Search for emerging jets in the ATLAS detector and reinterpretation of the LHC results

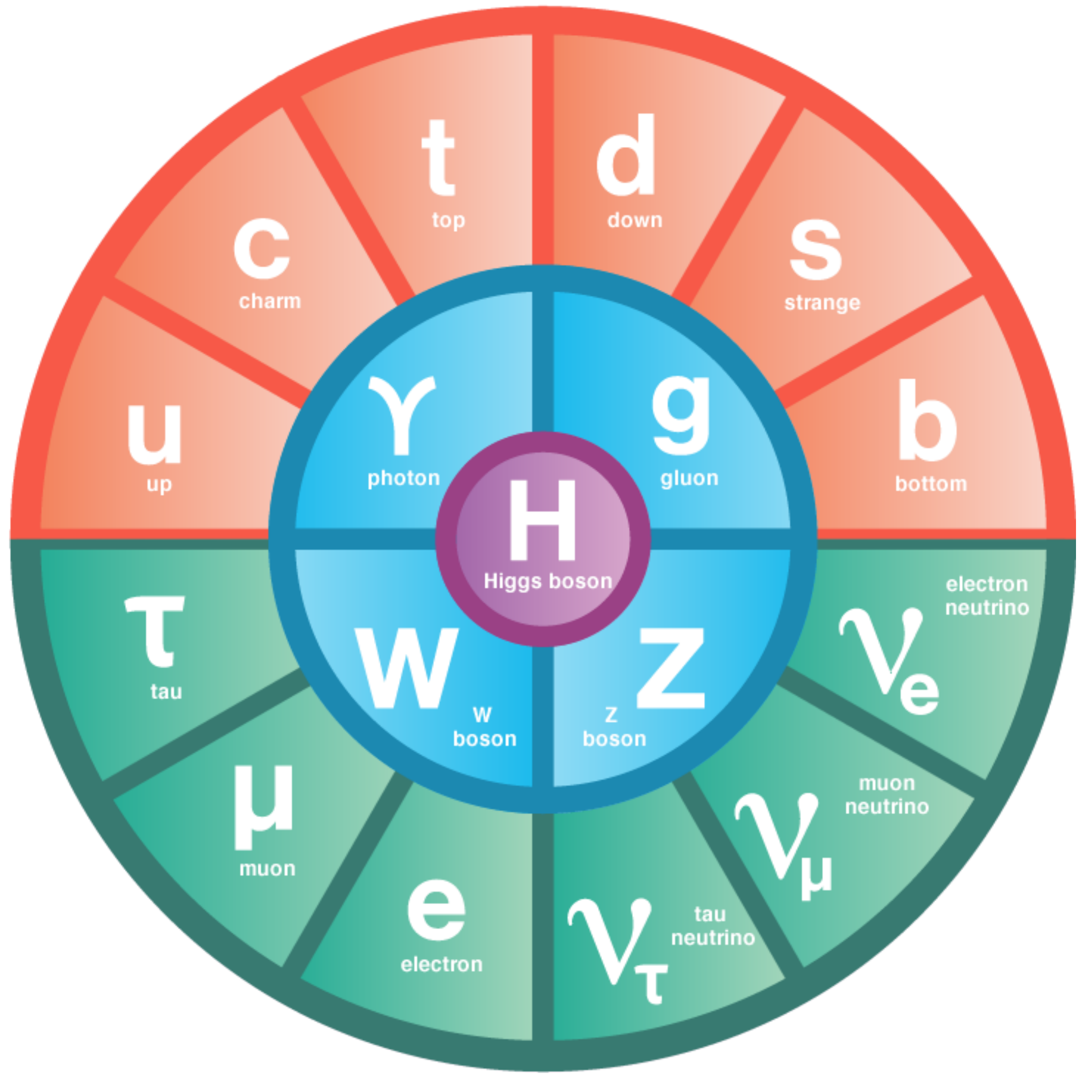
PhD seminar

Thomas WOJTKOWSKI

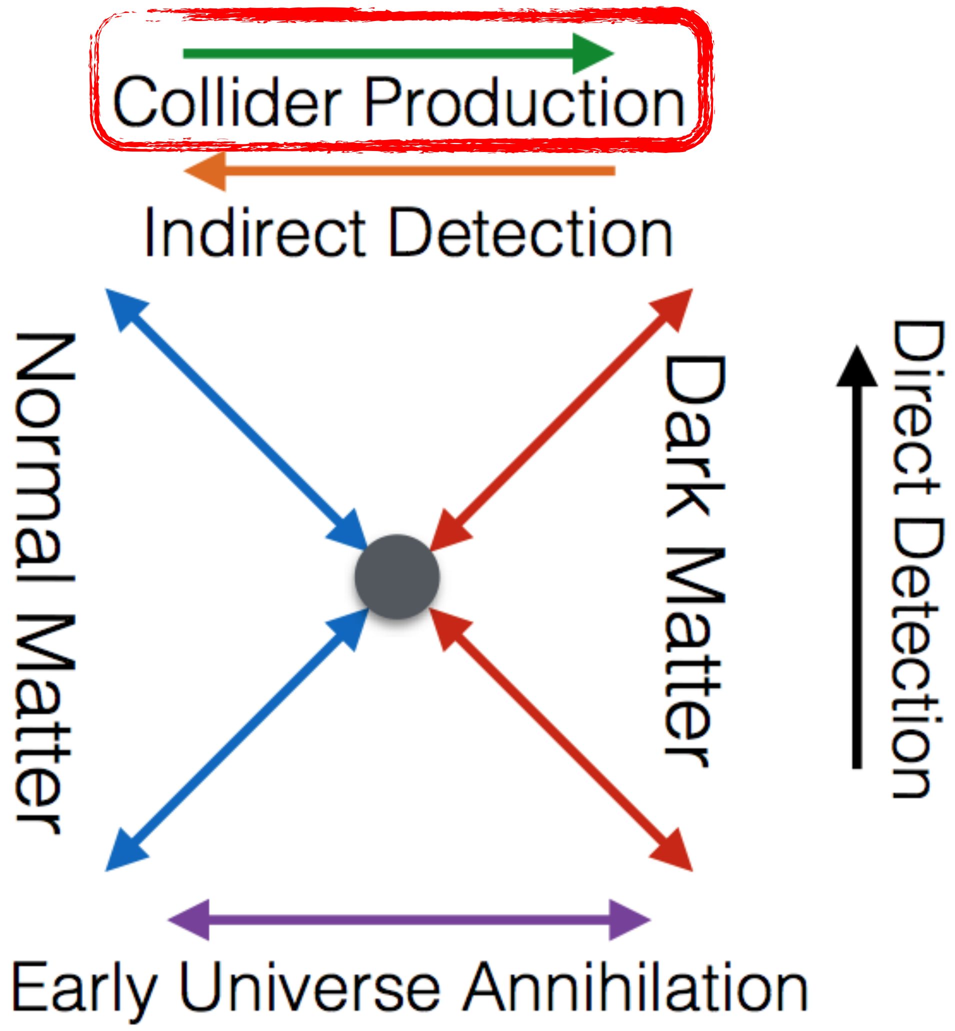
Supervisors : Pierre-Antoine DELSART and Marie-Hélène GENEST

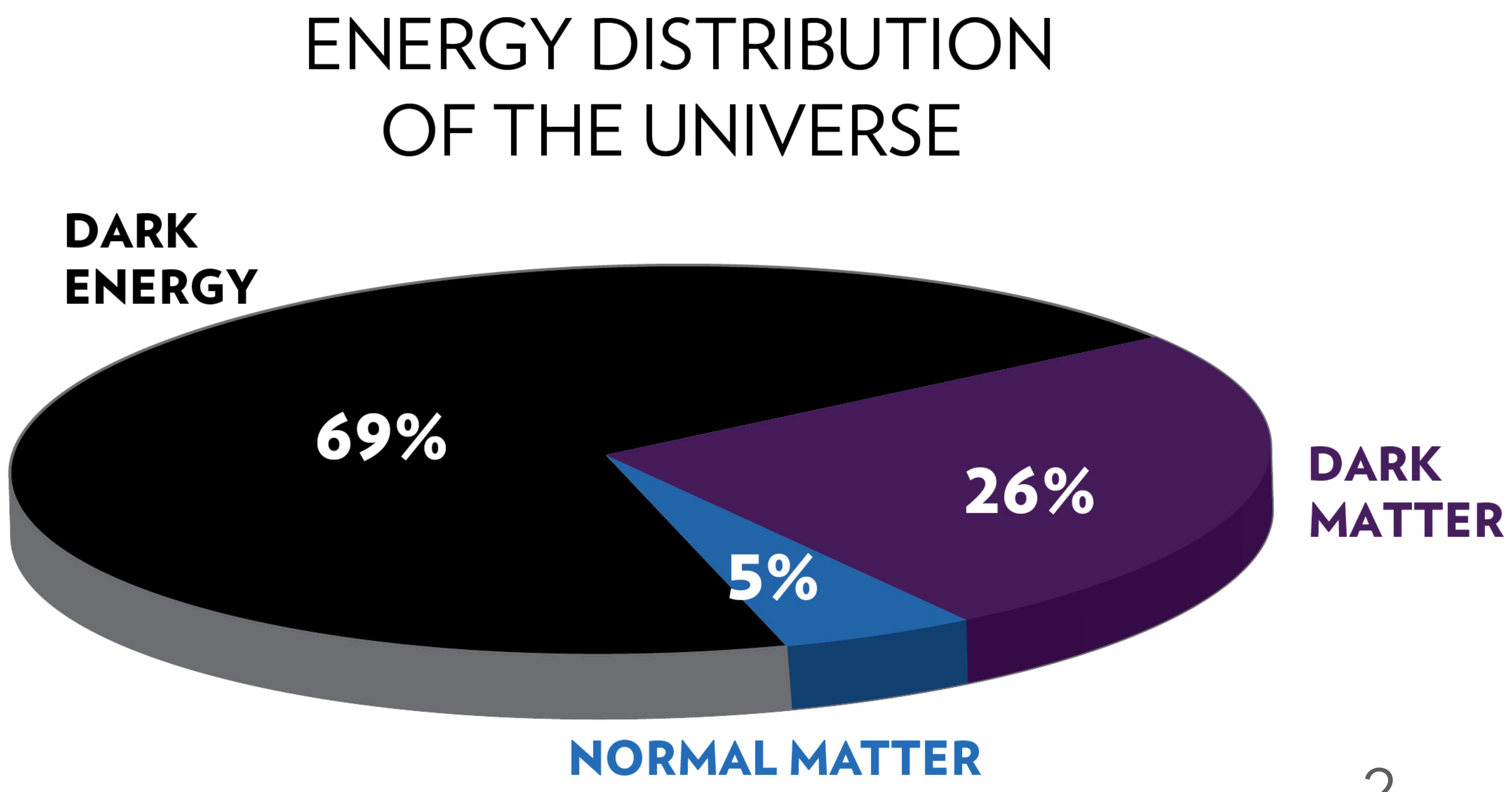
Context

Search for physics beyond Standard Model at the LHC



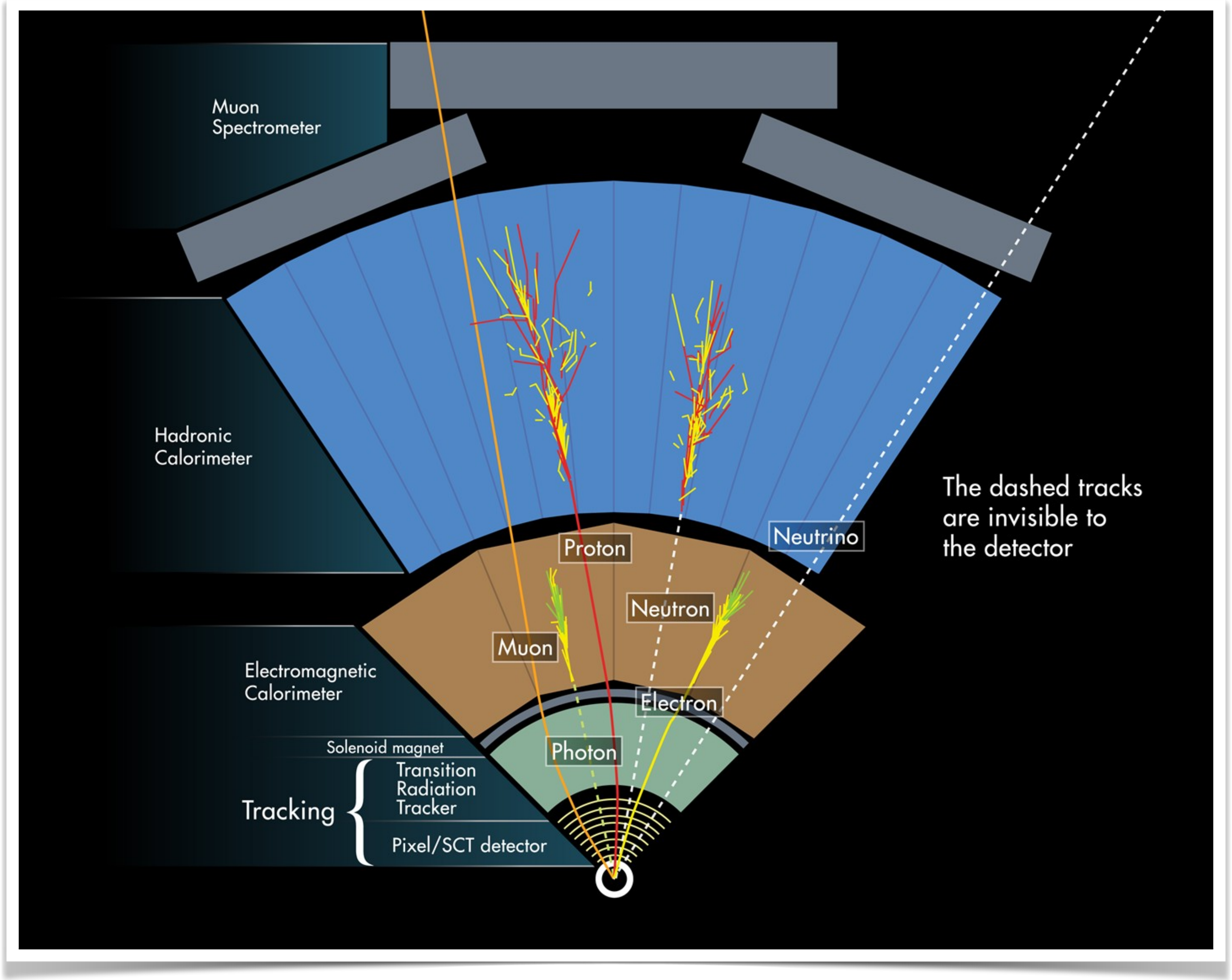
- Standard Model (SM) : only 5% of the energy-mass in the Universe
- **Dark Matter** :
  - manifestation through gravitational effects, but unknown nature
  - hypothesis : could be new massive particles weakly interacting with SM ones
- Large Hadron Collider : p-p collisions at 13.6 TeV total energy
  - possible production of dark matter particles through very rare process

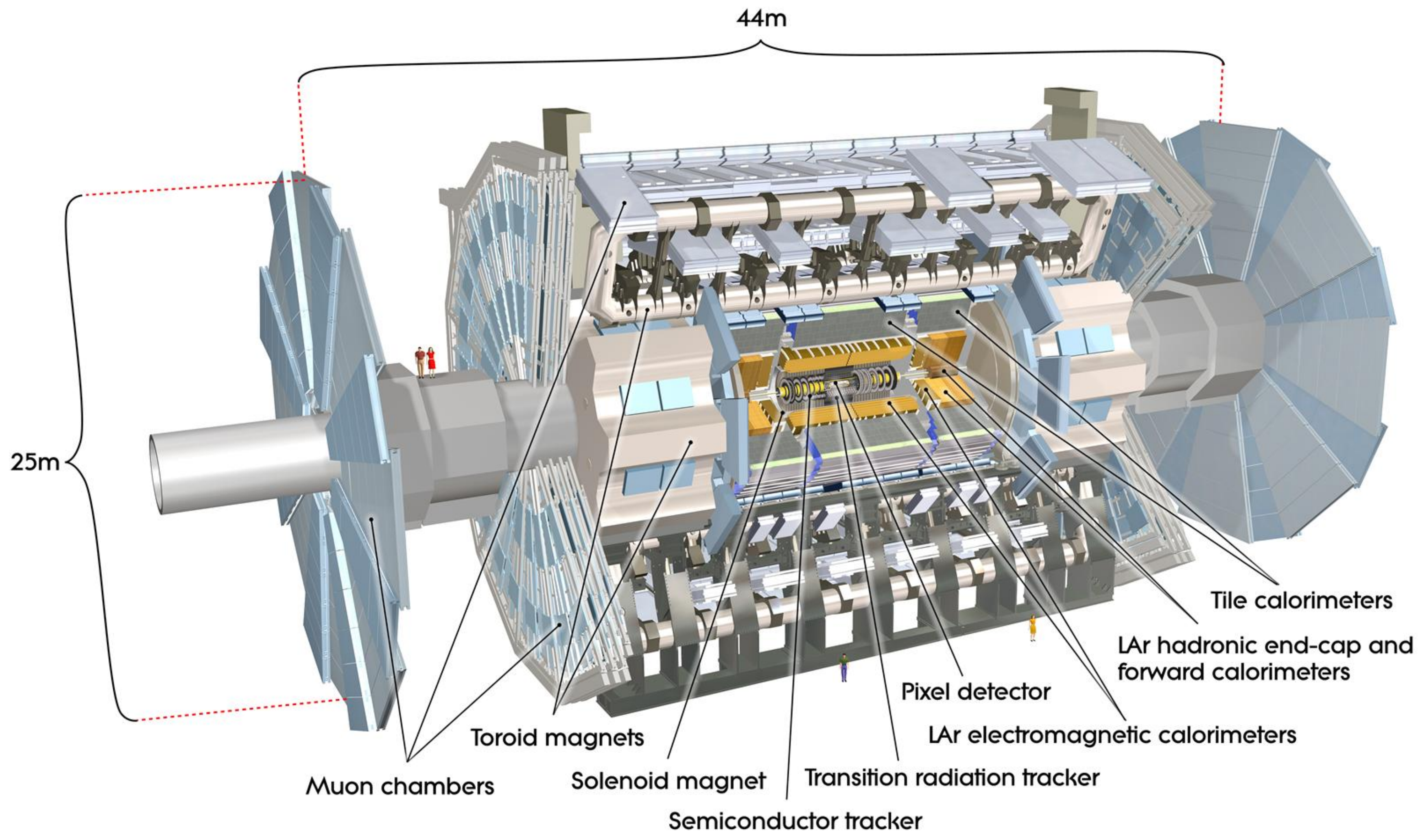




ATLAS detector

- General purpose detector : SM, search for new physics
- Structure in layers :
  - **inner detector** : *track* (trajectory and momentum of a charged particle curved by magnet system)
  - **calorimeters** : *cluster* (particle energy deposition except for μ and ν)
  - **muon spectrometer** : muon trajectory and momentum

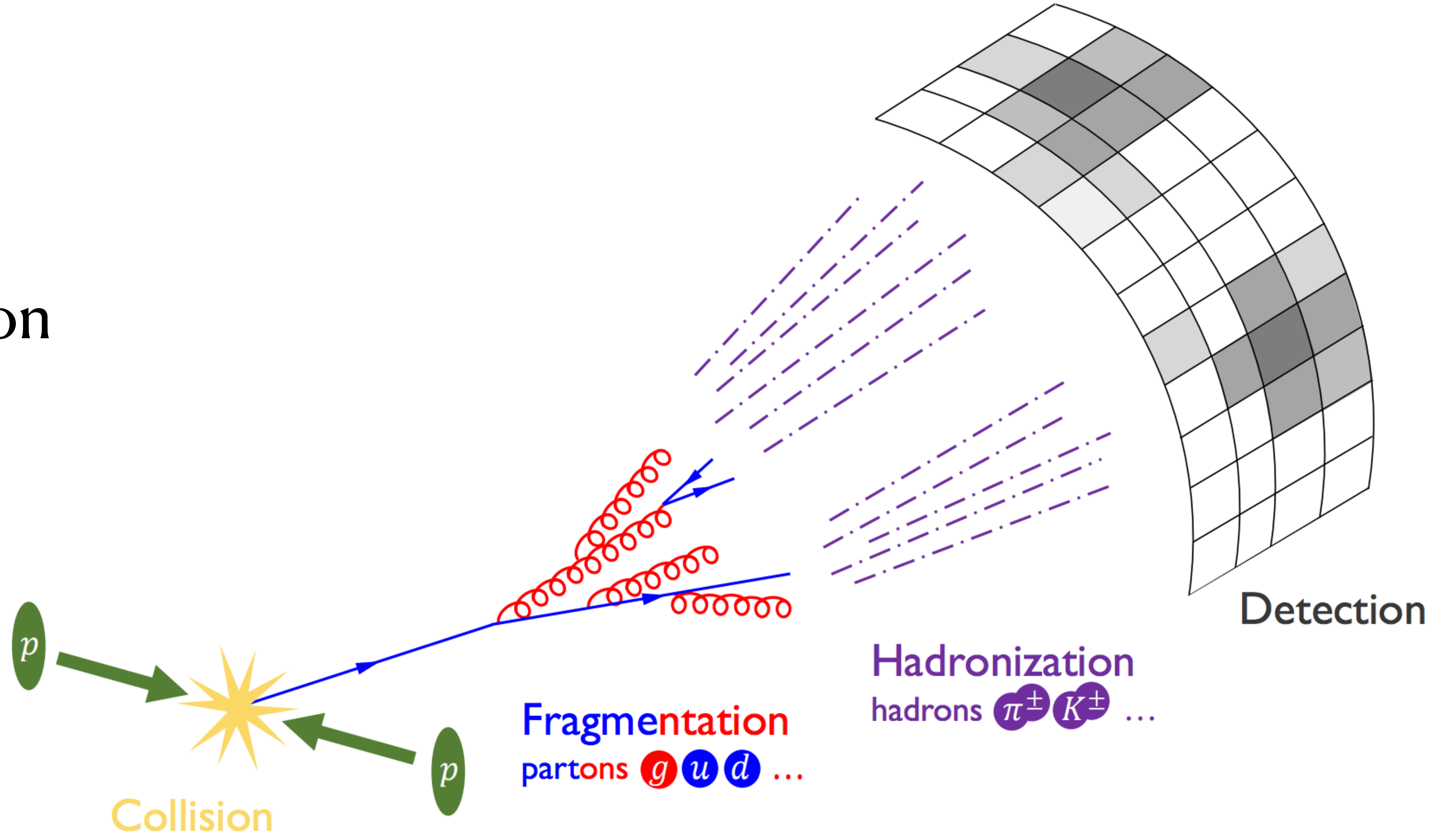


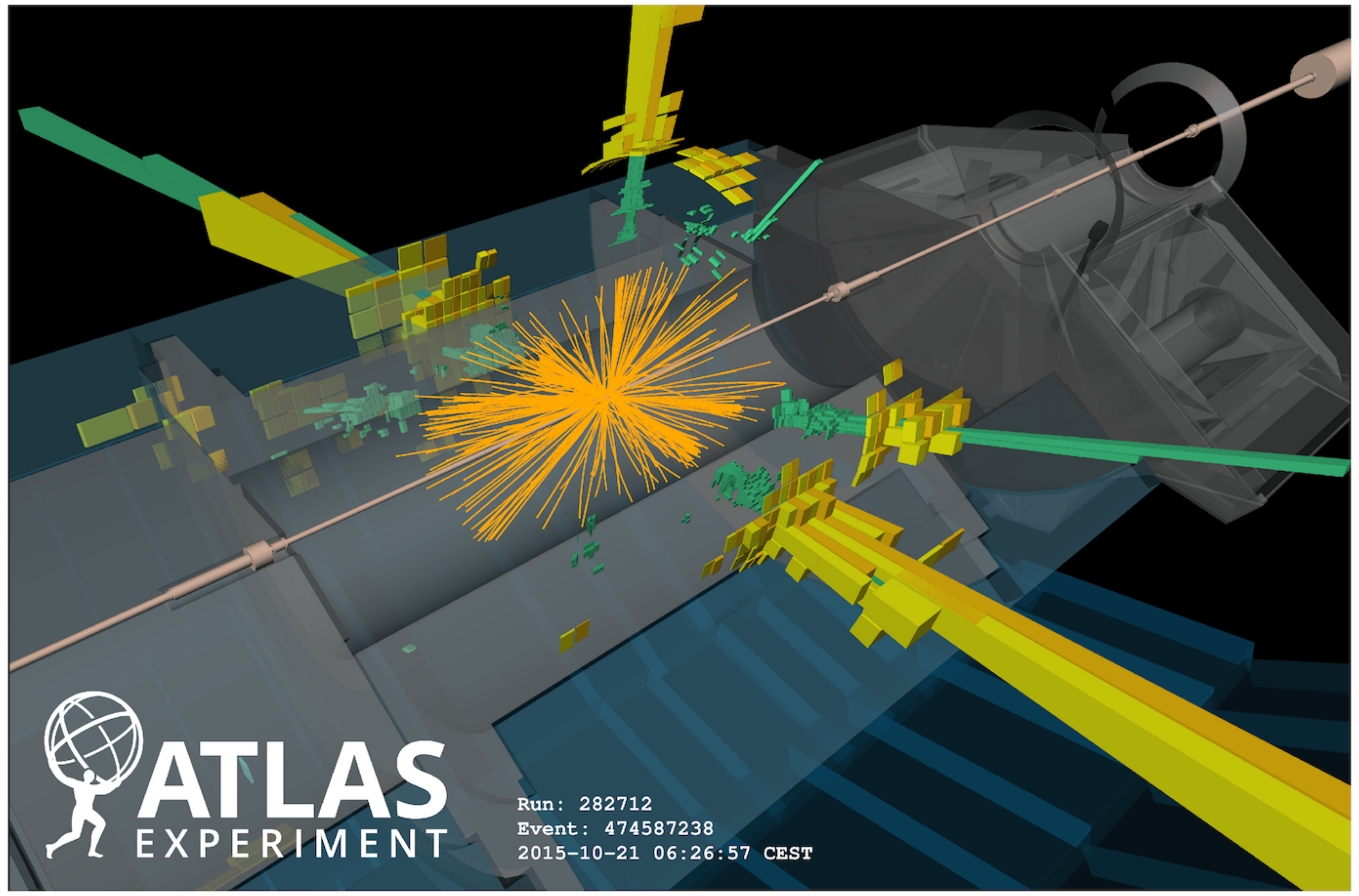


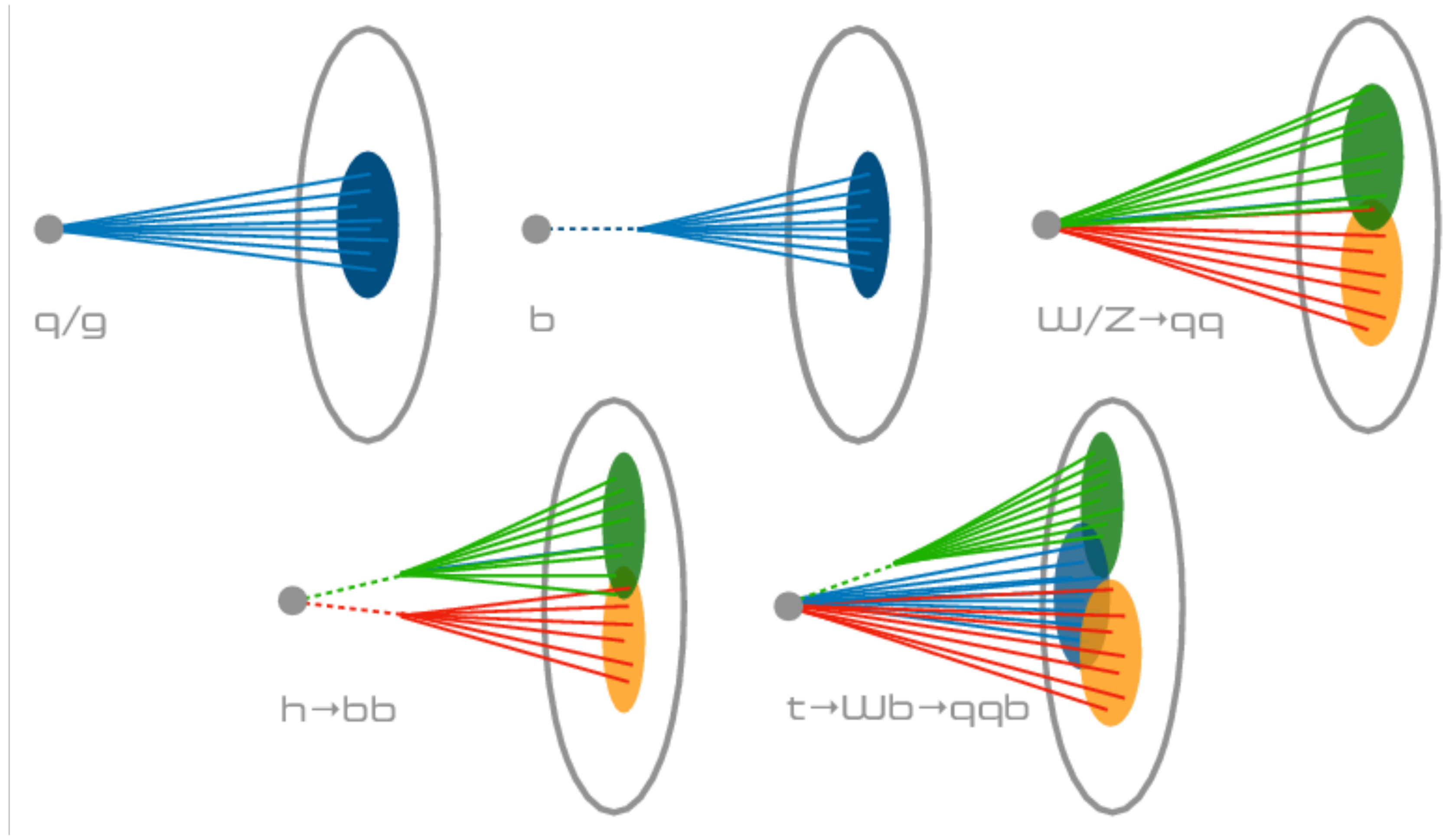
- 40.10^6 beam crossings /s : **trigger system**, 1000 events /s stored for analysis
  - must be very well configured
- Offline event reconstruction : signals turn into physical object (jets, leptons, photons ...)

Hadronic jets

- **QCD processes** :
  - pp collision, emission of high energy parton
  - parton shower : collinear partons emitted
  - hadronization : gathering of partons to form hadrons
- Jet : cone of produced hadrons
- Different jet topologies (q/g, top ...)







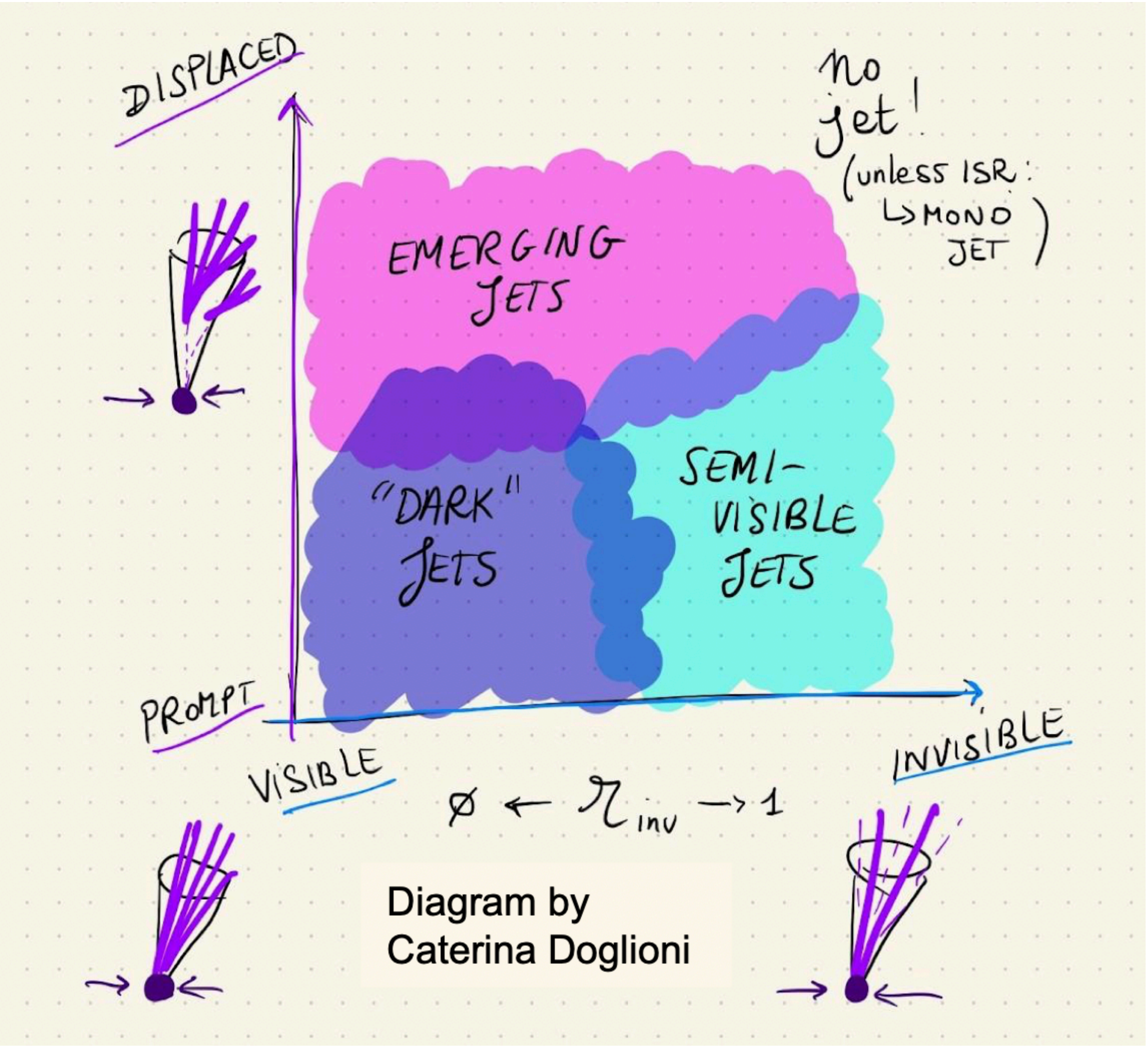
- In ATLAS, different ways to reconstruct jet constituents using tracks and clusters
- Jet algorithms regroup constituents

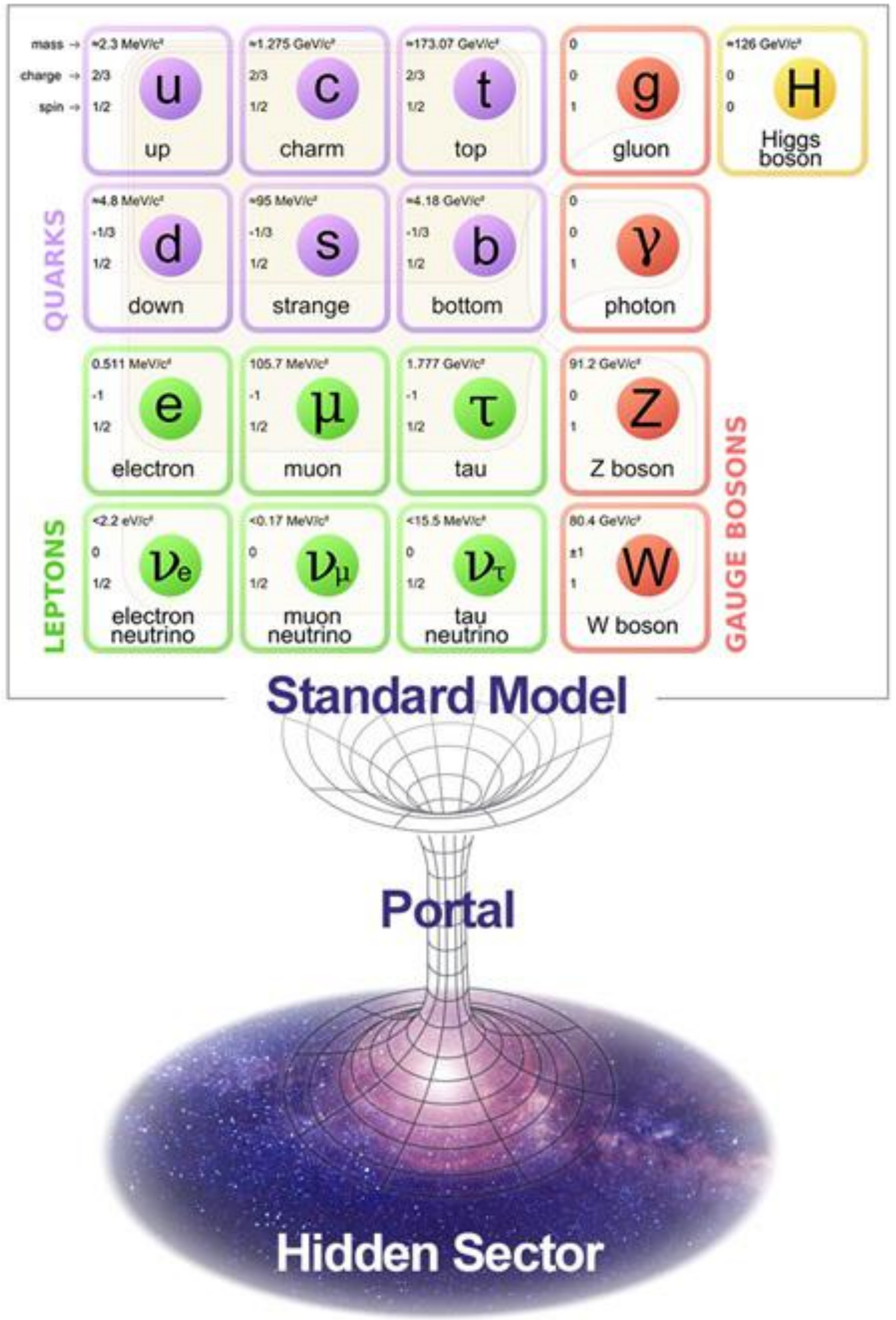
Hidden sector

- Extension SM : **QCD-like hidden dark sector**
  - dark quarks q_d
  - dark gluons g_d

$$\mathcal{L}_d = \bar{q}'_i(i\not{D} - m_{q'_i})q'_i - \frac{1}{4}G'^{\mu\nu}G'_{\mu\nu}$$

- Parton shower and hadronization in dark sector → jet of dark hadrons
- Stable particle : DM candidates
- **Portal SM - hidden sector,** new interaction
  - q_d production in pp collisions
  - dark hadrons decay to SM quarks, forming jets : dark, semi-visible or **emerging**



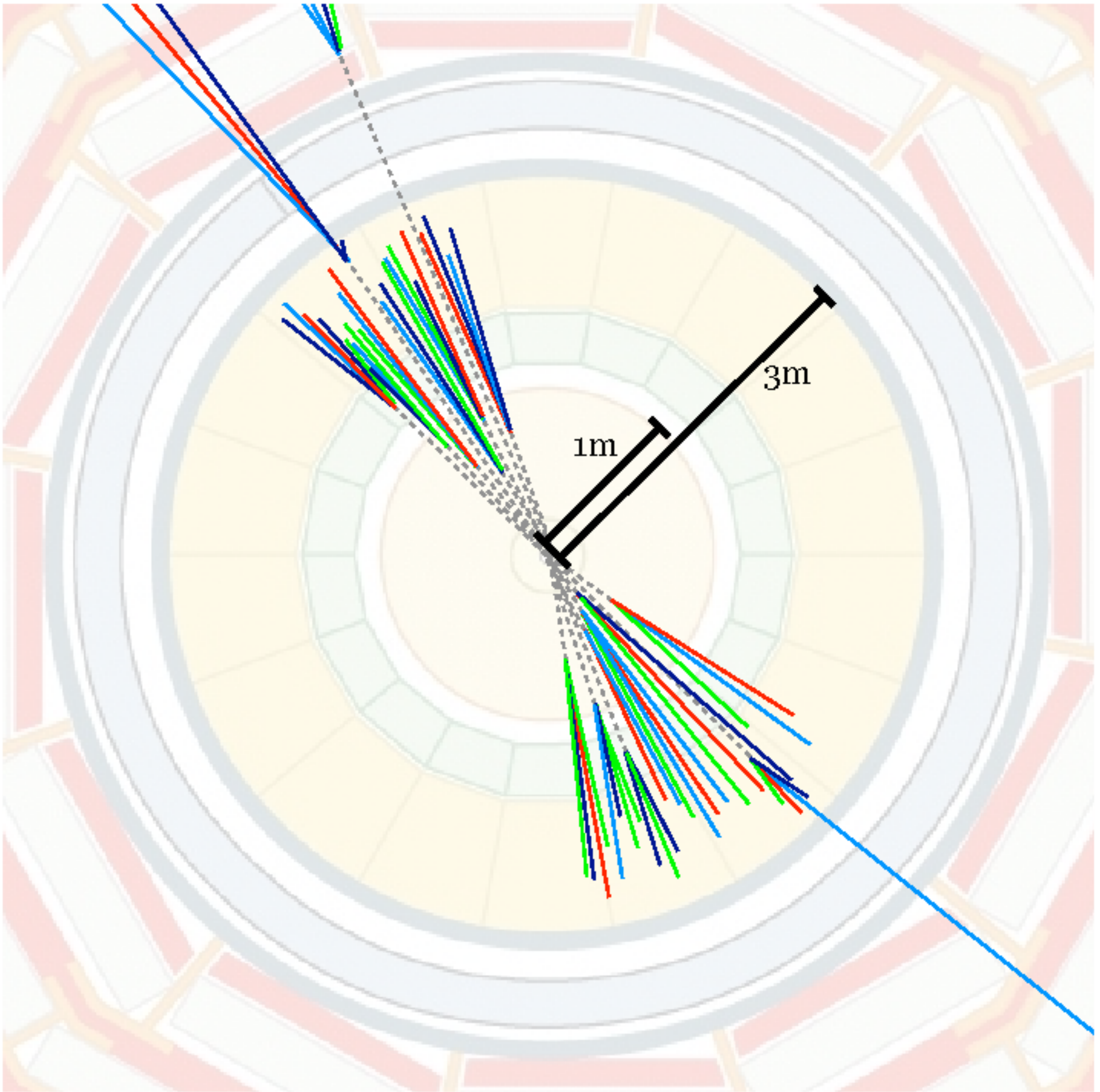


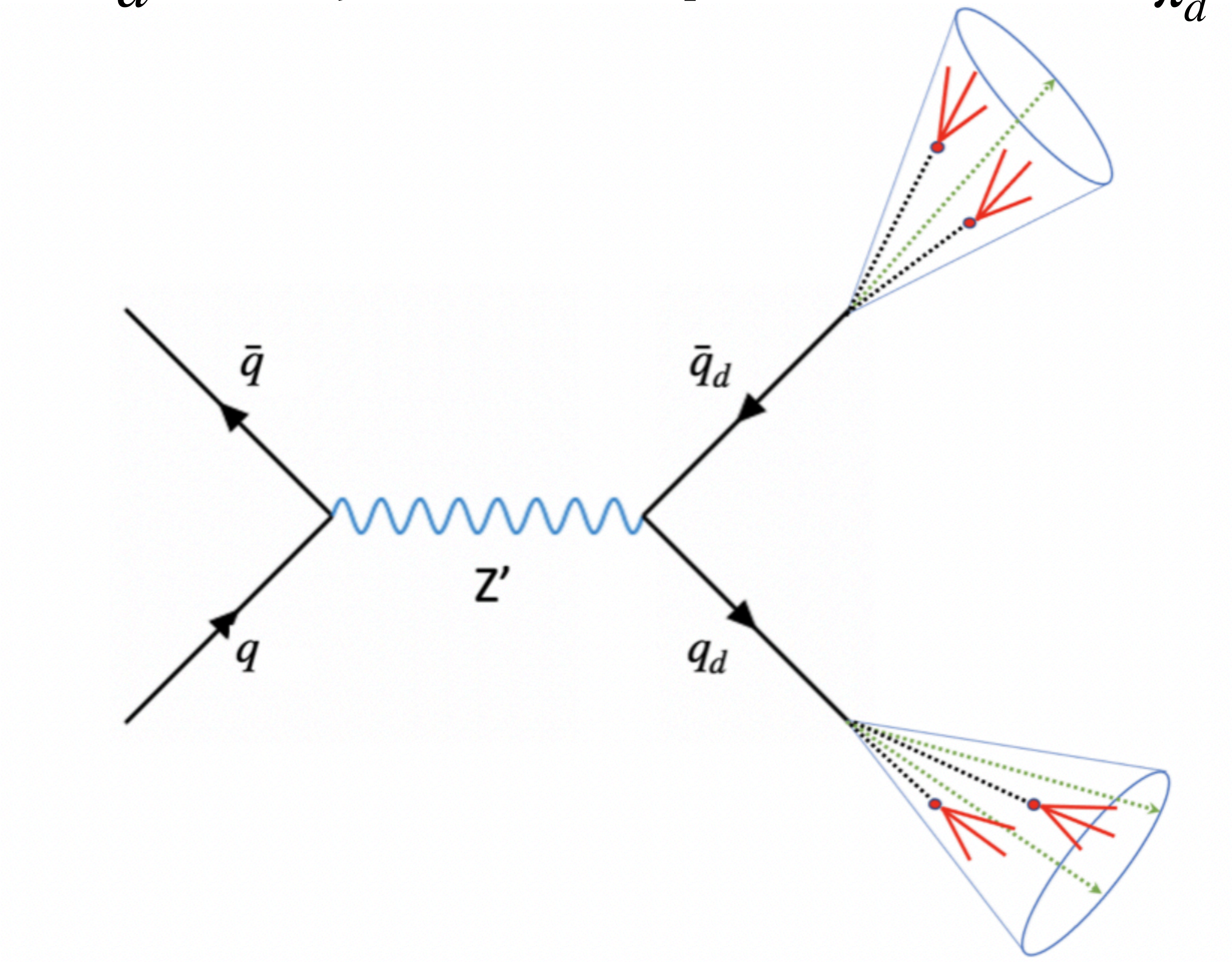
Emerging jets

- Model considered :
  - q_d production via **new Z' mediator**

$$\mathcal{L}_{\text{med}} = -\frac{1}{4}Z'^{\mu\nu}Z'_{\mu\nu} - \frac{1}{2}M_{Z'}^2 Z'^{\mu}Z'_{\mu} + Z'_{\mu}(\bar{q}'_i\gamma^{\mu}q'_i + \bar{q}_j\gamma^{\mu}q_j)$$

  formation of jet containing unstable ρ_d and π_d :
  - ρ_d decays to π_d
  - π_d decays to SM quarks with $c\tau_{\pi_d} \sim$ mm



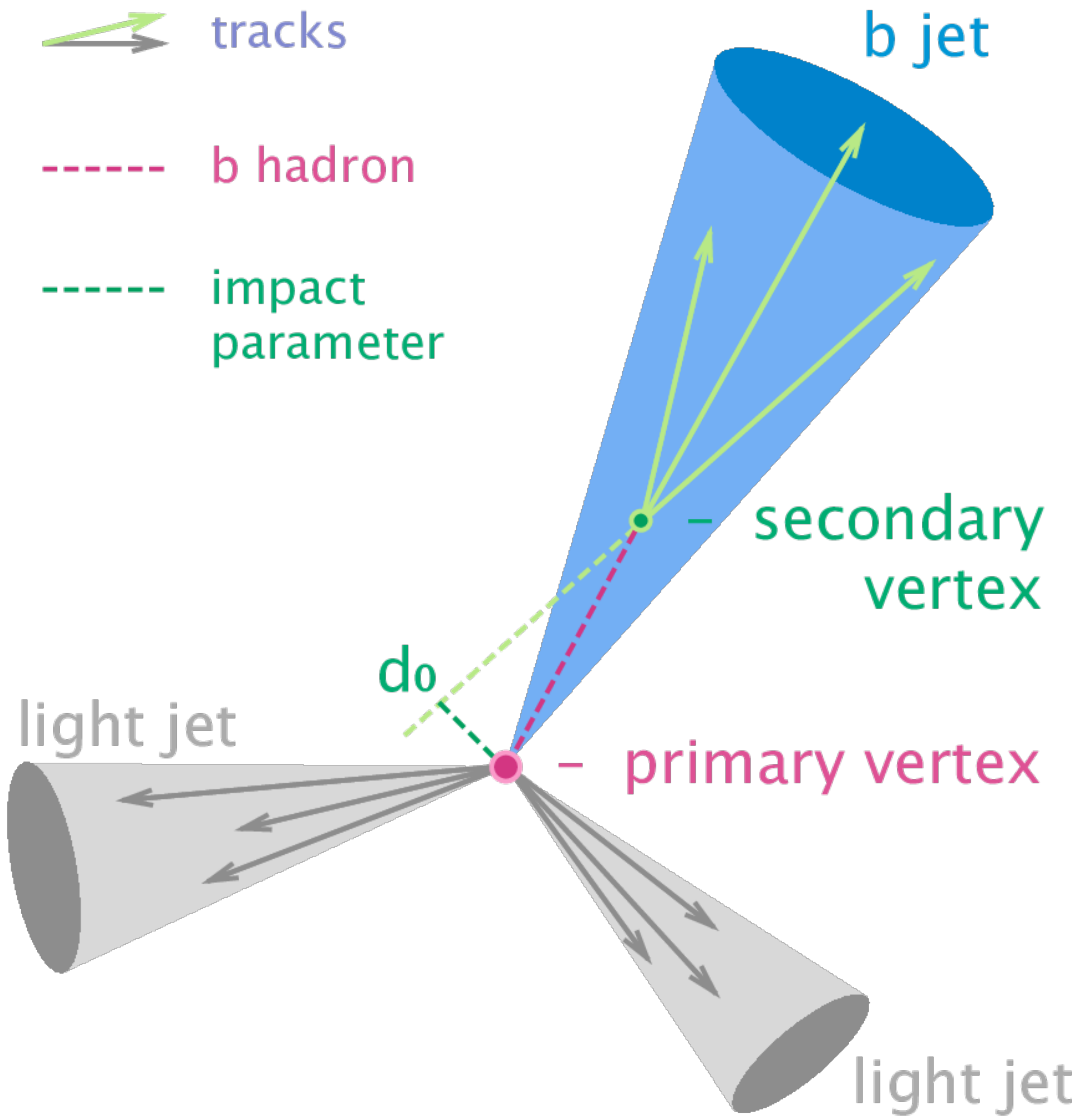


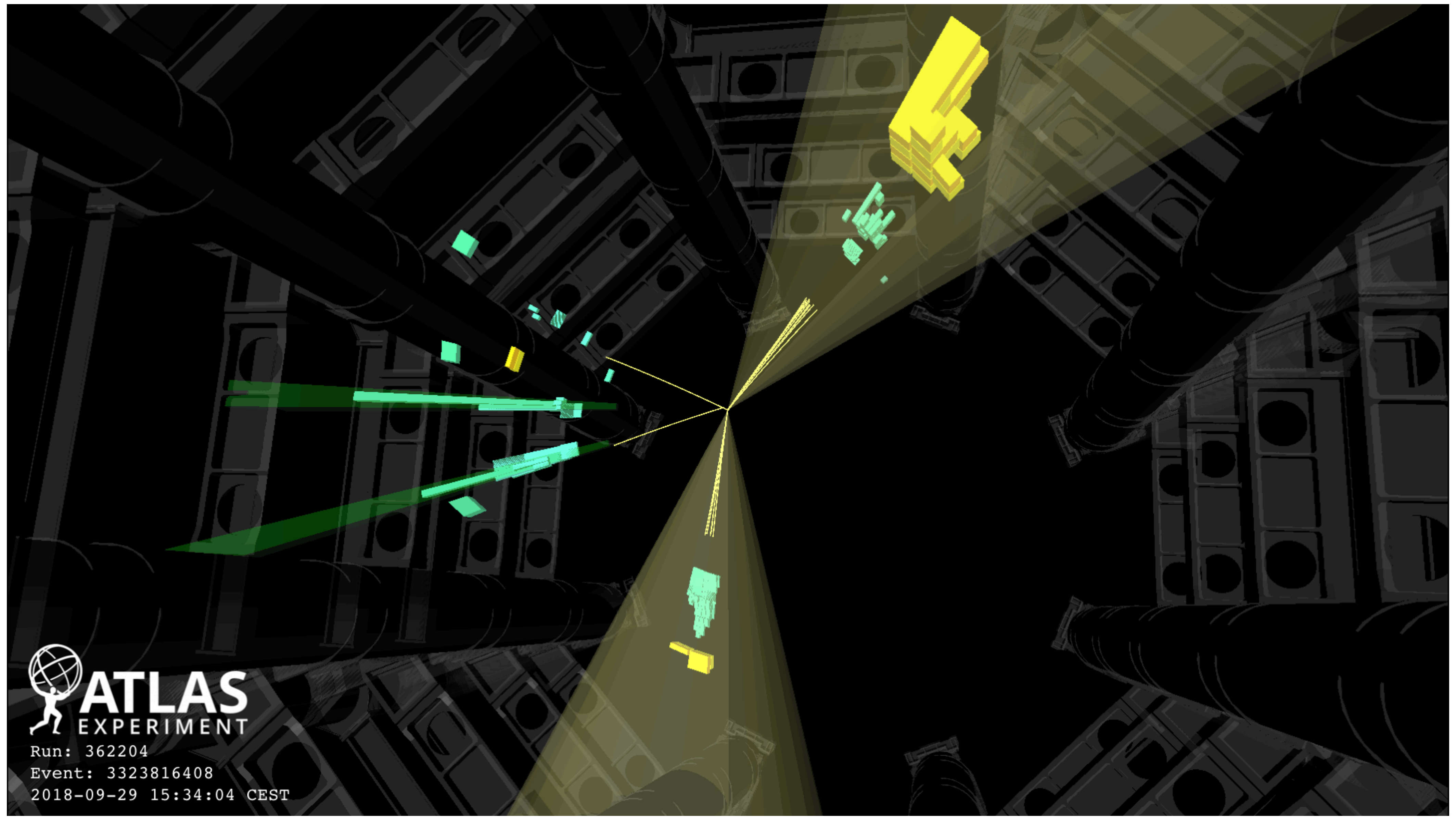
- **Double hadronization** (in both hidden and visible sectors)
- Signal appearance at some distance from the interaction point : emerging jet (EJ)

Emerging jets analysis

- Final state :
  2 energetic jets, **displaced tracks** and **secondary vertices**

- Main background : di-jet events from QCD processes
  - can reproduce EJ signature : neutral B mesons, photons (pair production)



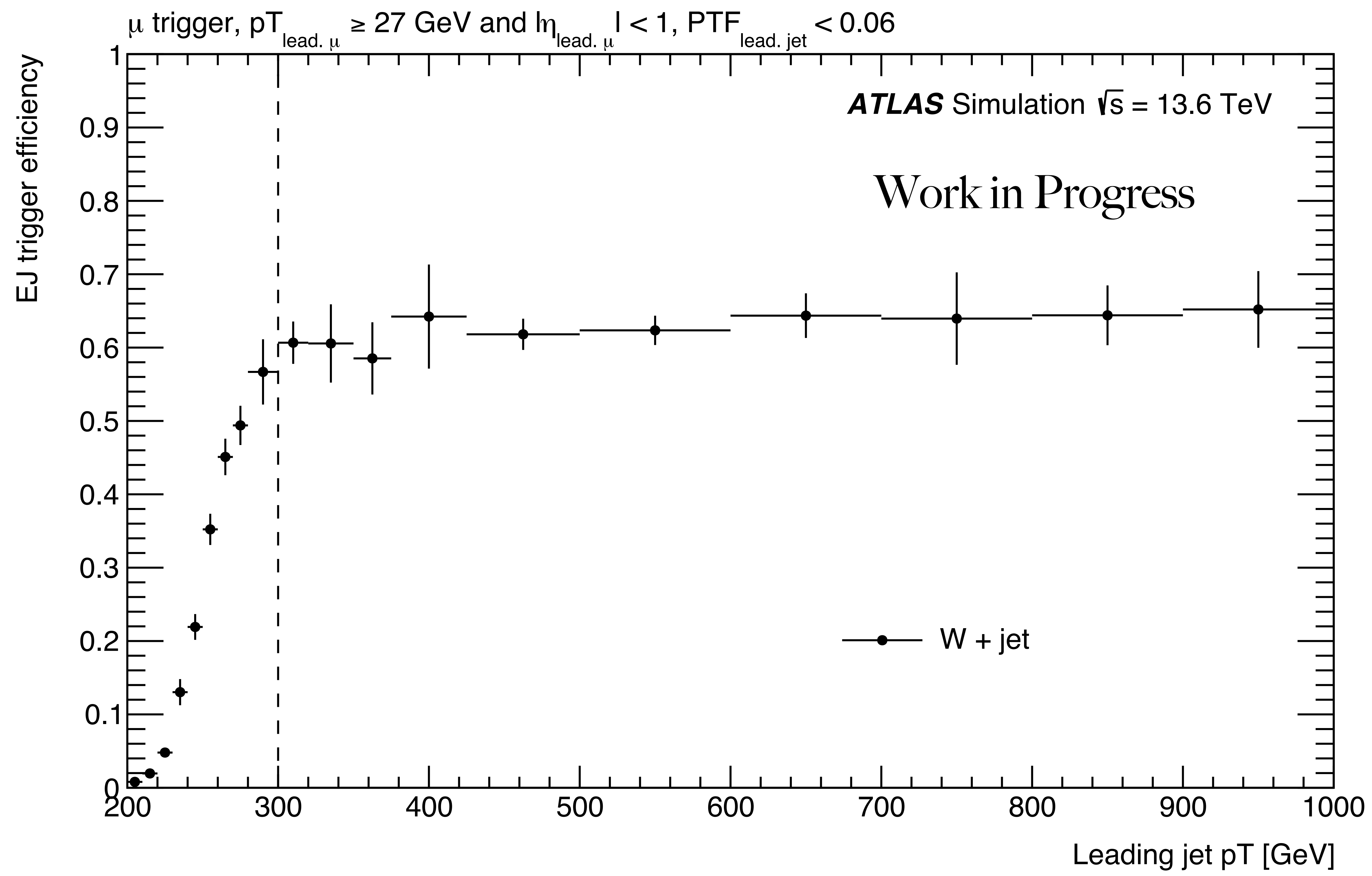


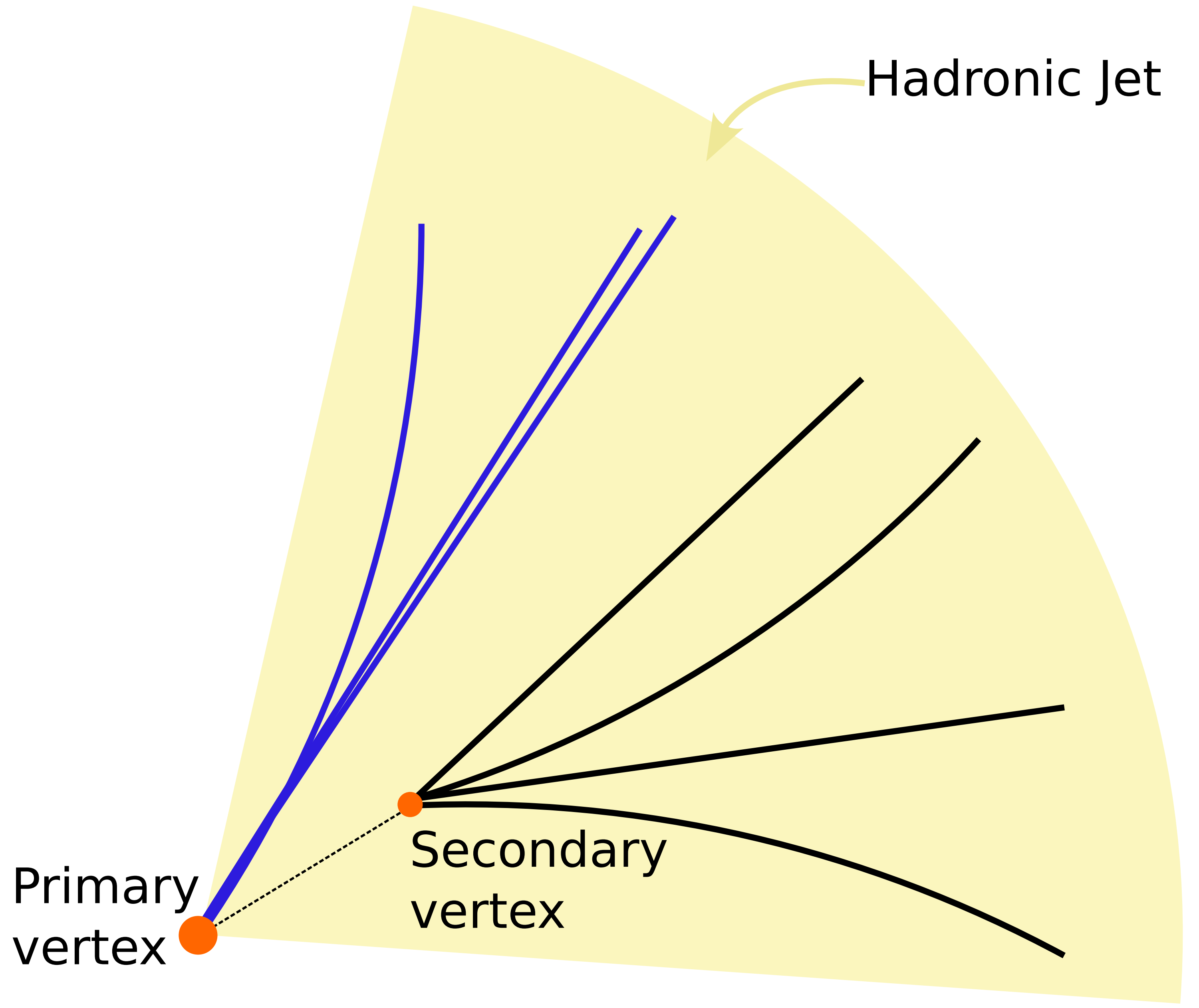
ATLAS event with 2 pair-producing photons (green cones)

- Analysis principle :
  - events selection optimisation (based on simulated events), signal region (SR) definition
  - SM contribution estimation in SR
  - events count in SR, comparison to SM expectation
  - statistical interpretation, constraint on model parameters (if no excess)

Trigger

- **High p_T jet trigger** : jet with $p_T > 460$ GeV
- **Emerging jet trigger** : jet with $p_T > 200$ GeV, $PTF^{jet} < 0.08$
(Prompt Track p_T Fraction)





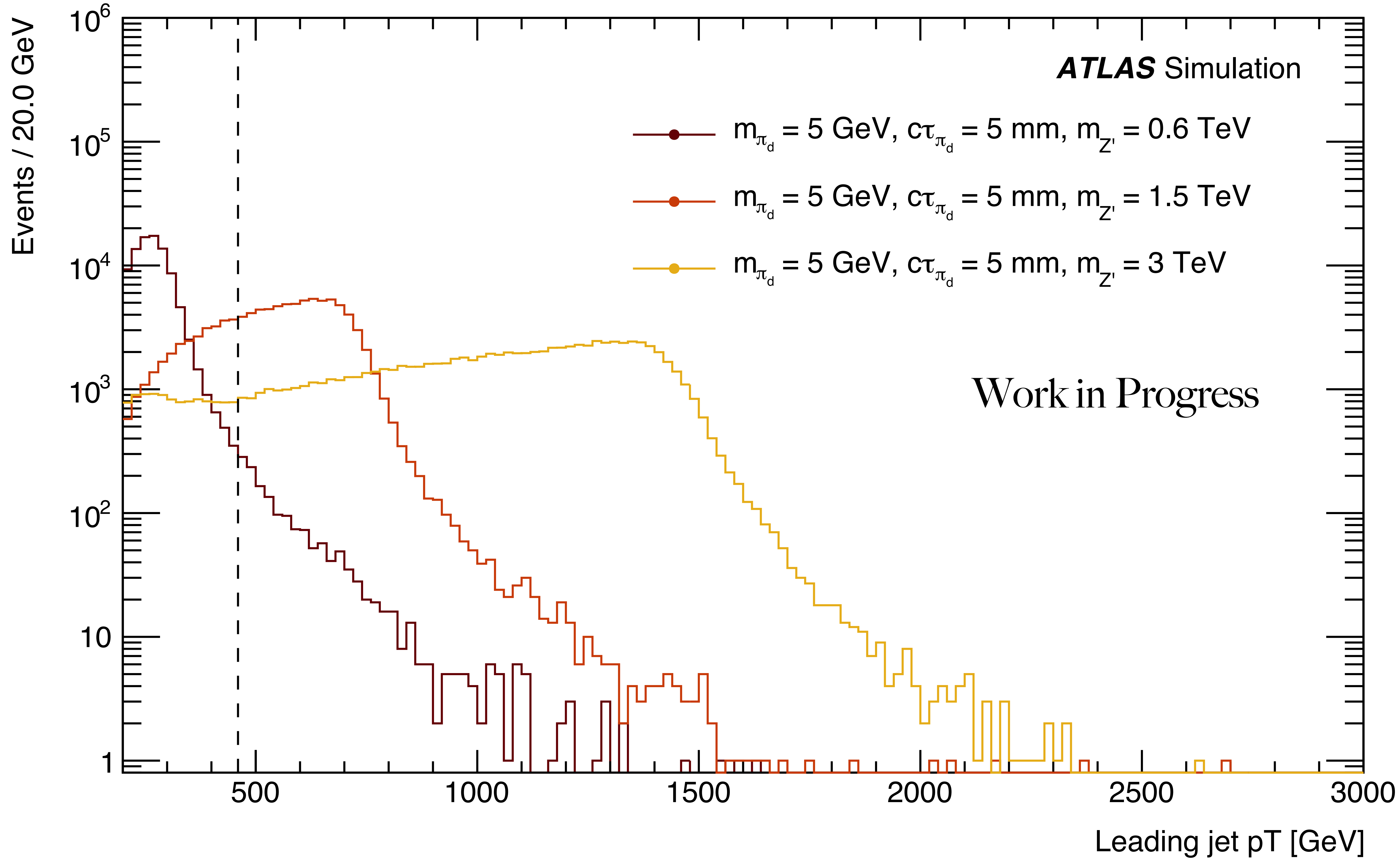
$$\text{PTF} = \frac{p_{T,\text{ prompt}}}{p_{T,\text{ total}}}$$

- **Trigger efficiency**
  offline selections :
  - on jet p_T (and PTF for EJ trigger) in order to be on the plateau

Benchmark EJ model

- Signal events MC simulation :
  finite number of parameter values

m_{π_d} (GeV)	5	10	20
$c\tau_{\pi_d}$ (mm)	5, 50		
$m_{Z'}$ (GeV)	600, 1500, 3000		



High p_T jet trigger can't be used to search for $m_{Z'} = 600$ GeV signal

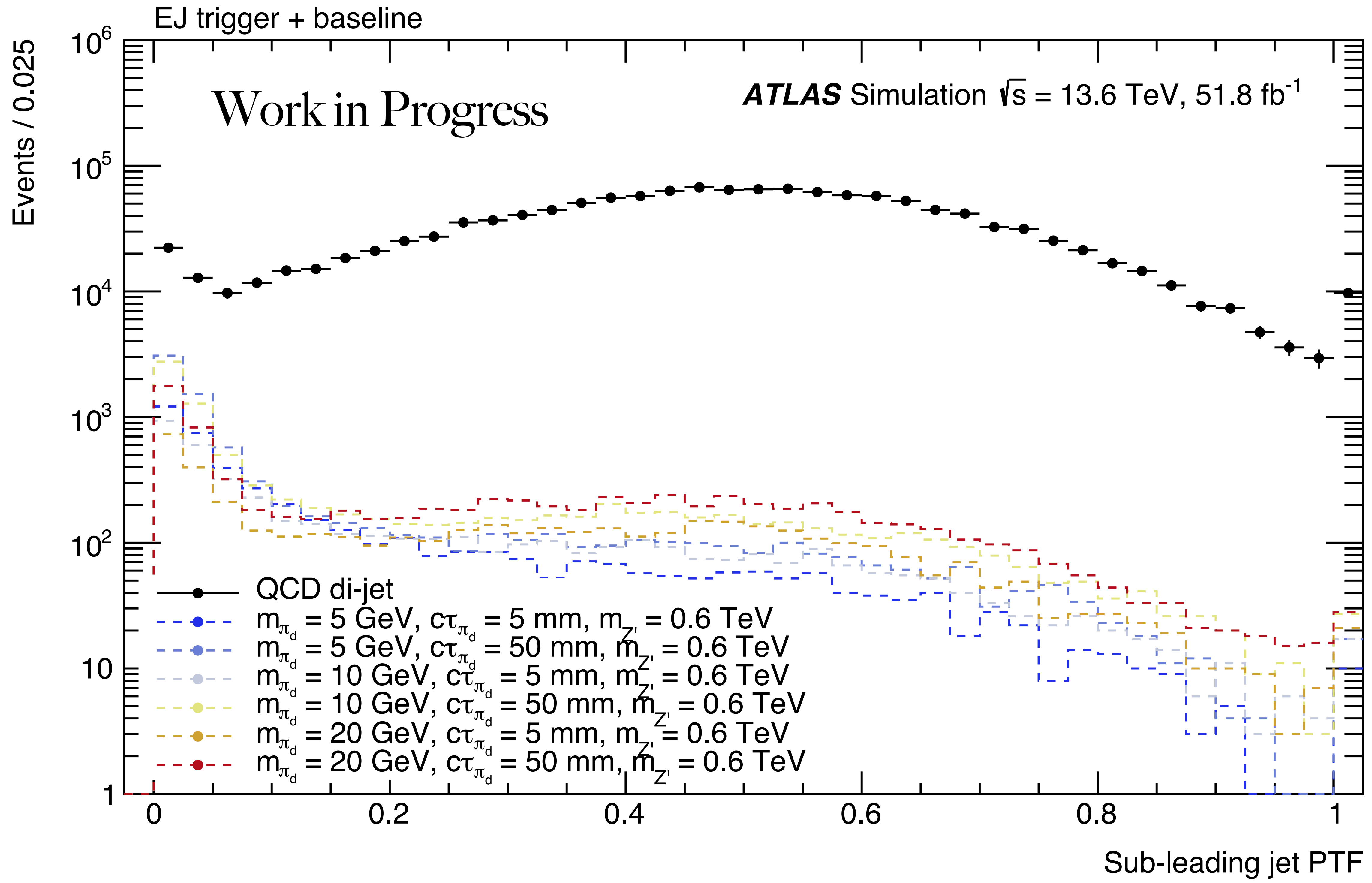
- Strategy : **2 separate event selections**
  - one using the high p_T jet trigger
  - the other **using the emerging jet trigger**, sensibility to low $m_{Z'}$ signal (my focus)

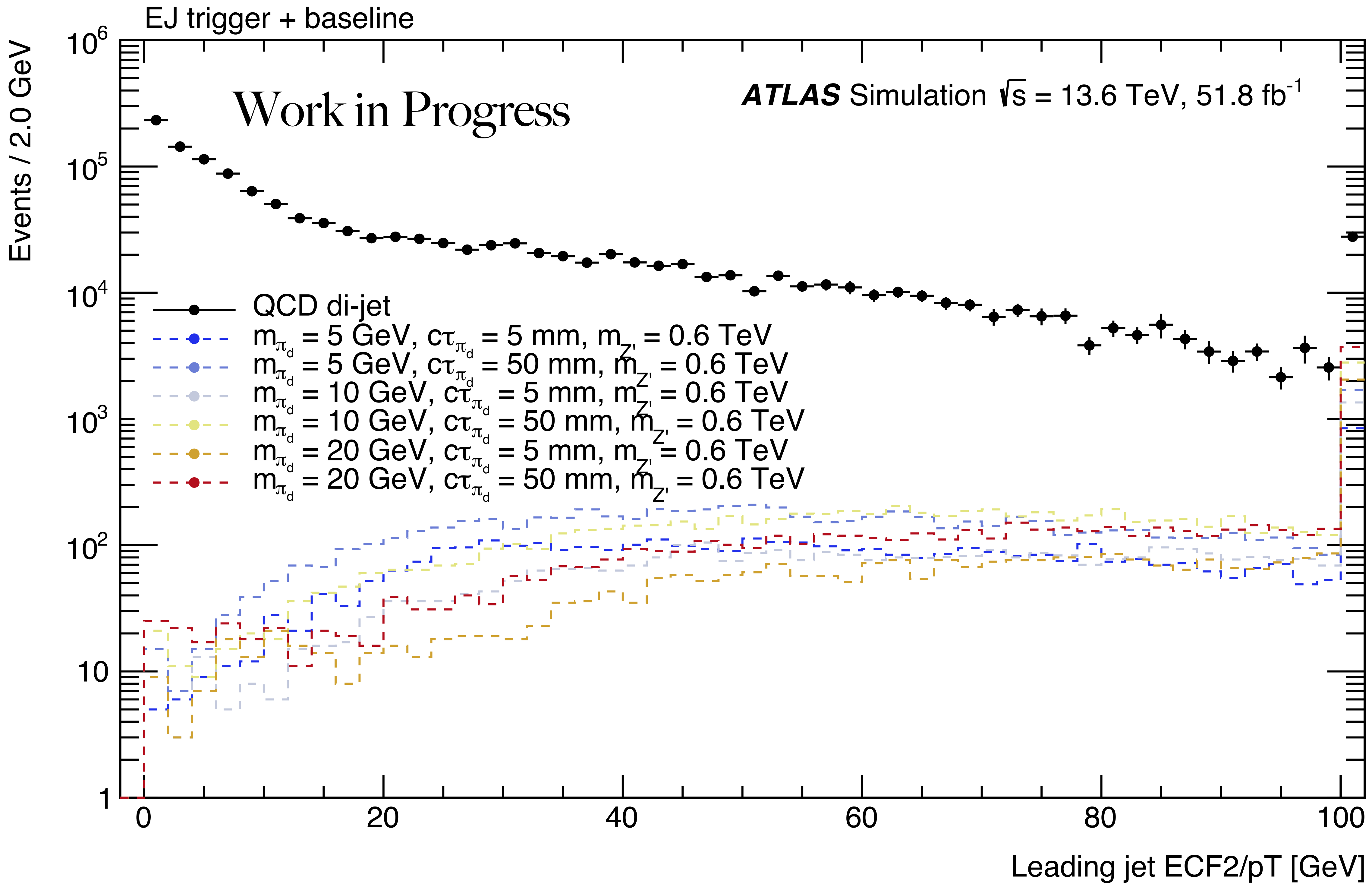
Signal selection

- Baseline selection :
  - $N_{jet} \geq 2$
  - $p_{T,\ lead.\ jet} > 300$ GeV and $PTF_{lead.\ jet} < 0.025$
- Pre-selection :
  - number of **reconstructed secondary vertex**

- Discriminating jet variables :
  - **track variable** : PTF
  - **substructure** : $ECF2 = \sum_{i,j \in j} p_{T_i}\ p_{T_j}\ \Delta R_{ij}$

  (quantify energy distribution within the jet)

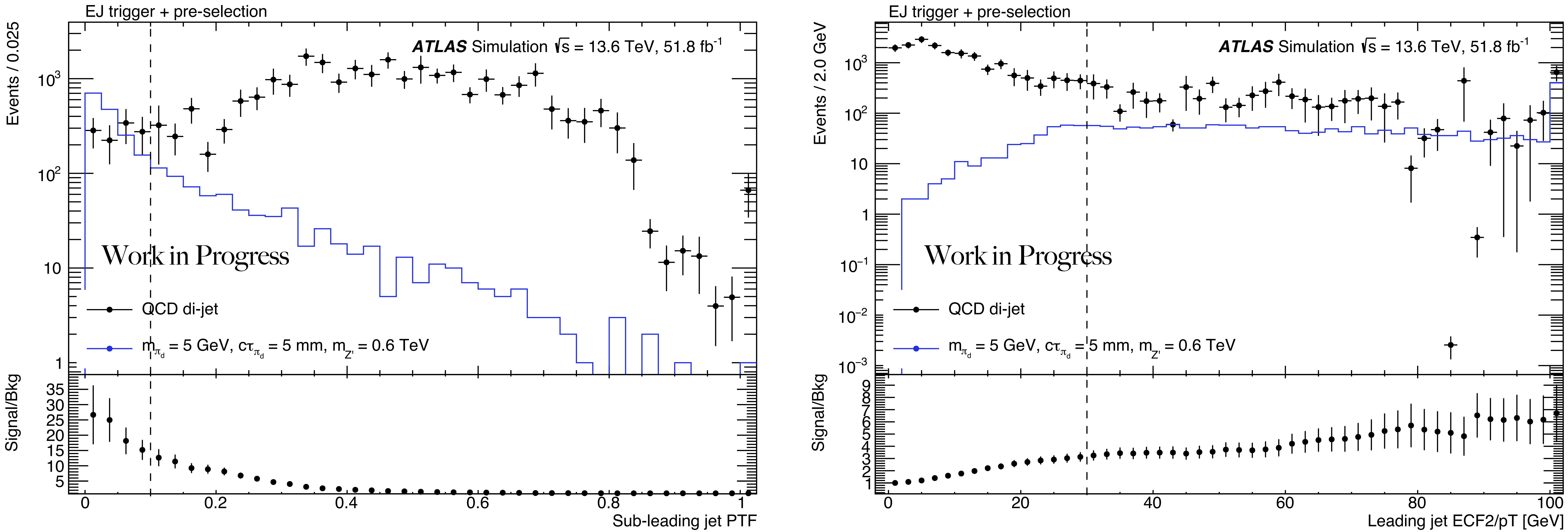




Signal selection

- SR defined by $ECF2_{\ lead.\ jet}$ and $PTF_{sub-lead.\ jet}$ cuts
  - **decorrelated** variables
  - **complementary effects** on background elimination
- What cut values ? : gain on signal/background ratio

Choice :
- $ECF2/p_{T\ lead.\ jet} > 30$ GeV
- $PTF_{sub-lead.\ jet} < 0.1$



Background estimation

- 4 regions in a $(PTF_{sub-lead.\ jet}, ECF2/p_{T\ lead.\ jet})$ plane delimited by cut values

- **Data-driven** background estimation in A (SR):

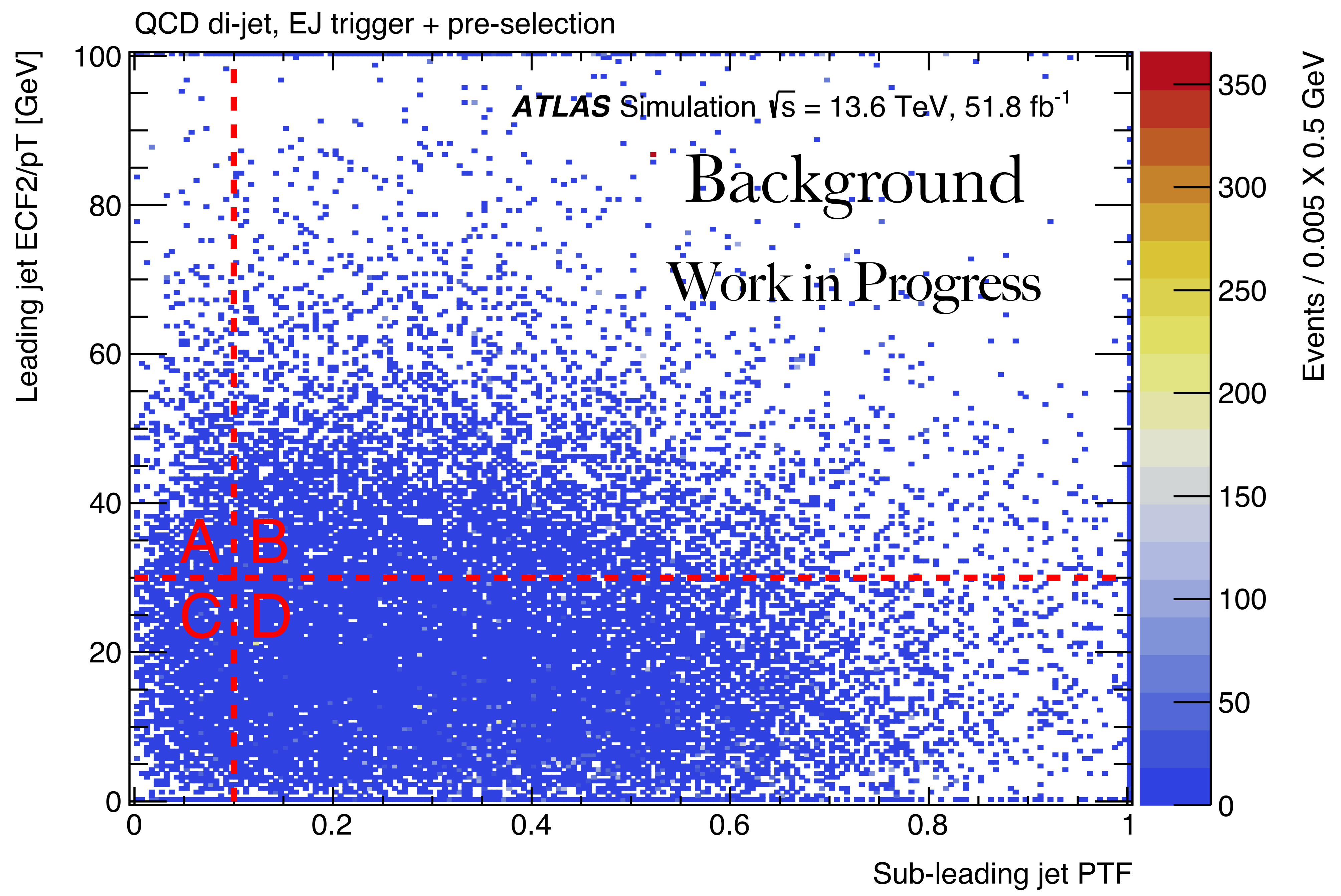
$$N_A^{bkg} = \frac{N_C^{bkg}}{N_D^{bkg}} \times N_B^{bkg} \approx \frac{N_C}{N_D} \times N_B$$

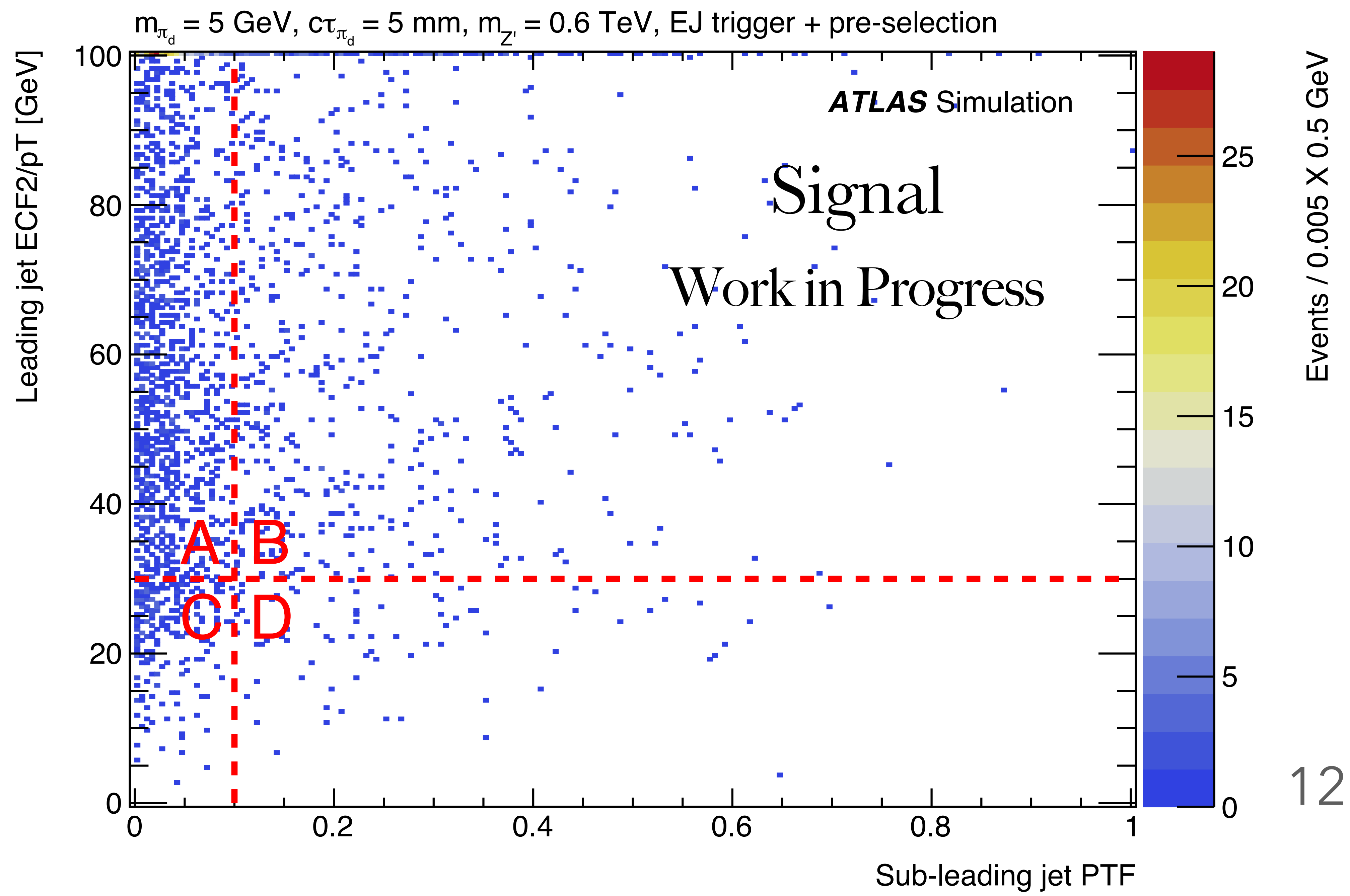
- decorrelated variables for background events
- negligible signal presence in B, C and D (likelihood-fit can take it into account)

- First check on simulated background events :

QCD di-jet	$N_{events} \pm$ MC stat. uncertainty
A	305 ± 141
B	6324 ± 730
C	818 ± 182
D	17462 ± 1003

$n_A^{bkg} = (n_B^{bkg} \times n_C^{bkg})/n_D^{bkg} = 296 \pm 76$ (MC stat.)

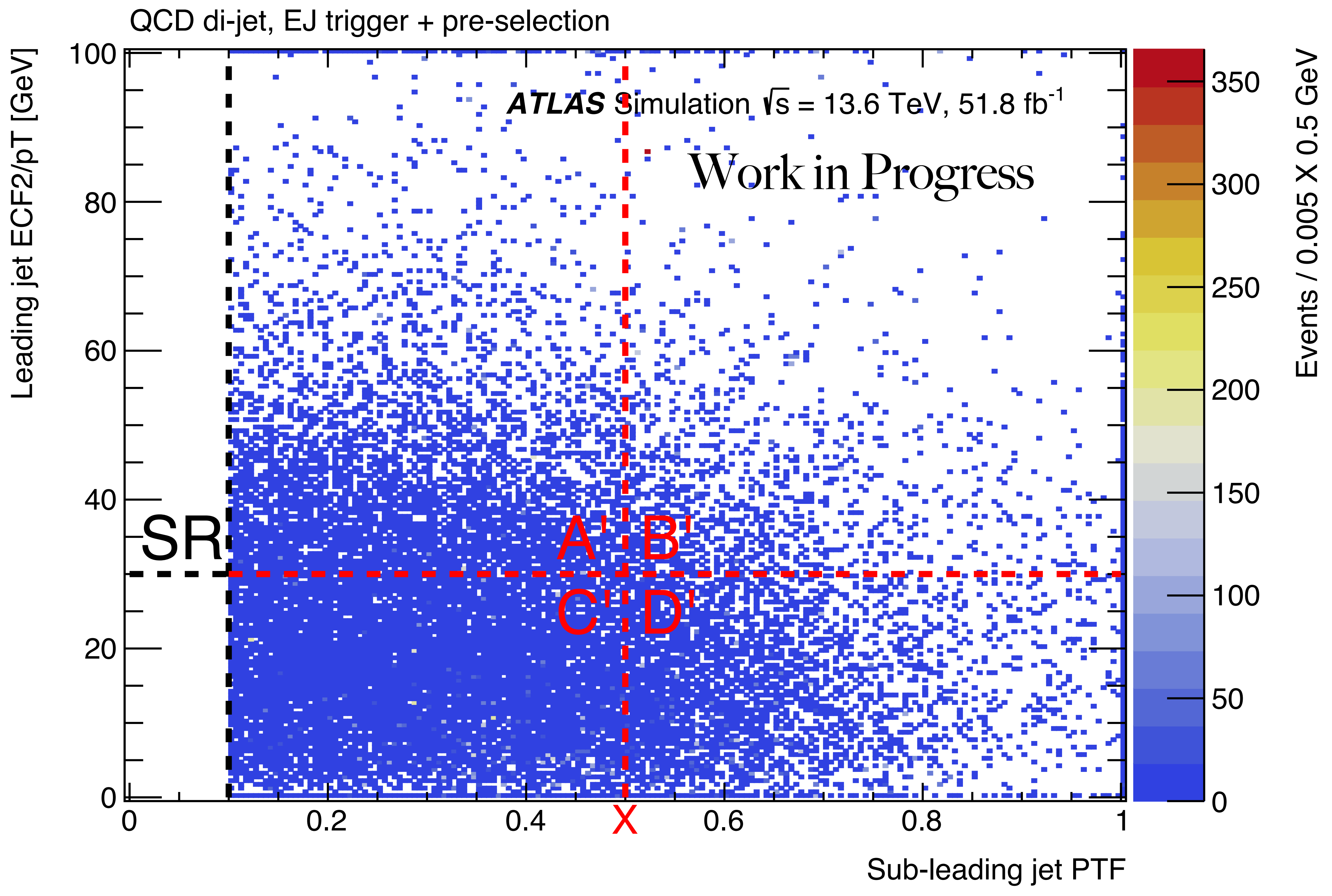




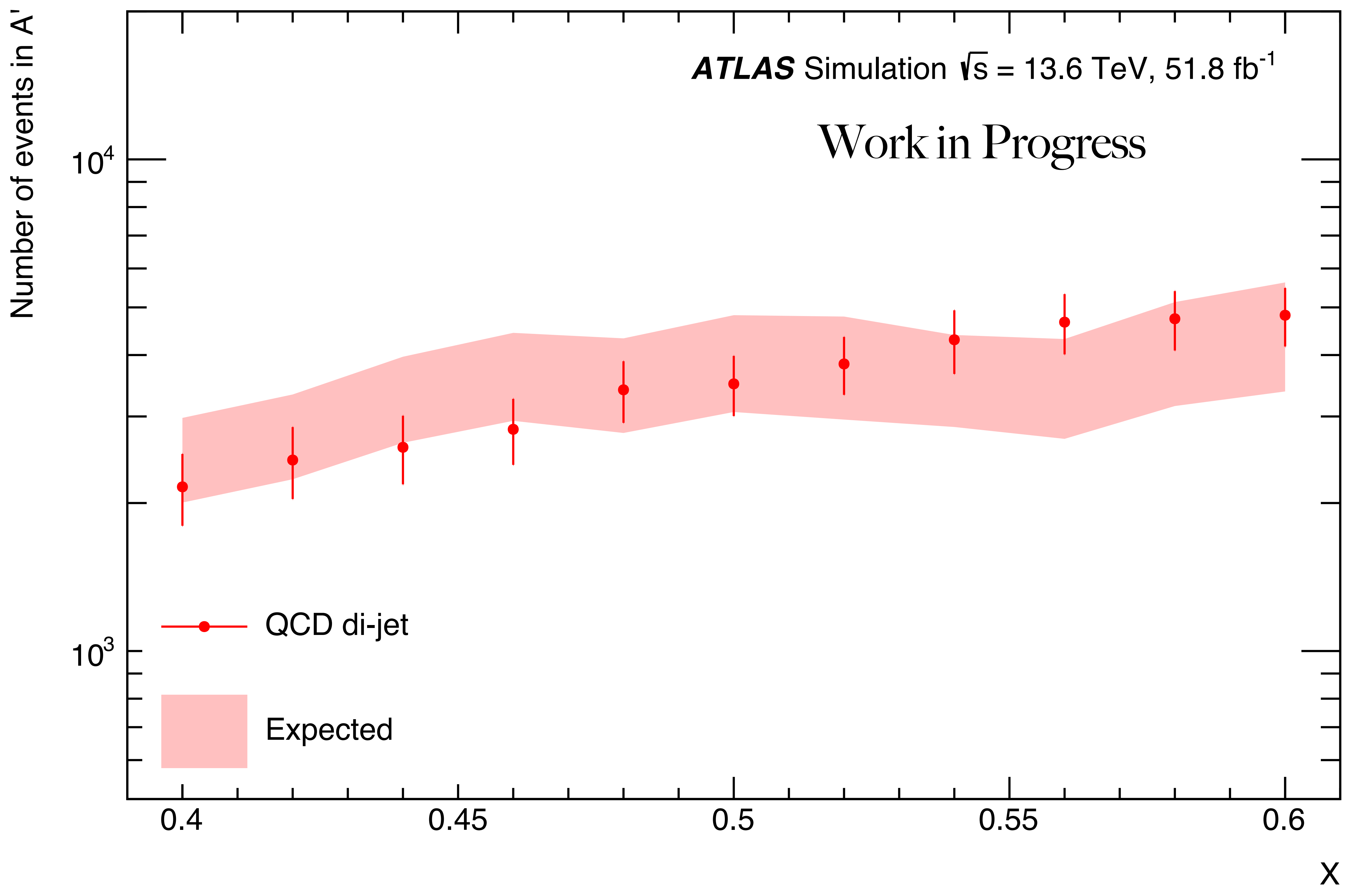
ABCD method in validation regions

- 4 new regions with X varying :

	$0.1 < PTF_{sub-lead.\ jet} < X$	$PTF_{sub-lead.\ jet} \geq X$
$ECF2/p_{T_{lead.\ jet}} \geq 30$ GeV	A'	B'
$ECF2/p_{T_{lead.\ jet}} < 30$ GeV	C'	D'



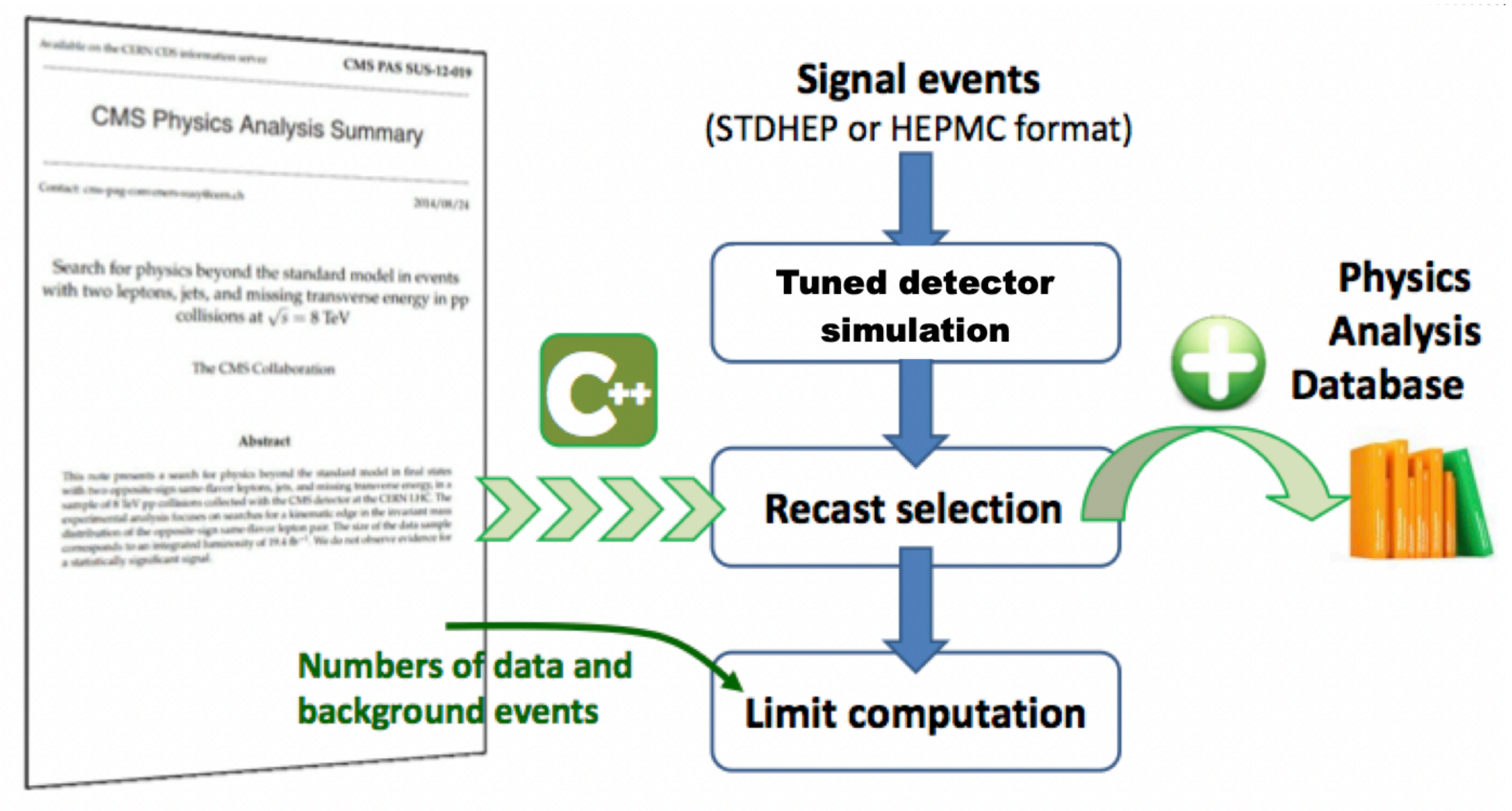
$$n_{A'}^{exp} = (n_{B'}^{bkg} \times n_{C'}^{bkg})/n_{D'}^{bkg} \pm \text{(MC stat.)}$$



Reinterpretation of LHC results

- BSM search : no sign of new physics for now, only constraints on theoretical models
- Many relevant theories should be tested in LHC analysis, but not possible in practice
- Importance of **reinterpretation** : published ATLAS or CMS analysis results used to study unexplored theories
- Example of a reinterpretation tool : **MadAnalysis**

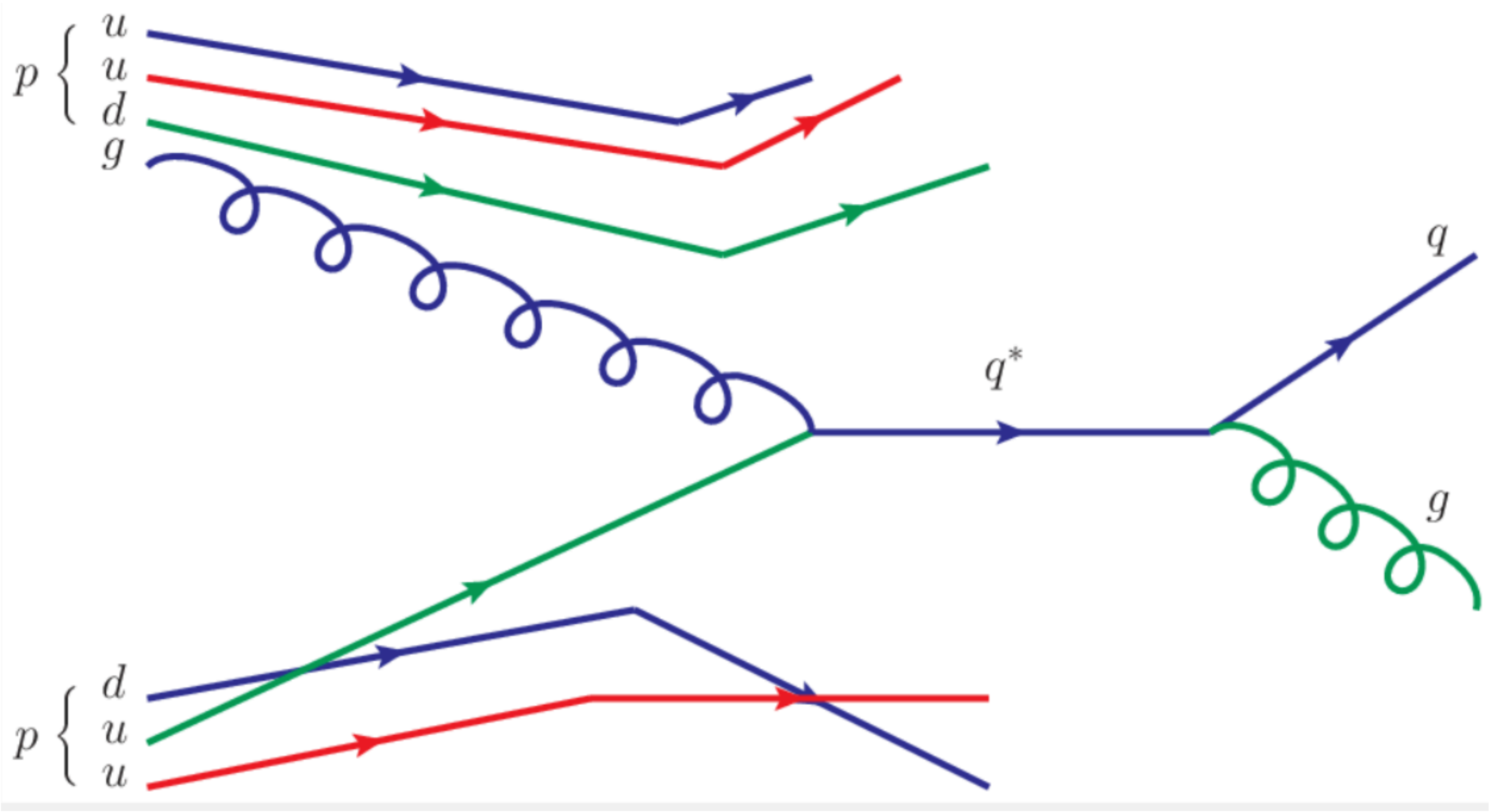
  - analysis of interest implemented in C++ code : reproduction of the analysis strategy
  - simplified detector simulation and analysis code applied to new physics events (at hadron level)
  - prediction of number of signal events in SRs, comparison with data and background expectation
  - derivation of signal cross-section limits



Implementation in MadAnalysis

- Example : **Search for new resonances in mass distributions of jet pairs using 139 fb^{-1} of pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector**

- Search for new heavy particles produced in pp collisions, and decaying into a pair of partons : di-jet final state



- Different BSM models considered : excited quarks q^*, new Z', W' ...

- Examination of the invariant di-jet mass m_{jj} : search for local excess above SM contribution

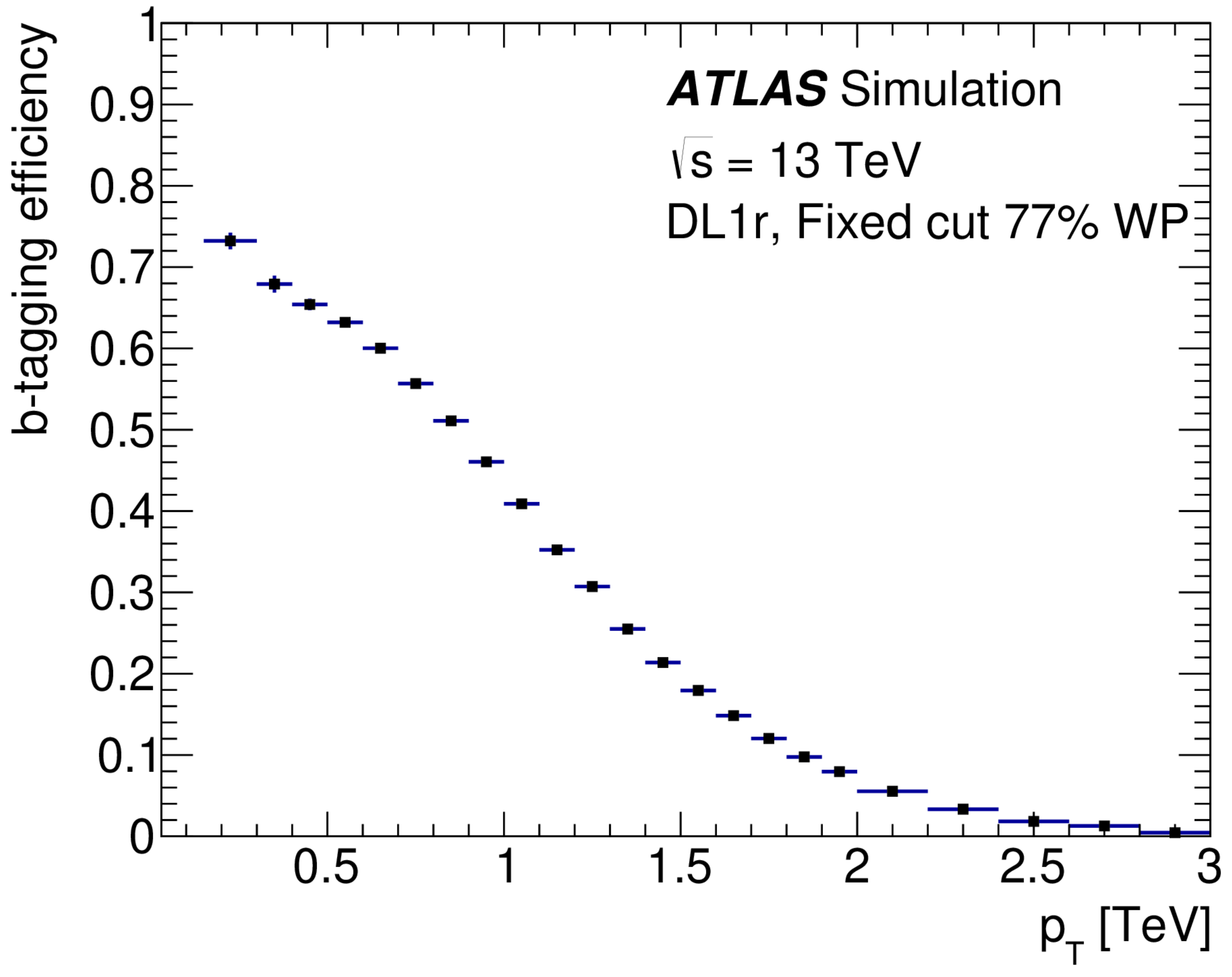
Implementation in MadAnalysis

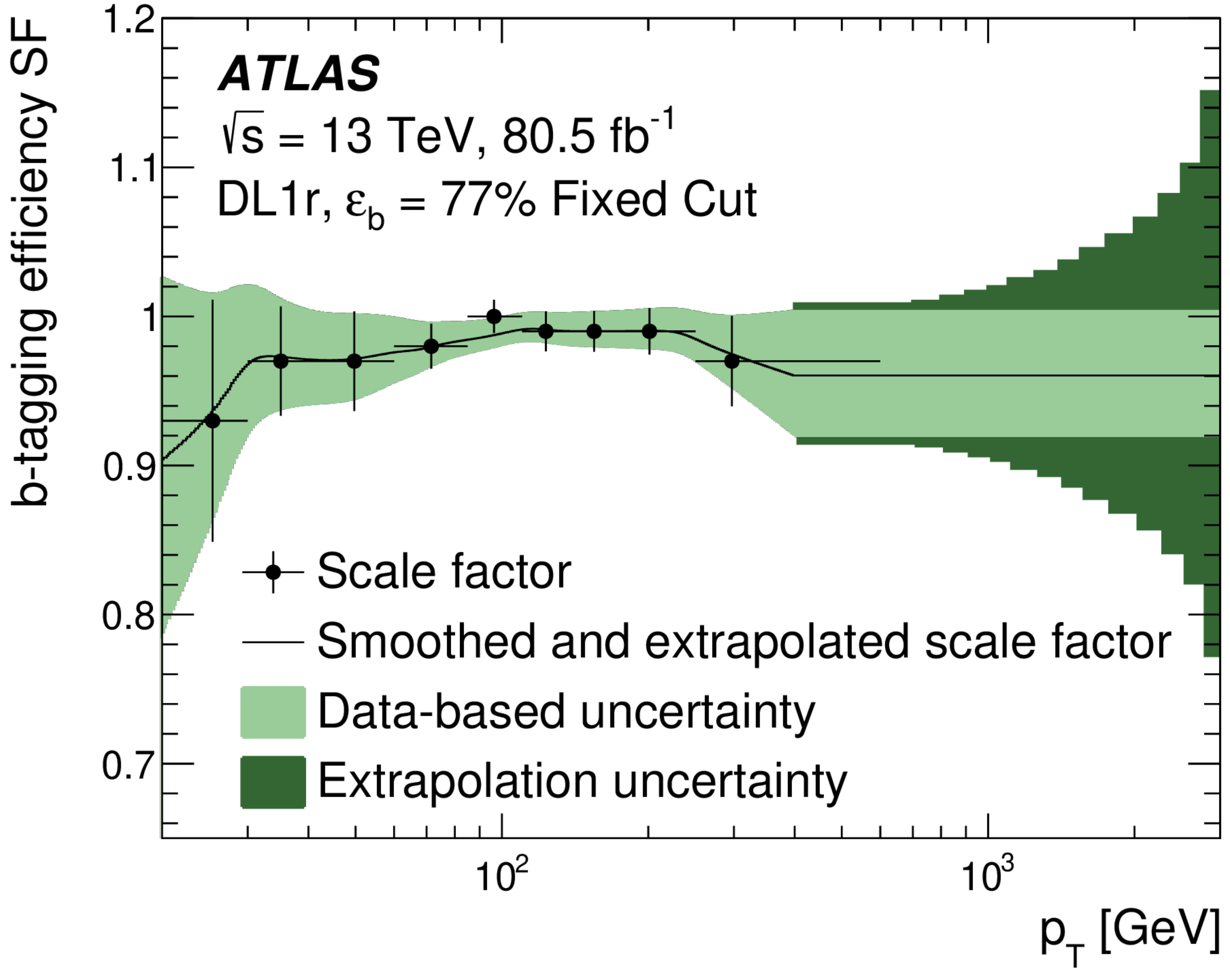
- Several steps :
  1) Writing of the **analysis code** based on article information : trigger, event pre-selection, cleaning procedure, cuts, SR definition

Category	Inclusive		$1b$	$2b$
Jet p_{T}	> 150 GeV			
Jet ϕ	$\lvert\Delta\phi(jj)\rvert > 1.0$			
Jet $\lvert\eta\rvert$	-		< 2.0	
$\lvert y^*\rvert$	< 0.6	< 1.2	< 0.8	
m_{jj}	> 1100 GeV	> 1717 GeV	> 1133 GeV	
b-tagging	no requirement		$\geqslant 1$ b-tagged jet	2 b-tagged jets
Signal	DM mediator Z' W' q^* QBH Generic Gaussian	W^*	b^* Generic Gaussian	DM mediator Z' ($b\bar{b}$) SSM Z' ($b\bar{b}$) graviton ($b\bar{b}$) Generic Gaussian

Implementation in MadAnalysis

2) Configuration of the **simplified detector simulation** : object (lepton, jet …) reconstruction and tagging efficiencies, energy smearing (detector resolution), jet-clustering algorithm parameters

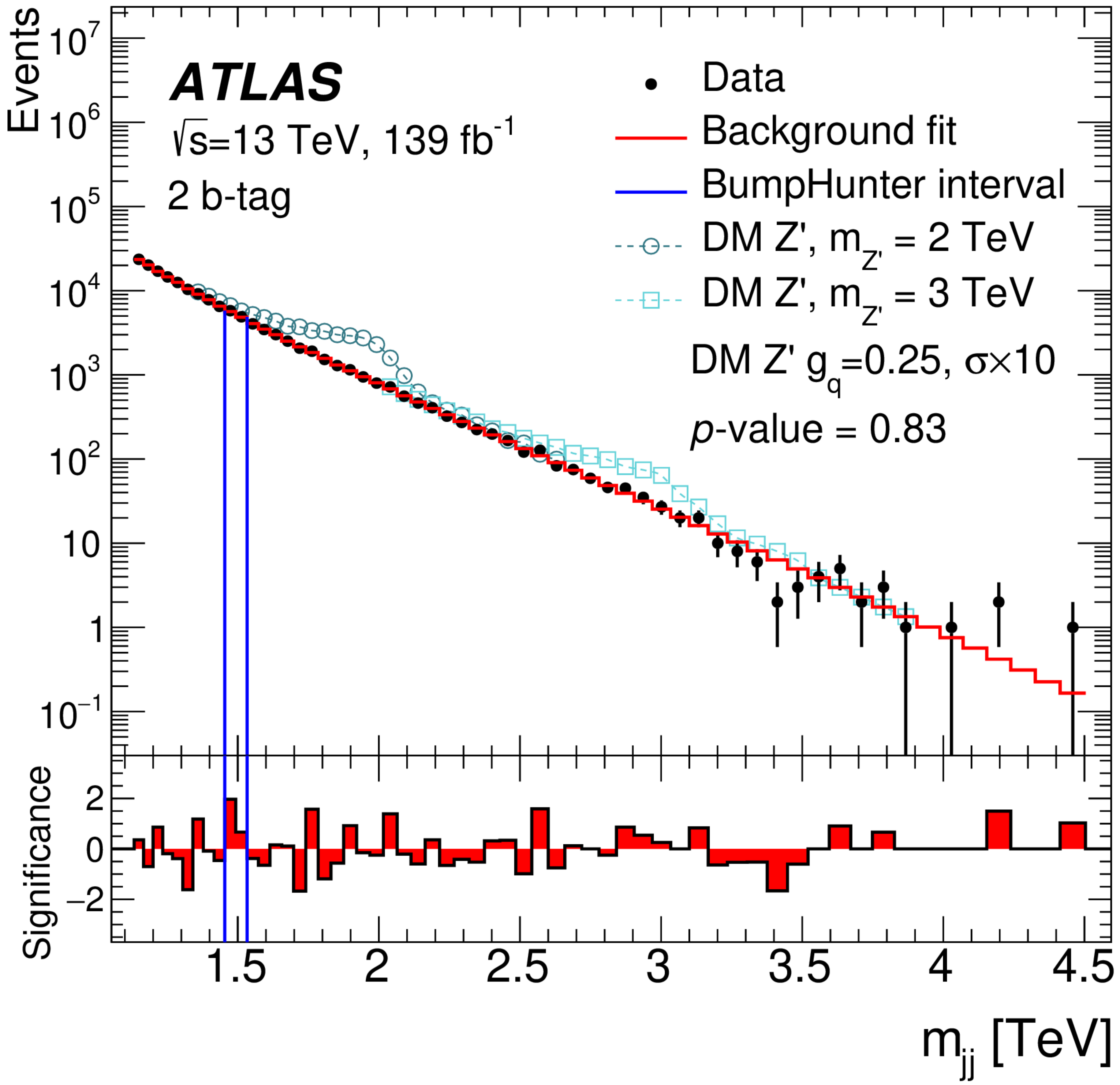




Implementation in MadAnalysis

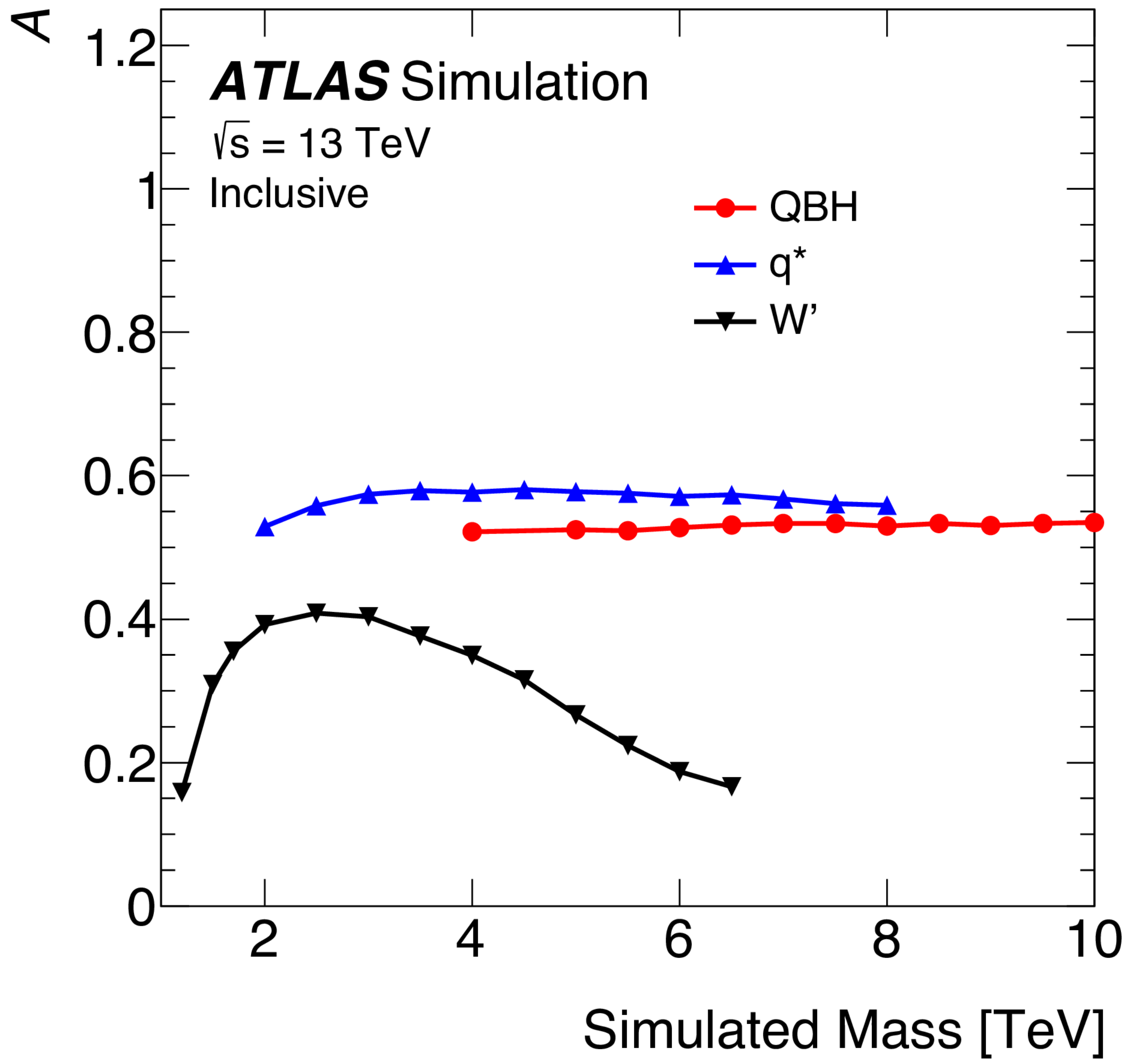
3) Retrieval of number of **data** and **expected background** events in SRs : available in HEPData, published together with the article

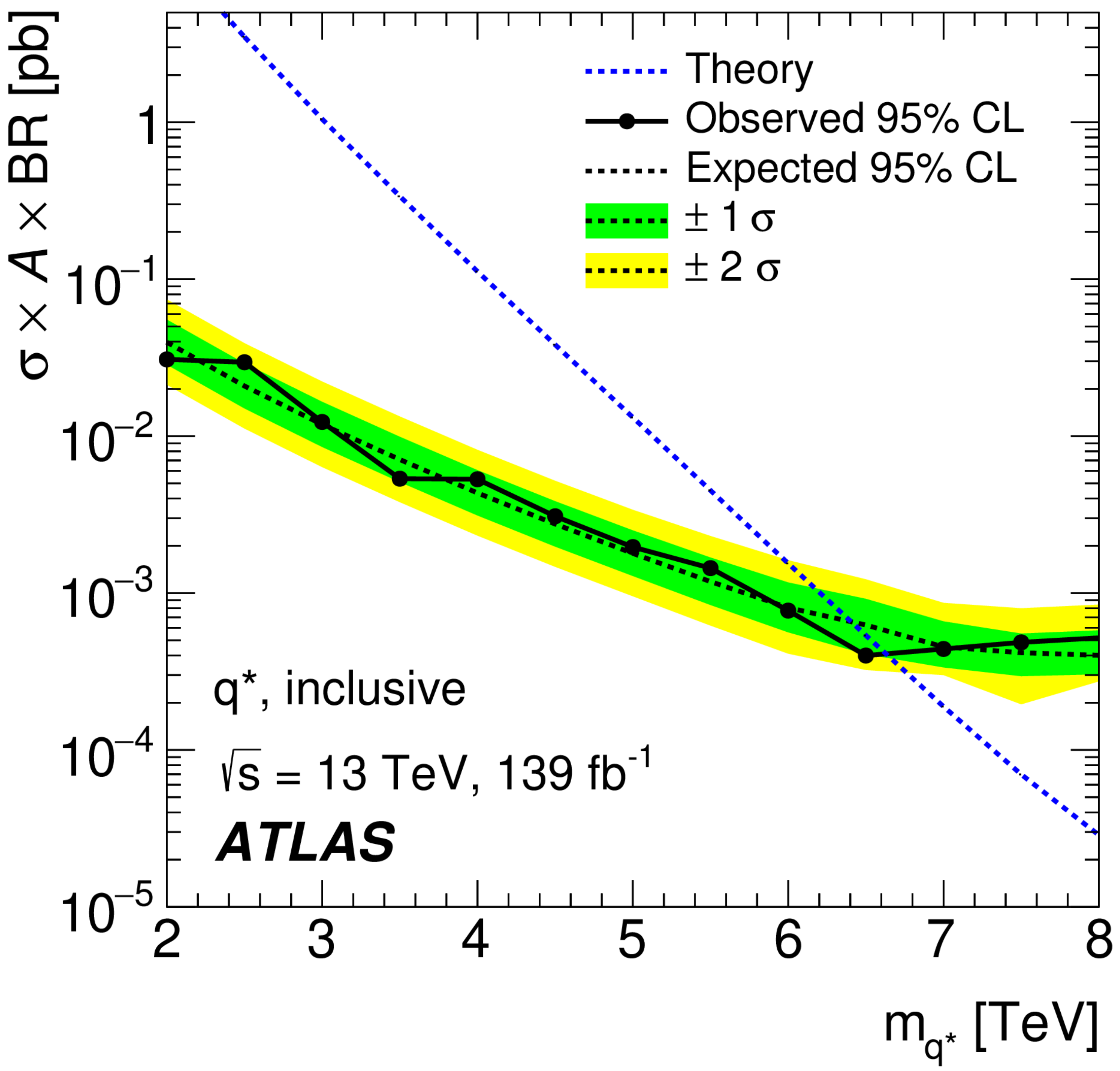
Number of Events	Observed	Fit
SQRT(S)	13 TeV	
LUMINOSITY	139 fb^{-1}	
m_{jj} [GeV]	high mass Z' 2 b-tag	
1133 - 1166	23572	2.351709e+04 ±1.533528e+02
1166 - 1200	20137	2.023711e+04 ±1.422572e+02
1200 - 1234	17041	1.692838e+04 ±1.301091e+02
1234 - 1269	14576	1.459987e+04 ±1.208299e+02
1269 - 1305	12523	1.256678e+04 ±1.121017e+02
1305 - 1341	10363	1.052901e+04 ±1.026110e+02



Implementation in MadAnalysis

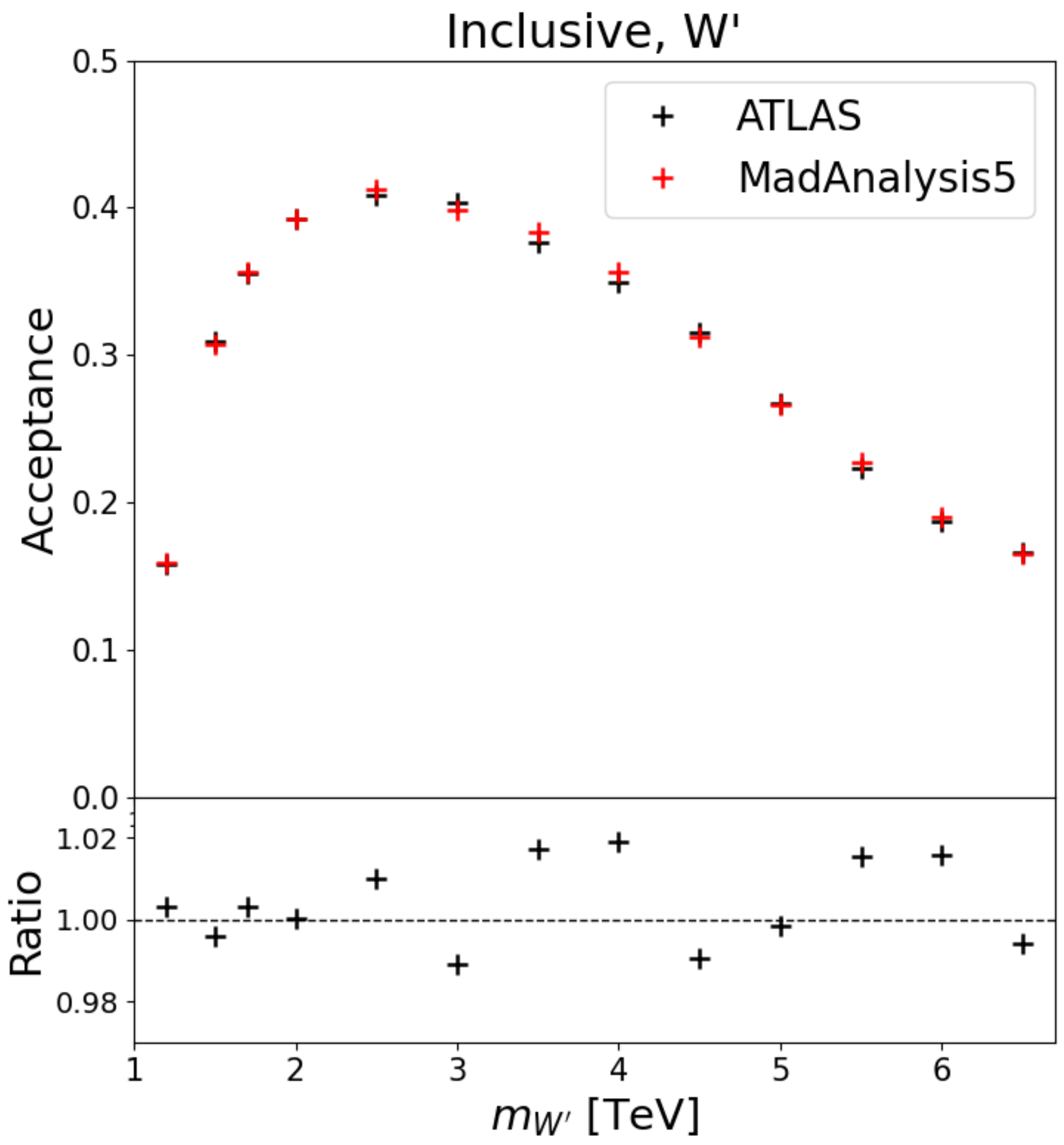
4) Validation :
- generation of signal events for several models that have been tested in the analysis
- application of the MadAnalysis code : **signal efficiencies** and **cross-section upper limits** computation
- **comparison** with official ATLAS results

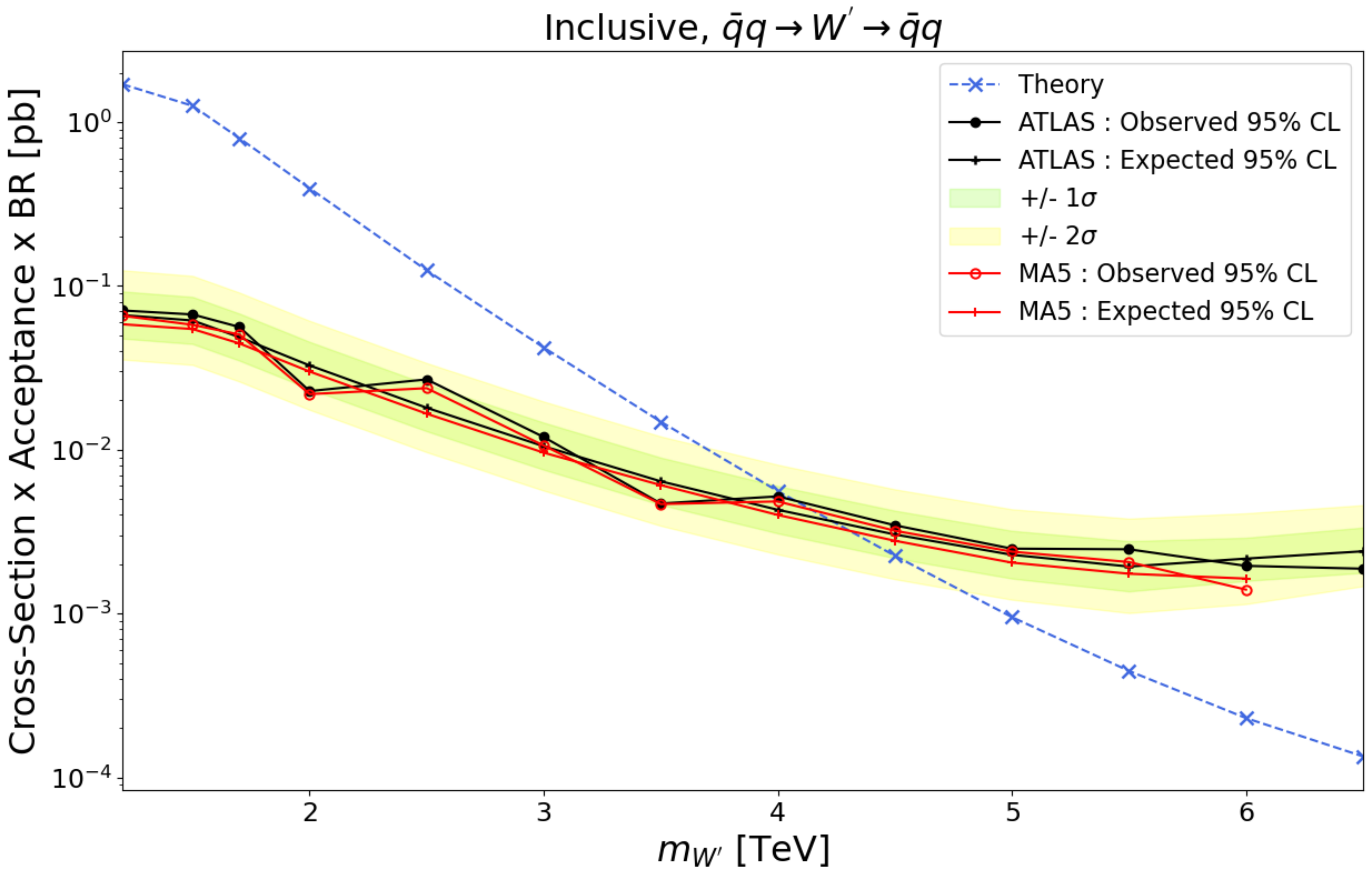




Implementation in MadAnalysis

4) Validation :
- redaction of a note describing the analysis and the validation results
- merging of the implementation to the MadAnalysis database : **available for recasting** (https://dataverse.uclouvain.be/dataset.xhtml?persistentId=doi:10.14428/DVN/KHJ1MW)





Conclusions - Prospects

- Analysis :
  - **ABCD validation in data** in control regions far away from SR
- Reinterpretation :
  - 2 ATLAS analysis added to the MadAnalysis database (http://madanalysis.irmp.ucl.ac.be/wiki/PublicAnalysisDatabase) :
  « Search for new resonances in mass distributions of jet pairs using 139 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector »
  &
  « Search for non-resonant production of semi-visible jets using Run 2 data in ATLAS »
  - next implementation : ATLAS dark jet analysis
  - **phenomenology article** : recasting these analysis to constrain dark QCD model
- **ATLAS technical work completed** :
  - related to jet constituent reconstruction, improvement on jet performances

Thanks for your attention

Back-up

Cutflow

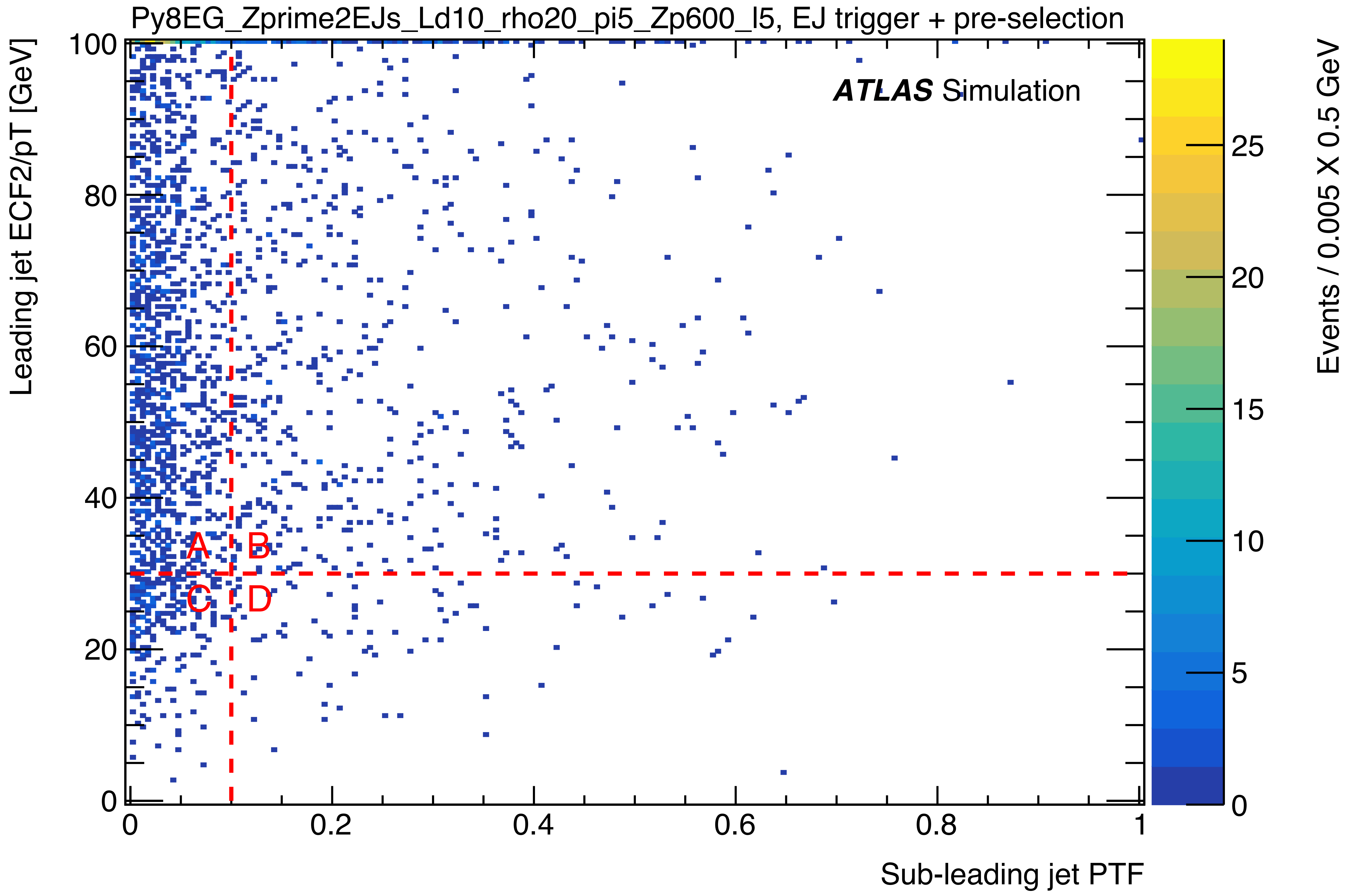
	m_pi_d = 5 GeV c tau_pi_d = 5 mm	m_pi_d = 5 GeV c tau_pi_d = 50 mm	m_pi_d = 10 GeV c tau_pi_d = 5 mm	m_pi_d = 10 GeV c tau_pi_d = 50 mm	m_pi_d = 20 GeV c tau_pi_d = 5 mm	m_pi_d = 20 GeV c tau_pi_d = 50 mm	QCD di-jet (MC weight, Lumi. = 51.8 fb-1)
Trigger	30 125	35 211	28 100	34 846	24 794	30 489	34 618 173.3 +/- 98 461.8
Baseline	4 506	8 185	4 451	8 898	4 442	8 012	1 329 036.7 +/- 11 003.6
Secondary vertex >= 1	2 317	2 129	2 245	2 518	1 783	1 865	24 908.2 +/- 1 261.4
ECF2/pT > 30 GeV	2 003	1 819	2 090	2 351	1 696	1 794	6 629.0 +/- 743.4
PTF < 0.1	1 349	1 388	1 287	1 565	878	928	305.3 +/- 140.9

Signal contamination in VR and CR

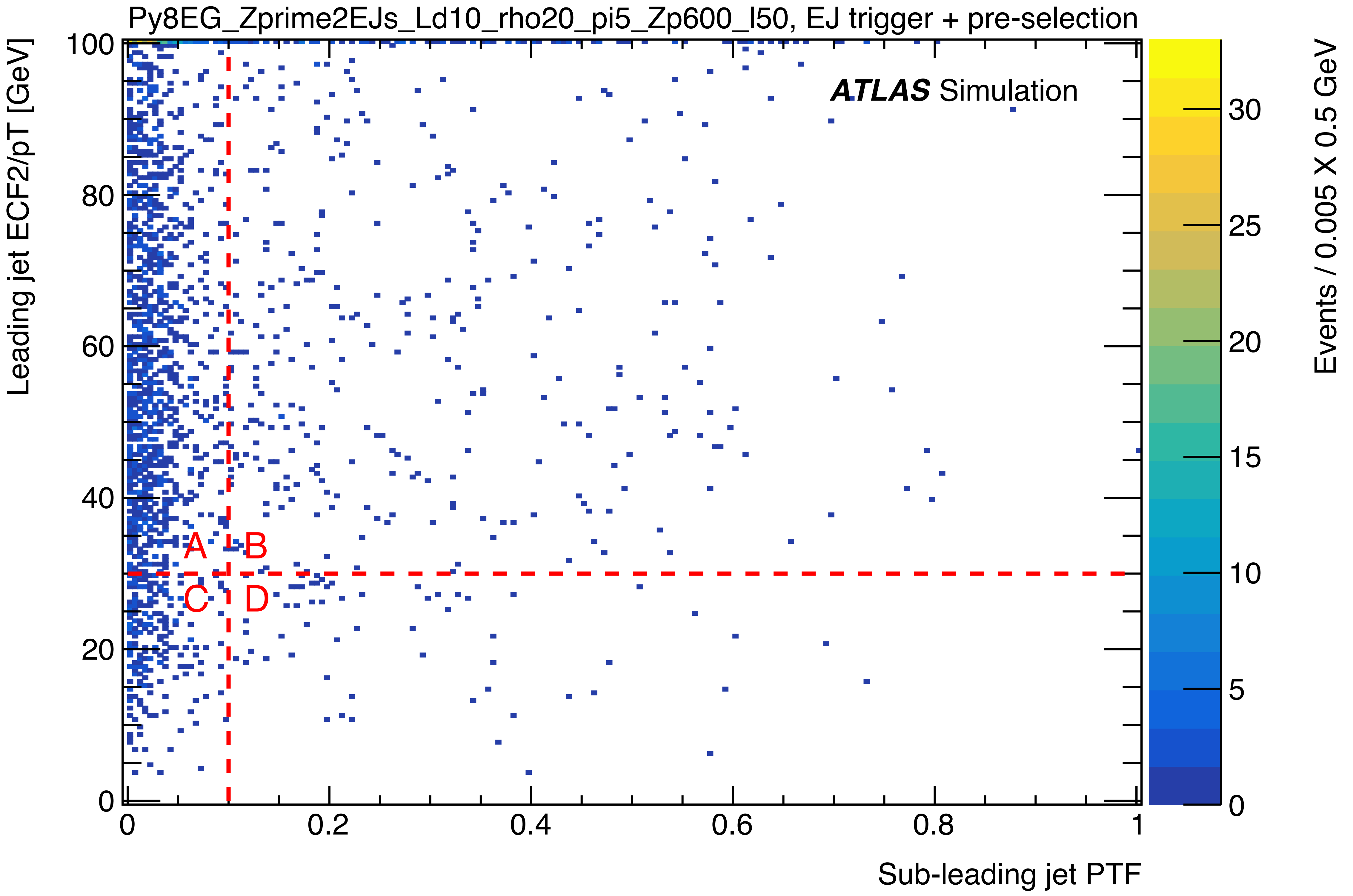
Signal/Background ratio between CR/VR and SR : $\frac{N_S(R)/N_B(R)}{N_S(A)/N_B(A)}$ with $R \in \{B, C, D\}$

Signal/Bkg ratio	B/A	C/A	D/A
m_pi_d = 5 GeV c tau_pi_d = 5 mm	0.02 +/- 0.01	0.07 +/- 0.03	0.0010 +/- 0.0005
m_pi_d = 5 GeV c tau_pi_d = 50 mm	0.010 +/- 0.007	0.07 +/- 0.03	0.0008 +/- 0.0004
m_pi_d = 10 GeV c tau_pi_d = 5 mm	0.03 +/- 0.01	0.03 +/- 0.01	0.0008 +/- 0.0004
m_pi_d = 10 GeV c tau_pi_d = 50 mm	0.02 +/- 0.01	0.03 +/- 0.02	0.0003 +/- 0.0002
m_pi_d = 20 GeV c tau_pi_d = 5 mm	0.04 +/- 0.02	0.02 +/- 0.01	0.0007 +/- 0.0003
m_pi_d = 20 GeV c tau_pi_d = 50 mm	0.05 +/- 0.02	0.020 +/- 0.009	0.0005 +/- 0.0003

Signal events repartition

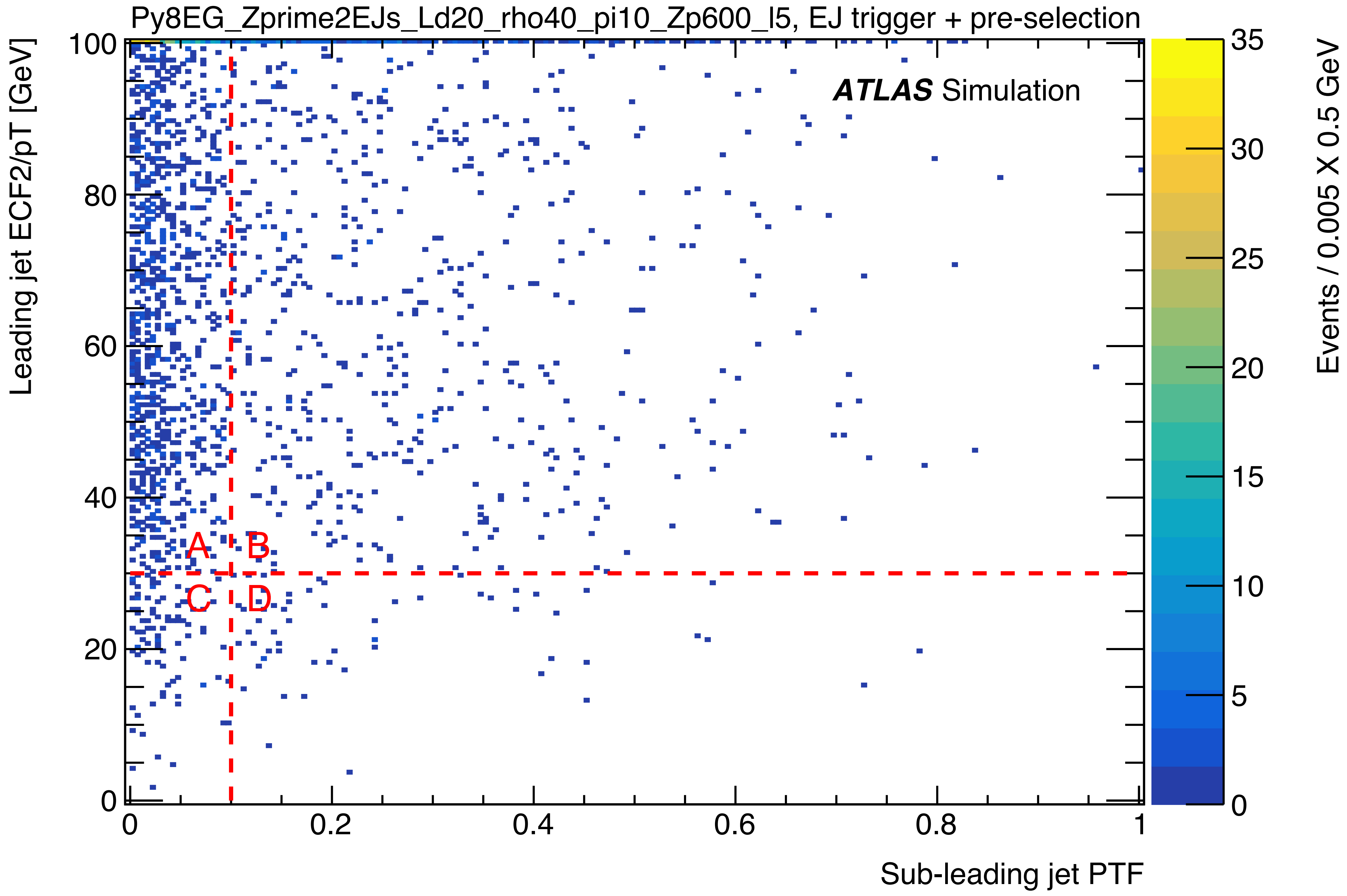


$m_{\pi_d} = 5$ GeV, $c\tau_{\pi_d} = 5$ mm

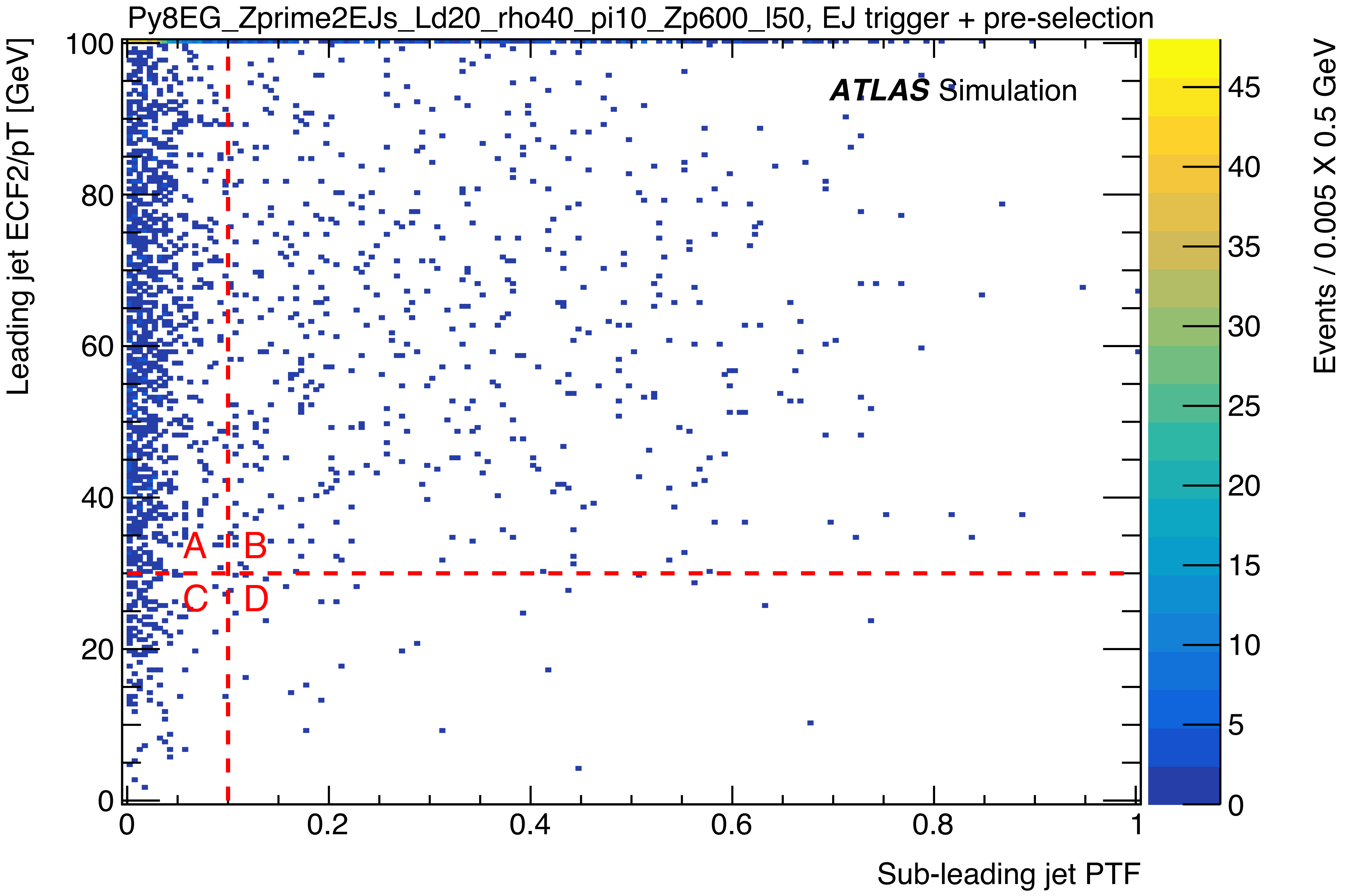


$m_{\pi_d} = 5$ GeV, $c\tau_{\pi_d} = 50$ mm

Signal events repartition

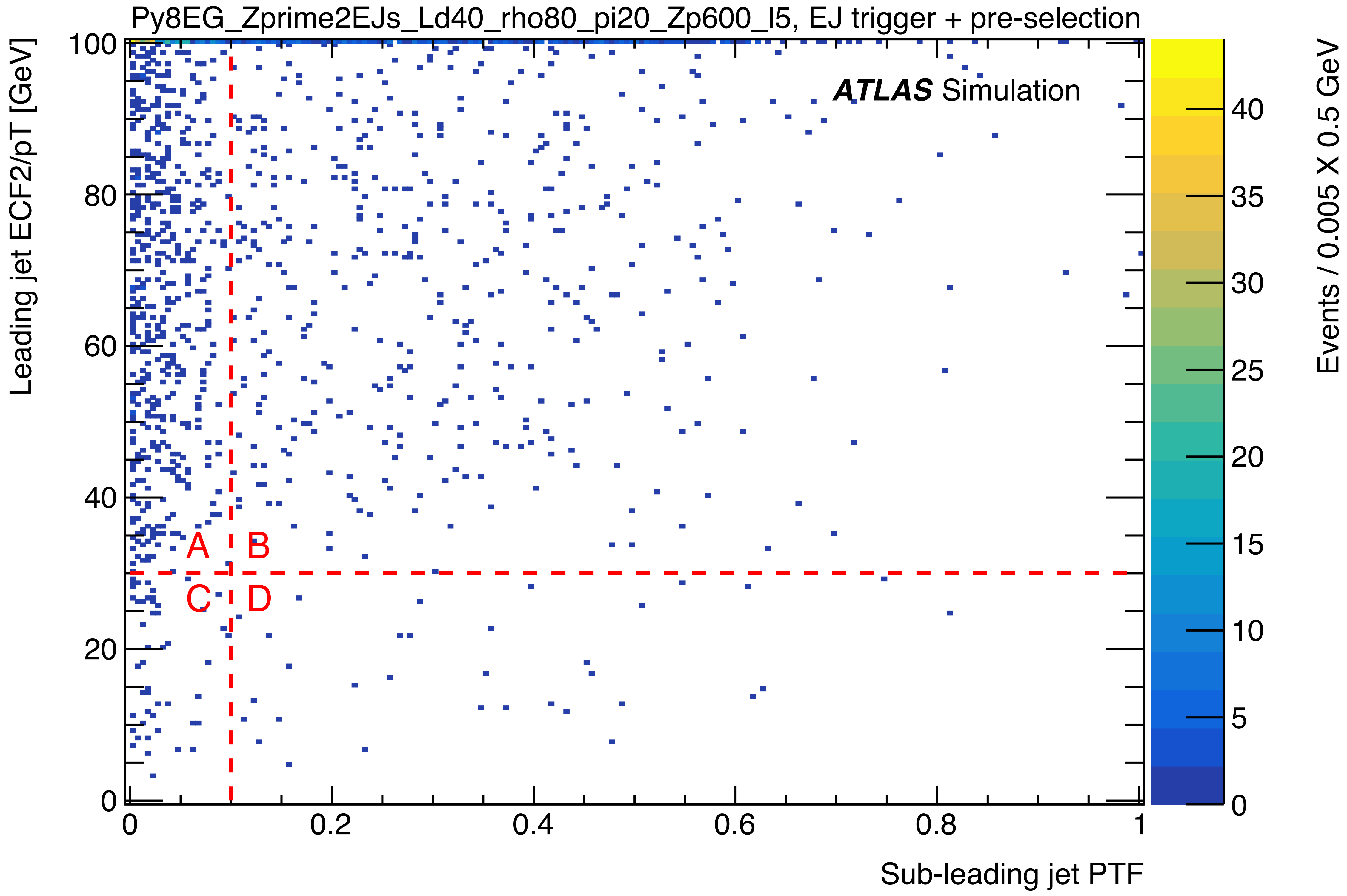


$m_{\pi_d} = 10$ GeV, $c\tau_{\pi_d} = 5$ mm

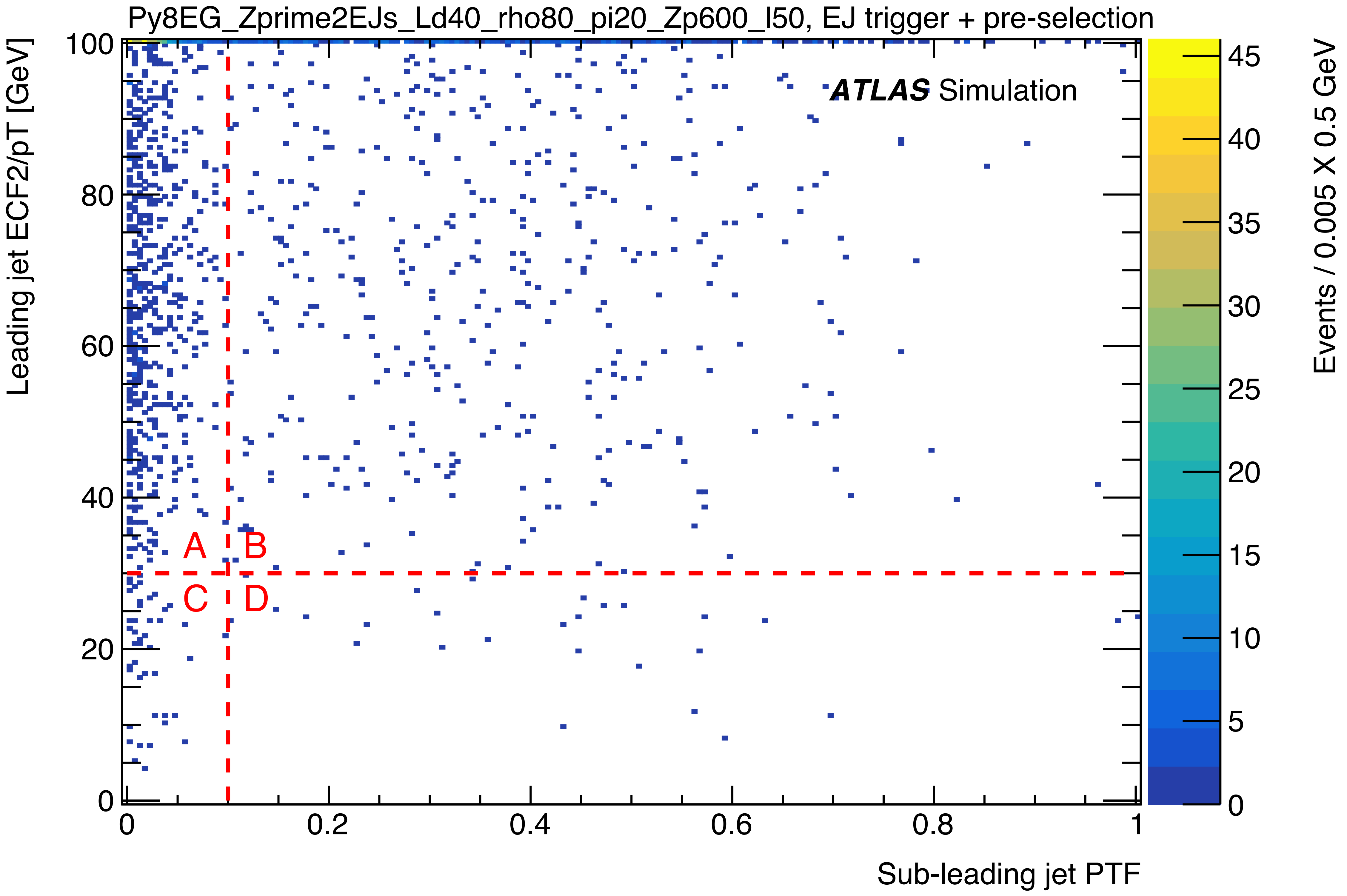


$m_{\pi_d} = 10$ GeV, $c\tau_{\pi_d} = 50$ mm

Signal events repartition



$m_{\pi_d} = 20$ GeV, $c\tau_{\pi_d} = 5$ mm

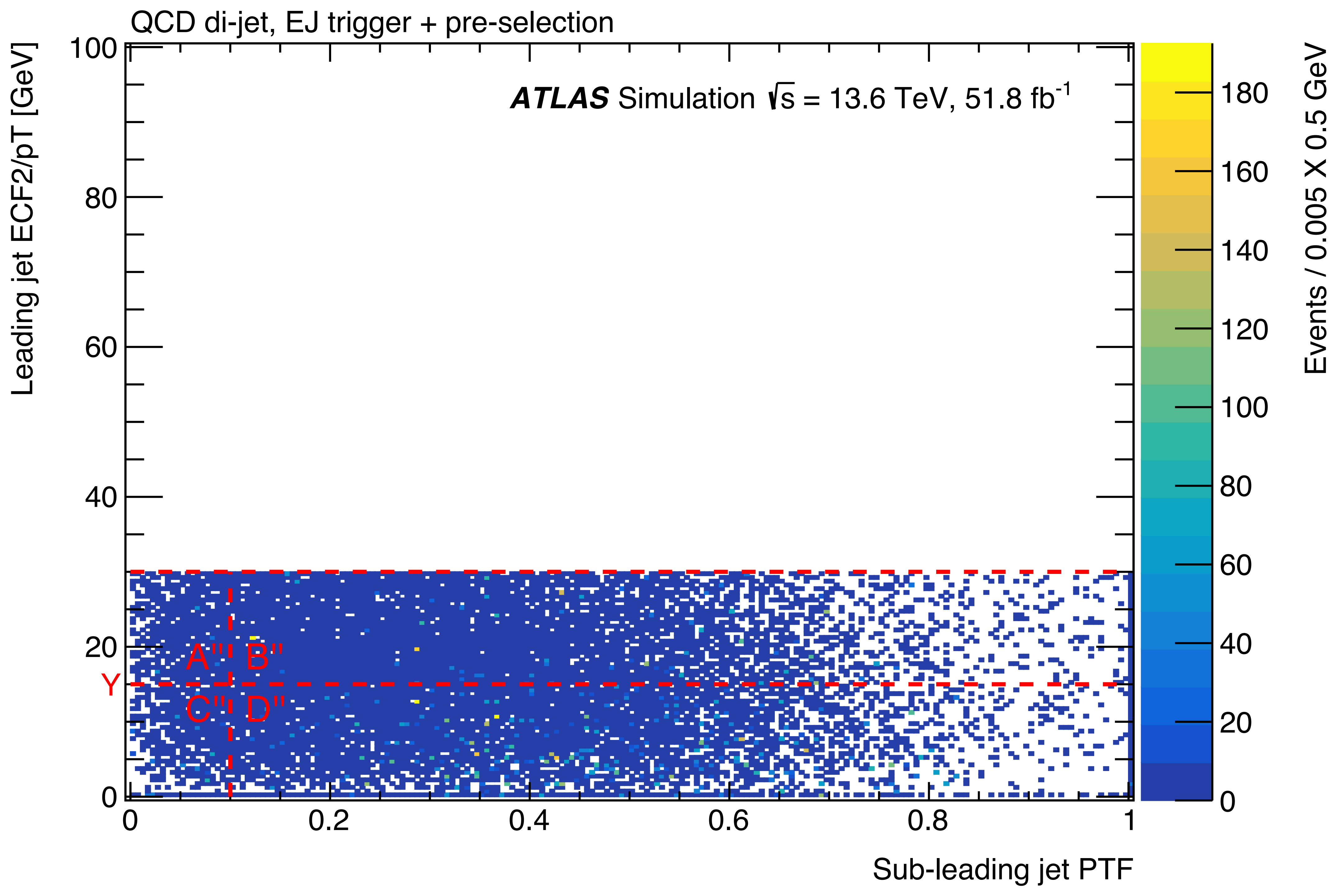


$m_{\pi_d} = 20$ GeV, $c\tau_{\pi_d} = 50$ mm

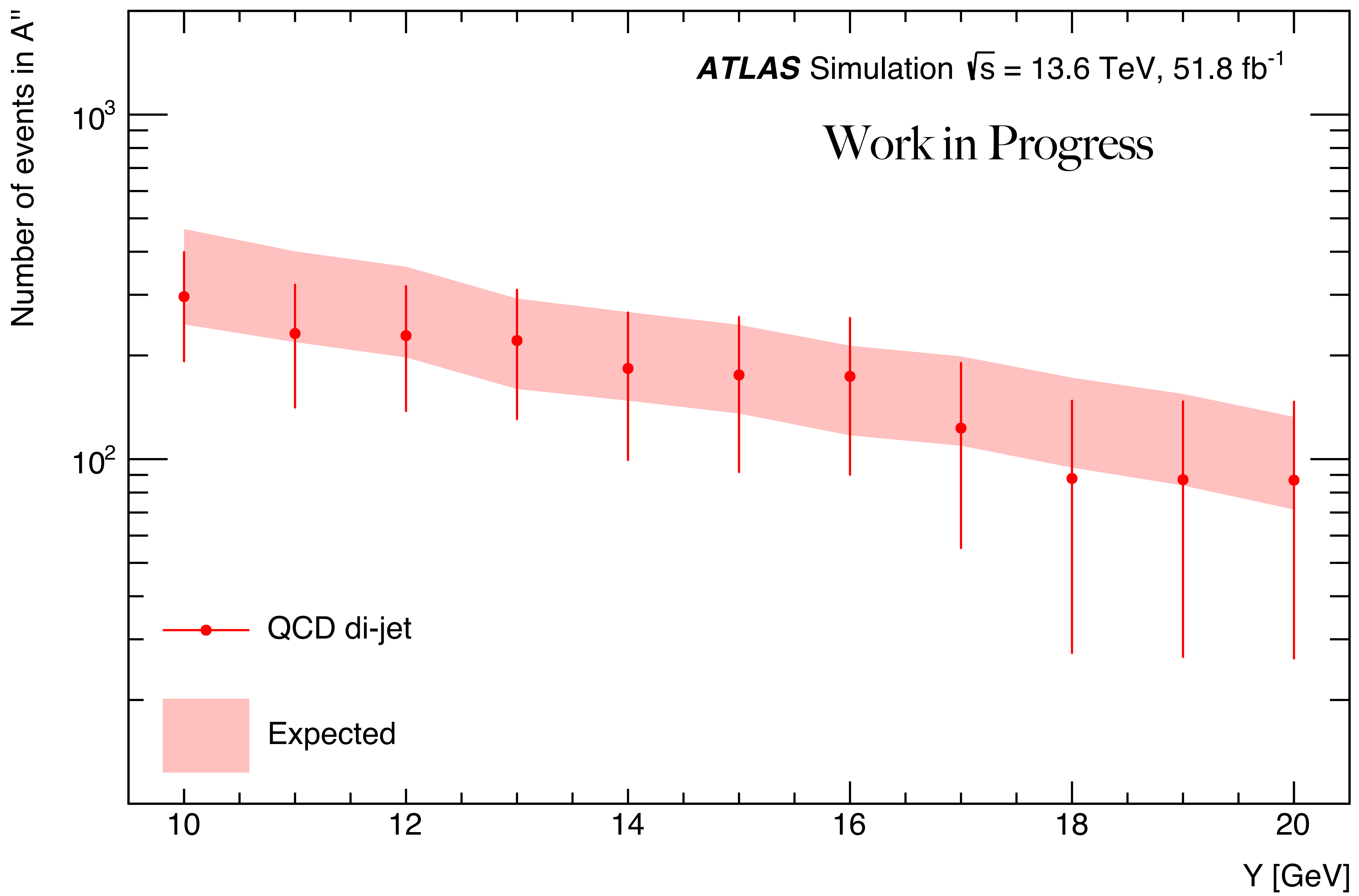
ABCD method in validation regions

- 4 new regions with Y varying

	$PTF_{sub-lead.\ jet} < 0.1$	$PTF_{sub-lead.\ jet} \geq 0.1$
$Y \leq ECF2/p_{T_{lead.\ jet}} < 30$ GeV	A”	B”
$ECF2/p_{T_{lead.\ jet}} < Y$	C”	D”



$$n^{exp}_{A''} = (n^{bkg}_{B''} \times n^{bkg}_{C''})/n^{bkg}_{D''} \pm \text{(MC stat.)}$$



Signal contamination in VR

Signal/Background ratio between A'/A'' and SR : $\frac{N_S(R)/N_B(R)}{N_S(A)/N_B(A)}$ with $R \in \{A', A''\}$