SPS/ÖPG Annual Joint Meeting 2021

FAKT session
31th August 2021

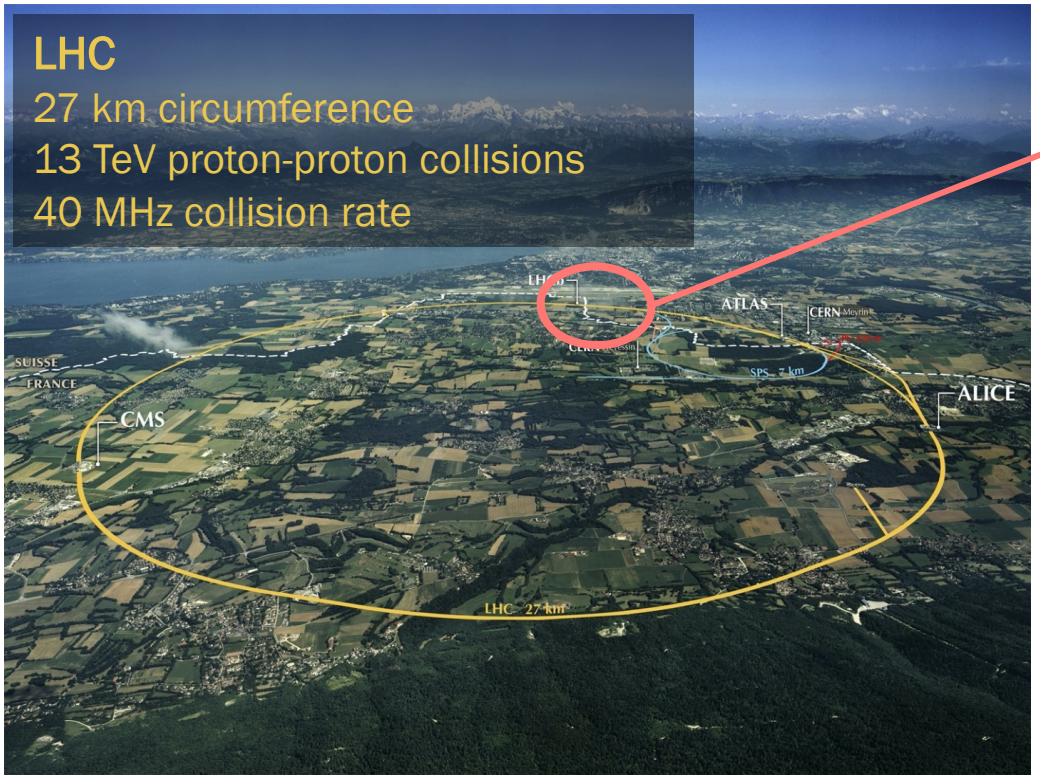
Search for the multilepton final state @LHCb
$B^+ \to K^+ J/\psi \gamma^*(\to \mu^+\mu^-)$

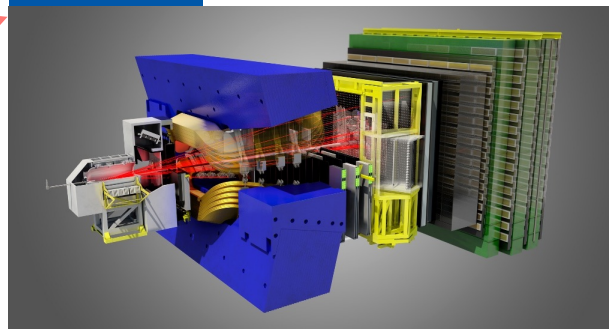
Sonia Bouchiba, Vitalii Lisovskyi, Lesya Shchutska

Large Hadron Collider @ CERN

Construction of knowledge on elementary particles and their interaction
= constant interconnection between theory buildings and experimental observations

Observing new phenomena is a nice goal... but one must first make them happen!





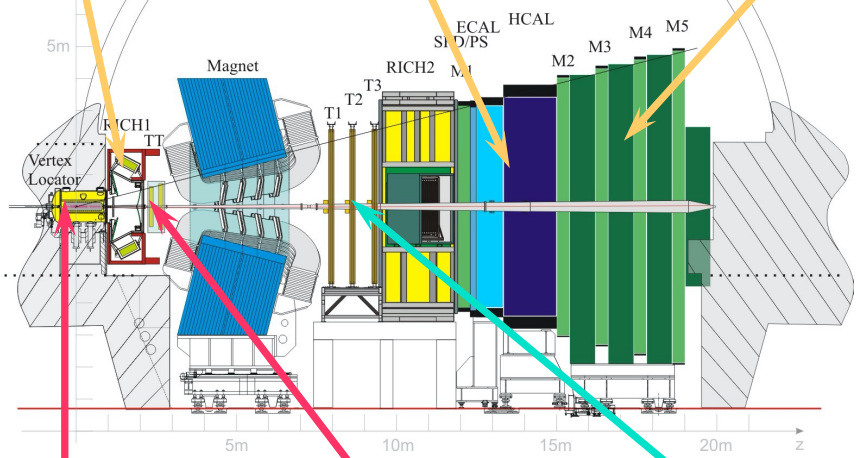
LHCb experiment

UPGRADED

RICH
Replace HPD and R/O

Calo PMTs
Replace R/O

Muons MWPC
almost compatible



VELO
Micro Si strips → pixels

UT Si strips
→ replaced

Tracking stations
Si strips/gas tube
→ SciFi

REPLACED

Largely developed @ EPFL!

A *B*-HADRON SPECIALISED DETECTOR

- Single-arm, forward spectrometer, $2<\eta<5$
- $b\bar{b}$ production peaks (back-)forward
- Aiming for $\sqrt{s} \gtrsim$ **13 TeV** for 3rd data taking run (*planned* 2022-2024)

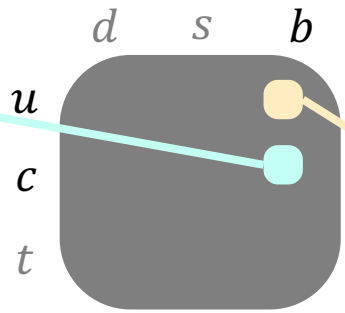
MAJOR UPGRADE FOR RUN 3 (2022)

- Event readout rate ×40
- Upgrade of ~all detectors
- Full software trigger, with increased computing power

Ok... but why?

- **FLAVOUR ANOMALIES IN SEMILEPTONIC B-DECAYS**

$b \to c\ell\nu_\ell$ shows anomaly @ 3σ using yields of τ vs. light leptons (μ, e) [HFLAV summary, 2019]



$b \to u\ell\nu_\ell$ can be probed in $B^+ \to \ell\nu_\ell$ decays

Only measurement [Belle Collab,2004]

$$\frac{\mathrm{BR}(B^+ \to \tau\nu_\tau)\mathrm{exp}}{\mathrm{BR}(B^+ \to \tau\nu_\tau)\mathrm{SM}} = 1.31 \pm 0.27 \ @1.5\sigma$$

- **HOW CAN LHCb CONTRIBUTE?** $B^+ \to \ell\nu_\ell$ is difficult or impossible at LHCb due to single-track signature and helicity suppression ☹

  However, adding an **Final State Radiation (FSR) tagging** with $\gamma^* \to \mu^+\mu^-$ allows to:

  - increase BF to available dataset (~10^{-8}) [Bharucha et al., 2021]
  - reconstruct a decay vertex of $B^+ \to \mu^+\nu_\mu\gamma^*(\to \mu^+\mu^-)$ and $B^+ \to e^+\nu_e\gamma^*(\to \mu^+\mu^-)$ **@LHCb** and probe Lepton Flavour Universality by comparing their yields

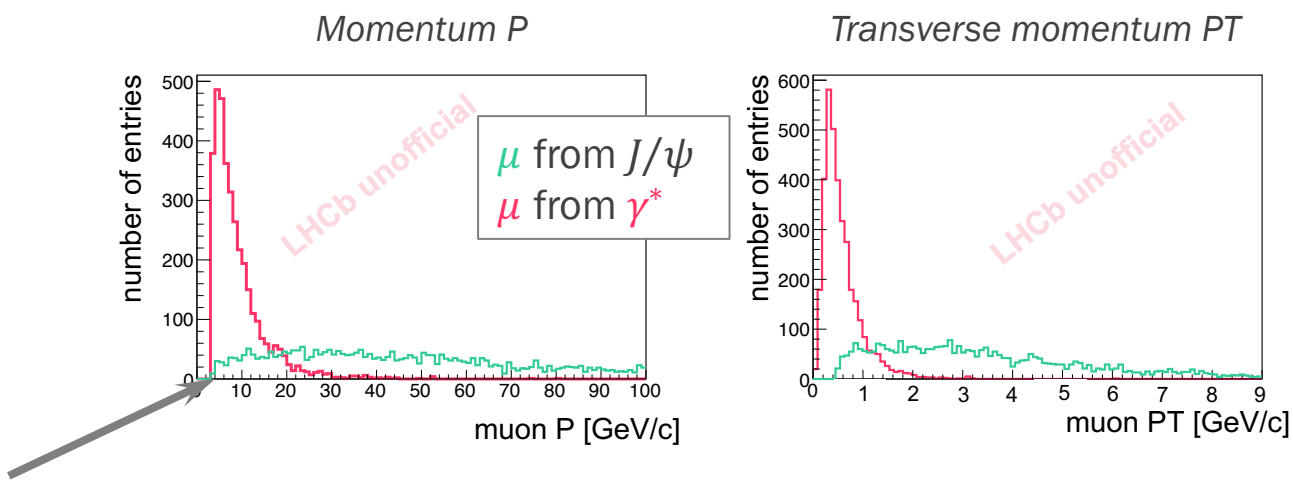
Experimental challenge 💪 **→ use of *very soft muons***

Experimental challenge 💪 ⟶ use of *very soft muons*

Goal: Develop a dedicated treatment of soft muon selection and identification with a reference channel $B^+ \to K^+ J/\psi\, \gamma^*(\to \mu^+\mu^-)$

Benefits:

- visible with available dataset (though have never been observed so far)
- usage of $J/\psi \to \mu^+\mu^-$ data stream allows to get an unbiased sample $\gamma^* \to \mu^+\mu^-$



- imposes minimum P of 3 GeV on soft muons ⟶ make sure they reach the muon stations and can be identified

The challenge: dealing with different backgrounds!

MISIDENTIFIED MUONS ⟶ 1 or 2 hadrons (K or π) decaying to muons, tricking the muon stations, are misID as muons

Dominant background
To be studied with a dedicated Boosted Decision Tree (BDT) only with Particle Identification (PID) vars

RESONANT DECAYS ⟶ $\omega, \phi, \rho \to \mu^+\mu^-$ with $\mathcal{B} \sim 10^{-8}$ and $\chi_{c1}, \chi_{c2} \to J/\psi\, \mu^+\mu^-$ contributions;

Peaking backgrounds
(dangerous!) Add in fit model / veto

COMBINATORIAL BACKGROUND ⟶ Random tracks put together by mistake by the selection algos

Exponential distribution
Supress with dedicated BDT

PARTIALLY RECONSTRUCTED DECAYS ⟶ Non-signal decays with an extra particle not entering the selection (e.g. $B^+ \to J/\psi\ K^+\eta(\to\ \mu^+\mu^+\gamma)$, with missed γ)

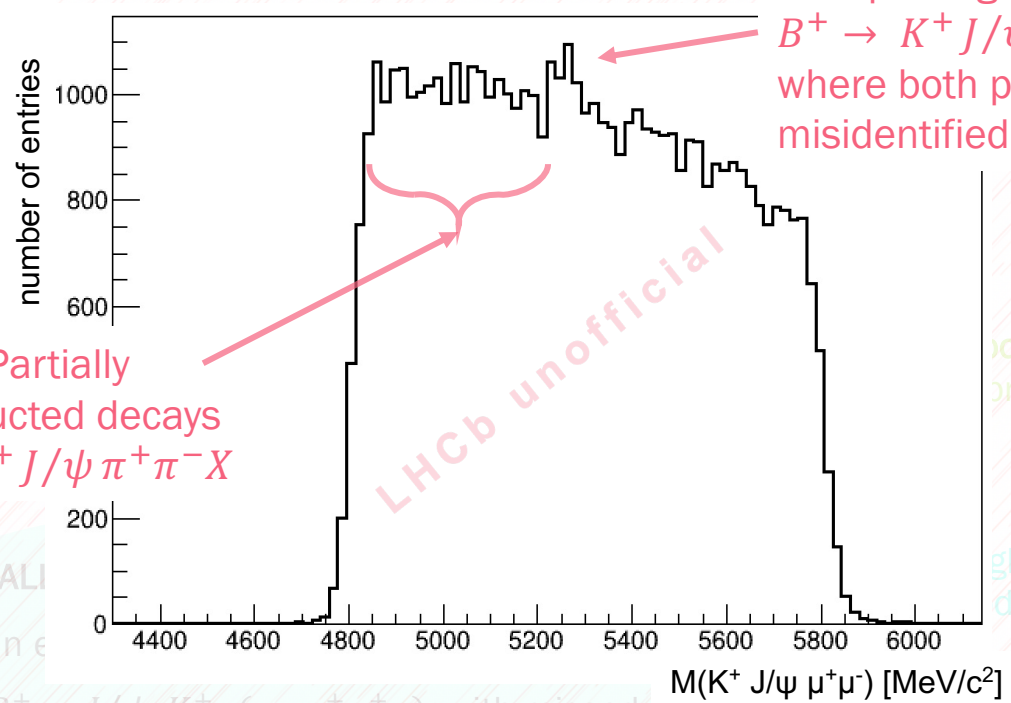
Neglect / add in fit model

The challenge: dealing with different backgrounds!

First look at a small subset of data with the preselection applied :

Double MisID peak on top of signal ☹
$B^+ \rightarrow K^+ J/\psi \pi^+\pi^-$
where both pions are misidentified as muons

MisID + Partially reconstructed decays
$B^+ \rightarrow K^+ J/\psi \pi^+\pi^- X$



Combinatorial background Boosted Decision Tree (BDT)

Goal: have a very efficient selection with ~100% signal efficiency

TRAINING VARIABLES

B-meson's (mother)

- impact parameter
- direction angle of flight
- quality of the constrained track fit
- position of decay (secondary vertex)

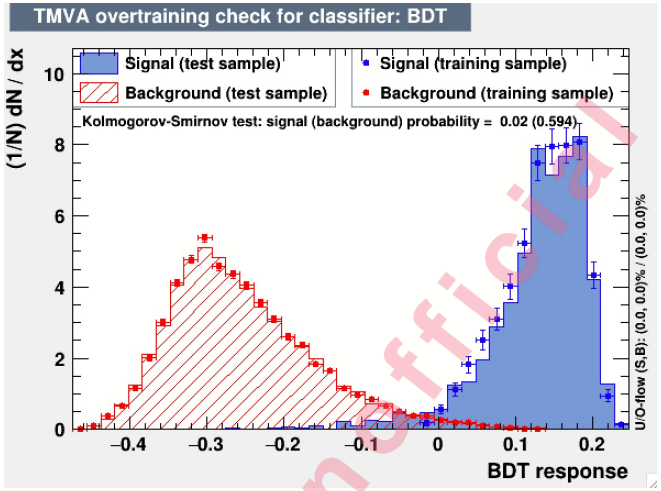
J/ψ flight distance

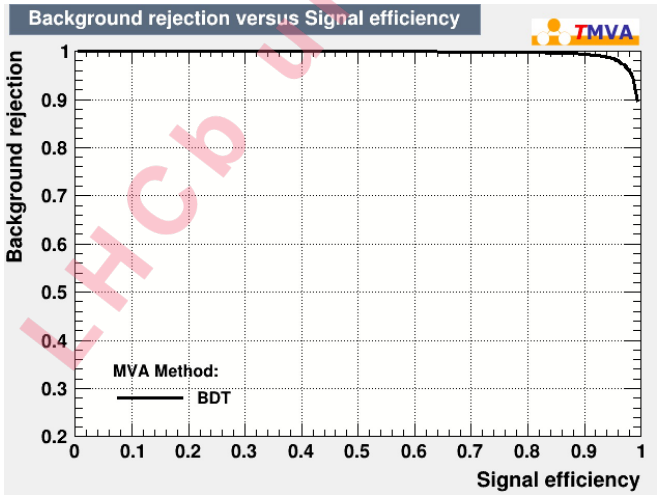
Minimum transverse momentum of J/ψ muons

USED SAMPLES

- **Background:** inclusive data sample 2018 (Magnet Down), high mass sideband (M(B)>5.4 GeV)

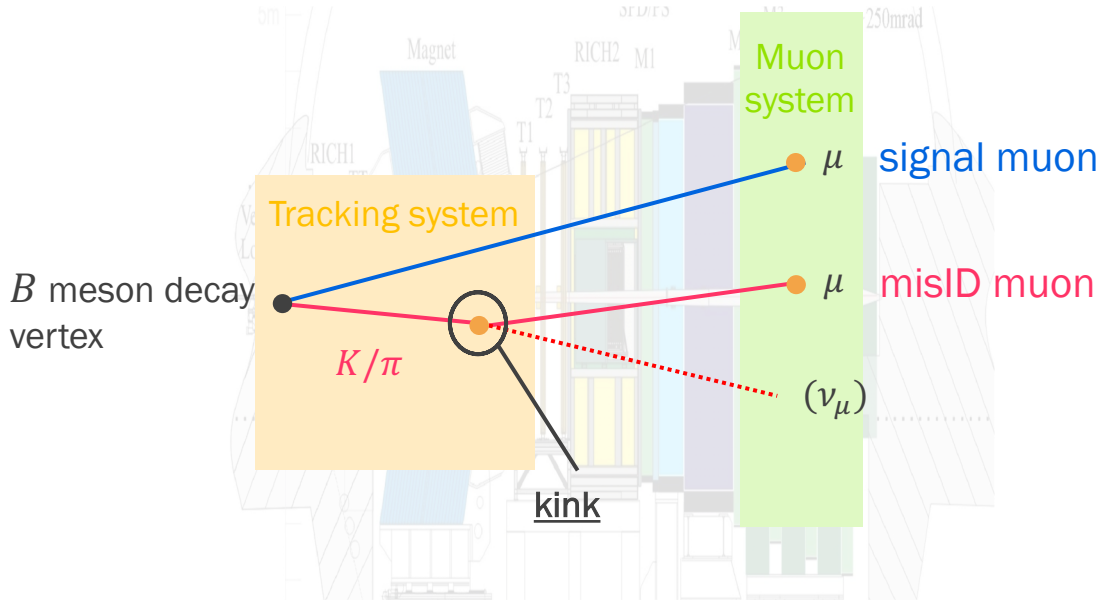
- **Signal:** MC simulation 2018 (Magnet Down), $B \to (\chi_{c1} \to J/\psi\, \mu^+\mu^-)K$ with normal χ_{c1} mass distribution Cut efficiencies





Rejecting misID soft dimuon background

The situation:



Learn with simulation: $D^0 \to K^- \pi^+ (\to \mu^+ \nu_\mu)$, with reconstructed K^- and μ^+, and designed to access to generator level truth!

2 "types" of features
- muon **identification** parameters
- track quality (kink finder) to discard $\pi \to \mu\nu_\mu$ decays in flight

MisID soft muon BDT

Let's mix all this!

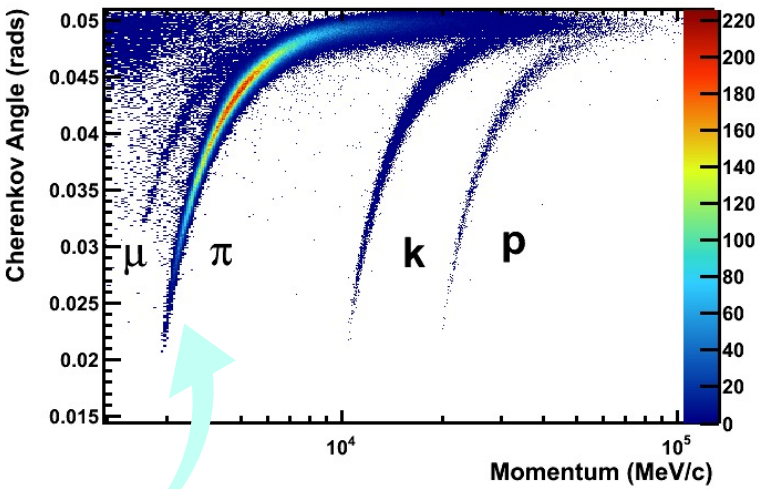
TRAINING VARIABLES (single soft muon only)

PID variables

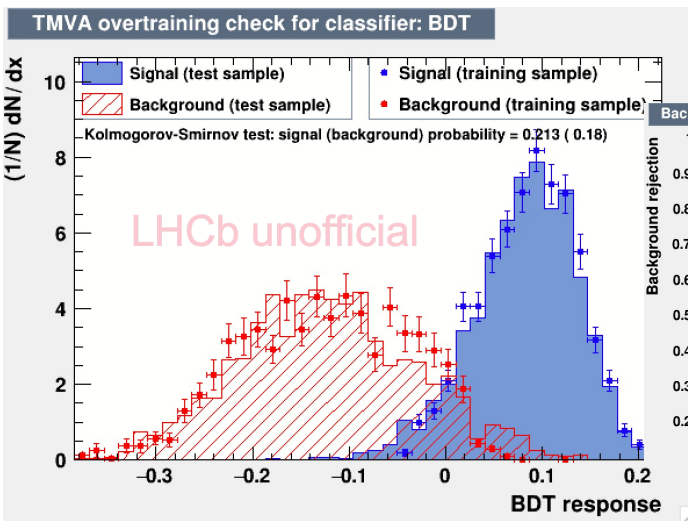
- Global NN-computed muon/kaon probability
- Global pion likelihood:
  PID(X) = log(Likelihood(X)) - log(Likelihood(pi))
- RICH pion likelihood

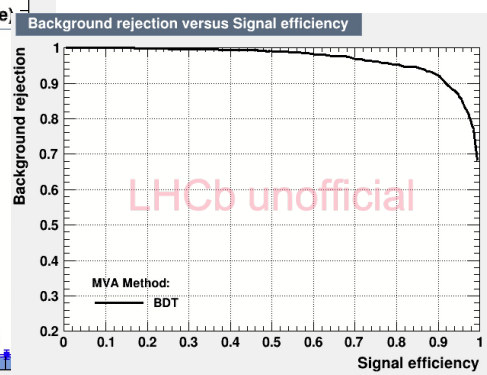
Kink-tracker

+kinematics to add dependency of the PID in the soft muon momentum



from [LHCb Rich Collab., 2012]





SAMPLES

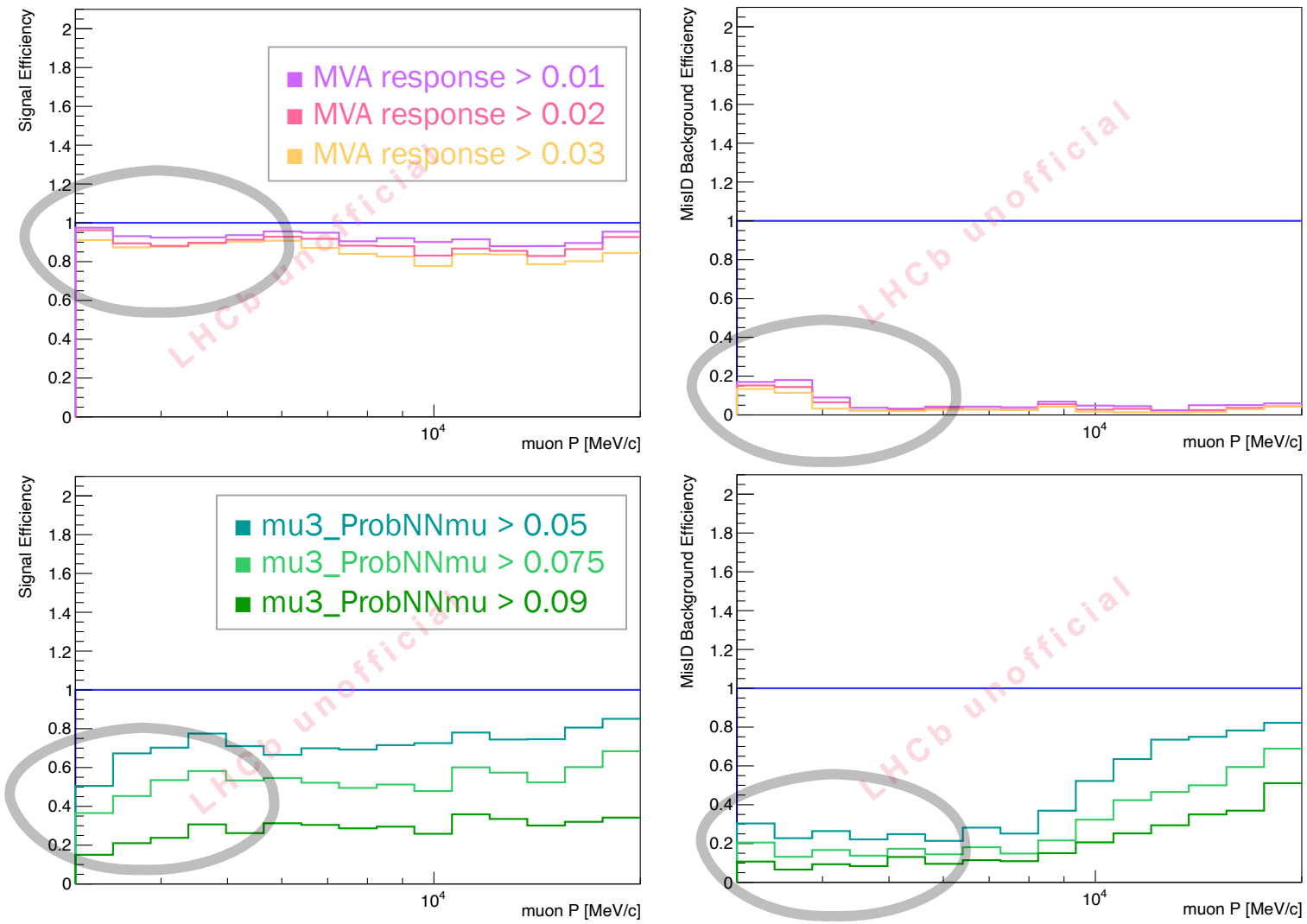
- **Background:** $D^0 \to K\pi(\to \mu\nu)$ sample weighted in momentum and in transverse momentum sample wrt simulation
- **Signal:** MC simulation 2018 MD, $B \to (\chi_{c1} \to J/\psi\, \mu^+\mu^-)K$ with normal χ_{c1} mass distribution

Is our method more powerful than a standard PID cut?

↓ Signal efficiency ↓ ↓ MisID background efficiency ↓

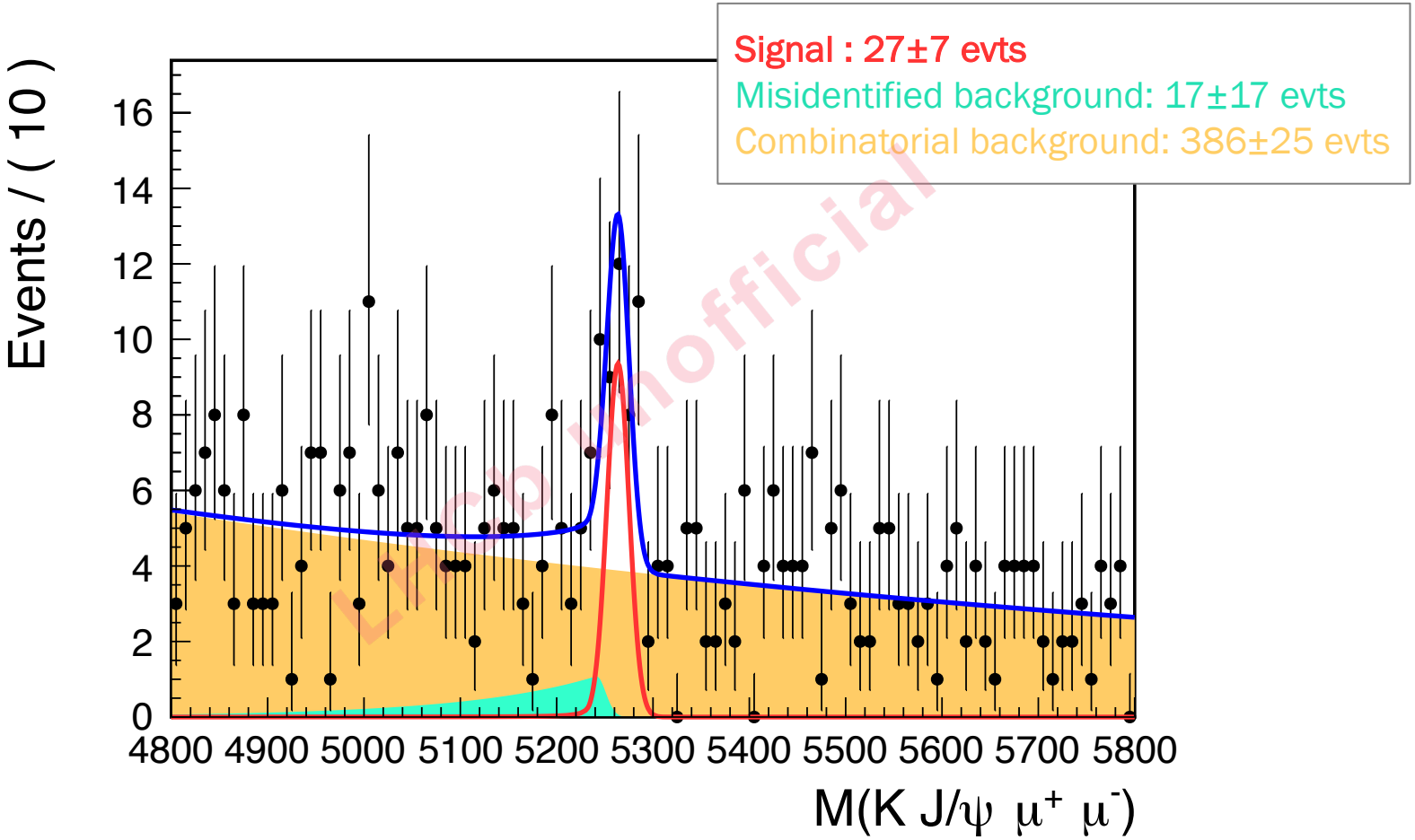
↓ soft muon BDT selection ↓

↓ standard muon ID algo ↓



Cuts applied on data sample

Small data sample: inclusive J/ψ data sample.



🏡 Take-home messages

- The decay channel $B^+ \to K^+ J/\psi\, \gamma^*(\to \mu^+\mu^-)$ is currently searched for at LHCb

- Particular attention is paid to soft muons background rejection.

- A dedicated momentum-dependent BDT is trained with specific variables.

  ☞*It shows a **better rejection of decays in flight** in background compared to the standard muon identification algorithm* ☺

- This study will be extended to FSR muons used in **other channels**, in order to probe Leptonic Flavour Universality in $B^+ \to \ell\nu_\ell$ decays without suffering of helicity suppression

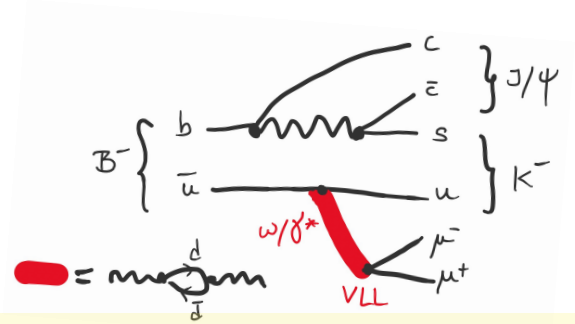
- Results will be used to improve data-taking efficiency for soft muons

Thank you!

Backup

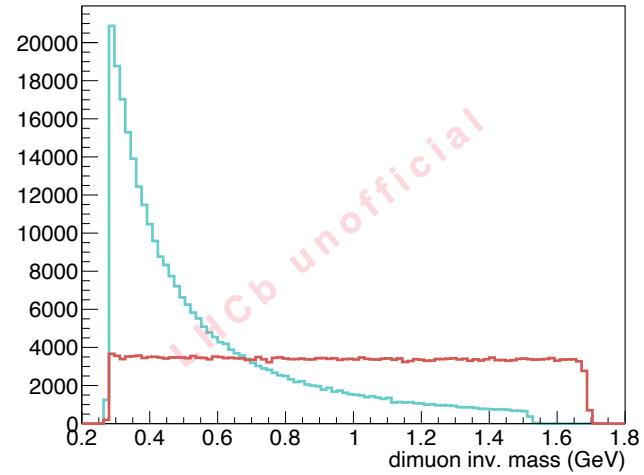
How to simulate $\gamma^* \to \mu^+\mu^-$ signal?

- No decay model in EvtGen for such γ^*
- **Assumption:** dimuon coming from **Initial State Radiation (ISR)** (choice among other possible processes)
- Identify γ^* as $\boldsymbol{\omega}$ and use model "vector to 2 leptons" (VLL)



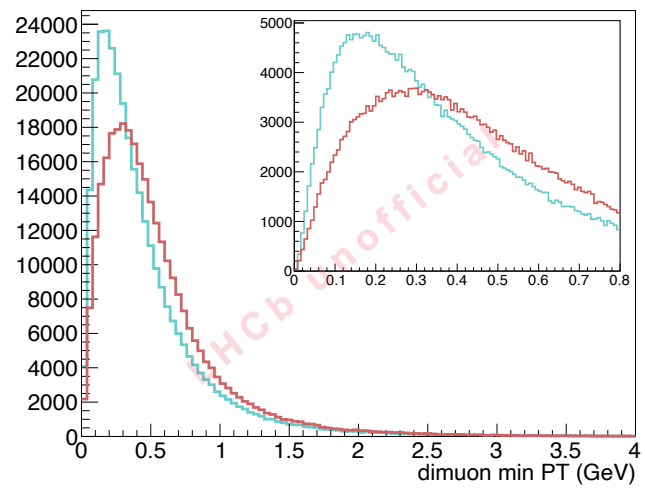
$$B^+ \xrightarrow{PHSP} K^+ \left(J/\psi \xrightarrow{VLL} \mu^+\mu^-\right)\left(\boldsymbol{\omega} \xrightarrow{VLL} \mu^+\mu^-\right)$$

- Fast simulation tool `RapidSim` (global LHCb response)
- Requested full-sim sample (~0.5M) for flat mass distribution, can be reshaped later

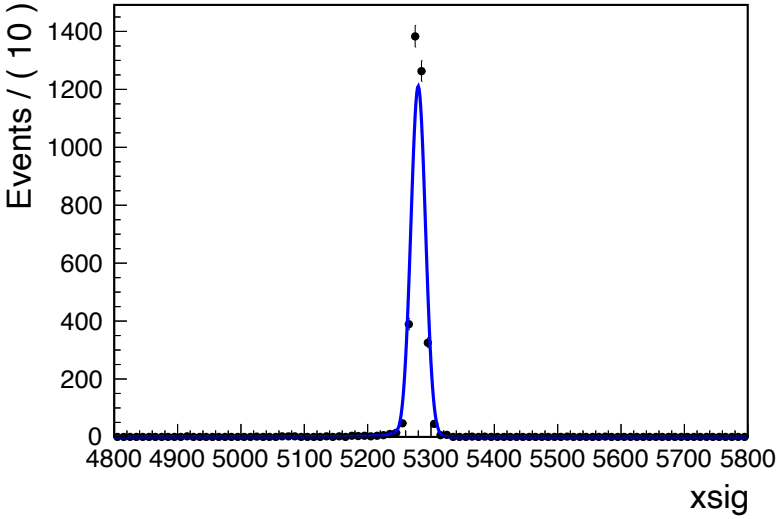


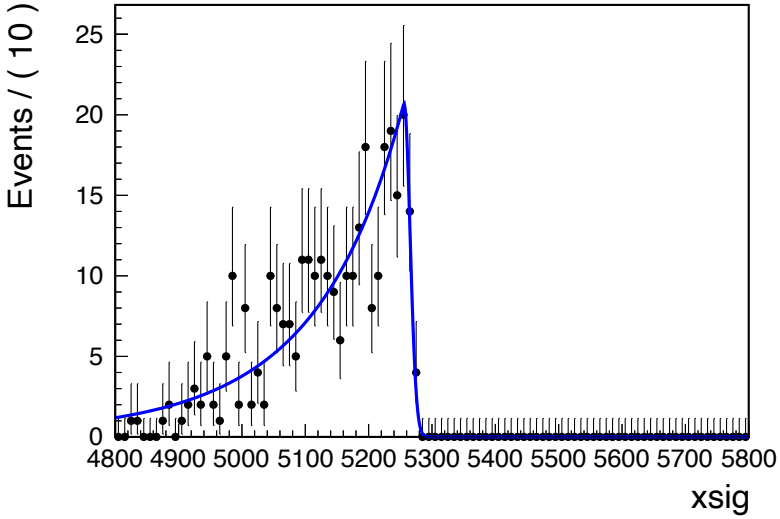
FLAT ω/γ^* mass distribution

WIDE B-W ω/γ^* mass distribution



Data fitting





Simulation sample

$$B \rightarrow (\chi_{c1} \rightarrow J/\psi\, \mu^+\mu^-)K$$

Fitted with a Crystal Ball shape
Tail parameters fixed

MidID muons sample

$$B^+ \rightarrow K^+ J/\psi\, (\pi^+ \rightarrow \mu^+ \upsilon_\mu)(\pi^- \rightarrow \mu^- \upsilon_\mu)$$

Fitted with a Crystal Ball shape
Tail parameters fixed

Preselection details

- Preselection:

B	Kaon	Jpsi (-> mu1 mu2)	Soft muons (mu3, mu4)
TOS L0 Muon **OR** DiMuon decision	PIDK > -2	50 MeV window around Jpsi mass	Soft muons Tracks' GhostProb < 0.3
B_DIRA_OWNPV > 0.9999	Track's GhostProb < 0.3	mu1,2 ProbNNmu > 0.05	

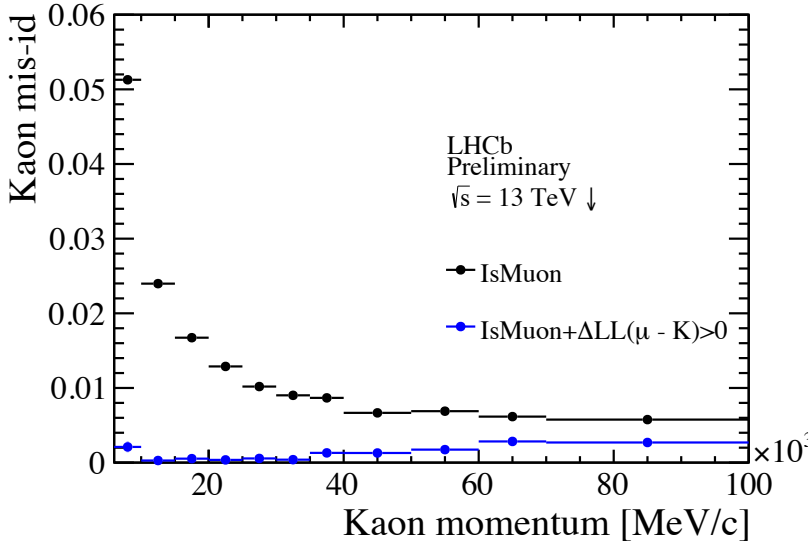
- Overall: require soft muons to both pass "**isMuon**" -> cut @P=3GeV
- (Correlated background cuts)

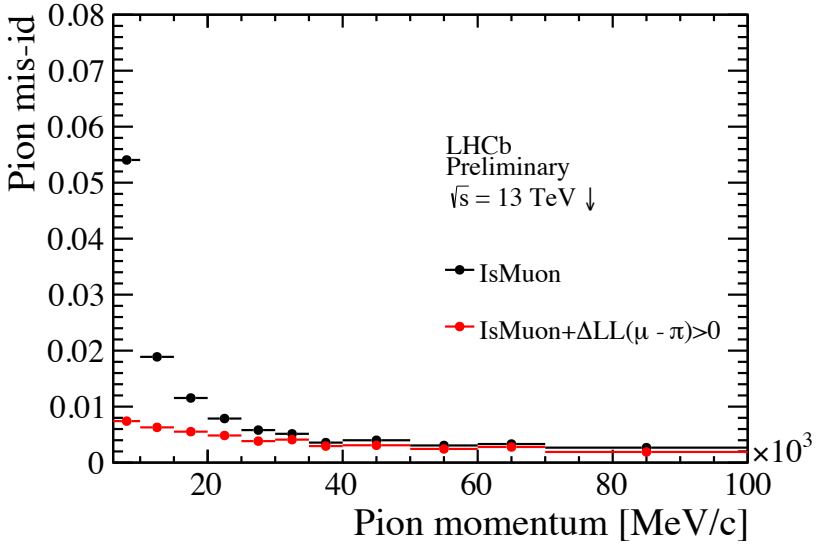
p [GeV/c]	Required stations	
	`IsMuon`	`IsMuonLoose`
$p < 3$	*Always false*	*Always false*
$p < 6$	M2 & M3	At least two of M2–M4
$6 < p < 10$	M2 & M3 & (M4 ‖ M5)	At least three of M2–M5
$p > 10$	M2 & M3 & M4 & M5	At least three of M2–M5

`isMuon` ALGORITHM

- tracks are extrapolated to the **muon stations**
- Requires 2 or 3 hits in the muon stations depending on particle momentum (backup)

↓ **Misidentification probabilities for kaons and pions** → more likely to miss muons at **low momenta**



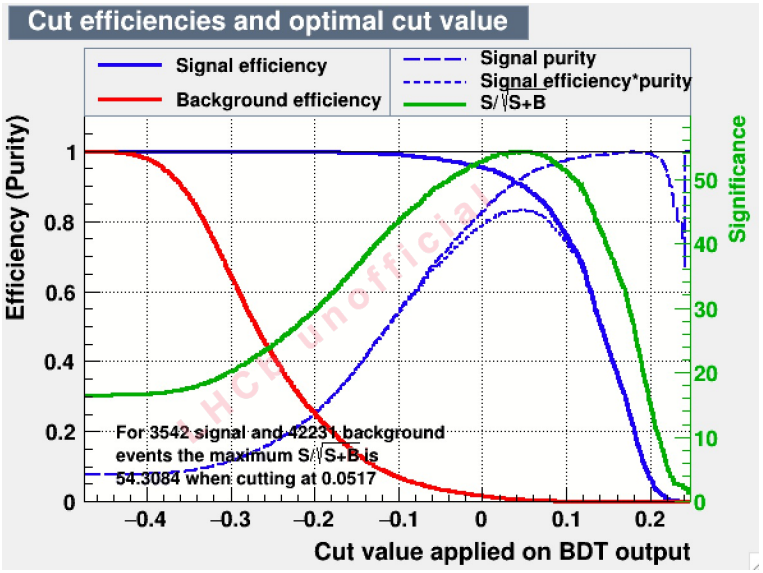


Taken from [LHCb Collab., 2019]

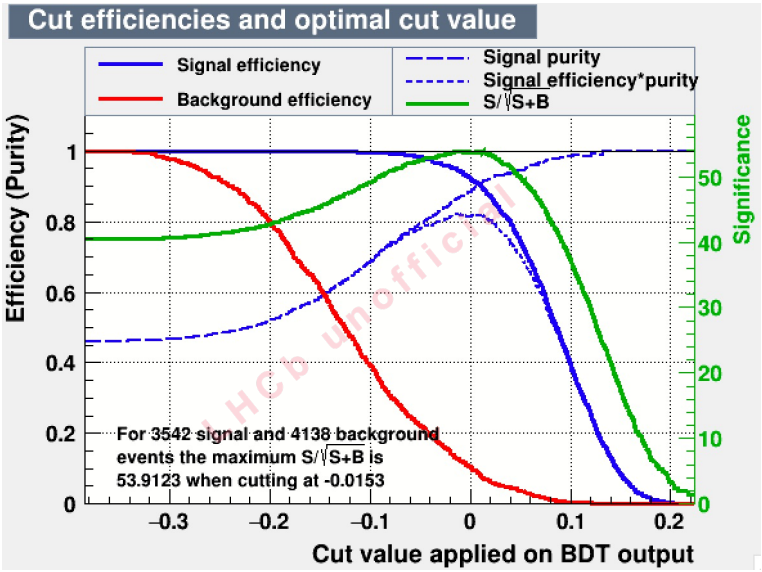
ΔLL(X-Y) = log(Likelihood(X)) - log(Likelihood(pi))
gives the likelihood for the probed particle X to be an Y

BDTs cut efficiencies

Benchmark Combinatorial BDT

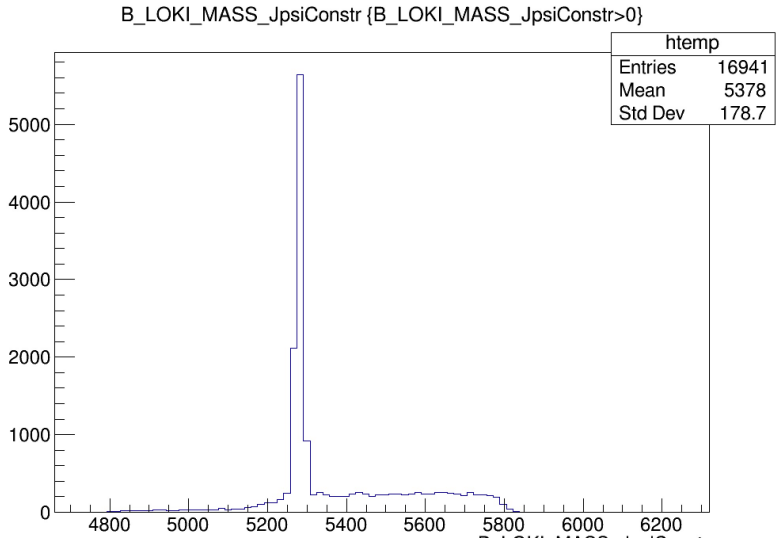


Benchmark soft muon BDT

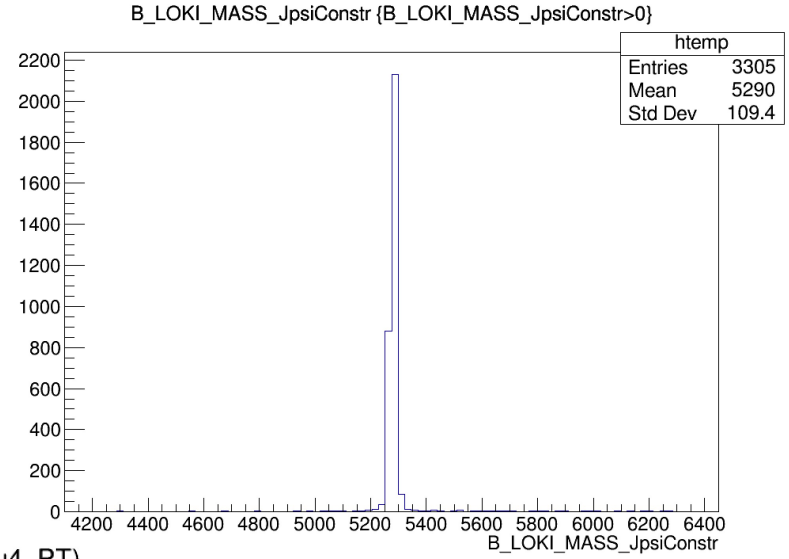


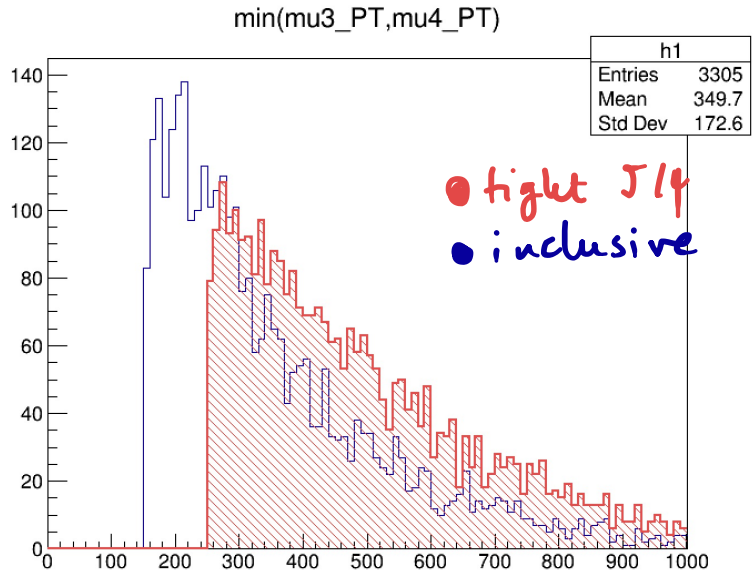
Comparison of stripping lines

inclusive line



tight J/ψ line





Category BDT2 with different conditions on mu3_P

Category in TMVA

- separate the training data (and accordingly the application data) into disjoint sub-populations exhibiting significantly different properties.
- The separation into phase space regions is done by applying requirements on the input and/or spectator variables (in this case mu3_P is spectator)
- In each of these disjoint regions (each event must belong to one and only one region), an **independent** training is performed using the most appropriate MVA method, training options and set of training variables in that zone.
- According to TMVA devs, the division into categories in presence of distinct sub-populations reduces the correlations between the training variables, improves the modelling, and hence increases the classification and regression performance.

Category BDT2 with different conditions on mu3_P

Background: weighted sample of $D^0 \to K\pi(\to \mu\nu)$

Signal: presel. inclusive MC 2018 MD, Evn 12145067 $B^+ \xrightarrow{\text{PHSP}} K^+ (J/\psi \xrightarrow{\text{VLL}} \mu^+\mu^-)(\omega \xrightarrow{\text{VLL}} \mu^+\mu^-)$ w/ flat ω mass distr.

Categories:

Cat 1: mu3_P< 6000

Cat2: mu3_P > 6000 && mu3_P < 12000

Cat3: mu3_P>12000

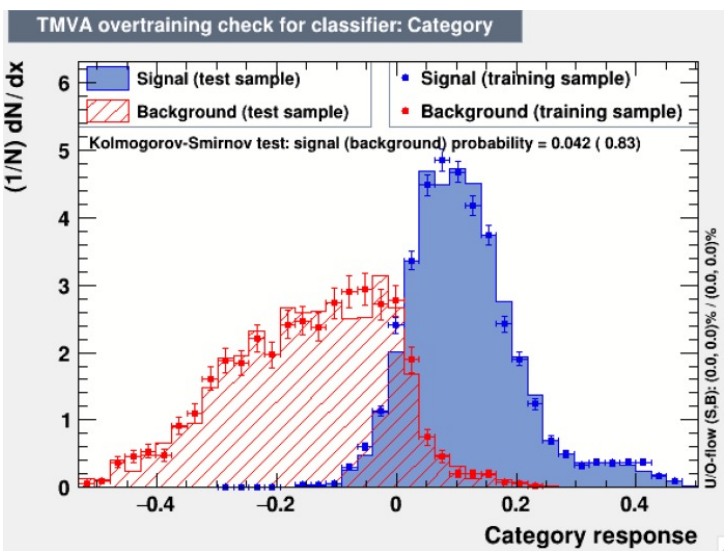
Training Variables (first attempt, will need the $\boldsymbol{var(p, p_T)}$ performance to target more systematically what is best)

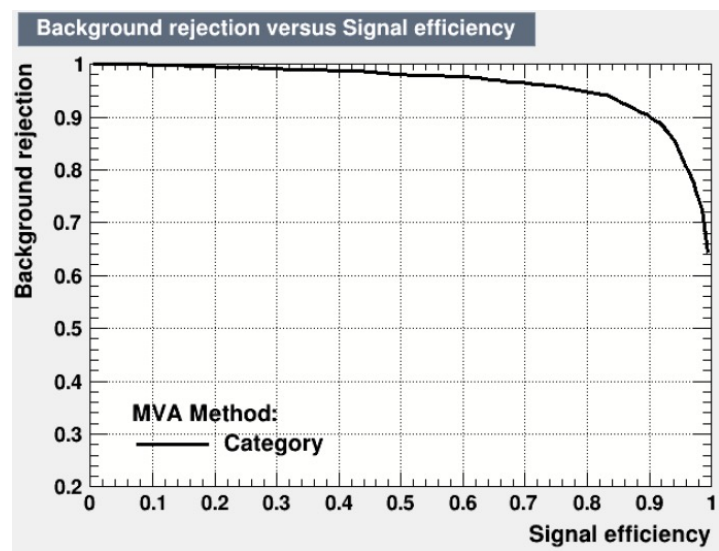
Cat1: mu3_TRACK_CHI2:log10(abs(mu3_ProbNNmu)):log10(abs(mu3_ProbNNk)):mu3_PIDmu:mu3_PIDK

Cat2 = vars of Cat1

Cat3 = vars of Cat1 + mu3_RichDLLmu:mu3_RichDLLk

Results:

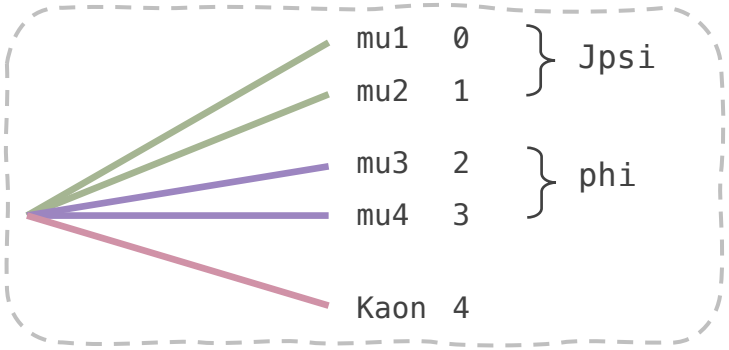




Correlated background suppression
Motivation

Motivation: "combinatorial" background,

- with 5 tracks random (not likely), IPChi2 already gives an information about compatibility (already cut on that)



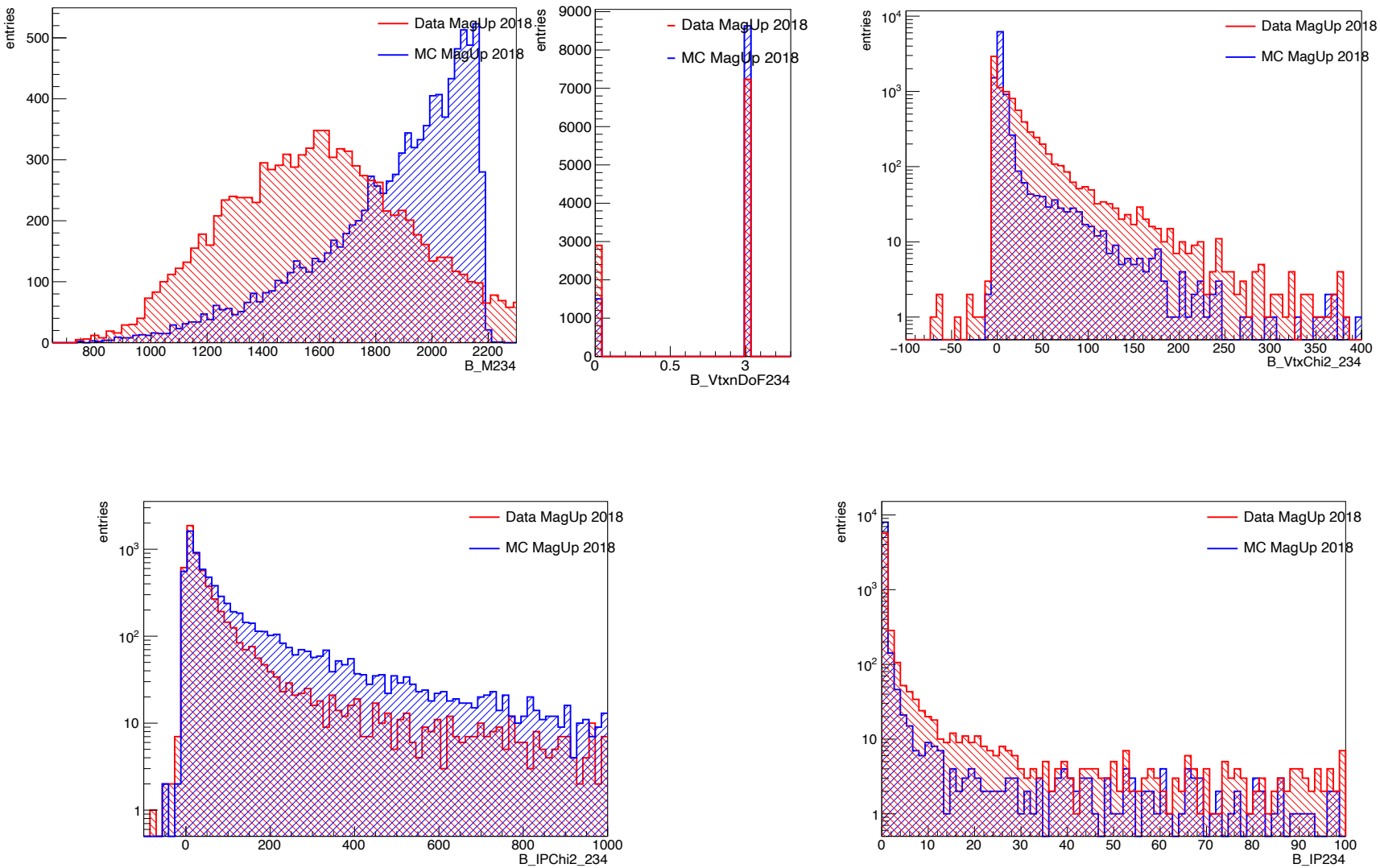
Aim: check quality of [Kmu3mu4] or [Jpsimu3mu4] constrained vertex to discard the candidates with double misID

=> take the sum of the IP chi2 or the **max**

Correlated background suppression

$K - \mu_3\mu_4$ correlations [234]

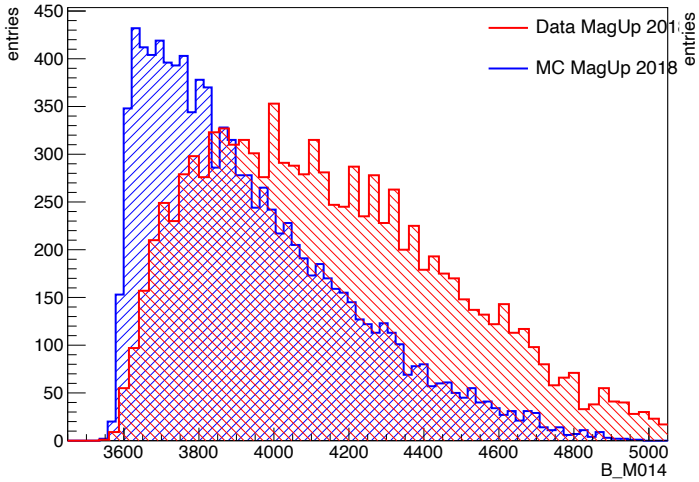
Temporary

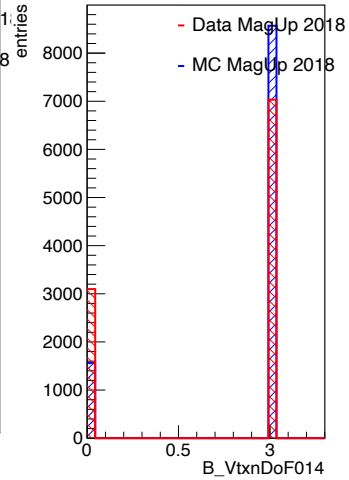


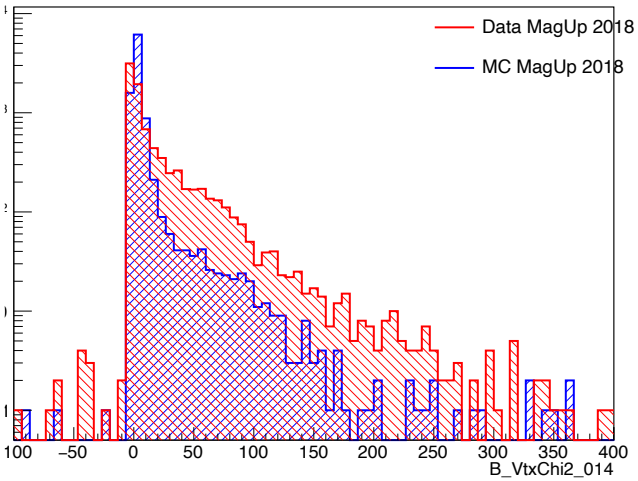
Correlated background suppression

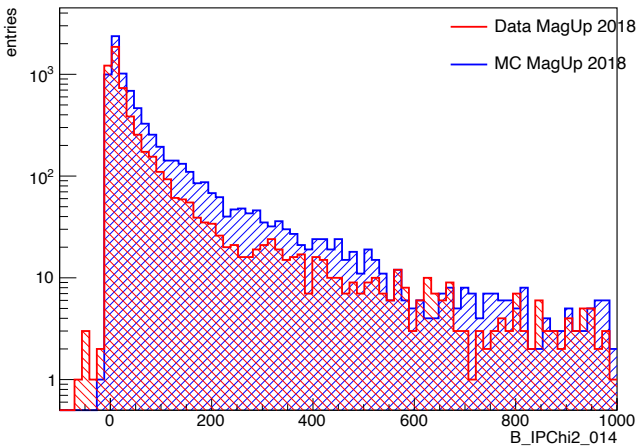
$J/\psi - K$ correlations [014]

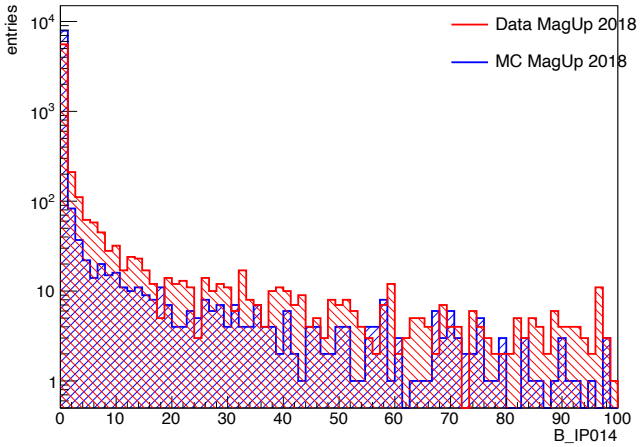
Temporary







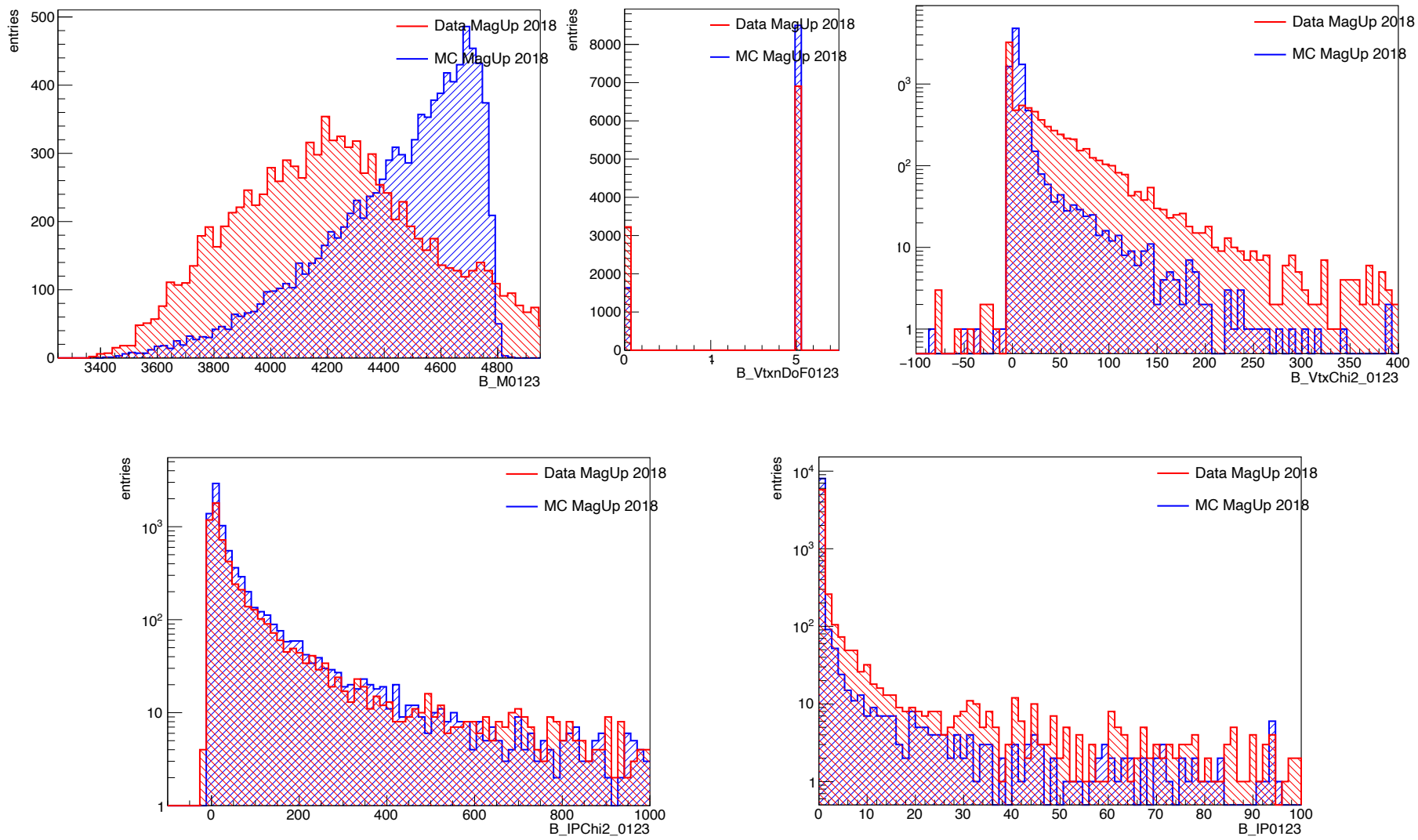


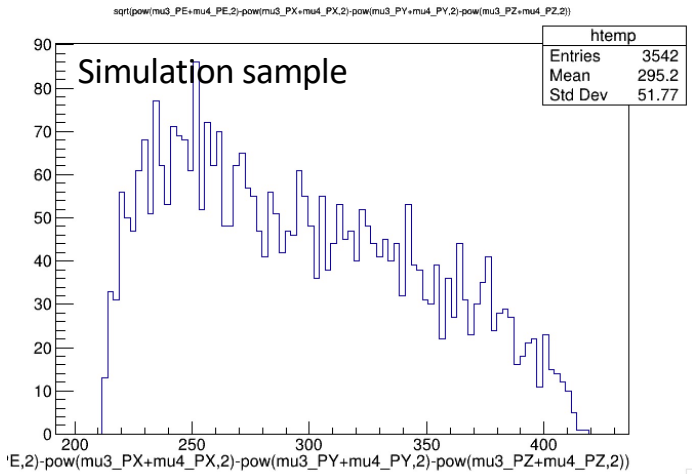


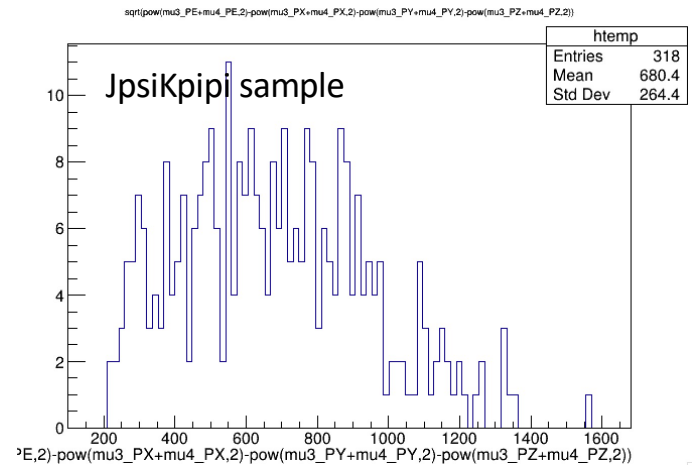
Correlated background suppression

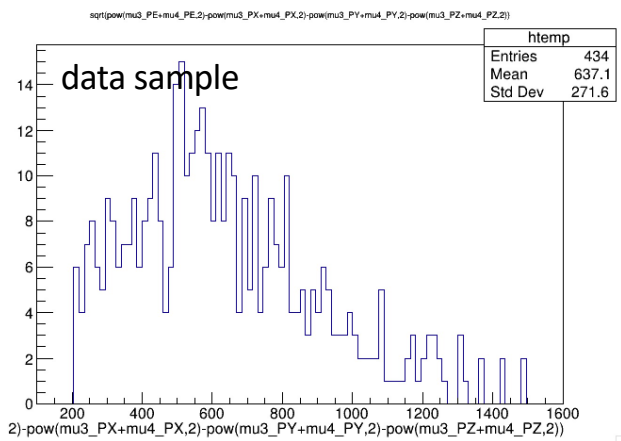
$J/\psi - \mu_3\mu_4$correlations [014]

Temporary




sqrt(pow(mu3_PE+mu4_PE,2)-pow(mu3_PX+mu4_PX,2)-pow(mu3_PY+mu4_PY,2)-pow(mu3_PZ+mu4_PZ,2))
htemp
Entries 3542
Mean 295.2
Std Dev 51.77
Simulation sample
'E,2)-pow(mu3_PX+mu4_PX,2)-pow(mu3_PY+mu4_PY,2)-pow(mu3_PZ+mu4_PZ,2))


sqrt(pow(mu3_PE+mu4_PE,2)-pow(mu3_PX+mu4_PX,2)-pow(mu3_PY+mu4_PY,2)-pow(mu3_PZ+mu4_PZ,2))
htemp
Entries 318
Mean 680.4
Std Dev 264.4
JpsiKpipi sample
PE,2)-pow(mu3_PX+mu4_PX,2)-pow(mu3_PY+mu4_PY,2)-pow(mu3_PZ+mu4_PZ,2))

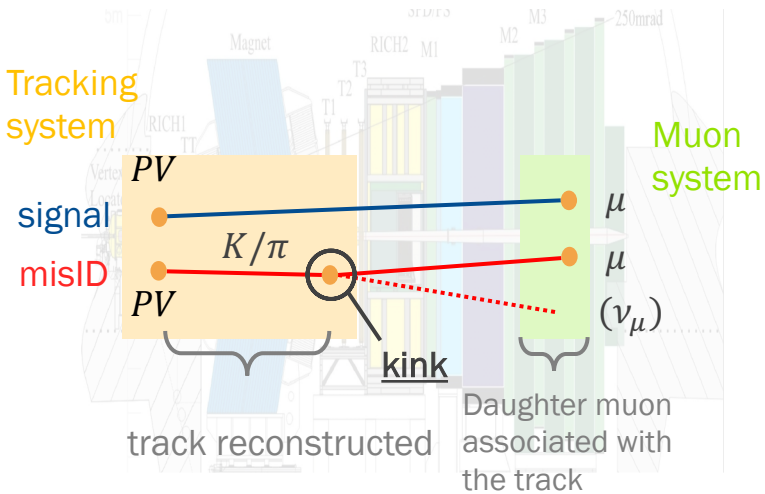

sqrt(pow(mu3_PE+mu4_PE,2)-pow(mu3_PX+mu4_PX,2)-pow(mu3_PY+mu4_PY,2)-pow(mu3_PZ+mu4_PZ,2))
htemp
Entries 434
Mean 637.1
Std Dev 271.6
data sample
2)-pow(mu3_PX+mu4_PX,2)-pow(mu3_PY+mu4_PY,2)-pow(mu3_PZ+mu4_PZ,2))

Soft dimuon misID background rejection

Kink calculator tool: principle

Case A: "Fake signal muons" are **muons, but coming from decays-in-flight of hadrons in the tracking stations.**

- 2 body decay, so change of direction of the muon track. This change of slope is the "kink".



- Interesting parameters

  `TRACK_kinkChi2:` for each point on the track, measure the track chi2 difference between forward and backward tracks. Return the maximum difference along one track.

  `TRACK_kinkChi2z:` returns the z-coord of the maximum track chi2 difference.

  `MATCHChi2 : same` as `TRACK_kinkChi2` but only for kinks in the **magnet**

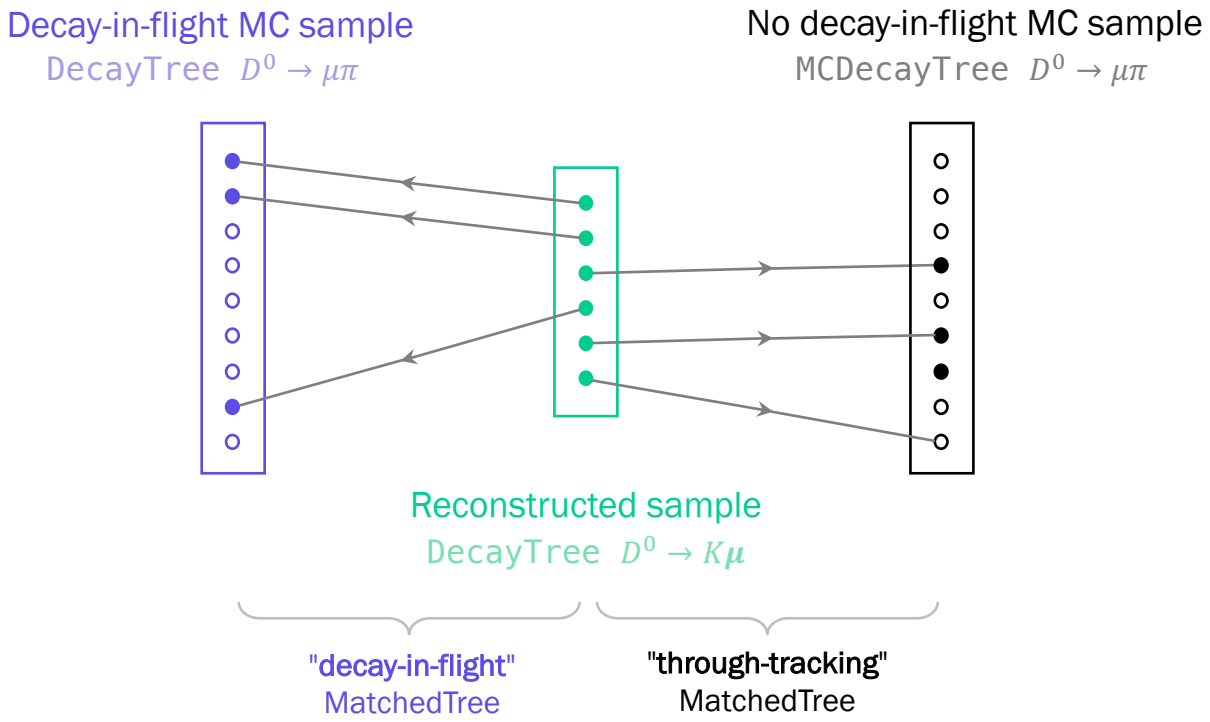
too sensitive to multiple scattering! ☹

Soft dimuon misID background rejection

Gen-level matching

Motivation:

- MC truth gives the TRUEID but **gives no info about the nature of the decay in flight.** In particular, where did it happen?
- Crosscheck Kink algorithm. Can we trust it? Can we use it in a MVA?

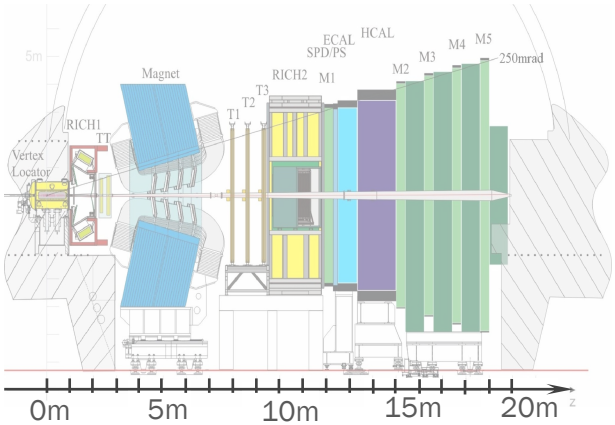


Soft dimuon misID background rejection

Kink calculator tool: does it tell the (MC)truth?

Where are the kinks happening?

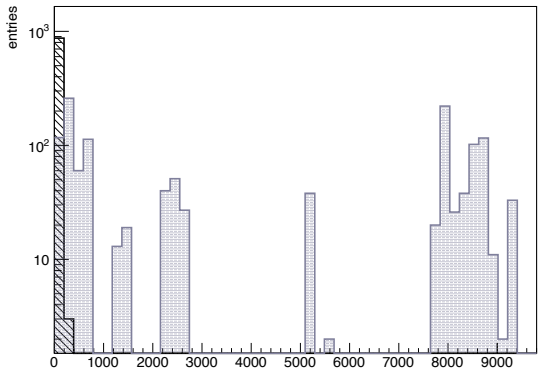
- **kinkZ** is the z-coord. of the greatest change in χ^2 along a track
- Compare {MC TRUTH z} to {z-coord of max kinkChi2}
  - Signal muon: **prompt,** created in VELO + multiple scattering in T1-3
  - misID samples: muons created in-flight mostly



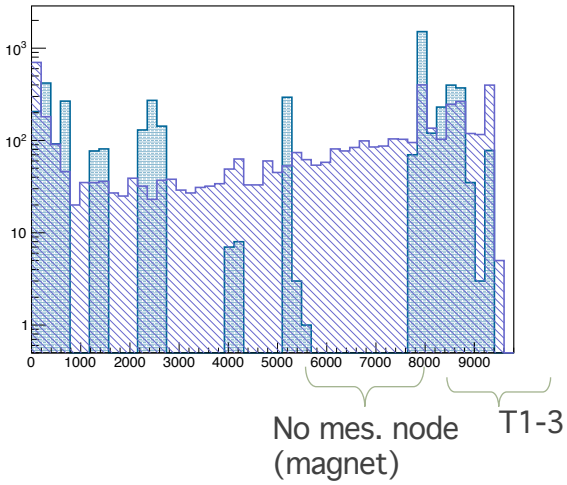
Mu soft TRUEORIGINVERTEX_Z **(MC TRUTH !)**

Mu soft TRACK_kinkChi2z

Signal $\gamma^* \rightarrow \mu\mu$

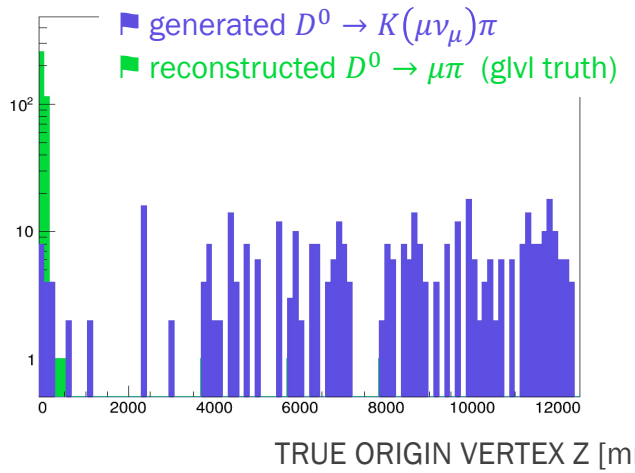


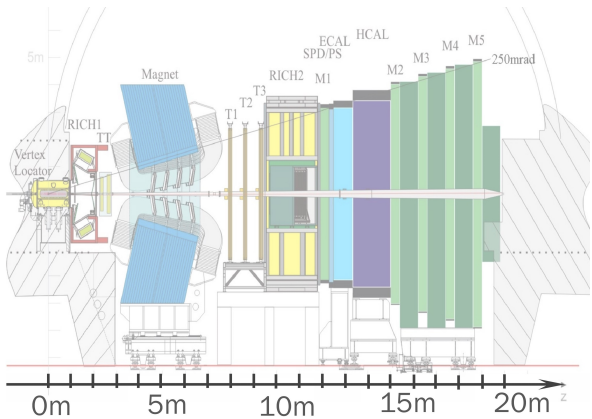
$D^0 \rightarrow (\pi \rightarrow \mu\, \nu_\mu) K$



Soft dimuon misID background rejection

Kink calculator tool: sounds good, doesn't work





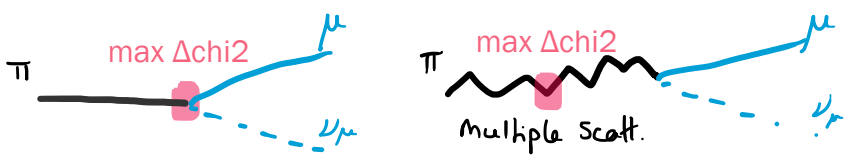
uncorrelated!

Plot the scatter plot

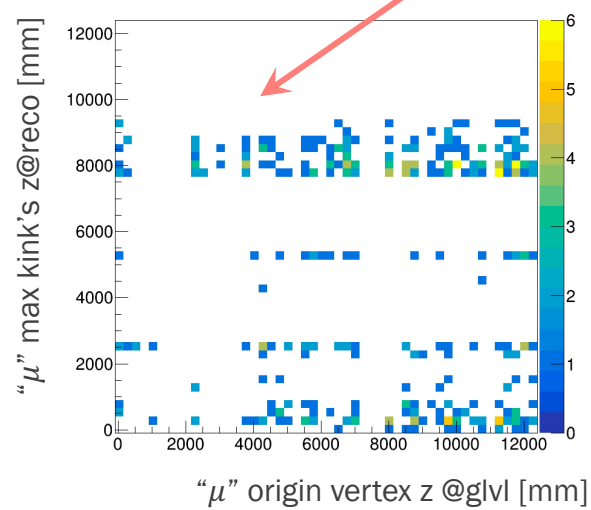
If kink algo works: kinkchi2z = TRUEORIGINVERTEX

But: we observe random distribution ☹

Probably due to multiple scattering that "steals" the max chi2 difference



Expectation *Reality*



=> Not adding kink tool in the MVA since not trusting them

Kink code

More sophisticated: **kink** cmputer of the track, could spot on-flight creation of muons

$\mathcal{B}(K \to \mu\, \nu_\mu) =$ 63.56 % [PDG 2019], $\tau_K = 1.2 \cdot 10^{-8}$ s
$\mathcal{B}(\pi \to \mu\, \nu_\mu) =$ 99.99 % [PDG 2019], $\tau_\pi = 2.6 \cdot 10^{-8}$s

<u>Algorithm</u>:
For each track, collect the batch of mes. ***nodes*** (collection of track *states*)
For each node (with step *s*):

```
chi2kink = ( chi2tot - chi2forward[i] - chi2backward[i + 1] ) / ndof;
```

Return the **maximum** one along the track.

S, ndof are hardcoded (4 and 5)

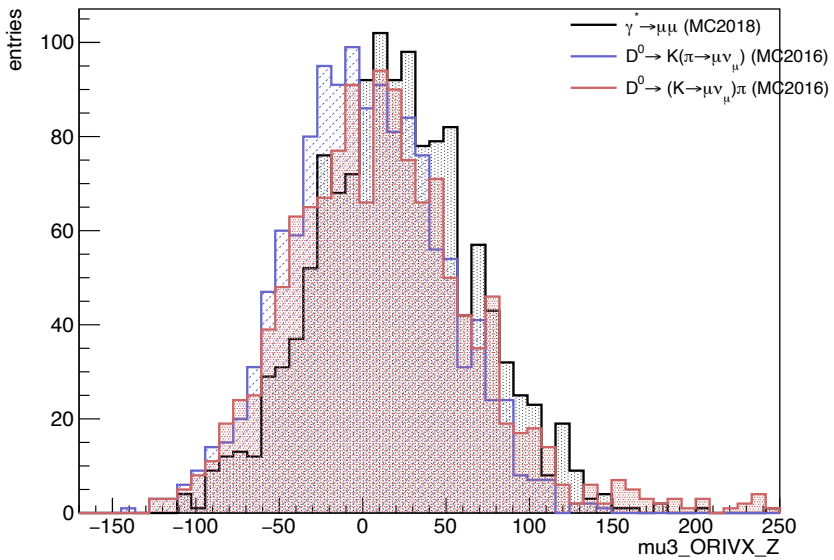
And zkink = 0.5 * (measnodes[ikink]->z() + measnodes[ikink + 1]->z());

.

Fake muons
Tracks stdy

- Muons are creatin on-flight from hadrons
  - Though, the *reconstrcuted* SV (ORIVX) are pretty similar
  - … but the TRUE z-coord of muon creation is very different
  - →much longer flight of K and pi -> muons tends to be created until T1-3 stations.. (then harsh cut)

Reco



Reality