Studies of R(D) & R(D*) using an inclusive tagging method at Belle II

Boyang Zhang

Department of Physics, University of Hawaii at Manoa

Introduction

B meson semileptonic decays involving b->c transitions provide rich information on the Standard Model (SM) and physics beyond it. The ratios of branching fractions R(D) and R(D*) are sensitive probes for testing lepton flavor universality in the SM, and their q^2 distribution has great potential to reveal new physics.

$$R(D) = \frac{BR(B \to D\tau\nu)}{BR(B \to D\ell\nu)}, \qquad R(D^*) = \frac{BR(B \to D^*\tau\nu)}{BR(B \to D^*\ell\nu)}$$

Although limited by statistical uncertainty, the world average of R(D) and R(D*) to date is 3.2σ from the SM. Better precision can be achieved with more data or a new analysis method that does not require high purity.

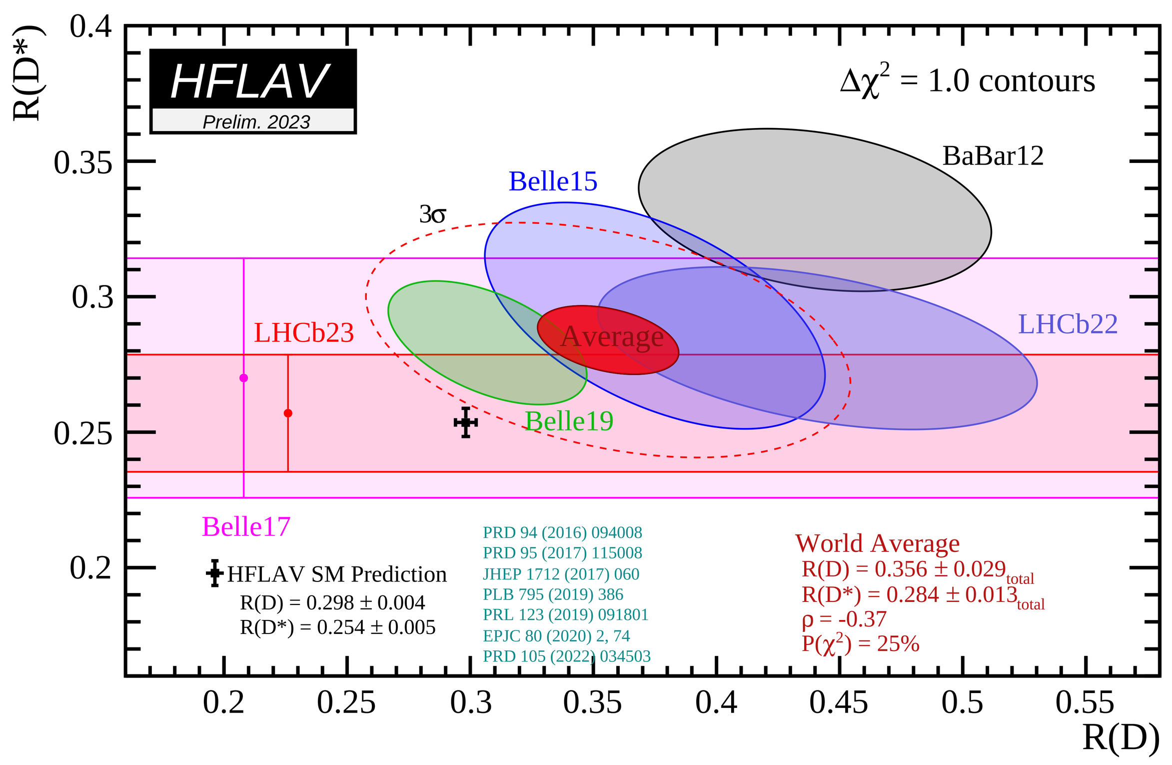


Figure 1. R(D) and R(D*) exceed the SM predictions by 1.98σ and 2.15σ respectively. Considering the R(D)-R(D*) correlation of -0.37, the resulting combined χ^2 is 13.0 for 2 degree of freedom, corresponding to a p-value of 1.47 x 10^{-3}. The difference from SM predictions corresponds to about 3.2σ. [1]

This poster presents an MC sensitivity study of a new method based on inclusive tagging, which has about 10 times higher reconstruction efficiency than in previous measurements. We plan to use the Belle II pre-LS1 (Long Shutdown 1) dataset, which corresponds to 364/fb.

Kinematics Tagging at Belle II

At the Belle II experiment, e^+e^- pairs are collided at the center of mass energy of the $\Upsilon(4S)$ resonance, producing pairs of B mesons. Since neutrinos are not directly detected at Belle II, it is not possible to fully reconstruct the kinematics of a B meson whose final state involves neutrinos. However, Belle II uses a technique called kinematic tagging to retrieve this missing information. The method relies on measuring the kinematics of the other B meson, called B_{tag}, and the beam energy precisely. By doing so, the kinematics of the B meson decaying to channels of interest, called B_{sig}, can be inferred and constrained.

- **Exclusive Tagging** is an algorithm [3] that automatically and exclusively reconstructs the B_{tag} in thousands of decay channels. Since both B mesons are exclusively reconstructed, the background can be significantly reduced by discarding events with unused tracks or clusters. However, although exclusive tagging provides high purity, the reconstruction efficiency is often low.
- **Inclusive Tagging** combines all particles that are not used in B_{sig} to be the B_{tag}, discarding events where the kinematics of B_{tag} is very different from that of a nominal B meson. This results in a reconstruction efficiency that is typically one order of magnitude higher than exclusive tagging. However, the tagged sample is less pure.

B_{sig} Reconstruction and Inclusive Tagging

- B_{sig}: Use well-identified $e^\pm$, $\mu^\pm$, $\pi^\pm$ and $K^\pm$ particles to reconstruct the signal decay channel $B^0 \to [D^- \to K^+\pi^-\pi^-][\tau^+ \to \ell^+\nu\bar{\nu}]\nu$ as well as its charge-conjugate channel.
- B_{tag}: Collect all particles not used in the reconstruction of B_{sig}, but exclude poorly measured charged particles, possible beam background, and fake photons. If V0 particles are detected, update the kinematics through vertex fitting.

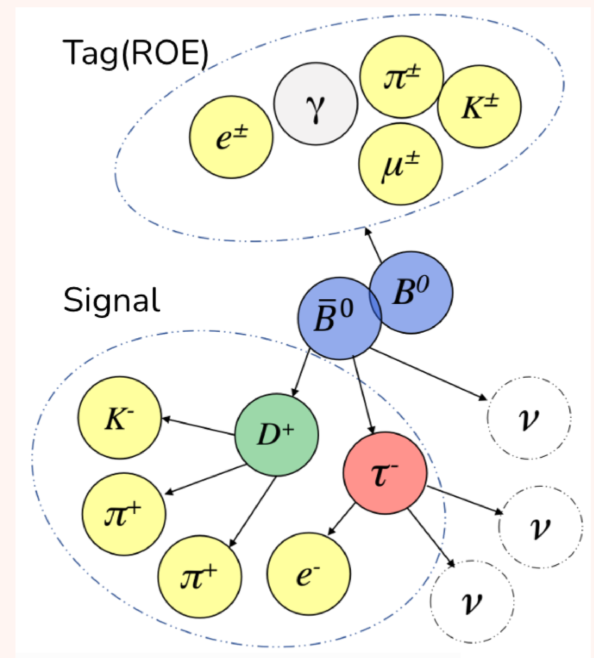


Figure 2. An example signal event with an inclusive tagging method

All events with final states involving $e^\pm$, $\mu^\pm$, $\pi^\pm$, and $K^\pm$ are reconstructed, and the major backgrounds are listed below:

- **Continuum**: A $e^+e^- \to q\bar{q}$ event is reconstructed, instead of a $e^+e^- \to \Upsilon(4S)$ event.
- **Fake D**: In a $e^+e^- \to \Upsilon(4S)$ event, $\pi^\pm$s and $K^\pm$s that do not all originate from a D meson are combined to form the D meson.
- **Fake B**: In a $e^+e^- \to \Upsilon(4S)$ event, $D^\pm$ and $\ell^\pm$ that do not originate from a B meson (or the same B meson) are combined to form the B meson.

Background Suppression

A multi-class classifier was trained using LightGBM [2], a Boosted Decision Tree (BDT) model, to suppress the backgrounds mentioned in the previous section. The input features are listed below:

- **Event Shape:** The topology of a $e^+e^- \to q\bar{q}$ event is more jet-like, whereas a $e^+e^- \to \Upsilon(4S)$ event is more spherical.
- **D Vertex and Daughters:** A correctly reconstructed D meson has a better vertex quality and its decay vertex is farther from the interaction point (IP). Information about daughters, such as their track quality, PID, and Dalitz variables, helps distinguish the Fake D background.
- B_{tag} **Kinematics and Multiplicity:** In a Fake B background event, the kinematics and multiplicity of B_{tag}, consisting of extra particles, differ from those of a correctly reconstructed B meson.

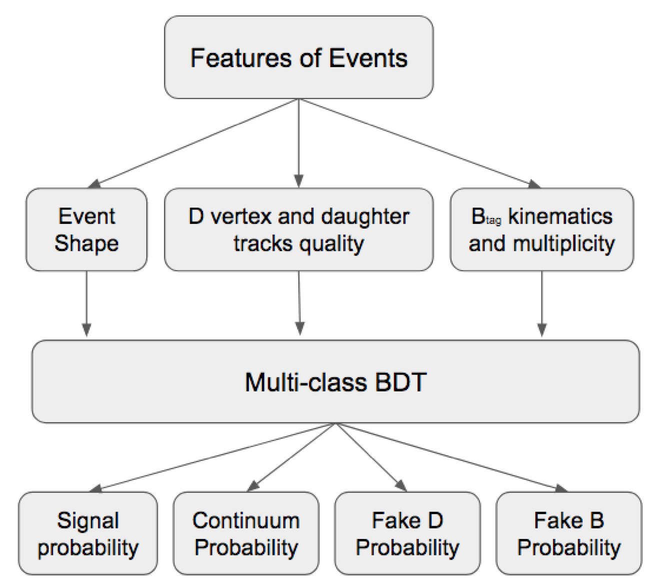
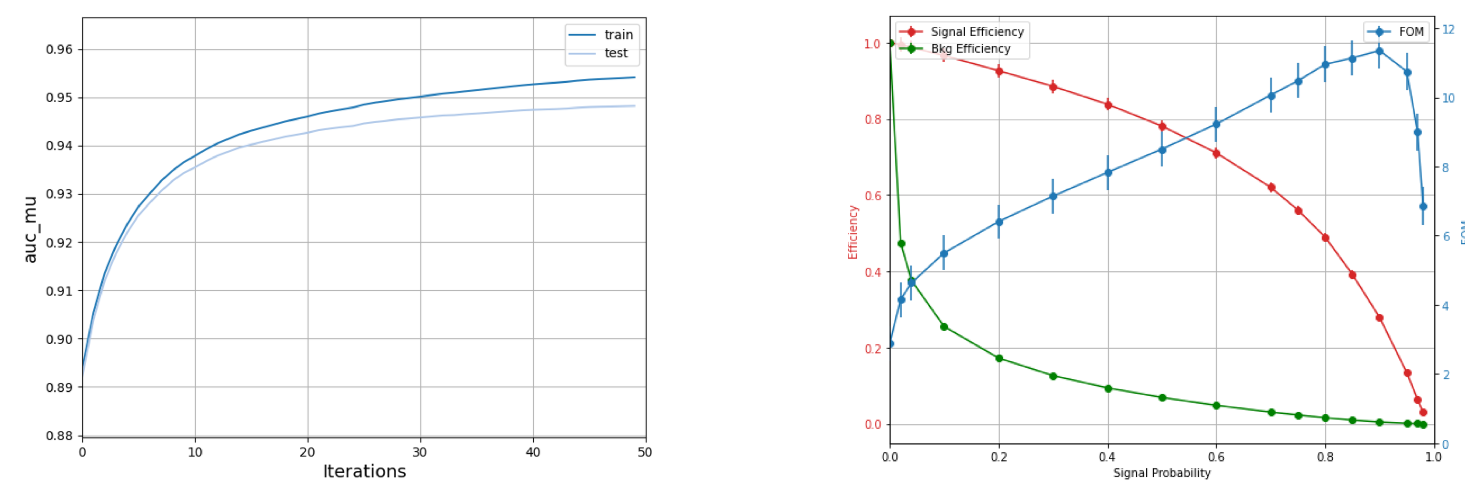


Figure 3. A block diagram demonstrating the Background Suppression BDT



(a) AUC ROC during training

(b) The figure of merit of the signal probability

Figure 4. BDT Performance

Signal Extraction

The amount of background becomes more manageable after implementing the Background Suppression BDT. A 2D template-fitting in $|\vec{p_D}| + |\vec{p_\ell}|$ $vs.$ M^2_{miss} is performed to extract the yields of signal and background components.

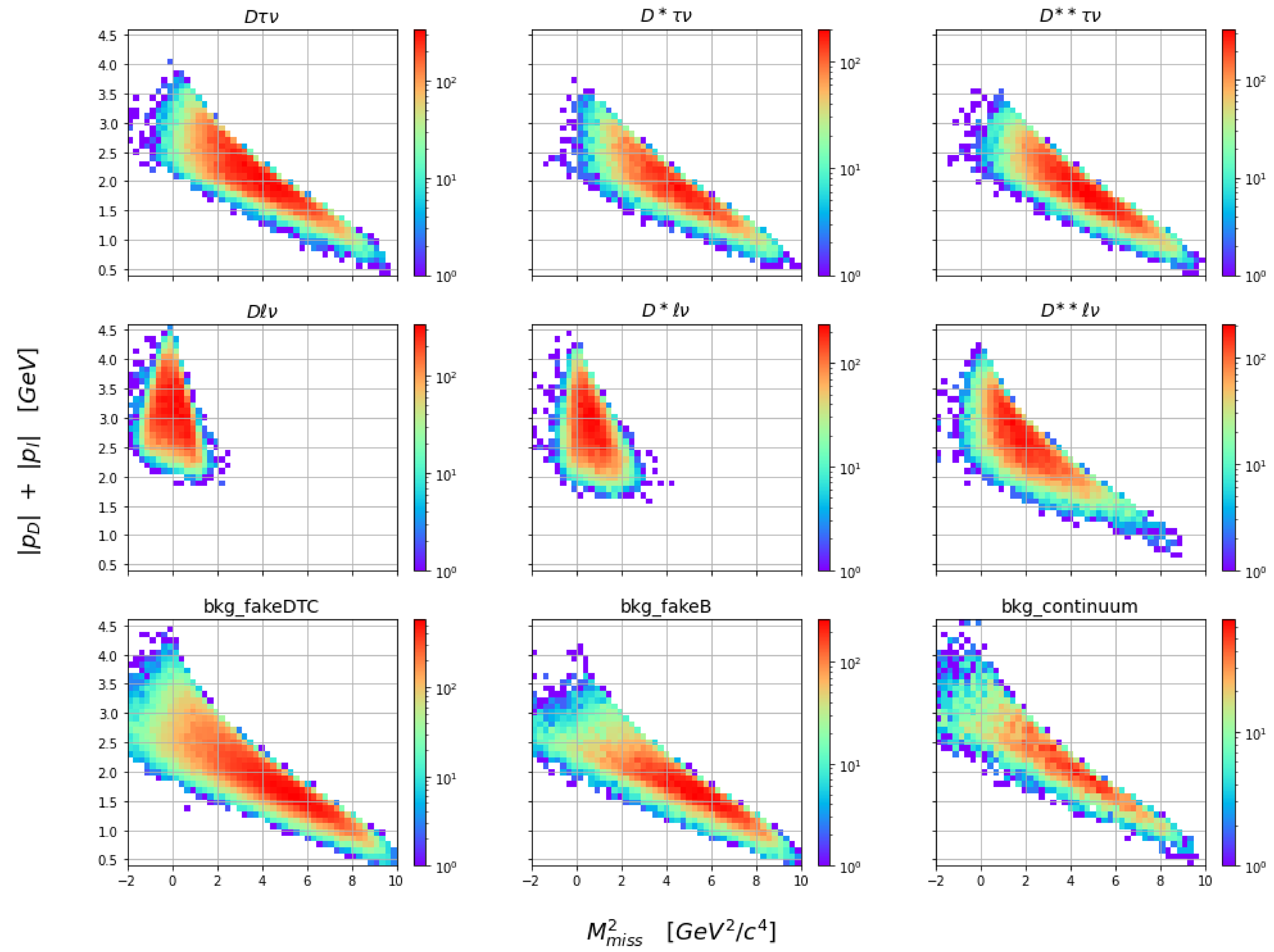


Figure 5. Templates of signals and backgrounds in $|\vec{p_D}|+|\vec{p_\ell}|$ vs M^2_{miss}

During a fitting practice where the background components are temporarily masked, an estimate of the statistical uncertainty of R(D) is approximately 20%. However, this fitting step is still in progress, and a recent test has shown some biases. Once these issues are addressed, the uncertainty is expected to be much smaller.

Summary and Conclusion

- **Inclusive tagging** provides one order of magnitude higher reconstruction efficiency than exclusive tagging methods, although it suffers from low purity.
- **Background Suppression BDT** significantly reduces the continuum, fake D, and fake B backgrounds.
- **Template fitting** is still in progress, and further investigation is needed to resolve biases observed during recent tests.

Acknowledgements

I would like to express my sincere gratitude to the U.S. Department of Energy Office of High Energy Physics for supporting my research through Award No. DE-SC0010504. I would also like to extend my appreciation to my advisor, Professor Tom Browder, for his invaluable guidance and mentorship. Thank you also to the Belle II collaboration and the University of Hawaii at Manoa for its support of this research endeavor.

References

[1] Heavy Flavor Averaging Group (HFLAV). R(D) and R(D*). https://hflav-eos.web.cern.ch/hflav-eos/semi/winter23_prel/html/RDsDsstar/RDRDs.html, 2023.

[2] Guolin Ke, Qiushi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, and Qing Ye. Lightgbm: A highly efficient gradient boosting decision tree, 2017. Software available from https://github.com/Microsoft/LightGBM.

[3] Keck T., Abudinén F., Bernlochner F.U., and et al. The full event interpretation. *Comput Softw Big Sci*, 3(6), 2019.