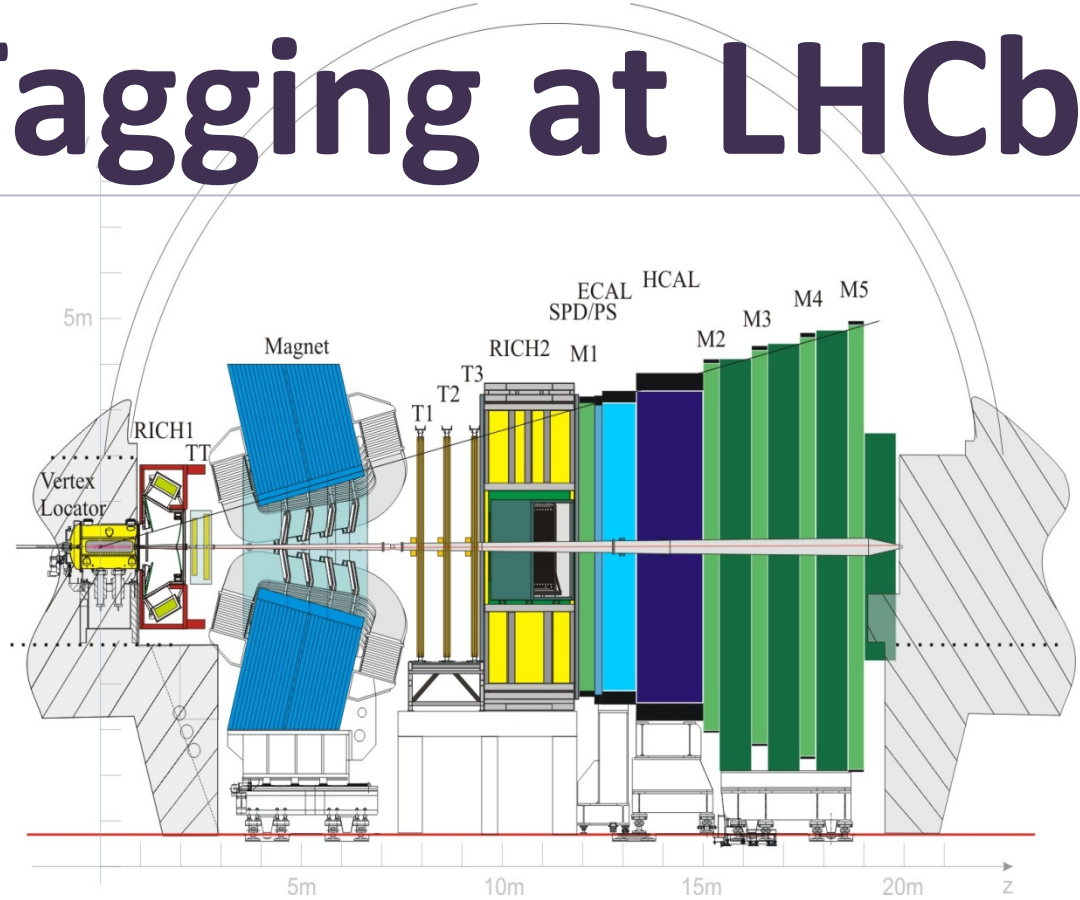
Flavour Tagging at LHCb

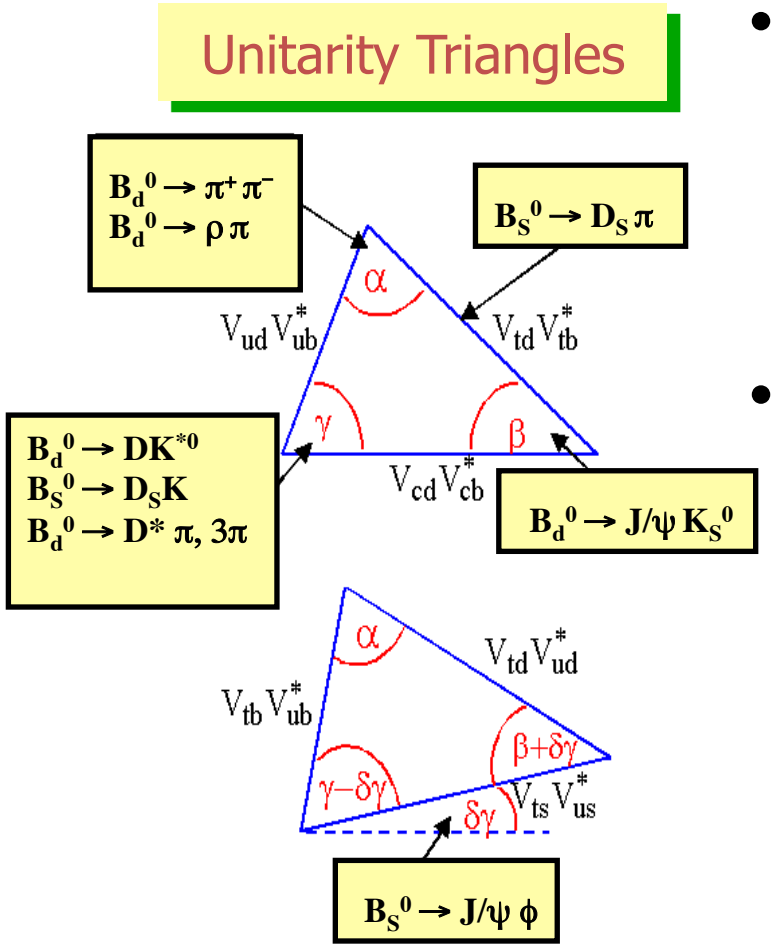
Marc Grabalosa Gándara



Institut de Ciències del Cosmos

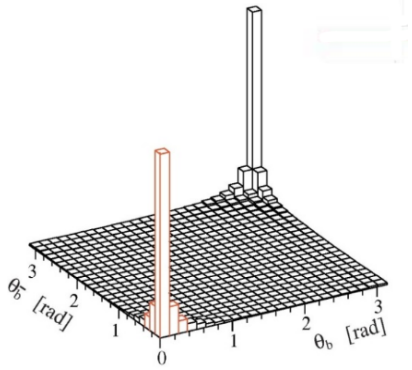
UNIVERSITAT DE BARCELONA

Motivation

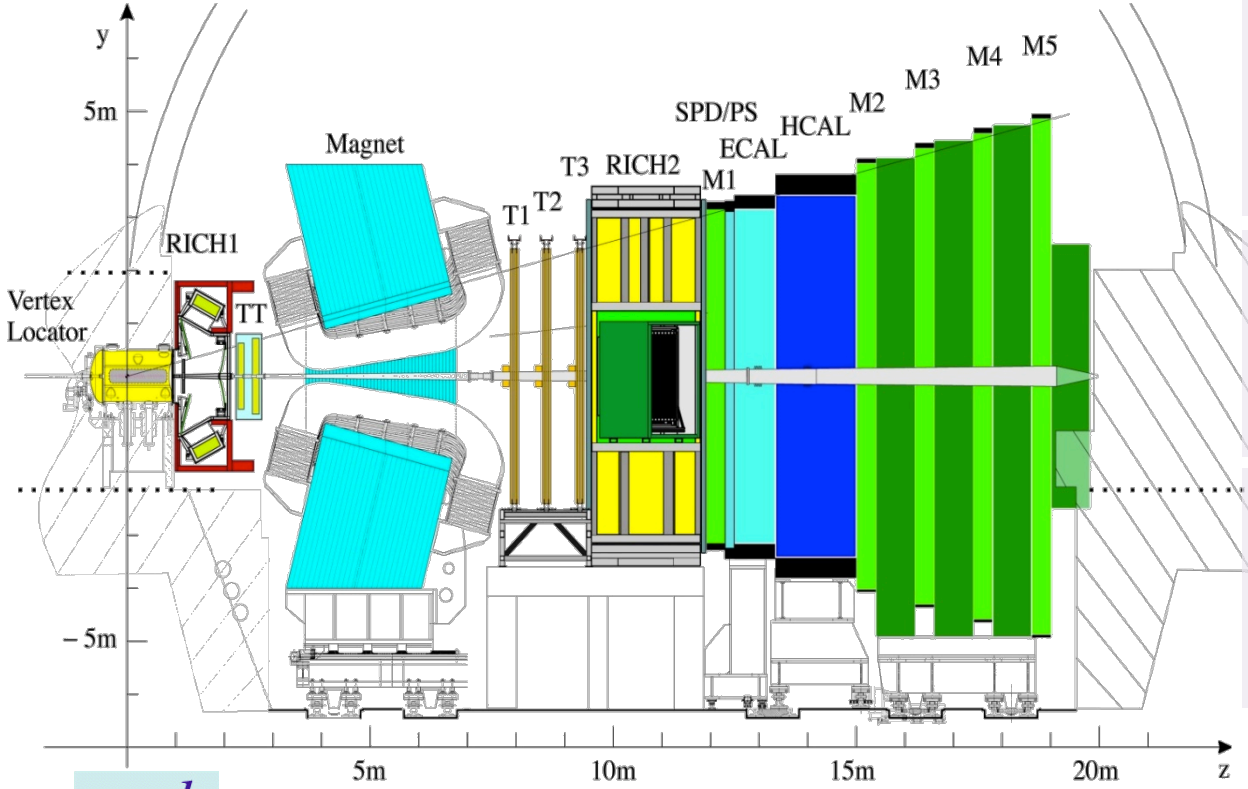


- The LHCb will improve the precision on CKM parameters and will look for new physics in rare decays
- Most of the key measurements require the knowledge of the initial flavour of the B meson
  - B_s oscillation frequency, phase and $\Delta\Gamma_s$ ($B_s \to D_s\pi$, $J/\Psi\Phi$, $J/\Psi\eta$, $\eta_c\Phi$)
  - Measure the CKM parameters
    - β with $B_d \to J/\psi K_S$ as a proof of principle

LHCb overview



$\mathcal{L}$ = 2(5)·10^{32} , **2fb^{-1}/year**
10 MHz visible interaction
10 KHz bb (**10^{12} bb / year**)



VELO:
Reconstruct vertexes
Proper time measurement

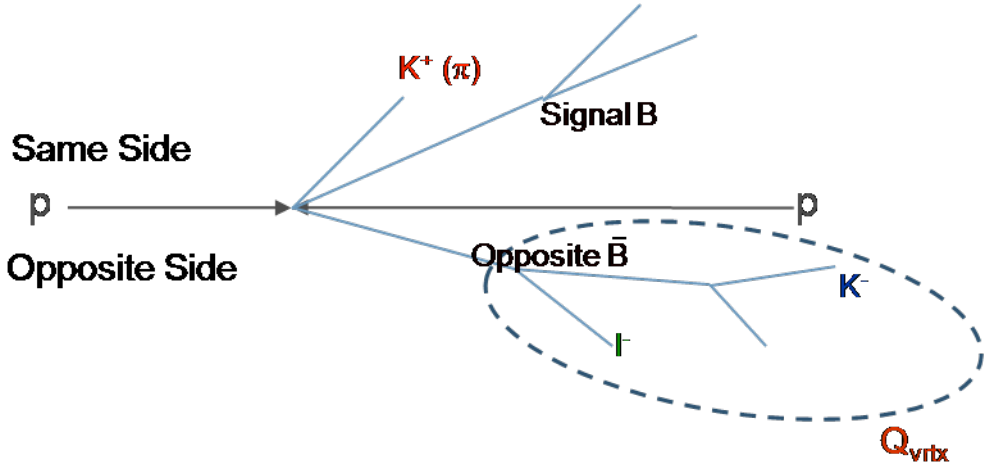
Tracking & magnet:
Excellent tracking reconstruction

Calorimeter:
Energy measurement
Identify γ, π^0

Muon and RICH detectors:
For Particle Identification

Flavour Tagging

- Flavour Tagging is the procedure to determine the flavour of the reconstructed B meson at production time
- Tagging algorithms are usually classified into two groups: Same Side and Opposite Side



Tagging efficiency $$\varepsilon_{tag} = \frac{N_R + N_W}{N_R + N_W + N_U}$$

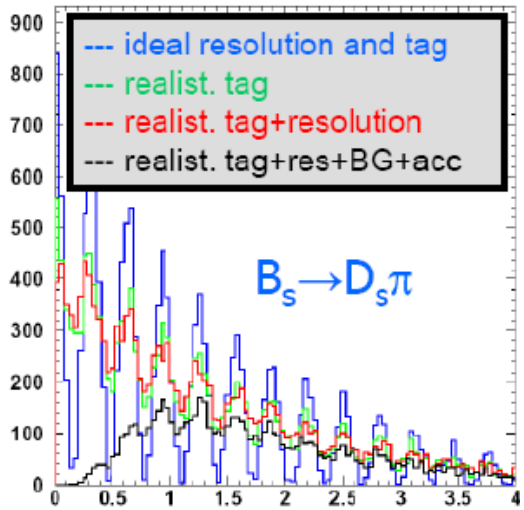
Wrong tag fraction $$\omega = \frac{N_W}{N_R + N_W}$$

R = right tagged
W = wrong tagged
U = untagged

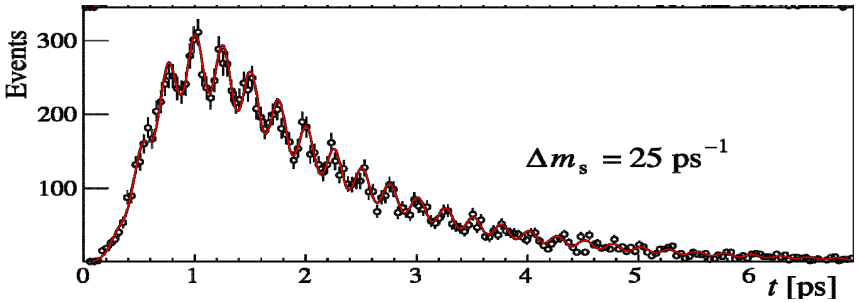
Taggers
OS: muons, electrons, kaons, vertex charge
SS: kaons (B_s) or pions ($B_{d,u}$)

Tagging and dilution

- The tagging algorithms may fail in identifying the flavour so that the observed CP-asymmetry is diluted wrt the true asymmetry



Expected unmixed $B_s \to D_s^- \pi^+$ sample in one nominal year of LHCb data taking

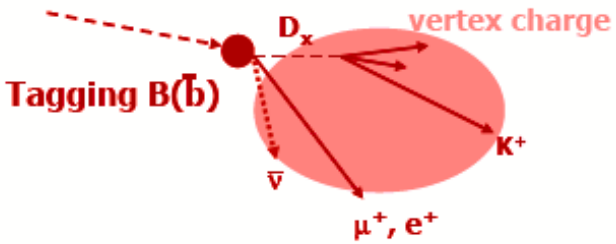


Effective efficiency $\varepsilon_{eff} = \varepsilon_{tag}(1-2\omega)^2$

- The statistical uncertanty on the measured CP asymmetry is directly related to the effective efficiency

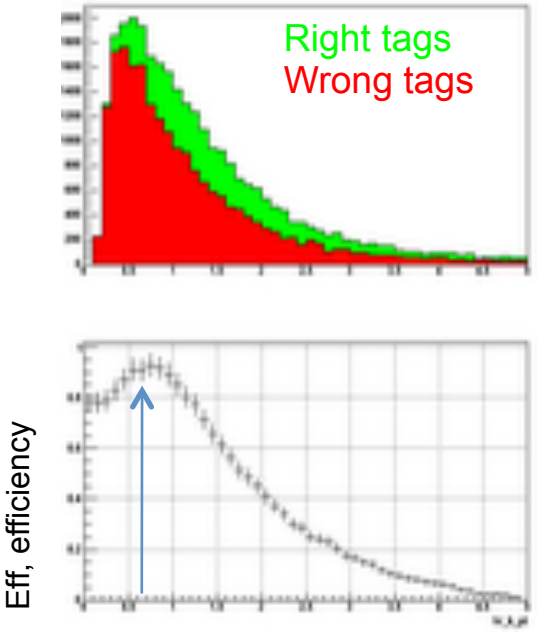
Opposite Side taggers

- Opposite Side taggers exploits the anti-correlation in the flavour of the two B produced in the same event by looking at the decay products of the opposite B



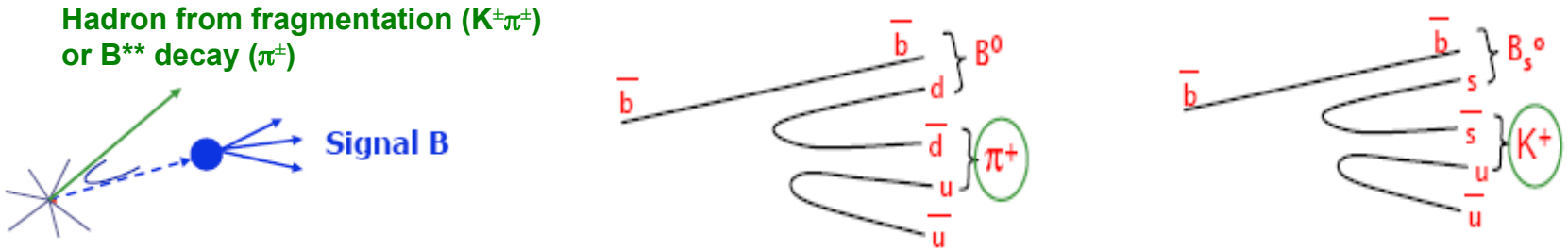
Kinematic and geometrical variables (significance of impact parameter, P, Pt,...) show a dependence in purity of right vs wrong tags. Find the optimal cut which maximizes the effective efficiency.

Pt of K(OS)

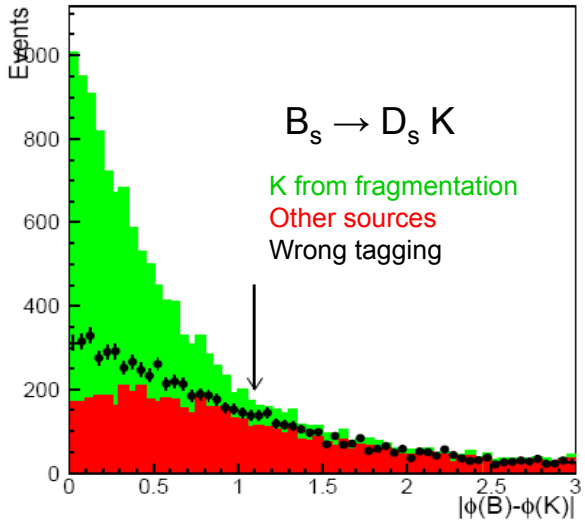


Same Side taggers

- Same Side taggers exploit the correlation in the charge of the mesons produced in the fragmentation chain of the B signal



Particle selection and proximity cuts to signal B are applied



Combination of taggers

- The final decision is taken combining all the individual tagger decisions (if there is more than one)
  - **Neural Net**: a wrong tag probability is computed for each of the tagger candidates using a Neural Net (NN). We combine them using a probability function or another NN
  - **Particle Identification**: Sort the events based on the particle identification of the track
- For each event, a wrong tag (ω) probability is computed.
- The events are sorted in 5 tagging categories of decreasing ω, to avoid mixing events with different dilutions. This improves the overall tagging performances and the corresponding precision in the CP fits

Monte Carlo Results

- Expected results from Monte Carlo

Tagger	ε%	ω%	ε_{eff} %
OS Muon	5.1	29.9	0.8
OS Electron	2.5	28.1	0.5
OS Kaon	18.8	36.4	1.4
SS Kaon	19.0	35.4	1.6
OS Vertex Charge	23.2	39.4	1.0
Combined	43.1	31.8	5.7

$B_s \rightarrow J/\psi\ \phi$

Combined tagging performance for different channels (Differences can be due to B spectra, trigger, etc)

	ε %	ω %	ε_{eff} %
$B_s \rightarrow D_s\pi$	60.7	30.9	8.6 ± 0.2
$B_d \rightarrow J/\psi\ K^*$	53.2	35.8	4.3 ± 0.1
$B_d \rightarrow \pi\pi$	56.8	34.4	5.5 ± 0.2
$B_u \rightarrow J/\psi\ K^+$	53.1	36.1	4.1 ± 0.1

Calibration & optimization of taggers

- Test tagging algorithms with real data
- Use control channels to optimize and calibrate the taggers
- **2010** : collected 37pb^{-1} in different trigger/pile up conditions.

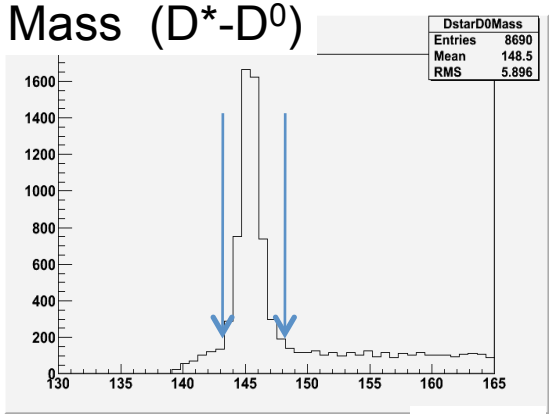
Channel	Statistcis (37pb^{-1})	B/S	ε_{tag}	ω
$B^+ \rightarrow J/\Psi K^+$	14k	2-8 (MC)	Count tagged events	Count n. of right and wrong tagged signal events
$B^0_d \rightarrow D^{*-}\mu^+\nu$	60k	0.3	Count tagged events	Fit the time evolution. Obtain mistag from B^0_d oscillation amplitude

Taggers calibration

- Optimize Flavour Tagging cuts per tagger in one channel ($B^0_d \rightarrow D^{*-}\mu^+\nu$)
  - Study trigger dependencies
- Split in tagging categories to improve performances. Measure the tagging performances in statistical independent samples ($B^0_d \rightarrow D^{*-}\mu^+\nu$ and $B^+ \rightarrow J/\Psi K^+$) and calibrate NN, and mistag probability per event
  - Verify calibration in a statistical independent sample ($B^0 \rightarrow J/\Psi K^*$)
- Use the calibrated mistag per event in the extraction of physics parameters
  - (e.g. the weak phase ϕ_s in $B^0_s \rightarrow J/\Psi\phi$)

Taggers optimitzation

- Tagging cuts have been optimized in MC
- Running on real data give worse performances, induced by some variables which are not calibrated
- First fit on oscillation was done for first data ($3.5pb^{-1}$)
  - Background not taken into account
  - Consider only events $B^0_d \rightarrow D^{*-}[D^0(K\pi)\pi^-]\mu^+\nu$
  - Consider all events in mass window of D^*-D^0 as signal

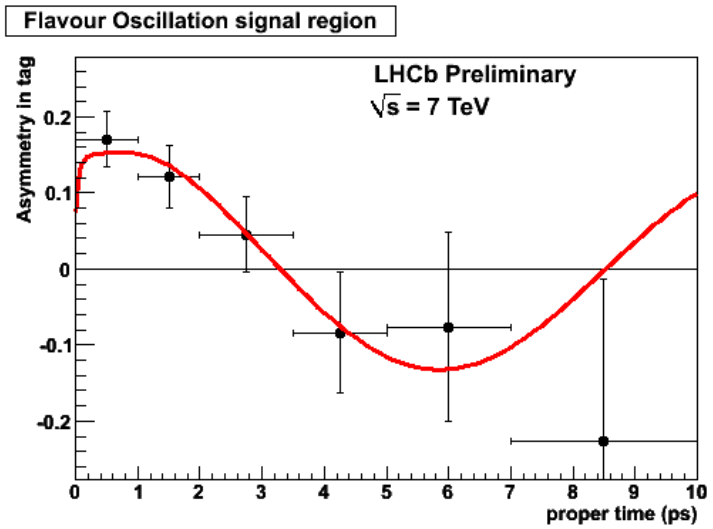


$3.5pb^{-1}$	ε%	ω%	εD^2 %
μ	5.6	31	0.8±0.3
e	3.0	29	0.5±0.3
K	22.4	38	1.3±0.3
Vtx	44.8	46	0.3±0.1

Opposite Side performances

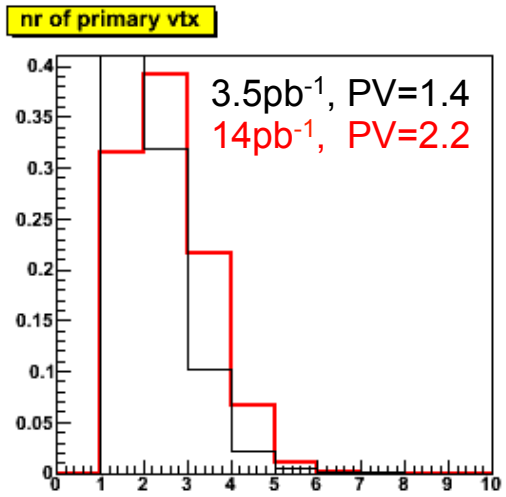
- Use $B^0_d \rightarrow D^{*-}\mu^+\nu$ channel
- More data has been collected
- Measure tagging performances with a detailed fit to account for different background sources

14pb^{-1}	ε%	ω%	εD^2 %	εD^2 % (mc)
μ	6.8±0.2	33±3	0.8±0.2	0.9±0.1
e	2.0±0.1	41±4	0.1±0.1	0.4±0.1
K	20.2±0.4	40±1.6	0.9±0.3	1.7±0.2
Vtx	28.4±0.4	44±1	0.4±0.2	1.0±0.1
All OS	40.7±0.5	41±1	**1.3±0.3**	**2.7±0.2**



2010 Runing conditions

- Evolving trigger conditions to accomodate increasing luminosity
- More primary vertices than in original design
  - Higher number of PV induces a lower tagging performance



$B^0_d \rightarrow D^{*-}\mu^+\nu$		ε%	ω %	εD^2 %
μ	PV = 1	6.4	31.9	0.8±0.4
	PV > 1	6.5	38.1	0.4±0.2
K	PV = 1	18.9	40.2	0.7±0.3
	PV > 1	21.5	41.8	0.6±0.3
Vtx	PV = 1	26.5	45	0.2±0.2
	PV > 1	30.6	46	0.2±0.2
All OS	PV = 1	38.4	41	1.4±0.3
	PV > 1	42.6	44	0.6±0.3

Combination of taggers

- The final decision is taken combining all individual tagger decisions

30 pb^{-1}	ε%	ω%	εD^2%
One sample	42.0±0.3	41.8±0.6	1.13±0.17
5 categories	41.8±0.3	39.8±0.6	1.73 ±0.20

  - Still lower than the Monte Carlo expectations ($\varepsilon D^2 = 2.5\%$)
  - Already above current CDF opposite side tagging performance ($\varepsilon D^2 = 1.2 \pm 0.2\%$)
  - Sorting events in tagging categories enhances the tagging power by ~50%
- The future introduction of same side tagging should nearly double the total effective efficiency

Conclusions

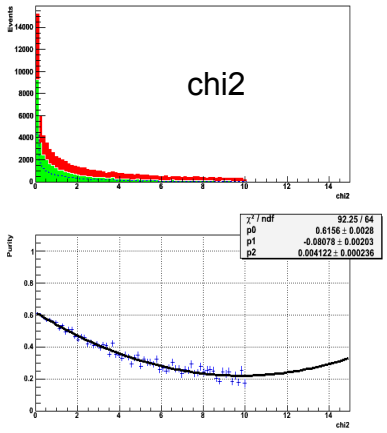
- Flavour tagging is a fundamental ingredient for B physics measurements in LHCb
- LHCb can already tag the flavour of B particles!
  Although effective efficiencies are still ~50% the MC expectation
- We are currently working on improving the tagging performances and calibrate the tagging response
- Tagging depends on a very large number of variables (track multiplicities, number of primary vertices, etc)

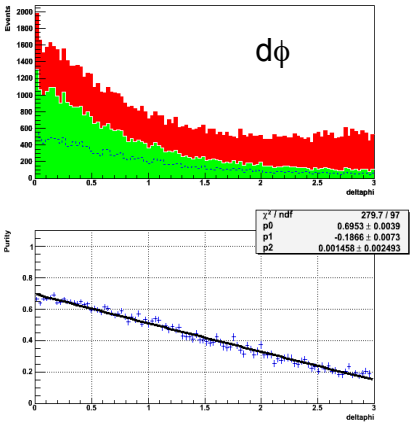
Thanks to: M. Musy, M. Calvi, S. Vecchi, S. Furcas, O. Leroy, et al.

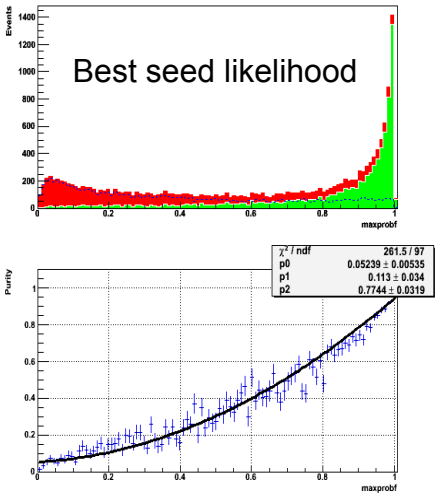
Back-up

Vertex Charge tagger

- The charge of secondary vertex will give as the flavour of the B
  - Select tracks from OS and build a 2-track seed



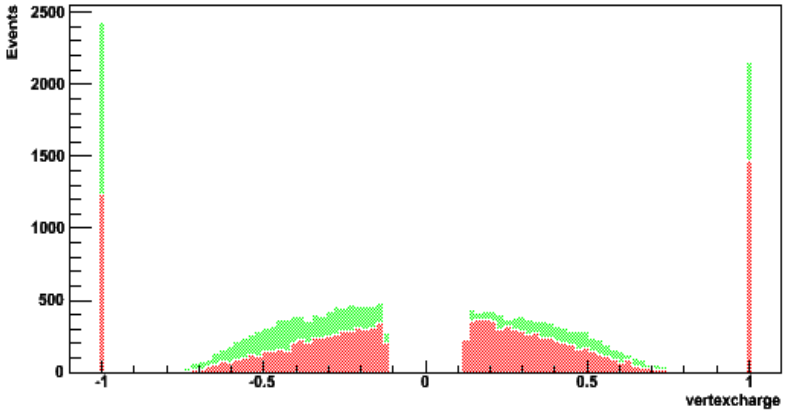




  - Make a likelihood
  - Add tracks to the seed based on kinemaical cuts
  - Evaluate a pt weighted charge

Vertex Charge tagger

- The tagger decision is given according the weighted charge



- Treat like the other taggers using a NN
- A correlation exist between vertex charge tagger and the other OS taggers as it uses several tracks

MC2010 cuts

	IPPU_cut = 3.0		
μ	Muon_Pt_cut = 1.1 Muon_P_cut = 0.0 Muon_IPs_cut = 0.0 Muon_Ics_cut = 2.2 Muon_PIDm_cut = 2.0 ProbMin_muon = 0.0 AverageOmega = 0.33	K	Kaon_Pt_cut = **0.4** Kaon_P_cut = 4.0 Kaon_IPs_cut = **3.8** Kaon_IP_cut = 1.5 Kaon_LCS_cut = **2.5 chi2/ndf** Kaon_PIDk = 0.0 Kaon_PIDkp = -1.0 Kaon_ghost_cut= **-14.** Kaon_ipPU_cut = 4.0 ProbMin_kaon = 0.0 AverageOmega = 0.33
e	Ele_Pt_cut = 1.1 Ele_P_cut = 0.0 Ele_Ics_cut = **2.5** Ele_IPs_cut = 0 Ele_ghost_cut = **-15.** VeloChargeMin = 0.0 VeloChargeMax = 1.3 EoverP = 0.85 Ele_PIDe_cut = **4.0** ProbMin_ele = 0.0 AverageOmega = 0.33	Vtx	PowerK = 0.35 MinimumVCharge= 0.12 ProbMin_vtx = 0.52 P0 = 5.255669e-01 P1 =-3.251661e-01 Gt075 = 0.35 TracksEq2 = 0.4141 TracksGt2 = 0.3250 AverageOmega = 0.41

Strip9 cuts

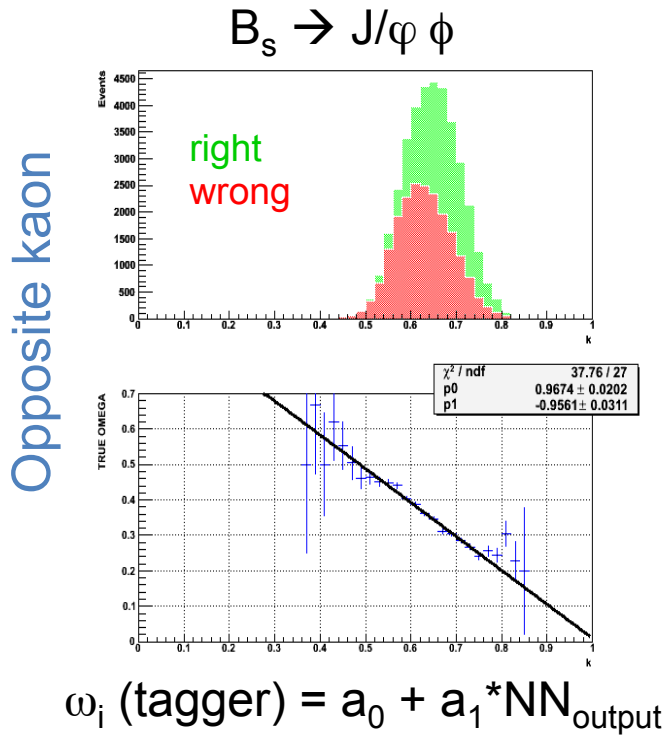
	IPPU_cut = 3.0		
μ	Muon_Pt_cut = 1.1 Muon_P_cut = 0.0 Muon_IPs_cut = 0.0 Muon_Ics_cut = 2.2 Muon_PIDm_cut = 2.0 ProbMin_muon = 0.0 AverageOmega = 0.33	K	Kaon_Pt_cut = **0.9** Kaon_P_cut = 4.0 Kaon_IPs_cut = **4.** Kaon_IP_cut = 1.5 Kaon_LCS_cut = 5 Kaon_PIDk = 0.0 Kaon_PIDkp = -1.0 Kaon_ghost_cut= **-999.** Kaon_ipPU_cut = 4.0 ProbMin_kaon = 0.0 AverageOmega = 0.33
e	Ele_Pt_cut = 1.1 Ele_P_cut = 0.0 Ele_Ics_cut = **5** Ele_IPs_cut = 0 Ele_ghost_cut = **-999.** VeloChargeMin = 0.0 VeloChargeMax = 1.3 EoverP = 0.85 Ele_PIDe_cut = **5.5** ProbMin_ele = 0.0 AverageOmega = 0.33	Vtx	PowerK = 0.35 MinimumVCharge= 0.12 ProbMin_vtx = 0.52 P0 = 5.255669e-01 P1 =-3.251661e-01 Gt075 = 0.35 TracksEq2 = 0.4141 TracksGt2 = 0.3250 AverageOmega = 0.41

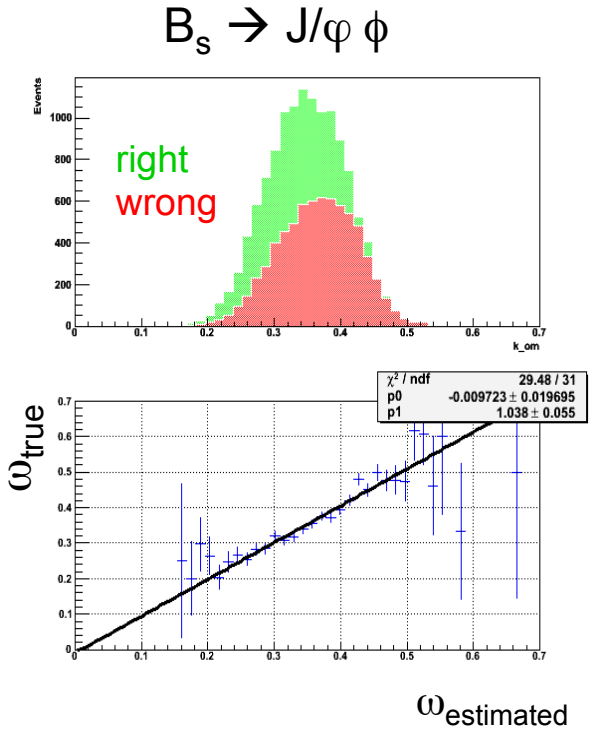
Strip10 cuts

	IPPU_cut = 3.0		
μ	Muon_Pt_cut = 1.1 Muon_P_cut = 0.0 Muon_IPs_cut = 0.0 Muon_lcs_cut = 2.2 Muon_PIDm_cut = 2.0 ProbMin_muon = 0.0 AverageOmega = 0.33	K	Kaon_Pt_cut = **0.8** Kaon_P_cut = 4.0 Kaon_IPs_cut = **2.8** Kaon_IP_cut = 1.5 Kaon_LCS_cut = **2.7** Kaon_PIDk = **2.0** Kaon_PIDkp = -1.0 Kaon_ghost_cut = -999. Kaon_ipPU_cut = **6.0** ProbMin_kaon = 0.0 AverageOmega = 0.33
e	Ele_Pt_cut = **0.8** Ele_P_cut = 0.0 Ele_lcs_cut = **5** Ele_IPs_cut = 0 Ele_ghost_cut = -999. VeloChargeMin = 0.0 VeloChargeMax = **1.6** EoverP = **0.6** Ele_PIDe_cut = **4.** ProbMin_ele = 0.0 AverageOmega = 0.33	Vtx	PowerK = 0.35 MinimumVCharge = **0.09** ProbMin_vtx = 0.52 P0 = 5.255669e-01 P1 = -3.251661e-01 Gt075 = 0.35 TracksEq2 = 0.4141 TracksGt2 = 0.3250 AverageOmega = 0.41

Neural Net method

- For each event, a wrong tag probability is computed for each of the tagger candidates using a Neural Net (NN).
- Each tagger gives an ω_i as a function of the NN output



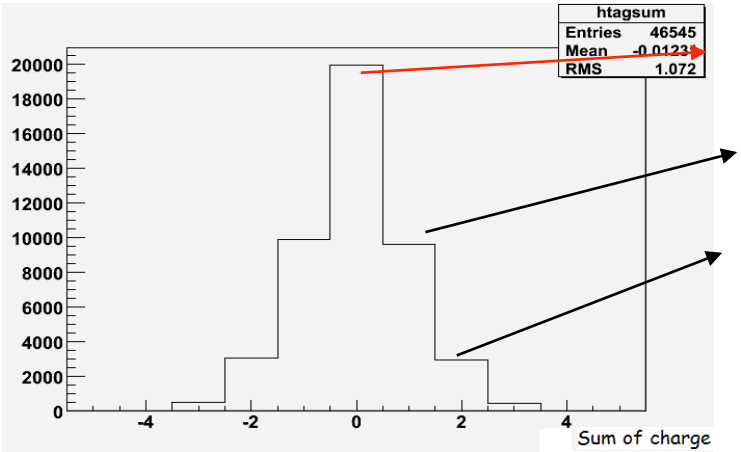


Combination of taggers

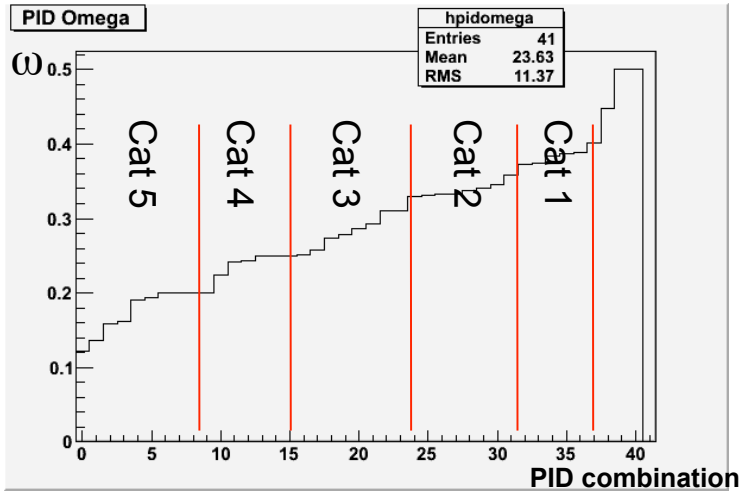
- The final probability for the event will be a combination of the taggers wrong tag fraction
  - Combine the taggers as independent
    - $P^{+1} = (1-\omega_k)\,\omega_e$ ... $P^{-1} = \omega_k\,(1-\omega_e)$ ...
    - $P(B) = P^{+1}/(P^{+1} + P^{-1})$
    - If $P(B) < 0.55$ event is left untagged
  - Use a NN to take into account correlation between taggers
- Calculate the final combined effective efficiency, we bin the events in 5 categories (and treat them separetly in the CP fits)

PID based combination of taggers

- We form possible combinations according to particle identification (muons, electrons, kaons, vertex, SS), and the sum of the individual taggers decision

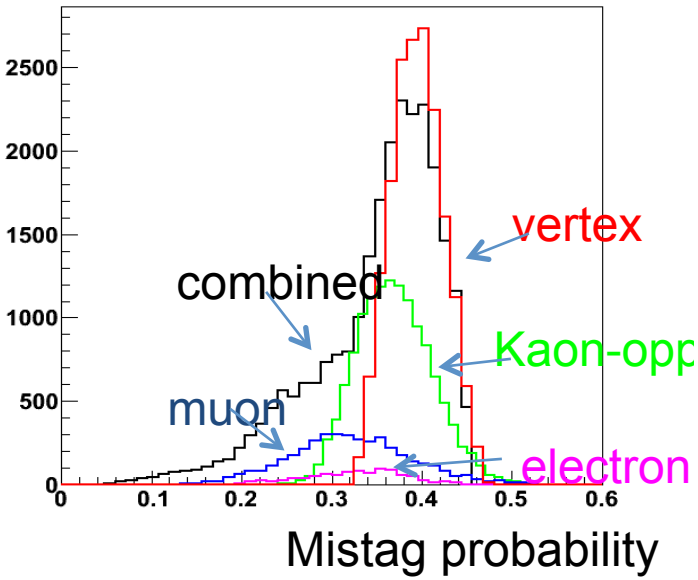


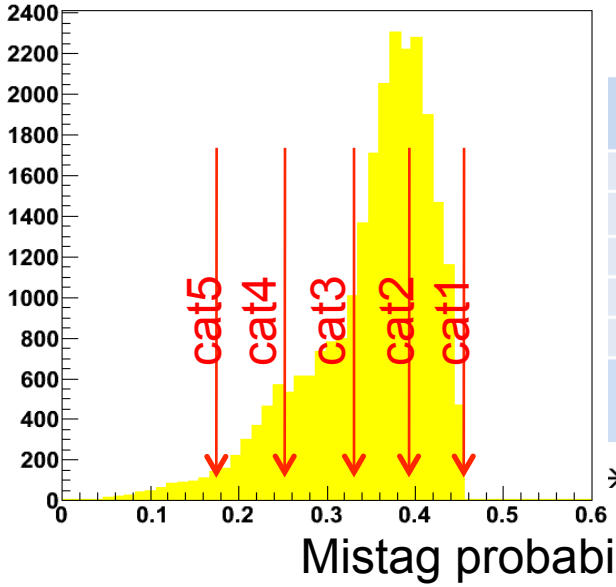
- Sort all possible combinations, according to the estimated omega in a control channel, and bin the events in 5 categories



Combination of taggers

- Use event mistag probability to split into 5 independent samples and fit the B^0 oscillation in each
- Using the NN combination in real data:





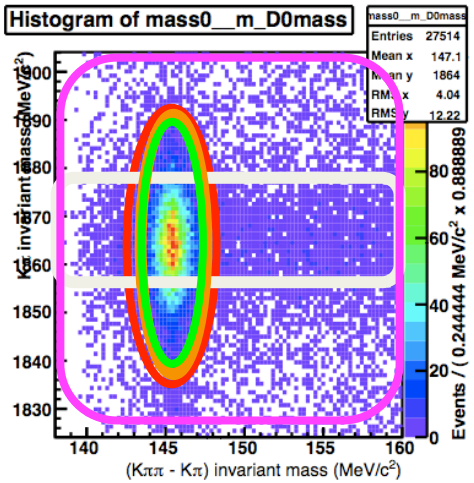
	ε (%)	ω (%)	εD² (%)
All OS	42.0 ±0.3	41.8 ±0.6	1.13 ±0.17
CAT 1	16.6 ±0.2	45.1 ±1.0	0.16 ±0.06
CAT 2	15.0 ±0.2	43.7 ±1.0	0.24 ±0.08
CAT 3	6.4 ±0.1	37.2 ±1.6	0.42 ±0.10
CAT 4	2.9 ±0.09	27.9 ±2.1	0.57 ±0.11
CAT 5	1.02 ±0.05	20.9 ±3.5	0.35 ±0.08
Combined	41.8 ± 0.3	39.8 ±0.6	1.73 ± 0.20

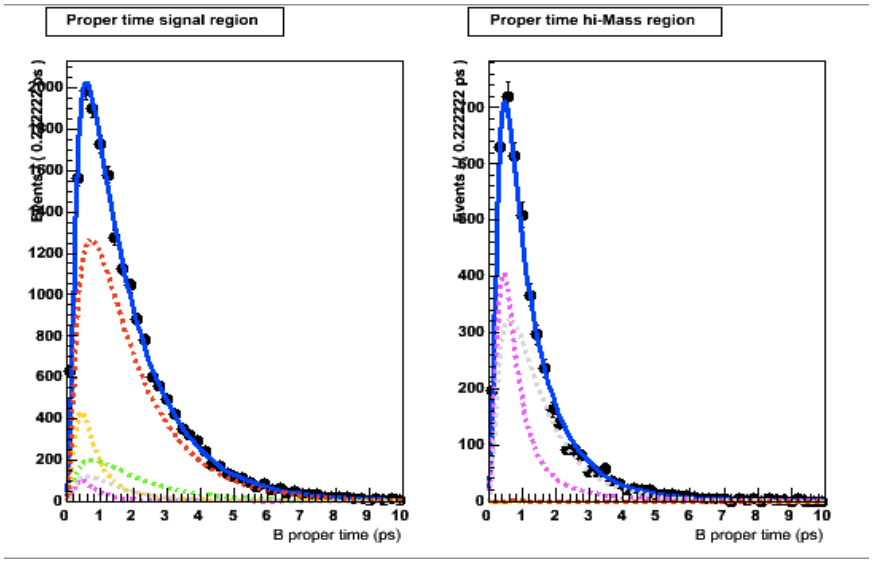
→ +50% in tagging power from the use of categories.

Fit $B^0_d \rightarrow D^{*-}\mu^+\nu$

(thanks to S.Vecchi)

- Measure tagging performances with a more detailed fit to account for different background sources
  - **Signal:** $B^0 \rightarrow D^{*-}\mu^+\nu$
  - **Background B^+:** $pp \rightarrow B^+ \rightarrow D^{*-}\mu^+\nu X$ true D* from B, not oscillating
  - **Background D*:** $pp \rightarrow D^{*-}\mu^+ X$ true D* from PV
  - **Background comb:** $pp \rightarrow (K\pi\pi)^-\mu^+ X$ not D* from PV
  - **Background D0fB:** $pp \rightarrow b\bar{b} \rightarrow D^0\mu^+\nu\pi^- X$ true D^0 from B, oscillating (yes/no at choice)
- Use tagged/untagged data, observables:
- Background has independent tagging parameters (some asumptions are needed)

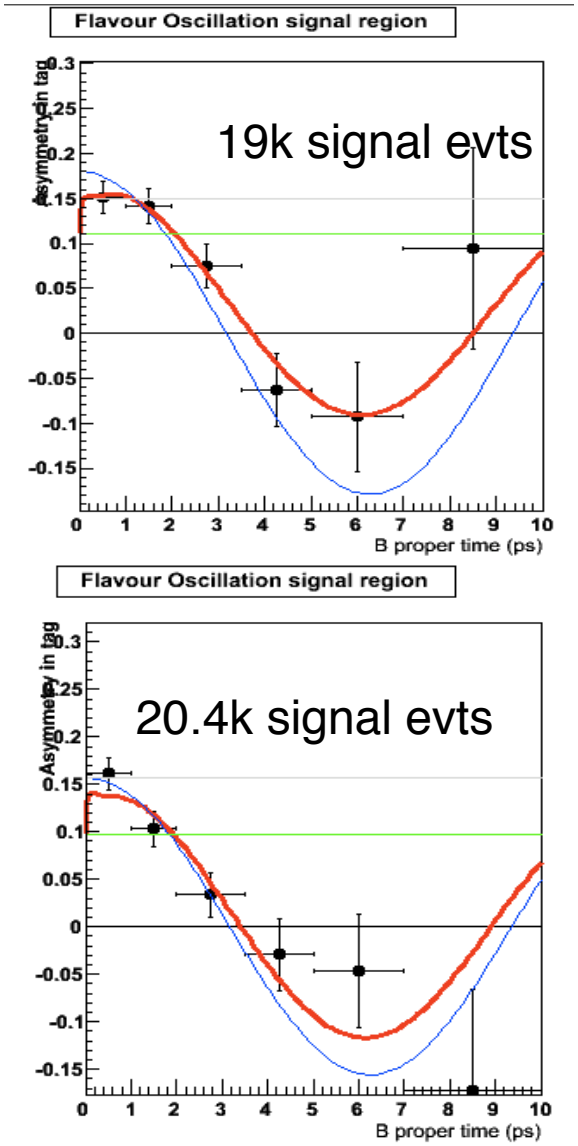




OS performances ($B^0_d \to D^{*-}\mu^+\nu$)

Strip10	ε%	ω %	εD^2 %
μ	6.8±0.2	33.3±2.7	0.76±0.24
e	2.0±0.1	41.1±4.2	0.06±0.06
K	20.2±0.4	39.5±1.6	0.89±0.26
Vtx	28.4±0.4	44.3±1.3	0.37±0.16
All OS	40.7±0.5	40.9±1.1	**1.34±0.31**

Strip11	ε%	ω %	εD^2 %
μ	6.4±0.2	34.3±2.6	0.64±0.21
e	1.6±0.1	36.2±4.8	0.12±0.08
K	20.3±0.4	41.6±1.4	0.57±0.20
Vtx	28.6±0.4	44.3±1.2	0.37±0.15
All OS	41.4±0.5	42.1±1.0	**1.03± 0.25**



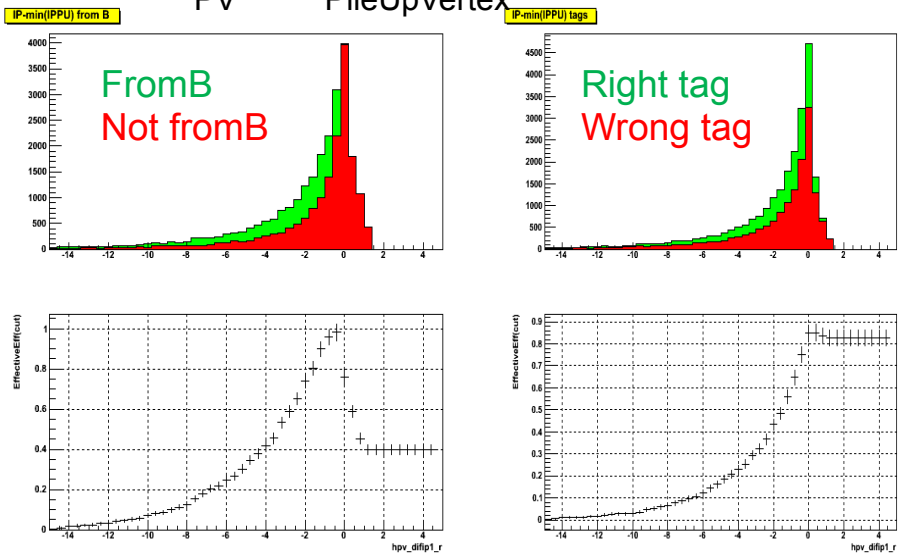
New tuning

- New tuning has been done for strip10 data, and tested for strip11

Strip11	ε%	ω %	εD² %
μ	4.3±0.2	28.4±2.8	0.64±0.21
e	3.0±0.1	38.9±3.4	0.15±0.09
K	17.6±0.3	39.7±1.4	0.75±0.20
Vtx	30.1±0.4	43.0±1.1	0.60±0.20

- Still a lot to improve
  - MC(ν=3) has been studied. New variables seems to discriminate quite good between the tagging particles which are coming or not from the B. Unfortunately are not so good to discriminate right/rong tags
  - Proper time of B opposite can help as to determine w.

IP_{PV} - $IP_{PileUpVertex}$



Comparison

B-factories vs. b-factory

	$e^+e^- \to \Upsilon(4S) \to B\bar{B}$ PEPII, KEKB	$pp \to b\bar{b}X$ ($\sqrt{s}$ = 14 TeV, Δt_{bunch}=25 ns) LHC (LHCb–ATLAS/CMS)	
Production $\sigma_{b\bar{b}}$	1 nb	~500 μb	☺
Typical $b\bar{b}$ rate	10 Hz	100–1000 kHz	
$b\bar{b}$ purity	~1/4	$\sigma_{b\bar{b}}/\sigma_{inel}$ = 0.6% Trigger is a major issue !	☹
Pileup	0	0.5–5	
b-hadron types	B^+B^- (50%) $B^0\bar{B}^0$ (50%)	B^+ (40%), B^0 (40%), B_s (10%) B_c (< 0.1%), b-baryons (10%)	☺
b-hadron boost	Small	Large (decay vertexes well separated)	
Production vertex	Not reconstructed	Reconstructed (many tracks)	
Neutral B mixing	Coherent $B^0\bar{B}^0$ pair mixing	Incoherent B^0 and B_s mixing (extra flavour-tagging dilution)	☹
Event structure	$B\bar{B}$ pair alone	Many particles not associated with the two b hadrons	

Pdf parametrization B

Pdf(Δm,m,t,mix) = **S**(Δm,m,t,mix) +
B⁺(Δm,m,t,mix) + **D***(Δm,m,t,mix) +
comb(Δm,m,t,mix) + **D0fB**(Δm,m,t,mix)

Signal: $B^0 \to D^{*-} \mu^+ \nu$
background B⁺: $pp \to B^+ \to D^{*-} \mu^+ \nu X$ true D* from B, not oscillating
Background D*: $pp \to D^{*-} \mu^+ X$ true D* from PV
Background comb: $pp \to (K\pi\pi)^- \mu^+ X$ not D* from PV
Background D0fB: $pp \to b\bar{b} \to D^0 \mu^+ \nu \pi^- X$
true D^0 from B, oscillating (yes/no at choice)

Pdf parametrization B (1/4)

Pdf(Δm,m,t,mix) = **S**(Δm,m,t,mix) +
B+(Δm,m,t,mix) + **D***(Δm,m,t,mix) +
comb(Δm,m,t,mix) + **D0fB**(Δm,m,t,mix)

Signal B^0: S(Δm) * **S**(m) **S**(t,mix)

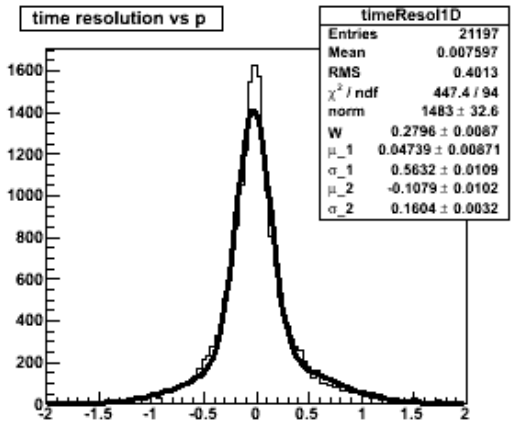
S(Δm) = double gaussian $(m_{\Delta m}, \sigma^1_{\Delta m}, \sigma^2_{\Delta m}, f^1_{Dm})$

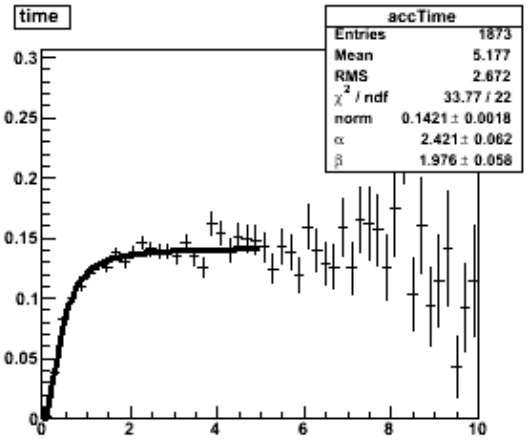
S(m) = double gaussian $(m_{D0}, \sigma^1_m, \sigma^2_m, f^1_m)$

S(t,mix) = A(t) * RooBDecay(τ^{B0} , ε^S, ω^S ; ΔM_{B0}=0.507, $\Delta\Gamma$=0.)
with time resolution from MC double gaussian $(\mu_t, \sigma^1_t, \sigma^2_t, f^1_t)$ (*)
A(t) = $(\alpha t)^\beta / (1 + \alpha t)^\beta$ α, β fixed from montecarlo (**)

(*) assumptions to be further verified

(**)Montecarlo doesn't include the trigger cuts





Pdf parametrization B (2/4)

Pdf(Δm,m,t,mix) = **S**(Δm,m,t,mix) +
B+(Δm,m,t,mix) + **D***(Δm,m,t,mix) +
comb(Δm,m,t,mix) + **D0fB**(Δm,m,t,mix)

background B+: B+ (Δm) * **B+** (m) * **B+** (t) ***B+**(mix) f_{B+}=free parameter

B+ (Δm) = same as the signal

B+ (m) = same as the signal

B+ (t) = A(t) * RooDecay(τ^{B+}=**1.088***τ^{B0})

with time resolution from MC double gaussian (μ_t, σ^1_t, σ^2_t,f^1_t)

B+ (mix) = ε^S (1- ω^S) mix=+1 ; ε^S ω^S mix=-1 or 1 - ε^S mix=0 (*)

(*) assumptions to be further verified

Pdf parametrization B (3/4)

Pdf(Δm,m,t,mix) = **S**(Δm,m,t,mix) +
B+(Δm,m,t,mix) + **D***(Δm,m,t,mix) +
comb(Δm,m,t,mix) + **D0fB**(Δm,m,t,mix)

Background D*: D* (Δm) * **D*** (m) * **D*** (t) ***D*** (mix) f_{D^*}=0.05 free parameter

D* (Δm) = same as the signal

D* (m) = same as the signal

D* (t) = same as **comb**(t) = RooLandau(μ^{bg}_t, σ^{bg}_t) (*)

D* (mix) = same as **comb** (mix) =

ε^{comb} (1- ω^{comb}) mix=+1;

ε^{comb} ω^{comb} mix=-1 or 1 - ε^{comb} mix=0 (*)

Background comb: comb(Δm)* **comb**(m)* **comb**(t)***comb**(mix)

f_{comb}=free parameter

comb(Δm) = RooDstD0BG(Δm,m_{th},c)

comb(m) = Constant(m)

comb(t) = RooLandau(μ^{bg}_t, σ^{bg}_t)

comb(mix) = ε^{comb} (1- ω^{comb}) mix=+1;

ε^{comb} ω^{comb} mix=-1 or 1 - ε^{comb} mix=0

Pdf parametrization B (4/4)

Pdf(Δm,m,t,mix) = **S**(Δm,m,t,mix) +
B+(Δm,m,t,mix) + **D***(Δm,m,t,mix) +
comb(Δm,m,t,mix) + **D0fB**(Δm,m,t,mix)

Background D0fB: D0fB(Δm)***D0fB**(m)***D0fB**(t)***D0fB**(mix)

f_{D0fB}=free parameter

D0fB(Δm) = RooDstD0BG(Δm,m_{th},c) (like comb)
D0fB(m) = double Gaussian(m,m_{D0}, σ^1_m, σ^2_m,f^1_m)
D0fB(t) = A(t) * RooDecay(t^{D0fB}=ft*τ^{B0}) (**) assumption

D0fB(mix) = ε^{D0fB} (1- ω^{D0fB}) mix=+1 , $\varepsilon^{D0fB}\omega^{D0fB}$ mix=-1
or 1 – ε^{D0fB} mix=0

OR
D0fB(t,mix) = A(t) * RooBDecay(t^{D0fB}=ft*τ^{B0}) (**) assumption