Evidence for top quark production in PbPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV

CMS PAS HIN-19-001

marc dünser (CERN)
on behalf of the CMS collaboration

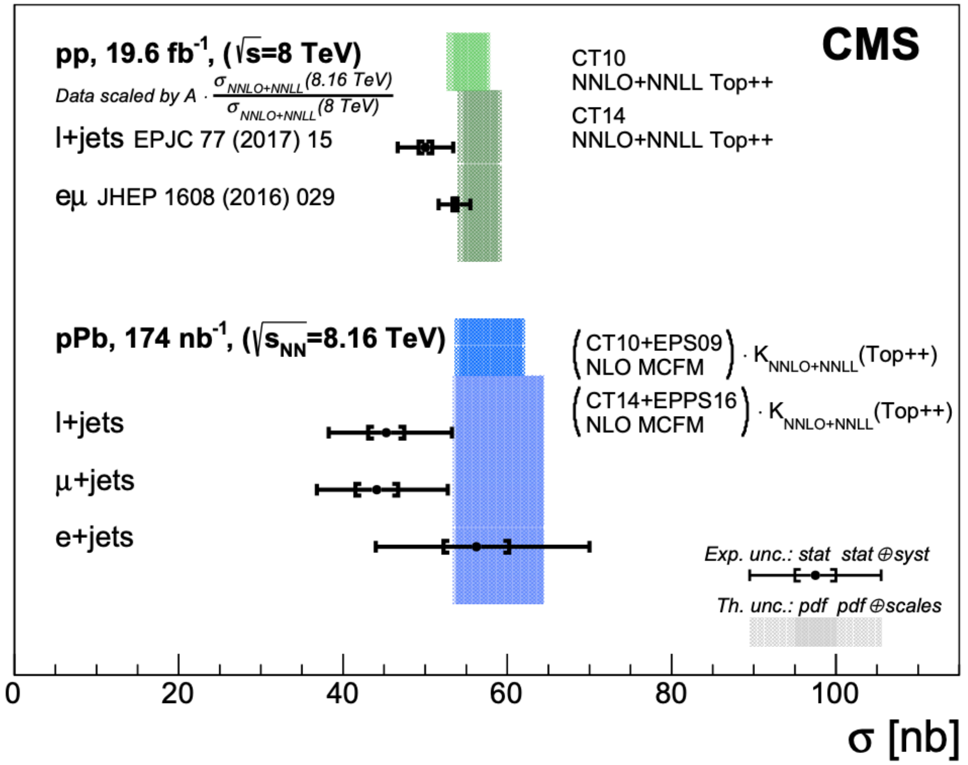
LHC TOP WG meeting, CERN
15th of november 2019

top quarks in non-pp collisions

this is not the first time tops have been seen outside of pp

published in PRL
[link to public]

-> CMS had the first observation of ttbar in pPb in 2016 with $174\ \mathrm{nb}^{-1}$ @ $\mathrm{sqrt}(s_{NN}) = 8.16$ TeV



in 2018 we took $1.7\ \mathrm{nb}^{-1}$ @ $\mathrm{sqrt}(s_{NN}) = 5.02$ TeV of PbPb collision data

-> about $74\ \mathrm{pb}^{-1}$ pp equivalent

-> expected cross section of $73.05 * 208^2$ pb

-> inclusively some few thousand ttbar pairs
some hundred events of dileptonic ttbar. need to retain efficiency!

top quark production has so far been unobserved in ion-ion collisions!

-> CMS just made the first such result public!
CDS record: [https://cds.cern.ch/record/2699428]
public pages: [CMS-PAS-HIN-19-001]

physics motivation

so far unobserved in ion-ion collisions!!

-> and the first viable dataset came less than a year ago!

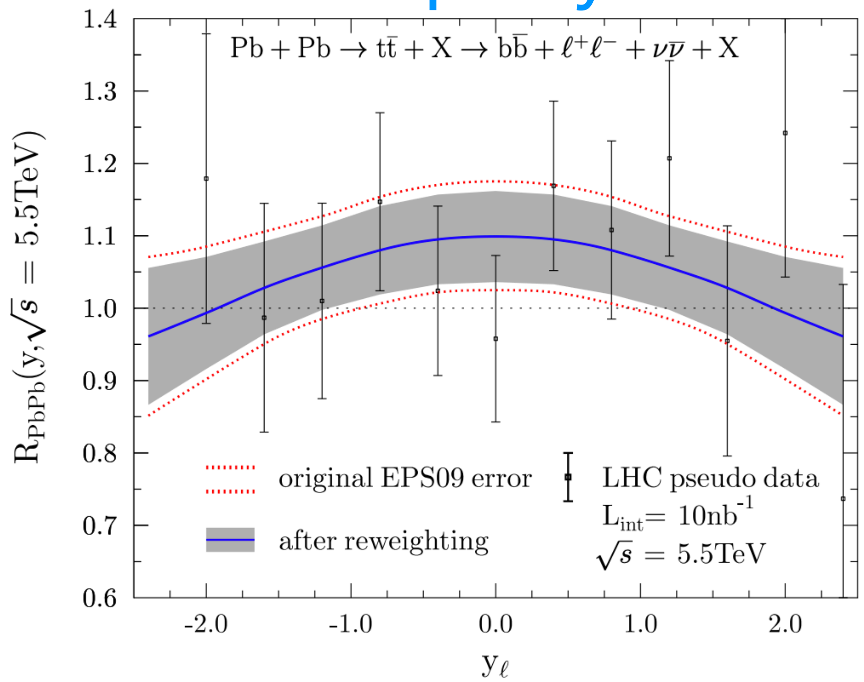
ttbar production is a very high-x gluon process

-> ~80% produced via gluon-fusion

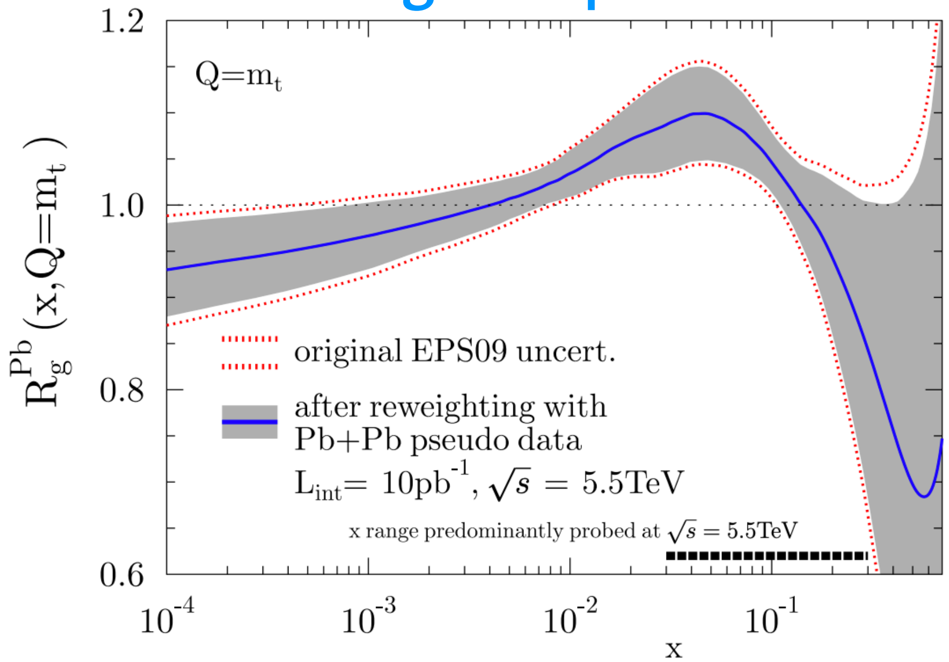
-> leptonic decays unaffected by final state effects in the QGP

-> constraints on nuclear pdfs through σ_{tt}

lepton-y



gluon-pdf



more details nPDFs: [arXiv:1501.05879]

physics motivation

ttbar sensitive to the time structure of QGP

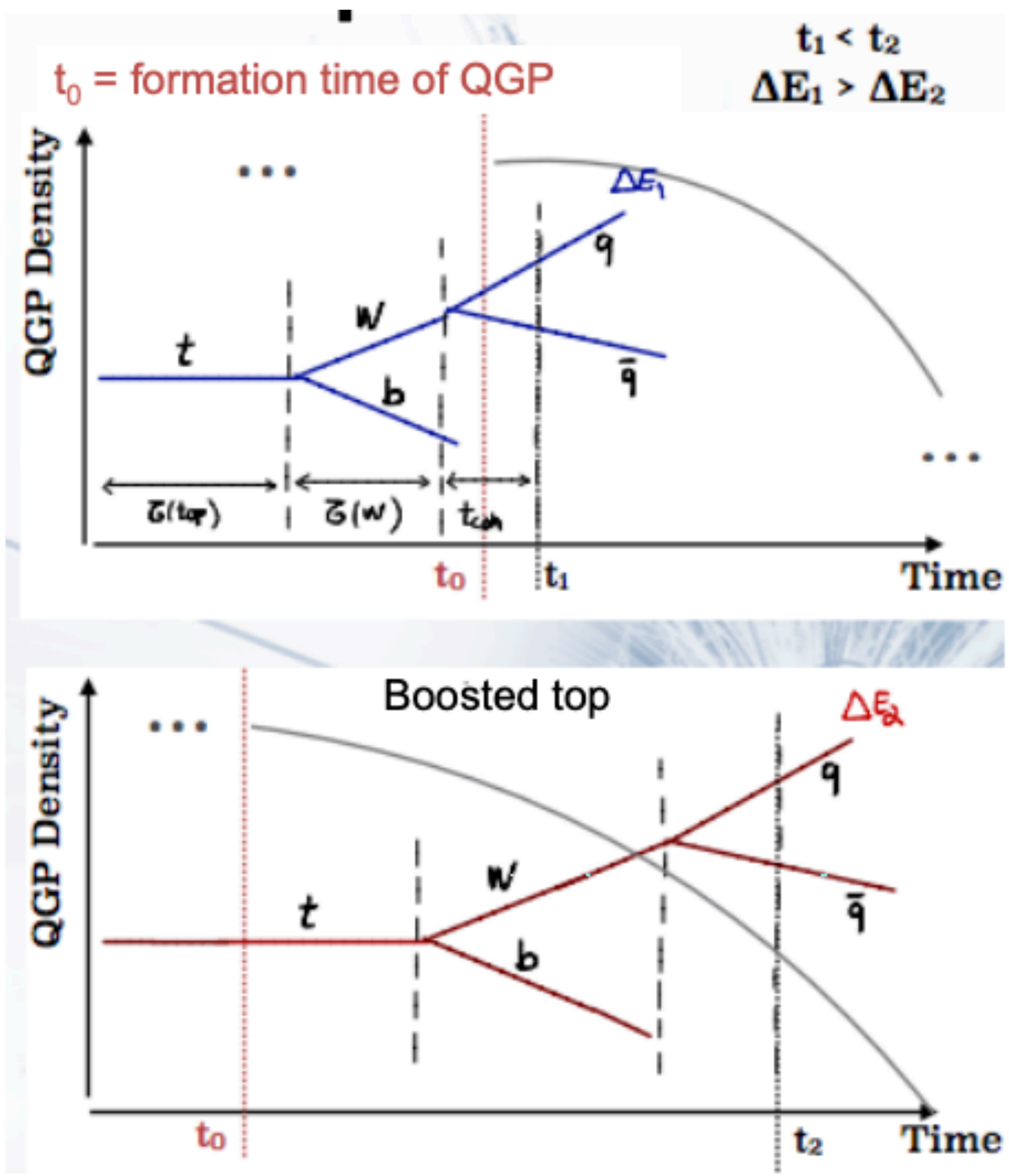
-> at rest the top quarks decay well before the QGP is formed

-> hadronic products of Ws pass through the QGP medium

-> boosted top quarks can live longer than the formation time of the QGP and probe the medium

this will require much larger datasets at higher CM energies

-> but goes to show that the top quark is a very interesting probe also in heavy ion collisions



more details time structure: arXiv:1711.03105

analysis strategy

for highest possible purity we look into fully leptonic decays

-> most sensitivity from ***eμ***, but also ***μμ*** and ***ee*** (exclude on-Z for SF)
-> $p_T > 20$ (25) GeV and $|eta| < 2.4$ (2.1) for μ(e)
-> ID and special heavy-ion isolation applied

we extract two signal strengths in the analysis

-> **one with only leptonic** information from a BDT
-> **one including counting $n_{b\text{-tags}}$** with $p_T > 30$ GeV and $|eta| < 2.0$

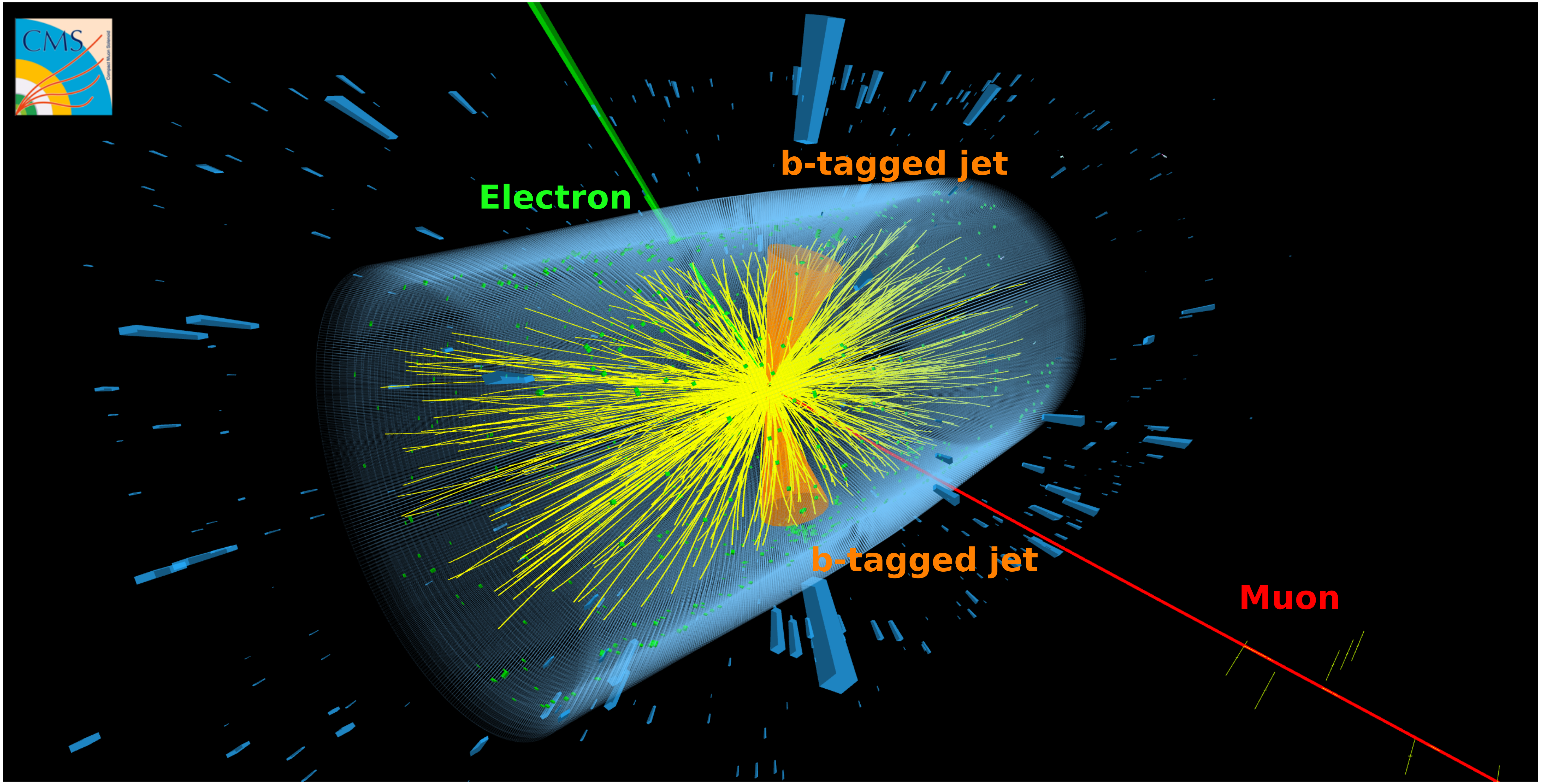
to maximise sensitivity we train a BDT to discriminate ttbar v. DY

-> use only *leptonic* observables!

(b-)jets in heavy-ion collisions are different compared to pp

-> there's many more, for one. most from UE
-> they are affected by the QGP in some way that isn't modeled by the MC
-> employing standard use of jets in pp top analyses is tricky

how does it look like?



not vastly different from pp ttbar production...

backgrounds - overview

not too many backgrounds left
- -> in the SF: mostly DY+jets
- -> in the OF: DY->tautau and ‘nonprompt’ backgrounds
- -> smaller contributions from tW, VV

DY, tW and VV from MC with aforementioned corrections
- -> we find good pre-fit agreement in DY
- -> special care taken for the DY background in a few ways (p_T, jets)

non-prompt background more tricky
- -> mostly W+jets (not enough MC) with jets passing ID/iso
- -> but also QCD multi-jets with heavy flavor decays
- -> derive this background from event-mixing
- -> normalize it to the same-sign control region

systematic uncertainties

systematic uncertainties considered of a few different types

-> applied as either a shape uncertainty or lnN, including MC stats

theory:

-> nPDFs + QCD scales. very small changes in shape of the signal

-> p_T^{top} modeling, derived from pp. very small effect

$$w = \sqrt{SF[p_{\mathrm{T}}(t)] \cdot SF[p_{\mathrm{T}}(\bar{t})]}, \quad SF(p_{\mathrm{T}}) = e^{0.199 - 0.00166 p_{\mathrm{T}}}$$

-> m_{top} uncertainty. also small effect

-> Z-p_T modeling: up/down variations data/MC in dilepton-p_T

experimental

-> background normalization: 30% for VV, tW

-> nonprompt background 20% (equivalent to stat power of SS sample)

-> DY background normalization unconstrained

-> luminosity: 5% flat log-normal

-> trigger/RECO/ID/iso efficiency uncertainties propagated (2-5%)

systematic uncertainties - $n_{\text{b-tags}}$

a number of systematic uncertainties are considered for the $n_{\text{b-tags}}$ spectrum

-> b-tagging efficiencies (b and udsg separate)
-> JER+JES applied as is standard

heavy-ion specific: jet quenching uncertainty on the jets

-> (b)-quarks moving through the medium lose some momentum
-> results in jets going below the threshold of $p_T > 30$ GeV
-> not modeled in our MC

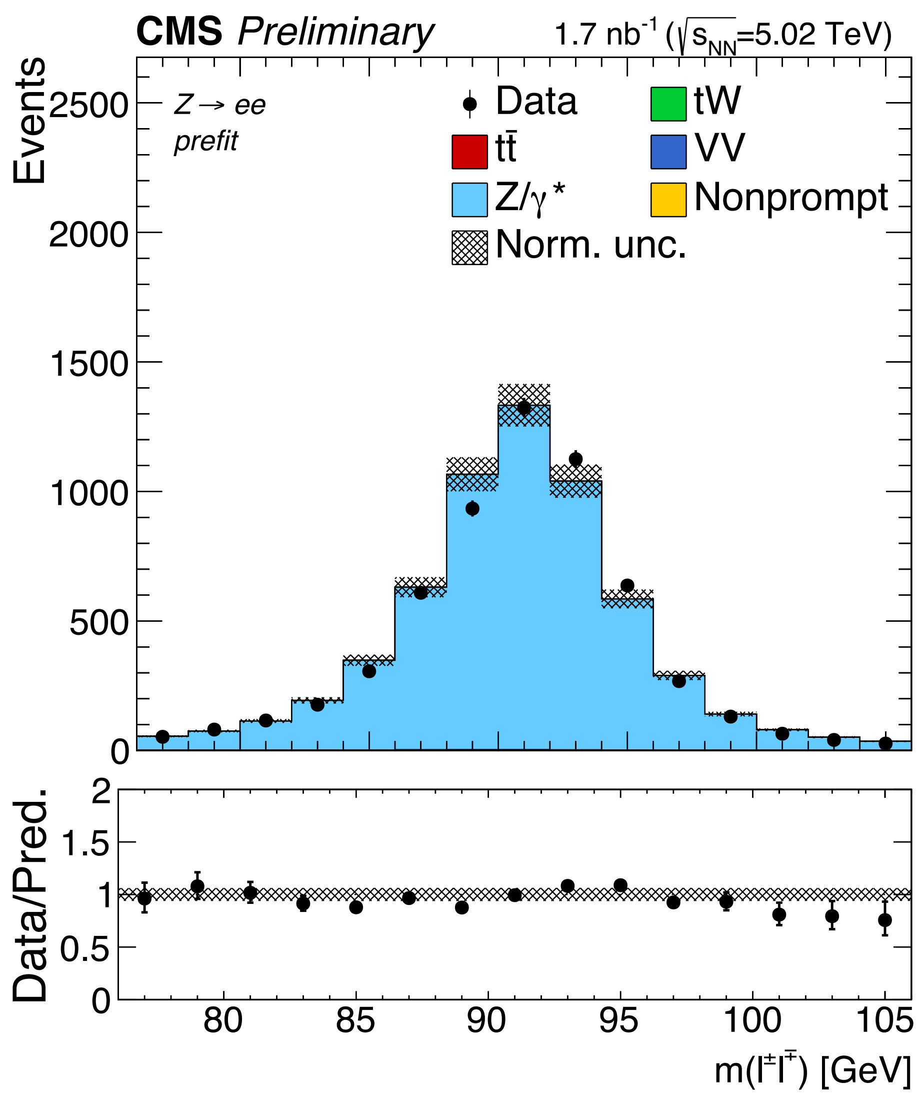
we use a theoretical model which is fit to CMS data

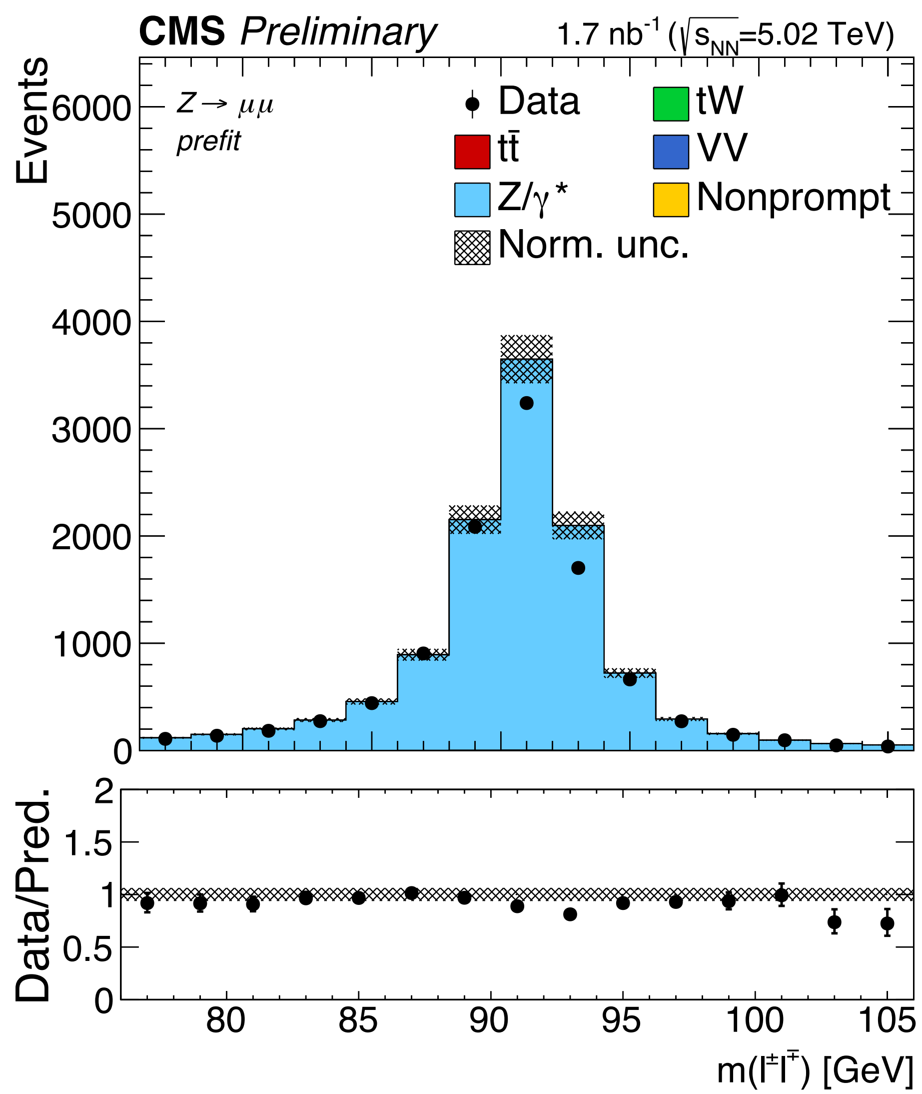
-> F. Arleo: *"Quenching of Hadron Spectra in Heavy Ion Collisions at the LHC"* https://arxiv.org/pdf/1703.10852.pdf
-> implemented with centrality dependence
-> overall effect of around 5-10% in the $n_{\text{b-tags}}$ spectrum

pre-fit distributions - DY region

on-Z region excluded from the signal region

-> ideal place to check data-MC agreement



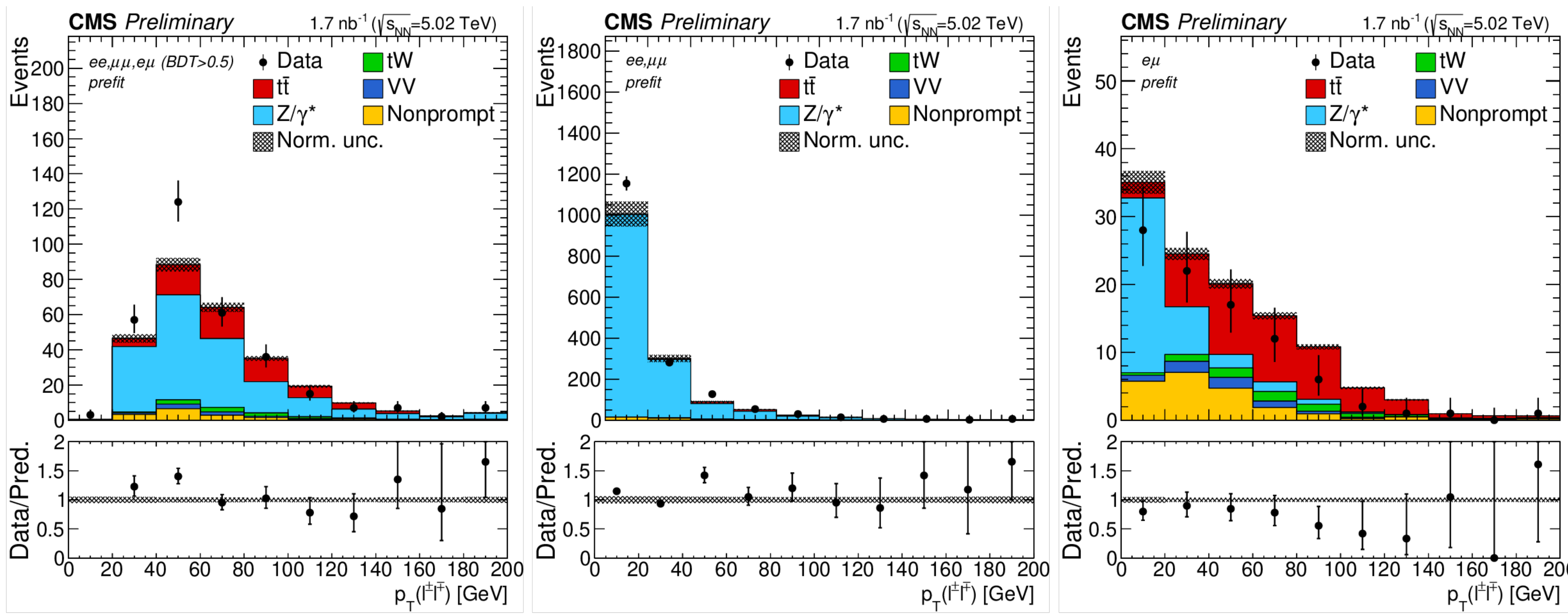


good agreement seen overall

-> NB: these distributions are all pre-fit!

pre-fit distributions - signal region

signal region pre-fit (off-Z SF + full range OF)



data slightly below the prediction for high sensitivity region

-> will extract a signal strength < 1

event yields

we observe 1768 events in the data

-> of which **43.2 ± 11.1 ttbar signal** extracted from the fit

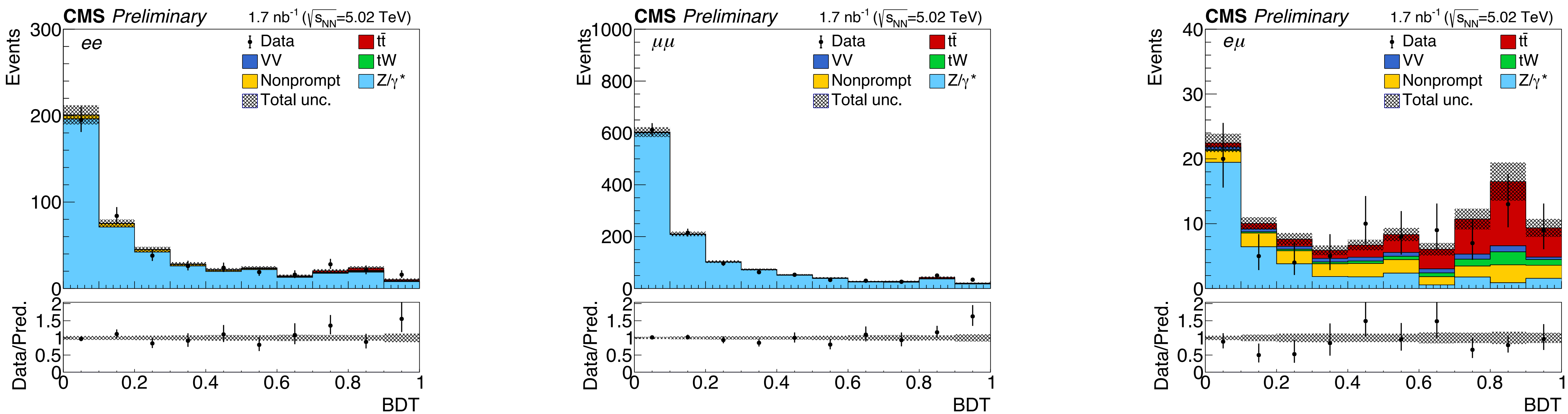
Process	Final state e^+e^- 0b	1b	2b	$\mu^+\mu^-$ 0b	1b	2b	$e^\pm\mu^\mp$ 0b	1b	2b
Z/γ^*	389.8±15.4	40.4±2.7	4.4±0.8	1027.5±27.3	136.1±5.7	14.1±1.7	35.1±1.7	4.4±0.9	0.7±0.2
Nonprompt	17.3±2.2	1.4±0.2	≤ 0.1	7.6±1.0	0.8±0.1	≤ 0.1	17.1±1.9	4.0±0.4	≤ 0.1
tW	1.1±0.2	0.9±0.2	≤ 0.1	1.8±0.4	1.3±0.3	0.2±0.1	3.4±0.7	2.5±0.5	0.4±0.1
VV	1.9±0.3	0.2±0.1	≤ 0.1	3.3±0.6	0.4±0.1	≤ 0.1	5.4±0.9	0.6±0.1	≤ 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

best S/B in eμ. SF only matters in >1 b-tags

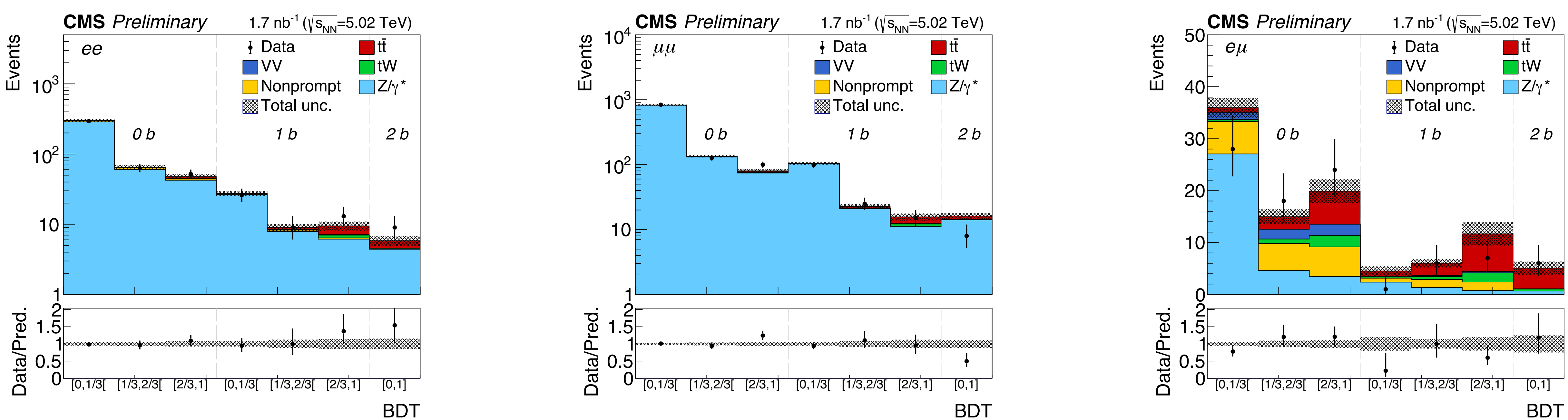
-> very high purity in 2b channel

results

leptonic BDT only



leptonic BDT + binning in $n_{\text{b-tags}}$



results

interpret the result of the fit in terms of two signal strengths and cross sections of ttbar

-> reference nucleon-nucleon cross section is **73.05 pb @ 5.02 TeV** from top++ calculation with EPPS16+CT14 pdf

in BDT+$n_{b\text{-}tags}$ we find significance of 4.0 sigma (6.0 expected)!

-> signal strength **μ = 0.64 ± 0.21**

leptonic-BDT only **results compatible:**

-> **μ = 0.81 ± 0.25** with a significance of **3.8 (4.8 expected)**

uncertainty very dominated by statistical power of the dataset!

-> 85% of total uncertainty is of statistical nature

first strong evidence for ttbar in PbPb!!

-> first time with leptonic variables only!

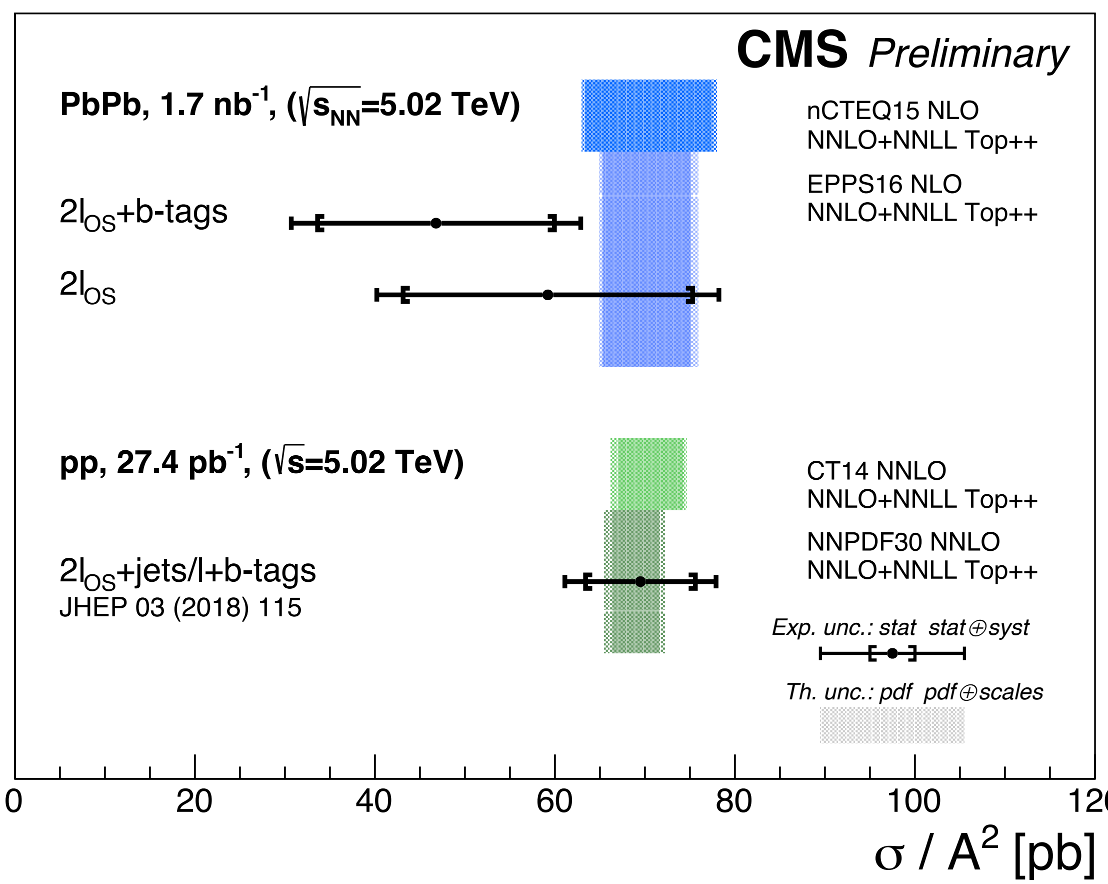
more results

very statistically dominated result

-> need to take more PbPb data!

all extracted signal strengths < 1

-> could be a statistical effect only



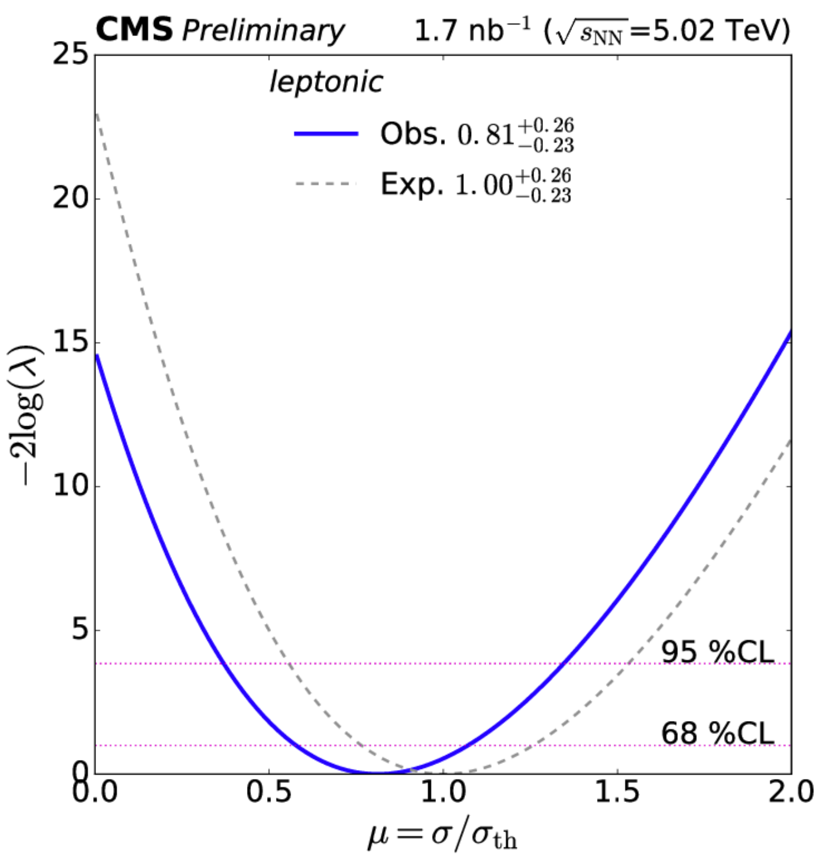
Source	$\Delta\mu/\mu$	
	leptonic-only	leptonic+b-tagged
Total statistical uncertainty	0.27	0.28
Total systematic experimental uncertainty	0.17	0.19
Background normalization	0.12	0.12
Background and $t\bar{t}$ signal distribution	0.07	0.08
Lepton selection efficiency	0.06	0.06
Jet energy scale and resolution	—	0.02
b tagging 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

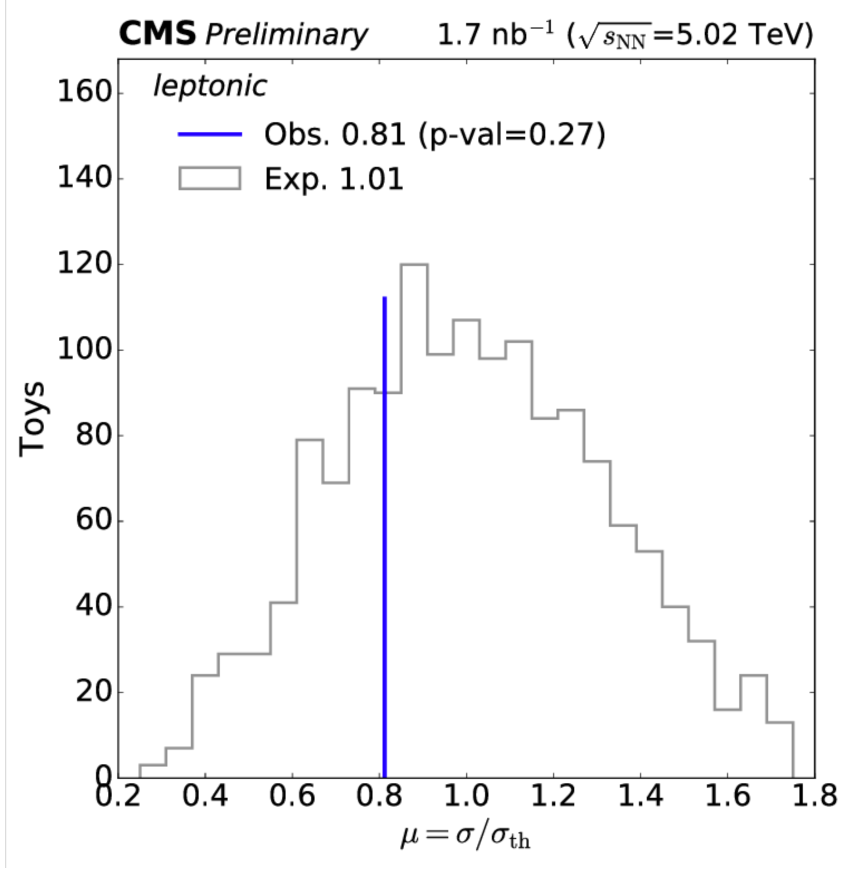
a bit of statistics

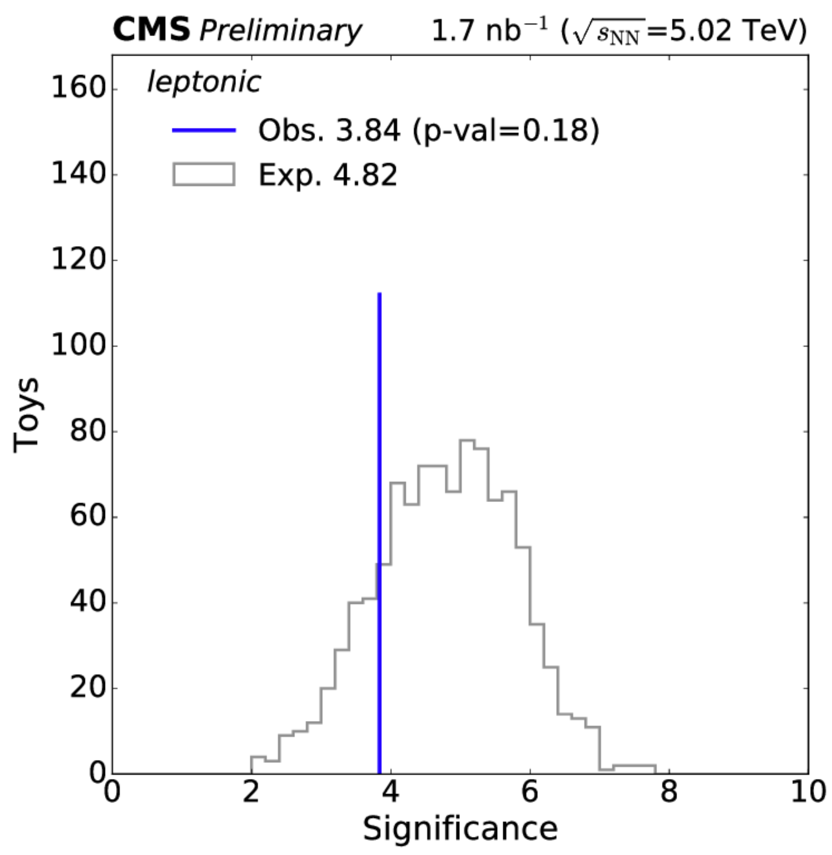
likelihood scans of the leptonic BDT

-> expected significance somewhat below 5 sigma

BDT+only







observed significances almost coincide

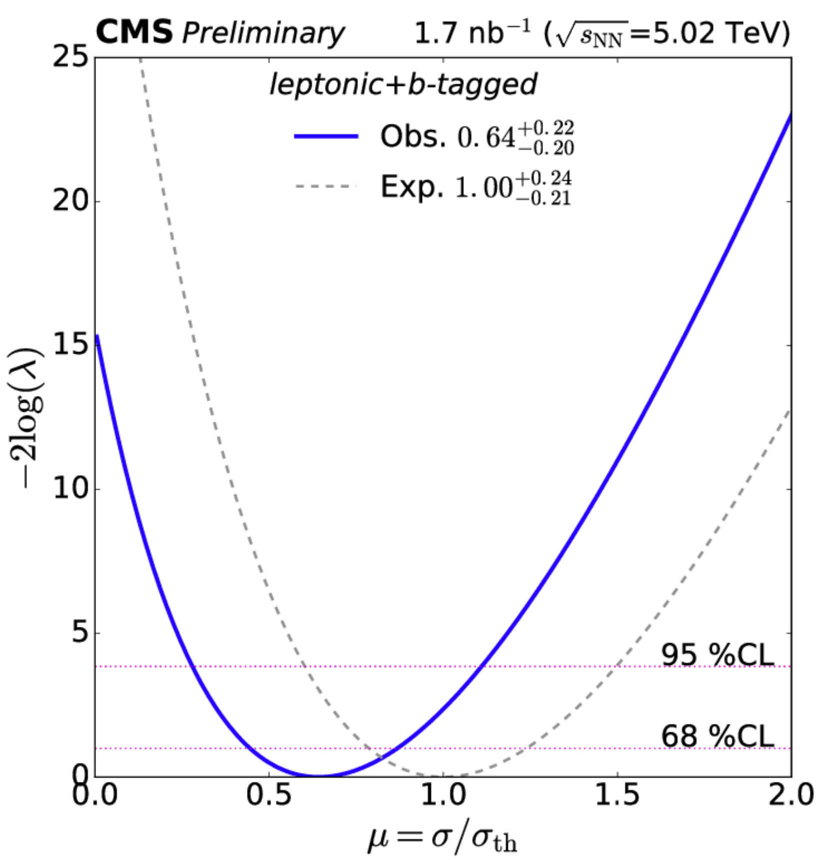
-> uncertainties are very similar in all variants

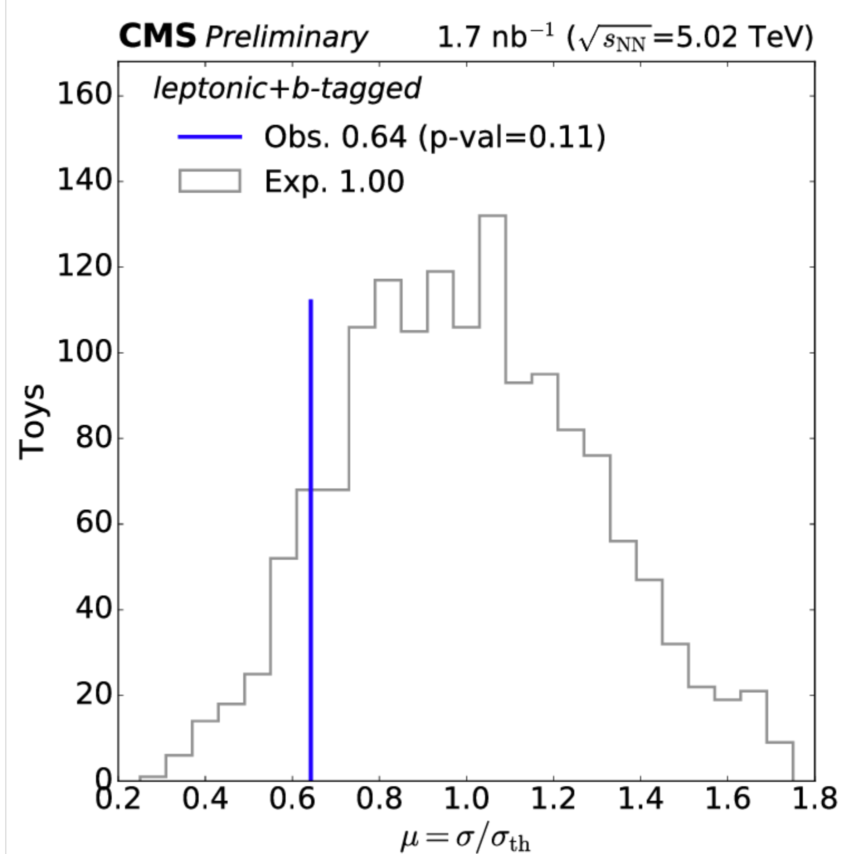
a bit of statistics

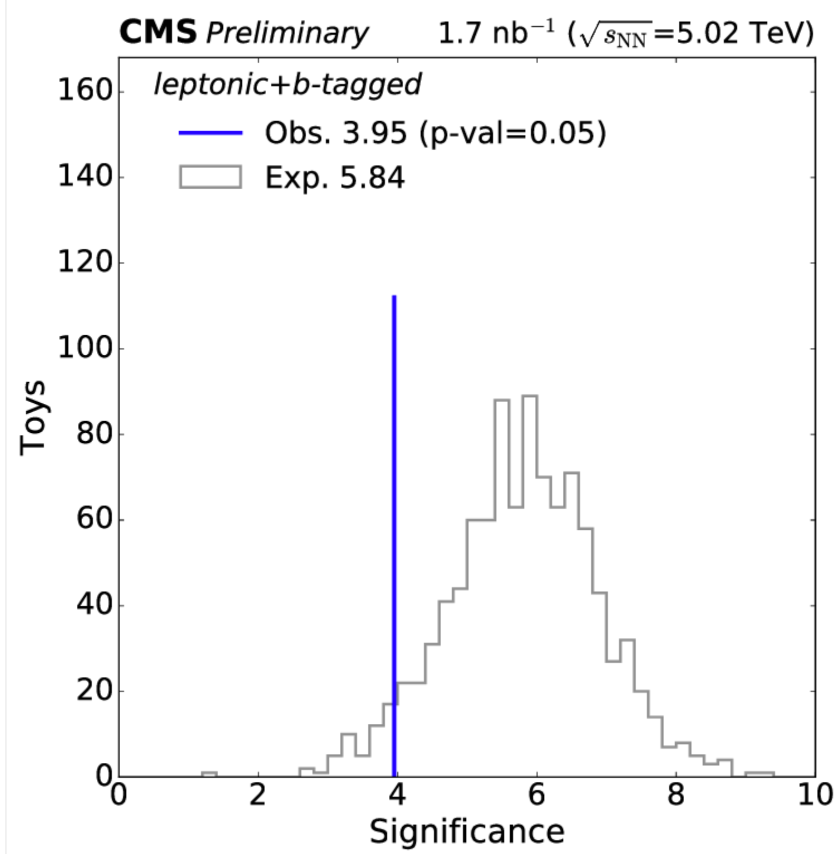
likelihood scans of the leptonic BDT +n$_{\text{b-tags}}$

-> expected significance drastically improved by inclusion of b-tag information

BDT+n$_{\text{b-tags}}$







observed significances almost coincide

-> uncertainties are very similar in all variants

summary

this result opens the door for exciting top physics in PbPb!

-> a first step towards using the top quark to probe heavy ion and QGP behavior

we see strong evidence for ttbar in our dataset

-> the statistical power is such that it's difficult to say whether $\mu < 1$ is due to a real physics effect or just a statistical fluctuation

-> a good case to take much more PbPb data soon!

would be great if ATLAS could perform a similar analysis!

-> would expect a gain of about sqrt(2) by combining the experiments and claim 5σ observed discovery!

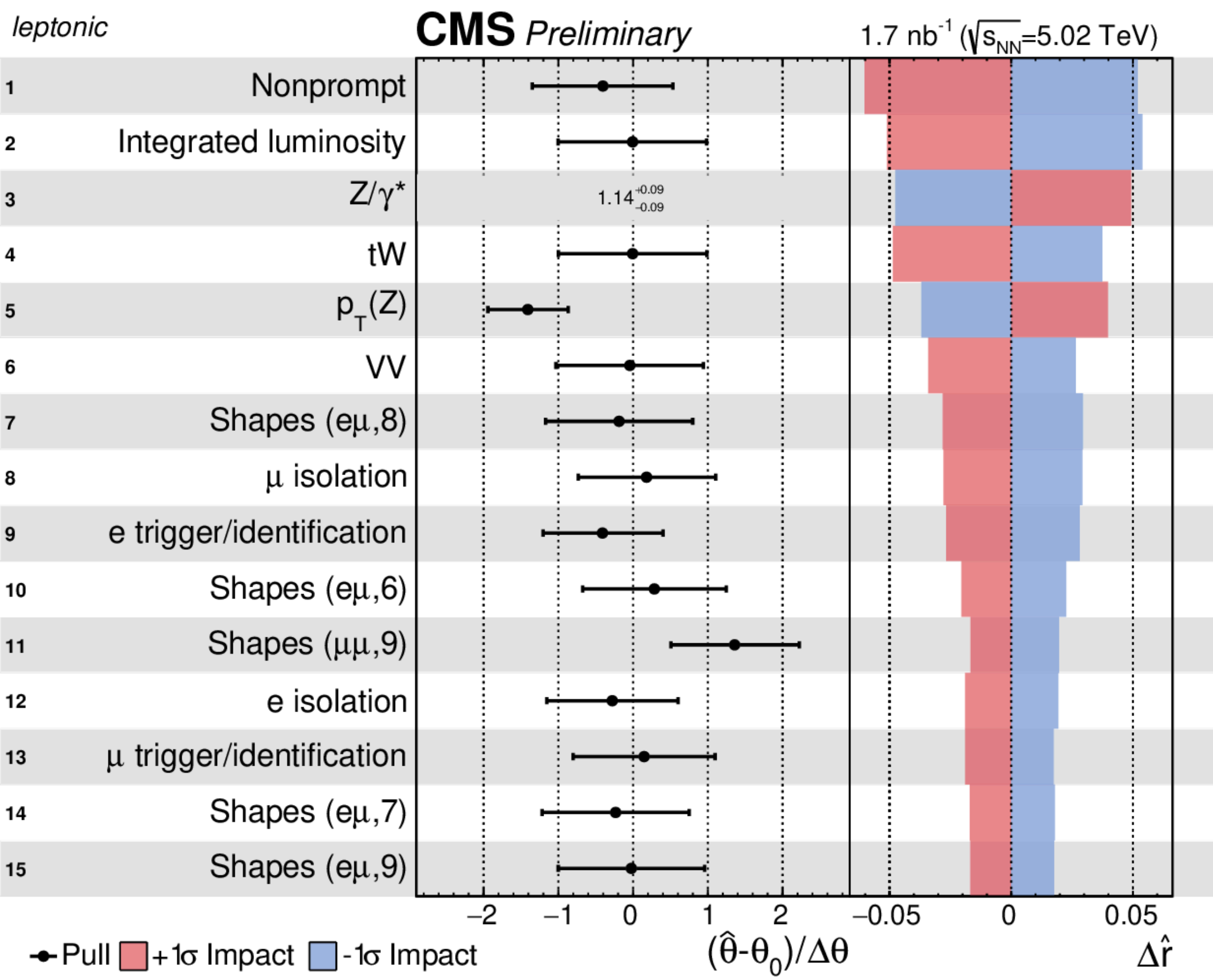
the end

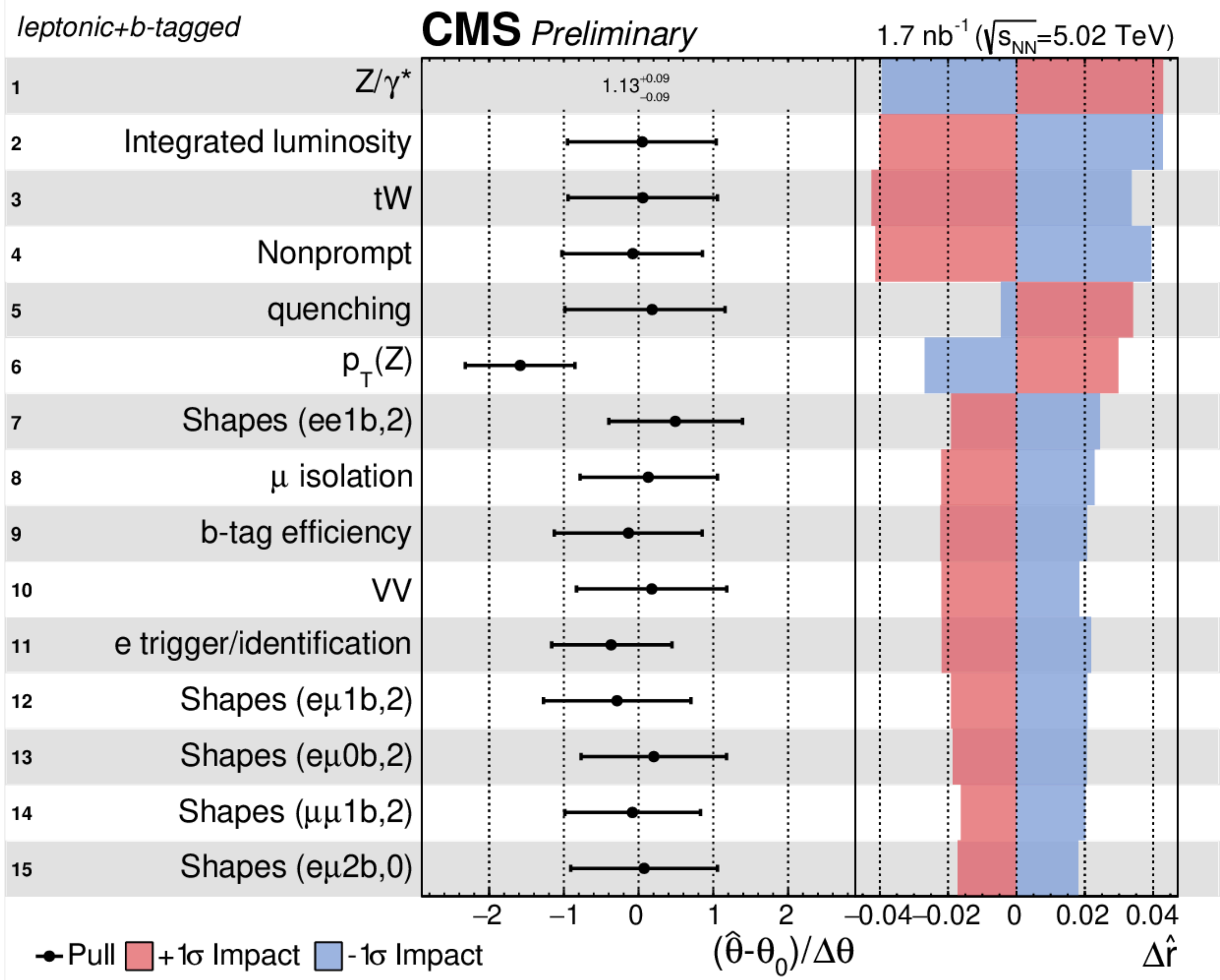
marc dünser (CERN)
on behalf of the CMS collaboration

extras

impacts

NB: largest (by far) statistics: ~0.2





leptonic only

BDT + $n_{b\text{-tags}}$

signal BDT

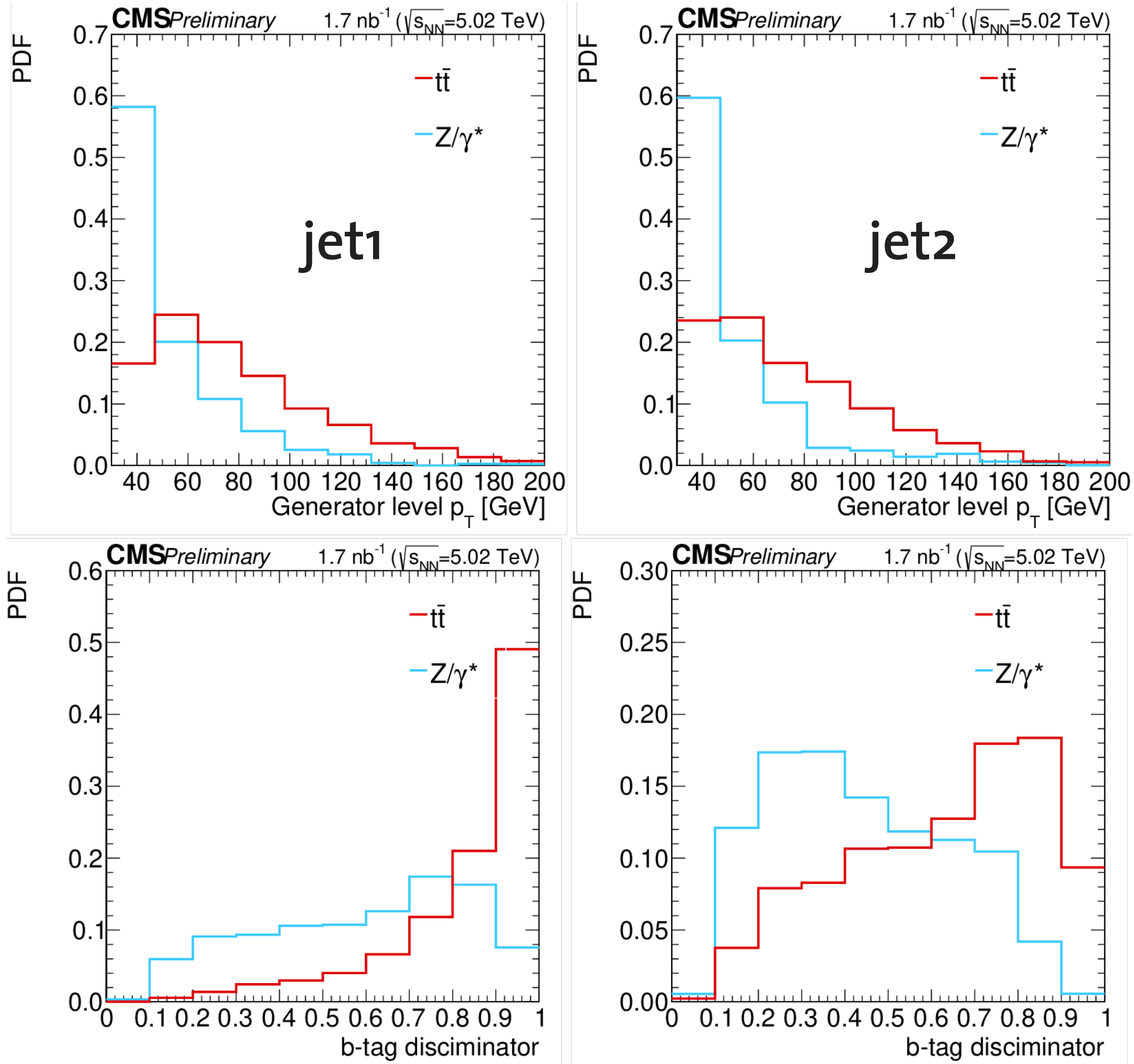
to maximize the sensitivity we train a leptonic BDT

-> no hadronic information used for this at all

- $p_{\mathrm{T}}(\ell_1)$ the p_{T} of the highest-p_{T} lepton,
- $A_{p_{\mathrm{T}}}$ the asymmetry in the lepton-p_{T}'s, namely $\frac{p_{\mathrm{T}}(\ell_1)-p_{\mathrm{T}}(\ell_2)}{p_{\mathrm{T}}(\ell_1)+p_{\mathrm{T}}(\ell_2)}$,
- $p_{\mathrm{T}}(\ell\ell)$ the p_{T} of the dilepton system,
- $|\eta(\ell\ell)|$ the absolute η of the dilepton system,
- $|\Delta\phi(\ell\ell)|$ the absolute value of the separation in ϕ of the two leptons, and
- $\Sigma|\eta_i|$ the sum of the absolute ηs of the leptons.

more on b-jets

pure MC studies



nonprompt background derivation

event mixing of two leptons from two different events in data!

-> same flavor combination mixed

-> pick 100 events, exclude the same event and repetitions

-> each combination gets a distance assigned

-> distance based on kNN algorithm
features are centrality, rho, iso, p_T^l, p_T^{ll}

use set of closest events in this distance as central shape

-> furthest distance as alternative shape
for systematic treatment in the fit

normalize the distribution to the same-sign data sample yield