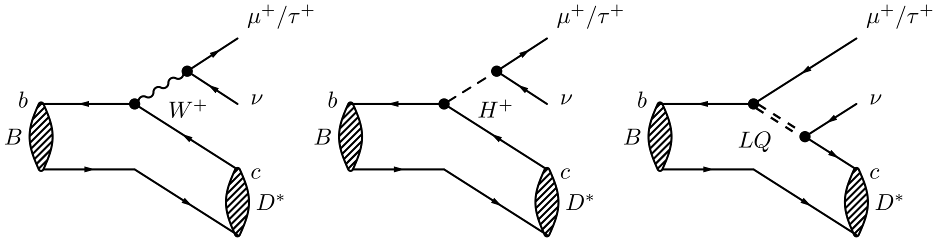
$B \to X\tau\nu$ measurements at LHCb

Greg Ciezarek,
on behalf of the LHCb collaboration

April 10, 2018

$B \to D^{(*)}\tau\nu$



- In the Standard model, the only difference between $B \to D^{(*)}\tau\nu$ and $B \to D^{(*)}\mu\nu$ is the mass of the lepton
    - Form factors mostly cancel in the ratio of rates (except helicity suppressed amplitude)
- Ratio R($D^{(*)}$) $= \mathcal{B}(B \to D^{(*)}\tau\nu)$ / $\mathcal{B}(B \to D^{(*)}\mu\nu)$ is sensitive to e.g charged Higgs, leptoquark

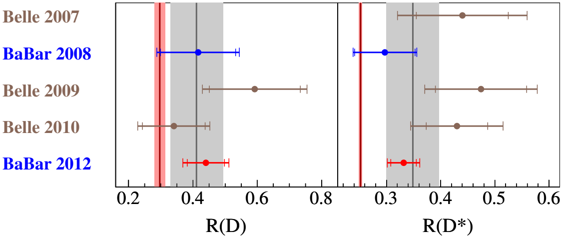
History

Phys. Rev. D. (2010) 88 072012

Phys. Rev. Lett. 99 (2007) 191807

arxiv:0910.4301

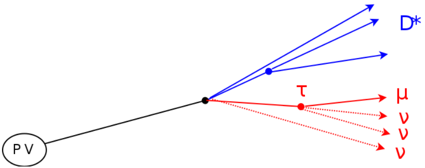
Phys. Rev. D. (2010) 82 072005



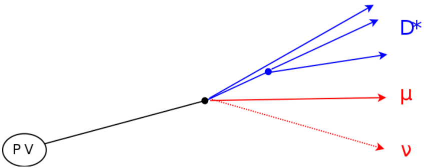
- How this started: measurements from B factories in $\tau \to \ell\nu\nu$ channel
  - Final measurement from BaBar (Phys. Rev. D. 88 072012) claimed 3 σ excess over SM expectation
  - Status at the time of the Babar measurement

Experimental challenge

$B \rightarrow D^*\tau\nu$

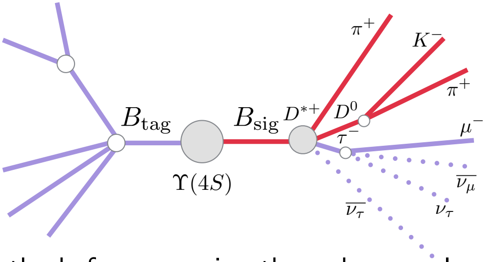


$B \rightarrow D^*\mu\nu$



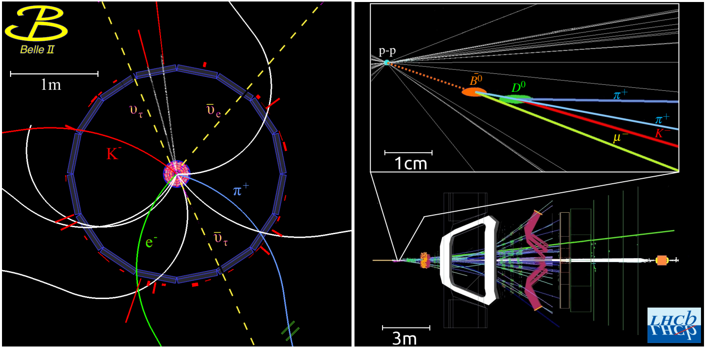
- Difficulty: neutrinos - 2 for $(\tau \rightarrow \pi\pi\pi\nu)\nu$, 3 for $(\tau \rightarrow \mu\nu\nu)\nu$
  - No narrow peak to fit (in any distribution)
- Main backgrounds: partially reconstructed B decays
  - $B \rightarrow D^*\mu\nu$,$B \rightarrow D^{**}\mu\nu$, $B \rightarrow D^*D(\rightarrow \mu X)X$ ...
  - $B \rightarrow D^*\pi\pi\pi X$, $B \rightarrow D^*D(\rightarrow \pi\pi\pi X)X$ ...
- Also combinatorial, misidentified background

B Factory method



- Traditional methods for measuring these decays rely on $e^+e^- \to B\overline{B}$ event properties
  - Centre of mass fixed
  - Nothing else produced in event
- "Tag reconstruction"
  - Fully reconstruct other B → measurement of signal B kinematics
  - Signal B + other B should be entire event → strong rejection against other missing reconstructable particles
- Penalty: sub percent efficiency

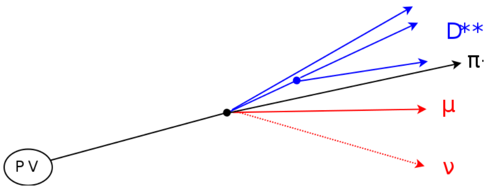
Can you do this at a hadron collider?

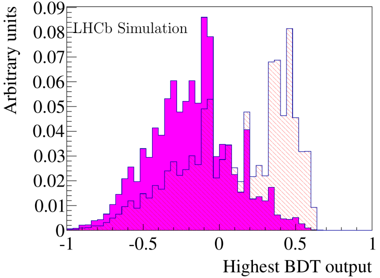


- In a hadron collider the $B\overline{B}$ centre of mass isn't fixed $\rightarrow$ rest of event provides little constraint on the signal B kinematics
    - Event also contains a lot of junk from the proton-proton interaction $\rightarrow$ reconstructing the whole event is meaningless
- Needed completely different methods

Isolation

Phys. Rev. Lett. 115 (2015) 111803

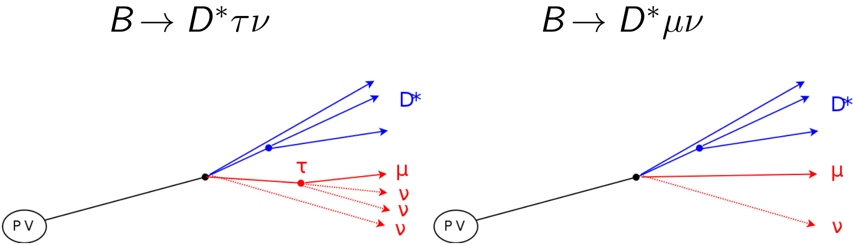




- Reject physics backgrounds with additional charged tracks
- MVA output distribution for $B \to D^{**}\mu^+\nu$ background (hatched) and signal (solid)
- Inverting the cut gives a sample hugely enriched in background $\to$ control samples

Fit strategy

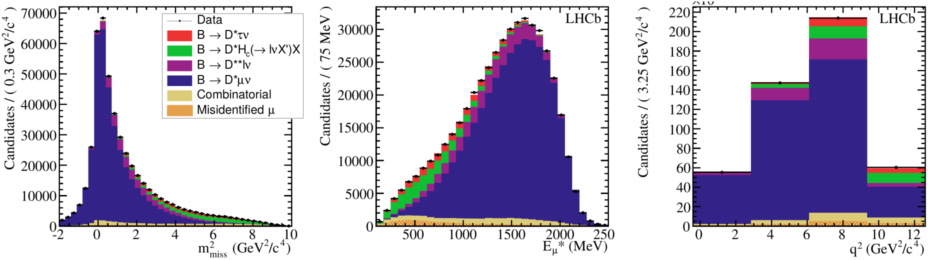
Phys. Rev. Lett. 115 (2015) 111803



- Can use B flight direction to measure transverse component of missing momentum
- No way of measuring longitudinal component $\to$ use approximation to access rest frame kinematics
  - Assume $\gamma\beta_{z,visible} = \gamma\beta_{z,total}$
  - $\sim$20% resolution on B momentum, long tail on high side
- Can then calculate rest frame quantities - $m^2_{missing}$, E_μ, q^2

Fit strategy

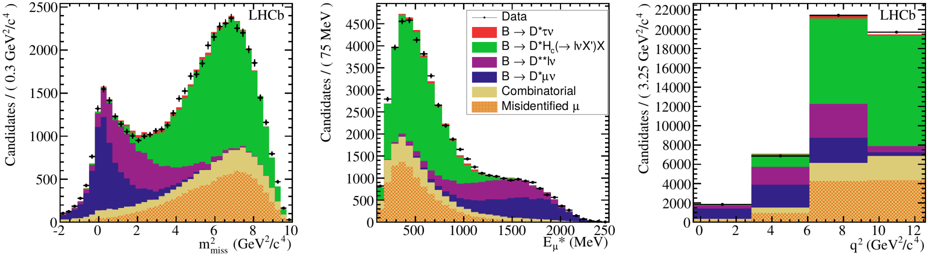
Phys. Rev. Lett. 115 (2015) 111803



- Three dimesional template fit in E_μ (left), $m^2_{missing}$ (middle), and q^2
  - Projections of fit to isolated data shown
- All uncertainties on template shapes incorporated in fit:
  - Continuous variation in e.g different form factor parameters

Background strategy

Phys. Rev. Lett. 115 (2015) 111803

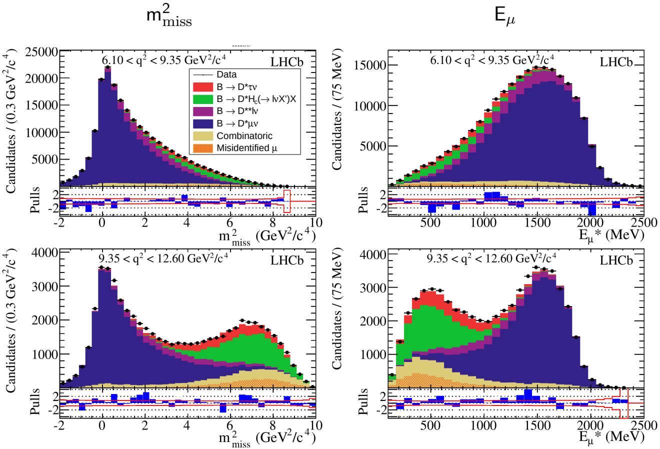


- All major backgrounds modelled using control samples in data
  - Dedicated samples for different backgrounds ($D^*\pi$, $D^*\pi\pi$, D^*DX
  - Quality of fit used to justify modelling
  - Data-driven systematic uncertainties
- All combinatorial or misidentified backgrounds taken from data
- More details on everything in backups

Signal fit

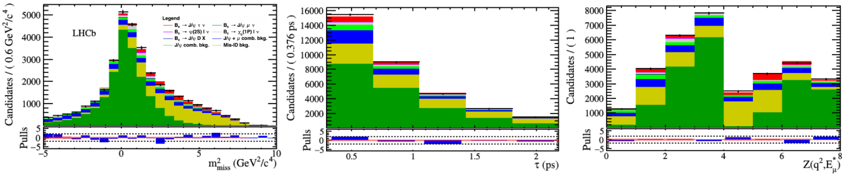
$9.35 < q^2 < 12.60\,\mathrm{GeV}^2/c^4$

$9.35 < q^2 < 12.60\,\mathrm{GeV}^2/c^4$



- Fit to isolated data, used to determine ratio of $B \to D^*\tau\nu$ and $B \to D^*\mu\nu$
- Model fits data well
- We measure $\mathcal{R}(D^*) = 0.336 \pm 0.027 \pm 0.030$, consistent with SM at 2.1σ level
  - Phys. Rev. Lett. 115 (2015) 111803(Run 1 data)

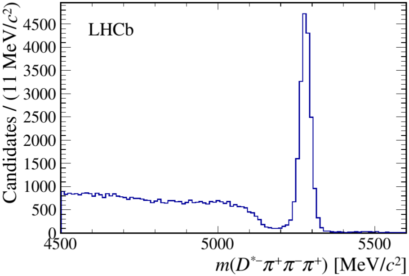
$B_c \to J/\psi\, \tau\nu$



- $R_{J/\psi} \equiv B_c \to J/\psi\, \tau\nu / B_c \to J/\psi\, \mu\nu$
- Measured using very similar techniques to $\mathcal{R}(D^*)$, on run 1 data
- $R_{J/\psi} = 0.71 \pm 0.17 \pm 0.18$
  - $\sim 2\sigma$ from SM
  - But nearly as far from consistency with $\mathcal{R}(D^*)$
- LHCb-PAPER-2017-035(Run 1 data)

$\mathcal{R}(D^*)$ with $\tau \to \pi\pi\pi\nu$

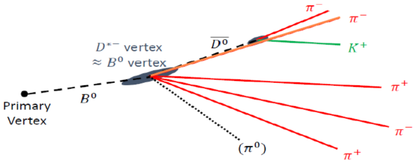
LHCb-PAPER-2017-017, LHCb-PAPER-2017-027

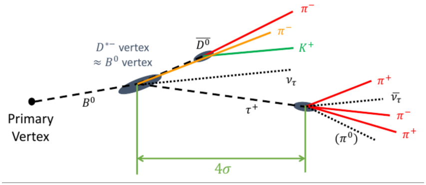


- Compared to muonic $\mathcal{R}(D^*)$:
  - Large $B \to D^*\mu\nu$, $B \to D^{**}\mu^+\nu$ backgrounds absent
  - Additional $B \to D^*\pi\pi\pi X$ backgrounds
  - $B \to D^*DX$ with $D \to \pi\pi\pi X$
- Control experimental efficiencies by measuring rate relative to $B \to D^*\pi\pi\pi$

Removing $B \to D^*\pi\pi\pi X$

LHCb-PAPER-2017-017, LHCb-PAPER-2017-027

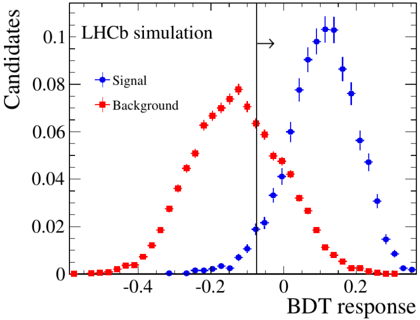




- Can use decay topology to remove direct $B \to D^*\pi\pi\pi X$ decays:
- If the $\pi\pi\pi$ vertex is displaced from the B vertex, cannot be direct $B \to D^*\pi\pi\pi X$
- Can remove a large, poorly measured background
  - And control the remainder
- $B \to D^*DX$ major physics background remaining

Dealing with $B \to D^* DX$

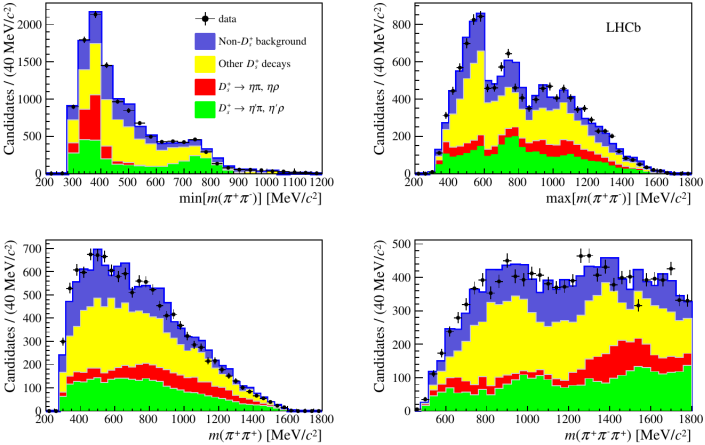
LHCb-PAPER-2017-017, LHCb-PAPER-2017-027



- $[\pi\pi\pi]$ lifetime discriminates between tau and $B \to D^* DX$
- Can use partial reconstruction techniques to reconstruct D peak in $B \to D^{*+} D$ (not $B \to D^* DX$)
- $\tau \to \pi\pi\pi\nu$ is mostly $a1(1260)$, $D \to \pi\pi\pi X$ mostly isn't
  - Use the $\pi\pi\pi$ (sub) structure to separate $B \to D^* \tau\nu$ from $B \to D^* DX$
  - Shown: control region for $D_s \to \pi\pi\pi X$
- Put everything in an MVA: kinematics, Dalitz, partial reconstruction,

$D \to \pi\pi\pi X$

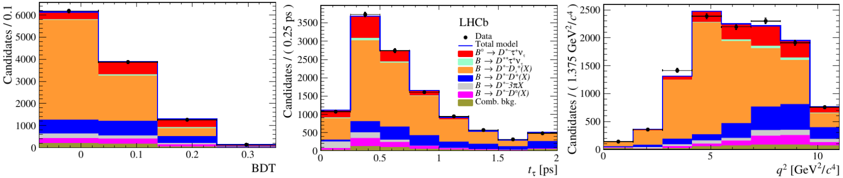
LHCb-PAPER-2017-017, LHCb-PAPER-2017-027



- Again, use data to control background modelling
- Use low BDT region to control $D_s \to \pi\pi\pi X$ substructure

Fit

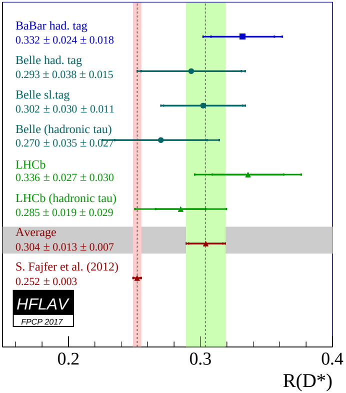
LHCb-PAPER-2017-017, LHCb-PAPER-2017-027



- 3D template fit in BDT,q^2, tau lifetime to determine signal yield

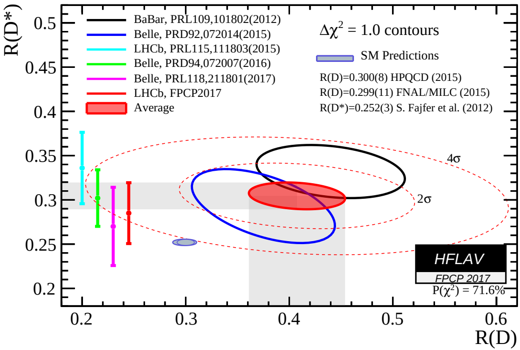
Result

LHCb-PAPER-2017-017, LHCb-PAPER-2017-027



- Result equally compatible with SM, world average
- More precise than our past result (still only run 1 data)
- New average gives a slightly lower value, but higher precision → significance increases very, very slightly
- LHCb-PAPER-2017-017, LHCb-PAPER-2017-027(Run 1 data)

Where do we stand?



- Official HFLAV combination of $\mathcal{R}(D)$and $\mathcal{R}(D^*)$
- Excellent consistency between results
- Combined: **4.1σ tension with SM**
  - (Before considering more conservative $B \to D^* \tau \nu$ form factors..))

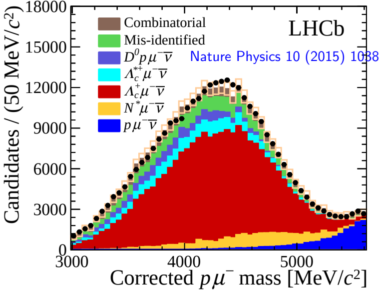
Where next?

- Next step from muonic $\mathcal{R}(D^*)$: $D^0\mu X$ vs $D^{*+}\mu X$
  - Backgrounds not so much worse than in $D^{*+}\mu X$
  - Significant improvement in precision
- Ongoing:$B_s \to D_s^{(*)}\tau\nu$
  - Similar situation to $\mathcal{R}(D^{(*)})$
  - Main difference to $B \to D^{(*)}\tau\nu$: feed-down mostly via neutrals

Where next?

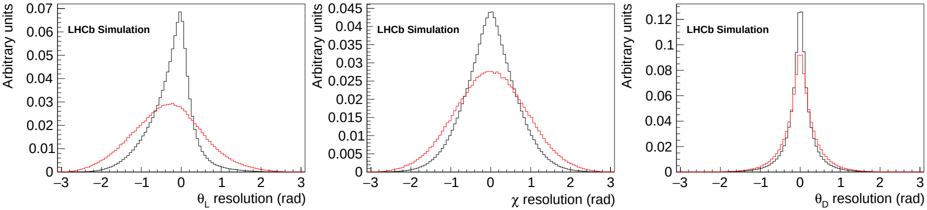
- Ongoing: $\Lambda_b \to \Lambda_c^{(*)}\tau\nu$
  - Different spin structure to meson modes $\to$ different physics sensitivity
  - In particular, would help discriminate tensor contributions
- Potential: $B \to D^{**}\tau\nu$
  - Samples of $D^{**}\mu X$ not so small: control sample for $\mathcal{R}(D^*)$ measurement shown
  - To interpret results, need to split measurements between different D^{**} states
  - More work needed first on $B \to D^{**}\mu\nu$ modes

$b \to u\tau\nu$



- If we establish a new physics signal in $b \to c\tau\nu$, would really want to test the flavour structure: $b \to u\tau\nu$
  - $b \to c\tau\nu$ hard enough to measure, before extra suppression $\to$ background levels challenging
  - Requires very careful choice of channel to give us any hope
- $B \to p\bar{p}\tau\nu$ with $\tau \to \mu\nu\nu$
  - Experimentally the cleanest, Theoretically not so good...
  - Will make detailed measurements of corresponding $B \to p\bar{p}\mu\nu$ mode
- $\Lambda_b \to p\tau\nu$ with $\tau \to \pi\pi\pi\nu$?
  - Lattice calculations used to measure $|V_{\rm ub}|$ with equivalent $\Lambda_b \to p\mu\nu$ mode $\to$ already have a good theory prediction

Angular resolutions for $B \to D^* \tau \nu$



- Angular resolution for $B \to D^* \mu \nu$, $B \to D^* \tau \nu$ $(\tau \to \mu \nu \nu)$
- Tau decay results in loss of information
  - θ_ℓ and χ degraded, θ_D a bit less
- These resolutions aren't horrific $\to$ we can make a measurement (with unknown sensitivity)
- These resolutions aren't insignificant $\to$ needs massive care

What can we do?

- Unfolding this seems a nightmare (as does background subtraction) $\rightarrow$ we are unlikely to publish corrected q^2 / angular distributions for signal
- But we can fit the data
  - Templates we fit already include effects of resolution, acceptance ...

What to measure

- First need to see if the excess holds up!
- Afterwards:
  - Does measured value change allowing NP operators?
  - Can enhancement be accommodated by theory uncertainty?
  - Pure vector/axial/tensor/...?
  - Or a combination of operators?
  - Can we fit the full matrix element?

Scalar form factor

- Trying to measure (pseudo)scalar form factor directly from $B \to D^{(*)}\tau\nu$ doesn't seem so implausible
  - If no new (pseudo)scalar physics, and form factor agrees with prediction $\to$ model independent SM exclusion
  - Uncertainty from QED corrections?
- Testing SM only hypothesis $\to$ constrain other form factors from $B \to D^{(*)}\mu\nu$
- Not yet sure when we become sensitive enough

Tau polarisation?

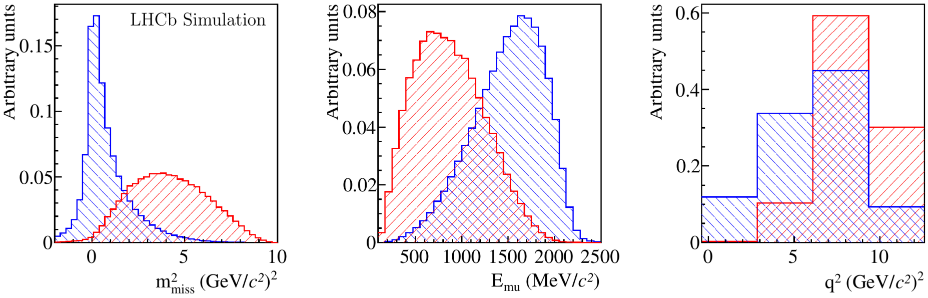
- With $\tau \to \mu\nu\nu$:
  - Some sensitivity to polarisation, but probably can't disentangle from angular distribution?
- With $\tau \to \pi\pi\pi\nu$:
  - Combined $\pi\pi\pi$ momentum has little sensitvity to polarisation
  - But some information in substructure $\to$ exploring this
  - Thesis of Laurent Duflot (LAL 93-09)
- Measurement of polarisation and angular information correlated
- Physics of polarisation and angular information correlated
- We should consider both together

Conclusion

- World average for $\mathcal{R}(D^{(*)})$still in tension with SM
- LHCb has established techniques to measure $B \to X_c \tau \nu$ with both $\tau \to \mu\nu\nu$ and $\tau \to \pi\pi\pi\nu$
  - Relatively independent systematics, important as precision improves
- Wide program underway with a full range of charm hadrons
- Plans for how to go beyond branching fractions
  - Overlaps with measurements in $B \to X_c \mu\nu$
- Lots to look forward to

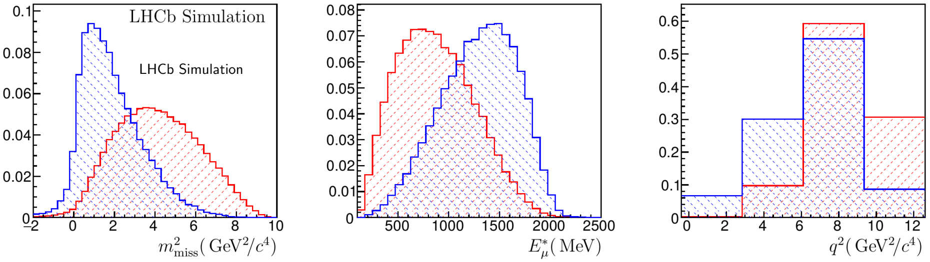
Backups

$B \to D^* \mu\nu$



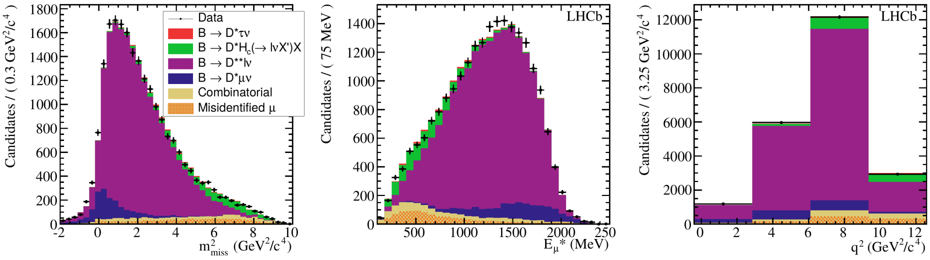
- $B \to D^* \mu\nu$ (black) vs $B \to D^* \tau\nu$ (red)
- $B \to D^* \mu\nu$ is both the normalisation mode, and the highest rate background ($\sim 20 \times B \to D^* \tau\nu$)
  - Use CLN parameterisation for form factors
  - Float form factors parameters in fit $\to$ uncertainty taken into account

$B \to D^{**}\mu^+\nu$



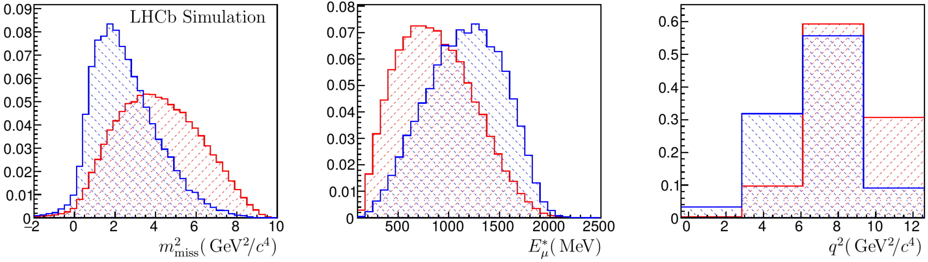
- $B \to D^{**}\mu^+\nu$ refers to any higher charm resonances (or non resonant hadronic modes)
- Not so well measured
  - Set of states comprising D^{**} known to be incomplete
  - Decay models not well measured
- For the established states (shown in black):
  - Separate components for each resonance (D_1, D_2^*, D_1')
  - Use LLSW model (Phys. Rev. D. (1997) 57 307), float slope of Isgur-wise function

$B \to D^{**}(\to D^{*+}\pi)\mu\nu$ control sample



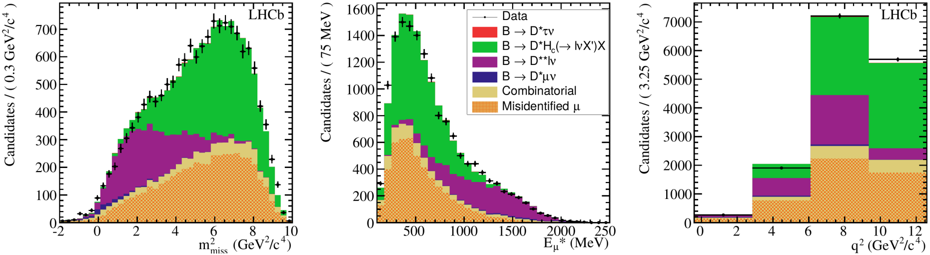
- Isolation MVA selects one track, $M_{D^{*+}\pi}$ around narrow D^{**} peak $\to$ select a sample enhanced in $B \to D^{**}\mu^+\nu$
  - Use this to constrain, justify $B \to D^{**}\mu^+\nu$ shape for light D^{**} states
  - Also fit above, below narrow D^{**} peak region to check all regions of $M_{D^{*+}\pi}$ are modelled correctly in data

Higher $B \to D^{**}\mu^{+}\nu$ states



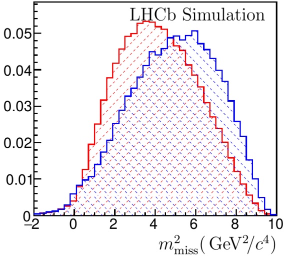
- Previously unmeasured $B \to D^{**}(\to D^{*+}\pi\pi)\mu\nu$ contributions recently measured by BaBar
  - Too little data to separate individual (non)resonant components
  - Single fit component, empirical treatment
- Constrain based on a control sample in data
  - Degrees of freedom considered: D^{**} mass spectrum, q^2 distribution
  - Effect of D^{**} mass spectrum negligible

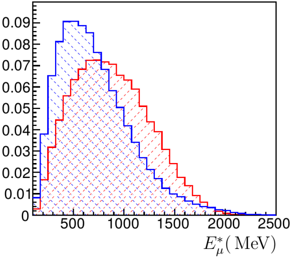
$B \to D^{**}(\to D^{*+}\pi\pi)\mu\nu$ control sample

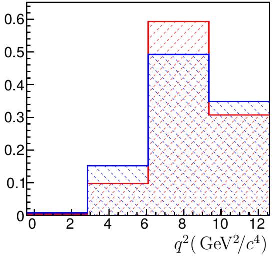


- Also look for two tracks with isolation MVA → study $B \to D^{**}(\to D^{*+}\pi\pi)\mu\nu$ in data
- Can control shape of this background

$B \to D^* DX$

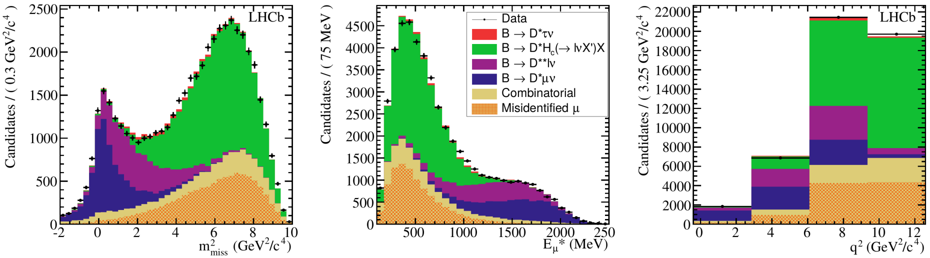






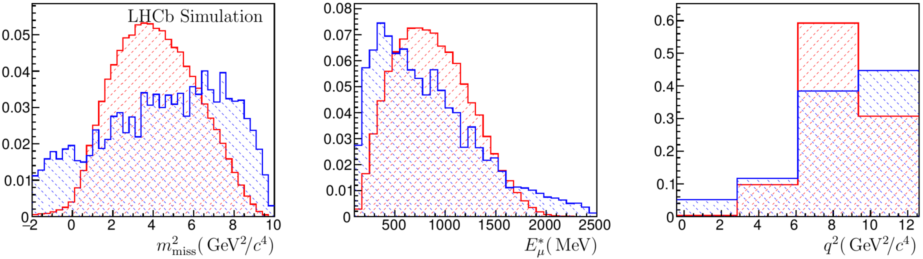
- $B \to D^* DX$ consists of a very large number of decay modes
  - Physics models for many modes not well established
- Constrain based on a control sample in data
- Single component, empirical treatment
  - Consider variations in M_{DD}
  - Multiply simulated distributions by second order polynomials
  - Parameters determined from data

$B \to D^* D X$ control sample



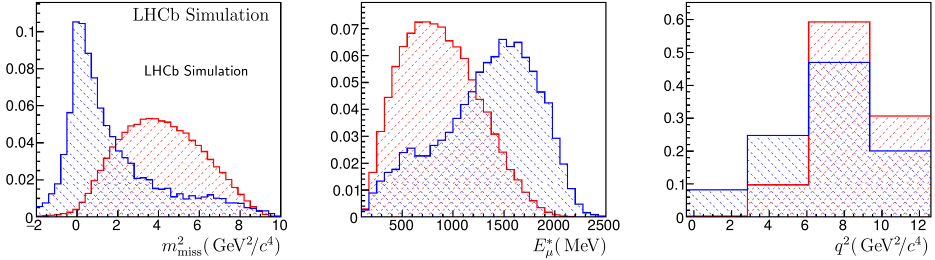
- Isolation MVA selects a track with loose kaon ID $\to$ select a sample enhanced in $B \to D^* D X$
- Use this to constrain, justify $B \to D^* D X$ shape

Combinatorial backgrounds



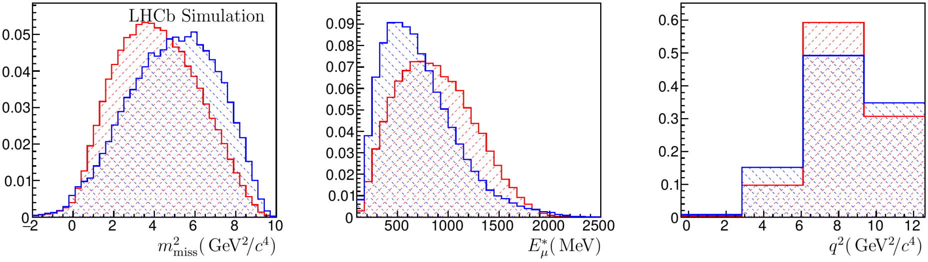
- Combinatorial background modelled using same-sign $D^{*+}\mu^+$ data
- Two sources of combinatorial background are treated separately (shown on next slide)

Combinatorial backgrounds



- Non D^{*+} backgrounds (fake D^*) template modelled using $D^0\pi^-$ data (shown)
    - Yield determined from sideband extrapolation beneath D^{*+} mass peak
- Hadrons misidentified as muons (fake muons)
    - Controlled using $D^{*+}h^{\pm}$ sample
    - Both template and expected yield can be determined
- Both of these are subtracted from $D^{*+}\mu^+$ template to avoid double counting

$D^{*+}\tau X$ backgrounds



- Two small backgrounds containing taus, each $<\sim$ 10% of the signal yield: $B \to D^{**}\tau^+\nu$ (shown) and $B \to D^*(D_s \to \tau\nu)X$
  - Both too small to measure
- $B \to D^{**}\tau^+\nu$ constrained based on measured $B \to D^{**}\mu^+\nu$ yield, theoretical expectations ($\sim$50% uncertainty)
- $B \to D^*(D_s \to \tau\nu)X$ constrained based on $B \to D^* DX$ yield, and measured branching fractions ($\sim$30% uncertainty)

Systematics / efficiencies

Model uncertainties	**Size ($\times 10^{-2}$)**
Simulated sample size	2.0
Misidentified μ template shape	1.6
D^* form factors	0.6
$B \to D^* D X$ shape	0.5
$\mathcal{B}(B \to D^{**}\tau\nu)/\mathcal{B}(B \to D^{**}\mu\nu)$	0.5
$B \to [D^*\pi\pi]\mu\nu$ shape	0.4
Corrections to simulation	0.4
Combinatoric background shape	0.3
D^{**} form factors	0.3
$B \to D^*(D_s \to \tau\nu)X$ fraction	0.1
Total model uncertainty	**2.8**

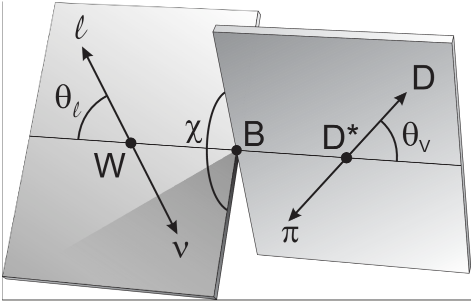
Multiplicative uncertainties	**Size ($\times 10^{-2}$)**
Simulated sample size	0.6
Hardware trigger efficiency	0.6
Particle identification efficiencies	0.3
Form-factors	0.2
$\mathcal{B}(\tau \to \mu\nu\nu)$	< 0.1
Total multiplicative uncertainty	**0.9**
Total systematic uncertainty	**3.0**

- Largest systematic from simulation statistics $\rightarrow$ reducible in future
- Next largest systematic from choice of method used to construct fake muon template
- Other systematic from background modelling depend on control samples in data
  - No uncertainties limited by external inputs
- Systematics from ratio of $B \to D^*\mu\nu$ and $B \to D^*\tau\nu$ efficiencies small

Other hadronic analyses

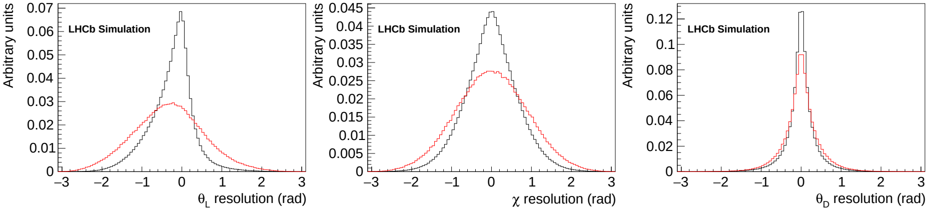
- After $\mathcal{R}(D^*)$, expect full program of measurements with hadronic tau
- $\mathcal{R}(\Lambda_c)$already underway
- Key issue: normalisation channels
  - Hadronic $\mathcal{R}(D^*)$ measurement relies on precise external measurement of $B \to D^{*+}\pi^-\pi^+\pi^-$
  - These do not exist for e.g $\Lambda_b \to \Lambda_c\pi^-\pi^+\pi^-$
  - Plan to use theory calculation for $\mathcal{B}(\Lambda_b \to \Lambda_c\mu\nu)/\mathcal{B}(B \to D^*\mu\nu)$ to avoid dependence on Λ_b production fraction

Beyond Rs



- Ratios of branching fractions are only the first observable
  - q^2, angles, τ/D^* polarisation have different sensitivity to new physics
- Variables fitted in $\tau \to \mu\nu\nu$ analyses already have some sensitivity to this
  - For now, measurements assume SM distributions (+ uncertainties)

Angular resolutions for $B \to D^* \tau\nu$ ($\tau \to \mu\nu\nu$)

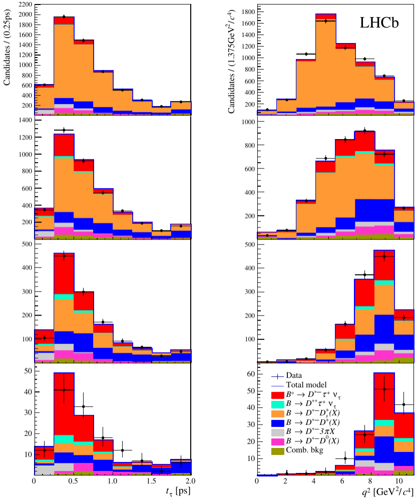


- Angular resolution for $B \to D^* \mu\nu$ (black) and $B \to D^* \tau\nu$ (red)
- Tau decay results in degredation of resolution
- Pretty wide, but have something to work with
  - Interesting mesurements also possible in muonic modes
- Ideas for how to exploit this, some tools already exist
- Sensitivity not yet known, may need larger samples to really pin things down..

Future

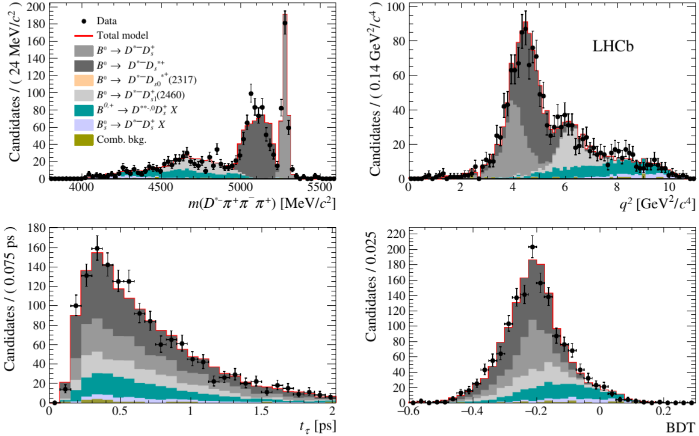
- What we have analysed now is a tiny fraction of the sample we will eventually collect
  - With 50 fb^{-1} (2021-2030), samples will grow by a factor ~ 30
  - With 300 fb^{-1}, (2034) samples will grow by a factor ~ 200
  - No sign that we hit a systematic limit
  - O(10 million) $B \to D^{*}\tau\nu$ ($\tau \to \mu\nu\nu$) events $\to$ huge power for angular analysis
  - Need to work together with theory to understand all contributions to the needed precision $\to$ continuous process
  - Even more suppressed signals ($B_c \to J/\psi\tau\nu X$, $B \to D^{**}\tau\nu$, $b \to u\tau\nu$ modes?) can have high statistical precision

Fit



- Now in slices of BDT output

Dealing with $B \to D^* D X$



- Use data to control $B \to D^* D X$ modelling
- Can use $D_{(s)} \to \pi\pi\pi$ mass peak to select a pure $B \to D^* D X$ sample
- This controls the $B \to D^* D X$ modelling, but not the $D \to \pi\pi\pi X$

Unfolding isn't fundamentally sound

- Unfolding doesn't have good statistical properties
- See e.g R. D. Cousins, S.J. May, Y. Sun "Should unfolded histograms be used to test hypotheses?"
  - Spoilers: probably not
  - Even before biases introduced by regularisation
  - Going in the other direction is a fundamentally well defined procedure
- Describing the full space will require O(1000) bins → not practical to unfold
- Uncertainty from background shapes difficult to reproduce accurately as a simple "background subtraction"
  - Often just ignored, we really cannot do this

Forward folding

- Don't deconvolute data to theory, convolute theory to data
  - Best convolution: MC simulation
- This is exactly what we are already doing!
  - Can build on what we already have...
- Problem: model dependence - need to choose functional form
  - We will explore all possibilities

Histogram expansion PDF

- What we want to do: reweight MC, reproduce histogram PDF
  - Event-by-event $\rightarrow$ slow
- Weight for each event can be written as
  $\sum [(\text{Combination of fit coefficients}) \times (\text{Stuff invariant in fit})]$
  - (or expand it until it can be..)
  - Loop through events once, for each term generate a histogram
  - Adding up histograms, scaled by fit coefficients, exactly equivalent to fully reweighted histogram
- Only need to sum up histograms $\rightarrow$ fast
  - Already using for muonic $\mathcal{R}(D^{(*)})$