Analysis of fake leptons background for the study of interference between single top and $t\bar{t}$ production with the ATLAS detector

Giulia Maineri

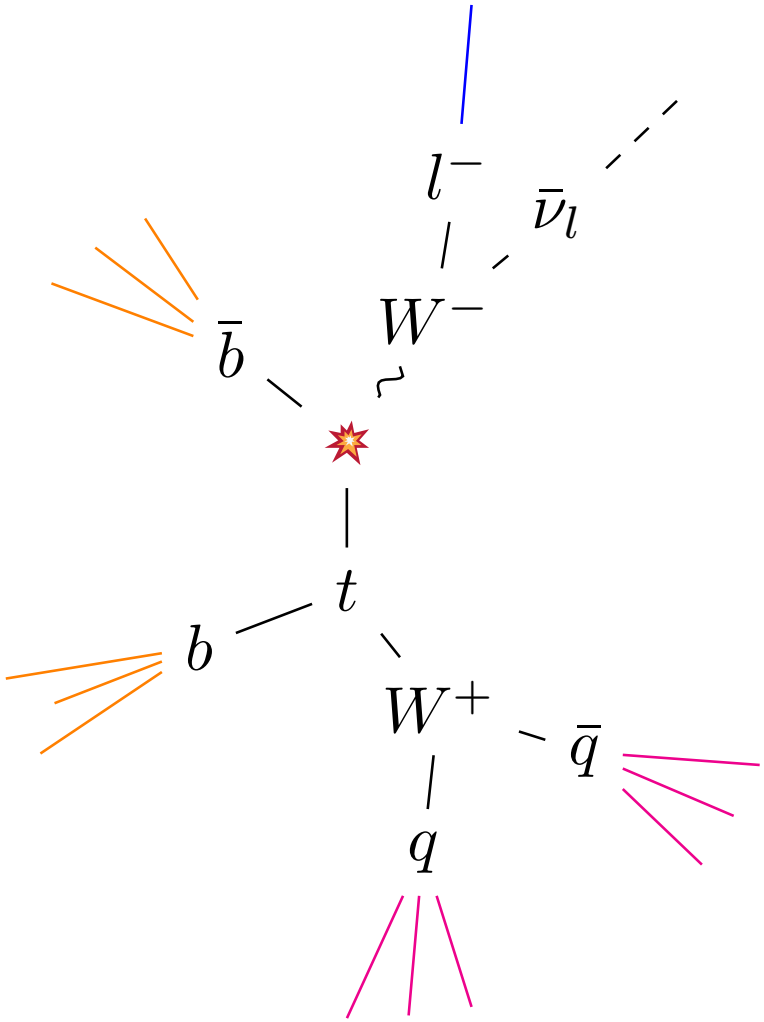
Summer Student Project, supervised by: T. Novak, F. Meloni

6 September 2023

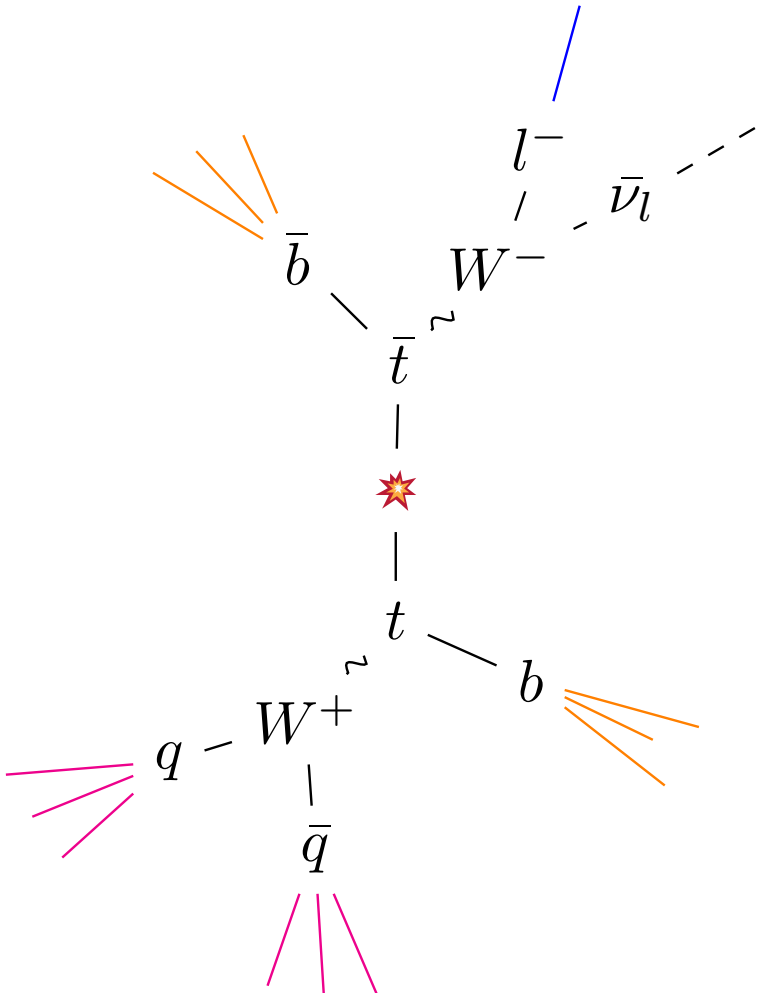
Signal events

= final states *WbWb* which come from:

- single-top production, in association with a W-boson ad a b-quark

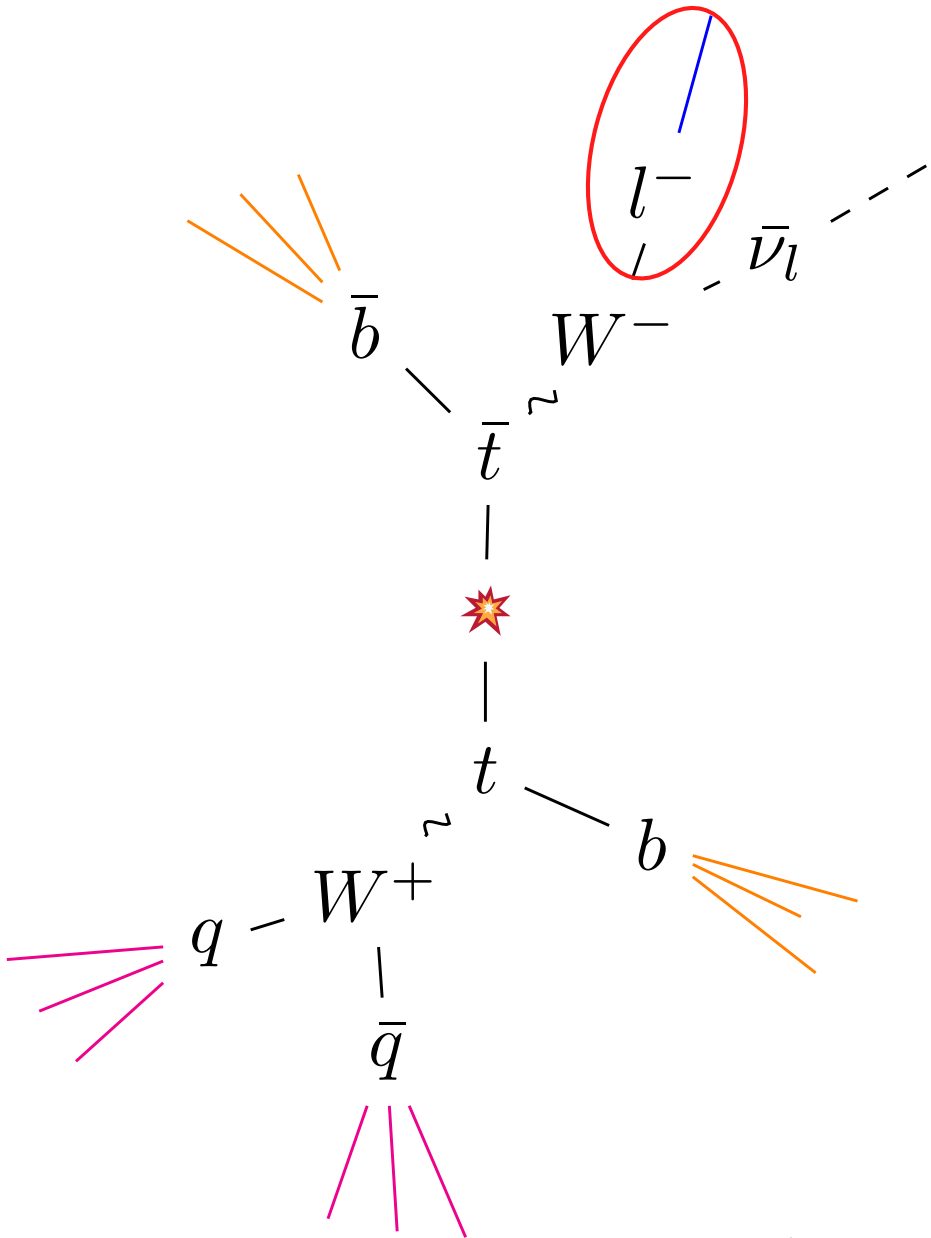


- $t\bar{t}$ production



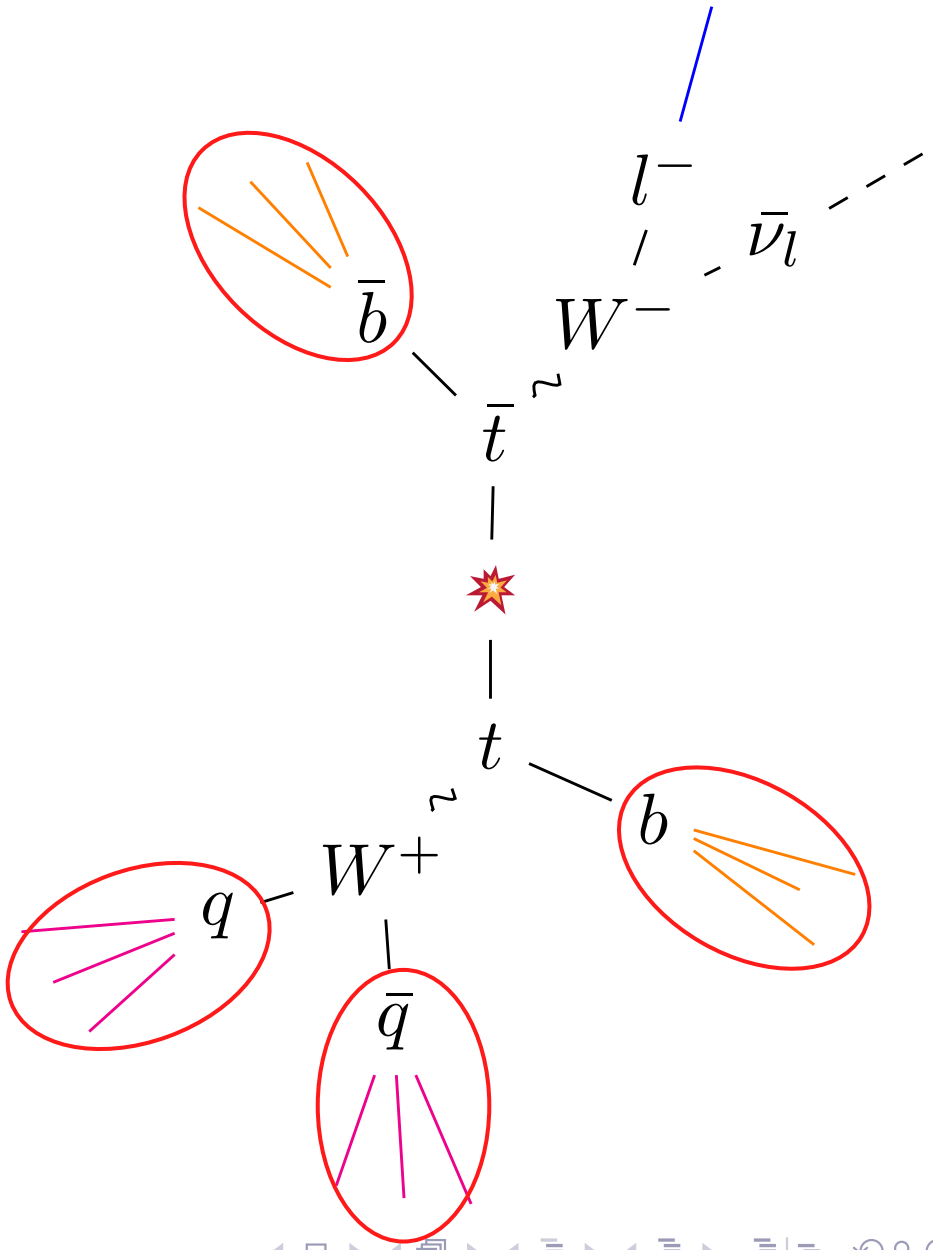
Signal Region

- $N_l = 1,\ l \in \{e, \mu\}$;
- $N_{\text{jets}} \geq 4$;
- $N_{\text{b-jets}} \geq 2$;
- $E_T^{miss} + m_T^l > 60\,\text{GeV}$;
- $p_T^l > 30\,\text{GeV}$;
- $p_T^j > 25\,\text{GeV}$.



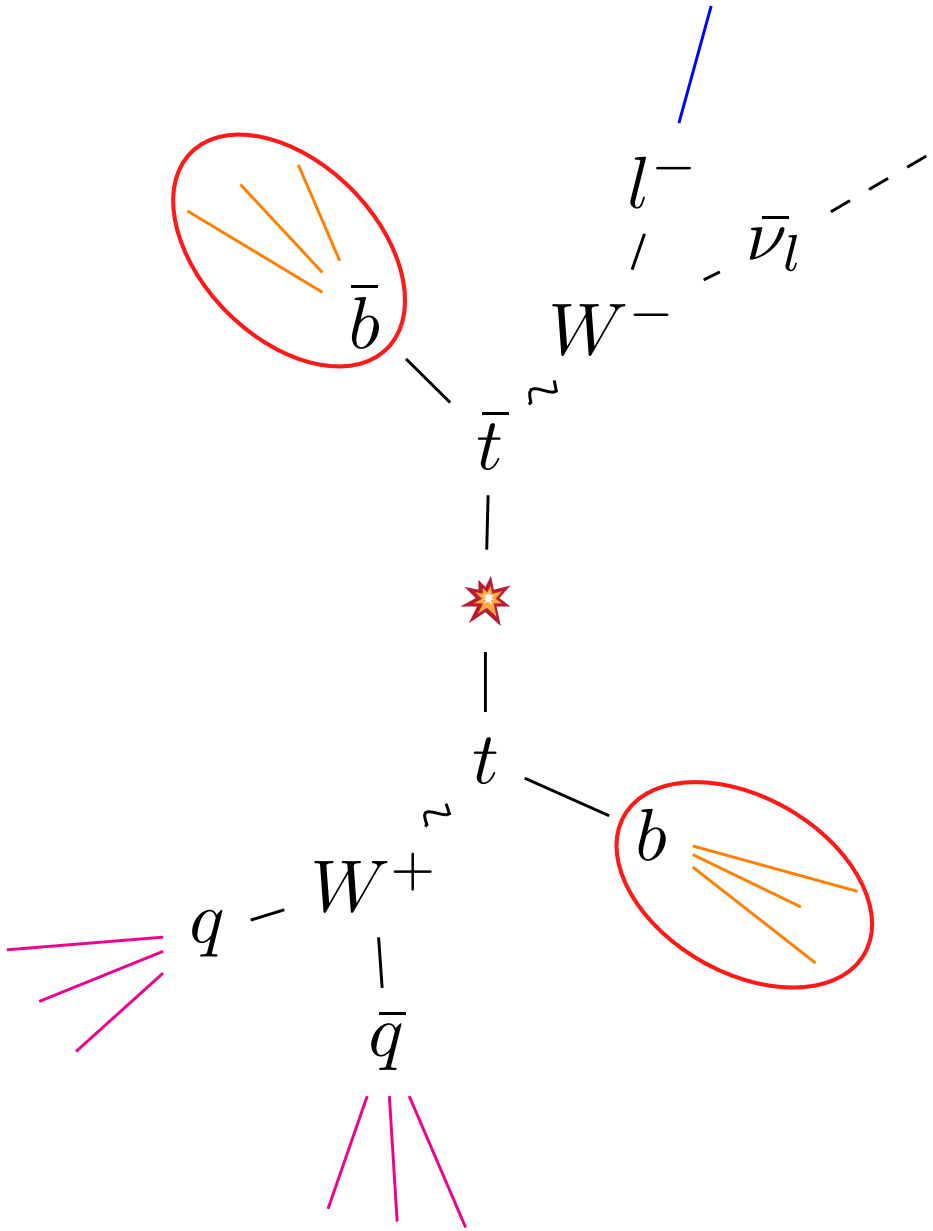
Signal Region

- $N_l = 1,\ l \in \{e, \mu\}$;
- $N_{\text{jets}} \geq 4$;
- $N_{\text{b-jets}} \geq 2$;
- $E_T^{miss} + m_T^l > 60\,\text{GeV}$;
- $p_T^l > 30\,\text{GeV}$;
- $p_T^j > 25\,\text{GeV}$.



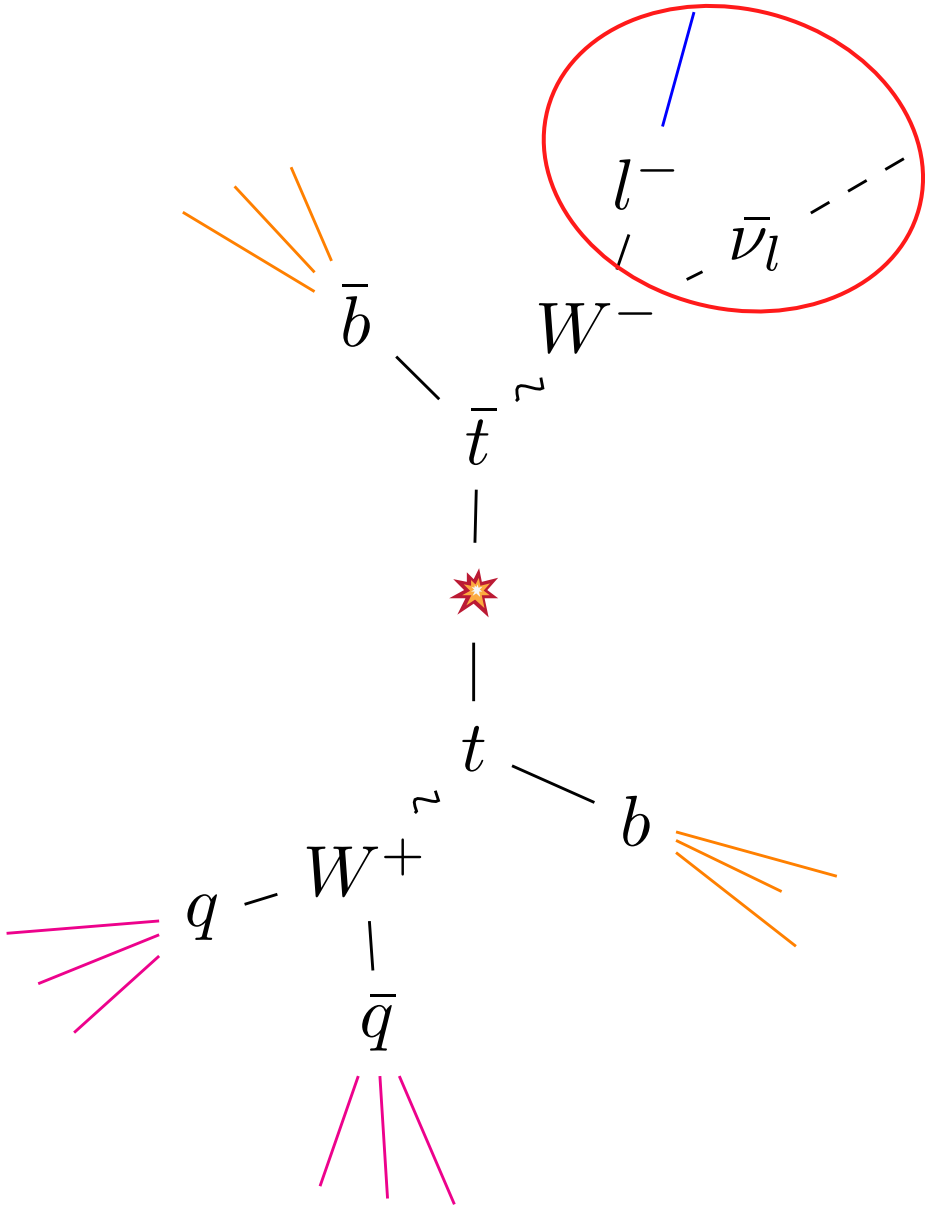
Signal Region

- $N_l = 1,\ l \in \{e, \mu\}$;
- $N_{\text{jets}} \geq 4$;
- $N_{\text{b-jets}} \geq 2$;
- $E_T^{miss} + m_T^l > 60\,\text{GeV}$;
- $p_T^l > 30\,\text{GeV}$;
- $p_T^j > 25\,\text{GeV}$.



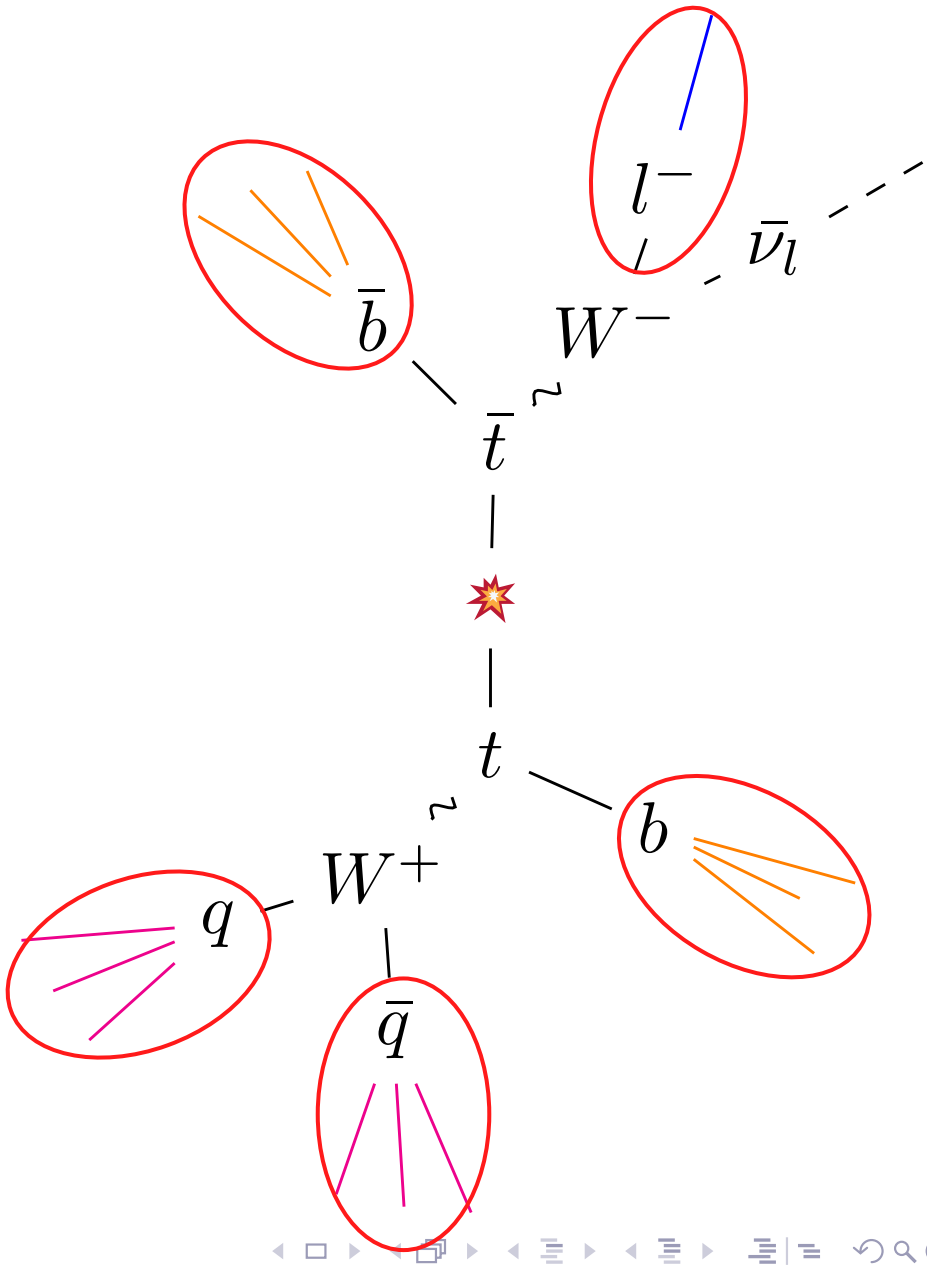
Signal Region

- $N_l = 1,\ l \in \{e, \mu\}$;
- $N_{\text{jets}} \geq 4$;
- $N_{\text{b-jets}} \geq 2$;
- $E_T^{miss} + m_T^l > 60\,\text{GeV}$;
- $p_T^l > 30\,\text{GeV}$;
- $p_T^j > 25\,\text{GeV}$.



Signal Region

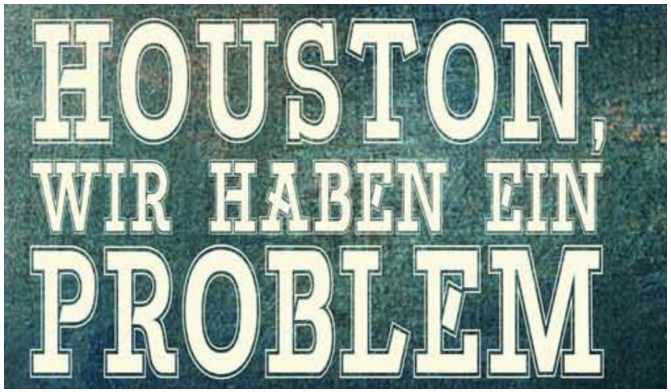
- $N_l = 1,\ l \in \{e, \mu\}$;
- $N_{\text{jets}} \geq 4$;
- $N_{\text{b-jets}} \geq 2$;
- $E_T^{miss} + m_T^l > 60\,\text{GeV}$;
- $p_T^l > 30\,\text{GeV}$;
- $p_T^j > 25\,\text{GeV}$.



Signal Region

- $N_l = 1,\ l \in \{e, \mu\}$;
- $N_{\text{jets}} \geq 4$;
- $N_{\text{b-jets}} \geq 2$;
- $E_T^{miss} + m_T^l > 60\,\text{GeV}$;
- $p_T^l > 30\,\text{GeV}$;
- $p_T^j > 25\,\text{GeV}$.

$\Longrightarrow$ lepton can be fake!

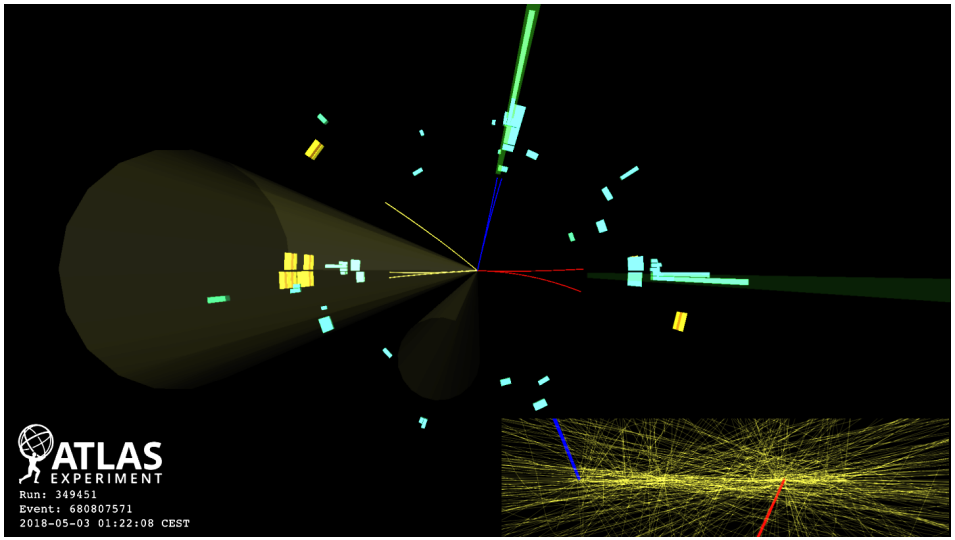


Fake leptons background analysis

Fake Leptons Sources

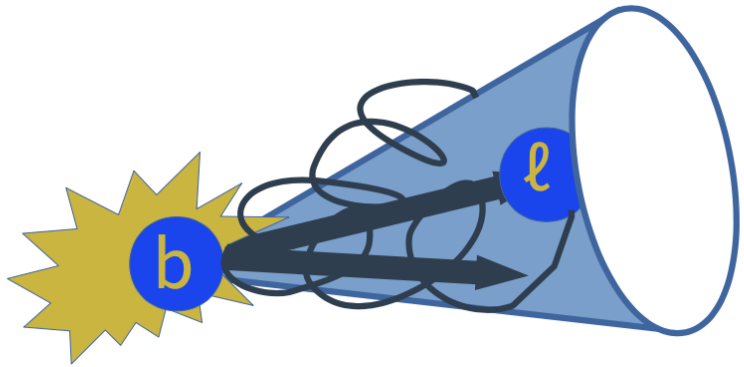
Fake electrons sources

- jets with large EM energy;
- photon conversions;
- semileptonic decays of charm and bottom hadrons.



Fake muons sources

- pions or kaons decays;
- energetic particles that reach the muon spectrometer;
- semileptonic decays of charm and bottom hadrons.



Fake Factors method: a data-driven technique

low-statistic
tighter signal region

high-statistic looser
control region

$$F = \left(\frac{N^{data}_{tight} - N^{MC}_{tight}}{N^{data}_{loose} - N^{MC}_{loose}} \right)_{CR}$$

Fake Factors method: a data-driven technique

low-statistic
tighter signal region

high-statistic looser
control region

$$F = \left(\frac{N^{data}_{tight} - N^{MC}_{tight}}{N^{data}_{loose} - N^{MC}_{loose}} \right)_{CR}$$

background enriched
Control Region

Fake Factors method: a data-driven technique

- Tight electrons: isolation + identification
- Tight muons: only isolation

low-statistic
tighter signal region

high-statistic looser
control region

$$F = \left(\frac{N^{data}_{tight} - N^{MC}_{tight}}{N^{data}_{loose} - N^{MC}_{loose}} \right)_{CR}$$

looser requirements

background enriched
Control Region

Fake Factors method: a data-driven technique

- Tight electrons: isolation + identification
- Tight muons: only isolation

low-statistic
tighter signal region

high-statistic looser
control region

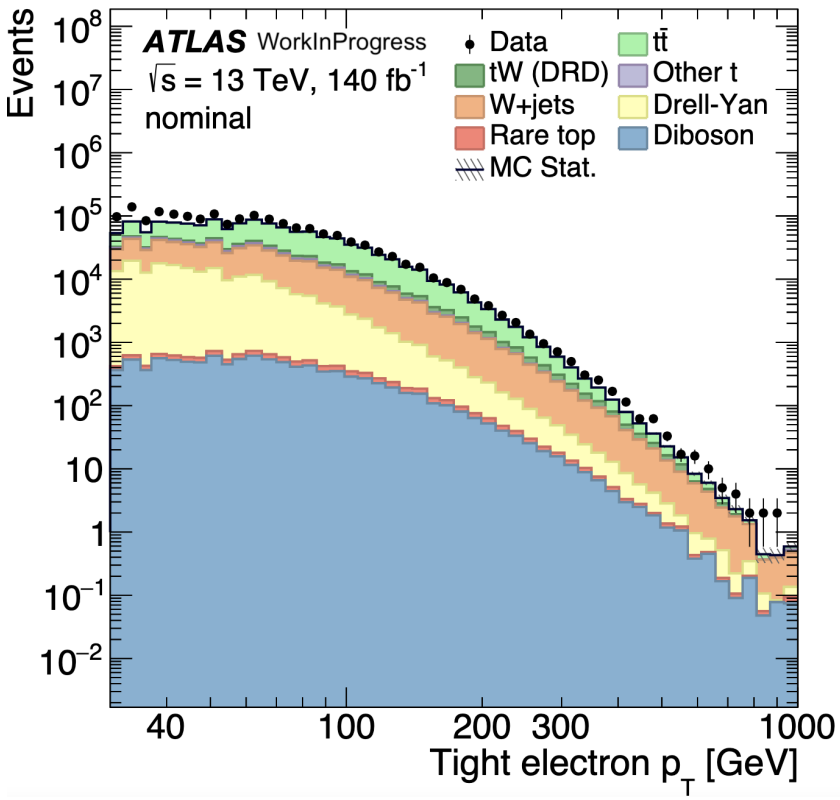
$$F = \left(\frac{N^{data}_{tight} - N^{MC}_{tight}}{N^{data}_{loose} - N^{MC}_{loose}} \right) CR$$

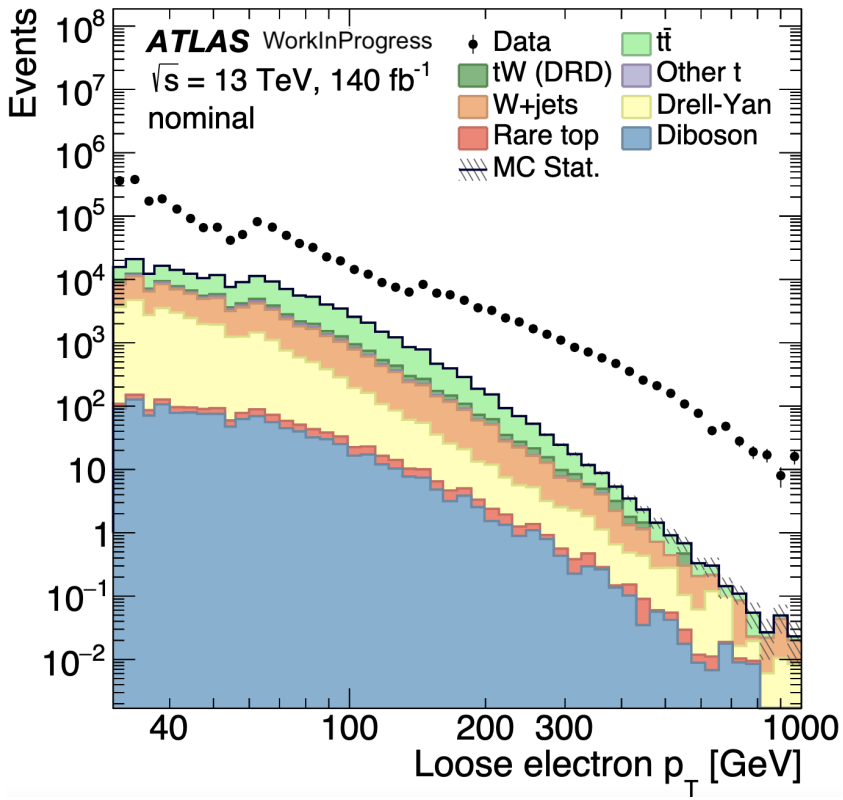
subtraction of MC predictions

background enriched
Control Region

looser requirements

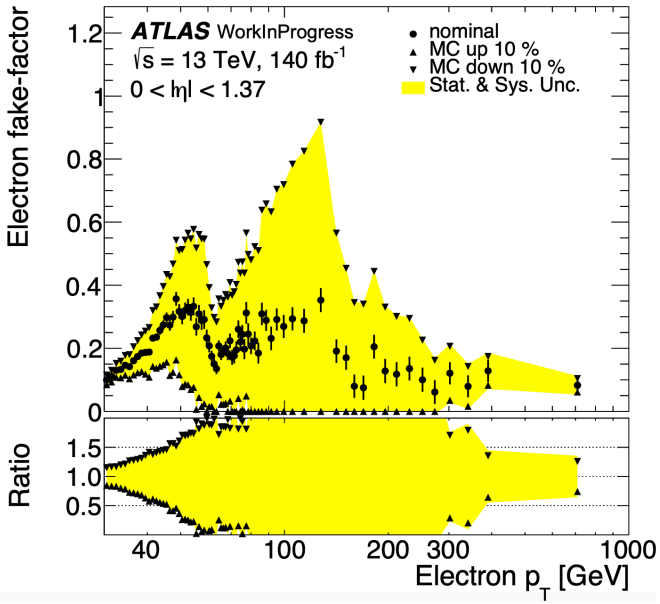
Fake Factors method: a data-driven technique



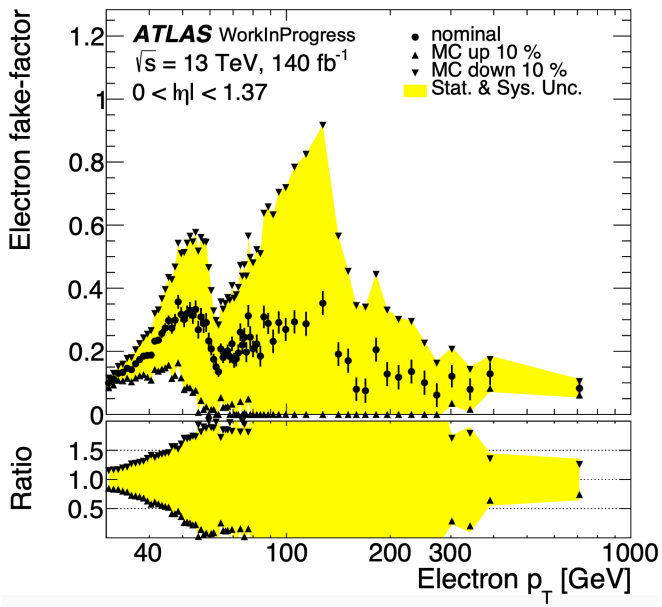


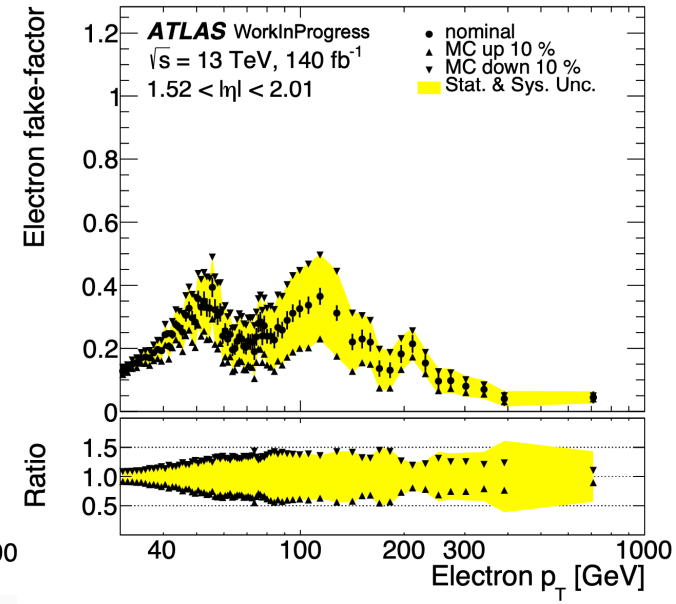
p_T distributions of tight (lhs) and loose (rhs) electrons in nominal Control Region

Nominal Control Region

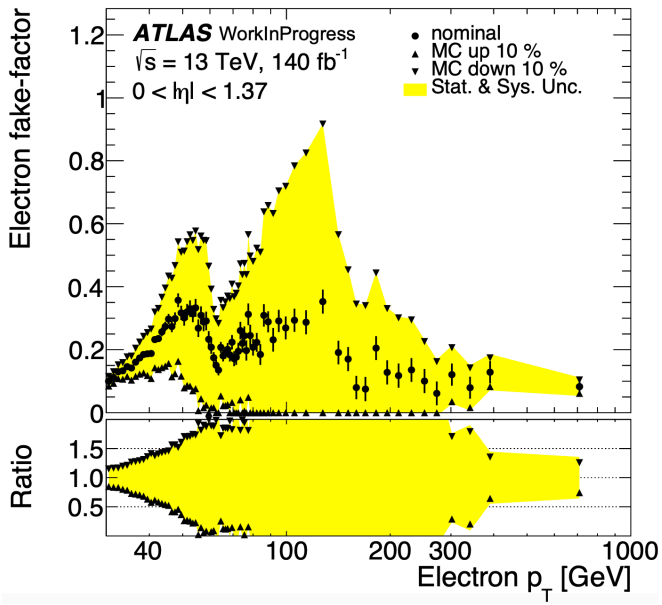


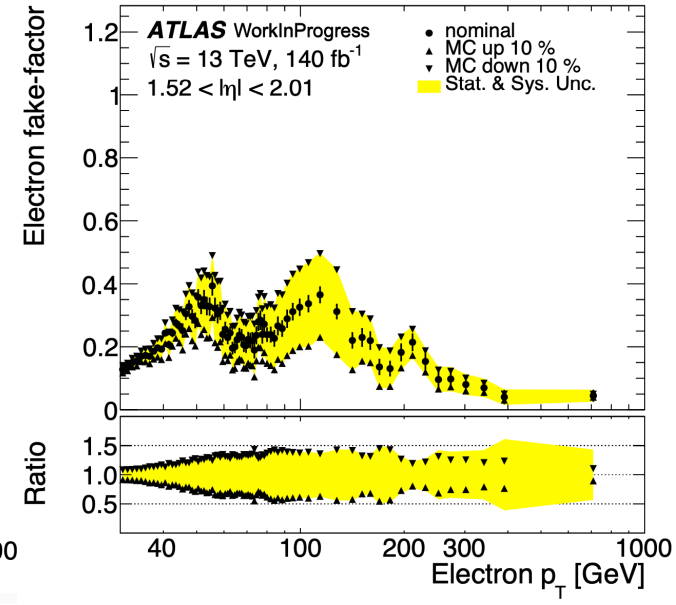
Nominal Control Region

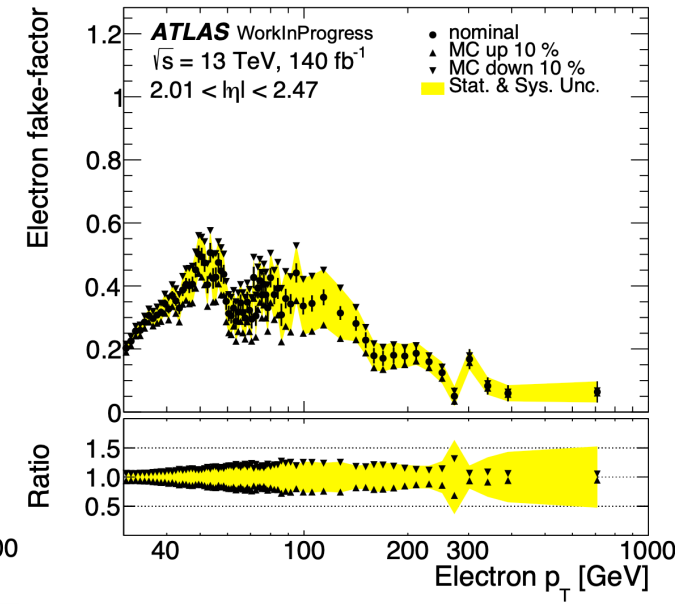




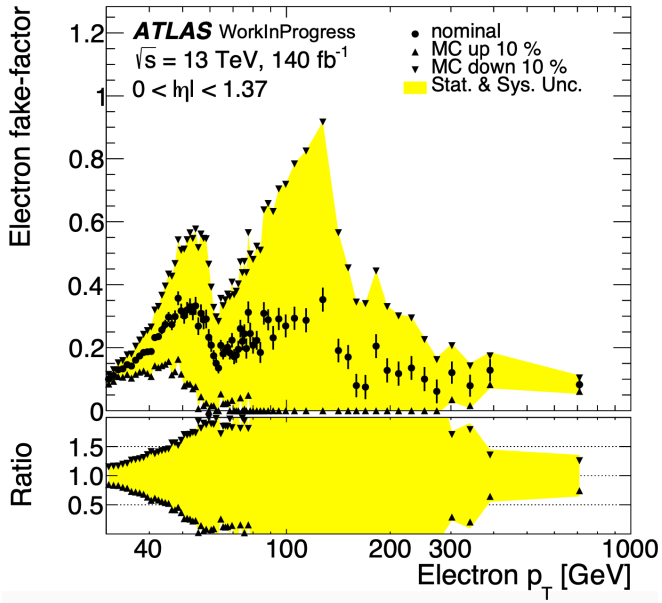
Nominal Control Region

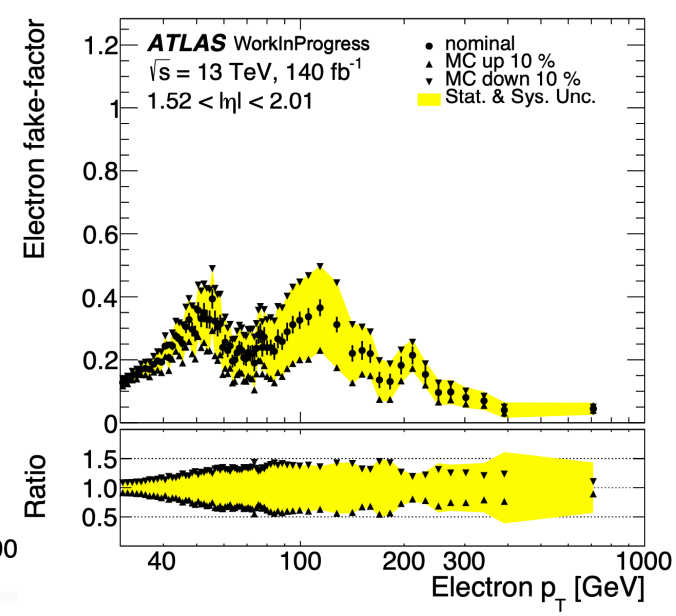


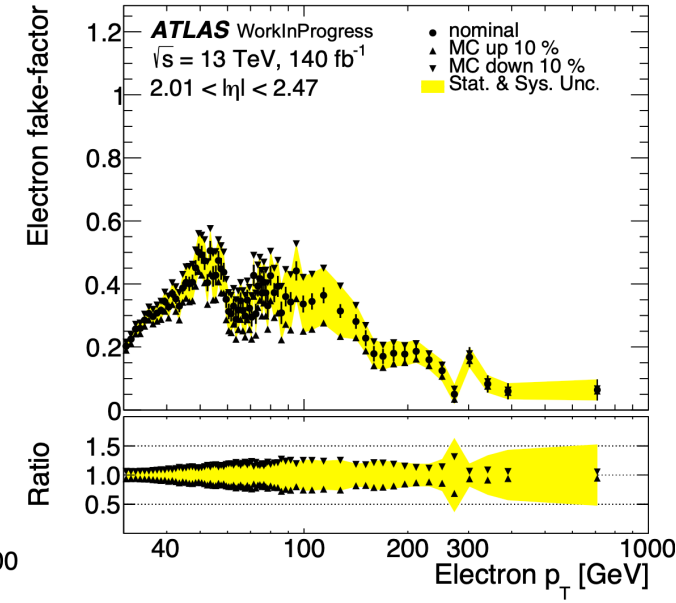




Nominal Control Region



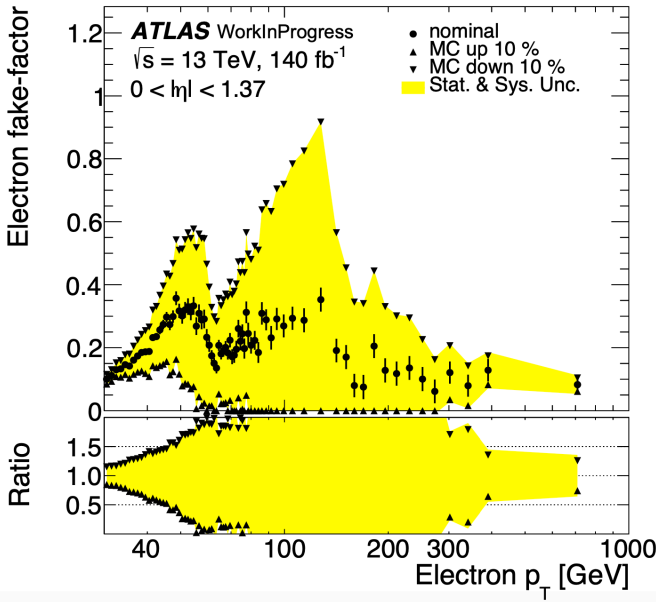


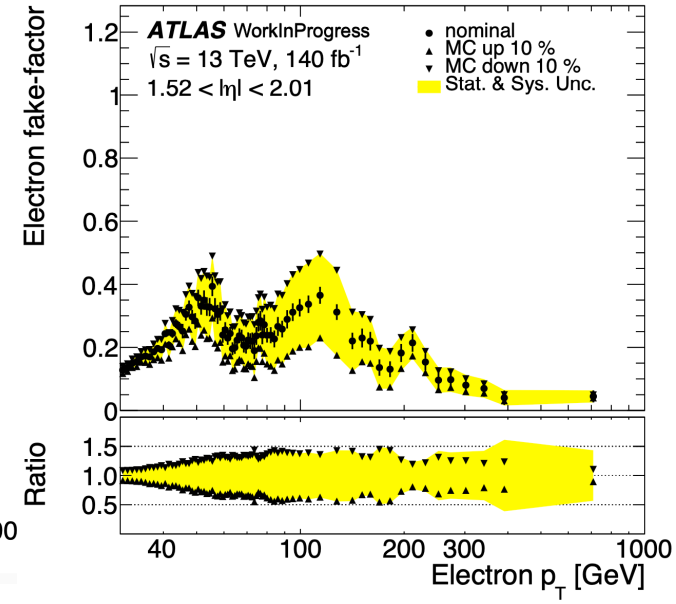


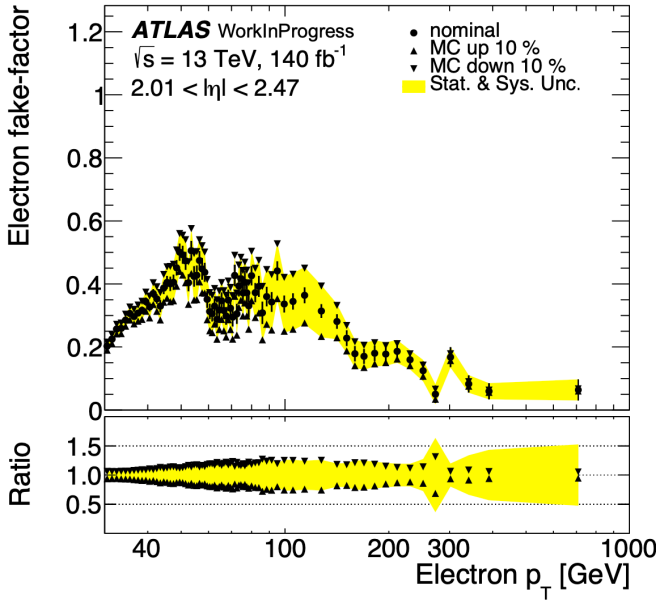
lower η
$\rightarrow$ lower number of fakes
$\rightarrow$ MC variation higher effect

higher η
$\rightarrow$ higher number of fakes
$\rightarrow$ MC variation smaller effect

Nominal Control Region



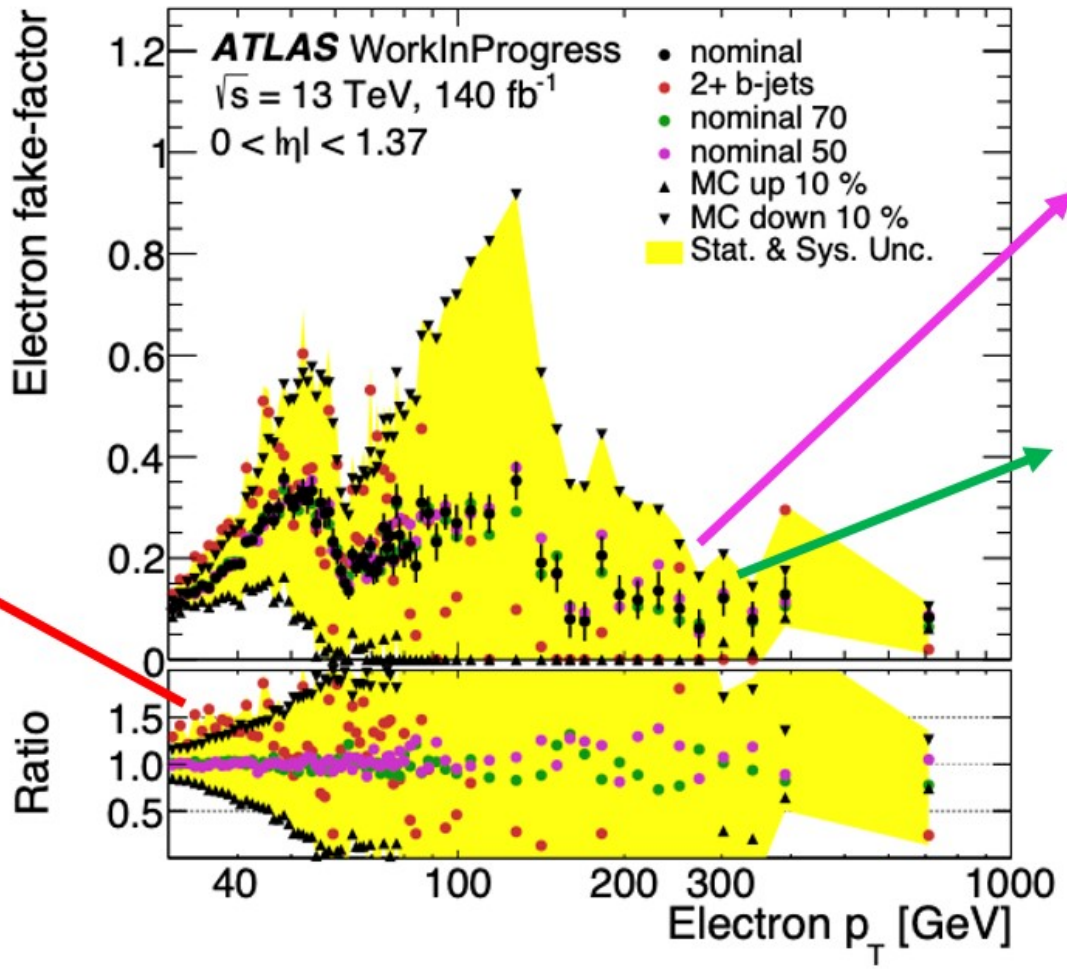




Nominal Selection:

- $N_e = 1$;
- $N_{jets} \geq 3$;
- $N_{b-jet} \geq 1$;
- $p_T^l > 30\,\text{GeV}$;
- $E_T^{miss} + m_T^l < 60\,\text{GeV}$.

Modified Control Region

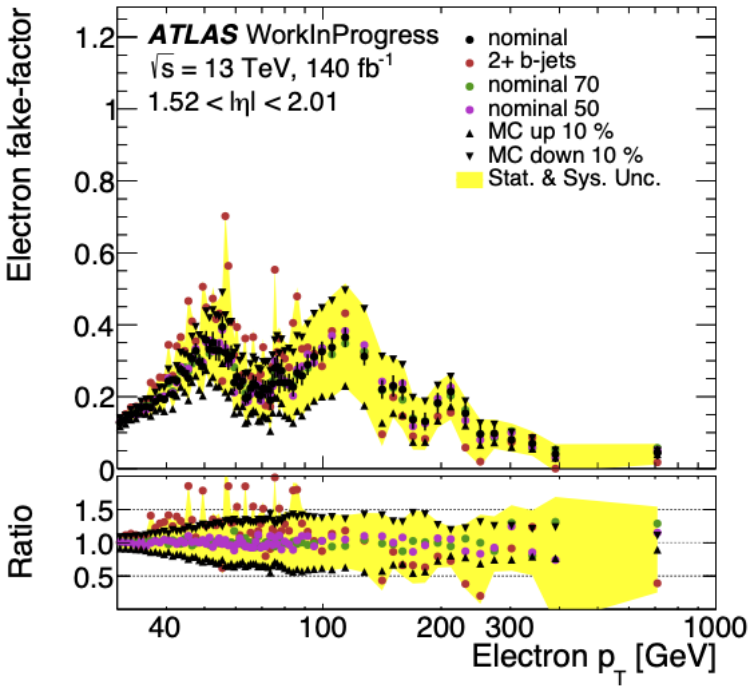


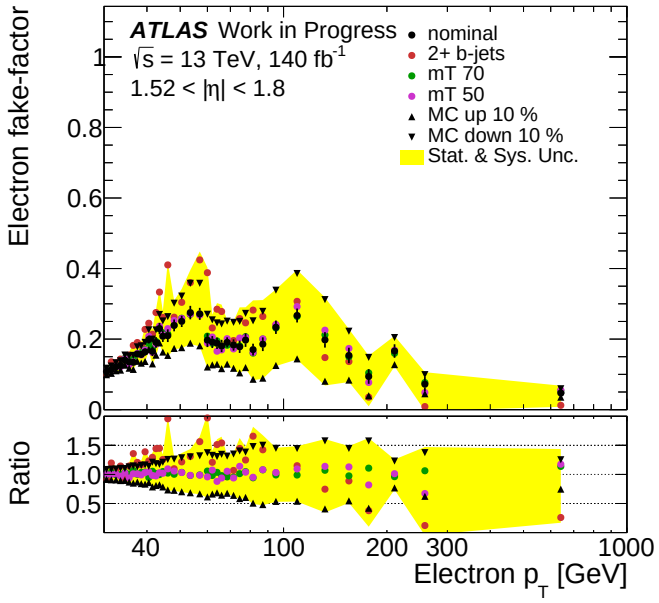
$N_{b-jet} \geq 2$

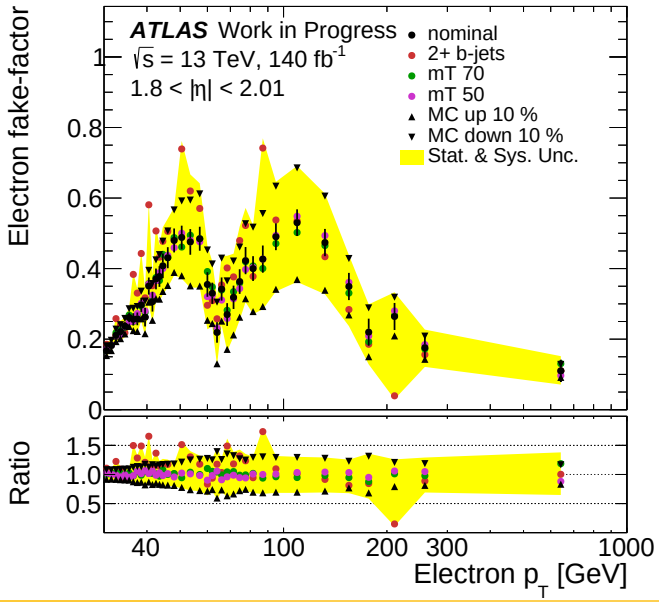
$E_T^{miss} + m_T^l < 50\,\text{GeV}$

$E_T^{miss} + m_T^l < 70\,\text{GeV}$

Split Control Region

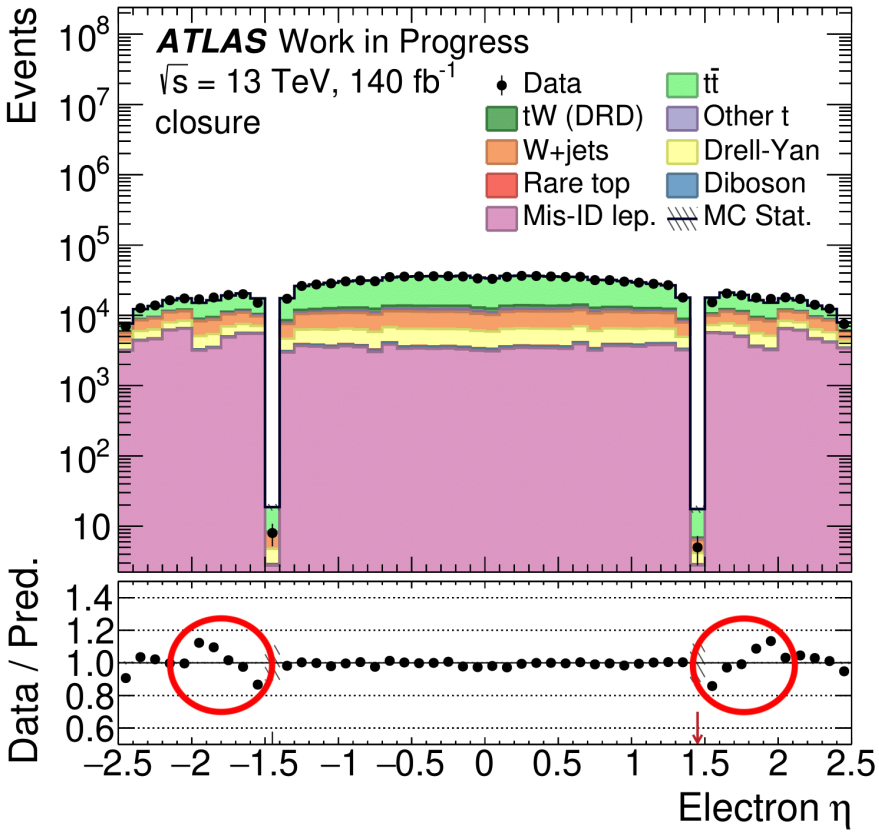




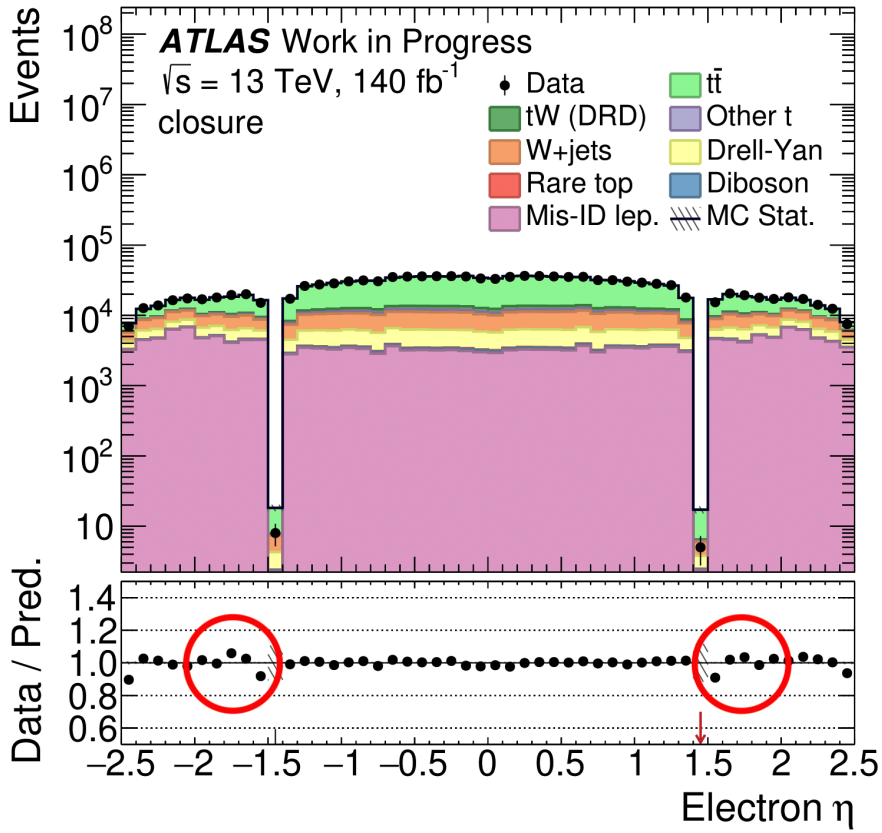


Closure Test: η distribution

η distributions of electrons in Closure Region for Fake Factors



3 η bins



4 η bins

Summary

- The **fake leptons background** for the analysis of *WbWb* final states have been studied;

Summary

- The **fake leptons background** for the analysis of *WbWb* final states have been studied;
- A **data-driven** technique has been used, in order to achieve **higher precision** estimations;

Summary

- The **fake leptons background** for the analysis of *WbWb* final states have been studied;
- A **data-driven** technique has been used, in order to achieve **higher precision** estimations;
- **Fake Factor** distributions have been computed in the Control Regions defined with nominal selection;

Summary

- The **fake leptons background** for the analysis of *WbWb* final states have been studied;
- A **data-driven** technique has been used, in order to achieve **higher precision** estimations;
- **Fake Factor** distributions have been computed in the Control Regions defined with nominal selection;
- The stability of the Fake Factor method has been tested by modifying the definition of the **Control Regions**;

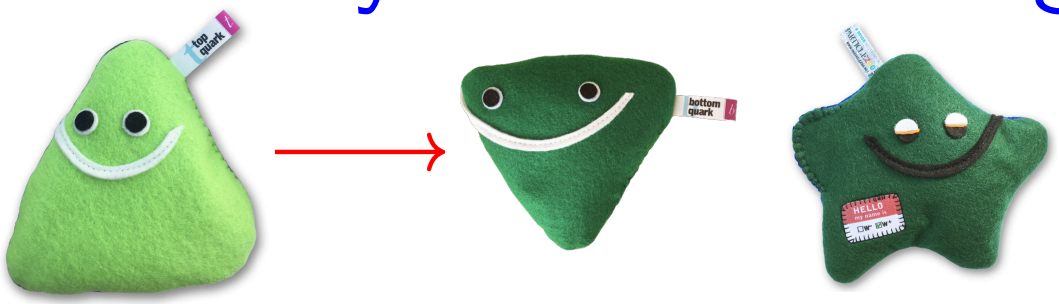
Summary

- The **fake leptons background** for the analysis of *WbWb* final states have been studied;
- A **data-driven** technique has been used, in order to achieve **higher precision** estimations;
- **Fake Factor** distributions have been computed in the Control Regions defined with nominal selection;
- The stability of the Fake Factor method has been tested by modifying the definition of the **Control Regions**;
- A **binning optimization** has been performed to obtain a better agreement between data and Monte Carlo in the Closure test.

Summary

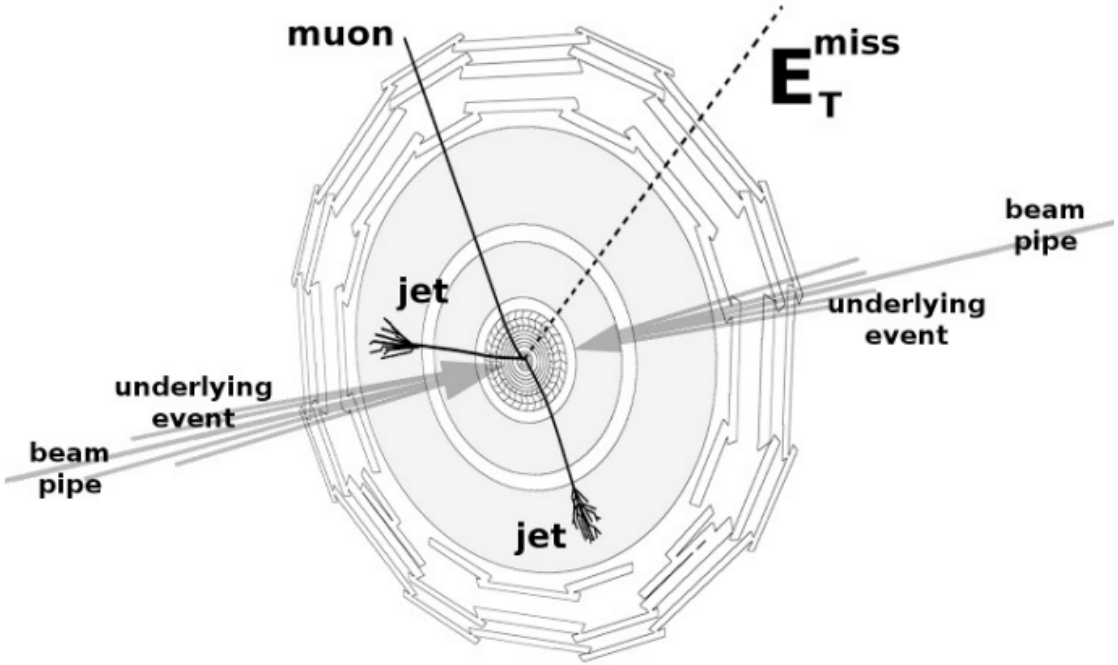
- The **fake leptons background** for the analysis of *WbWb* final states have been studied;
- A **data-driven** technique has been used, in order to achieve **higher precision** estimations;
- **Fake Factor** distributions have been computed in the Control Regions defined with nominal selection;
- The stability of the Fake Factor method has been tested by modifying the definition of the **Control Regions**;
- A **binning optimization** has been performed to obtain a better agreement between data and Monte Carlo in the Closure test.

Thank you for listening



Backup

Missing transverse momentum

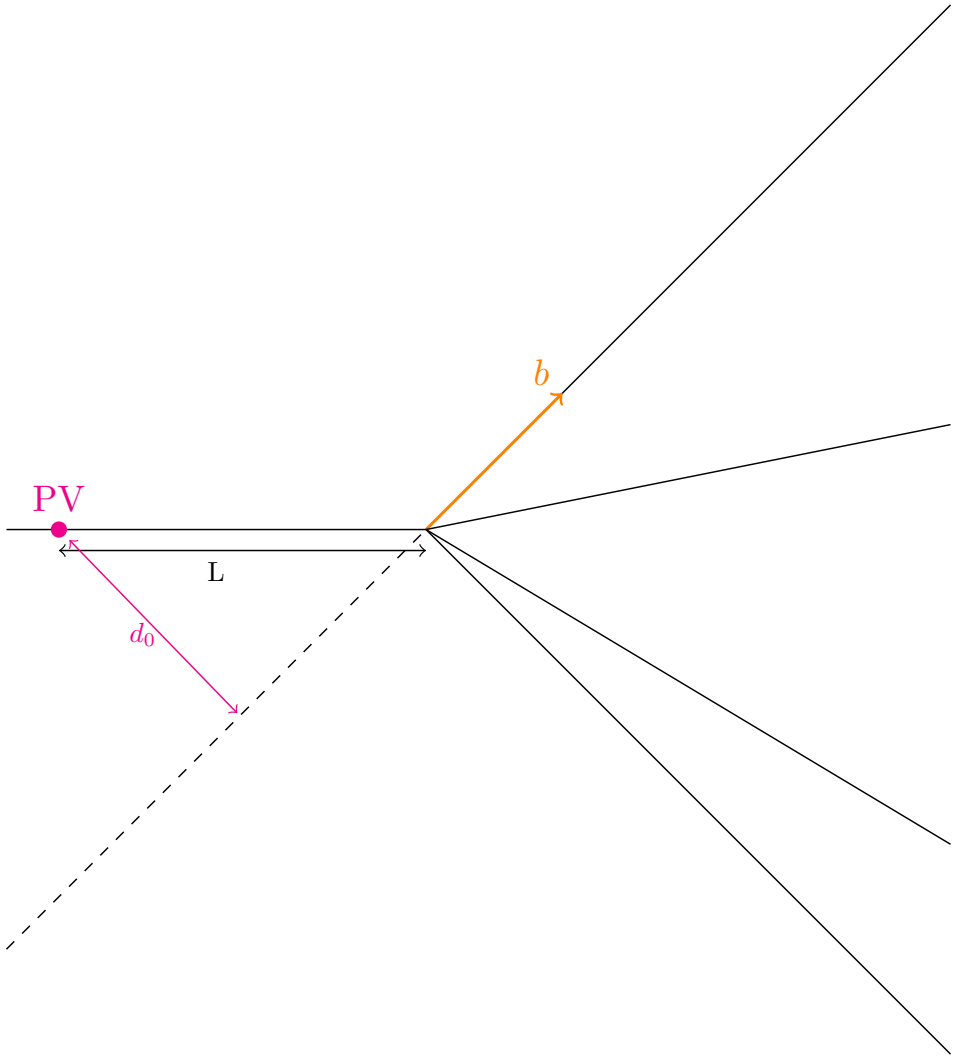


$$\vec{E}_T^{miss} = -\left[\sum_e \vec{p}_T^{(e)} + \sum_\mu \vec{p}_T^{(\mu)} + \sum_\gamma \vec{p}_T^{(\gamma)} + \sum_\tau \vec{p}_T^{(\tau)} + \sum_{jet} \vec{p}_T^{(jet)} + \sum_x \vec{p}_T^{(x)}\right]$$

B-tagged jets

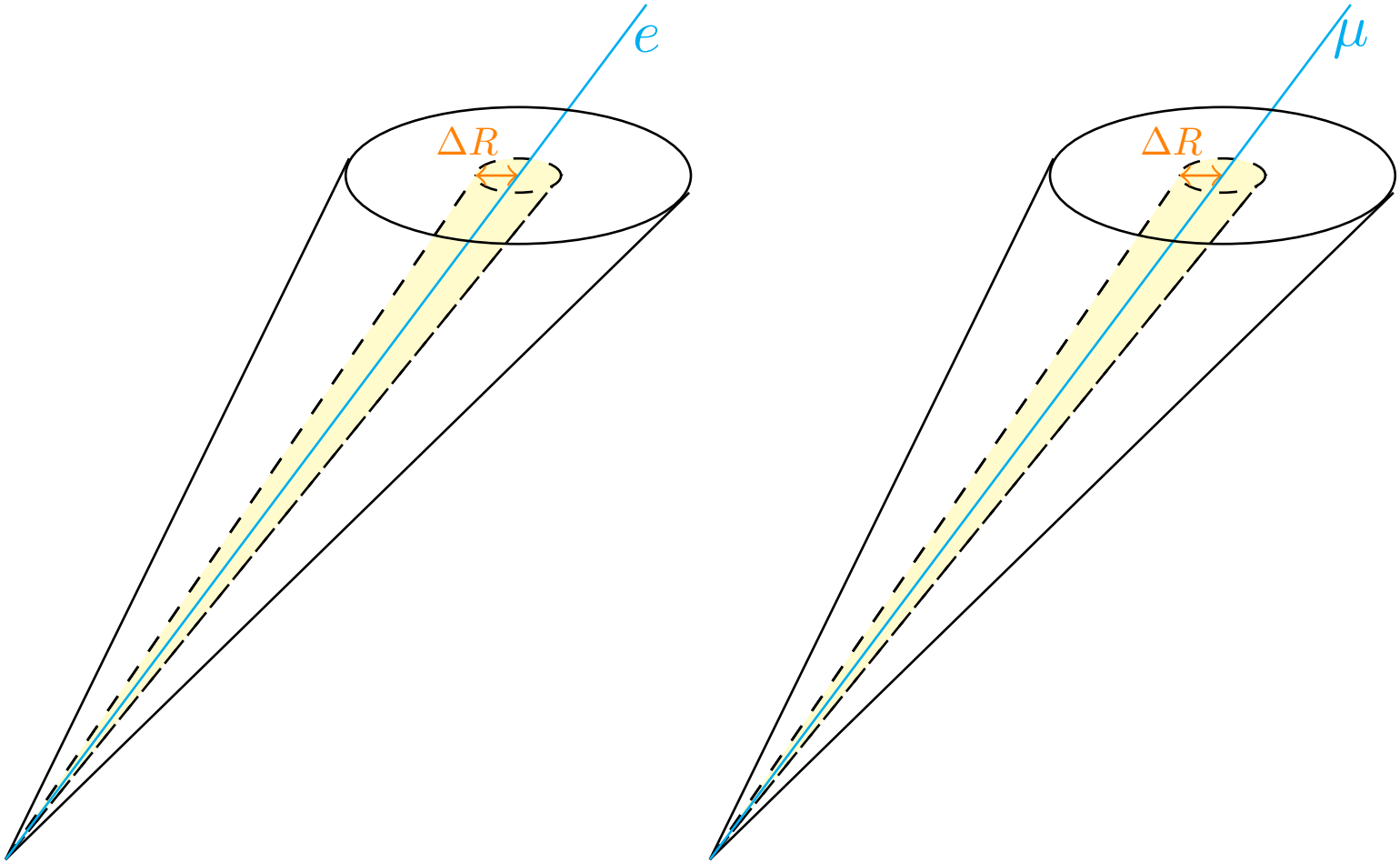
How to recognize a b-jets?

$\Longrightarrow$ **b-tagging**, with methods based on lifetimes, masses and decay topologies



Tight and loose leptons

- isolation requirement



- tight: tracks+calorimeter clusters
- loose: only tracks

Tight and loose leptons

- identification requirement
  $\Longrightarrow$ "quality" of the particle reconstruction process
  $\Longrightarrow$ likelihood based
  - Loose objects, i.e. objects required to have:
    - $p_T > 10\,\text{GeV}$;
    - $|\eta| < 2.37$
    - narrow cluster in the middle layer of the ECal;
    - little energy in the HCal.
  - Tight object, i.e. objects required to have:
    - $p_T > 10\,\text{GeV}$;
    - $|\eta| < 1.37$ or $1.52 < |\eta| < 2.37$ (avoid crack regions);
    - cluster with a well-defined shaped in the middle layer of the ECal;
    - almost no energy leakage in the HCal.