

A Supplementary Material

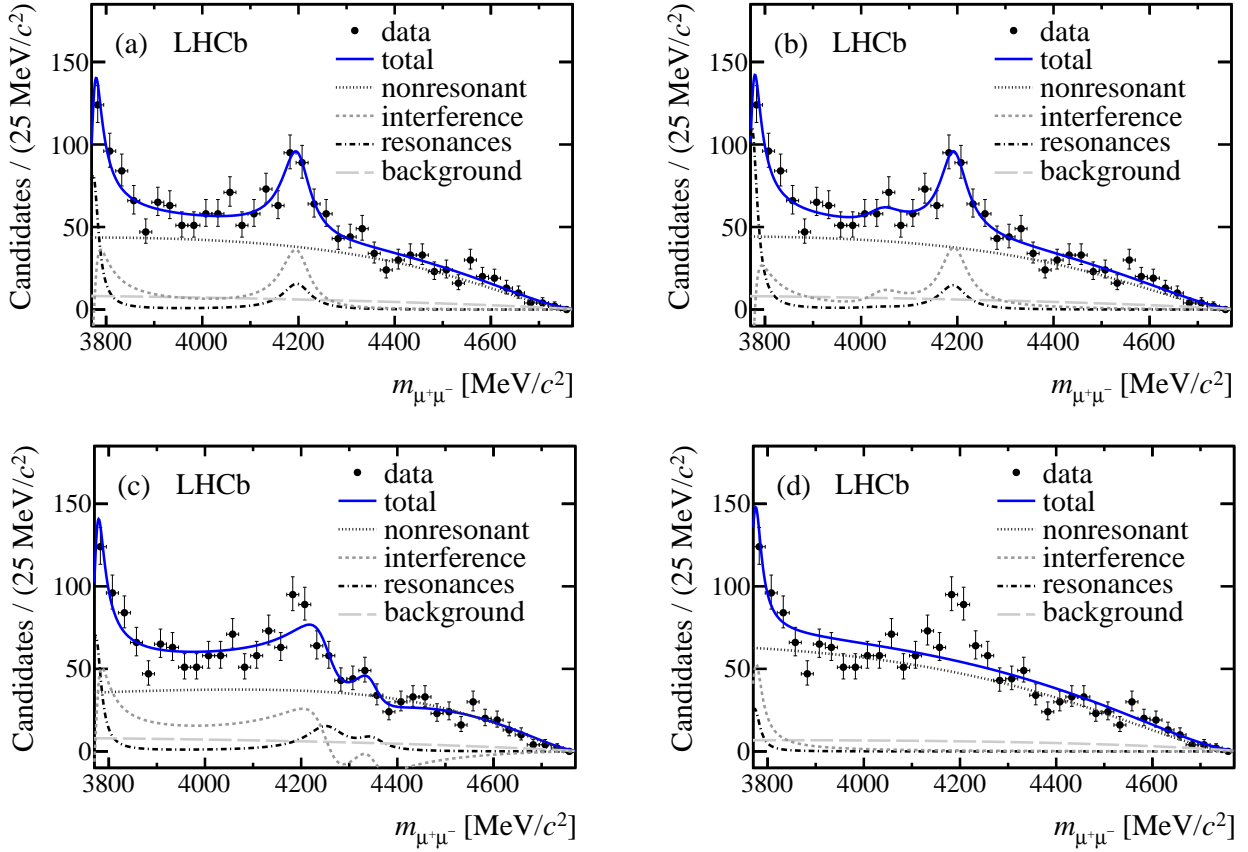


Figure 1: Dimuon mass distribution of data with fit results overlaid for the fits with: (a) a single extra resonance whose mean width and phase are left free in the fit; (b) the $\psi(4040)$ and $\psi(4160)$ resonances with the relative strong phases free in the fit; (c) the $Y(4260)$ and $Y(4350)$ with the relative strong phases free in the fit; (d) no additional resonance. In all fits the contributions from the non-resonant vector and axial components, the $\psi(3770)$ and the interference terms are included.

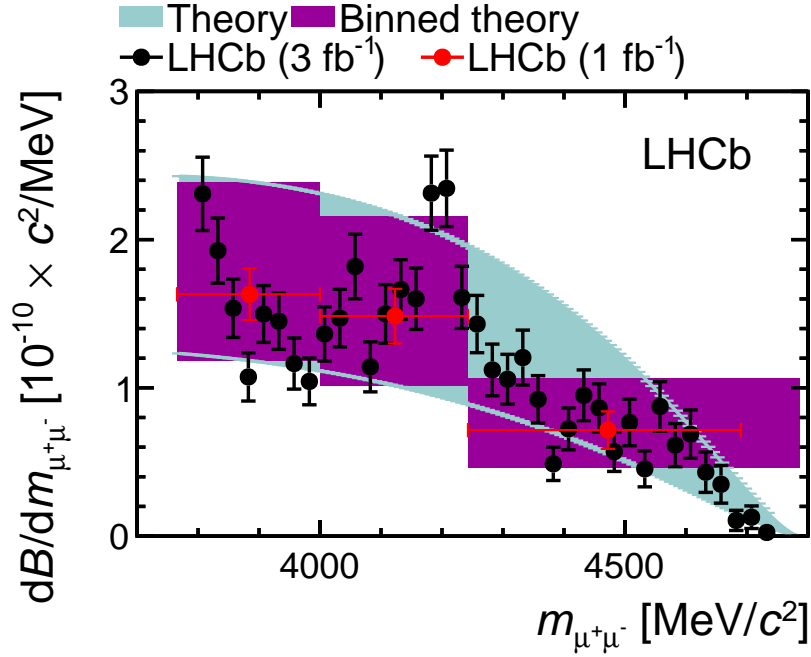


Figure 2: Differential branching fraction of the $B^+ \rightarrow K^+ \mu^+ \mu^-$ decay in the low recoil region. The 3 fb^{-1} result used for this analysis is overlaid on the data from the LHCb analysis based on 1 fb^{-1} [?]. For the 3 fb^{-1} result only statistical uncertainties are shown. The bands shown indicate the theoretical prediction for the differential branching fraction and are calculated using input from Ref. [?].

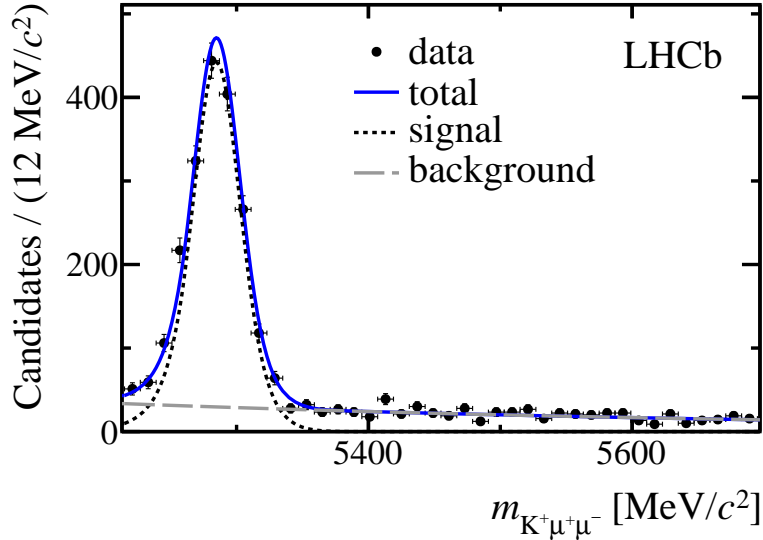


Figure 3: Mass fit to $B^+ \rightarrow K^+ \mu^+ \mu^-$ decays for $m_{\mu^+ \mu^-} > 3770 \text{ MeV}/c^2$. The signal region defined as $5240 < m_{K^+ \mu^+ \mu^-} < 5320 \text{ MeV}/c^2$ contains 1830 candidates with a signal to background ratio of about 8.

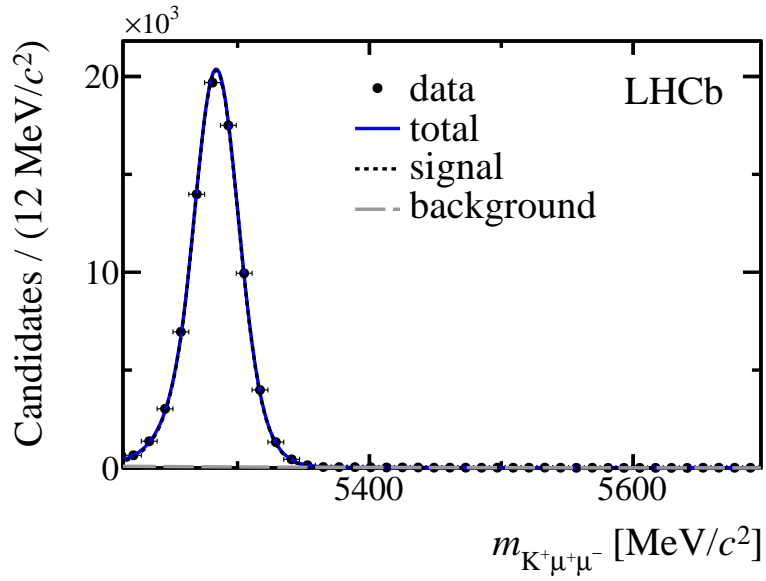


Figure 4: Mass fit to $B^+ \rightarrow \psi(2S)K^+$ decays.

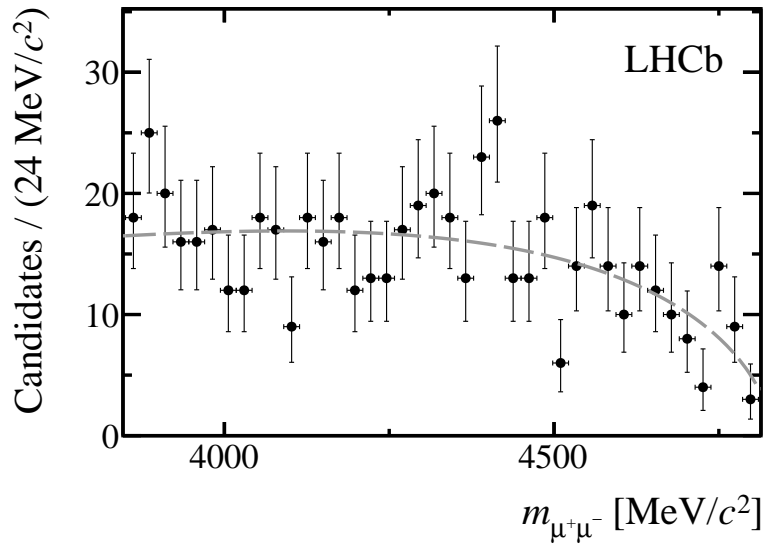


Figure 5: Fit to the $m_{\mu^+\mu^-}$ distribution in the $B^+ \rightarrow K^+\mu^+\mu^-$ sideband region defined as $5350 < m_{K^+\mu^+\mu^-} < 5500 \text{ MeV}/c^2$. The shape is parametrised by an ARGUS function.

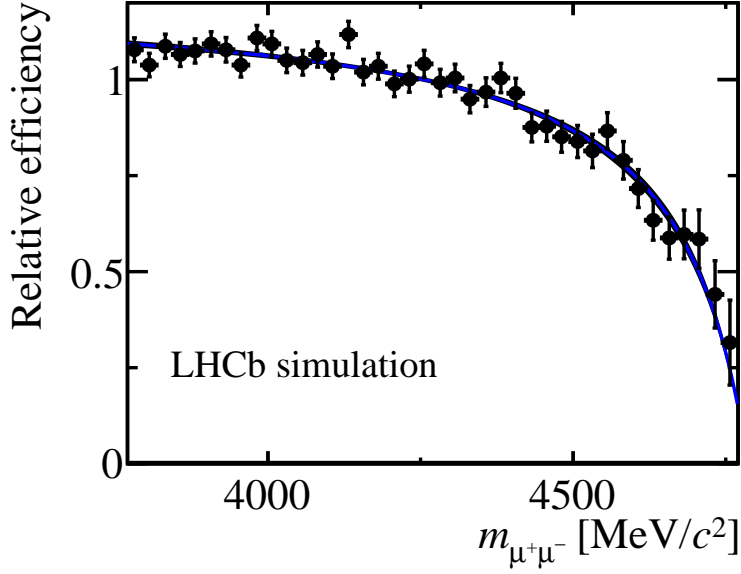


Figure 6: Relative efficiency between $B^+ \rightarrow K^+ \mu^+ \mu^-$ and $B^+ \rightarrow J/\psi K^+$ as a function of $m_{\mu^+\mu^-}$. The shape is parametrised with a single parameter threshold model $\varepsilon(m_{\mu^+\mu^-}) \propto 1 - 1/[1 + \kappa \cdot \delta(m_{\mu^+\mu^-})]$, where $\delta(m_{\mu^+\mu^-}) = (m_{B^+} - m_{K^+} - m_{\mu^+\mu^-}) / (m_{B^+} - m_{K^+})$ and κ is a free parameter in the fit.

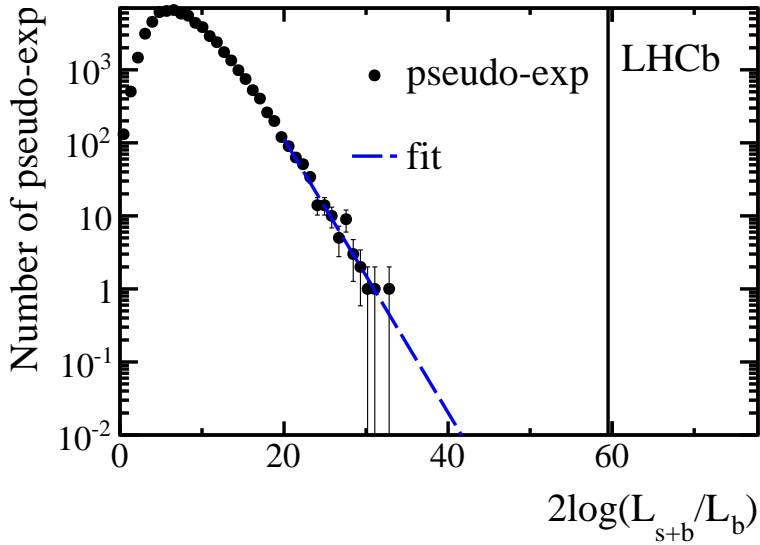


Figure 7: Distribution of the log likelihood ratio from fits that include and exclude a single additional resonance component for 6×10^5 pseudo-experiments that only include the non-resonant, $\psi(3770)$ and background components. The black points denote the number of pseudo-experiments binned in $2 \log(L_{s+b}/L_b)$ with a bin size of 0.875. The dashed blue line corresponds to the fit of an exponential function which is used to extrapolate to larger values of $2 \log(L_{s+b}/L_b)$. The vertical black line denotes the observed likelihood ratio. A signal significance exceeding six standard deviations is observed.