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

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

Recherche de bosons de Higgs supplémentaires de haute masse se désintégrant en paire de taus dans l'expérience CMS au LHC à l'aide du *machine learning*

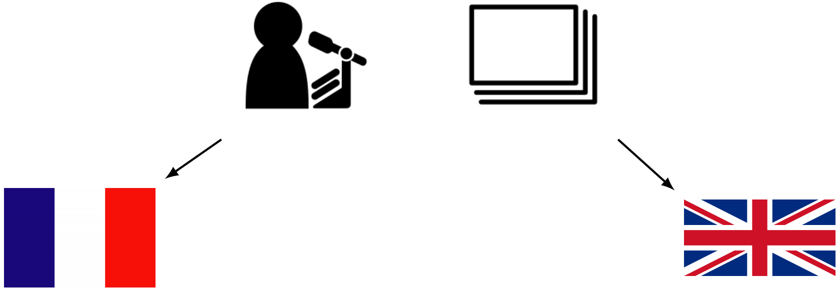
Soutenance de thèse de doctorat

Lucas TORTEROTOT

Institut de Physique des deux Infinis – Lyon

8 juillet 2021

Lang(u)age



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

Search for additional heavy Higgs bosons decaying to tau lepton pair in the CMS experiment at LHC with machine learning techniques

Ph.D. thesis defense

Lucas TORTEROTOT

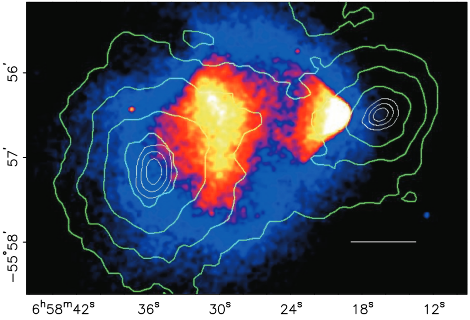
Institut de Physique des deux Infinis – Lyon

July 8, 2021

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

Current standard model status

- Robust and predictive (top quark, W, Z and one Higgs boson...)
- Still not good enough, unable to explain some observations such as:
  - dark matter $\longrightarrow$
  - matter vs antimatter asymmetry
  - naturalness problem
  - ...
- 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** **heavy** **Higgs bosons** **decaying to tau lepton pair** in the **CMS experiment** at **LHC**

Keywords in title

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

type of particle

already one!

heavier

why this case?

why this one?

not lighter?

Part I
Phenomenology

Keywords in title

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

already one!

heavier

not lighter?

type of particle

why this one?

why this case?

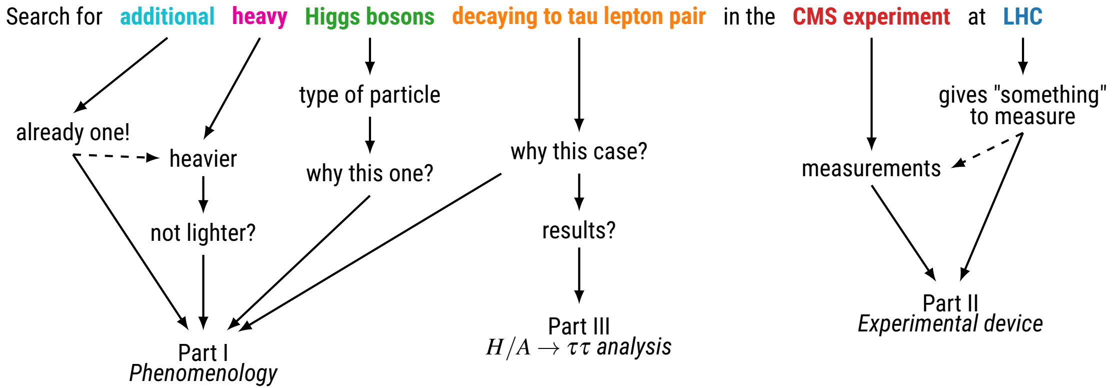
Part I
Phenomenology

measurements

gives "something" to measure

Part II
Experimental device

Keywords in title



Keywords in title

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

already one!

heavier

not lighter?

type of particle

why this one?

why this case?

results?

gives "something" to measure

measurements

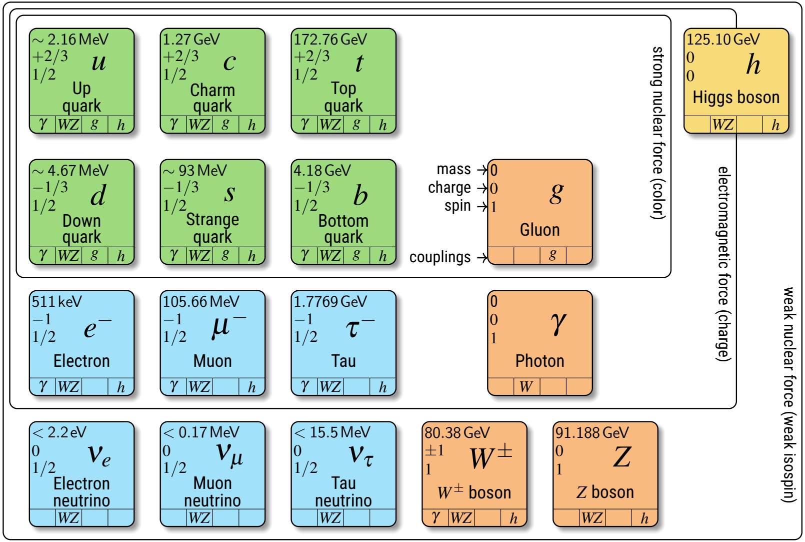
Part I
Phenomenology

Part III
$H/A \to \tau\tau$ *analysis*

Part II
Experimental device

+ Part IV: **with machine learning techniques**

The Standard Model



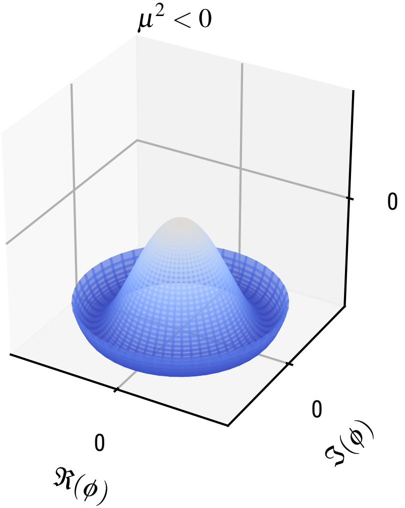
Higgs boson in the Standard Model

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_3 + i\phi_4 \\ \phi_1 + i\phi_2 \end{pmatrix}$$

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad \lambda > 0$$

$$\boxed{\langle \phi \rangle_0 = \frac{v}{\sqrt{2}} = \sqrt{\frac{-\mu^2}{2\lambda}} \neq 0}$$

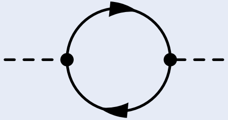
$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}$$



The Standard Model and naturalness problem

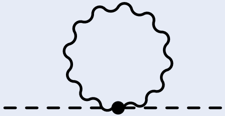
- Higgs mass measured: $m_h = 125.10 \pm 0.14\,\mathrm{GeV}$
- Higgs mass derivation: $m_h^2 = m_{h0}^2 - \frac{3}{8\pi^2} y_t^2 \Lambda^2 + \frac{1}{16\pi^2} g^2 \Lambda^2 + \frac{1}{16\pi^2} \lambda^2 \Lambda^2 + \ldots$

top quark



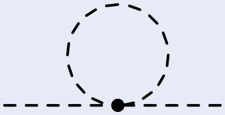
$$-\frac{3}{8\pi^2} y_t^2 \Lambda^2 \sim -(2\,\mathrm{TeV})^2$$

vector bosons



$$+\frac{1}{16\pi^2} g^2 \Lambda^2 \sim +(0.7\,\mathrm{TeV})^2$$

Higgs itself

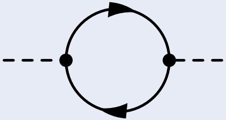


$$+\frac{1}{16\pi^2} \lambda^2 \Lambda^2 \sim +(0.5\,\mathrm{TeV})^2$$

▷ Particle Data Group. "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* **8** (Aug. 2020). DOI: 10.1093/ptep/ptaa104.

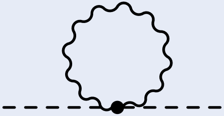
Supersymmetry

top quark



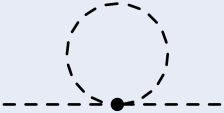
$$\sim -(2\,\mathrm{TeV})^2$$

vector bosons



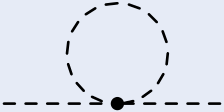
$$\sim +(0.7\,\mathrm{TeV})^2$$

Higgs itself



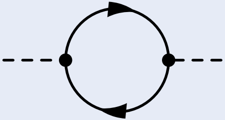
$$\sim +(0.5\,\mathrm{TeV})^2$$

stop quark



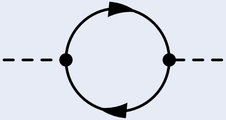
$$\sim +(2\,\mathrm{TeV})^2$$

bosinos



$$\sim -(0.7\,\mathrm{TeV})^2$$

Higgsinos



$$\sim -(0.5\,\mathrm{TeV})^2$$

2 Higgs doublets models for supersymmetry

$$\langle \phi_1 \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_1 \end{pmatrix}, \quad \langle \phi_2 \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_2 e^{i\xi} \end{pmatrix}$$

$$\boxed{\tan\beta = \frac{\langle \phi_2 \rangle_0}{\langle \phi_1 \rangle_0} = \frac{v_2}{v_1}}$$

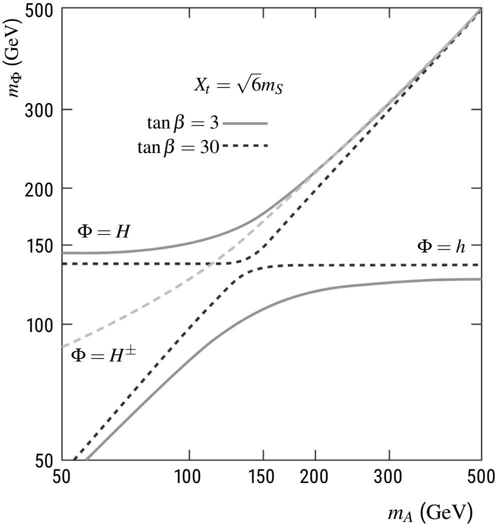
▷ J. F. Gunion et al. *The Higgs hunter's guide*. T. **80**. Upton, NY: Brookhaven Nat. Lab., 1989. URL: `https://cds.cern.ch/record/425736`.

Higgs bosons in the MSSM

Minimal Supersymmetric extension of Standard Model

5 Higgs bosons

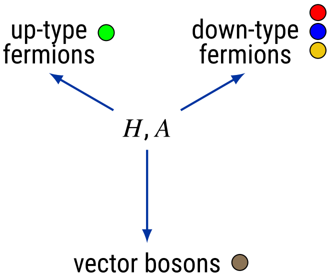
light scalar	h	SM or MSSM
heavy scalar	H	MSSM or SM
pseudo-scalar	A	MSSM
$+$ charged	H^+	MSSM
$-$ charged	H^-	MSSM

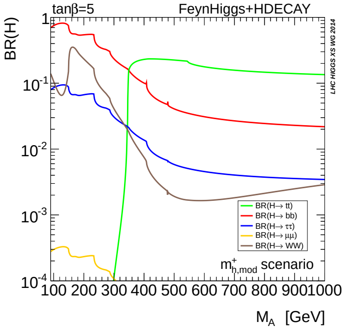


▷ The CMS Collaboration. "Search for additional neutral MSSM Higgs bosons in the di-tau final state in pp collisions at $\sqrt{s} = 13\,\mathrm{TeV}$". *Journal of High Energy Physics* **09**.007 (Sept. 2018). DOI: 10.1007/JHEP09(2018)007.

▷ Y. Nagashima. *Beyond the Standard Model of Elementary Particle Physics*. Weinheim: Wiley-VCH, June 2014. URL: http://cds.cern.ch/record/1620277.

$H/A \to \tau\tau$?

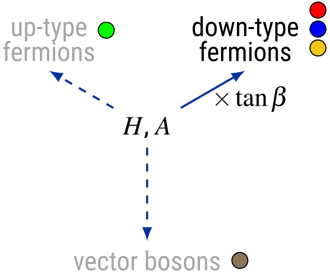


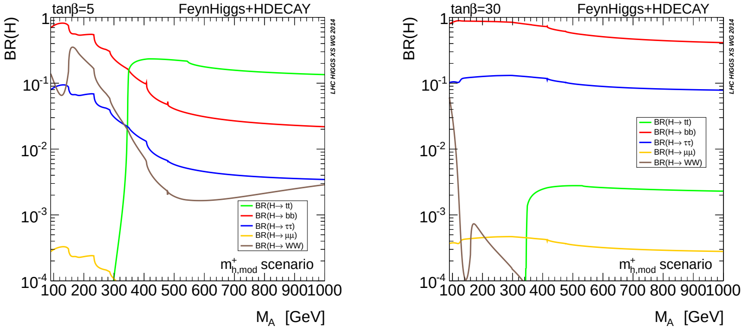


▷ The CMS Collaboration. "Search for additional neutral MSSM Higgs bosons in the di-tau final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$". *Journal of High Energy Physics* **09**.007 (Sept. 2018). DOI: 10.1007/JHEP09(2018)007.

▷ LHC Higgs Cross Section Working Group. "Higgs Properties". *Handbook of LHC Higgs Cross Sections*. **3**. CERN Yellow Reports: Monographs. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1559921.

$H/A \to \tau\tau$? – enhanced and suppressed couplings

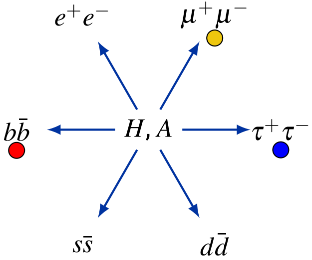


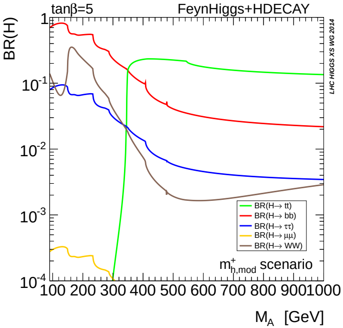


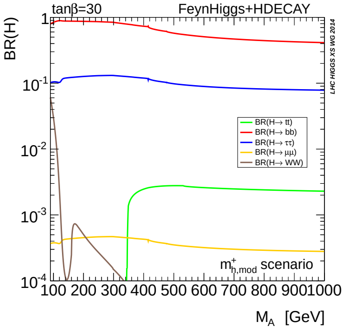
▷ The CMS Collaboration. “Search for additional neutral MSSM Higgs bosons in the di-tau final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$”. *Journal of High Energy Physics* **09**.007 (Sept. 2018). DOI: 10.1007/JHEP09(2018)007.

▷ LHC Higgs Cross Section Working Group. “Higgs Properties”. *Handbook of LHC Higgs Cross Sections*. **3**. CERN Yellow Reports: Monographs. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1559921.

$H/A \to \tau\tau$?



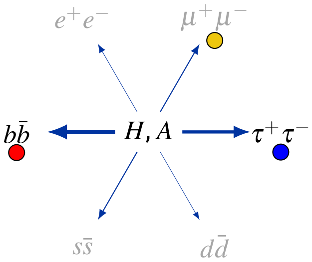


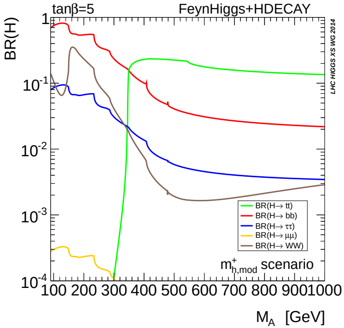


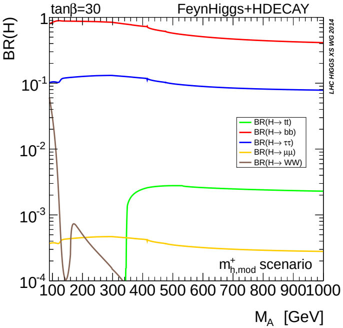
▷ The CMS Collaboration. "Search for additional neutral MSSM Higgs bosons in the di-tau final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$". *Journal of High Energy Physics* **09**.007 (Sept. 2018). DOI: 10.1007/JHEP09(2018)007.

▷ LHC Higgs Cross Section Working Group. "Higgs Properties". *Handbook of LHC Higgs Cross Sections*. **3**. CERN Yellow Reports: Monographs. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1559921.

$H/A \to \tau\tau$? – Higgs couplings and particules masses



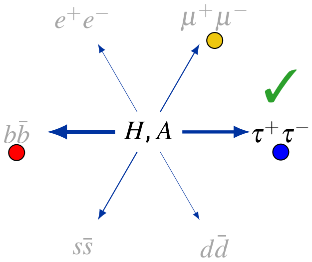


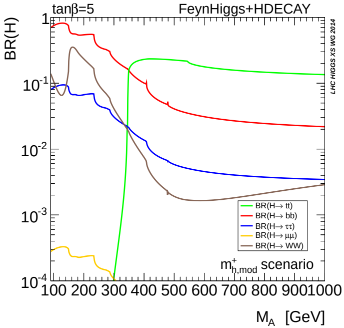


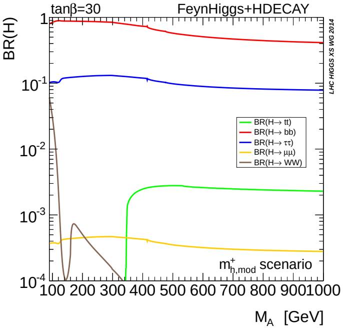
▷ The CMS Collaboration. "Search for additional neutral MSSM Higgs bosons in the di-tau final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$". *Journal of High Energy Physics* **09**.007 (Sept. 2018). DOI: 10.1007/JHEP09(2018)007.

▷ LHC Higgs Cross Section Working Group. "Higgs Properties". *Handbook of LHC Higgs Cross Sections*. **3**. CERN Yellow Reports: Monographs. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1559921.

$H/A \to \tau\tau$? – avoid hadronic background



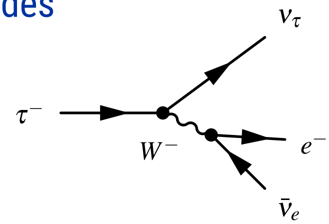




▷ The CMS Collaboration. “Search for additional neutral MSSM Higgs bosons in the di-tau final state in pp collisions at $\sqrt{s} = 13\,\text{TeV}$”. *Journal of High Energy Physics* **09**.007 (Sept. 2018). DOI: 10.1007/JHEP09(2018)007.

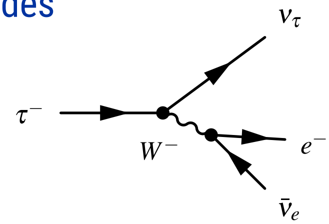
▷ LHC Higgs Cross Section Working Group. “Higgs Properties”. *Handbook of LHC Higgs Cross Sections*. **3**. CERN Yellow Reports: Monographs. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1559921.

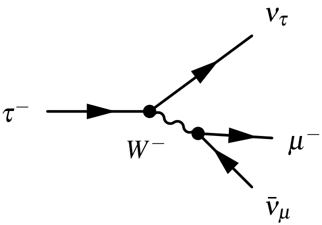
τ decay modes



▷ Particle Data Group. "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* **8** (Aug. 2020). DOI: 10.1093/ptep/ptaa104.

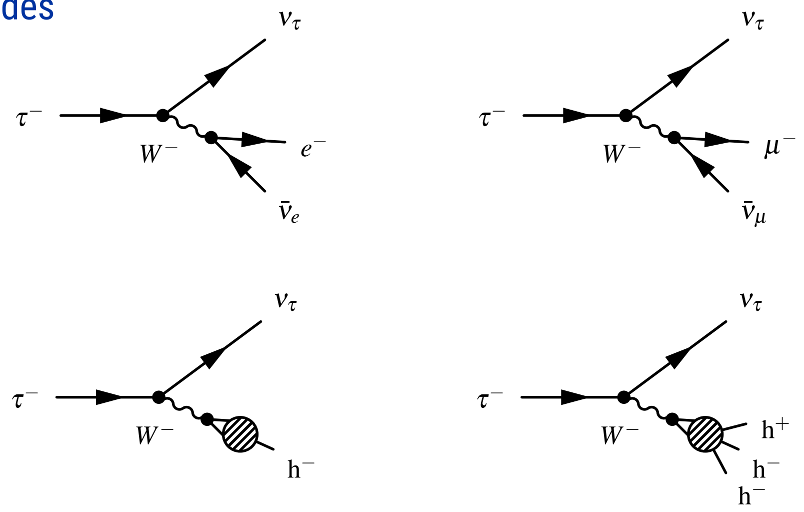
τ decay modes





▷ Particle Data Group. "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* **8** (Aug. 2020). DOI: `10.1093/ptep/ptaa104`.

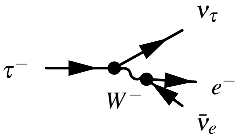
τ decay modes



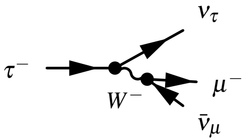
▷ Particle Data Group. "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* **8** (Aug. 2020). DOI: `10.1093/ptep/ptaa104`.

$H/A \to \tau\tau \to L_1 L_2$

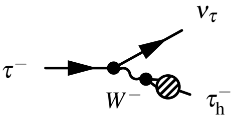
$\tau \to e + \nu_e + \nu_\tau \Rightarrow e$
17.8 %



$\tau \to \mu + \nu_\mu + \nu_\tau \Rightarrow \mu$
17.4 %



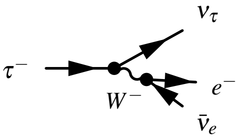
$\tau \to \text{hadrons} + \nu_\tau \Rightarrow \tau_\mathrm{h}$
64.8 %



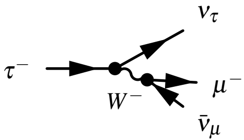
▷ Particle Data Group. "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* **8** (Aug. 2020). DOI: 10.1093/ptep/ptaa104.

$H/A \to \tau\tau \to L_1 L_2$

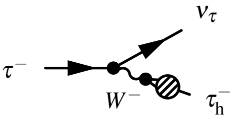
$\tau \to e + \nu_e + \nu_\tau \Rightarrow e$
17.8 %

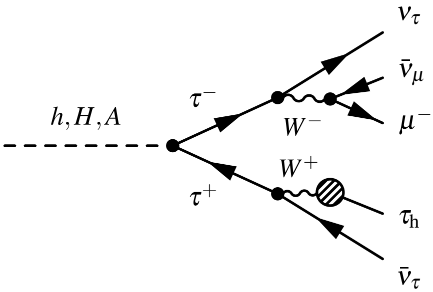


$\tau \to \mu + \nu_\mu + \nu_\tau \Rightarrow \mu$
17.4 %



$\tau \to \text{hadrons} + \nu_\tau \Rightarrow \tau_\mathrm{h}$
64.8 %

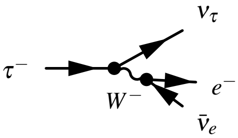




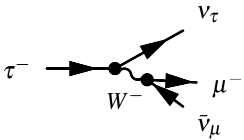
▷ Particle Data Group. “Review of Particle Physics”. *Progress of Theoretical and Experimental Physics* **8** (Aug. 2020). DOI: 10.1093/ptep/ptaa104.

$H/A \to \tau\tau \to L_1 L_2$

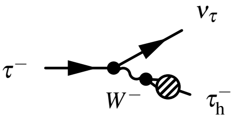
$\tau \to e + \nu_e + \nu_\tau \Rightarrow e$
$17.8\,\%$

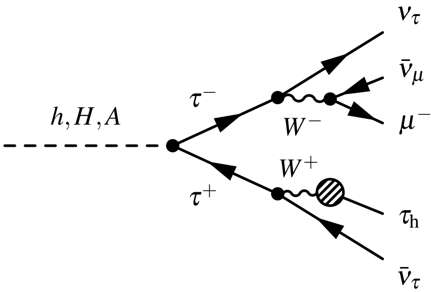


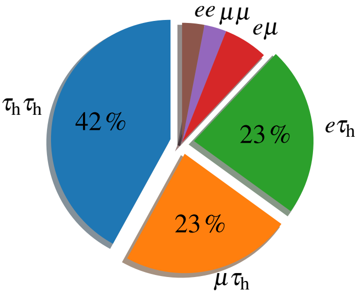
$\tau \to \mu + \nu_\mu + \nu_\tau \Rightarrow \mu$
$17.4\,\%$



$\tau \to \text{hadrons} + \nu_\tau \Rightarrow \tau_\mathrm{h}$
$64.8\,\%$







▷ Particle Data Group. "Review of Particle Physics". *Progress of Theoretical and Experimental Physics* **8** (Aug. 2020). DOI: 10.1093/ptep/ptaa104.

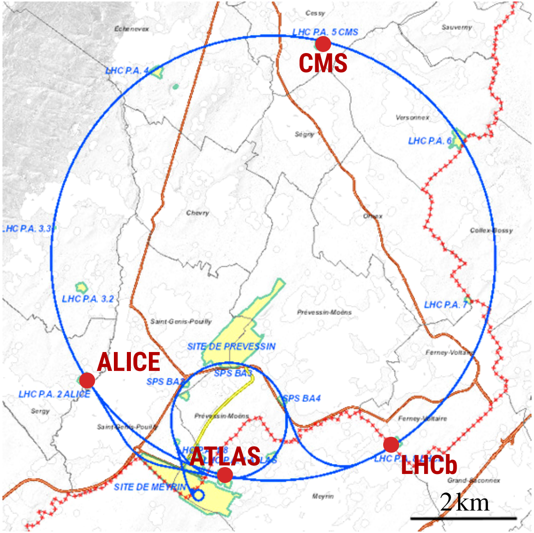
Principle

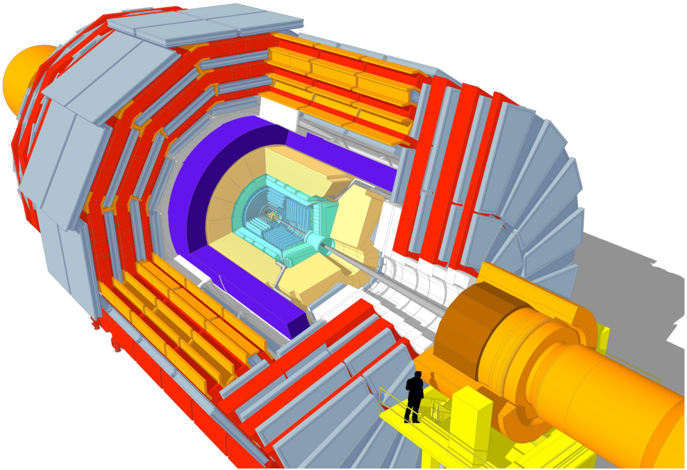
$$E = mc^2$$

mass (new particles) from the collision energy

CERN LHC


Londres
Bristol
Cardiff
Brighton
Southampton
Plymouth
Manche (mer)
Guernesey
Jersey
Brest
Rennes
Nantes
Tours
Paris
Anvers
Bruxelles
Lille
Belgique
Cologne
Dortmund
Allemagne
Leipzig
Francfort-sur-le-Main
Luxembourg
Mannheim
Nuremberg
Stuttgart
Strasbourg
Augsbourg
Munich
Salzbourg
Zurich
Liechtenstein
Suisse
CERN
France
La Rochelle
Limoges
Clermont-Ferrand
Lyon
Golfe de Gascogne
Bordeaux
Grenoble
Turin
Milan
Vérone
Venise
Padoue
Gênes
Bologne
Toulouse
Montpellier
Monaco
Marseille
Cannes
Pise
Saint-Marin
Florence
Italie
Santander
Saint-Sébastien
Bilbao
Vitoria-Gasteiz
Andorre

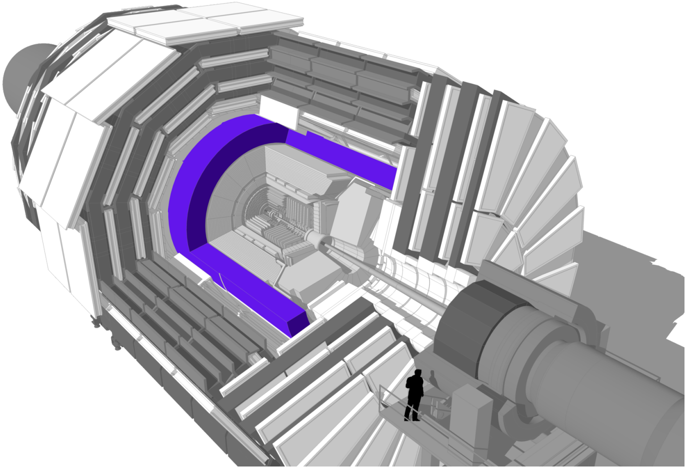

CMS
LHC P.A. 5 CMS
LHC P.A. 4
LHC P.A. 6
LHC P.A. 3.3
LHC P.A. 3.2
LHC P.A. 7
ALICE
LHC P.A. 2 ALICE
SITE DE PREVESSIN
SPS BA3
SPS BA4
ATLAS
LHCb
SITE DE MEYRIN
Cessy
Echenevex
Sauverny
Versonnex
Ségny
Chevry
Ornex
Collex-Bossy
Prévessin-Moëns
Saint-Genis-Pouilly
Ferney-Voltaire
Sergy
Meyrin
Grand-Saconnex
2 km



CMS detector

- Mass: $\sim 14{,}000\,\mathrm{t}$
- Diameter: $15\,\mathrm{m}$
- Length: $28.7\,\mathrm{m}$

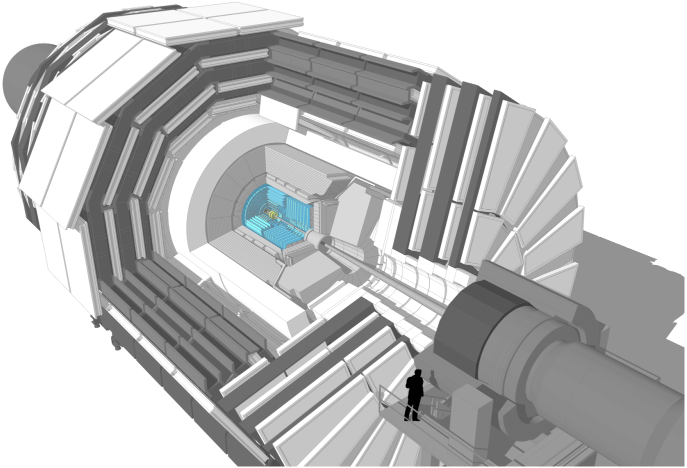
$\Rightarrow$ How to *see* the particles?



Solenoid

- Niobium titanium coil
- Superconducting
- $\sim 18{,}000\,\mathrm{A}$
- $4\,\mathrm{T}$ in the inner volume

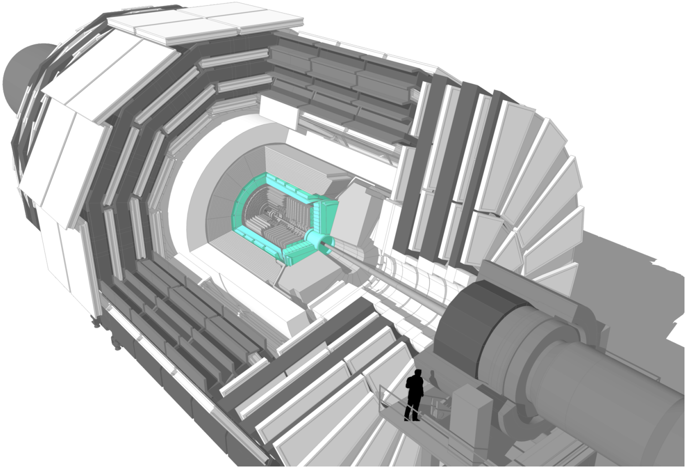
$\Rightarrow$ Bends charged particles trajectories in the transverse plane



Tracker

- Made of Silicon
- Inner: pixels ($100 \times 150\,\mu\text{m}^2$, $\sim 1.9\,\text{m}^2$, $\sim 124\,\text{M}$ channels
- Outer: microstrips ($80 - 180\,\mu\text{m}$) $\sim 200\,\text{m}^2$ $\sim 9.6\,\text{M}$ channels

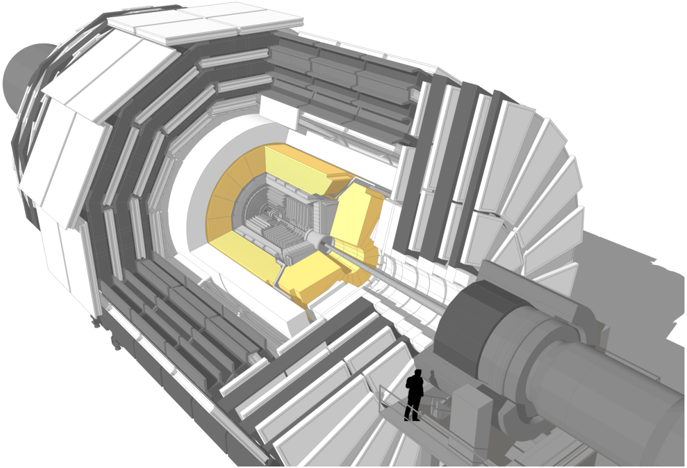
$\Rightarrow$ Charged particles leave hits when going through



Electromagnetic CALorimeter

- $\sim 76{,}000$ scintillating $PbWO_4$ crystals

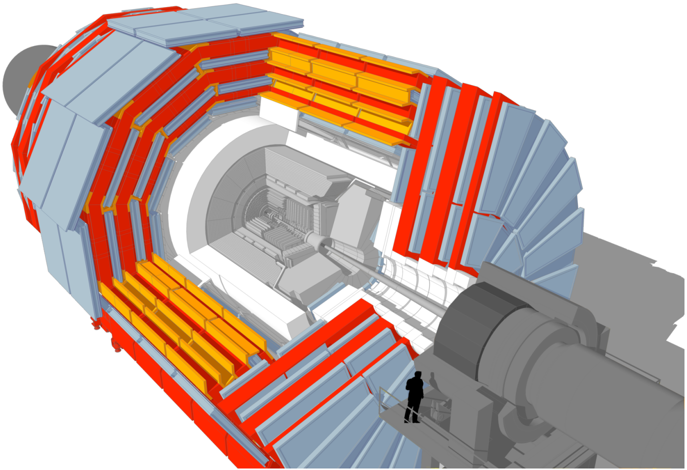
$\Rightarrow$ Electrons and photons are stopped, energy deposits



Hadronic CALorimeter

- Brass + plastic scintillator, ~ 7000 channels

$\Rightarrow$ Hadrons are stopped, energy deposits



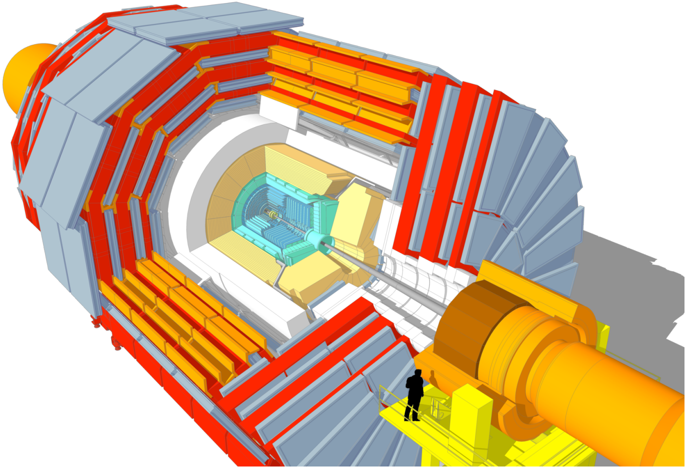
Steel return yoke (red)

- Allows for $2\,\mathrm{T}$ magnetic field around the solenoid

Muon chambers (blue-gray)

- Barrel: 250 drift tubes, 480 resistive plate chambers
- Endcaps: 540 cathode strips, 576 resistive plate chambers

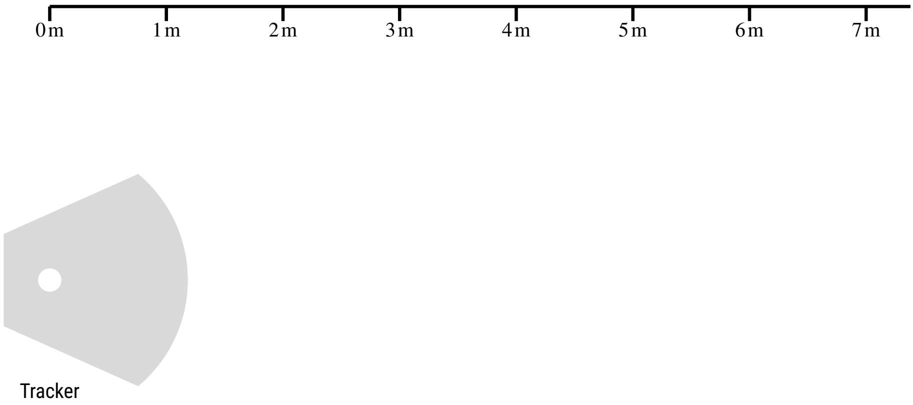
$\Rightarrow$ Charged particles leave hits when going through (only muons do)

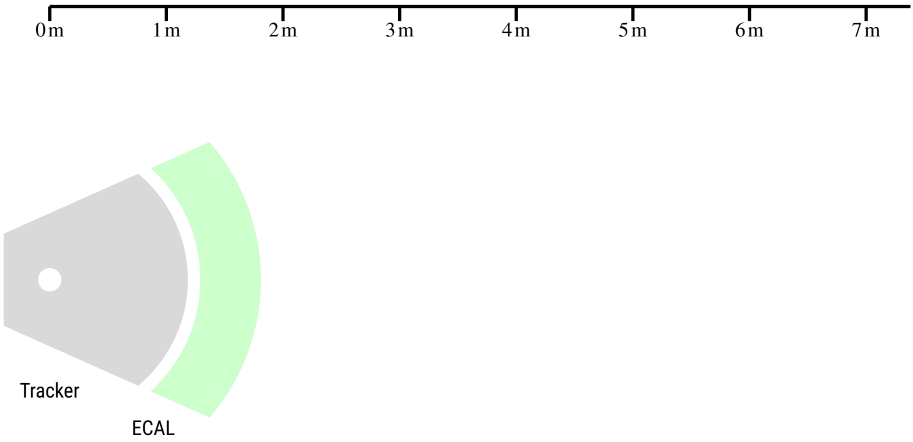


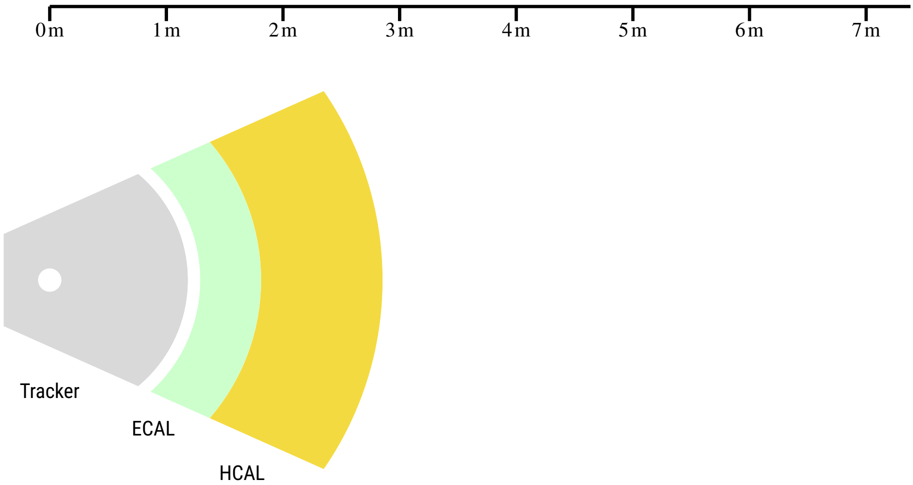
Sensitive parts of CMS

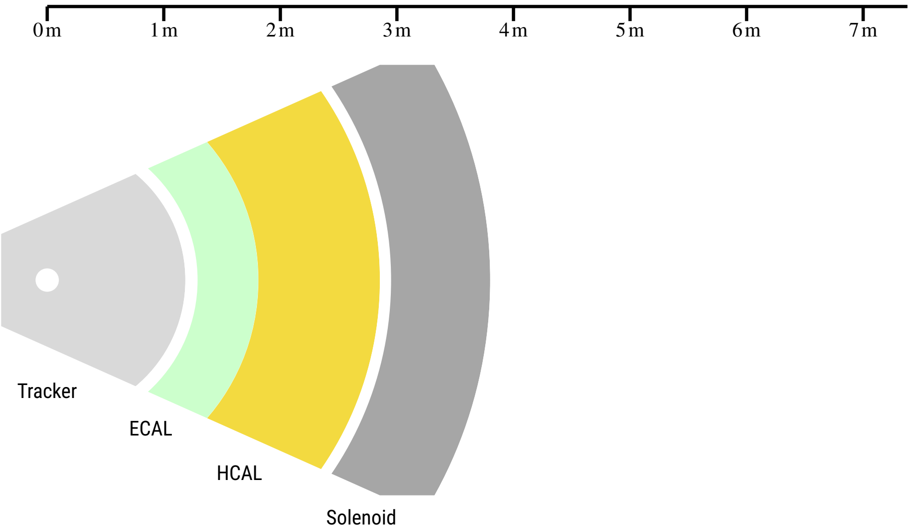
Combine sub-detectors signals to determine which particles were there!

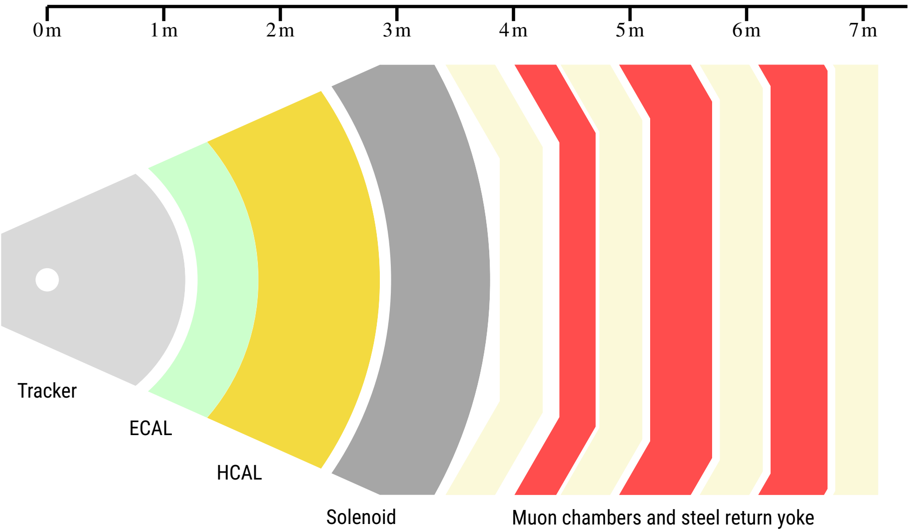

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m

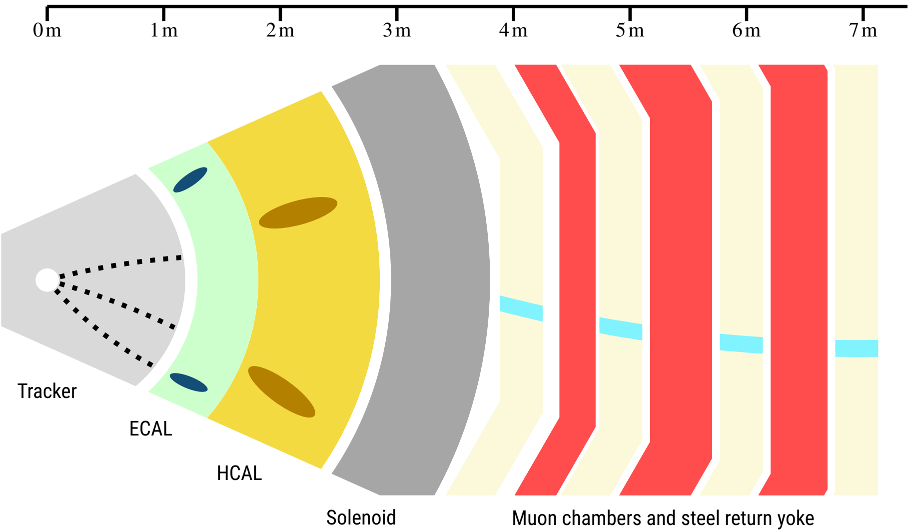

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Tracker

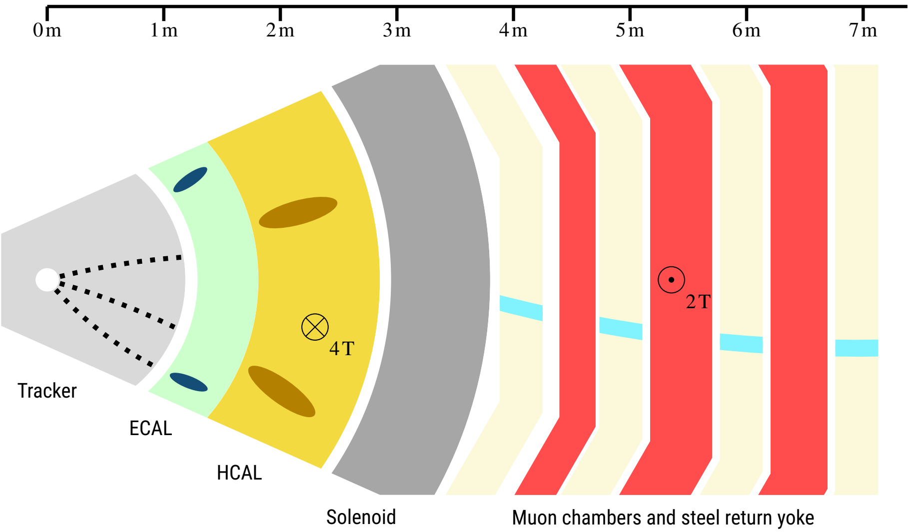

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Tracker
ECAL

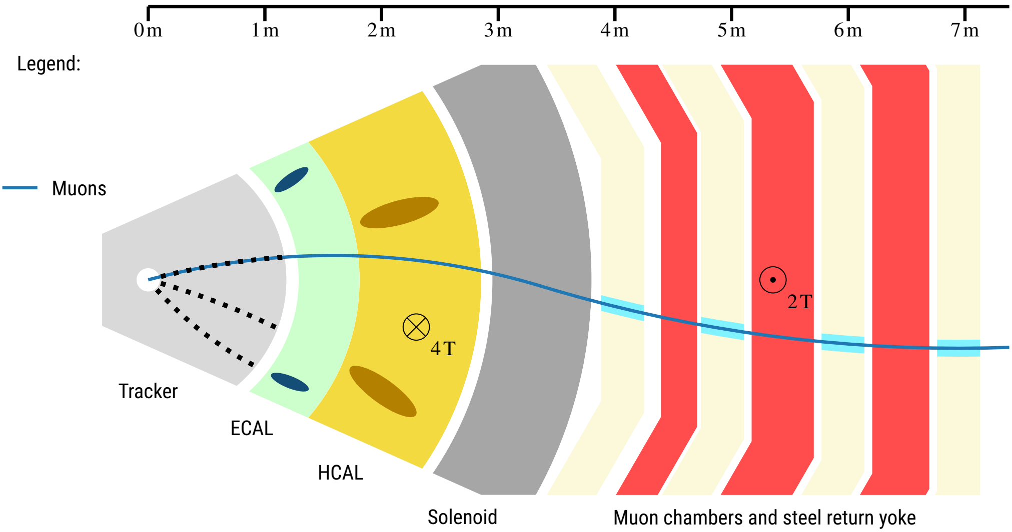

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Tracker
ECAL
HCAL

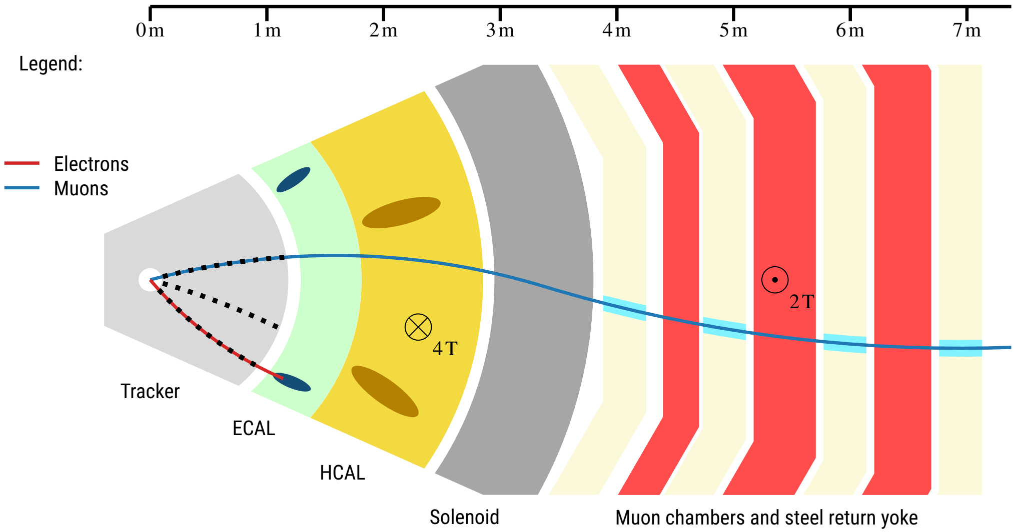

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Tracker
ECAL
HCAL
Solenoid

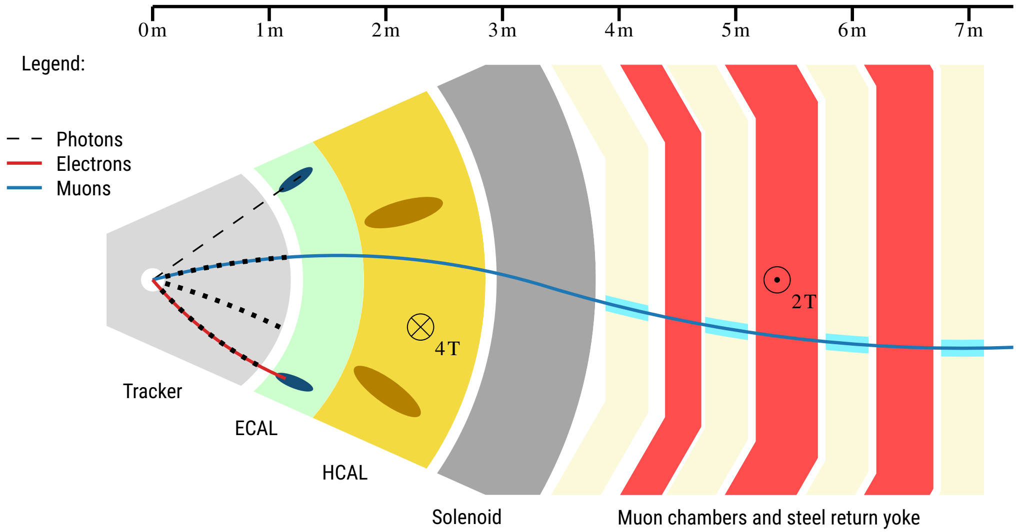

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Tracker
ECAL
HCAL
Solenoid
Muon chambers and steel return yoke

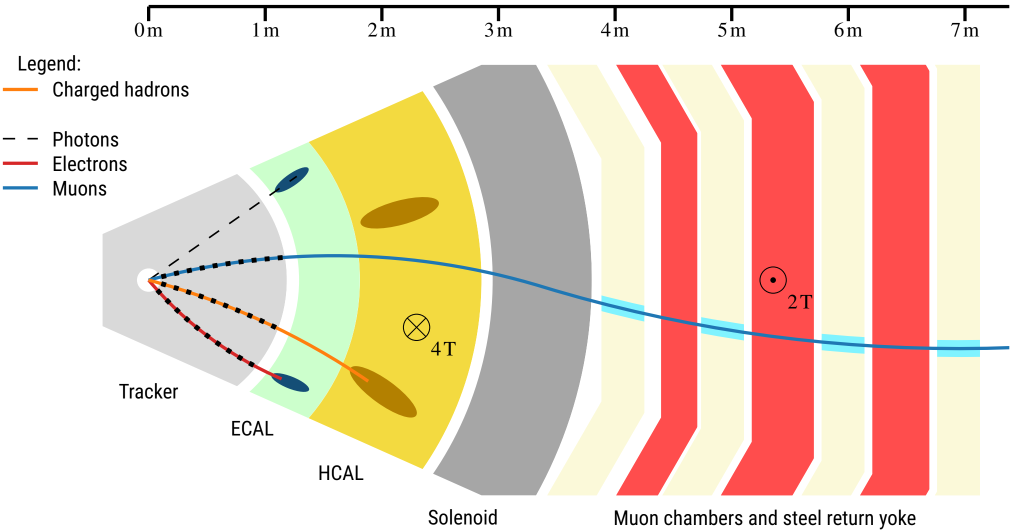

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Tracker
ECAL
HCAL
Solenoid
Muon chambers and steel return yoke

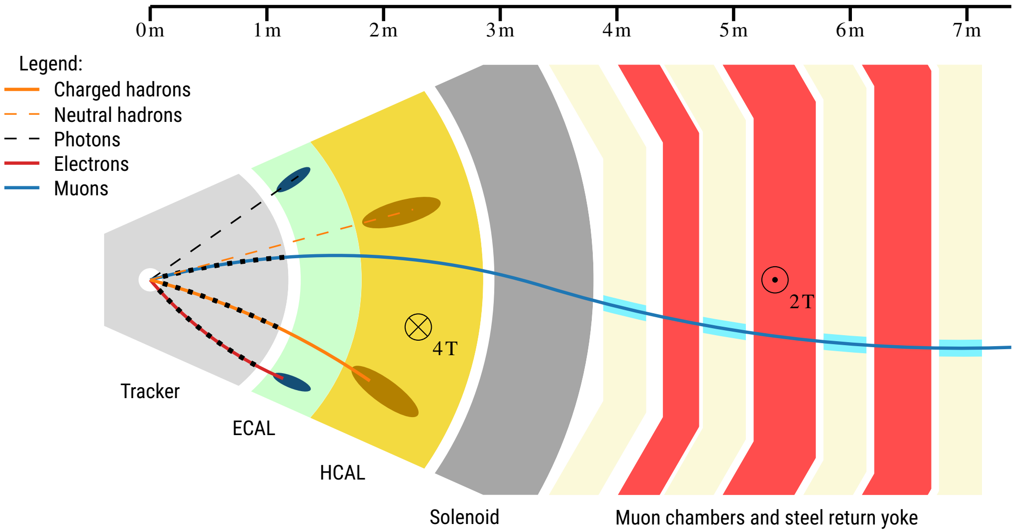

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
4 T
2 T
Tracker
ECAL
HCAL
Solenoid
Muon chambers and steel return yoke


0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Legend:
Muons
Tracker
ECAL
HCAL
Solenoid
Muon chambers and steel return yoke
4 T
2 T


0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Legend:
Electrons
Muons
4 T
2 T
Tracker
ECAL
HCAL
Solenoid
Muon chambers and steel return yoke

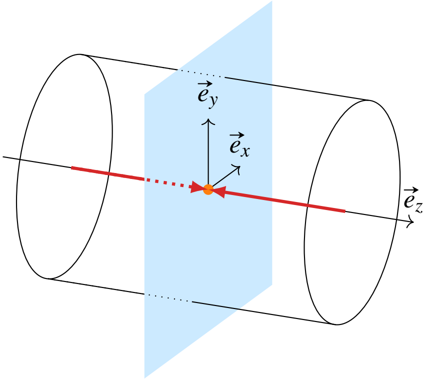

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Legend:
Photons
Electrons
Muons
4 T
2 T
Tracker
ECAL
HCAL
Solenoid
Muon chambers and steel return yoke

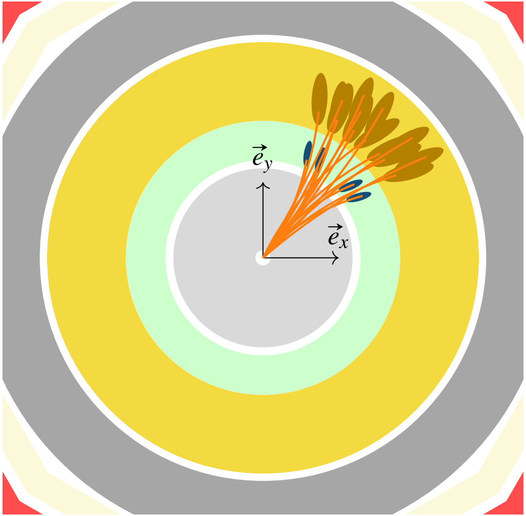

0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Legend:
Charged hadrons
Photons
Electrons
Muons
4 T
2 T
Tracker
ECAL
HCAL
Solenoid
Muon chambers and steel return yoke


0 m
1 m
2 m
3 m
4 m
5 m
6 m
7 m
Legend:
Charged hadrons
Neutral hadrons
Photons
Electrons
Muons
4 T
2 T
Tracker
ECAL
HCAL
Solenoid
Muon chambers and steel return yoke

Neutrinos and missing transverse energy (MET)

$(\vec{e}_x, \vec{e}_y)$ = transverse plane ($\eta = 0$)

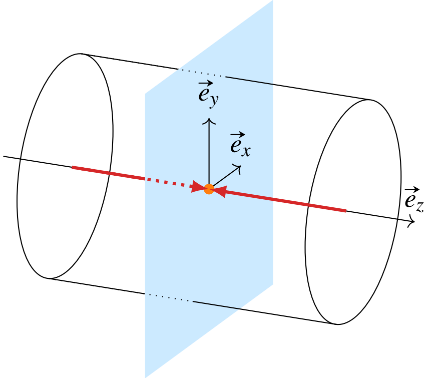


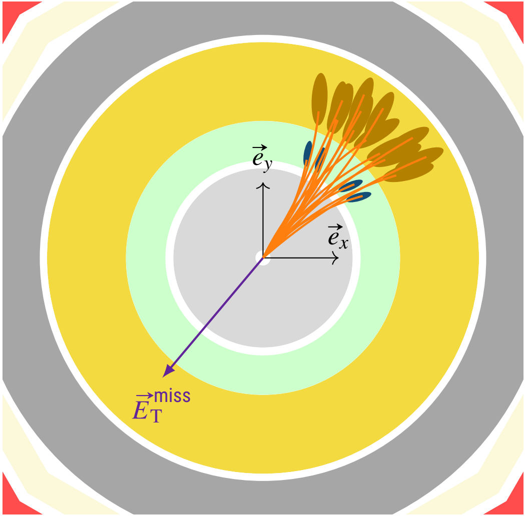


Conserved momentum: $\sum_{\text{initial state}} \vec{p}_\mathrm{T} = \sum_{\text{final state}} \vec{p}_\mathrm{T} = \vec{0}$

Neutrinos and missing transverse energy (MET)

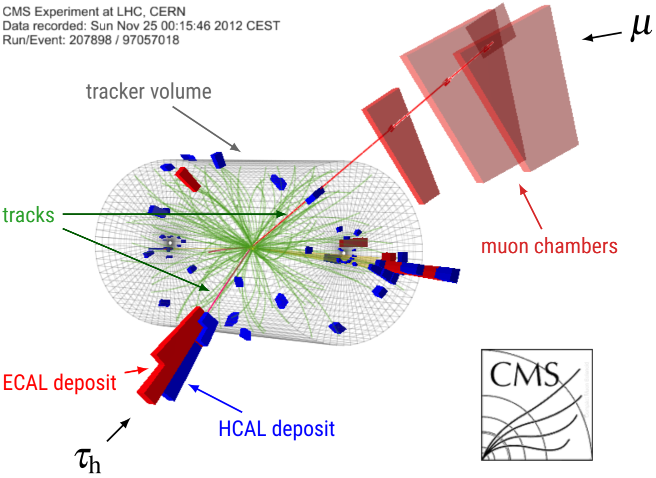
$(\vec{e}_x, \vec{e}_y)$ = transverse plane $(\eta = 0)$





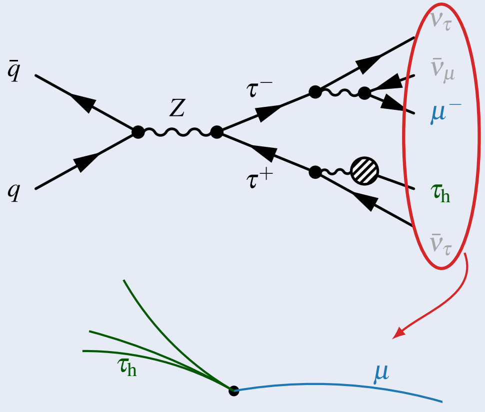
$$\sum_{\text{final state}} \vec{p}_\mathrm{T} = \vec{0} \Rightarrow \boxed{\vec{E}_\mathrm{T}^\mathrm{miss} = - \sum_{\text{visible particles}} \vec{p}_\mathrm{T}}$$

Event display: $h \to \tau\tau \to \mu\tau_{\mathrm{h}}$ candidate from real data

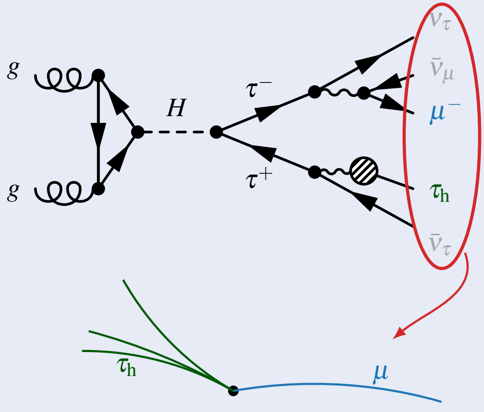


Backgrounds?

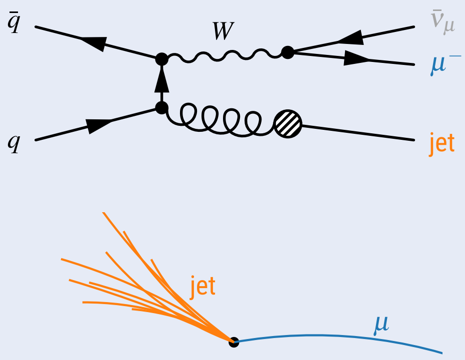
Drell-Yan (especially $Z \to \tau\tau$)



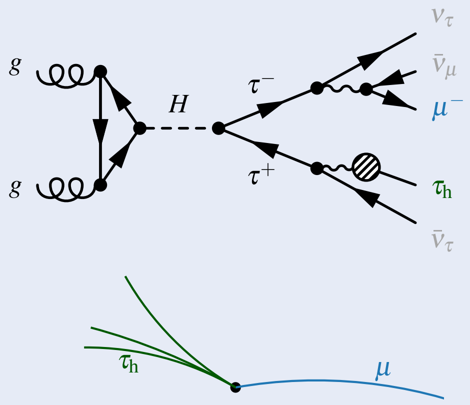
$H \to \tau\tau \to \mu\tau_\mathrm{h}$



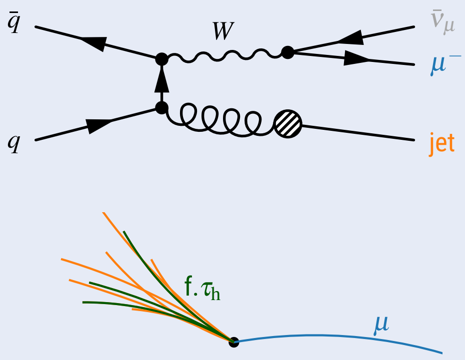
$W + \text{jets}$



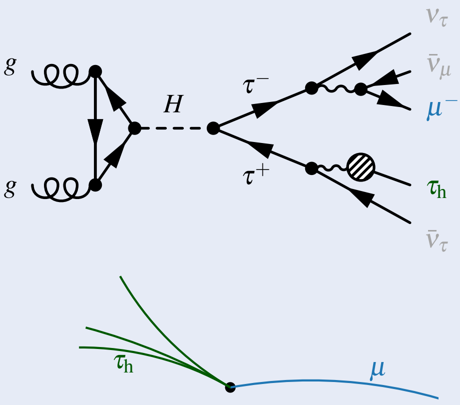
$H \to \tau\tau \to \mu\tau_h$



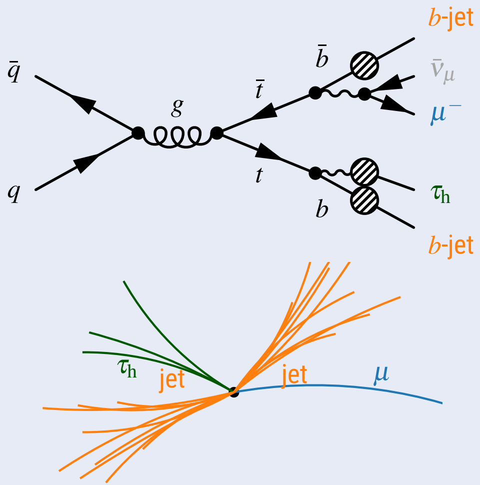
$W+$jets, jet $\to$ fake τ_h



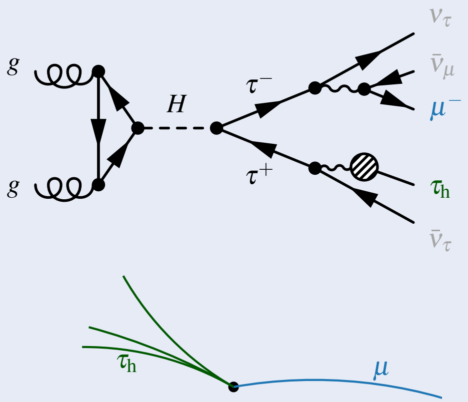
$H \to \tau\tau \to \mu\tau_\mathrm{h}$



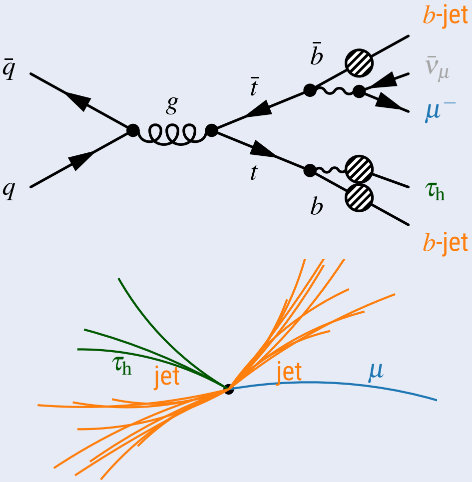
$t\bar{t}$



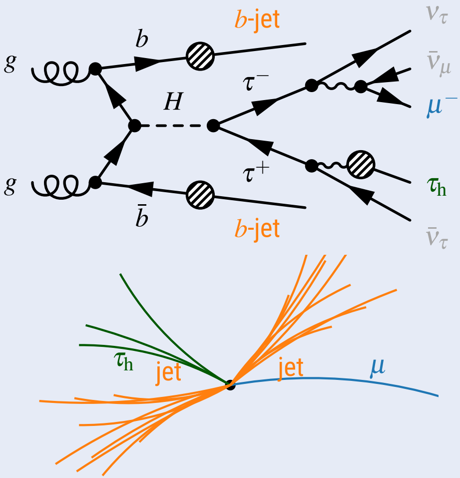
$H \to \tau\tau \to \mu\tau_h$



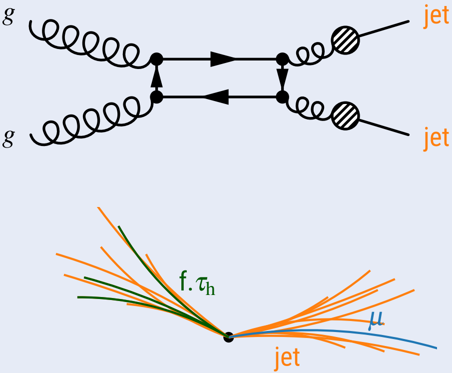
$t\bar{t}$



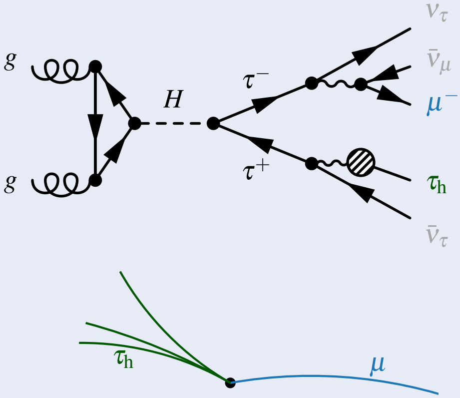
H production with b-jets



QCD

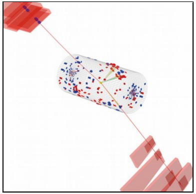


$H \to \tau\tau \to \mu\tau_h$



Embedded events

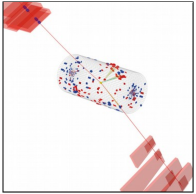
$Z \to \mu\mu$ data



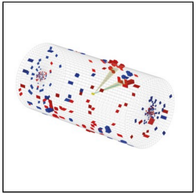
▷ The CMS Collaboration. "An embedding technique to determine $\tau\tau$ backgrounds in proton-proton collision data". *Journal of Instrumentation* **14**.06 (June 2019). DOI: 10.1088/1748-0221/14/06/p06032.

Embedded events

$Z \to \mu\mu$ data



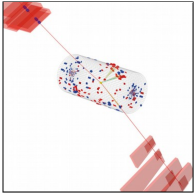
Remove $\mu\mu$ system



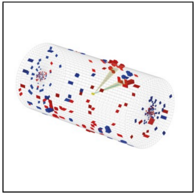
▷ The CMS Collaboration. "An embedding technique to determine $\tau\tau$ backgrounds in proton-proton collision data". *Journal of Instrumentation* **14**.06 (June 2019). DOI: 10.1088/1748-0221/14/06/p06032.

Embedded events

$Z \to \mu\mu$ data

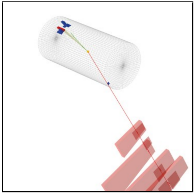


Remove $\mu\mu$ system



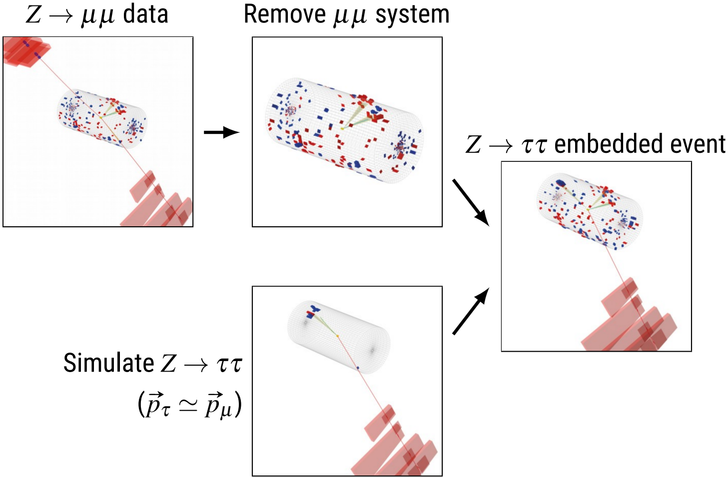
Simulate $Z \to \tau\tau$

$(\vec{p}_\tau \simeq \vec{p}_\mu)$



▷ The CMS Collaboration. “An embedding technique to determine $\tau\tau$ backgrounds in proton-proton collision data”. *Journal of Instrumentation* **14**.06 (June 2019). DOI: 10.1088/1748-0221/14/06/p06032.

Embedded events



▷ The CMS Collaboration. "An embedding technique to determine $\tau\tau$ backgrounds in proton-proton collision data". *Journal of Instrumentation* **14**.06 (June 2019). DOI: 10.1088/1748-0221/14/06/p06032.

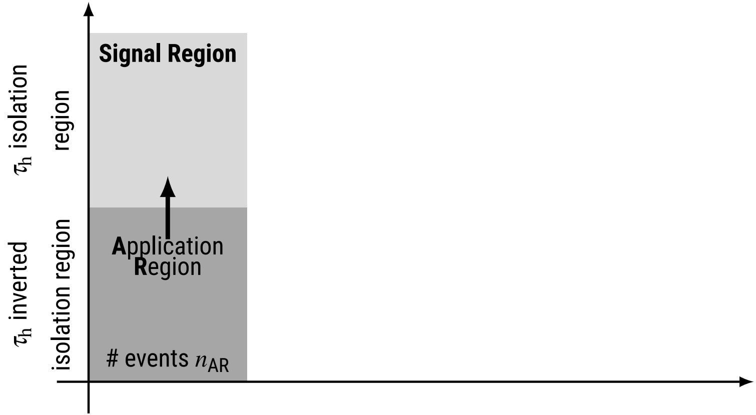
The Fake Factor method

- How many events contain misidentified τ_h? (fake taus)

▷ J. Andrejkovic & J. Bechtel. “Data-driven background estimation of fake-tau backgrounds in di-tau final states with the full Run-II dataset”. *CMS analysis Note* (June 2020). URL: `https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%5C%20AN-2019/170`.

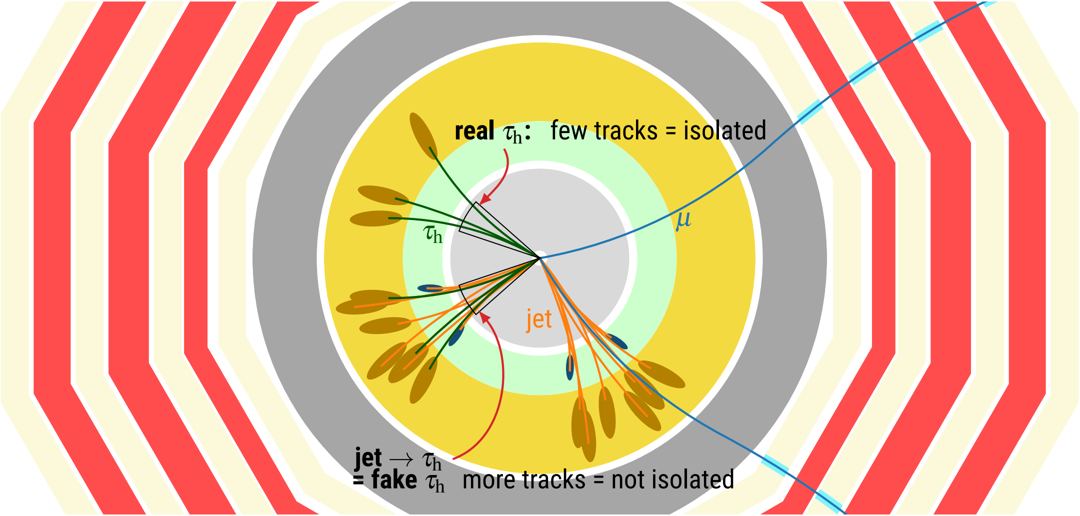
The Fake Factor method

- How many events contain misidentified τ_h? (fake taus)



▷ J. Andrejkovic & J. Bechtel. “Data-driven background estimation of fake-tau backgrounds in di-tau final states with the full Run-II dataset”. *CMS analysis Note* (June 2020). URL: `https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%5C%20AN-2019/170`.

Particles isolation – qualitatively



The FF method: determination regions definitions

$t\bar{t}$

Estimation from simulated samples, same selection as in SR.

▷ J. Andrejkovic & J. Bechtel. “Data-driven background estimation of fake-tau backgrounds in di-tau final states with the full Run-II dataset”. *CMS analysis Note* (June 2020). URL: https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%5C%20AN-2019/170.

The FF method: determination regions definitions

$t\bar{t}$

Estimation from simulated samples, same selection as in SR.

$W+$jets

Same as SR, except:

- transverse mass $m_T^{(\ell)} > 70\,\mathrm{GeV}$ ($m_T^{(\ell)} < 70\,\mathrm{GeV}$ in the SR);
- no b-jet (allowed in the SR).

▷ J. Andrejkovic & J. Bechtel. “Data-driven background estimation of fake-tau backgrounds in di-tau final states with the full Run-II dataset”. *CMS analysis Note* (June 2020). URL: `https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%5C%20AN-2019/170`.

The FF method: determination regions definitions

$t\bar{t}$

Estimation from simulated samples, same selection as in SR.

$W+$jets

Same as SR, except:

- transverse mass $m_T^{(\ell)} > 70\,\mathrm{GeV}$ ($m_T^{(\ell)} < 70\,\mathrm{GeV}$ in the SR);
- no b-jet (allowed in the SR).

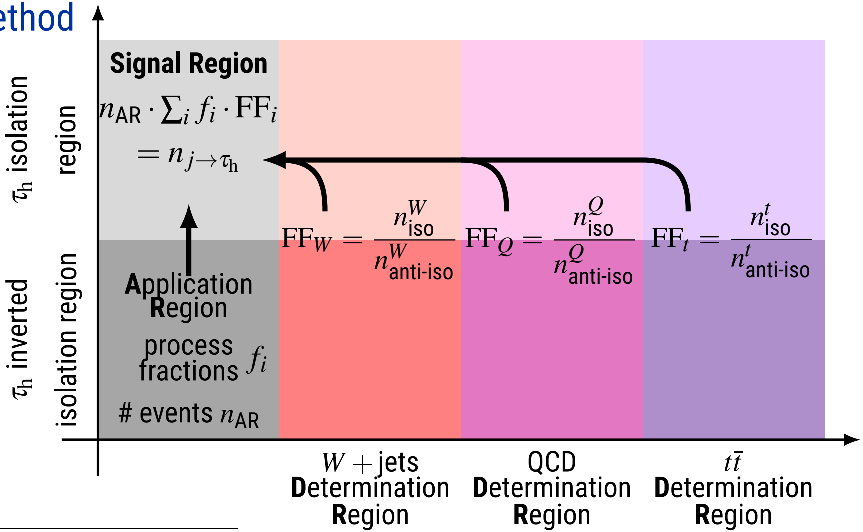
QCD multijet

Same as SR, except:

- same signs for L_1 and L_2 electric charges (opposite signs in the SR).

▷ J. Andrejkovic & J. Bechtel. “Data-driven background estimation of fake-tau backgrounds in di-tau final states with the full Run-II dataset”. *CMS analysis Note* (June 2020). URL: https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%5C%20AN-2019/170.

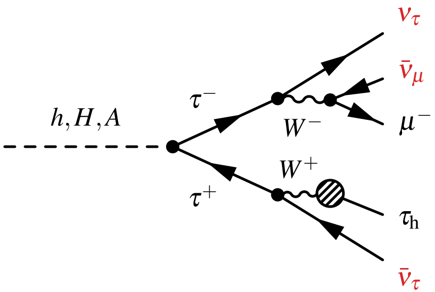
The FF method



▷ J. Andrejkovic & J. Bechtel. "Data-driven background estimation of fake-tau backgrounds in di-tau final states with the full Run-II dataset". *CMS analysis Note* (June 2020). URL: https://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%5C%20AN-2019/170.

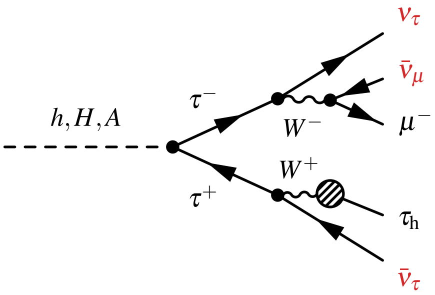
Discriminant variable?

- $E_{\mathrm{T}}^{\text{miss}}$ due to neutrinos.
  - No invariant mass!

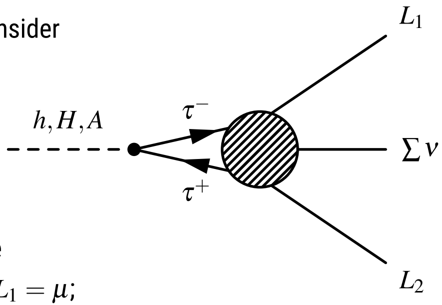


Discriminant variable?

- E_T^{miss} due to neutrinos.
  - No invariant mass!



- Consider



where

- $L_1 = \mu$;
- $L_2 = \tau_h$;
- $\sum \nu \simeq E_T^{\text{miss}}$;

with respect to the left side.

Discriminant variable: $m_{\mathrm{T}}^{\mathrm{tot}}$

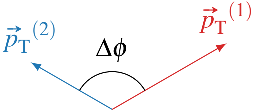
- For L_1, L_2 and ${E_{\mathrm{T}}}^{\text{miss}}$ system,
  - in the transverse plane (use ${E_{\mathrm{T}}}^{\text{miss}}$),
  - for $E_i \gg m_i$ (highly relativistic case),

deriving the "invariant" mass would then lead to

the **total transverse mass**, $m_{\mathrm{T}}^{\mathrm{tot}}$

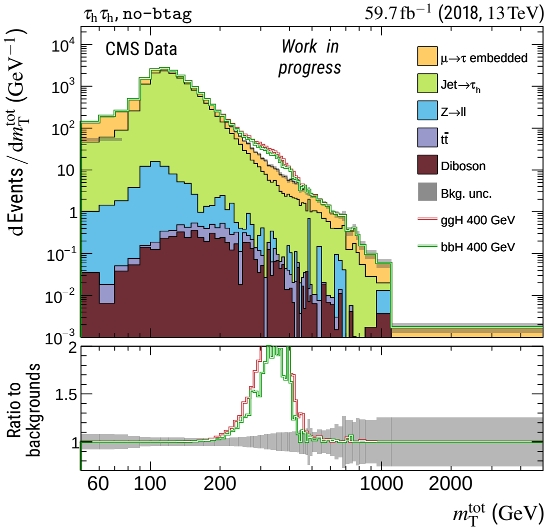
$$\boxed{m_{\mathrm{T}}^{\mathrm{tot}} = \sqrt{m_{\mathrm{T}}^2(L_1, {E_{\mathrm{T}}}^{\text{miss}}) + m_{\mathrm{T}}^2(L_2, {E_{\mathrm{T}}}^{\text{miss}}) + m_{\mathrm{T}}^2(L_1, L_2)}}$$

$$m_{\mathrm{T}}(1,2) = \sqrt{2 p_T^{(1)} p_T^{(2)} (1 - \cos\Delta\phi)}$$



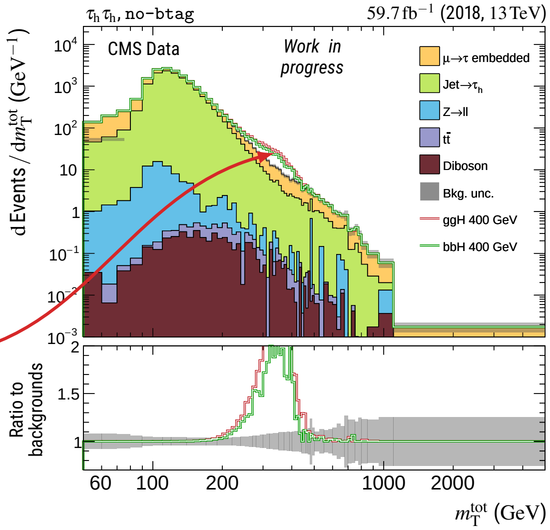
m_T^{tot} distributions

- Backgrounds = SM expectations:
  - DY $Z \to \tau\tau$ and some $t\bar{t}$ in $\mu \to \tau$ embedded
  - QCD, $W +$ jets and some $t\bar{t}$ in Jet $\to \tau_h$
  - $Z \to ee$ + $Z \to \mu\mu$ in $Z \to \ell\ell$
  - Remaining $t\bar{t}$ in $t\bar{t}$
  - Other small backgrounds in Diboson



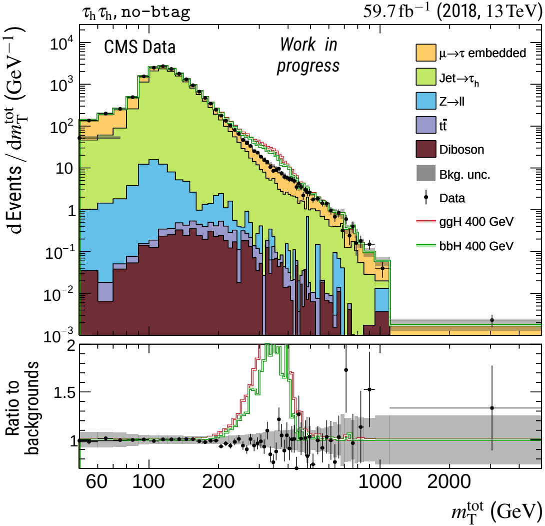
m_T^{tot} distributions

- Backgrounds = SM expectations:
  - DY $Z \to \tau\tau$ and some $t\bar{t}$ in $\mu \to \tau$ embedded
  - QCD, $W+$jets and some $t\bar{t}$ in Jet $\to \tau_h$
  - $Z \to ee$ + $Z \to \mu\mu$ in $Z \to \ell\ell$
  - Remaining $t\bar{t}$ in $t\bar{t}$
  - Other small backgrounds in Diboson
- H at 400 GeV expected $\sigma \times \mathcal{BR} = 1$ pb signal.



m_T^{tot} distributions

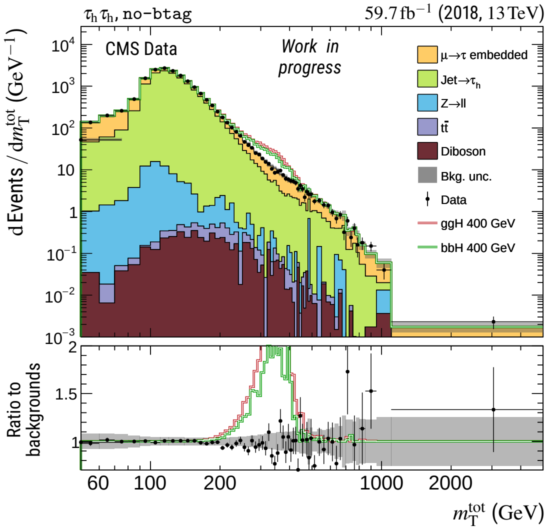
- Backgrounds = SM expectations:
  - DY $Z \to \tau\tau$ and some $t\bar{t}$ in $\mu \to \tau$ embedded
  - QCD, $W+$jets and some $t\bar{t}$ in Jet $\to \tau_h$
  - $Z \to ee$ + $Z \to \mu\mu$ in $Z \to \ell\ell$
  - Remaining $t\bar{t}$ in $t\bar{t}$
  - Other small backgrounds in Diboson
- H at $400\,\mathrm{GeV}$ expected $\sigma \times \mathcal{BR} = 1\,\mathrm{pb}$ signal.
- Compare to observed events (black dots).



m_T^{tot} distributions

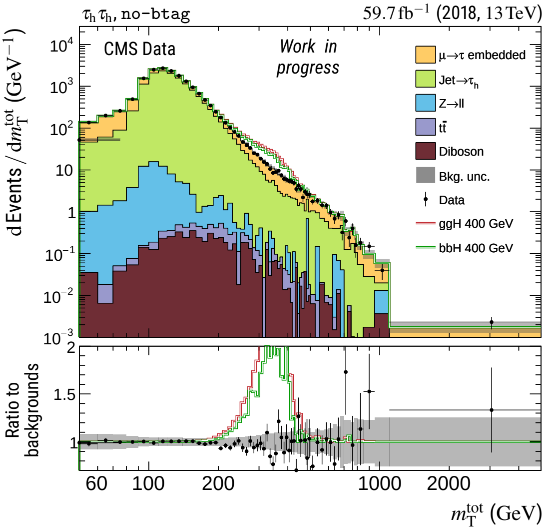
Not *just* a plot!

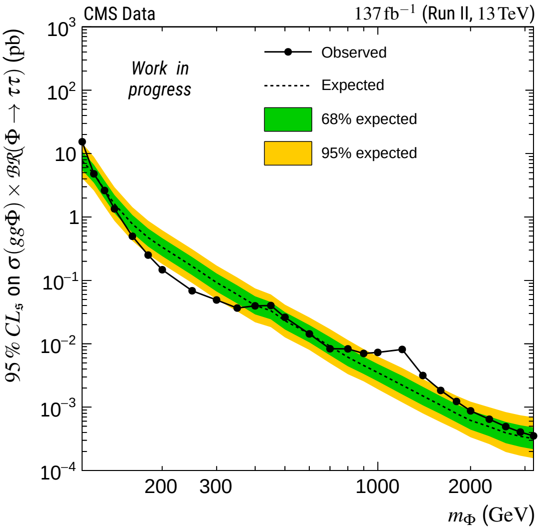
- Lot of hard work to obtain this:
  - simulated events
  - detector issues
  - uncertainties measured
- Collaborative work:
  - Karlsruhe Institute of Technology (DE)
  - Imperial College (UK)
  - DESY (DE)
  - IP2I (FR)

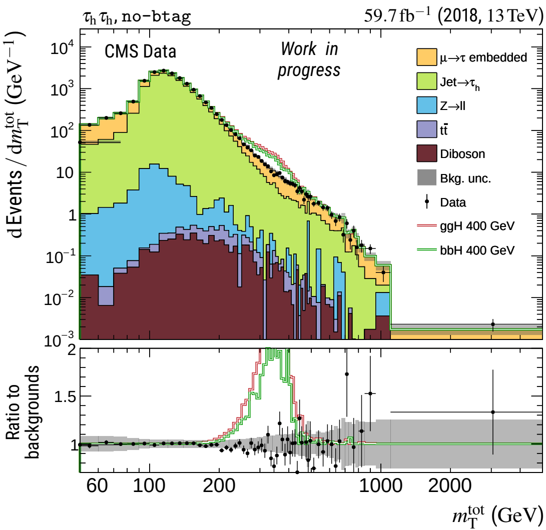


m_T^{tot} distributions

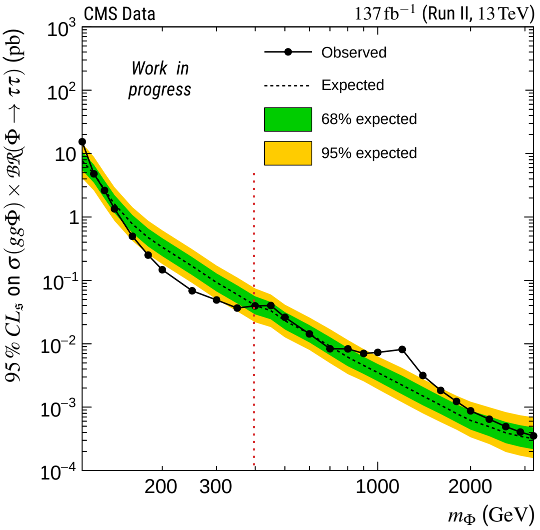
- Backgrounds = SM expectations:
  - DY $Z \to \tau\tau$ and some $t\bar{t}$ in $\mu \to \tau$ embedded
  - QCD, $W +$jets and some $t\bar{t}$ in Jet $\to \tau_h$
  - $Z \to ee$ + $Z \to \mu\mu$ in $Z \to \ell\ell$
  - Remaining $t\bar{t}$ in $t\bar{t}$
  - Other small backgrounds in Diboson
- H at $400\,\mathrm{GeV}$ expected $\sigma \times \mathcal{BR} = 1\,\mathrm{pb}$ signal.
- Compare to observed events (black dots).
- Data/Bkg agreement → **exclusion limits** on $\sigma \times \mathcal{BR}$

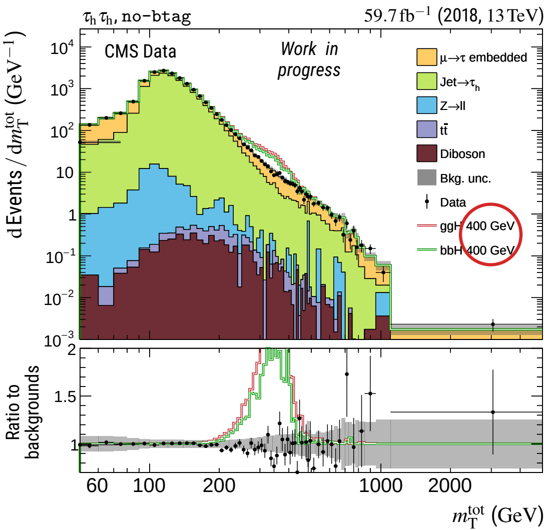




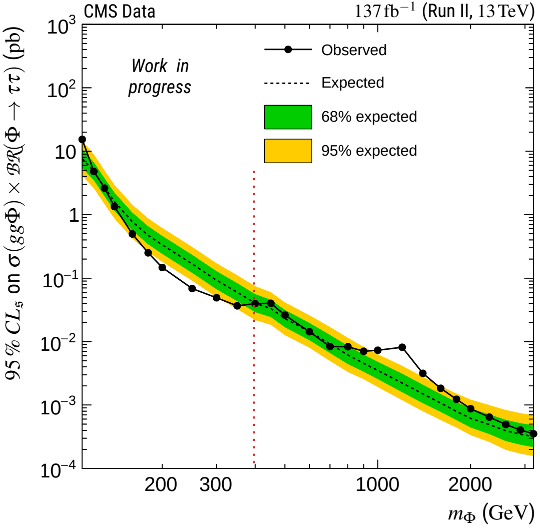


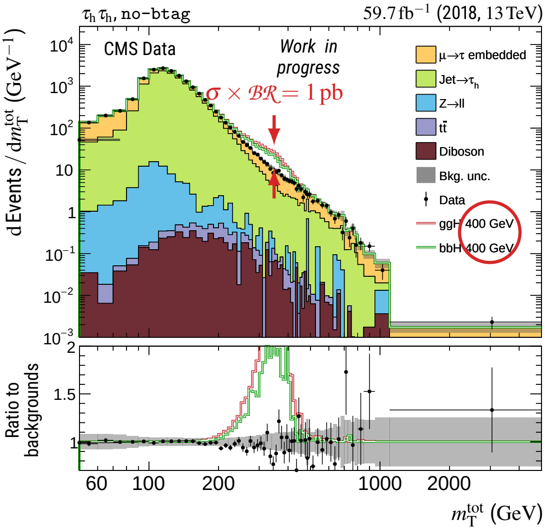
▷ A. L. Read. "Modified frequentist analysis of search results (the CL_s method)". *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. CERN-OPEN-2000-205. May 2000. URL: http://cds.cern.ch/record/451614.



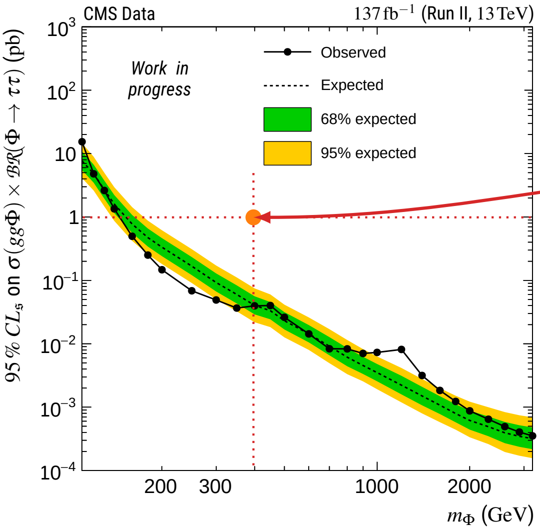


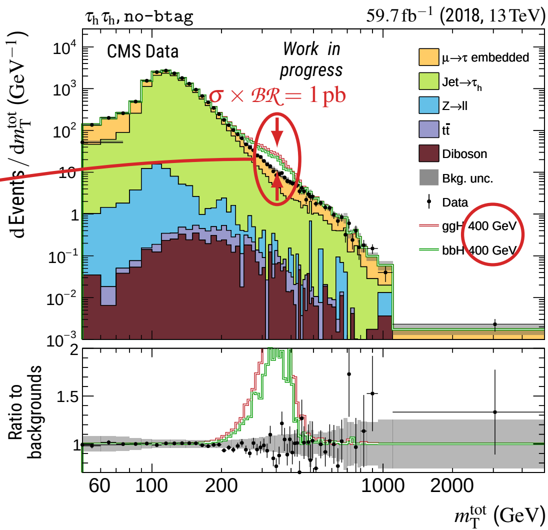
▷ A. L. Read. "Modified frequentist analysis of search results (the CL_s method)". *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. CERN-OPEN-2000-205. May 2000. URL: http://cds.cern.ch/record/451614.



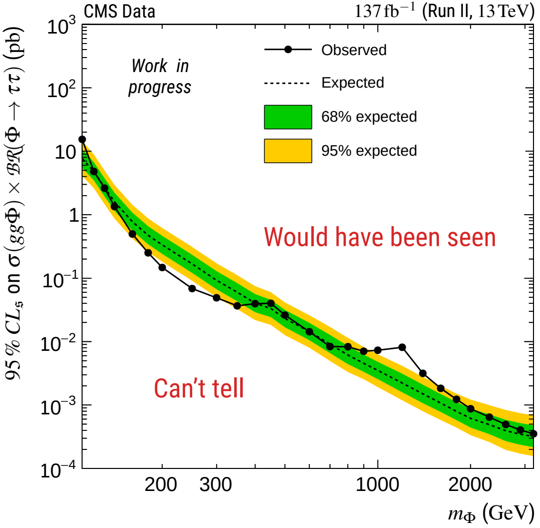


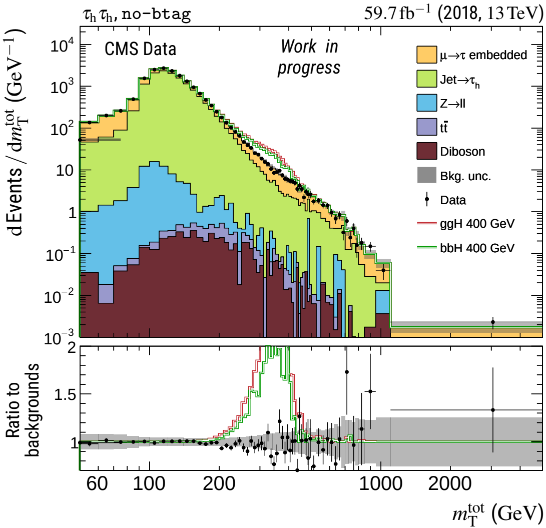
▷ A. L. Read. "Modified frequentist analysis of search results (the CL_s method)". *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. CERN-OPEN-2000-205. May 2000. URL: http://cds.cern.ch/record/451614.



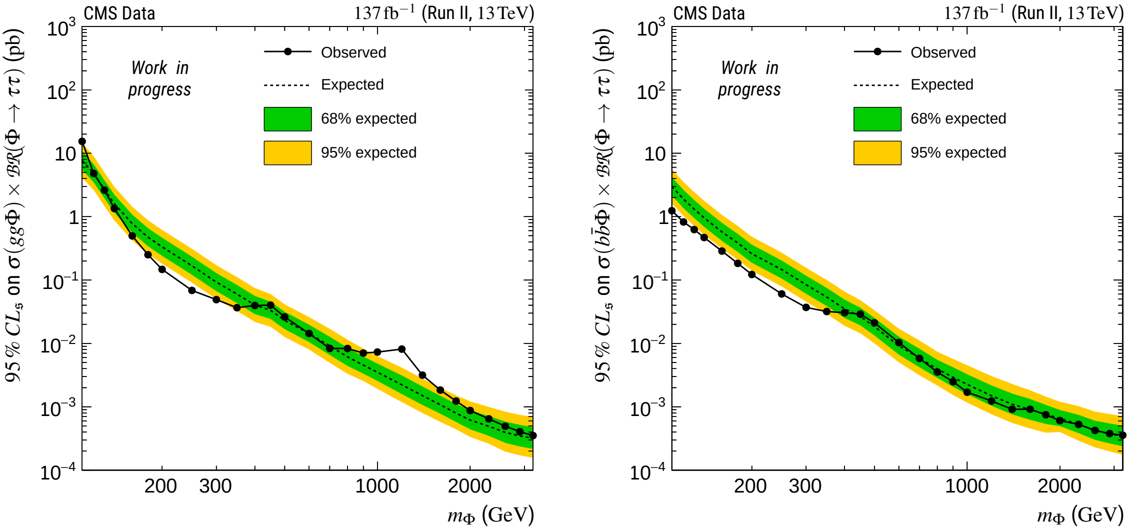


▷ A. L. Read. "Modified frequentist analysis of search results (the CL_s method)". *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. CERN-OPEN-2000-205. May 2000. URL: http://cds.cern.ch/record/451614.

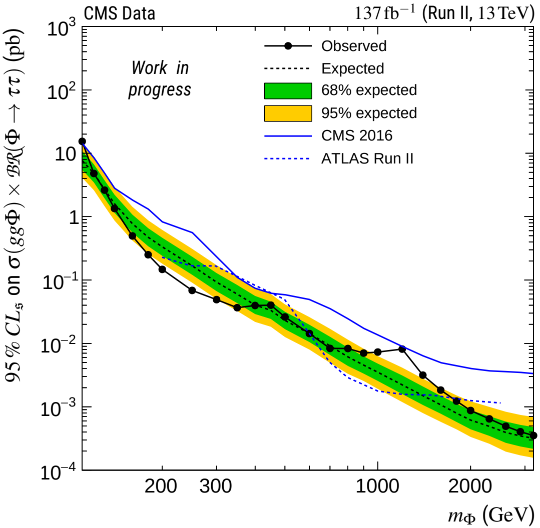


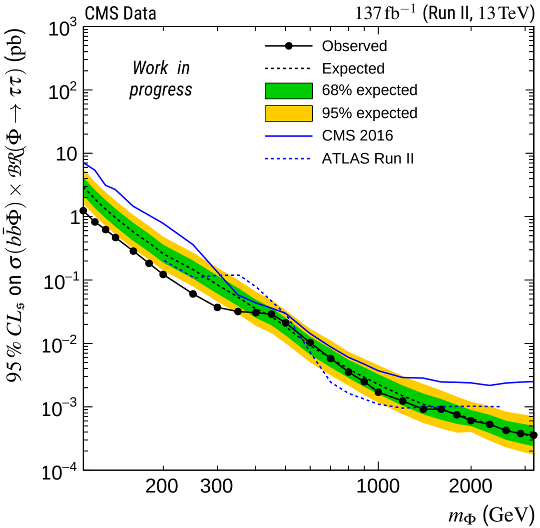


▷ A. L. Read. "Modified frequentist analysis of search results (the CL_s method)". *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. CERN-OPEN-2000-205. May 2000. URL: http://cds.cern.ch/record/451614.



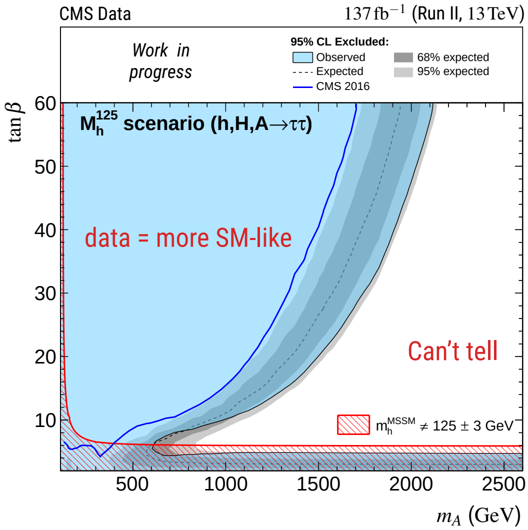
▷ A. L. Read. "Modified frequentist analysis of search results (the CL_s method)". *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. CERN-OPEN-2000-205. May 2000. URL: http://cds.cern.ch/record/451614.

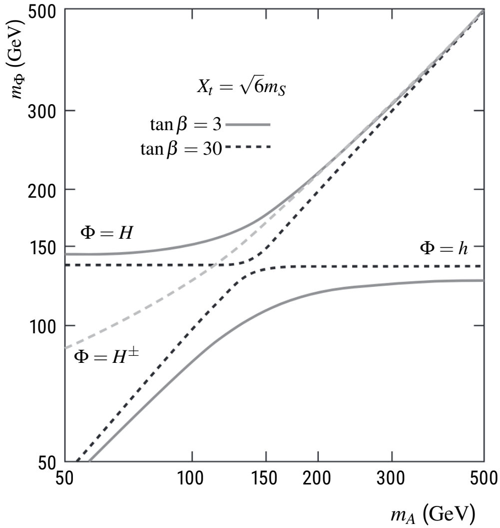




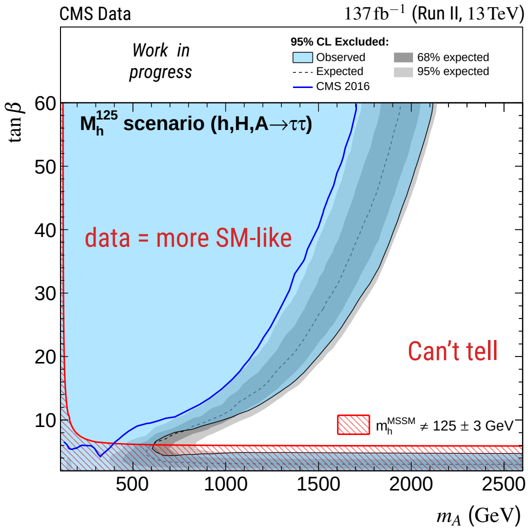
▷ A. L. Read. “Modified frequentist analysis of search results (the CL_s method)”. *Workshop on confidence limits, CERN, Geneva, Switzerland, 17-18 Jan 2000: Proceedings*. CERN-OPEN-2000-205. May 2000. URL: http://cds.cern.ch/record/451614.

- Model dependent limits:
  - Fix high-order MSSM parameter,
  - Explore $(m_A, \tan\beta)$ plane,
  - Do data stick more to SM or MSSM?





$X_t = \sqrt{6} m_S$
$\tan\beta = 3$
$\tan\beta = 30$
$\Phi = H$
$\Phi = h$
$\Phi = H^\pm$
m_Φ (GeV)
m_A (GeV)



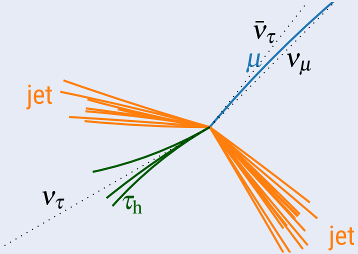
CMS Data
$137\,\mathrm{fb}^{-1}$ (Run II, 13 TeV)
Work in progress
95% CL Excluded:
Observed
Expected
CMS 2016
68% expected
95% expected
M_h^{125} scenario (h,H,A→ττ)
data = more SM-like
Can't tell
$m_h^{MSSM} \neq 125 \pm 3$ GeV
$\tan\beta$
m_A (GeV)

- Remember: invariant mass not fully available:
  - neutrinos in di-τ events.

- Remember: invariant mass not fully available:
  - neutrinos in di-τ events.

What's here

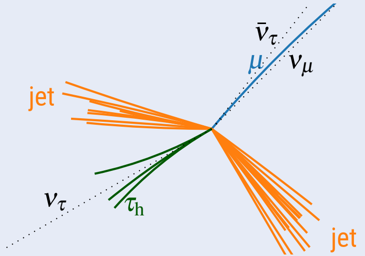
(e.g. VBF Higgs production + decay to $\tau\tau$, $\mu\tau_\mathrm{h}$ channel)



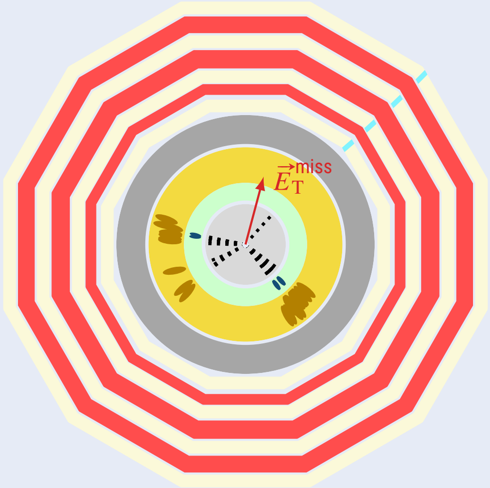
- Remember: invariant mass not fully available:
  - neutrinos in di-τ events.

What's here

(e.g. VBF Higgs production + decay to $\tau\tau$, $\mu\tau_h$ channel)



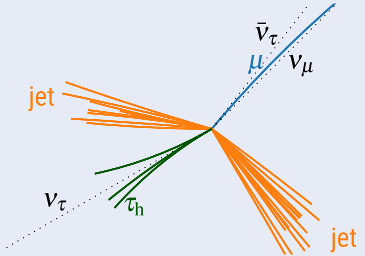
What CMS sees: no neutrinos but E_T^{miss}



- Remember: invariant mass not fully available:
  - neutrinos in di-τ events.

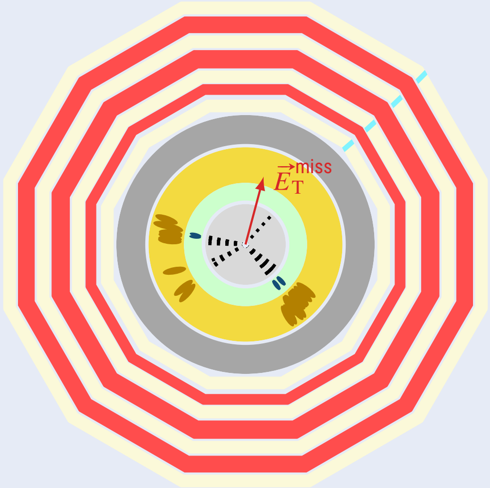
What's here

(e.g. VBF Higgs production + decay to $\tau\tau$, $\mu\tau_h$ channel)

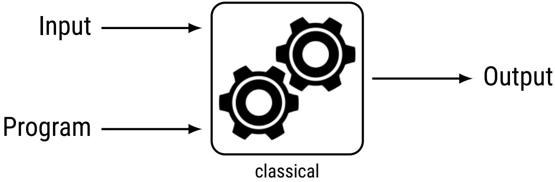


- It would be great to have a di-τ mass estimator!
  - What about ***machine learning***?

What CMS sees: no neutrinos but E_T^{miss}

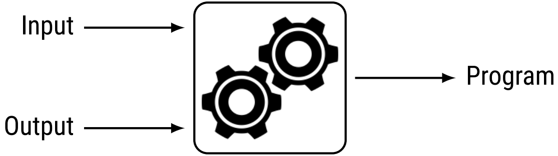


What is *machine learning*? – A brief introduction



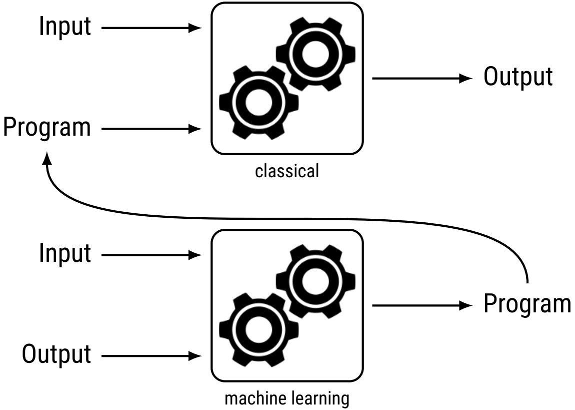
What is *machine learning*? – A brief introduction

classical



machine learning

What is *machine learning*? – A brief introduction



What is *machine learning*? – A brief introduction

Aim: find a function (program) mapping features (input) to a target (output)

What is *machine learning*? – A brief introduction

Aim: find a function (program) mapping features (input) to a target (output)

- Categorical target ⇒ Classification

e.g. cat or dog on the image

▷ C. Bernet. *The Data Frog – Image Recognition: Dogs vs Cats!* URL: https://thedatafrog.com/en/articles/dogs-vs-cats/.

What is *machine learning*? – A brief introduction

Aim: find a function (program) mapping features (input) to a target (output)

- Categorical target ⇒ Classification

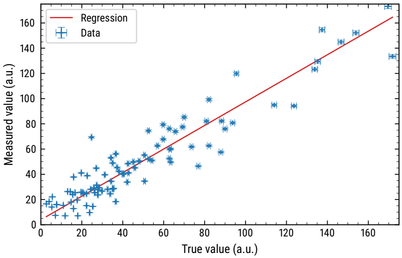
e.g. cat or dog on the image

▷ C. Bernet. *The Data Frog – Image Recognition: Dogs vs Cats!* URL: https://thedatafrog.com/en/articles/dogs-vs-cats/.

- Continuous target ⇒ Regression

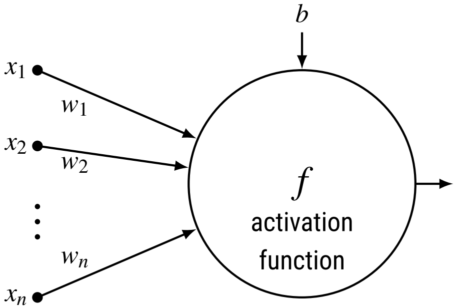
e.g. discriminating variable!

Linear case:



What if the target is not linear wrt. input?

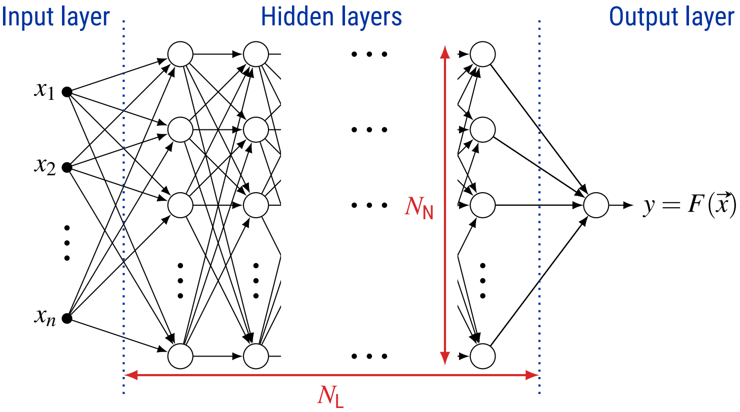
Neurons in ML



Output $= f\left(\sum_{i=1}^{n} w_i x_i + b\right)$

- Parameters: $w_1, w_2, \ldots, w_n, b$
- Equivalent to linear regression for $f = \mathbb{1} : x \mapsto x$

(Deep) Neural Networks

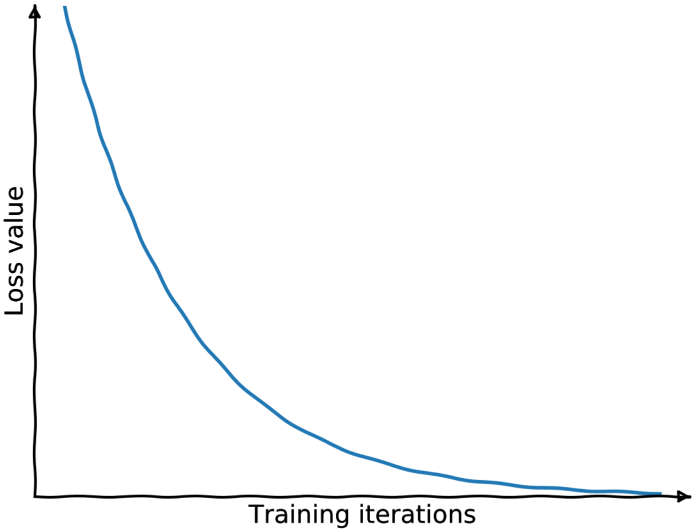


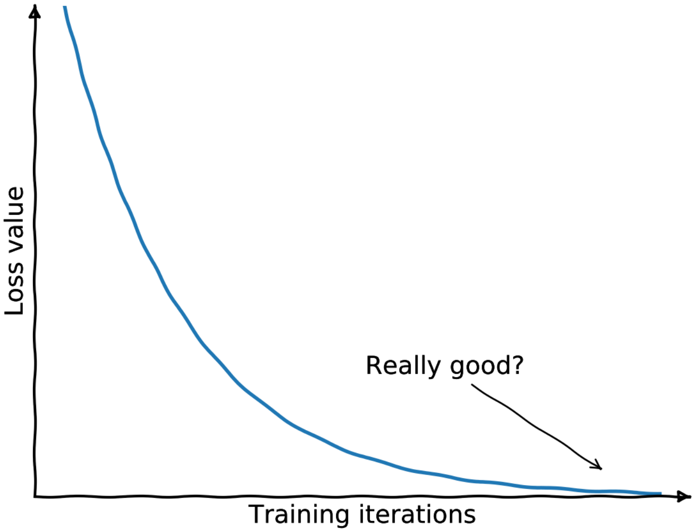
How to train a neural network?

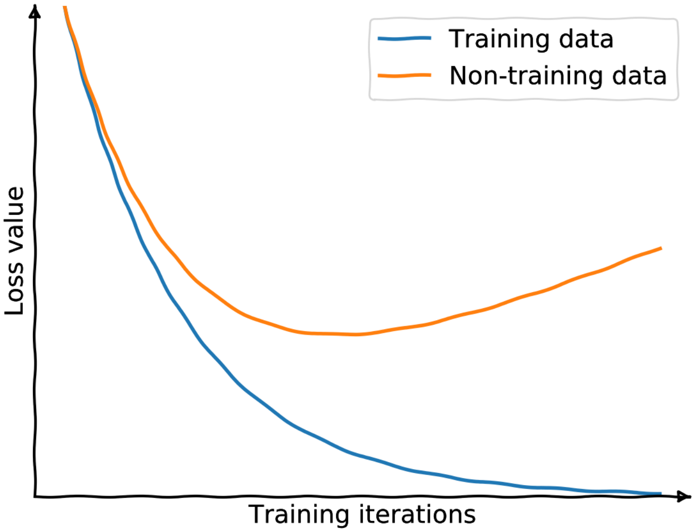
Train = optimize parameters $(w_1, w_2, \ldots, w_n, b)$ for each neuron.

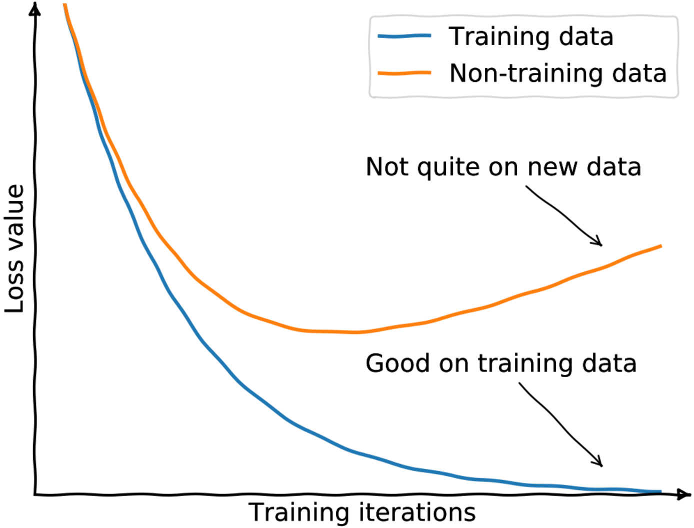
- Get a **training dataset** = examples of inputs $\vec{x}_i$ with corresponding outputs y_i
- Compare the model predictions $F(\vec{x}_i)$ to the true values y_i
  - Define a **loss function** $\mathfrak{L}$ such that its minimum is reached when $F(\vec{x}_i) = y_i$
  - Change the parameters a bit, aiming at minimizing $\mathfrak{L}(F(\vec{x}_i), y_i)$
  - Repeat

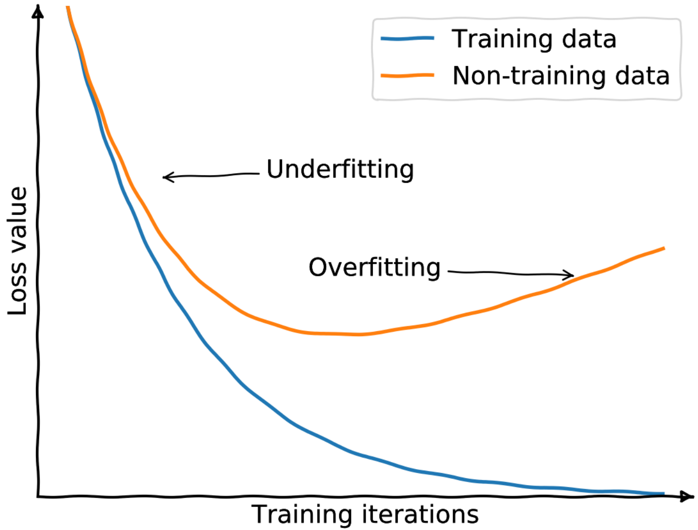
When to stop training?

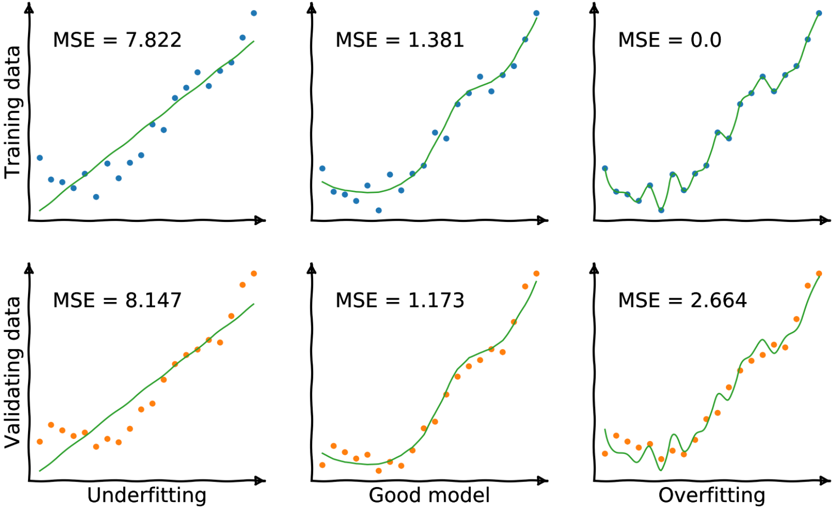

Loss value
Training iterations


Loss value
Really good?
Training iterations

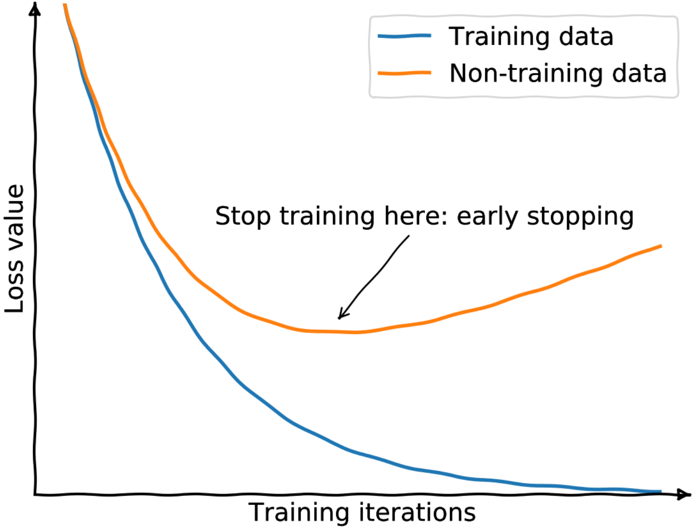

Training data
Non-training data
Loss value
Training iterations

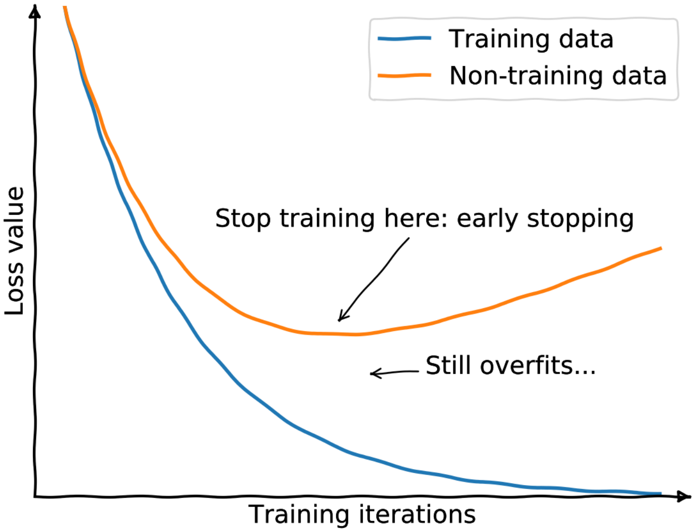

Training data
Non-training data
Not quite on new data
Good on training data
Loss value
Training iterations

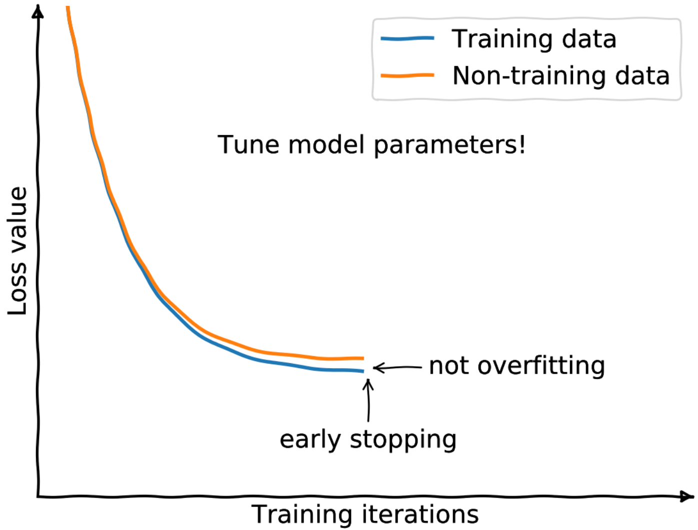

Training data
Non-training data
Underfitting
Overfitting
Loss value
Training iterations

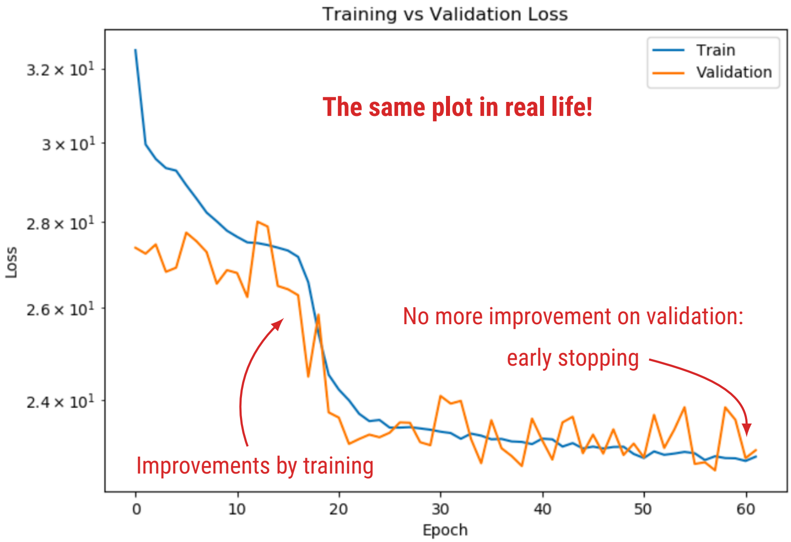


Training data
Validating data
MSE = 7.822
MSE = 1.381
MSE = 0.0
MSE = 8.147
MSE = 1.173
MSE = 2.664
Underfitting
Good model
Overfitting


Training data
Non-training data
Stop training here: early stopping
Loss value
Training iterations


Training data
Non-training data
Loss value
Stop training here: early stopping
Still overfits...
Training iterations

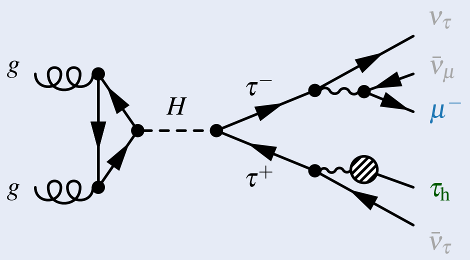

Training data
Non-training data
Tune model parameters!
Loss value
not overfitting
early stopping
Training iterations


Training vs Validation Loss
Train
Validation
The same plot in real life!
No more improvement on validation:
early stopping
Improvements by training
Loss
Epoch
3.2×10^1
3×10^1
2.8×10^1
2.6×10^1
2.4×10^1
0
10
20
30
40
50
60

Build a neural network: target and inputs?

- Model target: generated Higgs mass.
- Model inputs:
  - τ_1 (here = μ^-) and τ_2 (here = τ_h) p_T, η, ϕ;
  - PuppiMET p_T, ϕ;
  - METcov xx, xy and yy;
  - Number of neutrinos from *reco* tau decays;
  - $m_\mathrm{T}^{(1,MET)}$, $m_\mathrm{T}^{(2,MET)}$, $m_\mathrm{T}^{(1,2)}$, $m_\mathrm{T}^{\mathrm{tot}}$ (Puppi);
  - jet 1, jet 2 p_T, η, ϕ;
  - Additionnal Hadronic Activity $p_\mathrm{T}, \eta, \phi, N_{\text{jets}}^{\text{AHA}}$;
  - $\mathtt{npvsGood} \to$ how much PU.

$gg \to H \to \tau\tau \to \mu\tau_\mathrm{h}$



$$\boxed{m_\mathrm{T}^{\mathrm{tot}} = \sqrt{m_\mathrm{T}^2(\tau_1, E_\mathrm{T}^{\text{miss}}) + m_\mathrm{T}^2(\tau_2, E_\mathrm{T}^{\text{miss}}) + m_\mathrm{T}^2(\tau_1, \tau_2)}}, \quad m_\mathrm{T}(1,2) = \sqrt{2 p_T^{(1)} p_T^{(2)} (1 - \cos\Delta\phi)}$$

Build a neural network: hyperparameters?

- NN hyperparameters (and other tested values):
  - **Adam** optimizer (Adadelta, SGD),
  - Weight initialized with **Glorot uniform** (Glorot normal, normal, uniform),
  - Custom $\mathfrak{L}_{\mathrm{MA}\sqrt{\mathrm{PE}}\times\mathrm{b}}$ loss ($\mathfrak{L}_{\mathrm{MAPE}}$, $\mathfrak{L}_{\mathrm{MAE}}$, $\mathfrak{L}_{\mathrm{MSE}}$),
  - **Softplus** activation function (ReLU, ELU, SELU, Exponential),

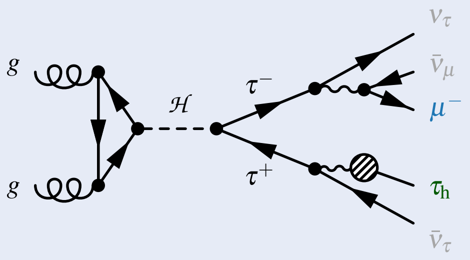
$$\mathit{softplus}(x) = \ln(1 + e^x)$$

  - 3 hidden layers (2 to 5),
  - 1000 neurons per hidden layer (200 to 2000 per steps of 100).

Datasets?

- Generate $\mathcal{H} \to \tau\tau$ events:
  - $\mathcal{H}$ is SM Higgs with a different mass,
  - $\mathcal{H}$ produced by gluon fusion,
  - set $\mathcal{BR}(\mathcal{H} \to \tau\tau) = 1$ to avoid non di-τ events.
- All final states used simultaneously for training:
  - $\tau_h\tau_h$, $\mu\tau_h$, $e\tau_h$, $\mu\mu$, $e\mu$, ee.

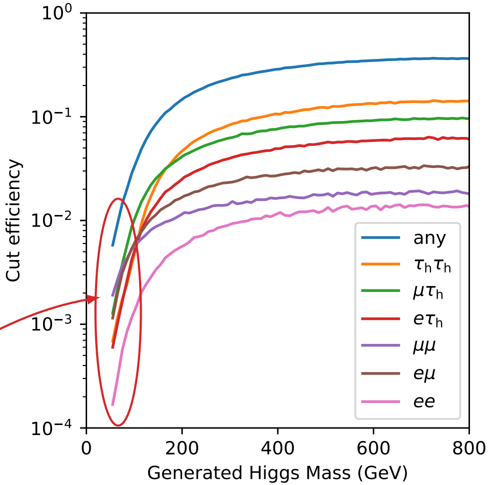
$gg \to \mathcal{H} \to \tau\tau \to \mu\tau_h$



Datasets: low mass boundary

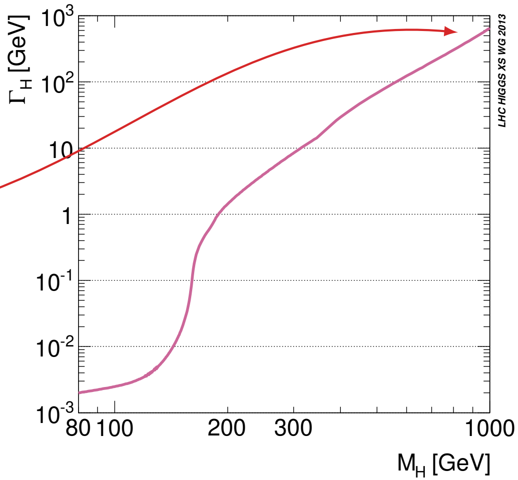
- Event selection:
  - Same as in the MSSM $H/A \to \tau\tau$ analysis,
  - Add $\mu\mu$ and ee channels.
- Events amount drops at low mass:
  - due to p_{T} cuts,
  - go down to $50\,\mathrm{GeV}$ for $m_{\mathcal{H}}$.

Less than $1\,\%$ pass the selection



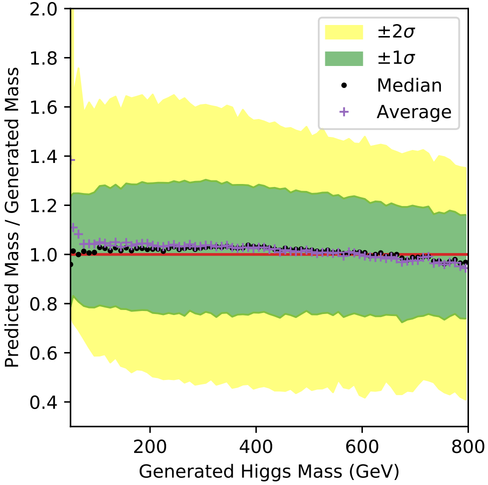
Datasets: high mass boundary

- Higgs width $\simeq$ Higgs mass around $1\,\mathrm{TeV}$:
  - can't have coherent mass points,
  - go up to $800\,\mathrm{GeV}$ for $m_{\mathcal{H}}$.



▷ LHC Higgs Cross Section Working Group. "Higgs Properties". *Handbook of LHC Higgs Cross Sections*. **3**. CERN Yellow Reports: Monographs. Geneva: CERN, 2013. URL: https://cds.cern.ch/record/1559921.

Model's performances

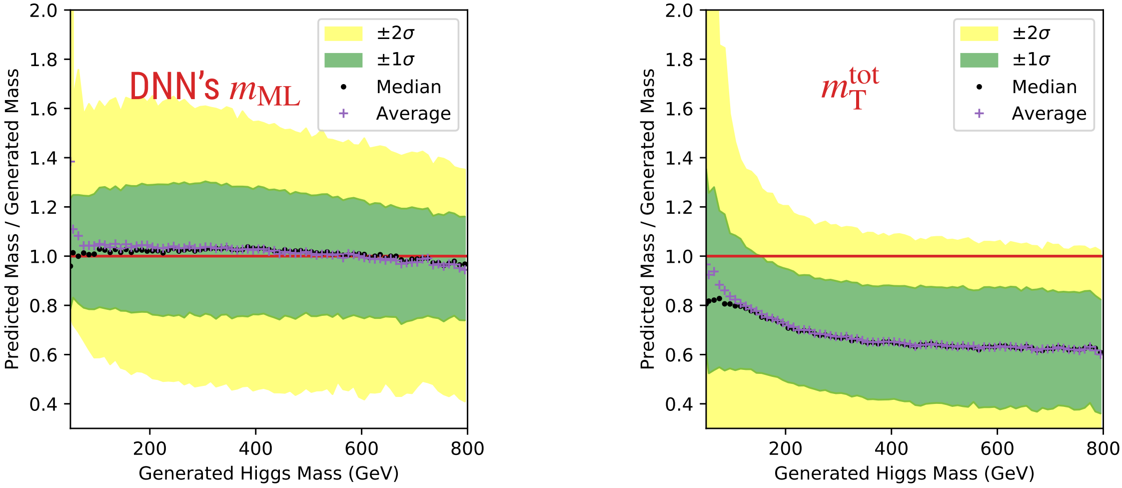


- Model's response:

$$r = \frac{\text{prediction}}{\text{true value}} = \frac{m_{\mathrm{ML}}}{m_{\mathcal{H}}}$$

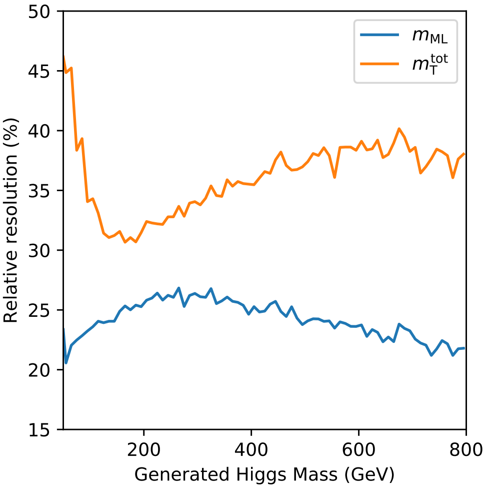
  - Closer to 1 is better.
  - $r = 1.00 \pm 0.05$ from 80 to $800\,\mathrm{GeV}$
  - $\mathcal{H}$ mass reconstruction **achieved** ✓

DNN's m_{ML} predictions *vs* $m_{\mathrm{T}}^{\mathrm{tot}}$

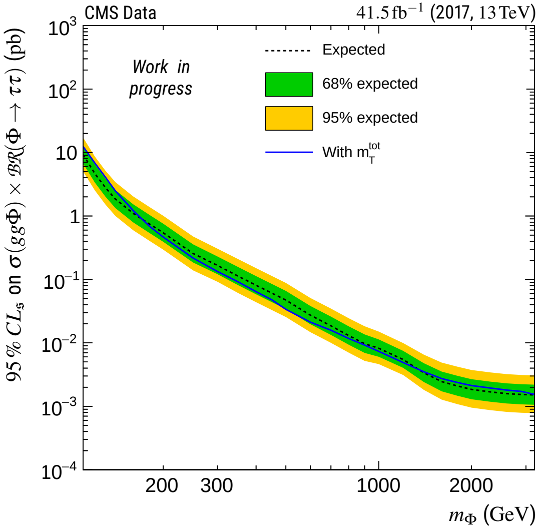


Using the model to get a discriminating variable

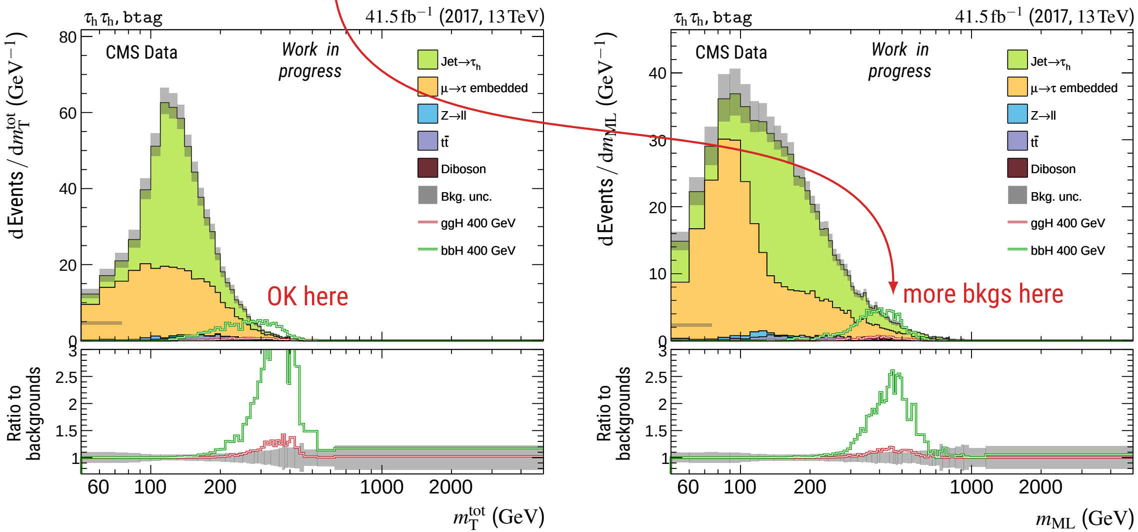
- In the $H/A \to \tau\tau$ analysis, discriminating variable $= m_{\mathrm{T}}^{\mathrm{tot}}$.
- $m_{\mathrm{T}}^{\mathrm{tot}}$ is equal to the invariant mass assuming:
  - all neutrinos are a single particle with $\vec{p}_{\mathrm{T}} = \vec{E}_{\mathrm{T}}^{\mathrm{miss}}$,
  - all is going on in the transverse plane (any $p_z = 0$).
- **Our model has a better resolution** on $m_{\mathcal{H}}$ than $m_{\mathrm{T}}^{\mathrm{tot}}$.



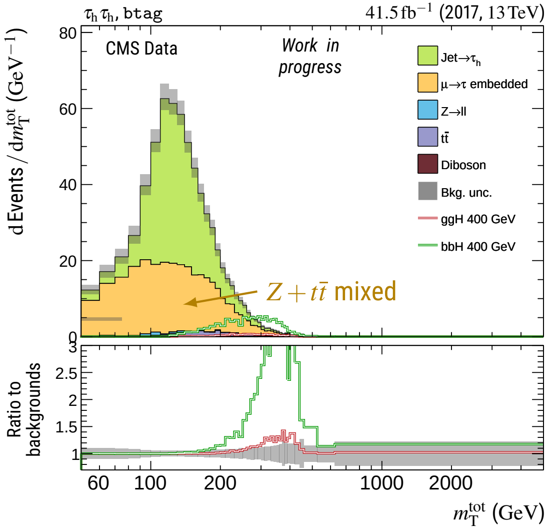
- Proceed to the search for massive Higgs boson Φ with di-τ events on the 2017 era.
- Use m_{ML} as discriminating variable.
- Not really better than with $m_{\mathrm{T}}^{\mathrm{tot}}$... Why?

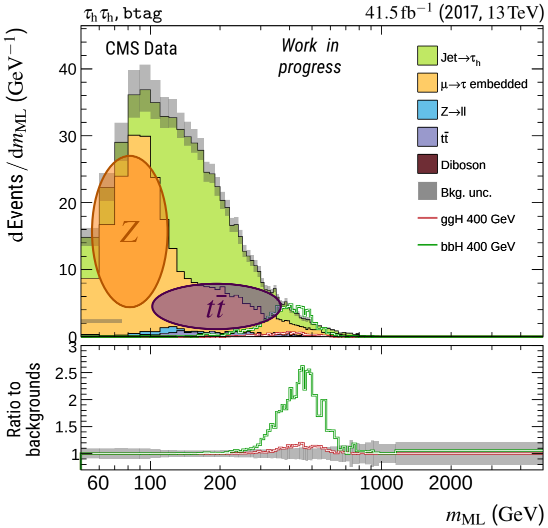


- Large fakes τ_h high mass tails falling into the signal region $\Rightarrow$ lowered signal to background ratio.

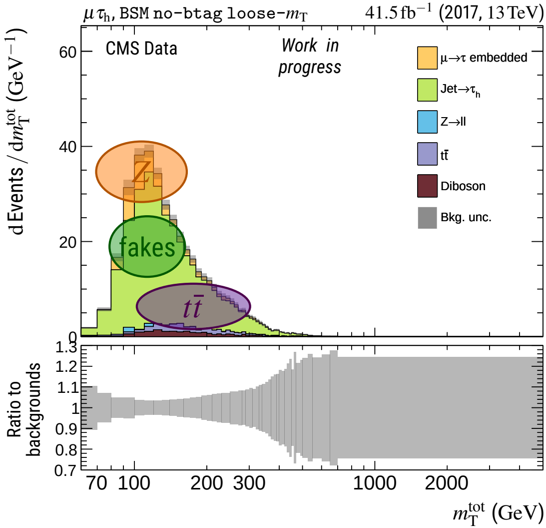


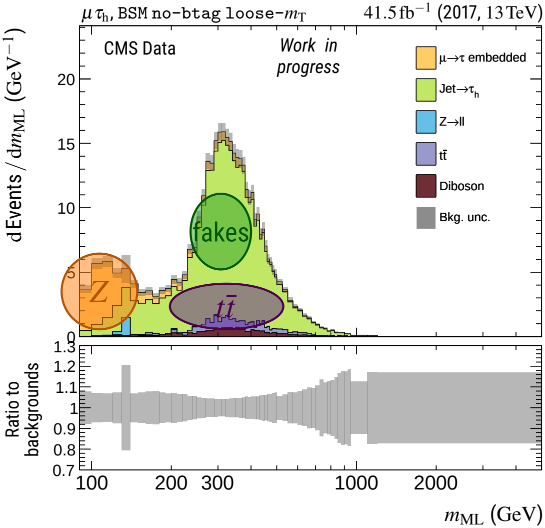
▶ **BUT** Z / $t\bar{t}$ separation! See the two m_{ML} embedded components, not present with $m_{\mathrm{T}}^{\mathrm{tot}}$.





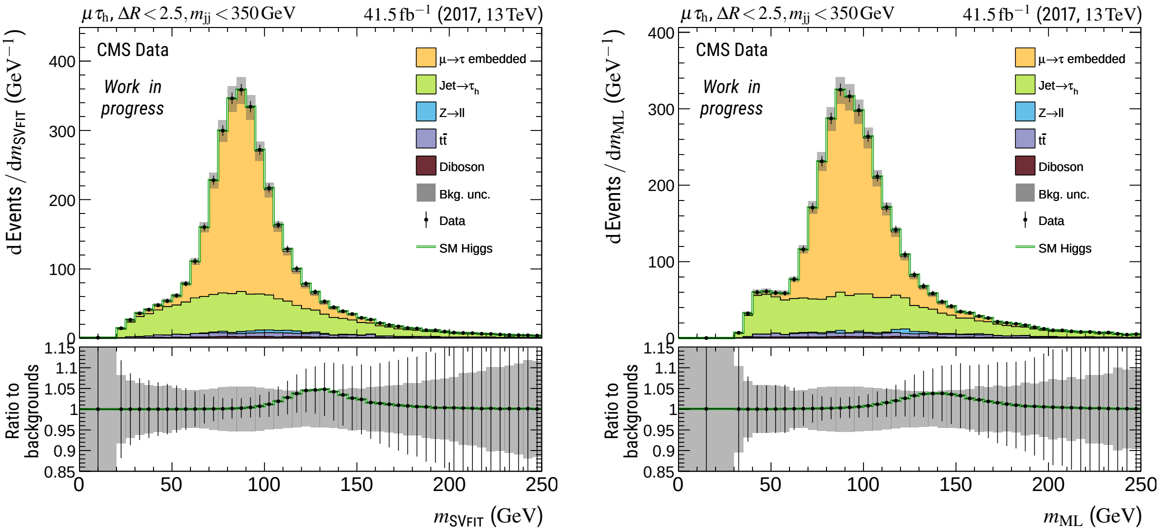
▶ Visible in other categories too. But here, the SVFIT mass is above $250\,\mathrm{GeV}$ and our model sees a Z signal!



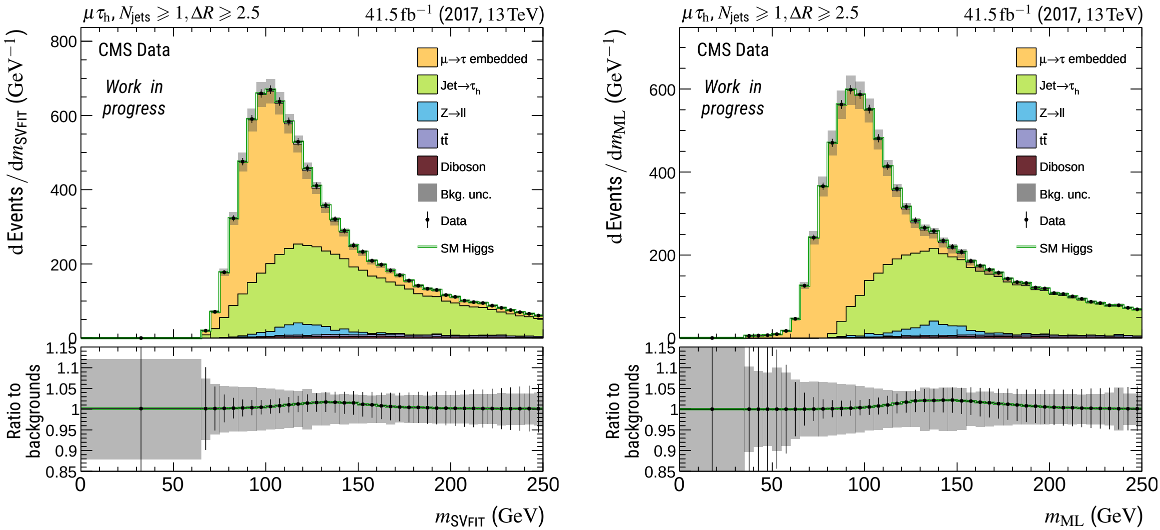


Let's compare the SVFIT predictions m_{SVFIT} to our model's ones m_{ML}.

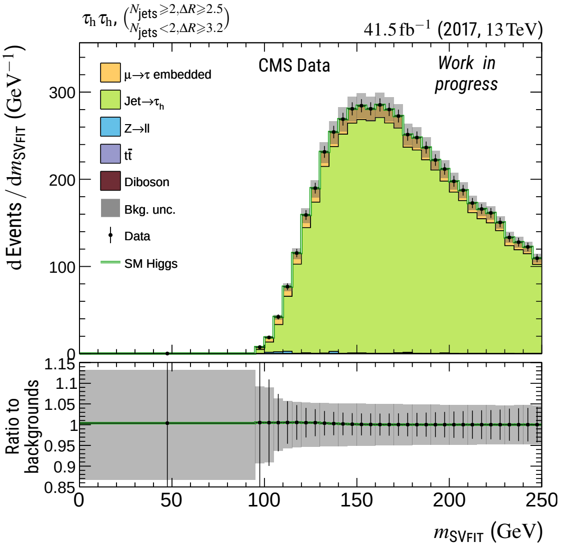
- **Similar SM Higgs signal sensitivity**, small (expected) overestimation from our model.

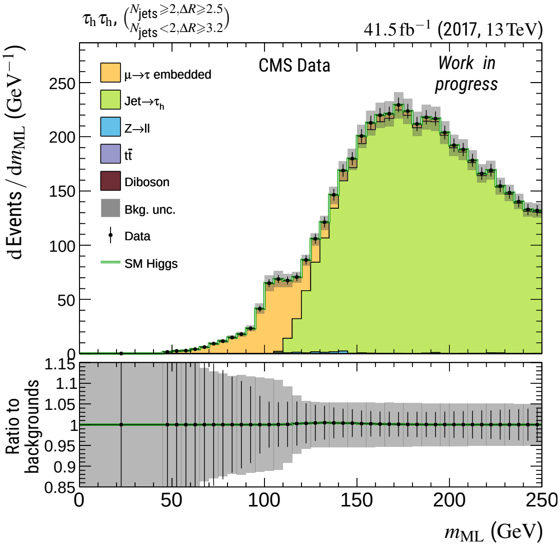


▶ **Better DY estimations** (peak at $100\,\mathrm{GeV}$ for m_{SVFIT}, $92\,\mathrm{GeV}$ for m_{ML}) and fakes at higher masses!



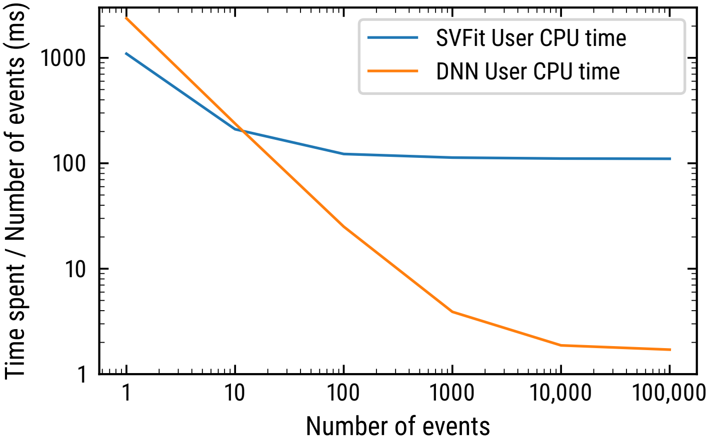
- Our model finds $Z \to \tau\tau$ events when SVFIT does not!





- Computing time: DNN (Python) is $\sim 60\times$ faster than SVFIT (C++)!

- SVFIT:
  - fit to find the best mass
  - for each event
- DNN:
  - fit done once (training)
  - apply the DNN formula



Conclusion & prospects: $H/A \to \tau\tau$

TODO

Conclusion & prospects: ML project

- Successful $m_{\mathcal{H}}$ reconstruction in di-τ events.
  - Not only MSSM $H/A \rightarrow \tau\tau$ but any $X \rightarrow \tau\tau$ analysis could benefit.
- m_{ML} vs $m_{\mathrm{T}}^{\mathrm{tot}}$:
  - A good mass estimator is not equivalent to a good discriminating variable.
  - Still, we already have the same performances at this point.
- m_{ML} vs m_{SVFIT}:
  - Similar Higgs sensitivity for some event topologies.
  - Better Z estimation observed (the model has been trained on $\mathcal{H} \rightarrow \tau\tau$ with various masses only).
  - Faster (about 60 times!).
  - Could be improved by updating the training datasets (other kinds of events).
  - Very promising as a SVFIT successor.

Merci!

Thank you for your attention!

JERC

custom loss