Analysis of the charmless decay $B^0 \to \rho\pi$ in the LHCb experiment

Diego Alejandro Roa Romero
O. Deschamps, R. Lefèvre, P. Perret

Laboratoire de Physique Corpusculaire - Université Blaise Pascal

December 2011

Contents

Quarks in the Standard Model

- In the Standard Model, we find six quarks coming in three generations:

u c t

d s b

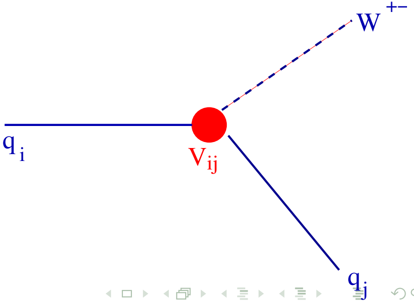
- These are the mass eigenstates composing the hadrons (valence quarks), except for the top quark which weakly decays before hadronizing
- As the weak interaction eigenstates are different from the mass eigenstates, the W bosons couple quarks of different generations

CKM matrix

- The transformation from the mass eigenstates basis (q) to the weak interaction one (q') can be represented by a 3×3 unitary matrix, the Cabbibo-Kobayashi-Maskawa matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}$$

The weak coupling between two mass eigenstates (ij) then depends on the matrix element V_{ij} of the CKM matrix.



Unitary Triangle

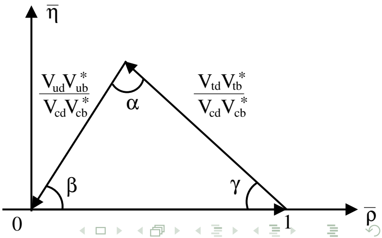
- Unitarity implies that the matrix elements satisfy

$$\sum_j V_{ij} V^*_{jk} = \delta_{ik} \quad \forall\ i,k = 1,2,3$$

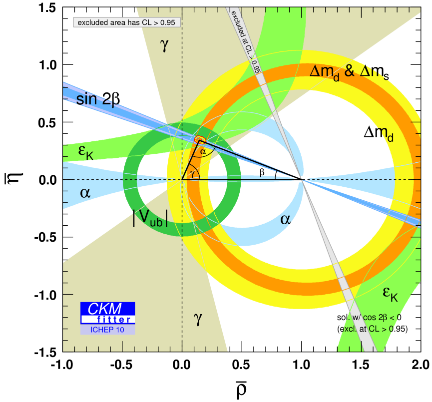
- We are particularly interested in one of those relations

$$V_{cd} V^*_{cb} + V_{td} V^*_{tb} + V_{ud} V^*_{ub} = 0$$

The representation of this relation in the complex plane is a triangle. The angle α is related with CP violation in B mesons: no CP violation would mean a flat triangle, i.e. $\alpha = \pi$.



Constraints on the unitary triangle



Concerning α

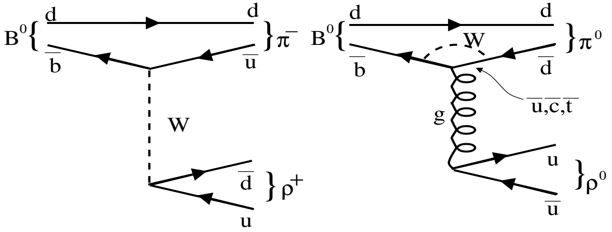
- The combination of the measurements gives: $\alpha = (89.0 \pm^{4.4}_{4.2})^\circ$
- The global fit, excluding α measurements, gives: $\alpha = (92.9 \pm^{3.6}_{5.1})^\circ$

$b \longrightarrow u$ transitions

- To measure α we have to use decays involving $b \rightarrow u$ transitions:

$$\alpha = arg\left(\frac{-V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right)$$

- In these processes the main contributions come from diagrams at tree level of weak origin and penguin diagrams involving QCD and weak factors:



Decay channels

- The channels involving $b \to u$ transitions are $B^0 \to \pi\pi$, $B^0 \to \rho\pi$ and $B^0 \to \rho\rho$ (branching ratios between $7 \cdot 10^{-7}$ and $2.4 \cdot 10^{-5}$)
- We focus on $B^0 \to \rho\pi \to \pi^+\pi^-\pi^0$ which should lead to the best experimental sensitivity
- We will see in the following slides:
  - how the decay amplitudes can be written in terms of α,
  - how the phase space can be expressed,
  - how the time evolution of a $|B^0>$ state can be written according to $B^0 - \bar{B}^0$ mixing,
  - and finally how to extract α.

Decay amplitude

- We can express the total amplitude of $B^0 \longrightarrow \pi^+\pi^-\pi^0$ as the combination of the amplitudes of $B^0 \longrightarrow \pi^{+-0}\rho^{-+0}$
- Factorizing the penguin and tree parts, the amplitudes can be expressed as

$$\mathbf{A}^j = V_{ub}^* V_{ud} \mathbf{T}^j - V_{tb}^* V_{td} \mathbf{P}^{\mathbf{j}}$$

where "j" represents the decay to $\rho^+\pi^-$, $\rho^-\pi^+$ or $\rho^0\pi^0$

- In terms of α

$$e^{i\beta}\mathbf{A}^j = e^{-i\alpha}\mathbf{T}^j - \mathbf{P}^j$$

- Isospin decomposition leads to

$$-\frac{1}{2}\left(\mathbf{P}^{+-} + \mathbf{P}^{-+}\right) = \mathbf{P}^{00}$$

$B^0 \longrightarrow 3\pi$ amplitude

- The total amplitude $\mathbf{A}_{3\pi}$ of the $B^0 \longrightarrow 3\pi$ decay is

$$\mathbf{A}_{3\pi} = \sum_i f^j \mathbf{A}^j$$

- The factors f^j account both for pure form factors and the angular distributions associated to the spin of the ρ vector meson
- This point is very important and determines the way α will be extracted

Parametrization of the phase space

- Initially, there are 12 degrees of freedom corresponding to the 4-momentums of the 3 pions
- 4-momentum conservation between the initial B meson and the decay products imposes 4 relations
- The nature of the decay products being known, their invariant masses give 3 more relations
- As the B meson is a scalar, the orientation of the decay plane is isotropic and any choice of the 3 Euler angles is equivalent
- The phase space can then be represented by only 2 parameters:

$$s^+ = m^2_{\pi^+\pi^0},\ s^- = m^2_{\pi^-\pi^0}$$

and the factors f^j can be expressed as functions of s^+ and s^-

$B^0 - \bar{B}^0$ mixing

- We can describe the B^0 system by the flavour eigenstates $|B^0>=|\bar{b}d>$ and $|\bar{B}^0>=|b\bar{d}>$ that can be written as linear combinations of the mass eigenstates:

$$\begin{aligned} |B_L> &= p|B^0> + q|\bar{B}^0> \\ |B_H> &= p|B^0> - q|\bar{B}^0> \end{aligned}$$

  with $|p|^2 + |q|^2 = 1$

- The time evolution of a $|B^0>$ state, prepared as such at $t=0$, is given by (the formula for $\bar{B}^0(t)$ is similar)

$$|B^0(t)>= e^{-imt}e^{-\frac{\Gamma t}{2}} \times \left[cos\left(\frac{\Delta mt}{2}\right)|B^0> + i\frac{p}{q}sin\left(\frac{\Delta mt}{2}\right)|\bar{B}^0> \right]$$

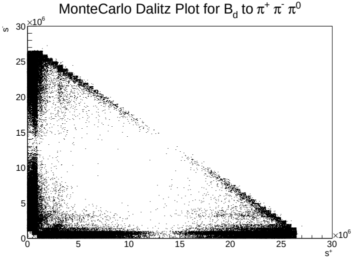
  with: $m = (M_H + M_L)/2$, $\Delta m = M_H - M_L$ and $\Gamma = (\Gamma_H + \Gamma_L)/2$
  assuming: $\Delta\Gamma = \Delta\Gamma_H - \Gamma_L << \Gamma$ and Δm

Amplitude distribution

The decay amplitude distribution as a function of phase space and proper time can be expressed as (here for an initial B^0)

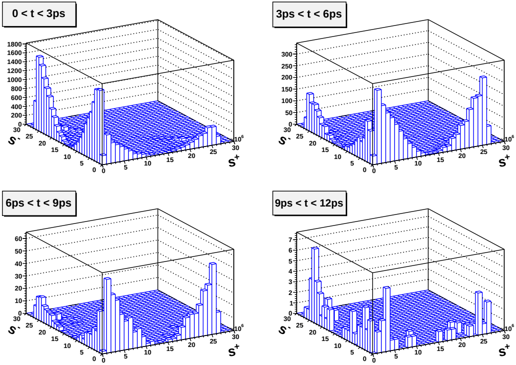
$$\begin{aligned} \mathbf{M}(t, s^+, s^-) &= e^{-\Gamma t/2} cos(\frac{\Delta m t}{2}) \mathbf{A}_{3\pi}(s^+, s^-) \\ &+ i e^{-\Gamma t/2} \frac{q}{p} sin(\frac{\Delta m t}{2}) \bar{\mathbf{A}}_{3\pi}(s^+, s^-) \end{aligned}$$

- The distribution as a function of $(s^+; s^-)$ is called a Dalitz plot
- The strategy to extract α is to fit the time dependent Dalitz plot obtained on flavour tagged (initial B^0 or initial $\bar{B}^0$) decays

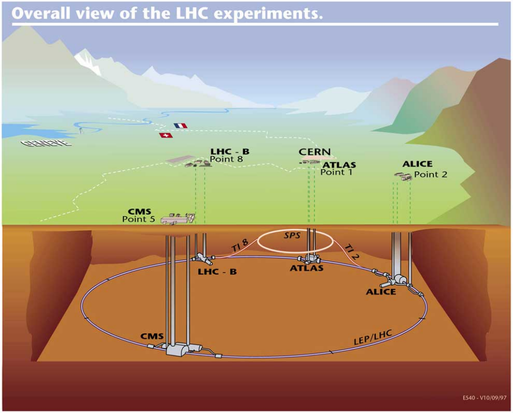


Example of time dependent Dalitz plots

Example of Dalitz plots for an initial B^0 and various ranges of proper time:



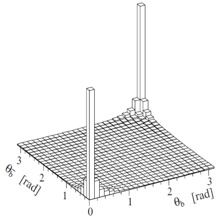
The Large Hadron Collider



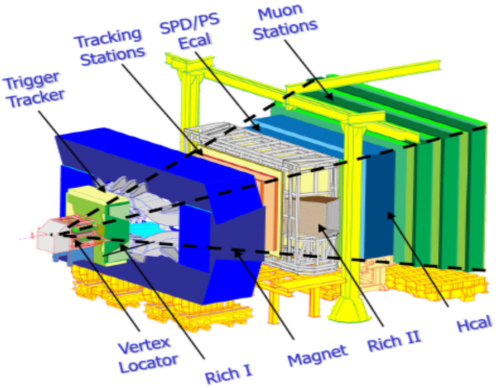
- *pp* collisions
- $\sqrt{s} = 7\ \mathrm{TeV}$
- 1.1 fb^{-1} recorded in LHCb in 2011

The LHCb Experiment

LHCb is a single arm spectrometer covering the region between $1.9 < \eta < 4.9$



$b\bar{b}$ mostly produced close to the beam pipe



Trigger

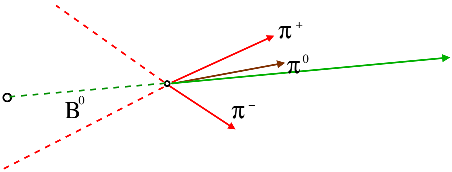
A very efficient trigger is required: even if the $b\bar{b}$ cross section is high at the LHC ($\sigma_{b\bar{b}} \sim 300\ \mu$b at $\sqrt{s} = 7$ TeV), the rate of background events is much higher ($\sigma_{inel} \sim 60$ mb); in addition, the branching ratios of channels of interest are small ($Br(B^0 \to 3\pi) = 2.4 \times 10^{-5}$).

- L0 uses custom electronics: fully synchronous (40 MHz), 4 μs fixed latency
  - High p_T candidates from calorimeters (hadron, e, γ) and from muon system (μ, di-μ); veto high occupancy events (Global Events Cuts)
- High Level Trigger (HLT) uses a farm of about 2000 CPUs
  - HLT1 $\to$ fast tracking
  - HLT2 $\to$ full event reconstruction

Key elements

The following are key elements for the extraction of α fitting the time dependent Dalitz plot of flavour tagged $B^0 \longrightarrow \rho\pi$ decays:

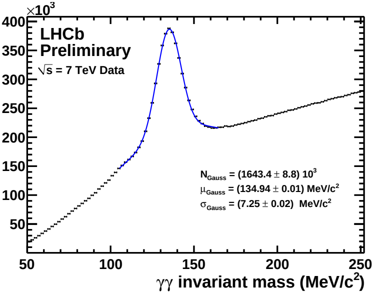
- π^0 reconstruction
- Kaon identification
- Propertime measurement
- Flavour tagging



π^0 reconstruction

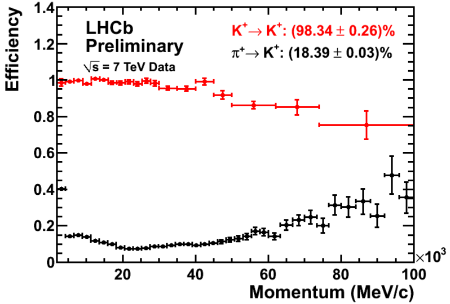
- π^0 mostly decays in two photons (99% of the cases)
- In LHCb, photons are reconstructed as calorimeter clusters made of 3×3 calorimeter cells
- π^0 can be merged or resolved, i.e. whether or not the clusters of the 2 photons overlap

$\gamma\gamma$ invariant mass for resolved π^0 (first 3 nb^{-1}) $\rightarrow \sigma = 7.25\ \mathrm{MeV/c^2}$



Kaon identification

- Kaon identification is essential to distinguish similar decays such as $B^0 \to K^-\pi^+\pi^0$ and $B^0 \to \pi^+\pi^-\pi^0$
- This identification is mainly made by RICH detectors
- Calibration samples:
  - K from $\phi \to K^+K^-$
  - π from $K_s \to \pi\pi$
- Plots for $dLL(K-\pi) > 0$

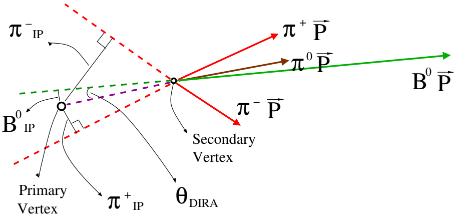


$B^0 \longrightarrow \rho\pi$ topology

We are looking for:

- Well reconstructed tracks: $\pi^\pm$ with low track $\chi^2/ndof$
- Tracks not coming from the primary vertex: $\pi^\pm$ with large IP significance
- Tracks coming from the B^0 decay vertex: end vertex with low χ^2
- B^0 coming from the primary vertex: low IP significance, low θ_{DIRA}

- Decay products from a B meson: relatively high p_T because of the high B mass
- Decay from a B meson: high flight distance, 3-body invariant mass in the B^0 mass range

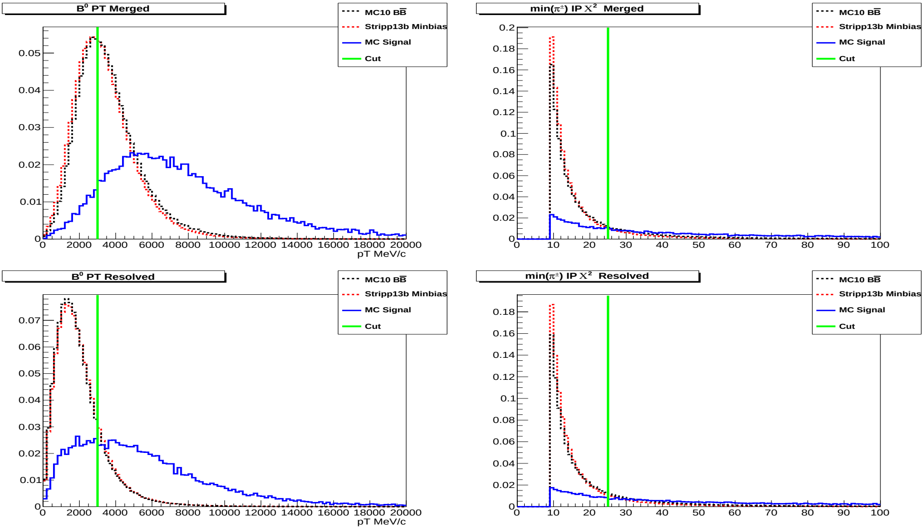


Event selection

- The total sample of LHCb is so BIG. It is divided in groups depending on each set of channels (stripping)
- The cut based selection is driven by the stripping selection we designed to select $B^0 \to hh\pi^0$ final states
- This stripping selection grants access to B_d^0 and B_s^0 decays to $\pi\pi\pi^0$, $K\pi\pi^0$ and $KK\pi^0$
- no Kaon identification cut applied to the tracks, large B mass window

We will now discuss briefly the stripping and trigger selections

Illustration of stripping cuts



Stripping selection for $B \rightarrow hh\pi^0$

$\pi^{\pm}$ cuts

$p_T > 500\ \mathrm{MeV/c}$	$p > 5000\ \mathrm{MeV/c}$
Track χ^2 probability $> 10^{-6}$	IP χ^2 >25

π^0 cuts

p_T >1500 $\mathrm{MeV/c}$ (Resolved), 2500 $\mathrm{MeV/c}$ (Merged)
$CL(\gamma^1)$ and $CL(\gamma^2) > 0.2$ (Resolved π^0 only)

B^0 cuts

$p_T > 2500\ \mathrm{MeV/c}$ (Resolved), 3000 $\mathrm{MeV/c}$ (Resolved)	
End vertex χ^2 probability $> 10^{-3}$	IP $\chi^2 < 9$
$\theta_{DIRA} < 10$ mrad	Flight distance $\chi^2 > 64$
$4200 < m_{B^0} < 6400\ \mathrm{MeV/c^2}$	

Trigger selection: L0 and HLT1

- L0 and HLT1 selections based on standard trigger lines
- L0: hadron, γ and electron lines are the most relevant ones
- HLT1
  - Hlt1Track: single detached high momentum track ($IP\chi^2$ cut ~ 36; p_T cut $\sim 1.5\ \mathrm{GeV/c}$)
  - Hlt1Track + Photon: looser momentum cuts on the single detached high momentum track in the case of a L0 photon trigger (p_T cut $\sim 0.8\ \mathrm{GeV/c}$)
- To reduce the background, with a very limited loss on signal efficiency, we require that the $\pi^+\pi^-\pi^0$ combination selected offline is enough to fire the HLT1 trigger

Trigger selection: HLT2

- HLT2 selection relies on both a standard trigger line (Hlt2Topo2Body) and a dedicated line we designed to improve the trigger efficiency (Hlt2B2HHPi0)
- The purpose of the Hlt2Topo2Body line is to trigger on 3-body decays for which only two tracks have been reconstructed in the HLT2 ($3^{\rm rd}$ particle = neutral or low momentum track)
- The Hlt2B2HHPi0 line implements in the HLT2 similar cuts to the ones we use for the $B^0 \to hh\pi^0$ stripping selection
- To reduce the background, with a very small cost on signal efficiency, we require that the $\pi^+\pi^-\pi^0$ combination selected offline is enough to fire at least one of these 2 HLT2 lines

HLT2 line dedicated to $B \to hh\pi^0$

$\pi^\pm$ cuts

$p_T > 500\ \mathrm{MeV/c}$	$p > 5000\ \mathrm{MeV/c}$
Track $\chi^2/\mathrm{ndof} < 2.4$	IP $\chi^2 > 9$
Distance of closest approach of the 2 tracks < 0.2 mm	

π^0 cut: $p_T > 1500\ \mathrm{MeV/c}$ (Resolved), $2500\ \mathrm{MeV/c}$ (Merged)

B^0 cuts

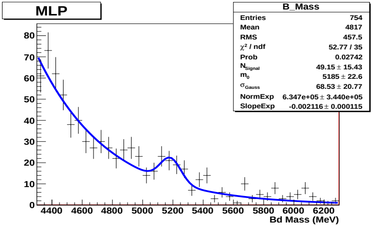
$p_T > 2500\ \mathrm{MeV/c}$ (Resolved), $3000\ \mathrm{MeV/c}$ (Merged)	
End vertex $\chi^2 < 10$	IP $\chi^2 < 25$
$\theta_{DIRA} < 16$ mrad	Flight distance $\chi^2 > 100$
$4200 < m_{B^0} < 6400\ \mathrm{MeV/c^2}$	

Additional cuts for $B^0 \to \rho\pi$ study

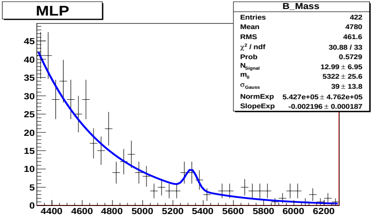
- The two tracks are associated to pions: $dLL(K-\pi) < 0$ for both tracks
- The B^0 decay go through the intermediate ρ resonance: $400 < m_{\pi\pi}^{min} < 1200\ \mathrm{MeV/c^2}$ with $m_{\pi\pi}^{min}$ the minimum invariant mass among $m_{\pi^+\pi^0}$, $m_{\pi^-\pi^0}$ and $m_{\pi^+\pi^-}$
- $4300 < m_{B^0} < 6300\ \mathrm{MeV/c^2}$

Results for 2010 data ($\sim 35pb^{-1}$)

- In order to analyze the 2010 data, the stripped data was further purify using a Multivariate Analysis: Fisher, Neural Network and Boosted Decision Tree methods were tried
- Those expected significance should increase by at least a factor 5 over the 2011 data sample ($1.1fb^{-1}$) The $B^0 \longrightarrow \pi^+\pi^-\pi^0$ should clearly be observable



B^0 Mass using Merged π^0



B^0 Mass using Resolved π^0

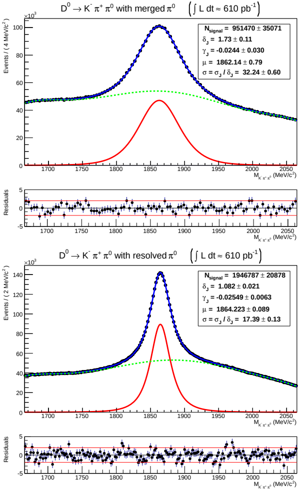
π^0 and γ Confidance Level

- The confidance level is defined for photons and π^0. It is a tool to distinguish good neutral particles from background.

- It uses information from the SPD, Preshower and ECAL clusters and the possible matching between those clusters and tracks.

$D^0 \to K^-\pi^+\pi^0$ control sample

- $Br(D^0 \to K^-\pi^+\pi^0) \sim 14\%$
- Similar stripping selection
- Good resolutions in D^0 mass
  - Resolved π^0: about 14 $\mathrm{MeV/c^2}$
  - Merged π^0: about 30 $\mathrm{MeV/c^2}$
- This sample is being used to study photon and pi0 identification with very high statistics



Conclusions

- The $B^0 \to \rho\pi$ decay should allow to precisely measure the angle α of the unitary triangle in LHCb
- The extraction of α will be done through a Dalitz time dependent analysis of flavoured tagged decays
- Trigger and stripping selections have been implemented
- We benefit from a nice $D^0 \to K^-\pi^+\pi^0$ control sample
- We have 1.1 fb^{-1} of recorded data waiting to be analyzed

THANK YOU

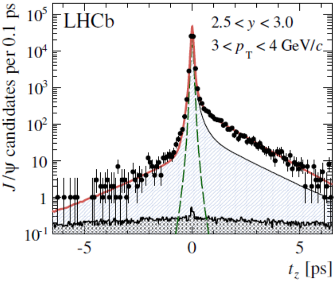
Propertime measurement

- The measurement of the propertime is of major importance for a lot of analyses in LHCb
- It has been used for instance to extract the $b\bar{b}$ cross section using $J/\psi \to \mu\mu$ events

- Pseudo-propertime defined as:

$$t_z = \frac{(z_{J/\psi} - z_{PV}) \times M_{J/\psi}}{p_z^{J/\psi}}$$

- $\sigma_{b\bar{b}} = 288 \pm 4 \pm 48\ \mu$b



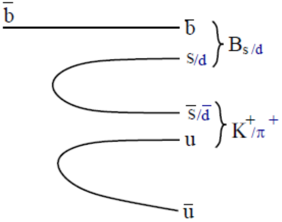
- The propertime resolution is around 40 to 50 fs depending on the final state

Flavour tagging

Flavour tagging is the identification of the initial state ($t = 0$) of the B meson (B or $\bar{B}$)

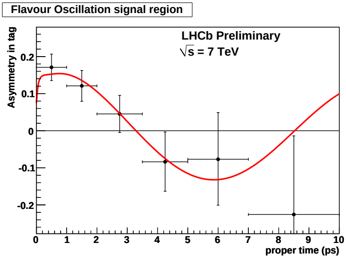
- Opposite side tagging: identifies the flavour of the partner b-hadron ($b\bar{b}$ pair produced at $t = 0$)
  - Lepton tagging: $b \to l^- X$ (warning: $b \to cX \to l^+ X'$)
  - Vertex charge tagging: $B^+ = \bar{b}u$ / $B^- = b\bar{u}$
  - Kaon tagging: $b \to cX \to sX'$ ($K^+ = \bar{s}u$ / $K^- = s\bar{u}$)

- Same side tagging: fragmentation track close to the B meson
  - Kaon in the case of B_s^0: K^+ for B_s^0 / K^- for $\bar{B}_s^0$
  - Pion in the case of $B_{(d)}^0$: π^+ for $B_{(d)}^0$ / π^- for $\bar{B}_{(d)}^0$



B^0 oscillation

- Tagging efficiency $= \epsilon_{tag}$
- Dilution: $D = 1 - 2\omega$ where ω is the wrong tagging probability
- Effective statistics after tagging: $N_{eff} = N_{total} \times \epsilon_{tag} D^2$
- First signal of flavour oscillation observed for $B^0 \to D^{*-}\mu^+\nu_\mu$
  - Data sample of 1.9 pb^{-1}
  - "Out of the box" tagging algorithm: $\epsilon_{tag} D^2 \sim 2\%$ ("already" 60% of expected nominal performance)
  - $\Delta m_d = 3.8 \pm 0.5\ \mathrm{MeV/c^2}$ (PDG: $3.34 \pm 0.03\ \mathrm{MeV/c^2}$)



Computation of expected signal yield

- The number of signal events is given by:

$$S = 2 \times \sigma_{b\bar{b}} \times f(b \to B^0) \times Br(B^0 \to 3\pi) \times \epsilon_{tot} \times \int L dt$$

- ϵ_{tot} accounts for all the efficiencies:

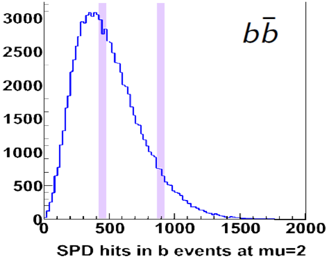
$$\epsilon_{tot} = \epsilon_{gen} \times \epsilon_{sel} \times \epsilon_{GEC} \times \epsilon_{trig}$$

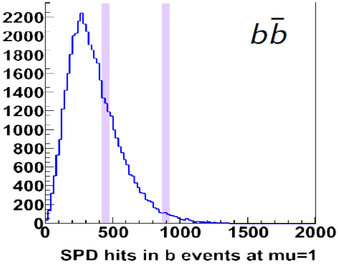
- Some numbers

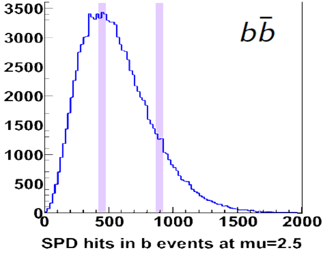
$\sigma_{b\bar{b}} = 292\ \mu$b	$f(b \to B^0) = 0.41$
$Br(B^0 \to 3\pi) = 2.4{\cdot}10^{-5}$	$\int L dt = 33\ pb^{-1}$
$\epsilon_{gen} = 15.8\%$ (acceptance)	$\epsilon_{GEC} = 60\%$

Global Event Cuts (GEC)

- High occupancy events are more difficult to reconstruct and take more time in the HLT
- They are vetoed using the numbers of SPD hits and the clusters in the trackers







Efficiencies

$$\epsilon_{gen} = \frac{\text{number of events generated in the acceptance}}{\text{number of events generated}}$$

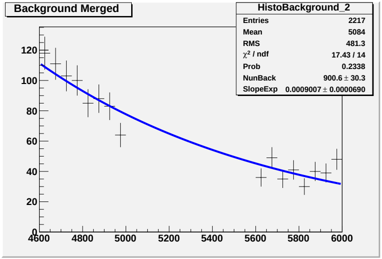
$$\epsilon_{sel} = \frac{\text{number of events selected}}{\text{number of events generated in the acceptance}}$$

$$\epsilon_{trig} = \frac{\text{number of events selected passing the trigger}}{\text{number of events selected}}$$

Efficiency	Merged	Resolved
ϵ_{sel}	6.1×10^{-3}	5.2×10^{-3}
ϵ_{trig}	0.43	0.25

Background and signal expectations

- To estimate the background contribution in the signal region ($5000 < m_{B^0} < 5600\ \mathrm{MeV/c^2}$), the data are fitted by an exponential
- This leads to the following expectations



π^0 type	S	B	S/B	$S/\sqrt{S+B}$
Merged	47	901	0.05	1.51
Resolved	23	1015	0.02	0.73

- To improve those performances we use a multivariate analysis

Multivariate analysis

- Multivariate classifiers combine correlated input variables into a discriminant output
- We use TMVA (Toolkit for MultiVariate Analysis) which provides a ROOT-integrated environment and implements a variety of multivariate clasification algorithms through a common interface

- The results of two classifiers are reported here:
  - Fisher: projection of the data over the hyperplane of best separation
  - Multi-Layer-Perceptron (MLP): artificial neural network interconecting layers of artificial neurons through non-linear functions

Training method

- Signal sample: MC events passing the offline selection as well as the trigger requirements and in the mass window $5000 < m_{B^0} < 5600$
- Background sample: data events passing the offline selection as well as the trigger requirements and in the mass windows $4300 < m_{B^0} < 5000$ or $5600 < m_{B^0} < 6300\ \mathrm{MeV/c^2}$
- To make sure there is no over training, each of those two samples is divided into a training sample (half of the statistics) and a test sample (other half of the statistics)

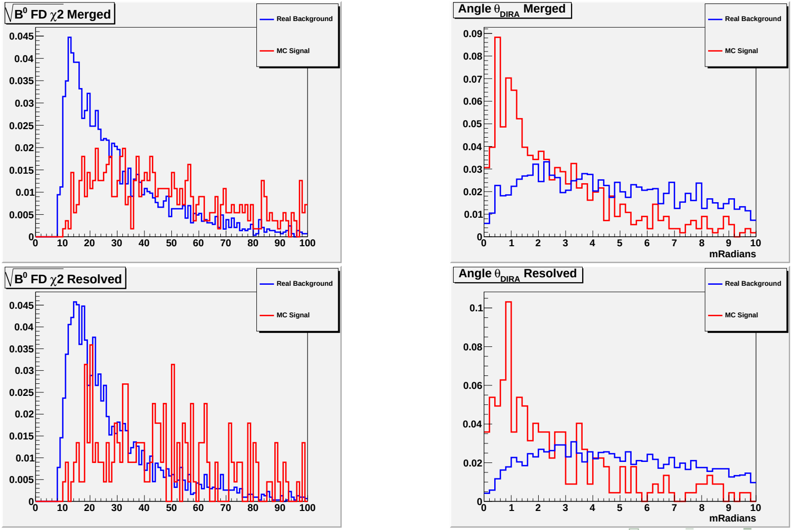
Variables used in TMVA

Small set of variables providing good discrimination between signal and background:

- max $[p_T(\pi^+),\ p_T(\pi^+)]$
- min $[p_T(\pi^+),\ p_T(\pi^+)]$
- $\sqrt{\mathrm{IP}\chi^2}$ of the $\pi^\pm$ with max p_T
- $\sqrt{\mathrm{IP}\chi^2}$ of the $\pi^\pm$ with min p_T
- $p_T(\pi^0)$
- min $[CL(\gamma^1),\ CL(\gamma^2)]$ (resolved π^0 only)

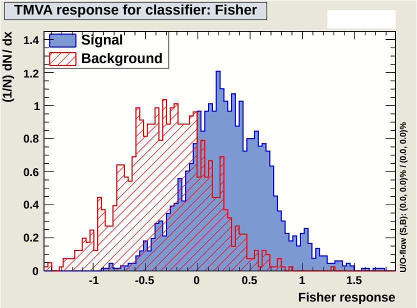
- $-\log_{10}[\mathrm{End\ vertex}\ \chi^2\ \mathrm{prob.}(\mathrm{B}^0)]$
- $\sqrt{\mathrm{IP}\chi^2(\mathrm{B}^0)}$
- θ_{DIRA}
- $\sqrt{\mathrm{Flight\ distance}\ \chi^2(\mathrm{B}^0)}$
- min $[p_{\pi^+} \perp \vec{p}_{B^0},\ p_{\pi^-} \perp \vec{p}_{B^0},\ p_{\pi^0} \perp \vec{p}_{B^0}]$
- $cos[max(\theta_{\pi^+B^0},\ \theta_{\pi^-B^0},\ \theta_{\pi^0B^0})]$ in the B^0 rest frame

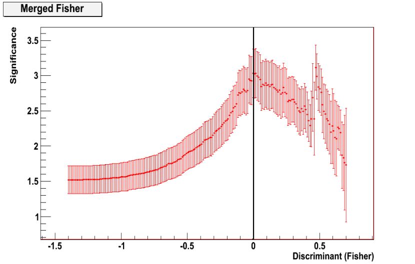
Some distributions



Choise of the cut on the discriminant

The criteria used was to select the cut that gives the best expected significance $(S/\sqrt{S+B})$

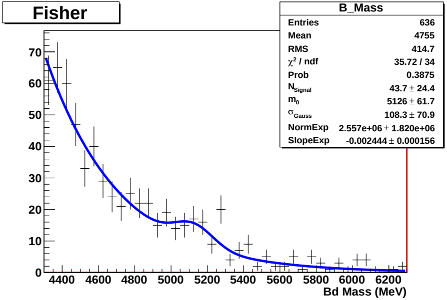




The best expected significances we obtained are

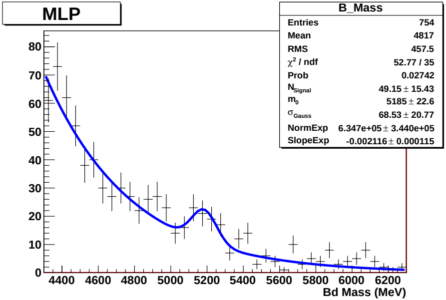
π^0 type	Fisher	MLP
Merged	3.0	2.6
Resolved	1.4	1.3

Merged π^0 results

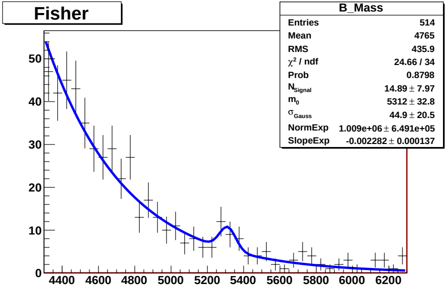


- Fisher results
  - $S_{exp} = 39$
  - $B_{exp} = 108 \pm 11$
  - $S_{fit} = 43.7 \pm 24.4$
  - $(S+B)_{obs} = 120$

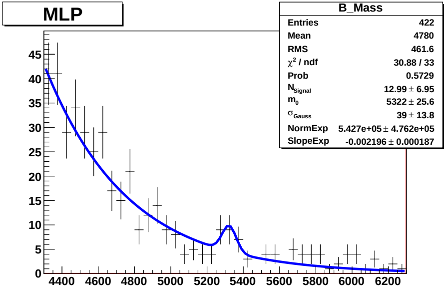
- MLP results
  - $S_{exp} = 35$
  - $B_{exp} = 142 \pm 11$
  - $S_{fit} = 49.2 \pm 15.4$
  - $(S+B)_{obs} = 156$



Resolved π^0 results



- Fisher results
  - $S_{exp} = 14$
  - $B_{exp} = 89 \pm 11$
  - $S_{fit} = 15.0 \pm 8.0$
  - $(S+B)_{obs} = 82$



- MLP results
  - $S_{exp} = 14$
  - $B_{exp} = 98 \pm 10$
  - $S_{fit} = 13.0 \pm 7.0$
  - $(S+B)_{obs} = 61$