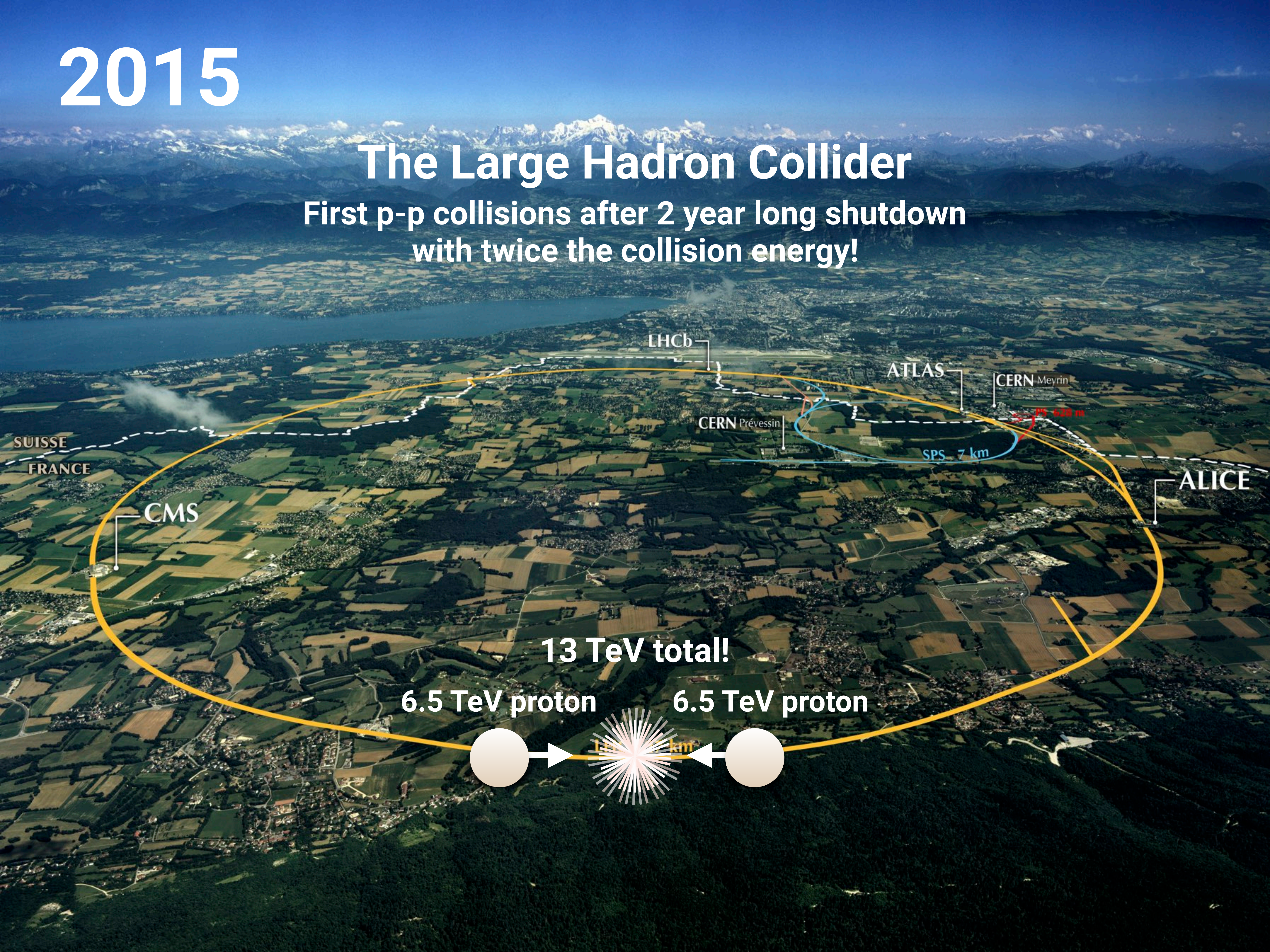
A Novel Multidimensional Search for Diboson Resonances
and
Encoding Jet Substructure with a Deep Neural Network



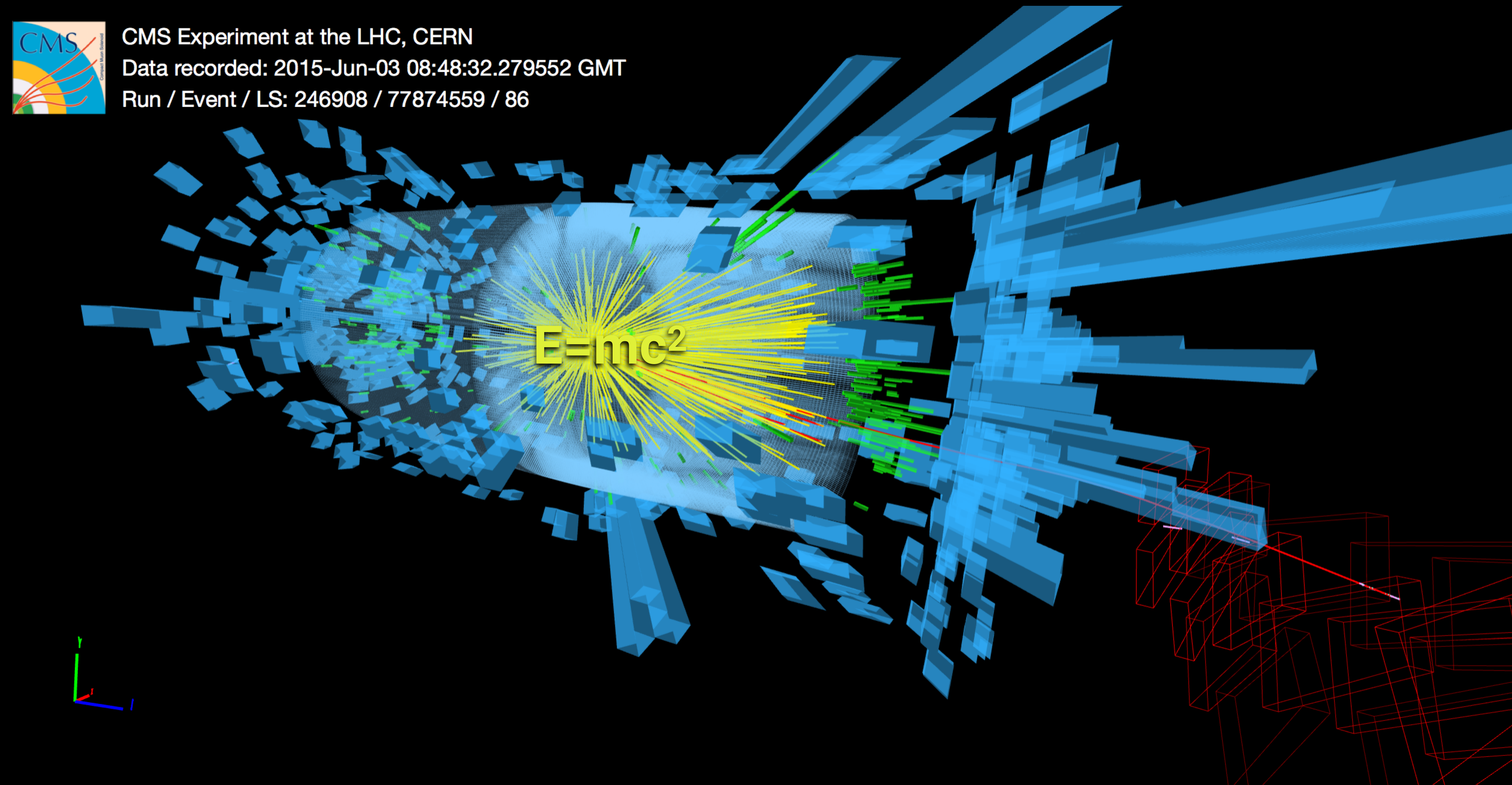
<u>PhD committee:</u>
Prof. Jesse Thaler (MIT)
Dr. Andreas Hinzmann (HUU)
Prof. Florencia Canelli (UZH)
Prof. Ben Kilminster (UZH, advisor)

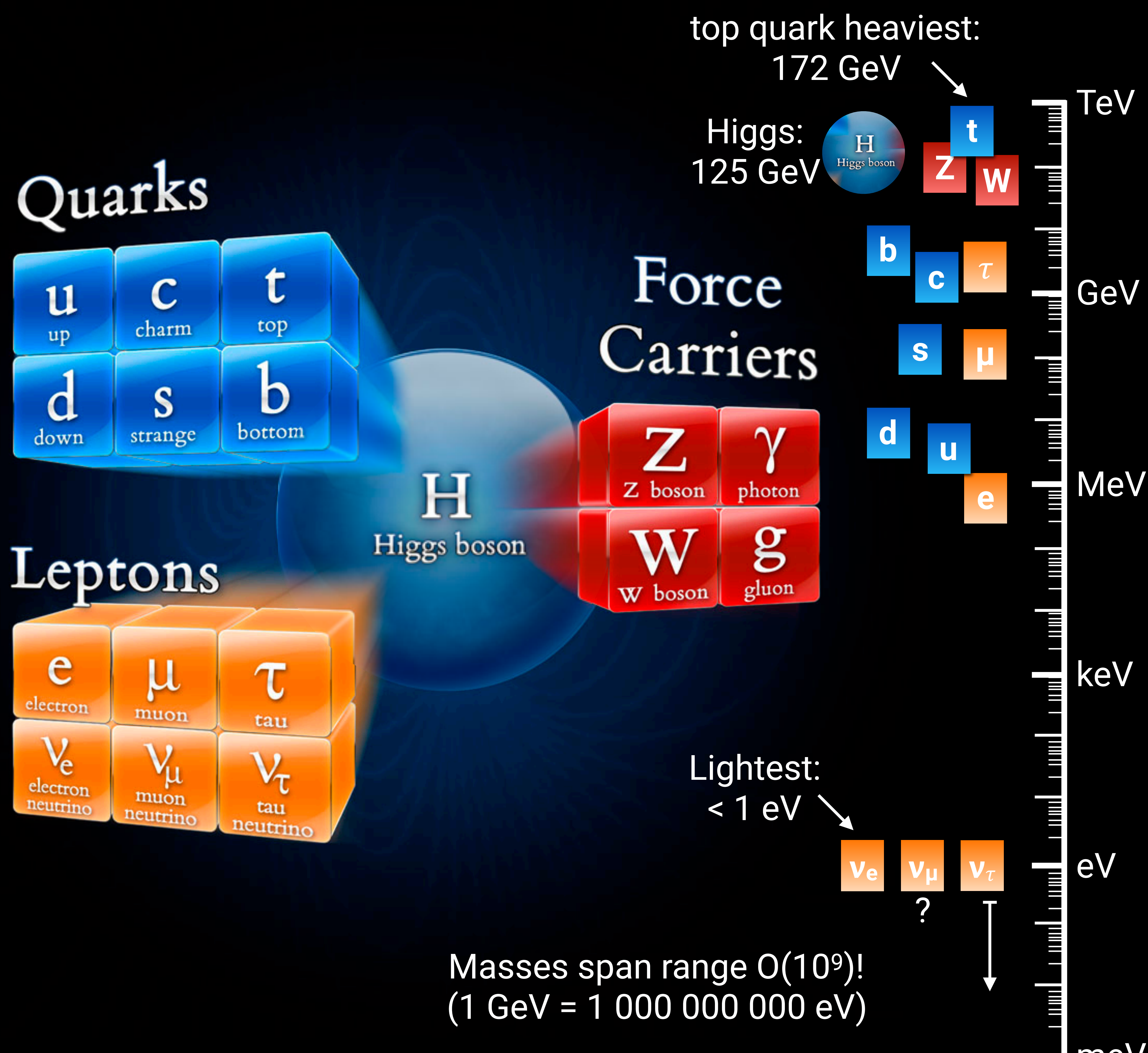
Thea Klæboe Årrestad

PhD defense, UZH
March 14th 2019
University of Zurich

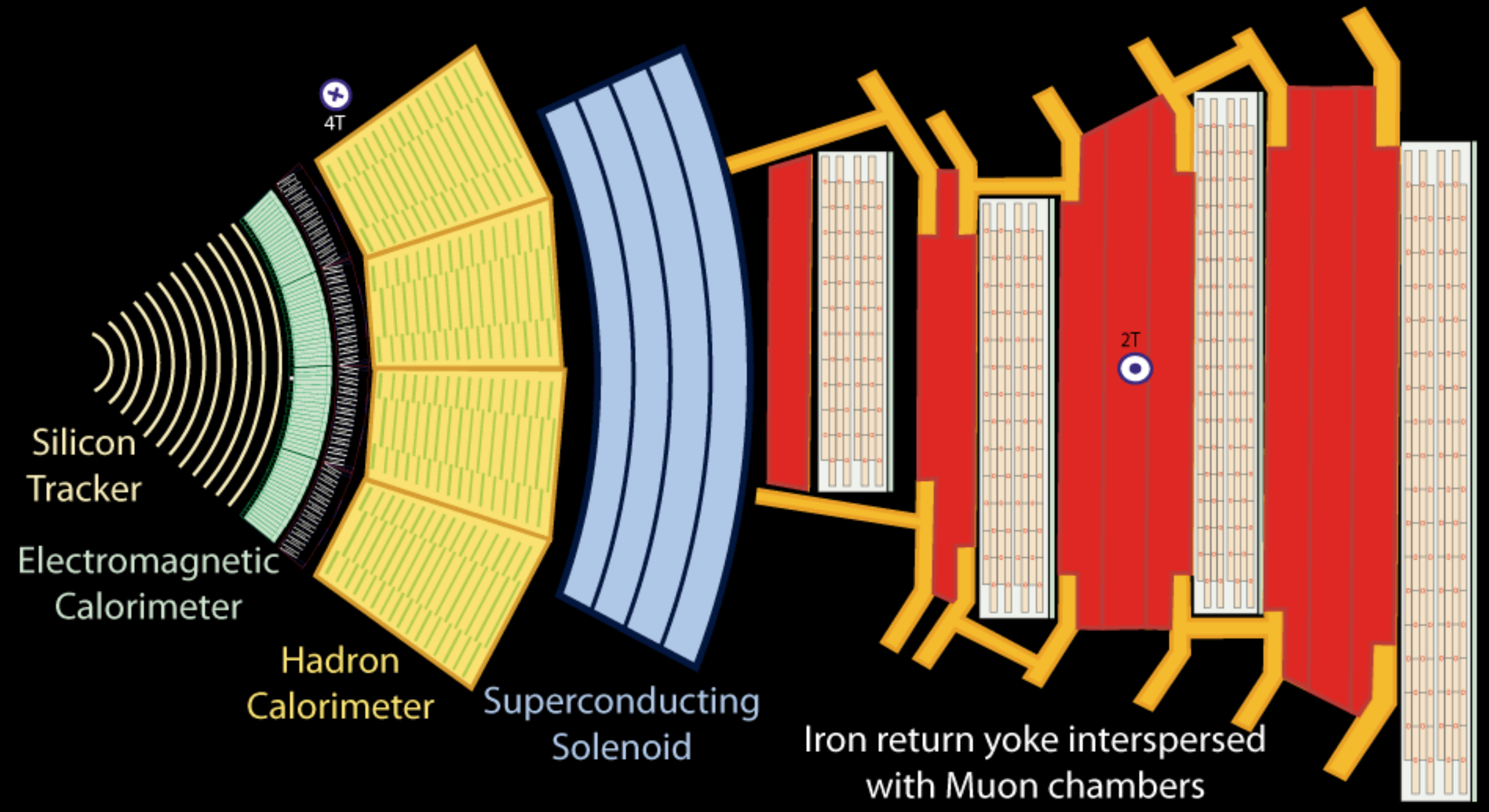

2015
The Large Hadron Collider
First p-p collisions after 2 year long shutdown
with twice the collision energy!
LHCb
ATLAS
CERN Meyrin
CERN Prévessin
SPS 7 km
SUISSE
FRANCE
CMS
ALICE
13 TeV total!
6.5 TeV proton
6.5 TeV proton

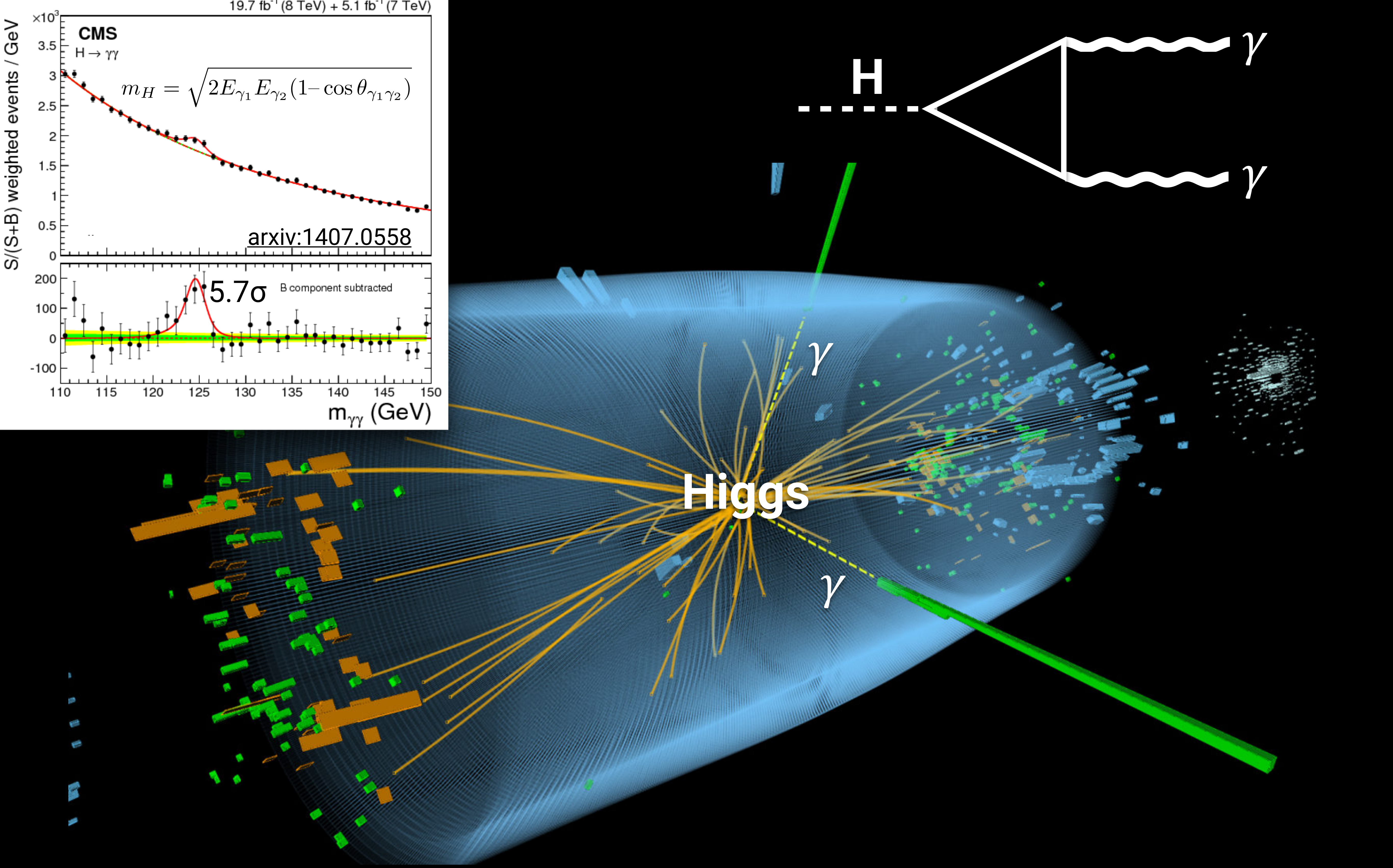

LHCb
ATLAS
CERN Meyrin
CERN Prévessin
PS 628 m
SPS 7 km
ALICE
SUISSE
FRANCE
CMS
LHC 27 km

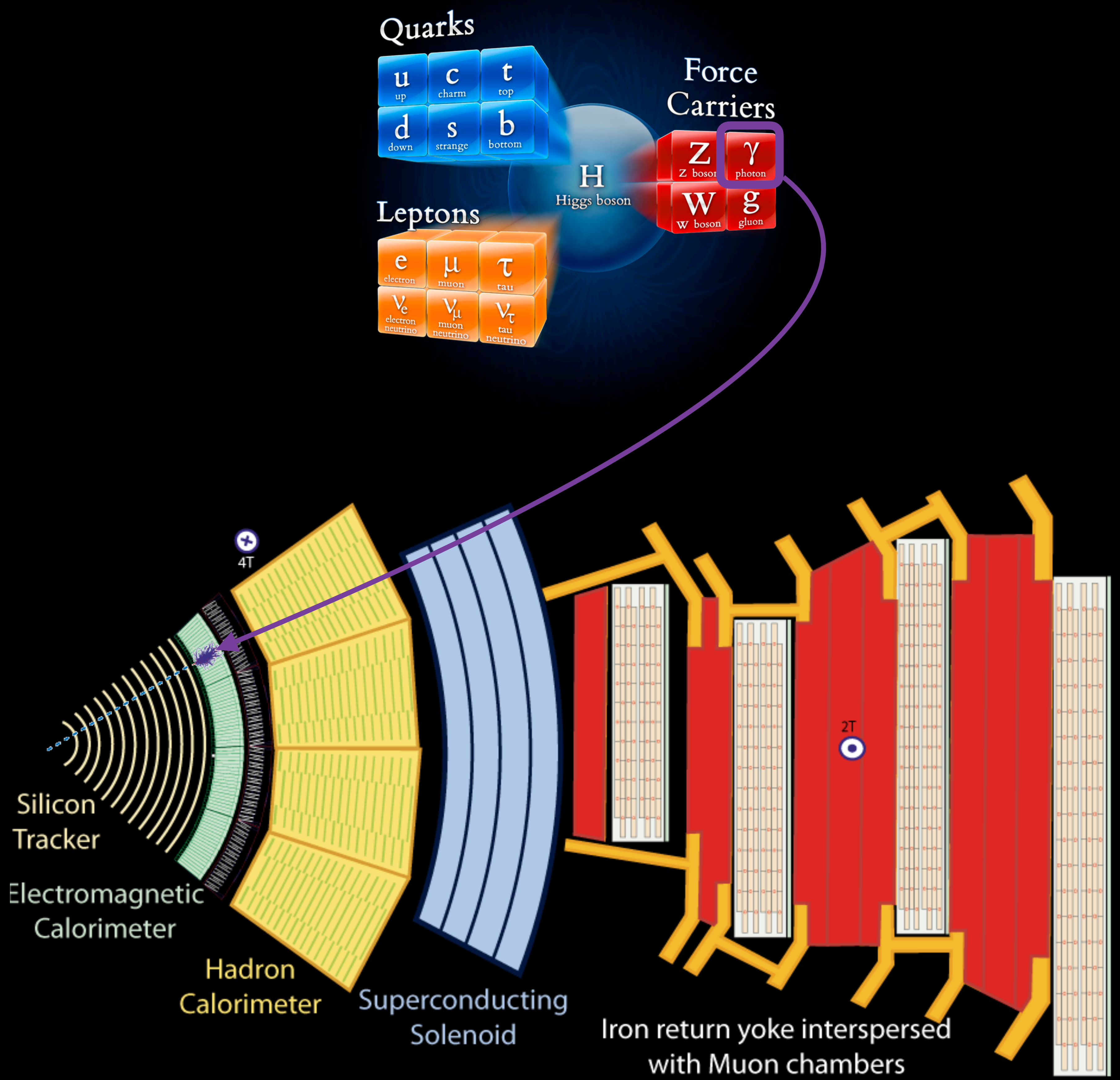

CMS
CMS Experiment at the LHC, CERN
Data recorded: 2015-Jun-03 08:48:32.279552 GMT
Run / Event / LS: 246908 / 77874559 / 86
E=mc²


Quarks
u up
c charm
t top
d down
s strange
b bottom
Leptons
e electron
μ muon
τ tau
νe electron neutrino
νμ muon neutrino
ντ tau neutrino
H Higgs boson
Force Carriers
Z Z boson
γ photon
W W boson
g gluon
top quark heaviest:
172 GeV
Higgs:
125 GeV
t
Z
W
b
c
τ
s
μ
d
u
e
Lightest:
< 1 eV
νe
νμ
ντ
?
TeV
GeV
MeV
keV
eV
meV
Masses span range O(10⁹)!
(1 GeV = 1 000 000 000 eV)


Quarks
u
up
c
charm
t
top
d
down
s
strange
b
bottom
Force
Carriers
Z
γ

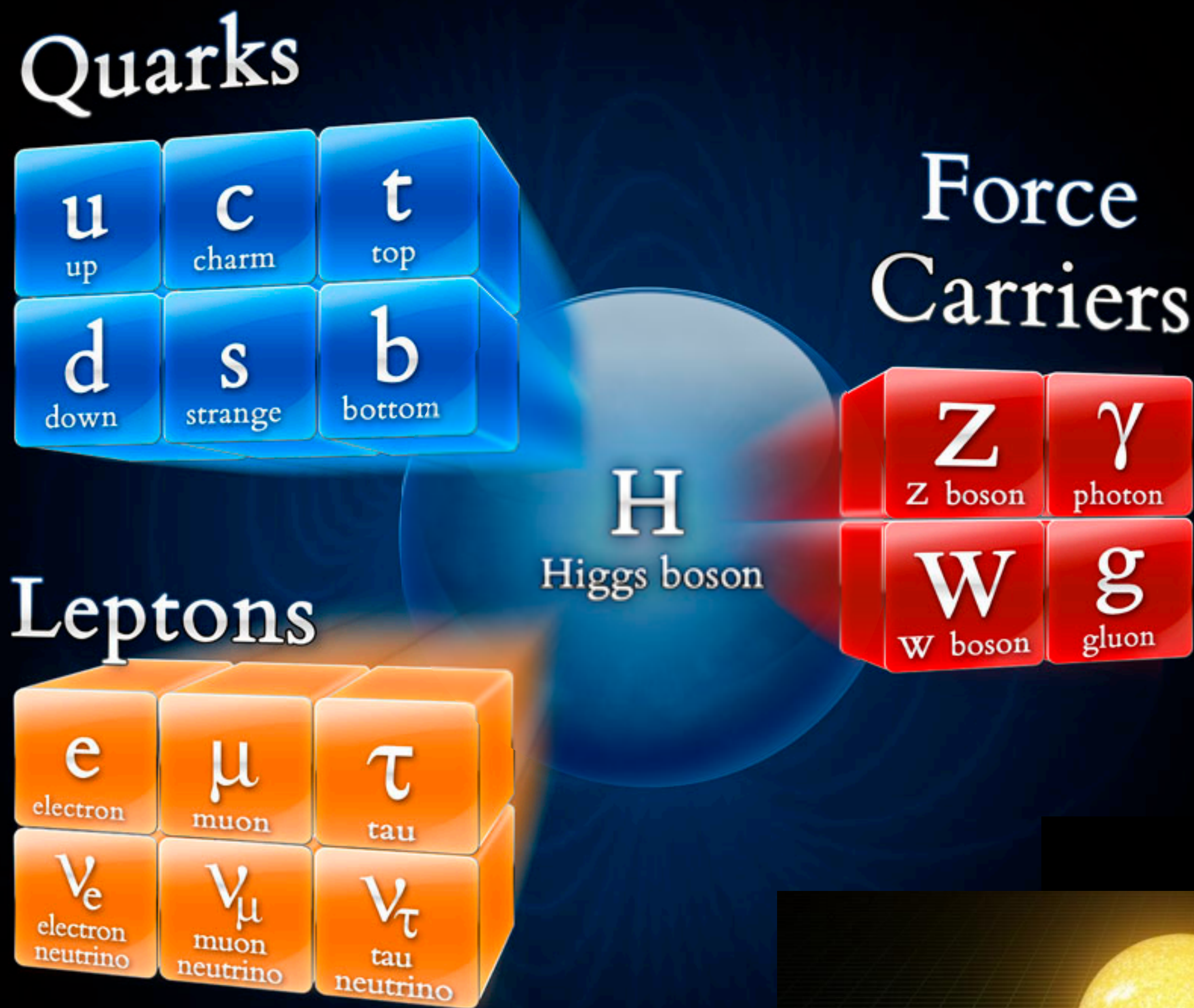

4T
2T
Silicon
Tracker
Electromagnetic
Calorimeter
Hadron
Calorimeter
Superconducting
Solenoid
Iron return yoke interspersed
with Muon chambers


19.7 fb⁻¹ (8 TeV) + 5.1 fb⁻¹ (7 TeV)
CMS
H → γγ
$m_H = \sqrt{2E_{\gamma_1}E_{\gamma_2}(1-\cos\theta_{\gamma_1\gamma_2})}$
arxiv:1407.0558
S/(S+B) weighted events / GeV
5.7σ
B component subtracted
$m_{\gamma\gamma}$ (GeV)
H
γ
γ
γ
γ
Higgs


Quarks
u up
c charm
t top
d down
s strange
b bottom
H Higgs boson
Force Carriers
Z Z boson
γ photon
W W boson
g gluon
Leptons
e electron
μ muon
τ tau
νe electron neutrino
νμ muon neutrino
ντ tau neutrino
4T
2T
Silicon Tracker
Electromagnetic Calorimeter
Hadron Calorimeter
Superconducting Solenoid
Iron return yoke interspersed with Muon chambers

Why is the Higgs mass we MEASURED so much smaller ($\times 10^{16}$) than the Higgs mass we CALCULATED

?

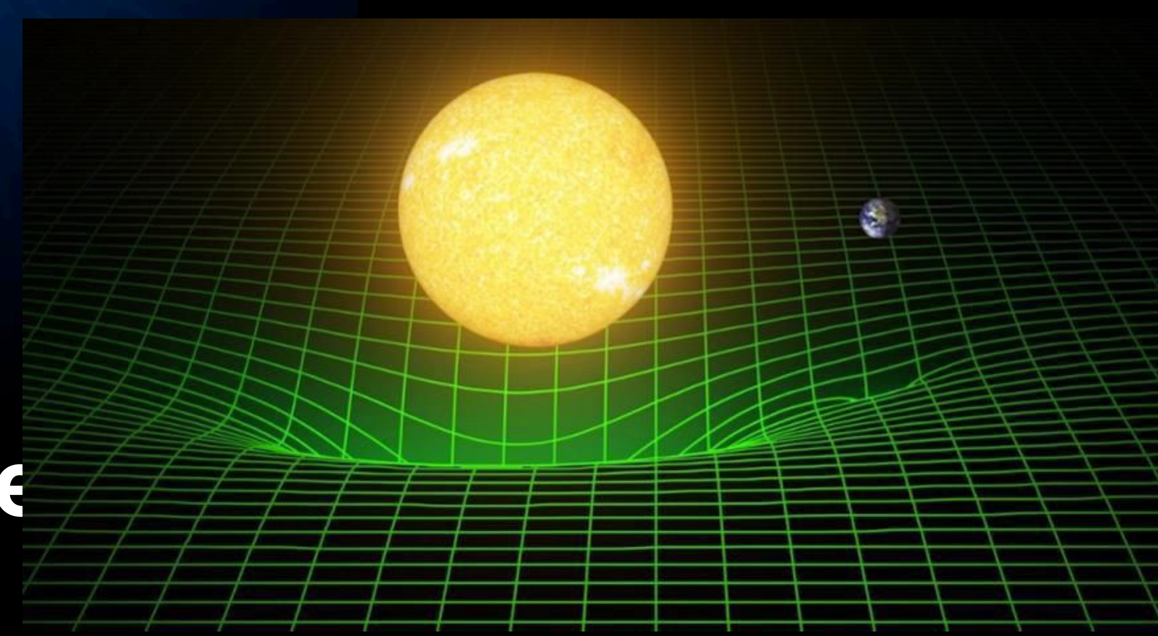


Why is gravity 10,000,000, 000,000,000, 000,000,000, 000,000,000, 000,000 weaker than the other forces*

?

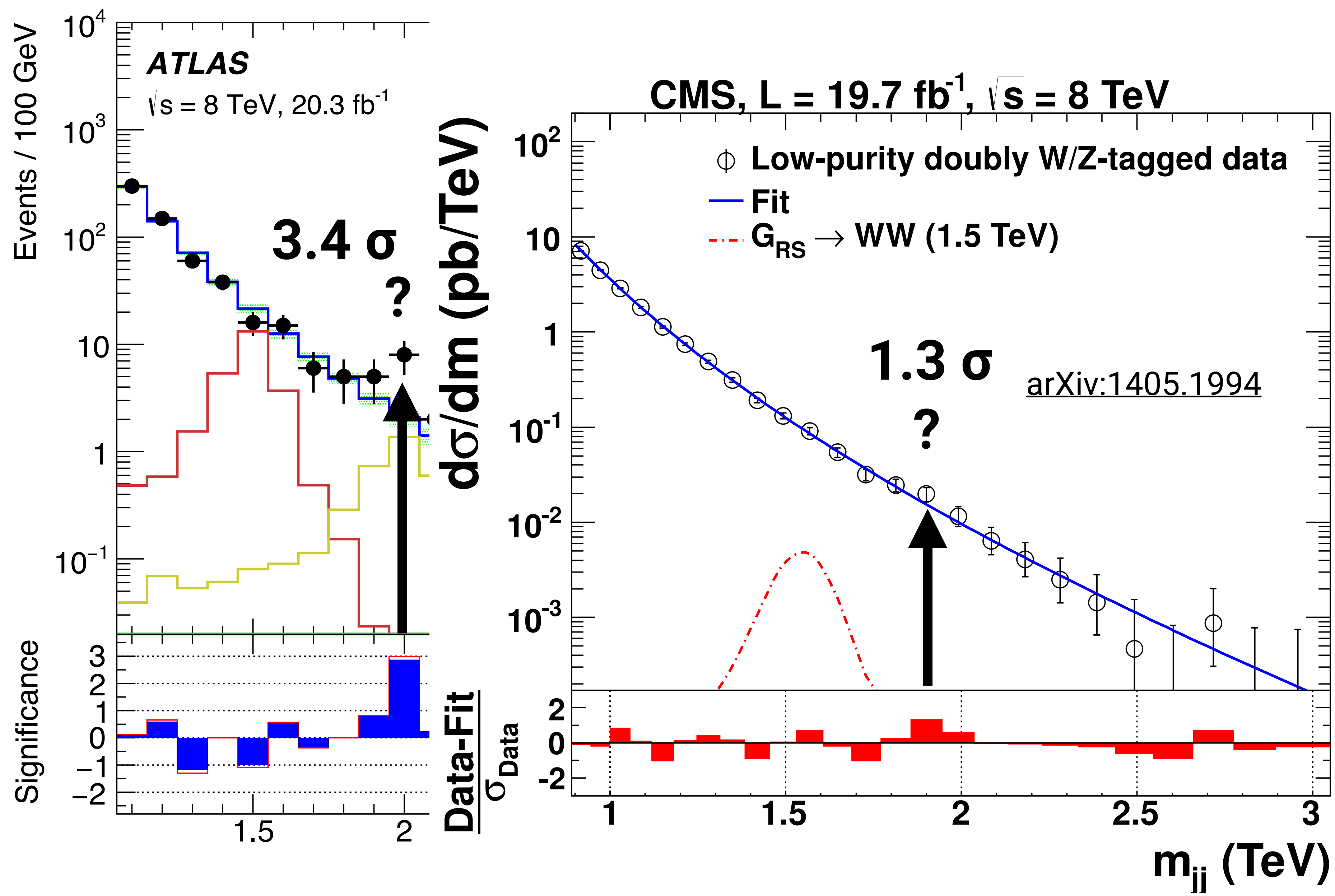
Higgs presicion measurements

be



“Search for high-mass diboson resonances with boson-tagged jets“

ATLAS Collaboration



A **new particle** with a mass of 2 TeV?

Compatible observations in ATLAS and CMS!

What could it be?

10. arXiv:1506.07511 [pdf, other] hep-ph hep-ex doi 10.1103/PhysRevD.92.055030

G221 Interpretations of the Diboson and Wh Excesses

Authors: Yu Gao, Tathagata Ghosh, Kuver Sinha, Jiang-Hao Yu

Submitted 17 November, 2015; v1 submitted 24 June, 2015; originally announced June 2015.

Comments: 16 pages, 6 figures, revised version 2. With a 2.9 TeV Z' event observed in CMS, we updated our paper by discussing the possible 2.9 TeV Z' signature together with the 2 TeV W' excess

Report number: UTTG-11-15, MI-TH-1522, CETUP2015-007

Journal ref: Phys. Rev. D 92, 055030 (2015)

11. arXiv:1507.00268 [pdf, other] hep-ph doi 10.1103/PhysRevD.92.095025

Simple non-Abelian extensions of the standard model gauge group and the diboson excesses at the LHC

Authors: Qing-Hong Cao, Bin Yan, Dong-Ming Zhang

Submitted 27 November, 2015; v1 submitted 1 July, 2015; originally announced July 2015.

Comments: Publish version; title changed as suggested by journal Editor

Journal ref: Phys. Rev. D 92, 095025 (2015)

5. arXiv:1507.03098 [p

A scalar hint from the diboson excess?

Authors: Giacomo Cacciapaglia, Aldo Deandrea, Michio Hashimoto

cussions are clarified. References added. Matches publishe

6. arXiv:1507.03553 [pdf, other] hep-ph doi 10.1155/2016/3279568

On the compatibility of the diboson excess with a gg-initiated composite sector

1. arXiv:1506.06739 [pdf, ps, other] hep-ph hep-ex

Triboson interpretations of the ATLAS diboson excess

Authors: J. A. Aguilar-Saavedra

Submitted 25 September, 2015; v1 submitted 22 June, 2015; originally announced June 2015.

Comments: LaTeX 17 pages. v2: Enlarged discussion to address CMS WH excess. v3: Added discussion of diboson helicities. Final version to appear in JHEP

other] hep-ph doi 10.1103/PhysRevD.92.055001

he ATLAS Diboson Resonances

tsumi Nagata, Yuji Omura

5; v1 submitted 12 June, 2015; originally announced June 2015.

es; version accepted for publication in Physical Review D

Report number: IPMU15-0083, FTPI-MINN-15/31

Journal ref: Phys. Rev. D 92, 055001 (2015)

7. ar

Interpretations of the ATLAS Diboson Anomaly

Authors: Kingman Cheung, Wai-Yee Keung, Po-Yan Tseng, Tzu-Chiang Yuan

Submitted 17 November, 2015; v1 submitted 19 June, 2015; originally announced June 2015.

Comments: v4: match the published version; v3: 18 pages, 6 figures, change to leptophobic Z' model to take into account the EW constraints, and some updates to the analysis and text; v2: 17 pages, 7 figures; a new section and a new figure are added; correct the statement about the WH; references are also added

tensions of the standard model gauge group and the diboson excesses at the LHC

Yan, Dong-Ming Zhang

submitted 1 July, 2015; originally announced July 2015.

007/JHEP02(2016)084

Prospects for Spin-1 Resonance Search at 13 TeV LHC and the ATLAS Diboson Excess

Authors: Tomohiro Abe, Teppei Kitahara, Mihoko M. Nojiri

Submitted 22 January, 2016; v1 submitted 7 July, 2015;

7. arXiv:1507.04431 [pdf, ps, other] hep-ph hep-ex doi 10.1016/j.physl

2 TeV Higgs boson and diboson excess at the LHC

Authors: Chuan-Hung Chen, Takaaki Nomura

Submitted 18 August, 2015; v1

Comments: 12 pages, 7 figure

16. arXiv:1604.03578 [pdf, other] hep-ph

A model for the LHC diboson excess

Authors: Manuel Buen-Abad, Andrew G. Cohen, Martin Schmaltz

Submitted 25 April, 2016; v1 submitted 12 April, 2016; originally announced April 2016.

12. arXiv:1507.00900 [pdf, other] hep-ph hep-ex

Unitarity implications of diboson resonance physics

Authors: Giacomo Cacciapaglia, Mads T. Frandsen

Submitted 3 July, 2015; originally announced July 2015.

Comments: 5 pages, 2 figures

Journal ref: Phys. Rev. D 92, 055035 (2015)

Left-Right Dark Matter

Authors: Debasish Borah, Arna

Submitted 5 December, 2017; v1 su

Comments: Version 2: 26 pages, 8 fi

Journal ref: Phys. Rev. D 96, 115019

11. arXiv:1507.06018 [pdf, other] hep-ph

Low Scale Composite Higgs Model and 1.8 ~ 2 TeV Diboson E

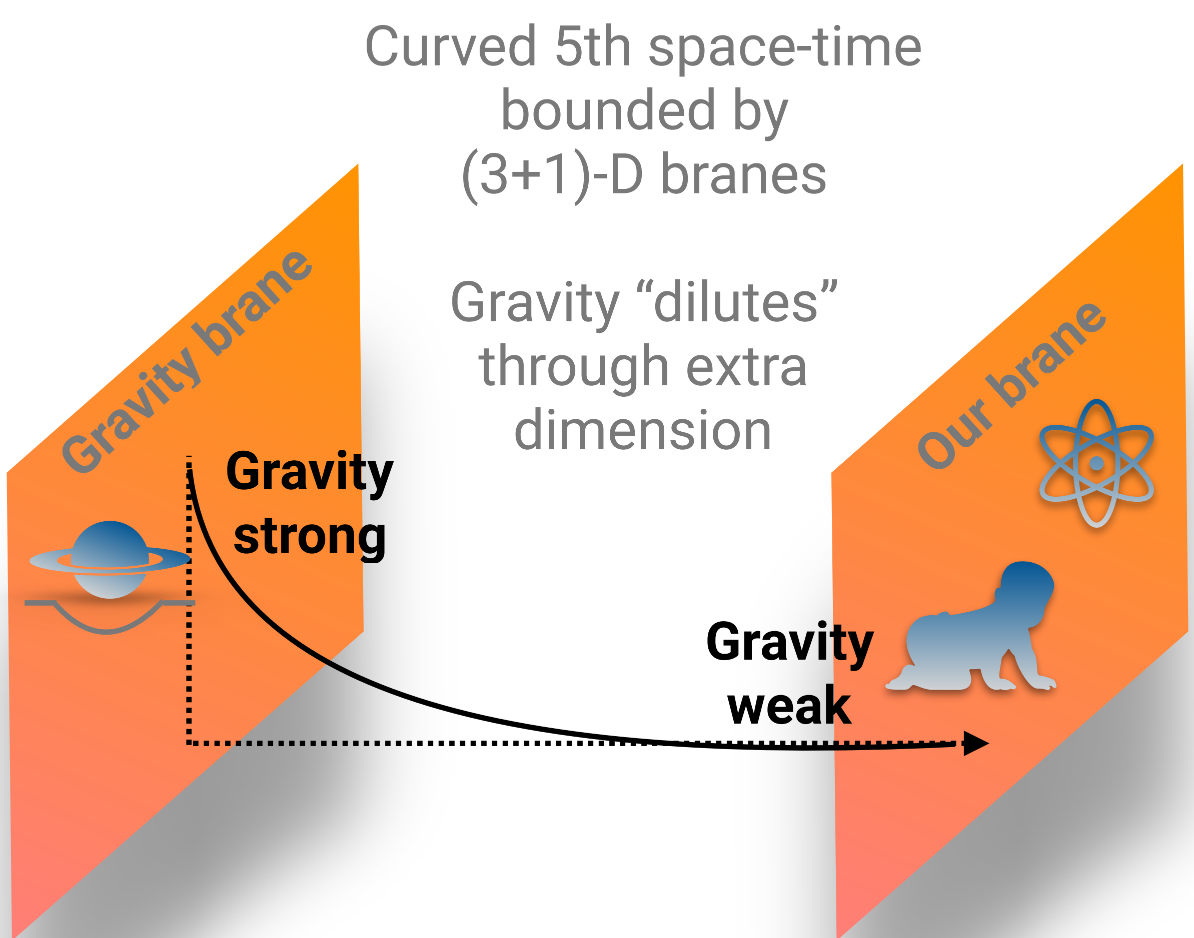
Authors: Ligong Bian, Da Liu, Jing Shu

Submitted 21 July, 2015; originally announced July 2015.

What could it be?

Why is gravity weak?

Warped Extra Dimensional theories



Predicts:

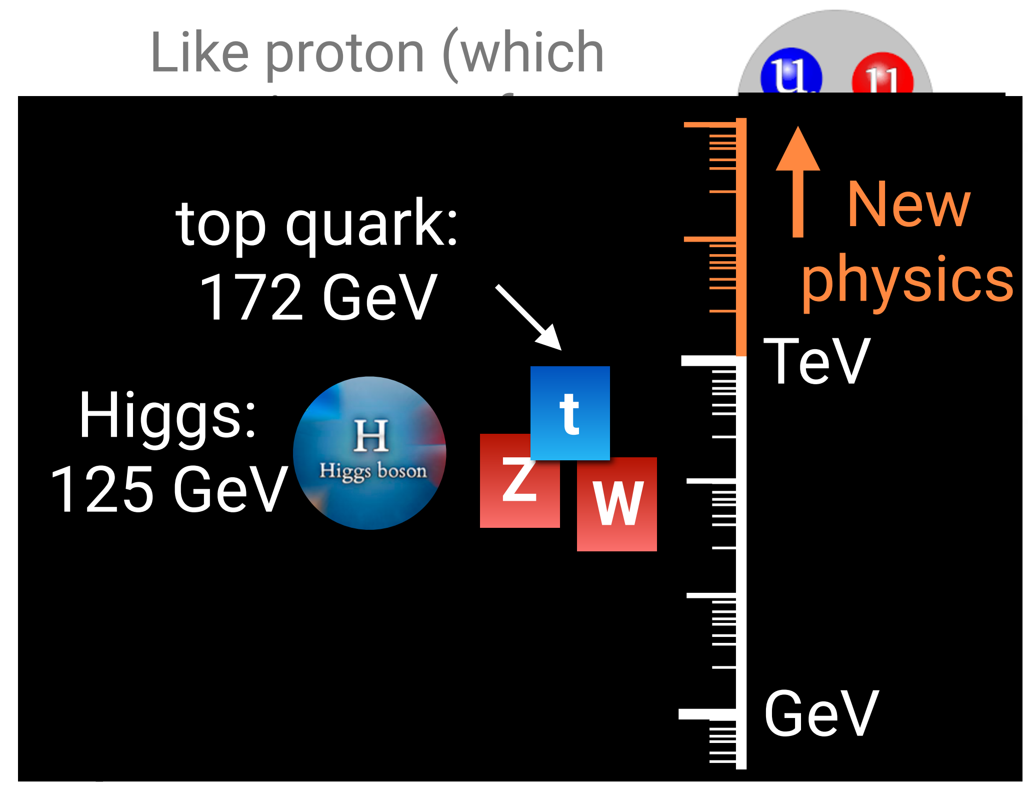
Heavy graviton resonances:

Signature: G_{Bulk}→WW and G_{Bulk}→ZZ

Why is Higgs light?

Composite Higgs models

Like proton (which



$\rho^{\pm} \rightarrow \pi^{\pm} \pi^{0}$ ⟷ W' → WZ

Predicts:

Heavy (~TeV) copies of SM particles:

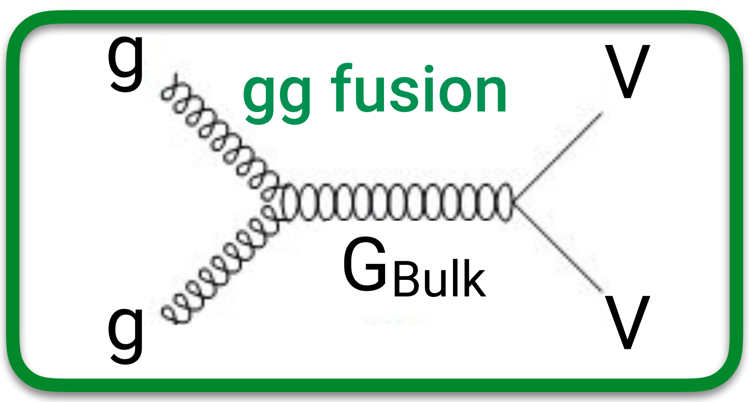
Signature: Z'→WW and W'→WZ

Partonic luminosity

Going from 8→13 TeV: partonic luminosity increases!

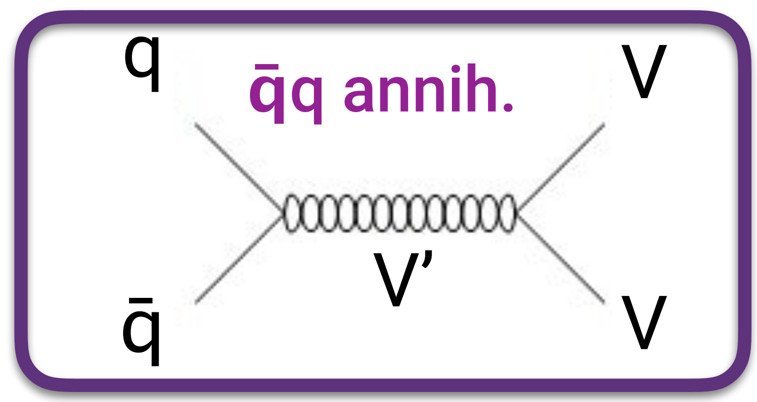
G_{Bulk}→WW/ZZ

mainly produced through gg fusion!



V'→WW/WZ

mainly produced through $\bar{q}q$ annihilation



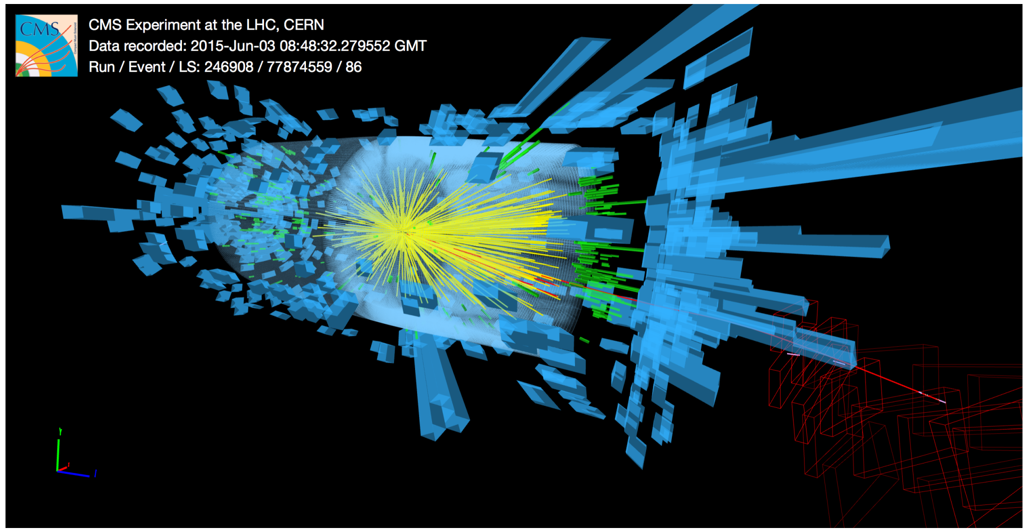
m_X	Production	13/8 TeV signal yield
2 TeV	$\boxed{gg}$	x 15
2 TeV	$\boxed{qq}$	x 8

15 x
8 x
more signal events at 13 TeV!

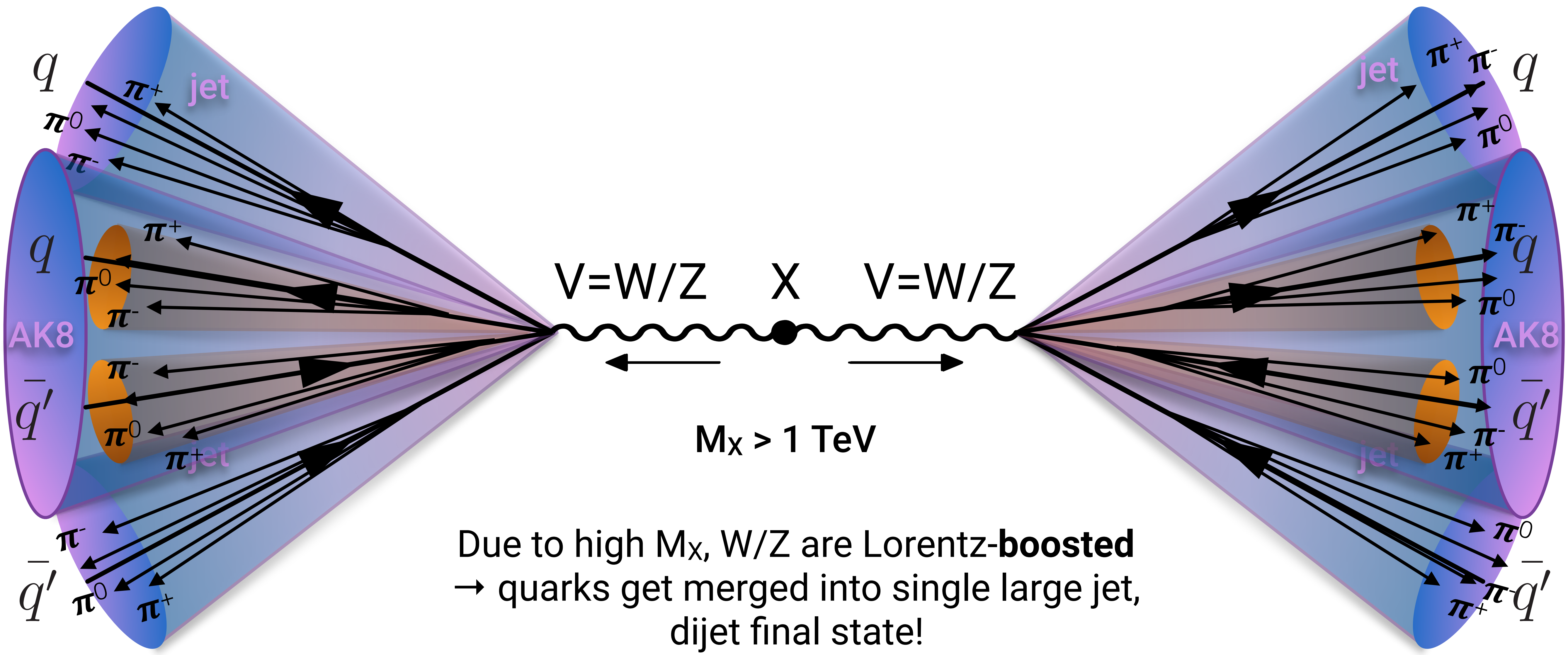
Same discovery potential as 8 TeV dataset with only 1/7th of 13 TeV data!

Thesis work:
Diboson resonance searches at 13 TeV with CMS

- I : First search for diboson resonances at 13 TeV
- II : A new pileup-resistant, perturbative robust tagger
- III: A novel framework for multi-dimensional searches
- IV: Encoding jet substructure with a deep neural network



Signature: X→VV→4q

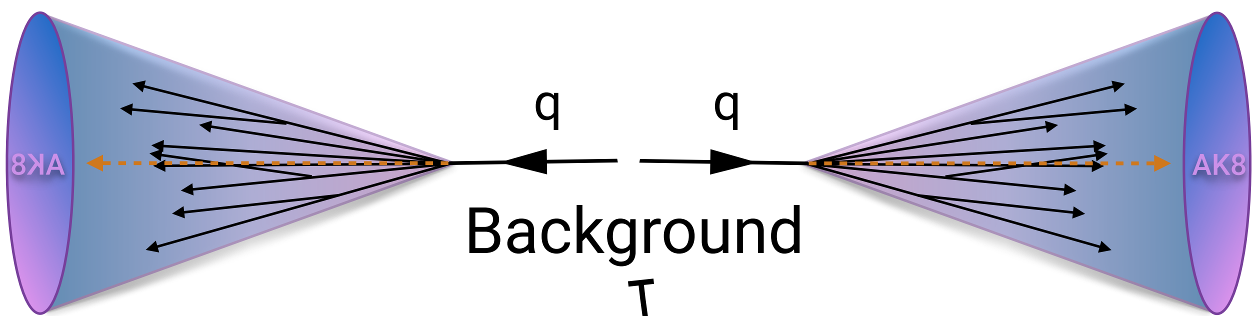


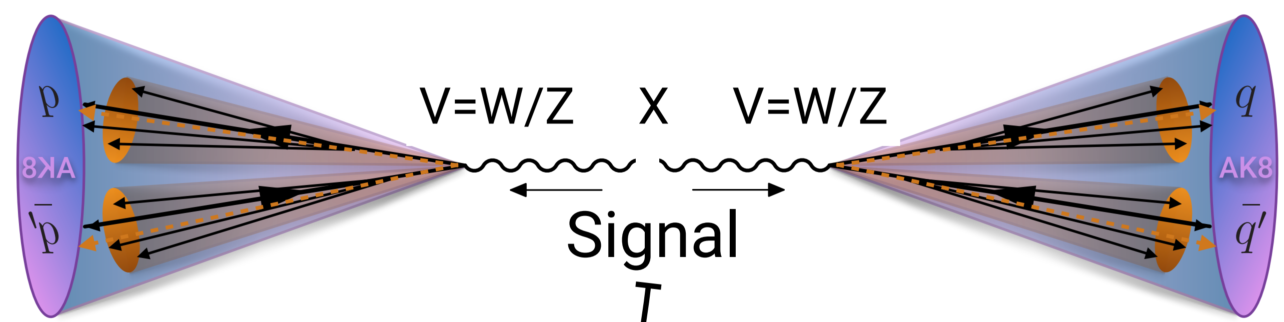
M_X > 1 TeV

Due to high M_X, W/Z are Lorentz-**boosted**
→ quarks get merged into single large jet,
dijet final state!

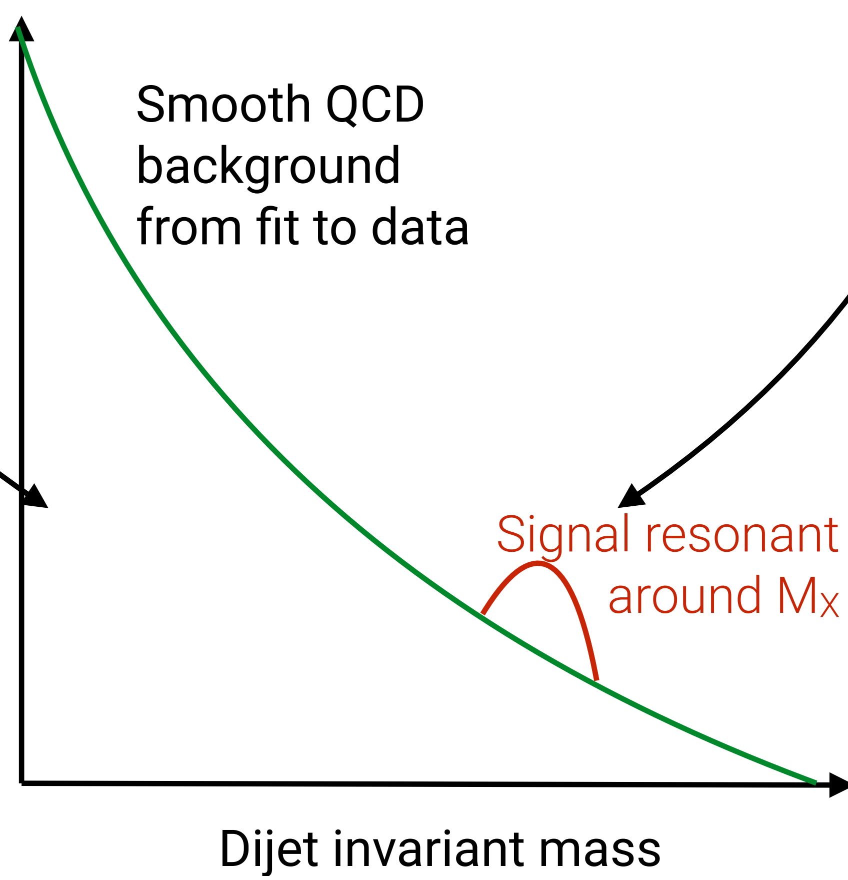
Analysis strategy

1. Reconstruct two W/Z jets, discriminate them from the quark/gluon jet QCD background





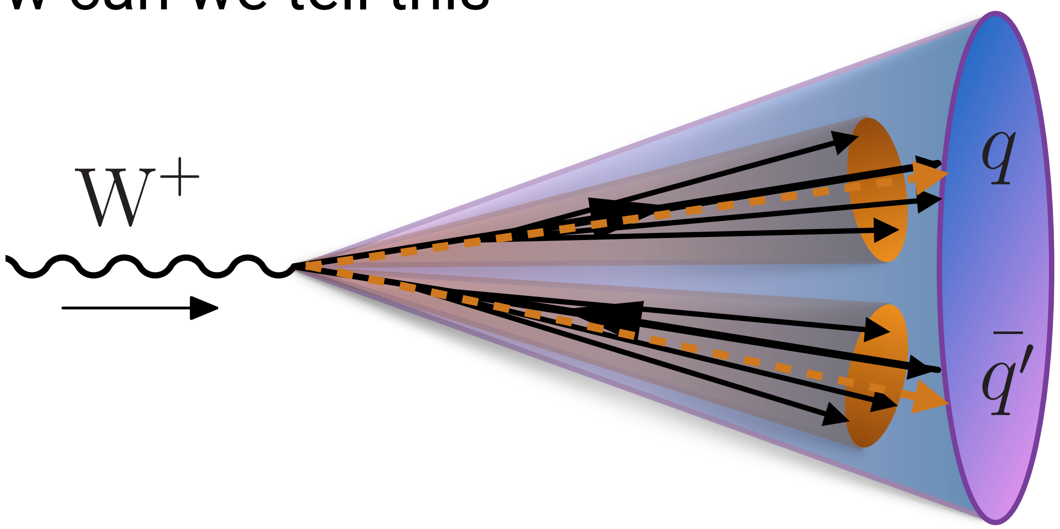
2. “Bump hunt” in dijet invariant mass spectrum
   - signal resonant, QCD smoothly falling



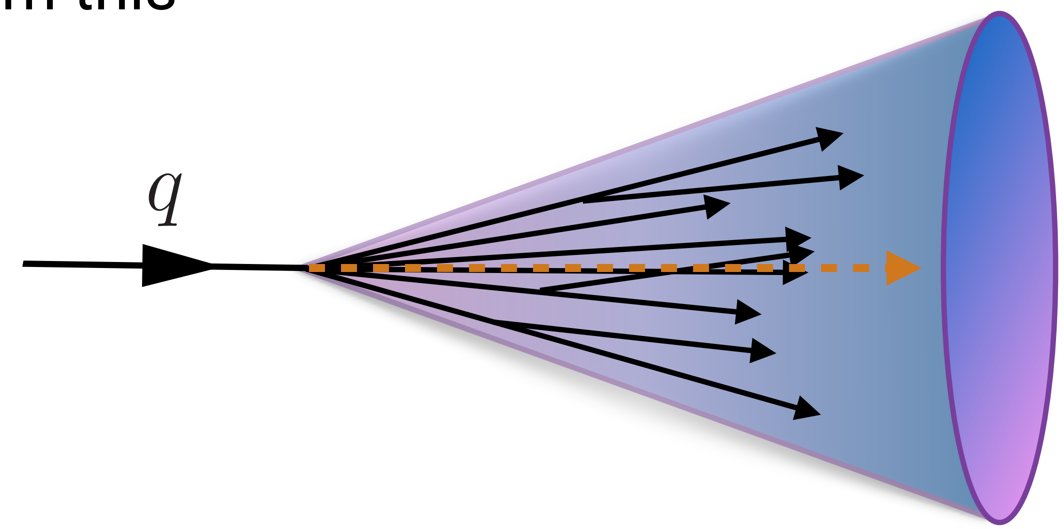
3. Estimate background from fit to data in signal region
   - smoothness test of observed data (no MC!)

Jet substructure methods

How can we tell this

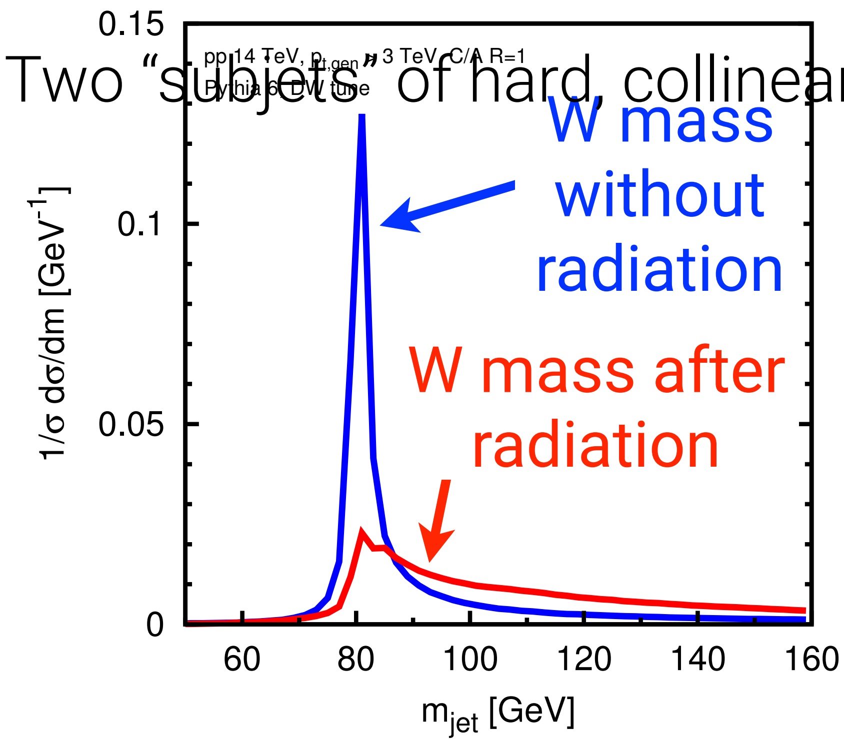


from this

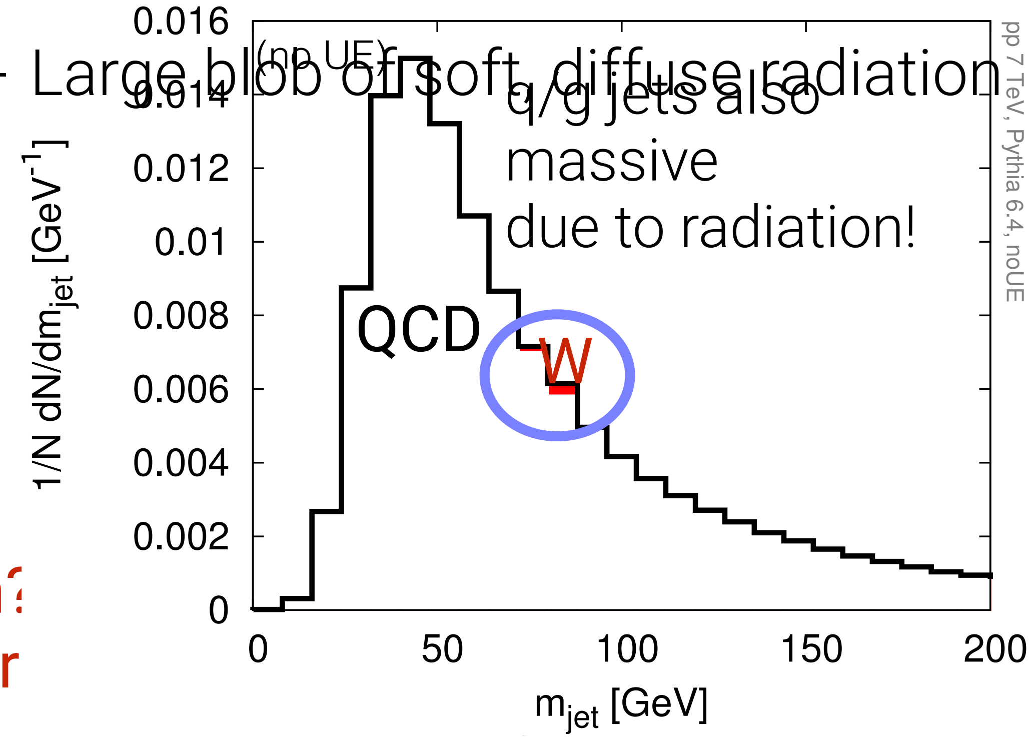


?

- Mass = 80 GeV?
- Two “subjets” of hard, collinear particles

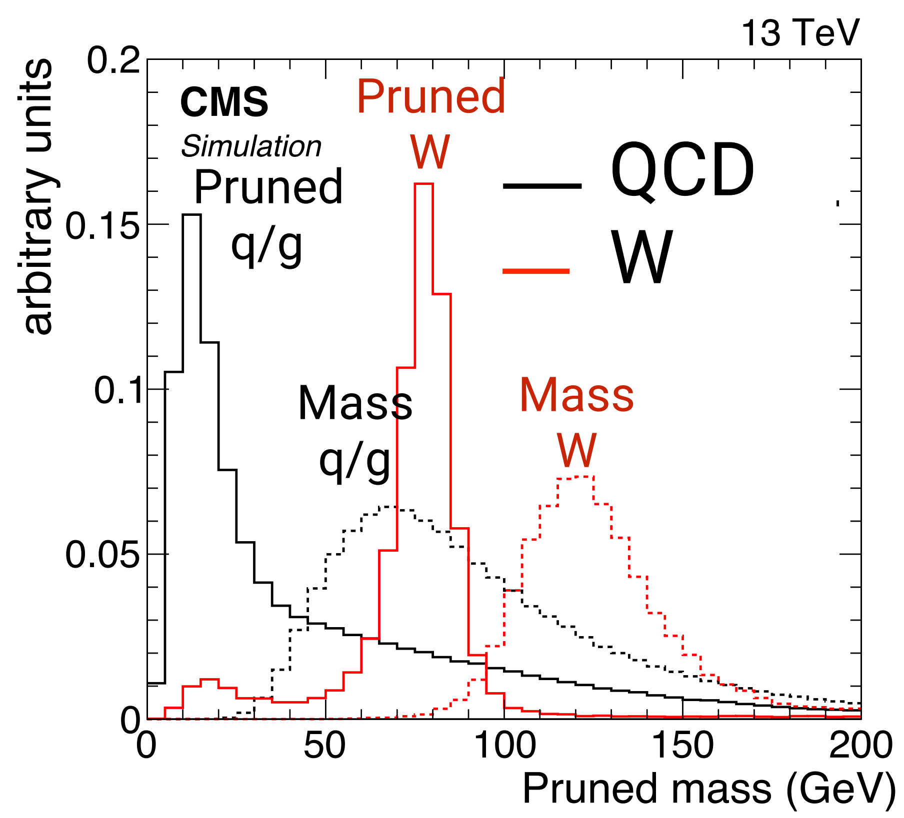


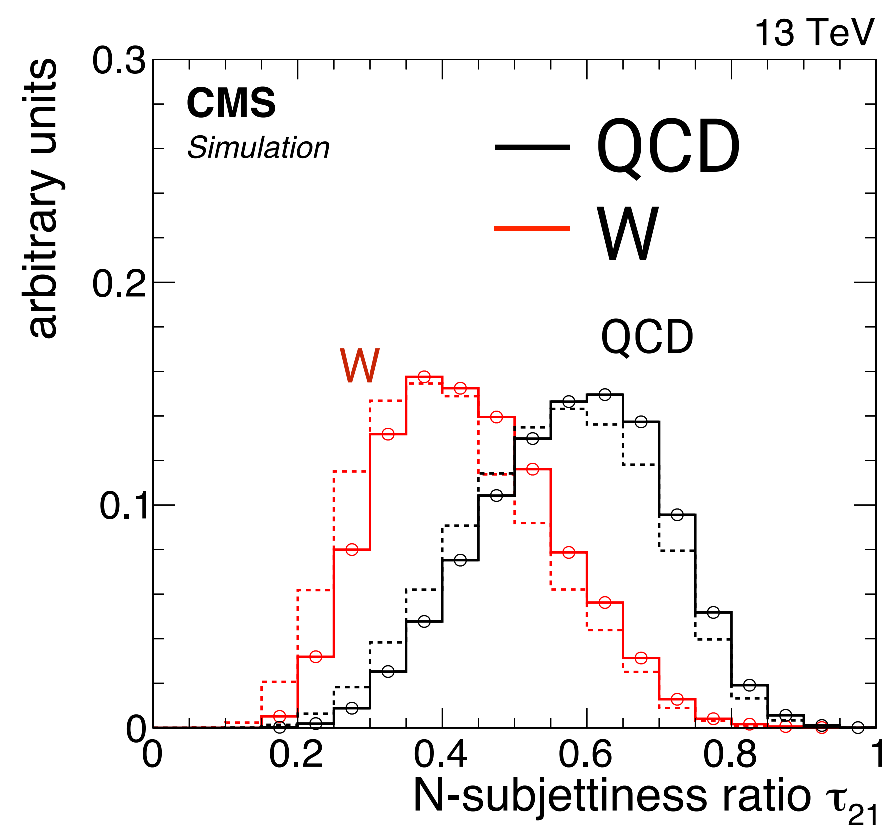
- Mass ≃ 0 GeV?
- Large blob of soft, diffuse radiation



Radiation?
Substructur

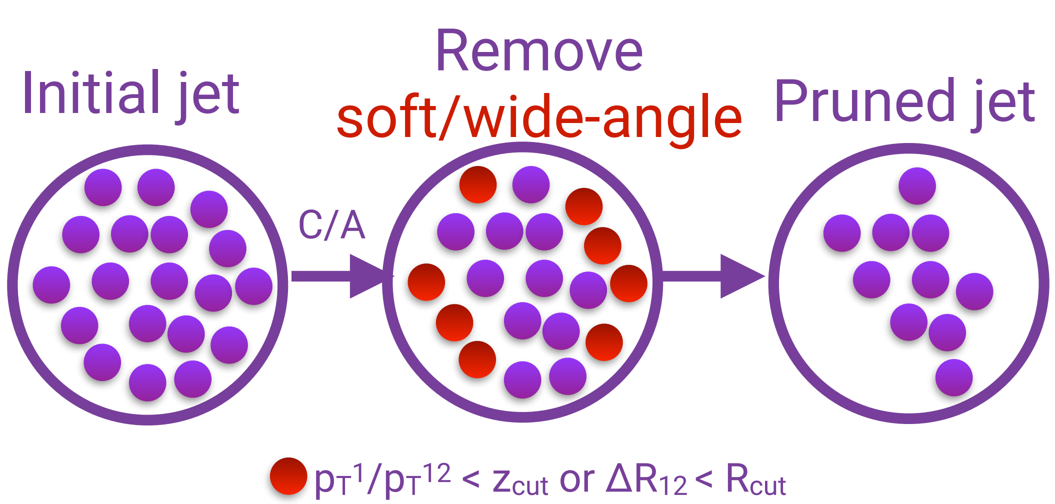
Jet substructure methods





Jet mass resolution → Pruning

arxiv:0912.0033

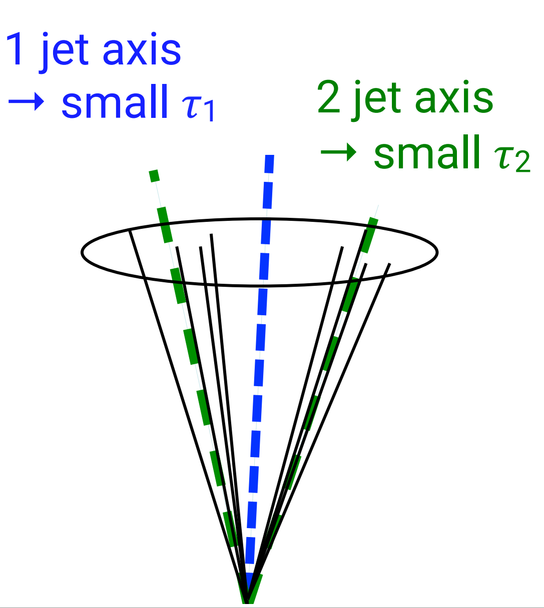


Are there "subjets" → n-subjettiness

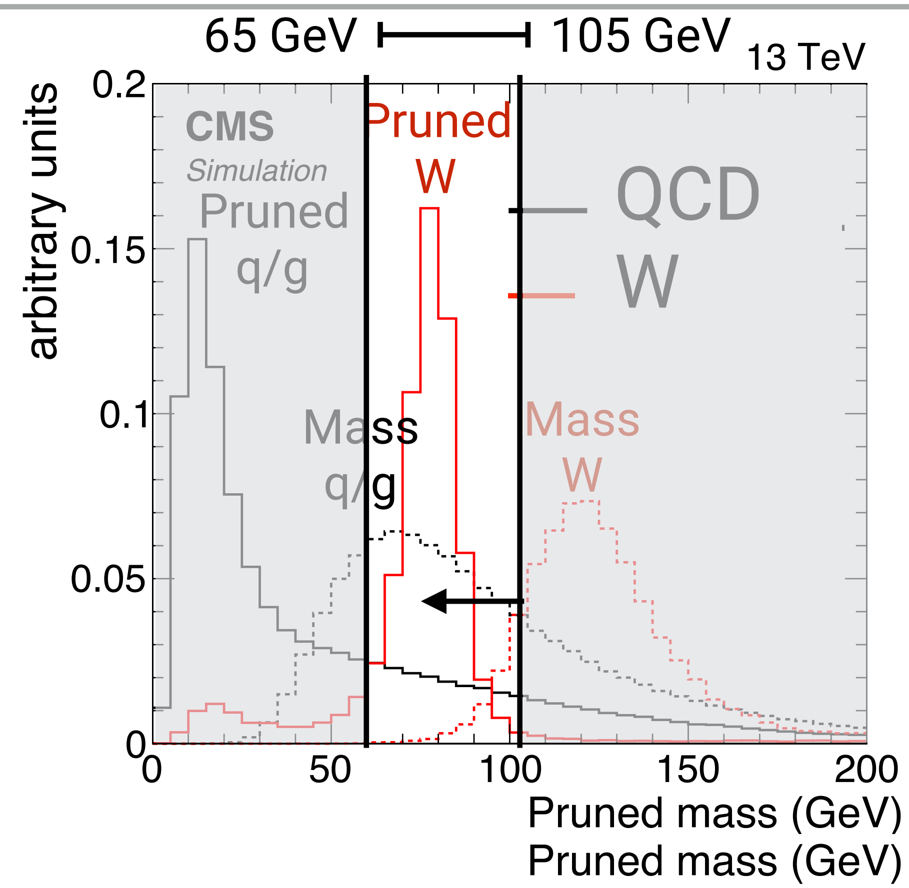
arxiv:1011.2268

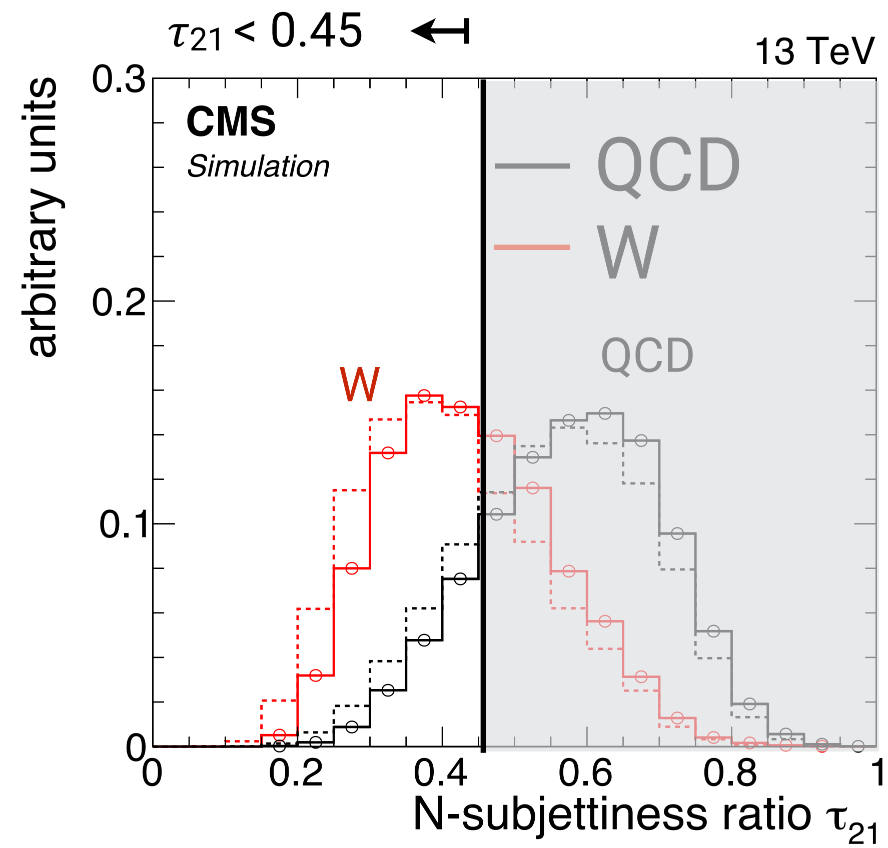
Probability of jet having N subjets, τ_N

- use ratio τ_2/τ_1



Jet substructure methods





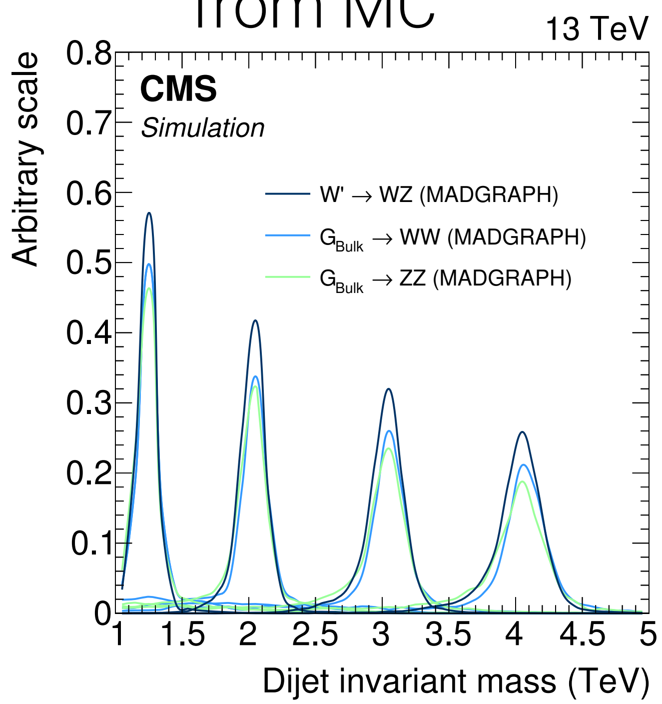
W/Z-tagging:

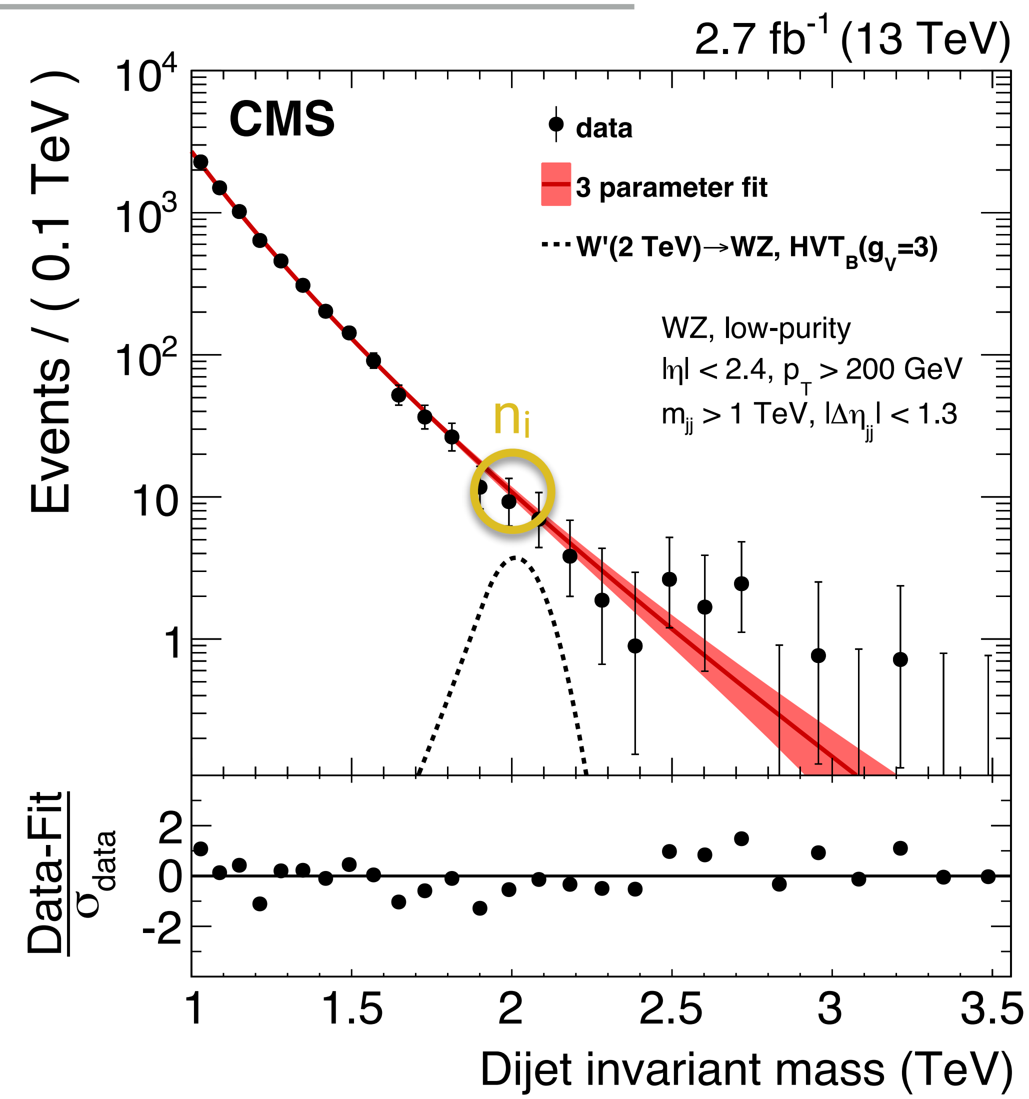
65 GeV < M_{pruned} < 105 GeV
+
τ_{21} < 0.45

~55% efficiency
at 1-2% mistag rate

Statistical interpretation

Signal PDF extracted from MC





Background

described by smooth fit to data, yield is estimated from B comp. of best S+B fit

$$\mu_i = \sigma \cdot N_i(S) + N_i(B)$$

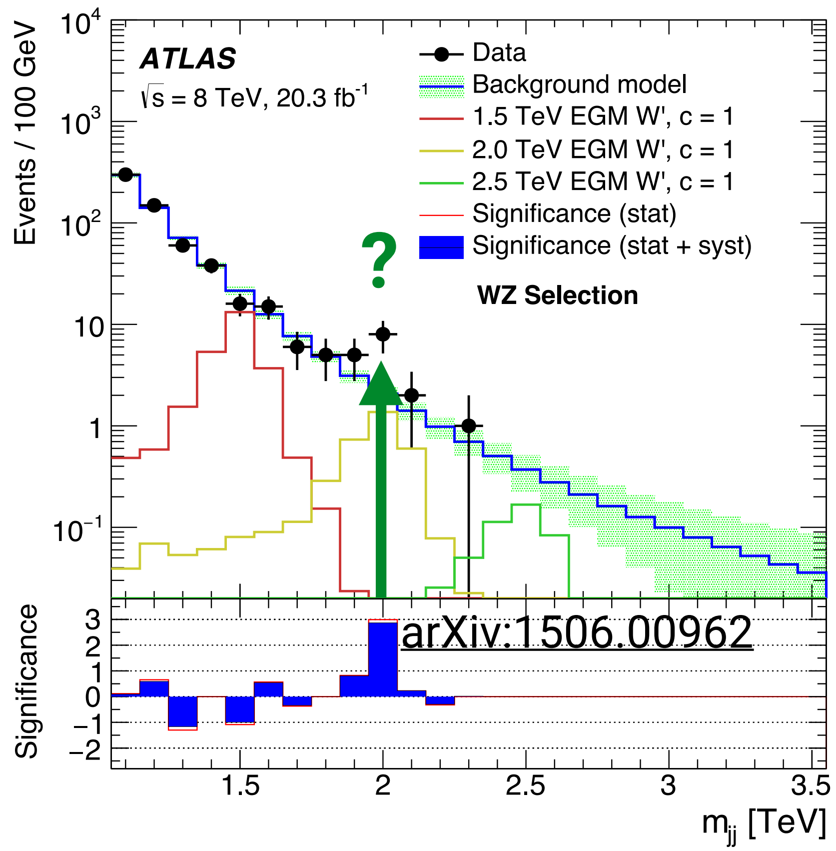
Observed events in bin n_i

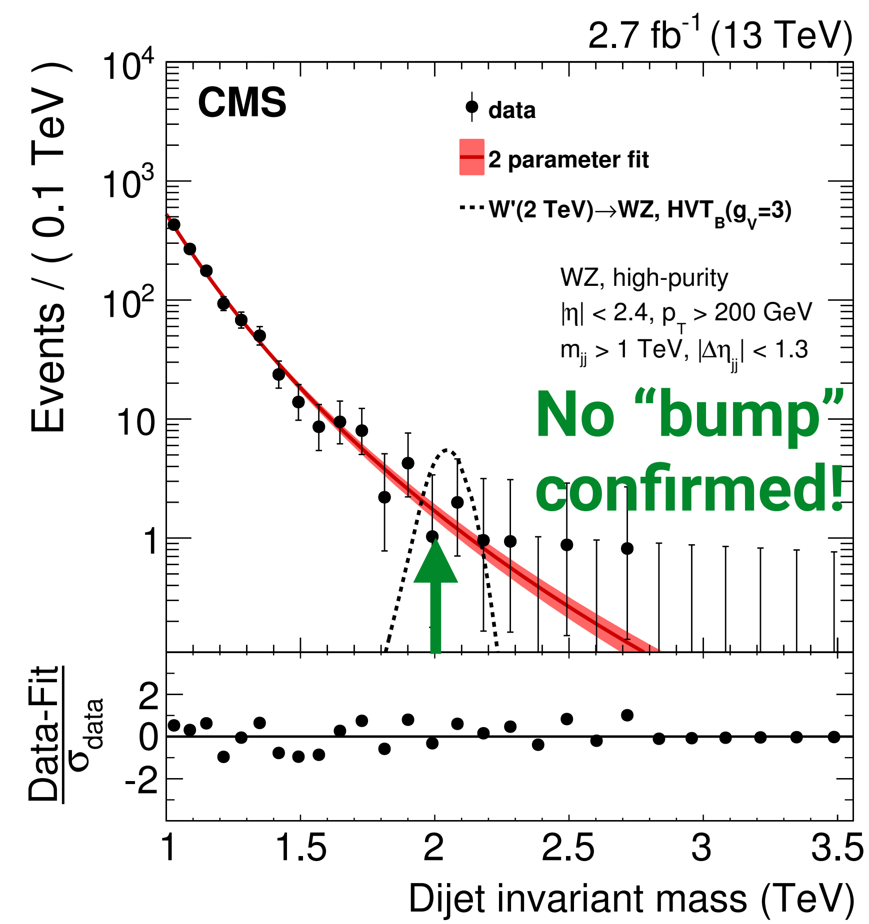
Limits on signal strength σ by comparing fits of observed data with "background-only" and "background + signal" function.

$$L = \prod_i \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}$$

Results

And?

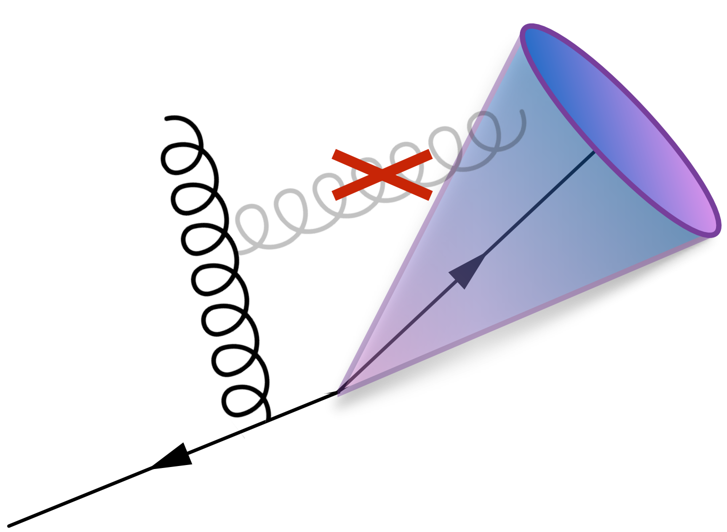




Just exclude 2 TeV excess for W'→WZ! However, other signals far from excluded

Thesis work:
Diboson resonance searches at 13 TeV with CMS

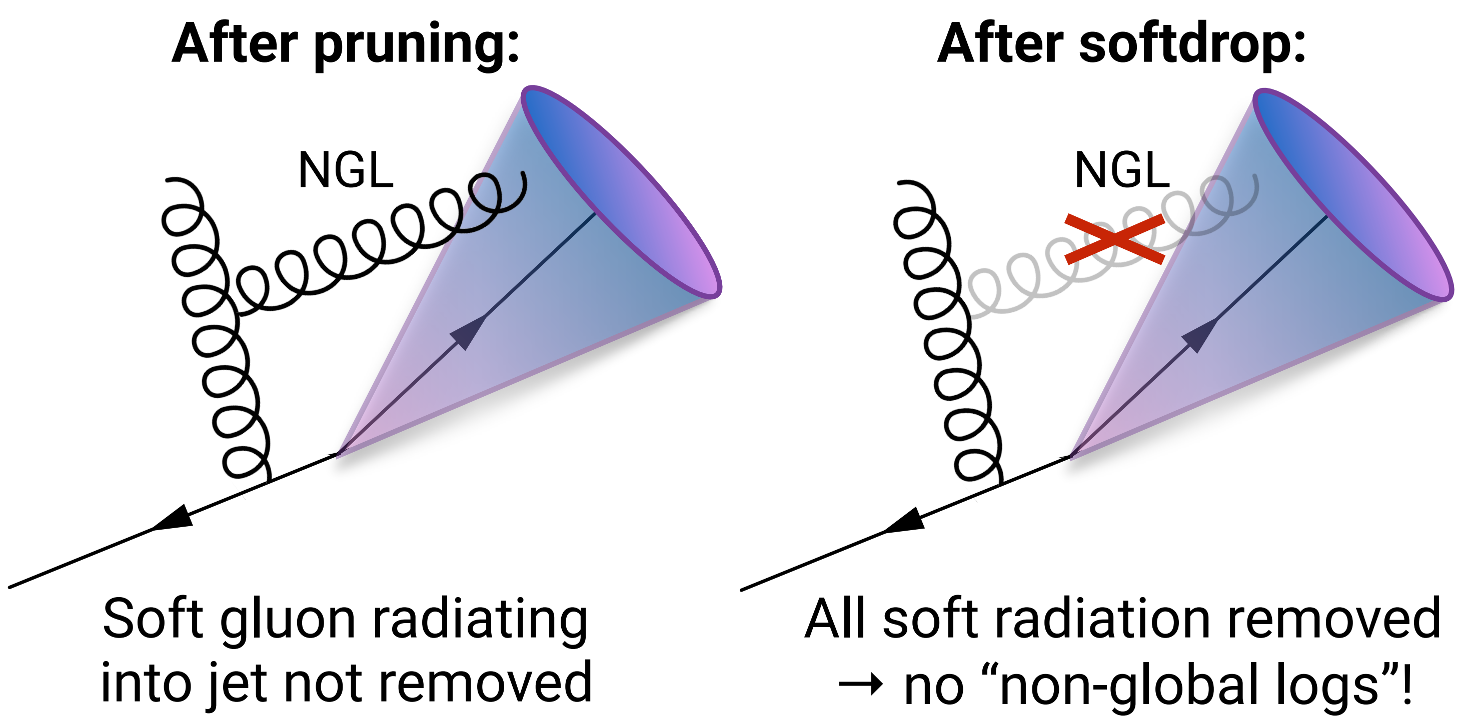
- I : First search for diboson resonances at 13 TeV
- **II : A new pileup-resistant, perturbative robust tagger**
- III: A novel framework for multi-dimensional searches
- IV: Encoding jet substructure with a deep neural network



New jet mass algorithm: Softdrop

Softdrop (β=0)

arXiv:1307.0007 arxiv:1402.2657

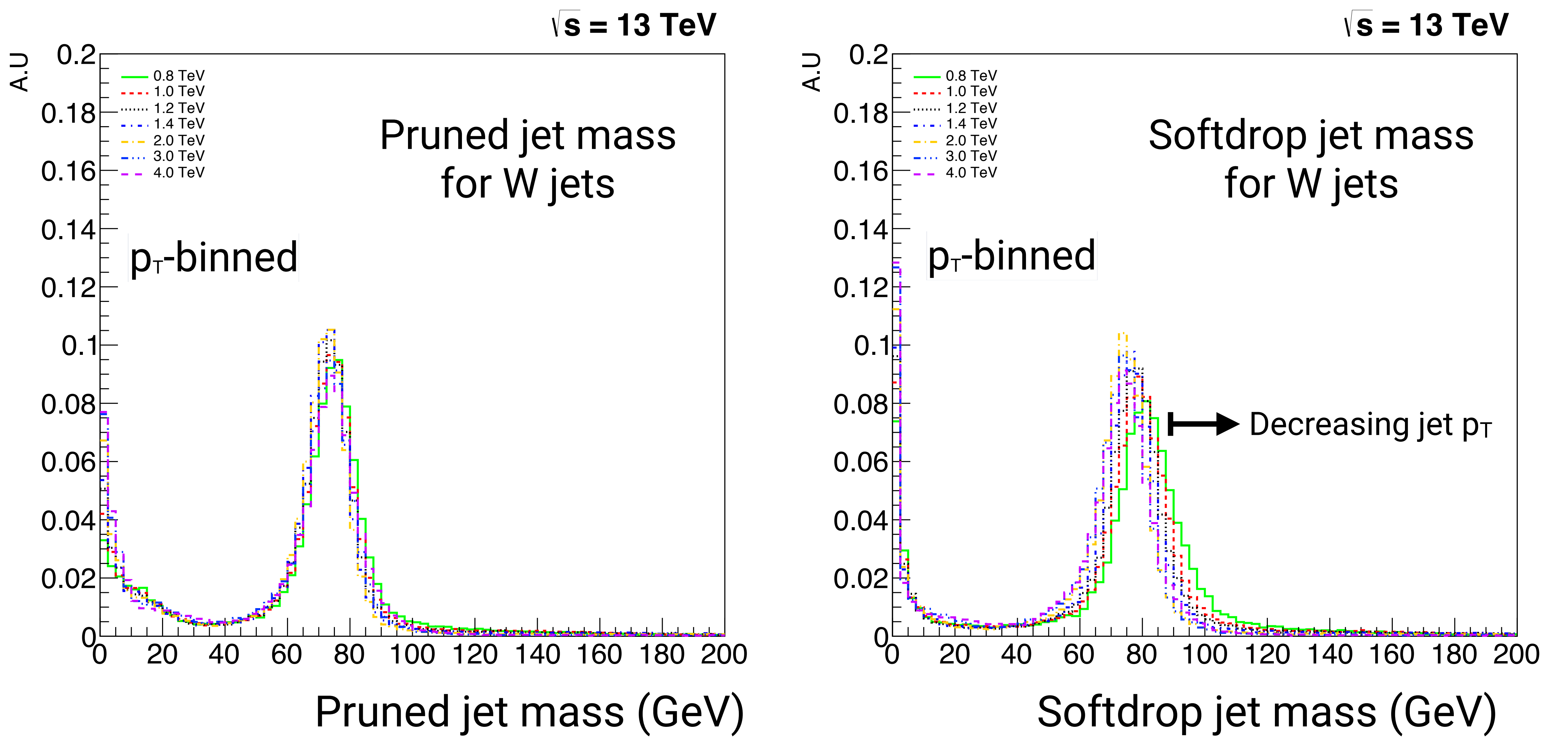


Pruning has "non-global logarithmic terms" in mass → not “perturbatively robust”. Softdrop has no such terms.

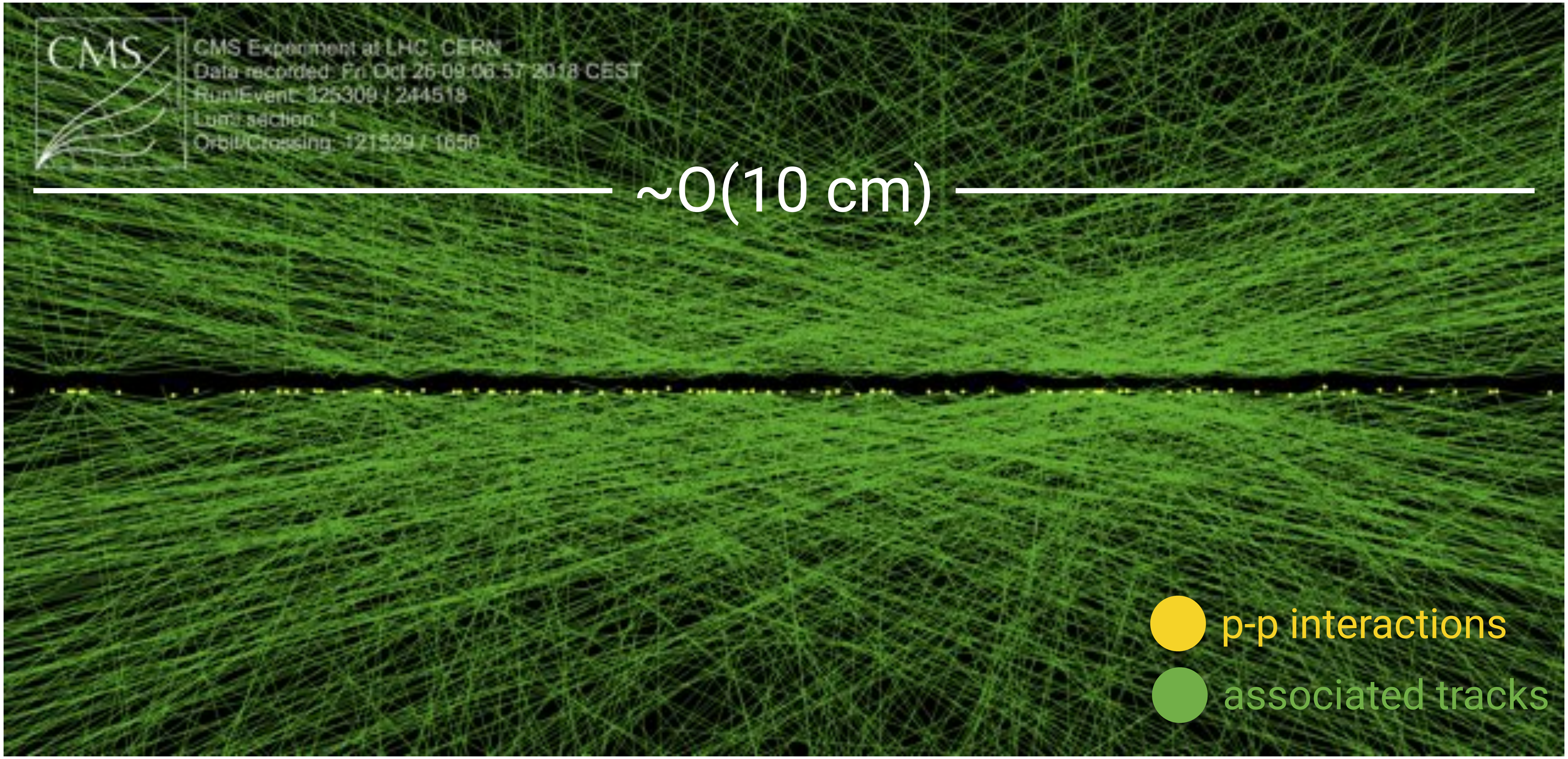
- can be calculated to higher precision than what is possible for other groomed or plain jet mass variables

Problems with softdrop

However, softdrop mass for W jets highly p_T dependent:



Pileup



Never only ~1 pp collision per event, but several!

Pileup

Pileup in 2016 double that of 2015!

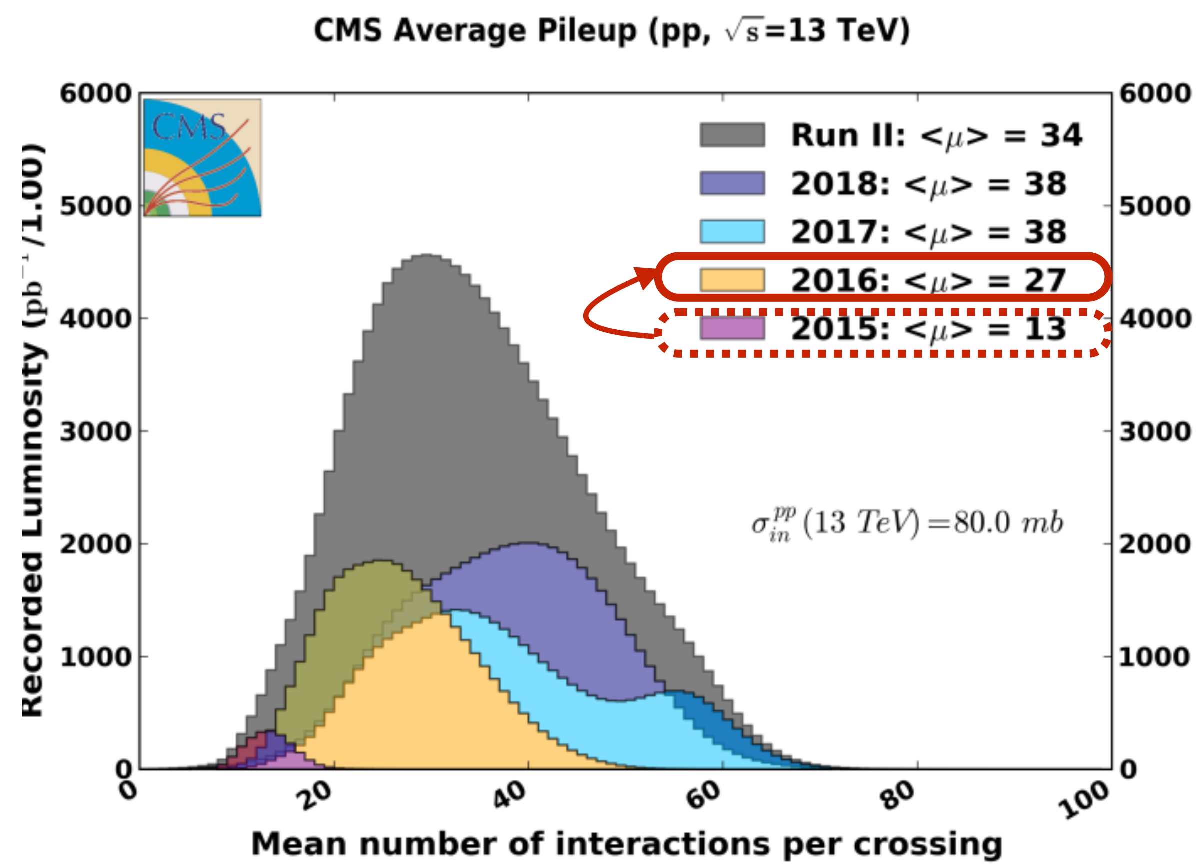
Fortunately,
PileUp Per Particle Identification (PUPPI)

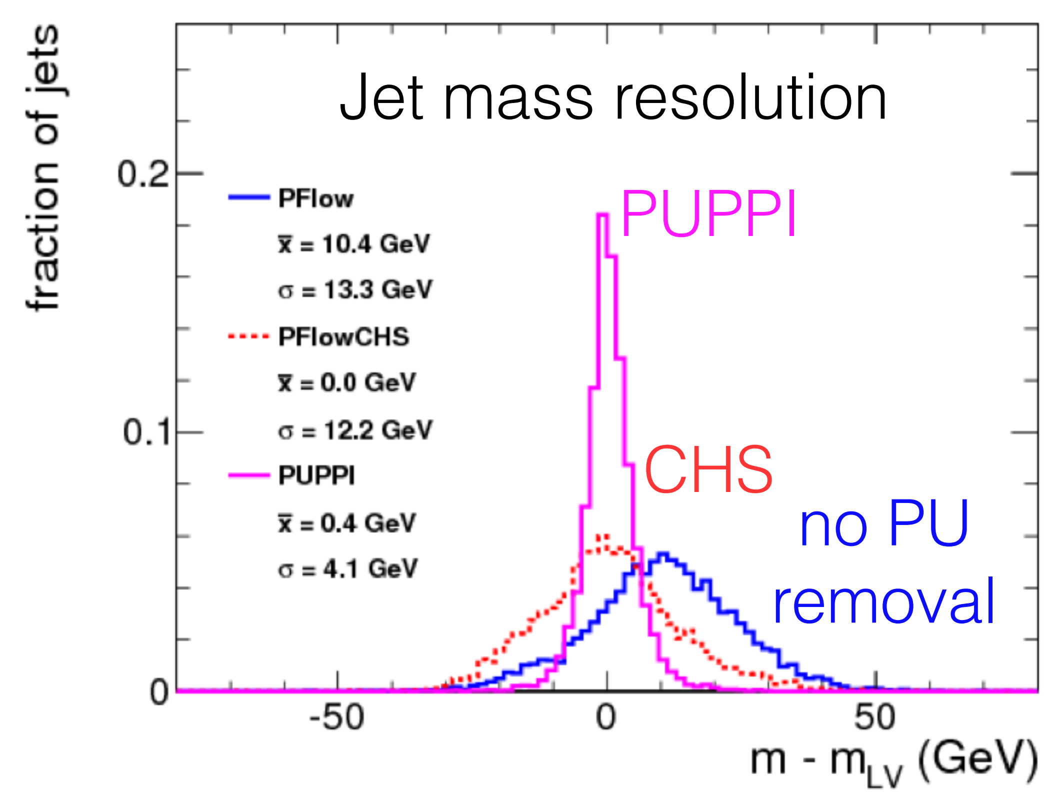
- CHS (old): remove charged particles not associated with primary vertex

- PUPPI (new): probability for ANY particle (neutral+charged) to be from pileup, reweights each accordingly

Huge resolution improvement for jet observables in large-cone jets

Tagger based on PUPPI and softdrop!





Developing a new V-tagger: Softdrop jet mass corrections

How does PUPPI softdrop look??

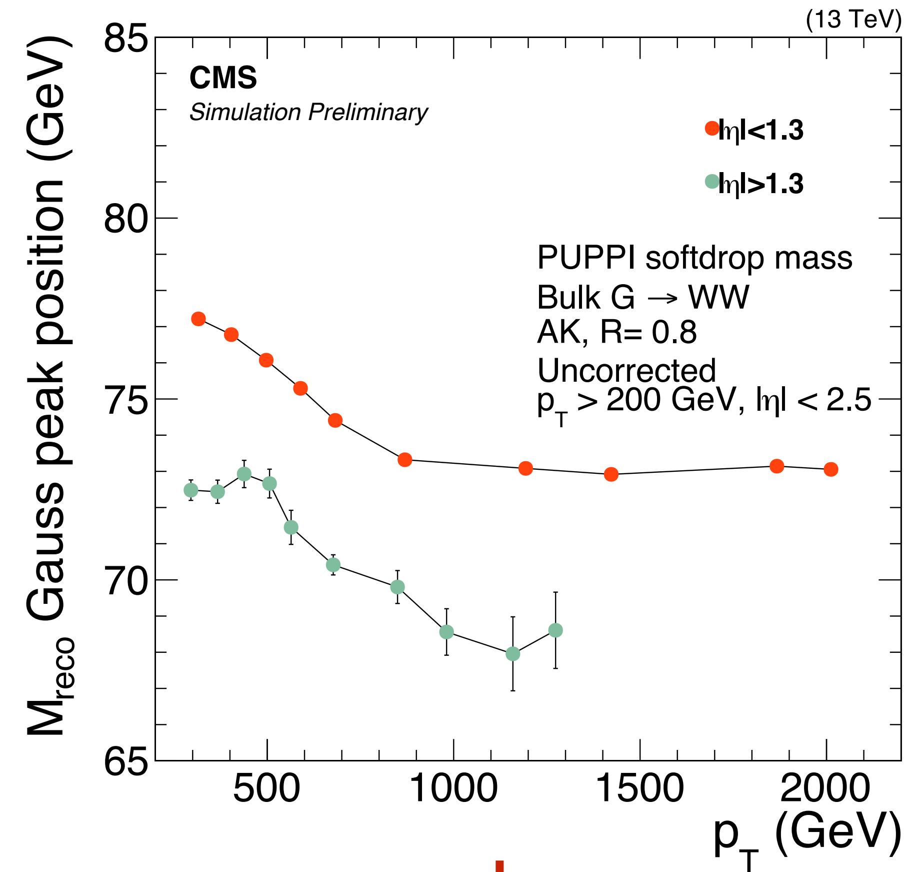
- p_T-dependence still present!

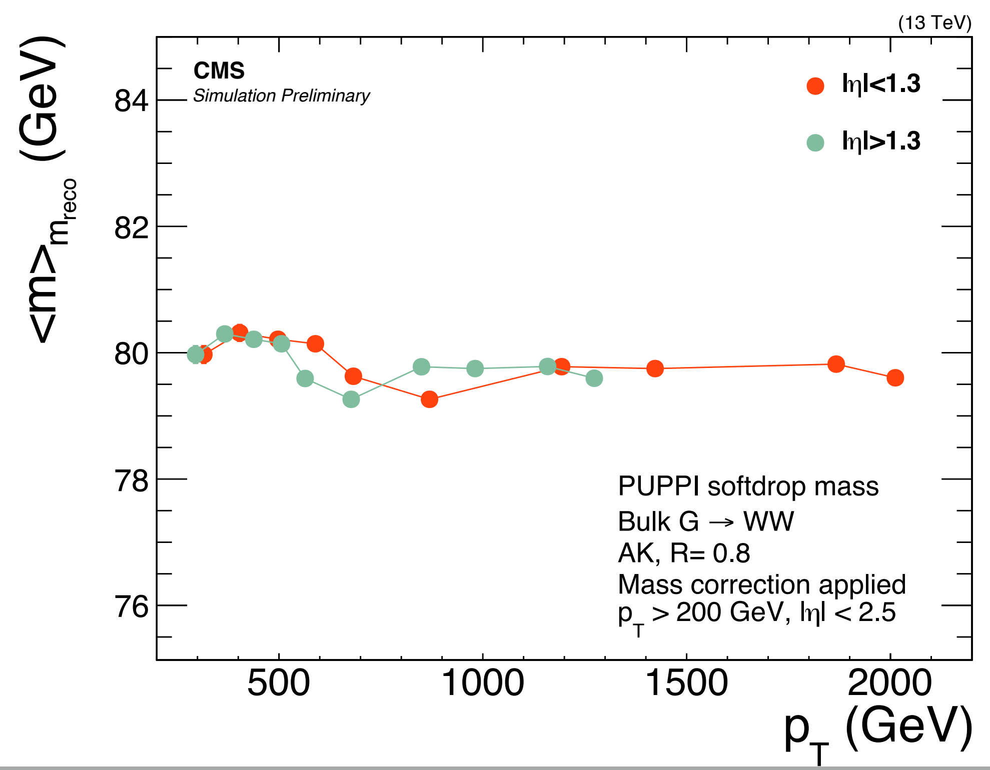
Solution: Compute dedicated PUPPI softdrop jet mass corrections!

- remove p_T/η-dependence, shift mass to 80 GeV

Finally stable softdrop mass peak with p_T

(Aside: not a problem with softdrop algorithm, must develop dedicated softdrop jet corrections!)



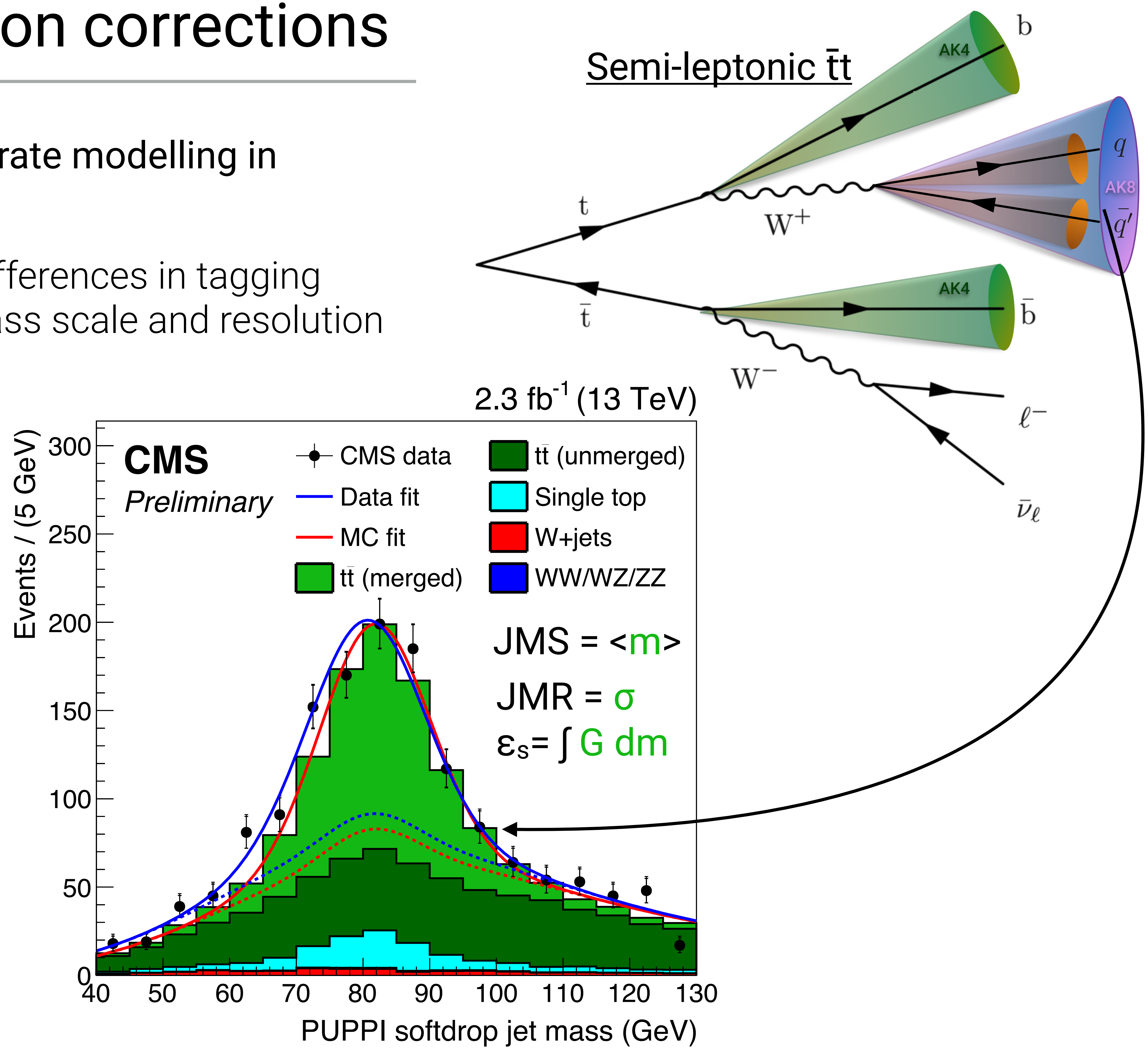


Developing a new V-tagger: Data/simulation corrections

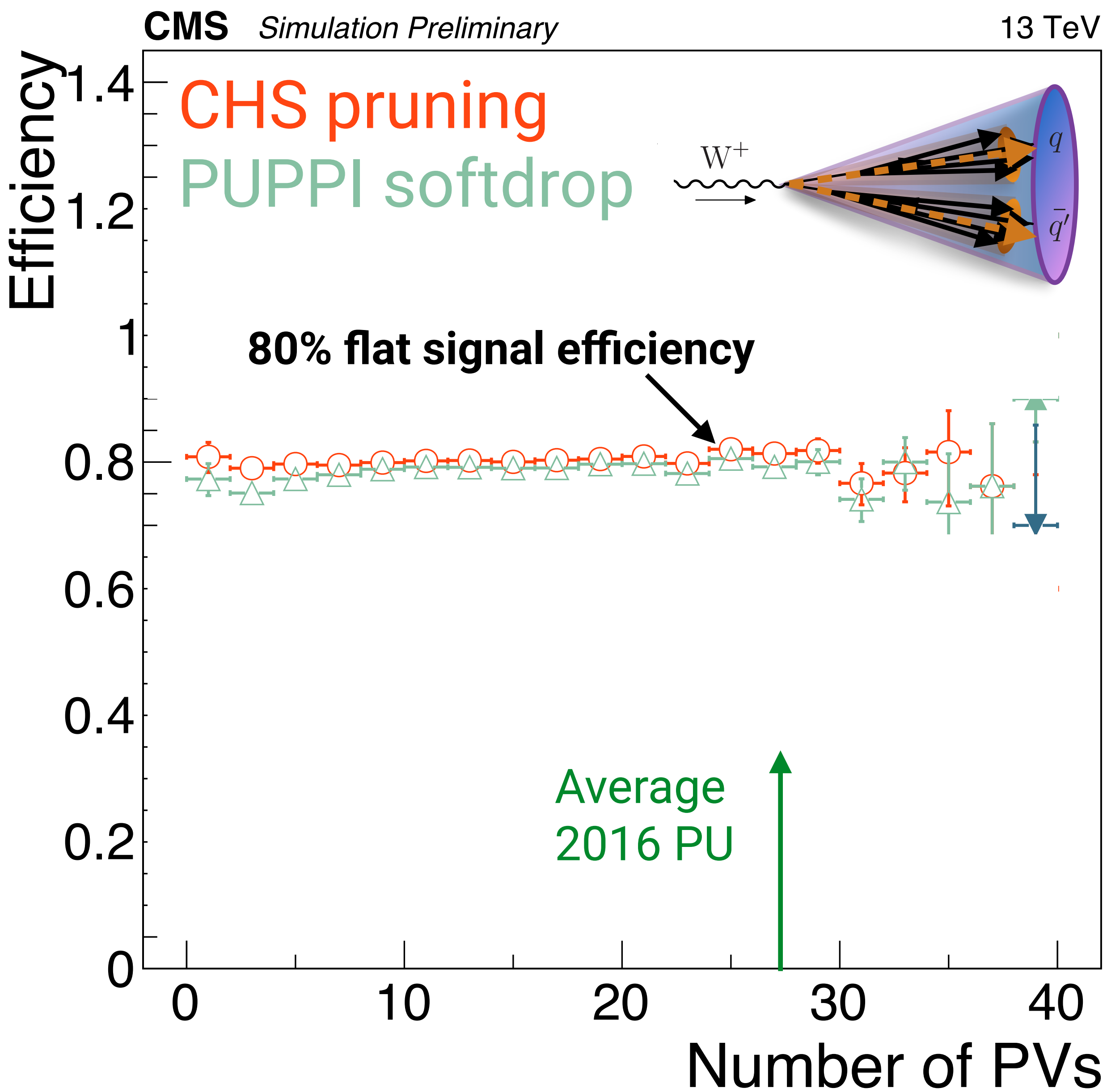
To account for inaccurate modelling in simulation

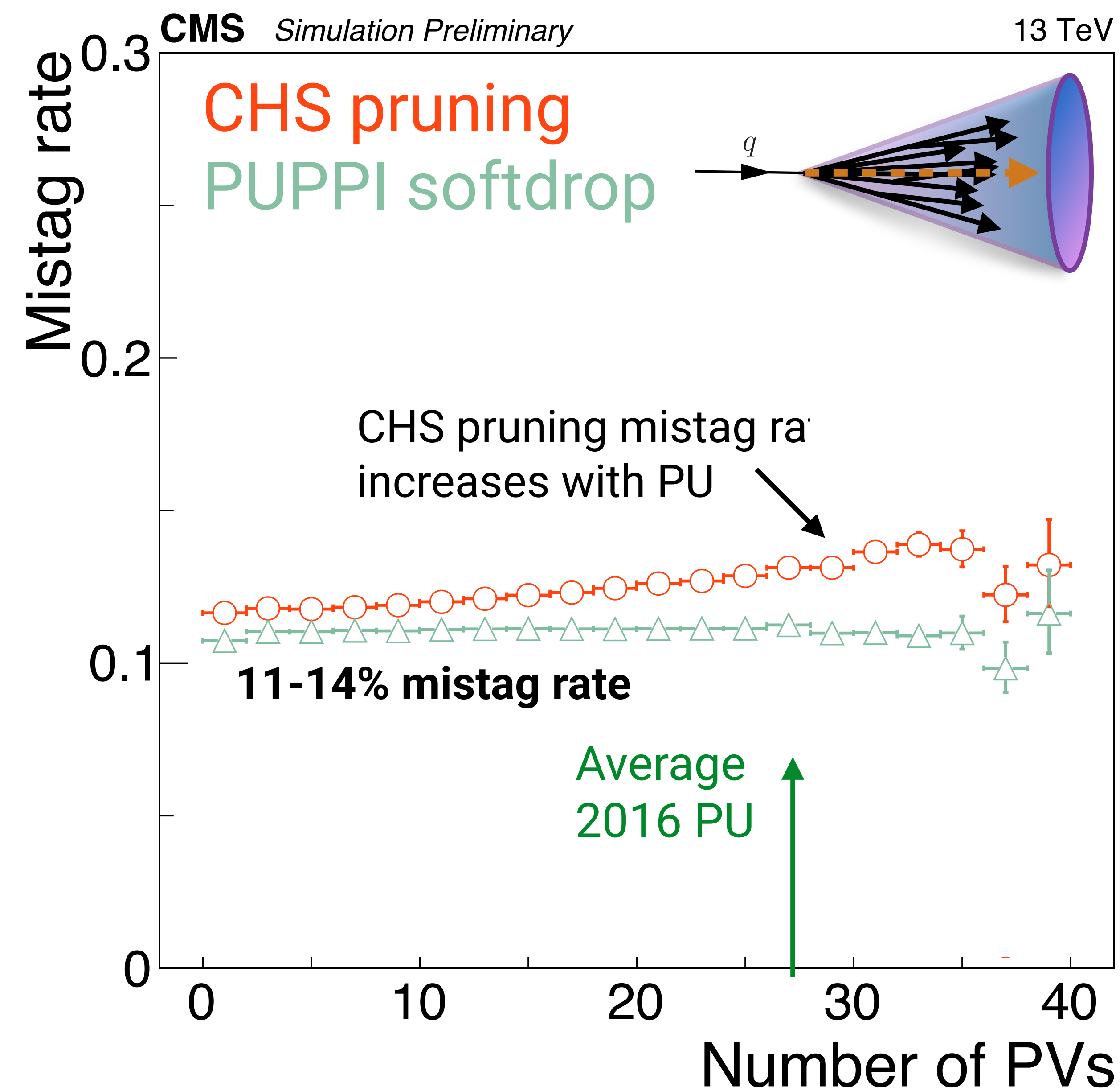
evaluate data/MC differences in tagging efficiency (ε_S), jet mass scale and resolution

Important!
Affects the estimated signal yield in all analyses using W-tagging!



Developing a new V-tagger: Performance

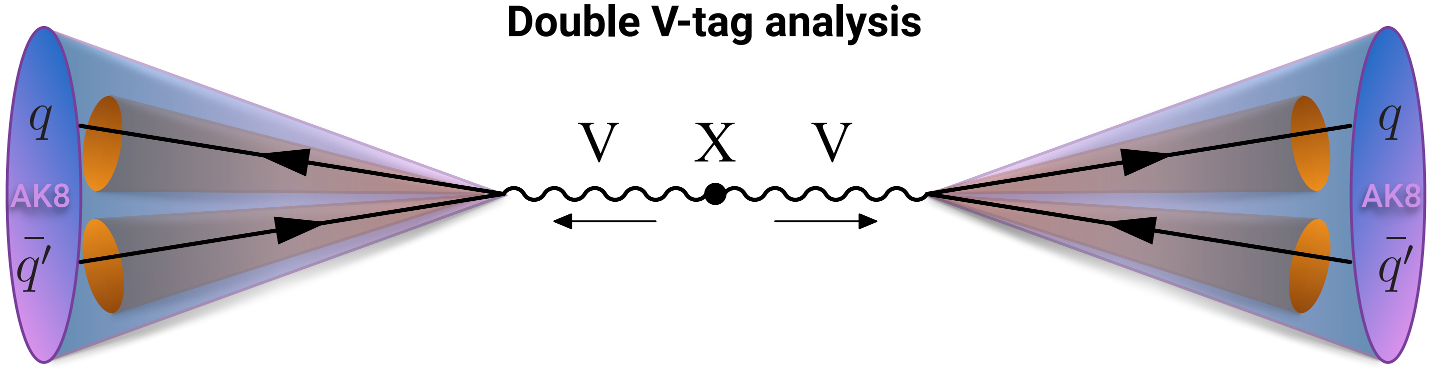


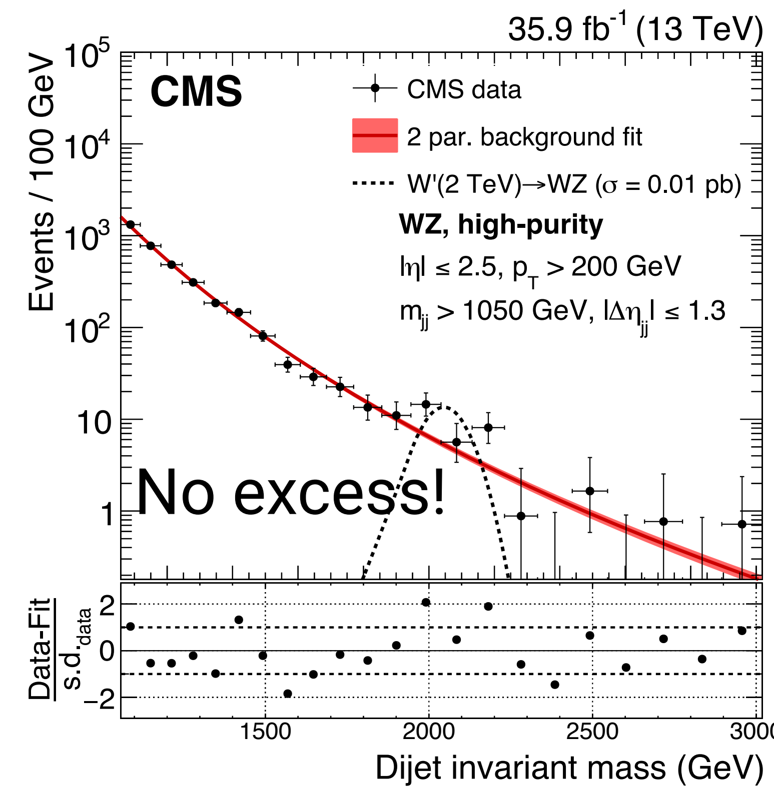


PUPPI softdrop also better performance than CHS pruning for the expected pileup in 2016.
CHS pruning 15% higher mistag rate than PUPPI softdrop!

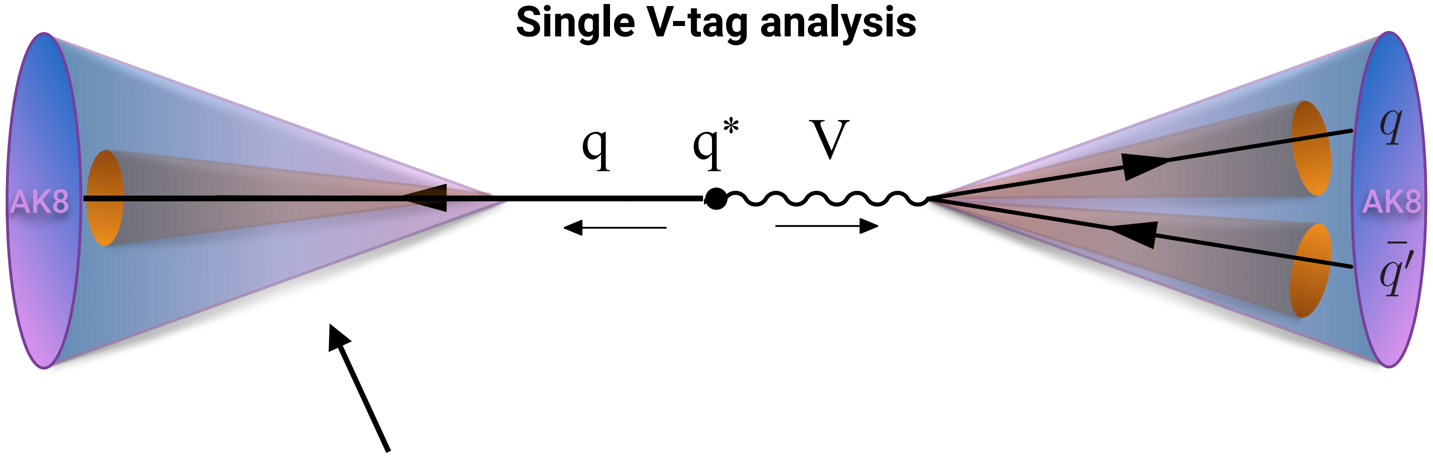
Results

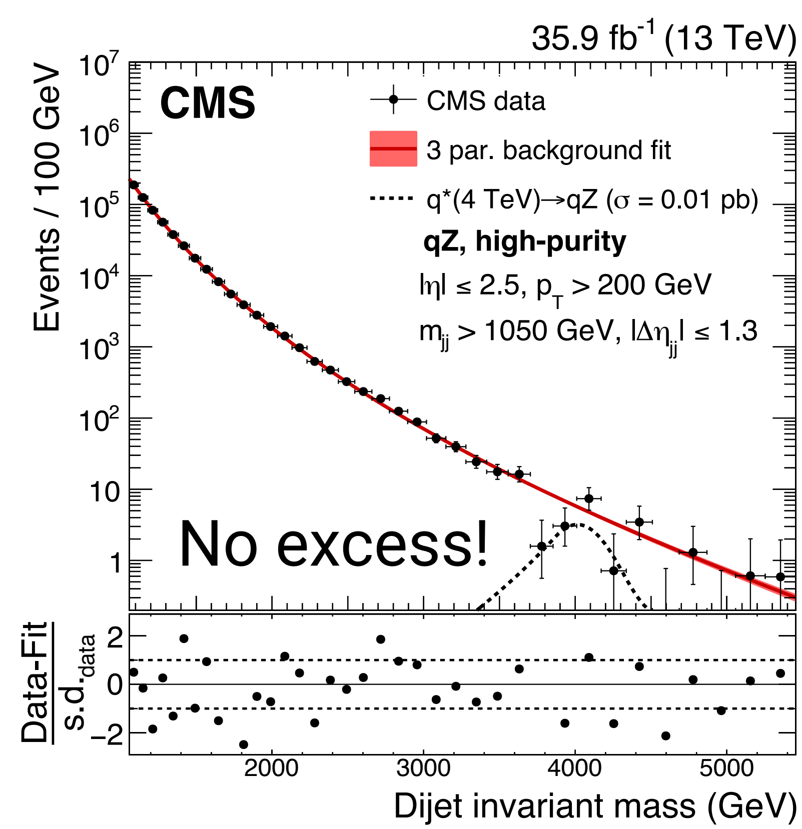
Double V-tag analysis





Single V-tag analysis



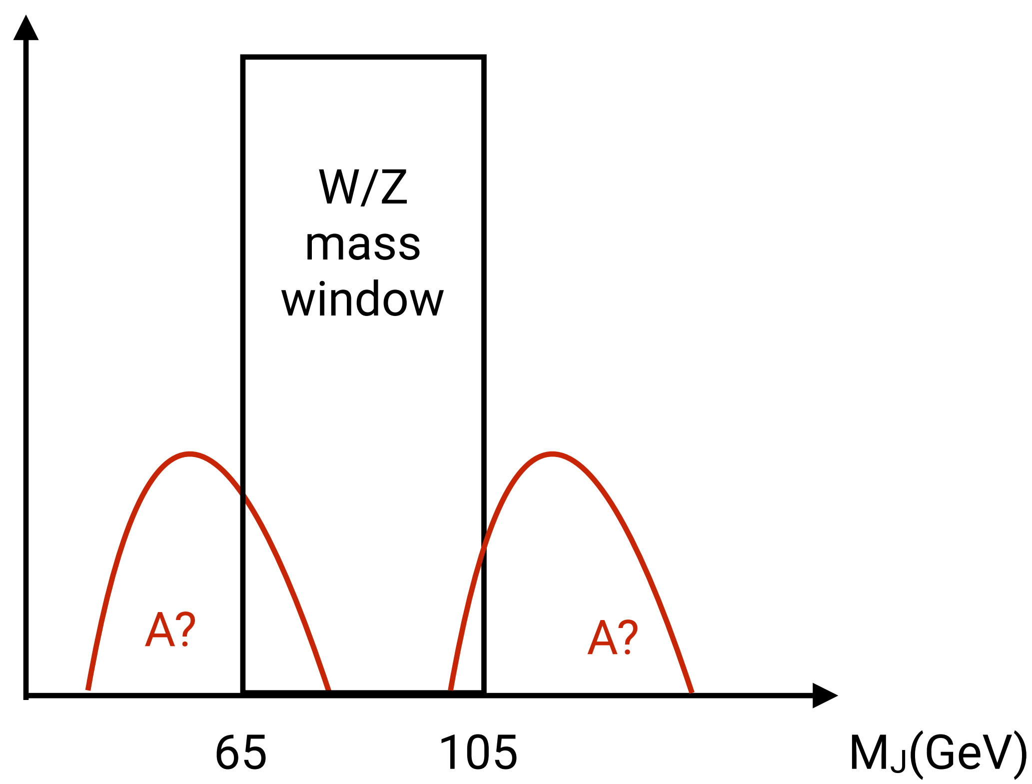


Adding search for excited quarks decaying to qV by removing W-tag. Never before analysed channel at 13 TeV!

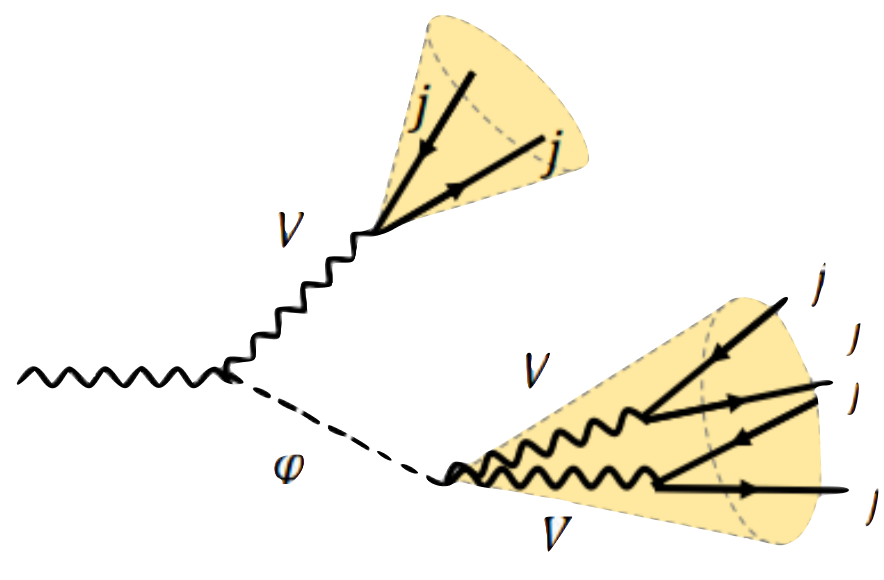
What could we be missing?

Signals could still be present in our data, but may look different!

- catching tail of other non-SM boson?



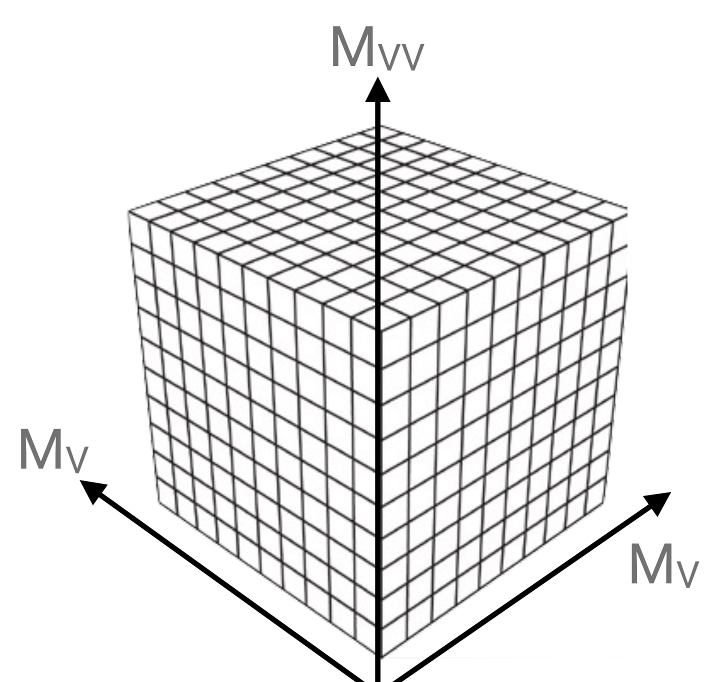
- not necessarily 2-, but N-pronged?



→ Idea: Lets make a **generic framework** allowing us to easily scan full jet mass and dijet invariant mass spectrum!

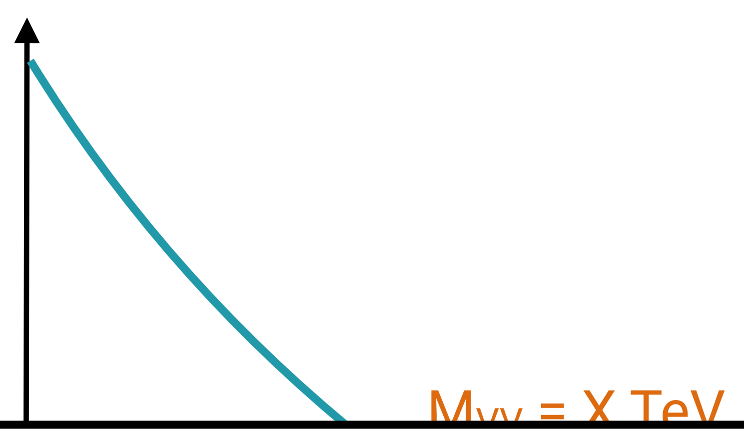
Thesis work:
Diboson resonance searches at 13 TeV with CMS

- I : First search for diboson resonances at 13 TeV
- II : A new pileup-resistant, perturbative robust tagger
- III: A novel framework for multi-dimensional searches
- IV: Encoding jet substructure with a deep neural network

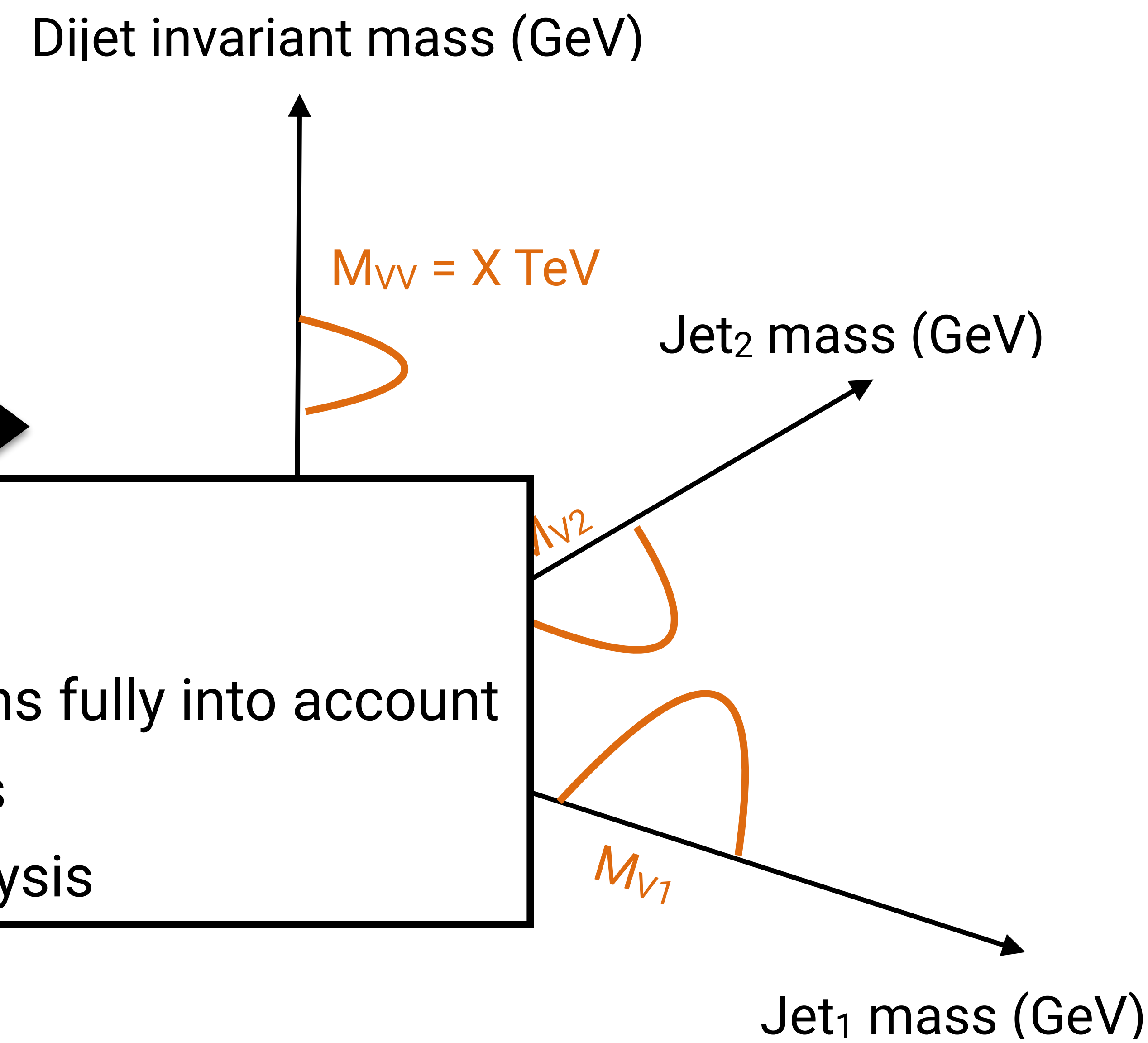


3D fit strategy

1D



3D



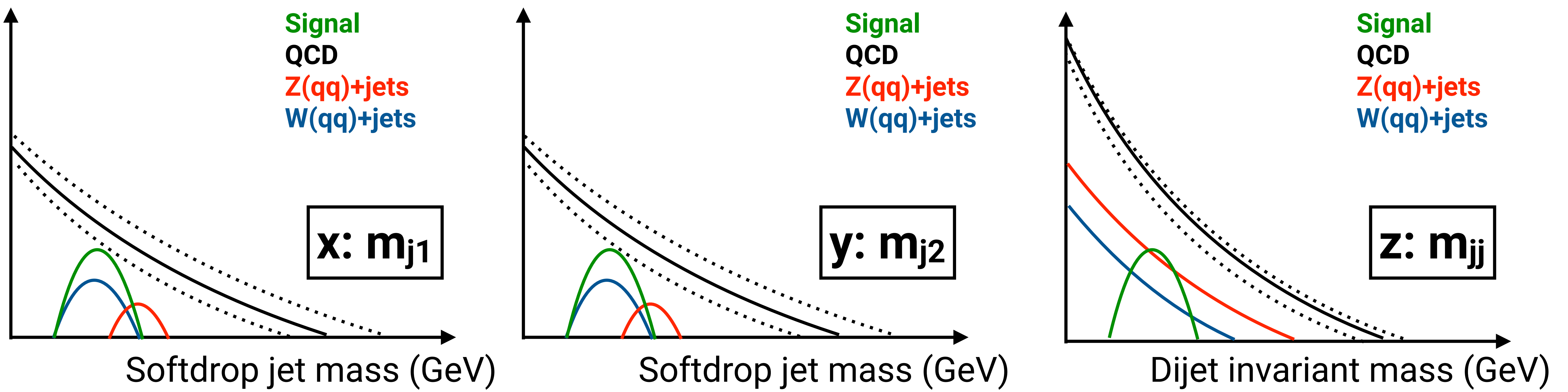
The benefits of this:

1. Can take jet mass and dijet mass correlations fully into account
2. Larger signal acceptance without mass cuts
3. Can scan full jet mass spectrum in one analysis

Take advantage of the fact that signal is resonant in 3D: M_V, M_V and M_{VV}

- scan M_{V1}-M_{V2}-M_{VV} hyperplane!

Building PDFs



4 ingredients to full 3D model, derived from MC

1. Signal 3D PDF
   - Resonant in x, y and z
2. Non-resonant background
   - QCD, main background
   - Non-resonant in x, y and z
3. Resonant background
   - W/Z+jets, resonant in x+y
4. Alternate PDFs
   - 5 additional shape uncertainties

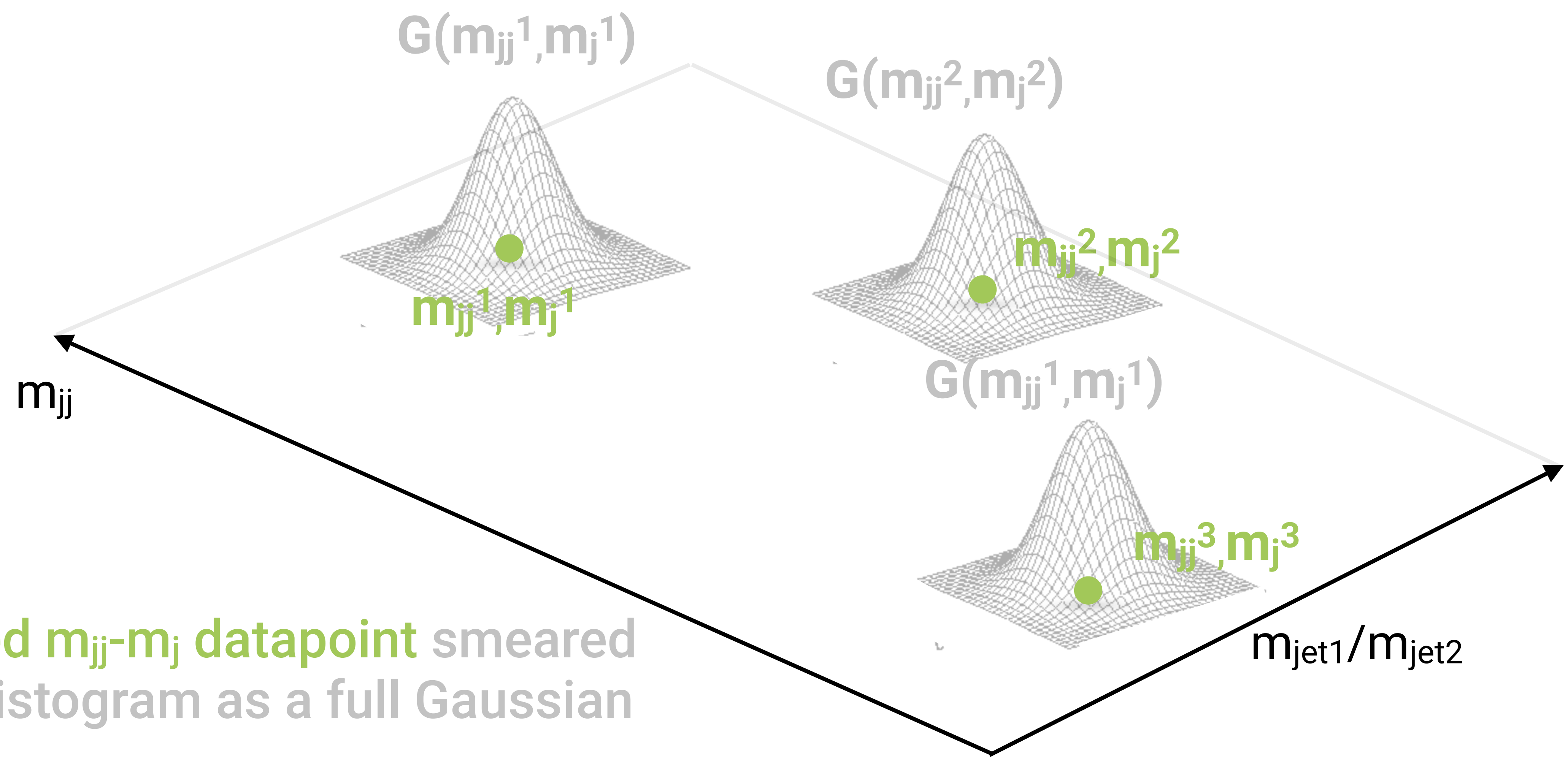
Non-resonant background

Conditional PDFs to account for m_j-m_{jj} correlations!

$$\mathbf{P_{NR}(m_{jj}, m_{jet1}, m_{jet2})} = \underbrace{P_{jj}(m_{jj} \mid \theta_1)}_{\text{1D histogram}} \times \boxed{\underbrace{P_j(m_{jet1} \mid m_{jj}, \theta_2)}_{\text{2D histogram}}} \times \underbrace{P_j(m_{jet2} \mid m_{jj}, \theta_3)}_{\text{2D histogram}}$$

Final 3D histogram has 250k bins, how to get smooth shape from QCD MC?

→ Start from generated jet mass/m_{jj} and use "forward smearing kernel" approach

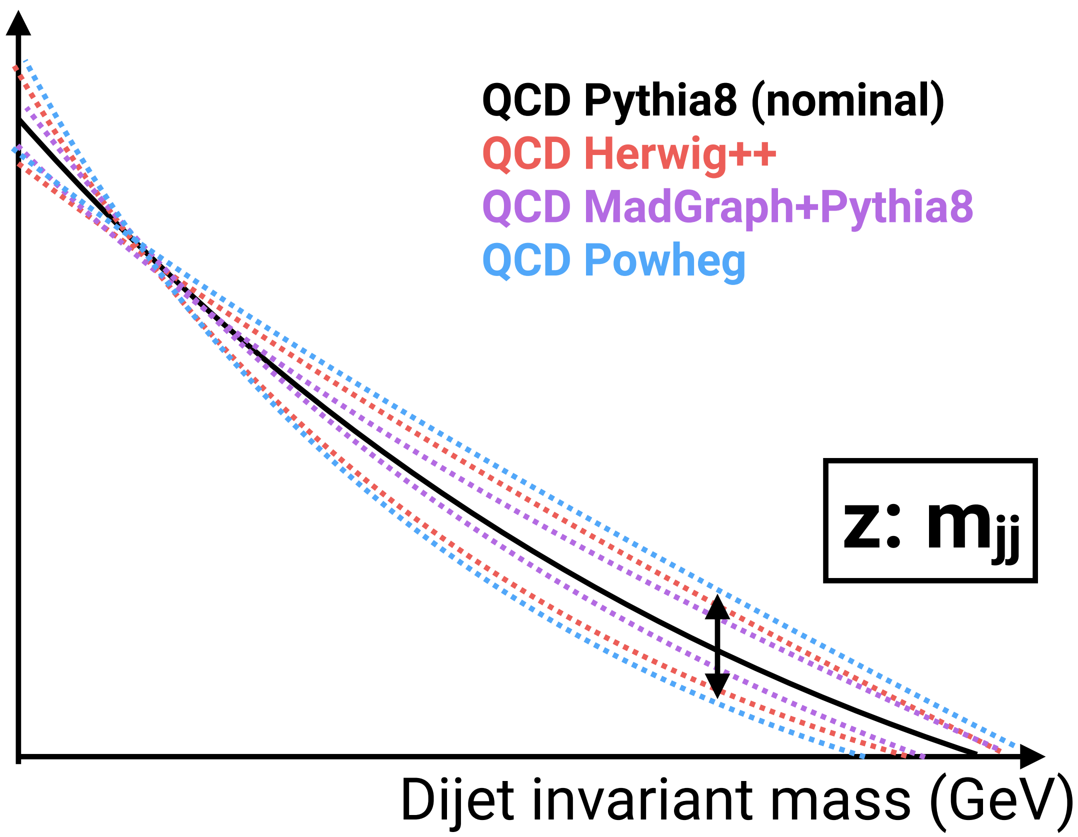


Generated m_{jj}-m_j datapoint smeared and fill histogram as a full Gaussian

Alternative shapes

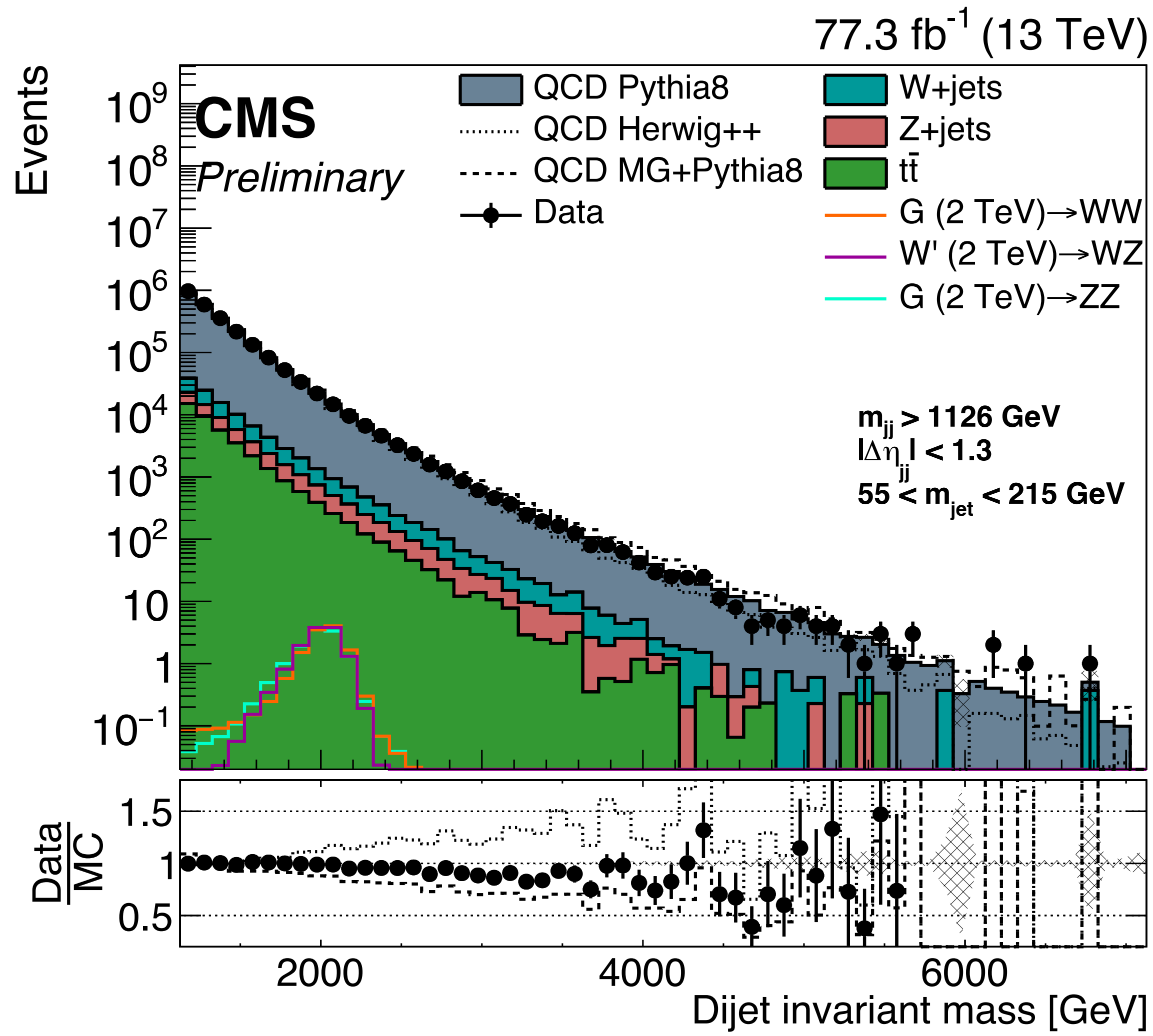
Is Nature Herwig++, MadGraph or Pythia?
LO(Pythia) or NLO (Powheg)?

- predictions disagree, let's allow it to be all!



Add alternate shapes based on different MC

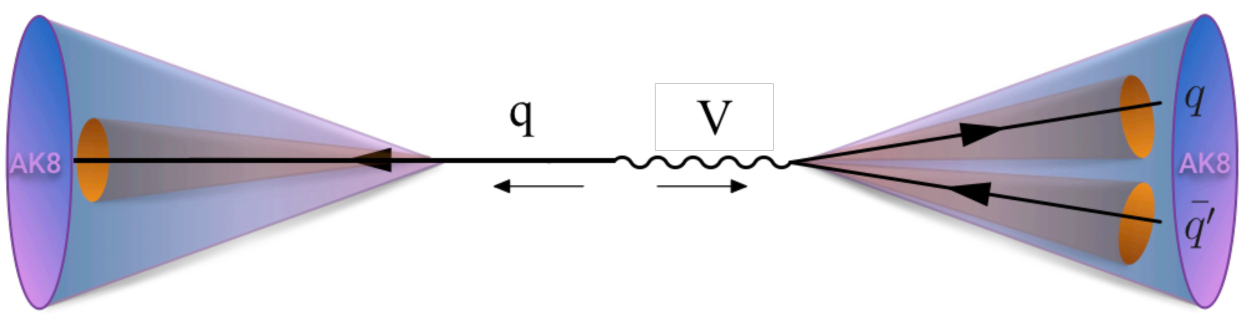
- large pre-fit uncertainties, fit can adjust to match data



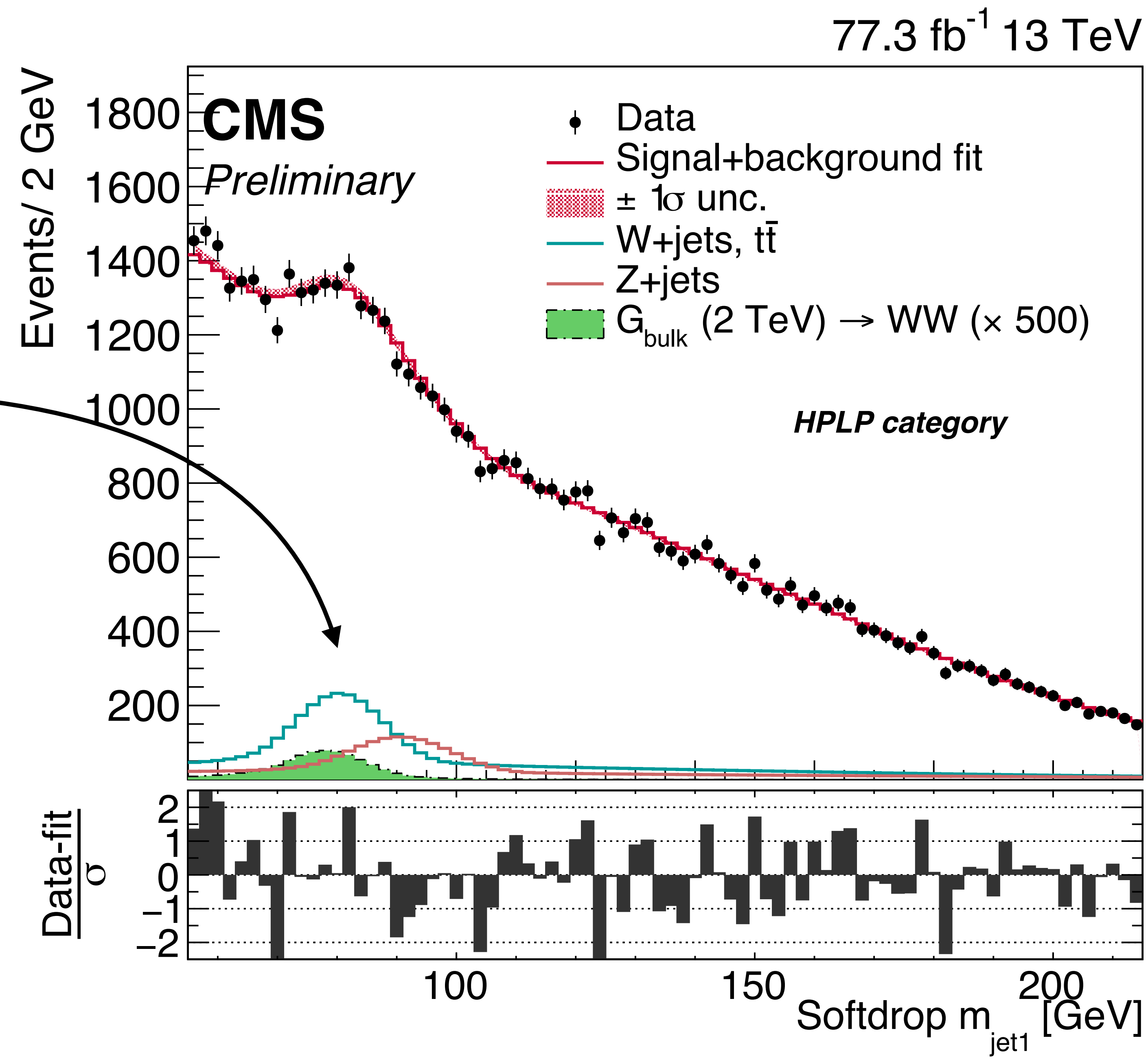
First results with 3D fit

Results first shown in public today!
(to be presented at Moriond next week)

First measurement of boosted W/Z(qq)+jets in diboson searches!

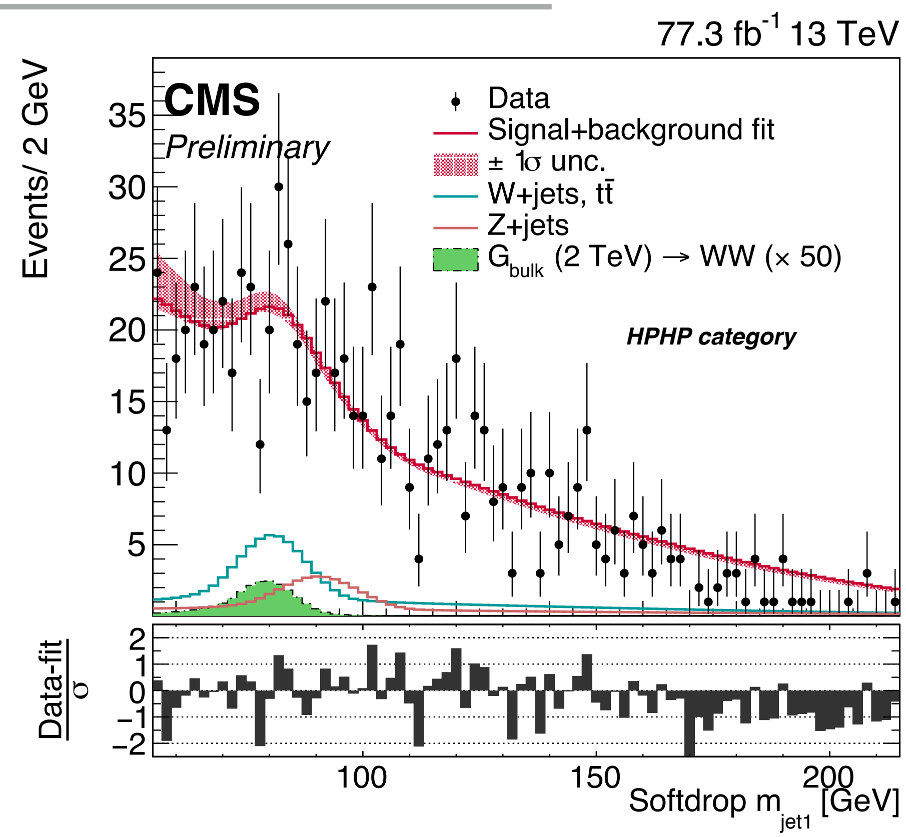


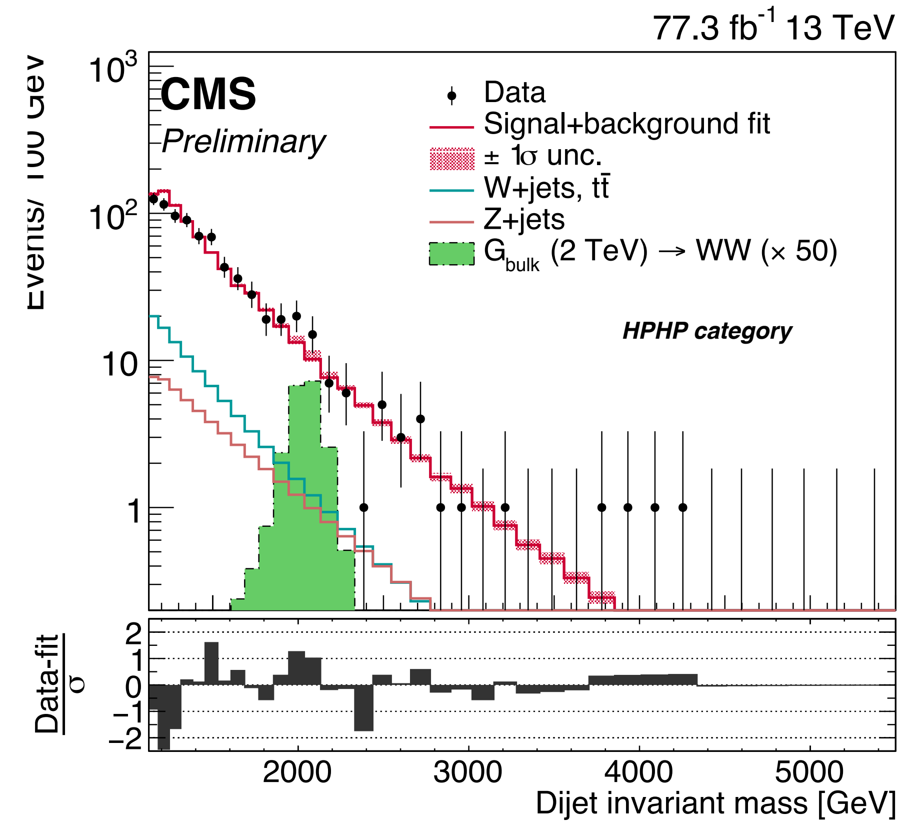
- improves sensitivity to m_X, constrain signal systematic uncertainties (mass scale, resolution, tag eff. etc.)



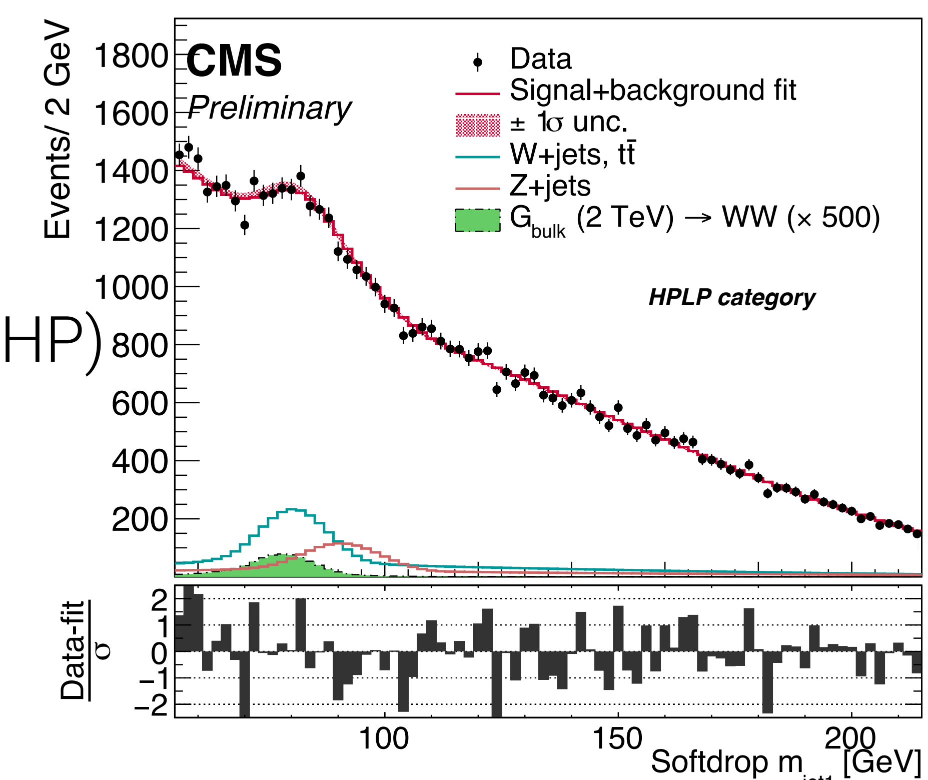
First results with 3D fit

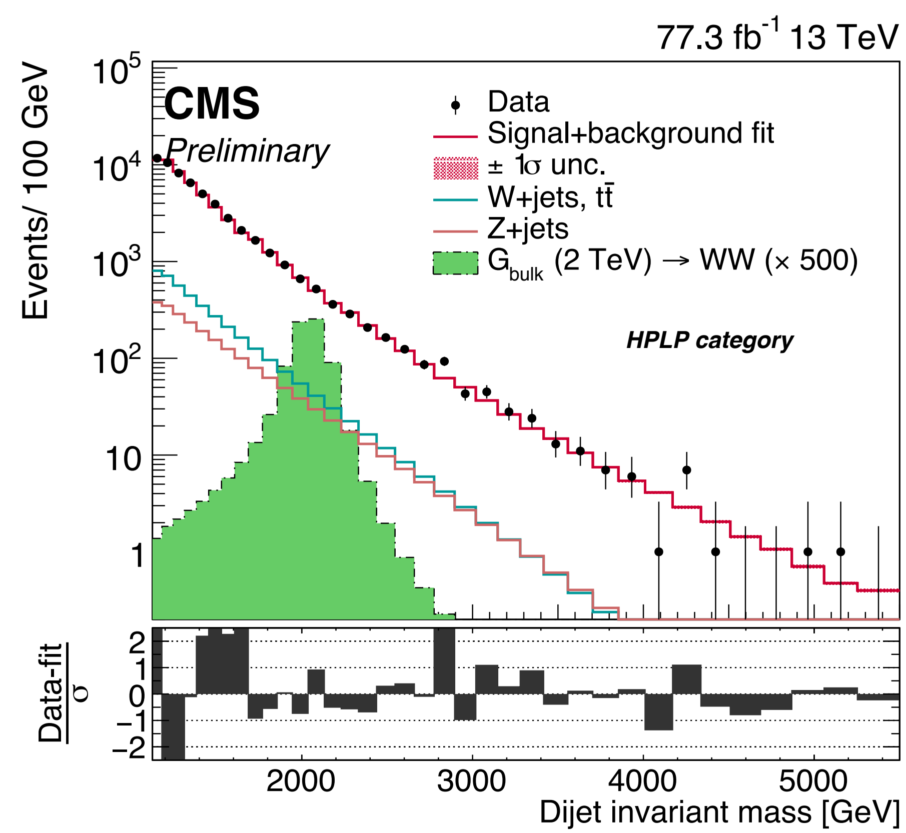
High-purity category
- best signal/background



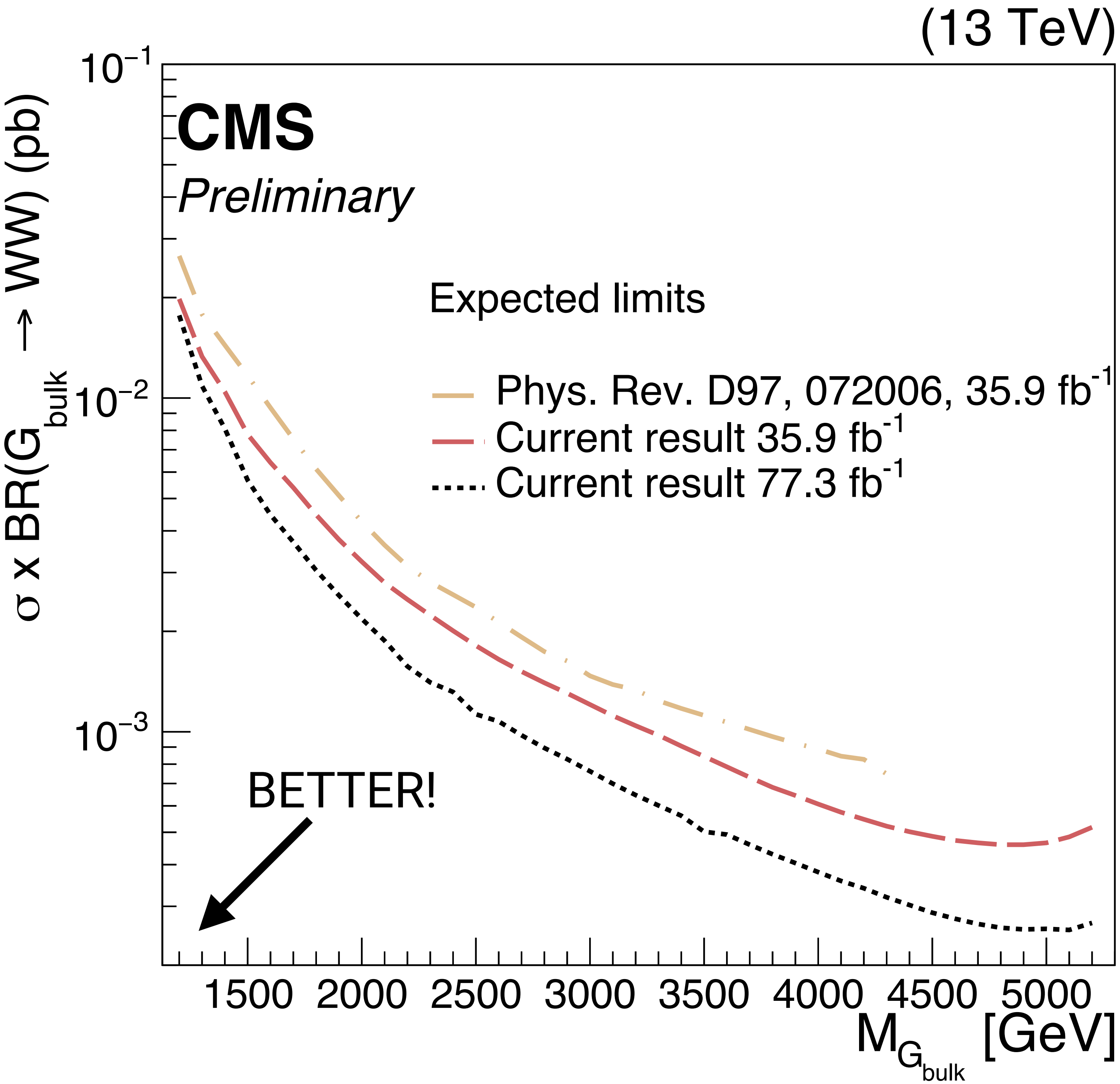


Low-purity category
- high statistics, constrains HP)



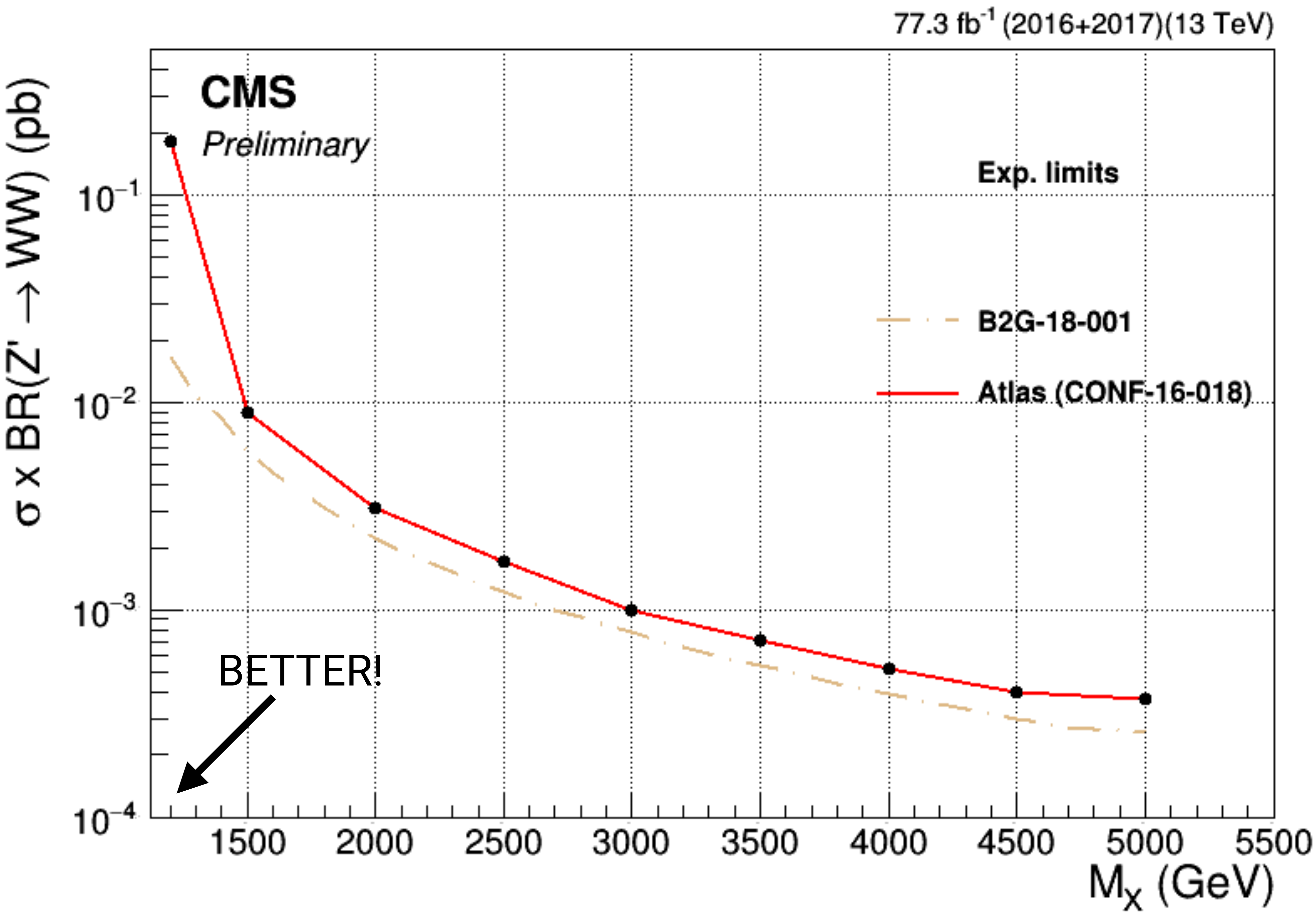


Comparison to old method



3D fit method yields 20-30% improvement with respect to 1D search!
Adding 2017 data yields ~40% performance improvement

Comparison to ATLAS

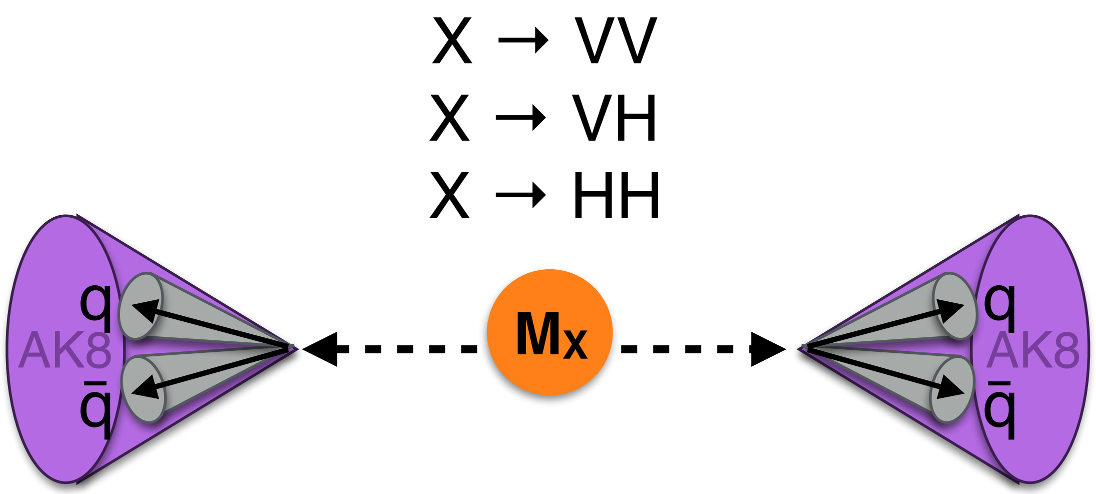


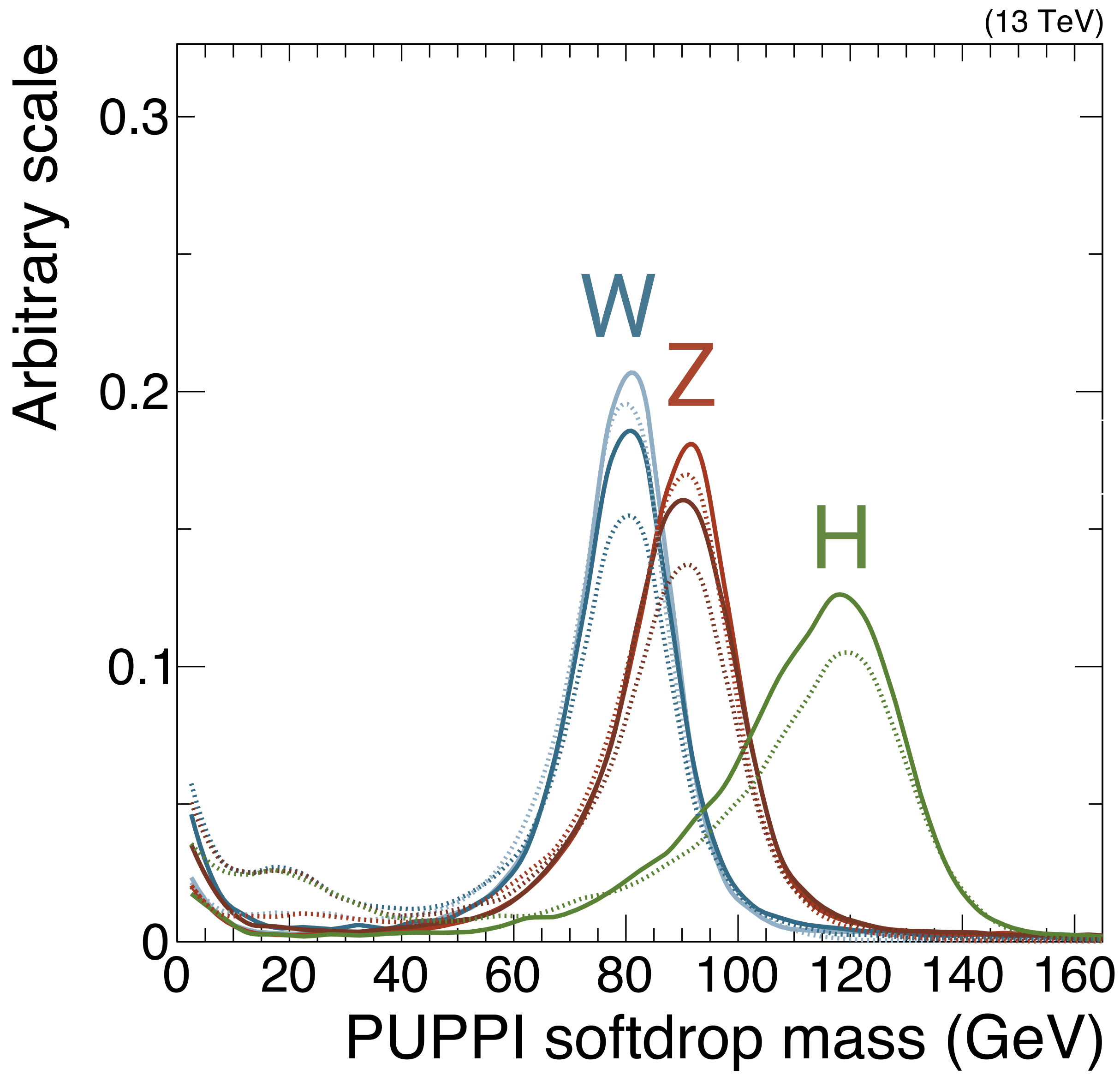
Up to 35% better than ATLAS equivalent search!

Next steps

For full 13 TeV dataset of 150 fb^{-1}:

VV, VH(bb) + HH in one single analysis!





Future plans

Going further, important to improve analysis sensitivity as no more C-O-M energy increase (after 14 TeV)

- need better taggers

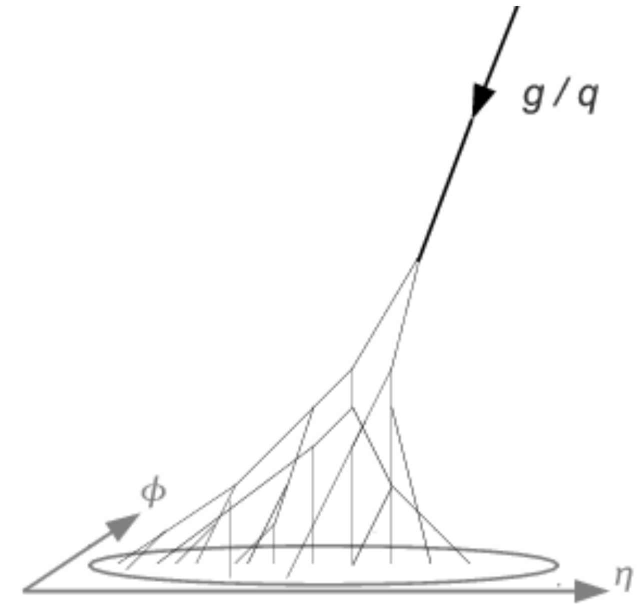
⟼ Deep Neural Networks!

So far, assume 2-prong signals and discriminate from background using τ_{21}

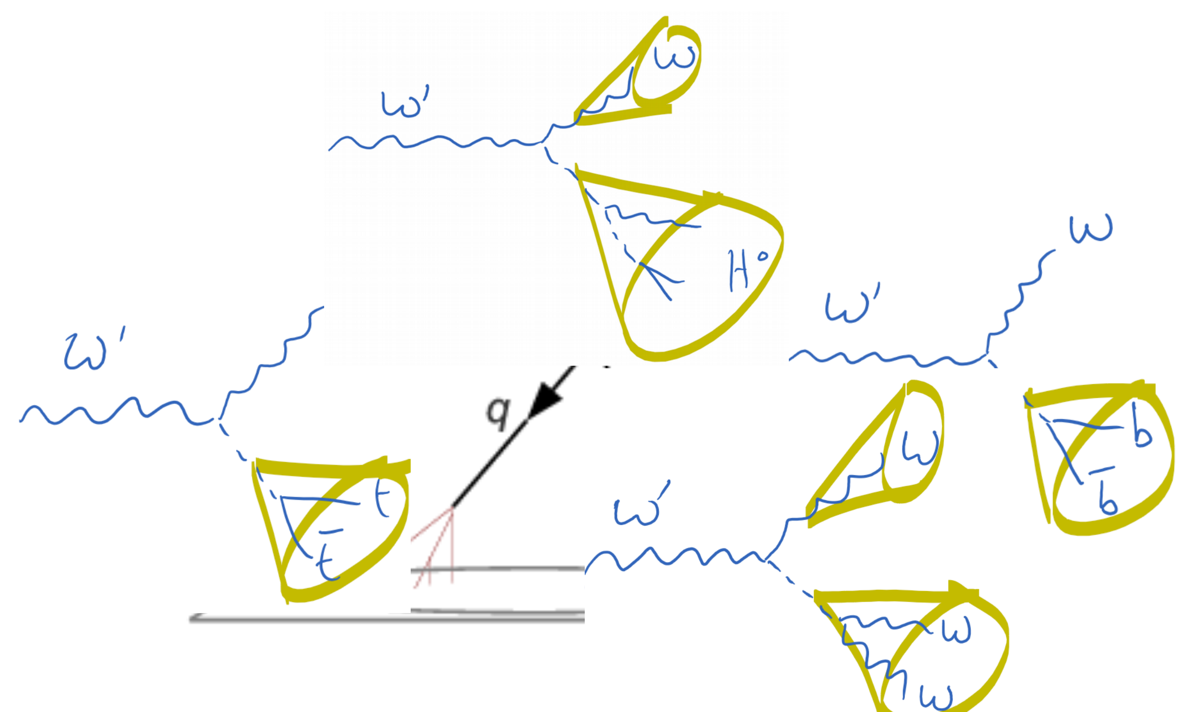
- how to stay model-independent to look for any signal?
- need tagger that "knows" what substructure looks like

⟼ Deep Neural Networks!

quark/gluon jet



vs.

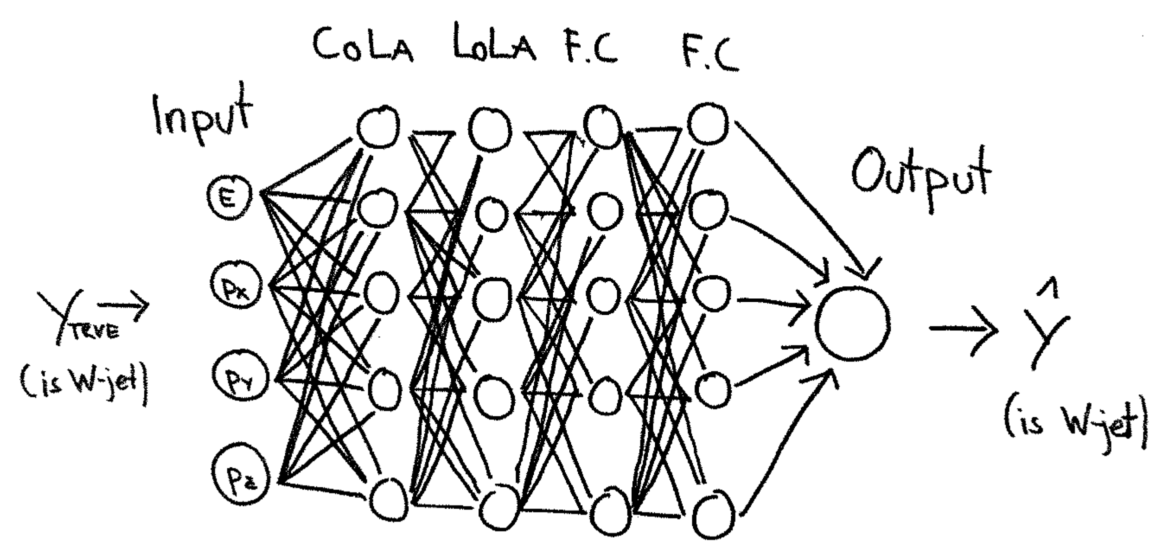


any W-jet

any of these?

Thesis work:
Diboson resonance searches at 13 TeV with CMS

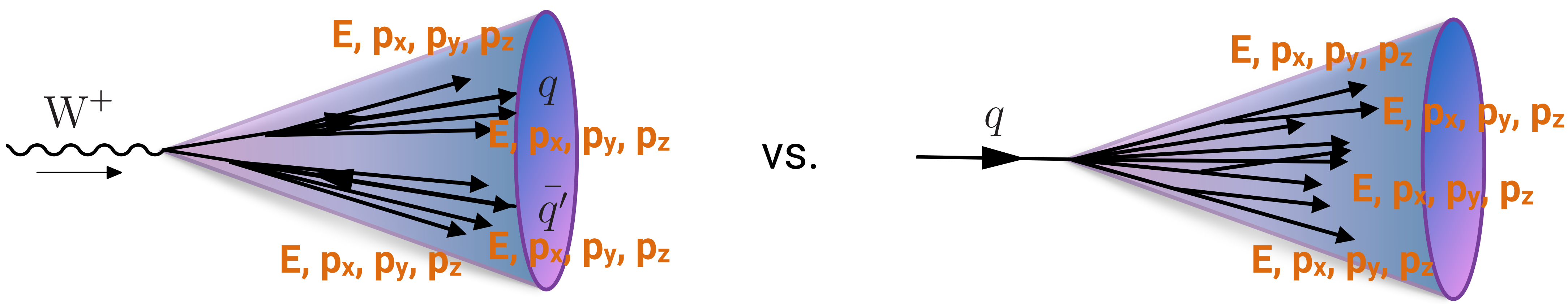
- **I** : First search for diboson resonances at 13 TeV
- **II** : A new pileup-resistant, perturbative robust tagger
- **III**: A novel framework for multi-dimensional searches
- IV: Encoding jet substructure with a deep neural network

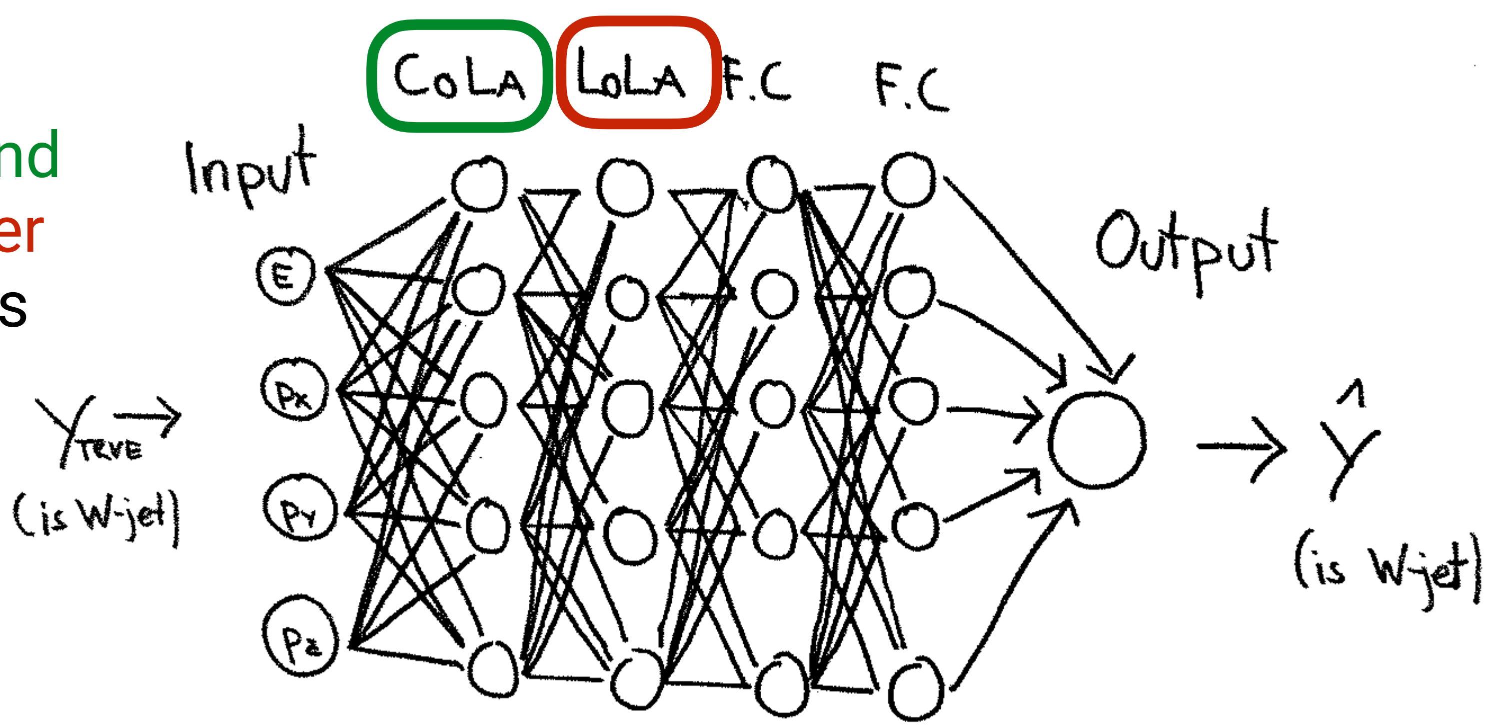


LoLa: DNN for W-tagging

Physics-based deep neural network to discriminate q/g from W jets
(introduced for top tagging by G. Kasieczka et. Al)

- only inputs are jet constituent 4-vectors!



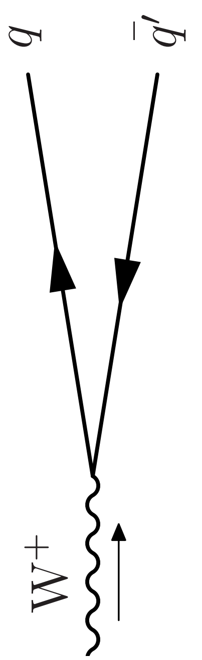


Jet clustering layer and Minkowski space layer ensure network learns substructure!

Custom layers

<u>Input:</u>

$$\begin{pmatrix} E^1 & E^2 \\ p_x^1 & p_x^2 \\ p_y^1 & p_y^1 \\ p_z^1 & p_z^2 \end{pmatrix}$$



E.g for 2 particles
(in reality 40)

Custom layers

5 (#CoLa) x 7 (#LoLa) matrix!

CoLa output:

$$\begin{pmatrix} E^1+E^2 & E^1 & E^2 & w_{1,4}E^1+w_{2,4}E^2 & w_{1,5}E^1+w_{2,5}E^2 \\ p_x^1+p_x^2 & p_x^1 & p_x^2 & w_{1,4}p_x^1+w_{2,4}p_x^2 & w_{1,5}p_x^1+w_{2,5}p_x^2 \\ p_y^1+p_y^1 & p_y^1 & p_y^1 & w_{1,4}p_y^1+w_{2,4}p_y^1 & w_{1,5}p_y^1+w_{2,5}p_y^1 \\ p_z^1+p_z^2 & p_z^1 & p_z^2 & w_{1,4}p_z^1+w_{2,4}p_z^2 & w_{1,5}p_z^1+w_{2,5}p_z^2 \end{pmatrix}$$

Lorentz Layer:

$$m^2(x^C_{\mu,i}) = g^{\mu\nu} x^C_{\mu,i} x^C_{\nu,i}$$

$$p_T(x^C_{\mu,i})$$

$$w^E_{ij} E(x^C_{\mu,j})$$

$$w^{s1}_{ij} \sum d^2(x^C_{\mu,i}, x^C_{\mu,j})$$

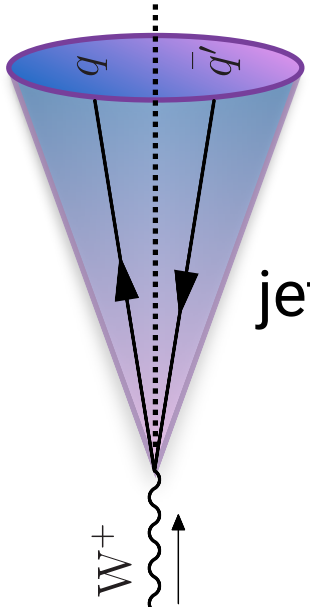
$$w^{s2}_{ij} \sum d^2(x^C_{\mu,i}, x^C_{\mu,j})$$

$$w^{m1}_{ij} \min d^2(x^C_{\mu,i}, x^C_{\mu,j})$$

$$w^{m2}_{ij} \min d^2(x^C_{\mu,i}, x^C_{\mu,j})$$

For each CoLa column, calculate 7 LoLa quantities:

$$d^2(x^C_{\mu,i}, x^C_{\mu,j}) = (x^C_{\mu,i} - x^C_{\mu,j}) g^{\mu\nu} (x^C_{\mu,i} - x^C_{\mu,j})_\nu$$



d^2

jet axis - constituents

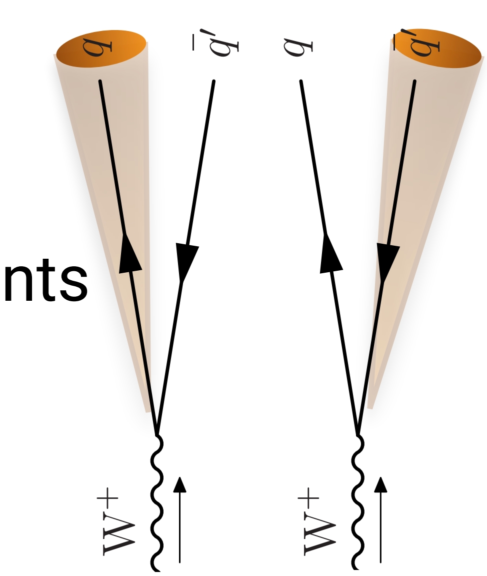
"τ_1"-like!

Minkowski metric!

d^2

subjets - constituents

"τ_2"-like!



Performance

55% signal efficiency per jet increase compared to τ_{21} at given mistag rate

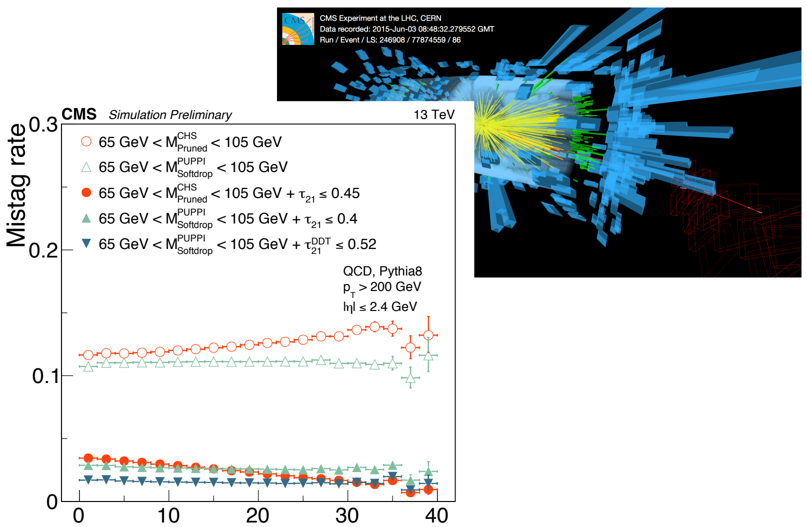
- for analysis requiring 2 tagged jets → 2*x signal efficiency!

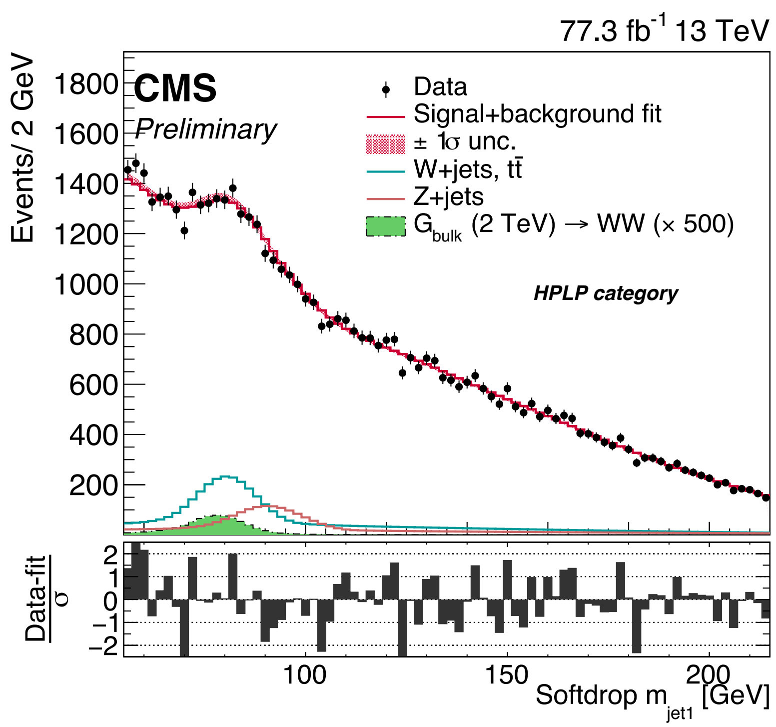
Could lead to large improvement in sensitivity for future VV analyses!

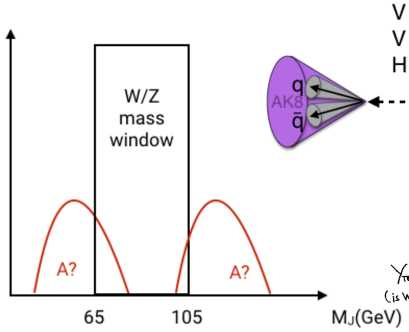
Train and use as "generic" anti-q/g tagger in the future?!*

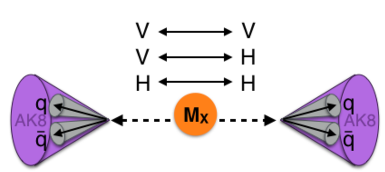
*https://arxiv.org/abs/1808.08979

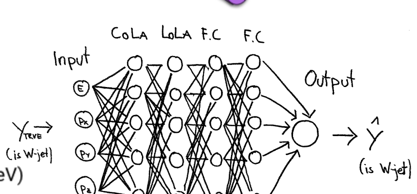
Summary & outlook











- I : First search for diboson resonances at 13 TeV

  Published in Journal of High Energy Physics (2017), DOI: 10.1007/JHEP03(2017)162

- II : A new pileup resistant, perturbative robust tagger

  Published in PRD, DOI: 10.1103/PhysRevD.97.072006; CMS-PAS-B2G-16-021; CMS-PAS-JME-16-003

- III: A novel framework for multi-dimensional searches

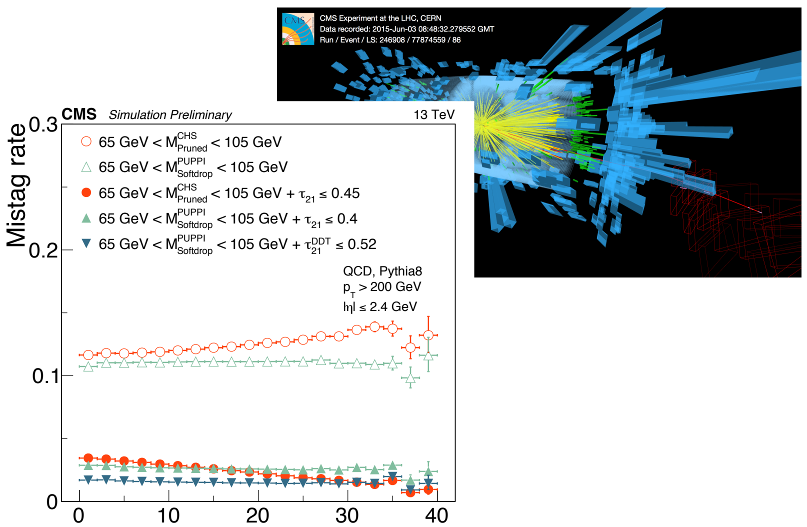
  In progress. To be submitted to The European Physical Journal C

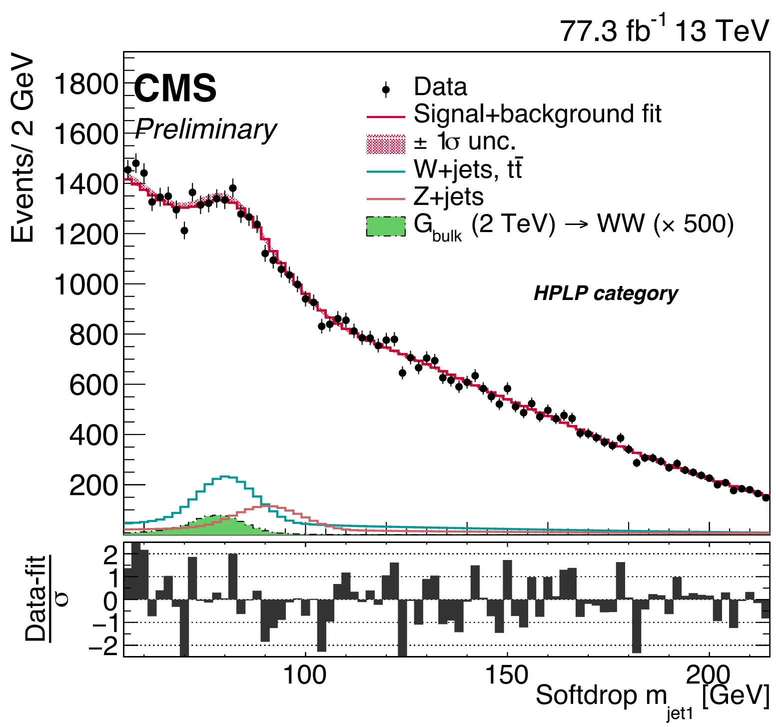
- IV: Encoding jet substructure with a deep neural network

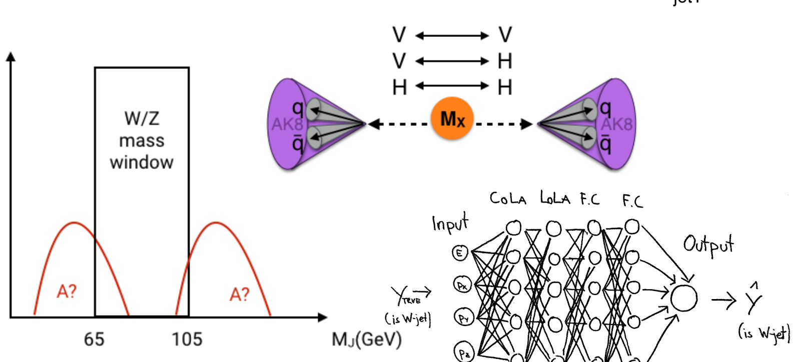
  Work in progress.

Summary & outlook

Not presented today!







- I : First search for diboson resonances at 13 TeV

  Published in Journal of High Energy Physics (2017), DOI: 10.1007/JHEP03(2017)162
  - Double Higgs tagging algorithm training/performance
  - W-tag scale factor method improvement for 2015 data
  - Study of CHS softdrop, optimization of W-tagging algorithm
- II : A new pileup resistant, perturbative robust tagger

  Published in PRD, DOI: 10.1103/PhysRevD.97.072006; CMS-PAS-B2G-16-021; CMS-PAS-JME-16-003
  - W-tag scalefactors for 2016 data, jet mass corrected+pruning
  - Study of other taggers (dichroic, τ_{21} with grooming)
  - CMS tracker offline quality monitoring
- III: A novel framework for multi-dimensional searches

  In progress. To be submitted to The European Physical Journal C
  - W-tag scalefactors for 2017 data (two taggers)
  - Measurement of W-tag p_T-dependence, jet mass scale and resolution
  - Barrel pixel gain calibrations for 2018 data taking
- IV: Encoding jet substructure with a deep neural network

  Work in progress.

Backup

3D fit

Analysis strategy

Select two high-p_T AK8 jets, $|\Delta\eta_{jj}| < 1.3$

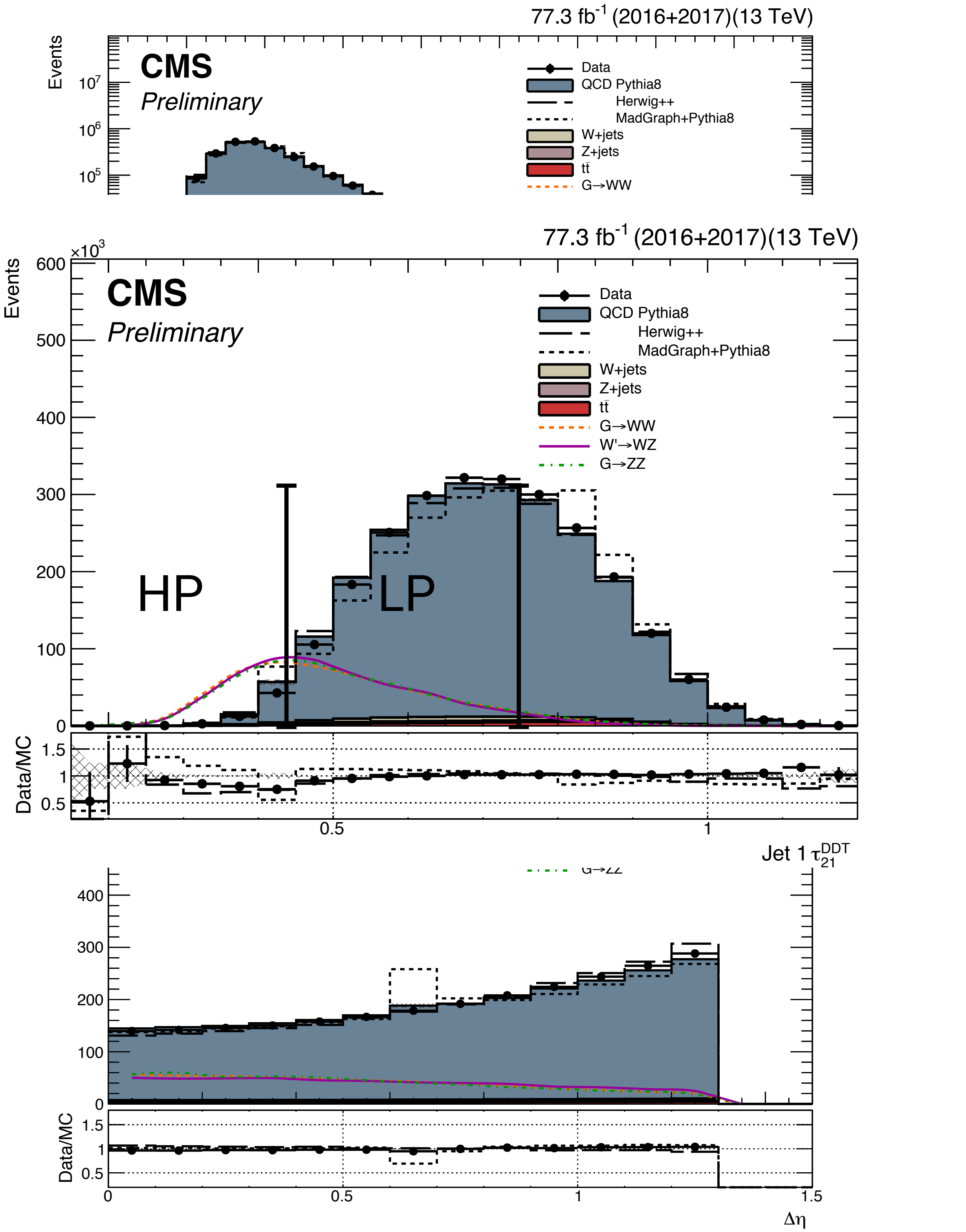
- random sorting of m_{jet1} and m_{jet2} to avoid bias in jet mass shapes

Substructure: τ_{21}^{DDT} with two signal categories

- HPHP: $jet_1 + jet_2$ with $\tau_{21}^{DDT} < 0.43$:
- HPLP: $jet_{1/2}$ with $\tau_{21}^{DDT} < 0.43$
  $jet_{2/1}$ with $0.43 < \tau_{21}^{DDT} < 0.79$:

Bump hunt in m_{jj} -m_{jet1}-m_{jet2} mass plane

- 55 GeV < m_{jet} < 215 GeV (limited by PF reco)
- 1126 GeV < m_{jj} < 5.5 TeV (trigger-limited, see later)



Trigger turn-on

Combination of HT/substructure triggers

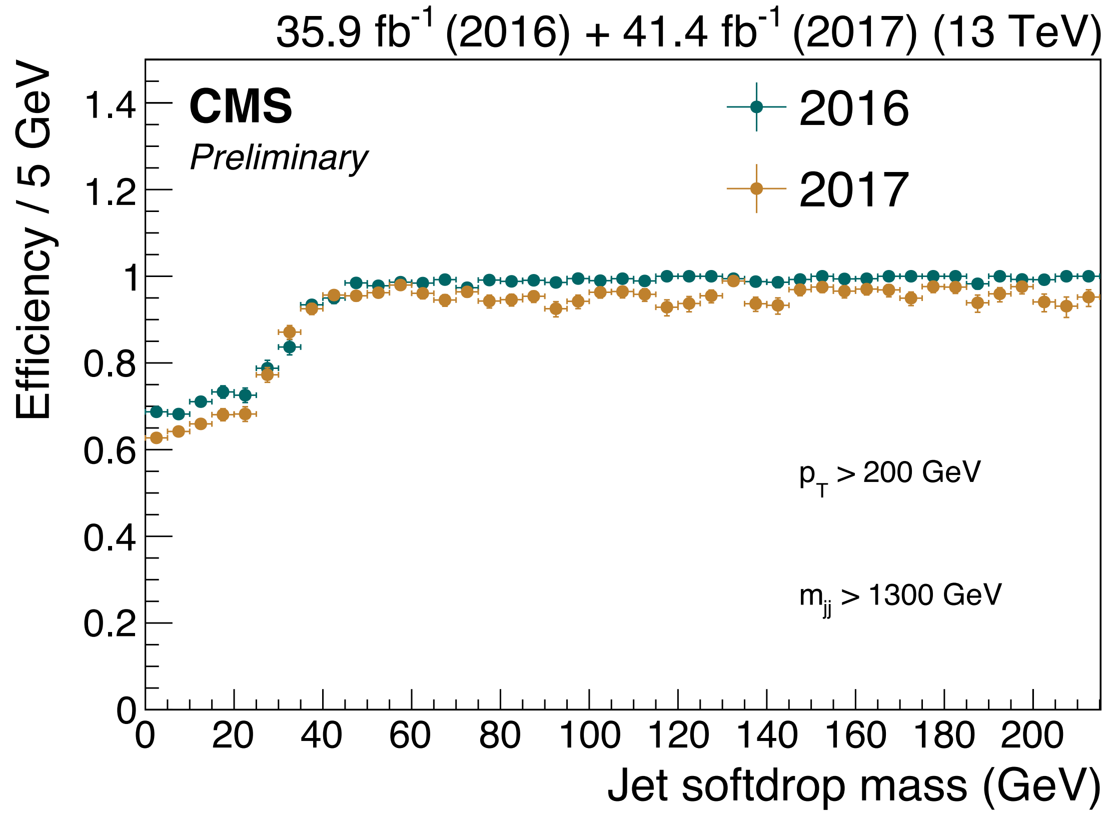
- AK8PFJet500 (p_T 450 in 2016)
- AK8PFHT*_TrimMass50 (Trim 30 in 2016)
- PFHT1050 (HT 800 on 2016)

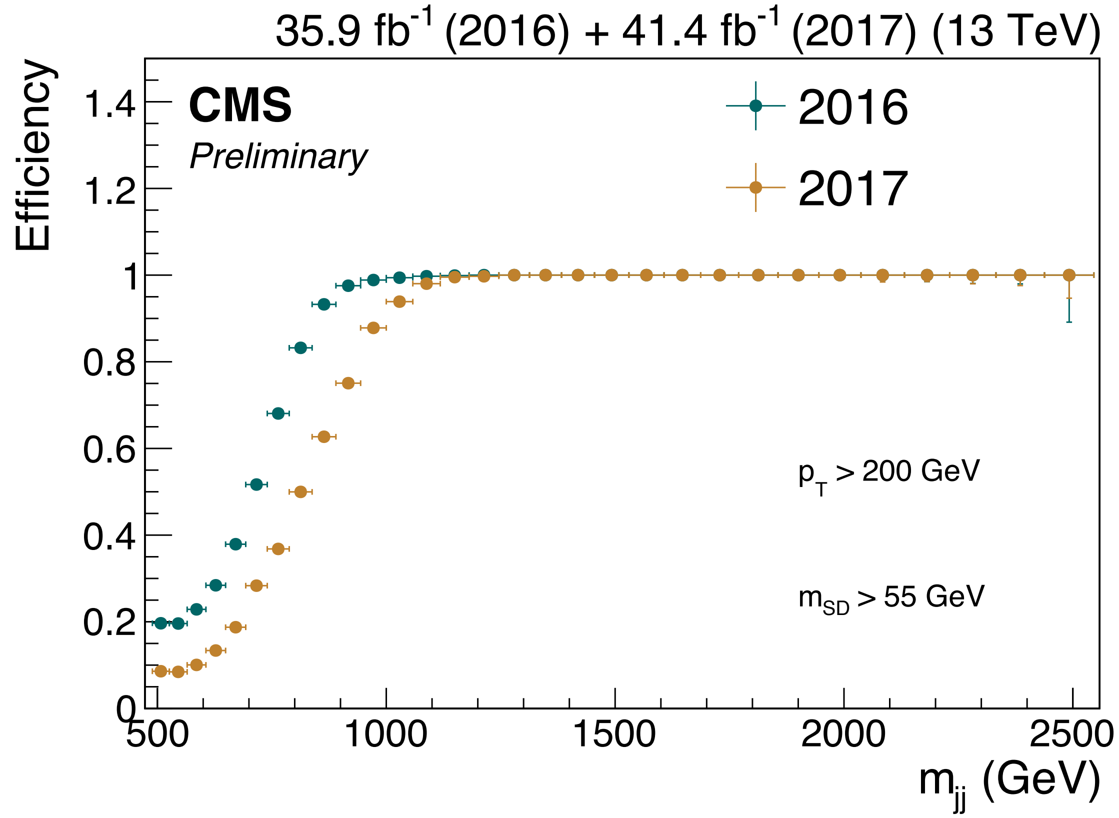
Evaluated in unbiased Single Muon dataset with reference triggers: IsoMu27 or Mu50

Trigger fully efficient at

- m_{jj} > 990 GeV (2016)
- m_{jj} > 1126 GeV (2017)

Sets analysis threshold at 1126 GeV



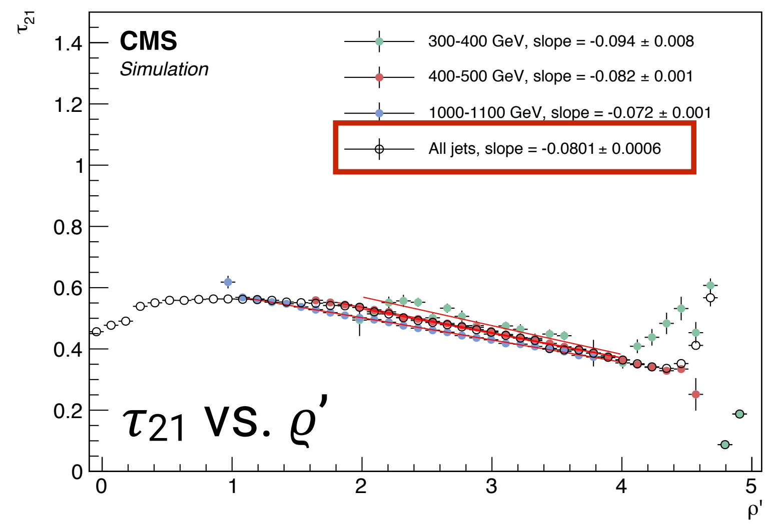


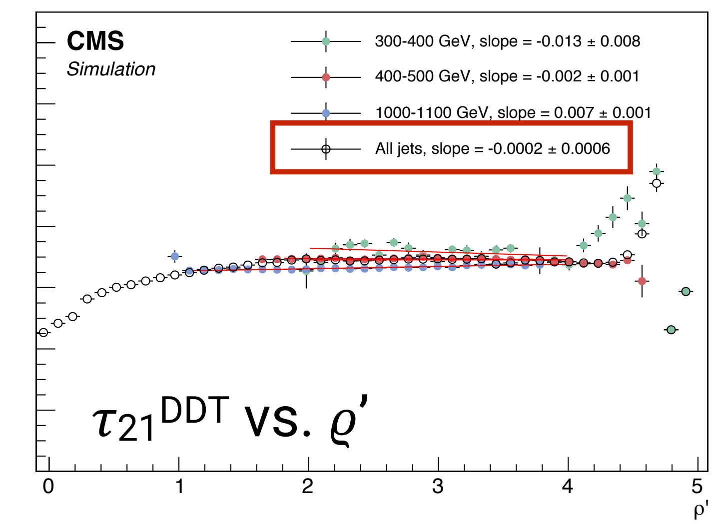
Decorrelating τ_{21}

To improve statistical power of τ_{21}, <u>decorrelate</u> from softdrop mass/p_T

Fit linear part of ϱ' = $m^2/(p_T$*1 GeV) vs. τ_{21} in QCD MC

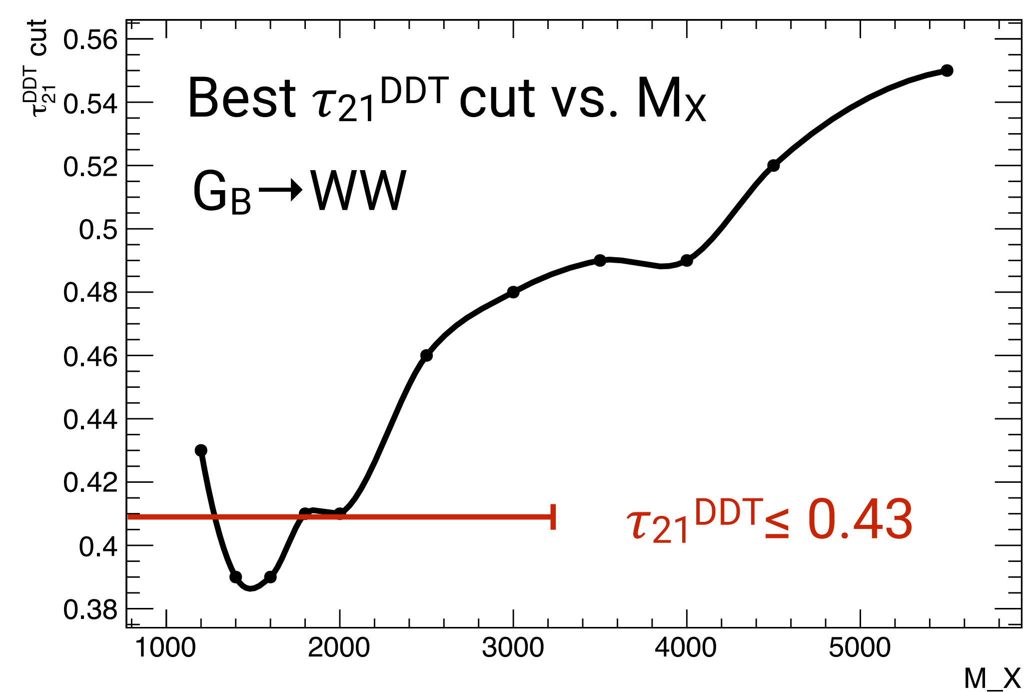
$$\tau_{21}^{DDT} = \tau_{21} + 0.080 \times \log\left(\frac{m^2}{p_T \times 1\ \mathrm{GeV}}\right)$$





Working point optimised for

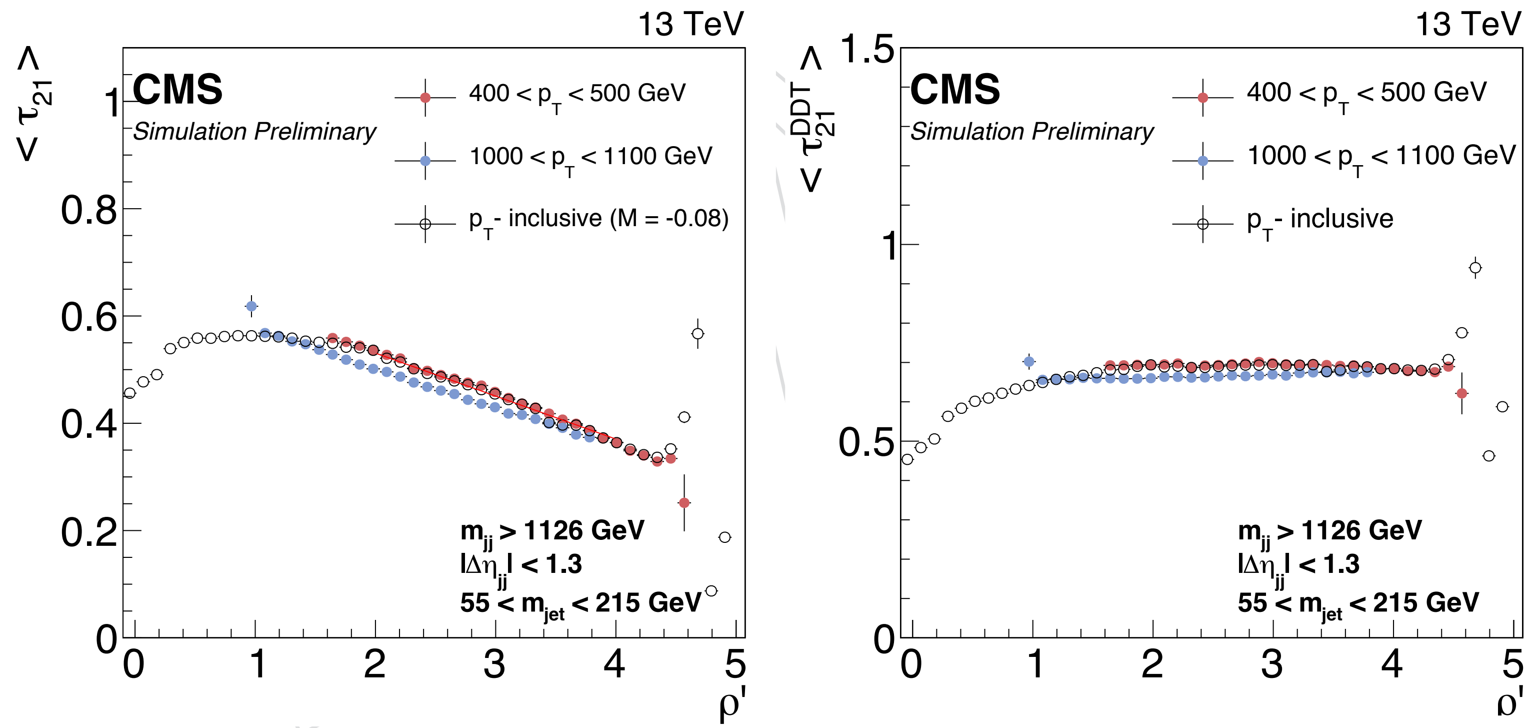
- HP: highest Punzi significance
- LP: contain >95% of signal
  + highest significance



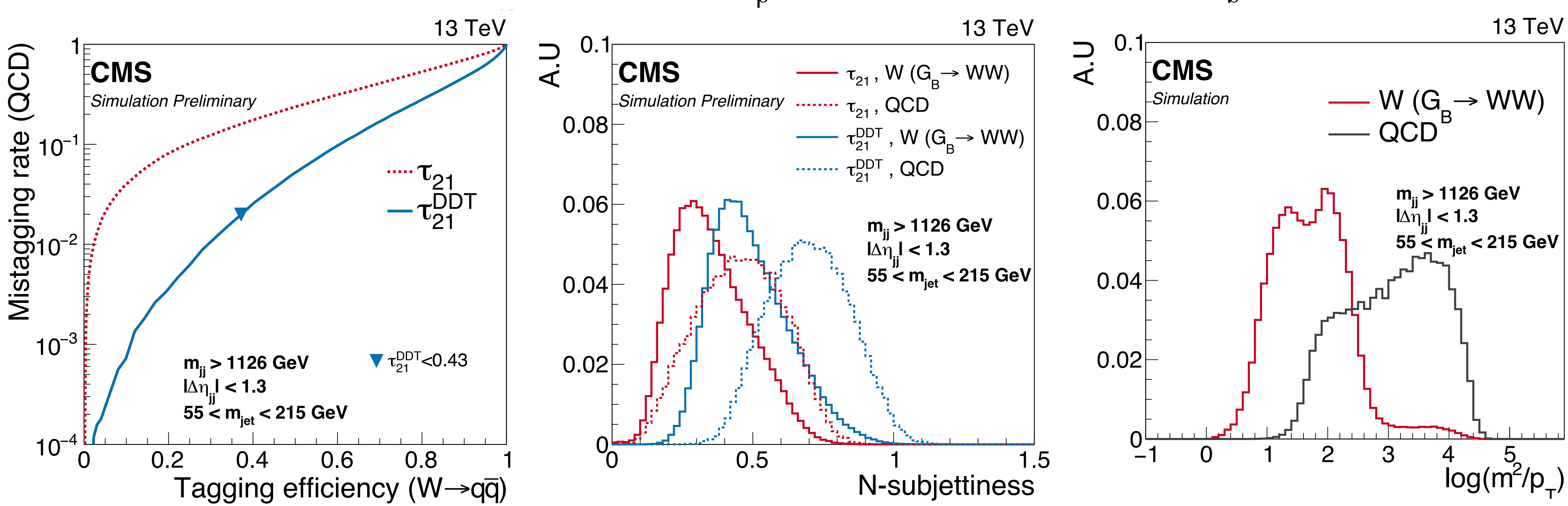
Extremely tight due to high background

- compared to default W-tagger (τ_{21}<0.35) ε_S 50% lower, background reduced by 90%

Tau21DDT

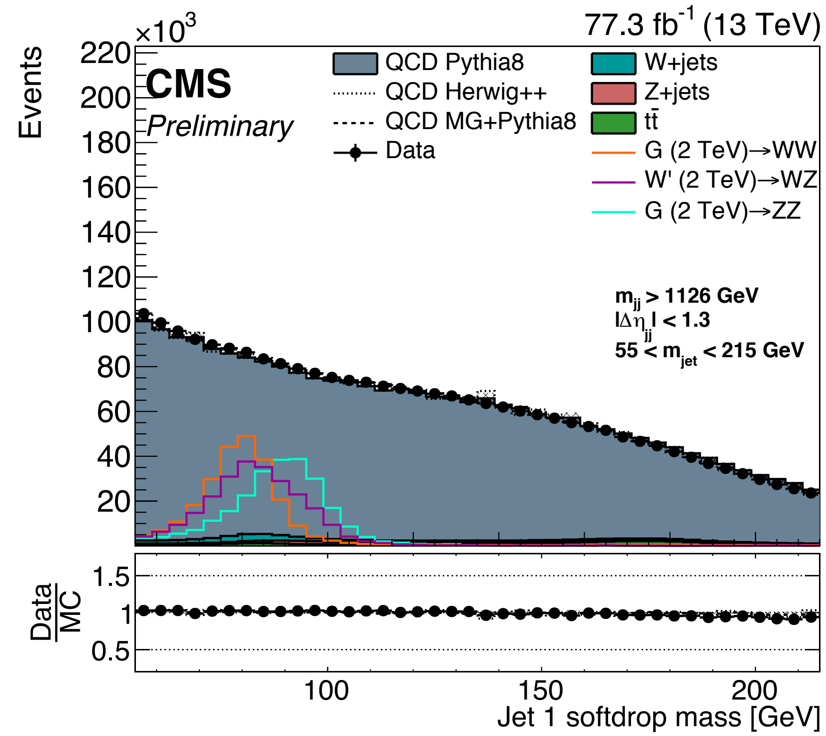


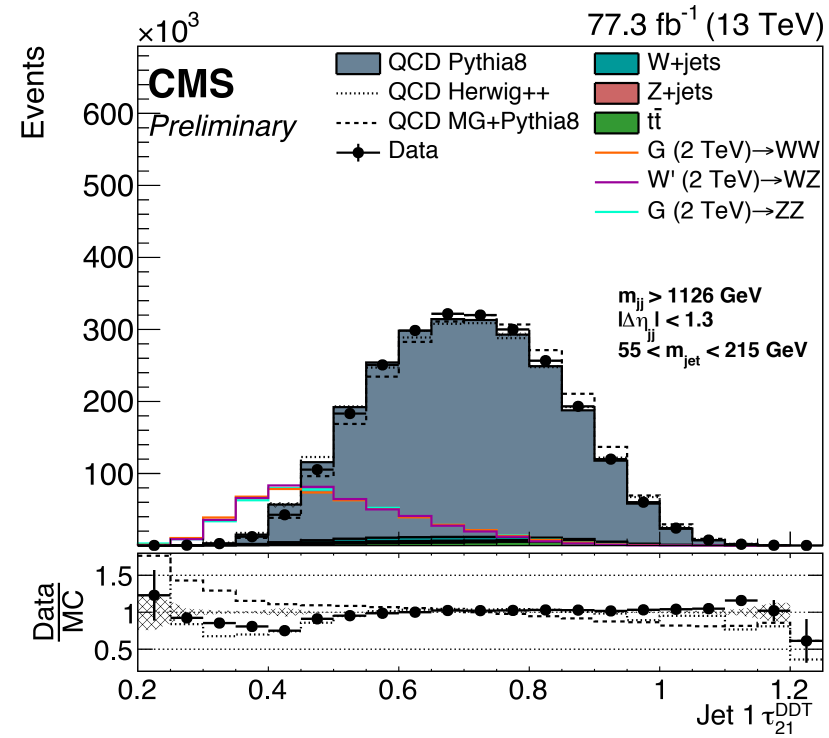
CMS
Simulation Preliminary
13 TeV
$< \tau_{21} >$
$400 < p_T < 500$ GeV
$1000 < p_T < 1100$ GeV
p_T- inclusive (M = -0.08)
$m_{jj} > 1126$ GeV
$|\Delta\eta_{jj}| < 1.3$
$55 < m_{jet} < 215$ GeV
ρ'
$< \tau_{21}^{DDT} >$
p_T- inclusive

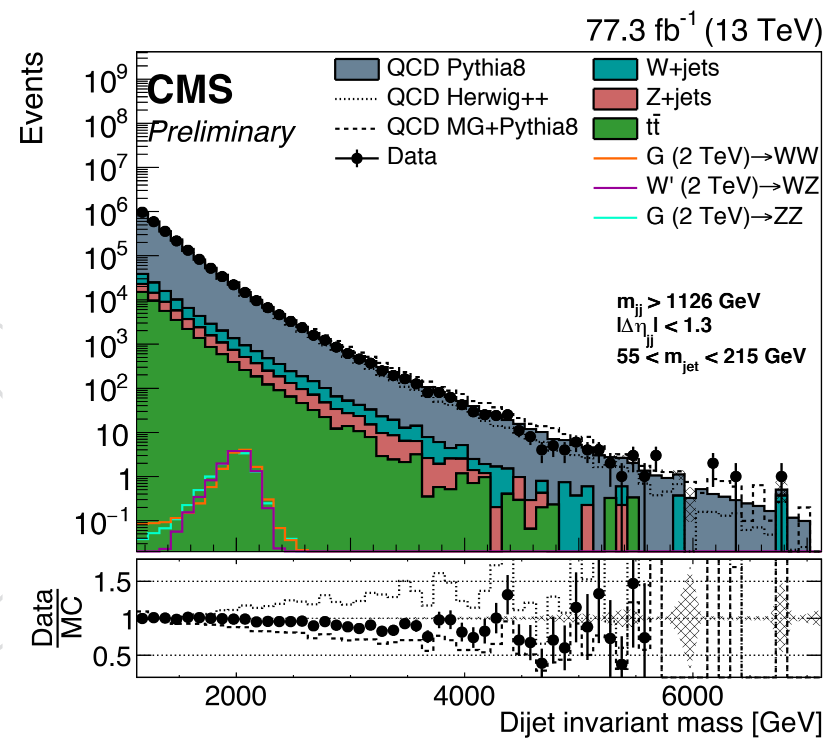


CMS
Simulation Preliminary
13 TeV
Mistagging rate (QCD)
Tagging efficiency ($W \to q\bar{q}$)
τ_{21}
τ_{21}^{DDT}
$\tau_{21}^{DDT} < 0.43$
$m_{jj} > 1126$ GeV
$|\Delta\eta_{jj}| < 1.3$
$55 < m_{jet} < 215$ GeV
A.U
τ_{21}, W ($G_B \to$ WW)
τ_{21}, QCD
τ_{21}^{DDT}, W ($G_B \to$ WW)
τ_{21}^{DDT}, QCD
N-subjettiness
Simulation
W ($G_B \to$ WW)
QCD
$\log(m^2/p_T)$

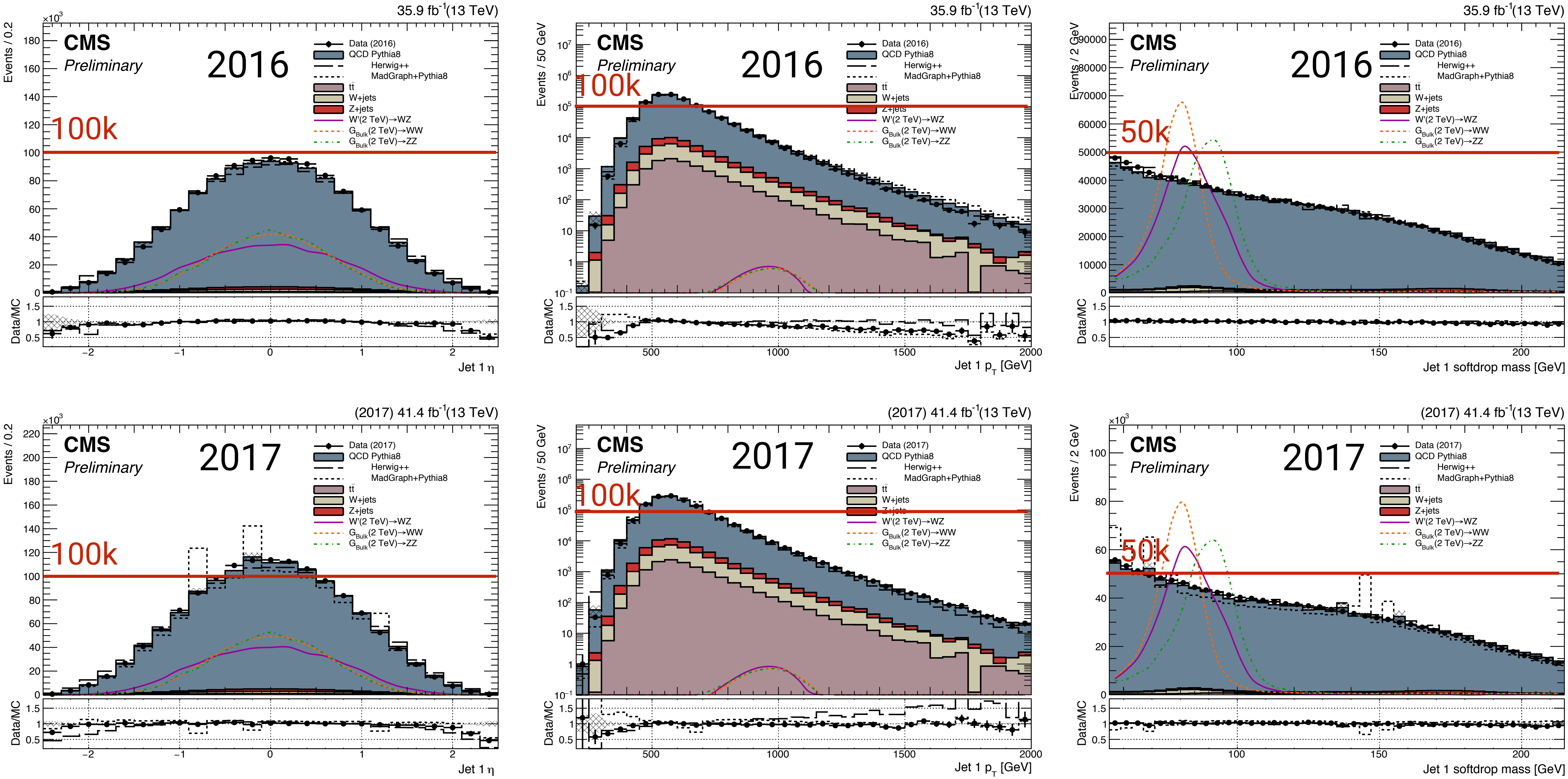
Controlplots in 16+17 data







2016 versus 2017



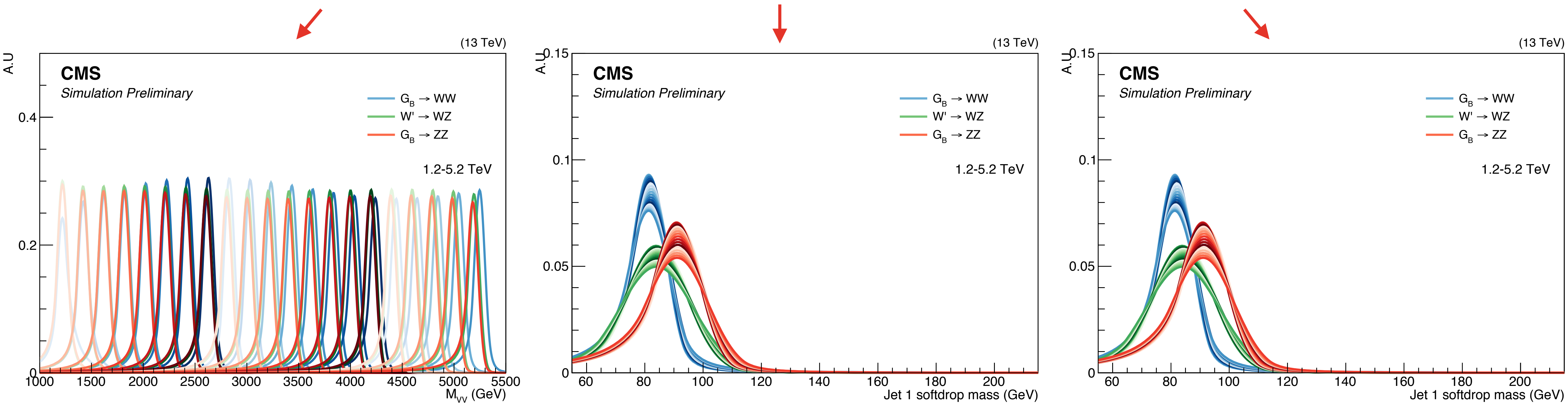
Similar spectra in 2016 vs. 2017, ~15-20% higher yield in 2017 dataset

Signal modelling

Templates are product of resonance mass and jet mass shapes (m_j/m_{jj} uncorrelated)

$$\mathbf{P_{SIG}(m_{jj}, m_{jet1}, m_{jet2})} =$$

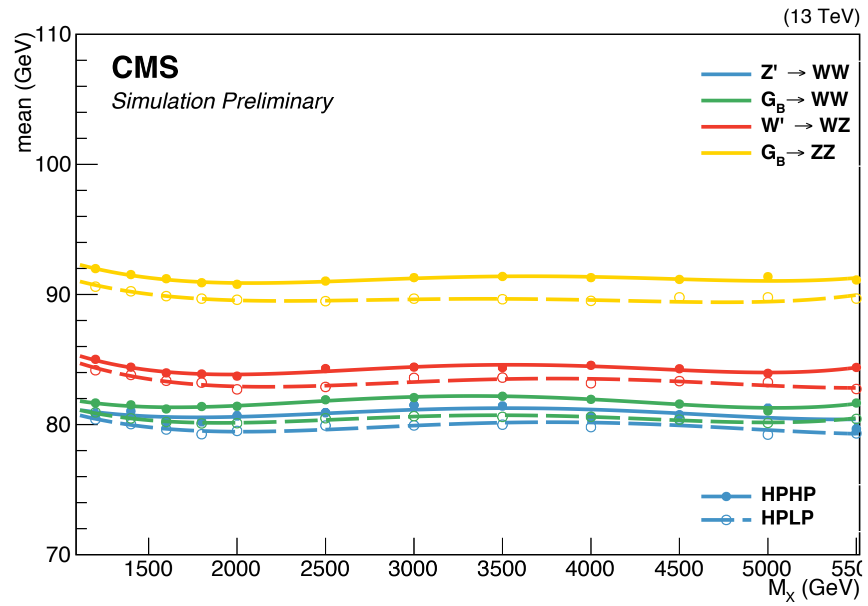
$$P_{jj}\,(\,m_{jj}\,|\,\theta_1(m_X)\,)\;\; x\;\; P_j\,(\,m_{jet1}\,|\,\theta_2(m_X)\,)\;\; x\;\; P_j\,(\,m_{jet2}\,|\,\theta_2(m_X)\,)$$

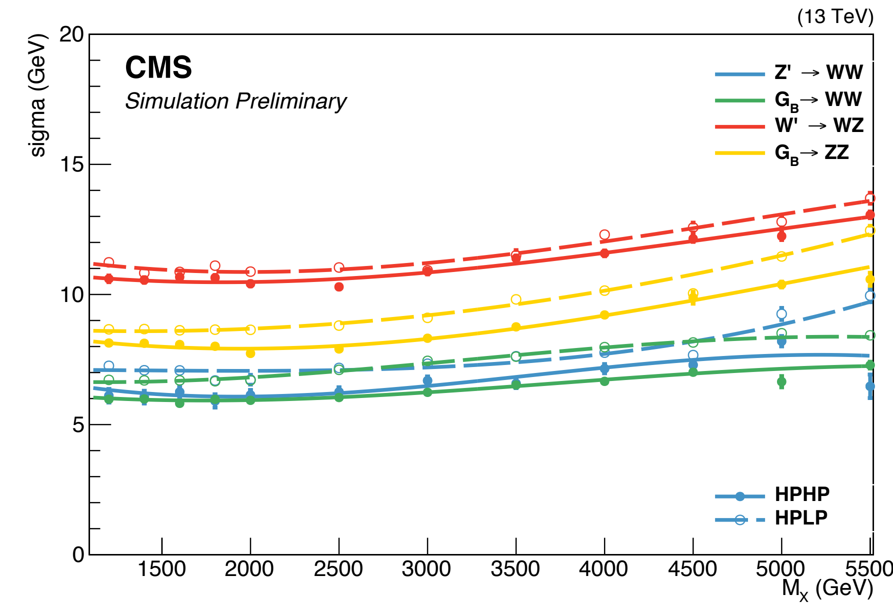


Fit softdrop jet mass and dijet mass with double CB, parametrise as function of m_X

Signal modelling

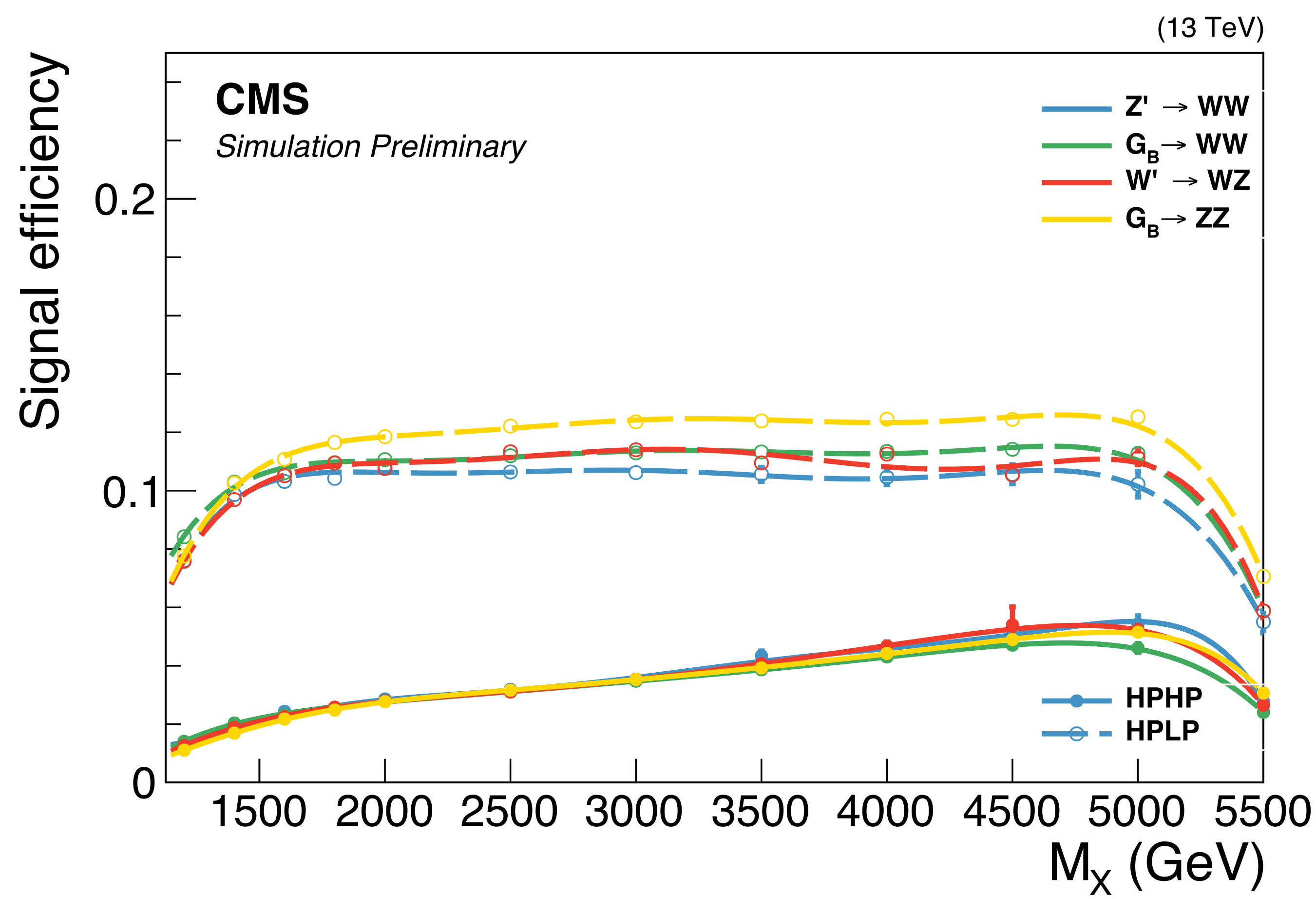
Mean and width stable due to decorrelated τ_{21}



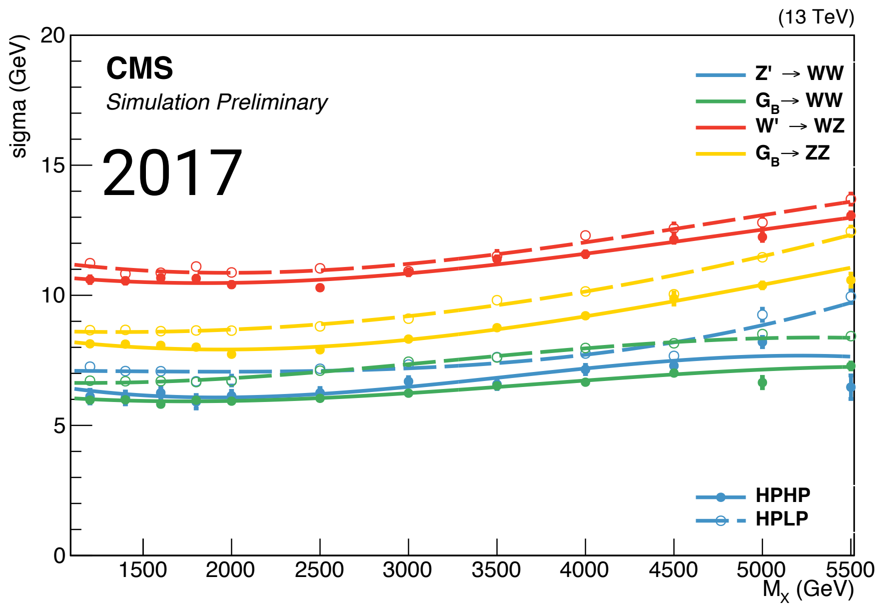


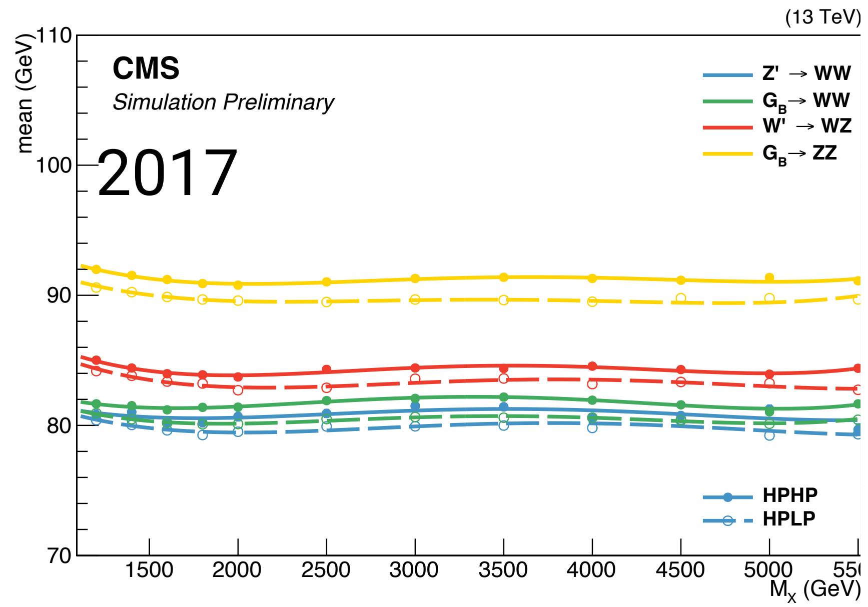
Yield from integral of m_{VV} histogram

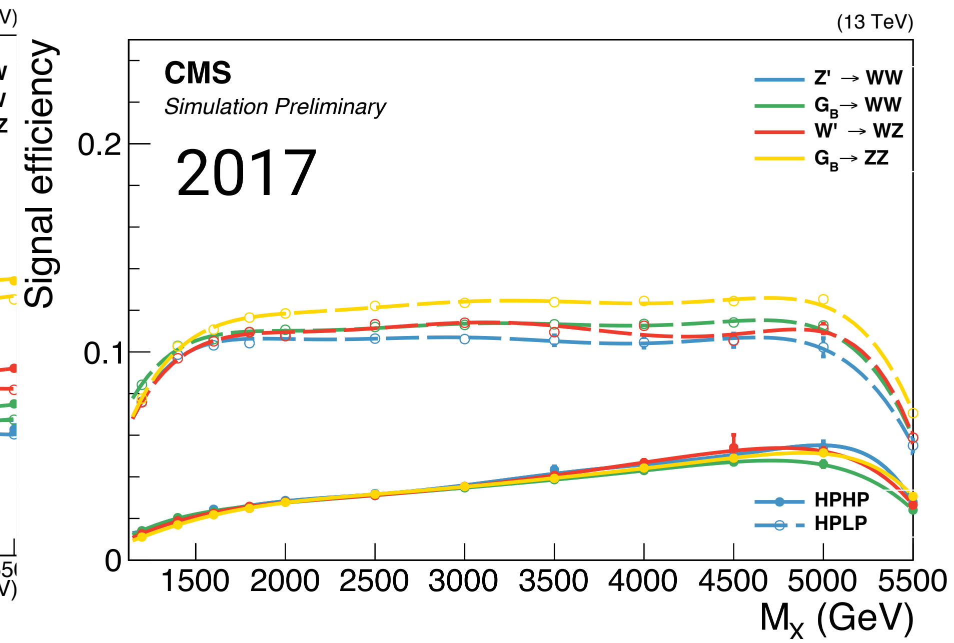
- parametrised as a function of m_X for smooth signal efficiency versus p_T
- efficiency lower at edges due to bin edge cut off
- lower signal efficiency in HP, but background strongly reduced

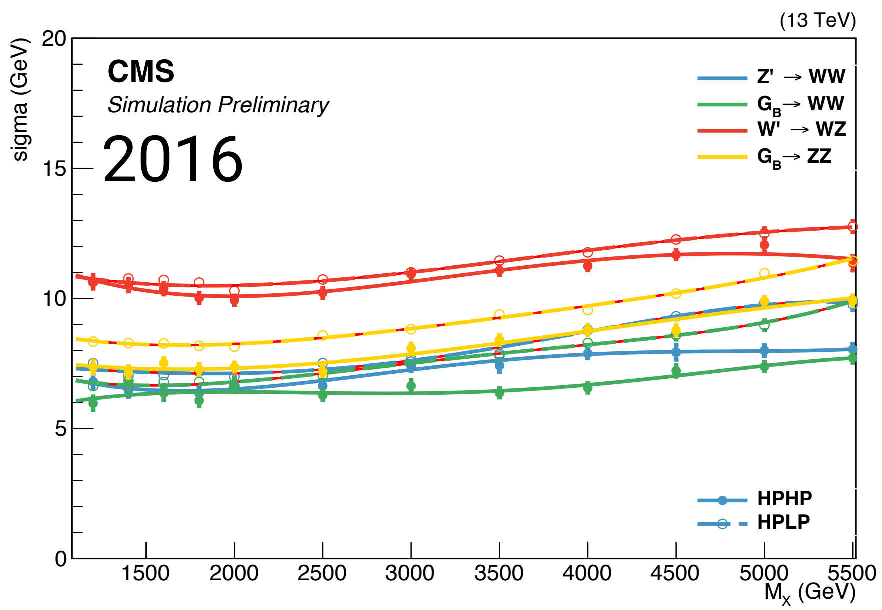


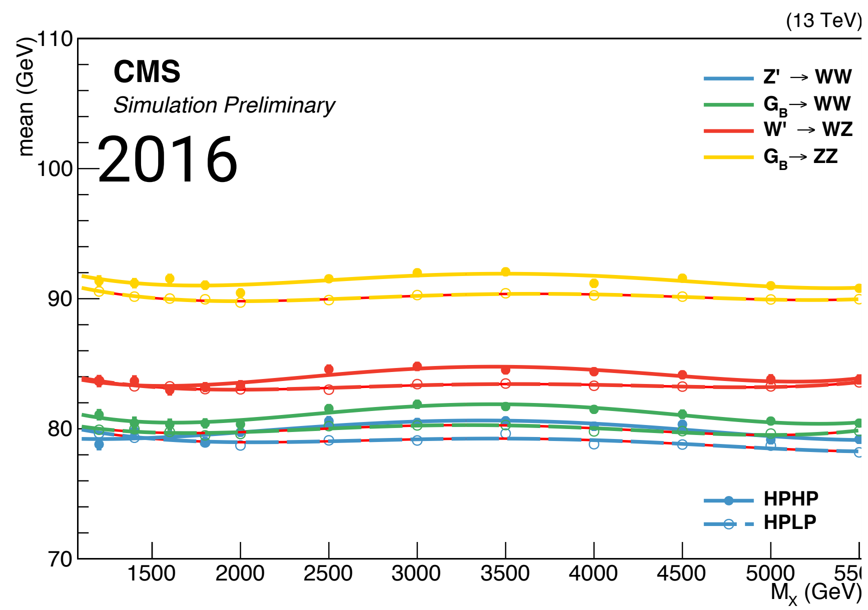
Signal 2016 versus 2017

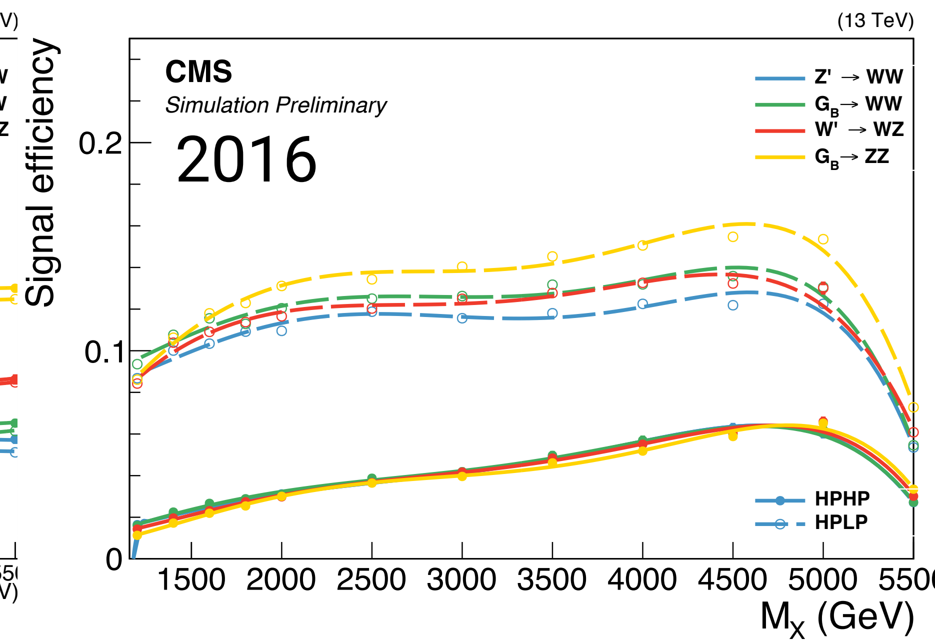












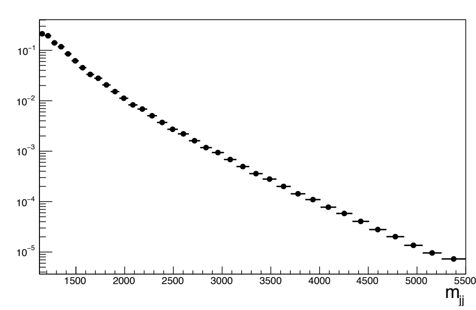
Non-resonant background

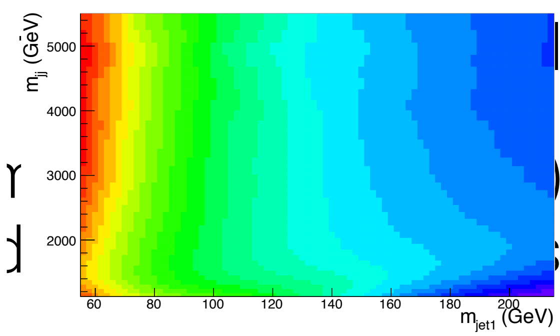
Universität Zürich[UZH]

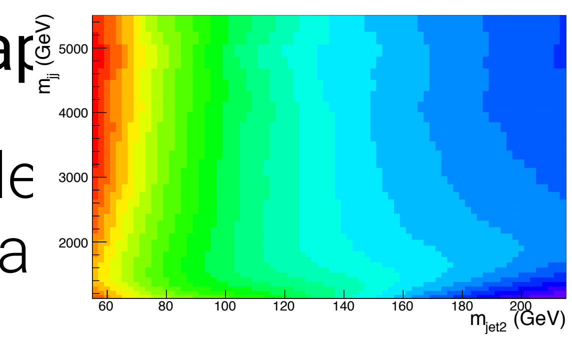
To account for correlations m_{jet}/m_{jj}, non-resonant background modelled conditionally

- $P_{non\text{-}res}(m_{jj}, m_{jet1}, m_{jet2}) = P_{jj}(m_{jj} \mid \theta_1) \times P_j(m_{jet1} \mid m_{jj}, \theta_2) \times P_j(m_{jet2} \mid m_{jj}, \theta_2)$

With 250k bins, need to e



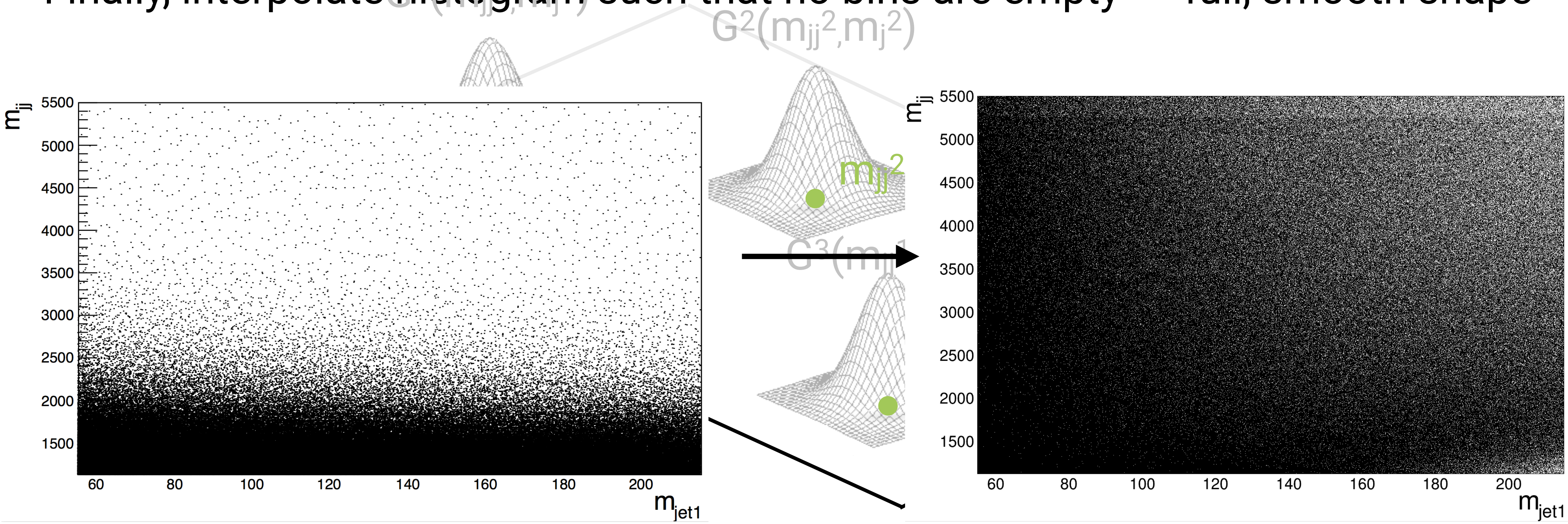




l ap

- rather than filling 1D/2 ...), le
  contribute 1D/2D gau ... sca

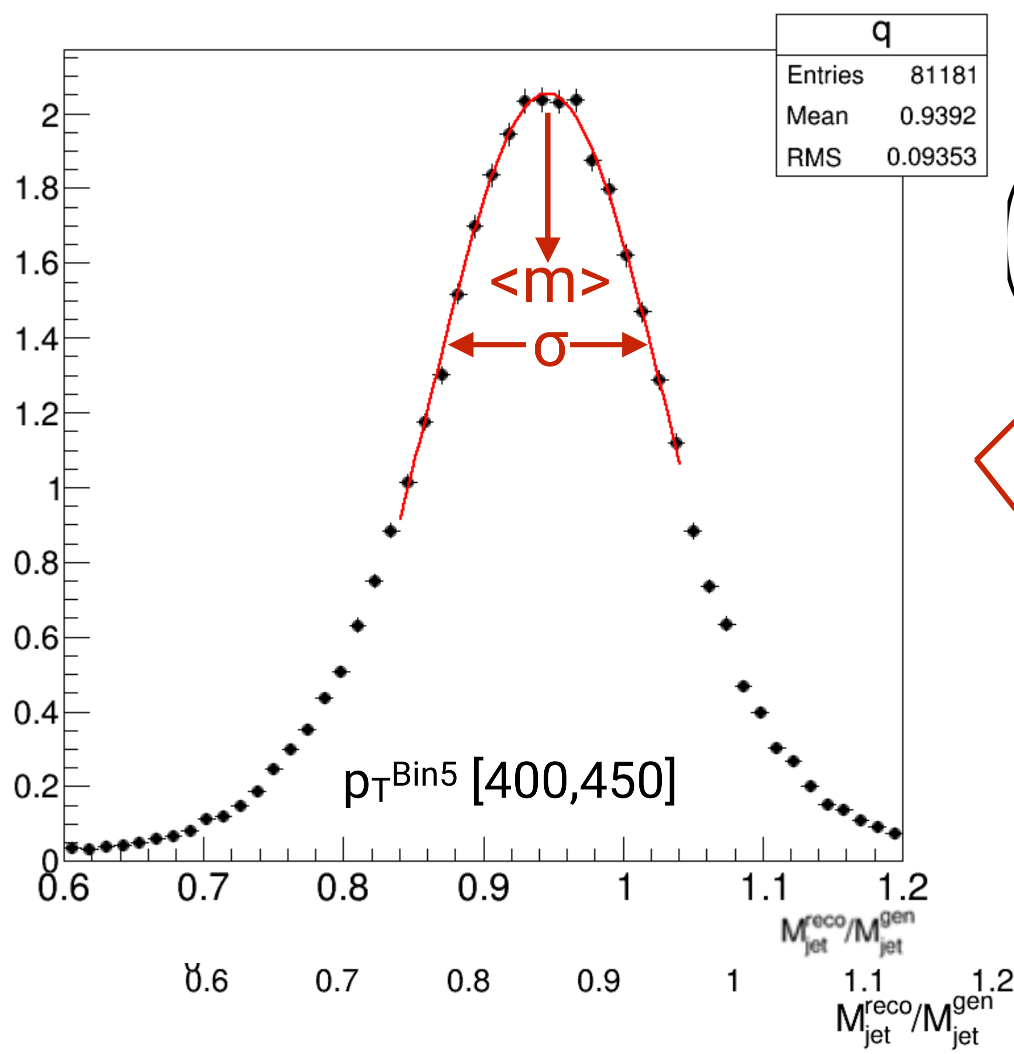
Finally, interpolate histogram such that no bins are empty → full, smooth shape

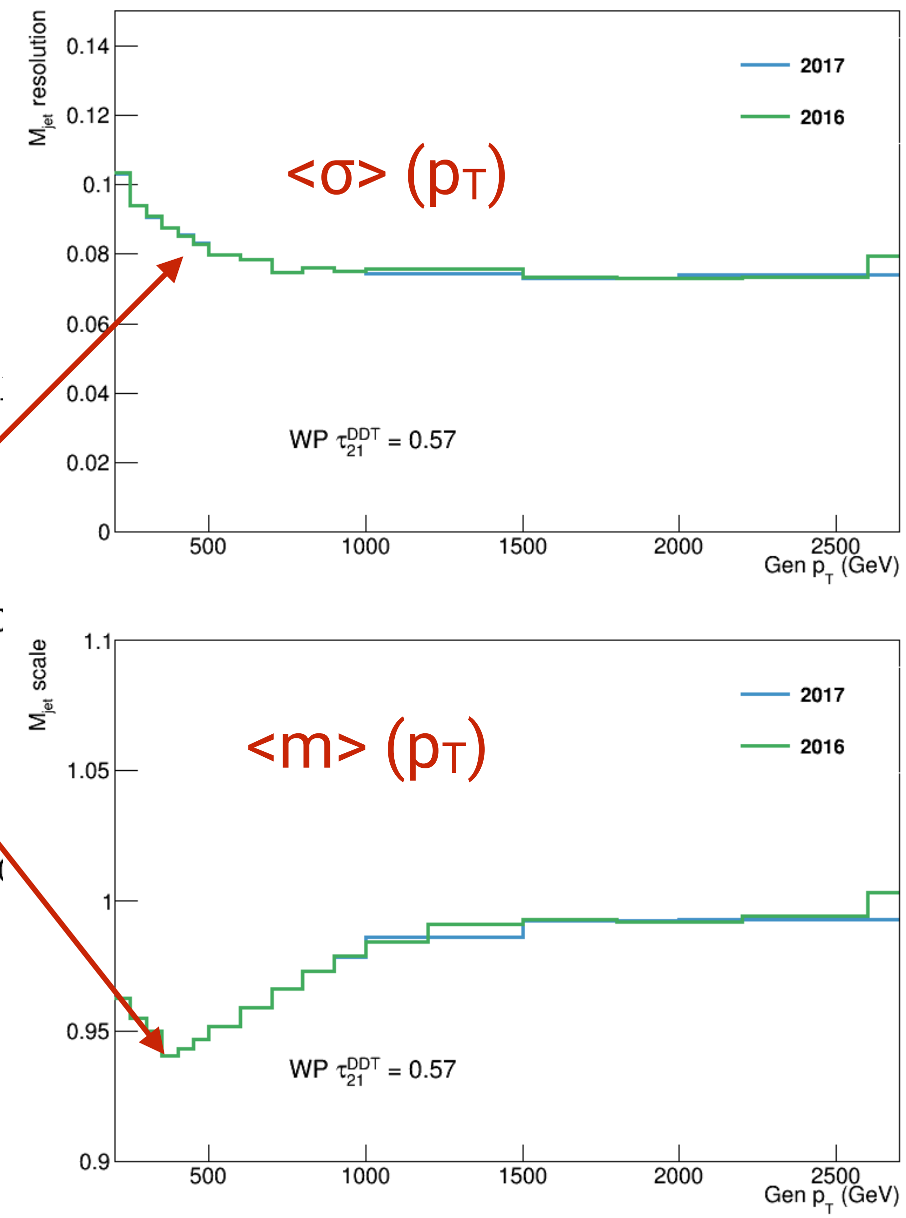


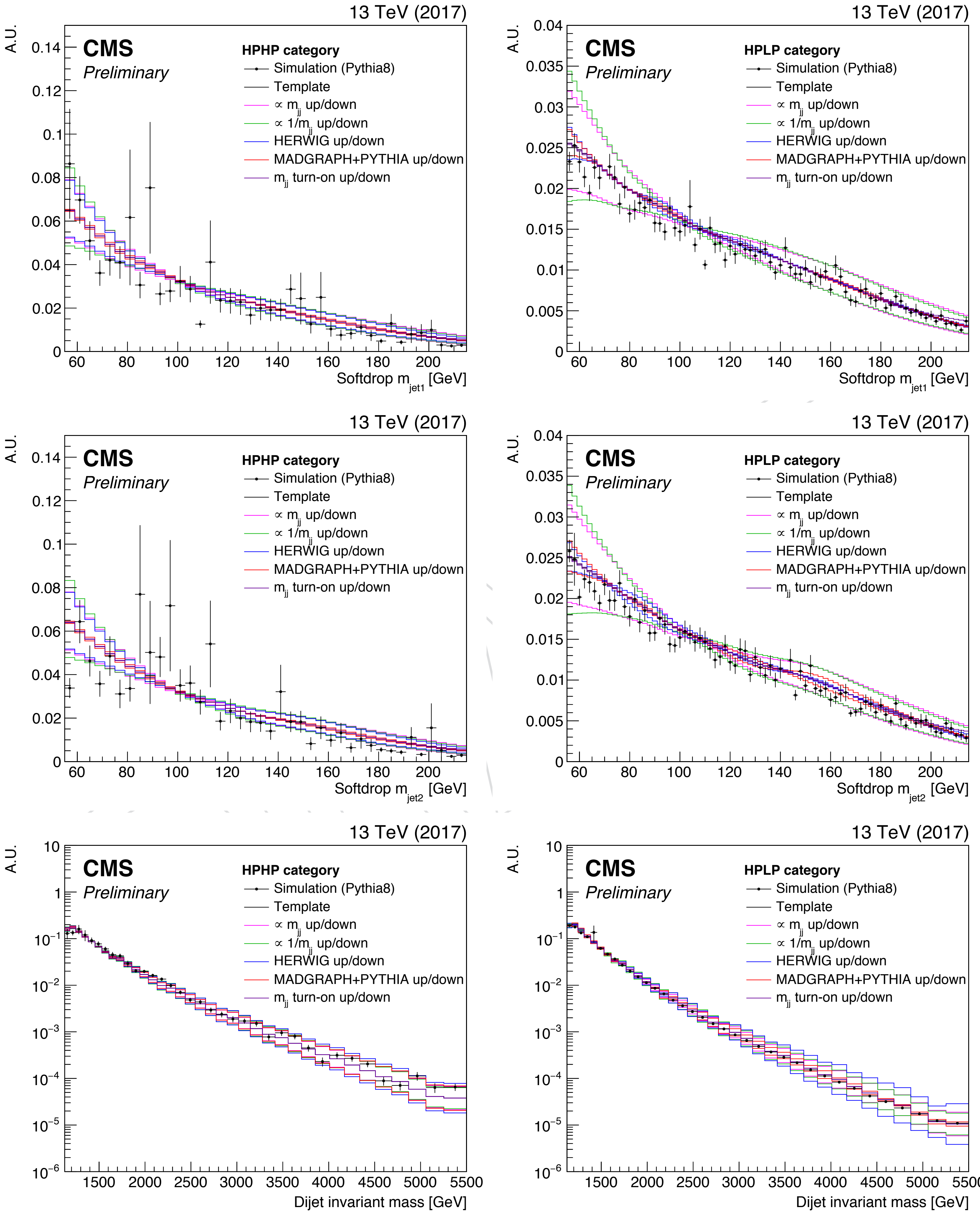
Defining Gaussian kernel

In bins of gen jet p_T, derive jet mass scale and resolution from Gaussian fit to $M_j(reco)/M_j(gen)$

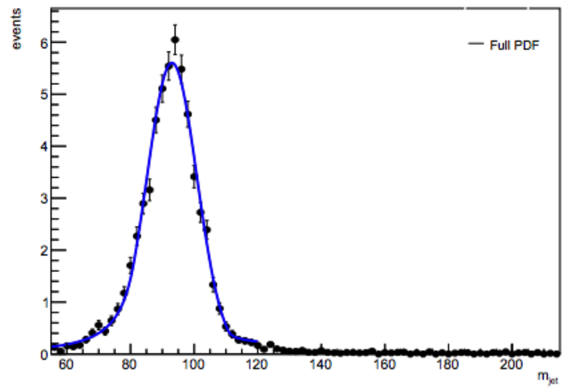
Each event smeared with 1D(M_{JJ})/2D(M_J) Gaussian kernel



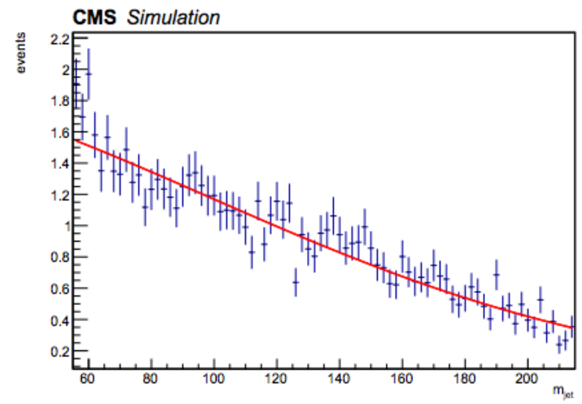



13 TeV (2017)
CMS
Preliminary
HPHP category
Simulation (Pythia8)
Template
∝ m_{jj} up/down
∝ $1/m_{jj}$ up/down
HERWIG up/down
MADGRAPH+PYTHIA up/down
m_{jj} turn-on up/down
A.U.
Softdrop m_{jet1} [GeV]
13 TeV (2017)
CMS
Preliminary
HPLP category
Simulation (Pythia8)
Template
∝ m_{jj} up/down
∝ $1/m_{jj}$ up/down
HERWIG up/down
MADGRAPH+PYTHIA up/down
m_{jj} turn-on up/down
A.U.
Softdrop m_{jet1} [GeV]
13 TeV (2017)
CMS
Preliminary
HPHP category
A.U.
Softdrop m_{jet2} [GeV]
13 TeV (2017)
CMS
Preliminary
HPLP category
A.U.
Softdrop m_{jet2} [GeV]
13 TeV (2017)
CMS
Preliminary
HPHP category
A.U.
Dijet invariant mass [GeV]
13 TeV (2017)
CMS
Preliminary
HPLP category
A.U.
Dijet invariant mass [GeV]

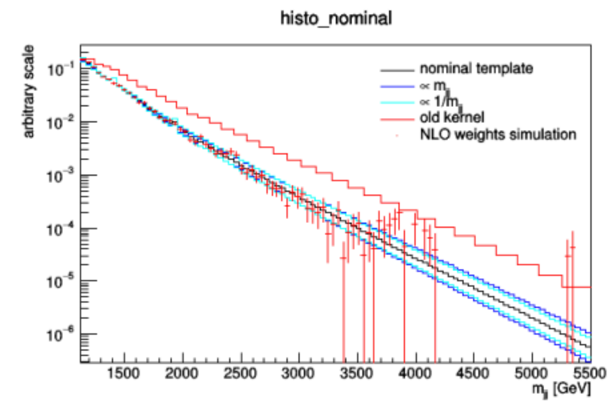
Resonant background



$\times$



$\times$



$$
\begin{aligned}
P_{Vjets}\left(M_{jet1}, M_{jet2}, M_{jj}\right) = f \times &\left(P_{dijet}\left(M_{jj}\right) \times P_{res}(M_{jet1}) \times P_{non-res}(M_{jet2})\right) \\
+(1-f) \times &\left(P_{dijet}\left(M_{jj}\right) \times P_{res}(M_{jet2}) \times P_{non-res}(M_{jet1})\right)
\end{aligned}
$$

fit resonant part of M_{jet} with signal function $\Rightarrow$ fully correlated systematics

model QCD-jet with simple Gaussian

M_{jj} shape, same kernel approach as for QCD

two contribution in final fit: Z+jets and W+jets plus $t\bar{t}$

two p_T dependent corrections applied to V+jets background:

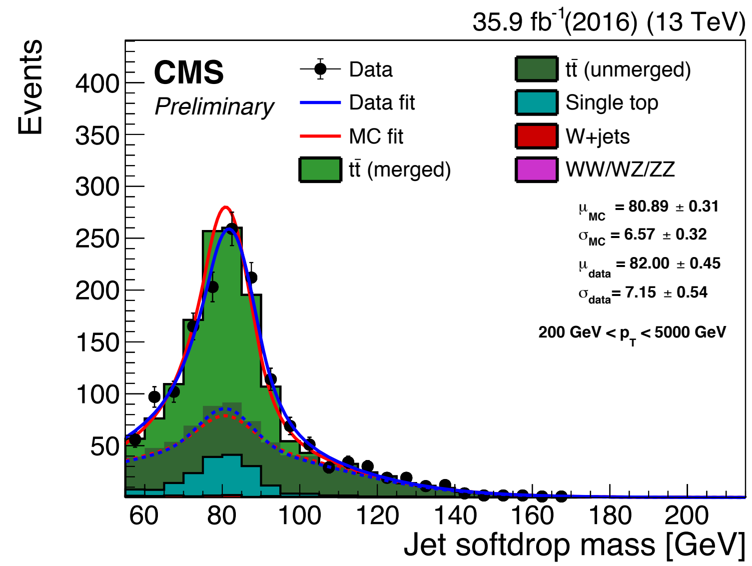
- NLO-kfactor: correction p_T distribution of LO sample to NLO
- electroweak correction: for higher order electroweak processes

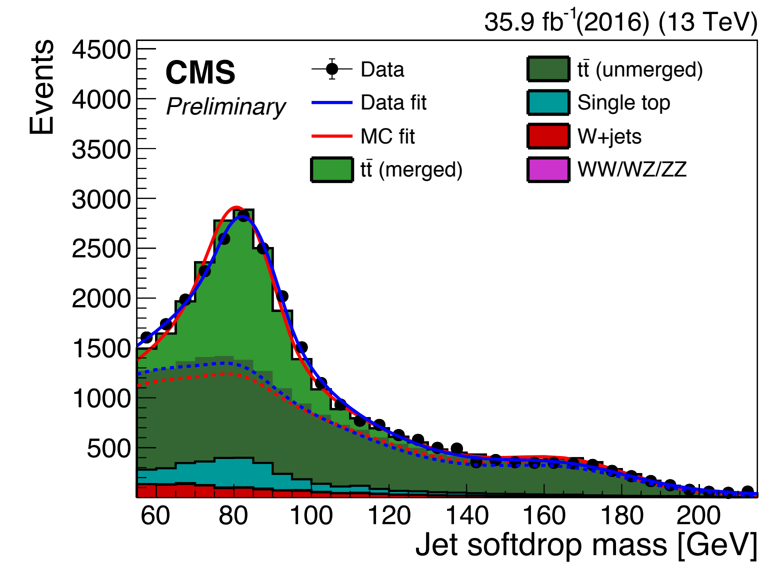
Uncertainties

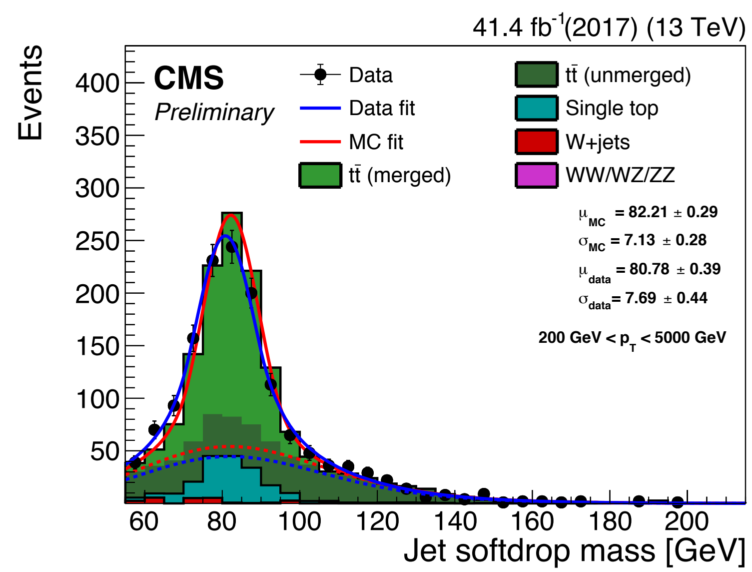
Source	Relevant quantity	HPHP unc. (%)	HPLP unc. (%)
PDFs	Signal yield	3	
W-tagging efficiency	Signal+ V+jets yield	25 (21)	13 (11)
W-tagging p_T-dependence	Signal+ V+jets yield	8-23	9-25
Integrated luminosity	Signal+ V+jets yield	2.3 (2.6)	
QCD normalisation	Background yield	50	
V+jets normalisation	Background yield	10	
V+jets ratio	Migration	10	
PDFs	Signal M_{VV}/M_j mean and width	< 1	
Jet energy scale	Signal M_{VV} mean	2	
Jet energy resolution	Signal M_{VV} width	5	
Jet mass scale	Signal + V+jets M_j mean	1	
Jet mass resolution	Signal + V+jets M_j width	8	
QCD HERWIG++	QCD shape	–	
QCD MADGRAPH+PYTHIA8	QCD shape	–	
p_T-variations	QCD shape	–	
Scale-variations	QCD shape	–	
High-M_j turn-on	QCD shape	–	
p_T-variations	V+jets M_{VV} shape	–	

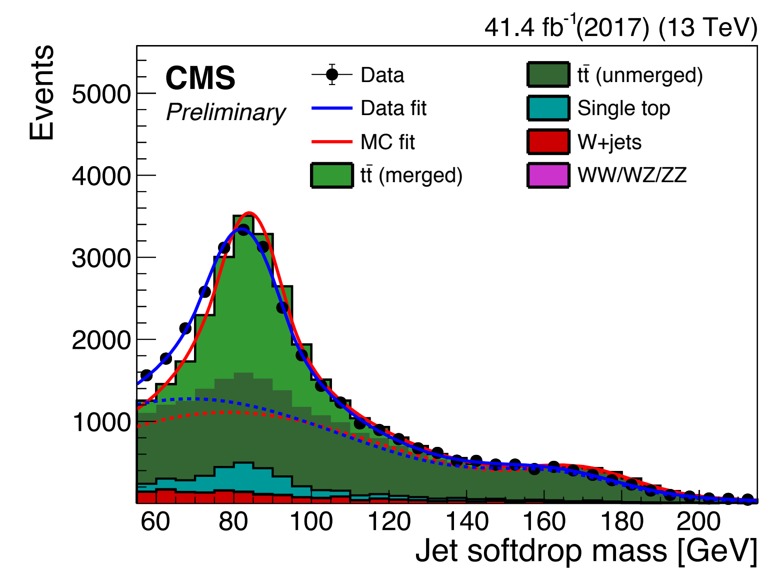
W-tagging scalefactors

	m [GeV]	σ [GeV]	W-tag efficiency
2016			
$\tau_{21}^{DDT} < 0.43$			
Data	82.0± 0.5 (stat.)	7.1 ± 0.5 (stat.)	0.080 ± 0.008 (stat.)
Simulation	80.9± 0.2 (stat.)	6.6 ± 0.2 (stat.)	0.085 ± 0.003 (stat.)
Data/simulation	1.014± 0.007 (stat.+sys.)	1.09 ± 0.09 (stat.+sys.)	0.94 ± 0.10 (stat.+sys.)
$0.43 < \tau_{21}^{DDT} < 0.79$			
Data			0.920 ± 0.008 (stat.)
Simulation			0.915 ± 0.003 (stat.)
Data/simulation			1.006 ± 0.009 (stat.+sys.)
2017			
$\tau_{21}^{DDT} < 0.43$			
Data	80.8± 0.4 (stat.)	7.7 ± 0.4 (stat.)	0.060 ± 0.006 (stat.)
Simulation	82.2± 0.3 (stat.)	7.1 ± 0.3 (stat.)	0.070 ± 0.005 (stat.)
Data/simulation	0.983± 0.007 (stat.+sys.)	1.08 ± 0.08 (stat.+sys.)	0.96 ± 0.12 (stat.+sys.)
$0.43 < \tau_{21}^{DDT} < 0.79$			
Data			0.935 ± 0.006 (stat.)
Simulation			0.932 ± 0.005 (stat.)
Data/simulation			1.003 ± 0.008 (stat.+sys.)

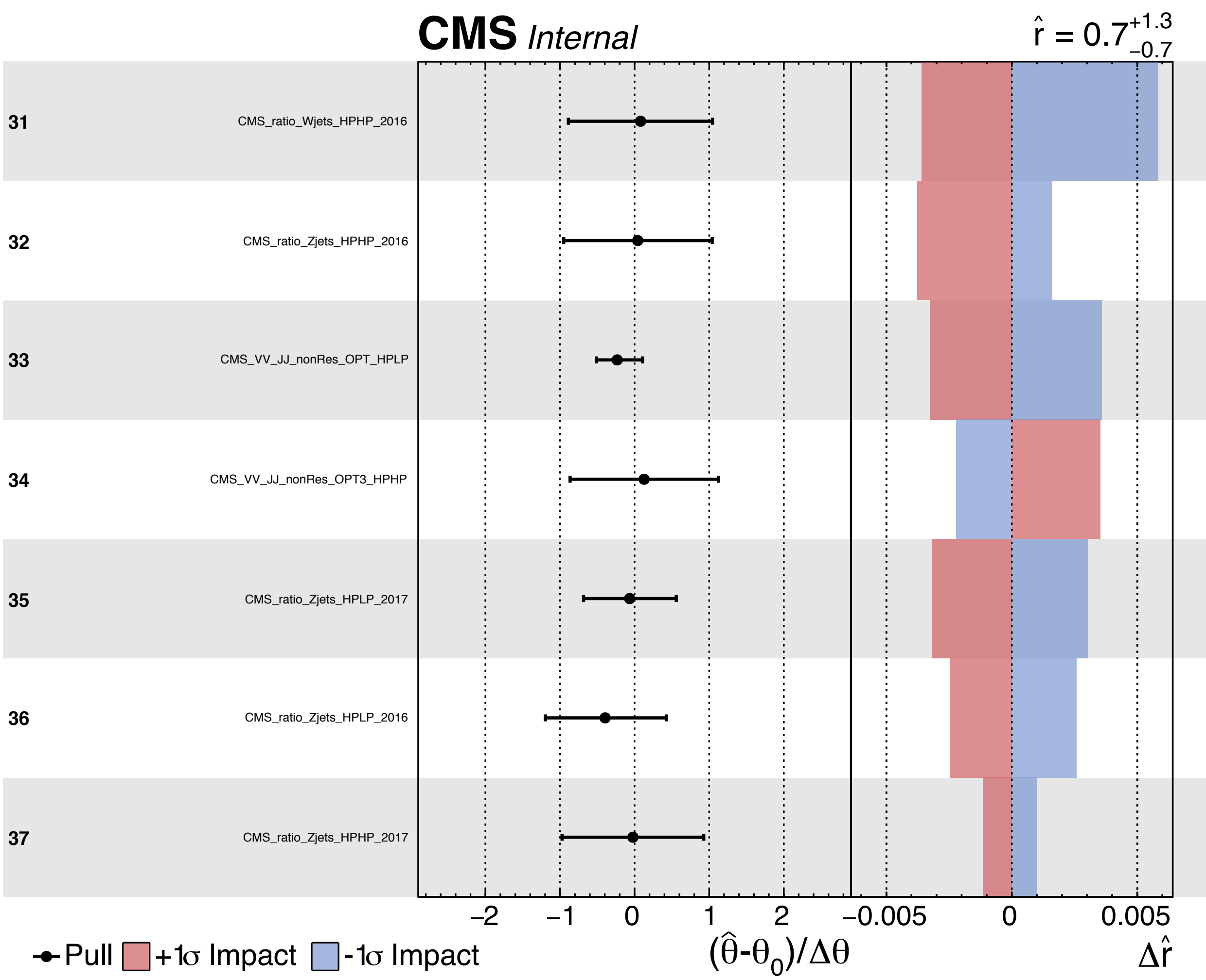




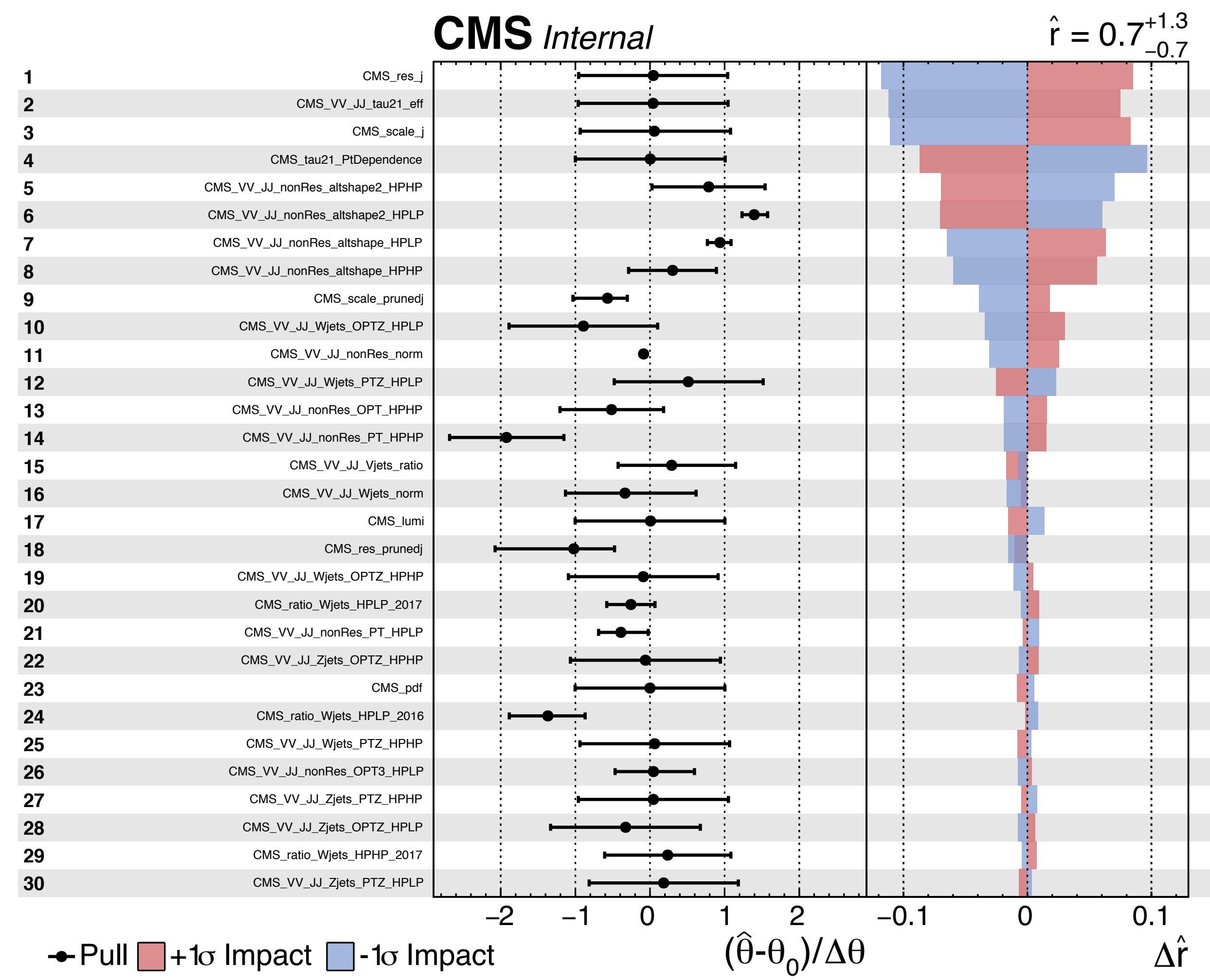




Impacts

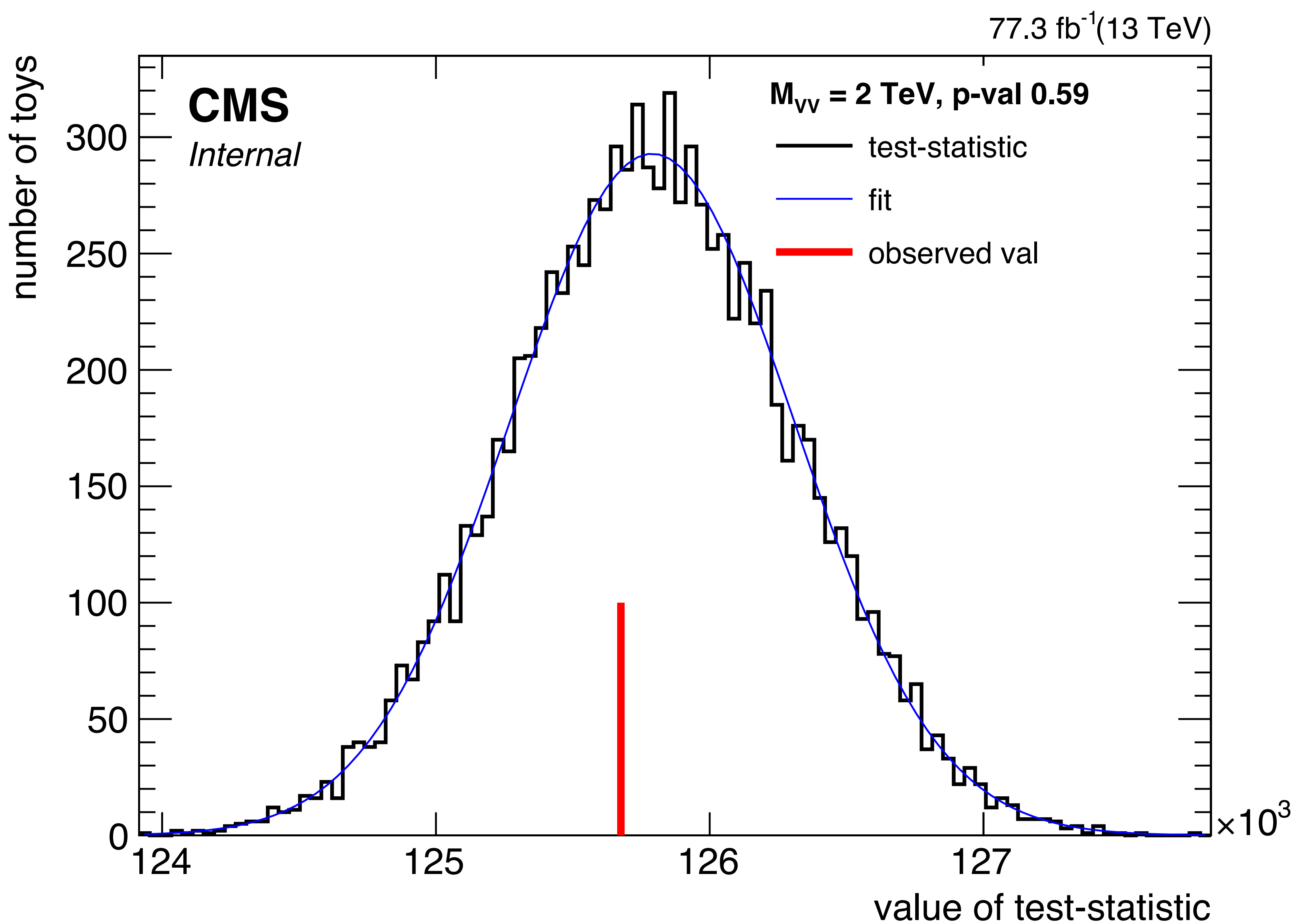


CMS Internal
$\hat{r} = 0.7^{+1.3}_{-0.7}$
31 CMS_ratio_Wjets_HPHP_2016
32 CMS_ratio_Zjets_HPHP_2016
33 CMS_VV_JJ_nonRes_OPT_HPLP
34 CMS_VV_JJ_nonRes_OPT3_HPHP
35 CMS_ratio_Zjets_HPLP_2017
36 CMS_ratio_Zjets_HPLP_2016
37 CMS_ratio_Zjets_HPHP_2017
−2 −1 0 1 2
−0.005 0 0.005
Pull +1σ Impact -1σ Impact
$(\hat{\theta}-\theta_0)/\Delta\theta$
$\Delta\hat{r}$



CMS Internal
$\hat{r} = 0.7^{+1.3}_{-0.7}$
1 CMS_res_j
2 CMS_VV_JJ_tau21_eff
3 CMS_scale_j
4 CMS_tau21_PtDependence
5 CMS_VV_JJ_nonRes_altshape2_HPHP
6 CMS_VV_JJ_nonRes_altshape2_HPLP
7 CMS_VV_JJ_nonRes_altshape_HPLP
8 CMS_VV_JJ_nonRes_altshape_HPHP
9 CMS_scale_prunedj
10 CMS_VV_JJ_Wjets_OPTZ_HPLP
11 CMS_VV_JJ_nonRes_norm
12 CMS_VV_JJ_Wjets_PTZ_HPLP
13 CMS_VV_JJ_nonRes_OPT_HPHP
14 CMS_VV_JJ_nonRes_PT_HPHP
15 CMS_VV_JJ_Vjets_ratio
16 CMS_VV_JJ_Wjets_norm
17 CMS_lumi
18 CMS_res_prunedj
19 CMS_VV_JJ_Wjets_OPTZ_HPHP
20 CMS_ratio_Wjets_HPLP_2017
21 CMS_VV_JJ_nonRes_PT_HPLP
22 CMS_VV_JJ_Zjets_OPTZ_HPHP
23 CMS_pdf
24 CMS_ratio_Wjets_HPLP_2016
25 CMS_VV_JJ_Wjets_PTZ_HPHP
26 CMS_VV_JJ_nonRes_OPT3_HPLP
27 CMS_VV_JJ_Zjets_PTZ_HPHP
28 CMS_VV_JJ_Zjets_OPTZ_HPLP
29 CMS_ratio_Wjets_HPHP_2017
30 CMS_VV_JJ_Zjets_PTZ_HPLP
−2 −1 0 1 2
−0.1 0 0.1
Pull +1σ Impact -1σ Impact
$(\hat{\theta}-\theta_0)/\Delta\theta$
$\Delta\hat{r}$

GOF



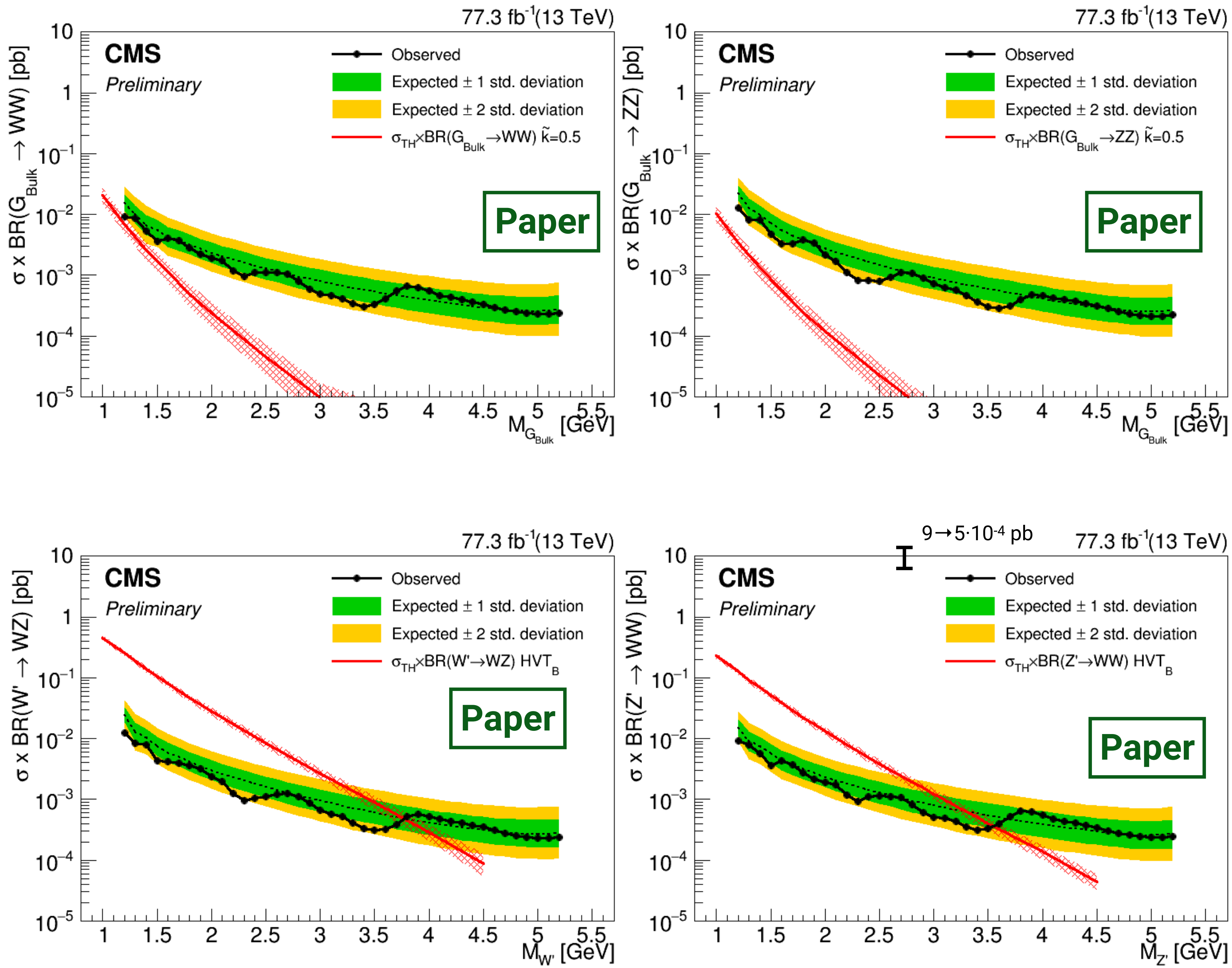
77.3 fb^{-1}(13 TeV)
CMS
Internal
M_{VV} = 2 TeV, p-val 0.59
test-statistic
fit
observed val
number of toys
300
250
200
150
100
50
0
124
125
126
127
$\times 10^3$
value of test-statistic

Limits

Limits set on 4 different signal hypothesis

Expect to exclude

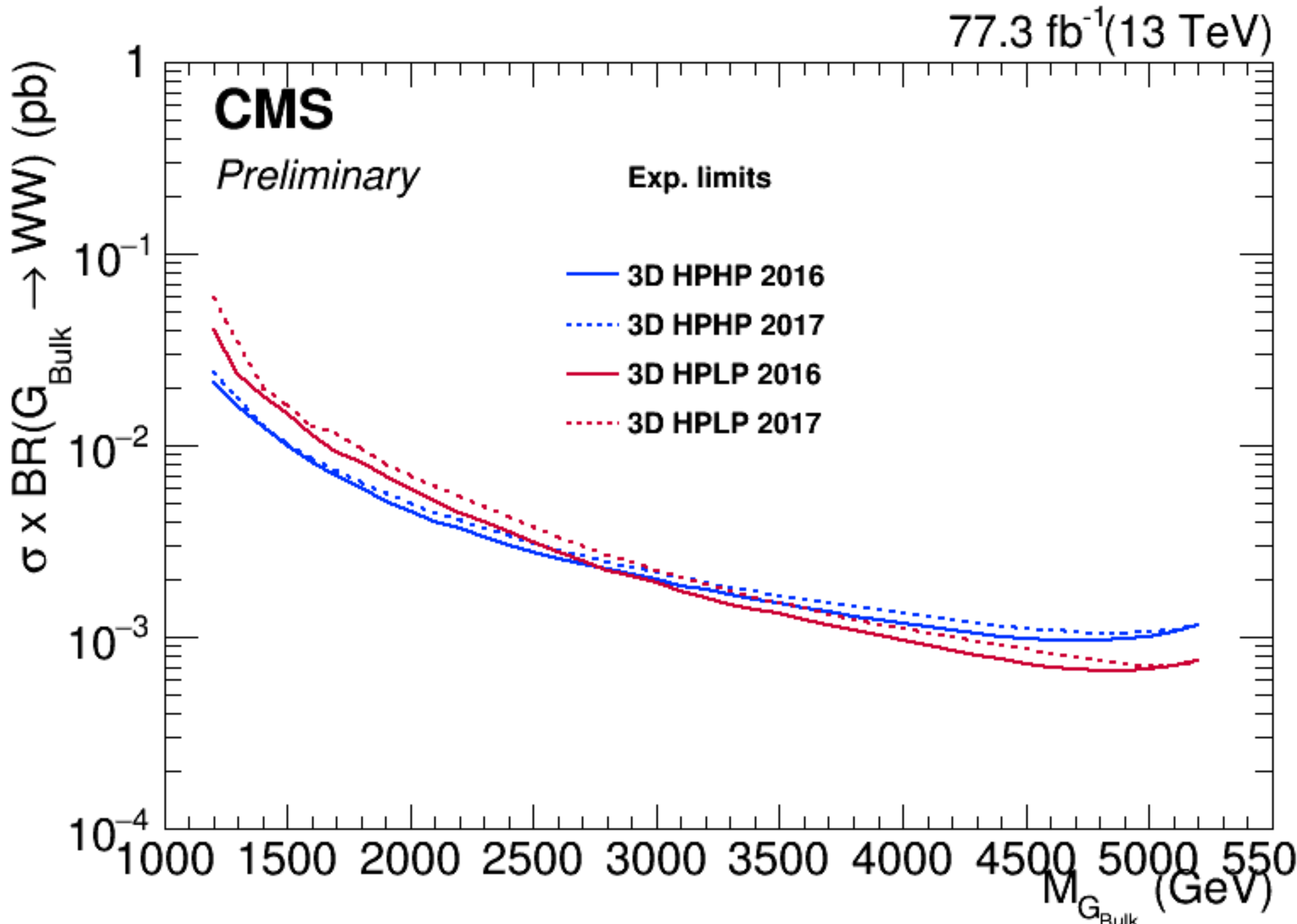
- W' → WZ: up to 3.8 TeV
- Z' → WW: up to 3.6 TeV
- Approaching G_{Bulk} ($\tilde{k}$=0.5)



Limits

HPHP versus HPLP

- HPHP dominate at low masses, while HPLP dominate at high



low masses ← HPHP HPLP → high masses

Loose DDT postfit

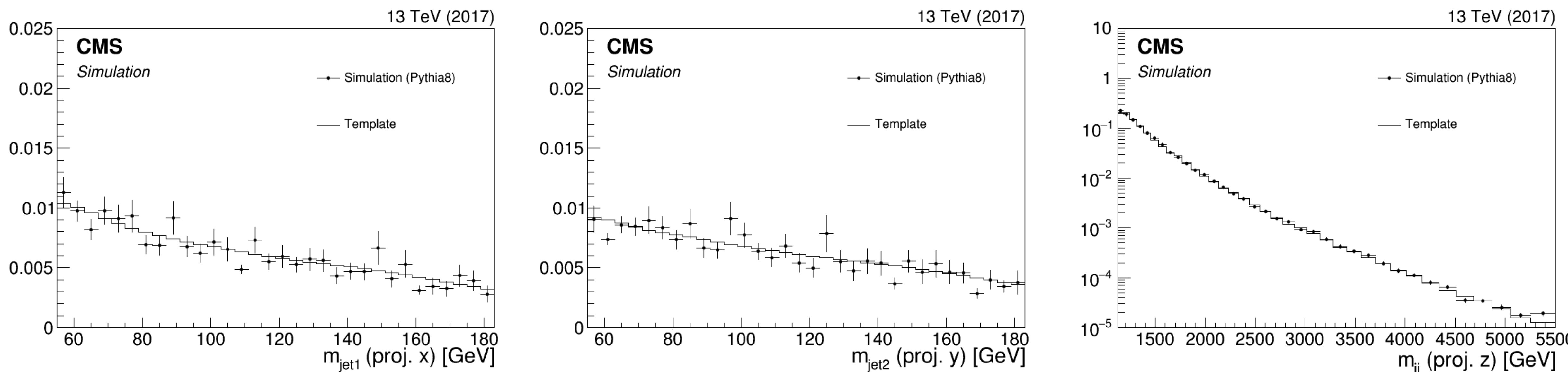


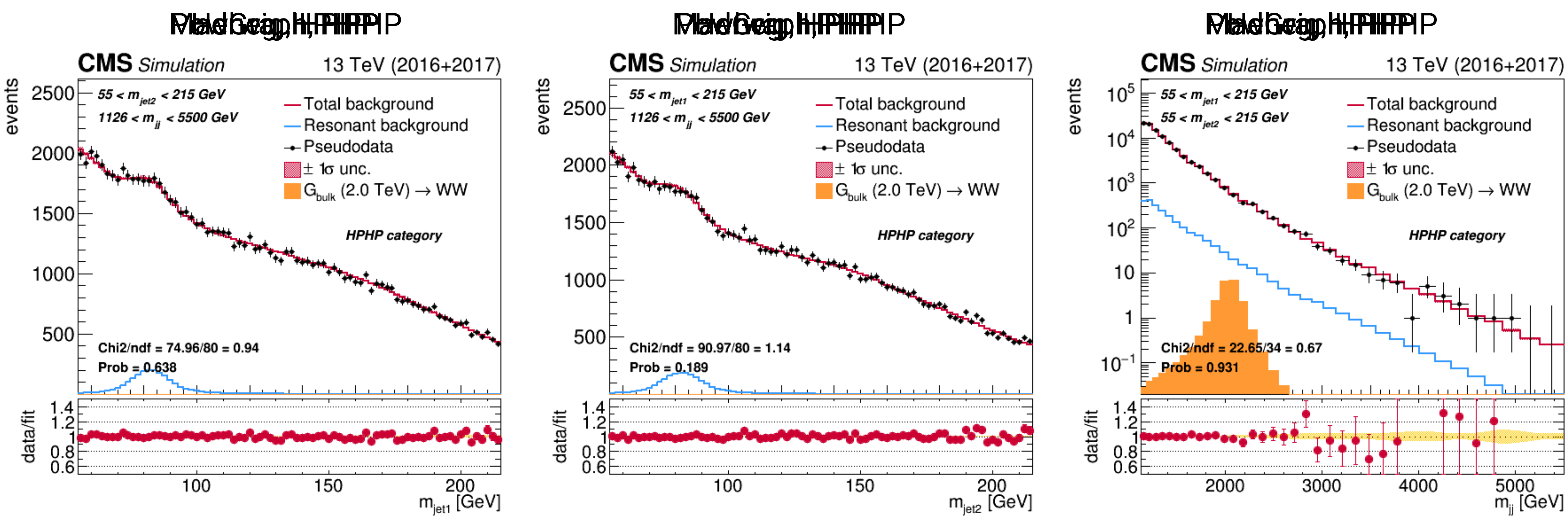
Figure 33: Comparison between QCD MC simulation (markers) and kernels derived from generator level quantities (lines) in the HPHP category, using a looser cut on τ_{21}^{DDT}.

To validate kernel transfer method, we check that we can fit a higher-statistics 387
HPHP region by loosening the τ^{DDT} cut to 0.49. Results in 12 times more background

Kernels: validation in MC

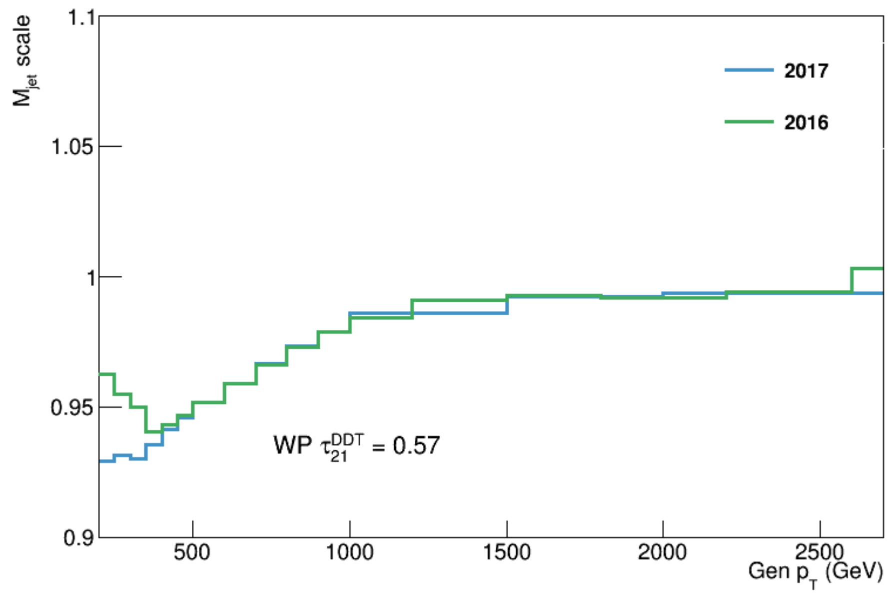
Several MC checks performed to validate kernels. Use kernel produced with nominal Pythia8 MC as starting point and fit pseudo data generated under

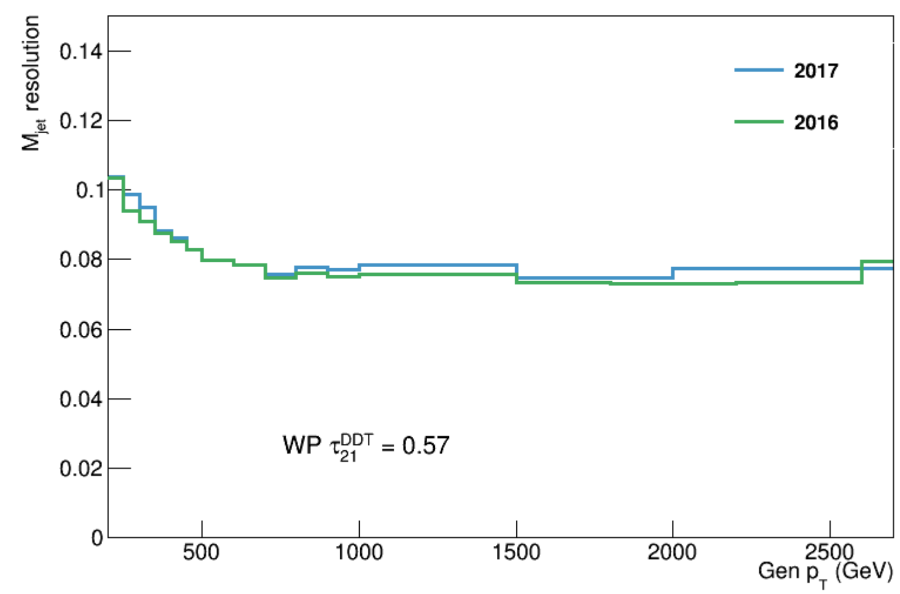
- Herwig - check kernels can account for variations in showering
- MadGraph - check kernels can account for variations in matrix elemenent
- Powheg NLO - check kernels can account for variations in perturbative predictions

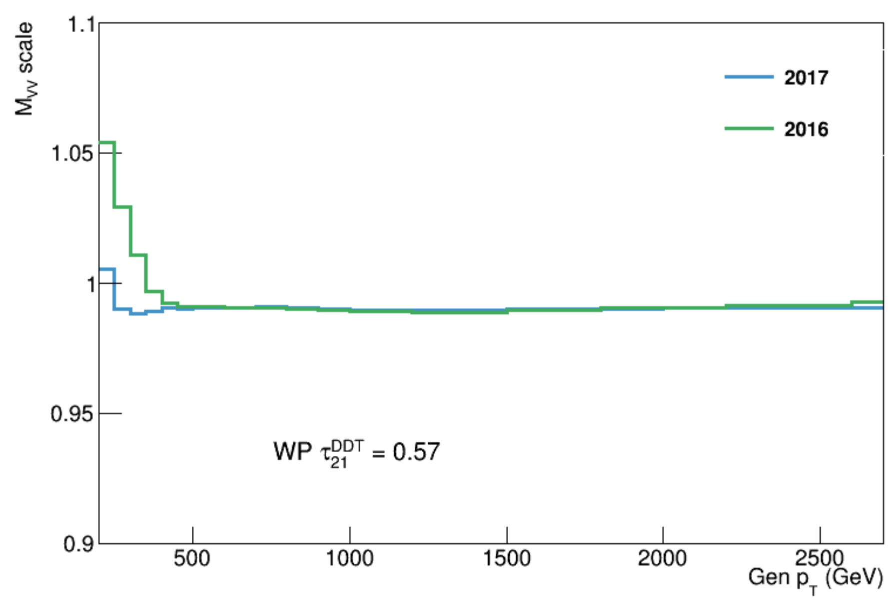


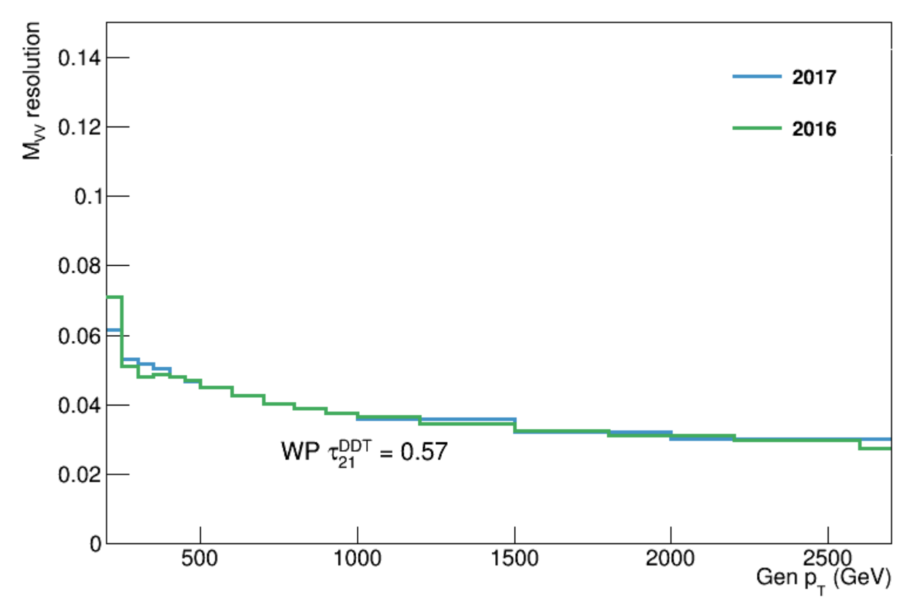
Detector resolution

gen-p_T bins = [200,250,300,350,400,450,500,600,700,800,900,1000,1500,2000,2700,3500,5000]

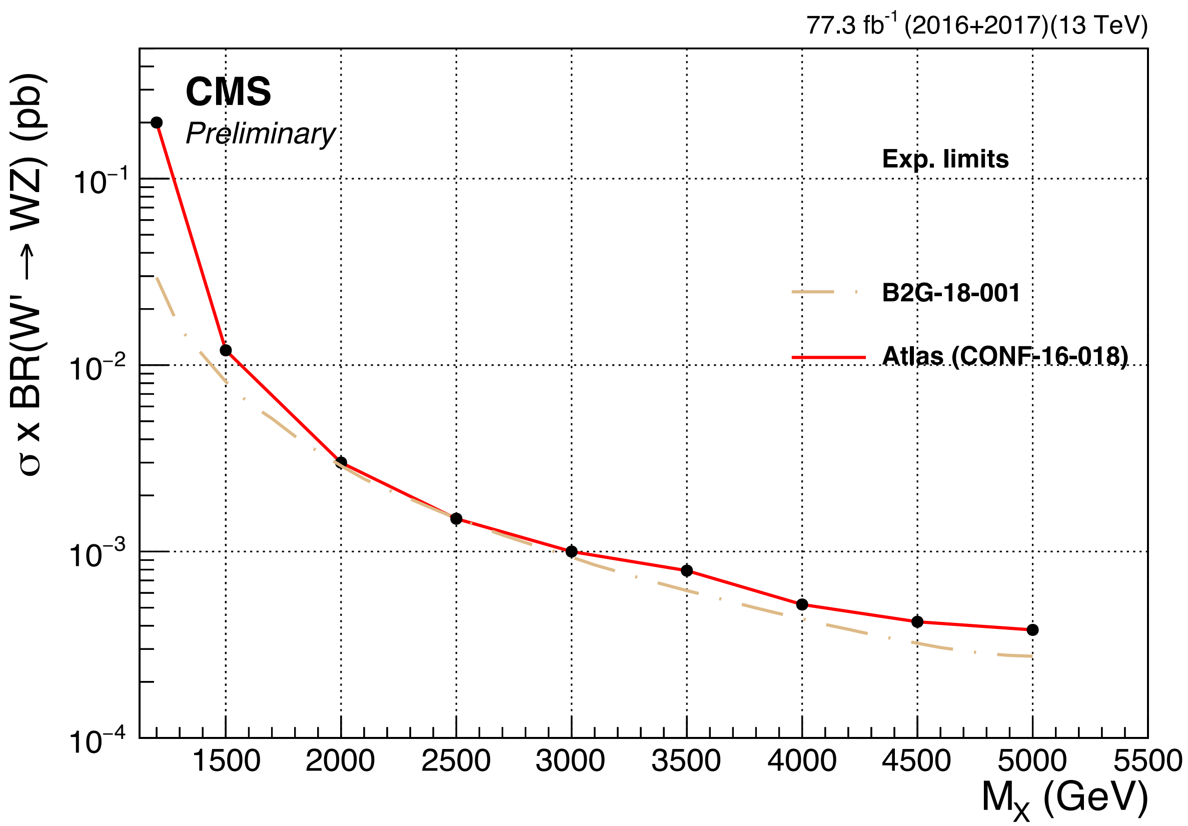


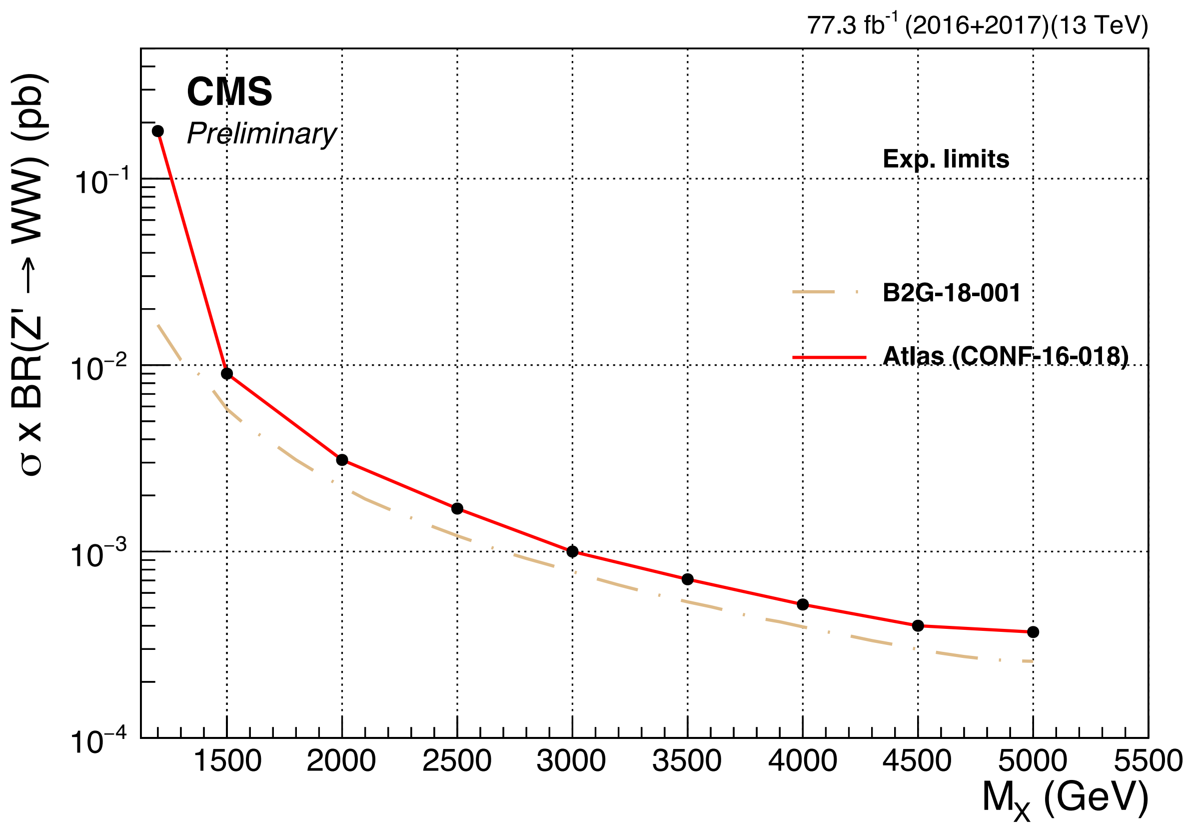




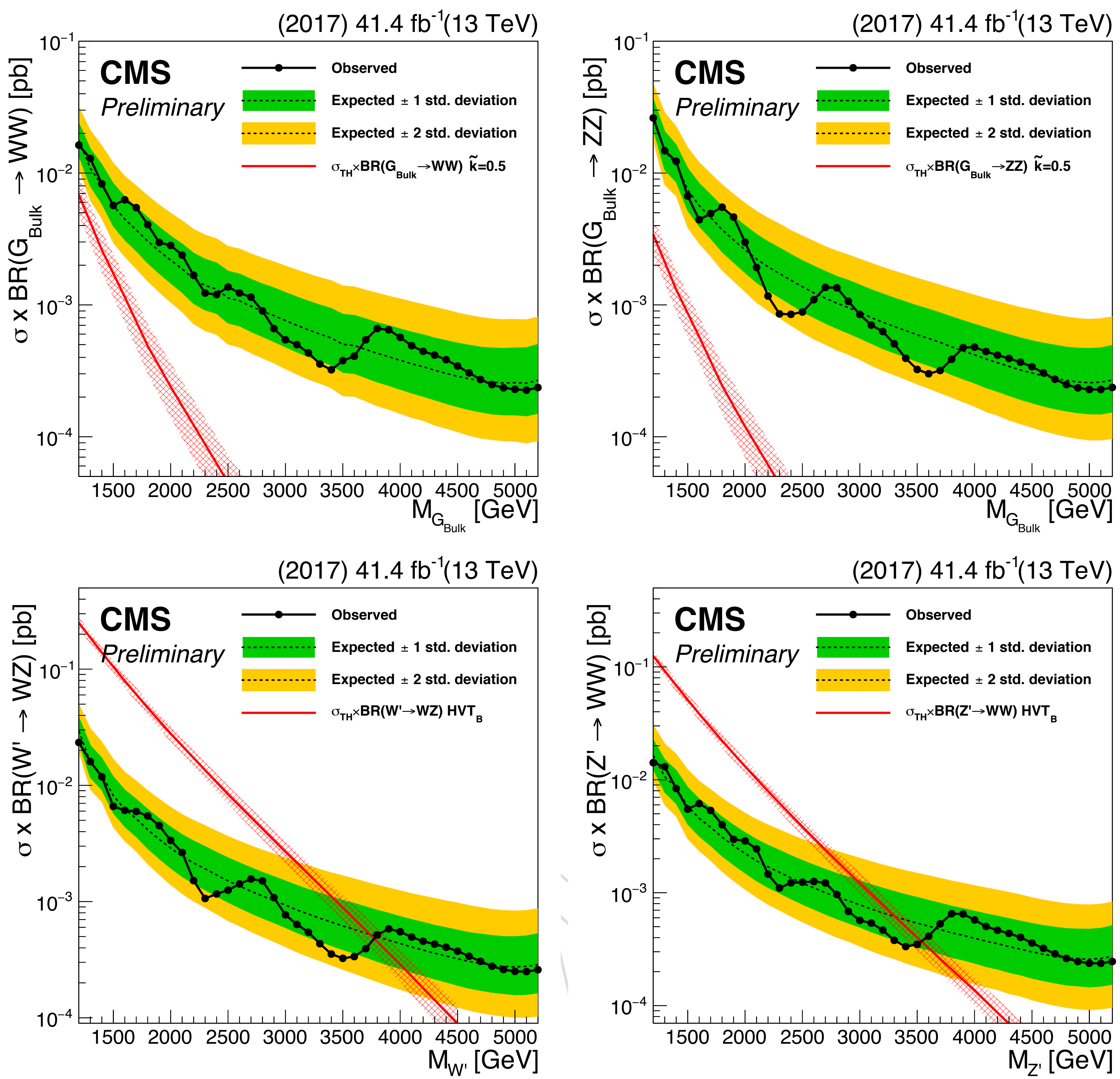


Comparisons

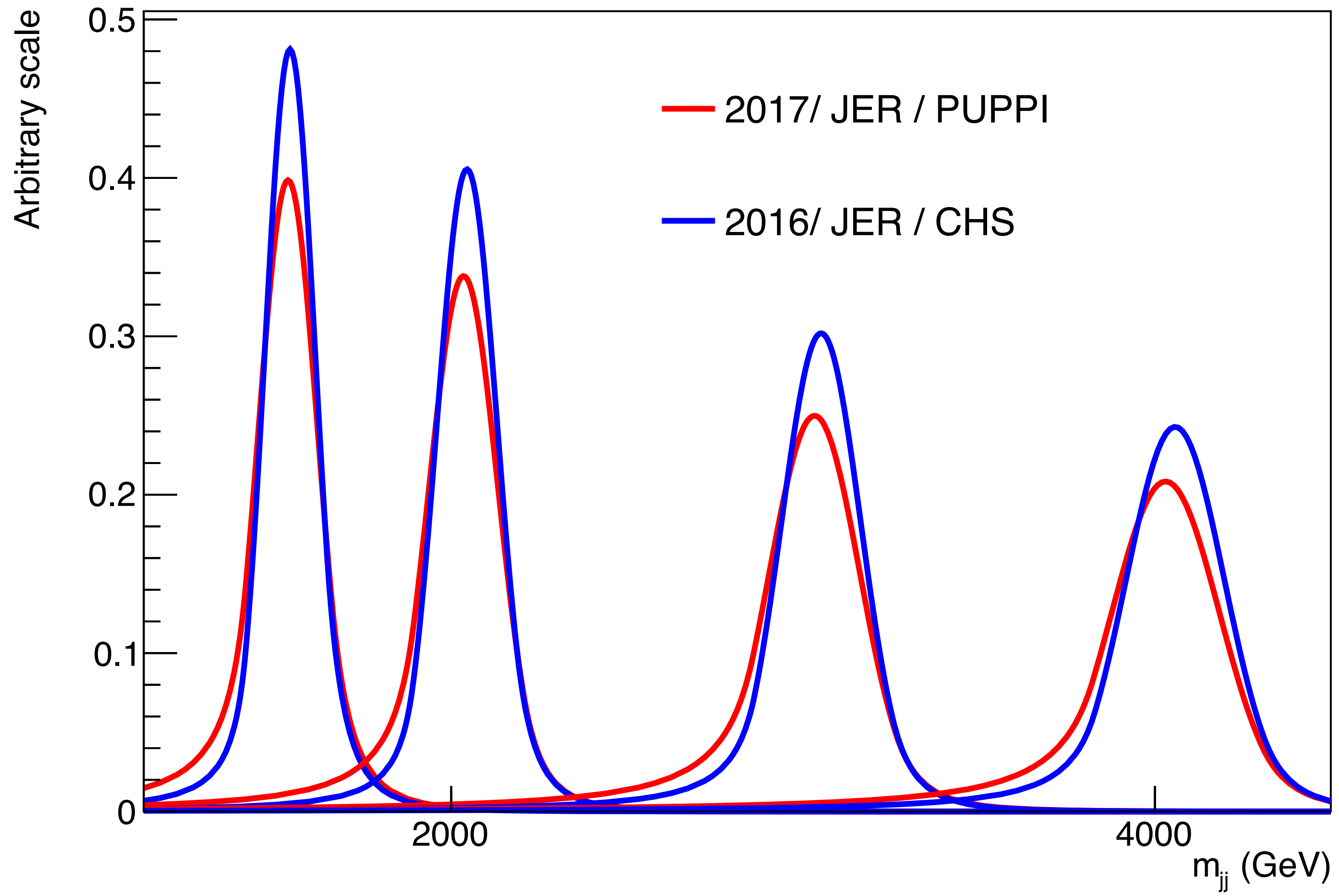




3D limits

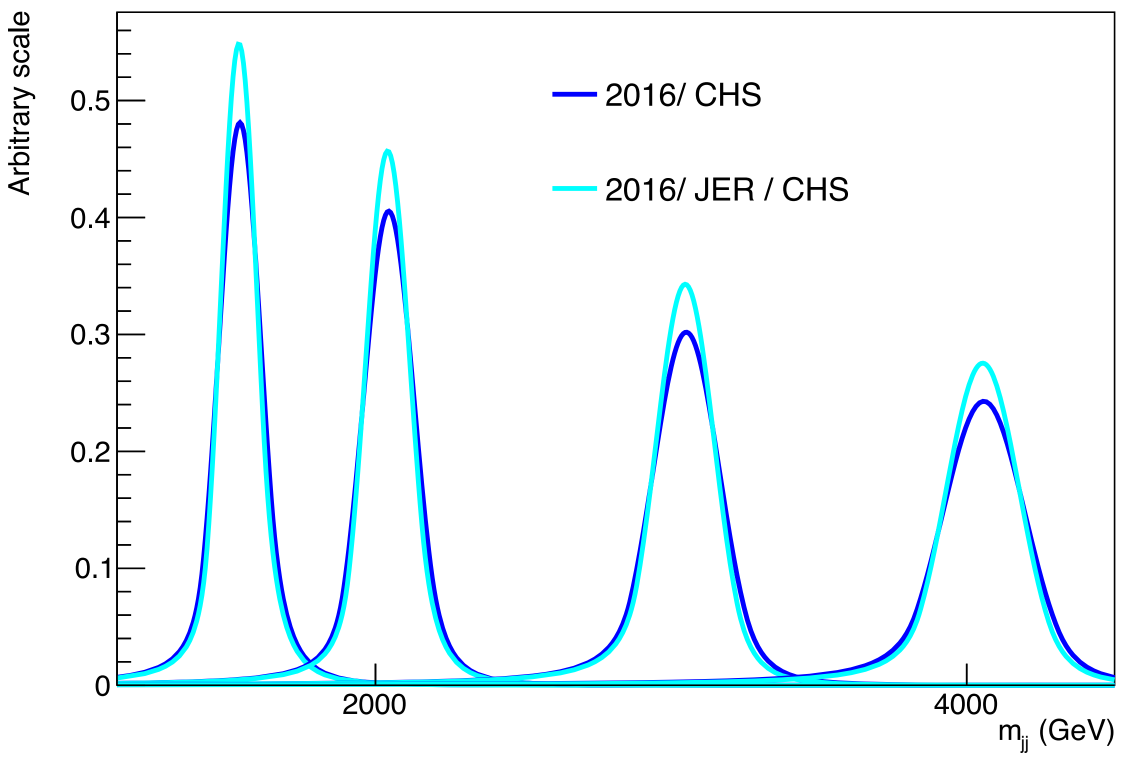

(2017) 41.4 fb-1(13 TeV)
CMS
Preliminary
Observed
Expected ± 1 std. deviation
Expected ± 2 std. deviation
σTH×BR(GBulk→WW) k̃=0.5
σ x BR(GBulk → WW) [pb]
MGBulk [GeV]
σTH×BR(GBulk→ZZ) k̃=0.5
σ x BR(GBulk → ZZ) [pb]
σTH×BR(W'→WZ) HVTB
σ x BR(W' → WZ) [pb]
MW' [GeV]
σTH×BR(Z'→WW) HVTB
σ x BR(Z' → WW) [pb]
MZ' [GeV]
1500 2000 2500 3000 3500 4000 4500 5000

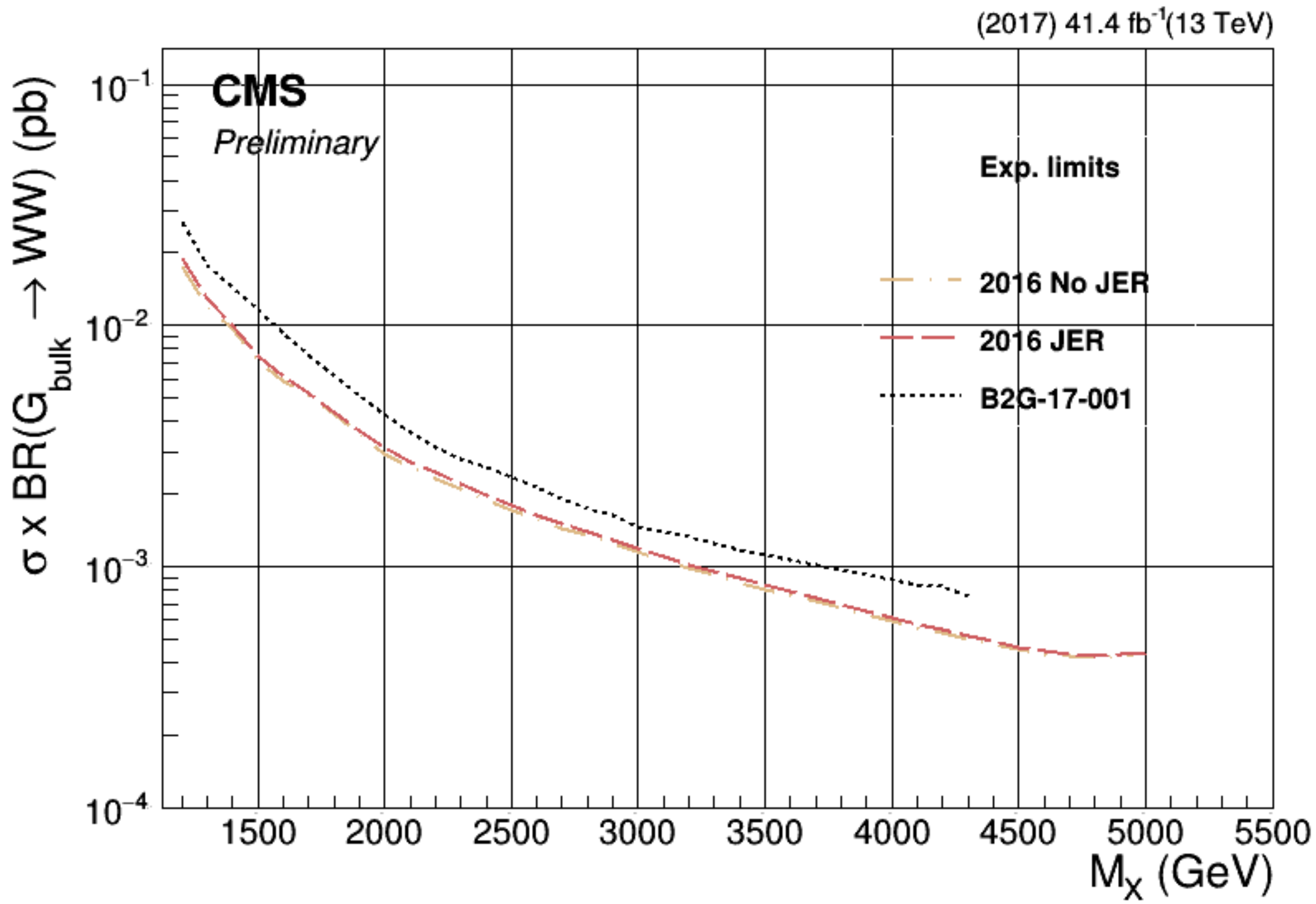
3D: 16 vs 17



Arbitrary scale
0.5
0.4
0.3
0.2
0.1
0
2000
4000
m_{jj} (GeV)
2017/ JER / PUPPI
2016/ JER / CHS

With/without JER

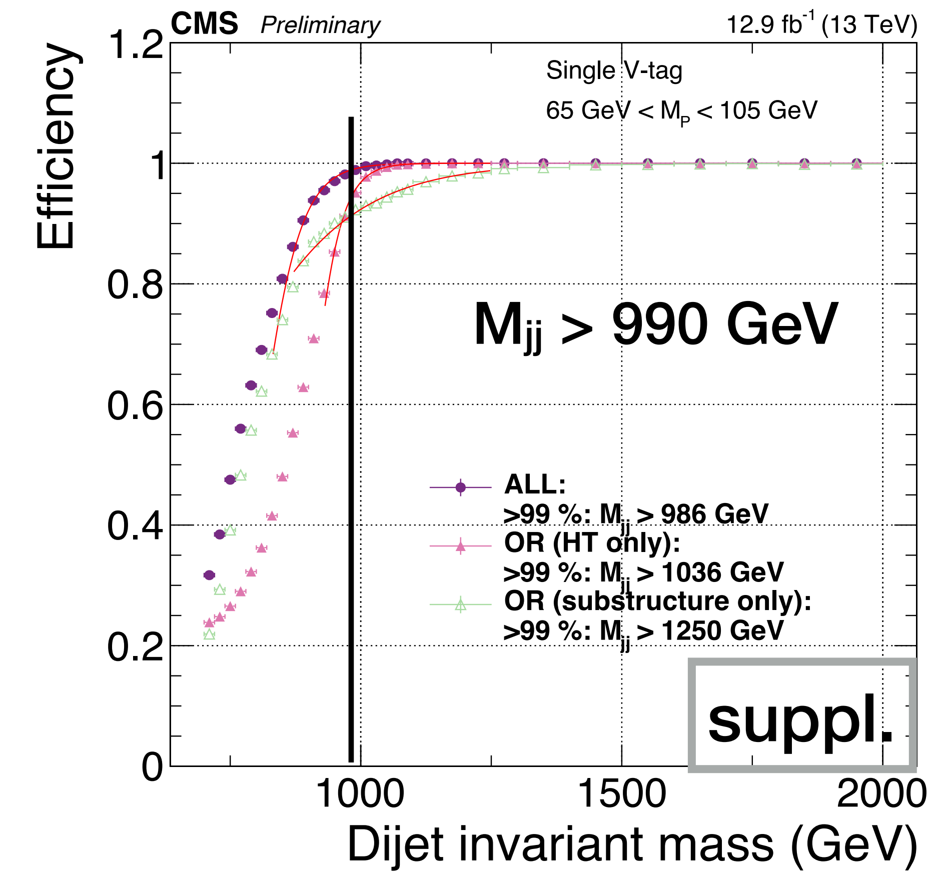


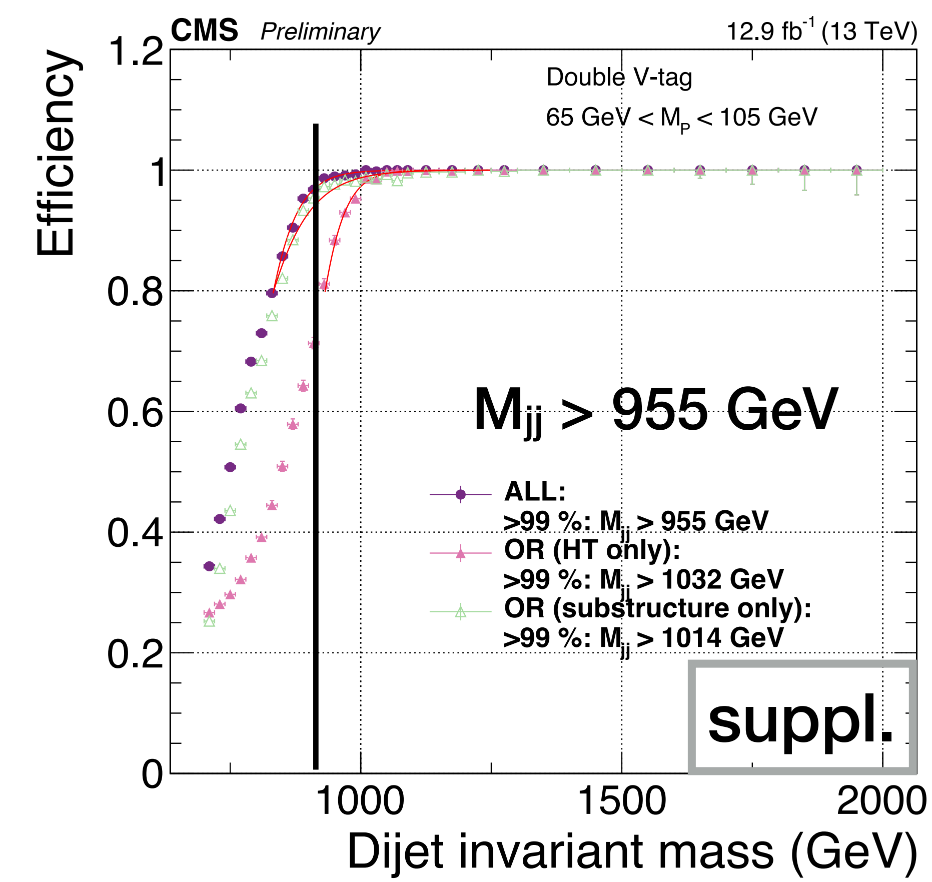


Search I

Trigger and dijet mass cut

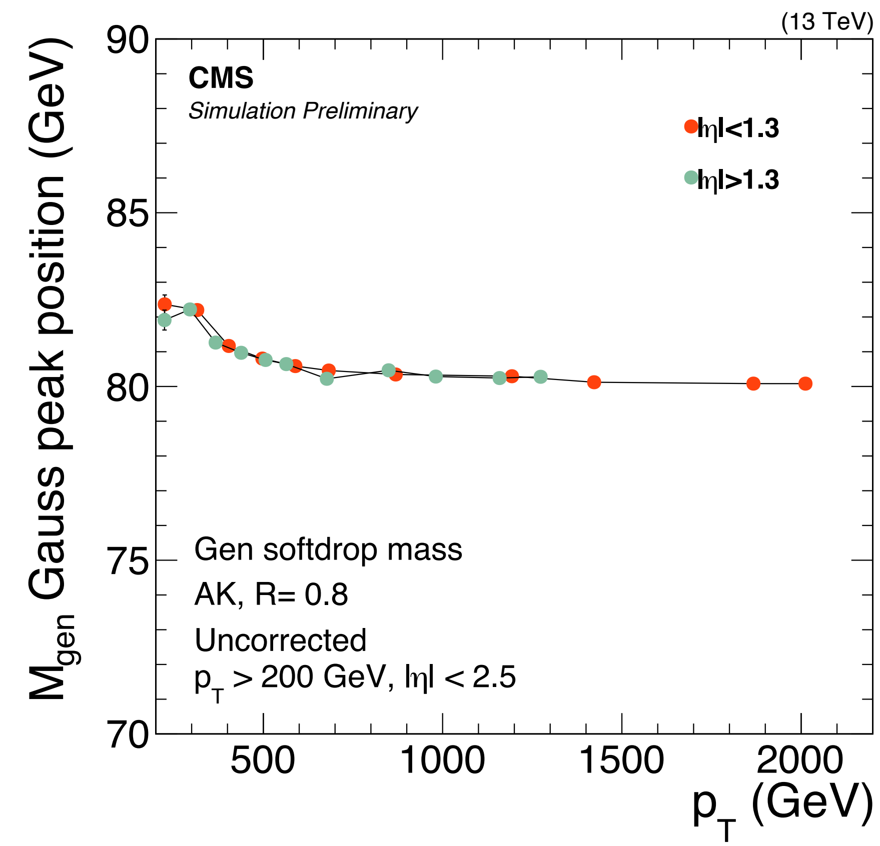
- **Background estimate depends on smoothly falling M_{jj} spectrum**
  - Start analysis where trigger efficiency > 99%
- **Combination of HT + substructure based triggers**
  - AK8PFJet360_TrimMass30
  - AK8PFHT700_TrimR0p1PT0p03Mass50
  - PFHT650_WideJetMJJ900DEtaJJ1p5
  - PFHT800
- **> 99% efficient for:**
  - Single-tag : M_{jj} > 986 GeV → start at 990 GeV
  - Double-tag: M_{jj} > 955 GeV → start at 955 GeV
- **For control plots, require Mjj > 1020 GeV (no jet mass cut applied, higher turn-on)**

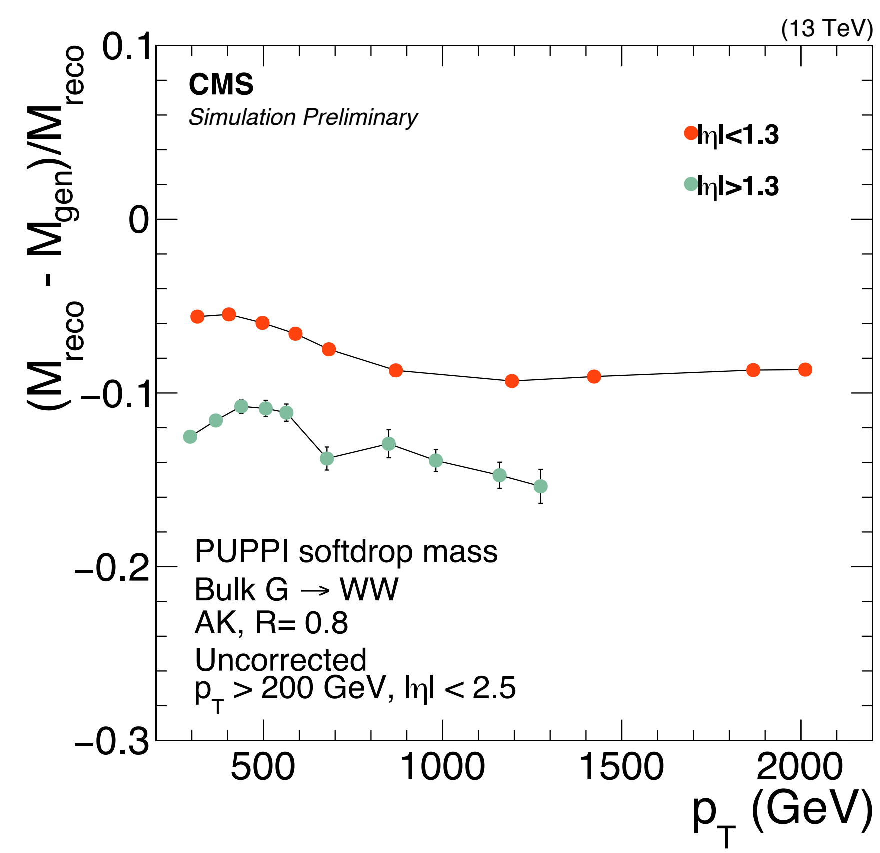




Mass correction for PUPPI softdrop mass

- **Generator level correction**
  - p_T-dependent shift in softdrop mass introduced at generator level
  - Correct for this jet mass shift (JMS) effect by fit to $M_{PDG=80.4\ GeV} / M_{gen}$
- **Reconstruction level correction**
  - After applying corrections above, $(M_{reco}-M_{gen})/M_{reco}$ jet mass shift is a 5-15% effect
  - Correct for residual effect by fit to M_{gen}/M_{reco}
- **Residual data/MC correction due to detector effects estimated in semi-leptonic $t\bar{t}$ (see slide 11)**
- **Potential difference due to simulation of tt topology accounted for as systematic uncertainty by comparing Pythia8+Powheg(NLO) with Pythia8+Madgraph(LO)**

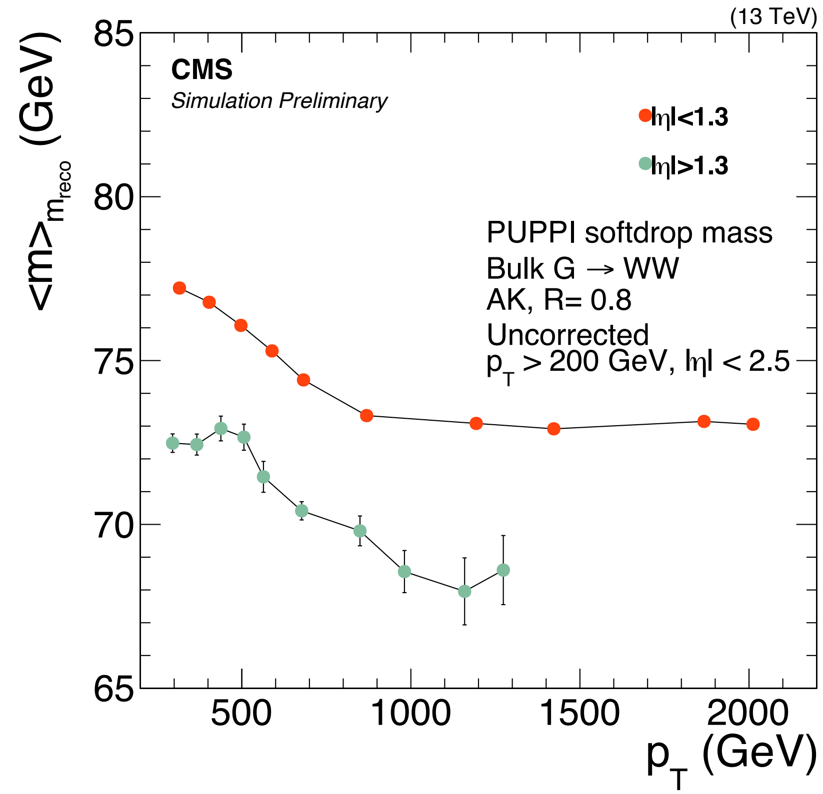




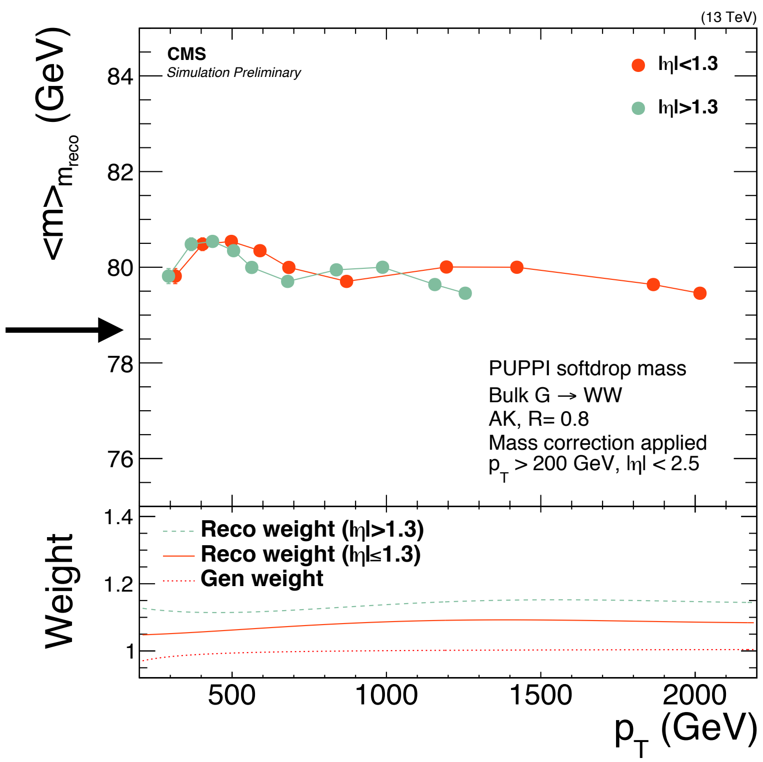
Closure test

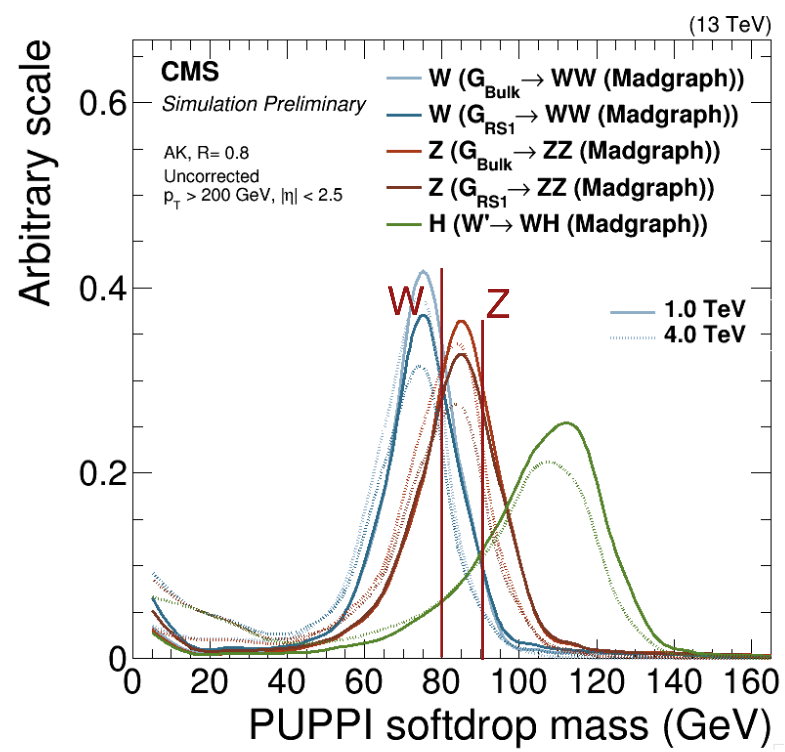
- **After mass corrections applied, mass stable as a function of p_T around 80 GeV**
- **Work for Z and H as well**
- **Additionally validated in semi-leptonic $t\bar{t}$**
  - jet mass scale and resolution close to unity (see next slide)

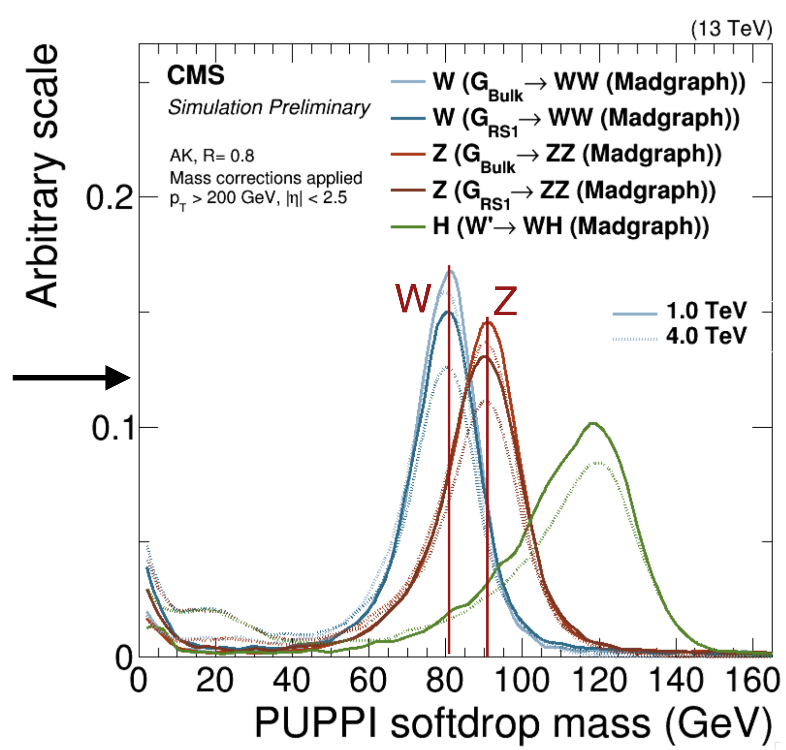
Uncorrected



New mass corrections

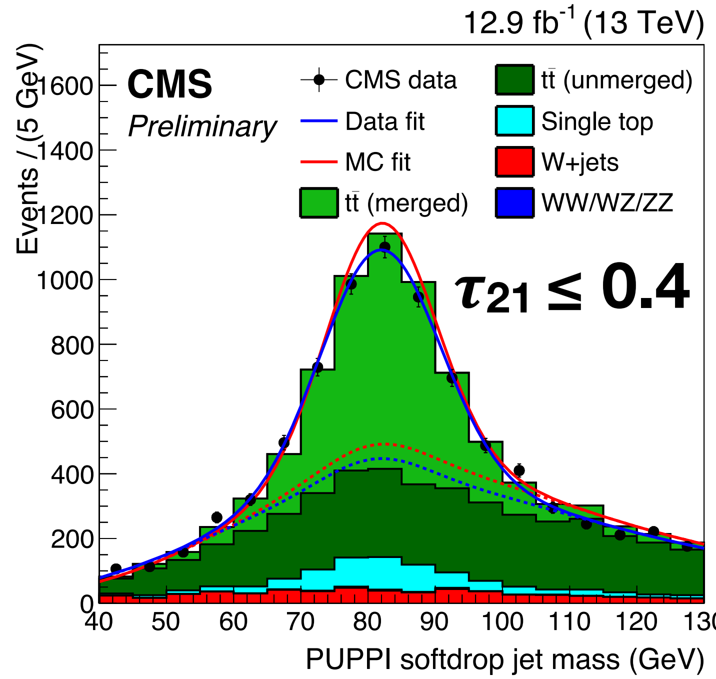


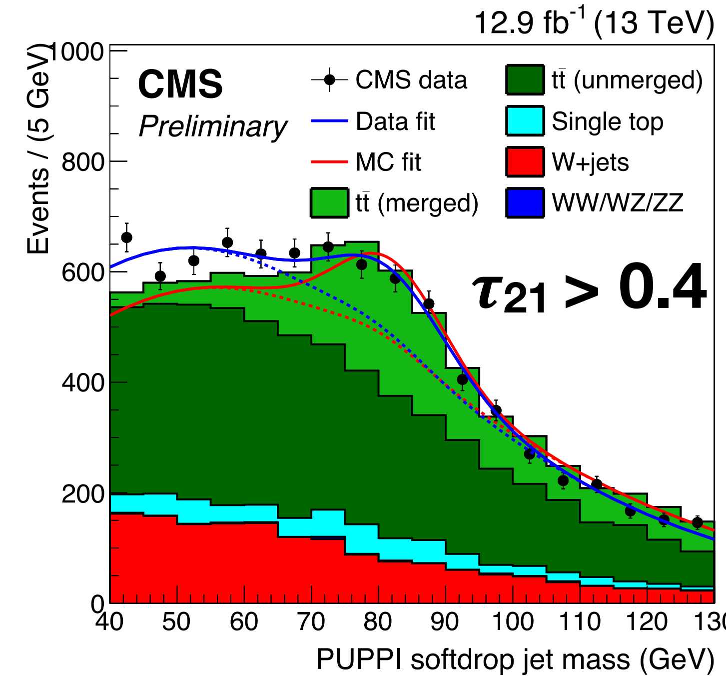




W-tagging scalefactor and jet mass scale

- **Estimated in merged W-jet enriched sample in semi-leptonic $t\bar{t}$ (p_T~200 GeV):**
  - Simultaneous fit in pass ($\tau_{21} \leq 0.4$) and fail ($\tau_{21} > 0.4$) category for data and MC
  - Extract W-tagging efficiency as integral of Gaussian fit component → data/MC efficiency yields SF
  - Jet mass scale/resolution from Gaussian mean and width
- **Jet mass resolution used to smear MC and additionally inserted as systematic uncertainties (scaling up/down within unc.)**
- **Scalefactor inserted as scale of signal yield and as systematic uncertainty**
- **Further documentation here**

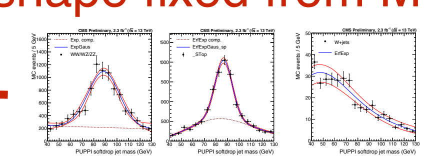




$\tau_{21} \leq 0.4$	m [GeV]	σ [GeV]
Data	81.9 ± 0.3 GeV	8.9 ± 0.4 GeV
Simulation	82.0 ± 0.2 GeV	8.3 ± 0.3 GeV
Data/simulation	0.999 ± 0.004 (stat) ± 0.0006 (sys)	1.08 ± 0.07 (stat) ± 0.08 (sys)

τ_{21} selection	Efficiency scale factor
$\tau_{21} < 0.4$	1.03 ± 0.03 (stat) ± 0.04 (sys) ± 0.06 (sys)
$0.4 < \tau_{21} < 0.75$	0.88 ± 0.12 (stat) ± 0.17 (sys) ± 0.12 (sys)

Yields+shape fixed from MC



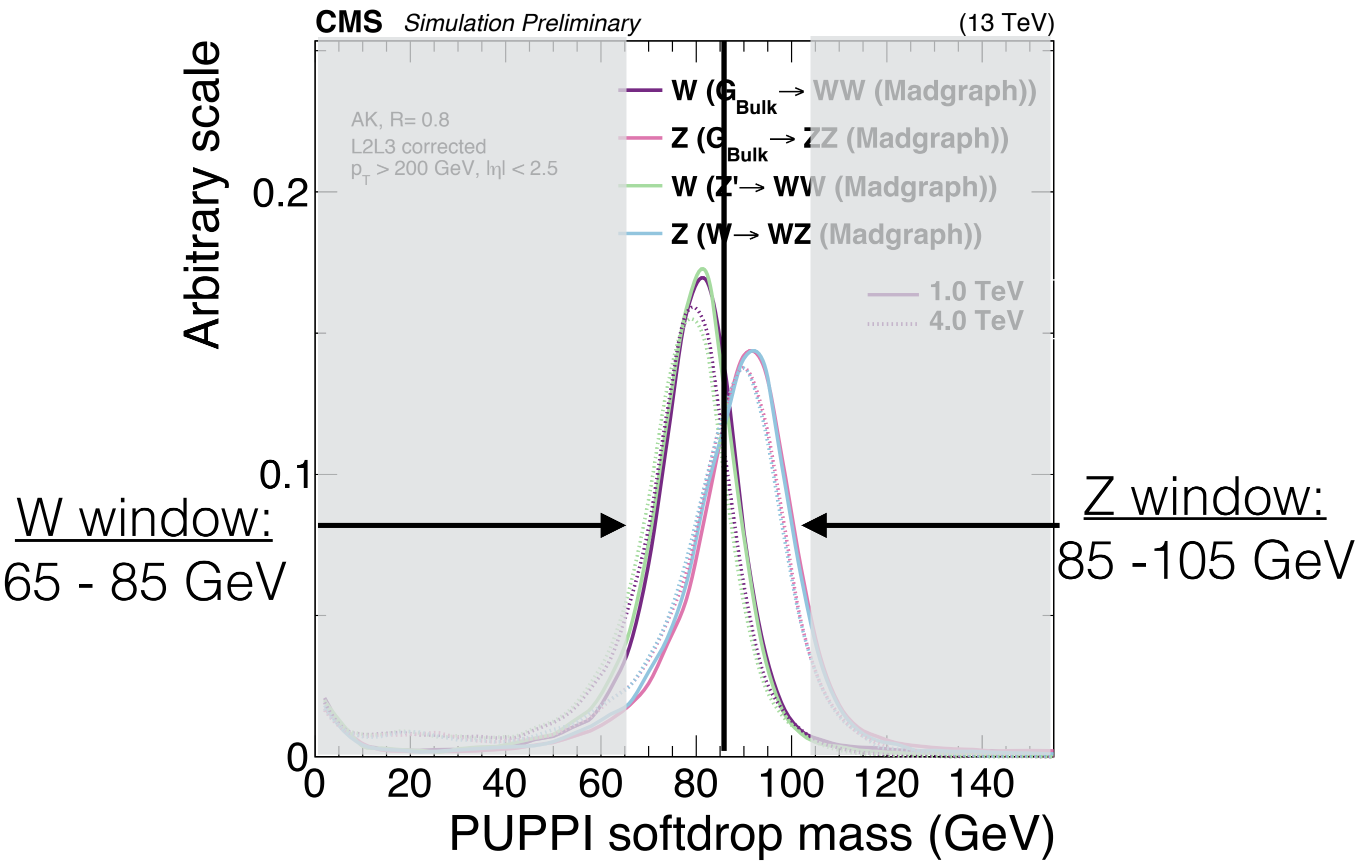
$$L_{\text{pass}} = \prod_{i}^{N_{\text{evt}}^{pass}} \left[N_{\text{W}} \cdot \epsilon_{HP} \cdot f_{\text{pass}}^{\text{sig}}(m_j) + N_2 \cdot f_{\text{pass}}^{\text{bkg}}(m_j) + N_{\text{pass}}^{\text{sTop}} \cdot f_{\text{pass}}^{\text{sTop}} + N_{\text{pass}}^{\text{VV}} \cdot f_{\text{pass}}^{\text{VV}} + N_{\text{pass}}^{\text{wjet}} \cdot f_{\text{pass}}^{\text{wjet}} \right]$$

$$L_{\text{fail}} = \prod_{i}^{N_{\text{evt}}^{fail}} \left[N_{\text{W}} \cdot (1 - \epsilon_{HP}) \cdot f_{\text{fail}}^{\text{sig}}(m_j) + N_3 \cdot f_{\text{fail}}^{\text{bkg}}(m_j) + N_{\text{fail}}^{\text{sTop}} \cdot f_{\text{fail}}^{\text{sTop}} + N_{\text{fail}}^{\text{VV}} \cdot f_{\text{fail}}^{\text{VV}} + N_{\text{fail}}^{\text{wjet}} \cdot f_{\text{fail}}^{\text{wjet}} \right]$$

$$SF = \frac{\epsilon_{HP}(\text{Data})}{\epsilon_{HP}(\text{MC})}$$

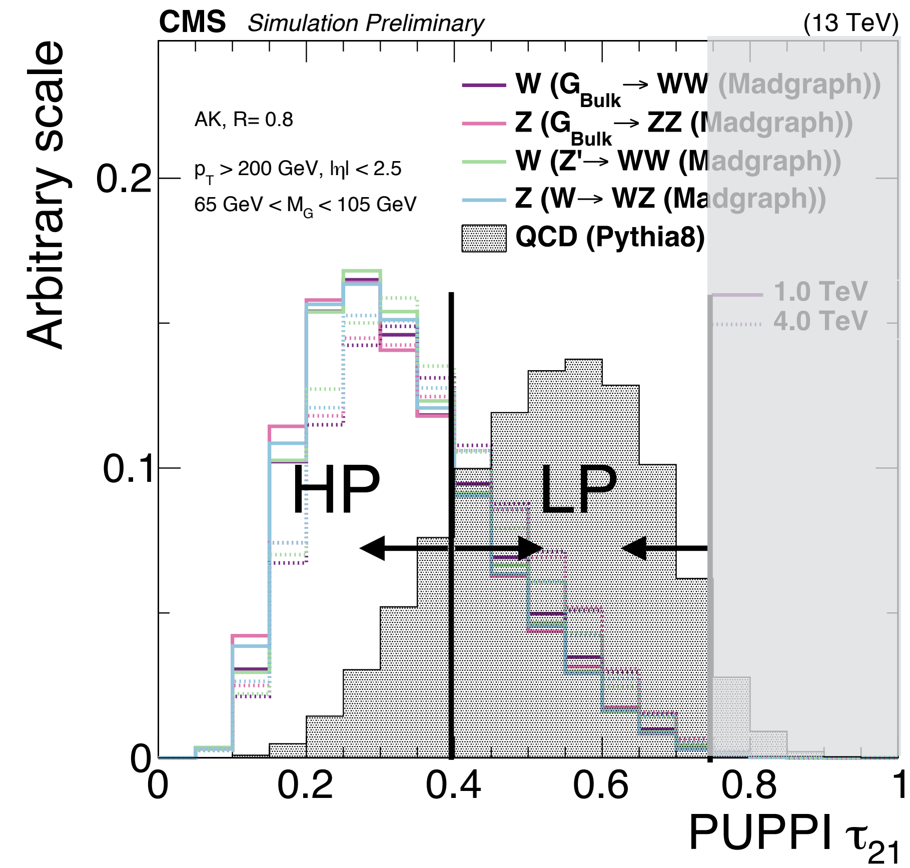
Mass and purity categorisation

- **Mass categories:**
  - To enhance sensitivity, split mass window



- **5 mass categories: WW/ZZ/WZ and qW/qZ:**
  - Combined into one VV/qV limit ~ slight gain in sensitivity
  - Expect more events in WZ channel for W'(→WZ) than G_{Bulk}(→WW/ZZ)

- **N-subjettiness categories:**



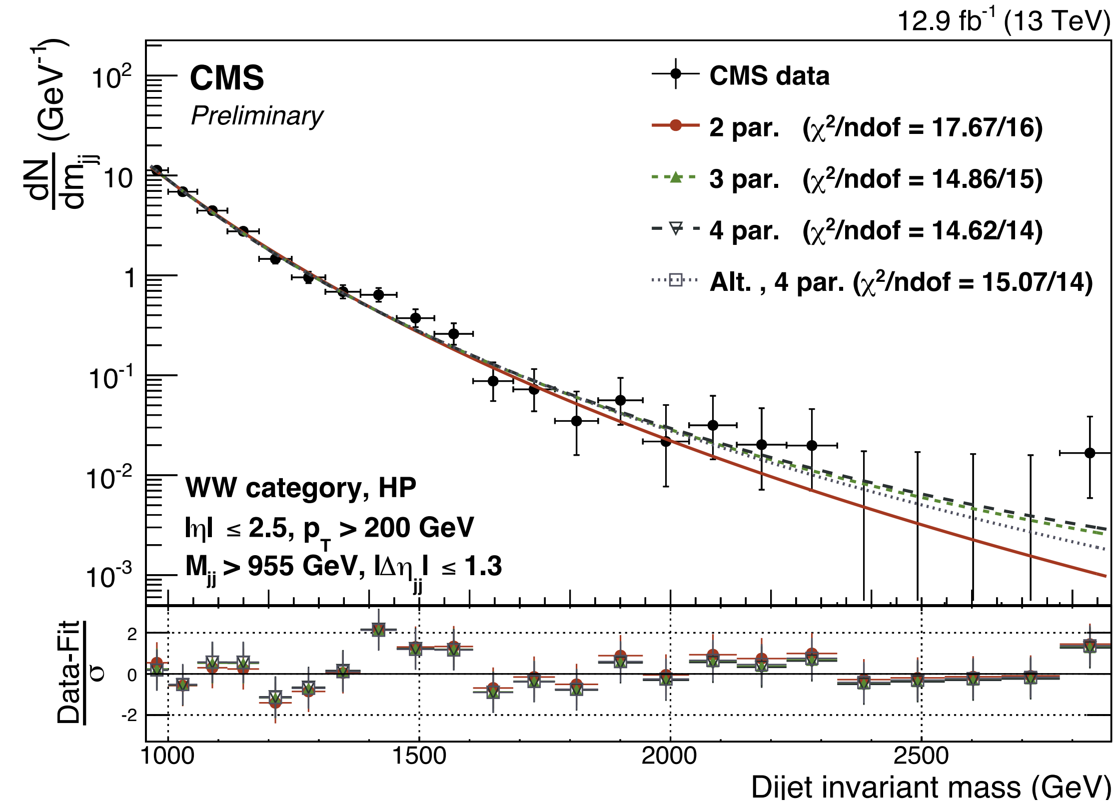
- **Two categories:**
  - High-purity: PUPPI $\tau_{21} \leq 0.4$ (best S/B)
  - Low-purity: $0.4 <$ PUPPI $\tau_{21} \leq 0.75$ (enhance sensitivity at high M_X)

- **All categories combined for final limits**

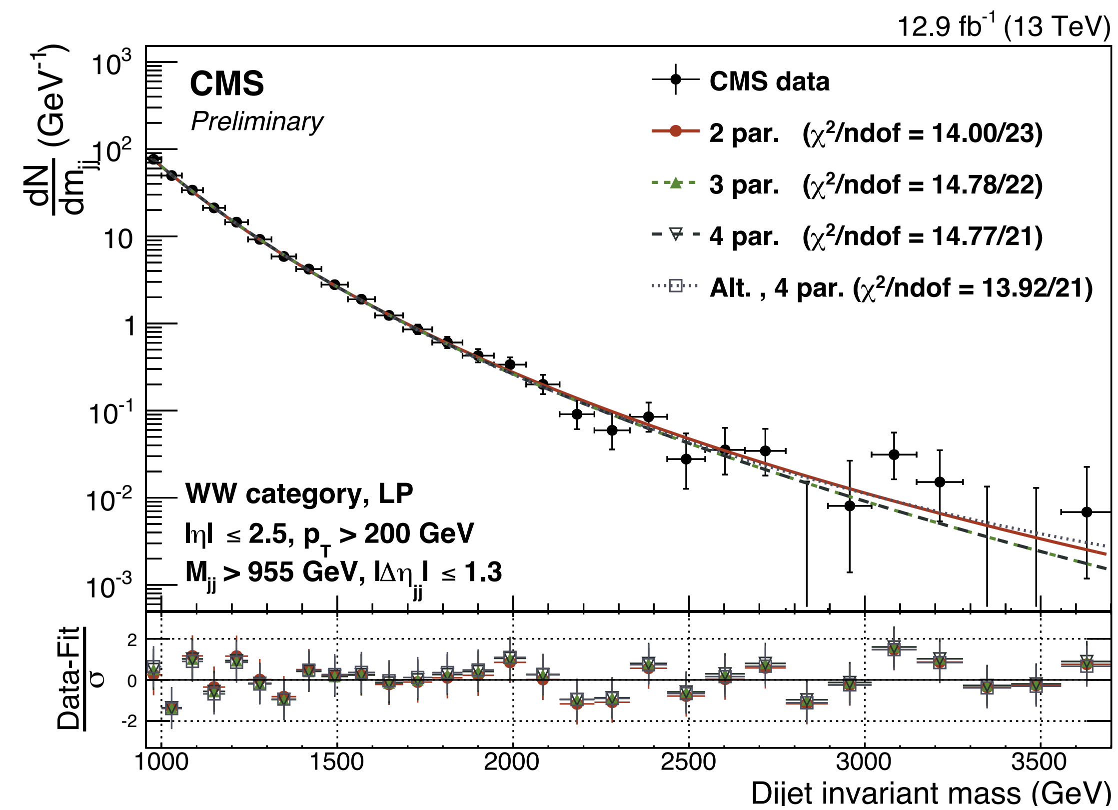
Double-tag: WW

- **Tested fit functions:**
  - 2 par: $\frac{dN}{dm} = \frac{P_0}{(m/\sqrt{s})^{P_2}}$
  - 3 par: $\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2}}$
  - 4 par: $\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2+P_3\times\log(m/\sqrt{s})}}$
  - Alt.: $\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s}+P_3(m/\sqrt{s})^2)^{P_1}}{(m/\sqrt{s})^{P_2}}$

WW category, HP			
Function	Residuals	χ^2	ndof
2 par	0.251	17.673	16
3 par	0.187	14.863	15
4 par	0.183	14.618	14
Fishers23	5.454	CL	0.033
Fishers34	0.391	CL	$\boxed{0.541}$



WW category, LP			
Function	Residuals	χ^2	ndof
2 par	2.974	13.997	23
3 par	3.082	14.775	22
4 par	3.080	14.768	21
Fishers23	-0.805	CL	$\boxed{1.000}$
Fishers34	0.015	CL	0.905



WW HP: 3 parameters
WW LP: 2 parameters

Single-tag: qZ

- **Higher order fit functions necessary for single-tag category (statistics higher, distributions more complex):**
  - 3 par: $\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2}}$
  - 4 par: $\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2+P_3\times\log(m/\sqrt{s})}}$
  - 5 par: $\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2+P_3\times\log(m/\sqrt{s})+P_4\times\log(m/\sqrt{s})^2}}$
  - Alt. 4: $\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s}+P_3(m/\sqrt{s})^2)^{P_1}}{(m/\sqrt{s})^{P_2}}$
  - Alt. 5: $\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s}+P_3(m/\sqrt{s})^2)^{P_1}}{(m/\sqrt{s})^{P_2+P_4\times\log(m/\sqrt{s})}}$

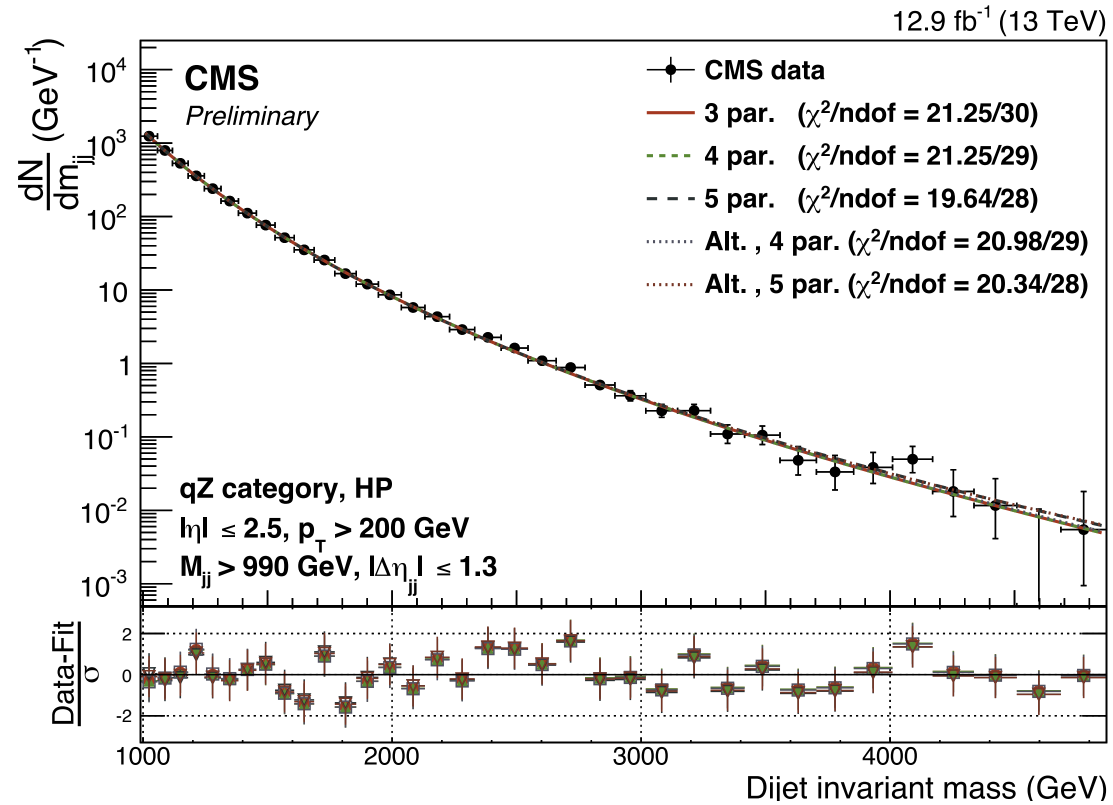
qZ category, HP

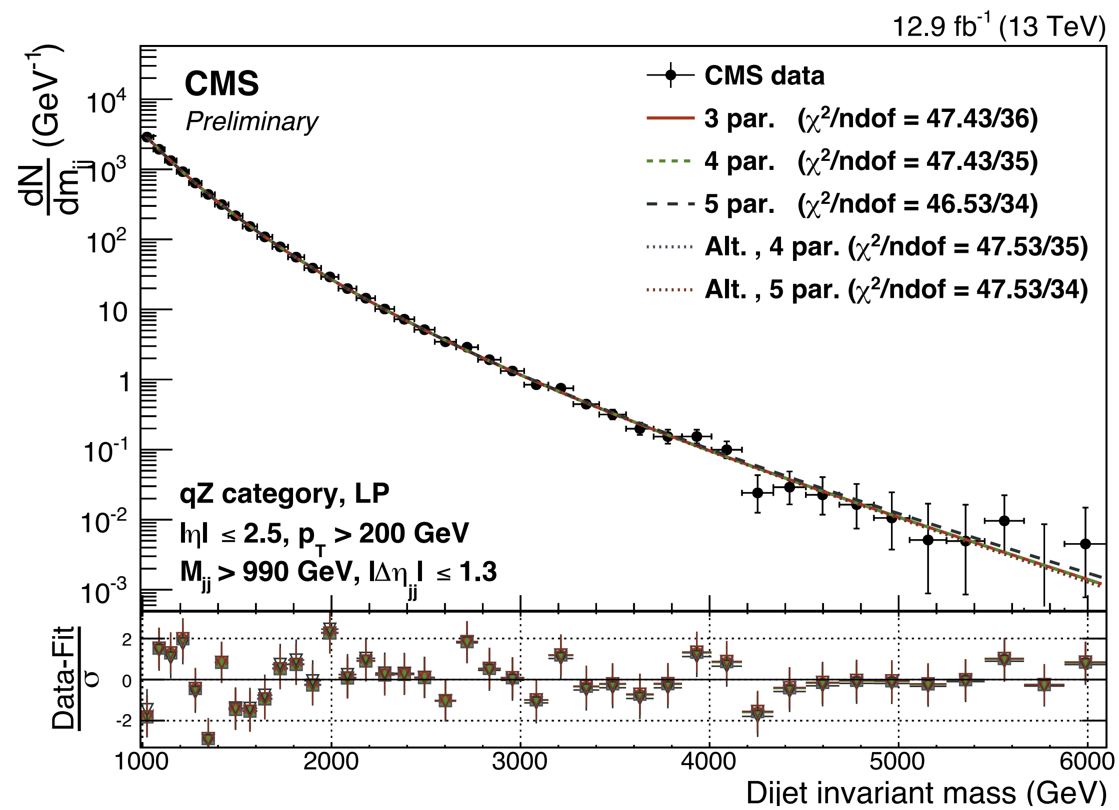
Function	Residuals	χ^2	ndof
3 par	12.963	21.252	30
4 par	12.961	21.252	29
5 par	9.256	19.644	28
Alt. 4 par	13.931	20.977	29
Alt. 5 par	9.739	20.344	28
Fishers34	0.004	CL	$\boxed{0.948}$
Fishers45	11.609	CL	0.002
FishersAlt4Alt5	12.484	CL	0.001

qZ category, LP

Function	Residuals	χ^2	ndof
3 par	369.554	47.426	36
4 par	369.554	47.426	35
5 par	298.358	46.525	34
Alt. 4 par	379.111	47.531	35
Alt. 5 par	379.120	47.531	34
Fishers34	0.000	CL	$\boxed{0.994}$
Fishers45	8.352	CL	0.007
FishersAlt4Alt5	-0.001	CL	0.000

→ **qZ HP: 3 parameters**
qZ LP: 3 parameters



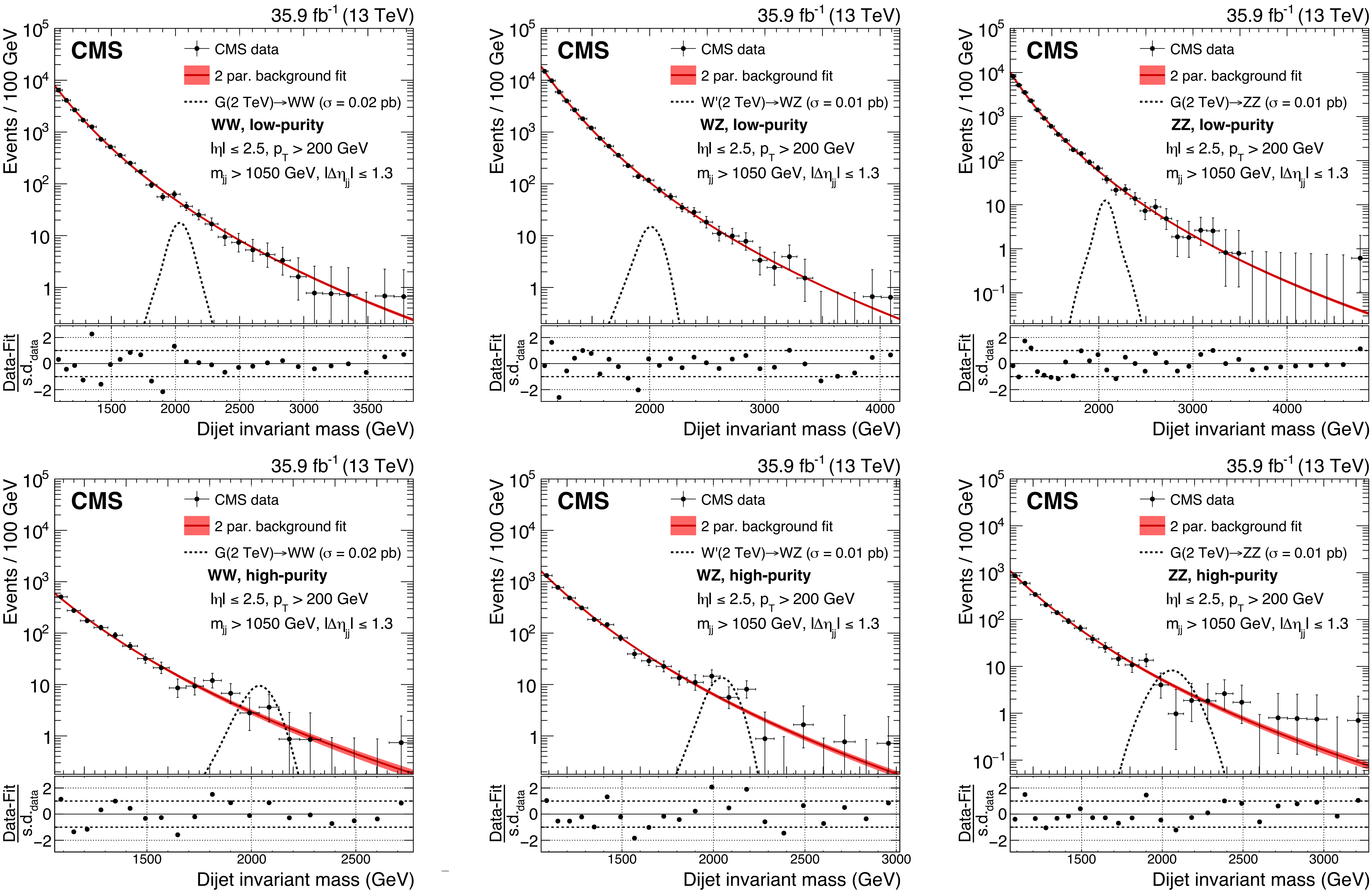


Systematic uncertainties

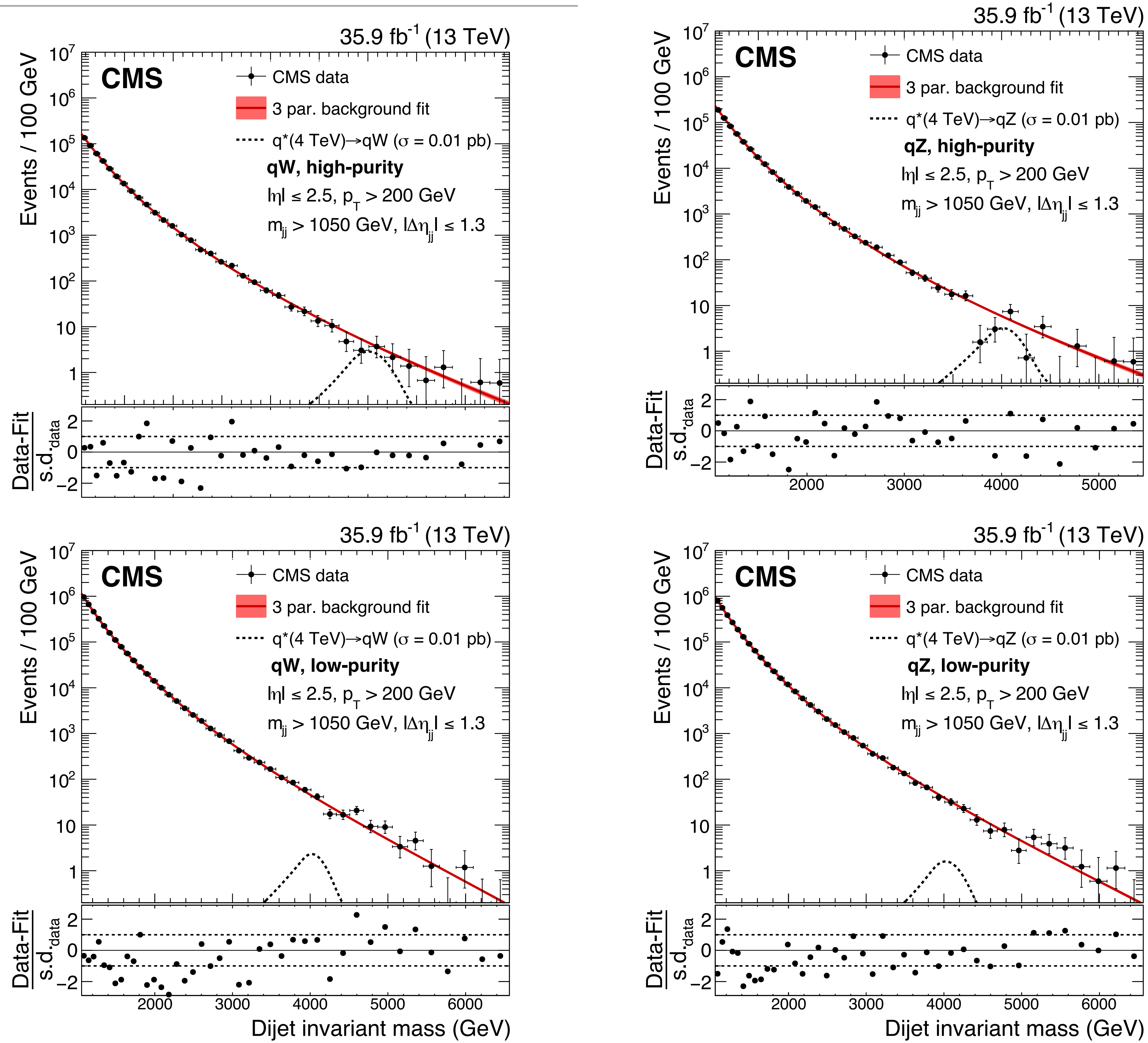
- **Main systematic uncertainties related to signal modelling**
  - Tagging efficiency of W-tagging scalefactor (include statistical uncertainty, τ_{21} cut efficiency measurement, simulation of the tt topology and extrapolation to high p_T)
  - Jet energy/mass scale and resolution (uncertainties due to jet mass scale and resolution are evaluated by scaling the PUPPI softdrop jet mass up and down within uncertainties listed on slide 12
- **Uncertainties related to jet mass and τ_{21} are correlated among categories (e.g events migrating out of HP move into LP)**

Source	Relevant quantity	HP+HP unc. (%)	HP+LP unc. (%)
Jet energy scale	Resonance shape	2	2
Jet energy resolution	Resonance shape	10	10
Jet energy scale	Signal yield	<0.1–4.4	
Jet energy resolution	Signal yield	<0.1–1.1	
Jet mass scale	Signal yield	0.02–1.5	
Jet mass resolution	Signal yield	1.3–6.8	
Pileup	Signal yield	2	
Integrated luminosity	Signal yield	6.2	
PDF and scales (W′ and Z′)	Signal yield	2–18	
PDF and scales (G_{bulk})	Signal yield	8–78	
Jet mass scale	Migration	<0.1–16.8	
Jet mass resolution	Migration	<0.1–17.8	
W-tagging τ_{21}	Migration	15.6	21.9
W-tagging p_T-dependence	Migration	7–14	5–11

Double-tag: Background-only fits

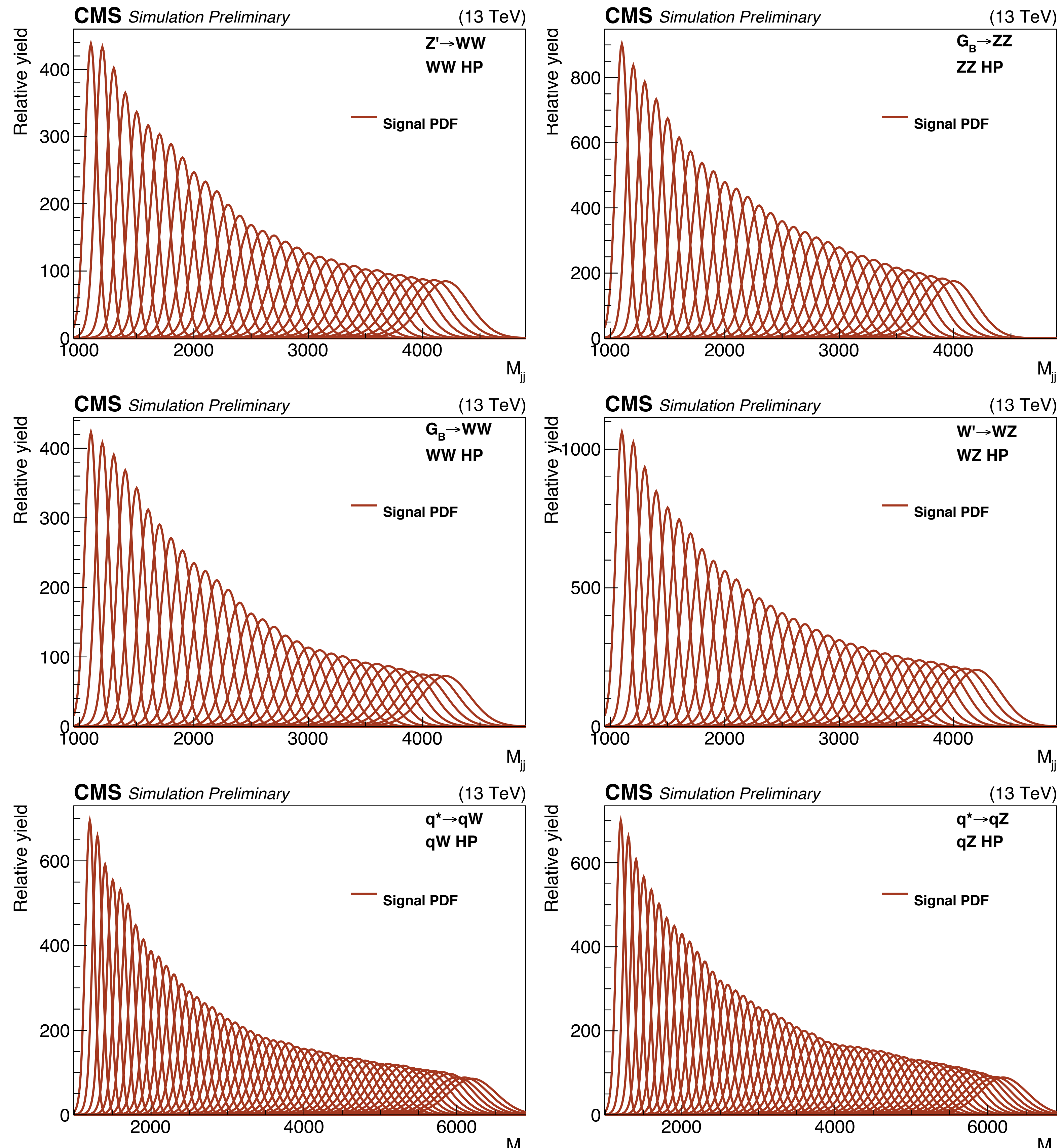


Single-tag: Background-only fits

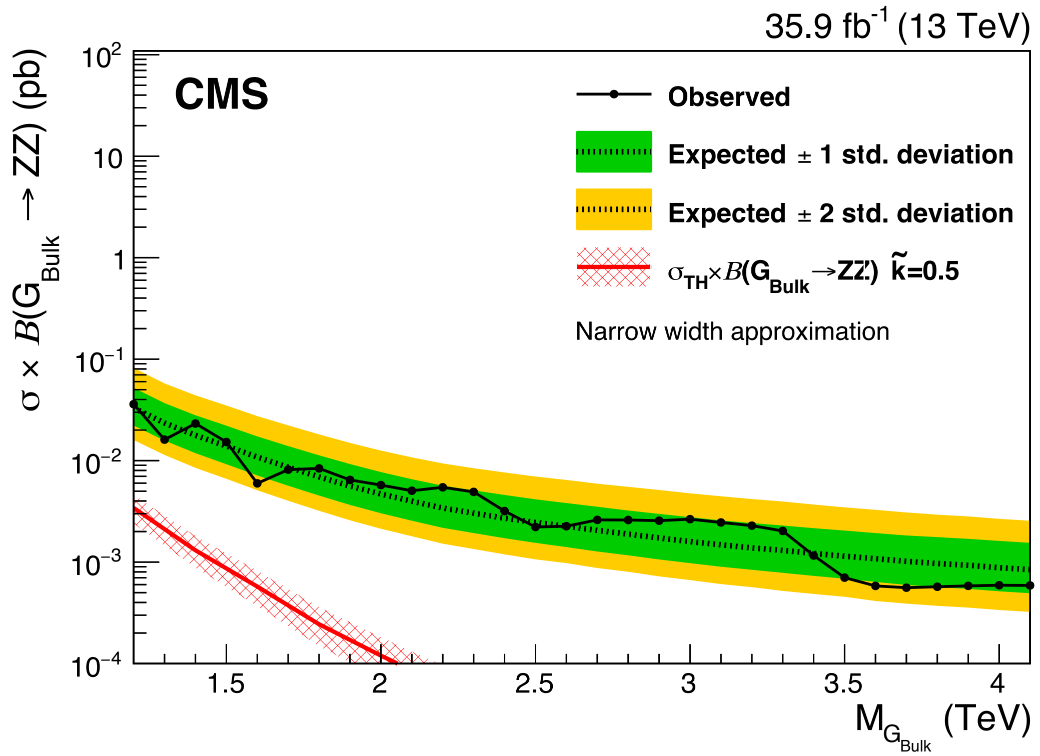


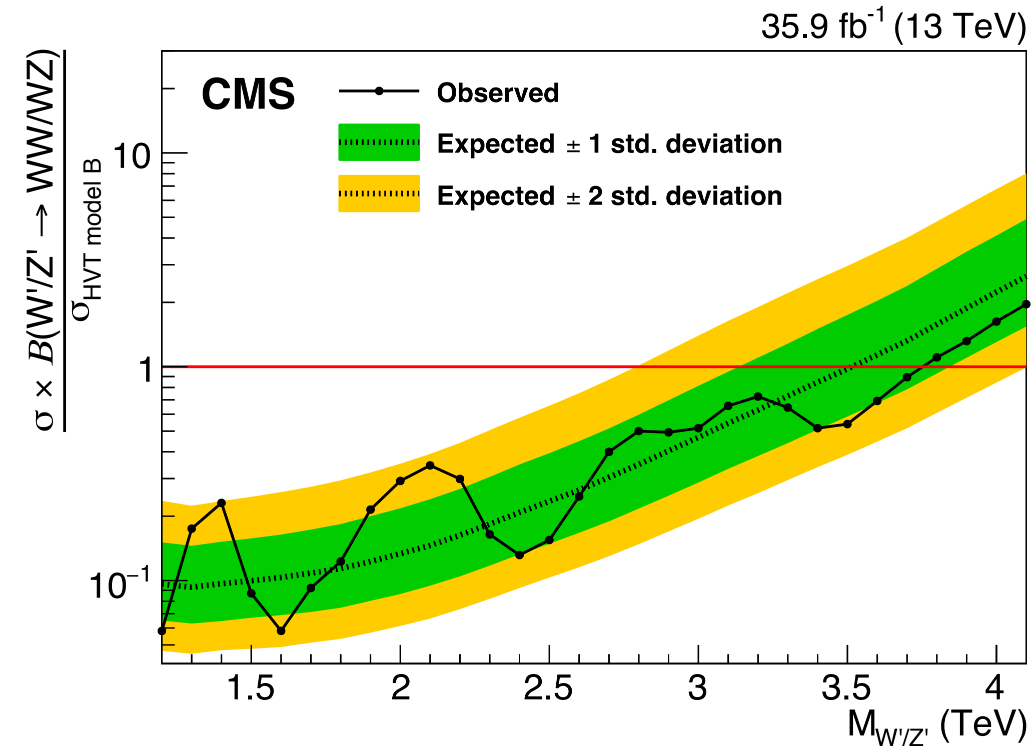
Signal interpolation

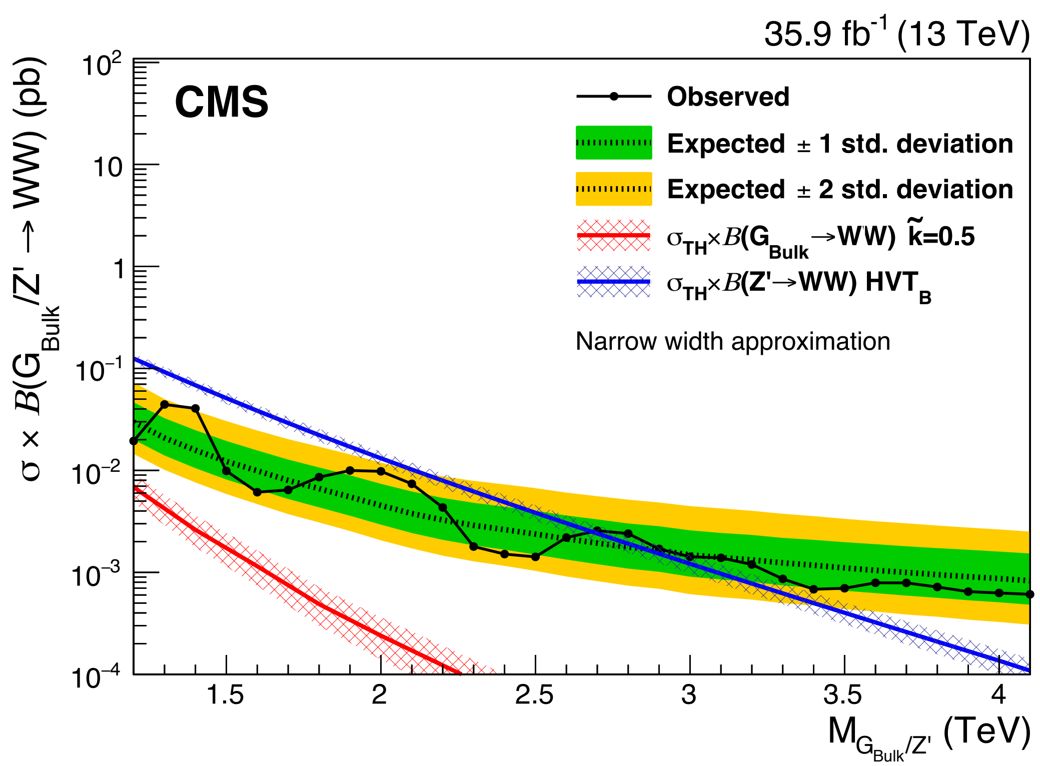
- **Signal shape extracted from MC**
  - P.D.F models constructed as composite models with Gaussian core and an exponential tail.
- **Interpolated in steps of 100 GeV**
- **Hypothesis test by comparing fits of observed data with “background-only”function and “background + signal" function.**
- **To ensure full containment of signal peak, setting limits for**
  - Single-tag: 1.2 - 6.2 TeV
  - Double-tag: 1.1 - 4.2 TeV

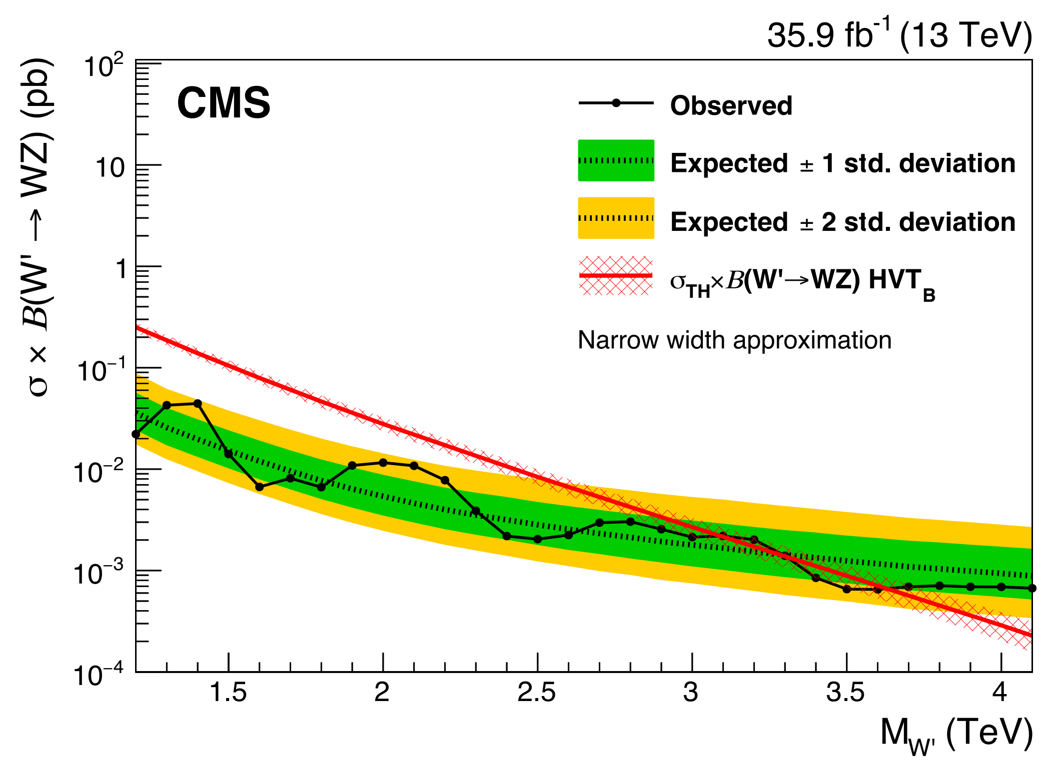


Double-tag: HVT final limits

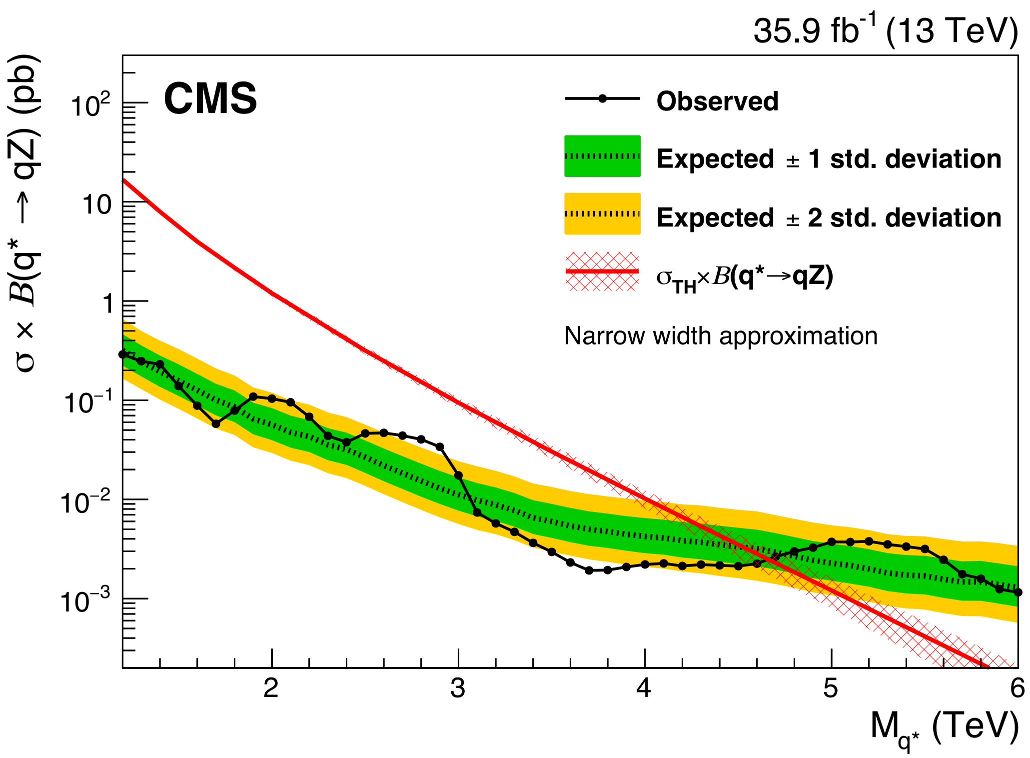


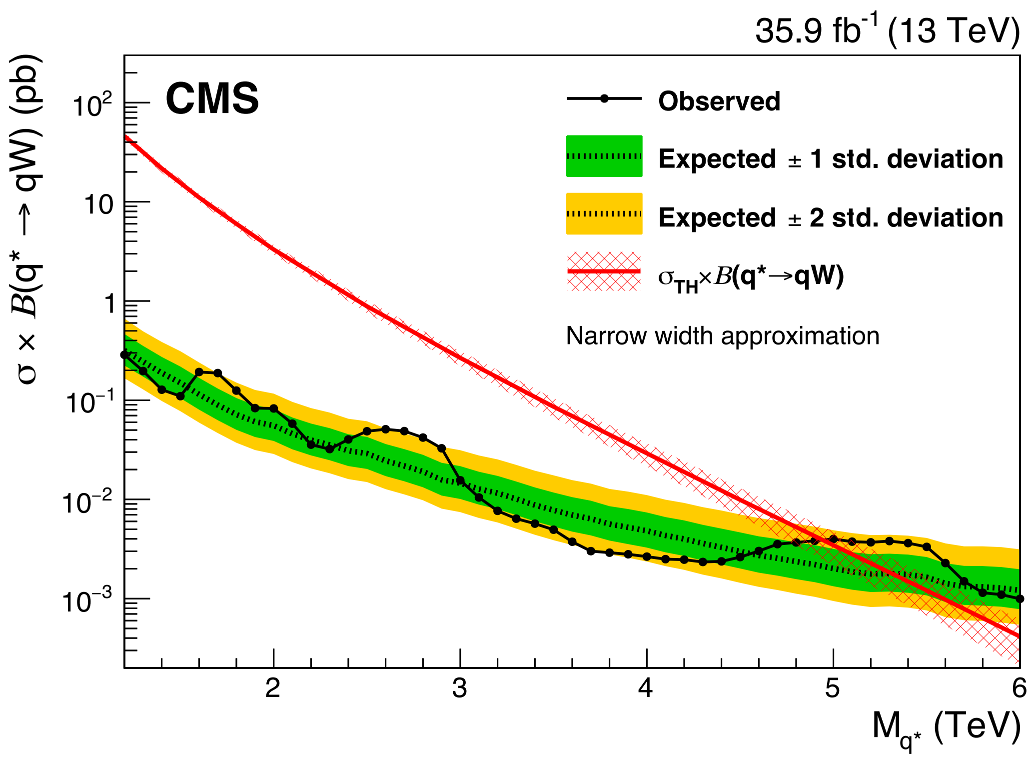






Single-tag: q* final limits





1D fitting procedure

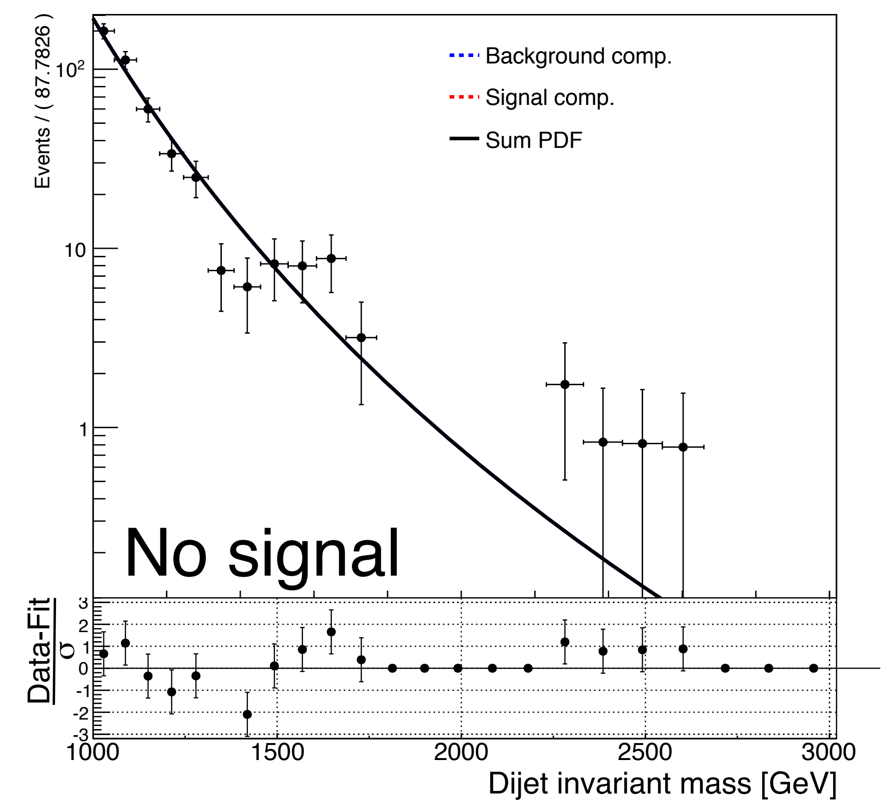
- Use binned maximum likelihood fit

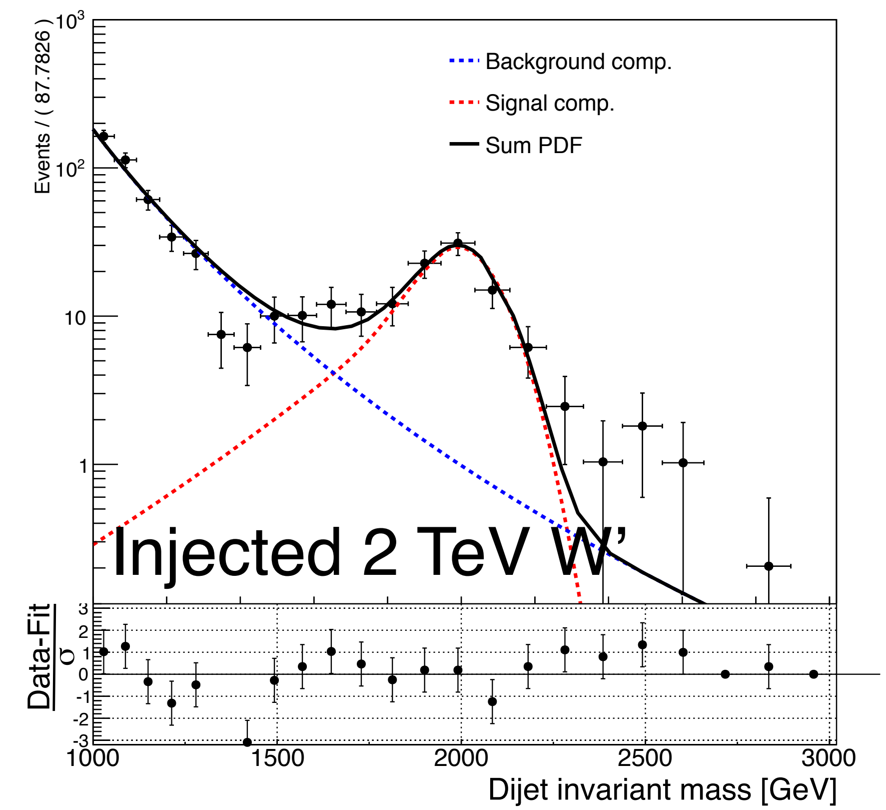
$$L = \prod_i \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!} \quad \text{with} \quad \mu_i = \sigma \cdot N_i(S) + N_i(B)$$

- Background Ni(B) is described by smooth distribution

$$\frac{d\sigma}{dm} = \frac{P_0(1 - m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2}} \longleftarrow \text{3 parameter fit used in Run 1}$$

- Is function sufficient to describe background?
  F-test: Increase number of parameters until fit shows no significant improvement
  - Only estimates **how many** parameters are needed, **NO parameters fixed**
- In limit setting, **simultaneously** fit signal yield and background function
- While maximising likelihood as a function of resonance mass, μ and parameters of background function left floating
- Observed data compared with “background-only” function and “background + signal" function





F-test

- Compute residuals RSS (DATA - fit) and DOF (nBins-nPar-1) for bins with bin content > 0 for each fit function
- If simpler fit function is correct: Relative increase in sum of squares = rel. increase in DOF

$$\frac{RSS1 - RSS2}{RSS2} \approx \frac{n1 - n2}{n2}$$

- If more complicated model is correct:

$$\frac{RSS1 - RSS2}{RSS2} > \frac{n1 - n2}{n2}$$

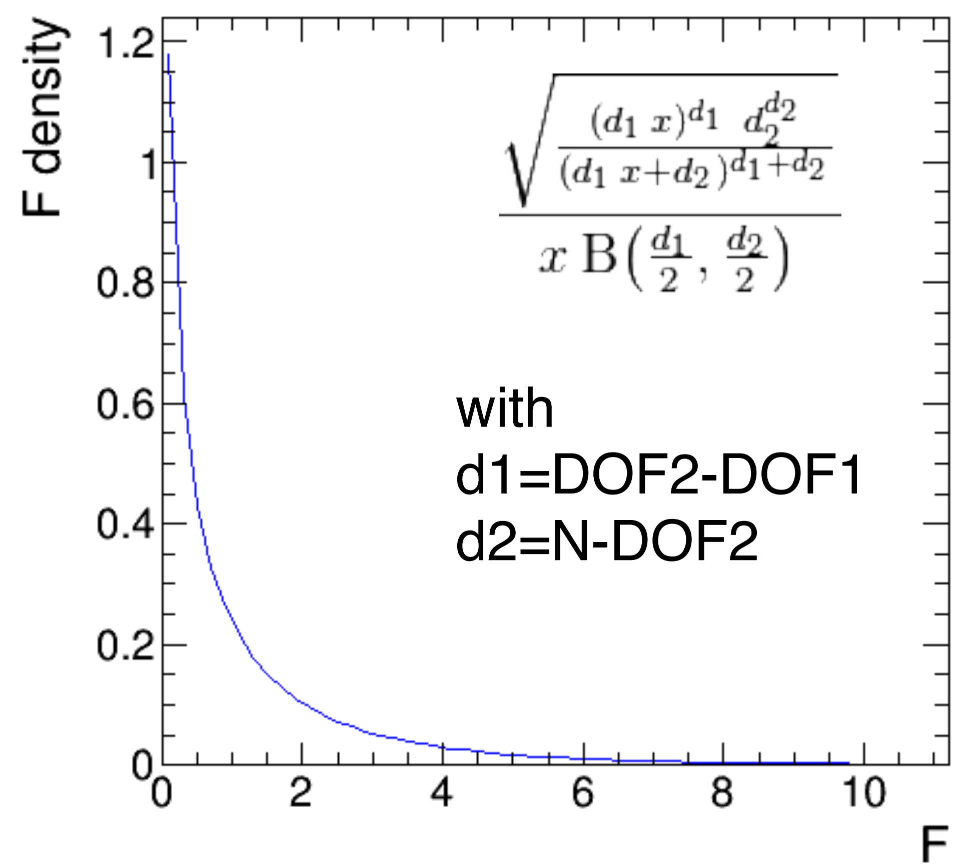
- To quantify

$$F = \frac{\frac{RSS1 - RSS2}{n2 - n1}}{\frac{RSS2}{N - n2}}$$

Only use residuals, errors not considered

- Construct F-distribution PDF

$$\frac{\sqrt{\frac{(d_1\, x)^{d_1}\, d_2^{d_2}}{(d_1\, x + d_2)^{d_1 + d_2}}}}{x\, \mathrm{B}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)}$$



- CL = 1 - Fdistr.Integral (0. , F) → gives CL under null hypothesis of simpler function being sufficient. If CL > 10%: simpler function is sufficient

Tested functions

A. *3 parameter default fit function*

A) $$\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2}}$$

B. *2 parameter*

B) $$\frac{dN}{dm} = \frac{P_0}{(m/\sqrt{s})^{P_2}}$$

C. *4 parameter*

C) $$\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2+P_3\times\log(m/\sqrt{s})}}$$

D. *5 parameter*

D) $$\frac{dN}{dm} = \frac{P_0(1-m/\sqrt{s})^{P_1}}{(m/\sqrt{s})^{P_2+P_3\times\log(m/\sqrt{s})+P_4\times\log(m/\sqrt{s})^2}}$$

LoLa

The basic setup

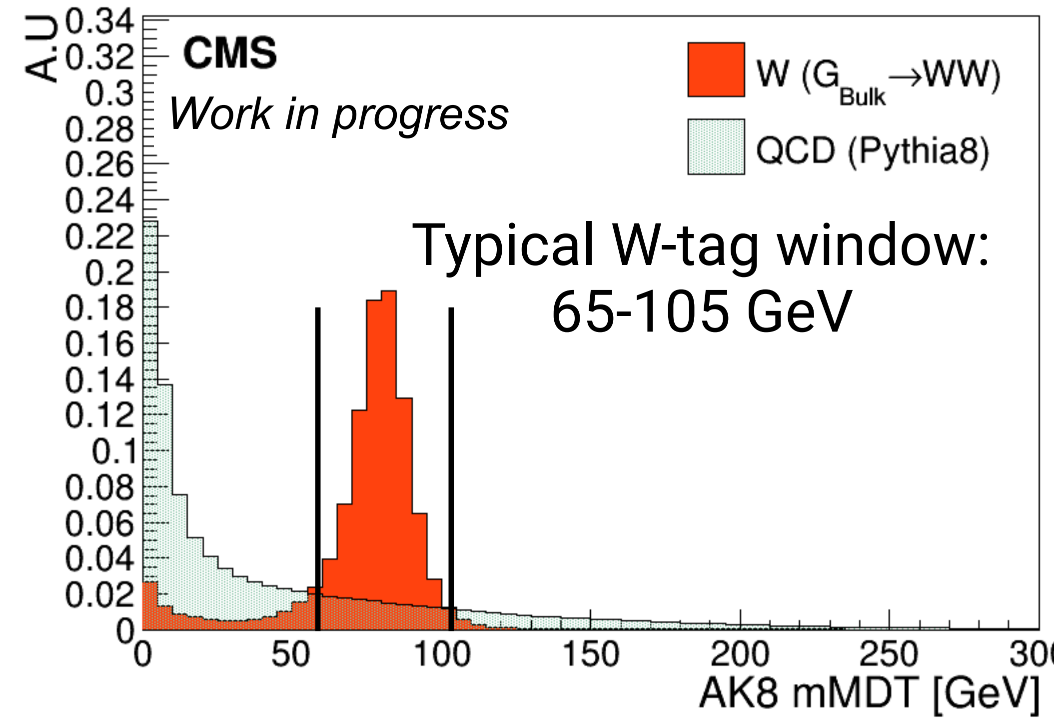
Signal

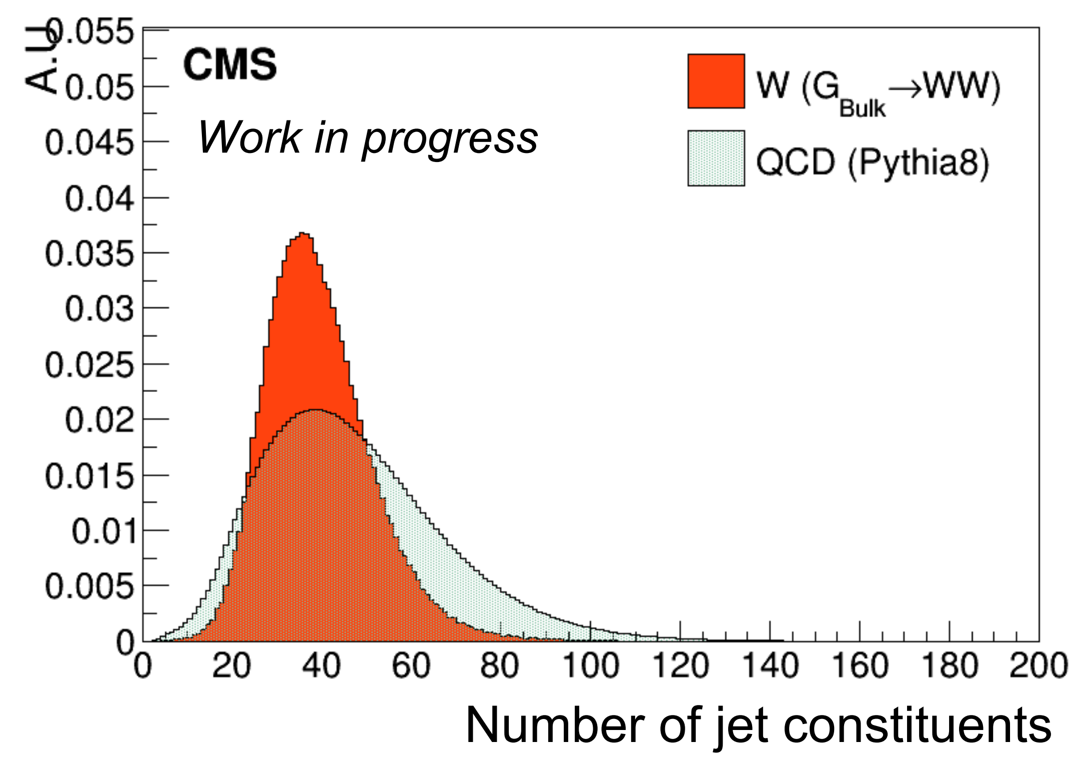
- 320k fully merged hadronic W-jets (AK8) from W'→WZ →4q ($M_{W'}$ = 0.6-4.5 TeV)
- why small training set?→ Do not mix signal samples until one is understood (can change with W polarisation etc.)

Background

- QCD Pythia8 non-W jets
- Danger: Jet substructure strongly depends on shower generators (different description of gluon radiation). Different QCD MC might yield different results

Disclaimer: The following contains student work in progress studies and not CMS approved results





Input

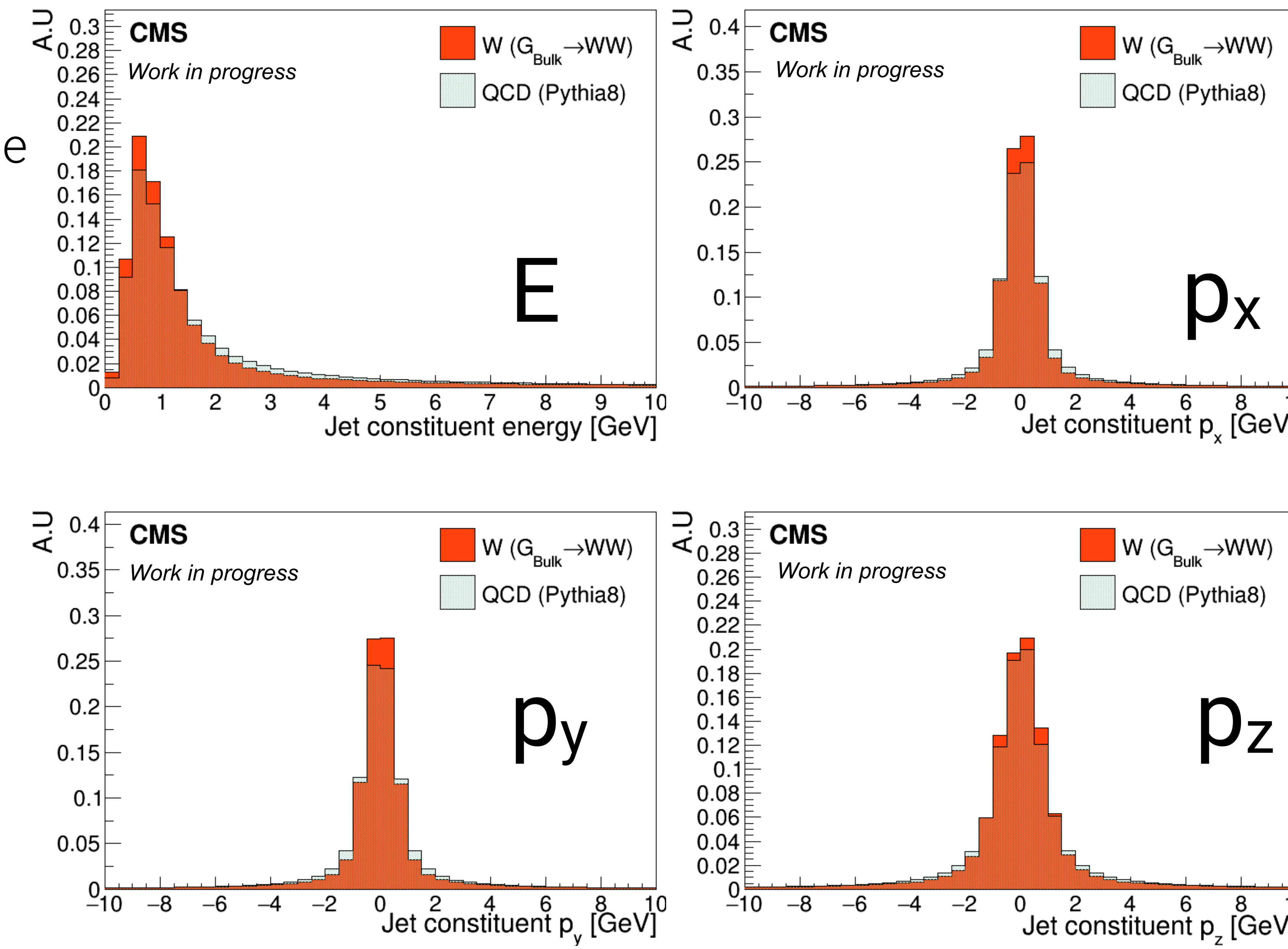
Four input features

- 4-vectors of the N=20 highest-p_T jet constituents of PUPPI (PileUp Per Particle Removal) AK8 jets

4x20 matrix $k_{\mu,i}$ for each jet

$$(k_{\mu,i}) = \begin{pmatrix} k_{0,1} & k_{0,2} & \cdots & k_{0,N} \\ k_{1,1} & k_{1,2} & \cdots & k_{1,N} \\ k_{2,1} & k_{2,2} & \cdots & k_{2,N} \\ k_{3,1} & k_{3,2} & \cdots & k_{3,N} \end{pmatrix}$$

(4 Features x 20 constituents)

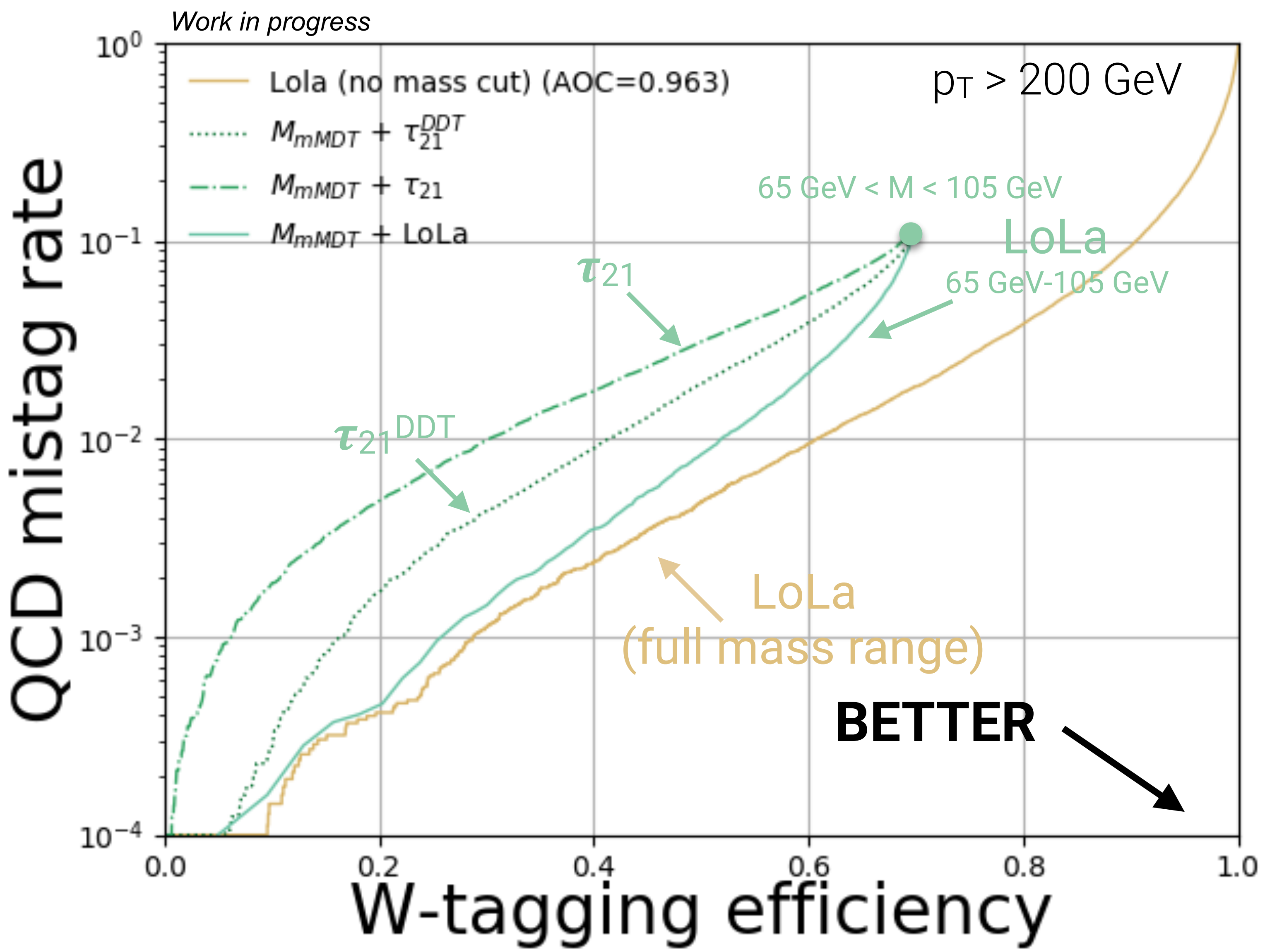


Overall performance

Compare performance to most commonly used V-taggers

- Softdrop mass + τ_{21}
- Softdrop mass + τ_{21}^{DDT} (mass/p_T decorrelated τ_{21})

- LoLa performs significantly better than current baseline
  - 20% higher ε_S at given ε_B compared to best cut-based
  - no need for mass window, increased signal acceptance



For two-W final state, 43% increase in signal efficiency for same mistag rate as current baseline (B2G-17-001)

Beyond performance

LoLa is naturally learning that p_T and mass are discriminating variables

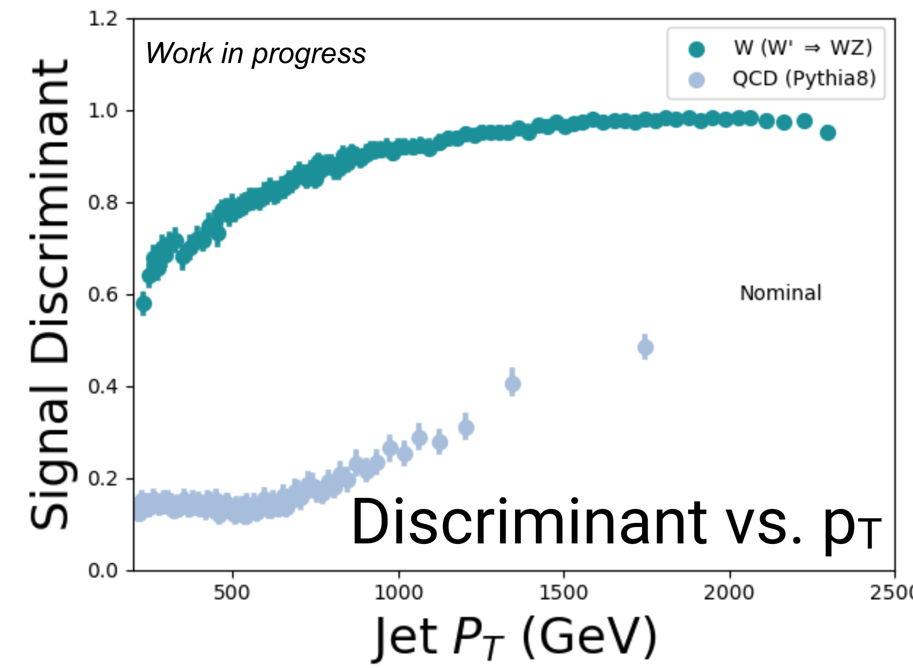
p_T-dependence is a problem because

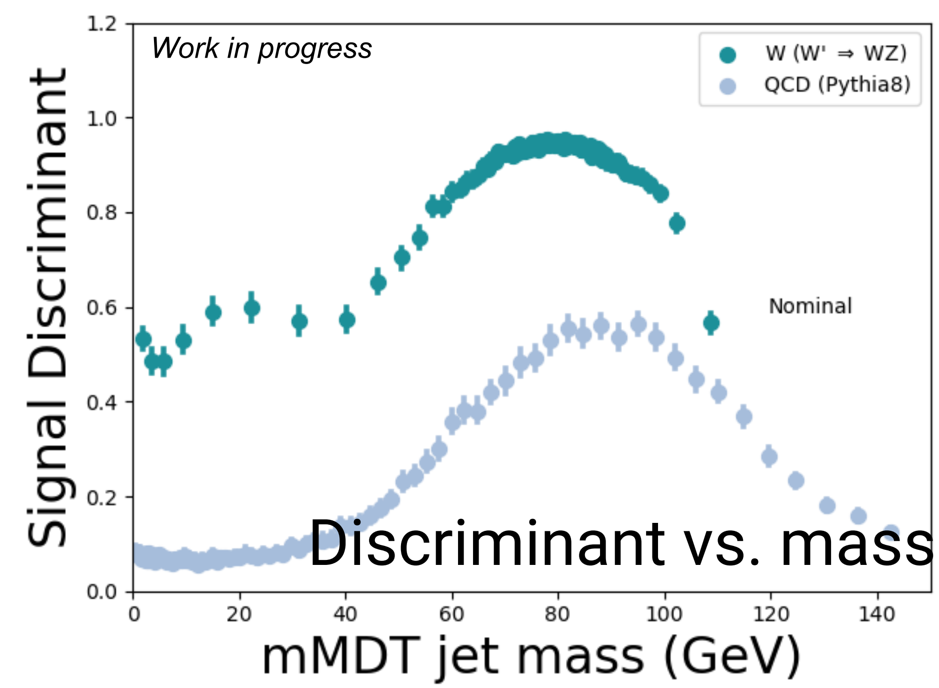
- signal efficiency is variable, requires working point scaling with p_T
- p_T (tagger validation region) != p_T (signal region)

Mass-dependence in itself not a problem, but could introduce large background rate uncertainties if using mass sidebands

- ultimately trade-off between efficiency and (analysis-dependent) systematics
- can not decide before checking in analysis

Output strongly correlated with mass/p_T





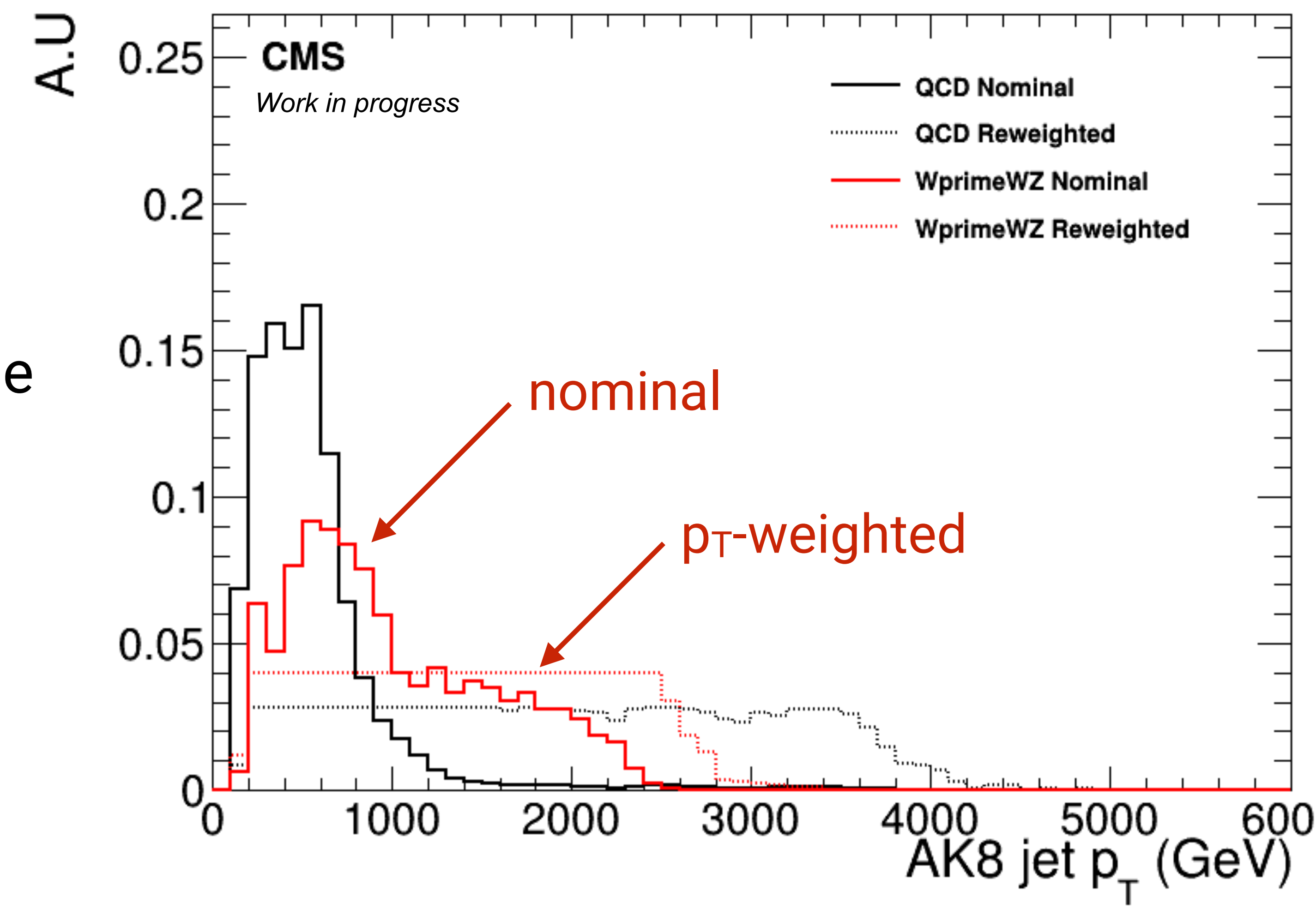
Coping with p_T

LoLa uses full jet-p_T range (> 200 GeV) in training and validation

- want tagger that offers discrimination where there is most background and W-boost not extreme (p_T < 600 GeV)

Reweight training set event-by-event to be flat in p_T-space

- passed as sample weights to training

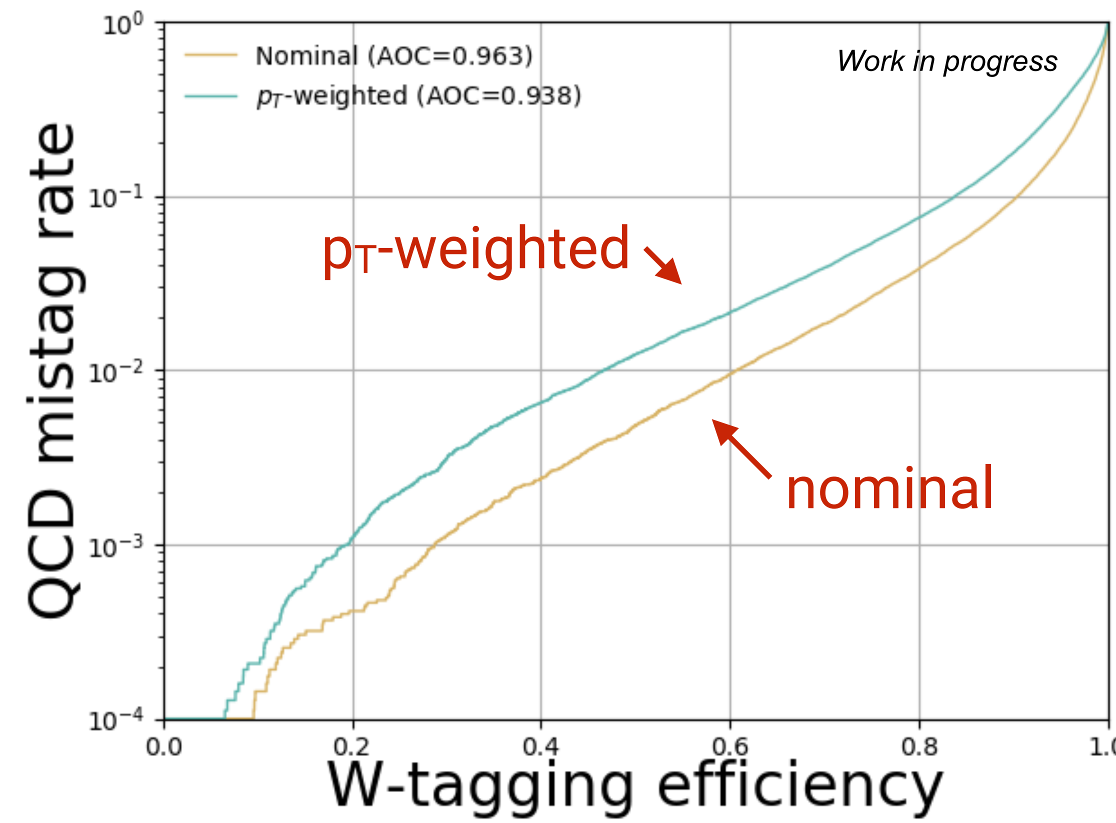


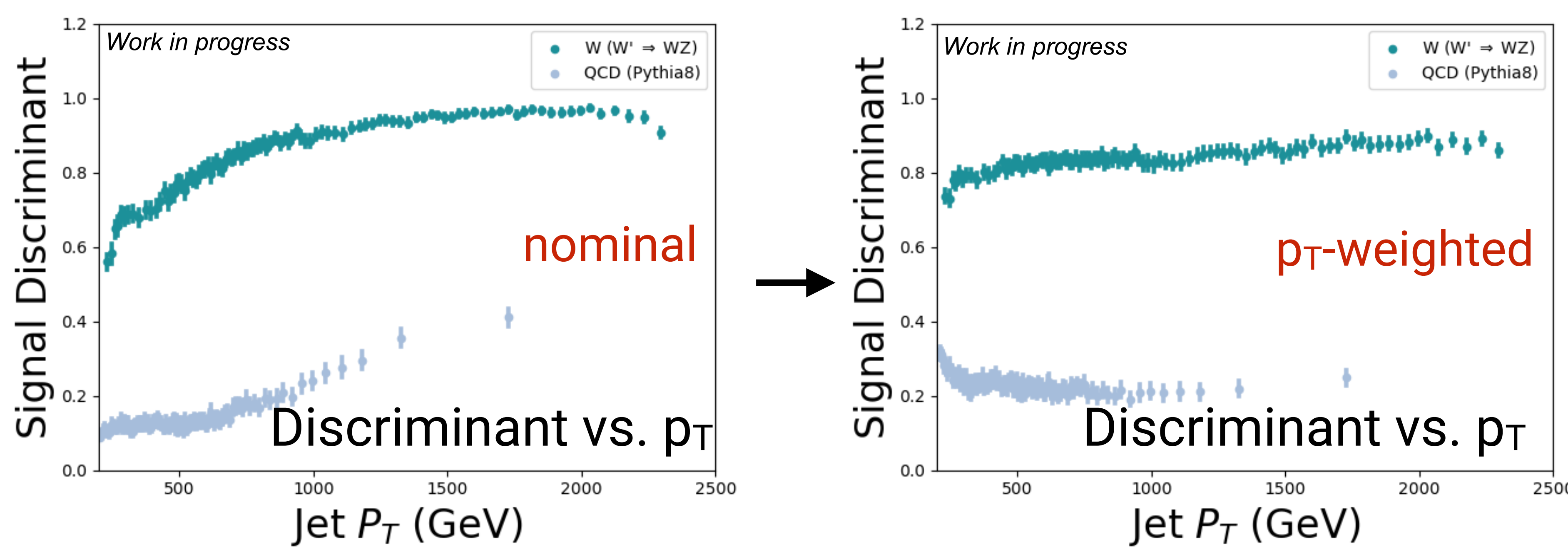
Coping with p_T

Such strategies yields loss in overall performance, but reduced p_T-dependence

- a boost of statistics in extreme bins could improve performance for a p_T-weighted training

No “truth” for which solution is better before running full analysis including systematics for p_T-dependent tagging





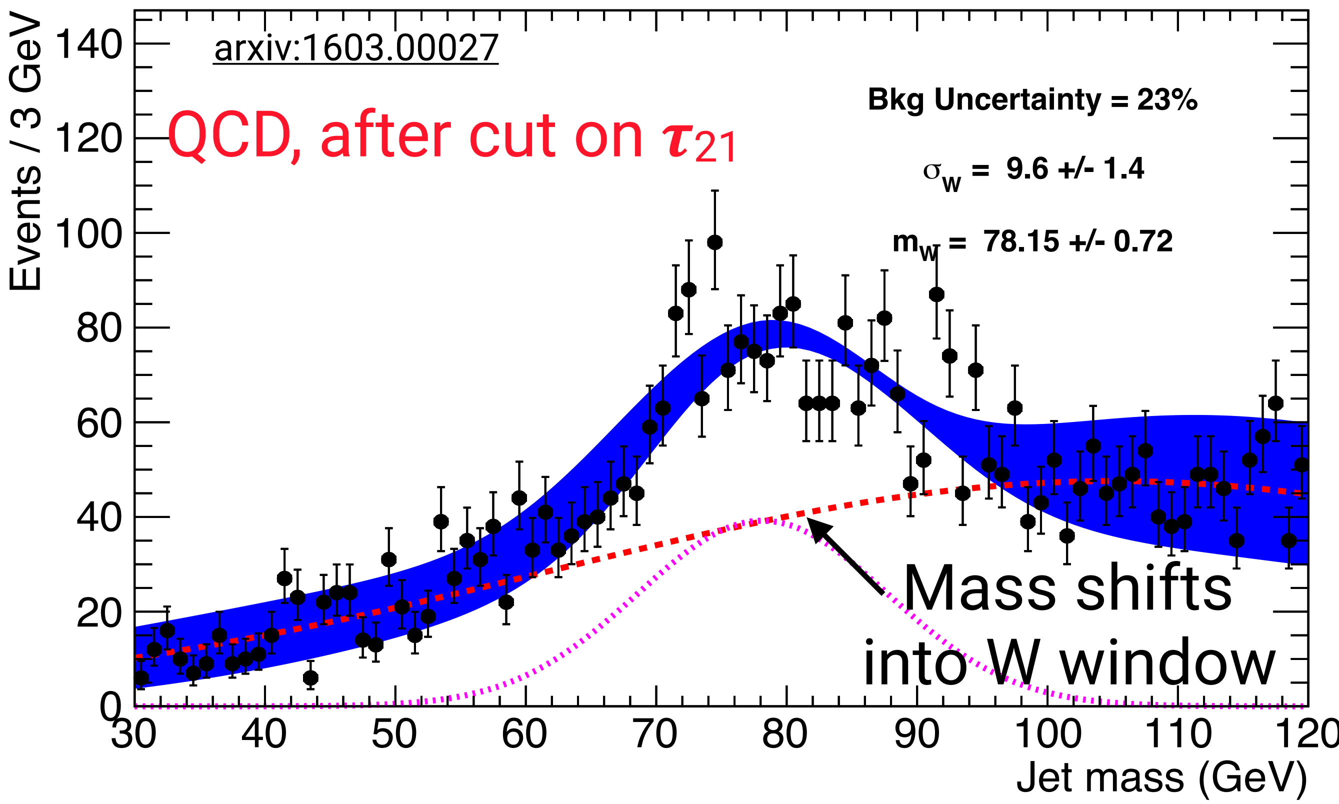
Mass sculpting

If you feed a DNN W-jet constituent 4-vectors it will inevitably learn W-mass

- good! Clearly W-mass != q/g-jet mass

Unfortunately, we often estimate background in mass sidebands

- bad! After cut on tagger, jet mass is sculpted making background spectrum difficult to constrain



Mass sculpting

! LoLa removes low-mass QCD,
→ increased signal acceptance
w.o mass window

At 1% mistag rate, see significant mass sculpting with LoLa

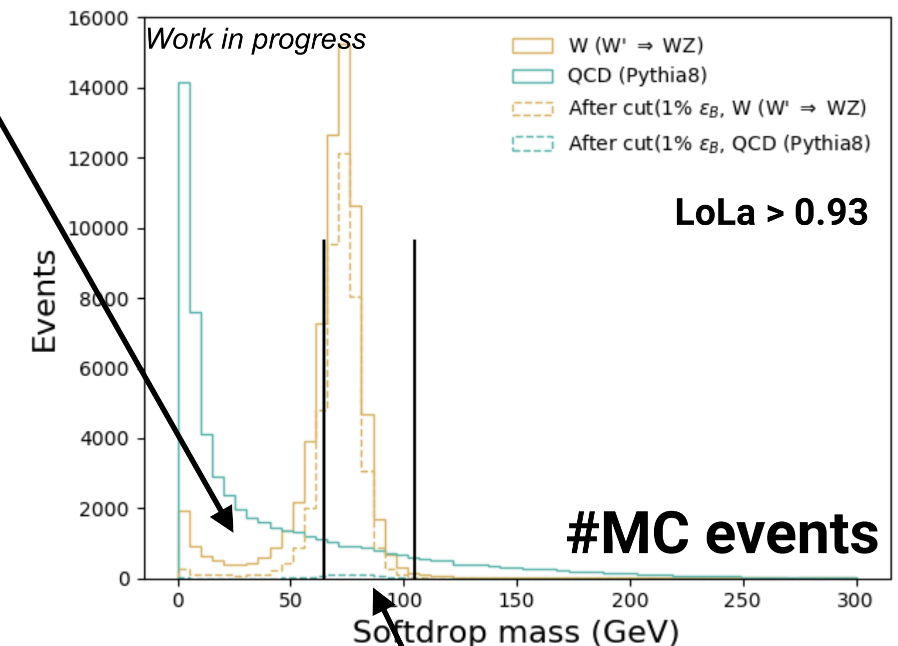
- bulk of remaining QCD jets after cut are in signal region

Hot topic in ML for jets: adversarial DNNs that penalise loss if mass is learned (nicely shown in C. Shimmin et. Al)

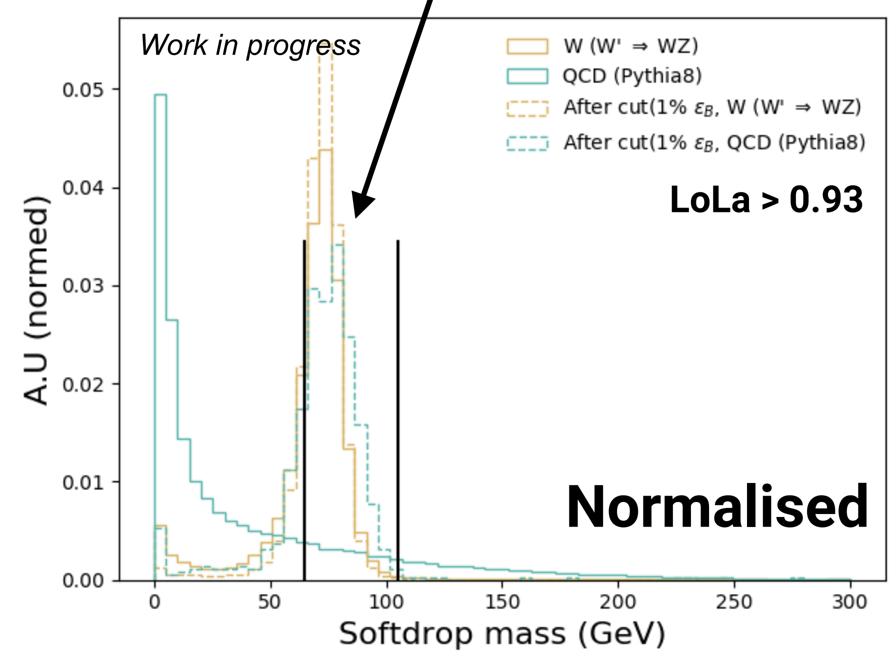
- loss in efficiency, but overall improvement due to reduced uncertainties when relying on mass-sidebands

Adversarial LoLa in progress, but best to offer both (for non-sideband based and sideband-based analyses)

- unconstrained, high-efficiency LoLa
- mass-decorrelated LoLa (adversarial)

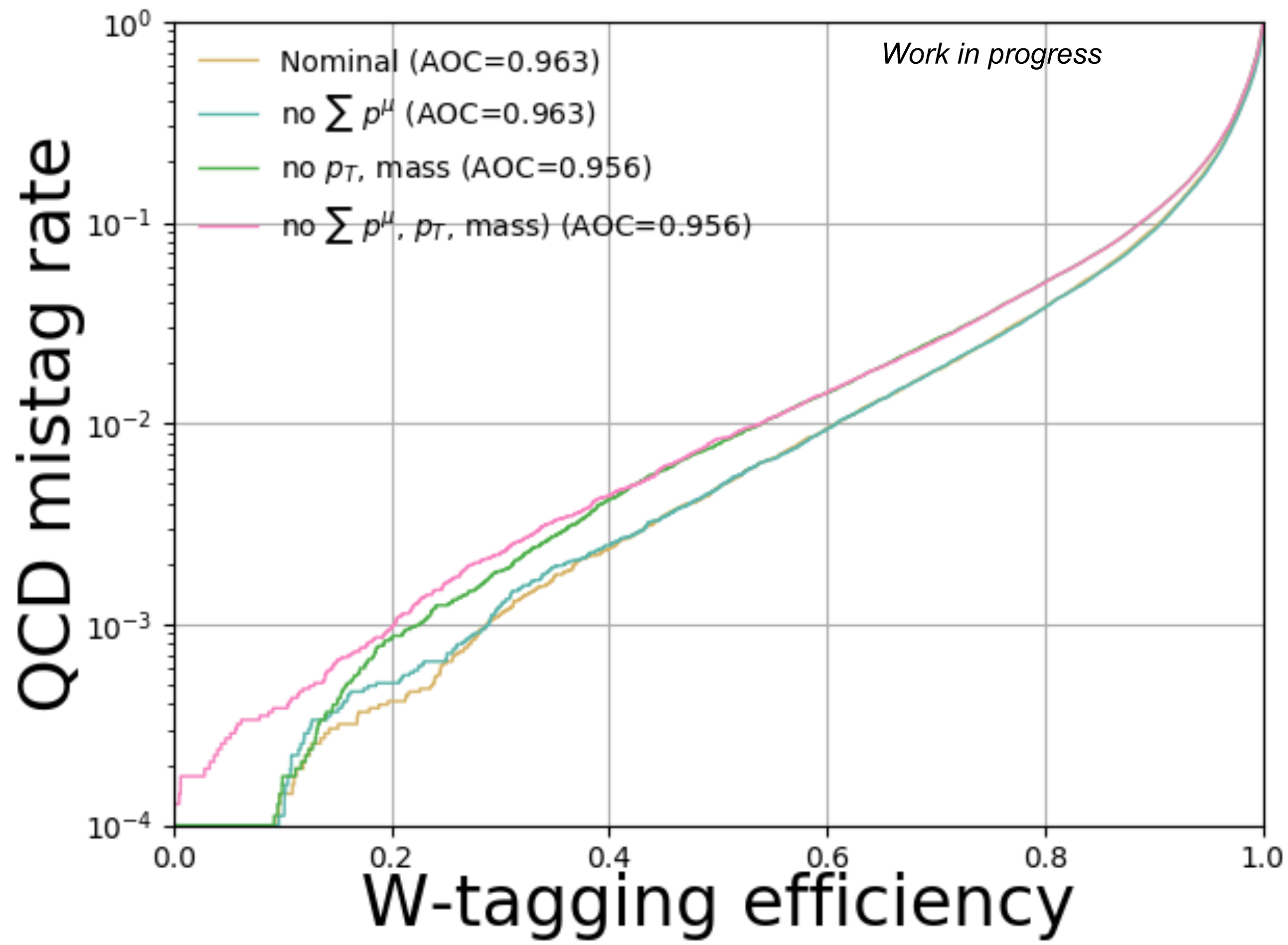


Bulk of remaining 1% background jets are in W-mass window

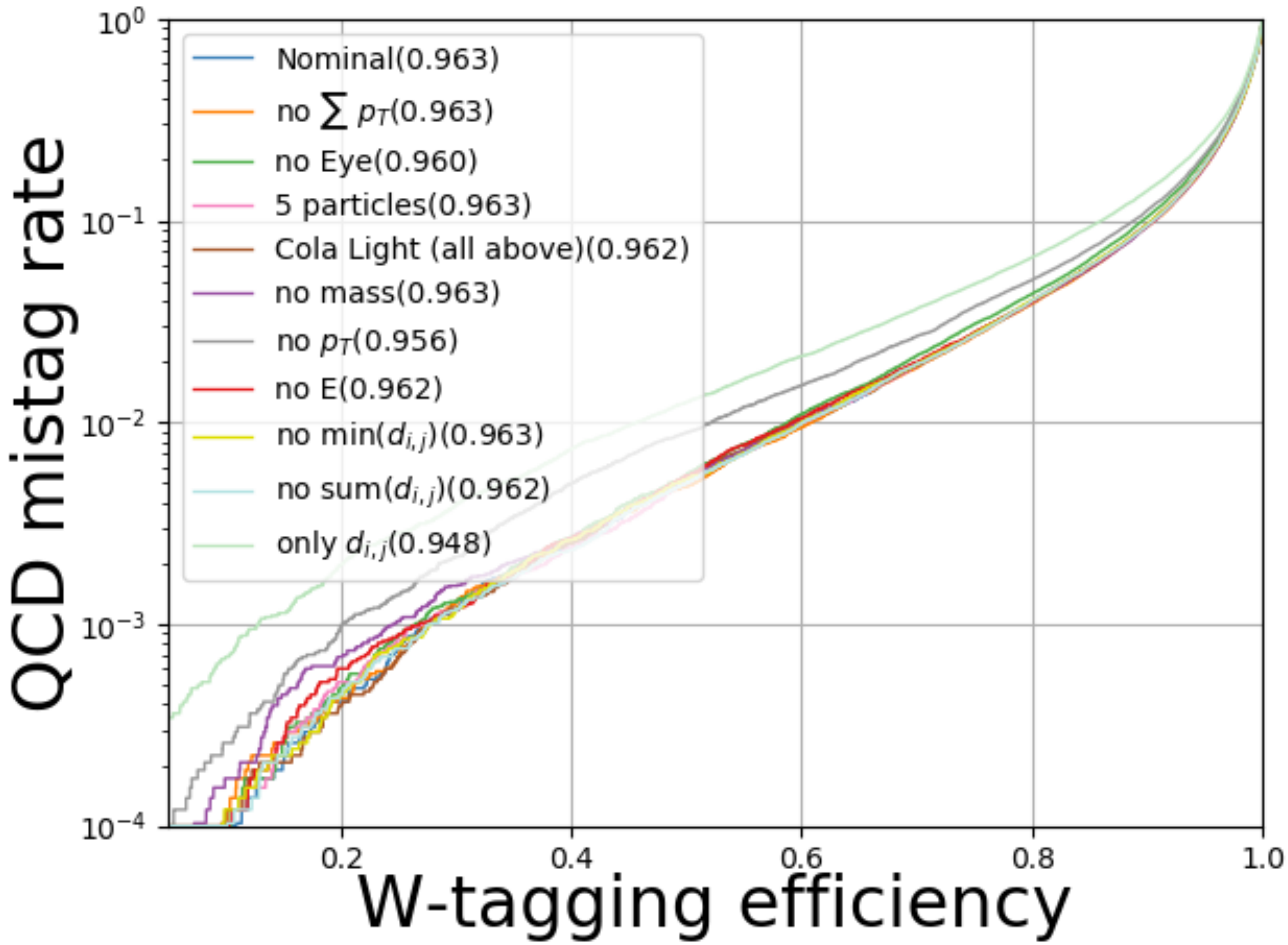


What does LoLa learn?

- Compare nominal training to training after removing variables sensitive to mass and p_T
- Remove CoLa column that passes sum of all 4-momentum (“jet” 4-vector)
  - not much impact on overall performance
  - not much information taken from LoLa “n-subjettiness”
- Remove Lola mass and p_T variables reduce performance significantly
  - worst when removing jet 4-vector, mass and p_T

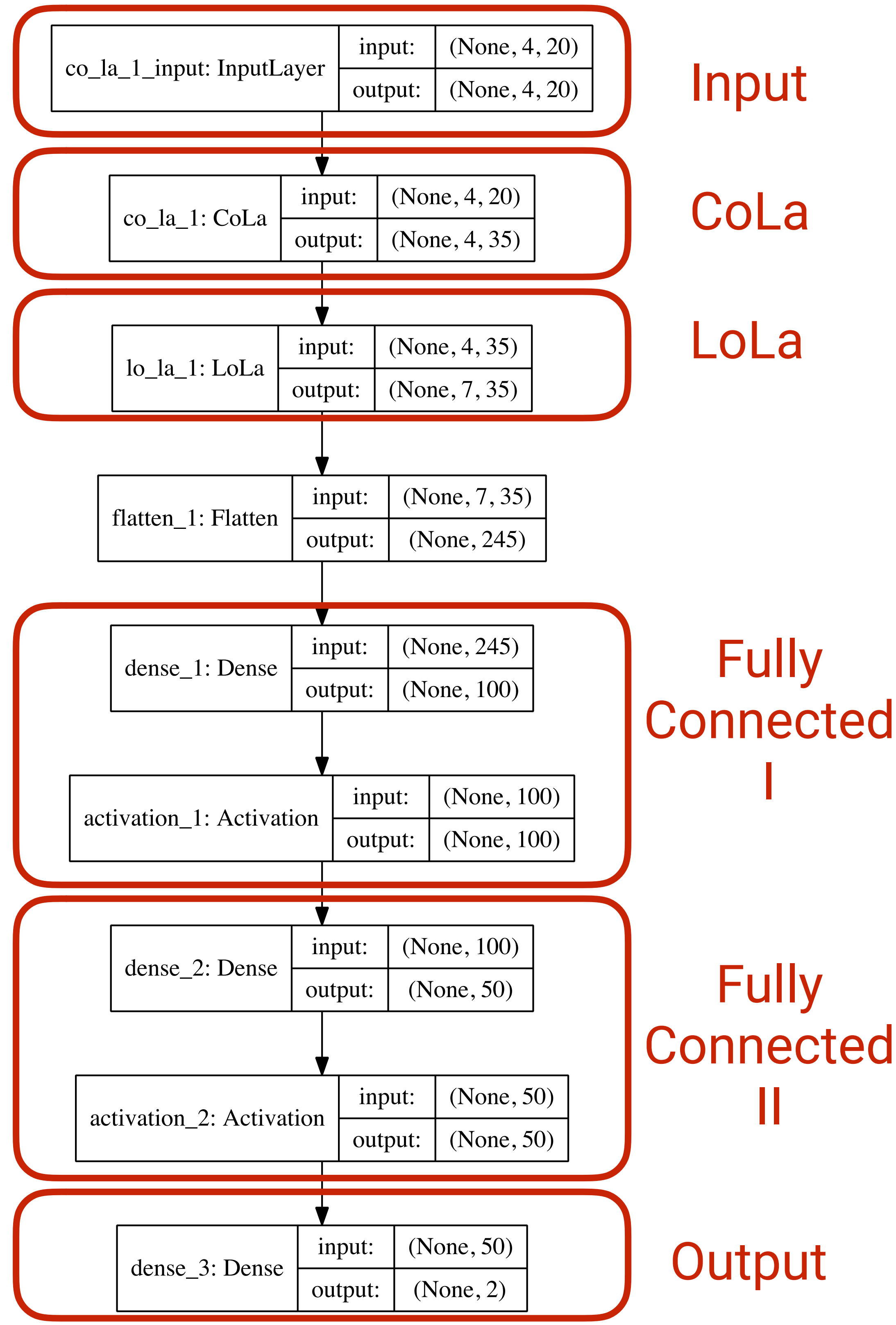


What does LoLa learn?



Model

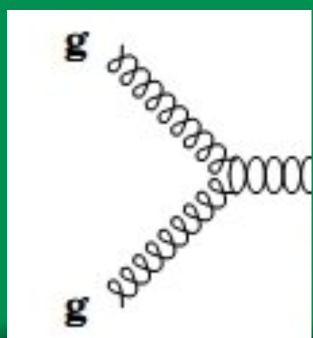
- 4 layer DNN doing supervised learning with fixed-size input vectors
  - feed forward sequential network
  - Two novel layers (CoLa and LoLa) implementing Minkowski metric and “substructure” calculations (see later) and two fully connected layers
- Technicalities
  - Keras with Theano backend (rewriting to Tensorflow)
  - Loss function: categorical crossentropy
  - ADAM optimiser (adapt learning rate of model parameters during training)
- Train 200k + Test 60k + Val 60k on AWS (CMS Christmas Wishlist: GPU cluster!)



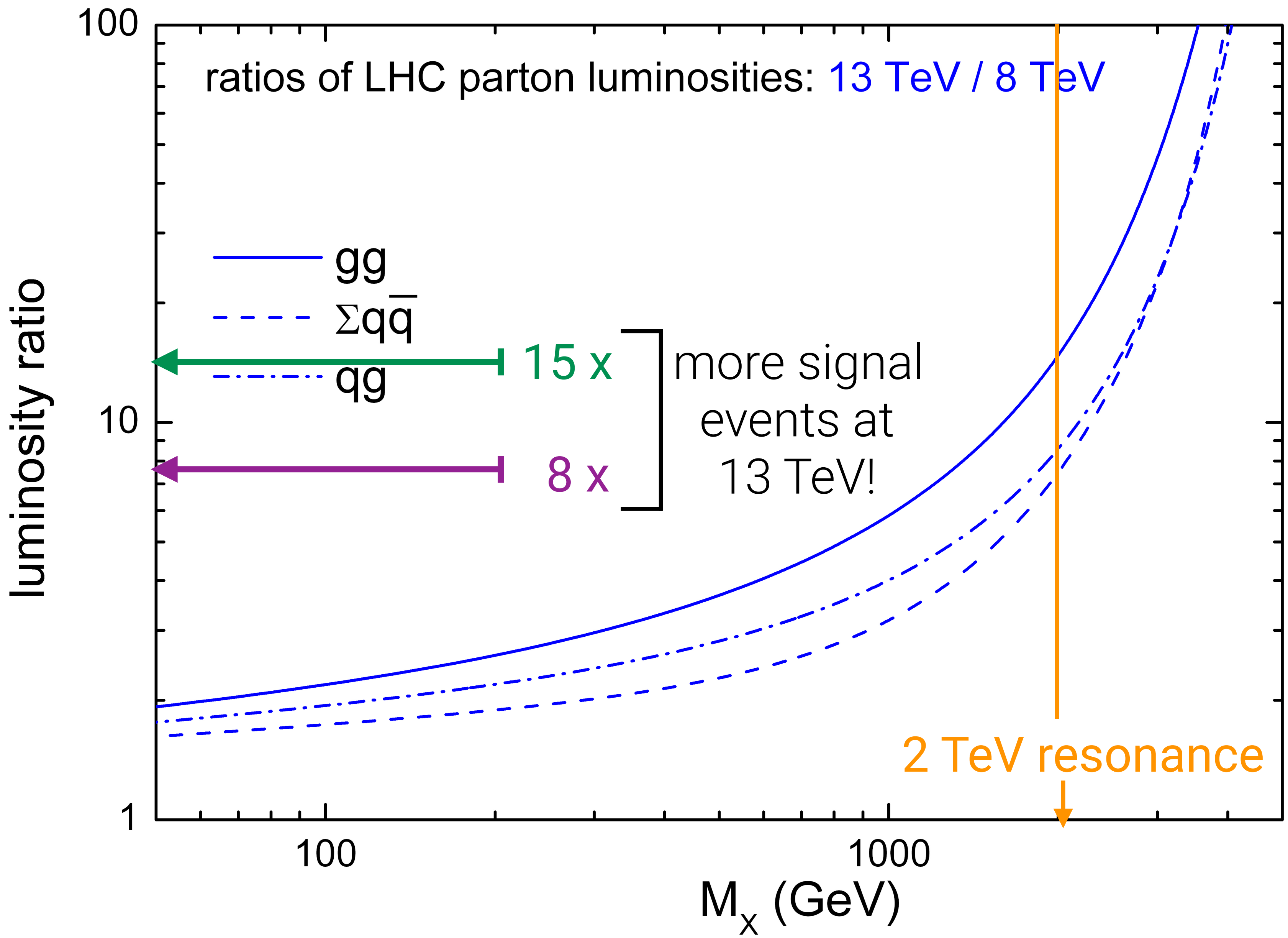
Universität Zürich[UZH]

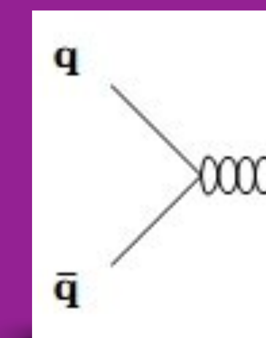
Intro

What could it be?



Mainly produced through gg fusion





Mainly produced through q̄q annihilation

With only 3 fb^{-1} of 13 TeV data, same discovery potential as 8 TeV dataset of 20 fb^{-1}

Signature: $G_{Bulk} \rightarrow WW$ and $G_{Bulk} \rightarrow ZZ$

Signature: $Z' \rightarrow WW$ and $W' \rightarrow WZ$