Suppressing beam background using `FastBDT` at Belle II

Priyanka Cheema [1] Racha Cheaib [2]

[1]University of Sydney [2]DESY

WHAT IS BELLE II?

The Belle II detector is situated inside an **electron-positron collider** in Tsukuba, Japan called **SuperKEKB**. It began taking data in 2019 and has recorded over 200 fb^{-1} of data to date. SuperKEKB is built to collide positrons and electrons at energies of **4 GeV and 7 GeV** respectively, with the aim of producing an $\boldsymbol{\Upsilon(4S)}$ meson at the interaction point (IP). The high energies are required to ensure that the CM energy $\sqrt{s}$ exactly equals the $\Upsilon(4S)$ rest mass of **10.58 GeV**. Once created, the $\Upsilon(4S)$ then decays **almost exclusively into a pair of** B **mesons** (denoted $B\bar{B}$). The **Belle II detector** surrounds the IP, and particles produced in collisions at the IP are reconstructed using data gathered by the different sub-detectors of Belle II. A diagram of SuperKEKB and Belle II along with its sub-detectors is given in Fig. 1.

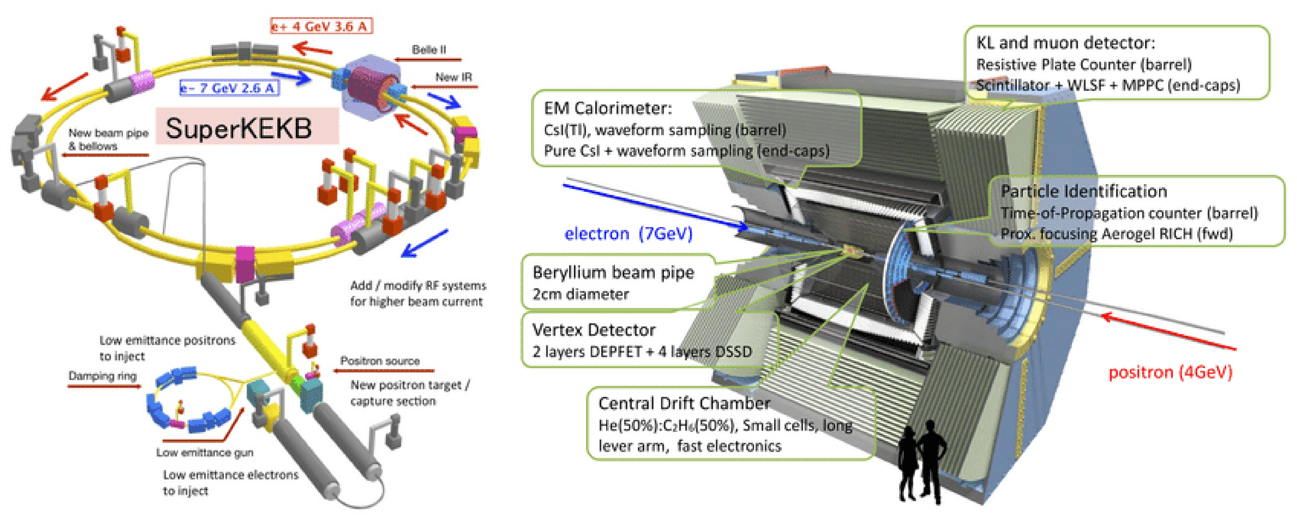

Figure 1: Schematic view of the SuperKEKB collider (left) and the Belle II detector (right) [2]

SuperKEKB and Belle II were built to study **CP violation in the** B **meson system**, rare B decays and to perform precision measurements for Standard Model parameters. It is also used as a τ and charm factory, with many studies on τ and charm physics being done alongside searches for dark matter.

WHAT IS BEAM BACKGROUND?

Beam backgrounds are induced at Belle II by electron and positron **beam interactions** that occur as they travel around the beam pipe and crossover at the IP. There are 3 different main sources of beam backgrounds: Touschek scattering, Bhabha scattering and beam-gas scattering

- These beam interactions create particle showers that leave signal-like signatures in sub-detectors - our focus is on those that get reconstructed as photons by the **electromagnetic calorimeter (ECL)**
- **Problematic** as they can be incorrectly reconstructed as photons that originate from a $B\bar{B}$ event at the IP **(a signal photon)**

Beam background is **highly dependent on the instantaneous luminosity**, and as this is continually increasing for SuperKEKB, so will the level of beam background. Therefore, analysis tools that **suppress beam background** are incredibly important in ensuring this problem is manageable.

Impact of beam background on $\mathrm{E_{ECL}}$

For many B decay analyses, a **key variable** used for signal extraction and background suppression is E_{ECL}. It is equal to the total energy sum of clusters in the ECL not accounted for during the event reconstruction. For events where all the decay products of the B meson are properly reconstructed (signal), E_{ECL} is **peaked close to 0**. However, for events where additional background processes or decays occur, E_{ECL} **shift from zero**. This is what provides a clean separation between signal and background events. Improving the resolution of the E_{ECL} signal peak is advantageous for suppressing background and improving the significance of measurements. E_{ECL} is **significantly impacted by beam background** as it broadens the signal distribution and shifts its peak away from 0. This is because clusters resulting from beam background processes remain unaccounted for during reconstruction. Thus, **suppressing beam background is key** as it helps to improve the signal E_{ECL} resolution which helps improve signal reconstruction.

THE BEAM BACKGROUND MVA

The architecture chosen for our beam background MVA is the `FastBDT` which is an **ensemble consisting of stochastic gradient boosted decision trees** [1]. The MVA was trained on 1,167,178 photons, equally split between signal and beam background photons. Signal photons were reconstructed from **Monte Carlo (MC) generated** $\mathbf{B\bar{B}^0}$ **events**. For beam background clusters, $\mathbf{e^+e^- \to \mu^+\mu^-}$ events were reconstructed from 62.8 fb^{-1} of Belle II data, and all the clusters found in the **rest of the event (ROE)** were labelled as beam background. This sample has ROE clusters come from both beam background and **initial state/final state radiation (ISR/FSR) or split muon clusters**. Current studies are working to remove this component from the sample.

Two different MVA's were trained. The **base MVA** was trained on 7 features that contained the ECL cluster information:

- energy: clusterE
- distribution of energy: clusterE1E9 (ratio of energy in centre crystal to energy in surrounding 9 crystals) , clusterSecondMoment (distribution of energy in a plane perpendicular to the photon trajectory), clusterLAT (lateral energy distribution), clusterZernikeMVA (output of MVA that characterises cluster shapes)
- cluster position in the ECL: clusterPhi (azimuthal angle), clusterTheta (polar angle)

An **improved MVA** trained on the same 7 features in addition to:

- minC2TDist: distance between cluster and nearest track
- clusterTiming: time of ECL cluster in relation to the event time t_0

These additional features were expected to improve the discrimination power of the MVA when comparing their distributions. This is presented in Fig. 2.

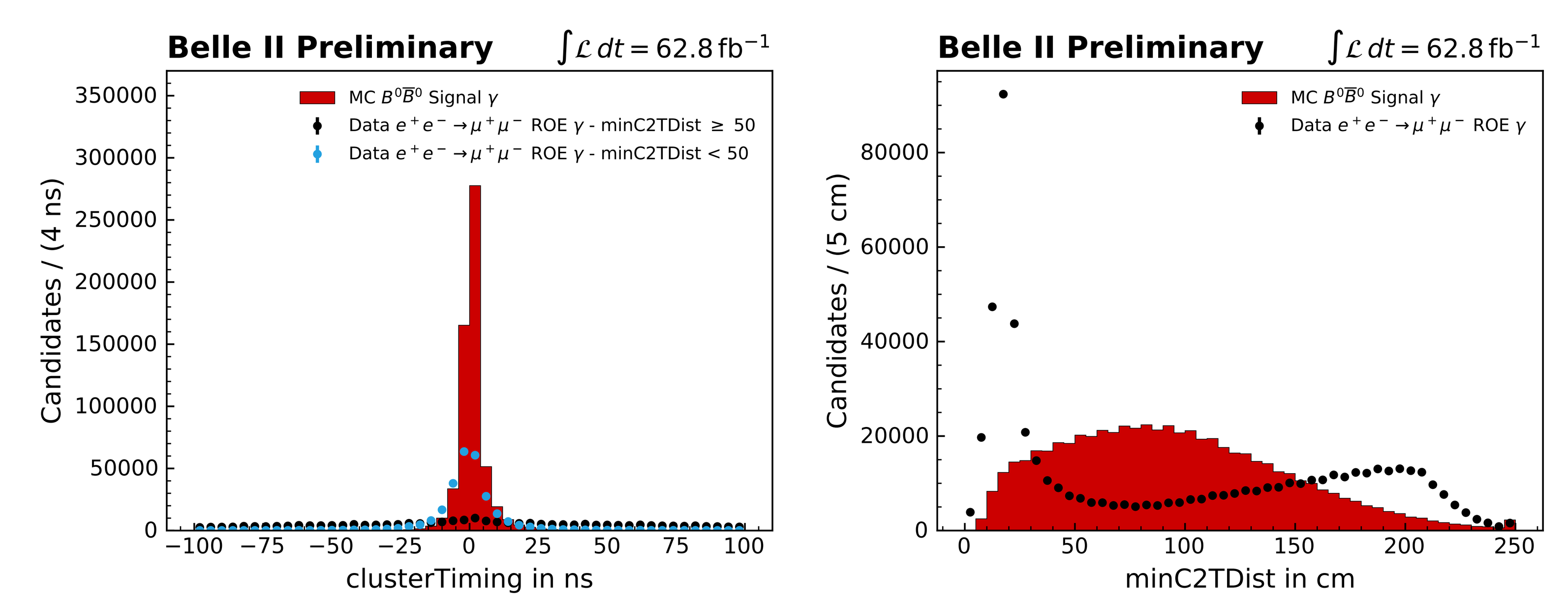

Figure 2: Signal and beam background photon distributions for clusterTiming and minC2TDist. The left distribution is split to show the beam background component (minC2TDist $\geq$ 50) and ISR/FSR component (minC2TDist < 50)

The `FastBDT` MVA outputs a variable called `BDTgamma` which gives the probability that a photon is signal-like. A cut on `BDTgamma` can be used as the threshold for photons to be identified as signal. This cut will need to be tuned depending on the B decay in question.

CLASSIFICATION PERFORMANCE

Receiver operating characteristic (ROC) curves are useful for comparing performance between classifiers. The ROC **area under curve (AUC)** is a metric that evaluates the classifier's separation power. A value of 1 for the AUC indicates perfect separation and 100% accuracy. The ROC and AUC for the base and improved MVA are given in Fig. 3.

- The base MVA has an **AUC of 0.87** $\to$ information on energy and energy distribution plus hit position of ECL clusters is able to provide significant separation power - makes sense:
  - Scatterings that create beam background typically produce low energy photons
  - Origin points are different for signal photons (the IP) and beam background photons (whole beam pipe)
- The improved MVA has an **AUC of 0.97** $\to$ near perfect separation

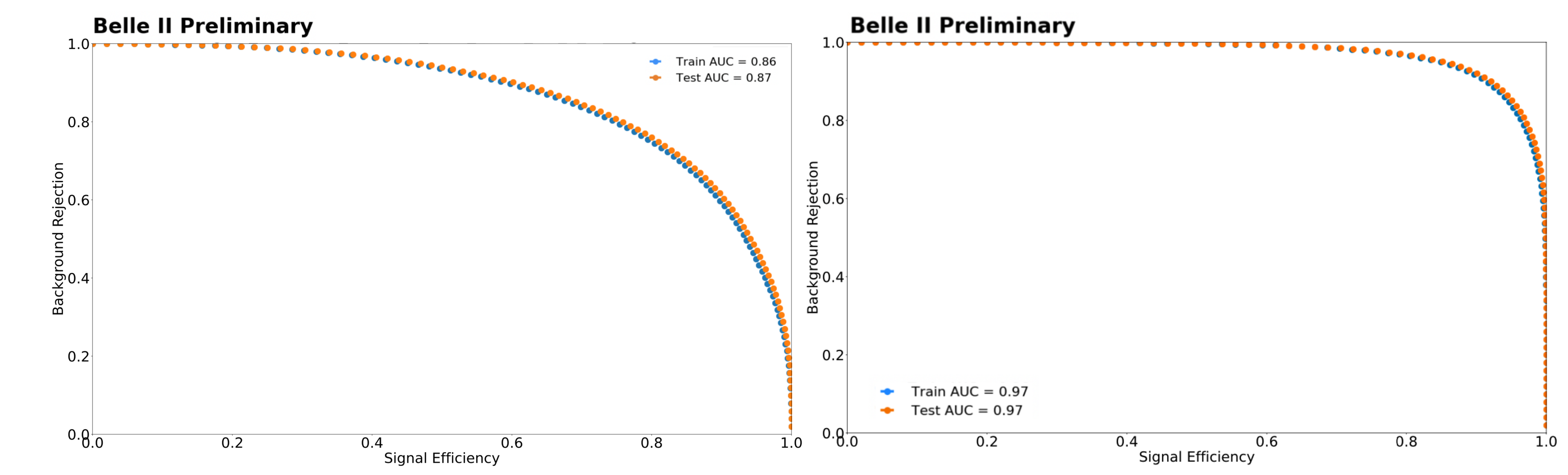

Figure 3: ROC curves and AUC values for the (left) base and (right) improved MVA. The test AUC is shown to demonstrate no overfitting.

TESTING THE MVA ON $\mathrm{B^0 \to D^*\ell\nu}$

To see if the separation power of the **improved MVA** is enough to suppress beam background in the signal E_{ECL} distribution, the improved MVA was tested on a sample of $B^0 \to D^*\ell\nu$ signal events. A cut of `BDTgamma` > 0.3 was chosen for this B decay. The E_{ECL} distribution corresponding to this cut is shown in Fig. 4. Note that the `BDTgamma` cut is quite loose, and further improvements on E_{ECL} can be made by tightening this cut.

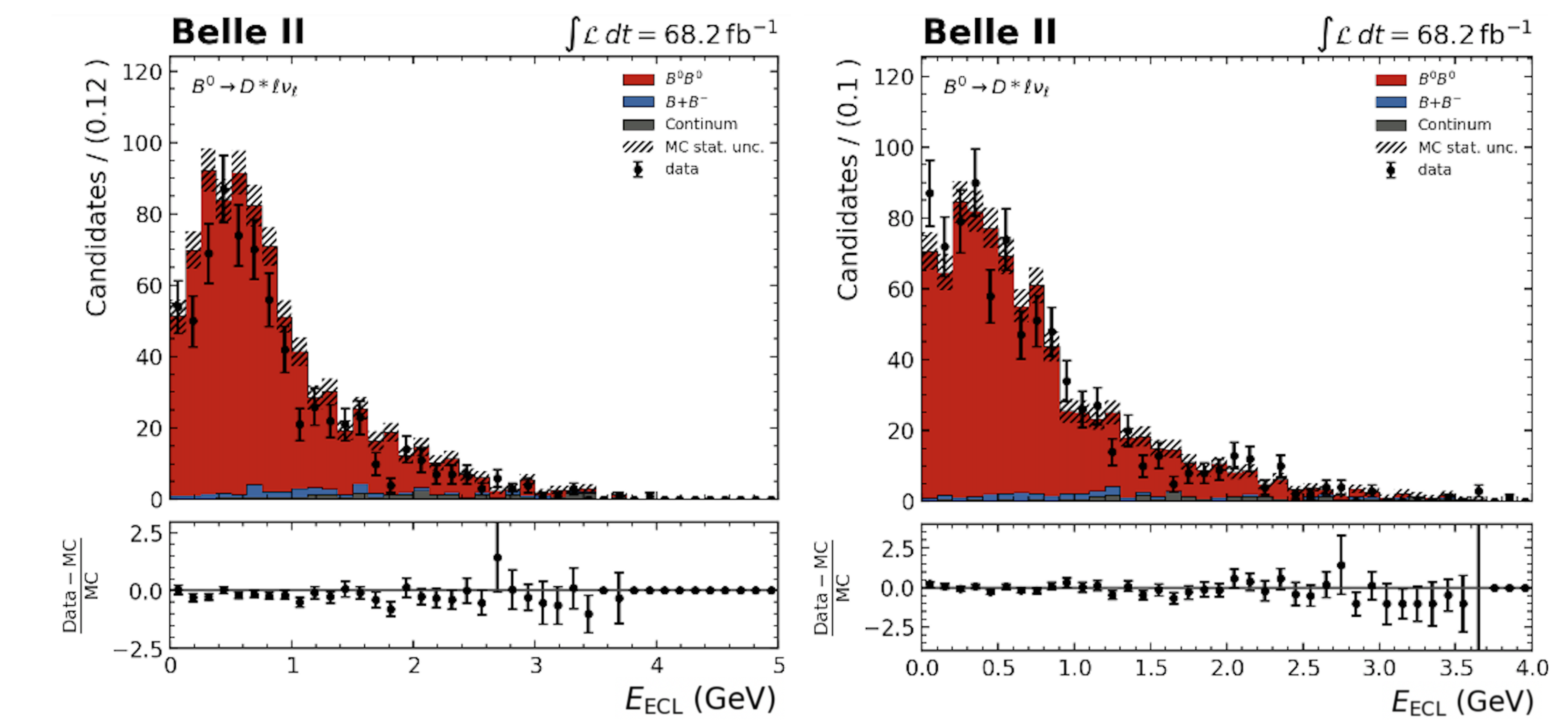

Figure 4: E_{ECL} for $B^0 \to D^*\ell\nu$ (labelled $B^0\overline{B^0}$) (left) before and (right) after the improved MVA `BDTgamma` > 0.3 cut

It is clear that the signal E_{ECL} distribution is **peaked closer to 0** after using the improved MVA, showing that beam background has been suppressed.

IS THIS THE WHOLE PICTURE?

Recent studies have shown that the **source distributions for beam background photons** need to be adjusted. This is for two reasons:

1. Current reconstruction method for beam background from $e^+e^- \to \mu^+\mu^-$ incorrectly labels **initial state/final state radiation** as beam background $\to$ distribution shapes will change
2. **minC2TDist** for beam background will look different in $B\bar{B}$ events than $e^+e^- \to \mu^+\mu^-$ as they have **more tracks** $\to$ $B\bar{B}$ events will not have minC2TDist peak at high values

We are currently working to correct the distributions and **further improve** the MVA training.

REFERENCES

[1] Thomas Keck. *FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification*. 2016. arXiv: `1609.06119 [cs.LG]`.

[2] Bianca Scavino. ``Exotic and conventional quarkonium physics prospects at Belle II". In: *Journal of Physics: Conference Series* 1667 (Oct. 2020), p. 012037.