Search for $B_{s,d} \rightarrow \mu\mu$ decays

Beauty 2014
Edinburgh, 14-18 July, 2014

Arno Heister
on behalf of the CMS collaboration

For further reference:

- "Measurement of the B(s) to mu+ mu- branching fraction and search for B0 to mu+ mu- with the CMS Experiment" (arXiv:1307.5025)
- PhysicsResultsBPH13004

Rare B_s and B_d decays

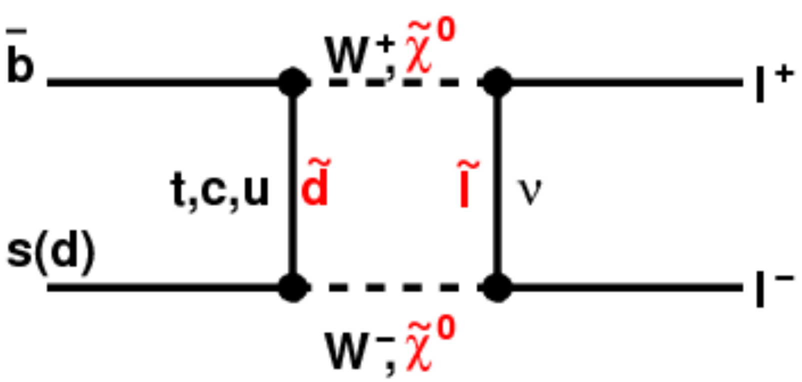
Arno.Heister@cern.ch

Standard Model predictions:

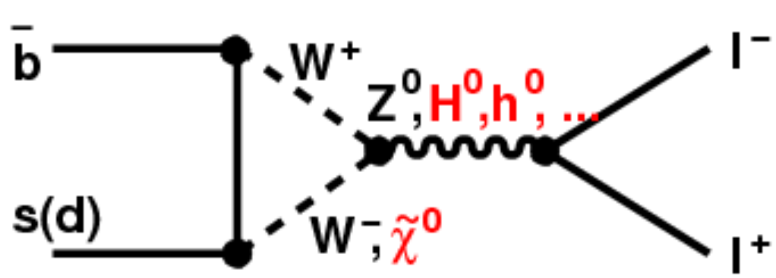
$$BR(B_s \to \mu^+\mu^-) = (3.56 \pm 0.18) \times 10^{-9}$$ arXiv:1303.3820

$$BR(B_d \to \mu^+\mu^-)_{SM} = (1.07 \pm 0.10) \times 10^{-10}$$ arXiv:1208.0934

“Box diagram”



“Z penguin”

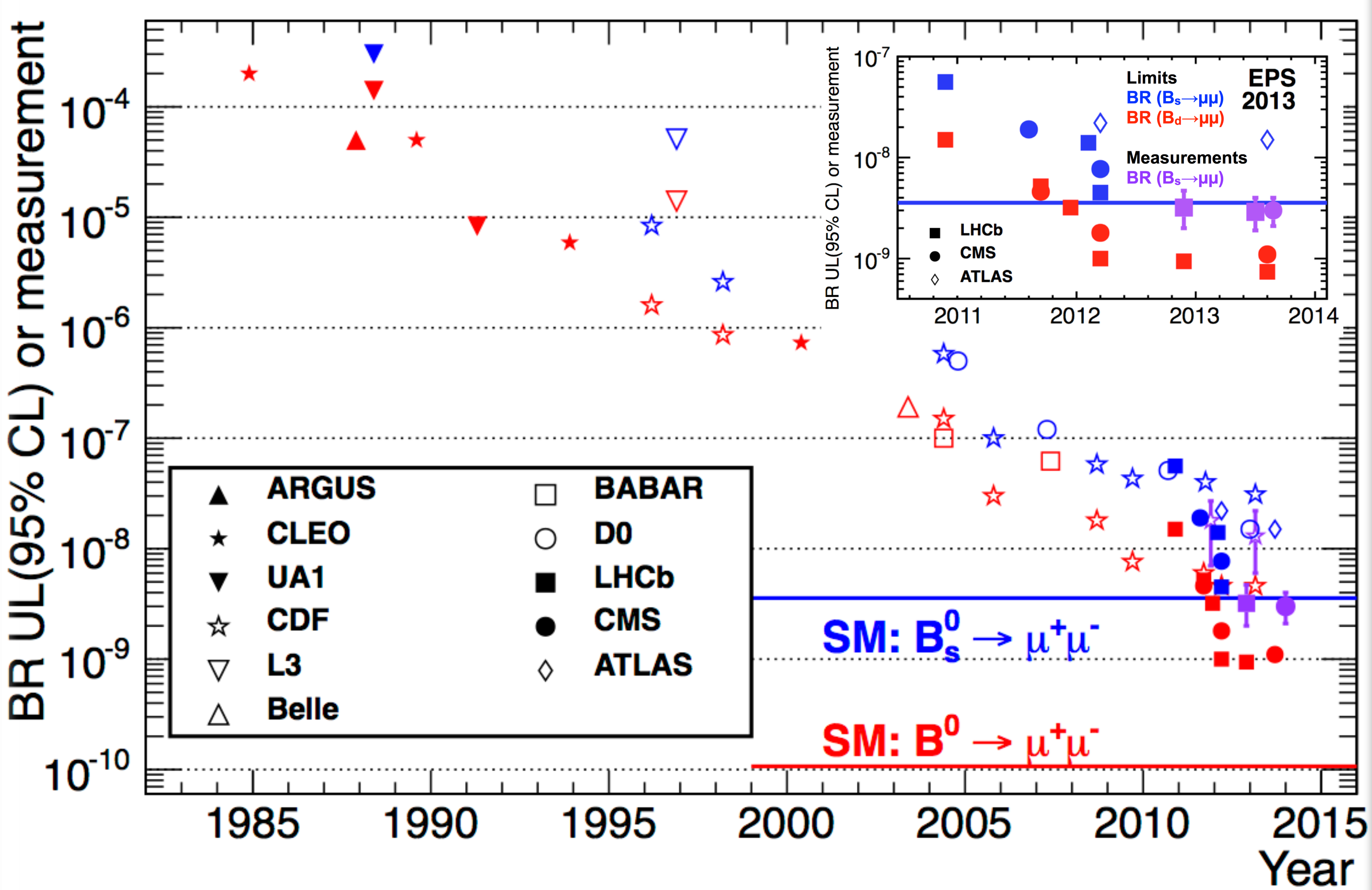


black: Standard Model particles
red: new particles (here MSSM)

- In Standard Model decays of B_s and B_d into muons are highly suppressed
  - $b \to s(d)$ is a flavor changing neutral current (FCNC) transition
  - decay through higher-order loop diagrams
  - helicity suppressed
  - Cabibbo enhancement of $B_s \to \mu\mu$ over $B_d \to \mu\mu$ due to $|V_{td}| < |V_{ts}|$
- Enhancements of the branching fractions possible due to new physics.

B_s and B_d: "State of the art"

Arno.Heister@cern.ch



- More than a 25 years quest.
- Until 2012 only limits could be achieved.
- 2012 LHCb first evidence (3.5 σ) for B_s arXiv:1211.2674

Analysis overview

Arno.Heister@cern.ch

- Analysis of CMS data from LHC run 1: $5fb^{-1}$ @ $\sqrt{s}$ = 7TeV, $20fb^{-1}$ @ $\sqrt{s}$ = 8TeV
- Simultaneous search for the decay $B_s \rightarrow \mu^+\mu^-$ and $B_d \rightarrow \mu^+\mu^-$ in a mass window:
  - Candidate events (two muons from a common decay vertex, etc.) selected by multivariate analysis techniques (Boosted Decision Tree)
  - Fit to invariant mass of the finally selected dimuon sample
  - Efficiencies, acceptances taken from Monte Carlo (MC) simulations and validated with data
    - B_s mesons in data and MC validated and compared utilizing: $B_s \rightarrow J/\psi\phi \rightarrow \mu^+\mu^-K^+K^-$
  - Backgrounds:
    - *combinatorial* estimated from sidebands,
    - rare single B decays by Monte Carlo (MC) simulation,
      - "non-peaking": $B \rightarrow h\mu\nu$, $B \rightarrow \mu\mu\nu$, $\Lambda_b \rightarrow p\mu\nu$, (h,h': charged hadrons misidentified as muons)
      - "peaking": $B \rightarrow hh'$.
- Relative normalization to: $B^\pm \rightarrow J/\psi K^\pm$ to avoid dependance on b-production cross section and luminosity measurement:

$$BR(B_s \to \mu^+\mu^-) = \frac{N_s}{N_{obs}^{B^+}} \cdot \frac{f_u}{f_s} \cdot \frac{\epsilon_{tot}^{B^+}}{\epsilon_{tot}} \cdot BR(B^+)$$

  - f_s/f_u: ratio of probabilities for a b-quark hadronizing into a B^+ or B_s meson
  - measured by LHCb: $f_s/f_u = 0.256 \pm 0.020$ (arXiv:1301.5286)

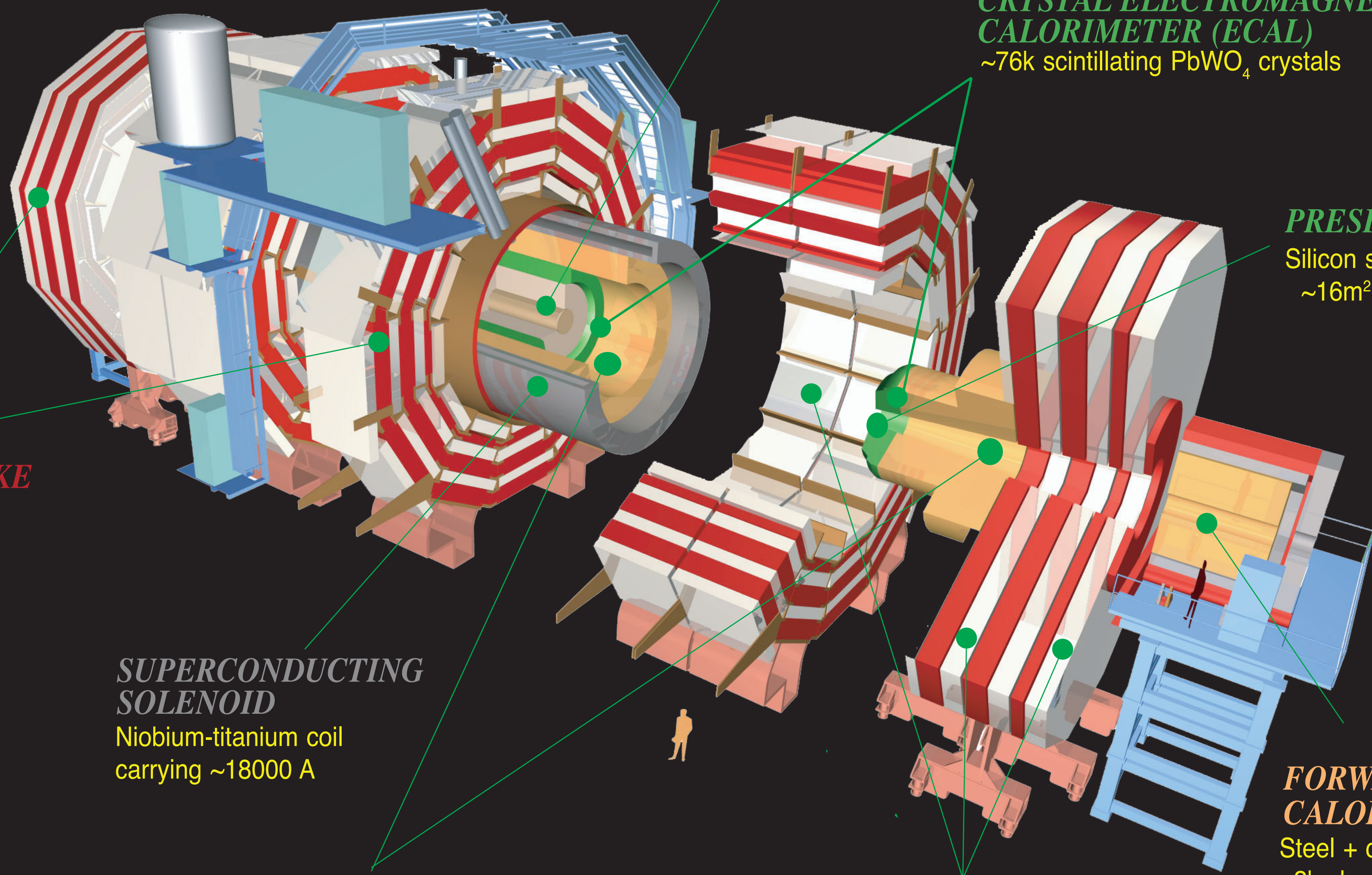
CMS Detector

Pixels
Tracker
ECAL
HCAL
Solenoid
Steel Yoke
Muons

SILICON TRACKER
Pixels (100 x 150 μm²)
~1m² ~66M channels
Microstrips (80-180μm)
~200m² ~9.6M channels

CRYSTAL ELECTROMAGNETIC CALORIMETER (ECAL)
~76k scintillating $PbWO_4$ crystals

PRESHOWER
Silicon strips
~16m² ~137k channels



STEEL RETURN YOKE
~13000 tonnes

SUPERCONDUCTING SOLENOID
Niobium-titanium coil carrying ~18000 A

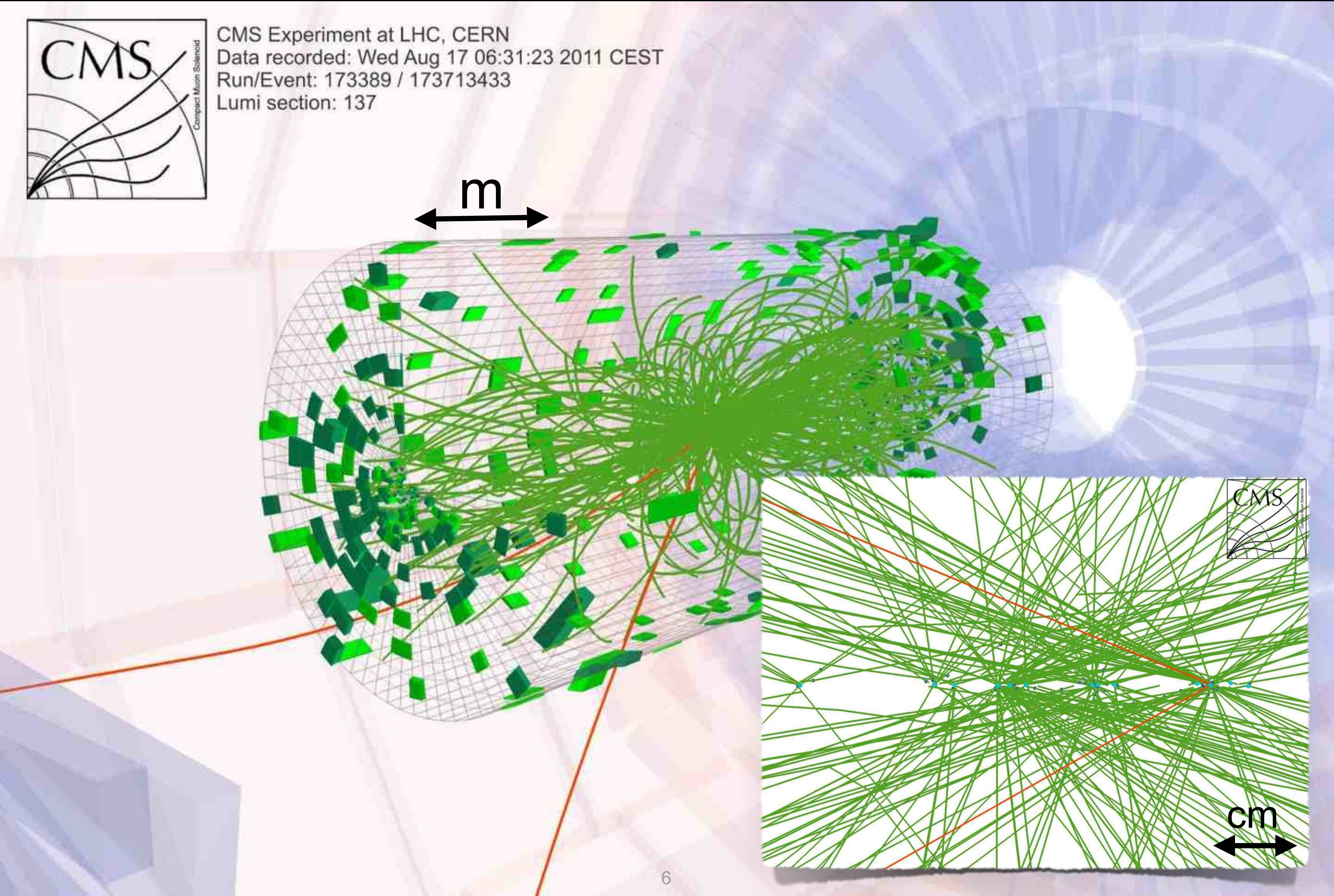
FORWARD CALORIMETER
Steel + quartz fibres
~2k channels

HADRON CALORIMETER (HCAL)
Brass + plastic scintillator
~7k channels

MUON CHAMBERS
Barrel: 250 Drift Tube & 480 Resistive Plate Chambers
Endcaps: 473 Cathode Strip & 432 Resistive Plate Chambers

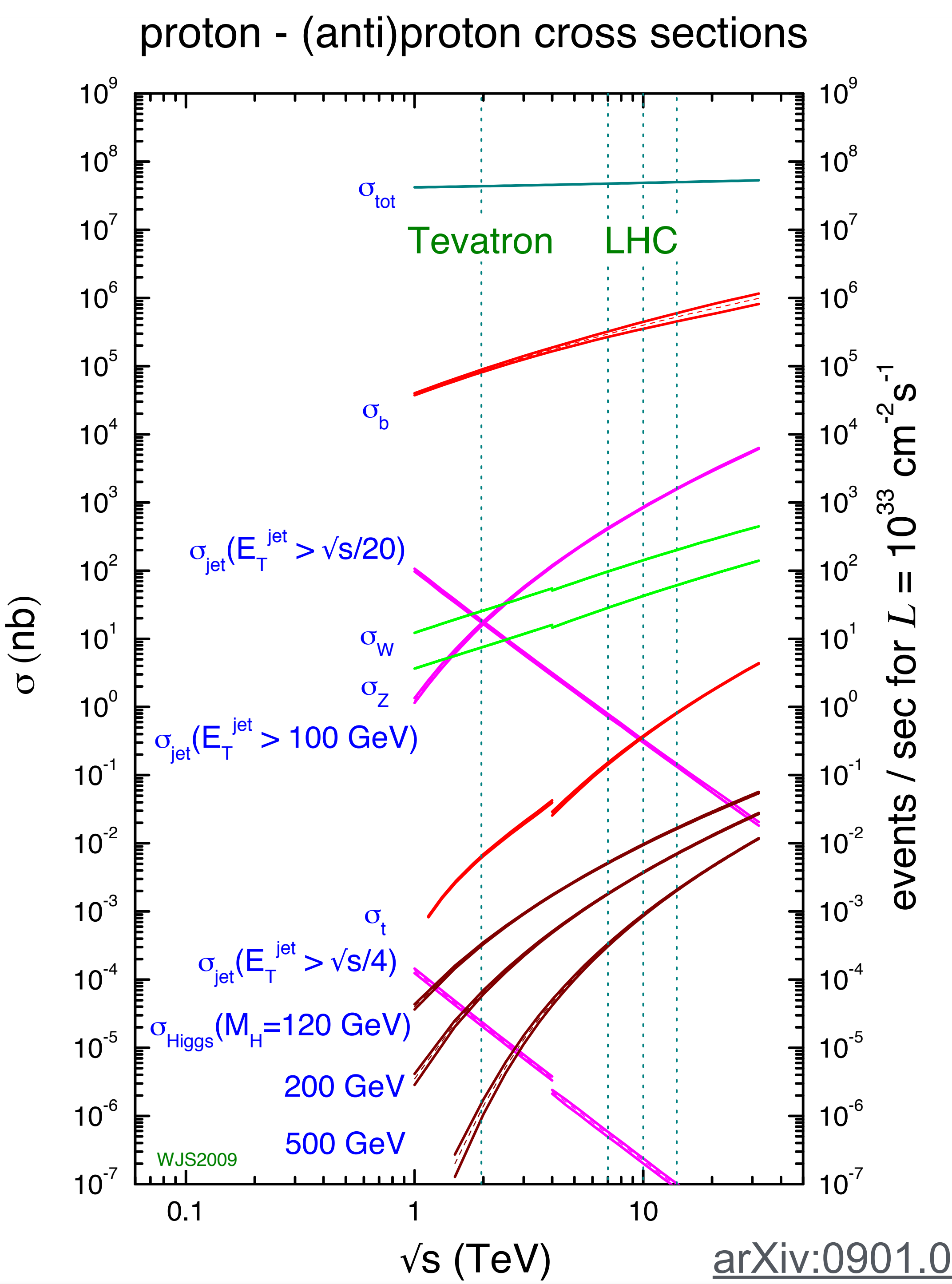
Total weight	**: 14000 tonnes**
Overall diameter	**: 15.0 m**
Overall length	**: 28.7 m**
Magnetic field	**: 3.8 T**

$B_s \rightarrow \mu\mu$ candidate event



CMS dimuon trigger overview

Arno.Heister@cern.ch



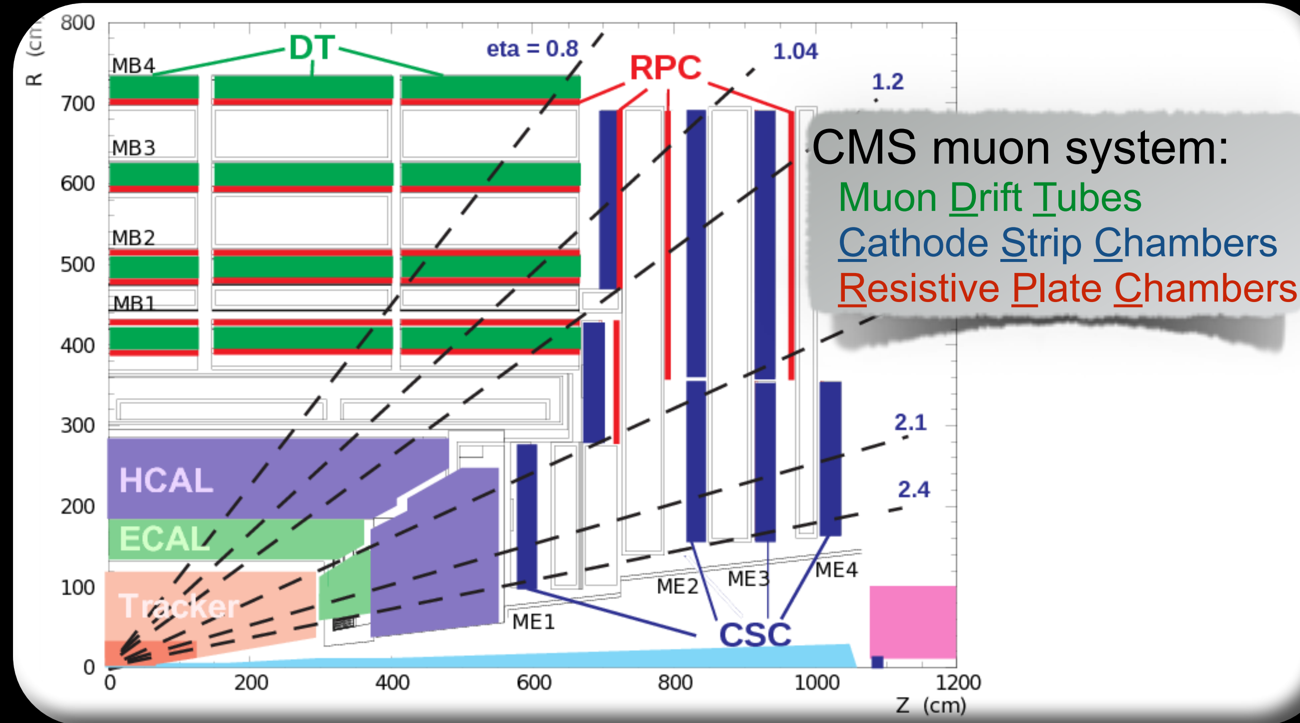
arXiv:0901.0002

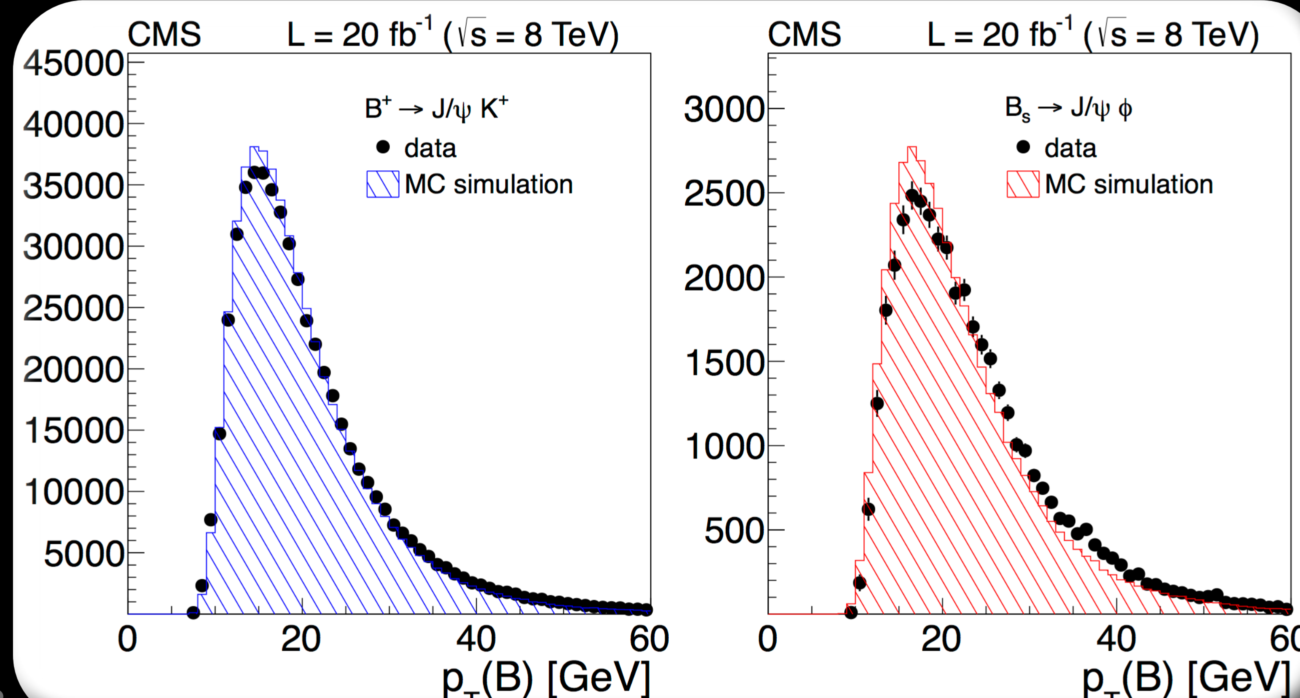
- $\sigma_b \approx 300$ µb ⇒ ≈10^6 Hz b-events
- maximum CMS archiving rate O(100) Hz
- CMS trigger split in 2 steps:
  - L1 Trigger (max. rate 100 kHz):
    - custom electronics
    - pre-selection of events based on information from the muon systems and calorimeters.
  - High Level Trigger (HLT) (max. rate O(100) Hz):
    - a multi-processor computer farm running standard CMS reconstruction software
    - reconstruction algorithms are performance optimized
    - uses the information of all detectors incl. silicon tracker to select the final data used for analysis.
- Di-muon event candidate selection:
  - L1 trigger: two muon candidates in the muon chambers ⇒ 10 kHz at the end of 2012
  - HLT: due to varying LHC conditions the selection (p_T, η, mass window) had to be optimized (rate and efficiency) ⇒ 10-30 Hz at the end of 2012

Dimuon reconstruction

Arno.Heister@cern.ch

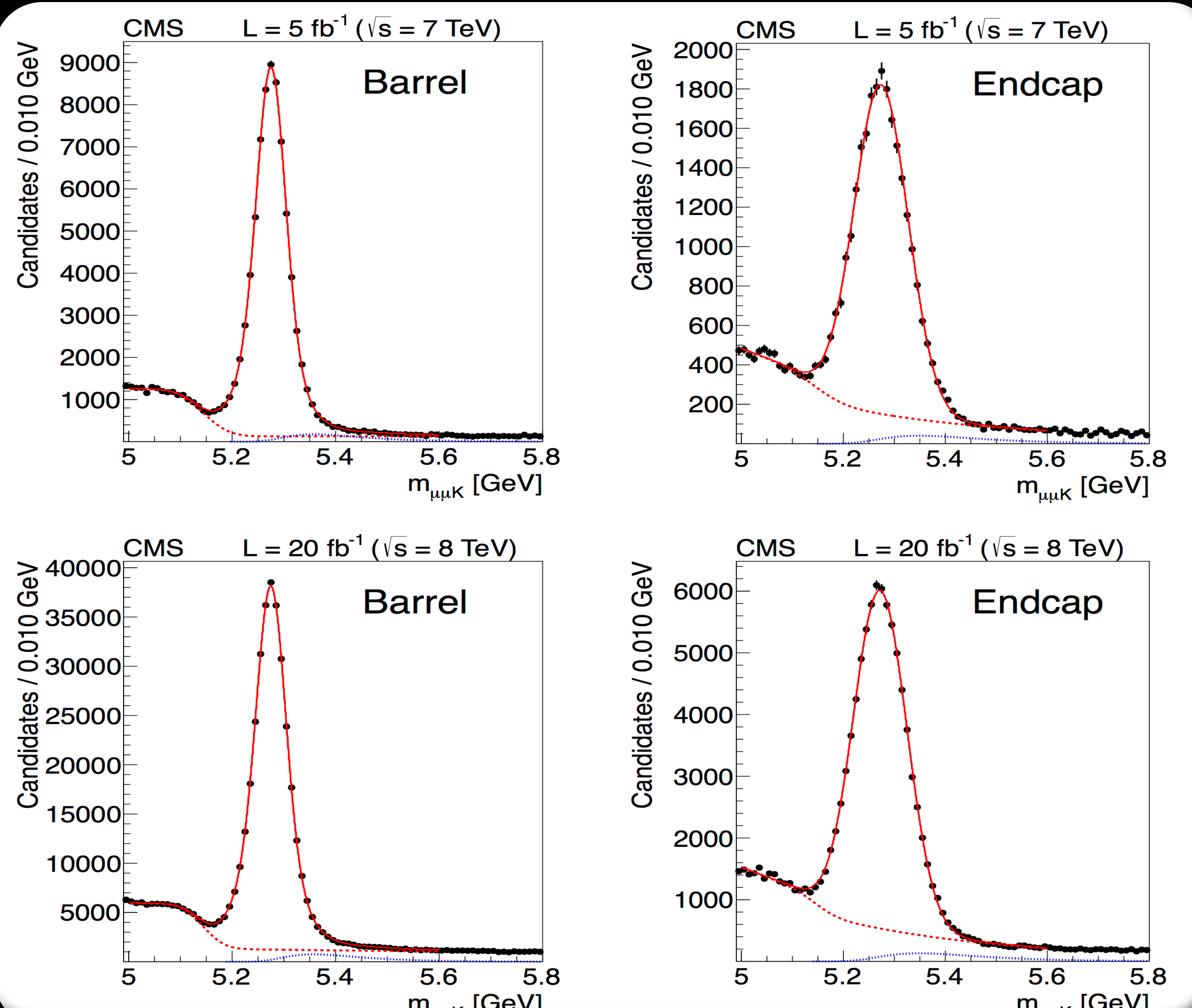
- *Offline* reconstruction of muons by means of silicon tracker and muon chamber hits in combination.
- Well reconstructed muon candidates candidates (p_T>4 GeV) accepted by means of a BDT.
  - BDT trained on MC simulation samples of B-meson decays to kaons and muons.
  - BDT further reduces the hadron-to-muon misidentification.
- Another set of BDTs is used for the final signal (background) selection.
  - background level and mass resolution depend significantly on the B-candidate $\eta_{\mu\mu}$, e.g. $\delta m_{\mu\mu}$ 32 MeV @ $\eta_{\mu\mu} \approx 0$ and 75 MeV @ $|\eta_{\mu\mu}|$> 1.8
  - events categorized in "barrel": $|\eta_{\mu\mu}|$< 1.4 and "endcap": $|\eta_{\mu\mu}|$> 1.4.





Normalization channel $B^{\pm} \rightarrow J/\psi K^{\pm}$

Arno.Heister@cern.ch

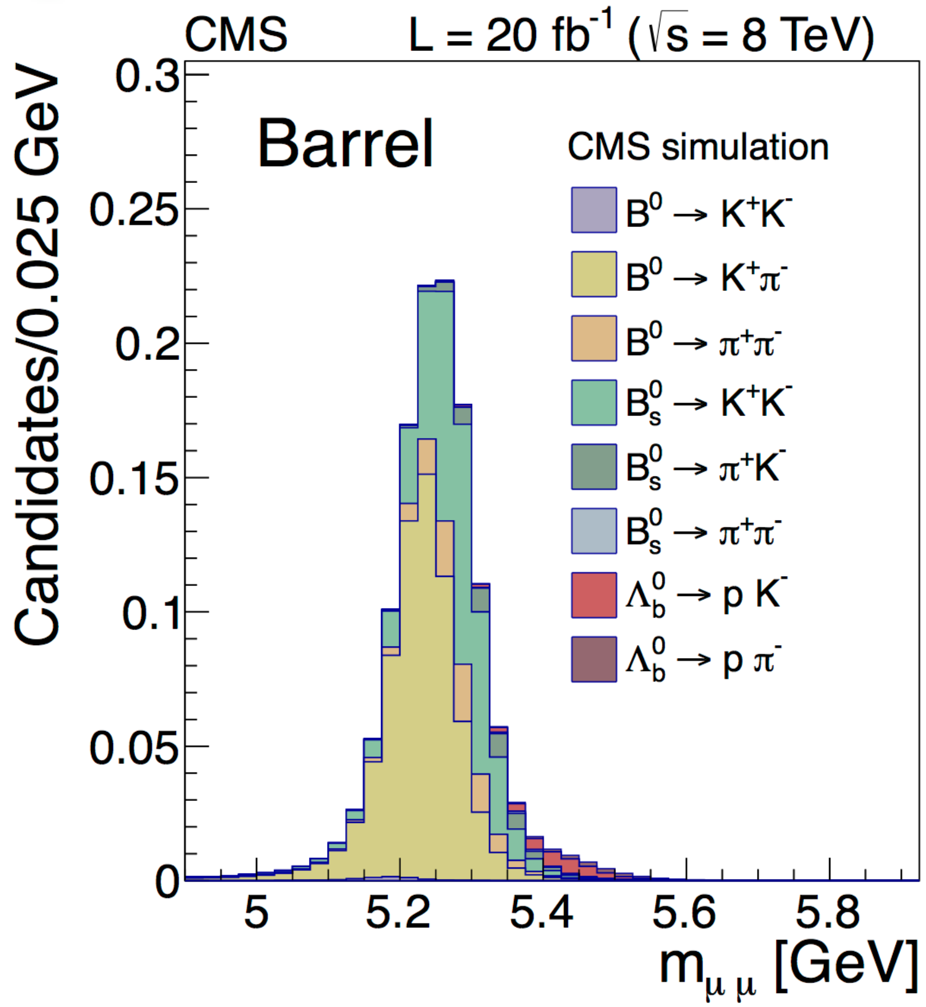


- Similar selection as for signal events, but
  - different mass window,
  - changed p_T thresholds,
  - all (decay) tracks used in vertex reconstruction.
- Yield extraction:
  - shape templates extracted from MC for signal and backgrounds,
  - estimated systematic error on the event yield, considering alternative fitting functions: 5%.
- Different mass resolution for different categories (7/8 TeV + barrel endcap) visible.

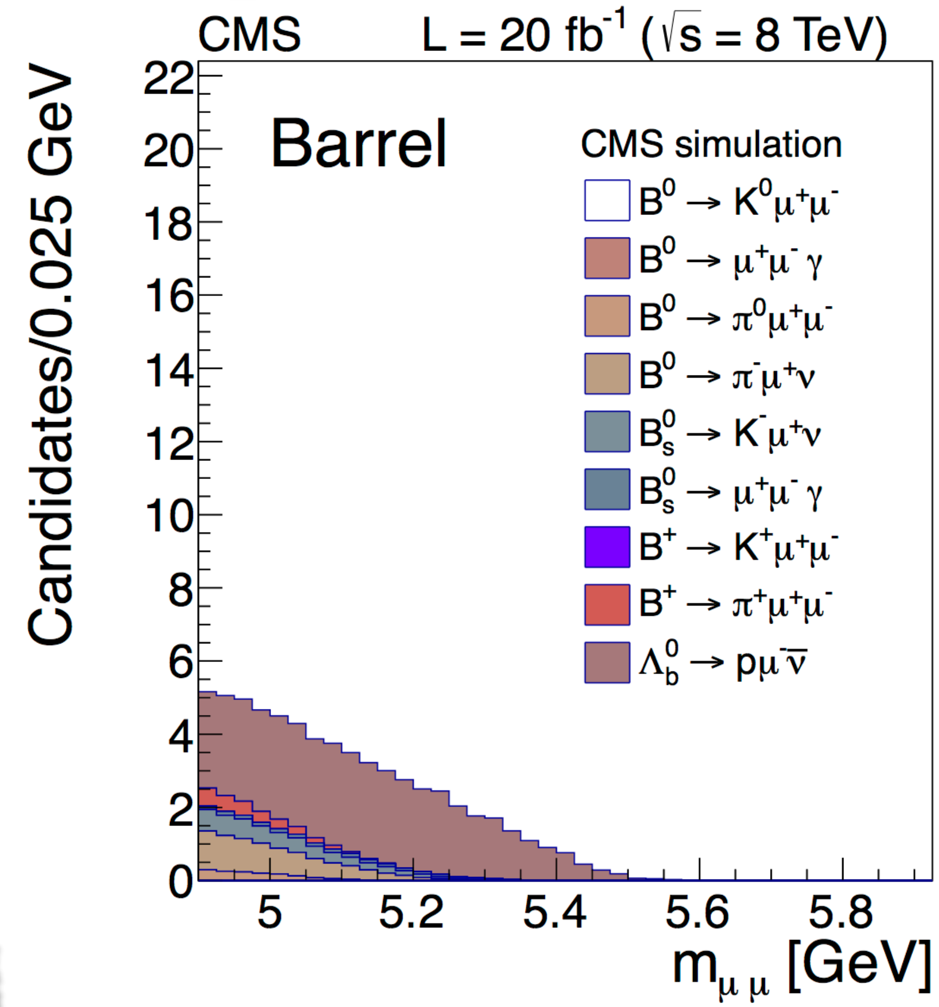
Rare backgrounds

Arno.Heister@cern.ch

“peaking”



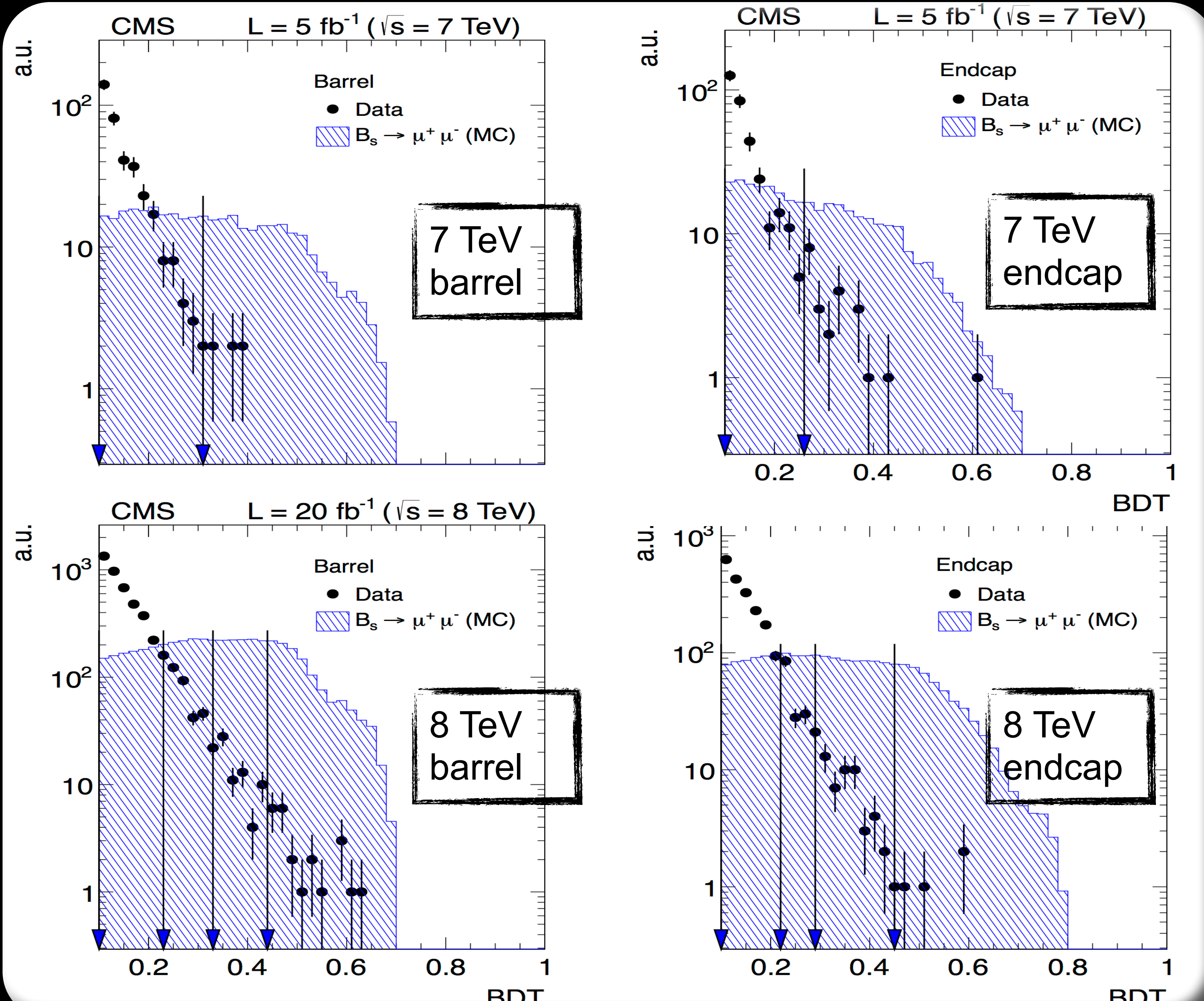
“non-peaking”



- Expected number of events in each channel normalized to $B^{\pm}$ in the analysis.
- Numbers weighted with muon mis-identification probability measured on data

BDT categories

Arno.Heister@cern.ch



1D-BDT categories:

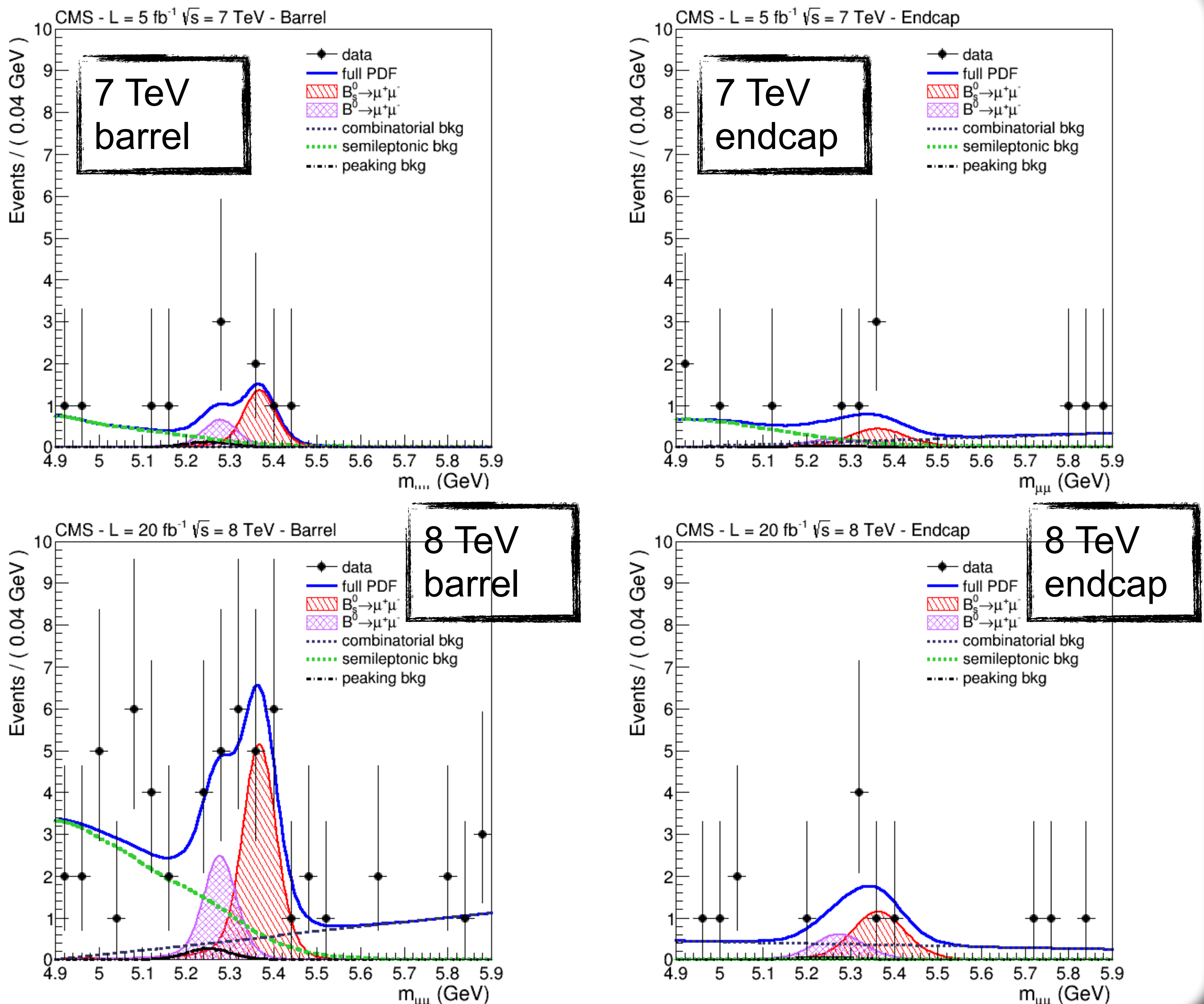
b >	barrel	endcap
2011	0.29	0.29
2012	0.36	0.38

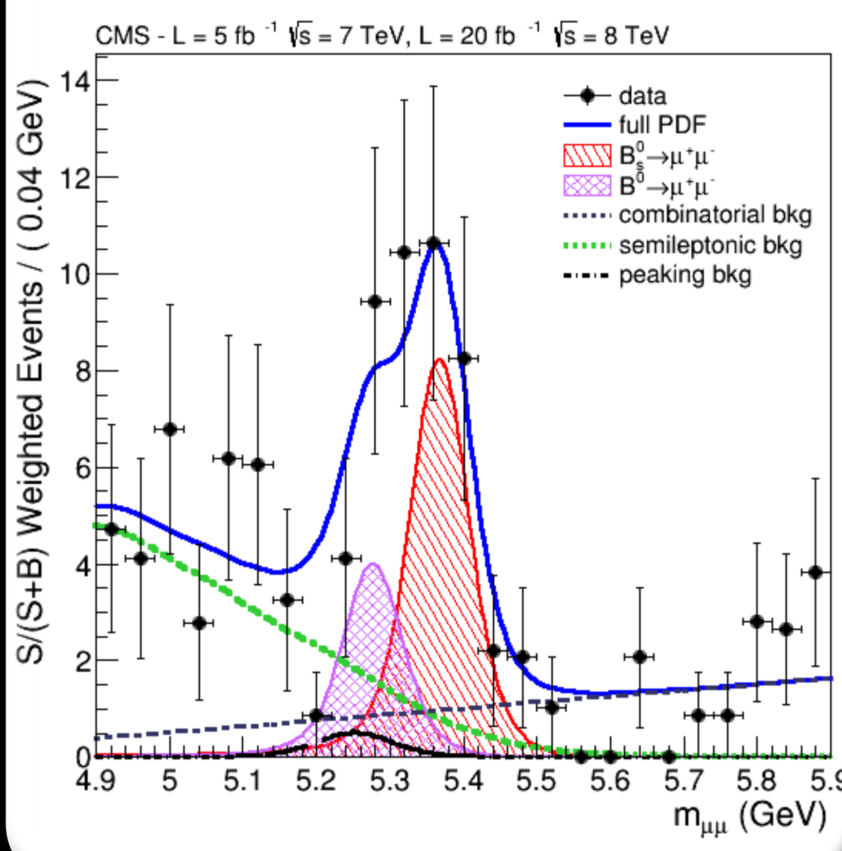
Output discriminant b of the signal selection BDT used in two ways:

- 1D-BDT method: final selection based on optimizing b for best $S/\sqrt{(S+B)}$
- categorized BDT method: final selection based on twelve categories based on, categories have same expected signal yield

1D-BDT: fit results

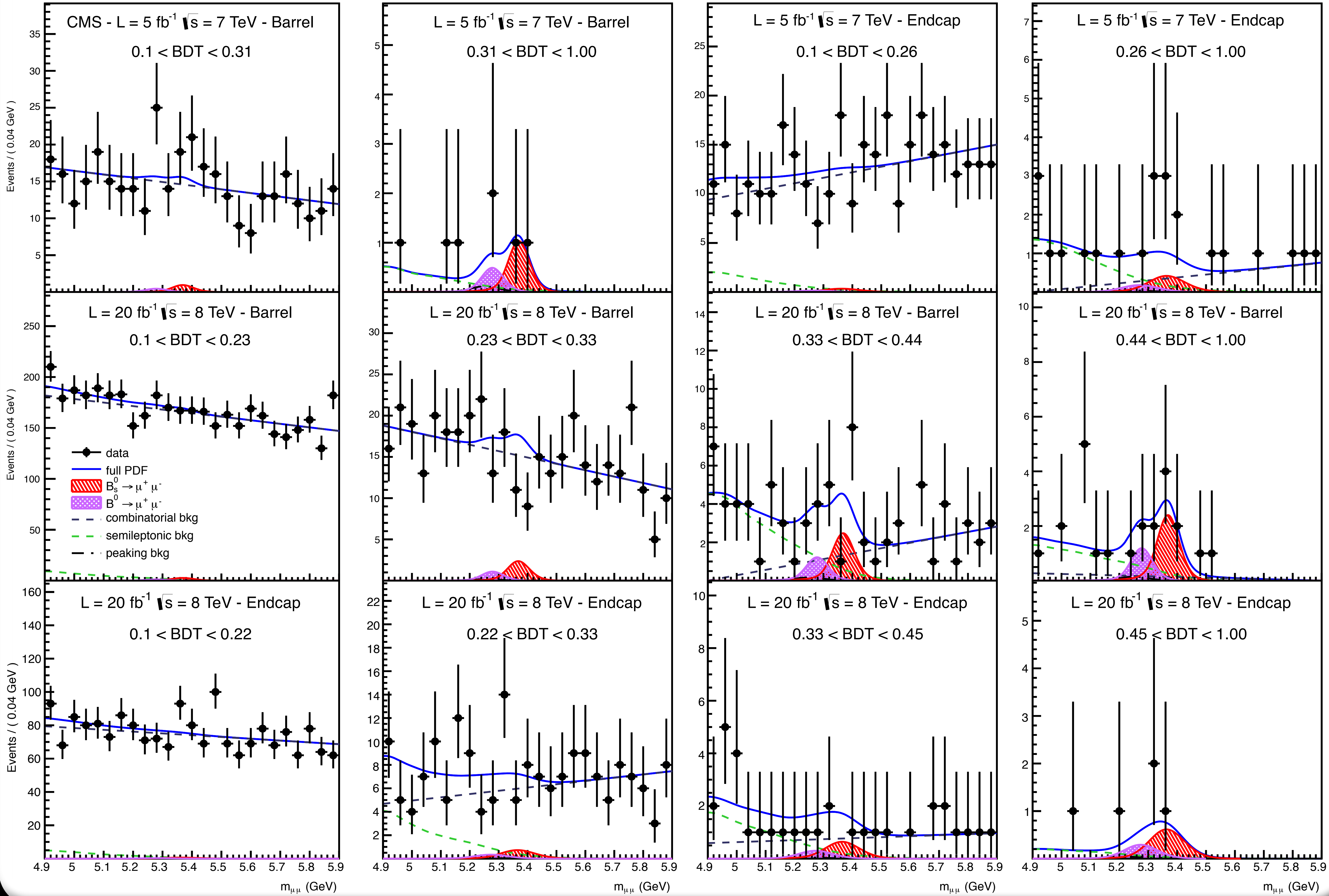
Arno.Heister@cern.ch





- $B_s \rightarrow \mu\mu$ significance: 4.8 σ (expected 4.7 median)
- individual categories are weighted with $S/\sqrt{(S + B)}$
- normalization: fitted B_s signal corresponds to the total yield of the individual contributions

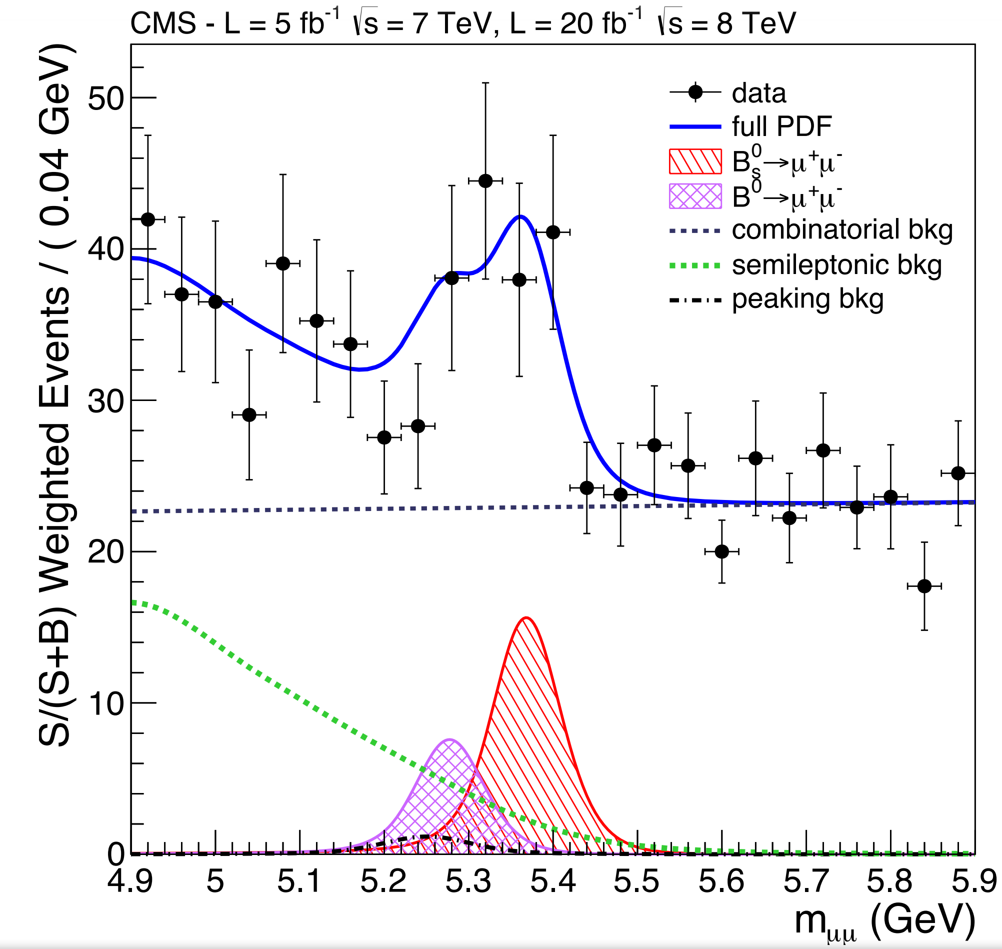
Categorized BDT: fit results



BR extraction

Arno.Heister@cern.ch

- B_s and B_d fitted simultaneously
- Signal
  - Crystal Ball
  - Normalization floating
- Peaking background
  - Sum of Gaussian and Crystal ball (same mean)
  - Constrained (Log-Normal) to expectation and normalized to the measured B^+ yield incl. check on independent data set
- Rare semi-leptonic background
  - Fixed shape, normalization floating constrained with 75% of nominal value
  - Constrained Gaussian kernels from MC data
- Combinatorial background
  - First degree polynomial incl. validation on independent data-set
- Per event mass resolution included



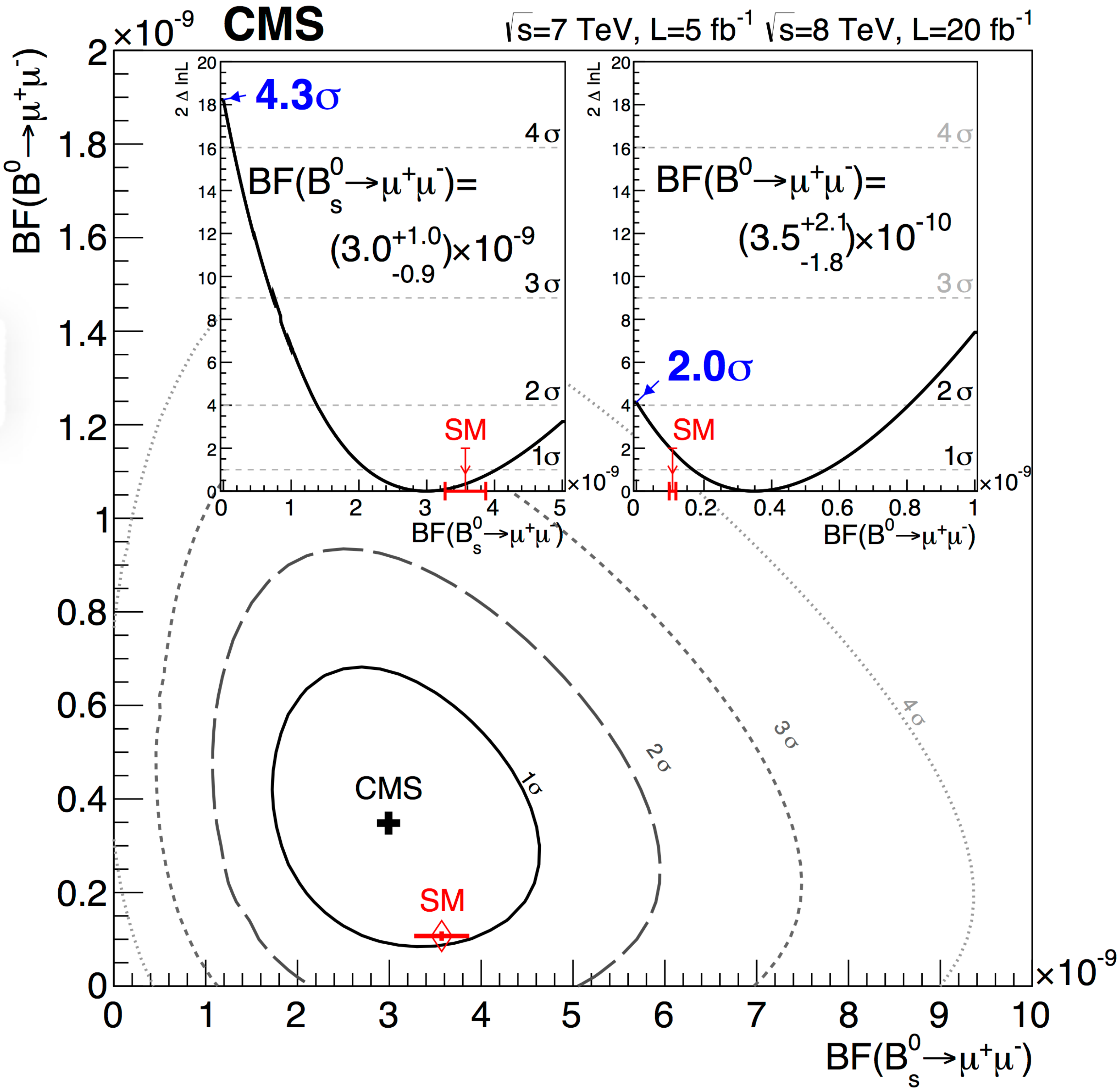
- $B_s \to \mu\mu$ significance: 4.3 σ (expected 4.8 median)
- individual categories are weighted with $S/\sqrt{(S + B)}$
- normalization: fitted B_s signal corresponds to the total yield of the individual contributions

Systematics

Arno.Heister@cern.ch

- Hadron-to-muon mis-identification
  - studied with $D^* \rightarrow D^0\pi$, $D^0 \rightarrow K\pi$, $K_s \rightarrow \pi\pi$, $\Lambda \rightarrow p\pi$
  - 50% uncertainty, conservatively assumed to be uncorrelated
- BR uncertainties
  - dominated by $\Lambda_b \rightarrow p\mu\nu$ (6.5×10^{-4}) with 100% uncertainty
- $f_s/f_u = 0.256 \pm 0.020$ (arXiv:1301.5286)
  - additional 5% to account for possible p_T and η dependance
  - no p_T dependance from ratios of $B^{\pm} \rightarrow J/\psi K^{\pm}$ vs $B_s \rightarrow J/\psi\varphi$
- Normalization channel
  - yields 5%
  - $BR(B_d \rightarrow J/\psi K^{\pm}) \times BR(J/\psi \rightarrow \mu\mu) = (6.0 \pm 0.2) \times 10^{-5}$ (PDG)

BR results

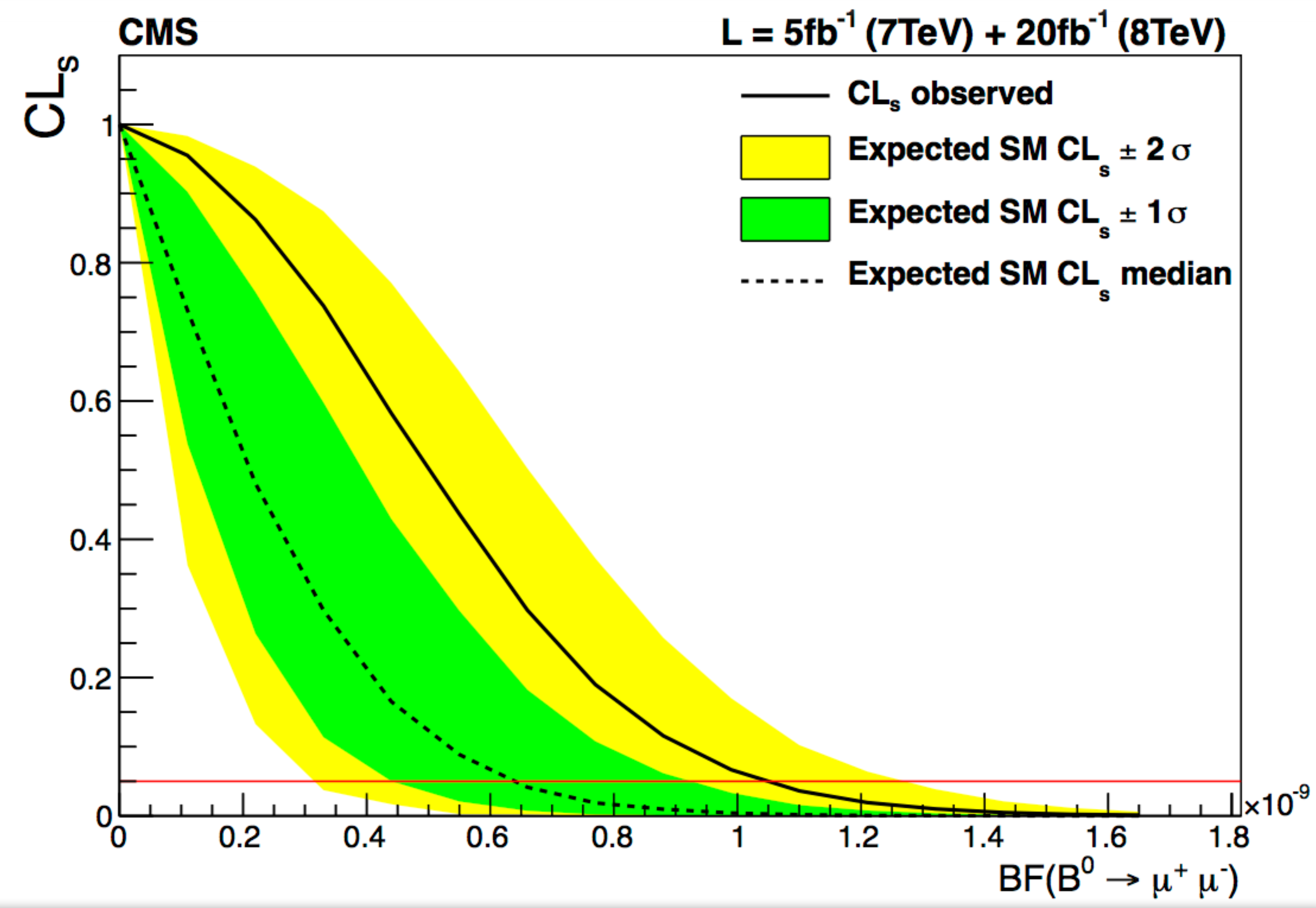


Arno.Heister@cern.ch

Upper limits on $B_d \rightarrow \mu\mu$

Arno.Heister@cern.ch

$BR(B_d \rightarrow \mu^+\mu^-) < 1.1 \times 10^{-9}$ @ 95% CL
(expected 6.3×10^{-10} in presence of SM+bkg)



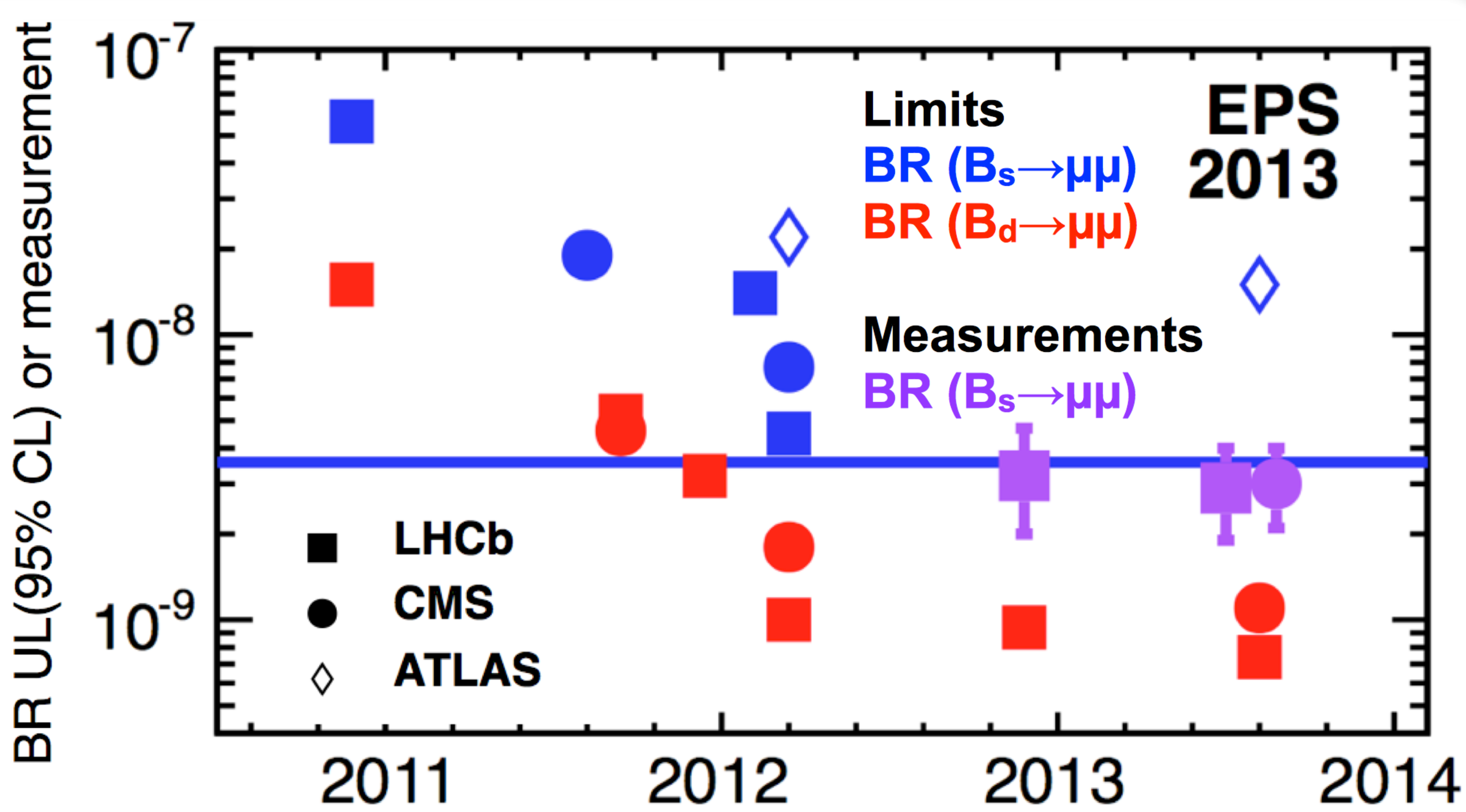
- Upper limit computed using the CLs method.
- No significant excess observed.

Summary and outlook

Arno.Heister@cern.ch

$$BR(B_s \to \mu^+\mu^-) = (3.0^{+0.9}_{-0.8}(\text{stat.})^{+0.6}_{-0.4}(\text{syst.})) \cdot 10^{-9}$$

$$BR(B_d \to \mu^+\mu^-) < 1.1 \cdot 10^{-9} \quad @\, 95\,\%\, CL$$

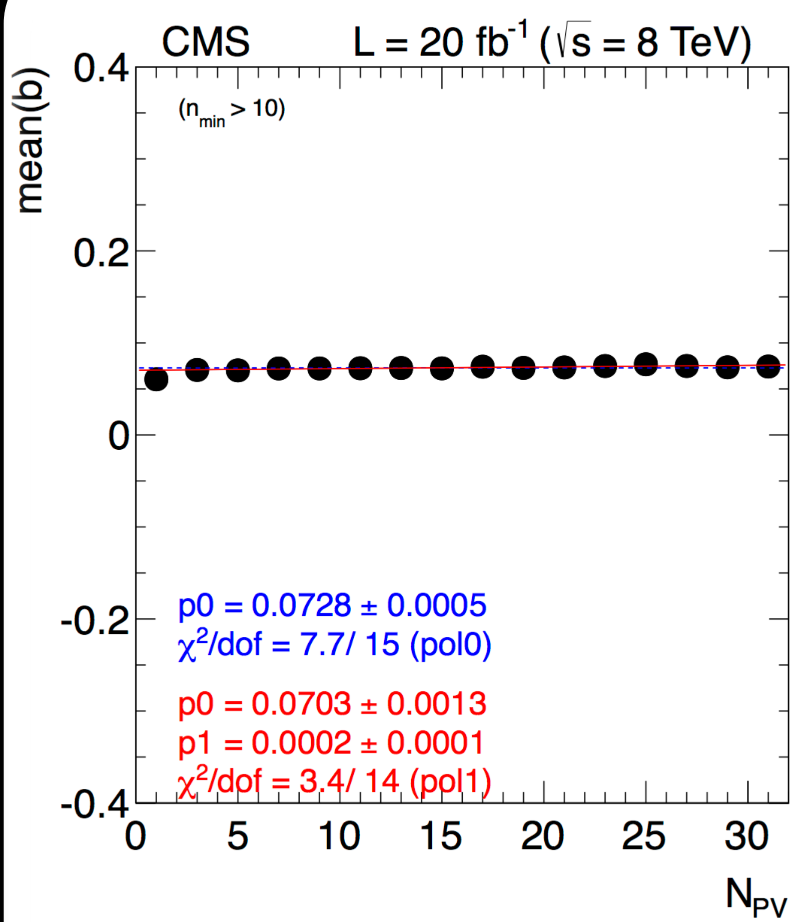


- Analysis statistically dominated, final LHC Run 1 statistics ~25 fb^{-1}
  - 4.3 σ significance (4.8 σ expected)
  - consistent with the Standard Model expectations
- LHC Run 2:
  - 300 fb^{-1} expected,
  - optimized analysis for B_d➔μμ,
  - trigger conditions challenging, e.g. 4x higher L1T rate in case of same selection, etc.

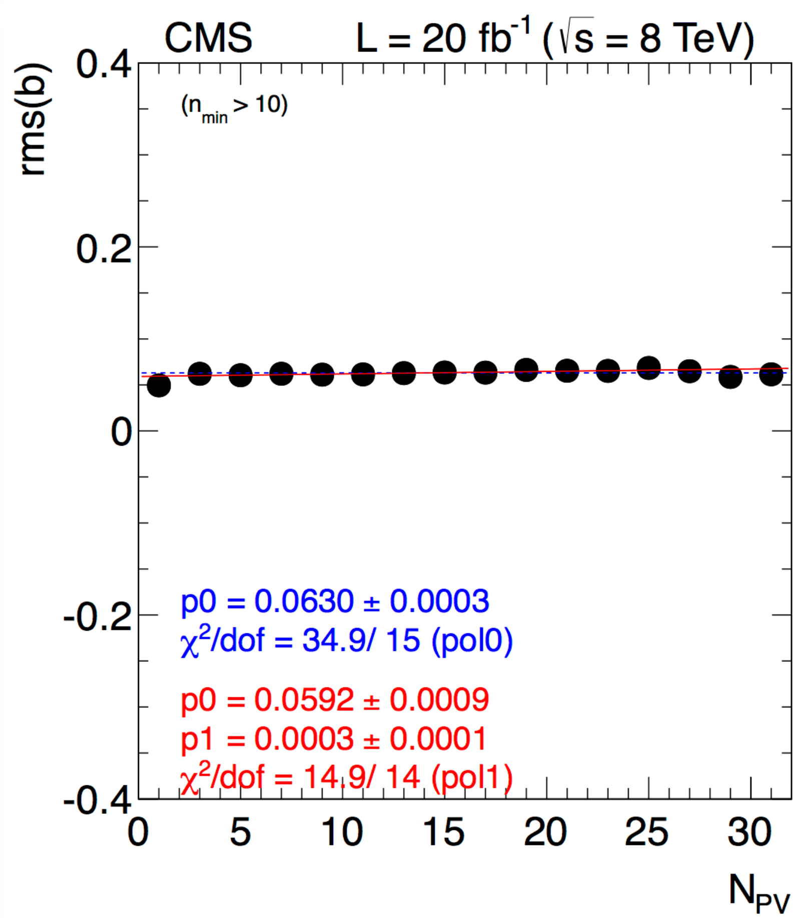
Backup slides ...

Pile-up

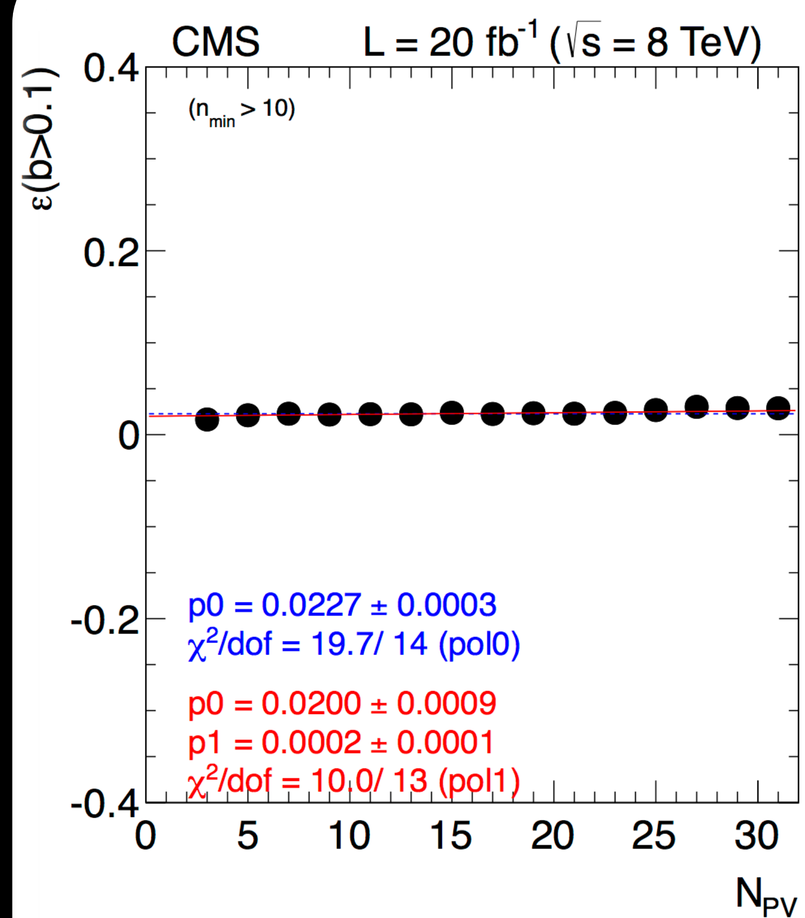
<9> @ 2011, <21> @ 2012



- Selections have been tuned to be pile-up independent, e.g. isolation searches only for tracks coming from the same primary vertex or not associated with any



- Every input variable has been evaluated to be insensitive versus the number of reconstructed primary vertices



- All selections are compatible with a constant efficiency up to at least 30 PV (~40 PU)