UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “GALILEO GALILEI”
Corso di Laurea in Fisica

TESI DI LAUREA

Metodi di ripesamento e applicazioni all’esperimento LHCb

Reweighting methods and applications at the LHCb experiment

Laureando: **Pietro Bernardi**
Matricola: 591527

Relatore: **Dott. Alessandro Bertolin**

Anno Accademico 2015 - 2016

Contents

Chapter 1

Introduction

This thesis focuses on Monte Carlo reweighting methods. The typical approach will be presented first, while a novel method [1] will be shown later. Both approaches will be applied to two cases of interest related to an ongoing analysis of B_s^0 decay data from the LHCb detector [2, 3] at the LHC hadron collider [4] at CERN.

The LHCb detector is a single-arm forward spectrometer covering the pseudorapidity range $2 < \eta < 5$, designed for the study of particles containing beauty or charm quarks.
The B_s^0 decay of interest is:

$$B_s^0 \to D_s^{*-}\pi^+, \tag{1.1}$$

the π^+ track from the B_s^0 decay is called the "bachelor" track. Monte Carlo[1] samples are available for it. The data samples used for this analysis corresponds to these decay modes:

$$B_s^0 \to D_s^{*-}\pi^+ \tag{1.2}$$

and:

$$B_s^0 \to D_s^{-}\pi^+. \tag{1.3}$$

The two reweighting cases of interest are:

- Reweighting of data to background discriminating variables.
- Reweighting of tagging related variables.

Data to background discrimination is an essential procedure that allows to extract useful information from the collected data, the signal. It can be performed by means of a Boosted Decision Tree [5, 6](BDT). The BDT is a machine learning software device that is trained on a particular sample by feeding a certain number of variables to its inputs. The trained BDT can then be used on real data to classify background and signal events. The variables chosen as input are those that show the highest discriminating power.

In general, flavour tagging in B physics is the process by which the flavour of the beauty meson, at production time, is determined independently of its decay time decay pattern. Hence if a B^0 to $\overline{B^0}$, or vice versa, oscillation occurs it can be recognized as such. Flavour tagging at LHCb is performed looking to the event topology on the so called Opposite Side (OS) of the event with respect to the direction of the B_s^0 meson to be tagged. In this case the decision achieved is clearly independent of the B_s^0 meson decay time decay pattern.
Using appropriate variables also same side (SS) flavour tagging is possible. These variables

[1] From here on Monte Carlo will be substituted by MC for brevity.

must be independent from the B_s^0 meson to be tagged.

Results obtained with different reweighting methods will be compared for each of the two physics cases.
In Chapter 2, a description of the two reweighting methods is given. In Chapter 3, the two cases of interest described above are presented. In Chapter 4, conclusions are drawn.

Chapter 2

Reweighting Methods

Reweighting is a procedure by which the distribution of a certain variable is changed in shape by applying suitable *weights*. In general, the goal of this procedure is to make the MC simulation to show better agreement with data. These weights can be calculated following different approaches. In this work, MC weights will be calculated in two ways: the "bin by bin" method and a new approach recently developed.

2.1 Typical Approach

The typical approach to reweighting requires MC and data to be binned in histograms. The MC distribution for the variable being reweighted is called the *original* distribution, while the data distribution for the *same* variable is named the *target* distribution. Once both distributions have been binned, an histogram division is performed. This constitutes the typical approach to reweighting, sometimes called "bin by bin" method.

If N bins are used in the histograms, this procedure will return N weights, given by the following equation:

$$w_i = \frac{T_i}{O_i} \tag{2.1}$$

where T_i stands for *Target i-th bin content* and O_i stands for *Original i-th bin content.* The w_i are the weights, obtained for each of the N bins, related to the variable for which data and MC distributions were divided. The ensemble of these variable-related weights will be called *weighting rule* (WR) from now on. For instance the η[1] WR will be written as $WR(\eta)$.

Once this procedure is completed, the calculated WR can also be used to reweight other MC distributions.

The WR calculated using a certain variable to reweight other variables does not guarantee that the reweighted MC distributions will show a better agreement with their corresponding data distributions. This is due to the fact that, in general, *variables are not completely independent* but there might be correlations. So reweighting with respect to a given variable could introduce biases and distortions in other variables' distributions that share some correlation with the WR variable. This ultimately can increase the disagreement between MC and data.

One example of this situation could be the B_s^0 meson transverse momentum, Ptr, and its pseudorapidity. These two variables are intimately connected so reweighting MC using only the η

[1] Here and in the following η indicates the pseudorapidity of a particle, defined as:

$$\eta = -\log\tan\frac{\theta}{2}$$

WR, will increase the disagreement in the transverse momentum's distribution and viceversa. A technique to lessen this effect that still uses histogram division will be presented too.

Moreover, when the typical approach is used, the binning has to be chosen such that a sufficiently large number of events is contained in each bin. If not, this process will introduce large fluctuations in the reweighted distribution. Some *artifacts* in the reweighted distribution might appear.

2.2 GBR Approach

The new approach to reweighting called Gradient Boosted Reweighter (GBR) uses machine learning algorithms [1, 7]. It is implemented in Python programming language.
The GBR algorithm generates Boosted Decision Trees [5, 6], that are iteratively built during the training stage.
The GBR is trained on a subset of both MC and data distributions. It's also possible to run the training on multi-dimensional distributions to handle many variables at once. This is welcome when it's not clear whether correlations exist between those variables. With the typical approach instead the difficulty increases as the number of variables increases. After the training stage, the GBR can predict weights for a MC distribution given a target data distribution. Different weights are provided for individual events. So the problem of unstable WR due to small number of events in a given bin is ruled out because the GBR algorithm is unbinned.

Chapter 3

Practical Applications

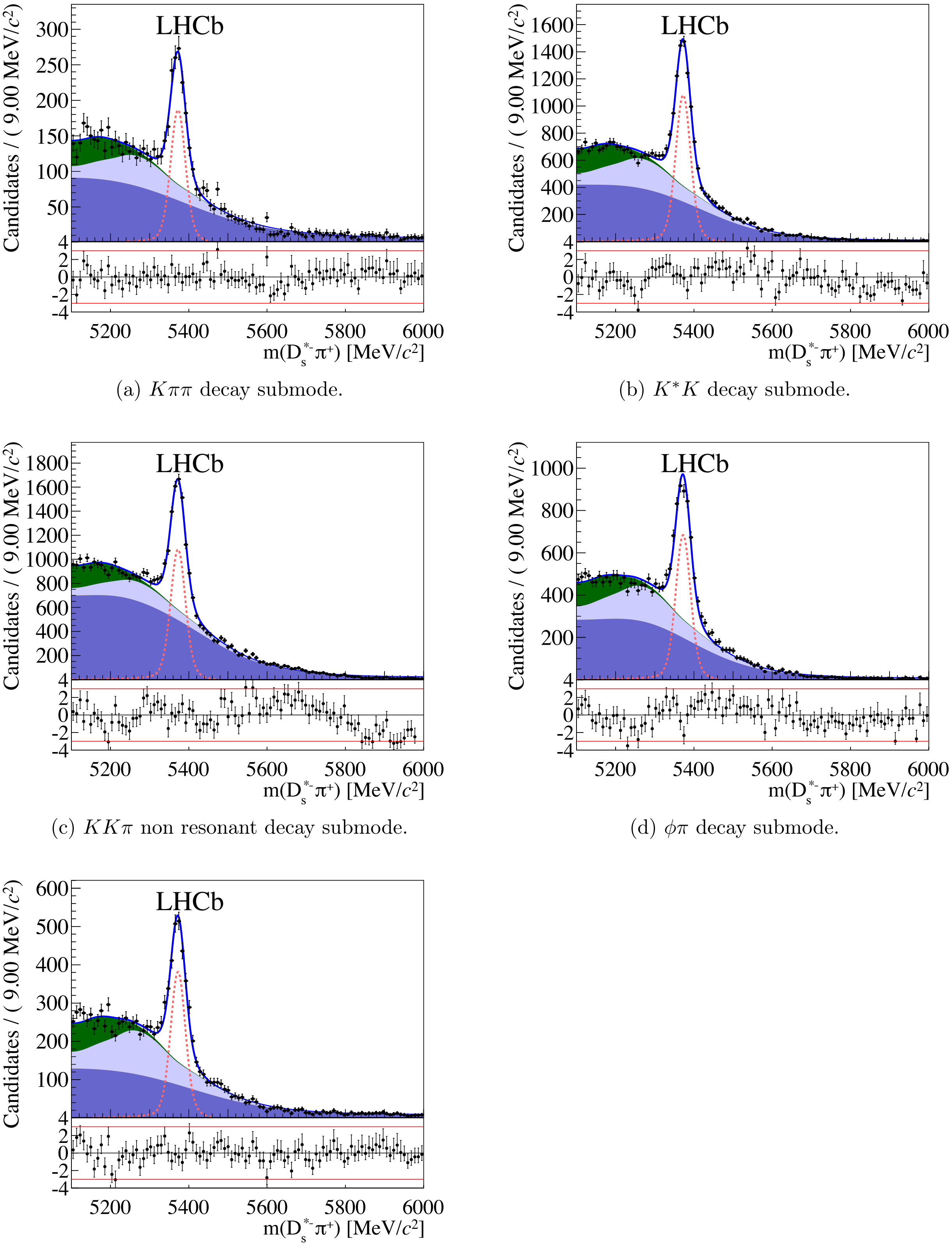
3.1 Data to background discriminating variables

The first case of interest that was analyzed involves the BDT input variables used for the analysis of the (1.1) decay mode. The full decay can be written as:

$$\begin{array}{cccccc} B_s^0 & \to & D_s^* & \pi^+ & & \\ & & D_s^* & \to & D_s & \gamma \\ & & & & D_s & \to & KK\pi(\phi\pi, K^*K, non-resonant) \\ & & & & D_s & \to & K\pi\pi \\ & & & & D_s & \to & \pi\pi\pi \end{array}$$

At present time the BDT used to separate signal and background in that particular decay, trained on the $KK\pi$ sample, has 14 input variables. These variables are listed below:

- Photon related:
    - *Ph_Ptr*: the transverse momentum.
    - *PtrRel*: the transverse momentum with respect to the D_s flight direction[1].
    - *Ph_CL*: the confidence level.
    - *Ph_Eta*: the pseudorapidity η.
    - *Ph_isNotE*: the probability of photon not being an electron[1].
    - *Cos_ThetaS*: the cosine of the D_s polar angle in the D_s^* rest frame.
- Transverse momentum of the final state charged particles: bachelor, *Km*, *Kp*, *Pi*.
- B_s^0 related:
    - *Ds_DIRAOri*: the angle between the D_s momentum vector and the vector connecting its origin and decay vertices.
    - *Bs_DIRAOwn*: the angle between the B_s^0 momentum vector and the vector connecting its production and decay vertices.
    - *Bs_IpChi2Own*: the χ^2_{IP} defined as the difference in χ^2 of the associated primary vertex, *PV*, reconstructed with and without the considered particle.
    - *Bs_RFD*: the radial flight direction.



(a) $K\pi\pi$ decay submode.

(b) K^*K decay submode.

(c) $KK\pi$ non resonant decay submode.

(d) $\phi\pi$ decay submode.

(e) $\pi\pi\pi$ decay submode.

Figure 3.1: B_s^0 mass fits for each of the five decay submodes.

3.1.1 sWeight calculation

In this chapter the $D_s^{*-}\pi^+$MC will be reweighted using $D_s^{*-}\pi^+$data as target. However, since data always comes with background, it is mandatory to get rid of these background contributions i.e. data *have to be sWeighted* [8].
The sWeight procedure assigns weights based on the data invariant mass distribution. Candidate events close to the invariant mass peak are getting larger weights. Conversely, where the signal is expected to be low and background dominates, smaller weights are applied.
The *sWeights* are obtained from invariant mass fits to the B_s^0 spectra for each of the five decay submodes shown in figure 3.1. These decays modes are defined as being mutually exclusive.
So, for each submode, a set of sWeights is available and by keeping track of which decay submode a given data point belongs to, it is possible to assign the right sWeight to it.

3.1.2 Chosen Monte Carlo weighting rules

Several WR are calculated using the following variables: the pseudorapidity and the transverse momentum of the B_s^0 and the number of tracks, $nTracks$. Given the fact that η and Ptr are not independent an additional $2-dim$ WR in the (η, Ptr) plane is calculated.
In order to use all informations available, a Global Weighting Rule[2] is computed as the product between the $nTracks$ WR and the $2-dim$ one:

$$\boxed{GWR(nTracks, \eta, Ptr) = WR(nTracks) \cdot WR(\eta, Ptr)} \tag{3.1}$$

The $nTracks$ and (η, pt) WR are calculated following the two approaches outlined in Chapter 2. The reason why the GWR is expressed as a product and not as a $3-dim$ WR is a reasonable compromise: the $nTracks$ variable is not strongly correlated to η and Ptr, so one can separate these variables and obtain a product of two terms, each independent with respect to the other.

Moreover, a $3-dim$ WR would be quite difficult to calculate in practice following the typical approach. However, it is worth noticing that such a difficulty would not arise when using the GBR approach, given the algorithm's native multi-dimensional handling capabilities.

Typical approach

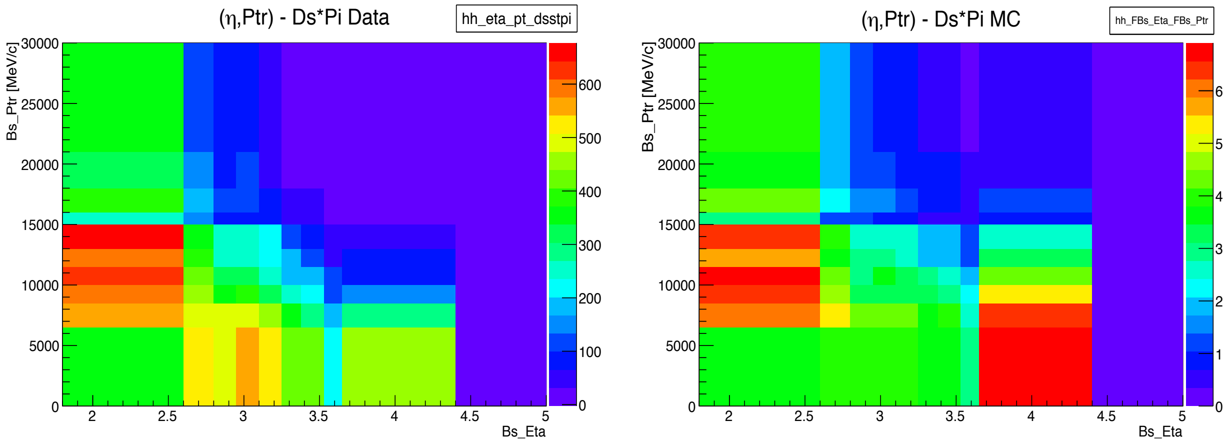
In this case, the $WR(nTracks)$ is calculated by simple $1-dim$ histogram division with variable sized bins to improve the WR stability given the small number of events with a large number of tracks. The $WR(\eta, Ptr)$ is calculated by binning the η variable on the abscissa and the Ptr on the ordinate. To improve stability of the resulting WR, variable bins were used so that each bin would not have an excessively small number of events. Bins were adjusted manually, no automatic procedures were used.

Figure 3.2 shows $D_s^{*-}\pi^+$data and MC respectively, binned in a $2-dim$ histogram. The region at high η and high Ptr is scarcely populated because of kinematic constraints. These histograms are normalized, divided and the $WR(\eta, Ptr)$ is obtained.

The data to MC comparisons for the 14 reweighted BDT input variables, described in the top part of Section 3.1, are shown in figures 3.3 and 3.4. Reweighting results will be discussed in the next section.

[1] As will became clear in the following, this variable will be excluded in the next version of the BDT.

[2] From now on GWR.



(a) $D_s^{*-}\pi^+$data (η, Ptr)

(b) $D_s^{*-}\pi^+$MC (η, Ptr)

Figure 3.2: Binned B_s^0 η and Ptr distributions using variable size bins for both variables.

GBR approach

In this case while the $WR(nTracks)$ is still calculated following the typical approach, the (η, Ptr) WR is predicted using the GBR algorithm. So the resulting GWR is a *hybrid* rule[3]. This approach is chosen because of the absence of correlation between the number of tracks and the other two B_s^0 kinematic variables.

The GBR algorithm is trained on a subset of the B_s^0 η and Ptr data and MC distributions. The MC is treated as the *original* distribution while the data is the *target*. The reweighted MC distributions are similar to those shown in figures 3.3 and 3.4. A comparison between these reweighting approaches will be given in the next section.

3.1.3 Observations

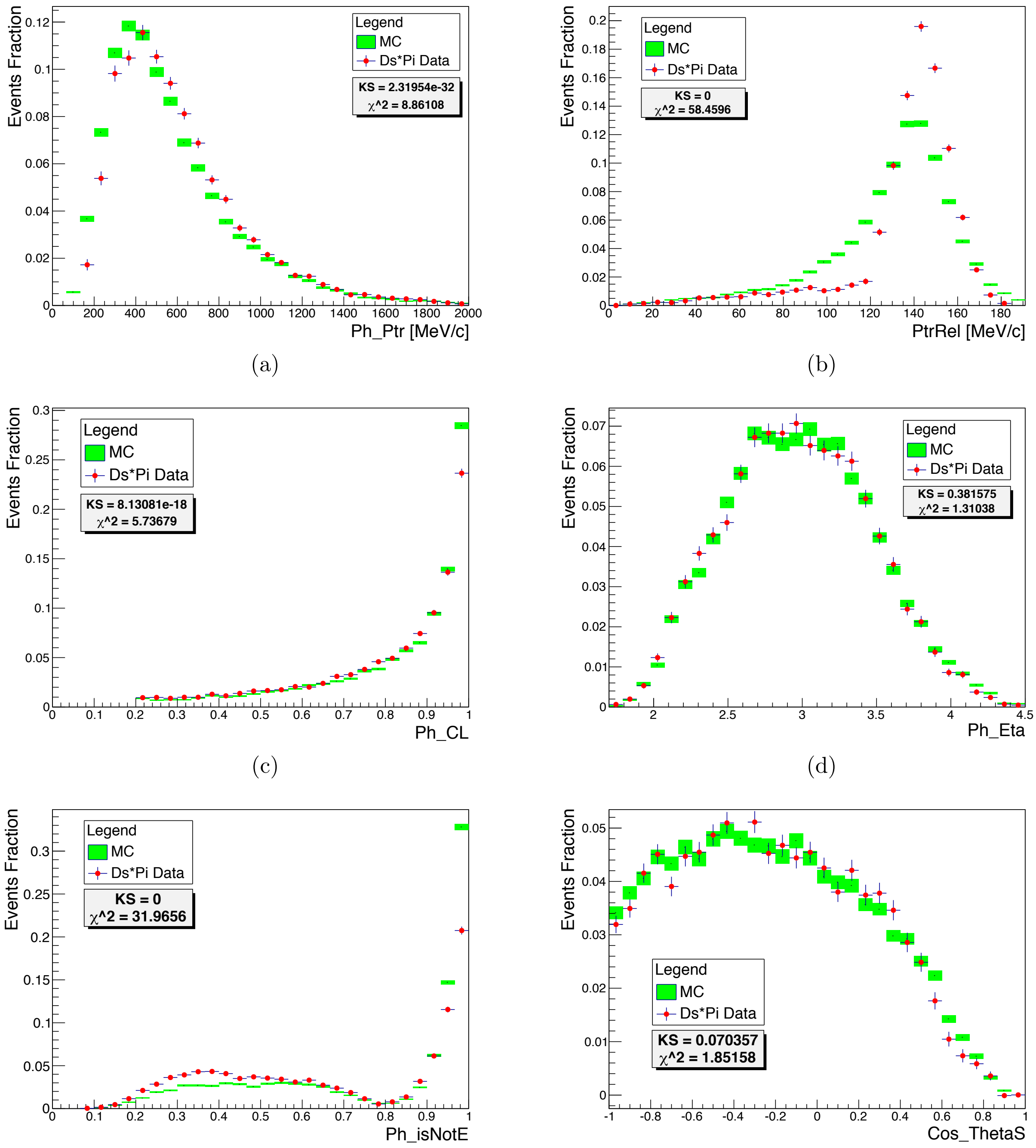
Reweighting results are summarized, in figure 3.5, using the χ^2/NDF calculated between $D_s^{*-}\pi^+$MC and $D_s^{*-}\pi^+$sWeighted data for all the 14 BDT input variables' distributions. Here NDF indicates the number of degrees of freedom as returned by the $ROOT$ analysis package. Also 5 additional distributions are added:

- the BDT output variable, BDT_Var.
- the η, Ptr and ϕ of the B_s^0 , Bs_Eta, Bs_Ptr and Bs_Phi, respectively.
- the number of tracks, $nTracks$.

It is evident that reweighting greatly improves agreement between data and simulation, generally for all variables.
The typical approach and the hybrid one, seem to produce nearly identical results, but some exceptions are found:

- The typical approach leads to a better agreement for:
  - Bac_Ptr
  - Km_Ptr

[3] A pure GBR approach is easily feasible by feeding $(nTracks, \eta, Ptr)$ as input for the GBR algorithm and thus the $3 - dim$ GWR is obtained.



(a)

(b)

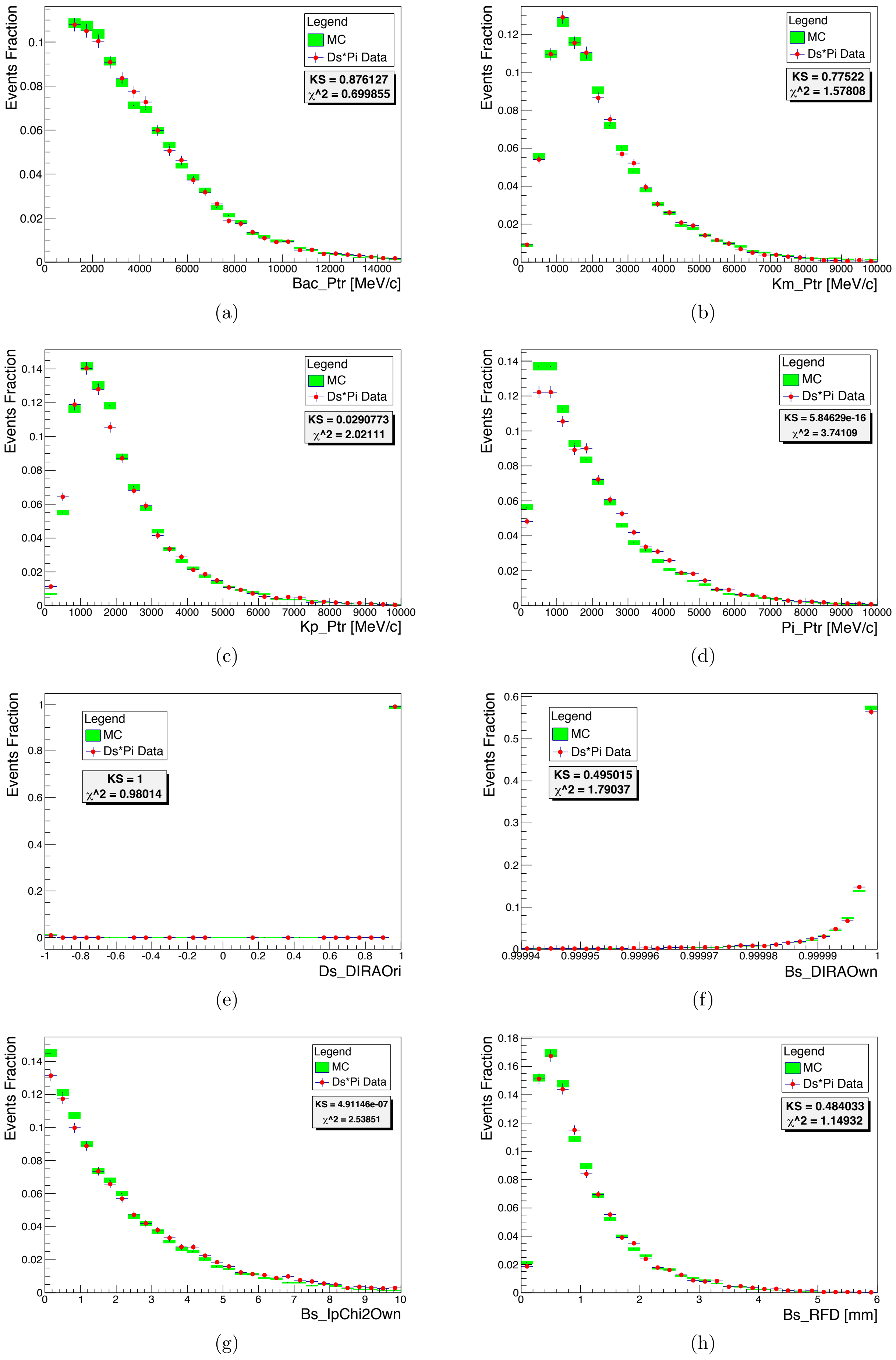
(c)

(d)

(e)

(f)

Figure 3.3: BDT input variables (1 of 2).



(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

Figure 3.4: BDT input variables (2 of 2).

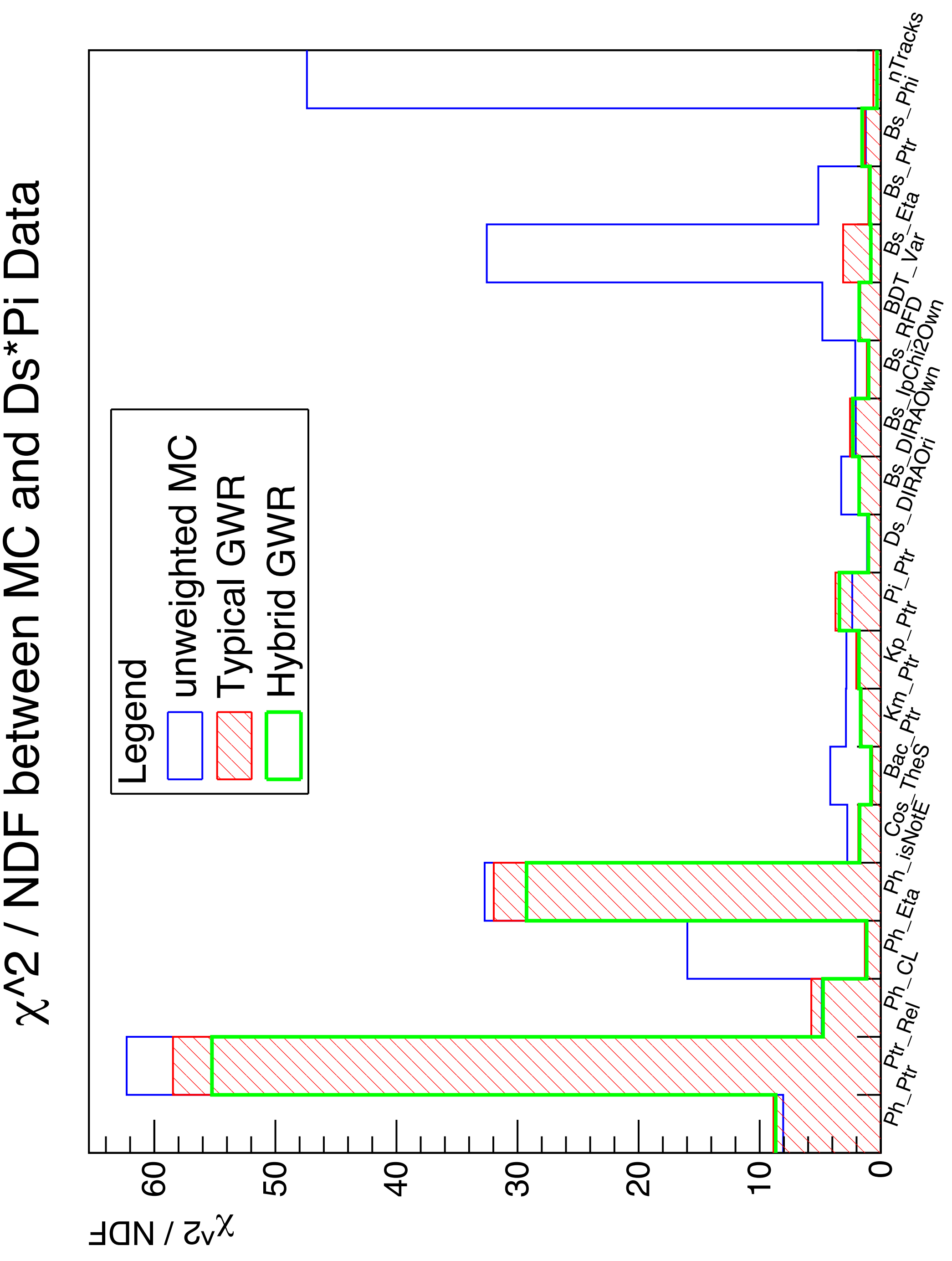


Figure 3.5: χ^2/NDF between $D_s^{*-}\pi^+$data and $D_s^{*-}\pi^+$MC with the two approaches to reweighting.

Variable	KS typical	KS GBR
Ph_Eta	0.3816	0.1235
Cos_TheS	0.0704	0.1200
Bac_Ptr	0.8761	0.5868
Km_Ptr	0.7752	0.7174
Kp_Ptr	0.0291	0.0348
$Ds\_DIRAOri$	1	1
$Bs\_DIRAOwn$	0.4950	0.7990
Bs_RFD	0.4840	0.5494
BDT_Var	0.0017	0.0010
Bs_Eta	0.7852	0.6328
Bs_Ptr	0.1017	0.1856
Bs_Phi	0.3236	0.3285
$nTracks$	0.1111	0.8445

Table 3.1: Kolmogorov - Smirnov tests results for all the BDT variables.

  – $Ds\_DIRAOri$
  – BDT_Var
  – Bs_Phi

- The GBR approach performs better for all the other variables. Bs_Eta is the variable that shows the greatest difference in χ^2 between the two reweighting styles.

Results from the Kolmogorov-Smirnov test between data and MC distributions are shown in table 3.1. The distributions for which the KS test returned zero are not shown. These are: Ph_Ptr, $PtrRel$, Ph_CL, Ph_isNotE, Pi_Ptr, $Bs\_IpChi2Own$. The different tests behaviour of the χ^2 and the KS for some variables deserves further investigations.

As an outcome of this study it results clearly that the variables $PtrRel$ and Ph_isNotE are not reproduced at all by the MC simulation. And furthermore a reweighting in the key variables $nTracks$ and B_s^0 η and Ptr is not significantly improving the agreement. Hence these variables will be discarded in the next iteration of the B_s^0 BDT training.

3.2 Flavour tagging related variables

The second case of interest involves the following variables:

- η, Ptr and ϕ of the B_s^0 .
- The number of tracks $nTracks$.

These have a large impact on B_s^0 tagging performances [9]. The goal of this analysis is to check the so called *portability* of the calibration of the taggers, for both the opposite side, OS, and the same side, SS, taggers. In order to accomplish this task, $D_s^{*-}\pi^+$MC is reweighted using $D_s^-\pi^+$data as target. These are the data that were originally used to calibrate both taggers.

3.2.1 Adopted procedure

Mistag probability for the Right and Wrong tag samples

The mistag probability for the Right Tag (RT) and Wrong Tag (WT) samples are built as histograms separately, for the OS and the SS cases, by checking the values of these variables:

Bs_TRUEID	Tagger Decision	Output
$+531$	$+1$	Rightly Tagged B_s^0
$+531$	-1	Wrongly Tagged B_s^0
-531	$+1$	Wrongly Tagged $\overline{B_s^0}$
-531	-1	Rightly Tagged $\overline{B_s^0}$

Table 3.2: Bs_TRUEID and tagger decision combinations.

- Bs_TRUEID: this variable assumes two possible values: ±531, where +531 stands for B_s^0 and −531 for $\overline{B_s^0}$.
- $Bs\_TAGDECISION\_OS$: this variable holds the OS tagger's decision. It assumes three values: 0 for untagged events, 1 for the B_s^0 and −1 for the $\overline{B_s^0}$.
- $Bs\_SS\_nnetKaon\_DEC$: this variables holds the SS tagger's decision. The values are as for the OS tagger.

The values of these variables are checked following the scheme given in table 3.2. The mistag probabilities are obtained from the variables:

- $Bs\_TAGOMEGA\_OS$ for the OS
- $Bs\_SS\_nnetKaon\_PROB$ for the SS

The $\omega(\eta)$ distribution

The ratio:

$$\frac{WT}{RT + WT}$$

as a function of the mistag probability is called the $\omega(\eta)$ distribution. In the LHCb tagging jargon, the mistag probabilities are unfortunately labelled with the symbol η that was up to now used to identify the pseudorapidity. It can be calculated separately for each of the OS and SS cases.

For each of the OS and SS cases, the ω distribution as a function of the mistag probability, η, is computed. The $\omega(\eta)$ distribution is obtained by taking the WT distribution and dividing it by the sum of the WT and RT distributions:

$$\omega(\eta) = \frac{WT}{RT + WT} \tag{3.2}$$

Following the procedure described in [9], each ω distribution is fitted with the linear function:

$$\omega(\eta) = p_0 + p_1 \cdot (\eta - <\eta>) \tag{3.3}$$

where p_0 and p_1 are fit parameters and $<\eta>$ is the mean η of the $RT + WT$ histogram. As shown in [9], for a "well calibrated tagger", the following results are expected:

- $p_1 \equiv 1$
- $p_0 - p1 < \eta > \equiv 0$

when fitting the $\omega(\eta)$ histogram.

3.2.2 Chosen weights

The GWR is the same as in the previous application, see section 3.1.2 and equation (3.1). The $WR(nTracks)$ is calculated with the typical approach, while the $WR(\eta, Ptr)$ is obtained both with the typical and the hybrid methods. Unweighted MC will also be used to compare results between the two scenarios.

In the OS case histograms have fixed size bins. In the SS case they have less bins of variable sizes to cope with the fact that there is small statistics at low values of η.

3.2.3 Calibration Fits

Usage of centroids instead of bin centers

The $\omega(\eta)$ distribution is built using binned MC given the fact that both RT and WT distributions are binned too. To fit these distributions correctly, one needs to calculate the centroids for each bin of the $\omega(\eta)$ histogram. This is achieved by adding a significant number of sub-bins in each bin. From the histogram with the sub-bins the average value of η in each bin is easily obtained.

The centroids are generally shifted from the center of the bin towards the region of the bin with higher statistics. They shift back to the center when the bin is small but contains a sufficiently large number of events.

In general, the centroid is a more representative point than the center of the bin, since its position encodes the information on the distribution of the events in the bin. This information is lost if only the center of the bin is considered.

The centroid positions are then plotted on the x axis, that holds the mistag probability η. The y axis holds the ω distribution given by (3.2).

Fitting procedure

Fits were performed both for the OS and SS cases using the function (3.3). In the OS case the full range $0 - 0.5$ was used. In the SS cases two different fit ranges are used because of very low statistics at low values of η that significantly shifted the first bin[4] centroid towards the upper bin edge.

In the following, the term *intercept* will refer to the quantity:

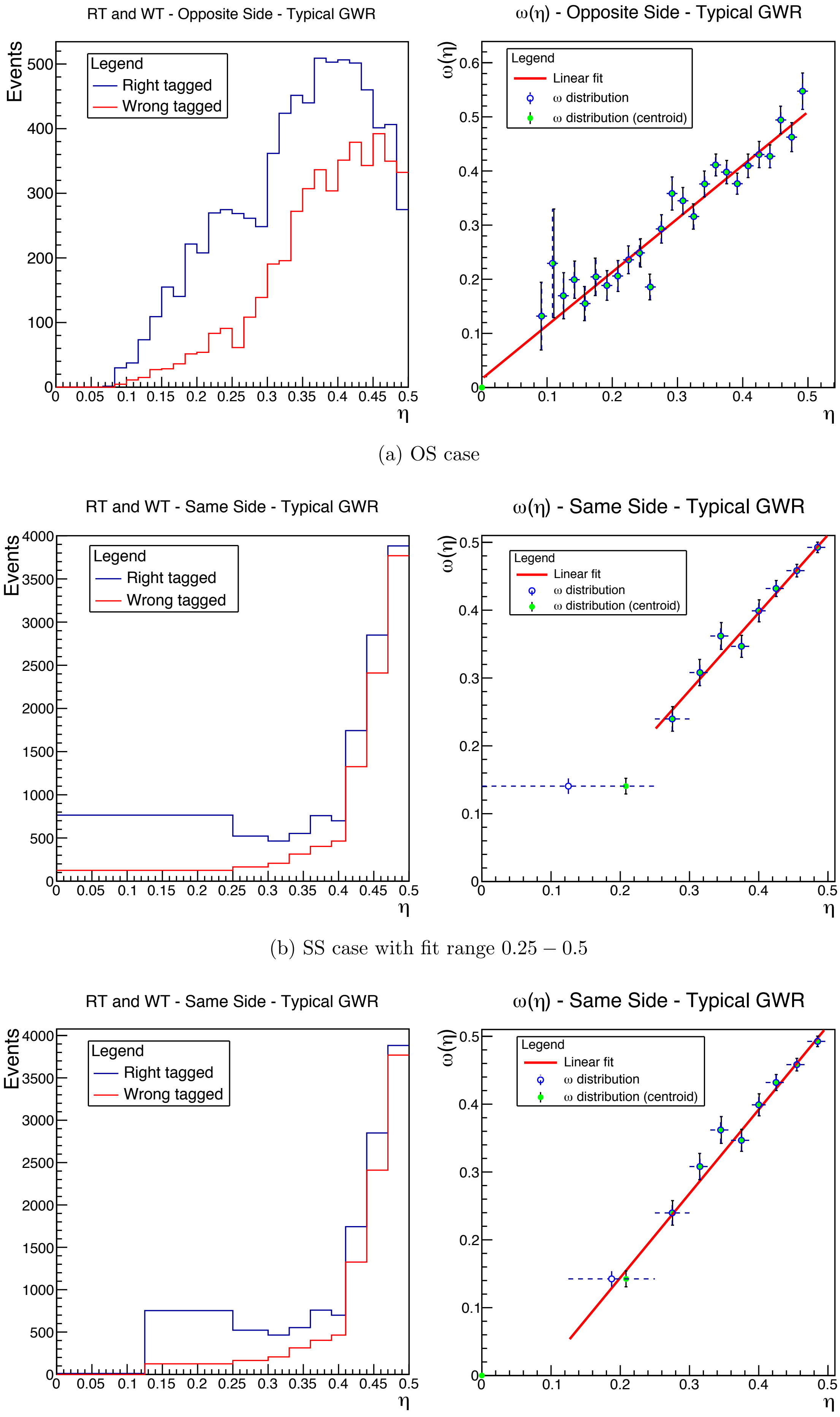
$$p0 - p1 < \eta > \tag{3.4}$$

where $p0$ is the usual fit line's intersection with the plot's ordinate.

Results

Results for the typical GWR are shown in figures 3.6a, 3.6b and 3.6c: empty circles indicate the bin centers, while the green points represent the centroids. Similar results are obtained when the MC is reweighted using the hybrid GWR. Plots are omitted. Results are summarized in table 3.3 and discussed in the next section. A comparison with the unweighted MC will also be given in the next section.

[4] First bin for SS has the range: $[0, 0.25]$



(a) OS case

(b) SS case with fit range $0.25 - 0.5$

(c) SS case with fit range $0.125 - 0.5$

Figure 3.6: The $\omega(\eta)$ distributions, with calibration fits, applying the typical GWR to the $D_s^{*-}\pi^+$MC.

Opposite Side - OS			
GWR	Slope	Intercept	χ^2/NDF
Typical	0.9863 ± 0.0526	0.0111 ± 0.0056	1.50
Hybrid	0.9742 ± 0.0542	0.0048 ± 0.0059	1.81
Same Side - SS - Fit range $[0.25, 0.5]$			
GWR	Slope	Intercept	χ^2/NDF
Typical	1.1472 ± 0.0709	0 ± 0.0045	0.82
Hybrid	1.142 ± 0.0736	-0.0043 ± 0.0048	0.59
Same Side - SS - Fit range $[0.125, 0.5]$			
GWR	Slope	Intercept	χ^2/NDF
Typical	1.2390 ± 0.0450	-0.0012 ± 0.0044	1.10
Hybrid	1.2068 ± 0.0488	-0.0049 ± 0.0048	0.71

Table 3.3: Calibration fit results with reduced χ^2 with reweighted MC. The "intercept" on third column refers to the quantity defined in equation (3.4).

3.2.4 Observations

In the OS case, fit results show a higher χ^2 when the hybrid GWR is used to reweight the MC (1.81) with respect to the typical approach (1.50). But in both cases the fit parameters are the ones expected for a "well calibrated tagger". It is worth noticing that when using the hybrid GWR in the OS case, some large spikes appeared in the WT and RT distributions. These spikes were traced down to very large weights predicted by the GBR algorithm. These, in turn, were the results of single "pathological" events. These very large weights, being nothing more than artifacts, were arbitrarily set to 1.0 to stabilize the reweighted distributions.

In the SS case, fit results show that:

- For the fit range $[0.25 - 0.5]$: the goodness of fit[5] is excellent both with the usage of typical and hybrid GWR to reweight the MC. The fitted parameters are close to the expected values i.e. 1 and 0.

- For the fit range $[0.125 - 0.5]$: this range was chosen by splitting the interval $[0, 0.25]$ in two equal sized bins, the first ranging from 0 to 0.125 and the second from 0.125 to 0.25. Because the $0 - 0.125$ bin is unpopulated, it was discarded and only the $0.125 - 0.20$ bin was kept. Results here show a slight worsening of GOF being slightly more evident when the typical GWR is used. The slope of the fit increases with respect to the previous fit range while the intercept remains close to 0. This will require additional investigations.

It is interesting to investigate whether reweighting increases or decreases the GOF with respect to the unweighted case. Results for the unweighted case are as follows:

Opposite Side			
Range	Slope	Intercept	χ^2/NDF
$[0, 0.5]$	0.9483 ± 0.0425	0.0076 ± 0.0044	1.79
Same Side			
Range	Slope	Intercept	χ^2/NDF
$[0.25, 0.5]$	1.1836 ± 0.0555	-0.0127 ± 0.0035	0.63
$[0.125, 0.5]$	1.2283 ± 0.0374	-0.0131 ± 0.0034	0.71

[5] From now on GOF.

In the OS case the GOF is intermediate between the two reweighting methods and it's even better than the result obtained when the hybrid GWR is used. Reweighting leads to a better agreement with 1 of the measured slope value. An improvement on the intercept is obtained only with the hybrid GWR.

In the SS case the GOF is again intermediate between the two reweighting methods for each of the two fit intervals: $[0.25, 0.5]$ and $[0.125, 0.5]$. In the $[0.25, 0.5]$ fit range the slope is greater than the ones obtained both with the typical and hybrid GWR. In the $[0.125, 0.5]$ fit range, instead, the slope is intermediate between the hybrid and the typical GWR. The intercepts of the unweighted case are less compatible with 0 than the intercepts obtained before. Also for the slopes a smaller improvement is observed after reweighting.

Chapter 4

Conclusions

In the analysis of the variables used in input to the $D_s^{*-}\pi^+$BDT, it was found that two of the 14 input variables, the $PtrRel$ and the Ph_isNotE, are quite insensitive to the MC reweighting attempts. Both approaches (typical and hybrid GWR) fail to increase the agreement between these two variables' MC and their corresponding data distributions. For this reason, they will be excluded in the next training iteration of the BDT.
It is worth noticing that the usage of the GBR algorithm, resulting in the hybrid GWR, outperformed the typical approach to reweighting in the majority of the considered distributions. It might be interesting to study the case in which a pure GBR weighting rule is used to further explore the advantages or disadvantages between these to approaches.

In the flavour tagging study, in the OS tagger case the slopes with and without reweighting are close to the expected values. The same behaviour is found also for the intercept.
In the SS case, the slope is close to the expected value of 1 and the unweighted intercept is less compatible with the expected value of 0 than those obtained when the MC is reweighted. Moreover, because of very low statistics at small values of the mistag probability η, two fitting attempts were conducted by varying the fit range. Results show that fitting in the η range $[0.25, 0.5]$ gives values closer to the expected ones than fitting in the $[0.125, 0.5]$ range. Fitting with a hybrid GWR reweighted MC seems to produce better results.

The GBR algorithm seems to be an interesting alternative to the typical reweighting method. It is worth noticing that the algorithm does not seem to be protected towards single "pathological" events. These make the algorithm to predict very large weights that, in turn, will induce large fluctuations in the reweighted samples.

List of Figures

List of Tables

Glossary

BDT Boosted Decision Tree.. 1, 5, 7–10, 12, 19, 21, 23

GBR Gradient Boosted Reweighter.. 4, 7, 8, 12, 16, 19, 23

GOF Goodness of Fit.. 16, 17, 23

GWR Global Weighting Rule for Monte Carlo reweighting, given by the product between the $nTracks$ weighting rule and the $2 - dim$ one.. 7, 8, 14–17, 19, 21, 23

MC Monte Carlo.. 1, 3, 4, 7, 8, 11, 12, 14–16, 19, 21, 23

OS Opposite Side.. 1, 12–17, 19, 23

RT Right Tagged.. 12–14, 16, 23

SS Same Side.. 1, 12–17, 19, 23

WR Weighting Rule. 3, 4, 7, 8, 23

WT Wrong Tagged.. 12–14, 16, 23

References

[1] Alex Rogozhnikov. *Reweighting with Boosted Decision Trees.* Oct. 2015. URL: http://arogozhnikov.github.io/2015/10/09/gradient-boosted-reweighter.html.

[2] A. A. Alves Jr. et al. “The LHCb detector at the LHC”. In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[3] R. Aaij et al. “LHCb detector performance”. In: *Int. J. Mod. Phys.* A30 (2015), p. 1530022. DOI: 10.1142/S0217751X15300227. arXiv: 1412.6352 [hep-ex].

[4] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[5] Leo Breiman et al. *Classification and Regression Trees.* Belmont, California, U.S.A.: Wadsworth Publishing Company, 1984.

[6] Yoav Freund and Robert E Schapire. “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting”. In: *Journal of Computer and System Sciences* 55.1 (1997), pp. 119–139. ISSN: 0022-0000. DOI: http://dx.doi.org/10.1006/jcss.1997.1504. URL: http://www.sciencedirect.com/science/article/pii/S002200009791504X.

[7] Alex Rogozhnikov. *Reweighting Algorithms.* 2015. URL: https://arogozhnikov.github.io/hep_ml/reweight.html.

[8] R. Aaij et al. “First observation and measurement of the branching fraction for the decay B s 0 → D s ∗ ± K ±”. In: *Journal of High Energy Physics* 2015.6 (2015), pp. 1–16. ISSN: 1029-8479. DOI: 10.1007/JHEP06(2015)130. URL: http://dx.doi.org/10.1007/JHEP06(2015)130.

[9] LHCb collaboration et al. *A new algorithm for identifying the flavour of B_s^0 mesons at LHCb.* 2016. eprint: arXiv:1602.07252.