$B_s \rightarrow \mu^+\mu^-$
AT THE LHC

Diego Martínez Santos (Universidade de Santiago de Compostela)
on behalf of ATLAS, CMS and LHCb collaborations

PROGRAMA NACIONAL DE BECAS FPU

OVERVIEW

- Motivation for the study of $B_s \rightarrow \mu\mu$ as an indirect probe of NP

- Analyses at the LHC: ATLAS/CMS/LHCb

    - How to find such a rare decay and disentangle from background
    - Normalization and Calibration to get a correct BR

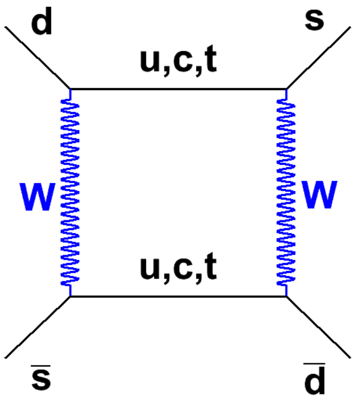
- Conclusions

INDIRECT APPROACH

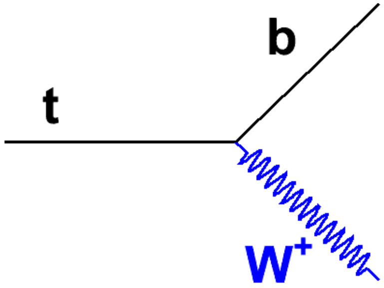
• $B_s \rightarrow \mu\mu$ can access NP through new virtual particles entering in the loop → indirect search of NP

• Indirect approach can access higher energy scales and see NP effects earlier:

- Some examples:
  - 3^{rd} quark family inferred by Kobayashi and Maskawa (1973) to explain CP V in K mixing (1964). Directly observed in 1977 (b) and 1995 (t)
  - Neutral Currents discovered in 1973, Z^0 directly observed in 1983



~30 years till the direct observation…



INDIRECT APPROACH

- $B_s \rightarrow \mu\mu$ can access NP through new virtual particles entering in the loop → indirect search of NP

- Indirect approach can access higher energy scales and see NP effects earlier:

  - A very early example of how indirect measurements give information about higher scales ☺:
    - **Ancient Greece**: Earth must be some round object, Eratosthenes measurement of Earth's radius in **c. III BC** (using differences in shadows at different cities)
    - Roundness of Earth not directly observed until **~1946-61**

Eratosthenes

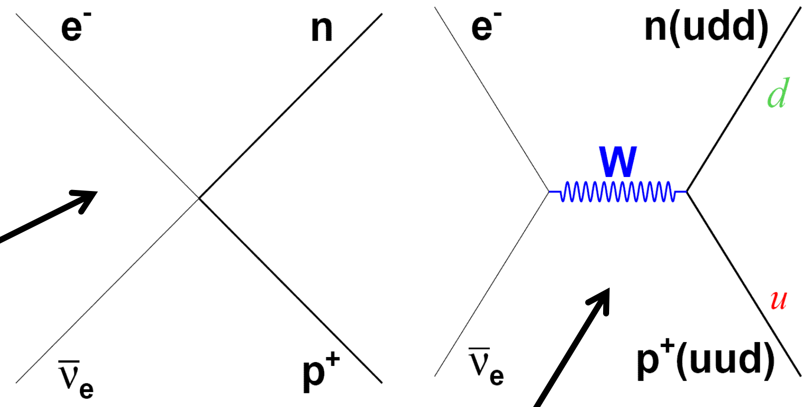
~2.3 K years till the direct observation…

WILSON COEFFICIENTS

Hadronic weak decays are often studied in terms of effective hamiltonians of local operators Q_i:

$$H_{eff} \propto \sum_i C_i \hat{Q}_i$$

effective local theory



underlying "fundamental" theory (SM)

Degrees of freedom of exchanged particles are integrated out giving rise to the **Wilson coefficients C_i**.

An **example** of similar approach: **Fermi's theory of neutron decay**

BR($B_s \rightarrow \mu\mu$) expressed in eff. th. as:

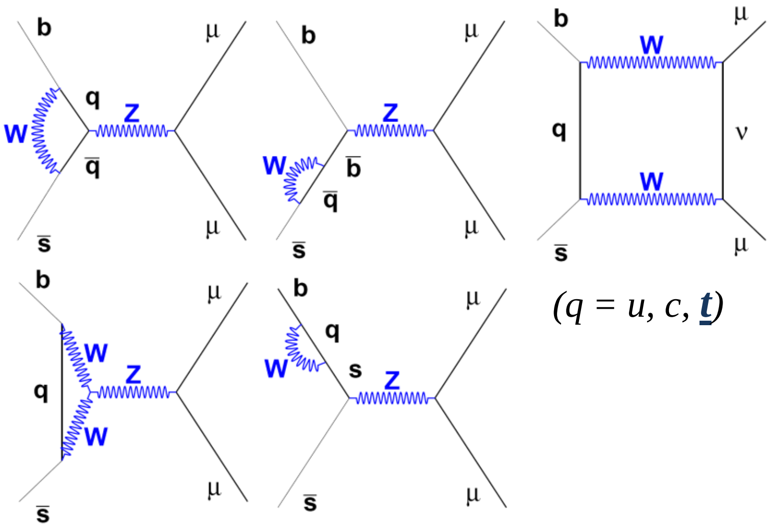
$$BR(B_q \to \mu^+\mu^-) = \frac{G_F^2\alpha^2}{64\pi^3} |V_{tb}^* V_{tq}|^2 \tau_{Bq} M_{Bq}^3 f_{Bq}^2 \sqrt{1-\frac{4m_\mu^2}{M_{Bq}^2}} \times$$
$$\times \left\{ M_{Bq}^2 \left(1-\frac{4m_\mu^2}{M_{Bq}^2}\right) C_S^2 + \left[M_{Bq} C_P + \frac{2m_\mu}{M_{Bq}} C_{10} \right]^2 \right\}$$

$C_{P,S,10}$ (pseudoscalar, scalar and axial) **depend on the underlying model (SM, SUSY…)**

DECAY PHYSICS (SM)

$$BR(B_q \to \mu^+\mu^-) = \frac{G_F^2\alpha^2}{64\pi^3}|V_{tb}^* V_{tq}|^2 \tau_{Bq} M_{Bq}^3 f_{Bq}^2 \sqrt{1-\frac{4m_\mu^2}{M_{Bq}^2}} \times$$

$$\times\left\{M_{Bq}^2\left(1-\frac{4m_\mu^2}{M_{Bq}^2}\right)|C_S|^2 + \left[M_{Bq}C_P + \frac{2m_\mu}{M_{Bq}}C_{10}\right]^2\right\}$$



$(q = u, c, t)$

$C_{S,\,P} \to$ scalar and pseudo scalar are negligible in SM

C_{10} gives the only relevant contribution

This decay is very suppressed in SM:

$BR(B_s \to \mu\mu) = (3.35 \pm 0.32)\mathrm{x}10^{-9}$ $BR(B_d \to \mu\mu) = (1.03 \pm 0.09)\mathrm{x}10^{-10}$

M.Blanke et al., JHEP 10 003,2006

Current experimental upper limit (CDF, $3.7\mathrm{fb}^{-1}$) still one order of magnitude to reach such values. @ 90% CL:

$BR(B_s \to \mu\mu) < 3.6\mathrm{x}10^{-8}$ $BR(B_d \to \mu\mu) < 6.0\mathrm{x}10^{-9}$

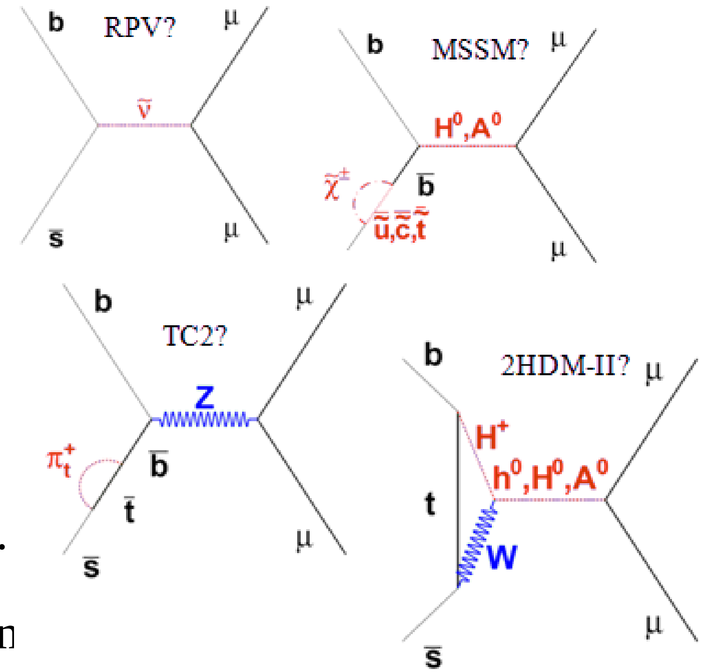
CDF collab., CDF Public Note 9892

NEW PHYSICS EFFECTS

NP can contribute to this decay rate (specially SUSY at high tanβ ($tan\beta = v_u/v_d$)):

- More than one Higgs → contributions to $\mathbf{C_{S,P}}$
  - 2HDM-II : BR proportional to $\tan^4\beta$
  - SUSY (MSSM): above + extra $\tan^6\beta$ +…

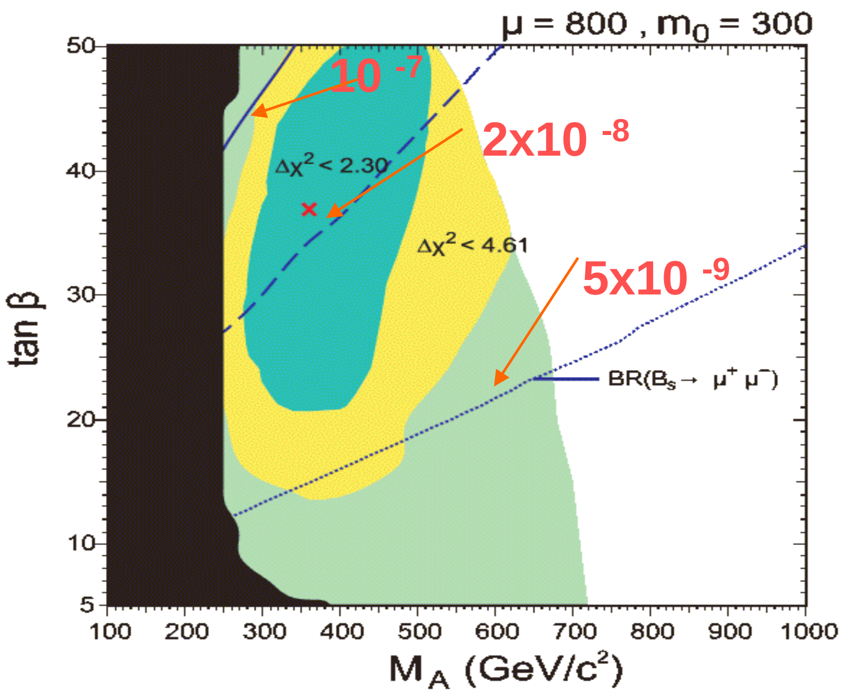
- RPV SUSY: tree level diagrams
- Technicolor (TC2), Little Higgs (LHT) … modify $\mathbf{C_{10}}$.

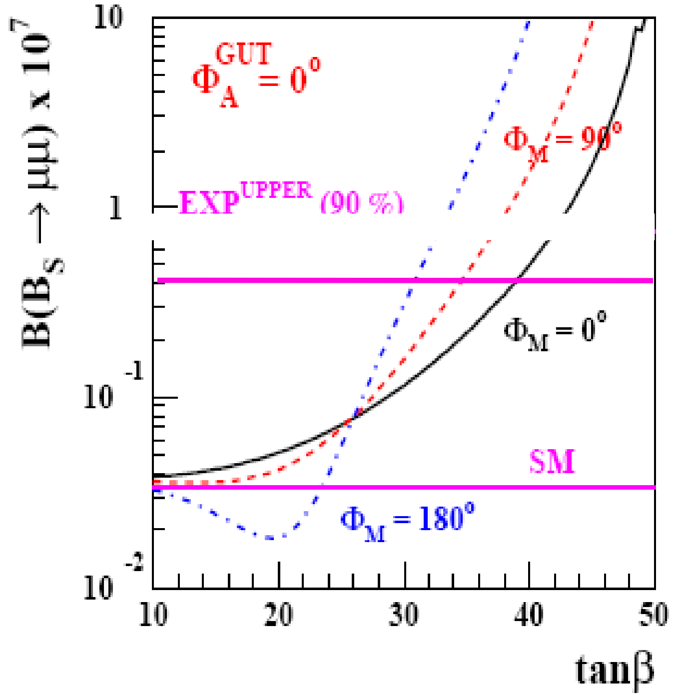


NP can modify the BR from < SM up to current experin

→ ***Whatever the actual value is, it will have an impact on NP searches***

NP (II)





J.Ellis et. al. Phys.Rev.D76:115011, 2007[arXiv:0708.2079v4 [hep-ph]] (2008)

NUHM: best χ^2 of the fit → BR ~$2x10^{-8}$

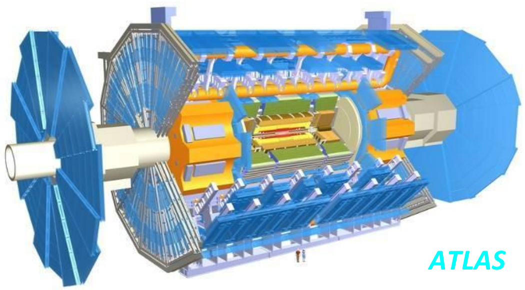
MCPVMFV: Enhancements up to current u.l, but also < SM depending on the phases

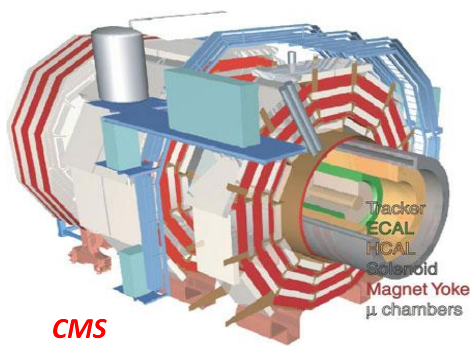
	CMSSM	mGMSB	mAMSB
BR($B_s \rightarrow \mu^+\mu^-$)	~$4.5x10^{-8}$	~$3.2x10^{-8}$	~$0.4x10^{-8}$

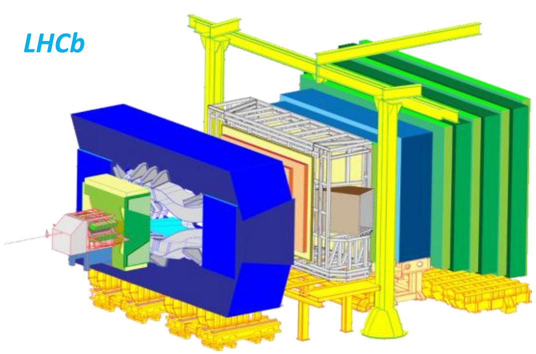
S. Heinemeyer et al., arXiv:0805.2359v2 [hep-ph]

LHC Sensitivity to $B_s \rightarrow \mu^+\mu^-$

LHC EXPERIMENTS

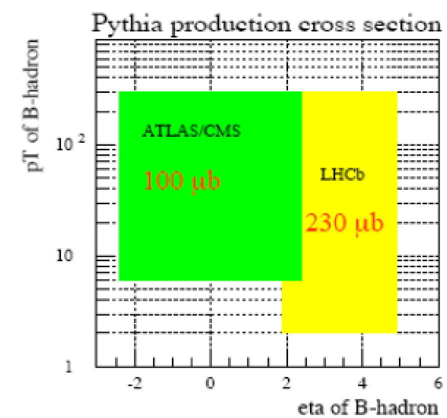






ATLAS & CMS:

- General purpose experiments
- Central detectors $|\eta| < 2.5$
- High pt physics at $L = 10^{33}$ - 10^{34} $cm^{-2}s^{-1}$
- B – physics: high pt muon triggers



LHCb:

- B – physics dedicated experiment
- Forward spectrometer $1.9 < \eta < 4.9$
- Lower pt triggers. Efficient also for purely hadronic channels (see talk of Leandro de Paula)
- Instant Luminosity 2-5 $x10^{32}$ $cm^{-2}s^{-1}$

ANALYSIS OVERVIEW

Triggered and offline reconstructed (incl. muon identification) **signal** events per fb^{-1} (i.e., effective $B_s \rightarrow \mu\mu$ cross section)

	ATLAS	CMS	LHCb
# evts/fb^{-1}	13.3	13.39	36.2
For trigger strategy	$L = 10^{33}$	$L > 10^{32}$	$L = 2x10^{32}$

$\sigma_{b\bar{b}}$ assumed to be 500 μbarn, BR($B_s \rightarrow \mu\mu$) = (SM)
ATLAS/LHCb: 3.35 $x10^{-9}$ CMS: 3.9 $x10^{-9}$

M. Artuso et al.
Eur. Phys. J. C (2008) 57: 309–492 (see expr. 128)

Main issues:

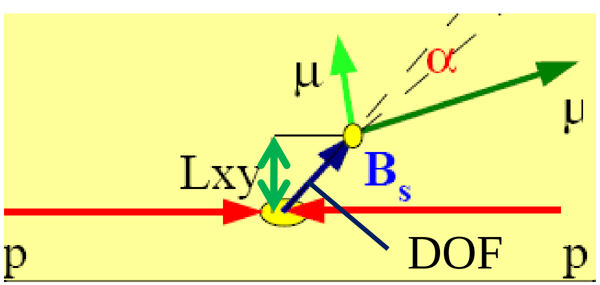
- Background discrimination: offline cuts/ multivariate analysis
- Normalization to another B channel with well known BR
  - It avoids needing the knowledge of xsections & integrated luminosity
  - Cancelation of systematic uncertainties

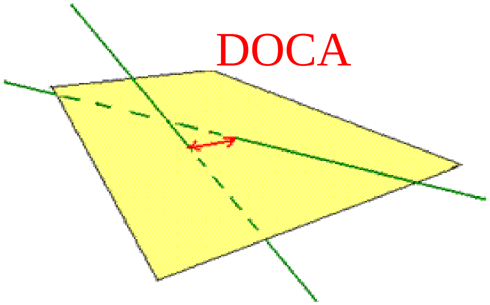
ATLAS analysis: CERN-OPEN-2008-020 [arXiv:0901.0512] (B-physics chapter)
CMS analysis: CMS PAS BPH-07-001 (2009)
LHCb analysis: LHCb-PUB-2007-033 (2007) , LHCb-PUB-2008-018 (2008)

USEFUL VARIABLES

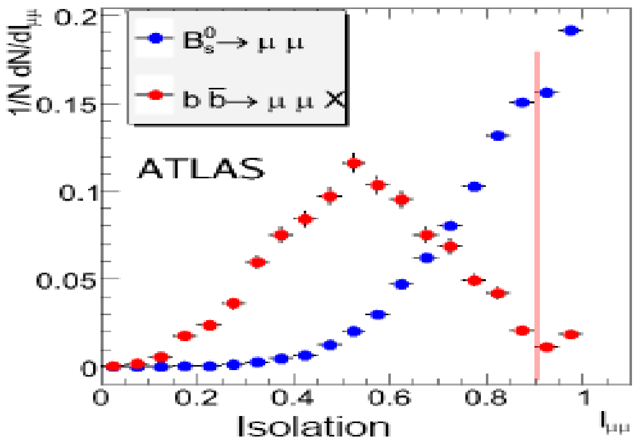
- Usual signatures of a given B decay:
  - Detached Secondary Vertex: large lifetime, distance of flight (DOF), Impact Parameter (IP) of daughters…
  - B coming from Primary Vertex: small B IP, small momentum-to-flight direction ("pointing")
  - Good quality Secondary Vertex: small χ^2, small DOCA (Distance Of Closest Approach)
  - Isolation
    - ATLAS / CMS: $Iso = \frac{p_T(B)}{p_T(B) + \sum p_T^i(\Delta R_i < 1.0)}$ — > 1GeV (ATLAS), > 0.9 GeV (CMS); $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$
    - LHCb: Above definition not suitable for LHCb geometry. Isolation is defined per muon as the no. of tracks compatible with a common μ-track SV
  - Invariant Mass around Bs: For combinatorial bkg. sensitivity scales as $1/\sqrt{\sigma_M}$

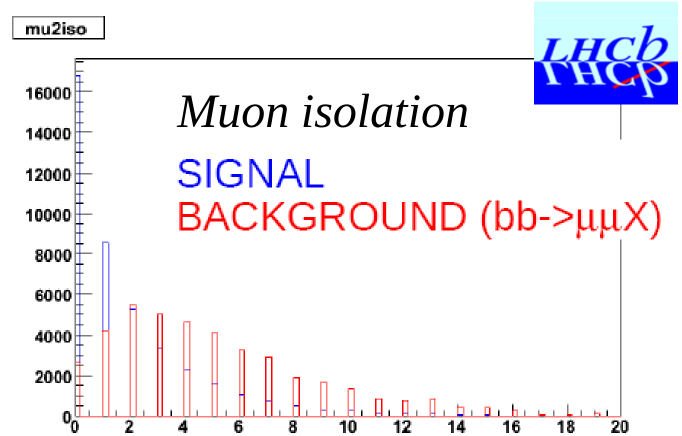
σ_{ATLAS} ~ 90 MeV, σ_{CMS} ~ 53 MeV, σ_{LHCb} ~ 22 MeV

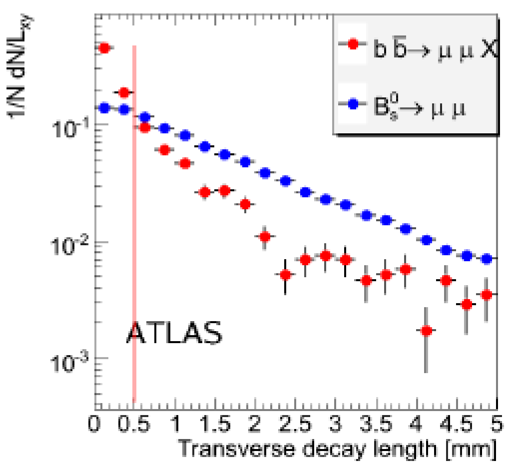


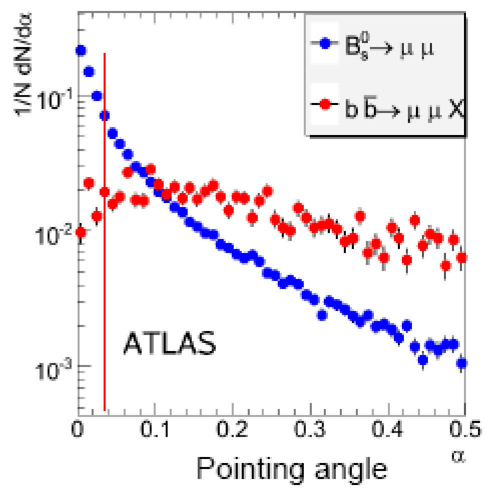


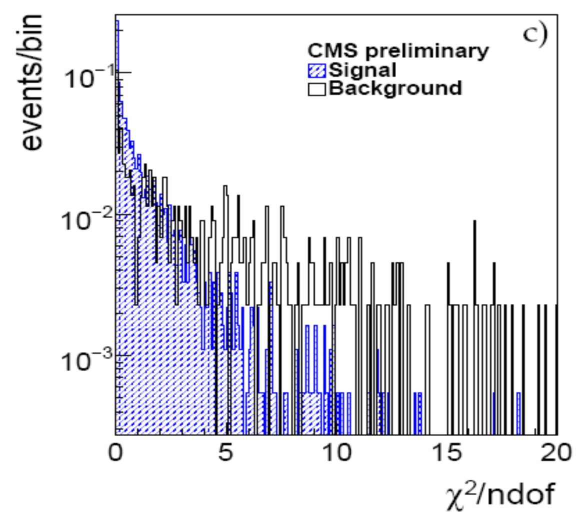
USEFUL VARIABLES



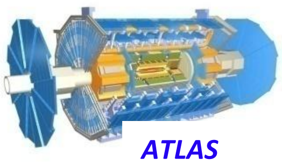








ATLAS / CMS

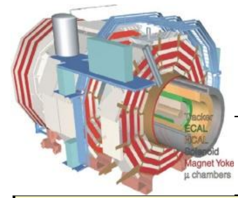


$\sigma_{b\bar{b}}$ assumed to be 500 μbarn

	$B_s \to \mu\mu$ efficiency	$bb \to \mu\mu X$ efficiency
Isolation>0.9	0.24	$(2.6\pm0.3)\mathrm{x}10^{-2}$
$L_{xy} > 0.5$ mm	0.26	$(1\pm0.3)\mathrm{x}10^{-3}$ *)
$\alpha < 0.017$ rad	0.23	
$M=M_{Bs}{}^{+140}_{-70}$MeV	0.76	0.079
$Evts/10fb^{-1}$ BR = $3.35\mathrm{x}10^{-9}$	**5.6**	**14^{+13}_{-10}**

(Efficiencies w.r.t following preselection criteria: 4 <M <7.3 GeV, $\chi^2 < 10$, $L_{xy} < 2$ cm. Isolation cut in signal also includes a factor 0.46 from trigger efficiency. This cuts are for analysis with L >~ $10fb^{-1}$)

ATLAS is also preparing an analysis based on a boosted decision tree

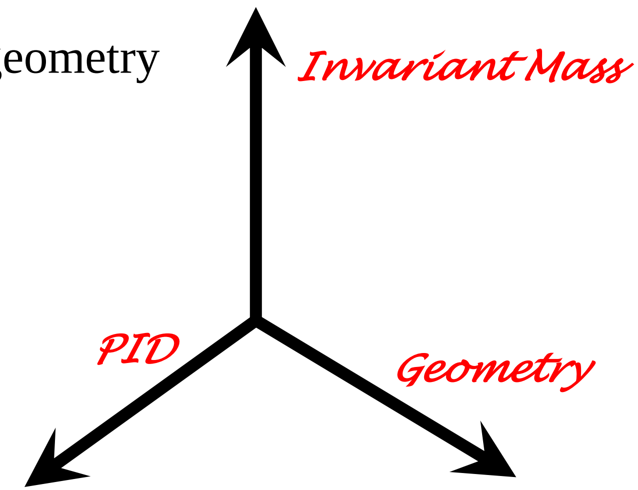


CMS	$B_s \to \mu\mu$	$bb \to \mu\mu X$
4.8 <M <6 GeV	~1.	0.048
cos(α)>0.9985	0.73	0.11
DOF > 17 σ	0.58	0.092
$\chi^2 < 5$	0.94	0.411
Isolation > 0.85	0.47	0.018
$\|M - M_{Bs}\| < 100$ MeV	0.94	0.17
$Evts/fb^{-1}$ **BR* =$3.9\mathrm{x}10^{-9}$**	**2.36**	**$2.5^{+0.7}_{-0.6}$**

*M. Artuso et al. Eur. Phys. J. C (2008) 57: 309–492 (see expr. 128)

CMS estimates total bkg as ~6.53

- LHCb uses cuts just to get a reasonable rate of events to analyze
- Selected signal candidates are classified in a 3D parameter space, according to:
  - **Invariant mass** (in a window of 60 MeV around B_s peak)
  - **PID likelihood** with info from different subdetectors, to get rid of possible remaining misid
  - **Geometry likelihood**:
    - Combines several variables related candidate geometry
    - Best separation power

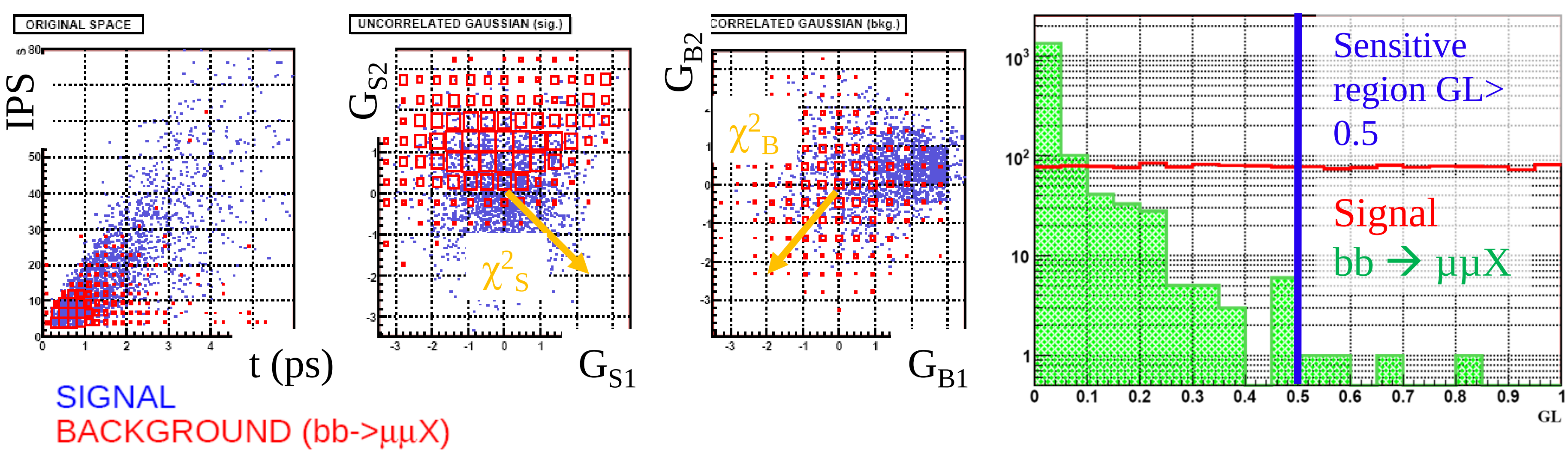


- 3D space is binned, so that **each bin is treated as an independent experiment**
- Results are combined using **Modified Frequentist Approach**.

How the Geometry likelihood is built:

1. Input variables: min Impact Parameter Significance (μ^+,μ^-), DOCA, Impact Parameter of B, lifetime, iso - μ^+, iso- μ^-
2. They are transformed to Gaussian through cumulative and inverse error function
3. In such space correlations are more linear-like → rotation matrix, and repeat 2
4. Transformations under signal hyp. → χ^2_S, under bkg. → χ^2_B.
5. Discriminating variable is χ^2_S -χ^2_B, made flat for better visualization.

lifetime



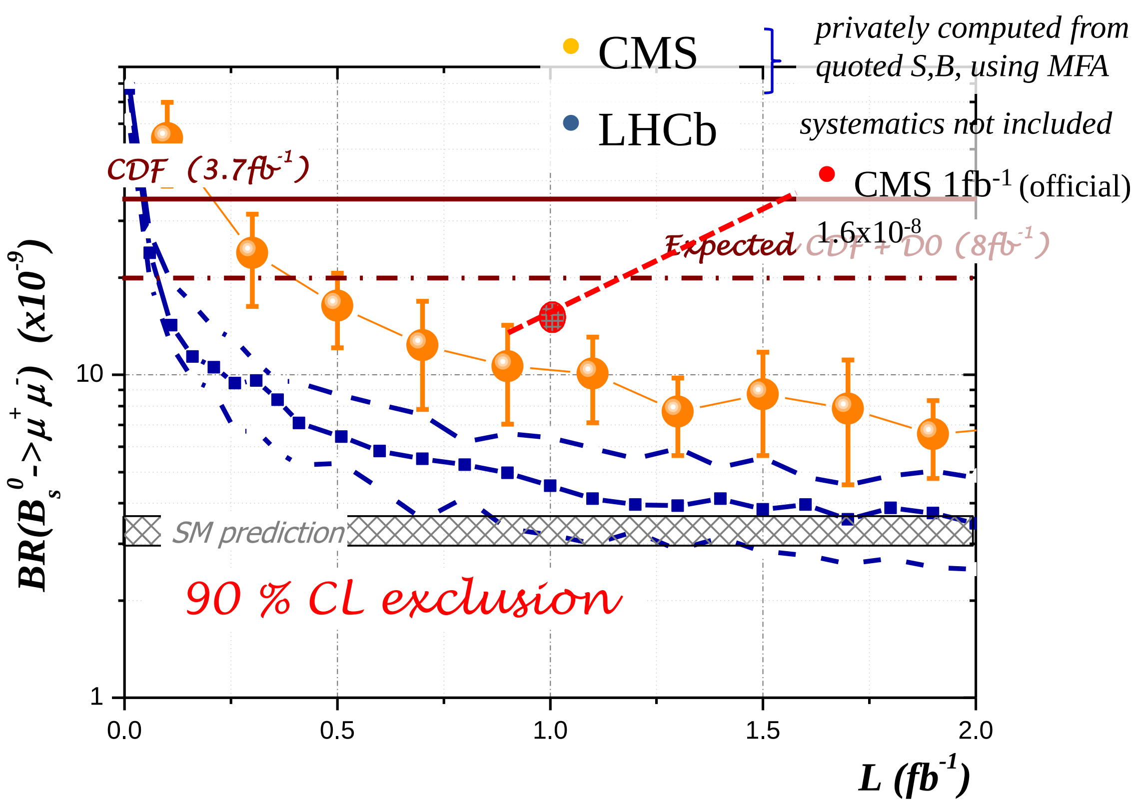
SENSITIVITIES

(expected S (for BR = 3.35e-9) & B per fb^{-1} in each experiment LHCb bins parameter space → N experiments)

CMS: S (BR = 3.35e-9) = 2.05
B = 6.53

LHCb	GL	
Mass (MeV)	**0.5–0.65**	**0.65-1**
5406.6 - 5429.6	$S = 0.13$ $B = 8^{+10}_{-5}$	$S = 0.3$ $B = 8^{+10}_{-5}$
5384.1 - 5406.6	$S = 0.55$ $B = 8^{+10}_{-5}$	$S = 1.4$ $B = 8^{+10}_{-5}$
5353.4 - 5384.1	$S = 1.6$ $B = 11^{+15}_{-7}$	$S = 3.8$ $B = 11^{+15}_{-7}$
5331.5 – 5353.4	$S = 0.6$ $B = 8^{+10}_{-5}$	$S = 1.5$ $B = 8^{+10}_{-5}$
5309.6 – 5331.5	$S = 0.2$ $B = 8^{+10}_{-5}$	$S = 0.45$ $B = 8^{+10}_{-5}$

- 90% CL exclusion sensitivity as a function of L
  - (Only bkg is observed)



SENSITIVITIES

(expected S (for BR = 3.35e-9) & B per fb^{-1} in each experiment LHCb bins parameter space → N experiments)

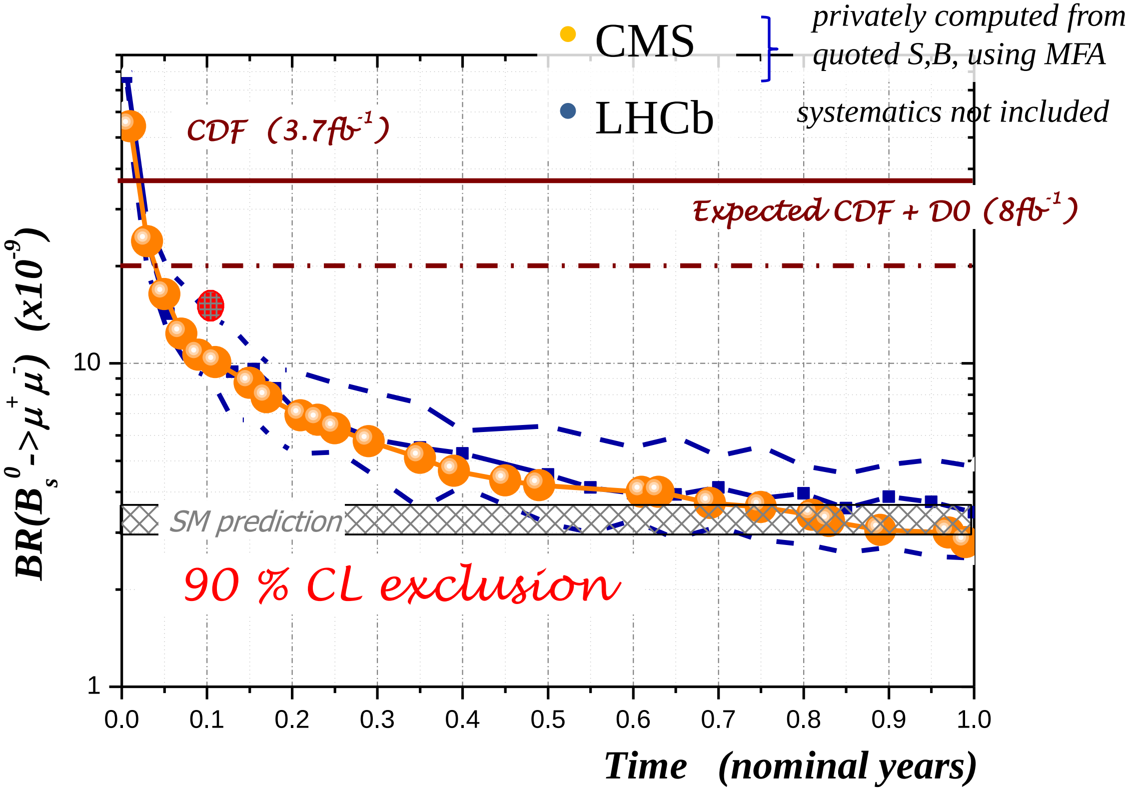
CMS

S (BR = 3.35e-9) = 2.05
B = 6.53

LHCb	GL	
Mass (MeV)	**0.5–0.65**	**0.65-1**
5406.6 - 5429.6	$S = 0.13$ $B = 8^{+10}_{-5}$	$S = 0.3$ $B = 8^{+10}_{-5}$
5384.1 - 5406.6	$S = 0.55$ $B = 8^{+10}_{-5}$	$S = 1.4$ $B = 8^{+10}_{-5}$
5353.4 - 5384.1	$S = 1.6$ $B = 11^{+15}_{-7}$	$S = 3.8$ $B = 11^{+15}_{-7}$
5331.5 – 5353.4	$S = 0.6$ $B = 8^{+10}_{-5}$	$S = 1.5$ $B = 8^{+10}_{-5}$
5309.6 – 5331.5	$S = 0.2$ $B = 8^{+10}_{-5}$	$S = 0.45$ $B = 8^{+10}_{-5}$

• 90% CL exclusion sensitivity as a function of time

Assuming nominal luminosities since the beginning
CMS → L = 10^{33} $cm^{-2}s^{-1}$
LHCb → L = $2x10^{32}$ $cm^{-2}s^{-1}$



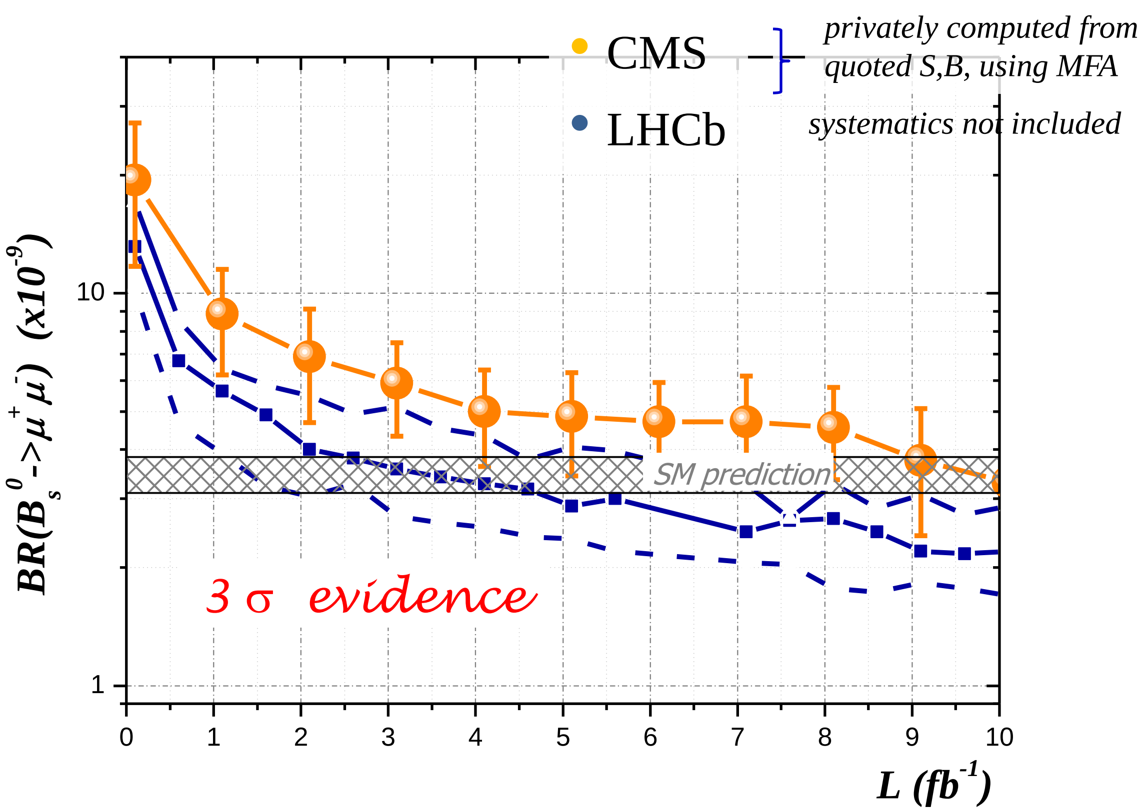
SENSITIVITIES

(expected S (for BR = 3.35e-9) & B per fb^{-1} in each experiment LHCb bins parameter space → N experiments)

S (BR = 3.35e-9) = 2.05
B = 6.53

LHCb	GL	
Mass (MeV)	0.5–0.65	0.65-1
5406.6 - 5429.6	$S = 0.13$ $B = 8^{+10}_{-5}$	$S = 0.3$ $B = 8^{+10}_{-5}$
5384.1 - 5406.6	$S = 0.55$ $B = 8^{+10}_{-5}$	$S = 1.4$ $B = 8^{+10}_{-5}$
5353.4 - 5384.1	$S = 1.6$ $B = 11^{+15}_{-7}$	$S = 3.8$ $B = 11^{+15}_{-7}$
5331.5 – 5353.4	$S = 0.6$ $B = 8^{+10}_{-5}$	$S = 1.5$ $B = 8^{+10}_{-5}$
5309.6 – 5331.5	$S = 0.2$ $B = 8^{+10}_{-5}$	$S = 0.45$ $B = 8^{+10}_{-5}$

- Signal evidence sensitivity as a function of L
  - (Signal + Background observed)



SENSITIVITIES

(expected S (for BR = 3.35e-9) & B per fb^{-1} in each experiment LHCb bins parameter space → N experiments)

CMS: $S_{(BR = 3.35e-9)} = 2.05$

$B = 6.53$

LHCb:

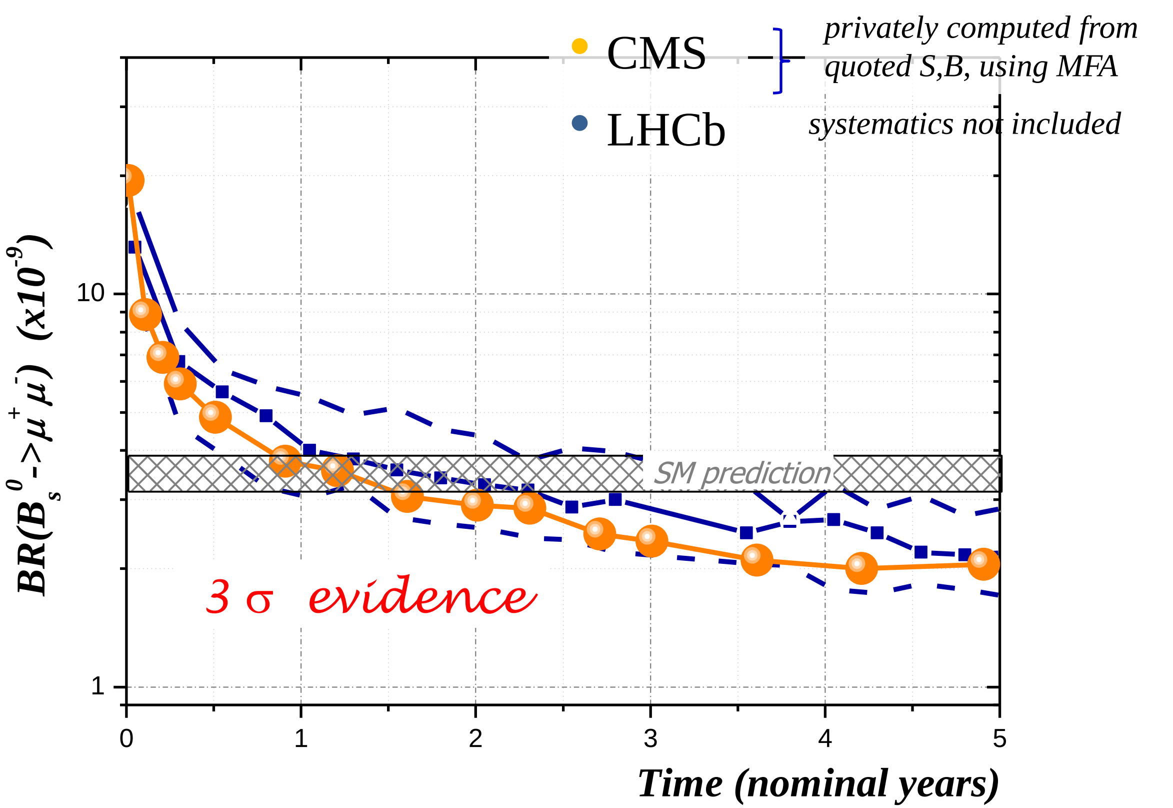
	GL	
Mass (MeV)	0.5–0.65	0.65-1
5406.6 - 5429.6	$S = 0.13$ $B = 8^{+10}_{-5}$	$S = 0.3$ $B = 8^{+10}_{-5}$
5384.1 - 5406.6	$S = 0.55$ $B = 8^{+10}_{-5}$	$S = 1.4$ $B = 8^{+10}_{-5}$
5353.4 - 5384.1	$S = 1.6$ $B = 11^{+15}_{-7}$	$S = 3.8$ $B = 11^{+15}_{-7}$
5331.5 – 5353.4	$S = 0.6$ $B = 8^{+10}_{-5}$	$S = 1.5$ $B = 8^{+10}_{-5}$
5309.6 – 5331.5	$S = 0.2$ $B = 8^{+10}_{-5}$	$S = 0.45$ $B = 8^{+10}_{-5}$

- Signal evidence sensitivity as a function of time

Assuming nominal luminosities since the beginning

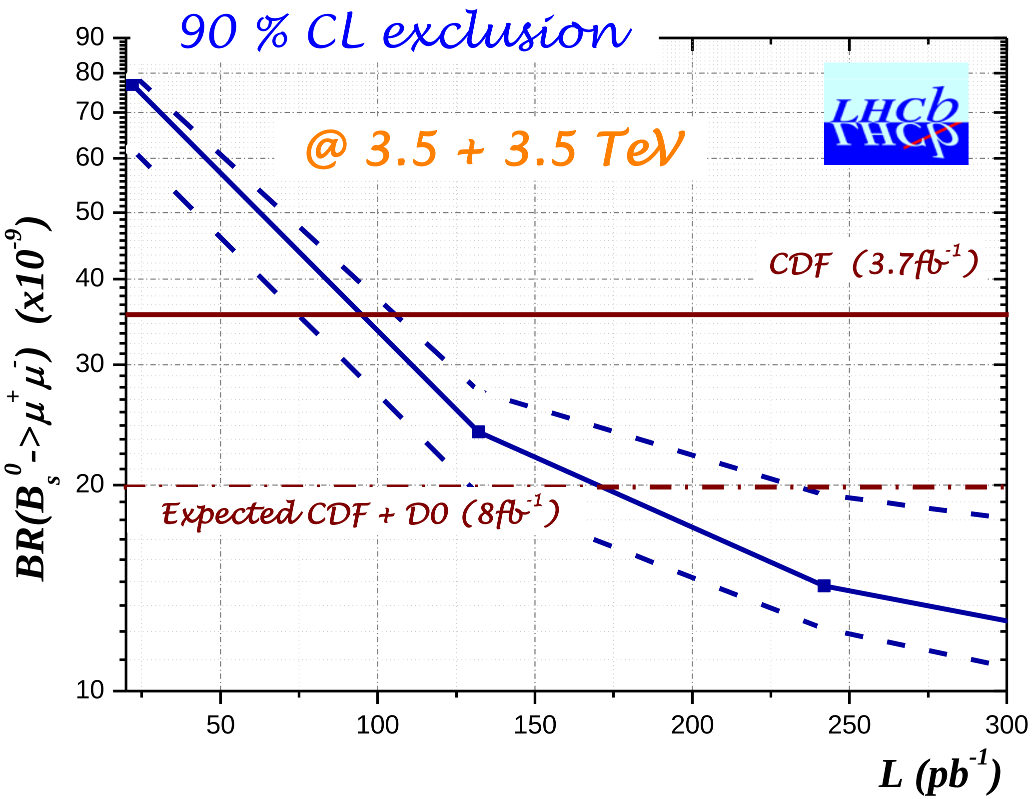
CMS → $L = 10^{33}\ cm^{-2}s^{-1}$

LHCb → $L = 2x10^{32}\ cm^{-2}s^{-1}$



LHC STARTUP

- LHC first data:
  - •Less energy (3.5 + 3.5 TeV)
  - •Less instant luminosity
- Exclusion sensitivity for
  - •45% of σ_{bb} w.r.t. 14 TeV (Pythia ratio $\sigma_{bb\_7TeV}/\sigma_{bb\_14TeV}$), so 225 μb
  - •First 10 months after LHC startup (assumed 300 pb^{-1})
- This data could allow LHCb to overtake Tevatron limits and impose new constraints on SUSY models



NORMALIZATION & CALIBRATION

NORMALIZATION

- Normalization is needed to convert # events into a BR w/o relying on knowledge of σ_{bb}, integrated luminosity or absolute efficiencies

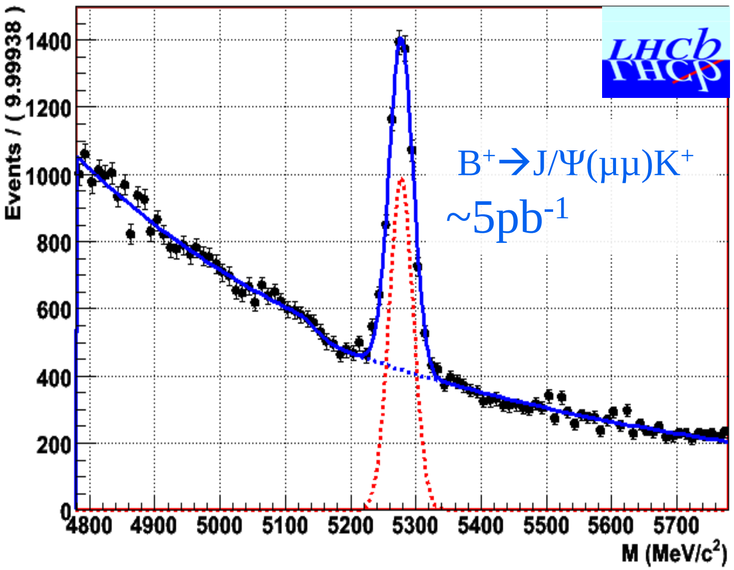
$$BR = BR_n \frac{\varepsilon_n}{\varepsilon} \cdot \frac{P(b \to B_n)}{P(b \to B_s)} \cdot \frac{N}{N_n}$$

- $P(b \to B^+, B_d)/P(b \to B_s)$ implies a **~14 % systematic**. Normalization to a B_s mode would introduce larger errors because of poorly known B_s BR's

- The fraction of efficiencies (acceptance, trigger, selection, PID…) needs to be computed/cancelled.

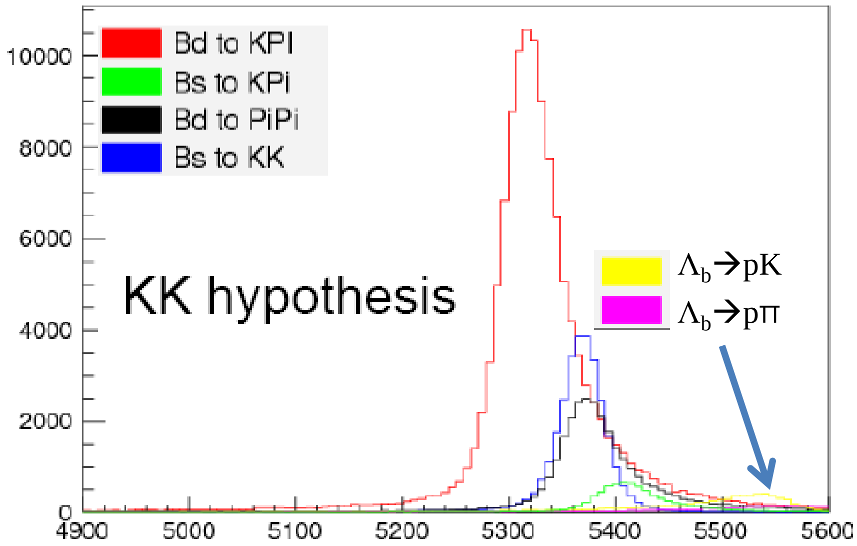
- **ATLAS/CMS/LHCb** : to $B^+ \to J/\Psi(\mu\mu)K^+$
  - Similar trigger and muon ID
  - The selection can be made similar to signal
  - But: Extra track to be reconstructed

  $B_d \to J/\Psi K^*$ / $B^+ \to J/\Psi(\mu\mu)K^+$ or other similar ratios allow to study this



NORMALIZATION ($B \to k\pi$)

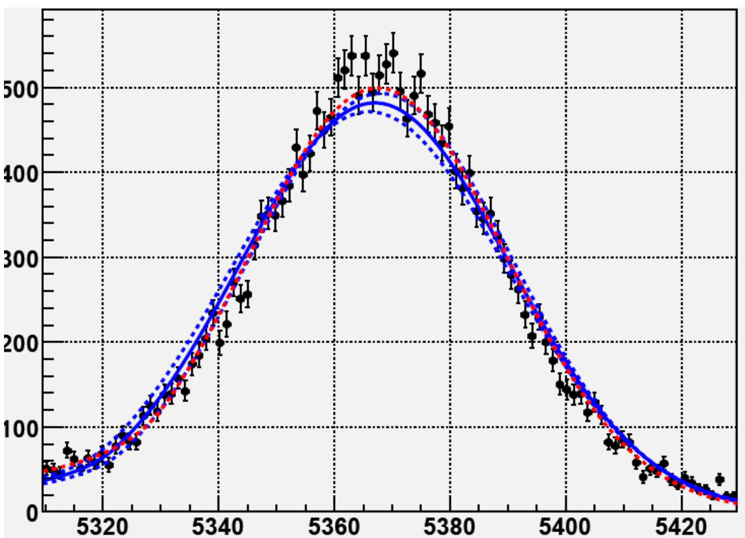
- **LHCb** also uses normalization to $B \to h^+h^-$ ($B_{d,s} \to K\pi$, $B_d \to \pi\pi$, $B_s \to KK$…)
- Same geometry & kinematics than signal, different trigger (hadronic) and PID
- How to get rid of the differences:
  - Use $B \to hh$ events **Triggered Independently of Signal**
    - Several thousands of such events per fb^{-1} will be available
  - Use $b \to J/\Psi X$ to **emulate muon ID and trigger** on that sample as a function of p/pt



- The most suitable mode: $B_d \to K\pi$ (well known BR, largest statistics…)
- It can be separeted from the inclusive sample using the RICH (see talk of Laurence Carson)

CALIBRATION

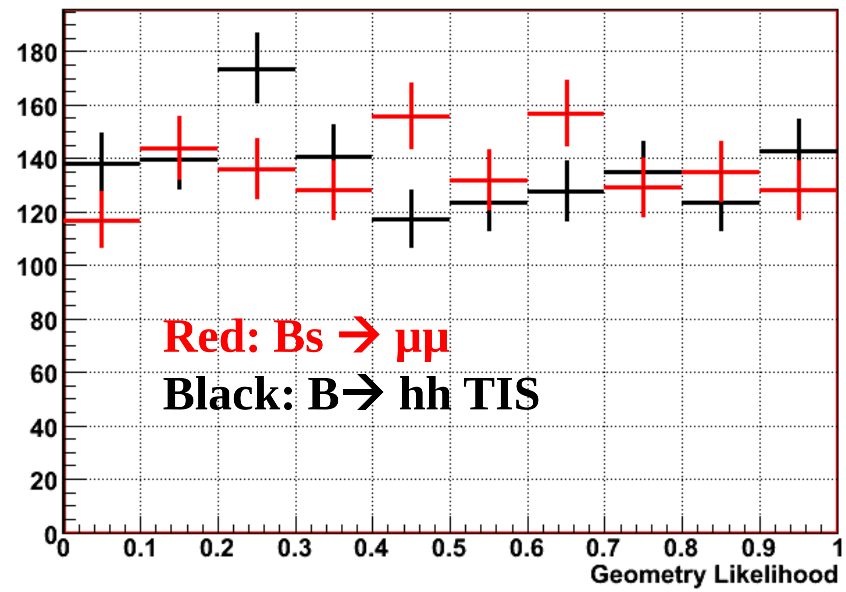
- **LHCb**: signal is distributed in several bins of a 3D space
- We need to know not only overall normalization, also the fraction of signal in each bin
  - **Invariant mass** → Can be calibrated with B_s → KK
  - **GL** → (inclusive) B→hh triggered independent of signal (TIS)
  - **PID likelihood** → J/Ψ taking p, pt distributions from B→ hh TIS



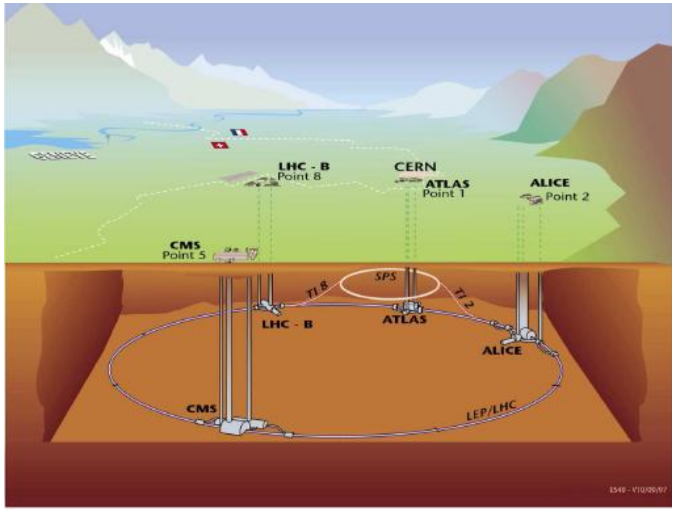
Data: Bs → μμ

Red: Fit to data itself

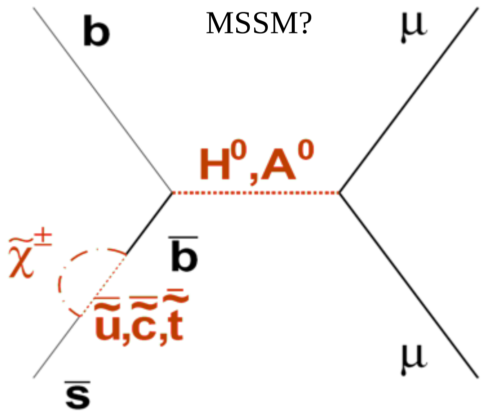
Blue: Function from calibration



CONCLUSIONS



- A measurement/exclusion of BR($B_s \rightarrow \mu\mu$) will have an important impact on NP searches
- LHC offers exceptional conditions for this study, scanning from current upper limit to < SM prediction
- LHCb takes advantage of its B-physics dedicated trigger, as well as good invariant mass resolution, having the best sensitivity for a given luminosity
- ATLAS/CMS benefit from their capabilities to run at higher luminosities
- The use of control channels such as $B^+ \rightarrow J/\Psi(\mu\mu)K^+$ and $B \rightarrow hh$ allows to perform a MC free analysis



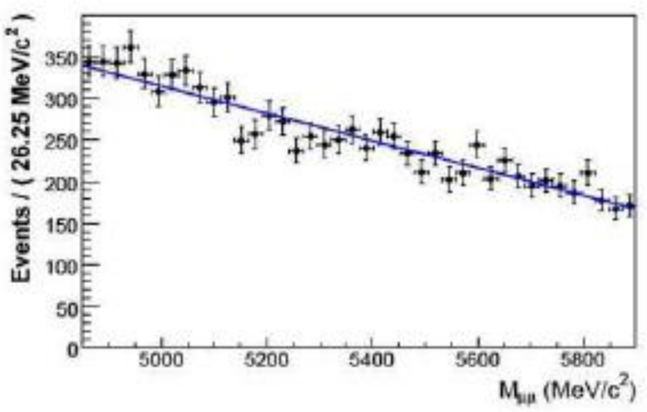
Title of the Slide

BACKGROUND LEVEL

- ATLAS/CMS/LHCb: amount of bkg in the signal region has to be known

- Bkg is dominated by combinatorial (bb→μμX) and hence can be understood from sidebands

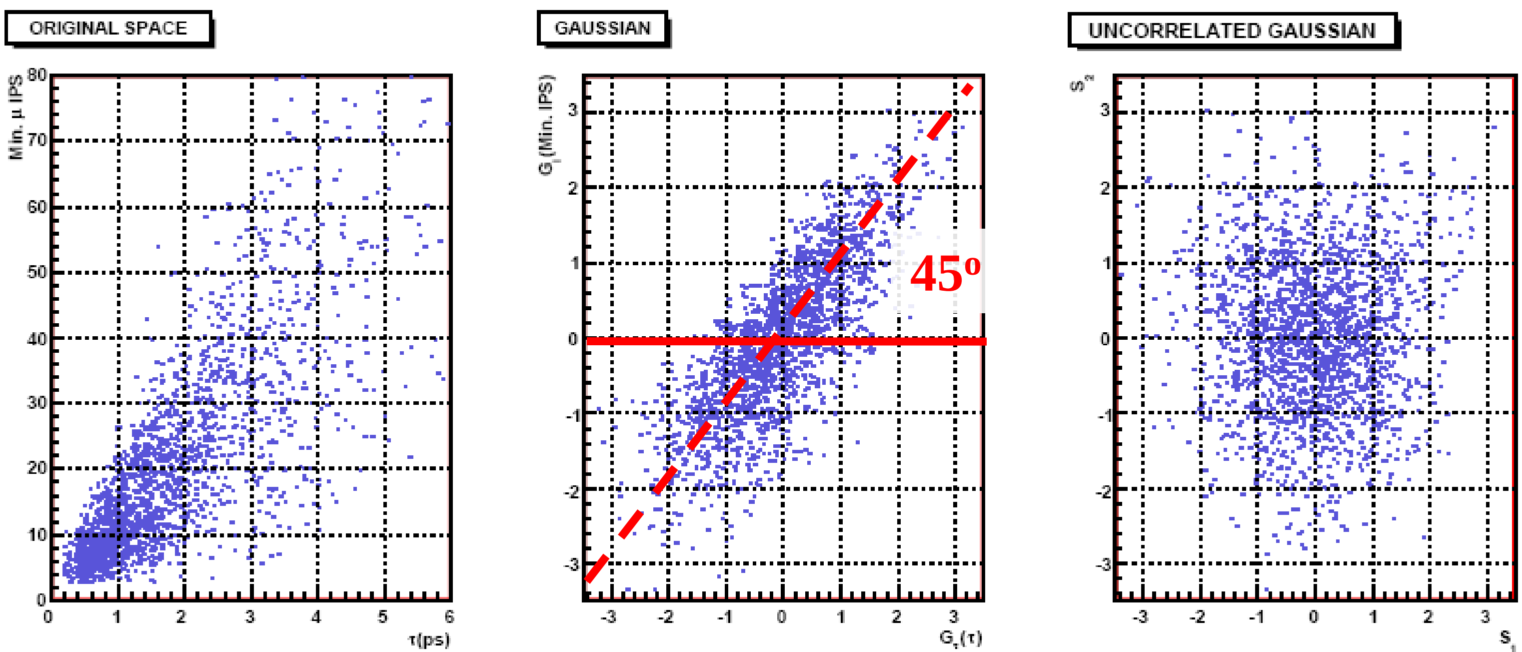
- Linear or exponential fit gives the bkg level in the signal region



- Specific/peaking bkg is negligible in current simulations

How the Geometry likelihood is built:

1. Input variables: min IPS (μ^+,μ^-), DOCA, IP of B, lifetime, iso - μ^+, iso- μ^-
2. They are transformed to gaussian through cumulative and inverse error function
3. In such space correlations are more linear-like → rotation matrix, and repeat 2



Sensitivity to B0

Supposing bb → mumu is also the dominant bkg at the
Bd window, for each luminosity you can access to 3-4 times smaller BR for Bd than for Bs.

ROUGH SENSITIVITY CALCULATION

- Signal yield → $\sigma^{eff}*L$
- bkg under the peak scales linearly with invariant mass resolution σ_M

$$S/\sqrt{B} \propto \frac{\sigma_{sig}^{eff}}{\sqrt{\sigma_{bkg}^{eff}\sigma_M}}\sqrt{L}$$

NORMALIZATION ($B \rightarrow k\pi$)

- $B_d \rightarrow K\pi$ has to be separated from the inclusive sample → Use of the RICH system → Extra efficiency factor to account for

- $B \rightarrow hh$ can self-calibrate this eff. using ratio $B_d \rightarrow K\pi$ / $B_d \rightarrow \pi\pi$ (very well known ratio of xsections) and the number of inclusive $B \rightarrow hh$, as well as the good B_s-B_d mass separation in LHCb

- Alternatively, $D^* \rightarrow D^0(K\pi)\,\pi$ reweighting by p,pt, can be also used (see Laurence Carson talk)

f(Bd → Kπ) = 0.677 ± 0.039
(MC = 0.681)

f(Bd → π π) = 0.169 ± 0.015
(MC = 0.172)

f(Bs → Kπ) = 0.0401 ± 0.0012
(MC = 0.0435)

f(Bs → KK) = 0.114 ± 0.011
(MC = 0.102)

Output of a MC experiment using $B_d \rightarrow K\pi$ / $B_d \rightarrow \pi\,\pi$ to calibrate RICH effs.

Title of the slide

Full expression (μ_q the ratio of masses m_q/m_b)

$$BR(B_q \to \mu^+\mu^-) = \frac{G_F^2 \alpha^2}{64\pi^3 \sin^4\theta_W} |V_{tb}^* V_{tq}|^2 \tau_{Bq} M_{Bq}^3 f_{Bq}^2 \sqrt{1 - \frac{4m_\mu^2}{M_{Bq}^2}} \times$$

$$\times \left\{ M_{Bq}^2 \left(1 - \frac{4m_\mu^2}{M_{Bq}^2}\right) \left(\frac{C_S - \mu_q C_S^{'}}{1+\mu_q}\right)^2 + \left[M_{Bq} \left(\frac{C_P - \mu_q C_P^{'}}{1+\mu_q}\right) + \frac{2m_\mu}{M_{Bq}} \left(C_A - C_A^{'}\right) \right]^2 \right\}$$

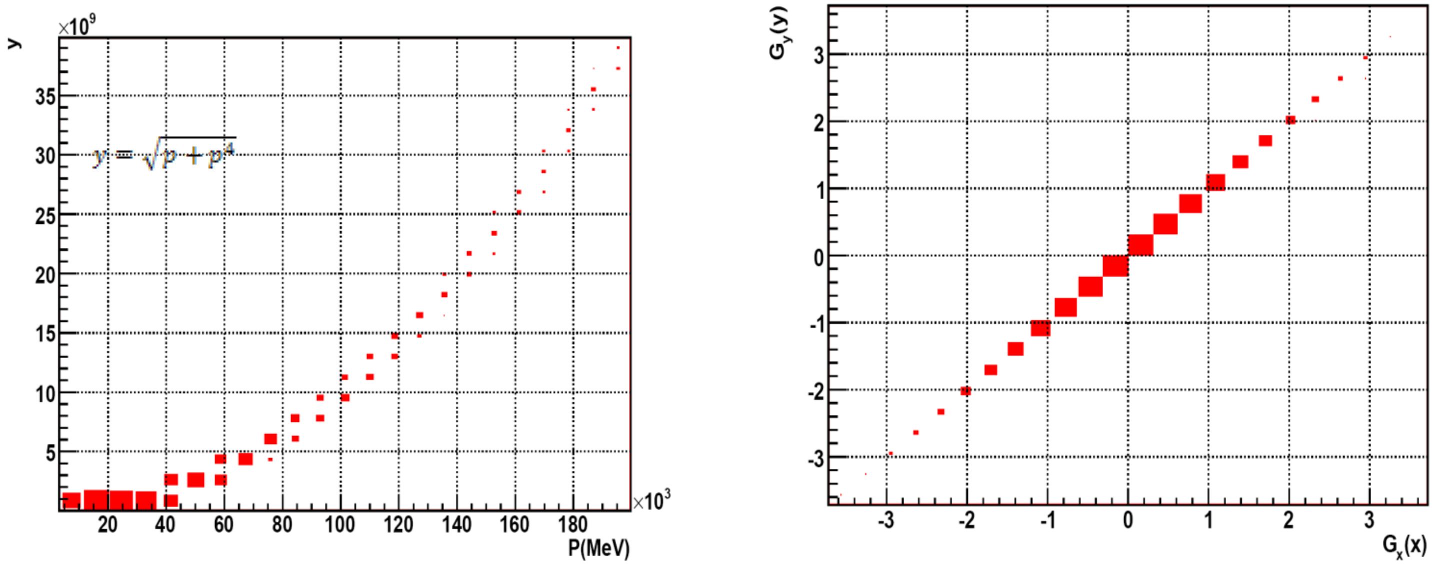


Figure -: Correlation in initial and Gaussian space.

Separation of Bd□Kπ

Extract the **fraction** of different components of B □ hh, without relying on MC PID efficiencies:

1. Measure those fractions in a "high purity" limit (PID cuts > X):

(Example for X = 20):

KK □ $N'_{kk} = 502$ → $f'_{kk} = 0.109$

Kπ □ $N'_{k\pi} = 3292$ → $f'_{k\pi} = 0.712$

ππ □ $N'_{\pi\pi} = 827$ → $f'_{\pi\pi} = 0.179$

Not necessary the same as in the nonPID B□ hh sample !!!

(Then the true fraction should be):

$$f_{K\pi} = \frac{f'_{K\pi} / \varepsilon_K \varepsilon_\pi}{f'_{KK} / \varepsilon_K^2 + f'_{K\pi} / \varepsilon_K \varepsilon_\pi + f'_{\pi\pi} / \varepsilon_\pi^2} = \frac{f'_{K\pi}}{f'_{K\pi} + f'_{KK}\left(\varepsilon_\pi / \varepsilon_K\right) + f'_{\pi\pi}\left(\varepsilon_K / \varepsilon_\pi\right)}$$

(Separate Bs □ Kπ and Bd □ Kπ is not an issue because of the mass resolution)

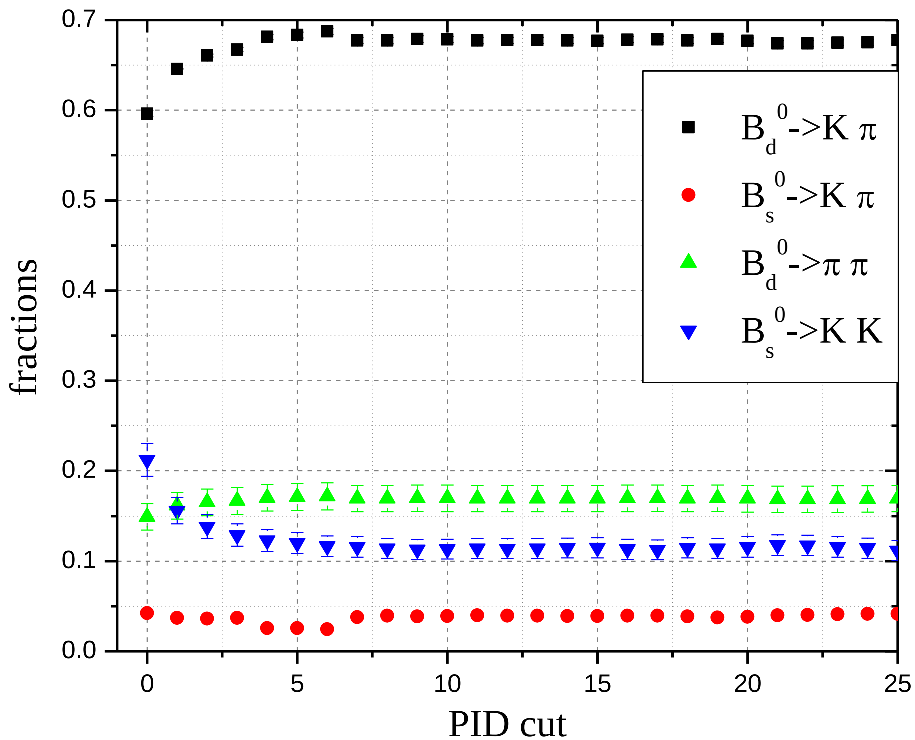
35

Separation of Bd□Kπ (II)

2. The ratio ($\varepsilon_\pi/\varepsilon_K$) □ thus the right fractions can be easily extracted from Bd modes, where the BR's are known.

$$\frac{N(B_d^0 \to K\pi)}{N(B_d^0 \to \pi\pi)} = \frac{BR(B_d^0 \to K\pi)}{BR(B_d^0 \to \pi\pi)} = 3.96 \pm 0.36 \Rightarrow \boxed{\frac{\varepsilon_\pi}{\varepsilon_K} = (3.96 \pm 0.36) \cdot \frac{N'_{\pi\pi}}{N'^{(d)}_{K\pi}}}$$

3. To ensure the high purity limit, repeat 1 & 2 until a plateau on the results is reached



f(Bd □ Kπ) = 0.677 ± 0.039
(MC = 0.681)

f(Bd □ π π) = 0.169 ± 0.015
(MC = 0.172)

f(Bs □ Kπ) = 0.0401 ± 0.0012
(MC = 0.0435)

f(Bs □ KK) = 0.114 ± 0.011
(MC = 0.102)

36