Data quality assessment and search for new $t\bar{t}$ resonances in ATLAS

Séminaire Doctorant 2^{e} Année

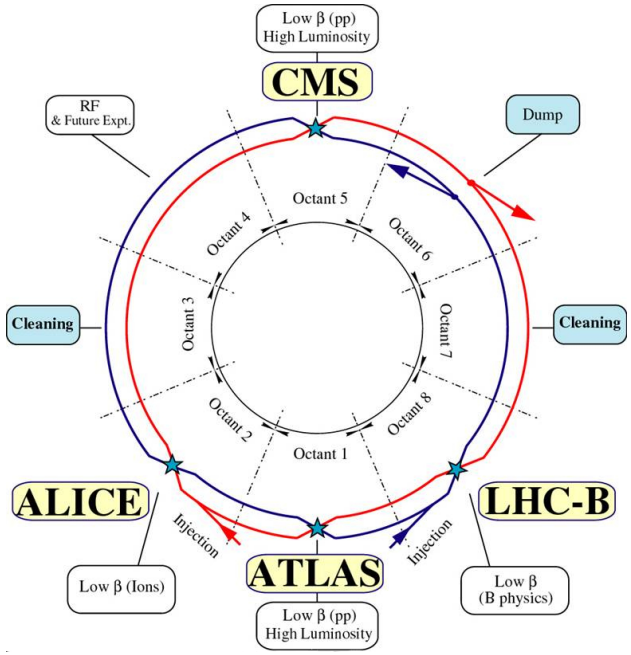
Clément Camincher

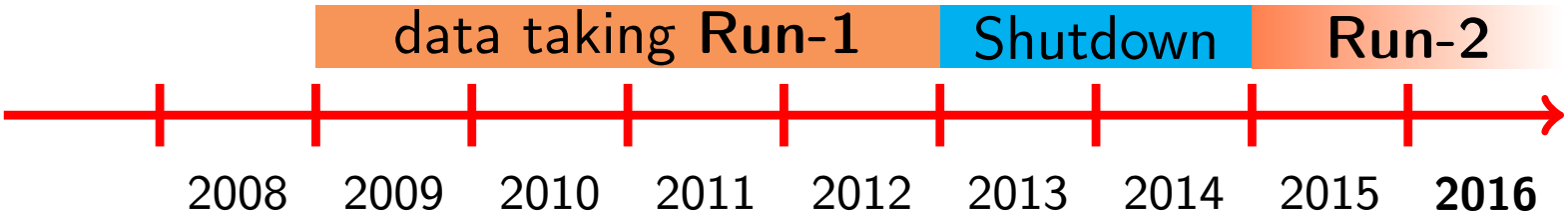
Sabine Crépé and Benjamin Trocmé

22/03/2016

The LHC

- The LHC is an accelerator of 27 km of circumference
- Proton-proton (or Pb) collision in 4 points

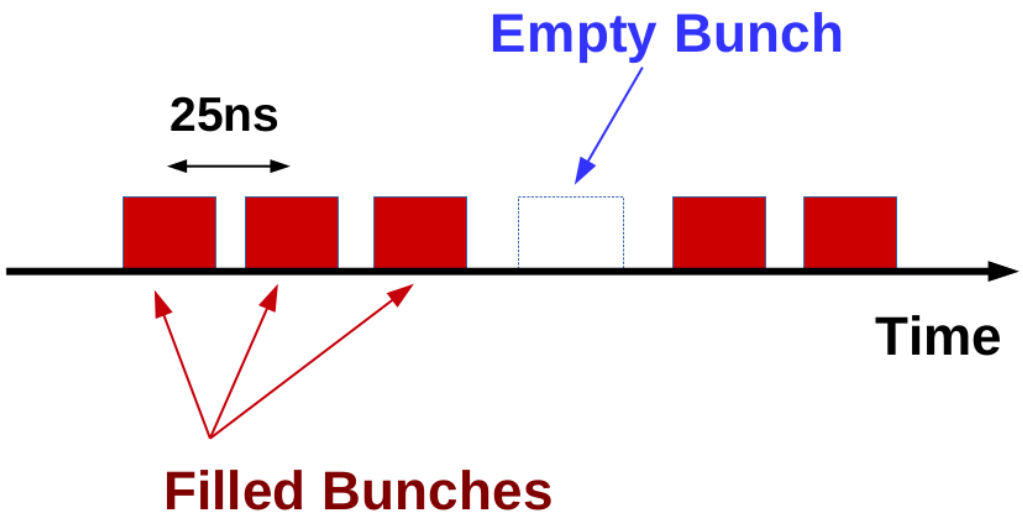




- After two years of shutdown, the LHC restarted for the Run-2
- In 2015 the proton-proton center of mass energy = 13 TeV

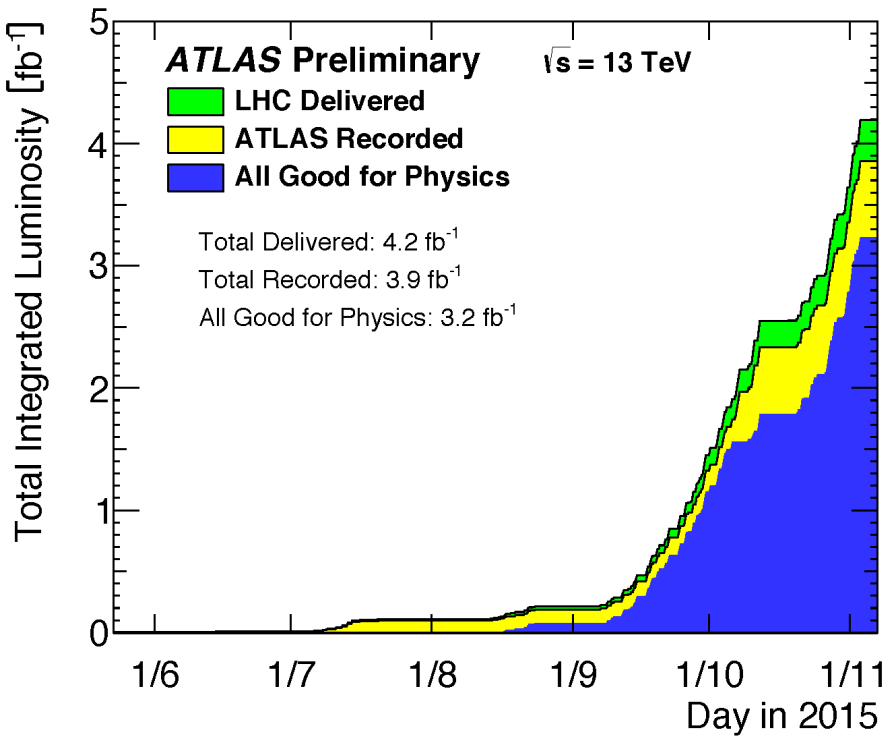
Filling the LHC

- The LHC is filled with bunches (packets of protons)
- Ideally 3500 bunches could be filled (each separated by 25 ns)
- Some remain empty in the filling scheme



- Some positions are **EMPTY**
- Useful to assess the stability of the detector
- No collision expected during empty bunch crossing

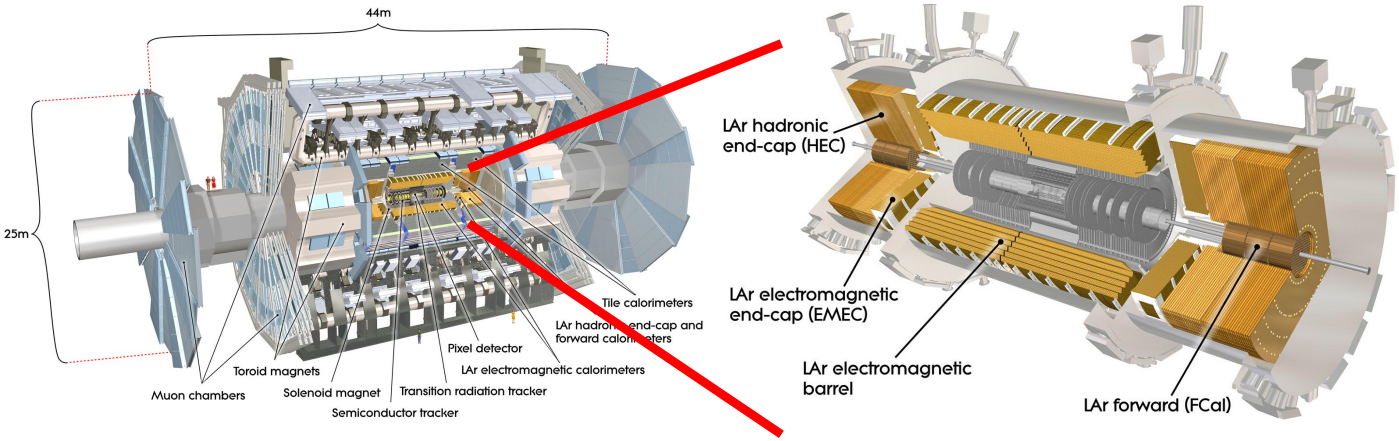
The data quality in ATLAS



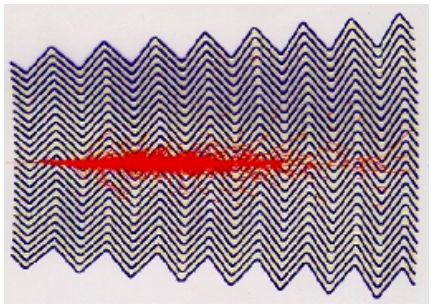
- Collecting data (luminosity) is one of the main goal of ATLAS
- In many cases the more is the best
- Analyses have to be done only on reliable data
- Hence, it is needed to assess the data quality and remove the problematic events

The electromagnetic Liquid Argon (LAr) Calorimeters of ATLAS

- Composed of an electromagnetic Barrel (EMB) + 2 end-caps (EMEC) on each side

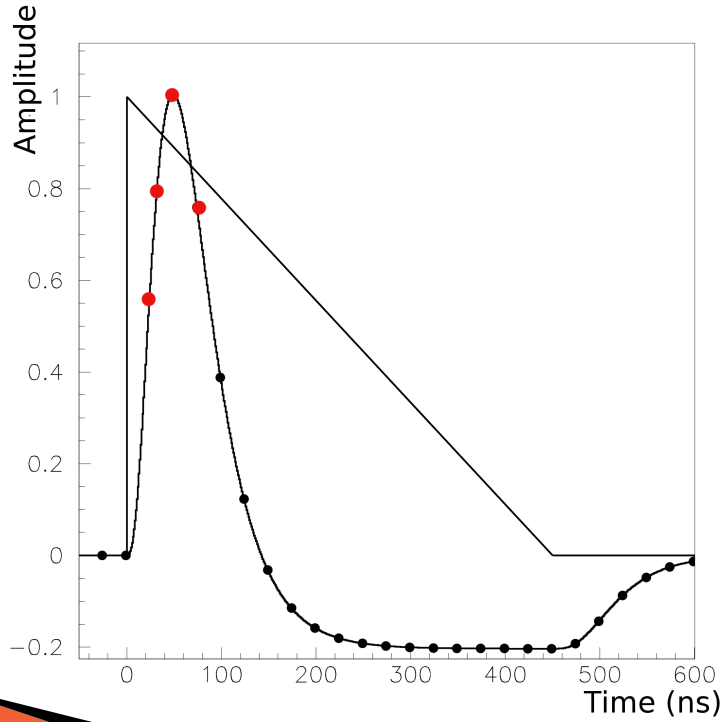


- Sampling calorimeters :
  - A shower is induced by lead plates (in accordion)
  - Particles ionize the liquid argon (in the gap between the plates)
  - The charges are collected with electrodes within the argon



Reading out the LAr Calorimeters

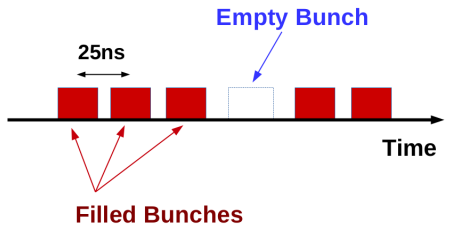
- The created intensity pulse is triangular
- Once shaped, an electromagnetic deposit becomes a bipolar pulse
- 4 samples of the measured signal are digitized and fitted to a calibration reference

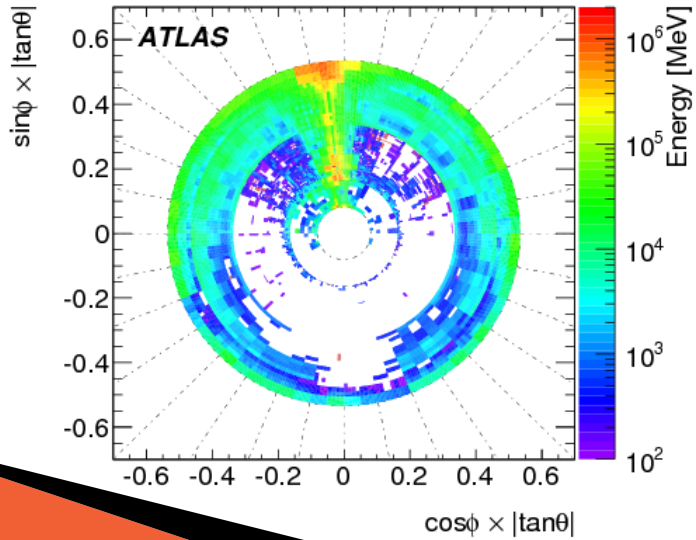


- The height of the pulse is proportional to the energy deposit
- A simplified χ^2 test is also computed $\Rightarrow$ Quality-factor (Q-factor)
- For a Q-factor $>$ 4000 the measured signal is considered as an instrumental background

The noise in the LAr Calorimeters

- In barrel $\approx$ 110 000 cells , 31 000 in each end-cap
- Each cell is independent and characterized by an electronic noise
- This noise is Gaussian distributed with a standard deviation σ
- In empty bunch no collision are expected
  $\Rightarrow$ almost no signal measured
- Sometimes a largely coherent noise is observed (noise burst)
- The reco energy of such event corresponds to several TeV





- Since the Run-1, many studies tried to find the source without success
- Need to characterize and remove this noise with as less data loss as possible

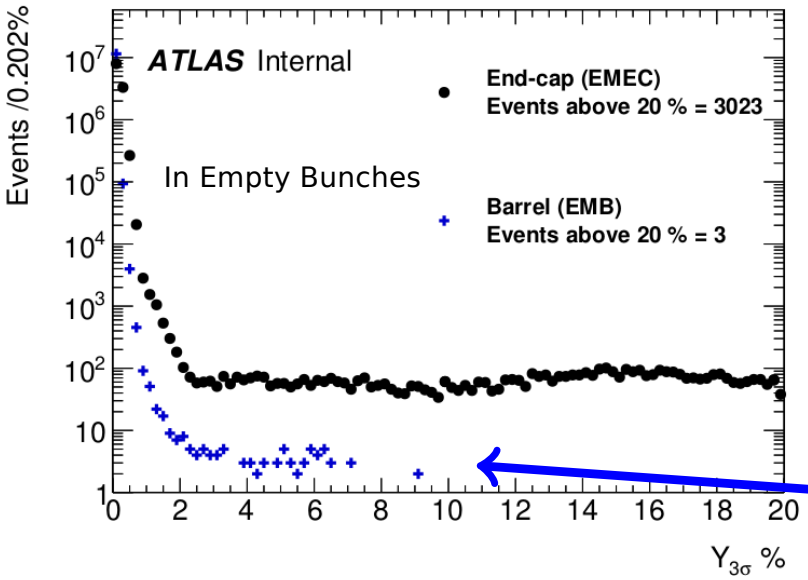
Characterizing the coherence

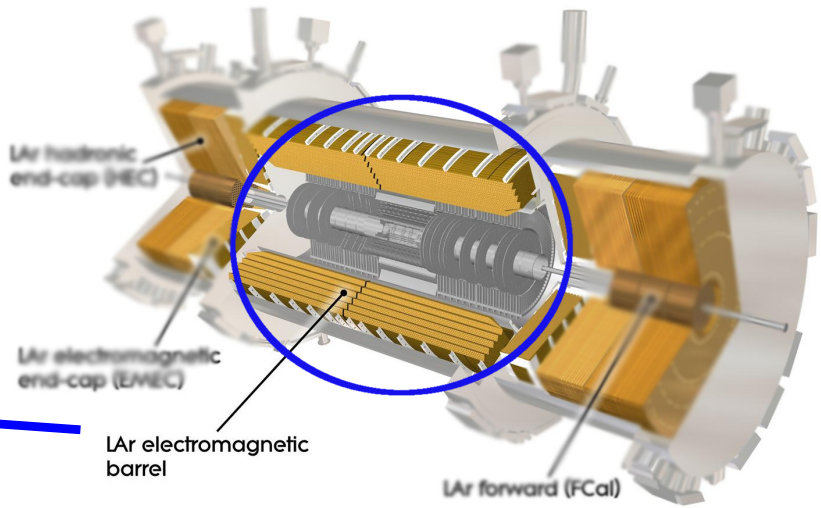
- To characterize the coherence, the $Y_{3\sigma}$ variable is defined

$$Y_{3\sigma}^{detector} = \frac{N_{Cells}^{detector}(E_{measured} > 3 \times \sigma_{Noise})}{N_{Cells}^{detector}}$$

- In the hypothesis that all the cells are independent and have only Gaussian noise
- $Y_{3\sigma}$ should be distributed around 0.13%

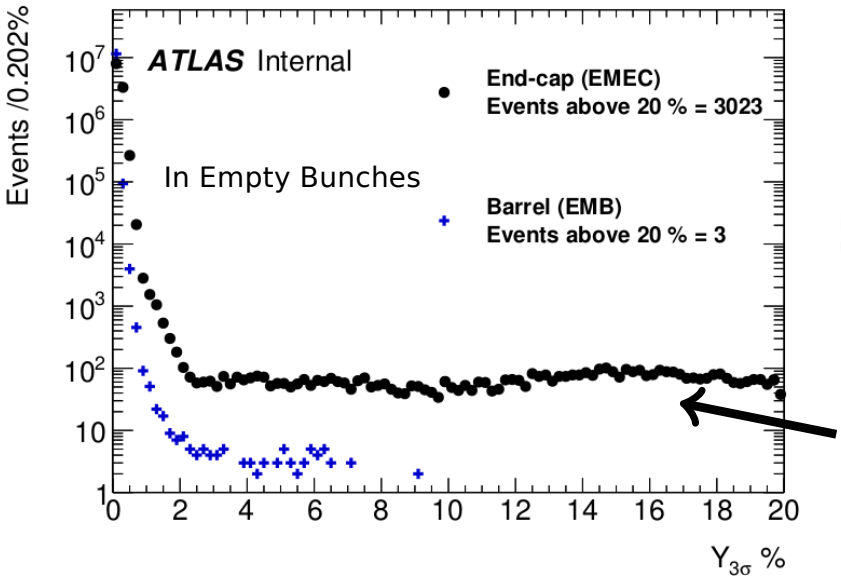
The $Y_{3\sigma}$ distributions in the EMB and EMEC

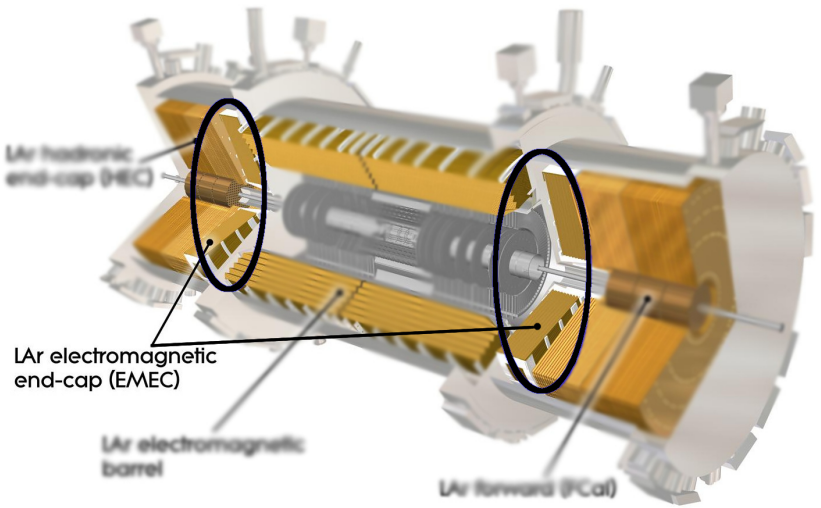




- In the barrel the distribution is as expected
- In the end-cap : presence of a long tail
- It signs a high occupancy when NO collision are expected !

The $Y_{3\sigma}$ distributions in the EMB and EMEC

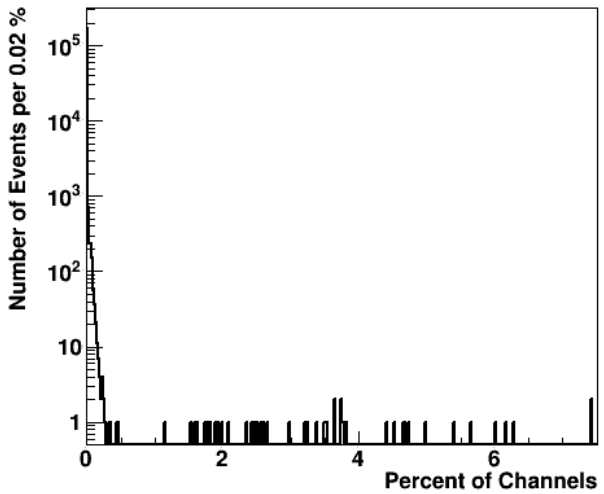




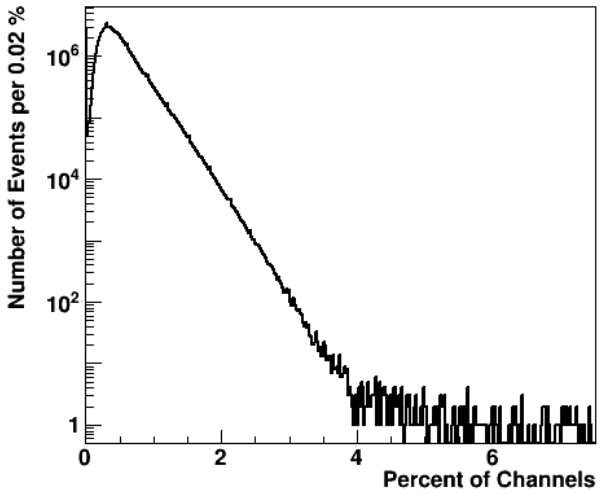
- In the barrel the distribution is as expected
- In the end-cap : presence of a long tail
- It signs a high occupancy when NO collision are expected !

$Y_{3\sigma}$ in empty bunches vs collisions

- When a collision is expected, $Y_{3\sigma}$ is naturally enhanced
- The flow of particles increases the occupancy



$Y_{3\sigma}$ **in empty bunches**

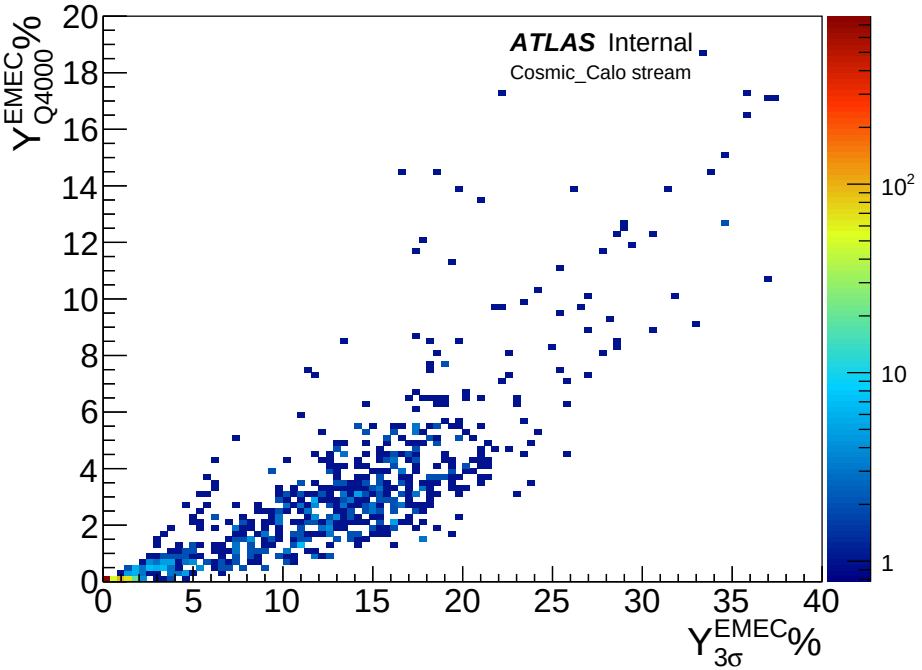


$Y_{3\sigma}$ **in collision bunches**

- $Y_{3\sigma}$ **is not usable to discriminate coherent noise when collisions happen in ATLAS**

How to select noisy events during collisions

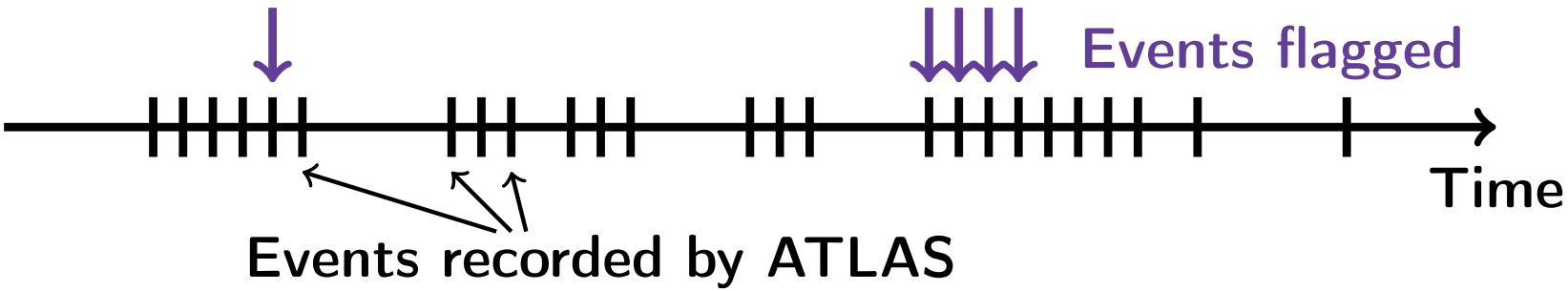
- The source is not known, but has likely an instrumental origin
- Another discriminating variable is the Q-factor
- Similarly to $Y_{3\sigma}$ the Y_{Q4000} variable is computed.
- Y_{Q4000} is the proportion of cells with $Q > 4000$



- The correlation indicates that a selection based on the Q-factor would also select noisy events

Procedure to remove noise bursts 1

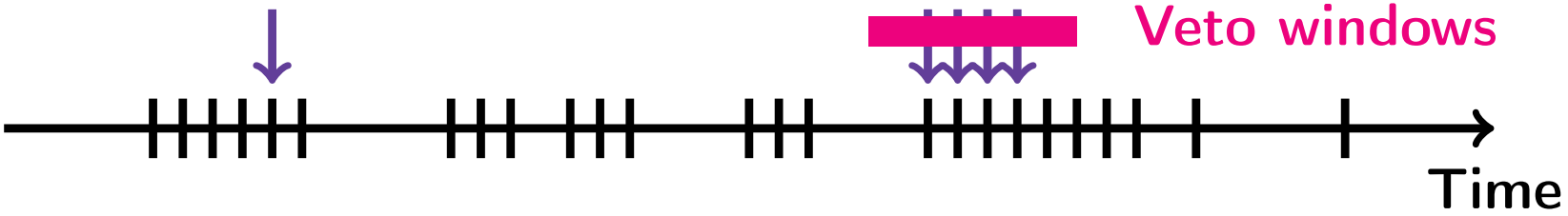
- Based on the Q-factor, the noisy events are flagged.



- **Why not just removing those problematic events ?**

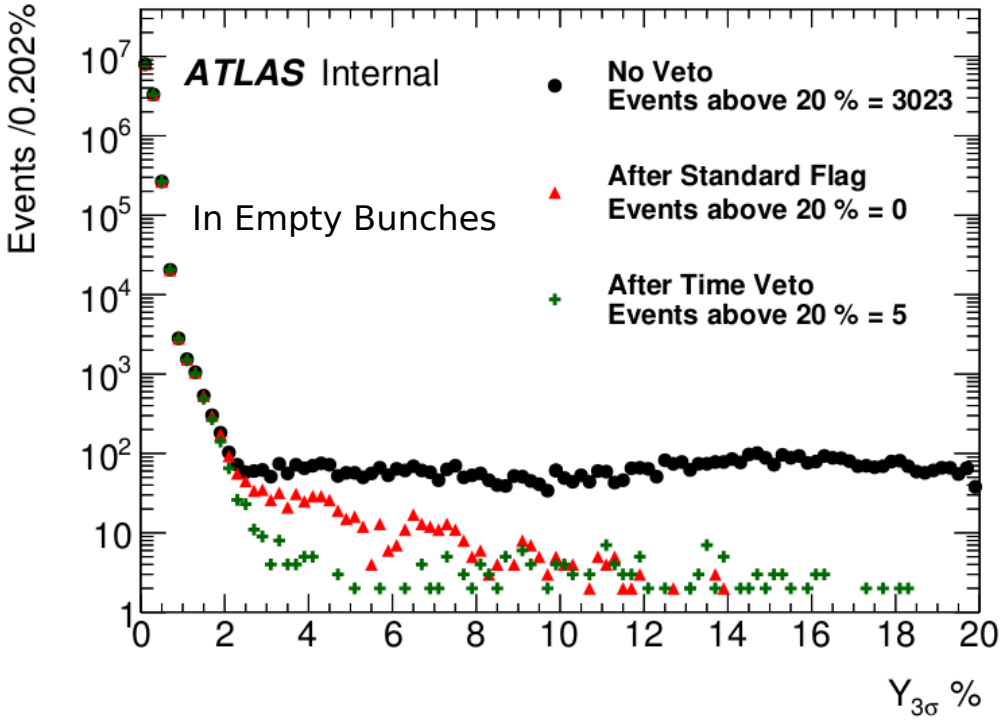
Procedure to remove noise bursts 2

- ATLAS is a machine of discoveries
- Some long living particles could induce delayed signal
  $\Rightarrow$ high Q-factor
- However, having two of such event in the vicinity is extremely unlikely
- When at least 2 flags are clustered, the noisy events are considered as coming from the same noise burst
- All the events in this period **are not used** for physics
- This induced a loss of **0.2%** in 2012 and **0.03%** in 2015



The $Y_{3\sigma}$ distribution after the removal procedure

- The $Y_{3\sigma}$ distribution allows to control the efficiency of the procedure



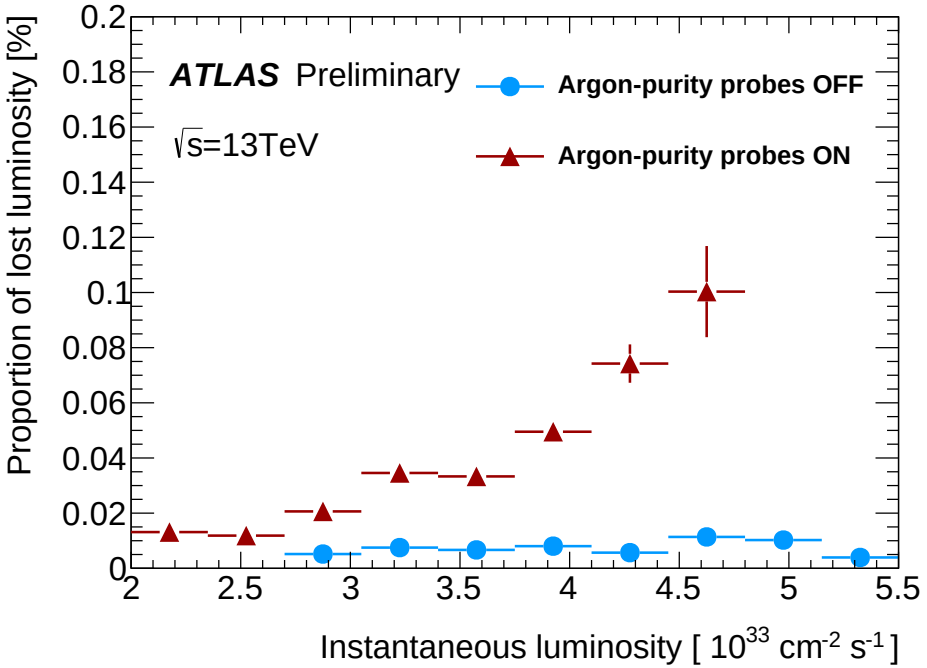
- The tail is well cleaned
- The remaining tail may come from isolated events

Trying to understand the source

- Despite this efficient procedure, some tests have been performed to understand the source of noise-bursts
- In October 2015, it was decided to switch OFF the probes measuring the purity of the liquid argon
- Those probes are inside the argon cryostat
- Argon purity probes = α particles from a radioactive source detected via a string chamber
- The high voltage (HV) lines of the string chamber were turned OFF.

Impact of purity probes HV lines on noise-bursts

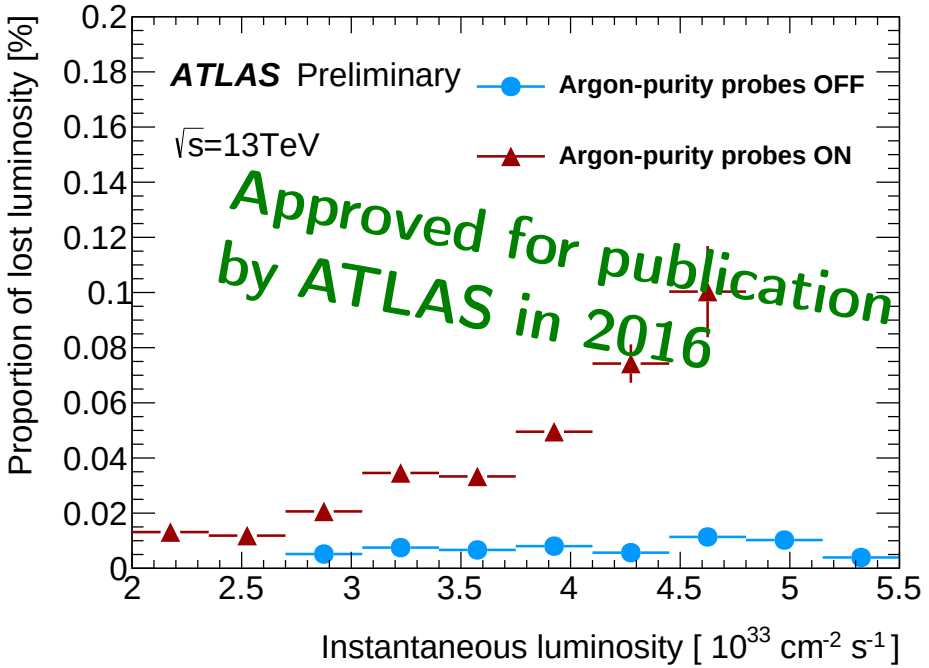
- Since the Run-1, a strong dependence with the instantaneous luminosity was observed
- Problematic as the LHC inst lumi is planned up to $30.10^{33} cm^{-2} s^{-1}$



- **Turning OFF the high voltage line reduced the rate and made it independent of the instantaneous luminosity**

Impact of purity probes HV lines on noise-bursts

- Since the Run-1, a strong dependence with the instantaneous luminosity was observed
- Problematic as the LHC inst lumi is planned up to $30.10^{33} cm^{-2} s^{-1}$



- **Turning OFF the high voltage line reduced the rate and made it independent of the instantaneous luminosity**

Data quality summary

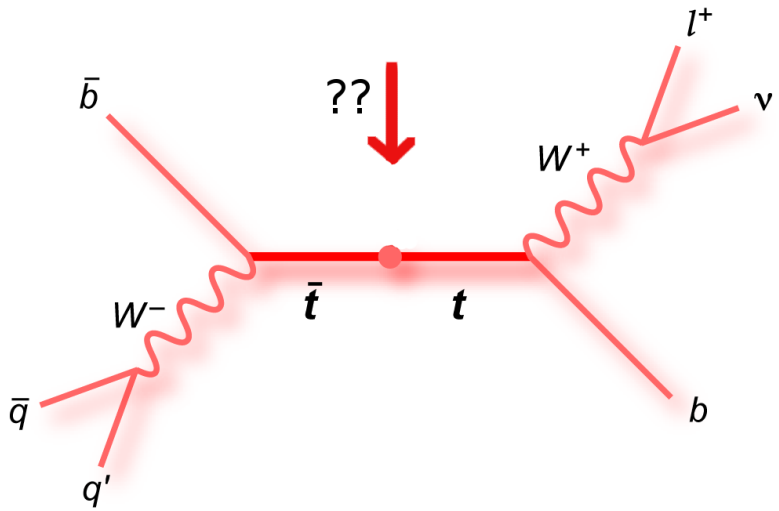
- The noise-burst are removed with a high efficiency and the rate is now very low.
- It won't impact much the data taking in the future
- This work as well as the validation of the 2015 data contributed to increase LAr efficiency to 99.4% in 2015 (99.1% in 2012)
- An internal note is in preparation
- 2 posters presented on behalf of the LAr community

ATLAS pp 25ns run: August-November 2015										
Inner Tracker			Calorimeters		Muon Spectrometer				Magnets	
Pixel	SCT	TRT	LAr	Tile	MDT	RPC	CSC	TGC	Solenoid	Toroid
93.5	99.4	98.3	99.4	100	100	100	100	100	100	97.8
All Good for physics: 87.1% (3.2 fb^{-1})										

Search of new physics in the top anti-top final state

The signal we are looking for

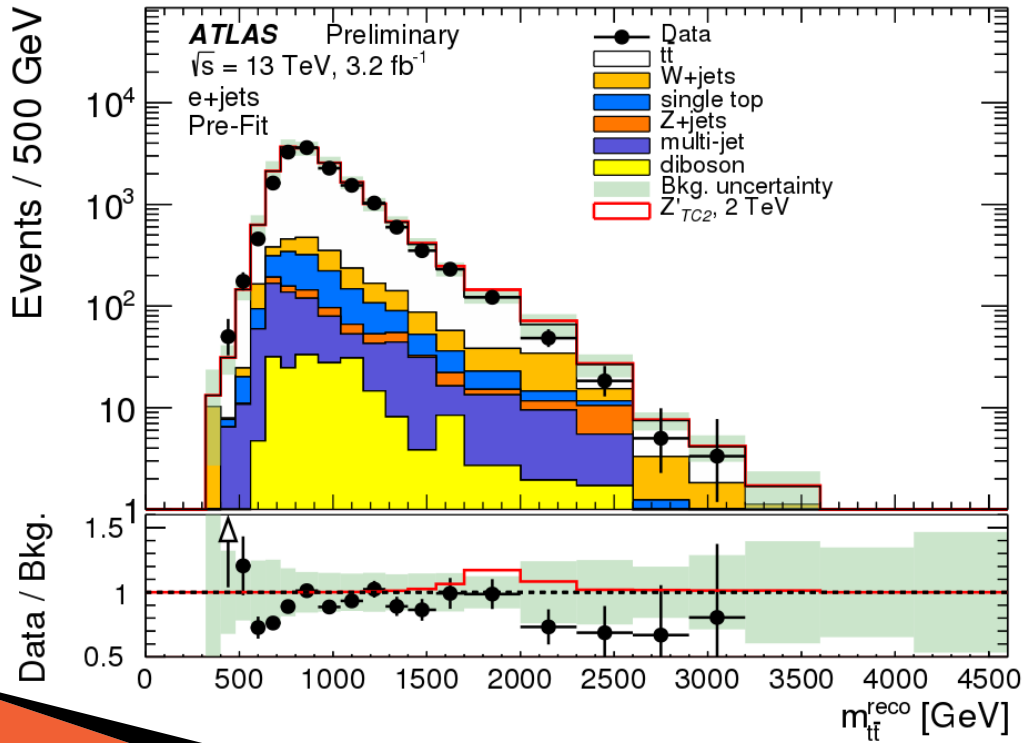
- Top quark is very heavy (175 GeV)
- It may couple strongly with new physics in many models (Extra dimensions, technicolor...)
- Try to be as model independent as possible
- Looking especially (but not only) for heavy resonances



- Decay in $t \rightarrow W + b$ before hadronization
- Here $t\bar{t}$ semi-leptonic decay studied as it is the best (statistic/background-rejection) ratio

The analysis methodology

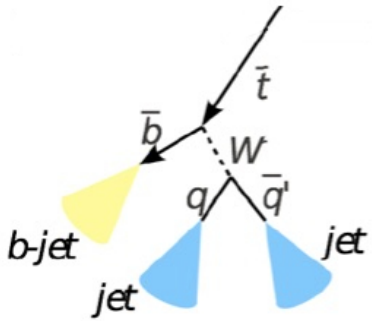
- Select the good reconstructed particles
- Select the events coming from a $t\bar{t}$ decay
- Reconstruct the top pair invariant mass
- Invariant mass $\Rightarrow$ mass of the new particle decaying in 2 tops



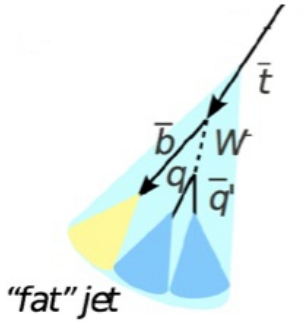
- Search for deviations between the data and the predicted backgrounds
- If no deviation found a limit is set on benchmark models

Difficulties to select the particles

- The quarks (except top) hadronize $\Rightarrow$ radiate gluons and quarks to create composite particles
- In high energy physics a shower of particles signs the emission of a quark $\Rightarrow$ a **JET**
- If the top has a high momentum (ex: from heavy resonance), the decay products will merge

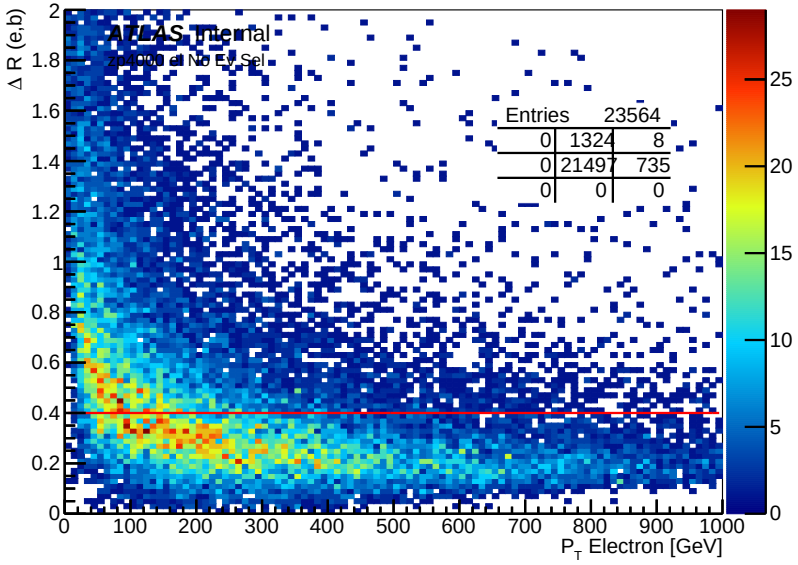


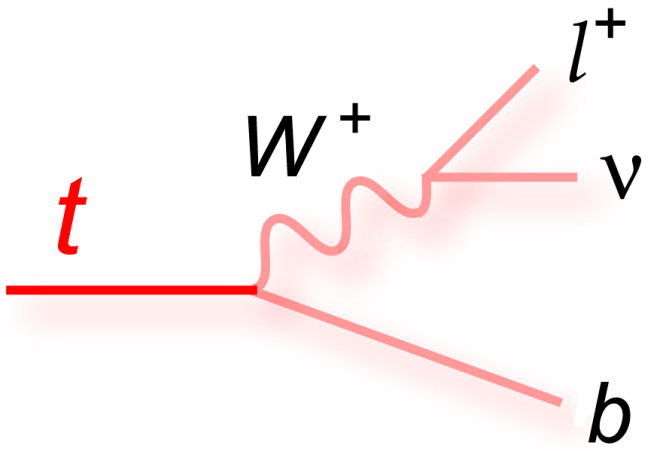
un-boosted



boosted

The leptonic boost

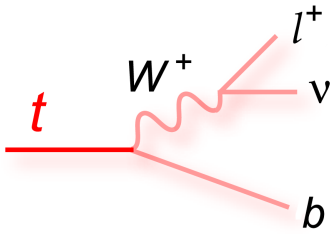


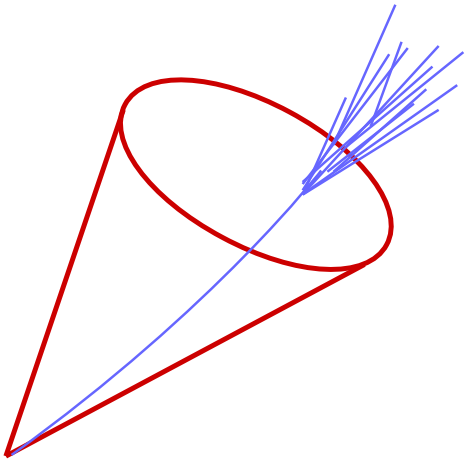


- Angular distribution of decay products vs the lepton energy
- The higher is the energy the smaller is the angular distance (kinematic effect)
- All these objects are measured by the same detectors
  $\Rightarrow$ delete events when lepton is in b-jet cone

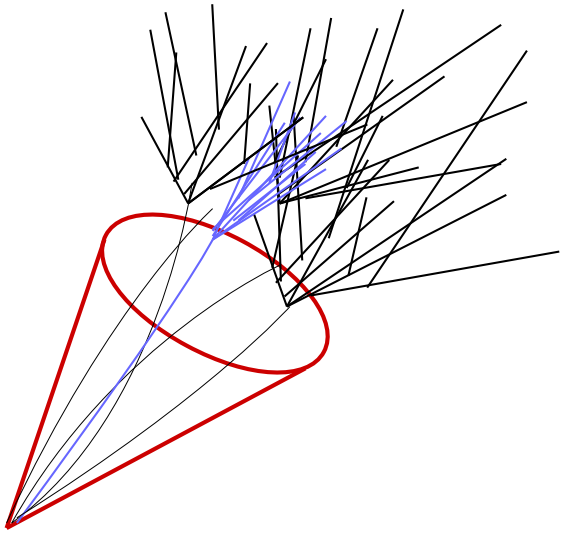
Leptons and b-jets

- But b-quark decay may also produce a lepton ($\approx 40\%$)





lepton from W

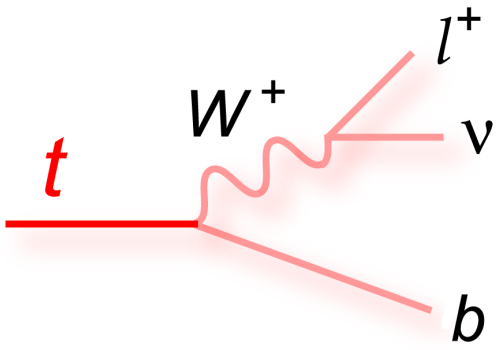


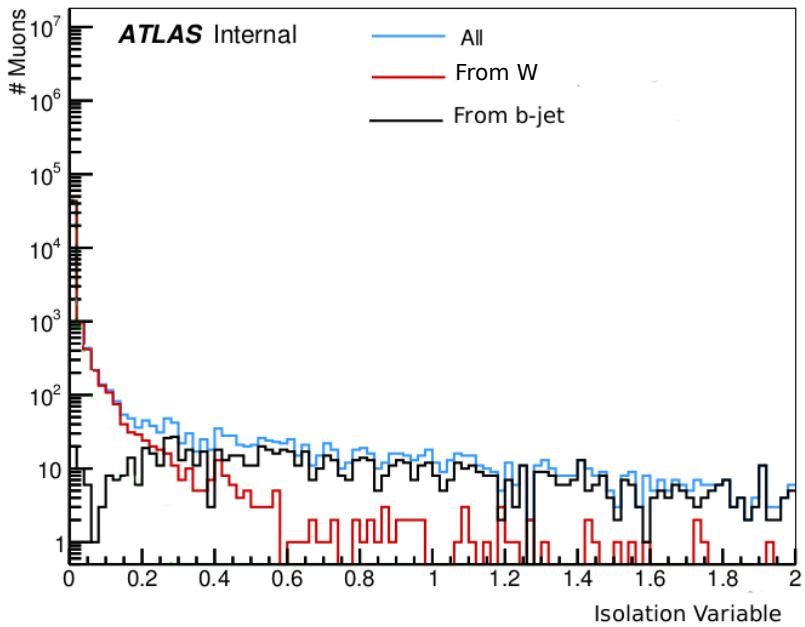
lepton in b-jet

- An isolation variable is computed: energy (activity) in a cone around the lepton

Isolation variable

- Due to the boost the b-quark gets closer to the lepton
- The cone size must decrease with the lepton energy $\Rightarrow$ new variable

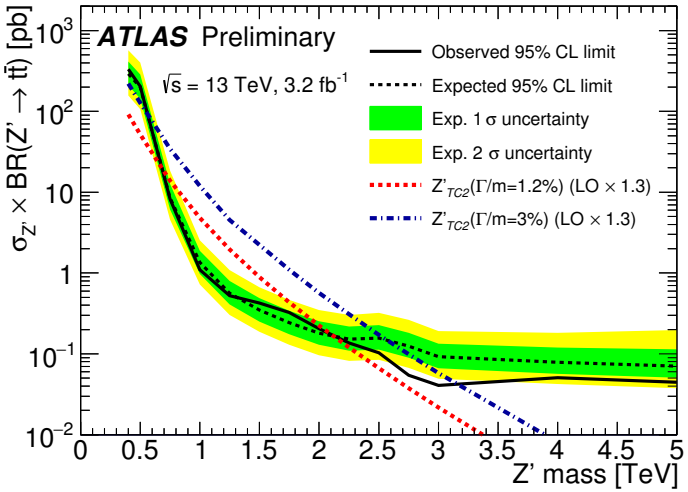


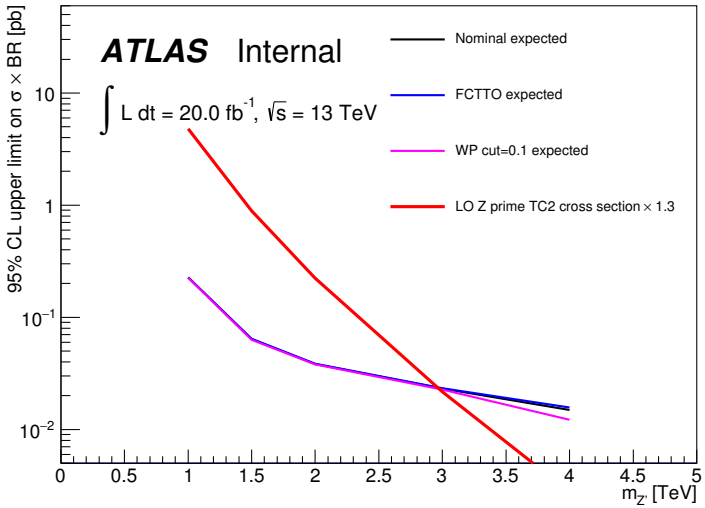


- Close to 0 $\Rightarrow$ the lepton is isolated
- Example on simulation of heavy resonance
- Try to optimized the selection based on such variable

Impact of isolation on the global analysis

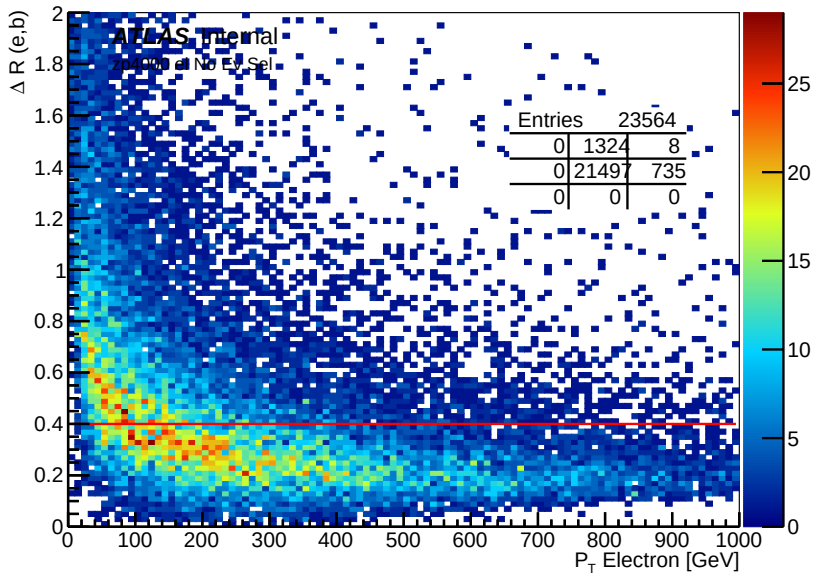
- Trying several working points
- Running the whole analysis to compare them
- New signal could be detected if its cross-section greater than measured cross-section limit





- Very small effect of isolation observed on global analysis

Perspectives for the analysis



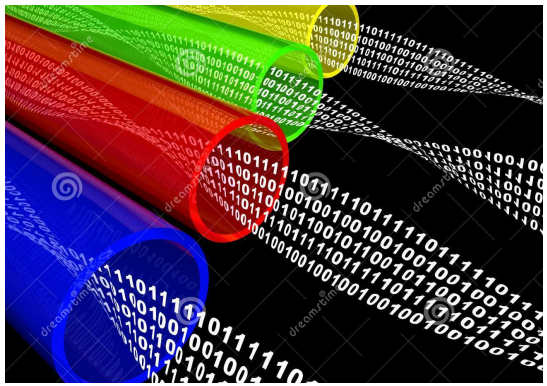
- To improve analysis selection efficiency
  - We will study techniques to recover lepton inside the b-jet
  - Study techniques related to boosted topologies (top tagging)

Concerning the data quality

- Keep in touch with LAr community to check if noise bursts remain under control

Concerning the search

- Work on boosted topology which is a very hot topic
- Possibilities to increase analysis sensibility
- More data will be soon available

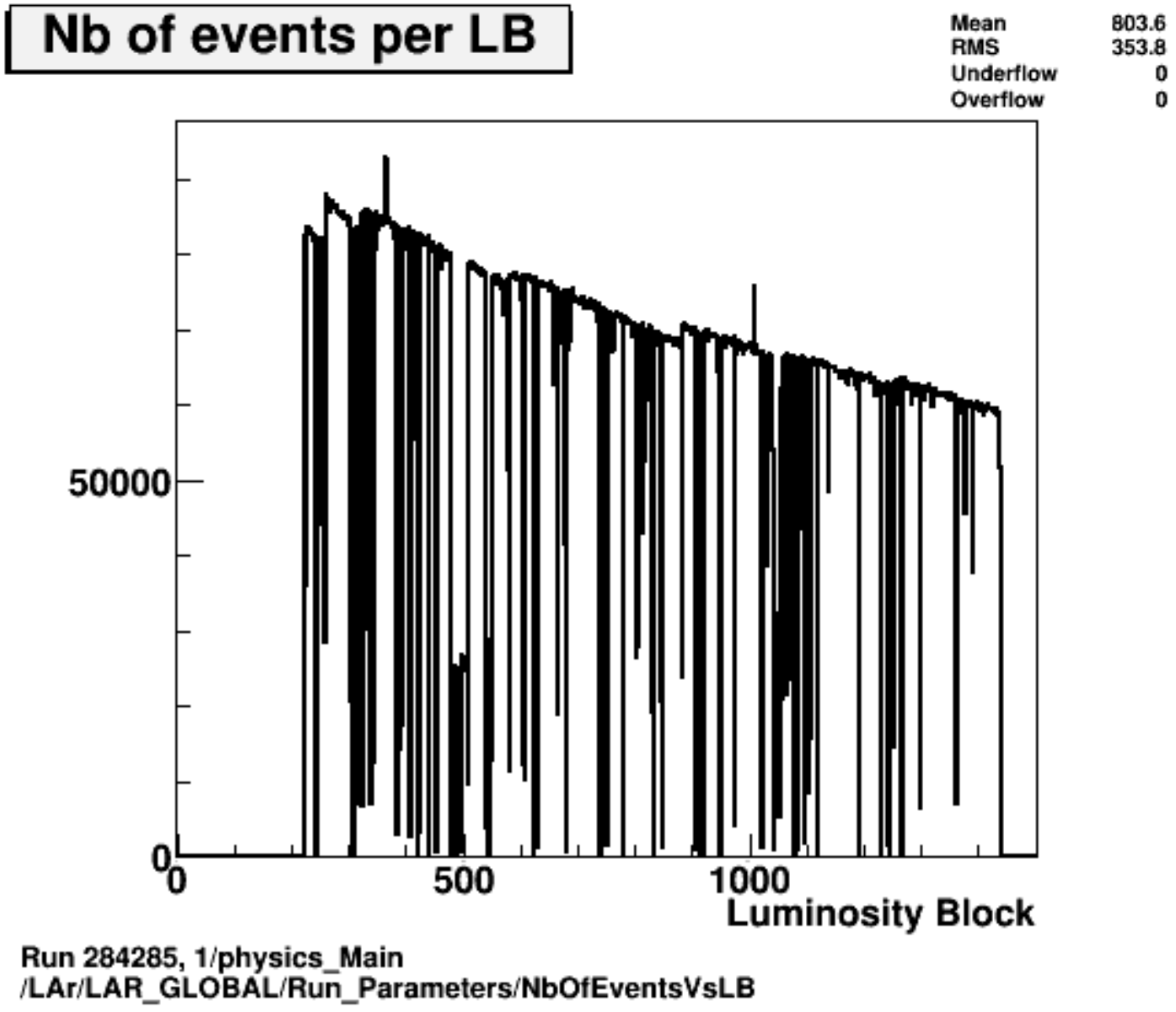


Ready to analyze the new data !

THANK YOU
QUESTIONS ?

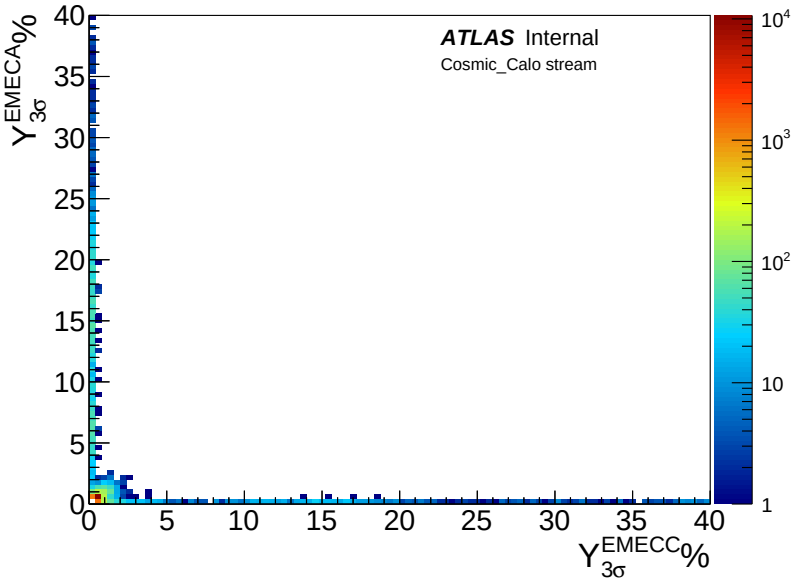
BACKUP

Rate in Main stream



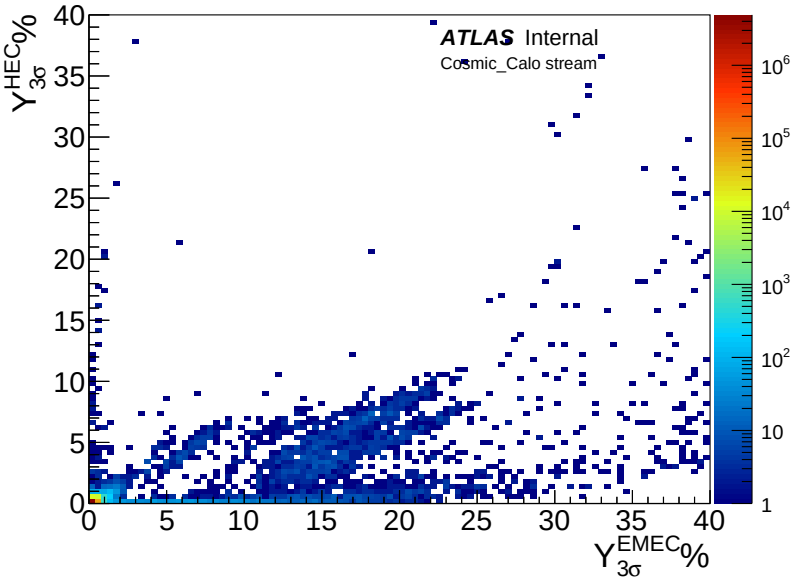
Space characterisation
Correlations

A side vs C side in EMEC



- No correlation between A and C side
- Same feature for FCal/HEC

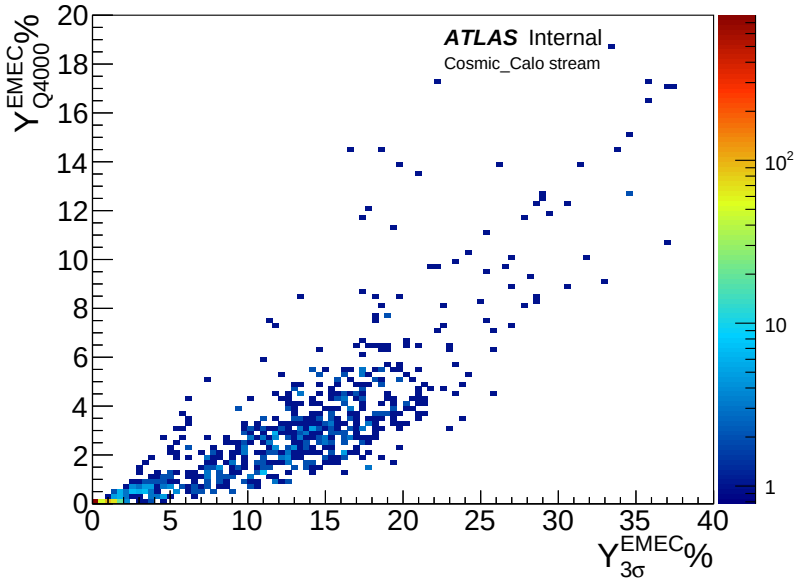
EMEC vs HEC



- Strong correlation between end-cap partitions
- NB in HEC $\Rightarrow$NB in EMEC
- The reverse is false
- Same feature for FCal

Selection of critical events

NOT relevant to use $Y_{3\sigma}$ in collision events Q factor should be used



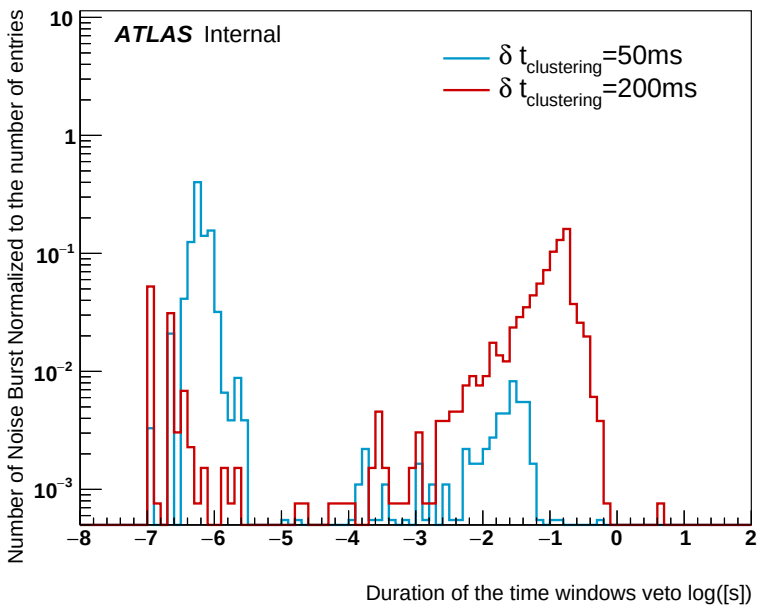
- Y_{Q4000} is the proportion of cells with $Q > 4000$
- Strong correlation between $Y_{3\sigma}$ and Y_{Q4000}

We define the Flags

- **Standard** : 5 FEB with more than 30 channel with Q>4000
- **Standard Optimized**: Experience show peculiar noisy FEB. A double weight is set on them. Similar to **Standard** but require the sum on FEB weight = 5
- **Saturated**: 20 Channel with saturated Qfactor (65536)& E>1GeV

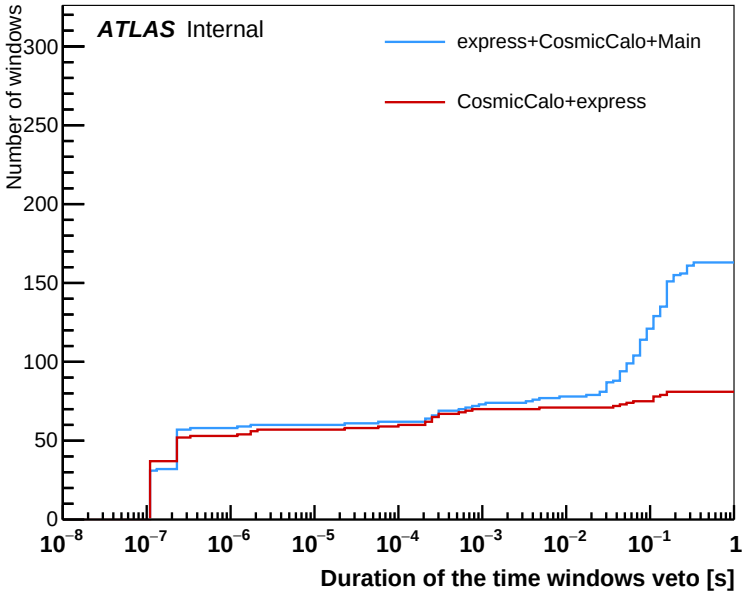
Time characterisation
Longer Noise burst

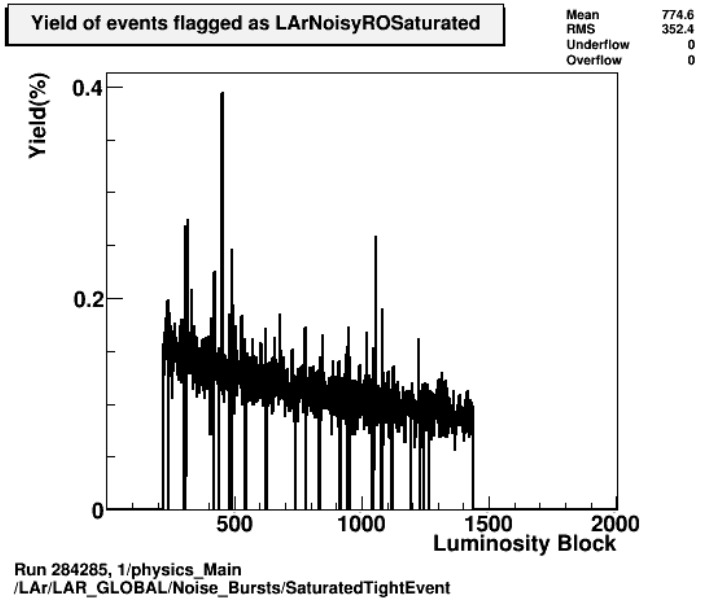
- The flags are gathered in time
- Cluster algorithm with characteristic time $\delta t_{Clustering}$
- Length of NB define as the time between the first and las flag in a NB.
- Much longer than in Run-1 with $\delta t_{Clustering} = 200\ ms$
- More similar if reduce clustering time to $\delta t_{Clustering} = 50\ ms$



Time characterisation
Impact of the physics streams

- Cumulative length distribution with or w/o the collision stream
- Strong increase $\approx 0.1\ s$ when collisiosn stream included
- Moreover a strong rate of sat flag is observed in the collisiosn stream ($0.2\% \times 80\ kevents/LB \approx 160 flag/min \approx 2/3 flag/s$

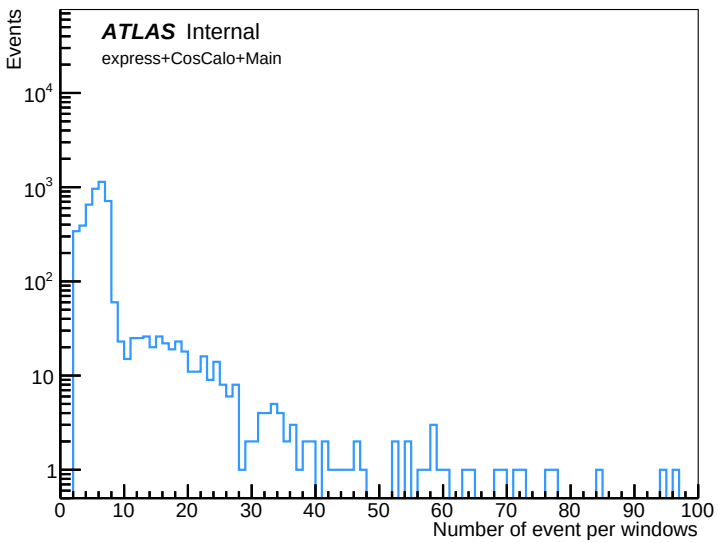




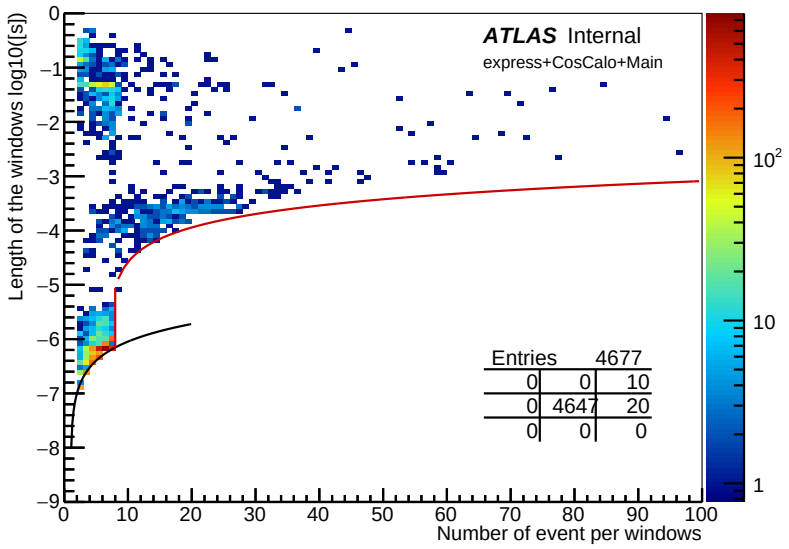
- The clustering methods catch "fake" saturated flag
- It increase artificially the length of NB

Flag repartition in the windows

Number of Flag/NB



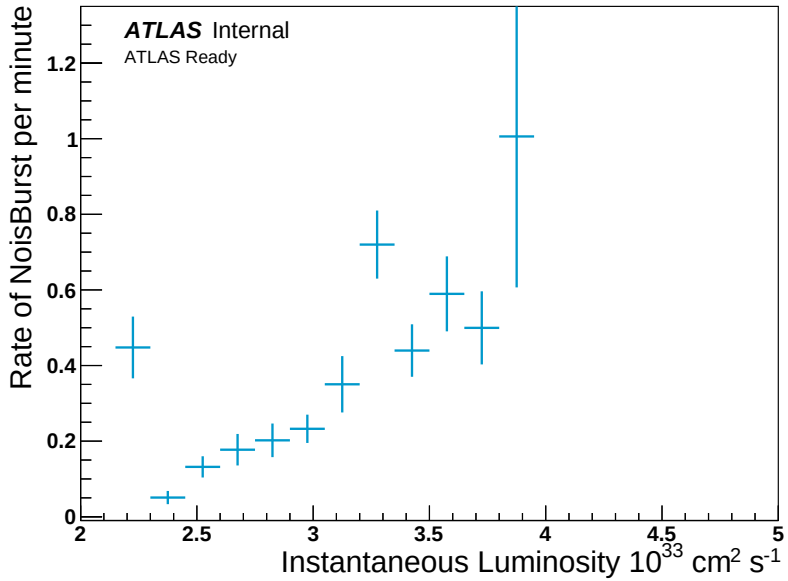
Number of flag vs length of NB



- Three populations are well visible
- Effect of **Simple** and Complex dead time visible

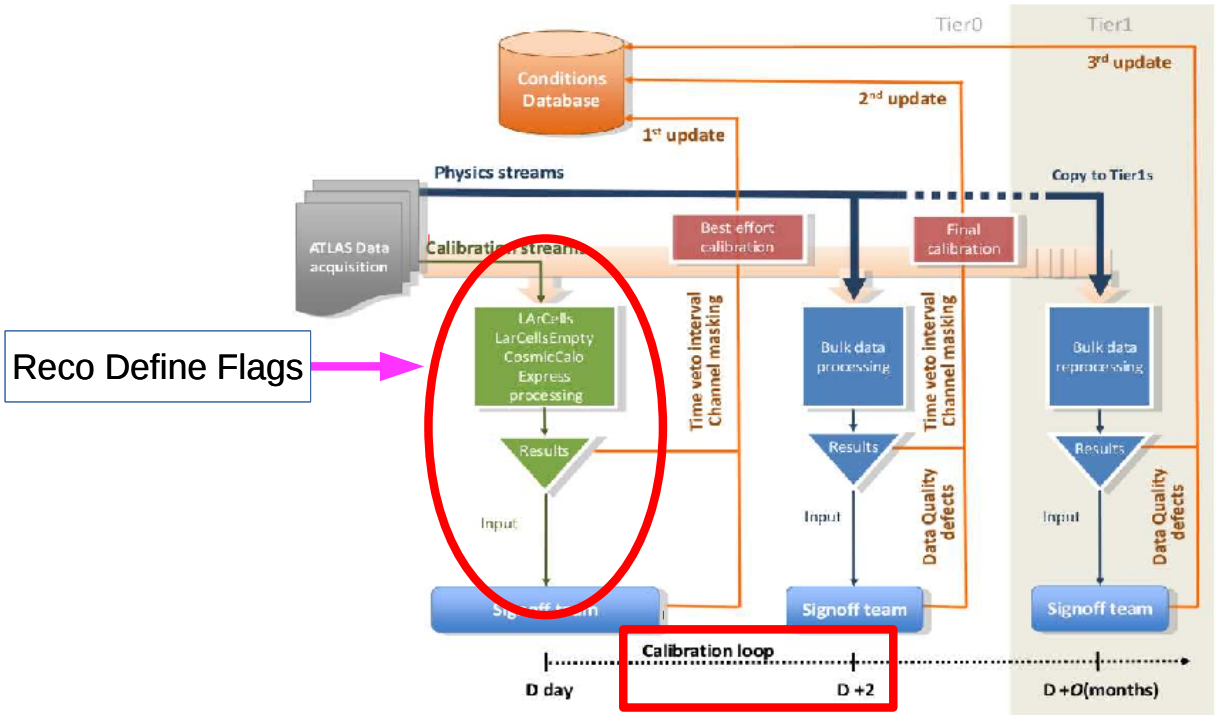
Dependance with luminosity

- A strong dependance with the instantaneous luminosity is observed

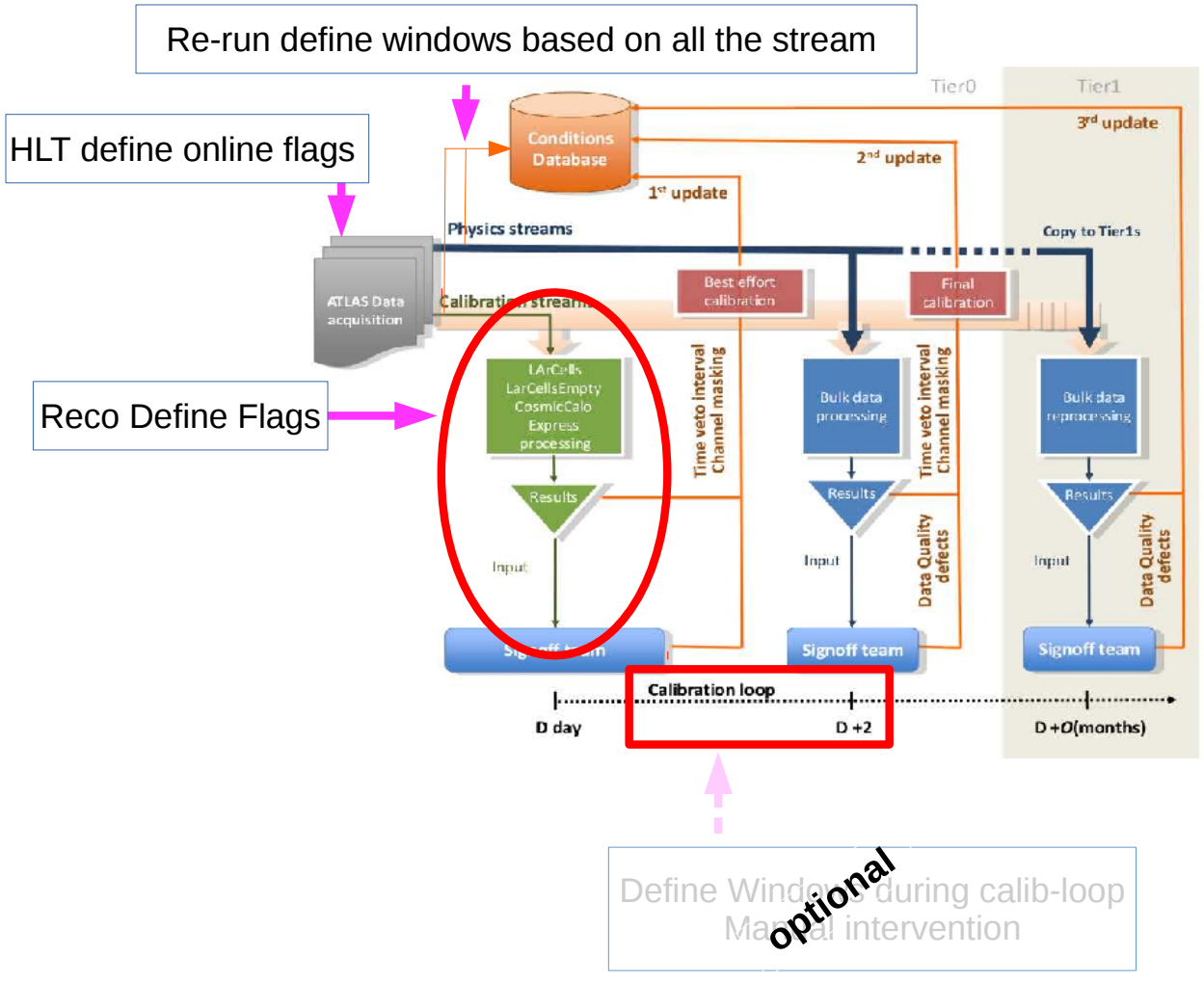


The online "Rerun algorithm"

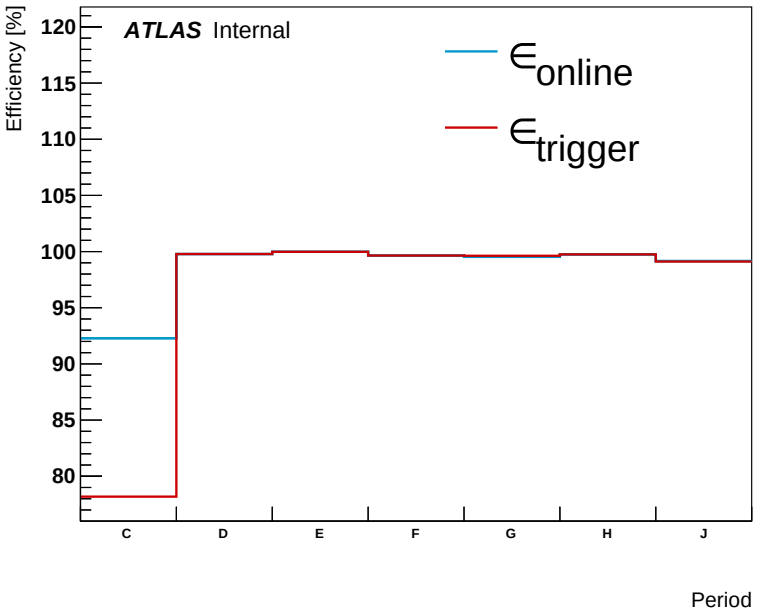
Online definition
Walter's Presentation

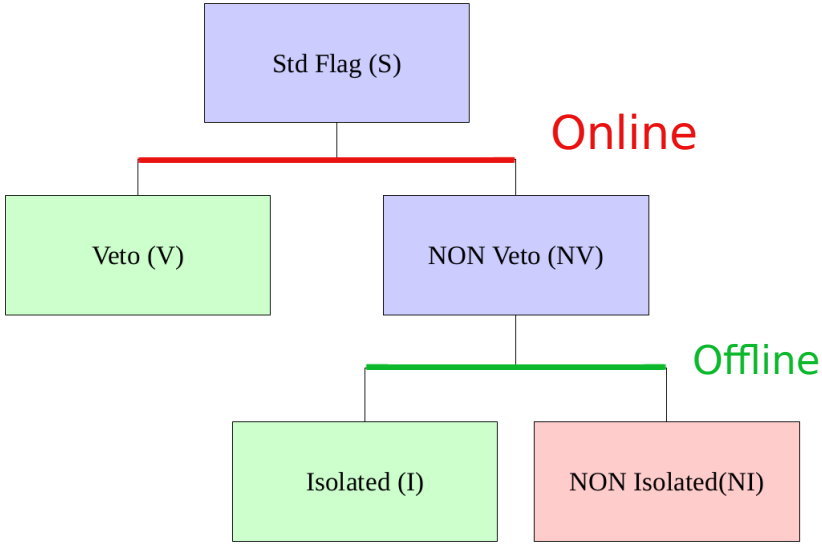


The online "Rerun algorithm"



Efficiency of the online procedure



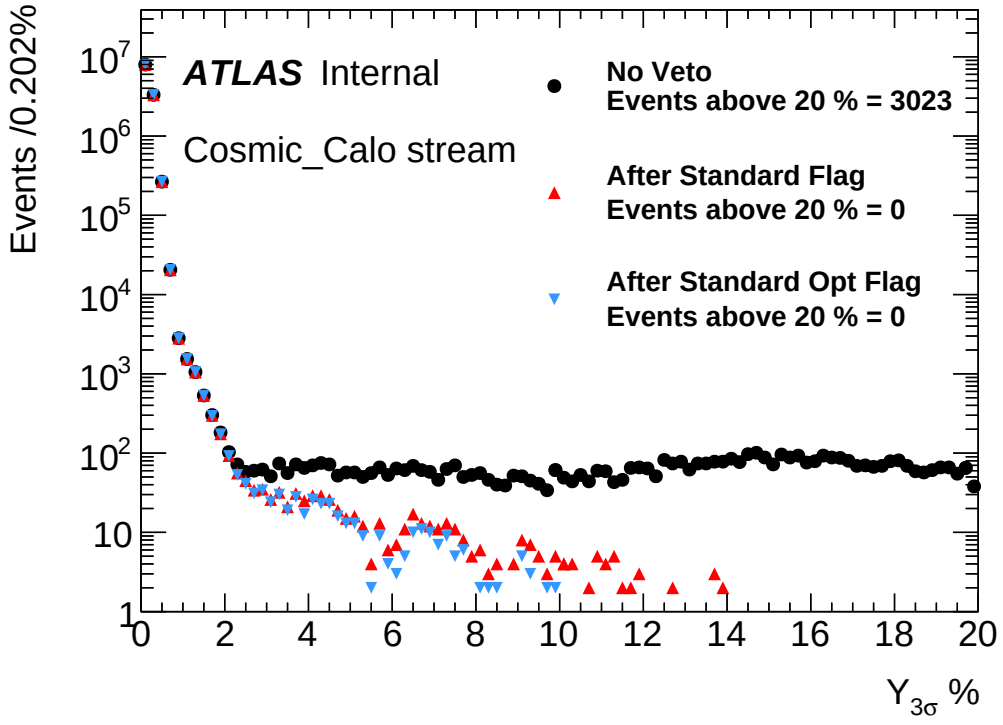


$\epsilon_{online} = \frac{V}{V+NI}$ It quantify if the online algorithm works as well as the Run-1 algo.

$\epsilon_{trigger} = \frac{V+NI}{V+NI+I}$ Assuming it exists no isolated events, it quantify how the trigger mangae to catch NB.

Flagging efficiency

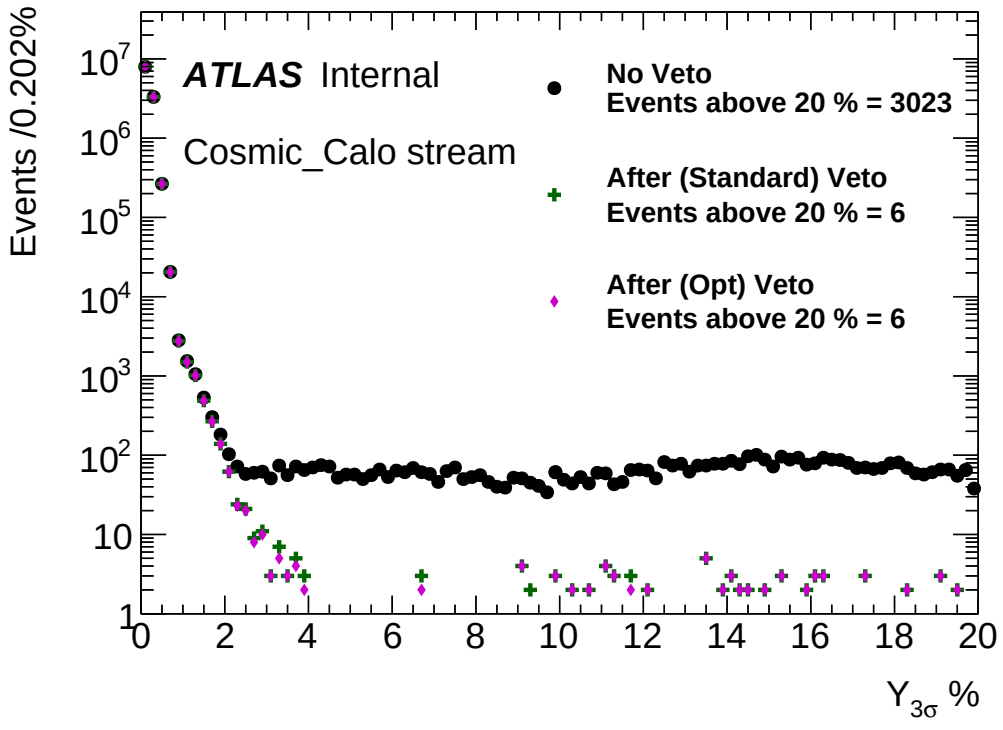
- The $Y_{3\sigma}$ distribution after the flag represent how the noise burst are found



- The **optimized flag** clean the tail a bit better

Veto efficiency

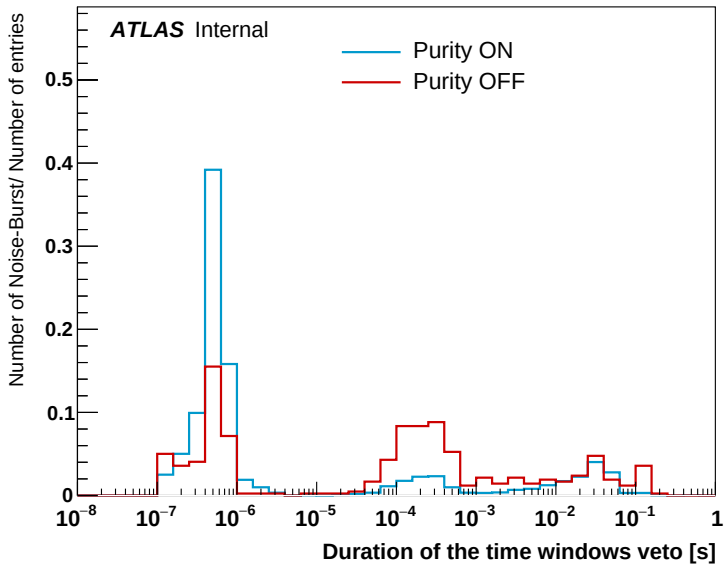
- The $Y_{3\sigma}$ distribution after the veto is the effective removal of the Noise-Burst



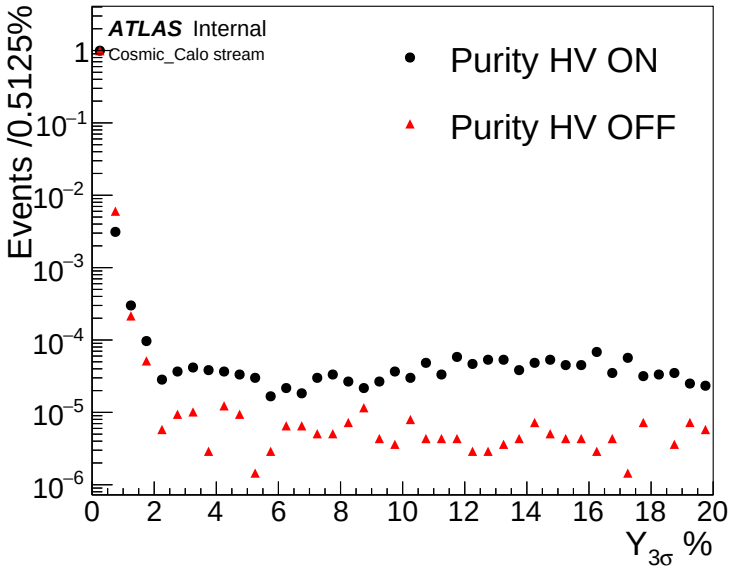
- No significant change is observed with the **optimized flag**
- Between 2-10% the tails are better cleaned
- But it remains a small tails up to few percent (trigger issue ?)

Characteristic of the remaining NB

With purity probes of it is reduced BUT NON ZERO



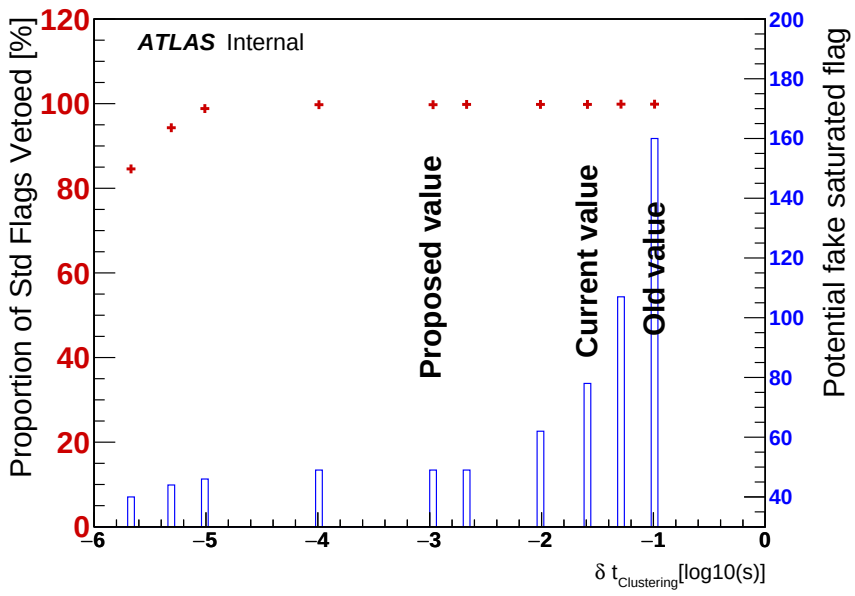
- The short NB are well supressed



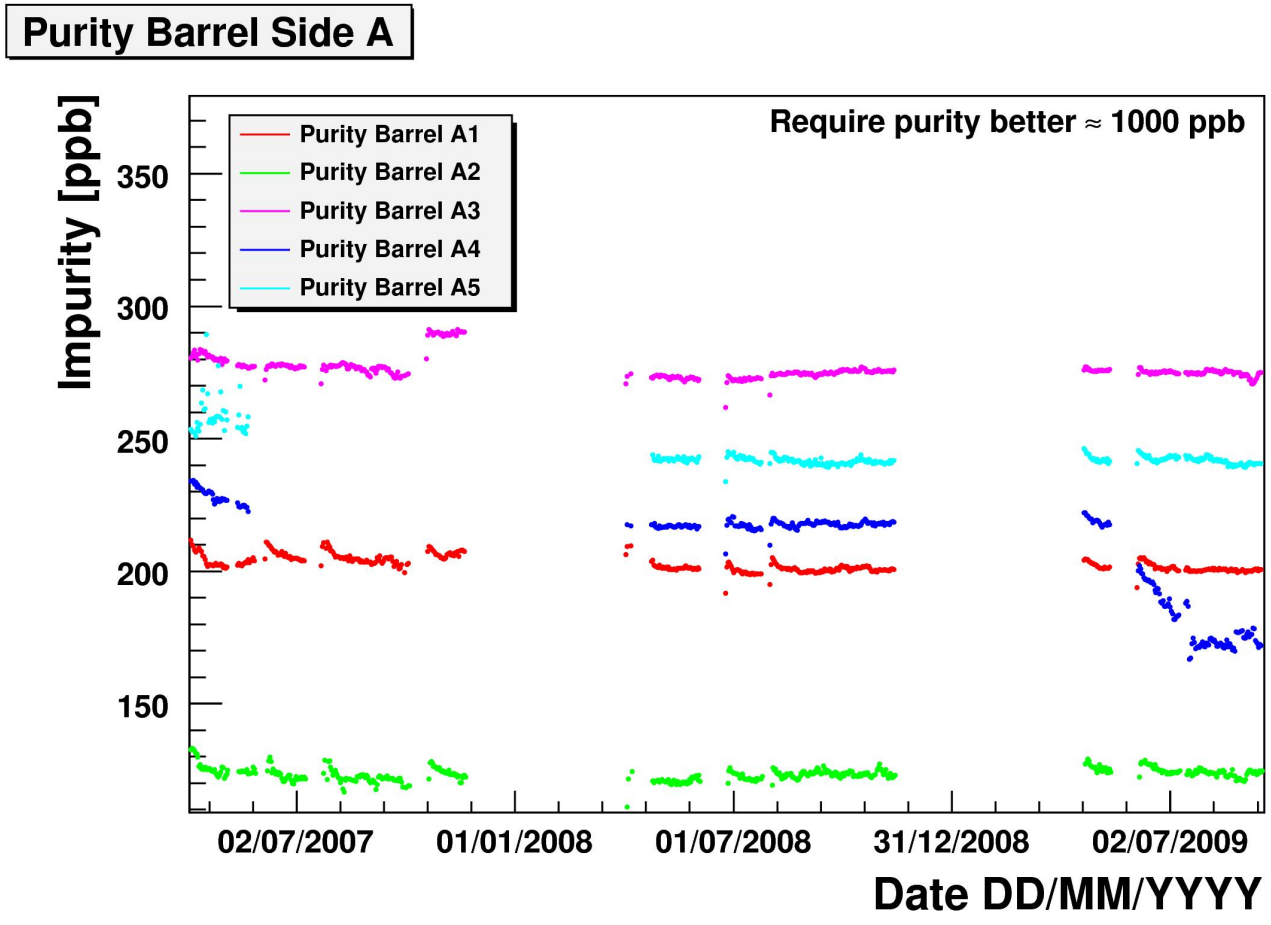
- The amplitude is still large

Next step/Recommandation

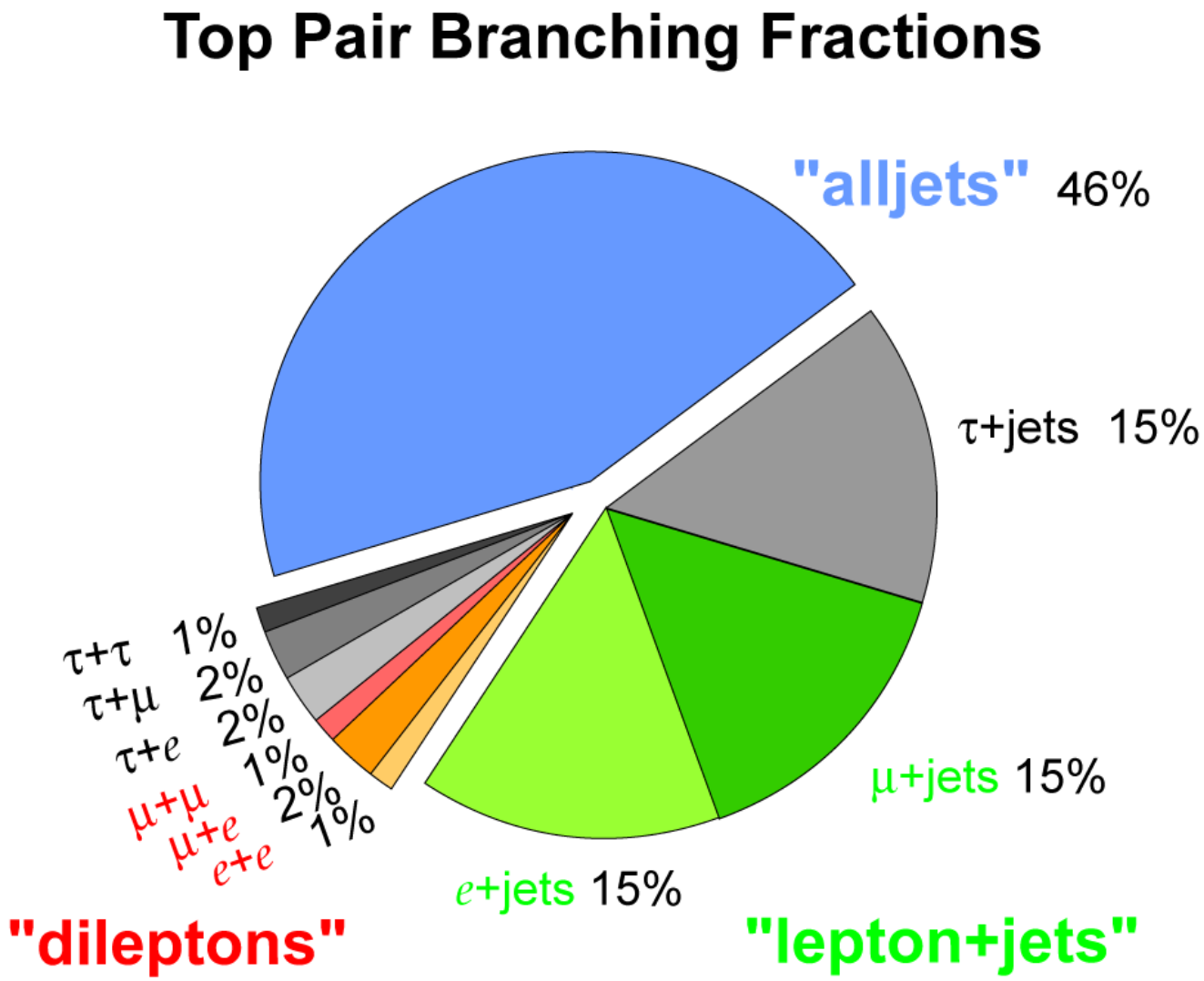
- Study saturated flag -> Maybe find better definition
- Keep an eye on teh NB rate -> Include the plot on weekly DQ plot
- May study $\frac{N_{Cells}^{Q>4000}}{N_{cells}^{Qavailable}}$
- Reduce clustering time $\delta t_{Clustering} \rightarrow 1\ \mathrm{ms}$



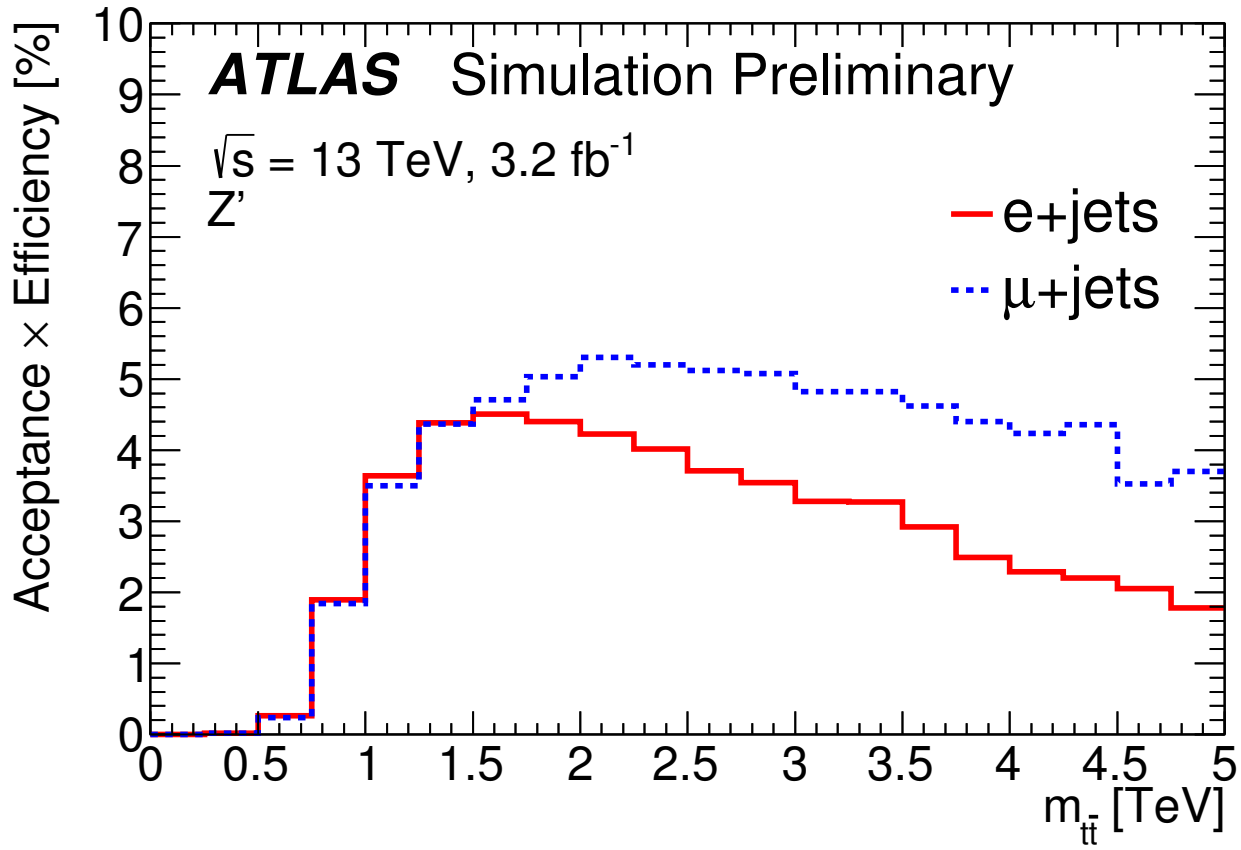
Purity on LAR calorimeters



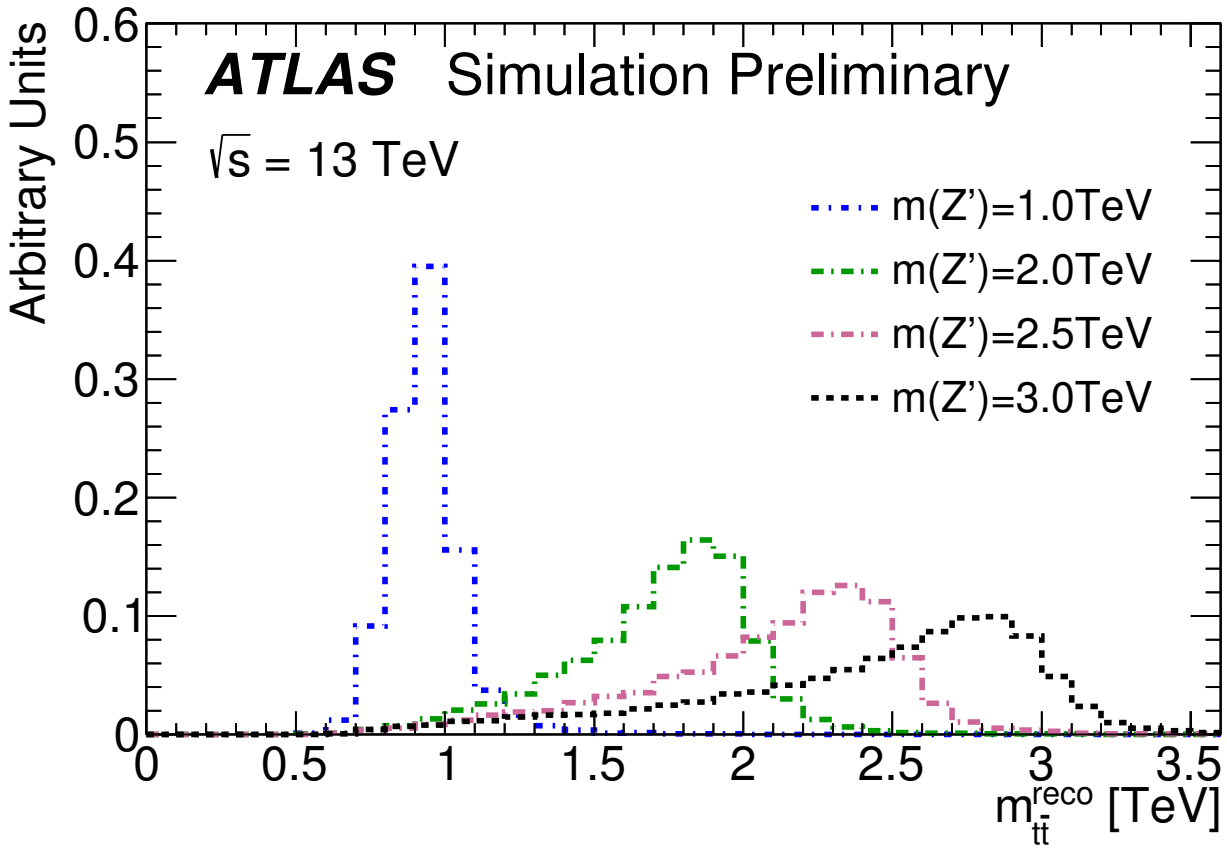
ttbar decays



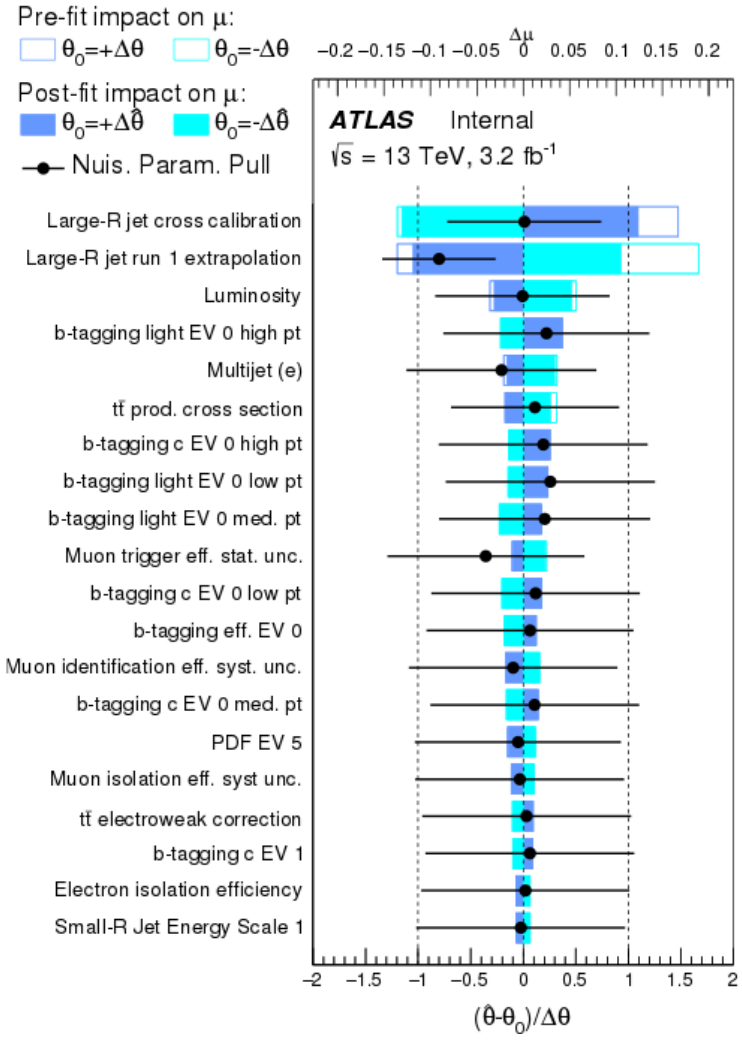
efficiency


ATLAS Simulation Preliminary
√s = 13 TeV, 3.2 fb-1
Z'
e+jets
μ+jets
Acceptance × Efficiency [%]
m_tt [TeV]
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
0
1
2
3
4
5
6
7
8
9
10

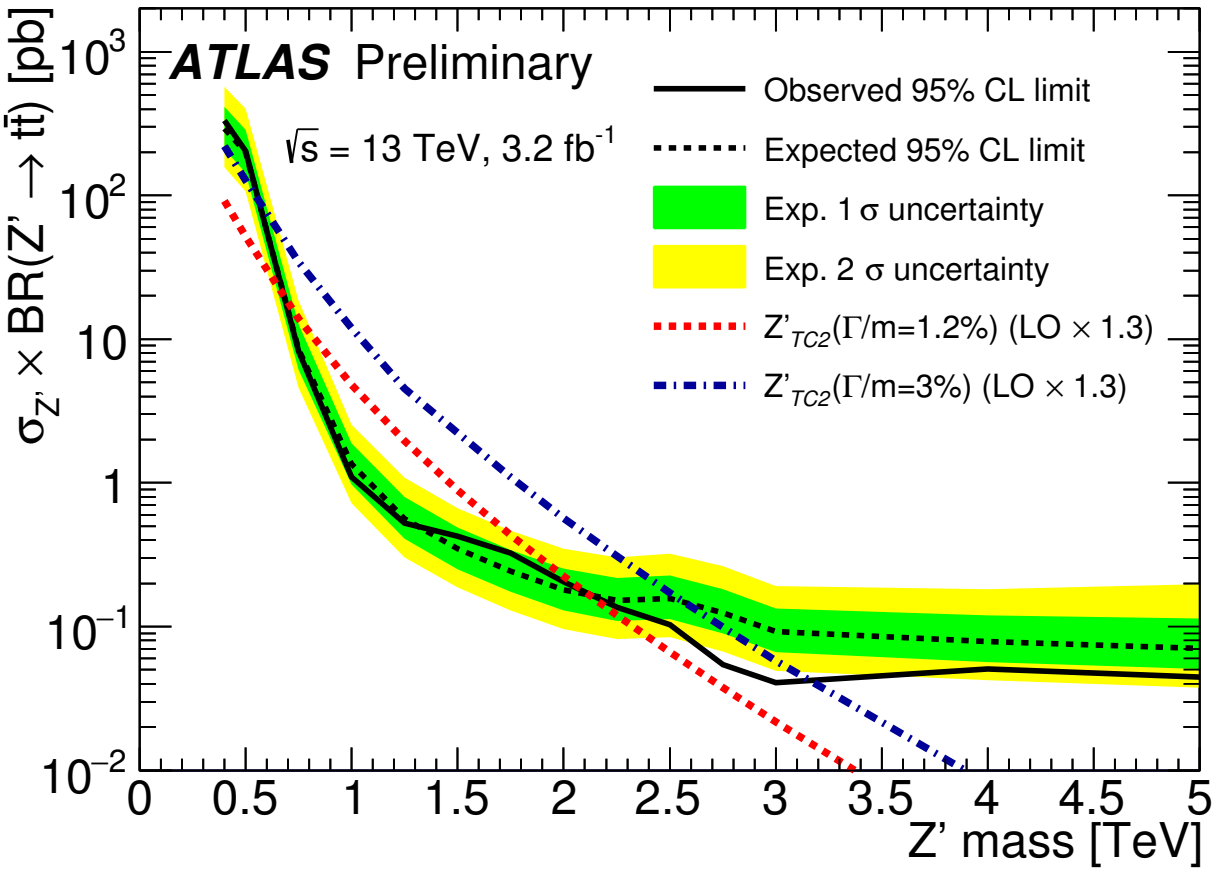
Signal shape



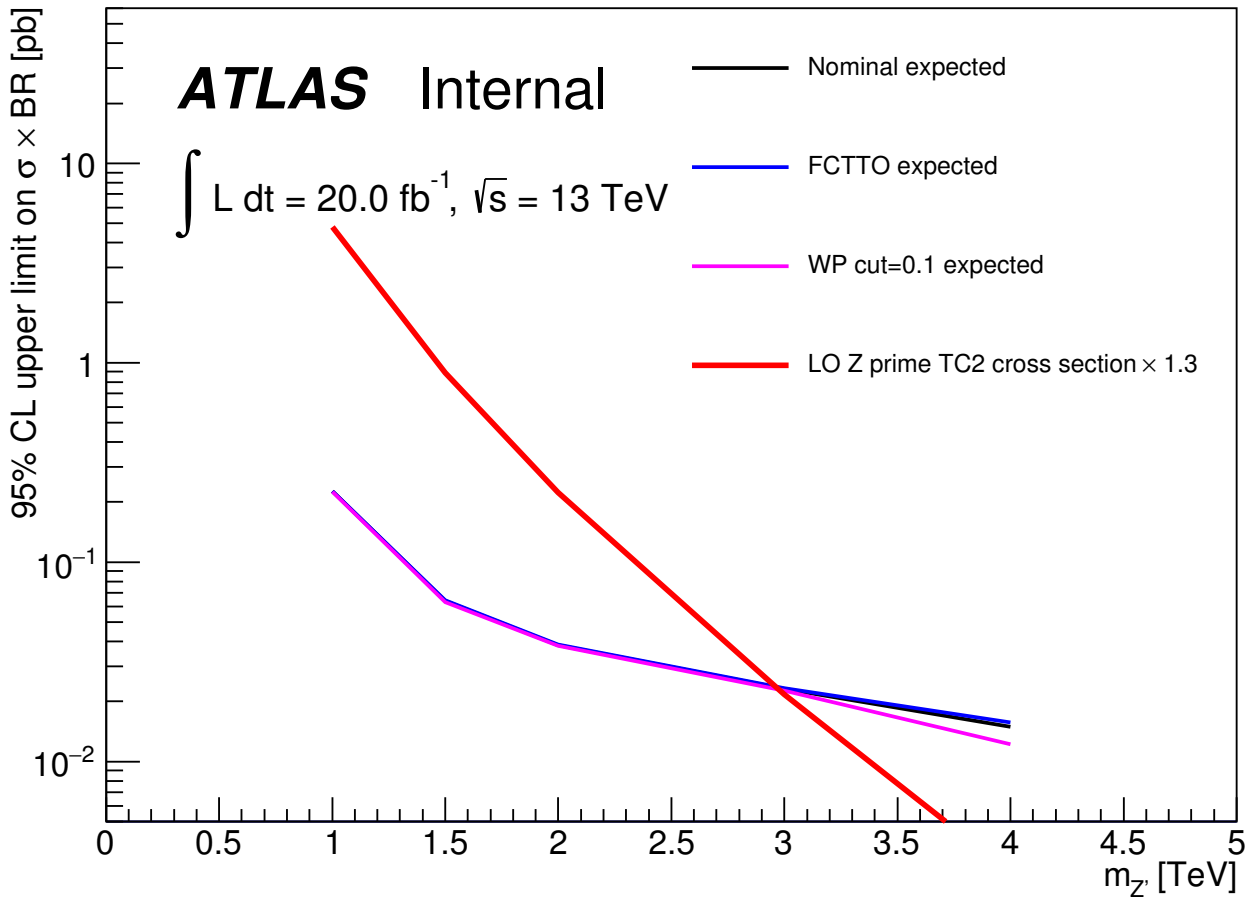
Uncertainties



Limits on ttbar



IMpact of isolation on ttbar



Limits on ttbar