Physics Letters B 703 (2011) 475–478

Contents lists available at ScienceDirect

Physics Letters B

www.elsevier.com/locate/physletb

Bi-Event Subtraction Technique at hadron colliders

Bhaskar Dutta [a], Teruki Kamon [a,b], Nikolay Kolev [c], Abram Krislock [a,*]

[a] Department of Physics & Astronomy, Mitchell Institute for Fundamental Physics, Texas A&M University, College Station, TX 77843, USA
[b] Department of Physics, Kyungpook National University, Daegu 702-701, South Korea
[c] Department of Physics, University of Regina, Regina, SK S4S 0A2, Canada

ARTICLE INFO

Article history:
Received 1 June 2011
Received in revised form 5 August 2011
Accepted 10 August 2011
Available online 16 August 2011
Editor: M. Cvetič

ABSTRACT

We propose the Bi-Event Subtraction Technique (BEST) as a method of modeling and subtracting large portions of the combinatoric background during reconstruction of particle decay chains at hadron colliders. The combinatoric background arises when it is impossible to know experimentally which observed particles come from the decay chain of interest. The background shape can be modeled by combining observed particles from different collision events and be subtracted away, greatly reducing the overall background. This idea has been demonstrated in various experiments in the past. We generalize it by showing how to apply BEST multiple times in a row to fully reconstruct a cascade decay. We show the power of BEST with two simulated examples of its application towards reconstruction of the top quark and a supersymmetric decay chain at the Large Hadron Collider.

The Large Hadron Collider (LHC) is up and running since 2009. Many models of particle physics beyond the Standard Model (SM) predict new particles which can be tested at the LHC. Heavy colored objects are expected to be produced at the LHC, followed by a chain of subsequent decays, according to such new models. Thus, we must fully or partially reconstruct these cascade decays from the particles which can be detected. However, reconstructions of these decays become experimentally difficult because it is impossible to know which particles come from the cascade decay we wish to reconstruct. The inevitable inclusion of particles which do not come from the cascade decay of interest is referred to as combinatoric background.

This combinatoric background can be removed easily in some cases by powerful subtraction techniques. For instance, the Z boson can decay into oppositely charged, same flavored leptons: $Z \to e^+e^-/\mu^+\mu^-$. Leptons are easy to detect in the collider setting, and their charges can easily be measured. To reconstruct the Z boson from these leptons, it is easy to collect a sample of Opposite-Sign Same-Flavor (OSSF) lepton pairs and construct the dilepton invariant mass for each pair. To model the combinatoric background, a sample of Opposite-Sign Opposite-Flavor (OSOF) lepton pairs is selected as well. These OSOF lepton pairs cannot possibly both come from a single Z boson, and so they model the combinatoric background well. Performing the OSSF–OSOF subtraction of the invariant mass distributions (possibly using some normalization factor c), $h^{\mathrm{OSSF-OSOF}}(m_{\ell\ell}) = h^{\mathrm{OSSF}}(m_{\ell\ell}) - ch^{\mathrm{OSOF}}(m_{\ell\ell})$, yields a distribution which shows a clear peak of the Z boson mass.

However, such subtraction techniques are not available for jets, whose charges and flavors cannot so easily be determined. Thus, we introduce the Bi-Event Subtraction Technique (BEST) in which the combinatoric background of jets is modeled by combining jet information from a different event (or bi-event). This technique of modeling the combinatoric background by combining information from different events has been used before [1]. However, here we generalize it, by applying it to jets. Moreover, we have shown that it can be used multiple times for the same decay chain reconstruction.

The basic idea of BEST can be demonstrated for the reconstruction of the W boson decaying into two jets. For this case, a signal may be seen if a sample of jet pairs is collected for each event to construct the dijet invariant mass distribution, $h^{\mathrm{same}}(m_{jj})$. Here, the "same" suggests that the jet pairs come from the same event. Some of the jet pairs in the same event distribution may come from a single W boson decay in the events, while other jet pairs will be combinatoric background. By collecting another sample of jet pairs where each jet comes from a *different* event, the bi-event distribution, $h^{\mathrm{bi}}(m_{jj})$, can be formed. This bi-event distribution will have no jet pairs which come from a single W boson. Thus, this bi-event distribution models a large amount of the combinatoric background well. The $h^{\mathrm{bi}}(m_{jj})$ distribution can be normalized to the $h^{\mathrm{same}}(m_{jj})$ distribution in the region of pure background (well away from the W boson mass peak). For instance,

* Corresponding author.
E-mail address: crysanthemus@gmail.com (A. Krislock).

0370-2693/$ – see front matter
doi:10.1016/j.physletb.2011.08.029

the normalization factor can be calculated as

$$C_{jj}^{\mathrm{BEST}} = \frac{\int_{150\ \mathrm{GeV}}^{500\ \mathrm{GeV}} h^{\mathrm{same}}(m_{jj})\, dm_{jj}}{\int_{150\ \mathrm{GeV}}^{500\ \mathrm{GeV}} h^{\mathrm{bi}}(m_{jj})\, dm_{jj}}. \quad (1)$$

This normalization factor can be used when the shapes of these distributions are very close in this region. If the shapes of these distributions are not close, it could be due to some new physics. For instance, and additional resonance in the $h^{\mathrm{same}}(m_{jj})$ distribution could cause a mismatch in the shapes. However, it would be easy enough to recalculate the normalization taking an overall range which excludes the additional resonance. It should be noted that one needs a detailed systematic study of the shape from different physics processes. This is beyond the scope of this Letter.

Finally, the BEST is performed:

$$h^{\mathrm{BEST}}(m_{jj}) = h^{\mathrm{same}}(m_{jj}) - C_{jj}^{\mathrm{BEST}} h^{\mathrm{bi}}(m_{jj}). \quad (2)$$

The resulting dijet distribution shows a W boson mass peak with most of the combinatoric background removed.

If we wish to reconstruct decay chains involving these W bosons, we can take BEST even further. For instance, we can completely reconstruct the top quark from the decay chain $t \to bW \to bjj$. We can apply BEST again while combining the b jets with the reconstructed W bosons in order to reconstruct the top quark. However, this requires a more general application of BEST than has been used before.

For this example, we will refer to the same-event histograms by denoting the jets in the subscript as j and b for jets and b-jets respectively. For the bi-event histograms, we denote the jets in the subscript as j' and b'. Thus we now denote our histograms and normalization factor from Eqs. (1) and (2) as:

$$h^{\mathrm{same}}(m_{jj}) \equiv h_{jj}(M_{jj}), \quad (3a)$$
$$h^{\mathrm{bi}}(m_{jj}) \equiv h_{jj'}(M_{jj}), \quad (3b)$$
$$C_{jj}^{\mathrm{BEST}} \equiv C_{jj}^{\mathrm{BEST\#1}}, \quad (3c)$$
$$h^{\mathrm{BEST}}(m_{jj}) \equiv h_{jj}^{\mathrm{BEST\#1}}(m_{jj}). \quad (3d)$$

To combine the reconstructed W bosons with the b-jets to reconstruct the top quarks, we will need the following four additional histograms in order to perform two applications of BEST: $h_{bjj}(m_{bjj})$, $h_{bjj'}(m_{bjj})$, $h_{b'jj}(m_{bjj})$, and $h_{b'jj'}(m_{bjj})$. We perform the first BEST using the normalization factor calculated above in Eq. (1):

$$h_{bjj}^{\mathrm{BEST\#1}}(m_{bjj}) = h_{bjj}(m_{bjj}) - C_{jj}^{\mathrm{BEST\#1}} h_{bjj'}(m_{bjj}), \quad (4a)$$
$$h_{b'jj}^{\mathrm{BEST\#1}}(m_{bjj}) = h_{b'jj}(m_{bjj}) - C_{jj}^{\mathrm{BEST\#1}} h_{b'jj'}(m_{bjj}). \quad (4b)$$

Next we calculate another normalization factor for the second BEST which involves the combinatoric background of the b-jets. Once again, the range of this normalization factor is aimed at the region of pure background away from the top quark mass peak. Thus, it is calculated as:

$$C_{bjj}^{\mathrm{BEST\#2}} = \frac{\int_{200\ \mathrm{GeV}}^{500\ \mathrm{GeV}} h_{bjj}^{\mathrm{BEST\#1}}(m_{bjj})\, dm_{bjj}}{\int_{200\ \mathrm{GeV}}^{500\ \mathrm{GeV}} h_{b'jj}^{\mathrm{BEST\#1}}(m_{bjj})\, dm_{bjj}}. \quad (5)$$

With this normalization factor, we can finally perform the second BEST:

$$h_{bjj}^{\mathrm{BEST\#2}}(m_{bjj}) = h_{bjj}^{\mathrm{BEST\#1}}(m_{bjj}) - C_{bjj}^{\mathrm{BEST\#2}} h_{b'jj}^{\mathrm{BEST\#1}}(m_{bjj}). \quad (6)$$

Here, the resulting histogram will show a clean top quark mass peak with most of the combinatoric background removed. To clean up the resulting distribution even more, other subtraction techniques can also be employed, such as a sideband subtraction for the W boson reconstruction. Each additional subtraction will double the number of initial histograms which are needed for all of the subtractions.

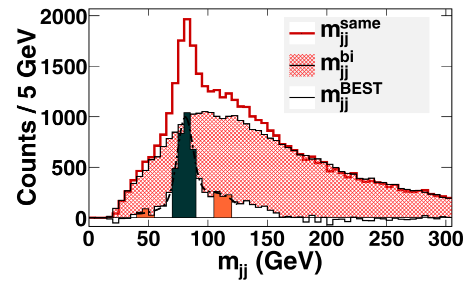

Fig. 1. The dijet invariant mass distribution, m_{jj}. This plot shows the same-event (m_{jj}^{same}), bi-event (m_{jj}^{bi}), and BEST (m_{jj}^{BEST}) distributions as described in the text. The BEST distribution is fitted with a Gaussian plus cubic function, to find the W boson mass peak and surrounding background. The BEST distribution is also split up into regions for a sideband subtraction used for reconstructing an invariant mass between a W boson and a b tagged jet. The W region is dark cyan filled, while the sidebands are orange filled. For an integrated luminosity of 2 fb^{-1}, we find the W boson mass, $m_W = 81.11 \pm 0.32$ GeV. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this Letter.)

We demonstrate this powerful technique by using it to extract $W \to jj$ for (i) $t\bar{t}$ events at $\sqrt{s} = 7$ TeV and (ii) SUSY events at $\sqrt{s} = 14$ TeV within LHC simulations.

For the $t\bar{t}$ events, we generate hard scattering LHC collision events using ALPGEN [2], perform the cascade decays with PYTHIA [3], and perform a LHC detector simulation using PGS4 [4]. The W + jets events are the main source of background for finding the top quark, so we generate these events in the same way. This background is mixed in randomly, according to production cross-sections, with our $t\bar{t}$ events. After PGS4 is finished with these events, we select events for analysis with the following cuts [5]: (i) Number of leptons, $N_\ell = 1$, where $p_T^{(\ell)} \geqslant 20$ GeV and $p_{T,\mathrm{iso}}^{(\ell)} \leqslant 0.1 \times p_T^{(\ell)}$; (ii) Missing transverse energy, $\not{E}_T \geqslant 20$ GeV; (iii) Number of jets, $N_j \geqslant 3$, where $p_T^{(j)} \geqslant 30$ GeV and at least one jet has been tightly b-tagged [4]; (iv) Number of taus, $N_\tau = 0$ for taus with $p_T^{(\tau)} \geqslant 20$ GeV [4].

With our events selected in this way, we pair up jets (which are not b-tagged) to fill the same-event and bi-event $h(m_{jj})$ distributions as described above. Each jet pair must have $\Delta R \geqslant 0.4$. To fill the bi-event distribution, we refer to jets from the previous event which has passed the same cuts as listed above. Once the distributions are filled with all events, we normalize the shape of the $h^{\mathrm{bi}}(m_{jj})$ distribution as described by Eq. (1). Then we perform our BEST. The result of this subtraction can be seen in Fig. 1, which shows a drastic reduction in the background obscuring the W boson reconstruction. Note that the bi-event distribution models the combinatoric background of any jet pairs which are not correlated by decay chains or event kinematics. Thus, BEST in this case removes (i) the combinatoric background from events with W bosons (coming from t decays) and (ii) uncorrelated jet pairs coming from our W + jets background sample.

Once we have found the W boson with this first application of BEST, we can combine the W boson with a b-jet to find the top quark. To remove additional background from the W signal, we perform a sideband subtraction. To do this, we split up the dijet signal into a W boson mass region, where 70 GeV $\leqslant m_{jj} \leqslant$ 90 GeV, and two sideband regions, 40 GeV $\leqslant m_{jj} \leqslant$ 55 GeV and 105 GeV $\leqslant m_{jj} \leqslant$ 120 GeV. We form the dijet (W) plus b invariant mass, keeping track of whether the dijet system was in the W window or sideband windows. In this way we make the W band ($h_{W\ \mathrm{band,\ BEST}}(m_{bW})$) and sideband ($h^{\mathrm{SB,\ BEST}}(m_{bW})$) distributions.