Testing the Standard Model of Elementary Particle Physics II

3rd lecture

Dr. Dominik Duda

6th May 2021

4.2 The physics program at the Large Hadron Collider

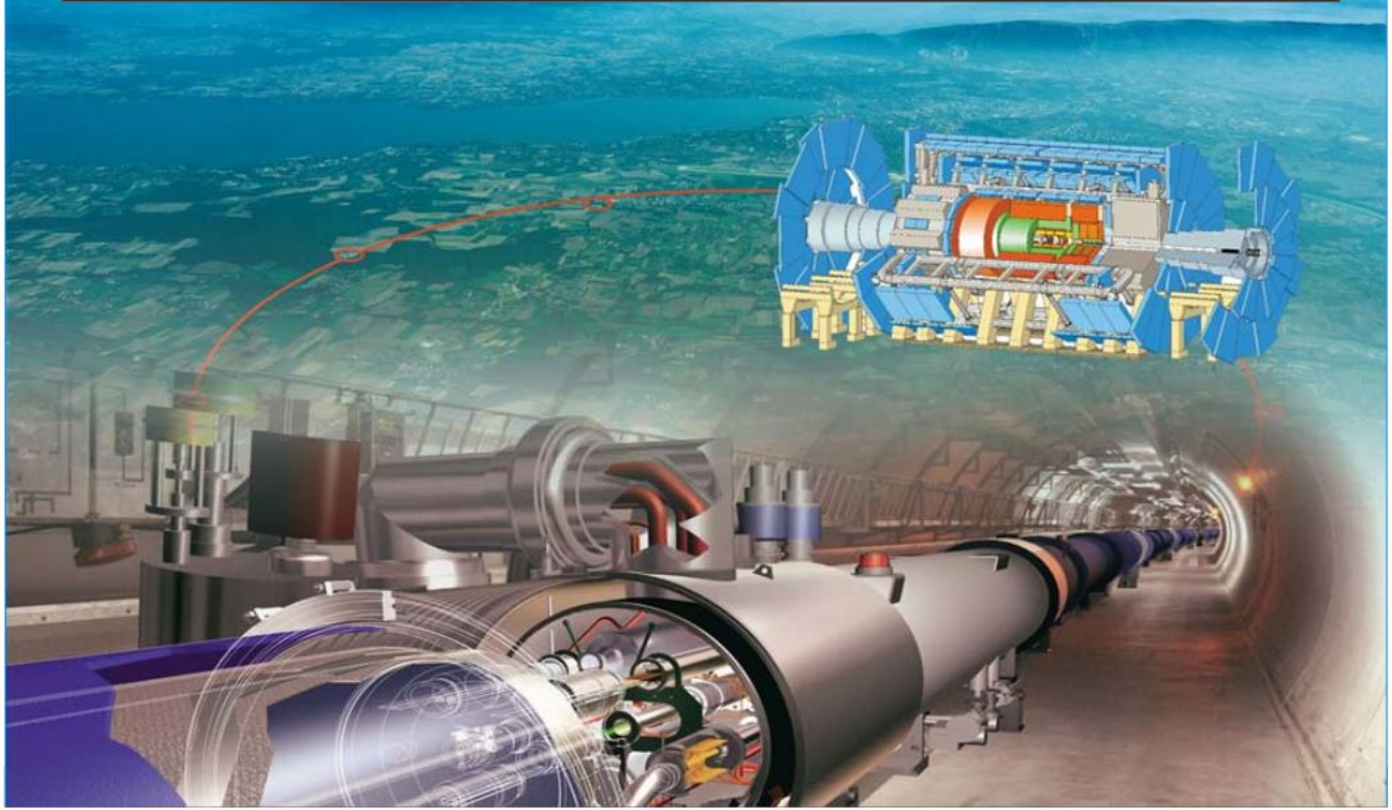


Table of content

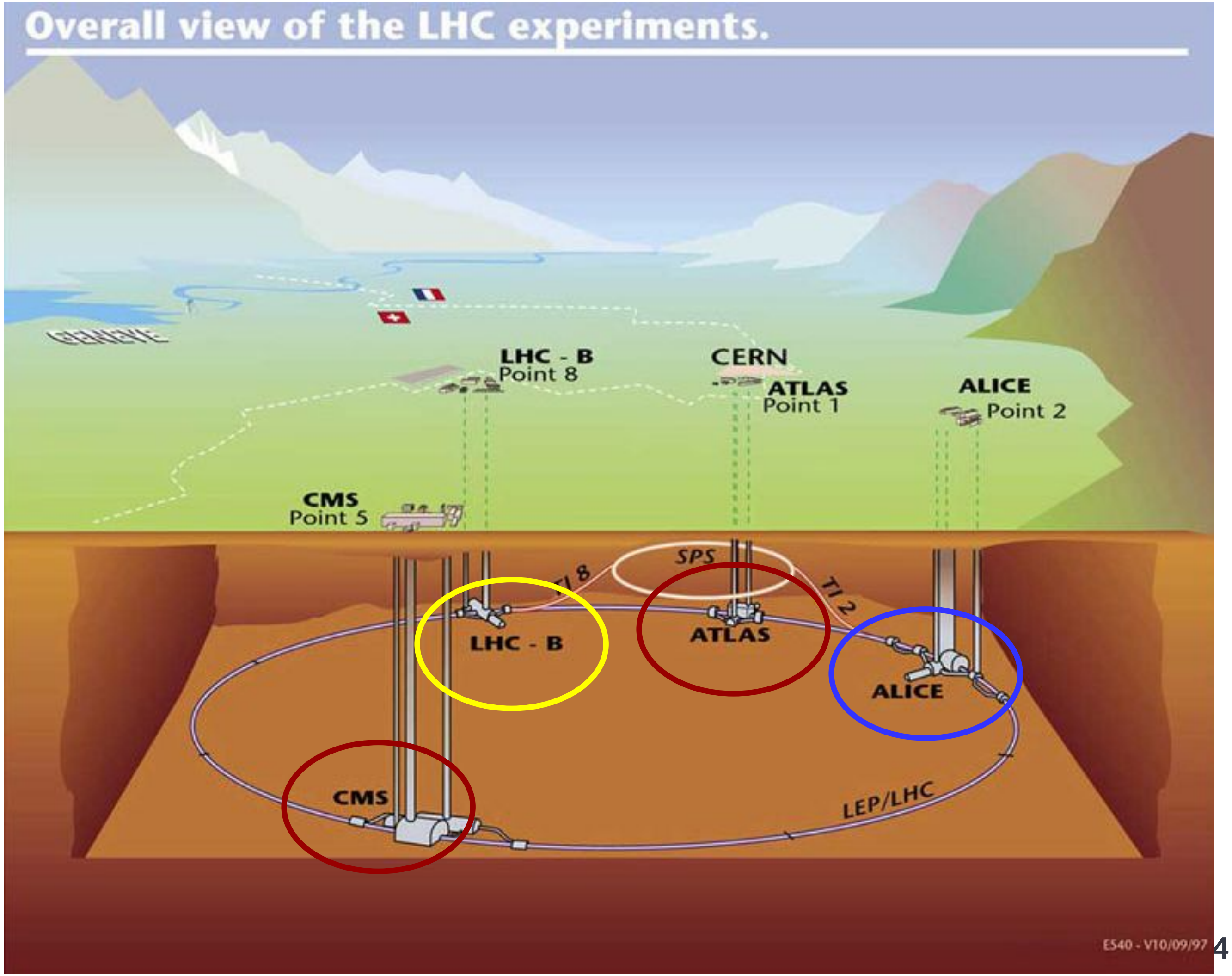
- **The physics program at the Large Hadron Collider**
    - A short overview
    - High energy physics (HEP) data analysis
        - Basics of Monte Carlo simulations
        - Statistics for Pedestrians

- **Literature**
    - Monte Carlo simulations:
      Introduction to parton-shower event generators
    - Statistics:
      LHC Statistics for Pedestrians
      R. J. Barlow: **Statistics**, WILEY, 1997

B-hadron physics

Heavy Ions

Multi-purpose



Physics program at ATLAS

B-Physics and light states

Standard Model physics

Top quark physics

Higgs boson physics

Heavy Ion Physics

Exotics Searches

Extended Higgs sector searches

Supersymmetry searches

ATLAS EXPERIMENT

Public web page listing all recent results: https://twiki.cern.ch/twiki/bin/view/AtlasPublic

Physics program at CMS

B-Physics and Quarkonia

Standard Model physics

Top quark physics

Higgs boson physics

Heavy Ion Physics

Exotics Searches

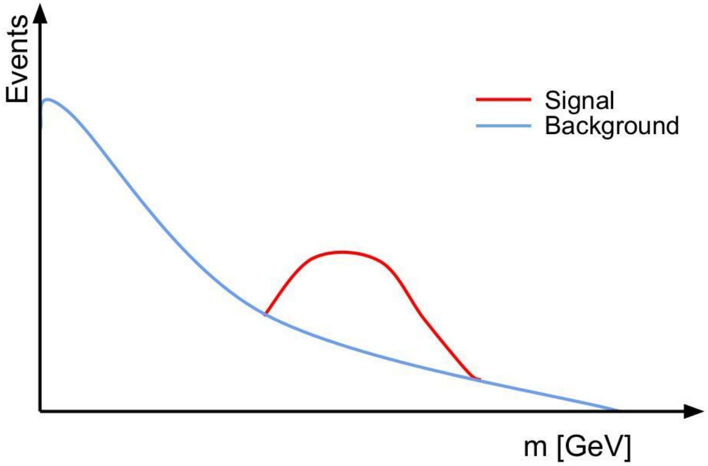
Supersymmetry searches

CMS

Public web page listing all recent results: http://cms-results.web.cern.ch/cms-results/public-results/publications/

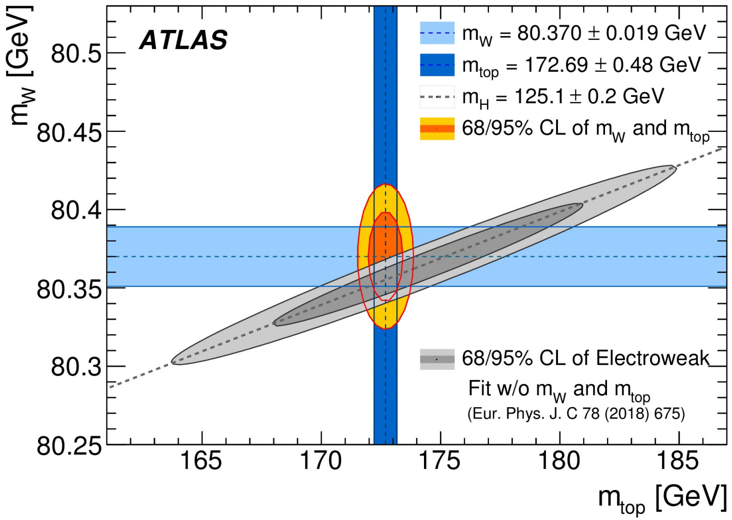
Measurements vs. Searches

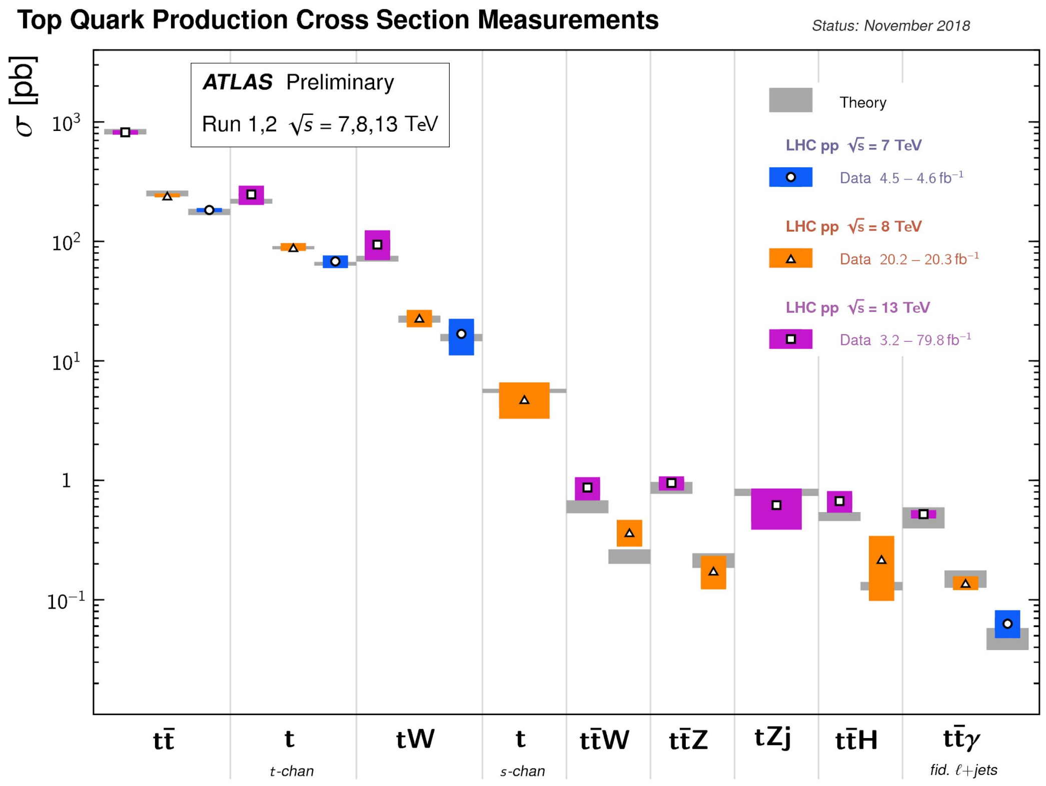
- **Measurements**: Study of properties of already know particles
  - Long-term efforts involving in-depth examinations of all relevant systematic uncertainties
- **Searches:** Probe for excess (or peak from resonant production) in data due to contribution from new particle
  - Often performed in "extreme" phase space regions:
    - At high p_T
    - Large (b-)jet multiplicities
    - Large amount of missing transverse momentum
    - Multi-lepton final states



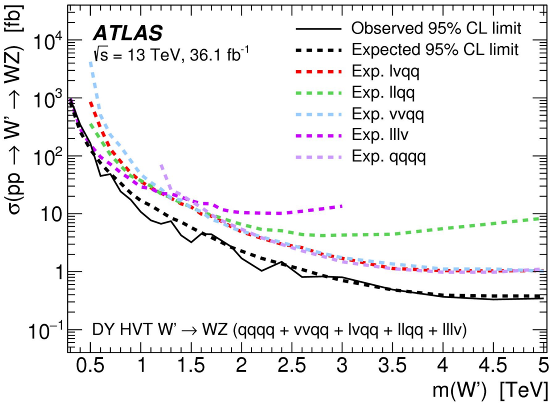
Measurements

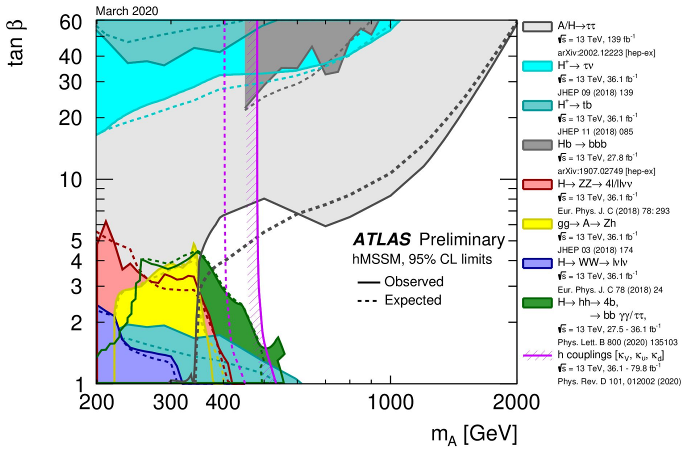
- Aim to determine:
  - Cross sections
  - Couplings
  - Masses
  - Spin/CP
  - Polarisation

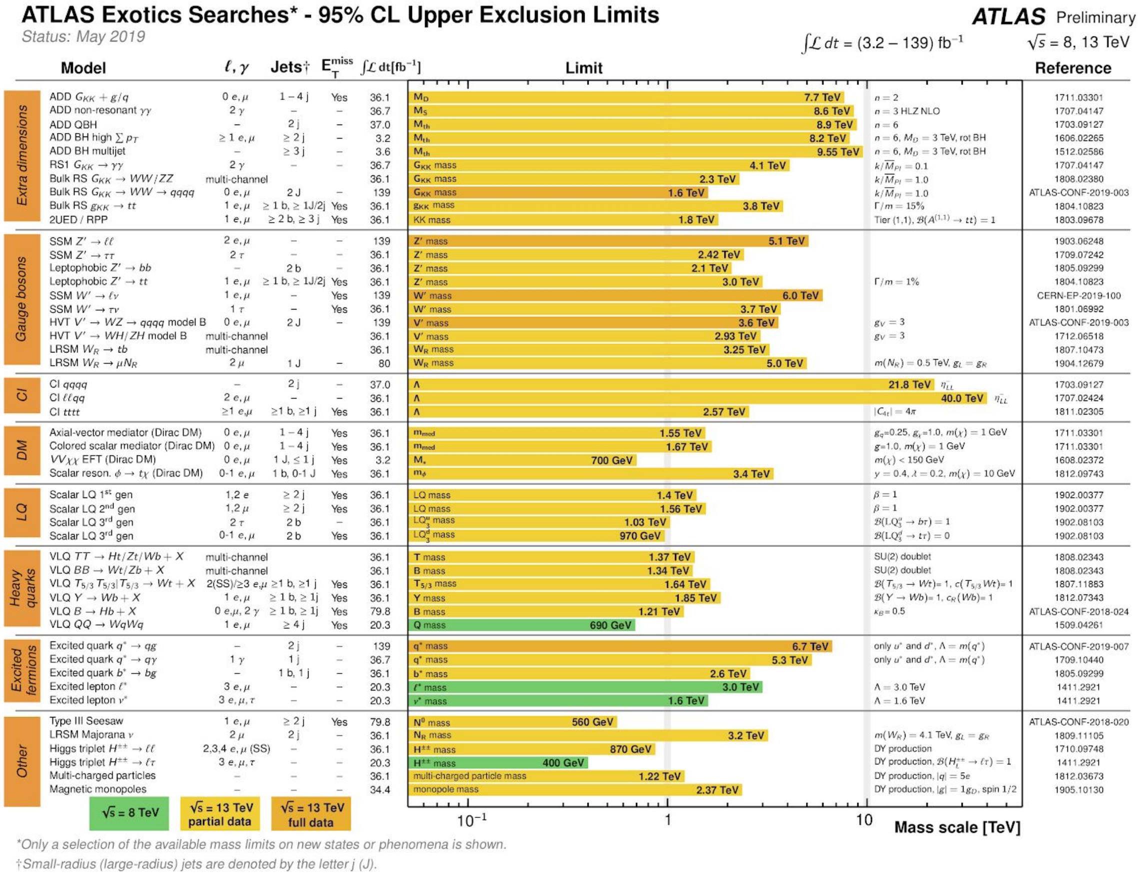




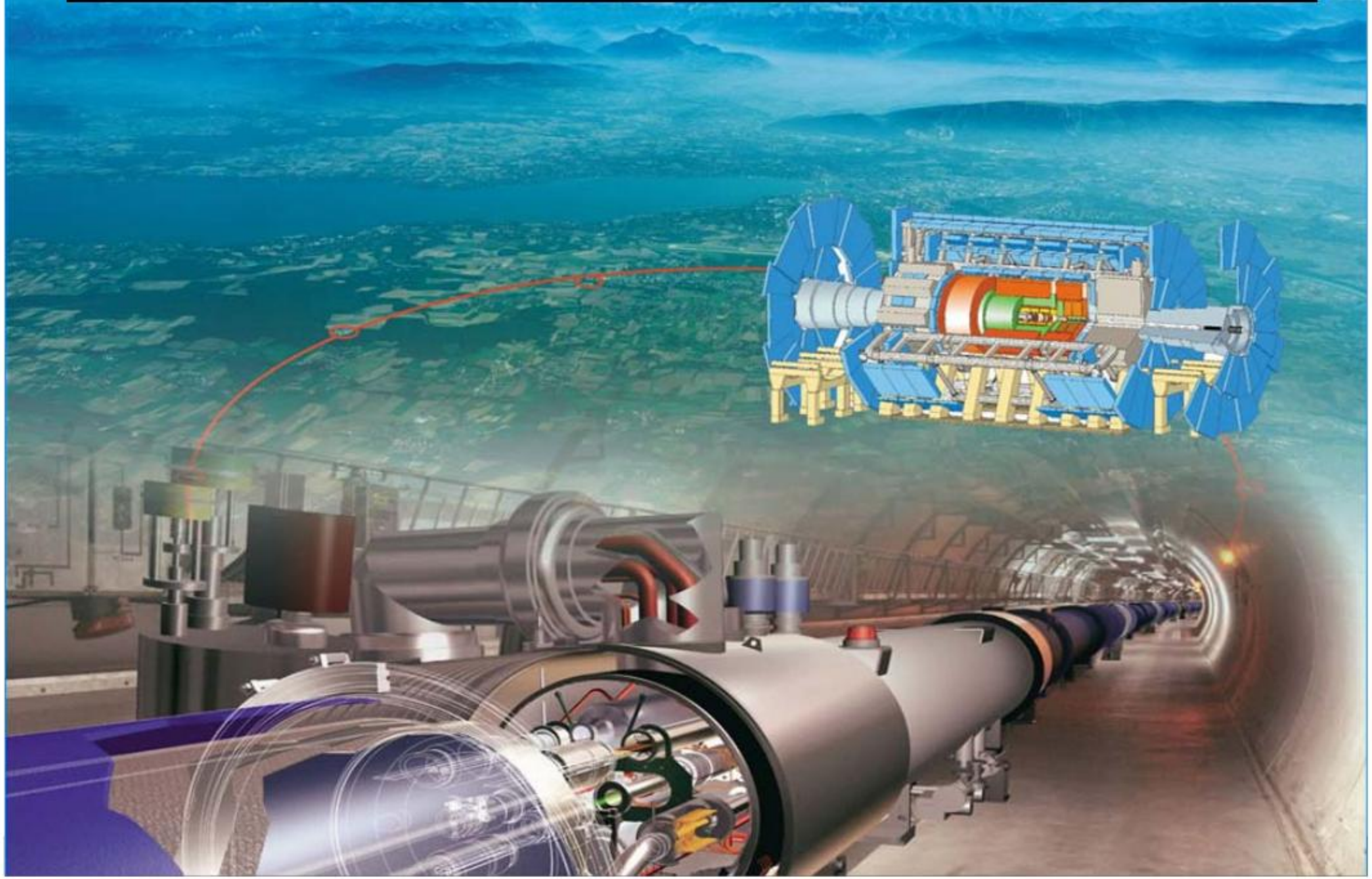
Searches


ATLAS
√s = 13 TeV, 36.1 fb-1
Observed 95% CL limit
Expected 95% CL limit
Exp. lvqq
Exp. llqq
Exp. vvqq
Exp. lllv
Exp. qqqq
DY HVT W' → WZ (qqqq + vvqq + lvqq + llqq + lllv)
σ(pp → W' → WZ) [fb]
m(W') [TeV]


March 2020
tan β
m_A [GeV]
ATLAS Preliminary
hMSSM, 95% CL limits
Observed
Expected
A/H→ττ
√s = 13 TeV, 139 fb-1
arXiv:2002.12223 [hep-ex]
H+→ τν
√s = 13 TeV, 36.1 fb-1
JHEP 09 (2018) 139
H+→ tb
√s = 13 TeV, 36.1 fb-1
JHEP 11 (2018) 085
Hb → bbb
√s = 13 TeV, 27.8 fb-1
arXiv:1907.02749 [hep-ex]
H→ ZZ→ 4l/llvv
√s = 13 TeV, 36.1 fb-1
Eur. Phys. J. C (2018) 78: 293
gg→ A→ Zh
√s = 13 TeV, 36.1 fb-1
JHEP 03 (2018) 174
H→ WW→ lvlv
√s = 13 TeV, 36.1 fb-1
Eur. Phys. J. C 78 (2018) 24
H→ hh→ 4b,
→ bb γγ/ττ,
√s = 13 TeV, 27.5 - 36.1 fb-1
Phys. Lett. B 800 (2020) 135103
h couplings [κV, κu, κd]
√s = 13 TeV, 36.1 - 79.8 fb-1
Phys. Rev. D 101, 012002 (2020)


ATLAS Exotics Searches* - 95% CL Upper Exclusion Limits
Status: May 2019
ATLAS Preliminary
∫L dt = (3.2 – 139) fb−1
√s = 8, 13 TeV
Model
ℓ, γ
Jets†
E_T^miss
∫L dt[fb−1]
Limit
Reference
Extra dimensions
Gauge bosons
CI
DM
LQ
Heavy quarks
Excited fermions
Other
√s = 8 TeV
√s = 13 TeV partial data
√s = 13 TeV full data
Mass scale [TeV]
*Only a selection of the available mass limits on new states or phenomena is shown.
†Small-radius (large-radius) jets are denoted by the letter j (J).

4.2.1 Basics of Monte Carlo simulations



Introduction

- Hadron colliders are discovery machines.
  - **Advantage:**
    - Proton beams can be accelerated to higher kinetic energies than electron beams
  - **Disadvantage:**
    - The event structure at hadron colliders is significantly more complex than at lepton colliders (due to the composite nature of the beam particles)

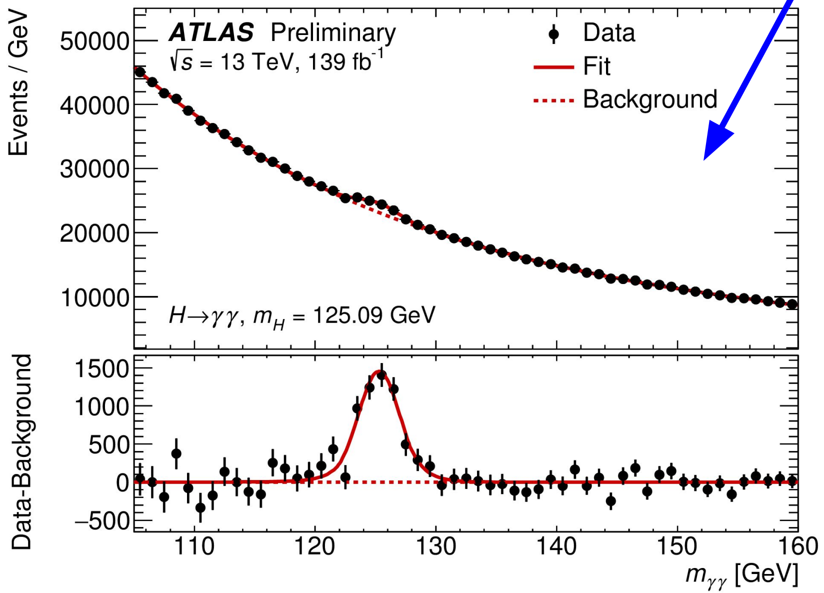
- **The description of full final states (from parton collisions) necessitates involved multi-particle calculations**
  - The high-dimensional phase space leaves Monte-Carlo integration as the only viable option.

→ Monte Carlo Method

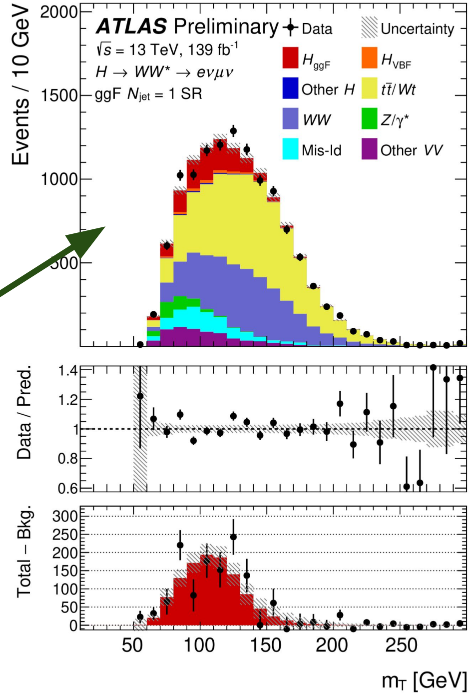
Monte Carlo simulations

- Almost all data analyses at the LHC experiments rely (at least to some extent) on Monte Carlo (MC) simulations

Validation of signal and/or background estimations



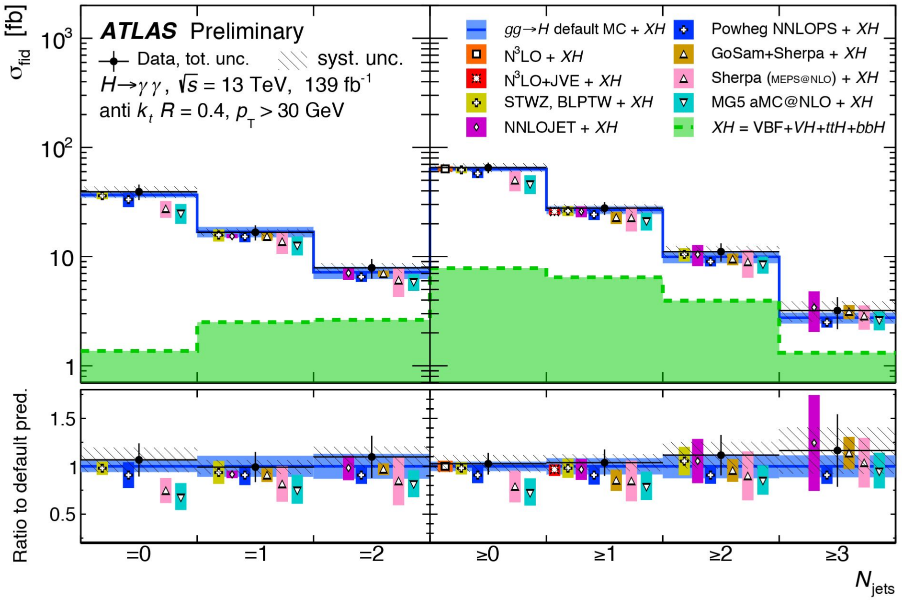
Estimation of contributions from signal and background processes

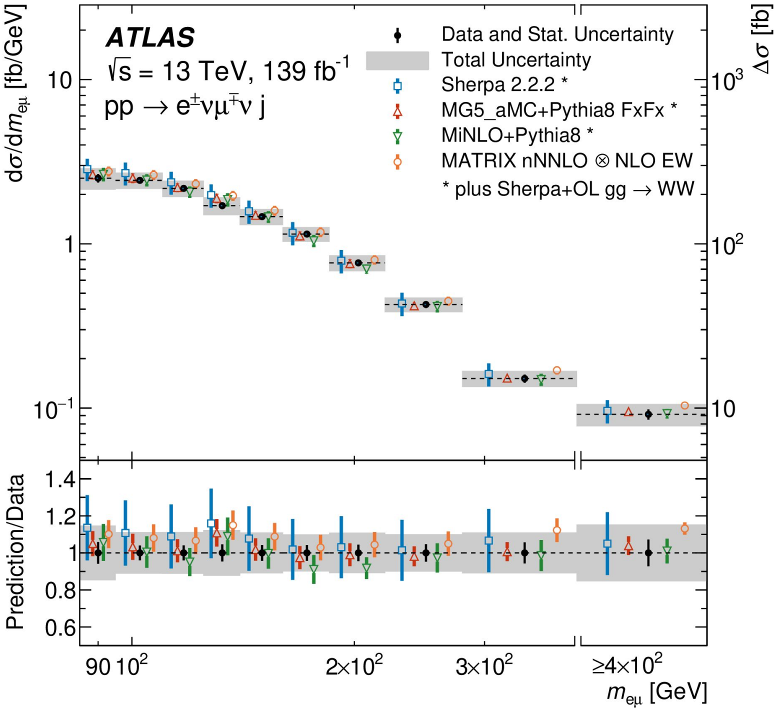


Monte Carlo simulations

- Need good description of SM processes
  - To be sensitive to new physics

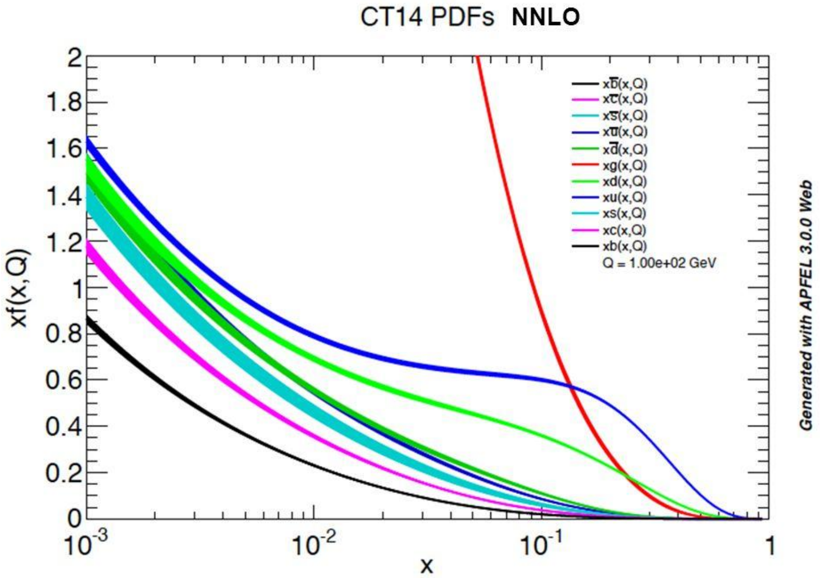
→ Tune Monte Carlo models

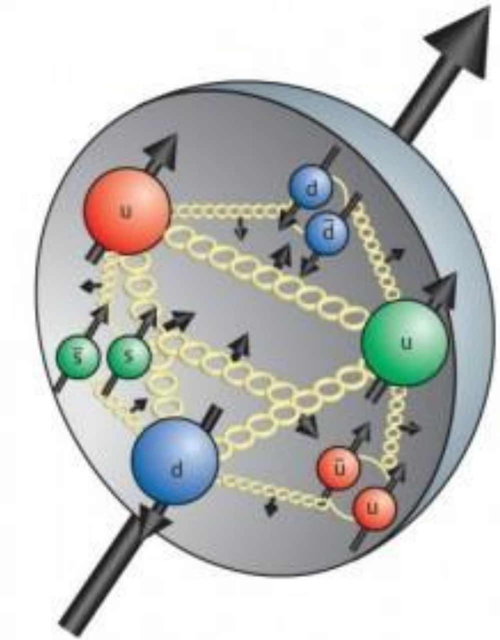




Monte Carlo simulations

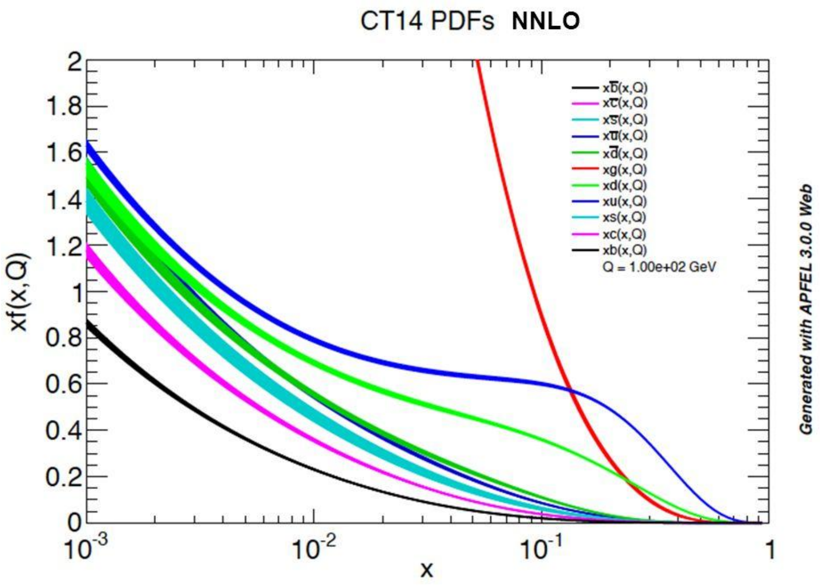
- Observations in data are compared to SM predictions (Monte Carlo simulations)
- **Use factorisation approach**:
  - Parton distribution functions (PDF)
  - Hard process (matrix element/scattering amplitude)
  - Parton shower (fragmentation, hadronization, decay of unstable particles)
  - Detector simulation (including overlay with pile-up)

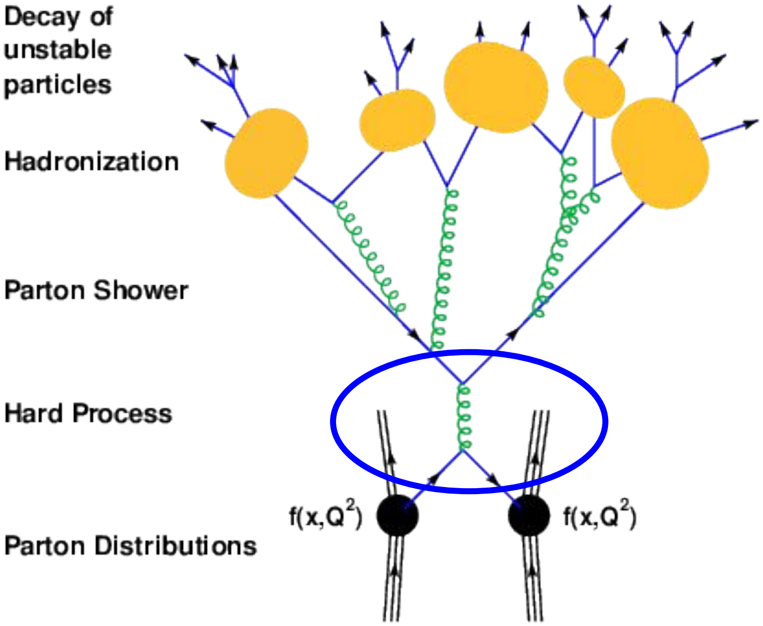




Monte Carlo simulations

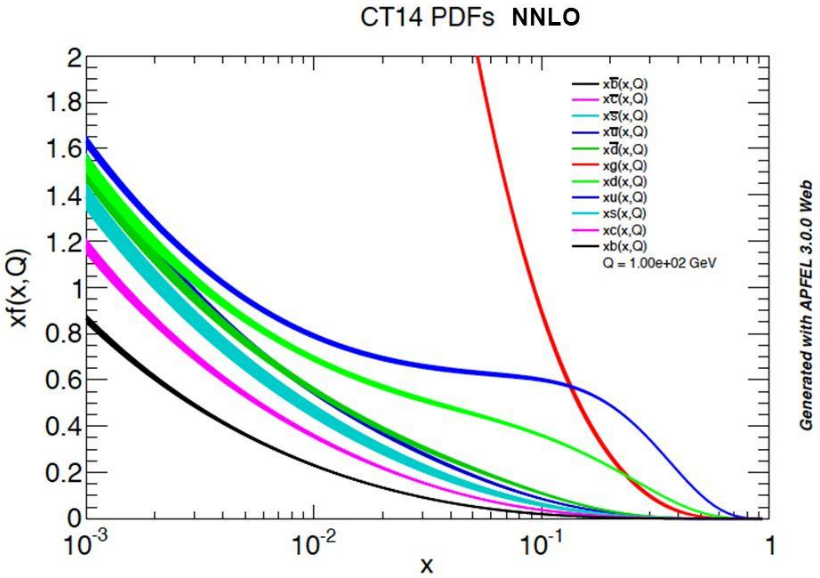
- Observations in data are compared to SM predictions (Monte Carlo simulations)
- **Use factorisation approach**:
    - Parton distribution functions (PDF)
    - Hard process (matrix element/scattering amplitude)
    - Parton shower (fragmentation, hadronization, decay of unstable particles)
    - Detector simulation (including overlay with pile-up)

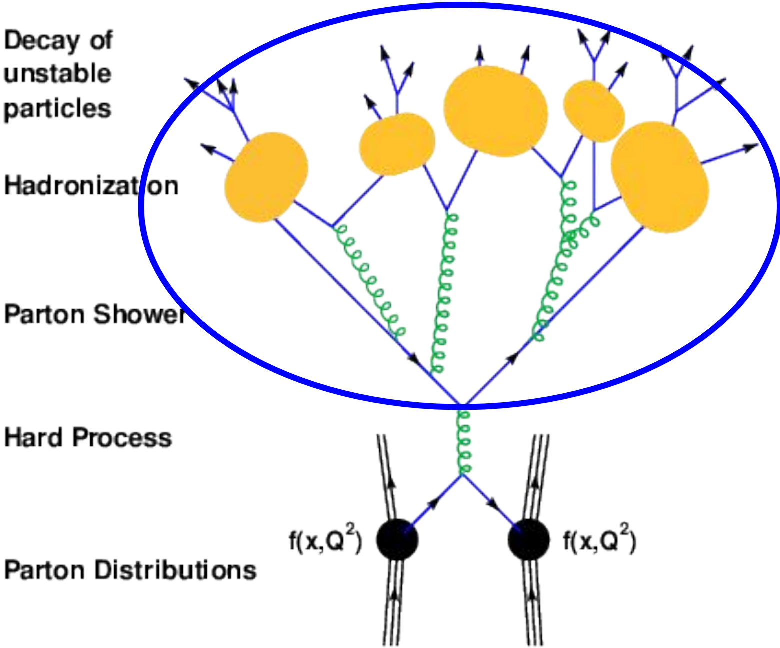




Monte Carlo simulations

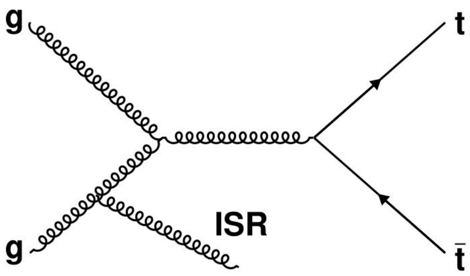
- Observations in data are compared to SM predictions (Monte Carlo simulations)
- **Use factorisation approach**:
    - Parton distribution functions (PDF)
    - Hard process (matrix element/scattering amplitude)
    - Parton shower (fragmentation, hadronization, decay of unstable particles)
    - Detector simulation (including overlay with pile-up)

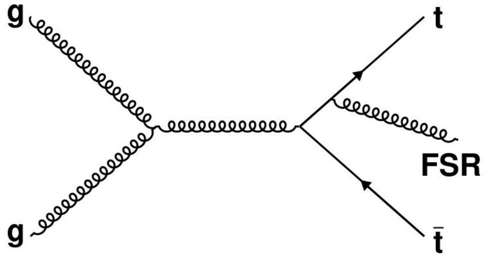




Monte Carlo simulations

- **Initial state radiation:**
  - Radiative emissions from the initial state partons
- **Final state radiation:**
  - Radiative emissions from the final state partons
- **Multiparton interactions (MPI) / underlying event (UE):**
  - Particle production not associated with the leading hardest parton-parton interaction (same protons)
- **Pile-up:**
  - Particle production not associated with the leading hardest parton-parton interaction (different protons)





Monte Carlo simulations

$\mu \approx \lambda_{QCD}$ Decay of unstable particles

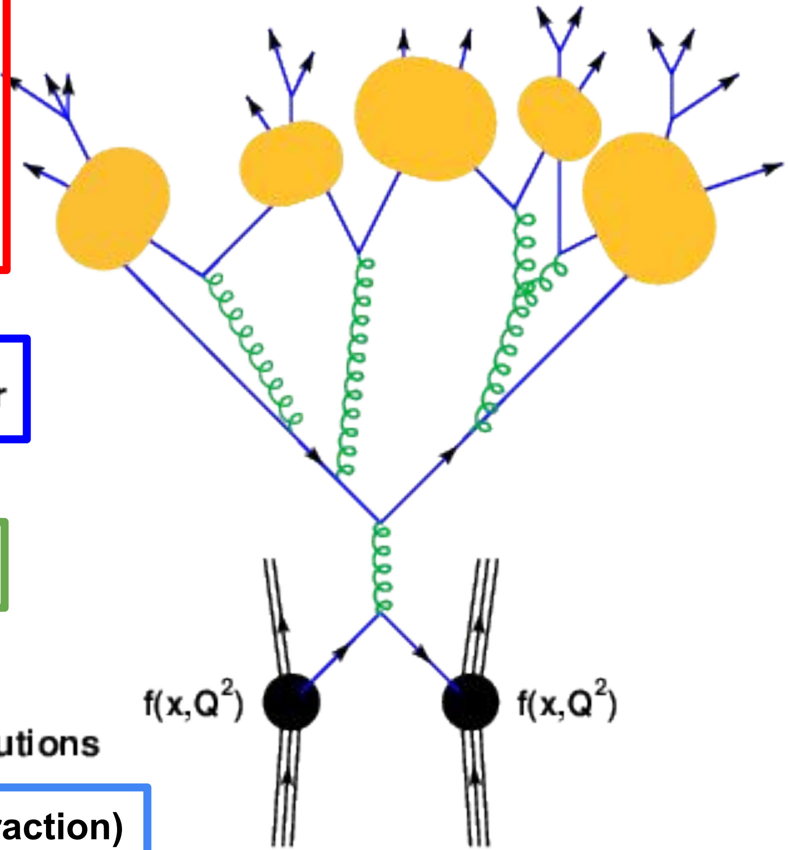
Hadronization

$\lambda_{QCD} < \mu < Q$ Parton Shower

$\lambda_{QCD} \ll \mu \approx Q$ Hard Process

Parton Distributions

(Multiparton Interaction)



Monte Carlo simulations

Parton distribution functions (PDFs)

Determined from data (as PDF sets)

Hard scattering

$\lambda_{QCD} \ll \mu \approx Q$

Hard process computed using fixed order (FO) or mixed order perturbation theory at LO, NLO, NNLO, N3LO,…

Parton shower

$\lambda_{QCD} < \mu < Q$

Hierarchy of scales appearing as large logarithms in the calculation. Computed using resummation or MC parton showers (PS)

Hadronization and hadron decay

$\mu \approx \lambda_{QCD}$

Multiparton interaction

Cannot be computed directly: non-perturbative models tuned to e^+e^- data

Monte Carlo simulations

Parton distribution functions (PDFs) → **Determined from data (as PDF sets)**

Hard scattering

$\lambda_{QCD} \ll \mu \approx Q$

→ **Hard process computed using fixed order (FO) or mixed order perturbation theory at LO, NLO, NNLO, N3LO,…**

+ ME and PS matching (different order → different schemes)

Parton shower

$\lambda_{QCD} < \mu < Q$

→ **Hierarchy of scales appearing as large logarithms in the calculation. Computed using resummation or MC parton showers (PS)**

Hadronization and hadron decay

$\mu \approx \lambda_{QCD}$

Multiparton interaction

→ **Cannot be computed directly: non-perturbative models tuned to e^+e^- data**

Master formula for hadron collisions

$$\sigma_{h_1+h_2\to X} = \sum_{a,b}\int \underbrace{dx_1\,dx_2\,d\Phi}_{\text{phase - space integral}}\ \underbrace{f_a(x_1,\mu_F)\,f_b(x_2,\mu_F)}_{\text{Parton distribution function}}\ \underbrace{\hat{\sigma}_{a+b\to X}(\hat{s},\mu_F,\mu_R)}_{\text{Parton-level cross section}}$$

- The parton-level fixed order cross section can be computed as a series in perturbation theory, using the coupling constant as an expansion parameter

$$\hat{\sigma} = \sigma^{\mathrm{Born}}\left(1 + \frac{\alpha_S}{2\pi}\sigma^{(1)} + \left(\frac{\alpha_S}{2\pi}\right)^2\sigma^{(2)} + \left(\frac{\alpha_S}{2\pi}\right)^3\sigma^{(3)} + \ldots\right)$$

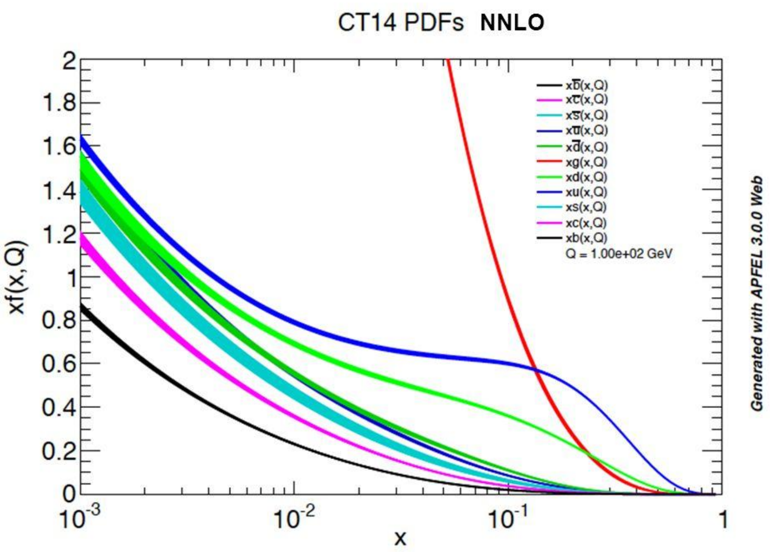
Coefficients of the perturbative series

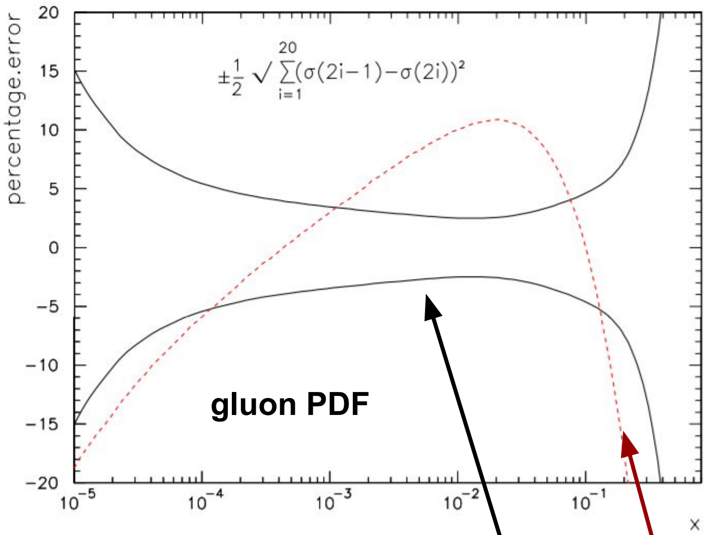
- Including higher corrections improves the predictions and reduces theoretical uncertainties

More information can be found e.g. via: https://arxiv.org/pdf/1207.2389v4.pdf

Particle distribution functions (PDFs)

- Quantify the probability density for finding a parton with a certain flavour and momentum fraction
- Obtained from fits to data
- Crucial source of uncertainties for both searches and measurements





Relative variations of the MSTW set

Variation of CTEQ6 set wrt central values of the MSTW set

Taken from https://arxiv.org/pdf/1207.2389v4.pdf

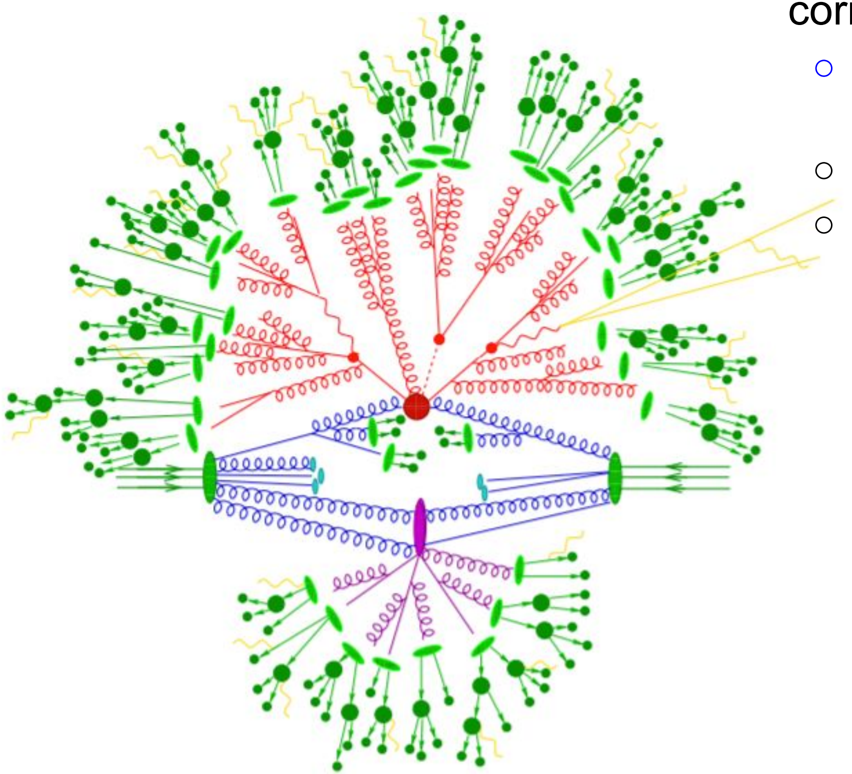
Particle distribution functions (PDFs)

- PDF sets can be downloaded from **LHAPDF** page (https://lhapdf.hepforge.org/pdfsets):

LHAPDF ID	Set name	Number of set members
251	GRVPI0	1
252	GRVPI1	1
270	xFitterPI_NLO_EIG	8
280	xFitterPI_NLO_VAR	6
10000	cteq6	41
10042	cteq6l1	1
10150	cteq61	41
10550	cteq66	45
10770	CT09MCS	1
10771	CT09MC1	1
10772	CT09MC2	1
10800	CT10	53

260000	NNPDF30_nlo_as_0118	101
260200	NNPDF30_nlo_as_0118_nf_3	101
260400	NNPDF30_nlo_as_0118_nf_4	101
260600	NNPDF30_nlo_as_0118_nf_6	101
260800	NNPDF30_nlo_as_0118_mc	101
261000	NNPDF30_nnlo_as_0118	101
261200	NNPDF30_nnlo_as_0118_nf_3	101
261400	NNPDF30_nnlo_as_0118_nf_4	101
261600	NNPDF30_nnlo_as_0118_nf_6	101
261800	NNPDF30_nnlo_as_0118_mc	101

Parton shower

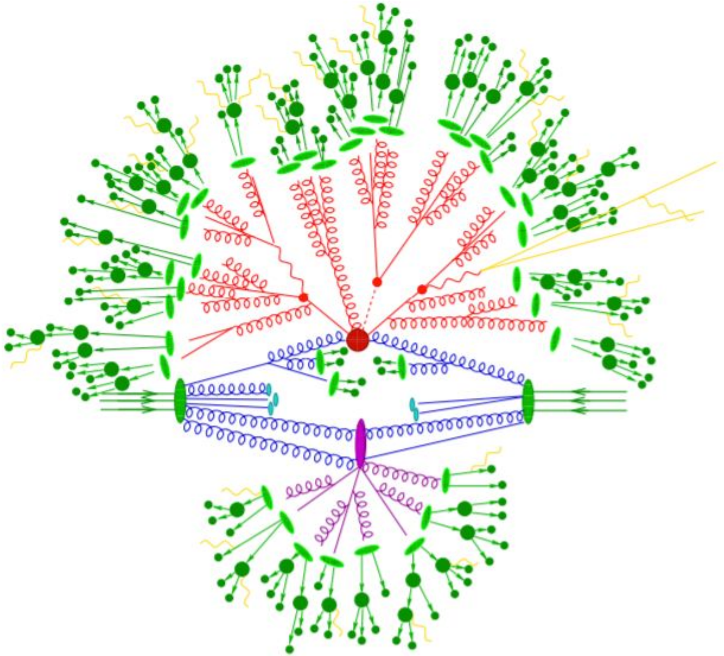


- Parton showers approximate higher-order real-emission corrections to the hard scattering
  - Simulation of the branching of a single external parton into two partons.
  - Local conservation of flavor and four momenta
  - Respect unitarity (i.e. a parton may either split into two partons, or not.

Hadronisation/Fragmentation

- To complete the simulation of realistic event topologies:
  - Quarks and gluons from the hard scattering simulations, parton showers and multiple scattering simulations must be transformed into **color-neutral final states.**

- Two different models are used nowadays:
  - **String model** (Lund model)
  - **Cluster model**



Monte Carlo generators

- **The Monte Carlo Method:**
  - Monte Carlo (MC) techniques are based on a repeated random sampling of numerical estimations of variables following complicated probability density functions
    - **Based on the implementation of (B)SM predictions**
- **Monte Carlo Event Generators** try to give the **best full description** (according our current knowledge) of a collision combining theoretical predictions for the **different stages** of an event and providing a **fully exclusive final state in terms of hadrons and leptons** which is as close as possible to what is measured in a real experiment
- Predictions are usually fed into a detector-simulation software to emulate the reconstruction effects of our real world detectors

See also: https://pdg.lbl.gov/2019/reviews/rpp2018-rev-monte-carlo-techniques.pdf

Some Monte Carlo generators

- MadGraph_aMC@NLO (https://launchpad.net/mg5amcnlo):
  - Tool for calculation of cross sections for SM and BSM processes and event generation (LO or NLO)
- POWHEG (http://powhegbox.mib.infn.it/):
  - MC generator for hard processes at NLO
- Sherpa (https://sherpa-team.gitlab.io/):
  - MC event generator for the simulation of ℓℓ, ℓγ, γγ, ℓh and hh collisions
- Pythia (http://home.thep.lu.se/Pythia/):
  - *Multi-purpose MC generator (for event generation and/or parton shower)*
    - *Supports* Lund string fragmentation model
- ALPGEN (https://arxiv.org/pdf/hep-ph/0206293.pdf):
  - MC generator for hard multiparton processes in hadronic collisions
- HERWIG (https://herwig.hepforge.org/):
  - *Multi-purpose MC generator (for event generation and/or parton shower)*
    - *Supports angular-ordered and dipole showers as well as MPI*
- MCFM (https://mcfm.fnal.gov/):
  - Tool dedicated to calculate cross sections of various processes at NLO (and NNLO) in QCD
    - Can also be used as event generator for some of these processes
- JHU (https://spin.pha.jhu.edu/):
  - Event generator for $pp \to X \to VV$, VBF, X+JJ, $pp \to VX$, $ee \to VX$

Detector simulations

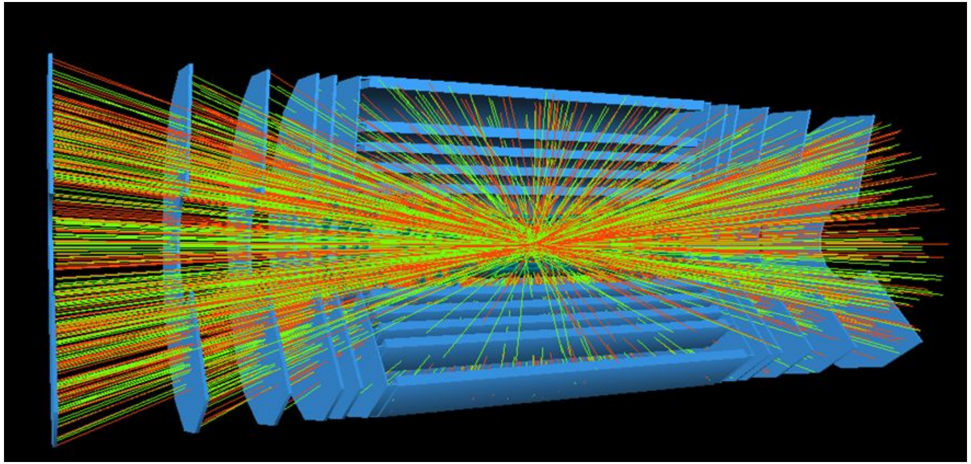
Taken from: https://cds.cern.ch/record/1300517/files/ATL-PHYS-PUB-2010-013.pdf

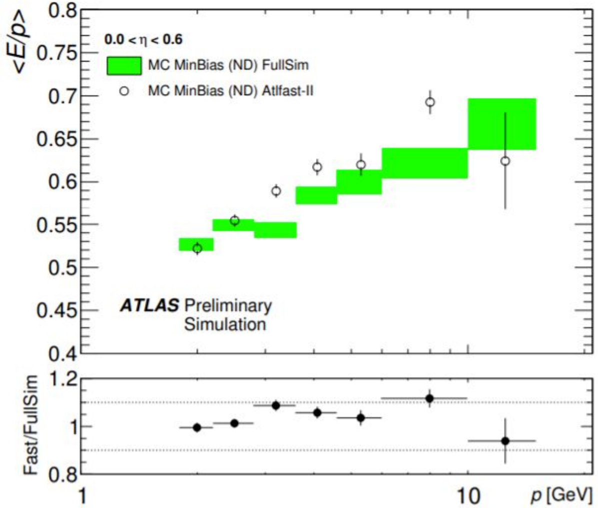
- Two different approaches:

  1) Full simulation: Includes a detailed description of the detector geometry and of the simulation of particle interactions in the detector material with **GEANT4**

  - The drawback of such a detailed simulation is a CPU time requirement of several minutes per event, of which more than 90% is spent inside the calorimeter systems.

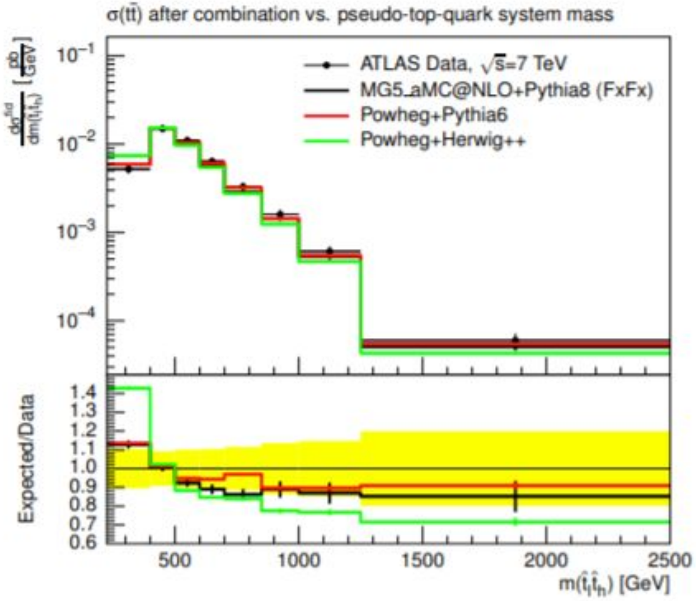
  2) Approximation (or better parameterized simulation) of the particle energy response and of the energy distribution in within the calorimeter system of (ATLAS, CMS,...)

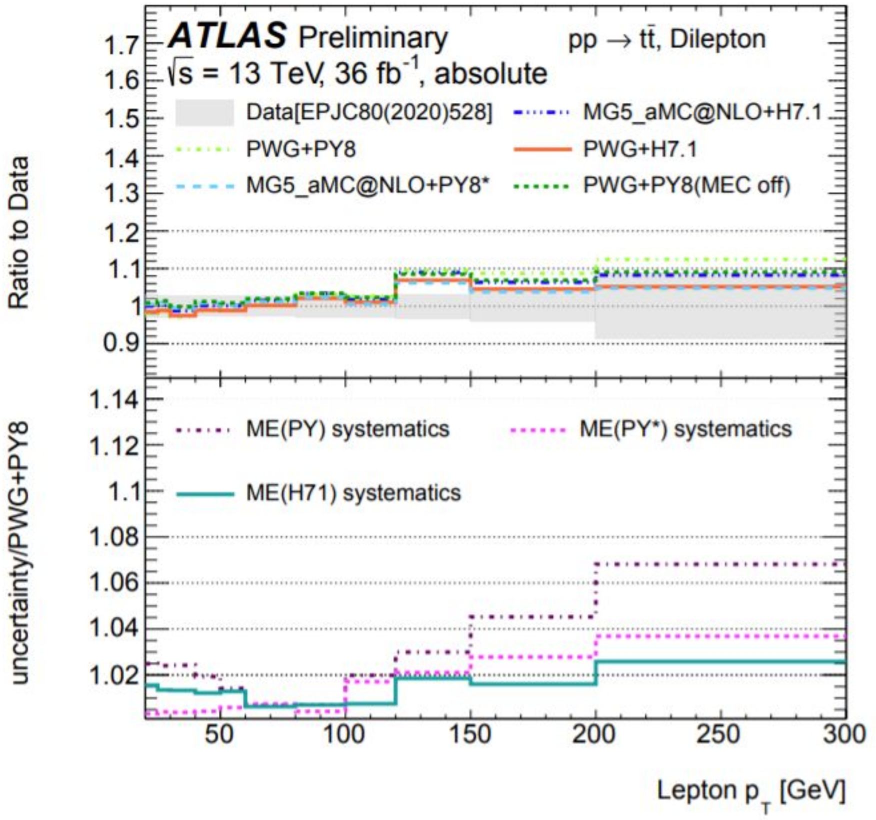




Tuning of Monte Carlo simulations

- Differential measurements are used:
  - To Compare different MC generator setups
  - To tune Monte Carlo models
    - In particular parton shower models



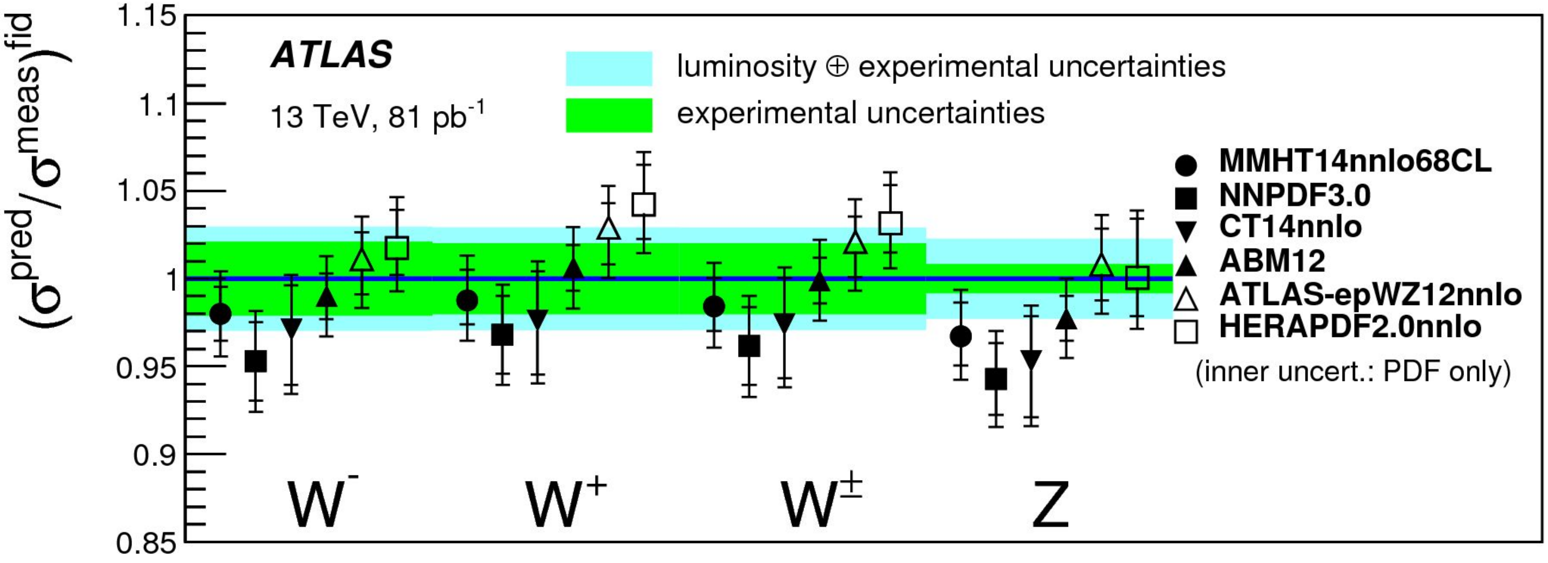


From https://cds.cern.ch/record/2730443/files/ATL-PHYS-PUB-2020-023.pdf

Tuning of Monte Carlo simulations

- Predictions from PDF sets are compared to data
  - To motivate choice of one particular set over the other

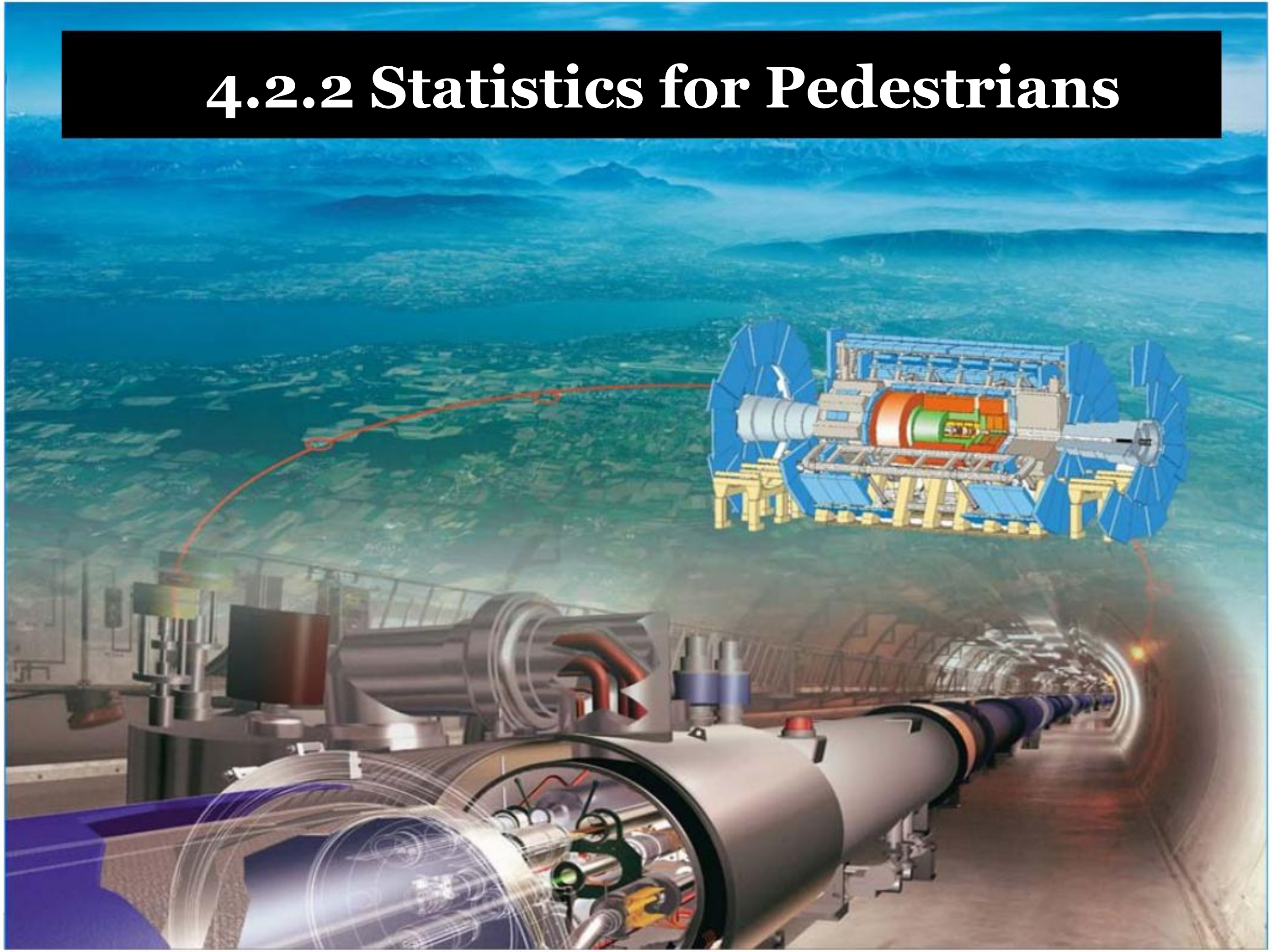
From: https://arxiv.org/abs/1603.09222



W & Z boson production cross section measurements are sensitive to the PDF sets

→ Can constrain them

4.2.2 Statistics for Pedestrians



Probability

- **What is probability?**
  - Frequentist approach:
    - **Definition:** If an experiment is performed N times with a specific outcome A occuring M times, then:

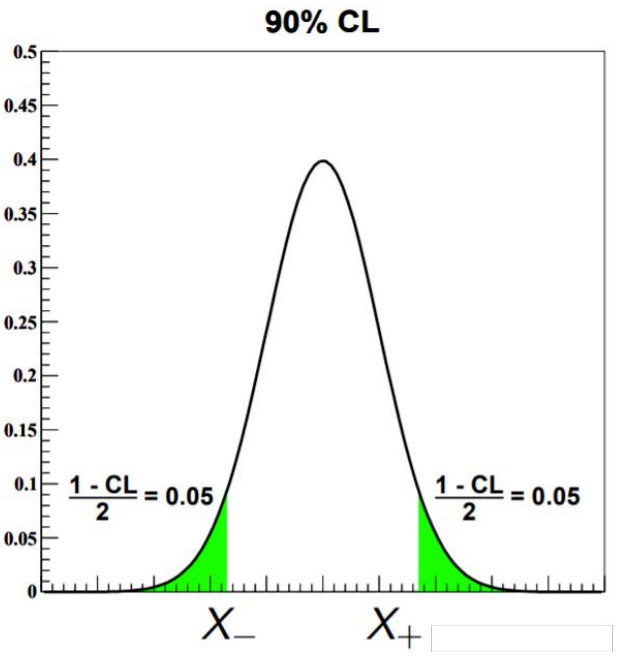
$$P(A) = \lim_{N \to \infty} \frac{M}{N}$$

    - **Interpretation:** If the experiment is performed once more, the outcome A will occur with the probability P(A).
  - Bayesian approach (subjective probability):
    - **Definition of Conditional probability P(A|B)**: The probability of A to occur under the condition that B has occurred.

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

Confidence Levels

- **Confidence level intervals:**
  - Given a precisely known true value μ of a certain property we can ask:
    - What is the range into which a certain amount (e.g. 90%) of measurements x_i will fall?



Example:

- Suppose cereal packets are produced according to a Gaussian distribution of mean 520g and standard deviation 10g.
  - 68% of all packets will weigh more than 510g and less than 530g
  - If we say that the weight of a packet lies in the interval 510g to 530g, we will be correct 68% of the time.

→ We can make that statement with 68% confidence

Confidence Levels

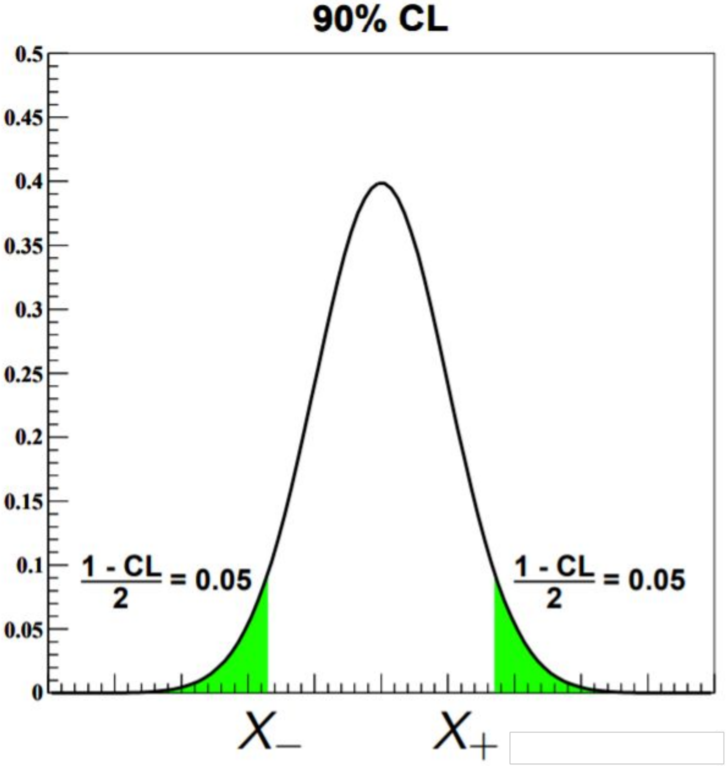
- **Central confidence intervals for a Gaussian distributions:**

$$P(X_- \geq x \geq X_+) = \int_{X_-}^{X_+} P(x)dx = CL$$

x: measurement
$X_\pm$: limits of the confidence interval.

- **Common values:**

1σ	$\hat{=}$	68.27%	1.6449σ	$\hat{=}$	90%
2σ	$\hat{=}$	95.45%	1.9600σ	$\hat{=}$	95%
3σ	$\hat{=}$	99.73%	2.5758σ	$\hat{=}$	99%
5σ	$\hat{=}$	99.99994%			



Confidence Levels

- **Central confidence intervals for a Gaussian distributions:**

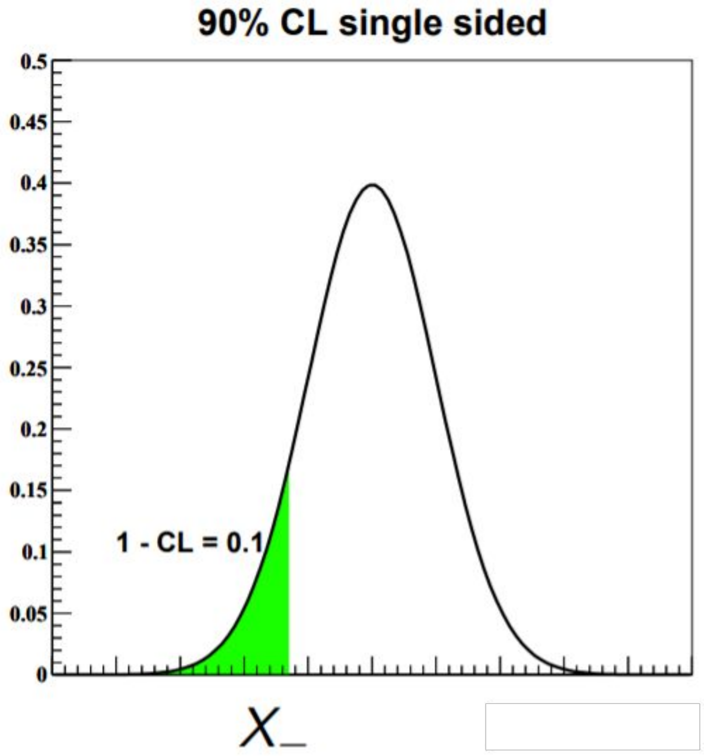
$$P(x \geq X_+) = \int_{-\infty}^{X_+} P(x)dx = CL_{\text{upper}}$$

$$P(X_- \geq x) = \int_{X_-}^{\infty} P(x)dx = CL_{\text{lower}}$$

x: measurement
$X_\pm$: limits of the confidence interval.

- **Common values:**

1σ	$\hat{=}$	84.13%	1.2816σ	$\hat{=}$	90%
2σ	$\hat{=}$	97.72%	1.6449σ	$\hat{=}$	95%
3σ	$\hat{=}$	99.87%	2.3263σ	$\hat{=}$	99%
5σ	$\hat{=}$	99.99997%			



Use a **one-sided confidence interval** to obtain the tightest upper (lower) bound on a sample mean

Test statistics & p-values

- A test statistic is a quantity calculated from our sample of data. Its value can be used to estimate how probable is the result that we observe with respect to some null hypothesis.
  - Usually: Null hypothesis = ’background only’ hypothesis
- Definition of tests statistics (simplified version in absence of uncertainties):

alternative hypothesis

$$q = \frac{L(H_1)}{L(H_0)} = \frac{L(s+b)}{L(b)}$$

via Likelihood ratio

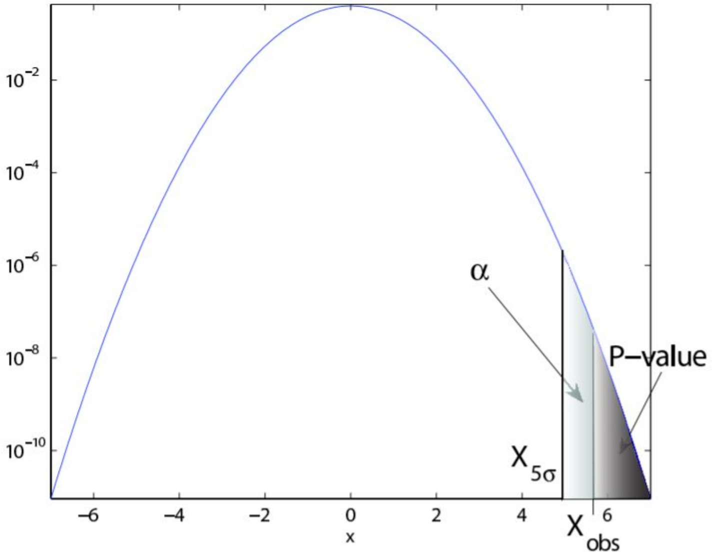
  - In a counting experiment μs and b would be the average number of the expected signal and background events and the Likelihoods would be derived from the data using Poisson statistics.

Test statistics & p-values

- Given the probability density function of the test statistic x for the background only hypothesis: g(x | H0)
- If the observed result $x_{obs} > x_{5\sigma}$ then the probability to get a result which is as or less compatible with the background hypothesis is given by:

$$p = \int_{x_{\rm obs}}^{\infty} g(x|H_0)dx$$

This probability is called the p − value



> In the presence of a signal, the background-only hypothesis is rejected with a probability of 1 − p.

Search procedure

- **The standard (frequentist) procedure to search for new physics processes follows:**
  1. Calculate a p value to test the null hypothesis that the data were generated by standard model processes alone.
  2. If $p \leq \alpha_1$ claim discovery and calculate a two-sided, 68% confidence level interval on the production cross section of the new process.
  3. If $p > \alpha_1$ calculate a 95% confidence level upper limit on the production cross section of the new process.
     - The purpose of reporting an upper limit when failing to claim a discovery is to exclude cross sections that the experiment is sensitive to and did not detect.

- Typical confidence levels are $\alpha_1 = 2.9 \times 10^{-7}$ (corresponds to 5σ)

Upper limit

- With the upper limit on a model parameter of interest (POI) μ, for a given observation N_{obs} we search for the largest value of μ for which the probability to make an observation of $N \leq N_{obs}$ is less than the value α.
  - In particle physics, we usually use $\alpha = 0.05$ (→ 95% CL limit)

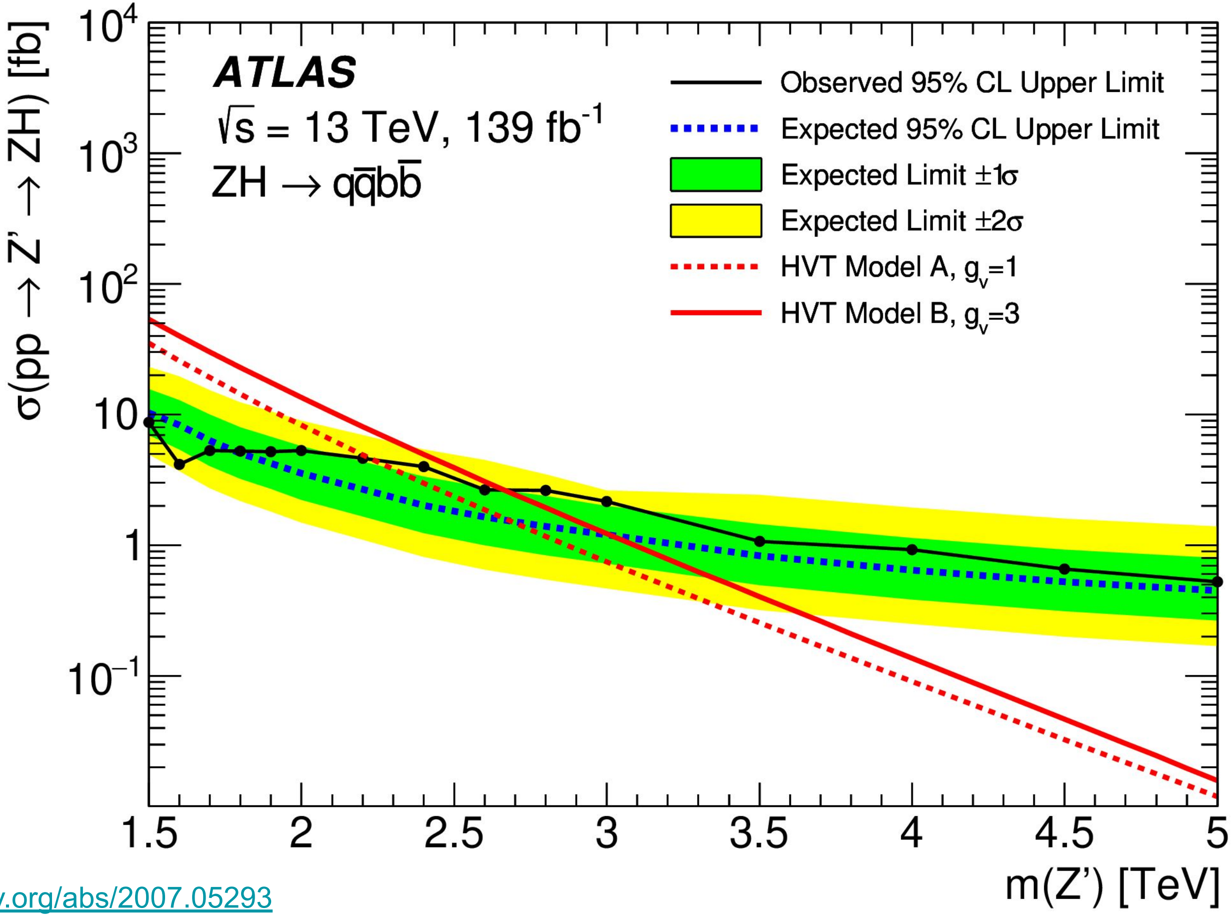
- For a given observation find the largest value for μ where N_{obs} is still contained in the interval

- μ is usually the signal strength parameter:

$$\mu = \frac{\sigma_{\text{observed}}}{\sigma_{\text{theory}}}$$

but could also be any other model parameter

Limits on the upper production cross section at the 95% C.L.



Taken from: https://arxiv.org/abs/2007.05293