Measurement of $\mathcal{B}(B^0_{(s)} \rightarrow \mu^+\mu^-)$ at LHC(b)

Mathieu Perrin-Terrin*

Seminar at CEA Saclay

Monday, January 6th 2014

*CERN

Outlines

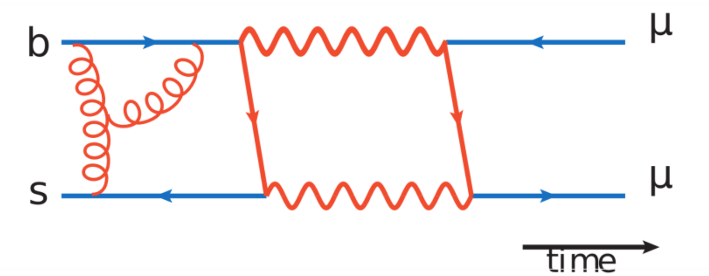
- Motivations to search for $B^0_{(s)} \to \mu^+\mu^-$
- Searching for $\mathrm{B}^0_{(s)} \to \mu^+\mu^-$ at LHCb
- Combination with CMS
- Conclusions

Outlines

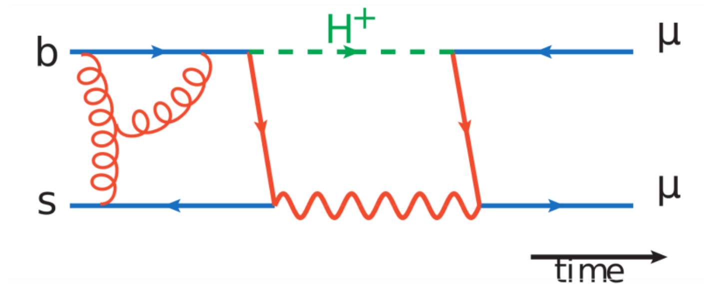
- Motivations to search for $B^0_{(s)} \rightarrow \mu^+\mu^-$
- Searching for $\mathrm{B}^0_{(s)} \rightarrow \mu^+\mu^-$ at LHCb
- Combination with CMS
- Conclusions

Searching for NP with Flavour Physics

- Decay amplitude **sums SM and NP** (if any) contributions
- For example: $A(B_s^0 \rightarrow \mu^+\mu^-) =$

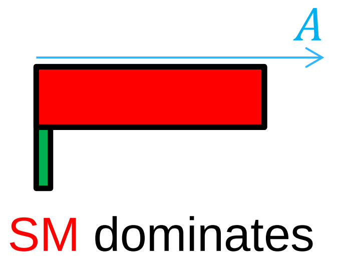


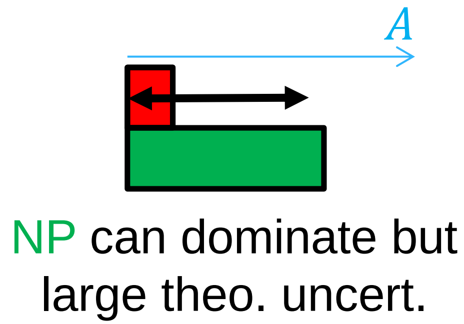
+

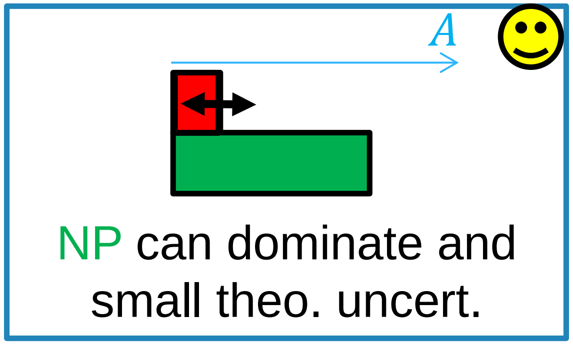


+ etc…

- Depending on the decay, several scenarios possible for A:





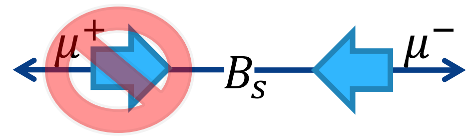


- Golden channel = **rare decay in SM, precisely predicted**

The $B^0_{(s)} \to \mu^+\mu^-$ Case: Rare & Precise

- **F**lavour **C**hanging **N**eutral **C**urrent:
  - No SM tree diagram, only higher order
  - Suppressed in SM
- Spin 0 weakly decaying into $\mu^+\mu^-$:
  - **Helicity suppression** in SM $BR_{SM} \propto m_\mu$

Helicity Suppressed



- Main source of uncertainties: QCD
  - Here, **leptonic final** state and…
  - …initial state **decay constant F_B well known**:

  F_B uncertainty strongly improved recently $20 \to 1.3\%$*

Golden channel = **rare decay in SM, precisely predicted**

*HPQCD, PRL 110, 222003, 2013

The $B^0_{(s)} \to \mu^+\mu^-$ Observables

- Simplest obs.: untagged time-integrated branching fraction:

$$BF = \frac{1}{2}\int dt \left[\Gamma\left(B^0_{(s)}(t) \to \mu^+\mu^-\right) + \Gamma\left(\bar{B}^0_{(s)}(t) \to \mu^+\mu^-\right) \right]$$

- Neutral B mesons oscillate in admixture of mass eigenstates:

$$\langle \Gamma(B^0_{(s)}(t) \to \mu^+\mu^-) \rangle = R_H e^{-\Gamma_H^{(s)} t} + R_L e^{-\Gamma_L^{(s)} t}$$

- For $\boldsymbol{B^0 \to \mu^+\mu^-}$ the decay widths are similar, $\Gamma_H \sim \Gamma_L$ so:

$$BF(B^0 \to \mu^+\mu^-) = BF(B^0(t) \to \mu^+\mu^-)_{t=0}$$
$$= \frac{\tau_{B^0}}{2}(R_H + R_L)$$
$$BF(B^0 \to \mu^+\mu^-) \stackrel{\text{SM}}{=} \mathbf{1.07 \pm 0.10 \times 10^{-10}}$$

Updated from Buras et al. EPJ 2012

- For $B_s^0 \to \mu^+\mu^-$ the decay widths are different, $\Gamma_H^s \neq \Gamma_L^s$, hence:

$$BF(B_s^0 \to \mu^+\mu^-) = \underbrace{\frac{\tau_{B_s^0}}{2}(R_H + R_L)}_{BF(B_s^0(t) \to \mu^+\mu^-)_{t=0}} \times \frac{1 + A_{\Delta\Gamma} y_s}{1 - y_s^2}$$

with:

$$A_{\Delta\Gamma} = \frac{R_H - R_L}{R_H + R_L} \in [-1;1] \qquad y_s = \frac{\Gamma_L - \Gamma_H}{\Gamma_L + \Gamma_H} \overset{\text{meas.}}{=} 0.0613 \pm 0.0059$$

- SM predictions

$$A_{\Delta\Gamma} \overset{\text{SM}}{=} 1$$

$$BF(B_s^0 \to \mu^+\mu^-) \overset{\text{SM}}{=} 3.57 \pm 0.30 \times 10^{-9}$$

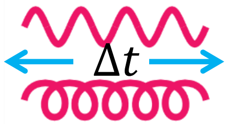
Updated from Buras et al. EPJ 2012

Uncertainty Budget	
F_{Bs}	72.5%
$\lvert V_{tb}^* V_{ts}\rvert$	22.8%
m_t	3.7%
τ_{Bs} and y_s	1.1%

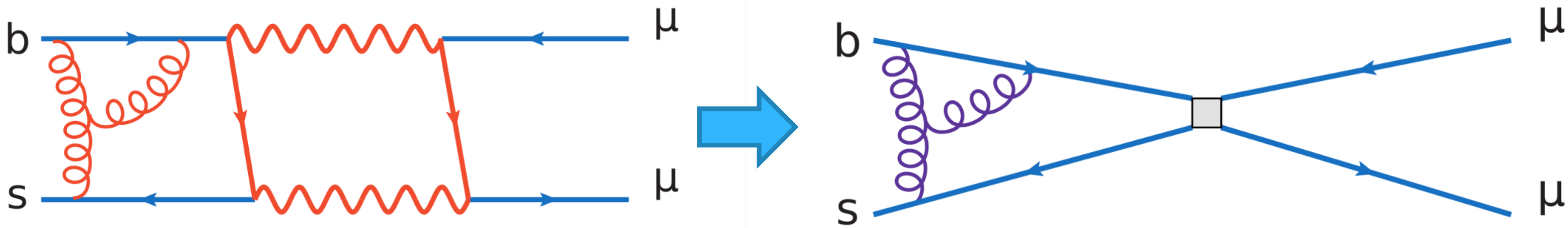
What About New Physics?

Model Independent Approach

- **Recall:** virtual particles are **off-shell**: $E^2 - p^2 \neq m^2_{shell}$
- Uncertainty principle, $\boldsymbol{\Delta t \propto 1\ /\Delta E}$ **:**



- **QCD** energy scale << **EW** energy scale, hence:



- **Energy scales separate** in amplitude expression:

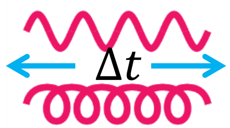
$$A(I \to F) \propto \sum_i \langle F|Q_i|I\rangle \times C_i$$

Matrix elements:
- encode **low energy**
- Main source of uncertainty

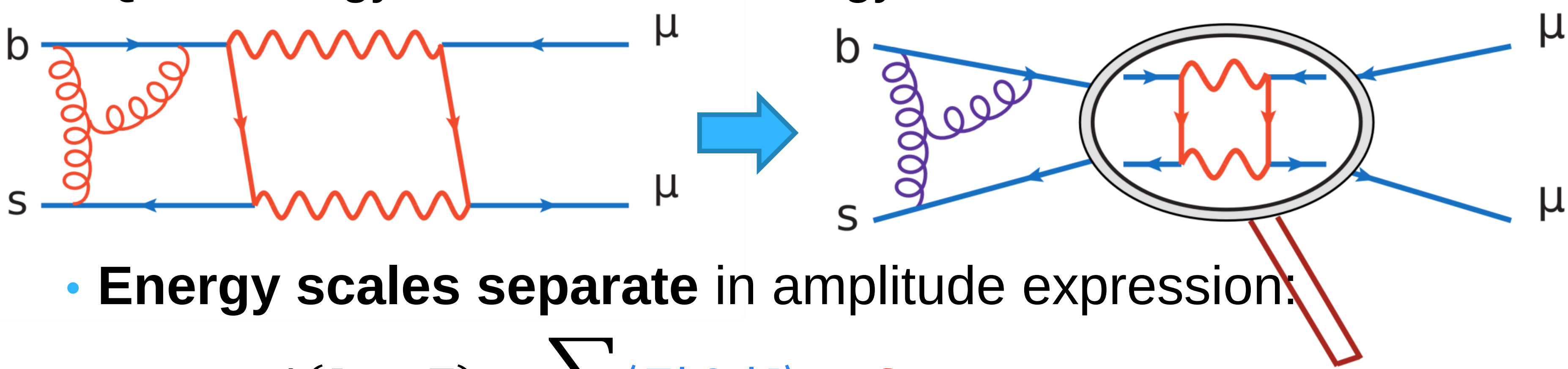
Wilson coefficents:
- 10 universal coefficients
- Encode **high energy**

Model Independent Approach

- **Recall:** virtual particles are **off-shell**: $E^2 - p^2 \neq m^2_{shell}$
- Uncertainty principle, $\boldsymbol{\Delta t \propto 1\,/\Delta E}$ **:**



- **QCD** energy scale << **EW** energy scale, hence:



- **Energy scales separate** in amplitude expression:

$$A(I \to F) \propto \sum_i \langle F|Q_i|I\rangle \times C_i$$

Matrix elements:

- encode **low energy**
- Main source of uncertainty

Wilson coefficents:

- 10 universal coefficients
- Encode **high energy**

New Physics in $BF(B^0_{(s)} \to \mu^+\mu^-)_{t=0}$

$$BF\left(B^0_{(s)}(t) \to \mu^+\mu^-\right)_{t=0}{}^{*}$$

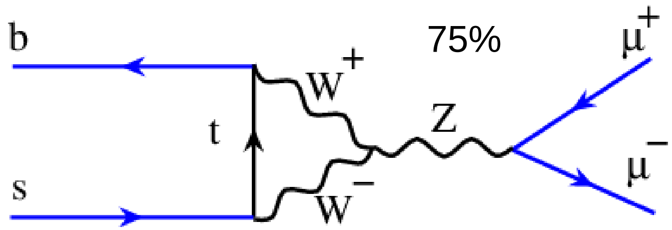
$$\propto \left(1 - \frac{4m_\ell^2}{m_B^2}\right) |C_S - C'_S|^2 + \left|(C_P - C'_P) + 2\frac{m_\mu}{m_B^2}(C_{10} - C'_{10})\right|^2$$

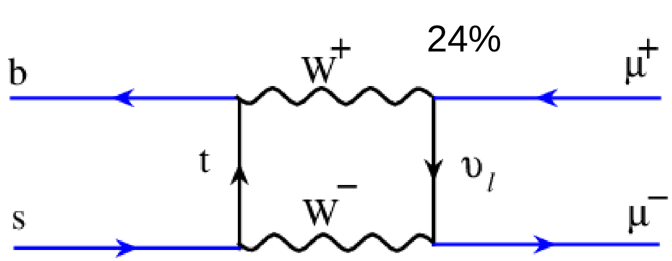
Helicity Supp.

Sensitive to (**P**seudo-)**S**calar NP

- Not affected by helicity supp.
- Scalar adds up with SM
- Pseudo-scalar can interfere destructively with SM

SM contributions:

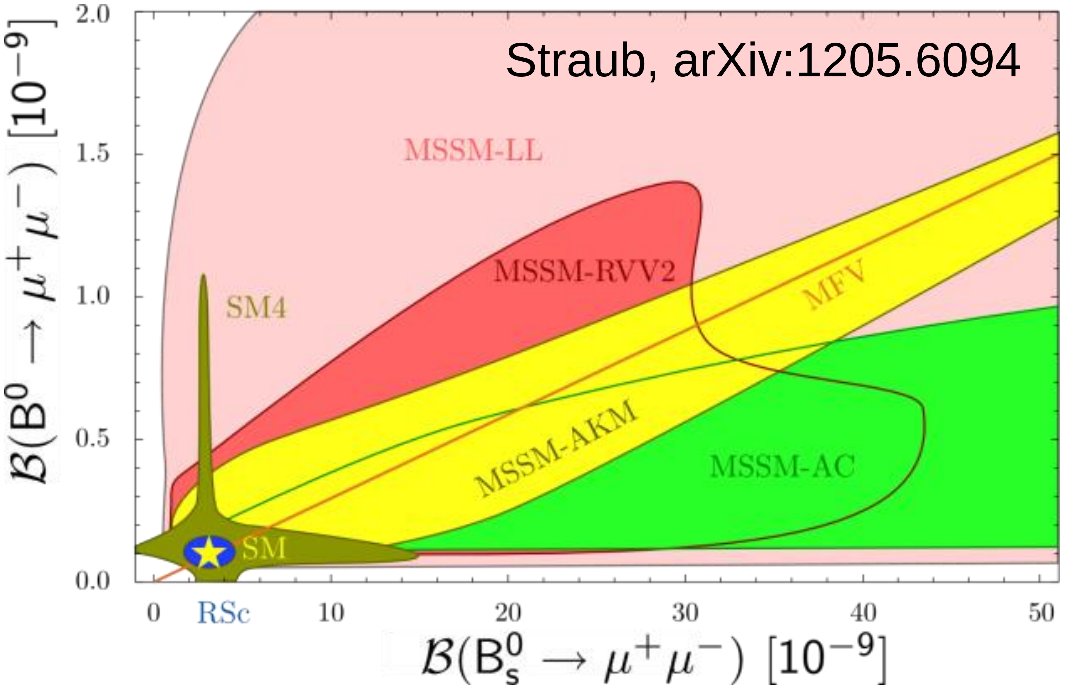




*$A_{\Delta\Gamma}$ phenomenology not accounted for

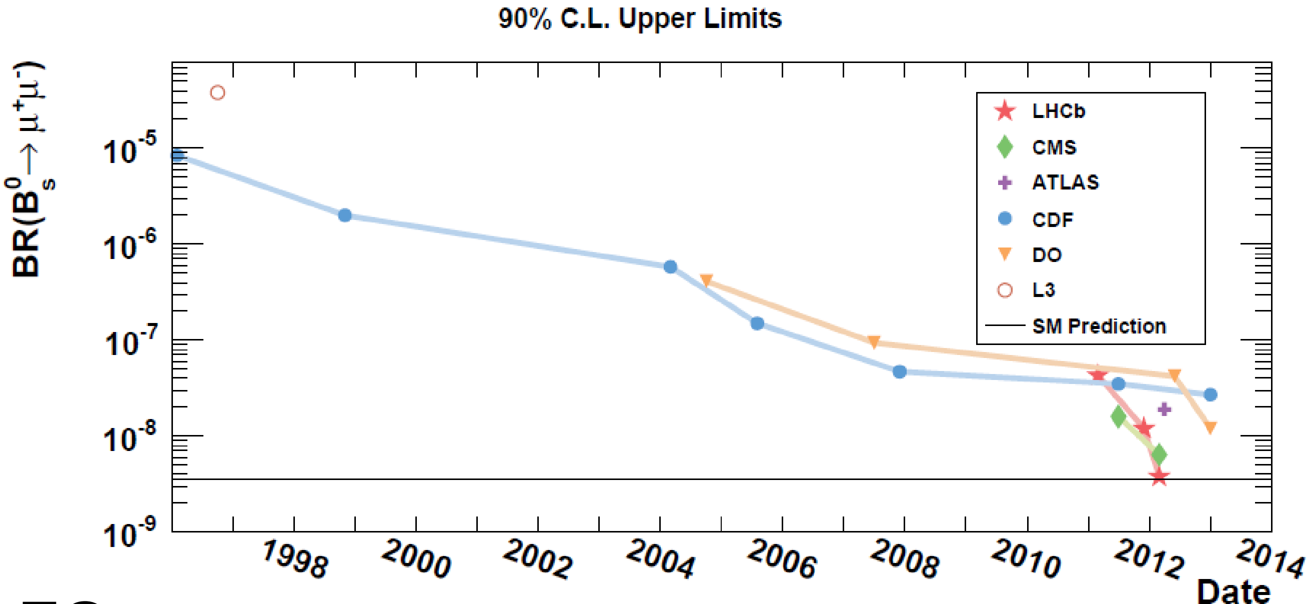
Which Models are Probed?

- $BF(B_s^0 \to \mu^+\mu^-)$:
  - Models with extended Higgs sector, e.g SuperSymmetry with large $\tan\beta$ as $C_{S,P} \propto \tan^3\beta$
  - Lepto-quarks
  - Z' models
- $BF(B^0 \to \mu^+\mu^-)$:
  - MFV hypothesis
  - Fourth generation

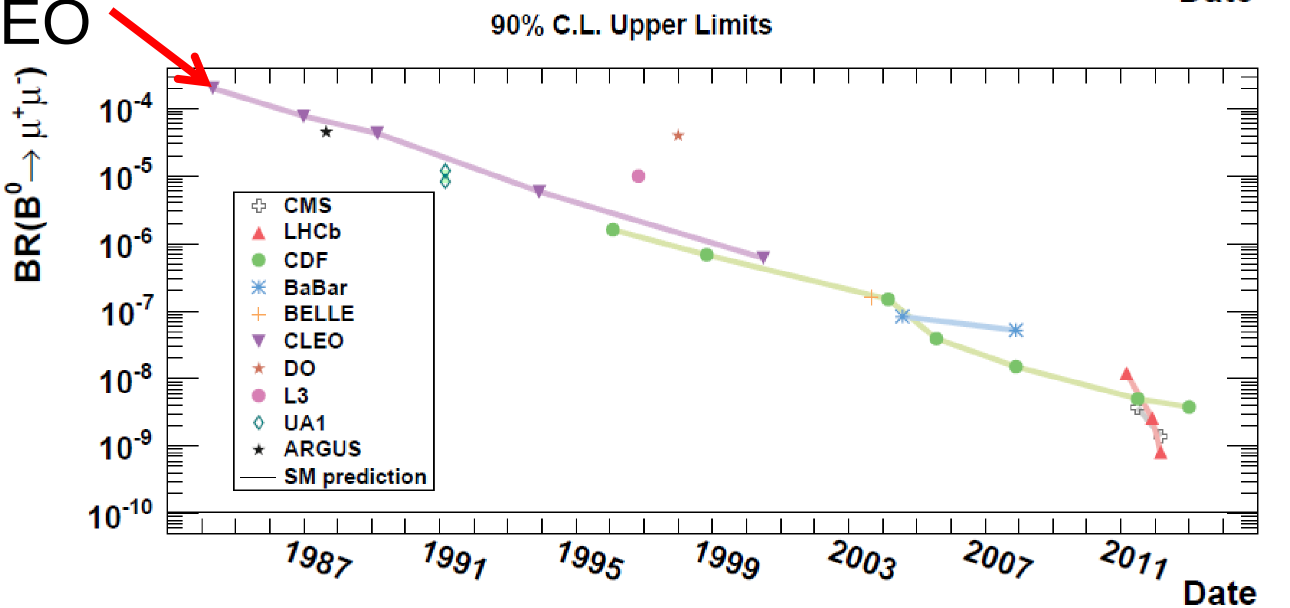


Experimental Picture

A long hunting…



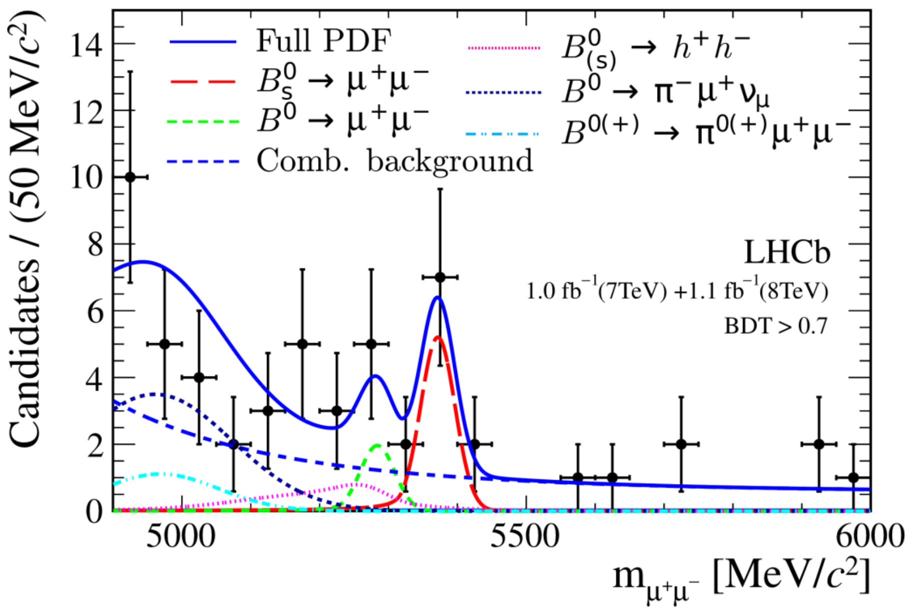
1984 CLEO



…before the first evidence!

- Nov 2012 LHCb: First evidence with 1 (7 TeV) + 1 (8 TeV) fb^{-1}

PRL 110, 021801, **2013**



$$B\left(B^0 \to \mu^+\mu^-\right) < 9.4 \times 10^{-10}$$ at 95%CL

$$B\left(B_s^0 \to \mu^+\mu^-\right) = 3.2^{+1.5}_{-1.2} \times 10^{-9}$$

Significance of 3.5 σ !

- **Today:** update with the full dataset: 1 (7 TeV) + 2 (8 TeV) fb^{-1}
- All data consistently reprocessed
- Data in $m(B^0_{(s)}) \pm 60$ MeV/c^2 are blind til analysis completion.

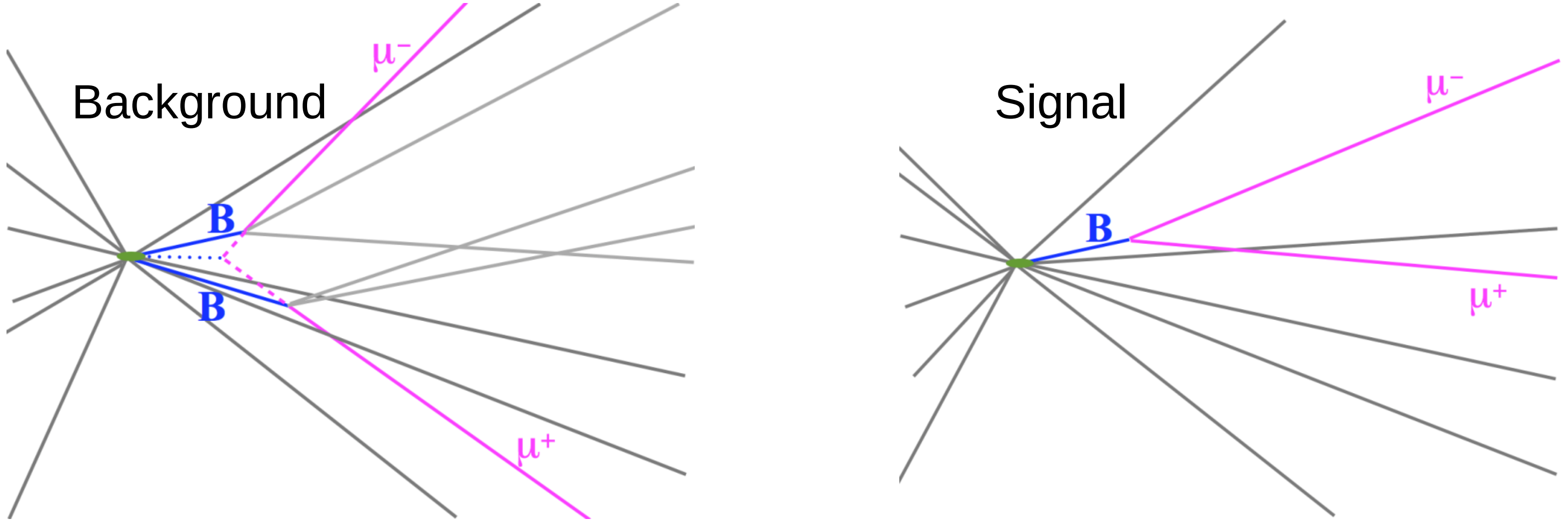
Outlines

- Motivations to search for $B^0_{(s)} \to \mu^+\mu^-$
- Searching for $\mathrm{B}^0_{(s)} \to \mu^+\mu^-$ at LHCb
- Combination with CMS
- Conclusions

2 Key Points for $B^0_{(s)} \to \mu^+\mu^-$ with LHCb

1. Production of $B^0_{(s)}$ mesons: (x-section and trigger)
2. Separation Signal/Background (detector performance)

- Combinatorial background: $b\bar{b} \to \mu\mu X$

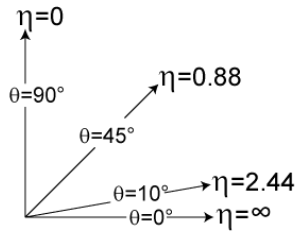


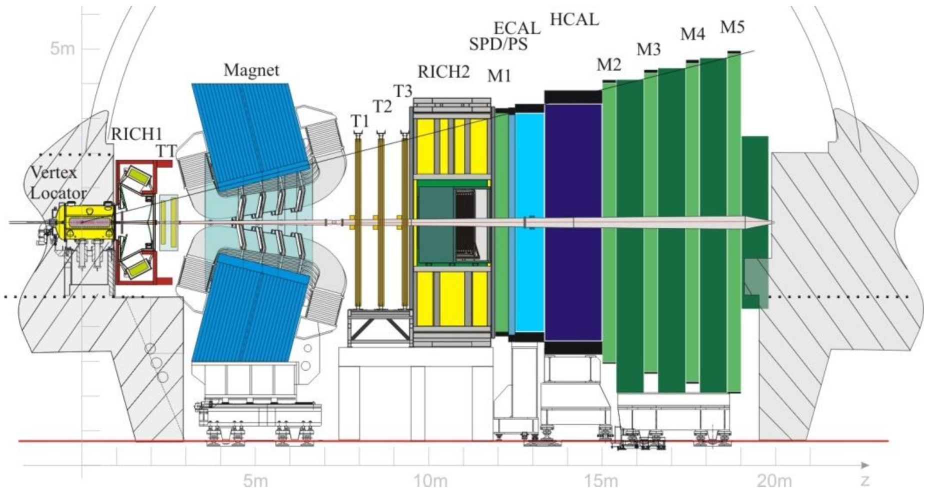
- Physical backgrounds:

e.g. $B \to K\pi, KK, \pi\pi$ where K, π decay in flight to μ

Key Point 1: $B^0_{(s)}$ production

- LHCb: a single arm forward spectrometer: $b\bar{b}$ are produced forward
- Instantaneous luminosity $4 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$





Exp.	Accept.	$\sigma_{b\bar{b}}$	$b\bar{b}$ pairs
ATLAS CMS	$\|\eta\| < 2.2$	$\sim 75\mu b$	$\sim 4 \times 10^{11}$
LHCb	$2 < \eta < 6$	$\sim 94\mu b$	$\sim 9 \times 10^{10}$

Albrecht, arXiv:1207.4287

- Not all $p\bar{p}$ interactions produce B: $\sigma_{b\bar{b}} \approx 1\% \, \sigma_{tot}$
- Record only interesting events with a trigger system

Trigger

Trigger is a 2-level system:

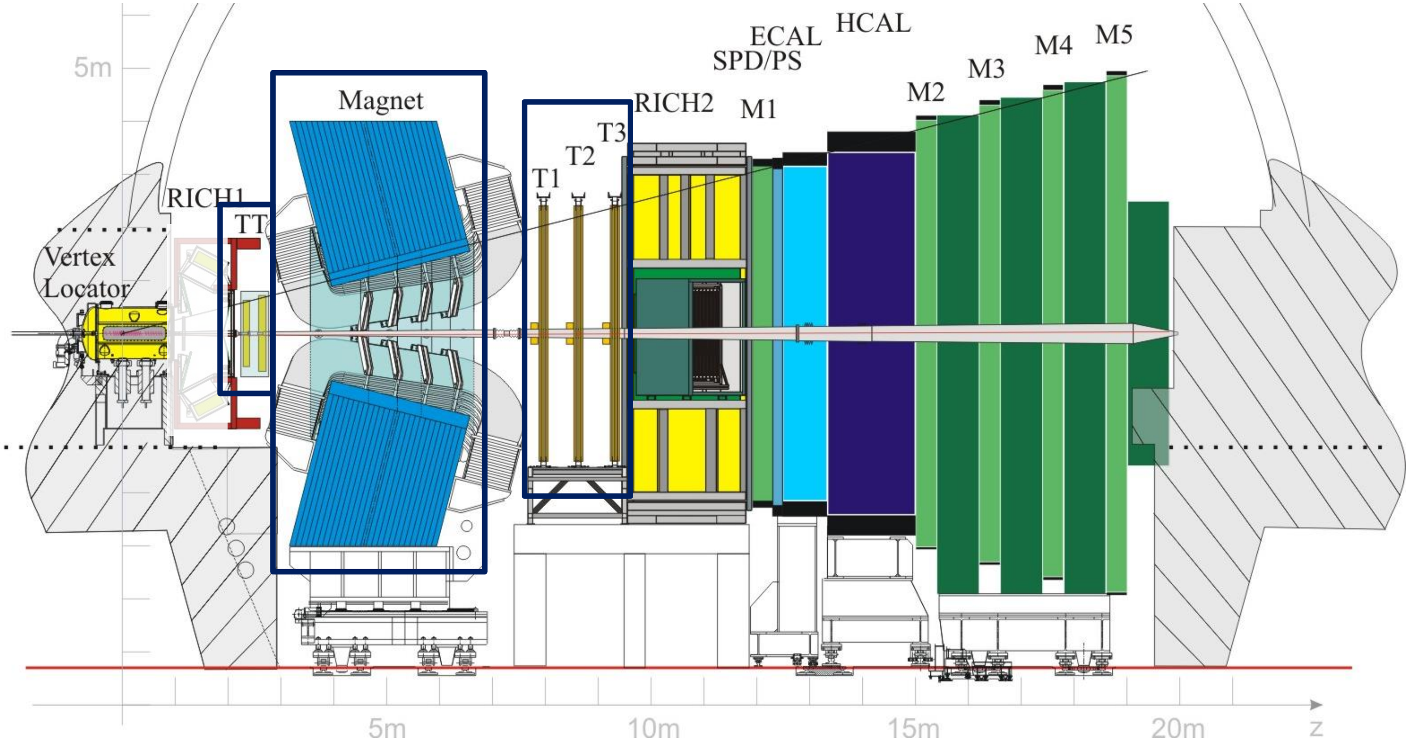
- L0 Trigger:
  - Made with custom electronics
  - Use fast available sub-detectors information
  - Rate reduction: 40 MHz → 1 MHz
- HLT Trigger
  - Software trigger
  - Refine selection based on partial reconstructions
  - Rate reduction : 1 MHz → 2-6 kHz

- Trigger Efficiency for $B_s^0 \to \mu^+\mu^-$: 90%

Key Point 2: Separation Sig/Bkg

Signal separated from combinatorial Bkg thanks to:

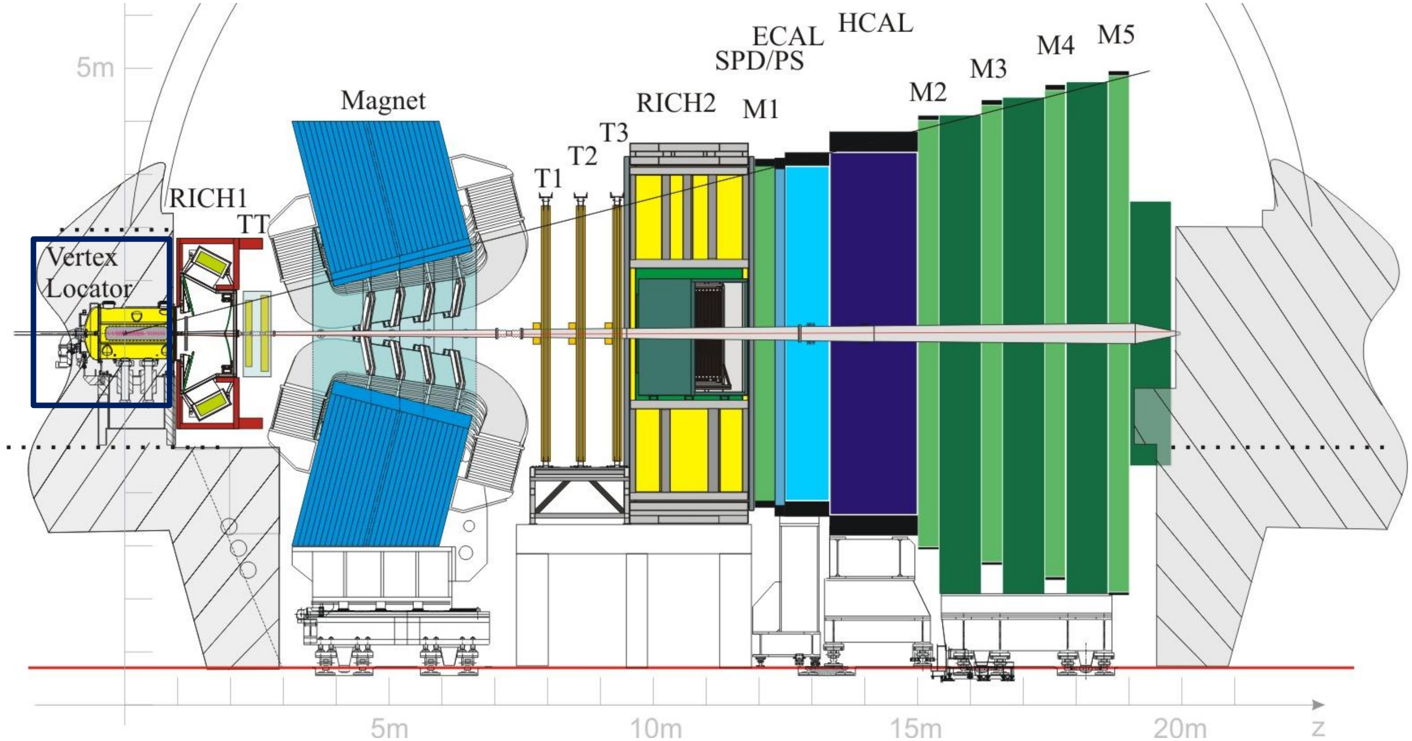
- Mass and momentum resolution (magnet, tracking)
- $\frac{\delta p}{p} \sim 0.4 \rightarrow 0.6\%$ for $p = 5 \rightarrow 500\ GeV/c$
- $\Delta m_{\mu\mu} \sim 25 \mathrm{MeV/c^2}$ (2 [3-4] times better than CMS [ATLAS])



Key Point 2: Separation Sig/Bkg

Signal separated from combinatorial Bkg thanks to:

- Secondary vertex resolution: (high boost and tracking)
- B average flight distance $10\ mm$
- $\sigma_{IP} = 25\mu m$ at $p_t = 2\mathrm{GeV/c}$



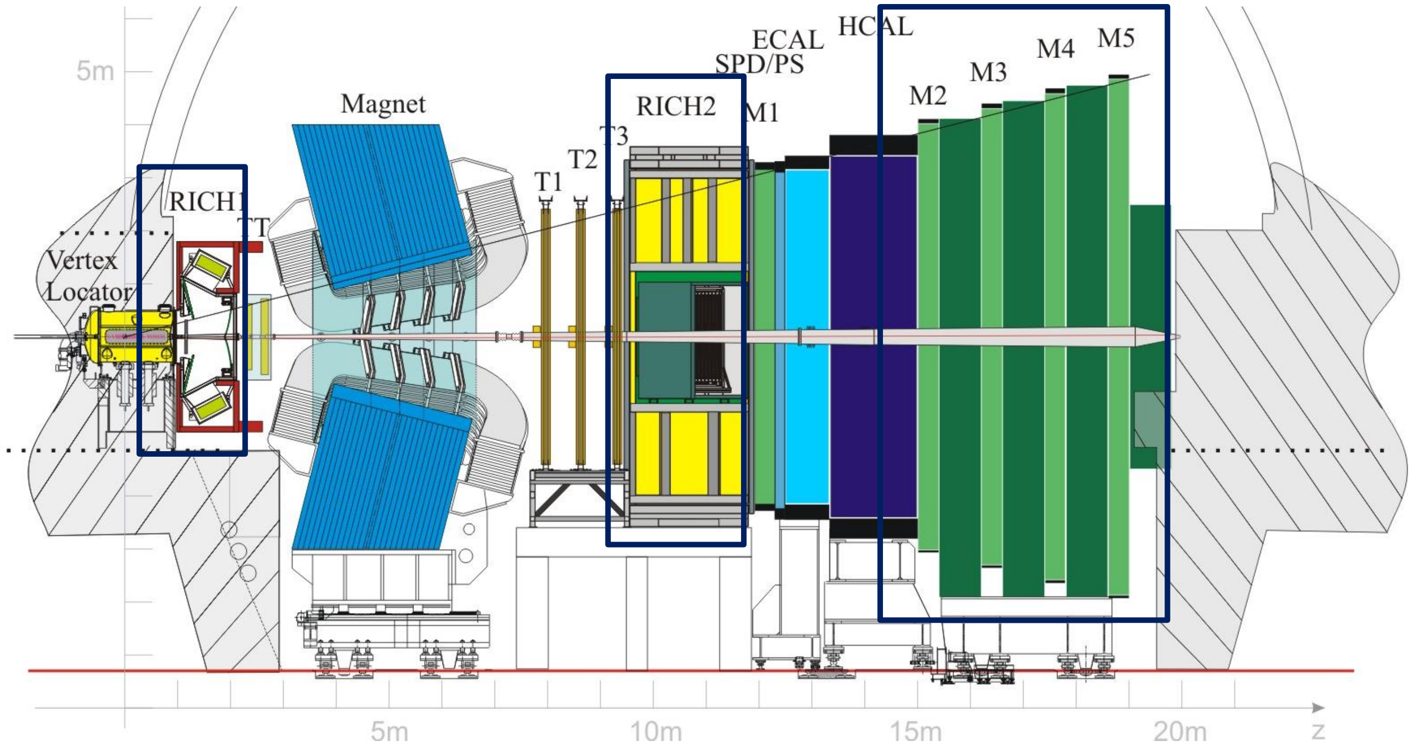
Key Point 2: Separation Sig/Bkg

Signal separated from physical Bkg thanks to:

Particle identification info (RICH – muons chambers)

$\epsilon(\mu \rightarrow \mu) \sim 98\%$ $\epsilon(K \rightarrow \mu) \sim 0.3\%$

$\epsilon(\pi \rightarrow \mu) \sim 0.6\%$ $\epsilon(p \rightarrow \mu) \sim 0.3\%$



LHCb Analysis

Analysis Overview

- Answering the question:

  Which BF is (in-)compatible with the observed data?
- Observed data = $N(B^0_{(s)} \to \mu^+\mu^- + \text{Bkg})$

1. Increasing sensitivity:
   - Events Selection and Classification in categories
2. Derive Signal and Bkg expectations from control channels
3. Compare expectations with observation and infer about BF
   - CLs method and Fitting procedure if signal evidence

- **Data Set:**

2011	$1.0\ \text{fb}^{-1}$	7TeV	
2012	$+1.1\ \text{fb}^{-1}$	8TeV	PRL 110, 021801, 2013
2012	$+0.9\ \text{fb}^{-1}$	8TeV	PRL 111, 101805, 2013

Analysis Overview

- Answering the question:

  Which BF is (in-)compatible with the observed data?
- Observed data = $N(B^0_{(s)} \rightarrow \mu^+\mu^- + \mathrm{Bkg})$

1. Increasing sensitivity:
   - Events Selection and Classification in categories
2. Derive Signal and Bkg expectations from control channels
3. Compare expectations with observation and infer about BF
   - CLs method and Fitting procedure if signal evidence

- **Data Set:**

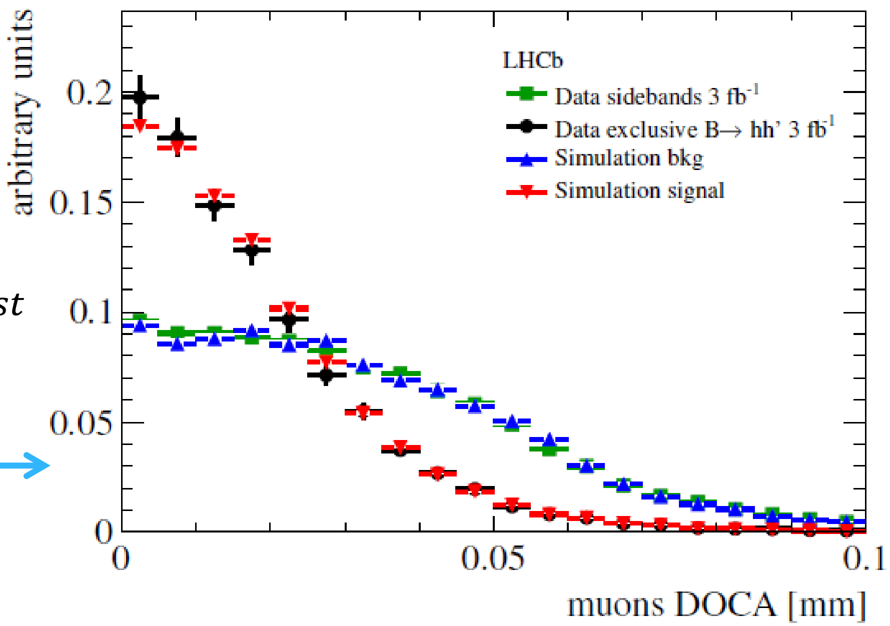
2011	$1.0\ \mathrm{fb}^{-1}$	7TeV	
2012	$+1.1\ \mathrm{fb}^{-1}$	8TeV	PRL 110, 021801, 2013
2012	$+0.9\ \mathrm{fb}^{-1}$	8TeV	PRL 111, 101805, 2013

Selection and Classification

- After loose selection, classify evt based on $m_{\mu\mu}$ and geometry
- Train Boosted Decision Tree (BDT) to recognize signal from combinatorial background based on 12 geometrical variables

B candidate:
- proper time
- impact parameter
- transverse momentum
- B isolation
- Angle between $\vec{p}_B$ and $\vec{p}_{thrust}$
- Angle in B rest frame between p_{μ^+} and $\vec{p}_{thrust}$

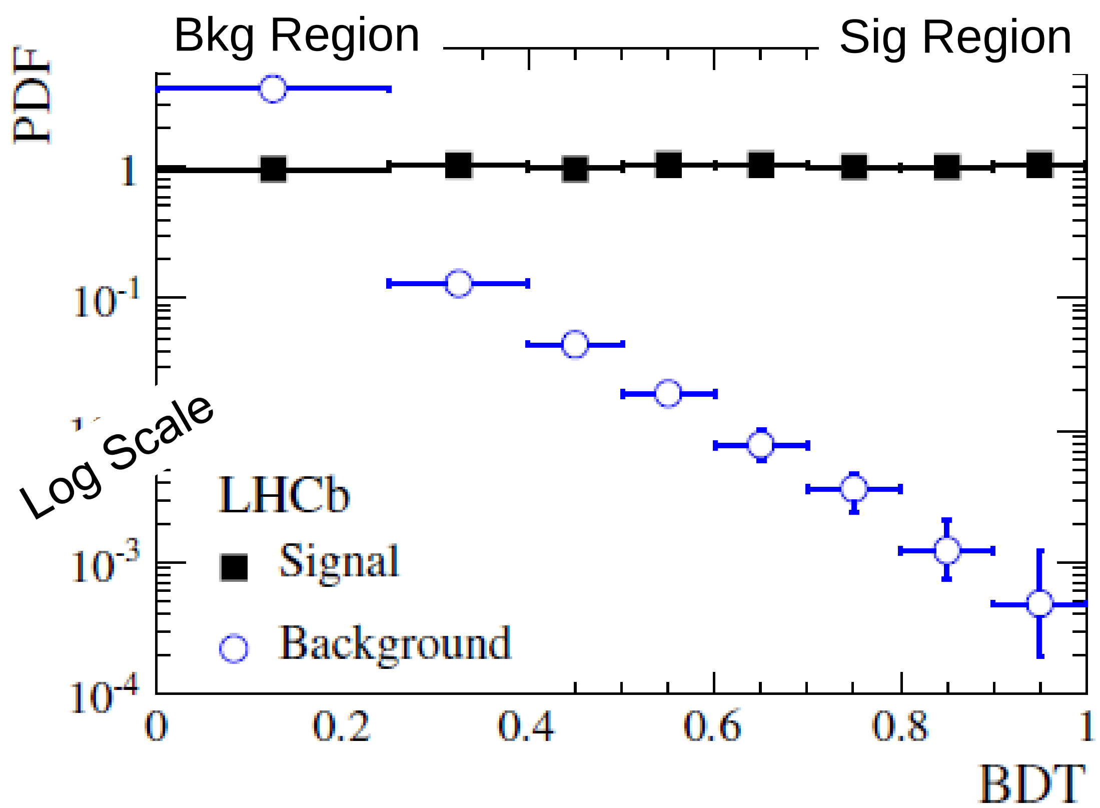
Muons:
- min IP significance
- distance of closest approach
- muon isolation
- polarisation angle
- $|\eta(\mu_1)-\eta(\mu_2)|$
- $|\varphi(\mu_1)-\varphi(\mu_2)|$



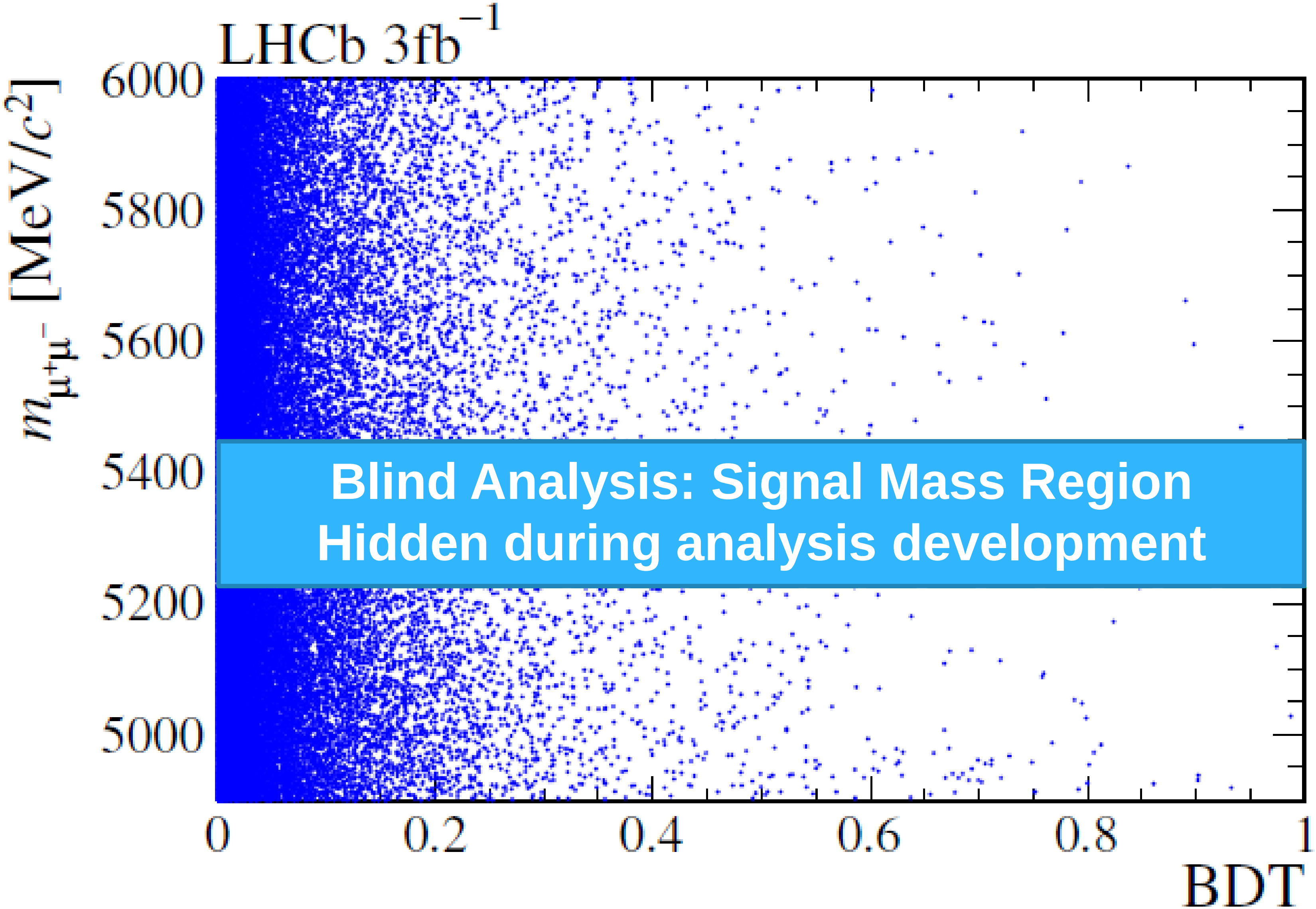
$\vec{p}_{thrust}$ is the sum of momenta of all tracks consistent with originating from the decay of the other b hadron

Selection and Classification

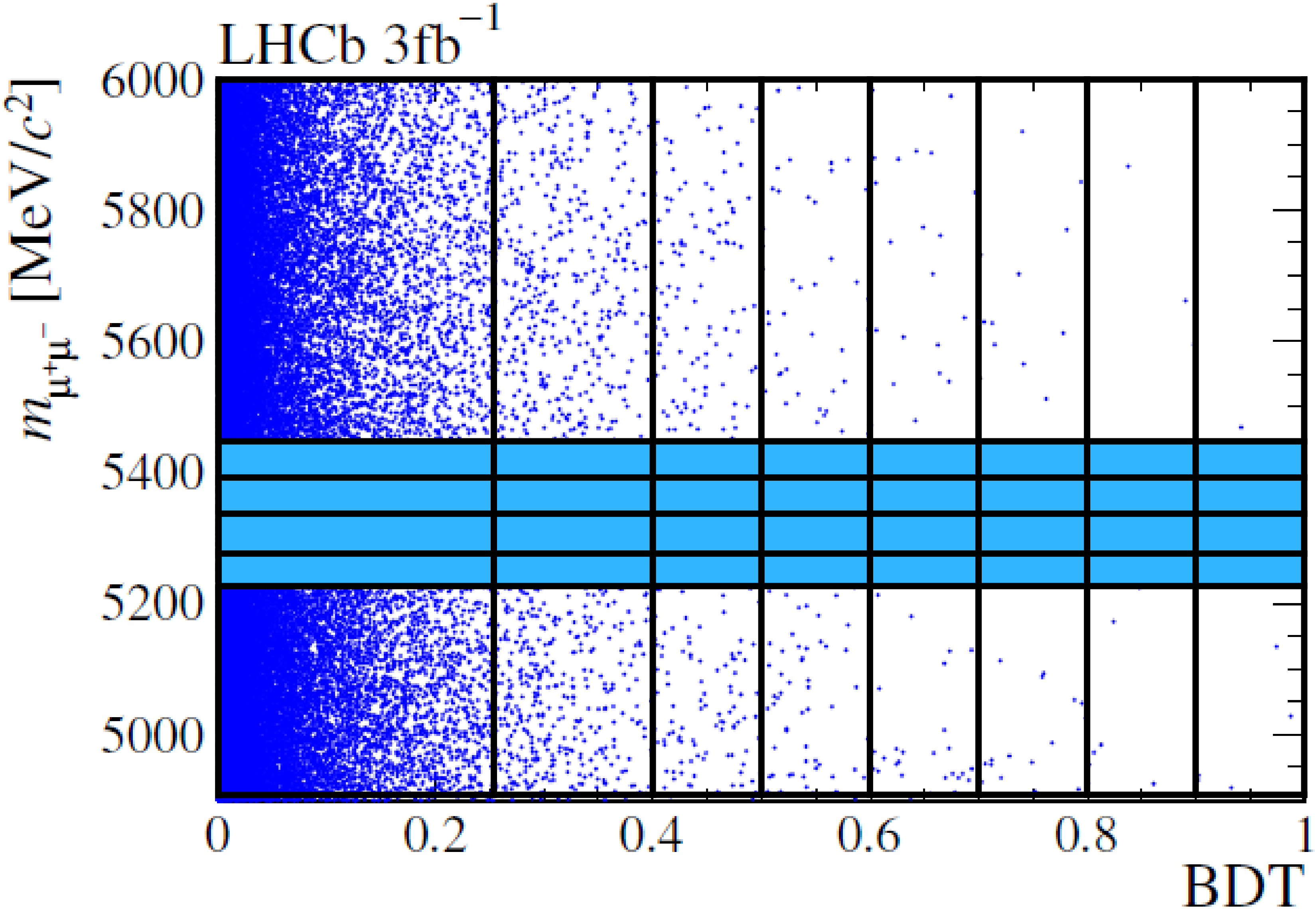
- After loose selection, classify evt based on $m_{\mu\mu}$ and geometry
- Train Boosted Decision Tree (BDT) to recognize signal from combinatorial background based on 12 geometrical variables



2-D Binned Classification



2-D Binned Classification



Analysis Overview

- Answering the question:

  Which BF is (in-)compatible with the observed data?

- Observed data = $N(B^0_{(s)} \to \mu^+\mu^- + \mathrm{Bkg})$

1. Increasing sensitivity:
   - Events Selection and Classification in categories
2. Derive Signal and Bkg expectations from control channels
3. Compare expectations with observation and infer about BF
   - CLs method and Fitting procedure if signal evidence

- **Data Set:**

2011	$1.0\ \mathrm{fb}^{-1}$	7TeV	
2012	$+1.1\ \mathrm{fb}^{-1}$	8TeV	PRL 110, 021801, 2013
2012	$+0.9\ \mathrm{fb}^{-1}$	8TeV	PRL 111, 101805, 2013

Bkg Expectations

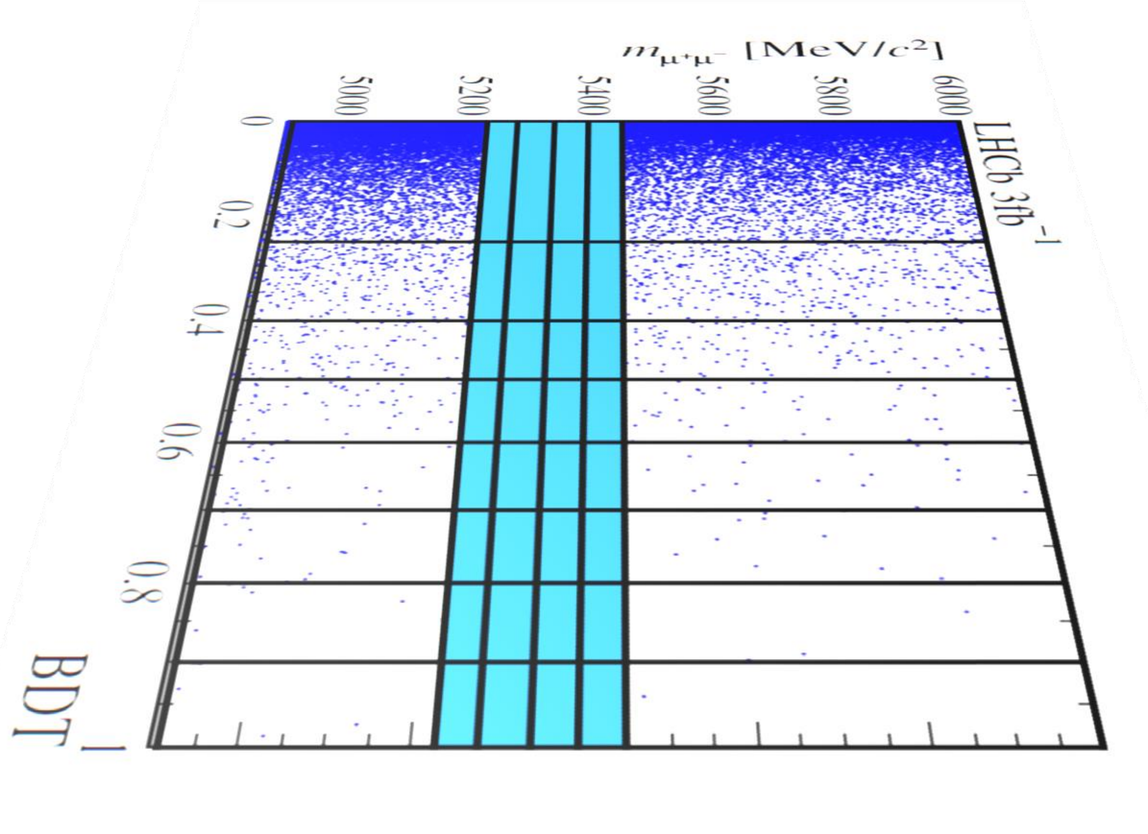
- Extrapolate Bkg from side-bands
- Fit components:
  - Combinatorial
  - MisId

    $B_s^0 \to h^+ h'^-$

    $B^0 \to \pi, K^+ \mu^- \overline{\nu}_\mu$
  - Partially Reco

    $B^{0,+} \to \pi^{0,+} \mu^+ \mu^-$

    $\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$
  - **Total**



Bkg Expectations

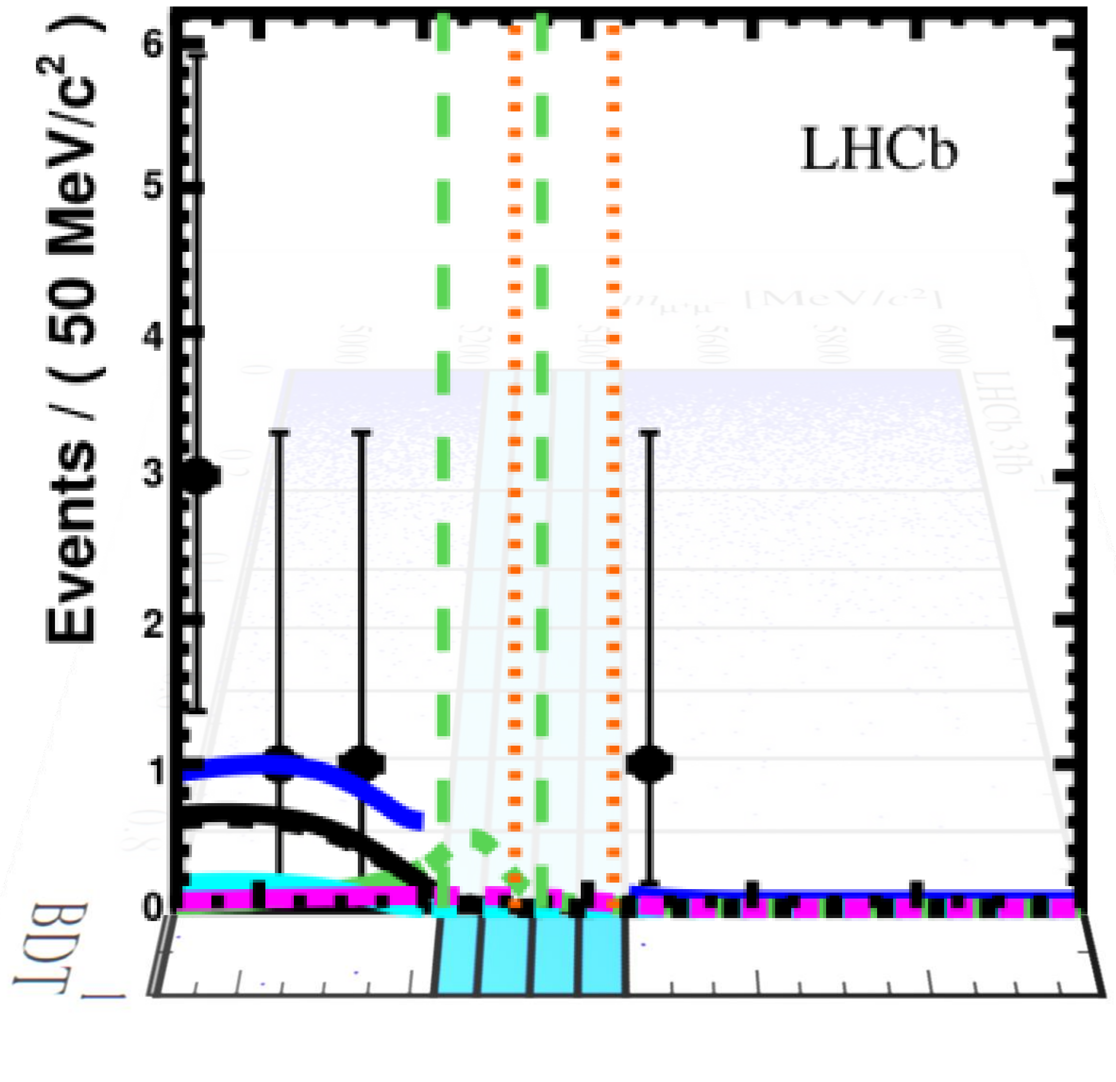
- Extrapolate Bkg from side-bands
- Fit components:
  - Combinatorial
  - MisId

    $B_s^0 \to h^+ h'^-$

    $B^0 \to \pi, K^+ \mu^- \overline{\nu}_\mu$
  - Partially Reco

    $B^{0,+} \to \pi^{0,+} \mu^+ \mu^-$

    $\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$
  - **Total**



Bkg Expectations

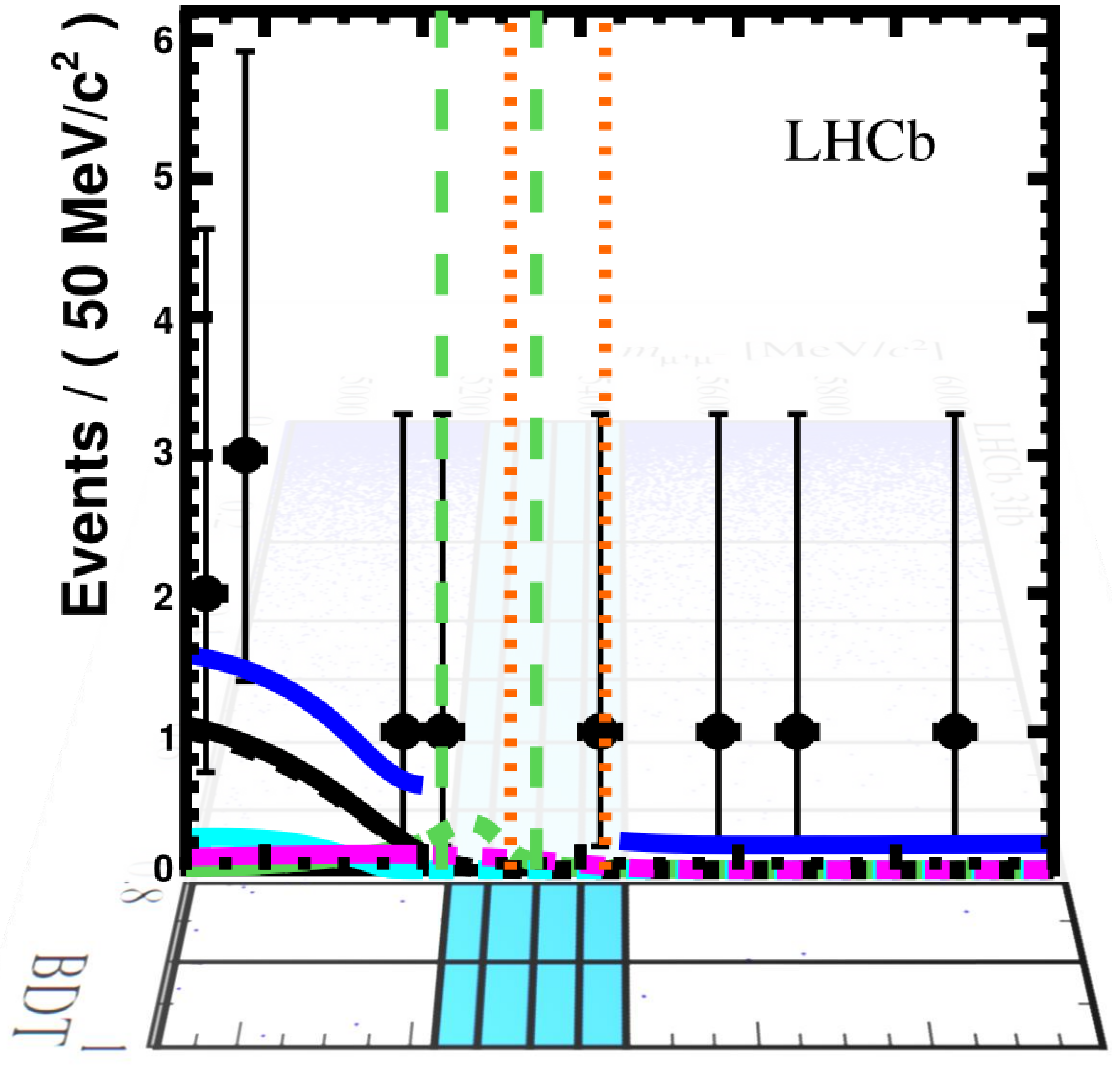
- Extrapolate Bkg from side-bands
- Fit components:
  - Combinatorial
  - MisId

  $B_s^0 \to h^+ h'^-$

  $B^0 \to \pi, K^+ \mu^- \overline{\nu}_\mu$
  - Partially Reco

  $B^{0,+} \to \pi^{0,+} \mu^+ \mu^-$

  $\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$
  - **Total**



Bkg Expectations

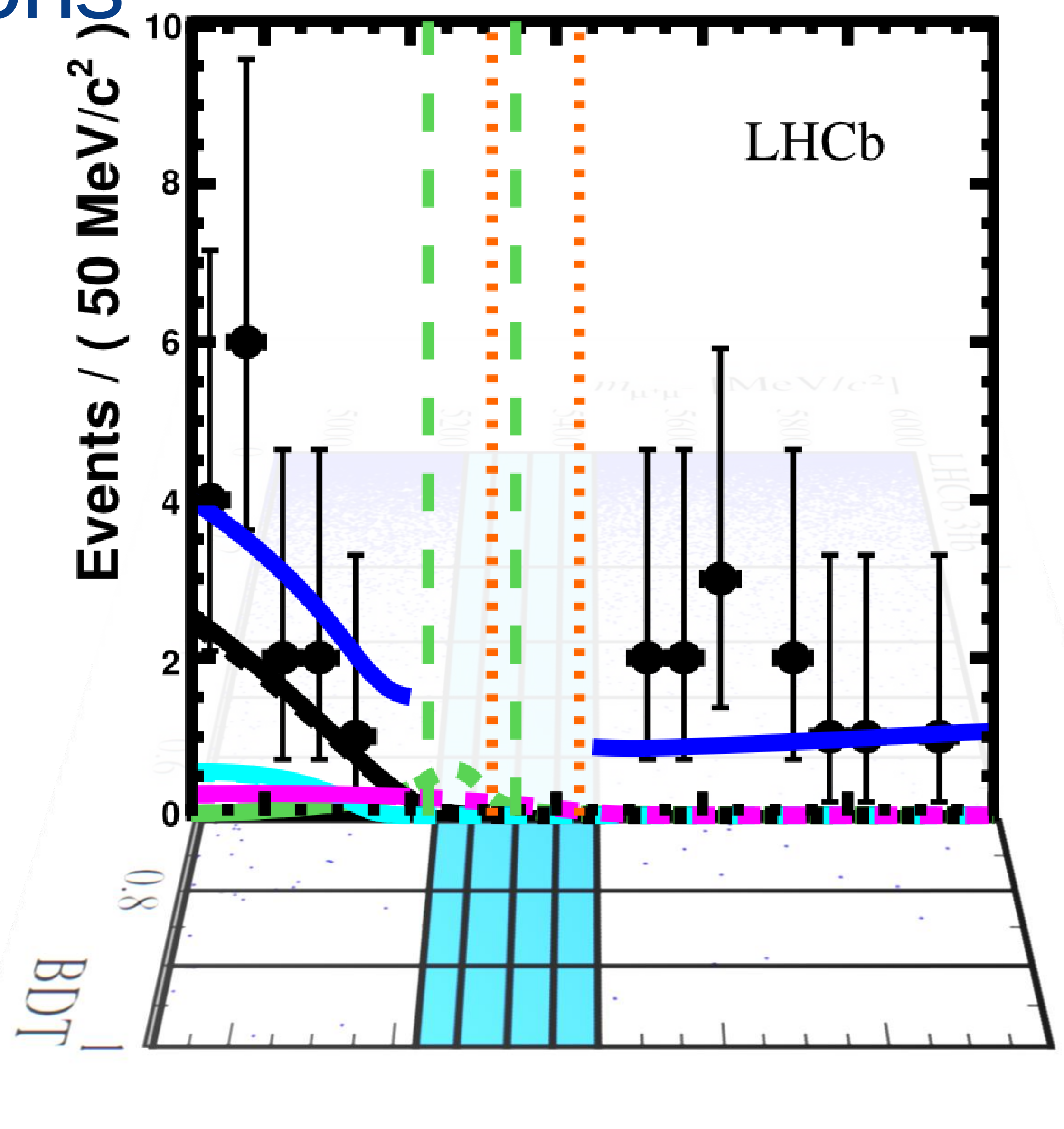
- Extrapolate Bkg from side-bands
- Fit components:
  - Combinatorial
  - MisId

    $B_s^0 \to h^+ h'^-$

    $B^0 \to \pi, K^+ \mu^- \overline{\nu}_\mu$
  - Partially Reco

    $B^{0,+} \to \pi^{0,+} \mu^+ \mu^-$

    $\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$
  - **Total**



Bkg Expectations

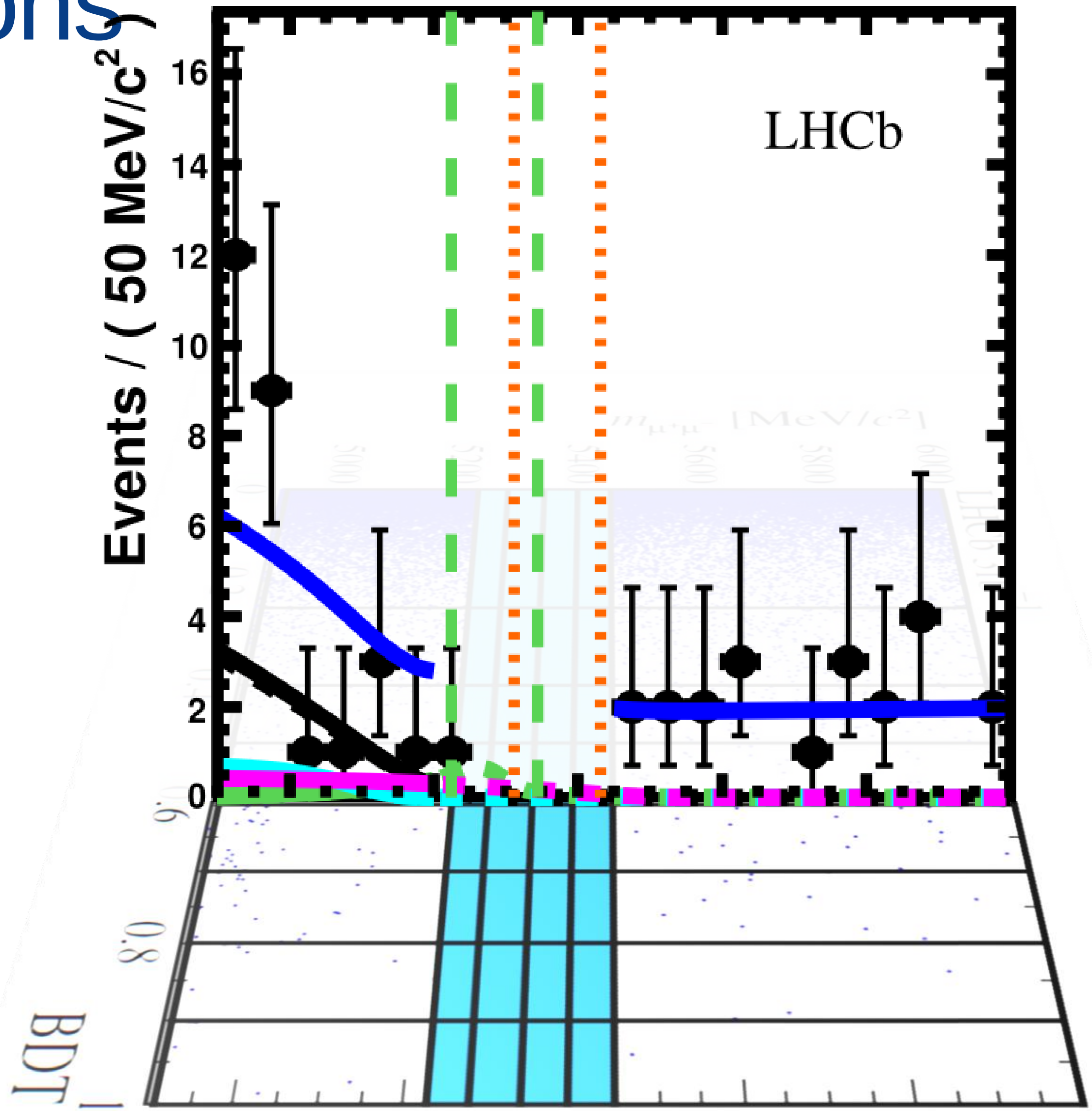
- Extrapolate Bkg from side-bands
- Fit components:
  - Combinatorial
  - MisId

    $B_s^0 \to h^+ h'^-$

    $B^0 \to \pi, K^+ \mu^- \overline{\nu}_\mu$
  - Partially Reco

    $B^{0,+} \to \pi^{0,+} \mu^+ \mu^-$

    $\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$
  - **Total**



Bkg Expectations

- Extrapolate Bkg from side-bands
- Fit components:
  - Combinatorial
  - MisId

$B_s^0 \to h^+ h'^-$

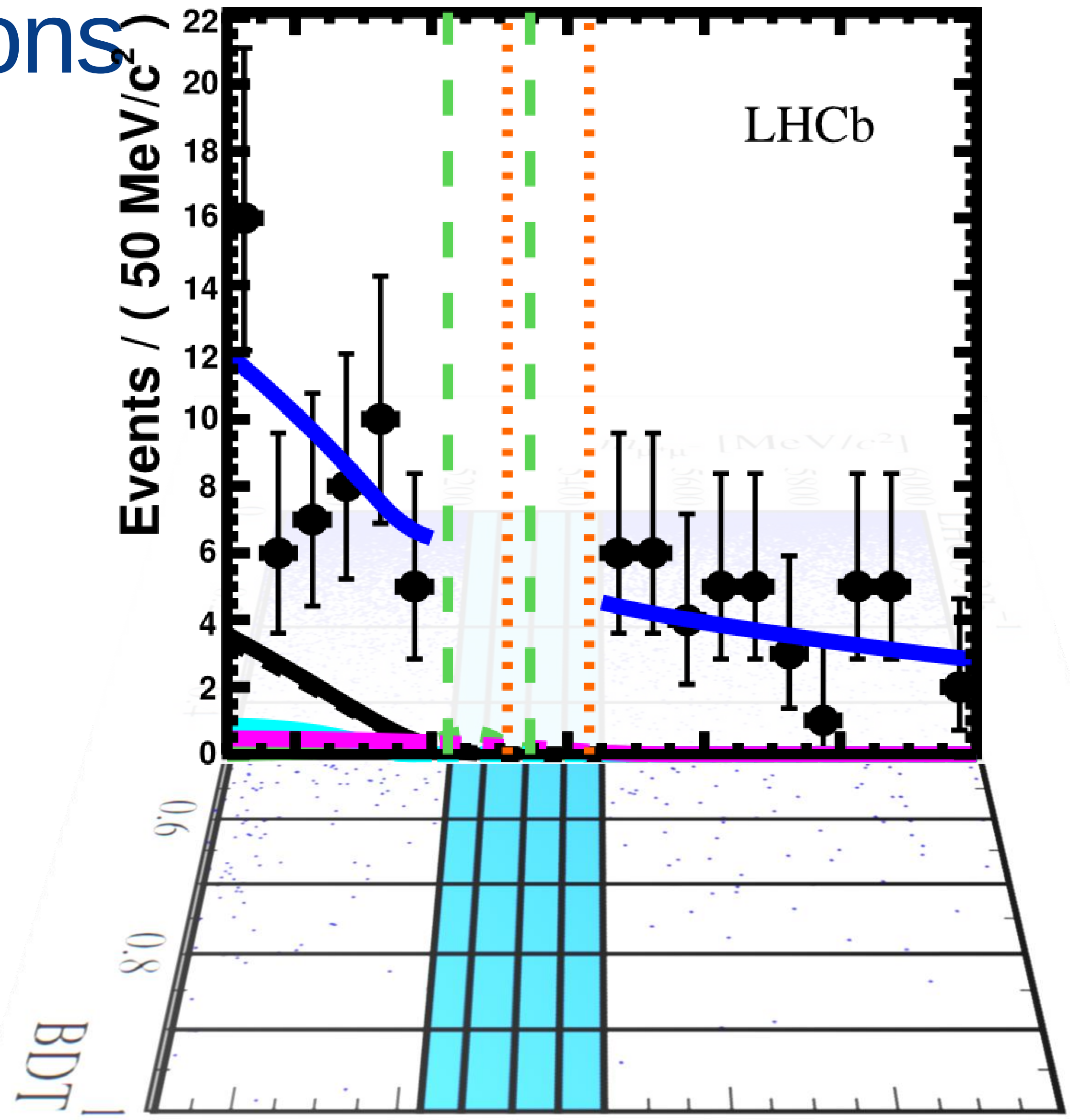
$B^0 \to \pi, K^+ \mu^- \overline{\nu}_\mu$

  - Partially Reco

$B^{0,+} \to \pi^{0,+} \mu^+ \mu^-$

$\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$

  - **Total**



Bkg Expectations

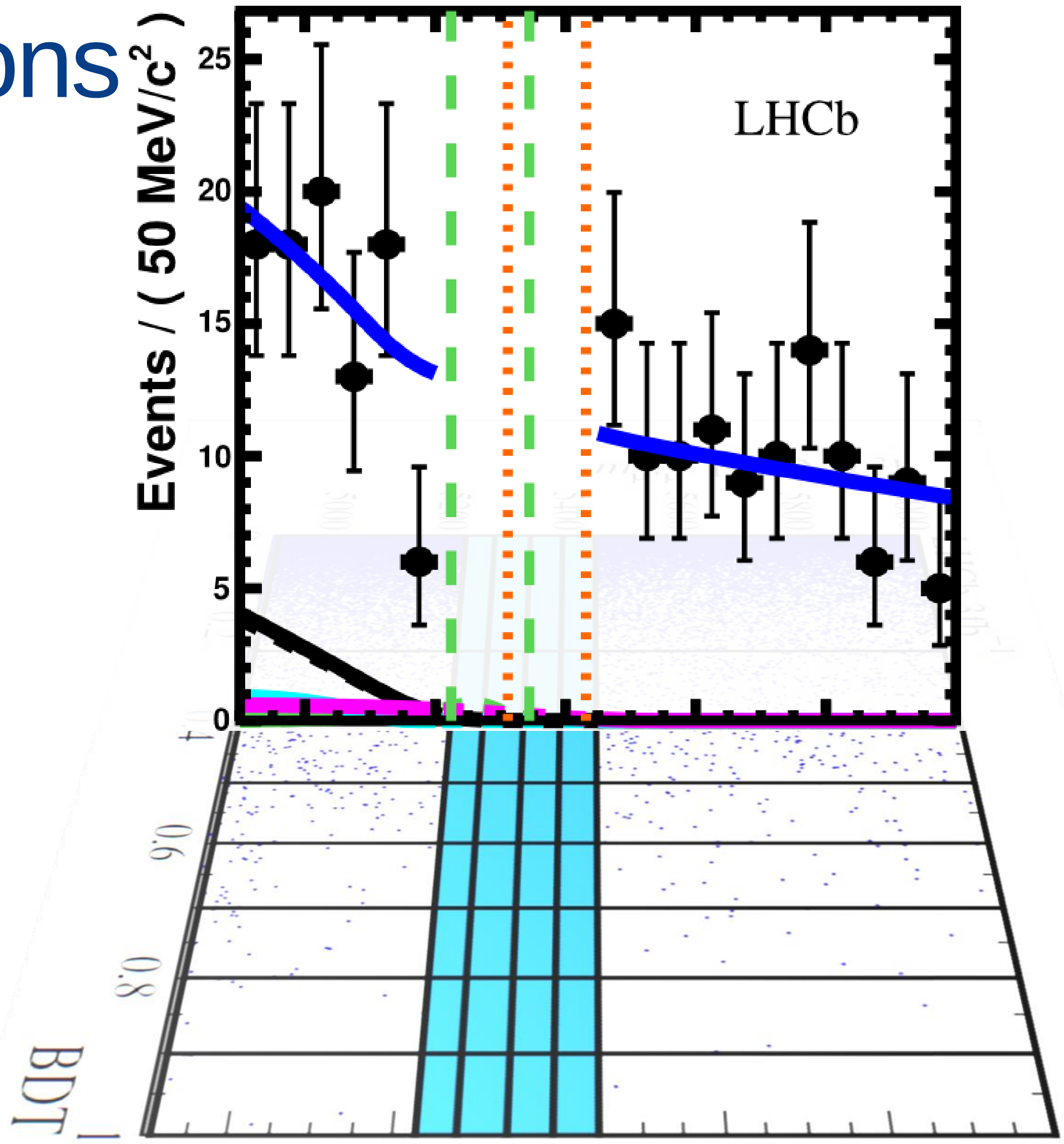
- Extrapolate Bkg from side-bands
- Fit components:
  - Combinatorial
  - MisId

    $B_s^0 \rightarrow h^+ h'^-$

    $B^0 \rightarrow \pi, K^+ \mu^- \overline{\nu}_\mu$
  - Partially Reco

    $B^{0,+} \rightarrow \pi^{0,+} \mu^+ \mu^-$

    $\Lambda_b^0 \rightarrow p \mu^- \overline{\nu}_\mu$
  - **Total**



Bkg Expectations

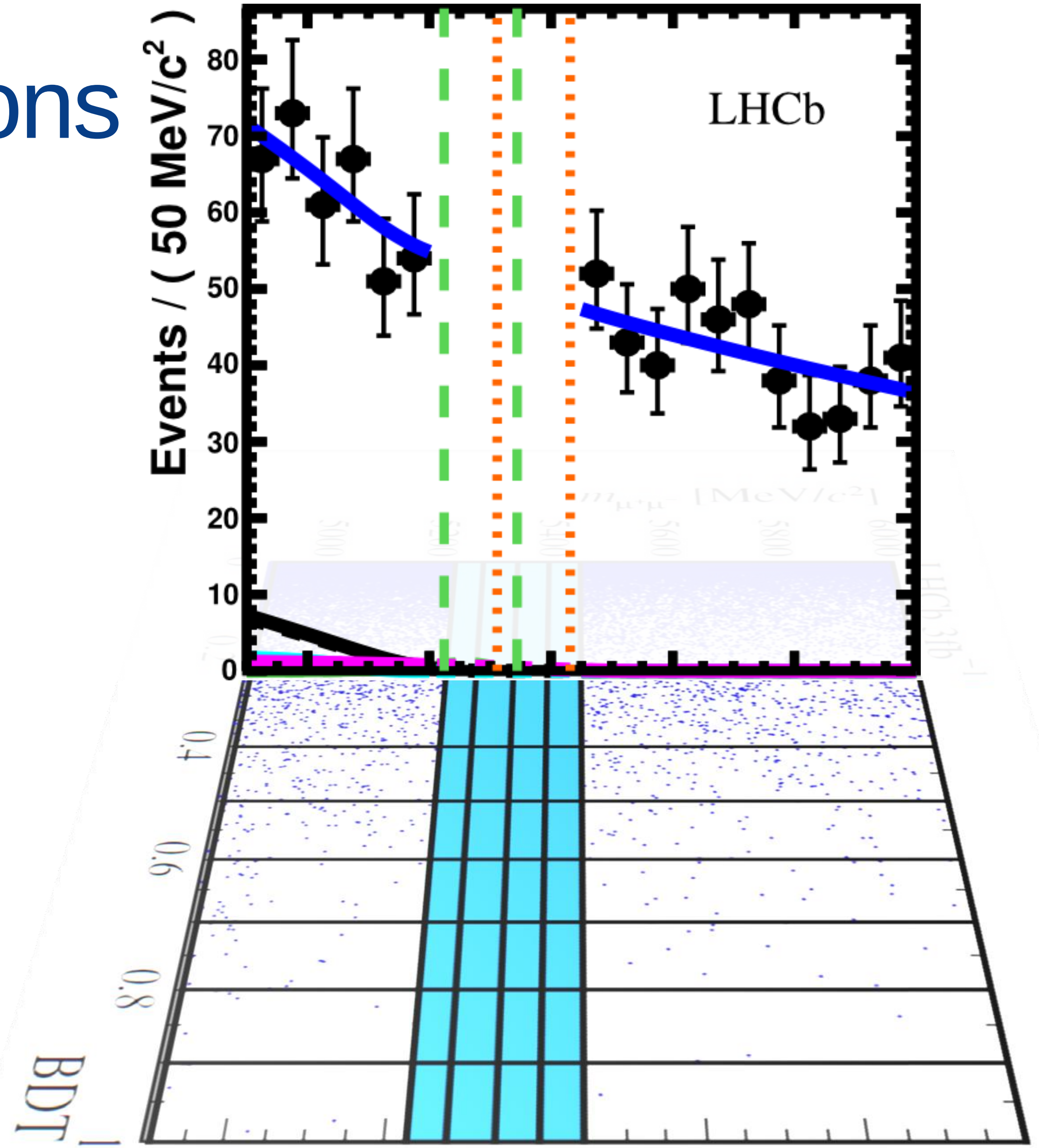
- Extrapolate Bkg from side-bands
- Fit components:
  - Combinatorial
  - MisId

    $B_s^0 \to h^+ h'^-$

    $B^0 \to \pi, K^+ \mu^- \overline{\nu}_\mu$
  - Partially Reco

    $B^{0,+} \to \pi^{0,+} \mu^+ \mu^-$

    $\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$
  - **Total**



Bkg Expectations

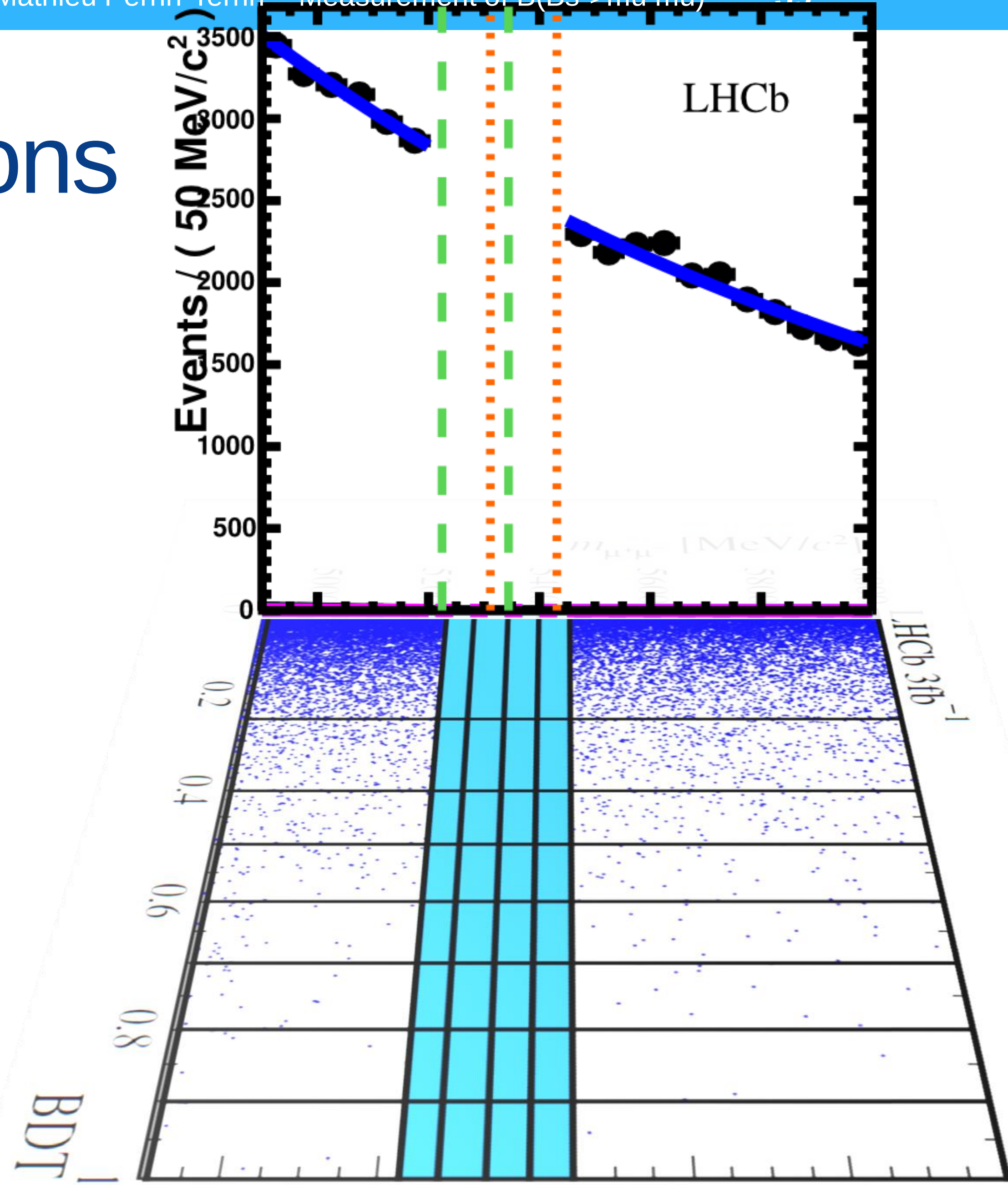
- Extrapolate Bkg from side-bands
- Fit components:
  - Combinatorial
  - MisId

    $B_s^0 \to h^+ h'^-$

    $B^0 \to \pi, K^+ \mu^- \overline{\nu}_\mu$
  - Partially Reco

    $B^{0,+} \to \pi^{0,+} \mu^+ \mu^-$

    $\Lambda_b^0 \to p \mu^- \overline{\nu}_\mu$
  - **Total**



Misidentified Backgrounds - $B^0_{(s)} \to h^+ h'^-$

1. MisID probabilities measured on data in bins of p and p_T
   - $\pi \to \mu$ and $K \to \mu$ measured with $D^0 \to K^+\pi^-$ from $D^{*+} \to D^0\pi^+$,
   - $p \to \mu$ measured with $\Lambda \to p\,\pi^-$

2. Probabilities then convoluted with $B^0_{(s)} \to h^+ h'^-$ MC spectra to get the average double misID efficiency $\varepsilon_{\mu\mu \to hh}$ (~10^{-5})

3. Rate obtained applying $\varepsilon_{\mu\mu \to hh}$ to the $B^0_{(s)} \to h^+ h'^-$ yield

4. Mass shape is evaluated from MC

5. $B^0_{(s} \to h^+ h'^-$ included as fit component with rate constrained to expected yield

Other exclusive backgrounds

- Yields obtained by normalising to $B^+ \to J/\psi\, K^+$

	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

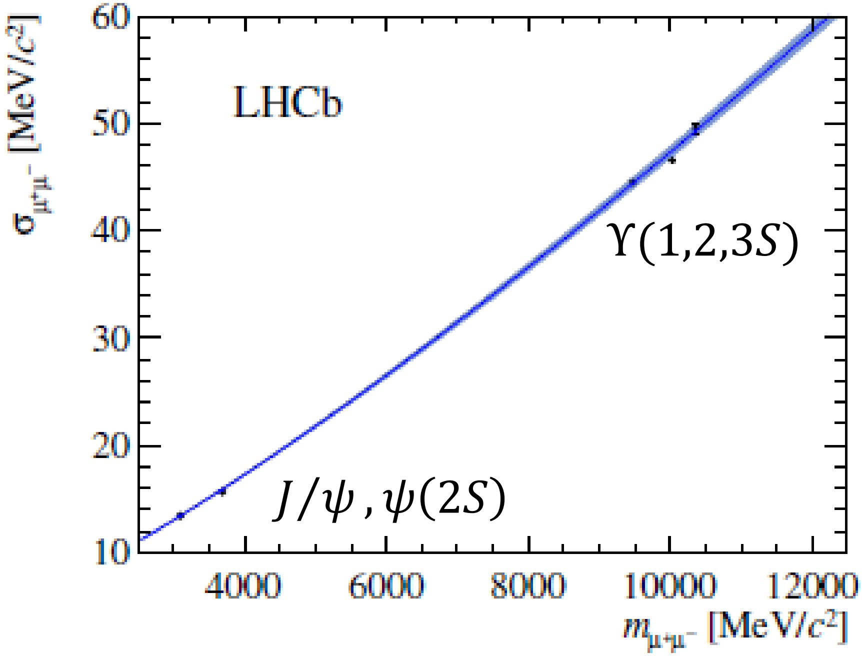
Expected background yield in [4.9-6] GeV/c^2

- $B^0 \to \pi, K^+ \mu^- \bar{\nu}_\mu$, $B^{0,+} \to \pi^{0,+} \mu^+ \mu^-$: included in fit
  - Mass PDF in each BDT bin determined from MC
  - normalization fixed to expected yield
- $\Lambda^0_b \to p \mu^- \bar{\nu}_\mu$: treated as a systematic

Signal Expectations - PDF

BDT PDF:

- $B^0_{(s)} \to h^+ h'^-$ same geometry as signal
- PDF by extracting yields with a $m_{\mu\mu}$ fit in each bin



Mass PDF, Crystal Ball

- Mean $B^0_{(s)}$ mass, from $B^0_{(s)} \to h^+ h'^-$
- Mass resolution, from J/ψ, $\psi(2S)$, $\Upsilon(1,2,3S)$ and $B^0_{(s)} \to h^+ h'^-$
- Radiative tail from MC

$$\sigma_{B^0} = 22.8 \pm 0.4 \text{ MeV}$$

$$m_{B_s} - m_{B_d} \sim 3.5\sigma_{B^0}$$

Signal Expectations - Yields

- Number of signal events corresponding to a BF:

$$N_{B^0_{(s)}\to\mu^+\mu^-} \propto BF(B^0_{(s)} \to \mu^+\mu^-) \times N_{B_s}$$

- N_{B_s} obtained by normalising to channels of known BF:

$B^+ \to J/\psi K^+$ and $B^0 \to K\pi$

- Correcting for efficiencies and hadronisation probability:

$$\boldsymbol{N_{B^0_{(s)}\to\mu^+\mu^-} = BF\left(B^0_{(s)} \to \mu^+\mu^-\right) \times \frac{N_{norm}}{BF_{norm}} \times \frac{f_{B^0_{(s)}}}{f_{norm}} \times \frac{\epsilon_{sig}}{\epsilon_{norm}}}$$

Ratio of probability for a *b* quark to hadronise into a $\boldsymbol{B^0_{(s)}}$ or into the normalisation initial state

Normalisation Ingredients

$$N_{B^0_{(s)}\to\mu^+\mu^-} = BF\left(B^0_{(s)} \to \mu^+\mu^-\right) \times \frac{N_{norm}}{BF_{norm}} \times \frac{f_{B^0_{(s)}}}{f_{norm}} \times \frac{\epsilon_{sig}}{\epsilon_{norm}}$$

- N_{norm}: Obtained by fitting control channels invariant mass
- $f_{B^0_{(s)}} / f_{norm}$: Measured in LHCb*
- $\epsilon_{sig} / \epsilon_{norm}$: Measured with simulations

Averaging the results from the two control channels:

SM expectations in signal mass windows:
$40 \pm 4 \;\; B_s^0 \to \mu^+\mu^-$ and $4.5 \pm 0.4 \;\; B^0 \to \mu^+\mu^-$

*LHCb-CONF-2013-011

Time Acceptance Correction - 1

- Recall, decay rate depends on $A_{\Delta\Gamma}$:

$$\langle\Gamma(B^0_{(s)}(t) \to \ell^+\ell^-)\rangle = R_H e^{-\Gamma_\mathrm{H}^{(s)}\mathrm{t}} + R_L e^{-\Gamma_\mathrm{L}^{(s)}\mathrm{t}} \quad \text{and} \quad A_{\Delta\Gamma} = \frac{R_H - R_L}{R_H + R_L}$$

- So $B^0_s \to \mu^+\mu^-$ signal efficiency ϵ_{sig} depend on $A_{\Delta\Gamma}$:

$$\epsilon_{sig}(A_{\Delta\Gamma}) = \frac{\int \Gamma(A_{\Delta\Gamma}, t) \times \epsilon(t)\, dt}{\int \Gamma(A_{\Delta\Gamma}, t) dt}$$

- Signal efficiency used in normalisation…

$$BF_{sig} = BF_{norm} \times \frac{N_{sig}}{N_{norm}} \times \frac{f_{sig}}{f_{norm}} \times \frac{\epsilon_{sig}(A_{\Delta\Gamma})}{\epsilon_{norm}}$$

…is obtained form MC and must be corrected to match latest y_s and $\tau_{B^0_s}$ values:

$$\epsilon^{SM}_{B^0_s \to \mu\mu} / \epsilon^{MC} = 1.0457 \qquad \epsilon^{SM}_{B^0 \to \mu\mu} / \epsilon^{MC} = 1.015$$

Time Acceptance Correction - 2

- BDT uses decay-time as input variable
- BDT PDF obtained with mainly $B^0 \to h^+h'^-$
- $B^0 \to h^+h'^-$ and $B_s^0 \to \mu^+\mu^-$ have different lifetime distributions
- Need corrections:

$$PDF^i_{corr} \Big/ PDF^i_{raw} = \delta^i_{PDF}$$

Bin	PDF Correction $\delta^i_{PDF} - 1$ (%)
1	-3.1061 ± 0.0196
2	-1.3778 ± 0.0290
3	-0.3887 ± 0.0392
4	$+0.2701 \pm 0.0423$
5	$+0.7193 \pm 0.0447$
6	$+1.3650 \pm 0.0457$
7	$+2.5423 \pm 0.0463$
8	$+4.7365 \pm 0.0433$

- Nota: all these corrections depend on the value assumed for $A_{\Delta\Gamma}$

Analysis Overview

- Answering the question:

  Which BF is (in-)compatible with the observed data?
- Observed data = $N(B^0_{(s)} \to \mu^+\mu^- + \mathbf{Bkg})$

1. Increasing sensitivity:
   - Events Selection and Classification in categories
2. Derive Signal and Bkg expectations from control channels
3. Compare expectations with observation and infer about BF
   - CLs method and Fitting procedure if signal evidence

- **Data Set:**

2011	$1.0\ \text{fb}^{-1}$	7TeV	
2012	$+1.1\ \text{fb}^{-1}$	8TeV	PRL 110, 021801, 2013
2012	$+0.9\ \text{fb}^{-1}$	8TeV	PRL 111, 101805, 2013

Sensitivity Expectations (CLs method)

Invariant mass [MeV/c^2]		BDT 0.0 – 0.25	0.25 – 0.4	0.4 – 0.5	0.5 – 0.6	0.6 – 0.7	0.7 – 0.8	0.8 – 0.9	0.9 – 1.0
5311 – 5431	Exp. comb. bkg	6138^{+114}_{-112}	$121.6^{+4.8}_{-4.7}$	$28.2^{+2.2}_{-2.1}$	$11.9^{+1.5}_{-1.4}$	$4.77^{+1.11}_{-0.95}$	$2.17^{+0.79}_{-0.65}$	$0.79^{+0.48}_{-0.34}$	$0.29^{+0.32}_{-0.18}$
	Exp. peak. bkg	$0.330^{+0.105}_{-0.089}$	$0.210^{+0.068}_{-0.058}$	$0.140^{+0.045}_{-0.038}$	$0.148^{+0.048}_{-0.040}$	$0.147^{+0.047}_{-0.040}$	$0.140^{+0.045}_{-0.038}$	$0.130^{+0.042}_{-0.035}$	$0.111^{+0.035}_{-0.030}$
	Exp. signal	$8.78^{+1.09}_{-0.99}$	$5.40^{+0.75}_{-0.67}$	$3.52^{+0.46}_{-0.41}$	$3.75^{+0.47}_{-0.43}$	$3.76^{+0.47}_{-0.43}$	$3.61^{+0.46}_{-0.42}$	$3.68^{+0.46}_{-0.42}$	$3.79^{+0.46}_{-0.42}$

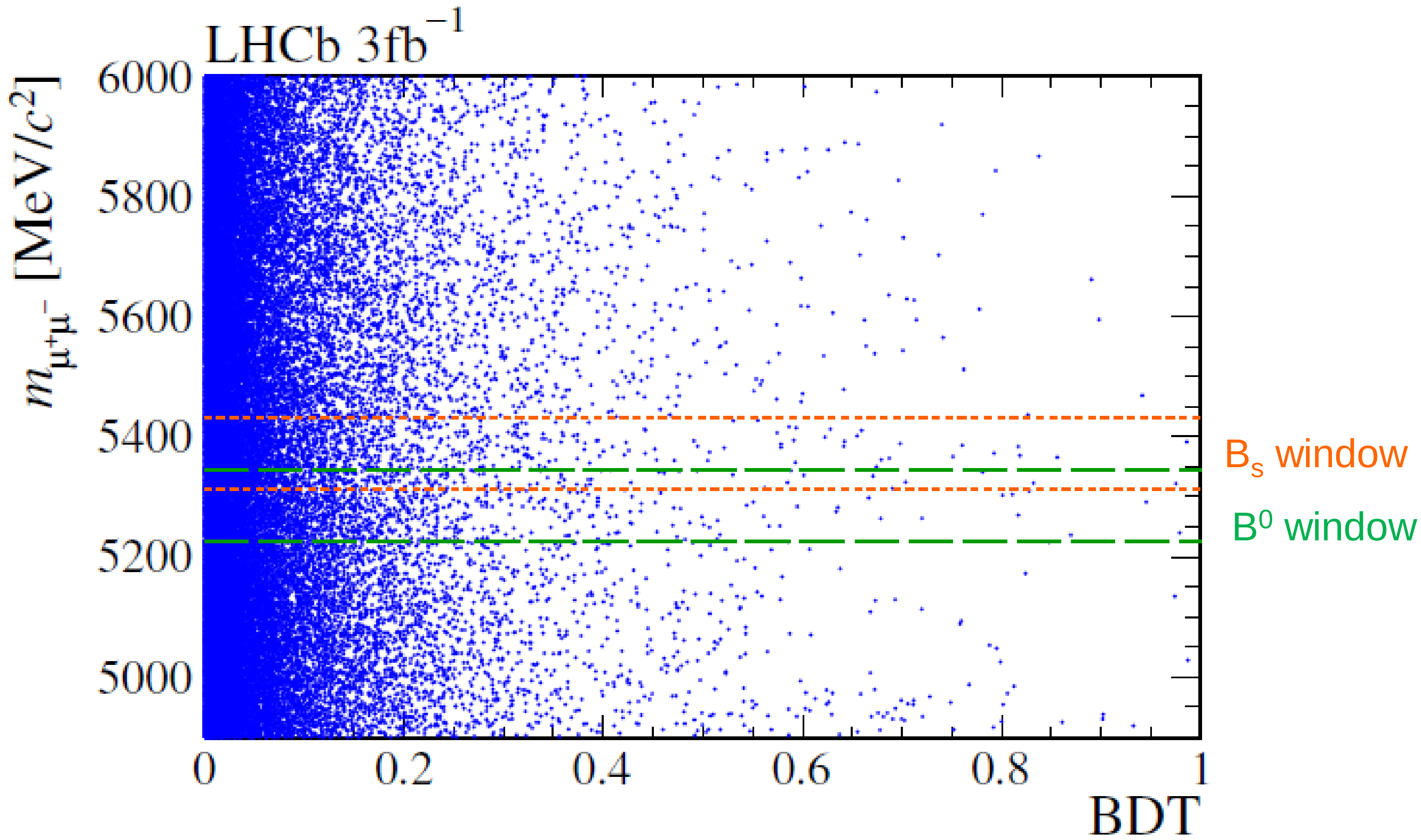
Expected significance for a $B_s^0 \rightarrow \mu^+\mu^-$ SM signal : 5.0 σ

Invariant mass [MeV/c^2]		BDT 0.0 – 0.25	0.25 – 0.4	0.4 – 0.5	0.5 – 0.6	0.6 – 0.7	0.7 – 0.8	0.8 – 0.9	0.9 – 1.0
5224 – 5344	Exp. comb. bkg	6520^{+119}_{-117}	$127.0^{+5.2}_{-5.0}$	$29.4^{+2.4}_{-2.3}$	$12.8^{+1.7}_{-1.5}$	$4.9^{+1.2}_{-1.1}$	$2.14^{+0.88}_{-0.70}$	$0.82^{+0.53}_{-0.37}$	$0.29^{+0.35}_{-0.19}$
	Exp. peak. bkg	$1.97^{+0.64}_{-0.47}$	$1.25^{+0.41}_{-0.31}$	$0.83^{+0.27}_{-0.20}$	$0.88^{+0.29}_{-0.21}$	$0.88^{+0.28}_{-0.21}$	$0.83^{+0.27}_{-0.20}$	$0.77^{+0.25}_{-0.18}$	$0.66^{+0.21}_{-0.16}$
	Exp. Cross-feed	$1.38^{+0.18}_{-0.16}$	$0.85^{+0.12}_{-0.11}$	$0.554^{+0.075}_{-0.067}$	$0.590^{+0.078}_{-0.070}$	$0.591^{+0.076}_{-0.070}$	$0.567^{+0.077}_{-0.069}$	$0.579^{+0.076}_{-0.069}$	$0.595^{+0.077}_{-0.069}$
	Exp. signal	$0.99^{+0.12}_{-0.11}$	$0.610^{+0.081}_{-0.075}$	$0.398^{+0.049}_{-0.046}$	$0.424^{+0.050}_{-0.047}$	$0.425^{+0.050}_{-0.047}$	$0.408^{+0.050}_{-0.047}$	$0.416^{+0.049}_{-0.046}$	$0.428^{+0.050}_{-0.046}$

No significant $B^0 \rightarrow \mu^+\mu^-$ SM signal expected:

$B(B^0 \rightarrow \mu^+\mu^-) < 5.4 \times 10^{-10}$ at 95% CL

Open the box


LHCb 3fb^{-1}
6000
5800
5600
5400
5200
5000
$m_{\mu^+\mu^-}$ [MeV/c^2]
0
0.2
0.4
0.6
0.8
1
BDT
B_s window
B^0 window

Fit Projection

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

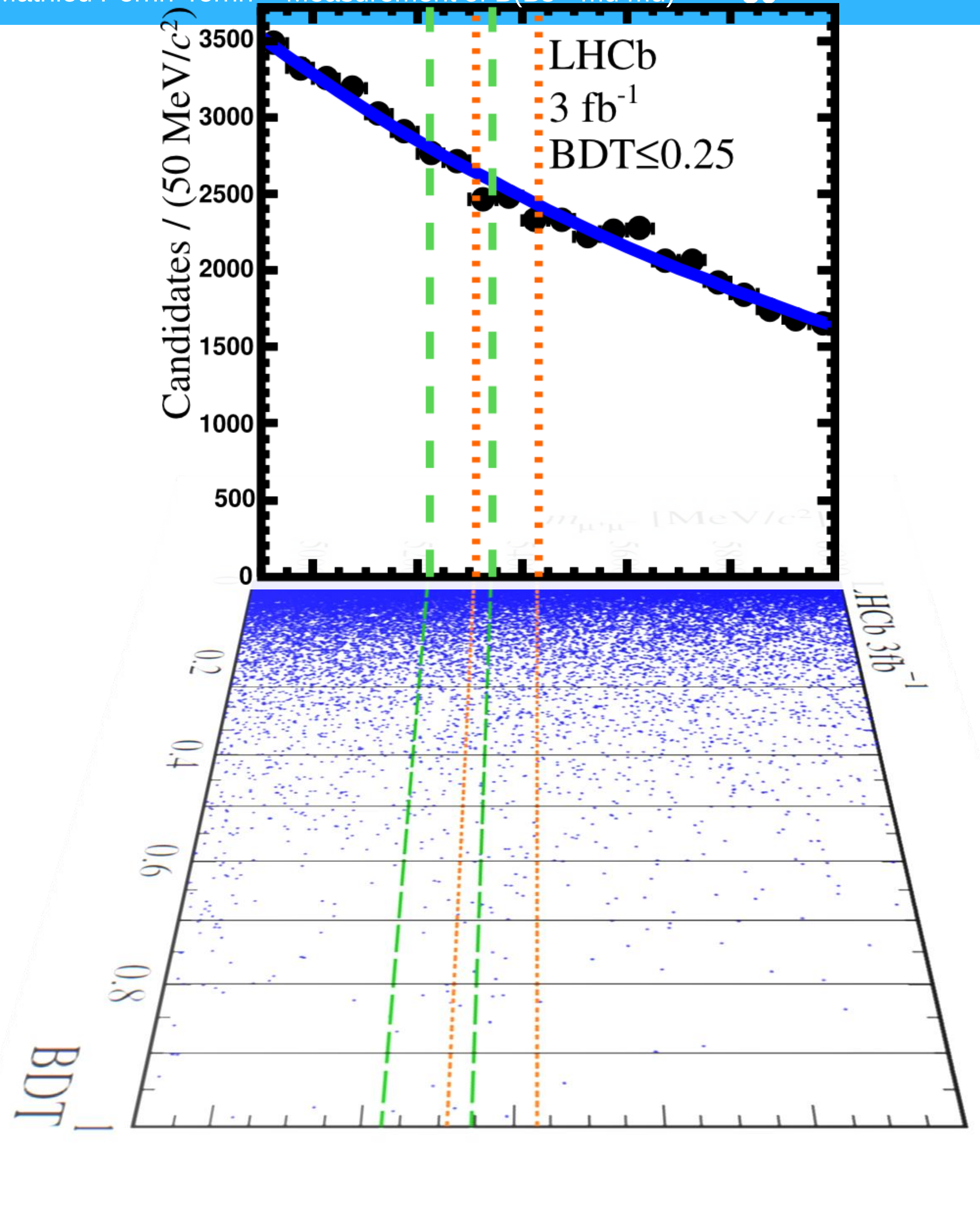
$B^0 \rightarrow \mu^+\mu^-$

$B^0 \rightarrow \pi, K^+\mu^-\overline{\nu}_\mu$

$B^{0,+} \rightarrow \pi^{0,+}\mu^+\mu^-$

$B_{(s)}^0 \rightarrow h^+h'^-$

Total



Fit Projection

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

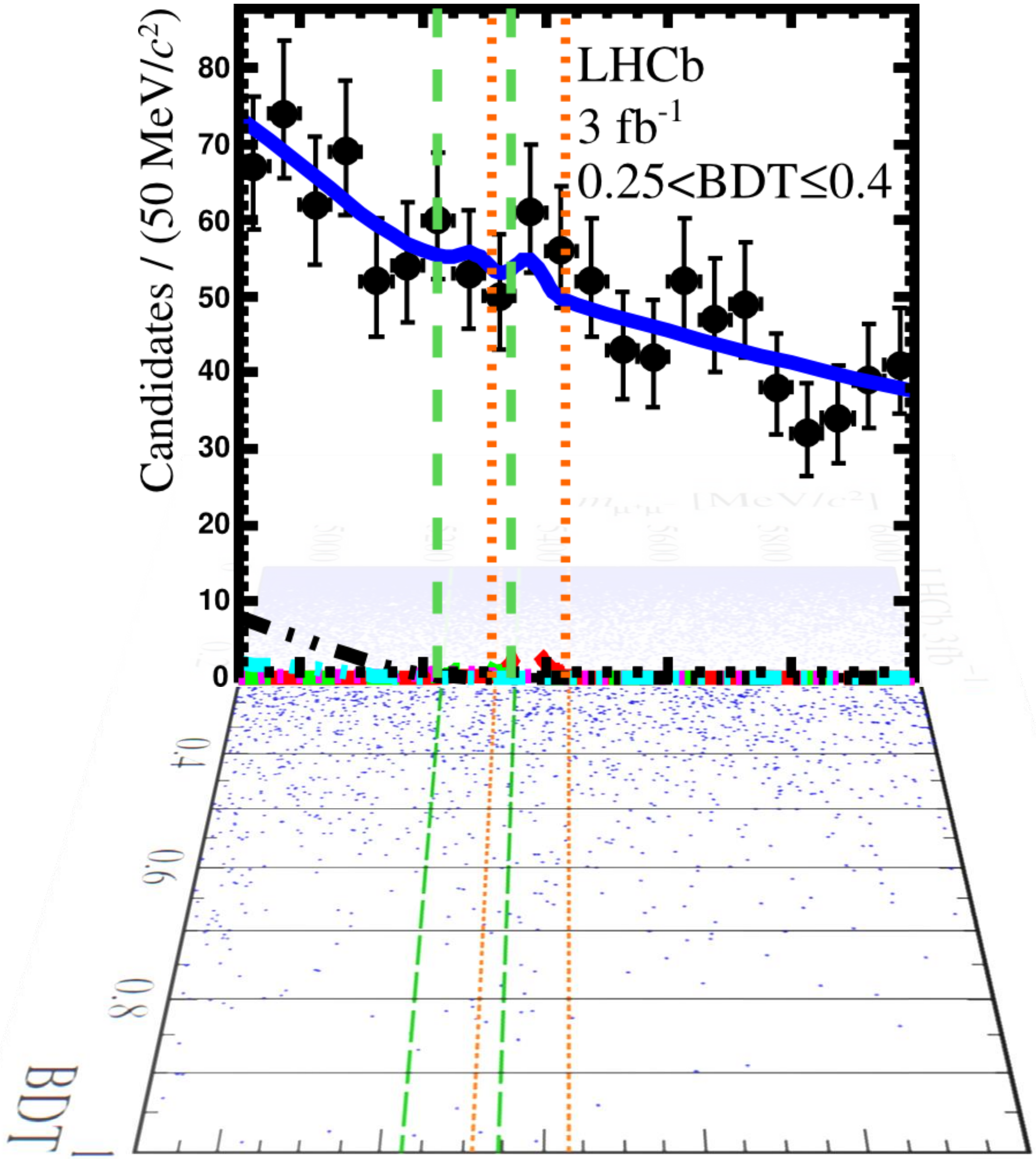
$B^0 \rightarrow \mu^+\mu^-$

$B^0 \rightarrow \pi, K^+\mu^-\overline{\nu}_\mu$

$B^{0,+} \rightarrow \pi^{0,+}\mu^+\mu^-$

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

Total



Fit Projection

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

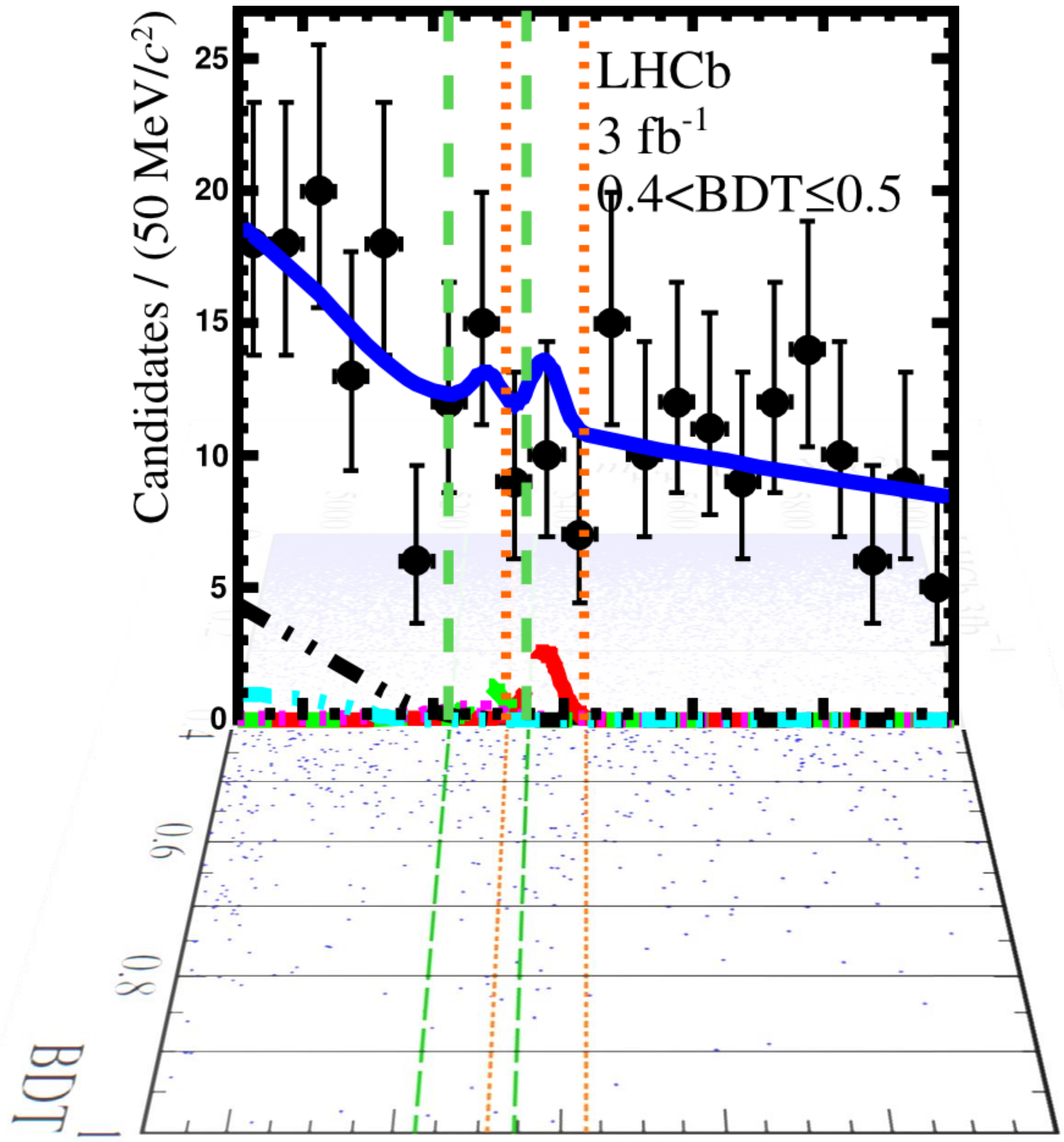
$B^0 \to \mu^+\mu^-$

$B^0 \to \pi, K^+\mu^-\overline{\nu}_\mu$

$B^{0,+} \to \pi^{0,+}\mu^+\mu^-$

$B_{(s)}^0 \to h^+h'^-$

Total



Fit Projection

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

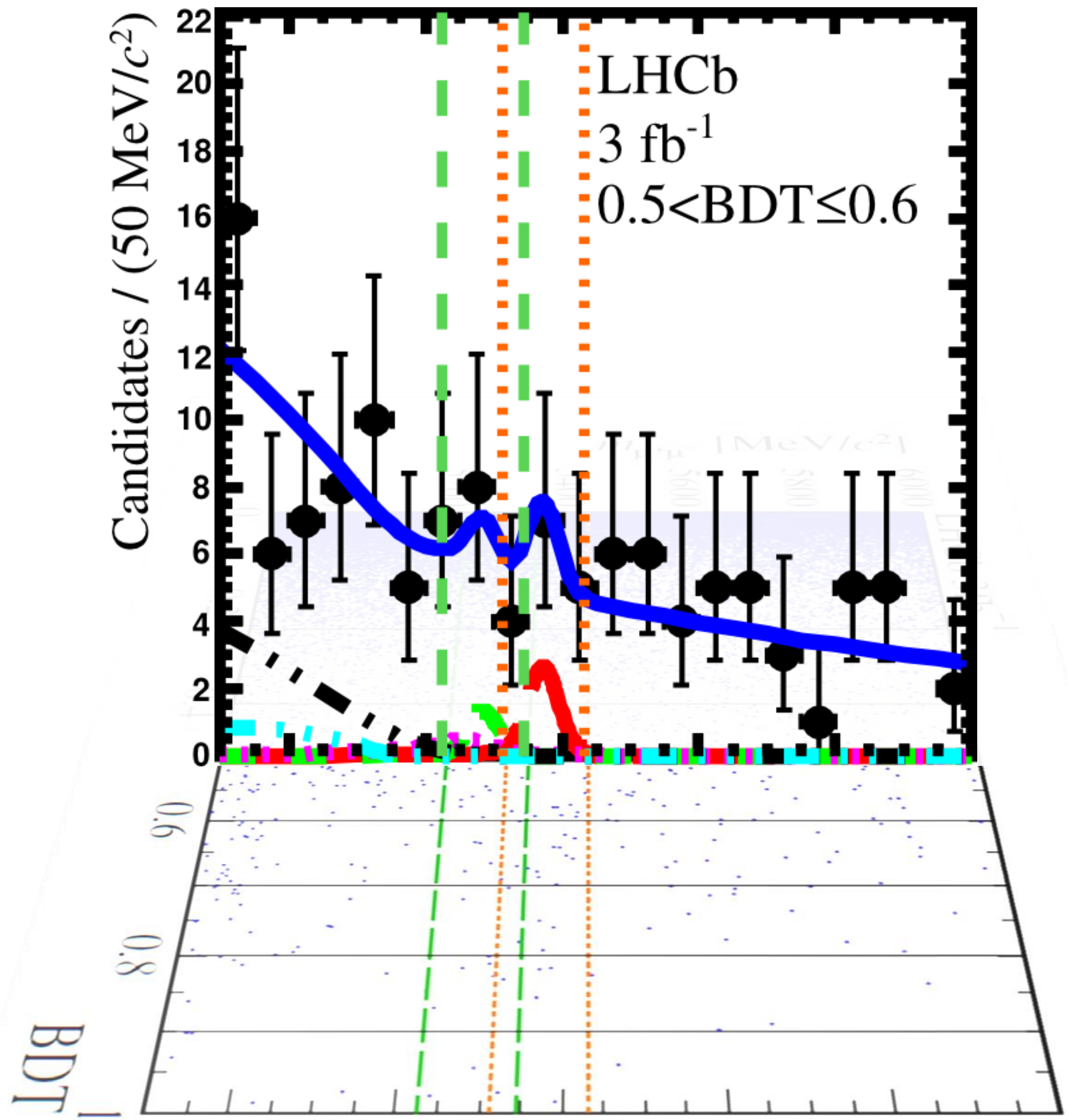
$B^0 \rightarrow \mu^+\mu^-$

$B^0 \rightarrow \pi, K^+\mu^-\overline{\nu}_\mu$

$B^{0,+} \rightarrow \pi^{0,+}\mu^+\mu^-$

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

Total



Fit Projection

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

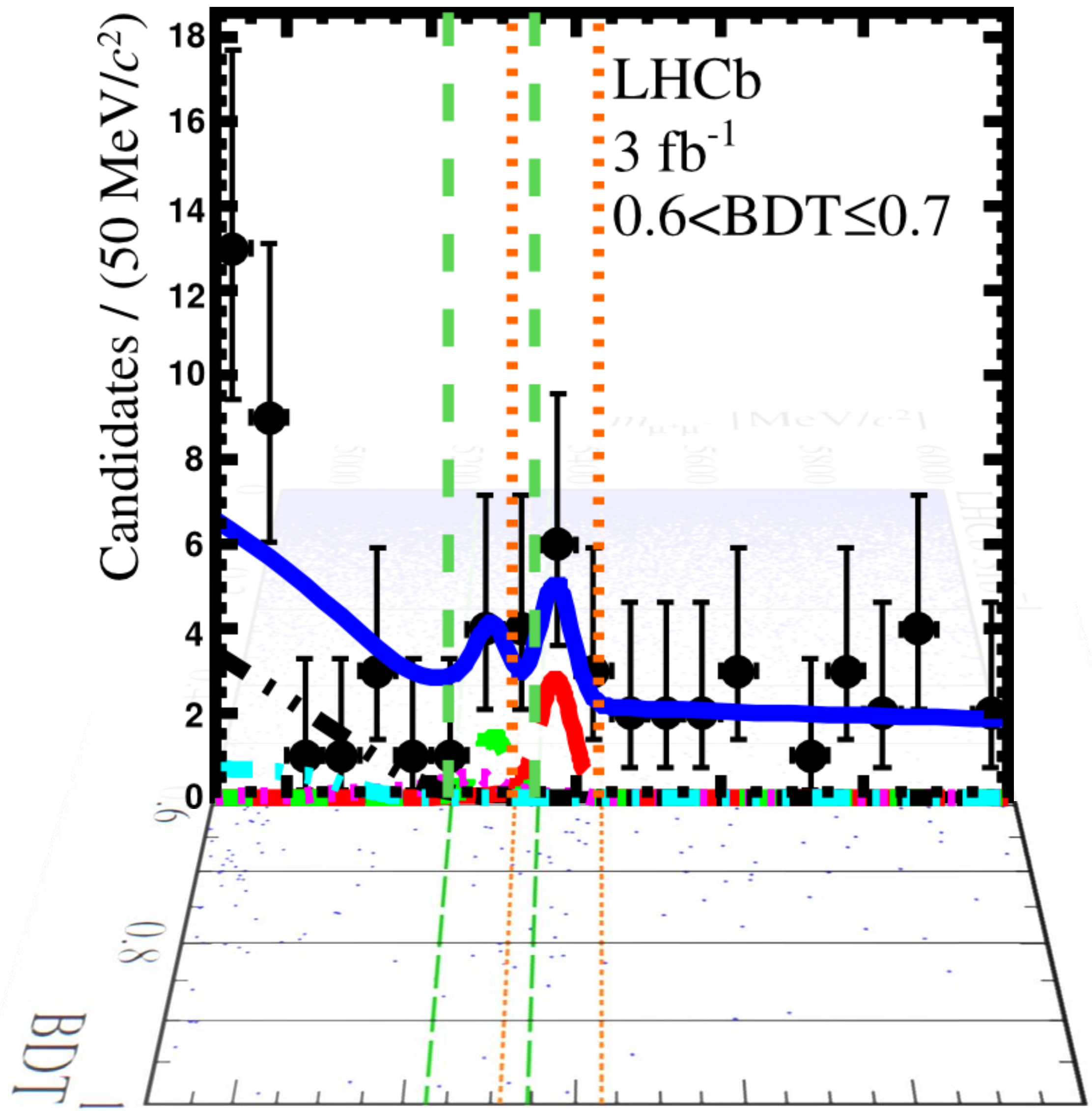
$B^0 \to \mu^+\mu^-$

$B^0 \to \pi, K^+\mu^-\overline{\nu}_\mu$

$B^{0,+} \to \pi^{0,+}\mu^+\mu^-$

$B^0_{(s)} \to h^+h'^-$

Total



Fit Projection

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

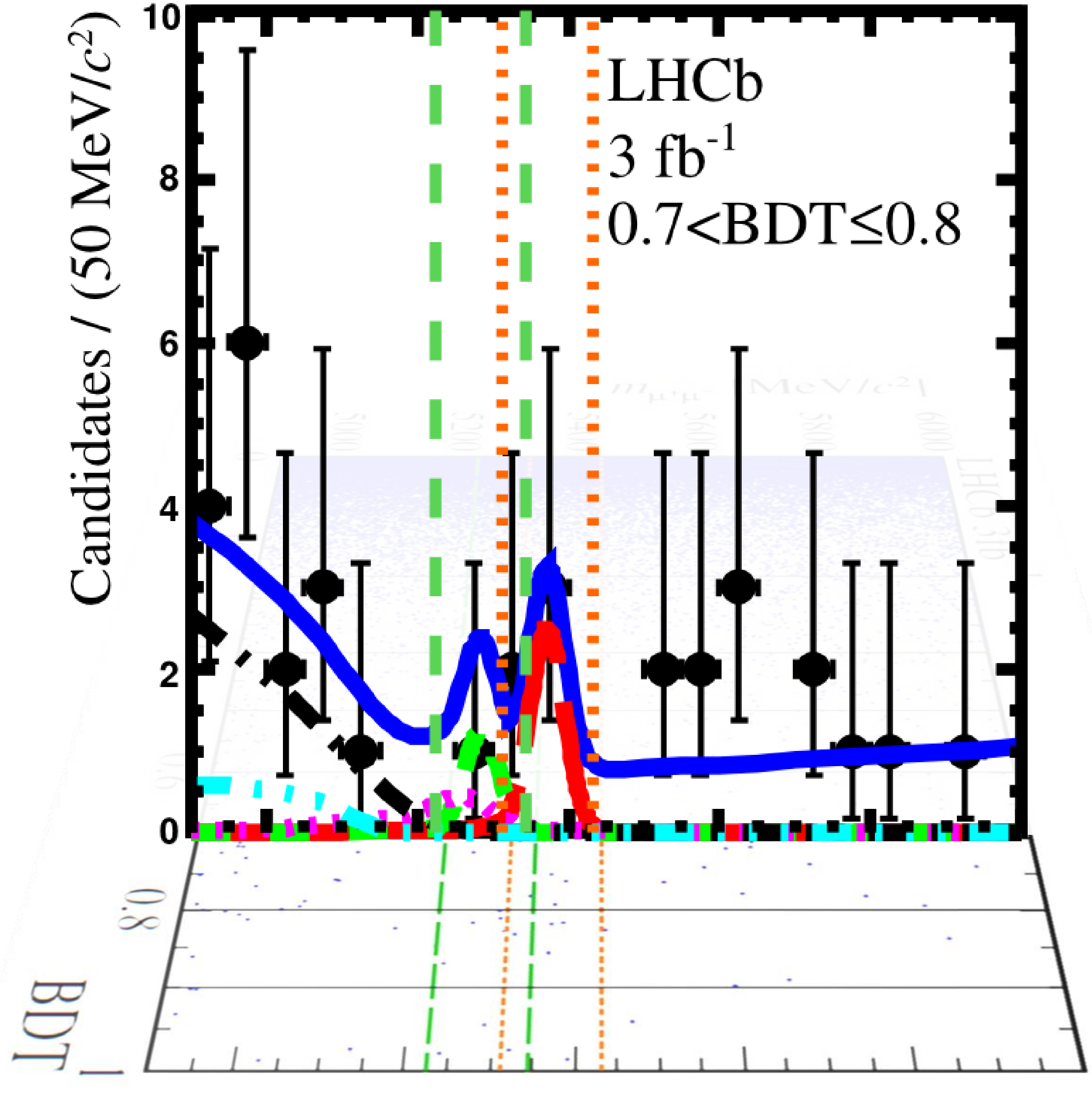
$B^0 \to \mu^+\mu^-$

$B^0 \to \pi, K^+\mu^-\overline{\nu}_\mu$

$B^{0,+} \to \pi^{0,+}\mu^+\mu^-$

$B^0_{(s)} \to h^+h'^-$

Total



Fit Projection

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

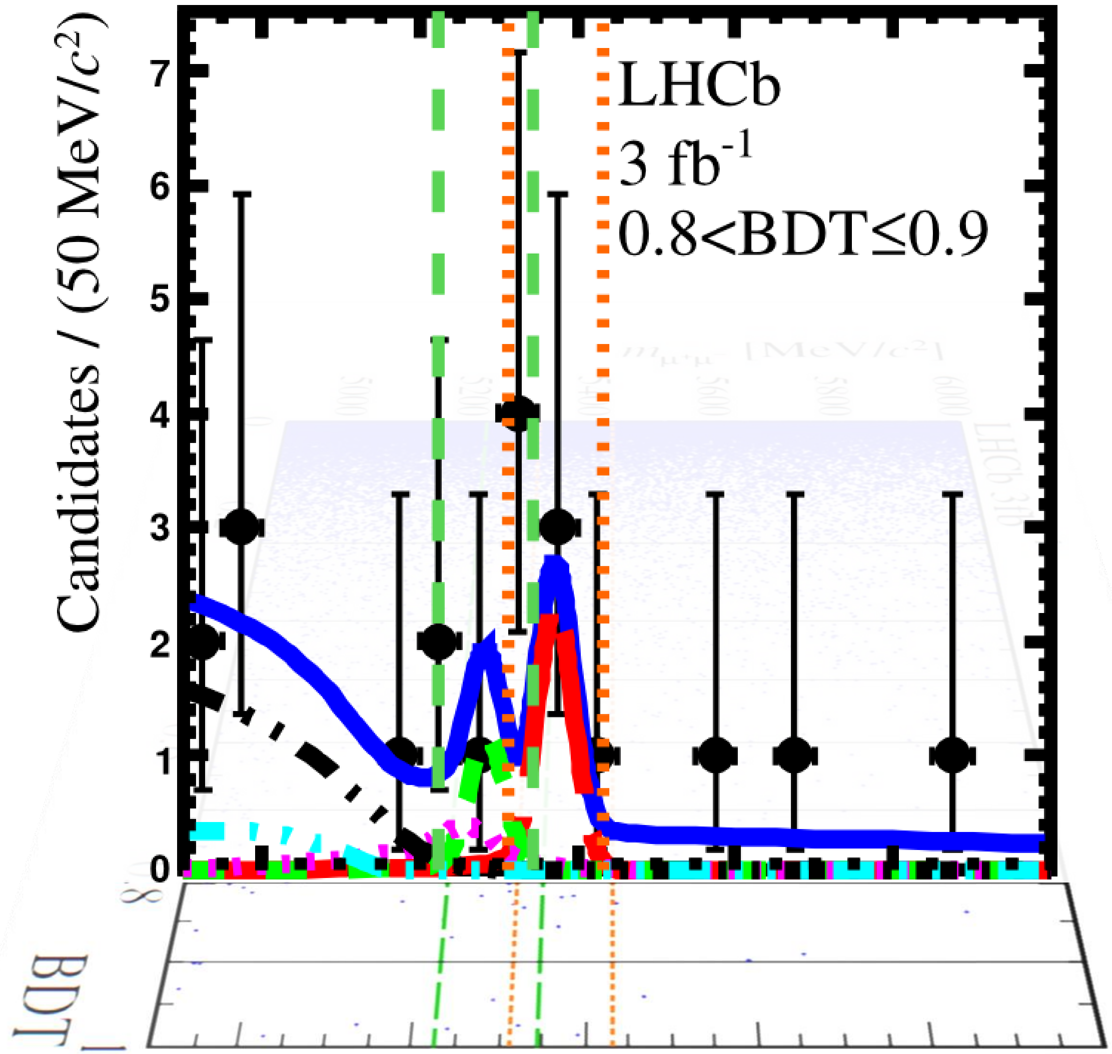
$B^0 \to \mu^+\mu^-$

$B^0 \to \pi, K^+\mu^-\overline{\nu}_\mu$

$B^{0,+} \to \pi^{0,+}\mu^+\mu^-$

$B^0_{(s)} \to h^+h'^-$

Total



Fit Projection

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

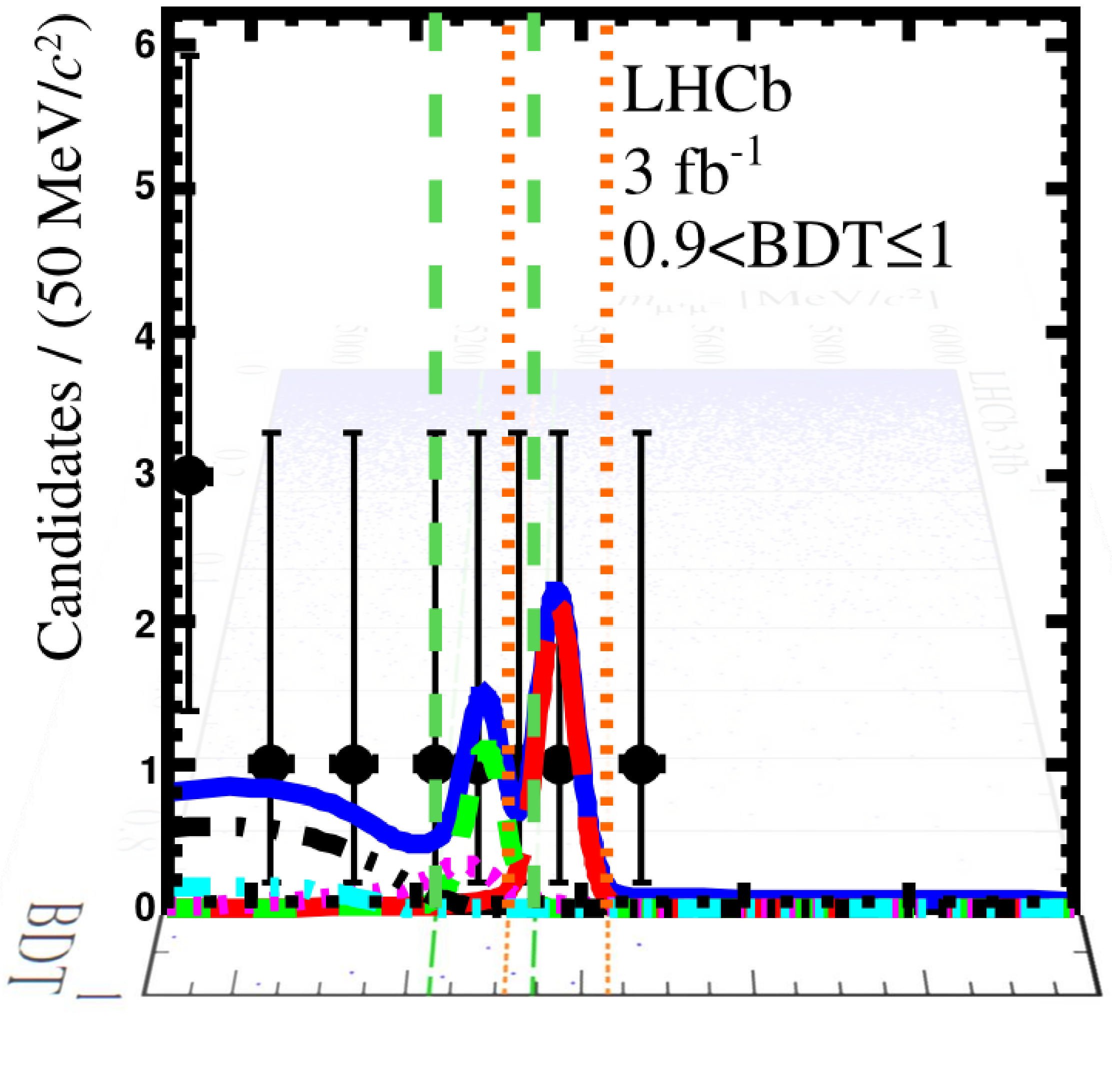
$B^0 \rightarrow \mu^+\mu^-$

$B^0 \rightarrow \pi, K^+\mu^-\overline{\nu}_\mu$

$B^{0,+} \rightarrow \pi^{0,+}\mu^+\mu^-$

$B_{(s)}^0 \rightarrow h^+h'^-$

Total



Fit Result

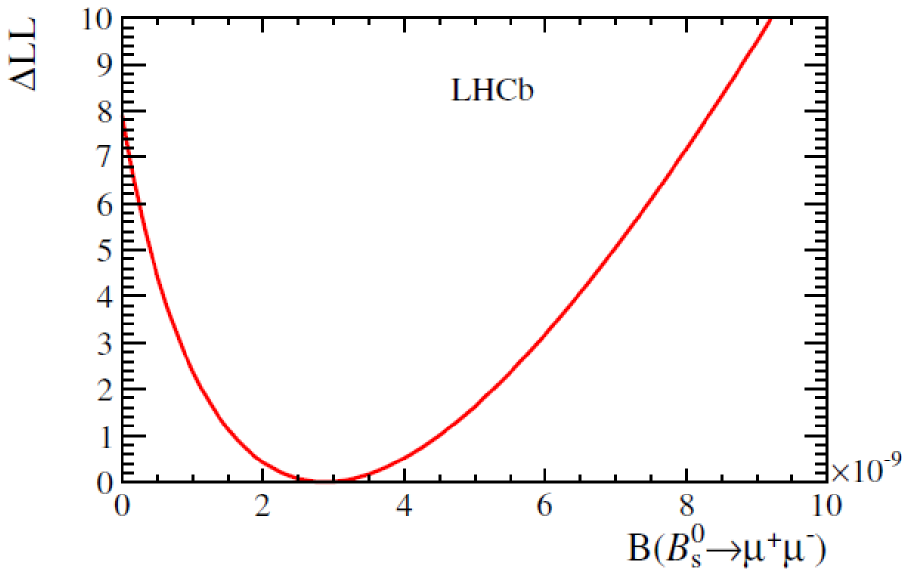
arXiv:1307.5024, Phys.Rev. Lett.111(2013) 101805

$$B(B_s^0 \to \mu^+\mu^-) = (2.9^{+1.1}_{-1.0}(stat)^{+0.3}_{-0.1}(syst)) \times 10^{-9}$$

Significance: 4.0 σ
expected 5.0 σ (median)

$$B(B^0 \to \mu^+\mu^-) = (3.7^{+2.4}_{-2.1}(stat)^{+0.6}_{-0.4}(syst)) \times 10^{-10}$$

Significance: 2.0 σ

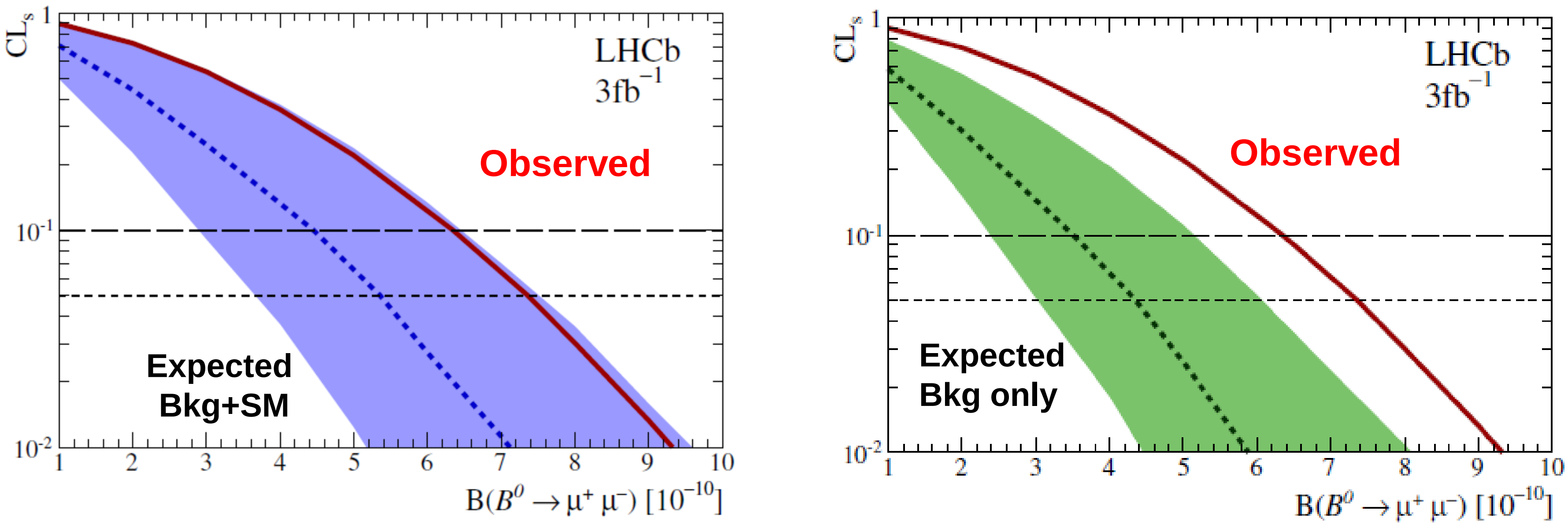


Correlation between BR($B^0 \to \mu^+\mu^-$) and BR($B_s \to \mu^+\mu^-$) : 3.3%

Profile Likelihood:
All parameters except $B(B_s^0 \to \mu^+\mu^-)$ are floated within their errors.

$B^0 \to \mu^+\mu^-$ upper limit

- Obtained with the CLs method



	Limit at 95%CL
Expected bkg only	4.4×10^{-10}
Expected bkg + SM	5.4×10^{-10}
observed	7.4×10^{-10}

Outlines

- Motivations to search for $B^0_{(s)} \to \mu^+\mu^-$
- Searching for $\mathrm{B}^0_{(s)} \to \mu^+\mu^-$ at LHCb
- Combination with CMS
- Conclusions

CMS Analysis in Brief

- Dataset: 5 (7 TeV) + 20 (8 TeV) fb^{-1}
- Trigger requirement :
  - (sub-)leading muon pT> 3 (4) GeV, dimuon pT>4.9 GeV for $|\eta_{\mu\mu}|<1.8$
  - (sub-)leading muon pT> 4 (4) GeV, dimuon pT>7 GeV for $|\eta_{\mu\mu}|>1.8$
  - $4.8<m_{\mu\mu}<6$ GeV
  - Vertex fit $p(\chi^2)>0.5\%$
- Discriminant variables: $m_{\mu\mu}$ and BDT
- Two categories and 2011 and 2012 data kept separated:
  - Both muons in barrel : $\sigma_{\mu\mu}\sim 40$ MeV
  - At least one muon in endcap: more events but $\sigma_{\mu\mu}\sim 60$ MeV

BDT

- Training on MC signal and data sidebands:
  - To avoid biases, use 3 separate samples:
  train on 1st, test on 2nd and apply on 3rd $\Rightarrow$ 3 BDT per categories

- 12 variables used, independent of pile-up conditions

- Signal BDT distribution taken from MC, systematics evaluated using control sample

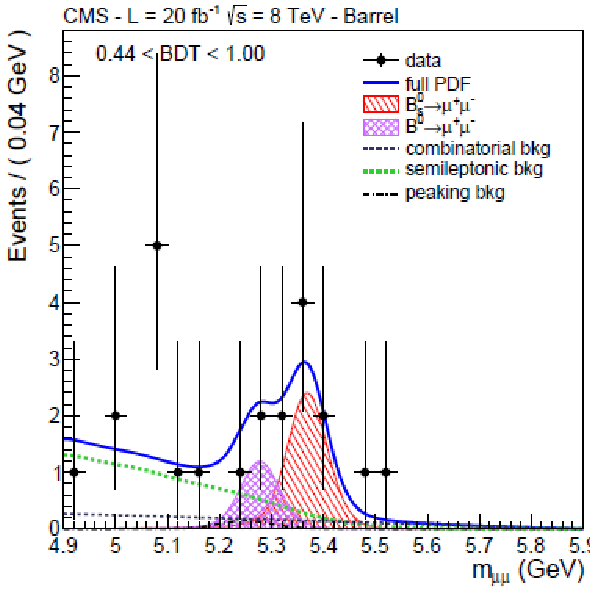
- Then, 2 possible methods:
  - Simple cut on BDT output, optimised for each sample:
  limit computation
  - Analysis in 3 bins of BDT vs mass (higher expected sensitivity):
  simultaneous maximum likelihood fit

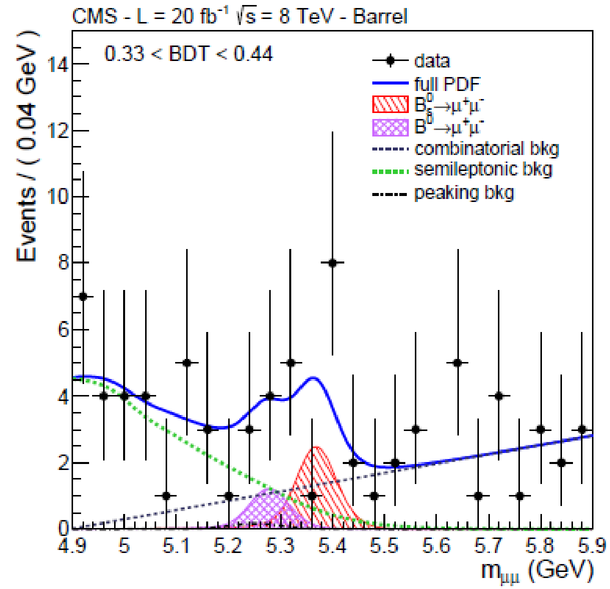
8 TeV data

Most Sensitive Bin

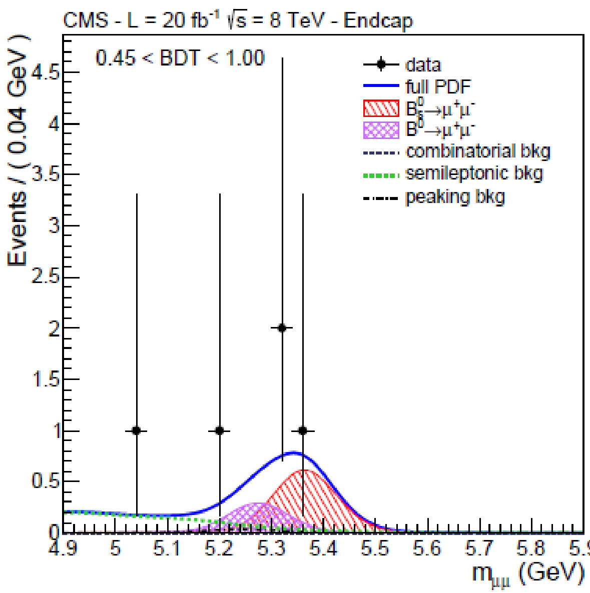
2nd Most Sensitive Bin

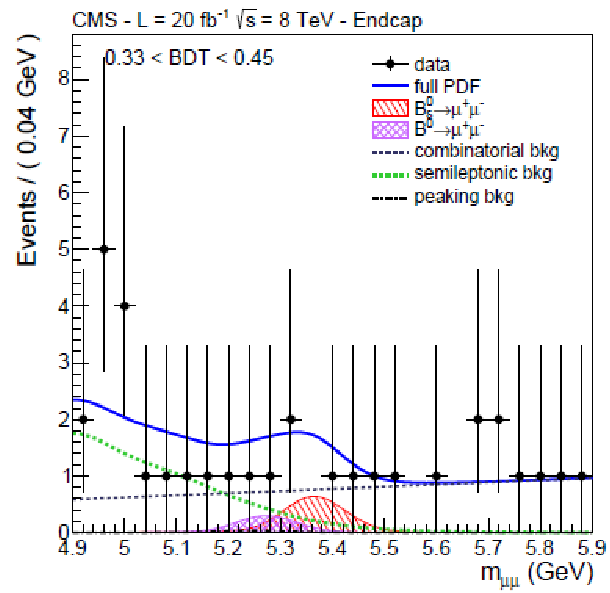
Barrel





End Cap





Results

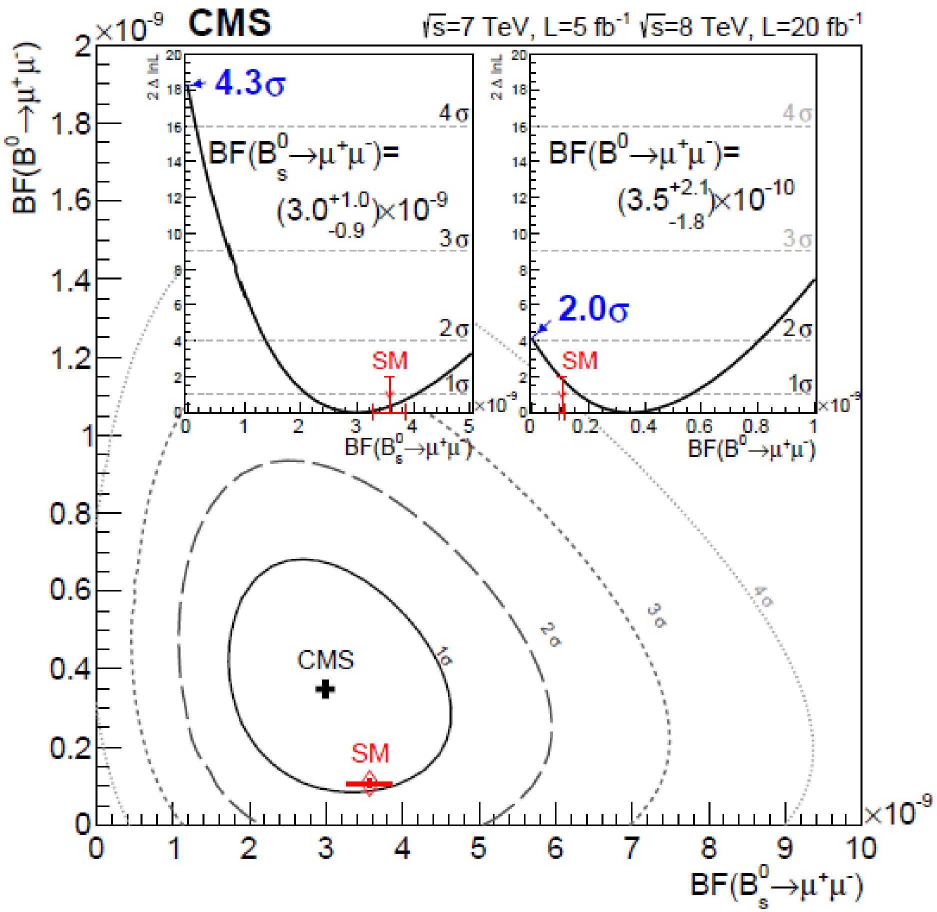
arXiv:1307.5025, Phys.Rev. Lett.111(2013) 101804

$$B(B_s^0 \to \mu^+\mu^-) = (3.0^{+0.9}_{-0.8}(stat)^{+0.6}_{-0.4}(syst)) \times 10^{-9}$$

Significance: 4.3 σ (exp. 4.8)

$$B(B^0 \to \mu^+\mu^-) = (3.5^{+2.1}_{-1.8}(stat + syst)) \times 10^{-10}$$

Significance: 2.0 σ



CLs Method	Limit at 95%CL
Expected bkg + SM	6.3×10^{-10}
observed	11×10^{-10}

Combination CMS LHCb

- Simple average technique (arXiv:physics/0406120)
- Treat correlated systematic uncertainty: ${f_s}/{f_d}$
- Update CMS results with latest ${f_s}/{f_d}$ *

$$B(B_s^0 \to \mu^+\mu^-) = 2.96^{+0.97}_{-0.85} \pm 0.17_{f_s/f_d} \quad \text{CMS}$$

$$B(B_s^0 \to \mu^+\mu^-) = 2.87^{+1.09}_{-0.95} \pm 0.17_{f_s/f_d} \quad \text{LHCb}$$

- Combined result:

$$\boldsymbol{B\left(B_s^0 \to \mu^+\mu^-\right) = (2.9 \pm 0.7) \times 10^{-9}} \quad \textbf{Signifiance} > \boldsymbol{5\sigma}$$

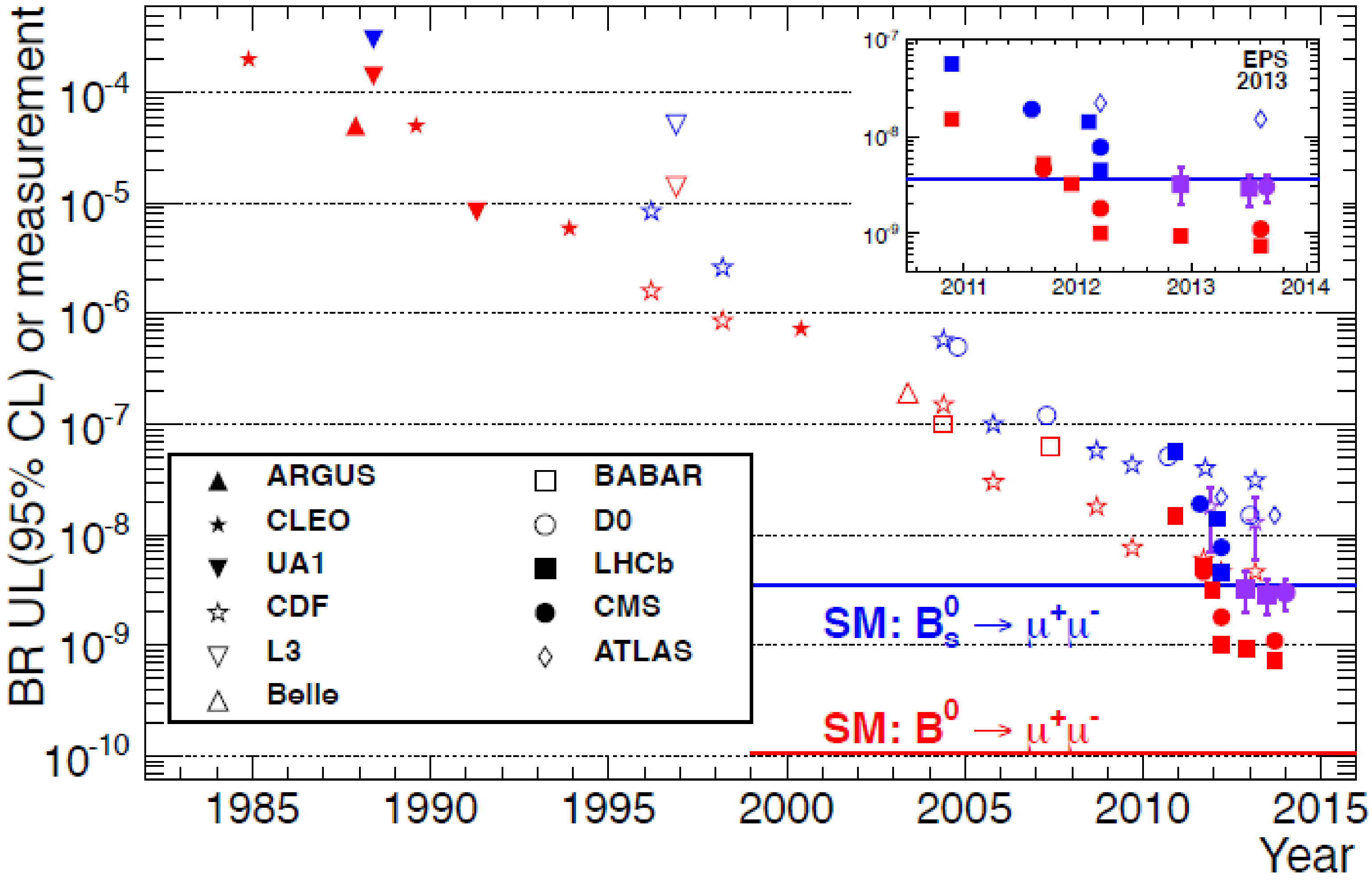
$$\boldsymbol{B\left(B^0 \to \mu^+\mu^-\right) = (3.6^{+1.6}_{-1.4}) \times 10^{-10}} \quad \textbf{Signifiance} < \boldsymbol{3\sigma}$$

*LHCb-CONF-2013-011

Outlines

- Motivations to search for $B^0_{(s)} \rightarrow \mu^+\mu^-$
- Searching for $\mathrm{B}^0_{(s)} \rightarrow \mu^+\mu^-$ at LHCb
- Combination with CMS
- Conclusions

From 1984 to now…



Summary

CMS 25 fb^{-1}

$$B(B_s^0 \to \mu^+\mu^-) = (3.0^{+1.0}_{-0.9}) \times 10^{-9}$$ **4.3 σ**

$$B(B^0 \to \mu^+\mu^-) = 3.5^{+2.1}_{-1.8} \times 10^{-10}$$ **2.0 σ**

$$B(B^0 \to \mu^+\mu^-) < 1.1 \times 10^{-9} \text{ @95\%CL}$$

LHCb 3 fb^{-1}

$$B(B_s^0 \to \mu^+\mu^-) = (2.9^{+1.1}_{-1.0}) \times 10^{-9}$$ **4.0 σ**

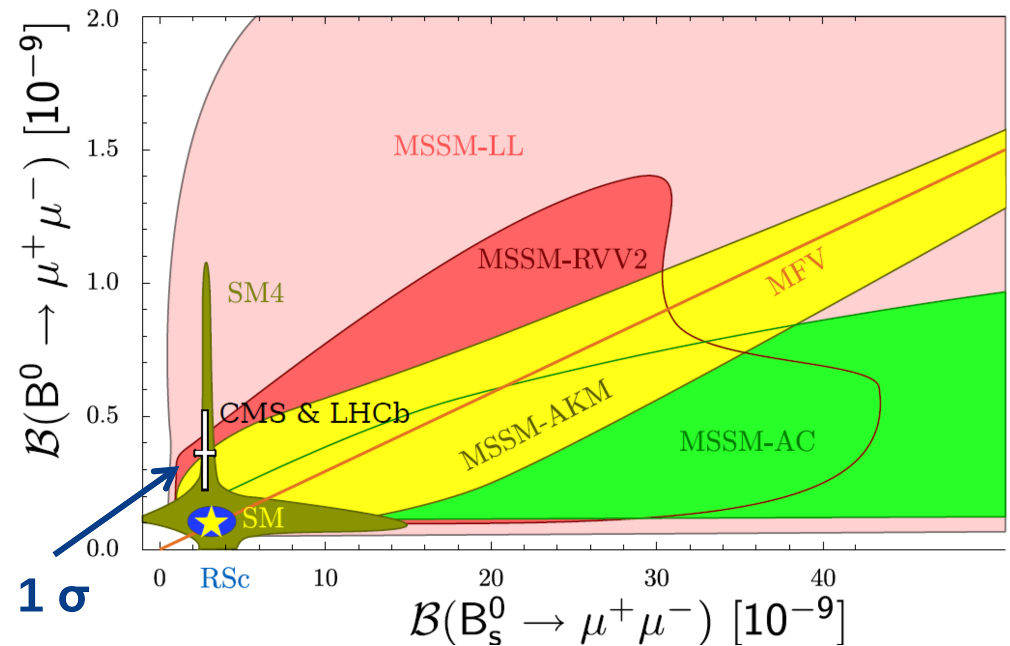
$$B(B^0 \to \mu^+\mu^-) = 3.7^{+2.4}_{-2.1} \times 10^{-10}$$ **2.0 σ**

$$B(B^0 \to \mu^+\mu^-) < 7.4 \times 10^{-10} \text{ @95\%CL}$$

CMS + LHCb :

First observation of BR($B_s \to \mu^+\mu^-$) !!

$$B(B_s^0 \to \mu^+\mu^-) = (2.9 \pm 0.7) \times 10^{-9}$$



Spares

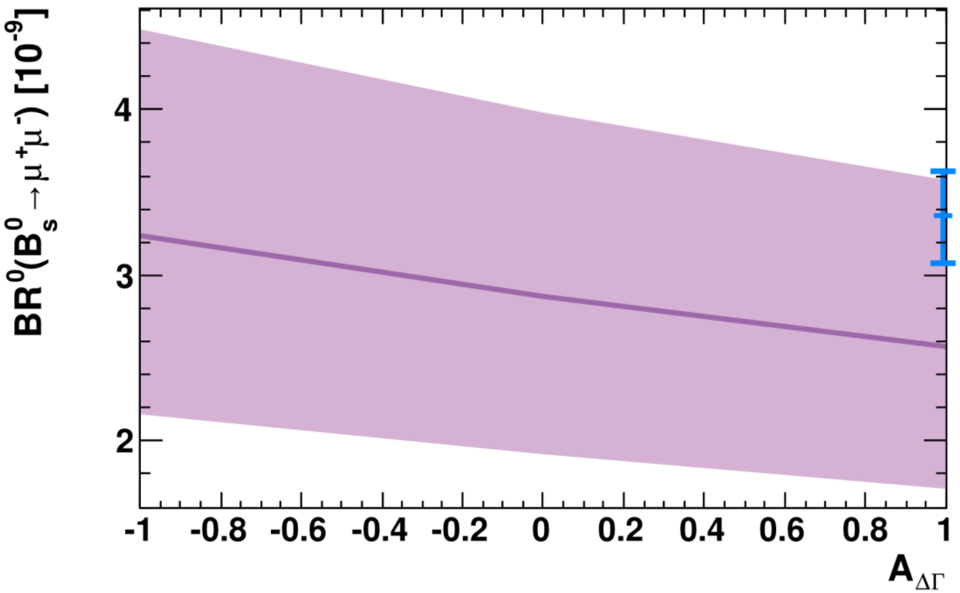
Some projections

- From LHCB-TDR-012:

Obs.	End 2018	LHCb upgrade $50 fb^{-1}$
$B(B_s^0 \to \mu^+\mu^-)$	0.5×10^{-9}	0.15×10^{-9}
$\frac{B(\mathrm{B}_s^0 \to \mu^+\mu^-)}{B(\mathrm{B}^0 \to \mu^+\mu^-)}$	100%	35%

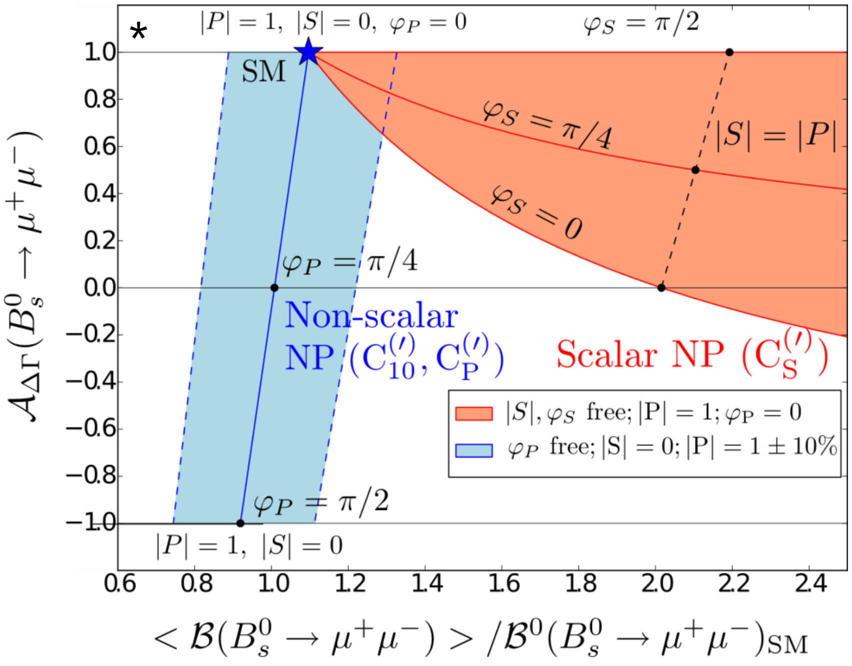
$A_{\Delta\Gamma}$ Dependency

- BR results valid only if $A_{\Delta\Gamma} = 1$
- In NP $A_{\Delta\Gamma}$ can take any value between -1 and +1
- Results depend on $A_{\Delta\Gamma}$:



Exploiting $A_{\Delta\Gamma}$ Dependency

- New Physics can enter both $BF_{t=0}$ and $A_{\Delta\Gamma}$
- The experi. BF dependency on $A_{\Delta\Gamma}$ must be accounted for
- The constraints are more accurate and more stringent

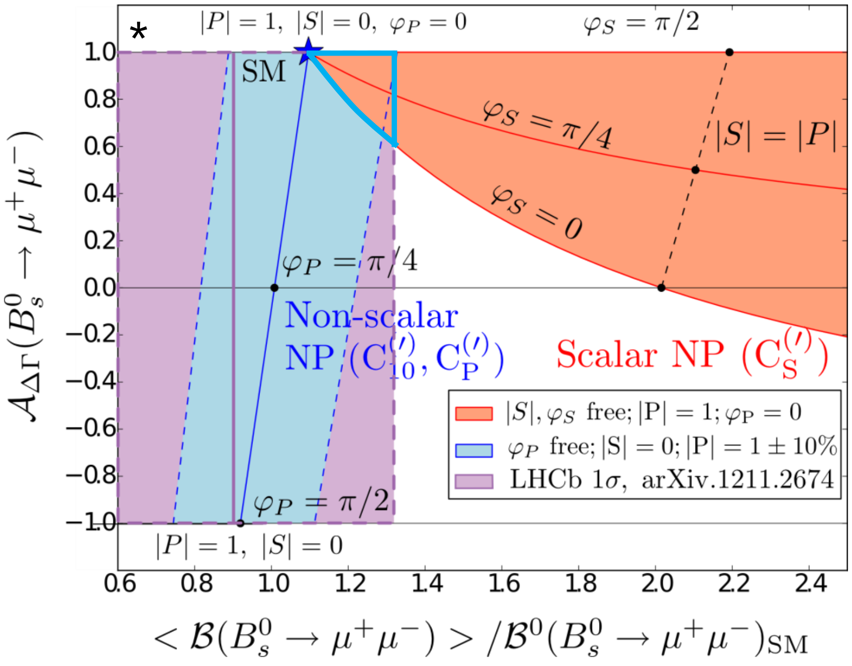


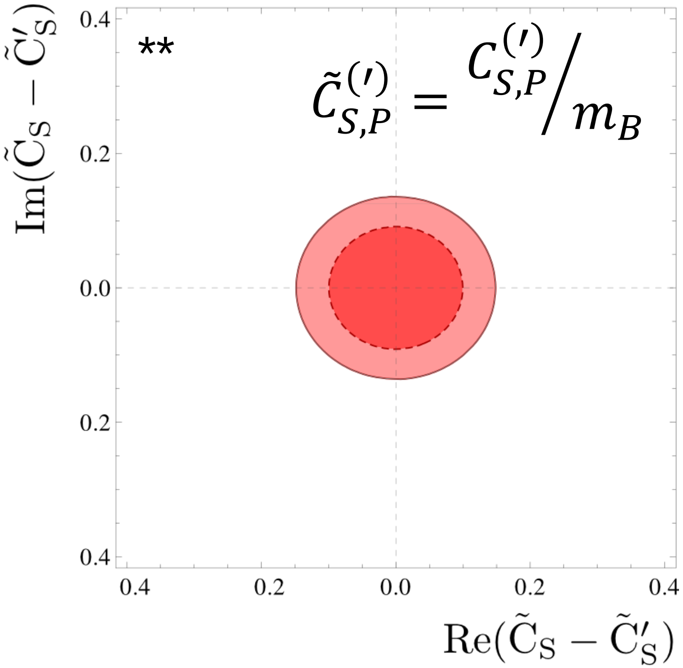
Different NP scenarios feature different non-trivial $A_{\Delta\Gamma} - BF$ correlations

* Modified from De Bruyn et al. PRL 109, 041801, 2012

Exploiting $A_{\Delta\Gamma}$ Dependency

- New Physics can enter both $BF_{t=0}$ and $A_{\Delta\Gamma}$
- The experi. BF dependency on $A_{\Delta\Gamma}$ must be accounted for
- The constraints are more accurate and more stringent



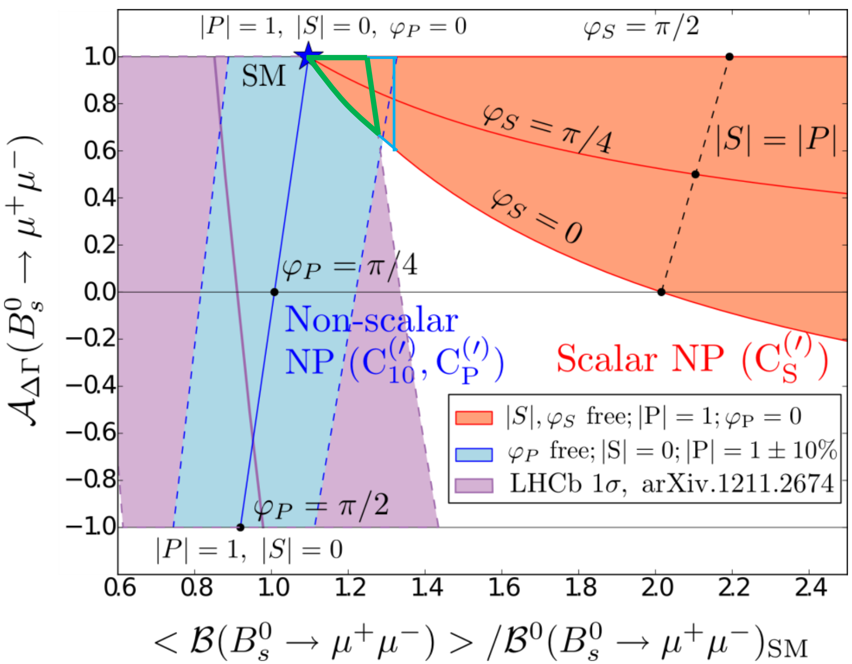


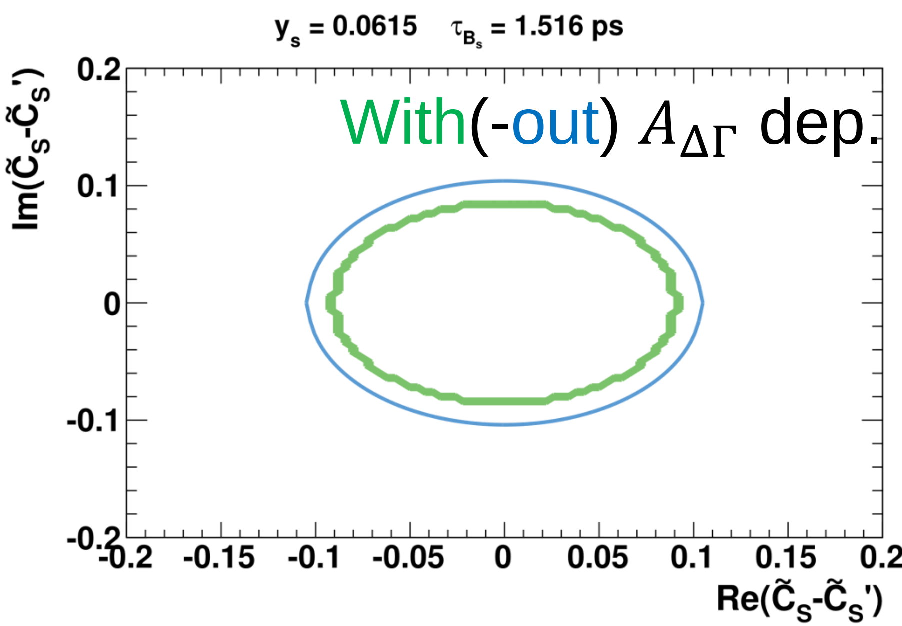
* Modified from De Bruyn et al. PRL 109, 041801, 2012

** Altmannshofer, arXiv:1306.0022

Exploiting $A_{\Delta\Gamma}$ Dependency

- New Physics can enter both $BF_{t=0}$ and $A_{\Delta\Gamma}$
- The expe. BF dependency on $A_{\Delta\Gamma}$ must be accounted for
- The constraints are more accurate and more stringent





- Constraints on scalar are 30% more stringent

Sensitivity Projection on $A_{\Delta\Gamma}$

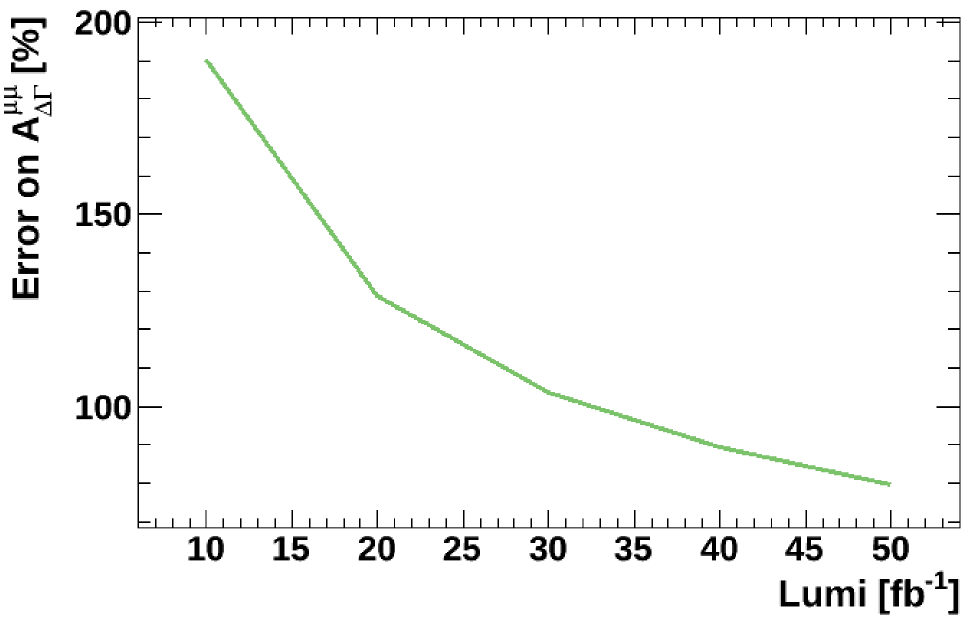
- $A_{\Delta\Gamma}$ can be obtained from effective lifetime,

$$\tau_{eff} = \frac{\int \langle \Gamma(B_s^0(t) \to \mu^+\mu^-)\rangle \times t\, dt}{\int \langle \Gamma(B_s^0(t) \to \mu^+\mu^-)\rangle\, dt} = \frac{\tau_{B_s}}{1-y_s^2} \times \frac{1+2\boldsymbol{A_{\Delta\Gamma}}\, y_s + y_s^2}{1+\boldsymbol{A_{\Delta\Gamma}}\, y_s}$$

- LHCb measured $\tau_{eff}(B_s^0 \to K^+K^-)$ at 7% precision with 552 signal events which give an uncertainty on $A_{\Delta\Gamma}$ of 100%

- Assuming the same sensitivity (S/B =10) for $B_s^0 \to \mu^+\mu^-$, with 50fb^{-1}:

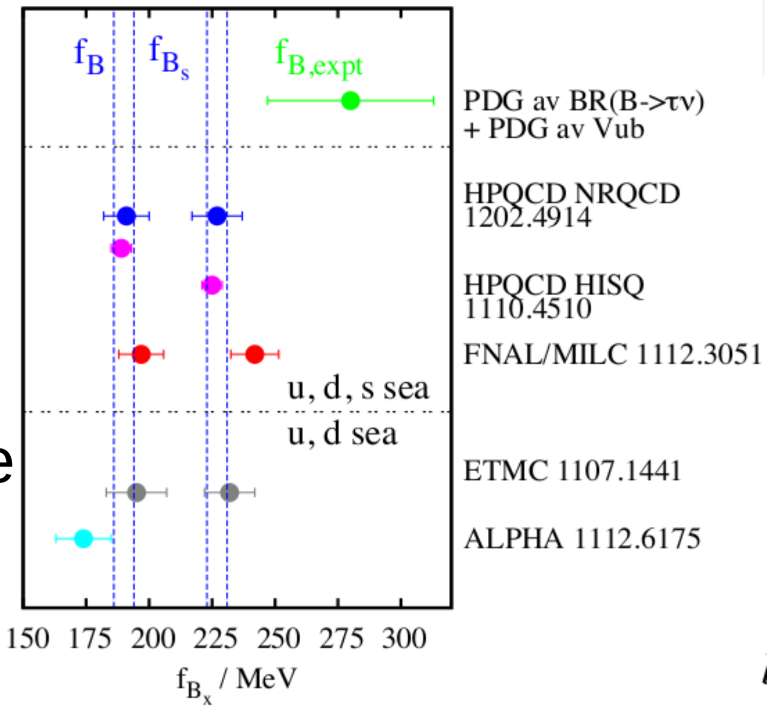
$$A_{\Delta\Gamma}^{\mu\mu} / \delta A_{\Delta\Gamma}^{\mu\mu} = 75\%$$



Uncertainty with new lattice F_B

- Recent works in Lattice QCD claims uncertainties at 1.3%
- Results still discussed
- 'Conservative' approach:

Central value from weighted average + uncertainty of 8 MeV



- If results confirmed:

$$BF(B_s^0 \to \mu^+\mu^-) = 3.57 \pm 0.18 \times 10^{-9}$$

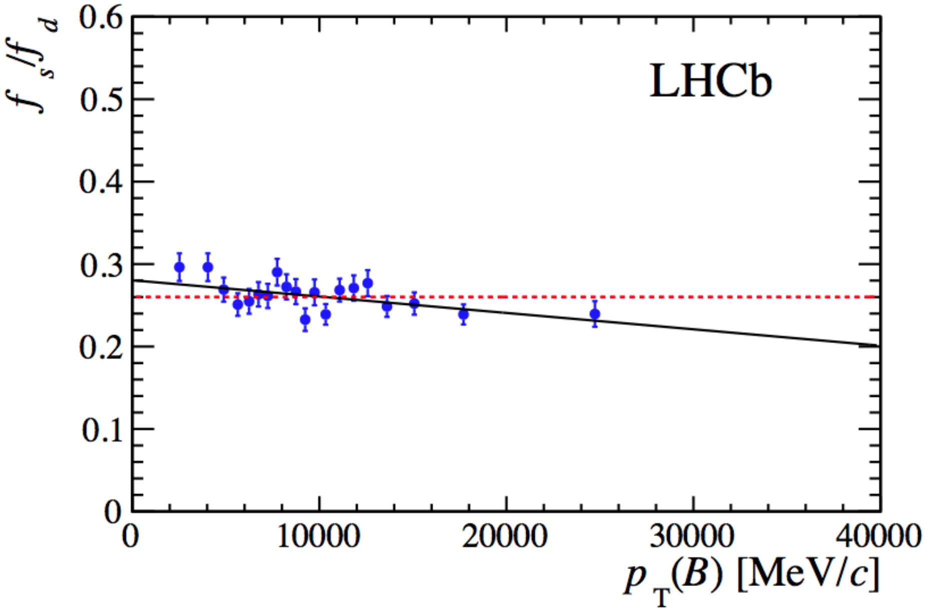
- Dominant uncertainty $|V_{tb}^* V_{ts}|$

Uncertainty Budget		
$\mathbf{F_{Bs}}$	**72.5%**	27.0%
$\lvert V_{tb}^* V_{ts}\rvert$	22.8%	**60.0%**
m_t	3.7%	9.8%
τ_{Bs} and y_s	1.1%	2.8%

Hadronisation Probablility f_s/f_d

- f_s/f_d is measured at LHCb by comparing abundances of:
  - $B_s^0 \to D_s^- \pi^+$, $B^0 \to D^- K^+$ and $B^0 \to D^- \pi^+$ arXiv:111.2357 aka PRD85 032008 (2012)
  - $B_s^0 \to D_s^- \mu^+ X$ and $B^0 \to D^- \mu^+ X$ LHCb-paper-2012-037 in preparation
- Use $B(D_s^- \to K^+ K^- \pi^-)$ and τ_B
- at 7 TeV: $f_s/f_d = 0.259 \pm 0.015$

- p_T dependency small enough to be negligible
- $\sqrt{s}$ dependency checked with $B^+ \to J/\psi K^+$ and $B_s^0 \to J/\psi \phi$: stable within 1σ



Exclusive Backgrounds : $B_s^0 \to K^+ \mu^- \bar{\nu}_\mu$ and $B^0 \to \pi^+ \mu^- \bar{\nu}_\mu$

- Lower contribution from $B_s^0 \to K^+ \mu^- \bar{\nu}_\mu$ explained by:
  - $f_s/f_d = 0.26$
  - $B(B_s^0 \to K^+ \mu^- \bar{\nu}_\mu)/B(B^0 \to \pi^+ \mu^- \bar{\nu}_\mu) = 0.88$
  - $\epsilon_{K\to\mu}/\epsilon_{\pi\to\mu} = 0.28$ (RICH efficiency and $B(K^- \to \mu^- \bar{\nu}_\mu)/B(\pi^- \to \mu^- \bar{\nu}_\mu)$)

BDT Variables

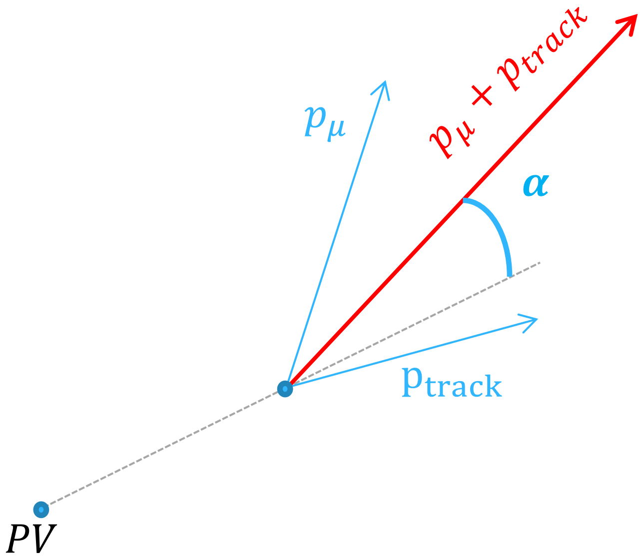
Muon isolation: number of other tracks with which the muon can make a good vertex

Other tracks requierement:

- Long track
- Impact Param Significance with PV > 3

Vertex requirement:

- Angle track-muon<0.27rad
- Distance of Closest Approach < 130 µm
- Distance to PV: 0.5cm<d<4cm
- Distance to SV: -0.15cm<d<30cm
- $\frac{|\overrightarrow{p_\mu}+\overrightarrow{p_{track}}| \sin\alpha}{|\overrightarrow{p_\mu}+\overrightarrow{p_{track}}| \sin\alpha + p_{T,\mu} + p_{T,track}} < 0.6$



BDT Variables

Polarisation Angle:

angle between the muon momentum in the B rest frame and the vector perpendicular to the B momentum and the beam axis

B Isolation:

$$I = \frac{p_{T,B}}{p_{T,B} + \sum_{tracks} p_{T,track}}$$

sum running on the tracks such that $\delta\eta^2 + \delta\phi^2 < 1.0$

MVA Selection Variables

- B Candidate
  - impact parameter*
  - impact parameter χ^2
  - χ^2 of the vertex
  - pointing angle
  - distance of closest approach*
- Muons
  - min IP

*common with BDT