Contents

6 B-meson reconstruction

Editors:
Paul Jackson (BABAR)
Anze Zupanc (Belle)
– (theory)

Introduction

The recoil analyses exploit large samples of $\Upsilon(4S)$ decays into $B\overline{B}$ pairs collected with the detectors of the asymmetric energy B–factories *BABAR* and Belle. Both detectors are optimized to exploit the $\Upsilon(4S)$ production in an asymmetric environment. The *BABAR* detector **?** is situated at the PEP-II facility consisting of asymmetric storage rings containing 3.1 GeV positrons and 9 GeV electrons. Moving radial outwards from the interaction region *BABAR* is composed of a tracking system, a five-layer silicon vertex tracker (SVT) and a 40-layer drift chamber (DCH), a Cherenkov detector for charged π–K discrimination, a CsI calorimeter (EMC) for photon and electron identification, and an 18–layer flux return (IFR) located outside of the 1.5 T solenoidal coil and instrumented with resistive plate chambers for muon and neutral hadron identification. For the most recent data taking, a portion of the IFR has been replaced with limited streamer tubes.

The Belle detector **?** operates at KEKB, colliding 8 GeV electrons onto 3.5 GeV positrons. Like *BABAR*, Belle is also a magnetic spectrometer consisting of a 4-layer silicon vertex detector and 50-layer central drift chamber. An aerogel threshold Cherenkov counter and time-of-flight scintillation counters provide with a measurement of dE/dx ionization energy loss in the drift chamber the particle identification, and an electromagnetic calorimeter (of CsI) inside a 1.5 T magnetic field provides electron and photon identification. An iron flux return is instrumented outside the coil to detect K_L^0 mesons and identify the muons. A minor part of data was recored using a 3-layer silicon detector.

Methodology and Motivation

Due to the presence of multiple neutrinos in the final state certain decay modes lack the kinematic constraints which are usually exploited in B decay searches in order to reject both continuum and $B\overline{B}$ backgrounds. The strategy adopted in these analyses is to reconstruct exclusively the decay of one of the B mesons in the event, referred to as the "tag"-B (B_{tag}). The remaining particle(s) in the event, referred to as the "recoil"-B ($B_{\rm recoil}$), are then compared with the signature expected for the target signal mode. This can be done in two ways depending on whether the $B_{\rm tag}$ is reconstructed semileptonically or hadronically.

We will describe the two techniques in general terms and then make specific reference to choices made in a variety of searches at the B-factories. It is important to realize that the two analysis techniques are complimentary and non-overlapping and, as such, can be readily combined to improve the sensitivity of any analysis. This has the effect of essentially doubling the available dataset which for a statistically limited search, which is of vital importance. At a proposed future Super Flavour Factory (SFF) **?** sufficient statistical precision on this channel may cause the systematic uncertainties to play a more prominent role in the overall sensitivity. We will discuss some of the challenges and methods for calculating systematic uncertainties.

Herein, specific reference will be made to the searches for $B^+ \to \tau^+\nu_\tau$, to elucidate the necessity of the recoil method, although the techniques of studying the system recoiling against a reconstructed tag-B meson can be applied to any analysis [citations]. In particular, recoil methods are crucial for studying those channels where the decay kinematics cannot otherwise be fully constrained.

Many decay modes where the B meson cannot be exclusively reconstructed rely on these methods to make measurements feasible. For the proposed high luminosity asymmetric e^+e^- Super Flavour Factory **?** almost all B decay measurements, not related to CP violation or the CKM picture of the Standard Model, will benefit from recoil methods. This corresponds to a large program of purely leptonic, semileptonic and radiative penguin B decays. Furthermore, with a huge dataset the recoil methods will provide a clean "single B beam" which will permit the extraction of hadronic B decay branching fractions using a missing mass technique **?**.

Techniques

0.0.1 Hadronic tag B reconstruction

The full reconstruction of one B meson, decaying purely hadronically, has been utilized in a multitude of analyses by the B factories. The approaches of BaBar and Belle (and CLEO before them **?**) differ somewhat, and their treatment provides samples which vary in efficiency and purity. The optimization of these choices depends primarily on the target signal in the recoil system and the available kinematic constraints which can be imposed.

BABAR opts for a semi-exclusive approach to B reconstruction. The starting point is the selection of a sample of events in which the hadronic decay of one of the two B mesons ($B_{\rm reco}$) is fully reconstructed. About 1000 different $B \to DY$ decay chains are selected, where D refers to a charm meson and Y represents a collection of hadrons composed of $n_1\pi^\pm + n_2K^\pm + n_3\pi^0 + n_4K^0_S$ ($n_1 = 1,...5$, $n_2 = 0,...2$, $n_3 = 0,...2$ and $n_4 = 0,1$) and having total charge equal to ±1. The following modes are reconstructed: $D^- \to K^+\pi^-\pi^-$, $K^+\pi^-\pi^-\pi^0$, $K^0_S\pi^-$, $K^0_S\pi^-\pi^0$, $K^0_S\pi^-\pi^-\pi^+$; $D^{*-} \to \overline{D}^0\pi^-$; $\overline{D}^0 \to K^+\pi^-$, $K^+\pi^-\pi^0$, $K^+\pi^-\pi^-\pi^+$, $K^0_S\pi^+\pi^-$; and $\overline{D}^{*0} \to \overline{D}^0\pi^0, \overline{D}^0\gamma$. The D^- and D^{*-} ($\overline{D}^0$ and $\overline{D}^{*0}$) decays are used as a "seed" to reconstruct B^0 (B^+) decays. Overall, they correctly reconstruct one B candidate in 0.3% (0.5%) of the $B^0\overline{B}^0$ (B^+B^-) events. The kinematic consistency of a $B_{\rm reco}$ candidate with a B meson decay is checked using two variables: the energy difference, ΔE

$$\Delta E = E^*_B - E_{\rm beam}\ , \qquad (1)$$

where E^*_B is the energy of the B meson and $E_{\rm beam}$ is the beam energy, both in the $\Upsilon(4S)$ frame; $m_{\rm ES}$ ($M_{\rm bc}$), the energy substituted mass (beam constrained mass), as defined by *BABAR* (Belle), is given by:

$$m_{\rm ES} = M_{\rm bc} = \sqrt{[(s/2 + \mathbf{p_i}\cdot\mathbf{p_B})^2/E_i^2] - |\mathbf{p_B}|^2}\ , \qquad (2)$$

where $\sqrt{s}$ is the total energy of the e^+e^- system in the $\Upsilon(4S)$ rest frame, and $(E_i, \mathbf{p_i})$ and $(E_B, \mathbf{p_B})$ are the four-momenta of the e^+e^- system and the reconstructed B candidate respectively, both in the laboratory frame. For ease of description this will be referred to as $m_{\rm ES}$ in what follows. For suitable $B_{\rm reco}$ candidates the requirements are $-0.1 < \Delta E < 0.08\,{\rm GeV}$ and $m_{\rm ES} > 5.21\,{\rm GeV}/c^2$, correctly reconstructed events should have the $m_{\rm ES}$ and ΔE distributions peak at the B meson mass and at zero, respectively.

The combinatorial background from $B\overline{B}$ events and $e^+e^- \to q\bar{q}$ ($q = u, d, s, c$) production, in the $B_{\rm reco}$ sample, can be subtracted by performing an unbinned likelihood fit to the $m_{\rm ES}$ distribution, using the following threshold function (Argus) **?**

$$\frac{dN}{dm_{\rm ES}} = N\cdot m_{\rm ES}\cdot\sqrt{1-x^2}\cdot\exp\left(-\xi\cdot(1-x^2)\right) \qquad (3)$$

for the background (where $x = m_{\rm ES}/m_{\rm max}$ and $m_{\rm max}$ is the endpoint of the curve) and a Gaussian function corrected for radiation losses (Crystal Ball) **?** peaked at the B meson mass for the signal.

For each reconstructed tag B mode i, the $m_{\rm ES}$ distribution of the reconstructed B candidates in the data sample is fit with the sum of a Crystal Ball or Gaussian (depending on whether or not neutrals are present in the tag B decay mode) and an Argus function; the purity of the mode is determined as $S_i/(S_i + B_i)$ where S_i is the area of the Crystal Ball or Gaussian and B_i is the area of the Argus function for some region in $m_{\rm ES}$, typically taken to be $m_{\rm ES} > 5.27\,{\rm GeV}/c^2$. In events with more than one reconstructed $B_{\rm tag}$ candidate, the selected candidate is commonly taken as the one with the highest purity or lowest $|\Delta E|$. Only modes with an integrated purity greater than some selection criteria (typically between 30% and 55%) are included. The signal region of the $B_{\rm tag}$ is then defined to be in a region around $-90 < \Delta E < 60$ MeV and $m_{\rm ES} > 5.27\,{\rm GeV}/c^2$ and the events contained in the sideband $5.21 < m_{\mathrm{ES}} < 5.26\,\mathrm{GeV}/c^2$ are used as a control sample for continuum and combinatorial backgrounds.

A somewhat different strategies for hadronic B reconstruction are used by the Belle. In the first approach Belle reconstructs a set of exclusive final states with a high purity, in order to extract a clean subsample of B decays from the data sample.The B_{tag} candidates are reconstructed in the following decay modes: $B^- \to D^{(*)0}(\pi,\ \rho,\ a_1,\ D_s^{(*)})^-$ and $B^0 \to D^{(*)+}(\pi,\ \rho,\ a_1,\ D_s^{(*)})^-$. Within the above decay modes the D mesons used in the reconstruction of the B_{tag} are $D^0 \to K^-\pi^+$, $K^-\pi^+\pi^0$, $K^-\pi^+\pi^+\pi^-$, $K_S^0\pi^0$, $K_S^0\pi^-\pi^+$, $K_S^0\pi^-\pi^+\pi^0$ and K^-K^+, $D^+ \to K^-\pi^+\pi^+$, $K^-\pi^+\pi^+\pi^0$, $K_S^0\pi^+$, $K_S^0\pi^+\pi^0$, $K_S^0\pi^+\pi^+\pi^-$ and $K^+K^-\pi^+$, and the D_s^+ mesons are reconstructed as $D_s^+ \to K_S^0K^+$ and $K^+K^-\pi^+$. The $\bar{D}^{*0}$ and D_s^{*+} mesons are reconstructed in $\bar{D}^{*0} \to \overline{D}^0\pi^0$, $\overline{D}^0\gamma$, and $D_s^{*+} \to D_s^+\gamma$ modes respectively. The selection of B_{tag} candidates is based on the beam-constrained mass, M_{bc}, as defined in equation 2, and ΔE as given in equation 1. The selection criteria for B_{tag} depend on the studied signal mode and are typically defined as $-0.08 < \Delta E < 0.06\,\mathrm{GeV}$ and $M_{\mathrm{bc}} > 5.27\,\mathrm{GeV}/c^2$. If an event has multiple B_{tag} candidates the one with the smallest χ^2 is selected based on deviations from the nominal values of ΔE, the D candidate mass and the $D^* - D$ mass difference, if applicable. Overall, they correctly reconstruct one B candidate in 0.1% (0.2%) of the $B^0\overline{B}^0$ (B^+B^-) events with purity around 60%.

In a second approach Belle increased the number of reconstructed exclusive B decay modes and used the neural network in their selection in order to increase the hadronic tag B reconstruction efficiency. The B_{tag} candidates are reconstructed in the following decay modes: $B^- \to D^{(*)0}(\pi^-, \pi^-\pi^0, \pi^-\pi^-\pi^+)$, $D^{*0}\pi^-\pi^-\pi^+\pi^0$, $D^+\pi^-\pi^-$, D^0K^-, $D^{(*)0}D_s^{(*)-}$ and $J/\psi(K^-,\ K^-\pi^0,\ K_S^0\pi^-,\ K^-\pi^-\pi^+)$, and $B^0 \to D^{(*)-}(\pi^+,\ \pi^+\pi^0,\ \pi^+\pi^+\pi^0)$, $D^{*-}\pi^+\pi^+\pi^-\pi^0$, $\overline{D}^0\pi^0$, $D^{(*)-}D_s^{(*)+}$ and $J/\psi(K_S^0,\ K^+\pi^-,\ K_S^0\pi^+\pi^-)$. The D mesons used in the reconstruction of the B_{tag} are $D^0 \to \pi^-\pi^+$, $K_S^0K^-K^+$, $D^+ \to K^+K^-\pi^+\pi^0$ and $D_s^+ \to K^+\pi^+\pi^-$, $K^+K^-\pi^+\pi^0$, $K_S^0K^+\pi^+\pi^-$, $K_S^0K^-\pi^+\pi^+$, $K^+K^-\pi^+\pi^+\pi^-$ and $\pi^+\pi^+\pi^-$ in addition to the modes used by Belle in the first approach, given above. The reconstruction proceeds in four stages: in the first stage $\pi^\pm$, $K^\pm$, K_S^0, γ and π^0 candidate lists are created, used in second and later stages for building up the D^0, $D_{(s)}^\pm$ and J/ψ candidates, and D^{*0} and $D_{(s)}^{*\pm}$ in the third stage, and finally in the last, fourth, stage the $B^\pm$ and B^0 candidates are created. At each stage all available information on given candidate is used to calculate a single scalar variable with NeuroBayes, which can be by construction interpreted as a probability that a given candidate is correctly reconstructed. Finally, the selection of B_{tag} candidates is based on the beam-constrained mass, M_{bc}, as defined in equation 2, and ΔE as given in equation 1. In case of multiple B_{tag} candidates in a single event the one with the highest probability is accepted. Overall, Belle correctly reconstructs using this approach one B candidate in 0.4% (0.6%) of the $B^0\overline{B}^0$ (B^+B^-) events with purity around 20% in the $|\Delta E| < 0.05\,\mathrm{GeV}$ and $M_{bc} > 5.27\,\mathrm{GeV}$ region.

The treatment and definition of signal regions commonly depends on the target signal in the recoiling system. The remaining charged tracks and neutral reconstructed objects in the event, after reconstructing the B_{tag} are associated to the recoiling B meson and are studied to search for the target signal. The main advantage of this recoil technique is to provide a clean environment of $B\bar{B}$ events with a strong suppression of the combinatorial and continuum backgrounds arising from wrong assignment of charged and neutral objects to the parent B. Due to the complete reconstruction of one of the B mesons, without the missing neutrino as in the semileptonic reconstruction, the signal B rest frame can be accurately estimated. This provides a additional constraint on the signal B kinematics. For example, the reconstruction of $B^+ \to \mu^+\nu$ is a two-body decay where only one visible signal particle is produced. In the signal B rest frame, the experimental signature is a mono-energetic muon which is smeared only by the detector resolution Knowledge of the signal B rest frame is a useful constraint for all analyses exploiting this technique

0.0.2 Semileptonic tag B reconstruction

This method of semi–exclusive B reconstruction involves the selection of a D meson and suitable lepton candidate, ℓ, which are then combined into a $D\ell$ candidate.

The B_{tag} is reconstructed in the set of semileptonic B decay modes $B^- \to D^0\ell^-\overline{\nu}_\ell X$, where ℓ denotes an e or μ, and X can be either nothing or a transition particle from a higher mass charm state decay, which one does not necessarily need to reconstruct. This methodology naturally includes the $B^- \to D^0\ell^-\overline{\nu}_\ell$ and $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$ modes and also retains those modes with other charm states which decay, via the emission of soft transitions particles, to the D^0. The technique can be similarly applied to the tagging of neutral B mesons where one would reconstruct $\bar{B}^0 \to D^{(*)+}\ell^-\overline{\nu}_\ell$ for a combination of all possible $\bar{B}^0 \to D^+\ell^-\overline{\nu}_\ell$ and $\bar{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$ states reconstructed exclusively or resulting from a cascade decay from a higher mass charm state. The main loss in efficiency arises from the B and charm decay branching fractions while further selection criteria must be applied in order to suppress non-B decay backgrounds (continuum) and fakes from hadronic B decays.

The D^0 decay is reconstructed in the four cleanest hadronic decay modes: $K^-\pi^+$, $K^-\pi^+\pi^-\pi^+$, $K^-\pi^+\pi^0$, and $K_s^0\pi^+\pi^-$. The K_s^0 is reconstructed only in the mode $K_s^0 \to \pi^+\pi^-$. In principle other D^0 final states such as K^-K^+, $K^-\pi^+\pi^-\pi^+\pi^0$ and $K_s^0\pi^+\pi^-\pi^0$ will increase the efficiency but may suffer from a lower purity, although if the π^0 resolution is sufficient these modes could be useful to include. The added benefit of reconstructing the low momentum transition daughter of D^{*0} decays is to provide a more complete and exclusive tag B selection. Indeed if one neglects to reconstruct these π^0 or γ daughters (from $D^{*0} \to D^0\pi^0/\gamma$) then they will populate those particles considered for the signal B target mode. However, it is observed that the semi-exclusive reconstruction of $B \to D^0\ell\nu X$ provides a higher efficiency with some loss of purity.

For neutral B tags the selection becomes that of either $\bar{B}^0 \to D^+\ell^-\overline{\nu}_\ell$ or $\bar{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$. The D^+ decays are reconstructed in the decay modes $K^-\pi^+\pi^+$ or $K^0_S\pi^+$. Further modes of study may include $K^-\pi^+\pi^+\pi^0$, $K^-K^+\pi^+$, $K^0_S\pi^+\pi^0$, $K^0_S\pi^+\pi^-\pi^+$ and $K^0_S\pi^+\pi^-\pi^+\pi^0$. The D^{*+} decays can be reconstructed as both $D^0\pi^+$ and $D^+\pi^0$. The mass difference between D^* and D provides a powerful constraint as does the invariant mass of the D^0 or D^+ candidate.

As datasets grow the use of $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$ and $\bar{B}^0 \to D^{(*)+}\ell^-\overline{\nu}_\ell$ tags will likely be the most suitable way to approach this method.

The center of mass lepton momentum (p^*_ℓ) for both electrons and muons is selected to be greater than 800 MeV/c. This is the lowest end of muon identification for the current B factories and there is commonly non-B background below $p^*_\ell \sim$1 GeV. The reconstructed D^0 mesons are required to be within $\pm 3\sigma$ of the nominal value. The cosine of the hypothetical angle produced by the B meson and $D\ell$ candidate, $\cos\theta_{B,D\ell}$, is a powerful discriminant, the expression is provided in Equation 4. In the event that the $D\ell$ and neutrino are the only decay products of the B then the $\cos\theta_{B,D\ell}$ must lie in the physical region between ± 1. If additional decay products from the cascade of a higher mass charm state down to the D^0 go unreconstructed then this will force the value of $\cos\theta_{B,D\ell}$ to be negative. In order to retain such candidates events where $\cos\theta_{B,D\ell}$ falls between -2.5 and +1.1 are commonly retained. The positive limit is allowed to be slightly outside of the physical region to account for detector and reconstruction effects. Of course, for the reconstruction of exclusive channels ($B^- \to D^0\ell^-\overline{\nu}_\ell$, $B^- \to D^{*0}\ell^-\overline{\nu}_\ell$, $\bar{B}^0 \to D^+\ell^-\overline{\nu}_\ell$ and $\bar{B}^0 \to D^{*+}\ell^-\overline{\nu}_\ell$), the selection should be tightened to only consider the physical region.

$$\cos\theta_{B,D\ell} = \frac{(2\,E_B E_{D\ell} - m_B^2 - m_{D\ell}^2)}{2\,|\boldsymbol{p}_B||\boldsymbol{p}_{D\ell}|} \;. \qquad (4)$$

A typical $B^- \to D^0\ell^-\bar{\nu}_\ell X$ selection yields an efficiency of approximately 6×10^{-3} with a mode dependent purity which averages to $\sim$ 60%. For neutral B reconstruction the efficiency is typical half that of a similar charged B selection. The loss of one neutrino in the tagging mode limits the constraints that can be imposed when all of the B meson decay products are reconstructed. Hence, the signal B direction cannot be found as is possible for hadronic B reconstruction. However, this constraint is not of paramount importance in the analysis of $B^+ \to \tau^+\nu_\tau$ however, as there are multiple additional neutrinos present in the signal mode. Furthermore, as we shall see in the following section, the hadronic tagged technique suffers from a somewhat lower efficiency although this is counterbalanced by providing a more background free signal region.

0.0.3 Double Tagging

There are typically several assumptions made in the initial use of the recoil method. The first is that the B reconstruction efficiency is well modeled by the Monte Carlo simulations of generic B decays and continuum events. The second is that for analyses with few reconstructed particles from the signal B, the extra energy used to discriminate signal from background events is also well-modeled. These are assumptions which can be checked, however, by using control samples which test both the tag B reconstruction efficiency and the description of extra energy in a fully-reconstructed event. Both *BABAR* and Belle use double-tagged samples, in which both Bs are fully reconstructed either in semileptonic or hadronic final states, as such a control.

The double-tag approach was pioneered by the *BABAR* experiment **?**, using double-semileptonic B decays. For the semileptonic B_{reco} technique described in section 0.0.2 this means the reconstruction of two, distinct and non-overlapping $B^- \to D^0\ell^-\bar{\nu}_\ell X$ candidates and little other detector activity. Both *BABAR* and Belle have also used "hybrid double-tags", where one B is reconstructed in a purely hadronic final state while the second B is reconstructed in a purely semileptonic final state ($B^- \to D^0\ell^-\overline{\nu}$ or $B^- \to D^{*0}\ell^-\overline{\nu}$). These samples vary in size, depending on the final states used, but given a semileptonic tag reconstruction efficiency (quoted by *BABAR*) of $\sim$ 0.7% and hadronic tag efficiency of $\sim$ 0.2%, one expects to find approximately 50 semileptonic double-tagged events per fb^{-1}, 30 hybrid tags per fb^{-1}, and 4 hadronic double-tagged events per fb^{-1}. Given the large datasets of current B-factories, and the expected datasets at a future super flavor factory, these are significant samples which can be used as important cross-checks of the assumptions in the recoil method.

The double-tagged events have several important features. The first is that one expects the yield to naively proceed as the ε^2_{tag}, which is the basis of the cross-check of the tag efficiency. The second is that the complete reconstruction of both Bs creates an environment in which the extra energy in a given event should represent the effect of energy deposits unassociated with the B decays themselves. This latter feature is an important ingredient in the cross-check of the extra energy modeling in signal events, where it is also assumed that all detected particles associated with the B decays have been reconstructed.

The cross-check of the tag efficiency is currently only used in the semileptonic approach, and only by *BABAR*. The approach to this check has changed somewhat since the first papers were published using the semileptonic recoil approach. While the earlier publications themselves do not document the method, papers presented at conferences on the same work describe details of the method. The early approach to the double-tag sample (Ref. **?**) was to make the following assumption. Given an efficiency, ε_{tag}, for reconstructing one of the two B_s in an event in a semileptonic final state, the number of double tags (N_2)

is given simply by

$$N_2 = \varepsilon_{tag}^2 \times N_{B^+B^-} \quad (5)$$

where $N_{B^+B^-}$ is the number of charged B pairs originally produced by the B-factory or generated in Monte Carlo simulations. The tag efficiency cross-check was performed by taking the ratio of the above equation in data and in MC simulation and assuming that the double-tag sample is dominated by charged Bs so that $N_{B^+B^-}$ cancels, yielding the correction factor (c_{tag}) for the tagging efficiency in MC,

$$c_{tag} = \frac{\varepsilon_{tag}^{data}}{\varepsilon_{tag}^{MC}} = \sqrt{\frac{N_2^{data}}{N_2^{MC}}}. \quad (6)$$

This correction is only valid under the original assumption that the sample is dominated by charged B mesons in data, and that the efficiency of selecting two such Bs is given by ε_{tag}^2. While MC studies of the double-tags suggested that the contamination from neutral B decays, or other backgrounds, was very small (Fig. **??**), the second assumption - that the reconstruction of the first B does not bias the reconstruction of the second - is not addressed. The closeness of the correction to 1.0, as cited by *BABAR*, does suggest that the second assumption is essentially correct.

A second approach to the efficiency correction attempts to address some of the potential deficiencies of the first method outlined above. In the alternative approach **?**, the data/MC comparison is performed using the ratio of single-tagged to double-tagged events. If the efficiency of reconstructing the first tag is $\varepsilon_{tag,1}$ and the efficiency of reconstructing the second tag is $\varepsilon_{tag,2}$, then the single-tag and double-tag yields, N_1 and N_2, are given by

$$N_1 = \varepsilon_{tag,1} \times N_{B^+B^-} \quad (7)$$

$$N_2 = \varepsilon_{tag,1} \times \varepsilon_{tag,2} \times N_{B^+B^-}. \quad (8)$$

The ratio of the two cancels some of the common factors, yielding the following quantity to be determined in both data and MC simulations,

$$\varepsilon_{tag,2} = \frac{N_2}{N_1} \quad (9)$$

The *BABAR* experiment determines the number of single-tagged events by subtracting the combinatoric component under the D^0 mass distribution using an extrapolation of events from the D^0 mass sideband. This leaves a sample of events containing correctly reconstructed events, misreconstructed events from neutral B semileptonic decay, and events from $e^+e^- \rightarrow c\bar{c}$ continuum background events with real D^0 mesons paired with a combinatoric lepton. The correction to the tag efficiency is assumed to be equal for either the first or second tag, and is computed from the data and MC as,

$$c_{tag} = \frac{\varepsilon_{tag,2}^{data}}{\varepsilon_{tag,2}^{MC}} = \frac{N_2^{data}/N_1^{data}}{N_2^{MC}/N_1^{MC}} \quad (10)$$

The correction is computed by *BABAR* using only events in which the D^0 meson in the first tag B decays only into the $K^-\pi^+$ final state. This is cross-checked using a sample in which the D^0 meson from the first tag decays into only the $K^-\pi^+\pi^-\pi^+$ final state, yielding complementary results.

In both of the above methods, and across several iterations of semileptonic recoil-based analyses, *BABAR* has found the correction to be very close to 1.0. This suggest both that the assumptions in the above two methods are largely accurate, and also that existing simulations of these and the background decays are adequate for the purposes of modeling the decays. The correction has an associated systematic error, which is typically determined by propagating the statistical sample sizes of the double-tag and single-tag samples and assuming that the knowledge of the correction will thus improve with more data. Currently, the uncertainty on the correction is typically about 4%.

The second application of the double-tagged sample is to test the modeling of extra particles left in the detector after both Bs have been as completely reconstructed as possible. In the case of signal events, this typically means that the tag B is reconstructed up to any neutrinos in the final state (as in semileptonic tags), and that the signal B is reconstructed up to the neutrinos in its final state. This typically leaves particles in the event that are assumed to come from several sources: neutrals, such as photons, which arise from the electron-positron collider but not the interaction point; some low momentum charged particles associated with interactions between the beam and the beampipe; neutral clusters from hadronic showering in the calorimeter which fail to associate with a track; and detector noise. These sources would typically lead to a few extra neutral particles left in a signal event in about 20-30% of the reconstructed events (see, for instance, the signal MC of extra energy in Fig. **??**).

Double-tagged events are used to test the simulation of these extra neutral particles by fully reconstructing both Bs either semileptonically, hadronically, or in a hybrid configuration. This is assumed to then leave neutral and charged particles in the event that arise from the same sources as in signal. One crucial difference between signal events and double-tagged events is the number of hadrons in the final state. This is expected to lead to an increased presence of neutral clusters in the calorimeter which result from broad hadronic showers which are only partially associated with the appropriate reconstructed charged hadron.

An example of the use of the double-tags to test the extra energy simulation is the Belle Collaboration's hadronic-tagged search for $B^+ \rightarrow \tau^+\nu_\tau$. Belle constructs a hybrid double-tag sample (one hadronic B and one semileptonic B per event in the sample), and assumes that the extra neutral clusters remaining in these events comes from the same sources as in signal events. They compare the extra energy in data and MC (Fig. **??**) and use the difference as a variation on their PDF model for signal events.

Several conclusions can be drawn from such a comparison. First, existing detector simulations at the flavor factories appear to handle the variety of sources of extra neutral clusters fairly well, even in a moderate to high multiplicity final states of B decay. However, relying on the use of extra energy for a final extraction of the signal yield - in this example, as a means to fit the shape of the signal - does come at a price. In this case, the signal model parameters must be varied by the amount determined by separately fitting the data and MC double-tag events. The statistical sample size of the double tags will enter as an uncertainty in the fit model. With the current sample size of hybrid tags in the Belle analysis, about 240 events in the first two extra energy bins, we expect a shape uncertainty of about 6%, depending on the difference between the data and MC. While double-tag samples should grow proportionally with the accumulated data at flavor factories, achieving a high-precision measurement ($\sim 1\%$) of a rare decay branching fraction will require a significant increase in the data set if the extra energy continues to be a key component of the signal extraction.

Bibliography: BaBar Publications
Bibliography: Belle Publications
Bibliography