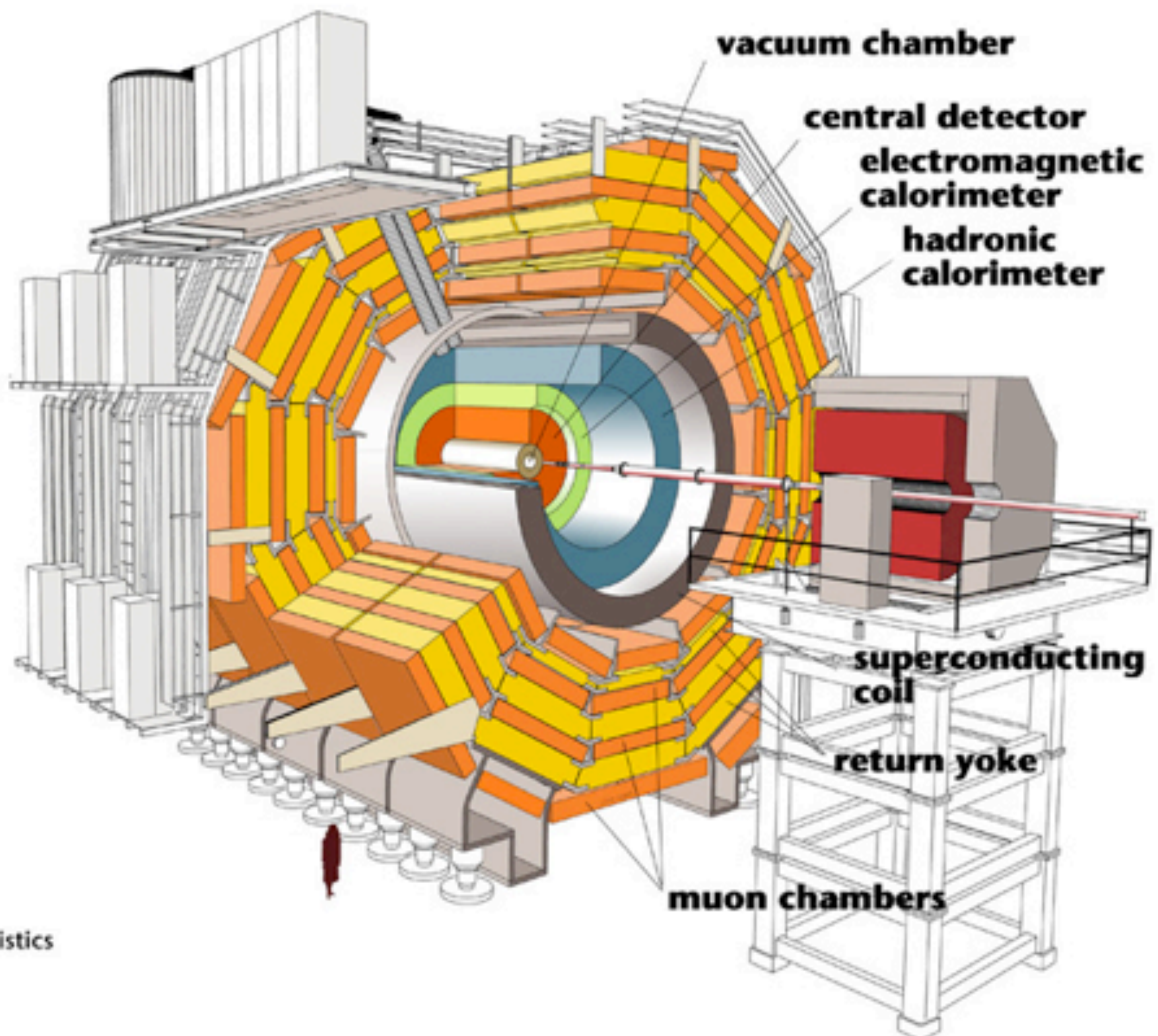
Heavy flavor quark production in association (or not) with vector bosons in CMS

*
* *

Simon de Visscher, Universität Zürich
on behalf of the CMS collaboration

Outline and references



Detector characteristics

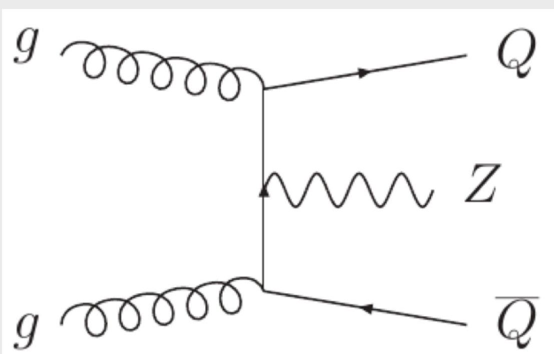
Width: 22m
Diameter: 15m
Weight: 14'500t

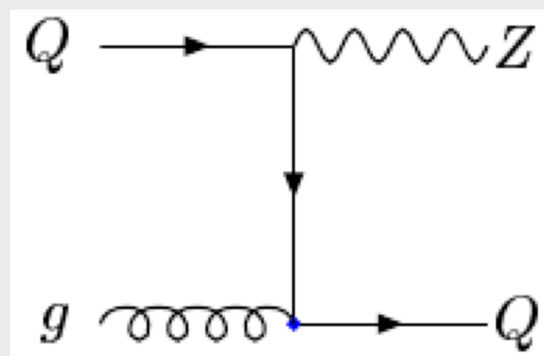
Introduction	
W+c:	CMS-EWK-11-013
Z+>=1b-jet:	arxiv:1204.1643
Z+1,2 b-jets(s):	CMS-SMP-12-003
Z+b-hadron:	CMS-EWK-11-015
bb→μμX:	CMS-BPH-10-015
conclusion	

QCD

HF/V+HF measurements provide stringent test of QCD predictions and MonteCarlo techniques.

-4-flavour vs 5-flavour





-tree-level vs NLO

-test of ME/PS jet matching predictions

-PDF

-MPI

Backgrounds for Higgs, BSM searches

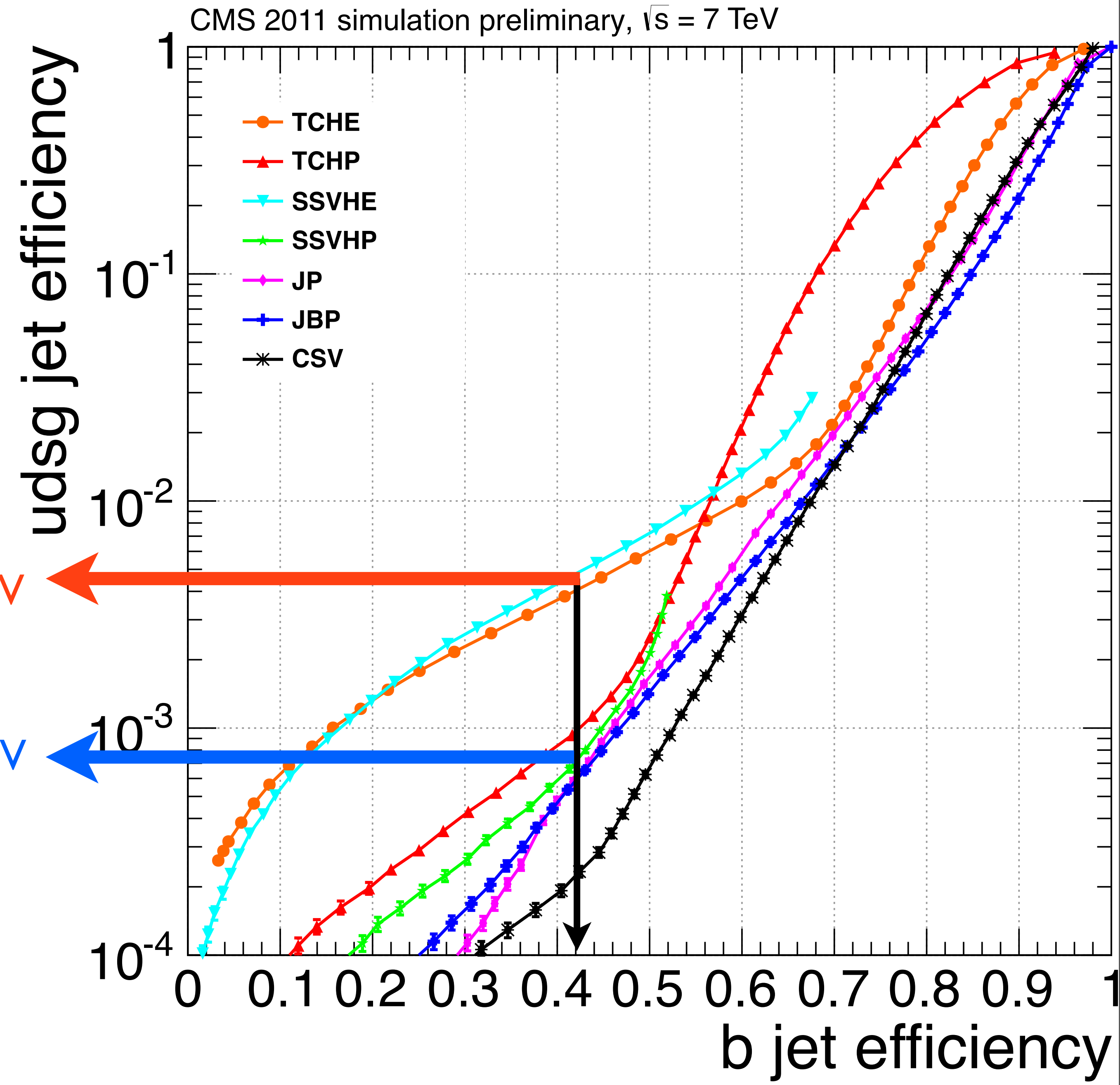
SM Higgs:VH, H→bb

2HDM : H→Zϕ, ϕ→bb

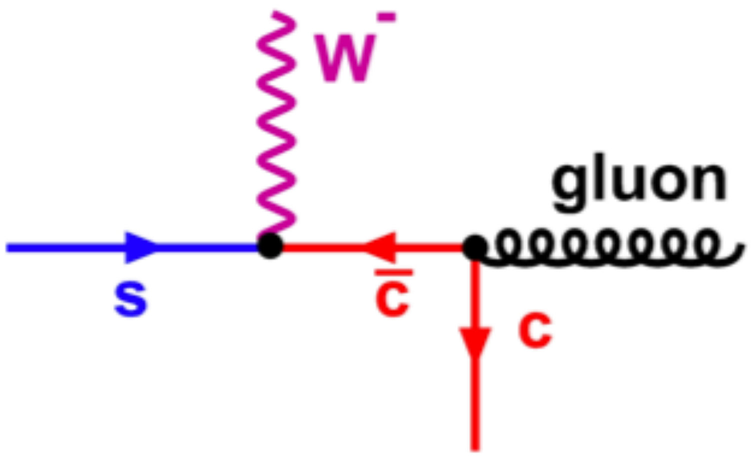
SUSY : sbottom

Heavy hadron tagging technique

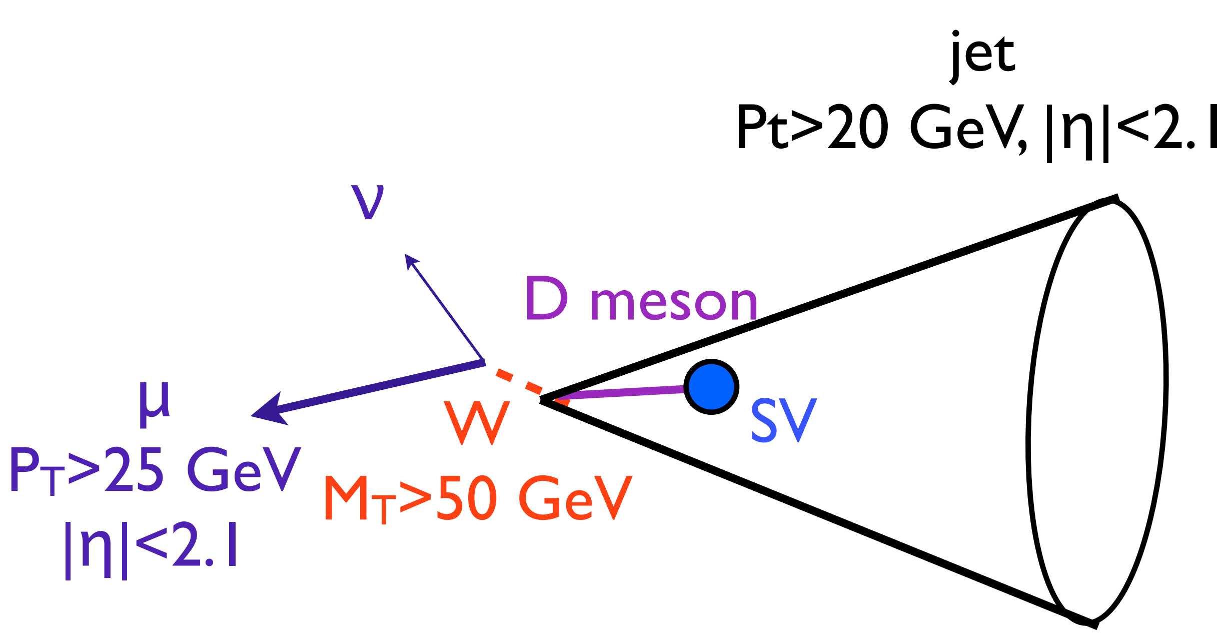
- Simple Secondary Vertex (SSV) tagger.
  - Detect the presence of a displaced secondary vertex (SV) inside a jet
  - Use the SV flight distance significance as discriminator $D_{SSV} = sign(S)log(1+|s|), S = \frac{L_{3D}(PV-SV)}{\sigma_{3D}(PV-SV)}$
  - Cut on discriminator defines the b-tagging efficiency
  - Two versions
    - High efficiency (HE): >=2 tracks attached to the SV
    - High purity (HP): >=3 tracks attached to the SV

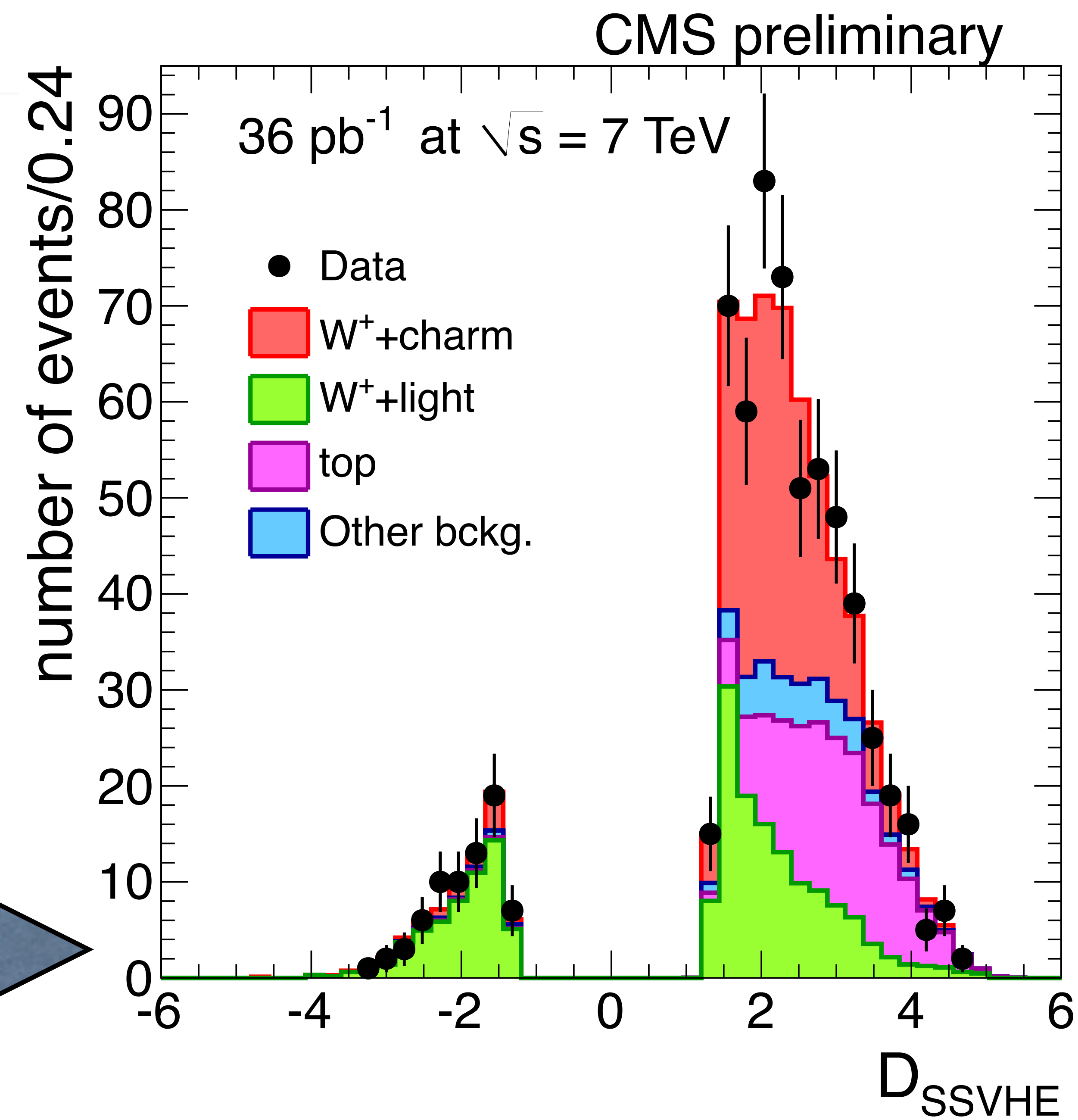


W+c relative cross-section measurement



- Check theory predictions about the following questions
  - Is the s-quark PDF antisymmetric under charge conjugaison
  - Proportion of Wc into W+jets? $R_c \sim \frac{s+\bar{s}}{\Sigma(q+\bar{q})}$





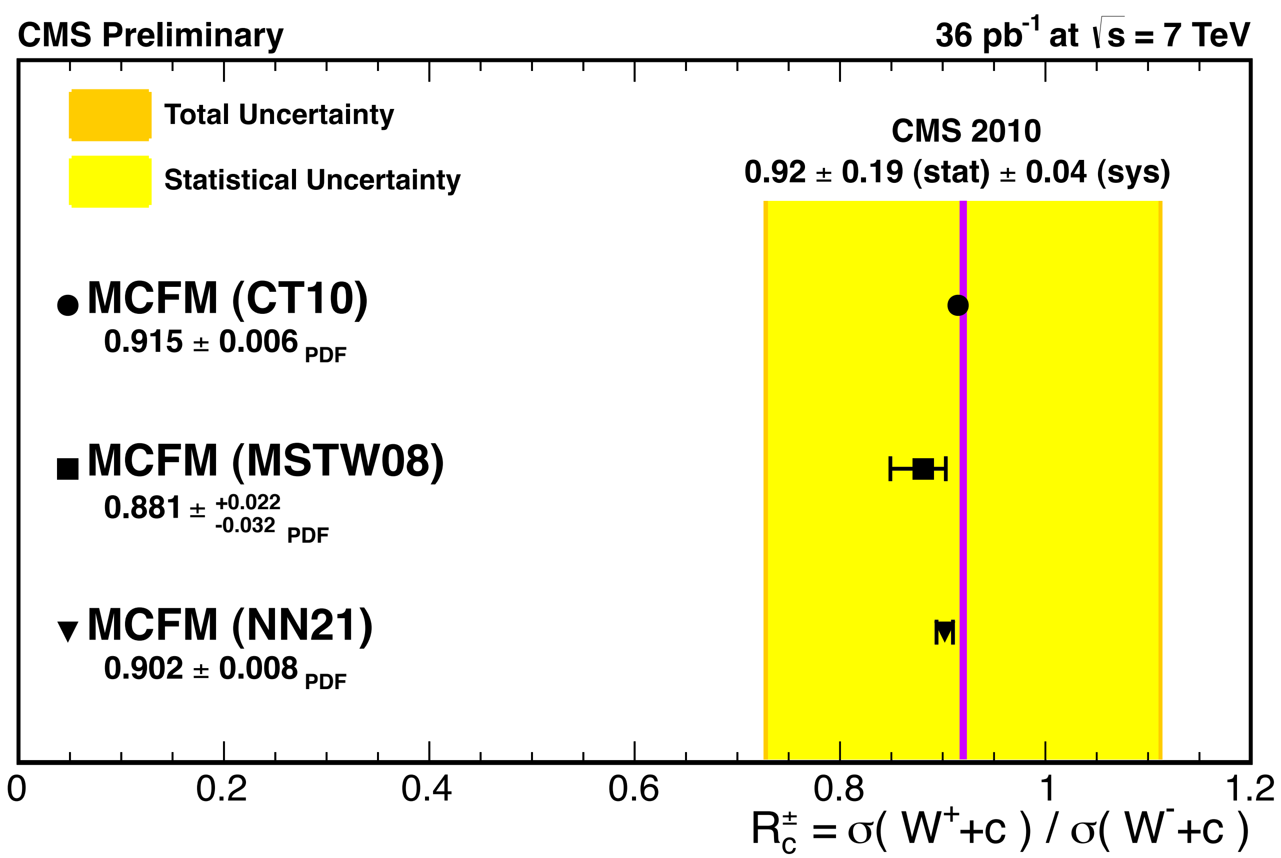
Fit the SSV(high efficiency) discriminator distribution to extract W+c content

only events witg flight distance<0.15 cm are kept.

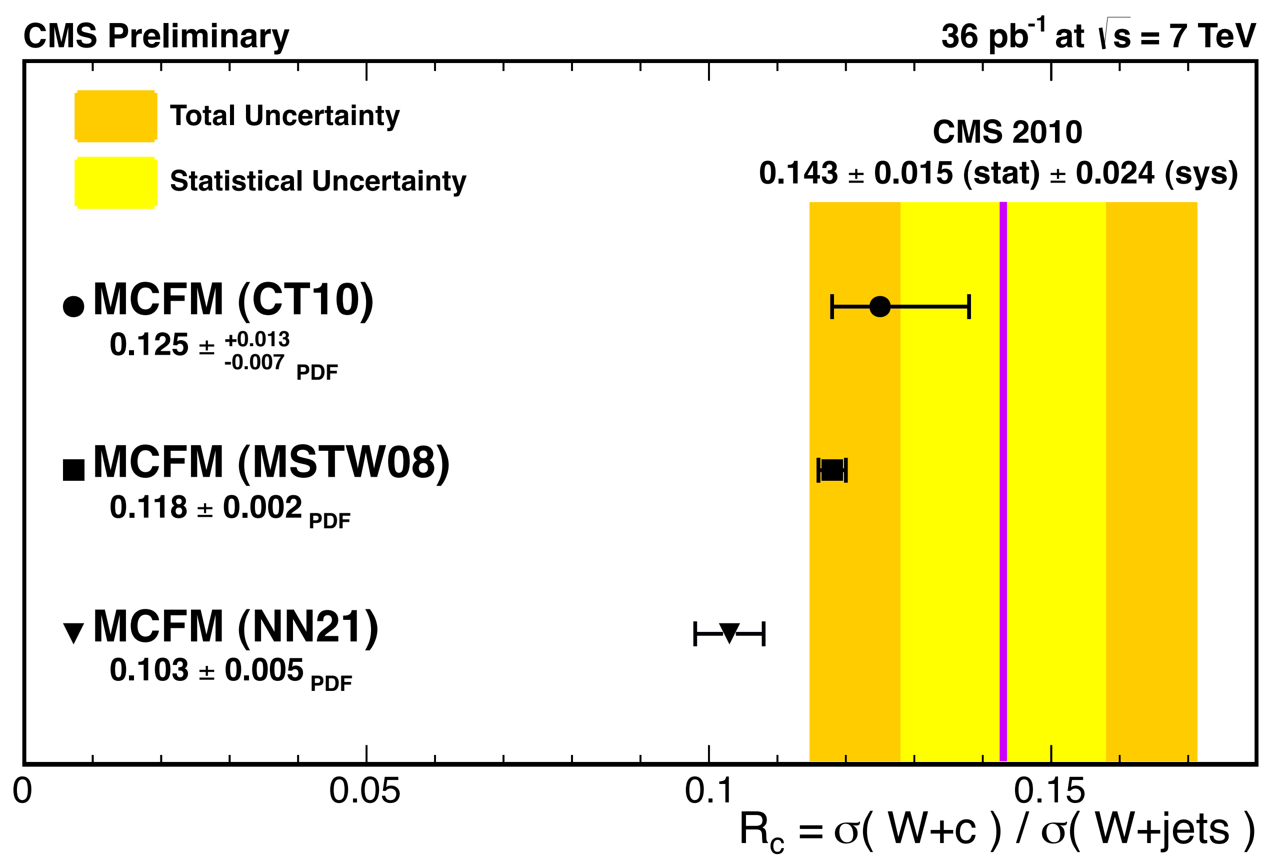
W+c: probe the s-quark pdf

Luminosity= 36/pb (2010 data)

$$\sigma(W^+\bar{c})/\sigma(W^-c)$$



$$\sigma(W+c)/\sigma(W+jets)$$



$R^{\pm}_c \neq 1$ because of d/dbar PDF difference, result compatible with s=sbar PDF see hep-ph/1203.6781

Overall good agreement between data and MC expectation. PDF dependence is visible

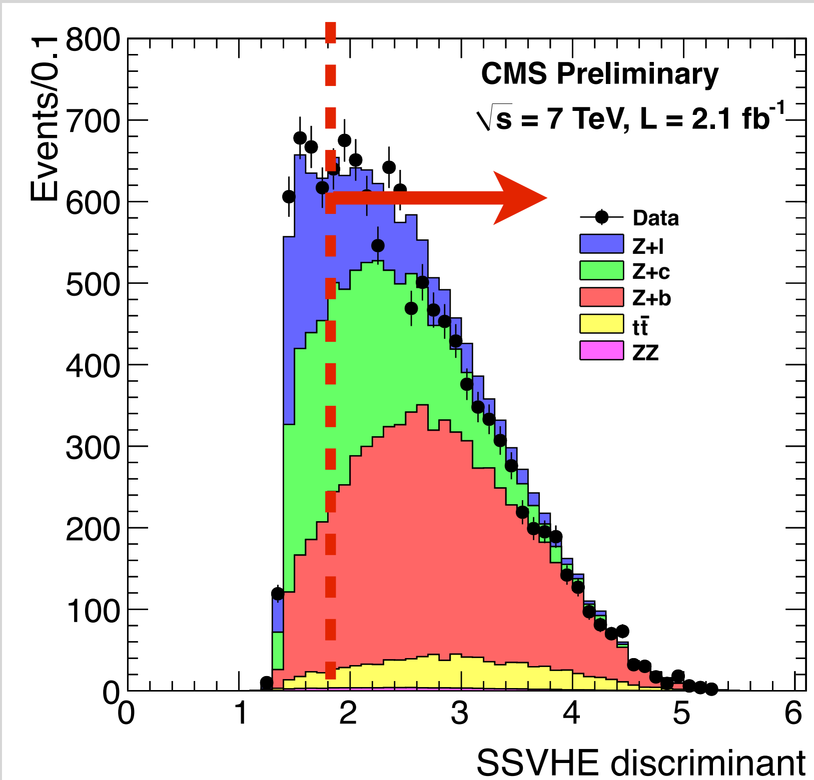
Z+b cross-section and angular correlation

- Z+>=1 b-jets: L=2.2/fb
- Z+1,2 b-jets: L=2.13/fb
- Z+2 b-hadrons: L=4.6/fb

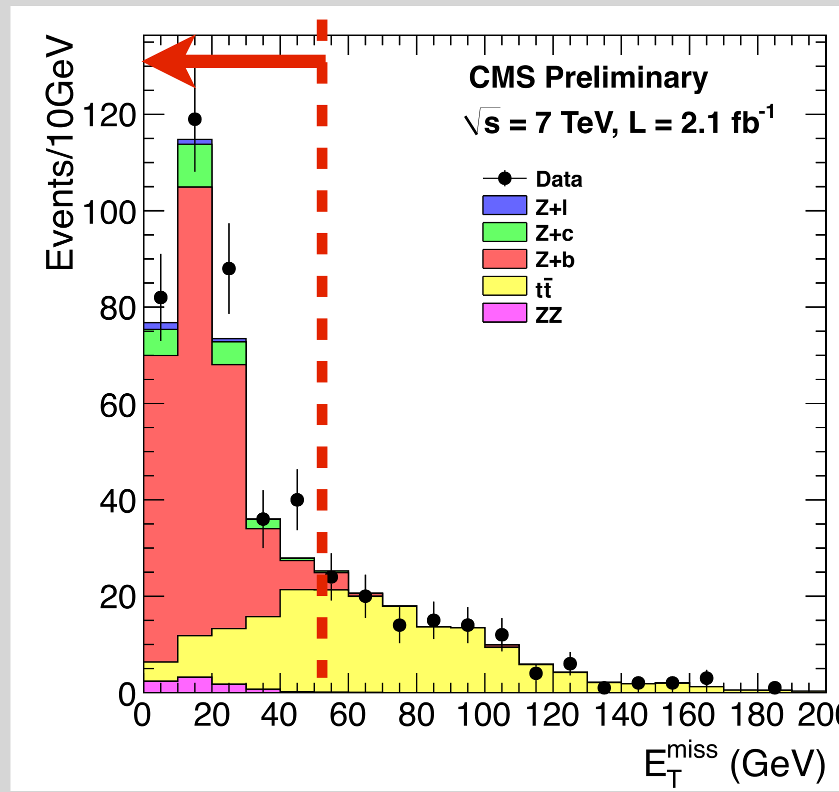
Z+b-jet(s): cross-section measurement

- Measure Z+>=1b and Z+2b independently. illustration for Z+2b

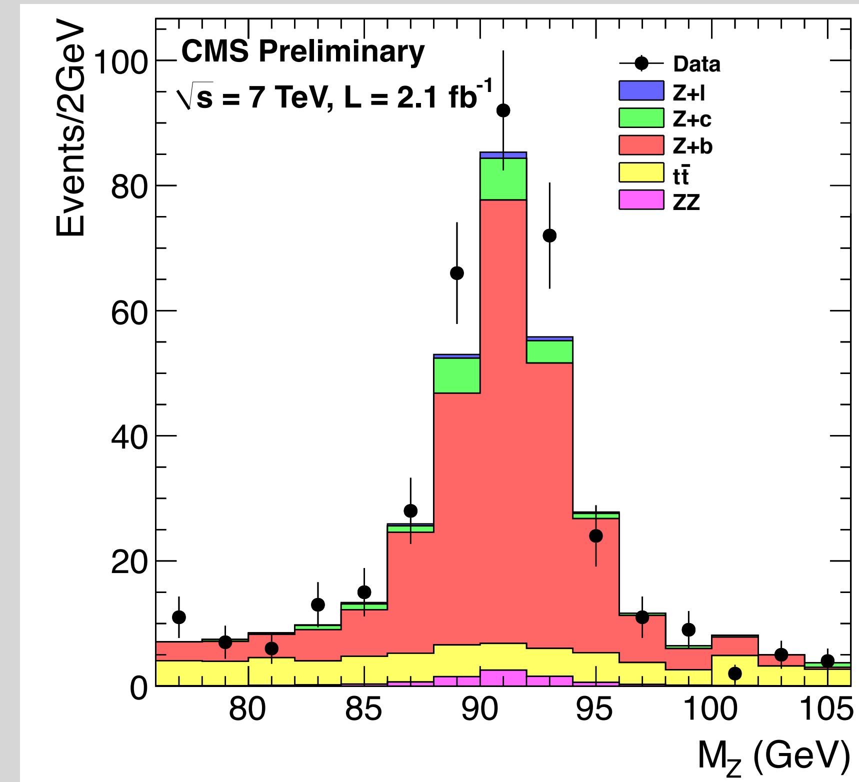
2 b-jets id, assuming dilepton selection



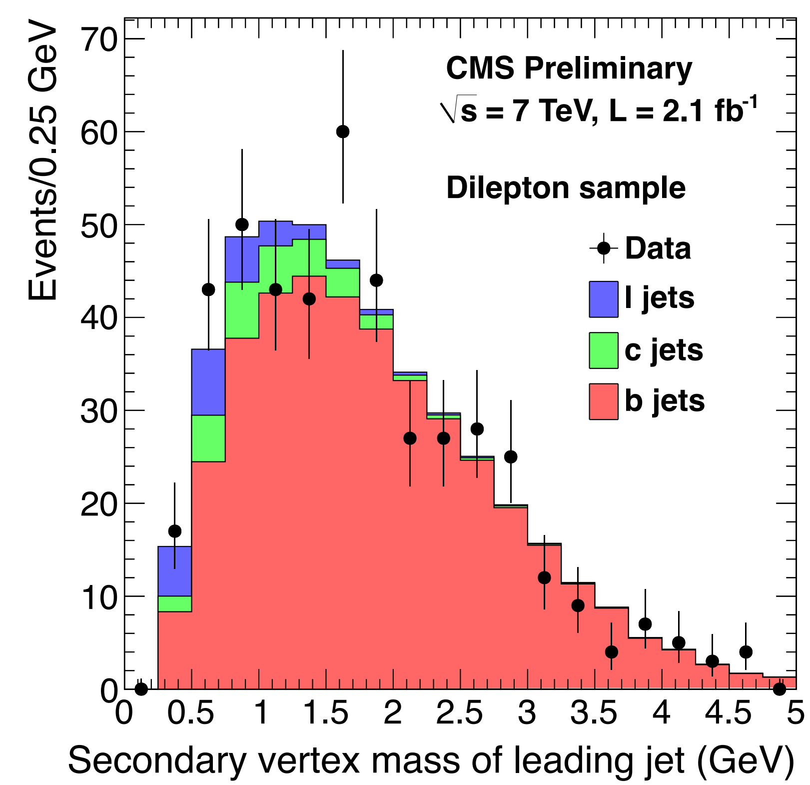
tt suppr., assuming dilepton+2 b-jets id.



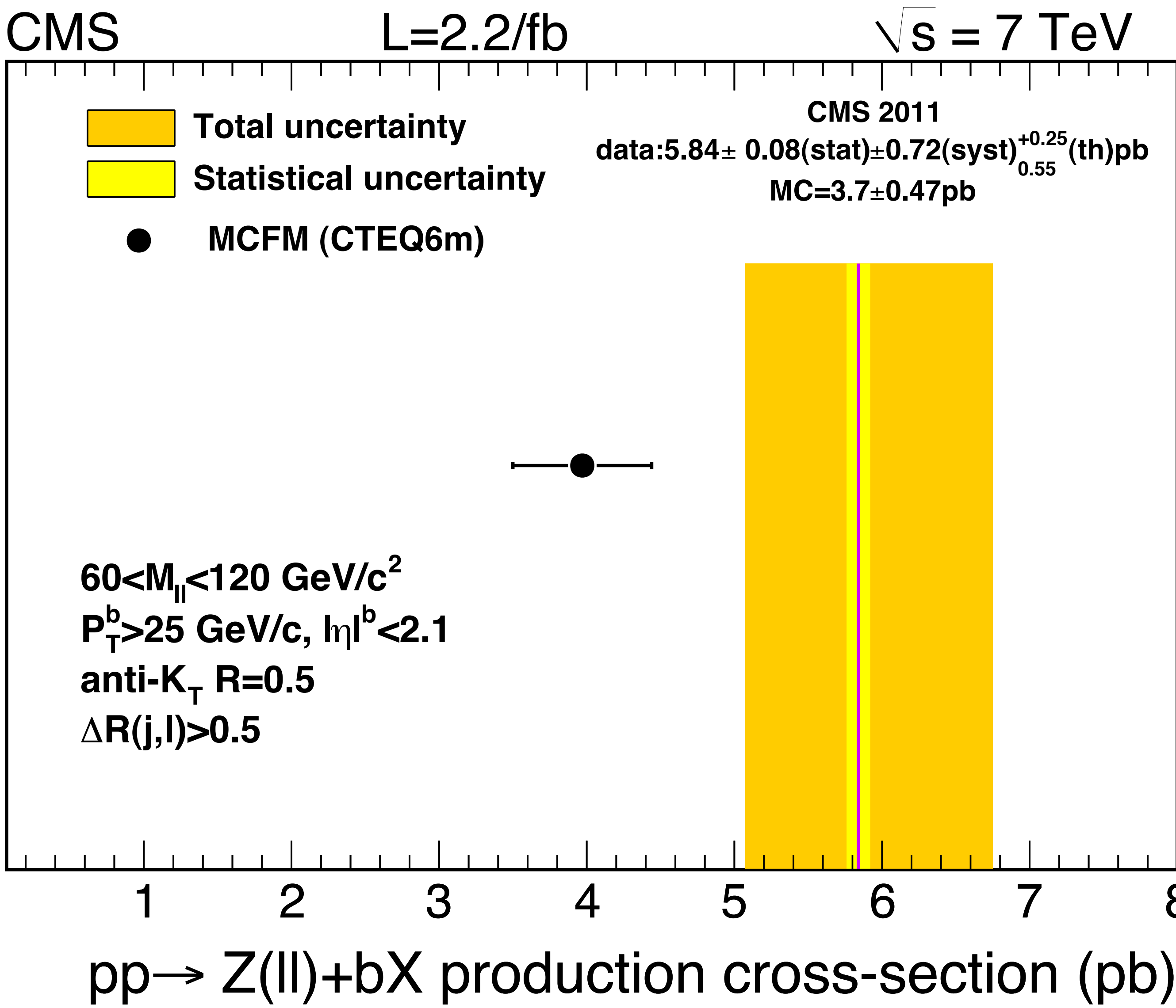
M(ll), assuming 2 b-jets + MET cut

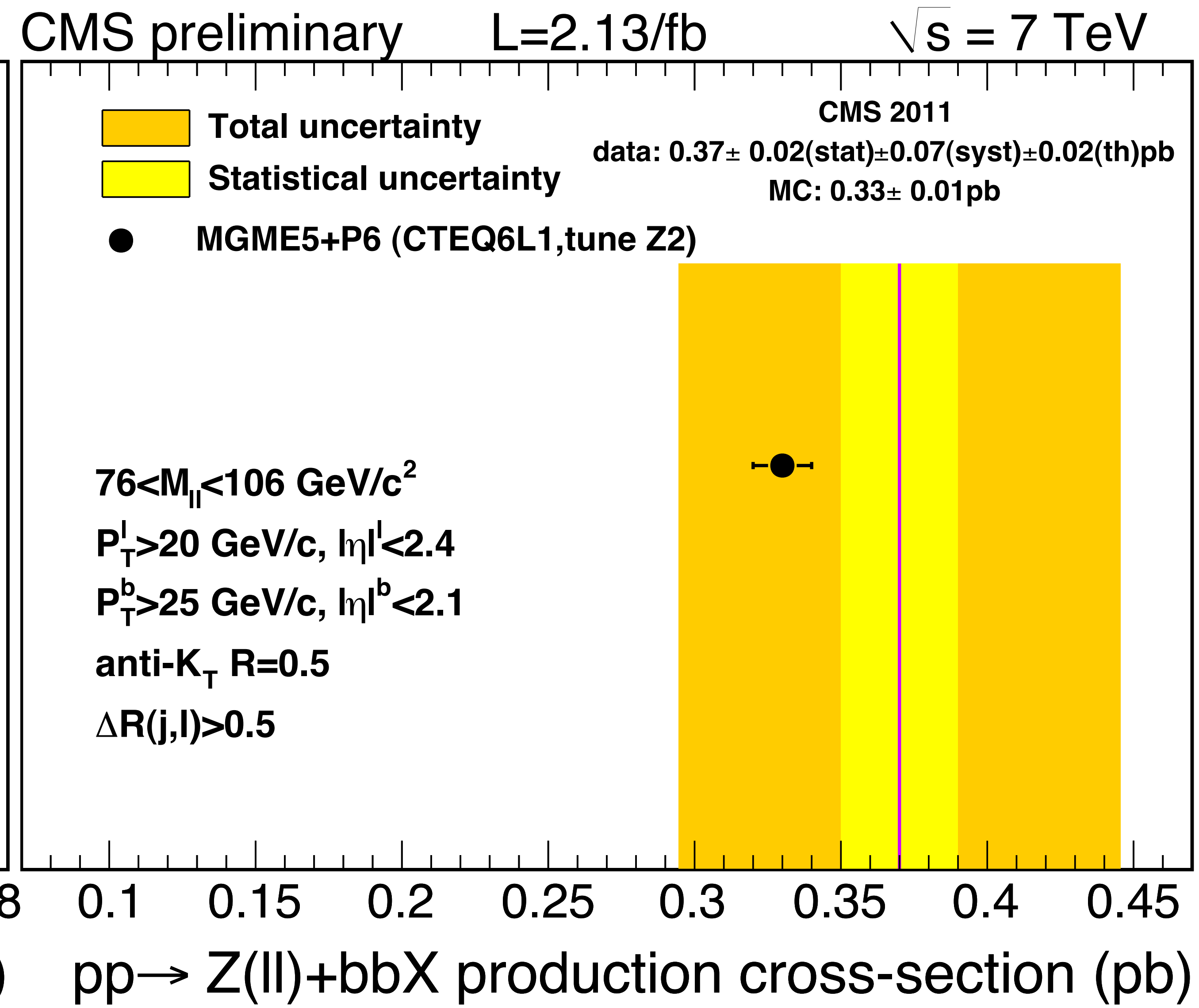


- Correct for
  - for ttbar contamination
  - purity →
  - lepton acceptance and efficiencies
  - b-tagging efficiency
  - detector and reconstruction effects



Z+b-jet(s): cross-section measurement



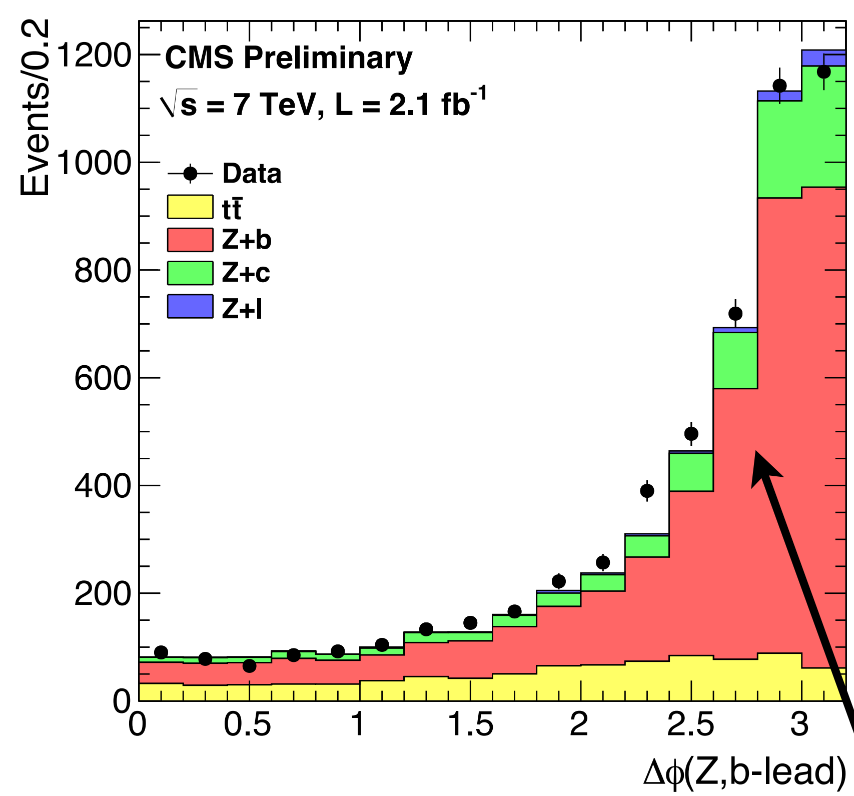


Z+bX:tension between MCFM and data
Z+bbX: good agreement with Madgraph

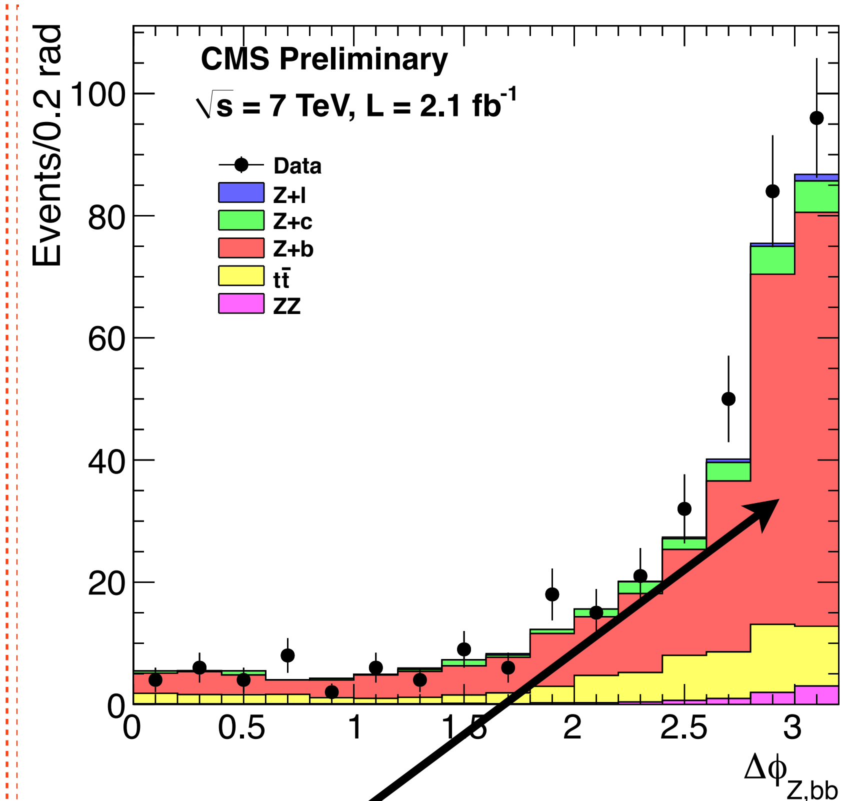
Z+b-jet(s): tensions

Distributions, before any correction

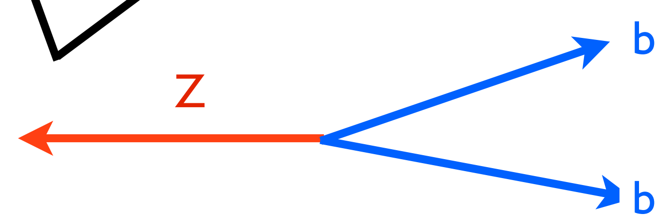
Z+>=1b



Z+2b



Data seems to favorize a larger

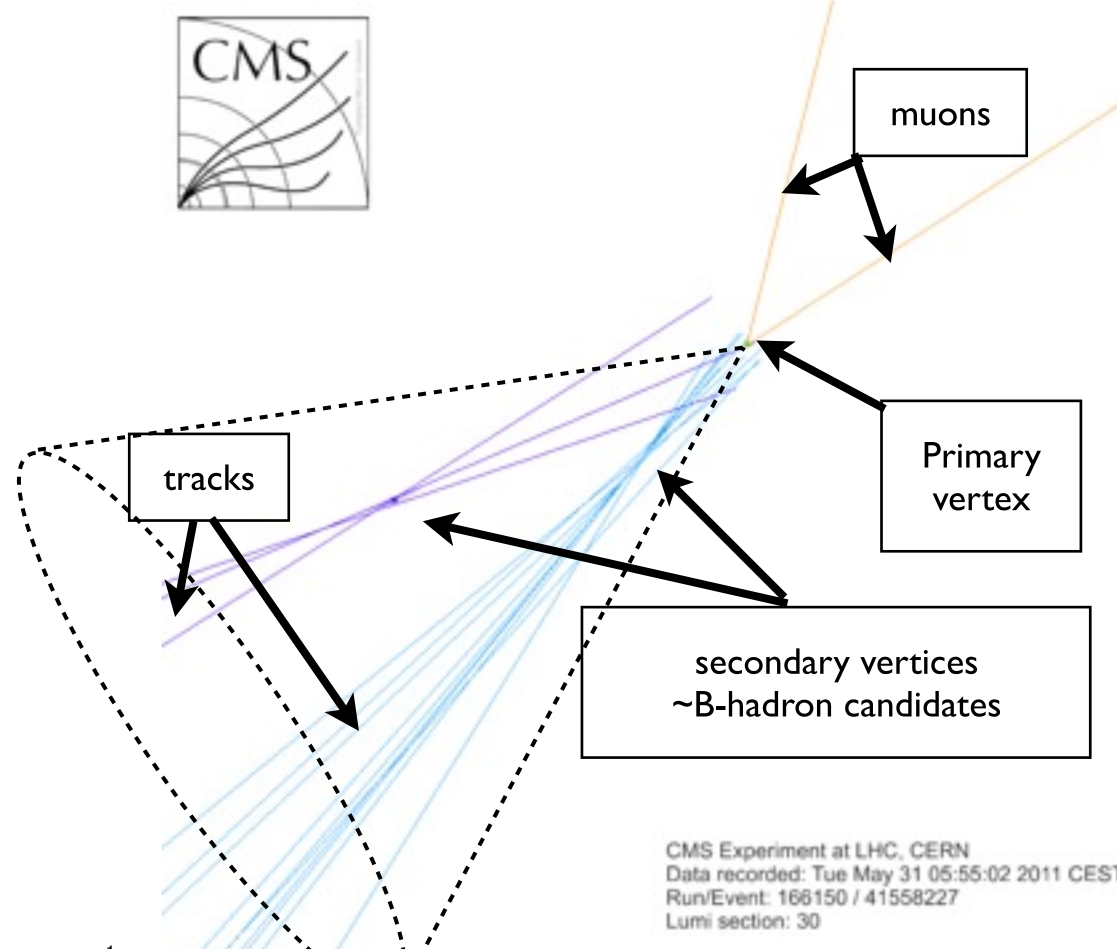


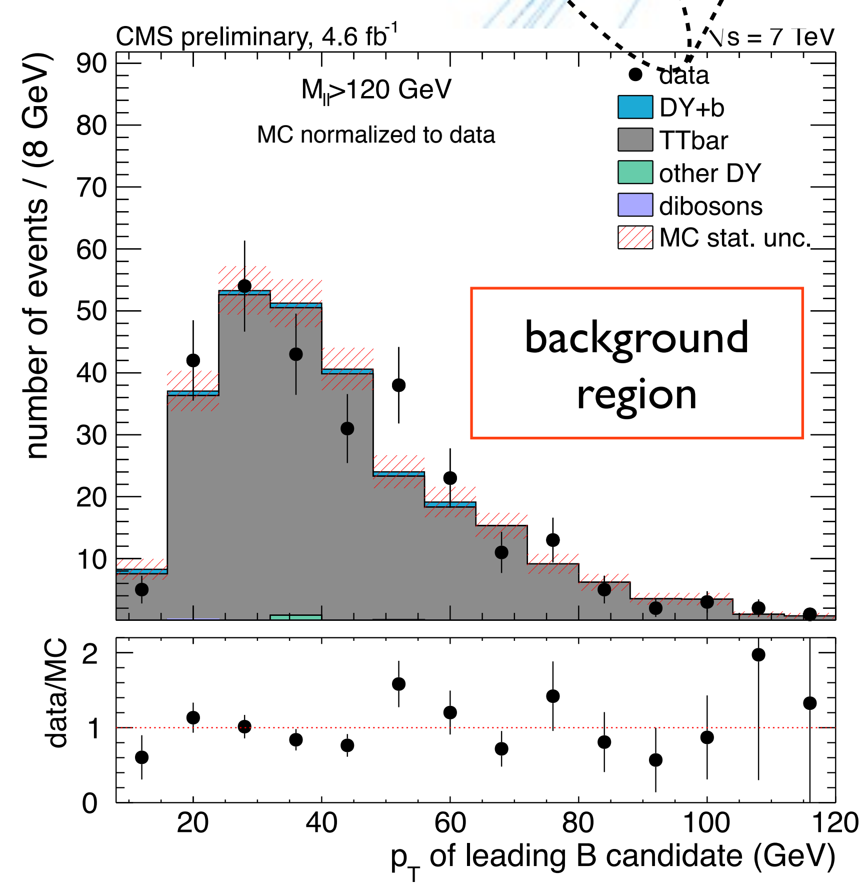
contribution

Here: Madgraph **5F**, need to compare with 4F prediction.

Z+bb: angular correlation

- Measure also «gluon splitting» bb production: in collinear configuration.
  - **No use of jet** because of radius limitation (~0.5-1.0).: use the **I**nclusive **V**ertex **F**inder (IVF)
    - Seed: tracks with large impact parameter,. Other tracks clustered around the seed
    - Vertex fitter: makes Secondary Vertices, merge if B→D-like decay (produces B-candidate)
    - Resolution $\Delta R(B_1,B_2)$~0.02!
    - Draw back: no absolute IVF efficiency
      - ⇒ no data/MC scale factors
      - ⇒must normalize results arbitrarily





Z+bb angular correlation: ΔR(B,B)

Phase-space definition

Pt(l)>20 GeV, |η|<2.4.
60<M(l^+l^-)<120 GeV
Pt(B-hadron)>15 GeV , |η|<2.0

Fit the number of Z+x vs TTbar

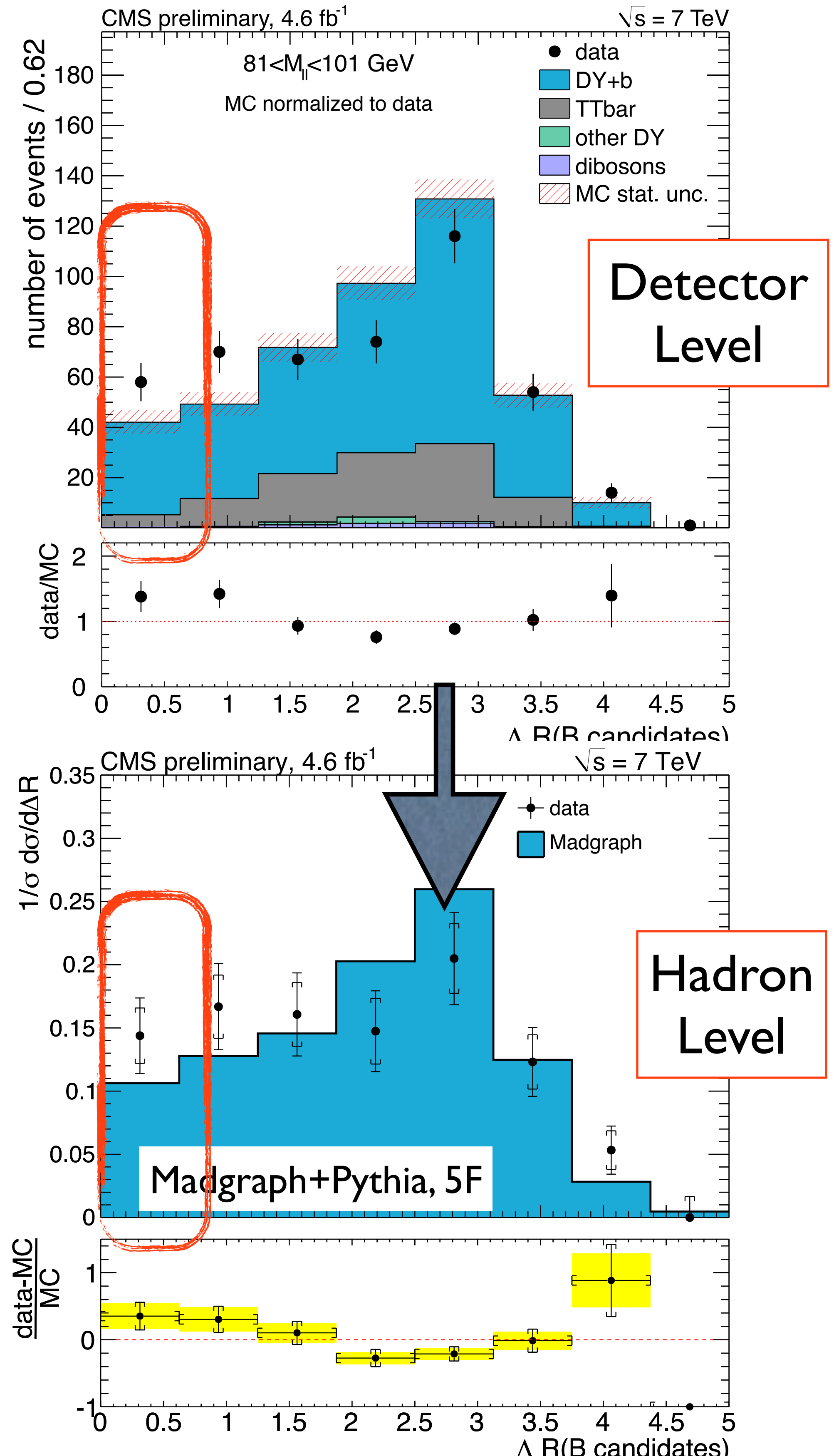
Purity

$$\sigma_{\text{visible}} = \sum_{i=1}^{8} \frac{N_i^{data,fit} \cdot \mathcal{P}_i}{\epsilon_i^{2SV} \cdot \epsilon_i^{\ell} \cdot \mathcal{A}^{\ell}}$$

IVF efficiency, assuming HPS

Lepton acceptance and efficiency

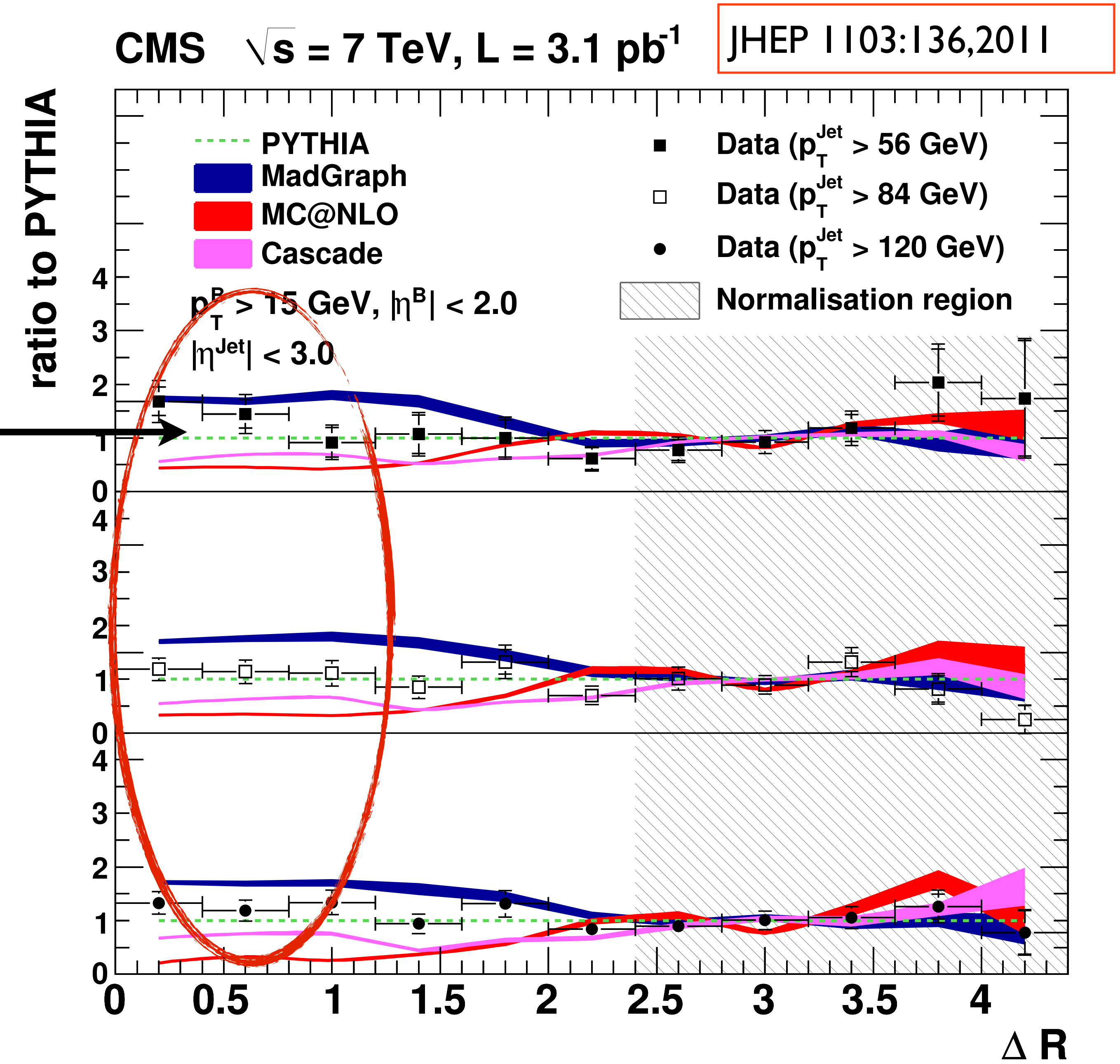
Data exhibits a flatter shape than Madgraph 5F.
Very collinear region probed for the first time!



BB→μμX cross-section measurement

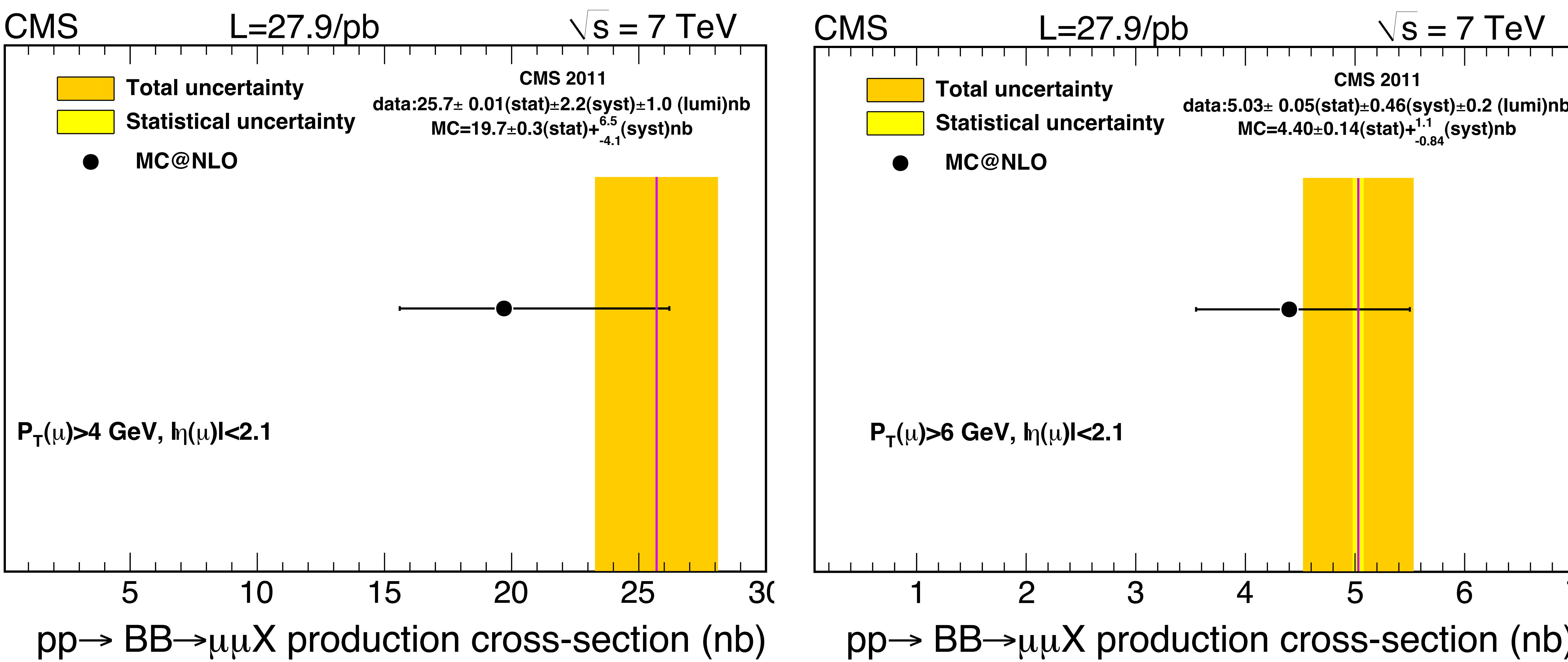
- Previous CMS studies about b and bb production (cross-section and angular correlation
  - MC@NLO fails at describing collinear configuration

→how about very soft P_T(b-hadron) configuration where gluon splitting contribution is small?



Production cross-section of BB→μμ events

- BB→μμX
  - allows to probe very soft region of the phase-space, where NLO predictions are less sensitive to gluon splitting contribution (and uncertainties)
  - high quality muons with >11 track hits, >1 pixel hits/track, $\chi^2<2$ and potentially large impact parameter with respect to primary vertex



MC@NLO in good agreement with data

Conclusion

- The CMS experiment has already provided and continues to perform HF/V+HF analyses in the context of SM studies.
  - W+c study had allowed to estimate the W+c production rate and charge independance.
  - Inclusive Z+b(b) cross-section has been measured with 2.1/fb and is found to be larger than MCFM predictions
  - Zbb angular correlation show discrepancies with Madgraph prediction in 5F scheme and aMC@NLO (4F). No possibility to conclude yet about MPI impact
  - BB→μμ study allows to probe very soft production of b-hadrons, the MC@NLO prediction is in agreement with data

Back-up slides

W+c

Sample	Muon	$M_T > 50\,\mathrm{GeV}$ & DY veto	≥ 1 jet $p_T^{jet} > 20\,\mathrm{GeV}$, $\lvert\eta^{jet}\rvert < 2.1$	< 3 jets with $p_T^{jet} > 40\,\mathrm{GeV}$ $\lvert\eta^{jet}\rvert < 2.1$	$N_{tk}^{vtx} \geq 2$	$\Delta l <$ 0.15 cm
Events with μ^+	112292	80259	18237	18010	768	636
Events with μ^-	83451	55749	13948	13725	653	563
Events with $\mu^\pm$	195743	136008	32185	31735	1421	1199

Source	$N_{bg}/(N_{sg}+N_{bg})$	N_{bg} in 36 pb^{-1}
$W + udsg$	$24.5 \pm 0.6\%$	220 ± 4
$W + b$	$4.1 \pm 0.2\%$	37 ± 2
$t\bar{t}$	$14.8 \pm 0.2\%$	133 ± 1
single-t	$5.2 \pm 0.1\%$	47 ± 1
$W \to \tau\nu$	$1.3 \pm 0.1\%$	12 ± 1
$Z \to \mu^+\mu^-$	$1.9 \pm 0.1\%$	17 ± 1
$Z \to \tau^+\tau^-$	$0.1 \pm 0.1\%$	1 ± 1
QCD	$1.1 \pm 0.1\%$	10 ± 1
Total Bckg.	$53.0 \pm 0.7\%$	476 ± 5
	$N_{sg}/(N_{sg}+N_{bg})$	N_{sg} in 36 pb^{-1}
$W + c$	$47.0 \pm 0.8\%$	423 ± 6

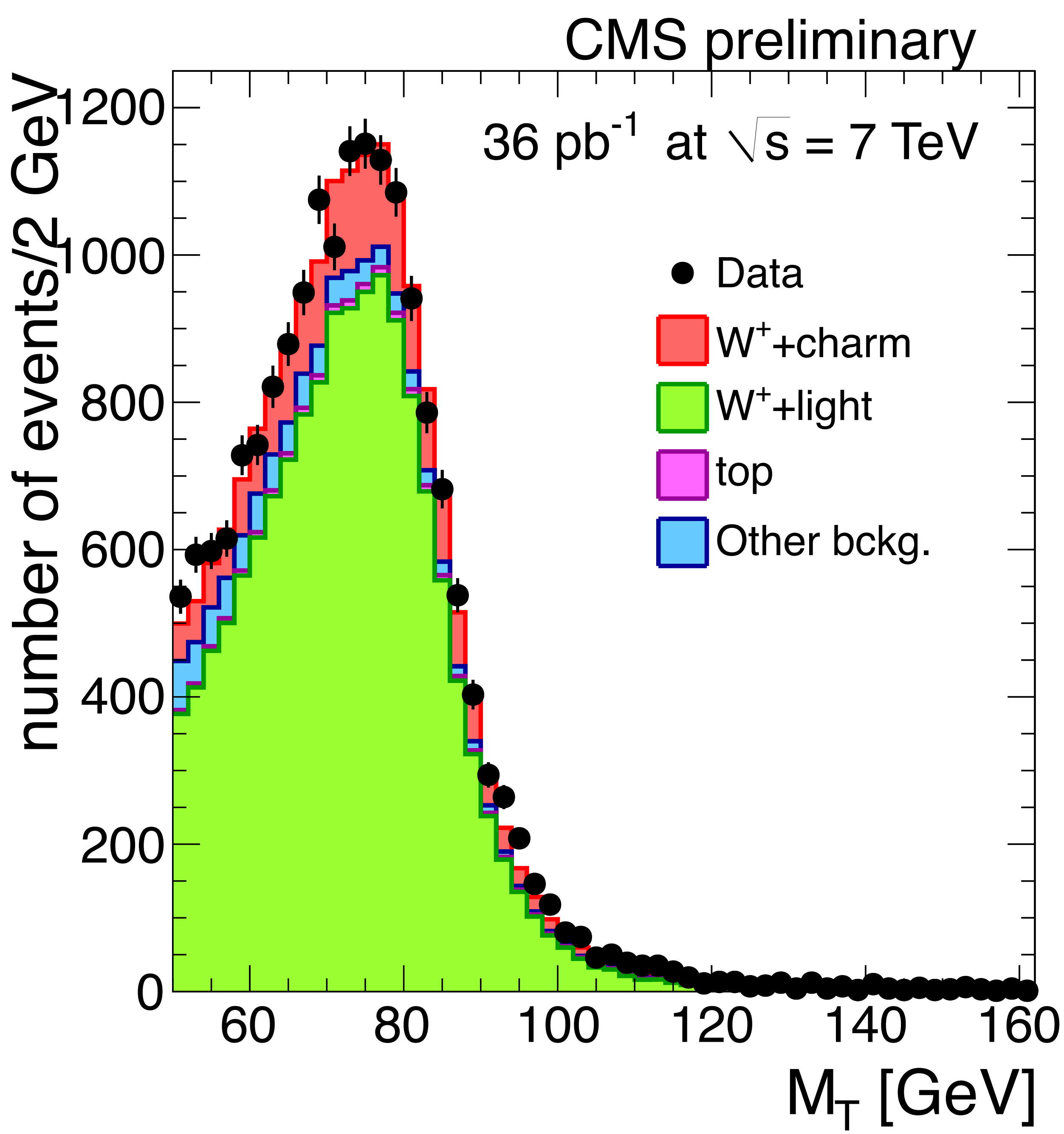


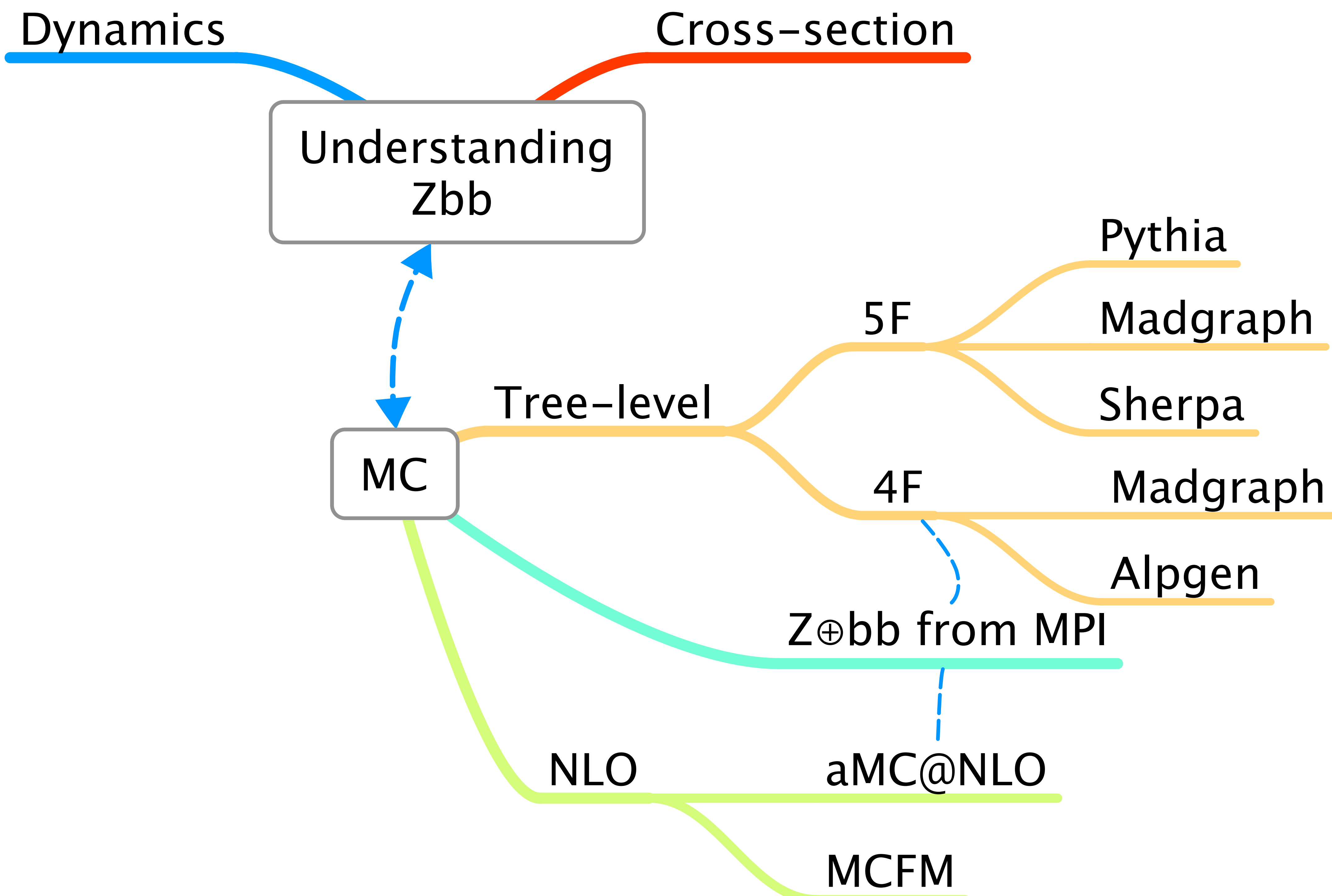
Table 3: Relative systematic uncertainties (%) in the measurement of *l*

Source	Relative uncertainty (%)
Charge asymmetry in efficiency	1.0
Muon resolution	<0.1
Pile-up effects	1.8
Jet energy scale/resolution	1.1
Jet multiplicity	0.7
Vertex reconstruction	0.3
Top templates	1.7
Light-quark contribution	1.1
W+b background	0.2
Other Monte Carlo backgrounds	1.4
PDF uncertainties	2.2
Charm fragmentation function	<0.1
Charm fragmentation BRs	0.1
TOTAL	4.1

Table 4: Relative systematic uncertainties (%) in the measurement of

Source	Relative uncertainty (%)
Charge asymmetry in efficiency	-
Muon resolution	0.7
Pile-up effects	2.5
Jet energy scale/resolution	2.3
Jet multiplicity	2.5
Vertex reconstruction	14.1
Top templates	6.2
Light-quark contribution	3.3
W+b background	2.4
Other Monte Carlo backgrounds	0.2
PDF uncertainties	0.2
Charm fragmentation function	0.2
Charm fragmentation BRs	0.2
TOTAL	16.5

Z+HF quarks


Dynamics
Cross-section
Understanding Zbb
MC
Tree-level
5F
Pythia
Madgraph
Sherpa
4F
Madgraph
Alpgen
Z⊕bb from MPI
NLO
aMC@NLO
MCFM

Z+b-jet(s): cross-section measurement

Phase-space definition

Pt(b-jet)>25 GeV, |η|<2.1

Pt(l)>20 GeV, |η|<2.4.

76<M(l^+l^-)<106 GeV (2 b-jet)

ΔR(jet,$l^{\pm}$)>0.5

Z+1,2 b

- 1b → 3.41±0.05(stat)±0.27(syst)±0.06(th)
- 2b → 0.37±0.02(stat)±0.07(syst)±0.02(th)
- Σ → 3.78±0.05(stat)±0.31(syst)±0.11(th)

Madgraph
0.33±0.01(stat)

Extending phase-space correction from
76<M_{ll}<106 GeV
to
60<M_{ll}<120 GeV
+no cut on leptons Pt and eta

→ 5.72±0.09(stat)±0.47(syst)±0.39(th) (ext.)

MCFM
3.97±0.47pb

Z+b-jet(s): cross-section measurement

Two independent measurements
- Z+>=1 b-jet
- Z+(1,)2 b-jet

MonteCarlo

Z+b(s): Madgraph, MCFM

Z+light: Madgraph

top: Madgraph

diboson: Pythia

Detector-level definition		
	Z+>=1b-jet	Z+2 b-jet
Lepton	$Pt^{e}>25$ GeV, $\lvert\eta\rvert^{e}<2.4$ $Pt^{\mu}>20$ GeV, $\lvert\eta\rvert^{\mu}<2.1$	
Z	$60<M(l^{+}l^{-})<120$ GeV	$76<M(l^{+}l^{-})<106$ GeV
b-jets	High purity Simple Secondary Vertex 3 tracks used: eff~35%, udscg-mistag~0.1%	High efficiency Simple Secondary Vertex 2 tracks used: eff~55%, udscg-mistag~1%

Z+>=1b

$$\sigma_{hadron}(Z/\gamma^{*}+b, Z/\gamma^{*}\to \ell\ell) = \frac{N_{\ell\ell+b}\times(\mathcal{P}-f_{t\bar{t}})}{\mathcal{A}_{\ell}\times\mathcal{C}_{hadron}\times\epsilon_{\ell}\times\epsilon_{b}\times\mathcal{L}}$$

N selected $l^{+}l^{-}b$ → $N_{\ell\ell+b}$

Purity → $\mathcal{P}$

tt fraction → $f_{t\bar{t}}$

Acceptance → $\mathcal{A}_{\ell}$

Correction factor for detector and reconstruction effects → $\mathcal{C}_{hadron}$

efficiencies → $\epsilon_{\ell}\times\epsilon_{b}$

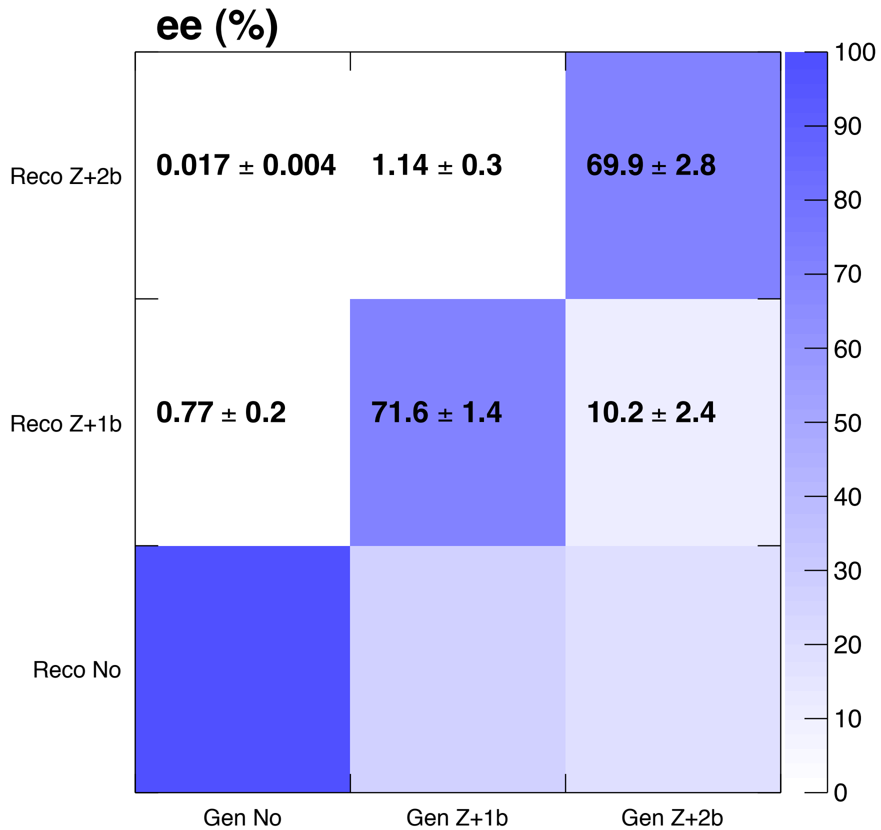
Z+2b

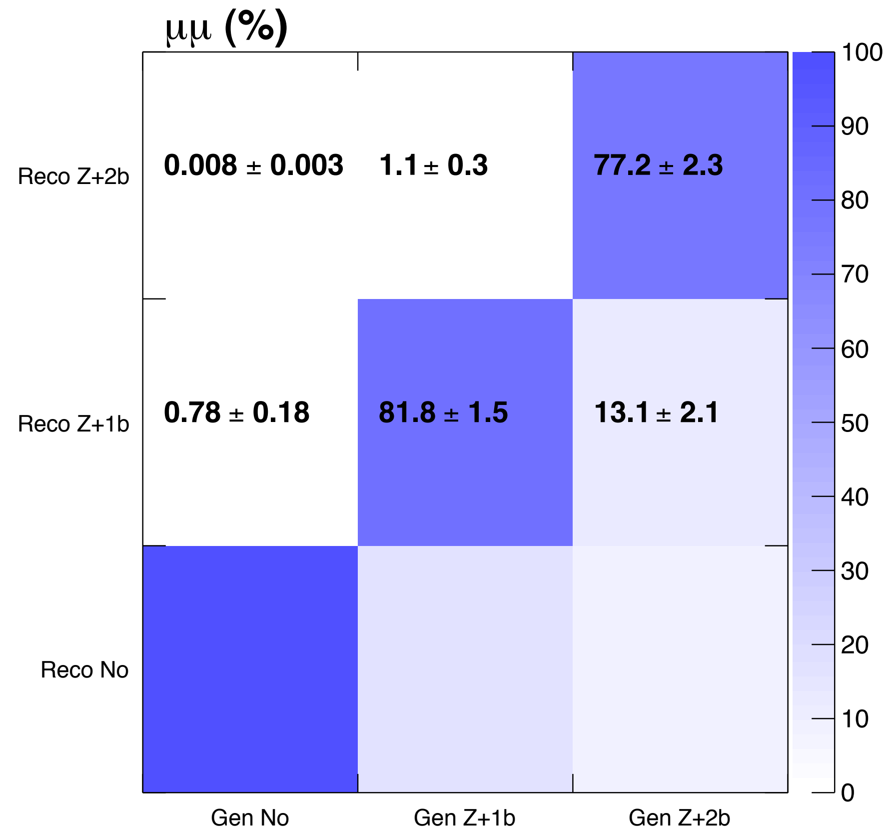
$$N^{sig}_{Z(ll)+bb} = N_{Z(ll)+bb}\times(f_{bb}-f_{t\bar{t}}) - N_{ZZ}$$

$$f_{bb} = 1 - f_{cc} - f_{bl} - f_{lb}$$

b-tagging efficiency correction: uses unfolding matrix

Z+1,2 b-jet(s) efficiencies





Migrations for b-tagging (%)	ee_{sel} + 1 b reco	ee_{sel} + (2+) b reco
ee_{sel} + (2+) b tag	-	26.8 ± 2.2
ee_{sel} + 1 b tag	49.7 ± 2.3	49.5 ± 2.4

Migrations for lepton selection (%)	ee + 1 b reco	ee + (2+) b reco
ee_{sel} + 1 b reco	62.9 ± 2.7	-
ee_{sel} + (2+) b reco	-	63.7 ± 2.9

Table 4: Migration factors for the b-tagging and lepton selection efficiencies, for the muon channel. Notations are the same as explained in the caption of Table 3, with muons instead of electrons.

Migrations for b-tagging (%)	$\mu\mu_{sel}$ + 1 b reco	$\mu\mu_{sel}$ + (2+) b reco
$\mu\mu_{sel}$ + (2+) b tag	-	26.2 ± 2.2
$\mu\mu_{sel}$ + 1 b tag	49.7 ± 2.3	50.4 ± 2.0

Migrations for lepton selection (%)	$\mu\mu_{sel}$ + 1 b reco	$\mu\mu_{sel}$ + (2+) b reco
$\mu\mu$ + 1 b reco	84.7 ± 1.6	-
$\mu\mu$ + (2+) b reco	-	84.0 ± 1.7

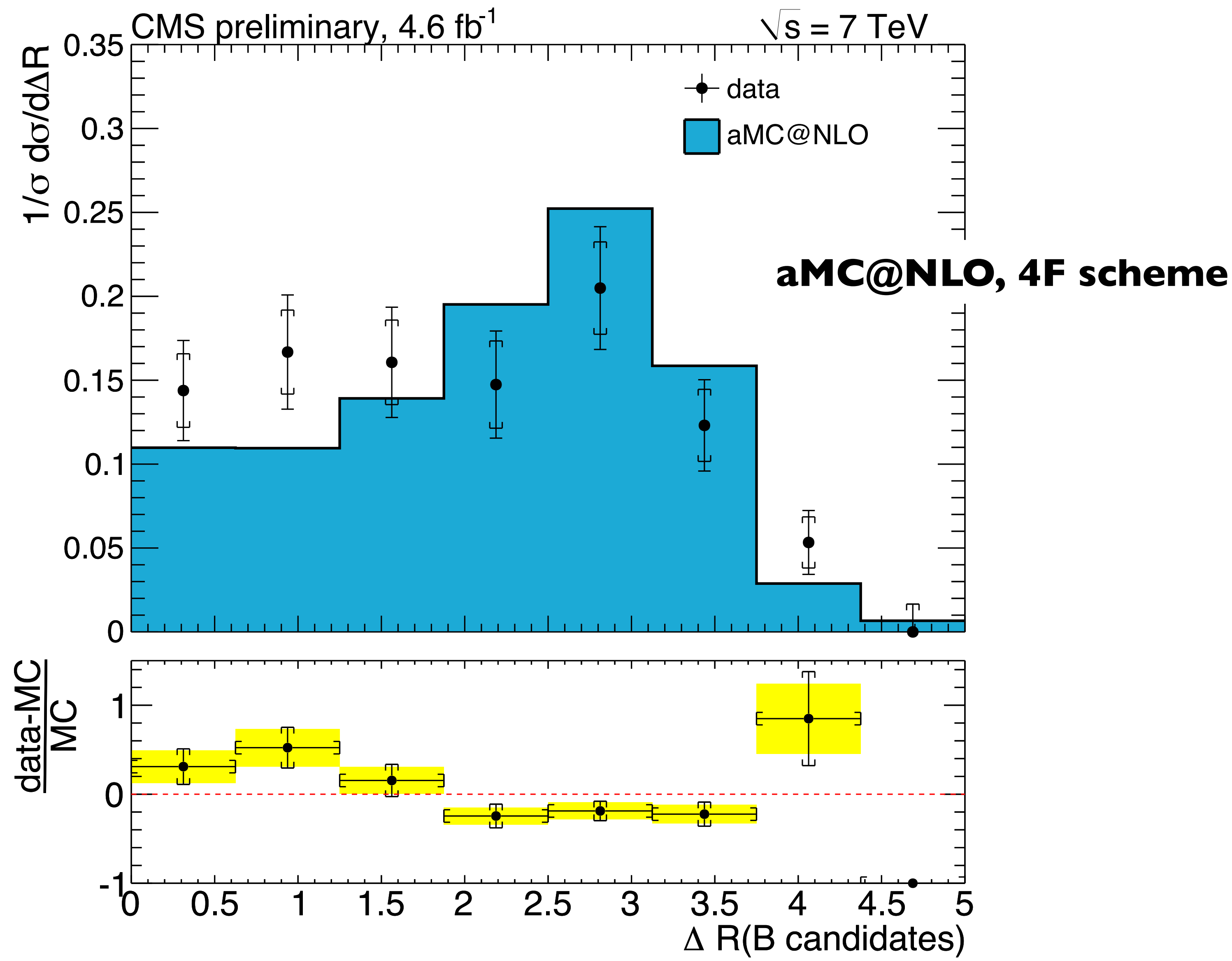
Z+b-jet(s) systematics

Z+>=1b

Correlated sources	Fractional uncertainty (%)	
b-tagging efficiency	10	
b-jet purity	5.6 (ee+b)	4.6 ($\mu\mu$+b)
$t\bar{t}$ contribution	2.9	
Jet energy scale	2.5	
Luminosity	2.2	
Jet energy resolution	0.5	
Pile-up	1.5 (ee+b)	0.5 ($\mu\mu$+b)
Mistagging rate	0.04	
Theory (via $\mathcal{A}_\ell$)	$^{+4.2}_{-6.5}$	
Theory (via $\mathcal{C}_{\text{hadron}}$)	$^{+0.7}_{-6.9}$	
Uncorrelated sources	**ee+b**	**$\mu\mu$+b**
Trigger and dilepton selection	4	2
$t\bar{t}$ contribution	1.9	2.2
Experimental systematic	13.0	12.3
Theoretical systematic	$^{+4.2}_{-9.5}$	$^{+4.2}_{-9.5}$
Statistical	2.2	1.7

Z+1,2b

	ee(%)		$\mu\mu$(%)	
Correlated sources	Z+1b	Z+2b	Z+1b	Z+2b
b-jet purity	3.5	10.3	2.5	11.0
$t\bar{t}$ contribution	0.9	8.9	0.5	9.4
b-tagging efficiency	4.0	7.4	3.9	7.5
Jet energy scale	3.9	6.9	3.8	6.4
Luminosity	4.5	4.5	4.5	4.5
E_T^{miss} selection	0.3	2.4	0.3	2.4
Pileup	1.7	1.8	0.3	0.3
ZZ contribution	0.1	0.5	0.1	0.7
Jet energy resolution	0.1	0.2	0.1	0.1
Mistagging rate	0.02	0.08	0.02	0.07
Theory (via $\mathcal{A}_l$)	1.8	5.9	3.0	6.4
Uncorrelated sources	**Z+1b**	**Z+2b**	**Z+1b**	**Z+2b**
MC sample stat.	1.2	5.1	0.9	4.2
Dilepton selection	4.0	4.0	1.9	1.9
Statistical	2.4	10.0	1.8	8.2
Experimental systematic	9.1	18.9	7.7	18.8
Theoretical systematic	1.8	5.9	3.0	6.4



- A (probably naive) estimation of MPI impact on Zbb/Z+SV analyses

$$\sigma_{(Z+light)_{hard}+(b\bar{b})_{MPI}} = \frac{\sigma_{pp\to Z+light} \times \sigma_{pp\to b\bar{b}}}{\sigma_{eff}}$$

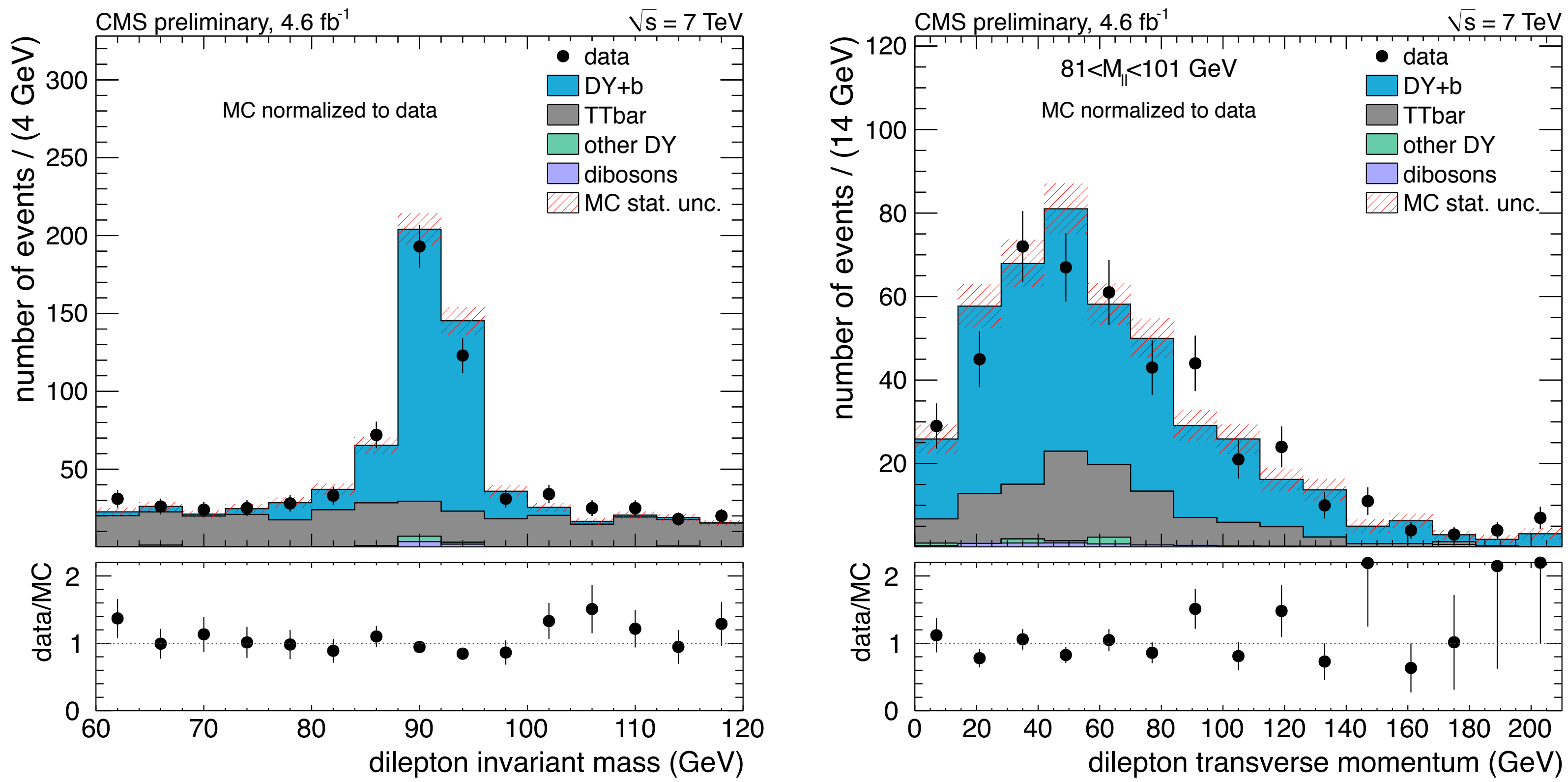
~1pb

~3nb

~12mb

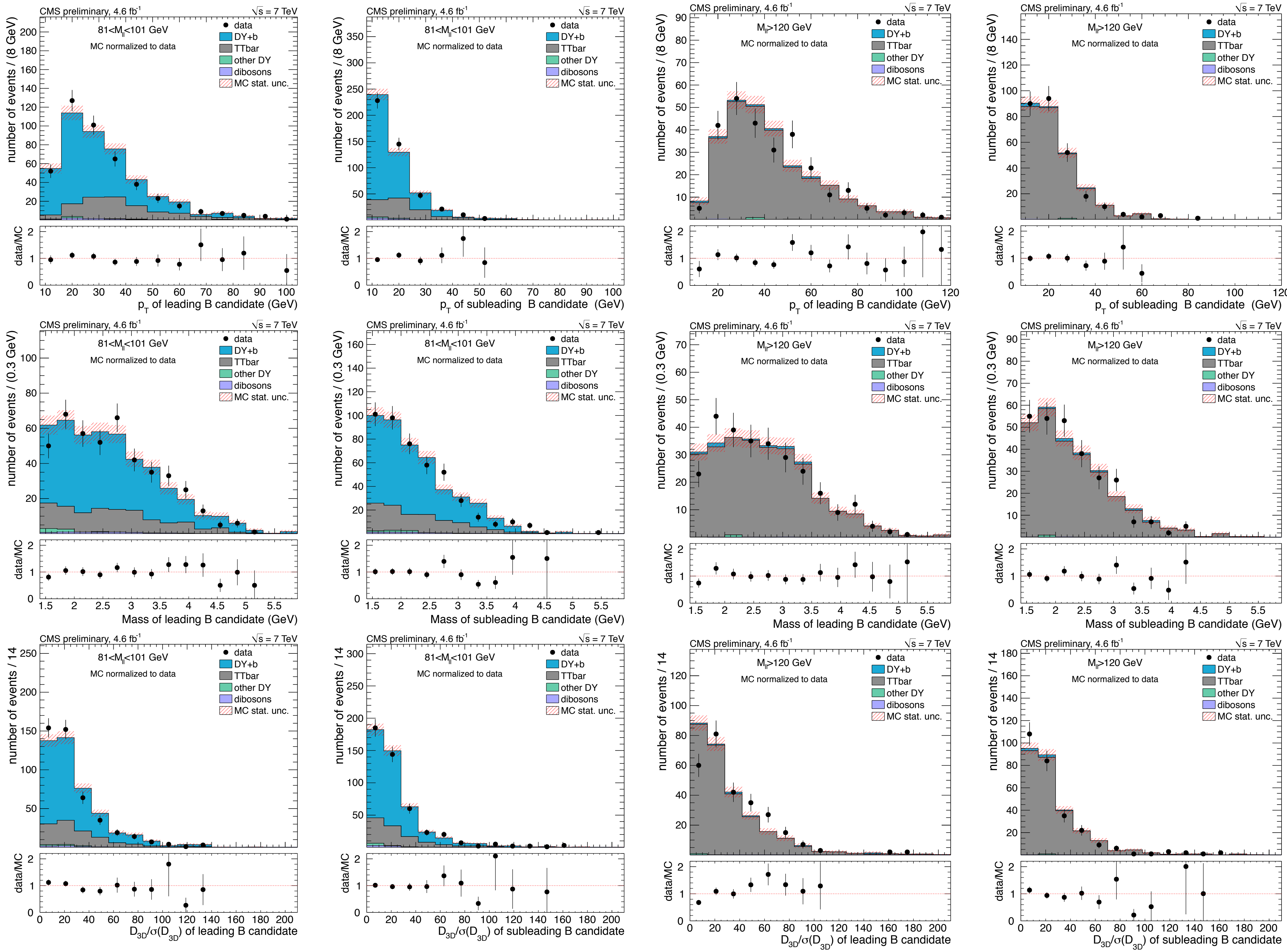
~5μb with PtB>15 GeV

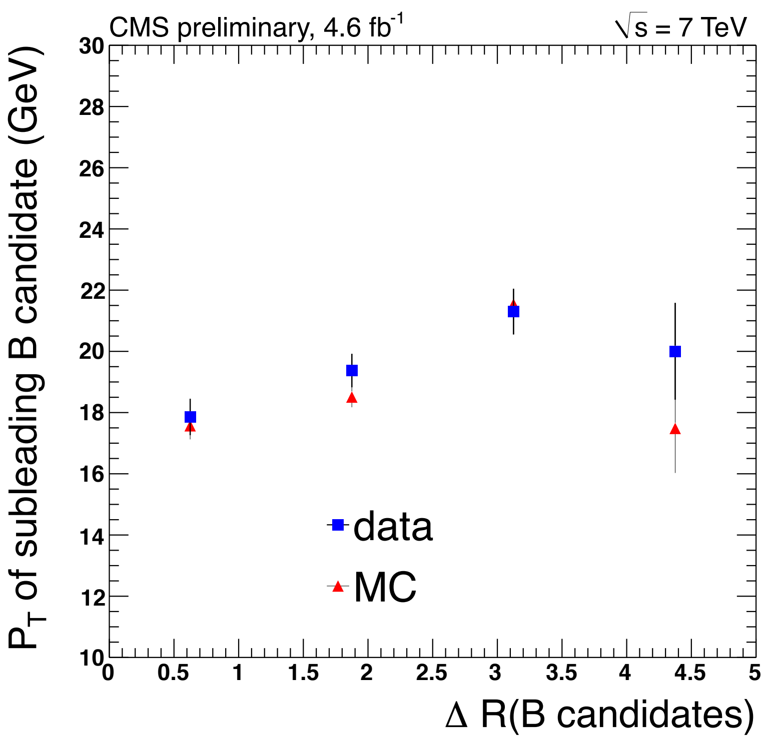
ZSV: yields and Z distributions

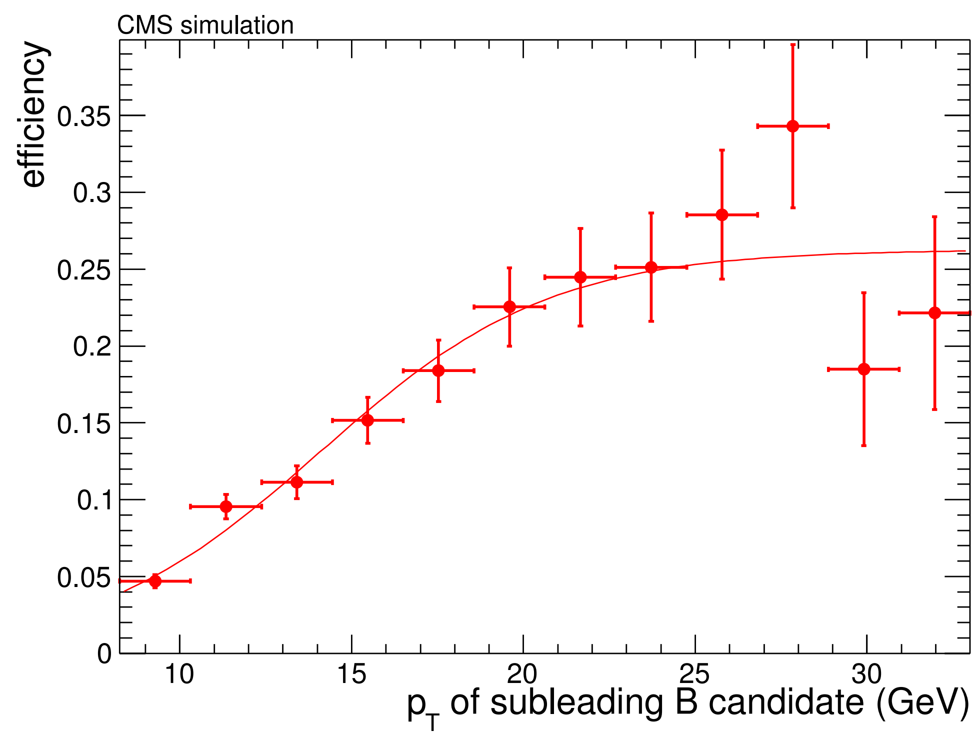


	$60 < M_{\mu\mu} < 120$ GeV	$M_{\mu\mu} \geq 120$ GeV	$60 < M_{ee} < 120$ GeV	$M_{ee} \geq 120$ GeV
data	454	154	254	120
DY+b	290.25 ± 11.73	3.90 ± 1.47	145.91 ± 8.36	0.69 ± 0.40
$t\bar{t}$	249.17 ± 7.34	167.21 ± 5.91	138.06 ± 5.53	115.21 ± 4.98
Other DY	3.43 ± 1.21	0.83 ± 0.58	3.87 ± 1.29	0.00 ± 0.00
WW	0.00 ± 0.00	0.18 ± 0.09	0.00 ± 0.00	0.12 ± 0.07
WZ	0.23 ± 0.07	0.00 ± 0.00	0.09 ± 0.04	0.00 ± 0.00
ZZ	5.67 ± 0.19	0.13 ± 0.03	2.78 ± 0.12	0.09 ± 0.02
sum MC	548.75 ± 13.89	172.24 ± 6.12	290.71 ± 10.1	116.11 ± 4.99
data/MC	0.83 ± 0.04	0.89 ± 0.08	0.87 ± 0.06	1.03 ± 0.10

ZSV: sanity checks







Source	Uncertainty
Softer B-hadron p_T and IVF phase-space correction	$\pm 9\%$
IVF purity	$\pm 4\%$
Fit uncertainty	$\pm(1\% - 2\%)$
Leptons kinematics	$\pm 0.5\%$
MC statistics	$\pm(6\% - 10\%)$