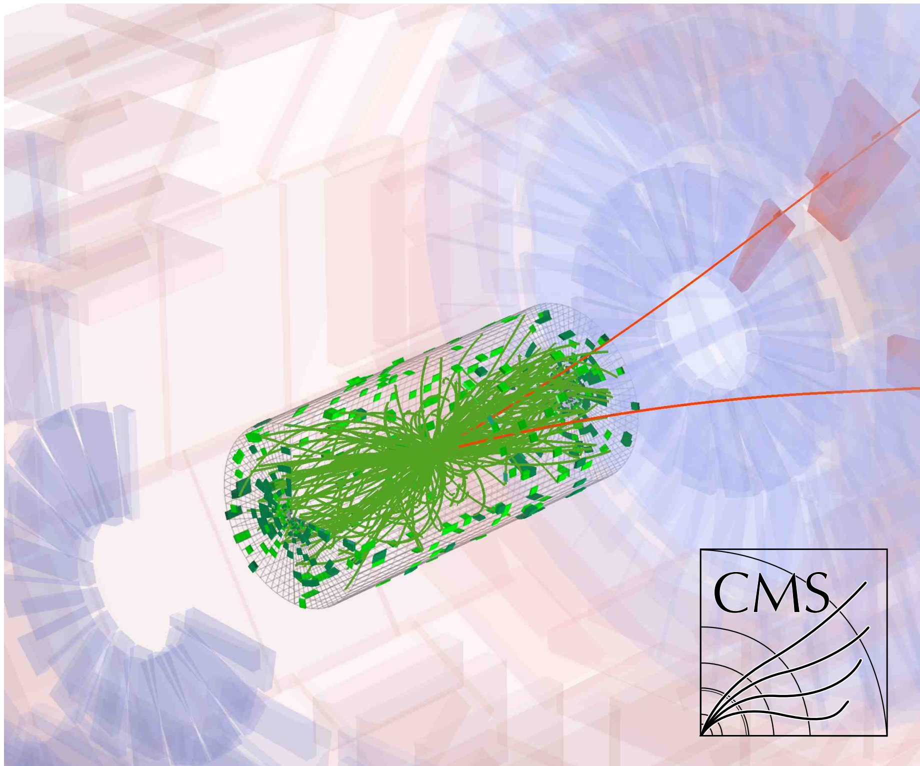
Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ in CMS (2011 dataset)

Urs Langenegger

(PSI)

on behalf of the CMS collaboration

2012/02/28



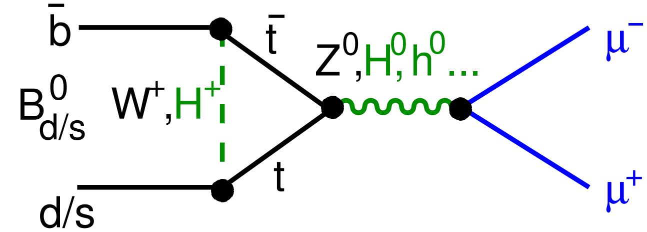
- Introduction
- Candidate(s) selection
- Data/MC validation
- Pileup independence
- Search Analysis
- Results

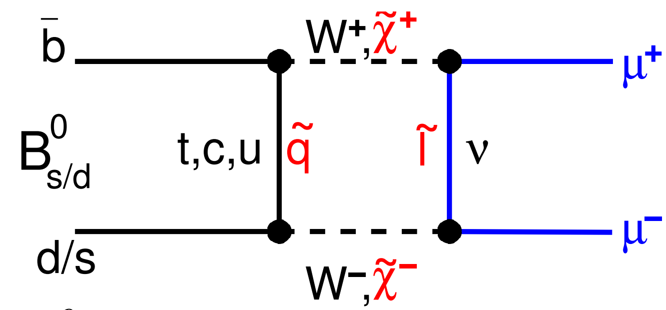
Motivation: Search for New Physics

- Decays highly suppressed in Standard Model (Buras 2010)
  - ▷ effective FCNC, helicity suppression
  - ▷ SM expectation:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.2 \pm 0.2) \times 10^{-9}$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.0 \pm 0.1) \times 10^{-10}$$

  - ▷ Cabibbo-enhancement $(|V_{ts}| > |V_{td}|)$ of $B_s^0 \to \mu^+\mu^-$ over $B^0 \to \mu^+\mu^-$ only in MFV models





- Sensitivity to new physics
  - ▷ 2HDM: $\mathcal{B} \propto (\tan\beta)^4, m_{H^+}$; MSSM: $\mathcal{B} \propto (\tan\beta)^6$
  - → sensitivity to extended Higgs boson sectors
  - → Constraints on parameter regions

- $B_s^0 \to \mu^+\mu^-$ (and $B^0 \to \mu^+\mu^-$) considered as golden channel(s)
  - ▷ high sensitivity to new physics
  - ▷ (very) small theoretical uncertainties
  - → comparable in sensitivity to $\mu \to e\gamma$, $B \to X\nu\bar{\nu}$

State of the art

- At the Tevatron

Upper limit	$B_s^0 \to \mu^+\mu^-$	$B^0 \to \mu^+\mu^-$
D0[1]	5.1×10^{-8}	n/a
CDF[2]	4.0×10^{-8}	6.0×10^{-9}

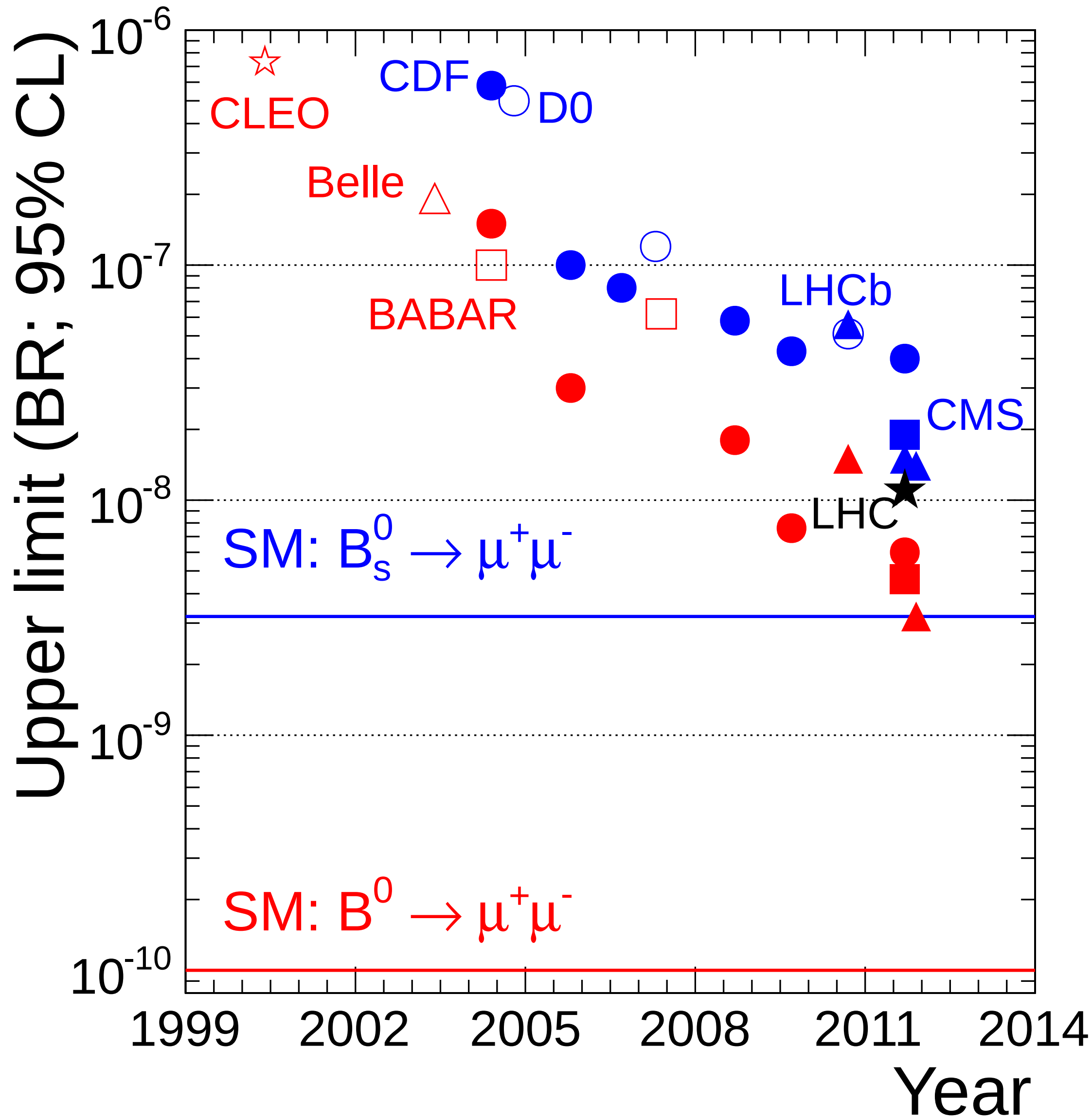
[1] 6.1 fb^{-1}, PL, B693, 539
[2] 7 fb^{-1}, PRL, 107, 191801

- At the LHC:

Upper limit	$B_s^0 \to \mu^+\mu^-$	$B^0 \to \mu^+\mu^-$
CMS[3]	1.9×10^{-8}	3.6×10^{-9}
LHCb[4]	1.4×10^{-8}	3.2×10^{-9}
CMS + LHCb	1.1×10^{-8}	n/a

[3] 1.1 fb^{-1}, PRL, 107, 191802
[4] 0.4 fb^{-1}, PL, B708, 55

(all upper limits at 95%CL)



- CDF[2] also has $\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (1.8^{+1.1}_{-0.9}) \times 10^{-8}$

A 'new' analysis

- Analysis was performed blind

  reblinded old data (1.1 fb^{-1})

  → total amount of data: 4.9 fb^{-1}

- Significant analysis modifications
  - ▷ tighter muon identification ($3\times$ smaller fake rate)
  - ▷ isolation variables

    primary vertex isolation (redefined)

    B vertex isolation: distance of closest track (redefined)

    B vertex isolation: track counting (new)

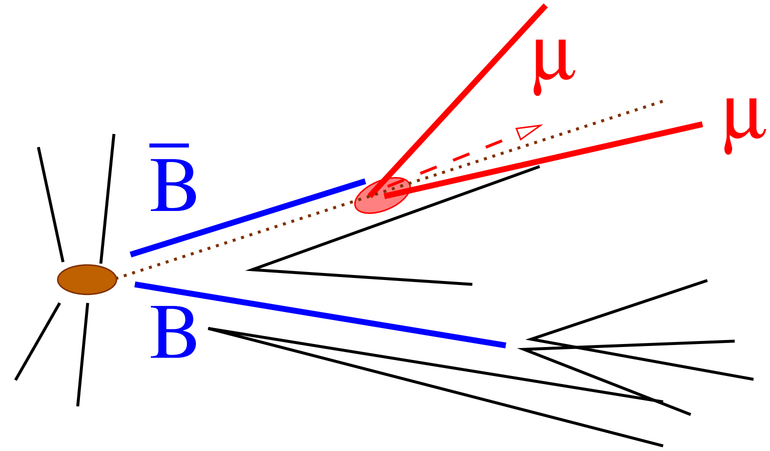
  → non-monotonous changes
  - ▷ 3D impact parameter and its significance (new)

⇒ Better analysis
  - ▷ pileup independence up to $N_{PV} \approx 30$
  - ▷ higher sensitivity
  - ▷ larger signal/background

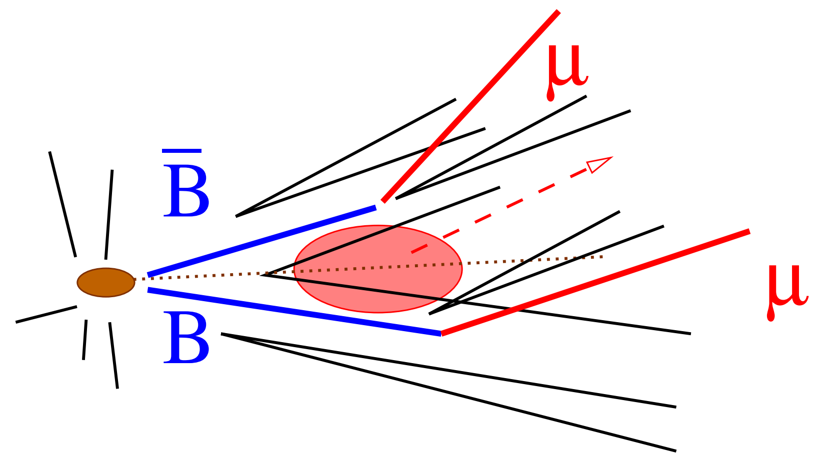
Analysis is (still) cut-n-count

Analysis overview

- Signal $B_s^0 \to \mu^+\mu^-$
  - ▷ two muons from one decay vertex
    - well reconstructed secondary vertex
    - momentum aligned with flight direction
    - long-lived B
    - mass around $m_{B_s^0}$



- Background
  - ▷ two semileptonic (B) decays (gluon splitting)
  - ▷ one semileptonic (B) decay and one misidentified hadron
  - ▷ rare single B decays
    - peaking, e.g. $B_s^0 \to K^+K^-$
    - non-peaking, e.g. $B_s^0 \to K^-\mu^+\nu$
  - $\to$ mass resolution
  - $\to$ not well-reconstructed secondary vertex
  - $\to$ pointing angle



$\Rightarrow$ **High signal efficiency and high background reduction**

Methodology

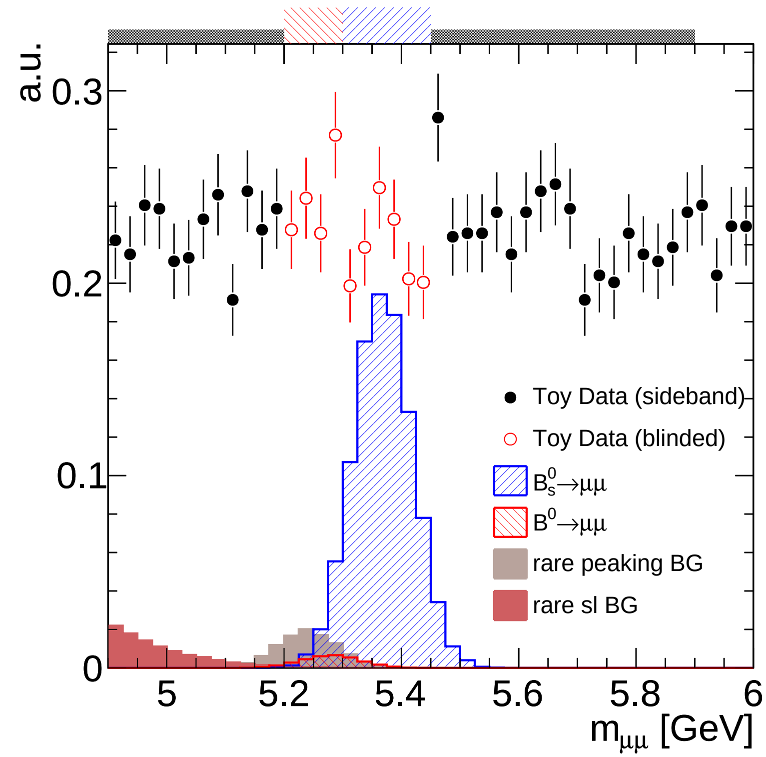
- Measurement of $B_s^0 \to \mu^+\mu^-$ relative to normalization channel:
  - ▷ similar trigger and selection to reduce systematic uncertainties

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-; 95\%\mathrm{C.}L.) = \frac{N(n_{obs}, n_B, n_S; 95\%\mathrm{C.}L.)}{\varepsilon_{B_s^0} N_{B_s^0}} = \frac{N(n_{obs}, n_B, n_S)}{\varepsilon_{B_s^0} \mathcal{L}\,\sigma(pp \to B_s^0)}$$

$$= \frac{N(n_{obs}, n_B, n_S)}{N(B^\pm \to J/\psi\, K^\pm)} \frac{A_{B+}}{A_{B_s^0}} \frac{\varepsilon^{ana}_{B+}}{\varepsilon^{ana}_{B_s^0}} \frac{\varepsilon^{\mu}_{B+}}{\varepsilon^{\mu}_{B_s^0}} \frac{\varepsilon^{trig}_{B+}}{\varepsilon^{trig}_{B_s^0}} \frac{f_u}{f_s} \mathcal{B}(B^+ \to J/\psi\,[\mu^+\mu^-]K)$$

- Calibration/validation of MC:
  - ▷ $B^\pm \to J/\psi\, K^\pm$ normalization with high statistics
  - ▷ $B_s^0 \to J/\psi\, \phi$ B_s^0 signal MC ($p_\perp$, isolation, ...)

- Analysis in two channels
  - ▷ barrel (both muons $|\eta| < 1.4$):
    better signal/background ratio
    good mass resolution (36 MeV)
  - ▷ endcap (at least one muon with $|\eta| > 1.4$):
    add more statistics [$\sigma(m) \approx 70$ MeV]



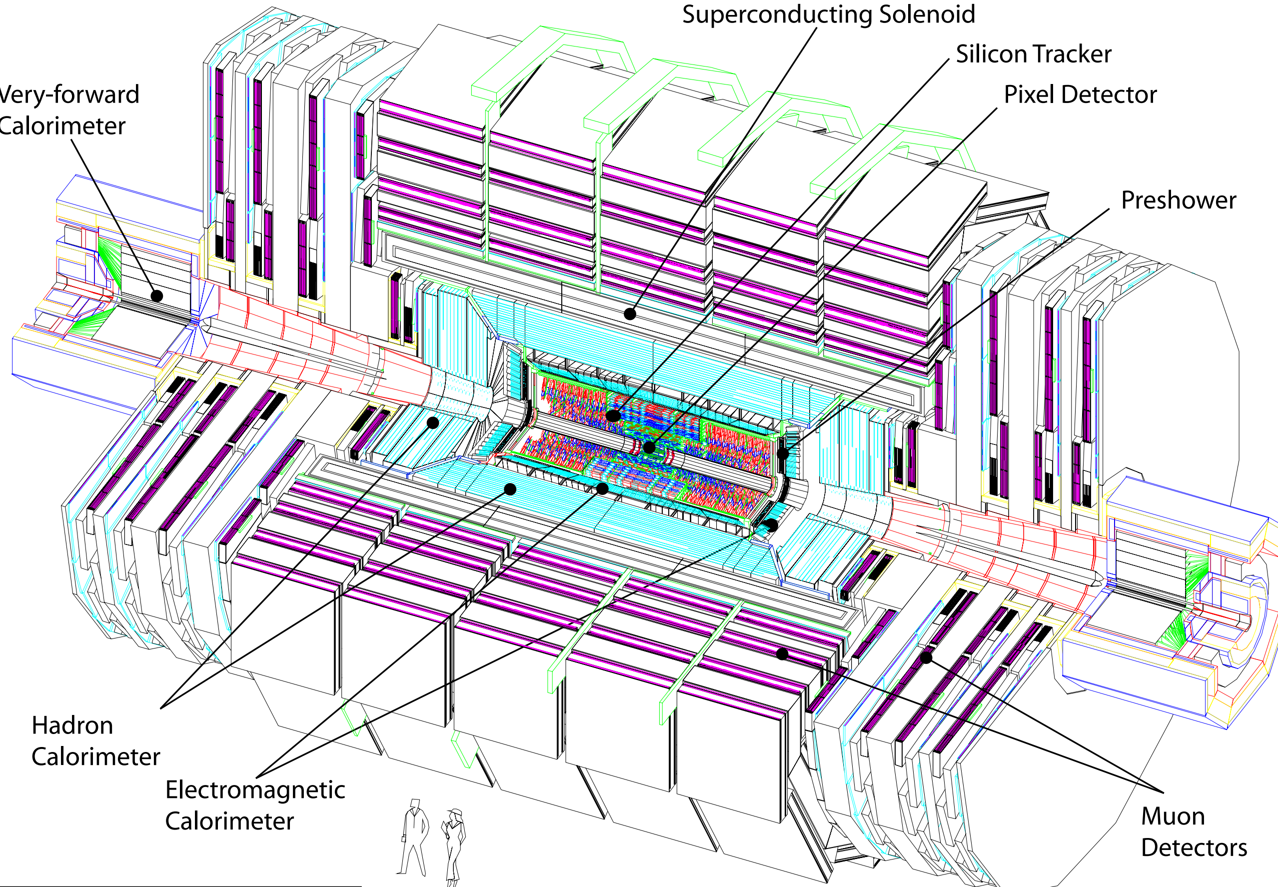
⇒ Blind analysis

JINST 3, S08004 (2008)

The CMS detector

- Design prioritization
  - ▷ lepton ID $\rightarrow$ muons
  - ▷ b/τ tagging $\rightarrow$ tracking
  - ▷ jets and $\not\!\!E_T$

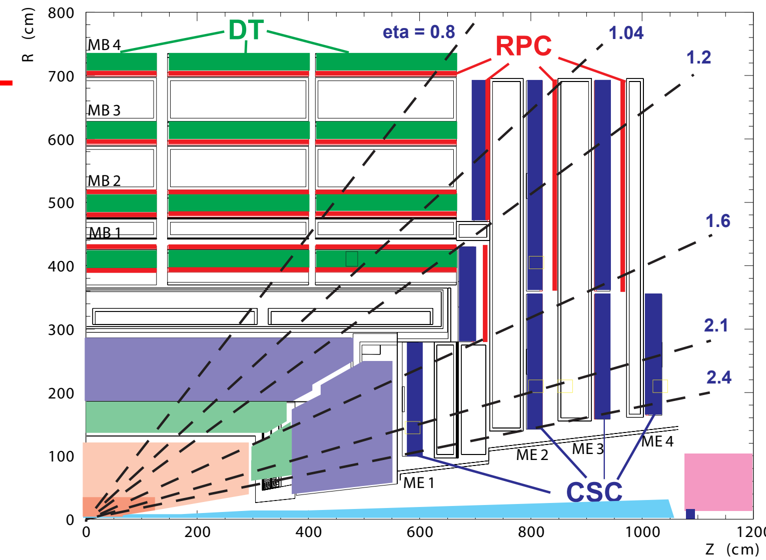
Weight	12'500 t
Length	21.6 m
Diameter	15 m
Magnetic field	3.8 T



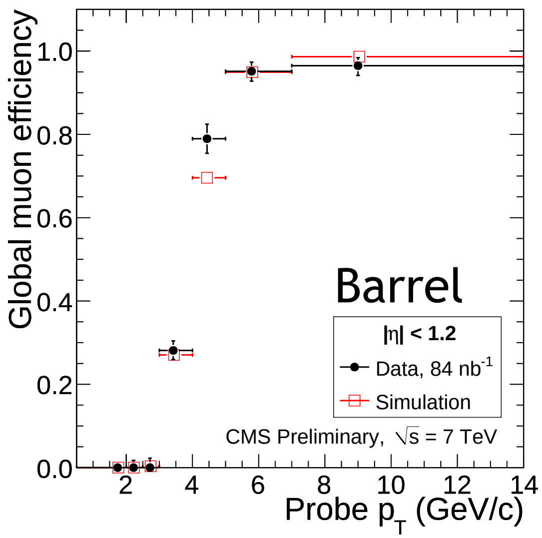
Component	Characteristics	Resolutions
Pixel	3/2 Si layers	$\delta_z \approx 20$ μm, $\delta_\phi \approx 10$ μm
Tracker	10/12 Si strips	$\delta(p_\perp)/p_\perp \approx 1\%$
ECAL	$PbWO_4$	$\delta E/E \approx 3\%/\sqrt{E} \oplus 0.5\%$
HCAL (B)	Brass/Sc, $> 7.2\lambda$	$\delta E/E \approx 100\sqrt{E}\%$
HCAL (F)	Fe/Quartz	$\delta(\not\!\!E_T) \approx 0.98\sqrt{\sum E_T}$
Magnet	3.8 T solenoid	
Muons	DT/CSC + RPC	$\delta(p_\perp)/p_\perp \approx 10\%$ (STA)

Tracking resolution:
impact parameter ≈ 15 μm

Muon reconstruction



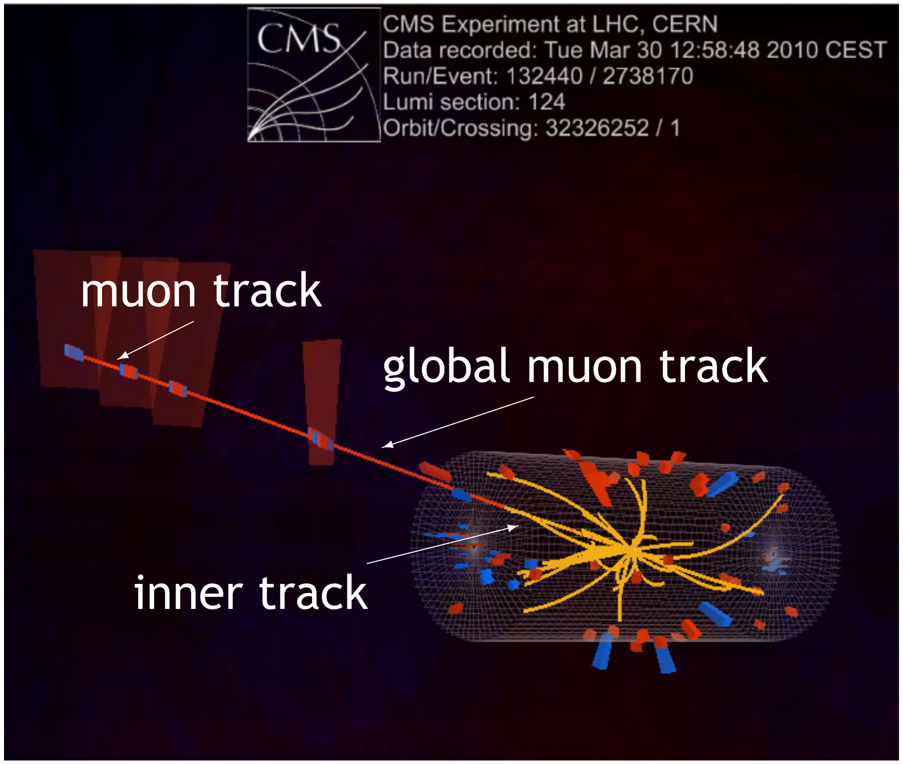
- Large muon acceptance $|\eta| < 2.4$
  - ▷ drift tubes
  - ▷ cathode strip chambers
  - ▷ resistive plate chambers
- 3 muon reconstruction algorithms
  - ▷ standalone muon: in muon system (trigger ingredient)
  - ▷ global muon ('GM'): outside-in standalone muon $\rightarrow$ to inner track
  - ▷ tracker muon ('TM'): inside-out inner track $\rightarrow$ muon detector



Muon misidentification

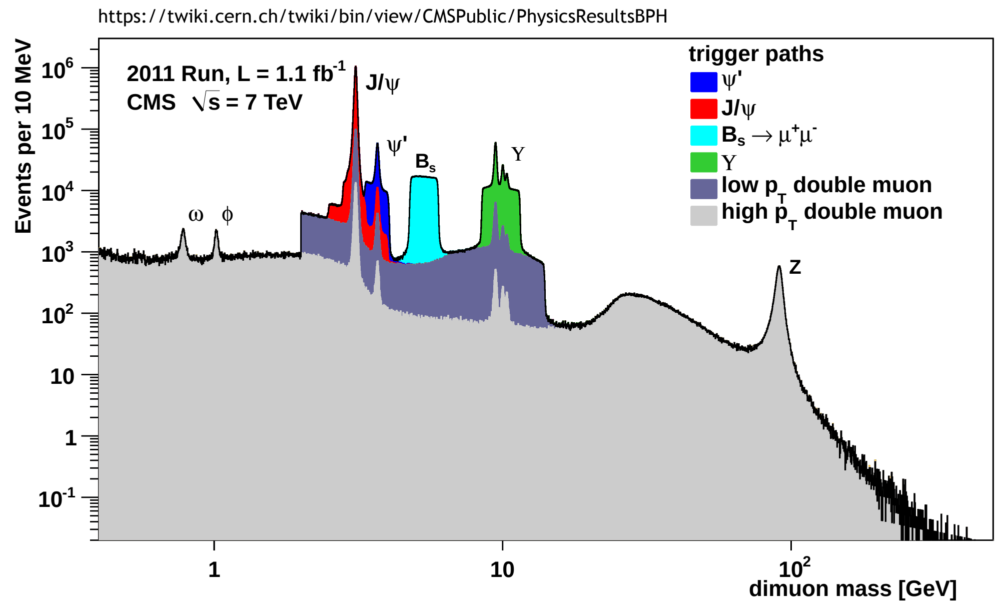
$$\varepsilon(\mu|\pi) \leq 0.1\%$$
$$\varepsilon(\mu|K) \leq 0.1\%$$
$$\varepsilon(\mu|p) \leq 0.05\%$$

measured in data:
$D^{*+} \rightarrow D^0\pi_s^+ \rightarrow K^-\pi^+\pi_s^+$
$\Lambda \rightarrow p\pi^-$



Trigger

- Dimuon trigger
  - L1 (hardware) trigger
  - High-level trigger
    full tracking and vertexing
  - requirements tightened over time

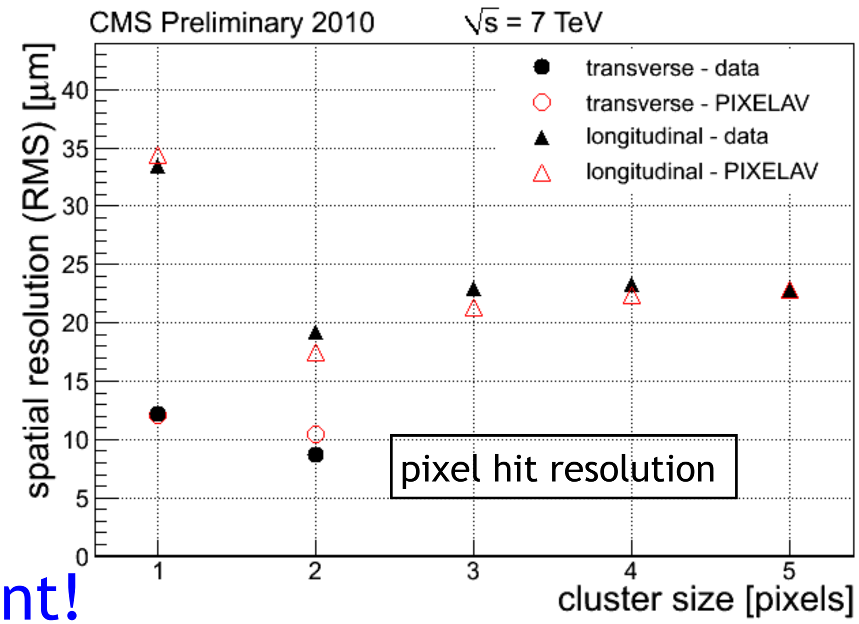


- HLT $B_s^0 \to \mu^+\mu^-$
  - inv. mass $4.8 < m_{\mu^+\mu^-} < 6.0$ GeV
  - dimuon vertex $\mathcal{P}(\chi^2, \mathrm{dof}) > 0.5\%$
  - distance of closest approach $d_{ca} < 0.5$ cm
  - single muon $p_\perp > 4$ GeV, dimuon $p_\perp > 3.9(5.9)$ GeV in barrel (endcap)

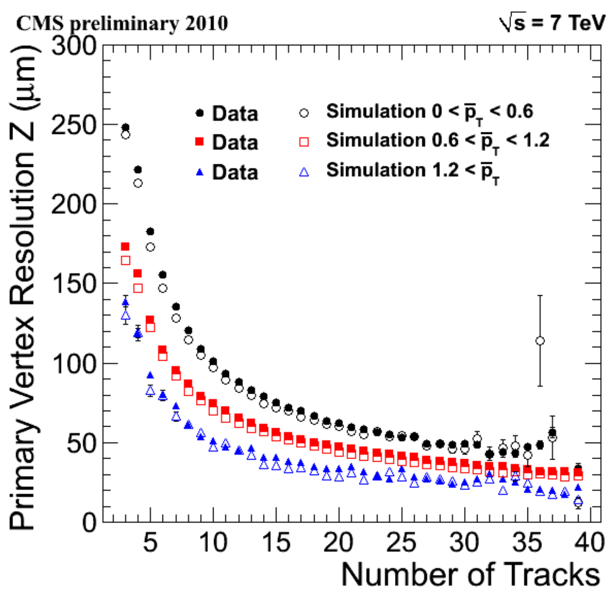
- HLT $B^\pm \to J/\psi\, K^\pm$ and $B_s^0 \to J/\psi\, \phi$
  - single muon $p_\perp > 4$ GeV, dimuon $p_\perp > 6.9$ GeV
  - distance of closest approach among muons $d_{ca} < 0.5$ cm
  - invariant dimuon mass $2.9 < m_{\mu^+\mu^-} < 3.3$ GeV
  - pointing angle $\cos\alpha_{xy} > 0.9$ and dimuon vertex $\mathcal{P}(\chi^2/dof) > 15\%$
  - 'displaced' J/ψ: flight length significance $\ell/\sigma(\ell) > 3$

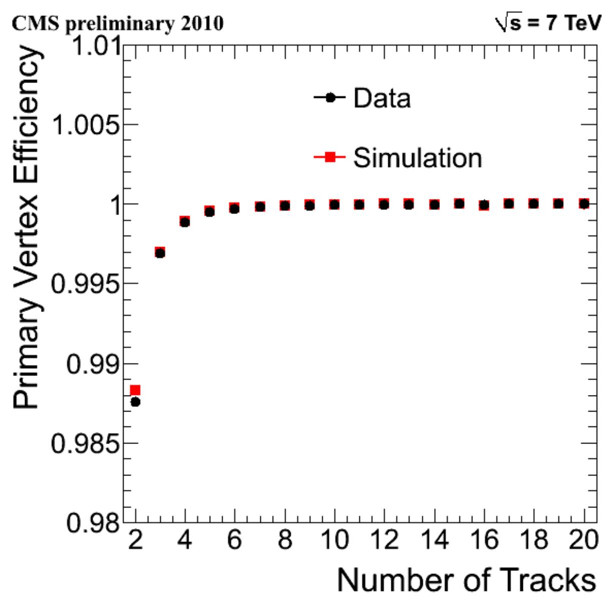
3D vertexing

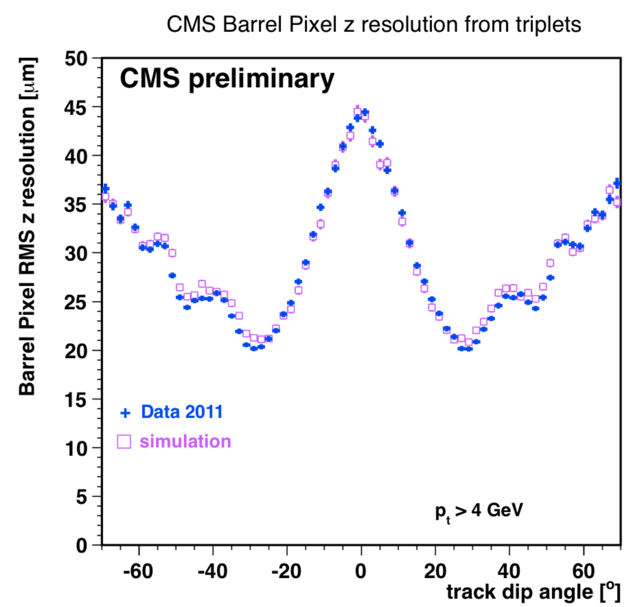
- All silicon tracker
  - ▷ high granularity, low occupancy
  - → very well described by MC simulation
- Pixel detector
  - ▷ $100 \times 150\ \mu\mathrm{m}^2$ pixel size
  - ▷ substantial charge sharing
  - ▷ excellent resolution



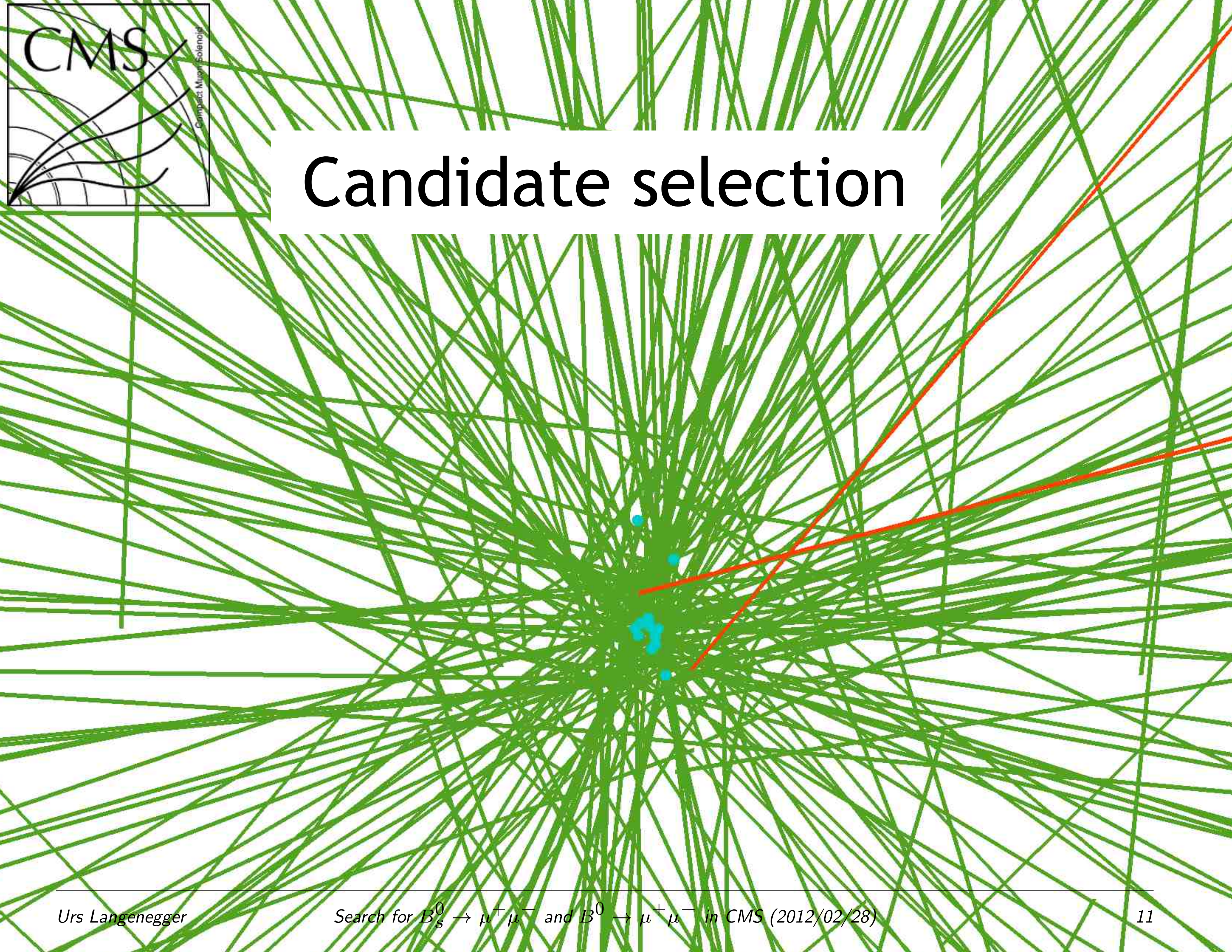
⇒ **Essential in high-pileup environment!**







Candidate selection

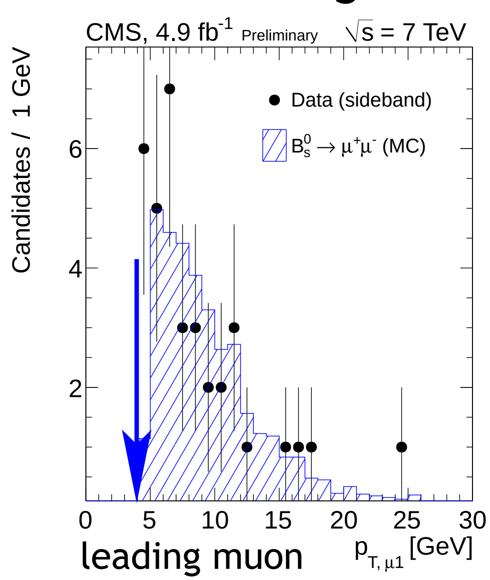


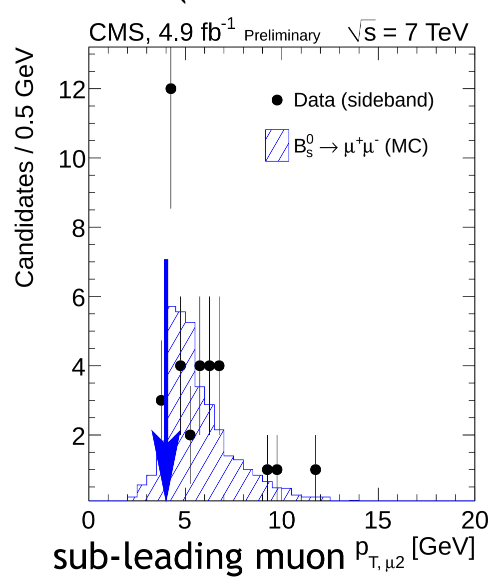
Two analyses

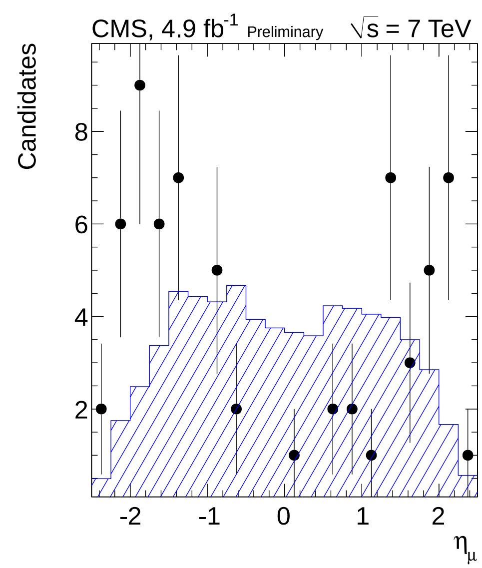
- 1. Search analysis $B \to \mu^+\mu^-$ in two channels
  - ▷ barrel (both muons $|\eta| < 1.4$):
  - ▷ endcap (≥ 1 muon with $|\eta| > 1.4$):
- 2. Validation analysis in one channel
  $B^\pm \to J/\psi K^\pm$ and $B_s^0 \to J/\psi \phi$ (and dimuons)

similar selection, but not identical

- Overlays of data and MC simulation (selection summary on p. 20)
  - ▷ 'all other' selection criteria are applied
    MC signal
    data background in sidebands ($4.9 < m < 5.2$ GeV and $5.45 < m < 5.9$ GeV)

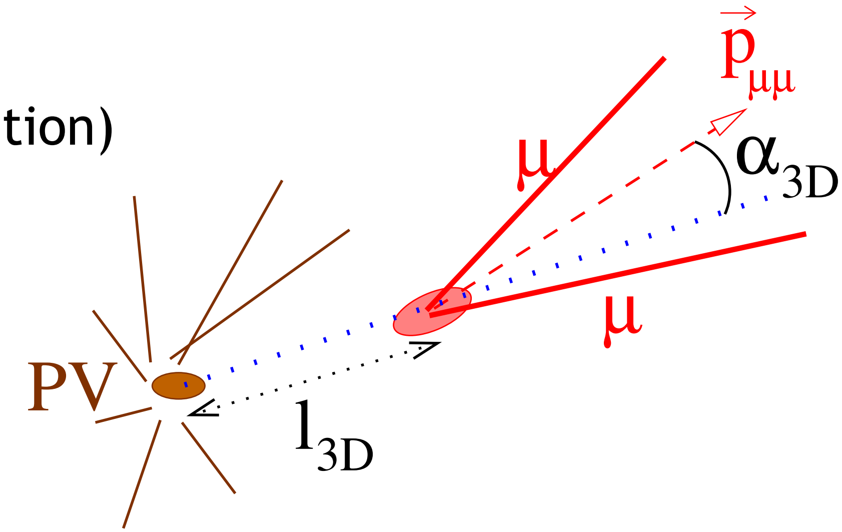


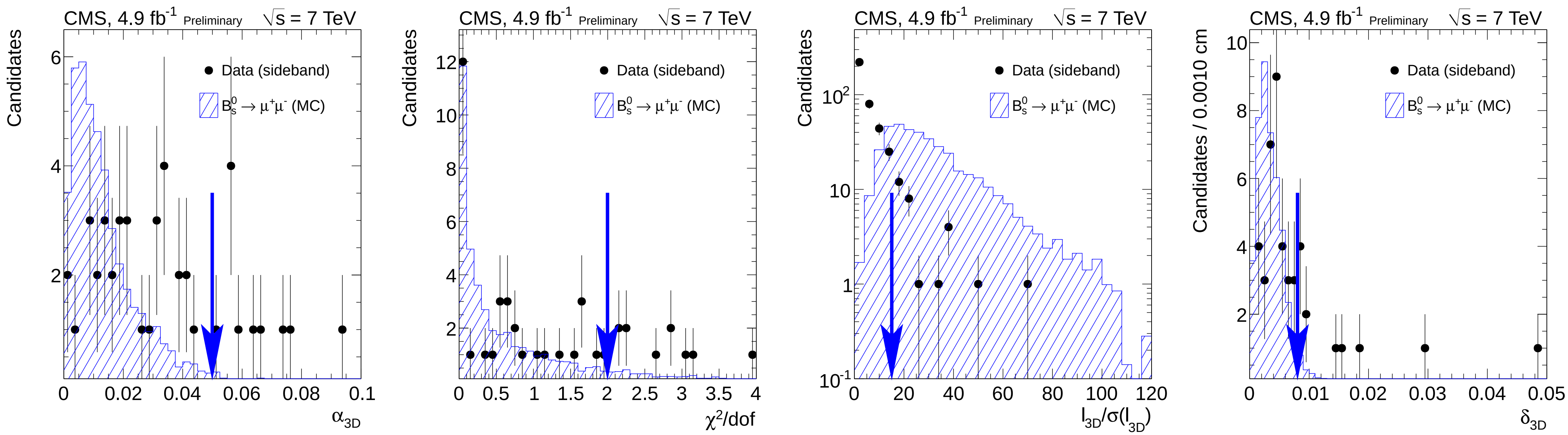




Signal selection: vertexing

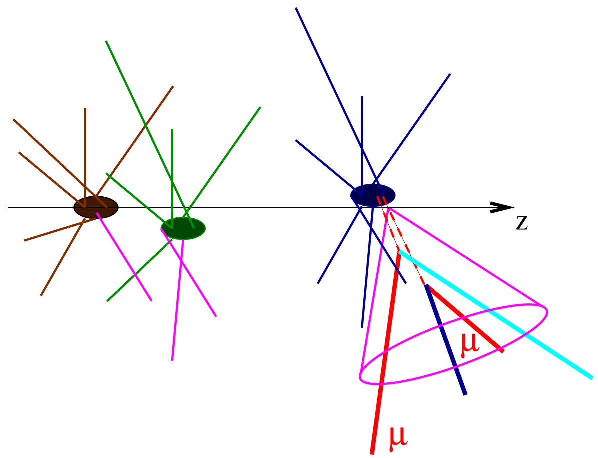
- Choose one primary vertex
  - ▷ longitudinal impact parameter (z position)
  - ▷ refit without signal tracks
- Discriminating variables
  - ▷ pointing angle α_{3D}
  - ▷ B vertex fit quality χ^2/dof
  - ▷ flight length significance $\ell_{3d}/\sigma(\ell_{3d})$
  - ▷ 3D impact parameter δ_{3D} and significance $\delta_{3D}/\sigma(\delta_{3D})$



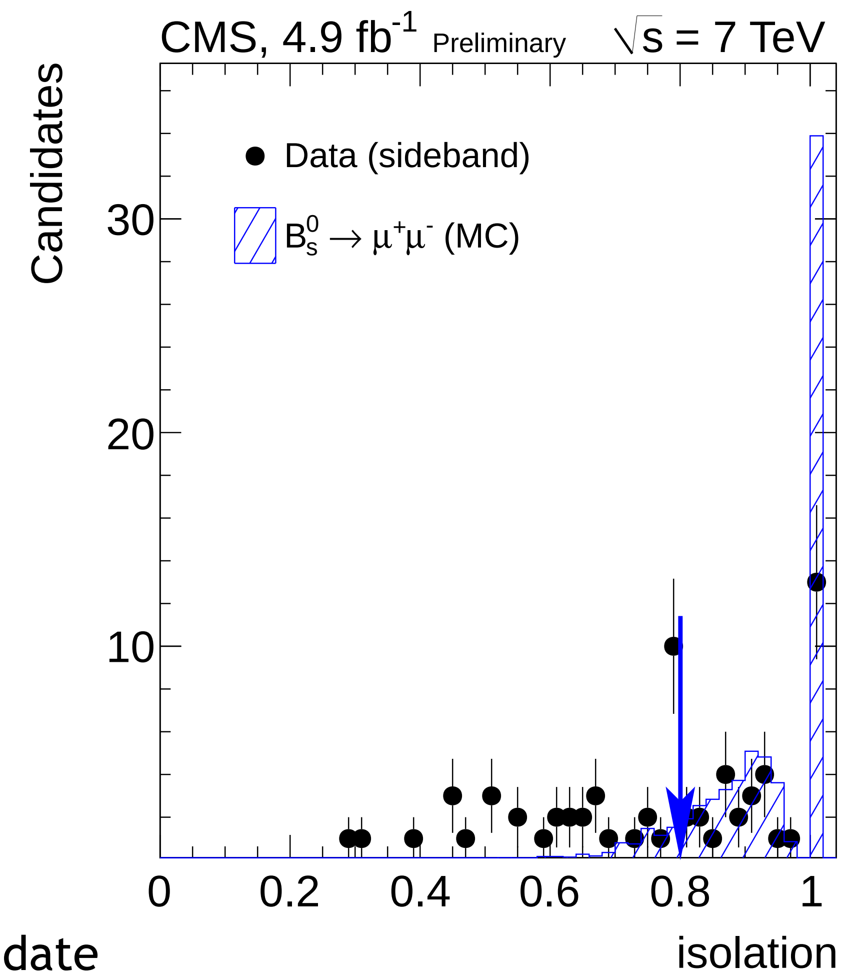


Isolation I

- Primary vertex isolation: Relative dimuon isolation
  - ‘classic’ variable



$$I = \frac{p_\perp(\mu^+\mu^-)}{p_\perp(\mu^+\mu^-) + \sum_{\Delta R<0.7} p_\perp}$$

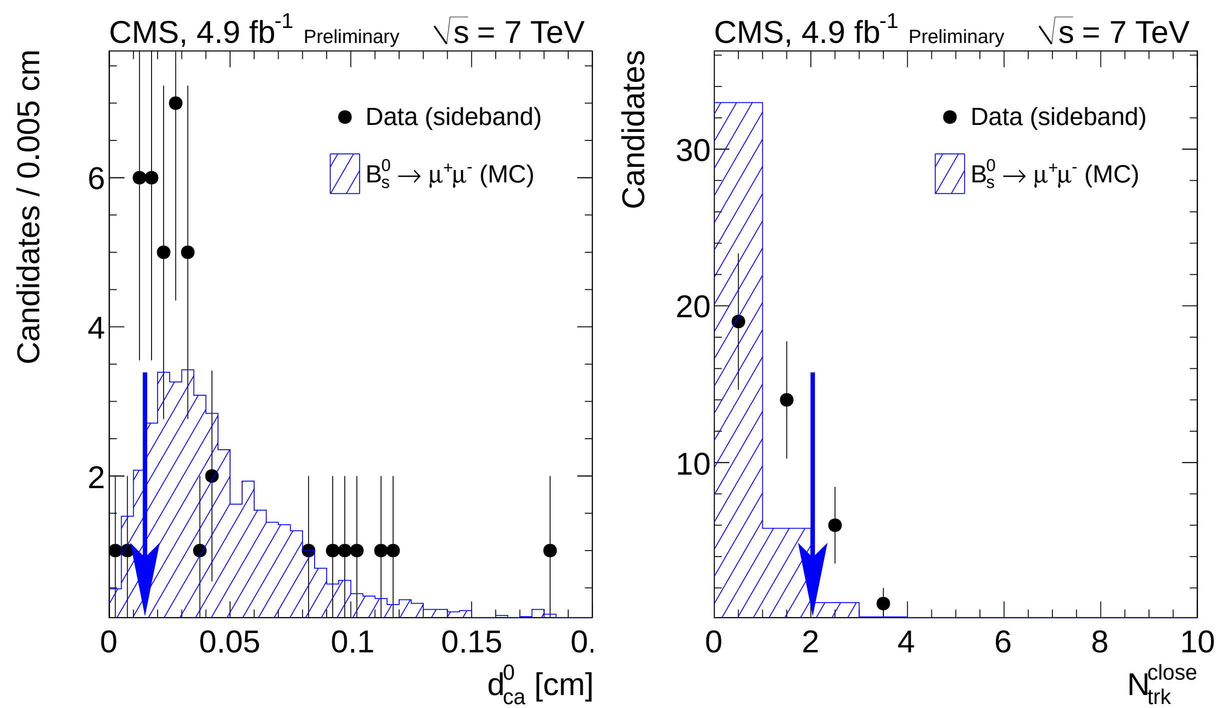


  - in cone around dimuon momentum
  - for tracks in cone with $\Delta R < 0.7$
    - with $p_\perp > 0.9$ GeV
    - either associated to same PV as candidate or with $d_{ca} < 500\,\mu$m and not associated to another PV

parameters tuned to minimize data/MC discrepancy ($B^\pm \to J/\psi\, K^\pm$) and maximize dimuon bg rejection

Isolation II

- B vertex isolation:
  - ▷ based on tracks reconstructed in the proximity of the secondary B vertex
  - → avoid pileup dependence:
    either tracks associated to no primary vertex
    or tracks associated to same vertex as B candidate
  - ▷ $d^0_{\rm ca}$: distance of closest track to B vertex
  - ▷ $N^{\rm close}_{\rm trk}$: number of close tracks
    $d_{ca} < 300\,\mu$m
    $p_\perp > 0.5$ GeV

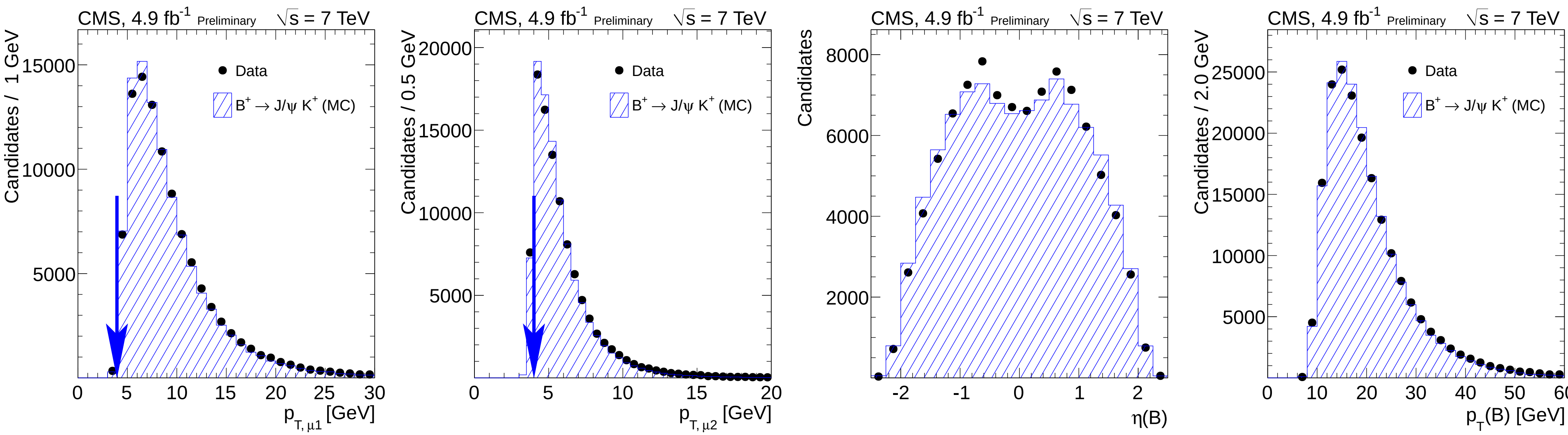


- Validation of B^0_s MC:
  - ▷ $B^0_s \to J/\psi\,\phi$!
    (see below)

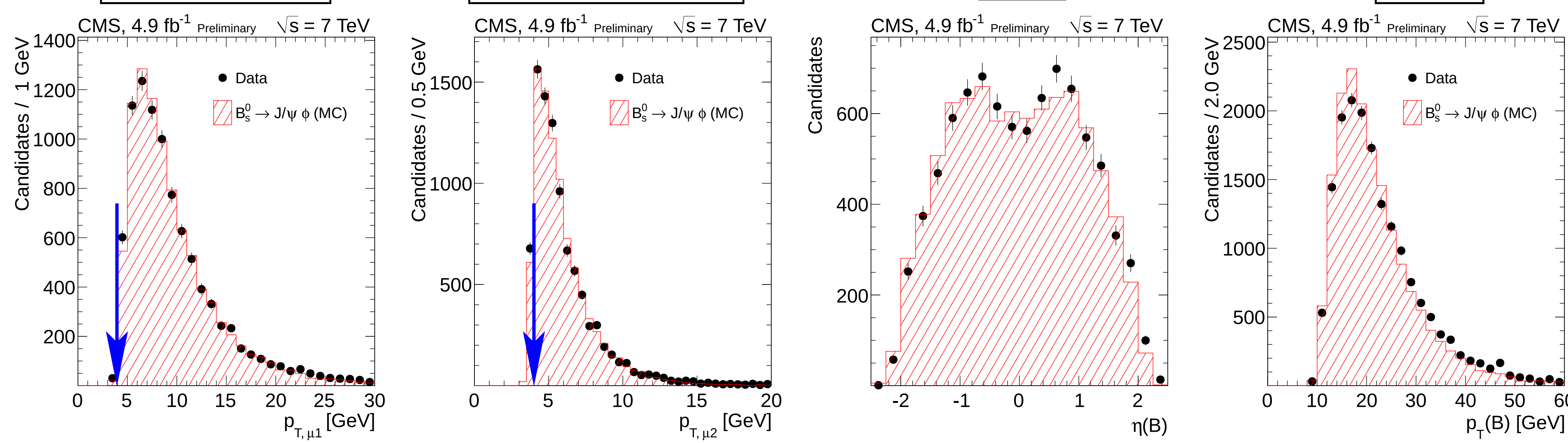
Normalization and control samples

- Normalization sample
  - $B^{\pm} \to J/\psi\, K^{\pm}$
  - validation of B^{+} MC
- Control sample
  - $B_s^0 \to J/\psi\, \phi$
  - validation of B_s^0 signal MC
- Combine J/ψ with 1 or 2 'kaons'
  - $3.0 < m(\mu\mu) < 3.2$ GeV
  - $p_{\perp}(\mu\mu) > 7$ GeV
  - $p_{\perp}(K) > 0.5$ GeV
  - additional selection for ϕ:
    $0.995 < m(KK) < 1.045$ GeV
    $\Delta R(K, K) < 0.25$
  - all 3 (4) tracks used in vertexing
- Comparison of (sideband-subtracted) data and MC simulation
  MC simulation normalized to data

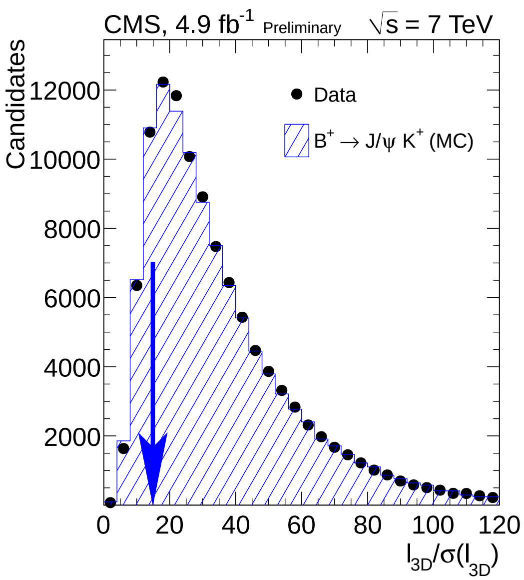
Kinematics

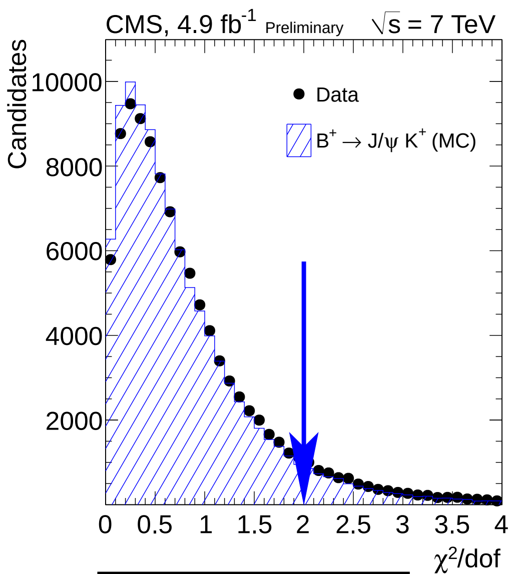


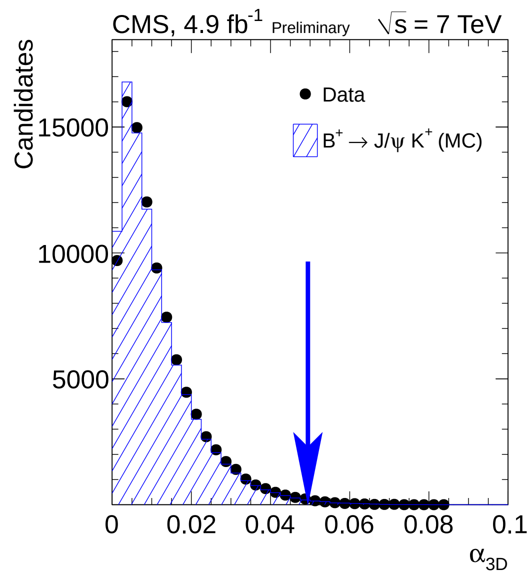
leading muon $p_\perp$	sub-leading muon $p_\perp$	$B\ \eta$	$B\ p_\perp$

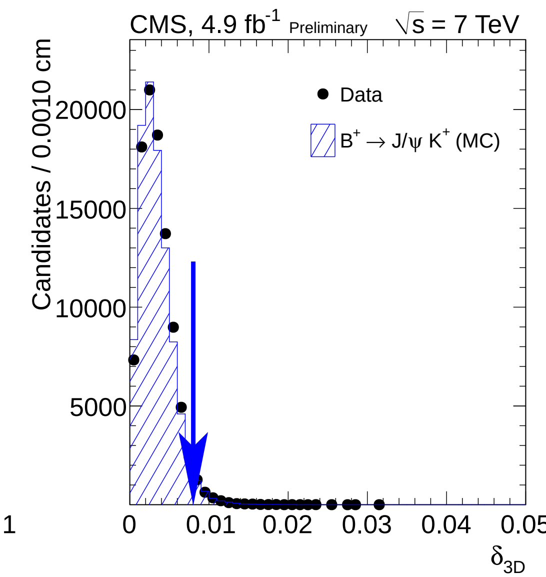


Vertexing

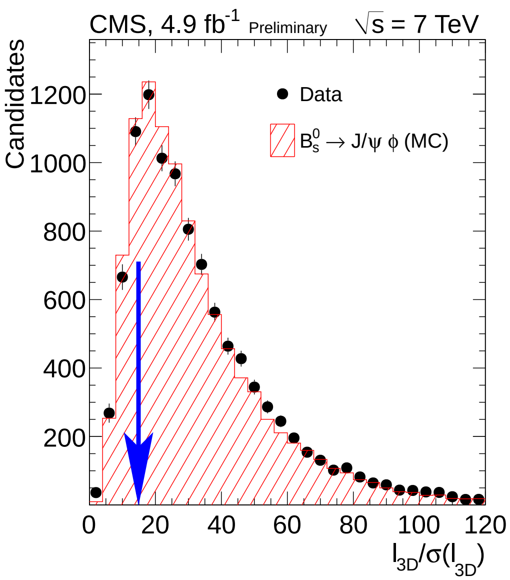


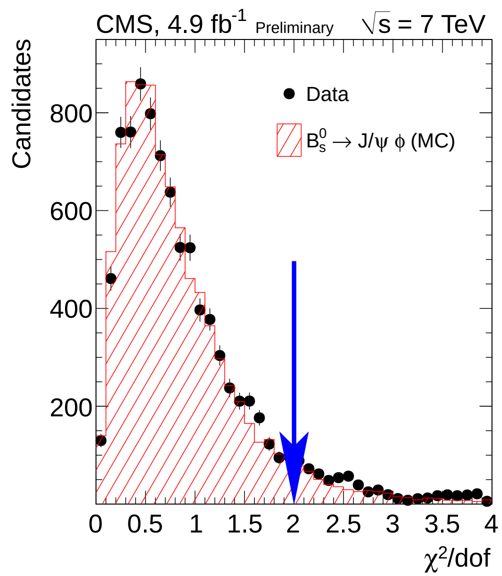


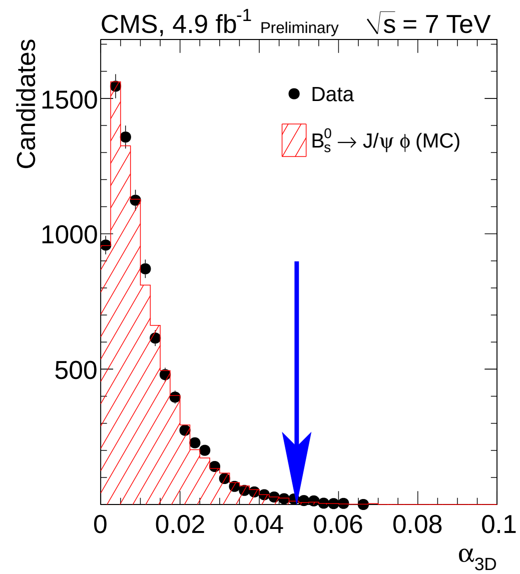


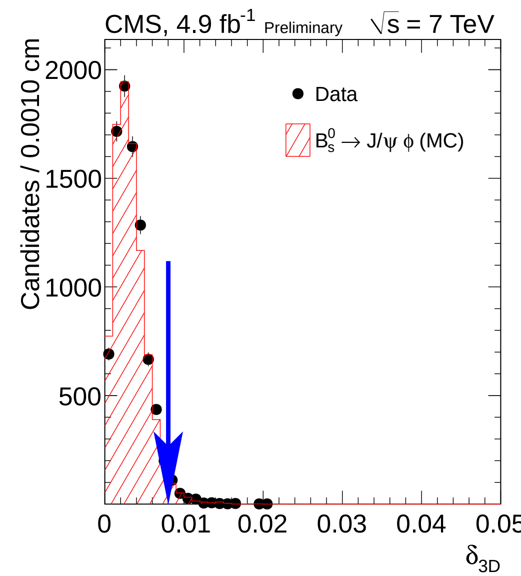


flight length sign. | B vertex χ^2/dof | 3D pointing angle | 3d impact param.

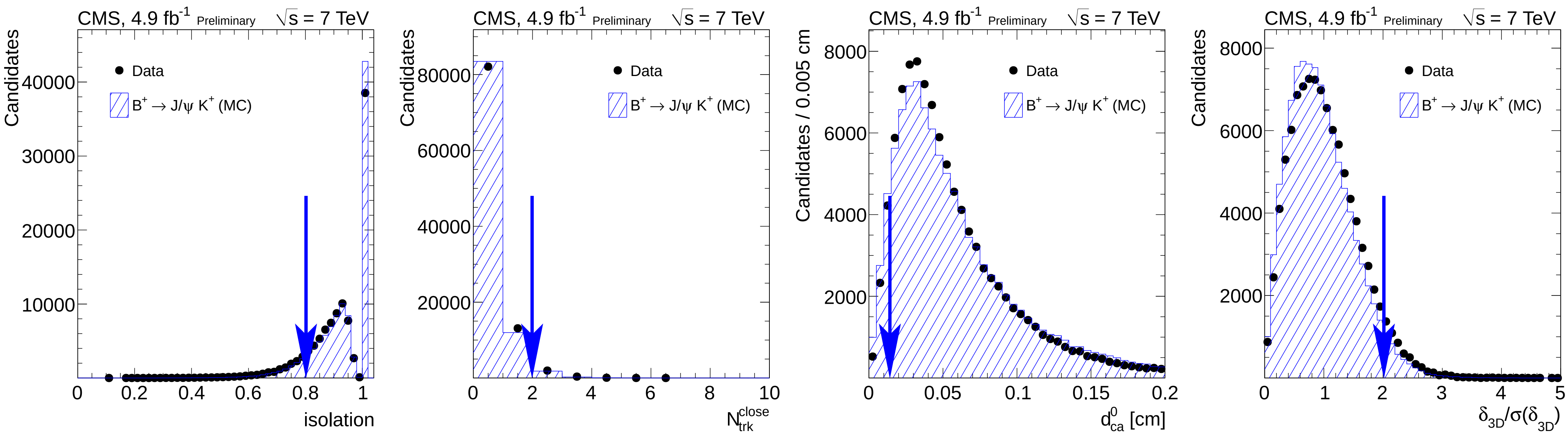








Isolation

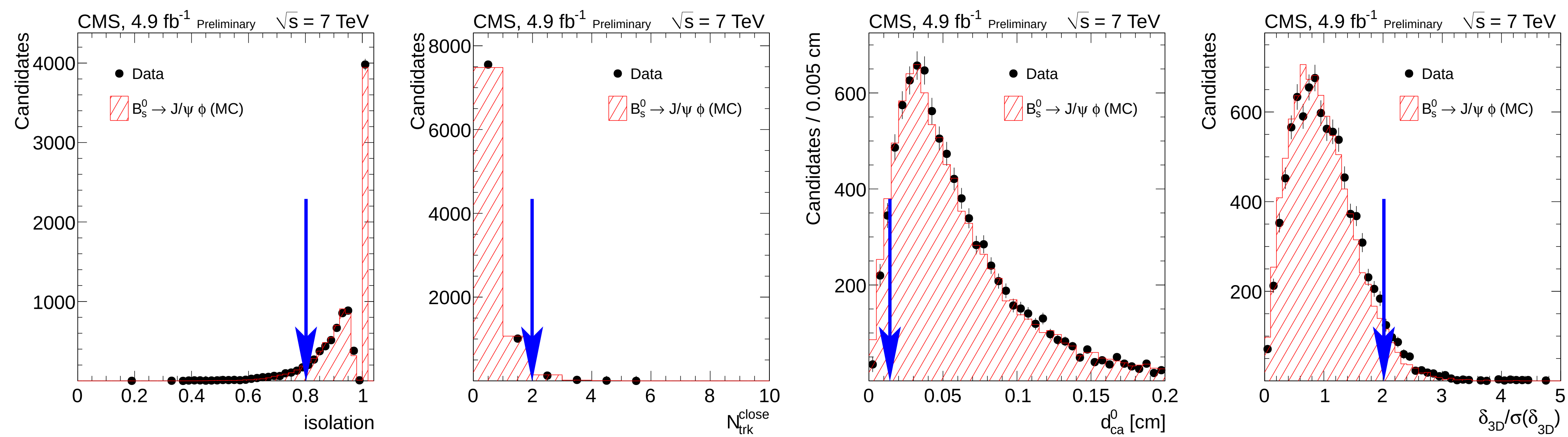


isolation

close tracks

distance of closest track

3d impact param. sign.



Selection efficiency (uncertainty)

- Determine selection efficiency in
  - data
  - MC simulation

  with respect to 'all other selection requirements', e.g. for $B^\pm \to J/\psi\, K^\pm$:

Variable	Selection	MC	Data	Difference
muon $p_\perp$	$p_\perp > 4.0$ GeV	0.927 ± 0.001	0.926 ± 0.001	-0.002 ± 0.001
pointing angle	$\alpha_{3D} < 0.0500\,\mathrm{rad}$	0.994 ± 0.000	0.995 ± 0.000	$+0.000 \pm 0.000$
vertex fit	$\chi^2/dof < 2.0$	0.936 ± 0.001	0.928 ± 0.001	-0.009 ± 0.001
impact parameter	$\delta_{3D} < 0.008$	0.972 ± 0.001	0.972 ± 0.001	$+0.001 \pm 0.001$
impact param. sign.	$\delta_{3D}/\sigma(\delta_{3D}) < 2.000$	0.959 ± 0.001	0.944 ± 0.001	-0.015 ± 0.001
flight length sig.	$\ell_{3d}/\sigma(\ell_{3d}) > 15.0$	0.923 ± 0.001	0.926 ± 0.001	$+0.004 \pm 0.001$
isolation	$I > 0.80$	0.893 ± 0.001	0.871 ± 0.001	-0.025 ± 0.002
close tracks	$N_{trk} < 2$	0.978 ± 0.000	0.975 ± 0.000	-0.003 ± 0.001
d^0_{ca}	$d^0_{ca} > 0.015$ cm	0.917 ± 0.001	0.929 ± 0.001	$+0.013 \pm 0.001$

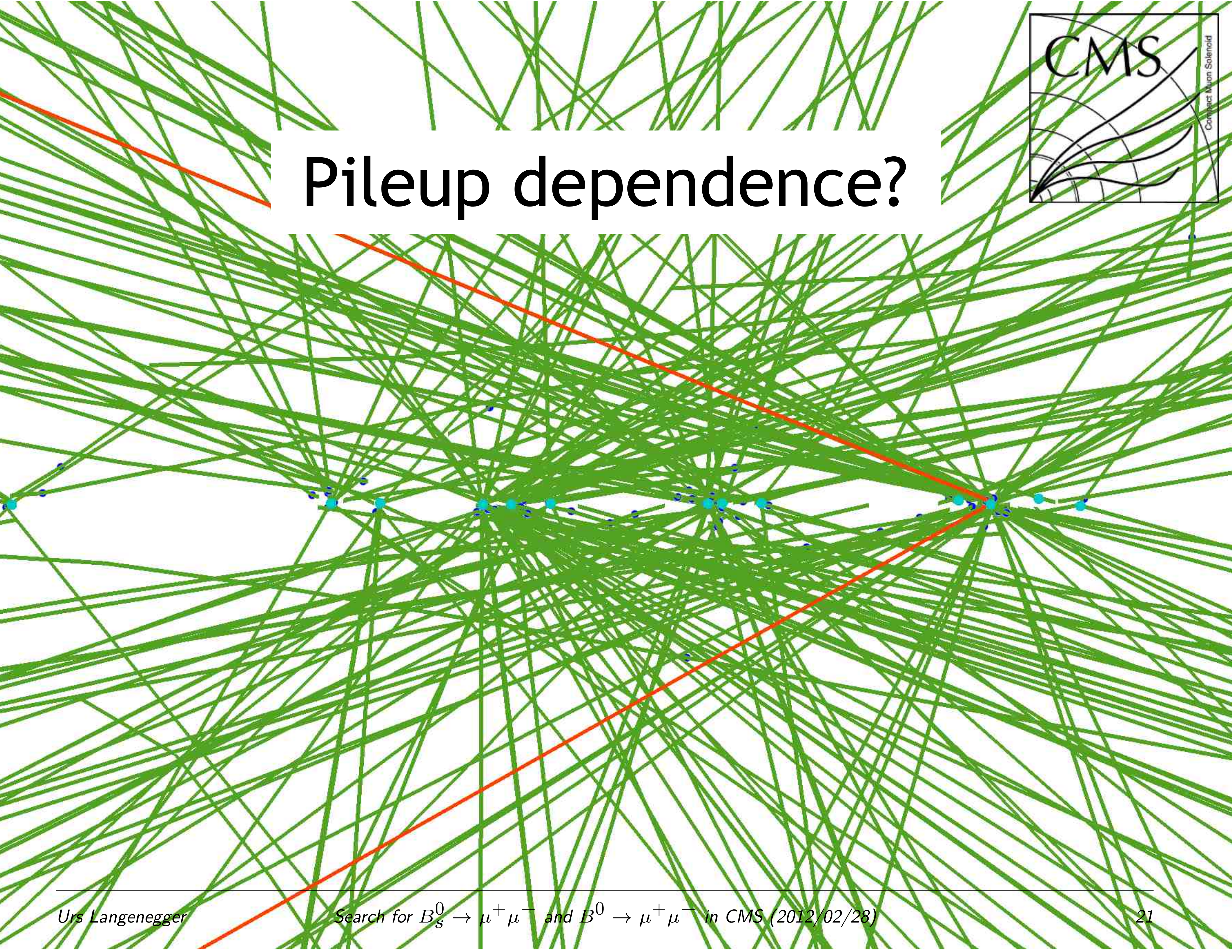
$\Rightarrow$ Systematic uncertainty from (quadr.) sum of relative differences

- $\to$ $B^\pm \to J/\psi\, K^\pm$: 4%

  (largest single deviation: 2.5% from isolation)
- $\to$ $B^0_s \to J/\psi\, \phi$: 3%

  (largest single deviation: 1.5% from B vertex χ^2/dof)
- $\to$ idem for signal selection efficiency uncertainty

Pileup dependence?



Pileup independence

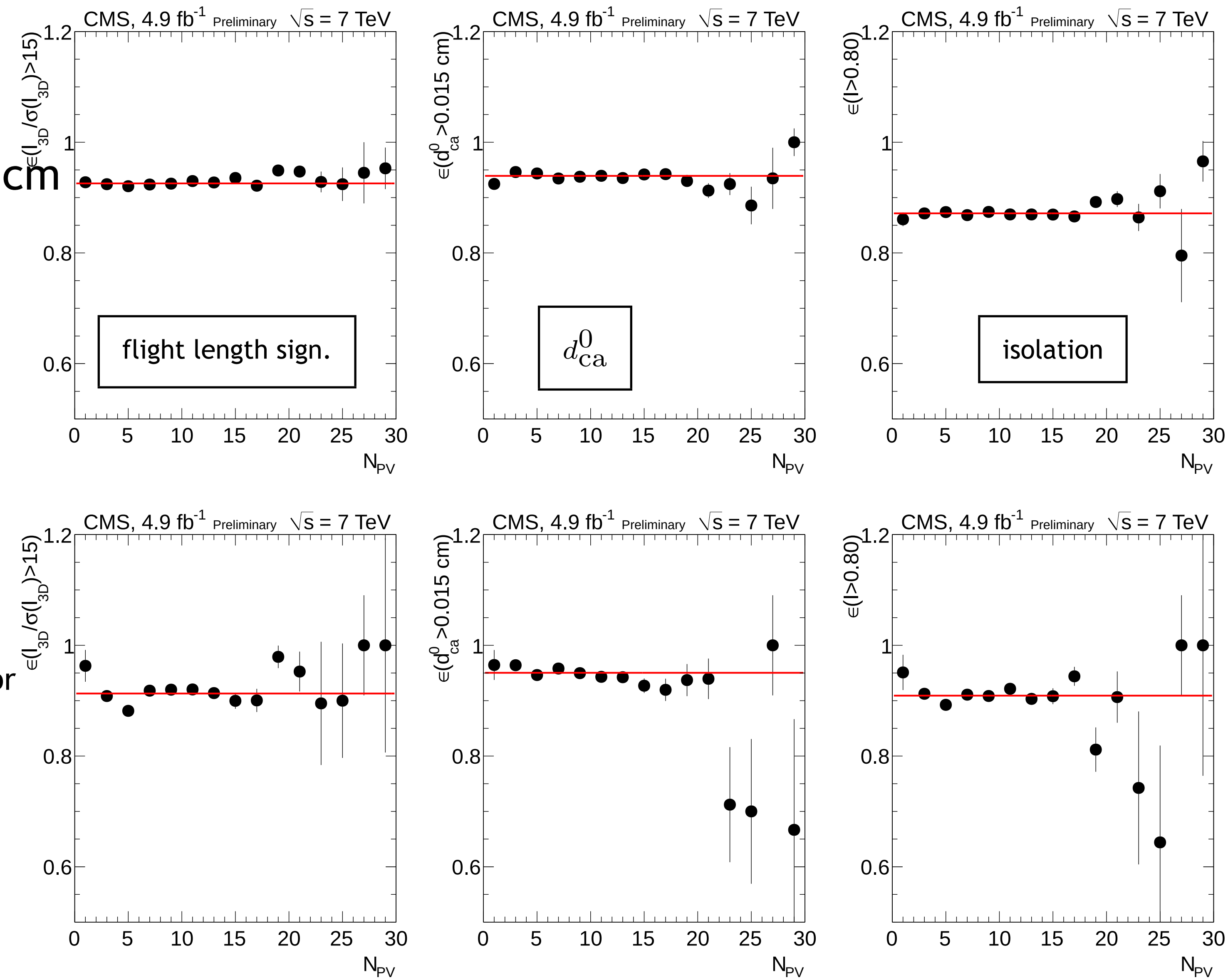
- Determine selection efficiency vs N_{PV} in data
- 2011 dataset:
  - ▷ $\langle N_{\mathrm{PV}} \rangle \approx 8$
  - ▷ RMS$(z) \approx 5.6\,$cm

$B^\pm \to J/\psi\, K^\pm$

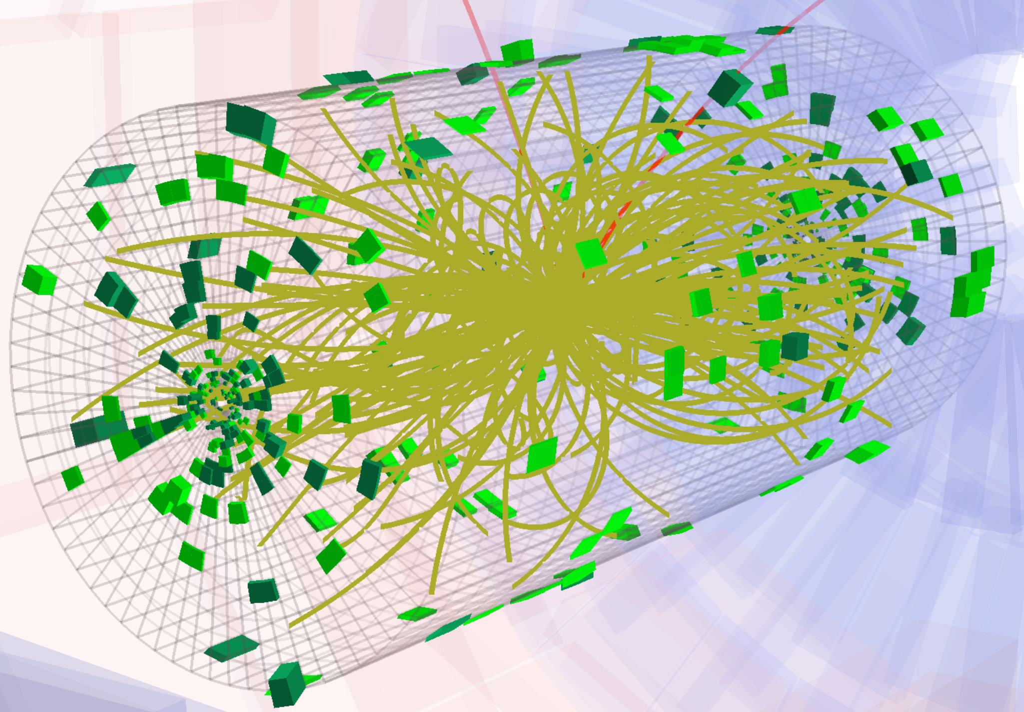
$B^0_s \to J/\psi\, \phi$

- MC: also checked ε for
  - ▷ $N_{\mathrm{PV}} < 6$
  - ▷ $N_{\mathrm{PV}} > 10$

$\Rightarrow$ no pileup dependence



Search Analysis



Selection for search analysis

- Random grid optimization
  - 14 variables included in 1.4×10^6 runs

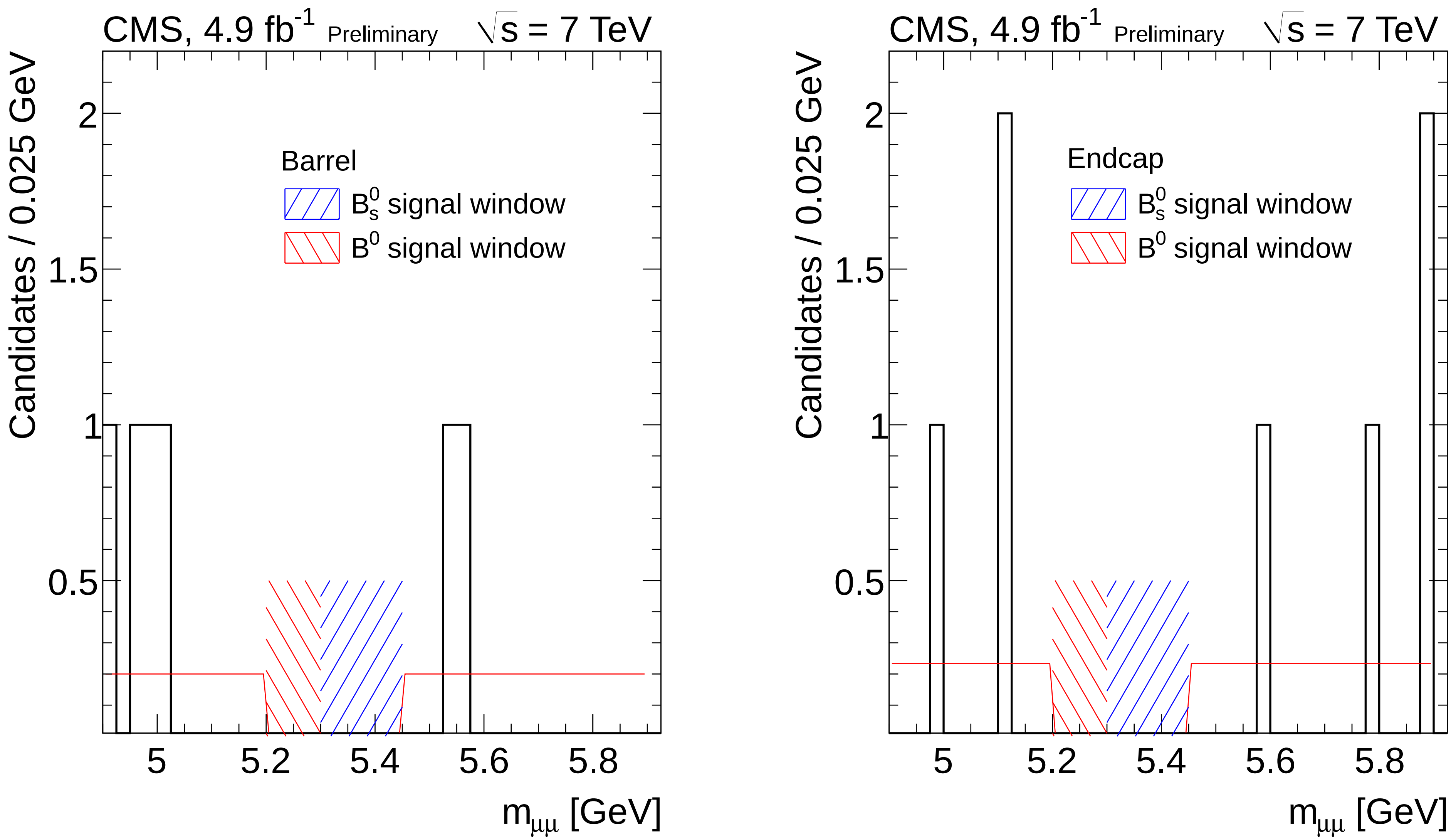
Variable	Barrel	Endcap	units	comparison to old analysis
$p_{\perp\mu,1} >$	4.5	4.5	GeV	same
$p_{\perp\mu,2} >$	4.0	4.2	GeV	tighter in endcap
$p_{\perp B} >$	6.5	8.5	GeV	tighter in endcap
$\ell_{3d} <$	1.5	1.5	cm	tighter
$\alpha <$	0.050	0.030	rad	looser
$\chi^2/dof <$	2.2	1.8		looser
$\ell_{3d}/\sigma(\ell_{3d}) >$	13.0	15.0		looser
$I >$	0.80	0.80		redefined
$d^0_{ca} >$	0.015	0.015	cm	redefined
$\delta_{3D} <$	0.008	0.008	cm	new
$\delta_{3D}/\sigma(\delta_{3D}) <$	2.000	2.000		new
$N_{trk} <$	2	2	tracks	new

- Total efficiency $\times$ acceptance

Efficiency	Barrel	Endcap
$B^0_s \to \mu^+\mu^-$	0.0029 ± 0.0002	0.0016 ± 0.0002
$B^\pm \to J/\psi K^\pm$	0.00110 ± 0.00009	0.00032 ± 0.00004

Dimuon mass distribution (blinded)

- Low background (sidebands shown only)

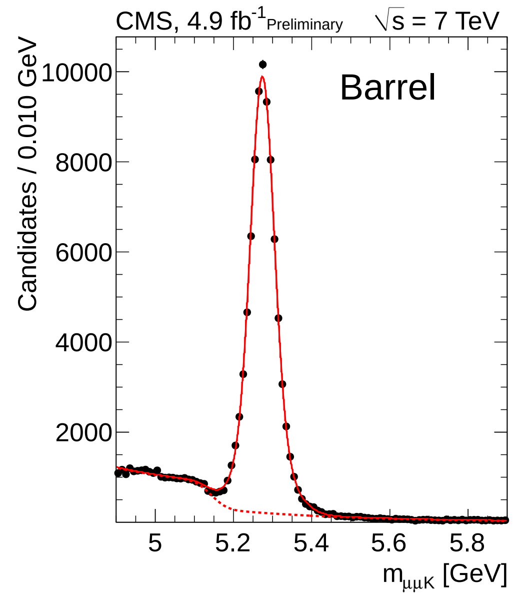


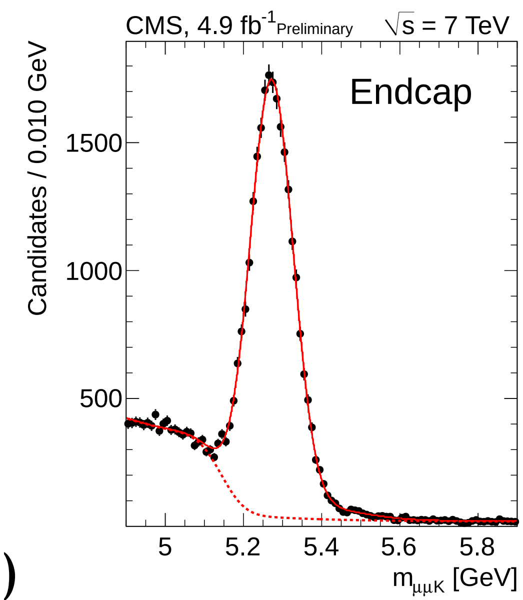
Measurement of $B^{\pm} \to J/\psi\, K^{\pm}$

- Use identical selection as for dimuon, plus
  - $3.0 < m(\mu\mu) < 3.2$ GeV
  - $p_{\perp}(\mu\mu) > 7$ GeV, $p_{\perp}(K) > 0.5$ GeV
  - all tracks used in vertexing
- Fit function
  - signal: double Gaussian
  - background: exponential + error function
    partially reconstructed B decays
    $B^0 \to J/\psi\, K^* \to \mu^+\mu^- K^-(\pi^+)$

	Barrel	Endcap
Acceptance	0.162 ± 0.006	0.111 ± 0.006
$\varepsilon_{\rm tot}$	0.00110 ± 0.00009	0.00032 ± 0.00004
$N_{\rm obs}$	82712 ± 4146	23809 ± 1203

- Systematic error on yield: 5%
  - variation of
    background pdf
    vary signal pdf
  - mass-constrain dimuons to J/ψ (better resolution)





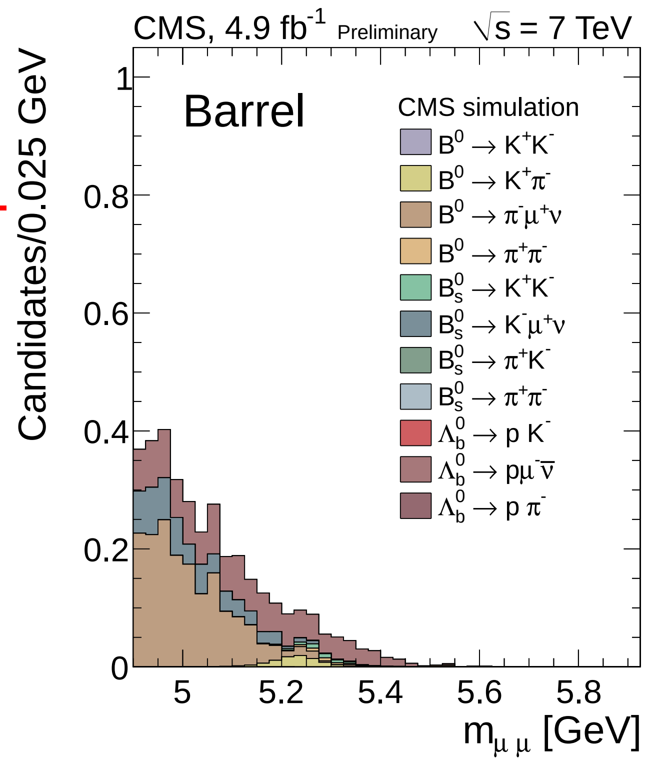
Rare backgrounds

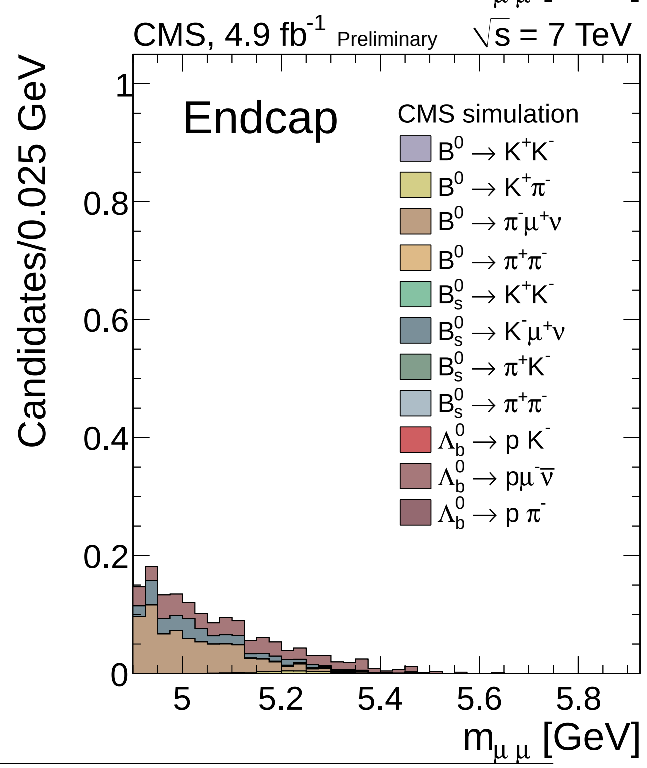
- Rare backgrounds
  - ▷ CKM-suppressed semileptonic decays
    e.g. $B_s^0 \to K^-\mu^+\nu$, one fake muon
    large $\mathcal{B}$, but mostly at low masses
  - ▷ 'peaking' hadronic decays
    e.g. $B_s^0 \to K^-K^+$, two fake muons
  - ▷ Normalization to B^+ yield in data

$$N(X) = \frac{\mathcal{B}(Y \to X)}{\mathcal{B}(B^\pm \to J/\psi\, K^\pm)} \frac{f_Y}{f_u} \frac{\varepsilon_{\rm tot}(X)}{\varepsilon_{\rm tot}(B^+)} N_{\rm obs}(B^+)$$

weighting with misid rate f (or ε_μ) and $\varepsilon_{\rm trig}$

- Note
  - ▷ B^0 more affected than B_s^0
  - ▷ endcap more diluted than barrel
    lower efficiency
- Systematic error varies
  - ▷ branching fraction uncertainties
  - ▷ $f_s/f_u = 0.267 \pm 0.021$ [LHCb, arxiv:1111.2357]





Systematic uncertainties

- Acceptance:
  - ▷ mixture of production processes
    gluon fusion
    flavor excitation
    gluon splitting
  - $\rightarrow$ half of acceptance variation
  - ▷ Studied variables sensitive to mixture
    muon vs B candidate:
    $\Delta R(B, \mu)$
    $p_{\perp}(\mu)$

Barrel	Endcap
3.5%	5%

- Selection efficiency
  - ▷ from data/MC comparisons
  - $\rightarrow$ quadratic sum for all selection criteria

Category	Barrel	Endcap
$\varepsilon_{\rm tot}$ (signal)	3%	3%
$\varepsilon_{\rm tot}$ (normalization)	4%	4%
kaon tracking	4%	4%

- Muon trigger and efficiency
  - $\rightarrow$ full variation, for thresholds $4 < p_{\perp} < 8$ GeV
  - $\rightarrow$ efficiency difference between data and MC

Category	Barrel	Endcap
μ trigger	3%	6%
μ ID	4%	8%

Cross checks

- Determination of $\mathcal{B}(B_s^0 \to J/\psi\,\phi)/\mathcal{B}(B^\pm \to J/\psi\,K^\pm)$
  - ▷ barrel vs. endcap
  - ▷ B^+ fitting
  - ▷ consistent definitions
    - acceptance
    - efficiency

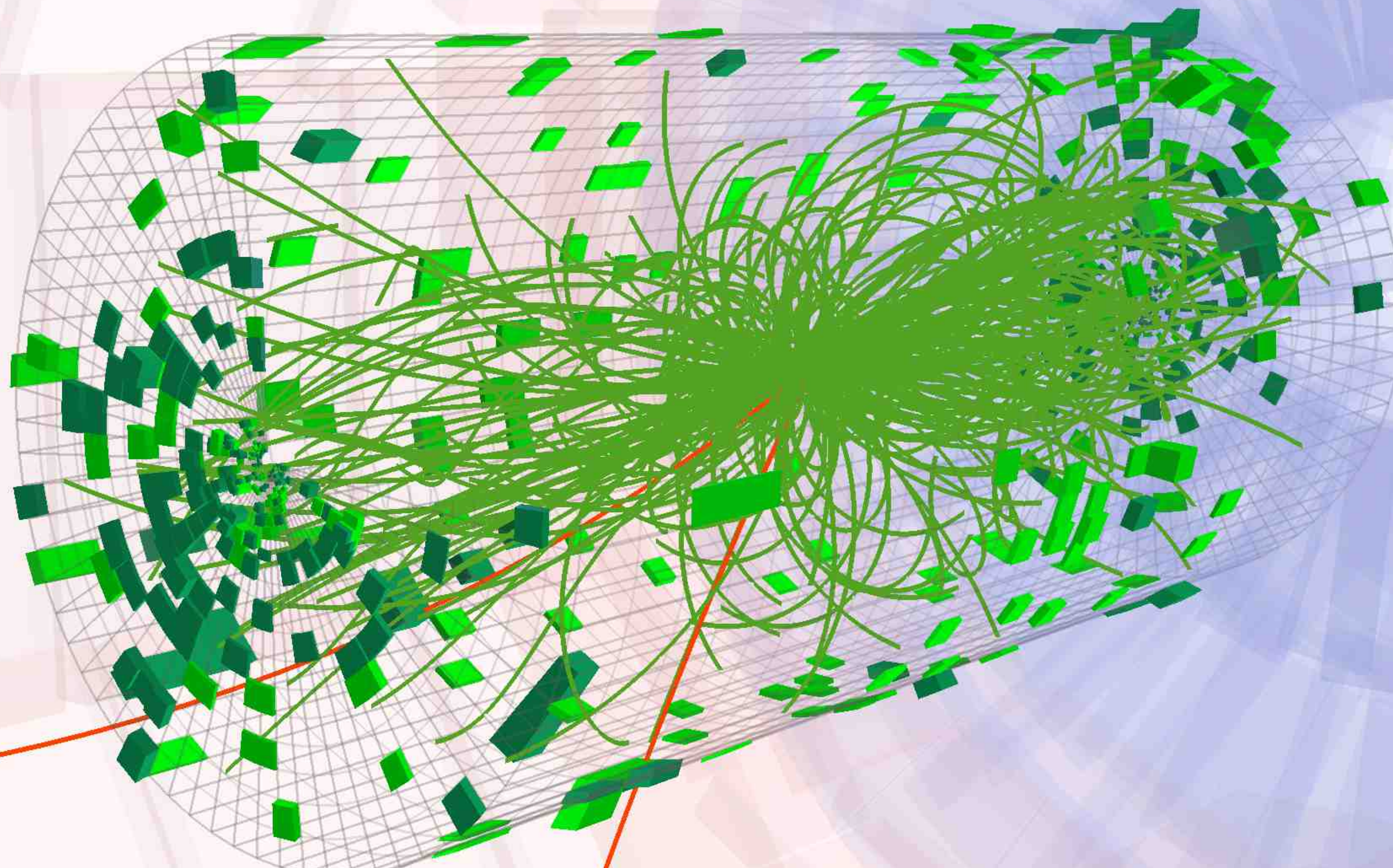
    (different number of daughters)

- Inverted isolation sample ($I < 0.7$, not blinded)

  $\to$ comparison of prediction vs. observation
  - ▷ validation of rare backgrounds
  - ▷ background interpolation

- Stability vs. time (HLT changes)
  - ▷ yields (dimuons, normalization and control sample)
  - ▷ yield ratios

Results



Upper limit calculation

- Methodology
  - ▷ $\mathrm{CL_s}$
  - ▷ Feldman-Cousins
  - ▷ statistical model:

$$
\begin{aligned}
N_s^B &\sim \mathrm{Pois}(\tau_s^B \nu_b^B + \nu_{s,\mathrm{rare}}^B + P_{ss}^B \mu_s \nu_s^B + P_{sd}^B \mu_d \nu_d^B) \\
N_d^B &\sim \mathrm{Pois}(\tau_d^B \nu_b^B + \nu_{d,\mathrm{rare}}^B + P_{ds}^B \mu_s \nu_s^B + P_{dd}^B \mu_d \nu_d^B)
\end{aligned}
$$

with $(i = s, d)$

τ_i^B	Ratio of $(B_i^0 \to \mu\mu)$-signal window size to size of background window
$\nu_{i,\mathrm{rare}}^B$	Expected number of rare background in $(B_i^0 \to \mu\mu)$-signal window.
ν_i^B	Expected number of reconstructed $(B_i^0 \to \mu\mu)$ decays in barrel region assuming the SM
P_{ij}^B	Probability for a reconstructed $B_j^0 \to \mu\mu$ decay to be in $(B_i^0 \to \mu\mu)$-signal window.
μ_i	Signal strength of $B_i^0 \to \mu\mu$, that is the ratio of true branching ratio to SM branching ratio.

- Systematic error on cross feed P_{ij}^B
  - ▷ mass scale and resolution
  - ▷ measure $J/\psi \to \mu^+\mu^-$, $B_s^0 \to \mu^+\mu^-$, $\Upsilon(1S) \to \mu^+\mu^-$
  - $\to$ compare MC resolution (and position) with prediction (interpolation)

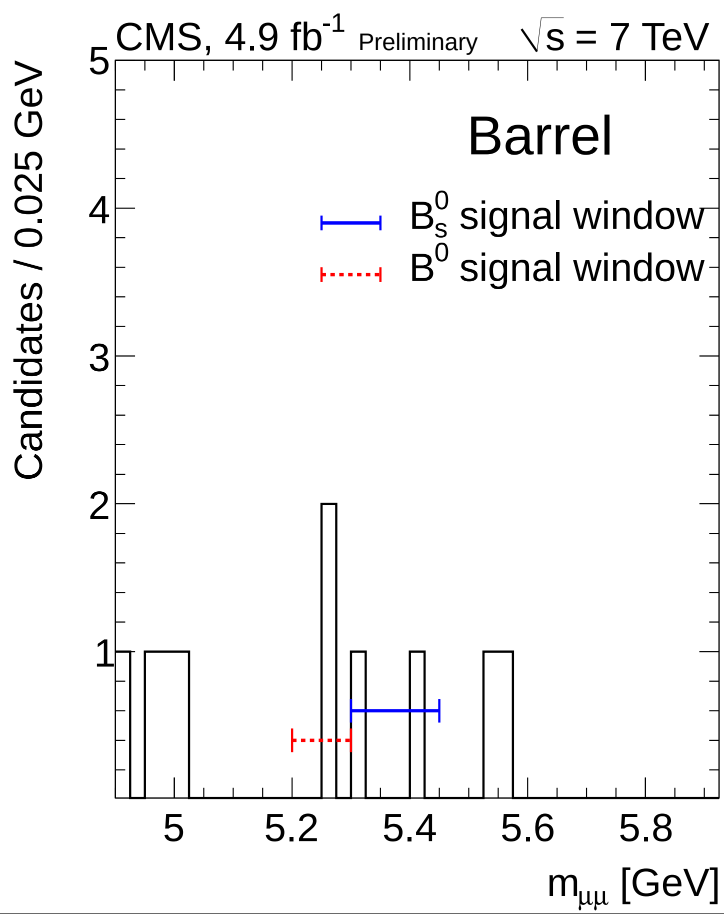
Summary of systematic errors

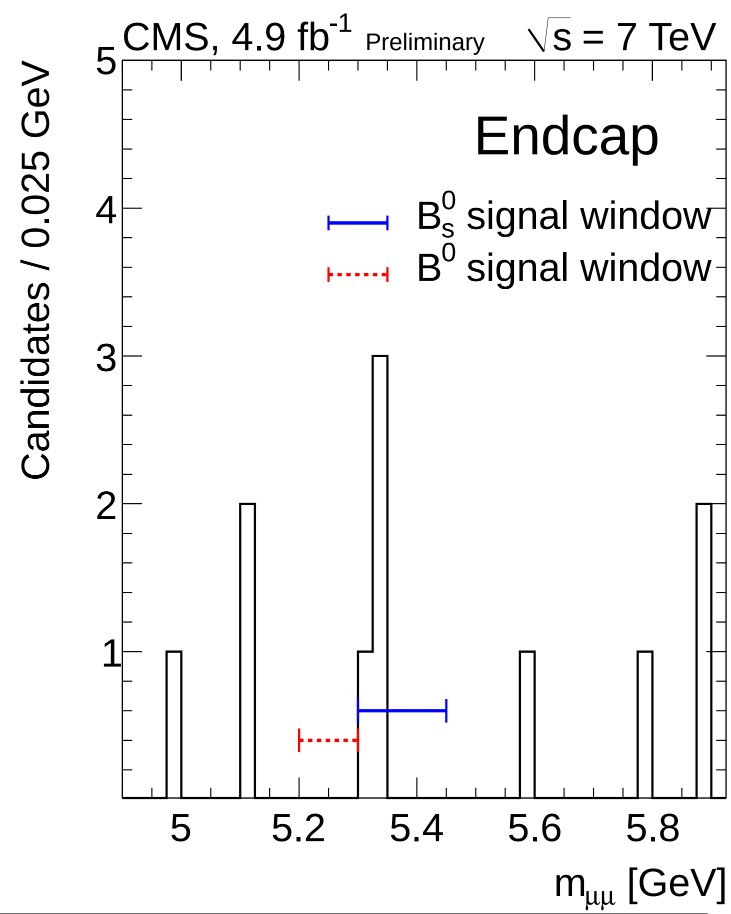
- Systematic uncertainties propagated into upper limit calculation
  all errors below in %

Category	Uncertainty	Barrel	Endcap
f_s/f_u	production ratio of u and s quarks	8.0	8.0
acceptance	production processes	3.5	5.0
P_{ij}^{B}	mass scale and resolution	3.0	3.0
efficiency (signal)	discrepancies data/MC simulation	3.0	3.0
efficiency (normalization)	discrepancies data/MC simulation	4.0	4.0
efficiency (normalization)	kaon track efficiency	4.0	4.0
efficiency	trigger	3.0	6.0
efficiency	muon identification	4.0	8.0
normalization	fit pdf	5.0	5.0
background	shape of combinatorial background	4.0	4.0
background	rare decays	20.0	20.0

Expectations and observation

Variable	$B^0 \to \mu^+\mu^-$ Barrel	$B_s^0 \to \mu^+\mu^-$ Barrel	$B^0 \to \mu^+\mu^-$ Endcap	$B_s^0 \to \mu^+\mu^-$ Endcap
Signal	0.24 ± 0.02	2.70 ± 0.41	0.10 ± 0.01	1.23 ± 0.18
Combinatorial bg	0.40 ± 0.34	0.59 ± 0.50	0.76 ± 0.35	1.14 ± 0.53
Peaking bg	0.33 ± 0.07	0.18 ± 0.06	0.15 ± 0.03	0.08 ± 0.02
Sum	0.97 ± 0.35	3.47 ± 0.65	1.01 ± 0.35	2.45 ± 0.56
Observed	2	2	0	4





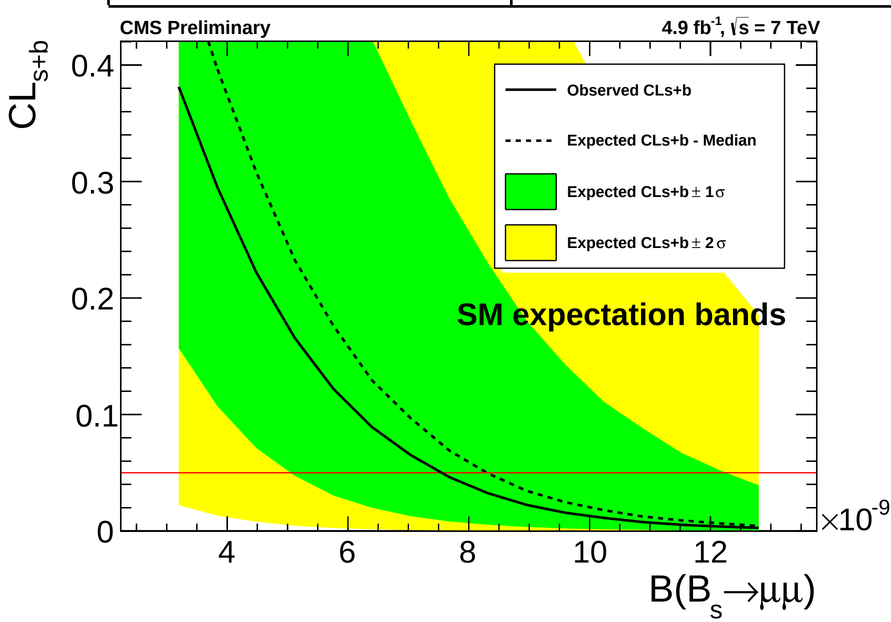
Results: upper limits

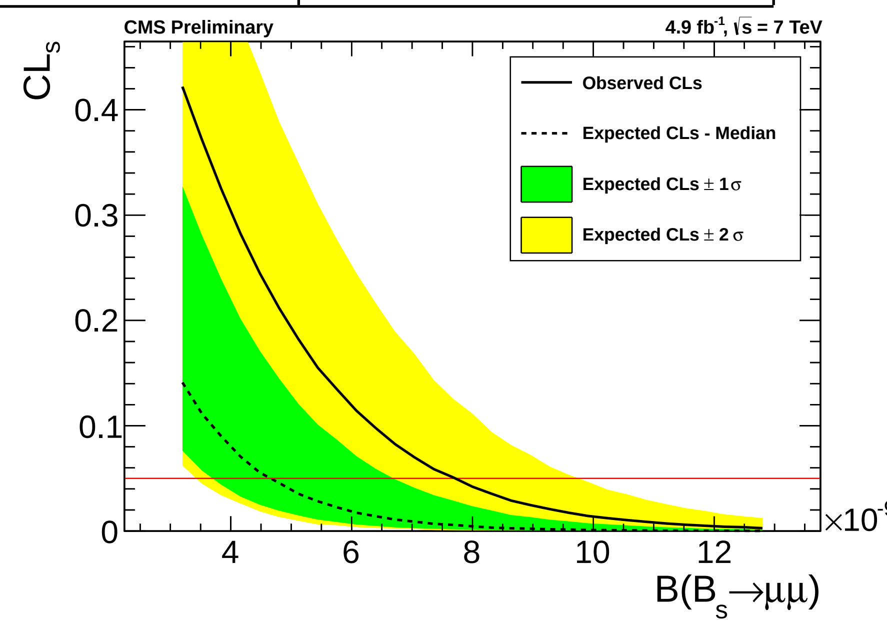
- Upper limit on $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$

upper limit (95%CL)	observed	(median) expected
$\mathcal{B}(B_s^0 \to \mu^+\mu^-)$	7.7×10^{-9}	8.4×10^{-9}
$\mathcal{B}(B^0 \to \mu^+\mu^-)$	1.8×10^{-9}	1.6×10^{-9}

- p-values (for background-only hypotheses)

p-values	background only	SM cross feed	floating cross feed
$\mathcal{B}(B_s^0 \to \mu^+\mu^-)$	0.06 (1.5σ)	0.07 (1.5σ)	0.11 (1.2σ)
$\mathcal{B}(B^0 \to \mu^+\mu^-)$	0.11 (1.2σ)	0.29 (0.6σ)	0.24 (0.7σ)

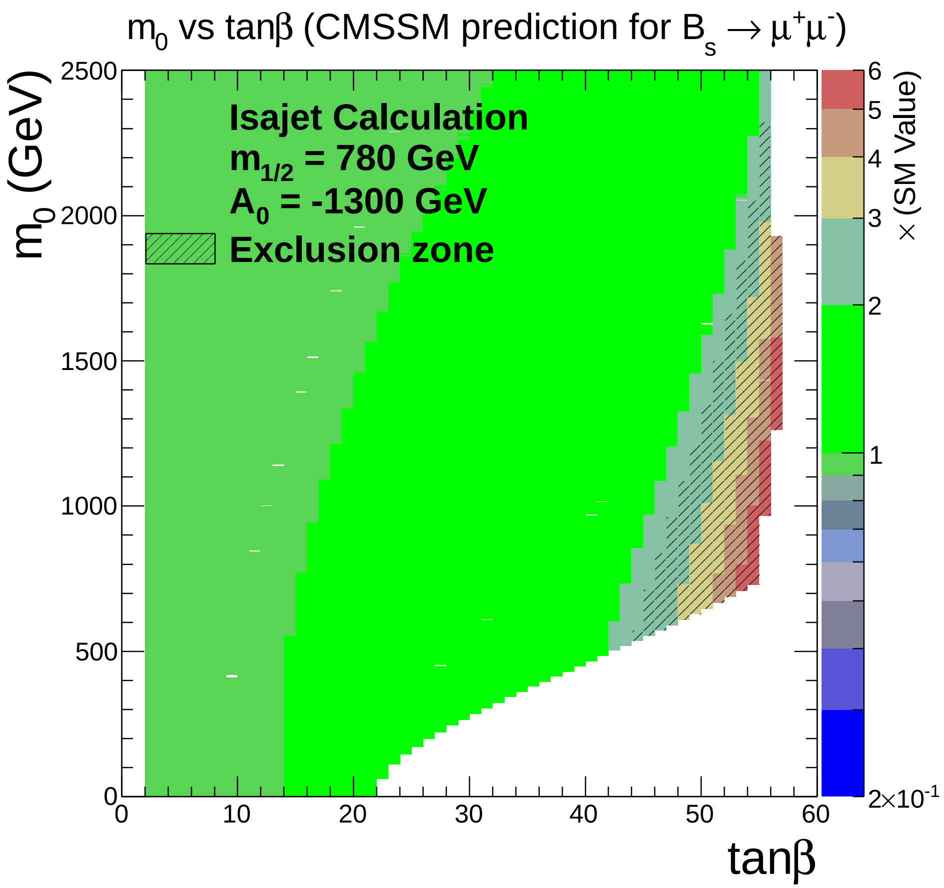


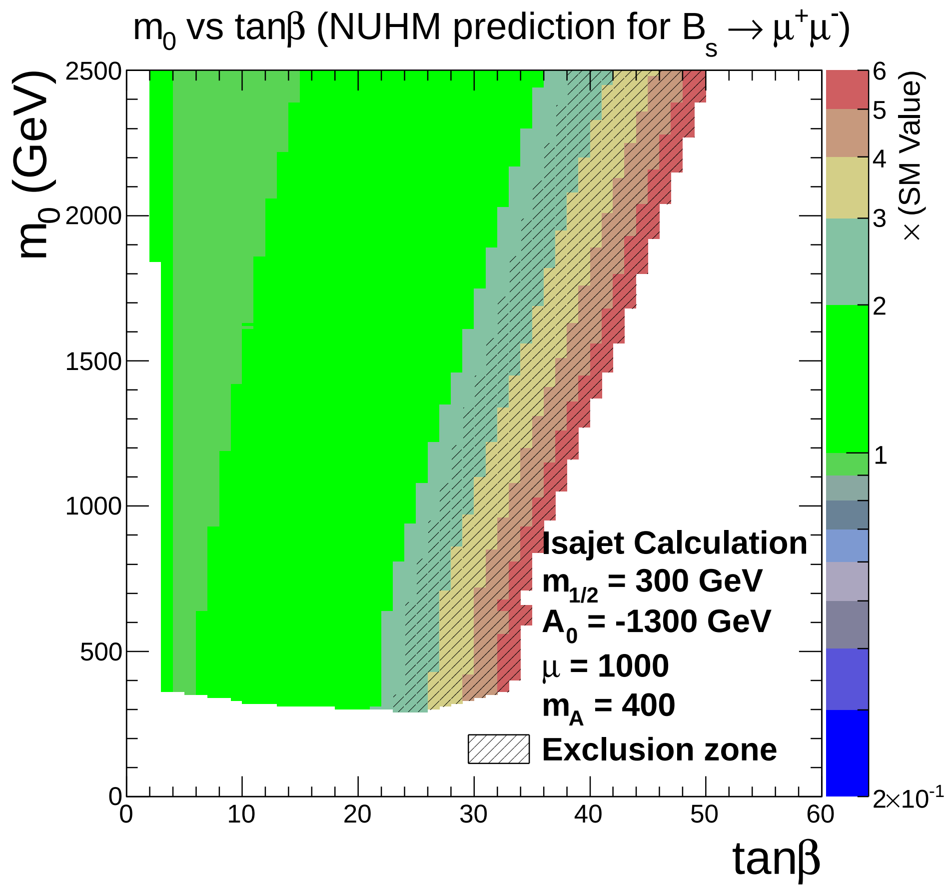


Interpretation examples

SuperIso V3.1
(CPC, 180, 1579)

- Empty region due to previous upper limit
  and other published data





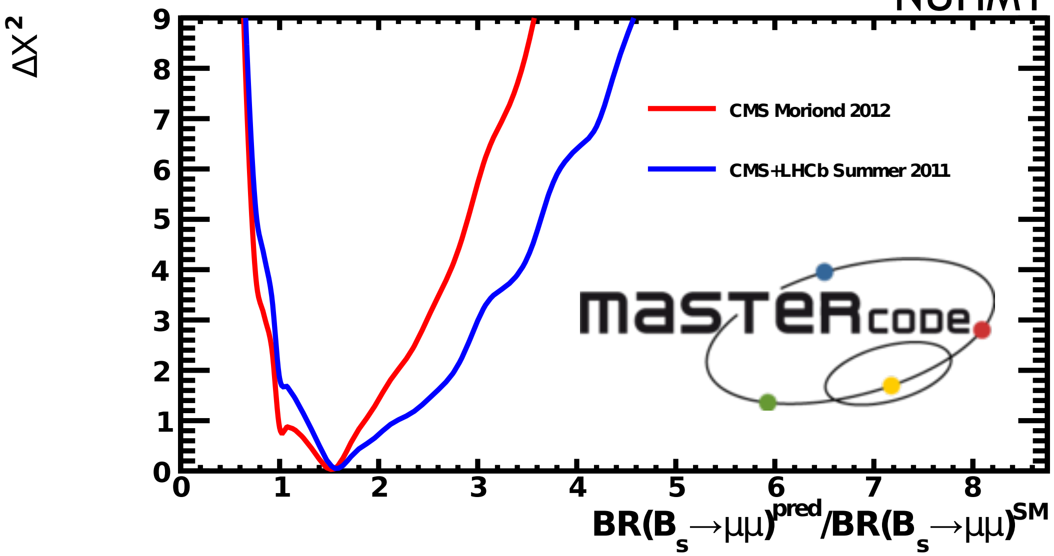
$\Rightarrow$ strongest impact at large $\tan\beta$

Interpretation examples (II)

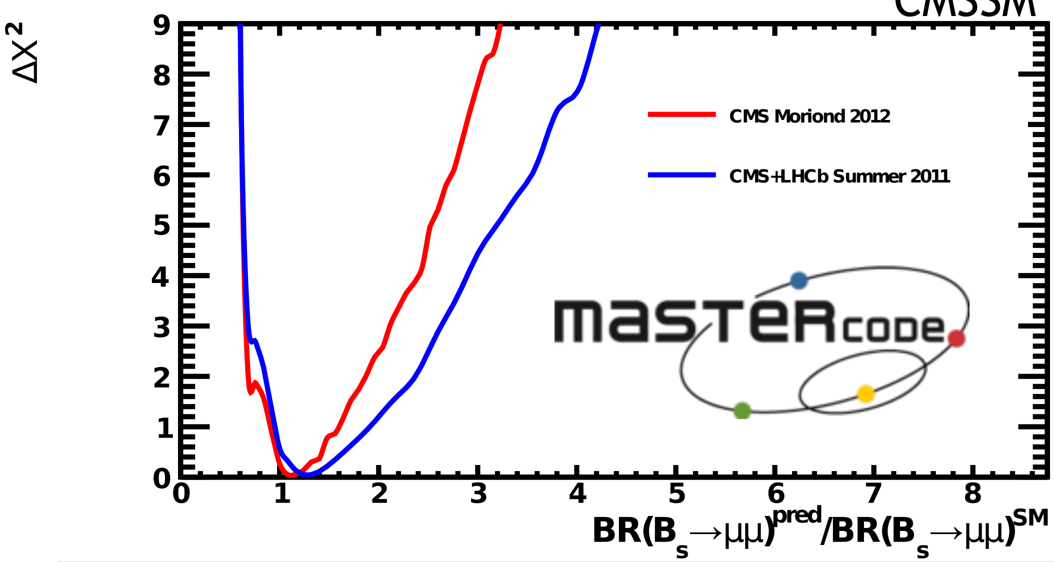
MasterCode
(arXiv:1112.3564)

- χ^2 difference

NUHM1

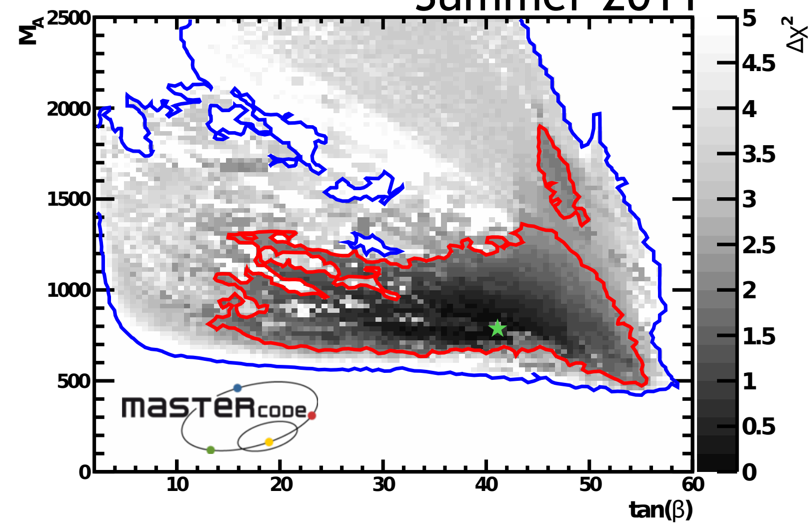


CMSSM

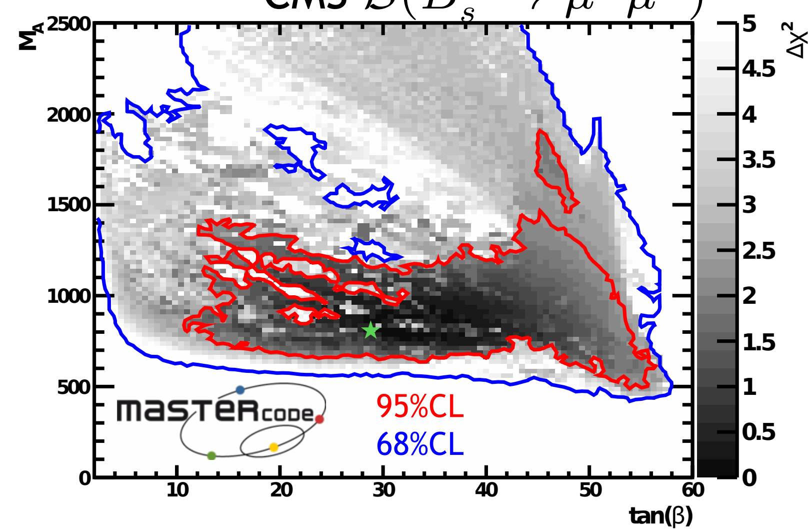


- 'best' fit for CMSSM

Summer 2011



CMS $\mathcal{B}(B_s^0 \to \mu^+\mu^-)$



Conclusions

- Search for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ in 2011 dataset

upper limit (95%CL)	observed	expected
$\mathcal{B}(B_s^0 \to \mu^+\mu^-)$	7.7×10^{-9}	8.4×10^{-9}
$\mathcal{B}(B^0 \to \mu^+\mu^-)$	1.8×10^{-9}	1.6×10^{-9}

- Significant improvement
  - ▷ EPS 2011: $\mathcal{B}(B_s^0 \to \mu^+\mu^-) < 1.9 \times 10^{-8}$
  - ▷ more/changed variables, e.g., better B vertex isolation
  - ▷ improved sensitivity
  - ▷ higher signal/background ratio
- Upper limit now approaching factor 2 of SM expectation
- Looking forward to 2012 dataset. Well prepared for
  - high instantaneous lumi (trigger)
  - high pileup (tracking and vertexing)