Lepton non-universality in $B \to D^* \tau \nu$?

Greg Ciezarek

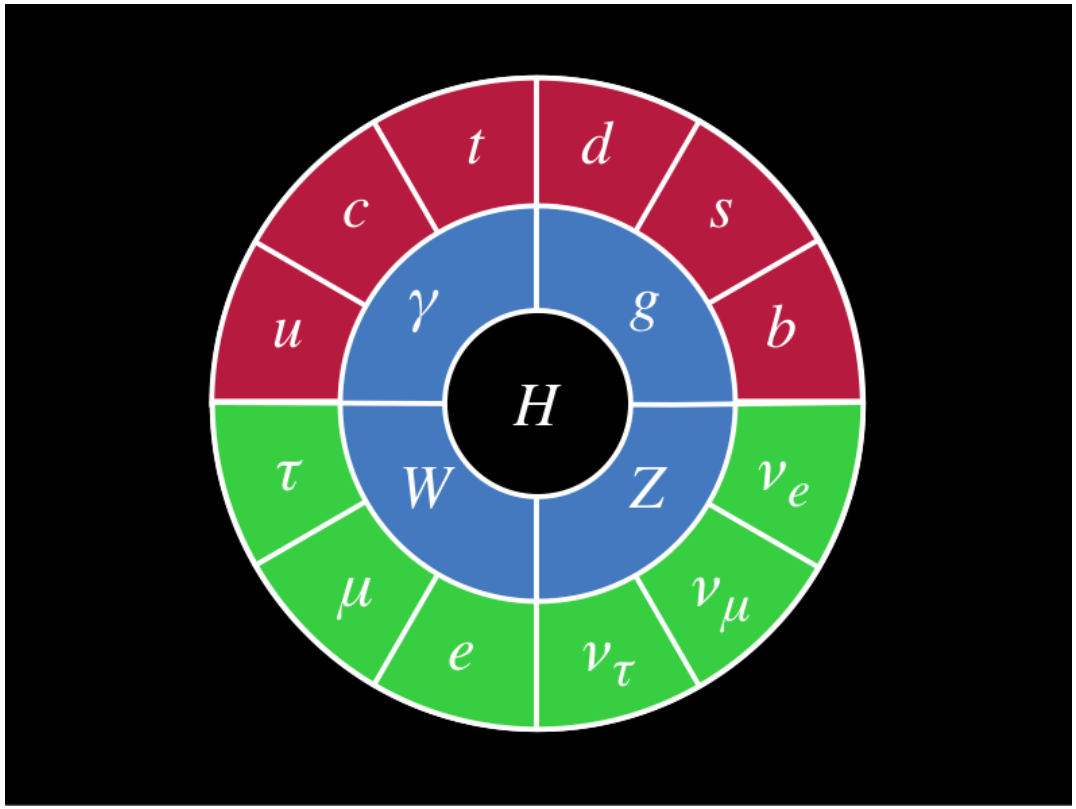
CERN

February 27, 2018

Overview

- Introduction to flavour
- Lepton universality
- $B \to D^* \tau \nu$ at LHCb with $\tau \to \mu\nu\nu$
- $B_c \to J/\psi \tau \nu$
- $B \to D^* \tau \nu$ at LHCb with $\tau \to \pi\pi\pi\nu$
- Interpretation

Flavour in the Standard Model



- Fermions are important (we're made of them)
- We have three (and only three) generations of fermions
  - And no clue why
- In the Standard Model (SM), they are distinguished only by their couplings to the Higgs (Yukawa couplings)

Flavour in the Standard Model

- Quark masses are generated by their Yukawa couplings- in the generational basis:

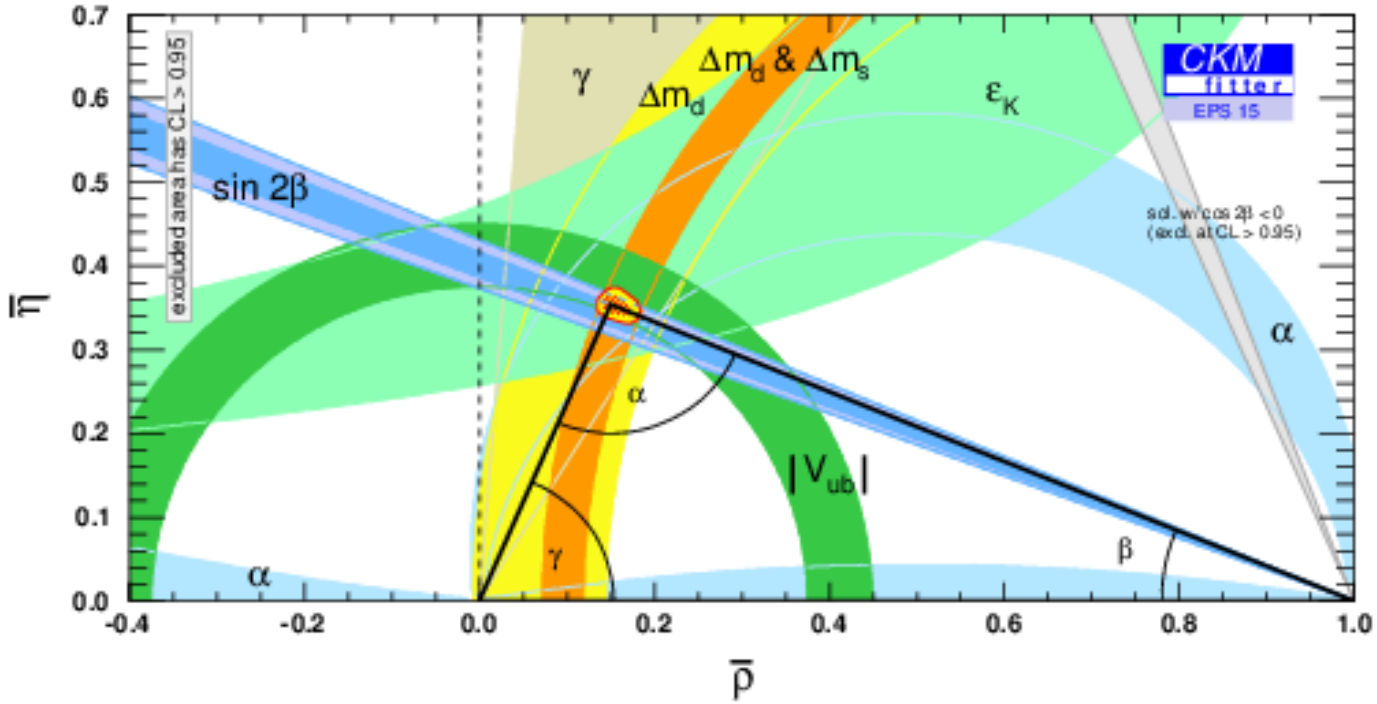
$$L_Y^q = \bar{u}_{Li} m_{ij}^u u_{Rj} + \bar{d}_{Li} m_{ij}^d d_{Rj} + h.c.$$

- These matrices can have off-diagonal terms $\rightarrow$ not in mass basis
- To move into mass basis, use four unitary diagonalisation matrices (V):

$$M_{diag}^{u,d} = V_{L,R}^{u,d} m_{u,d} V_{L,R}^{u\dagger,d\dagger}$$

- Product $V_L^u V_L^{d\dagger}$ appears in charged current interaction: this is the CKM matrix
- Only predictions are due to unitarity of V matrices:
  - Only get $V_L^u V_L^{u\dagger}, V_L^d V_L^{d\dagger}$ in neutral current interaction $\rightarrow$ no flavour changing Z interactions
  - CKM matrix must be unitary

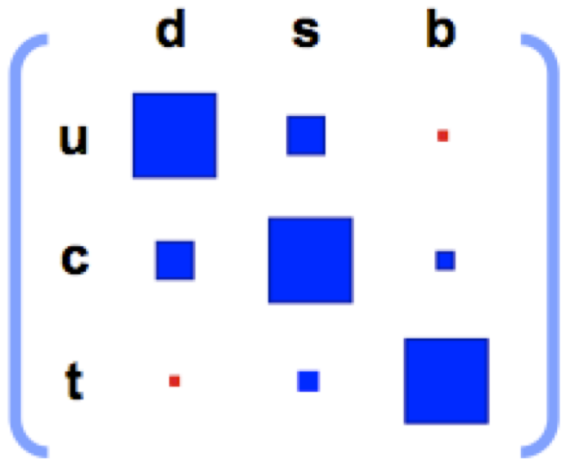
CKM unitarity



- Unitarity triangle fit consistent with CKM unitarity (so far)
  - The only SM prediction about flavour confirmed to first order
  - Flavour can be accommodated in the SM
  - Nobel prize for Kobayashi and Maskawa in 2008

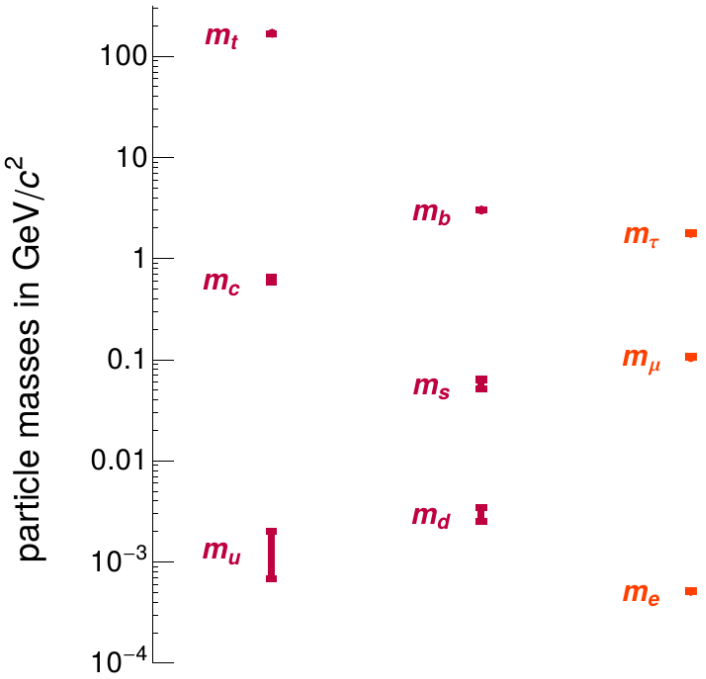
Quark flavour transitions

$$|V| \simeq \begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix}$$



- CKM matrix appears hierarchical
- Contains four fundamental parameters of SM
- PMNS matrix doesn't look remotely similar

Fermion masses



- Each one of these 9 masses is a fundamental parameter of the SM
- Electron and top masses differ by a factor $\sim 10^6$
- Even worse with neutrinos included?

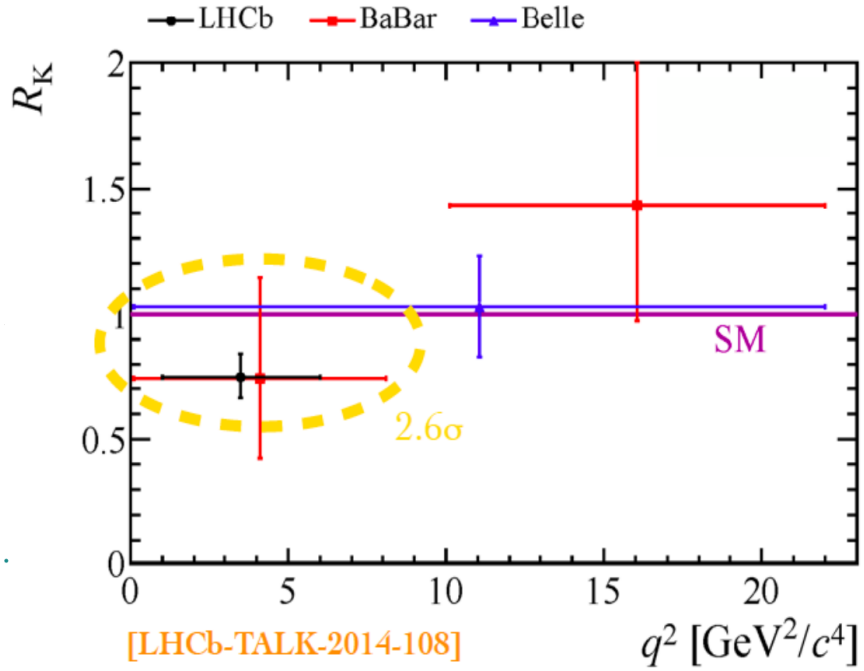
Yukawa couplings

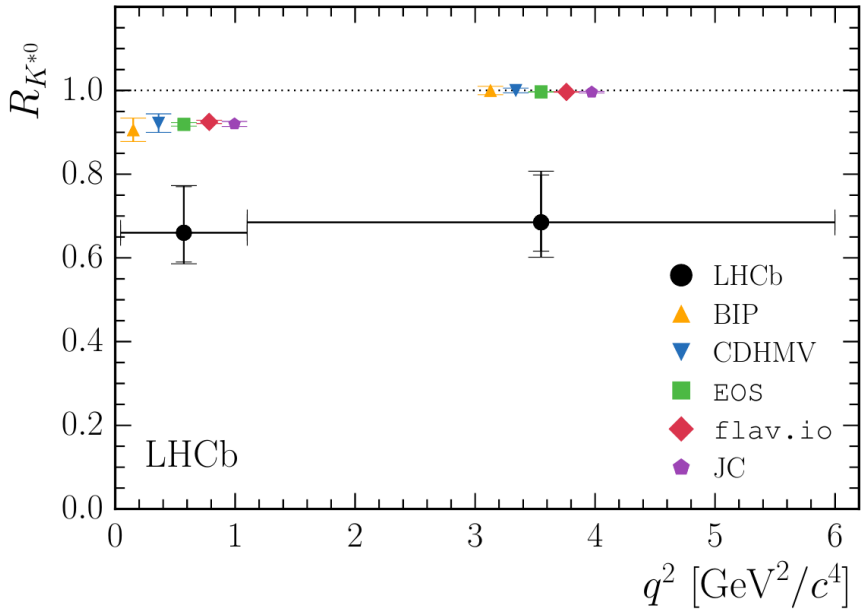
- 13 of 18 SM free parameters
- Only place which distinguishes generations
- Sole source of CP violation seen in nature
- No quark masses $\rightarrow$ no strong-CP problem
- Sole source of Higgs mass divergence
- All in all, deeply suspicious

Lepton universality

- Lepton universality: leptons only distinguished by their Yukawa couplings
- Tested to per-mille at Z pole by LEP,SLD
- Established to per-mille in pion decays, a few per mille in kaons
- Not so well tested in heavy flavour decays:
  - $e - \mu$ tested to $\sim 5\%$ in charm,$\sim 4\%$ in beauty
  - Hints of non-universality in beauty...

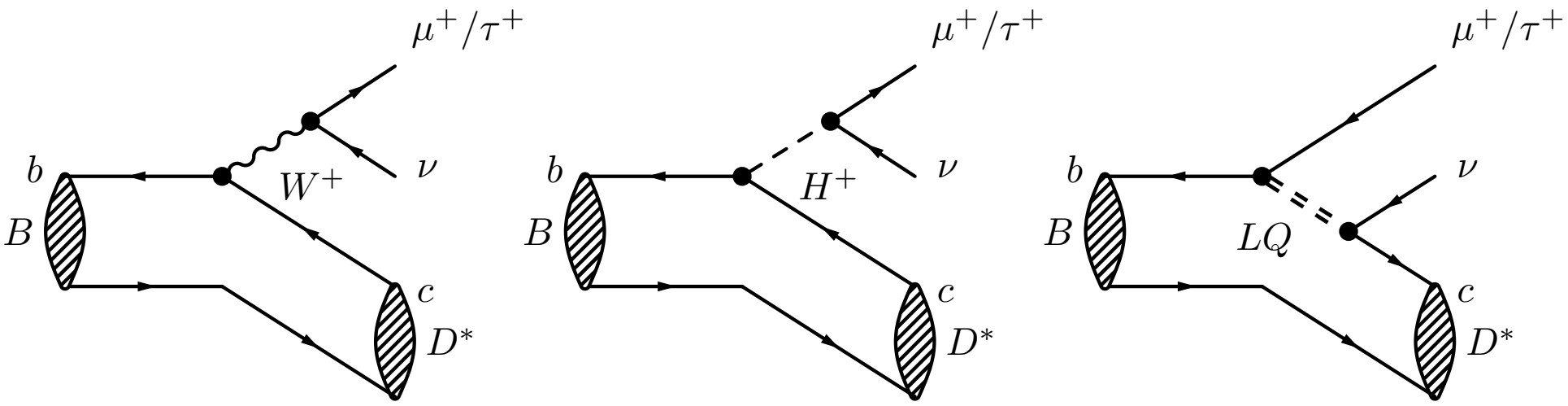
$b \to s\ell\ell$ anomalies





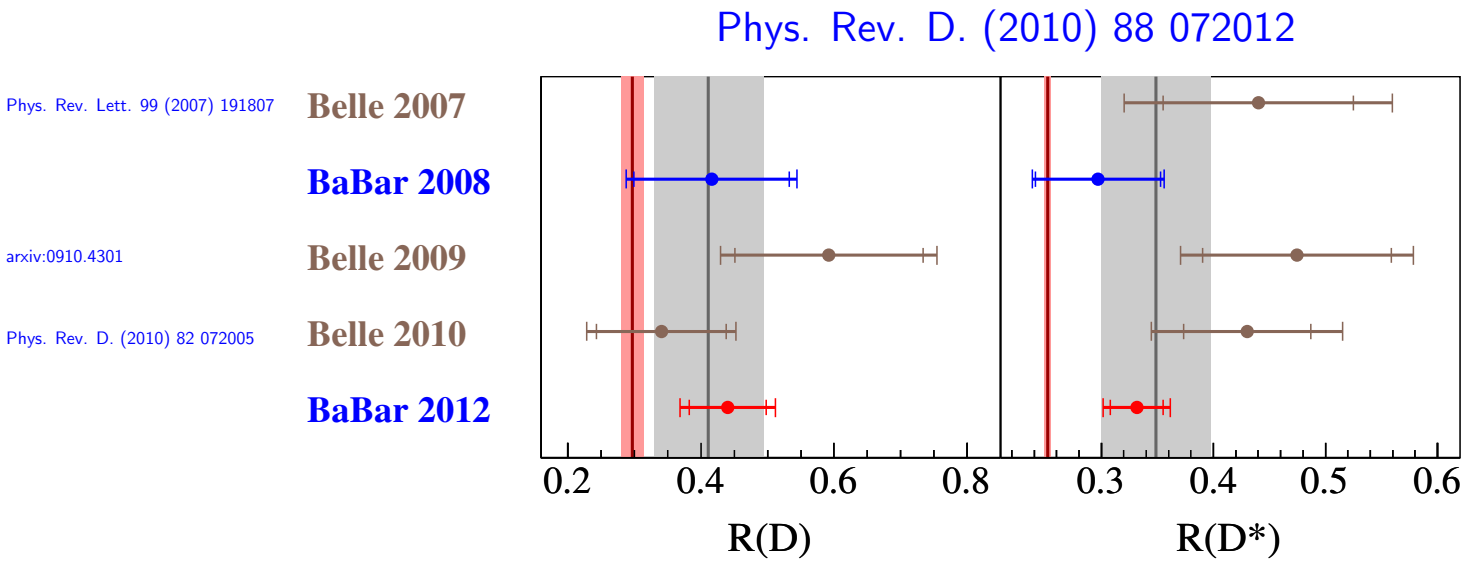
- $R_{K(*)} \equiv \mathcal{B}(B \to K^{(*)}\mu^+\mu^-)/\mathcal{B}(B \to K^{(*)}e^+e^-)$
  - Predicted to be unity with 1% uncertainty (away from very low masses)
  - R_K 2.6σ from universality (arxiv:1406.6482)
  - R_{K*} both bins 2σ from universality (arxiv:1705.05802)
- Interesting because it matches with tensions in $b \to s\mu\mu$ branching fraction measurements, "$P5'$ anomaly" in $B^0 \to K^{*0}\mu^+\mu^-$
  - Uncertainty about theory predictions for the lepton-universal observables, however

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



- 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
  - Theoretically clean: $\sim 2\%$ uncertainty
- Ratio R($D^{(*)} = \mathcal{B}(B \to D^{(*)}\tau\nu)$ / $\mathcal{B}(B \to D^{(*)}\mu\nu)$ is sensitive to lepton non-universality, or charged Higgs

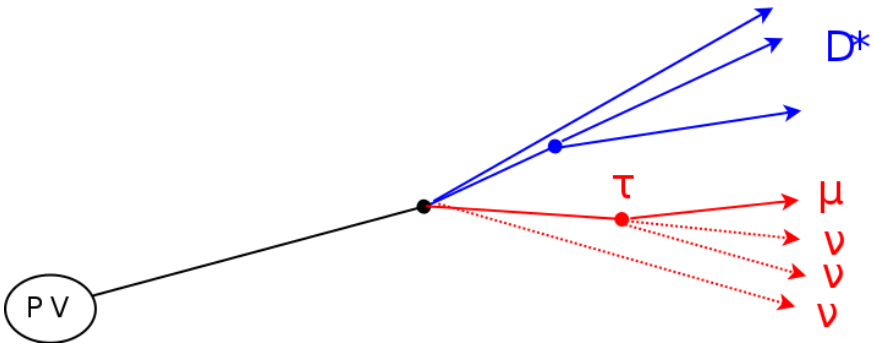
History



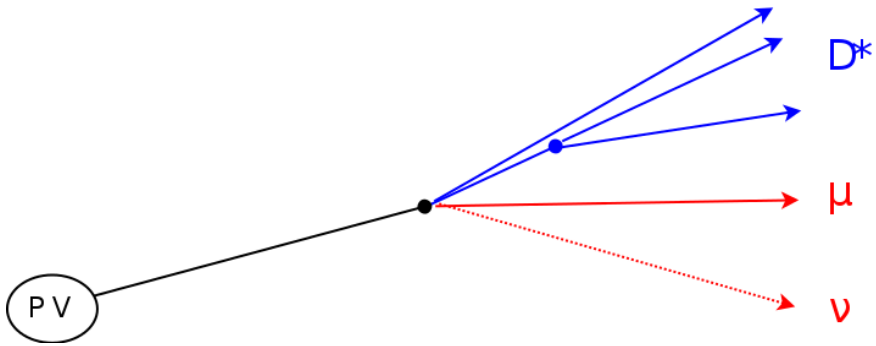
- Before 2015: measurements from B factories in $\tau \to \ell\nu\nu$ channel
- Final measurement from BaBar (arxiv:1303.0571) claimed 3 σ excess over SM expectation
- Strong motivation for more measurements
- (Will show more recent results later)

Experimental challenge

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

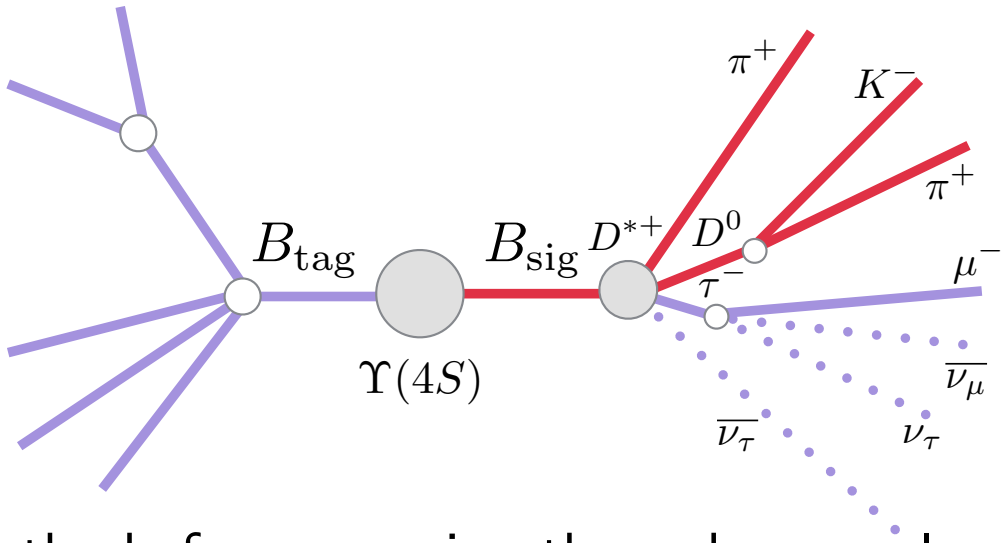


$B \to D^* \mu \nu$



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

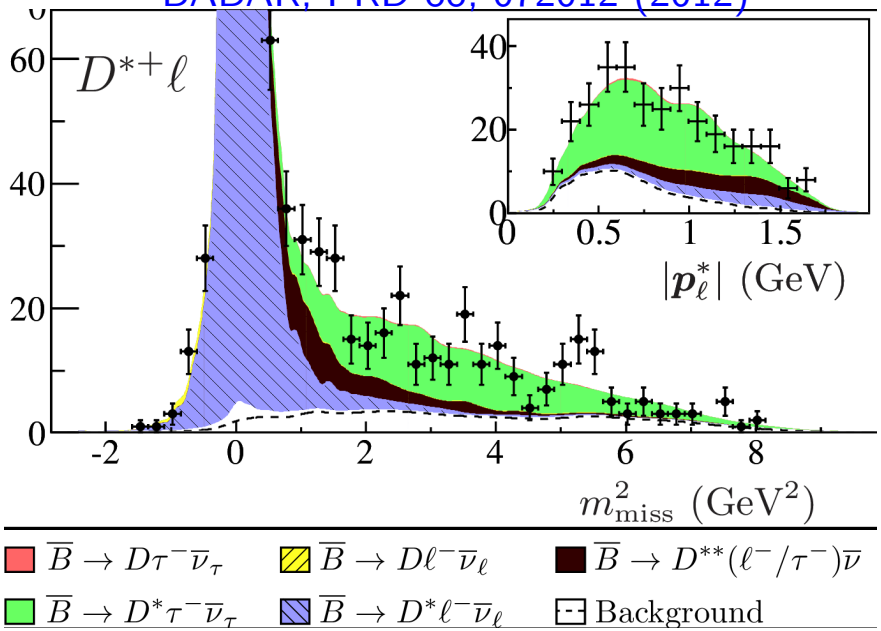
B Factory method



- 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 $B \to$ measurement of signal B kinematics
  - Signal B + other B should be entire event $\to$ strong rejection against other missing reconstructable particles
- Penalty: sub percent efficiency

Babar measurement

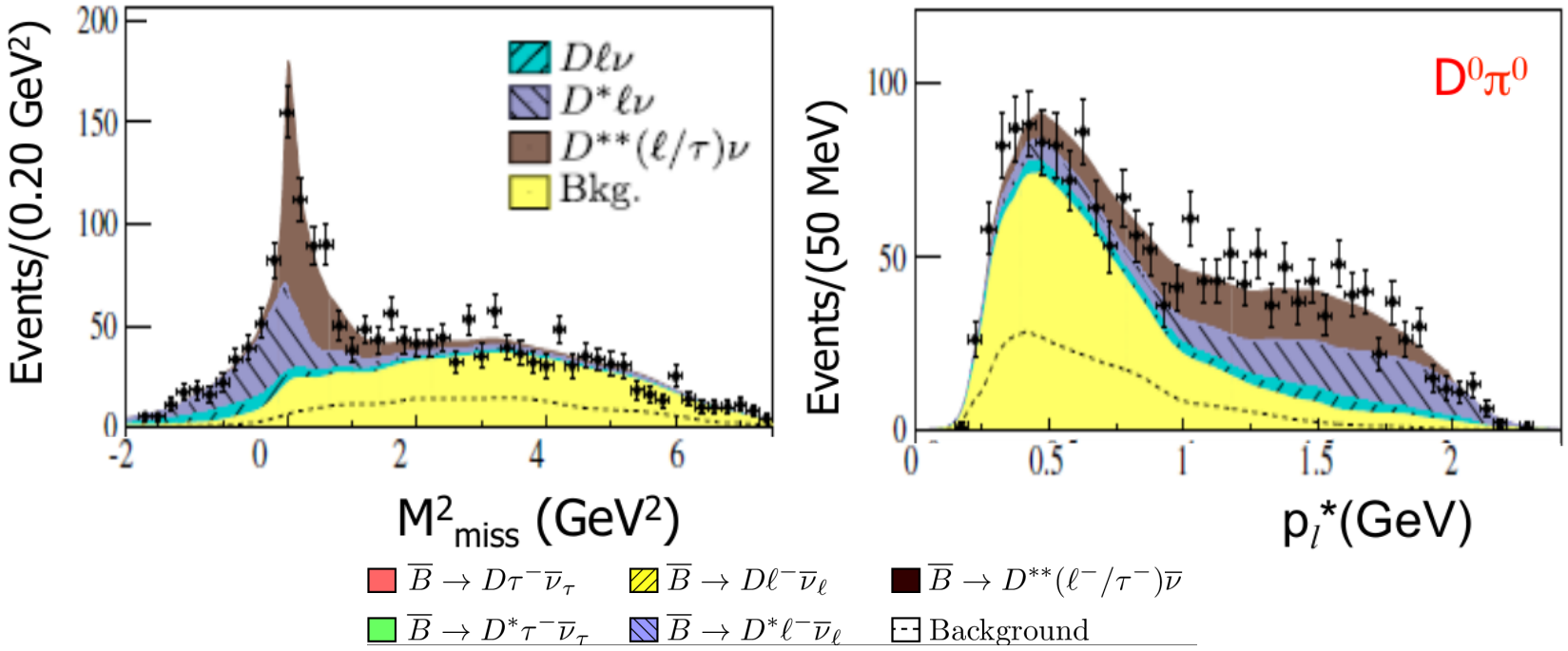
BABAR, PRD 88, 072012 (2012)



- Babar measurement strategy: 2D fit in missing mass^2 ($m^2_{\rm miss}$), lepton energy in B rest frame
- Narrow peak for $B \to D^*\mu\nu$, broader distributions for $B \to D^*\tau\nu$, $B \to D^{**}\mu^+\nu$
- (Cleanest sample shown)

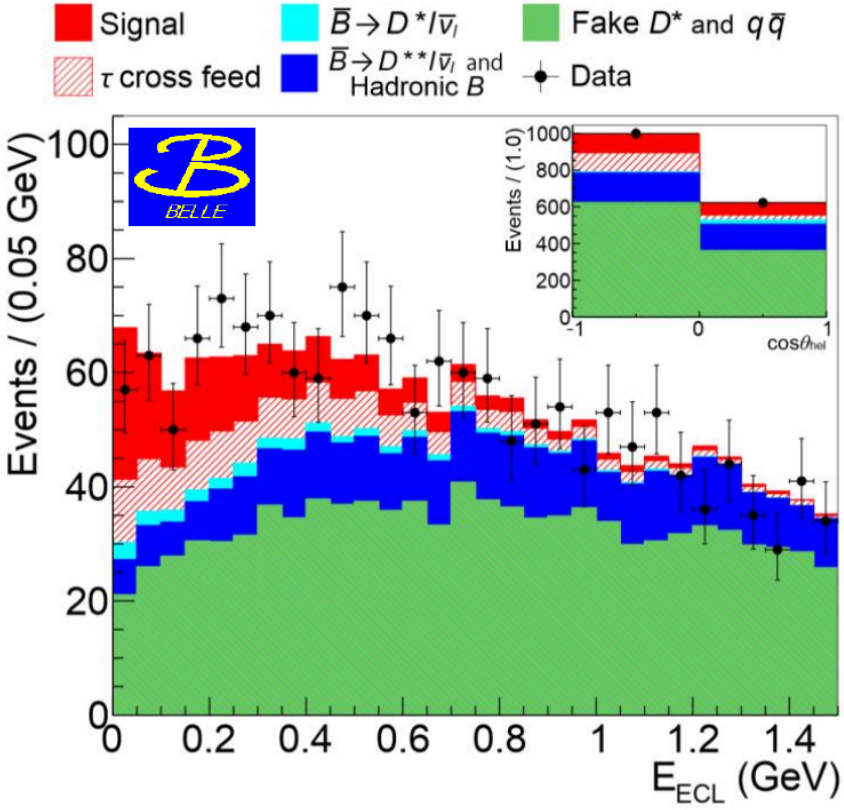
Babar control sample

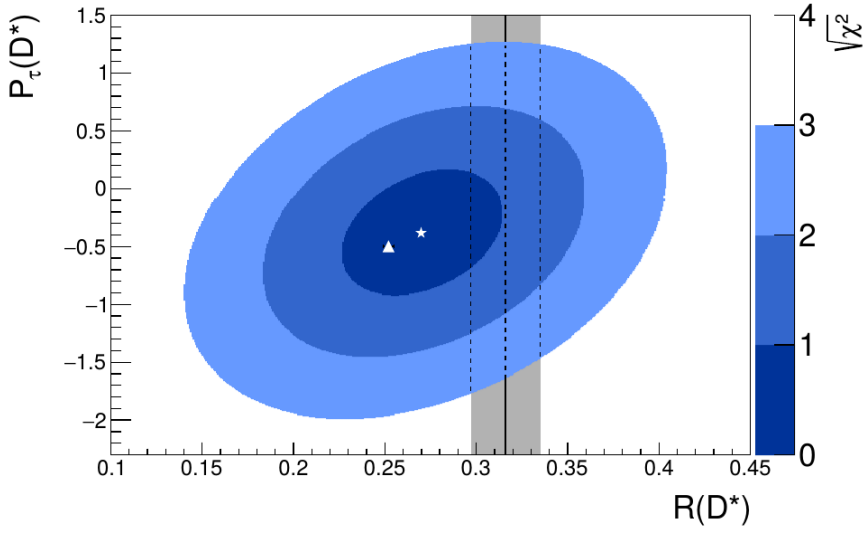
BABAR, PRD 88, 072012 (2012)



- $B \to D^{**}\mu^{+}\nu$ background controlled using $D^{(*)}\pi^{0}$ sample
  - Using data to control background crucial to all measurements
- Combined $\mathcal{R}(D)$and $\mathcal{R}(D^{*})$ 3 σ from SM: first hint of the hint
- Still the most precise single measurement

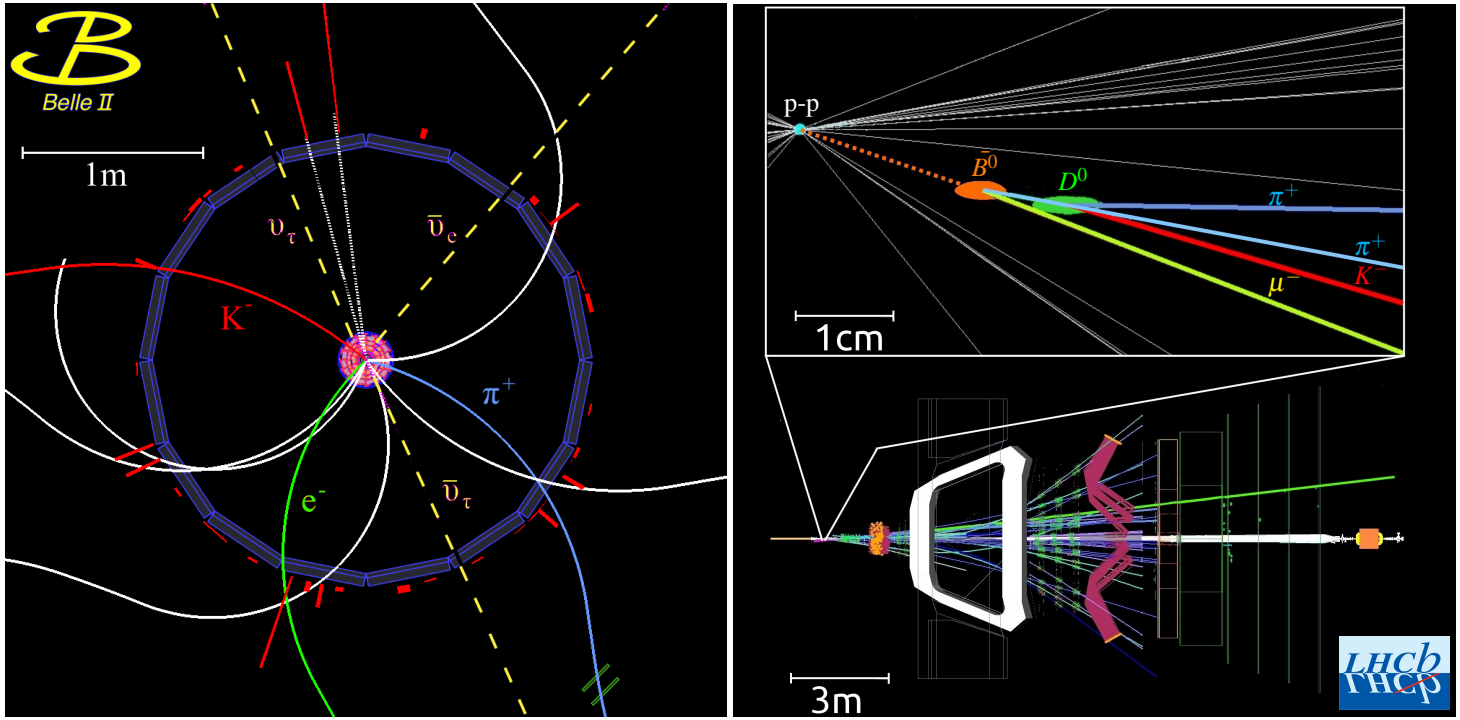
Belle $\tau^+ \to \pi^+(\pi^0)\nu$





- Recent from Belle: first measurement with $\tau^+ \to \pi^+(\pi^0)\nu$, and first measurement of τ polarisation
- Fit variable: energy left over in Calorimeters after tag+signal reconstruction (E_{ECL})
- Important proof of principle for Belle-II
- Phys. Rev. Lett. 118, 211801 (2017), Phys. Rev. D 97, 012004 (2018)

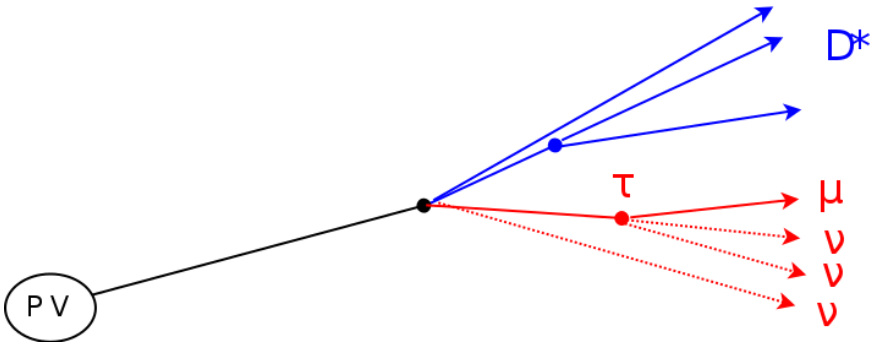
Can you do this at LHCb?



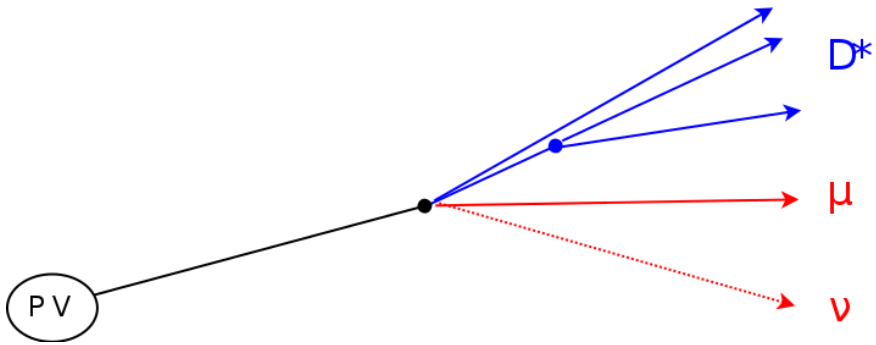
- Traditional methods for measuring these decays rely on $e^+e^- \to B\overline{B}$ event properties
  - In a hadron collider the $B\overline{B}$ centre of mass isn't fixed $\to$ rest of event provides little constraint on the signal B kinematics
  - Event also contains a lot of junk from the proton-proton interaction $\to$ reconstructing the whole event is meaningless
- Needed completely different methods

LHCb fit strategy

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

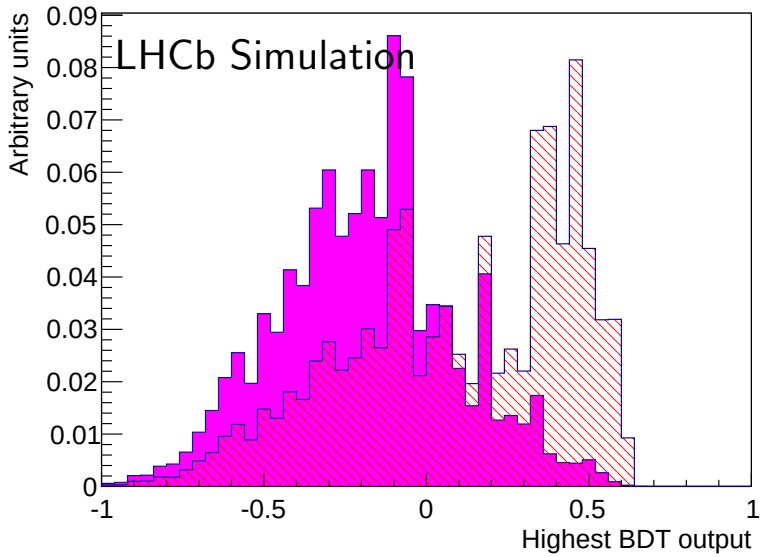


$B \to D^* \mu \nu$



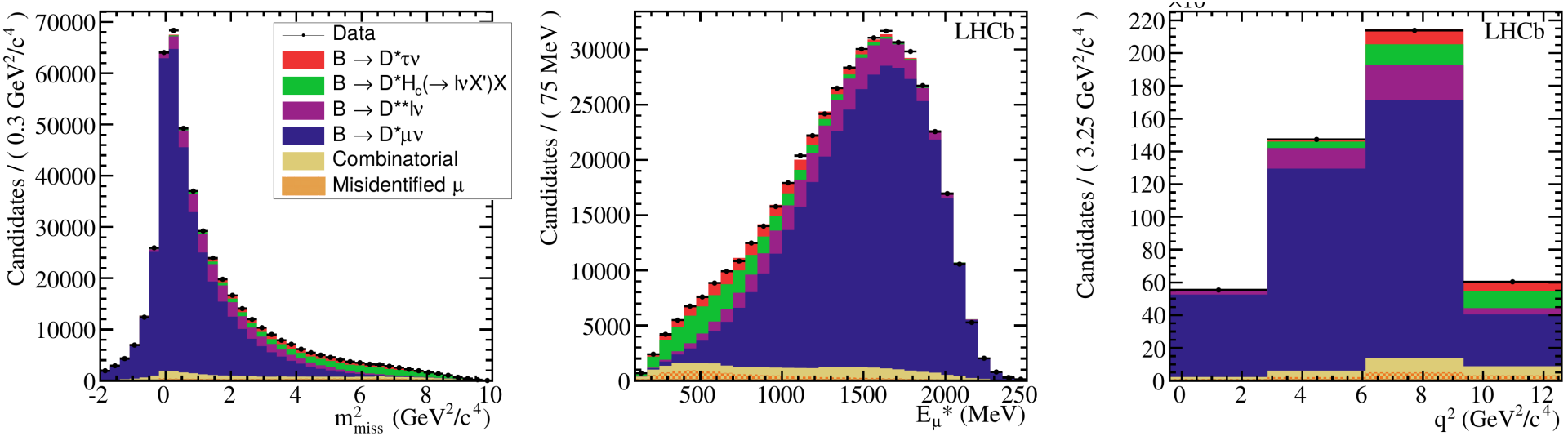
- Can use B flight direction to measure transverse component of missing momentum
- No way of measuring longitudinal component $\rightarrow$ use approximation to access rest frame kinematics
  - Assume $\gamma\beta_{z,visible} = \gamma\beta_{z,total}$
  - $\sim$20% resolution on B momentum, long tail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

Isolation MVA



- Strategy: use MVA to decide if each track is from the same B, or the rest of the event
- Highest MVA output distribution for $B \to D^{**}\mu^+\nu$ (hatched) and $B \to D^*\mu\nu$ (solid)
- Inverting the cut gives a sample hugely enriched in physics backgrounds $\to$ use this to control shapes

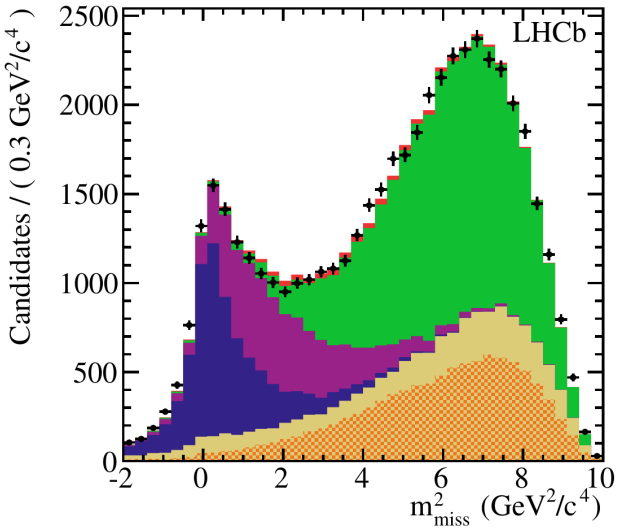
Fit strategy

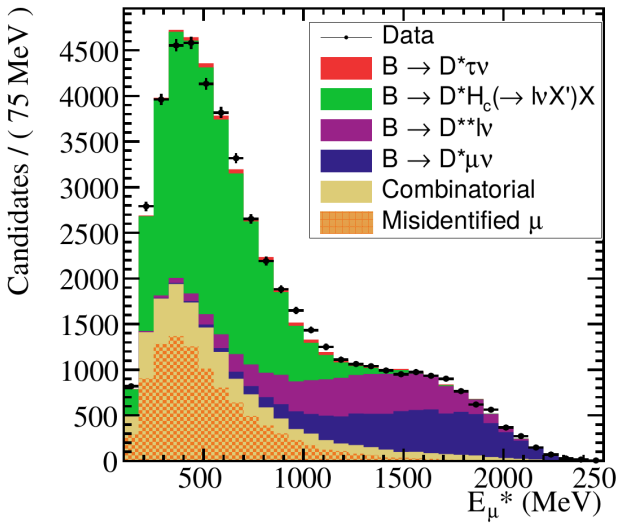


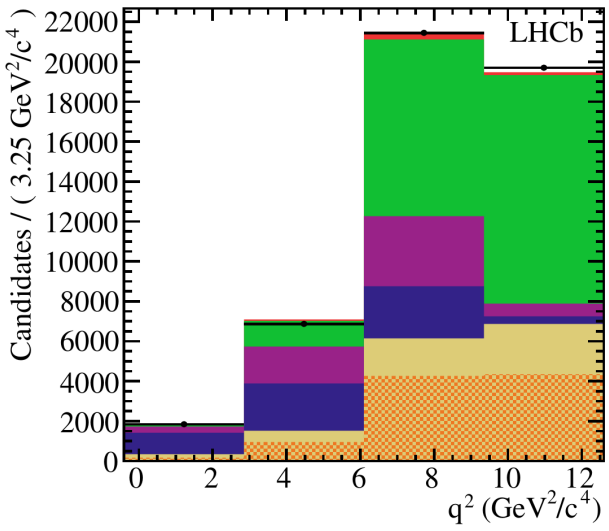
- Three dimesional template fit in E_μ (left), $m^2_{missing}$ (middle), and q^2
  - Projections of fit to isolated data shown
- Uncertainties on template shapes incorporated in fit:
  - Continuous variation in e.g $B \to D^* \mu\nu$ matrix element parameters

Background strategy

Phys. Rev. Lett. 115 (2015) 111803

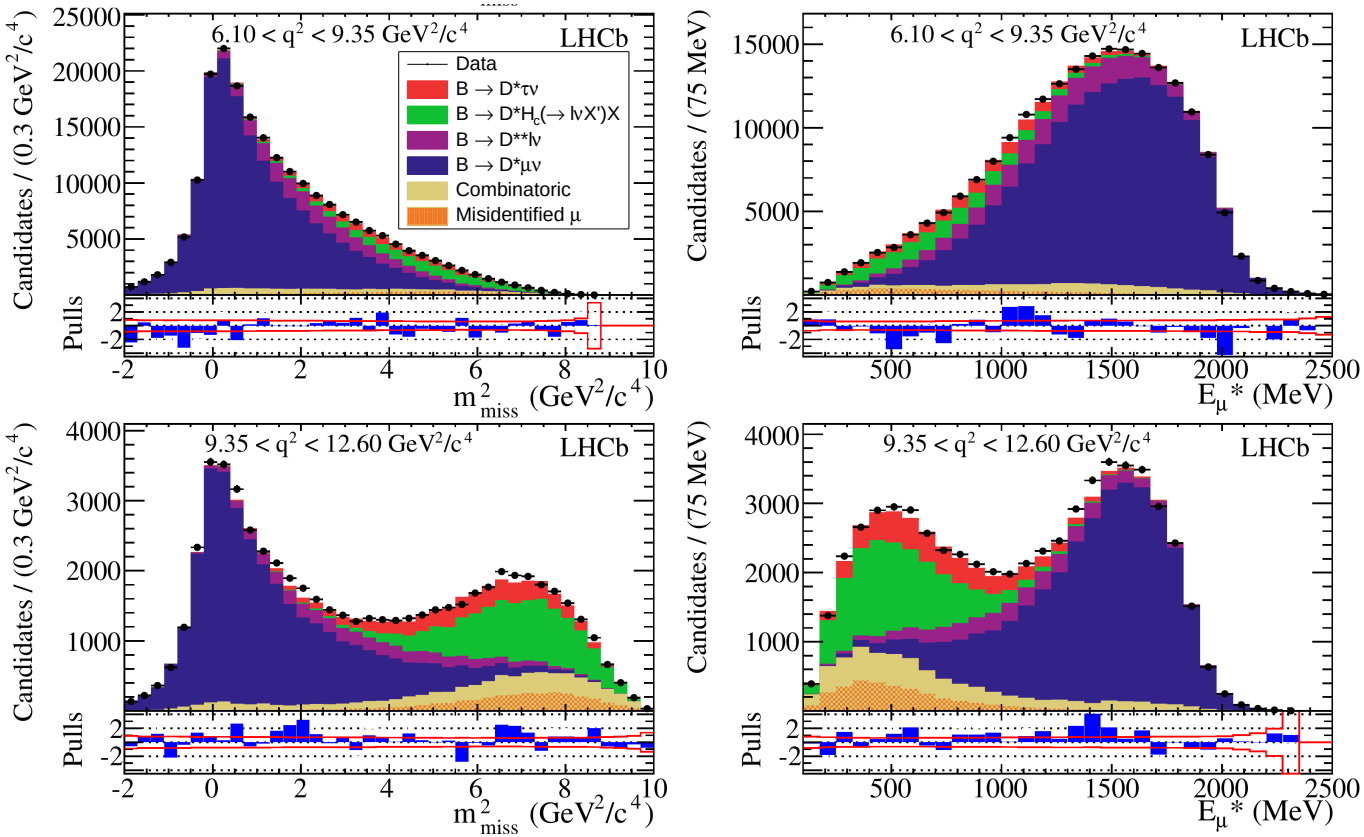






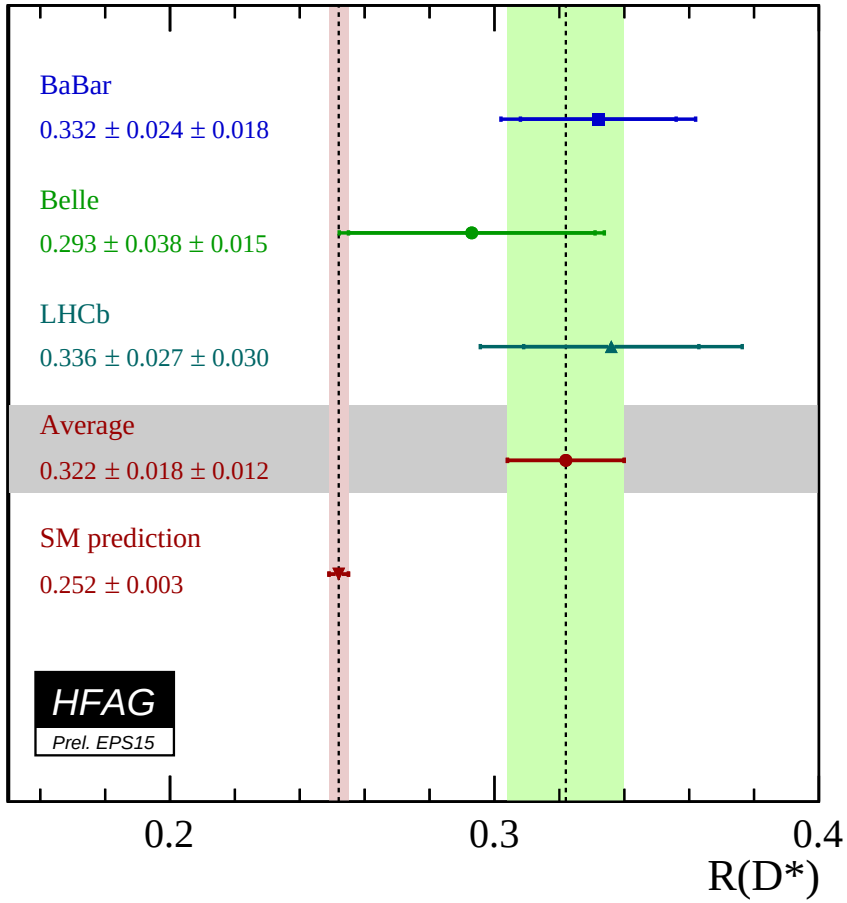
- All major backgrounds modelled using control samples in data
  - Dedicated samples for different backgrounds
  - Quality of fit used to justify modelling
  - Data-driven systematic uncertainties
- All combinatorial or misidentified backgrounds taken from data
- More details on everything in backups

Signal fit



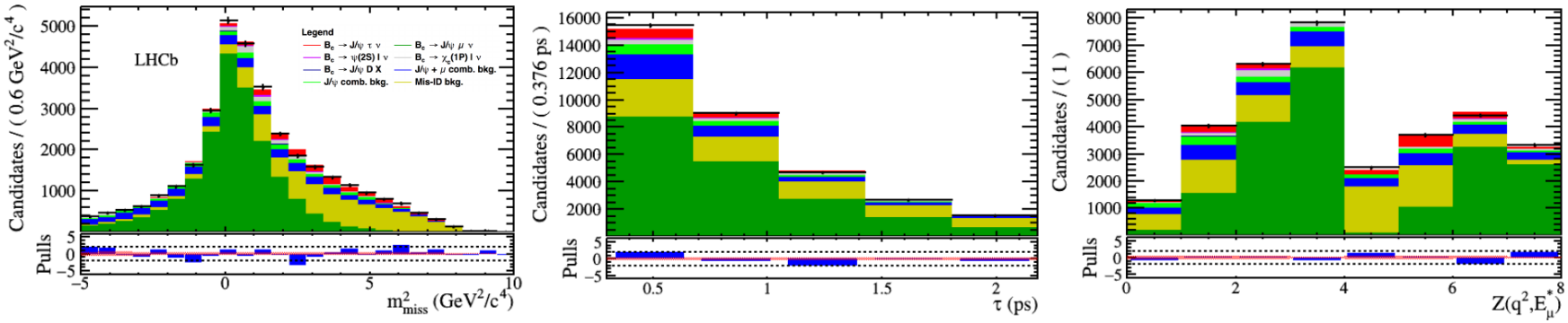
- Fit to isolated data, used to determine ratio of $B \to D^* \tau \nu$ and $B \to D^* \mu \nu$
- Model fits data well

Result



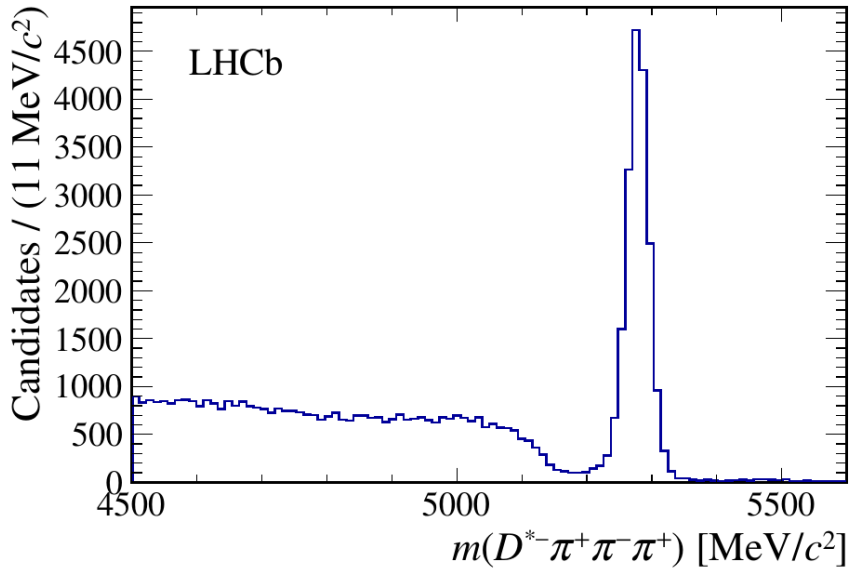
- We measure $\mathcal{R}(D^*) = 0.336 \pm 0.027 \pm 0.030$
  - In good agreement with past measurements
  - Agreement with SM at 2.1σ level
  - Only run 1 data!

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



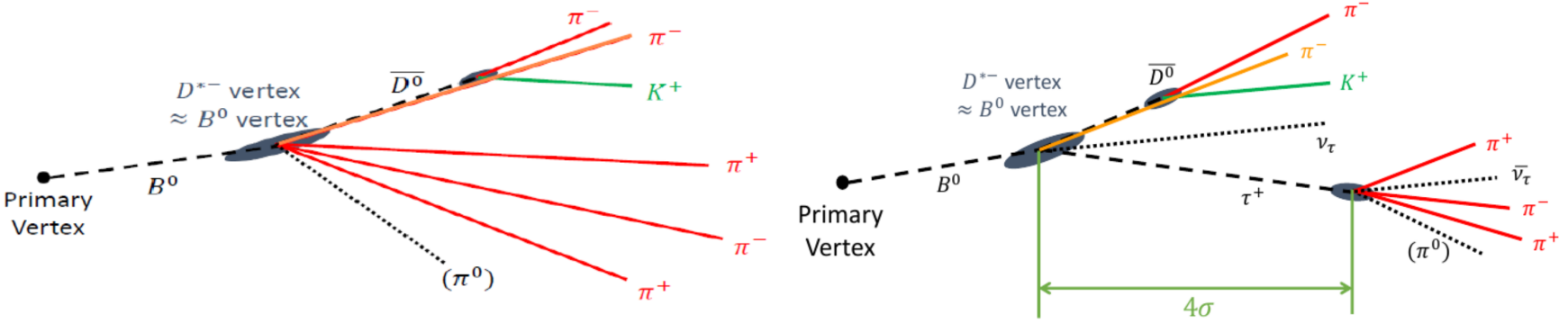
- $R_{J/\psi} \equiv B_c \to J/\psi\,\tau\nu / B_c \to J/\psi\,\mu\nu$
- Measured using very similar techniques to $\mathcal{R}(D^*)$, on run 1 data
- $R_{J/\psi} = 0.71 \pm 0.17 \pm 0.18$
  - $\sim 2\sigma$ from SM
  - But nearly as far from consistency with $\mathcal{R}(D^*)$
- LHCb-PAPER-2017-035(Run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

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



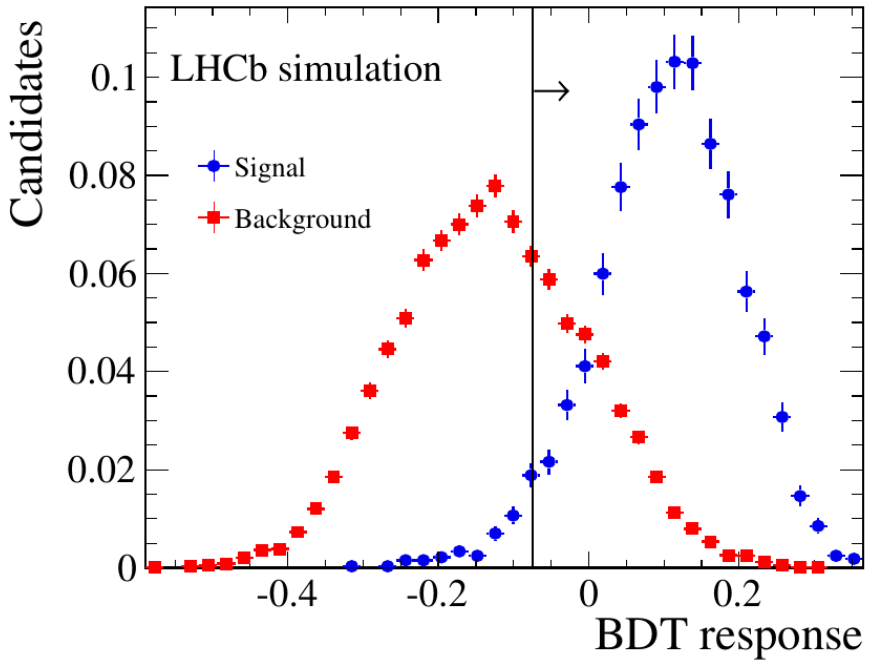
- Compared to muonic $\mathcal{R}(D^*)$:
  - Large $B \to D^*\mu\nu$, $B \to D^{**}\mu^+\nu$ backgrounds absent
  - Additional $B \to D^*\pi\pi\pi X$ backgrounds
  - $B \to D^* D X$ with $D \to \pi\pi\pi X$
- Control experimental efficiencies by measuring rate relative to $B \to D^*\pi\pi\pi$

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



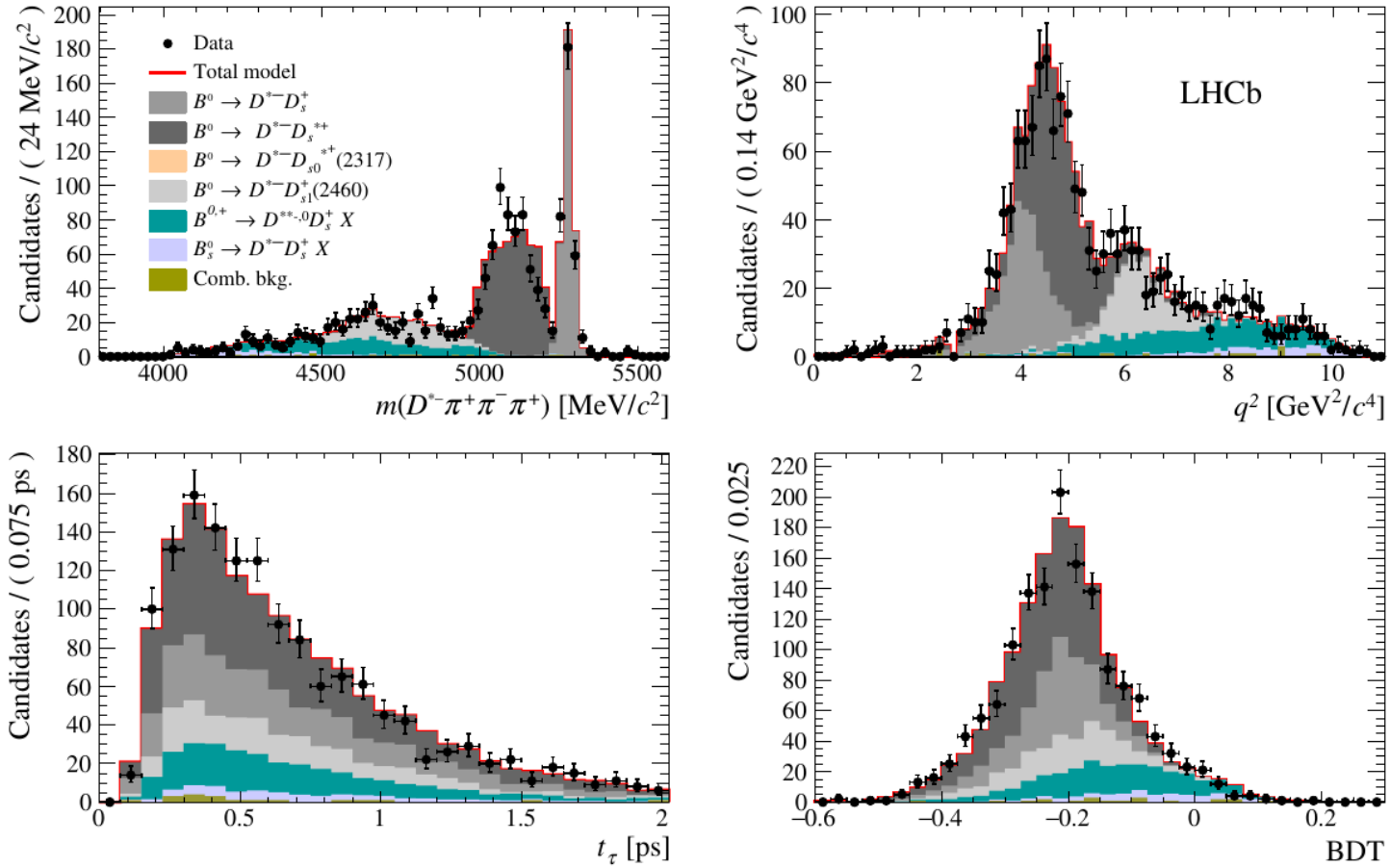
- Can use decay topology to remove direct $B \to D^* \pi\pi\pi X$ decays:
- If the $\pi\pi\pi$ vertex is displaced from the B vertex, cannot be direct $B \to D^* \pi\pi\pi X$
- Can remove a large, poorly measured background
  - And control the remainder
- $B \to D^* DX$ major physics background remaining

Dealing with $B \to D^* DX$



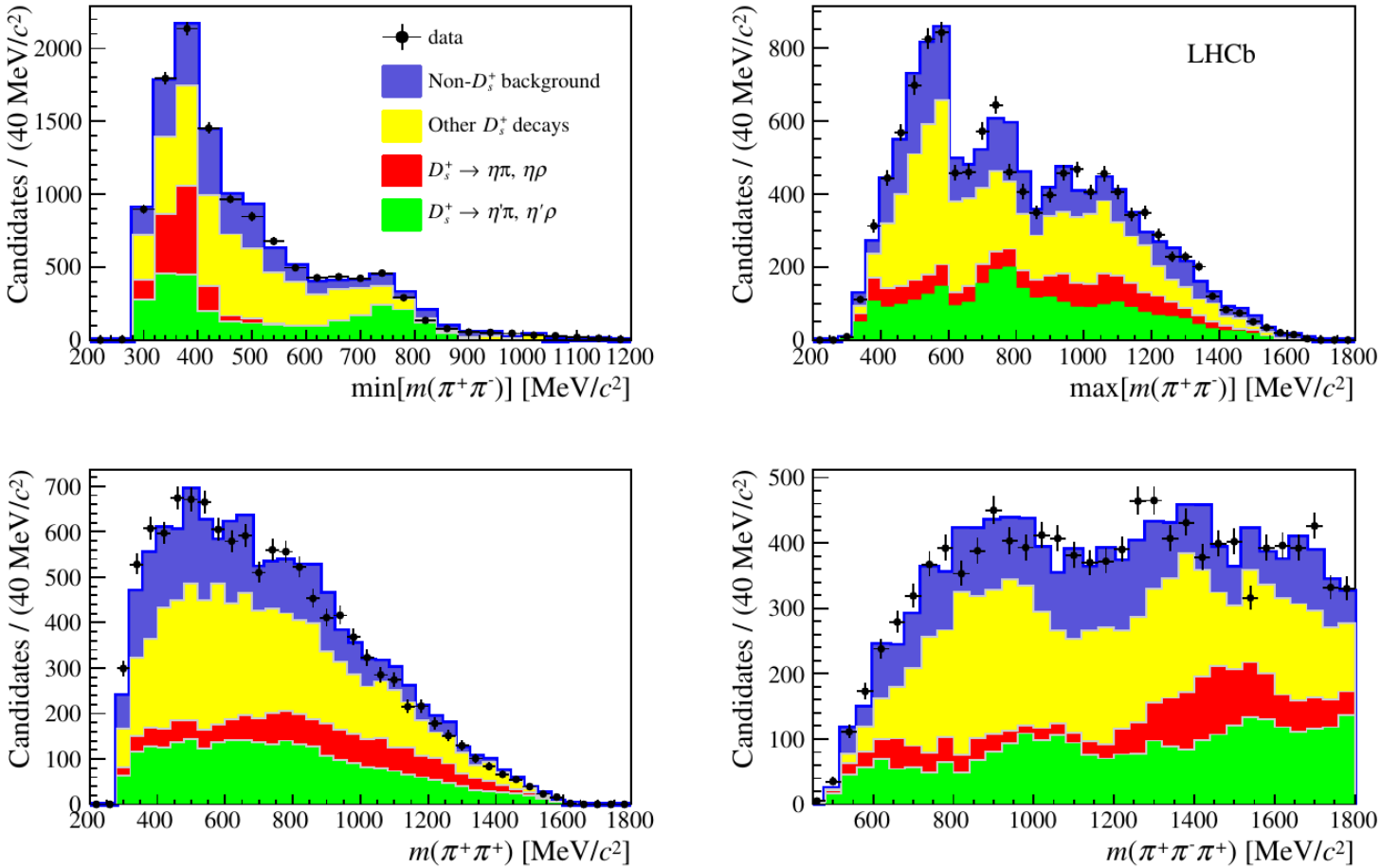
- $[\pi\pi\pi]$ lifetime discriminates between tau and $B \to D^* DX$
- Can use partial reconstruction techniques to reconstruct D peak in $B \to D^{*+} D$ (not $B \to D^* DX$)
- $\tau \to \pi\pi\pi\nu$ is mostly $a1(1260)$, $D \to \pi\pi\pi X$ mostly isn't
  - Use the $\pi\pi\pi$ (sub) structure to 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, neutral isolation

Dealing with $B \to D^* DX$



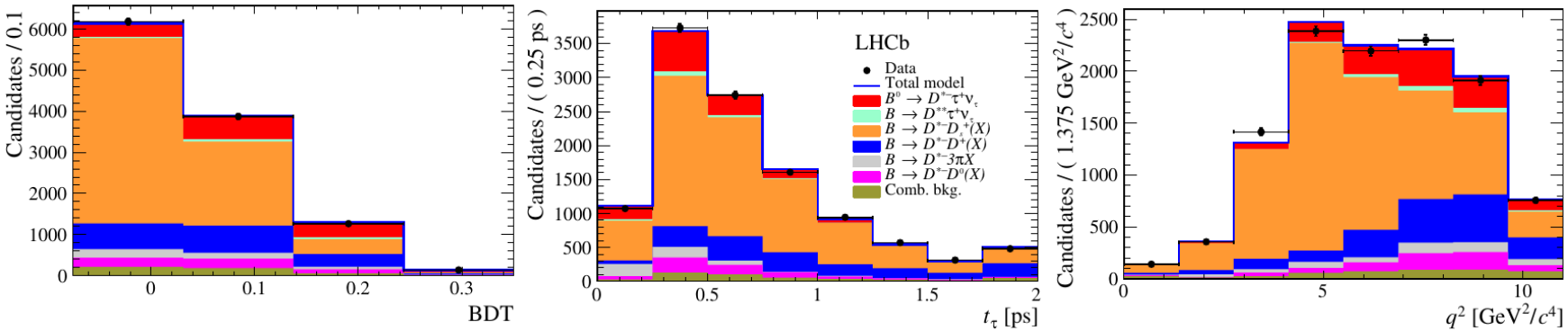
- Use data to control $B \to D^* DX$ modelling
- Can use $D_{(s)} \to \pi\pi\pi$ mass peak to select a pure $B \to D^* DX$ sample
- This controls the $B \to D^* DX$ modelling, but not the $D \to \pi\pi\pi X$

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



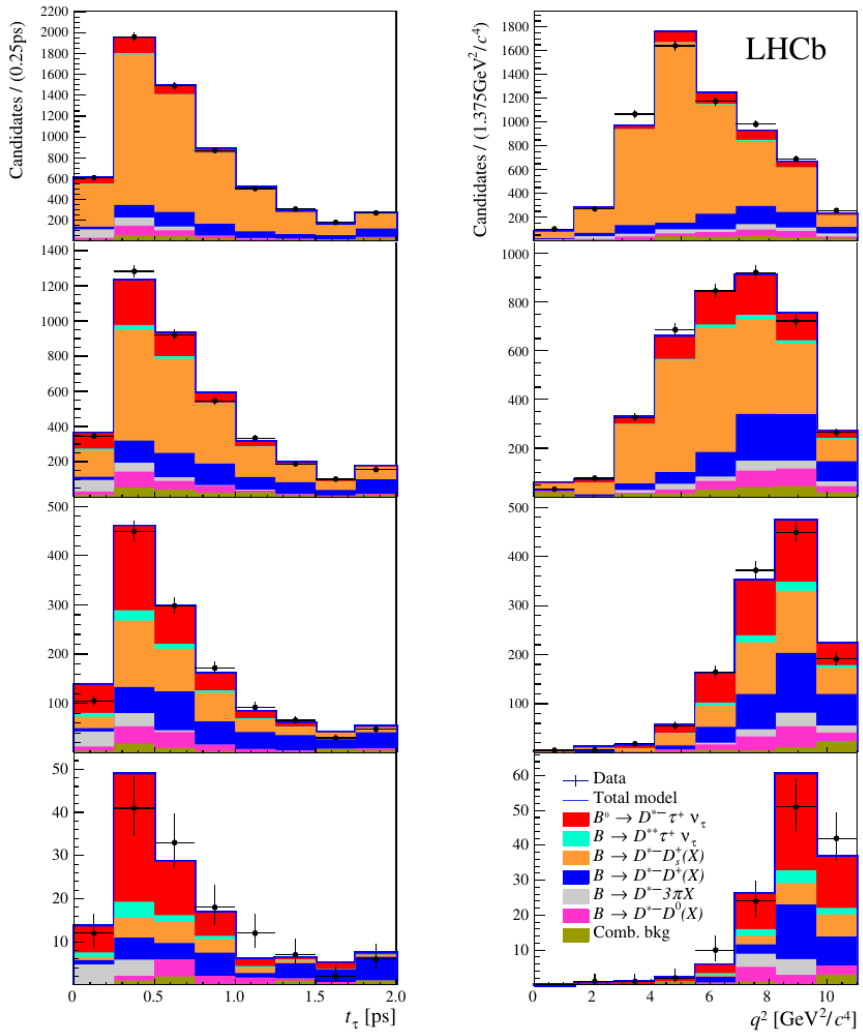
- Use low BDT region to control $D_s \to \pi\pi\pi X$ substructure
- Don't have such good control for $D^{+,0} \to \pi\pi\pi X$, so these float unconstrained

Fit



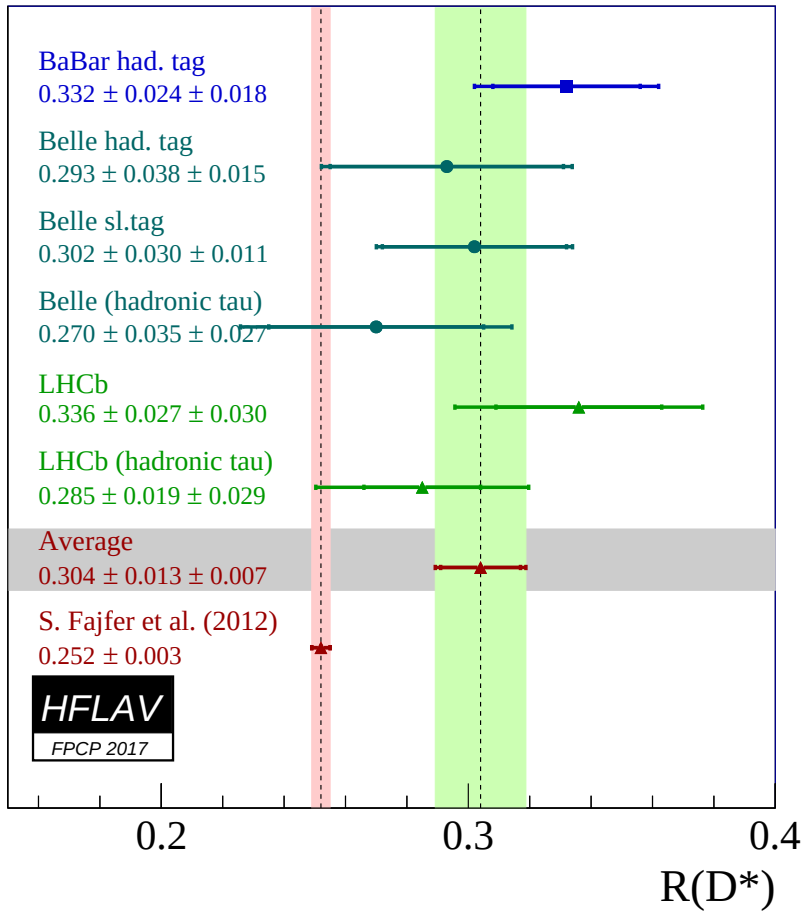
- 3D template fit in BDT, q^2, tau lifetime to determine signal yield
  - Only performed in high BDT range, previous control sample excluded

Fit



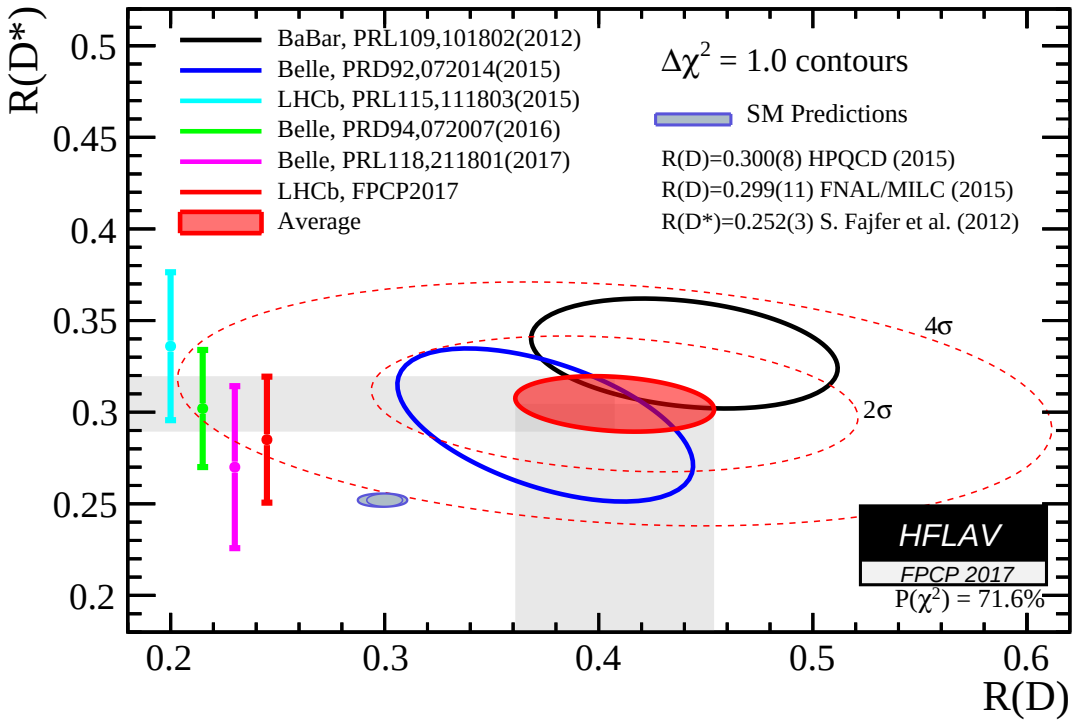
- Now in slices of BDT output

Result



- 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 $\rightarrow$ significance increases very, very slightly

Combination



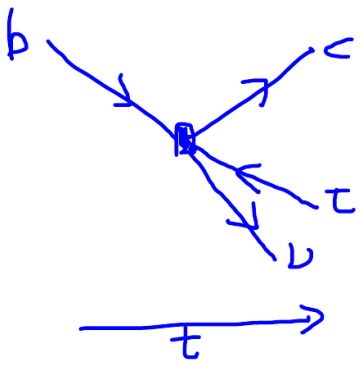
- Official HFLAV combination of $\mathcal{R}(D)$and $\mathcal{R}(D^*)$
- Excellent consistency between results
- Combined: **4.1σ tension with SM**

Is this possible??

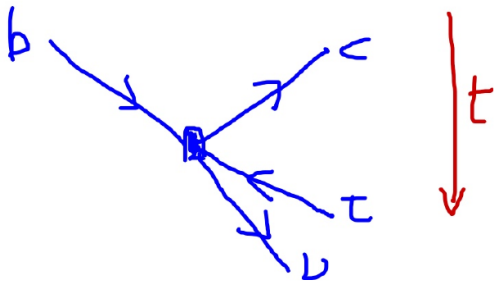
- Before introducing model dependence, only a few relevant constraints
- Could be any combination of (pseudo)scalar,(axial)vector,tensor mediator(s)
- $B^+ \to \tau\nu$ gives some constraint (particularly to scenarios with scalars)
  - Measurements not yet precise enough to be definitive
  - Uncertainty in theory prediction due to conflicting $|V_{\mathrm{ub}}|$ determinations

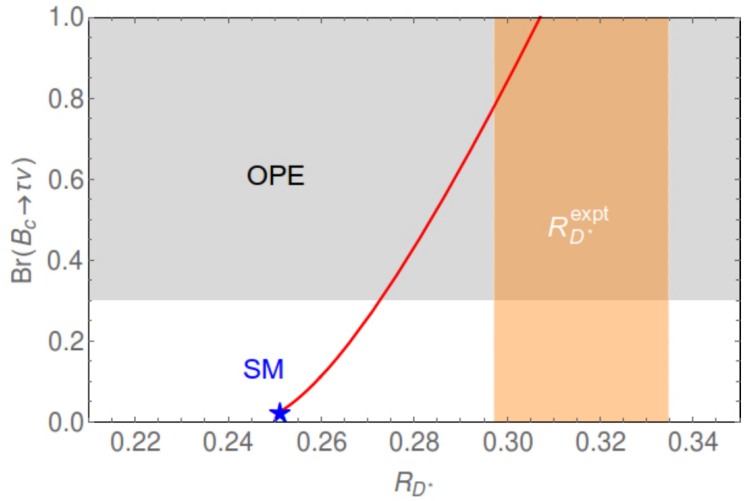
We have (sort of) measured $B_c \to \tau\nu$

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



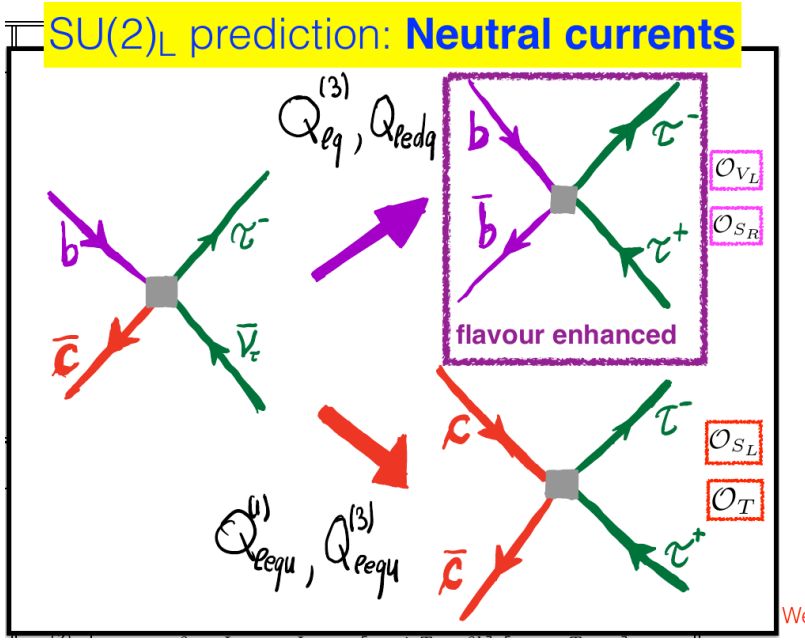
For $B_c \to \tau\nu$

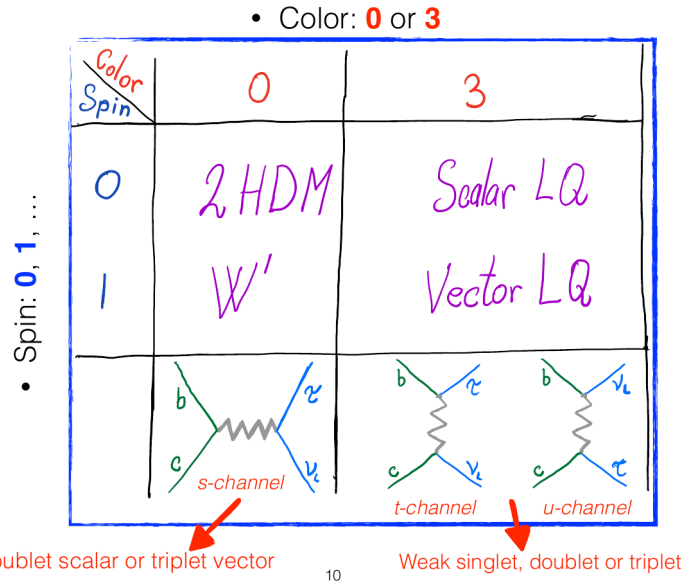




- $B_c \to \tau\nu$ is a significant fraction of total B_c width - 5% in SM
- Very sensitive to pseudoscalar operators
  - Pseudoscalar solution to $\mathcal{R}(D^{(*)})$would push up $B_c \to \tau\nu$ width enough to reduce lifetime!Alonso, Grinstein, Martin Camalich
  - Conservative bound from lifetime measurement: $\mathcal{B}(B_c \to \tau\nu) < 30\%$
  - Rules out pure pseudoscalar solution $\to$ probably not a charged Higgs

Constraints from direct searches





- Study of direct constraints by Faroughy, Greljo, Kamenik
- Strongest bounds come from recast $\tau\tau$ searches
  - For $W^{'}$ models: associated $Z^{'}$ runs into problems
  - For 2HDM: neutral Higgs' excluded
  - For (most) leptoquarks: t-channel exchange strong constraint IF suppressed by V_{cb} in $B \to D^{(*)}\tau\nu$

Models for both $\mathcal{R}(D^{(*)})$and $b \to s\ell\ell$

- Actual models also constrained by:
  - $B \to D^{(*)}\mu\nu$ / $B \to D^{(*)}e\nu$
  - $\tau \to \mu\nu\nu$ / $\tau \to e\nu\nu$
  - $B \to K^{(*)}\nu\nu$
  - Electroweak precision measurements
  - Absence of $\tau \to \mu\mu\mu$
- Is it possible for a model to evade all known constraints? Yes:
  - "B-physics anomalies: a guide to combined explanations" Buttazzo, Greljo, Isidori, Marzocca
  - General prediction - huge enhancement in $B \to K^{(*)}\tau\tau$
  - Relies on a flavour symmetry (breaking) scheme which produces large hierarchies $\to$ first hints of a theory of flavour?

What next

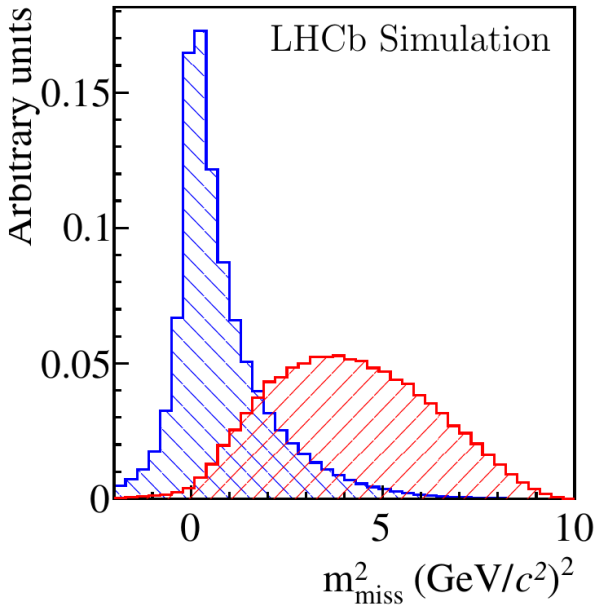
- Obvious next step: measure $\mathcal{R}(D)$
  - Underway
- Obvious next next step: run 2 data!
- Major systematics limited by MC / data size $\rightarrow$ measurements will continue to improve
- Work also underway with B_s,B_c,Λ_b
  - With muonic and hadronic tau
  - Growing program
- LHCb will have a lot more to say

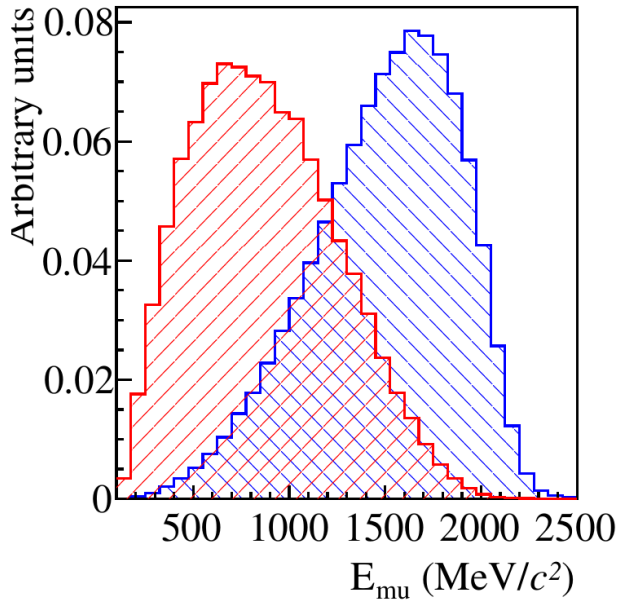
Conclusion

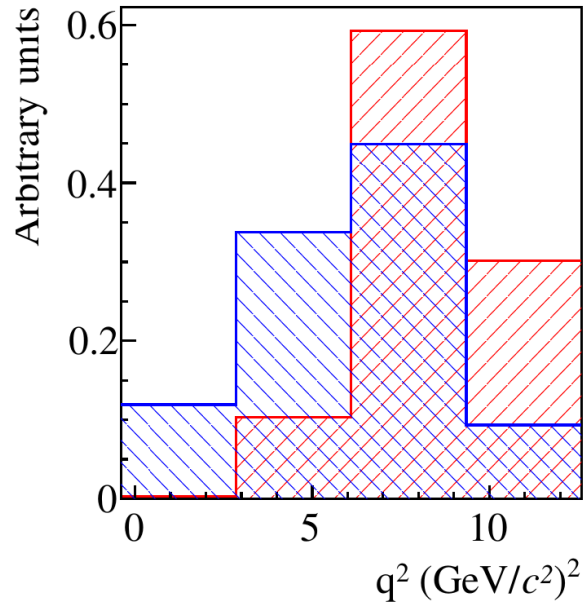
- First ever measurement of a $b \to \tau$ decay at a hadron collider: LHCB-PAPER-2015-025
- First ever $\mathcal{R}(D^*)$ measurement with $\tau \to \pi\pi\pi\nu$: LHCB-PAPER-2017-017 and LHCB-PAPER-2017-027 (soon to come)
- First ever measurement of $B_c \to J/\psi\tau\nu$: LHCb-PAPER-2017-035, soon to come
- World average for $\mathcal{R}(D^{(*)})$in $\sim 4\sigma$ tension with SM
  - Arguably the strongest current hint of new physics
- Possible to accommodate excess while evading other constraints
  - New physics "not much above" 1 TeV...
- Exciting times (not far) ahead

Backups

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

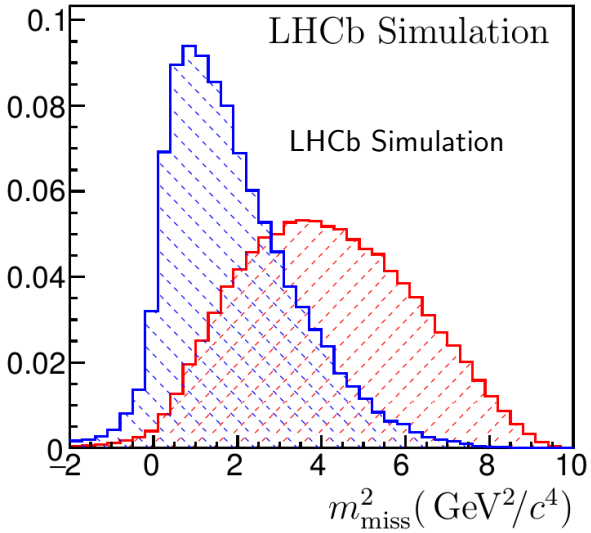


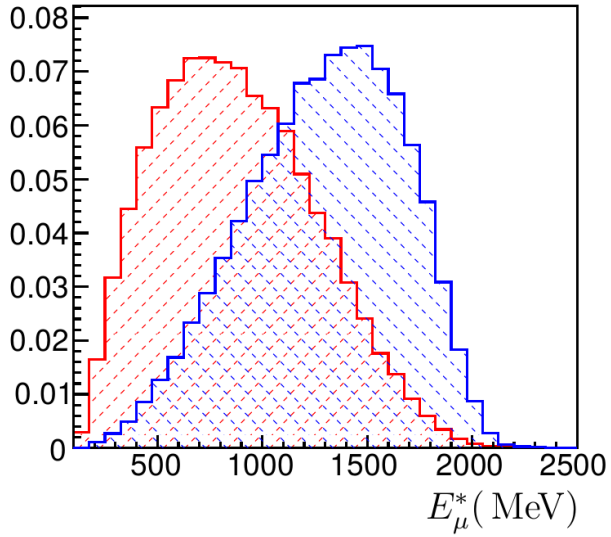


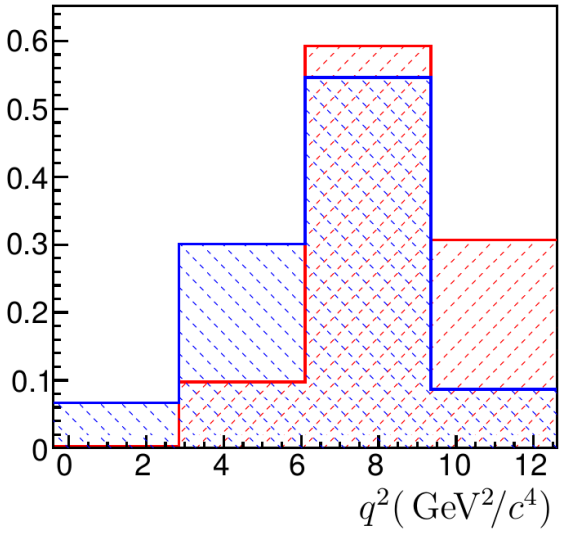


- $B \to D^* \mu \nu$ (black) vs $B \to D^* \tau \nu$ (red)
- $B \to D^* \mu \nu$ is both the normalisation mode, and the highest rate background ($\sim 20 \times B \to D^* \tau \nu$)
  - Use CLN parameterisation for form factors
  - Float form factors parameters in fit $\to$ uncertainty taken into account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

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

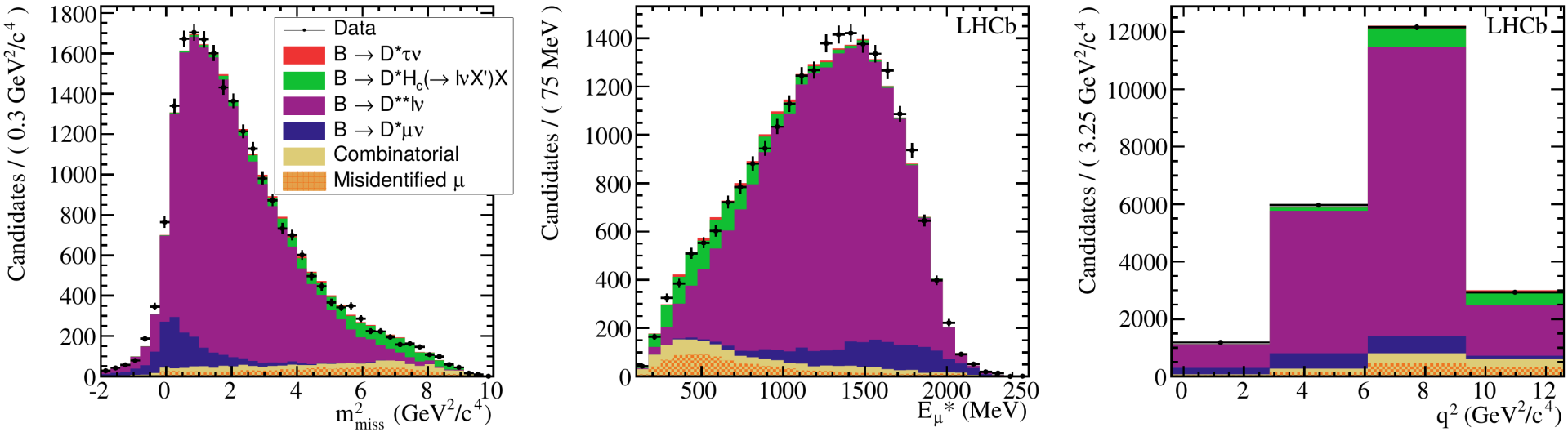






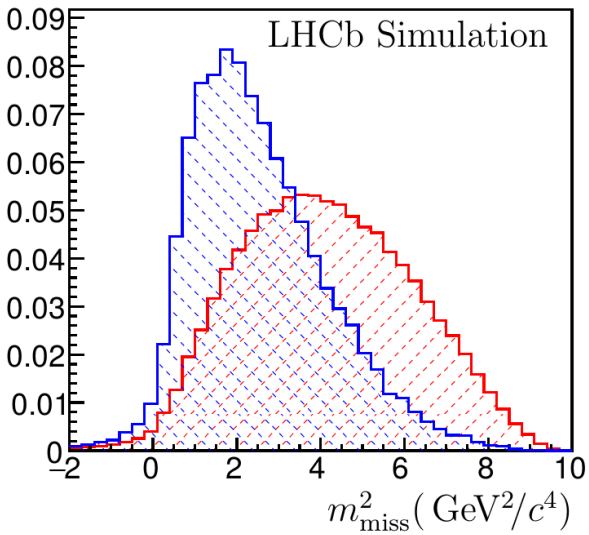
- $B \to D^{**}\mu^+\nu$ refers to any higher charm resonances (or non resonant hadronic modes)
- Not so well measured
  - Set of states comprising D^{**} known to be incomplete
  - Decay models not well measured
- For the established states (shown in black):
  - Separate components for each resonance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

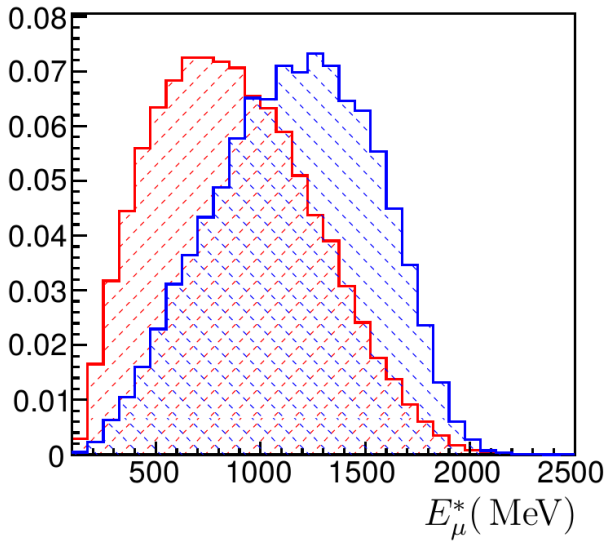
$B \to D^{**}(\to D^{*+}\pi)\mu\nu$ control sample

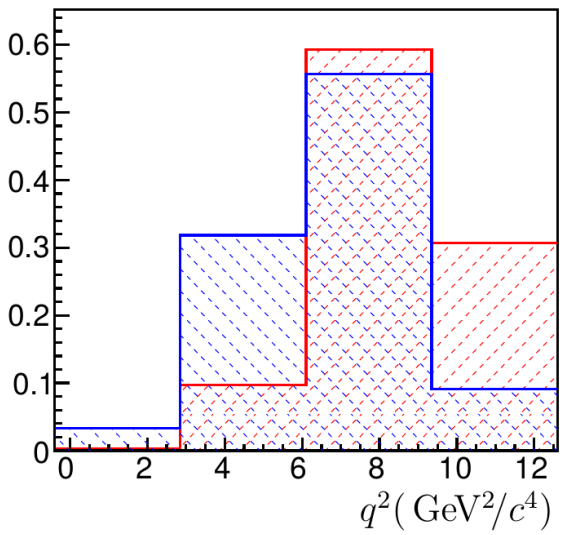


- Isolation MVA selects one track, $M_{D^{*+}\pi}$ around narrow D^{**} peak $\to$ select a sample enhanced in $B \to D^{**}\mu^+\nu$
  - Use this to constrain, justify $B \to D^{**}\mu^+\nu$ shape for light D^{**} states
  - Also fit above, below narrow D^{**} peak region to check all regions of $M_{D^{*+}\pi}$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

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

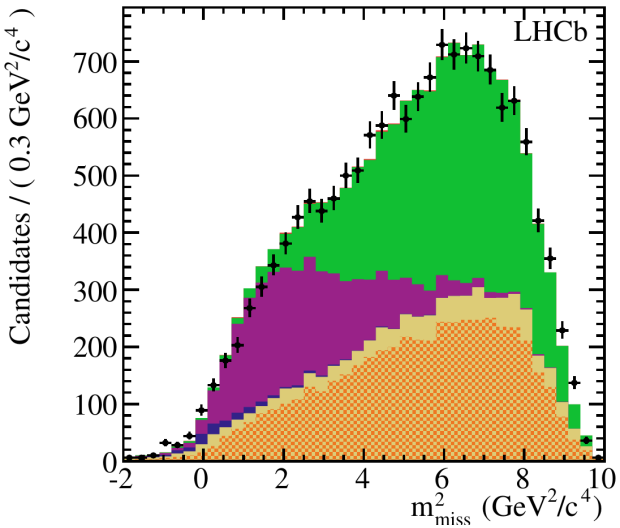


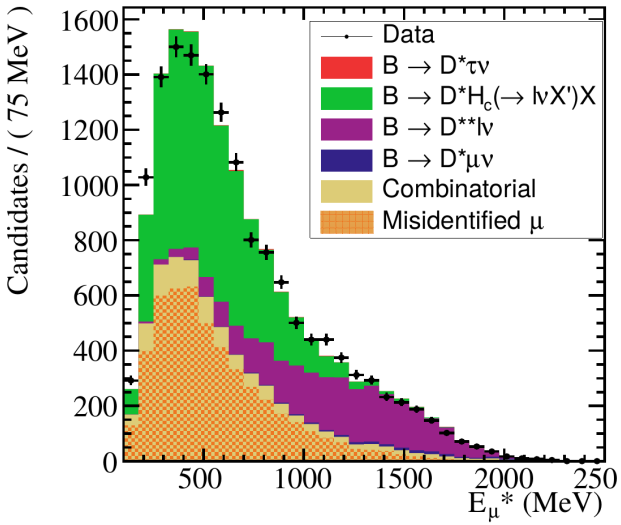


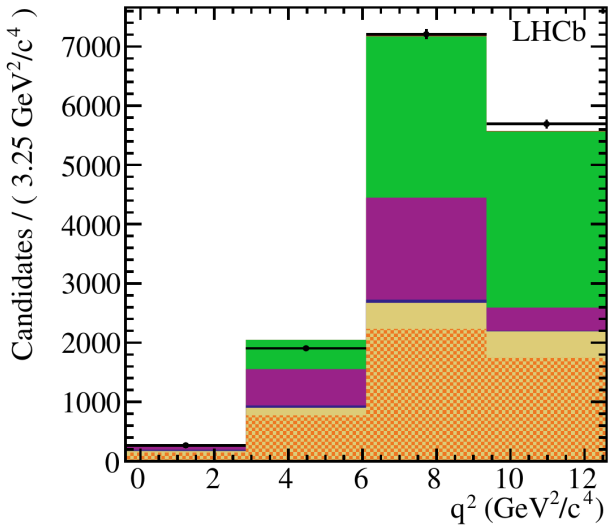


- Previously unmeasured $B \to D^{**}(\to D^{*+}\pi\pi)\mu\nu$ contributions recently measured by BaBar
  - Too little data to separate individual (non)resonant components
  - Single fit component, empirical treatment
- Constrain based on a control sample in data
  - Degrees of freedom considered: D^{**}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

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

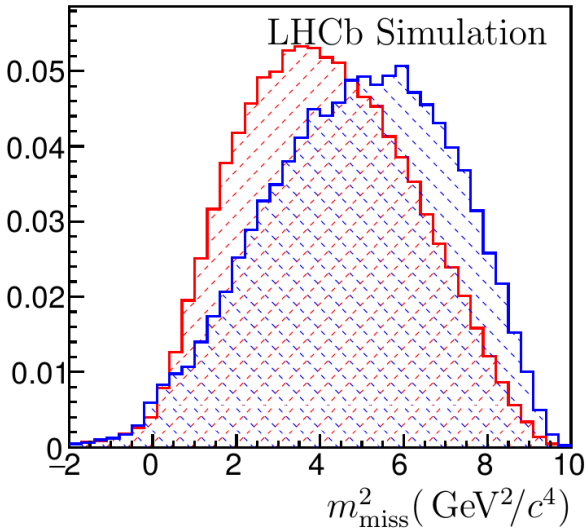


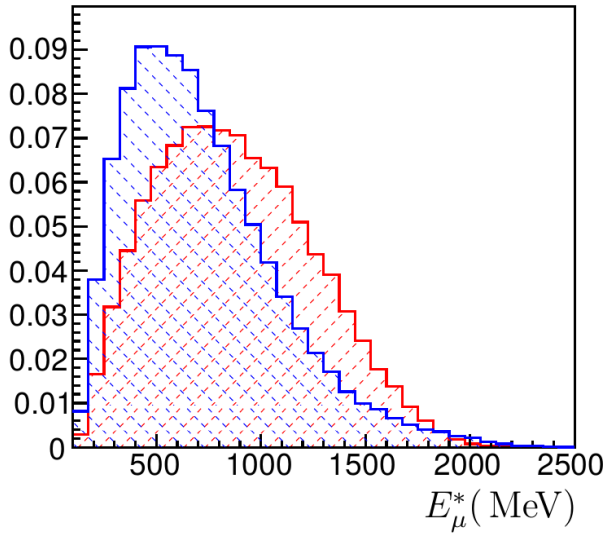


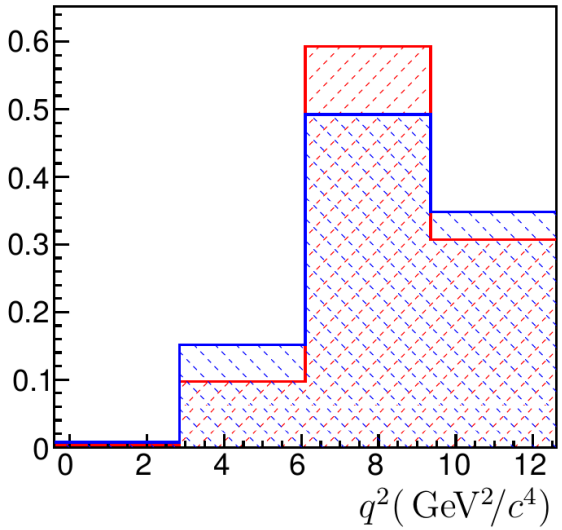


- Also look for two tracks with isolation MVA $\to$ study $B \to D^{**}(\to D^{*+}\pi\pi)\mu\nu$ in data
- Can control shape of this background

$B \to D^* DX$

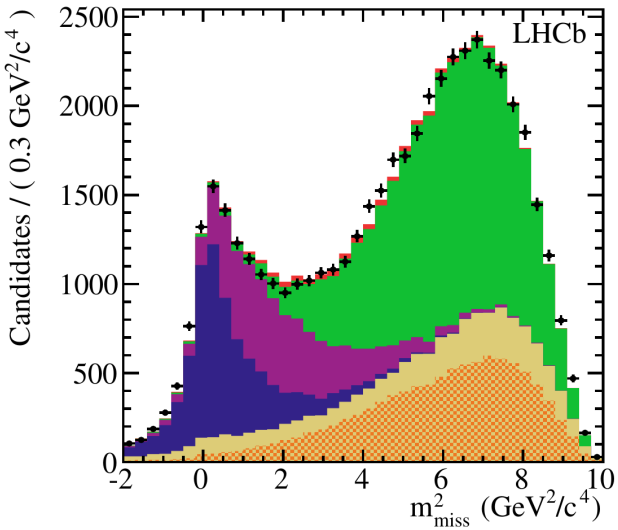


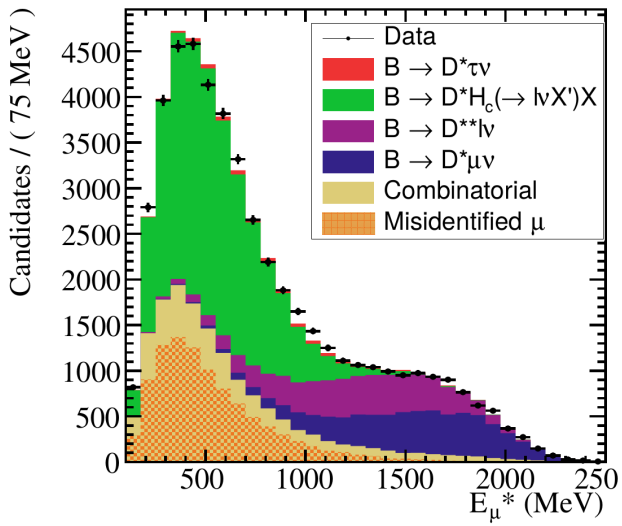


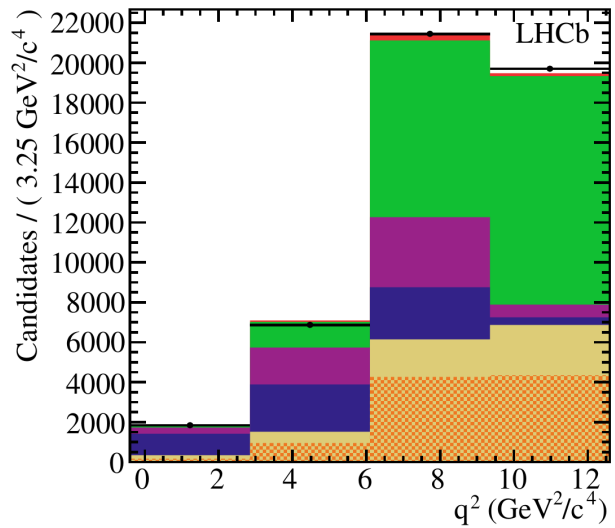


- $B \to D^* DX$ consists of a very large number of decay modes
  - Physics models for many modes not well established
- Constrain based on a control sample in data
- Single component, empirical treatment
  - Consider variations in M_{DD}
  - Multiply simulated distributions by second order polynomials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

$B \to D^* DX$ control sample

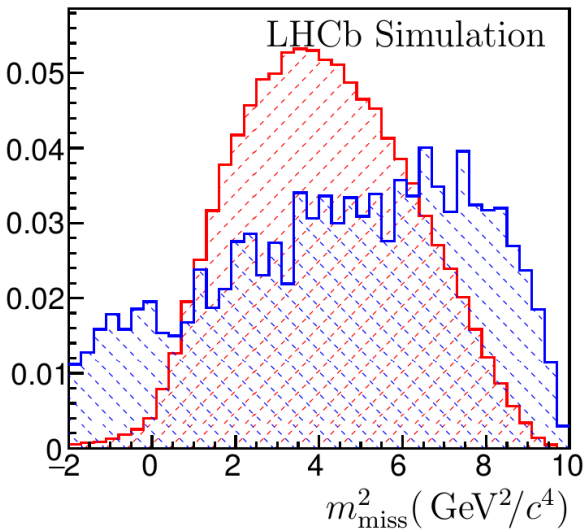


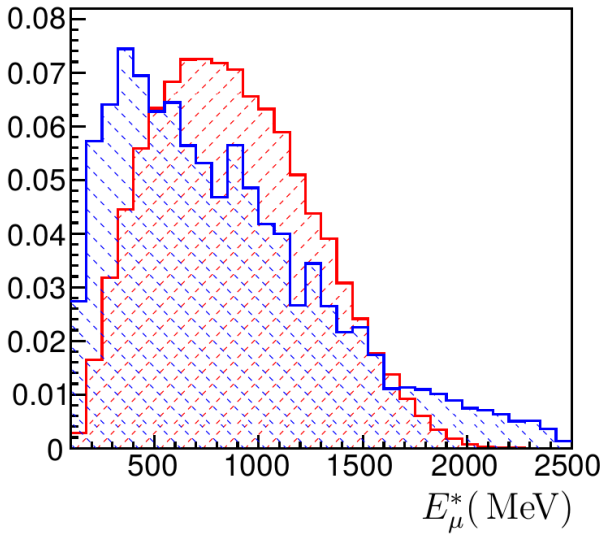


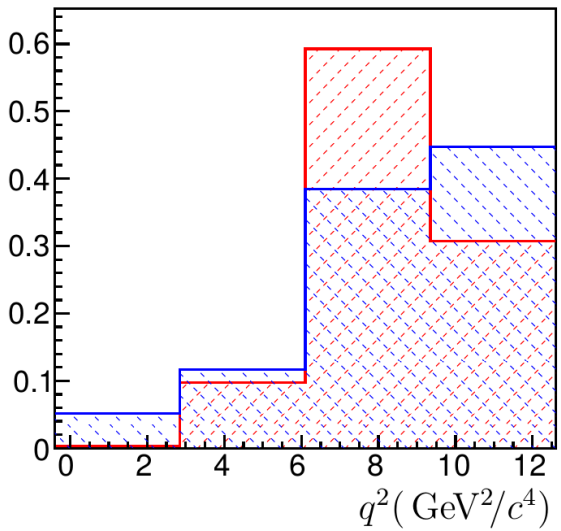


- Isolation MVA selects a track with loose kaon ID $\to$ select a sample enhanced in $B \to D^* DX$
- Use this to constrain, justify $B \to D^* DX$ shape

Combinatorial backgrounds

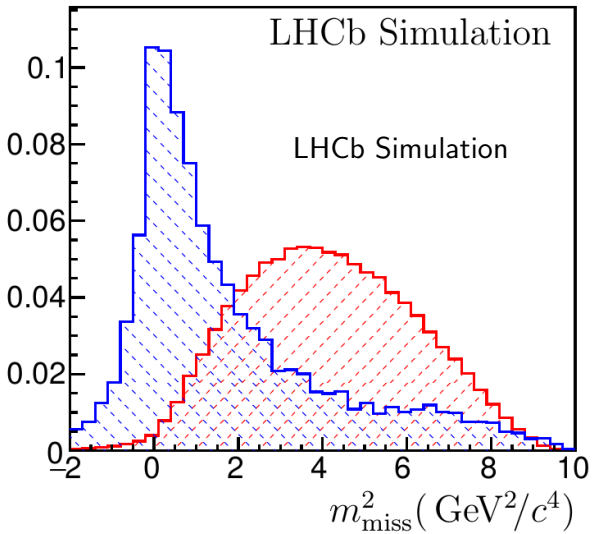


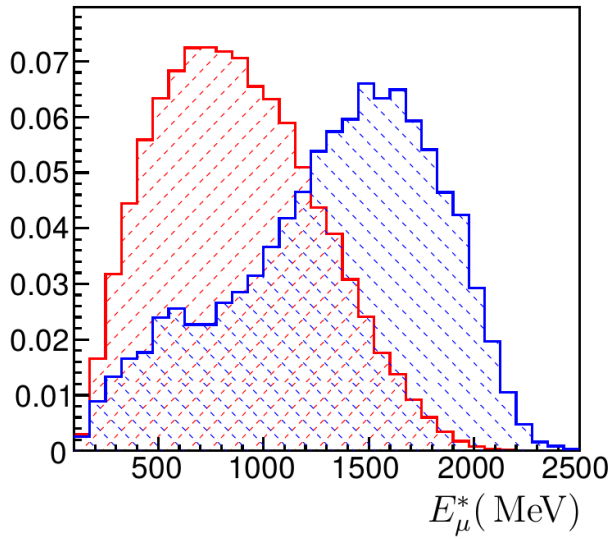


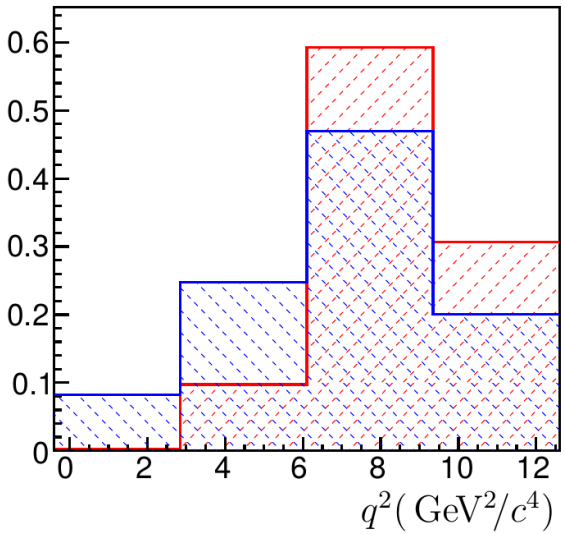


- Combinatorial background modelled using same-sign $D^{*+}\mu^+$ data
- Two sources of combinatorial background are treated separately (shown on next slide)

Combinatorial backgrounds

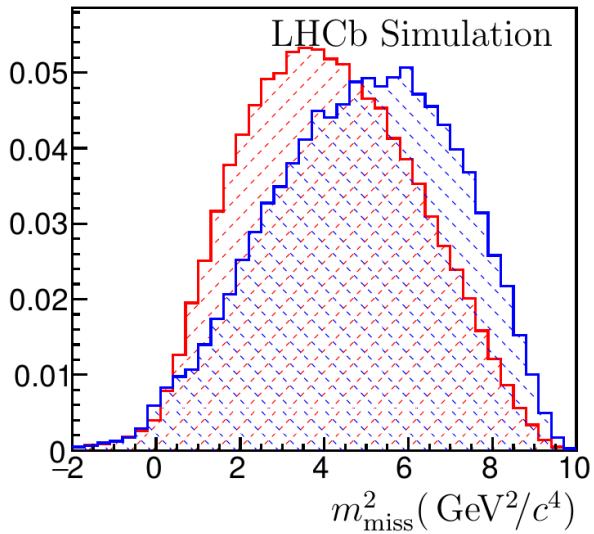


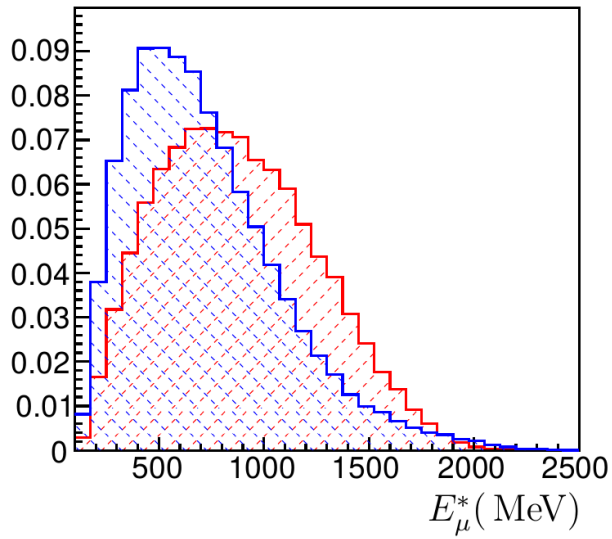


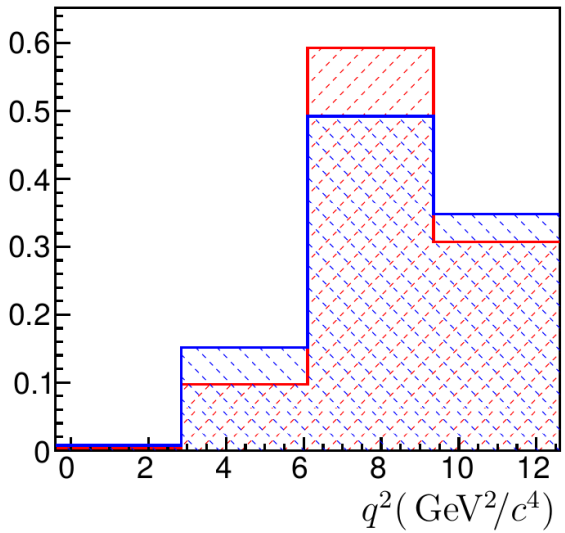


- Non D^{*+} backgrounds (fake D^*) template modelled using $D^0\pi^-$ data (shown)
  - Yield determined from sideband extrapolation beneath D^{*+} mass peak
- Hadrons misidentified as muons (fake muons)
  - Controlled using $D^{*+}h^\pm$ sample
  - Both template and expected yield can be determined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

- Statistical uncertainty on $\mathcal{R}(D^*)$ (fixing all templates to nominal shapes): 2.7% (absolute)
- 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