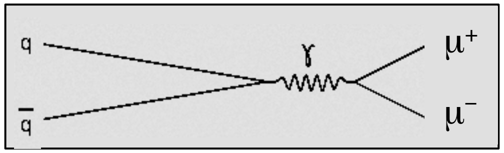
Sheldon Stone
Aug. 20, 2014

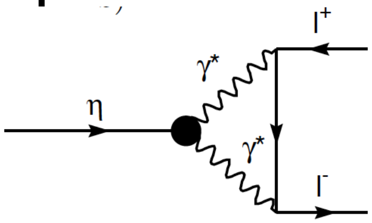
$B^0_{(s)} \rightarrow \mu^+\mu^-$

$P^0 \rightarrow \mu^+\mu^-$

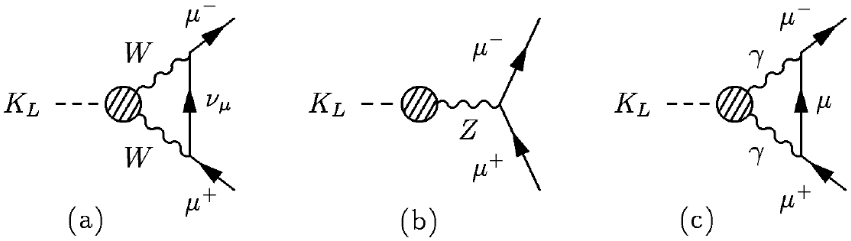
- What mesons do you know that decay into $\mu^+\mu^-$?
  - Spin-1 mesons formed of $q\bar{q}$, including ρ, ω, ϕ, ψ, Y…



  - Spin-0 mesons η, K^0_L, (note helicity supression)

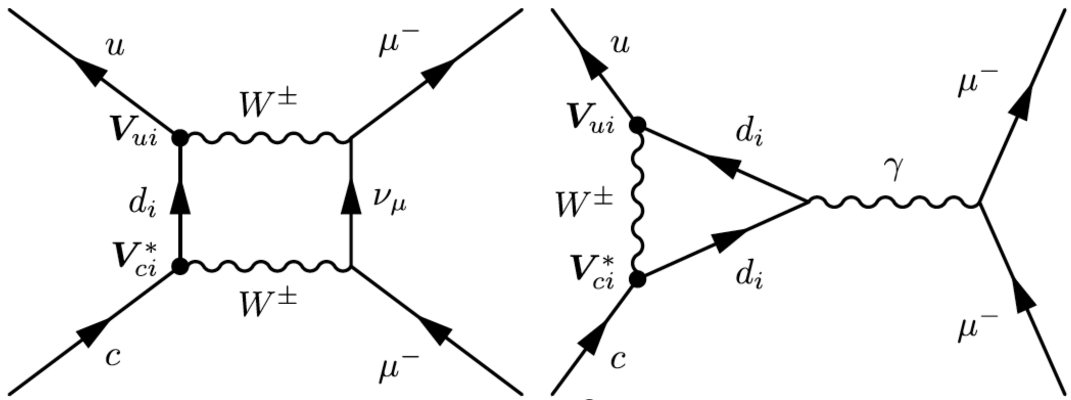


  - similar diagram for K_L, γ diagram dominates



$D^0 \to \mu^+\mu^-$

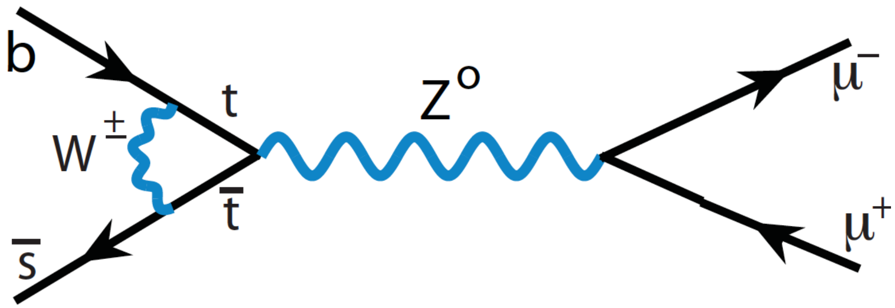
- The 2γ intermediate decay is highly suppressed ~few x 10^{-13} (hep-ph/0112235)
- Short distance diagrams are very small ~10^{-18}

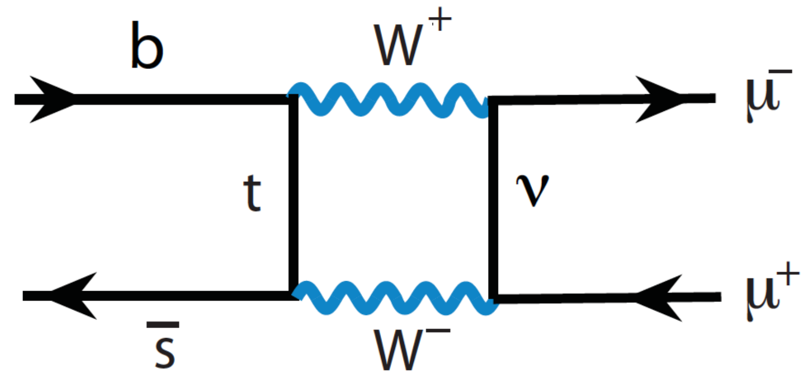


- Experimental limit: $\mathcal{B}(D^0 \to \mu^+\mu^-) < 6.2 \times 10^{-9}$ (LHCb arXiv:1305.5059)
- Good place to search for New Physics, but experimentally difficult; why?

SM Theory $B^0_{(s)} \to \ell^+\ell^-$

- Long distance 2γ contribution vanishes
- Short distance decay diagrams





- Predictions are

$$\overline{\mathcal{B}}[B_s \to \bar{e}e] = (8.54 \pm 0.55) \times 10^{-14}\,, \qquad \overline{\mathcal{B}}[B_d \to \bar{e}e] = (2.48 \pm 0.21) \times 10^{-15}\,,$$
$$\overline{\mathcal{B}}[B_s \to \bar{\mu}\mu] = (3.65 \pm 0.23) \times 10^{-9}\,, \qquad \overline{\mathcal{B}}[B_d \to \bar{\mu}\mu] = (1.06 \pm 0.09) \times 10^{-10}\,,$$
$$\overline{\mathcal{B}}[B_s \to \bar{\tau}\tau] = (7.73 \pm 0.49) \times 10^{-7}\,, \qquad \overline{\mathcal{B}}[B_d \to \bar{\tau}\tau] = (2.22 \pm 0.19) \times 10^{-8}\,,$$

- Includes NNLO QCD and NLO EW corrections

Questions

$$\overline{\mathcal{B}}[B_s \to \bar{e}e] = (8.54 \pm 0.55) \times 10^{-14}\,, \qquad \overline{\mathcal{B}}[B_d \to \bar{e}e] = (2.48 \pm 0.21) \times 10^{-15}\,,$$

$$\overline{\mathcal{B}}[B_s \to \bar{\mu}\mu] = (3.65 \pm 0.23) \times 10^{-9}\,, \qquad \overline{\mathcal{B}}[B_d \to \bar{\mu}\mu] = (1.06 \pm 0.09) \times 10^{-10}\,,$$

$$\overline{\mathcal{B}}[B_s \to \bar{\tau}\tau] = (7.73 \pm 0.49) \times 10^{-7}\,, \qquad \overline{\mathcal{B}}[B_d \to \bar{\tau}\tau] = (2.22 \pm 0.19) \times 10^{-8}\,,$$

- Why is e^+e^- rate so small?
- Why are the predictions different for the 3 leptons, does this violate lepton universality?
- Why isn't $\tau^+\tau^-$ easier than $\mu^+\mu^-$ as the predicted branching ratio is larger?

Experiment-overview

- Want to measure the branching ratio, the fraction of the time the B goes to $\mu^+\mu^-$
- Need to detect the $\mu^+\mu^-$
- Need to know how many B^0 or B_s we have
- Inclusive b production was measured by LHCb to be ~300 μb at 7x7 TeV
- So in 10^7 sec (1 year of running) at $\mathcal{L}=4\times10^{32}/cm^2\bullet s$, # b's is 10^{12}, (CMS~10x larger) but need to account for B fractions (f_d~1/3, f_s~1/10), acceptance, trigger …

Trigger for $\mu^+\mu^-$

LHCb

- Hardware level: One muon with p_T>1.76 GeV (also a track multiplicity cut), or two muons with $\sqrt{p_{T1}p_{T2}}$>1.6 GeV
- Higher level: Impact Parameter (IP) cut & invariant mass requirement
- Trigger eff ~90%

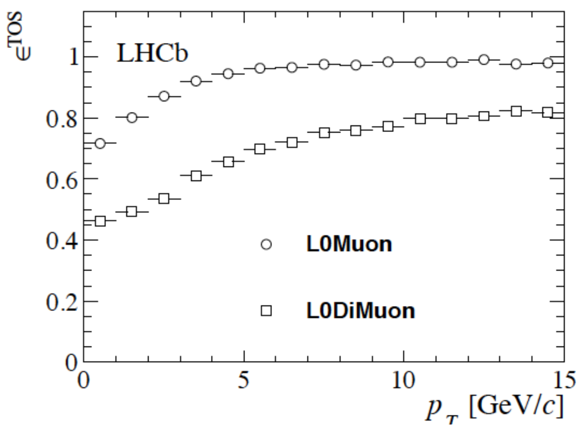
CMS

- Hardware level: Two muon candidates
- Higher level:
  - Dimuon mass cut
  - 7 GeV data: p_T>4 GeV for each muon, p_T(B)>3.9 GeV unless one μ has $|\eta|$>1.5 in which case p_T(B)>5.9 GeV
  - 8 GeV data: small changes
- Trigger efficency lower than for LHCb

Normalization modes

LHCb

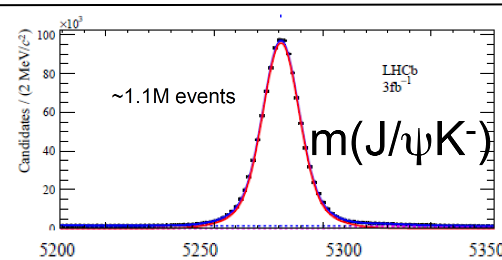
- $B^- \to J/\psi K^-$, $\psi \to \mu^+\mu^-$, similar trigger
- $B^0 \to K^-\pi^+$, same topology, different trigger
- Trigger eff of $B^- \to J/\psi K^-$



Efficiency ε^{TOS} of $B^+ \to J/\psi(\mu^+\mu^-)K^+$ as a function of p_T (J/ψ) for L0Muon and L0DiMuon

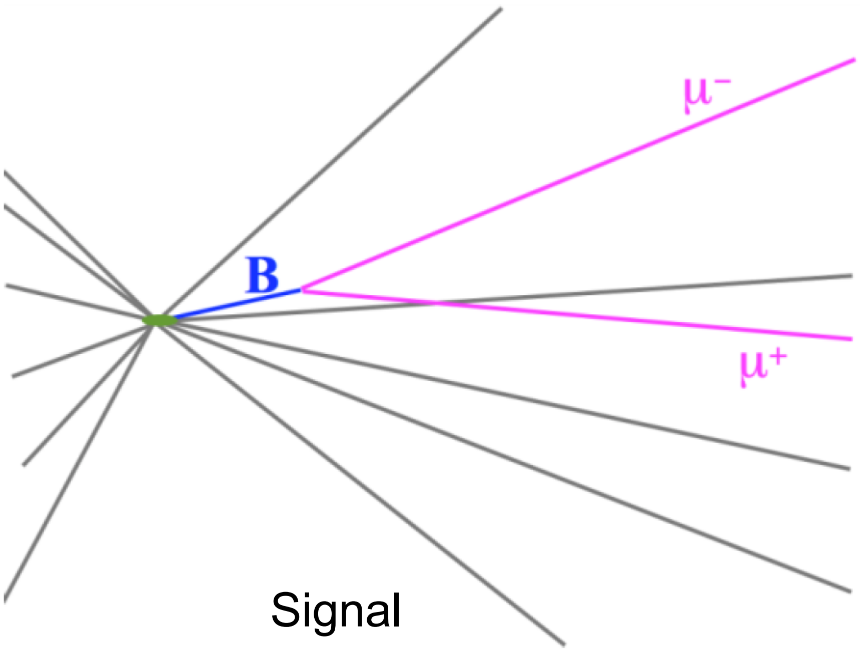
CMS

- $B^- \to J/\psi K^-$, $\psi \to \mu^+\mu^-$
- $B_s \to J/\psi\phi$, $\psi \to \mu^+\mu^-$, $\phi \to K^+K^-$ used for checking simulations
- BDT selection (neural network) – will discuss later, also LHCb
- Overall detection efficiencies for $B^0 \to \mu^+\mu^-$ is about 0.3%

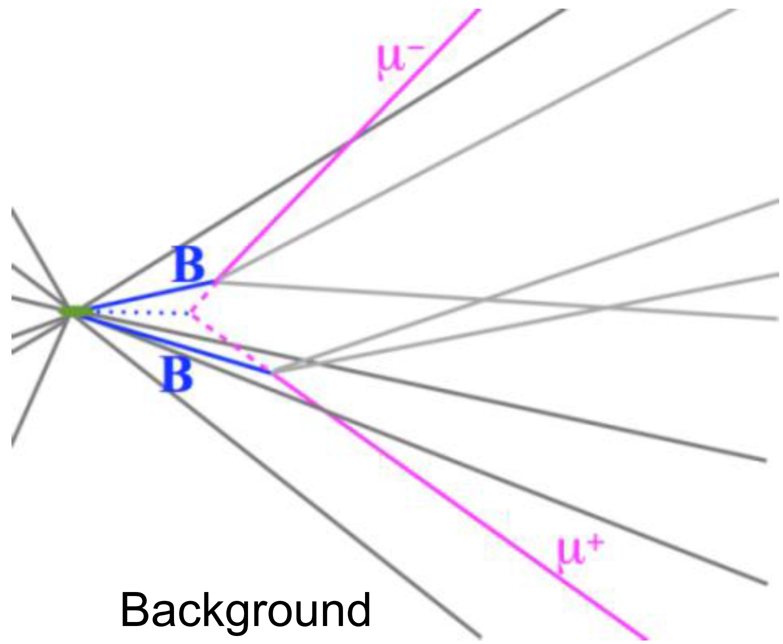


Main background

- $b \to X\mu\nu$ ~10%, $b \to cX$, $c \to Y\mu\nu$ ~10%
- So $b\bar{b} \to X'\mu^+\mu^-$ ~4×10^{-2}, compared with signal in SM ~4×10^{-9}.



Signal



Background

BDT selection for $\mu^+\mu^-$

- Idea of multivariate analyses is to use the variables & their correlations, rather than make rectangular cuts. Improves efficiency for a given background rejection

LHCb variables

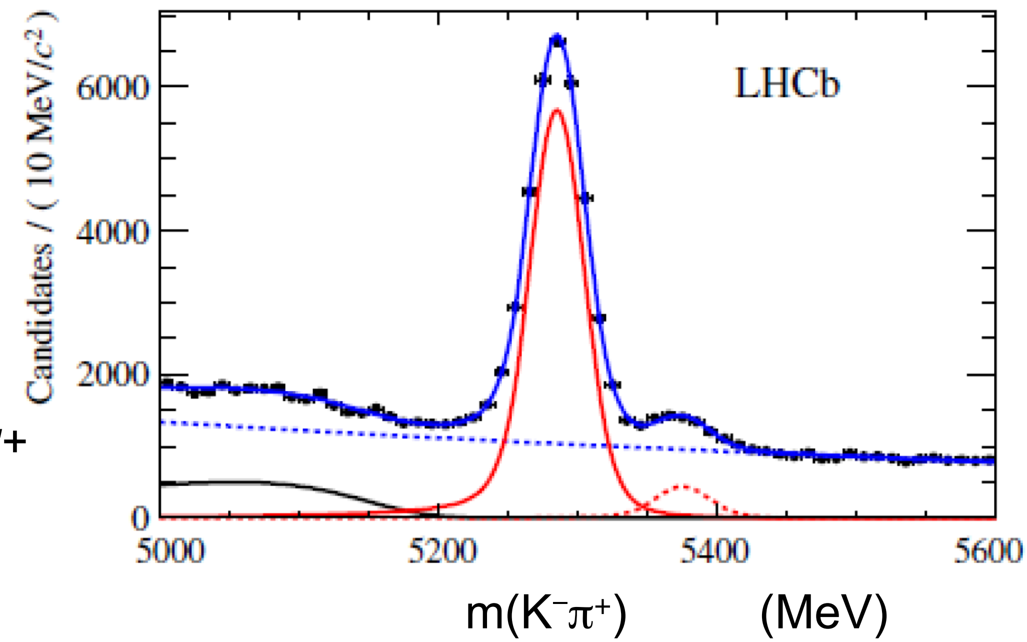
- Muons: IP significance, distance of closest approach of μ^+ & μ^-, isolation, polarization ∠, $\Delta\eta$ & $\Delta\phi$
- Define P_{thrust} as the $\Sigma \mathbf{p}_i$ of all tracks consistent with coming from the other B. Then for
- B candidate: decay time, IP, p_T, isolation, ∠ between p_B & P_{thrust}, & ∠ between μ^+ direction & P_{thrust} in B rest frame

CMS variables

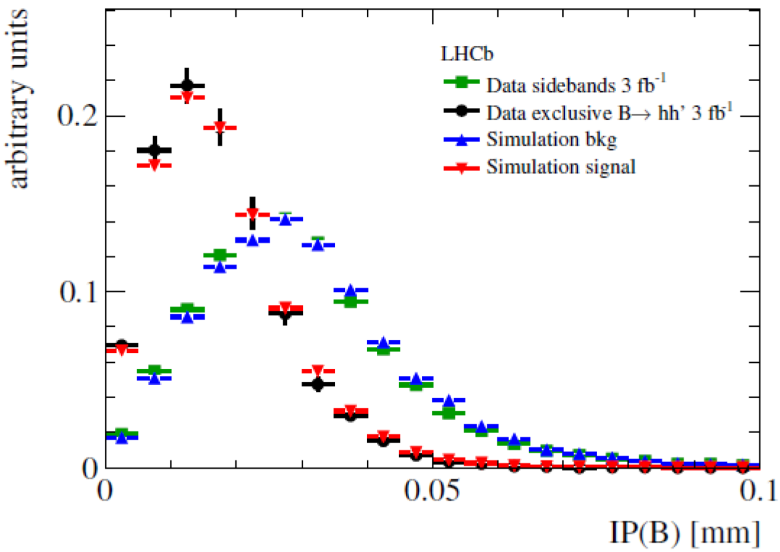
- B-vertex fit χ^2/ndof
- Distance of closest approach of μ^+ & μ^-
- the 3D pointing ∠ wrt pv
- 3D flight length significance
- 3D impact parameter (IP) of the B candidate
- IP significance

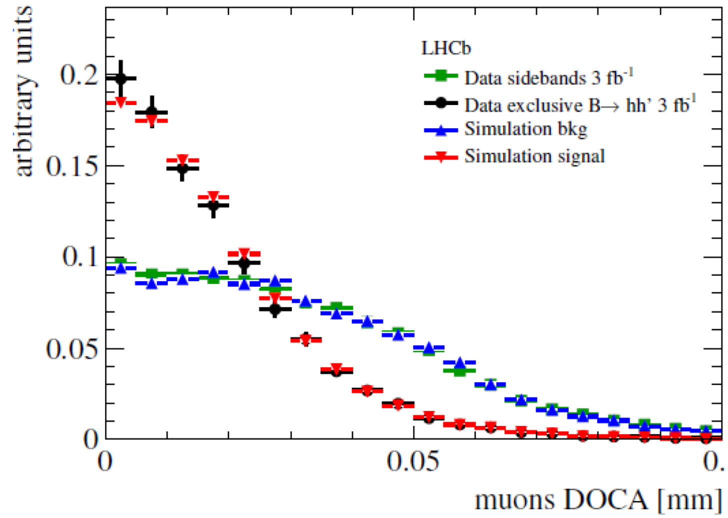
BDT discrimination

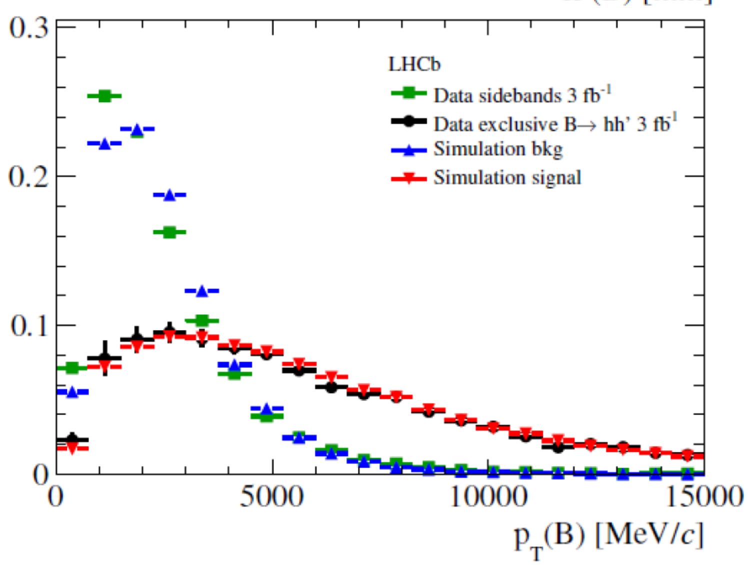
- Basic idea is to use a sample for signal & a separate sample for background. The program then figures out the best discrimination based on ONE variable
- Some examples from LHCb
  - Signal samples from simulation and $B \to h^- h'^+$
  - Background samples from simulation and sidebands of the dimuon mass

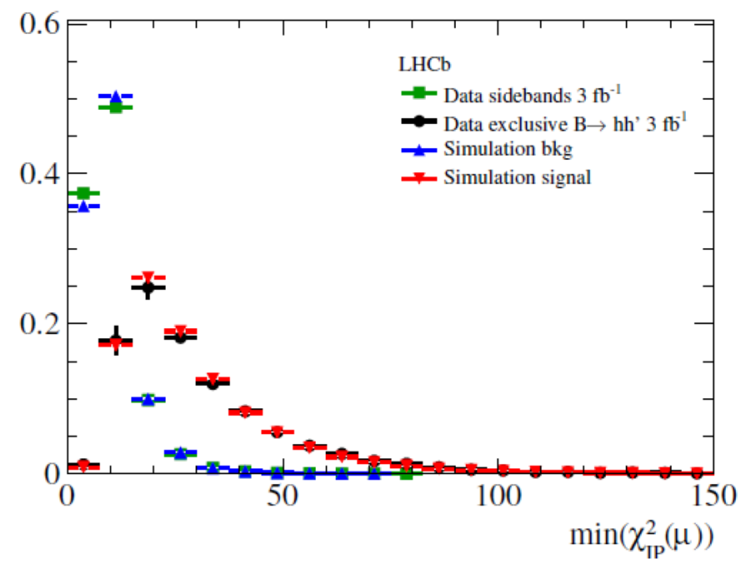


BDT variable studies



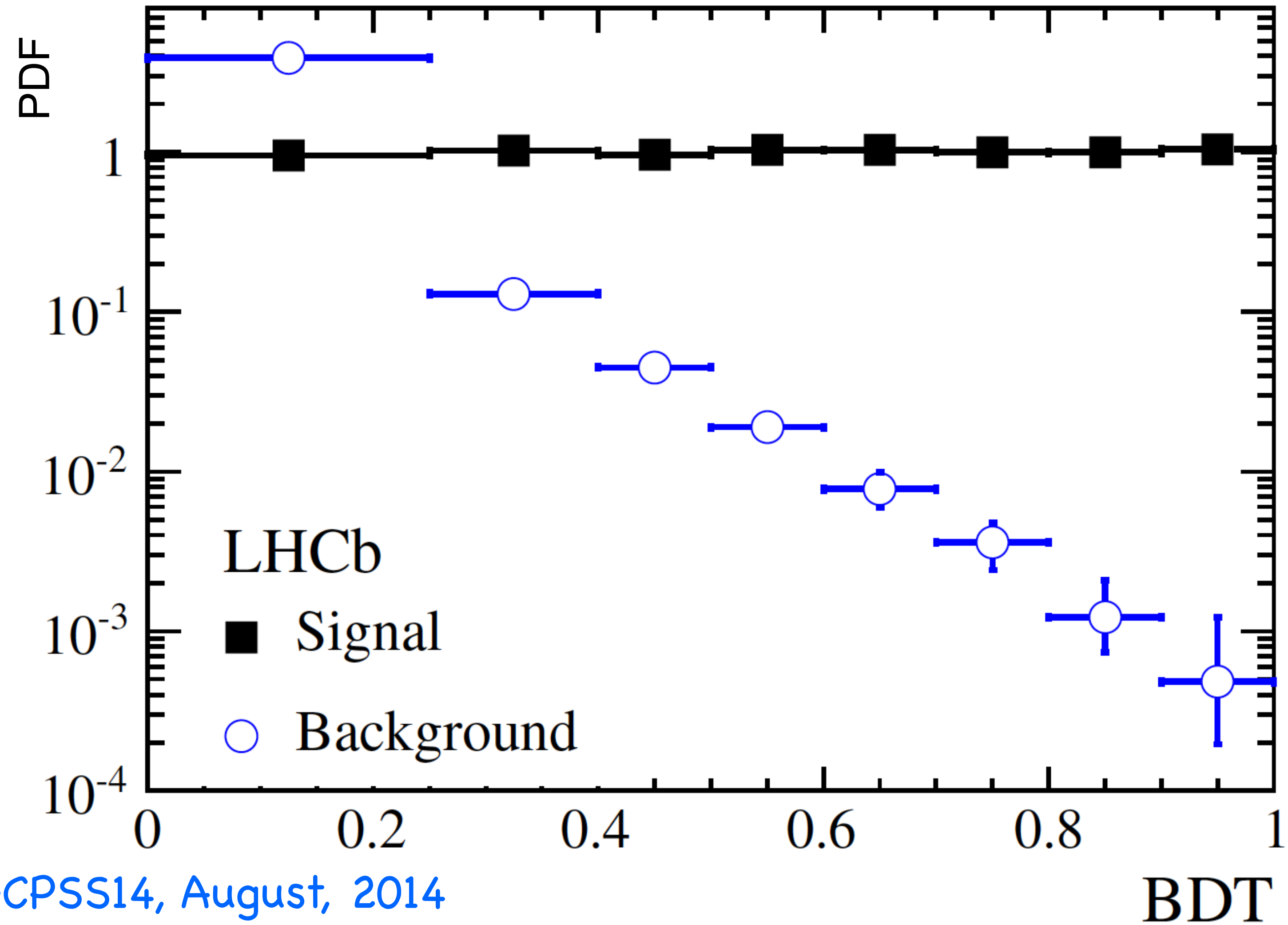






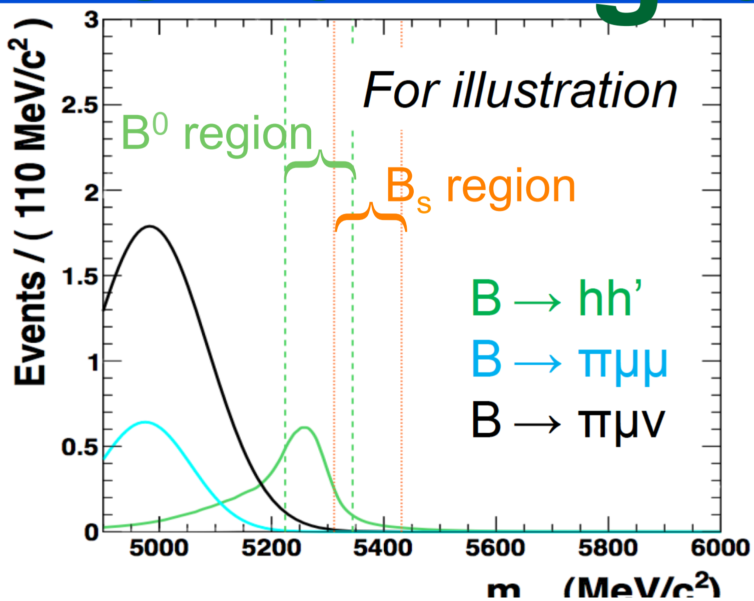
BDT output

- Tuned to be flat for signal



Remaining backgrounds

- LHCb

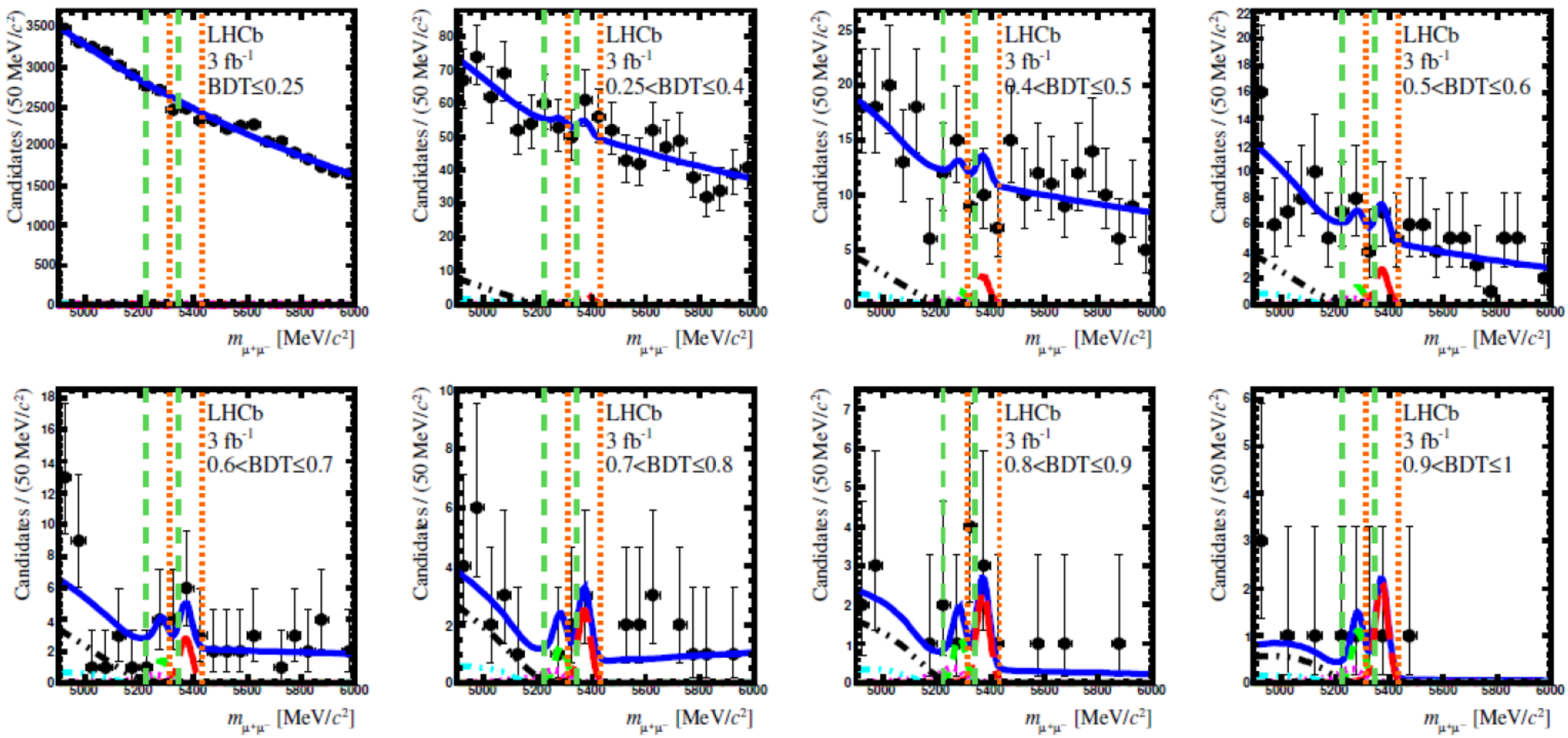


CMS also has $h^+h'^-$ & semileptonic background

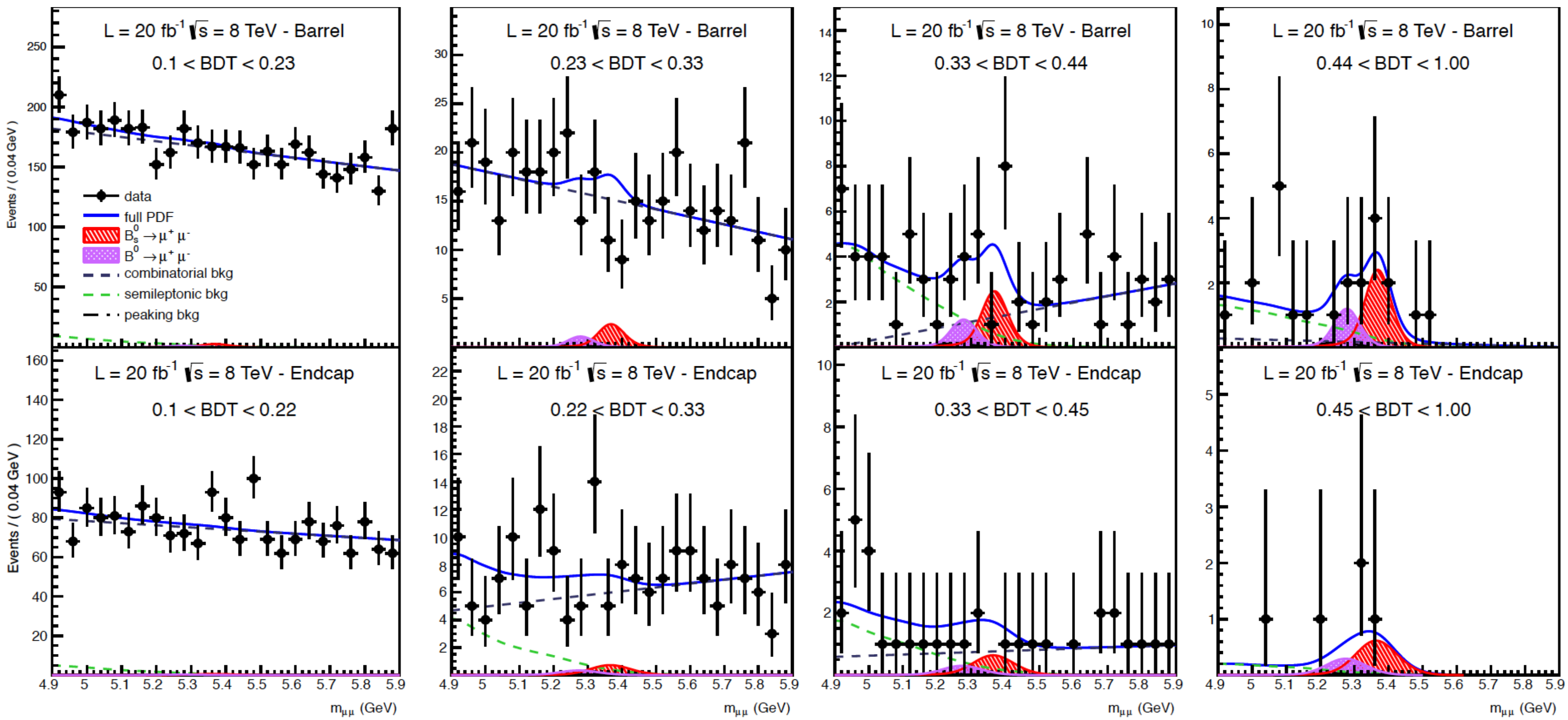
- Measured yields used for predictions where possible

	Yield in full BDT range	Fraction with BDT > 0.7 [%]
$B^0_{(s)} \to h^+h'^-$	15±1	28
$B^0 \to \pi^-\mu^+\nu_\mu$	115±6	15
$B^0_s \to K^-\mu^+\nu_\mu$	10±4	21
$B^{0(+)} \to \pi^{0(+)}\mu^+\mu^-$	28±8	15
$\Lambda^0_b \to p\mu^-\bar{\nu}_\mu$	70±30	11

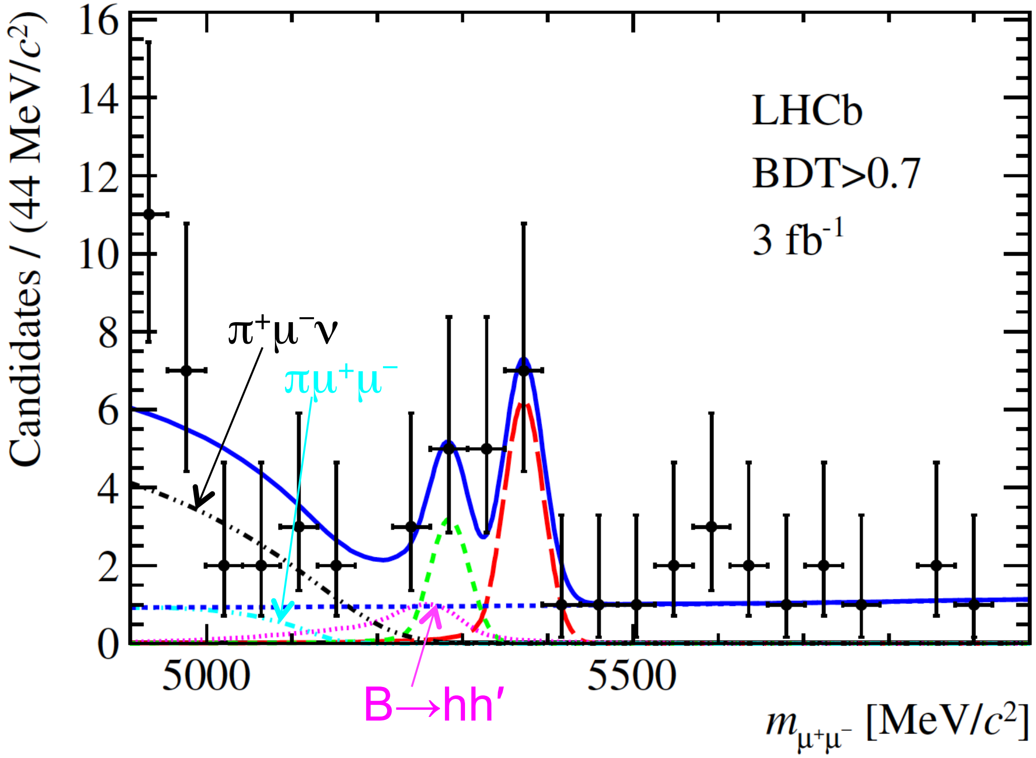
LHCb fit results

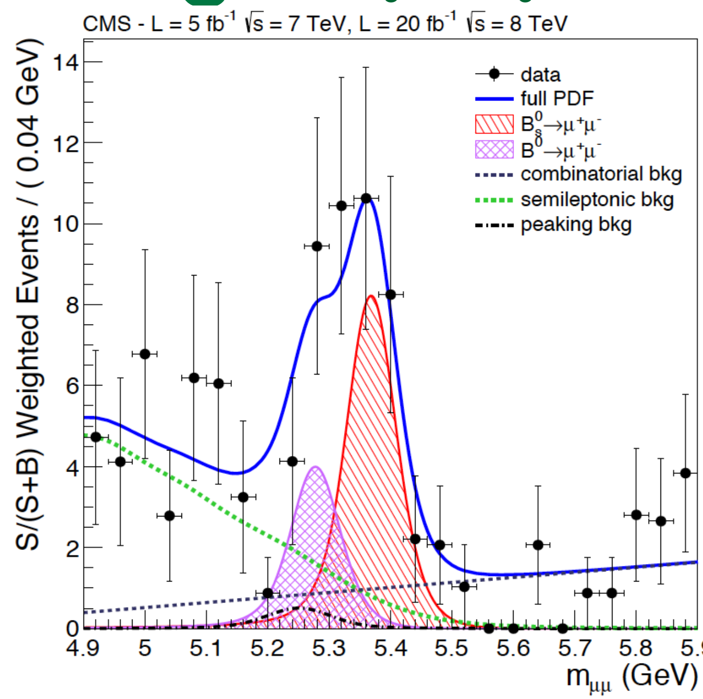


CMS fit results



Evidence for $B_s \to \mu^+\mu^-$





CMS: arXiv:1307.5025, PRL. 111.101804 (2013)

LHCb: arXiv:1307.5024, PRL.111.101805 (2013)

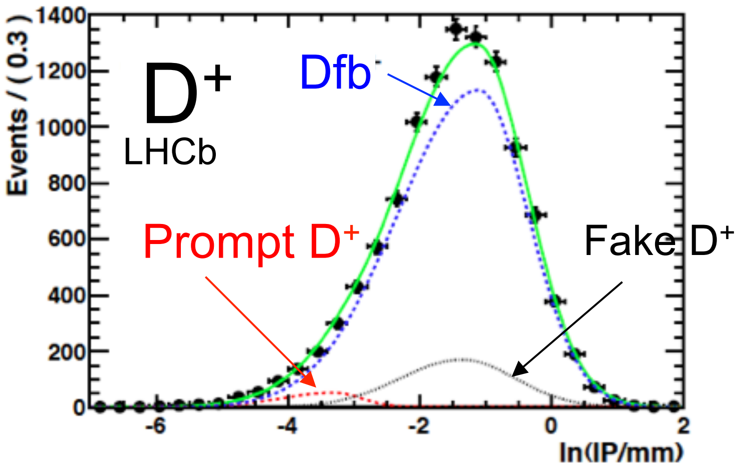
of B_s

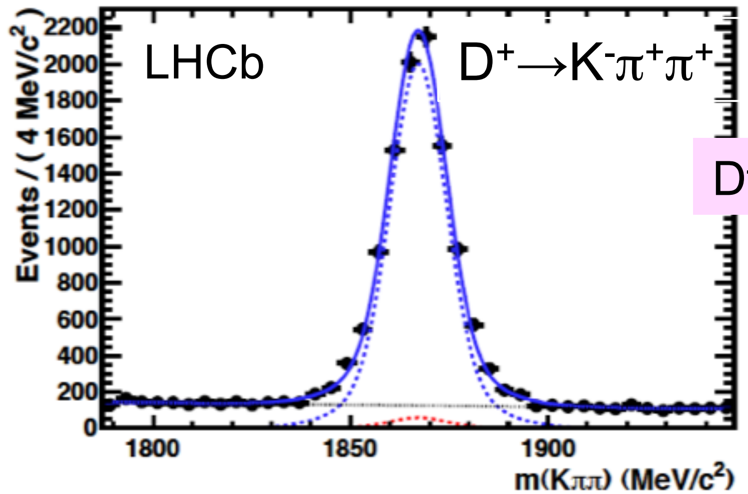
- We now have signal yields.
- Branching fraction requires # signal /# B_s
- # B_s determined by LHCb
- Semileptonic method – uses the fact that the semileptonic decay widths $\Gamma(b_i \rightarrow X_i \mu\nu)$ are equal for all b species. Since $\Gamma(b_i \rightarrow X_i \mu\nu)$=yield/$\tau_{bi}$ (known) measuring these modes gives production ratios, i.e. f_s/f_d

$$\frac{f_s}{f_u + f_d} = \frac{n_{\rm corr}(\bar{B}^0_s \rightarrow D\mu)}{n_{\rm corr}(B \rightarrow D^0\mu) + n_{\rm corr}(B \rightarrow D^+\mu)} \frac{\tau_{B^-} + \tau_{\bar{B}^0}}{2\tau_{\bar{B}^0_s}}$$

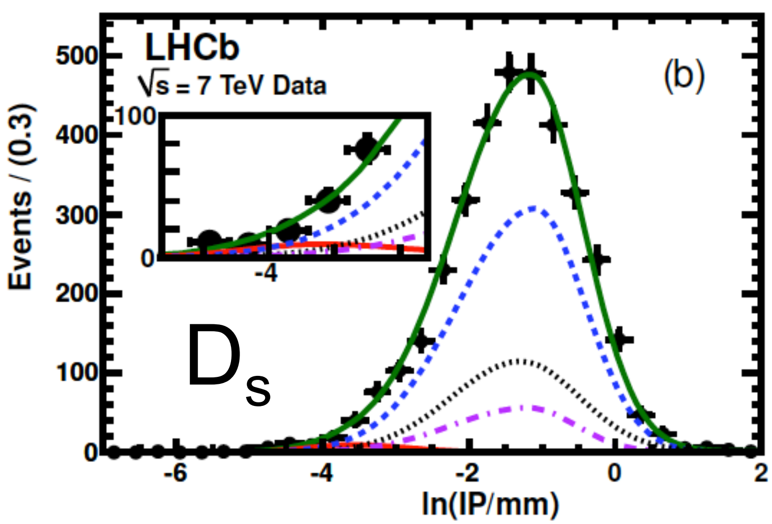
- Ncorr($B_s \rightarrow D\mu$) is $D_sX\mu$+DKXμ (arXiv:1111.2357)

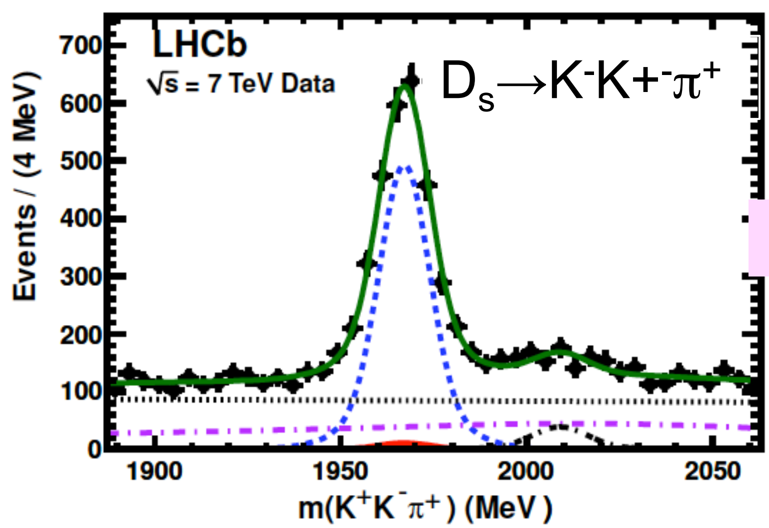
Production fractions: $B \to DX\mu\nu$ use equality of Γ_{sl} & known τ's





Dfb: 9406±110



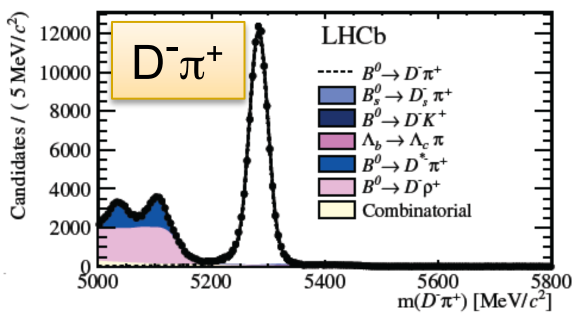


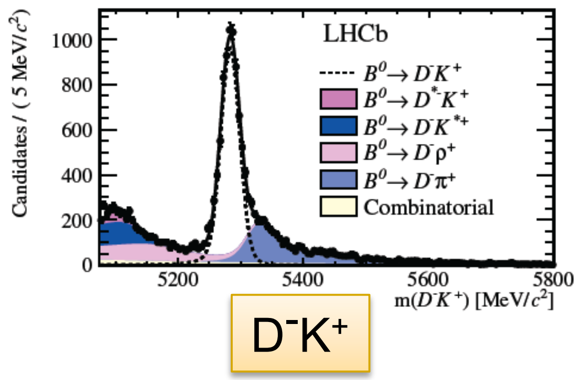
Dfb: 2446±60

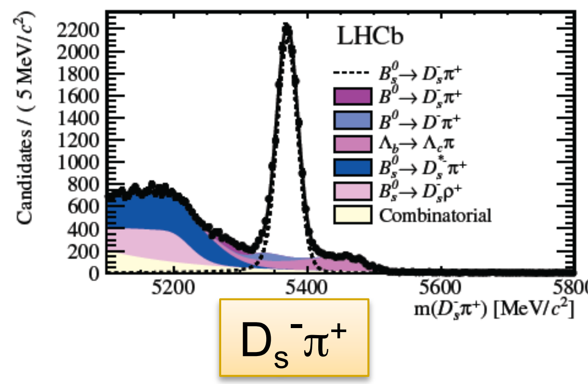
Also D^o, Λ_b

Hadronic

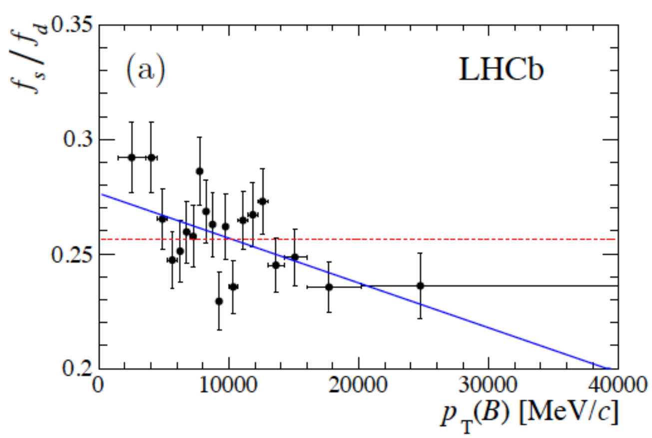
- Hadronic method – uses hadronic two-body decays: $B_s \rightarrow D_s\pi^-$, $B^0 \rightarrow D^0\pi^-$, $B^0 \rightarrow D^0K^-$ & form-factor ratio from theory (arXiv:1106.4435)

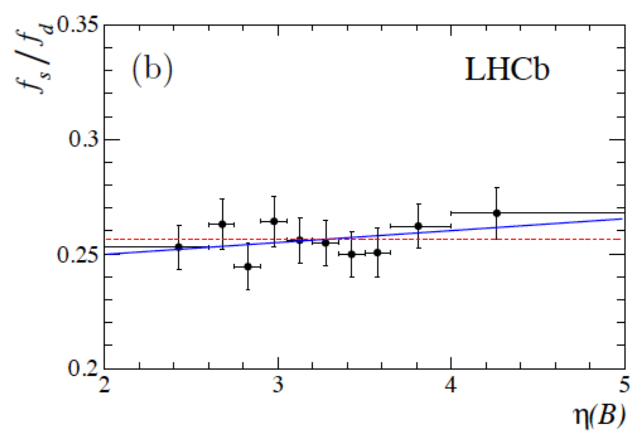






- Take ratios, use theory
- p_T & η dependences





Branching fraction

- Using measured $f_s/f_u=f_s/f_d=0.259\pm0.15$
- & relative $\mu^+\mu^-$ yields with respect to normalization modes

LHCb:

$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (2.9\,^{+1.1}_{-1.0}) \times 10^{-9},$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.7\,^{+2.4}_{-2.1}) \times 10^{-10}$$

--> 4.0σ

CMS:

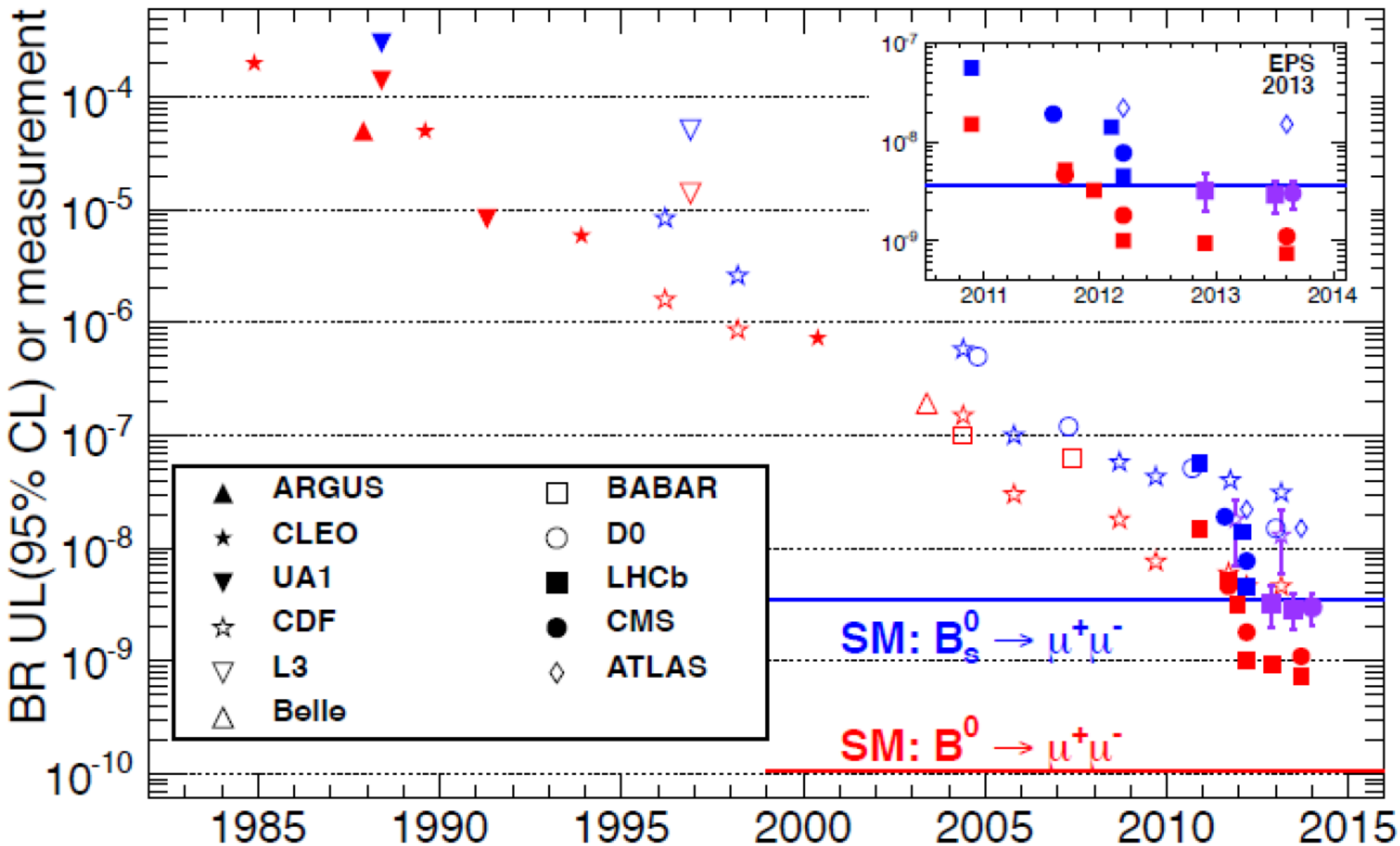
$$\mathcal{B}(B_s^0 \to \mu^+\mu^-) = (3.0\,^{+1.0}_{-0.9}) \times 10^{-9},$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (3.5\,^{+2.1}_{-1.8}) \times 10^{-10}$$

--> 4.3σ

- Avg: $\mathcal{B}(B_s\to\mu^+\mu^-)=(2.9\pm0.7)\mathrm{x}10^{-9}$
- Avg: $\mathcal{B}(B^0\to\mu^+\mu^-)=(3.6\,^{+1.6}_{-1.4})\mathrm{x}10^{-10}$ (not significant)

Upper limit < $5.7\mathrm{x}10^{-10}$ @ 90% c.l.

History



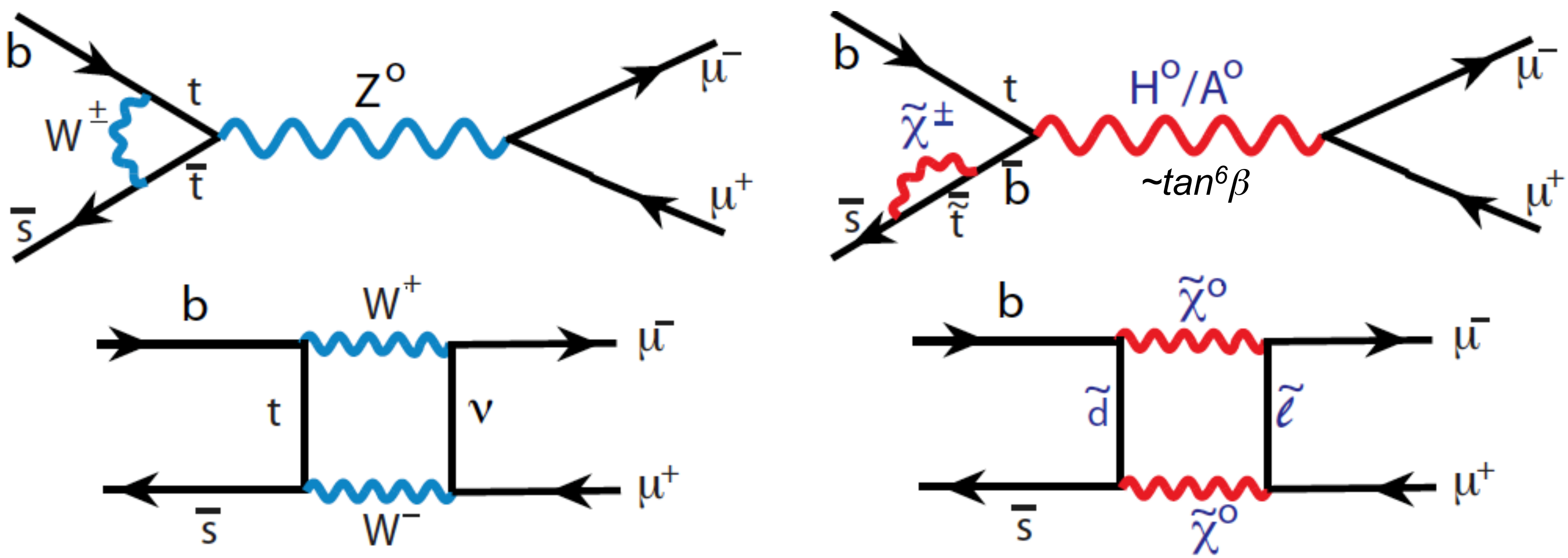
BR UL(95% CL) or measurement
10^-4
10^-5
10^-6
10^-7
10^-8
10^-9
10^-10
1985
1990
1995
2000
2005
2010
2015
EPS 2013
2011
2012
2013
2014
ARGUS
CLEO
UA1
CDF
L3
Belle
BABAR
D0
LHCb
CMS
ATLAS
SM: $B_s^0 \to \mu^+\mu^-$
SM: $B^0 \to \mu^+\mu^-$

Theory $B_s \rightarrow \mu^+\mu^-$

- SM branching ratio is $(3.65\pm0.23)\times10^{-9}$ [Buras arXiv:1012.1447], NP can make large contributions.

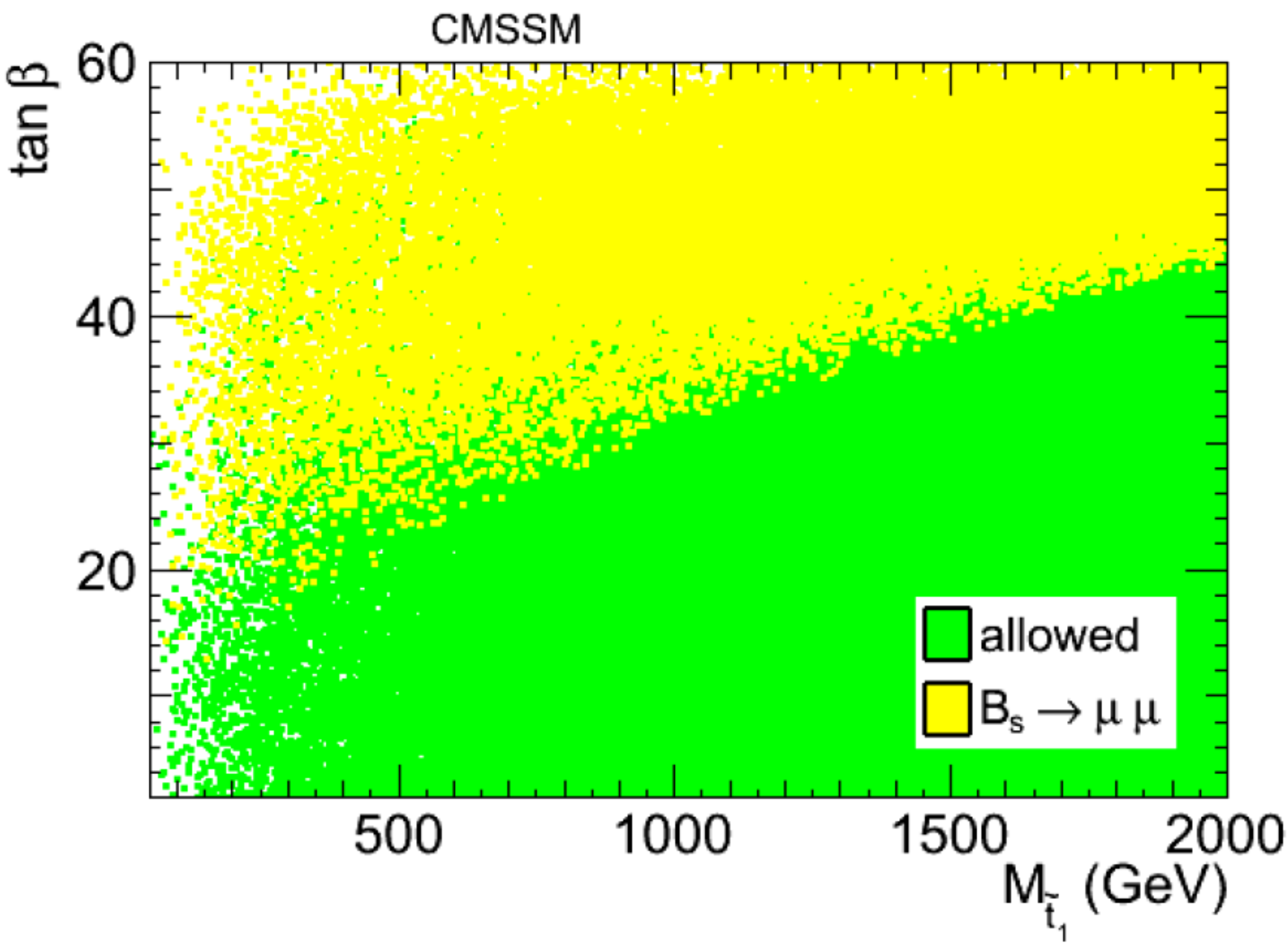
Standard Model

MSSM



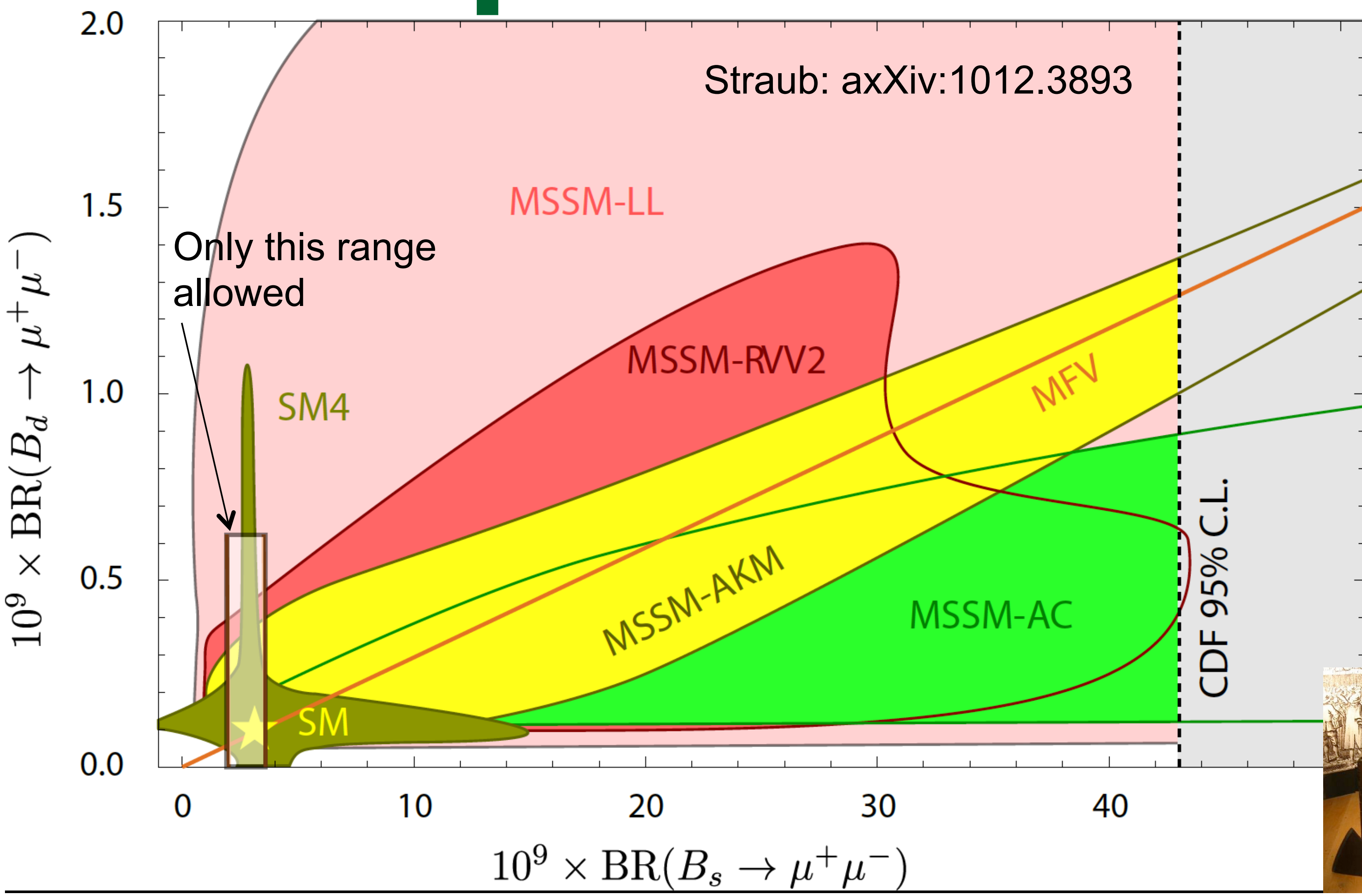
- Many NP models possible, not just Super-Sym

Implications



Mahmoudi et al

Implications II



An Aside on lifetimes

Γ(t) for neutral B decays

- Recall $\Gamma \cdot \tau = \hbar$

$$\Gamma[f,t] = \Gamma(B_s(t) \to f) + \Gamma(\overline{B}_s(t) \to f)$$

$$= \mathcal{N}_f \left[e^{-\Gamma_L t} \left| \langle f | B_L \rangle \right|^2 + e^{-\Gamma_H t} \left| \langle f | B_H \rangle \right|^2 \right].$$

$$= \mathcal{N}_f \left| A_f \right|^2 \left[1 + |\lambda_f|^2 \right] e^{-\Gamma t} \left\{ \cosh \frac{\Delta\Gamma\, t}{2} + \sinh \frac{\Delta\Gamma\, t}{2} \mathcal{A}_{\Delta\Gamma} \right\}$$

$$A_{\Delta\Gamma} \equiv -2\mathrm{Re}\left(\lambda_f\right) / \left(1 + \left|\lambda_f\right|^2\right), \quad \lambda_f = \frac{q}{p} \frac{\overline{A}_f}{A_f}$$

- Shape is not exponential & depends on decay mode. To 2nd order

$$\Gamma[f,t] \propto e^{-\Gamma t} \left[1 + \frac{1}{2} \left(\frac{\Delta\Gamma}{2} t \right)^2 + A_{\Delta\Gamma} \left(\frac{\Delta\Gamma}{2} t \right) \right]$$

B_s versus B^0

- For B^0 $\Delta\Gamma_d/\Gamma_d$ has been measured as 0.015±0.018 by B factories [PDG], so decay can be treated as purely exponential ($\Delta\Gamma_d$<0.032 ps^{-1} @ 95% cl) consistent with theoretical prediction of 2×10^{-3} ps^{-1} [arXiv:0412007]
- For B_s, $\Delta\Gamma$ is not small and $A_{\Delta\Gamma}$ depends on decay mode, mainly through $\overline{A}_f/A_f$ as q/p has been measured as being small
- For "flavor specific" B_s decay modes, where $B_s \rightarrow f$ & $\overline{B}_s \rightarrow \overline{f}$ the decay is the sum of two exponentials & here

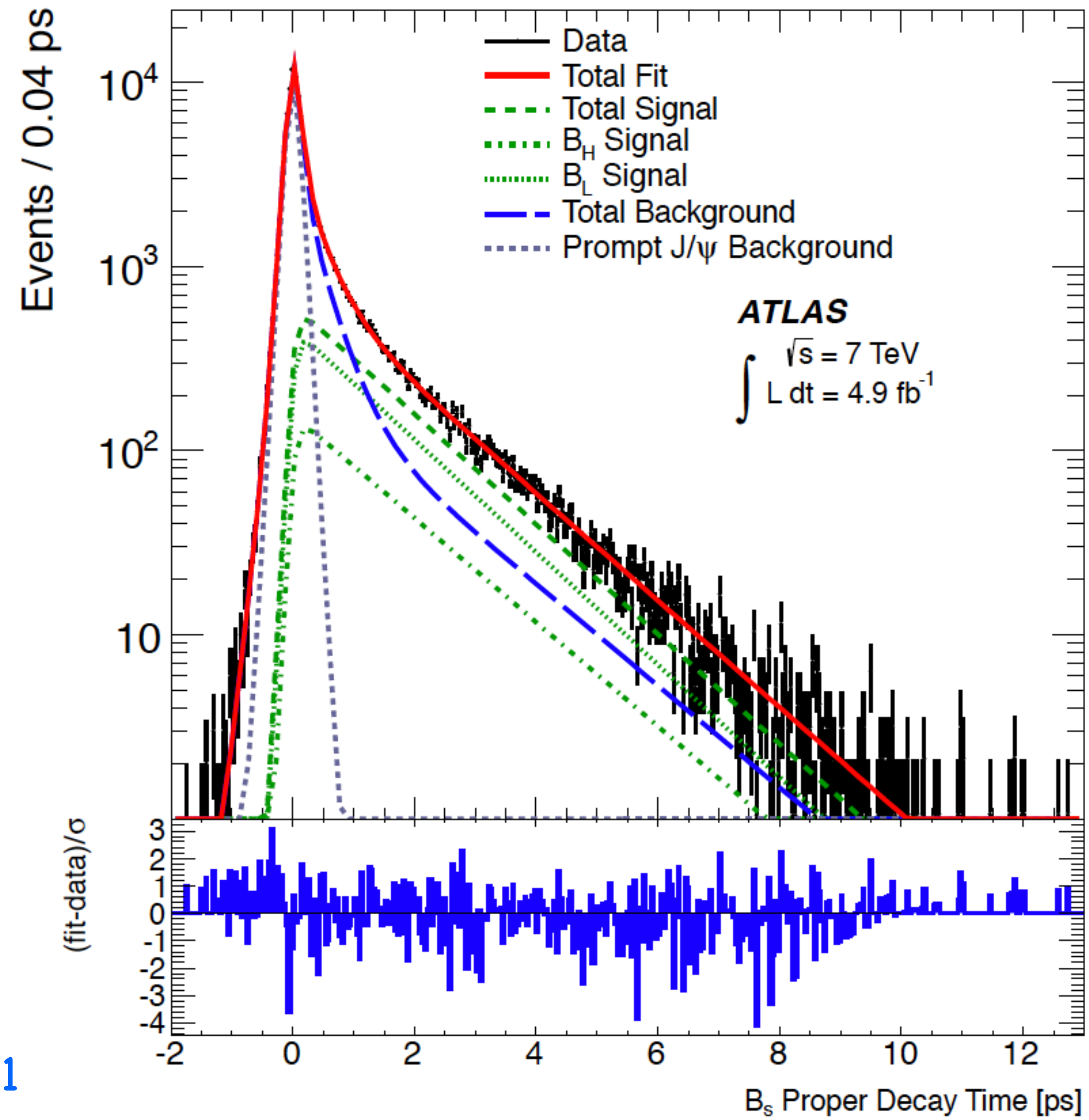
$$\Gamma_s = \Gamma_{flavor\ specific}\left(1-\left(\frac{\Delta\Gamma_s}{2\Gamma_s}\right)^2\right)\Bigg/\left(1+\left(\frac{\Delta\Gamma_s}{2\Gamma_s}\right)^2\right)$$

Measurement of Γ_s

- Here Γ_s is determined along with information on CP violation – direct measurements
- I use the measurements from $B_s \to J/\psi\phi$ from CDF, D0, ATLAS & LHCb (also $J/\psi\pi^+\pi^-$). Γ_s values are obtained from the lifetime fit along with the CPV measurement. (Both flavor tagged & untagged data are used)
- This differs from HFAG

Example: ATLAS

- Fit returns B_L & B_H distributions as well as a value for the CP violating phase
- G. Aad et al., ATLAS, JHEP 1212 (2012) 072

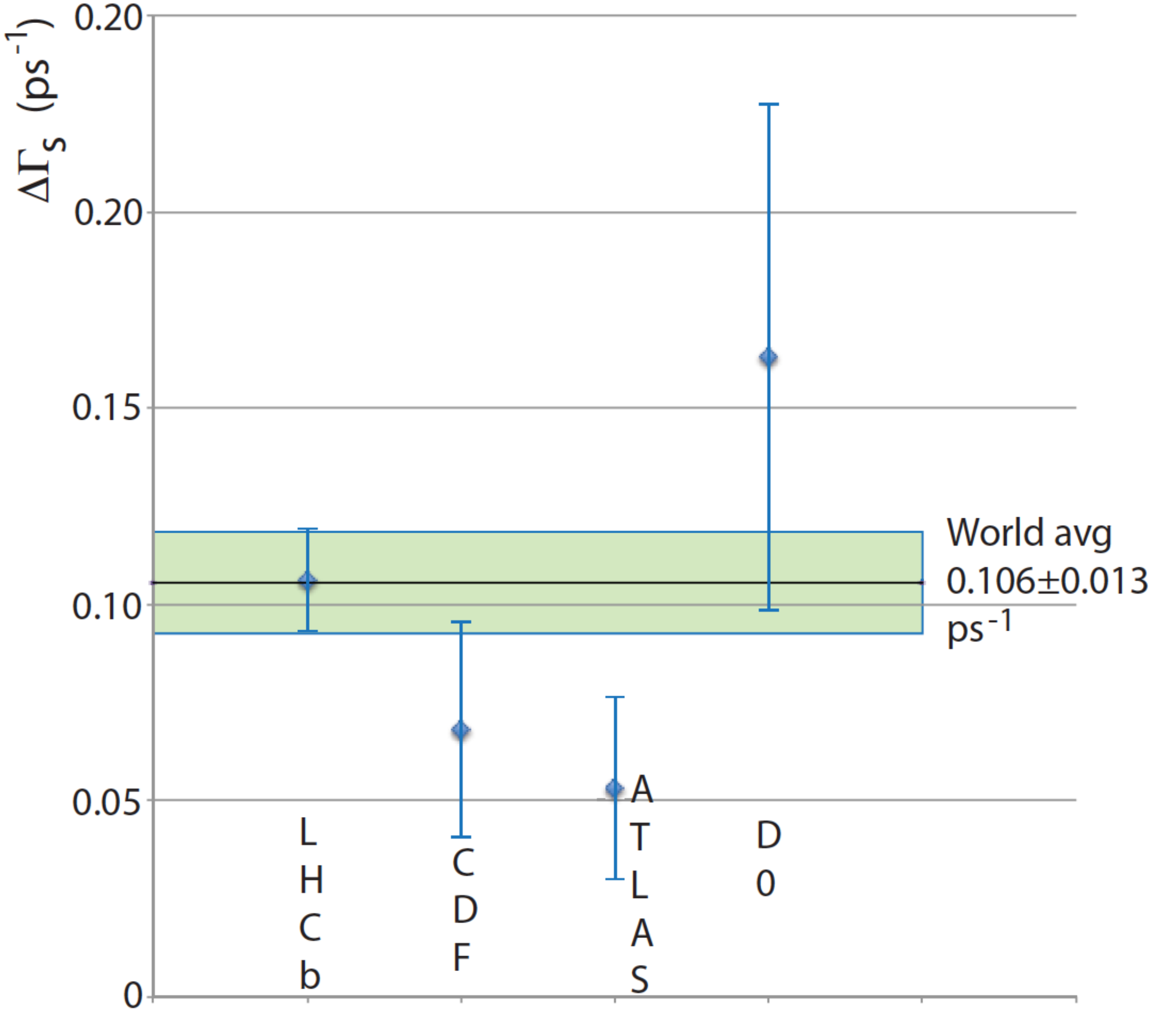


Γ_s values from $J/\psi(\phi$ & $\pi^+\pi^-)$

Exp.	$\int\mathcal{L}$ (fb^{-1})	Γ_s (ps^{-1})	ArXiv
ATLAS	4.9	0.6700±0.0070±0.0040	1208.0572
CDF	9.6	0.6545±0.0081±0.0039	1208.2967
D0	8.0	0.6930±0.0182	1109.3166
LHCb	1	0.6610±0.0040±0.0060	1304.2600
Average		0.666±0.0045	

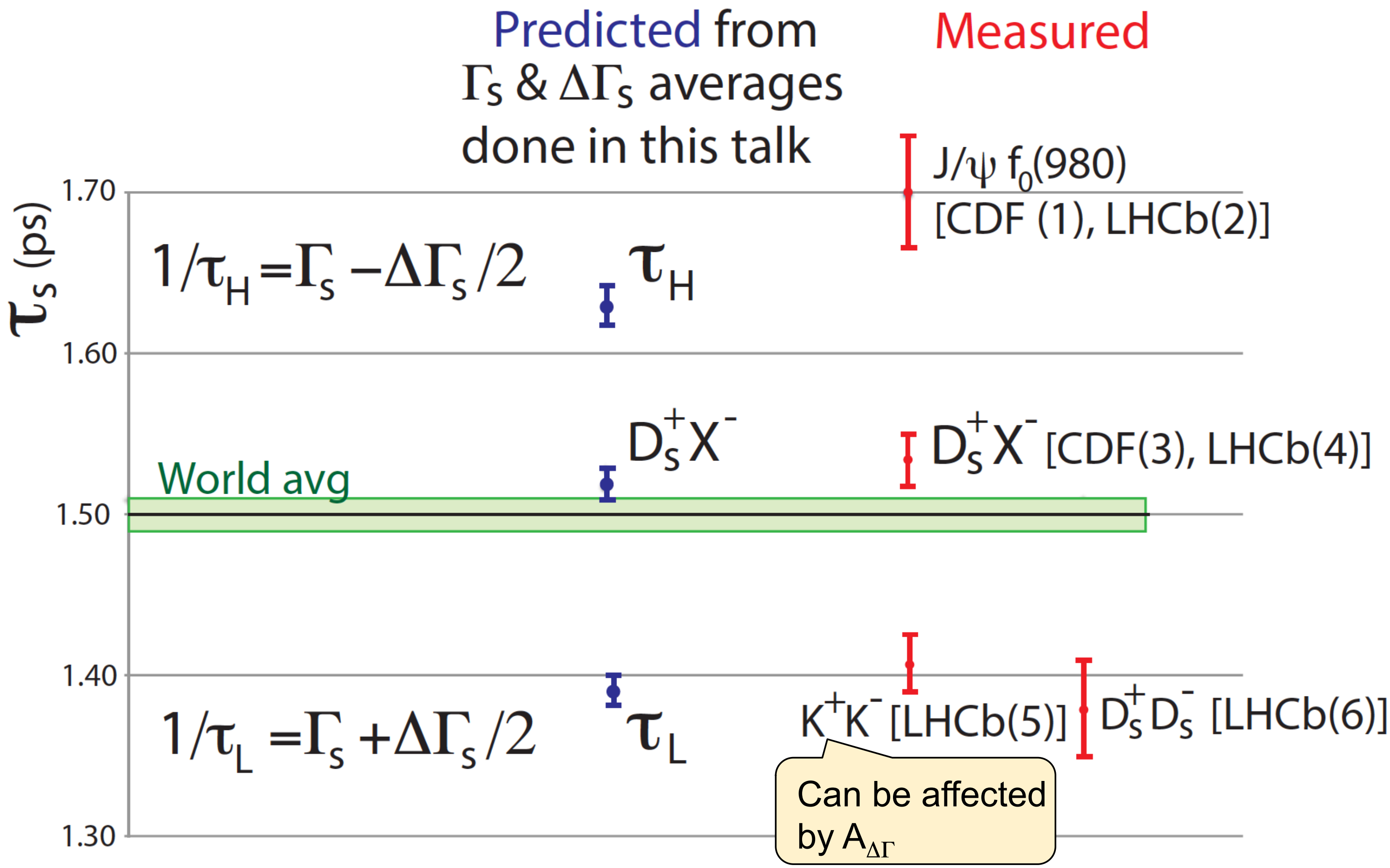
τ_s =1.500±0.010

$\Delta\Gamma_s$



	arXiv
ATLAS	1208.0572
CDF	1208.2967
D0	1109.3166
LHCb	1304.2600

τ_s from other measures



1-arXiv:1106.3682, 2-arXiv:1207.0878, 3-arXiv:1103.1864, 4-**arXiv:1407.5873** 5-LHCB-PAPER-2014-011, 6-arXiv:1312.1217

Rare Decays - Generic

- $$\mathcal{H}_{\text{eff}} = -\frac{4\,G_F}{\sqrt{2}} V_{tb} V_{ts}^* \frac{e^2}{16\pi^2} \sum_i (C_i O_i + C'_i O'_i) + \text{h.c.}\ .$$
- C_i are Wilson coefficients, O_i are 4-fermion operators. C_iO_i for SM, $C_i'O_i'$ are for NP. $P_{R,L} = (1\pm\gamma_5)/2$. $O' = O$ with $P_{R,L} \to P_{L,R}$

$$O_7 = \frac{m_b}{e} (\bar{s}\sigma_{\mu\nu} P_R b) F^{\mu\nu}, \qquad O_8 = \frac{g m_b}{e^2} (\bar{s}\sigma_{\mu\nu} T^a P_R b) G^{\mu\nu\, a},$$

$$O_9 = (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \ell)\,, \qquad O_{10} = (\bar{s}\gamma_\mu P_L b)(\bar{\ell}\gamma^\mu \gamma_5 \ell)\,,$$

$$O_S = m_b(\bar{s} P_R b)(\bar{\ell}\ell)\,, \qquad O_P = m_b(\bar{s} P_R b)(\bar{\ell}\gamma_5 \ell)\,,$$

- Each process depends on a unique combination.

Theory $B_s \to \mu^+\mu^-$ II

- For SM only have C_{10}, since C'_{10}, $C^{(\prime)}_S$ & $C^{(\prime)}_P$ are negligibly small
- Define new combination of Wilson coeff for further use in NP models

$$P \equiv \frac{C_{10} - C'_{10}}{C_{10}^{\rm SM}} + \frac{m_{B_s}^2}{2m_\mu}\left(\frac{m_b}{m_b + m_s}\right)\left(\frac{C_P - C'_P}{C_{10}^{\rm SM}}\right) \equiv |P|e^{i\varphi_P},$$

$$S \equiv \sqrt{1 - \frac{4\,m_\mu^2}{m_{B_s}^2}}\frac{m_{B_s}^2}{2m_\mu}\left(\frac{m_b}{m_b + m_s}\right)\left(\frac{C_S - C'_S}{C_{10}^{\rm SM}}\right) \equiv |S|e^{i\varphi_S}.$$

- In SM P=1, S=0

More definitions

$$\mathcal{S}_{\mu\mu} = \frac{|P|^2 \sin(2\varphi_P - \phi_s^{\rm NP}) - |S|^2 \sin(2\varphi_S - \phi_s^{\rm NP})}{|P|^2 + |S|^2},$$

$$\mathcal{A}_{\Delta\Gamma}^{\mu\mu} = \frac{|P|^2 \cos(2\varphi_P - \phi_s^{\rm NP}) - |S|^2 \cos(2\varphi_S - \phi_s^{\rm NP})}{|P|^2 + |S|^2}.$$

Time dependent rate

- In the $\mu^+\mu^-$ final state the sum of the final state helicities must be 0. Since helicities are difficult to measure, sum over L & R states
- Then we can construct the untagged lifetime as (see arXiv:1303.3820)

$$\langle\Gamma(B_s(t)\to\mu^+\mu^-)\rangle \equiv \Gamma(B_s^0(t)\to\mu^+\mu^-) + \Gamma(\bar{B}_s^0(t)\to\mu^+\mu^-)$$

$$= \frac{G_F^4\, M_W^4 \sin^4\theta_W}{4\pi^5}\left|C_{10}^{\rm SM} V_{ts}V_{tb}^*\right|^2 F_{B_s}^2 m_{B_s} m_\mu^2\, \sqrt{1-\frac{4m_\mu^2}{m_{B_s}^2}}$$

$$\times\ \left(|P|^2+|S|^2\right)$$

$$\times\ e^{-t/\tau_{B_s}}\left[\cosh\left(y_s\, t/\tau_{B_s}\right)+\mathcal{A}_{\Delta\Gamma}^{\mu\mu}\sinh\left(y_s\, t/\tau_{B_s}\right)\right].$$

Lifetime &CPV

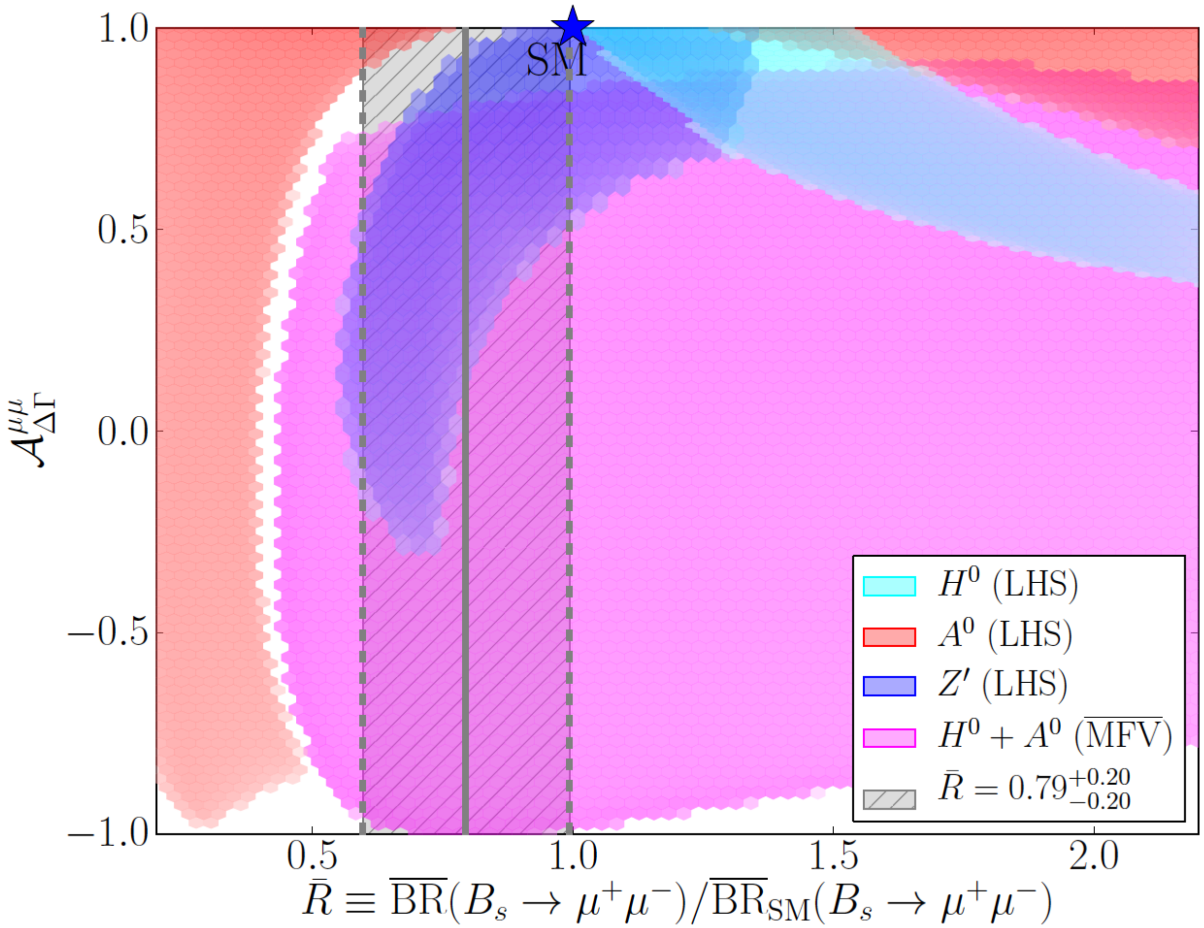
- So measuring the lifetime allows a determination of $\mathcal{A}_{\Delta\Gamma}^{\mu\mu}$ which is sensitive to NP
- Considering that we have about 30 events now in each experiment, this will take a while
- Can also hope to measure CPV

$$\frac{\Gamma(B_s^0(t) \to \mu^+\mu^-) - \Gamma(\bar{B}_s^0(t) \to \mu^+\mu^-)}{\Gamma(B_s^0(t) \to \mu^+\mu^-) + \Gamma(\bar{B}_s^0(t) \to \mu^+\mu^-)} = \frac{\mathcal{S}_{\mu\mu}\sin(\Delta M_s t)}{\cosh(y_s t/\tau_{B_s}) + \mathcal{A}_{\Delta\Gamma}^{\mu\mu}\sinh(y_s t/\tau_{B_s})}$$

- But this will take even more data

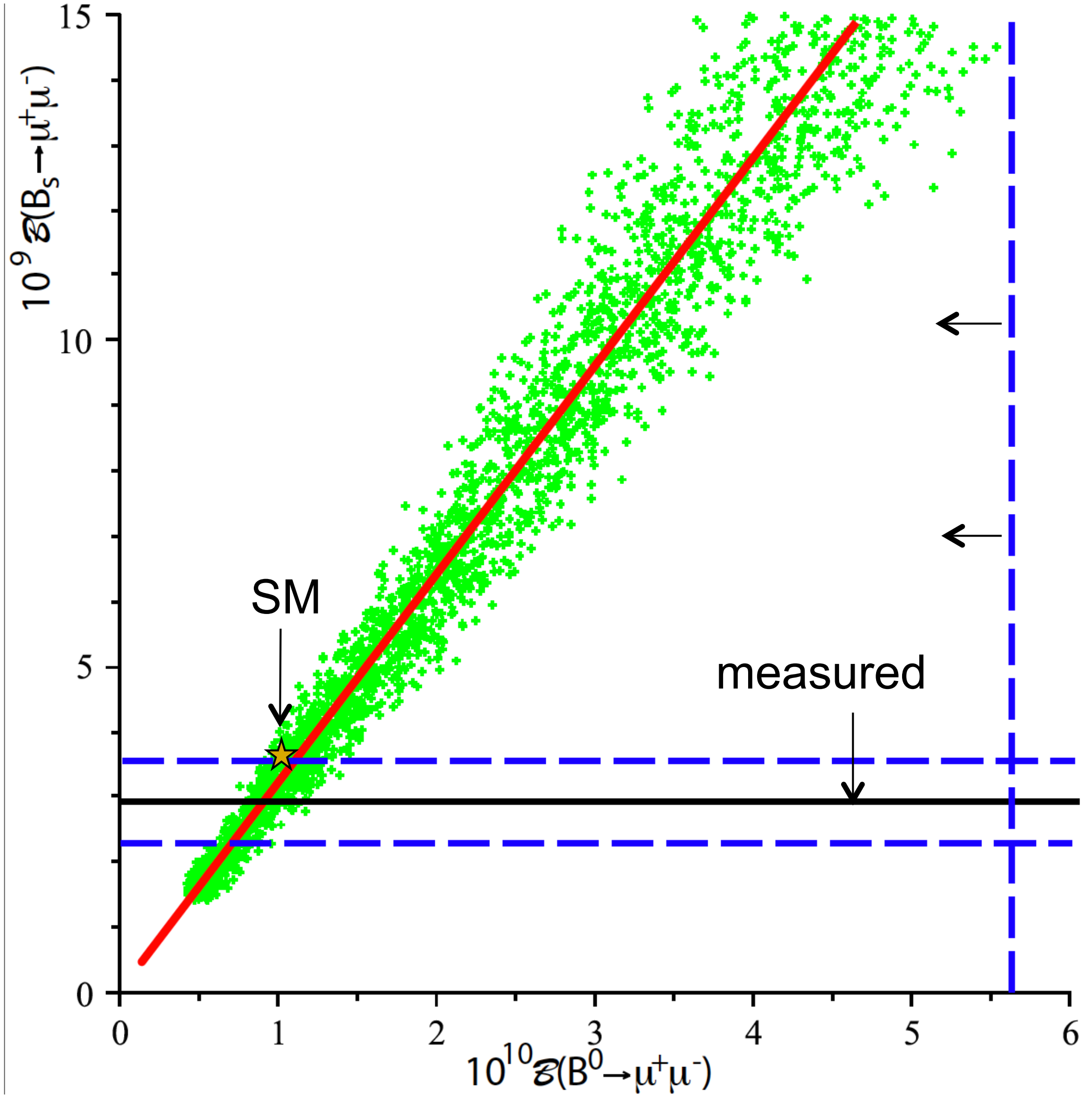
Different models

- See arXiv:1303:3820
- LHS≡Left handed scheme
- A^0 new pseudoscalar
- H^0 new scalar



What about MFV?

- In principle, ratio of B0/Bs can show if NP is consistent with MFV
- Correlation shown for a genic model with Higgs-mediated FCNC consistent with MFV. Green points give the uncertainties



Conclusions

- $\mathcal{B}(B^0_s \to \mu^+\mu^-)$ measured and consistent with SM
- More precise determination of $\mathcal{B}$ will limit models or show NP
- Other variables in the decay, the lifetime and CP asymmetry can also show NP, either generically or reflect specific models
- Much information also from a definitive determination of $\mathcal{B}(B^0 \to \mu^+\mu^-)$

The End