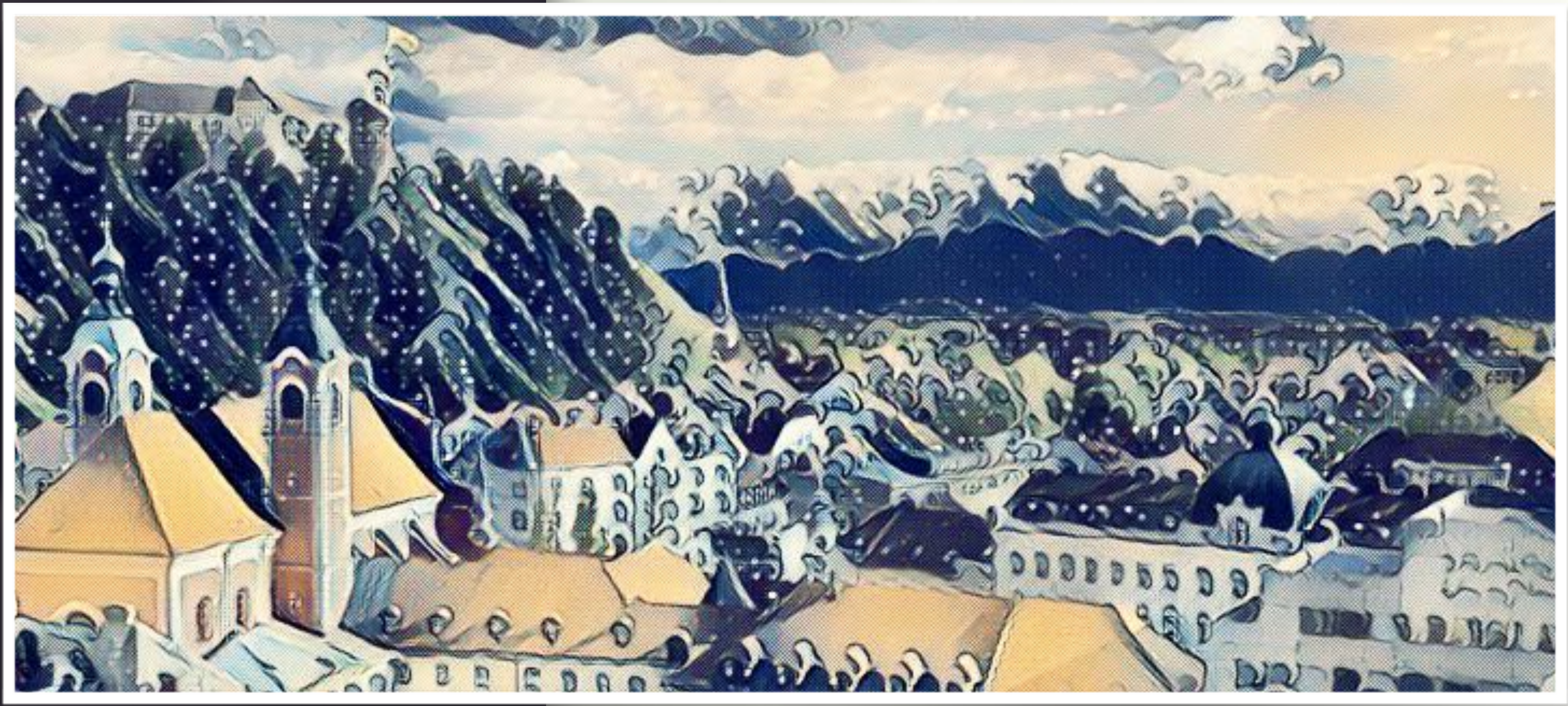
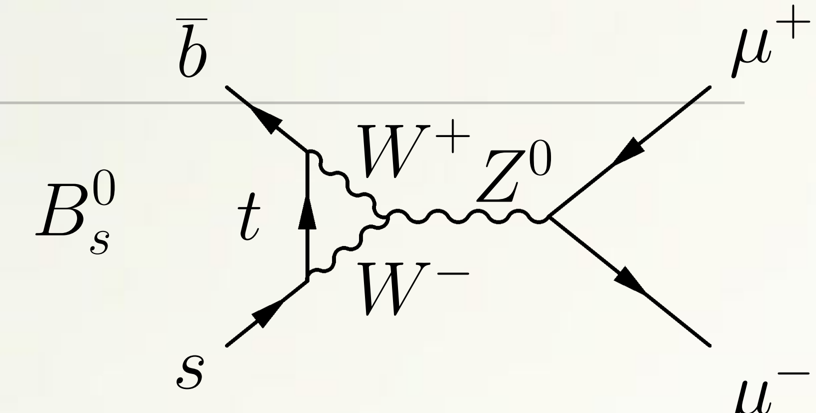


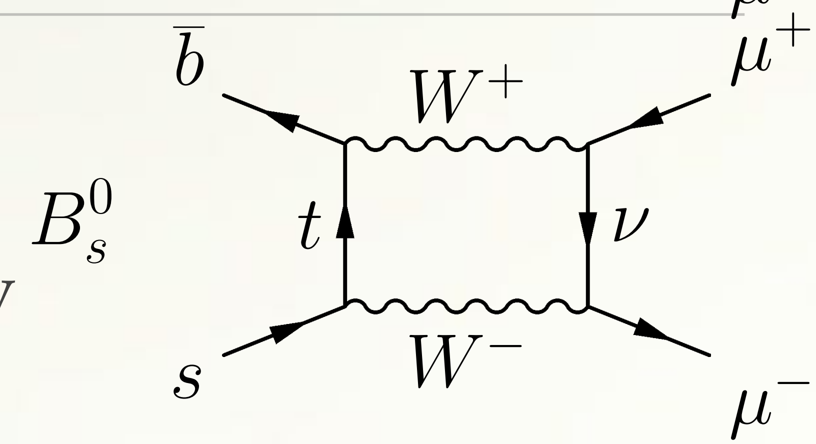
Beauty 2019, October 1st, 2019, Ljubljana, Slovenia

Measurement of $B_{s,d}\rightarrow\mu^+\mu^-$ & Search for $\tau\rightarrow 3\mu$ at CMS

Kai-Feng Chen
National Taiwan University
On behalf of the CMS Collaboration

Physics Motivation





- $B_{s,d} \to \mu^+\mu^-$ decays only proceed through FCNC processes and are highly suppressed in SM.
- Loop diagram + Suppressed SM + Theoretically clean = **an excellent place to look for NP**.
- What to measure:
  - **Branching fractions**: $B_s \to \mu\mu$ may start to enter precision regime, while first evidence of $B_d \to \mu\mu$ might emerge.
  - **Effective lifetime**: only the heavy B_s state can decay into dimuon in the SM; different composition of states may be allowed by NP.

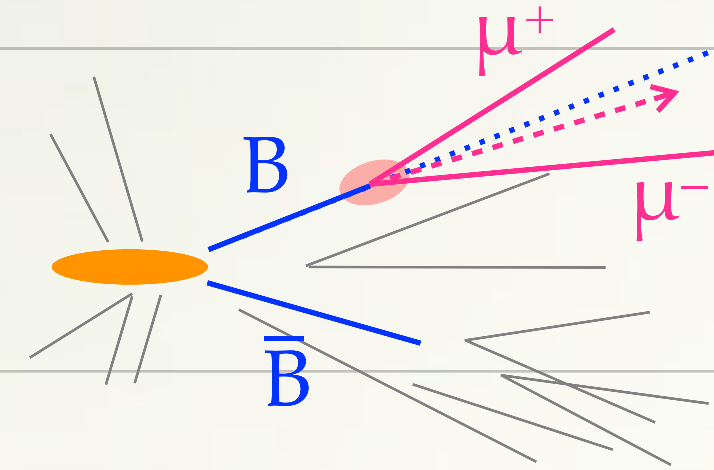
$$\tau_{\mu^+\mu^-} \equiv \frac{\int_0^\infty t\,\Gamma(B_s(t) \to \mu^+\mu^-)\,dt}{\int_0^\infty \Gamma(B_s(t) \to \mu^+\mu^-)\,dt} = \frac{\tau_{B_s^0}}{1-y_s^2}\left(\frac{1+2\mathcal{A}_{\Delta\Gamma}^{\mu^+\mu^-}y_s + y_s^2}{1+\mathcal{A}_{\Delta\Gamma}^{\mu^+\mu^-}y_s}\right) \quad \mathcal{A}_{\Delta\Gamma}^{\mu^+\mu^-}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

- SM predictions:

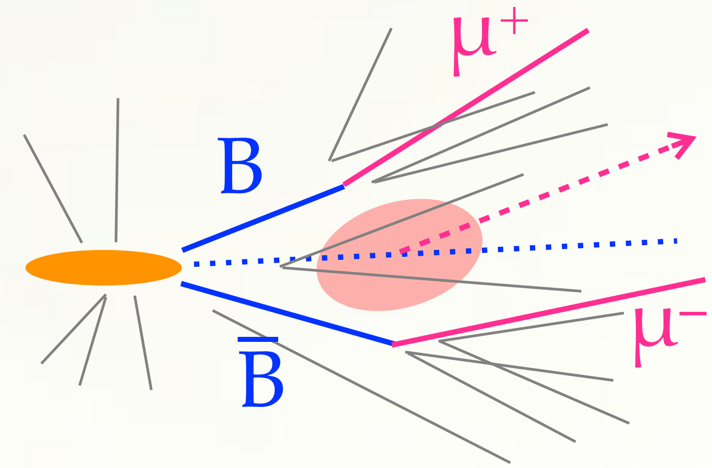
Ref:
Beneke et al, PRL 120, 011801 (2018)
Bobeth et al, PRL 112, 101801 (2014)

$\mathcal{B}(B_s \to \mu^+\mu^-) = (3.57 \pm 0.17) \times 10^{-9}$
$\mathcal{B}(B^0 \to \mu^+\mu^-) = (1.06 \pm 0.09) \times 10^{-10}$
$\tau(B_s \to \mu^+\mu^-) = 1.615$ ps

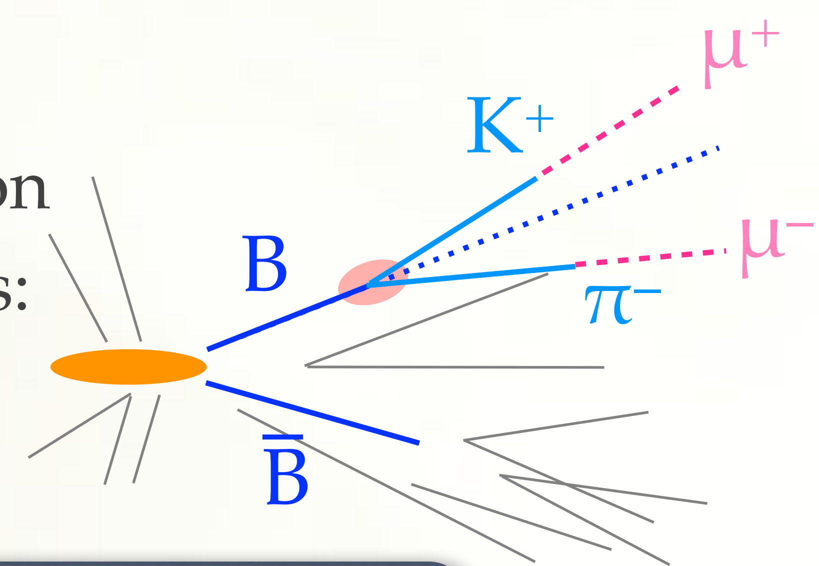
Analysis Aspects



- $B_{s,d} \to \mu^+\mu^-$ **signal signature:**
  - two muons from one displaced vertex; isolated from other activities; momentum aligned with its flight direction; invariant mass peaking at $M(B_{s,d})$.



- **Background sources:**
  - **Combinatorial background** consists of
    - two semileptonic B decays
    - one semileptonic B + a misidentified hadron
  - **Rare background** from single B meson decays:
    e.g. $B \to K\pi/KK/\pi\pi$ *(peaking)*,
    $B \to h^-\mu^+\upsilon$, $B \to h\mu^+\mu^-$ *(not peaking)*



Powerful background suppression reached by muon quality, well-reconstructed secondary vertex, isolation, pointing angle, and M(μμ) resolution.

Analysis Aspects (cont.)

- Background suppression achieved by
  - Strict **Boosted Decision Tree (BDT) muon identification** requirement, including tracking and muon related detector information;
  - **Event classification BDT**, which includes several topological and kinematical variables.
- Calibrations / validations with $\mathbf{B^+ \to J/\psi K^+}$ and $\mathbf{B_s \to J/\psi \phi}$.
- Normalized to the reference channel $\mathbf{B^+ \to J/\psi(\to \mu^+\mu^-)\, K^+}$:

$$\mathcal{B}(B_s \to \mu^+\mu^-) = \frac{N_S}{N(B^\pm \to J/\psi K^\pm)} \times \mathcal{B}(B^\pm \to J/\psi K^\pm) \times \frac{A(B^\pm)}{A(B_s)} \frac{\varepsilon^{ana}(B^\pm)}{\varepsilon^{ana}(B_s)} \frac{\varepsilon^{\mu}(B^\pm)}{\varepsilon^{\mu}(B_s)} \frac{\varepsilon^{trig}(B^\pm)}{\varepsilon^{trig}(B_s)} \frac{f_u}{f_s}$$

Acceptance → A

Selection efficiency → ε^{ana}

muon identification → ε^{μ}

Trigger efficiency → ε^{trig}

B-hadronization composition, for B_s only → f_u/f_s

Similar **trigger & selection** for reducing systematic uncertainties

Dataset & Trigger

- **Combining Run-1 data + 2016, 4 run periods in total:**
  - **5 fb^{-1} at 7 TeV from 2011**
  - **20 fb^{-1} at 8 TeV from 2012**

  } re-blind + re-analyzed

  - **36 fb^{-1} at 13 TeV from 2016**, *splits into 2 periods,* ***2016A*** *and* ***2016B****, based on detector conditions.*
- **Trigger:**
  - Identical L1 triggers for signal and normalization; very similar HLT criteria, except displacement requirement for normalization.
  - 2016 setup was very similar to 2012, except a tighter restriction on $|\eta_f|<1.4$ due to the trigger rate limitation in 2016.

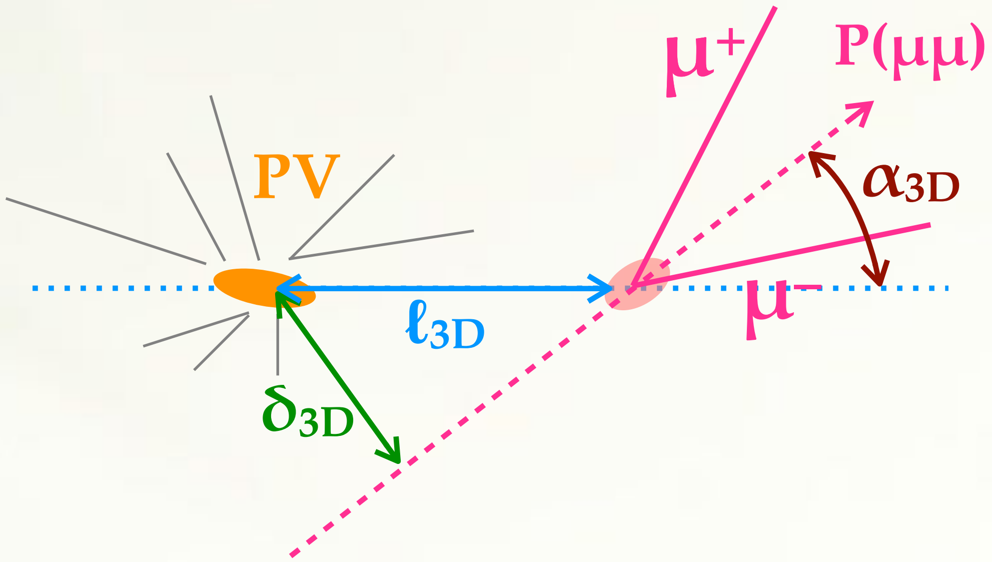
Year	Central	Forward
2011	$0 < \lvert\eta_f\rvert < 1.4$	$1.4 < \lvert\eta_f\rvert < 2.1$
2012	$0 < \lvert\eta_f\rvert < 1.4$	$1.4 < \lvert\eta_f\rvert < 2.1$
2016	$0 < \lvert\eta_f\rvert < 0.7$	$0.7 < \lvert\eta_f\rvert < 1.4$

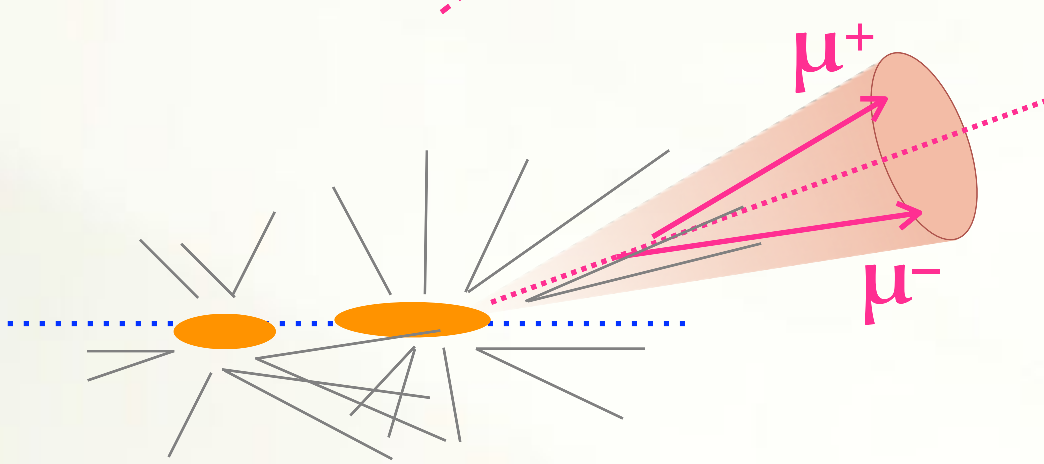
η_f : pseudo rapidity of the most forward muon

$\Rightarrow$ ie. No "endcap" in 2016

Background Suppression: Variables

- **Associated with B candidate:**
  - B and daughter μ kinematics: $\mathbf{p_T}$ and $\boldsymbol{\eta}$.
  - Minimal distance between two trajectories: $\mathbf{d_{ca}}$.
- **Associated with vertex:**
  - Impact parameters: δ_{3D}, $\delta_{3D}/\sigma(\delta_{3D})$
  - Flight length: ℓ_{3D}, $\ell_{3D}/\sigma(\ell_{3D})$
  - Pointing angle: α_{3D}
  - Quality of the fit: χ^2/dof
- **Associated with isolation:**
  - B candidate isolation: $\mathbf{I} = p_T(B)\ /\ [p_T(B) + \Sigma p_T(\text{tracks})]$ within $\Delta R<0.7$
  - muon isolation: $\mathbf{I_\mu} = p_T(\mu)\ /\ [p_T(\mu) + \Sigma p_T(\text{tracks})]$ within $\Delta R<0.5$
  - Track counting: N_{trk}^{close}
  - Minimal distance between nearby tracks to B vertex: d^0_{ca}





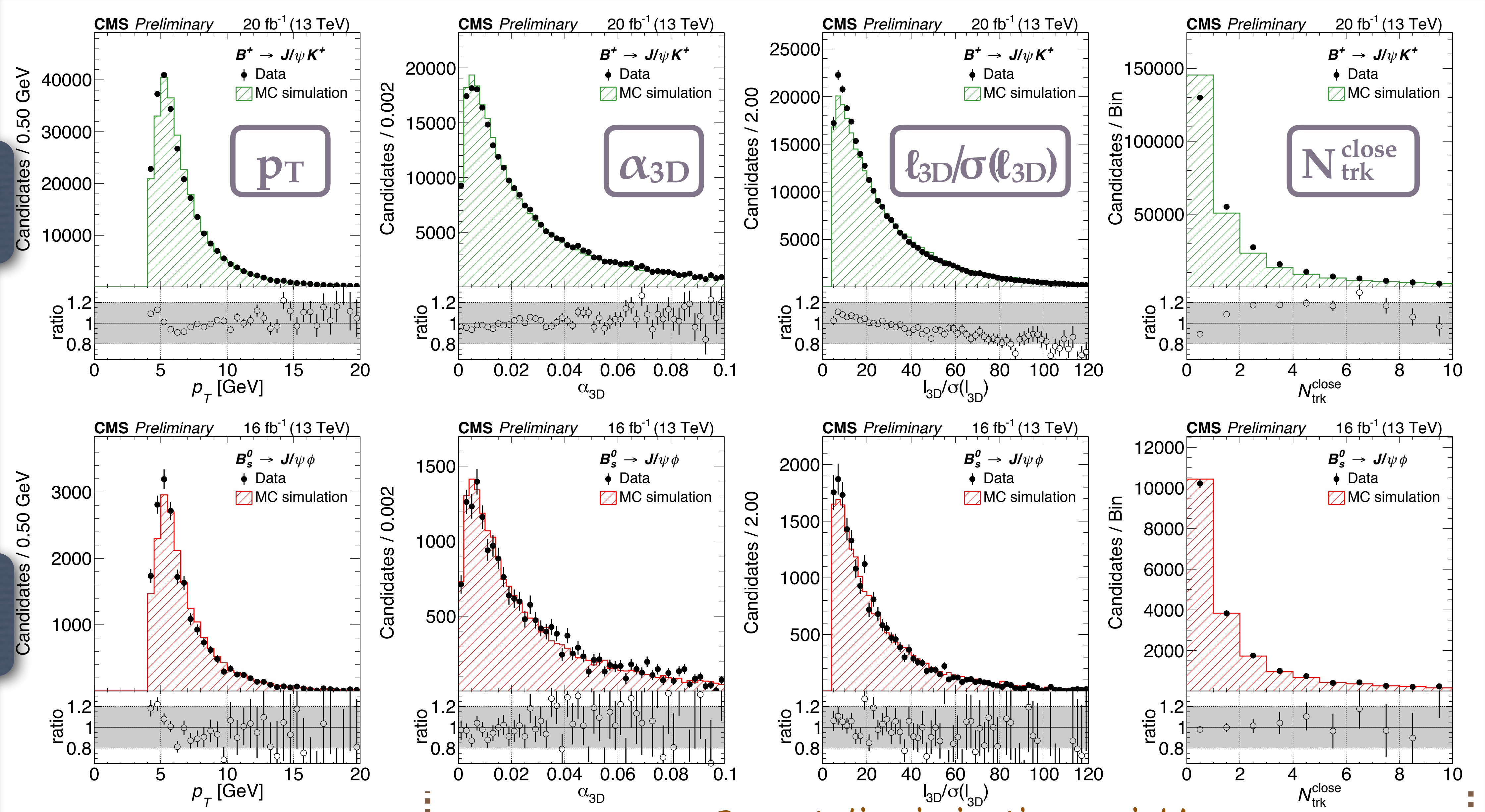
Simulation versus Data

- **Comparison between MC and data events:**
  - Background subtracted distributions from **normalization** & **control** channels.
  - Difference is accounted as a systematic uncertainty.

Ref. CMS-PAS-BPH-16-004

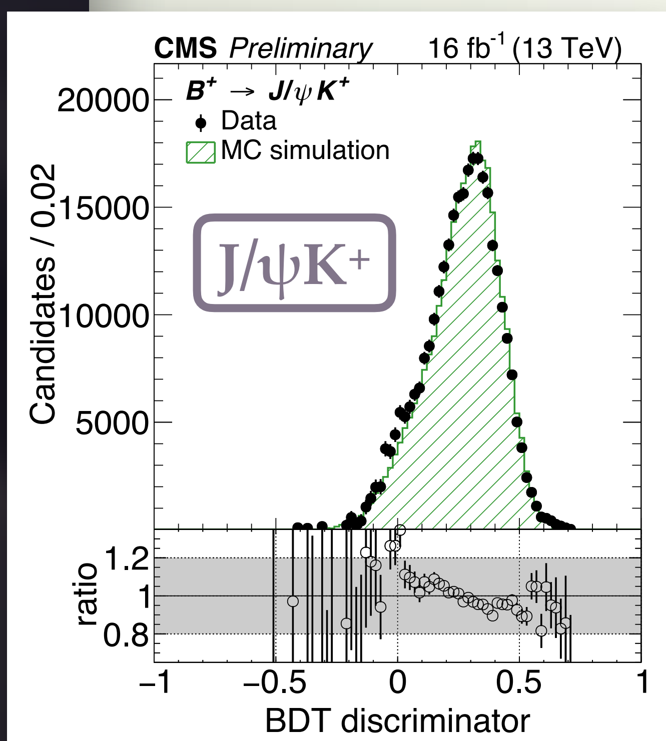
Normalization ($J/\psi K^+$)

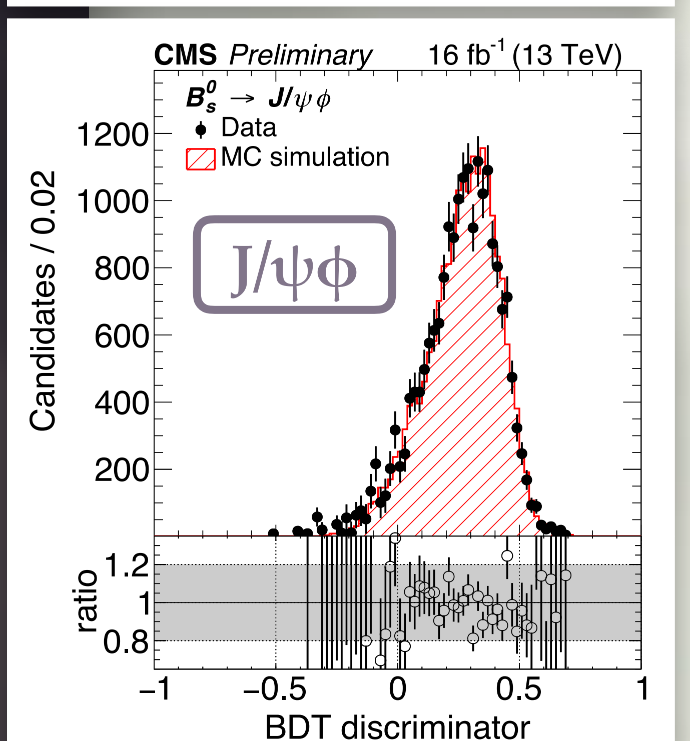
Control ($J/\psi\phi$)



3 most discriminating variables

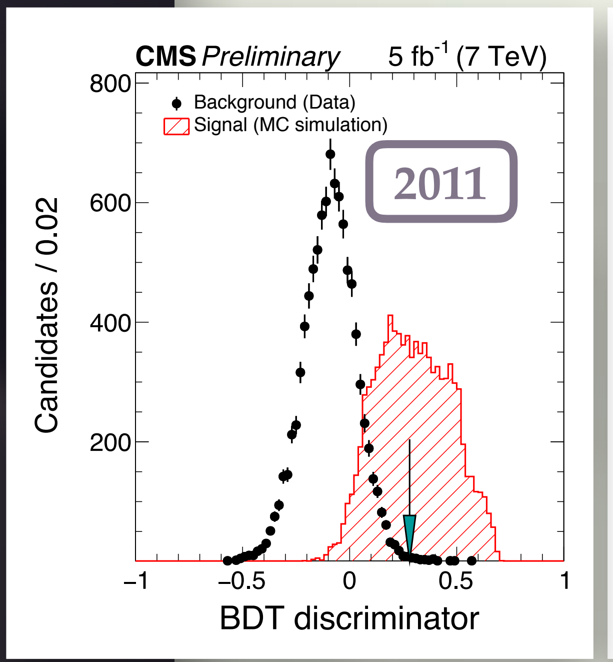
Background Suppression: MVA

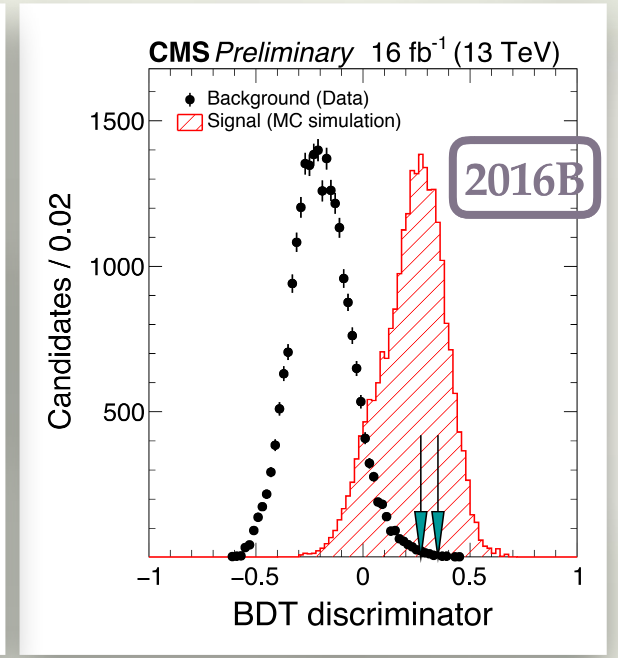




- **BDT Configurations:**
  - 2011 and 2012 samples: the same setup introduced for previous publication.
  - 2016 samples: newly trained BDT with the same input variables.
- **Training with TMVA:**
  - Signal source: **simulated $B_s \to \mu^+\mu^-$ events**.
  - Background source: **data sideband**.
  - Split the samples into 3 subsets: **train / test / application**;
    - ➡ *Resulting 3 BDTs for each channel.*
- **Systematic uncertainties:**
  - Many validation studies have been performed.
  - Measure the double ratio of MC/data and of control/normalization channels.
  - Resulting 5-10% uncertainty on the relative efficiencies;
  - Around ±0.07 ps on the effective lifetime.

BDT Categorizing





- **BDT binning optimized with:**
  - Maximum Likelihood estimator with asymptotic method (*w/ Asimov data*).
  - Run-1 channels re-optimized.
  - Look for the **best significance** for branching fraction measurement and **smallest uncertainty** for lifetime analysis.

Ref. CMS-PAS-BPH-16-004

Branching fractions: 14 categories

	2011		2012		2016A		2016B	
Category	Central	Forward	Central	Forward	Central	Forward	Central	Forward
Low	–	–	0.27	0.23	0.19	0.19	0.18	0.23
High	0.28	0.21	0.35	0.32	0.30	0.30	0.31	0.38

Effective lifetime: 8 categories

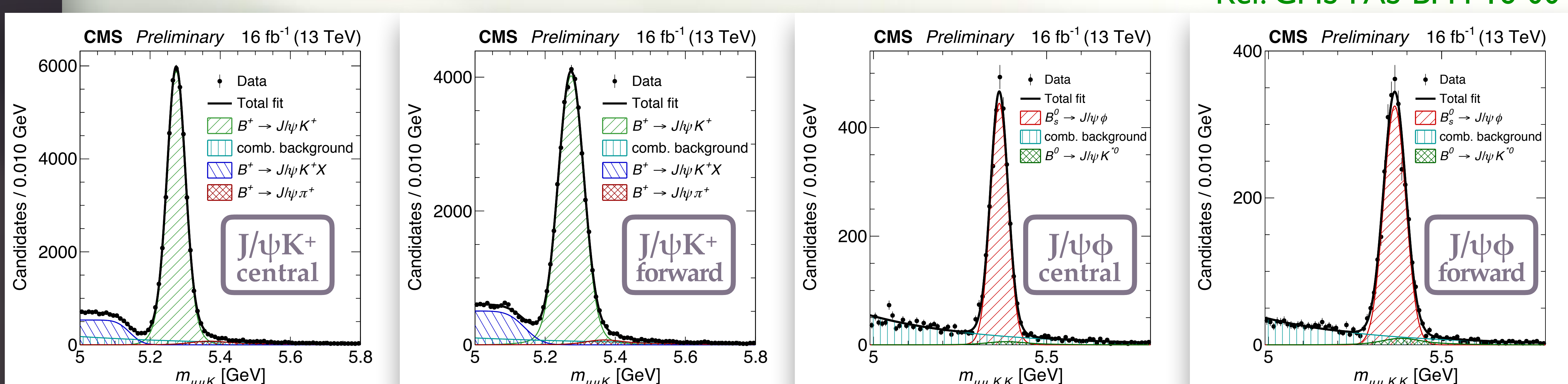
	2011		2012		2016A		2016B	
Category	Central	Forward	Central	Forward	Central	Forward	Central	Forward
BDT	0.22	0.19	0.32	0.32	0.22	0.30	0.22	0.29

Yield Extraction: J/ψK⁺ & J/ψϕ

- **Fits to invariant mass distributions**, with double-Gaussian as the signal model.
- **J/ψK⁺ background model:**
  - Combinatorial: exponential;
  - Partial reconstructed J/ψ+X: error function;
  - $B^+ \to J/\psi\pi^+$: triple-Gaussian from MC, fixed to 4% of $J/\psi K^+$ signal.
- **J/ψϕ background model:**
  - Combinatorial: exponential;
  - $B^0 \to J/\psi K^{*0}$: triple-Gaussian+linear from MC.

Systematic uncertainty of 4% on yield, mainly from the difference to fit w & w/o J/ψ mass constraint.

Ref. CMS-PAS-BPH-16-004



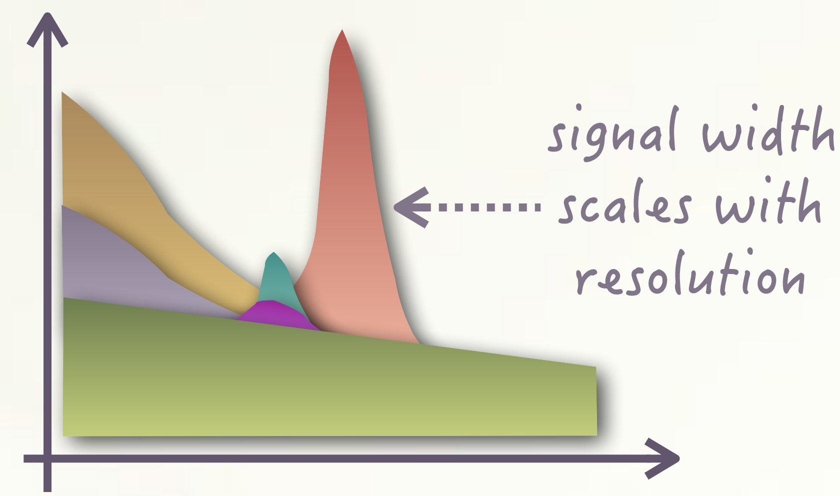
Branching Fraction Extraction

- The model: a 3D unbined ML fit

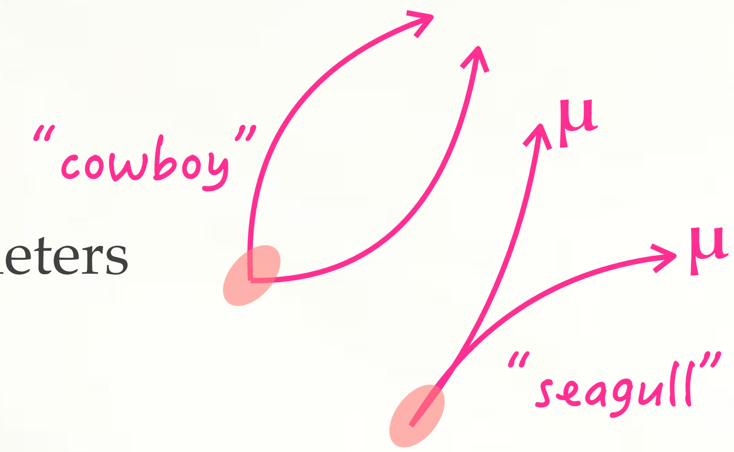
$$P(m_{\mu\mu};\sigma(m_{\mu\mu})) \times P(\sigma(m_{\mu\mu})/m_{\mu\mu}) \times P(\mathcal{C})$$

  - Invariant mass: $\mathbf{m_{\mu\mu}}$
  - Per-event mass resolution: $\boldsymbol{\sigma(m_{\mu\mu})}$
  - Dimuon bending configuration: C bending *towards* or *away* from each other

(⟹ against possible bias)



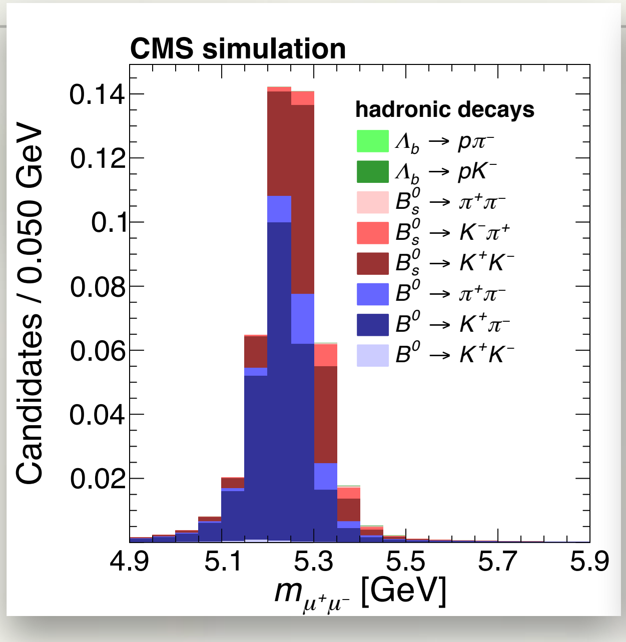
- Constrained nuisances:
  - With Gaussian: f_s/f_u, $\mathcal{B}(B^+ \to J/\psi K^+ \to \mu\mu K^+)$, ratio of efficiencies $\varepsilon(Bs)/\varepsilon(B^+)$, shape systematic parameters
  - With Lognormal: rare background yields

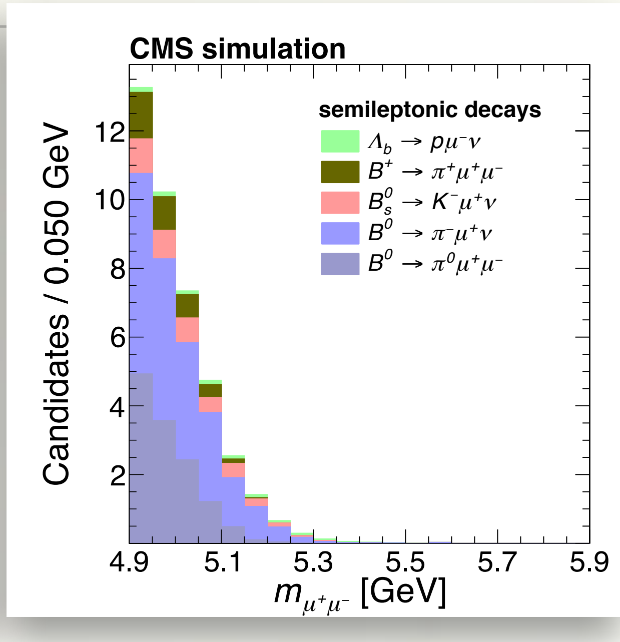


- PDF choice for **6 components**:

Component	$m_{\mu\mu}$	$\sigma(m_{\mu\mu})$
B_s, B_d signal	Crystal-ball, $\sigma_{CB} = \kappa \times \sigma(M_{\mu\mu})$	Keys PDF
$B \to hh$ background	Crystal-ball+Gaussian	Keys PDF
$B \to h\mu\mu$, $h\mu\nu$ background	Keys PDF	Keys PDF
Combinatorial	Linear (Bernstein)	Keys PDF

Expected Yields Calculation





- Contributions from
  - Signal B_s, B_d decays
  - Combinatorial background
  - Rare B decays, e.g. **B→hh**:

$$n^{\mathrm{exp}}_{B\to hh} = \frac{\epsilon^{\mathrm{tot}}_{B}}{\epsilon^{\mathrm{tot}}_{B^+}} \frac{f_u}{f_s} \frac{\mathcal{B}(B\to hh)}{\mathcal{B}(B^+\to J/\psi K^+)} \times N(B^+\to J/\psi K^+)$$

h→μ fake rates — *Acceptance & analysis efficiency*

$$\epsilon^{\mathrm{tot}}_{B} = w_+(p_T,\eta)\times w_-(p_T,\eta) \times A\times\epsilon_{\mathrm{ana}} \times \frac{1}{2}\epsilon^{\mathrm{signal}}_{\mathrm{trg}}$$

trigger efficiency w/ 100% uncertainty

- **Normalized with fitted N(B⁺→J/ψK⁺→μμK⁺)**
- Complete list of all possible rare 2-body hadronic decays.
- List of 3-body semileptonic decays is incomplete:
  - Scaling factor in low mass sideband
  - Grouping with # of true muons: **B→hμμ** vs. **B→hμυ**

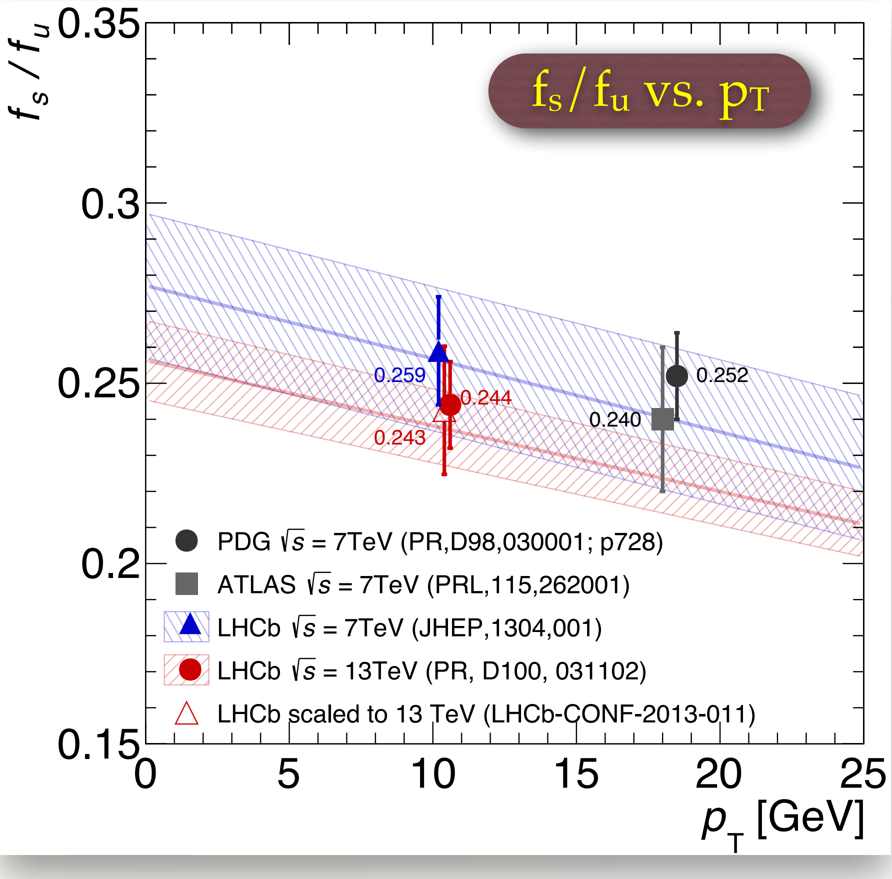
Extensive validations with **inverted muon ID** requirement.

A Comment on f_s/f_u

- The ratio of fragmentation fraction, **f_s/f_u, is an external input** to this measurement; experimental situation is not very clear:
  - **LHCb observes some non-trivial p_T dependence (ref. PRD 100, 031102)**.
  - Not the case for ATLAS **(ref. PRL 115, 262001)**, and not confirmed with CMS internal study using control samples.
- Take the value from PDG (*average of LHCb and ATLAS 7 TeV results*), plus additional uncertainties estimated from LHCb 13 TeV result:

  $f_s/f_u = 0.252 \pm 0.012$ (PDG)
  ± 0.015 (energy/pt dependence)

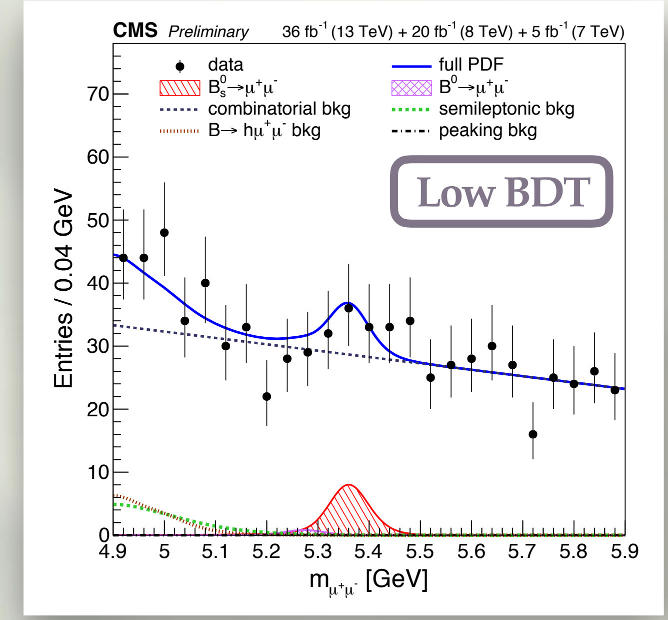
- **Resulting branching fraction can be rescaled**:
  - Average p_T of the CMS B_s events: 17.2 GeV. can be considered.
  - Treated as an external uncertainty (*not as a constrained nuisance parameter, not in likelihood contour scan either*).

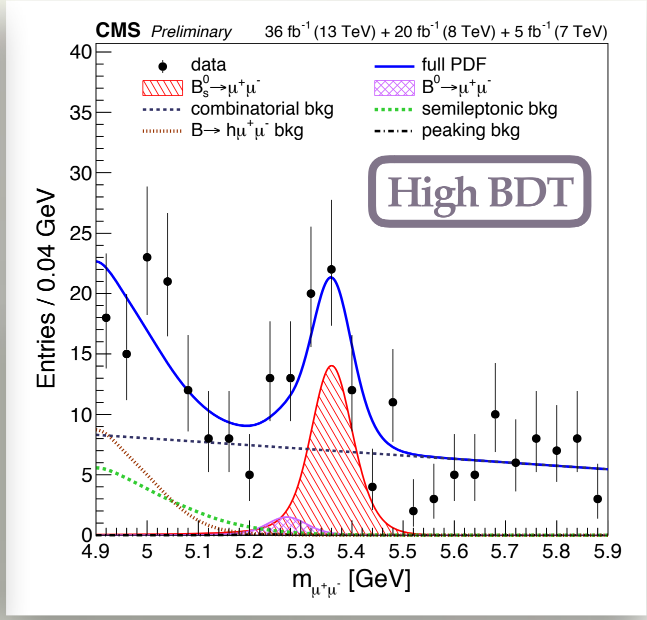


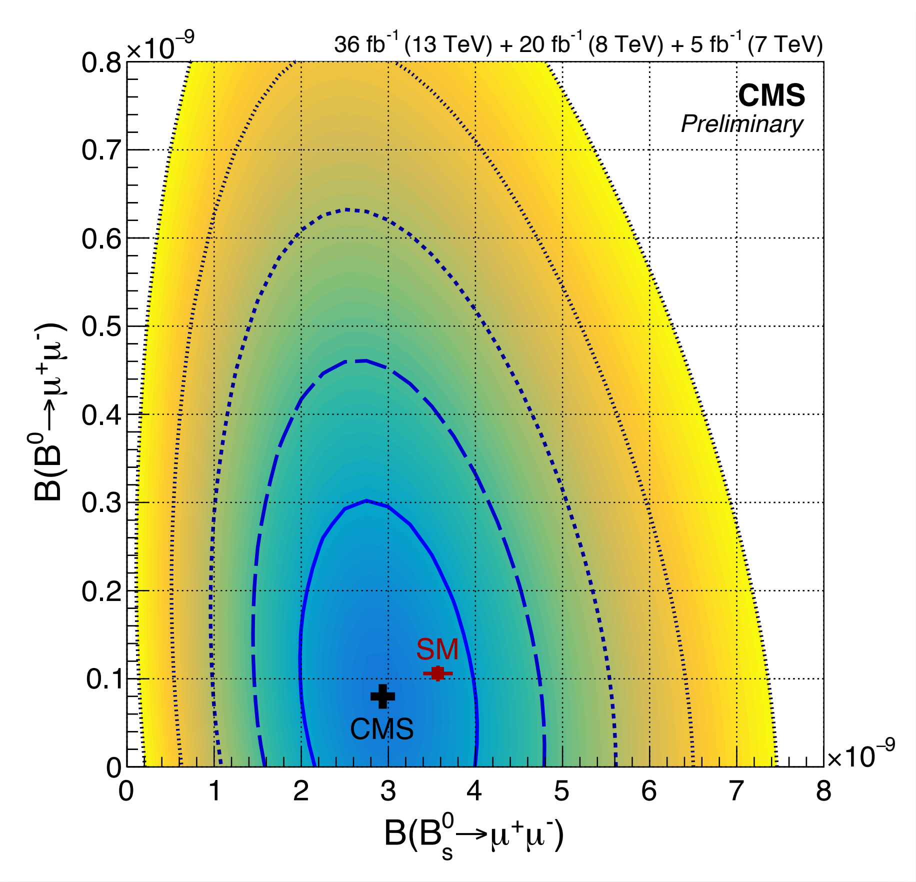
Results: Branching Fractions

Ref. CMS-PAS-BPH-16-004

Combined mass projection for low/high BDT categories







(Correlation: −0.181)

- **Fitted branching fractions:**

Channel	Branching fraction	Sign. (obs)	Sign. (exp)
$B_s \to \mu^+\mu^-$	$[2.9^{+0.7}_{-0.6}\,(\text{exp}) \pm 0.2 (f_s/f_u)\,] \times 10^{-9}$	5.6σ	6.5σ
$B^0 \to \mu^+\mu^-$	$(0.8^{+1.4}_{-1.3}) \times 10^{-10}$	0.6σ	0.8σ

→ Consistent with SM

- **Result with Run-1 subset data only:**

Channel	Branching fraction	Sign. (obs)	Sign. (exp)
$B_s \to \mu^+\mu^-$	$(2.3^{+1.0}_{-0.8}) \times 10^{-9}$	3.3σ	4.5σ

→ Consistent with Nature 522, 68 (CMS part)

Results: Branching Fractions (cont.)

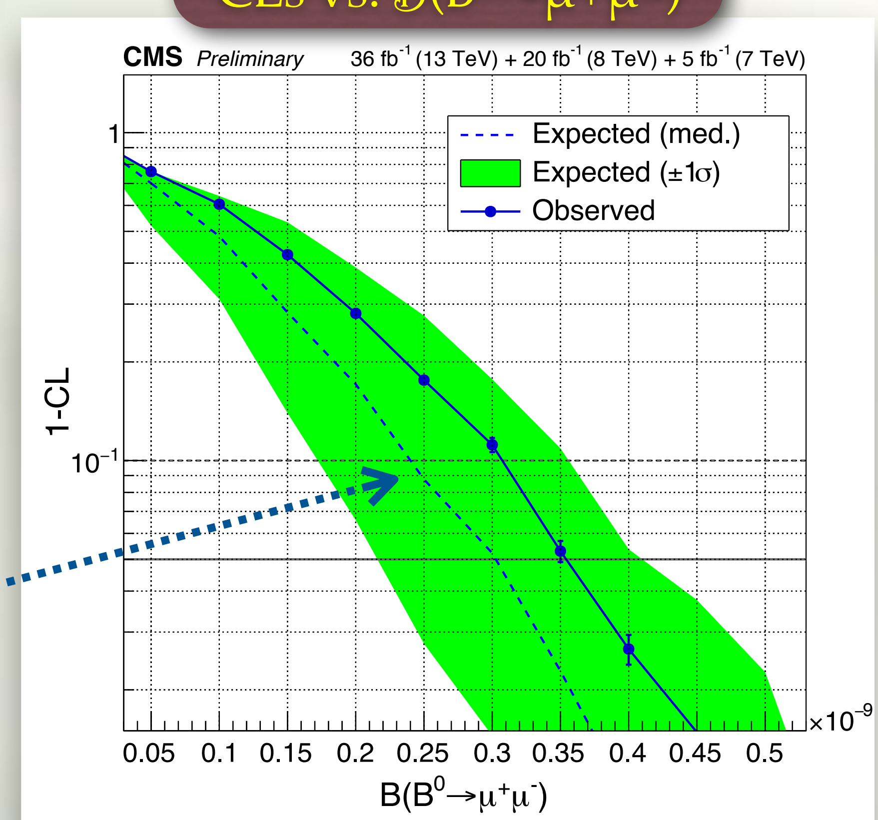
- **Feldman-cousins approach** has been performed:
  - Comparison with the result from profile likelihood scan.
- **One-sided upper limit**:
  - Full CLs prescription, using *LHC-style profile likelihood* as the test statistic.

This is the MAIN result for $B^0 \to \mu^+\mu^-$

Ref. CMS-PAS-BPH-16-004

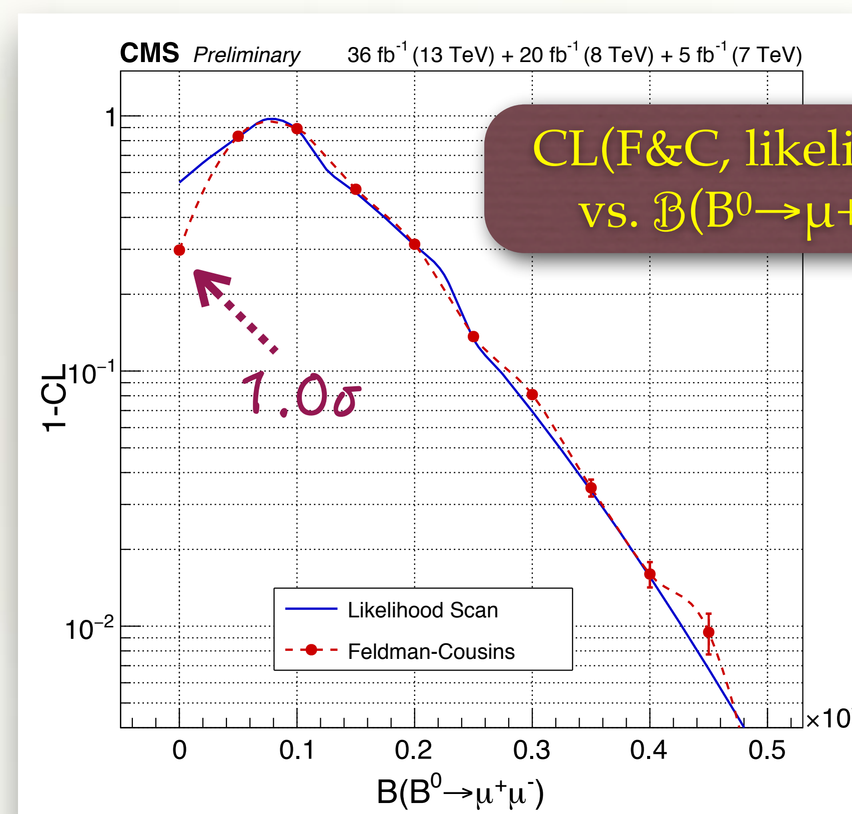
	UL on $B^0 \to \mu^+\mu^-$	
CL	Observed	Expected (null)
95%	**<3.6 × 10^{-10}**	<3.0 × 10^{-10}
90%	**<3.1 × 10^{-10}**	<2.4 × 10^{-10}

CLs vs. $\mathcal{B}(B^0 \to \mu+\mu-)$



Background only hypothesis (no SM $B^0 \to \mu\mu$)

CL(F&C, likelihood) vs. $\mathcal{B}(B^0 \to \mu+\mu-)$



Effective Lifetime

- The proper decay time is measured in 3D: $t = m\ell_{3D}/P$
- **Measurement with 2D unbined likelihood fit** (*primary method*)**:**
  - Fit to dimuon invariant mass $\mathbf{m_{\mu\mu}}$ and decay time **t** (1<t<11 ps); per-event decay time resolution $\boldsymbol{\sigma_t}$ as a conditional parameter in the resolution model.
  - Efficiency correction applied.

PDF Models

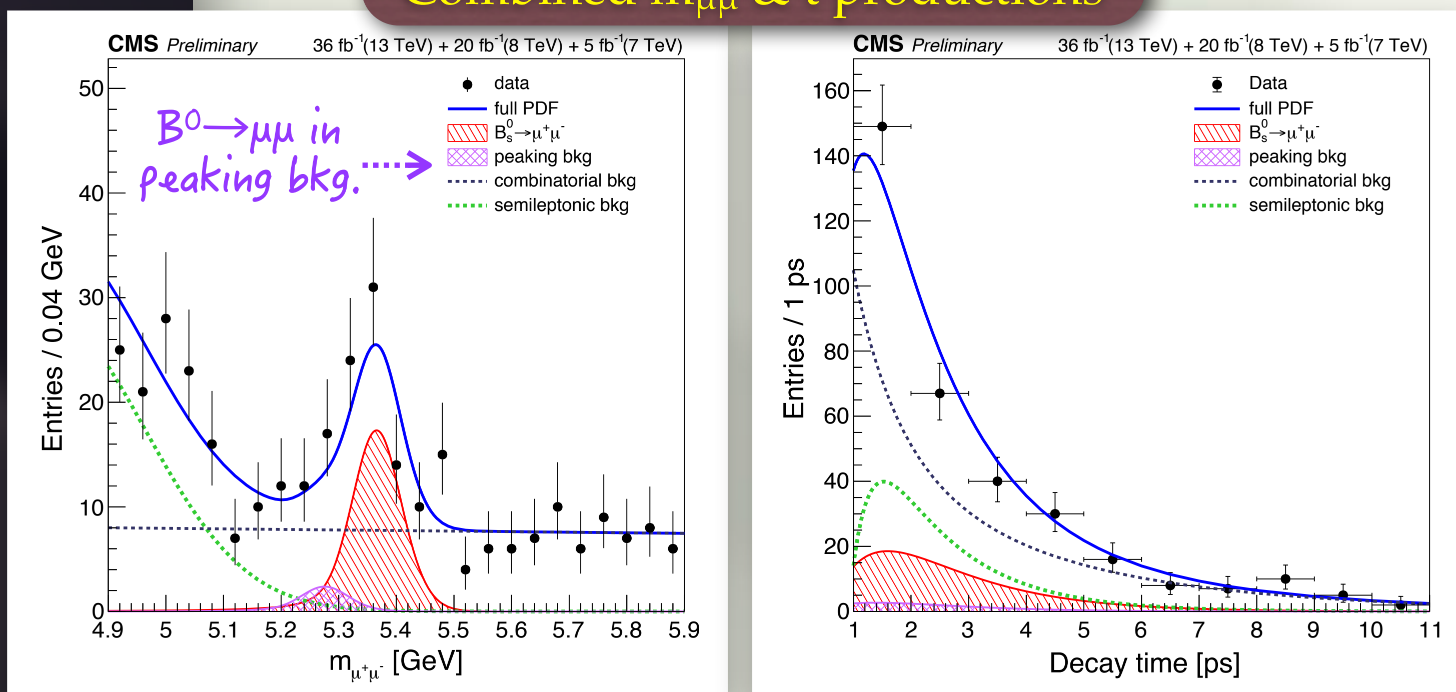
Component	$m_{\mu\mu}$	t
B_s signal	Crystal-ball	Exp(t) ⊗ Res(t;σ_t) ⊗ Eff(t)
$B^0 \to \mu\mu$, $B \to hh$ background	Crystal-ball+Gaussian	Exp(t) ⊗ Res(t;σ_t) ⊗ Eff(t)
$B \to h\mu\mu$, $h\mu\nu$ background	Gaussian	Exp(t) ⊗ Res(t;σ_t) ⊗ Eff(t)
Combinatorial	Linear (Bernstein)	Exp(t) ⊗ Res(t;σ_t) ⊗ Eff(t)

- **Measurement with sPlot:**
  - sPlot weights are derived from the BF model.
  - Weighted binned likelihood fit, with resolution function and efficiency correction in the PDF.
  - Custom implementation for asymmetric uncertainties.

Results from 2 methods are consistent.

Lifetime Results

Combined $m_{\mu\mu}$ & t productions



(projection with 1<t<11 ps, slightly better S/N comparing to BF fit.)

- **Primary result from 2D UML:**

Ref. CMS-PAS-BPH-16-004

$B_s\to\mu\mu$ Effective Lifetime	Expected Uncertainty
$1.70^{+0.61}_{-0.44}$ ps	$^{+0.39}_{-0.30}$ ps

Systematic uncertainty is small: 0.09ps

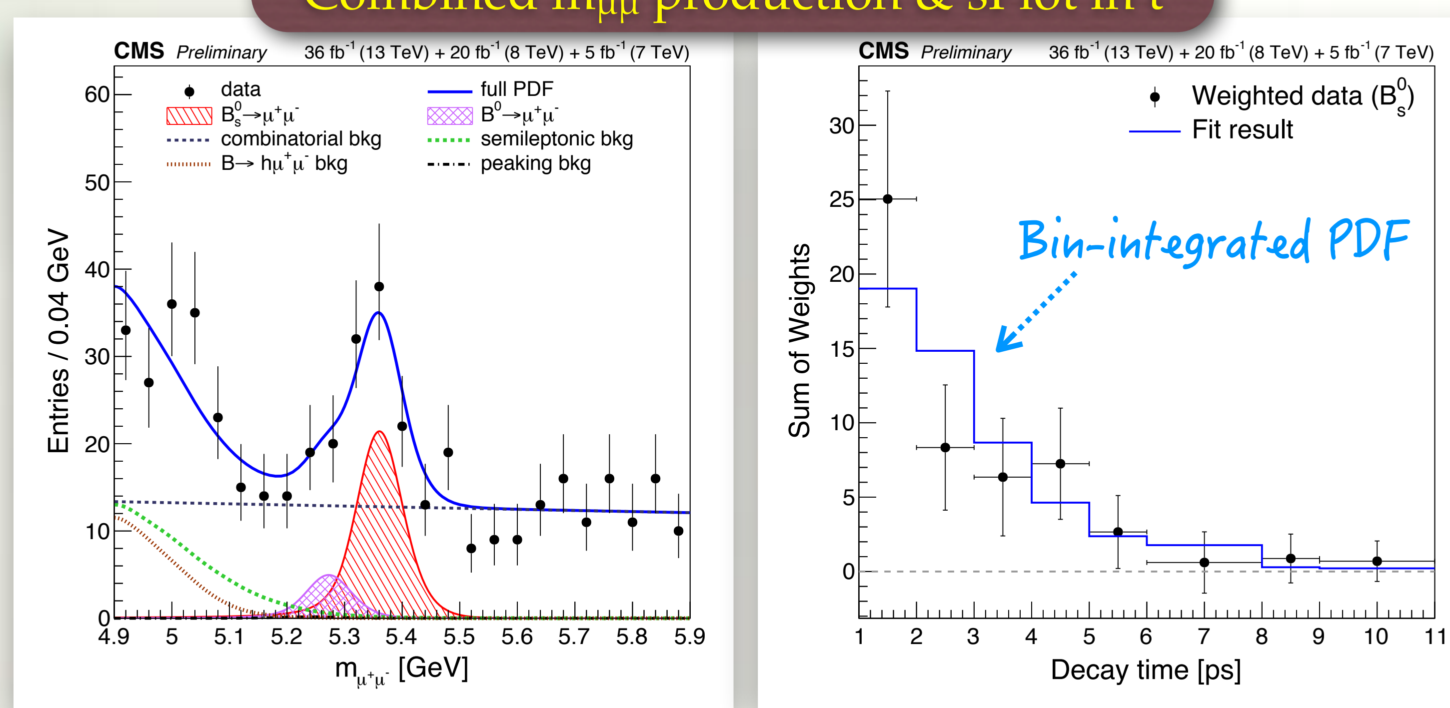
→ Consistent with SM

- **Result from sPlot fit:**

$B_s\to\mu\mu$ Effective Lifetime	Expected Uncertainty
$1.55^{+0.52}_{-0.33}$ ps	$^{+0.49}_{-0.31}$ ps

Consistent with the 2D UML.

Combined $m_{\mu\mu}$ production & sPlot in t



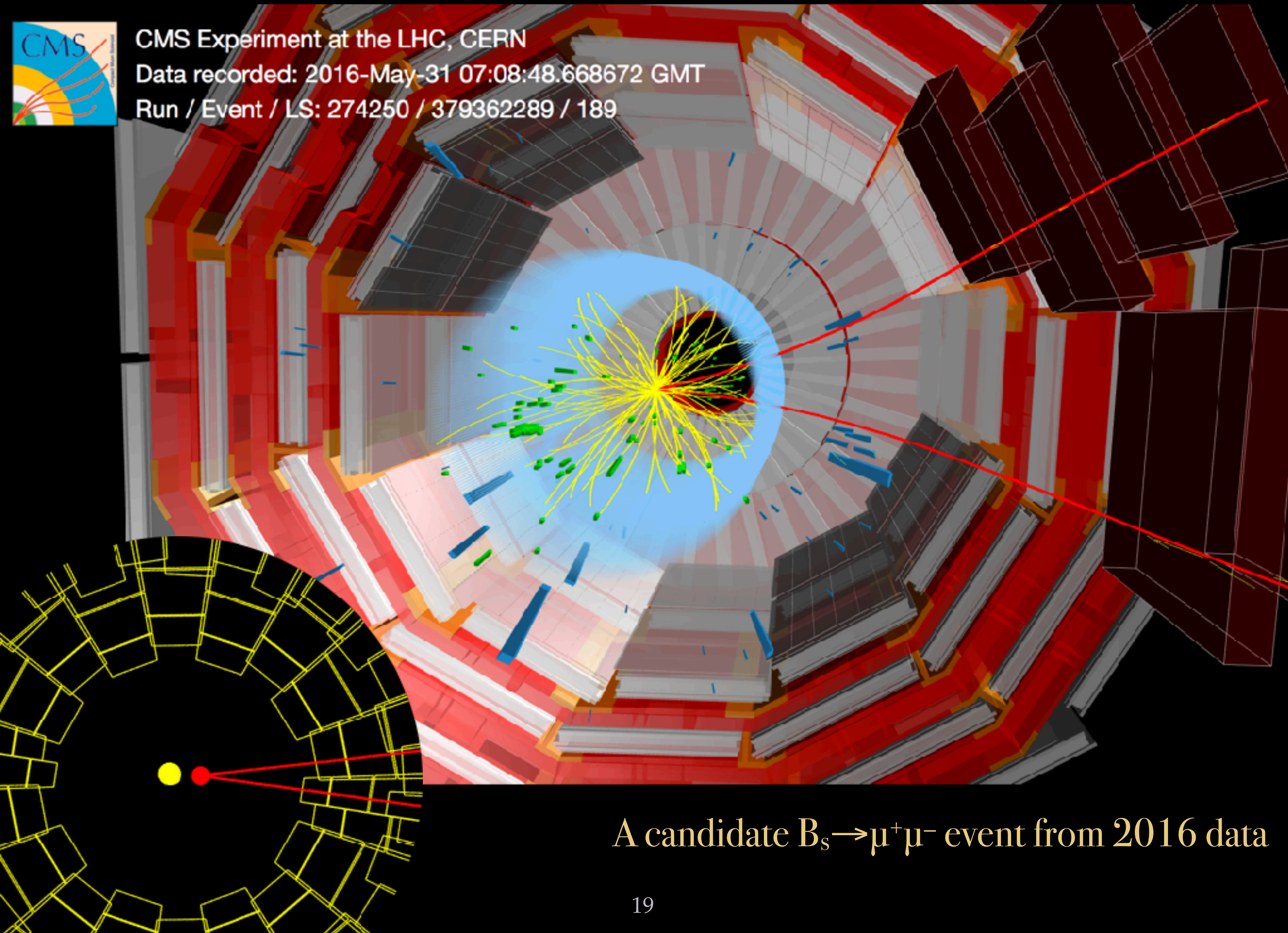
Systematic Uncertainties

	$\mathcal{B}(B_s \to \mu^+\mu^-)$ [%]	Effective Lifetime [ps]	
Sources		2D UML	sPlot
Kaon tracking	2.3-4	-	-
Normalization yield	4	-	-
Background yields	1	0.03	(*)
Production process	3	-	-
Muon ID	3	-	-
Trigger	3	-	-
Eff (data/MC)	5-10	-	(*)
Eff (func. model)	-	0.01	0.04
Eff (lifetime dep.)	1-3	(*)	(*)
Era dependence	5-6	0.07	0.07
BDT threshod	-	0.02	0.02
Tracker alignment	-	0.02	-
Finite size of MC	-	0.03	-
Fit bias	-	-	0.09
C correction	-	0.01	0.01
Total Systematics	(+0.3/–0.2) $\times 10^{-9}$	0.09 ps	0.12 ps
Total Uncertainties	(+0.7/–0.6) $\times 10^{-9}$	+0.61/–0.44 ps	+0.52/–0.33 ps

Ref. CMS-PAS-BPH-16-004

- Relative uncertainties for branching fractions; absolute uncertainties for lifetime.
- Systematic uncertainties are included as constrained nuisances for BF (*except* f_s/f_u).
- Cross checks with a study of $\mathcal{B}(B_s \to J/\psi\phi)$.

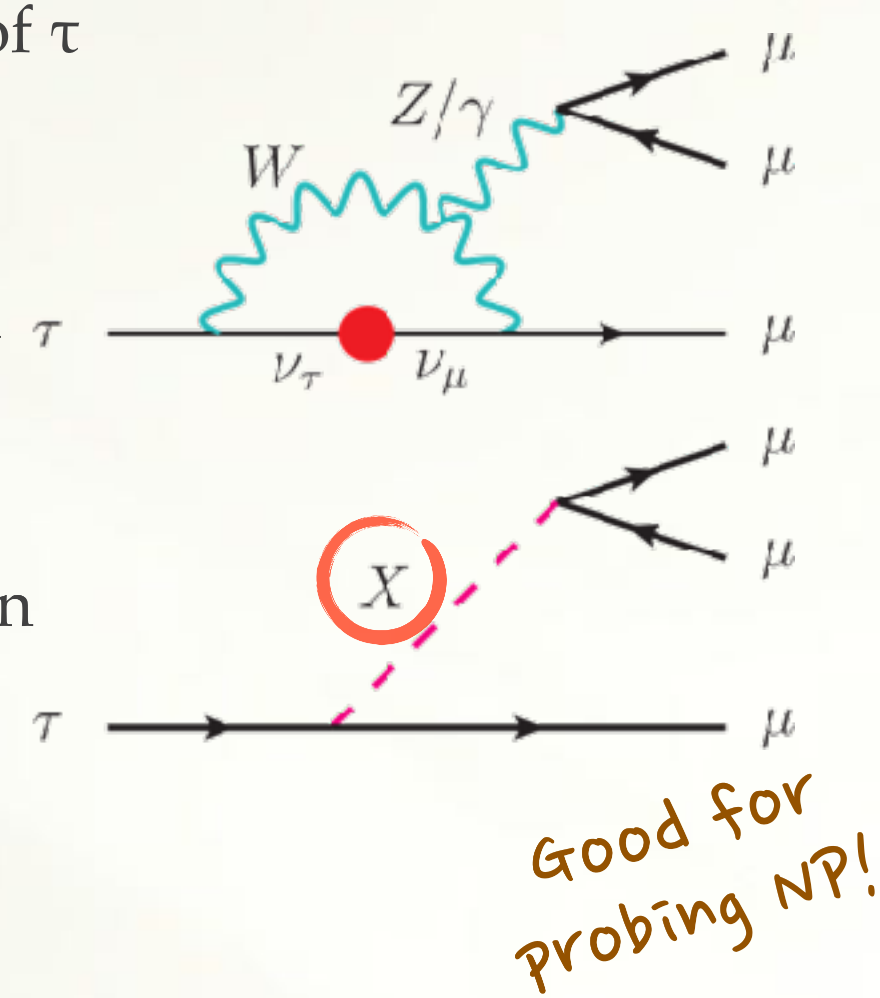
(*) contribution has been included in other items.



A candidate $B_s \rightarrow \mu^+\mu^-$ event from 2016 data

Search for τ→3μ: Introduction

- A charged lepton flavor violating (CLFV) decay of τ to 3 muons, no missing neutrinos.
- Allowed by neutrino oscillations, but with extraordinarily small branching fractions beyond experimental accessibility.
- The rate can be **strongly enhanced with New Physics scenarios**; experimentally the three-muon final state is accessible and clean.
- Searches have been performed by Belle, BaBar, LHCb, ATLAS, no hint of signal yet.
- Best limit from Belle: $\mathcal{B} < 2.1\times10^{-8}$ (@ 90% C.L.)
  [PLB 687 (2010) 139143]



CMS performed a search for the τ → 3μ, where τ leptons produced in D and B hadron decays, using the data collected in 2016 of 33 fb^{-1}.

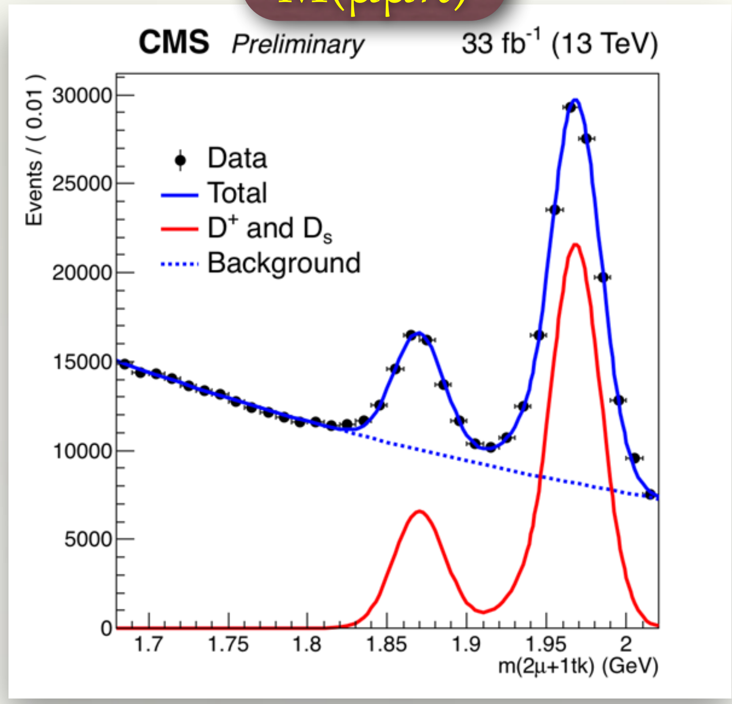
Event Selection

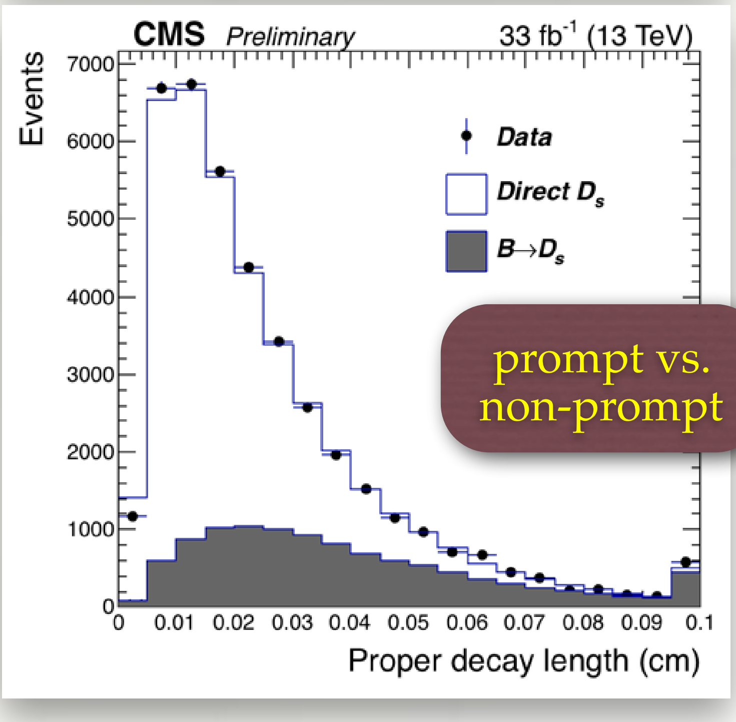
- **For τ candidate:**
  - Triggering with two muons plus a track, with vertex and mass requirement;
  - Requiring 3 global muons offline and sum of charges should be ±1.
- **For normalization:**
  - Select $D_s^+ \to \phi\pi^+ \to \mu^+\mu^-\pi^+$ with the same trigger and very similar momentum thresholds.
  - The fraction of (non-)prompt D_s estimated from a fit to the proper decay length distribution.

Yields for $BF(\tau \to 3\mu) = 10^{-7}$

	D→τ Signal	B→τ Signal	Data
Production	4.4×10⁵	1.5×10⁵	
3μ in fiducial volume	6.6×10³	2.3×10³	
Trigger	214	114	
3μ p_T>2GeV	88	47	1×10⁷
3μ candidate	64	29	1×10⁵

M(μμπ)



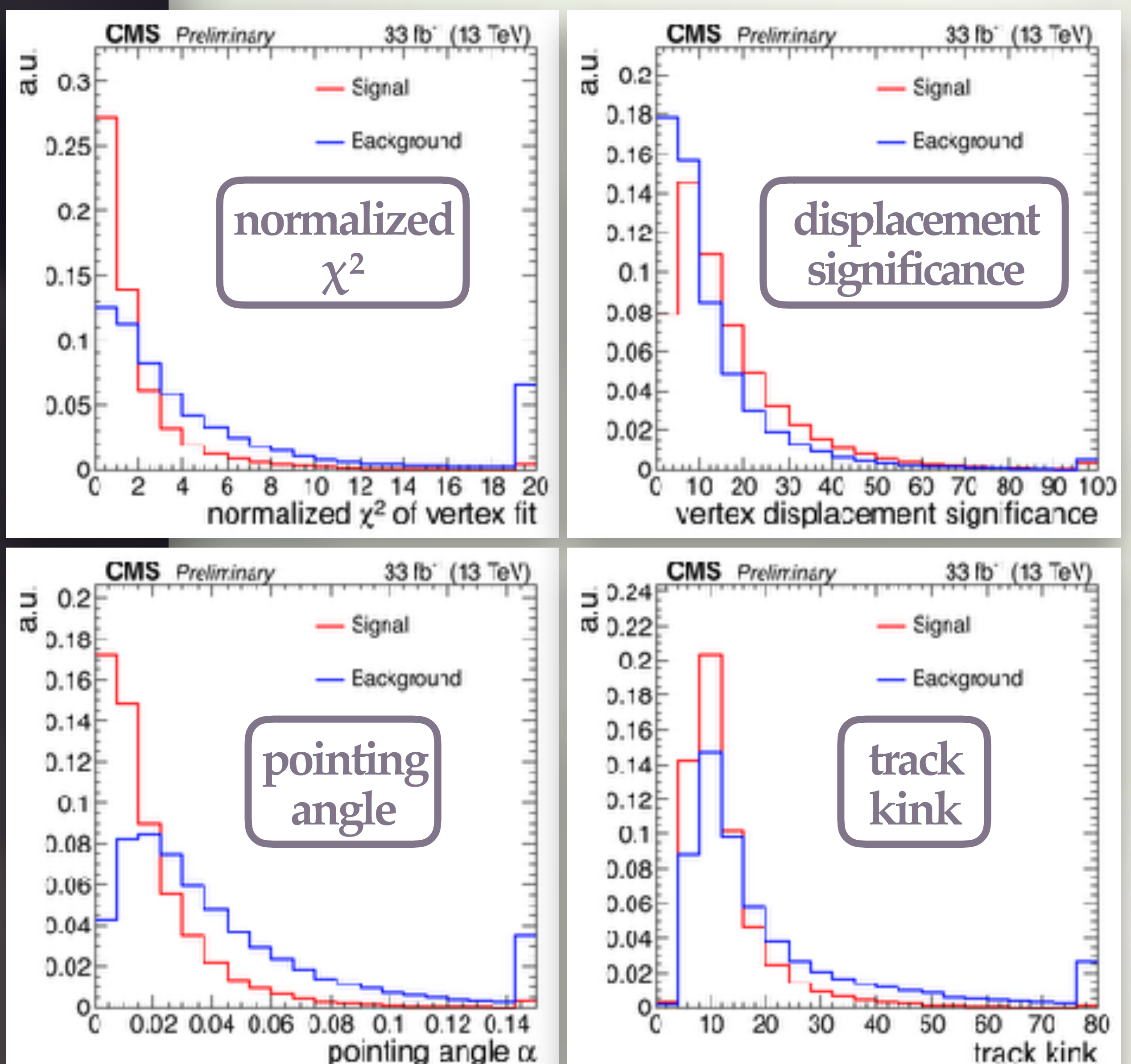


prompt vs. non-prompt

Ref. CMS-PAS-BPH-17-004

Background Suppression Variables

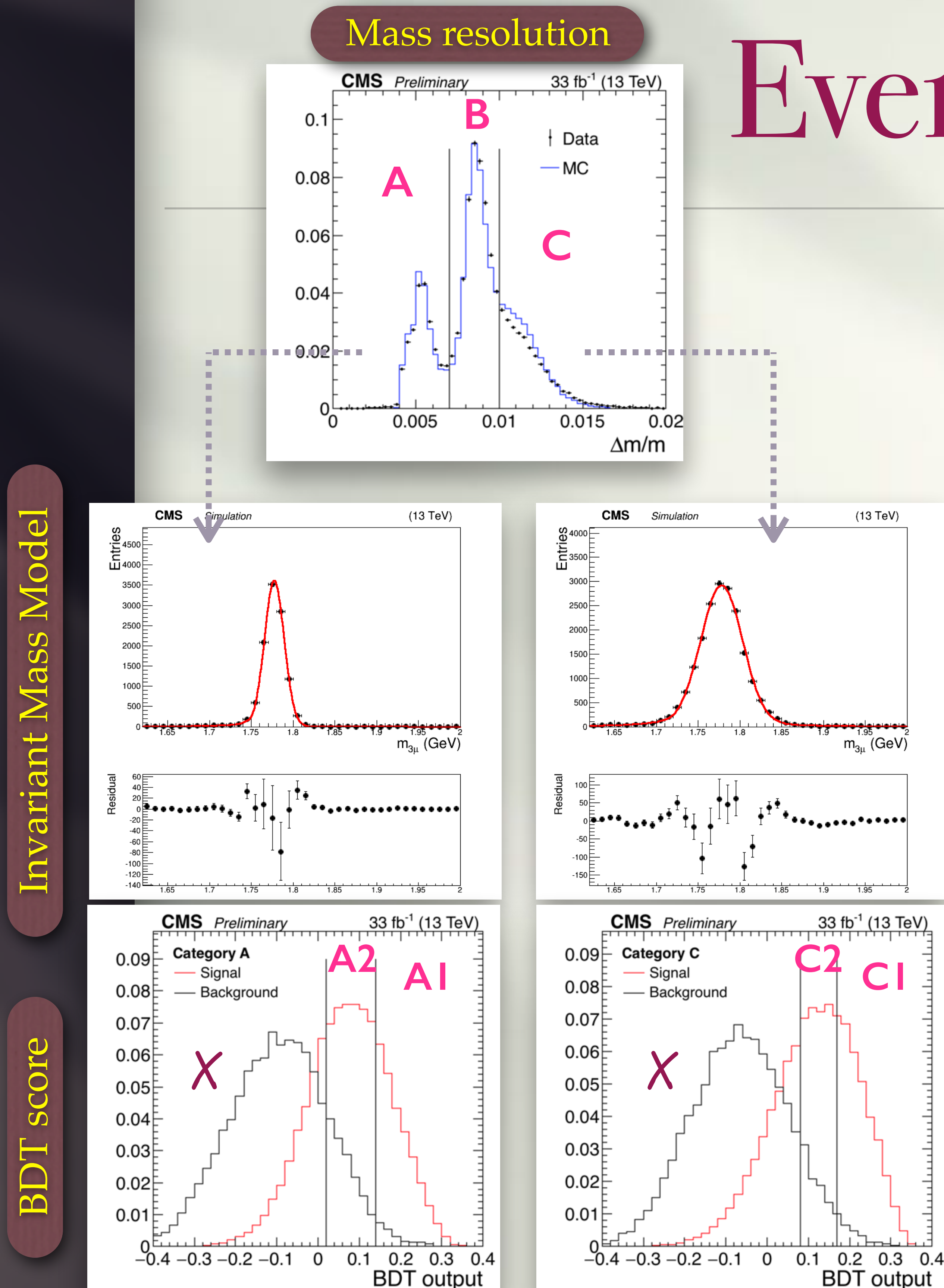
(⟹ 4 most discriminating variables)



Included in the Boosted Decision Tree

- **Associated with 3-μ vertex:**
  - normalized χ2 from 3-μ vertex fit;
  - significance of the 3-μ vertex 3D displacement w.r.t. IP;
  - pointing angle α;
  - closest distance between the 3-μ vertex and any other track;
- **Associated with fake-μ from hadron:**
  - kink parameter for the tracker track;
  - spatial compatibility of tracker track and muon track in muon system;
  - transverse impact parameter;
  - muon momentum;
  - track-based isolation;

Event Categorizing



- **#1 Categorized by mass resolution:**
  - 3-μ mass resolution has been evaluated event-by-event, ranging from 0.4 to 1.5%, depending on the muon rapidity;
  - Divided into 3 categories: <0.7%, 0.7–1%, >1%.
- **#2 Categorized by BDT score:**
  - BDT trained with signal MC and sideband data as background;
  - Further divided into 3 categories, 2 to be included, 1 dropped.
  - **Resulting 6 categories in total.**

→ *To be included in the fit*

Ref. CMS-PAS-BPH-17-004

Systematic Uncertainties

- Systematic uncertainties are incorporated via nuisance parameters.
- Small systematics, statistical uncertainties dominant.

Sources	Input uncertainty	Resulting variation
Ds normalization	10%	10%
$\mathcal{B}(D_s \to \tau\upsilon)$	4%	3%
$\mathcal{B}(D_s \to \phi\pi \to \mu\mu\pi)$	8%	8%
$\mathcal{B}(B \to D_s + X)$	16%	5%
$\mathcal{B}(B \to \tau + X)$	11%	3%
Uncertainty of B/D ratio	11%	3%
Uncertainty due to τ from D	100%	3%
Uncertainty due to τ from Bs	100%	4%
3-muon trigger	8%	2%
Acceptance ratio	1%	1%
Muon reconstruction	1.5%	1.5%
Pion reconstruction	2.3%	2.3%
BDT efficiency	5%	5%
Mass scale	0.07%	shape
Mass resolution	2.5%	shape

Stability check across different data taking era

Data/MC difference

Measured from 3μ mass sideband

Studied in $J/\psi \to \mu\mu$ with tag & probe

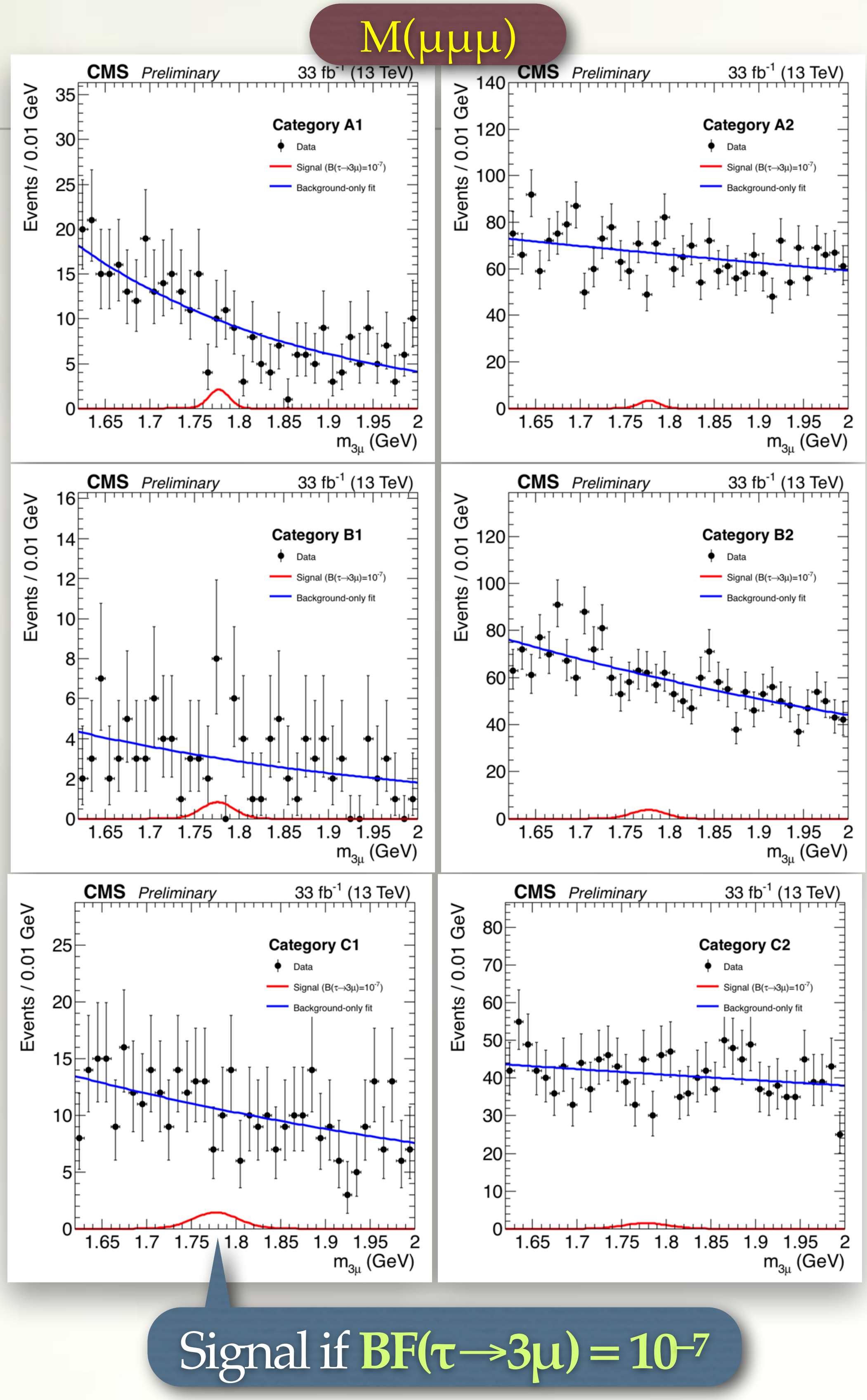
Estimated with $D_s \to \phi\pi \to \mu\mu\pi$ MC & data

Extraction of Limit

- Simultaneous maximum likelihood fit to 3μ invariant mass, performed with 6 resolution-BDT categories.
- Signal model is parametrized with Crystal Ball functions
- Background is modeled with an exponential plus a polynomial.
- **No hint of signal found**, observed (expected) limits are evaluated using CLs method:

$\mathcal{B}(\tau\to3\mu) < 8.8\ (9.9) \times 10^{-8}$ @ 90% C.L.

$\mathcal{B}(\tau\to3\mu) < 1.1\ (1.2) \times 10^{-7}$ @ 95% C.L.



Ref. CMS-PAS-BPH-17-004

Summary

CMS is an unique test bench for flavor physics predictions!

Ref. CMS-PAS-BPH-16-004

- **Measurement of the $B_s \to \mu\mu$ and search for $B^0 \to \mu\mu$**
  - **The rare decay $B_s \to \mu\mu$ has been measured** and the branching fractions have been updated. More data are required for probing $B^0 \to \mu\mu$ decays in the near future.
  - **Effective lifetime measurement with $B_s \to \mu\mu$ events** at CMS has been carried out for the first time.
  - All results are consistent with the SM; full Run-2 analysis is ongoing.

Ref. CMS-PAS-BPH-17-004

- **Search for $\tau \to \mu\mu\mu$ decay**
  - **Search of CLFV decay $\tau \to 3\mu$ has been conducted** at CMS. Using the τ leptons decaying from D and B mesons, no excess above the expected background is observed. Upper limits have been set.

More results are in the pipeline! Stay tuned!