

# 1 Supplementary material

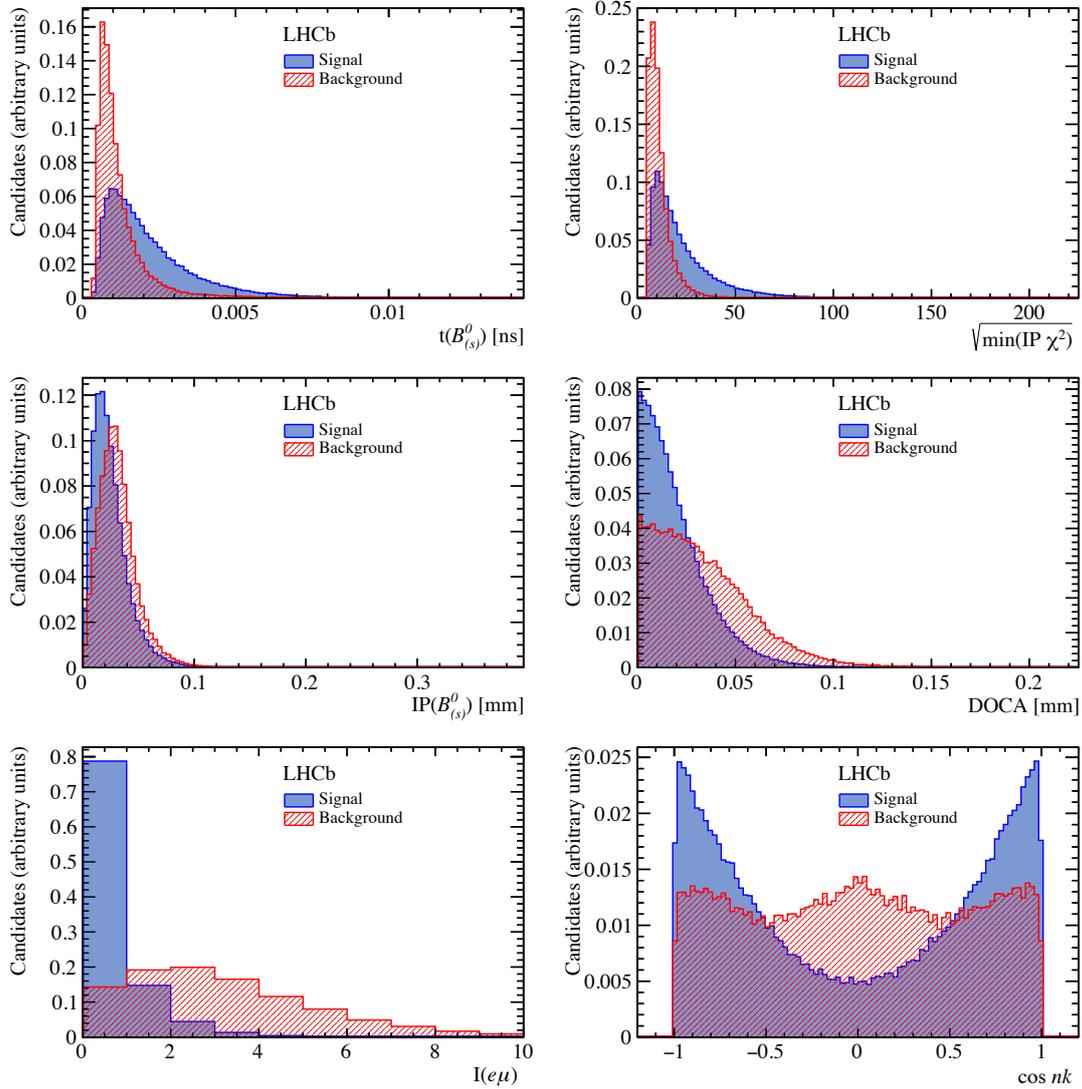


Figure 6: Distributions of six out of twelve variables used in the BDT training for simulated signal events (blue) and background events from opposite-sign data sidebands (red). From left to right and from top to bottom: the proper  $B_s^0$  lifetime,  $t(B_s^0)$ ; the square root of the minimum impact parameter  $\chi^2$  for the two tracks,  $\sqrt{\min(\text{IP}\chi^2)}$ ; the impact parameter of the  $B_s^0$ ,  $\text{IP}(B_s^0)$ ; the distance of closest approach between the two daughter tracks, DOCA; the isolation of the two tracks with respect to any other track in the event  $I(e\mu)$ , defined as the number of tracks in a cone of angle 0.27 rad around the reconstructed track (electron or muon); the cosine of the angle between the muon momentum in the B rest frame and the vector perpendicular to the  $B_s^0$  candidate momentum and the beam axis,  $\cos nk$ .

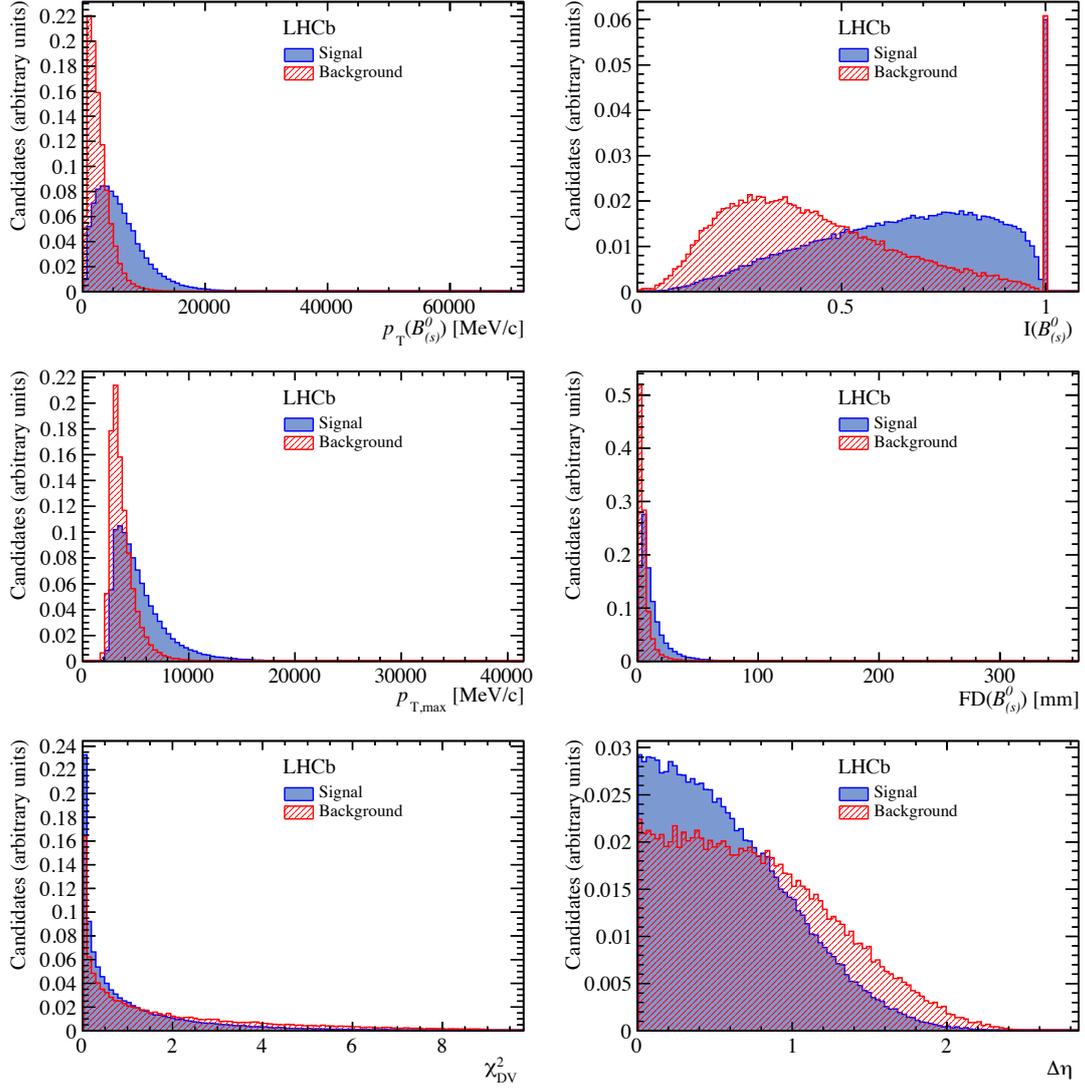


Figure 7: Distributions of six out of twelve variables used in the BDT training for simulated signal events (blue) and background events from opposite-sign data sidebands (red). From left to right and from top to bottom: the transverse momentum of the  $B_s^0$  candidate,  $p_T(B_{(s)}^0)$ ; the isolation of the  $B_s^0$  candidate,  $I(B_{(s)}^0)$ ; the maximum transverse momentum of the two daughter tracks,  $p_{T,max}$ ; the flight distance of the  $B_s^0$  candidate with respect to its primary vertex,  $FD(B_{(s)}^0)$ ; the  $\chi^2$  of the decay vertex of the  $B_s^0$  candidate,  $\chi^2_{DV}$ ; the difference of pseudo-rapidity between the two daughter tracks,  $\Delta\eta$ .

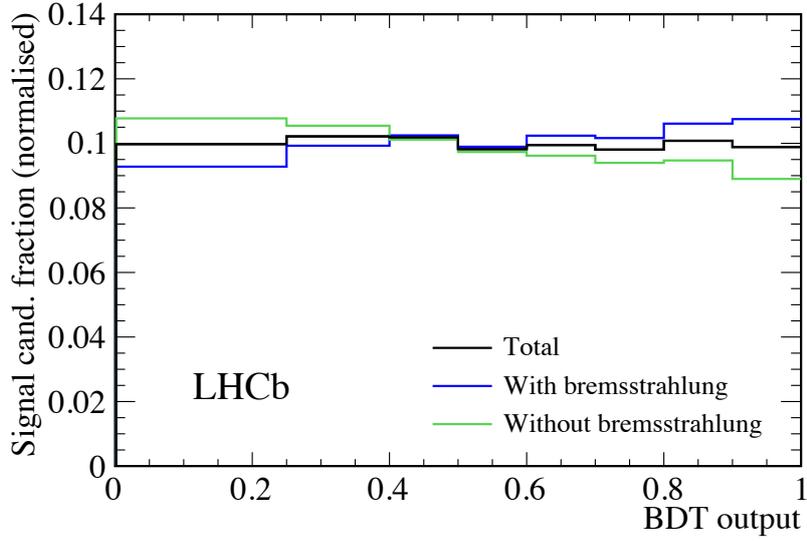


Figure 8: Shape of the BDT response for candidates with and without recovered bremsstrahlung photons and for the whole sample.

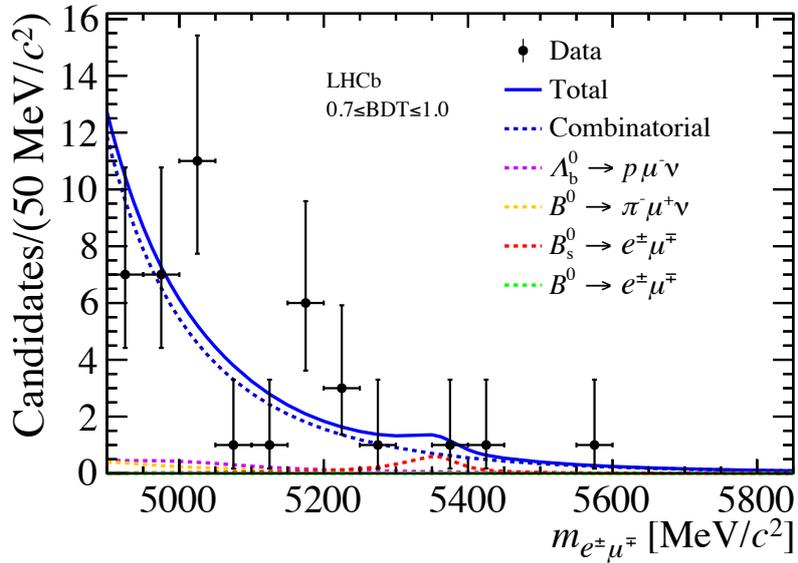


Figure 9: Invariant mass distributions of  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  candidates,  $m_{e^\pm \mu^\mp}$ , integrated over the three most sensitive BDT bins and the two bremsstrahlung categories.

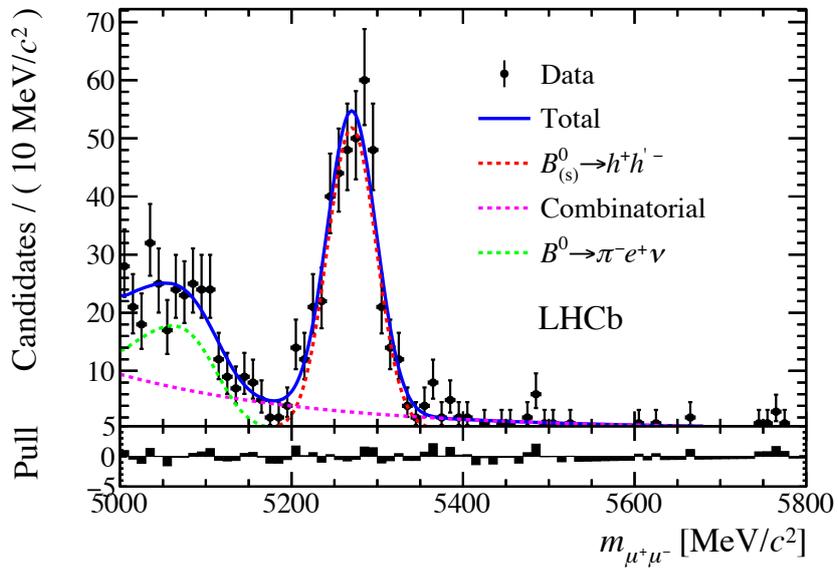


Figure 10: Distribution of two-track invariant mass with fit superimposed of the  $B \rightarrow h^+ h'^-$  sample where one track satisfies a pion particle identification requirement and the other track satisfies the  $B_{(s)}^0 \rightarrow e^\pm \mu^\mp$  electron particle identification requirement. The mass hypothesis used for the two tracks is the muon mass.