



MINISTÈRE  
DE L'ENSEIGNEMENT SUPÉRIEUR,  
DE LA RECHERCHE  
ET DE L'INNOVATION



PHAST  
PHYSIQUE  
ET ASTROPHYSIQUE  
UNIVERSITÉ DE LYON



Lyon 1

# Recherche d'un boson de Higgs de haute masse se désintégrant en paire de taus dans l'expérience CMS au LHC

Soutenance de thèse de doctorat

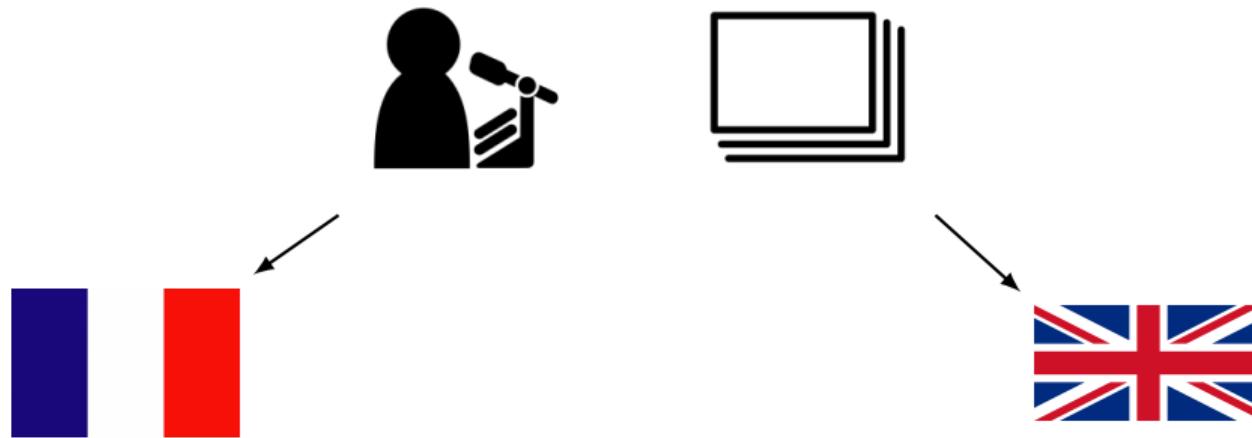
Lucas TORTEROTOT

Institut de Physique des deux Infinis – Lyon

XX xxxx 2021



# Lang(u)age





MINISTÈRE  
DE L'ENSEIGNEMENT SUPÉRIEUR,  
DE LA RECHERCHE  
ET DE L'INNOVATION  
RÉPUBLIQUE FRANÇAISE



PHAST  
PHYSIQUE  
ET ASTROPHYSIQUE  
UNIVERSITÉ DE LYON



# Search for additional neutral Higgs bosons decaying to tau lepton pair in the CMS experiment at LHC

Ph.D. thesis defense

Lucas TORTEROTOT

Institut de Physique des deux Infinis – Lyon

xxxx XX<sup>st/nd/rd/th</sup> 2021



# Keywords in title

Why do we **search for...?**

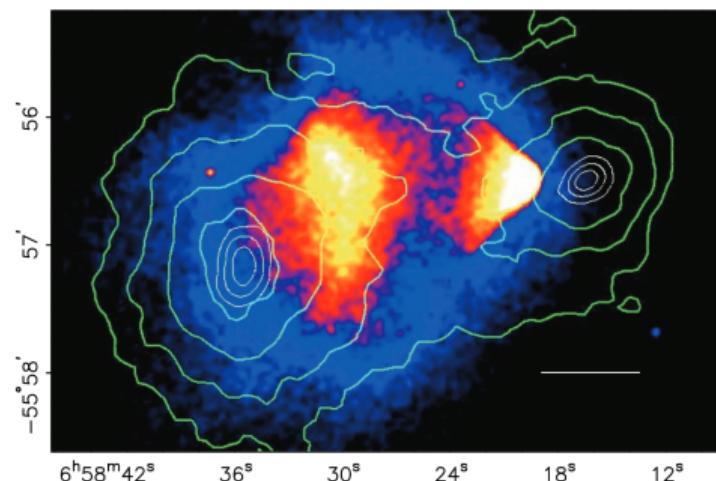
- ▷ D. Clowe et al. "A Direct Empirical Proof of the Existence of Dark Matter". *Astrophysical Journal* **648**.2 (Aug. 2006). DOI: 10.1086/508162.

# Keywords in title

Why do we **search** for...?

## Current model status

- Robust and predictive (top quark,  $W$ ,  $Z$  and Higgs bosons...)
- still not good enough, unable to explain some observations such as:
  - ▶ dark matter →
  - ▶ matter vs antimatter asymmetry
  - ▶ ...
- Go beyond with a new model!
- Consequences of this new model? **Test it!**



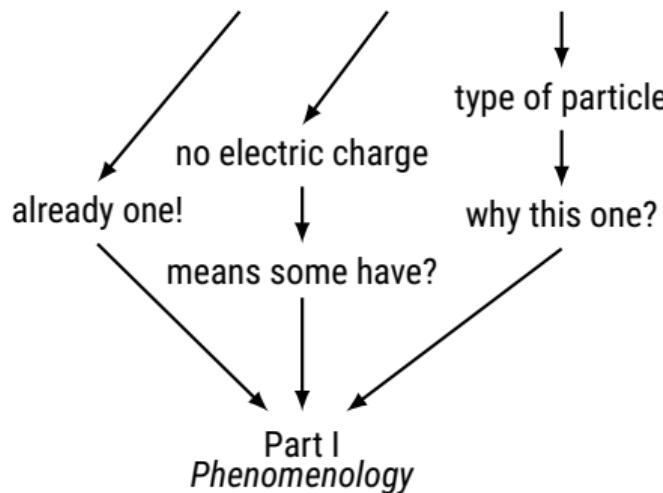
▷ D. Clowe et al. "A Direct Empirical Proof of the Existence of Dark Matter". *Astrophysical Journal* **648**.2 (Aug. 2006). DOI: 10.1086/508162.

# Keywords in title

Search for **additional neutral Higgs bosons decaying to tau lepton pair** in the **CMS experiment** at **LHC**

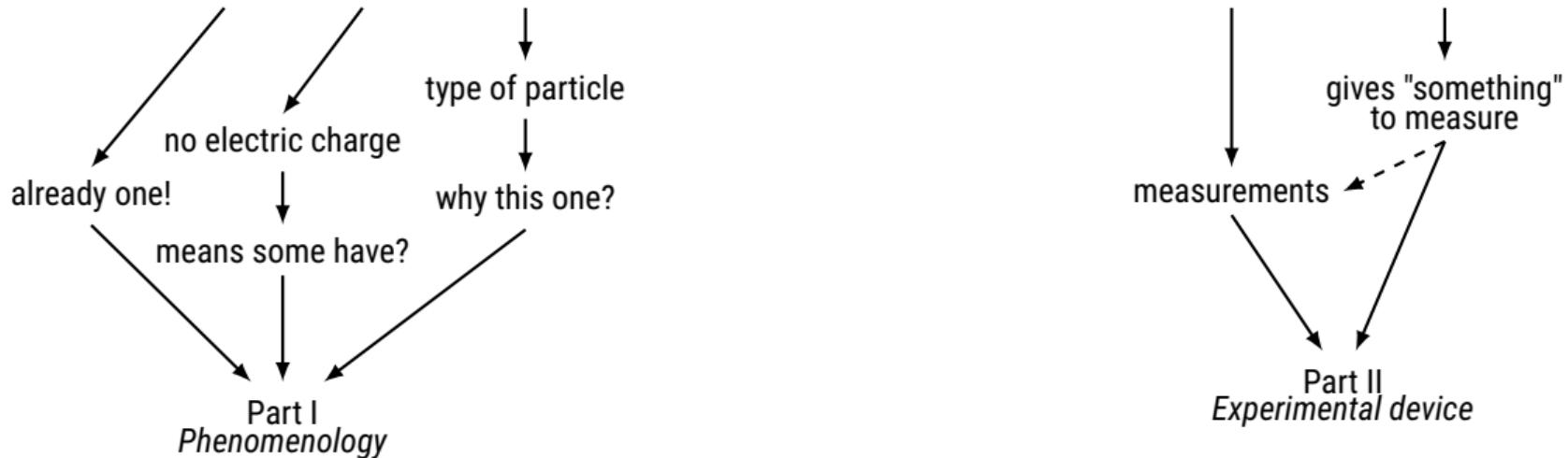
# Keywords in title

Search for **additional neutral Higgs bosons decaying to tau lepton pair** in the **CMS experiment** at **LHC**



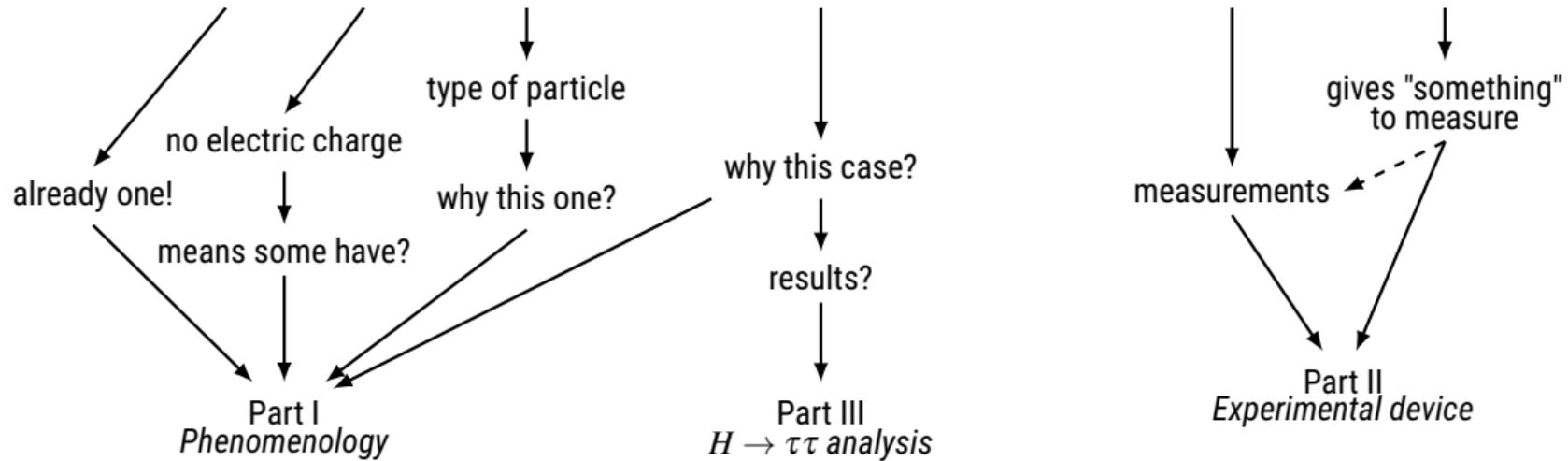
# Keywords in title

Search for **additional neutral Higgs bosons decaying to tau lepton pair** in the **CMS experiment** at **LHC**



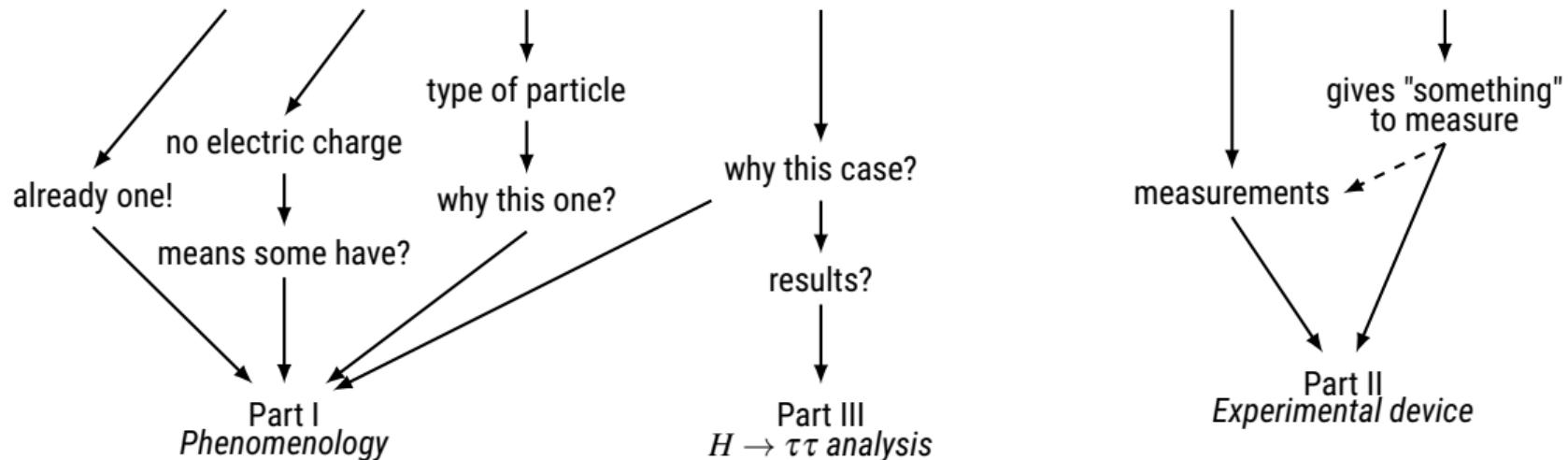
# Keywords in title

Search for **additional neutral Higgs bosons decaying to tau lepton pair** in the **CMS experiment** at **LHC**



# Keywords in title

Search for **additional neutral Higgs bosons decaying to tau lepton pair** in the **CMS experiment** at **LHC**



+ Part IV: *Machine Learning use in the  $H \rightarrow \tau\tau$  analysis*

## 1 Phenomenology

## 2 Experimental device

- CERN LHC
- The CMS detector
- Jet energy calibration

## 3 $H \rightarrow \tau\tau$ analysis

## 4 Machine Learning

- Event topology
- NN inputs
- NN structure
- NN training

# Principle

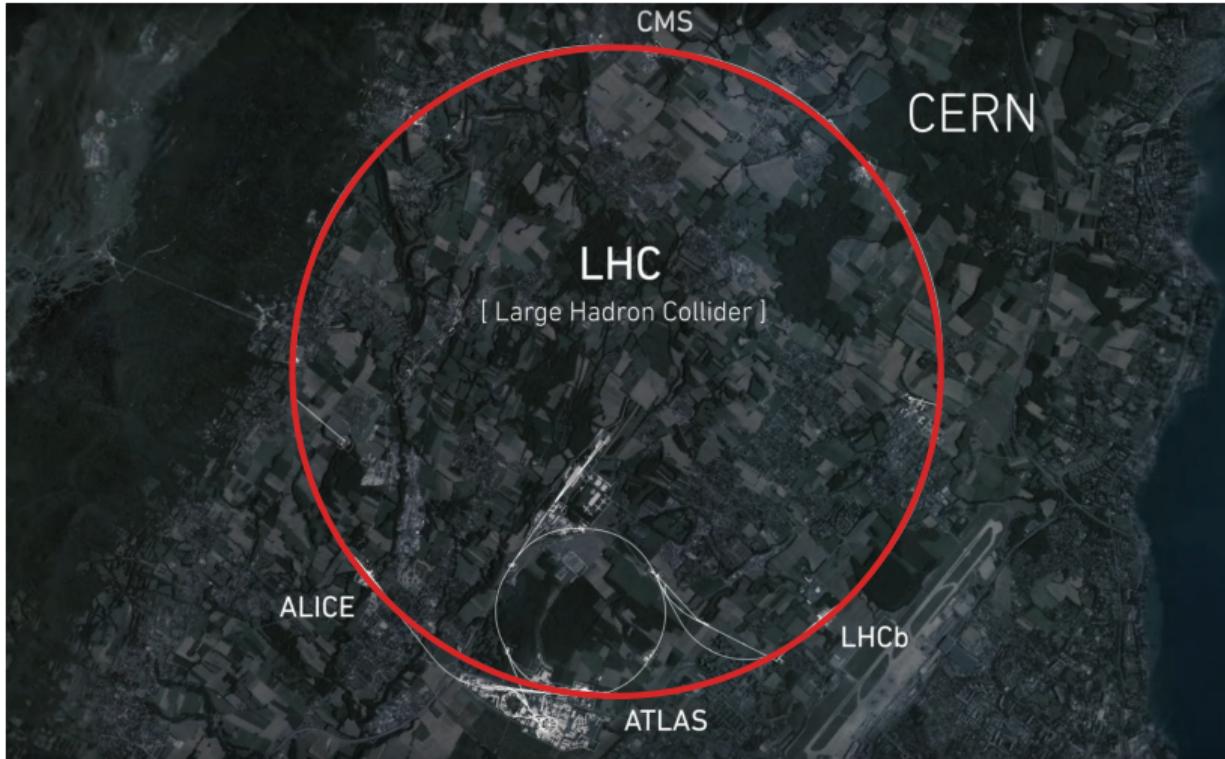
$$E = mc^2$$

mass (new particles) from the collision energy

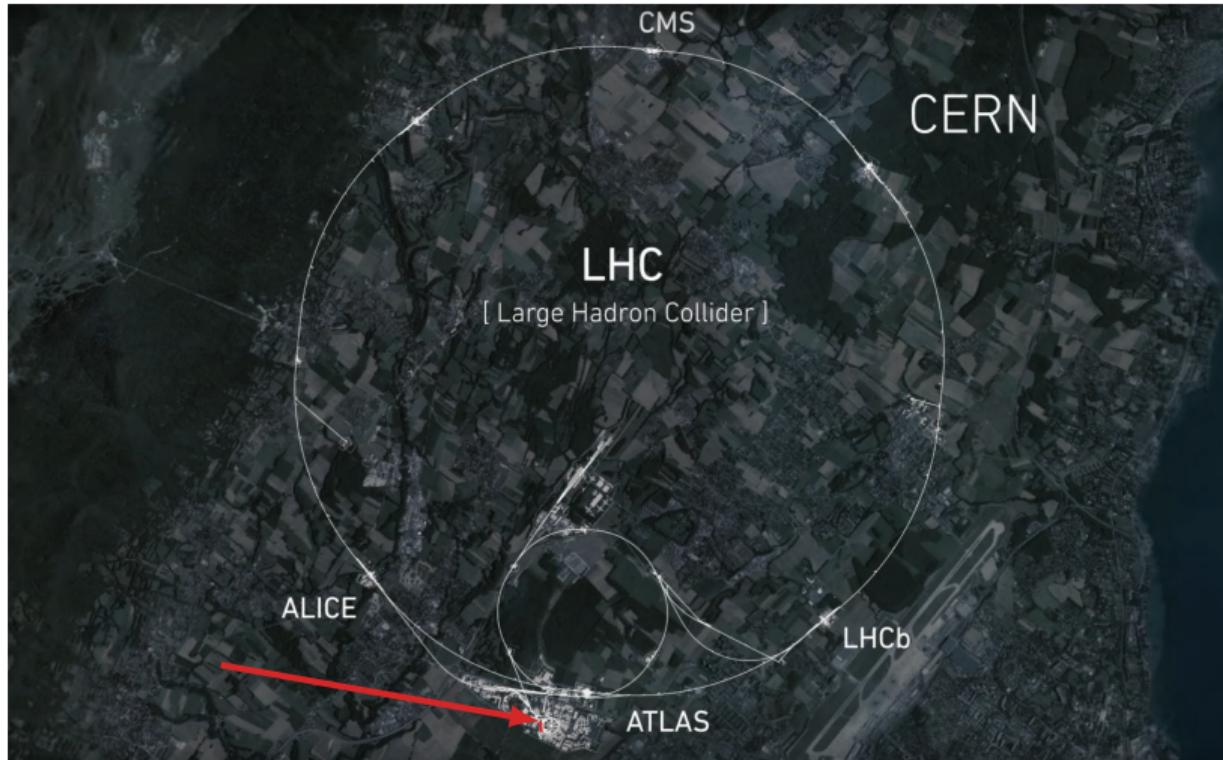
# CERN & LHC



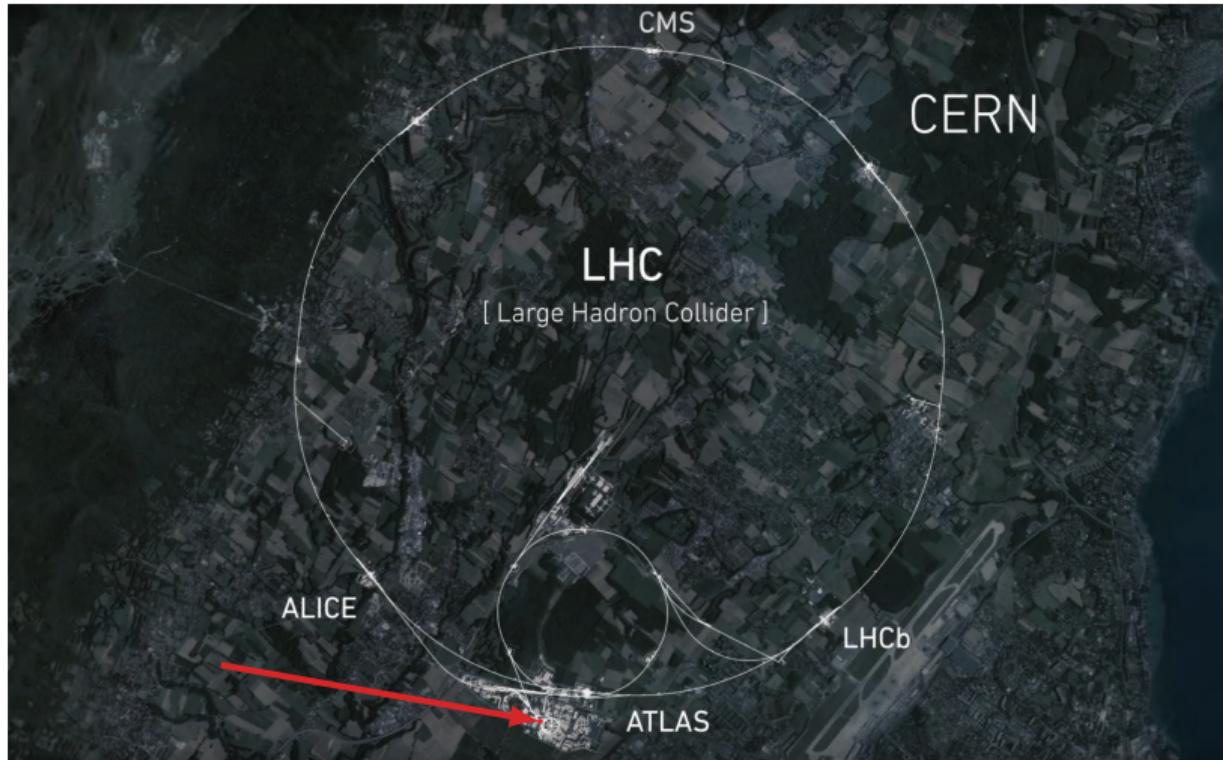
## LHC



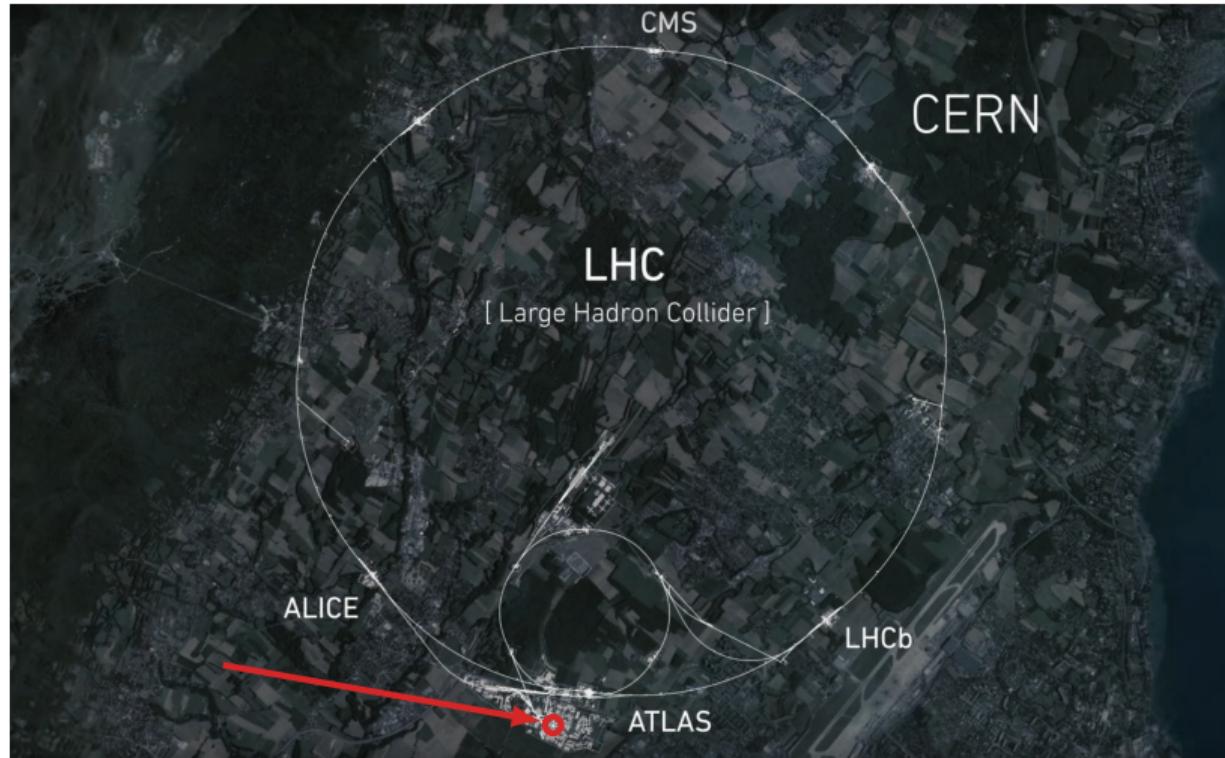
# LINAC2 (50MeV)



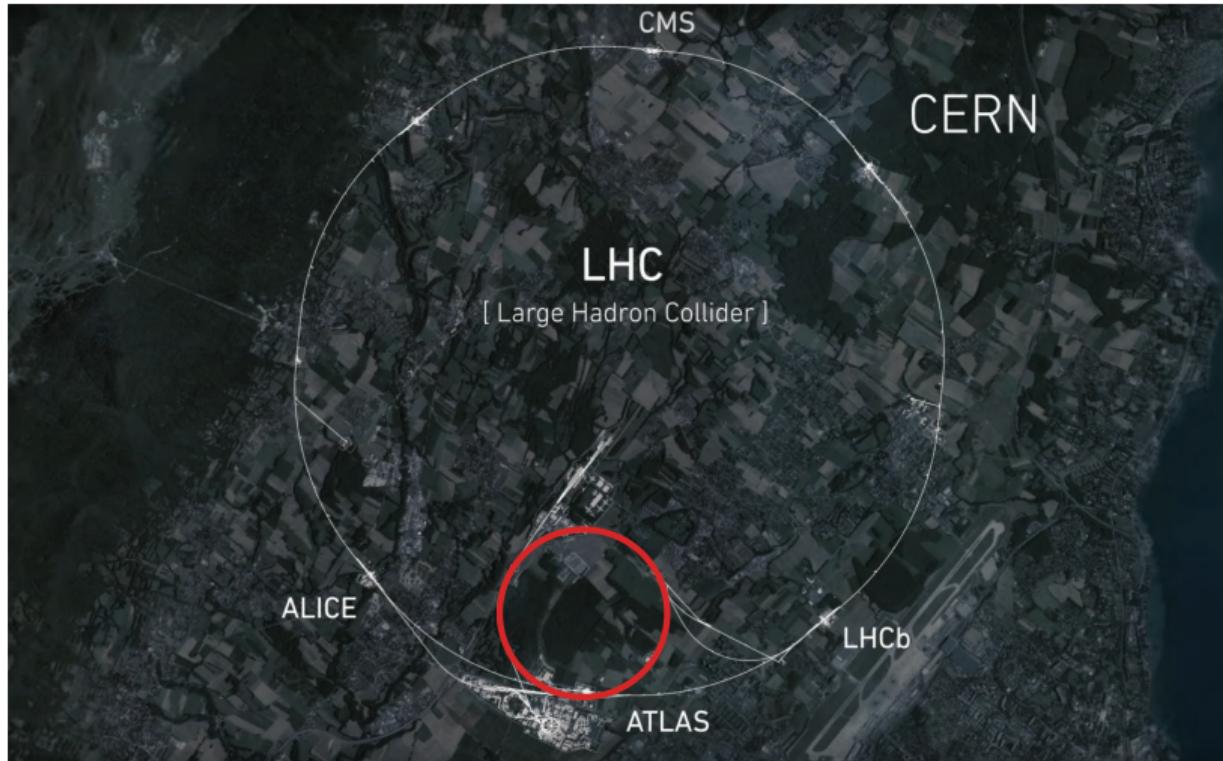
# Booster (1972, 157 m, 1.4 GeV)



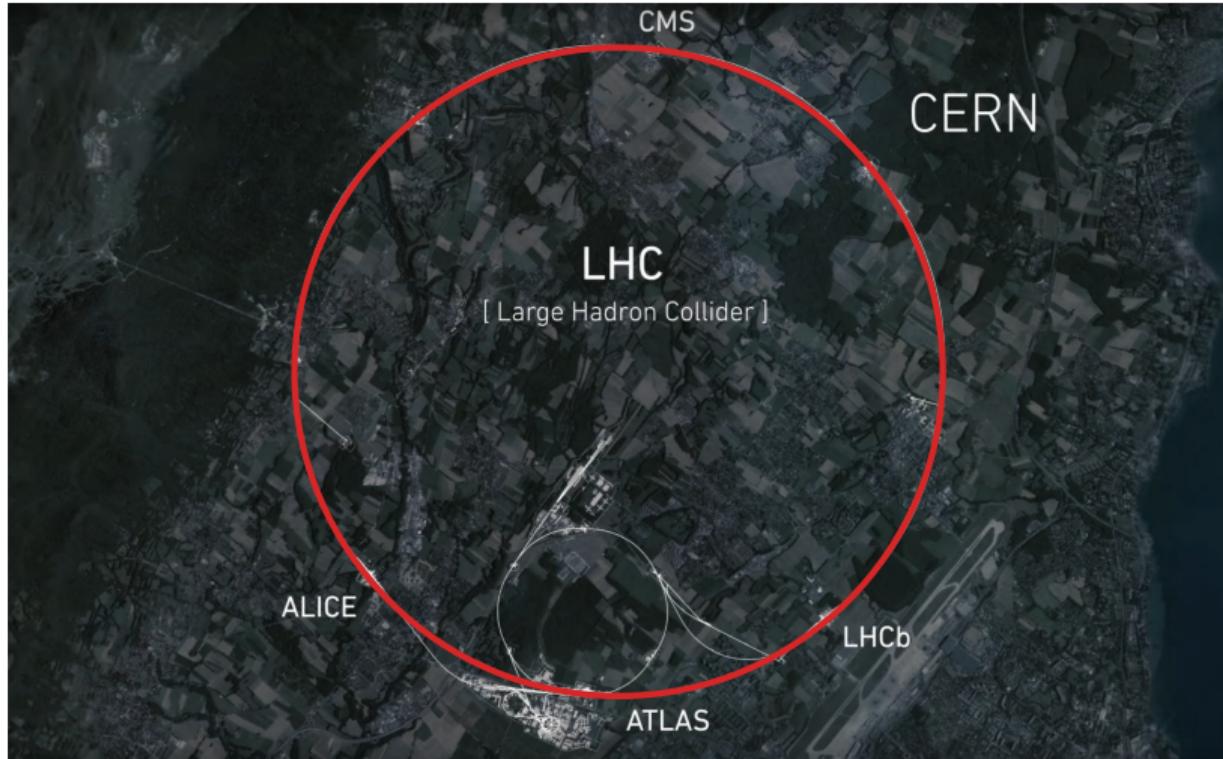
# PS (1959, 628 m, 25 GeV)

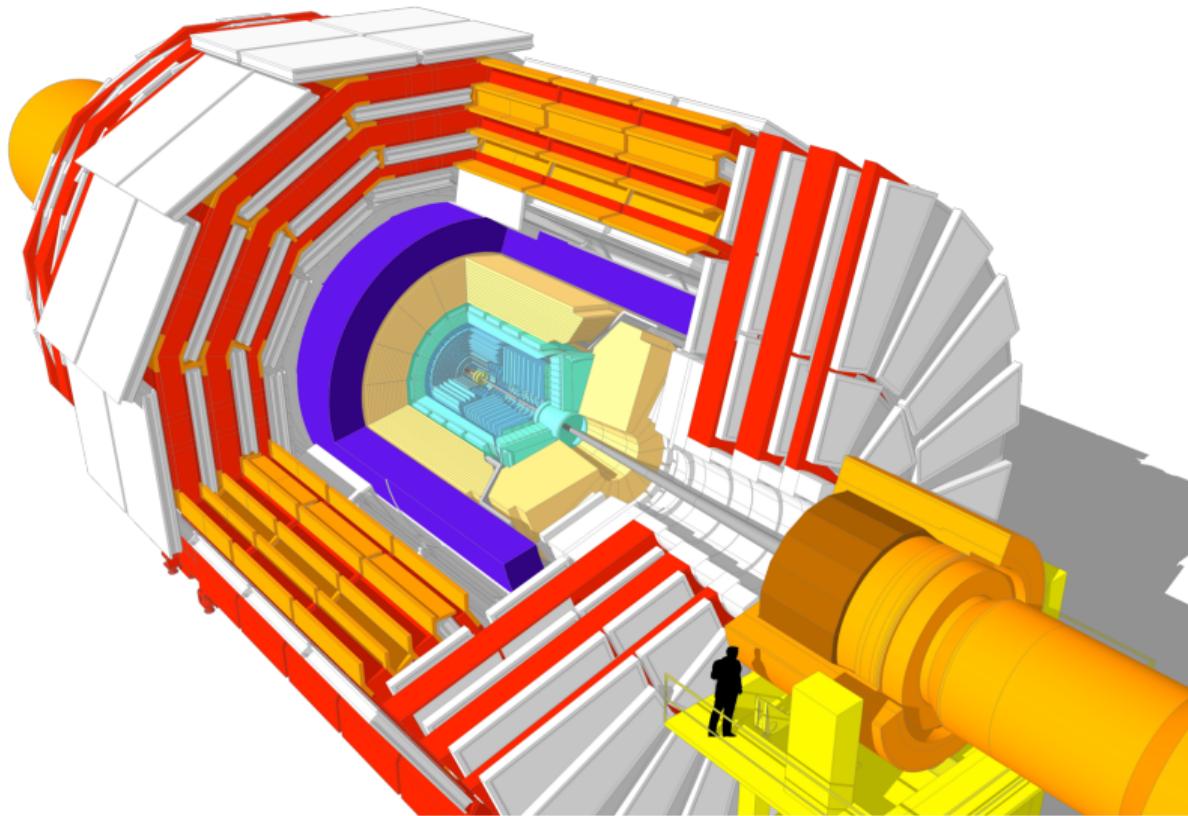


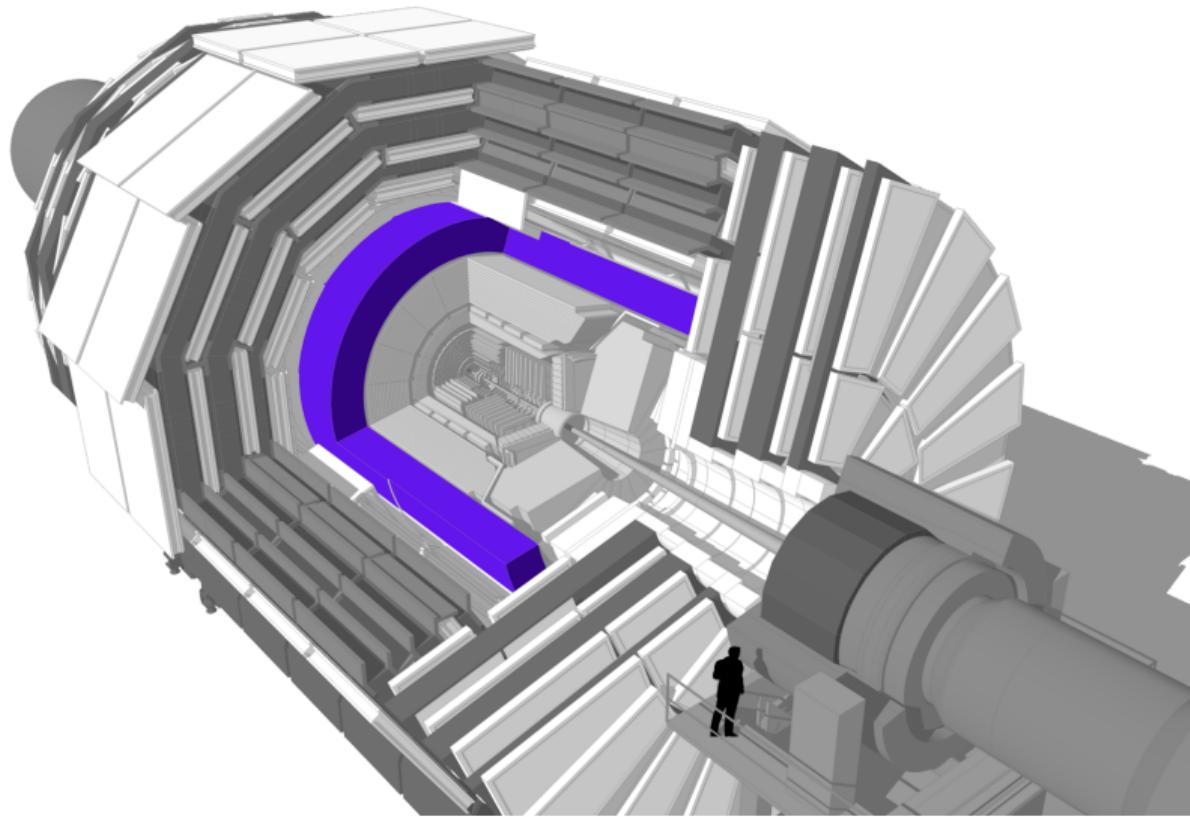
# SPS (1976, 7 km, 450 GeV)

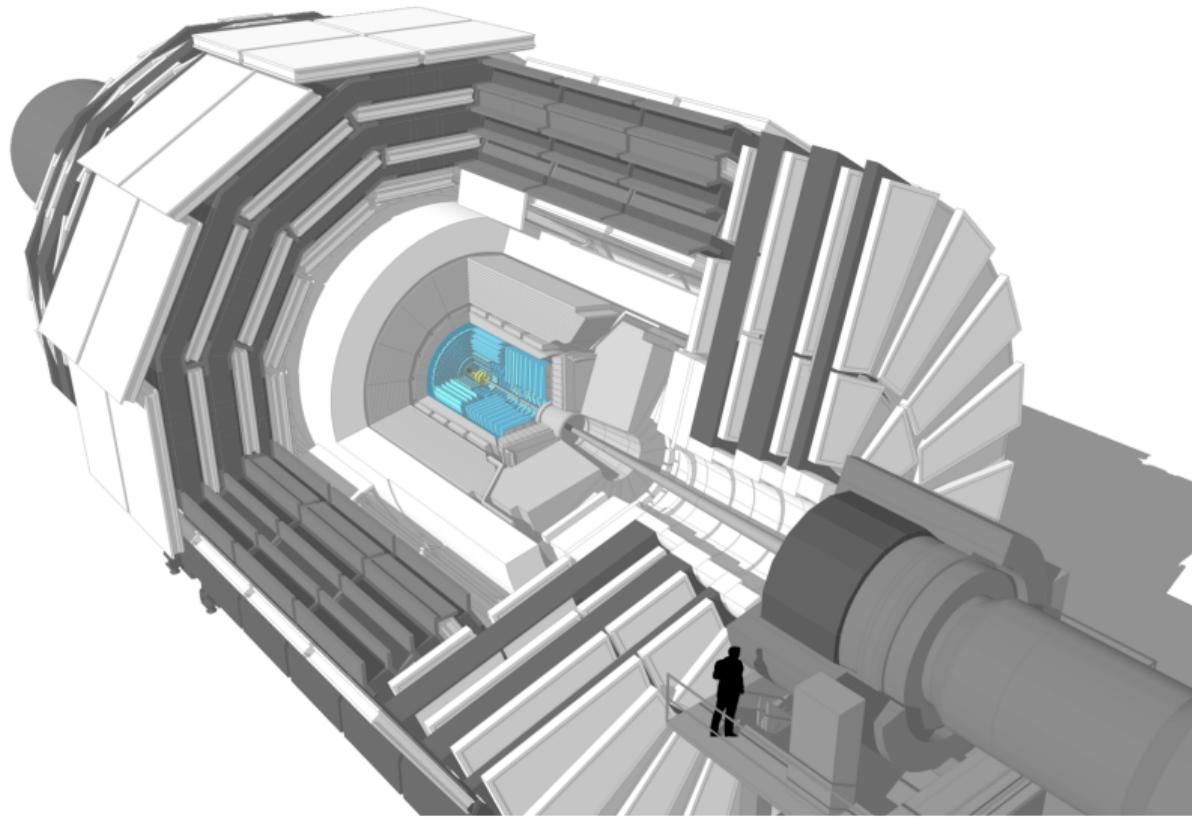


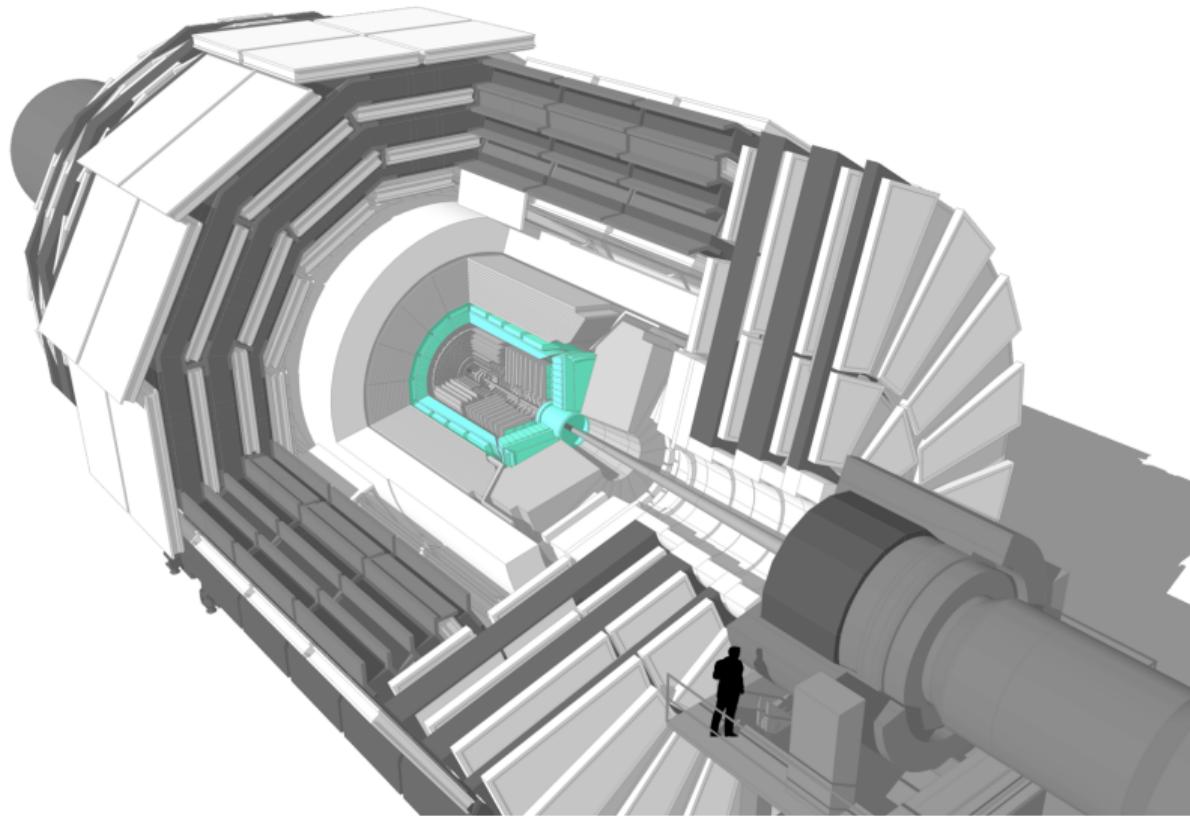
# LHC (2008, 27 km, $2 \times 7 \text{ TeV}$ )

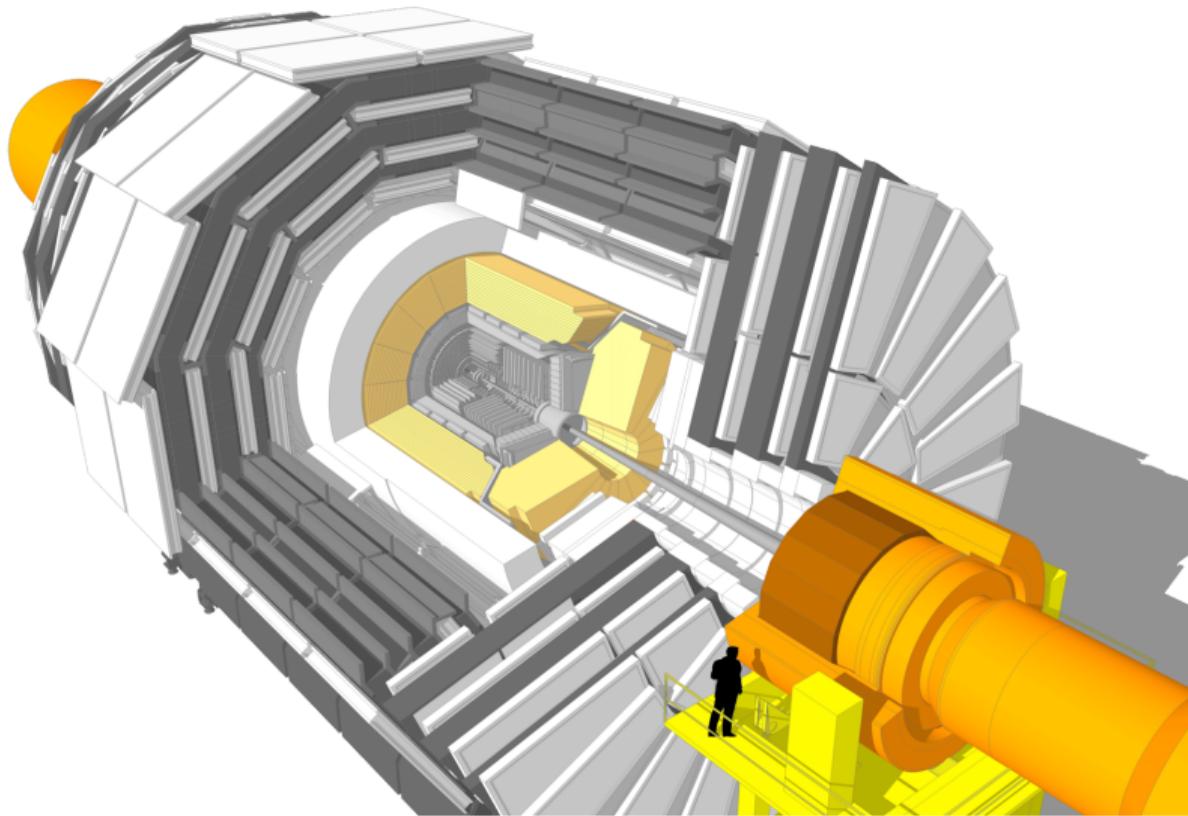


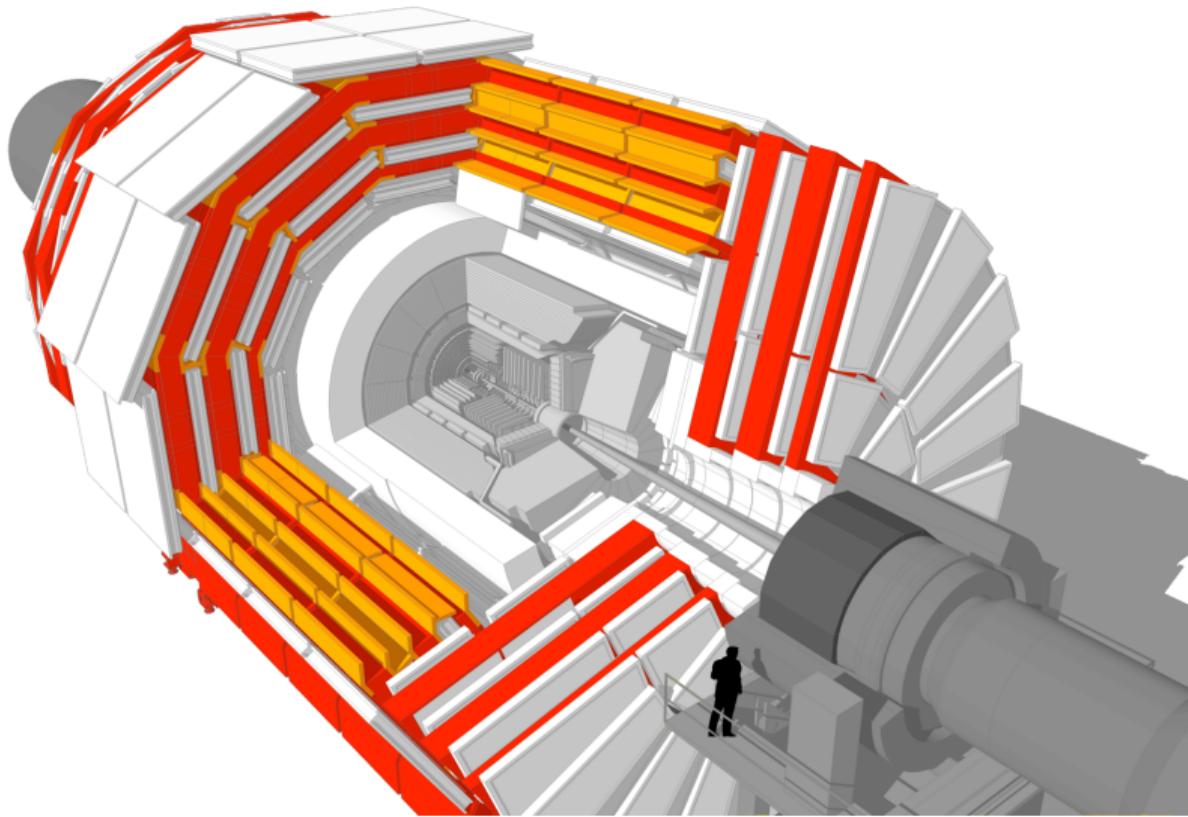


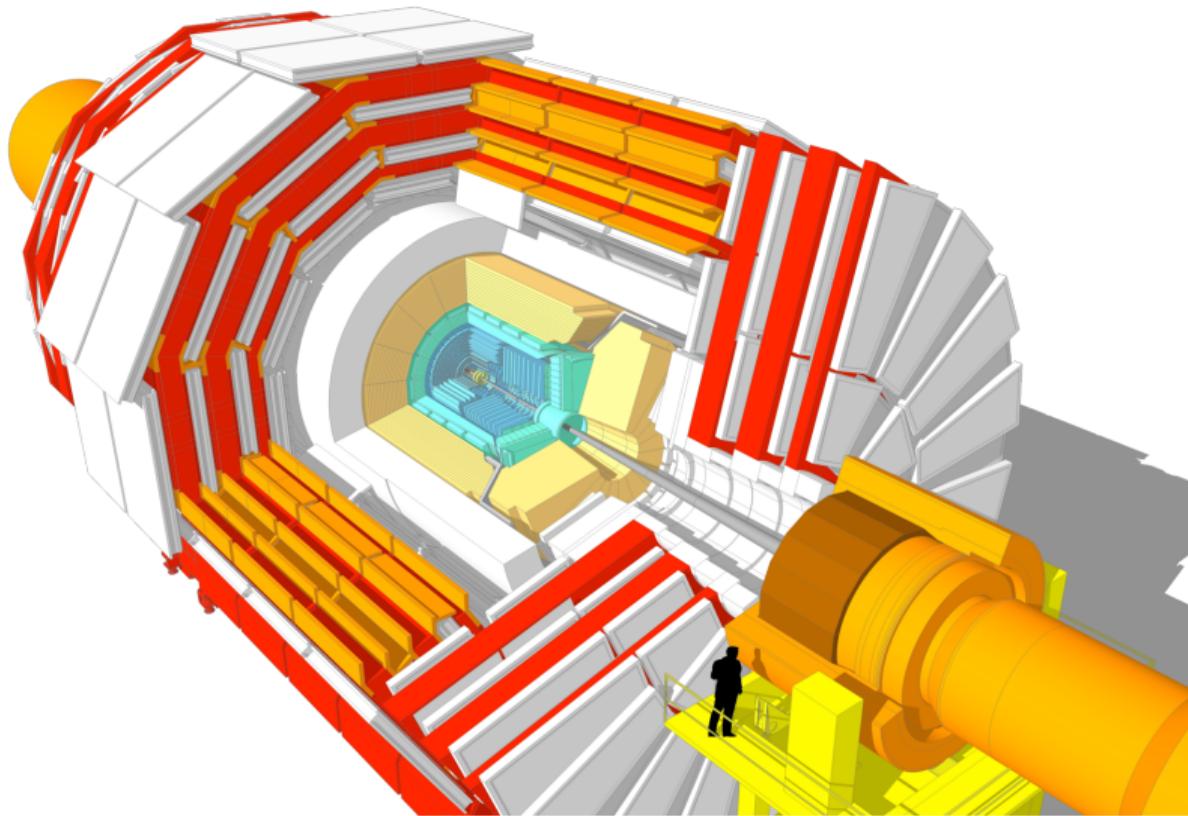






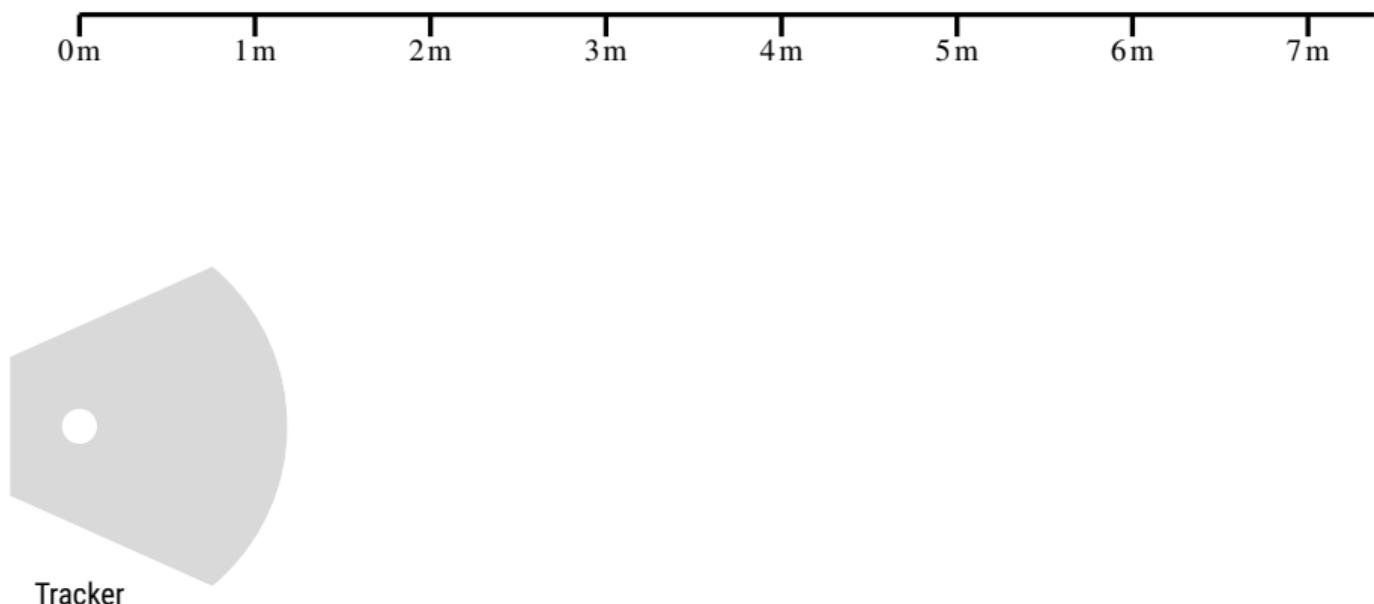




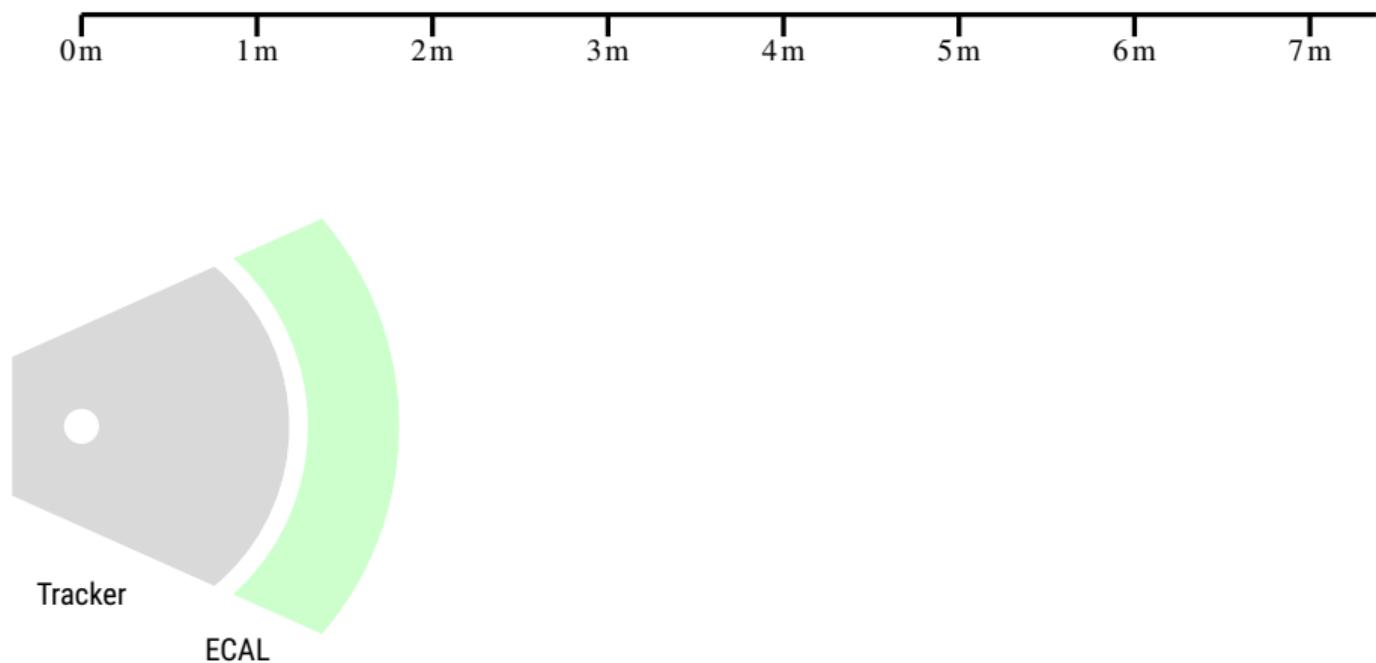


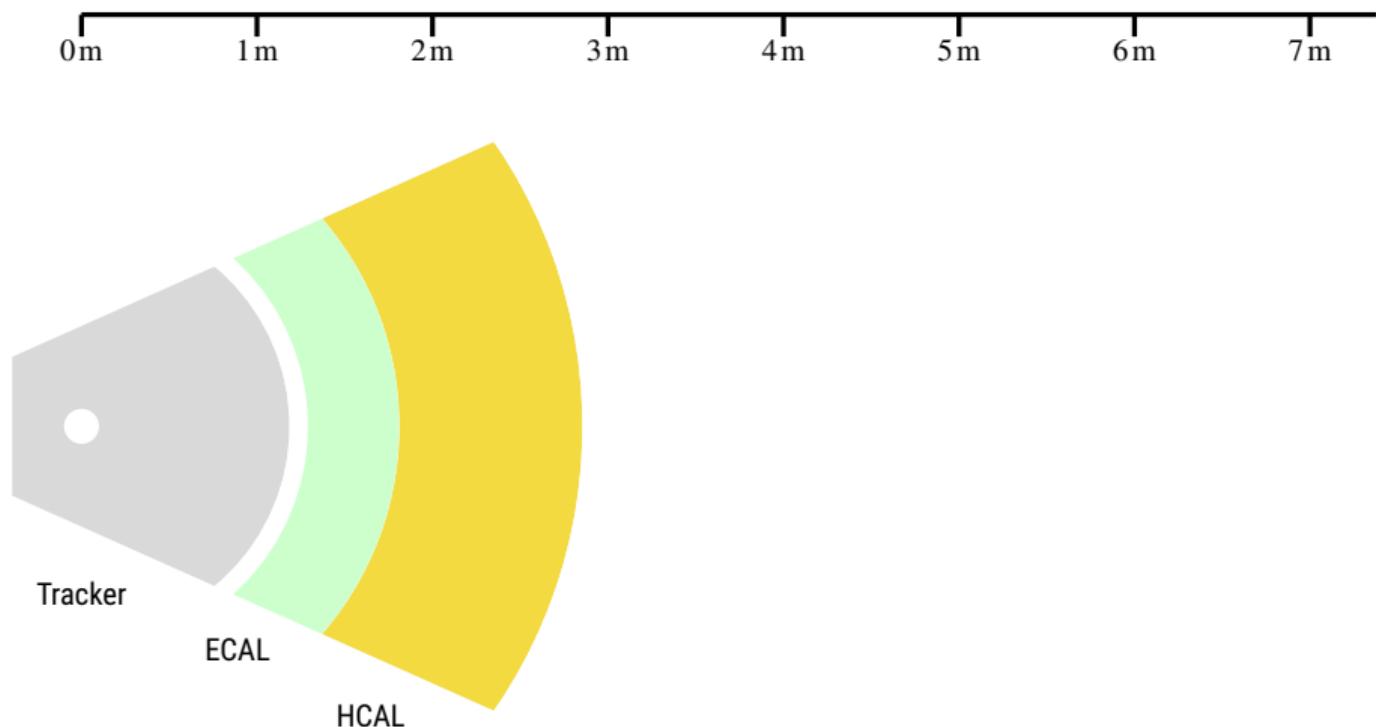


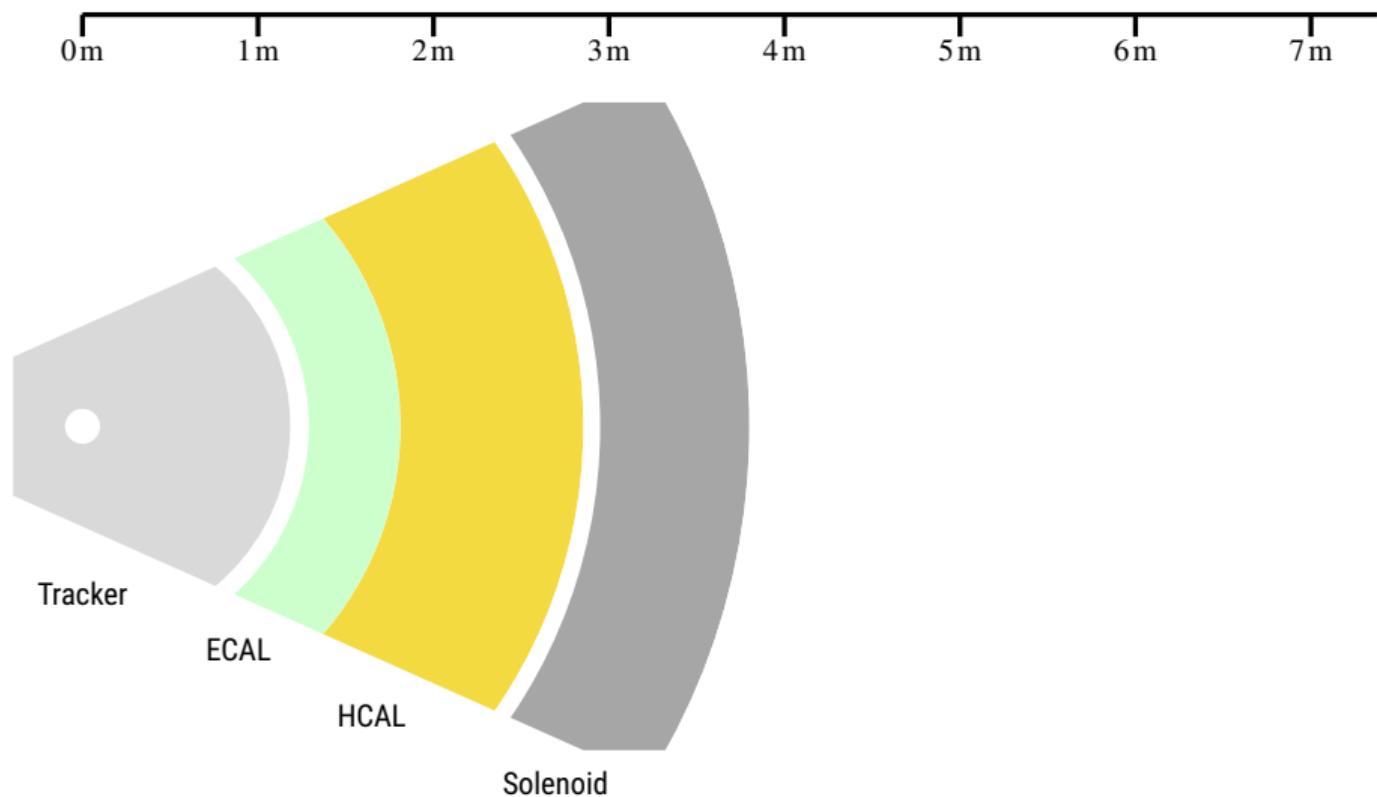


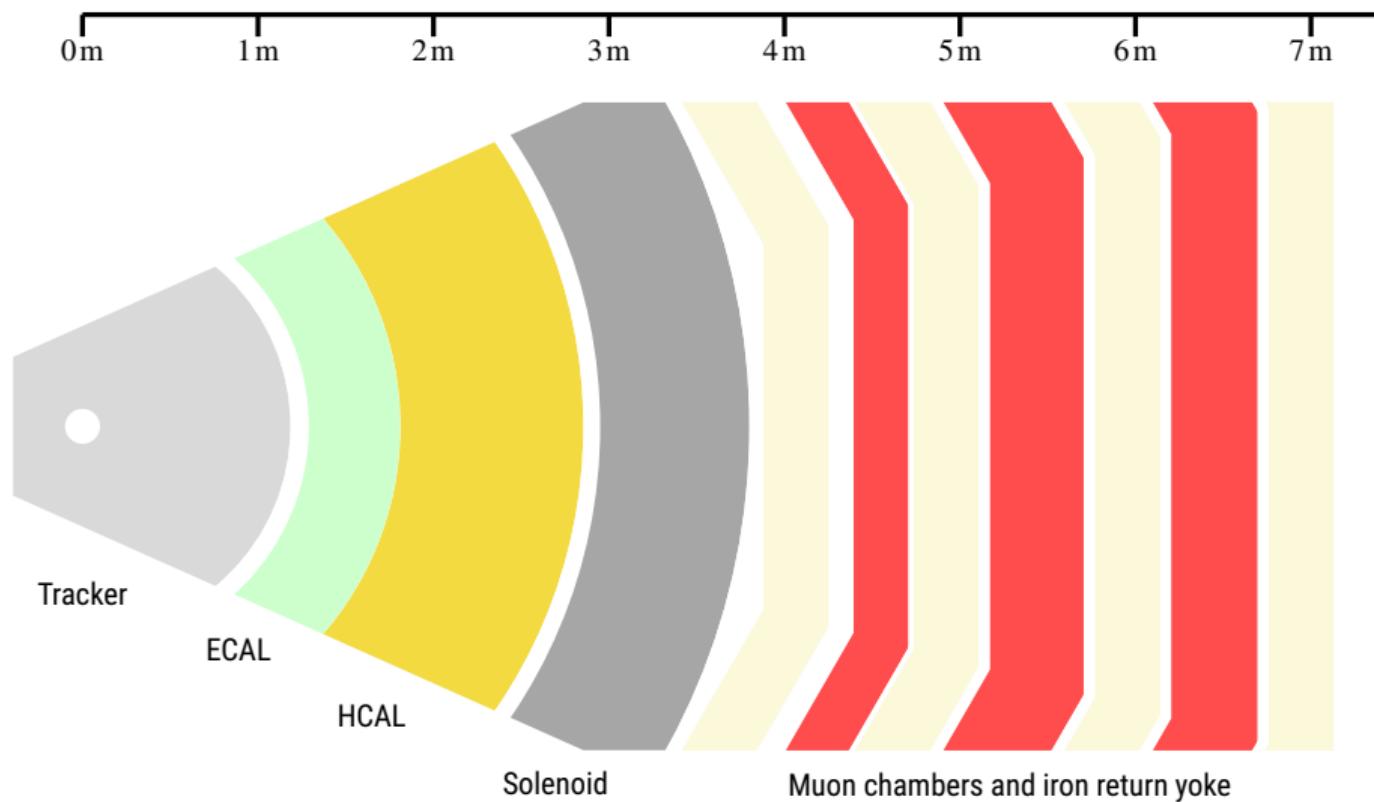


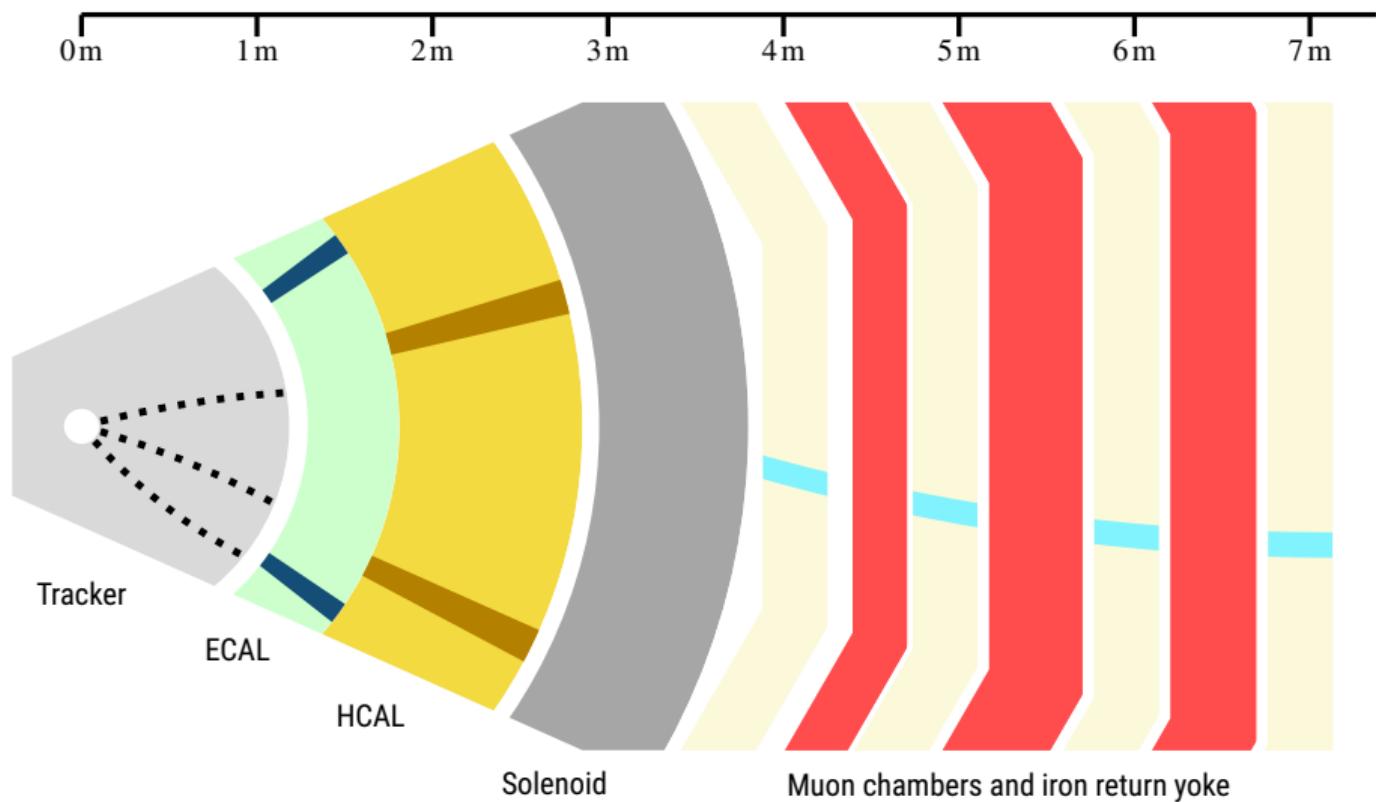
Tracker

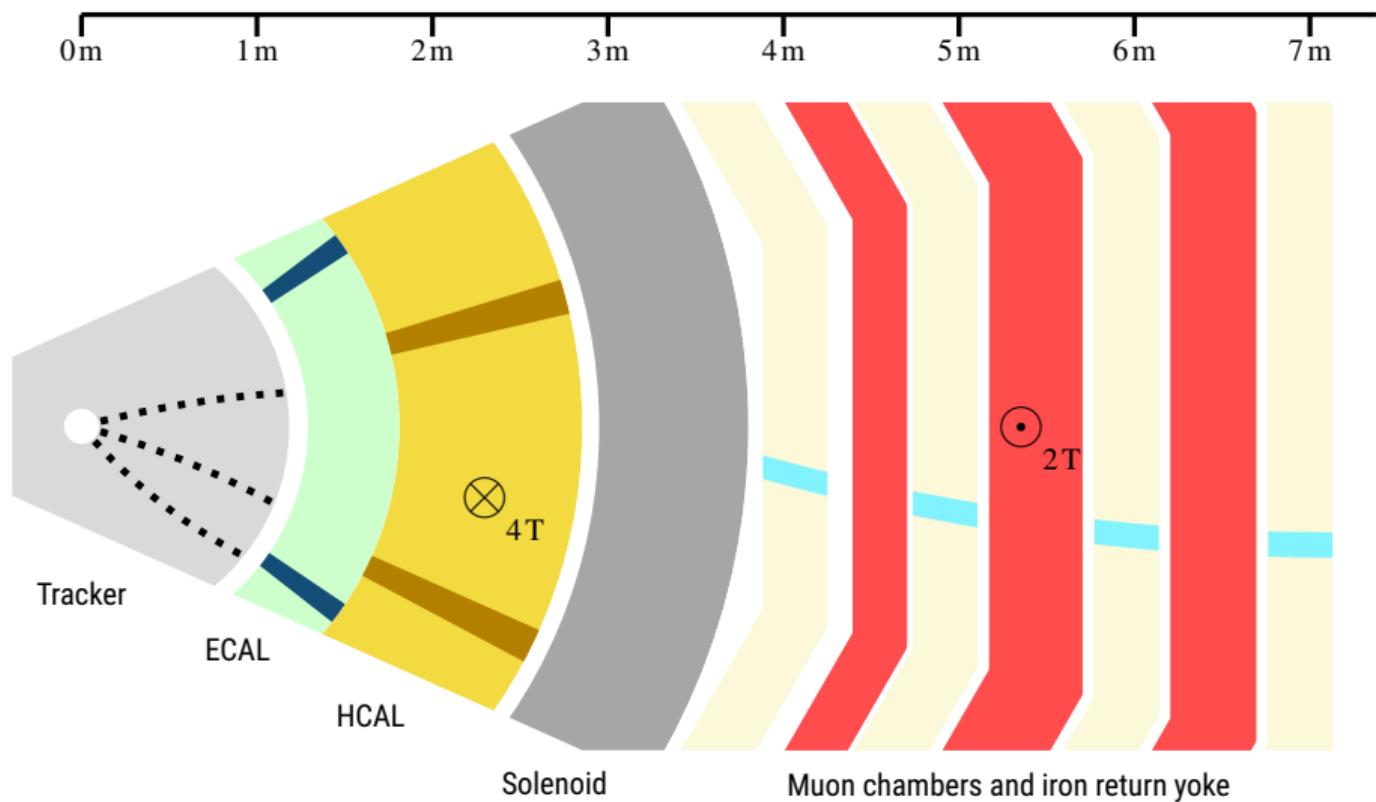


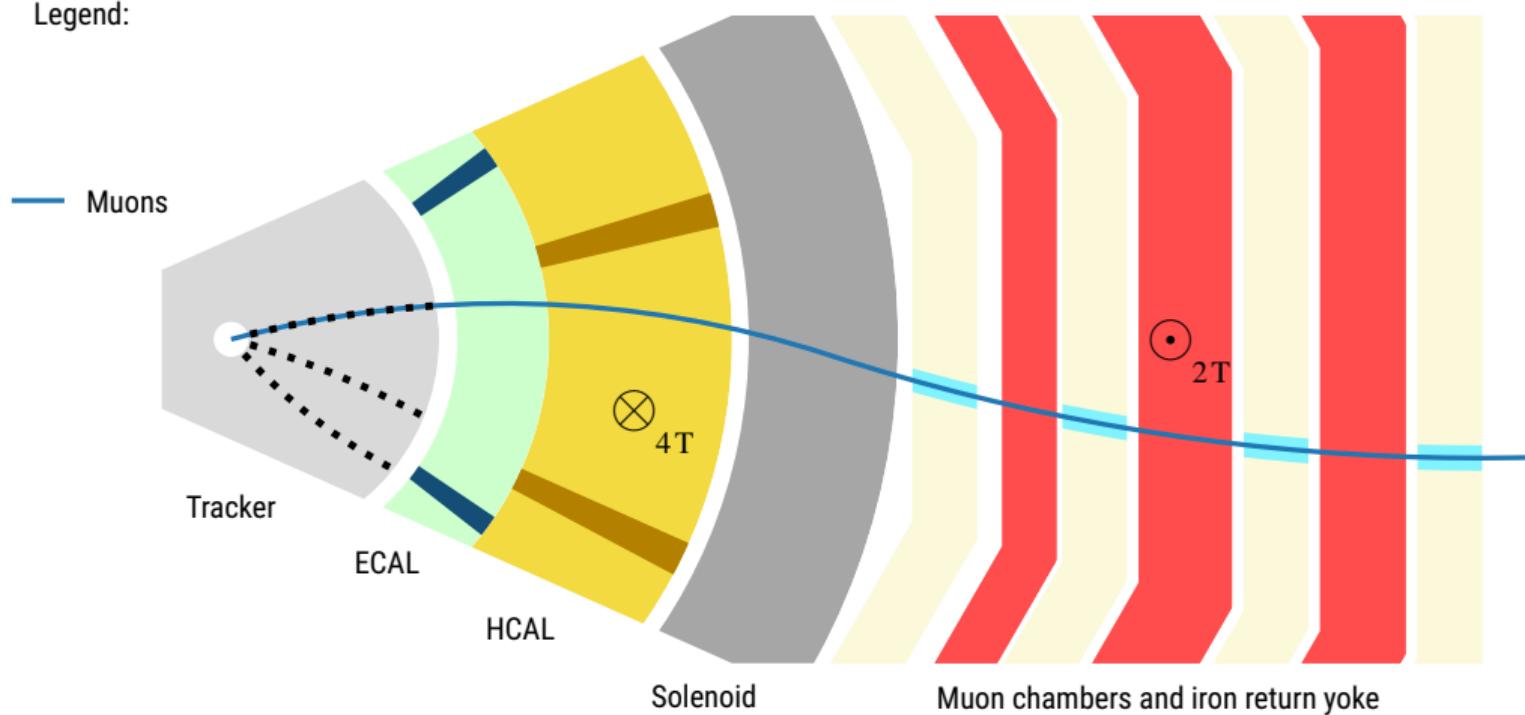


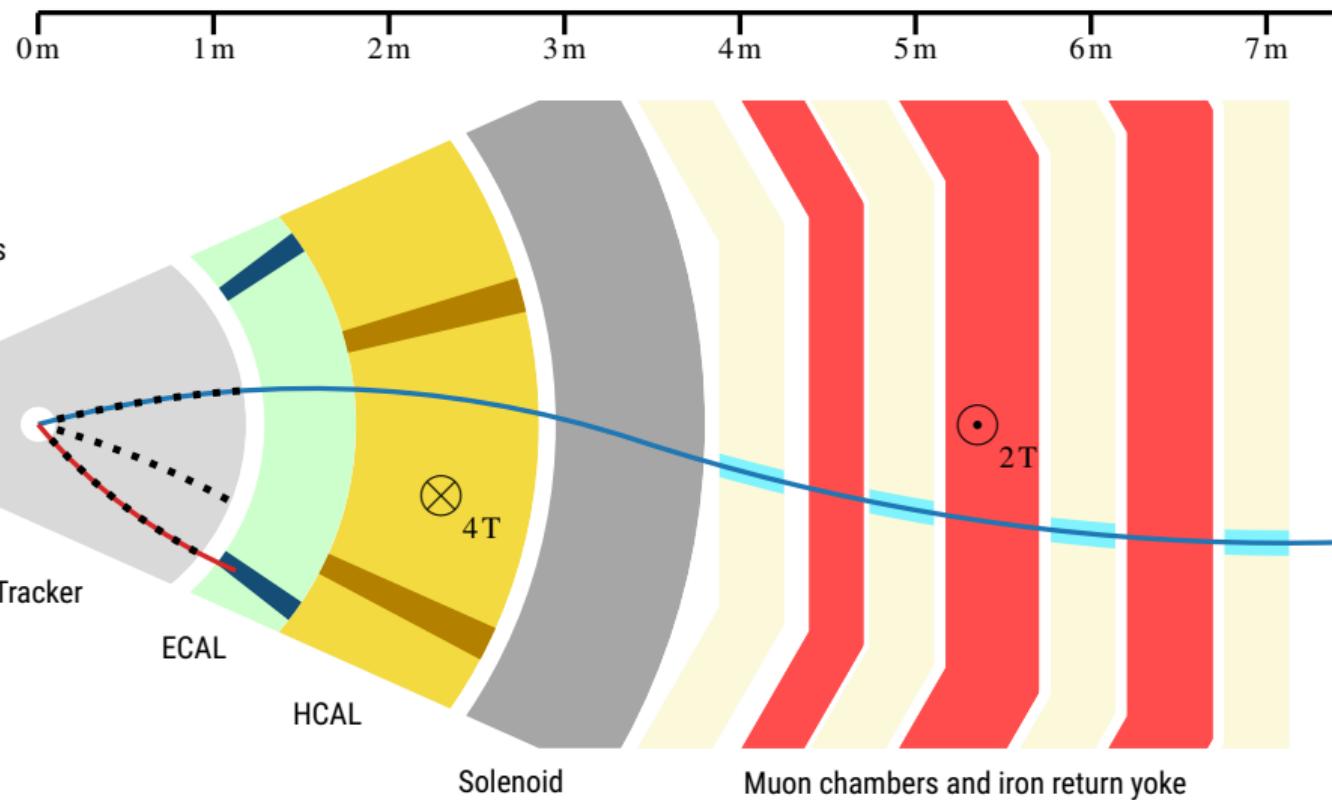


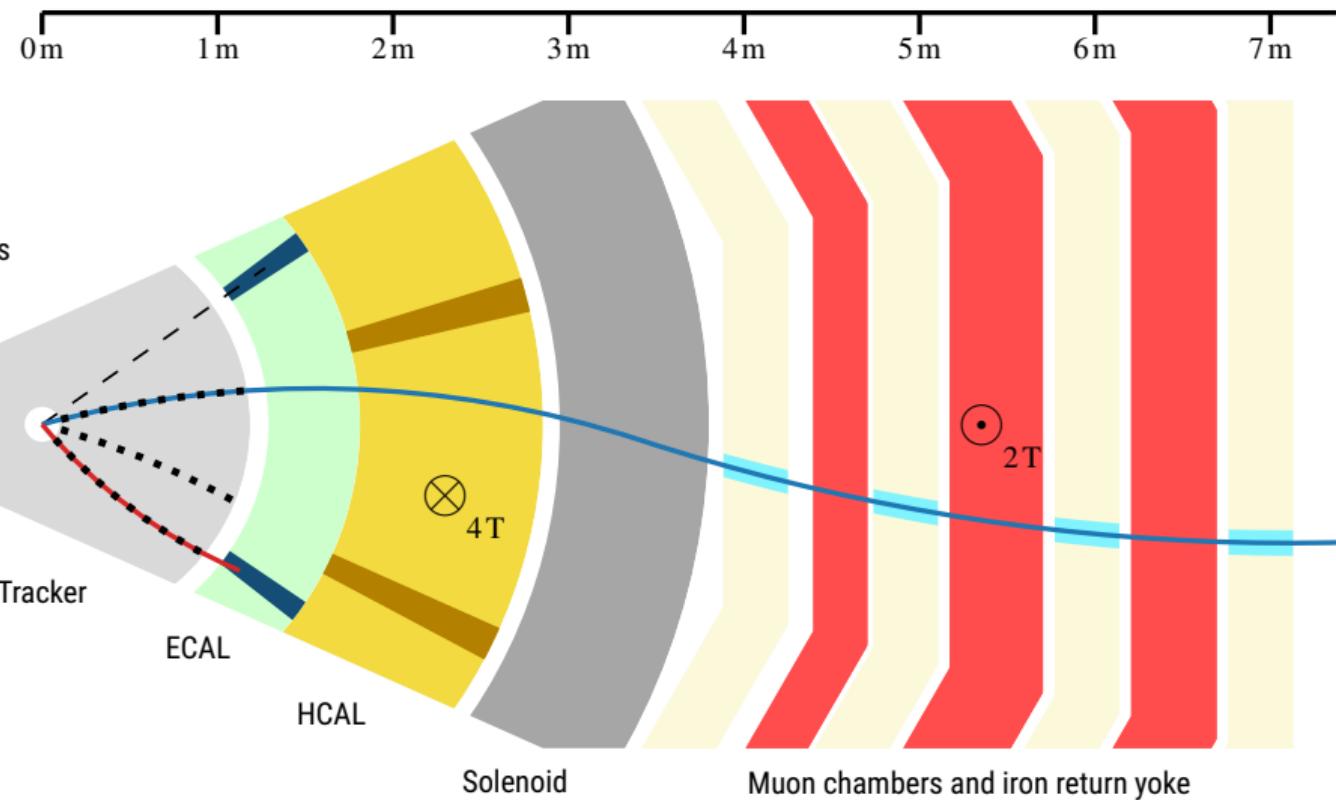


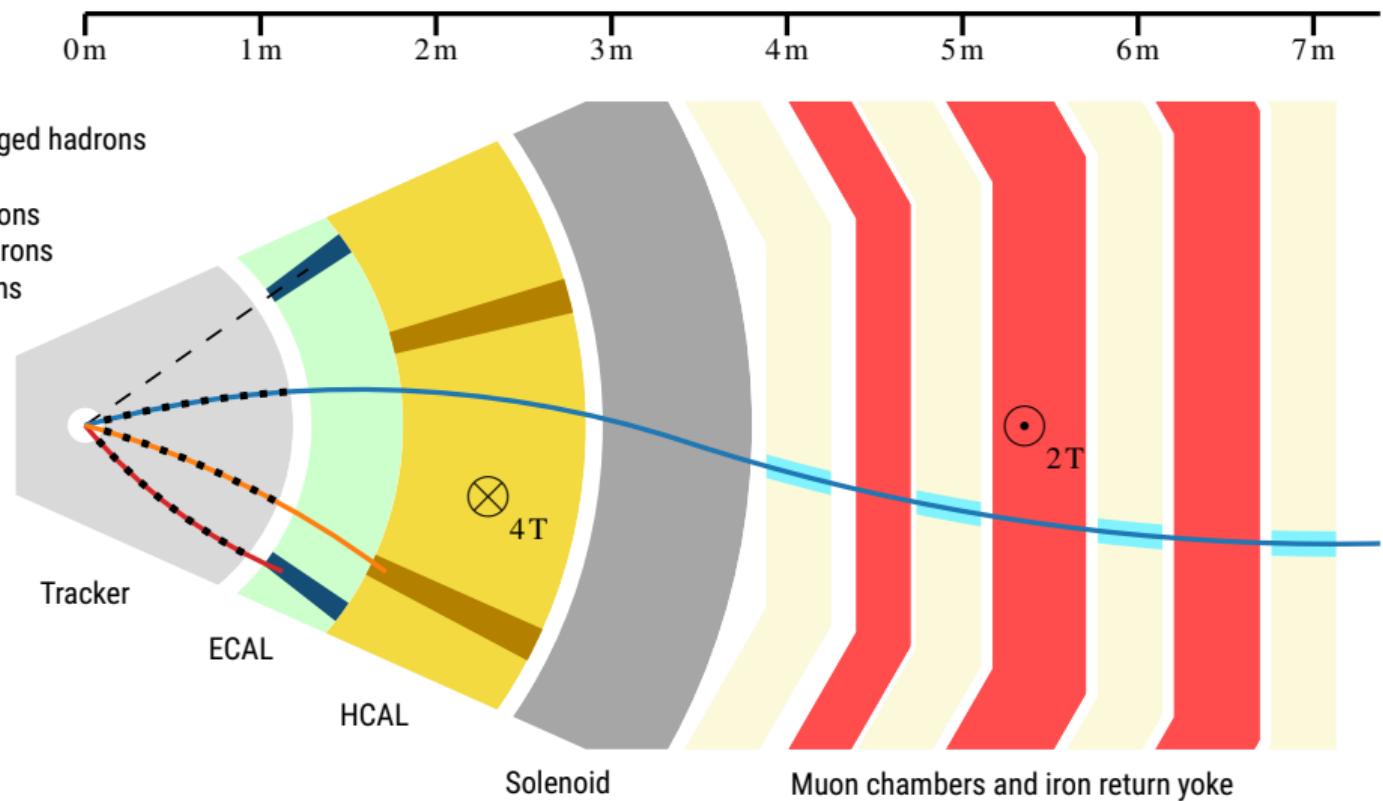


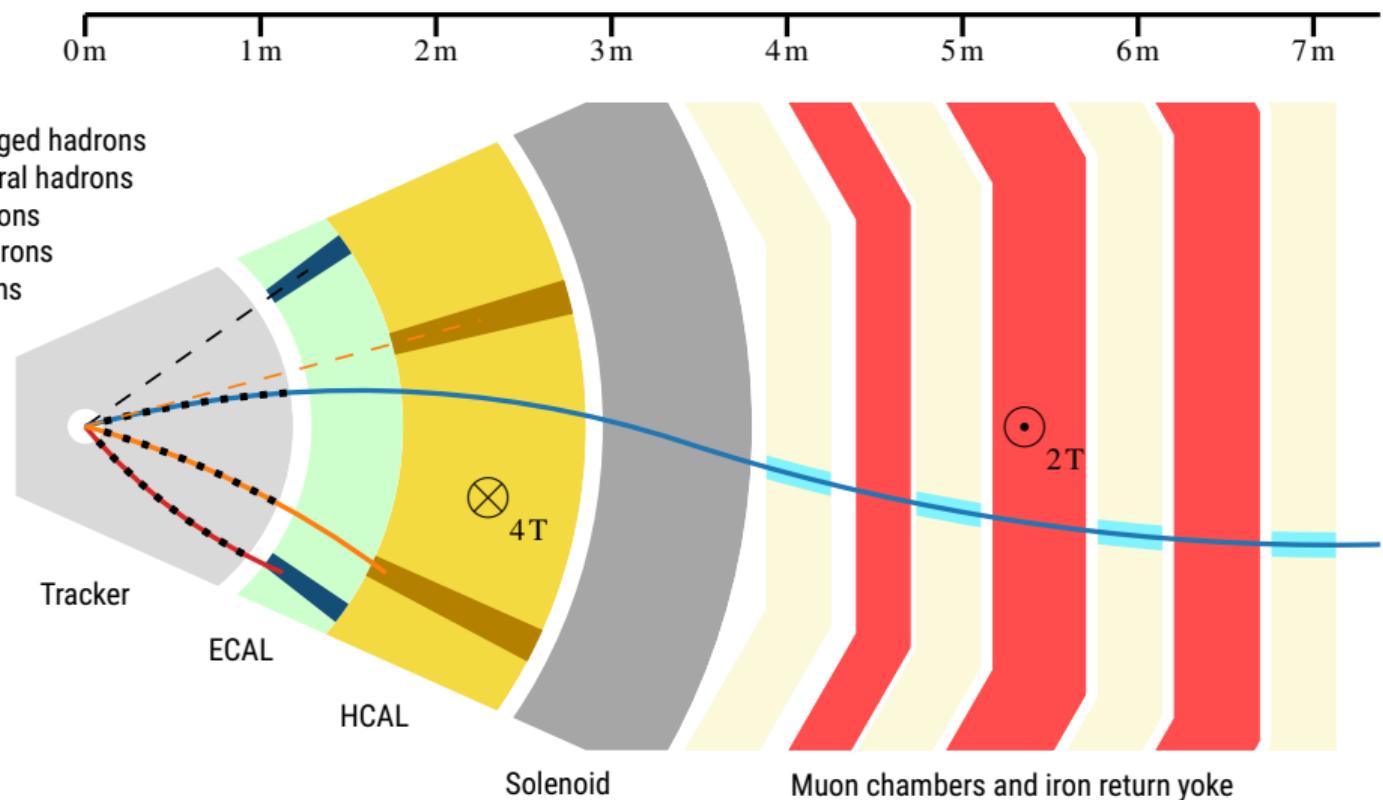


**Legend:**









## ► Niveaux de connaissance

particule (ptcl)  
reconstruit (reco)  
corrigé (corr)

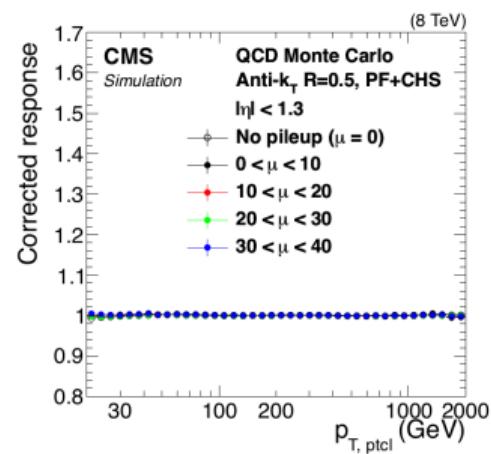
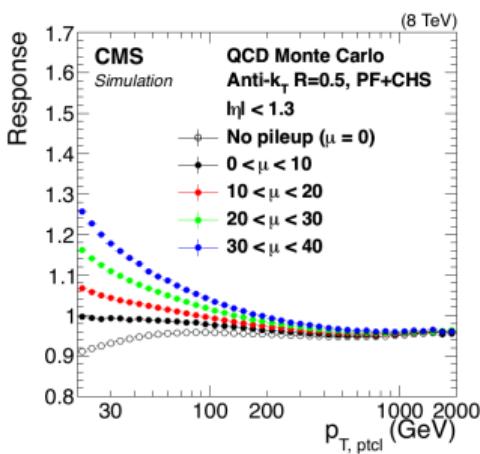
## ▶ Niveaux de connaissance

particule	(ptcl)
reconstruit	(reco)
corrigé	(corr)

## ▶ Réponse d'un jet

$$R = \frac{p_T}{p_{T\text{ptcl}}}$$

## Jets Reconstructs



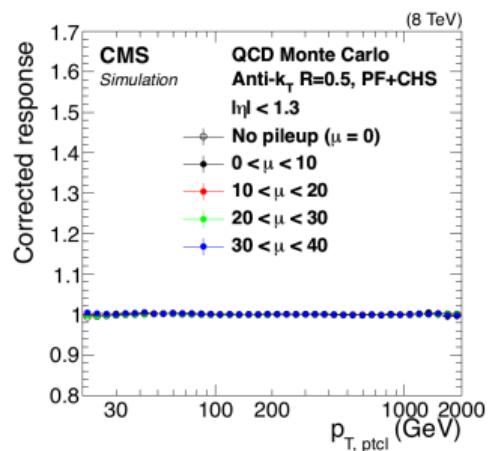
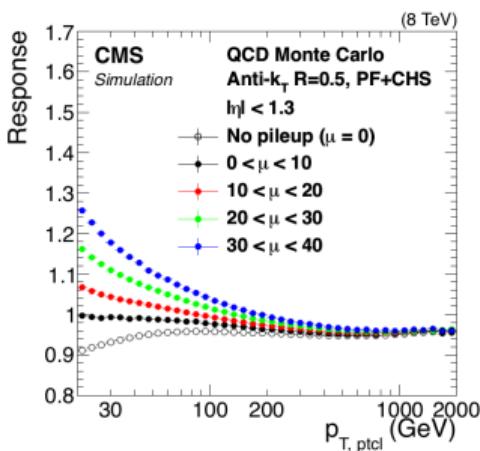
Appliqué aux données

Jets  
Reconstitués



Jets  
Calibrés

Appliqué aux simulations



## Appliqué aux données

Empilement

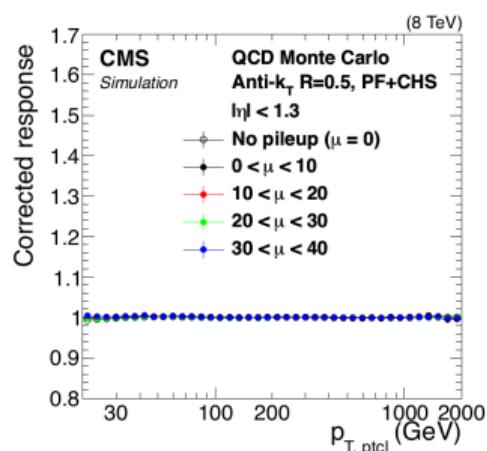
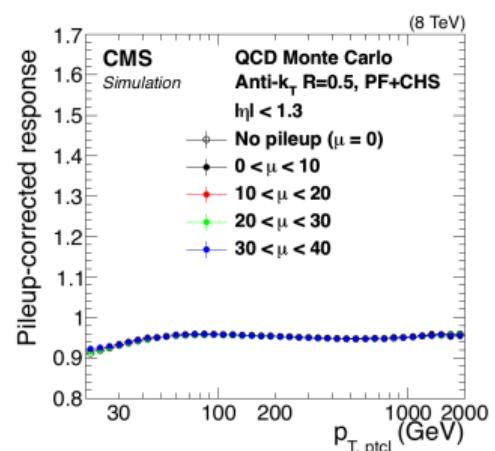
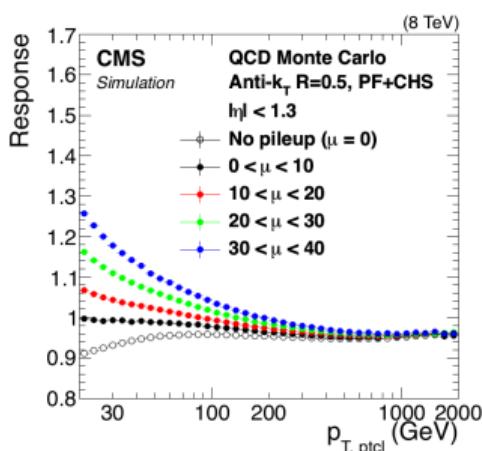
MC+RC

MC

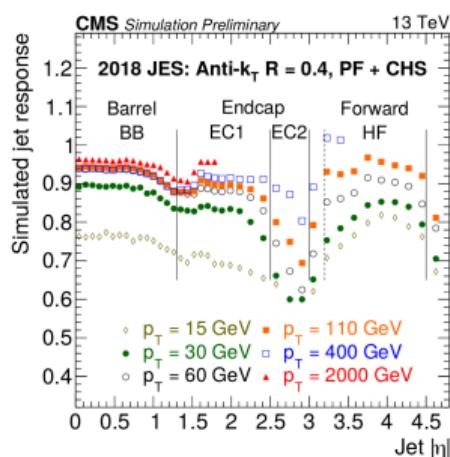
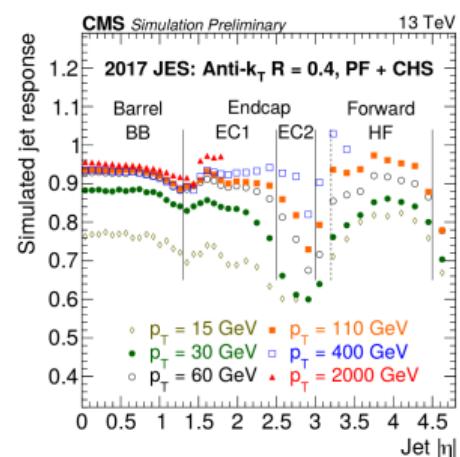
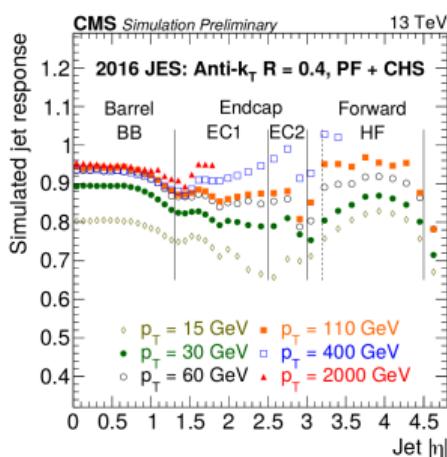
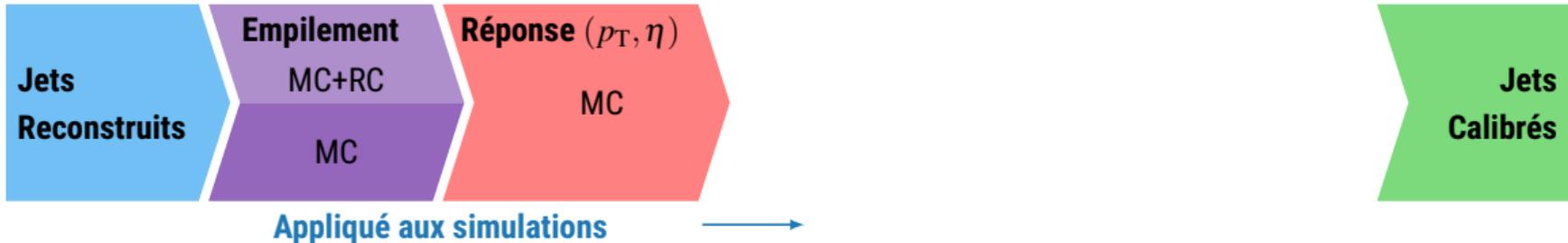
Jets

Reconstitués

## Appliqué aux simulations



## Appliqué aux données



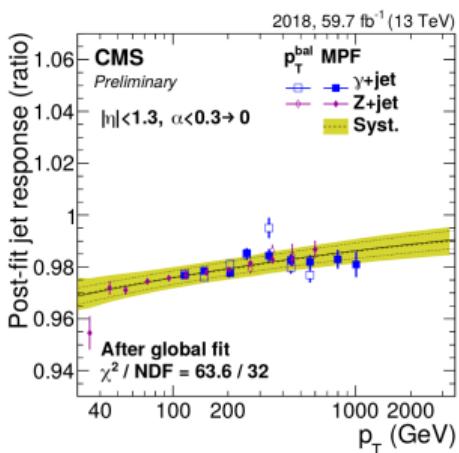
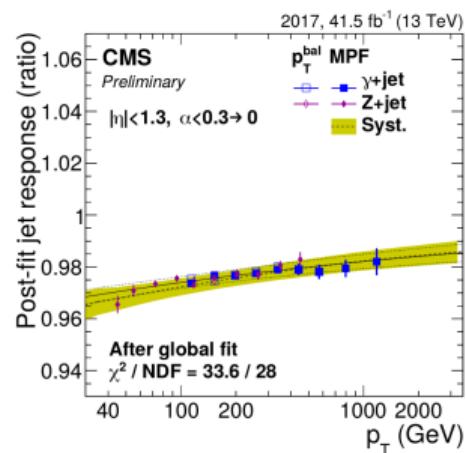
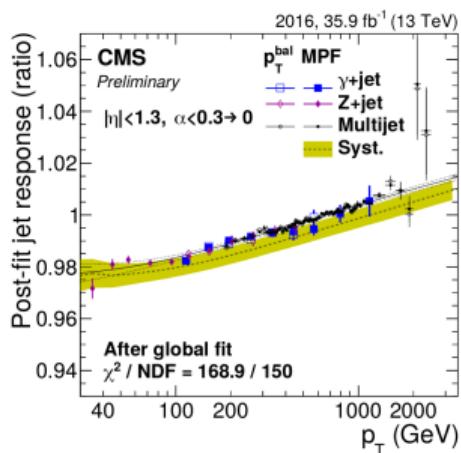
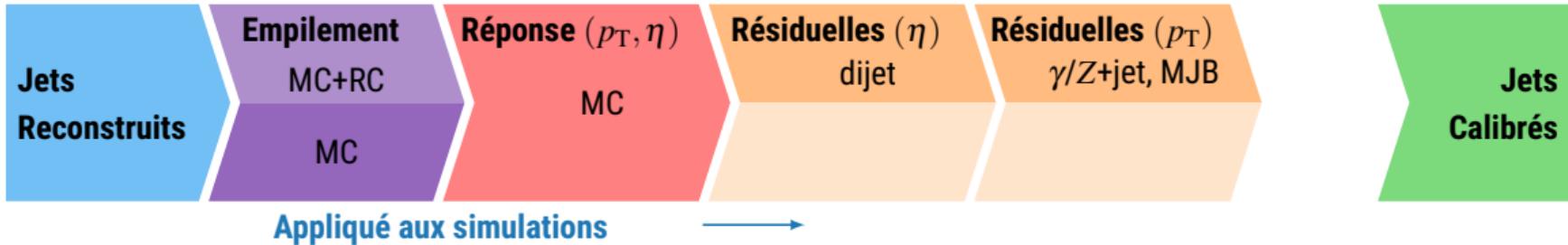
Appliqué aux données

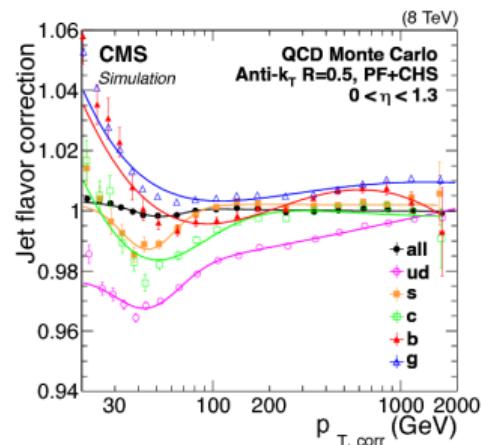
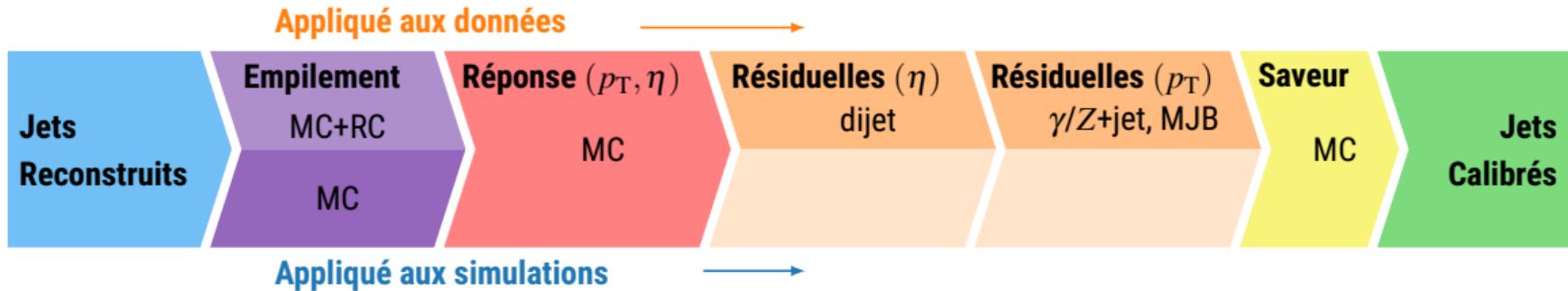


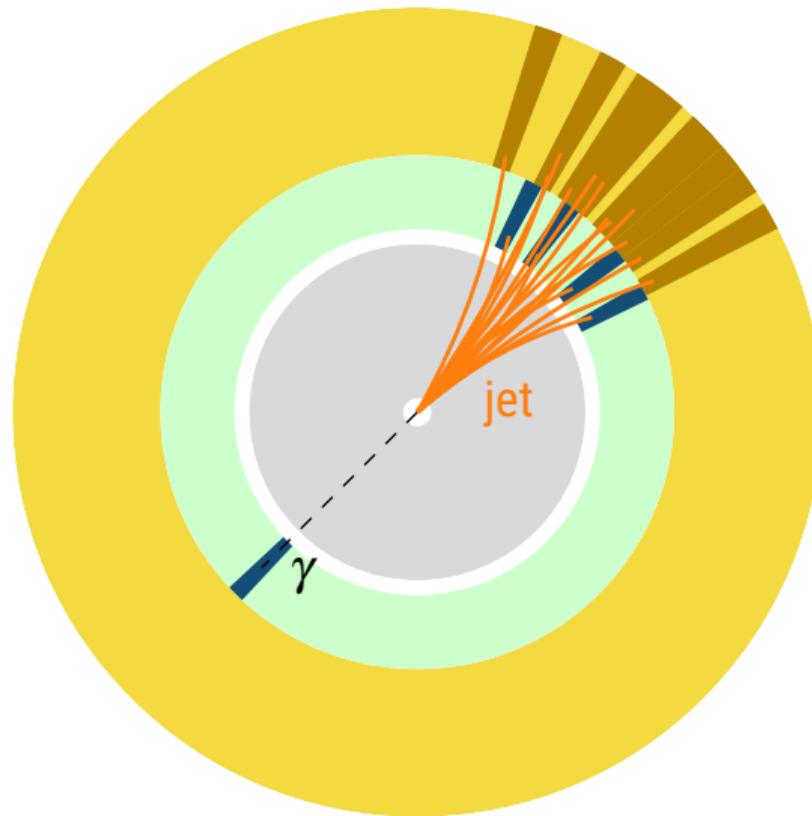
Appliqué aux simulations

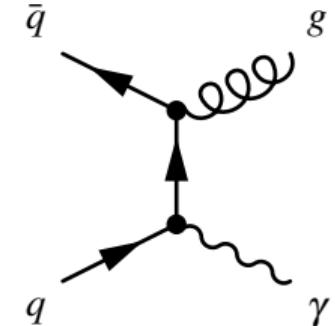
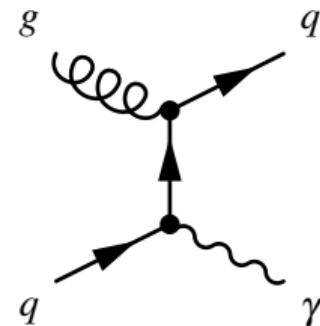
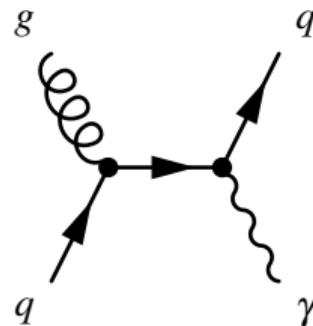


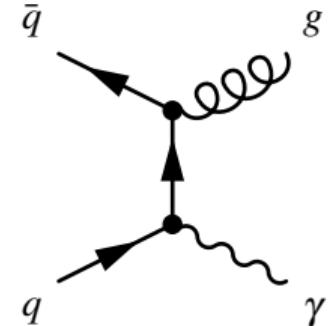
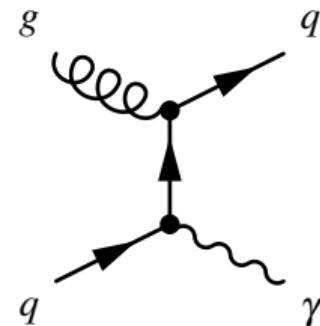
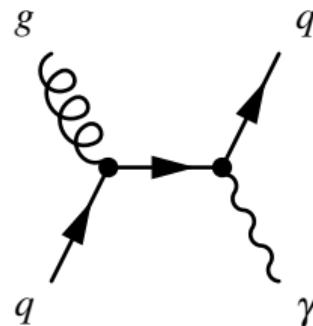
## Appliqué aux données



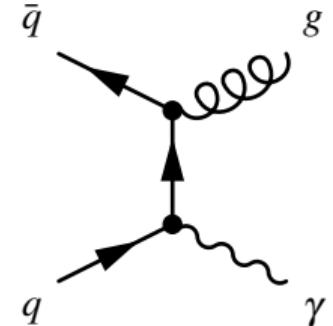
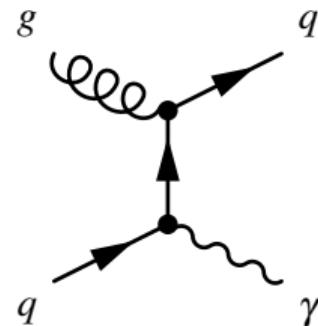
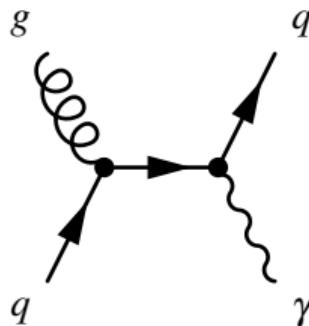






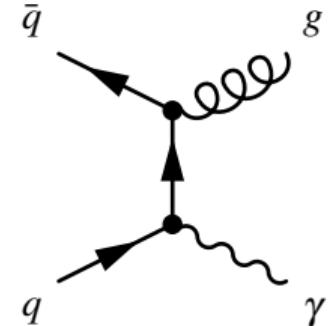
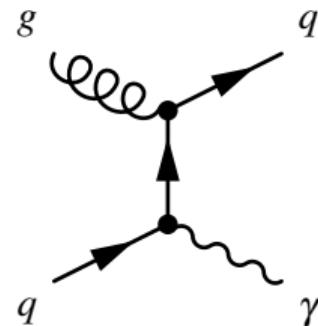
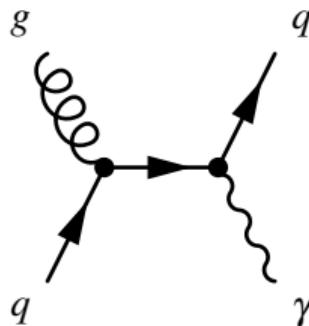


$$\vec{p}_{T,ptcl}^{\gamma} + \vec{p}_{T,ptcl}^{jet} = \vec{0} \Rightarrow p_{T,ptcl}^{\gamma} = p_{T,ptcl}^{jet}$$



$$\vec{p}_{\text{T ptcl}}^\gamma + \vec{p}_{\text{T ptcl}}^{\text{jet}} = \vec{0} \Rightarrow p_{\text{T ptcl}}^\gamma = p_{\text{T ptcl}}^{\text{jet}}$$

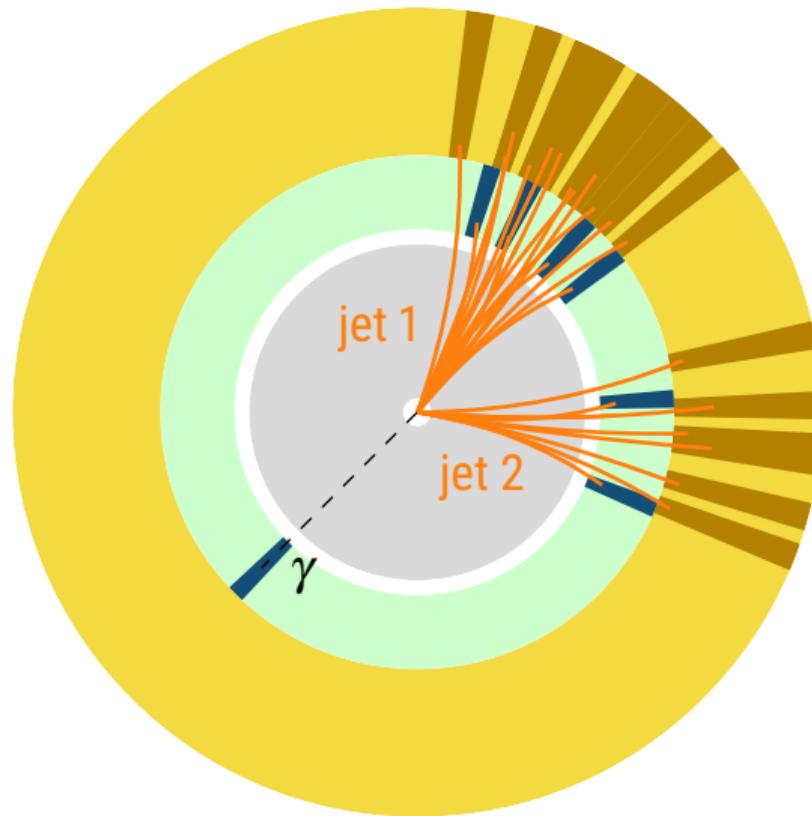
$$R = \frac{p_{\text{T ptcl}}^{\text{jet}}}{p_{\text{T ptcl}}^{\gamma}} = \frac{p_{\text{T ptcl}}^{\text{jet}}}{p_{\text{T ptcl}}^\gamma} \simeq \frac{p_{\text{T ptcl}}^{\text{jet}}}{p_{\text{T ptcl}}^{\gamma}}$$

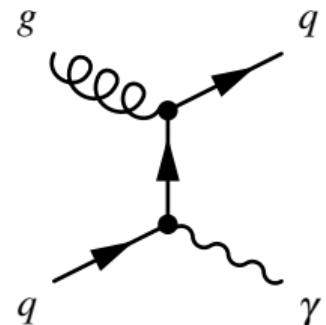


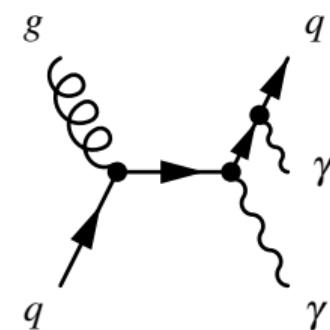
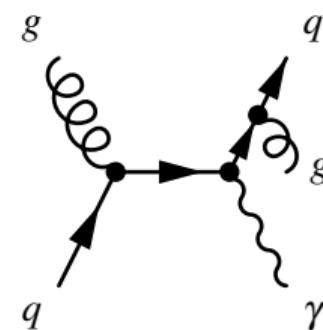
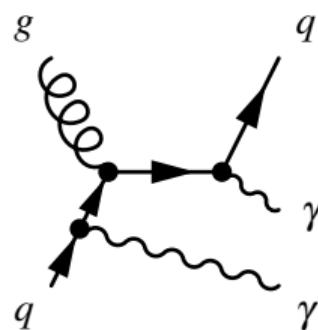
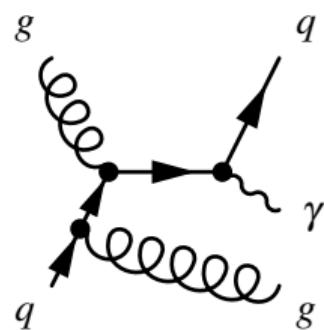
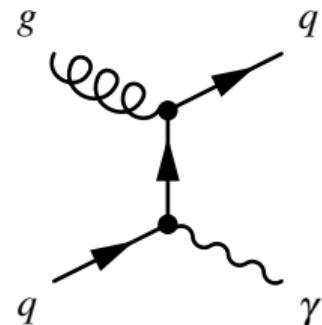
$$\vec{p}_{T,ptcl}^\gamma + \vec{p}_{T,ptcl}^{\text{jet}} = \vec{0} \Rightarrow p_{T,ptcl}^\gamma = p_{T,ptcl}^{\text{jet}}$$

$$R = \frac{p_{T,\text{reco}}^{\text{jet}}}{p_{T,\text{ptcl}}^{\text{jet}}} = \frac{p_{T,\text{reco}}^{\text{jet}}}{p_{T,\text{ptcl}}^\gamma} \simeq \frac{p_{T,\text{reco}}^{\text{jet}}}{p_{T,\text{reco}}^\gamma}$$

$$R_{bal} = \frac{p_{T,\text{reco}}^{\text{jet}}}{p_{T,\text{ptcl}}^\gamma}$$







$$R_{bal} = \frac{p_T^{\text{jet 1}}}{p_T^\gamma}$$

$$\alpha = \frac{p_T^{\text{jet 2}}}{p_T^\gamma}$$

$$\vec{p}_{T\text{ptcl}}^{\gamma} + \vec{p}_{T\text{ptcl}}^{\text{recul}} = \vec{0}$$

$$\vec{p}_T^\gamma + \vec{p}_T^{\text{recul}} = \vec{0}$$

$$\underbrace{\vec{p}_T^\gamma + R_{MPF} \vec{p}_T^{\text{recul}}}_{\vec{p}_T^{\text{reco}}} = -\vec{E}_T^{\text{miss}} \Rightarrow R_{MPF} = 1 + \frac{\vec{p}_T^\gamma \cdot \vec{E}_T^{\text{miss}}}{|\vec{p}_T^\gamma|^2}$$

# Jet Energy Resolution

- ▶ Remember  $R_{bal}$  definition,

$$R_{bal} = \frac{p_{T\text{reco}}^{\text{1st jet}}}{p_{T\text{reco}}^\gamma}$$

# Jet Energy Resolution

- ▶ Remember  $R_{bal}$  definition,

$$R_{bal} = \frac{p_{T\text{reco}}^{\text{1st jet}}}{p_{T\text{reco}}^\gamma}$$

Then

$$R_{bal} = \underbrace{\frac{p_{T\text{reco}}^{\text{1st jet}}}{p_{T\text{ptcl}}^{\text{1st jet}}}}_{\sigma_{\text{jet}} = \text{JER}} \times \underbrace{\frac{p_{T\text{ptcl}}^{\text{1st jet}}}{p_{T\text{ptcl}}^\gamma}}_{\text{PLI}} \times \underbrace{\frac{p_{T\text{ptcl}}^\gamma}{p_{T\text{reco}}^\gamma}}_{\sigma_\gamma \equiv 1}$$

- ▶ PLI: Particle Level Imbalance (pile-up, radiations, neutrinos...),  $\rightarrow 0$  when  $\alpha \rightarrow 0$ .

# Jet Energy Resolution

- ▶ Remember  $R_{bal}$  definition,

$$R_{bal} = \frac{p_{T\text{reco}}^{\text{1st jet}}}{p_{T\text{reco}}^\gamma}$$

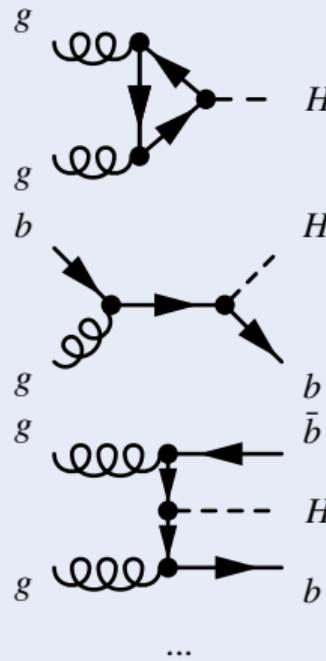
Then

$$R_{bal} = \underbrace{\frac{p_{T\text{reco}}^{\text{1st jet}}}{p_{T\text{ptcl}}^{\text{1st jet}}}}_{\sigma_{\text{jet}} = \text{JER}} \times \underbrace{\frac{p_{T\text{ptcl}}^{\text{1st jet}}}{p_{T\text{ptcl}}^\gamma}}_{\text{PLI}} \times \underbrace{\frac{p_{T\text{ptcl}}^\gamma}{p_{T\text{reco}}^\gamma}}_{\sigma_\gamma \equiv 1}$$

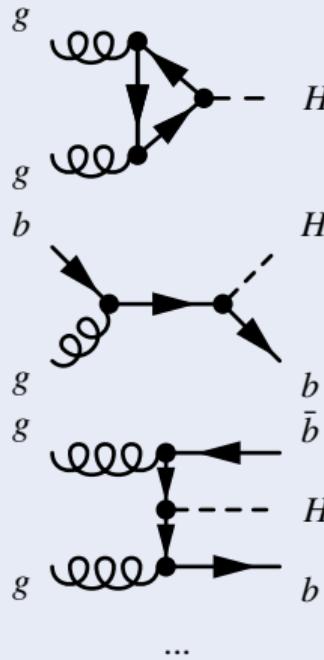
- ▶ PLI: Particle Level Imbalance (pile-up, radiations, neutrinos...),  $\rightarrow 0$  when  $\alpha \rightarrow 0$ .

$$\text{JER} = \sigma_{\text{jet}} = \sqrt{\sigma_{R_{bal}}^2 - \sigma_{\text{PLI}}^2}$$

## Higgs production

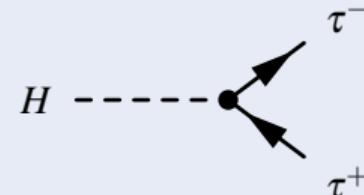


## Higgs production

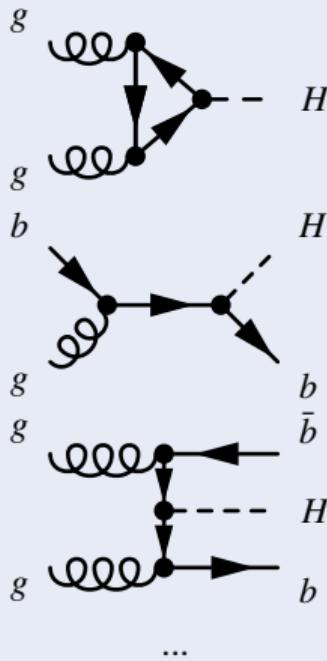


$H \rightarrow \tau\tau$

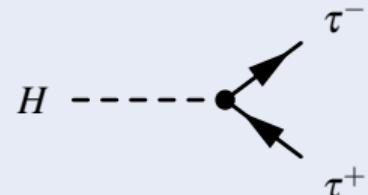
$\otimes$



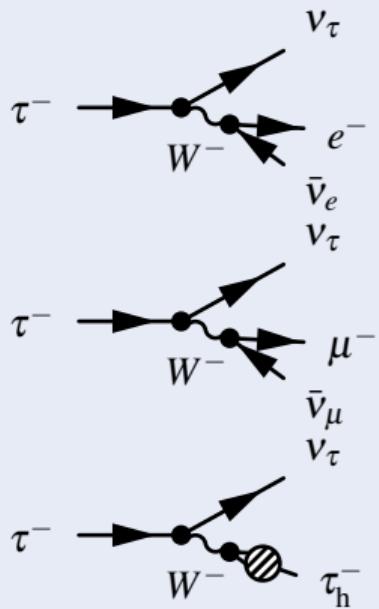
## Higgs production



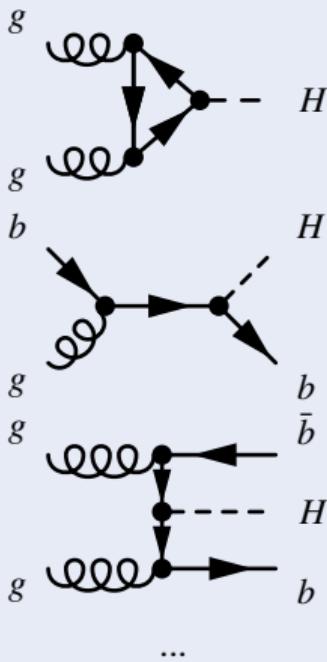
$H \rightarrow \tau\tau$



## $\tau$ decays

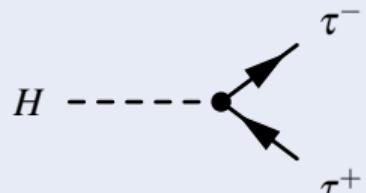


## Higgs production



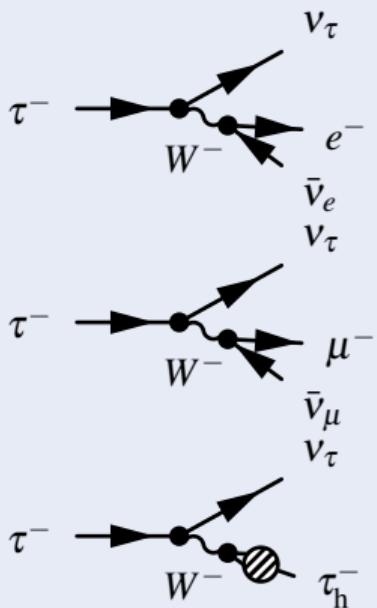
$\{0, 1, 2\}$  jets

## $H \rightarrow \tau\tau$



2 taus

## $\tau$ decays



$\{1, 2\}$  neutrinos per tau  
 $+ \{e, \mu, \tau_h\}$

- ▶ NN inputs are based on reconstructed variables.

- ▶ NN inputs are based on reconstructed variables.
- ▶  $H \rightarrow \tau\tau$  decays:
  - ▷ visible decay products  $\rightarrow p_T^{(1,2)}, \eta^{(1,2)}$  and  $\phi^{(1,2)}$ ,
  - ▷ MET to account for neutrinos  $\rightarrow E_T^{\text{miss}}, \phi^{\text{MET}}$ ;

- ▶ NN inputs are based on reconstructed variables.
- ▶  $H \rightarrow \tau\tau$  decays:
  - ▷ visible decay products  $\rightarrow p_T^{(1,2)}, \eta^{(1,2)}$  and  $\phi^{(1,2)}$ ,
  - ▷ MET to account for neutrinos  $\rightarrow E_T^{\text{miss}}, \phi^{\text{MET}}$ ;
- ▶ Higgs production:
  - ▷ two leading jets  $\rightarrow p_T^{(j1,j2)}, \eta^{(j1,j2)}$  and  $\phi^{(j1,j2)}$ ;

► NN inputs are based on reconstructed variables.

►  $H \rightarrow \tau\tau$  decays:

- ▷ visible decay products  $\rightarrow p_T^{(1,2)}, \eta^{(1,2)}$  and  $\phi^{(1,2)}$ ,
- ▷ MET to account for neutrinos  $\rightarrow E_T^{\text{miss}}, \phi^{\text{MET}}$ ;

► Higgs production:

- ▷ two leading jets  $\rightarrow p_T^{(j1,j2)}, \eta^{(j1,j2)}$  and  $\phi^{(j1,j2)}$ ;

► Higher level variables:

- ▷ transverse masses  $m_T^1, m_T^2, m_T^{\tau\tau}$ ,
- ▷ total transverse mass  $m_T^{\text{tot}}$ .

► NN inputs are based on reconstructed variables.

►  $H \rightarrow \tau\tau$  decays:

- ▷ visible decay products  $\rightarrow p_T^{(1,2)}, \eta^{(1,2)}$  and  $\phi^{(1,2)}$ ,
- ▷ MET to account for neutrinos  $\rightarrow E_T^{\text{miss}}, \phi^{\text{MET}}$ ;

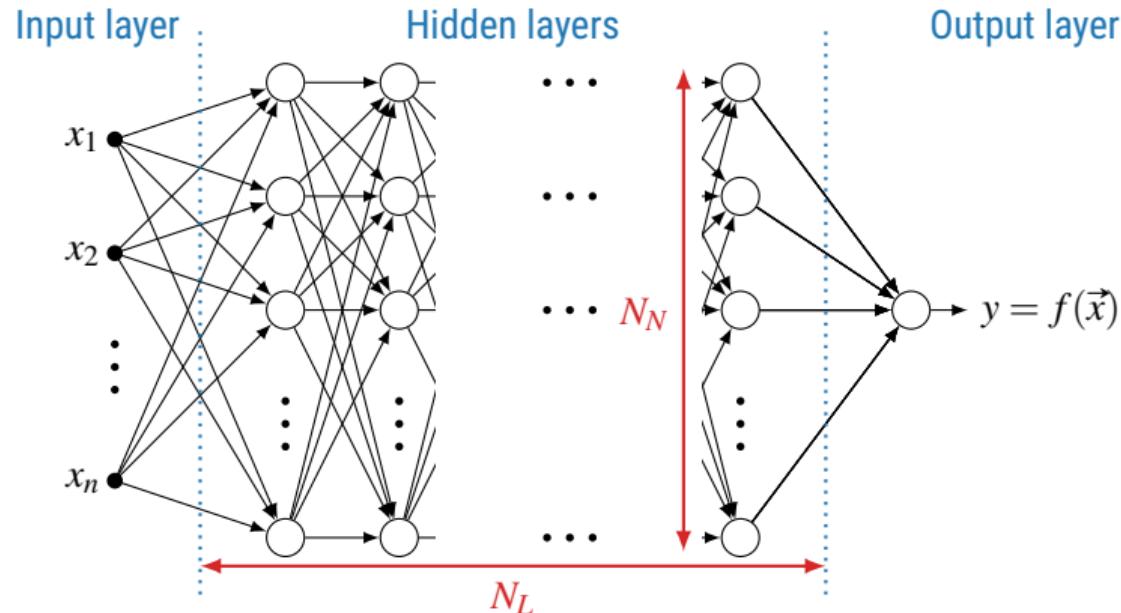
► Higgs production:

- ▷ two leading jets  $\rightarrow p_T^{(j1,j2)}, \eta^{(j1,j2)}$  and  $\phi^{(j1,j2)}$ ;

► Higher level variables:

- ▷ transverse masses  $m_T^1, m_T^2, m_T^{\tau\tau}$ ,
- ▷ total transverse mass  $m_T^{\text{tot}}$ .

► 18 inputs.



$$N_L \in [2, 10] \cup \{15\}$$

$$N_N \in \{500, 1000, 1500, 2000\}$$

► The "bottleneck" variant:

- ▷ Get a smoother reduction from  $\sim 1k$  neurons in hidden layers to 1 neuron for the output layer.
- ▷ Set a maximum value for  $N_N$  in the 3 last hidden layers: [1000, 500, 100].

► Example: neurons per layers with settings  $N_L = 4, N_N = 2000$  gives

Case	Hidden layers				Output layer
without bottleneck	2000	2000	2000	2000	1
with bottleneck	2000	1000	500	100	1

► Activation functions:

- ▷ relu for hidden layers;
- ▷ linear for output layer.

- ▶ GeV switched to TeV
- ▶ Target is  $m_H$
- ▶ Get a flat target distribution for the training, validating and testing sub-samples.
- ▶ Train a NN for:
  - ▷ all channels at once;
  - ▷ each channel separately;
  - ▷ full-hadronic, semi-leptonic, full-leptonic channels (categorize per amount of neutrinos in the final state).

Thank you for your attention!