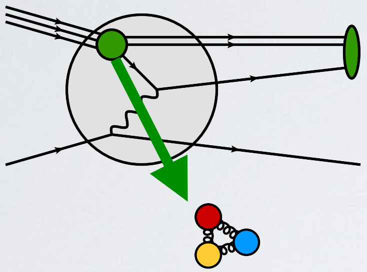
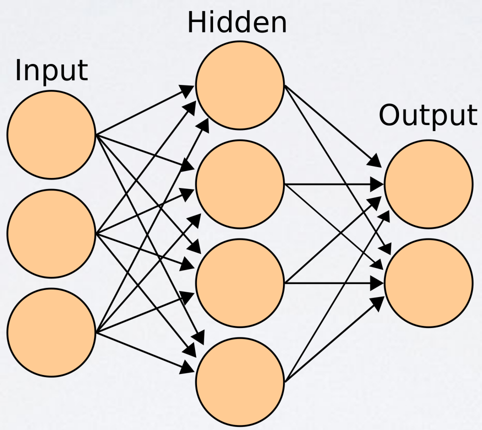
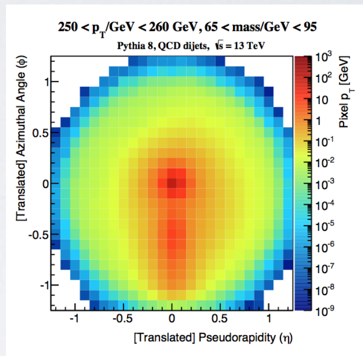
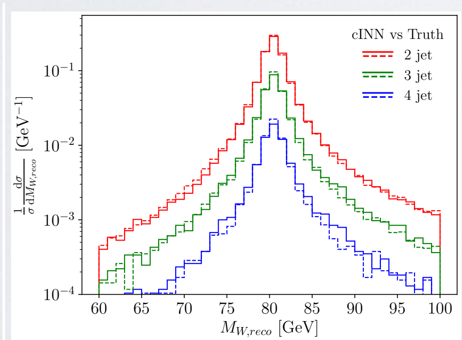
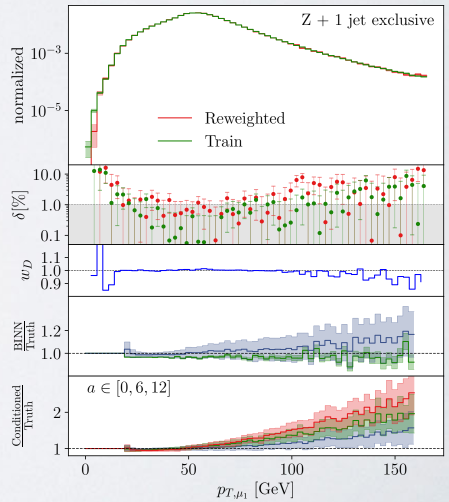
Machine Learning for applied and high energy physics











Prof. Simon Badger
EN 5.9
simondavid.badger at unito.it

Course overview

some motivation

High energy collider phenomenology

From field theory to experiment

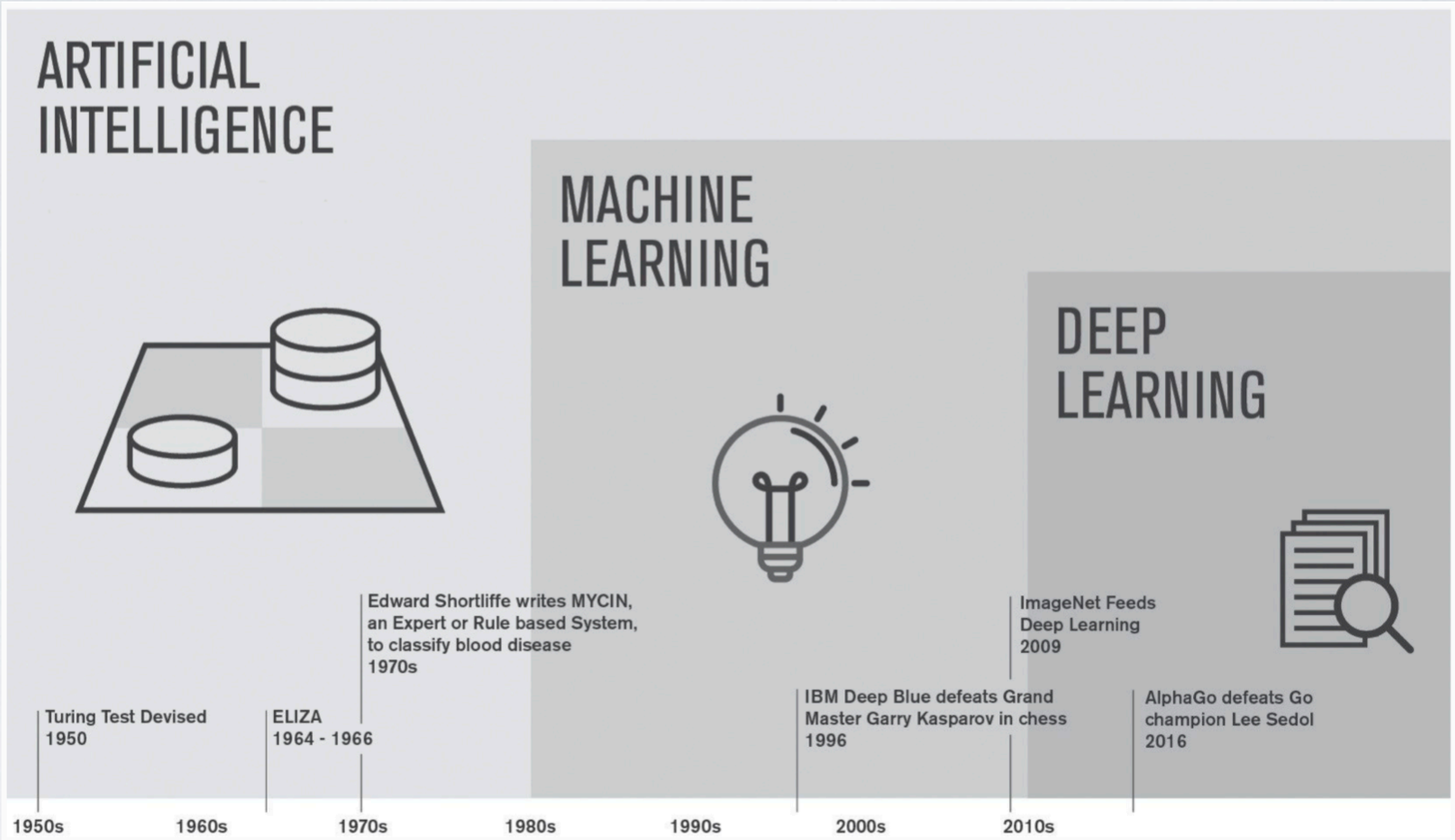
Simulations

The high cost of numeical integration

Machine Learning for fundamental physics

How can ML help us understand the laws of Nature

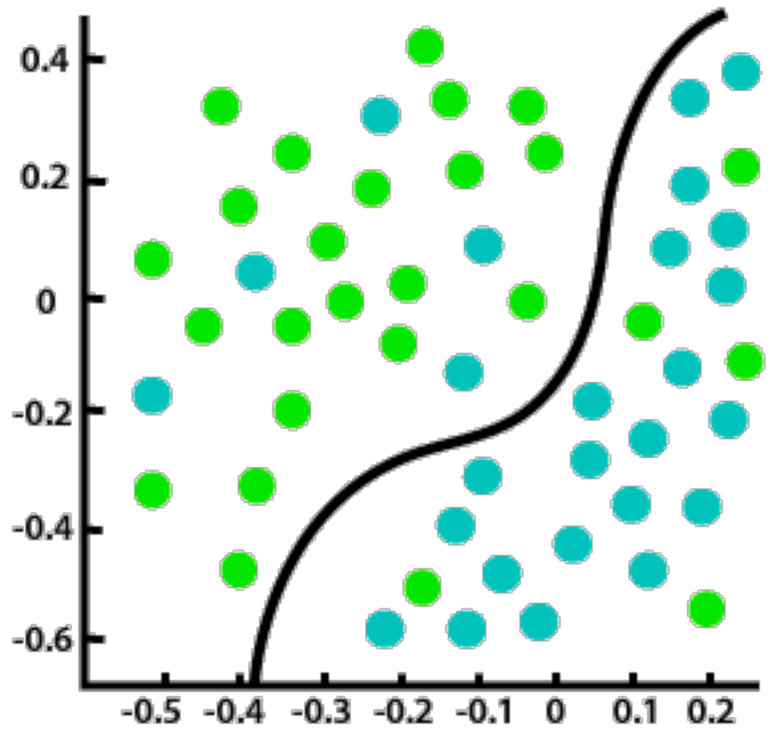
ML algorithms have been around for a while



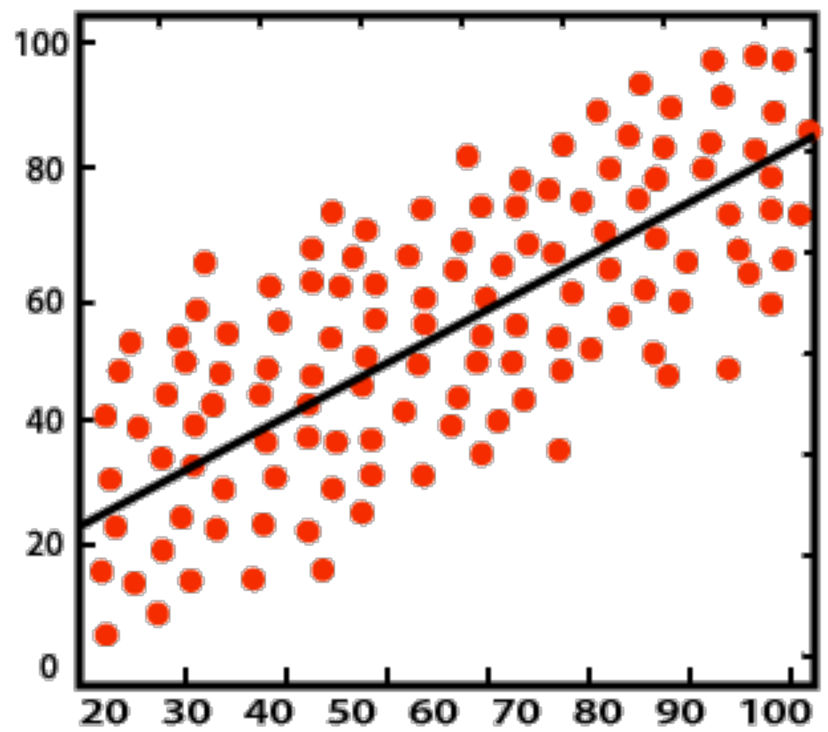
also in HEP

e.g. Finding Gluon Jets With a Neural Trigger,
Lonnblad, Petersen, Rognvaldsson (1990)

Supervised Learning



Classification



Regression

"Learn from examples"

the world of ML techniques can be a bit intimidating

Course overview

structure

- Introduction to collider simulations [1]

  What are the aims? What are the challenges?

- Overview of Machine Learning in HEP [1]

  What techniques are being used? Look at the Living Review. (Pandas dataframes)

- Non-linear regression: Neural Networks [2]

  Perceptron, analytic structure, training and backpropagation

- (Deep) Neural Networks [1]

  Optimised training methods, TensorFlow examples

Course overview

structure

- Logistic Regression for Classification [2]

  Cross-entropy, Soft-Max, ROC curves

- Application: Supersymmetry searches [1]

  signal/background discrimination

- Application: Optimising collider simulations [2]

  Phase space integration, Amplitude Neural Networks, Neural network errors

- Convolutional Neural Networks [1]

  2d image processing

- Application: Jet subtructure and Jet image classification [1]

  Jet algorithms, CNNs for boosted Higgs searches

Course overview

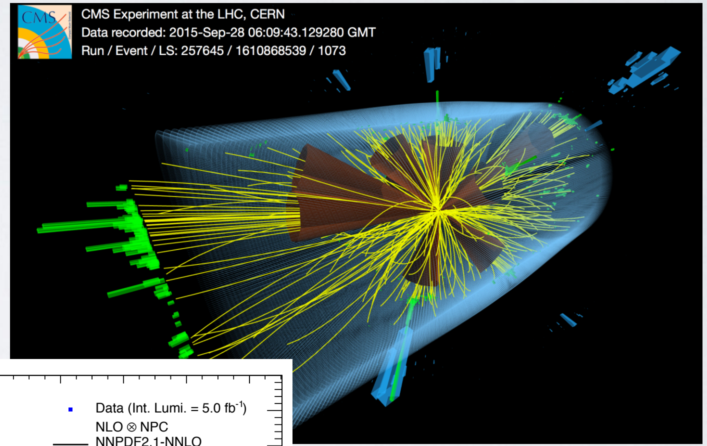
Exam

- Project based assessment
- Demonstrate understanding of course elements
  - choose the right tool for the problem at hand
  - technical skills: demonstrate ability to extend code from a given examples
- 15 minute presentation
  - physics background
  - methods
  - results

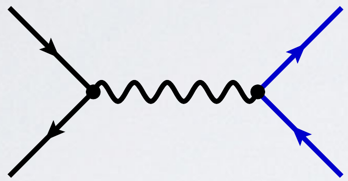
NB: no requirememnt to be original research - many examples in the 'Living Review' will be too much work

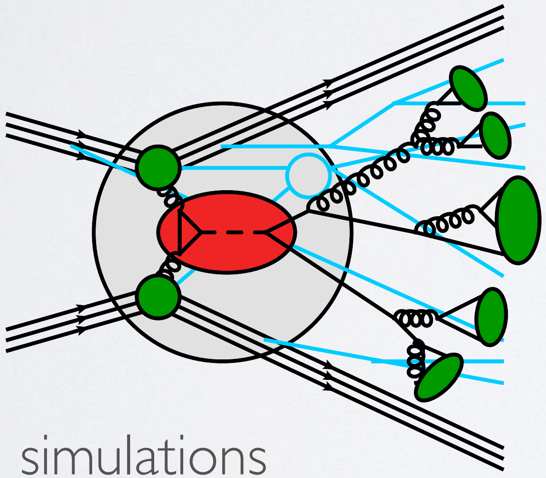
I. Introduction to collider physics

detectors

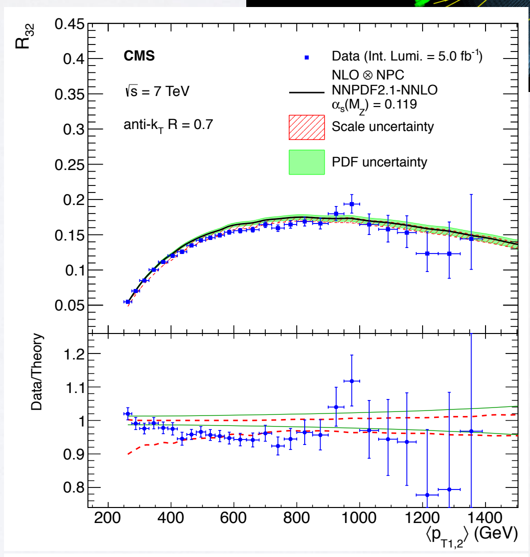


amplitudes

$\mathcal{A} =$ 

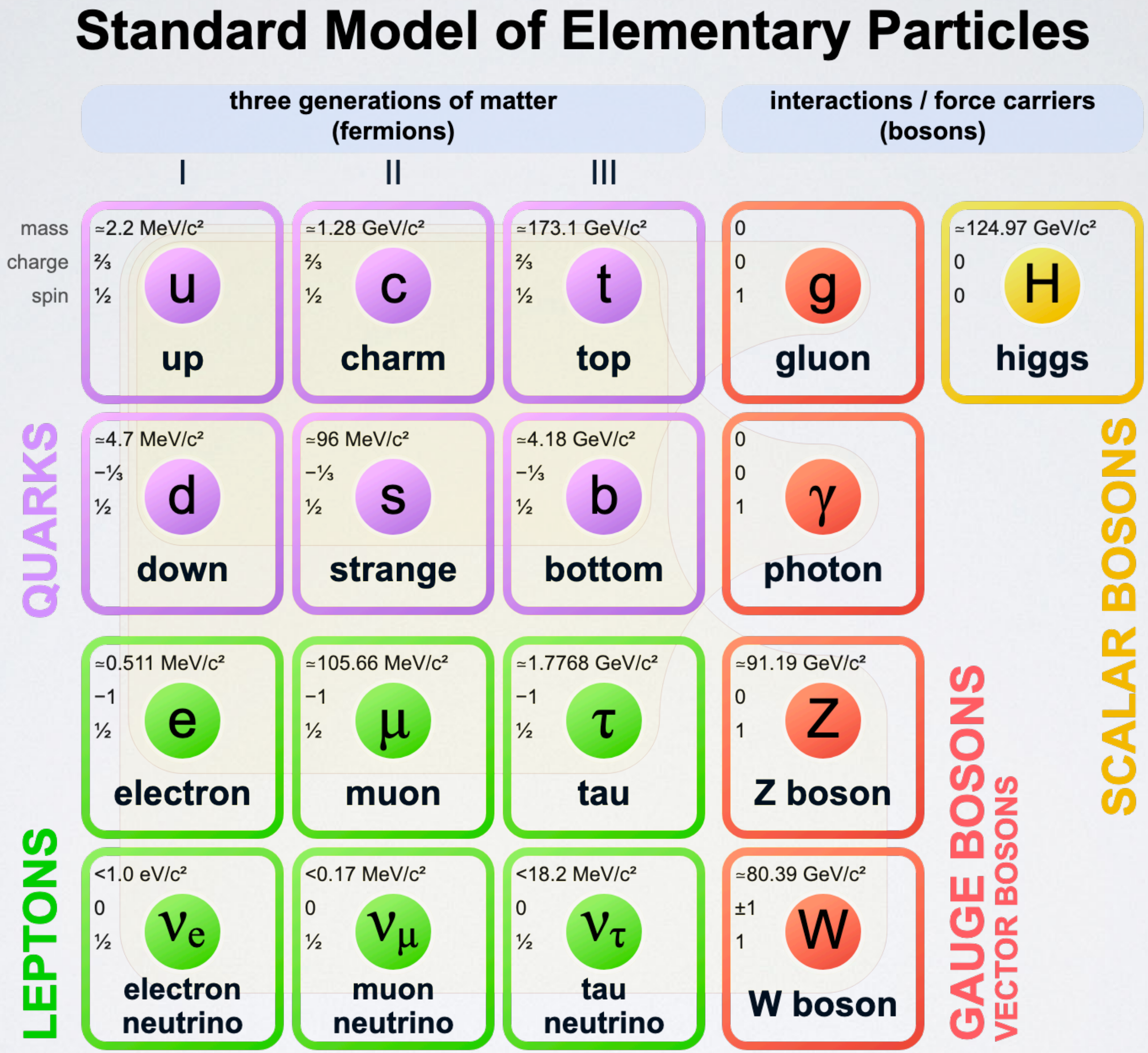


simulations

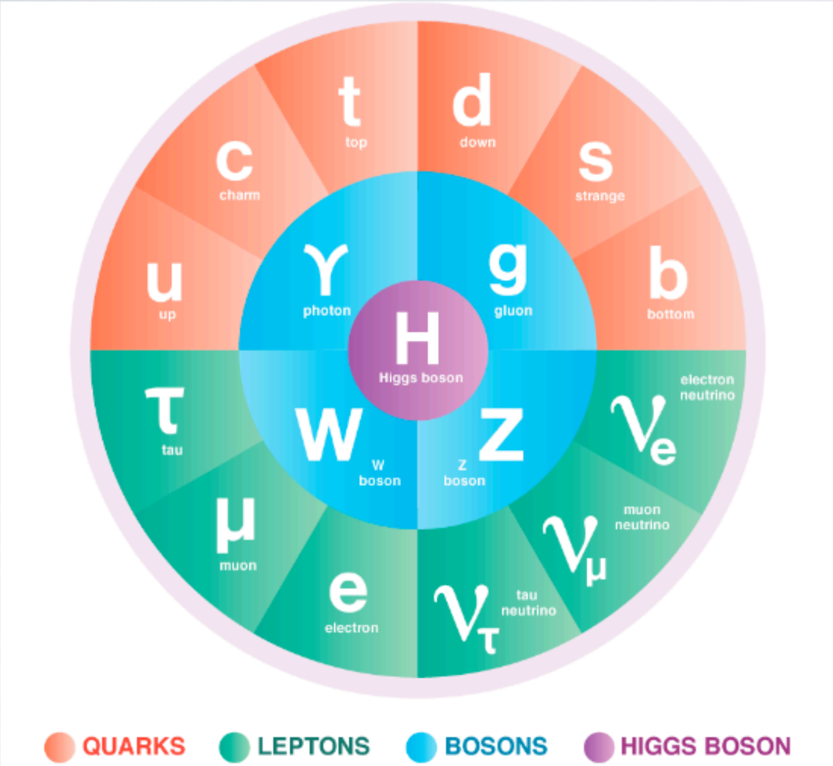


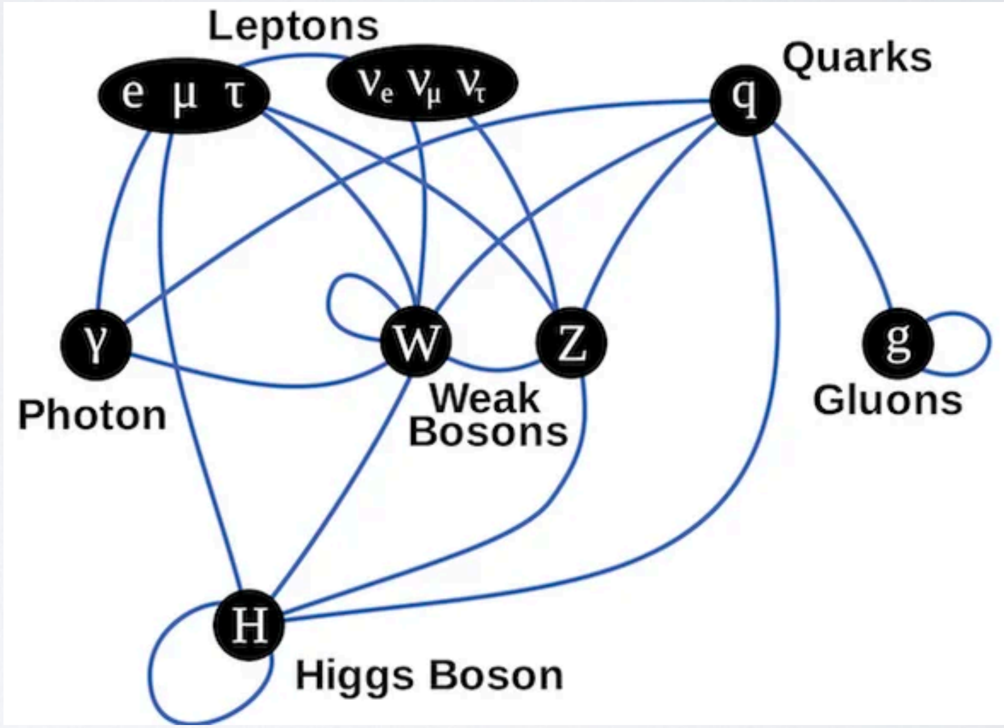
phenomenology

$SU(3)_C \times SU(2)_L \times U(1)_Y$



$SU(3)_C \times SU(2)_L \times U(1)_Y$

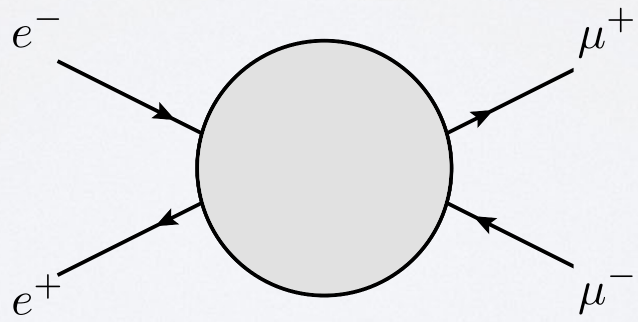




What do theorists see?
scattering cross sections

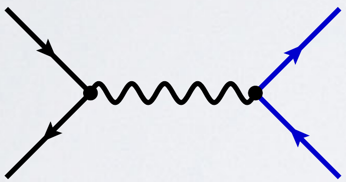
$$\sigma = \frac{1}{\text{flux}} \int \delta^{(4)}(Q - \sum_i p_i) \prod_i d^4 p \delta^{(+)}(p_i^2 - m_i^2) \langle |\mathcal{A}|^2 \rangle$$

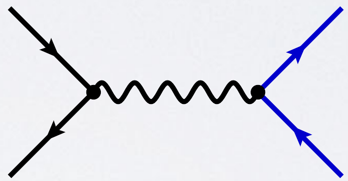
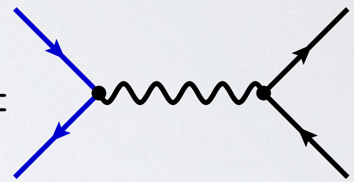
e.g.

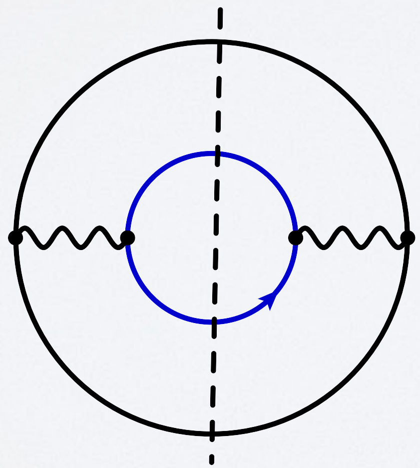


Scattering amplitudes:
Probabilities from the Lagrangian

$$\langle |\mathcal{A}|^2 \rangle = \sum_{\text{spins}} \mathcal{A}^\dagger \mathcal{A}$$

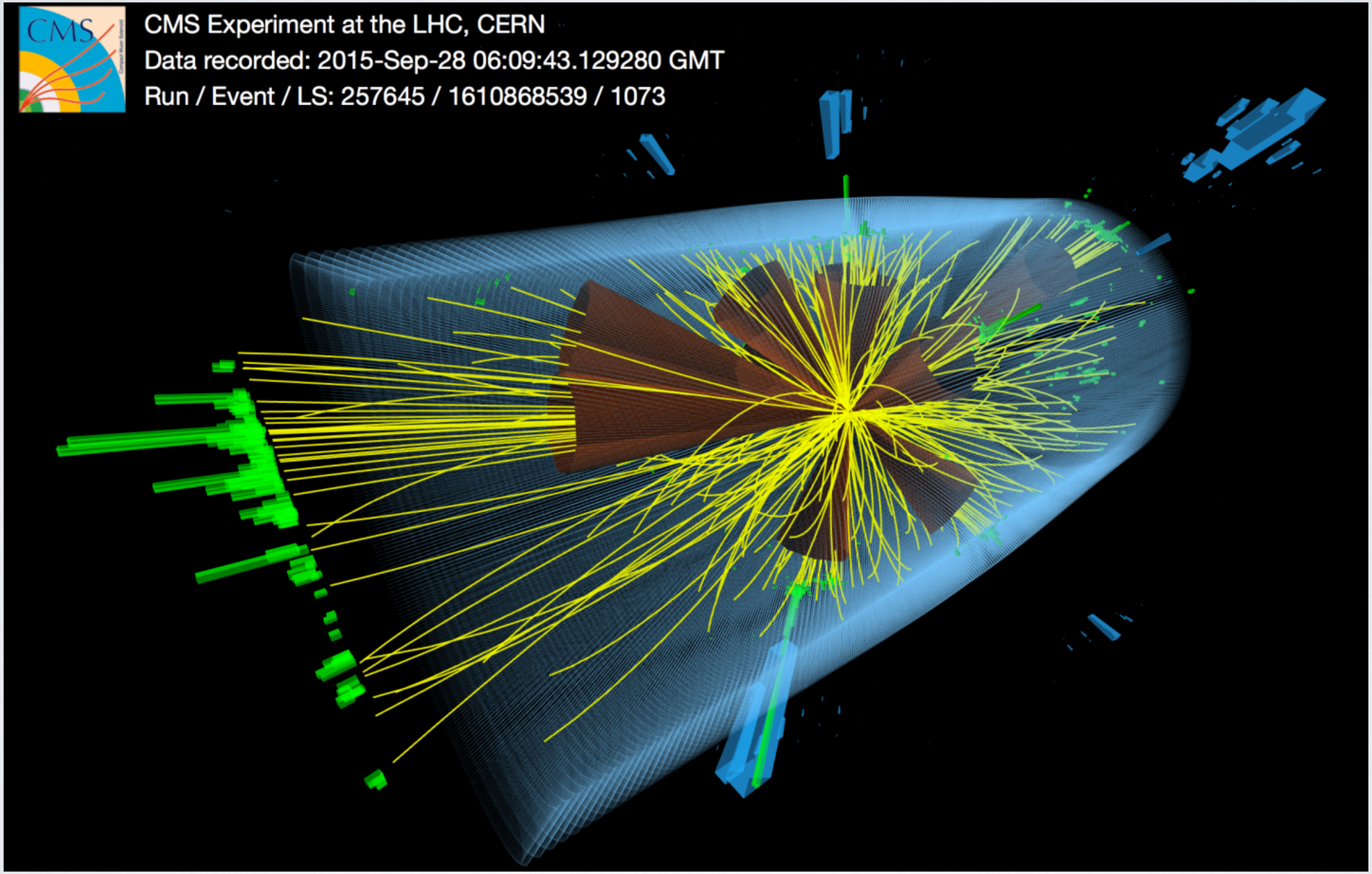
$\mathcal{A} =$ 

$\mathcal{A}^\dagger = \Big($  $\Big)^\dagger =$ 

$\langle |\mathcal{A}|^2 \rangle =$ 

diagrammatic
representation of the
squared amplitude

What do we see in the experiments?



What do we see in the experiments?

tracking detectors

calorimeters

exclusive particle identification

(see π, p, J/ψ **NOT** u,d,c,g)

unstable particle decay

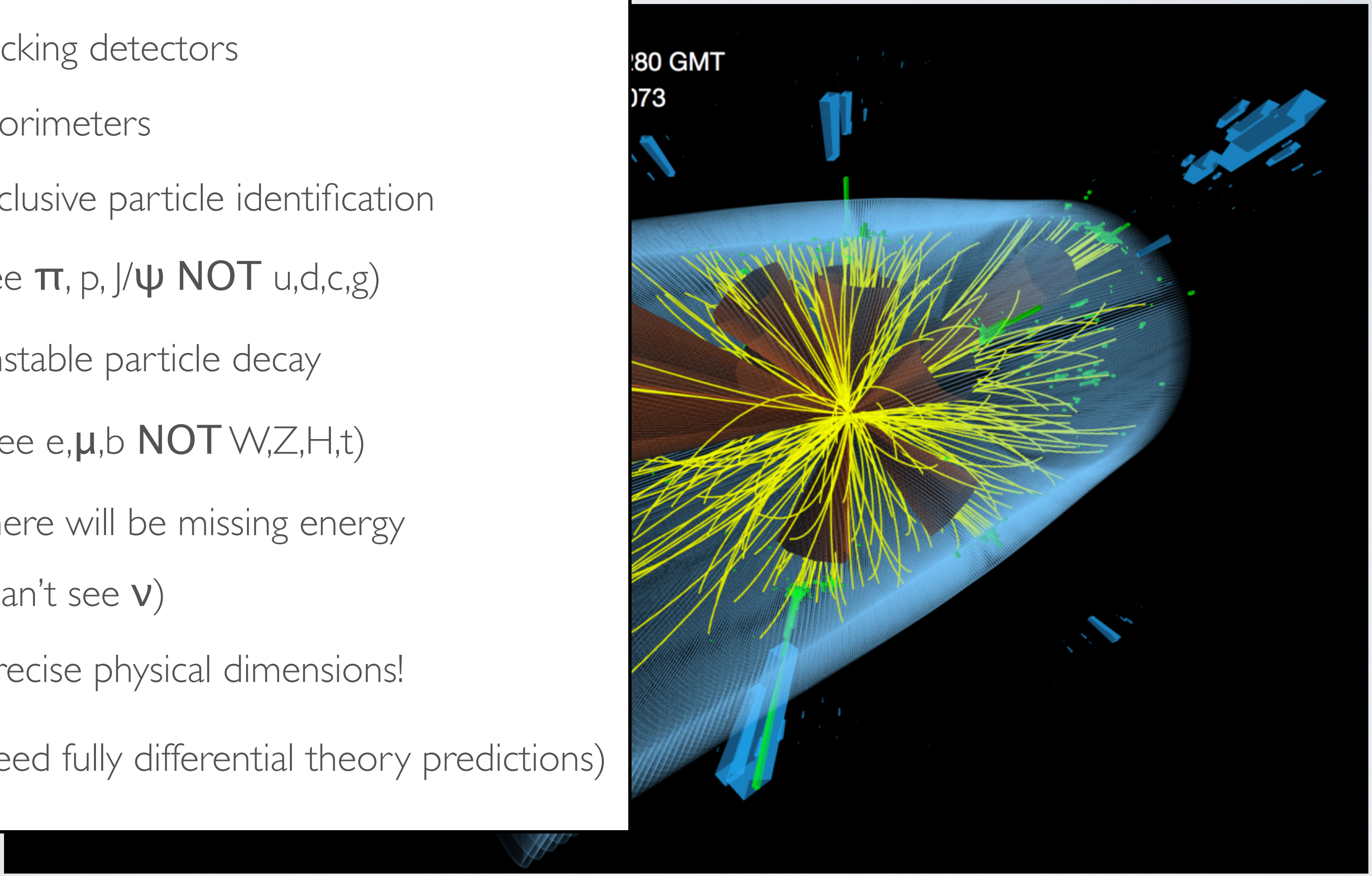
(see e,μ,b **NOT** W,Z,H,t)

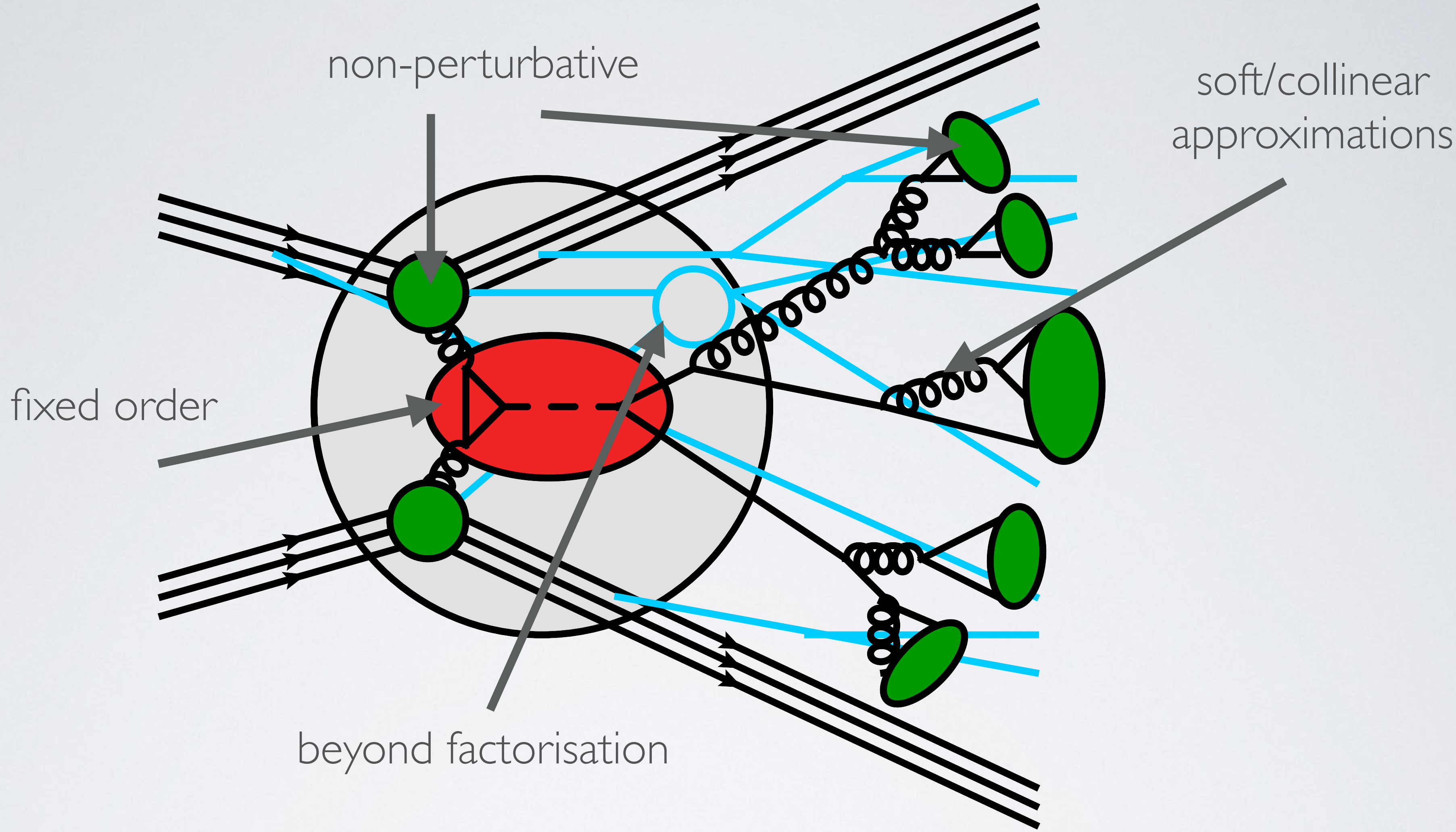
there will be missing energy

(can't see ν)

precise physical dimensions!

(need fully differential theory predictions)





numerical **event simulations** using Monte Carlo integration

'Monte Carlo Event Generators'

More on scattering at high energies
Factorisation at hadron colliders (e.g. LHC)

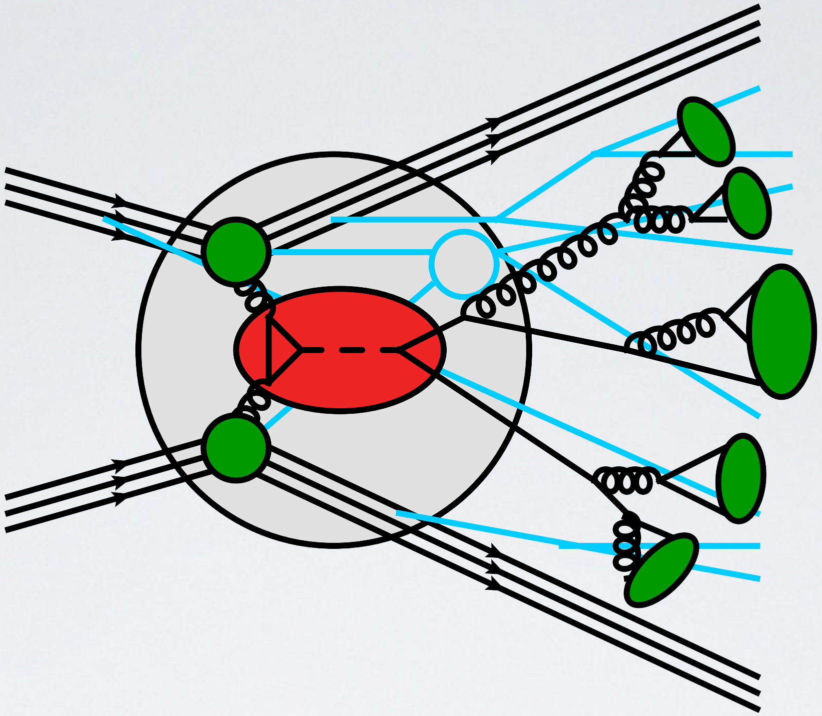
$$\widehat{\sigma} = \frac{1}{\text{flux}} \int \delta^{(4)}(Q - \sum_i p_i) \prod_i d^4 p \delta^{(+)}(p_i^2 - m_i^2) \langle |\mathcal{A}|^2 \rangle$$

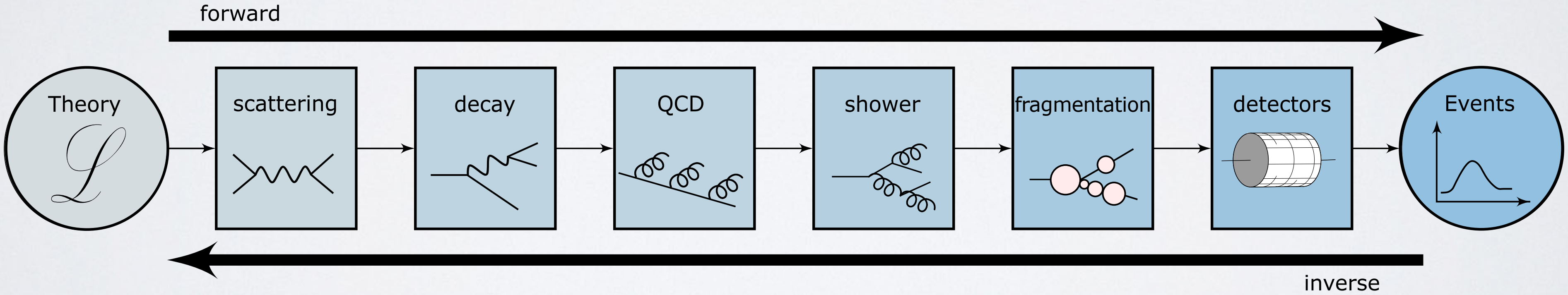
$$\sigma \sim \int dx_1 dx_2 f_i(x_1, Q^2) f_j(x_2, Q^2) \hat{\sigma}(ij \to \{p\}) \Theta + \mathcal{O}(Q/\Lambda)$$

PDF
parton distribution function

observable
(very broad definition)

beyond leading
order factorisation




forward
Theory
scattering
decay
QCD
shower
fragmentation
detectors
Events
inverse

what precision do we need?

determination of SM parameters

a lot!

general searches for BSM resonances

less…

Many experimental measurements approach 1% precision

e.g. Drell-Yan with ATLAS (2019)

	Data
$\sigma\ (W^+ \to \mu^+\nu)$ [pb]	3110 ± 0.5 (stat.) ± 28 (syst.) ± 59 (lumi.)
$\sigma\ (W^- \to \mu^-\bar{\nu})$ [pb]	2137 ± 0.4 (stat.) ± 21 (syst.) ± 41 (lumi.)
Sum [pb]	5247 ± 0.6 (stat.) ± 49 (syst.) ± 100 (lumi.)
Ratio	1.4558 ± 0.0004 (stat.) ± 0.0040 (syst.)

QCD is the largest coupling and so dominates the pertubative expansion

$$\alpha_s(M_z^2) = 0.117 \pm 0.0009$$

PDG world average (2021)

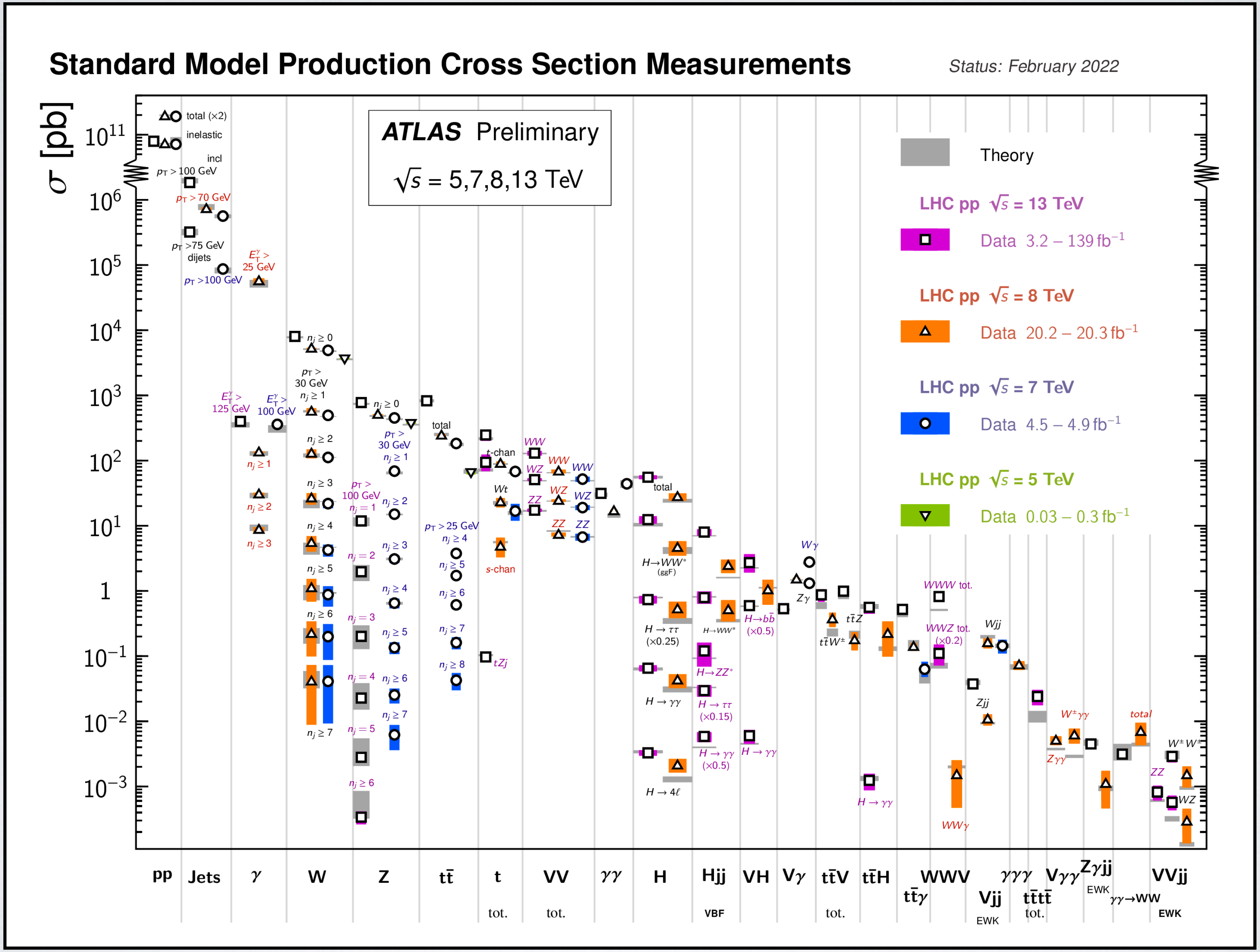
c.f. $\alpha^{(-1)}(0) = 137.035999150(33)$

$$d\sigma = \boxed{d\sigma^{\mathrm{LO}}} + \underbrace{\boxed{\alpha_s d\sigma^{\mathrm{NLO}}}}_{\sim 10\text{-}30\,\%} + \underbrace{\boxed{\alpha_s^2 d\sigma^{\mathrm{NNLO}}}}_{\sim 1\text{-}10\,\%}$$

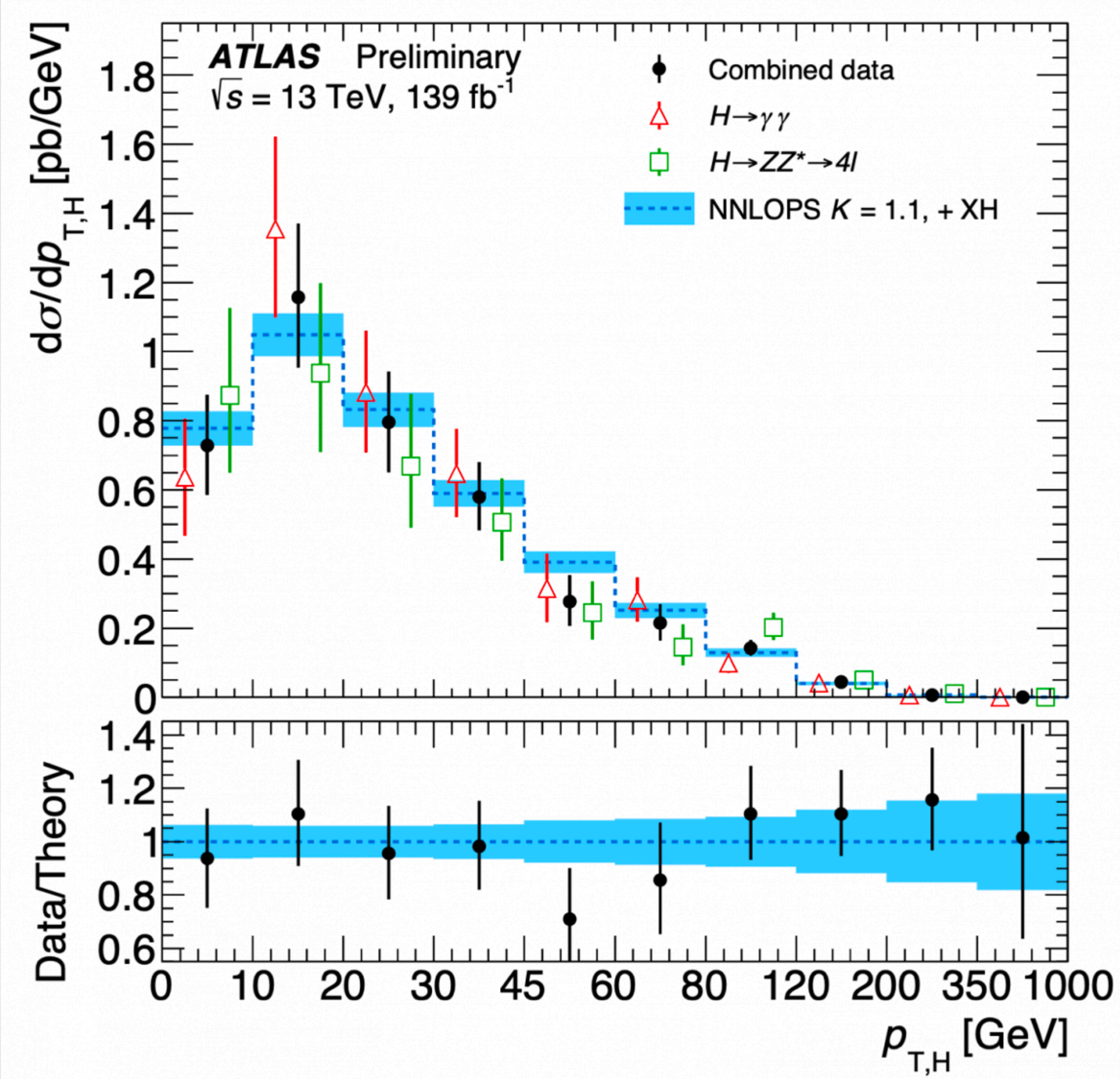
remember:
RG improved PT
is asymptotic

differentially many effects play a role: EW corrections, mass effects etc.

Current Standard Model Tests

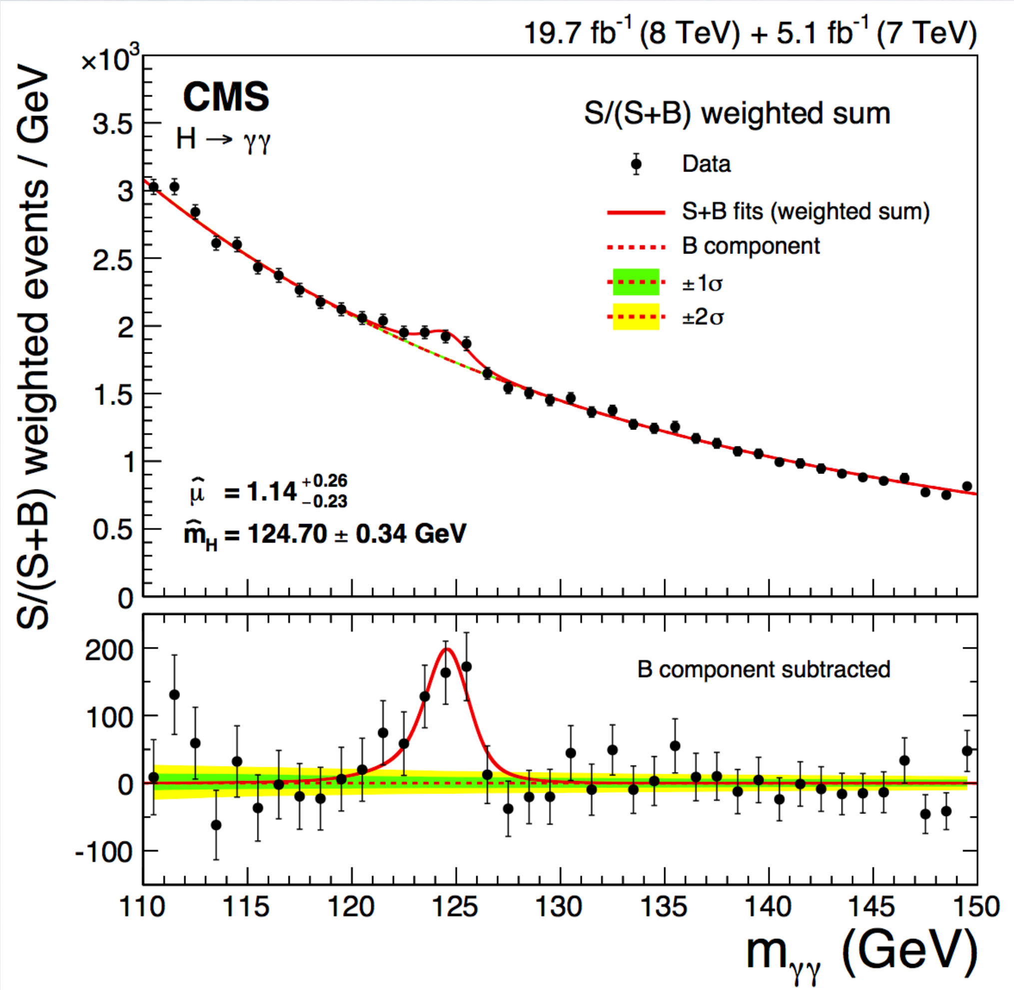


Differential tests offer more information



area under the curve is
the cross section
(Sezione d'urto)

Finding new particles



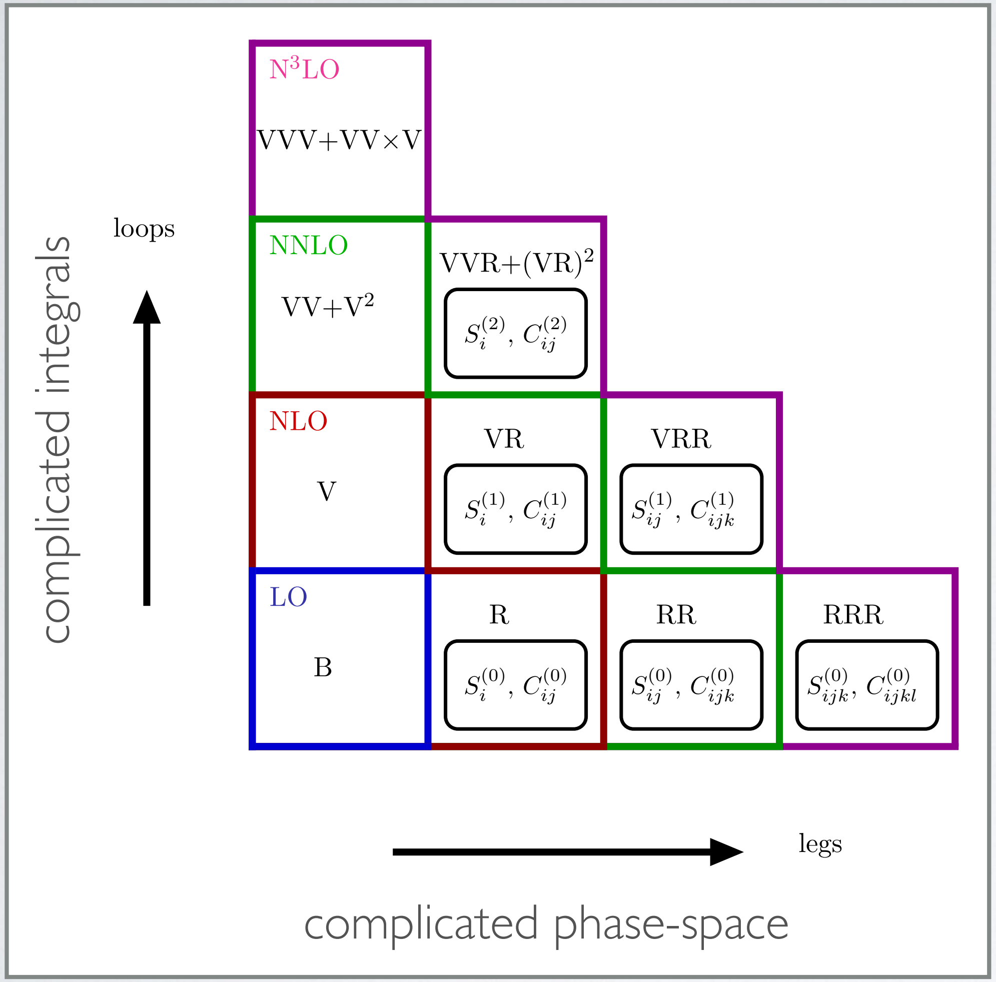
bump hunting is relatively easy

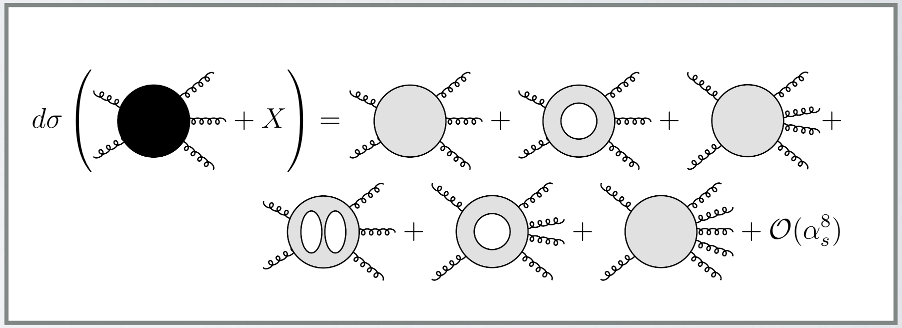
$$\mathcal{A} \sim \frac{1}{p^2 - m_h^2 - i m_h \Gamma_h}$$

leads to the Breit-Wigner distribution

$\Gamma_h =$ decay width

Precision theory for experiments:
quantum corrections from fixed order calculations



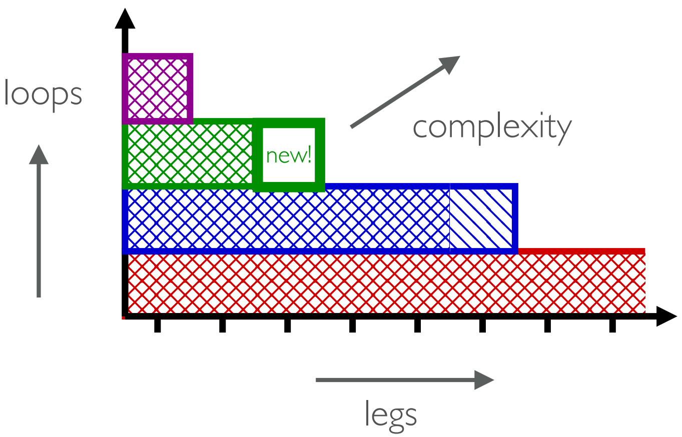


keeping theory in line with experiments takes **years** of dedicated effort

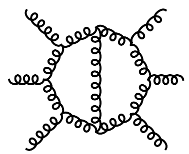
Growing Complexity

loops	1	2	3	4	5
diagrams	5	30	450	50,000	1.5×10^6
year	1973	1974	1980/1993	1997/2005	2016

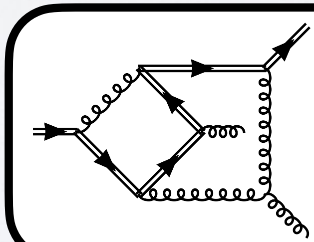
Diagrams contributiing to QCD β function up to 5 loops



more scales = more complicated



algebraic complexity
e.g. six-gluon scattering

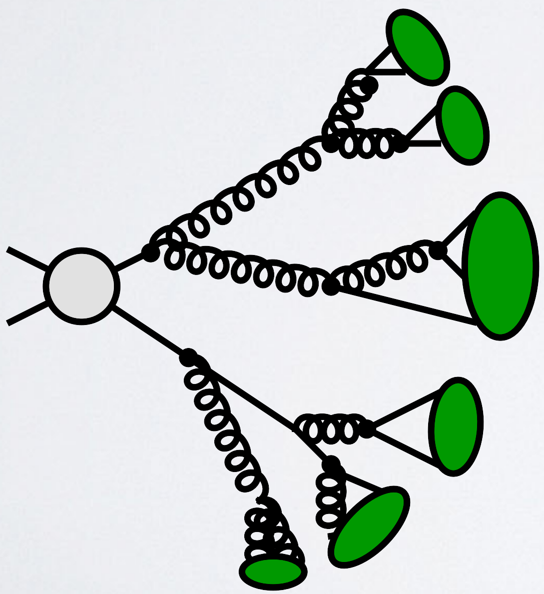


analytic complexity
e.g. pp⟶tt

More radiation: e.g. parton showers

The cancellation of IR divergences is connected to the universal singular behaviour of QCD

The probability of a parton emitting a soft or collinear gluon is independent of the hard scattering process



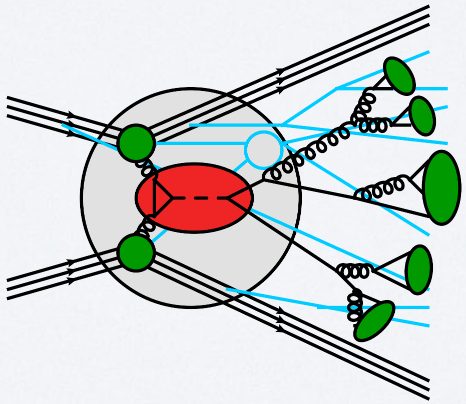
parton showers give a statistical evolution of the hard scattering process until the hadronisation scale is reached

the effect is an explicit re-summation of soft emissions **beyond fixed order** perturbation theory

General purpose event generators
(long term projects)

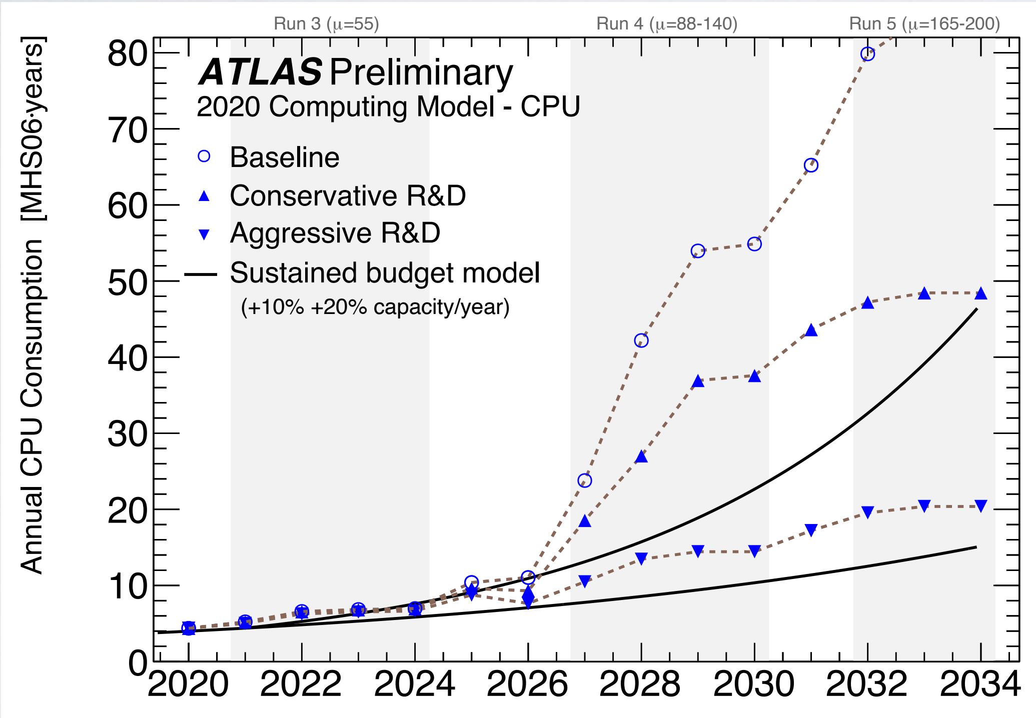
Pythia (1982-)
Herwig (1986-)
Sherpa (2002-)
Powheg (2004-)
MadGraph (1994-)

collect large numbers of different models (hadronisations, parton showers etc.) and interface with specialsed codes for PDFs, amplitudes etc.



what’s the computational cost?

LHC MC computing requirement projections



- Moore's Law no longer applies
- HL-LHC could have limitiations from MC statistics without new developments

from

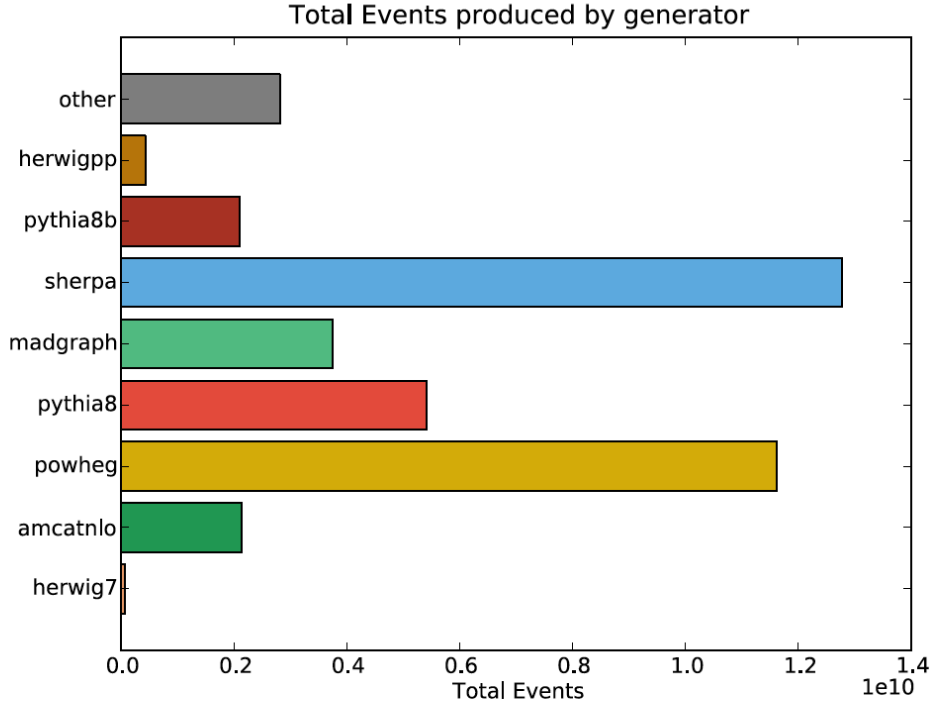
Christian Gütschow

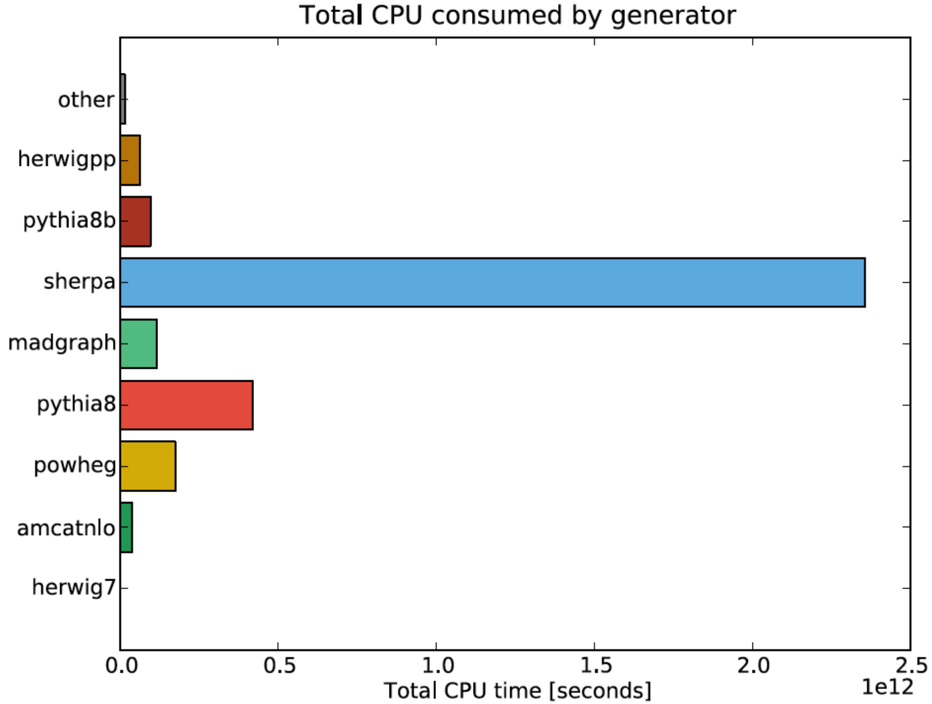
at

Taming the accuracy of event generators workshop CERN, June 2020

Breakdown by generator (bit outdated, but not too bad)

- left plot: does not account for alternative multi-leg setups
- right plot: most CPU spent on high-precision multi-leg calculations
  (e.g. for ATLAS: $V + 0, 1, 2j$@NLO$+3, 4j$@LO and $t\bar{t} + 0, 1j$@NLO$+2, 3, 4j$@LO)



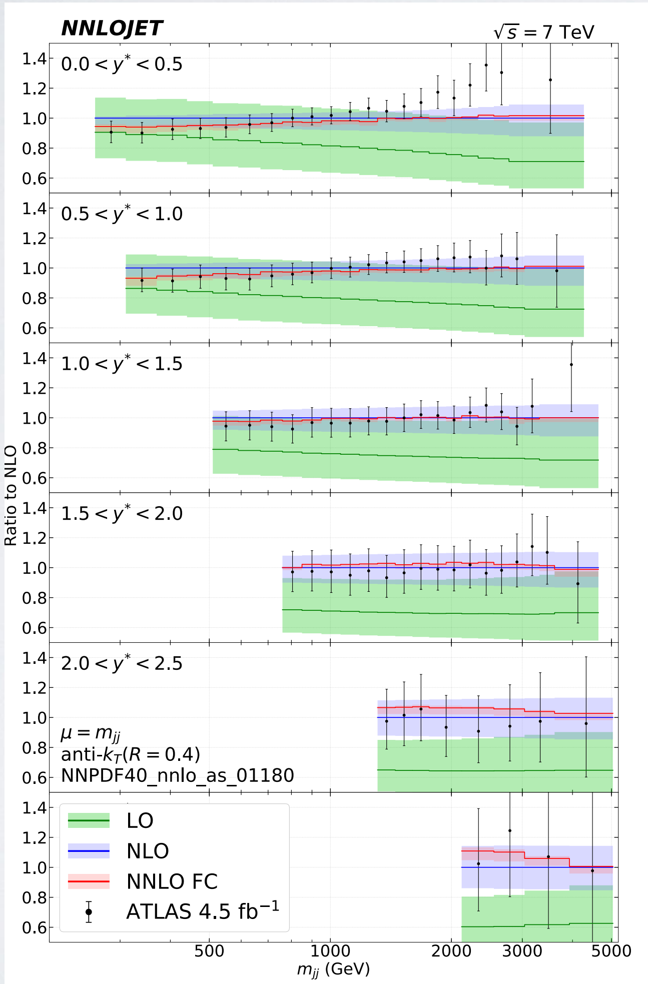


[courtesy of Josh McFayden]

- outlook: CPU spent on expensive setups expected to increase faster than for fast setups

and for NNLO precision?

many are still specialized (private) codes, NNLOjet, MATRIX, etc.



for some of the most expensive fixed order simulations

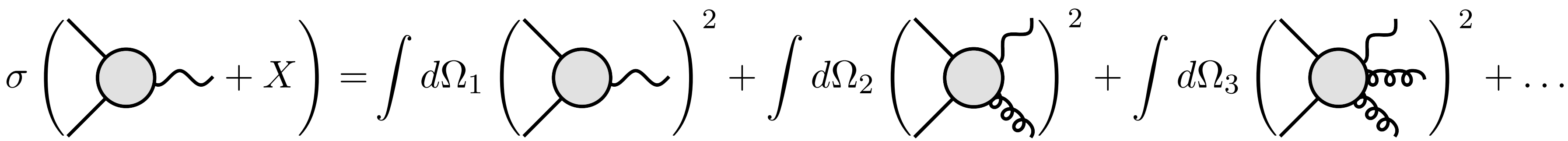
~ 200,000 CPU hours

example: full colour di-jets [Chen et al 2204.10173]

Making Precision Predictions

As mentioned before QFT contains divergences in both UV and IR

observables such as (differential) cross sections must be **inclusive** over the phase-space in order to capture the cancellations from **unresolved** radiation configurations



now consider the perturbative expansion of these amplitudes

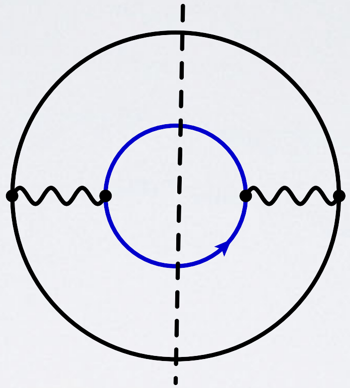
Making Precision Predictions

QCD corrections to $e^+e^- \to q\bar{q}$

IR divergences cancel between real and virtual corrections to physical observables.

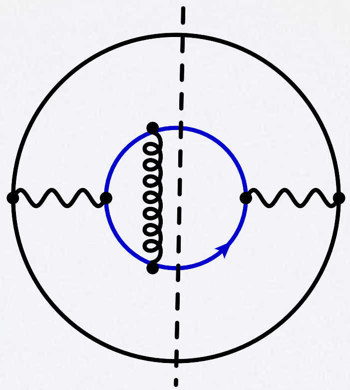
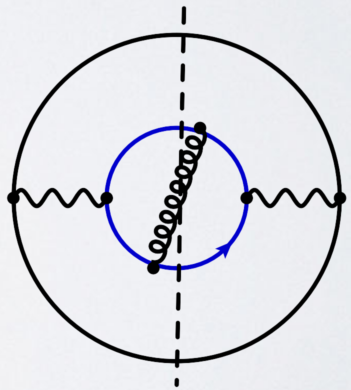
Kinoshita-Lee-Nauenberg (KLN) theorem

LO

$d\Phi_2$ 

'Born'

NLO

$d\Phi_2 d^4 l$  $+$ $d\Phi_3$ 

loop integration

'Virtual'

'Real'

Making Precision Predictions

The full calculation involves many steps and techniques…

$$\sigma^B(e^+e^- \to q\bar{q}) = \frac{4N_c\,\alpha^2\,q_q^2}{3\pi s}$$

$$\sigma^V(e^+e^- \to q\bar{q}) = \sigma^B C_F \frac{\alpha_s}{2\pi}\mathcal{H}(\epsilon)\left(-\frac{2}{\epsilon^2} - \frac{3}{\epsilon} - 7 - \delta^{CDR} + \pi^2\right)$$

$$\sigma^R(e^+e^- \to q\bar{q}) = \sigma^B C_F \frac{\alpha_s}{2\pi}\mathcal{H}(\epsilon)\left(\frac{2}{\epsilon^2} + \frac{3}{\epsilon} + 17/2 + \delta^{CDR} - \pi^2\right)$$

$$R = N_c \sum_{q=\{u,d,s,\dots\}} \Theta(\sqrt{s} - 2m_q)q_q^2\left(1 + C_F\frac{3\alpha_s}{4\pi}\right)$$

can use this to extract α_s from data $\Rightarrow \alpha_s(m_Z) = 0.1226 \pm 0.0038$

dimensional regularisation

$$d = 4 - 2\epsilon$$

scheme dependence

$$C_F = \frac{N_c^2 - 1}{2N_c}$$