Summer student report

Study of $B^+ \to \mu^+\nu_\mu$ and $K_s^0 \to \mu^+\mu^-$ decays

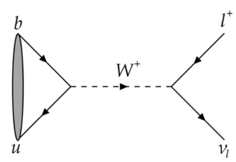
Author: Vitalii Lisovskyi

Supervisors: Xabier Cid Vidal, Jessica Prisciandaro

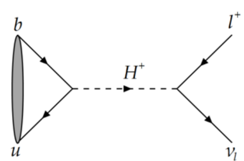
June 15 - August 7, 2015

$B^+ \to \mu^+\nu_\mu$ decay

In the Standard Model (SM) $B^+ \to \mu^+\nu_\mu$ decay is described by the following Feynman diagram:



This decay is suppressed by helicity conservation and the SM predicts its branching fraction to be about $5.6 * 10^{-7}$ [1]. In case of existence of unknown charged particles which can be mediators in this decay the branching fraction could be higher, so it is a golden mode for search for new physics. For example, instead of W boson we can have a charged Higgs boson as a mediator:



The current best limit of search of this decay was established by BaBar collaboration [1] and was about 2 times the SM prediction: $BR(B^+ \to \mu^+\nu_\mu) < 1 * 10^{-6} \quad (90\%CL)$. The analysis of this decay mode is challenging due to neutrino in the final state. It cannot be detected so the usual invariant mass reconstruction is not possible. In this case it was necessary to find another variable to fit. The best solution was to make a multivariate analysis and to fit the BDT variable.

Backgrounds

The first step of any analysis is the determination of possible background modes.
For the $B^+ \to \mu^+\nu_\mu$ decay, the dominant background modes are expected to be $b\bar{b}$ and $c\bar{c}$ production with further semileptonic decays including muons in the final state. Due to the small branching fraction of the studied decay, the expected number of signal events is smaller than the background level, so background must dominate in the data. So, our first purpose was to ensure that the combination of these two types of background describes the data well.
We used $78.6pb^{-1}$ of 2012 LHCb data after stripping 21. Stripping cuts included $p_t(\mu) > 5GeV/c$, $IP\chi^2(\mu) > 400$ and also selected events with low multiplicity (less than 150 tracks). To describe the signal and background mode we used appropriate Monte-Carlo simulations. We used MC stripped filtered production, and the number of events for signal, $b\bar{b}$ and $c\bar{c}$ was respectively 54818, 43563 and 6410.

We use variables of different types: global physics variables, including Particle Flow variables (contain all particles in the event including neutral), cylinder variables (build a cylinder around the muon and look at tracks inside in order to select an isolated muon), cone variables (build a cone around muon for the same purpose), and jet variables (build jets). We had to check if the combination of background modes explains the data well for all the variables. To build such plots MC distributions must be multiplied by expected number of each background type to match the number events in data, and be reweighted to match the multiplicity in data. To obtain the expected number of each background type we have made a fit of the muon transverse momentum distribution in data using MC for background modes.

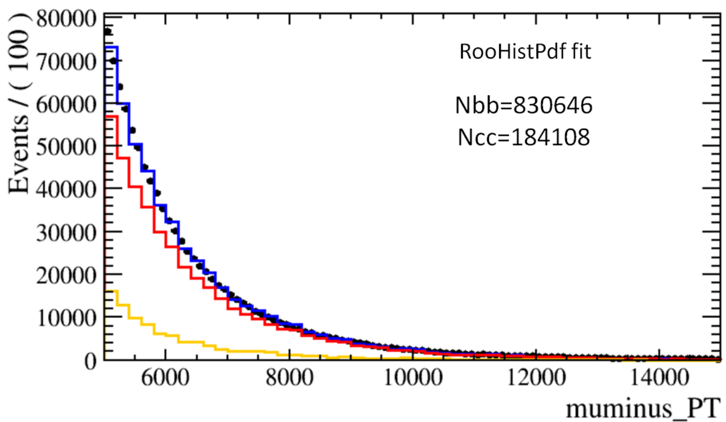


Figure 1: Fit on $p_T(\mu)$ to determine the ratio of $b\bar{b}$ and $c\bar{c}$ in data. $b\bar{b}$ and $c\bar{c}$ modes are described by red and yellow histograms respectively, and the blue one is sum of them

Obtained numbers of each background mode events were used to obtain further plots. For example, on Fig.2 plots are shown for some of these variables. Black points represent the data, and red and blue areas represent $b\bar{b}$ and $c\bar{c}$ background modes respectively.

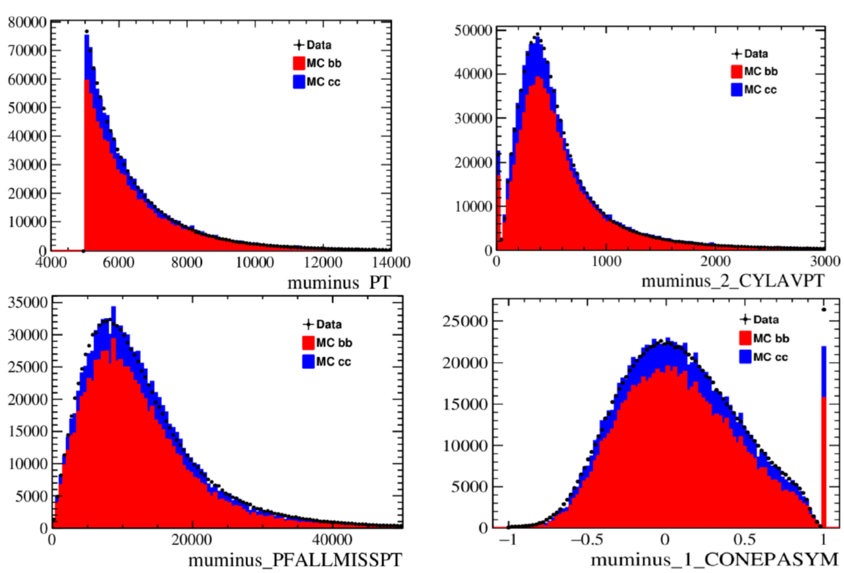


Figure 2: Comparison of data and background MC modes for $B^+ \to \mu^+\nu_\mu$ decay. Top left, muon p_T, top right, cylinder average p_T, bottom left, Particle Flow missed p_T, bottom right, cone asymmetry

For most of the variables sum of background modes looked similar to the data. The only exception were some of jet variables for which data and background looked completely different - this problem is in method of construction of these variables and is being studied now. Since all other variables look good, we can conclude that our understanding of the dominant background modes is correct. So we can try to search for the signal now. For this purpose we must find variables having different distributions for signal and background modes. Examples of such variables are shown on Fig.3:

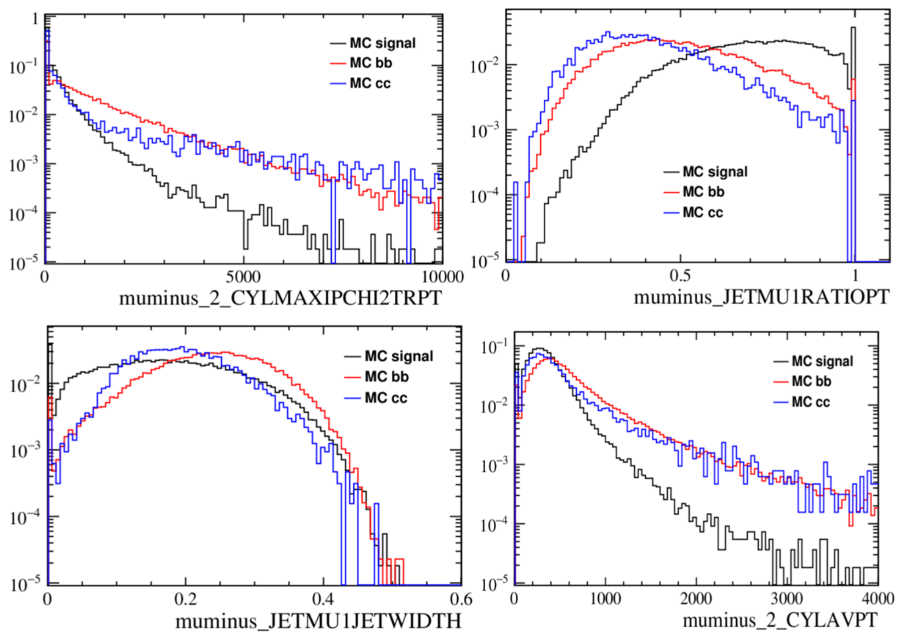


Figure 3: Distributions of some variables for signal and background modes. Top left, cylinder maximum impact parameter χ^2, top right, jet p_T, bottom left, jet width, bottom right, cylinder average p_T

Then we chose those variables that look most different to use for the multivariate analysis. One can notice that distributions of $b\bar{b}$ and $c\bar{c}$ background modes are different for some variables. So it was decided to train two different BDTs - BDT_{bb} and BDT_{cc}. As the signal sample we used signal MC in both cases, and as background sample - MC of the corresponding background mode. Signal and background MCs were split into training and testing samples. We also optimized the list of discriminating variables to choose those having a good agreement between data and background: this was necessary to be sure that fit on BDT will converge.

We trained several methods and chose the BDTG method as the one giving the best separation of signal and background, and the most convenient for the fit. Distributions of BDT_{bb} and BDT_{cc} variables for signal and background modes are shown on Fig.4.

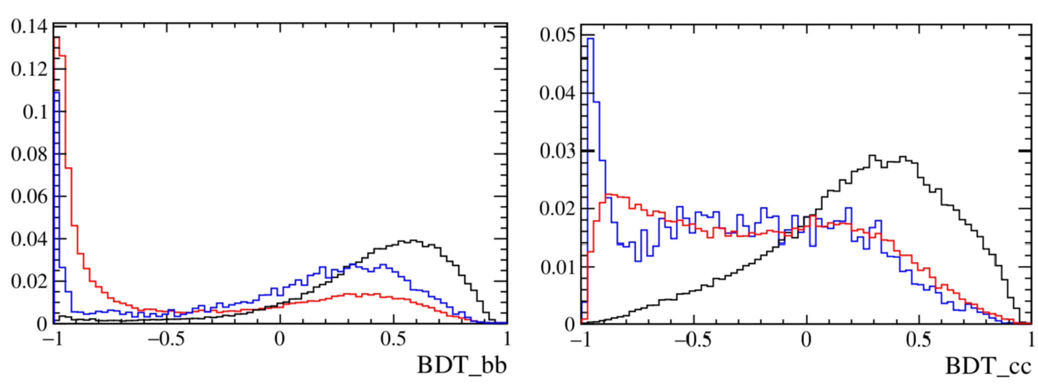


Figure 4: Distributions of BDT variables for signal and background modes. Black curves represent signal, red and blue - $b\bar{b}$ and $c\bar{c}$ background modes respectively

One can see that background distributions are peaking to -1, while signal is located more in the positive region of both BDT variables. This can also be represented by a 2-dimensional plot on Fig.5 where the signal is located in upper-right part of the plot.

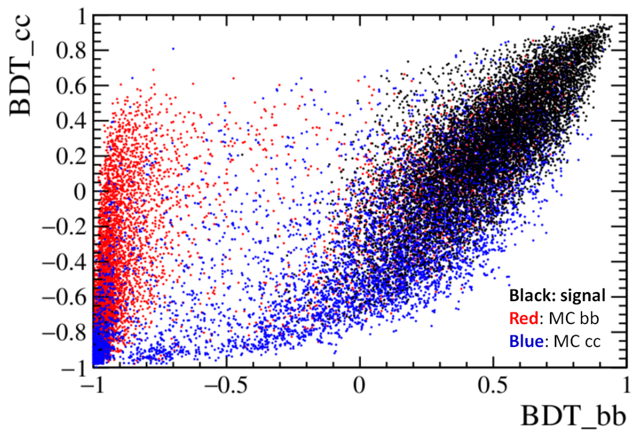


Figure 5: Distributions of signal and background modes in the frame of BDT variables

Then we applied these BDTs on the data and tried to make a fit. It was decided to start from the one-dimensional fit on each of two BDT variables. We used RooHistPdf method to describe shapes of signal and background BDT distributions. The fit was done in the whole BDT range from -1 to 1. The yields of each component were used as free fit parameters. Data points with fit curve are shown at the Fig.6 together with curves of signal and background components and fit parameters values.

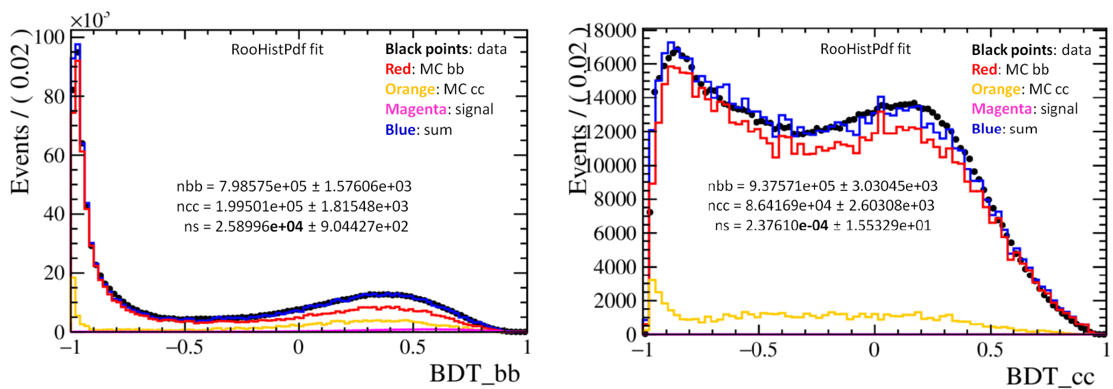


Figure 6: Fit of data on each of BDT variables

It should be mentioned that the signal yield in the first case is large (around 10^4 events) and in the second it is around zero (10^{-4} events), and the uncertainties of all fit parameters values are large. We can explain this difference by the fact that number of $c\bar{c}$ events in MC sample used for training was low so BDT_{cc} distributions have larger fluctuations. Moreover, the fit was unstable: moving boundaries of the fit region caused significant change of values of fit parameters. We also tried to make a two-dimensional fit on both BDT variables simultaneously, but it did not converge.

$K_s^0 \to \mu^+\mu^-$ decay

My contribution in this analysis consisted of two parts: calculation of stripping cuts efficiencies and obtaining the best fit of the dominant background mode.

Stripping cuts efficiencies

Stripping of the data included several cuts on different variables. I studied the efficiency of each cut for signal ($K_s^0 \to \mu^+\mu^-$) and normalization channel ($K_s^0 \to \pi^+\pi^-$). Results are shown in the following table. The first two pairs of columns represent efficiency of each cut if applied separately with its errors, and the last two pairs of columns show the efficiency of each cut in case when all the previous in the list cuts were already applied, also with errors. The little difference between the results of these two approaches is caused by the fact that in the second case some of the events were already rejected by previous cuts. In the last line the whole stripping efficiency is shown.

	$\pi\pi$ **separately**		$\mu\mu$ **separately**		$\pi\pi$ **all together**		$\mu\mu$ **all together**	
cuts	eff, %	err, %	eff, %	err, %	eff, %	err, %	eff, %	err, %
$\mu^\pm$_IP$\chi^2 > 100$	**76.65**	0.06	**77.90**	0.09	**76.65**	0.06	**77.90**	0.09
$AMAXDOCA(K_s^0) < 0.3$	**87.51**	0.05	**97.91**	0.03	**90.83**	0.05	**99.74**	0.0130
$DIRA(K_s^0) > 0$	**98.49**	0.02	**97.91**	0.03	**99.80**	0.01	**99.74**	0.01
$FD(K_s^0) * m(K_s^0)/p(K_s^0) > 8.953 * 0.29979$	**72.60**	0.07	**71.86**	0.10	**86.60**	0.06	**86.60**	0.09
$IP(K_s^0) < 0.4$	**88.88**	0.05	**91.07**	0.06	**90.89**	0.06	**91.60**	0.08
$p_T(\mu^\pm) > 250$	**34.17**	0.07	**31.65**	0.10	**35.86**	0.10	**33.77**	0.13
$400 < m(K_s^0) < 600$ or $m(K_s^0) > 465$	**99.18**	0.01	**99.86**	0.01	**99.99**	0.00	**99.99**	0.01
all stripping cuts	**-**	-	**-**	-	**19.61**	0.06	**18.99**	0.08

One can see that stripping efficiency and efficiency of each cut is similar for signal and normalization mode. Also the cut on p_T is the least efficient, and it is worth trying to improve it.

Fit of the dominant background mode

In the LHCb detector some pions can be misidentified as muons. This leads to a huge background from doubly misidentified ($K_s^0 \to \pi^+\pi^-$) with the mass peak shifted to the left. The tail of this peak is covering the signal window, so to estimate the amount of signal it was necessary to obtain the most accurate possible fit of this background mode in the signal window. For this purpose we used MC of ($K_s^0 \to \pi^+\pi^-$) with 0, 1 or both pions misidentified as muons, and tried different shapes to describe it. To obtain the best fit we compared the number of events in the signal window obtained from the fit with real number of events there. The best shape we found consisted of two parts: Bifurcated Crystal Ball function to describe the peak itself and Chebyshev polynomial of third order to describe combinatorial background (Fig.7):

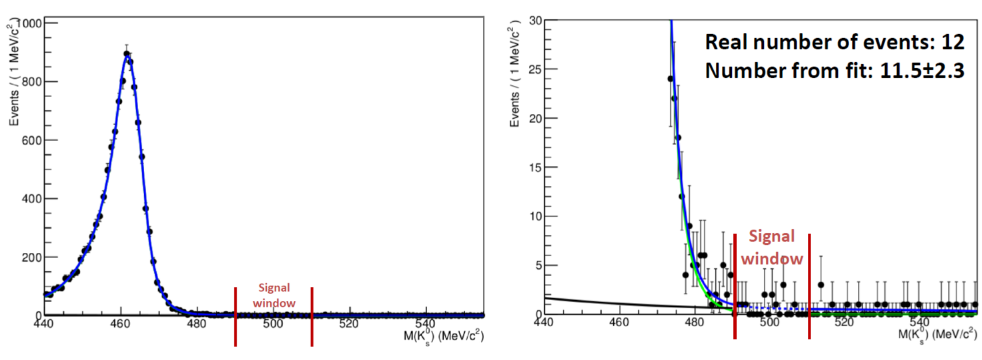


Figure 7: Fit of misID background mode. Left plot shows the full mass range while the right one zooms into signal window. Black points are data, green line is Bifurcated Crystal Ball function, black one is the Chebyshev polynomial of 3rd order and blue curve is sum of them.

We also checked applicability of this fit on the real data and looked on its behavior after application of different cuts. Generally, it describes the misID background mode well.

Conclusions

During my stay at CERN I worked on two topics within the LHCb collaboration: $B^+ \to \mu^+\nu_\mu$ and $K_s^0 \to \mu^+\mu^-$ analyses. For the first one we understood the dominant background modes, trained and fitted the BDT. But to continue analysis we need much machine and man power. For the second analysis I calculated the stripping cuts efficiencies and identified the least efficient cuts. Also I studied the dominant background mode and developed the best fit of it in the signal window.

References

[1] Bernard Aubert et al. Search for the Rare Leptonic Decays $B^+ \to l{+}\nu_l$ $(l = e, \mu)$. *Phys. Rev.*, D79:091101, 2009.