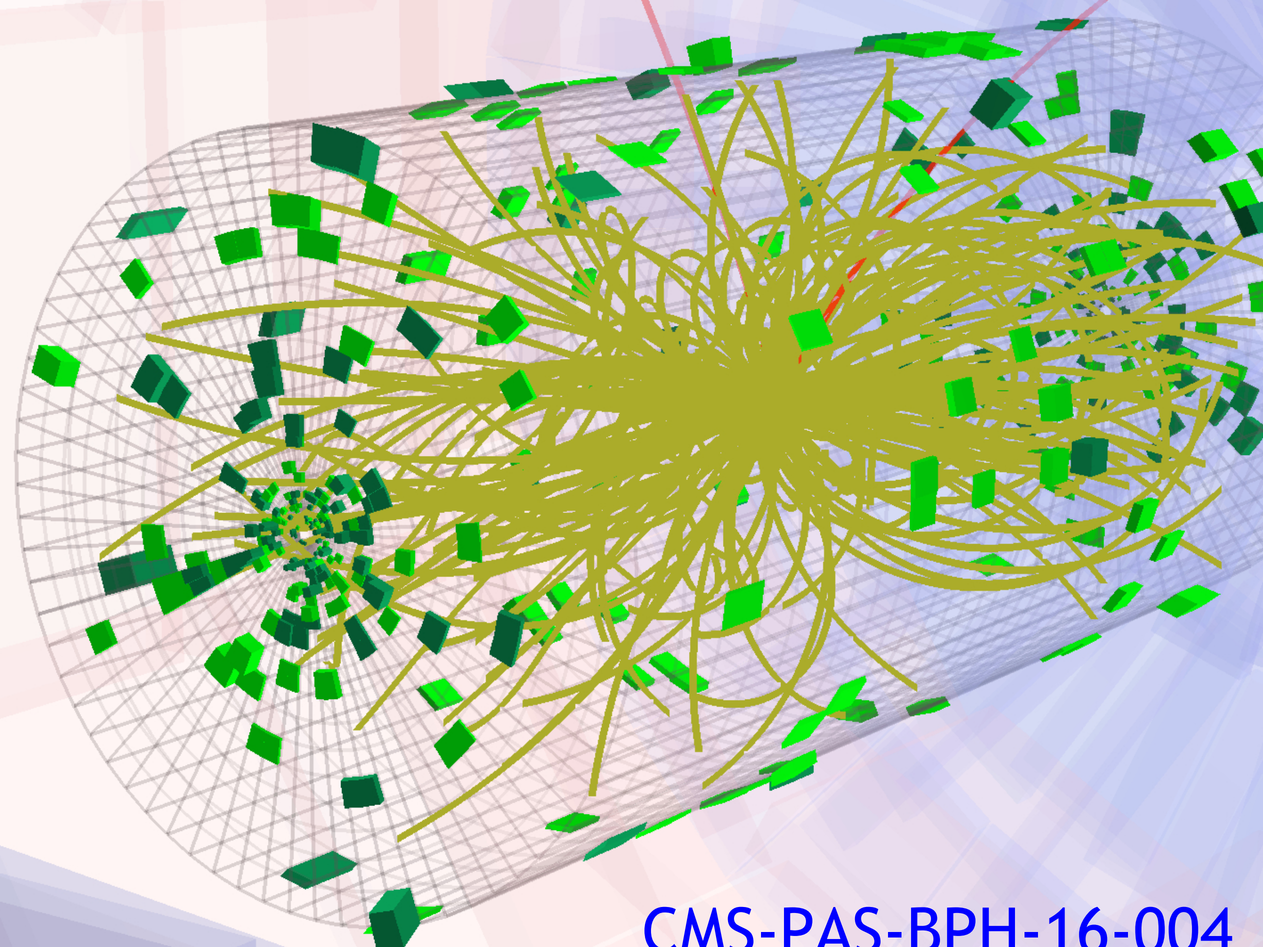
Measurement of $B \to \mu^+\mu^-$ at CMS

Urs Langenegger

(Paul Scherrer Institute)

- Why?
- How?
- And?

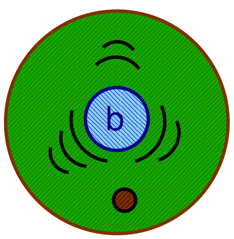


CMS-PAS-BPH-16-004

Introduction

- LHC is a proton-proton collider with high luminosity $\mathcal{L} \approx 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$
  - ▷ large $b\bar{b}$ production cross section $\approx 10\,\text{nb}^{-1}/\,\text{s} \times 500 \times 10^3\,\text{nb} \approx 5 \times 10^6\, b\bar{b}/\,\text{s}$
  - ▷ b quarks form hadrons (mesons and baryons)

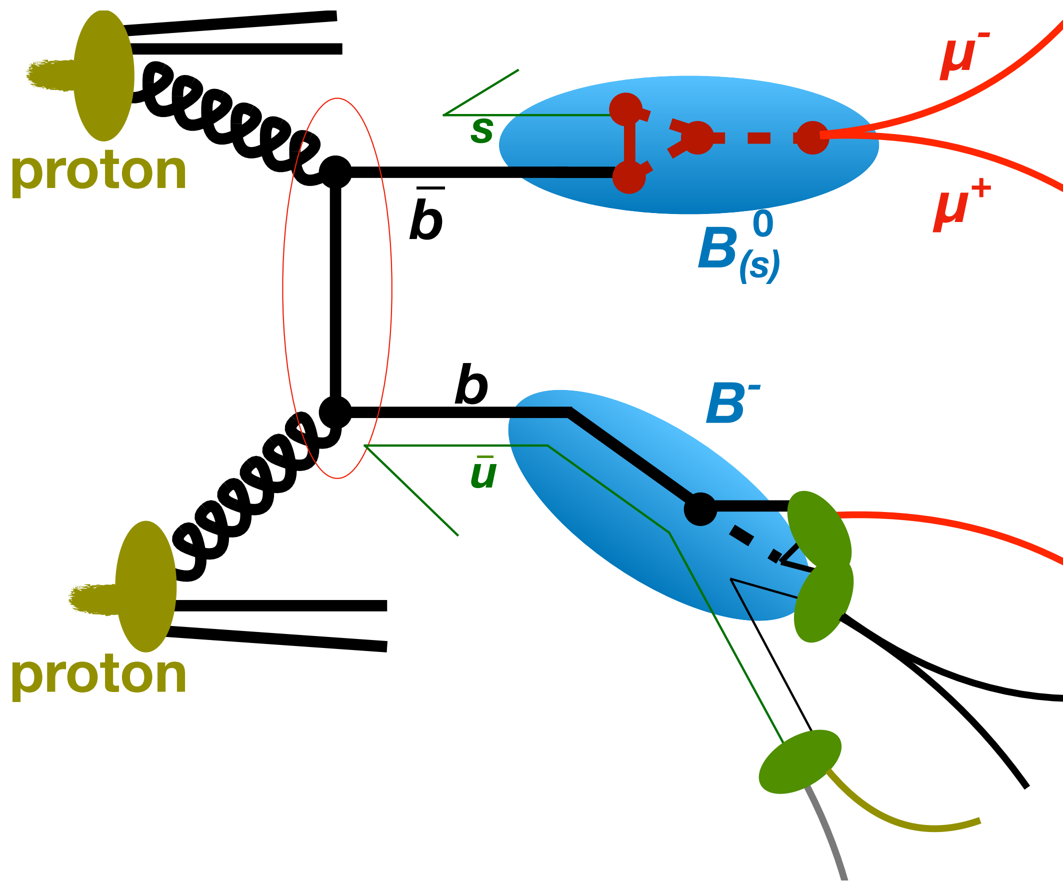
- B mesons
  - ▷ one beauty b quark (heavy)
  - ▷ one spectator quark (light)



  - ▷ mass: $m \approx 5.3\,\text{GeV}$
  - ▷ lifetime: $\tau \approx 1.5\,\text{ps}$

    $c\tau \approx 450\,\mu\text{m}$ → they fly!
  - ▷ 'botanics' with many states:

$$
\begin{aligned}
|B^0\rangle &= |\bar{b}d\rangle & |\bar{B}^0\rangle &= |b\bar{d}\rangle \\
|B^+\rangle &= |\bar{b}u\rangle & |B^-\rangle &= |b\bar{u}\rangle \\
|B^0_s\rangle &= |\bar{b}s\rangle & |\bar{B}^0_s\rangle &= |b\bar{s}\rangle
\end{aligned}
$$

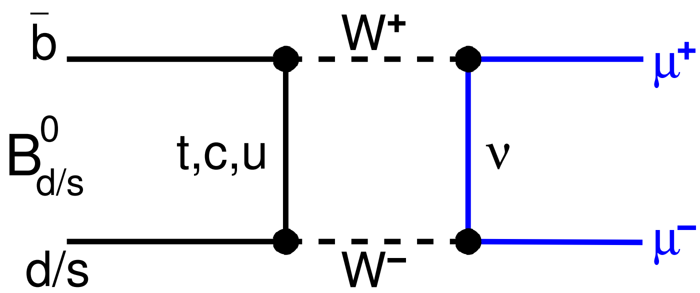


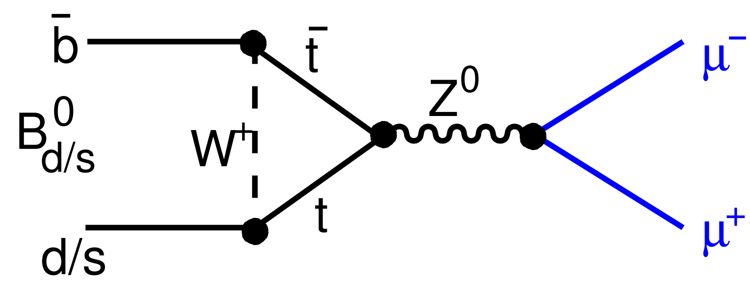
$B \equiv B^0, B^0_s, B^+$

Leptonic B decays

PRL, 112,101801
PRL, 120,111801
1908.07011

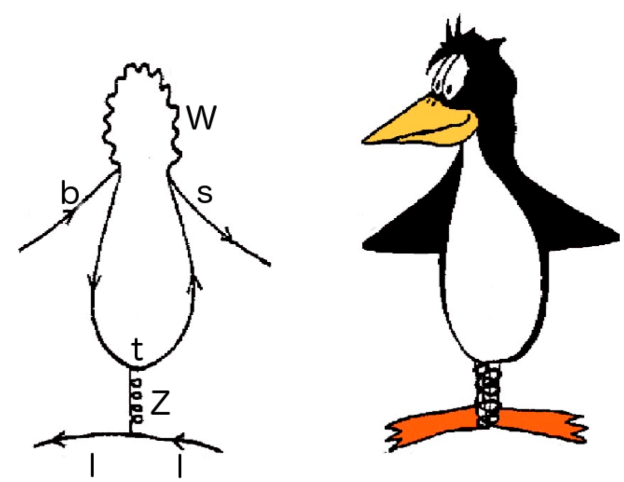
- Leptonic B decays have only leptons (e, μ, τ, ν) in final state
  - for example: $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$
  - many other modes possible (or ‘forbidden’) as well





- They are strongly suppressed
  - SM branching fractions are small

    (ignoring tiny contributions from Higgs boson exchanges)



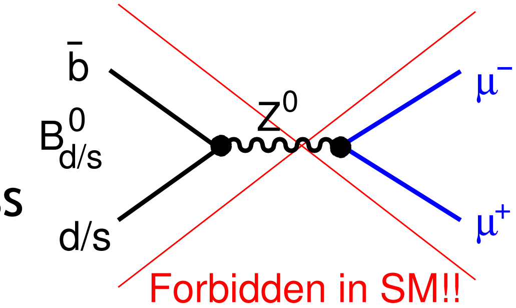
$$\bar{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)_{\mathrm{SM}} = (3.66 \pm 0.14) \times 10^{-9}$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-)_{\mathrm{SM}} = (1.03 \pm 0.05) \times 10^{-10}$$

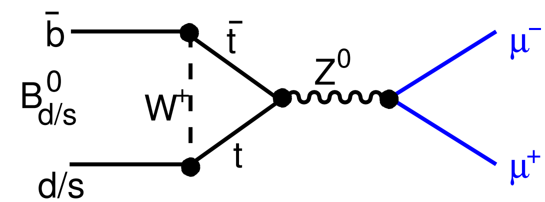
  - SM expectation: 4-5% theoretical uncertainty!
  - $\bar{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)$: decay time-integrated $\mathcal{B}$

$|B^0_{s\mathrm{H,L}}\rangle = p|B^0_s\rangle \pm q|\bar{B}^0_s\rangle$
with different lifetimes

Why are they suppressed in the SM?

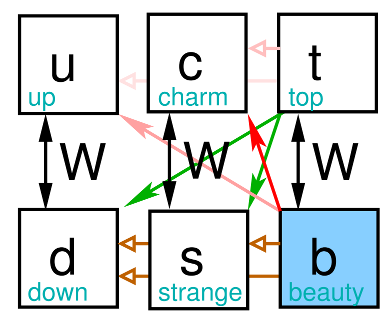
- Effective flavor-changing neutral currents
  - no flavor-changing neutral currents in SM
  - Penguin and box diagrams, but no tree-level process
  - CKM-suppression of $B^0 \to \mu^+\mu^-$ vs. $B^0_s \to \mu^+\mu^-$:
    $|V_{td}|^2 < |V_{ts}|^2$
  - Cabibbo-Kobayashi-Maskawa matrix



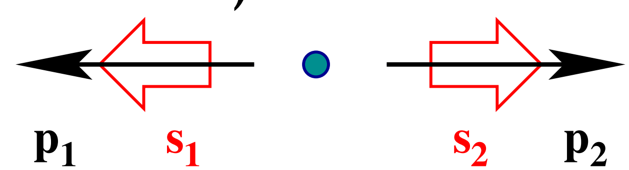


$$\begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{weak}} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{V_{\text{CKM}}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_{\text{mass}}$$

$$|V_{\text{CKM}}| = \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.224 & 0.974 & 0.042 \\ 0.009 & 0.041 & 0.999 \end{pmatrix}$$

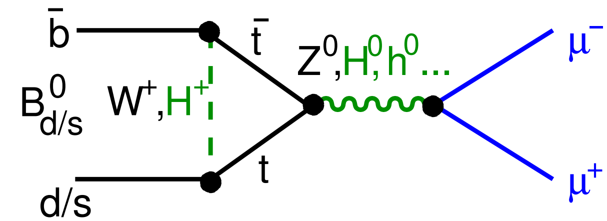


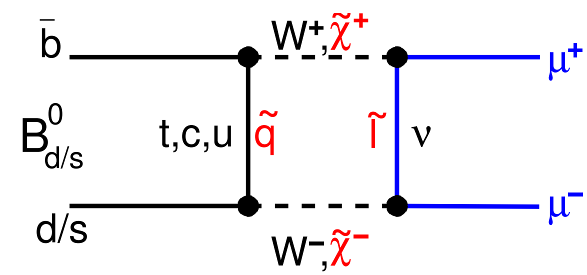
- Helicity suppressed ($V - A$ interaction in SM)
  - B mesons have no spin
  - μ have spin $1/2$
  - weak interaction is left handed



'Why?' Search for 'BSM' physics!

- Already small additional decay width contributions visible
  - ▷ because the SM decay width is so small
- Sensitivity to 'beyond-SM' physics
  - ▷ no helicity suppression
    - scalar couplings
    - pseudo-scalar couplings
  - ▷ 'extended' Higgs boson sectors
  - ▷ flavor 'violation': $B_s^0 \to \mu^+\mu^-$ vs $B^0 \to \mu^+\mu^-$



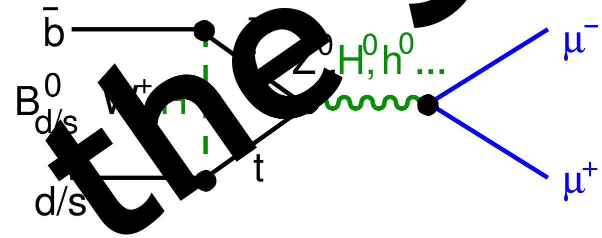


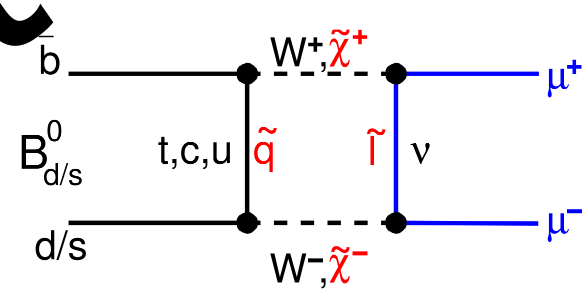
- Two approaches
  - ▷ model-independent: effective field theory
    - parametrize new physics with operators and (Wilson) coefficients
    - can include correlations to other processes $R_K^{(*)}, P_5', \ldots$
  - ▷ 'top-down': specific model
    - new particles extending the SM world
    - correlations between many processes precisely calculable
    - more specific than above, but very model dependent

Why? Search not for 'BSM' physics!

much better: Test the SM!

- Already small additional decay width contributions visible
  - ▷ because the SM decay width is so small
- Sensitivity to 'beyond-SM' physics
  - ▷ no helicity suppression
    - scalar couplings
    - pseudo-scalar couplings
  - ▷ 'extended' Higgs boson sectors
  - ▷ flavor 'violation': $B^0_s \to \mu^+\mu^-$ vs $B^0 \to \mu^+\mu^-$





- Two approaches
  - ▷ model-independent: effective field theory
    - parametrize new physics with operators and (Wilson) coefficients
    - can include correlations to other processes $R_K^{(*)}, P_5', \ldots$
  - ▷ 'top-down': specific model
    - new particles extending the SM world
    - correlations between many processes precisely calculable
    - more specific than above, but very model dependent

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

PRL, 109,041801
PRL, 118,191801

- A second independent observable: $B_s^0 \to \mu^+\mu^-$ effective lifetime
  - ▷ measure B_s^0 lifetime with $B_s^0 \to \mu^+\mu^-$ decays

$$\tau_{\mu^+\mu^-} \equiv \frac{\int_0^\infty t\,\Gamma(B_s(t)\to\mu^+\mu^-)\,dt}{\int_0^\infty \Gamma(B_s(t)\to\mu^+\mu^-)\,dt} = \frac{\tau_{B_s^0}}{1-y_s^2}\left(\frac{1+2\mathcal{A}_{\Delta\Gamma}^{\mu^+\mu^-}\,y_s+y_s^2}{1+\mathcal{A}_{\Delta\Gamma}^{\mu^+\mu^-}\,y_s}\right)$$

  - ▷ allows determination of $\mathcal{A}_{\Delta\Gamma}^{\mu^+\mu^-}$

    B_s^0 mean lifetime $\tau_{B_s^0} = 1.510 \pm 0.005\,$ps

    B_s^0 decay width difference $\Delta\Gamma_s$,

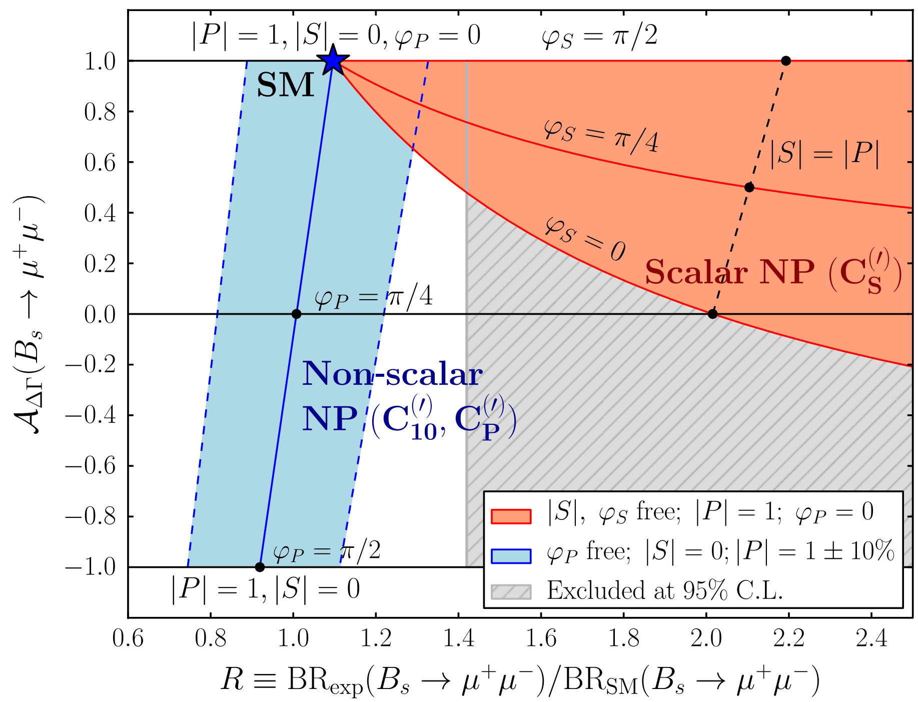
    $$\Delta\Gamma_s \equiv \Gamma_{sL} - \Gamma_{sH} = 0.088 \pm 0.006\,\text{ps}^{-1}$$

    $$y_s \equiv \tau_{B_s^0}\Delta\Gamma_s/2 = 0.062 \pm 0.006$$

  → scalar vs. non-scalar 'new physics'
  - ▷ SM prediction:

    only the $B_{s\,H}^0$ decays to dimuons:

    $$\mathcal{A}_{\Delta\Gamma}^{\mu^+\mu^-} = +1 \to \tau_{\mu^+\mu^-}^{\text{SM}} = 1.615\,\text{ps}$$

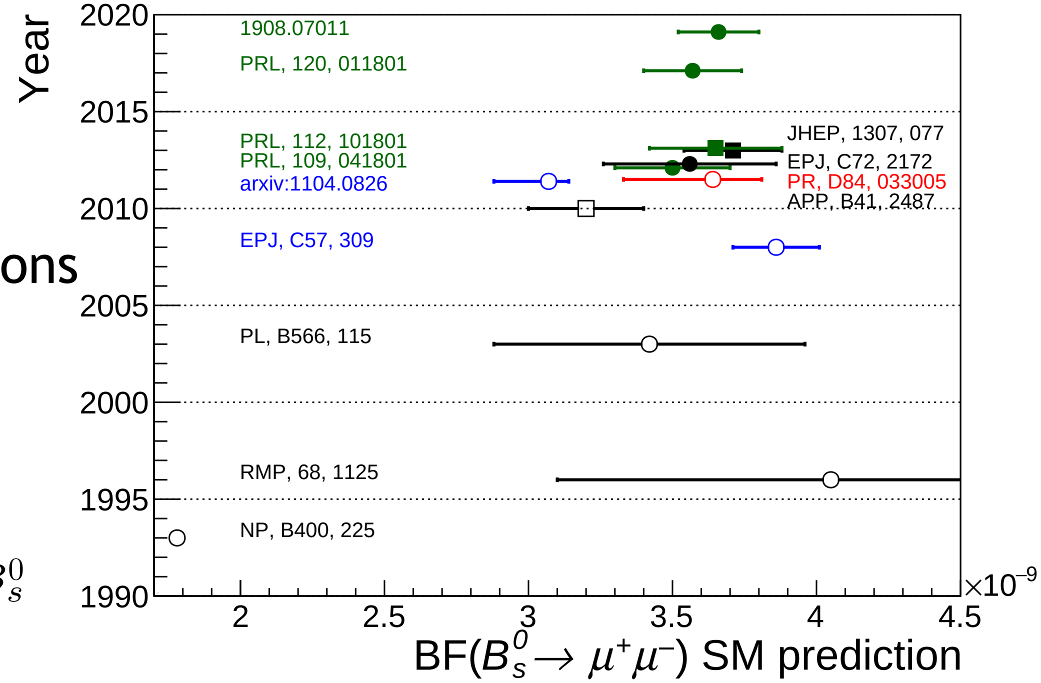


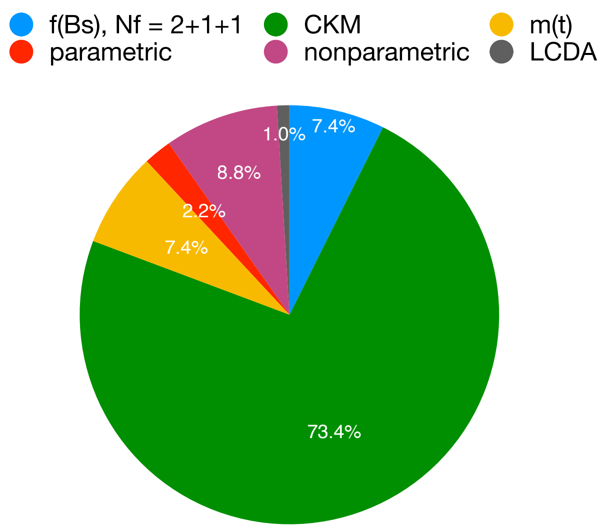
- One measurement to date:
  - ▷ LHCb 2017: $\tau_{\mu^+\mu^-} = 2.04 \pm 0.44 \pm 0.05\,$ps

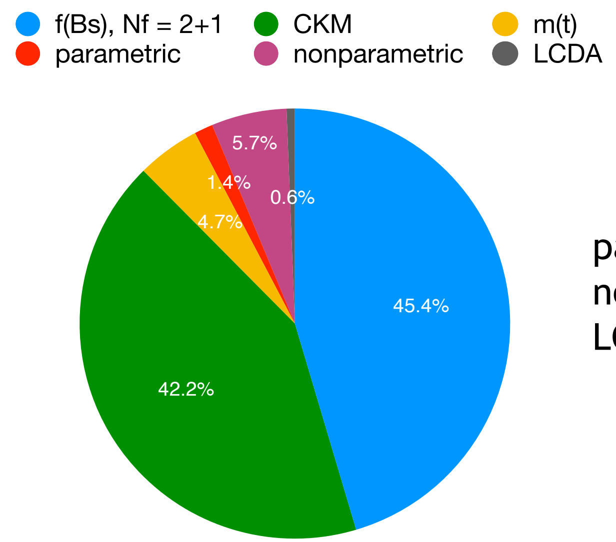
Theoretical context

1908.07011
1902.08191

- SM predictions based on
  - ▷ NNLO QCD
  - ▷ NLO EW
  - ▷ power-enhanced ll QED corrections (larger than anticipated)
  - ▷ external input (important!)
- Error budget (relative)
  - ▷ long focus on decay constant $f_{B^0_s}$
  - ▷ lattice QCD: $f_{B^0_s}$ at 0.6%! (FNAL/MILC, tbc)



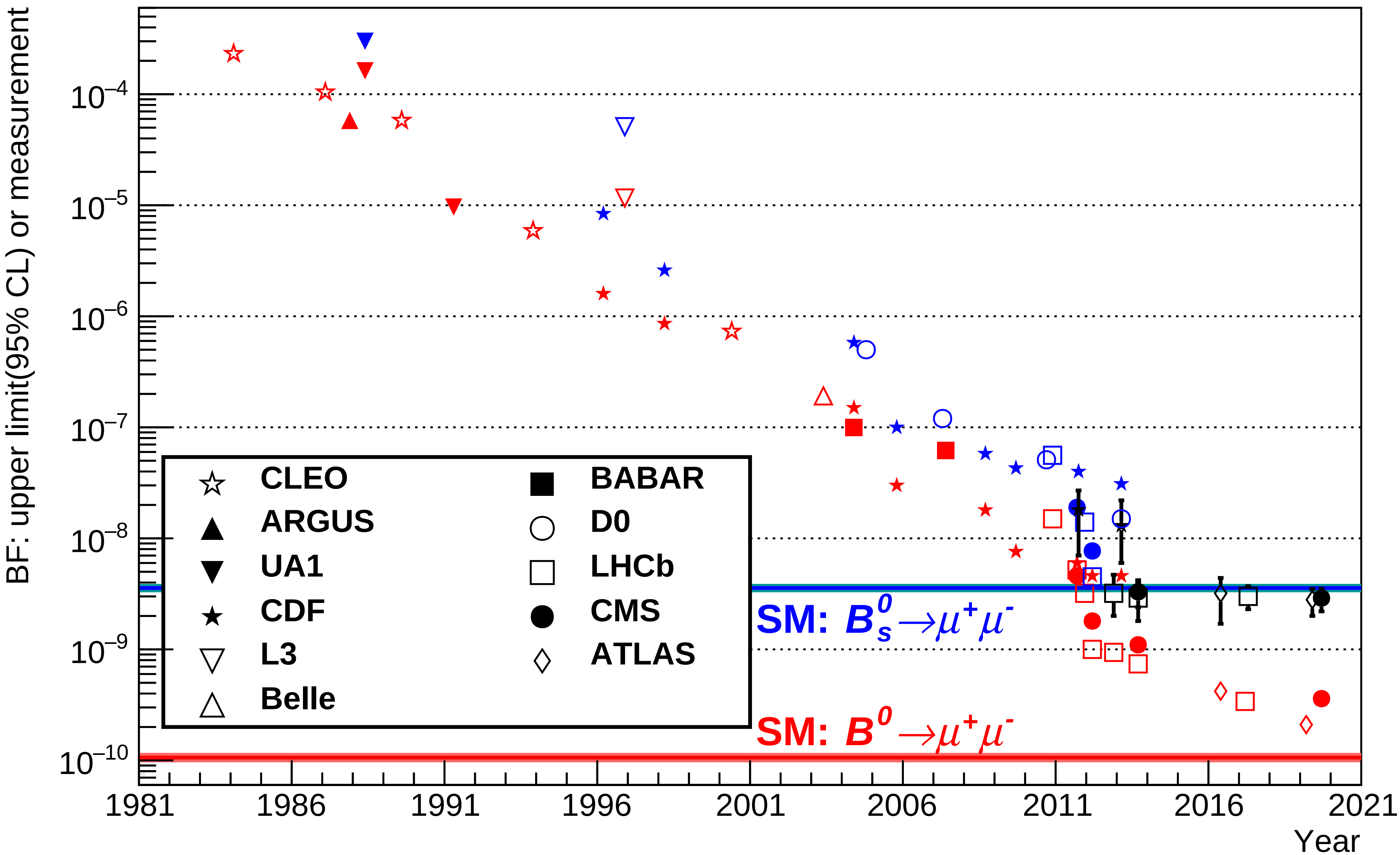




parametric: Γ_s, α_s
nonparametric: μ_W, μ_b, h.o.
LCDA: λ_B, $\sigma_{1,2}$

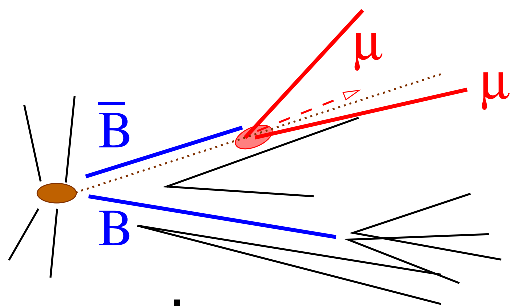
Experimental context

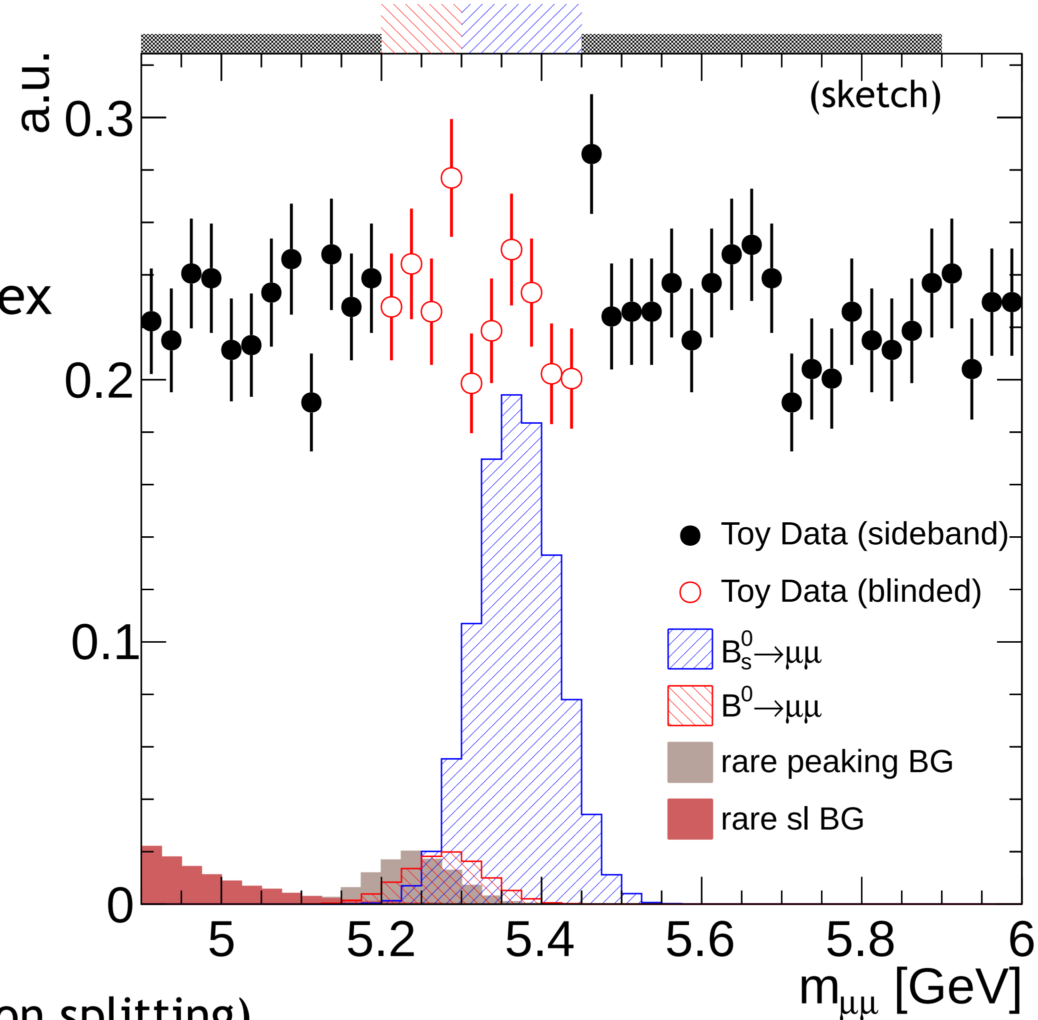
- $\overline{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$ with long history:



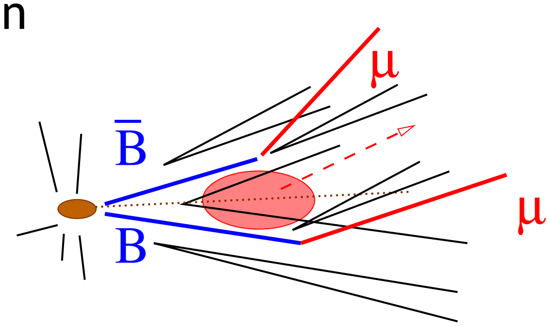
Analysis overview

- Signal $B^0_s \to \mu^+\mu^-$
  - ▷ two muons from one decay vertex
    - invariant mass
    - secondary vertex, isolated
    - momentum in flight direction





- Background
  - ▷ combinatorial (from sidebands)
    - two semileptonic (B) decays (gluon splitting)
    - one semileptonic (B) decay and one misidentified hadron
  - ▷ rare single B decays (from MC simulation)
    - non-peaking, e.g. $B^0_s \to K^-\mu^+\nu$, $\Lambda_b \to p\mu^+\nu$
    - peaking, e.g. $B^0_s \to K^+K^-$



- The goal: keep signal, very strongly reduce background!

Methodology

- Measurement of $B_s^0 \to \mu^+\mu^-$ relative to normalization channel:

$$\bar{\mathcal{B}}(B_s^0 \to \mu^+\mu^-) \quad = \quad \frac{n^{\rm obs}_{B_s^0}}{N(B^+ \to J/\psi\, K^+)} \frac{\varepsilon^{tot}_{B^+}}{\varepsilon^{tot}_{B_s^0}} \frac{f_u}{f_s} \mathcal{B}(B^+ \to J/\psi\,[\mu^+\mu^-]K)$$

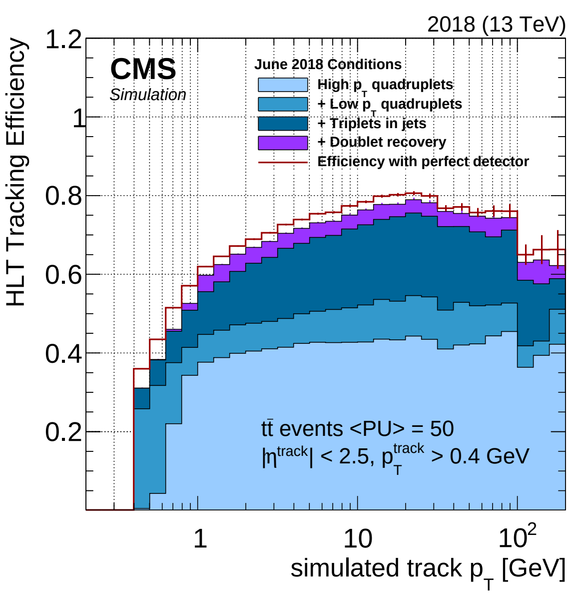
  - ▷ $B^+ \to J/\psi\, K^+$, $J/\psi \to \mu^+\mu^-$, with $\mathcal{B}(B^+ \to J/\psi\, K^+) = (1.01 \pm 0.03) \times 10^{-3}$

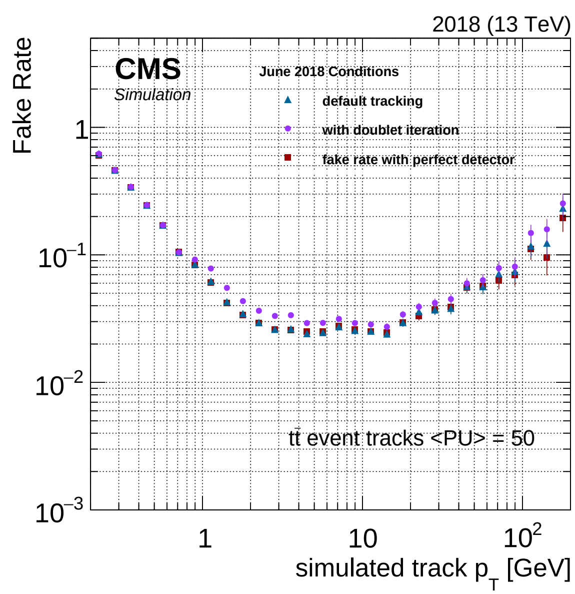
- Reconstructed decays for this result:
  - ▷ $B \to \mu^+\mu^-$: 'signal' sample
  - ▷ $B^+ \to J/\psi\, K^+$: 'normalization' sample
  - ▷ $B_s^0 \to J/\psi\, \phi$: 'control' sample for B_s^0 mesons

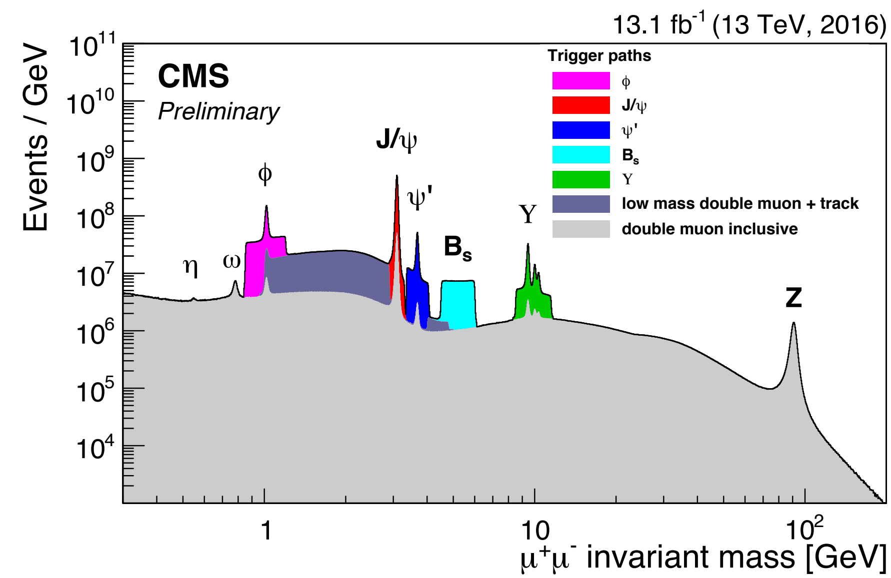
- Analysis steps
  - ▷ strict muon identification with boosted decision tree
  - ▷ tight candidate selection with (another) boosted decision tree
  - ▷ unbinned (extended) maximum likelihood fits to selected events
    - branching fractions $\bar{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$
    - effective lifetime $\tau_{\mu^+\mu^-}$

B physics trigger

- B physics mostly triggered with (displaced) dimuon $(+X)$ triggers
  (displaced from pp collisions; B hadrons have a lifetime of about 1.5 ps)
  - ▷ other setups are in progress/under analysis
- L1: hardware trigger based on muons
  - ▷ $< 4\,\mu$s latency
  - ▷ no explicit muon $p_\perp$ threshold (strong B field implies $p_\perp > 3$ GeV in barrel)
- HLT: high-level trigger
  - ▷ full tracking and vertexing
  - ▷ specific trigger paths

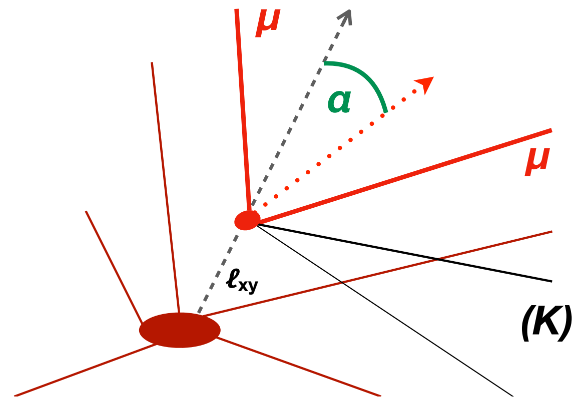


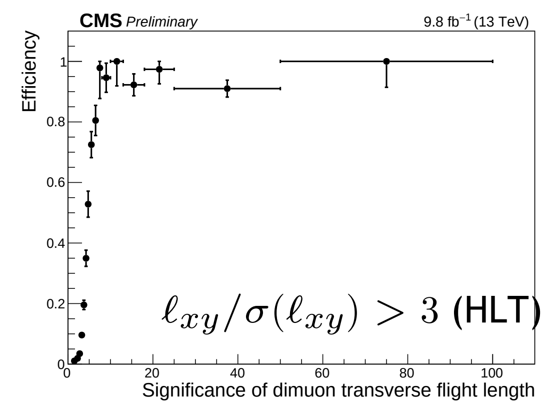




Displaced J/ψ and $B_s^0 \to \mu^+\mu^-$ triggers

- HLT 'displaced' J/ψ
  - two muons with opposite charge
  - $2.9 < m_{\mu\mu} < 3.3$ GeV
  - $\ell_{xy}/\sigma(\ell_{xy}) > 3$
  - $\cos\alpha > 0.9$, $\mathcal{P}(\chi^2/dof) > 10\%$
- HLT 'displaced' J/ψ + track(s)
  - two muons with opposite charge
    $2.9 < m_{\mu\mu} < 3.3$ GeV, $\ell_{xy}/\sigma(\ell_{xy}) > 3$
  - $\cos\alpha > 0.9$, $\mathcal{P}(\chi^2/dof) > 10\%$
  - invariant mass requirements on tracks
    (targeted towards $\phi \to K^+K^-$)
- HLT $B_s^0 \to \mu^+\mu^-$
  - two muons with opposite charge
  - inv. mass $4.8 < m_{\mu\mu} < 6.0$ GeV
  - $p_\perp > 4.0(3.5)$ GeV, $\mathcal{P}(\chi^2/dof) > 0.5\%$
  - no displacement requirement!

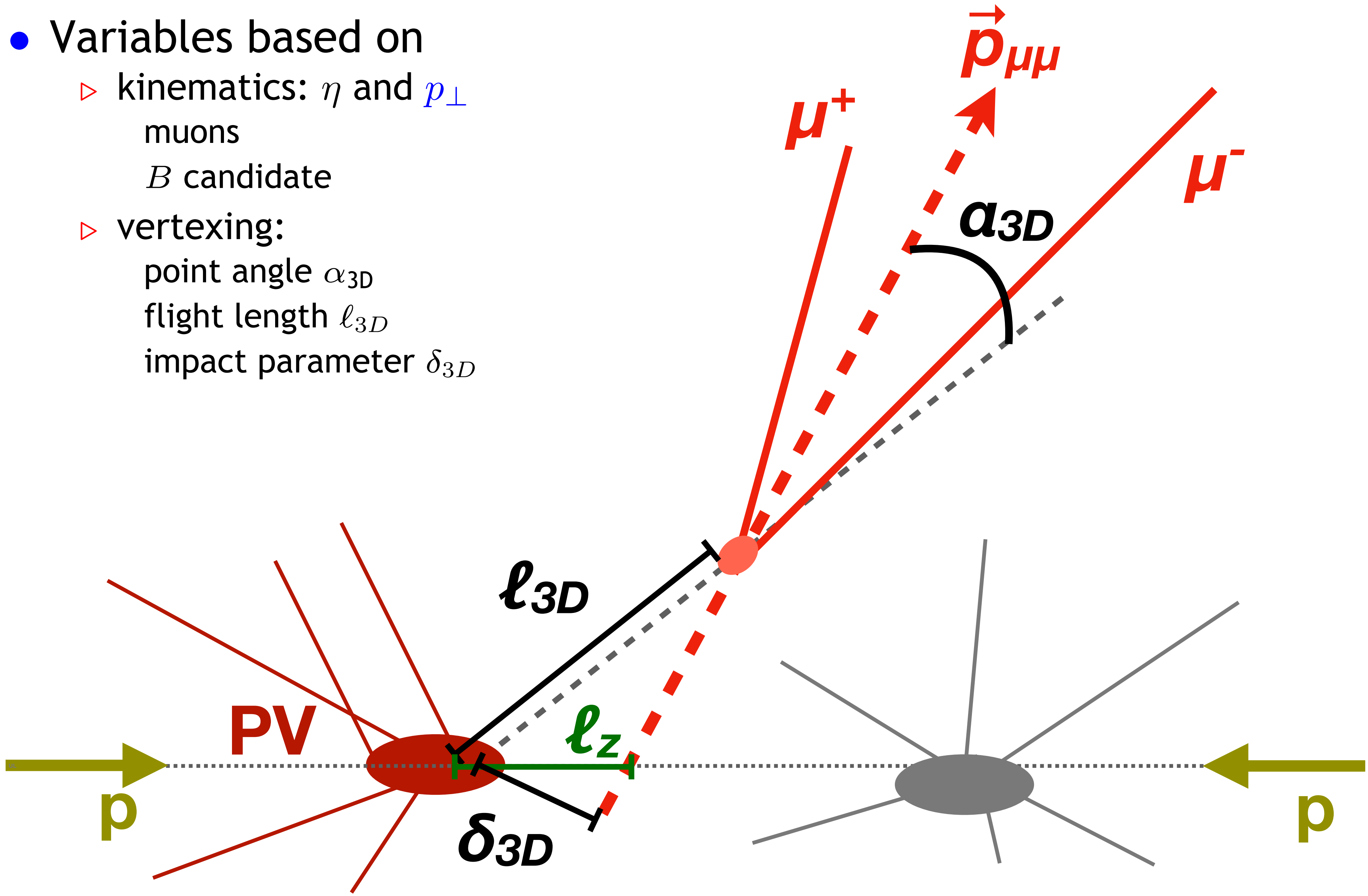




Dataset
- 2011: 5 fb^{-1} at $\sqrt{s} = 7$ TeV
- 2012: 20 fb^{-1} at $\sqrt{s} = 8$ TeV
- 2016: 36 fb^{-1} at $\sqrt{s} = 13$ TeV

Reconstruction I

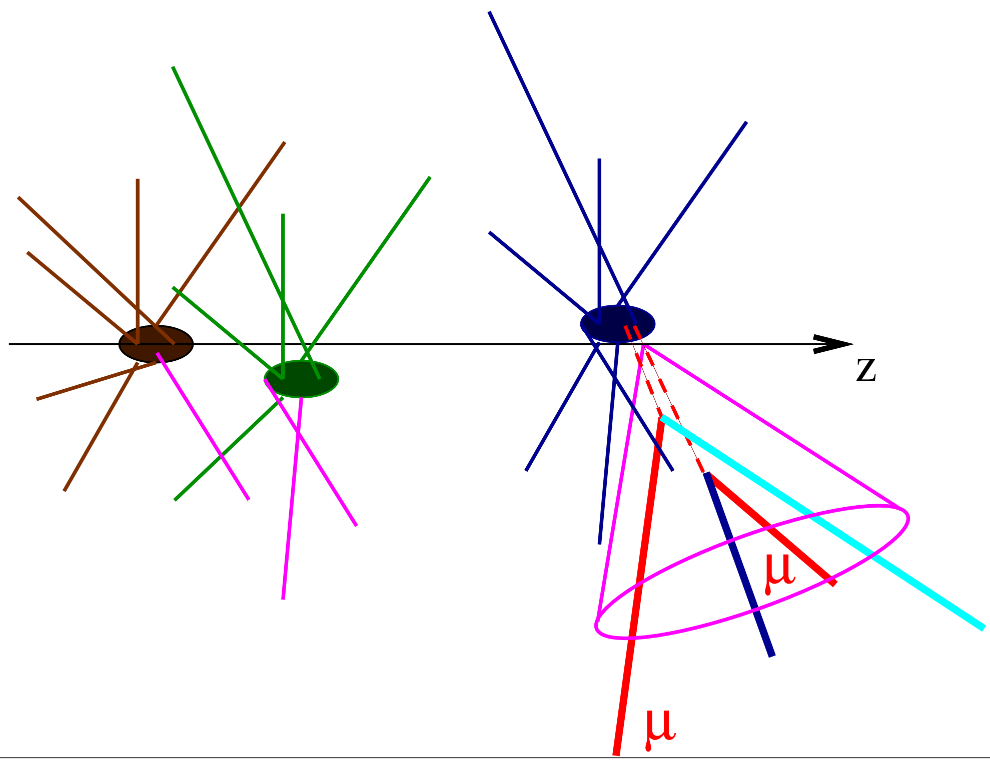
- Variables based on
  - kinematics: η and $p_\perp$
    - muons
    - B candidate
  - vertexing:
    - point angle α_{3D}
    - flight length ℓ_{3D}
    - impact parameter δ_{3D}



Reconstruction II

- **Isolation** (optimized cuts for background rejection and data/MC similarity)
  - $I \equiv p_{\perp B}/(p_{\perp B} + \sum_{\text{trk}} p_\perp)$: $p_\perp > 0.9, \Delta R < 0.7, d_{\text{ca}} < 0.05$ cm
  - $I_\mu \equiv p_{\perp\mu}/(p_{\perp\mu} + \sum_{\text{trk}} p_\perp)$: $p_\perp > 0.5, \Delta R < 0.5, d_{\text{ca}} < 0.1$ cm
  - $N_{\text{trk}}^{\text{close}}$: count tracks with $p_\perp > 0.5$ GeV and $d_{\text{ca}} < 0.03$ cm
  - d_{ca}^0: minimum d_{ca} of these tracks to B-SV

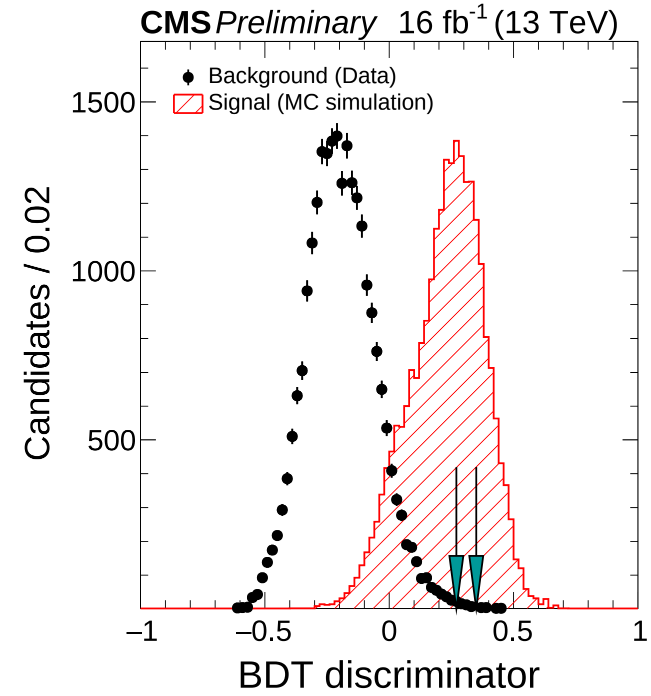
  ($\sum_{\text{trk}}$ w/ tracks from B-PV or no other PV, but passing d_{ca} requirement)

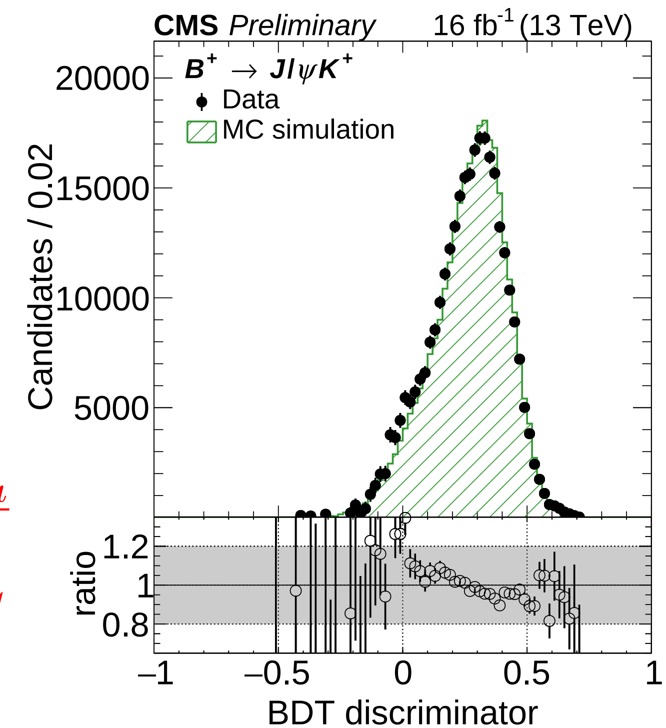


Multi-variate analysis

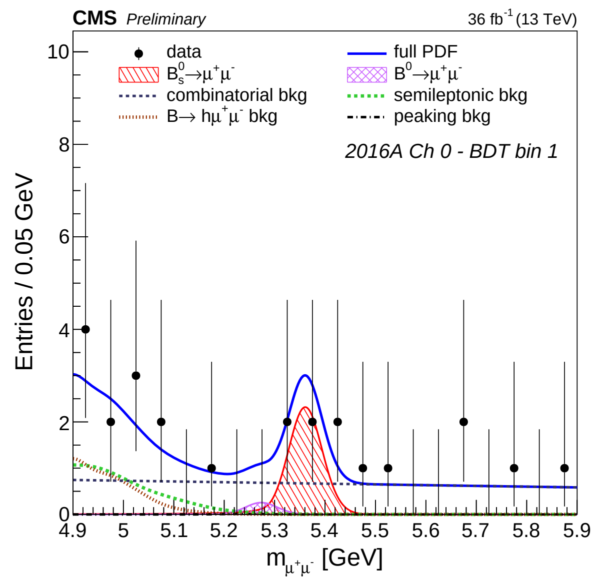
- Boosted decision tree
  - ▷ Run 1: BDT unchanged wrt PRL,111,101804
  - ▷ 2016: new BDT trained (same variables)
- BDT training (TMVA)
  - ▷ signal: $B_s^0 \to \mu^+\mu^-$ MC simulation
  - ▷ background: data dimuon sidebands
  - ▷ avoid selection bias
    - split data randomly into three subsets (0,1,2)
    - train on 0, test on 1, apply on 2. etc.
  - $\to$ in each channel, have 3 BDTs
  - ▷ many validation studies
  - ▷ defines categories for best sensitivity
- Systematic uncertainty
  - ▷ double ratio D
  - ▷ 5-10% on efficiency ratio
  - ▷ 0.07 ps on effective lifetime

$$D = \frac{\left[\frac{\varepsilon(B^+\to J/\psi\, K^+)}{\varepsilon(B_s^0\to J/\psi\,\phi)}\right]_{data}}{\left[\frac{\varepsilon(B^+\to J/\psi\, K^+)}{\varepsilon(B_s^0\to J/\psi\,\phi)}\right]_{MC}}$$





Fit model



- 3D Fit for $\overline{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$

$$P(m_{\mu\mu};\sigma(m_{\mu\mu})) \times P(\sigma(m_{\mu\mu})/m_{\mu\mu}) \times P(\mathcal{C})$$

  - ▷ dimuon mass $m_{\mu\mu}$
  - ▷ per-event dimuon mass resolution $\sigma(m_{\mu\mu})$
  - ▷ $\mathcal{C}$: binary distribution for dimuon bending configuration (against possible bias)
    $\mathcal{C}(\pm 1)$: bending towards (away from) each other

- Components of model

Component	Mass	Width	Mass resolution
Signal	CB	KEYS, $\sigma_{\text{CB}} = \kappa \times \sigma(m_{\mu\mu})$	KEYS
Background hh	CB+G	KEYS	KEYS
Background $h\mu\mu$, $h\mu\nu$	KEYS	n/a	KEYS
Combinatorial background	Bernstein pol1	n/a	KEYS (sideband)

(CB: crystal-ball, G: Gaussian)

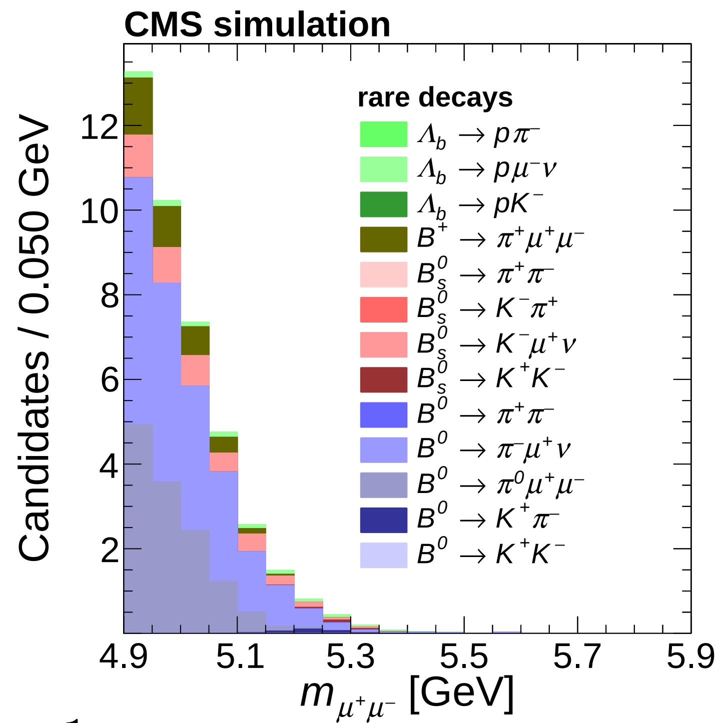
  - ▷ 2 parameters of interest: $\overline{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)$ and $\mathcal{B}(B^0 \to \mu^+\mu^-)$
  - ▷ constraints on nuisance parameters
    - gaussian: f_s/f_u, $B^+ \to J/\psi K^+$, efficiency ratios
    - lognormal: rare background yields

- 1-3% systematic error from unknown $B_s^0 \to \mu^+\mu^-$ eff. lifetime

Rare background yields

- Rare background yield expectations
  - ▷ known branching fractions
  - ▷ absolute yield from normalization sample:

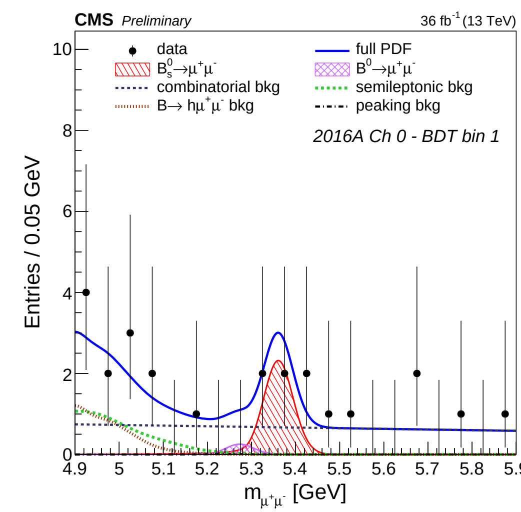
$$n^{\rm exp}_{B_x\to hh} = \frac{\varepsilon^{tot}_{B}}{\varepsilon^{tot}_{B^+}}\frac{f_u}{f_x}\frac{\mathcal{B}(B_x\to hh)}{\mathcal{B}(B^+\to J/\psi\,[\mu^+\mu^-]K)} \times N(B^+\to J/\psi\,[\mu^+\mu^-]K)$$



with

$$\varepsilon^{B\to hh}_{tot} = w_+(p_\perp,\eta)\times w_-(p_\perp,\eta)\times \varepsilon^{(BDT)}_{ana}\times A\times\frac{1}{2}\varepsilon^{signal}_{trig}$$

  - ▷ rare hadronic decays: complete set
  - ▷ rare sl decays: incomplete set/low statistics
    - $\rightarrow$ scale factor in low sideband
  - ▷ extensive validation with inverted muon ID selection
  - ▷ new muon ID: peaking background is very small



Results I

- Branching fractions from 3D UML fit:

$$\overline{\mathcal{B}}(B_s^0 \to \mu^+\mu^-) = [2.9^{+0.7}_{-0.6}(\text{exp}) \pm 0.2(f_s/f_u)] \times 10^{-9}$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) = (0.8^{+1.4}_{-1.3}) \times 10^{-10}$$
$$\mathcal{B}(B^0 \to \mu^+\mu^-) < 3.6 \times 10^{-10} \quad (95\%\ \text{CL})$$
$$(3.0 \times 10^{-10}\ \text{expected})$$

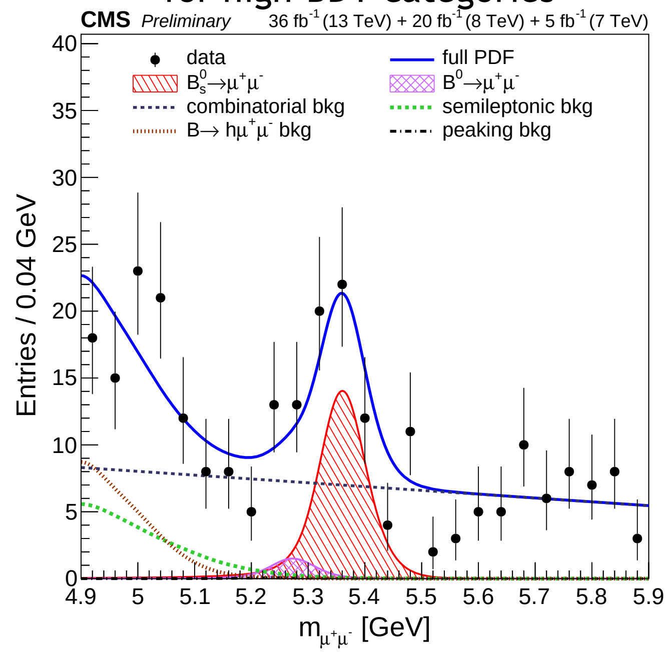
  - ▷ Observed (expected) significance:

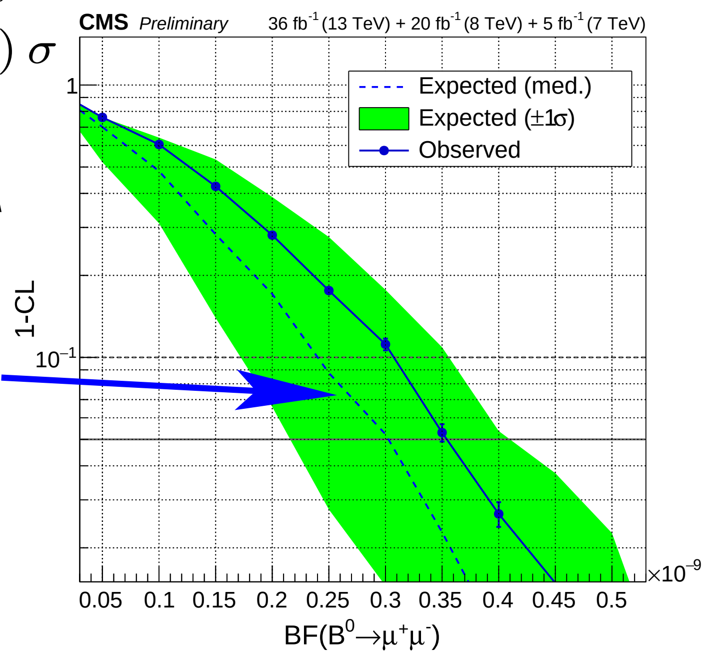
$B_s^0 \to \mu^+\mu^-$: $5.6(6.5)\,\sigma$

$B^0 \to \mu^+\mu^-$: $0.6(0.8)\,\sigma$

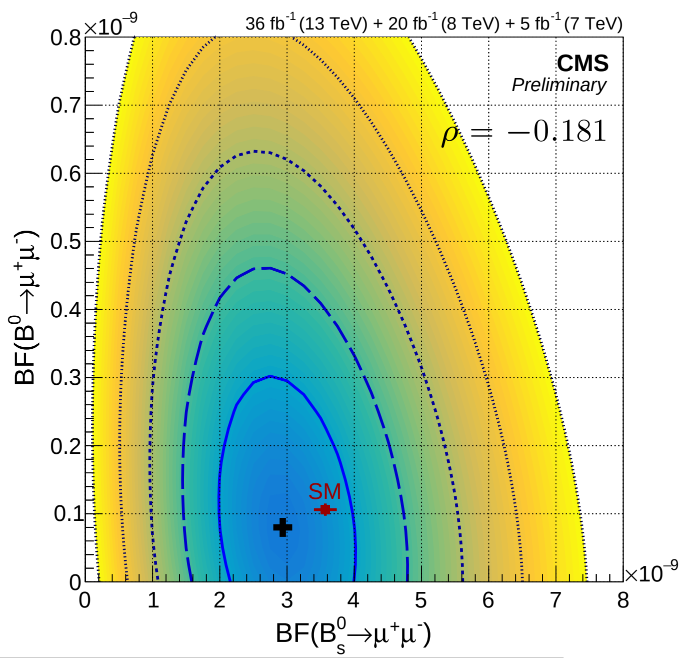
⇒ Consistent with SM

Combined mass projection for high-BDT categories





$B^0 \to \mu^+\mu^-$ not included



Results II

- Primary result from 2D UML fit:

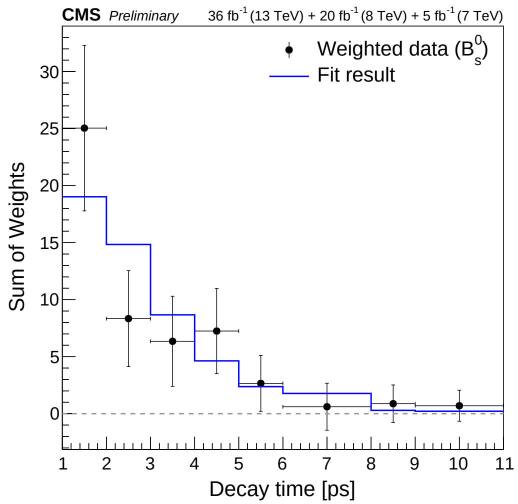
$$\tau(B_s^0 \to \mu^+\mu^-) \quad = \quad 1.70^{+0.61}_{-0.44}\ \text{ps}$$

  - ▷ systematic error small: 0.09 ps
  - ▷ expected error: $\binom{+0.39}{-0.30}$ ps

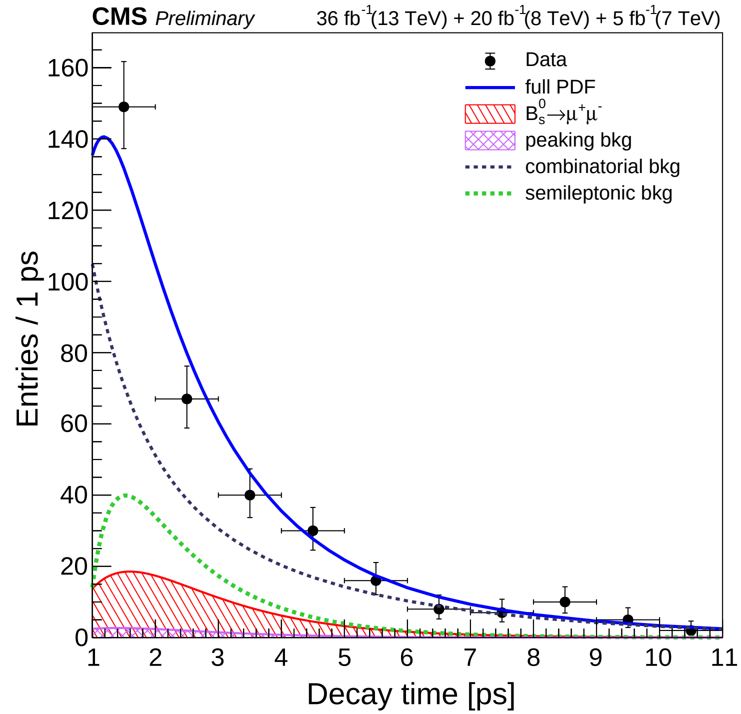
$\Rightarrow$ Consistent with SM

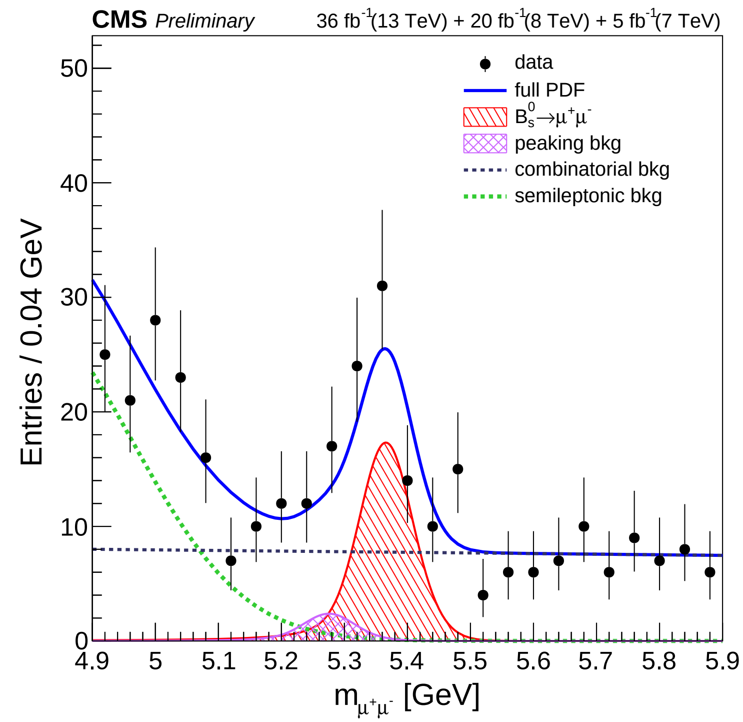
- Result from $sPlot$ method:

$$\tau(B_s^0 \to \mu^+\mu^-) \quad = \quad 1.55^{+0.52}_{-0.33}\ \text{ps}$$



expected error: $\binom{+0.49}{-0.31}$ ps





Conclusions

- $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ decays with Run 1 and 2016 data
  - ▷ update of branching fraction measurements

$$\overline{\mathcal{B}}(B_s^0 \to \mu^+\mu^-) = [2.9^{+0.7}_{-0.6}(\text{exp}) \pm 0.2(f_s/f_u)] \times 10^{-9}$$

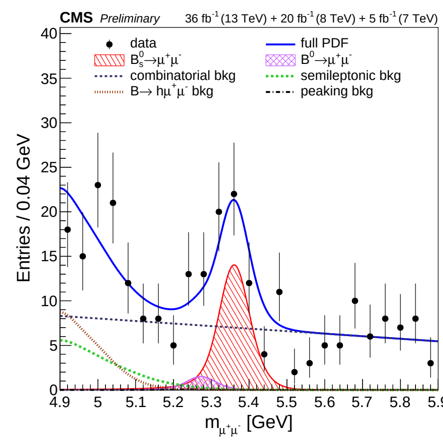
$$\mathcal{B}(B^0 \to \mu^+\mu^-) < 3.6 \times 10^{-10} \quad (95\%\text{CL})$$

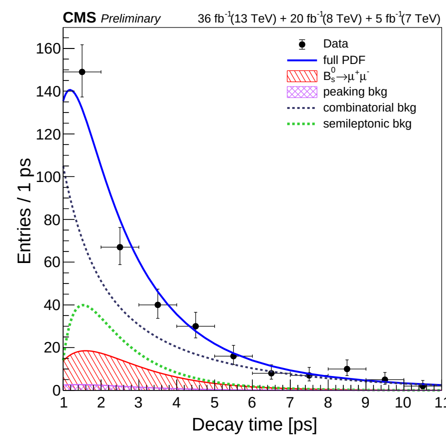
($B_s^0 \to \mu^+\mu^-$ significance: $5.6\,\sigma$ obs, $6.5\,\sigma$ exp, these results supersede PRL,111,101804)

  - ▷ first $\tau_{\mu^+\mu^-}$ measurement of CMS

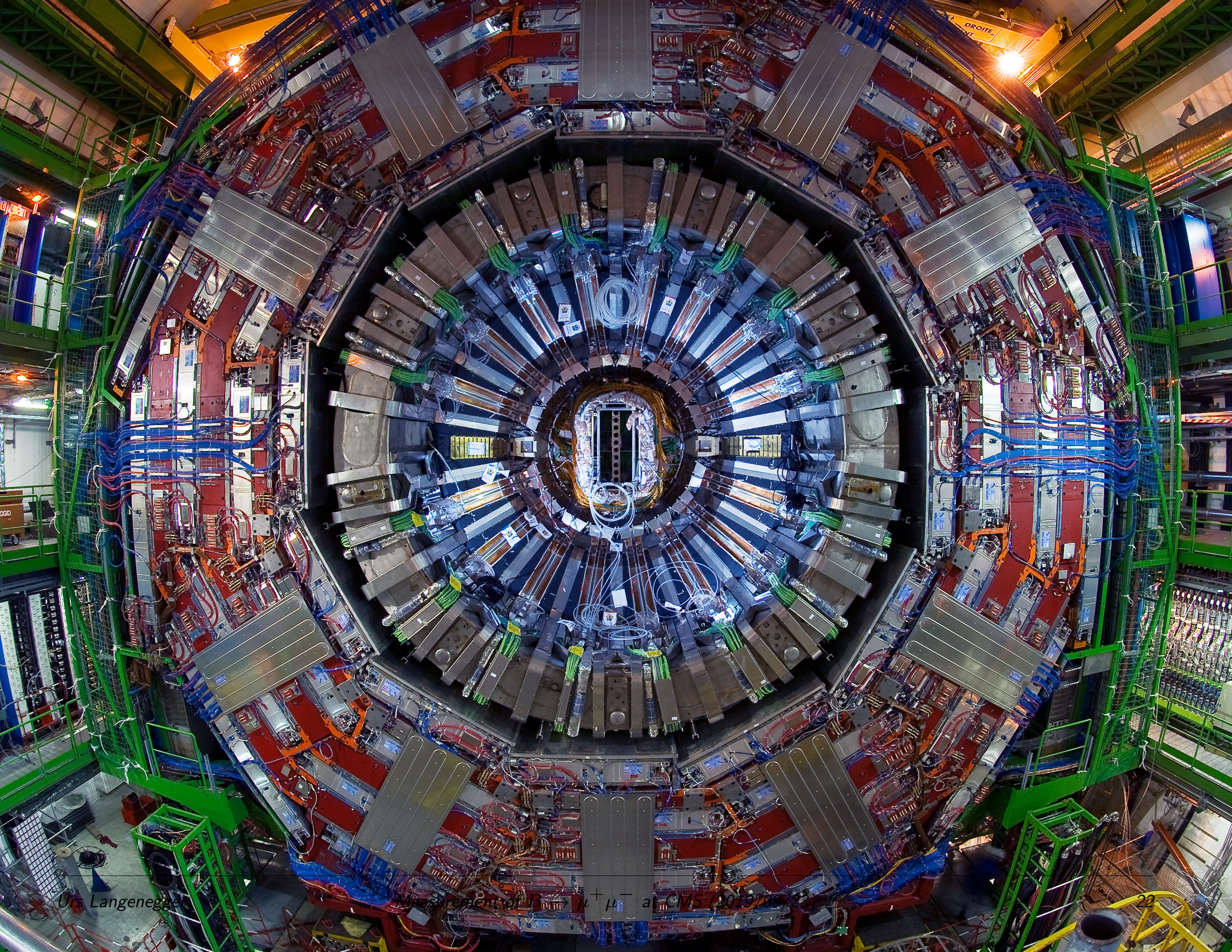
$$\tau_{\mu^+\mu^-} = 1.70^{+0.61}_{-0.44}\,\text{ps}$$

  - ▷ all results consistent with SM
    do not over-interpret 'low' result(s)!
  - ▷ (very) long delay due to MC issues
    'irrelevant' in statistics limited result





- End of PSI involvement in $B \to \mu^+\mu^-$
  - ▷ focus on other decays
    leptonic (forbidden) B decays with hadronic τ reconstruction
    maybe eventually $B_s^0 \to \tau^+\tau^-$??



Summary of systematic errors

- Uncertainties dominated by small signal sample size
  - ▷ relative errors for $\overline{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)$, absolute for $\tau_{\mu^+\mu^-}$

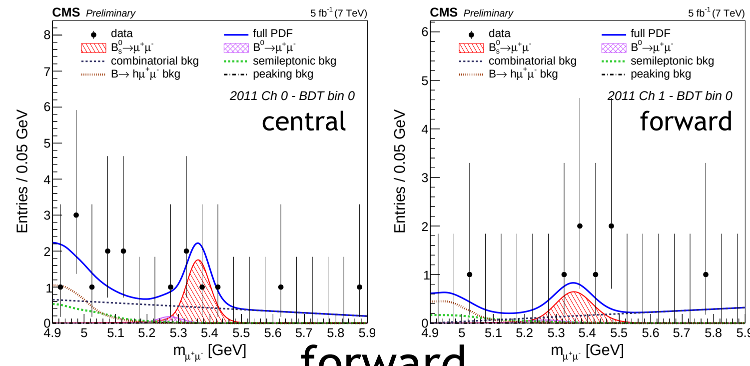
Source	$\overline{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)$ [%]	$\tau_{\mu^+\mu^-}$ [ps] 2D UML	*sPlot*
Kaon tracking	2.3 – 4	–	–
Normalization yield	4	–	–
Background yields	1	0.03	(*)
Production process	3	–	–
Muon identification	3	–	–
Trigger	3	–	–
Efficiency (data/MC simulation)	5 – 10	–	(*)
Efficiency (functional form)	–	0.01	0.04
Efficiency lifetime dependence	1 – 3	(*)	(*)
Era dependence	5 – 6	0.07	0.07
BDT discriminator threshold	–	0.02	0.02
Silicon tracker alignment	–	0.02	–
Finite size of MC sample	–	0.03	–
Fit bias	–	–	0.09
$\mathcal{C}$-correction	–	0.01	0.01
Total systematic uncertainty	$(^{+0.3}_{-0.2}) \times 10^{-9}$	0.09	0.12
Total uncertainty	$(^{+0.7}_{-0.6}) \times 10^{-9}$	$^{+0.61}_{-0.44}$	$^{+0.52}_{-0.33}$

(*) included in other item

  - ▷ successful cross check of σ_{syst} with measurement of $\mathcal{B}(B_s^0 \to J/\psi\,\phi)$

Fit details (plots)

2011



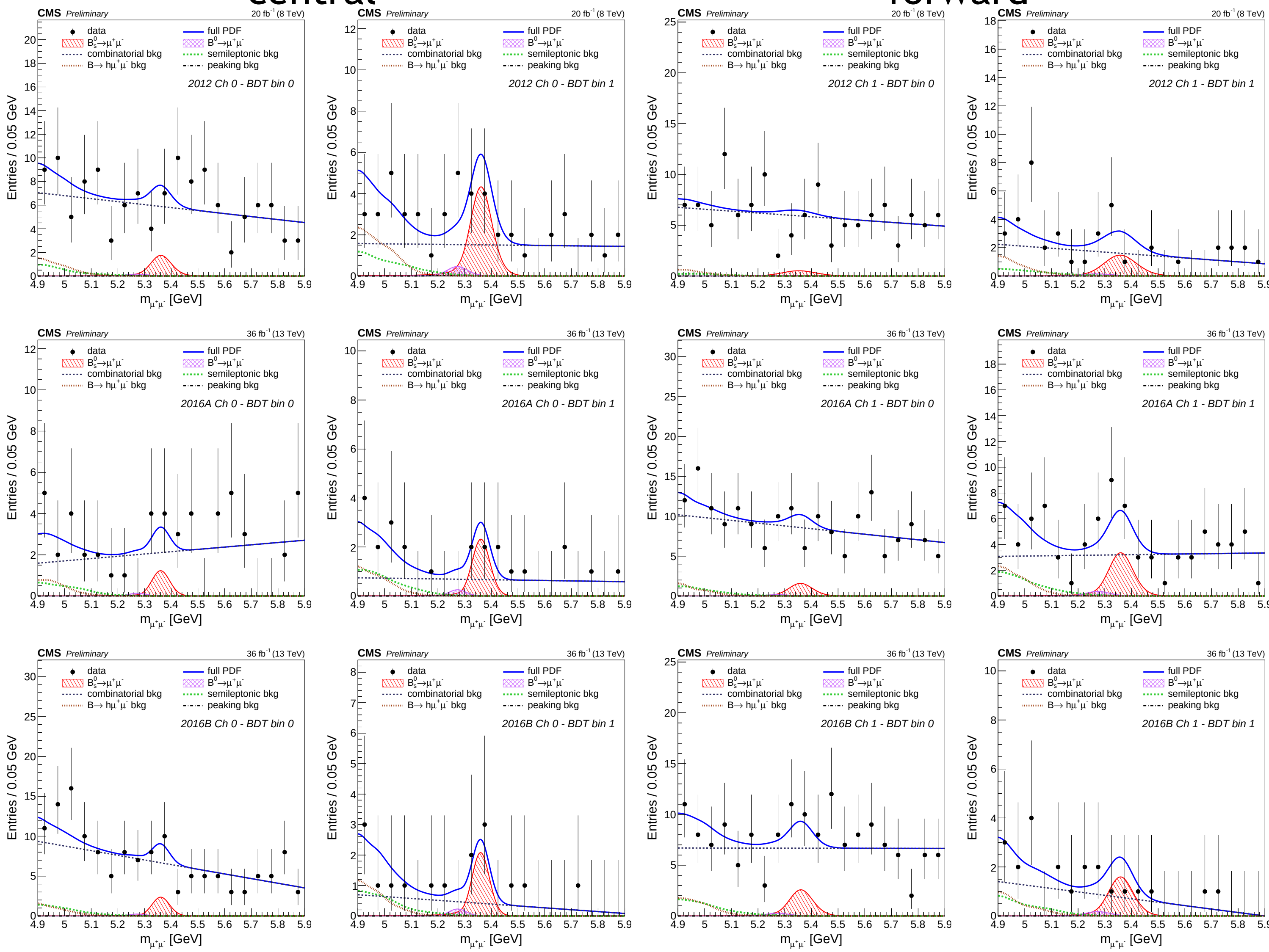
central

forward

2012

2016A

2016B



Fit details (numbers)

- Obs signal yield $60.8^{+14.5}_{-13.3}$ with $\langle p_\perp \rangle = 17.2\,\text{GeV}$
  - ▷ peaking background is $\approx 5-10\%$ of $B^0 \to \mu^+\mu^-$ yield
  - ▷ uncertainties include statistical and systematic errors
  - ▷ signal yields (and errors) determined from $\mathcal{B}$ (and include normalization errors)

Category	$N(B^0_s)$	$N(B^0)$	N_{comb}	$N^{B^+}_{\text{obs}}/100$	$\langle p_\perp(B^0_s)\rangle\,[\text{GeV}]$	$\varepsilon_{\text{tot}}/\varepsilon^{B^+}_{\text{tot}}$
2011/central/high	$3.6^{+0.9}_{-0.8}$	$0.4^{+0.7}_{-0.6}$	8.4 ± 3.8	750 ± 30	16.4	3.9 ± 0.5
2011/forward/high	$2.0^{+0.5}_{-0.4}$	$0.2^{+0.4}_{-0.3}$	3.2 ± 2.2	220 ± 12	14.9	7.5 ± 0.8
2012/central/low	$3.7^{+0.9}_{-0.8}$	$0.4^{+0.6}_{-0.6}$	115.8 ± 11.3	790 ± 32	16.1	3.8 ± 0.5
2012/central/high	$9.3^{+2.3}_{-2.1}$	$1.0^{+1.7}_{-1.6}$	30.2 ± 7.3	2360 ± 95	17.3	3.2 ± 0.4
2012/forward/low	$1.7^{+0.4}_{-0.4}$	$0.2^{+0.3}_{-0.3}$	116.7 ± 11.0	190 ± 9	14.3	7.3 ± 1.0
2012/forward/high	$4.7^{+1.2}_{-1.1}$	$0.5^{+0.9}_{-0.8}$	31.0 ± 6.5	660 ± 27	15.5	5.9 ± 0.8
2016BF/central/low	$2.2^{+0.5}_{-0.5}$	$0.2^{+0.4}_{-0.4}$	43.0 ± 7.1	580 ± 23	17.5	3.1 ± 0.4
2016BF/central/high	$4.0^{+1.0}_{-0.9}$	$0.4^{+0.8}_{-0.7}$	13.3 ± 4.7	1290 ± 57	19.3	2.5 ± 0.3
2016BF/forward/low	$3.7^{+0.9}_{-0.8}$	$0.4^{+0.7}_{-0.7}$	168.8 ± 13.5	780 ± 31	15.8	3.9 ± 0.5
2016BF/forward/high	$8.1^{+2.0}_{-1.8}$	$0.8^{+1.5}_{-1.4}$	64.2 ± 9.7	1920 ± 78	17.5	3.4 ± 0.4
2016GH/central/low	$4.1^{+1.0}_{-0.9}$	$0.4^{+0.8}_{-0.7}$	128.8 ± 12.0	1020 ± 44	17.2	3.3 ± 0.4
2016GH/central/high	$3.6^{+0.9}_{-0.8}$	$0.4^{+0.7}_{-0.6}$	7.8 ± 3.6	1320 ± 54	20.8	2.2 ± 0.2
2016GH/forward/low	$6.1^{+1.5}_{-1.4}$	$0.6^{+1.1}_{-1.0}$	133.4 ± 12.5	1260 ± 51	16.2	3.9 ± 0.4
2016GH/forward/high	$3.9^{+1.0}_{-0.9}$	$0.4^{+0.8}_{-0.7}$	14.1 ± 4.6	1180 ± 49	19.5	2.7 ± 0.3

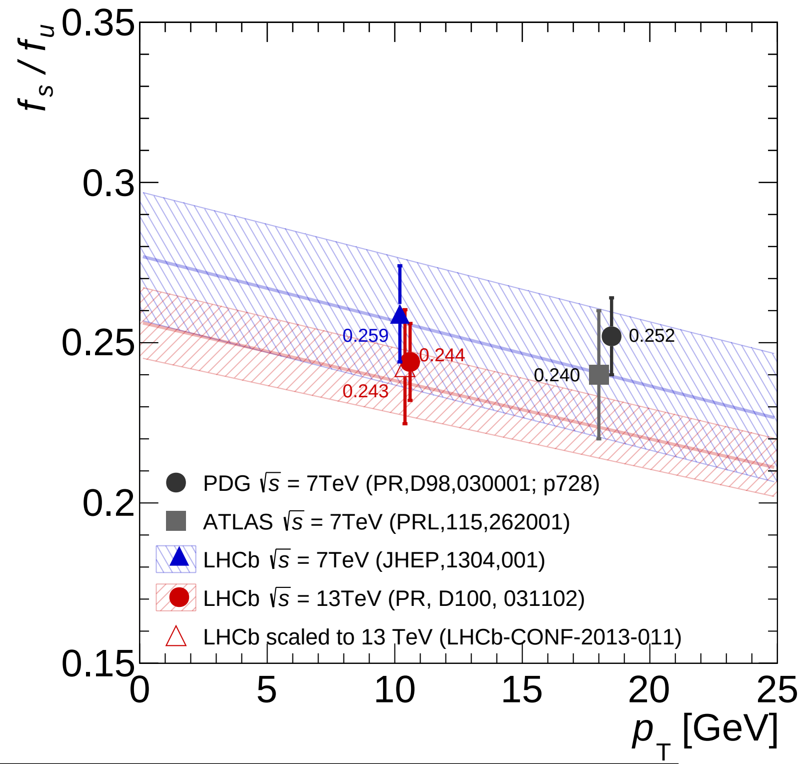
A note on f_s/f_u

PRL,115,262001
PR,D100,031102

- f_s/f_u is external input for $\overline{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)$
  - ▷ experimental situation not entirely clear
    - LHCb sees $p_\perp$-dependence (PR,D100,031102)
    - ATLAS does not see $p_\perp$-dependence (PRL,115,262001)
    - CMS does not see $p_\perp$-dependence (internal study with control sample)
  - ▷ fragmentation fraction x_B not measured
    (x_B: fraction of b momentum $\to B$)

- Ad hoc error added
  - ▷ PDG $f_s/f_u = 0.252 \pm 0.012$, based on
    $\sqrt{s} = 7$ TeV results of LHCb/ATLAS
  - ▷ additional ad-hoc error
    difference between PR,D100,031102 and PDG
    $p_\perp$ dependence from PR,D100,031102

$$f_s/f_u = 0.252 \pm 0.012(\text{exp.}) \pm 0.015(\text{CMS})$$



⇒ Our result can be rescaled
- ▷ $\sqrt{s}$ and $p_\perp$ of signal candidates provided for each category/channel/running period

Lifetime fitting

- Determination of proper decay time $t = m\,\ell_{3D}/p$ in 3D space
- 2D unbinned extended maximum likelihood fit to
  - B mass and t decay time in the range $1 < t < 11$ ps
    (σ_t as conditional parameter, complete propagation of uncertainties)
  - Efficiency correction (mostly HLT)
  - model components

mass	shape	source	fit params
Signal	CB	MC	fixed
BG $h\mu\nu$, $h\mu\mu$	G	w8-MC	fixed
BG hh, $B^0 \to \mu^+\mu^-$	CB+G	w8-MC	fixed
Combinatorial BG	Bernstein pol1	sideband	floating
decay time	**shape**	**source**	**fit params**
Signal	expo⊗res$^{(*)}$	MC	floating
BG $h\mu\nu$, $h\mu\mu$	expo⊗res	w8-MC	fixed
BG hh, $B^0 \to \mu^+\mu^-$	expo⊗res	w8-MC	fixed
Combinatorial BG	expo⊗res	sideband	floating

$(*)$ 'res' includes resolution and efficiency (no efficiency correction for the combinatorial background)

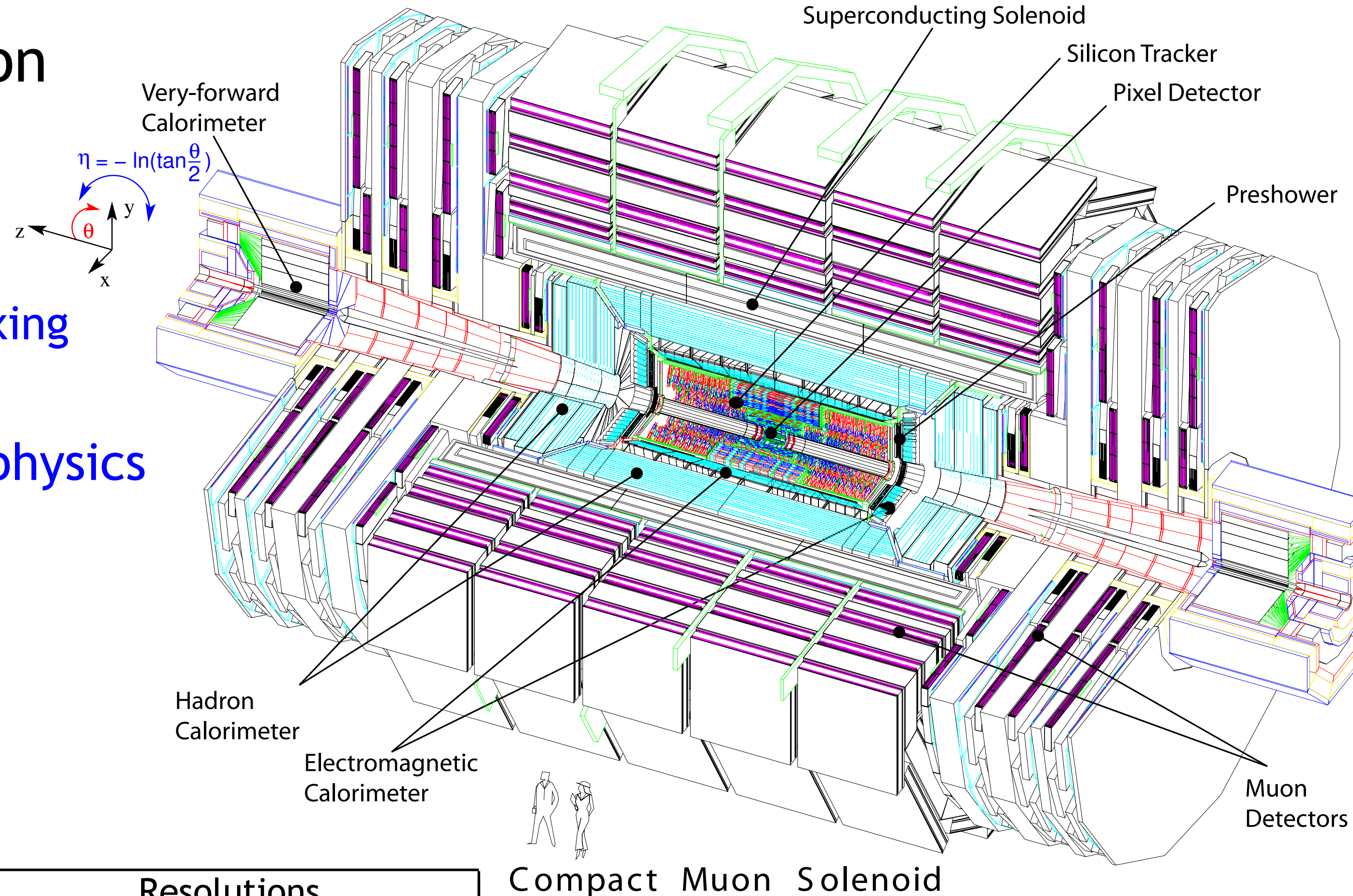
- $sPlot$ lifetime fit
  - $sPlot$ weights from $\overline{\mathcal{B}}(B_s^0 \to \mu^+\mu^-)$ model
  - binned maximum likelihood fit with resolution and efficiency modeling
  - custom algorithm for correct (asymmetric) uncertainties

$\Rightarrow$ Consistent results between the two setups

The CMS detector

- Design prioritization
  - ▷ lepton ID
    - → muons
  - ▷ b/τ tagging
    - → tracking/vertexing
  - ▷ jets and $\not{E}_T$
- → well suited for B physics

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



Component	Characteristics	Resolutions
Pixel	3/2 (4/3) Si layers	$\delta_z \approx 20\ \mu$m, $\delta_\phi \approx 10\ \mu$m
Tracker	10/12 Si strips	$\delta(p_\perp)/p_\perp \approx 1\%$
ECAL	$PbWO_4$	$\delta E/E \approx 3\%/\sqrt{E} \oplus 0.5\%$
HCAL (B)	Brass/Sc, $> 7.2\lambda$	$\delta E/E \approx 100\sqrt{E}\%$
HCAL (F)	Fe/Quartz	$\delta(\not{E}_T) \approx 0.98\sqrt{\sum E_T}$
Magnet	3.8 T solenoid	
Muons	DT/CSC + RPC	$\delta(p_\perp)/p_\perp \approx 10\%$ (STA)

Tracking resolution:

impact parameter $\approx 15\ \mu$m

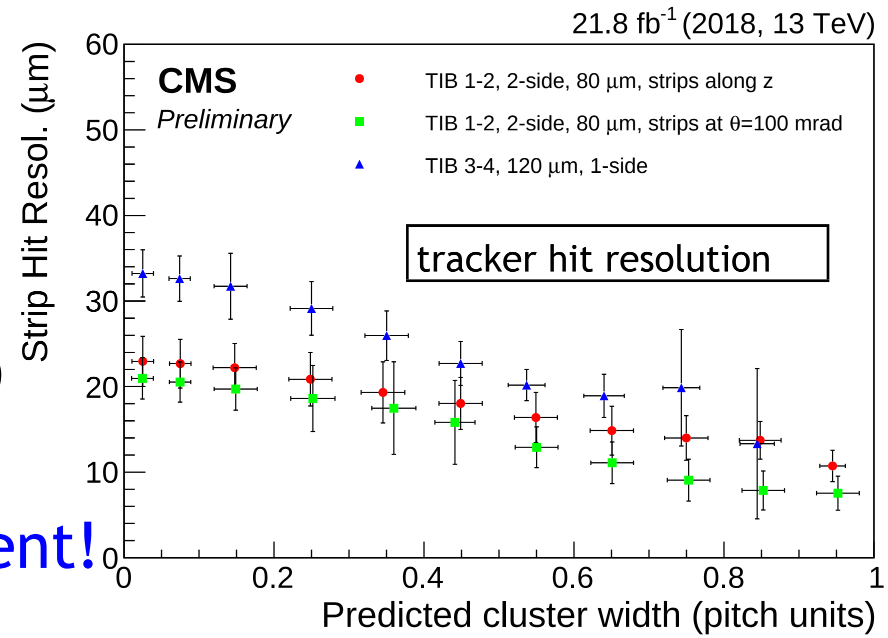
Primary vertex resolution:

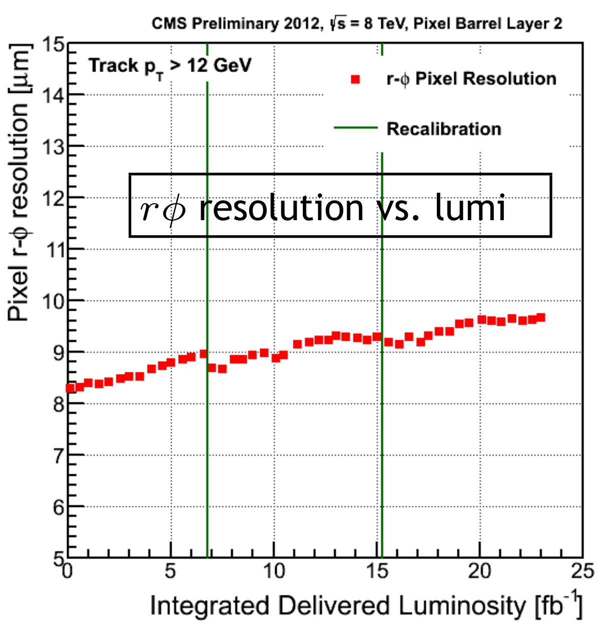
$\Delta z \approx 20 - 80\ \mu$m

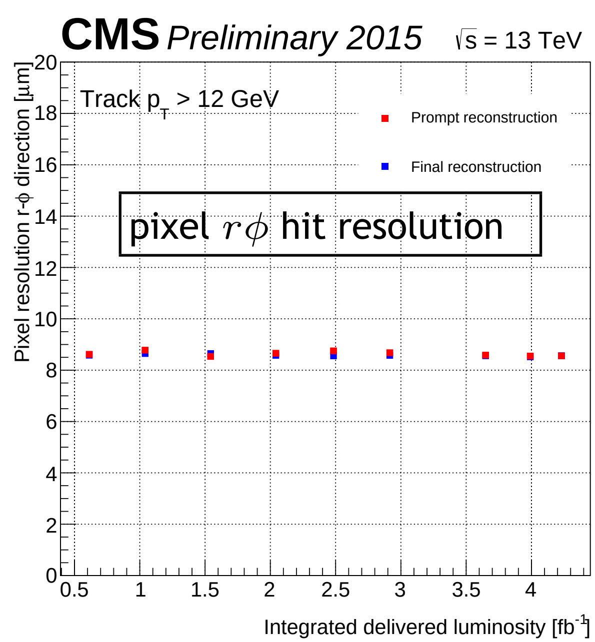
3D tracking and vertexing

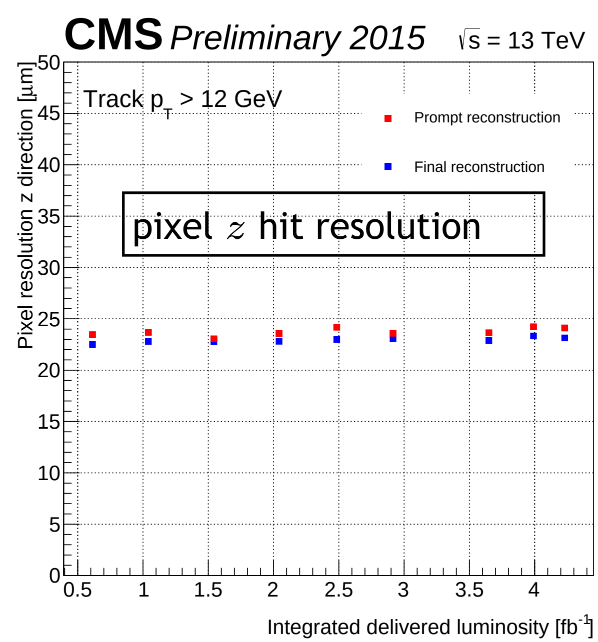
- All silicon tracker
  - ▷ high granularity, low occupancy
- Pixel detector
  - ▷ $100 \times 150\,\mu$m^2 pixel size
  - ▷ substantial charge sharing (low $V_{\rm bias}$)
  - $\rightarrow$ excellent resolution in $r\phi$ and z

$\Rightarrow$ Essential in high-pileup environment!





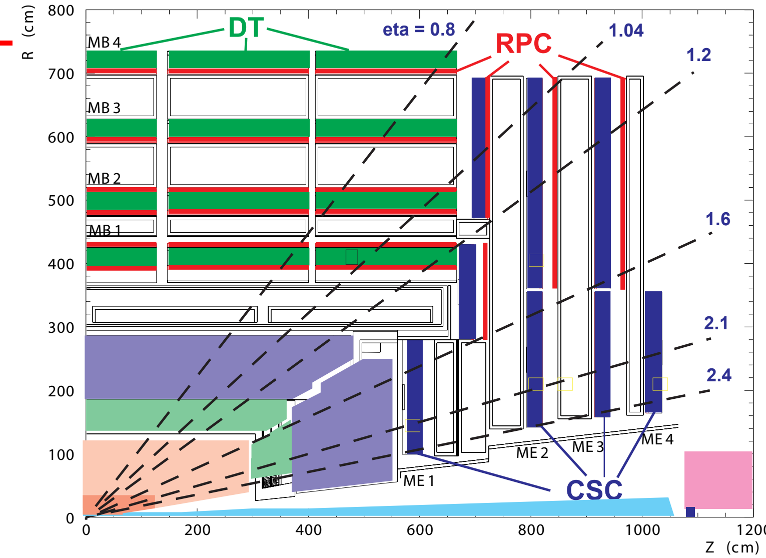




Muon reconstruction

JINST,13,P06015
JINST,7,P10002

- Large muon acceptance $|\eta| < 2.4$
  - ▷ drift tubes
  - ▷ cathode strip chambers
  - ▷ resistive plate chambers
- Muon reconstruction
  - ▷ standalone muon: in muon system (trigger ingredient)
  - ▷ 'soft': high efficiency for J/ψ analyses
  - ▷ 'BDT': low misidentification for $B \to \mu^+\mu^-$ analyses



Muon misidentification for BDT muons

$$\varepsilon(\mu|\pi) \approx 0.06\%$$
$$\varepsilon(\mu|K) \approx 0.10\%$$
$$\varepsilon(\mu|p) \leq 0.01\%$$

measured/validated in data:
$K^0_S \to \pi^+\pi^-$, $\phi \to K^-K^+$, $\Lambda \to p\pi^-$
$D^{*+} \to D^0\pi^+_s \to K^-\pi^+\pi^+_s$