LHCb-INT-2015-022
June 17, 2015

Using jets to identify semi-leptonic B decays

X. Cid Vidal[1].

[1] *CERN, Geneva, Switzerland*

Abstract

This note introduces a new method that uses jets to help identifying semi-leptonic b decays. In particular, this method exploits the characteristics of jets when such decays are present and explores different configurations for the jets reconstruction in terms of the use of the leptons. Specific applications of the method are shown in the context of the searches for $B^0 \to \mu^+\mu^-$ and $B^+ \to \mu^+\nu_\mu$, for which semi-leptonic b decays are the most important source of background.

1 Introduction

The presence of neutrinos in the final state, which are almost undetectable experimentally, makes very challenging the identification of semi-leptonic b decays ($b \to l\nu_l + X$). This is particularly true in the context of hadron colliders, in which the full reconstruction of the events is very difficult. Nevertheless, given the large branching fractions typically associated to semi-leptonic b decays, they are very important for several physics analysis. On the one hand, these decays become the most important background in several rare-decays searches, such as $B_s^0 \to \mu^+\mu^-$ [1] and $B^+ \to \mu^+\nu_\mu$ [2]. On the other, they have interest themselves for studying the CKM angles and possible new sources of CP–violation [3, 4]. Therefore, improving our capacity to detect semi-leptonic b decays would be very useful.

This work introduces a new method that goes in the direction explained above, helping to improve the identification of semi-leptonic b decays. This is achieved by treating the available information in a novel way and also by adding other not being used within other techniques. The underlying idea behind this is the fact jets are the the most generic way of reconstructing quarks. In this regard, in a $b \to l\nu_l + X$ decay, jets are the most inclusive method to get the X, whatever this is (see figure 1). The lepton can be alternatively inside or outside this jet, with some degree of freedom on the reconstruction side to force one or the other situation. This also affects the second b or $\bar{b}$ usually produced in the same process. Since jets include information also from neutrals, this approach has the additional advantage of adding information not taken into account by other methods, such as isolations [5, 6]. Specifically, for semi-leptonic b decays, the features of the jets will be very particular, and also the relation between the jet and the exclusively reconstructed charged lepton.

This note is divided as follows. The samples used are introduced in section 2. Then, generic examples of the types of variables that can be built are shown in section 3. More in particular, applications of this method are shown in the context of $B_s^0 \to \mu^+\mu^-$ and $B^+ \to \mu^+\nu_\mu$ analyses in sections 4 and 5. Finally, conclusions are drawn in section 6.

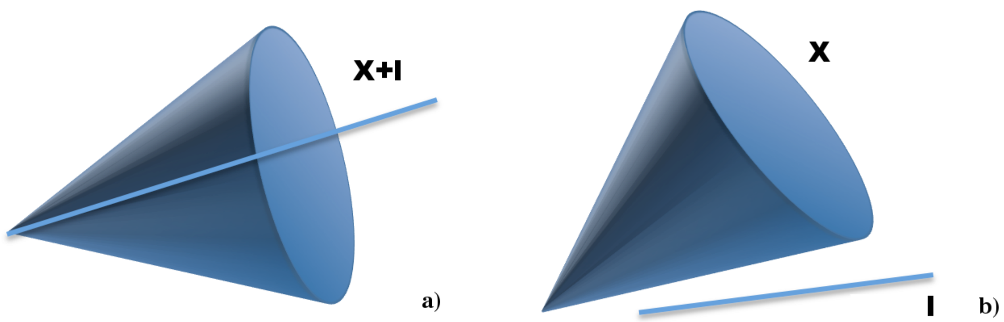

Figure 1: Generic reconstruction of a semi-leptonic b decay through jets. The lepton may be reconstructed, a) inside the jet; b) outside the jet.

Name	Event Type	Stats
$B_s^0 \to \mu^+\mu^-$	13112001	$\sim$ 20 M
$b\bar{b} \to \mu^-\mu^+X_1X_2$	10012009	$\sim$ 8 M
$b\bar{b}$	10011001 (stripping filtered)	$\sim$ 50K
$c\bar{c}$	10011001 (stripping filtered)	$\sim$ 12K
$B^+ \to \mu^+\nu_\mu$	12511002 (stripping filtered)	$\sim$ 50K

Table 1: MC2012 simulated samples used for this note.

2 Samples used

The simulated samples used for the analyses in this note, together with the relevant statistics, are presented in table 1. These samples are from the MC2012 generation.

3 Variable construction

As explained in the introduction, the properties of the jets will be different depending on the type of b decay under scrutiny. This can be also taken into account when performing the reconstruction of the jets.

Jet reconstruction at LHCb uses a *Particle Flow* [7] approach, similar to the one of CMS [8]. This implies that the event is fully reconstructed through all the particles that compose it, with this list of particles being fed later into the jet reconstruction algorithm. This procedure has advantages for dealing with semi-leptonic decays. In this regard, when a fully reconstructed lepton is selected, it can be removed from the list of particle flow inputs to force the reconstruction of the rest of the b decay (the X in figure 1 *b)*).

Once the jet reconstruction has been performed, it is possible to define different variables to exploit the properties of the jets when a semi-leptonic b decay is present. These properties will depend on how the jet was reconstructed (e.g., whether the lepton was left out of the jet reconstruction or not) and will be focused not only on the jets themselves but also on how they relate with the leptons. Examples of these variables are:

- ΔR between the jet and the lepton, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$
- Jet p_{T}: transverse momentum of the jet
- $Lepton(p_{\mathrm{T}})/jet(p_{\mathrm{T}})$: ratio of transverse momenta between the lepton and the jet that contains this lepton
- Jet tag: b-tagging of the jet. It usually relates to the presence of secondary vertices made out of detached particles (with high impact parameter) inside the jet. For this analysis, the tagging was made simply by matching the jets with LHCb's topological trigger objects [9]

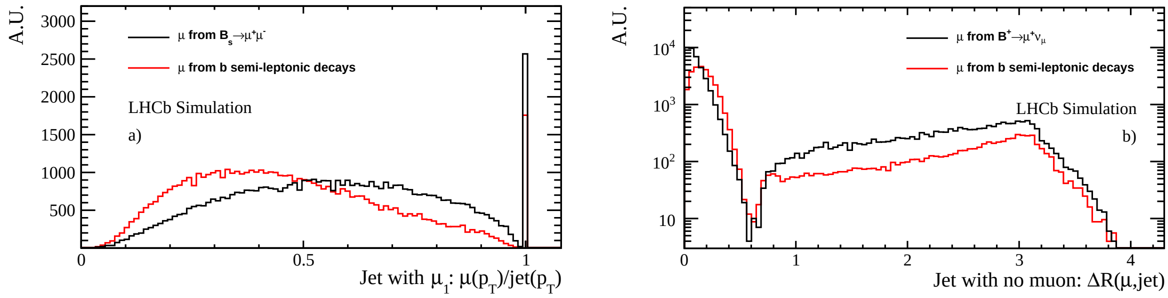

Figure 2: Examples of possible discriminating variables for $B_s^0 \to \mu^+\mu^-$ and $B^+ \to \mu^+\nu_\mu$. a) Muon p_{T} divided by the p_{T} of the jet containing the muon, for the opposite situation; b) ΔR between the muon and the jet for the case in which the lepton is left out of the jet reconstruction.

- Jet width: Average ΔR between the jet and its components weighted by the transverse momentum of each of the components

For the specific case of the tagging variable, the wide range of tagging possibilities available makes easy to conceive a better separation power than what is shown in this note. Specifically, since in most of the cases the semi-leptonic decays occur to charm final states, the use of a charm tagger [10] could improve the performance of the tagging variables.

Figure 2 shows examples of these discrimination variables for muons from $B_s^0 \to \mu^+\mu^-$ or $B^+ \to \mu^+\nu_\mu$ and muons from semi-leptonic b decays.

4 Use in $B_s^0 \to \mu^+\mu^-$

The searches for $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ offer an excellent chance to put into practice the method described in this note. One of the most important backgrounds in the search for both channels is the combinatorial $b\bar{b} \to \mu^-\mu^+X_1X_2$, i.e., a pair of b $\bar{b}$ quarks both decaying semi-leptonically to a muon and something else, so that these two muons together make the B candidate. All the machinery illustrated so far can be used, therefore, to separate this type of background from the signal. In this specific case, the fact that two semi-leptonic b decays are present increases the possibilities in terms of jet building. In this regard, apart from the possibilities seen in figure 1, it is also possible to force the presence of the B candidate as a single input for jet reconstruction. When this happens, both muons will be in the same jet. Therefore, for background, this jet will contain both b quarks at the same time, while for signal a second jet containing a second b quark is possible too. Figure 3 summarizes the situation in terms of jet reconstruction for the $B_s^0 \to \mu^+\mu^-$ analysis.

Figures 4, 5 and 6 show examples of the discrimination power between $B_s^0 \to \mu^+\mu^-$ and $b\bar{b} \to \mu^-\mu^+X_1X_2$ for variables built according to the different jet reconstruction possibilities shown in figure 3. Even if more variables could be possible, only the most discriminant examples are shown here.

In order to get an estimate of the overall discrimination power achieved by this method, the variables shown in figures 4, 5 and 6 were combined using the TMVA package [11], in a single quantity called $\mathrm{BDT_{jets}}$. Figure 7 shows the distribution of this variable for $B_s^0 \to \mu^+\mu^-$ and $b\bar{b} \to \mu^-\mu^+X_1X_2$. Even if correlations still exist at some degree between all the variables that enter into $\mathrm{BDT_{jets}}$ [12], they have been studied to be different in signal and background, so useful for the discrimination between these.

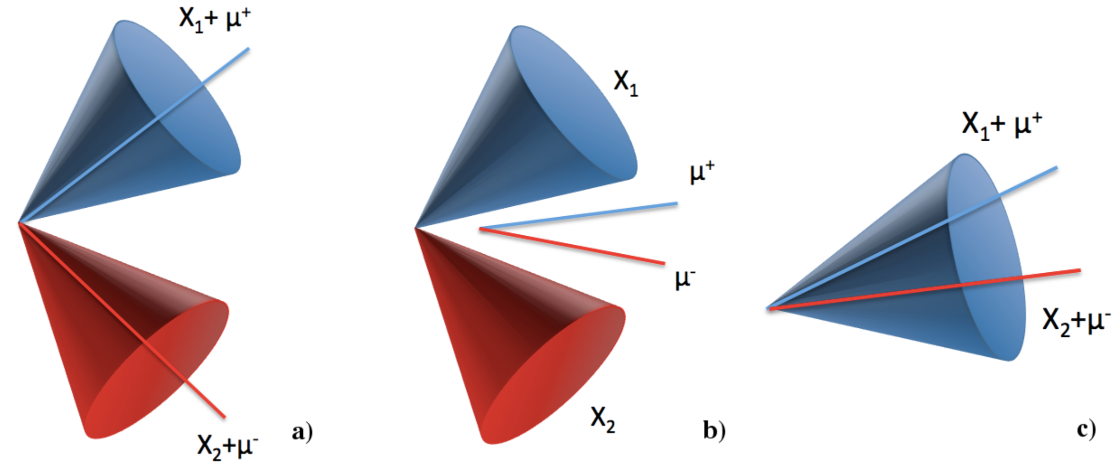

Figure 3: Jet reconstruction on $b\bar{b} \to \mu^-\mu^+X_1X_2$, main background in the search for $B_s^0 \to \mu^+\mu^-$. a) Both muons are reconstructed in different jets; b) both muons are left out of the list of inputs for jet reconstruction; c) the B candidate is forced as a single input to the jets reconstruction, so both muons are in the same jet.

A full study of the effect of these variables for the $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ analyses can be found in [12]. In this study, the use of $\mathrm{BDT_{jets}}$ was seen to improve the discrimination power of the MVA discriminant used in the analyses published so far. Furthermore, this variable was seen to be poorly correlated to the rest used to discriminate $B_s^0 \to \mu^+\mu^-$ from $b\bar{b} \to \mu^-\mu^+X_1X_2$, being among the best ranked in terms of separation between these, according to the TMVA package. This also proves the point made at the beginning that this method uses information not taken into account till now to discriminate signal from background. Finally, apart from MC, the improvement seen was checked to be coherent in data too, with a larger background rejection for the same signal efficiency.

5 Use in $B^+ \to \mu^+\nu_\mu$

In a similar way to $B_s^0 \to \mu^+\mu^-$, the search for $B^+ \to \mu^+\nu_\mu$ and similar states will have semi-leptonic b decays as main background. However, in this specific case, semi-leptonic c decays have also been seen to play an important role [13, 14]. With only one muon selected, the situation concerning jet reconstruction will be basically the one showed in figure 1. The only difference here is that the jet in this case could correspond to a charm quark, which will imply different properties of the associated jets.

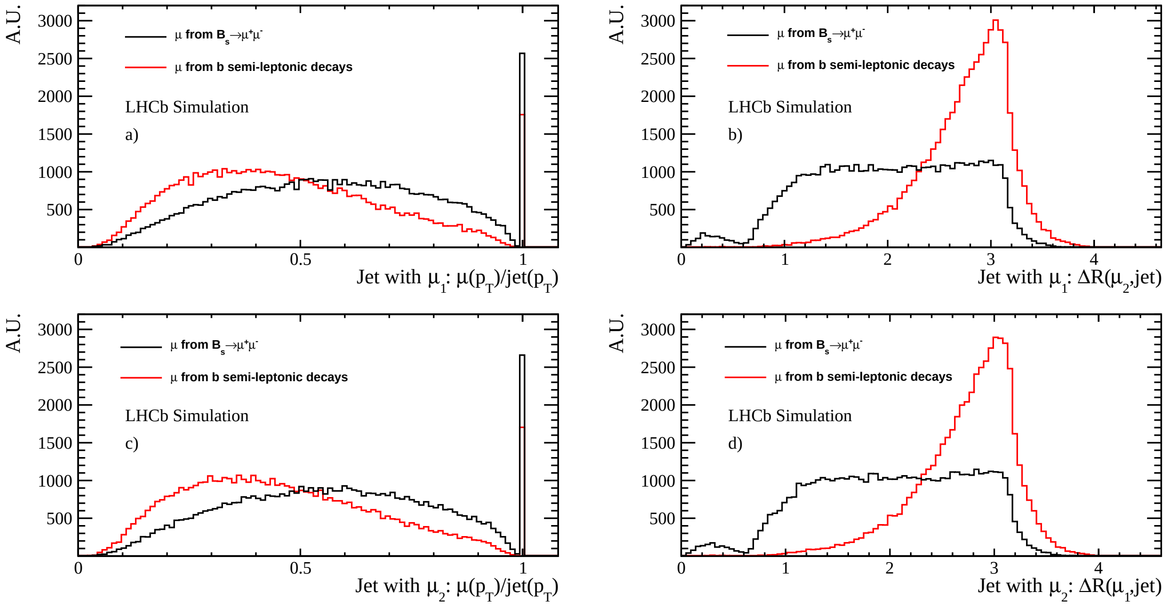

Figure 4: Discrimination between signal and background in the search for $B_s^0 \to \mu^+\mu^-$ in the case in which the muons are found within different jets (situation *a)* in figure 3). a) $\mu(p_{\mathrm{T}})/jet(p_{\mathrm{T}})$, i.e., ratio of transverse momenta between the muon and the jet that contains it; b) ΔR between the jet containing the muon and the other muon; c) and d), the same for the second muon.

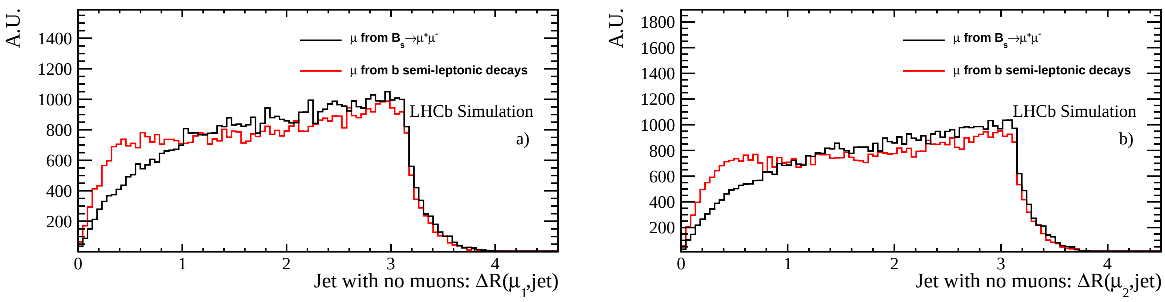

Figure 5: Discrimination between signal and background in the search for $B_s^0 \to \mu^+\mu^-$ in the case in which the muons are left out of the jet reconstruction (situation *b)* in figure 3). a) ΔR between the highest p_{T} jet and one of the muons; b) ΔR between the highest p_{T} jet and the second muon.

Figures 8 and 9 show examples of the discrimination power between signal, semi-leptonic b decays and semi-leptonic c decays for the variables built using the two jet reconstruction configurations showed in figure 1. From this figure it is also possible to conclude that this method can also be useful also to discriminate b and c semi-leptonic decays.

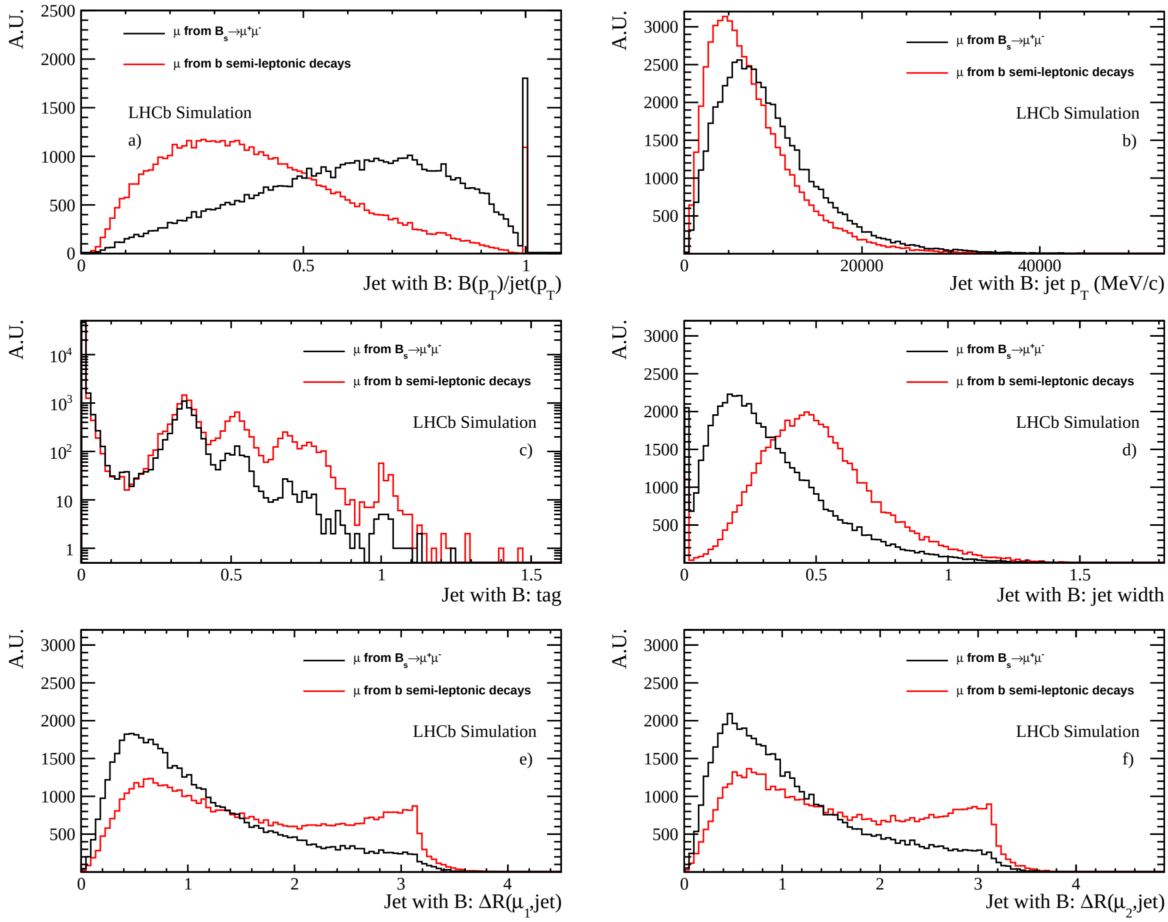

Figure 6: Discrimination between signal and background in the search for $B^0_s \to \mu^+\mu^-$ in the case in which the B candidate is used as a single input for the jet reconstruction (situation *c)* in figure 3). a) $B(p_{\rm T})/jet(p_{\rm T})$, i.e., ratio of transverse momenta between the B candidate and the jet that contains it; b) $p_{\rm T}$ of this jet; c) tagging of this jet; d) width of this jet; e) ΔR between this jet and the first muon; f) ΔR between this jet and the other muon.

6 Conclusions

This note has introduced a novel method to make use of jets also for flavour physics measurements. In particular, jets have been found to be useful to identify semileptonic b decays, and even to offer a chance to separate semi-leptonic b and c decays from each other and from other types of decays. The method proposed uses the information available in a novel way and it also adds other information from the detector not being used so far, in particular that of neutral particles.

Specific applications of the method have been shown in the context of the analyses of $B^0_s \to \mu^+\mu^-$ and $B^+ \to \mu^+\nu_\mu$. Variables have been proposed in both cases, but it is hard to define a completely general set of variables. Instead, the application of the method is

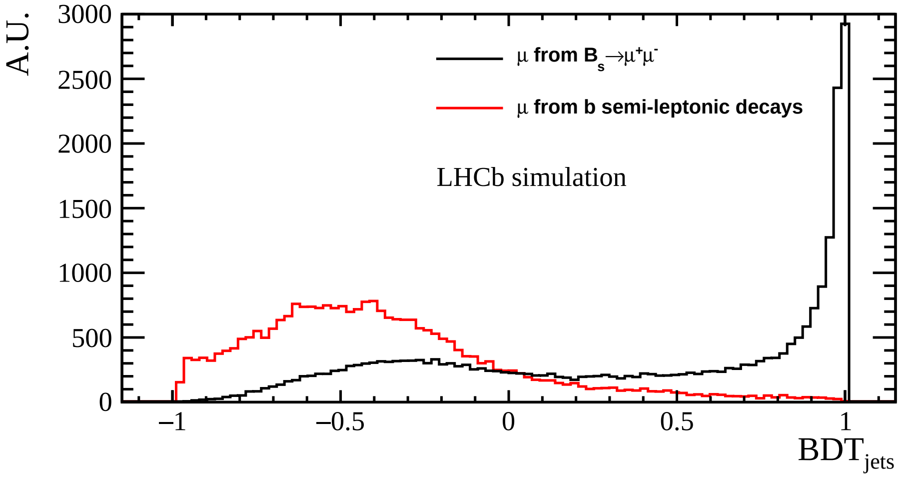

Figure 7: Overall discrimination between $B_s^0 \to \mu^+\mu^-$ and $b\bar{b} \to \mu^-\mu^+X_1X_2$ achieved combining all the variables shown in figures 4, 5 and 6. The BDT variable was obtained using the TMVA package [11].

analysis dependent. This also affects the reconstruction of the jets themselves, in particular regarding how leptons are treated. As an example, the fact that two semi-leptonic b decays are relevant for the $B_s^0 \to \mu^+\mu^-$ analysis offers more possibilities than in the case of $B^+ \to \mu^+\nu_\mu$.

More applications of this method seem possible for other analyses, particularly in the context of hadron colliders. As explained, while the idea is general, the implementation will be specific although several handles are already in place in terms of technical machinery.

Acknowledgements

I would like to acknowledge C. Potterat and V. Coco for many discussions that lead to this method. Also thanks for their kind support for the development of the technical tools needed to implement all these ideas.

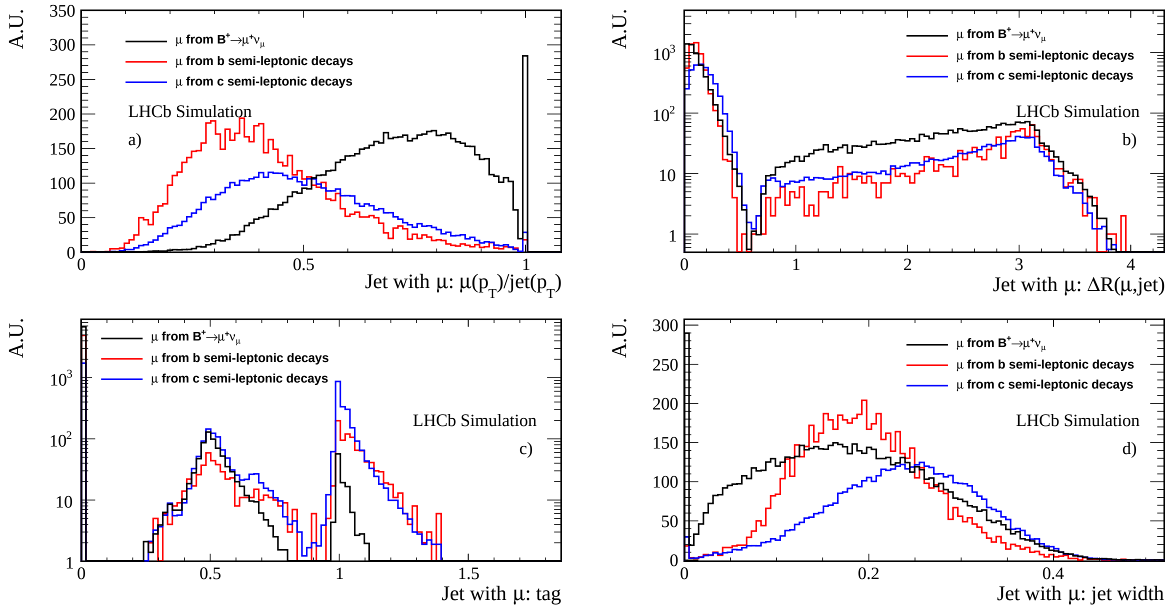

Figure 8: Discrimination between signal and background (both b and c semi-leptonic decays) in the search for $B^+ \to \mu^+\nu_\mu$ in the case in which the muon is found within a jet (situation *a)* in figure 1). a) $\mu(p_\mathrm{T})/jet(p_\mathrm{T})$, i.e., ratio of transverse momenta between the muon and the jet that contains it; b) ΔR between this jet the muon itself; c) tagging of this jet; d) width of this jet.

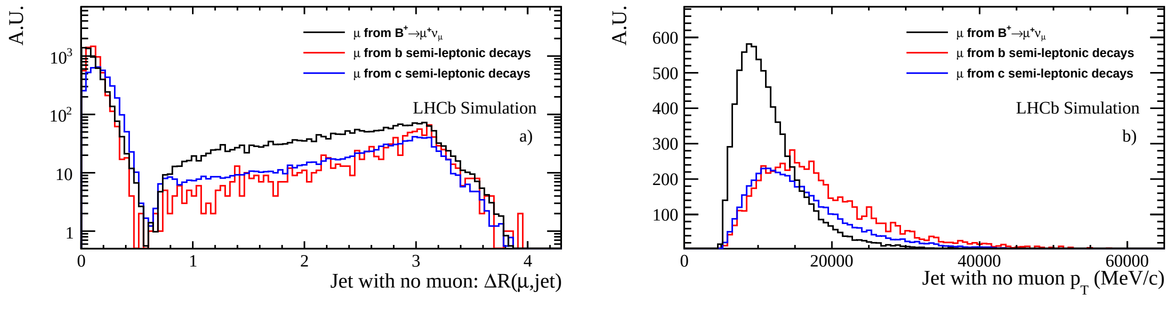

Figure 9: Discrimination between signal and background (both b and c semi-leptonic decays) in the search for $B^+ \to \mu^+\nu_\mu$ in the case in which the muon is not used as an input for the jet reconstruction (situation *b)* in figure 1). a) ΔR between the jet with highest p_T and the muon; b) transverse momentum of this jet.

References

[1] CMS, LHCb, V. Khachatryan *et al.*, *Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data*, Nature (2015) , arXiv:1411.4413.

[2] Belle, Y. Yook *et al.*, *Search for $B^+ \to e^+\nu$ and $B^+ \to \mu^+\nu$ decays using hadronic tagging*, arXiv:1503.0125.

[3] LHCb, R. Aaij *et al.*, *Determination of the quark coupling strength $|V_{ub}|$ using baryonic decays*, arXiv:1504.0156.

[4] LHCb, R. Aaij *et al.*, *Measurement of the semileptonic CP asymmetry in $B^0 - \overline{B}^0$ mixing*, Phys. Rev. Lett. **114** (2015) 041601, arXiv:1409.8586.

[5] G. Mancinelli and J. Serrano, *Study of Muon Isolation in the $B_s^0 \to \mu^+\mu^-$ Channel*, Tech. Rep. LHCb-INT-2010-011. CERN-LHCb-INT-2010-011, CERN, Geneva, Mar, 2010.

[6] C. Adrover *et al.*, *Search for the rare decays $B_s^0 \to \mu^+\mu^-$ and $B^0 \to \mu^+\mu^-$ with 1 fb $^{-1}$ of 2012 data*, Tech. Rep. LHCb-ANA-2012-081, CERN, Geneva, Aug, 2012.

[7] O. Augusto, *LHCb Jet Reconstruction*, Dec, 2012. Poster-2012-246.

[8] CMS, F. Beaudette, *The CMS Particle Flow Algorithm*, arXiv:1401.8155.

[9] V. V. Gligorov, C. Thomas, and M. Williams, *The HLT inclusive B triggers*, Tech. Rep. LHCb-PUB-2011-016. CERN-LHCb-PUB-2011-016. LHCb-INT-2011-030, CERN, Geneva, Sep, 2011. LHCb-INT-2011-030.

[10] LHCb, R. Aaij *et al.*, *Identification of beauty and charm quark jets at LHCb*, arXiv:1504.0767.

[11] A. Hocker *et al.*, *TMVA: Toolkit for multivariate data analysis*, PoS ACAT **184** (2007).

[12] X. Cid Vidal, *Update on Jet Isolations*, 2015. Talk at $B \to \mu^+\mu^-$ Workshop in Cambridge, April 14, University of Cambridge (UK).

[13] M. Vieites *et al.*, *Search for $B^+ \to \mu^+\nu_\mu$* , 2013. Talk at at VRD meeting, October 16, CERN (Switzerland).

[14] X. Cid Vidal, *$B^+ \to \mu^+\nu_\mu$ status*, 2014. Talk at at VRD meeting, October 22, CERN (Switzerland).

Appendix: Explanation of the variables in the ntuples

Tables 2 and 3 list the variables saved by the TupleTool *TupleToolJetsForB* (under *Phys/DecayTreeTupleJets*) and by the RelatedInfo tool *RelInfoJetsVariables* (under *Phys/IsolationTools*) in the LHCb software framework and their explanation.

JETNOMU1PX	p_x of the highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU1PY	p_y of the highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU1PZ	p_z of the highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU1PT	p_T of the highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU1JETWIDTH	Width (average ΔR between the jet and its components weighted by the p_T of each component, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$) of the highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU1NNTAG	b-tag of the highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU1MNF	Maximum neutral fraction (p_T of the highest p_T neutral particle within the jet normalized to the total jet p_T) of the highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU2PX	p_x of the second highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU2PY	p_y of the second highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU2PZ	p_z of the second highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU2PT	p_T of the second highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU2JETWIDTH	Width (average ΔR between the jet and its components weighted by the p_T of each component, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$) of the second highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU2NNTAG	b-tag of the second highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU2MNF	Maximum neutral fraction (p_T of the highest p_T neutral particle within the jet normalized to the total jet p_T) of the second highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU3PX	p_x of the third highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU3PY	p_y of the third highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU3PZ	p_z of the third highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU3PT	p_T of the third highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU3JETWIDTH	Width (average ΔR between the jet and its components weighted by the p_T of each component, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$) of the third highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU3NNTAG	b-tag of the third highest p_T jet when signal muons are left out of the jet reconstruction
JETNOMU3MNF	Maximum neutral fraction (p_T of the highest p_T neutral particle within the jet normalized to the total jet p_T) of the third highest p_T jet when signal muons are left out of the jet reconstruction

Table 2: Meaning of the variables produced by the LHCb tools to build jet variables.

JETMU1PX	p_x of the jet containing muon 1 when muons are used in the jet reconstruction
JETMU1PY	p_y of the jet containing muon 1 when muons are used in the jet reconstruction
JETMU1PZ	p_z of the jet containing muon 1 when muons are used in the jet reconstruction
JETMU1PT	p_{T} of the jet containing muon 1 when muons are used in the jet reconstruction
JETMU1JETWIDTH	Width (average ΔR between the jet and its components weighted by the p_{T} of each component, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$) of the jet containing muon 1 when muons are used in the jet reconstruction
JETMU1NNTAG	b-tag of the jet containing muon 1 when muons are used in the jet reconstruction
JETMU1MNF	Maximum neutral fraction (p_{T} of the highest p_{T} neutral particle within the jet normalized to the total jet p_{T}) of the jet containing muon 1 when muons are used in the jet reconstruction
JETMU2PX	p_x of the jet containing muon 2 when muons are used in the jet reconstruction
JETMU2PY	p_y of the jet containing muon 2 when muons are used in the jet reconstruction
JETMU2PZ	p_z of the jet containing muon 2 when muons are used in the jet reconstruction
JETMU2PT	p_{T} of the jet containing muon 2 when muons are used in the jet reconstruction
JETMU2JETWIDTH	Width (average ΔR between the jet and its components weighted by the p_{T} of each component, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$) of the jet containing muon 2 when muons are used in the jet reconstruction
JETMU2NNTAG	b-tag of the jet containing muon 2 when muons are used in the jet reconstruction
JETMU2MNF	Maximum neutral fraction (p_{T} of the highest p_{T} neutral particle within the jet normalized to the total jet p_{T}) of the jet containing muon 2 when muons are used in the jet reconstruction
JETBPX	p_x of the jet containing both muons when the B candidate is used as input for the jet reconstruction
JETBPY	p_y of the jet containing both muons when the B candidate is used as input for the jet reconstruction
JETBPZ	p_z of the jet containing both muons when the B candidate is used as input for the jet reconstruction
JETBPT	p_{T} of the jet containing both muons when the B candidate is used as input for the jet reconstruction
JETBJETWIDTH	Width (average ΔR between the jet and its components weighted by the p_{T} of each component, where $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$) of the jet containing both muons when the B candidate is used as input for the jet reconstruction
JETBNNTAG	b-tag of the jet containing both muons when the B candidate is used as input for the jet reconstruction
JETBMNF	Maximum neutral fraction (p_{T} of the highest p_{T} neutral particle within the jet normalized to the total jet p_{T}) of the jet containing both muons when the B candidate is used as input for the jet reconstruction

Table 3: Meaning of the variables produced by the LHCb tools to build jet variables.