Towards searches for D^{*0} and $B_{(s)}^{*0} \to \mu^+\mu^-$ decays

Fernando Abudinén, Tom Blake, Ulrik Egede, Tim Gershon

MWAPP meeting
February 2, 2022

- Paper draft on prospects
- Ongoing search for $D^{*0} \to \mu\mu$

MONASH University

WARWICK
THE UNIVERSITY OF WARWICK

Paper draft on prospects

Prospects for studies of $D^{*0} \to \mu^+\mu^-$ and $B^{*0}_{(s)} \to \mu^+\mu^-$ decays

F. Abudinén[1], T. Blake[1], U. Egede[2], T. Gershon[1]

[1] *Department of Physics, University of Warwick, Coventry, United Kingdom*
[2] *School of Physics and Astronomy, Monash University, Melbourne, Australia*

Abstract

Weak decays of the vector D^{*0} and $B^{*0}_{(s)}$ mesons to the $\mu^+\mu^-$ final state provide novel potential to test the Standard Model of particle physics. Such processes have extremely small branching fractions as the vector mesons are able to decay through electromagnetic and (for the D^{*0} meson) strong interactions. Nonetheless, the production of copious quantities of these particles in LHC collisions, and the ability to exploit experimental techniques that can suppress background to low levels, provides good potential to reach interesting sensitivity. The possibility to reconstruct these processes as part of the decay chain of B^- or B_c^+ mesons appears particularly attractive due to the clean experimental signature of the displaced vertex. Indeed, published LHCb data on $B^- \to \pi^-\mu^+\mu^-$ decays already implies a stringent limit on the branching fraction of $D^{*0} \to \mu^+\mu^-$. Estimates are made on the achievable sensitivity to $D^{*0} \to \mu^+\mu^-$ and $B^{*0}_{(s)} \to \mu^+\mu^-$ decays with the LHCb experiment.

Motivation

- Decays of heavy-flavored vector mesons into lepton pairs

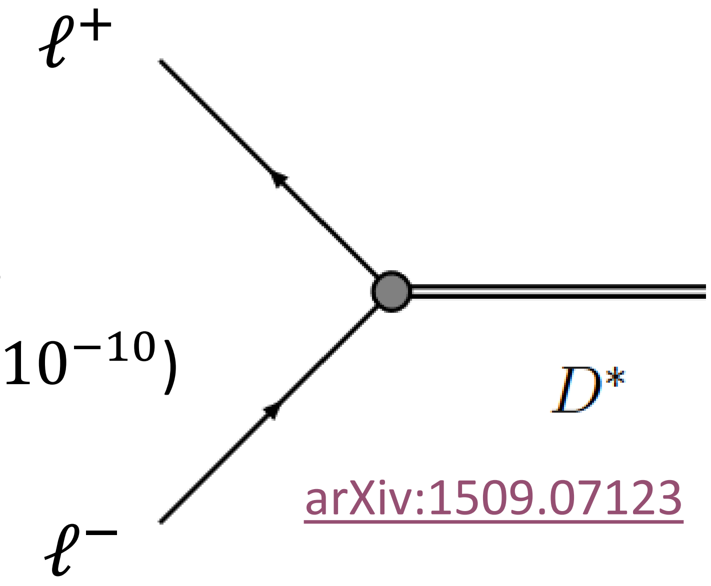
$\Rightarrow$ Probe same operators as the pseudoscalar decays

$\Rightarrow$ Not helicity suppressed and hence complementary

- D^{*0} and $B_{(s)}^{*0}$ mesons decay predominantly via strong or EM int.

$\Rightarrow$ Very small $V \to \mu\mu$ branching fractions predicted in the SM ($\lesssim 10^{-10}$)

$\Rightarrow$ Challenging to measure



arXiv:1509.07123

- Might be experimentally achievable at LHC (due to large number of produced D^{*0} and $B_{(s)}^{*0}$) if backgrounds can be kept low enough

$\Rightarrow$ Production through $B_{(c)}^{+} \to \mu^{+}\mu^{-}\pi^{+}$decays provide a way to keep low bkg. levels by exploiting the displaced vertex signature

Experimental status

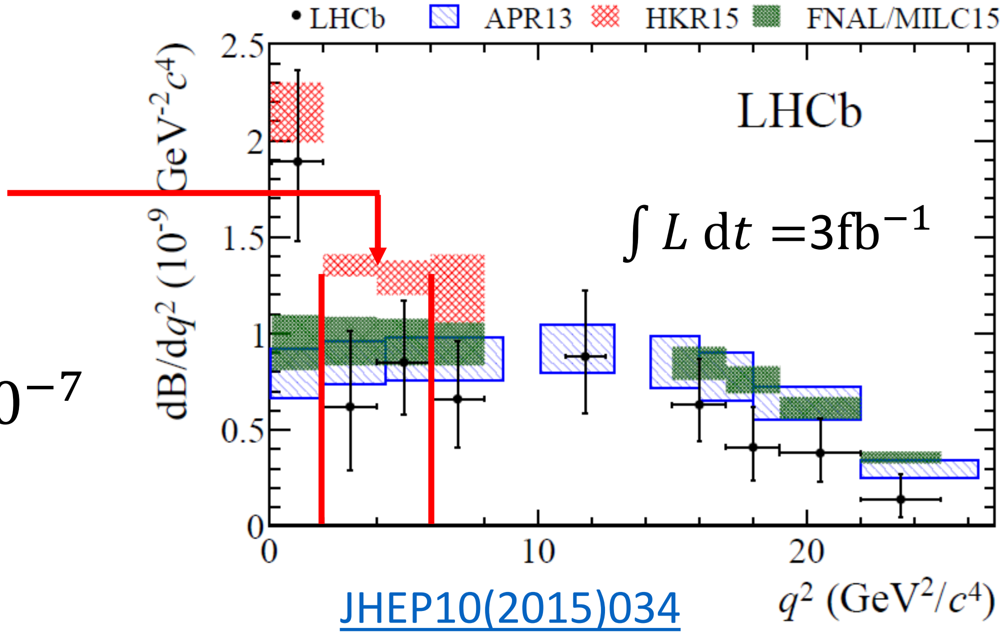
- Preliminary results for direct search of $e^+e^- \to D^{*0}$ at CMD-3:

$$\mathcal{B}(D^{*0} \to e^+e^-) < 1.7 \cdot 10^{-6} \text{ at } 90\% \text{ C.L.}$$

Phys. At. Nuc. 83, 954-957 (2020)

- Limit looking at previous measurements of $d\mathcal{B}(B^+ \to \mu^+\mu^-\pi^+)/dq^2$ at LHCb:

Assume that $\mathcal{B}(B^- \to [\mu^+\mu^-]_{D^{*0}}\, \pi^-)$ less than half the signal in the two bins around $M_{D^{*0}}^2$



JHEP10(2015)034

$$\Rightarrow \mathcal{B}(D^{*0} \to \mu^+\mu^-) = \frac{\mathcal{B}(B^- \to [\mu^+\mu^-]_{D^{*0}}\, \pi^-)}{\mathcal{B}(B^- \to D^{*0}\pi^-)} \lesssim 3 \cdot 10^{-7}$$

$\Rightarrow$ Expect at least ten times higher sensitivity with a dedicated search in 9fb^{-1}.

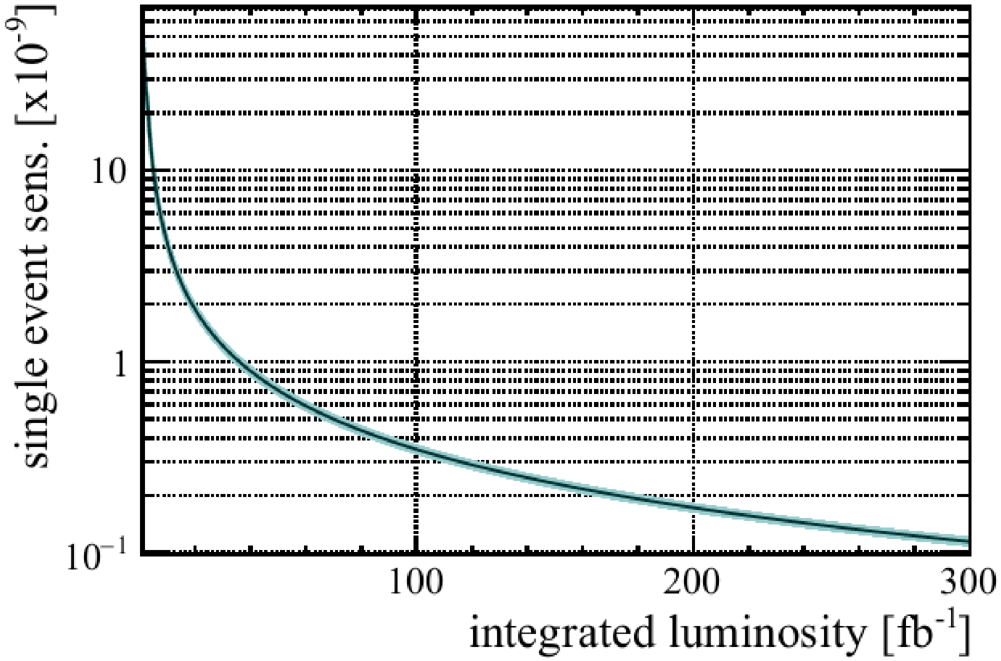
Search for $D^{*0} \rightarrow \mu^+\mu^-$

- Reconstruct $B^+ \rightarrow \mu^+\mu^-\pi^+$ candidates and search for peak in the $m(\mu^+\mu^-)$ distr.
- Normalize to $B^+ \rightarrow J/\psi(\rightarrow \mu^+\mu^-)K^+$decays

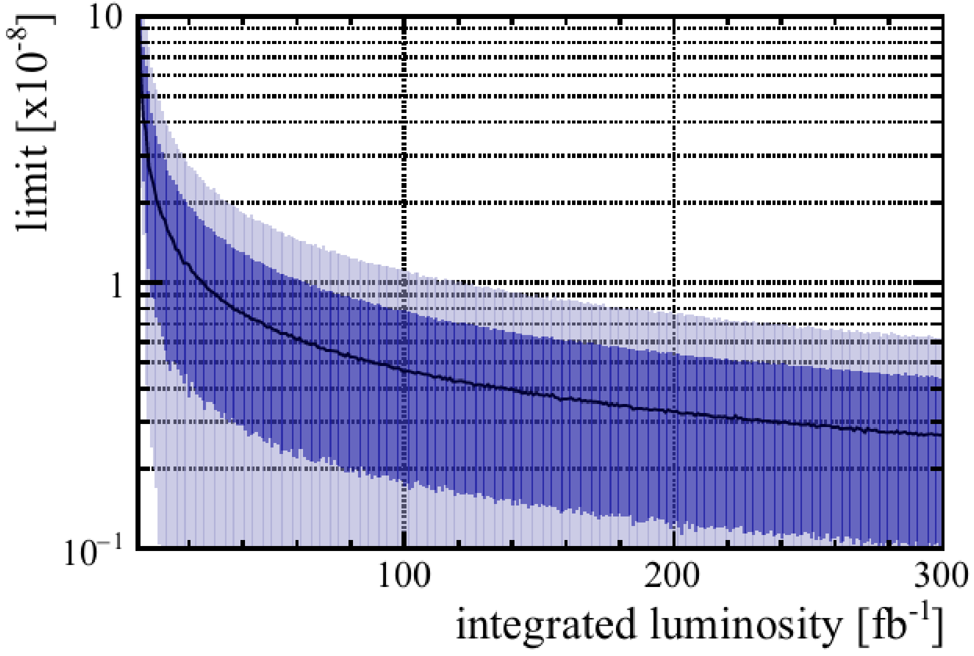
All known

$$\mathcal{B}\left(D^{*0} \rightarrow \mu^+\mu^-\right) = \frac{N_{B^- \rightarrow D^{*0}\pi^-}}{N_{B^- \rightarrow J/\psi K^-}} \frac{\epsilon_{B^- \rightarrow J/\psi K^-}}{\epsilon_{B^- \rightarrow D^{*0}\pi^-}} \frac{\mathcal{B}\left(B^- \rightarrow J/\psi K^-\right)}{\mathcal{B}\left(B^- \rightarrow D^{*0}\pi^-\right)} \mathcal{B}\left(J/\psi \rightarrow \mu^+\mu^-\right)$$
$$= \alpha_{D^{*0} \rightarrow \mu^+\mu^-} N_{B^- \rightarrow D^{*0}\pi^-}$$

Single event sensitivity



Limit from simulated experiments including expected combinatorial, and non-res. $\mu^+\mu^-\pi^+$and $\mu^+\mu^-K^+$ bkg.



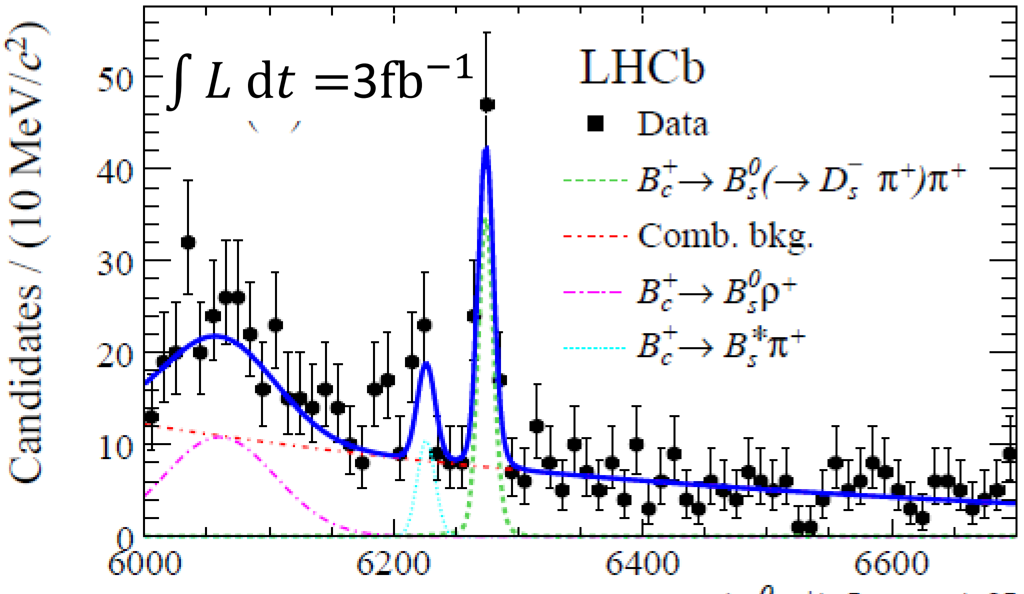
Search for $B^{*0}_{(s)} \rightarrow \mu^+\mu^-$

- Similar search but missing denominators $\mathcal{B}(B_c^+ \rightarrow B_s^*\pi^+)$ and $\mathcal{B}(B_c^+ \rightarrow B^{*0}\pi^+)$.

- Possible to measure $\mathcal{B}(B_c^+ \rightarrow B_s^*\pi^+)$ via partial reconstruction $\Rightarrow$

$$\mathcal{B}(B_s^* \rightarrow \mu^+\mu^-) = \frac{\mathcal{B}(B_c^+ \rightarrow [\mu^+\mu^-]_{B_s^*}\pi^+)}{\mathcal{B}(B_c^+ \rightarrow B_s^*\pi^+)}$$

$\Rightarrow$ Using as normalization $B_c^+ \rightarrow J/\psi\,\pi^+$ for numerator and $B_s^0 \rightarrow J/\psi\,\phi$ for denominator, and ratio $\sigma(B_s^0)/\,\sigma(B^+)$.



PRL.111.181801

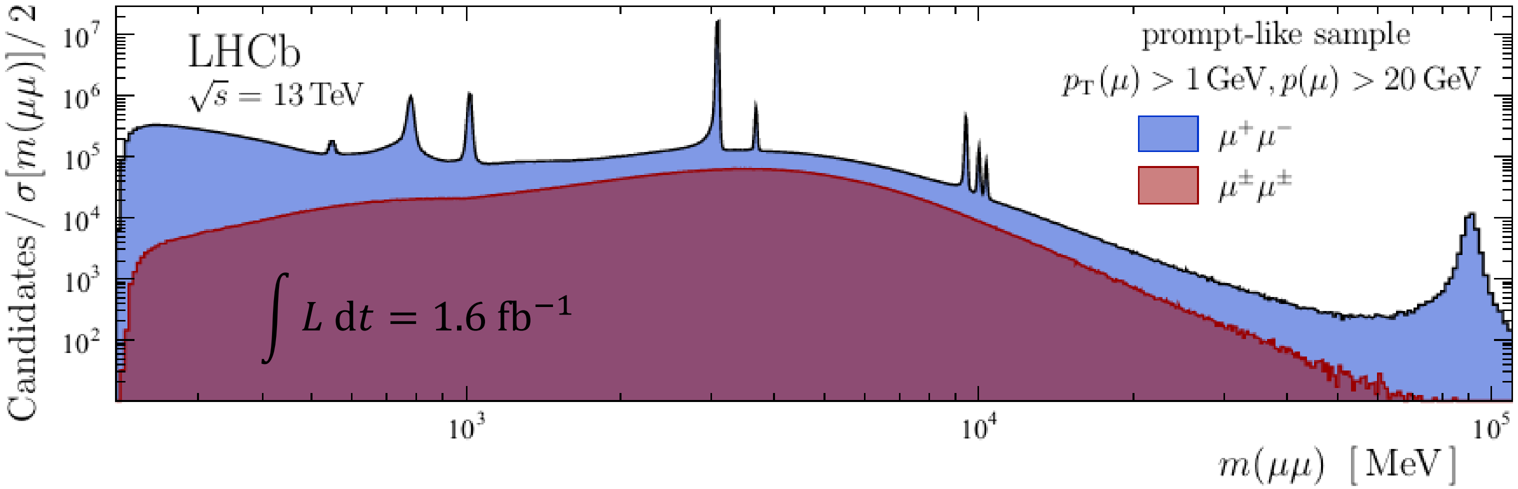
- $B_c^+ \rightarrow B^{*0}\pi^+$ Cabibbo-suppressed (and B^{*0} mass not measured yet) but possibly observable with a similar reconstruction technique in 9fb^{-1}.

Searches exploiting prompt decays

- D^{*0}, $B^{*0}_{(s)}$ produced directly in pp collisions

$\Rightarrow$ Requires respective cross sections (not yet measured)

Dimuon mass spectrum from dark photon search PRL120.061801



- Assuming same cross section for D^{*+} and D^{*0}, the exp. single event sensitivity is $\sim 10^{-11}$
- Assuming same background level as for dark photon search, the expected limit at $1.6\ \text{fb}^{-1}$ is

$$\mathcal{B}(D^{*0} \to \mu^+\mu^-) \lesssim 10^{-7}$$

$\Rightarrow$ Similar sensitivity as in search through $B^+_{(c)} \to \mu^+\mu^-\pi^+$ decays, but background dominated and missing cross section

Semi-inclusive searches

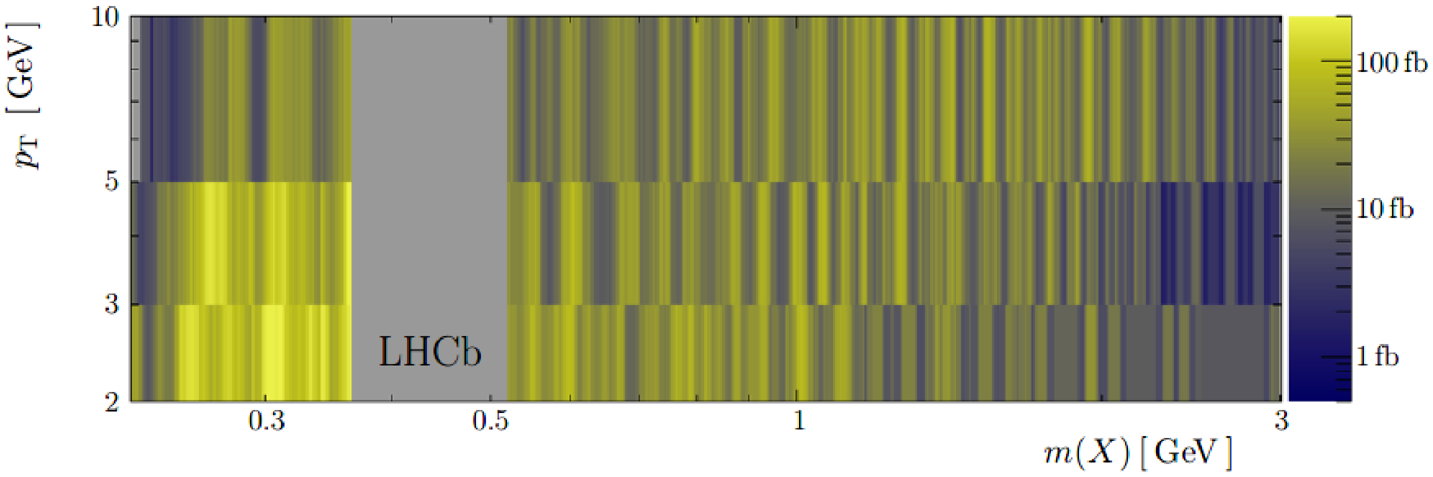
- Search for D^{*0}, $B^{*0}_{(s)} \to \mu\mu$ decays from a displaced (b-hadron) decay vertex
- Example:

$$N_{D^{*0}\to\mu^+\mu^-} = \mathcal{L}\,\mathcal{B}\left(D^{*0}\to\mu^+\mu^-\right)\sum_i \sigma_{B_i}\,\mathcal{B}\left(B_i\to D^{*0}X\right)\,\epsilon_{B_i\to D^{*0}X}$$

$\Rightarrow$ Production cross sections and inclusive branching fractions known for dominant $B^{\pm(0)}$ hadrons (in the D^{*0} case).

$\Rightarrow$ Analysis in principle possible, but worse dimuon mass resolution and higher irreducible backgrounds

Limit at 90% CL on $\sigma(X)\cdot\mathcal{B}(X\to\mu\mu)$ JHEP10(2020)156



$\Rightarrow$ Most likely less sensitive than exclusive searches

Ongoing search for $D^{*0} \rightarrow \mu^{+}\mu^{-}$

Goal

Perform a dedicated search for $D^{*0} \to \mu^+\mu^-$ in $B^+ \to \mu^+\mu^-\pi^+$ decays

$\Rightarrow$ Extend the search then to $B_{(s)}^{*0} \to \mu^+\mu^-$ in $B_c^+ \to \mu^+\mu^-\pi^+$ decays

Analysis overview

- Reconstruction and selection of B candidates in data and simulation
  (Start from analysis tools and selection for $B^+ \to \mu^+\mu^-\pi^+$)

- Determination of signal efficiency from MC and corrected for data/MC discrepancies
  (from which other $B^+ \to \mu^+\mu^-\pi^+$ analyses also profit)

- Fit of sample composition for signal and normalization channel ($B^+ \to J/\psi\, K^+$) to obtain yields

$\Rightarrow$ Obtain limit for the branching fraction (plan to follow Feldman-Cousins method)

Reconstruction and selection

J. Albrecht, G. Ciezarek,
C. Langenbruch, M. McCann,
M. Patel, K. Petridis, A. Shires,
T. Tekampe, M. Whitehead

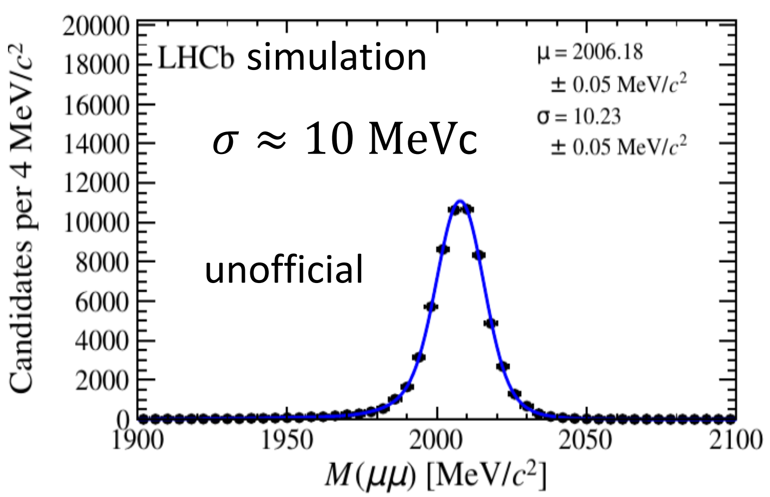
Currently aligned with the differential $\mathcal{B}(B^+ \to \mu^+\mu^-\pi^+)$ analysis:

- Baseline selection using topological and muon ID info
- Trigger selection
- Multivariate (BDT) selection exploiting topo., kinematic and vertex info
- Muon and hadronic particle ID selection

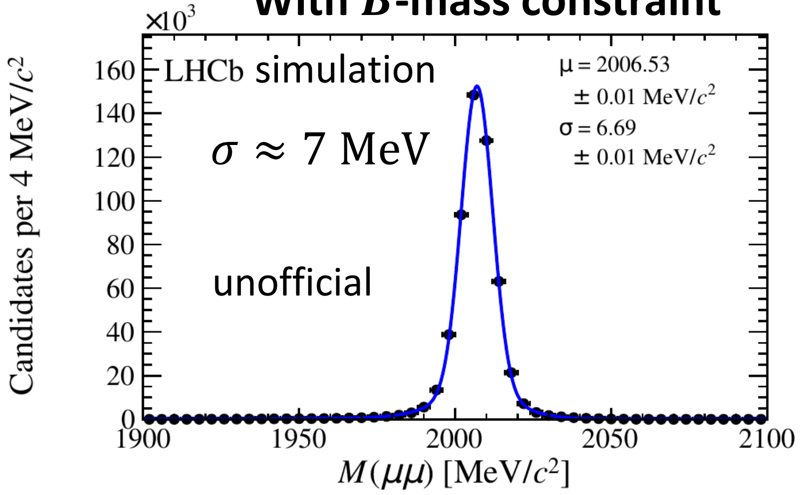
} Currently studying possible optimization

Apply B-mass constraint to improve dimuon mass resolution

No B-mass constraint

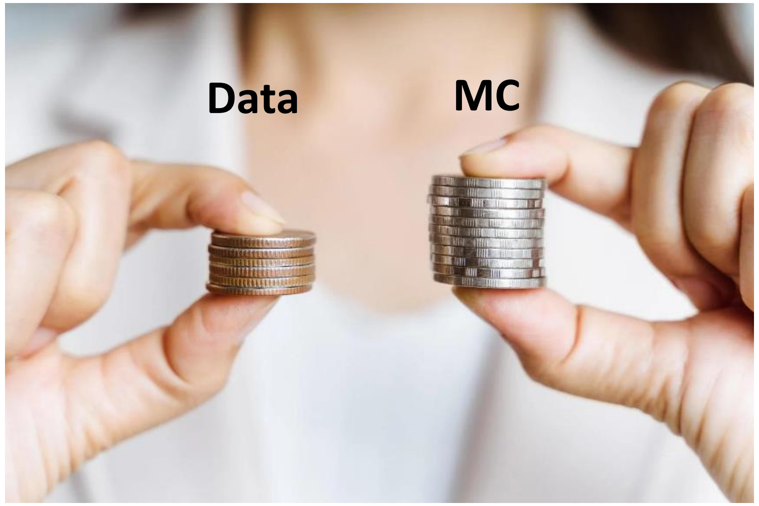


With B-mass constraint



$\Rightarrow$ Improvement in $M(\mu\mu)$ resolution by factor ≈ 1.5 using B-mass constraint.

Determination of efficiencies



- Obtain efficiencies from correctly associated candidates reconstructed in simulation.
- Correct for data/MC discrepancies associated with:
  - PID selection (PID Calib)
  - Muon-ID selection (dedicated MC and PID Calib)
  - Track reconstruction (Track Calib)
  - Trigger selection (TIS-TOS method)
  - p_{t} and track multiplicity

(Track reconstruction and Trigger selection:) using $B^{+} \to J/\psi\, K^{+}$ data

$\Rightarrow$ All corrections produced in this work in common framework for $h\mu\mu$ analyses.

PID Efficiency corrections

- Apply efficiency corrections obtained with PIDcalib
- Instead of applying PID cuts on MC, weight each event

$$\Rightarrow \quad \varepsilon = \frac{\sum_i^{N_{\text{sel}}} \varepsilon_i^{\text{Data}}}{N_{\text{gen}}} \qquad \varepsilon_i^{\text{Data}} = \varepsilon_\mu(\mu_i^+) \cdot \varepsilon_\mu(\mu_i^-) \cdot \varepsilon_h(h_i^+)$$

h : charged K or π

$\varepsilon_\mu, \varepsilon_h$: Efficiencies in data from PIDcalib histograms

N_{sel}: Entries after full selection excluding PID reqs.

N_{gen}: Entries in MCDecayTreeTuple

Example:

$B^+ \rightarrow \bar{D}^{*0}\,\pi^+$

Year\Polarity	Up	Down
2018 (cuts)	8.92 ± 0.04	8.95 ± 0.04
2018 (weights)	7.72 ± 0.03	7.76 ± 0.03

$\Rightarrow$ About 1% (absolute) difference between data and simulation

$\Rightarrow$ Effect of acceptance ($\sim$17%) to be included

Fit and toy MC studies

1D FIT

- Perform 1D fit to $M(\mu\mu)$ using candidates in signal region $|M(\mu\mu\pi) - M_B| < 3\sigma$
- 2 components: signal + background

2D FIT

- Perform 2D fit to $M(\mu\mu)$ and $M(\mu\mu\pi)$
- 4 components: signal, $B^+ \to \pi^+\mu^+\mu^-$, $B^+ \to K^+\mu^+\mu^-$ and combinatorial bkg.

$\Rightarrow$ Perform pseudo-experiments to assess sensitivity of each strategy

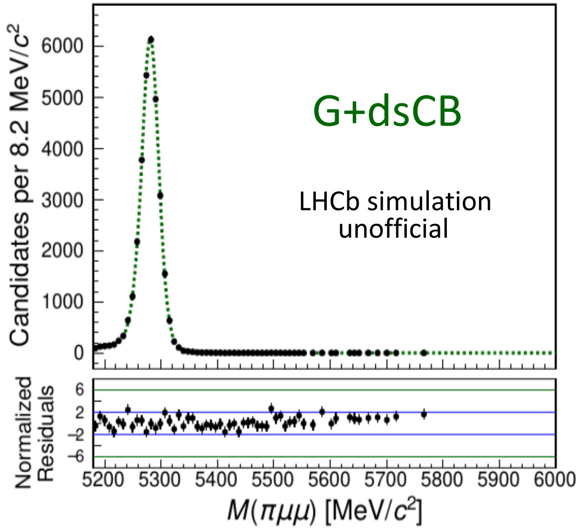
- Later: take resolution effects into account by determining global scaling factors using normalization channel

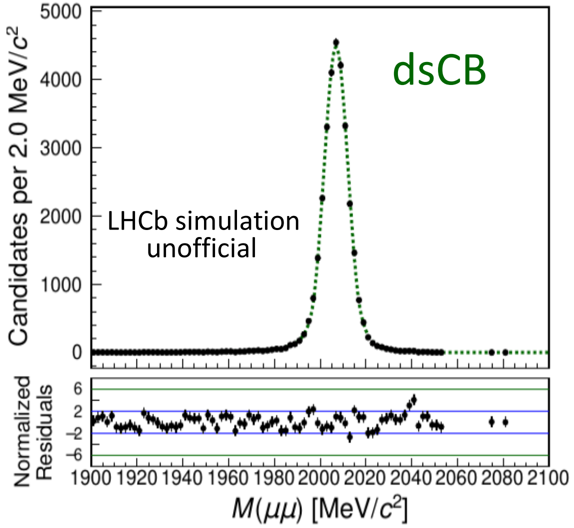
Toy experiments

- Develop 2D fit model and perform fit to sideband data to obtain realistic yields for pseudo-data sample generation
- No dependences between $M(\mu\mu\pi)$ and $M(\mu\mu)$ (as observed in MC and sideband data)
- Generate 2D pseudo-data samples from PDFs
- Fit full model (2D and 1D) including signal and backgrounds to pseudo-data sample (for 1D fit use only $M(\mu\mu\pi)$ signal region)
- Compare fit uncertainties for signal yield in each case

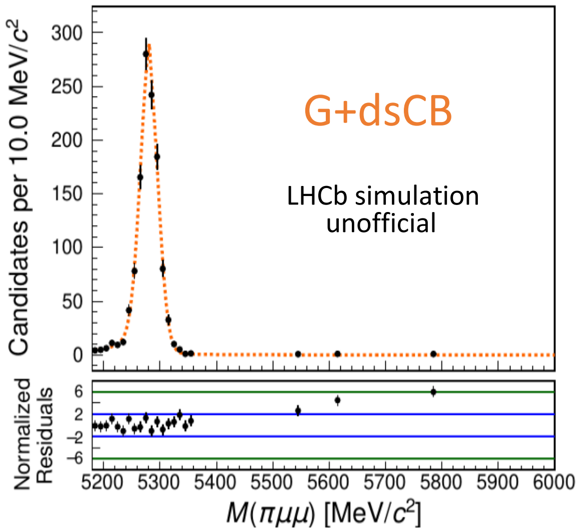
PDFs from simulation

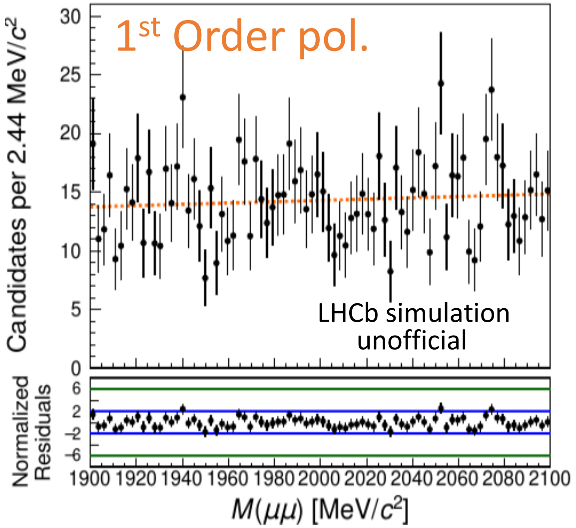
Signal



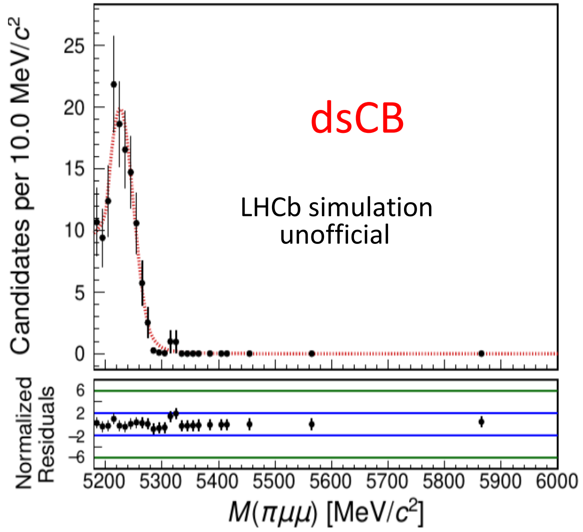


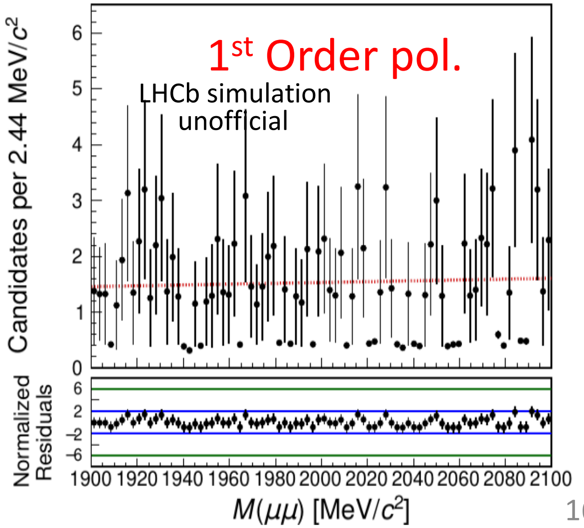
$B^+ \to \pi^+\mu^+\mu^-$



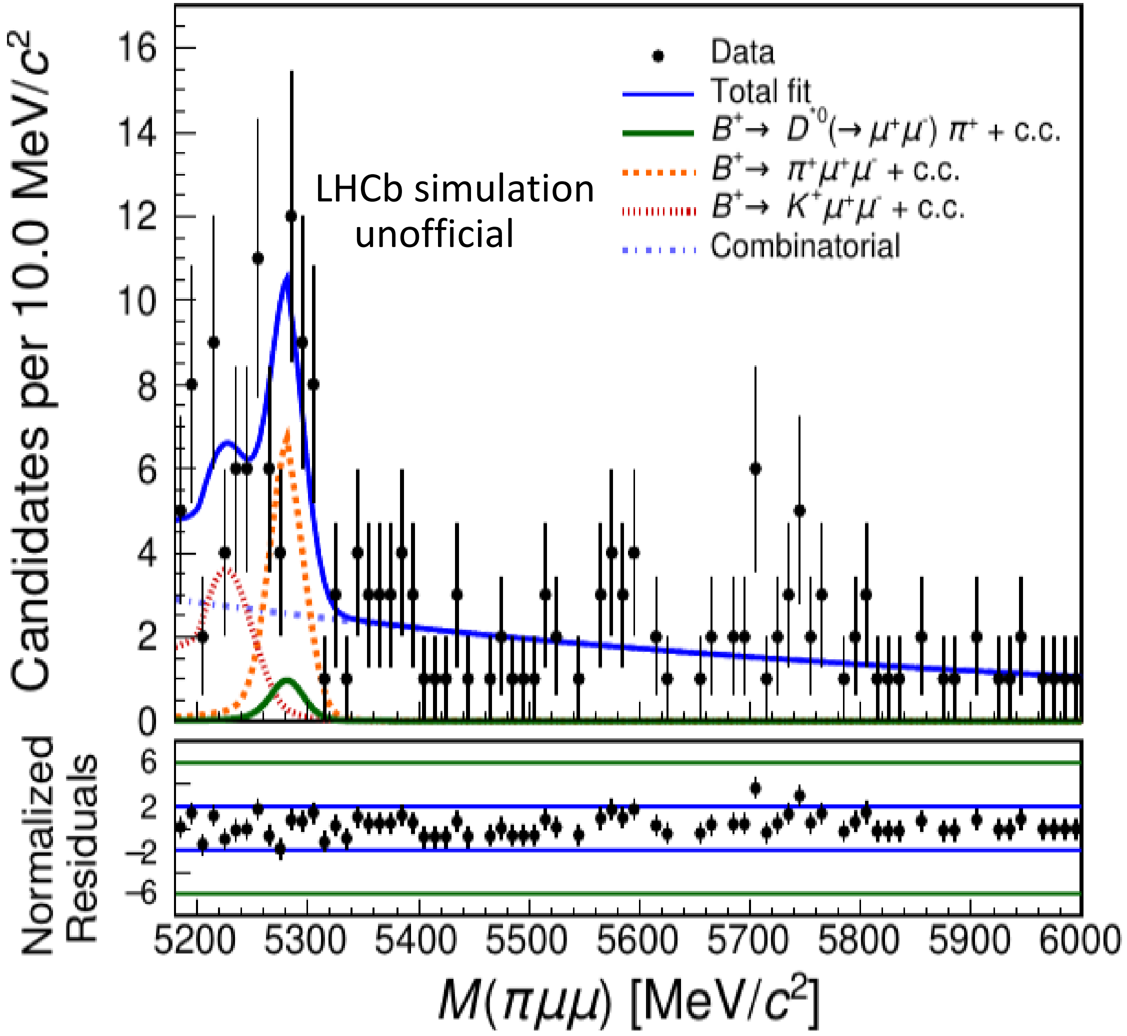


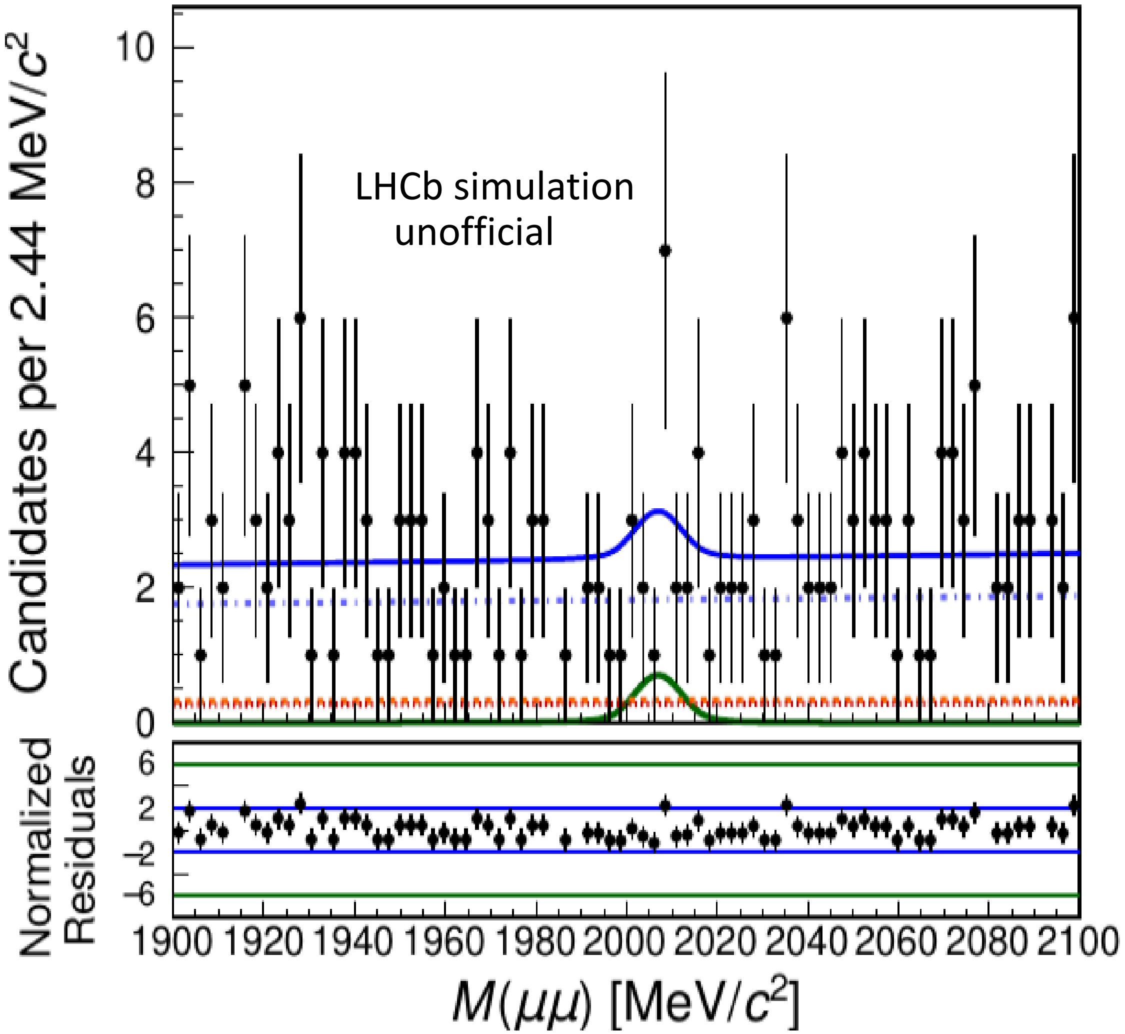
$B^+ \to K^+\mu^+\mu^-$





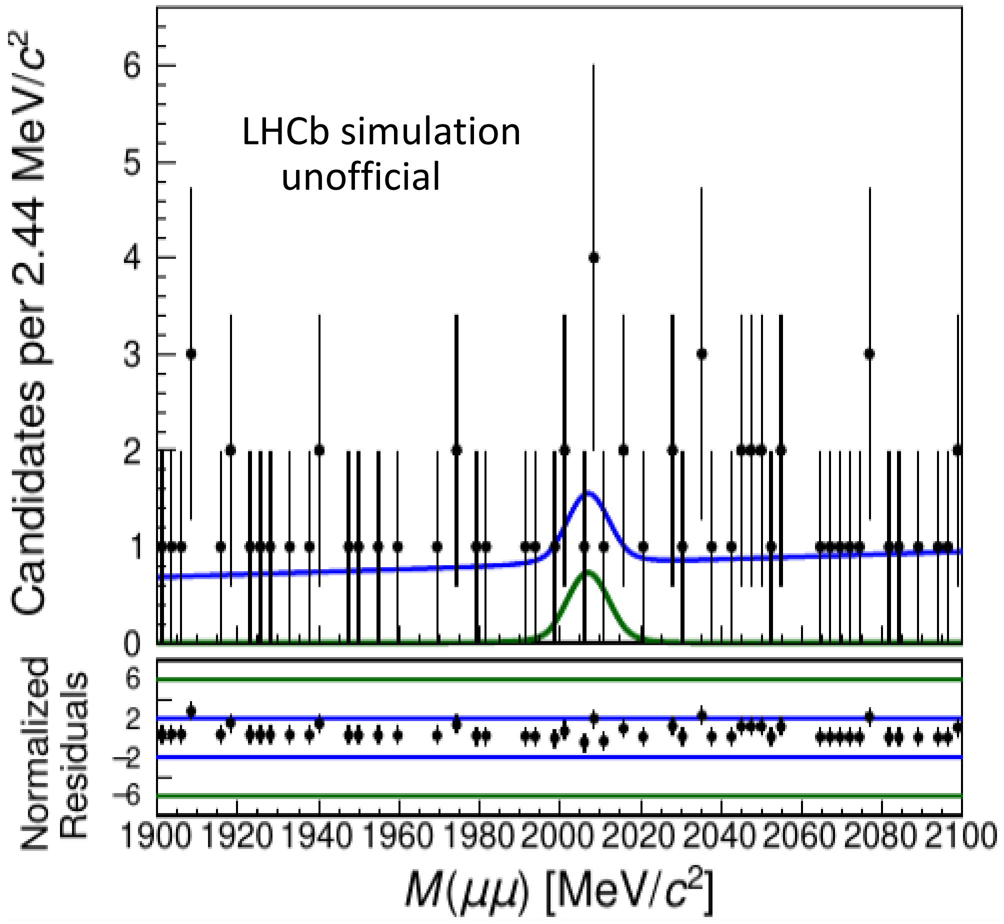
Example toy experiment (2D fit)





$\Rightarrow$ Take signal region and perform 1D fit to compare yield uncertainties

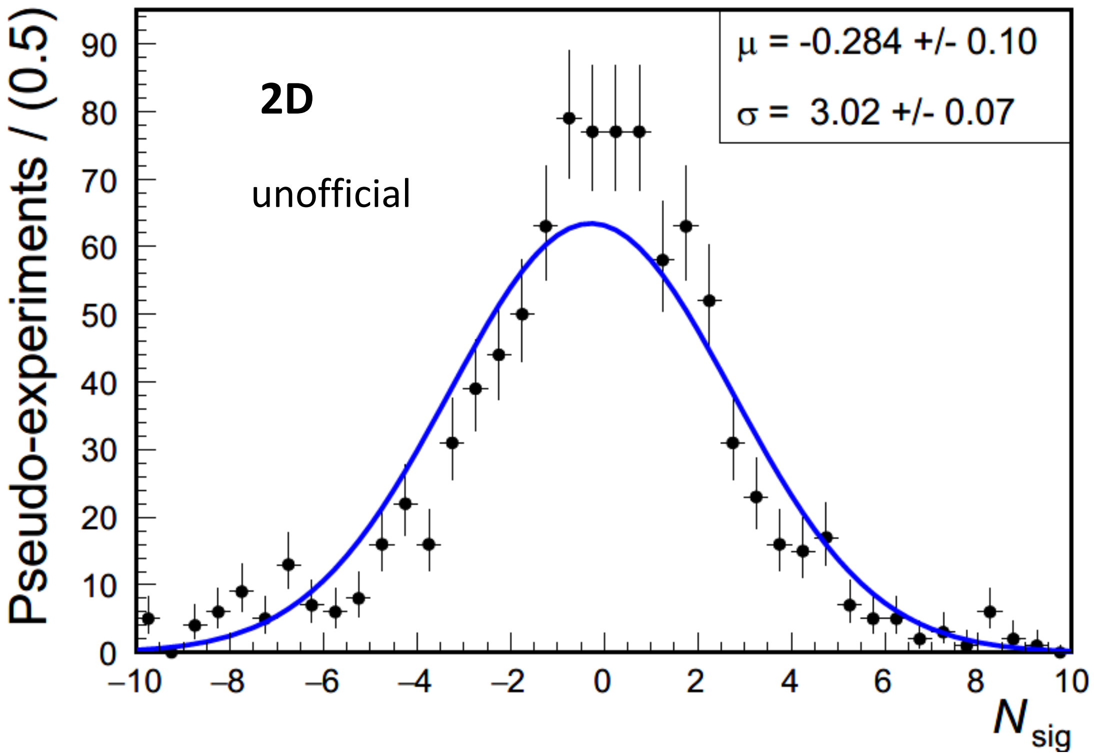
Example toy experiment (1D fit)

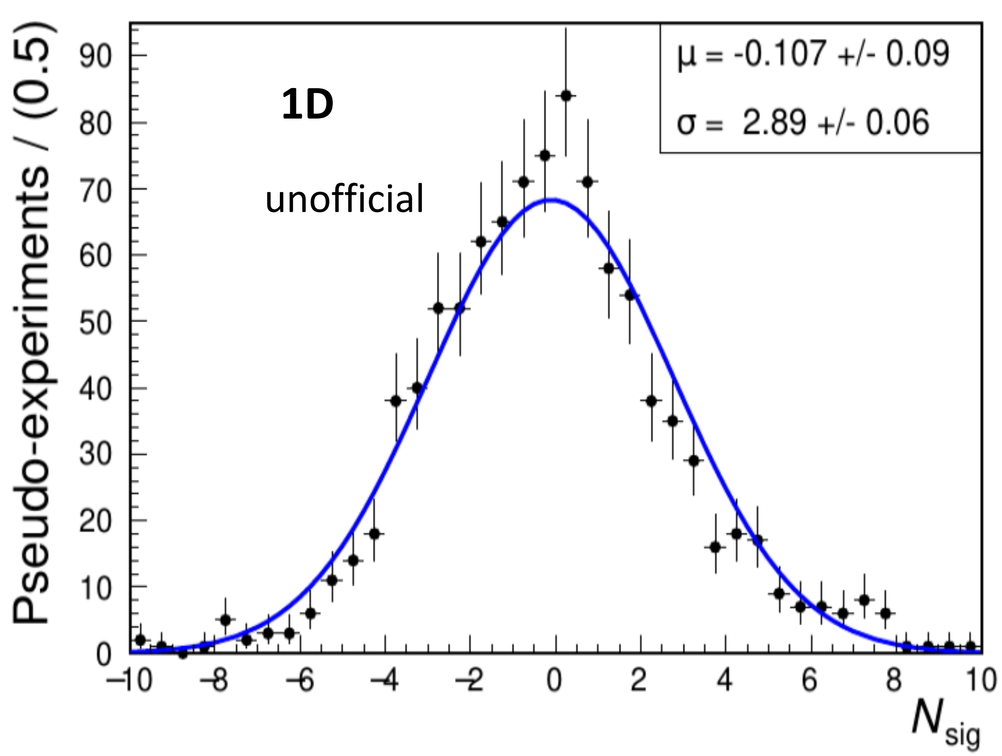


$\Rightarrow$ Compare yield uncertainties

2D vs 1D

Perform 1000 pseudo-experiments and compare distributions of residuals for signal yield





⇒ Unbiased pull distributions for all background yields in both cases (backup)

⇒ Distribution of signal yield residuals is slightly asymmetric

(probably associated with small amount of events in signal region)

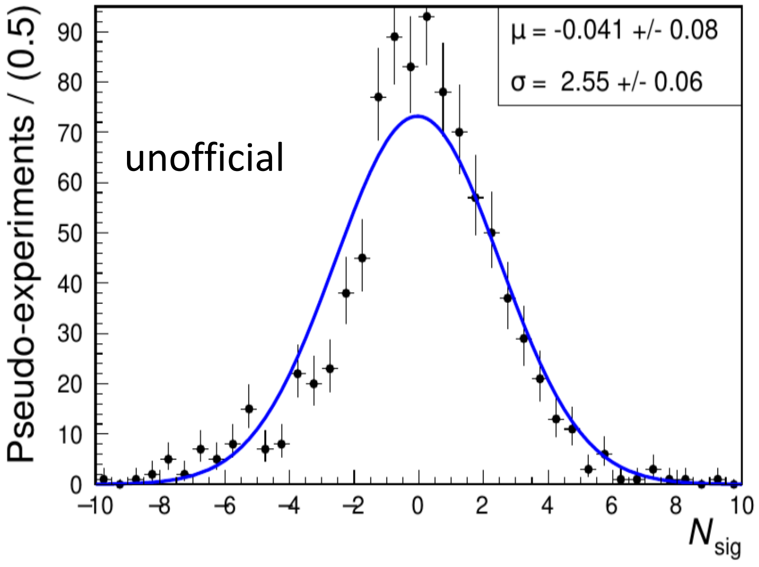
⇒ No significant difference in sensitivity

Optimize selection using toy MC

- Vary the BDT and hadron PID requirements
- Use width of signal yield distribution as figure of merit

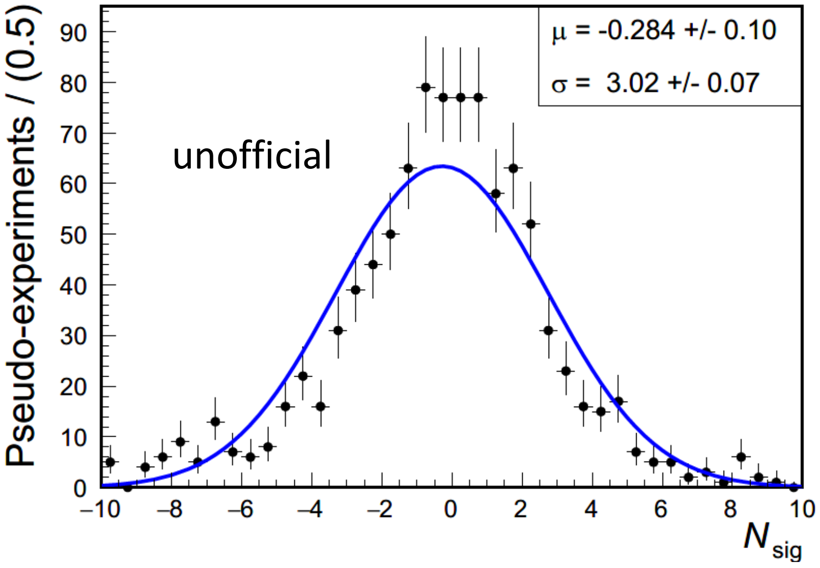
Rank 1
Best

BDT > 0.45, ProbNNpi > 0.4, ProbNNk < 0.15



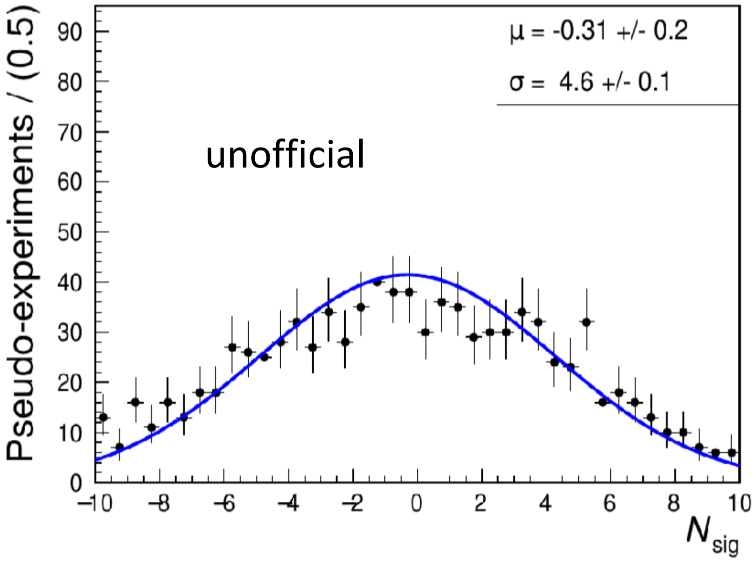
Rank 61
Default

BDT > 0.4, ProbNNpi > 0.2, ProbNNk < 0.05



Rank 216
Worst

BDT > 0.20, ProbNNpi > 0.0, ProbNNk < 0.3



$\Rightarrow$ All yield pulls unbiased

$\Rightarrow$ Will improve method by using width of $N_{\text{sig}}/\varepsilon$

Expected yields for different scenarios

	Rank 1 Best	Rank 43	Rank 58	Rank 61 Default	Rank 216 Worst
Expected yields	**BDT > 0.45 ProbNNpi > 0.4 ProbNNk < 0.15**	**BDT > 0.4 ProbNNpi > 0.4 ProbNNk < 0.15**	**BDT > 0.4 ProbNNpi > 0.2 ProbNNk < 0.15**	**BDT > 0.4 ProbNNpi > 0.2 ProbNNk < 0.05**	**BDT > 0.20 ProbNNpi > 0.0 ProbNNk < 0.3**
$B^+ \to \pi^+\mu^+\mu^-$	23	24	23	24	38
$B^+ \to K^+\mu^+\mu^-$	29	30	28	20	94
Combinatorial	80	184	196	175	1770
FOM	2.55	2.90	3.00	3.02	4.60

$\Rightarrow$ Improvement in sensitivity mainly driven by BDT selection

$\Rightarrow$ Currently exploring sensitivity of simultaneous fit in BDT bins

Summary and outlook

Paper on prospects for $\mathrm{V} \to \mu\mu$ searches:

- Sent to LHCb physics coordination
- Wish to have it on arXiv soon (to be sent then to a journal)

Ongoing $D^{*0} \to \mu\mu$ analysis:

- Most analysis tools in place
- Currently studying possible optimization of selection based on expected sensitivity
- Next step is to prepare fit for normalization channel and machinery for limit setting (most probably Feldman-Cousins method)
- Still to go through internal review (expect to have results by this summer)

Backup

Reconstruction and selection

J. Albrecht, G. Ciezarek,
C. Langenbruch, M. McCann,
M. Patel, K. Petridis, A. Shires,
T. Tekampe, M. Whitehead

Taken from current $B^+ \to \mu^+\mu^-\pi^+$ analysis:

- B2XMuMu_Line
- B-mass constraint (only for D^{*0} analysis)
- Trigger selection (backup)
- Preselection
- BDT output > 0.4
- $\mu^\pm$: ProbNNmu > 0.2, isMuon == 1
- $\pi^\pm$: isMuon ==0, ProbNNpi > 0.2, ProbNNk < 0.05
- $K^\pm$: isMuon ==0, ProbNNk > 0.4
- PID+Track Calib acceptances

Particle	Parameter	Preselection
All	PT	> 300 MeV
hadron	isMuonLoose	False
hadron	InAccMuon	True

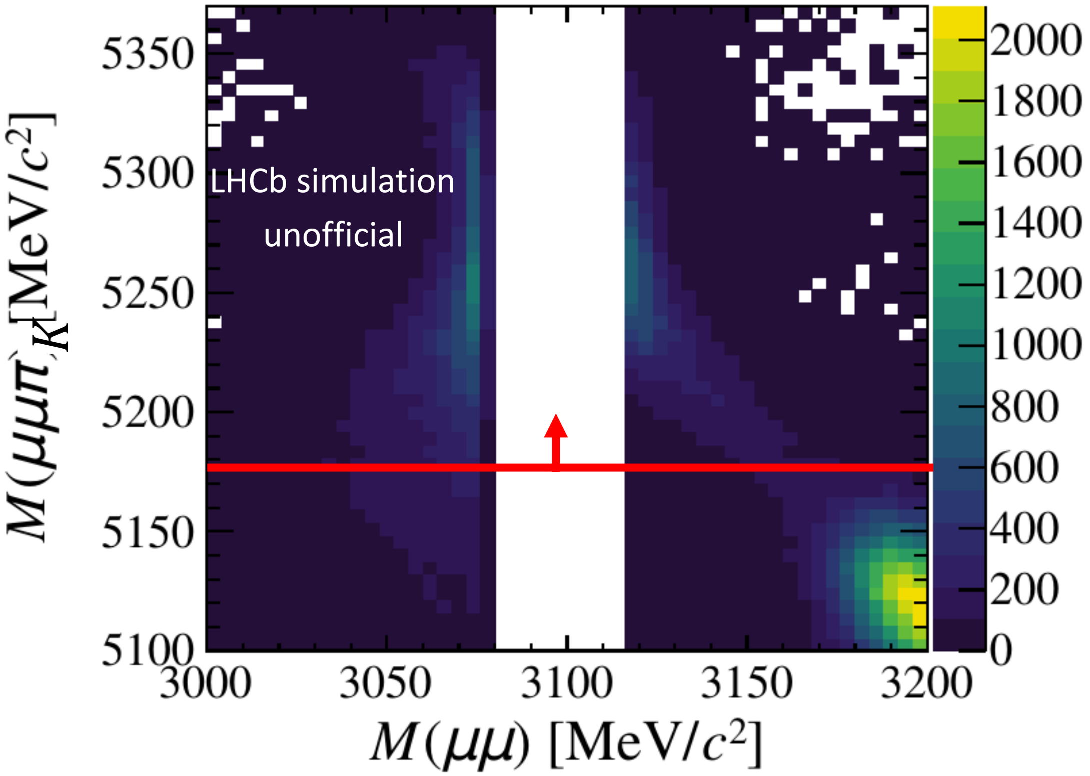
$\Rightarrow$ PID and BDT selection turn out to be already optimal (will show later)

Trigger selection

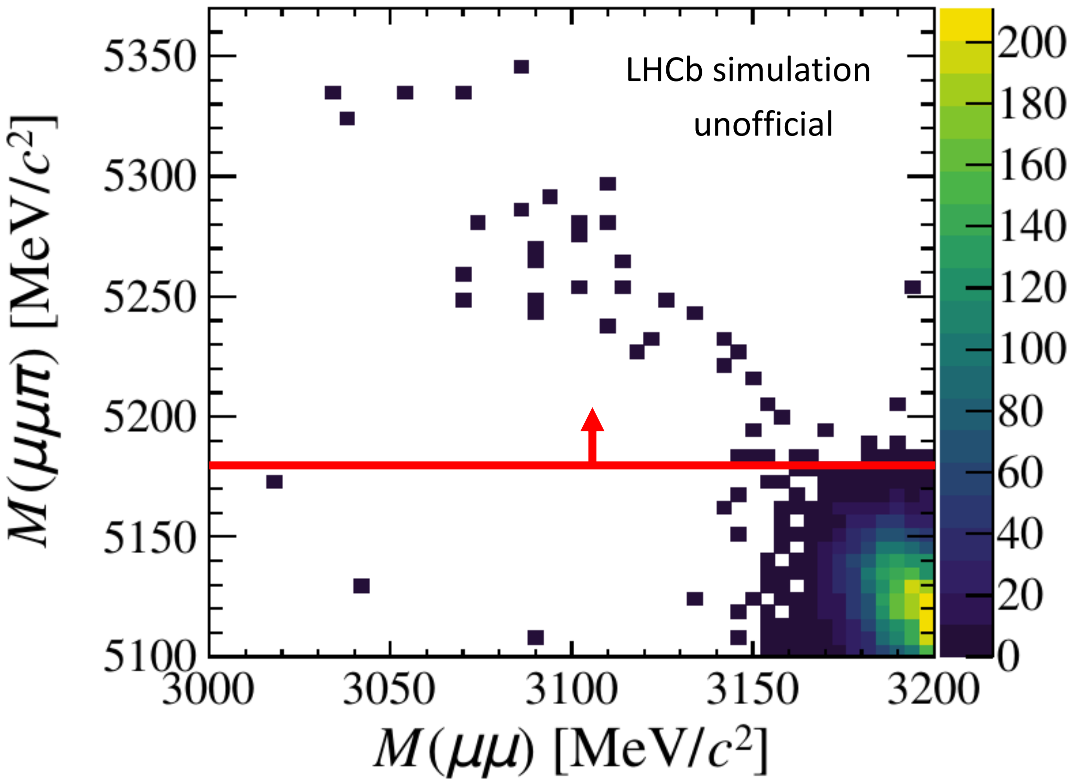
	Run I	Run II
L0	B MuonDecision TOS	B MuonDecision TOS
HLT1	B Hlt1TrackAllL0Decision TOS, or B Hlt1TrackMuonDecision TOS	B Hlt1TrackMVADecision TOS
HLT2	B Hlt2Topo2BodyBBDTDecision TOS, or B Hlt2Topo3BodyBBDTDecision TOS, or B Hlt2TopoMu2BodyBBDTDecision TOS, or B Hlt2TopoMu3BodyBBDTDecision TOS, or B Hlt2DiMuonDetachedDecision TOS	B Hlt2Topo2BodyDecision TOS, or B Hlt2Topo3BodyDecision TOS, or B Hlt2TopoMu2BodyDecision TOS, or B Hlt2TopoMu3BodyDecision TOS, or B Hlt2DiMuonDetachedDecision TOS

2D distributions for normalization channel

$B^+ \to J/\psi\, K^+$ data
$M(\mu\mu)$ from DTF with B-mass constraint



$B^0 \to J/\psi K^{*0}$ simulation
$M(\mu\mu)$ from DTF with B-mass constraint

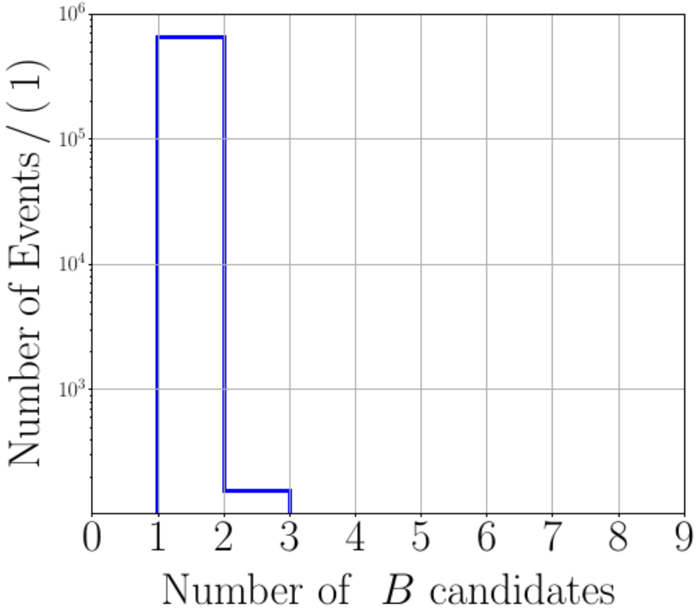


$\Rightarrow$ Bkg. from partially reconstructed decays mitigated by requiring $M(\mu\mu\pi) > 5180$ MeV/c^2
$\Rightarrow$ Removes only about 3% of signal (due to tail)

Candidate multiplicity

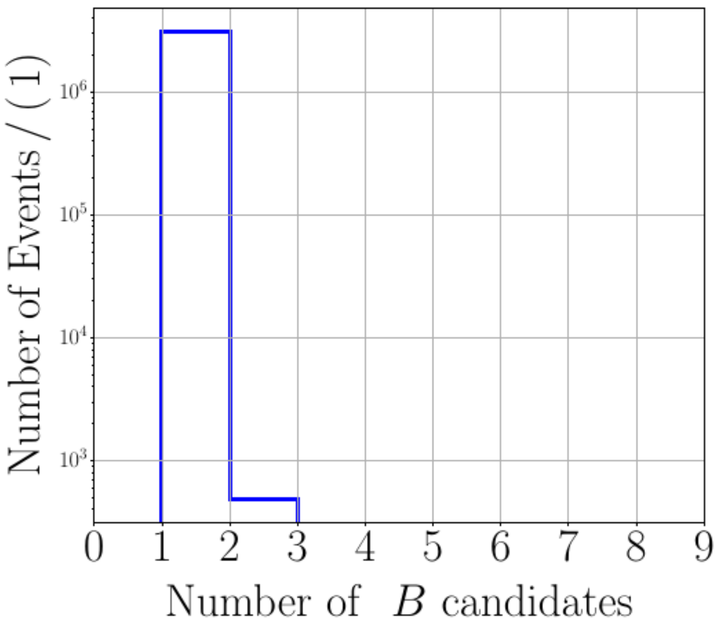
Check multiplicity after full selection (with $M(\mu\mu\pi)$ restriction)

$\boldsymbol{B^+ \to \overline{D}^{*0}\, \pi^+}$



Avg. mpty = 1.000173 +- 0.000018

$\boldsymbol{B^+ \to J/\psi\, K^+}$



Avg. mpty = 1.000162 +- 0.000007

$\Rightarrow$ No need for best candidate selection.

Efficiencies in simulation (uncorrected)

Signal: bkg. cats. 10 (or 0) and 50.

$$\varepsilon = \frac{\text{Signal candidates after full selection}}{\text{Entries in MCDecayTreeTuple}}$$

$\boldsymbol{B^+ \to \bar{D}^{*0}\, \pi^+}$

Year\Polarity	Up	Down
2011	9.58 ± 0.04	9.65 ± 0.04
2012	8.41 ± 0.04	8.28 ± 0.04
2015	6.55 ± 0.03	6.59 ± 0.03
2016	8.70 ± 0.04	8.70 ± 0.04
2017	9.48 ± 0.04	9.50 ± 0.04
2018	8.92 ± 0.04	8.95 ± 0.04

$\boldsymbol{B^+ \to J/\psi\, K^+}$

Year\Polarity	Up	Down
2011	11.95 ± 0.03	12.00 ± 0.03
2012	10.81 ± 0.03	10.78 ± 0.03
2015	9.23 ± 0.03	9.27 ± 0.03
2016	12.32 ± 0.01	12.34 ± 0.01
2017	13.06 ± 0.02	13.10 ± 0.02
2018	12.52 ± 0.02	12.52 ± 0.02

$\Rightarrow$ No large difference between polarities, but between years.

$\Rightarrow$ Effect of acceptance ($\sim$17%) to be included.

Muon ID study

Muon ID requirement (isMuon) used in B2XMuMu stripping line

⇒ Obtain efficiency of requirement in data using PID calib

⇒ Obtain efficiency of requirement in MC using dedicated ntuples

⇒ Obtain correction for each muon $w = \frac{\varepsilon_{\mathrm{Data}}}{\varepsilon_{\mathrm{MC}}}$

⇒ $\varepsilon_i^{\mu\mathrm{ID}} = w(\mu_i^+) \cdot w\,(\mu_i^-)$ (very close to 1)

Trigger efficiency

Use TISTOS method for L0Muon eff corrections (inspired by $\Lambda_b^0 \to J/\psi\Lambda$ analysis)

⇒ Using $B^+ \to J/\psi\, K^+$ sample (in p_{t} bins)

⇒ Very close to 1

p_t and track multiplicity

- Obtain background-subtracted data (sWeight method)
- Corrections determined using BDT trained using difference between data and MC (in 2D bins of p_t and track multiplicity)

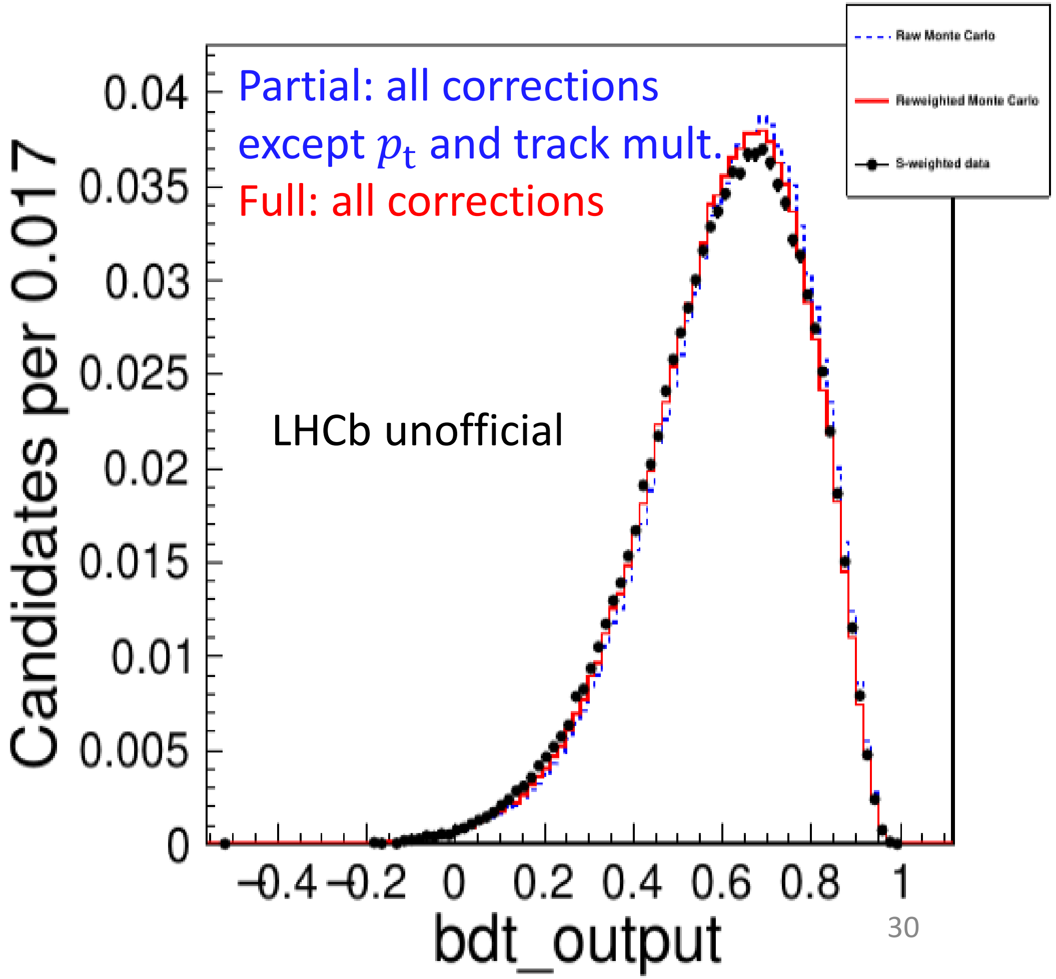
Closure test

Compare partially/fully corrected signal MC with **signal data** (sweighted)

Checked distributions for kinematic observables and BDT inputs/output

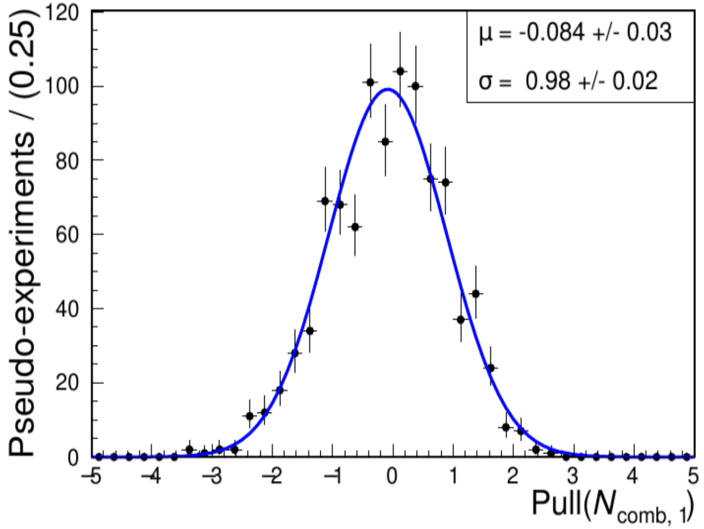
$\Rightarrow$ Good overall data/MC agreement

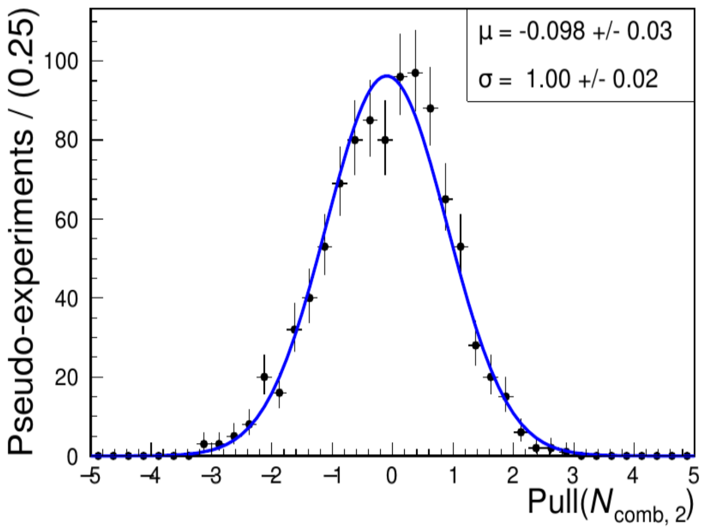
$\Rightarrow$ Ready for all years and polarizations

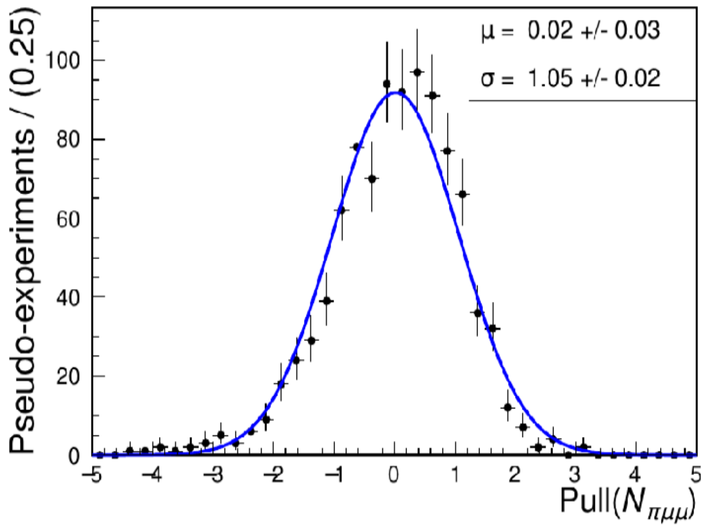


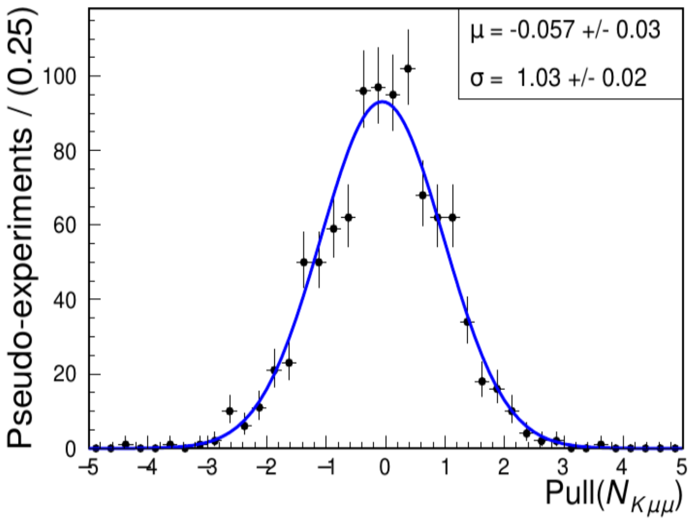
Pulls for backgrounds

Perform 1000 pseudo-experiments and obtain pulls









$\Rightarrow$ All looking good.