

SUPPLEMENTAL MATERIAL

π^0 BDT classifier

The BDT classifier used to identify B_{sig} candidates is trained using simulated samples consisting of approximately 180 000 true and misreconstructed π^0 candidates. These candidates are required to satisfy $|M_{\gamma\gamma} - 134 \text{ MeV}/c^2| < 27 \text{ MeV}/c^2$ and $p_{\gamma\gamma} > 420 \text{ MeV}/c$. The main background is fake π^0 candidates reconstructed from spurious photons. We use the following nine variables for training: the energies of both γ candidates; properties of the ECL cluster of the less energetic γ , namely, the azimuthal angle, number of hits, E_1/E_9 and E_9/E_{21} (the ratios of energy deposited in 1×1 , 3×3 , and $5 \times 5 - 4$ (at the corners) square bundle of crystals), and the MVA output for γ - K_L^0 separation using Zernike moments [1]; the helicity angle and the invariant mass of the π^0 . Figure 1 shows distributions of the BDT output for signal and background in an independent test sample.

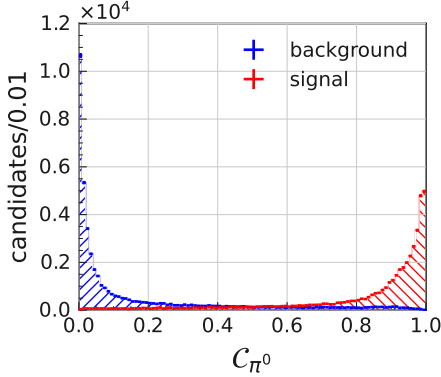


FIG. 1: The BDT output, C_{π^0} , distributions for signal and background.

$q\bar{q}$ suppression

To separate signal events from $q\bar{q}$ background events, we train a BDT classifier using approximately 770 000 B^0 candidates from signal and $q\bar{q}$ simulated samples. We apply the following selections on B_{tag} daughters: at least 1 CDC hit and $p^* < 3.2 \text{ GeV}/c$ for tracks, and $p > 50 \text{ MeV}/c$ and $p^* < 3.2 \text{ GeV}/c$ for ECL clusters. The following 29 variables are used as input [1]: the ratio between the second and zeroth Fox–Wolfram moments; the cosine of the angle between the thrust axis of B_{sig} and B_{tag} ; the thrust scalars of B_{sig} and B_{tag} ; the sum of the magnitudes of the transverse momenta of all the particles in the event; the missing mass squared of the event, defined as $M_{\text{miss}}^2 = (E_{e^+e^-} - \sum_i E_i)^2 - \sum_i |\mathbf{p}_i|^2$, where $E_{e^+e^-}$ is the sum of the e^+e^- beam energies, E_i and $\mathbf{p}_i$

are the energy and momentum of particle i , and the summation of i runs over all particles; 14 Super Fox–Wolfram moments; and 9 CLEO cones. As the flavor tagging is optimized for $B\bar{B}$ events, a high value of r indicates lower contamination from $q\bar{q}$ background. To maximize the selection performance, we set different thresholds on the BDT output for each r -bin. Figure 2 shows distributions of classifiers for MR1 and MR2 for signal and background for independent test samples.

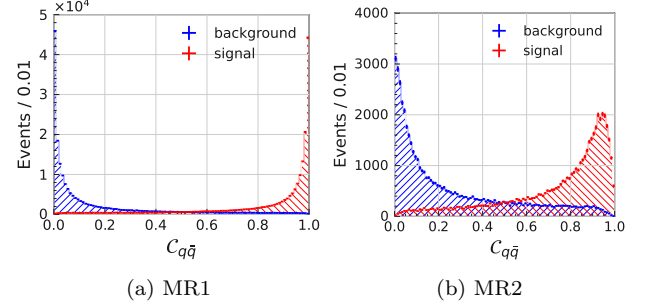


FIG. 2: Distributions of $q\bar{q}$ suppression BDT outputs for (a) MR1 and (b) MR2 regions, for signal and background.

Δt resolution functions

The resolution functions R_{sig} and $R_{B\bar{B}}$ are the convolutions of four functions [1],

$$R_{\text{vtx sig}} \otimes R_{\text{vtx tag}} \otimes R_{D \text{ tag}} \otimes R_{\text{boost}}, \quad (1)$$

where $R_{\text{vtx sig}}$ and $R_{\text{vtx tag}}$ are detector resolutions for B_{sig} and B_{tag} decay vertices, $R_{D \text{ tag}}$ represents the bias due to the secondary decay vertices of intermediate charm states in B_{tag} decay, and R_{boost} is the correction to the boost factors due to the small but nonzero momenta of the B mesons in the c.m. frame.

The detector resolutions $R_{\text{vtx sig}}$ and $R_{\text{vtx tag}}$ are modeled with the sum of two Gaussian functions:

$$R_{\text{vtx sig (tag)}}(\delta\ell) = (1 - f_{\text{tail}})G(\delta\ell; 0, s_{\text{main}}\sigma_\ell) + f_{\text{tail}}G(\delta\ell; 0, s_{\text{tail}}\sigma_\ell), \quad (2)$$

where $\delta\ell$ and σ_ℓ in $R_{\text{vtx sig}}$ ($R_{\text{vtx tag}}$) are the residual and the uncertainty of ℓ_{sig} (ℓ_{tag}), $G(x; \mu, \sigma)$ is a Gaussian function with mean μ and standard deviation σ , f_{tail} is the fraction of the tail component, and s_{main} and s_{tail} are scaling factors for the uncertainty. The parameters f_{tail} and s_{main} in $R_{\text{vtx sig}}$ (s_{main} and s_{tail} in $R_{\text{vtx tag}}$) are modeled with a linear dependence on $(\chi_{\text{tracks}}^2)^{\text{sig}}$ or $(\chi^2/\nu_{\text{d.f.}})^{\text{tag}}$, which are the B_{sig} vertex fit χ^2 calculated for the $K_S^0 \rightarrow \pi^+\pi^-$ tracks and the reduced χ^2 for B_{tag} vertex fit. These dependencies correct for underestimated vertex position uncertainty in case a track is reconstructed using incorrect hits.

The bias due to the secondary decay vertex $R_{D\text{tag}}$ is modeled by the sum of two exponential functions and a delta function:

$$R_{D\text{tag}}(\delta\ell_{\text{tag}}) = f_\delta\delta(\delta\ell_{\text{tag}}) + (1-f_\delta)f_p E_p(\delta\ell_{\text{tag}}; \tau_{\text{np}}\sigma_{\ell_{\text{tag}}}) + (1-f_\delta)(1-f_p)E_n(\delta\ell_{\text{tag}}; \tau_{\text{np}}\sigma_{\ell_{\text{tag}}}), \quad (3)$$

where E_p and E_n are one-sided exponential functions,

$$E_p(x; \tau) = \frac{1}{\tau} \exp\left(-\frac{x}{\tau}\right) \text{ for } x > 0, \text{ otherwise } 0, \quad (4)$$

$$E_n(x; \tau) = \frac{1}{\tau} \exp\left(+\frac{x}{\tau}\right) \text{ for } x \leq 0, \text{ otherwise } 0, \quad (5)$$

and f_δ , f_p , and τ_{np} are functions of $(\chi^2/\nu_{\text{d.f.}})^{\text{tag}}$ or $\sigma_{\ell_{\text{tag}}}$.

The boost factor correction R_{boost} is [1]:

$$R_{\text{boost}}(\Delta t - \Delta t_{\text{true}}) = \begin{cases} E_p(\Delta t - (a\Delta t_{\text{true}} + c_k|\Delta t_{\text{true}}|); |c_k|\tau_{B^0}) & (\cos\theta_B^* > 0) \\ \delta(\Delta t - a\Delta t_{\text{true}}) & (\cos\theta_B^* = 0) \\ E_n(\Delta t - (a\Delta t_{\text{true}} + c_k|\Delta t_{\text{true}}|); |c_k|\tau_{B^0}) & (\cos\theta_B^* < 0) \end{cases}, \quad (6)$$

where Δt_{true} is the true difference between the proper decay times of B_{sig} and B_{tag} , Δt is calculated with true ℓ_{sig} and ℓ_{tag} , and

$$a = \frac{E_B^*}{m_{B^0}}, \quad c_k = \frac{|\mathbf{p}_B^*| \cos\theta_B^*}{\beta m_{B^0}}, \quad (7)$$

with m_{B^0} being the nominal B^0 mass [1].

In total, we have 17 parameters: two in f_{tail} , two in s_{main} , and s_{tail} for $R_{\text{vtx sig}}$; f_{tail} , two in s_{main} , and two in s_{tail} for $R_{\text{vtx tag}}$; and two in f_δ , two in f_p , and three in τ_{np} for $R_{D\text{tag}}$. These are all fixed to values obtained from fitting MC samples.

We model the Δt PDF for $q\bar{q}$ background with the sum

of three Gaussian functions:

$$P_{q\bar{q}}(\Delta t; X^2, \sigma_{\Delta t}) = \left(\frac{1}{2}\right) \left[(1-f_2)(1-f_3)G(\Delta t; 0, (s_1^0 + s_1^1 X^2)\sigma_{\Delta t}) + f_2(1-f_3)G(\Delta t; 0, (s_2^0 + s_2^1 X^2)\sigma_{\Delta t}) + f_3G(\Delta t; 0, s_3\sigma_{\Delta t}), \right] \quad (8)$$

where X^2 is the average reduced χ^2 and $\sigma_{\Delta t}$ is the Δt uncertainty:

$$X^2 = \frac{(\chi_{\text{tracks}}^2)^{\text{sig}}/1.6 + (\chi^2/\nu_{\text{d.f.}})^{\text{tag}}}{2}, \quad (9)$$

$$\sigma_{\Delta t} = \frac{\sqrt{(\sigma_{\ell_{\text{sig}}})^2 + (\sigma_{\ell_{\text{tag}}})^2}}{\beta\gamma}. \quad (10)$$

The factor of 1.6 in Eq. (9) accounts for the effective number of degrees of freedom for the K_s^0 vertex in the B^0 vertex fit. The seven free parameters, namely f_2 , f_3 , s_1^0 , s_1^1 , s_2^0 , s_2^1 , and s_3 , are determined from a fit to events in the $M_{\text{bc}}-\Delta E$ sideband.

-
- [1] See A. Khotanzad and Y. Hong, *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 489 (1990) for information about the Zernike moments; G. C. Fox and S. Wolfram, *Phys. Rev. Lett.* **41**, 1581 (1978), Ed. A. J. Bevan, B. Golob, Th. Mannel, S. Prell, and B. D. Yabsley, *Eur. Phys. J. C* **74**, 3026 (2014), S. Brandt et al., *Phys. Lett.* **12**, 57 (1964), E. Farhi, *Phys. Rev. Lett.* **39**, 1587 (1977), S. H. Lee et al. (Belle Collaboration), *Phys. Rev. Lett.* **91**, 261801 (2003), and D. M. Asner et al. (CLEO Collaboration), *Phys. Rev. D* **53**, 1039 (1996) for additional information about the input variables for the $q\bar{q}$ suppression; H. Tajima et al., *Nucl. Instrum. Methods Phys. Res., Sect. A* **533**, 370 (2004) for additional information about the resolution functions; R. L. Workman et al. (Particle Data Group), *Prog. Theor. Exp. Phys.* **2022**, 083C01 (2022) for the nominal B^0 mass.