STANDALONE TRACK RECONSTRUCTION FOR THE LHCb UPGRADE SCINTILLATING FIBRE TRACKER
AND
STUDY OF DOUBLE CHARM B DECAYS.

PHENIICS Doctoral School Days 2016

RENATO QUAGLIANI
PHD STUDENT AT LAL ORSAY (FR) AND BRISTOL UNIVERSITY (UK).

MAY 9, 2016

OUTLINE

LHCb experiment at LHC and its upgrade

Hybrid seeding algorithm: a stand-alone track reconstruction algorithm for LHCb upgrade

Double charm decays at LHCb

Conclusions

LHCb EXPERIMENT AT LHC AND ITS UPGRADE

LHCb EXPERIMENT

LHCb is a high precision experiment designed to:

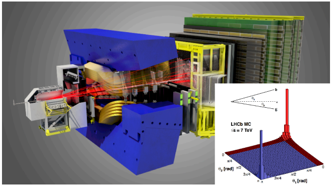
- Measure CP-violation using hadrons containing b and c quarks.
- Indirect searches of new physics trough SM strongly suppressed decay modes.

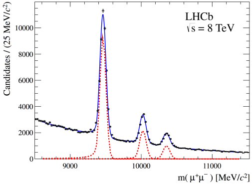
LHCb is a forward spectrometer:

- Heavy hadrons are produced boosted in the forward region: $1.9 < \eta < 4.9$.

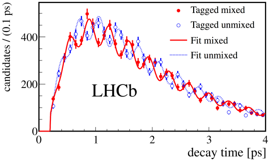
Requirements:

1. **Excellent tracking**(p, impact parameters and σ_{PV}):
    - $\sigma_{IP} \sim 20\ \mu\text{m}$.
    - $\epsilon_{\text{tracking}} > 96\%$.
    - $\frac{\delta p}{p} \sim 0.5 - 1.0\%$.
2. **Excellent τ_{decay} resolution**:
    - $\sigma_{\tau_{decay}} \sim 45\ \text{fs}$ for B mesons.
3. **Excellent particle identification (PID)**
    - ϵ_K-ID $\sim 95\%$.
    - ϵ_μ-ID $\sim 97\%$.

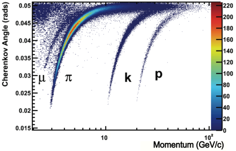




[J. High Energy Phys. 06 (2013) 064]

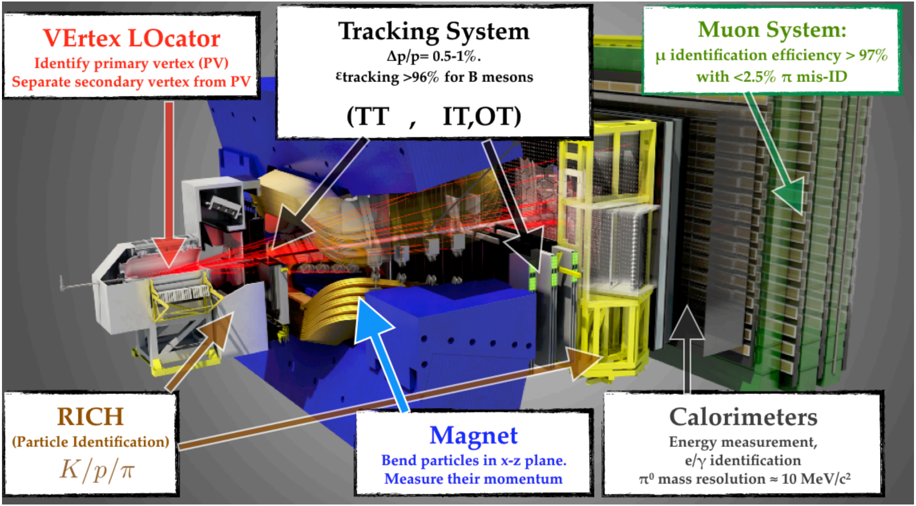


[New J. Phys. 15 (2013) 053021]



[Eur. Phys. J. C 73 (2013) 2431]

LHCb IN A NUTSHELL



VErtex LOcator

Identify primary vertex (PV)

Separate secondary vertex from PV

Tracking System

$\Delta p/p = 0.5\text{-}1\%$.

$\varepsilon_{tracking} > 96\%$ for B mesons

(TT , IT,OT)

Muon System:

μ identification efficiency > 97%

with <2.5% π mis-ID

RICH

(Particle Identification)

$K/p/\pi$

Magnet

Bend particles in x-z plane.

Measure their momentum

Calorimeters

Energy measurement,

e/γ identification

π^0 mass resolution ≈ 10 MeV/c^2

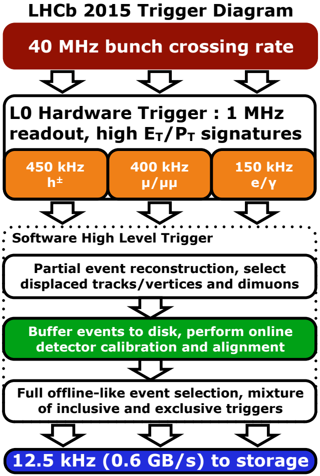
●LHCb ●VELO ●Tracking system ●Muon System ●RICH ●Magnet ●Calorimeters

LHCb TRIGGER IN RUN II AND RUN III

Conditions	Run II	Run III
$\sqrt{s}$	13 TeV	14 TeV
Bunch spacing	25 ns	25 ns
Pile-up	1.3-2.4	7.6
$\mathcal{L}$ $[\times 10^{32} cm^{-2} s^{-1}]$	**5**	**20**

RUN II (NOW) → RUN III (2020)

Hardware trigger replaced by a fully software based one.



LHCb 2015 Trigger Diagram

40 MHz bunch crossing rate

L0 Hardware Trigger : 1 MHz readout, high E_T/P_T signatures

450 kHz $h^\pm$

400 kHz $\mu/\mu\mu$

150 kHz e/γ

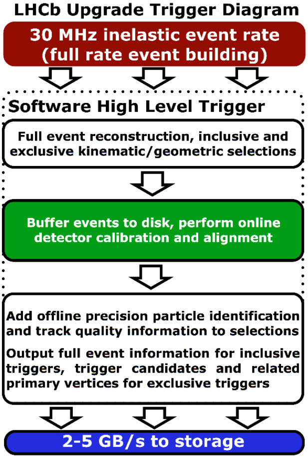
Software High Level Trigger

Partial event reconstruction, select displaced tracks/vertices and dimuons

Buffer events to disk, perform online detector calibration and alignment

Full offline-like event selection, mixture of inclusive and exclusive triggers

12.5 kHz (0.6 GB/s) to storage



LHCb Upgrade Trigger Diagram

30 MHz inelastic event rate (full rate event building)

Software High Level Trigger

Full event reconstruction, inclusive and exclusive kinematic/geometric selections

Buffer events to disk, perform online detector calibration and alignment

Add offline precision particle identification and track quality information to selections

Output full event information for inclusive triggers, trigger candidates and related primary vertices for exclusive triggers

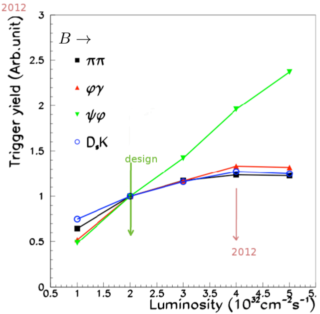
2-5 GB/s to storage

LHCb upgrade: from Run II to Run III

Higher $\mathcal{L} = 2 \times 10^{33} cm^{-2} s^{-1}$ (4× Run II) luminosity.

Physics

- Much more data (expected to collect at least 5 fb^{-1} per year)
- Improve precision down to theoretical level for $\cancel{CP}$ observable.



[CERN-LHCC-2011-001]

Problem

- Tighter p_{T} cut to stay within the 1 MHz read-out limitation in L0 (HW) hadronic trigger.
- Tighter cut → flattens of trigger yield for hadronic final states.
- Moreover, more complex event topology requires more information at trigger level.

Solution: upgrade 2020

1. **All sub-detector read-out at 40 MHz.**
   → Completely new tracking system.
2. **Fully implemented software trigger.**
   → Stringent timing limitations.

DETECTOR UPGRADE

Velo upgrade

- Hybrid pixel sensors closer to beam pipe $(5.5\ \mathrm{mm} \rightarrow 3.5\,\mathrm{mm})$
- Geometry simplify VELO, simpler track reconstruction.
- Better $\epsilon_{tracking}$ and σ_{IP-3D}.

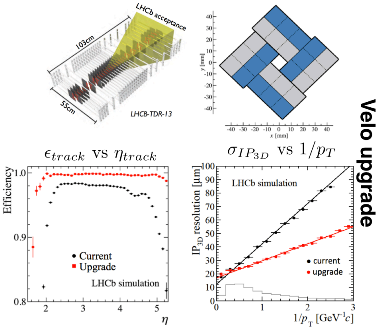
Upstream Tracker (UT)

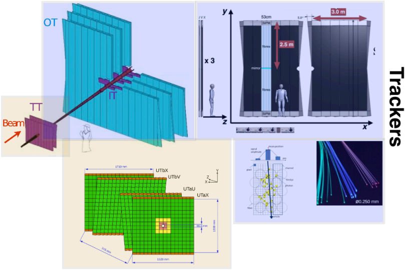
- Finer granularity and larger acceptance.
- 4 layers in *stereo* configuration.

Scintillating Fiber tracker (SciFi)

- Less complex than IT and OT.
- $2.5\ \mathrm{m}$ long fibers read-out by *SiPM*.
- $\sigma_{x-z} \sim 100\mu$m.

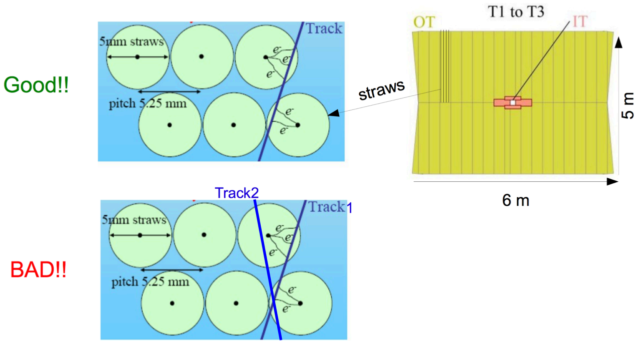
VELO upgrade TDR
Trackers upgrade TDR





WHY IS THE ACTUAL TRACKER NOT GOOD ENOUGH AND WE NEED THE SCIFI?

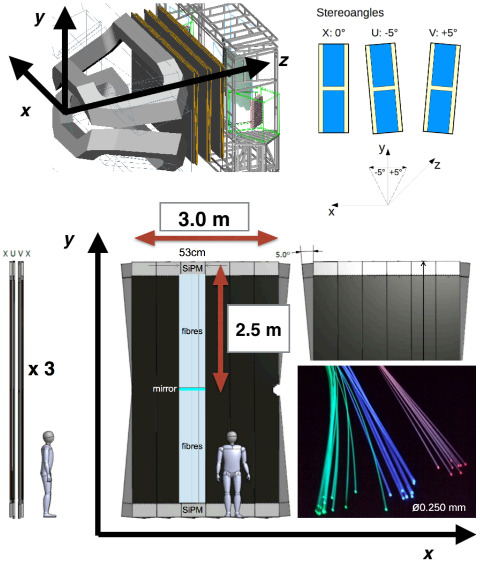
Outer tracker made by 5 mm straw gas drift tubes with a 35 ns drift time.
→ it becomes insensitive to multiple tracks per tube, also spillover (25 ns bunch spacing) is an issue.



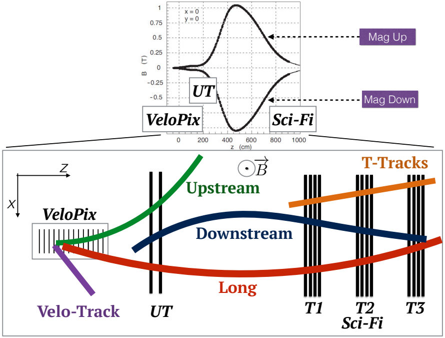
- In addition, at the upgrade conditions, the occupancy in the OT would be 40% leading to a large decrease of tracking efficiencies and higher fake track rate.

SCINTILLATING FIBER DETECTOR

- 3 stations of scintillating fibers read-out by *SiPM*.
- 4 layers each station in stereo configuration.
- 2.5 m long fibers mirrored on one end to increase light yield.
- Other end: read-out by *SiPM* arrays with single channel pitch of 250 m $\Rightarrow \sigma_{x-z} \sim 100\mu m$.
- Homogeneous detector, much simpler than IT+OT.
- *SiPM* and scintillating fibers allow fast read-out at collision rate.
- 360 m^2 active area and 10.000 km of fibers for full detector.



TRACK RECONSTRUCTION AT LHCb : SEEDING ALGORITHM



- Different track type are reconstructed by different algorithms.
- Seeding algorithm: **Standalone track reconstruction algorithm using solely the information from the SciFi**.
  Reconstruction of T-Tracks and provide T-segments.
- Provide input to reconstruct downstream and long tracks.
- Downstream: daughters of long-lived particles such as K^0_S and Λ^0.
- Long: particles produced close to the interaction point.

Starting point

The default version of the algorithm (TDR) had some critical issues:

- **Low tracking efficiencies** (how many good tracks are found by the algorithm).
- **High ghost rate** (how many fake tracks are found by the algorithm).
- **Slow** (how long does it take to run).

Just to give an idea:

Track type	Performance
from B	$(84.3 \pm 0.2)\,\%$
from long-lived	$(73.4 \pm 0.2)\,\%$
from B $P>5$ GeV/c	$(89.9 \pm 0.2)\,\%$
from long-lived $P>5$ GeV/c	$(88.6 \pm 0.2)\,\%$
from B <5 GeV/c	$(60.3 \pm 0.7)\,\%$
from long-lived $P<5$ GeV/c	$(57.1 \pm 0.4)\,\%$
Fake	$(25.6 \pm 0.1)\,\%$
Fake (evt.avg)	$(13.5 \pm 0.1)\,\%$
Avg. Timing [ms/evt]	90.5
Max Timing [ms/evt]	3208

GENERAL REMARKS

The `Hybrid Seeding Algorithm` is an evolution of a previous version (TDR). Both of them are based on a *tracking in projection* approach:

- Build tracks starting from *two-hit* combination from T1-x and T3-x , find T2-x (*three-hit* combination) and then complete the track on *x-z plane*.
- Add the *u/v-hits* with an *Hough-Like* transformation.

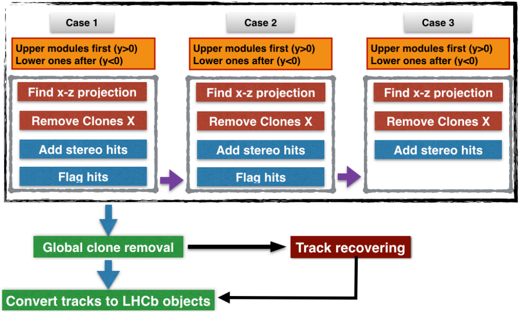
TDR seeding

- Simplified track model (parabola in *x-z*);
- All tracks looked in a single step;
- Span the full fiber length in y when looking to *u/v-layers*;
- Fixed tolerances and search windows;
- No "compensation" between *x-layers* and *u/v-layers*;

Hybrid seeding

- Track model updated to better describe impact of the $\overrightarrow{B}$ field.
- Track search divided in 3 `Cases` with a progressive tracking environment cleaning.
- *In-situ y-segmentation* of the detector: tracks with less number of hits are expected to be found mainly in the internal region.
- Variable tolerances and search windows.
- "compensation" between *x-layers* and *u/v-layers* (6+4, 5+5, 4+6) to recover hit inefficiencies.

SKELETON



Case 1

Upper modules first (y>0)
Lower ones after (y<0)

Find x-z projection

Remove Clones X

Add stereo hits

Flag hits

Case 2

Upper modules first (y>0)
Lower ones after (y<0)

Find x-z projection

Remove Clones X

Add stereo hits

Flag hits

Case 3

Upper modules first (y>0)
Lower ones after (y<0)

Find x-z projection

Remove Clones X

Add stereo hits

Global clone removal

Track recovering

Convert tracks to LHCb objects

Case 1	$P > 5$ GeV/c
Case 2	$P > 2$ GeV/c
Case 3	$P > 1$ GeV/c

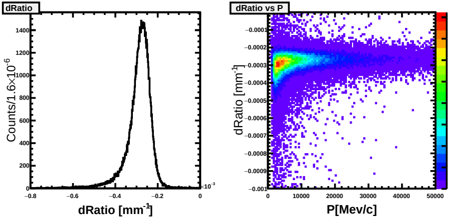
- Case structure used to “recover” for hit inefficiencies.
- Hit-flagging together with Case structure helps in looking for lower P tracks with a cleaner tracking environment.
- Cases are tuned depending on the momentum looked for.
- Main idea: first find ”easy” tracks, then look to harder ones with a cleaner environment (it helps also for timing).

TRACK MODEL I

Take into account the dependence of B_y on z at the first order ($dz = z - z_{ref}$):

$$
\begin{array}{rcl}
B_y(z) & \simeq & B_0 + B_1 \cdot dz \\
\rightarrow \quad x(z) & = & x_0 + t_x \cdot dz + \frac{q}{p}\left(B_0 \cdot dz^2/2 + B_1 \cdot dz^3/6\right) \\
 & = & x_0 + t_x \cdot dz + \frac{q}{p}\frac{B_0}{2} \cdot dz^2\left(1 + d_{ratio} \cdot dz\right) \\
 & = & a_x + t_x \cdot dz + c_x \cdot (dz^2) \cdot (1 + d_{ratio} \cdot dz) \\
y(z) & = & a_y + t_y \cdot dz
\end{array}
\qquad (1)
$$

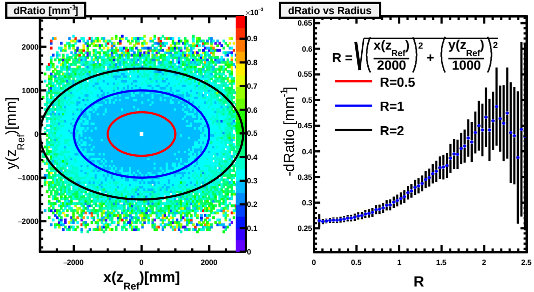
So, $d_{ratio} = B_1/3B_0$ which is roughly constant in the central region. From Monte-Carlo studies fitting reconstructible tracks for a cubic track model:



TRACK MODEL II

However, the assumption $\frac{B_1}{B_0} \sim Const$ can be further improved for the external regions once the information in y is known:

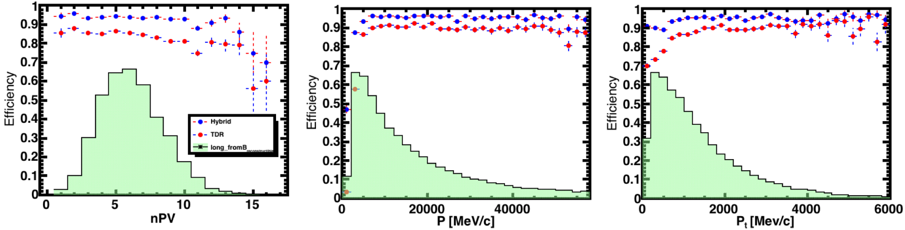
$$\begin{array}{ccccc} B_1 & = & B_1(x,y) & = & F(x,y) \\ B_0 & = & B_0(x,y) & = & F'(x,y) \\ d_{ratio} & \propto & \frac{F(x,y)}{F'(x,y)} & \propto & f(x,y) \end{array} \tag{2}$$



PERFORMANCES

Track Type	TDR seeding	Hybrid Seeding	Δ
from B	$(84.3 \pm 0.2)\,\%$	$(93.6 \pm 0.2)\,\%$	**+ 9.3 %**
from long-lived	$(73.4 \pm 0.2)\,\%$	$(89.8 \pm 0.2)\,\%$	**+ 16.4 %**
from B $P > 5\,\mathrm{GeV}/c$	$(88.2 \pm 0.1)\,\%$	$(95.7 \pm 0.1)\,\%$	**+ 7.5 %**
from long-lived $P > 5\,\mathrm{GeV}/c$	$(88.6 \pm 0.2)\,\%$	$(95.4 \pm 0.2)\,\%$	**+ 6.8 %**
from B $P < 5\,\mathrm{GeV}/c$	$(60.3 \pm 0.7)\,\%$	$(84.7 \pm 0.5)\,\%$	**+ 24.4 %**
from long-lived $P < 5\,\mathrm{GeV}/c$	$(57.1 \pm 0.4)\,\%$	$(83.8 \pm 0.3)\,\%$	**+ 26.7 %**
Fake	$(25.6 \pm 0.1)\,\%$	$(11.8 \pm 0.1)\,\%$	**- 54 %**
Fake (evt.avg)	$(13.5 \pm 0.1)\,\%$	$(6.5 \pm 0.1)\,\%$	**- 52 %**
Avg. Timing [ms/evt]	90.5	34.5	**∼3 times faster**
Max Timing [ms/evt]	3208	600	**∼5 times faster**

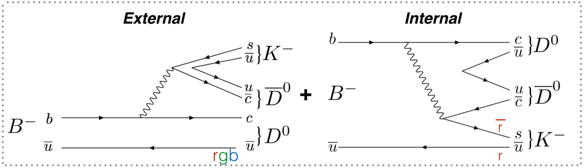
- More efficient
- Less fake tracks
- Faster

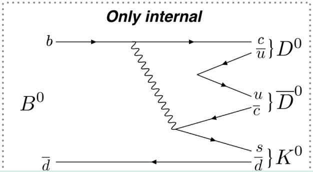


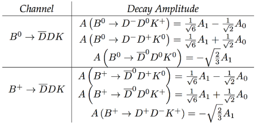
DOUBLE CHARM B HADRON DECAYS: $B^0 \to D^0\overline{D}^0K^{*0}$

1. $B^0 \to D^0\overline{D}^0K^{*0}$ has never been observed so far.
2. $\mathcal{B.R.}$ measurement is the first goal, reference channel $B^+ \to D^0\overline{D}^0K^+$ to control systematics.
   Expected events $\sim$ 100.

- Characterized by $b \to c(W^- \to \overline{c}s)$, which is an isospin conserving transition:
- If spectator quark doesn't play any role: isospin relations hold $\Rightarrow$ build relations of $\mathcal{B.R.}$:
  1. DK sub-system can have $I = 0$ (A_0, color favored) or $I = 1$ (A_1, color suppressed).
  2. Single measurement of B.R. and test of these relations can give insight into the **decay mechanism**: strength of FSI, QCD factorization of amplitudes in three body decays.



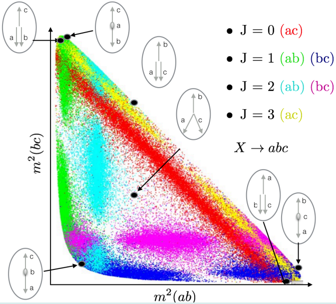




Channel	*Decay Amplitude*
$B^0 \to \overline{D}DK$	$A\left(B^0 \to D^-D^0K^+\right) = \frac{1}{\sqrt{6}}A_1 - \frac{1}{\sqrt{2}}A_0$ $A\left(B^0 \to D^-D^+K^0\right) = \frac{1}{\sqrt{6}}A_1 + \frac{1}{\sqrt{2}}A_0$ $A\left(B^0 \to \overline{D}^0D^0K^0\right) = -\sqrt{\frac{2}{3}}A_1$
$B^+ \to \overline{D}DK$	$A\left(B^+ \to \overline{D}^0D^+K^0\right) = \frac{1}{\sqrt{6}}A_1 - \frac{1}{\sqrt{2}}A_0$ $A\left(B^+ \to \overline{D}^0D^0K^+\right) = \frac{1}{\sqrt{6}}A_1 + \frac{1}{\sqrt{2}}A_0$ $A\left(B^+ \to D^+D^-K^+\right) = -\sqrt{\frac{2}{3}}A_1$

DOUBLE CHARM B HADRON DECAYS: $B^0 \to D^0\overline{D}^0 K^{*0}$

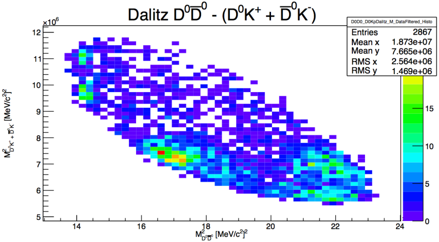
1. In $W^+ \to c\bar{s}$, $c\bar{s}$ do not "feel" the other quarks $\Rightarrow$ high chance to have $D_s^{(*)}$ final states, this requires at least one between D and K to be charged.
2. $D^0\overline{D}^0 K^{*0}$ all neutral. If any $D^0 K^{*0}$ resonance is observed, clear exotic (4-quark state).
3. If nothing, this channel is the best place to study $D^0\overline{D}^0$ resonant structures.
4. It's a $P \to PPV$ transition, so, many more quantum numbers are allowed for the PP $(D^0\overline{D}^0)$ pair.
5. Dalitz analysis needs to fit a multi-D space using coherent sum of amplitudes $\Rightarrow$ interferences can fake or hide peaking structures.



- $J = 0$ (ac)
- $J = 1$ (ab) (bc)
- $J = 2$ (ab) (bc)
- $J = 3$ (ac)

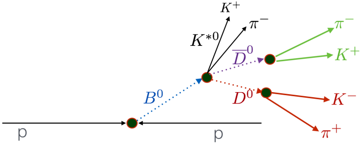
$X \to abc$

Preliminary plot in $D^0\overline{D}^0 K^+$



ANALYSIS STRATEGY

Decay topology



Selection

- K^{*0} decays strongly, necessity to separate $K\pi$ daughters from D^0 and $\overline{D}^0$ decay vertex.
- *Particle Identification*(PID) requirements are essential.
- Correct Monte Carlo for discrepancies (strong in PID) to properly evaluate selection efficiency.

SELECTION

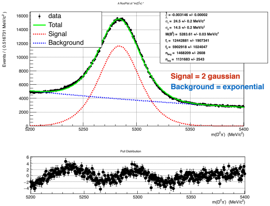
- Pre-selection of events, loose PID cuts to all final states.
- 2 step MVA selection:
  1. **D mesons selection D_{fromB}** :
     - D_{fromB} selection trained on S-weighted $3fb^{-1}$ data of $B^+ \to D^0\pi^+$.
     - D_{fromB} selection trained on MC $B^0 \to D^0\overline{D}^0K^{*0}$.
  2. **B meson and K^{*0} MVA selection using first stage output classifier**.
     - 2.1 Including/excluding PID in training.
     - 2.2 Different cut optimization in different scenarios.
  3. **Apply same selections to reference channel** $B^+ \to D^0\overline{D}^0K^+$

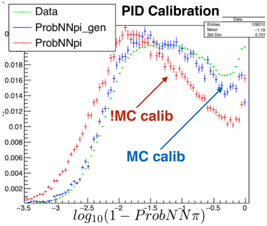
D_{fromB} MVA SELECTION (DATA DRIVEN)

Input variables

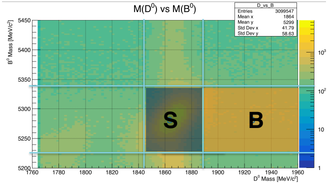
- Admixture of D^0 topological and kinematical variables (independent from B mass)
- π and K kinematic and topological variables.
- Adding (D_{fromB}^{PID}) or excluding (D_{fromB}^{NOPID}) π and K PID variables in training.

S-weight




PID Calibration

Training



- For MC based method just use the MC truth match as signal and background from Data defined as in Data driven method. $D_{fromB}^{NOPID,MC}$.

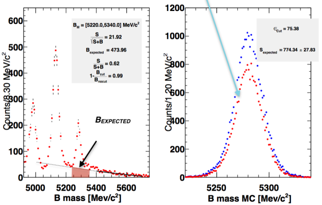
SECOND STAGE MVA SELECTION

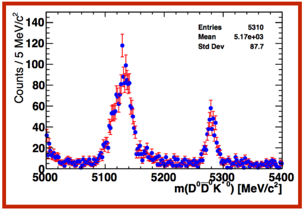
- Use output of D_{fromB} as training variable.
- Also kinematic and topological of B, D and K^{*0} mesons.
- Upper B mass sideband as background for training.
- Adding or excluding K from K^* PID variables.

Significance optimization achieved cutting on final classifier output together with PID variables (if was not used in MVA).

$$C = 2\cdot\int \mathcal{L}\cdot\sigma_{b\bar{b}}\cdot f_{u(d)}\cdot\mathcal{B.R.}(D \to DDh)\cdot(\mathcal{B.R.}(D \to K\pi))^2\cdot\epsilon_{Gen}\cdot\epsilon_{Presele}\cdot\epsilon_{Stripping}\cdot\epsilon_{Trigger}\cdot\epsilon_{Tracking}$$

$$S_{Expected} = \mathrm{C}\cdot\boxed{\epsilon_{Cut}}$$





DfromB type	2^{nd}*stageBDT*	Cut optimization	Max $\frac{S}{\sqrt{S+B}}$	$S_{expected}$	ϵ_{MC}	Purity ($\frac{S}{S+B}$)	Bkg. rejection
NOPID	no *ProbNNk* $K_{K^{*0}}$	4-D	19.31	745	72.6 %	50%	98%
	+ *ProbNNk* $K_{K^{*0}}$	3-D	19.65	782	76.2 %	49%	98%
NOPID$_{MC}$	no *ProbNNk* $K_{K^{*0}}$	4-D	24.09	705	68.6 %	82%	99%
	+ *ProbNNk* $K_{K^{*0}}$	3-D	24.58	756	73.6 %	80%	99%

CONCLUSIONS

TRACKING FOR UPGRADE

- New detector needs new software.
- We understand the correct way to handle the SciFi detector in pattern recognition.
- Huge gain in performance in all fields.
- Currently used as default for the upgrade physics simulations.
- Still new ideas to try out to further improve performance.

Analysis

- Double charm B decays covers a relatively wide variety of *HEP* topics.
- In particular, $B^0 \rightarrow D^0\overline{D}^0K^{*0}$ (never observed) provides a very clean environment to test different physics.
- Work ongoing to finalize the selection and move to mass fit.
- Once $\mathcal{B}.\mathcal{R}.$ done, may worth to try Dalitz analysis, in such case could be interesting to do a 4 body Dalitz analysis ($B^0 \rightarrow D^0\overline{D}^0K^+\pi^-$).