Particle Physics II
(LPHY2133)

Andrea Giammanco, UCL

Last week's homework

Channel	4e	4μ	$2e2\mu$	Total
ZZ background	2.7 ± 0.3	5.7 ± 0.6	7.2 ± 0.8	15.6 ± 1.4
Z + X	$1.2^{+1.1}_{-0.8}$	$0.9^{+0.7}_{-0.6}$	$2.3^{+1.8}_{-1.4}$	$4.4^{+2.2}_{-1.7}$
All backgrounds ($110 < m_{4\ell} < 160\,\text{GeV}$)	$3.9^{+1.1}_{-0.8}$	$6.6^{+0.9}_{-0.8}$	$9.5^{+2.0}_{-1.6}$	$20.0^{+3.2}_{-2.6}$
Observed ($110 < m_{4\ell} < 160\,\text{GeV}$)	6	6	9	21
Expected Signal ($m_{\text{H}} = 125\,\text{GeV}$)	1.37 ± 0.44	2.75 ± 0.56	3.44 ± 0.81	7.6 ± 1.1
All backgrounds (signal region)	$0.71^{+0.20}_{-0.15}$	$1.25^{+0.15}_{-0.13}$	$1.83^{+0.36}_{-0.28}$	$3.79^{+0.47}_{-0.45}$
Observed (signal region)	1	3	5	9

- The real analysis made use of fairly complex statistical methods (likelihood fit with profiling of systematics), but you can use the table above for a *cut-and-count analysis*
  - Q1: estimate the significance (⇔ p-value) of the excess in the signal region, ignoring any systematic uncertainty
  - Q2: as above, for the signal expectation in the 125 GeV hypothesis
  - Q3: as above, assuming a 50% uncertainty on the sum of backgrounds
  - Q4: propose some method to decrease the background uncertainty

Poisson distribution

$$P(n) = \mu^n \frac{e^{-\mu}}{n!}$$

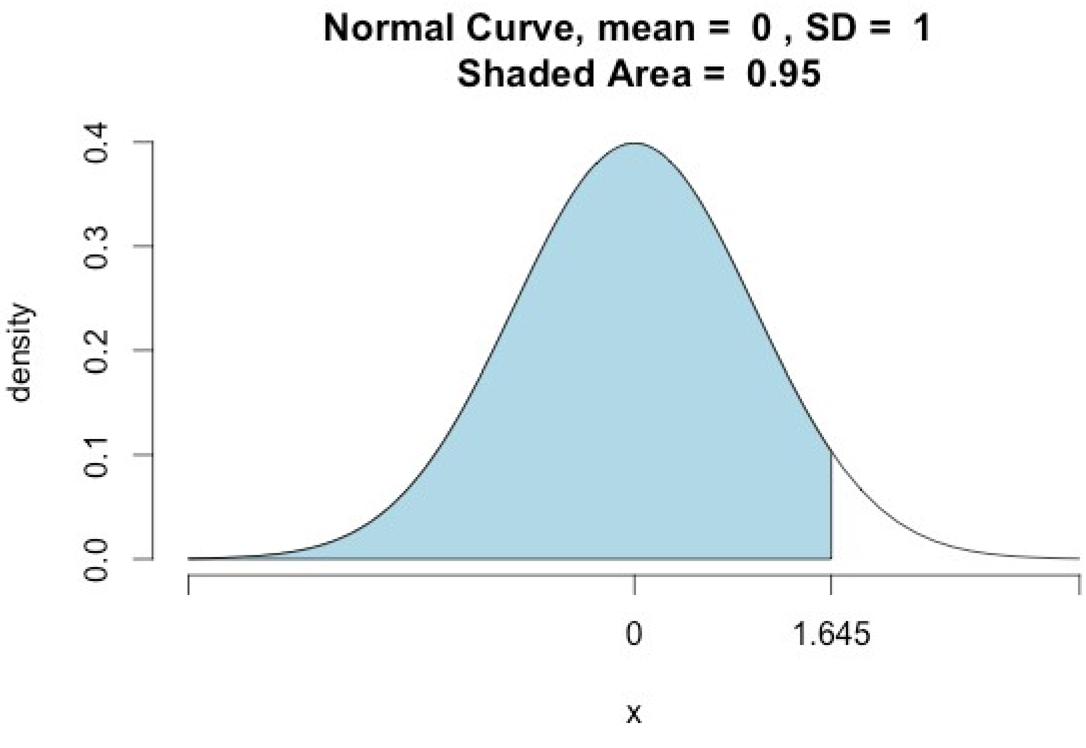
Here n is the number of observed events (9 in the signal region, summed over all channels), and μ is the expectation.

To calculate the p-value of the excess, consider the null hypothesis (B-only: $\mu = 3.8$)

Result: p-value ~ 1% of observing an excess as large as in the data (or superior) under the B-only hypothesis.

To answer Q2, consider n ~ S+B.

Correspondence between p-values and significances



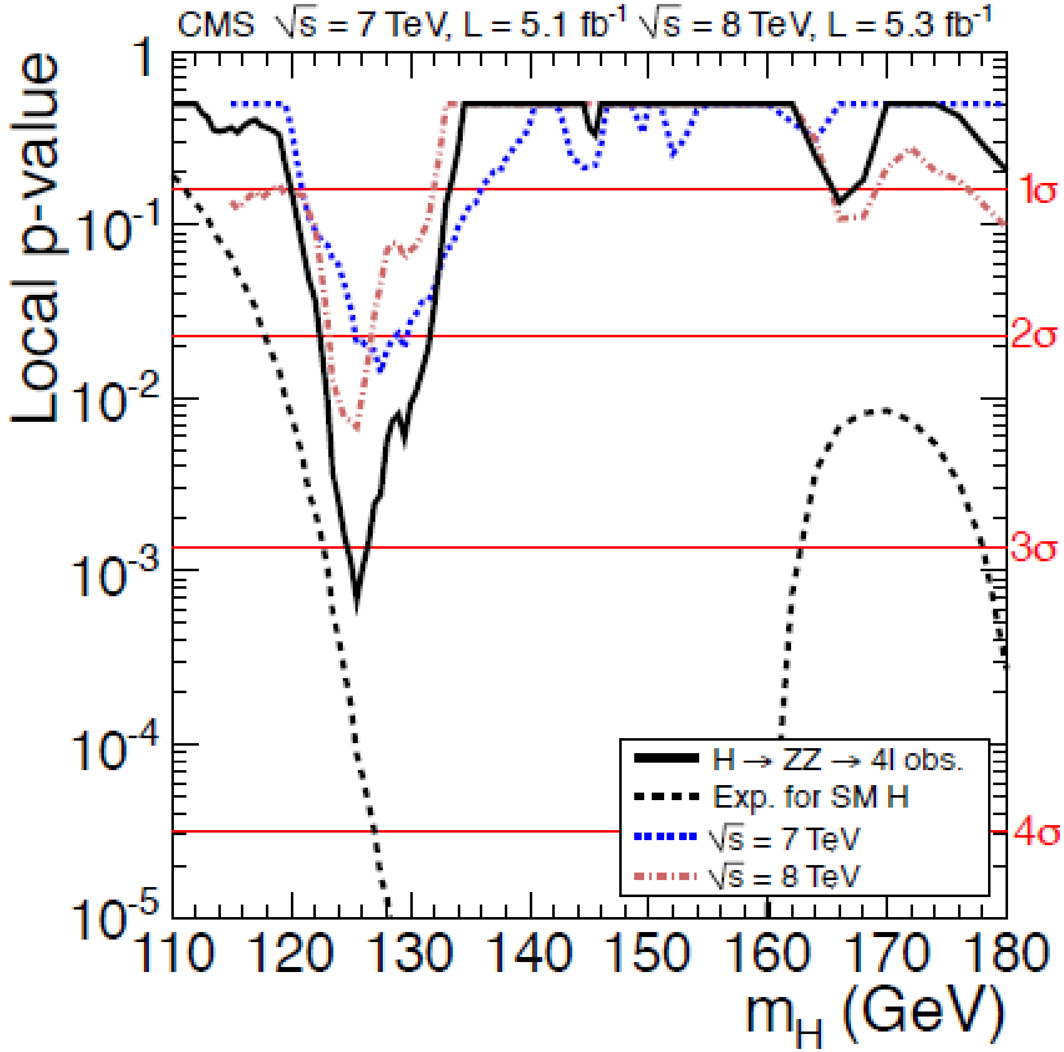
(In our case we get z~2.15)

Take a normalized (area=1) Gaussian centered on 0, with standard deviation =1.

Your p-value is the white area in the tail. If you know the significance (or *z score*) integrate from z to +∞ to get the p-value. If you know the p-value, inverse operation gives the z score.

Unfortunately a Gaussian can only be integrated numerically; trivial in ROOT and other programs; it's easy to find tables or free online tools (e.g.: link)

Fit result: why is it better?



A quick and dirty approximation

- Popular and easy-to-remember approximation: $z \sim S/\sqrt{B}$
- In case of an excess in data, $z \sim (n-B)/\sqrt{B}$
- Intuitive explanation: the signal has to "stick out" of the statistical fluctuations of the background
  - For a Poisson distribution, the standard deviation is the square root of the best estimator
- This approximation is valid for:
  - Relatively large statistics
  - Low purity: B>>S (this is not the case here!)
  - In our case this formula gives z~2.7 (compare with the number we obtained from the actual Poisson probability)

A quick and dirty approximation

- For an alternative derivation of formula z ~ (n-B)/√B, consider how a cross section is measured:

$$\sigma = \frac{n - B}{\epsilon L}$$

- Very rough approximation of significance: incompatibility of the measured cross section with 0 (= B-only hypothesis)
  - Consider error propagation from n to σ:

$$\delta\sigma = \frac{\partial\sigma}{\partial n}\delta n = \sigma \frac{\sqrt{n}}{(n - B)}$$

- Assuming z ~ σ/δσ, and considering low purity (n~B), we obtain our approximate formula

Now consider an uncertainty on B

$$(\delta\sigma)^2 = \left(\frac{\partial\sigma}{\partial n}\delta n\right)^2 + \left(\frac{\partial\sigma}{\partial B}\delta B\right)^2 = (\sigma)^2 \frac{n+\delta B^2}{(n-B)^2}$$

- Still approximating z ~ σ/δσ, and considering low purity (n~B), but without neglecting the systematic uncertainties on the expected background (e.g., theory uncertainties):

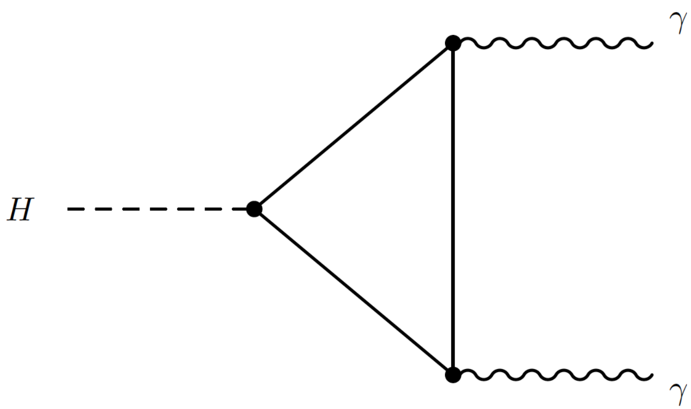
$$z \approx \frac{n-B}{\sqrt{B+\delta B^2}}$$

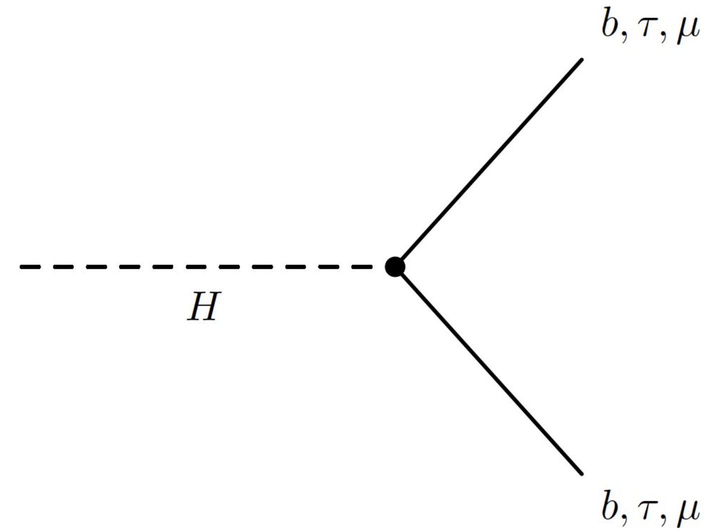
- You can use it for a semi-quantitative answer to Q3

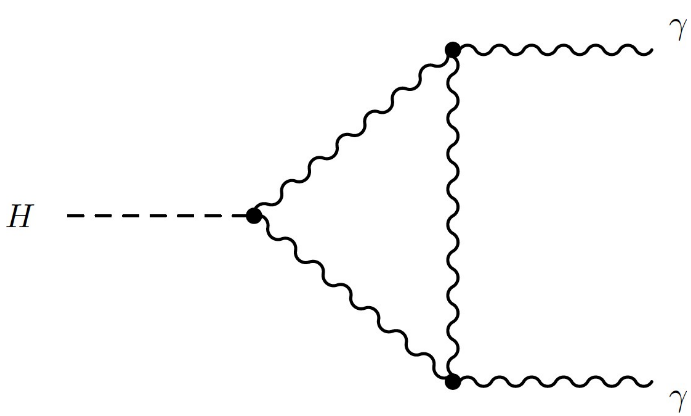
Section 4

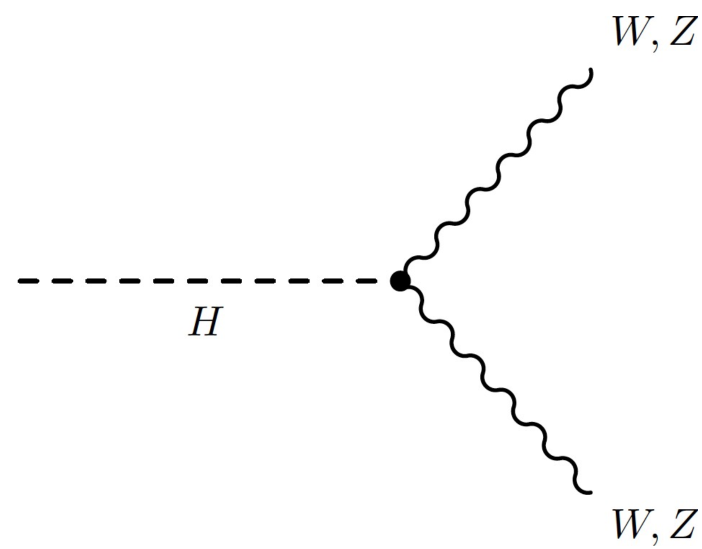
The search for the Higgs boson in the $\gamma\gamma$ channel

Higgs decays



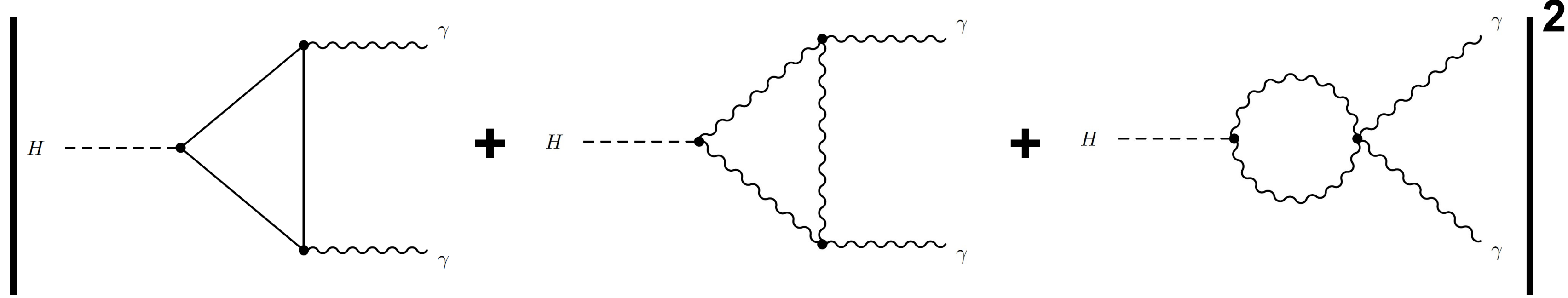






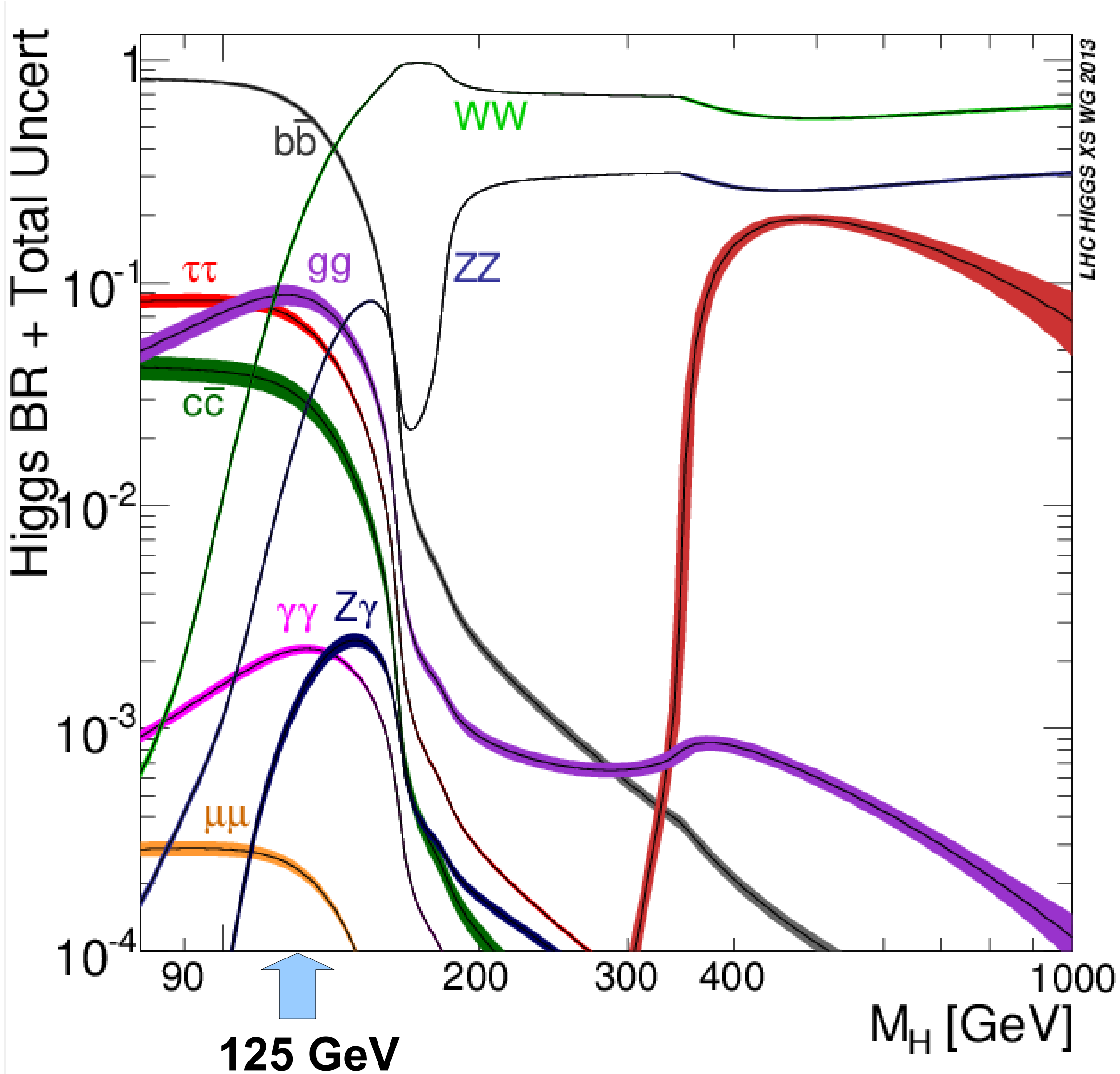
Higgs decay into $\gamma\gamma$

$\Gamma(H\rightarrow\gamma\gamma)\propto$

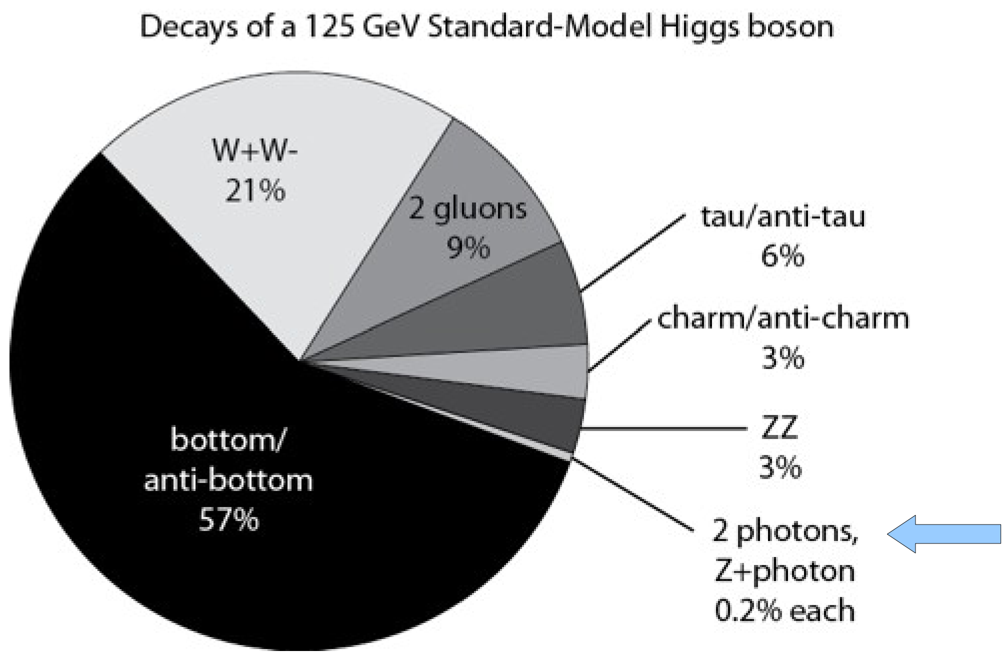


- Both the top and the V(=W,Z) couplings contribute
- Fermion loops and boson loops have amplitudes of opposite sign → destructive interference in SM
- This BR is small (but luckily not negligible) for a combination of this fact and of the large masses implied in the loops, that reduce the probability

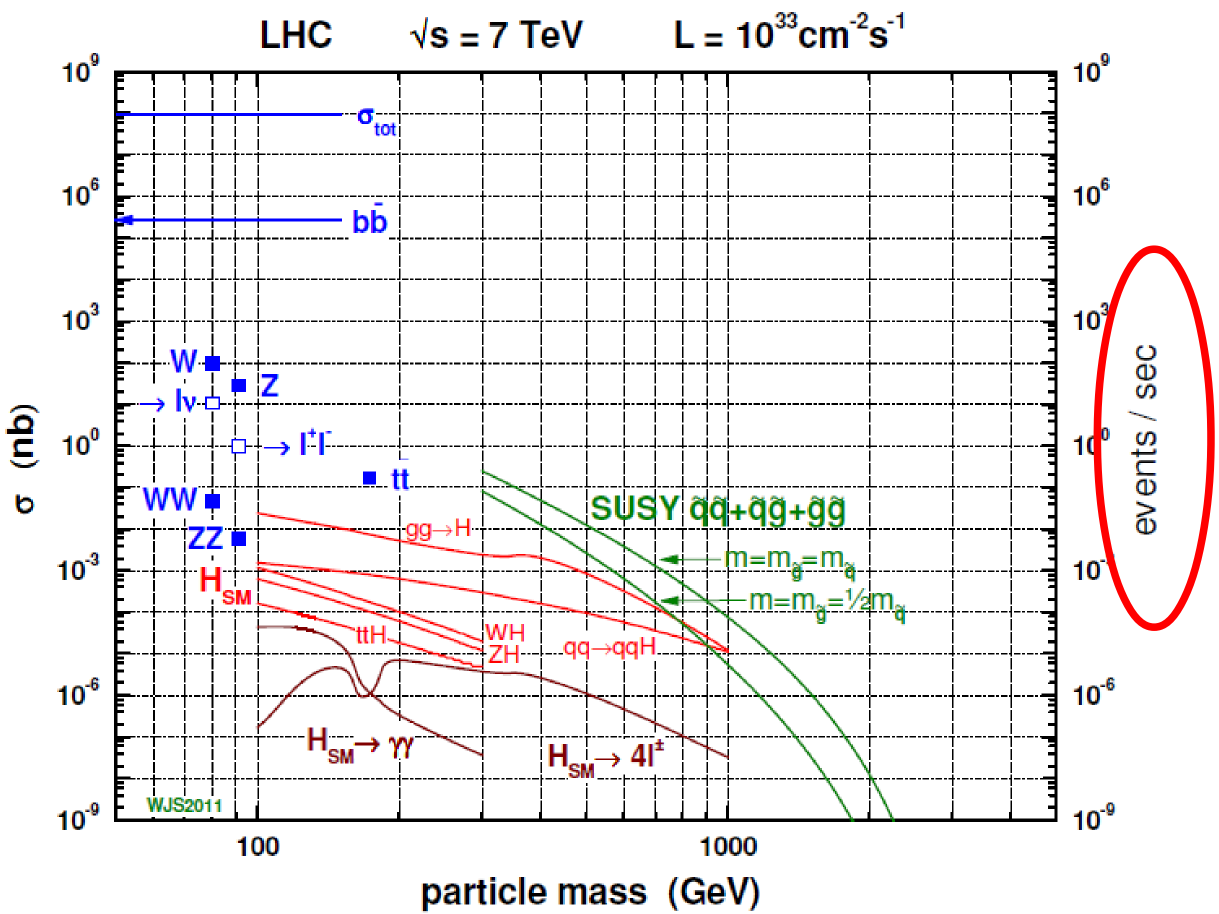
Branching ratios vs mass



Branching ratios @ 125 GeV

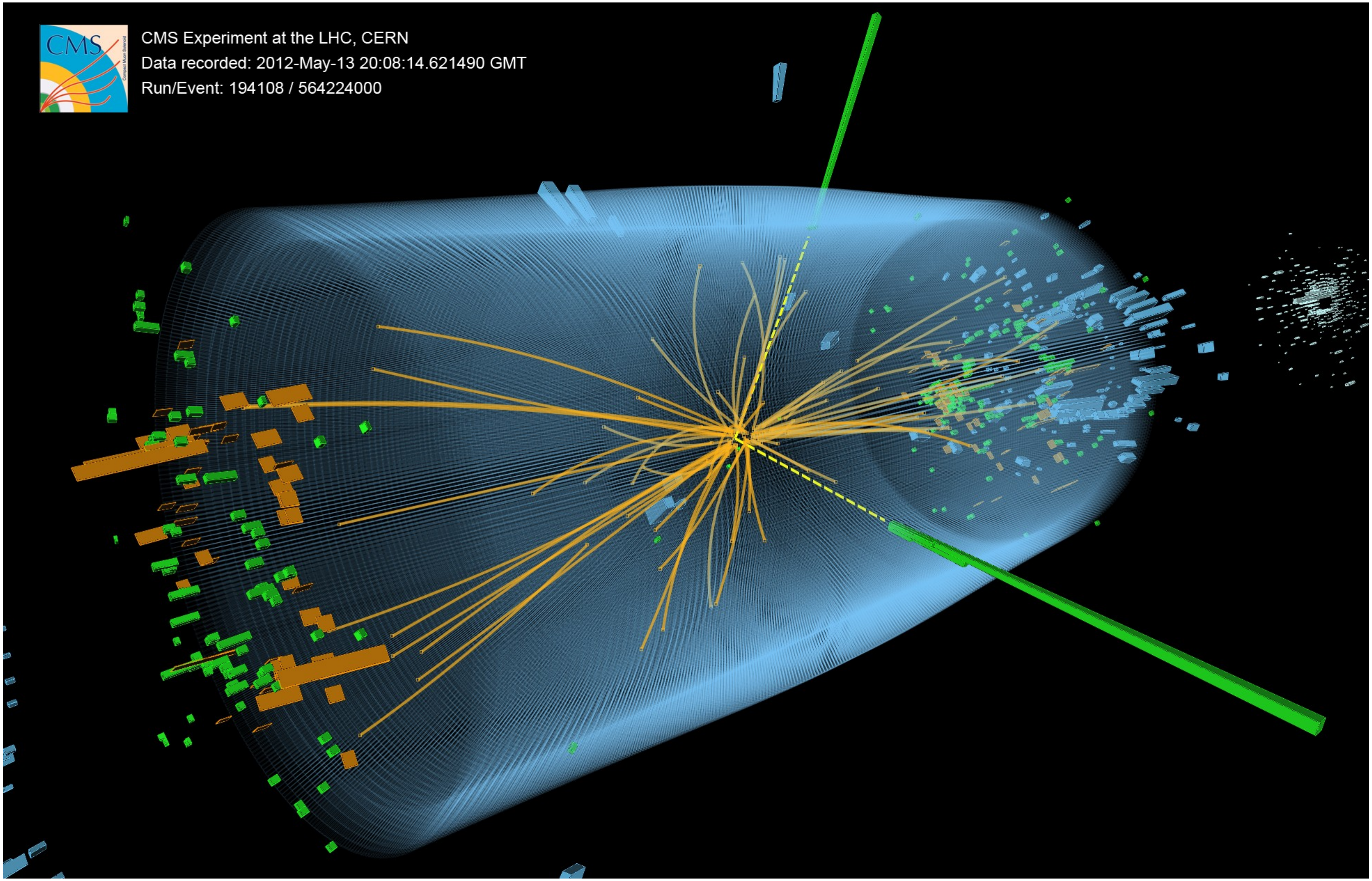


Event rates

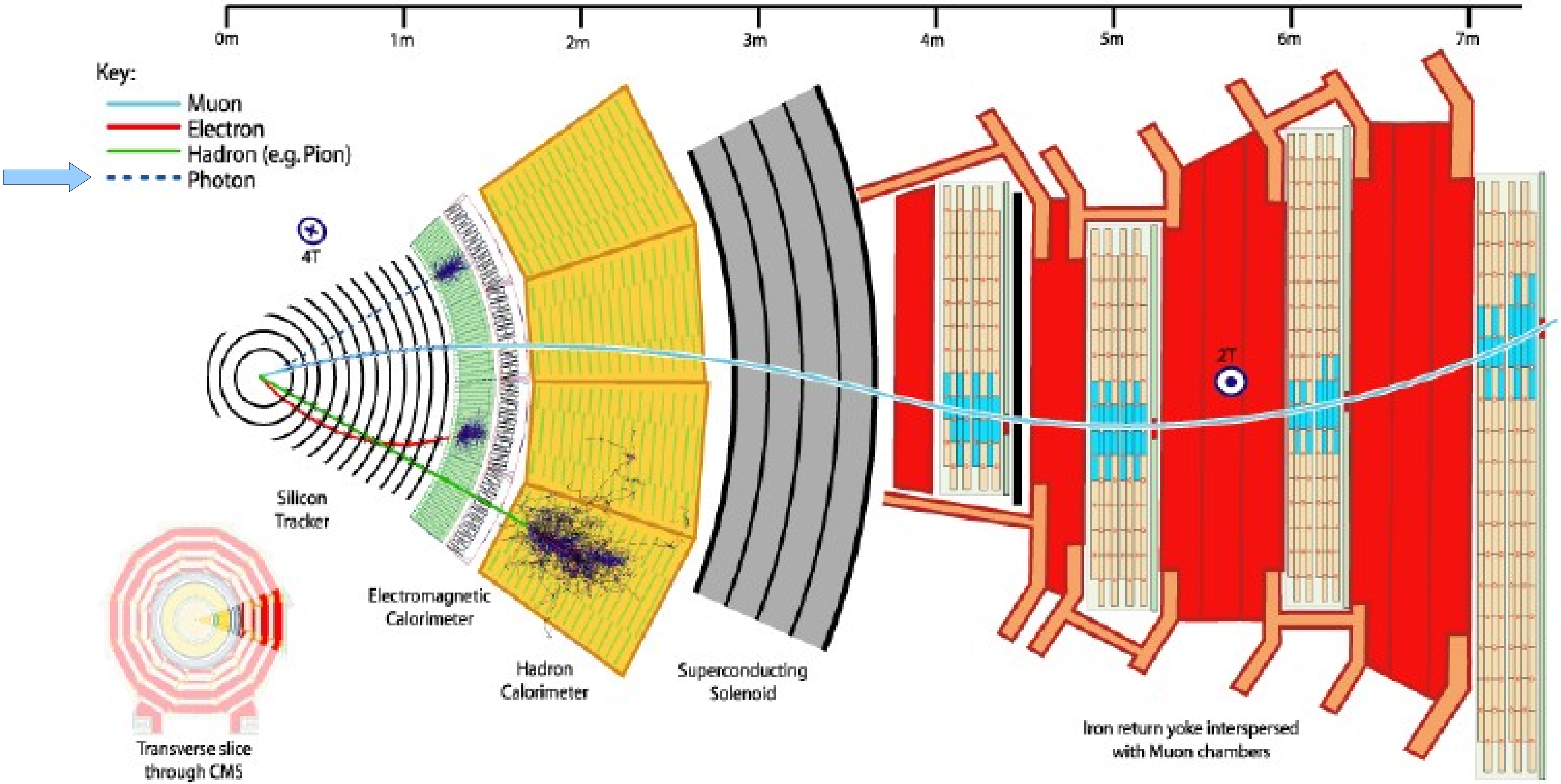


LHC √s = 7 TeV L = 10^33 cm^-2 s^-1
σ (nb)
events / sec
particle mass (GeV)
σtot
bb̄
W
Z
→ lν
→ l+l-
tt̄
WW
ZZ
HSM
gg→H
SUSY q̃q̃+q̃g̃+g̃g̃
m=mg̃=mq̃
m=mq̃=½mg̃
ttH
WH
ZH
qq→qqH
HSM→ γγ
HSM→ 4l±
WJS2011

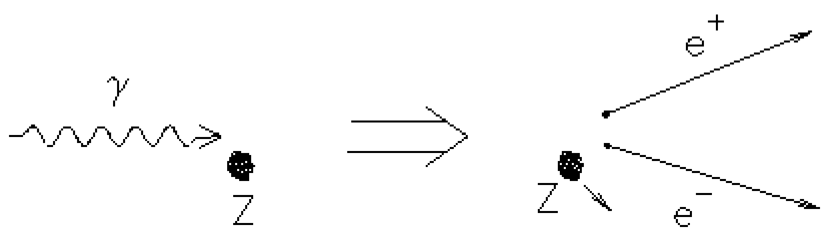
A candidate event

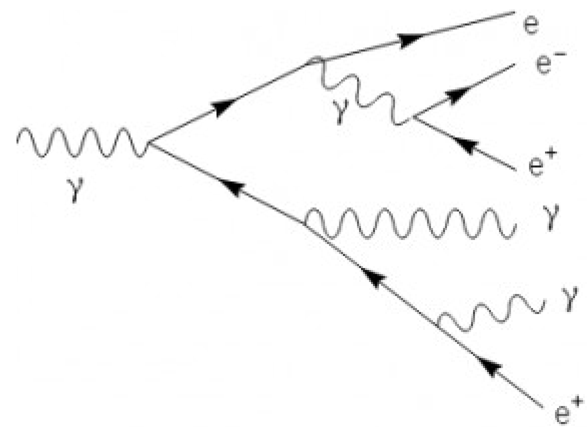


Photon identification



EM shower, initiated by photons

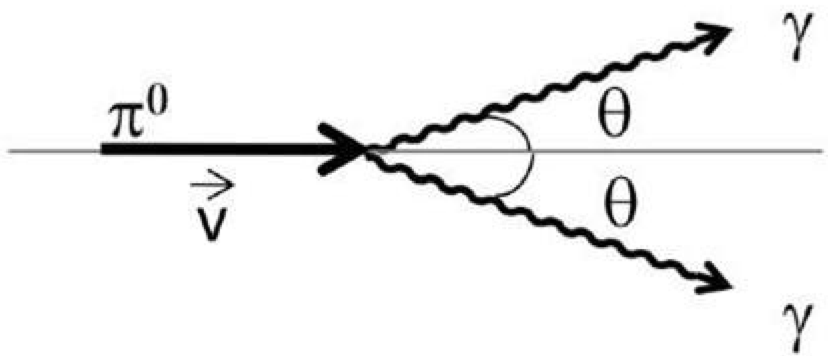




Basic principle of gamma-ray detection: for $E_\gamma >> 2m_e$, the main interaction mechanism for photons in matter is conversion into e+e- thanks to the intense EM fields in the proximity of nuclei.
(High-Z materials create more intense EM fields.)
In turn, electrons and positrons create EM showers by the same processes that we saw in the previous section.
And positrons, in addition, annihilate with the electrons in the material.

Neutral pion background

The π^0 has a lifetime of $\sim 10^{-16}$ s and a mass of 135 MeV.

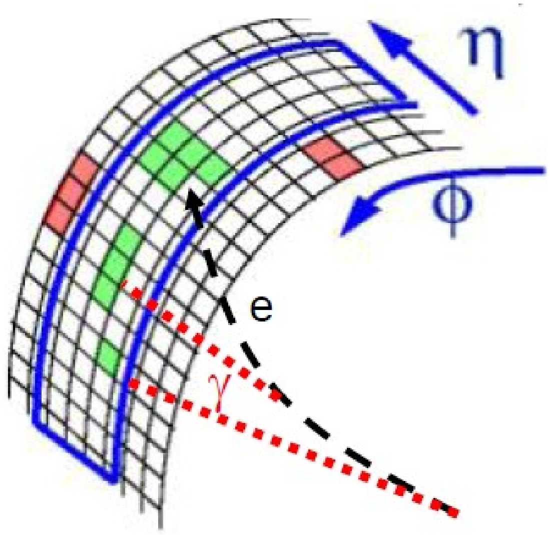


Exercise 1: calculate the decay length for E_π = 60 GeV; compare with the inner radius of CMS tracker (4 cm) and EM cal. (1.3 m).

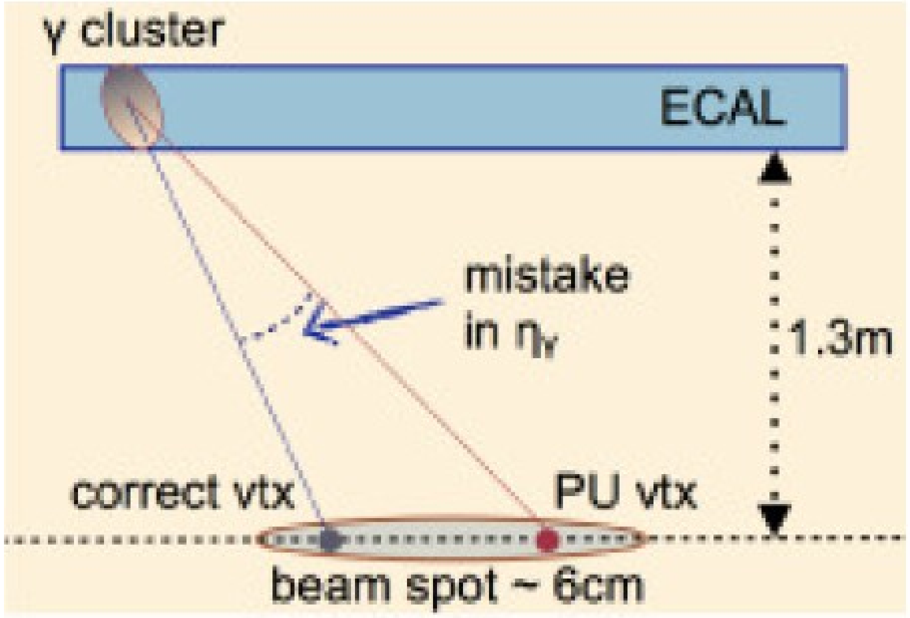
Exercise 2: calculate the angle θ, for E_π = 60 GeV; compare with the angular acceptance of an ECAL cell 2 cm wide.

Photon identification

- A particle is identified as a photon if:
  - Signal in the EM calorimeter does not match with any track; this rejects electrons
  - The energy seen in the hadronic calorimeter is ~0, or at least much less than in the EM calorimeter; this rejects most hadrons (but not π^0)
  - It is isolated; this rejects all kinds of hadrons
  - Spatial shape of the signal cluster is consistent with a single photon; this rejects π^0
- A complication:
  - The probability that a photon converts in e^+e^- before reaching the calorimeter is not so small
  - We don't want to reject all "converted photons", especially when searching for the rare $H\to\gamma\gamma$



Photon direction

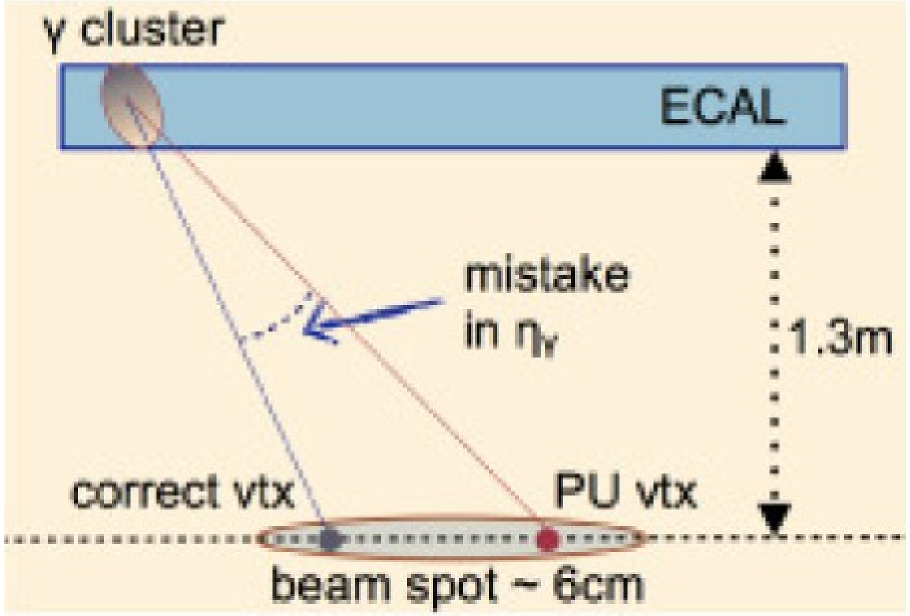


Remember the formula of invariant mass: $M_{12} = \sqrt{2E_1E_2(1-\cos\theta_{12})}$

When dealing with charged particles we neglect the uncertainty on the angle, but this is not negligible anymore for neutrals!

Exercise: assume 1 cm uncertainty on the origin of the photon along the z axis; what is the corresponding uncertainty on $m_{\gamma\gamma}$?

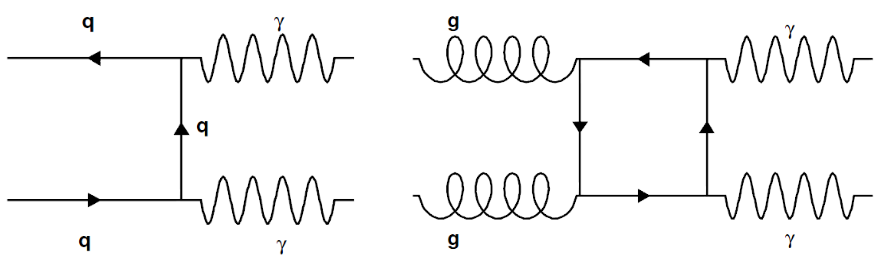
Photon direction



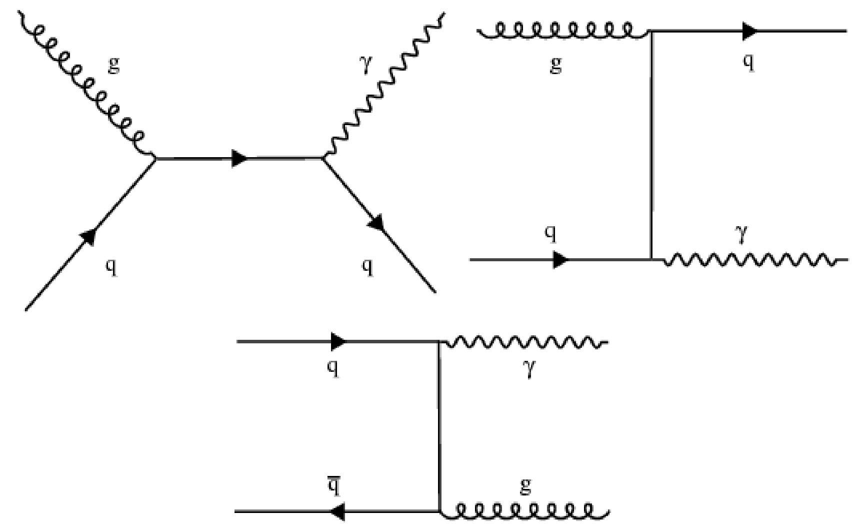
In practice, we first reconstruct the position of the „leading vertex“ as the one with most tracks associated (*Q: why?*), and then try to associate the photon direction with it.

By the way, the „converted photons“, that have the disadvantage of a worse background contamination, have at least one advantage: their trajectory is more precisely known (*Q: how?*)

Main backgrounds



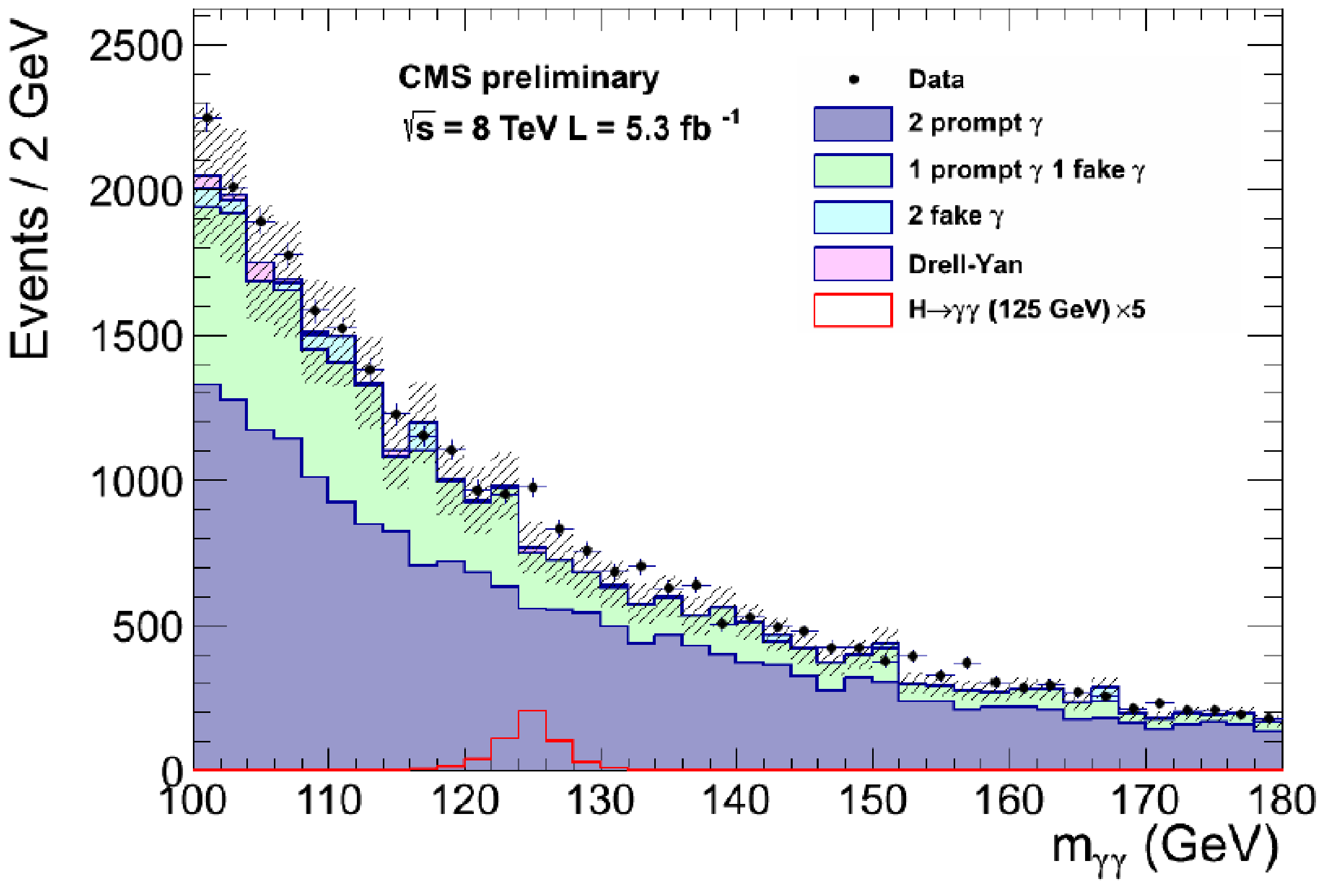
Irreducible background: both photons are real; rate is low but probability to pass selection is large



One photons is real, the other is fake

Non-negligible contribution from di-jet events with two fake photons; probability of selection is very small, but rate is huge

Invariant mass γγ, data vs MC (CMS 2012)

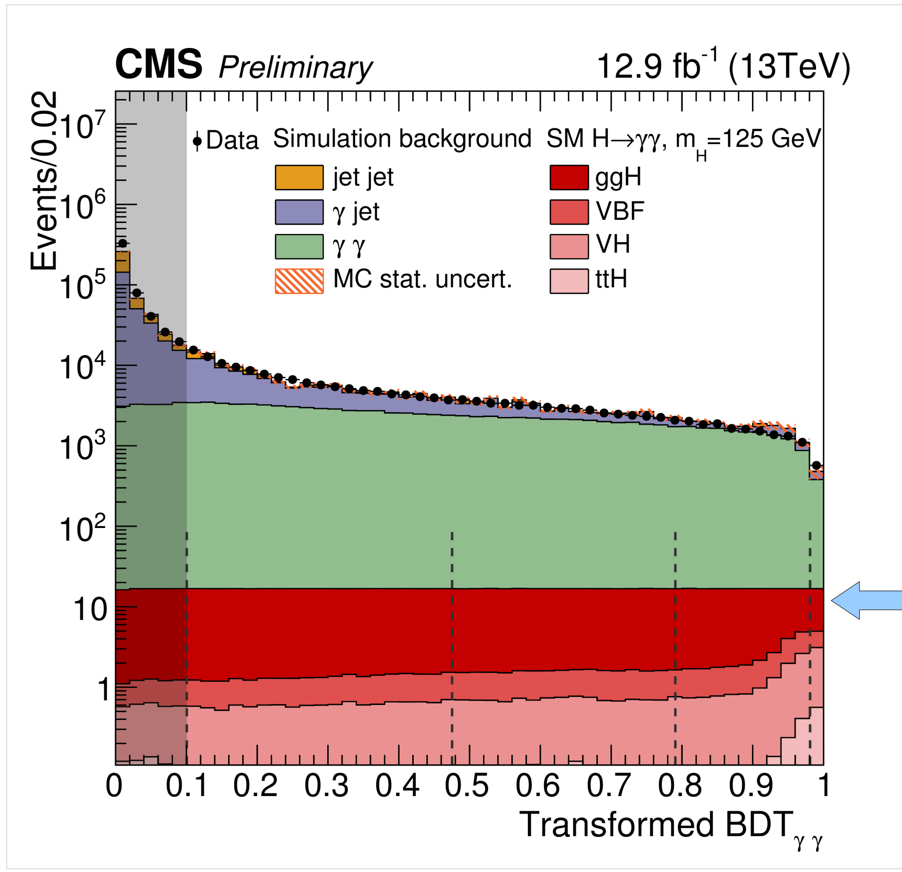


Multivariate discriminant

Object-level MVA is used to select good photons; its output is used as input to an event-level MVA that also includes:

- $p_T/m_{\gamma\gamma}$ of both photons
- Angular distribution of the two photons

  (*Q: why?*)
- Probability of having correctly identified the interaction vertex
- Estimated resolution on $m_{\gamma\gamma}$



(Ths variable was already used in 2012, but this more recent plot also shows the bkg fractions)

Flatness is by construction

Event weighting

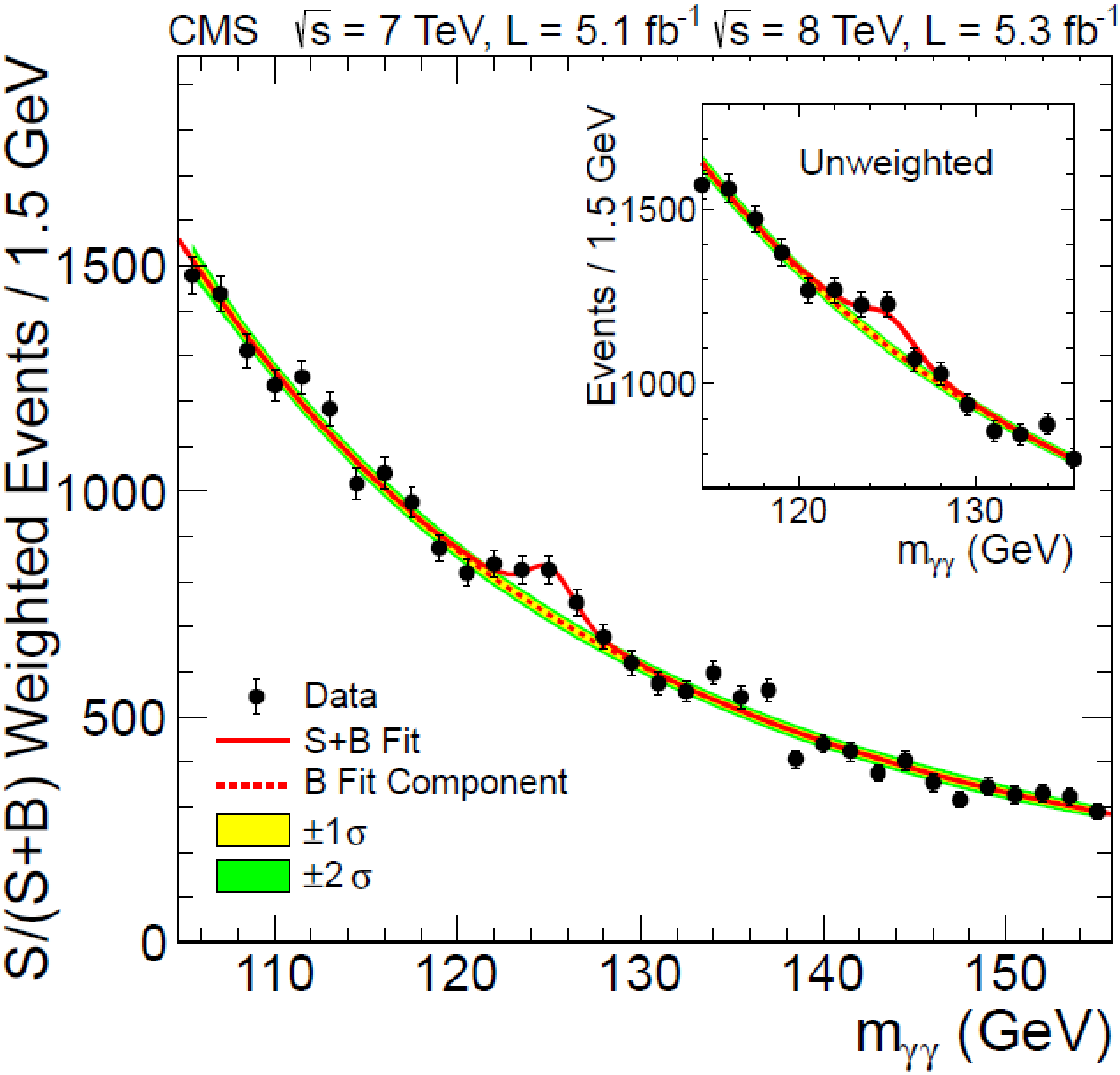
$$w_{\rm sig} = \frac{p_{\rm vtx}}{\sigma_m^{\rm correct}/m_{\gamma\gamma}} + \frac{1-p_{\rm vtx}}{\sigma_m^{\rm incorrect}/m_{\gamma\gamma}}$$

Here p_{vtx} is the estimated probability of having selected the correct vertex, and σ_m is the estimated mass resolution for correct/incorrect photon-vertex associations.

How to estimate those quantities:

- From $Z\rightarrow ee$ events we can extract σ_E/E
- From $Z\rightarrow\mu\mu$ events we have a reliable test of p_{vtx}

Invariant mass γγ (CMS 2012)

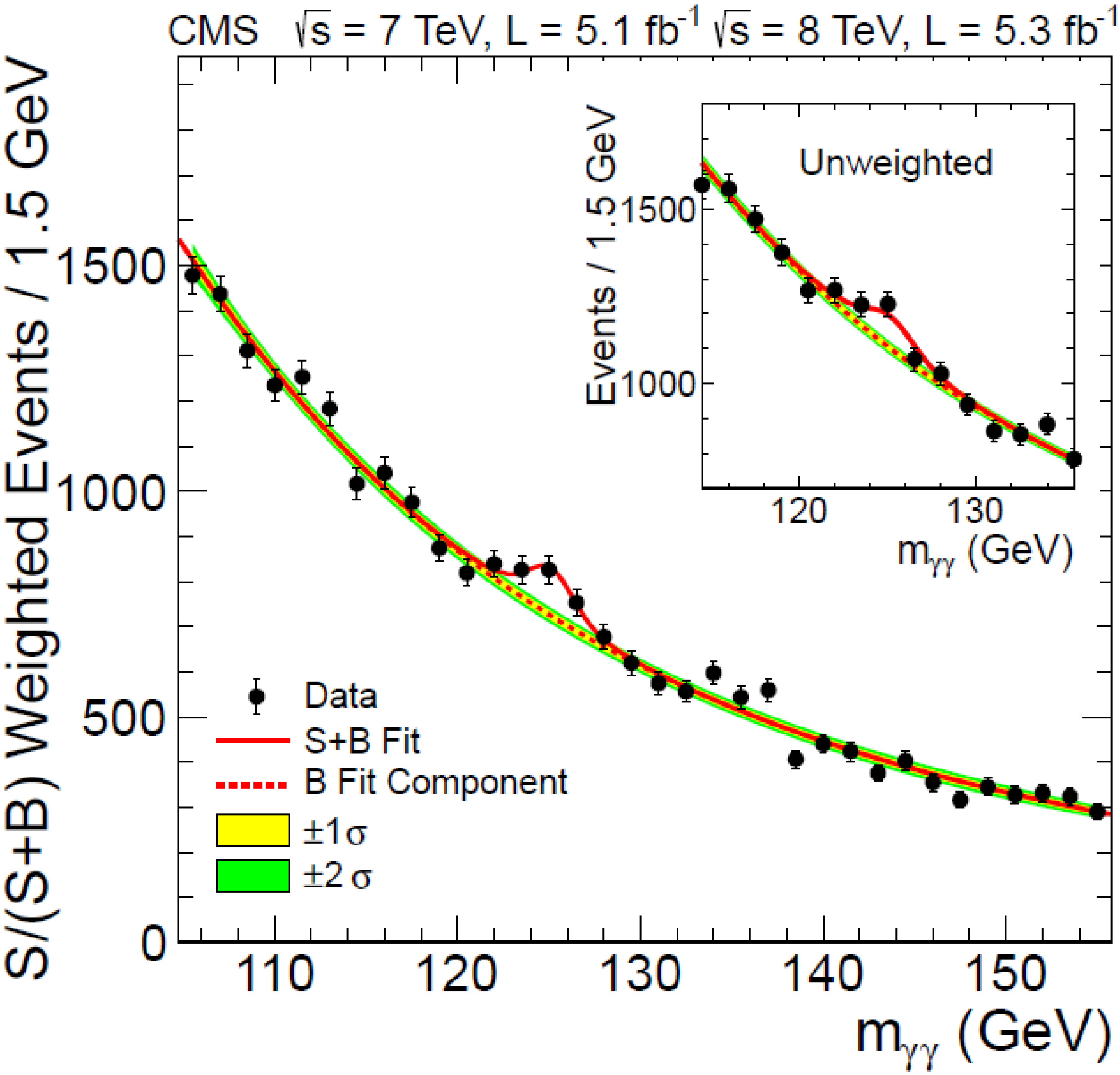


The data histogram is filled with weighted events (by categories with different selection purity.)

MC is not used, because it is difficult to reliably simulate all the effects that cause a jet to produce a fake photon: details of hadronization (number and spectrum of π^0's and other hadrons), details of detector response to hadrons.

Because S<<B, any error on B has a large effect on the result.

Final fit

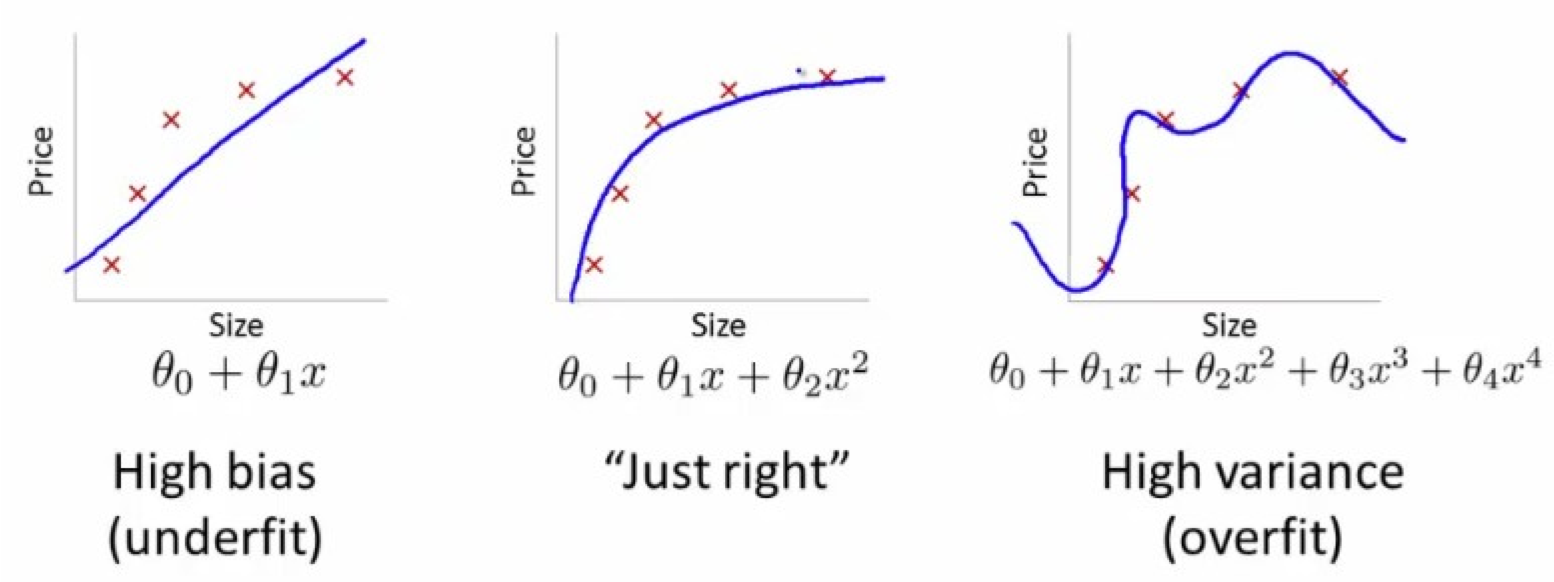


Function for signal is fitted on MC events.

A smooth function is considered for the sum of backgrounds.

Several classes of functions are tried (exponentials, power law, polynomials) with a number of free parameters chosen such to give a good χ^2 but without over-fitting.

Overfitting

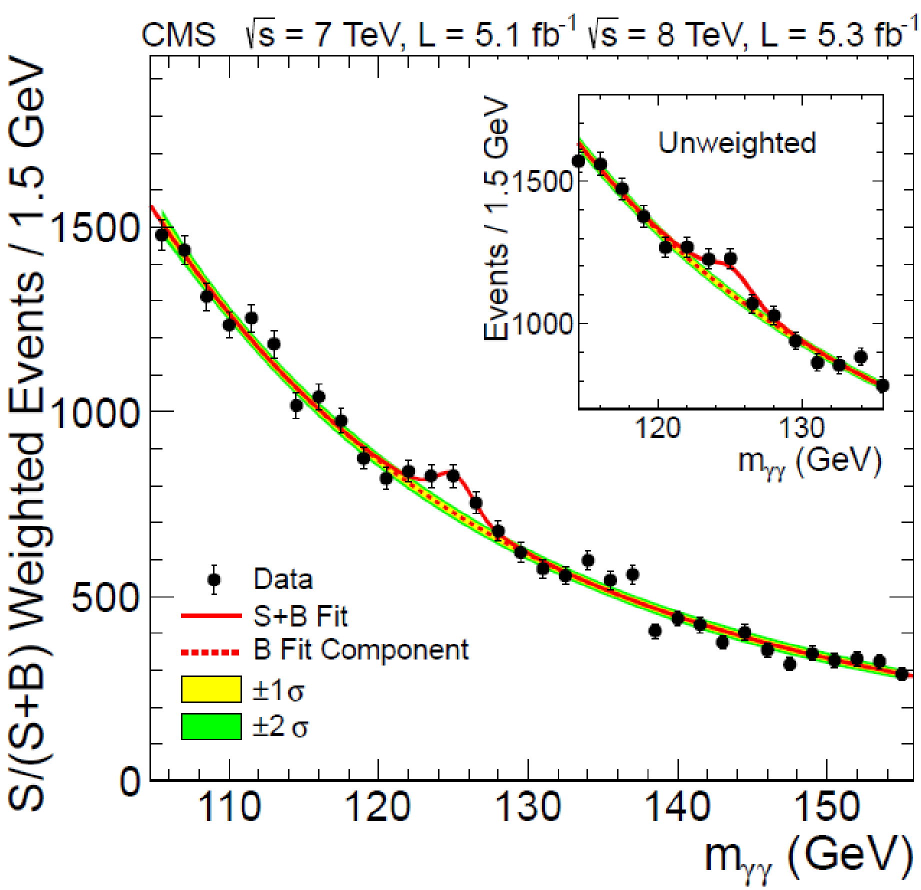


(Picture from link)

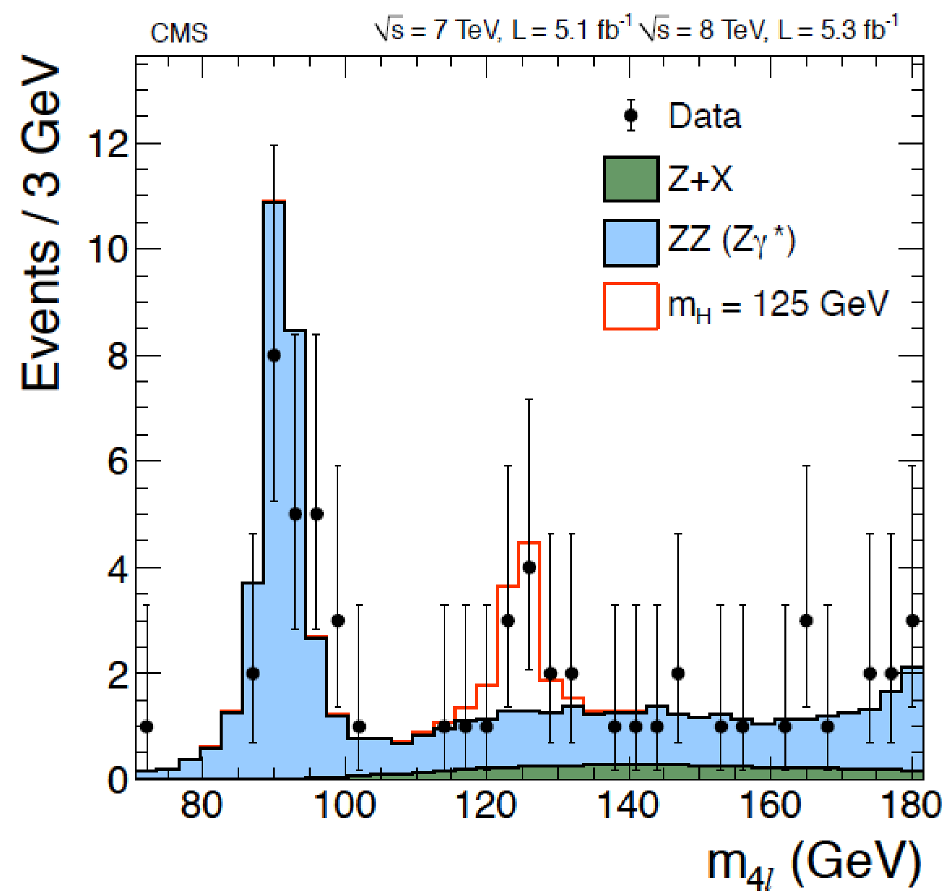
Compare the χ^2 obtained with n and n+1 parameters; by construction, the χ^2 will be smaller; to avoid overfitting, the extra parameter should be added only if it is *significantly* smaller.

There are several ways to choose the optimal number of free parameters. All these methods quantify the p-value that the improvement in χ^2 from n to n+1 is due to pure chance.

Compare with 4l analysis

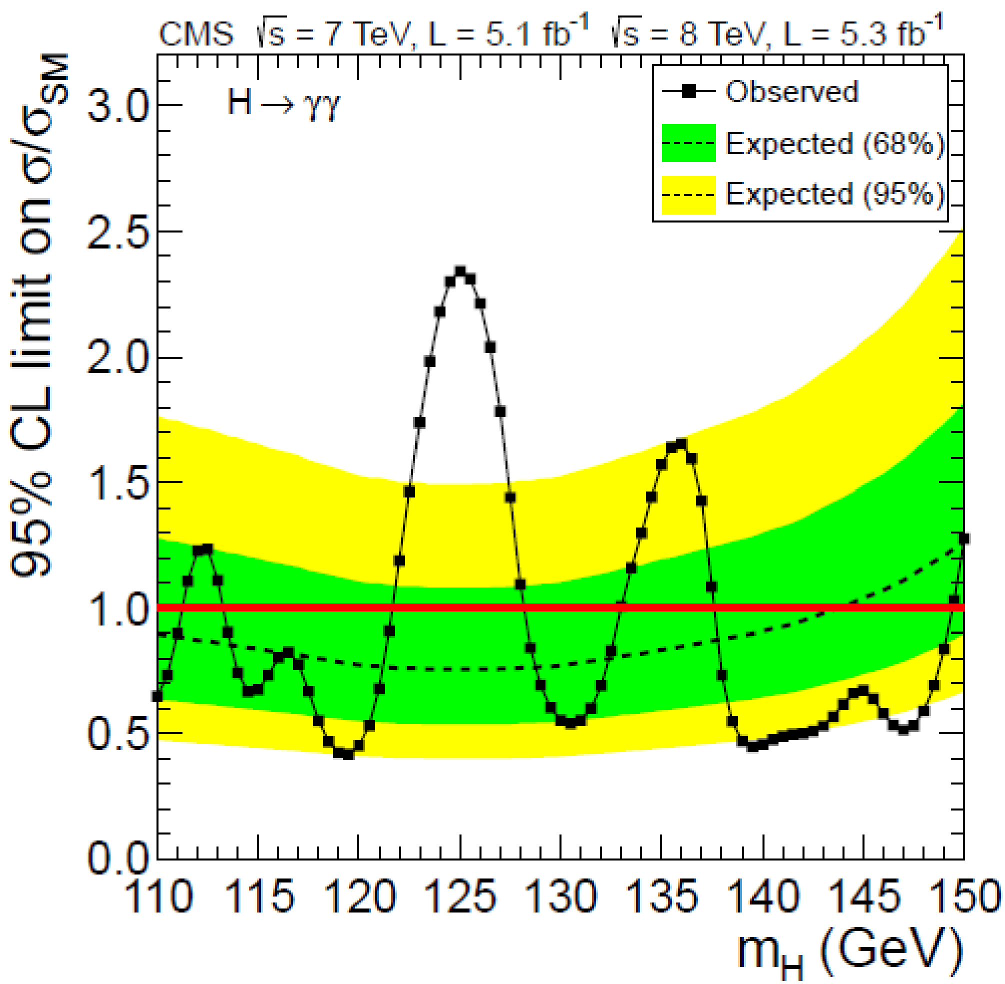


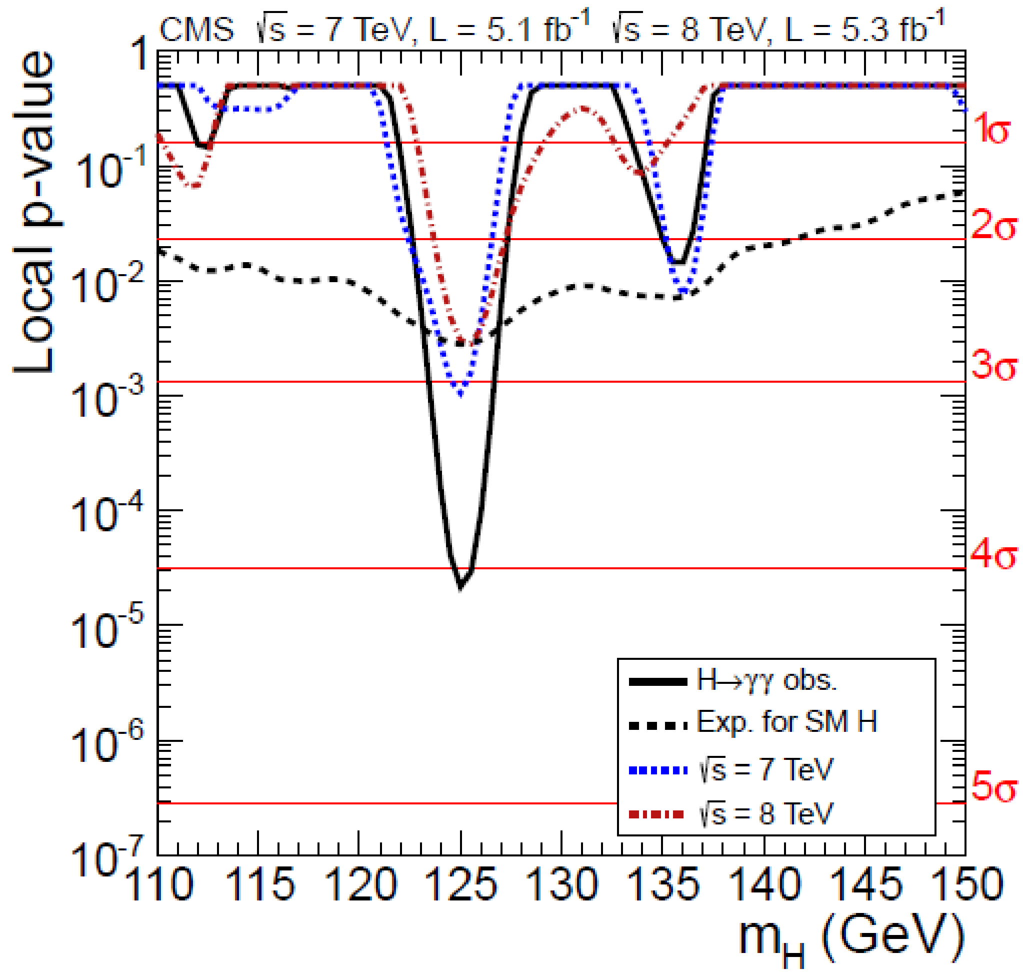
Higher statistics



Higher purity

Result


CMS $\sqrt{s}$ = 7 TeV, L = 5.1 fb^{-1} $\sqrt{s}$ = 8 TeV, L = 5.3 fb^{-1}
H → γγ
Observed
Expected (68%)
Expected (95%)
95% CL limit on σ/σ$_{SM}$
m$_H$ (GeV)


CMS $\sqrt{s}$ = 7 TeV, L = 5.1 fb^{-1} $\sqrt{s}$ = 8 TeV, L = 5.3 fb^{-1}
Local p-value
1σ
2σ
3σ
4σ
5σ
H→γγ obs.
Exp. for SM H
$\sqrt{s}$ = 7 TeV
$\sqrt{s}$ = 8 TeV
m$_H$ (GeV)

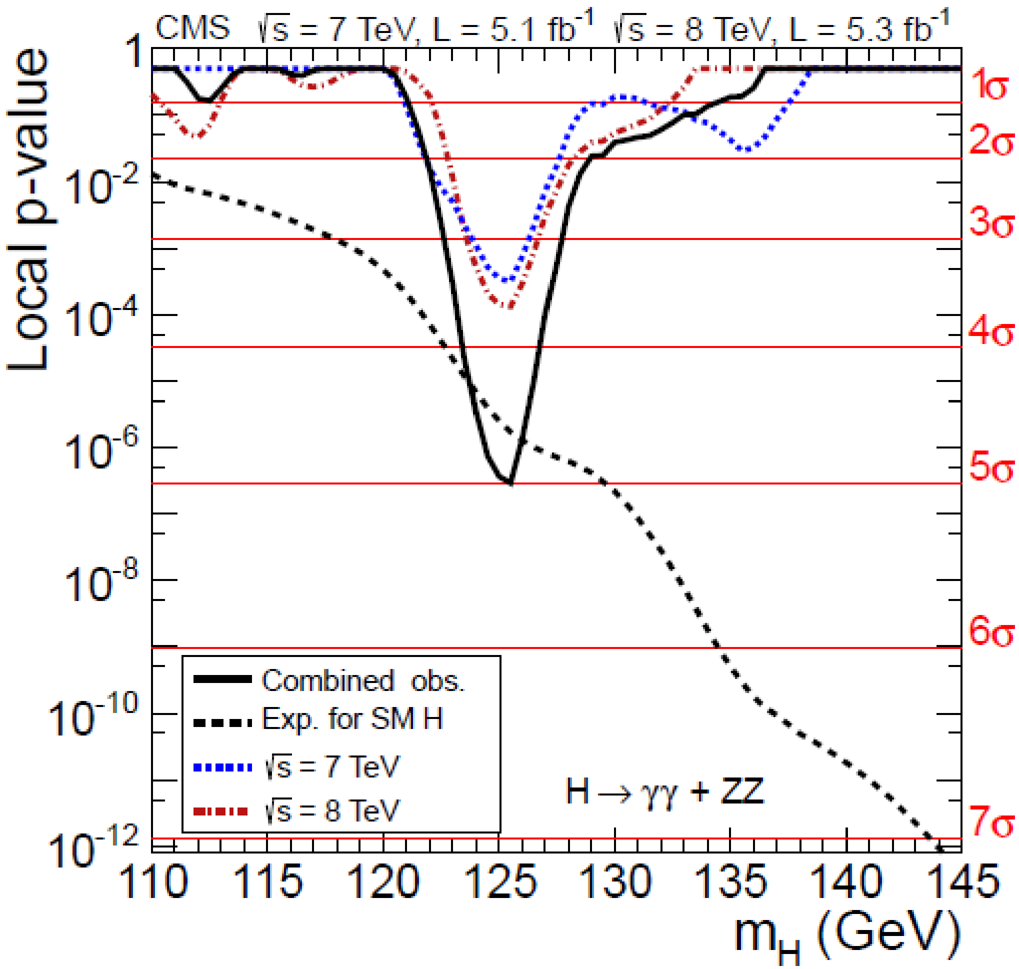
Exercise

- Answer the same questions as in last week's homework
- Consider an excess of ~200 events over an estimated background of ~750 in the most significant bin
- Which of those two formulas would you use?

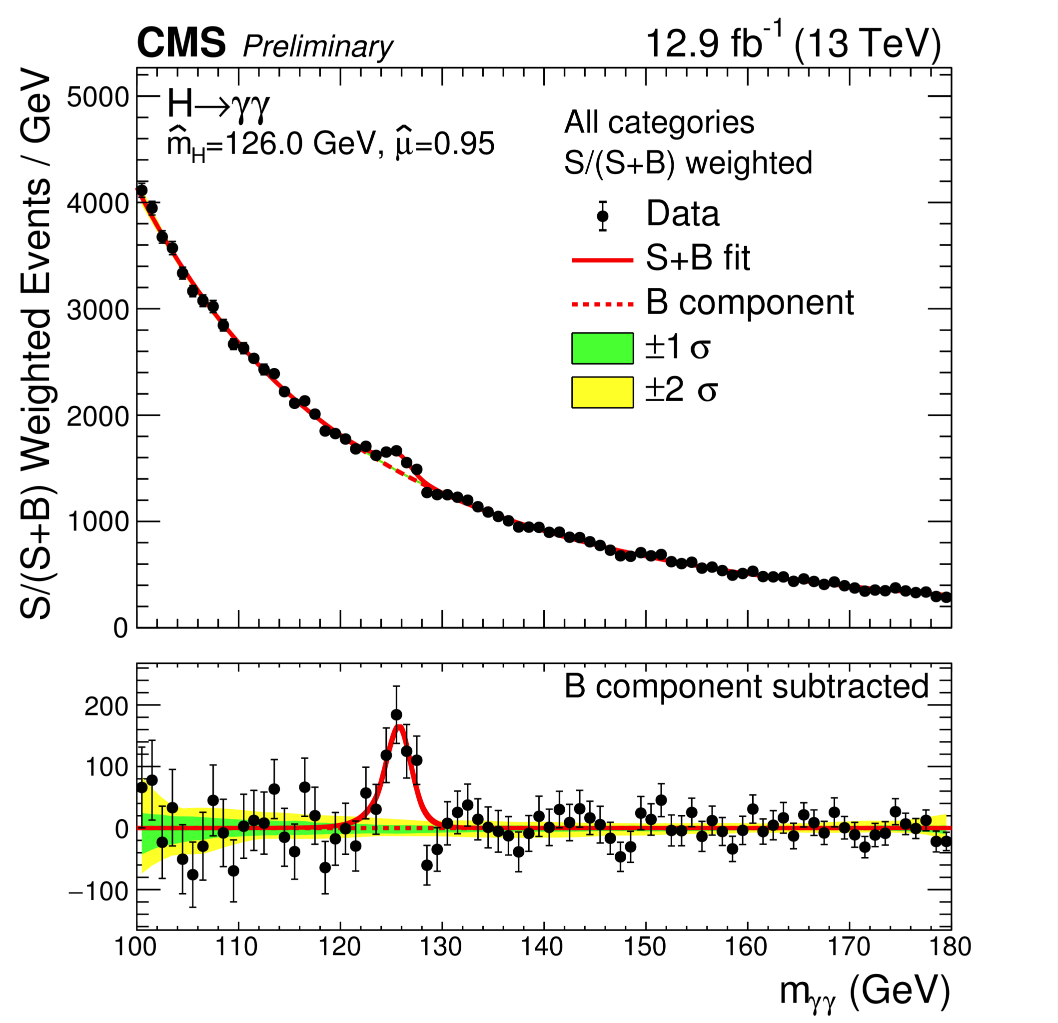
$$P(n) = \mu^n \frac{e^{-\mu}}{n!} \qquad\qquad z \approx \frac{n - B}{\sqrt{B + \delta B^2}}$$

- Make various hypotheses on the uncertainty of the expected background, to appreciate the importance of constraining it from the "sidebands"

Result of combination with 4l



How γγ looks like today



What comes next

- 9/3: other final states & study of Higgs properties
- 16/3 and 30/3: top-Higgs connection
- 23/3: mid-term evaluation
  - I am circulating four scientific articles about a different particle (the top quark)
  - Each of you has to choose one, and on 23/3 you will present in 15' (+15' Q&A) what you understand of it
  - In case of failure, it doesn't count in negative for your final evaluation; in case of success, it will count in positive
  - I will give individual feedback
- 27/4: start of the detector part of this course

For the mid-term evaluation

- Paper #1: $t\bar{t}\rightarrow 2l$ ($l=e,\mu$), CMS @ 7 TeV (early data), link
- Paper #2: $t\bar{t}\rightarrow 1l$ ($l=e,\mu$), CMS @ 7 and 8 TeV, link
- Paper #3: $t\bar{t}\rightarrow \tau+l$ ($l=e,\mu$), CMS @ 8 TeV, link
- Paper #4: $t\bar{t}\rightarrow 0l$, CMS @ 7 TeV, link
- Do not hesitate to contact me in case of doubts or difficulties. If I am not in the office (3^{rd} floor), send a mail.

Questions?

Caractérisation globale d'une collision hadronique, les variables cinématiques utilisées

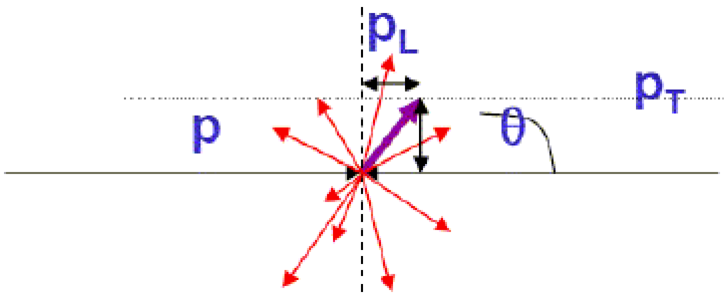
$$p_T = p_\perp = \sqrt{p_x^2 + p_y^2} = p \sin\theta$$ **Moment transversal**

Section efficace invariant

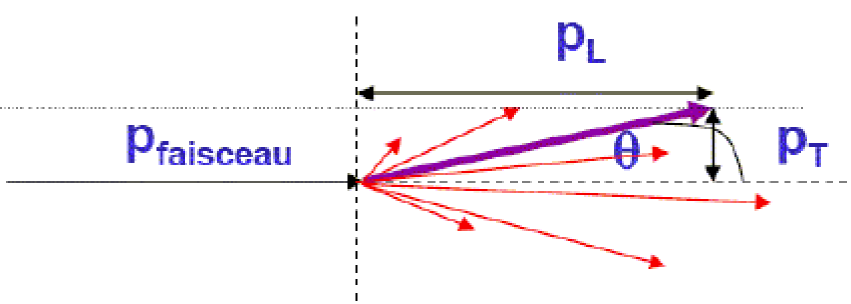
$$E\frac{d^3\sigma}{dp^3} = E\frac{d^3\sigma}{dp_x dp_y dp_z} = \frac{1}{2\pi}\frac{d^2\sigma}{p_T dp_T d(p_L/E)} \sim \underbrace{F(p_T)F'(p_L)}_{\text{(Feynman scaling)}}$$

$$F(p_T) \sim e^{-bp_T}\ ;\ <p_T>_{\text{particules sécondaires}} \approx 0.3 - 0.5\ GeV/c \approx \frac{\hbar}{R}$$

$$E_T = \sum_{i=\text{part.sécondaires}} E_i \sin\theta_i$$ **Energie transversal**



Système de centre de masse



Système de laboratoire

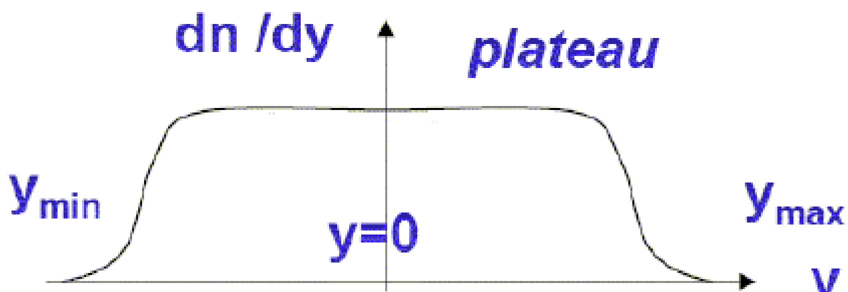
$$x_F = p_L / p_L^{\max} = p_L / \left(\sqrt{s}/2\right) \quad \text{(Feynman "}x\text{")}$$

Rapidité *y*, « invariante » de Lorentz

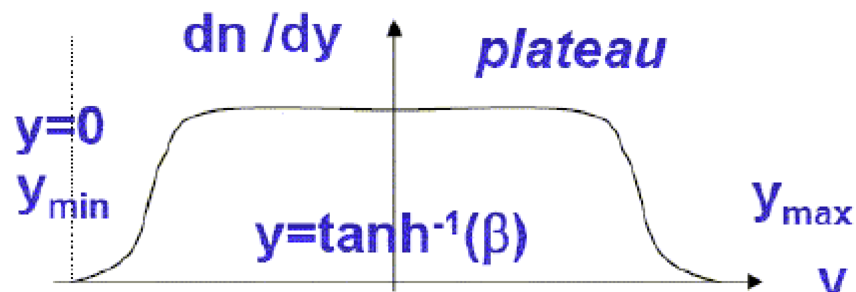
$$y = \frac{1}{2}\ln\left(\frac{E+p_L}{E-p_L}\right) \overset{\beta\to 1,\, m\to 0}{\approx} \eta = -\ln\left(\tan\frac{\theta}{2}\right)$$

η, pseudo-rapidité

$$y \to y + \tanh^{-1}(\beta) \qquad y_{\max} = \frac{1}{2}\ln\left(\frac{s}{m^2+p_T^2}\right)$$



Système de centre de masse



Système de laboratoire

38.5.2. ***Inclusive reactions***: Choose some direction (usually the beam direction) for the z-axis; then the energy and momentum of a particle can be written as

$$E = m_T \cosh y \ , \ p_x \ , \ p_y \ , \ p_z = m_T \sinh y \ , \tag{38.35}$$

where m_T is the transverse mass

$$m_T^2 = m^2 + p_x^2 + p_y^2 \ , \tag{38.36}$$

and the rapidity y is defined by

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right)$$

$$= \ln \left(\frac{E + p_z}{m_T} \right) = \tanh^{-1} \left(\frac{p_z}{E} \right) \ . \tag{38.37}$$

Under a boost in the z-direction to a frame with velocity β, $y \to y - \tanh^{-1} \beta$. Hence the shape of the rapidity distribution dN/dy is invariant. The invariant cross section may also be rewritten

$$E \frac{d^3\sigma}{d^3p} = \frac{d^3\sigma}{d\phi \, dy \, p_T \, dp_T} \Longrightarrow \frac{d^2\sigma}{\pi \, dy \, d(p_T^2)} \ . \tag{38.38}$$

For $p \gg m$, the rapidity [Eq. (38.37)] may be expanded to obtain

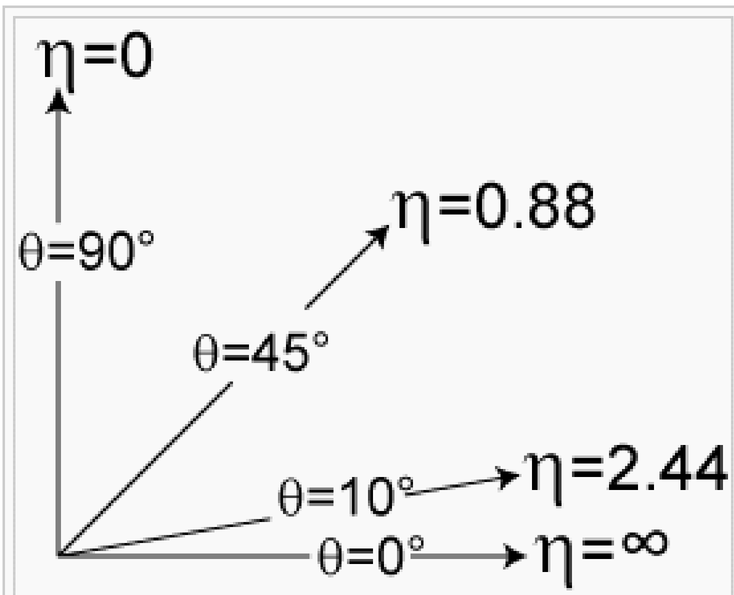
$$y = \frac{1}{2} \ln \frac{\cos^2(\theta/2) + m^2/4p^2 + \ldots}{\sin^2(\theta/2) + m^2/4p^2 + \ldots}$$

$$\approx -\ln \tan(\theta/2) \equiv \eta \tag{38.42}$$

where $\cos\theta = p_z/p$. The pseudorapidity η defined by the second line is approximately equal to the rapidity y for $p \gg m$ and $\theta \gg 1/\gamma$, and in any case can be measured when the mass and momentum of the particle is unknown. From the definition one can obtain the identities

$$\sinh\eta = \cot\theta \ , \quad \cosh\eta = 1/\sin\theta \ , \quad \tanh\eta = \cos\theta \ . \tag{38.43}$$

Pseudo-rapidity



$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$$

EM interactions

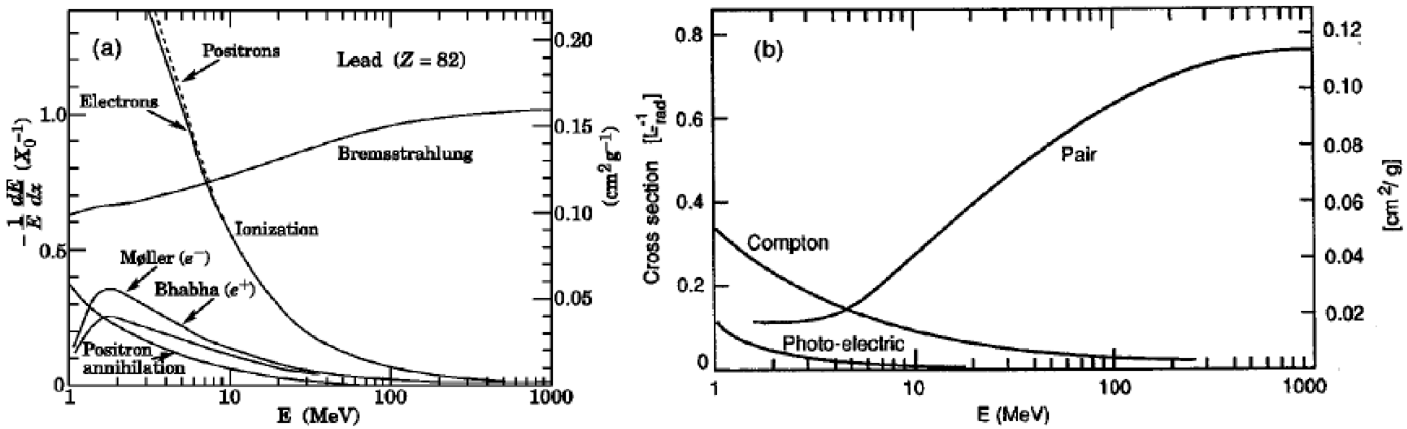


FIG. 1. (a) Fractional energy lost in lead by electrons and positrons as a function of energy (Particle Data Group, 2002). (b) Photon interaction cross section in lead as a function of energy (Fabjan, 1987).

To know more on calorimetry:

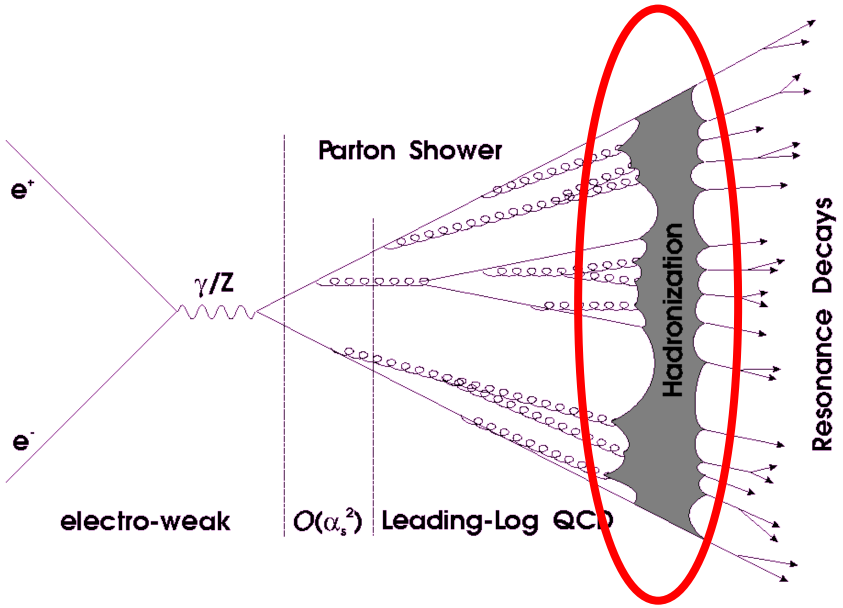
REVIEWS OF MODERN PHYSICS, VOLUME 75, OCTOBER 2003

Calorimetry for particle physics

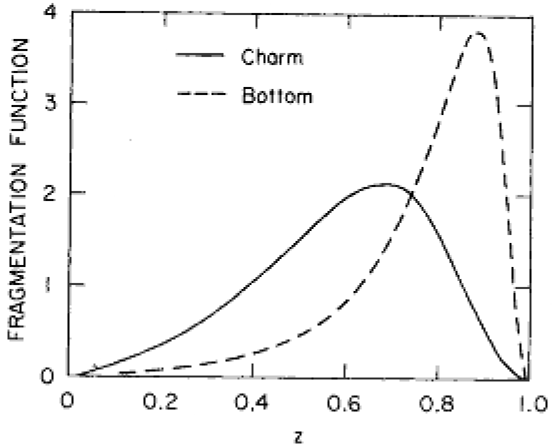
Christian W. Fabjan and Fabiola Gianotti

Fragmentation function

$$z = \frac{(E + p_{||})_{hadron}}{(E + p_{||})_{quark}}$$



$$x_B = \frac{E_{hadron}}{E_{beam}}$$



Peterson's function for heavy quarks:

$$D_Q^H(z) = \frac{N}{z[1-(1/z)-\epsilon_Q/(1-z)]^2} \; .$$