Measurements of associated production of vector bosons and heavy flavours with the CMS detector

Stefano Casasso[a,b]
for the CMS Collaboration

(a)INFN Sez. Torino,
(b)Universita' di Torino

QCD@Work 2012
Lecce, Italy
18-21 June 2012

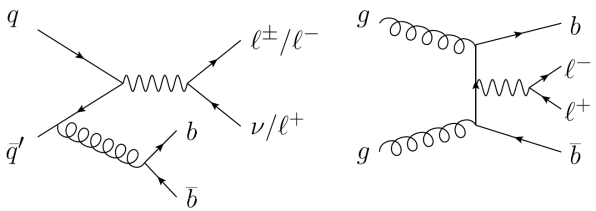
Outline

- The CMS experiment is well suited to look for V+heavy quarks signatures, thanks to its good b tagging performances and robust lepton reconstruction
- Analysis of V+heavy quarks performed in CMS:
  - $Z + b(b)$, jet based ($2.1\ \mathrm{fb}^{-1}$) (main focus of the talk)
  - $Z + bb$, secondary vertex based ($4.6\ \mathrm{fb}^{-1}$)
  - $W + c$ ($36\ \mathrm{pb}^{-1}$) (not inluded in the talk)
- Motivations for these channels:
  - $Z + b(b)$
    - Confirm/constrain NLO cross-section and kinematics predictions
    - Benchmark to $bb + \Phi$ discovery channel in MSSM with large $\tan\beta$
    - Background to Higgs boson searches (e.g. $H \to ZZ \to 4\ell$, $ZH(\to bb)$)
  - $W + c$
    - Put constraints on strange quark *PDF* of the proton in the intermediate x region

Motivations for detailed $Z + b(\bar{b})$ measurements

- Cross-section calculations: two different approaches:

Fixed flavor scheme
(arXiv:hep-ph/1106.6019)

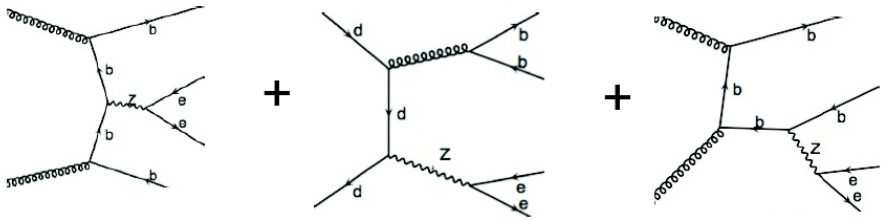


- Massive b
- Full event description $\longrightarrow$ aMC@NLO

Variable flavor scheme
(arXiv:hep-ph/0312024)

- Gluon splitting inside proton PDF
- Massless b $\longrightarrow$ Collinear approach

- Several production mechanisms at the LHC: b quarks fusion, gluon splitting, Z radiation, and Multiple Parton Interactions



- Angular correlations of bb pair $\longrightarrow$ comprehension of production mechanism

Cross section definition for $pp \to Z/\gamma^* + b$ (EWK-11-012)

The $pp \to Z/\gamma^* + b, Z/\gamma^* \to \ell\ell$ cross-section is calculated as:

$$\sigma_{hadron} = \frac{N_{\ell\ell+b} \times (\mathcal{P} - f_{t\bar{t}})}{\mathcal{A}_\ell \times \mathcal{C}_{hadron} \times \epsilon_\ell \times \epsilon_b \times \mathcal{L}}$$

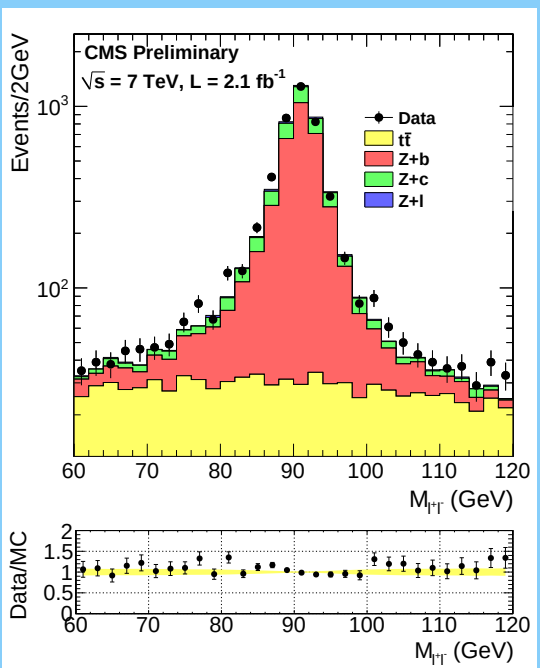
- $N_{\ell\ell+b}$: number of selected events
- $\mathcal{P}$: b-jet purity
- $f_{t\bar{t}}$: $t\bar{t}$ contamination
- ϵ_ℓ and ϵ_b: lepton and b-tagging efficiencies
- $\mathcal{A}_\ell$: lepton acceptance
- $\mathcal{C}_{hadron}$: detector and reconstruction effects
- $\mathcal{L}$: integrated luminosity

The cross section is defined with the following requirements:
(i) $\mathbf{p_T^b > 25}$ **GeV and** $\mathbf{|\eta^b| < 2.1}$
(ii) $\mathbf{60 < m_{\ell\ell} < 120}$ **GeV**
(iii) $\mathbf{\Delta R(jet, \ell) > 0.5}$

Event selection for $Z/\gamma^* + b$ jets

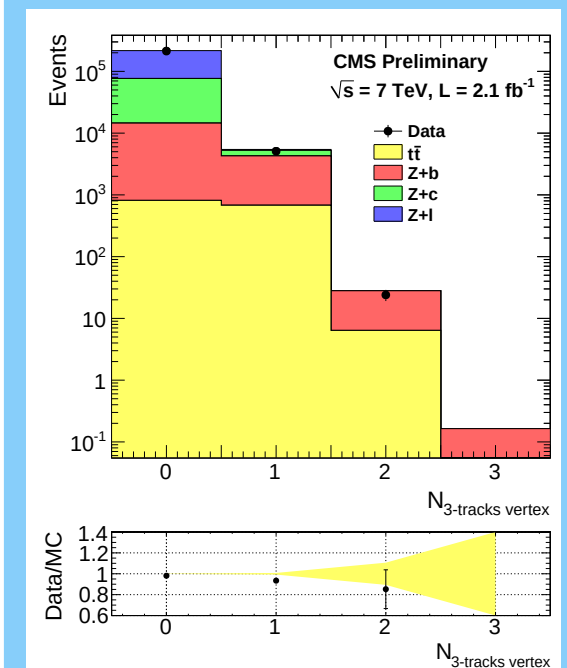
Z from isolated leptons:

- $p_T^{e(\mu)} > 25(20)$ GeV
- $|\eta_{e(\mu)}| < 2.5(2.1)$
- $60 < m_{\ell\ell} < 120$ GeV
- ID criteria + match with trigger objects



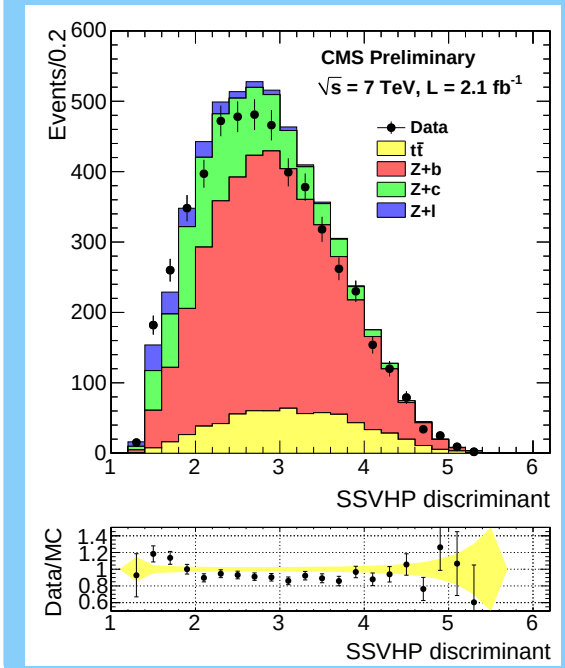
Jet selection:

- anti-$k_T(R = 0.5)$
- $p_T > 25$ GeV
- $|\eta| < 2.1$
- jet/lepton separation $\Delta R(\ell, jet) > 0.5$



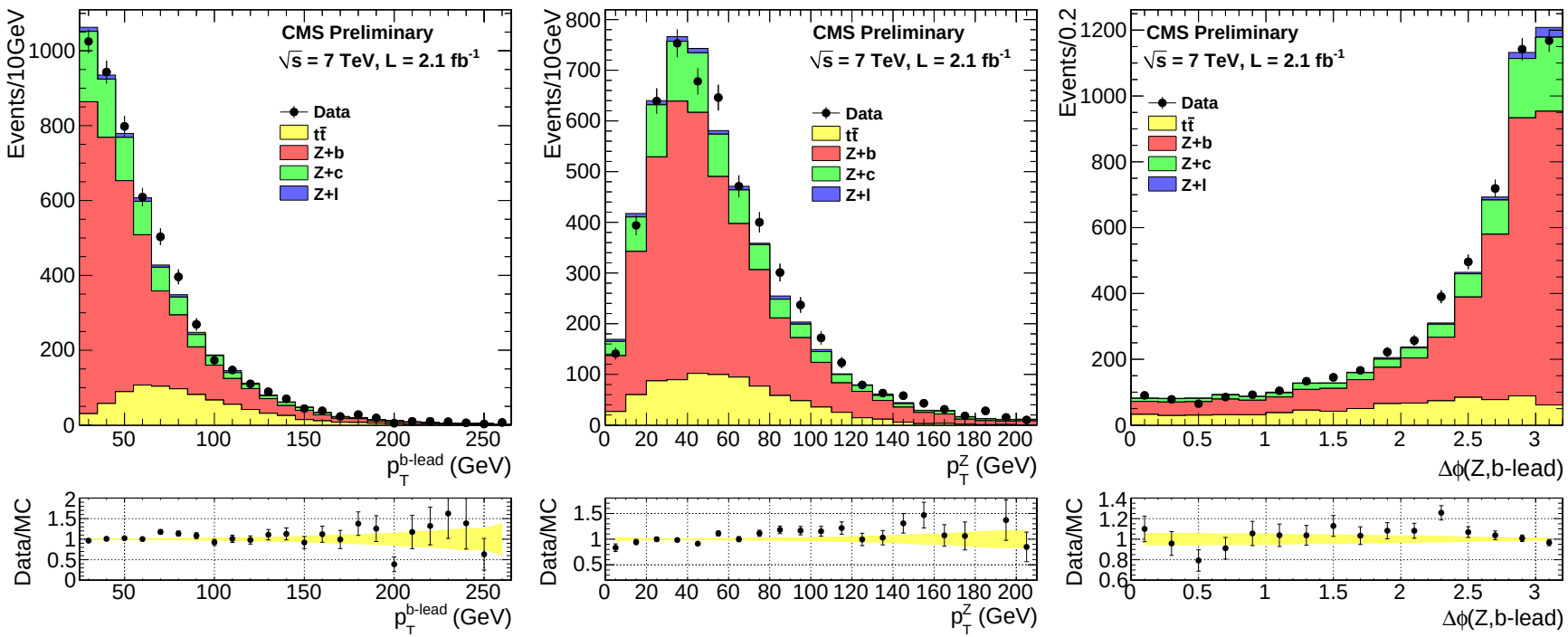
b-tagging:

- Detached secondary vertex (with at least 3 tracks)
- 1 b-tag eff. $\simeq 35\%$
- $udsg$-mistag $\simeq 0.1\%$



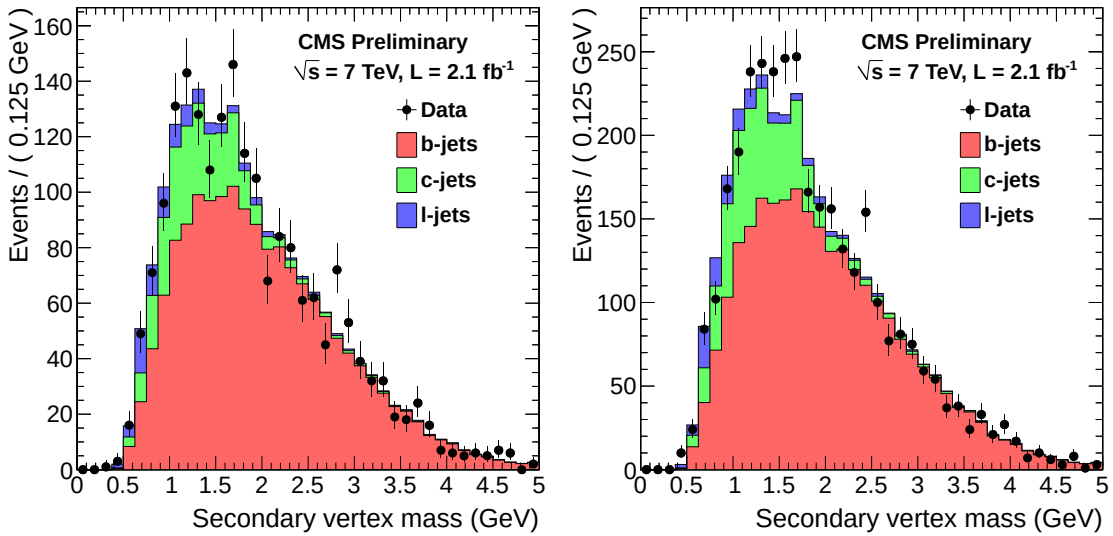
Data-MC comparison for $pp \rightarrow Z/\gamma^* + b$

- MC reweighted to match:
  - PU conditions in data
  - lepton and $b-\mathrm{tagging}$ efficiencies in data
- MC scaled to NLO cross-section
- Good agreement is found between Data and MC (`Madgraph`)
- *Some discrepancies in the shapes of kinematic variables*



Estimation of the backgrounds

- **Z + ucdsg**: extract the $b-$purity $\mathcal{P}$ with a (data-driven) template fit of the mass of the secondary vertex of the tagged jet



- **$t\bar{t}$**: extracted from extrapolation of upper sideband of $M(\ell\ell)$ under the signal region:

$$N^{est}_{t\bar{t}}(in) = \left(\frac{\mathcal{R}^{MC}_{t\bar{t}}}{\mathcal{R}^{MC}_{Z} - \mathcal{R}^{MC}_{t\bar{t}}} \right) \cdot (\mathcal{R}^{MC}_{Z} \cdot N_{obs}(out) - N_{obs}(in))$$

$$\mathcal{R}^{MC}_{t\bar{t}(Z)} = \frac{N^{MC}_{t\bar{t}(Z)}(in)}{N^{MC}_{t\bar{t}(Z)}(out)}$$

Extraction of the $Z/\gamma^* + b$ cross-section and comparison with theory

- Final cross-section calculated separately in the $Z \to \mu\mu$ and $Z \to ee$ channels:

	cross-section (pb)
$\sigma_{hadron}(Z + b, Z \to ee)$	$5.61 \pm 0.13\ (\text{stat}) \pm 0.73\ (\text{syst})\ ^{+0.24}_{-0.53}\ (\text{theory})$
$\sigma_{hadron}(Z + b, Z \to \mu\mu)$	$5.97 \pm 0.10\ (\text{stat}) \pm 0.73\ (\text{syst})\ ^{+0.25}_{-0.57}\ (\text{theory})$

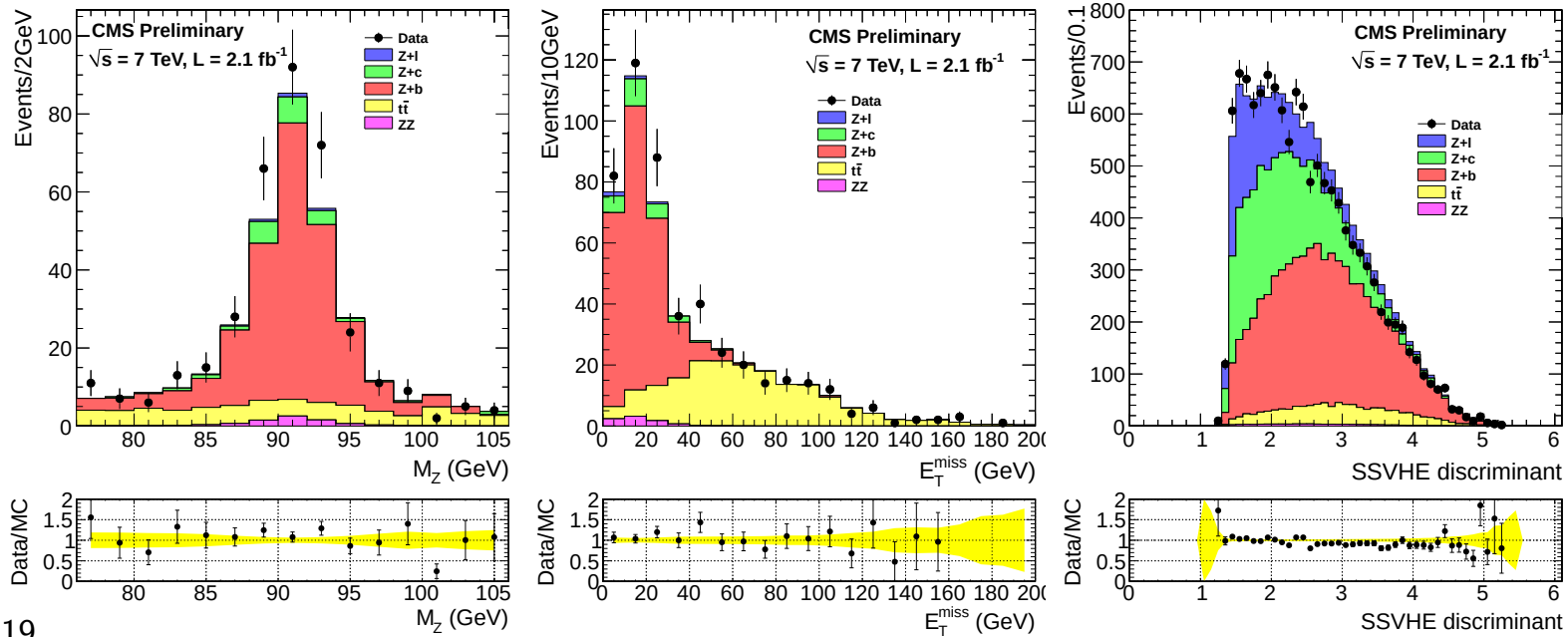
- Good agreement $\Rightarrow$ combine the two channels:

$$\sigma_{hadron}(Z + b, Z \to \ell\ell) = 5.84 \pm 0.08(stat) \pm 0.72(syst)^{+0.25}_{-0.55}(theory)\,\text{pb}$$

- MCFM NLO parton-level prediction:
  $\sigma^{parton}_{MCFM} = 4.73 \pm 0.54$ pb
- Parton-to-hadron level correction C_{NP}=0.84±0.03 is computed using `MADGRAPH+PYTHIA` and `aMC@NLO+HERWIG`:
  $\sigma^{hadron}_{MCFM} = 3.97 \pm 0.47$ pb
- *The measured cross-section is bigger than the hadron-level-corrected NLO prediction*

Event selection for $Z/\gamma^* + bb$ (SMP-12-003)

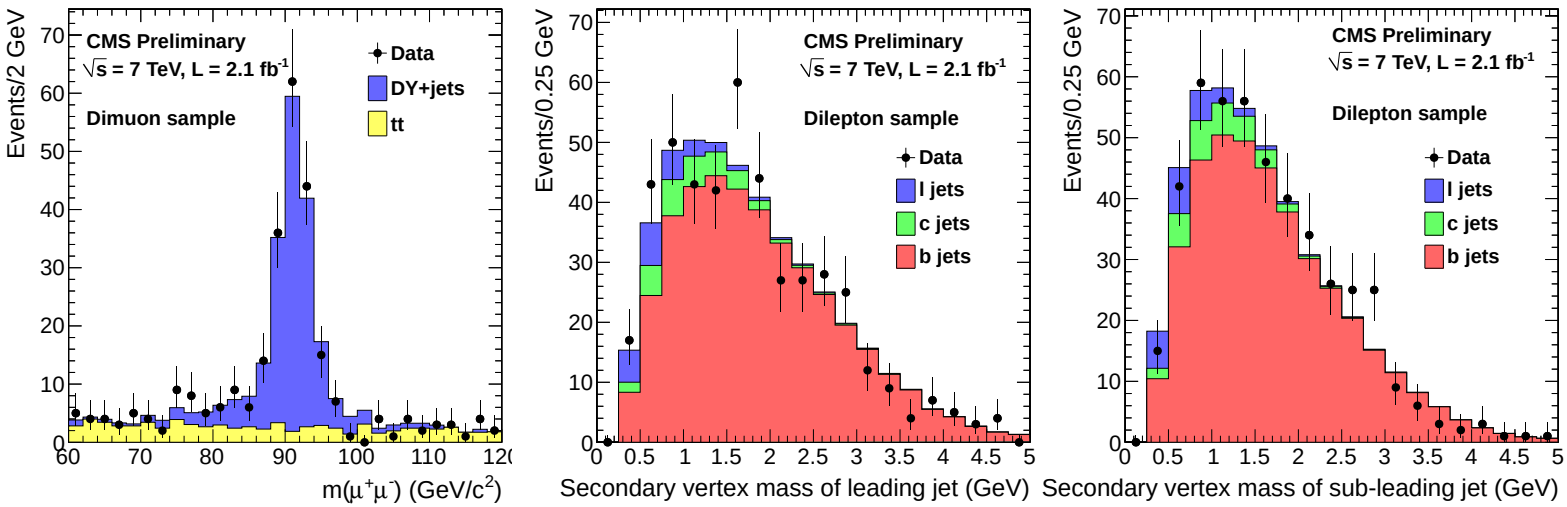
- Same lepton and jet selection as for the inclusive case
- Use 'high efficiency' discriminant (1 b-tag eff. $\simeq 55\%$) $\Rightarrow$ higher $t\bar{t}$ contamination
- Reduce $t\bar{t}$ background with:
  - Smaller di-lepton mass window $76 < m_{\ell\ell} < 106\,\mathrm{GeV}$
  - $\not\!\!E_T < 50\ \mathrm{GeV}$ cut



Background estimation for $Z/\gamma^* + bb$

- **$t\bar{t}$**: data driven estimated with fit to $M(\ell\ell)$
- **Z + ucdsg**: fits to secondary vertex mass distributions $\Rightarrow$ event purity

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



- **ZZ**: from MC sample normalized to the ZZ cross section σ_{ZZ} measured by CMS

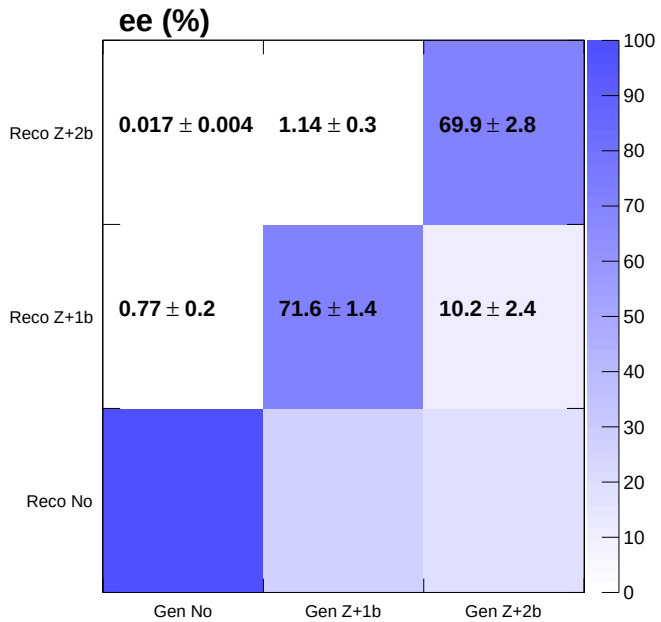
	$\mu\mu + bb$	$ee + bb$
Yields	219	148
bb-purity	$(83 \pm 6)\%$	$(83 \pm 6)\%$
$t\bar{t}$ frac.	$(20 \pm 5)\%$	$(17 \pm 5)\%$
$N_{Z(\ell\ell)Z(bb)}$	5.2 ± 0.2	3.0 ± 0.2

Measurement of the $Z/\gamma^* + bb$ cross-section in multiplicity bins

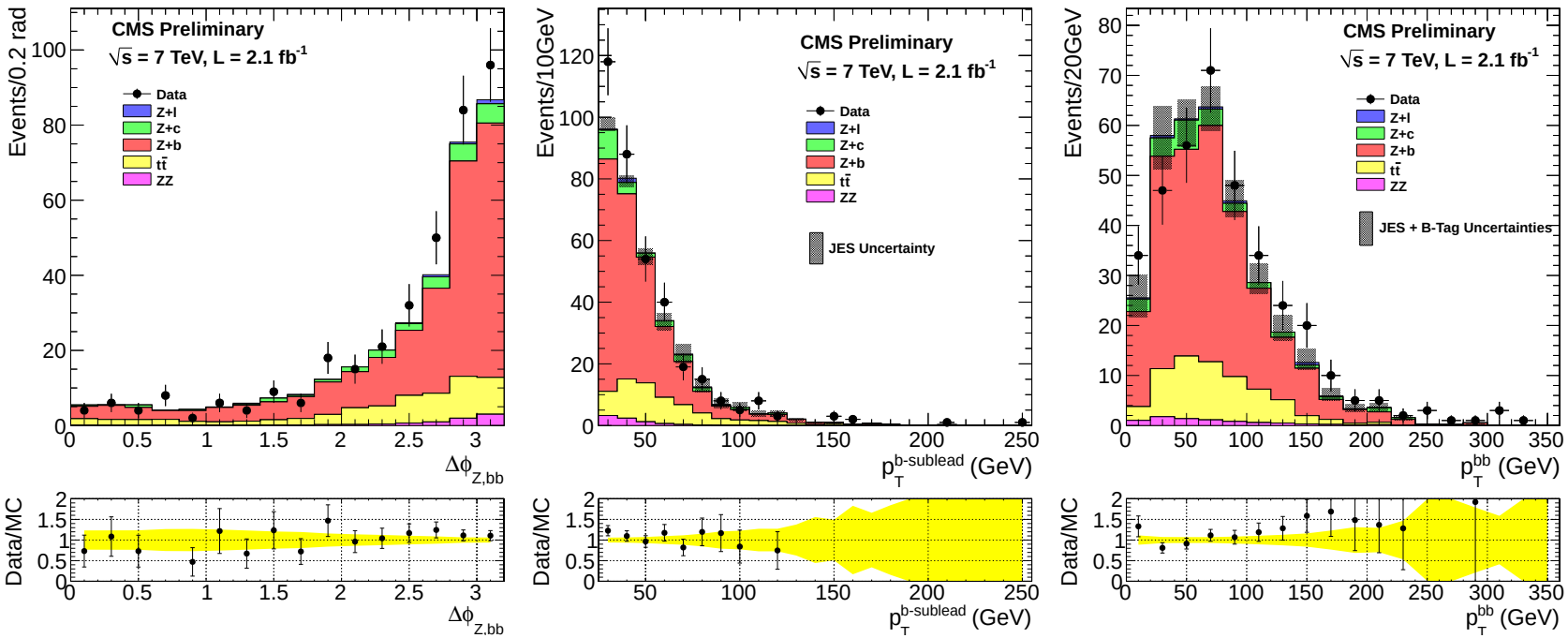
- Measure cross-sections in $b-\text{jet}$ multiplicity bins: $Z+1b$ (exclusive) and $Z+2b$ (inclusive)
- Correction factors take into account migrations between bins:
  - $b-\text{tagging}$ effects
  - jet resolution effects

Multiplicity bin	$\sigma \pm \delta_\sigma^{stat} \pm \delta_\sigma^{syst} \pm \delta_\sigma^{theo}$ (pb)
$\sigma_{hadr}(Z(\ell\ell)+1b)$	$3.41 \pm 0.05 \pm 0.27 \pm 0.06$
$\sigma_{hadr}(Z(\ell\ell)+2b)$	$0.37 \pm 0.02 \pm 0.07 \pm 0.02$
$\sigma_{hadr}(Z(\ell\ell)+b)$	$3.78 \pm 0.05 \pm 0.31 \pm 0.08$

- Common minimal lepton acceptance A_ℓ ($p_T^\ell > 20$ GeV and $|\eta|^\ell < 2.5$)



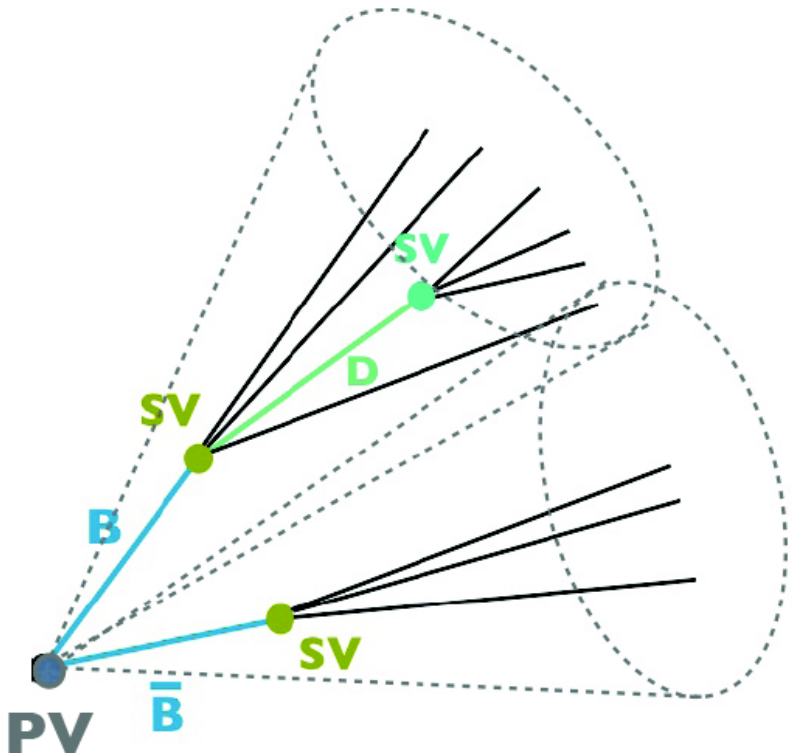
Data MC comparison of $Z/\gamma^* + bb$

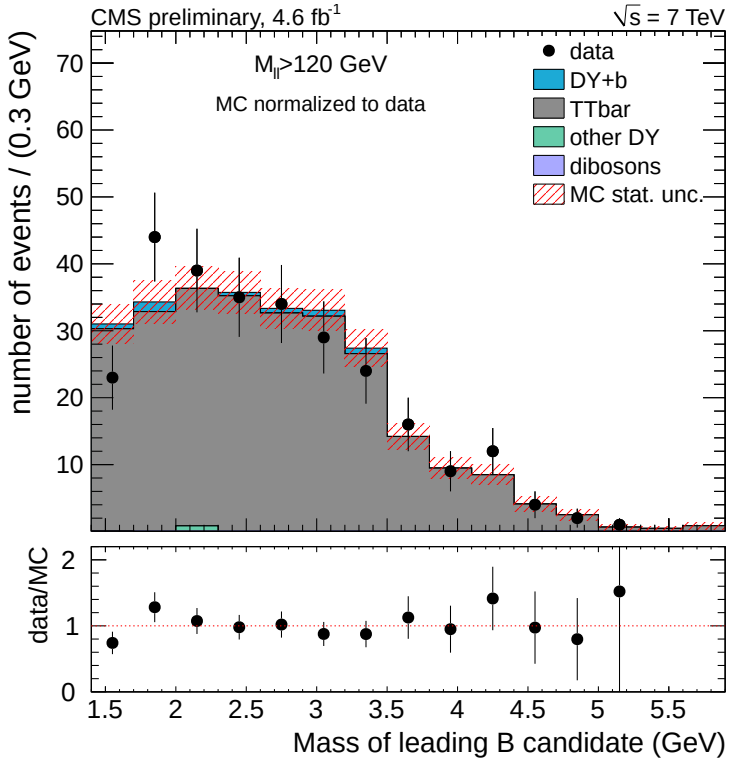


Some discrepancies in the data/MC comparison of kinematic properties $\Rightarrow$ need further studies, possibly use NLO MC simulations

BB angular correlation with Z/γ^* + secondary vertices (EWK-11-015)

- Identification of displaced secondary vertices with no use of jets $\Rightarrow$ study BB pair production also at small angular separation
- *Inclusive Vertex Finder (IVF)*: start from high impact parameter tracks seeds, cluster additional tracks in L_{3D}/σ_{3D} and ΔR
- Very good angular resolution in B hadron flight direction ($\simeq 0.02$ in ΔR)

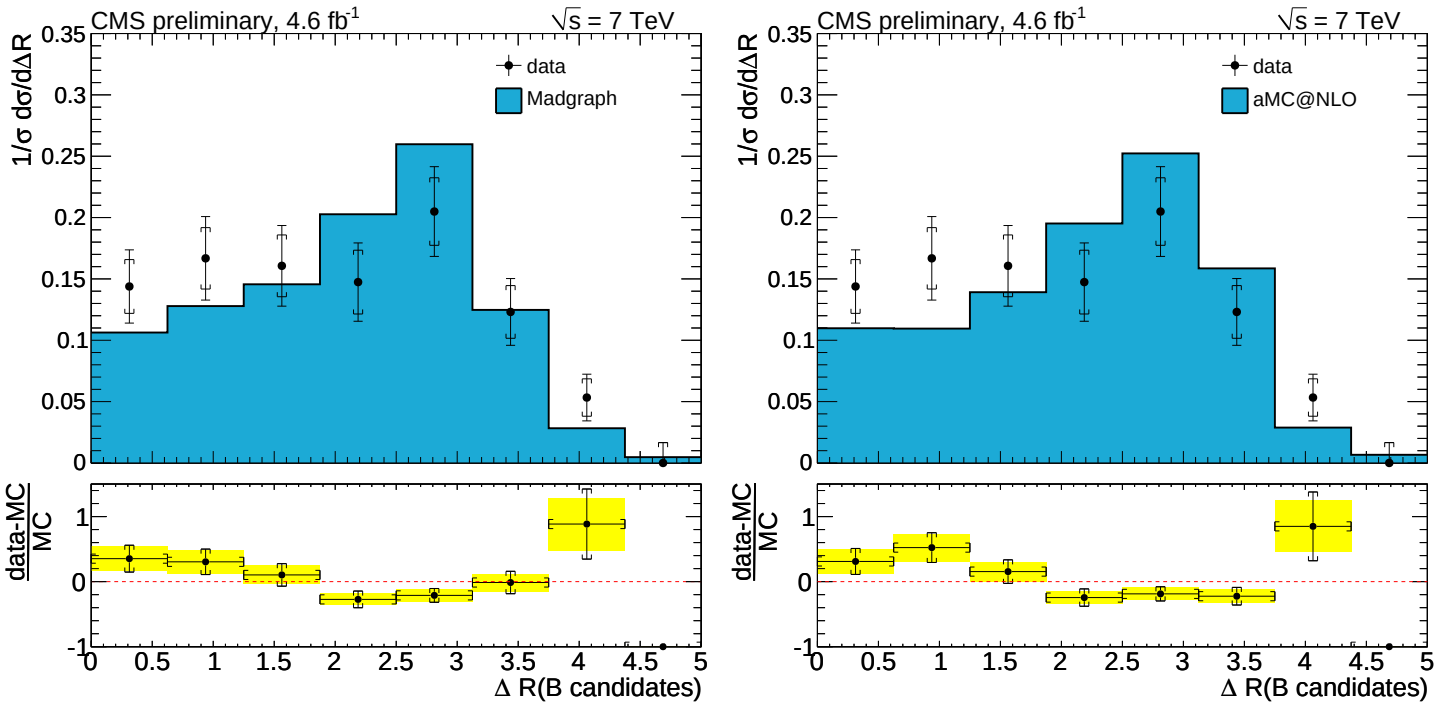




$\Delta R(BB)$ angular distribution in Z/γ^*+bb events

$$\frac{1}{\sigma}\frac{d\sigma}{d\Delta R(BB)} \Rightarrow \frac{1}{\sigma_{visible}}\frac{N_i^{data,fit}\cdot\mathcal{P}_i}{\epsilon_i^{2SV}\cdot\epsilon_i^{\ell}\cdot\mathcal{A}^{\ell}} \qquad i = \Delta R \text{ bin}$$

- $N^{data,fit}$ extracted from a M.L. fit to $M(\ell\ell)$
- $\mathcal{P}_i$ is the IVF purity, ϵ_i^{2SV} the IVF efficiency and $\epsilon_i^{\ell}\cdot\mathcal{A}^{\ell}$ the dilepton efficiency and acceptance
- Comparison with LO and NLO predictions show dicrepancies with both



Conclusions

- The CMS detector is powerful to study vector bosons plus heavy quarks final states
- Preliminary measurements of the $Z + b(b)$ cross sections and angular correlations have been presented, based on 2010 and 2011 collision data at $\sqrt{s} = 7$ TeV
- Reasonable agreement is found with theoretical predictions for absolute cross sections
- Some discrepancies in kinematic properties with Monte Carlo prediction have been observed


CMS EWK-11-012 Measurement of the $Z/\gamma^* + b$-jet cross section in pp collisions at $\sqrt{s} = 7$ TeV


CMS SMP-12-003 Measurement of the $Z\gamma* + bb$-jets cross section in pp collisions at $\sqrt{s} = 7$ TeV


CMS EWK-11-015 Angular correlation between B hadrons produced in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV


CMS EWK-11-013 Measurement of associated charm production in W final states at $\sqrt{s} = 7$ TeV

Backup slides

Calculation of lepton and b tagging efficiencies

Lepton efficiencies with Tag & Probe

- Z mass constraint from a pair of same flavor leptons: **one tag** (high purity), the other **probe** (to measure efficiency given a criterium)
- After di-lepton+jet event topology $\Rightarrow$ **Tag = tight selected lepton**

$$\epsilon_{lepton} = \epsilon_{trk} \times \epsilon_{(reco|trk)} \times \epsilon_{(id|reco)} \times \epsilon_{(iso|id)} \times \epsilon_{(trg|iso)}$$

- Double lepton trigger efficiency: on each leg separately, tag matched to loosest unprescaled single lepton trigger.

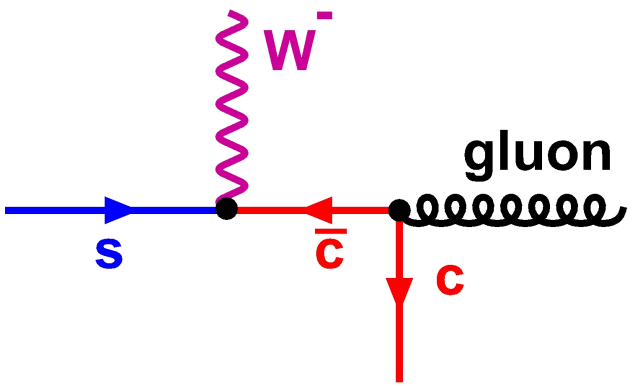
$\epsilon_{trig} = (\epsilon_{L1} \times \epsilon_{H2}) + (\epsilon_{L2} \times \epsilon_{H1}) - (\epsilon_{H1} \times \epsilon_{H2})$ L=Low threshold, H=high threshold

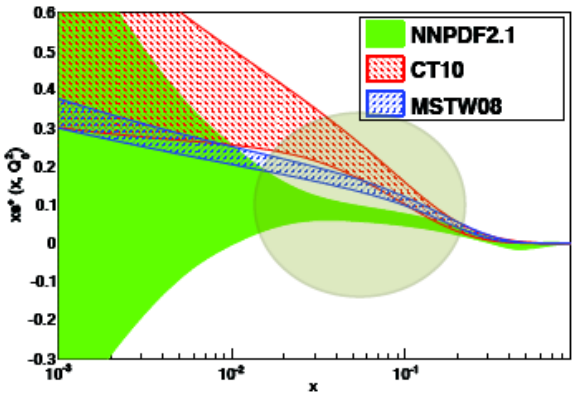
b-tag efficiencies

- Per jet efficiency estimated on signal MC sample reweigthed to reproduce per-jet btagging efficiency measured in data
- use data/MC scale factors SF_b,SF_c,SF_l; SF $= \epsilon_{DATA}/\epsilon_{MC}$
- MC b-tag efficiency and c-mistag: from $Z + b$ and $Z + c$ MC sample
- MC event weight calculated taking into accound all possible mistag combinations

W+charm measurement motivations (EWK-11-013)

- Process $pp \rightarrow W + c + X$ sensitive to the proton strange quark content
- At the LHC it is dominated by $\bar{s}g \rightarrow W^{+}\bar{c}$ and $sg \rightarrow W^{-}c$
- Processes like $\bar{d}g \rightarrow W + \bar{c}$ and $dg \rightarrow W^{-}c$ are Cabibbo disfavoured
- Processes with b quarks in the final state are even more suppressed ($1 \simeq 2\%$) $\Rightarrow$ more than 10% of the $W + jets$ events at the LHC with $p_T^j > 20$ GeV, contain c jets





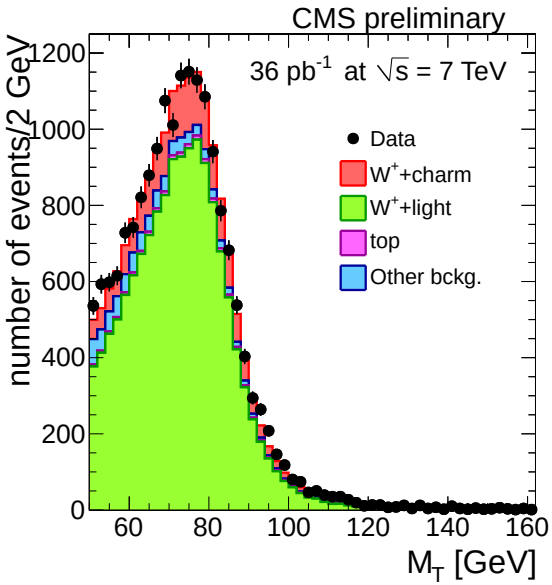
- The cross section ratios:

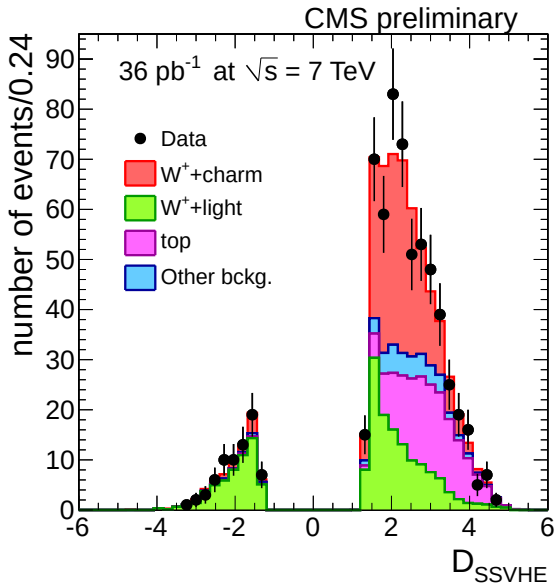
$$R_{+/-} = \frac{\sigma(W^{+} + \bar{c} + X)}{\sigma(W^{-} + c + X)} \quad R_c = \frac{\sigma(W + c + X)}{\sigma(W + jets + X)}$$

provide important information on the strange and anti-strange quark parton density functions of the proton at the electroweak scale

W+charm cross-section ratios measurements

- The ratios are measured using $W \rightarrow \mu\nu$ decays in the kinematic region p_j^T >20 GeV, $|\eta_j|$ <2.1 and with p_T^μ >25 GeV, $|\eta_\mu| < 2.1$, $M_T(\mu\nu) > 50$ GeV.





Discrimator based on the significance of a 2-track secondary vertex decay length

$$D_{SSVHE} = sign(S) \cdot \log(1 + |S|)$$
$$S = L_{3D}/\sigma_{L_{3D}}(SV)$$

- the charm fraction in the selected $W + jets$ sample extracted from a M.L. fit to the different components of the distribution of the D_{SSVHE} discriminator

$R_{+/-}$	0.92 $\pm$ 0.19 (stat.) $\pm$ 0.04 (syst.)
R_c	0.143 $\pm$ 0.015 (stat.) $\pm$ 0.024 (syst.)

- Results are in agreement with theoretical predictions at next-to-leading order based on available parton distribution functions.