Supplementary material

In Figs. 1 and 2 the fit to the fully selected normalization mode is displayed. The distributions of the BDT classifier are shown in Fig. 3. They illustrate on the one hand the good separation power of the classifier, and on the other hand a clear difference between the upper and lower mass sideband, which arises from specific signal-like background decays present in the lower mass sideband. The dielectron invariant-mass distribution of $B_s^0 \to e^+e^-$ and $B^0 \to e^+e^-$ decays in simulation is shown in Fig. 4. In Fig. 5 the data and fit model are displayed enhancing $\mathcal{B}(B_s^0 \to e^+e^-)$ to the maximally allowed value in Ref. [1]. The model shows a clear peak that is not present in the data.

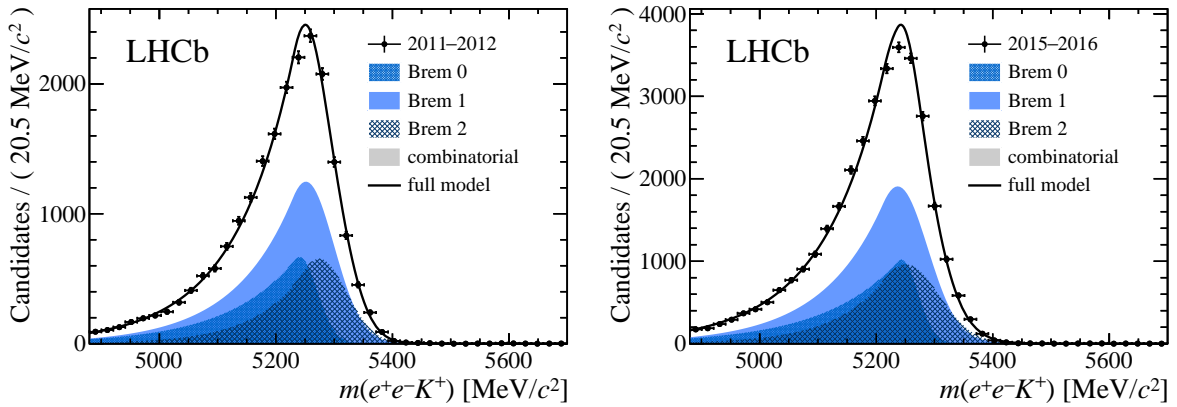Figure 1: Fits to the invariant-mass distribution of the normalization mode $B^+ \to J/\psi\, K^+$ in the (left) Run 1 and (right) Run 2 data sets. Background components are found to be negligible.
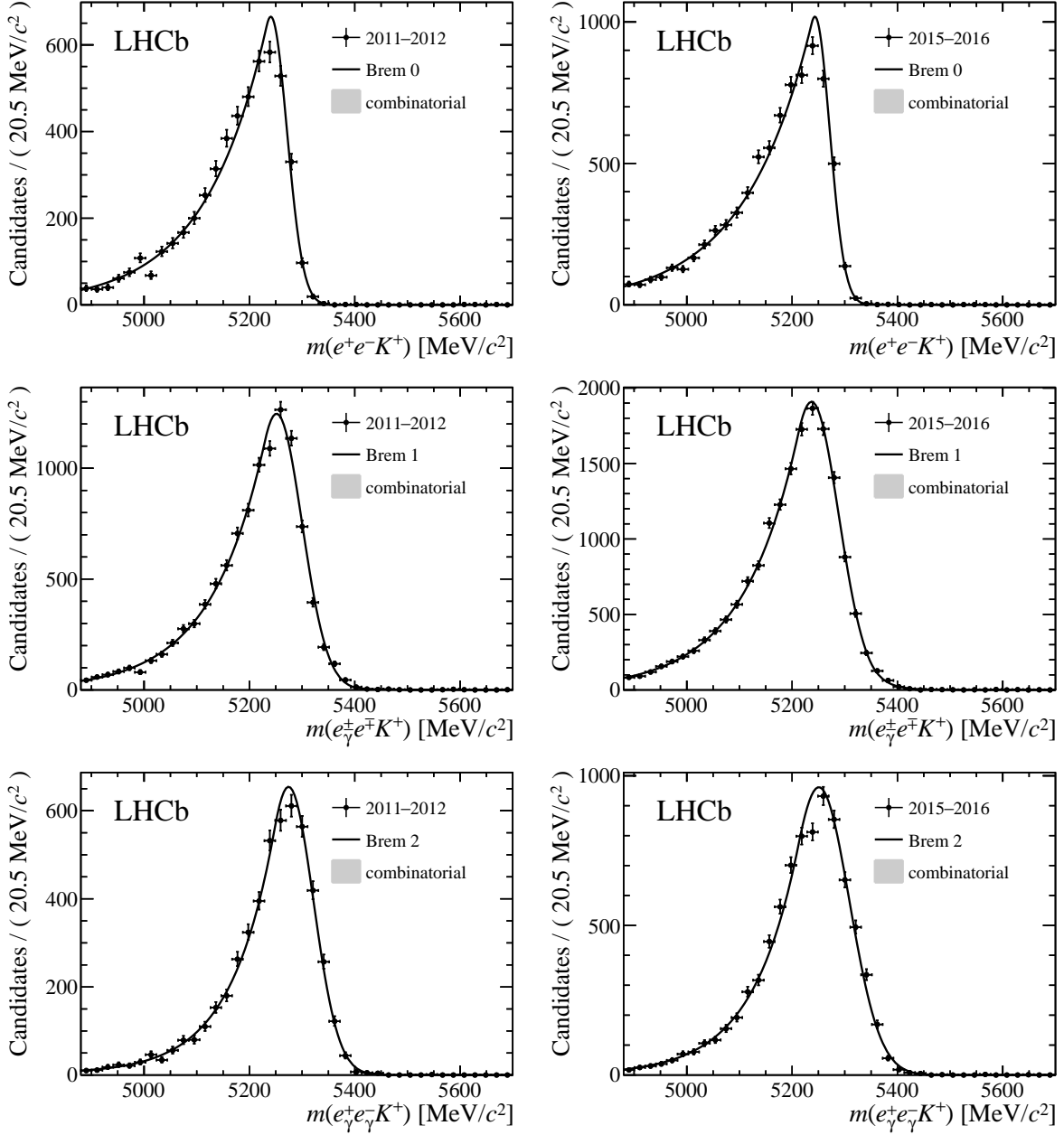
Figure 2: Fits to the invariant-mass distribution of the normalization mode $B^+ \to J/\psi\, K^+$ in the (left) Run 1 and (right) Run 2 data sets, split by bremsstrahlung category. From top to bottom, the data sets correspond to the bremsstrahlung correction category of no correction, correcting one electron and correcting both electrons. Background components are found to be negligible.
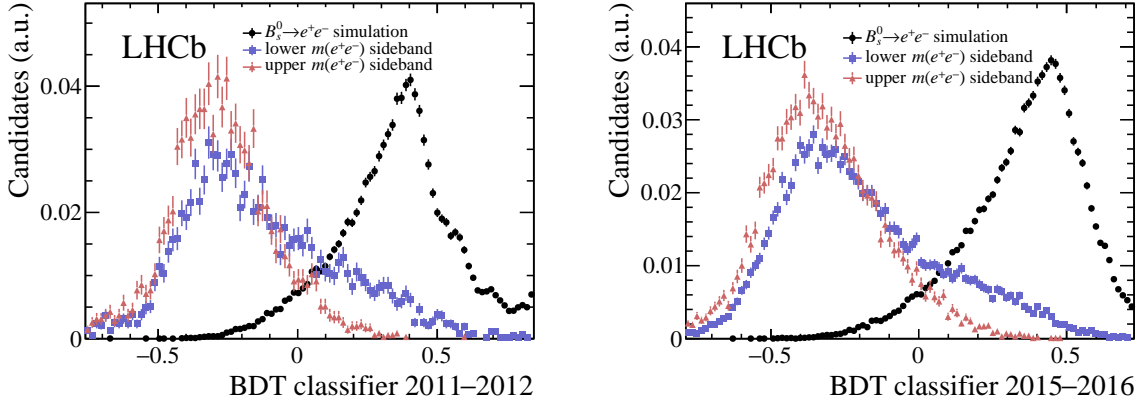
2

Figure 3: BDT distributions of corrected signal simulation and the dielectron mass sidebands for the (left) Run 1 and (right) Run 2 classifier. They illustrate the good separation power of the classifier. The classifier distribution is checked to not depend on the dielectron invariant-mass. Differences between the upper and lower dielectron mass sideband arise due to specific signal-like background decays present in the lower mass sideband.
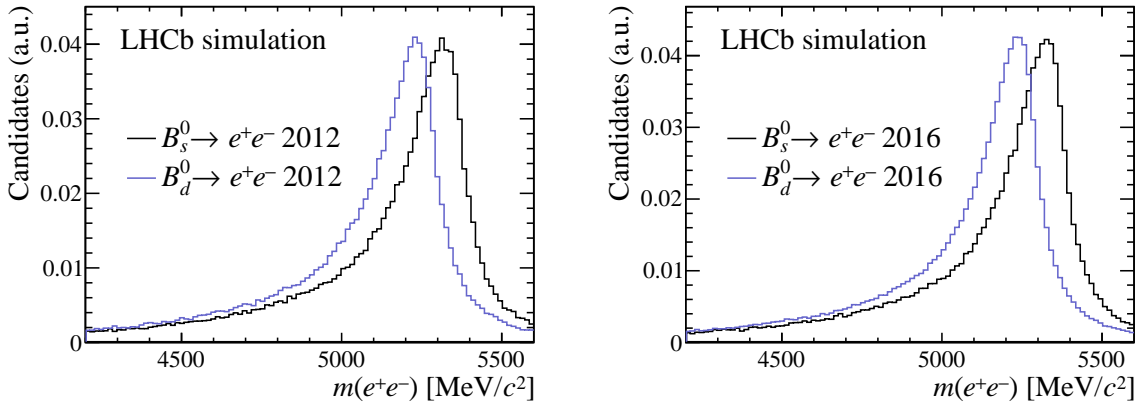

Figure 4: Dielectron invariant-mass distribution of $B_s^0 \to e^+e^-$ and $B^0 \to e^+e^-$ decays in simulation for (left) 2012 and (right) 2016 after the full selection.

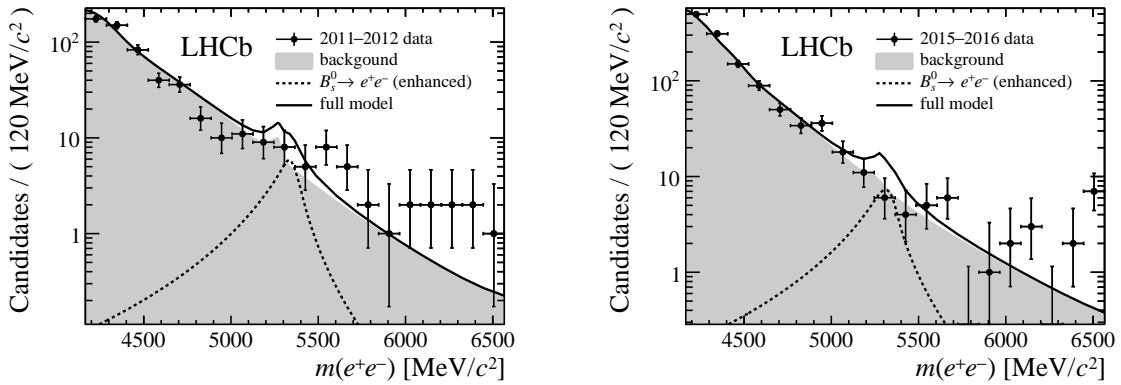Figure 5: Model of the dielectron invariant-mass distribution with $\mathcal{B}(B^0 \to e^+e^-)$ fixed to zero and $\mathcal{B}(B^0_s \to e^+e^-)$ enhanced to the maximally allowed value in Ref. [1]. The sum of bremsstrahlung categories is shown for (left) Run 1 and (right) Run 2 data sets. The relative proportions of background contributions change between Run 1 and Run 2 due to different performances of the particle-identification algorithms and BDT selections.

References

[1] R. Fleischer, R. Jaarsma, and G. Tetlalmatzi-Xolocotzi, *In pursuit of New Physics with $B_{s,d}^0 \to \ell^+\ell^-$*, JHEP **05** (2017) 156, `arXiv:1703.10160`.