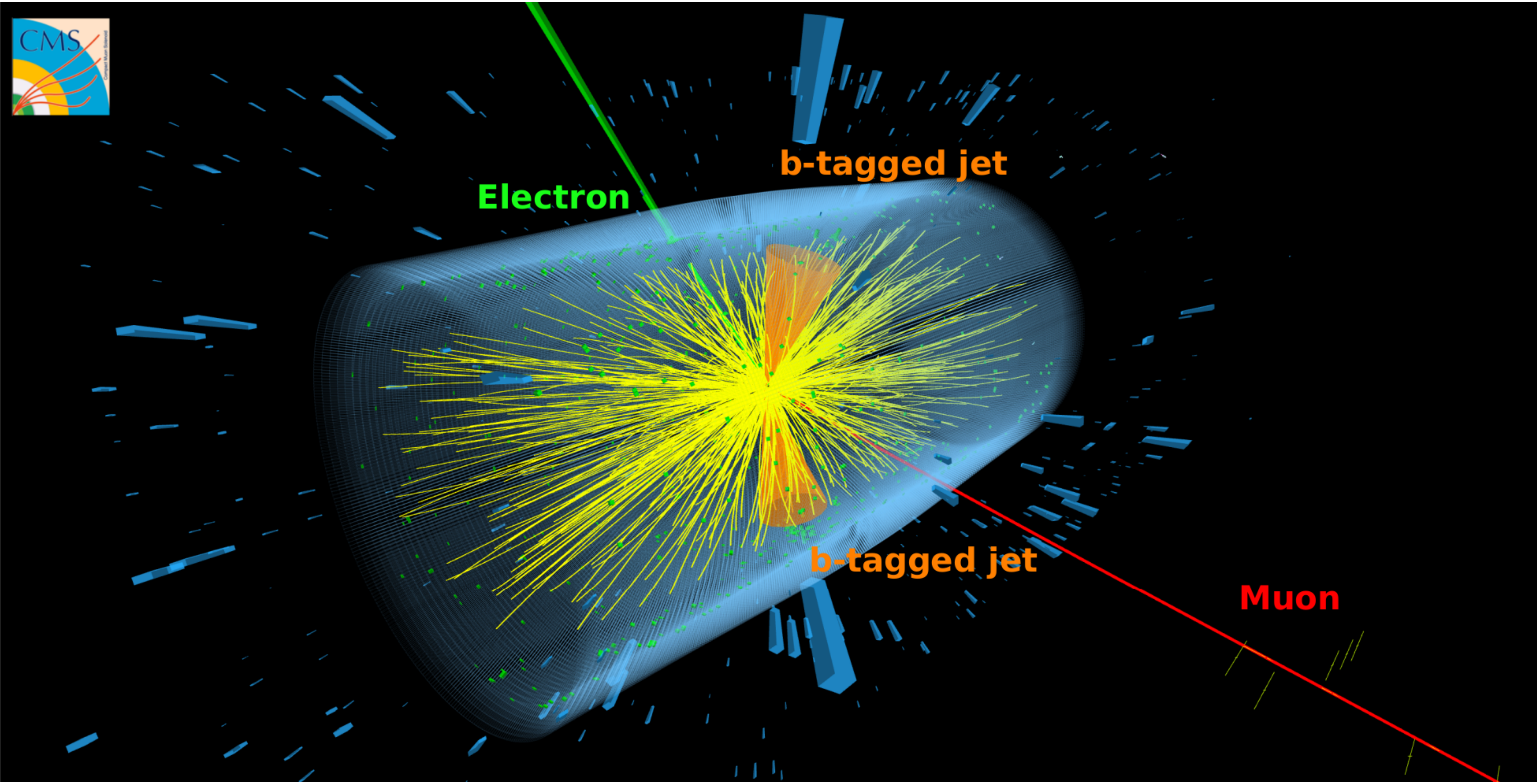
Top quark evidence in PbPb collisions at $\sqrt{s_{NN}}$ = 5.02 TeV in CMS

Inna Kucher

Laboratoire Leprince-Ringuet

IPNO seminar
13/02/2020

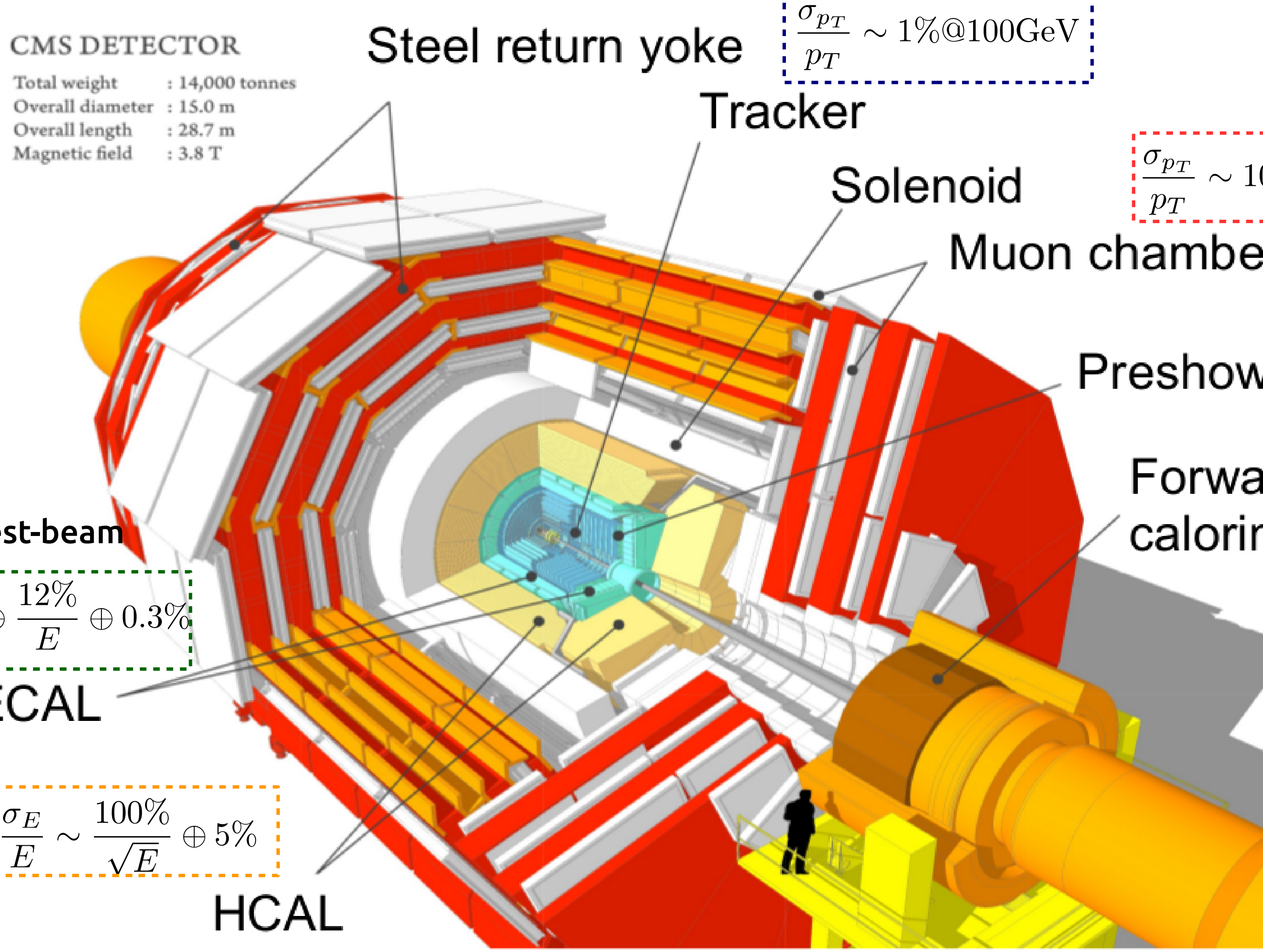
Top quark candidate in PbPb 2018 collisions



PbPb environment is much busier than pp one

(track multiplicity is ~ 10k in PbPb vs pp ~ 750)

CMS detector



$$\frac{\sigma_{p_T}}{p_T} \sim 1\%@100\mathrm{GeV}$$

$$\frac{\sigma_{p_T}}{p_T} \sim 10\%@1\mathrm{TeV}$$

From the test-beam

$$\frac{\sigma_E}{E} \sim \frac{2.8\%}{\sqrt{E}} \oplus \frac{12\%}{E} \oplus 0.3\%$$

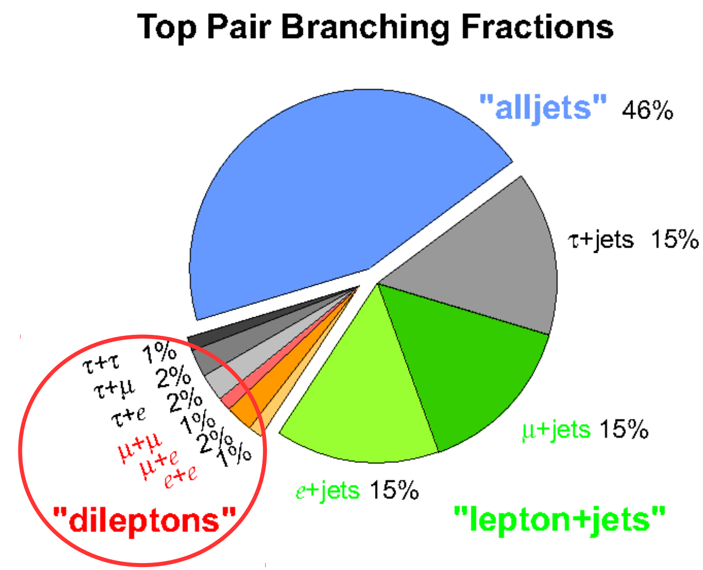
$$\frac{\sigma_E}{E} \sim \frac{100\%}{\sqrt{E}} \oplus 5\%$$

Pseudorapidity:

$$\eta = -\ln\left[\tan(\frac{\theta}{2})\right]$$

All subsystems are necessary to detect top quark decay particles

Analysis strategy



Fully leptonic decays of $t\bar{t}$ pair : **$t\bar{t} \rightarrow l\nu_l b\ l'\nu_{l'}\bar{b}$**

→ highest purity
→ small BR

Two analysis approaches :

1) Only lepton information : **$t\bar{t} \rightarrow l\nu_l b\ l'\nu_{l'}\bar{b}$**
2) Lepton + b-jet information : **$t\bar{t} \rightarrow l\nu_l b\ l'\nu_{l'}\bar{b}$**

Lepton only analysis : insensitive to b-jet quenching

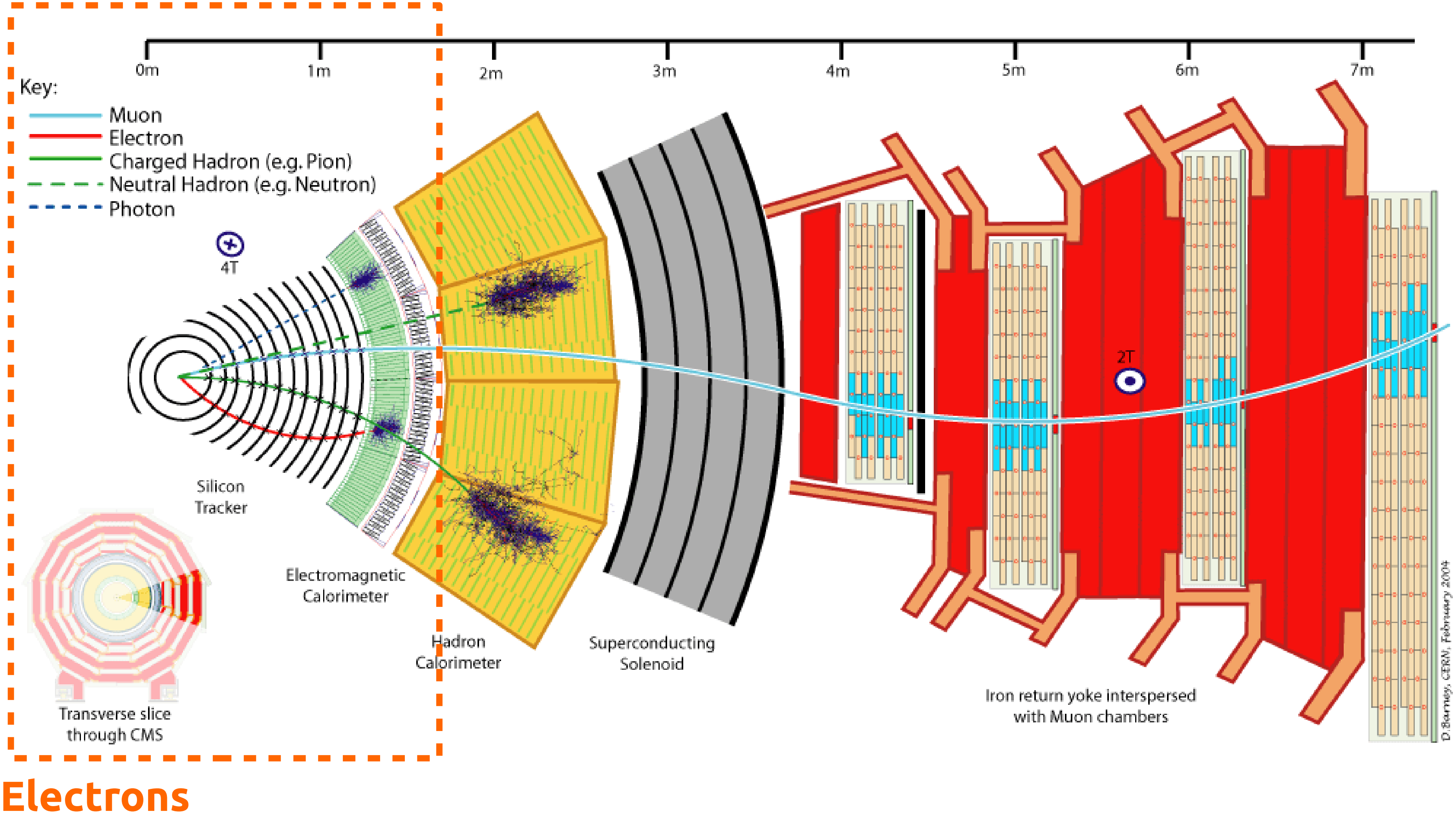
Backgrounds :

- ee/μμ : DY+jets (DY = ll' pairs coming from Z or γ*) with MG5_aMC@NLO
- eμ: Z→ ττ with MG5_aMC@NLO
  Non-prompt = W+jets and QCD multi-jets with heavy flavor decays
  (from data using event-mixing technique)
- Small contributions from tW with NLO POWHEG; ZZ, WW and WZ (VV) with POWHEG

Analysis was blinded to the mass region of interest in data

Electron reconstruction

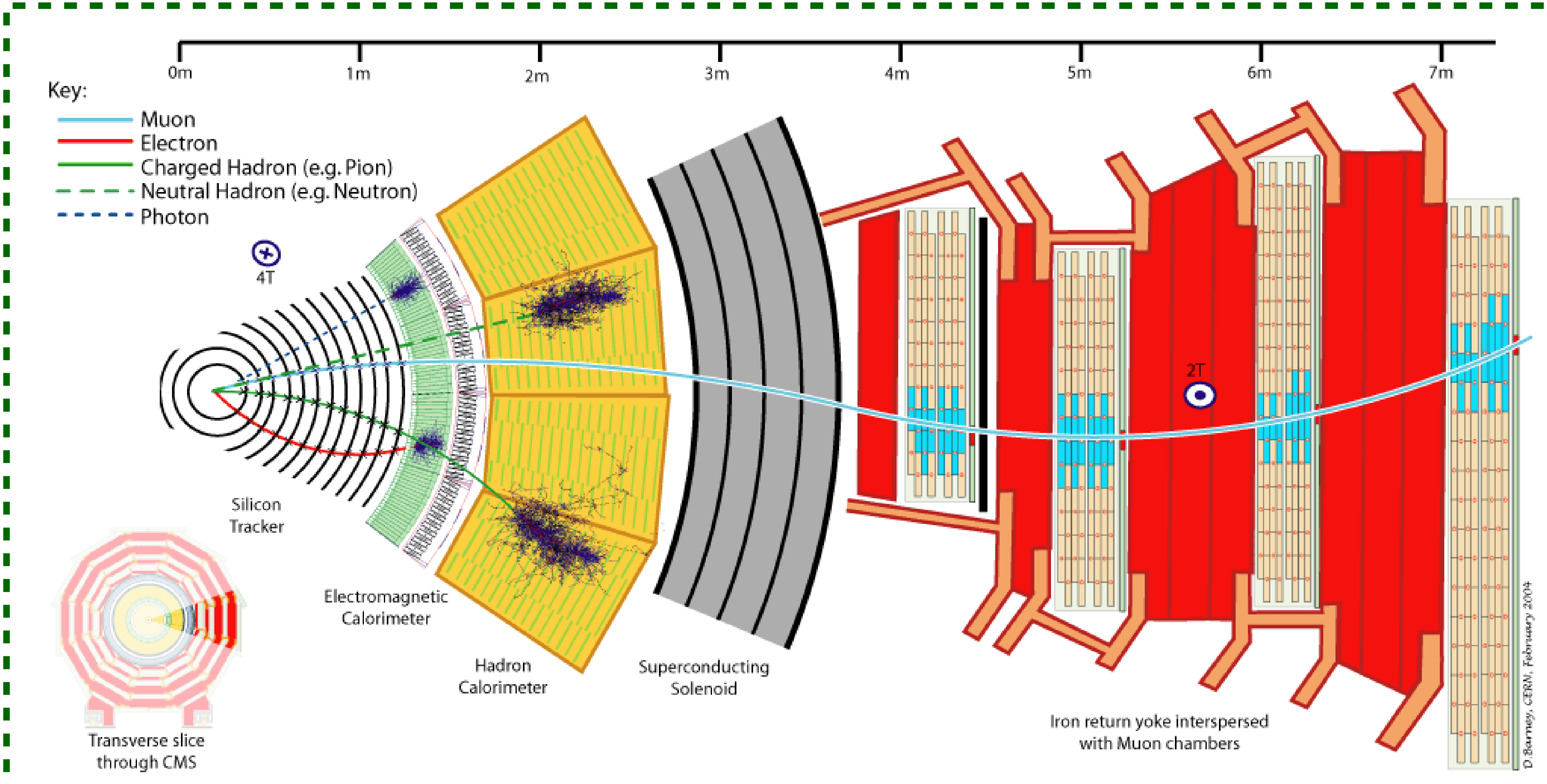
Electron (e) reconstruction : combines tracker and ECAL information



[JINST 10 (2015) P06005]

Muon reconstruction

Muon (μ) reconstruction : combines tracker and muon stations information



Muons

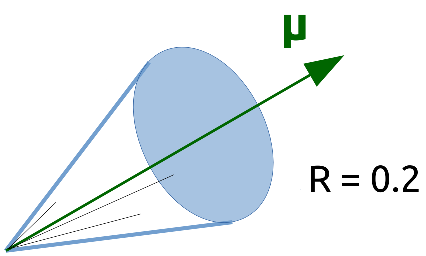
JINST 13 (2018) P06015

Lepton reconstruction and identification

Muon (μ) reconstruction : combines track and muon stations information

Electron (e) reconstruction : combines track and ECAL information

μ/e identification and selections :



→ identification criteria were optimized for PbPb environment

→ isolation criteria :

I_{rel} = [I – UE(ρ)]/p_T ; I – p_T sum of all particles inside the cone around the μ/e direction
UE(ρ) - median energy density of the underlying event

I_{rel} < 0.08 (-0.06) for μ(e) → flattens the dependence on the centrality

→ kinematic selections : p_T > 20(25) GeV and |η| < 2.4(2.1) for μ(e)
→ μ and e are opposite charged

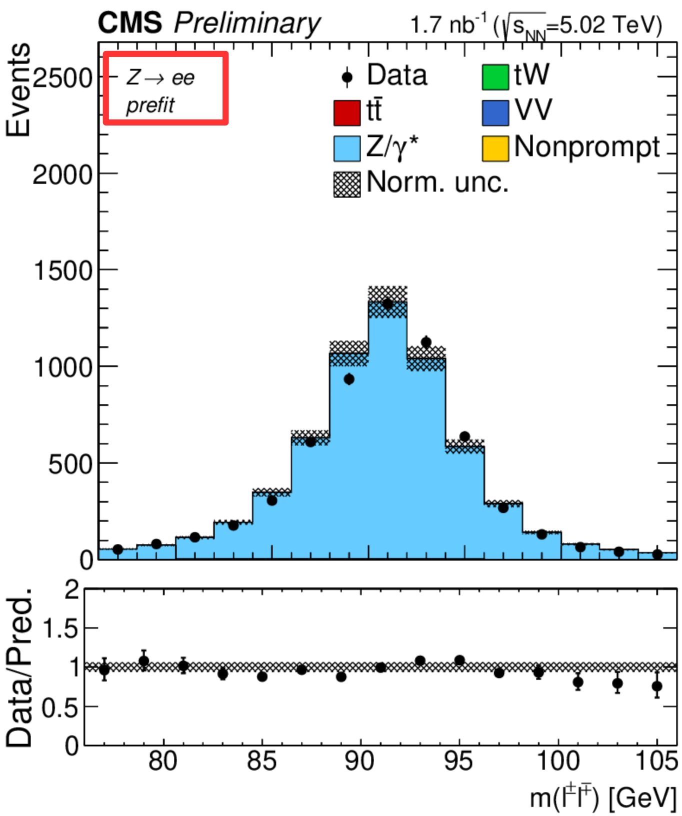
Dilepton mass

$t\bar{t} \to l\nu_{l}b\ l'\nu_{l'}\bar{b}$

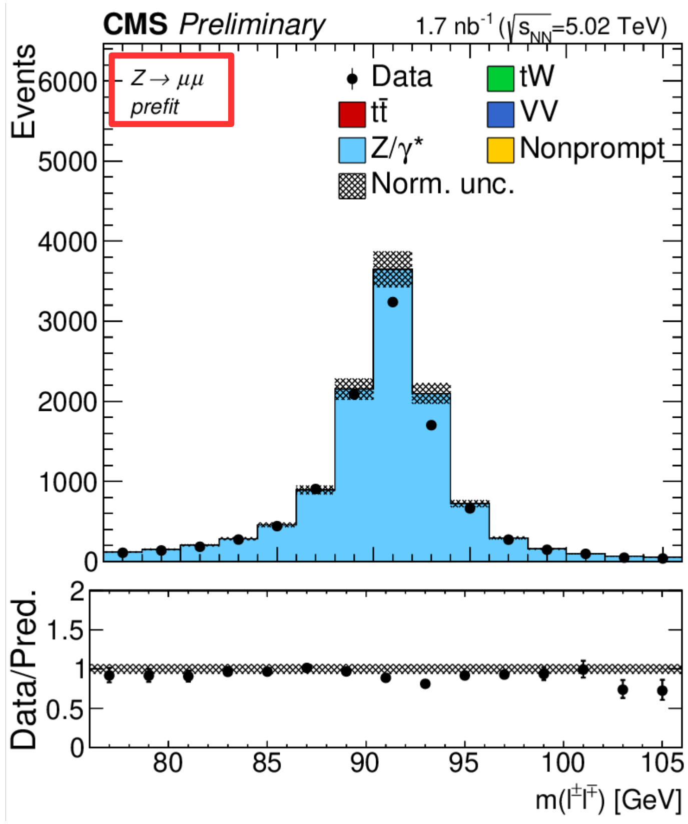
Three dilepton combinations are possible : ee, μμ, eμ

Distributions are **prefit** : MC represents the **expected** yields

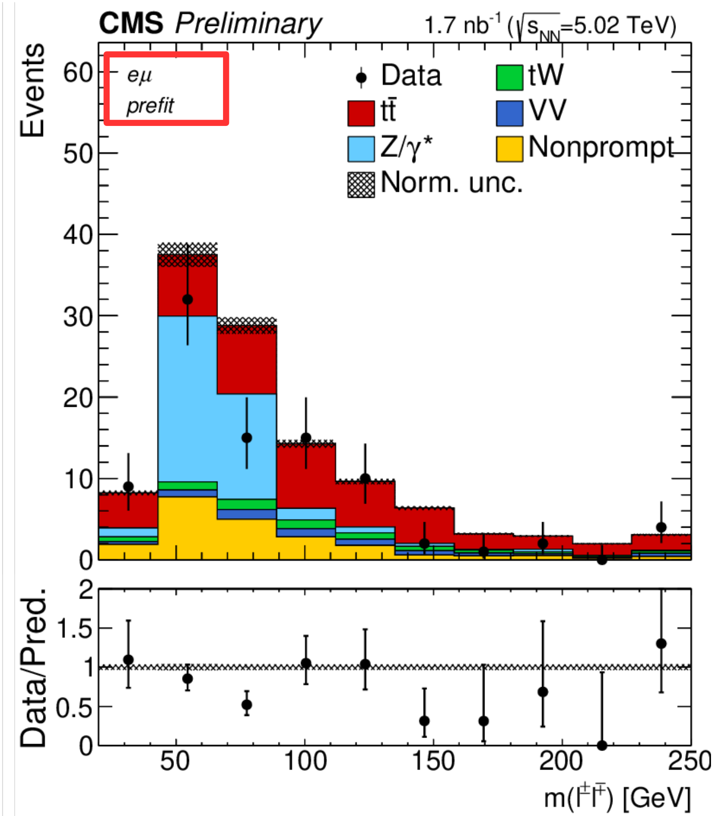
ee



μμ



eμ



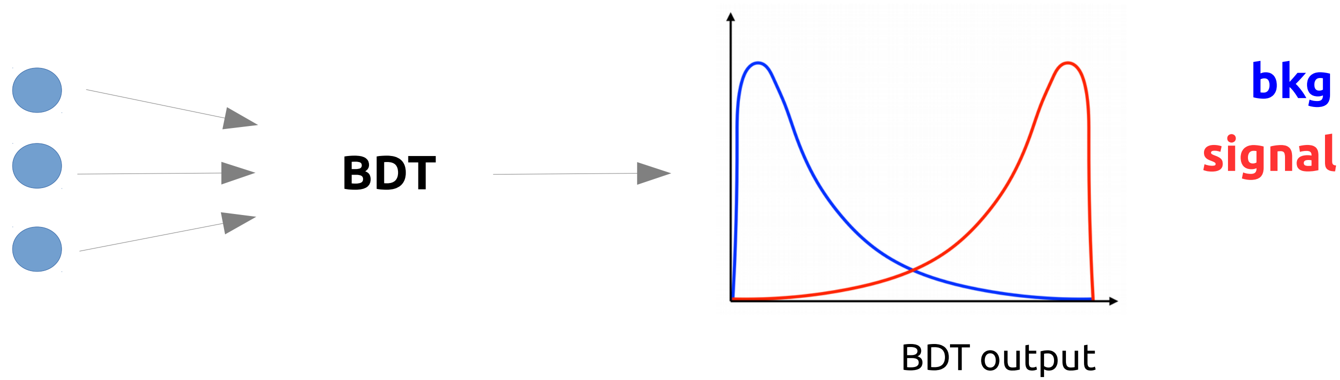
ee, μμ are dominated by Z boson production
To suppress Z : discard events with ($76 < m_{ll'} < 106$ GeV)

eμ – the cleanest channel to extract $t\bar{t}$

BDT discriminant

$t\bar{t} \rightarrow l\nu_l b\ l'\nu_{l'}\bar{b}$

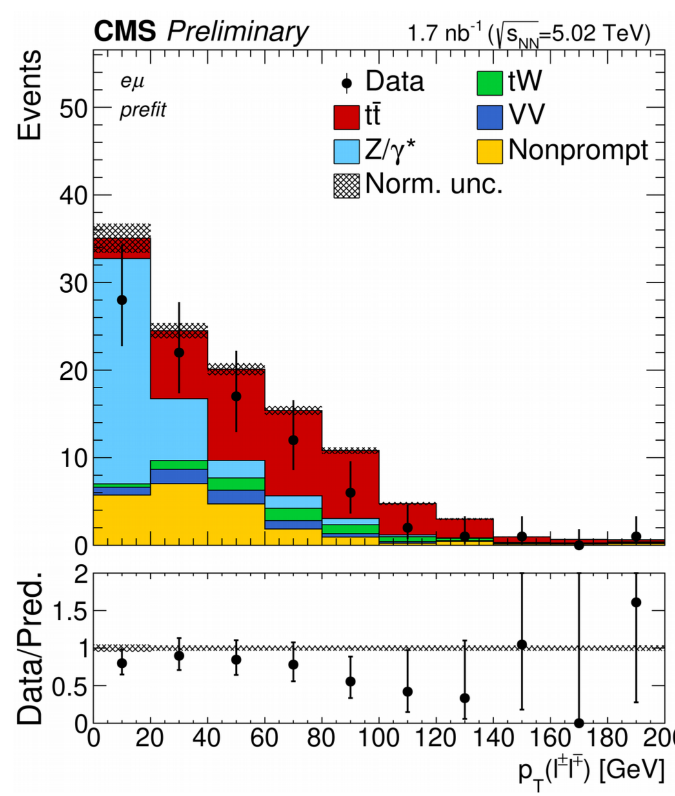
Boosted decision tree (BDT) algorithm - combines lepton information in one discriminant



Trained with signal = $t\bar{t}$ vs background = DY

Lepton information :

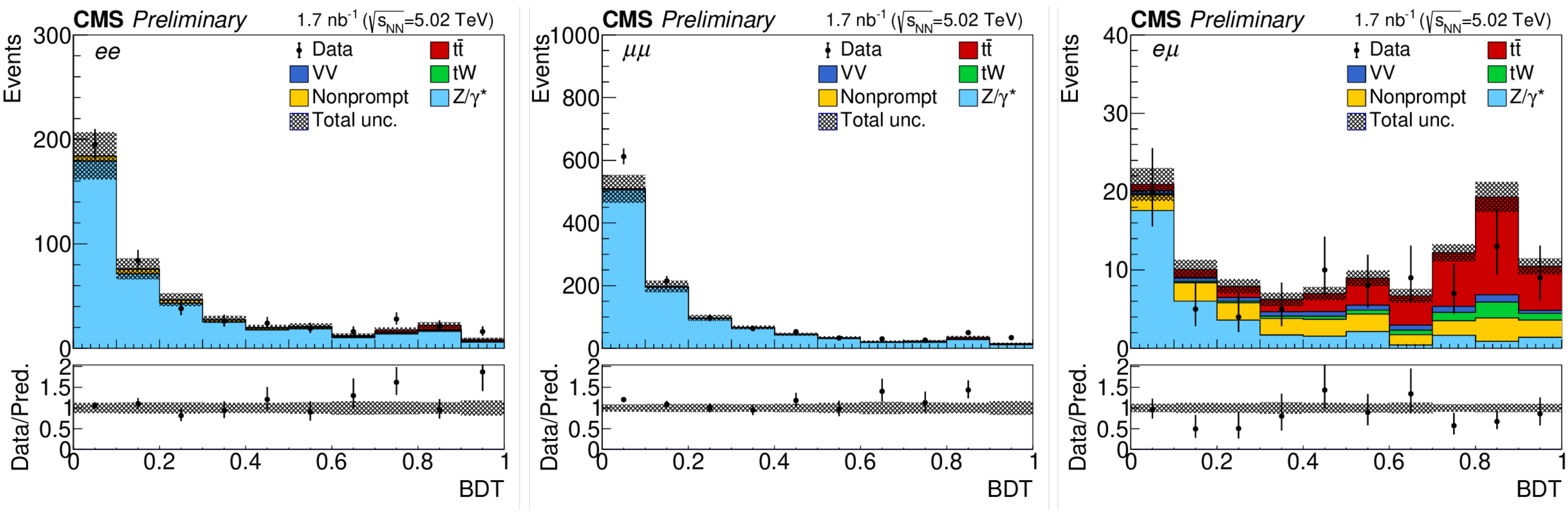
- leading lepton p_T
- momentum imbalance between leptons
- dilepton system p_T
- dilepton system $|\eta|$
- absolute azimuthal separation of the leptons
- scalar sum of the $|\eta|$ of the two leptons



BDT discriminant : pre-fit

$t\bar{t} \to l\nu_l b\ l'\nu_{l'}\bar{b}$

ee μμ eμ

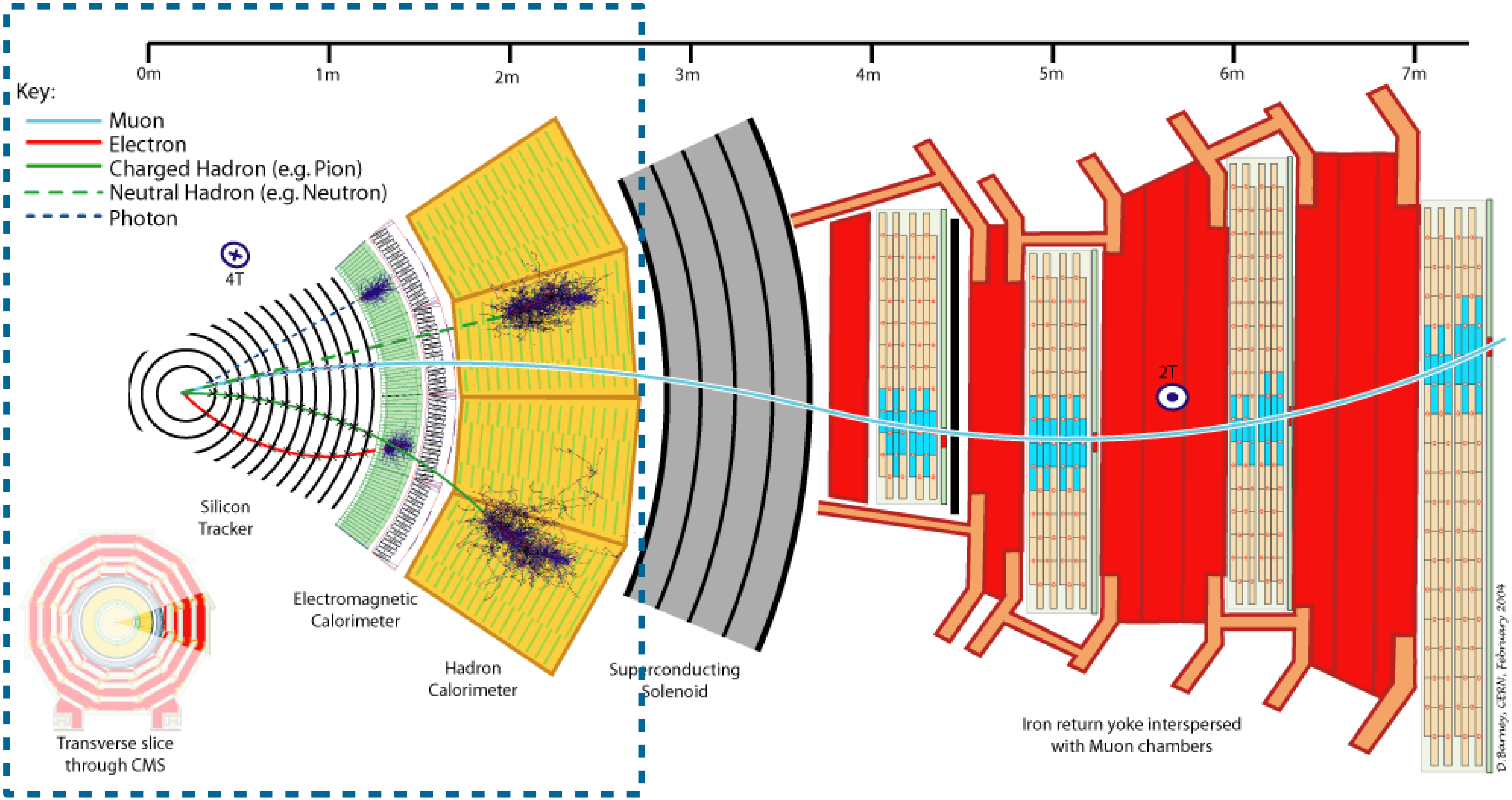


BDT discriminant : $t\bar{t}$ peaks at higher values of BDT ~ 0.8-1.0, DY peaks ~ 0

eμ channel : data points are lower than expectation

Jets in CMS

Jet reconstruction : combines tracker, ECAL and HCAL information

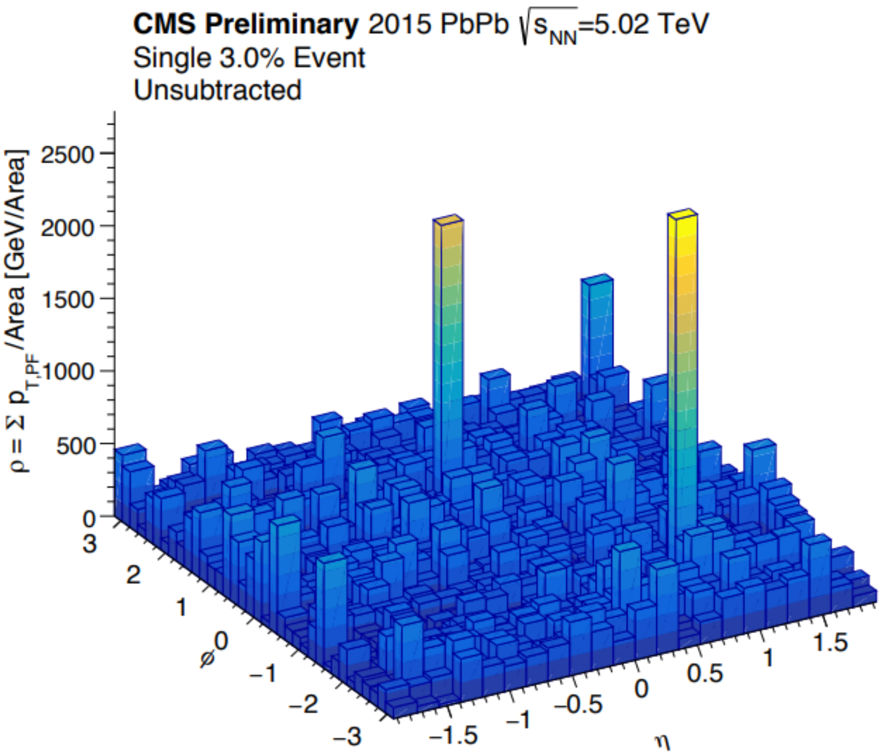


jets

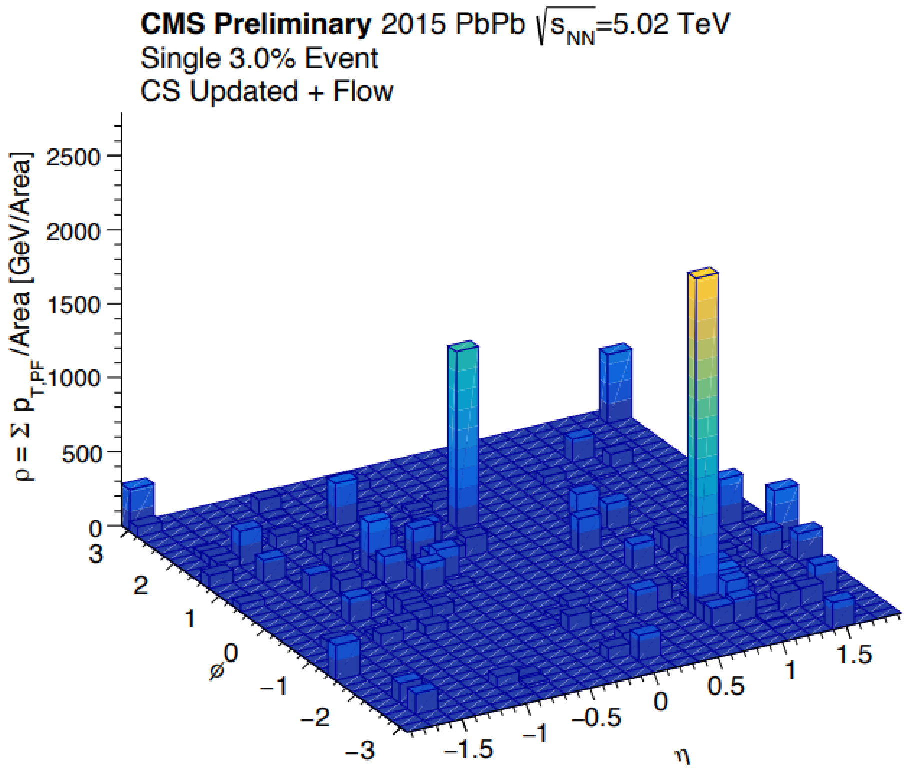
[JINST 12 (2017) P10003](#)

Jets in PbPb collisions

Before UE subtraction



After UE subtraction



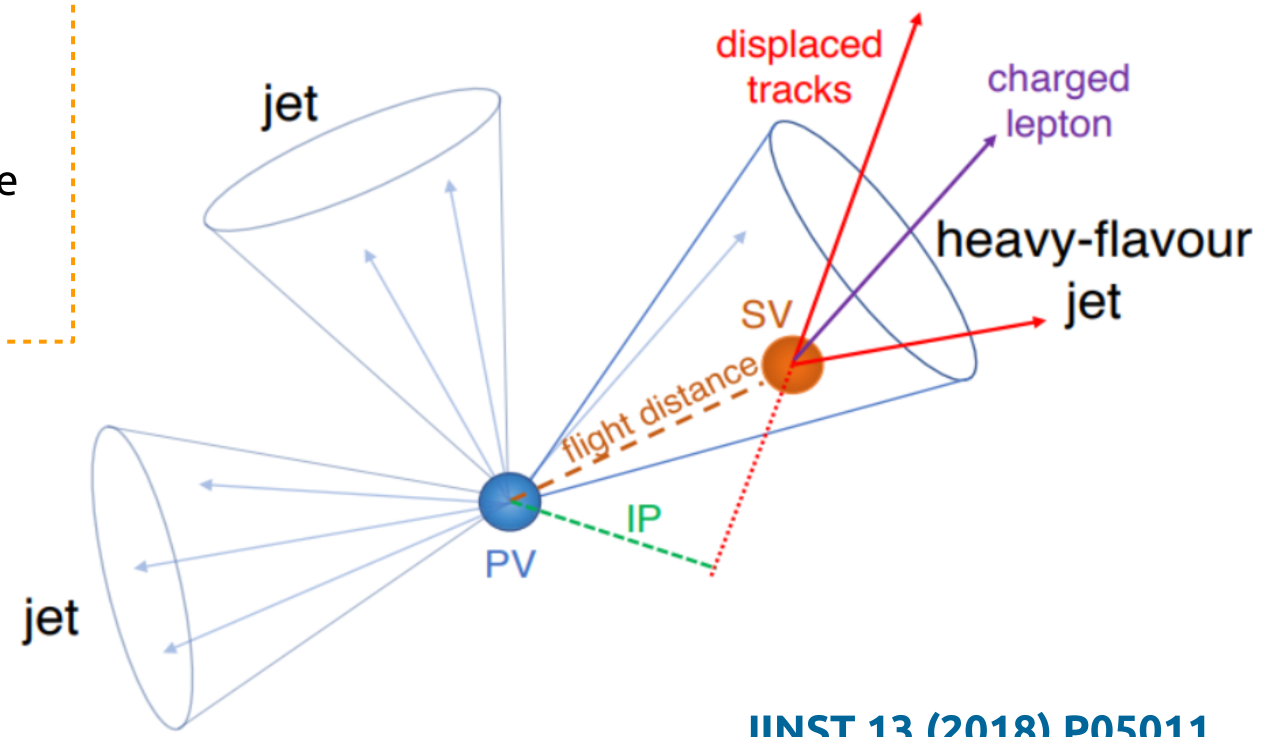
Particle-by-particle: correct the 4-momentum of a jet and substructure

[JHEP 1406 (2014) 092]

b-jet identification

b-hadrons

- Fragment hard, zb ~ 0.7 – 0.8
- Large decay multiplicity, $\langle n_{ch} \rangle \sim 5$
- Long-lived hadrons cτ ~ 500 μm → mm – cm displacement in lab frame
- Tend to decay semi-leptonically (20% for μ and e)



JINST 13 (2018) P05011

Method : exploit displaced vertices and tracks, both b-hadron and subsequent c-hadron decays

Method was re-optimized for PbPb environment
The working point was tuned to yield ~65%(5%) efficiency for b-(other-) jets

Lepton + b-jet analysis

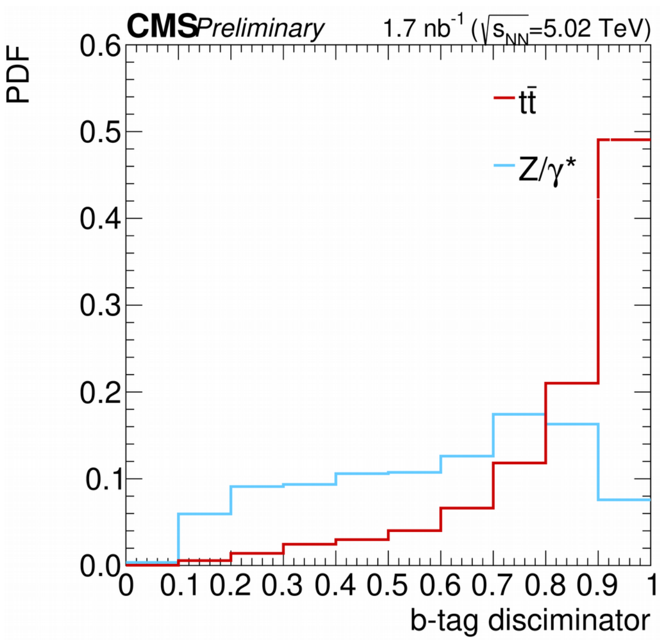
$t\bar{t} \rightarrow l\nu_l b\, l'\nu_{l'} \bar{b}$

Requiring b-jets in the analysis improves expected $t\bar{t}$ significance.

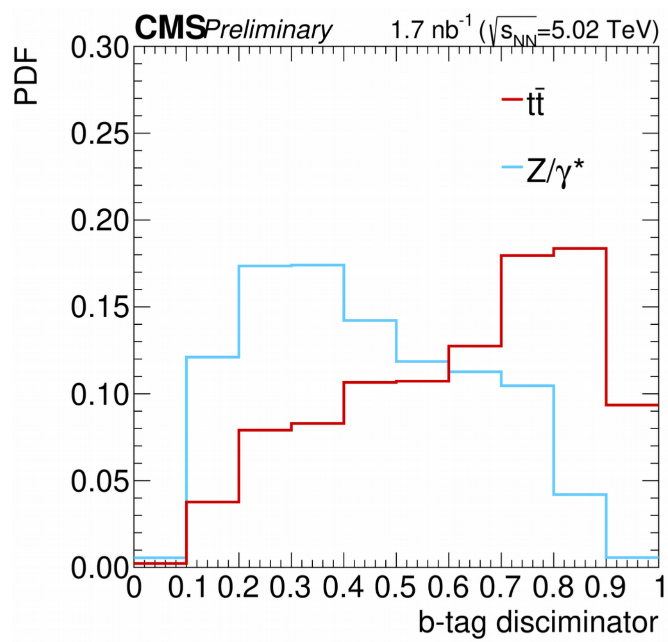
b-jet treatment :

- all jets with p_T > 30 GeV and $|\eta|$ < 2.0 are sorted by b-tag discriminant values
- two jets with the highest discriminant value are kept in the analysis

Leading jet



Sub-leading jet



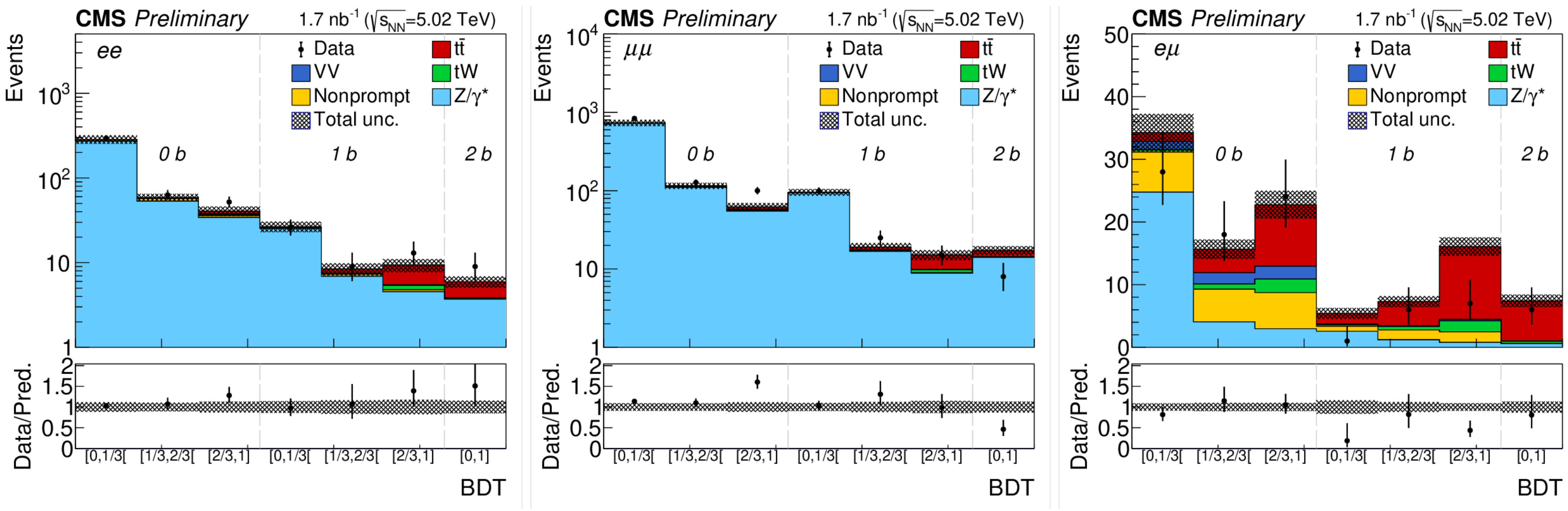
- count how many pass the b-tagging working point → categorize events in 0b, 1b,2b
- quenching: moves jets below p_T threshold leading to decreased probability of finding the b-jets → systematic uncertainty based on parametrization of the energy loss
  [Arleo, JHEP 0211 (2002) 044]

BDT pre-fit in lepton+bjet analysis

$t\bar{t} \rightarrow l\nu_l b\, l'\nu_{l'} \bar{b}$

BDT discriminant for 0b, 1b and 2b jets categories

ee μμ eμ



eμ channel : data points are lower than expectation
in the 1b and 2b categories, which have the highest S/B

BDT discriminant is an input for a statistical test : likelihood fit

Maximum likelihood method

Likelihood function: how theoretical assumption is compatible with observed data

Maximum likelihood method estimates the best values of the parameters to describe data

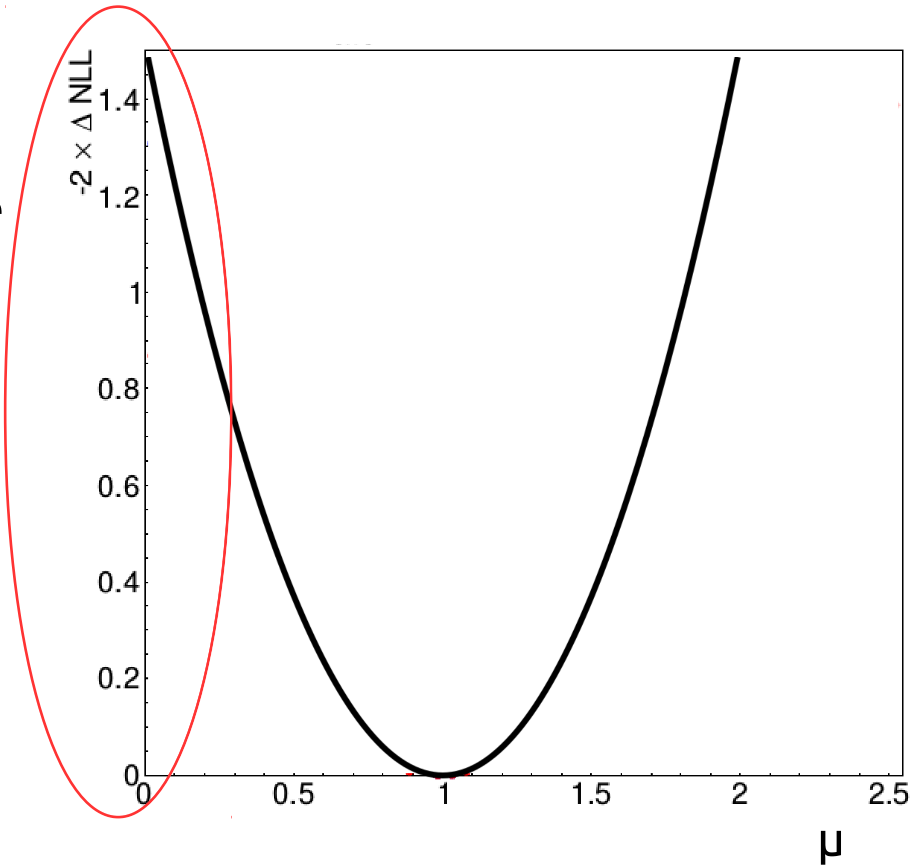
Parameters :

- Signal strength (μ)
- Signal contribution : yields and shape
- Background contribution : yields and shape
- Nuisance parameters : e.g. lumi, $p_T(Z)$, quenching, b-tag eff., lepton ID eff., …

Significance (σ) of an excess over the background-only expectation : ratio at $\mu = 0$

Expected : $\mu = 1$ and $\sigma = \sigma_{exp}$

Profile likelihood ratio example

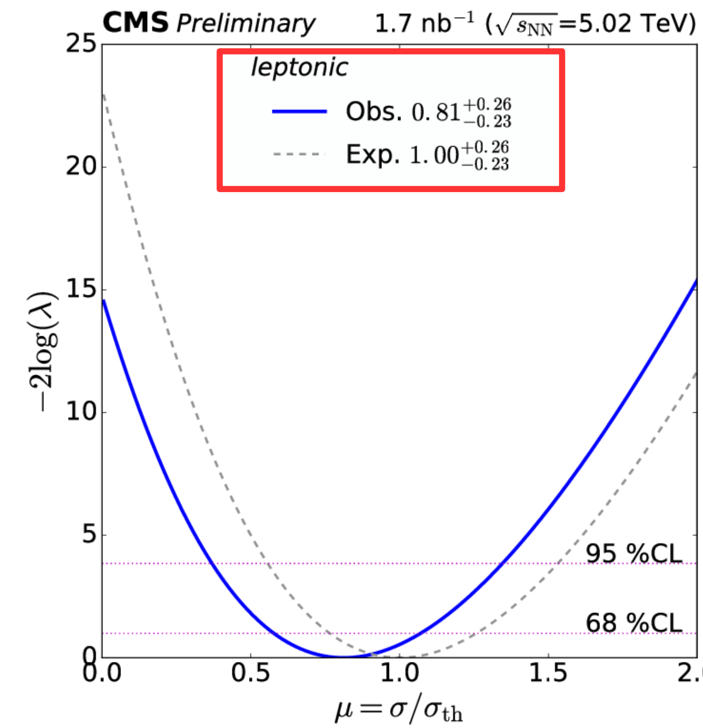


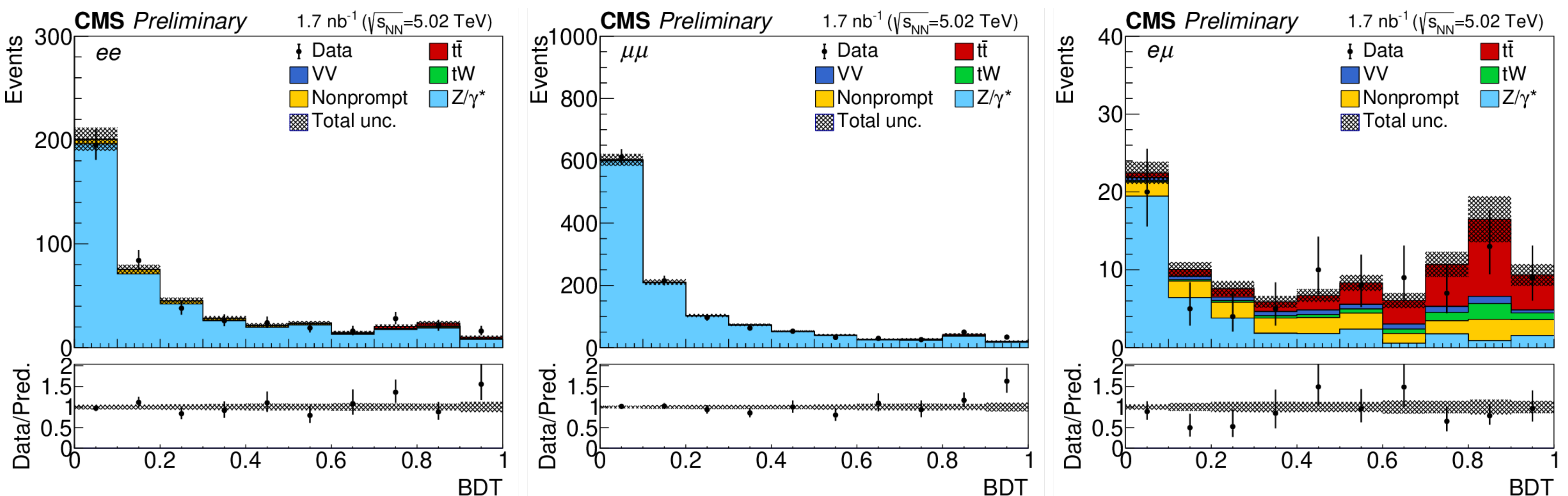
“Expected” includes nPDF effects, but not the jet quenching

Fit and results

$t\bar{t} \rightarrow l\nu_l b\, l'\nu_{l'}\bar{b}$

- Pre-fit deficits drive final μ=0.81±0.26
- Significance: 3.8σ(obs.), 4.8σ(exp.), 18% p-val
- Post-fit distributions in very good agreement

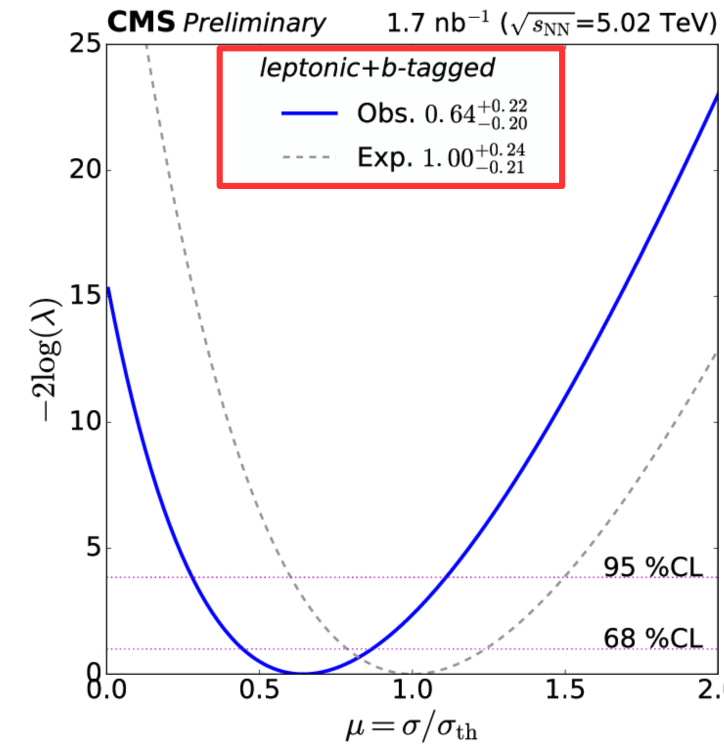


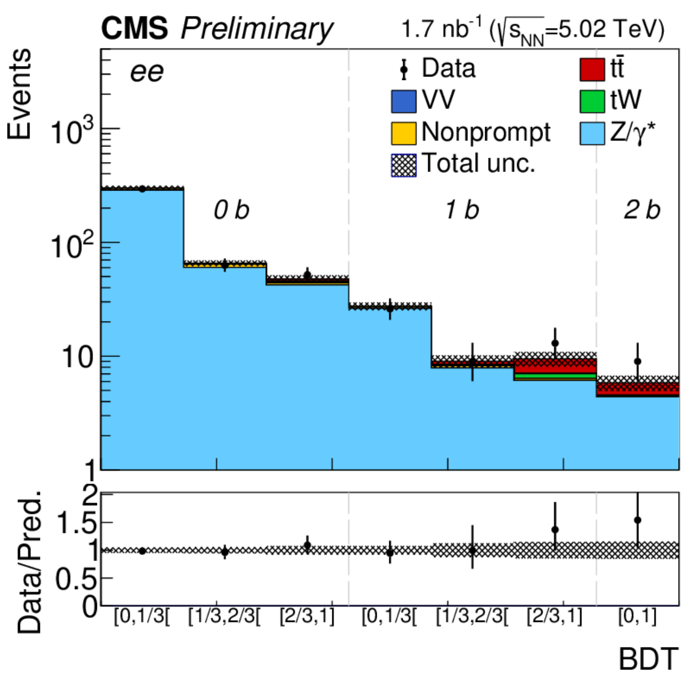


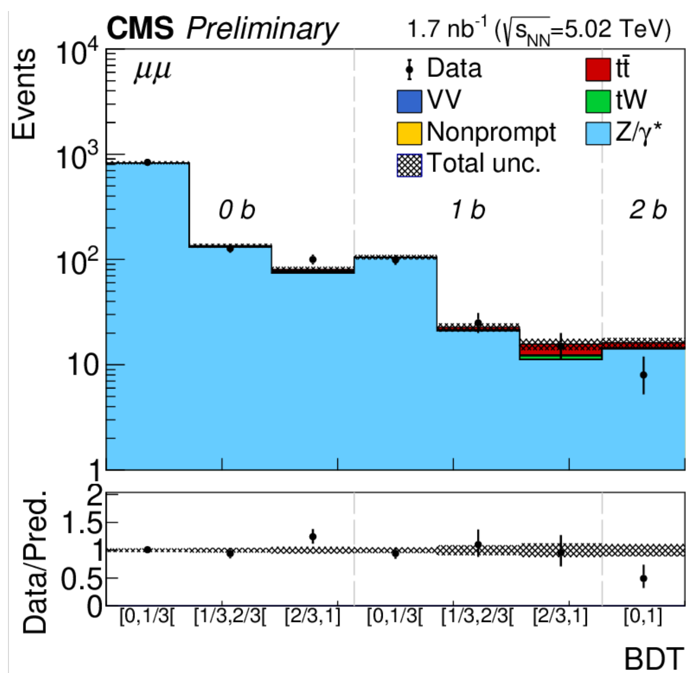
Fit and results

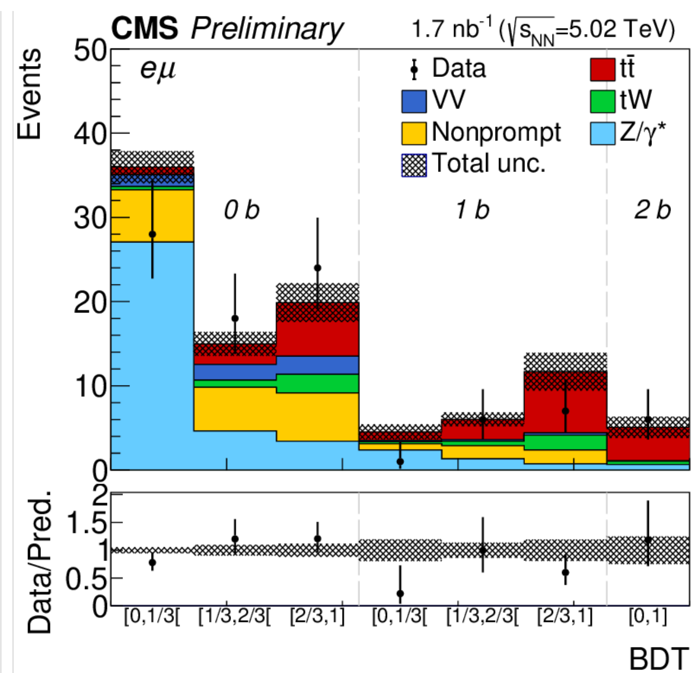
$t\bar{t} \rightarrow l\nu_l b\ l'\nu_{l'}\bar{b}$

- Deficit is slightly enhanced: μ=0.64±0.22
  - compatible with inclusive analysis
- Significance: 4.0σ(obs.), 5.8σ(exp.), 5% p-val
- Post-fit distributions in very good agreement









Event yields

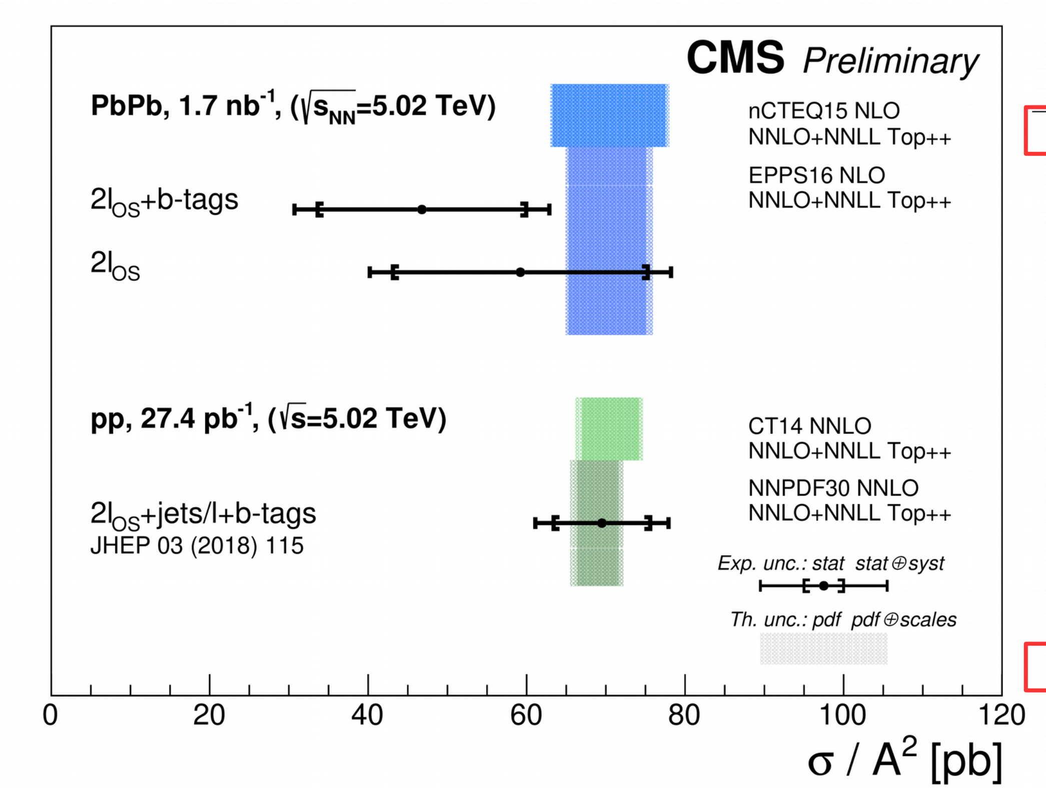
In total **1768** events were observed in the data :

→ **43 ± 11 $t\bar{t}$** events extracted from the likelihood fit

Process	Final state								
	e^+e^-			$\mu^+\mu^-$			$e^\pm\mu^\mp$		
	0b	1b	2b	0b	1b	2b	0b	1b	2b
Z/γ^*	389.8±15.4	40.4±2.7	4.4±0.8	1027.5±27.3	136.1±5.7	14.1±1.7	35.1±1.7	4.4±0.9	0.7±0.2
Nonprompt	17.3±2.2	1.4±0.2	$\leq$ 0.1	7.6±1.0	0.8±0.1	$\leq$ 0.1	17.1±1.9	4.0±0.4	$\leq$ 0.1
tW	1.1±0.2	0.9±0.2	$\leq$ 0.1	1.8±0.4	1.3±0.3	0.2±0.1	3.4±0.7	2.5±0.5	0.4±0.1
VV	1.9±0.3	0.2±0.1	$\leq$ 0.1	3.3±0.6	0.4±0.1	$\leq$ 0.1	5.4±0.9	0.6±0.1	$\leq$ 0.1
Total background	410.2±15.1	42.8±2.7	4.5±0.8	1040.2±27.1	138.6±5.7	14.4±1.8	61.1±2.9	11.5±1.3	1.1±0.2
$t\bar{t}$ signal	2.8±0.8	3.2±0.8	1.3±0.4	4.5±1.2	5.1±1.2	1.9±0.6	9.7±2.5	10.7±2.4	4.0±1.2
Observed (data)	410	48	9	1064	139	8	70	14	6

- **ee/μμ only matters in 1b and 2b categories**
- **best S/B in eμ channel**
- very high purity in 2b category

$t\bar{t}$ cross-section



Source	$\Delta\mu/\mu$ leptonic-only	$\Delta\mu/\mu$ leptonic+b-tagged
Total statistical uncertainty	0.27	0.28
Total systematic experimental uncertainty	0.17	0.19
Background normalization	0.12	0.12
Background and $t\bar{t}$ signal distribution	0.07	0.08
Lepton selection efficiency	0.06	0.06
Jet energy scale and resolution	—	0.02
btagging efficiency	—	0.06
Integrated luminosity	0.05	0.05
Total theoretical uncertainty	0.05	0.05
nPDF, μ_R, μ_F scales, and $\alpha_S(m_Z)$	<0.01	<0.01
Top quark and Z boson p_T modeling	0.05	0.05
Top quark mass	<0.01	<0.01
Total uncertainty	0.32	0.34

Two analyses yield consistent cross-sections

Statistical uncertainty dominates !

Summary

- CMS provide a strong evidence of the top quark production in PbPb collisions :

  $\mu = 0.81 \pm 0.25$; 3.8σ (4.8σ exp.) - lepton only analysis,
  $\mu = 0.64 \pm 0.21$; 4.0σ (6.0σ exp.) - lepton +b-jet analysis

- $t\bar{t}$ production cross-section in PbPb collisions :

  2.02 ± 0.69 μb - lepton only analysis,
  2.56 ± 0.82 μb - lepton + b-jet analysis

- The results are compatible between 2 analysis, as well as with expectations from scaled pp cross-section and QCD computations

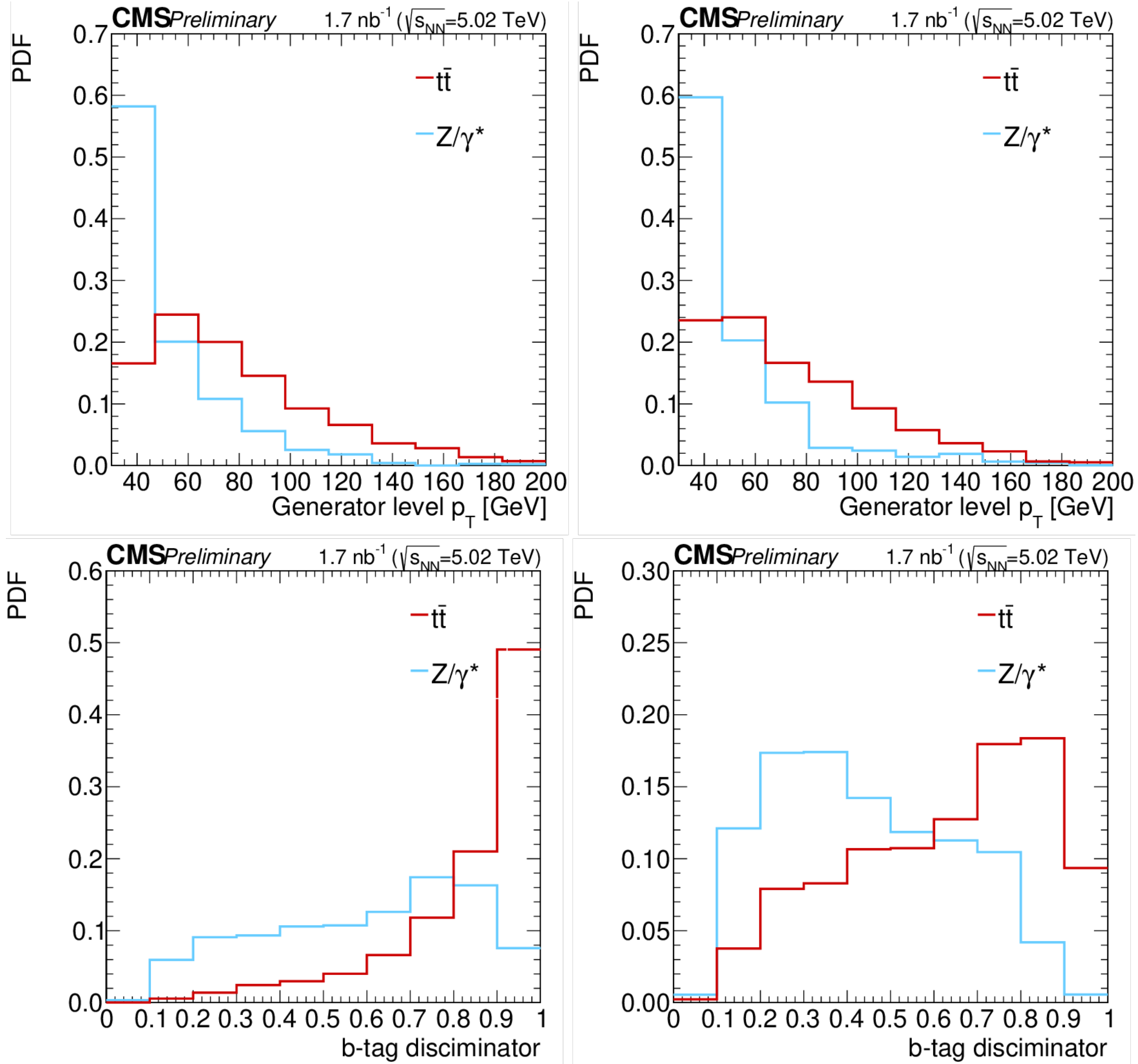
- First step towards the top quark as a tool to probe the QGP evolution

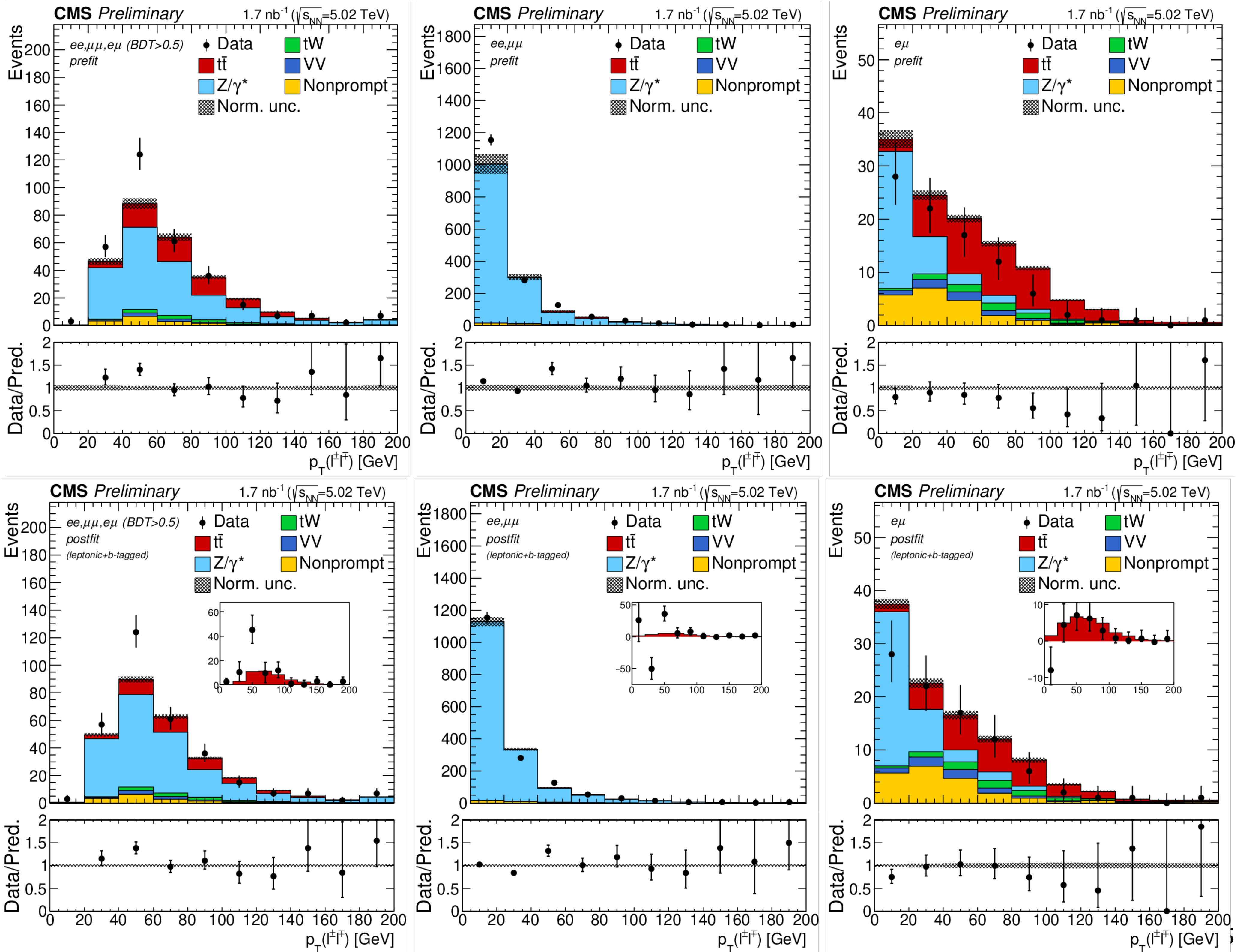
CMS-PAS-HIN-19-001
aiming "Nature Physics"

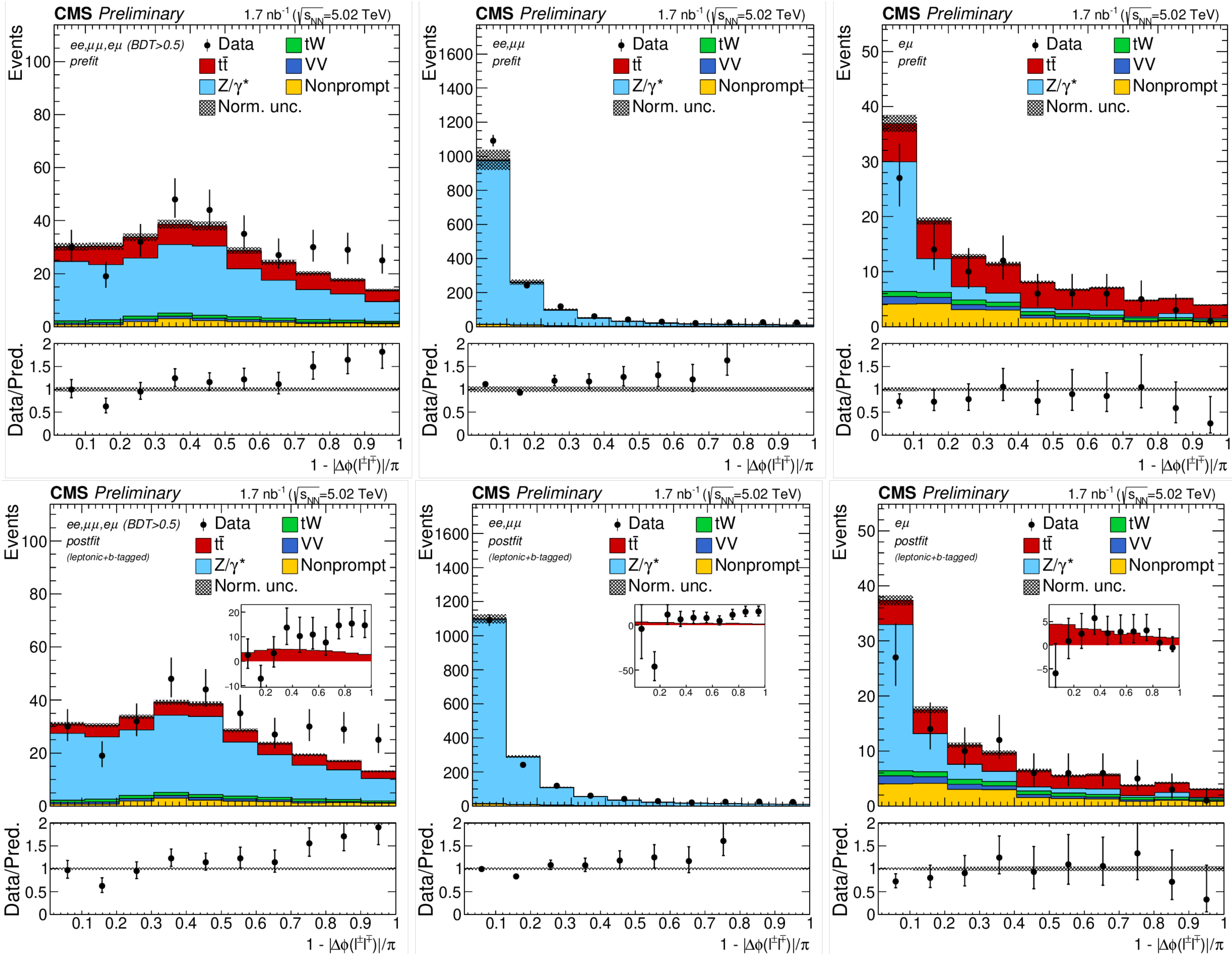
Backup slides

Signal strength and significance per channel

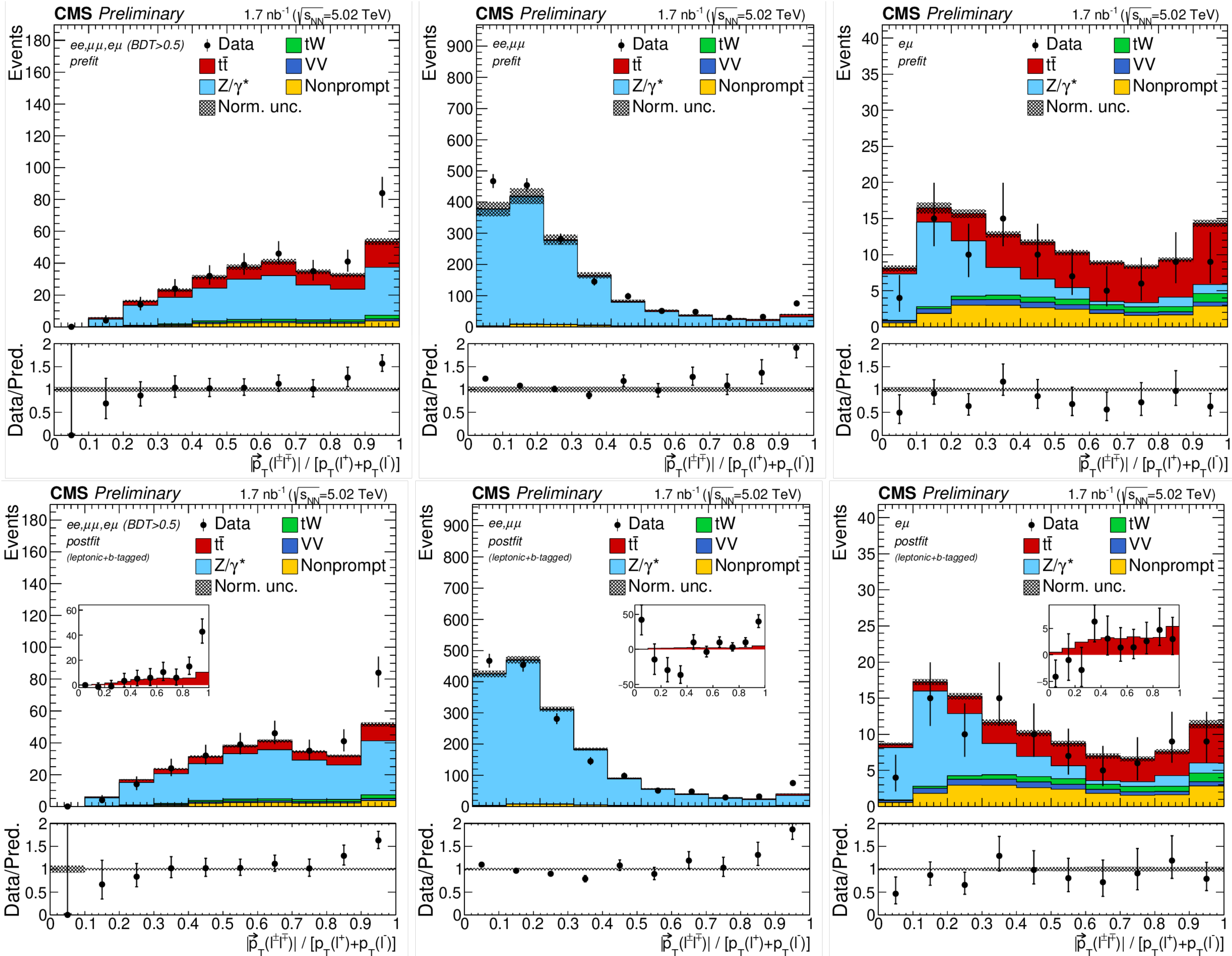
Fit alternative	Signal strength μ	Significance
$e^{\pm}\mu^{\mp}$ (leptonic only)	$0.66^{+0.24}_{-0.22}$ $(1.00^{+0.27}_{-0.25})$	3.3 (4.7)
$e^{+}e^{-}$, $\mu^{+}\mu^{-}$, and $e^{\pm}\mu^{\mp}$ (leptonic only)	$0.81^{+0.26}_{-0.23}$ $(1.00^{+0.26}_{-0.23})$	3.8 (4.8)
$e^{\pm}\mu^{\mp}$ (leptonic+b-tagged)	$0.61^{+0.23}_{-0.20}$ $(1.00^{+0.26}_{-0.23})$	3.8 (5.3)
$e^{+}e^{-}$, $\mu^{+}\mu^{-}$, and $e^{\pm}\mu^{\mp}$ (leptonic+b-tagged)	$0.64^{+0.22}_{-0.20}$ $(1.00^{+0.24}_{-0.21})$	4.0 (6.0)

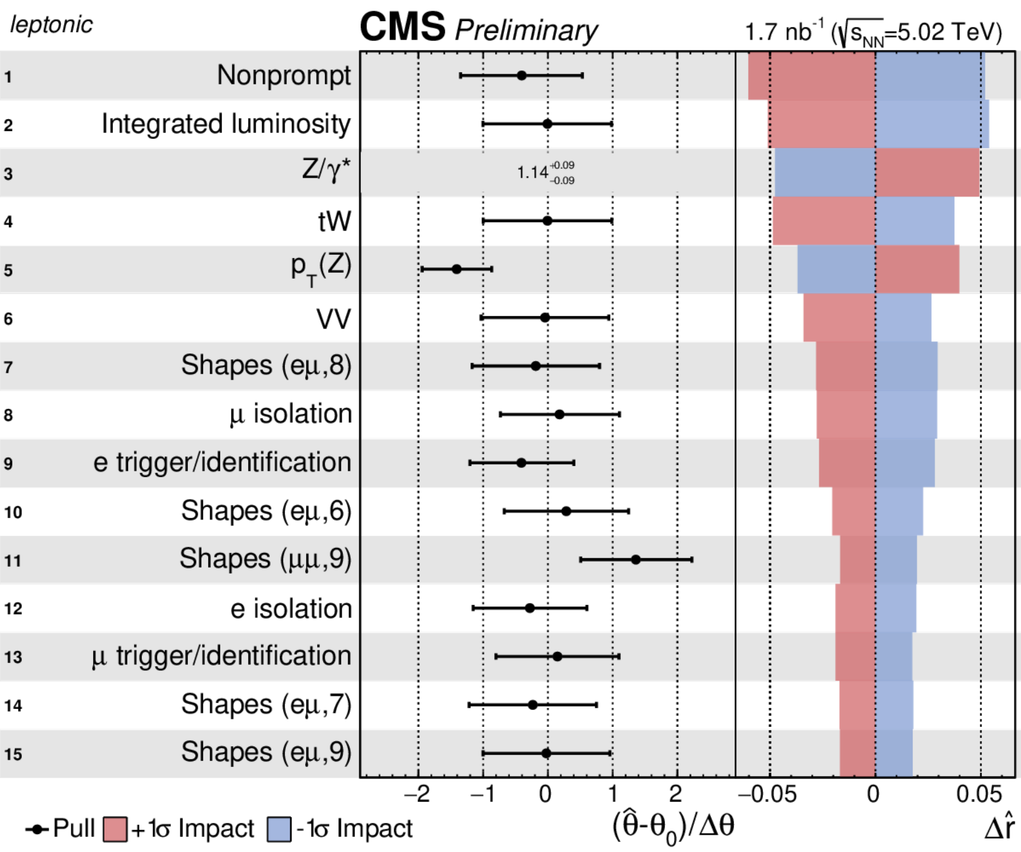

CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
PDF
$t\bar{t}$
Z/γ^*
Generator level p_T [GeV]
CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
PDF
$t\bar{t}$
Z/γ^*
Generator level p_T [GeV]
CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
PDF
$t\bar{t}$
Z/γ^*
b-tag discriminator
CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
PDF
$t\bar{t}$
Z/γ^*
b-tag discriminator

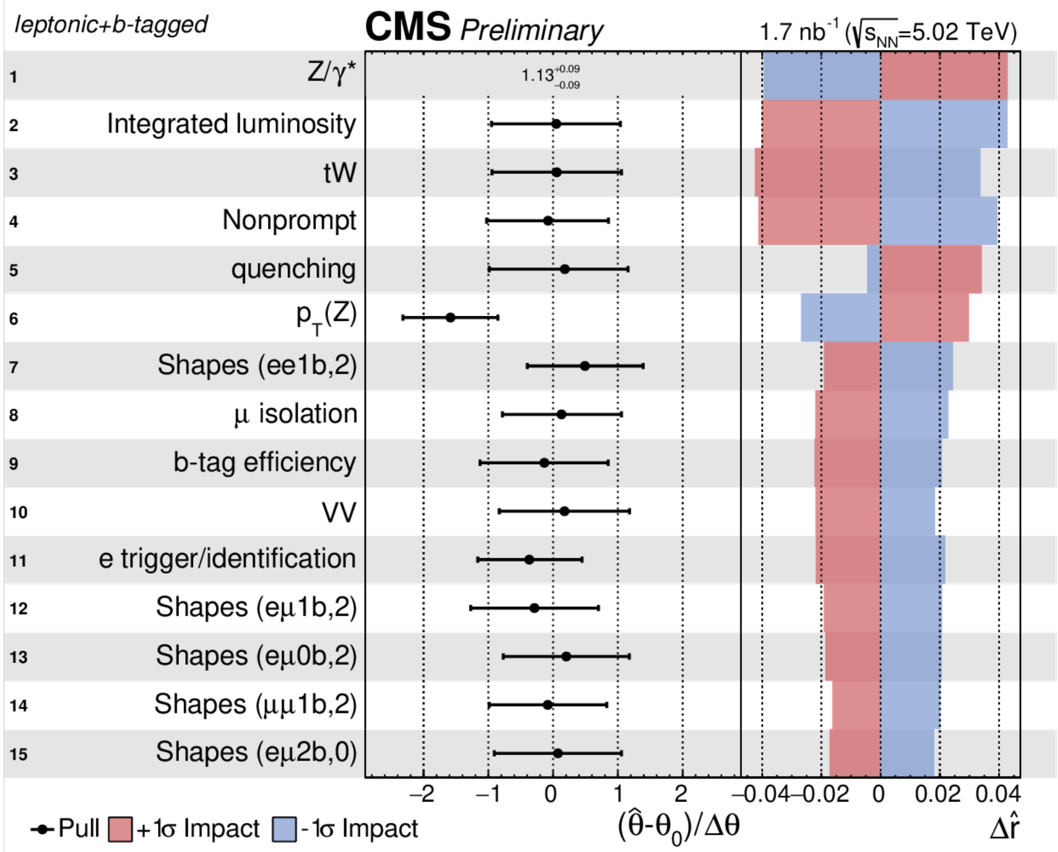

CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
ee,μμ,eμ (BDT>0.5)
prefit
Data
tW
$t\bar{t}$
VV
Z/γ*
Nonprompt
Norm. unc.
Events
Data/Pred.
$p_T(l^{\pm}l^{\mp})$ [GeV]
CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
ee,μμ
prefit
CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
eμ
prefit
CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
ee,μμ,eμ (BDT>0.5)
postfit
(leptonic+b-tagged)
CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
ee,μμ
postfit
(leptonic+b-tagged)
CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
eμ
postfit
(leptonic+b-tagged)

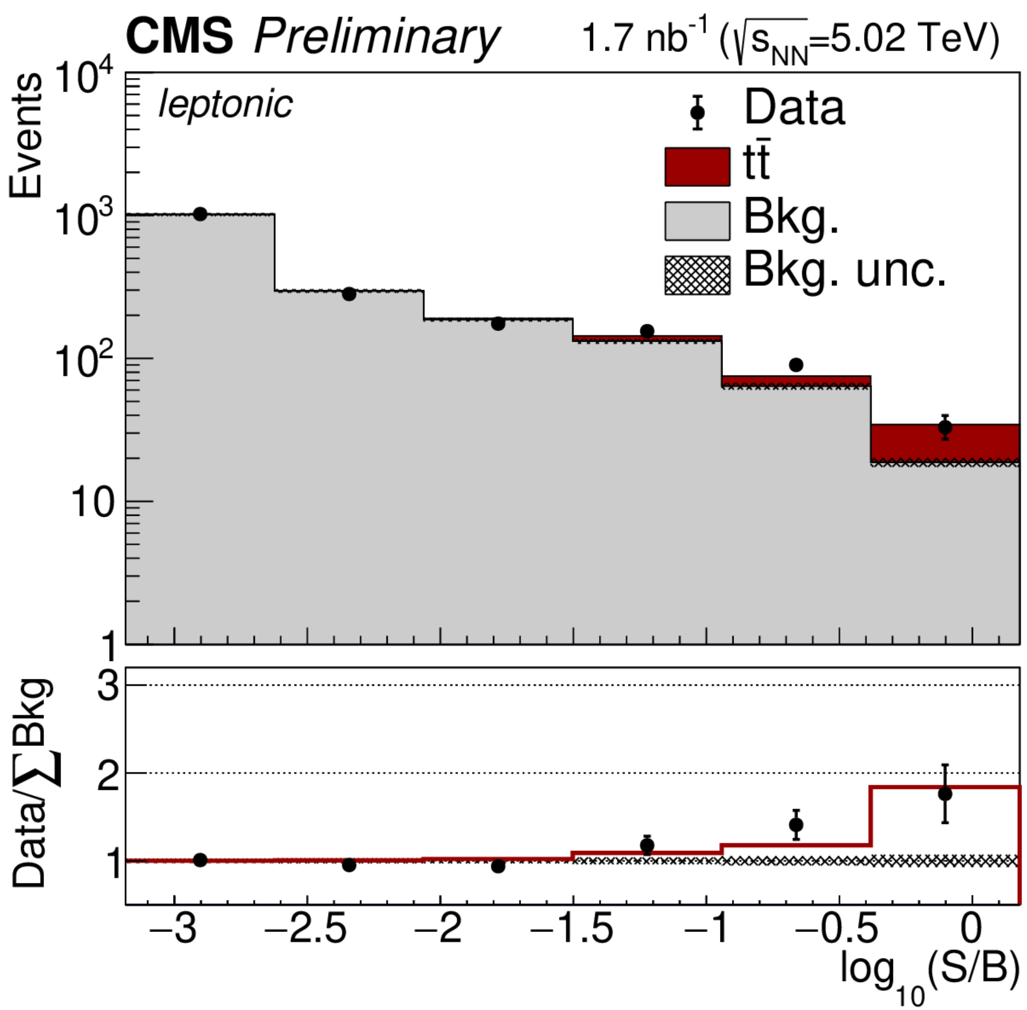


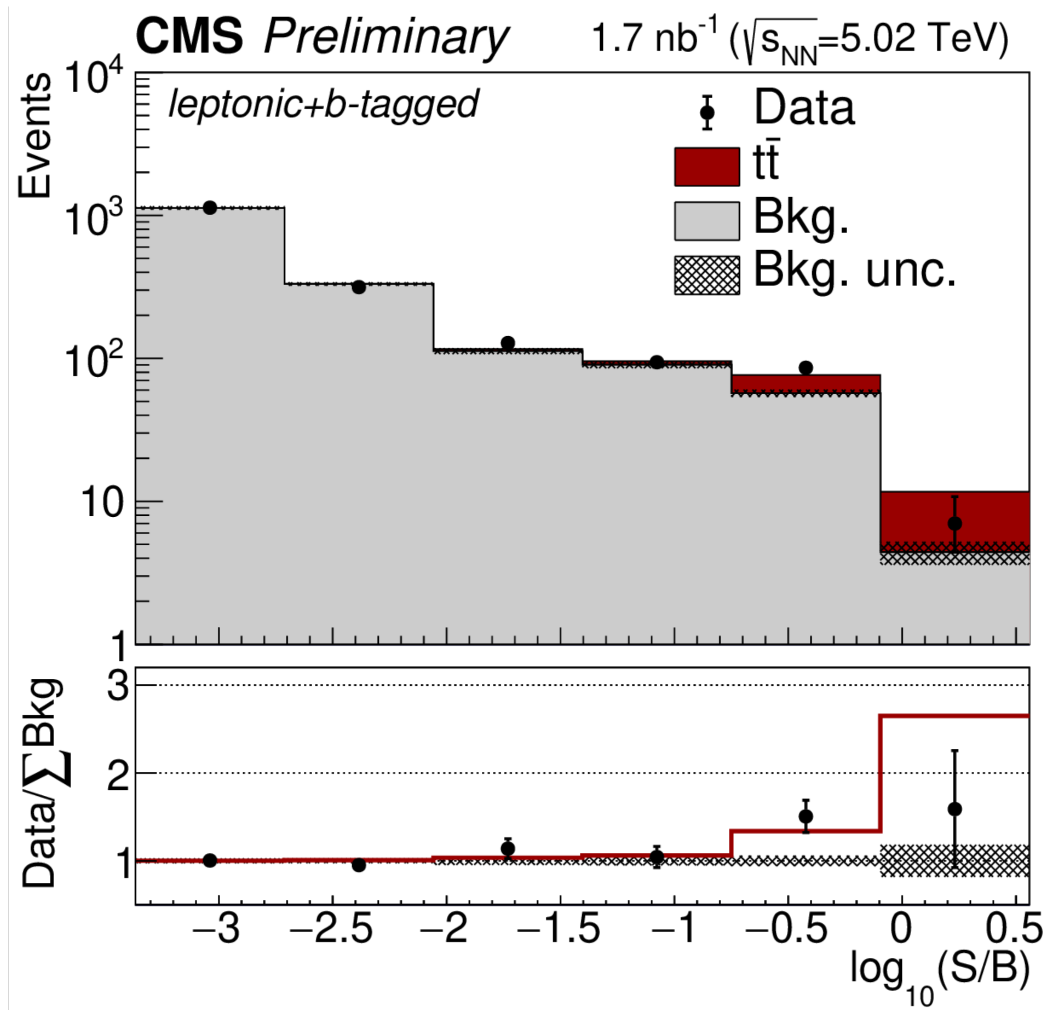
CMS Preliminary
1.7 nb^-1 (√sNN=5.02 TeV)
ee,μμ,eμ (BDT>0.5)
prefit
postfit
(leptonic+b-tagged)
ee,μμ
eμ
Data
tt̄
Z/γ*
Norm. unc.
tW
VV
Nonprompt
Events
Data/Pred.
1 - |Δφ(l±l∓)|/π


CMS Preliminary
1.7 nb^{-1} ($\sqrt{s_{NN}}$=5.02 TeV)
ee,μμ,eμ (BDT>0.5)
prefit
ee,μμ
eμ
postfit
(leptonic+b-tagged)
Data
tW
$t\bar{t}$
VV
Z/γ*
Nonprompt
Norm. unc.
Events
Data/Pred.
$|\vec{p}_T(l^{\pm}l^{\mp})| / [p_T(l^+)+p_T(l^-)]$


leptonic
CMS Preliminary
1.7 nb⁻¹ (√sNN=5.02 TeV)
1 Nonprompt
2 Integrated luminosity
3 Z/γ*
1.14 +0.09 −0.09
4 tW
5 pT(Z)
6 VV
7 Shapes (eμ,8)
8 μ isolation
9 e trigger/identification
10 Shapes (eμ,6)
11 Shapes (μμ,9)
12 e isolation
13 μ trigger/identification
14 Shapes (eμ,7)
15 Shapes (eμ,9)
−2 −1 0 1 2
−0.05 0 0.05
Pull +1σ Impact -1σ Impact
(θ̂-θ0)/Δθ
Δr̂


leptonic+b-tagged
CMS Preliminary
1.7 nb⁻¹ (√sNN=5.02 TeV)
1 Z/γ*
1.13 +0.09 −0.09
2 Integrated luminosity
3 tW
4 Nonprompt
5 quenching
6 pT(Z)
7 Shapes (ee1b,2)
8 μ isolation
9 b-tag efficiency
10 VV
11 e trigger/identification
12 Shapes (eμ1b,2)
13 Shapes (eμ0b,2)
14 Shapes (μμ1b,2)
15 Shapes (eμ2b,0)
−2 −1 0 1 2
−0.04 −0.02 0 0.02 0.04
Pull +1σ Impact -1σ Impact
(θ̂-θ0)/Δθ
Δr̂


CMS Preliminary
1.7 nb-1 (√sNN=5.02 TeV)
leptonic
Data
tt̄
Bkg.
Bkg. unc.
Events
Data/∑Bkg
log10(S/B)


CMS Preliminary
1.7 nb-1 (√sNN=5.02 TeV)
leptonic+b-tagged
Data
tt̄
Bkg.
Bkg. unc.
Events
Data/∑Bkg
log10(S/B)

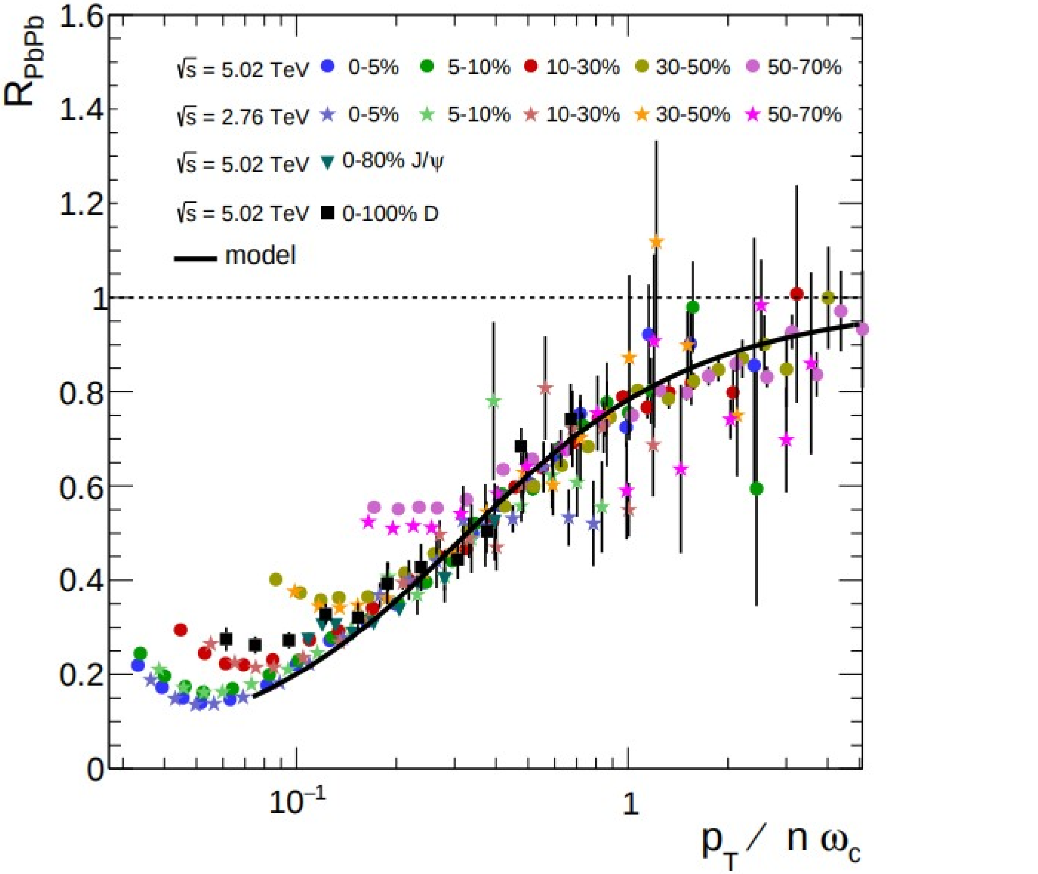
Analytical parametrization of the energy loss

[Arleo, PRL. 119, 062302]

- R_{AA} based fits to different spectra
- scaling behavior f(pT/wc) from →

- data indicates universal high-p_T behavior
- use to parametrize mean constituent energy loss (1-7 GeV depending on the centrality)
- use estimate to dampen jet energy in MC

- quenching effect around 7%



Maximum likelihood method

Likelihood function: how theoretical assumption is compatible with observed data

Maximum likelihood method estimates the values of the parameters

$$\mathcal{L}(\mu, \boldsymbol{\theta}) = \prod_{i=1}^{M} \frac{(\mu s_i + b_i)}{n_i!} e^{-(\mu s_i + b_i)}$$

- Signal strength = μ
- Signal contribution = s_i (according to a nominal model)
- Background contribution = b_i
- Nuisance parameters = θ

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}$$

$$\lambda(\mu = 0) = \frac{\mathcal{L}(0, \hat{\hat{\boldsymbol{\theta}}})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}})}$$

- $\hat{\hat{\boldsymbol{\theta}}}$ is the value of θ maximizing $\mathcal{L}$ for a certain μ
- $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ correspond to the true global maximum likelihood
- the profile likelihood ratio (λ) :

  → λ ~ 1 – data is compatible with signal expectation
  → λ ~ 0 – data is compatible with background-only expectation

$$s = \sqrt{-2Log\lambda(\mu = 0)}$$ - significance of an excess over the background-only expectation

Non-prompt bkg estimation

event mixing :

- same flavor combination mixed
- pick 100 events, exclude the same event
- exclude repetitions
- each combination gets a distance assigned
- distance based on kNN algorithm using : centrality, rho, iso, p_{Tl}, p_{Tll}

use set of closest events in this distance as central shape:

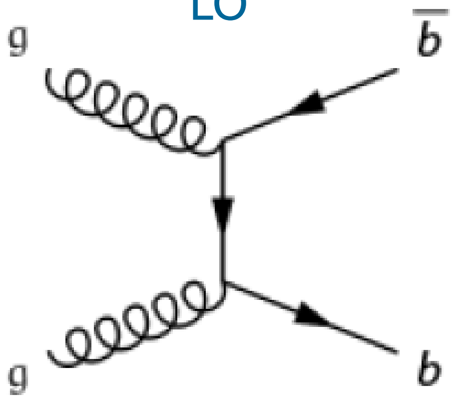
- furthest distance as alternative shape for systematic treatment

normalize the distribution to the same-sign data sample yield

b-jet production channels at LHC

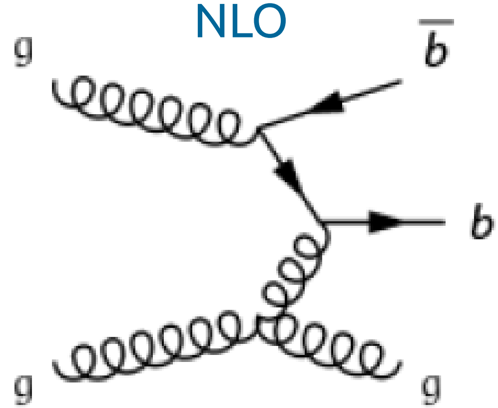
Flavor Creation (FCR)

LO



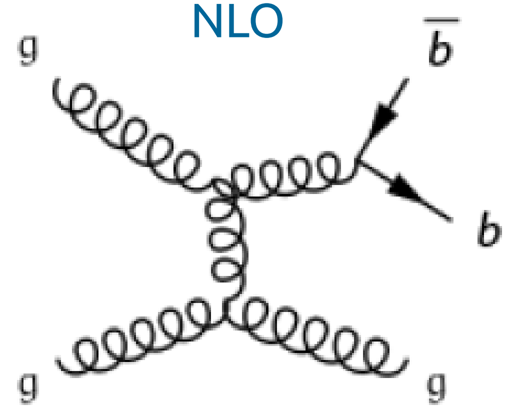
Flavor Excitation (FEX)

NLO



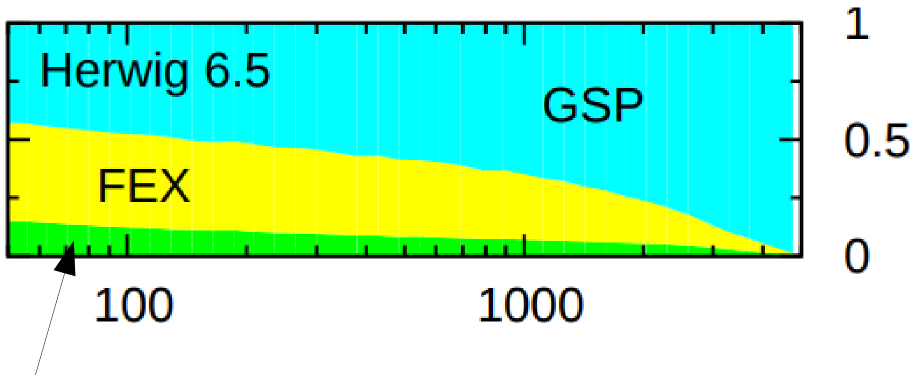
Gluon Splitting (GSP)

NLO



LHC, pp collisions at 14 TeV

JHEP 0707:026,2007



FCR is not dominant process

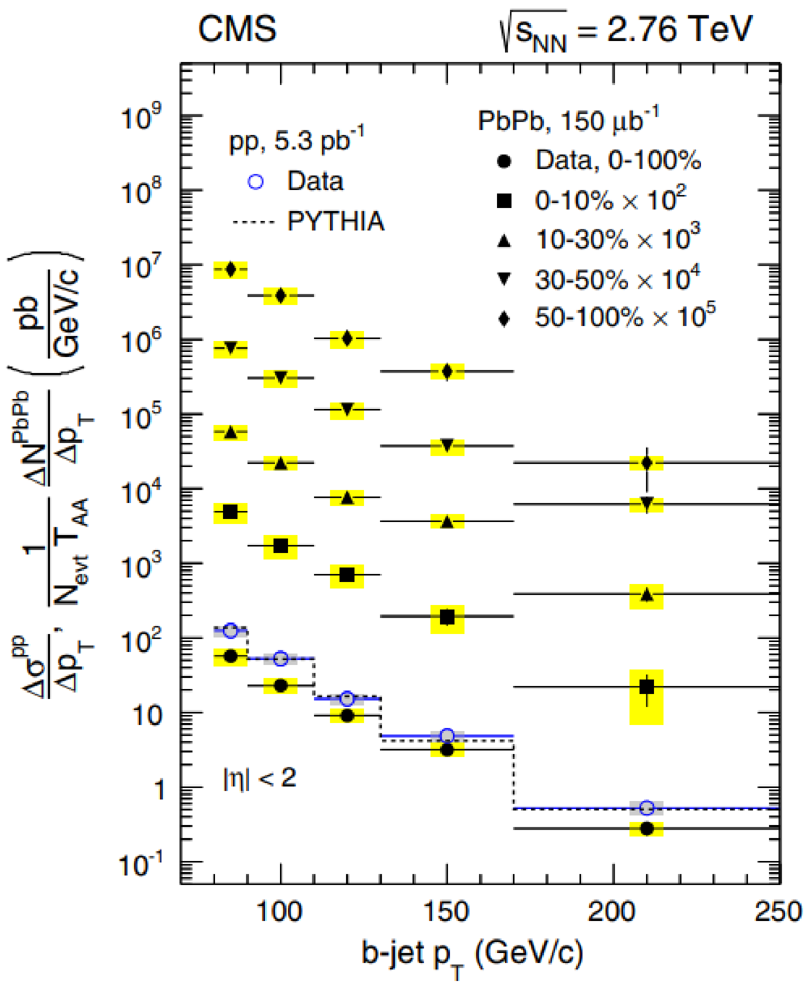
First Heavy Ion measurements convolute large contributions from NLO b-quark production processes

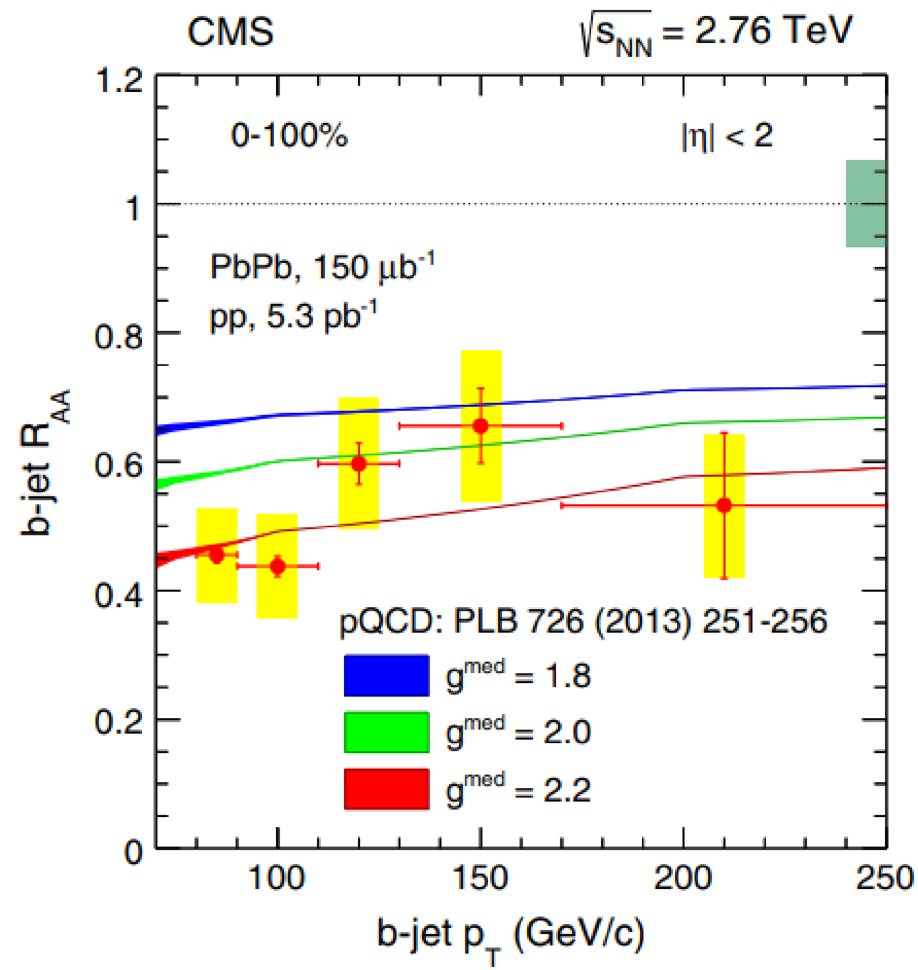
Energy loss is expected to depend on flavor
→ measure heavy flavor jets suppression

Quenching of b-jets

Jet spectra corrected for detector resolution effects for several centrality selections and pp

CMS, PRL 113 (2014) 132301



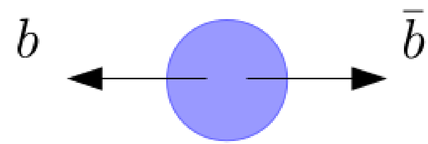


Suppression consistent with the one observed from inclusive jets

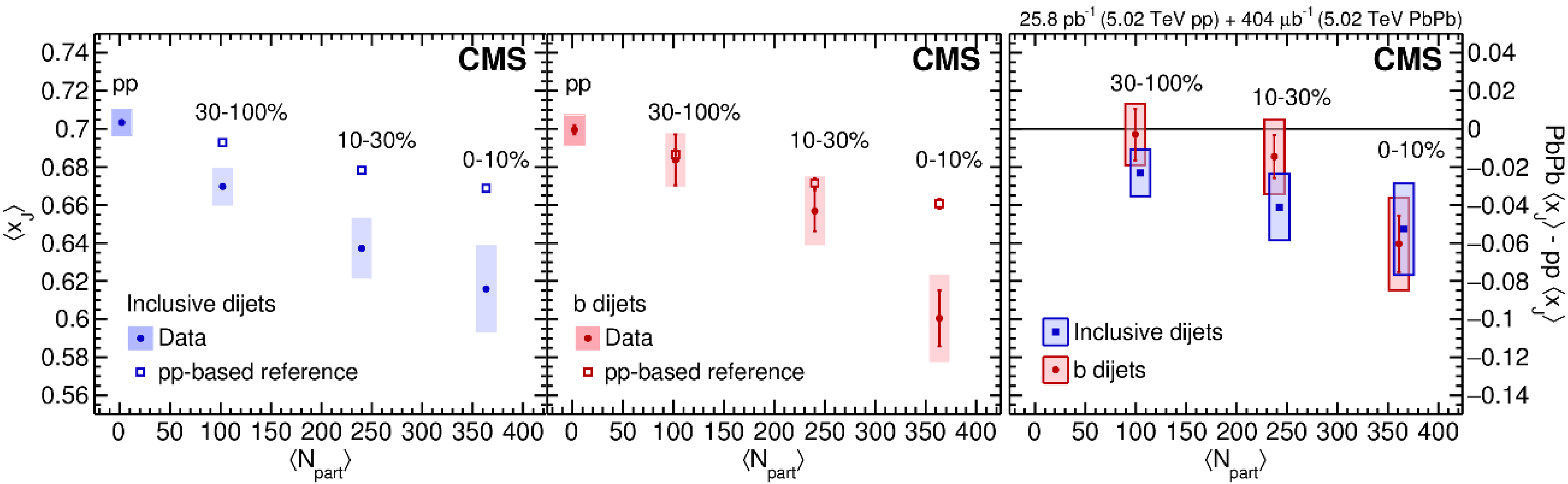
bb correlations

To suppress the contribution of gluon splitting and probe LO b-jet production :
look at pairs of b jets that are back-to-back in azimuth.

$$x_J = \frac{p_{T,2}}{p_{T,1}}$$



CMS , JHEP 1803 (2018) 181



No clear difference between pT balance of inclusive and b-dijets

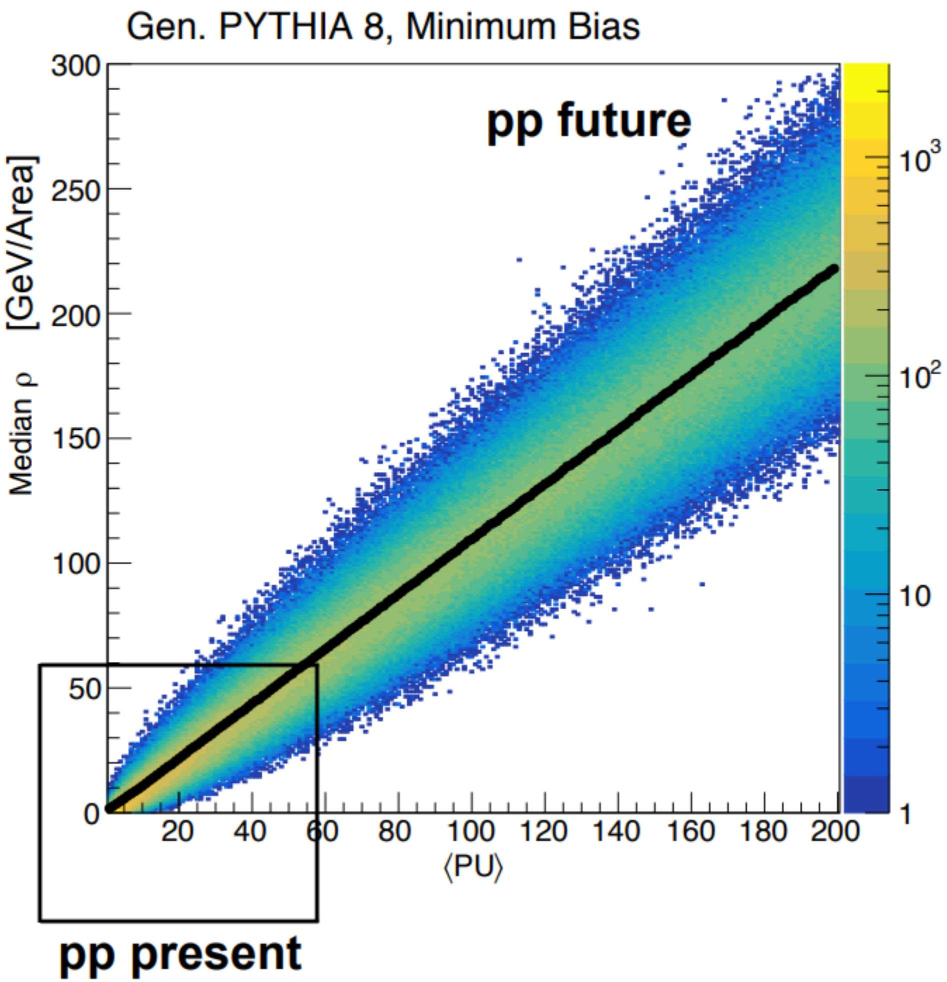
Data from Run 3 will allow to make a conclusive statement

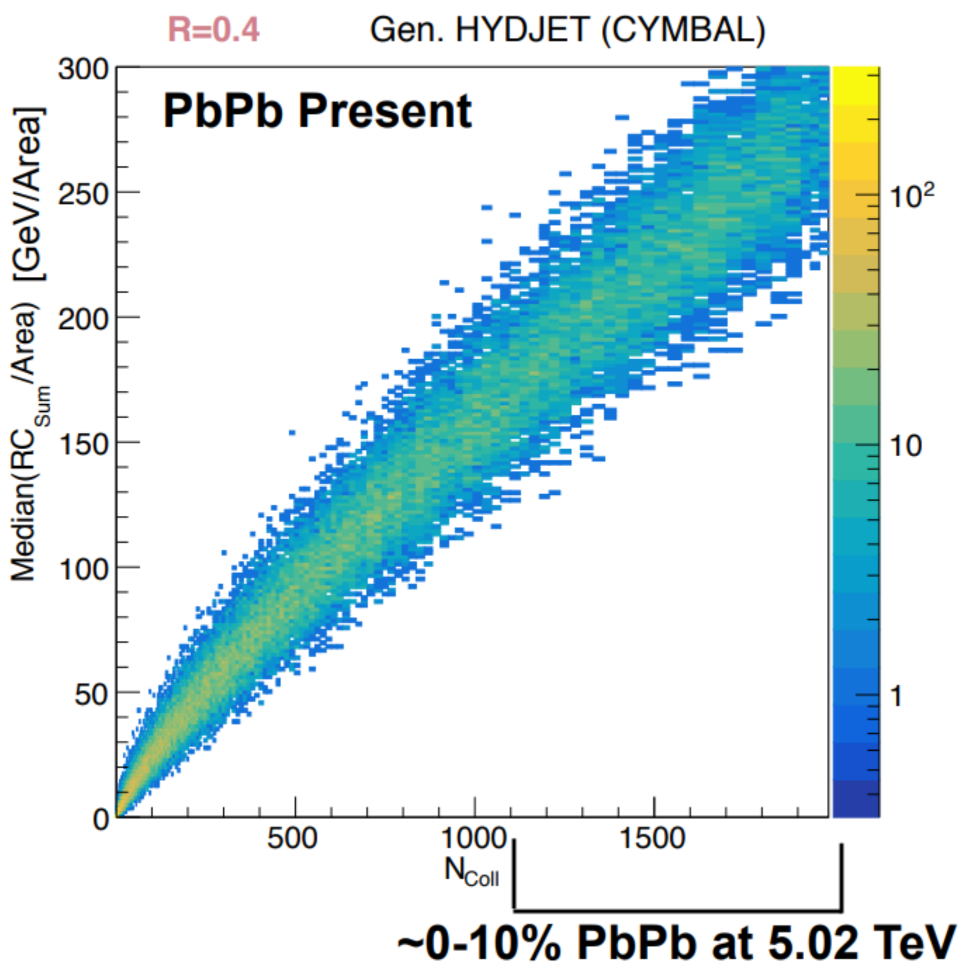
Underlying event in pp and PbPb collisions

Underlying Event (UE) - particles not associated with the hardest parton-parton process quantified as transverse momentum density (ρ)

PileUp (PU) – concurrent interactions coming from the same bunch crossing

BNL jet workshop '18, C.McGinn talk

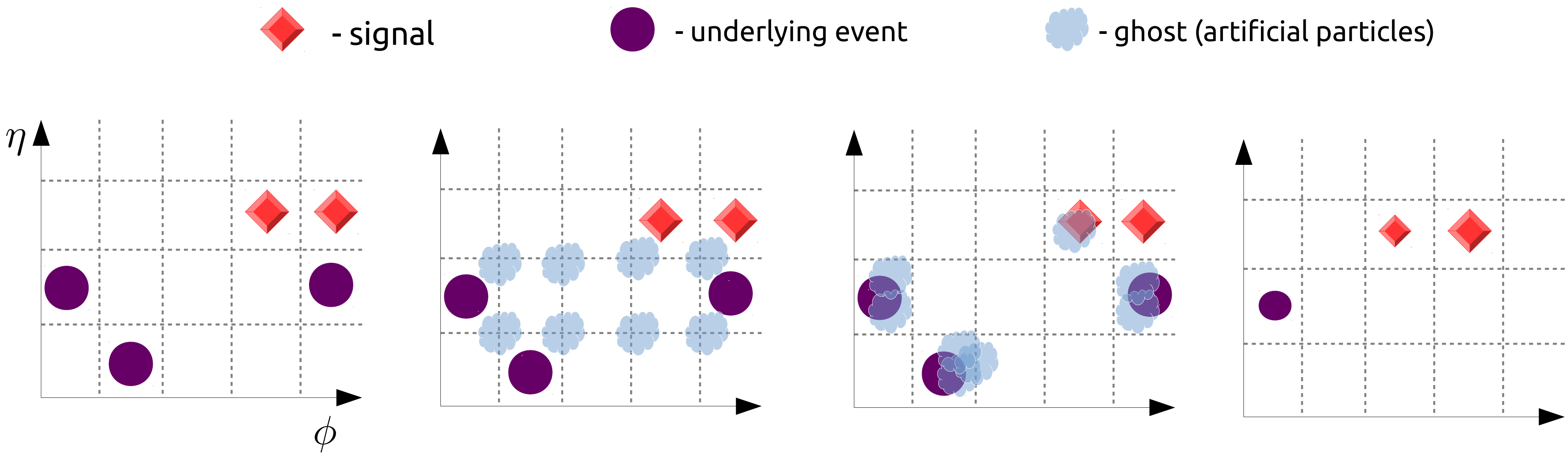




UE in pp with <PU> ~ 200 looks like central PbPb

UE subtraction : constituent subtraction

Particle-by-particle: correct the 4-momentum of a jet and substructure



Add ghosts with $p_T^{ghost} = A_{ghost} \cdot \rho$ in random locations;

A_{ghost} - area occupied

Combine them with the closest real particle

$p_T^{particle} > p_T^{ghost}$

$p_T^{particle} < p_T^{ghost}$

The largest pT particle/ghost survives

$p_T^{particle} = p_T^{particle} - p_T^{ghost}$

$p_T^{ghost} = p_T^{ghost} - p_{T\,particle}$

Repeat until no ghosts/particles left

Remaining particles get clustered into a jet