Search for $B_{(s,d)} \rightarrow \mu\mu$ in LHCb

Diego Martinez Santos (CERN)

(on behalf of the LHCb Collaboration)

CERN-PH-EP-2011-029, submitted to PLB
http://arxiv.org/abs/1103.2465

Outline

- Observables: Branching Ratios (BR) of $B_{s,d} \to \mu\mu$

1 Why do we want to measure that?

- indirect probe for New Physics(NP)

2 How do we do it? (Analysis strategy)

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

3 What did we get with 2010 data?

- Interesting results
- Amazing prospects

Indirect Approach

- $B_{s,d} \rightarrow \mu\mu$ can access NP through new virtual particles entering in the loop → indirect search

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

  - 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

  - ☺ Roundness of Earth (Eratosthenes, c.III B.C) discovered ~2300 years before direct observation

Eratosthenes

~2.3 k years till the direct observation…

SM and New Physics

This decay is very suppressed in SM :

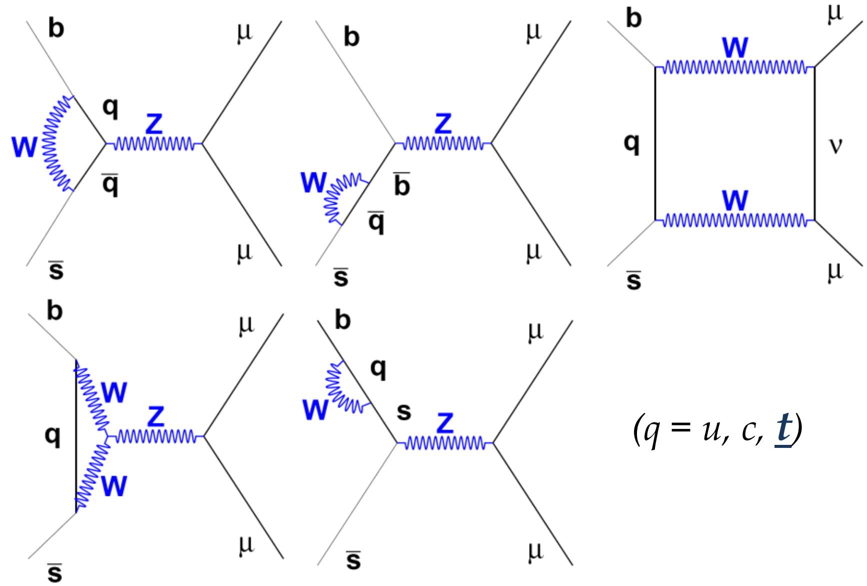
$$\mathbf{BR(B_s \to \mu\mu) = (3.2 \pm 0.2)x10^{-9}}$$
$$\mathbf{BR(B_d \to \mu\mu) = (1.0 \pm 0.1)x10^{-10}}$$

Experimental upper limit still one order of magnitude above such values. @ 95% CL:

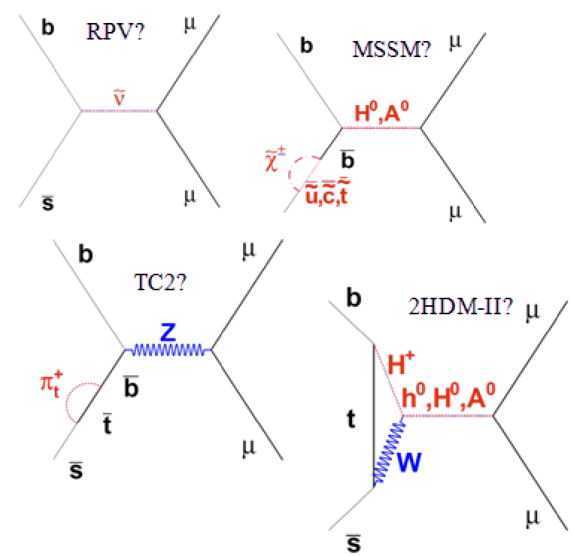
$BR(B_s \to \mu\mu) < 4.3x10^{-8}$
$BR(B_d \to \mu\mu) < 0.76x10^{-8}$ (**CDF**, 3.7 fb^{-1}, prel.)
$BR(B_s \to \mu\mu) < 5.1x10^{-8}$ (**D0**, 6.1 fb^{-1}, publ.)

But in NP models it can take any value from << SM (e.g, some NMSSM) up to current experimental upper limit (e.g. SUSY at high tanβ)

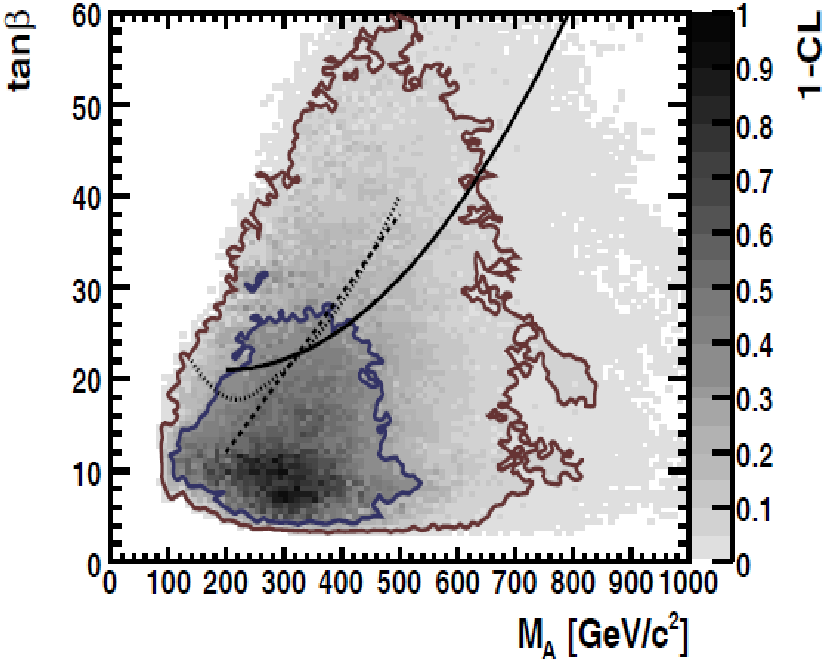
→ Whatever the actual value is, it will have an impact on NP searches



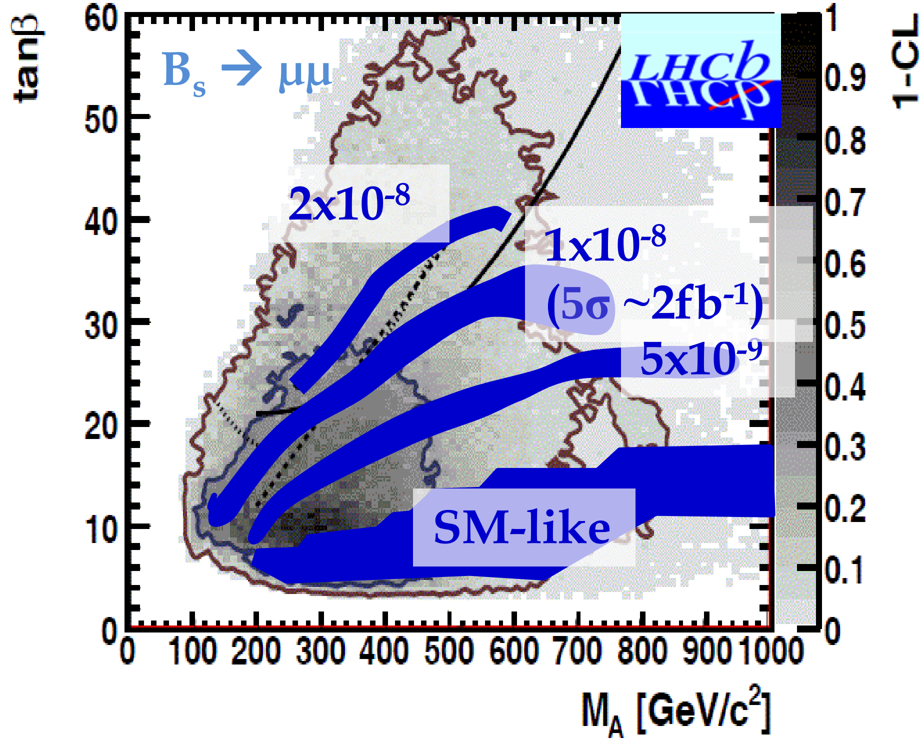
+?



Best fit contours in tanβ vs MA plane in the **NUHM1** model
[O. Buchmuller et al, Eur.Phys.J.C64:391-415,2009*]*



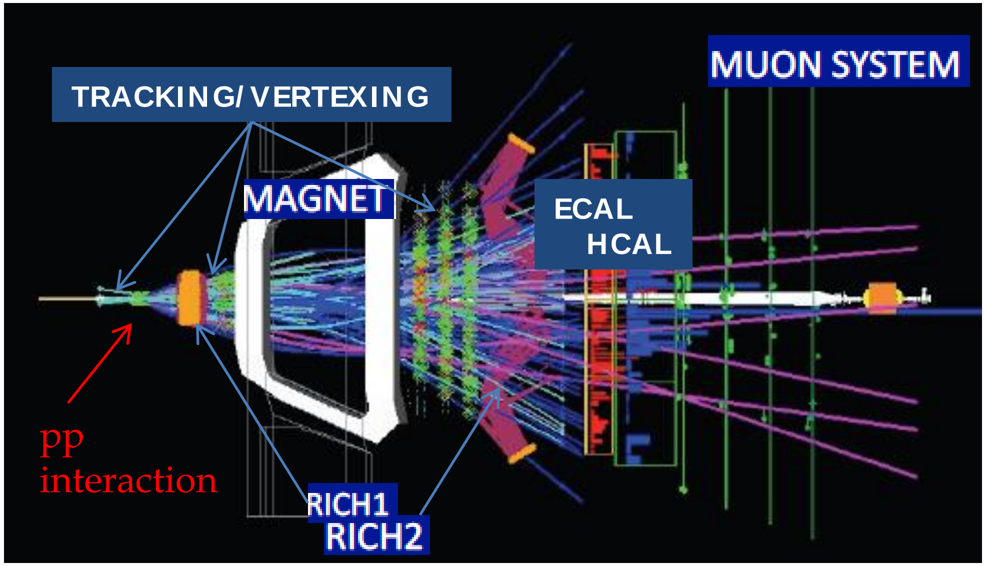
CMS
H/A $\rightarrow \tau\tau$,
5σ, ~30-60fb^{-1}



Regions compatible with
$BR(B_s \rightarrow \mu\mu) \sim 2\times10^{-8}$, 1×10^{-8},5×10^{-9} and SM-like.

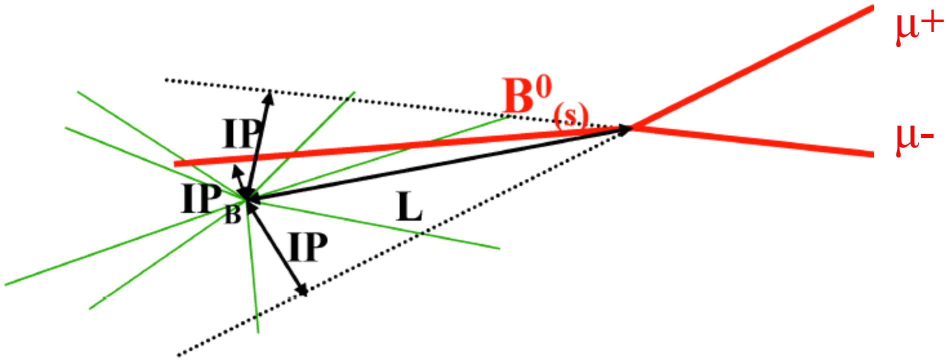
Private calculation using SuperIso program, (F. Mahmoudi, arXiv: 08083144)
and SoftSusy (B.C. Allanach, Comput. Phys. Commun. 143 (2002) 305-331)

LHCb



- Low angle spectrometer
- Very efficient trigger
- Good particle identification performance
- Precise reconstruction:
  - Separation production vertex – decay vertex σ(IP)~ 25 μm
  - Invariant mass Δp/p ~0.35-0.55%

- $B_{s,d} \to \mu\mu$ signature:
  - Hits in muon detector
  - μμ pair has B invariant mass
  - geometrical & kinematical signature: pt, detachment of decay vertex



Analysis strategy

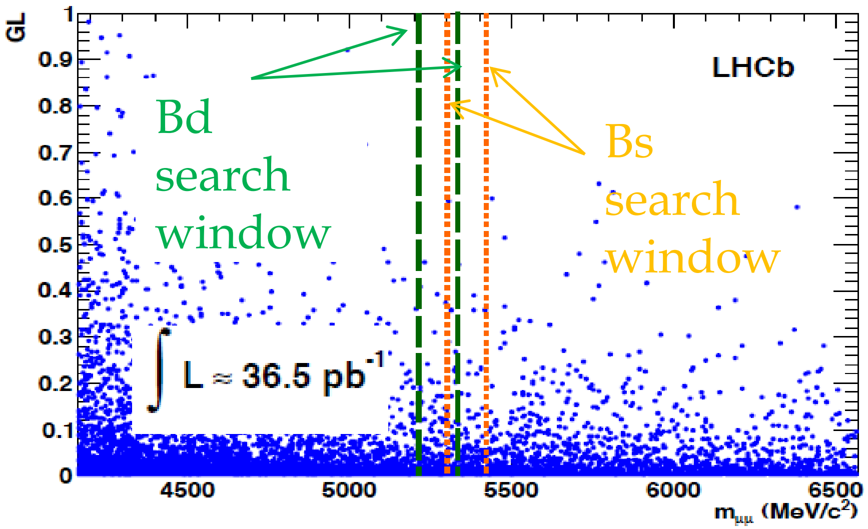
- Selection cuts in order to reduce the amount of data to analyze. LHCb trigger selects > 90% of the signal that is interesting for the offline analysis.

- Classification of $B_{s,d} \rightarrow \mu\mu$ events in bins of a 2D space

  - Invariant mass of the $\mu\mu$ pair

  - MultiVariate discriminant variable combining geometrical and kinematical information about the event: "Geometrical Likelihood" (GL)

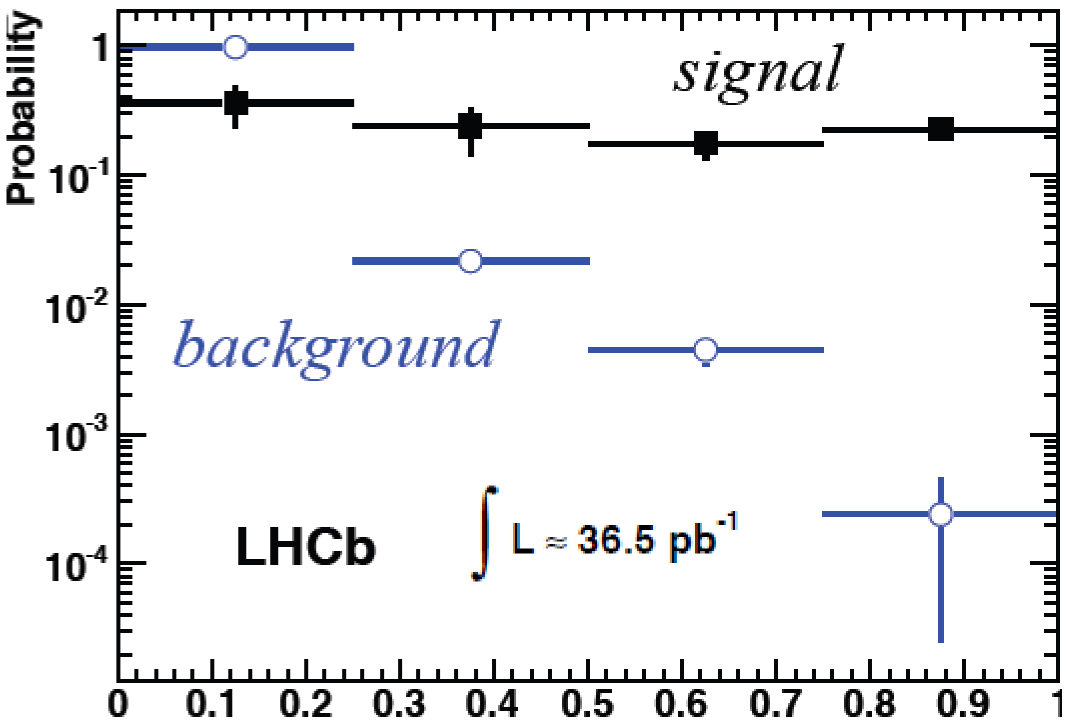
    - Flat distributed for signal, background peaks at 0



- Control channels to get signal and background expectations w/o relying on simulation

- Compare expectations with observed distribution. Results combined using CL_s method

Geometrical Likelihood



- S-B separation relies strongly on this variable
- Trained using MC samples of $B_s \rightarrow \mu\mu$ signal and bb $\rightarrow \mu\mu$ background.
- Distributions taken from data to not rely on the accuracy of the simulation

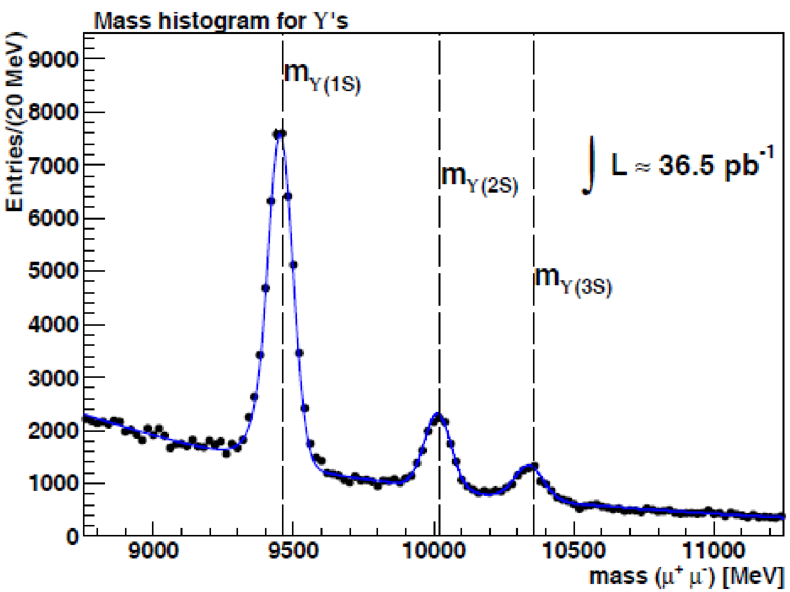
- Distribution of real signal obtained by looking at B $\rightarrow h^+h^-$ in real data. Similar to MC expectation.
- Background distribution is obtained from data by interpolating from mass sidebands in GL bins

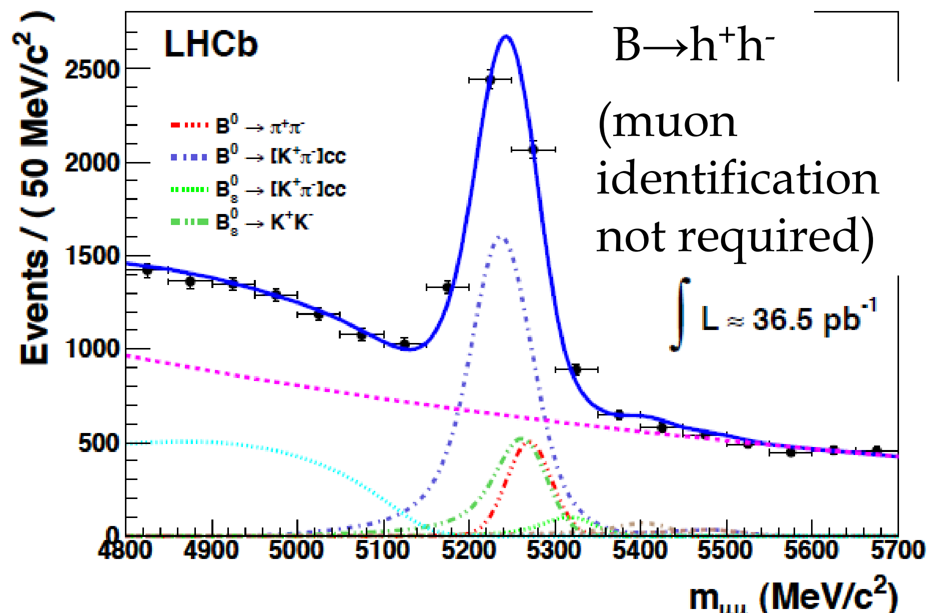
Invariant Mass

- Signal distribution depends on the actual mass resolution of LHCb in the B mass region (resolution depends on mass, almost linearly)

- Measured in data by **interpolating from dimuon resonances** (J/ψ (m<mB) , Υ (m>mB)…) **and** looking at $\mathbf{B \rightarrow h^+h^-}$ ($B_{d,s} \rightarrow K^+\pi^-$, $B_d \rightarrow \pi^+\pi^-$, $B_s \rightarrow K^+K^-$)

- $\mu\mu$ background yield in mass bins is interpolated from mass sidebands





$B \rightarrow h^+h^-$

(muon identification not required)

$\sigma = 26.71 \pm 0.95$ MeV/c^2

Normalization

$$\mathrm{BR} = \mathrm{BR}_{\mathrm{cal}} \times \frac{\epsilon^{\mathrm{REC}}_{\mathrm{cal}}\,\epsilon^{\mathrm{SEL|REC}}_{\mathrm{cal}}\,\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{cal}}}{\epsilon^{\mathrm{REC}}_{\mathrm{sig}}\,\epsilon^{\mathrm{SEL|REC}}_{\mathrm{sig}}\,\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{sig}}} \times \frac{f_{\mathrm{cal}}}{f_{B^0_s}} \times \frac{N_{B^0_s\to\mu^+\mu^-}}{N_{\mathrm{cal}}} = \alpha \times N_{B^0_s\to\mu^+\mu^-}$$

Ratio of probabilities of b quark to hadronize into the different mesons.
$f_d/f_+ = 1$, $f_d/f_s = 3.71 \pm 0.47$ (from HFAG)

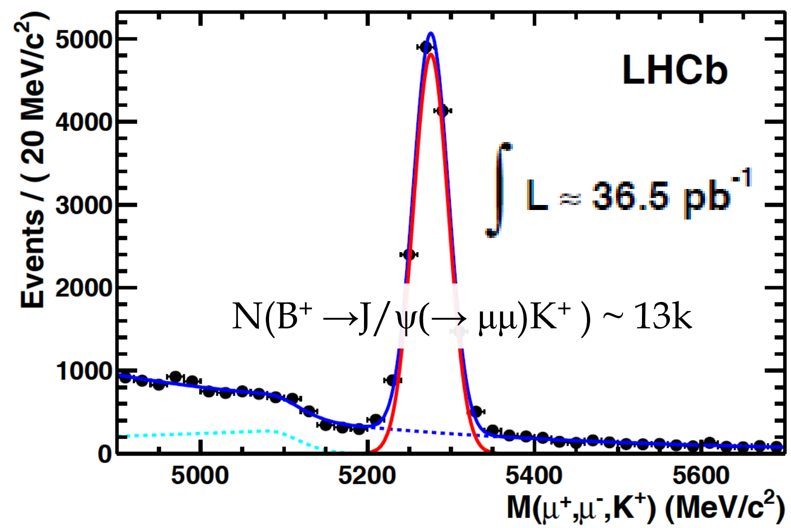
- Three channels are used, each one with different (dis)advantages:

  - $B^+ \to J/\psi(\to \mu\mu)K^+$:

    - Similar trigger (muon triggers) to the signal, similar particle identif.

    - Well known BR, but is B^+ and not B_s → ~13% systematic for $B_s \to \mu\mu$

    - Different number of tracks in the final state



Normalization factors:

$\alpha(B_d) = (2.27\ \ 0.18)\mathrm{x}10^{-9}$
$\alpha(B_s) = (8.2\ \ 1.3)\mathrm{x}10^{-9}$

Normalization

$$\mathrm{BR} = \mathrm{BR}_{\mathrm{cal}} \times \frac{\epsilon^{\mathrm{REC}}_{\mathrm{cal}}\,\epsilon^{\mathrm{SEL|REC}}_{\mathrm{cal}}\,\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{cal}}}{\epsilon^{\mathrm{REC}}_{\mathrm{sig}}\,\epsilon^{\mathrm{SEL|REC}}_{\mathrm{sig}}\,\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{sig}}} \times \frac{f_{\mathrm{cal}}}{f_{B^0_s}} \times \frac{N_{B^0_s\to\mu^+\mu^-}}{N_{\mathrm{cal}}} = \alpha \times N_{B^0_s\to\mu^+\mu^-}$$

Ratio of probabilities of b quark to hadronize into the different mesons.
$f_d/f_+ = 1$, $f_d/f_s = 3.71 \pm 0.47$ (from HFAG)

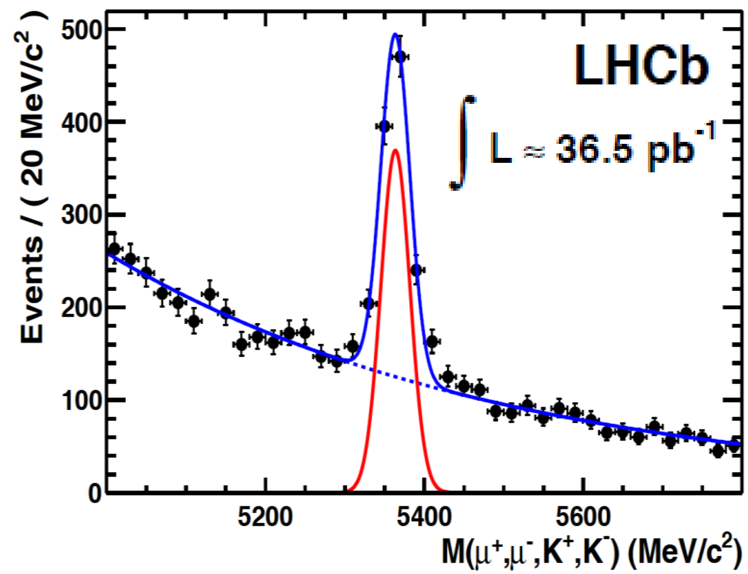
- Three channels are used, each one with different (dis)advantages:

  - $B_s \to J/\psi(\to \mu\mu)\varphi\ (\to K^+K^-)$:

    - Similar trigger (muon triggers) to the signal, similar particle identif.

    - It's a B_s, but BR known only with 26% precision

    - Different number of tracks in the final state



Normalization factors:

$\alpha(B_d) = (2.92\ \ 0.91)\mathrm{x}10^{-9}$
$\alpha(B_s) = (10.5\ \ 2.9)\mathrm{x}10^{-9}$

Normalization

$$\mathrm{BR} = \mathrm{BR}_{\mathrm{cal}} \times \frac{\epsilon^{\mathrm{REC}}_{\mathrm{cal}}\,\epsilon^{\mathrm{SEL|REC}}_{\mathrm{cal}}\,\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{cal}}}{\epsilon^{\mathrm{REC}}_{\mathrm{sig}}\,\epsilon^{\mathrm{SEL|REC}}_{\mathrm{sig}}\,\epsilon^{\mathrm{TRIG|SEL}}_{\mathrm{sig}}} \times \frac{f_{\mathrm{cal}}}{f_{B^0_s}} \times \frac{N_{B^0_s\to\mu^+\mu^-}}{N_{\mathrm{cal}}} = \alpha \times N_{B^0_s\to\mu^+\mu^-}$$

Ratio of probabilities of b quark to hadronize into the different mesons.
$f_d/f_+ = 1$, $f_d/f_s = 3.71 \pm 0.47$ (from HFAG)

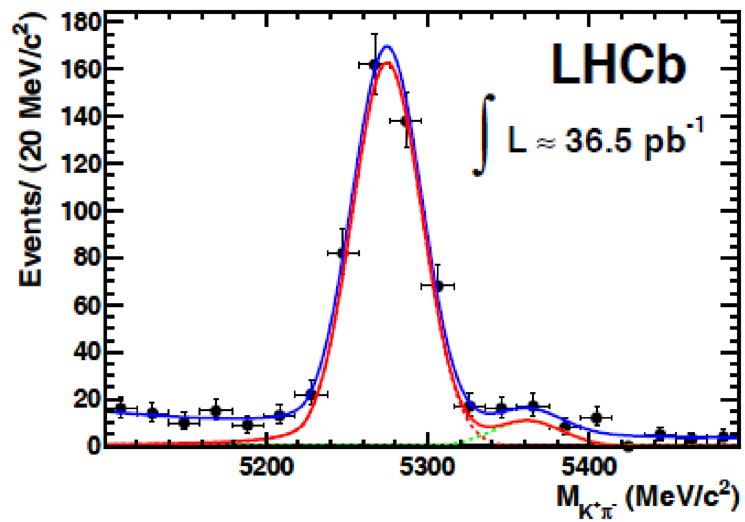
- Three channels are used, each one with different (dis)advantages:

  - $B_d \to K^+\pi^-$

    - Different trigger (used triggered on the underlying event/other b used)

    - Same kinematics, number of tracks in final state

    - Well known BR, but is B_d and not B_s → ~13% systematic for $B_s\to\mu\mu$

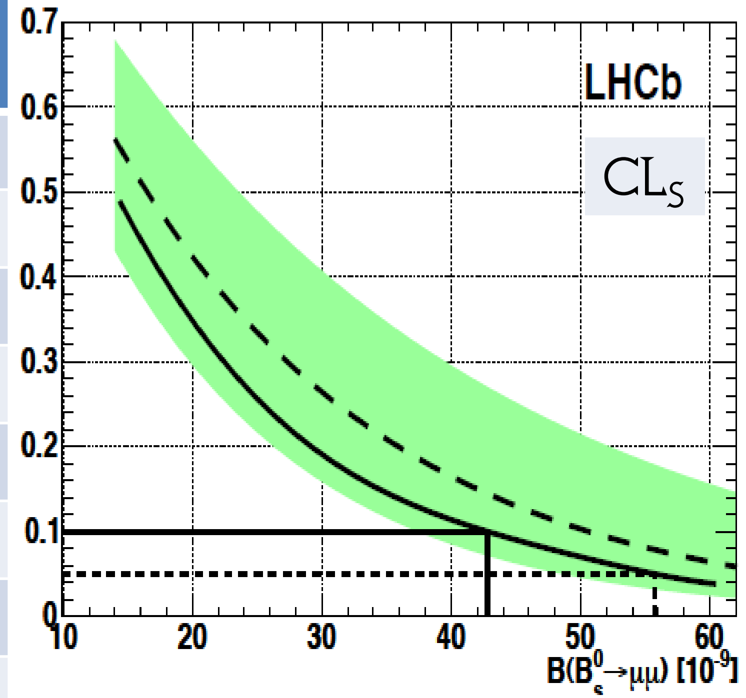


Normalization factors:

$\alpha(B_d) = (1.99\ \ 0.40)\mathrm{x}10^{-9}$
$\alpha(B_s) = (7.1\ \ 1.7)\mathrm{x}10^{-9}$

Observed pattern and Result (B_s)

B mass… (MeV/c2)	GL bin 1	GL bin 2	GL bin 3	GL bin 4
[-60, -40]	**39**	**2**	**1**	**0**
[-40,-20]	**55**	**2**	**0**	**0**
[-20,0]	**73**	**0**	**0**	**0**
[0,20]	**60**	**0**	**0**	**0**
[20,40]	**53**	**2**	**0**	**0**
[40,60]	**55**	**1**	**0**	**0**
(TOTAL)	*335*	*7*	*1*	*0*
(bkg exp.)	*329*	*7.36*	*1.51*	*0.081*



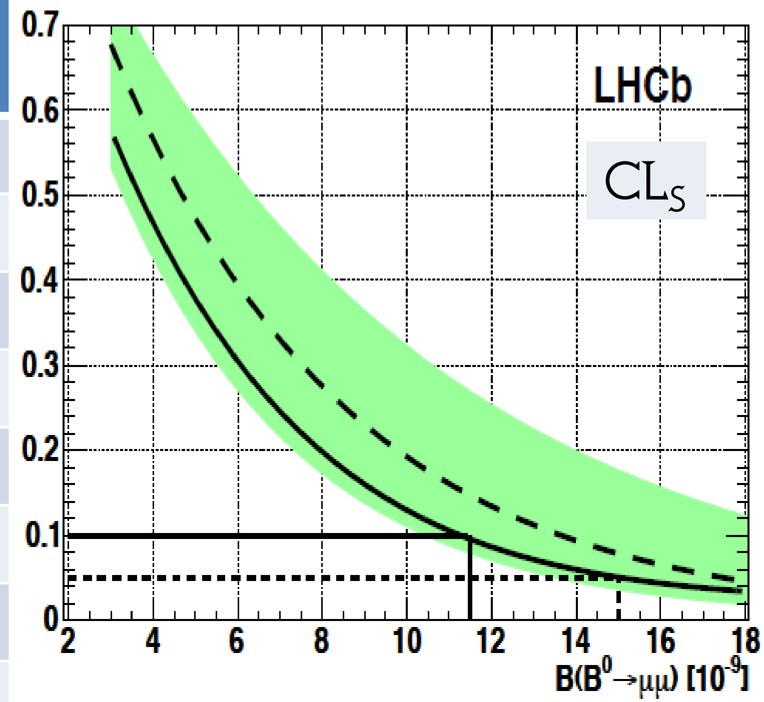
$BR(B_s \rightarrow \mu\mu) < 4.3\ (5.6)\ 10^{-8}$ @ 90 (95% CL)

$BR(B_s \rightarrow \mu\mu) < 4.3\times10^{-8}$ @ 95% CL(CDF, prelim)
$BR(B_s \rightarrow \mu\mu) < 5.1\times10^{-8}$ @ 95% CL(D0, publish.)

Expected are: 5.1 (6.5)

Observed pattern and Result (B_d)

B mass… (MeV/c2)	GL bin 1	GL bin 2	GL bin 3	GL bin 4
[-60, -40]	59	2	0	0
[-40,-20]	67	0	0	0
[-20,0]	56	2	0	0
[0,20]	60	0	0	0
[20,40]	42	2	1	0
[40,60]	49	2	0	0
(TOTAL)	*333*	*8*	*1*	*0*
(bkg exp.)	*352*	*8.29*	*1.85*	*0.118*

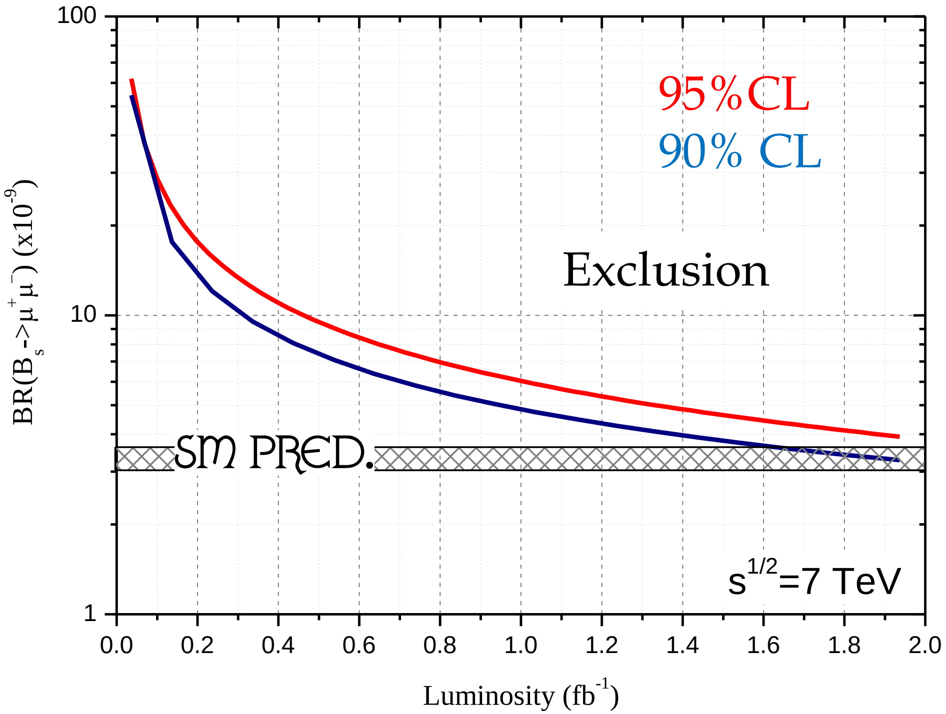


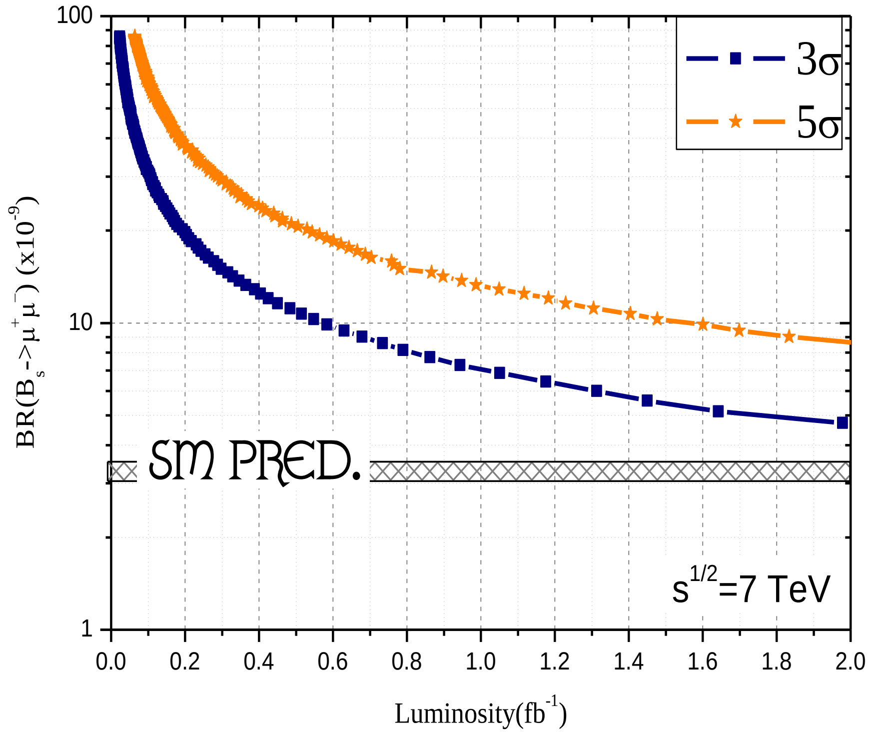
$BR(B_d \to \mu\mu) < 1.2\ (1.5)\ 10^{-8}$ @ 90 (95% CL)

$BR(B_s \to \mu\mu) < 0.76 \times 10^{-8}$ @ 95% CL(CDF, prelim)

Expected are: 1.4 (1.8)

Extrapolated sensitivity





LHCb can provide **VERY** interesting results in one year from now!

Conclusions

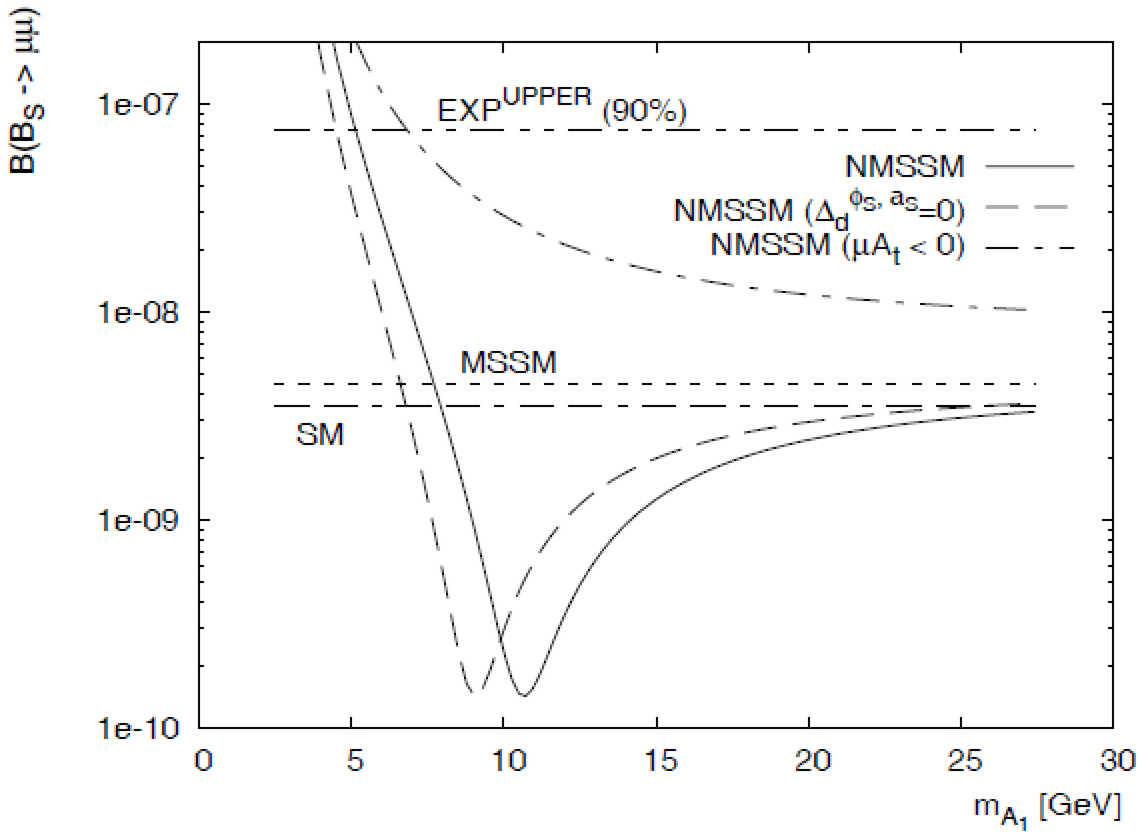
- $B_{s,d} \to \mu\mu$ is an interesting probe of physics beyond the Standard Model
- First LHCb result on BR($B_{s,d} \to \mu\mu$)

$$BR(B_s \to \mu\mu) < 4.3\ (5.6)\ \ 10^{-8} \text{ @ } 90\ (95\%\ CL)$$

$$BR(B_d \to \mu\mu) < 1.2\ (1.5)\ \ 10^{-8} \text{ @ } 90\ (95\%\ CL)$$

- Those are comparable with current best ones
- Extrapolation to 2 fb^{-1} shows that LHCb can find/exclude BR($B_{s,d} \to \mu\mu$) from $\sim 10^{-8}$ to quite close to SM prediction

Backup

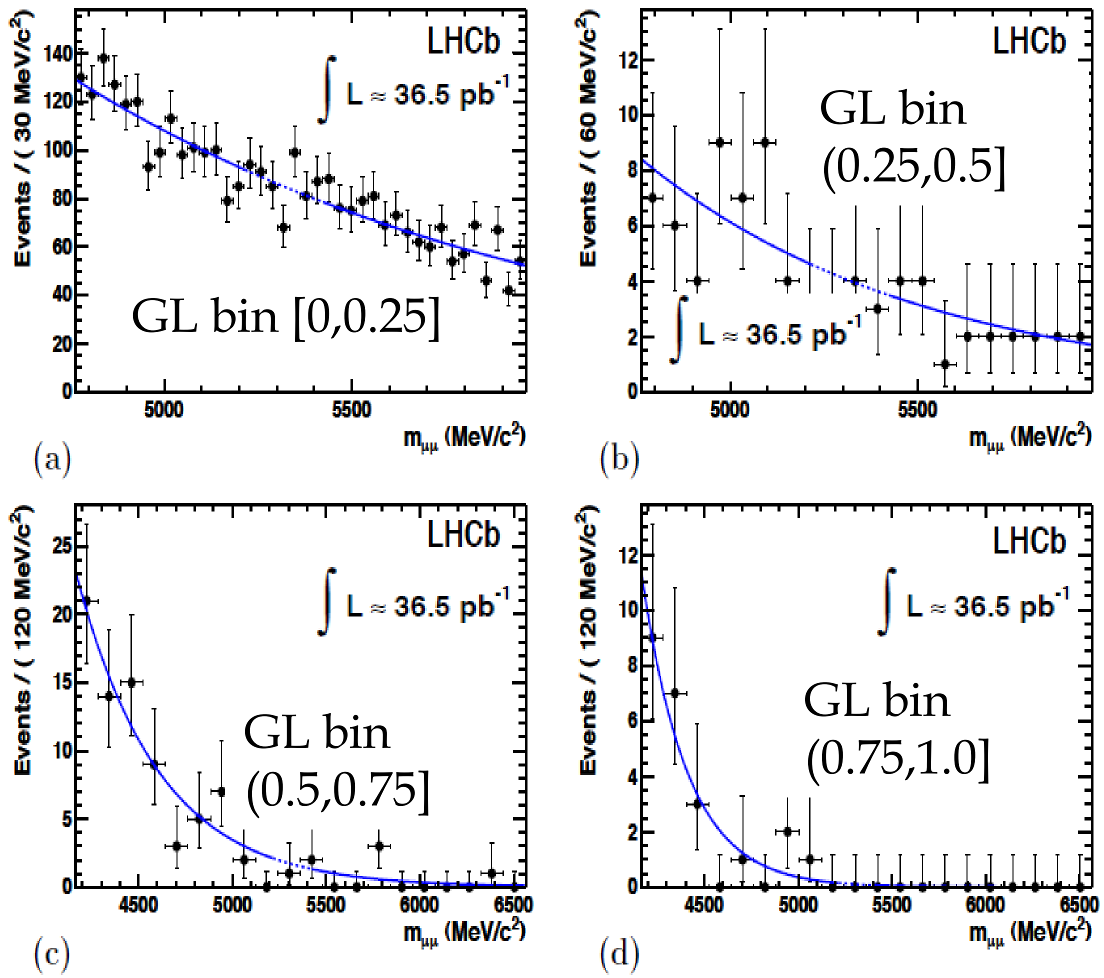


B(B_S -> μμ)
1e-07
1e-08
1e-09
1e-10
0
5
10
15
20
25
30
m_A1 [GeV]
EXP^UPPER (90%)
NMSSM
NMSSM (Δ_d^φS, aS=0)
NMSSM (μA_t < 0)
MSSM
SM

Background yield

- Interpolation in the 4 GL bins gives in "one shot" the 2D distribution GL vs Mass in the search window (60 MeV around the B mass)

- Peaking background ($B \rightarrow h^{+}h^{-}$ wrongly identified as muons) negligible for current amount of data



$B_s \rightarrow \mu^+\mu^-$ and $B \rightarrow h^+h^-$ $B \rightarrow J/\psi(\mu^+\mu^-)X$

Bs2MuMuNoMuID			Detached J/ψ		
	Cut	value		cut	value
μ / h	track χ^2/ndf IPS	<5 >5	μ	track χ^2/ndf IPS ISMUON	<5 >5 true
$B_{(s)}$	Δm DOCA vertex χ^2 VDS IPS	<600 MeV/c^2 <0.3 mm <15 >15 < 5	J/ψ	Δm DOCA vertex χ^2 VDS	<100 MeV/c^2 <0.3 mm <15 >13

We checked the stability of the results with different fit parameterizations
Single exponential, double exponential, linear fit

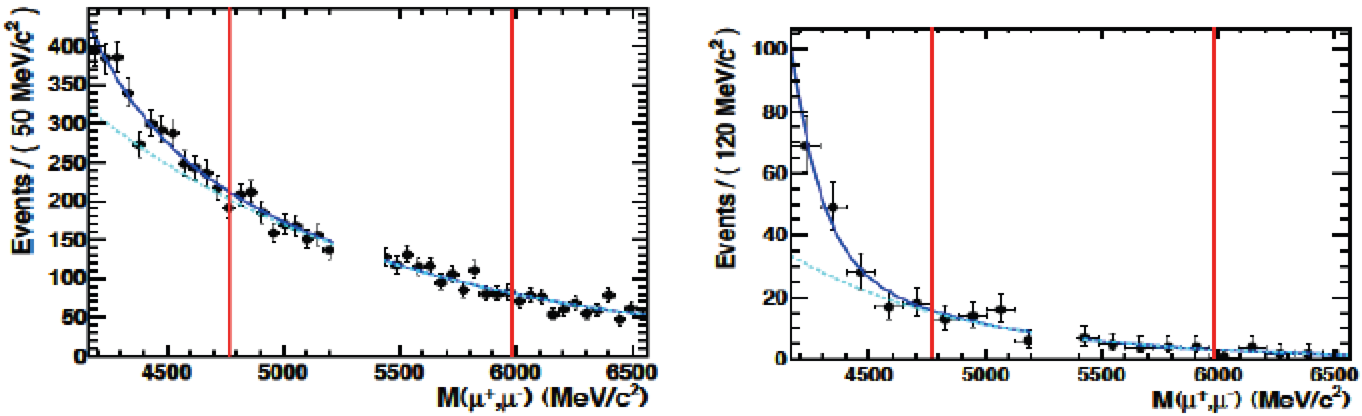


Table 45: Number of background events predicted by exponential, double exponential and linear interpolation in each GL_{KS} bin.

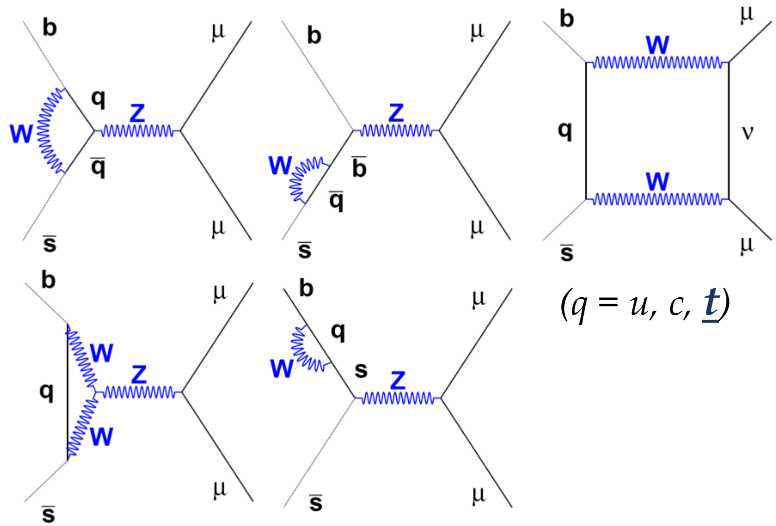
GL_{KS} bin	exp.	double exp.	linear in $5215.1-\Delta M$, $5423.1+\Delta M$ (MeV/c^2)			
			$\Delta M = 600$	$\Delta M = 300$	$\Delta M = 200$	$\Delta M = 100$
$0.0-0.25$	590 ± 11	580 ± 12	625 ± 10	596 ± 14	599 ± 19	582 ± 25
$0.25-0.5$	13.6 ± 1.7	12.8 ± 1.4	15.1 ± 1.6	15.6 ± 2.3	14.6 ± 2.8	$13.5^{+4.9}_{-3.6}$
$0.5-0.75$	$2.91^{+0.74}_{-0.64}$	$3.02^{+0.75}_{-0.66}$	3.98 ± 0.83	$2.8^{+1.4}_{-1.0}$	$2.1^{+1.6}_{-1.0}$	$2.1^{+2.7}_{-1.4}$
$0.75-1.0$	$0.17^{+0.20}_{-0.15}$	$0.31^{+0.31}_{-0.25}$	$0.69^{+0.54}_{-0.33}$	$0.69^{+0.90}_{-0.45}$	$0.5^{+1.2}_{-0.4}$	$0.0^{+1.9}_{-0.0}$

Decay Physics in SM

Branching Ratio (BR) as a function of Wilson Coefficients ("effective" theory) is:

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

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



($q = u, c, \underline{t}$)

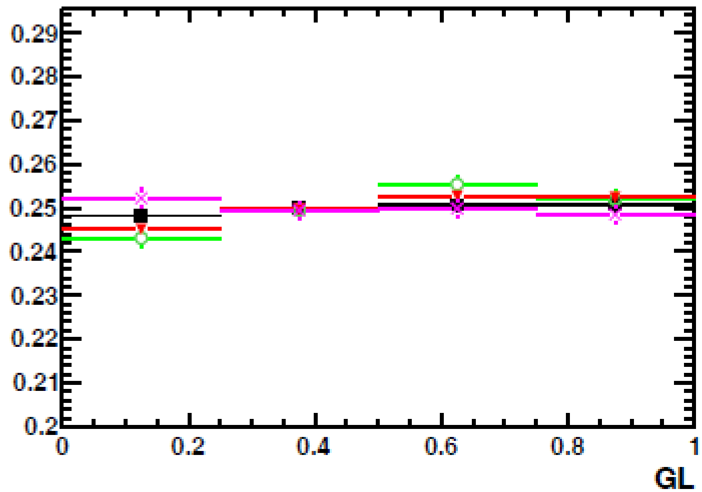
$C_{S,P} \rightarrow$ scalar and pseudo scalar are negligible in SM
C_{10} gives the only relevant contribution

This decay is very suppressed in SM (BR very small, but precisely predicted):

$BR(B_s \rightarrow \mu\mu) = (3.2 \pm 0.2)\text{x}10^{-9}$ **$BR(B_d \rightarrow \mu\mu) = (1.0 \pm 0.1)\text{x}10^{-10}$**

Experimental upper limit still one order of magnitude above such values. @ 95% CL:

$BR(B_s \rightarrow \mu\mu) < 4.3\text{x}10^{-8}$ $BR(B_d \rightarrow \mu\mu) < 0.76\text{x}10^{-8}$ (**CDF**, 3.7 fb-1, prelim.)
$BR(B_s \rightarrow \mu\mu) < 5.1\text{x}10^{-8}$ (**D0**, 6.1 fb-1, publ.)



$$N_{B^0_{(s)}\to h^+h^-}/N_{B^0_{(s)}\to\mu^+\mu^-} - 1$$

bin	$B^0 \to \pi^+\pi^-$	$B^0 \to K^+\pi^-$	$B^0_s \to K^+K^-$
$0.0 - 0.25$	$(-2.8 \pm 1.1)\%$	$(-1.1 \pm 0.9)\%$	$(0.7 \pm 1.1)\%$
$0.25 - 0.5$	$(0.8 \pm 1.1)\%$	$(0.4 \pm 0.9)\%$	$(0.3 \pm 1.1)\%$
$0.50 - 0.75$	$(2.4 \pm 1.1)\%$	$(0.4 \pm 0.9)\%$	$(-0.6 \pm 1.1)\%$
$0.75 - 1.0$	$(-0.4 \pm 1.1)\%$	$(0.3 \pm 0.9)\%$	$(-0.4 \pm 1.1)\%$

$\varepsilon^{REC}(B_s \rightarrow \mu^+\mu^-)=10.2\%$

MC estimates for the efficiency ratios:

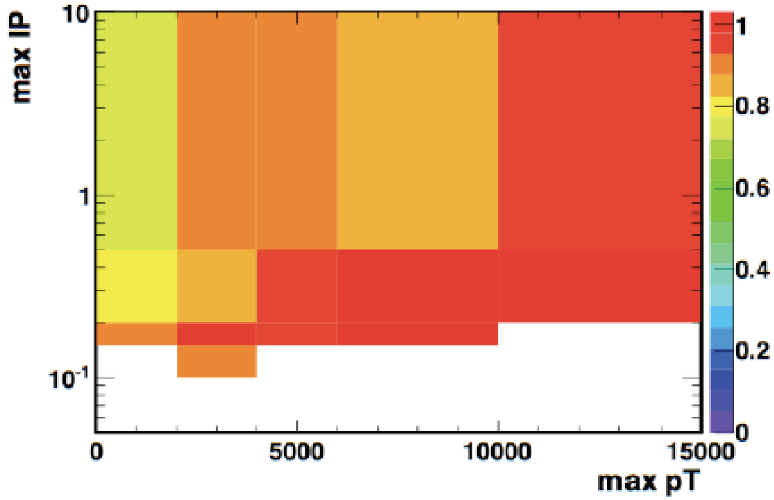
$$\frac{\epsilon^{\mathrm{REC}}_{\mathrm{B^+ \to J/\psi K^+}}}{\epsilon^{\mathrm{REC}}_{\mathrm{B^0_s \to \mu^+\mu^-}}} = 0.58 \pm 0.02 \qquad \frac{\epsilon^{\mathrm{REC}}_{\mathrm{B^0_s \to J/\psi \phi}}}{\epsilon^{\mathrm{REC}}_{\mathrm{B^0_s \to \mu^+\mu^-}}} = 0.39 \pm 0.03 \qquad \frac{\epsilon^{\mathrm{REC}}_{\mathrm{B^0_{(s)} \to h^+h^-}}}{\epsilon^{\mathrm{REC}}_{\mathrm{B^0_s \to \mu^+\mu^-}}} = 0.75 \pm 0.05$$

3body/2body 4body/2body 2body/2body

1) From $\varepsilon^{TRIG} = \varepsilon^{TIS} * N^{TRIG}/N^{TIS}$, we obtain

$$\epsilon_{Trg}^{J/\psi} = (85.9 \pm 0.9_{\rm stat} \pm 2.0_{\rm syst})\%$$

for $B \to J/\psi(\mu^+\mu^-)$ X events



2) Use the same events to produce a trigger efficiency map (p_T, IP); convolution with the harder muon spectra of $B_s \to \mu\mu$ gives:

$$\epsilon_{Trg}^{B_s^0 \to \mu^+\mu^-} = (89.9 \pm 0.8_{\rm stat} \pm 4.0_{\rm syst})\%$$

The $B \to J/\psi X / B_s \to \mu\mu$ ratio of efficiencies is then

$$\boxed{\frac{\epsilon_{\rm cal}}{\epsilon_{\rm sig}} = (95.6 \pm 1.3_{stat} \pm 4.8_{\rm syst})\%}$$

syst by comparing, on MC, true efficiency with method

-$B^0 \to K^+\pi^-$ decay mode known with 3.1%, but after stripping we're left with an inclusive sample of $B_{(s)} \to hh$ sample (no PID cuts) :

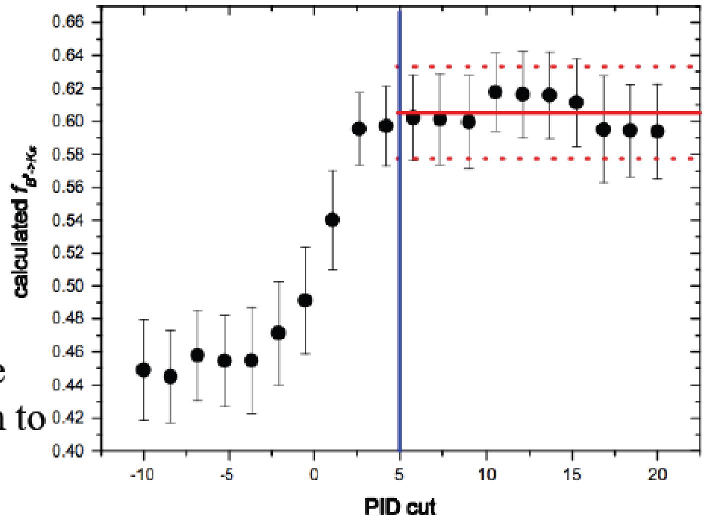
→need to know the fraction of $f_{B0 \to K+\pi-}$

-fit the number of each exclusive mode after tight PID cuts

-RICH efficiency factors cancel out when applying constraints on

$BR(B_d \to K\pi)/BR(B_d \to \pi\pi)$

- $f_{B0 \to K+\pi-}$ is obtained from a combination of the observed yields, when PID cuts are tight enough to guarantee negligible contamination



$f_{B0 \to K+\pi-} = 0.605 \pm 0027$

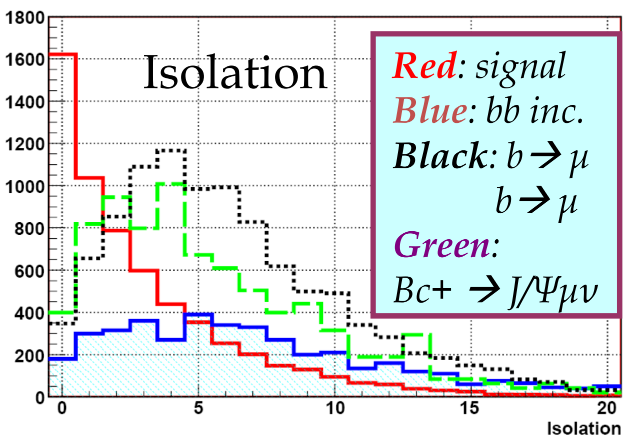
50

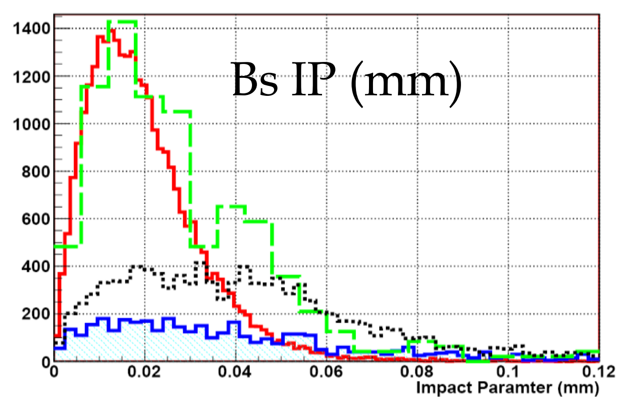
Geometrical Likelihood

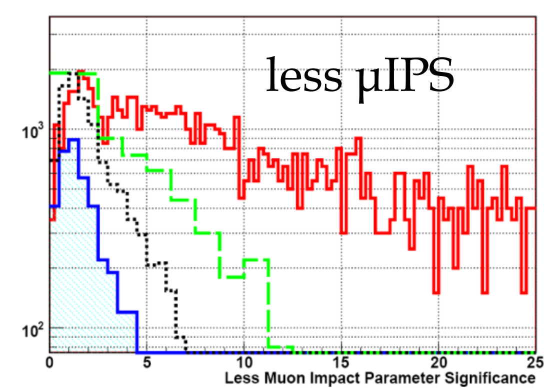
How the Geometry likelihood is built:

1. Input variables: min Impact Parameter Significance (μ^+,μ^-), DOCA, Impact Parameter of B, lifetime, iso - μ^+, iso- μ^-

- **Isolation:** Idea: muons making fake Bs→μμ might came from another SV's → For each muon; remove the other μ and look at the rest of the event: How many good - SV's (forward, DOCA, pointing) can it make? The precise criteria used is inherited from Hlt Generic



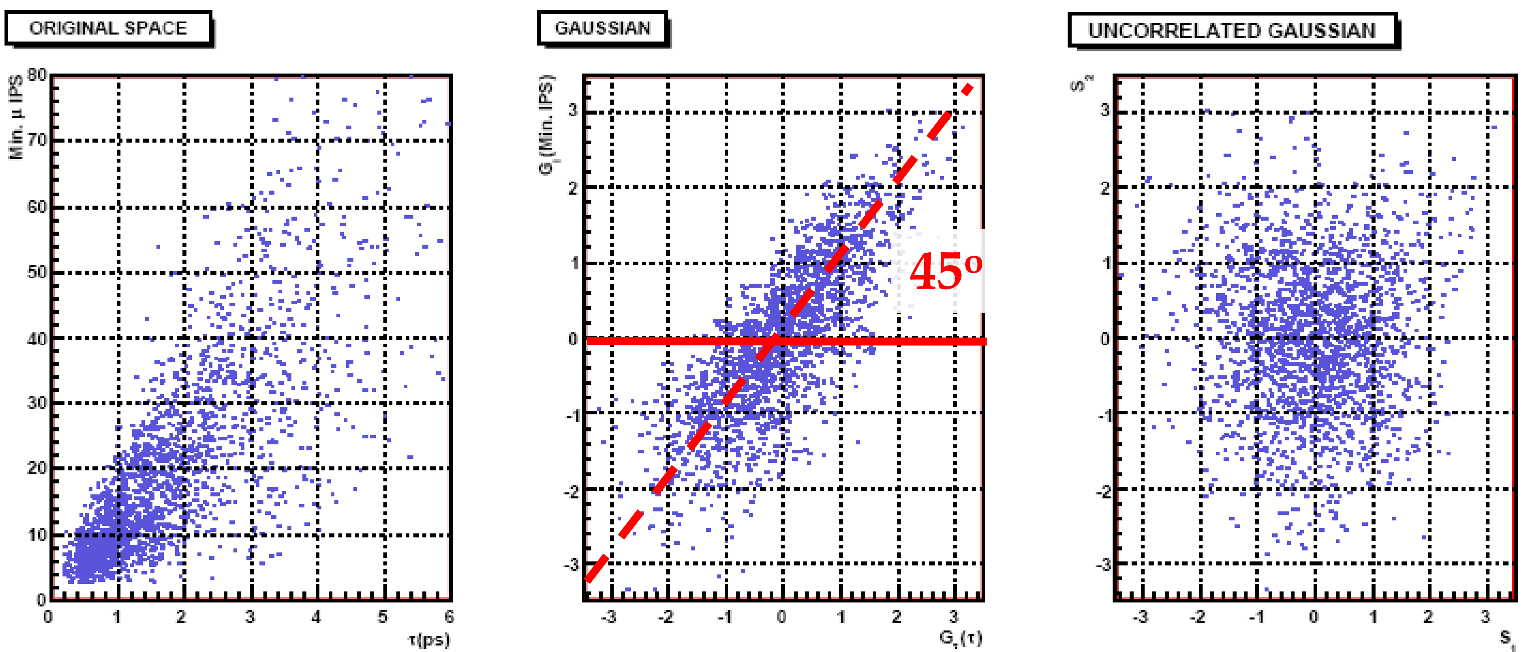




Geometrical Likelihood

How the Geometry likelihood is built:

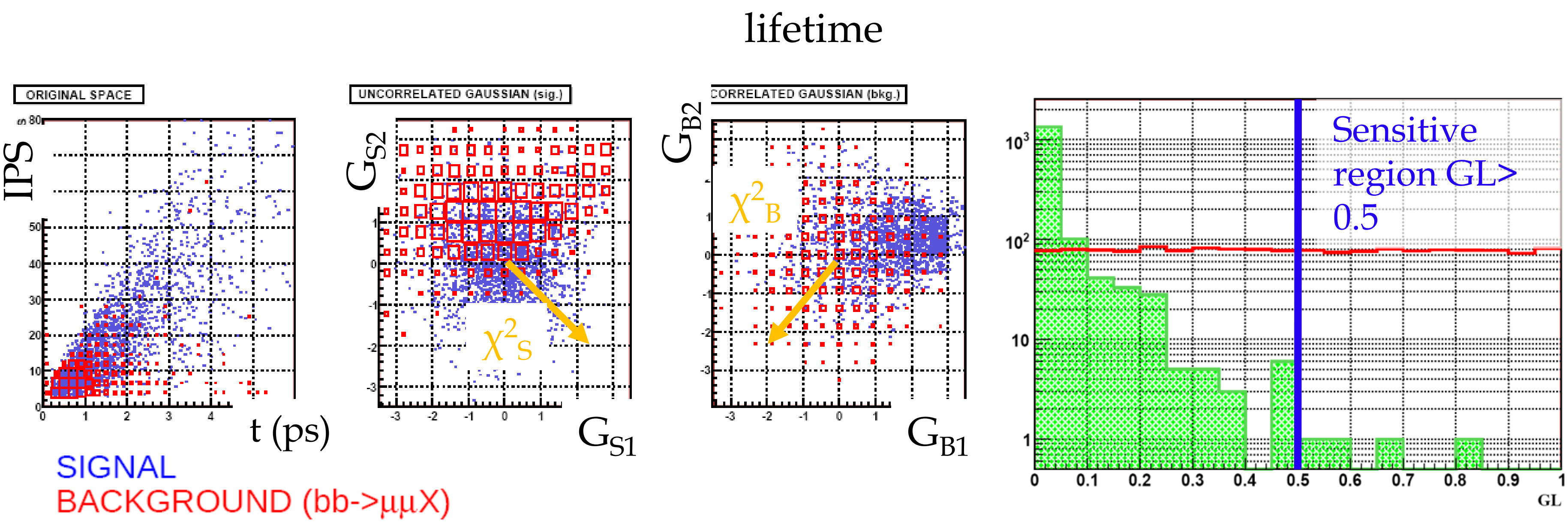
1. Input variables: min Impact Parameter Significance (μ^+,μ^-), DOCA, Impact Parameter of B, lifetime, iso - μ^+, iso- μ^-
2. They are transformed to Gaussian through cumulative and inverse error function
3. In such space correlations are more linear-like → rotation matrix, and repeat 2



Geometrical Likelihood

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.



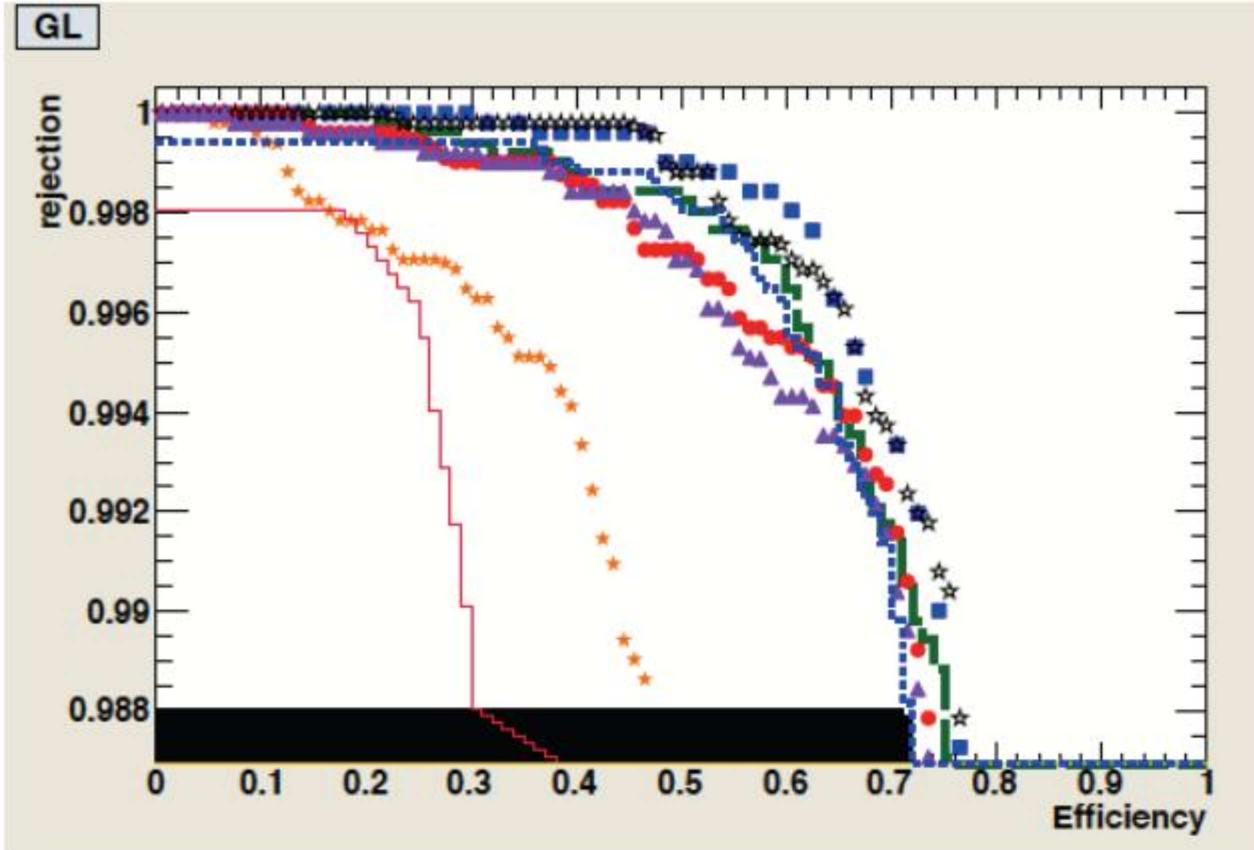
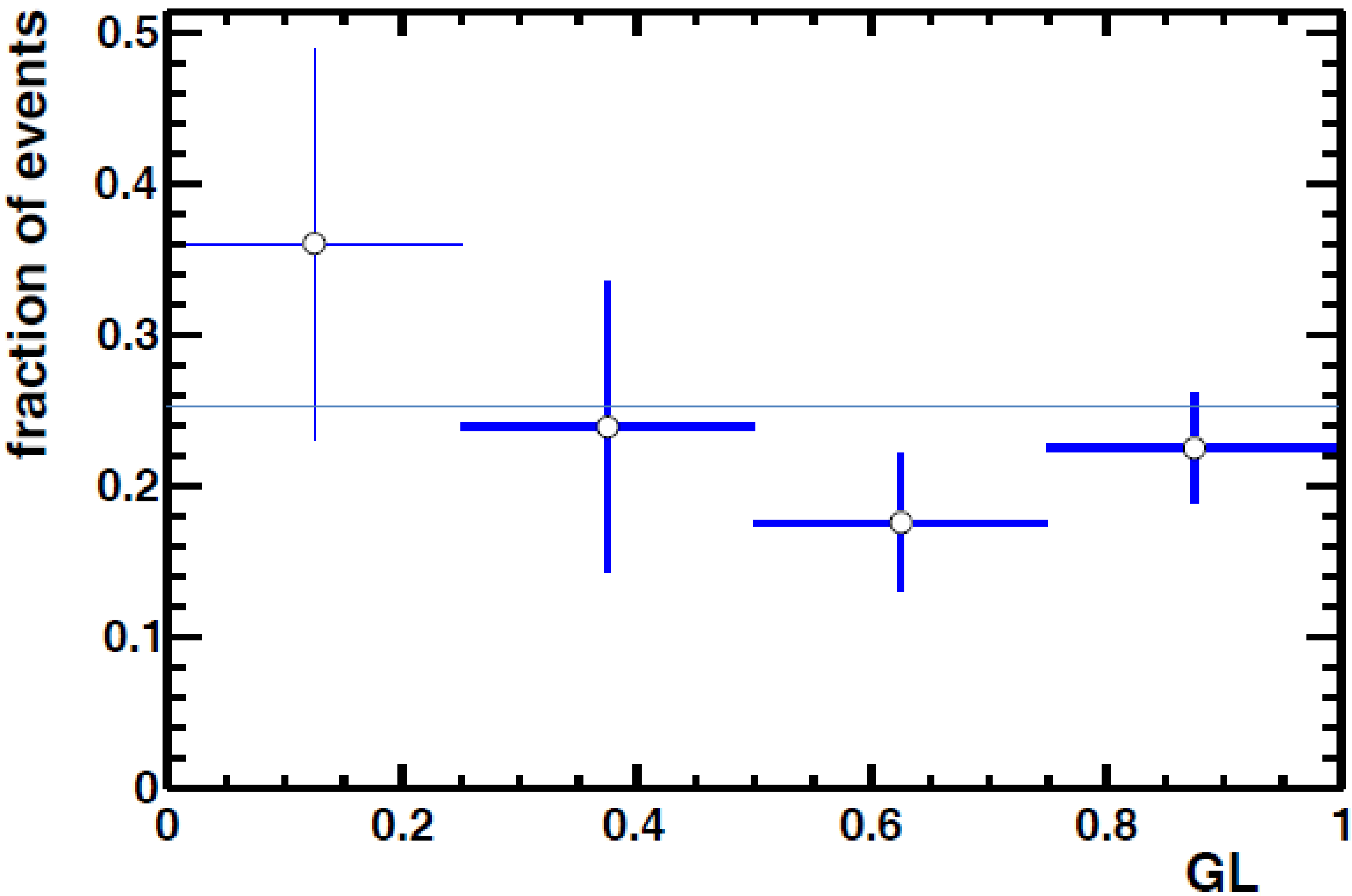


Figure 23: Performance of GL_K and set of other multivariate methods. The X axis shows the efficiency, and the Y axis the rejection. Blue squares: GL_K, Open stars: BDT. Short Dashed: PDERS. Violet triangles: Fisher Discriminant. Red cycles: Best performing NN. Green dashed line: Support Vector Machine.Red solid line: RuleFit. Orange stars: FDA. Black filled histogram: kNN.

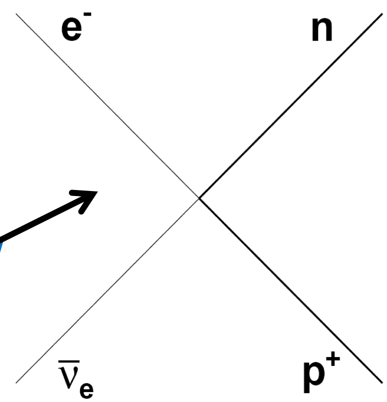

fraction of events
0.5
0.4
0.3
0.2
0.1
0
0
0.2
0.4
0.6
0.8
1
GL

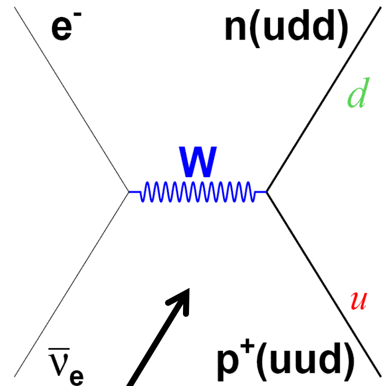
Wilson coefficients

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





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

underlying "fundamental" theory (SM)

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

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

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

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

Computing CLs

Reference: Thomas Junk, CERN-EP/99-041. 01 March 1999
(Used at LEP for Higgs searches)

For each bin:

si = expected signal events in bin
bi = expected bkg. events in bin
di = measured events in bin

$$X_i = \frac{Poisson(d_i, <d_i> = s_i + b_i)}{Poisson(d_i, <d_i> = b_i)}$$

For a configuration {Xi}: $X = \prod_i^N X_i$ (it is a binned likelihood ratio)

$$CL_{s+b} = P_{s+b}(X \le X^{OBSERVED})$$

$$CL_b = P_b(X \le X^{OBSERVED})$$

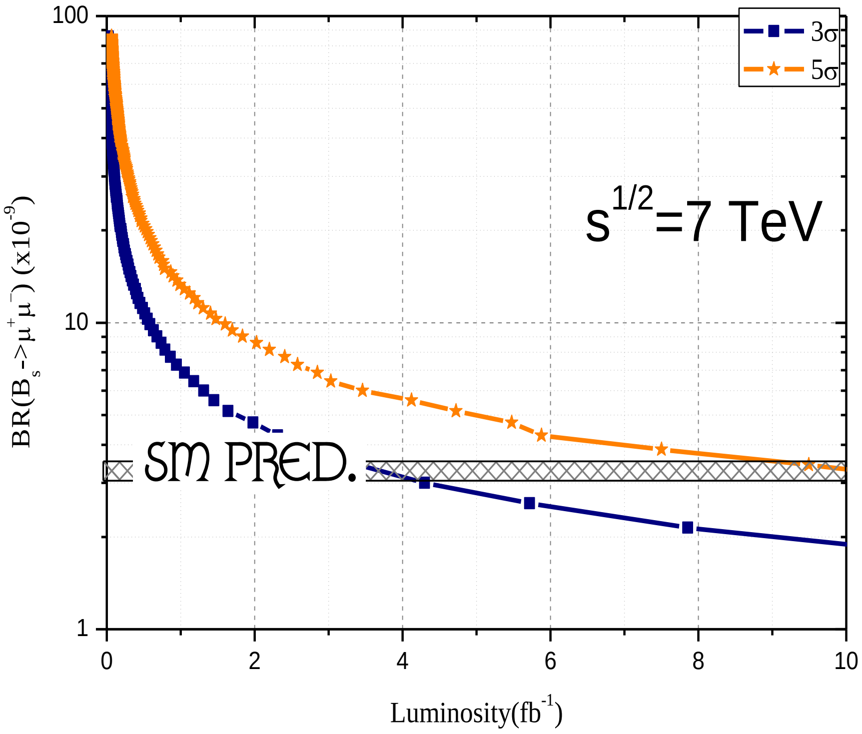
CLs = CLs+b/CLb

•High CLb → observed excess w.r.t bkg expectation→ signal (CLb>0.9973 → 3 sigma)

•Small CLs → too few events w.r.t prediction from signal hypothesis

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

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



100
10
1
0
2
4
6
8
10
3σ
5σ
s^1/2=7 TeV
SM PRED.
BR(B_s->μ^+μ^-) (x10^-9)
Luminosity(fb^-1)