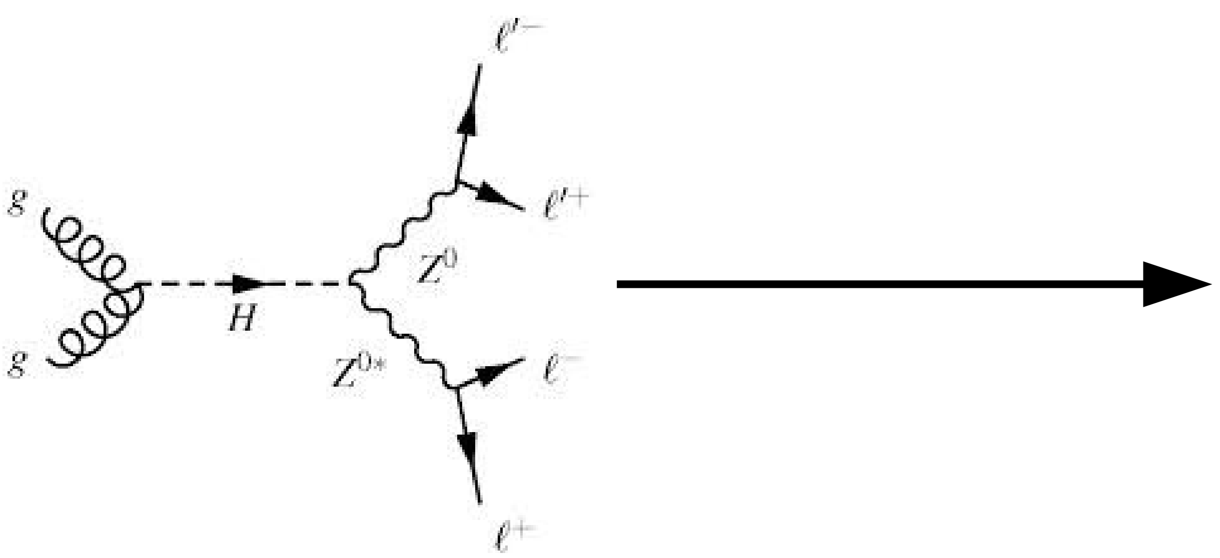
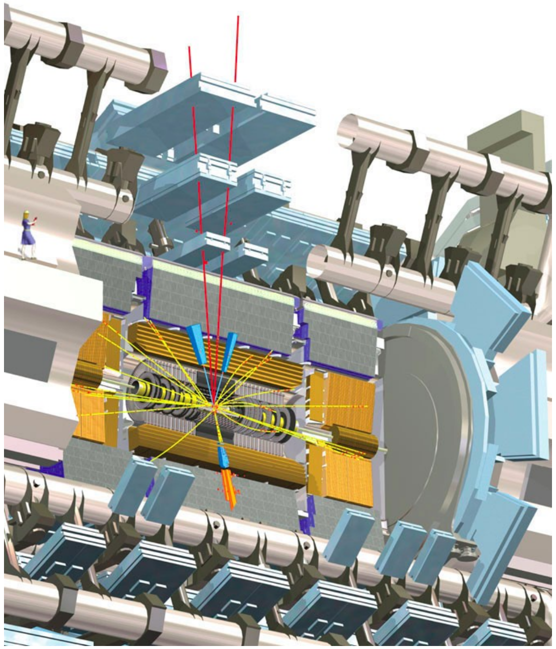
# Event Simulation II: Generators




Thorsten Kuhl

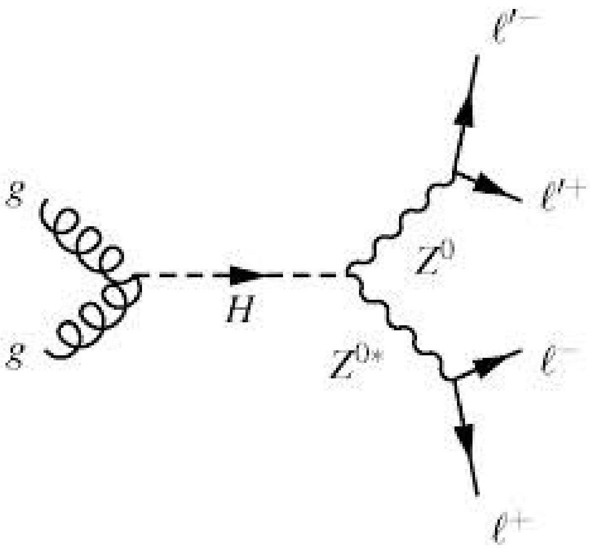

Graduate School, Mass-Spectrum-Symmetry
Autumn block course Berlin,
Sept. 29th-Oct. 2nd, 2014

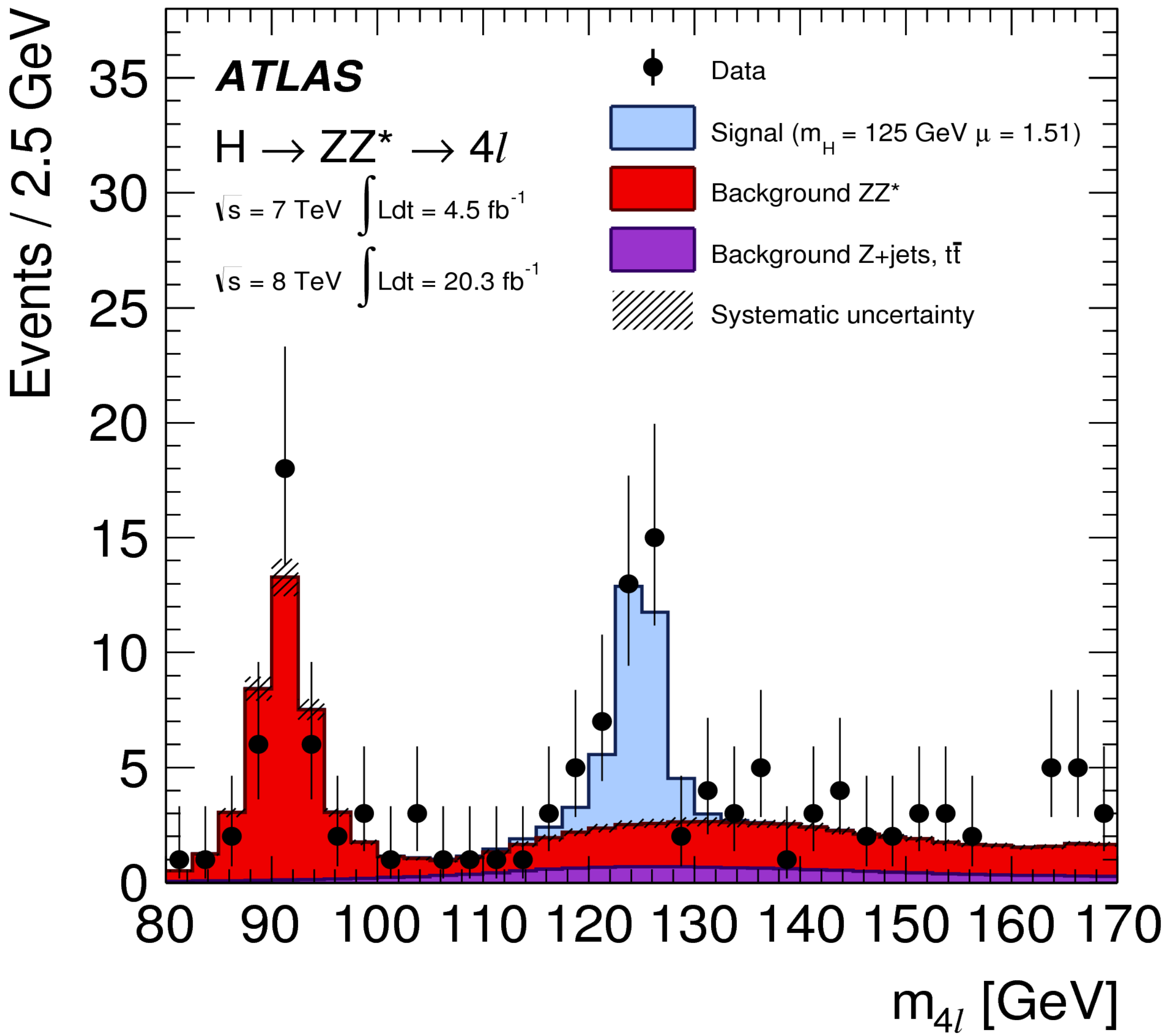
HELMHOLTZ GEMEINSCHAFT

DESY

# Why do we need simulation?

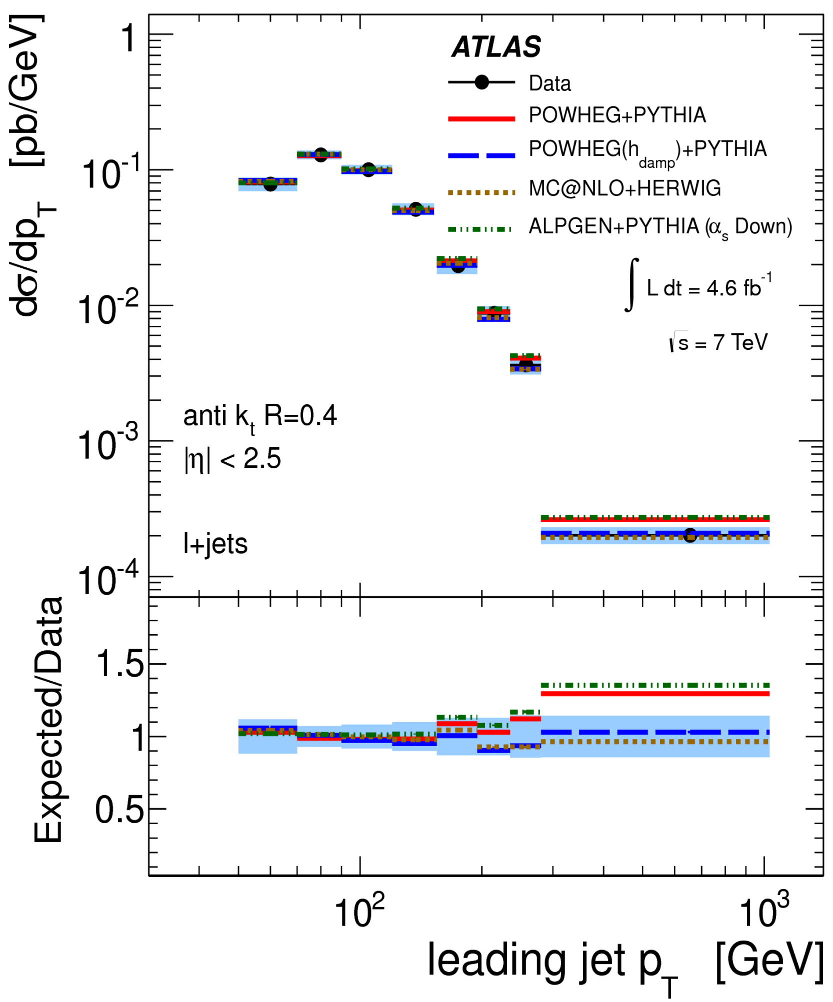
- Simulation of the known physics to discover new physics
  - Higgs → 4 leptons

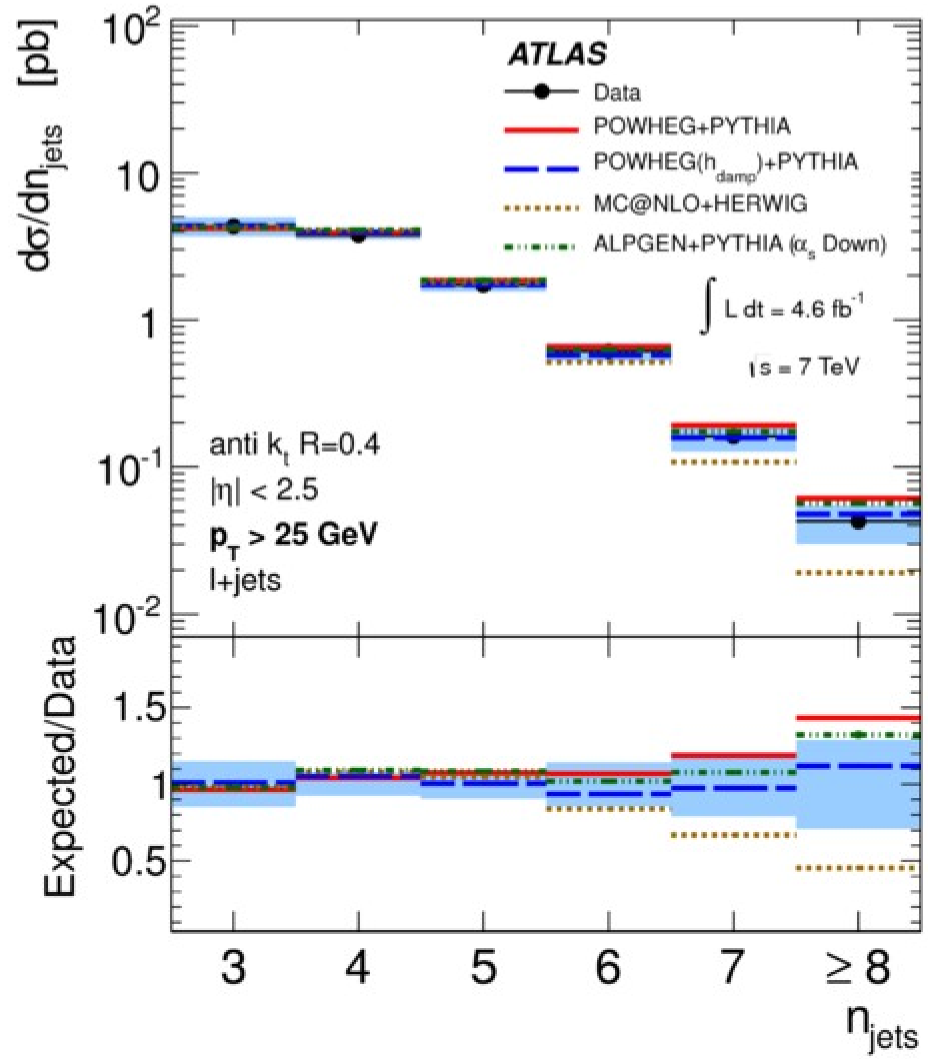


# Why event generation

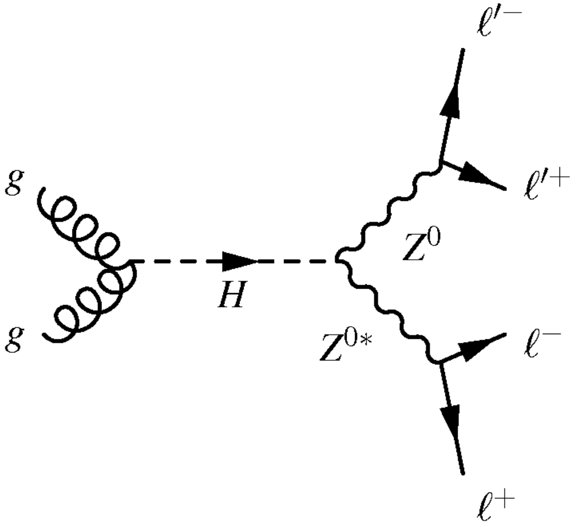
- Comparison of unfolded data (here ttbar plus jets)

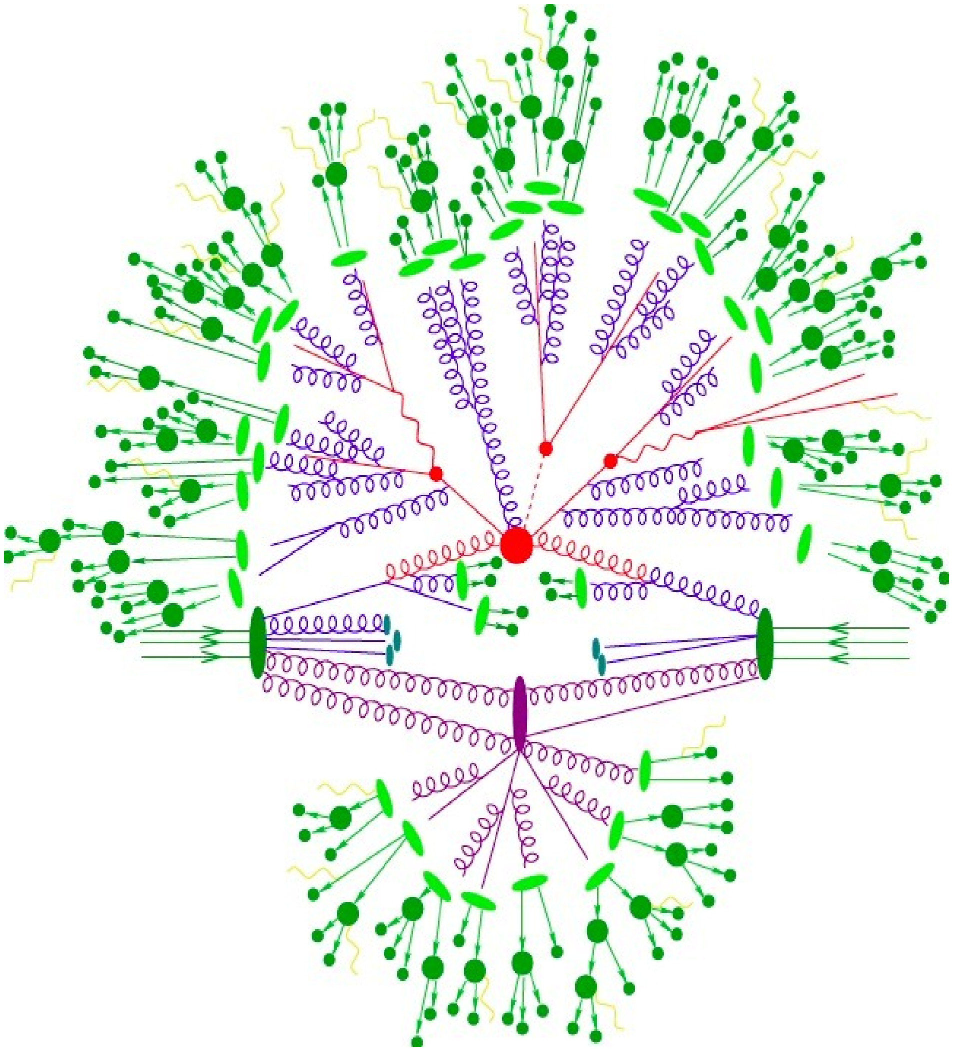


# Event Generation

- Theory → Experiment without Detector



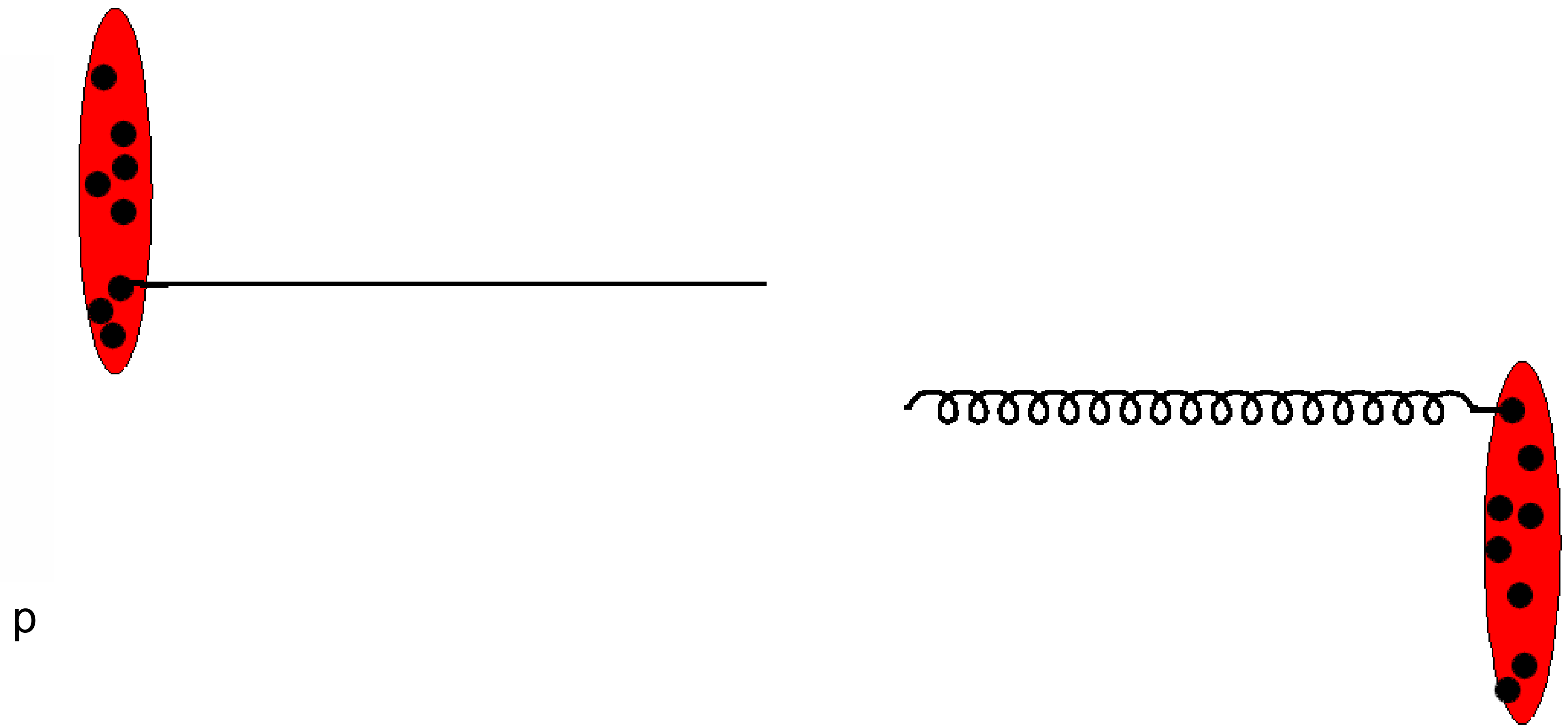
# Part II: Event Generation

- Overview:
  - Basic steps
  - Concept of factorization
- Incoming particles (PDFs)
- Matrix element
- Parton shower
- Matching/Merging
- Fragmentation hadronisation
- Decays/Life times

# A closer look to an particle interaction

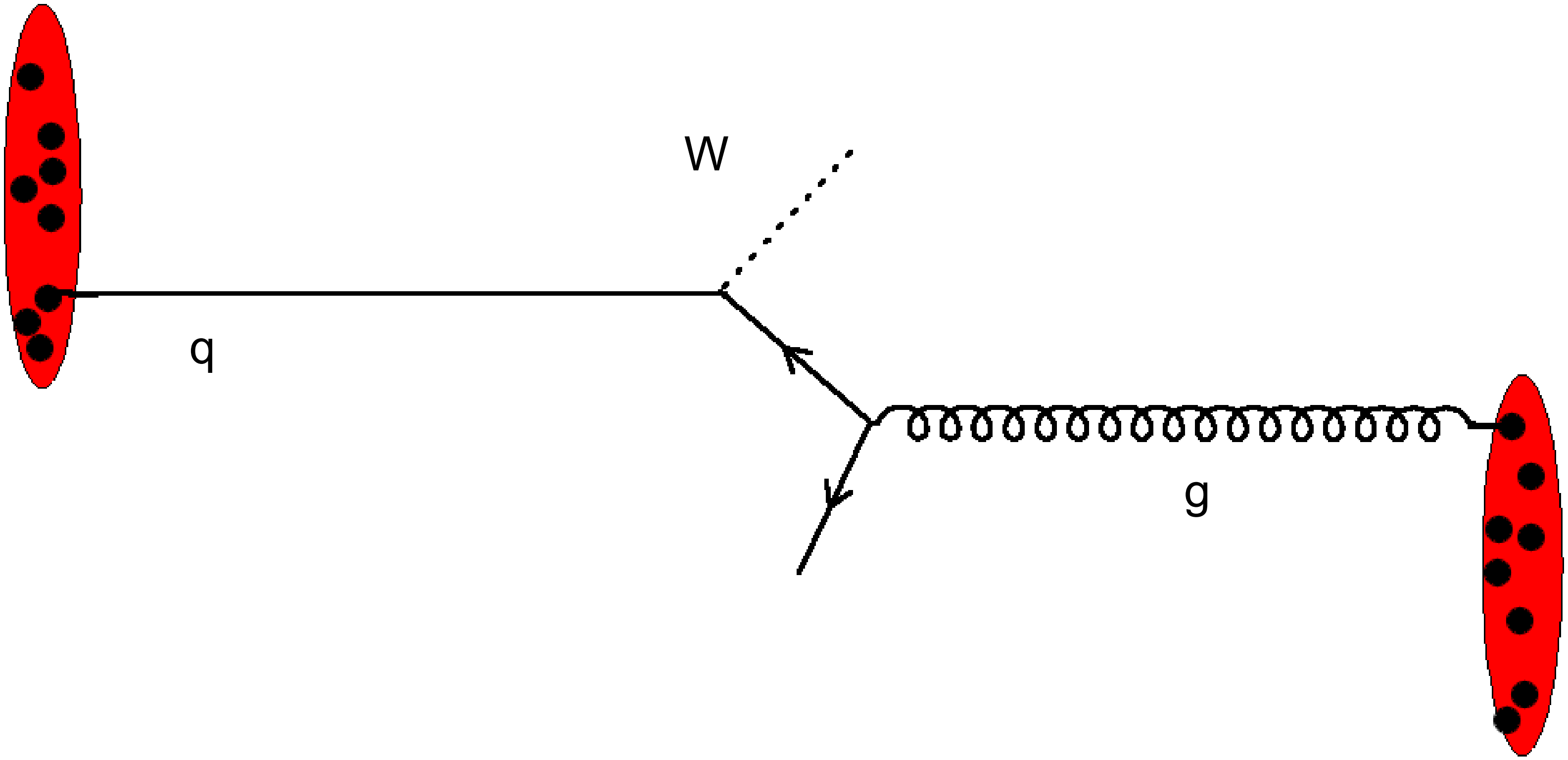
- Incoming Particles:
  - Only electrons, muons and neutrinos are fundamental particles
  - Protons, Neutrons: Sack of quarks and gluons

# Parton-Parton interaction



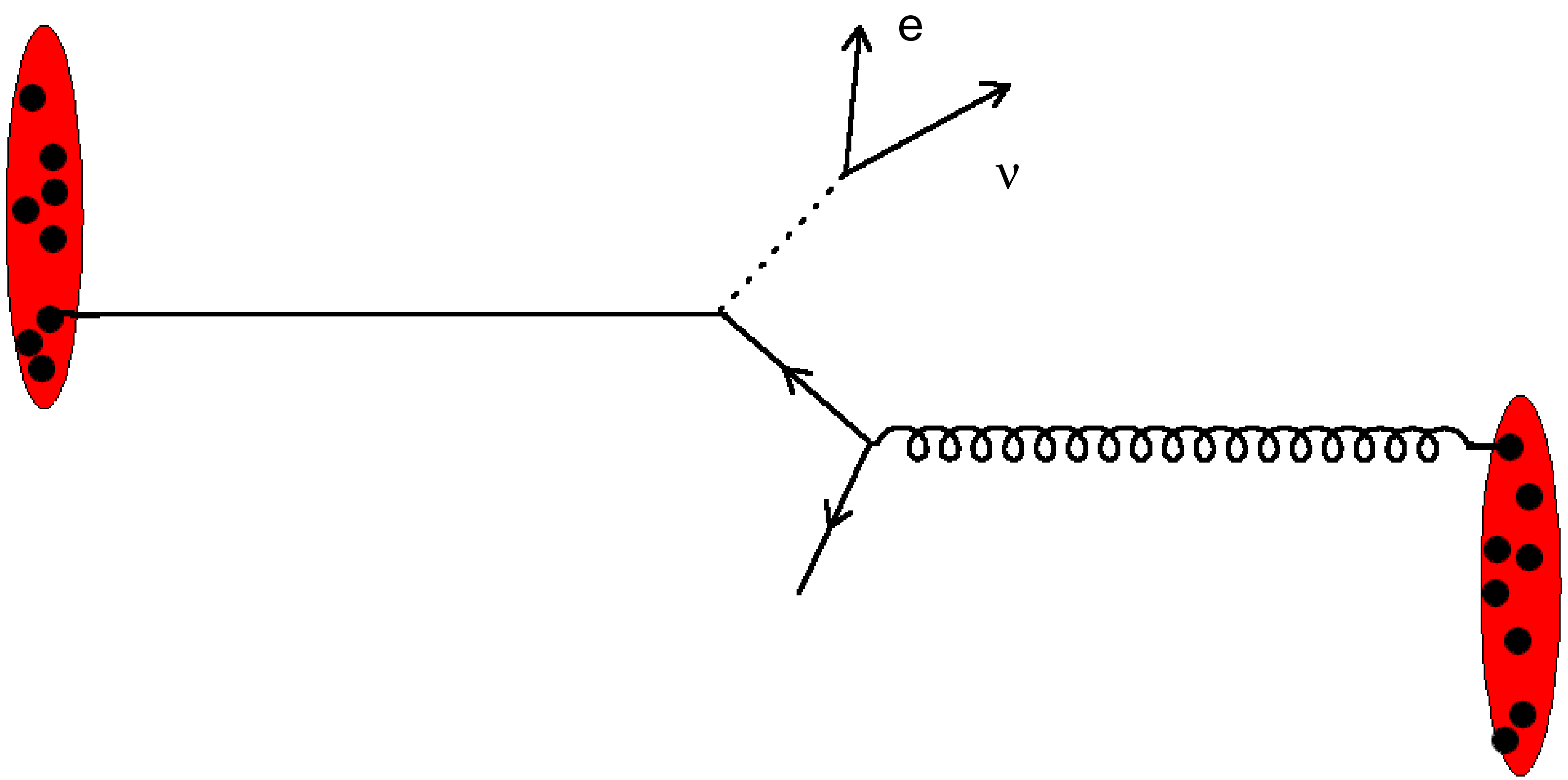

- Parton distribution function:
  - Substructure of the incoming proton
  - Momentum of the interacting quarks and gluons

# The hard scattering



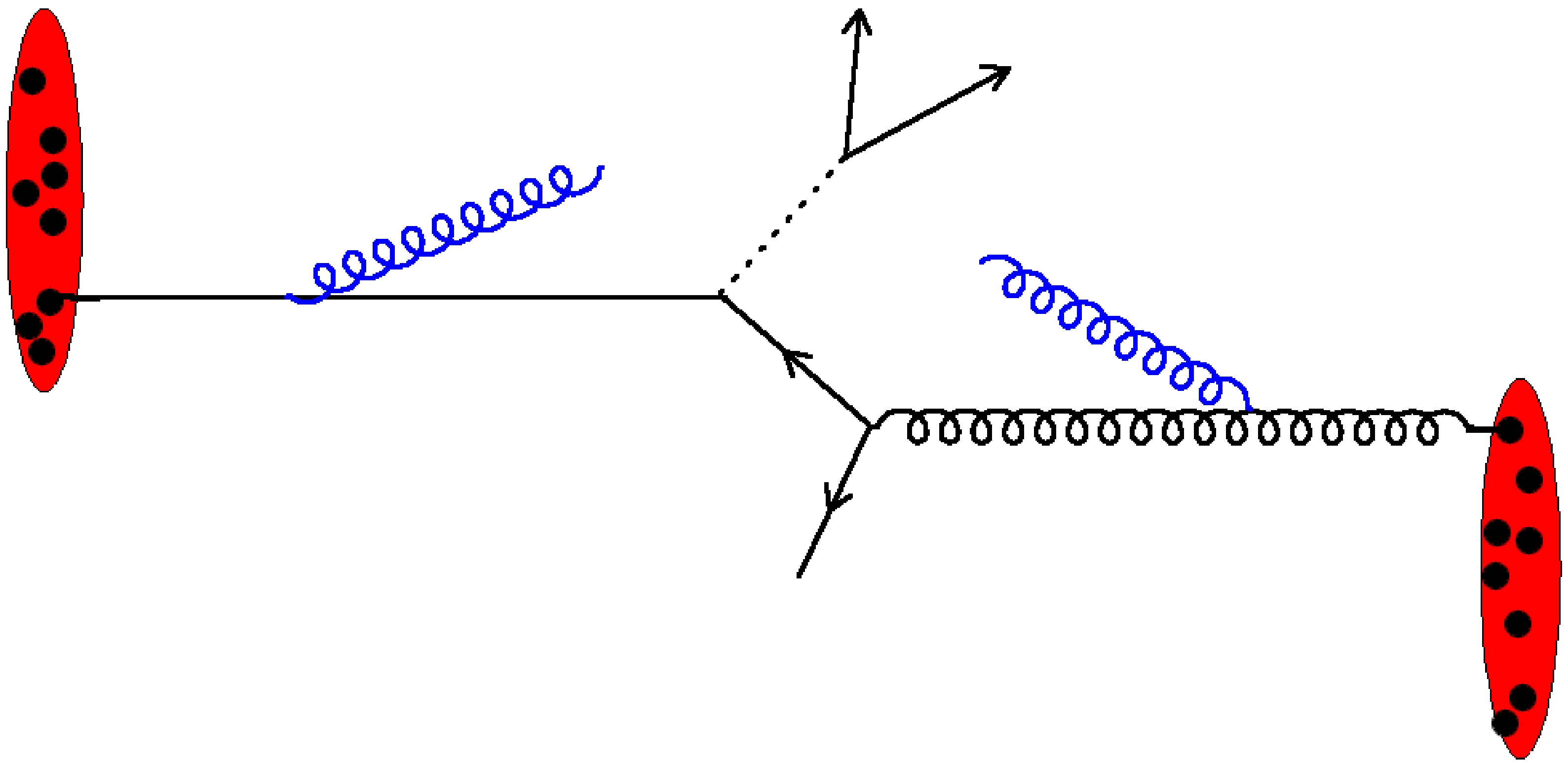

- Matrix element: Hard interaction

# EW-Boson-Decay




> Decay of short living resonances

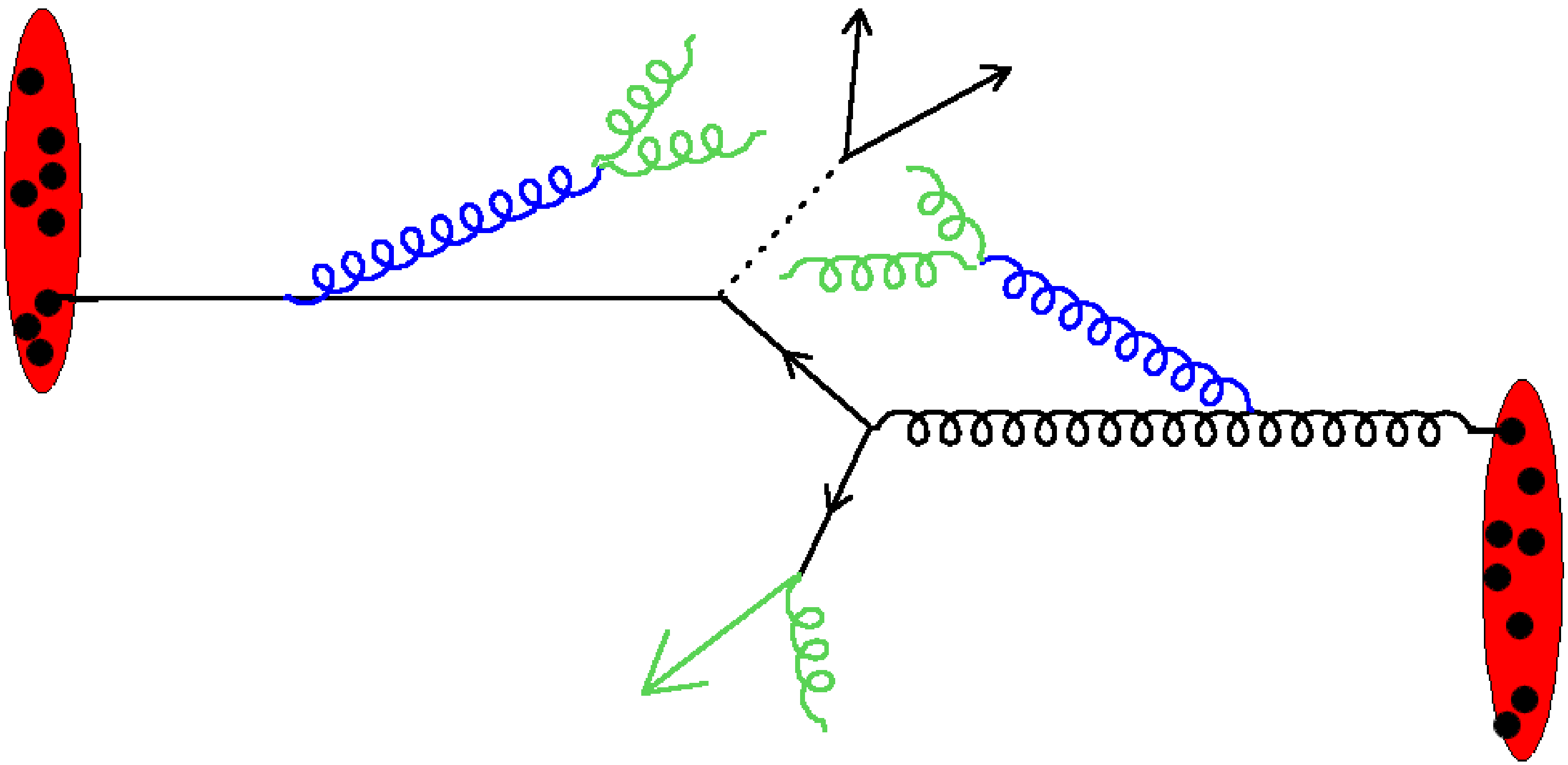
# Initial State Radiation



- Initial state radiation off the incoming partons

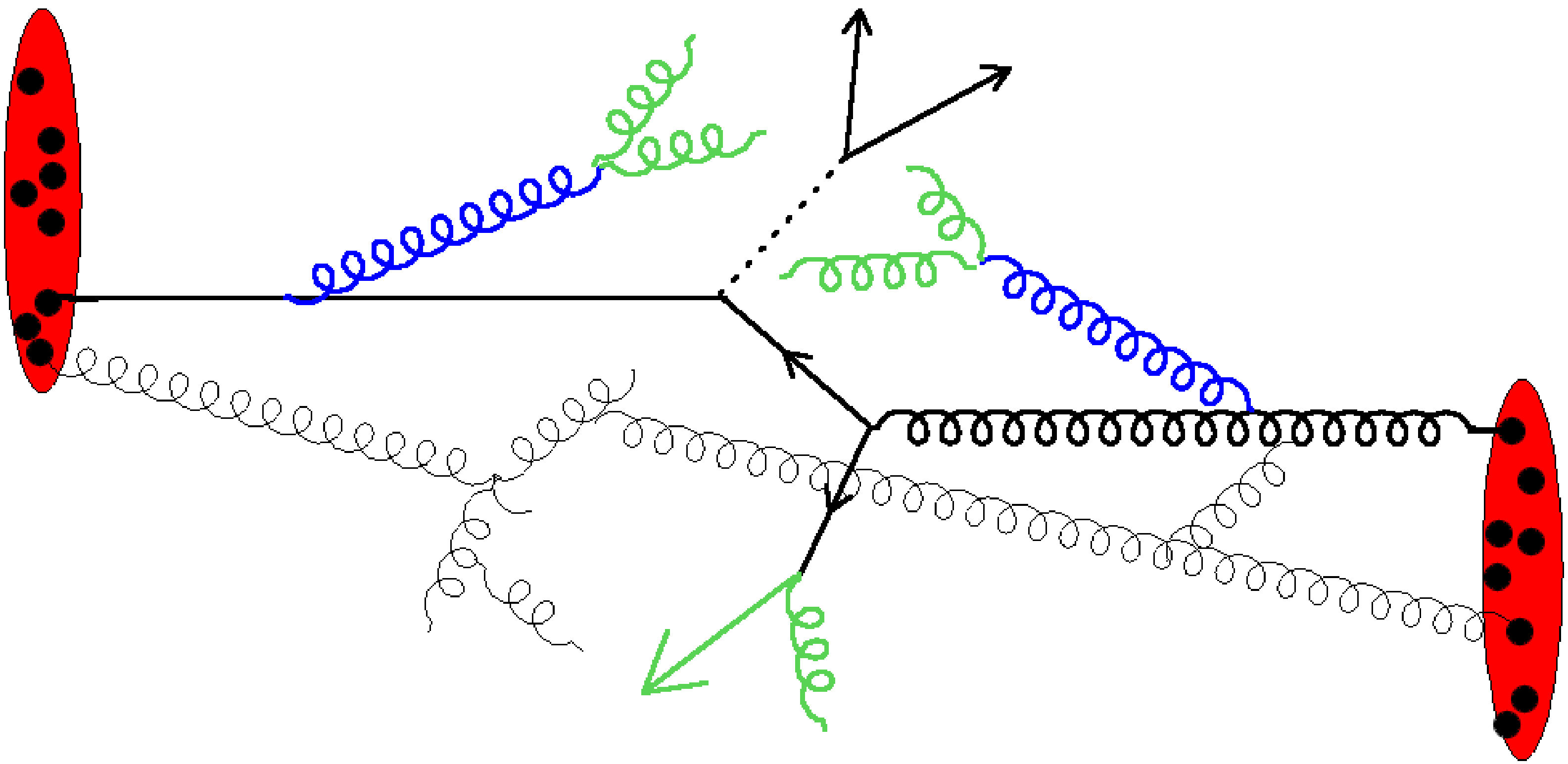
# Final state radiation: Parton shower





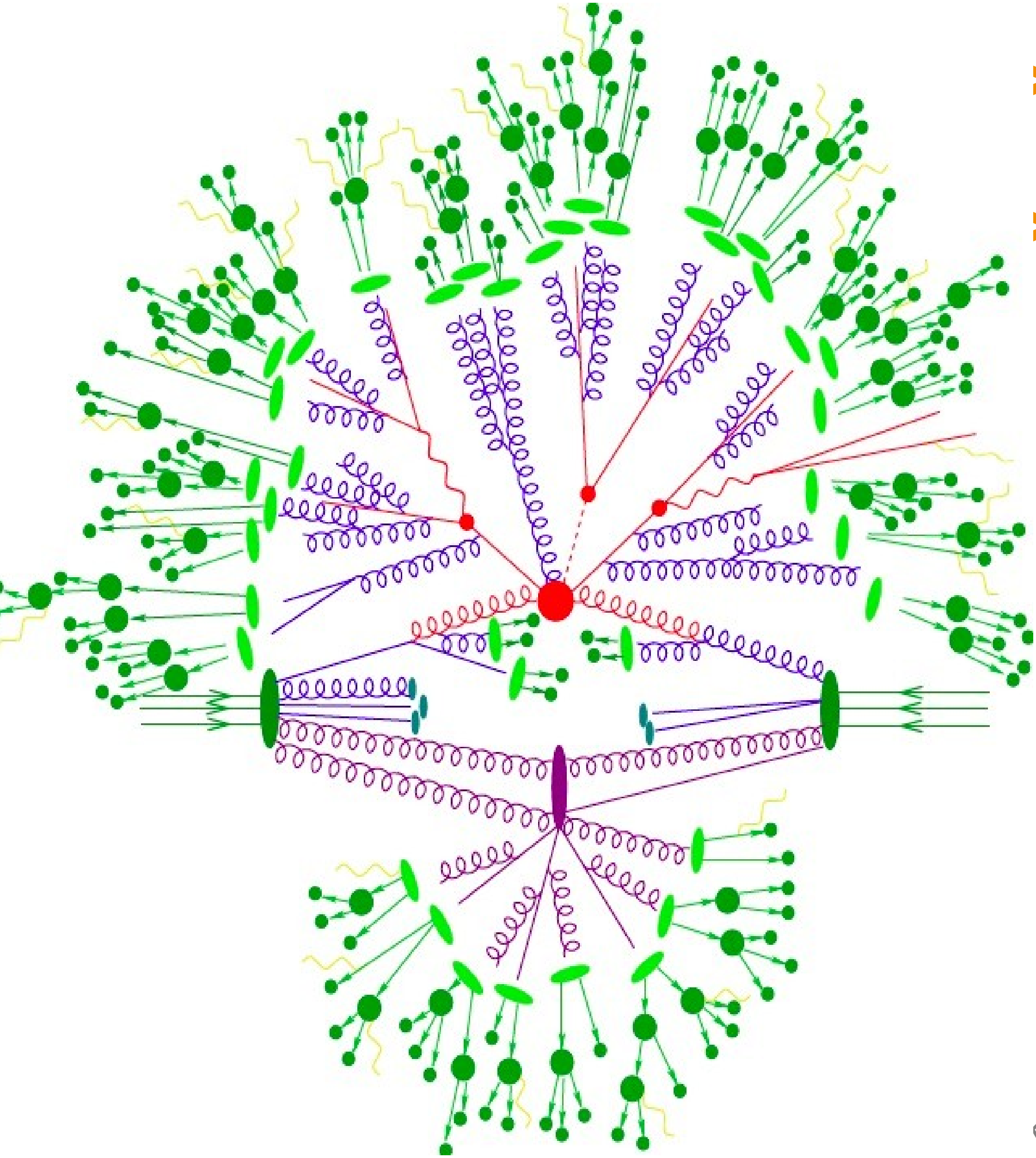
> Final state radiation: QCD radiation of “final state” partons

# Interaction of other constituents: Underlying event



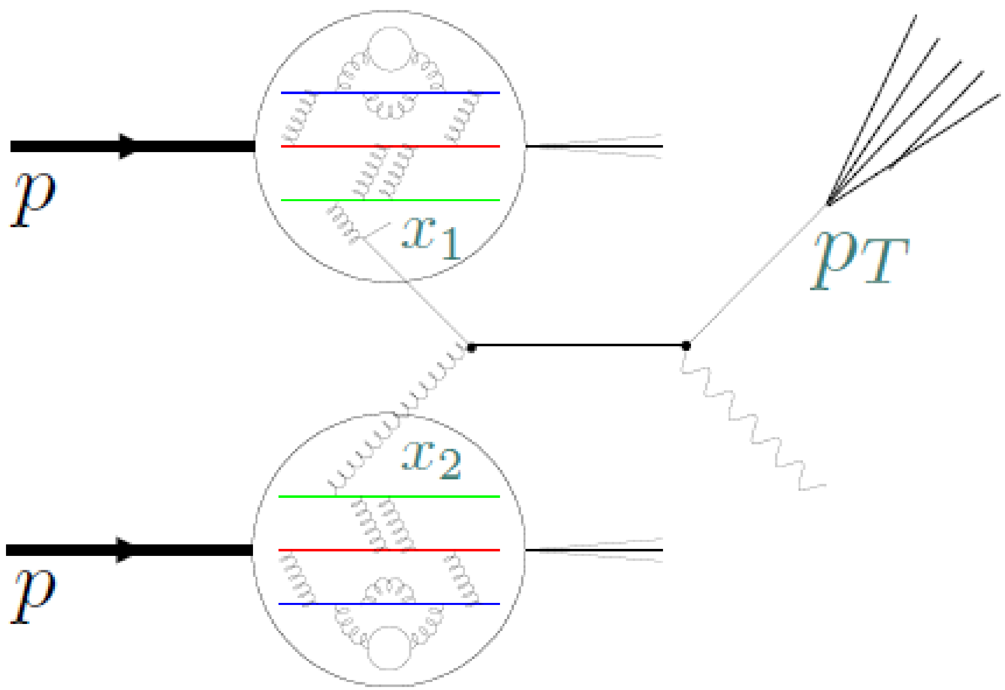
- Underlying event: Other soft interactions in the same proton/proton interaction

# A more realistic event: Sherpa W+3jet production



- More realistic event (display of an sherpa event)
- Hadronisation and decays:
  - Fragmentation: Conversion of partons into clusters or strings
  - Hadronization: cluster to Hadrons
  - Decay of strongly and electromagnetic decaying particles (everything which lives "short enough")

# Factorization theorem




- Generation of events heavily relies on factorization of the problem
- Different initial states can be added up (LO)
- Boundaries of the different steps (matching, merging) are current hot topic:
  - Different possibilities for producing the same final state

$$\sigma(AB \rightarrow X) = \Sigma_{a,b} \int dx_1 \int dx_2 \, f_{a/A}(x) \, f_{b/B}(x) \, \hat{\sigma}^{ab\rightarrow x} \, D_f^{x\rightarrow X}$$

PDF — Matrix element — Parton shower

# Overview summary

- Event generation done in many steps:
  - Incoming partons
  - Hard scattering/Matrix element
  - Initial and final state parton shower
  - Underlying events
  - Fragmentation
  - Hadronisation
  - Decay of short living particles
- Event Generation relies heavily on factorization: you can put one specialized module after the other
  - For simulation → you can put many modules one after the other, taken from different collaborations
- At the boundary and in case there are many possibilities to produce one result matching and merging algorithm are necessary → later

# First Module: Parton distribution function (PDF)



- Important input for the theory simulation (Generator)
- Distribution of quarks and gluons in the incoming proton
  - x momentum fraction of the interacting quarks and gluons
  - $Q^2$ : scale of the hard interaction (effective ECM energy)

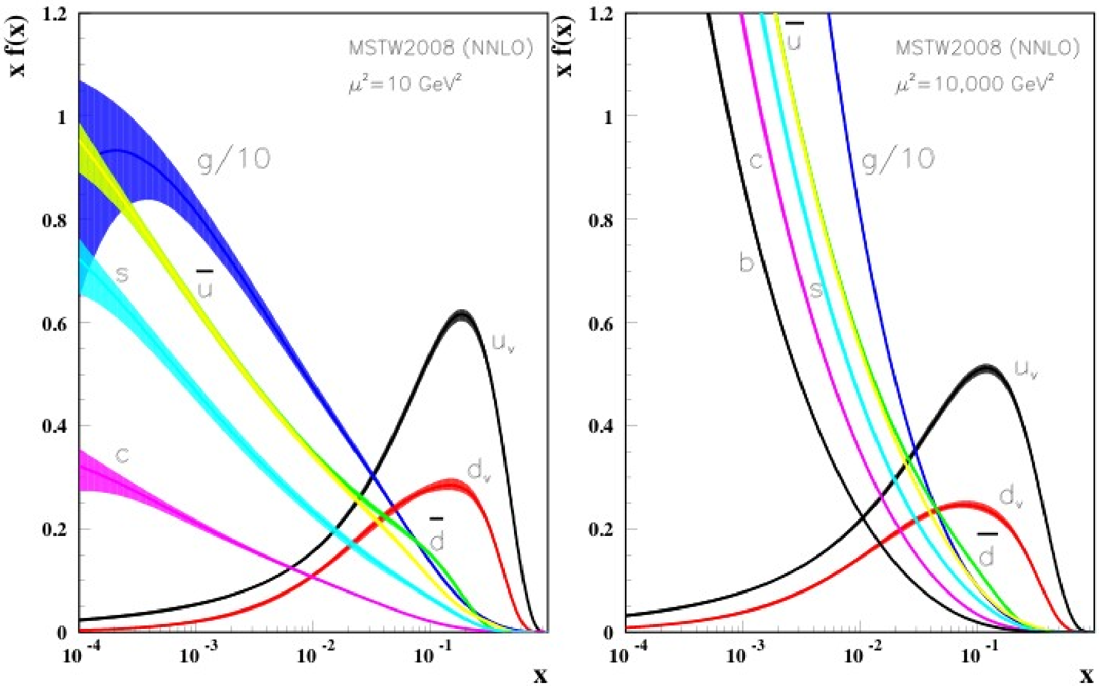



**Figure 16.4:** Distributions of $x$ times the unpolarized parton distributions $f(x)$ (where $f = u_v, d_v, \overline{u}, \overline{d}, s, c, b, g$) and their associated uncertainties using the NNLO MSTW2008 parameterization [13] at a scale $\mu^2 = 10\ \mathrm{GeV}^2$ and $\mu^2 = 10,000\ \mathrm{GeV}^2$.

# PDF- Fits: Inputs and evaluation

- PDF-Fits using a parametrisations in x and PDF sensitive data (multi-jets, W/Z/top productions, Hera ep scattering)

$$xu_v(x) = A_{uv} x^{Buv} (1-x)^{Cu\,v} (1+ D_{uv}\, x + E_{uv}\, x^2)$$

$$xubar(x) = A_u x^{Bu} (1-x)^{Cu} (1+ D_u\, x + E_u\, x^2)$$

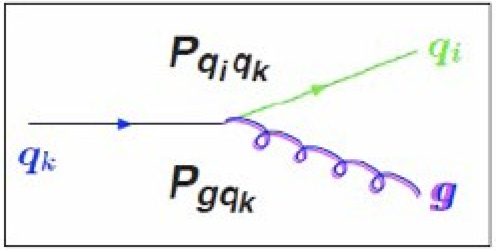
$$xg(x) = A_g x^{Bg} 1-x)^{Cg} (1+ D_g\, x + E_g\, x^2)$$

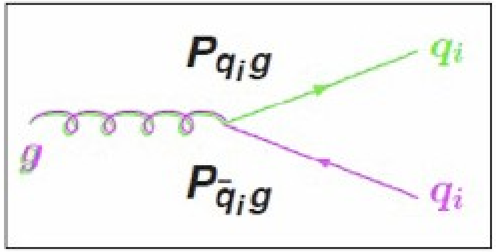
$$- A'_g\, x^{Bg'} (1-x)^{Cg'}$$

- DGLAP equation: evaluation from one scale $Q^2$ to an other scale

$$P_{q_i q_k}(z) = \delta_{ik}\left[\frac{4}{3}\frac{1+z^2}{(1-z)_+} + 2\delta(1-z)\right],$$

$$P_{q_i g}(z) = \frac{1}{2}\left[z^2 + (1-z)^2\right]$$



# PDF- Fits: Inputs and evaluation

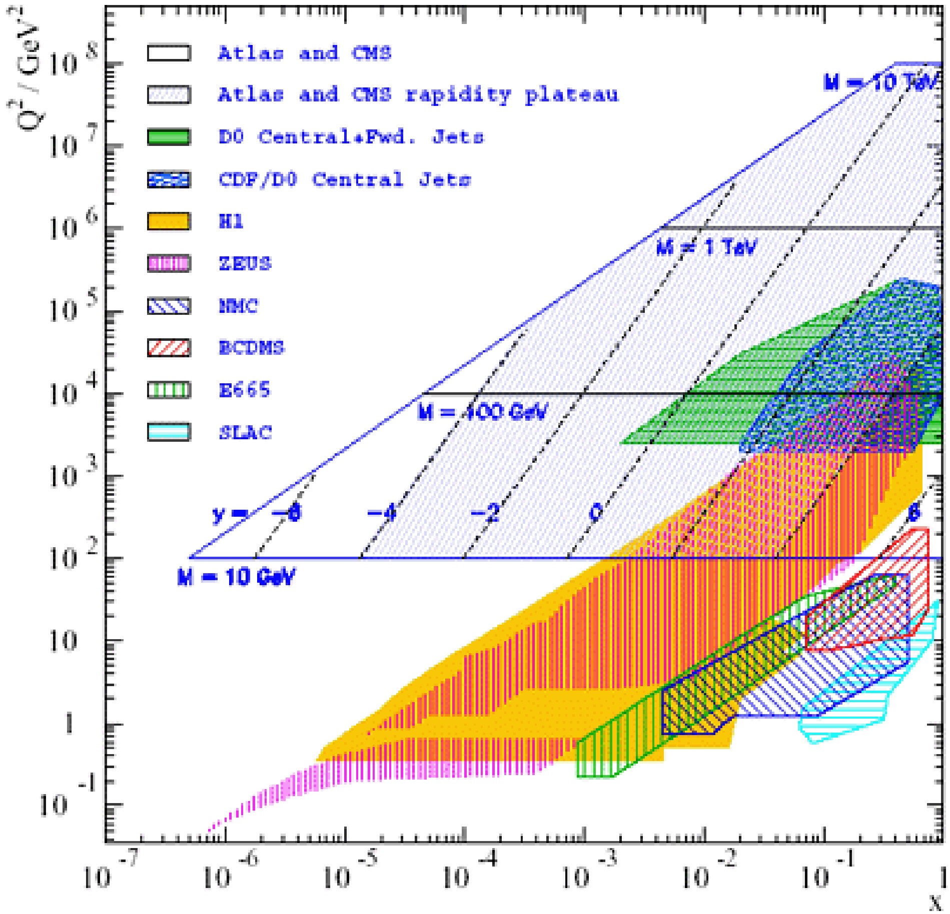



FIG. 1: Kinematic coverage of the DIS and collider $pp$-$p\bar{p}$ experiments. For $pp$ and $p-\bar{p}$ colliders, the Bjorken $x_1$ and $x_2$ of the interacting quarks are related to the mass $M$ of the Drell-Yan pair and its rapidity $y$ as $x_{1,2} = M/\sqrt{S}\exp(\pm y)$ where $S$ is the center of mass energy squared for the experiment.

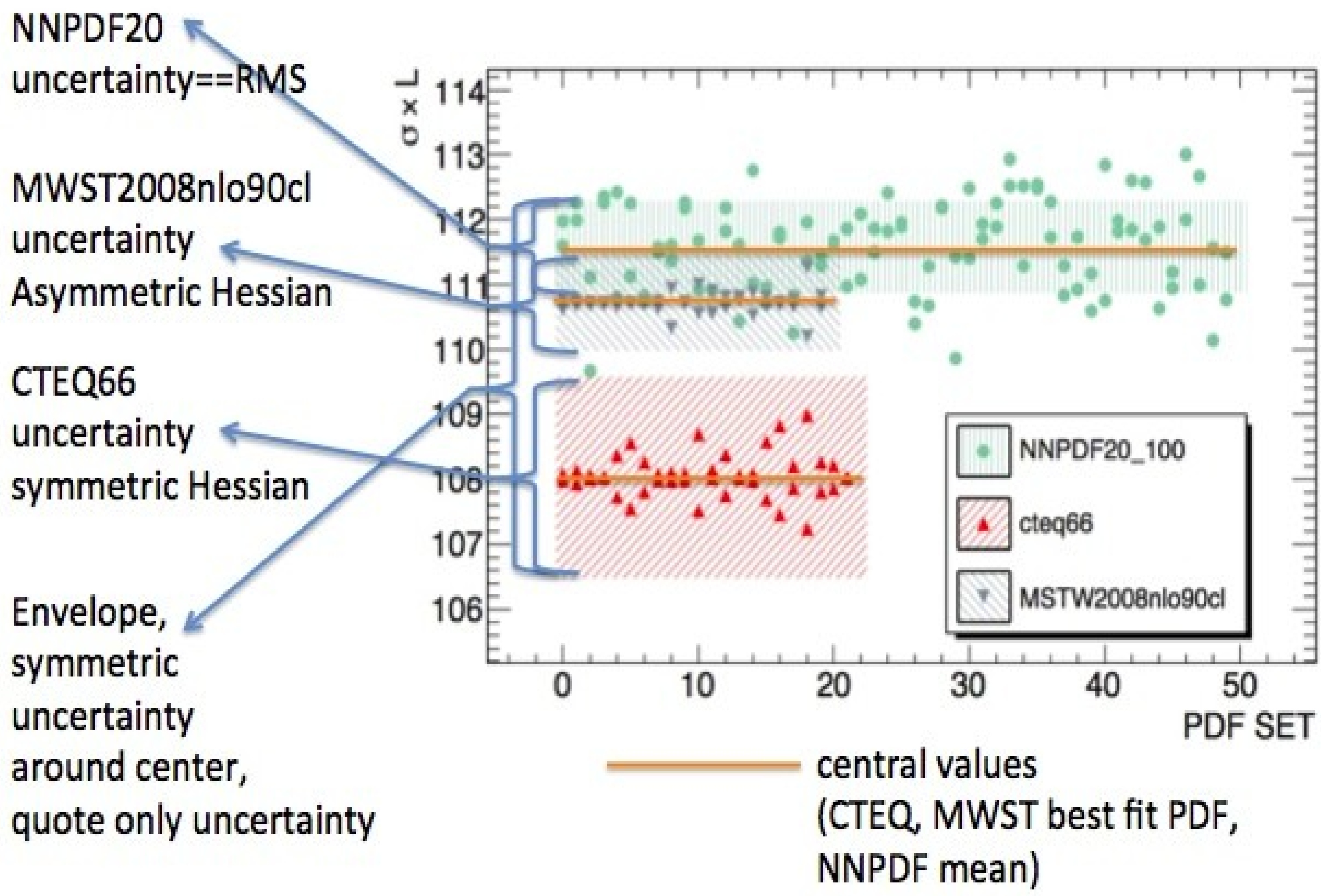
- Interpolation/Fit:
- Evolution:

# LHAPDF(6)

- Old generators have their own PDF format
  - You need to ask the author to includes the PDF
  - Sometimes fast, sometimes it takes months
- LHAPDF is a common format supported by all new generators
  - CTEQ
  - MRSTW
  - ABM
  - HeraPDF/AtlasPDF (no Fix target)
  - NNPDF (neural network instead fit)
- It includes:
  - All modern PDFs from the PDF groups
  - Variations of the PDFs (without fix target, only Hera)
  - Alphas variations, error vectors or pseudo experiments for doing systematics (inclusive prescription for the estimation)
- Errors are different estimated for every PDF
- PDF do not always agree inside errors (assumptions in the fits)

# LHAPDF(6)

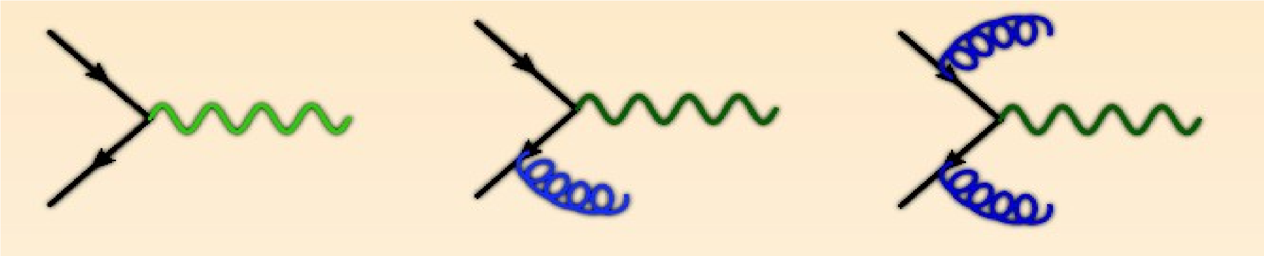
- ttbar x-section PDF errors:



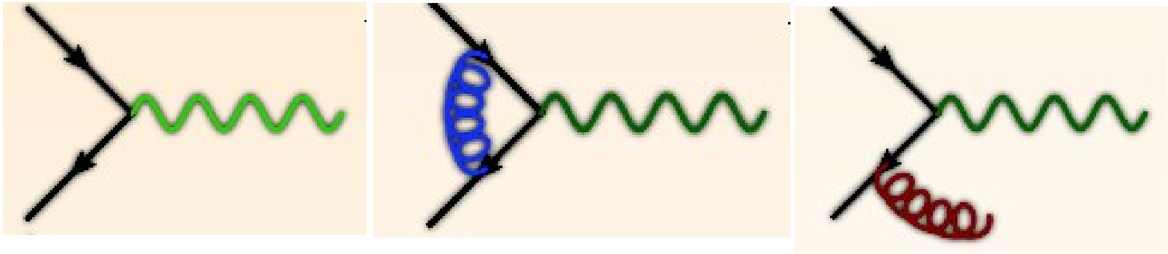
# Module II: Matrix Elements

- Matrix element: Simulation of the hard interaction
- Two main developments in new generators:
  - LHC has huge phase space, many jets in events
    → Multi-leg at Leading Order

    W + W+1jet + W+2jets + W+3jets + …..

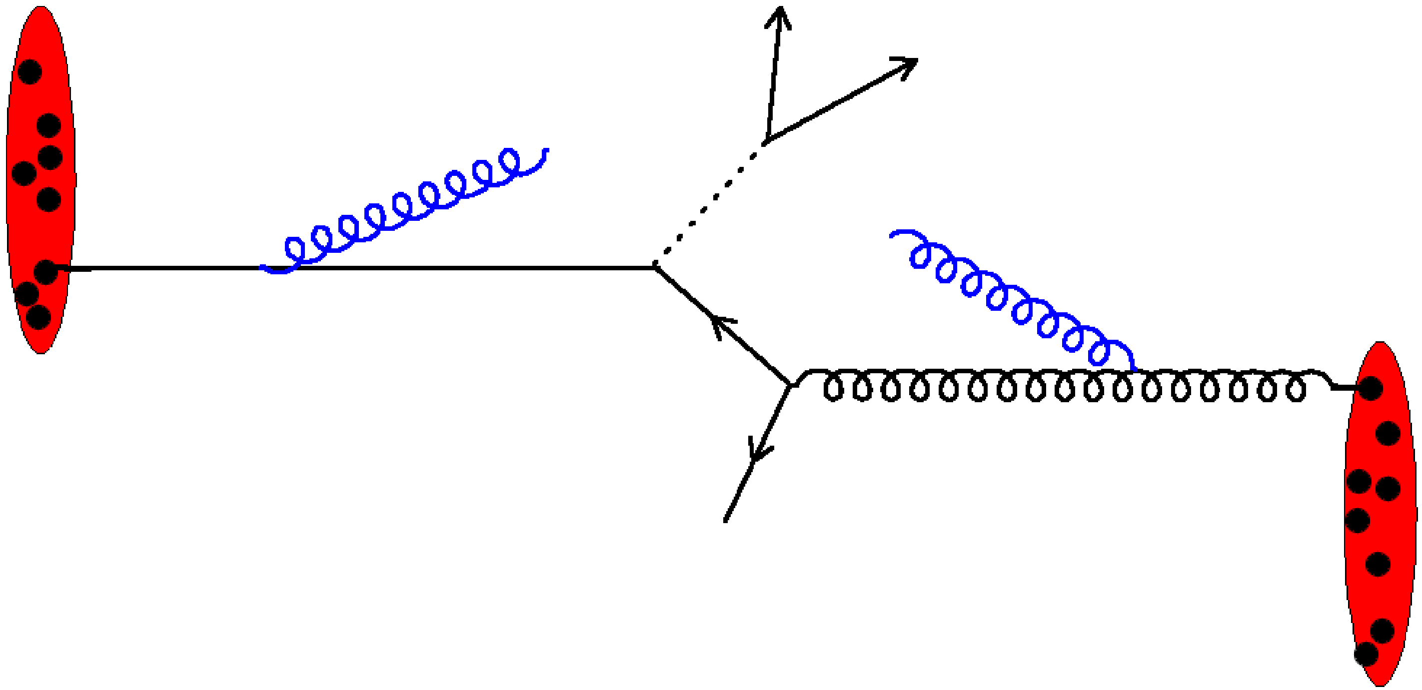


  - Fixed Order at higher precision: (N)NLO



- New development (last two years): Multi-Leg at NLO

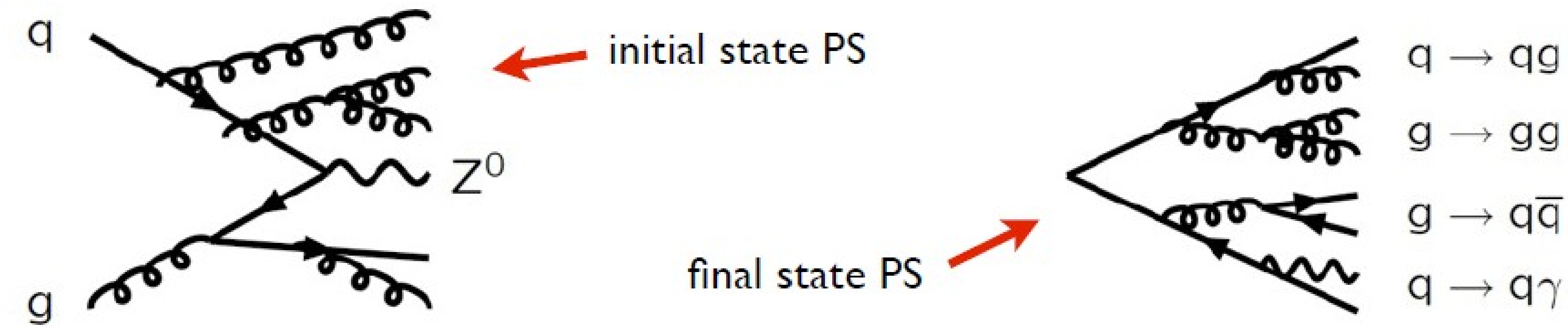
# Initial state radiation



- Radiation of gluons (photons) in the color (em) field of the other proton
- Usually part of the parton shower: Inverse engineering:
  - First do the hard interaction (PDF and Matrix Element)
  - Then do the embedding in a proton/proton interaction

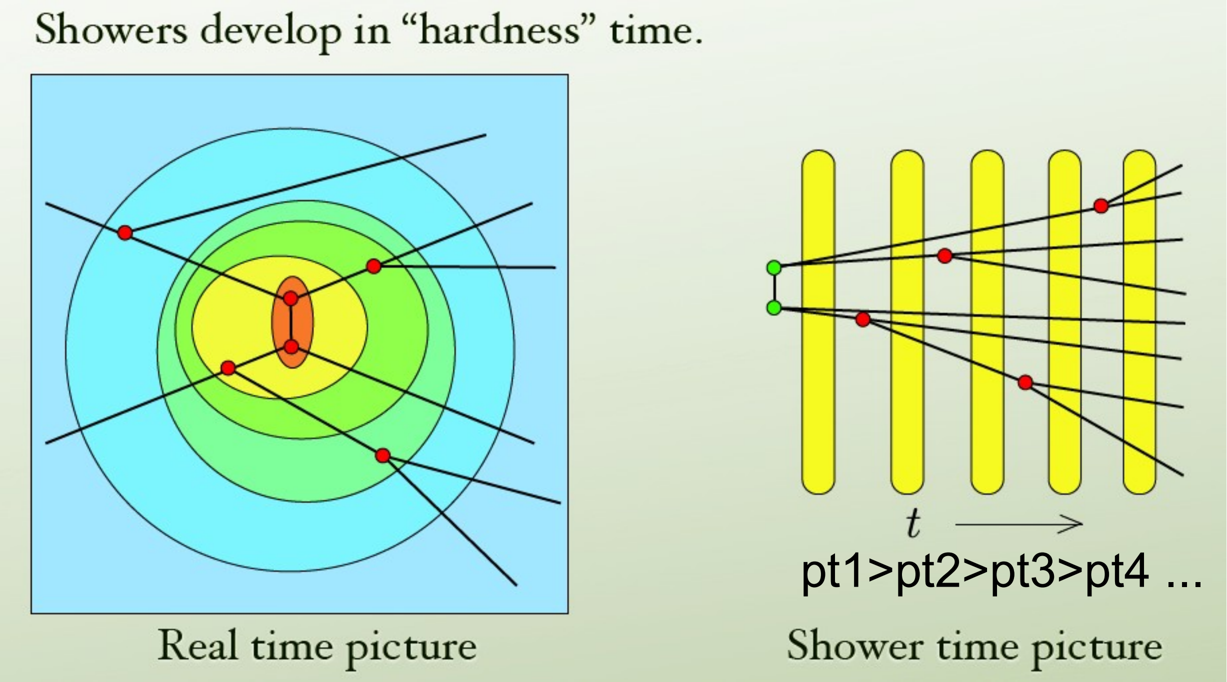
# Parton shower

- Hard and wide angle radiation of additional partons are well described by the partonshower → Multileg generators LO/NLO accuracy
  - Divergence for collinear and soft radiation (infra-red divergence)
- Parton shower describes collinear soft part:



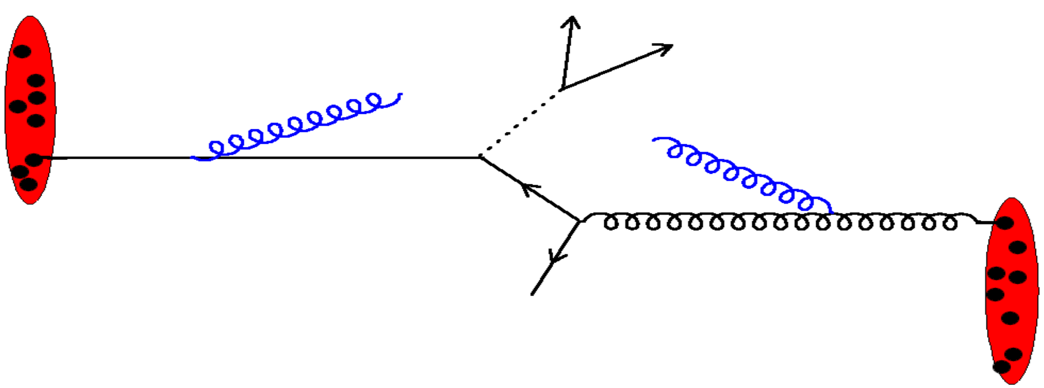
# Parton shower (II)

- Calculation evolute in time in one variables (angle, pt)
- Today two schemas used: pt ordered shower (Pythia), angular ordered (Herwig)
  - Pt-ordered: start with the highest momentum emission (ISR legs are mirrored in time)
  - Veto all emissions  with a pt higher then what is included in the ME: Matching



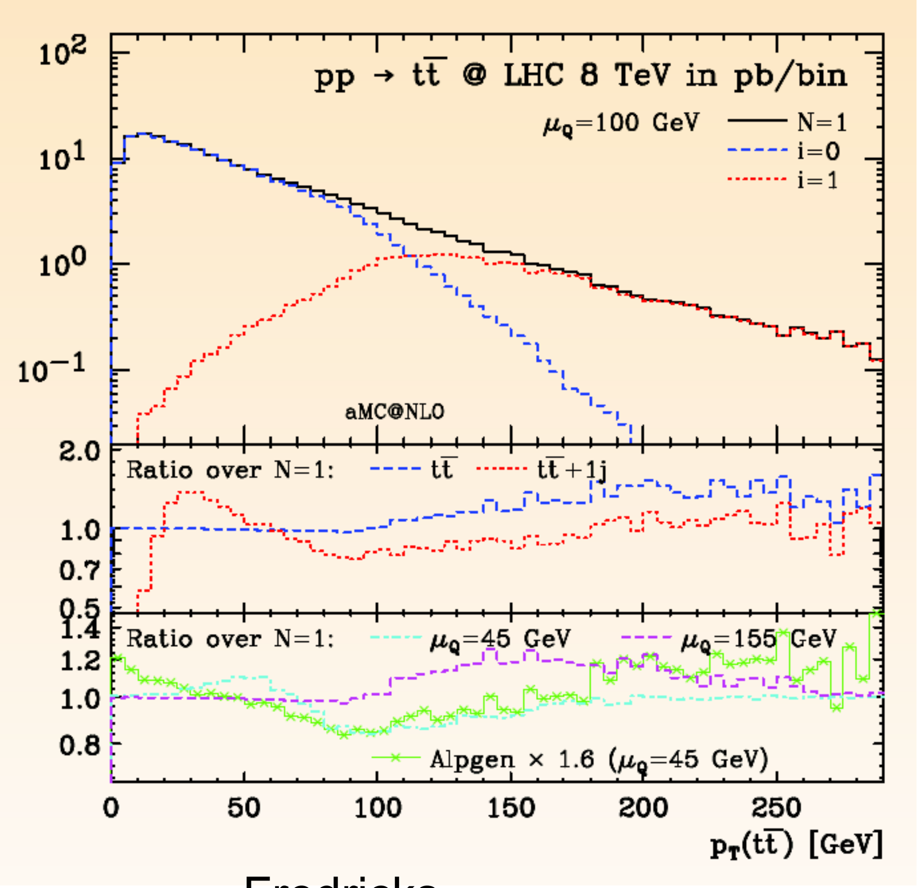
# ME to shower transition: Matching/Merging

- Parton shower does not change the cross section for an given event
- Matching → veto vs all radiation in PS harder then what is simulated in the matrix element (Powheg-NLO)
  - Some freedom: which particles are included in the matching (only incoming legs of the ME, incoming and outgoing, do we add the proton flight direction), still a lot of investigation, mostly by experimentalists (comparing what describes data best)
  - Better: full theoretical calculation instead of veto (subtract full overlap of shower in calculation, MC@NLO)



# ME to shower transition: Matching/Merging

- Merging: combination of different Multi-Leg histograms of different orders (MLM/CKKW)
  - Well established in LO (15 years), new in NLO (sherpa, aMC@NLO, FxFx, 2 years)

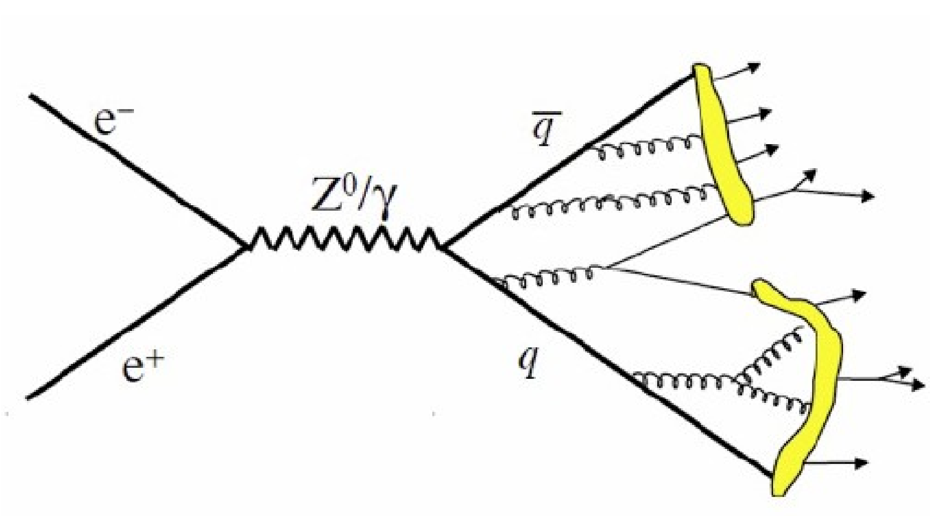


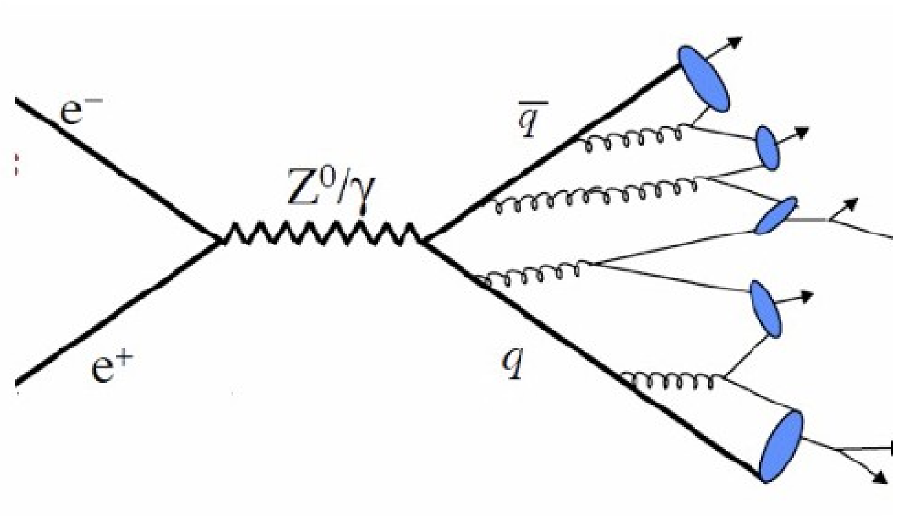

Fredricks

- aMC@NLO merged at pt(ttbar)= 100 GeV:
  - ttbar+0jet
  - ttbar+1jet
- Computing power limitation (o~1 day)
  - NLO: ttbar+2jets (2 days of phase space integration
  - LO: up to 4 jets (sherpa)

# Fragmentation/Hadronization

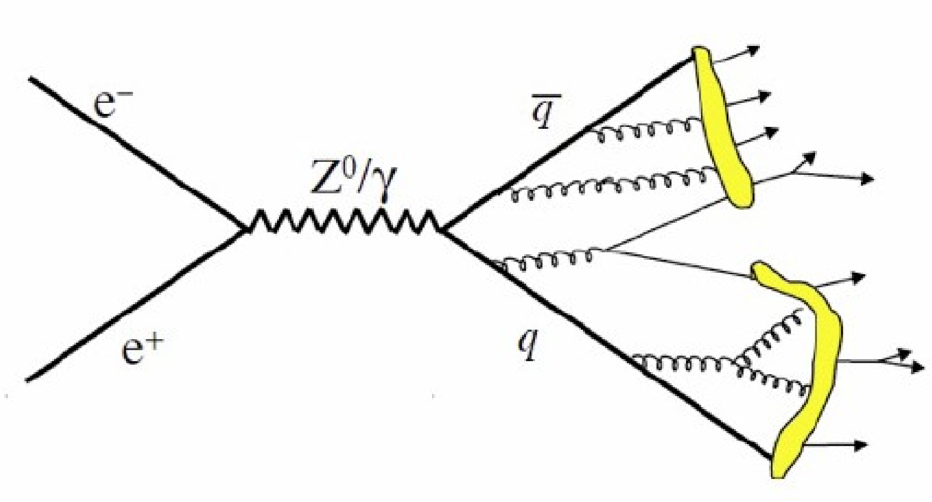
- We do not observed free quarks and gluons
- Parton shower stops when alphas gets big
- Phenomenological models to describe transition between quarks/gluons and hadrons
- String fragmentation                                   cluster fragmentation



# Fragmentation/Hadronization

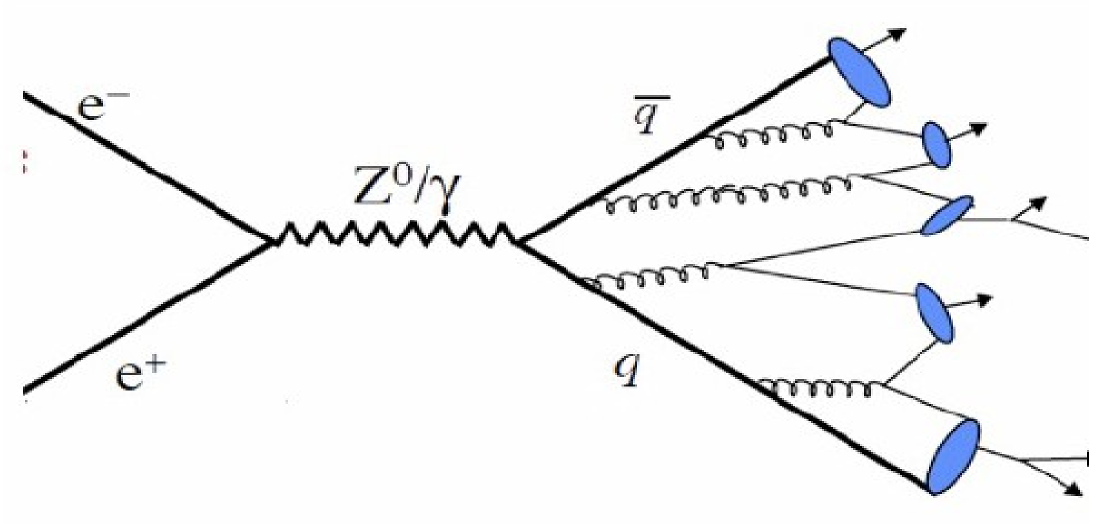
## String fragmentation




- Quarks are connected to colorless strings, gluons in between
- string produce tension
- if tension to big: new quark/antiquark pair
- at the end: hadrons formed by the quarks

# Fragmentation/Hadronization

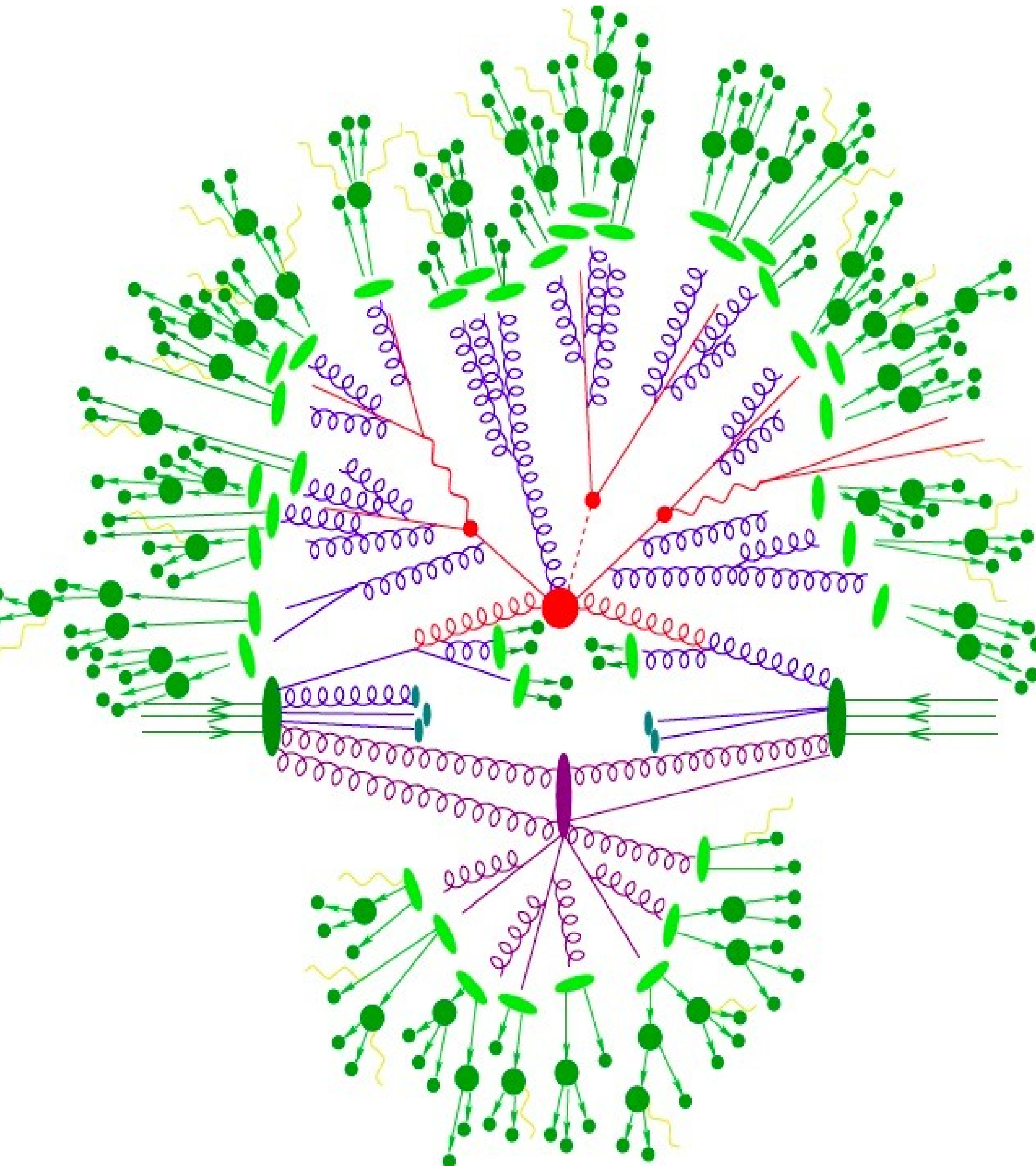
- Cluster fragmentation:




  - gluons are splitted into quark-pairs
  - clusters of quarks formed
  - hadrons formed dependent of cluster energy

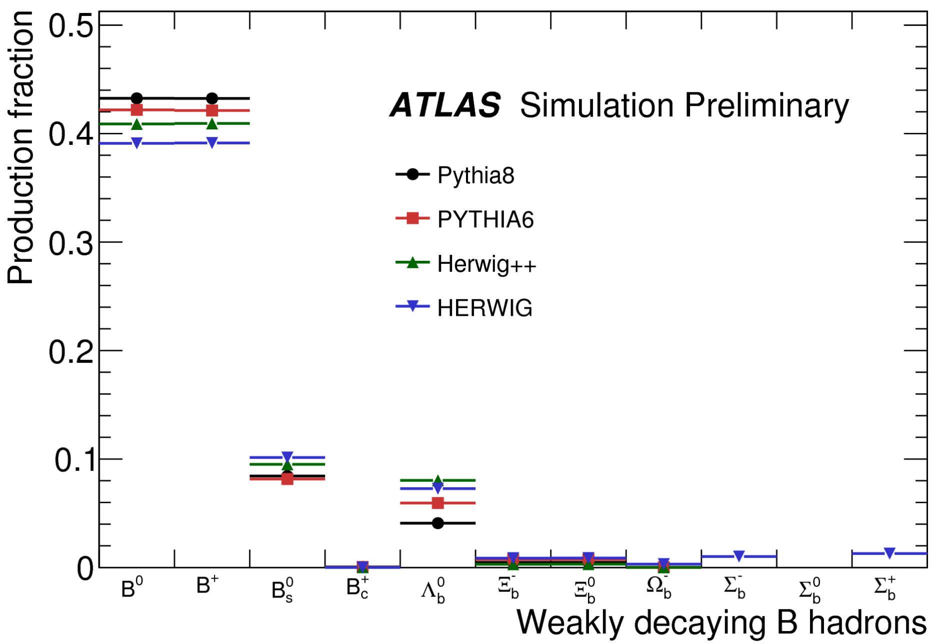
- Both models often compared to estimate systematics from Fragmentation

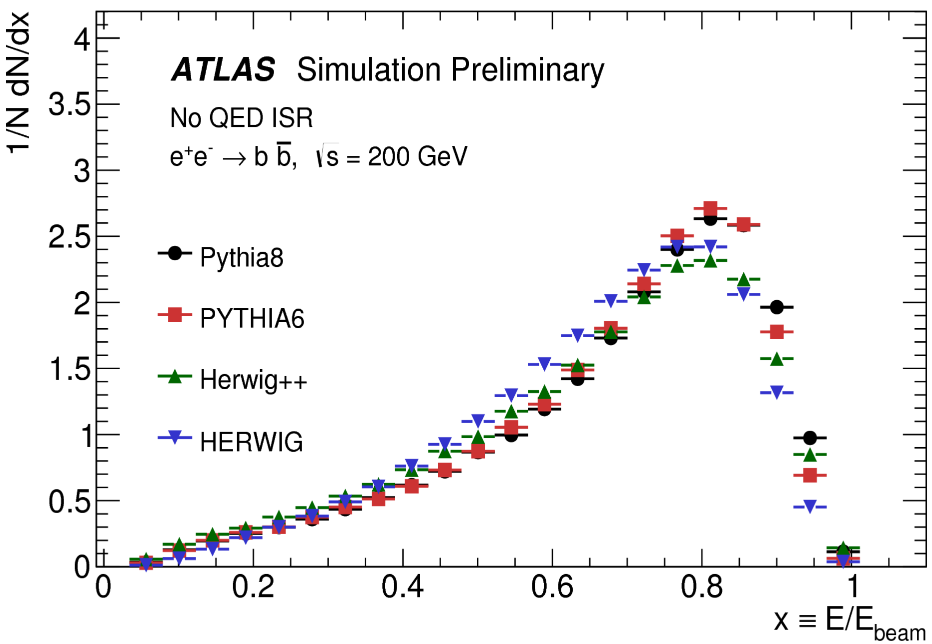
# A more realistic event



- More realistic event (display of an sherpa event)

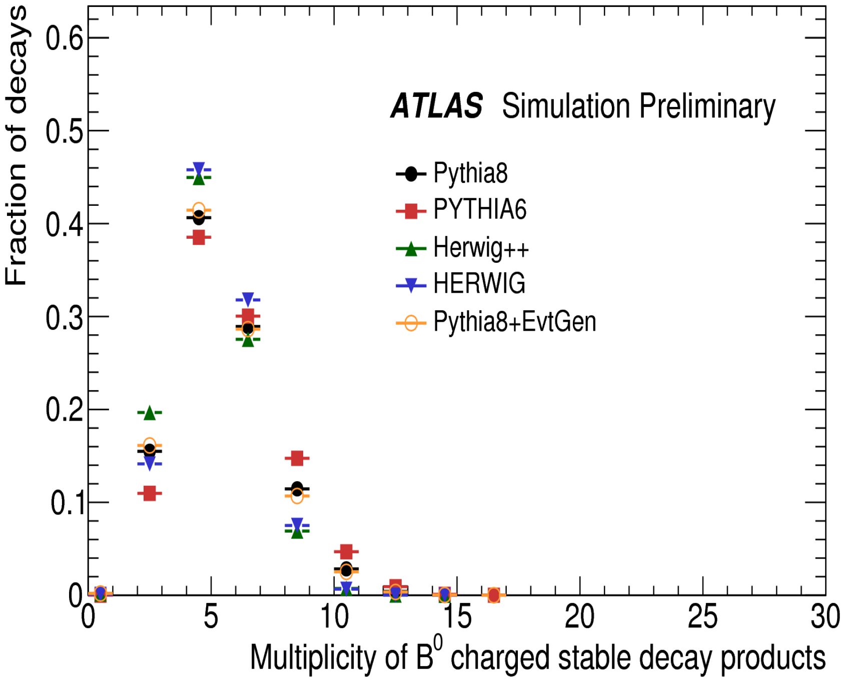
- Hadron spectrum are defined by the shower generator
  - Content sometimes differs much, dependent of fragmentation/hadronization model
  - Momentum defined by fragmentation function

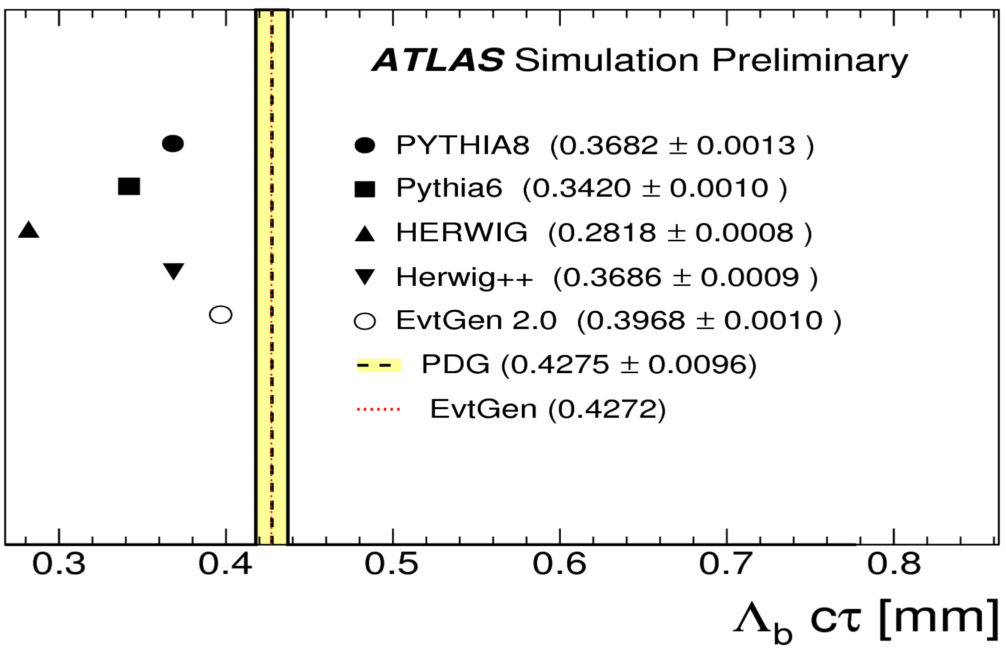


# Decay of short lived particles

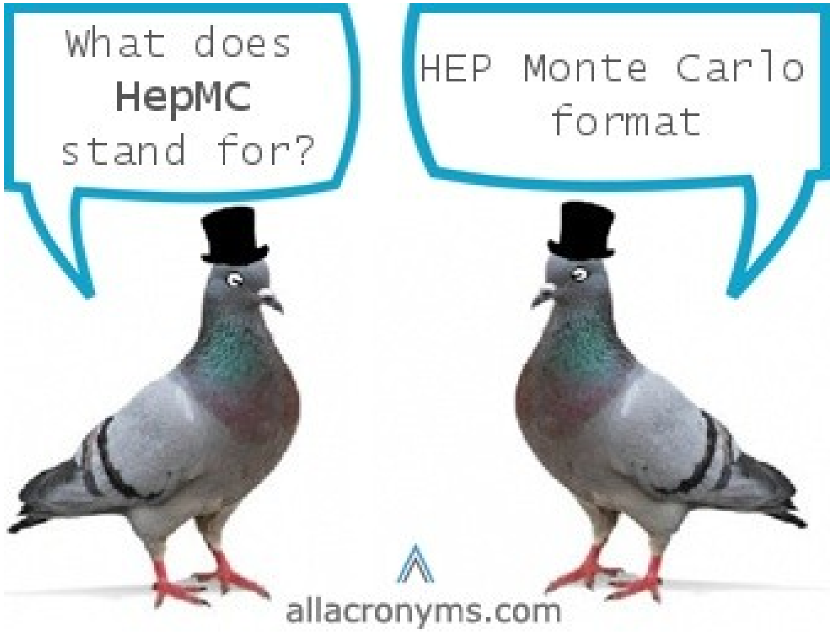
- EVTGEN: Data-Base tool to have a common description of electroweak B/D-Meson decays
- Best detector simulation does not help, if these numbers are wrong (different tagging performance)



# Interface to the detector simulation

- LHC: commen C++ structures to handle Generator events HepMC:
  - Classes containing the full event:
    - particles
    - vertices
    - general information and systematics
  - Easy navigation: bi-directional pointers



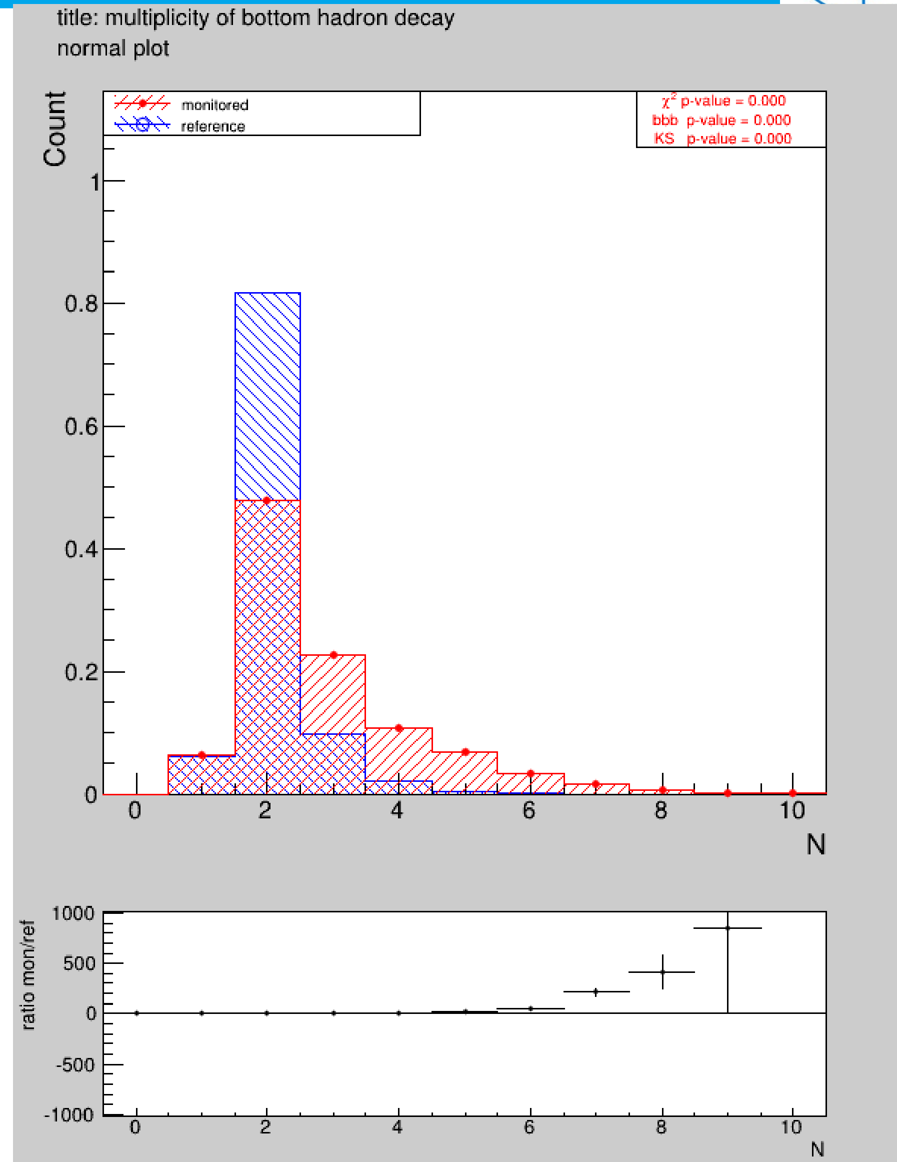
# Timing

- https://indico.desy.de/conferenceOtherViews.py?view=standard&confId=9240
- LO Generators completely dominated by infrastructure (LHAPDF, decay tables …)
- Very complex LO-ME generators are dominated by integration of the ME:
  - W/Z+4jets integration in one job: ~ 1 month
  - Very old: run every process with independent initial/final state configuration as separate process/production

    - very production system unfriendly, missing jobs can destroy physics
  - New (Madgraph5/aMC@NLO): Run every process as independent job but mix them in event generation step, still you have to wait for the slowest histogram (many gluons), throw away calculation, which effects the sample smaller then epsilon
  - Best (Sherpa 2.X): Run many jobs integration all processes in parallel, iterative collecting and resending, real parallel processing

# Validation

- Modern generator have hundreds of switches
- Validation is needed before large scale production
  - Change of default settings
  - “stable” gluons
  - B-Hadrons decaying into quarks
  - Here: change of format: double counting of b-decay products due to EVTGEN
- Auto-Validation by a job attached to production job based on analysis tools based on common HepMC format



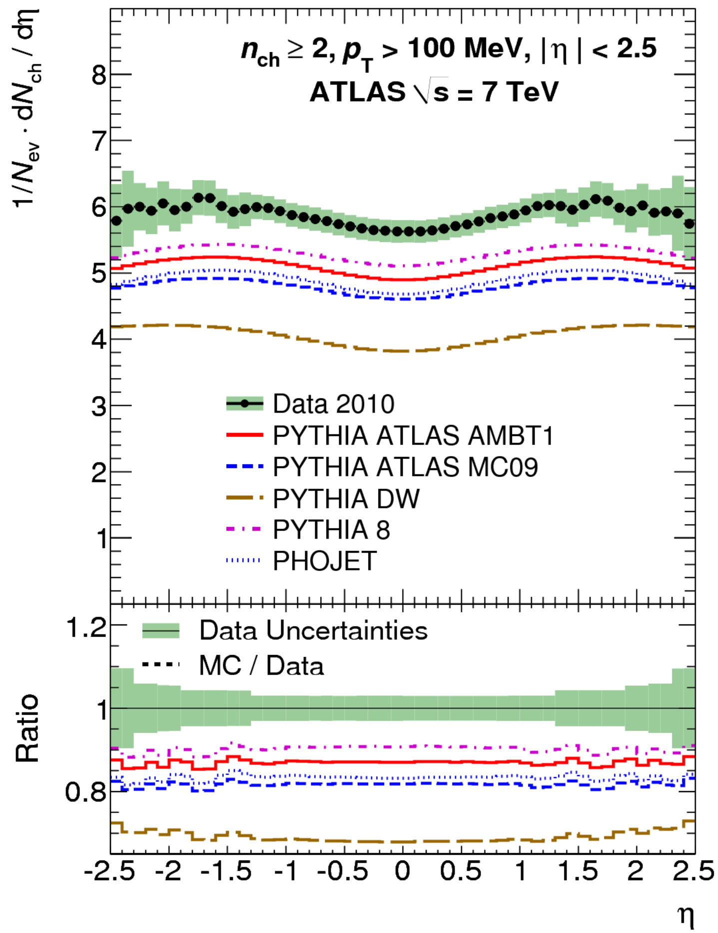
# The code armada

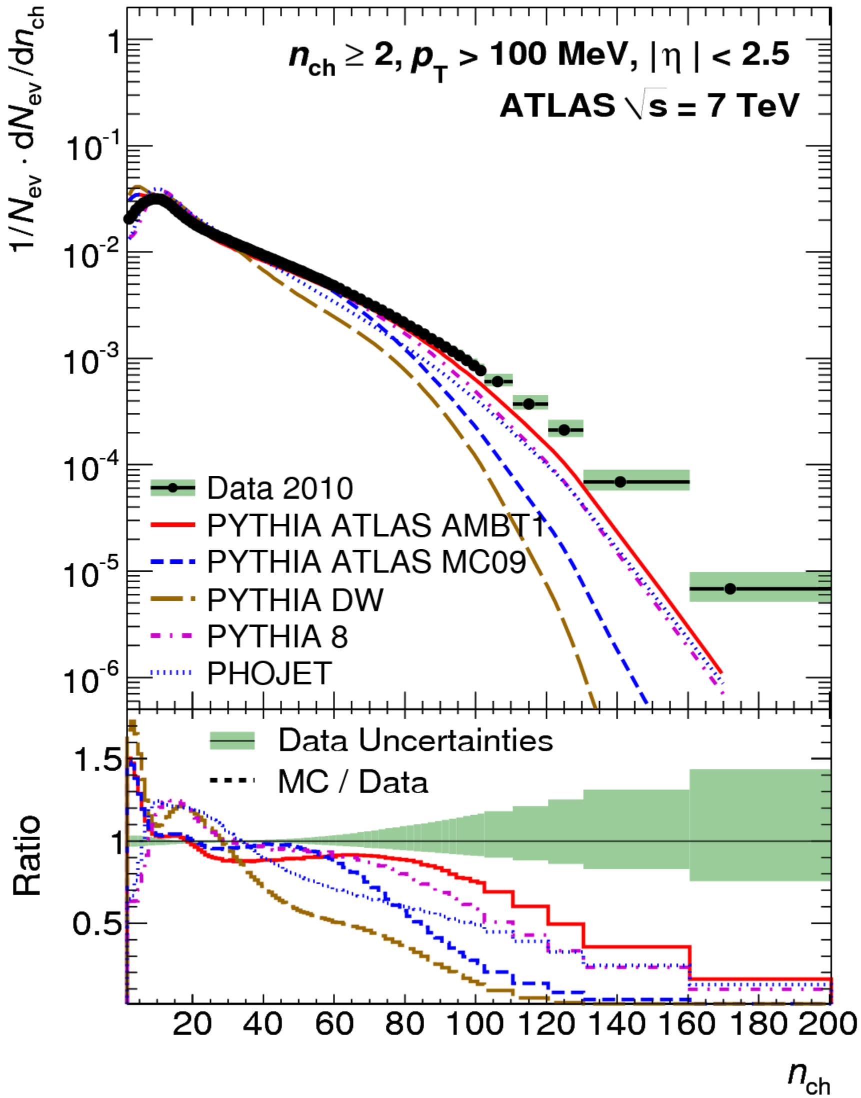
- PDFs: LHAPDF including CTEQ, MRSTW, NNPDF, ABK, HeraPDF/AtlasPDF with many different versions, alphas, flavour schemas, systematics
- Pure Matrix element generators: Alpgen (LO), Madgraph5/aMC@NLO(NLO/LO), Powheg (NLO), Acer(LO), gg2VV, Wizard
- LO and shower generators: Pythia6, Pythia8, fHerwig, Herwigpp
- All inclusive: Sherpa (LO/NLO+shower)
- Matching: Veto (Powheg), Subtraction (aMC@NLO)
- Merging: MLM, CKKW, FxFx
- Tau/Bdecay Models (afterburner): Tauola, EVTGEN
- Interfaces on position were you want to choose: LHAPDF for PDF, lhe_f for matrix elements, common event record for afterburner

→ Many possible combinations, but only few are valid (parameter settings etc

# Performance: Comparisons to unfolded data

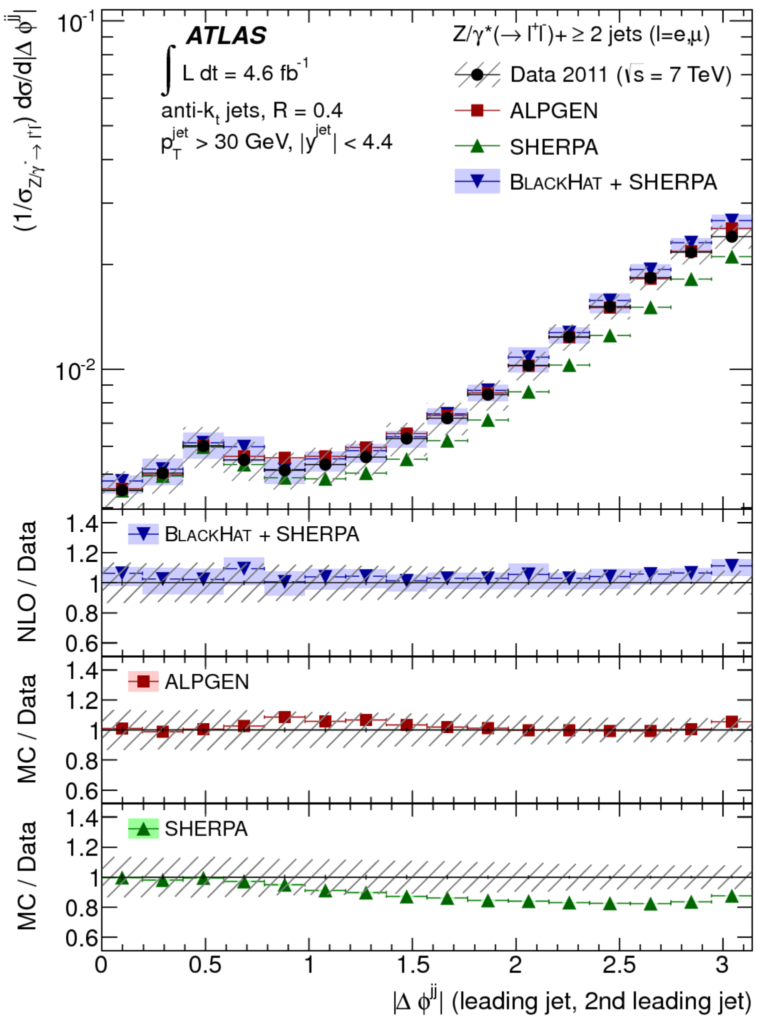
- Minimum Bias (most inclusive) (basically pure soft QCD)

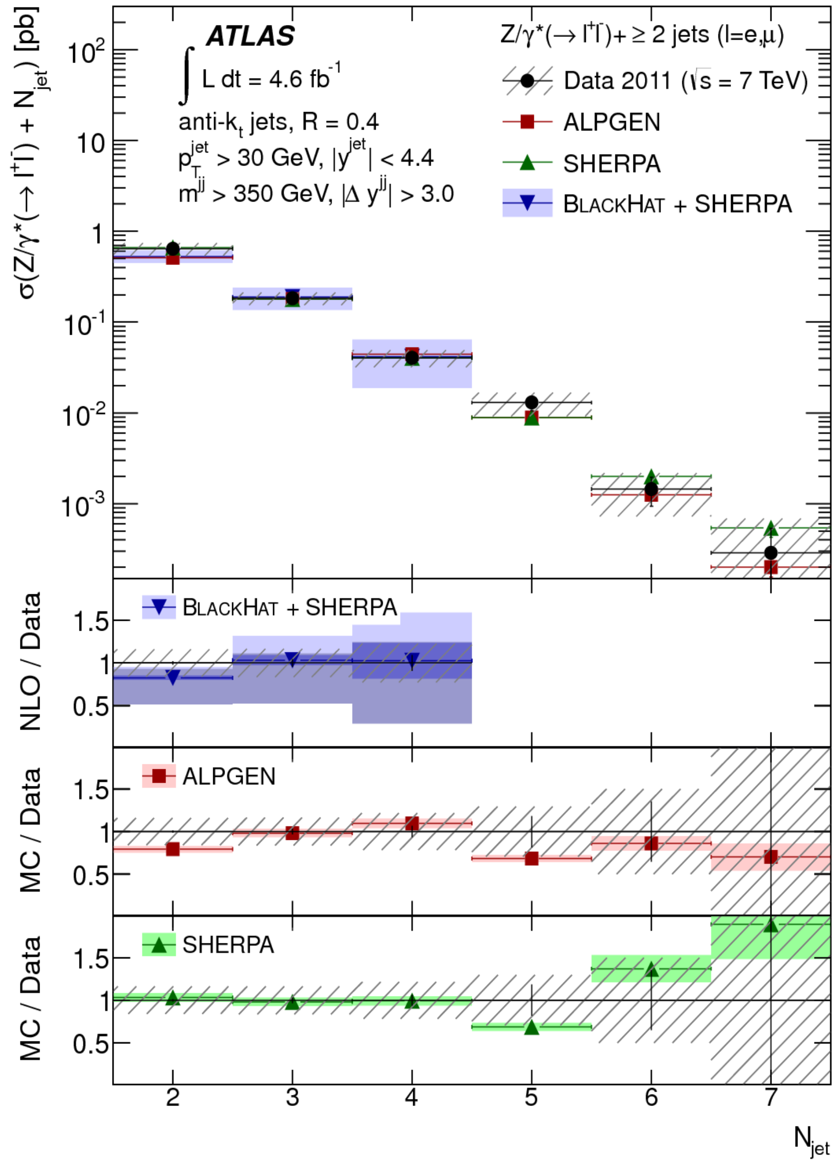


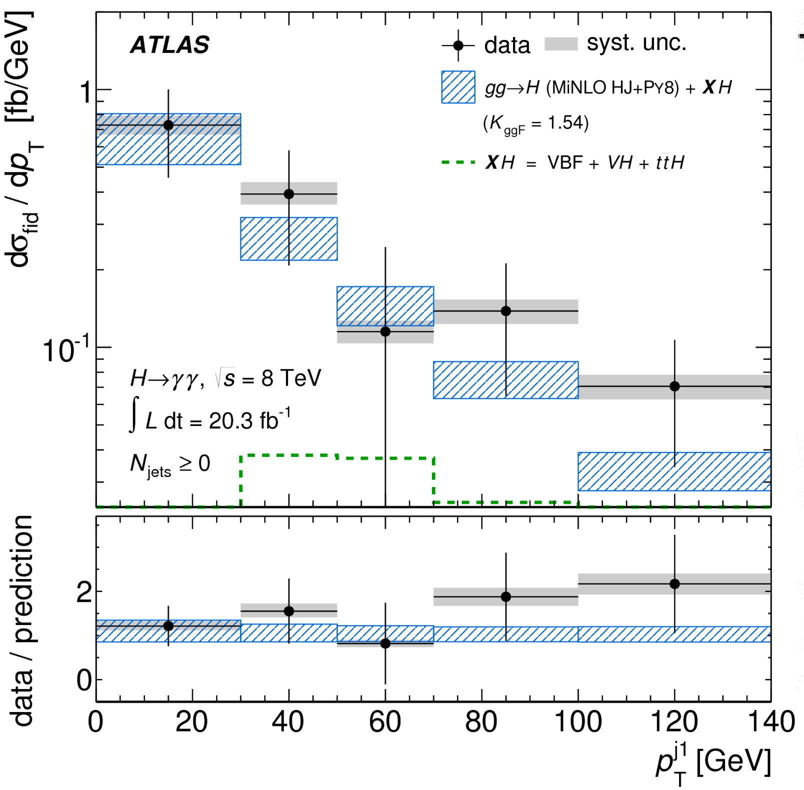
# Performance: Z+jet

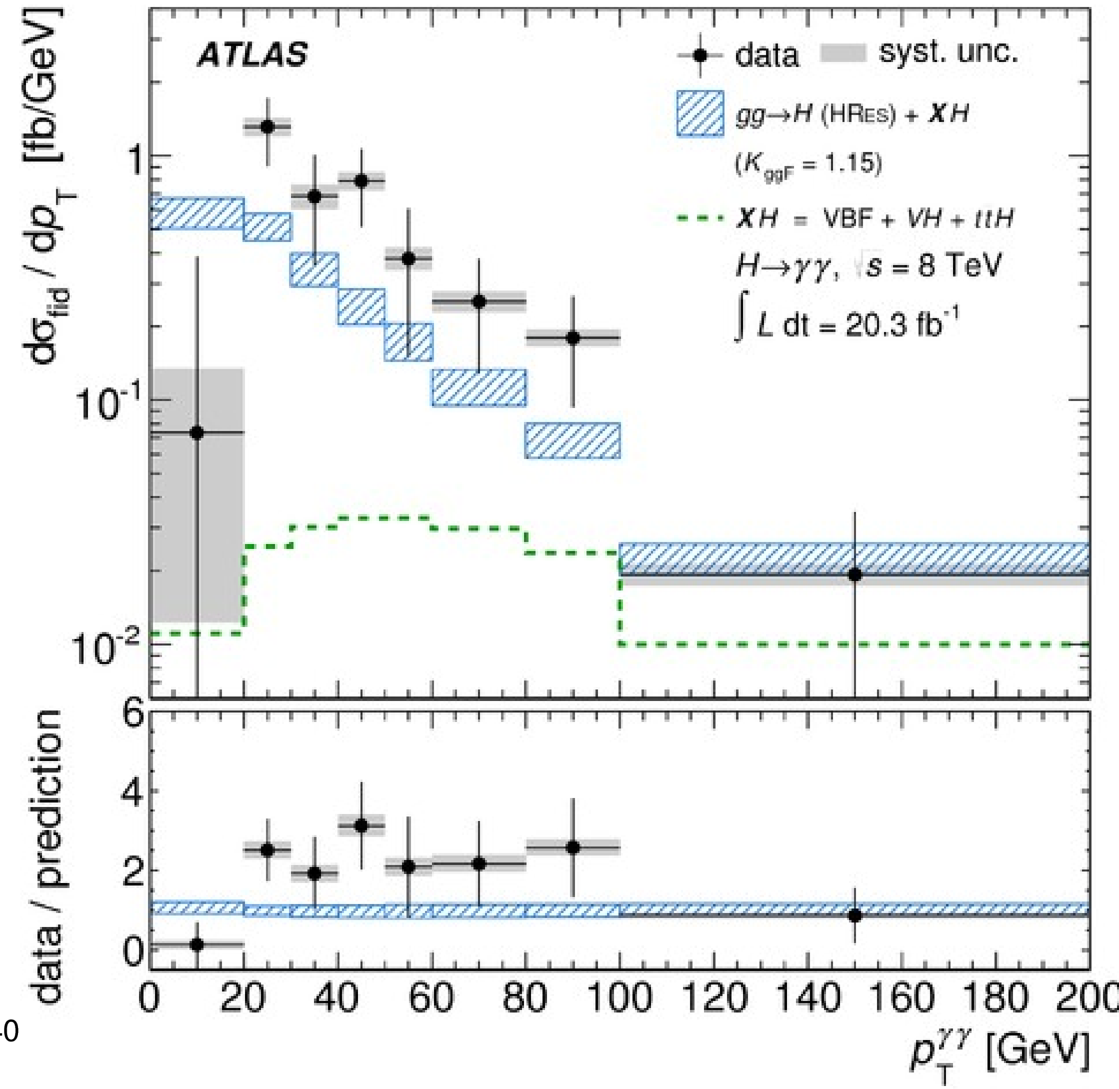
- Multileg bread and butter physics:



- The new player

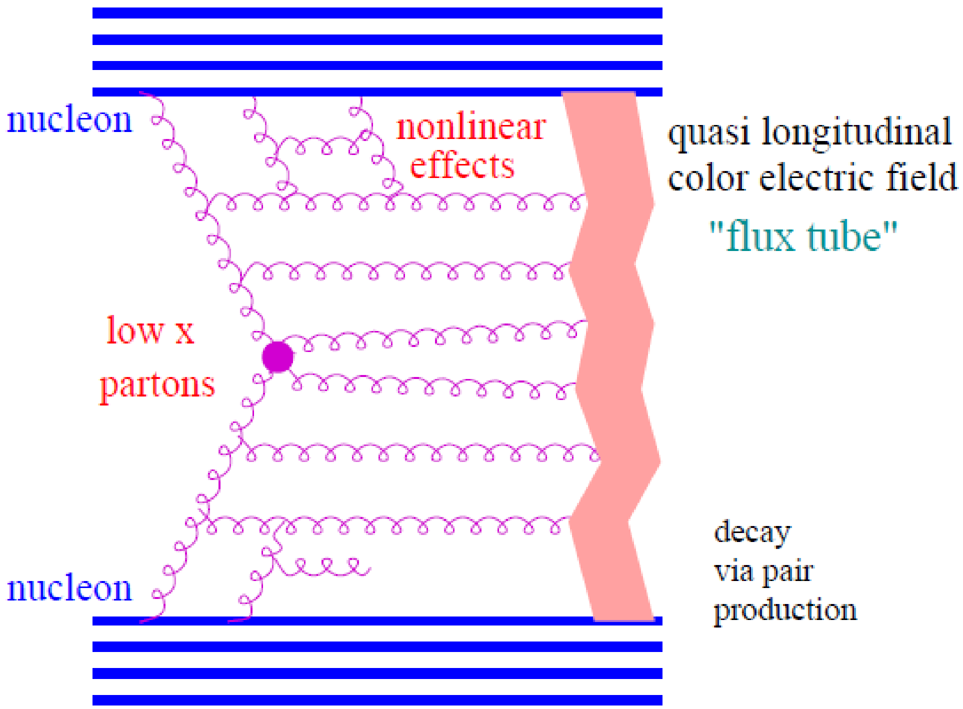


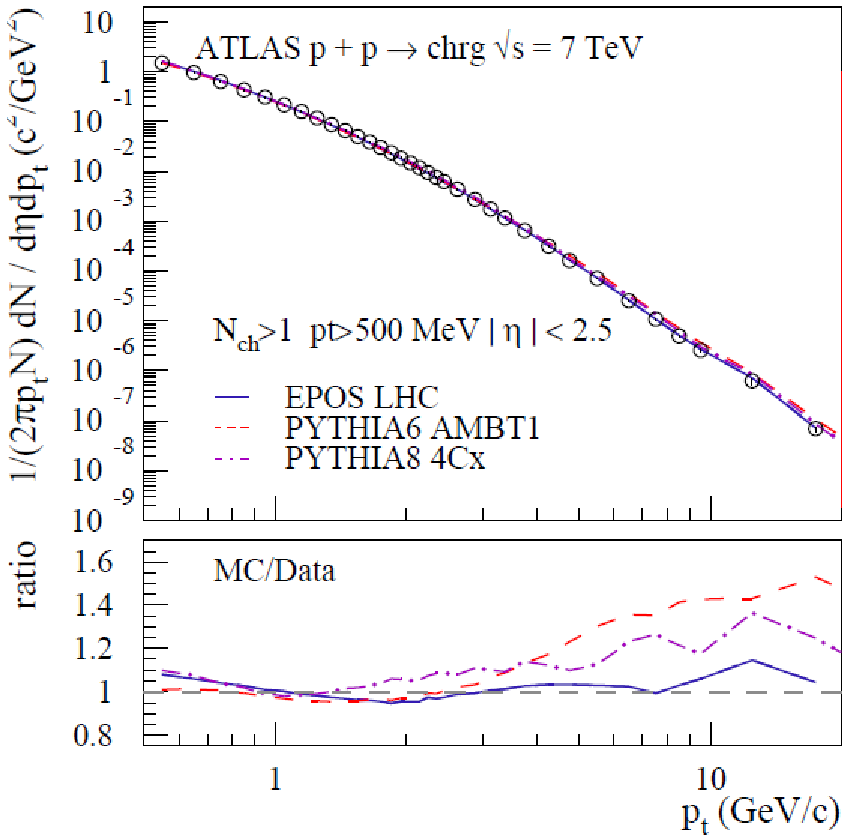



# Connections AstroParticle Physics and HEP

- General:
  - Particle physics: new physics high pt physics,
    → Events in central detector
  - Astropartice physics: Fix target, something hits air
    - first reaction highly boosted in forward direction
- Detector coveres different areas:
  - Particle physics: relative small eta, high Q2 (full energy), matching/merging is optimized for central events
  - Astropartilcle physics covers high eta, relative small Q2
- There are two connections:
  - Soft part of MinBias gets described well by Astroparticle generators (EPOS …)
  - Search for SUSY (new physics, high q2) in astroparticle physics often based on particle generators

# EPOS generator for MinBias in particle physics




- Improvement of Pile up simulation
- EPOS: Generator used in air shower and in heavy ion physics
- Best generator in forward direction and low multiplicity events
- Little drawback: no jets, try to combine with multijet generator

# Summary: Event Generation

- Event generation:
  - Modular description of the known physics on generator level
  - Incoming particles/PDF
  - Matrix-element
  - Partonshower/Fragmentation/Hadronization
  - Decays
- Common interfaces between steps
- State of the art (used in experiments:
  - LO-Multileg up to ~5 legs
  - Inclusive NLO calculation (Powheg/aMC@NLO/sherpa
- New developments:
  - NLO-Multileg (Madgraph5+aMC@NLO/sherpa)
- Time for simulation: 1/10 second to hours per event