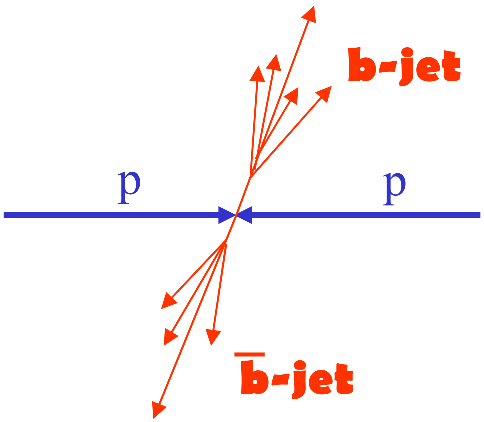
B quark production at CMS

Jérémy Andrea,
IPHC/Université de Strasbourg.
On behalf of the CMS collaboration

Introduction

- Measurement of the b production at LHC using the CMS detector.
- Test of the next-to-leading-order (NLO) QCD prediction at high energy collisions (underestimated at Tevatron).
- Dominant b-quark production mode at LHC: $b\bar{b}$ pair production (QCD). Other source of b quarks: top quark ($t \rightarrow Wb$) or possible new particles.
- B-tagging algorithms (identify b-jets): can be used to select $b\bar{b}$ events. It is also possible to estimate b-tagging performance from $b\bar{b}$ events.
- $b\bar{b}$ correlations: test high order processes,
  - For low order processes : $b\bar{b}$ back to back ($\Delta\varphi(b\bar{b}) \approx \pi$), balanced p_T,
  - For high order processes : spread $\Delta\varphi(b\bar{b})$ distribution, unbalanced p_T.



- Outline:
  - B-production mechanisms,
  - B-tagging algorithms,
  - Measurement of b production using b-tagging (14 TeV),
  - Measurement of $b\bar{b}$ angular correlations without b-tagging (10 TeV).

B production

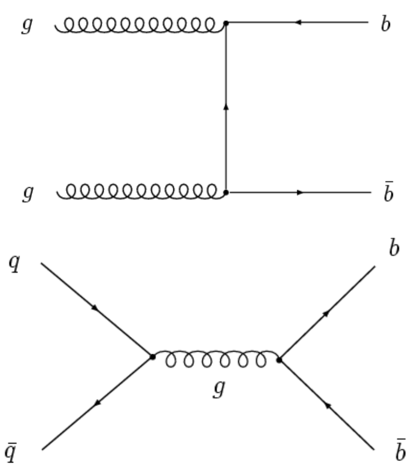
- Three different QCD mechanisms for $b\bar{b}$ production:
  1. Flavour Creation: gluons fusion and $q\bar{q}$ annihilation (low order diagrams)
  2. Flavour Excitation: $b\bar{b}$ pair from the quarks/gluons see with one b being part of the hard process (produce b quarks with asymmetric p_T),
  3. Gluon Splitting: $g \rightarrow b\bar{b}$ in initial or final states, can lead to small value of $\Delta\varphi$.

- Expected $b\bar{b}$ production rate (from PYTHIA, LO, 10 TeV) ≈ 438 μb.

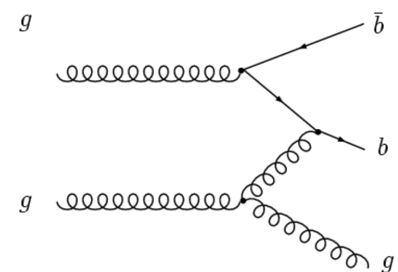
- Measure $b\bar{b}$ production as a function of b quarks p_T and azimuthal angle between the two b quarks.

- At Tevatron, the shapes of transverse momentum, angular distributions and azimuthal correlations are well described by perturbative QCD but the observed cross sections are larger than QCD predictions.

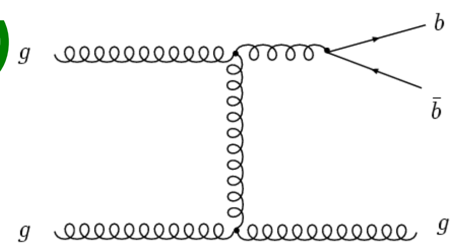
1)



2)

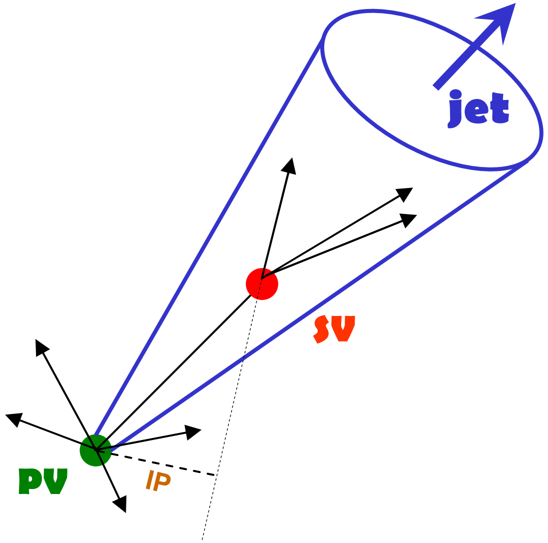


3)

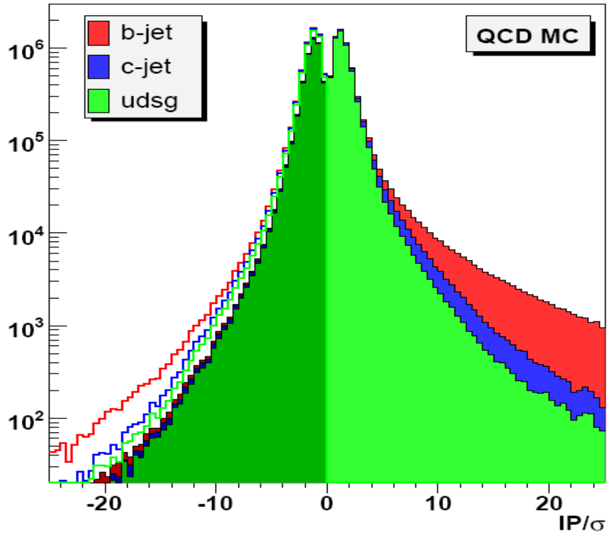


B-tagging at CMS

- Properties of b hadrons:
  - Large life time (≈ps, $c\tau \approx 500\ \mu m$),
  - Secondary vertex (SV),
  - Semi-leptonic decay Br(b→lX)≈20%, Br(b→cX→lX')≈20%,
  - "high" masses, B hadrons take away about 70% of the b quark energy.
- The SV can be shown by the impact parameter (IP) of tracks with respect to the jet axis.
- Sign of the IP: positive if the hadron decays "downstream" or negative it decays "upstream" (resolution on IP/σ) w.r.t. the primary vertex (sign of cos(IP-jet)).



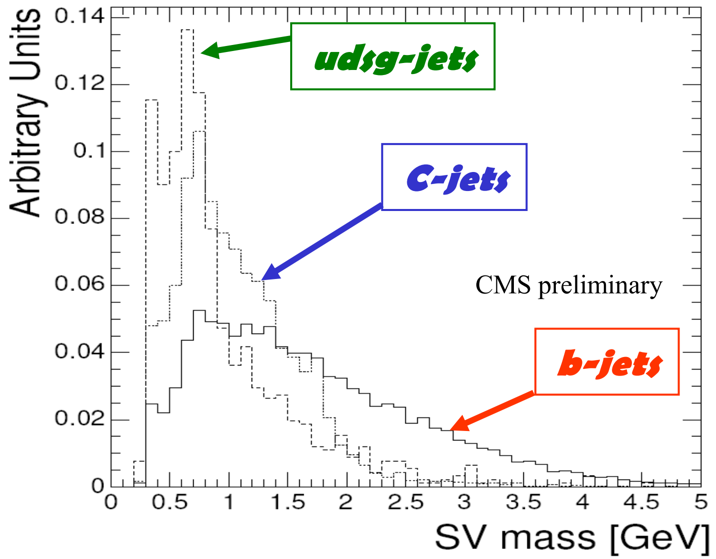
CMS PAS BTV-07-002

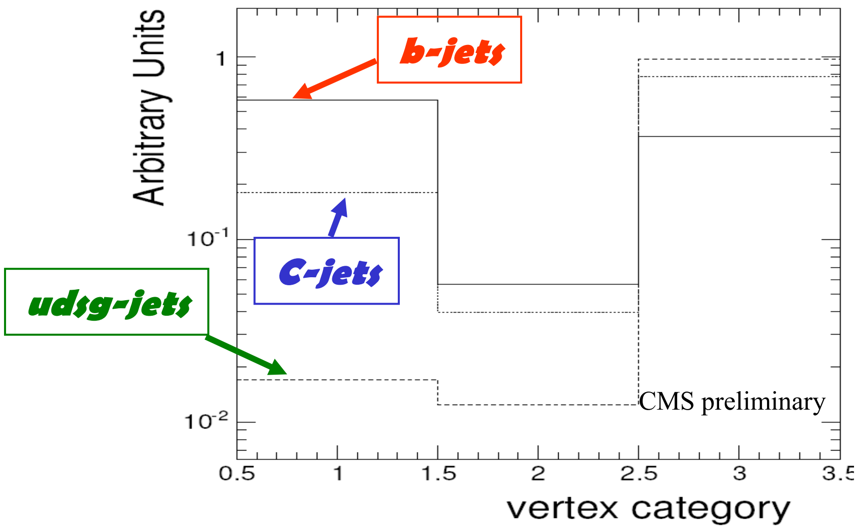


- To take into account resolution effects we use the impact parameter significance IP/σ.
- IP/σ of tracks are:
  - Almost symmetrical for light quark jets and gluon jets (udsg) and centred around 0,
  - Asymmetric for b(c) jets with a larger fraction of tracks with high positive impact parameters.
- Track Counting algorithms: count the number of tracks with an IP/σ > cut value in a jet.

B-tagging at CMS (2)

- Combined Secondary Vertex algorithm,
- Based on the secondary vertex reconstruction of the b-hadrons' weak decay,
- Combined topological and kinematical variables into a discriminator using a likelihood ratio method:
  - Invariant mass of tracks from SV, track multiplicity, distance between PV and SV, rapidity *y* w.r.t. the jet axis, IP/σ of the first tracks



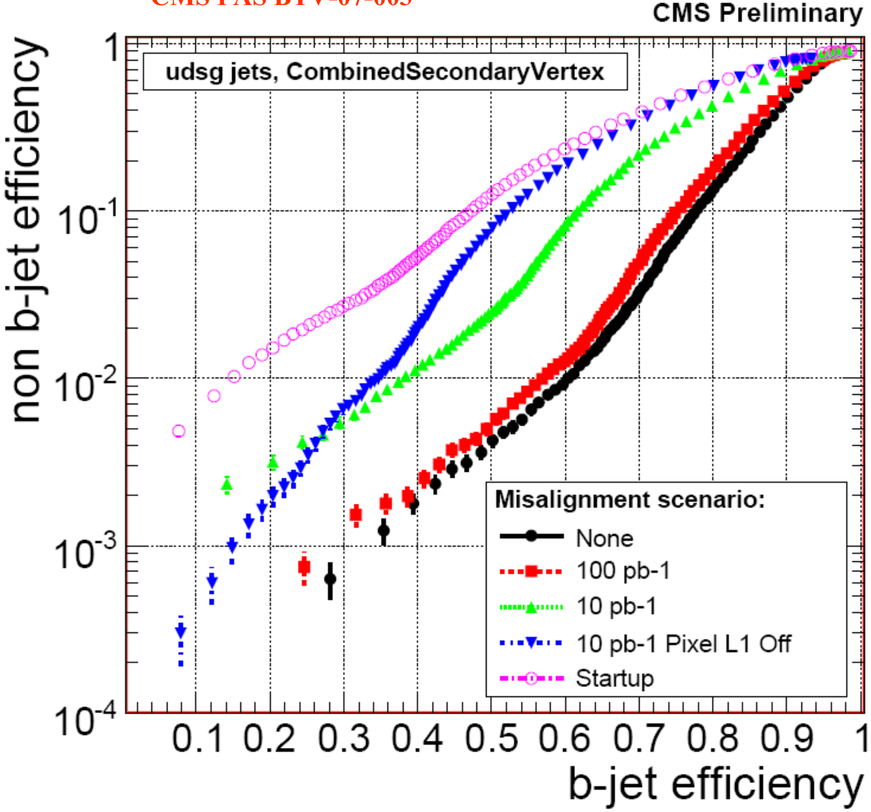


- Three different cases :
  1. "RecoVertex" => A SV is reconstructed,
  2. "PseudoVertex" => no SV reconstructed but there is at least 2 tracks with IP/σ>2,
  3. "NoVertex" => other cases

B-tagging at CMS (3)

- Performance curves for QCD events with $50 < p_T < 80$ GeV/c and for jets with $|\eta| < 1.4$.
- Probability to tag a non b–jet as a b-jet (mistag rate) vs b-tagging efficiency.
- For c jets, light quarks jets and gluon jets.
- For a mistag rate of light quark jets (udsg) of 1%, the b-tagging efficiency is about 60% for a perfect detector.

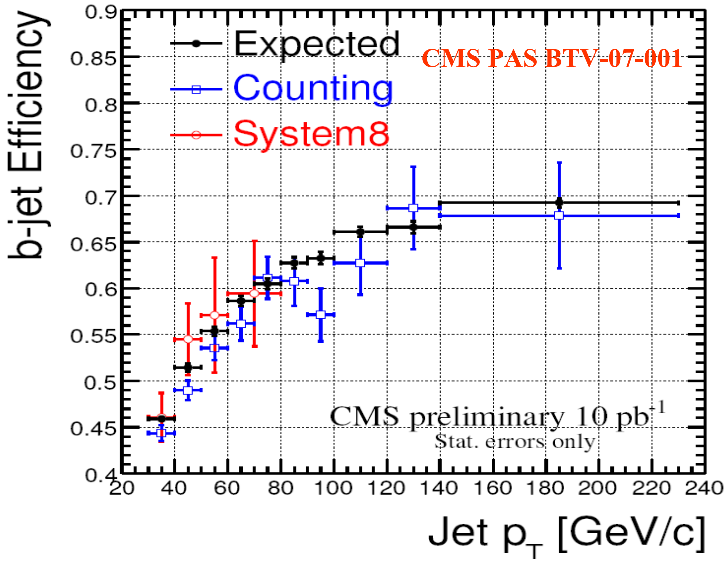
CMS PAS BTV-07-003



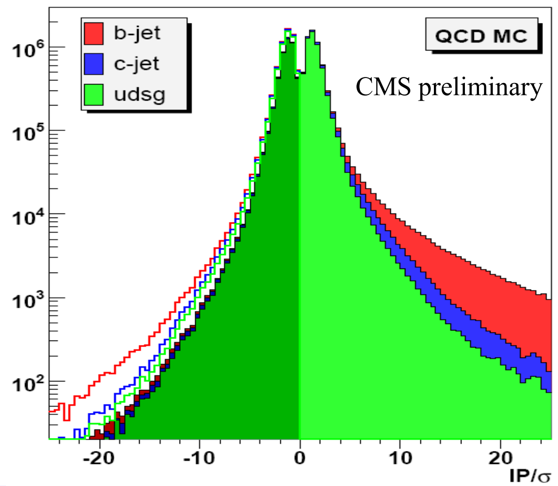
B-tagging performance from data

- B-tagging efficiency (Track Counting).
- From Top quark : selecting a pure sample of $t\bar{t}$ event and counting the number of b-tagged jets.
- From a "muon in jet" sample (QCD): Fit on the p_T of the muon relative to jet axis p_T^{rel},
- The system 8 ("muon in jet" sample QCD):
  - use 3 different identification criteria: the studied algorithm, a cut on the p_T^{rel} of the muon, an additional tagged jet b in the event.
  - Construct an equation system of 8 equations and 8 unknowns => estimate the b-tagging efficiency.

System 8 : For $L=100pb^{-1}$, mistag≈1%, uncertainty ≈8.6%



CMS PAS BTV-07-002



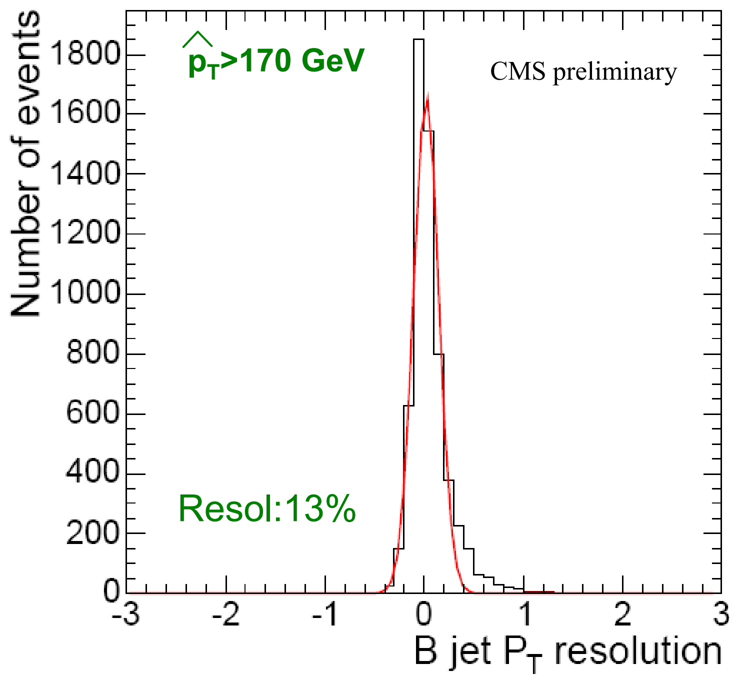
Mistag rate

- IP/σ distribution is almost symmetric for light quarks and gluon jets.
- Negative tagger which uses only tracks with IP/σ<0 ⇒ it gives an estimate of the mistag rate.
- Correct by a Monte-Carlo factor (sources of asymmetries in IP/σ for light jets like V^0, $\gamma \rightarrow ee$).

For $L=100pb^{-1}$, mistag≈1%, uncertainty ≈8.0%

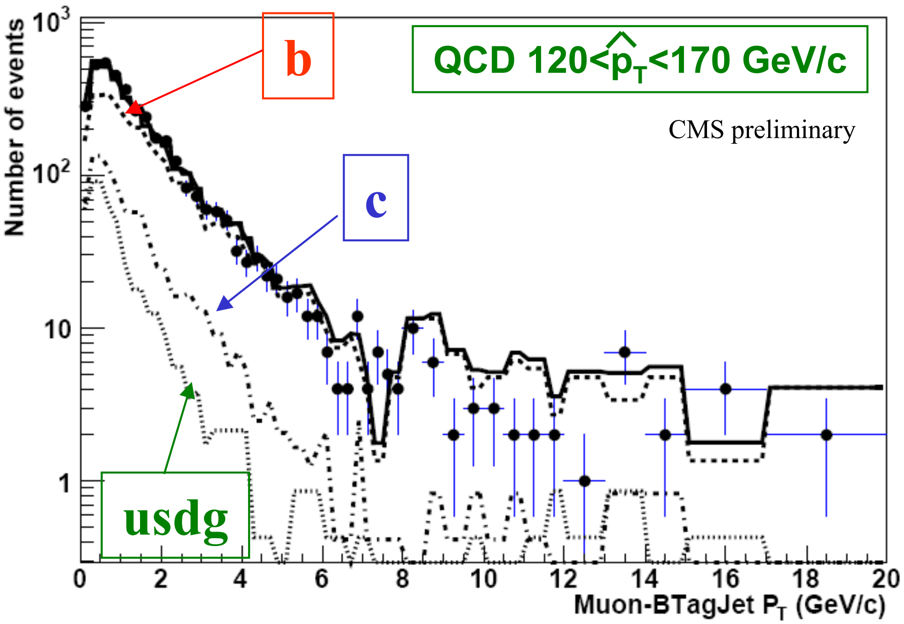
B production at CMS

- Inclusive b production measurement for 14 TeV collisions and high luminosity.
- Analysis strategy: select events with at least one b-tagged jet and at least one muon. p_T spectrum is determined by looking at the most energetic b-jet per event.
- Trigger selection:
  - Level 1 trigger "single muon" with p_T > 14 GeV/c,
  - High Level Trigger: "muon+b-jet" => 1 non-isolated muon (p_T > 19 GeV/c) + a b-tagging requirement on jet (E_T> 50 GeV and $|\eta|<2.4$, Track Counting from pixel-only tracks).
- Event selection:
  - At least 1 b-tagged jet in the event (using the Combined SV algorithms presented previously).
  - p_T > 50 GeV/c and $|\eta| < 2.4$
  - Only the most energetic jet is considered, measurement of $d\sigma/dp_T$, $d\sigma/d\eta$…
- B-tagging efficiency is about 65% in the barrel and 10% less in the end-caps.



B production at CMS (2)

- Additional requirement: the b jets candidates associated with the reconstructed muon (b-tagged jet is required to be the closest jet to the muon).
- Total selection efficiency ≈ 5% (≈ 25% when taking into account semi-leptonic branching ratio of b and c).
- p_T of the muon relative to the closest jet axis:
  - Shapes are different for b, c and uds jets (discriminating power),
  - Can be used to fit the data.



- Example of fit using the expected shapes of the muon p_T relative to jet axis for b, c and light quarks jets.

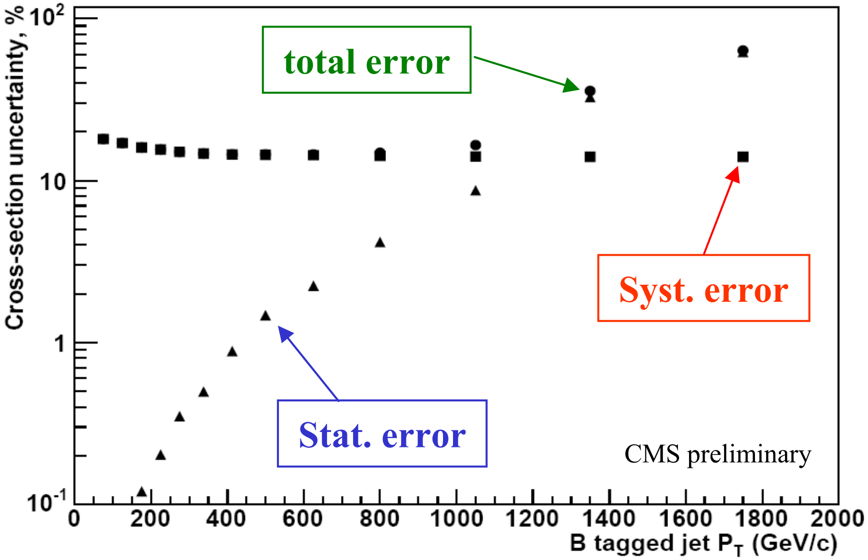
230<$\hat{p}_T$< 300 GeV/c	Monte-Carlo truth	Fit results
$b\bar{b}$	5250	5222±501
$c\bar{c}$	2338	2050±728
uds	1740	1778±341

- b purity changes from 70% to 55%.
- B-hadron p_T range accessible should goes up to 1.5 TeV/c.
- $t\bar{t}$ contamination < 1%.

B production at LHC (4)

- Systematic uncertainties for an integrated luminosity of 10 fb^{-1} at 14 TeV:
  - Dominated by JES uncertainty,
  - Fragmentation of b quarks,
  - Monte-Carlo modelling,
  - B-tagging, Luminosity, Trigger efficiency etc…

Source	uncertainty, %
jet energy scale	12
event selection	6
B tagging	5
luminosity	5
trigger	3
muon Br	2.6
misalignment	2
muon efficiency	1
$t\bar{t}$ background	0.7
fragmentation	9
total	18



- Estimated statistical (triangles), Systematic (squares) and total (dots) uncertainty of the cross section measurement as a function of the b-tagged jet p_T.

- For a bjet of 100 GeV/c, the total expected uncertainty is smaller than 20%.

$b\bar{b}$ correlations

CMS PAS BPH-08-004

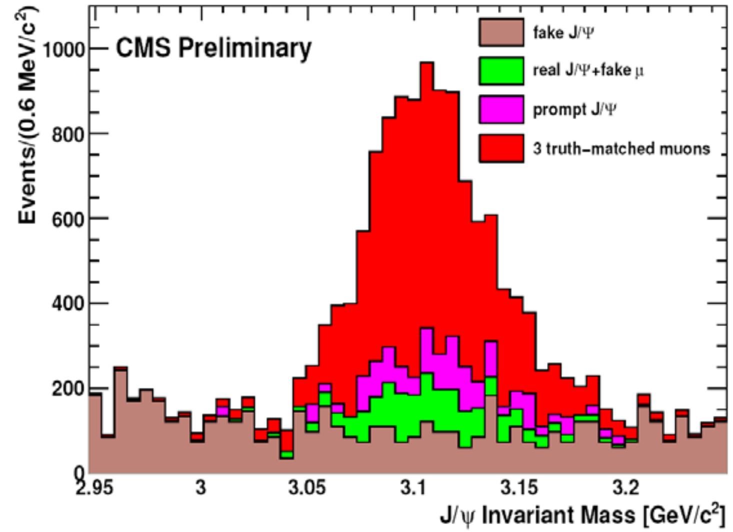
- Measure of the azimuthal correlations of $b\bar{b}$ production ($d\sigma/d\Delta\varphi$) for a luminosity of 50 pb^{-1}, 10 TeV collisions, without b-tagging (in the case where b-tagging does not perform well).
- Channel of interest: $b\bar{b}\rightarrow(J/\psi X)(\mu X')$ where the J/ψ particle decays into a pair of muons.

- Analysis strategy:
  - one b is tagged by reconstructing a J/ψ particle from 2 opposite sign muons and ask for an additional muon compatible with a semi-leptonic decay of the second b,
  - Measure the yield in each $\Delta\varphi$ bins by means of an unbinned likelihood,
  - Correct for resolution effects on the $\Delta\varphi$ distribution using an unfolding procedure.

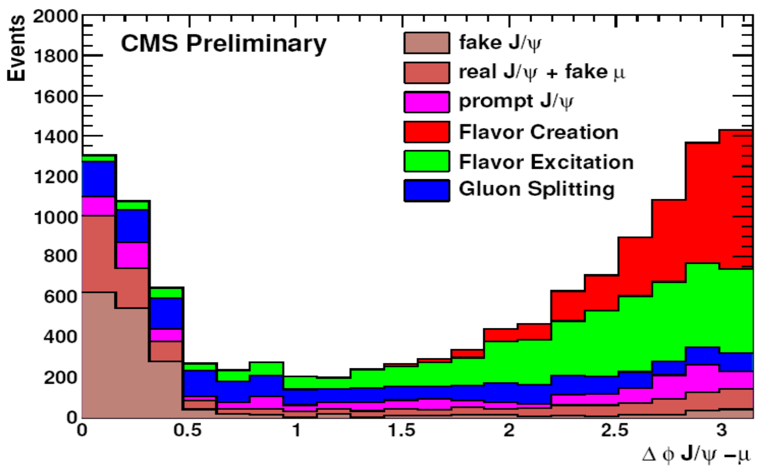
- Monte-Carlo samples :
  - Inclusive $pp\rightarrow bb\rightarrow J/\psi X$ produced by PYTHIA + EvtGen (for b hadrons generation),
  - Prompt J/ψ from primary vertex,
  - Minimum bias with at least one generated muon ($pp\rightarrow\mu X$).

$b\bar{b}$ correlations (2)

- Events selection:
  - Trigger requirements : 2 HLT muons with $p_T > 3$ GeV/c ($\varepsilon_{trig} \approx 22\%$),
  - From 2 reconstructed “tracker muons” (tracker tracks+some matching requirements with muon segments) the J/ψ vertex has to be reconstructed,
  - A third muon with $p_T > 3$ GeV/c and $|\eta| < 2.4$ is required.
  - Additional quality cuts on the 3rd muon (>10 silicon tracker hits + track fit χ^2/ndf).
- After the selection, the effective cross section is 145 pb.



Δφ (J/ψ-μ) after the selection

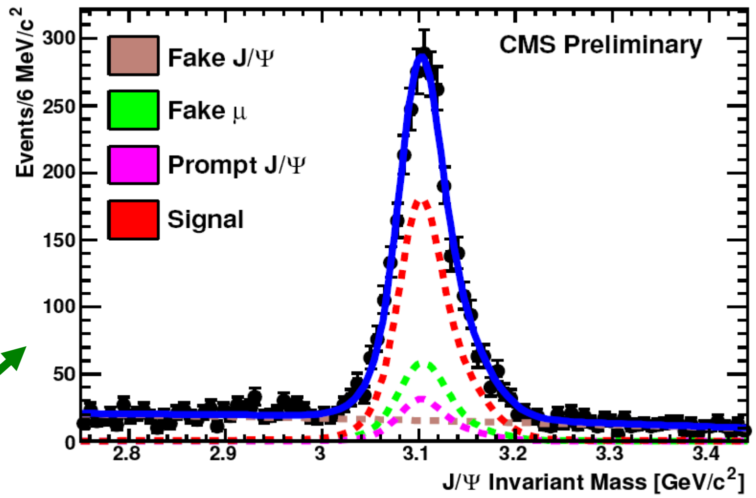


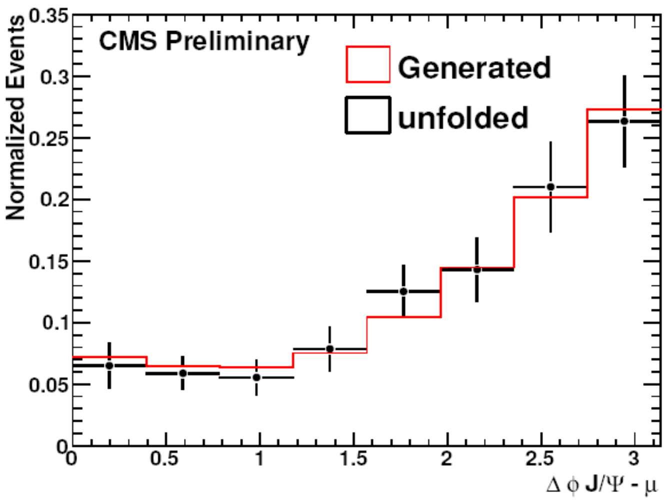
$\mu^+\mu^-$ invariant mass distribution for **signal** events, events with a **fake J/ψ**, a real J/ψ but a **fake 3rd muon** or a real J/ψ from **prompt interaction.**

Irreducible background from $B_c \rightarrow J/\psi\mu X$.

$b\bar{b}$ correlations (3)

- Signal extraction:
  - For the different Δφ bins => unbinned maximum-likelihood fit from $M_{\mu\mu}$, the transverse flight distance and the impact parameter of the 3rd muon.
  - PDFs for signal and backgrounds are extracted from Monte-Carlo sample.
  - Extract amplitudes for signal and backgrounds from the fit.
- Fit result for $13pb^{-1}$.





- Δφ resolution is of the same order than bin size,
- an unfolding procedure is applied to disentangle resolution effects in the Δφ distribution.
- Comparison between the generated and unfolded Δφ distributions.

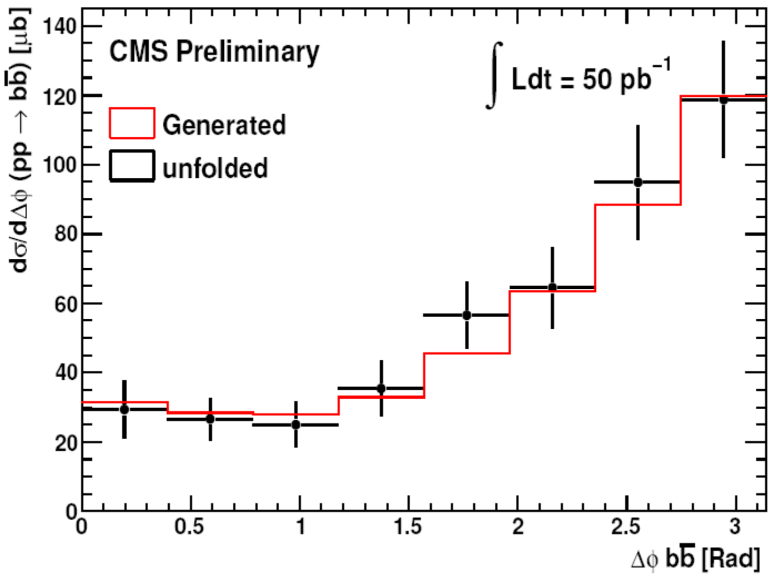
$b\bar{b}$ correlations (4)

- Sources of systematic uncertainties, estimated for each Δφ bins:
  - the fraction of muons produced by the cascades decay b→cX→ μX',
  - Bottom hadron lifetimes are correlated to the IP of their decay products: uncertainty on the fraction of bottom quark fragmenting in the different bottom hadrons,
  - Uncertainty on the PDF shapes,
  - J/ψ polarization and alignment effects,
  - Stat. of the Monte-Carlo and background from $B_c \rightarrow J/\psi \mu X$.

- Final results:
  - The final cross section is calculated according to the formula:

$$\frac{d\sigma}{d\Delta\varphi} = \frac{N_{fit}}{L \cdot \varepsilon_{trg} \cdot \varepsilon_{reco}}$$

  - Depending on the particular Δφ bin, the differential cross-section uncertainty is between 15 and 25% for an integrated luminosity of 50 pb^{-1}.
  - An uncertainty of 10% is expected for the integrated total bbX cross section.



Conclusion

- Studying b production at LHC will allow to test QCD predictions.

- Large numbers of analysis will depend on the $b\bar{b}$ production rate :
  - B-tagging efficiency estimated from data, Top physics, Higgs physics, search for new Physics,

- Already with early data (few pb^{-1} at 10 TeV collisions), CMS should be able to measure differential $b\bar{b}$ cross section ($\Delta\varphi$) with a good precision (without b-tagging).

- With higher luminosities (few fb^{-1} at 14 TeV collisions) and with the use of b-tagging, CMS should be able to measure differential $b\bar{b}$ cross section (p_T) up to 1.5 TeV.

- CMS commissioning is well ongoing with cosmic data. We are waiting for collisions...