Introduction to Monte Carlo generators: Matrix Elements

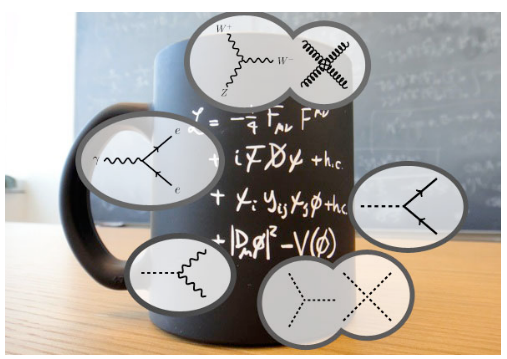
Rikkert Frederix
Lund University

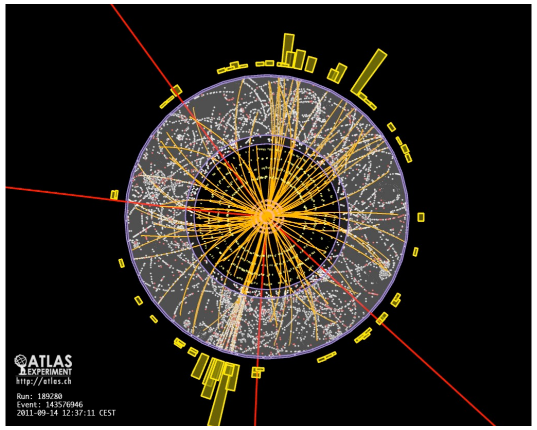
2024 CTEQ Summer School on QCD and Electroweak Phenomenology,
Hosted by the University of Münster, Germany

Bridging the gap

Theory

Lagrangian
Gauge invariance
QCD
BSM parameters
...





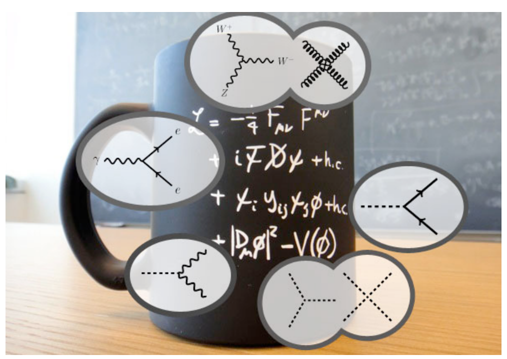
Detector calibration
Pions, Kaons, ...
Reconstruction
B-tagging efficiency
...

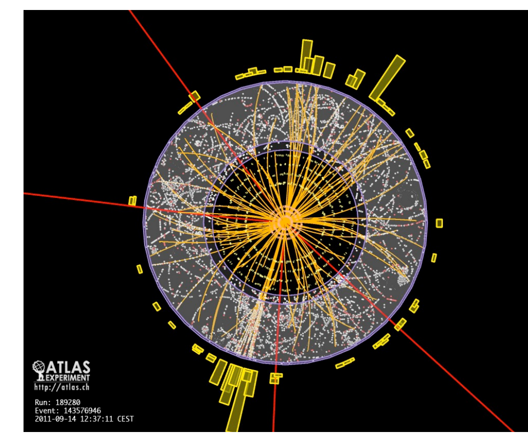
Experiment

Bridging the gap

Theory

Lagrangian
Gauge invariance
QCD
BSM parameters
...

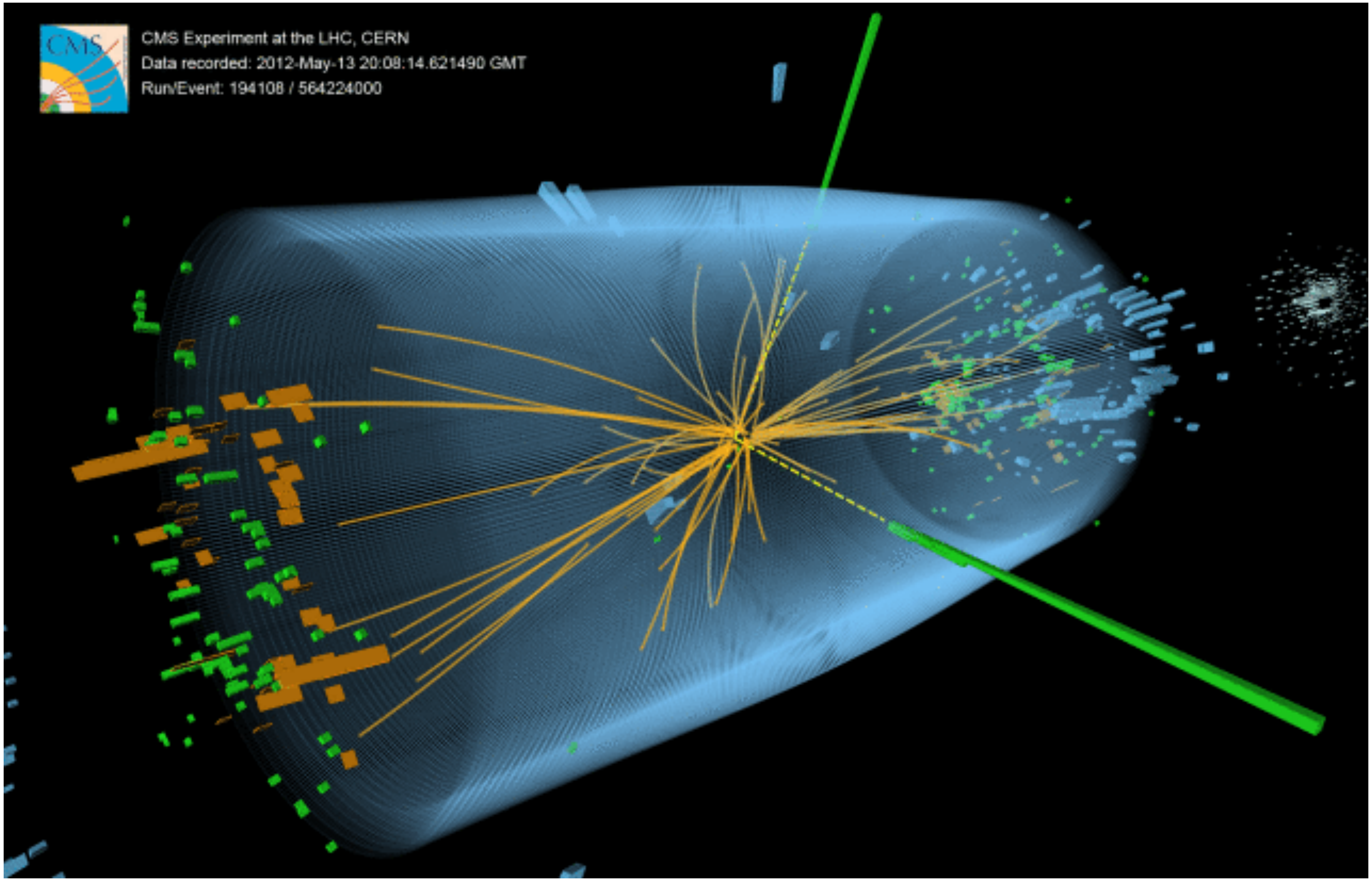




Detector calibration
Pions, Kaons, ...
Reconstruction
B-tagging efficiency
...

Experiment

Event display
Experimentalist's point of view



Event display: Theorist's point of view

LUNDS UNIVERSITET

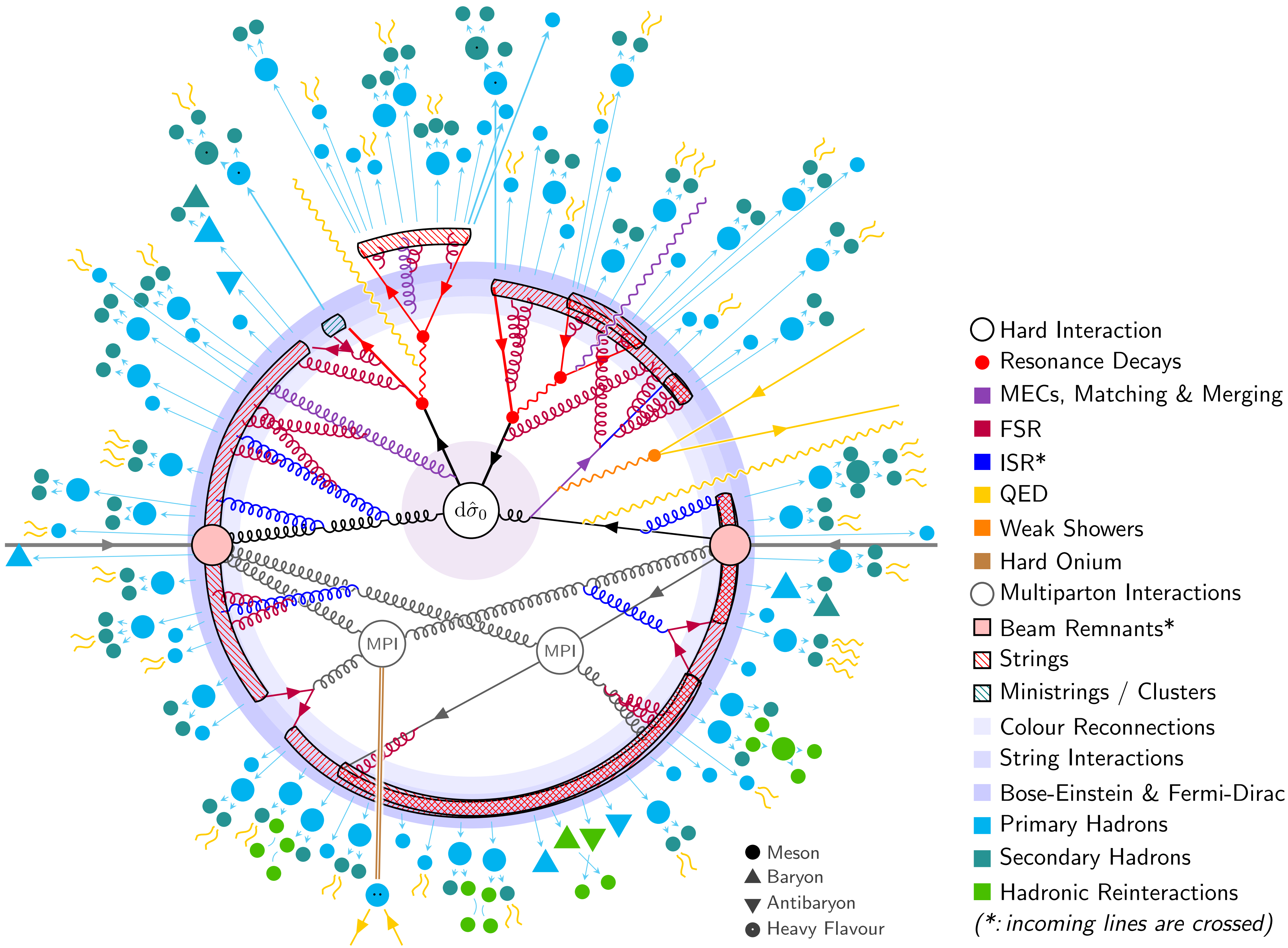
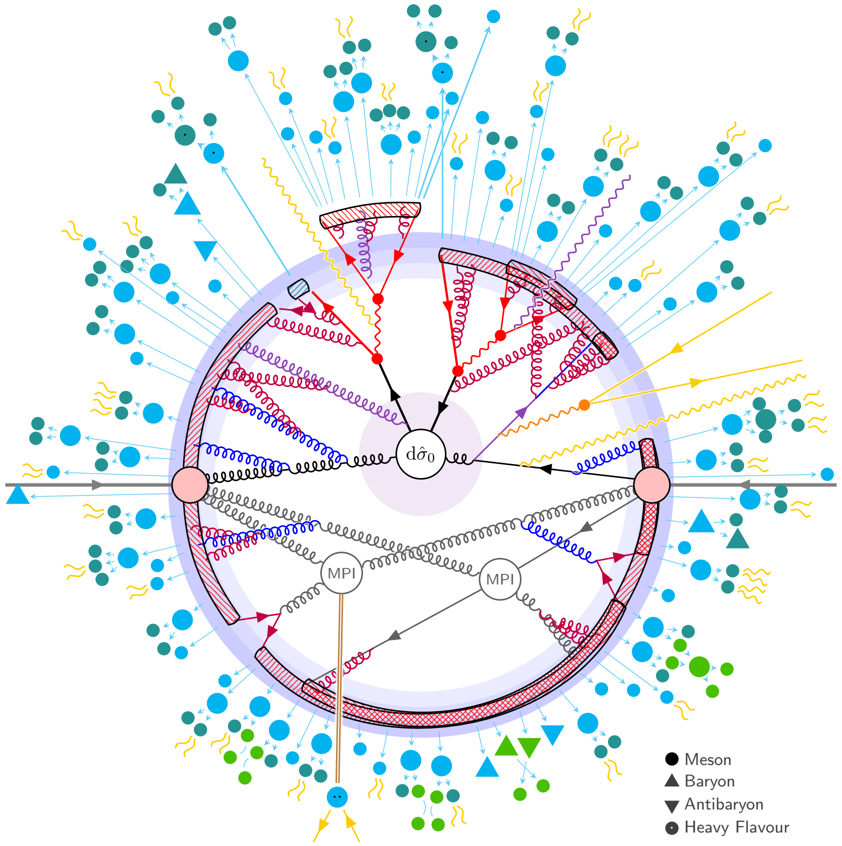
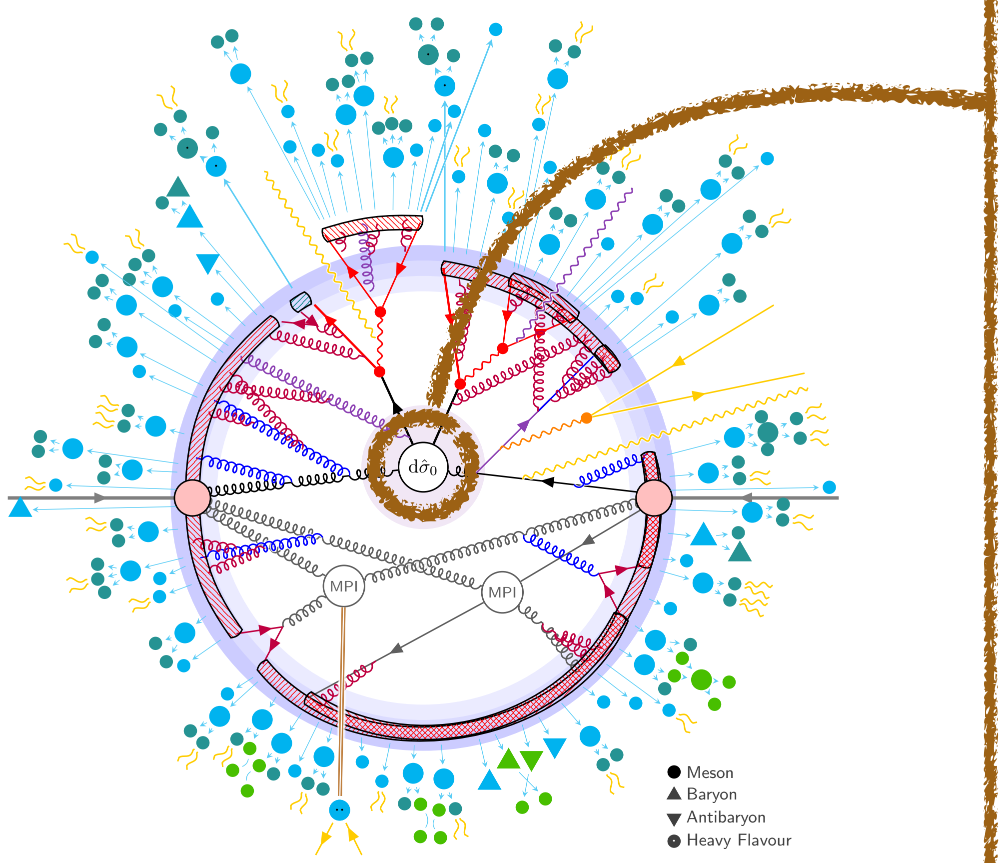


Figure taken from Bierlich et al., 2022 *(Pythia8.3 manual)*

The hard interaction

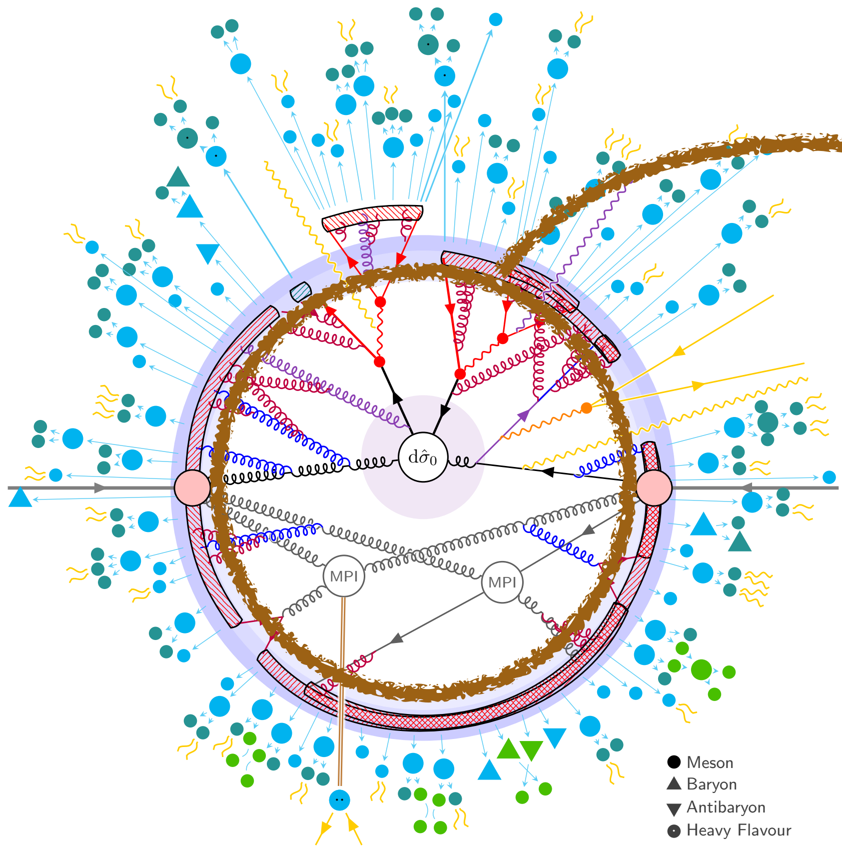


The hard interaction



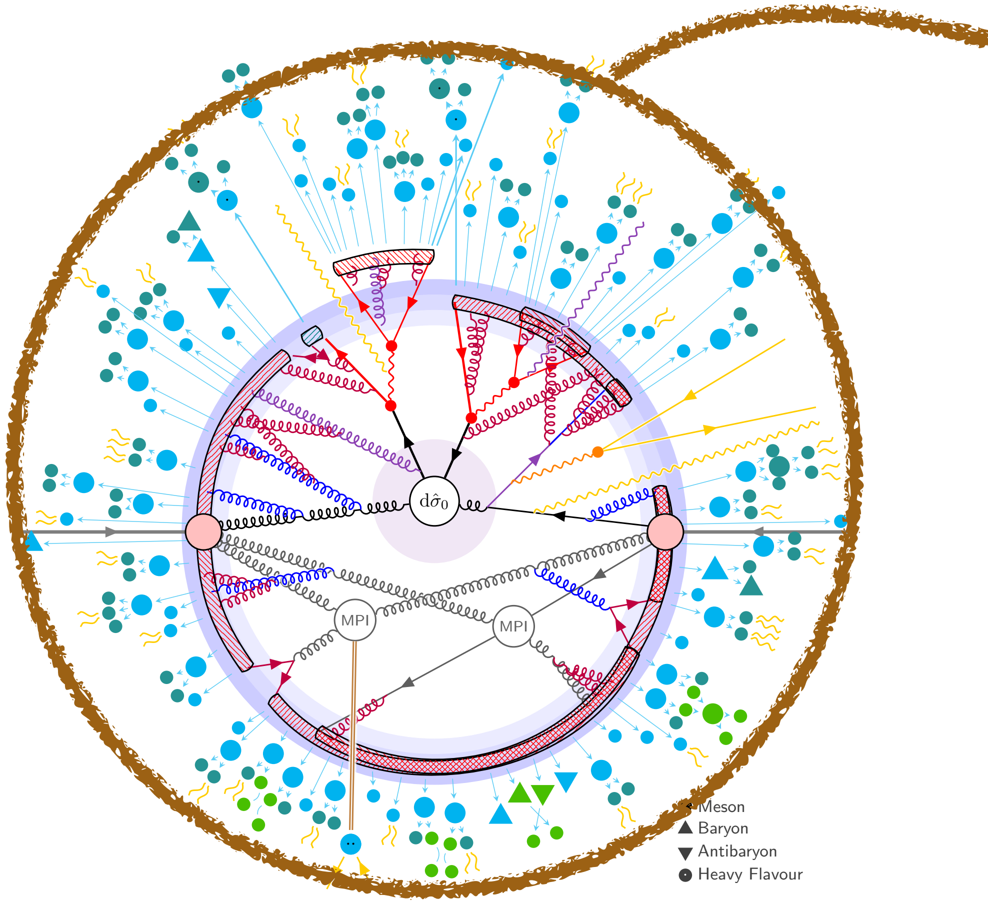
- Core of the event
- Process dependent
- First principles description
  - Largest energy transfers
  - New physics most-likely will appear here
  - Can use perturbation theory: LO, NLO, NNLO, etc.

The Parton Shower



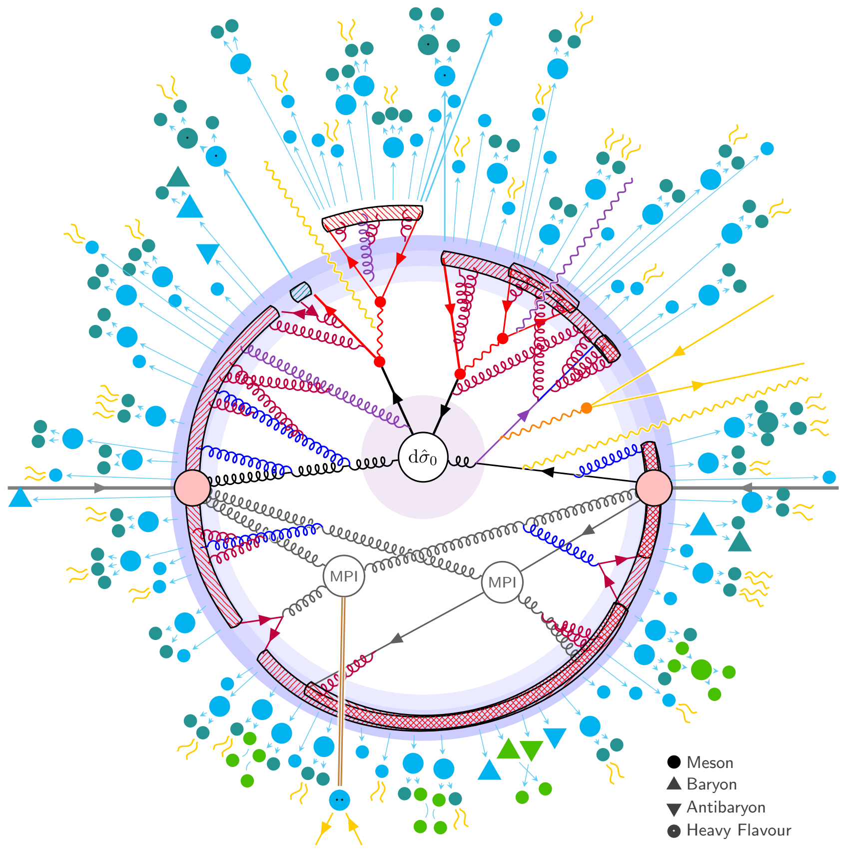
- Known QCD: first principles description
- Universal/process independent
- Can systematically be improved using perturbation theory

Hadronisation & Underlying event



- Low Q^2 physics
- Process and energy independent
- Based on models (motivated by physics)

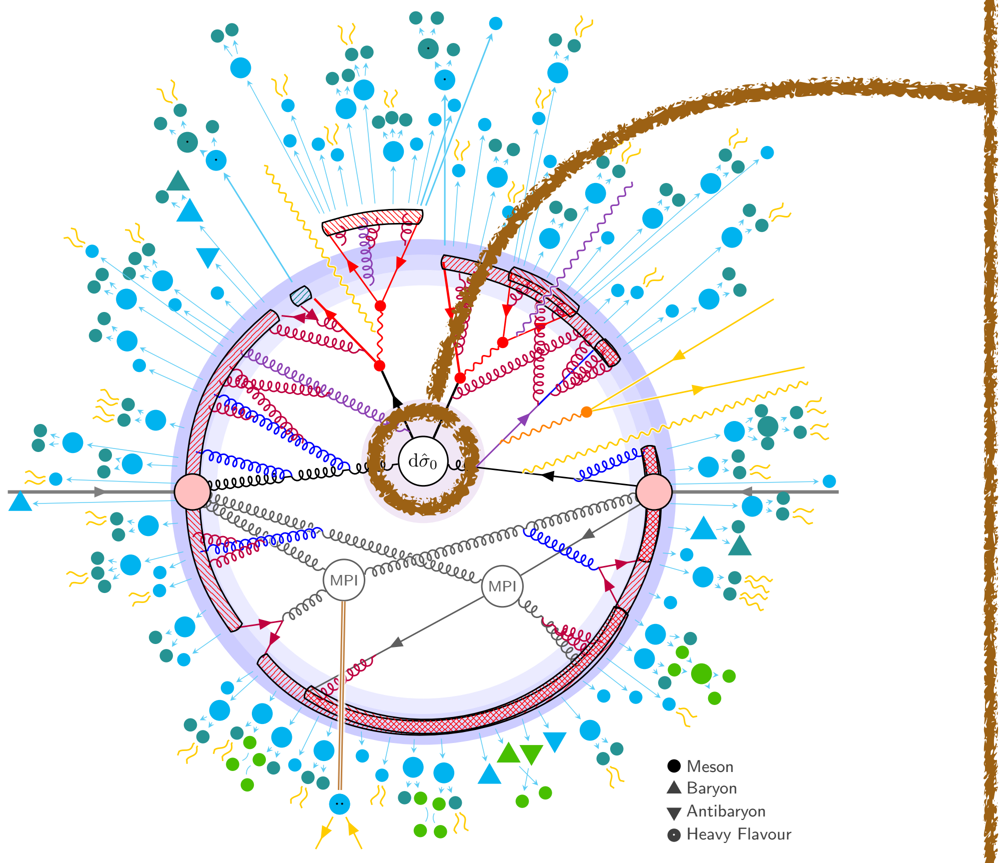
Lecture 1: the hard interaction



Lecture 2: Parton Showers & Hadronisation

Lecture 3: Matching & Merging

Lecture 1: the hard interaction



- Core of the event
- Process dependent
- First principles description
  - Largest energy transfers
  - New physics most-likely will appear here
  - Can use perturbation theory: LO, NLO, NNLO, etc.

Lecture 2: Parton Showers & Hadronisation

Lecture 3: Matching & Merging

Master equation for the hard interaction

$$\sum_{a,b} \int dx_1 dx_2 d\Phi_{\rm FS}\, f_a(x_1, \mu_F) f_b(x_2, \mu_F)\, \hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$$

Phase-space integral

Parton density functions

Parton-level cross section

1. Parton distribution functions
   - Universal/process independent
   - Extracted from experiment
   - Evolution from theory
     - Can be extract in one process, and applied to others

2. Parton-level cross section
   - Short distance coefficients as an expansion in α_S
   - From theory

Perturbative expansion

$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$ Parton-level cross section

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

Perturbative expansion

$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$ Parton-level cross section

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

LO predictions

Perturbative expansion

$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$ Parton-level cross section

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

LO predictions

NLO corrections

Perturbative expansion

$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$ Parton-level cross section

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

LO predictions

NLO corrections

NNLO corrections

Perturbative expansion

$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$ Parton-level cross section

$$\hat{\sigma} = \sigma^{\text{Born}}\left(1 + \frac{\alpha_s}{2\pi}\sigma^{(1)} + \left(\frac{\alpha_s}{2\pi}\right)^2\sigma^{(2)} + \left(\frac{\alpha_s}{2\pi}\right)^3\sigma^{(3)} + \dots\right)$$

LO predictions

NLO corrections

NNLO corrections

NNNLO corrections

Perturbative expansion

$$\hat{\sigma}_{ab\to X}(\hat{s},\mu_F,\mu_R) \quad \text{Parton-level cross section}$$

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

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

LO predictions | NLO corrections | NNLO corrections | NNNLO corrections

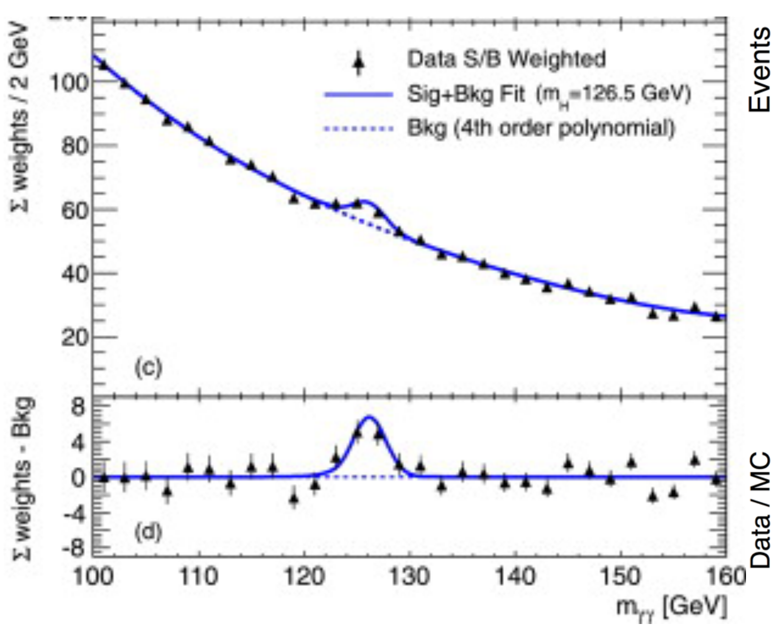
Higher-order computations

- Why?
  - *They improve the accuracy of our predictions*

Discoveries at hadron colliders

Peak

$H \to \gamma\gamma$

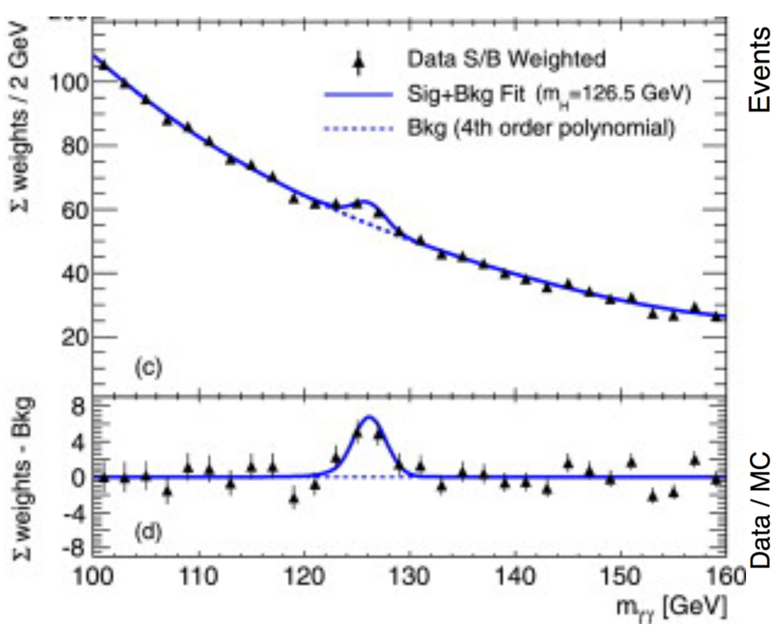


EASY

Background directly measured from **data**.
Theory needed only for parameter extraction

Discoveries at hadron colliders

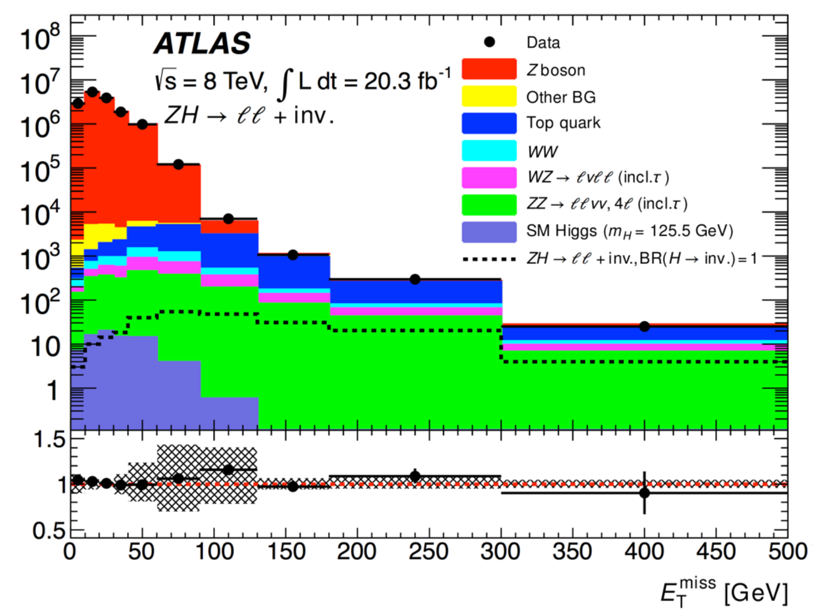
Peak
$H \rightarrow \gamma\gamma$



EASY

Background directly measured from **data**.
Theory needed only for parameter extraction

Shape
$Z\,H \rightarrow l^+l^- + inv.$

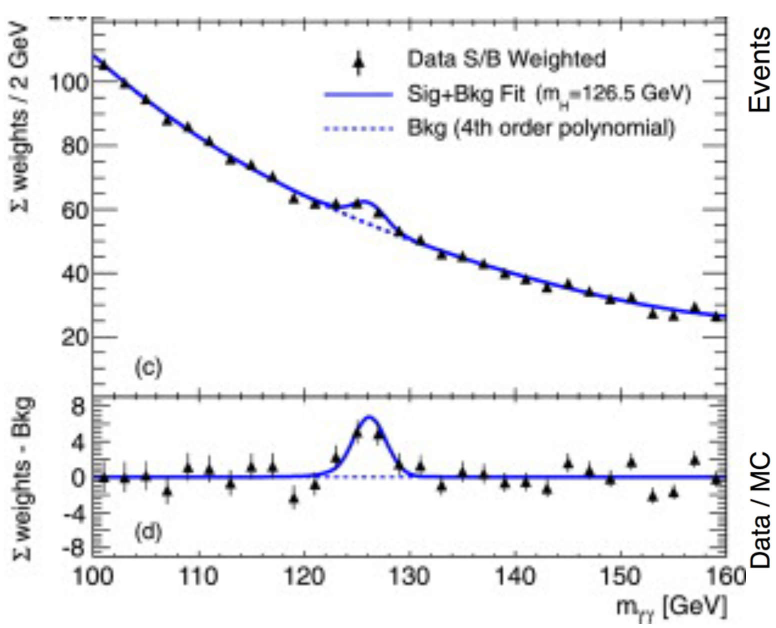


HARD

Background **SHAPE** needed.
Flexible MC for both signal and background validated and tuned to data

Discoveries at hadron colliders

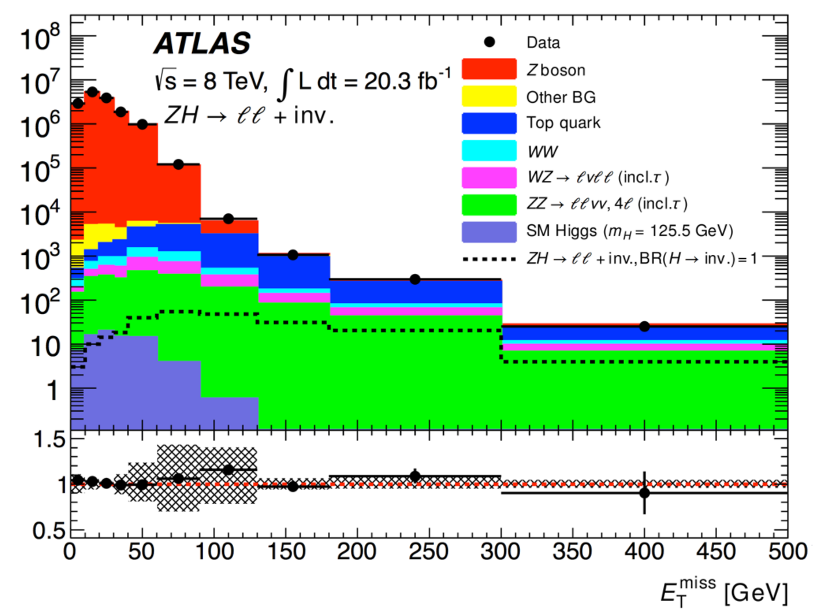
Peak
$H \rightarrow \gamma\gamma$



EASY

Background directly measured from **data**.
Theory needed only for parameter extraction

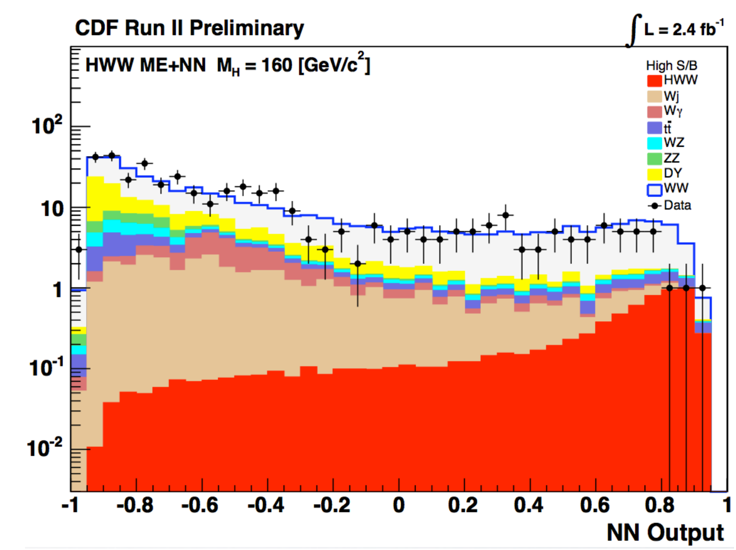
Shape
$Z\,H \rightarrow l^+ l^- + inv.$



HARD

Background **SHAPE** needed.
Flexible MC for both signal and background validated and tuned to data

Rate
$H \rightarrow W^+ W^-$

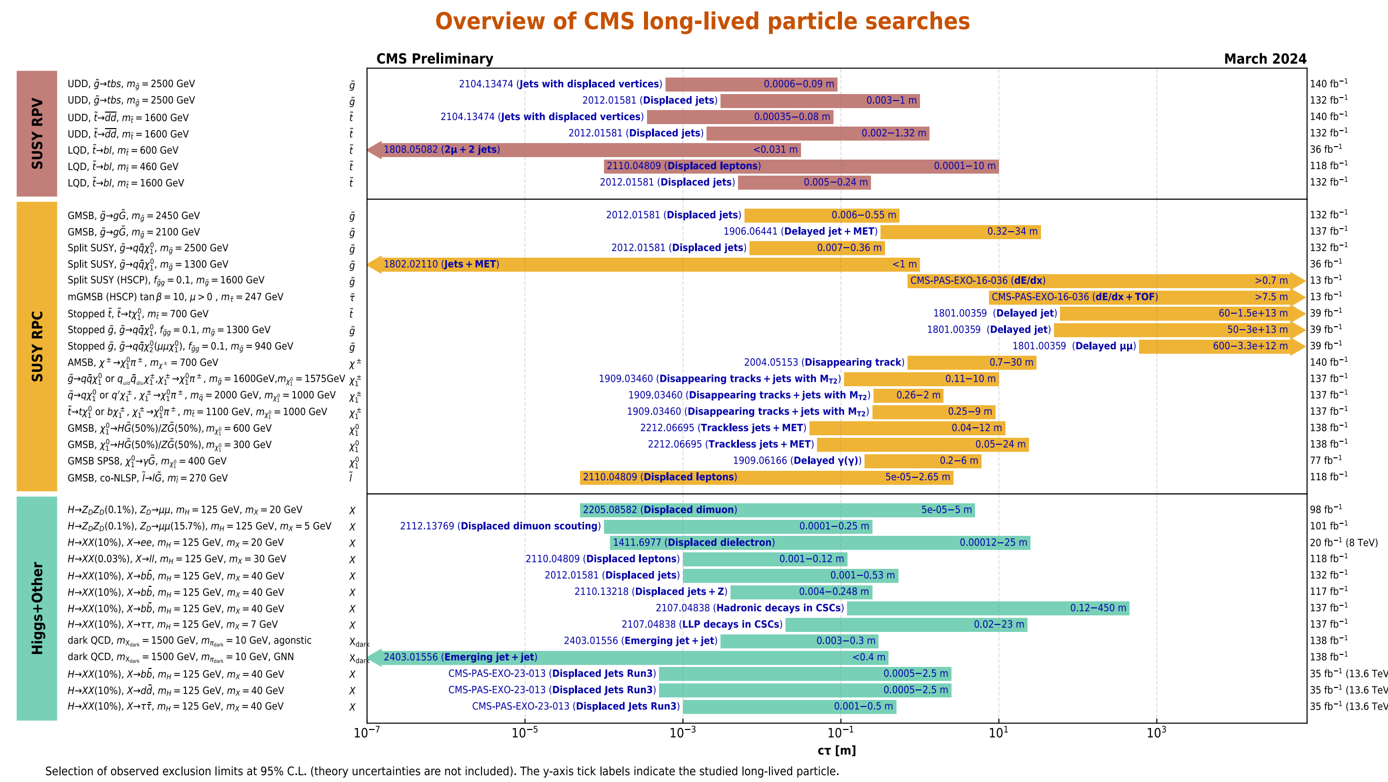


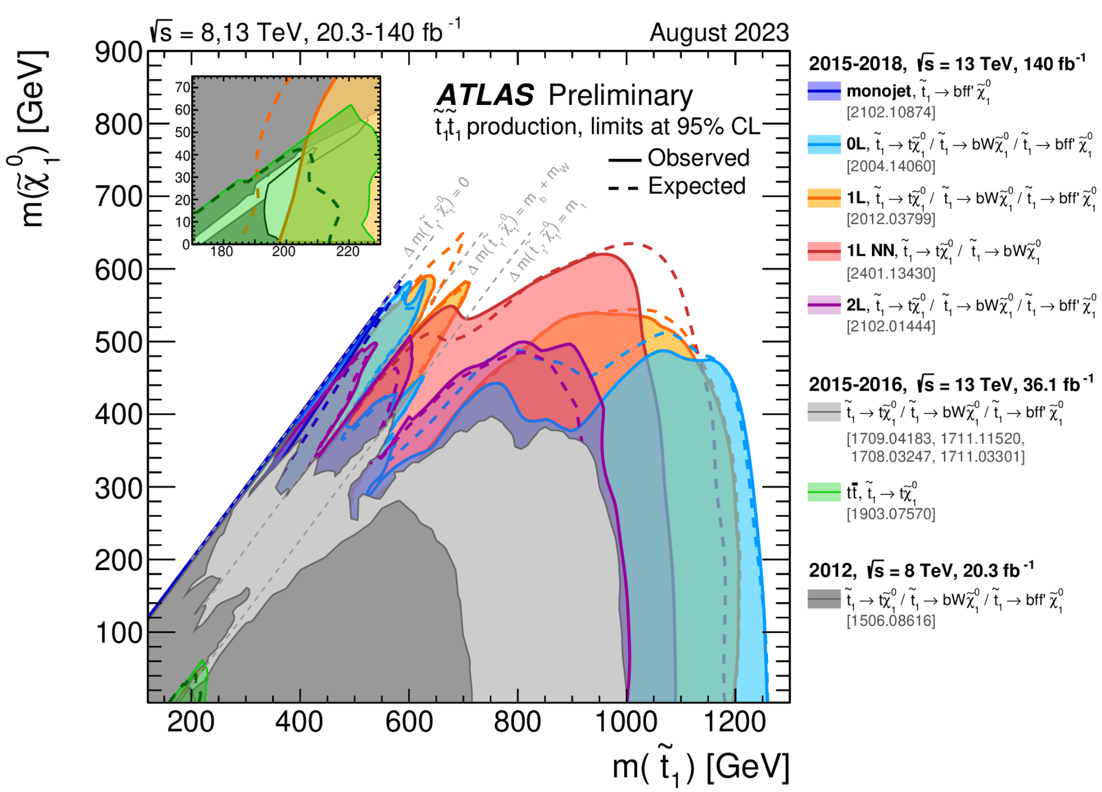
VERY HARD

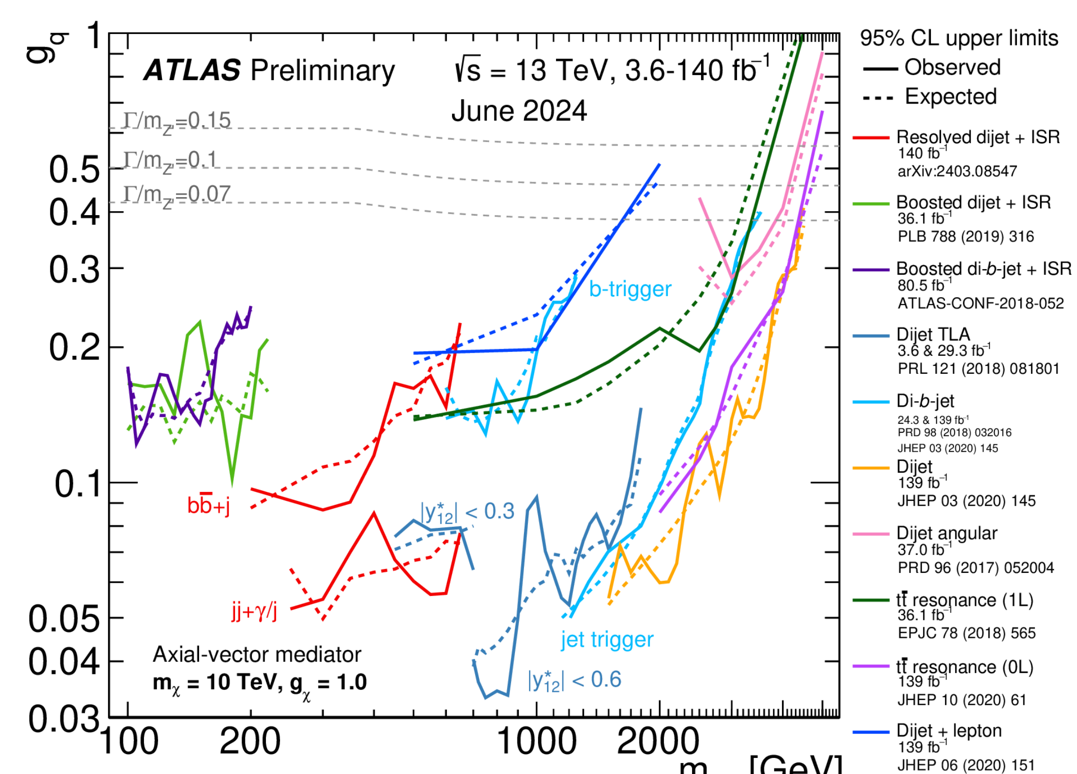
Relies on prediction for both **shape** and **normalization**.
Complicated interplay of best simulations and data

New Physics?

- No NP has been discovered yet
  - Either there is no NP, or it is hiding very well
  - If it is there, it will be a ‘Hard’ or ‘very Hard’ discovery
  - Need for accurate predictions for signal and background

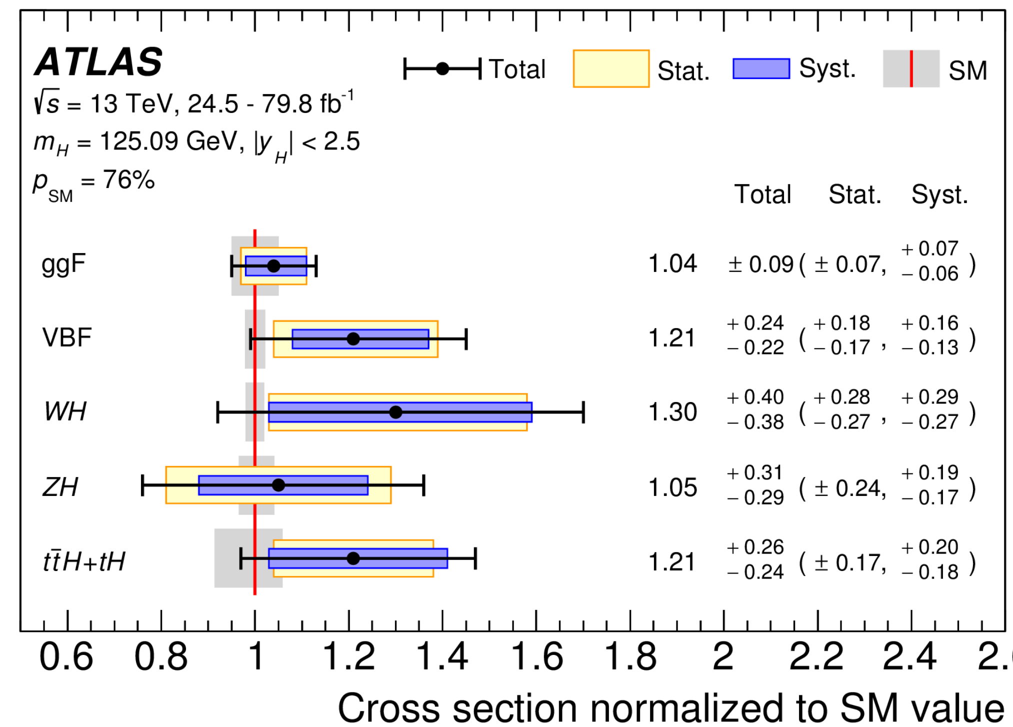


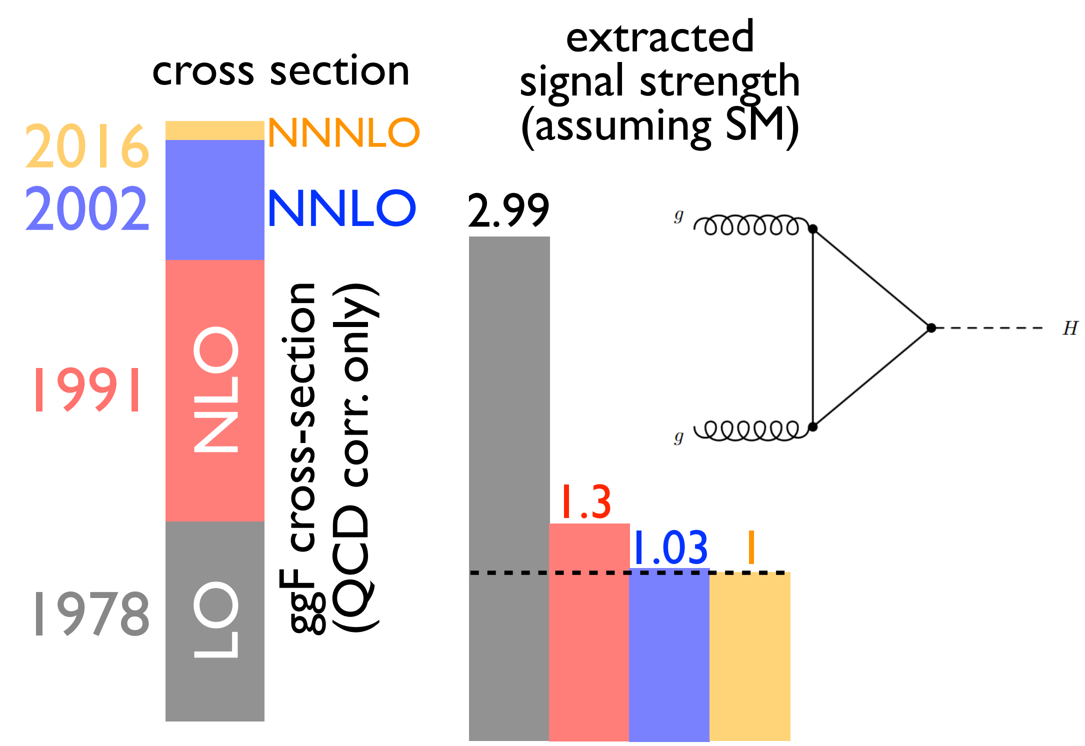




Standard Model measurements

- The measurement of the Higgs couplings is an emblematic example of the need for precision
  - Large perturbative corrections for the dominant channel (gluon fusion)
  - Without higher-order corrections, measured signal strength ~ 3 × SM
  - Very competitive experimental measurements!





Anastasiou et al, arXiv:1602.00695

Perturbative expansion

$$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R) \quad \text{Parton-level cross section}$$

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

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

LO predictions | NLO corrections | NNLO corrections | NNNLO corrections

- Including higher corrections improves predictions and reduces theoretical uncertainties

Perturbative expansion

$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$ Parton-level cross section

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

$$\hat{\sigma} = \sigma^{\text{Born}}\left(1 + \frac{\alpha_s}{2\pi}\sigma^{(1)} + \left(\frac{\alpha_s}{2\pi}\right)^2\sigma^{(2)} + \left(\frac{\alpha_s}{2\pi}\right)^3\sigma^{(3)} + \dots\right)$$

LO predictions | NLO corrections | NNLO corrections | NNNLO corrections

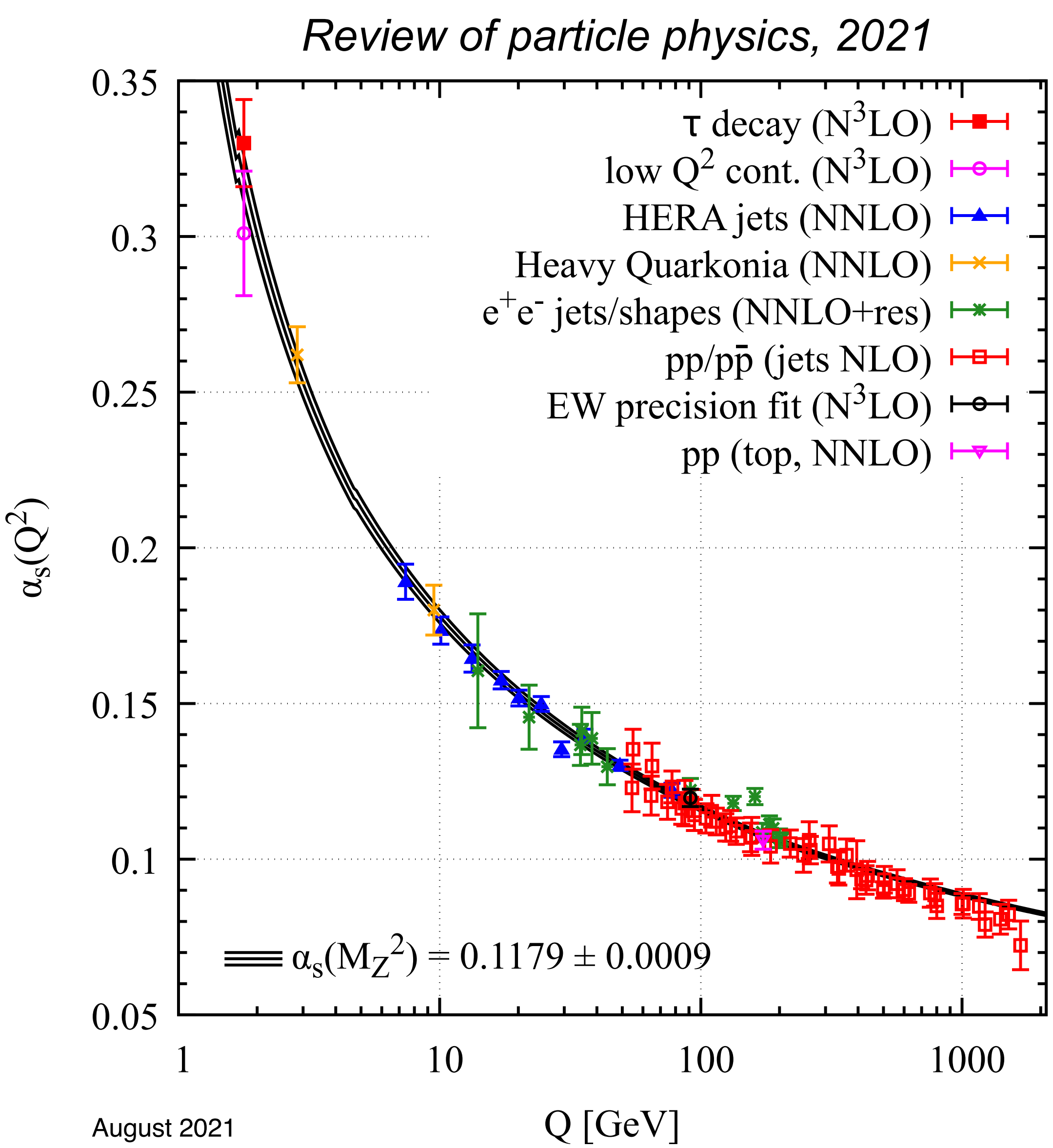
- Including higher corrections improves predictions and reduces theoretical uncertainties

Renormalisation scale

$$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$$ Parton-level cross section

- Introduced in perturbation theory and "is an unphysical parameter"
- Consequences:
  - The coupling constant depends on this parameter
  - Beyond LO, the matrix elements depend on this parameter
  - Only if on includes all orders in perturbation theory, this parameter drops out
- Which value should you give it?

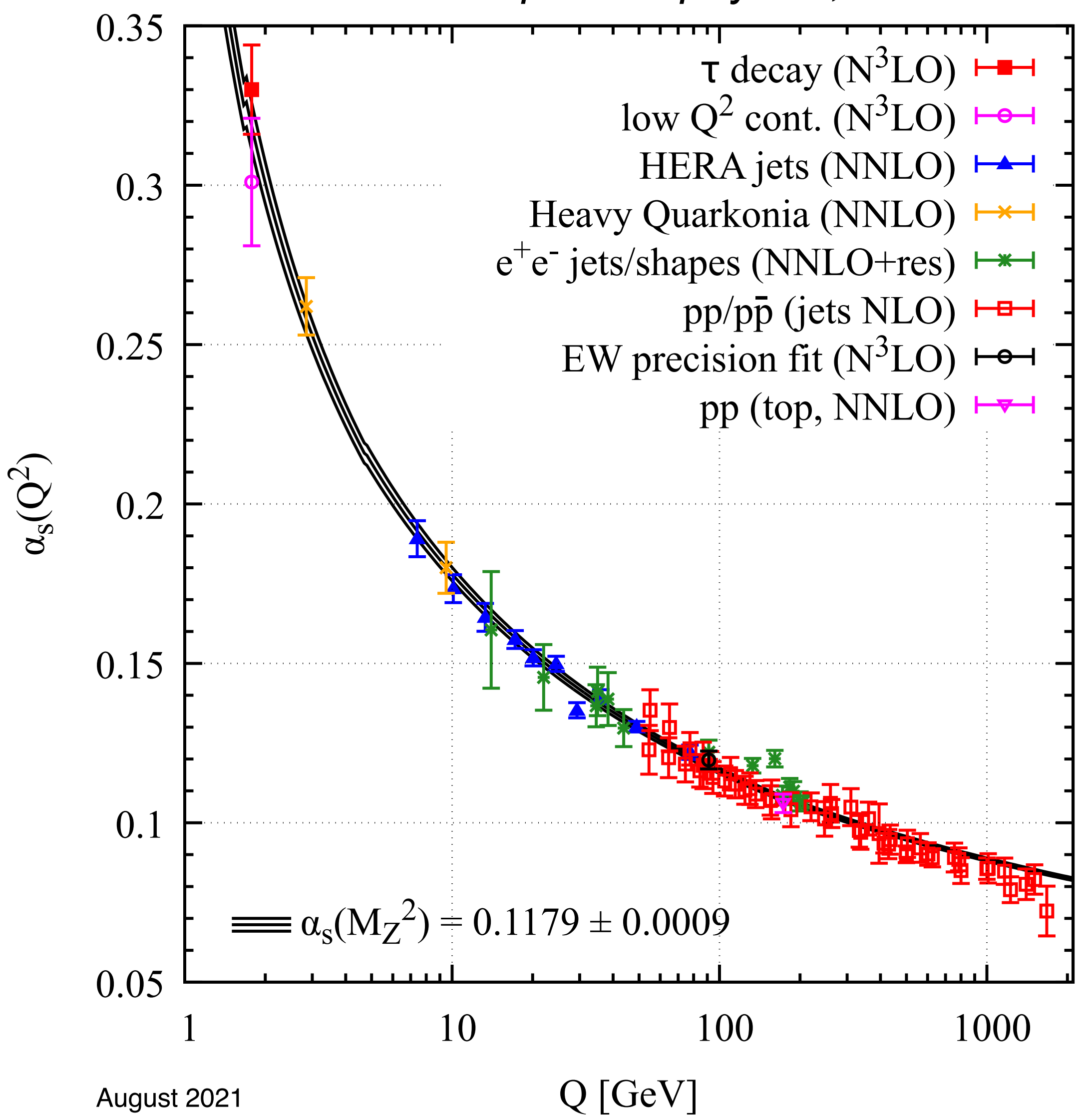
Running of the strong coupling



- The value of the strong coupling depends (logarithmically) on the renormalisation scale
- The larger the scale, the smaller the coupling
- Naively: choose large renormalisation scales, and perturbation theory will work well…?

Running of the strong coupling

Review of particle physics, 2021



- The value of the strong coupling depends (logarithmically) on the renormalisation scale
- The larger the scale, the smaller the coupling
- Naively: choose large renormalisation scales, and perturbation theory will work well…?

Not so simple

Renormalisation scale II

- Which value should you give it?

$$\hat{\sigma} = \sigma^{\mathrm{Born}} \left(1 + \sum_{k=1}^{\infty} \left(\frac{\alpha_s}{2\pi} \right)^k \sigma^{(k)} \right)$$

- For the theory to converge, both α_s and $\sigma^{(k)}$ should be small, and both depend on the renormalisation scale
- Just like α_s, also $\sigma^{(k)}$ depends logarithmically on the renormalisation scale: it contains (powers of) $\log(\mu_R^2/Q^2)$, with Q^2 any (relevant) invariant, such as particle masses, 2-body invariant masses, $\sqrt{\hat{s}}$, etc.
- For the best convergence, the renormalisation scale should be chosen such that it matches the typical Q^2 relevant to the process and or observable

What about the factorisation scale?

$\hat{\sigma}_{ab \to X}(\hat{s}, \mu_F, \mu_R)$ Parton-level cross section

What about the factorisation scale?

$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$ Parton-level cross section

What about the factorisation scale?

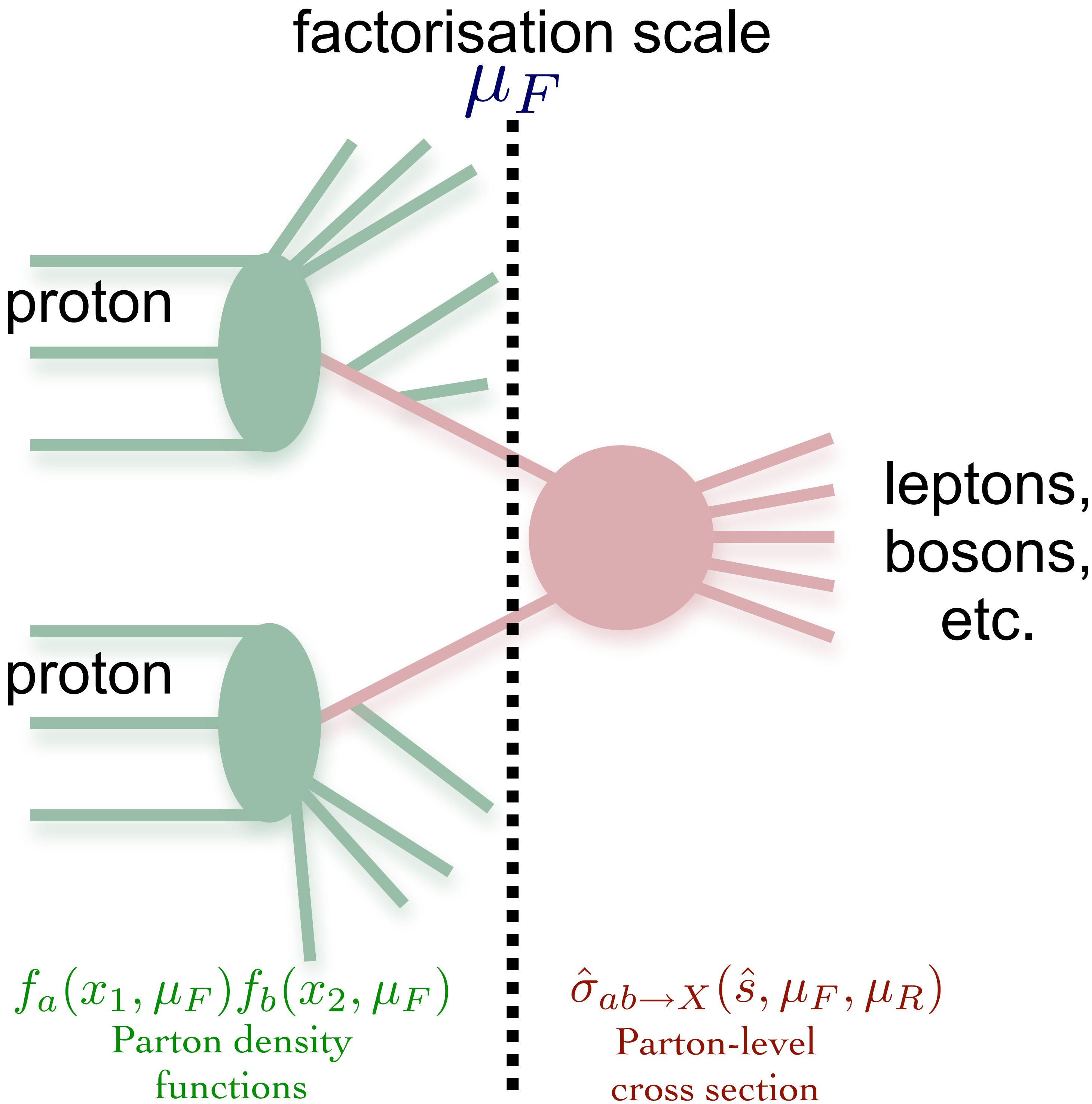
$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$ Parton-level cross section

- Has a clearer physical interpretation:
  - Separation scale between physics included in parton density functions and hard matrix elements
  - Just like renormalisation scale, should take a numerical value close to the relevant scale to the process
  - Including higher-orders reduces the dependence on the factorisation scale

$$\sum_{a,b} \int dx_1 dx_2 d\Phi_{\mathrm{FS}}\, f_a(x_1, \mu_F) f_b(x_2, \mu_F)\, \hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$$

Phase-space integral — Parton density functions — Parton-level cross section

Inclusiveness

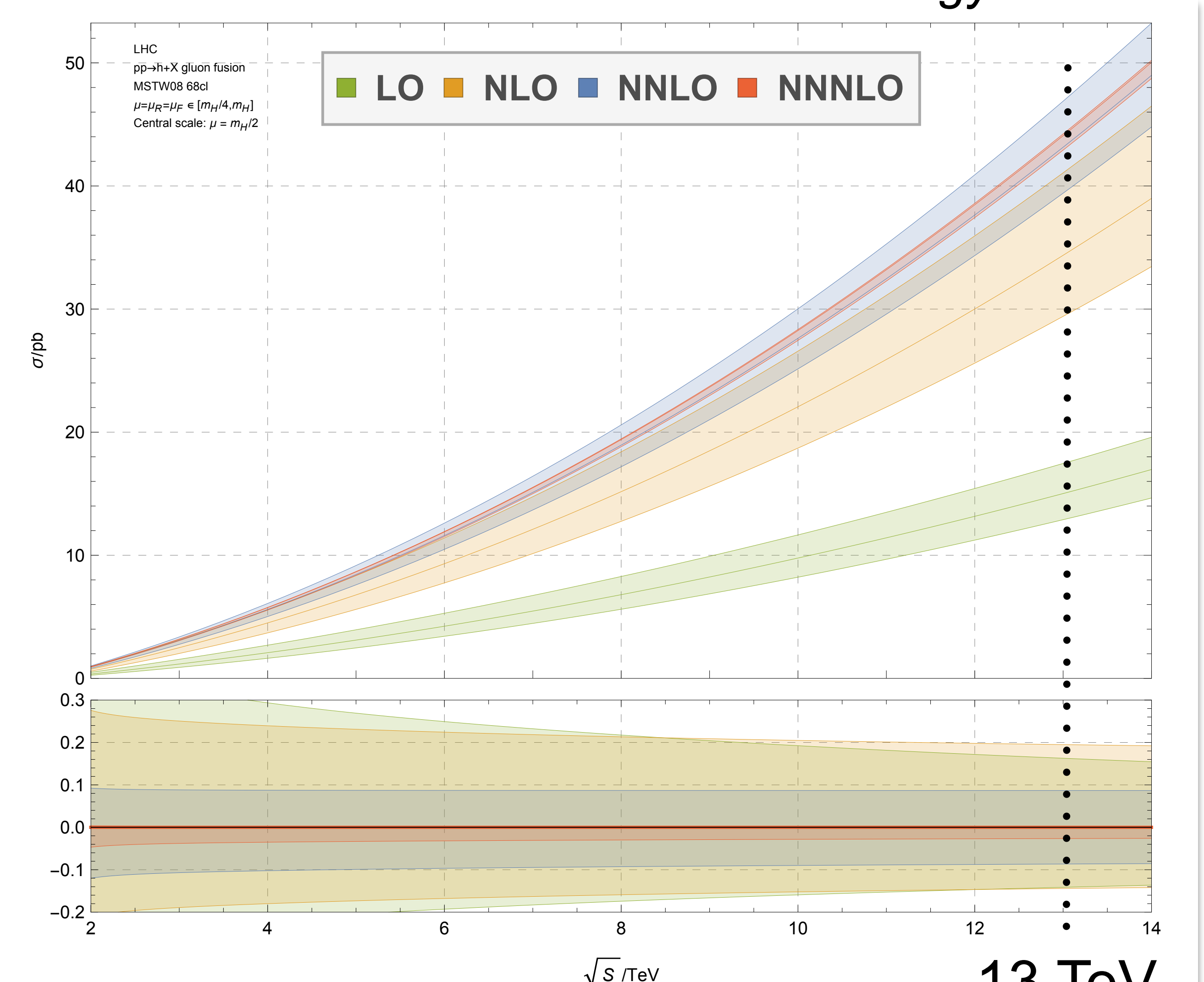


- All additional radiation, softer than the factorisation scale, is included in the computed cross section through the evolution of the parton density functions
  - Only exact in the collinear limit
- At N^{k}LO accuracy, up to k of these emissions are included exactly also outside the collinear limit
  - This reduces the dependence on the factorisation scale

Perturbation theory at work

- The inclusion of higher orders improves the reliability of a given computation
  - More reliable description of total rates and shapes
  - Residual uncertainties related to the arbitrary scales in the process decrease
  - The computational complexity grows exponentially
  - NLO is mandatory for LHC physics!

Cross section for Higgs production in gluon fusion as a function of the collider energy



Anastasiou, et al, arXiv:1503.06056

Master equation for the hard interaction

$$\sum_{a,b} \int dx_1 dx_2 d\Phi_{\mathrm{FS}}\, f_a(x_1, \mu_F) f_b(x_2, \mu_F)\, \hat{\sigma}_{ab \to X}(\hat{s}, \mu_F, \mu_R)$$

Phase-space integral

Parton density functions

Parton-level cross section

1. Parton distribution functions
   - Universal/process independent
   - Extracted from experiment
   - Evolution from theory
     - Can be extract in one process, and applied to others
2. Parton-level cross section
   - Short distance coefficients as an expansion in α_S
   - From theory

In practice: predictions at LO

How to calculate e.g. 3-jet production at the LHC?

- Identify all subprocesses (gg→ggg, qg→qgg....) in:

$$\sigma(pp \to 3j) = \sum_{ijk} \int f_i(x_1) f_j(x_2) \hat{\sigma}(ij \to k_1 k_2 k_3)$$

easy

- For each one, calculate the amplitude

$$\mathcal{A}(\{p\},\{h\},\{c\}) = \sum_i D_i$$

difficult

- Square the amplitude, sum over spin & colour, and integrate over the phase-space

$$\hat{\sigma} = \frac{1}{2\hat{s}} \int d\Phi_p \sum_{h,c} |\mathcal{A}|^2$$

quite hard

In practice: predictions at LO

How to calculate e.g. 3-jet production at the LHC?

- Identify all subprocesses (gg→ggg, qg→qgg....) in:

$$\sigma(pp \to 3j) = \sum_{ijk} \int f_i(x_1) f_j(x_2) \hat{\sigma}(ij \to k_1 k_2 k_3)$$

easy

- For each one, calculate the amplitude

$$\mathcal{A}(\{p\},\{h\},\{c\}) = \sum_i D_i$$

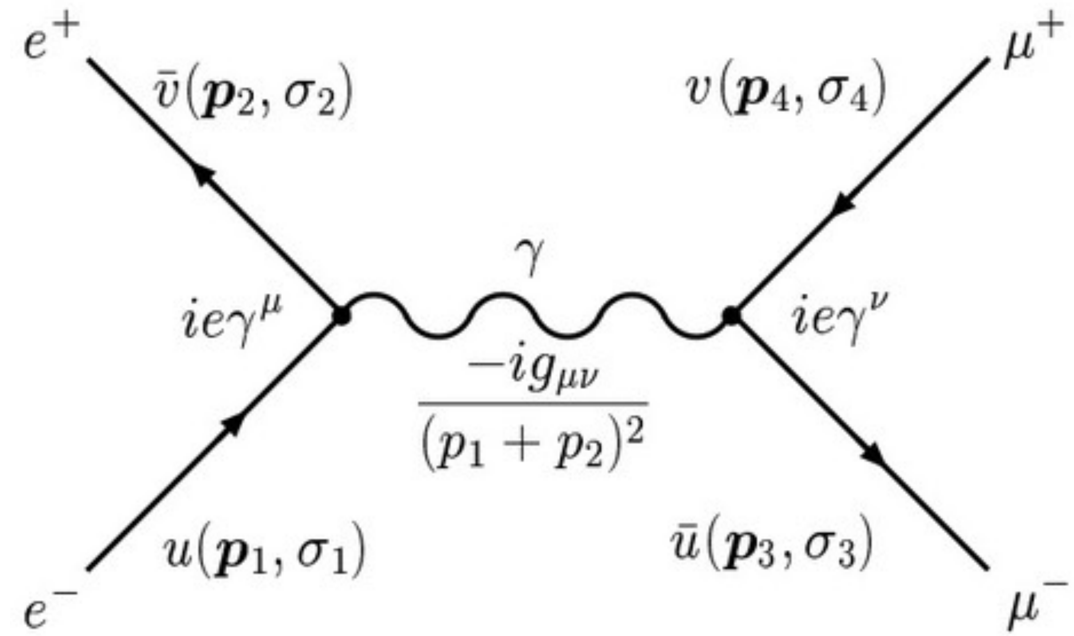
difficult

- Square the amplitude, sum over spin & colour, and integrate over the phase-space

$$\hat{\sigma} = \frac{1}{2\hat{s}} \int d\Phi_p \sum_{h,c} |\mathcal{A}|^2$$

quite hard

Feynman Rules



- Based on Feynman Rules: universal building blocks to create Feynman diagrams
- Feynman diagrams correspond to mathematical expressions
  - Tedious to do by hand, but no problem for a computer
- Using helicity amplitudes with explicit representations for the spinors/polarisation vectors can reduce the complexity in the numerical evaluation of the expressions
- Recycling identical sub-structures in multiple diagrams and/or using recursion relations, can further reduce the computation time

In practice: predictions at LO

How to calculate e.g. 3-jet production at the LHC?

- Identify all subprocesses (gg→ggg, qg→qgg....) in:

$$\sigma(pp \to 3j) = \sum_{ijk} \int f_i(x_1) f_j(x_2) \hat{\sigma}(ij \to k_1 k_2 k_3)$$

easy

- For each one, calculate the amplitude

$$\mathcal{A}(\{p\}, \{h\}, \{c\}) = \sum_i D_i$$

difficult

- Square the amplitude, sum over spin & colour, and integrate over the phase-space

$$\hat{\sigma} = \frac{1}{2\hat{s}} \int d\Phi_p \sum_{h,c} |\mathcal{A}|^2$$

quite hard

LUNDS UNIVERSITET

Phase-space integral

$$\mathrm{Dim}[\Phi(n)] = 3n - 2$$

$$\sigma = \frac{1}{2s} \int |\mathcal{M}|^2 d\Phi(n)$$

- Calculations of cross section or decay widths involve integrations over the phase space of rather peaked, multi-variate functions

Phase-space integral

$$\text{Dim}[\Phi(n)] = 3n - 2$$

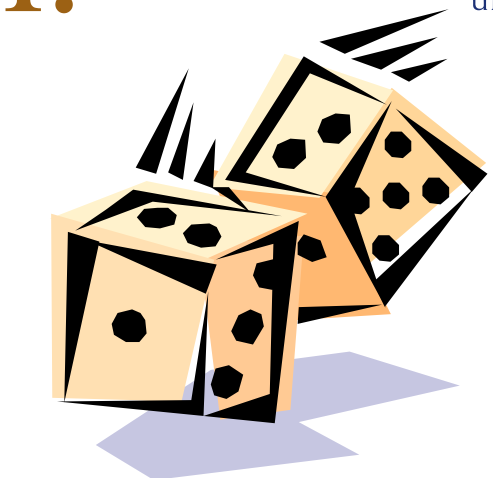
$$\sigma = \frac{1}{2s} \int |\mathcal{M}|^2 d\Phi(n)$$

- Calculations of cross section or decay widths involve integrations over the phase space of rather peaked, multi-variate functions

General and flexible method is needed:

Numerical (Monte Carlo) integration

Monte-Carlo integration: Integrals as averages



- Integral as a sum:

$$I = \int_a^b f(x)\,\mathrm{d}x \quad \Rightarrow \quad I_N = (b-a)\frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

$$V = (b-a)\int_a^b \left[f(x)\right]^2 \mathrm{d}x - I^2 \quad \Rightarrow \quad V_N = (b-a)^2\frac{1}{N}\sum_{i=1}^{N}\left[f(x_i)\right]^2 - I_N^2$$

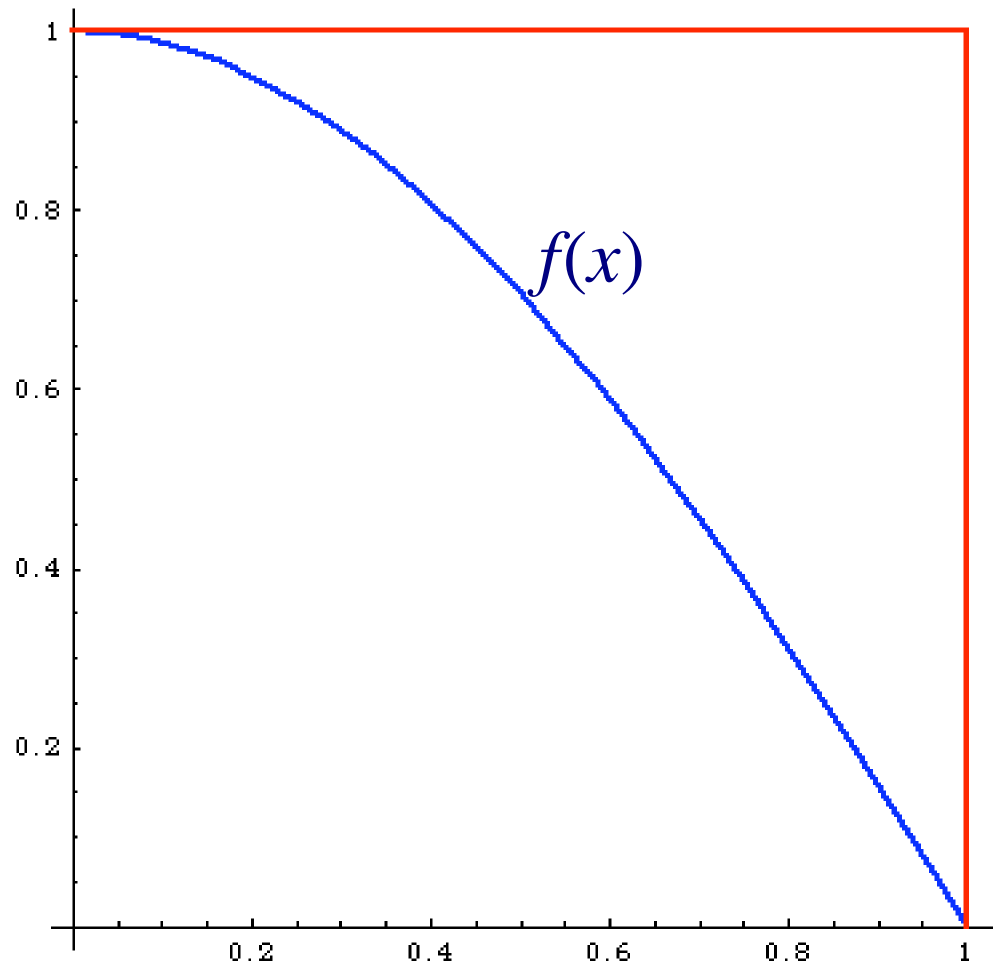
$$I = I_N \pm \sqrt{V_N/N}$$

- Convergence is slow but it can be estimated easily
- Scaling of the error does not depend on # of dimensions!
- Improvement by minimising V_N
- Optimal/Ideal case: $f(x) =$ constant $\Rightarrow$ $V_N = 0$

Event generation

- Every phase-space point computed in this way, can be seen as an event (=collision) in a detector
- However, they still carry the “weight” of the matrix elements:
  - events with large weights where the diff. cross section is large
  - events with small weights where the diff. cross section is small
- In nature, the events don’t carry a weight:
  - more events where the diff. cross section is large
  - less events where the diff. cross section is small
- How to go from weighted events to unweighted events?

Generation of unweighted events



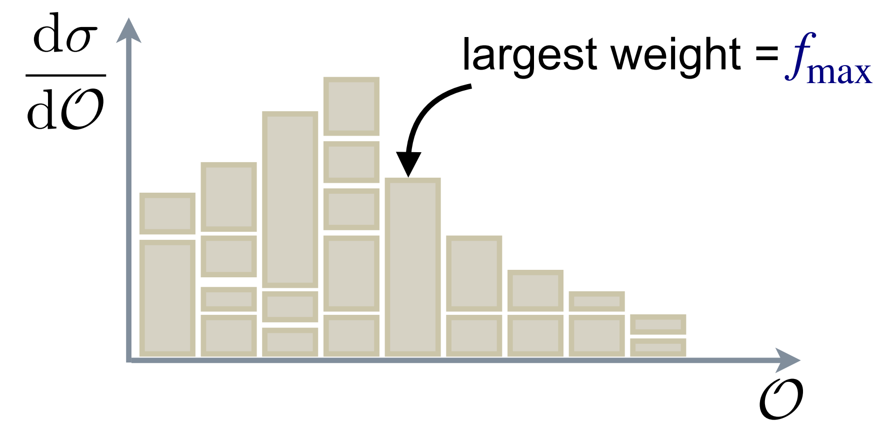
- Integral is area under a graph
  - Instead of picking a random x and compute $f(x)$,
  - pick a random x and y and check if $f(x) < y$. If so, keep event
  - Integral: 'total area' multiplied by fraction of events kept

- "Unweighted" events contain the maximum amount of statistical information in the least amount of events
  - Ideal if post-processing (slow detector simulation!) or storage is at a premium
- It requires knowledge on f_{max} to determine the 'total area'.

Unweighting in multiple dimensions

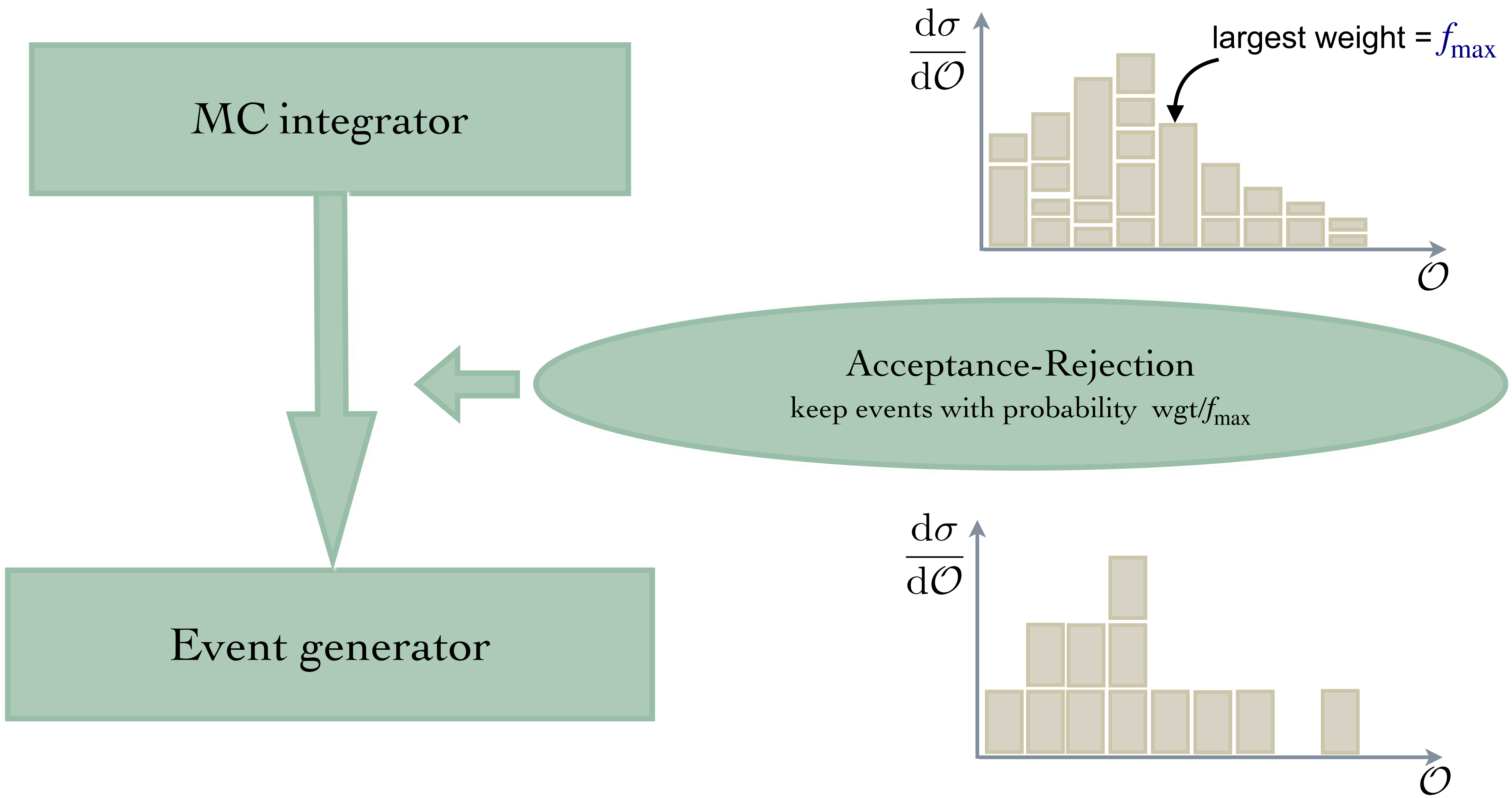
- Procedure works the same in multiple dimensions
- In practice, $f_{\max}$ is determined dynamically: event with largest weight encountered

MC integrator



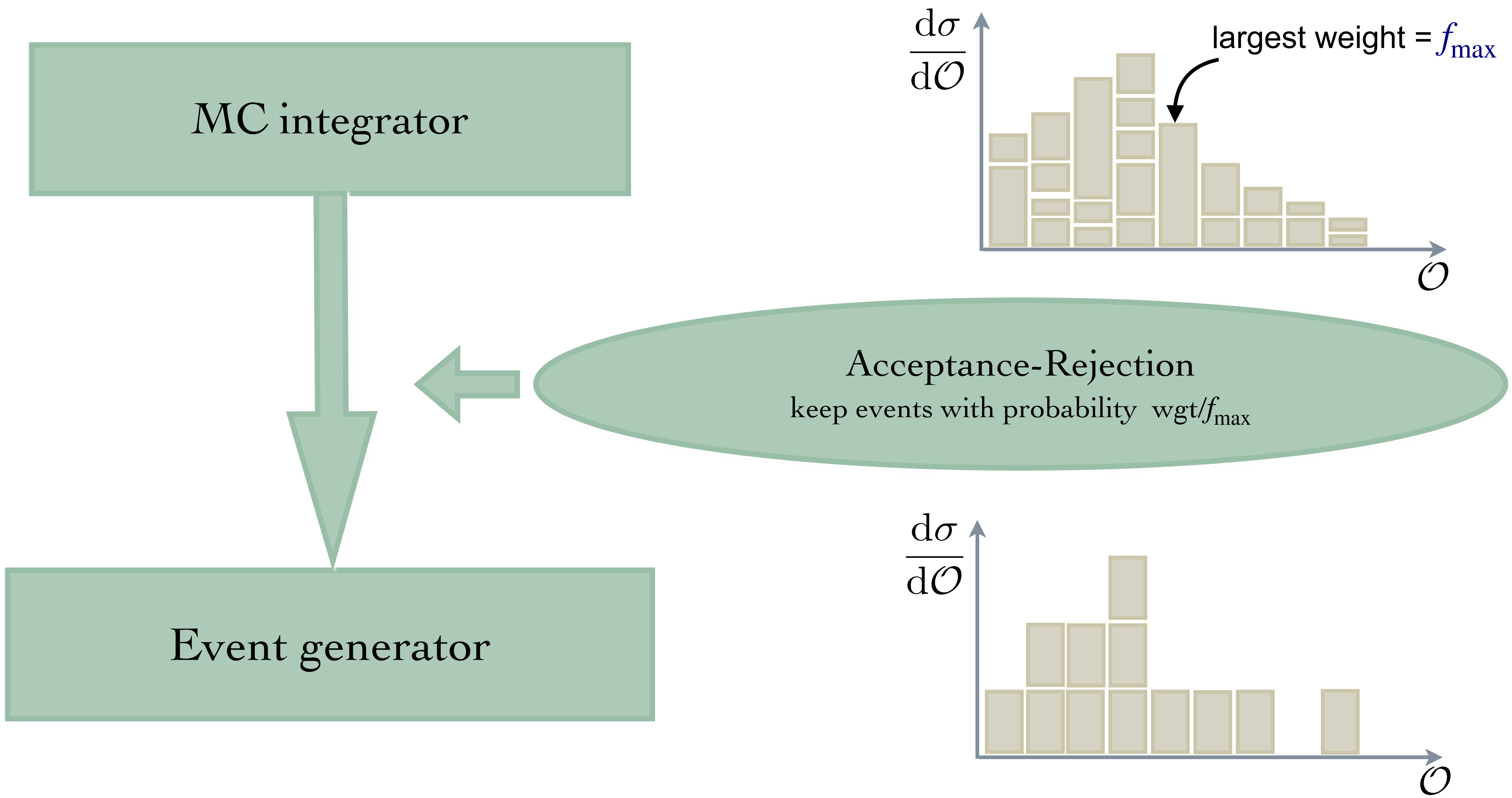
Unweighting in multiple dimensions

- Procedure works the same in multiple dimensions
- In practice, $f_{\max}$ is determined dynamically: event with largest weight encountered



Unweighting in multiple dimensions

- Procedure works the same in multiple dimensions
- In practice, $f_{\max}$ is determined dynamically: event with largest weight encountered



- This is possible only if $f(x)$ is bounded (and has definite sign)!

Curse of dimensionality

- Error in Monte-Carlo integration scales like $1/\sqrt{N}$, with N the number of sampling points, independently of the number of dimensions. However…
  - the variance among the points is (typically) much larger for high dimensions: more complicated integrands
  - *increasing the dimensions makes the available space much larger*
- This makes phase-space integration for multi-particle processes a very hard problem

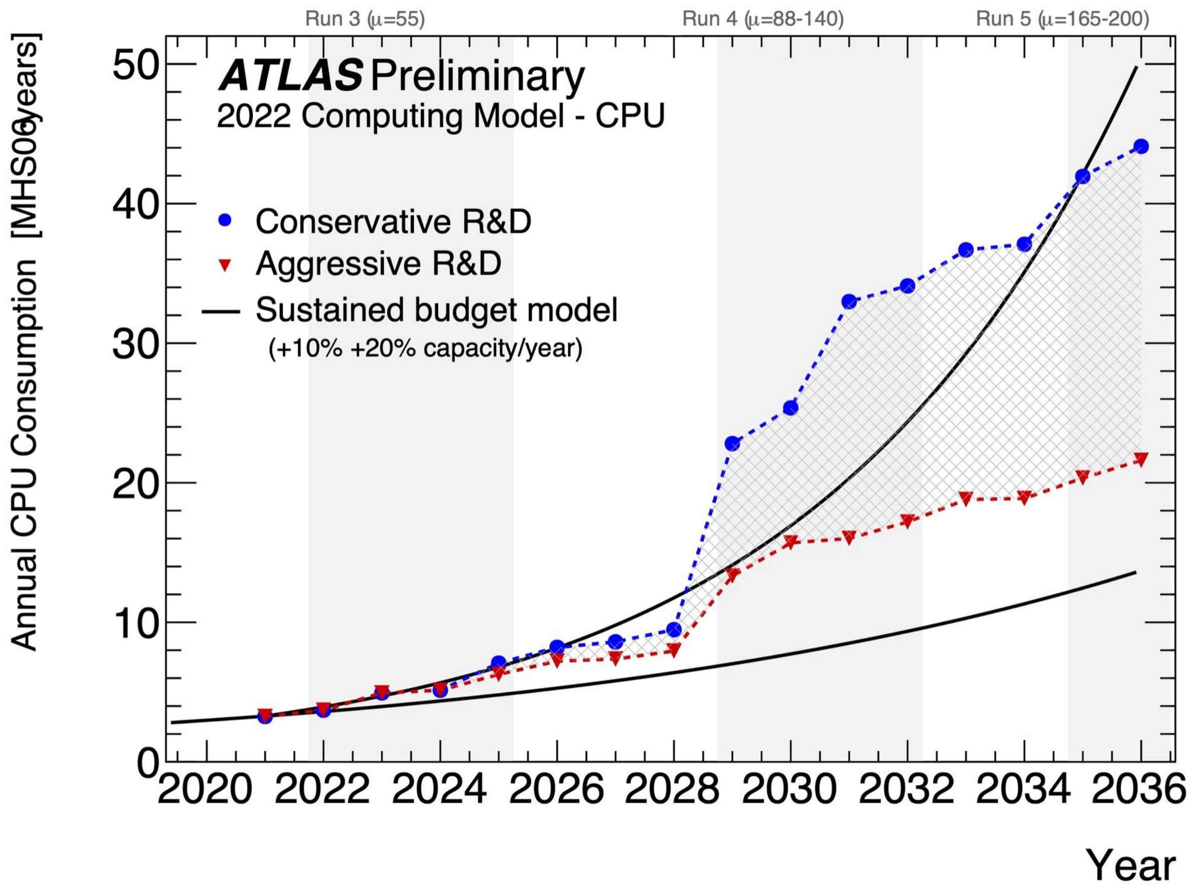
Number of Dimensions	Nearest Nghbr. Distance
1	5.0×10^{-5}
2	5.0×10^{-3}
3	2.6×10^{-2}
4	6.3×10^{-2}
5	0.11
7	0.23
10	0.39
25	1.1

(average) nearest-neighbour distance among 10000 randomly generated points in a unit hypercube

Optimisation

- Optimising phase-space integration and event unweighting can easily reduce the computation time by orders of magnitude
  - Typically much more than optimising the evaluation time of the matrix elements (at least for tree-level contributions)
- A very active area of research!
  - Some recent progress:
    - Optimised phase-space parametrisation
      [E. Bothmann *et al.* 2023]
    - Massively parallel setups
      [E. Bothmann *et al.* 2022, 2023]
    - Normalising flows and Machine Learning for efficient phase-space point generation
      [T. Heimel *et al.* 2022, 2023, 2024]
    - Reweighting low-accuracy events to high-accuracy
      [RF & T. Vitos, *in preparation*]

Expectation: computation needs HL-LHC



SM

subprocs handler

ME calculator

ME generators: general structure

Includes all possible subprocess leading to a given multi-jet final state automatically or manually (done once for all)

d~ d -> a d d~ u u~ g
d~ d -> a d d~ c c~ g
s~ s -> a d d~ u u~ g
s~ s -> a d d~ c c~ g
...

"Automatically" generates a code to calculate $|M|^2$ for arbitrary processes with many partons in the final state.

Use Feynman diagrams with tricks to reduce the factorial growth, others have recursive relations to reduce the complexity to exponential.

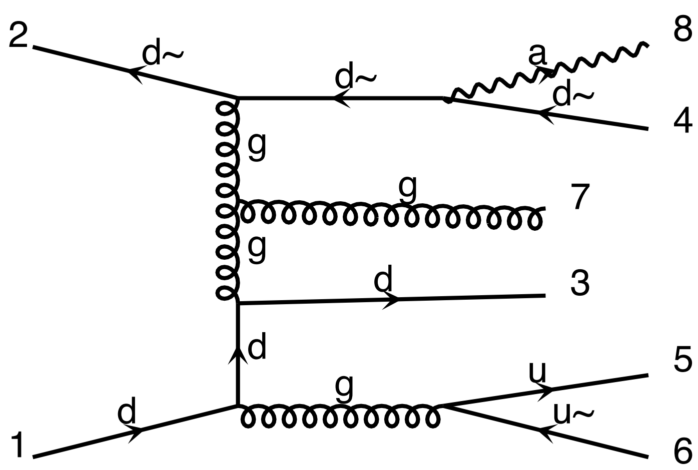
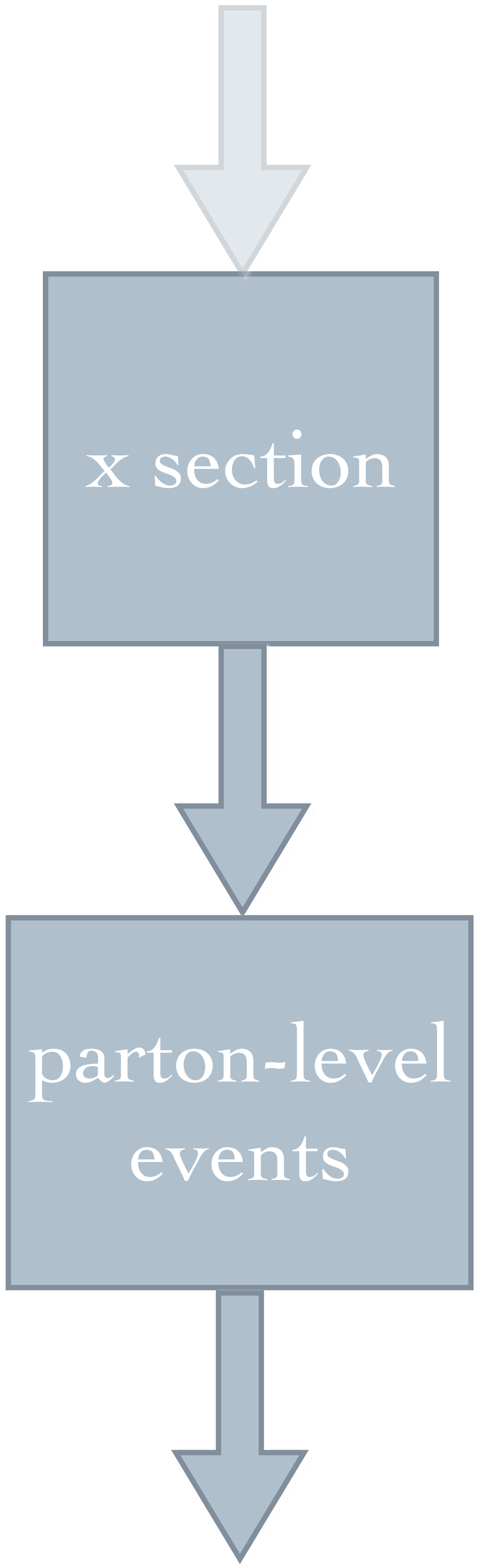
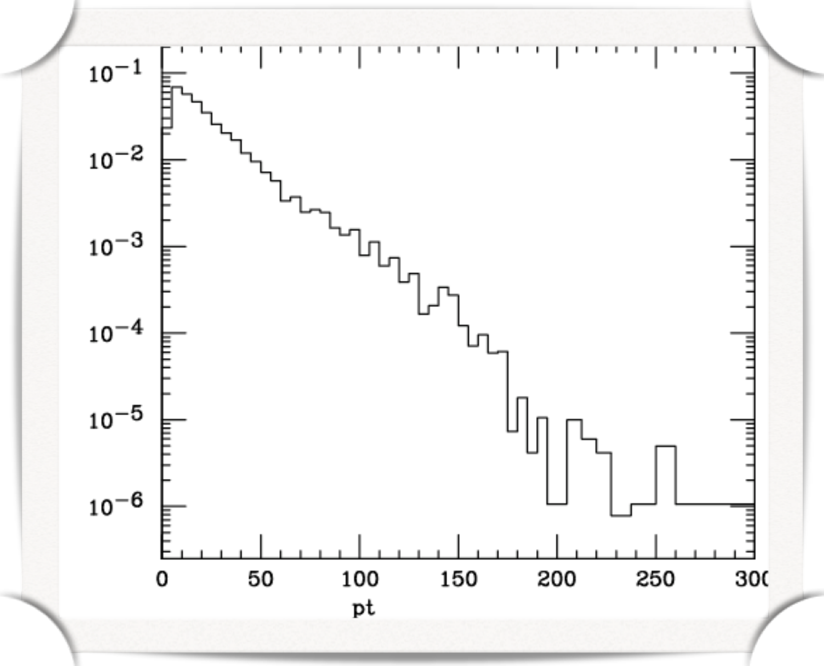


diagram 949 QCD=5, QED=1

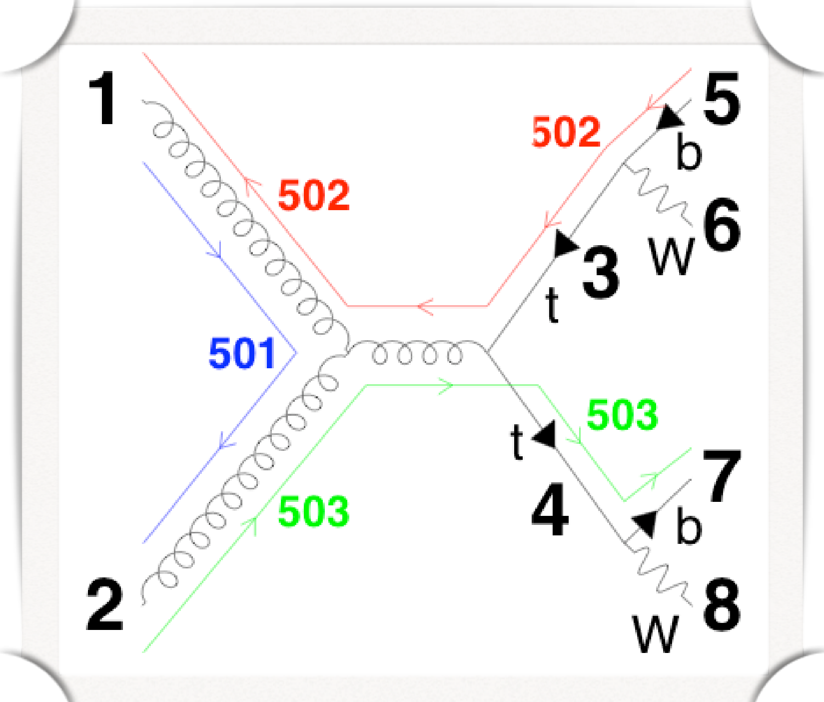
ME generators: general structure



Integrate the matrix element over the phase space using importance sampling and a multi-channel technique and using parton-level cuts.

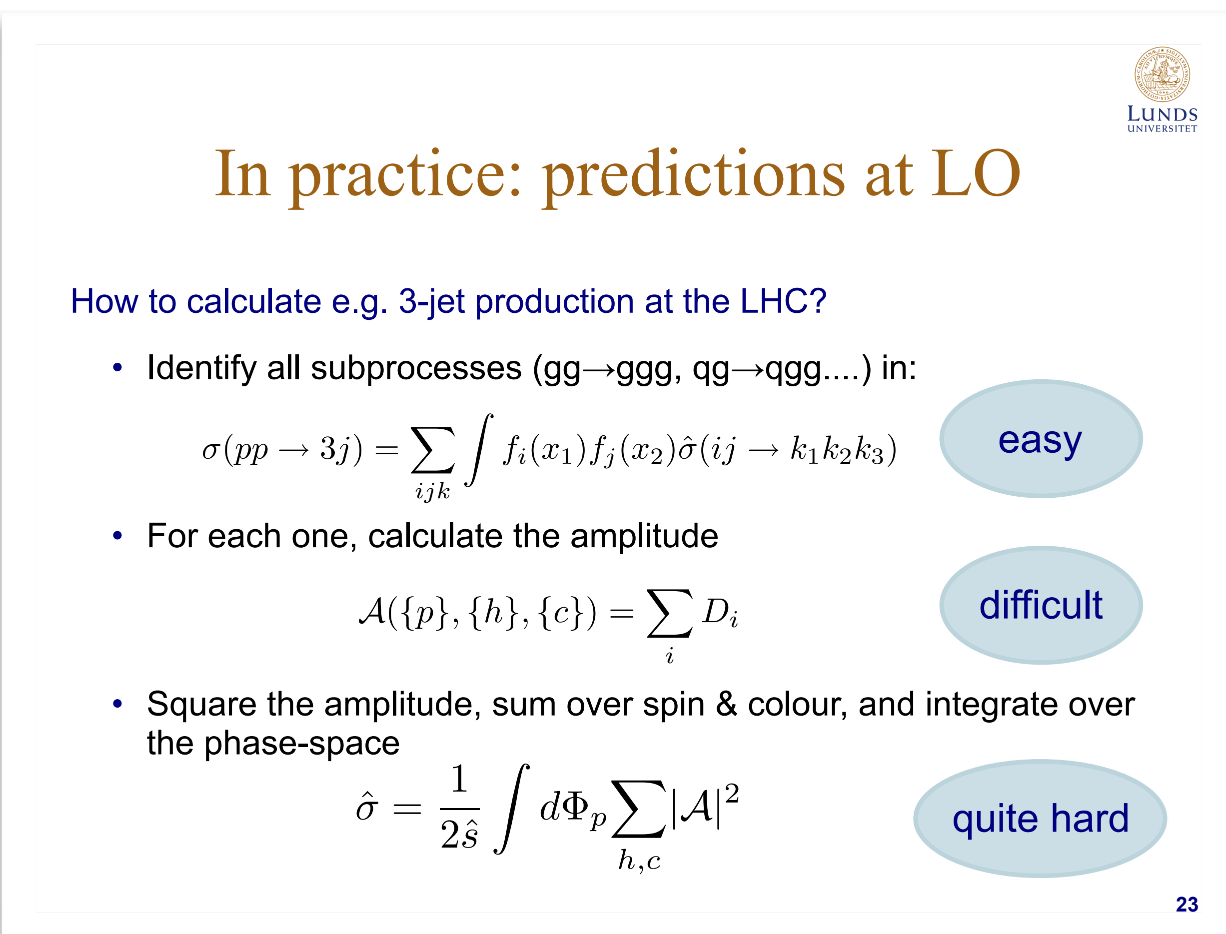


Events are obtained by unweighting.
These are at the parton-level.
Information on particle id, momenta, spin, color is given in the Les Houches Event (LHE) File format.



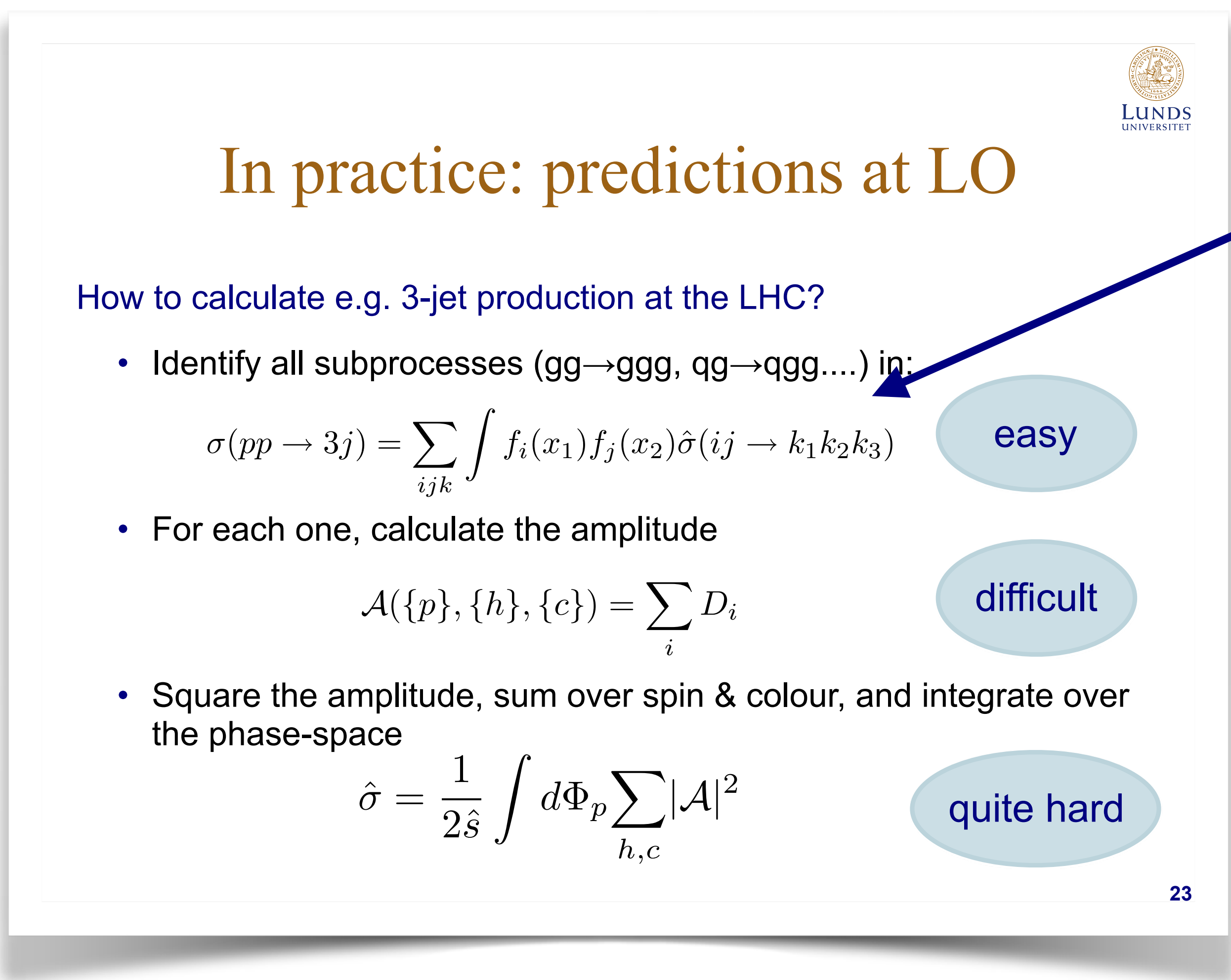
What about higher orders?

- All three steps change when including higher orders
- Let's focus on NLO.
  (NNLO and beyond imposes similar technical challenges, but orders of magnitude more complex)



What about higher orders?

- All three steps change when including higher orders
- Let's focus on NLO.
  (NNLO and beyond imposes similar technical challenges, but orders of magnitude more complex)

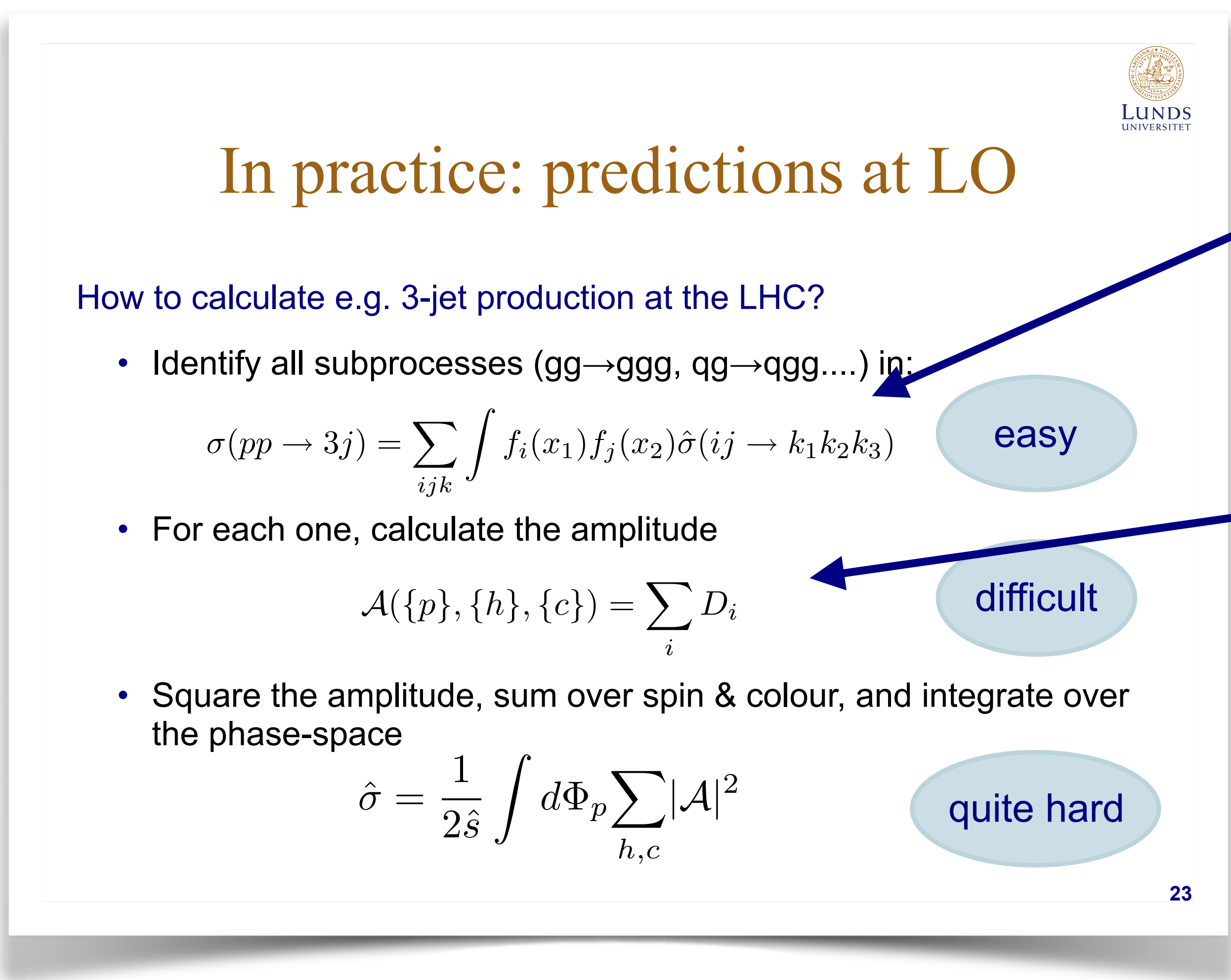


The same subprocesses contribute, and

- need also subprocesses with one more parton

What about higher orders?

- All three steps change when including higher orders
- Let's focus on NLO.
  (NNLO and beyond imposes similar technical challenges, but orders of magnitude more complex)



The same subprocesses contribute, and

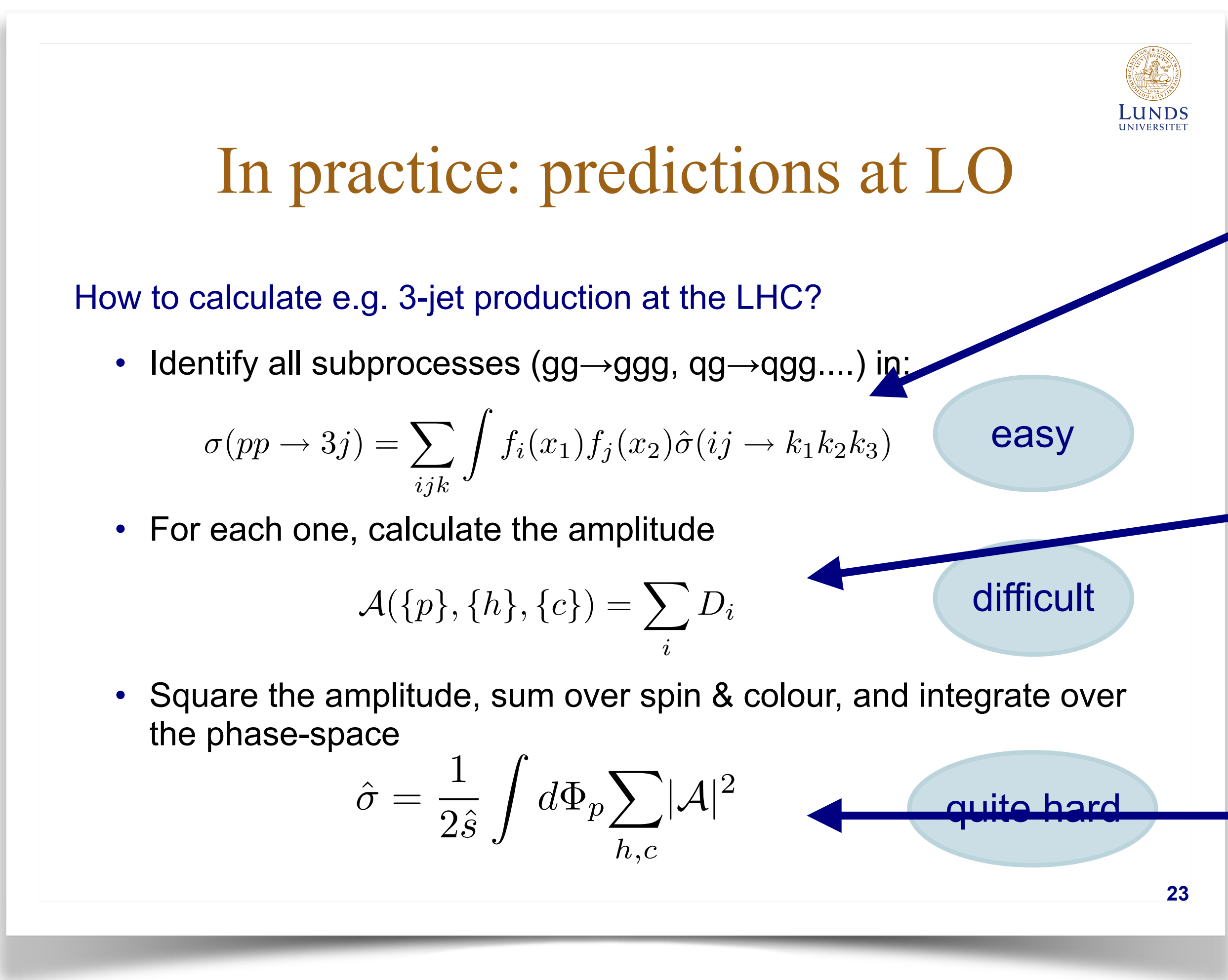
- need also subprocesses with one more parton

The same amplitudes need to be included, and

- need also generate amplitudes with particles going in a loop

What about higher orders?

- All three steps change when including higher orders
- Let's focus on NLO.
  (NNLO and beyond imposes similar technical challenges, but orders of magnitude more complex)



The same subprocesses contribute, and

- need also subprocesses with one more parton

The same amplitudes need to be included, and

- need also generate amplitudes with particles going in a loop

Still need to integrate over the phase-space,

- need also to cancel divergencies

NLO: how to?

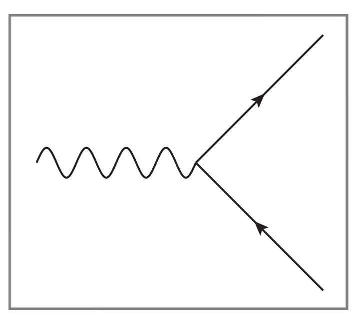
- Three ingredients need to be computed at NLO

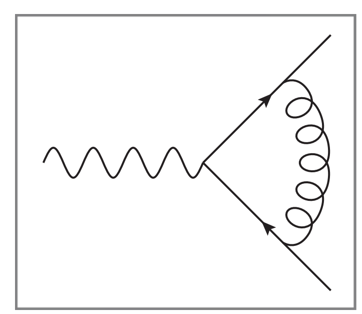
$$\sigma_{NLO} = \int_n \alpha_s^b d\sigma_0 + \int_n \alpha_s^{b+1} d\sigma_V + \int_{n+1} \alpha_s^{b+1} d\sigma_R$$

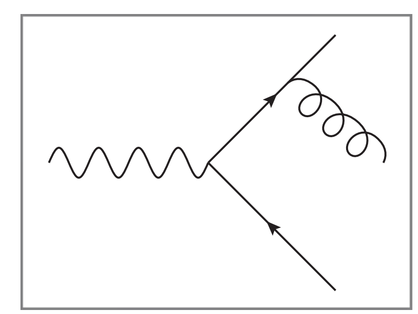
Born cross section

Virtual corrections

Real-emission corrections

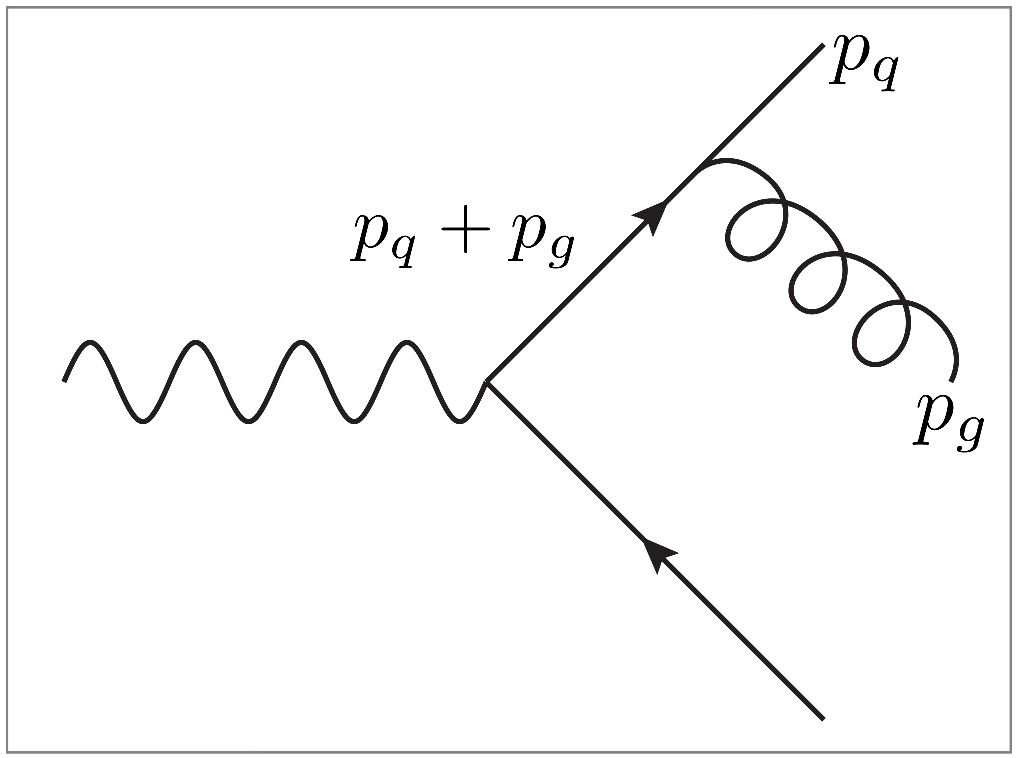






- Remember: virtual and reals are not separately finite, but their sum is (KLN theorem). Divergences have to be subtracted before numerical integration

IR-singularities in the real emission



$$\int_{n+1} \alpha_s^{b+1} d\sigma_R$$

- When the integral over the phase- space of the gluon is performed, one can have $(p_q + p_g)^2 = 0$
- Since $(p_q + p_g)^2 = 2E_q E_g (1 - \cos\theta)$, it can happen when $E_g = 0$ (soft) or $\cos\theta = 1$ (collinear)
- In both cases, the propagator diverges

IR-singularities in the virtual corrections

- The same IR singularities as in the real-emission corrections also appear in the (renormalised) virtual corrections, but with opposite sign. (Follows from KLN theorem!)
  - Virtual corrections: integration over the loop momenta gives poles in $1/\epsilon$, with ϵ the dimensional regulator
  - Real corrections: integration over the phase-space gives poles in $1/\epsilon$, with ϵ the dimensional regulator
- Problematic! Integration over the phase-space is performed numerically. Cannot be done in a non-integer number of dimensions!
- Note: observables must not be sensitive to collinear/soft real emission branching (i.e., for KLN to be applicable). Hence, must use "infrared-safe" observables, and cannot use infinite resolution
- No problem in the virtual corrections: integration over the loop momentum is typically done (semi-)analytically, so poles in ϵ and the finite remainder can be computed explicitly

Example

- Suppose we want to compute the integral
  $\int_0^1 f(x)\,\mathrm{d}x$, with $f(x) = \dfrac{g(x)}{x}$ and $g(x)$ a regular function

- Let's introduce a regulator, which renders the integral finite
  $$\int_0^1 f(x)\,\mathrm{d}x \longrightarrow \int_0^1 x^{\epsilon} f(x)\,\mathrm{d}x = \int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
  and in the end we take the limit $\epsilon \to 0$

- The divergence turns into a pole in ϵ. How can we extract the pole analytically, while doing the integral numerically?

Extraction of poles

$$\int_0^1 f(x)\,\mathrm{d}x \longrightarrow \int_0^1 x^\epsilon f(x)\,\mathrm{d}x = \int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$

Phase-space slicing

- Introduce a small parameter δ:
- $$\int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x = \int_0^\delta \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
  $$\simeq \int_0^\delta \frac{g(0)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
  $$= \left(\frac{1}{\epsilon} + \log\delta\right) g(0) + \int_\delta^1 \frac{g(x)}{x}\,\mathrm{d}x$$
  where we have taken the limit $\epsilon \to 0$ in the 2nd term

Extraction of poles

$$\int_0^1 f(x)\,\mathrm{d}x \longrightarrow \int_0^1 x^\epsilon f(x)\,\mathrm{d}x = \int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$

Phase-space slicing

- Introduce a small parameter δ:
- $$\int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x = \int_0^\delta \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
  $$\simeq \int_0^\delta \frac{g(0)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
  $$= \left(\frac{1}{\epsilon} + \log\delta\right) g(0) + \int_\delta^1 \frac{g(x)}{x}\,\mathrm{d}x$$

  where we have taken the limit $\epsilon \to 0$ in the 2nd term

Subtraction method

- Add and subtract $g(0)/x$:
- $$\int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x = \int_0^1 x^\epsilon \left(\frac{g(0)}{x} + \frac{g(x)}{x} - \frac{g(0)}{x}\right)\mathrm{d}x$$
  $$= \int_0^1 \frac{g(0)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_0^1 \frac{g(x)-g(0)}{x^{1-\epsilon}}\,\mathrm{d}x$$
  $$= \frac{g(0)}{\epsilon} + \int_0^1 \frac{g(x)-g(0)}{x}\,\mathrm{d}x$$

  where we have taken the limit $\epsilon \to 0$ in the 2nd term

Extraction of poles

$$\int_0^1 f(x)\,\mathrm{d}x \longrightarrow \int_0^1 x^\epsilon f(x)\,\mathrm{d}x = \int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$

Phase-space slicing

- Introduce a small parameter δ:
- $$\int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x = \int_0^\delta \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
$$\simeq \int_0^\delta \frac{g(0)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
$$= \left(\frac{1}{\epsilon} + \log\delta\right) g(0) + \int_\delta^1 \frac{g(x)}{x}\,\mathrm{d}x$$

where we have taken the limit $\epsilon \to 0$ in the 2nd term

Subtraction method

- Add and subtract $g(0)/x$:
- $$\int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x = \int_0^1 x^\epsilon \left(\frac{g(0)}{x} + \frac{g(x)}{x} - \frac{g(0)}{x}\right)\mathrm{d}x$$
$$= \int_0^1 \frac{g(0)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_0^1 \frac{g(x)-g(0)}{x^{1-\epsilon}}\,\mathrm{d}x$$
$$= \frac{g(0)}{\epsilon} + \int_0^1 \frac{g(x)-g(0)}{x}\,\mathrm{d}x$$

where we have taken the limit $\epsilon \to 0$ in the 2nd term

- Both methods have a simple *universal* integral to be done *analytically* (that yields the pole to be canceled against the pole in the virtual corrections);
and a complicated *finite* integral to be performed *numerically*
- Since no approximation in the subtraction method, this is the preferred method at NLO
- Since simpler structures in phase-space slicing, this is the preferred method at NNLO

Extraction of poles

$$\int_0^1 f(x)\,\mathrm{d}x \longrightarrow \int_0^1 x^\epsilon f(x)\,\mathrm{d}x = \int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$

Phase-space slicing

- Introduce a small parameter δ:
- $$\int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x = \int_0^\delta \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
$$\simeq \int_0^\delta \frac{g(0)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
$$= \left(\frac{1}{\epsilon} + \log\delta\right) g(0) + \int_\delta^1 \frac{g(x)}{x}\,\mathrm{d}x$$

where we have taken the limit $\epsilon \to 0$ in the 2nd term

Subtraction method

- Add and subtract $g(0)/x$:
- $$\int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x = \int_0^1 x^\epsilon \left(\frac{g(0)}{x} + \frac{g(x)}{x} - \frac{g(0)}{x}\right)\mathrm{d}x$$
$$= \int_0^1 \frac{g(0)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_0^1 \frac{g(x)-g(0)}{x^{1-\epsilon}}\,\mathrm{d}x$$
$$= \frac{g(0)}{\epsilon} + \int_0^1 \frac{g(x)-g(0)}{x}\,\mathrm{d}x$$

where we have taken the limit $\epsilon \to 0$ in the 2nd term

- Both methods have a simple *universal* integral to be done *analytically* (that yields the pole to be canceled against the pole in the virtual corrections); and a complicated *finite* integral to be performed *numerically*
- Since no approximation in the subtraction method, this is the preferred method at NLO
- Since simpler structures in phase-space slicing, this is the preferred method at NNLO

Extraction of poles

$$\int_0^1 f(x)\,\mathrm{d}x \longrightarrow \int_0^1 x^\epsilon f(x)\,\mathrm{d}x = \int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$

Phase-space slicing

- Introduce a small parameter δ:
- $$\int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x = \int_0^\delta \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
  $$\simeq \int_0^\delta \frac{g(0)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_\delta^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x$$
  $$= \left(\frac{1}{\epsilon} + \log\delta\right) g(0) + \int_\delta^1 \frac{g(x)}{x}\,\mathrm{d}x$$

  where we have taken the limit $\epsilon \to 0$ in the 2nd term

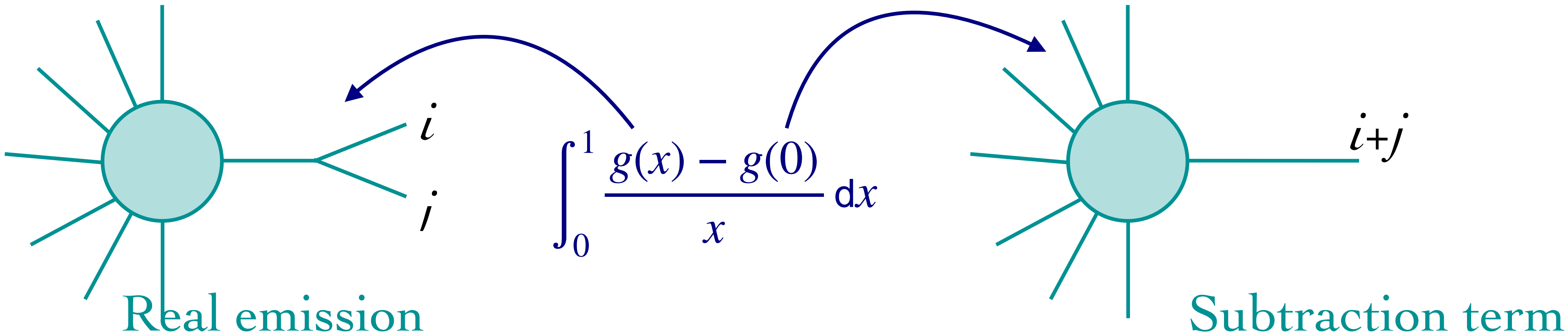
Subtraction method

- Add and subtract $g(0)/x$:
- $$\int_0^1 \frac{g(x)}{x^{1-\epsilon}}\,\mathrm{d}x = \int_0^1 x^\epsilon \left(\frac{g(0)}{x} + \frac{g(x)}{x} - \frac{g(0)}{x}\right)\mathrm{d}x$$
  $$= \int_0^1 \frac{g(0)}{x^{1-\epsilon}}\,\mathrm{d}x + \int_0^1 \frac{g(x)-g(0)}{x^{1-\epsilon}}\,\mathrm{d}x$$
  $$= \frac{g(0)}{\epsilon} + \int_0^1 \frac{g(x)-g(0)}{x}\,\mathrm{d}x$$

  where we have taken the limit $\epsilon \to 0$ in the 2nd term

- Both methods have a simple *universal* integral to be done *analytically* (that yields the pole to be canceled against the pole in the virtual corrections); and a complicated *finite* integral to be performed *numerically*
- Since no approximation in the subtraction method, this is the preferred method at NLO
- Since simpler structures in phase-space slicing, this is the preferred method at NNLO

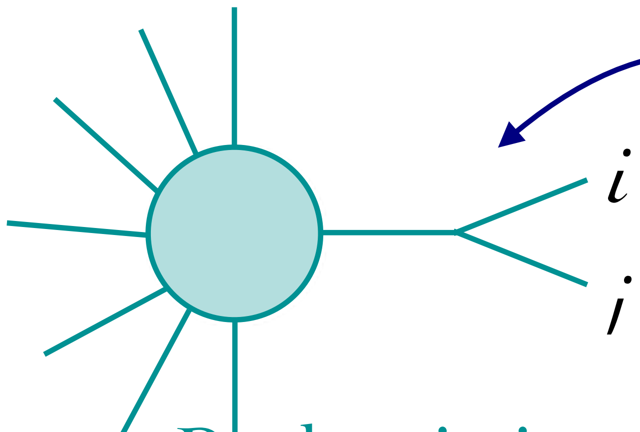
NLO: kinematics of subtraction terms



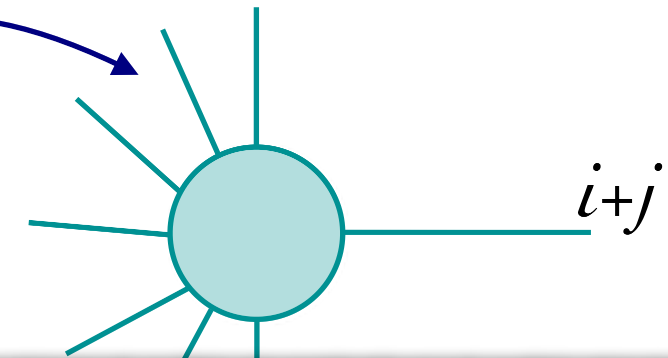
- Real emission and subtraction term cannot be separated (individually, they are divergent!)
- i and j are on-shell in the real emission, but $i + j$ is not: $x \sim m_{i+j}^2$

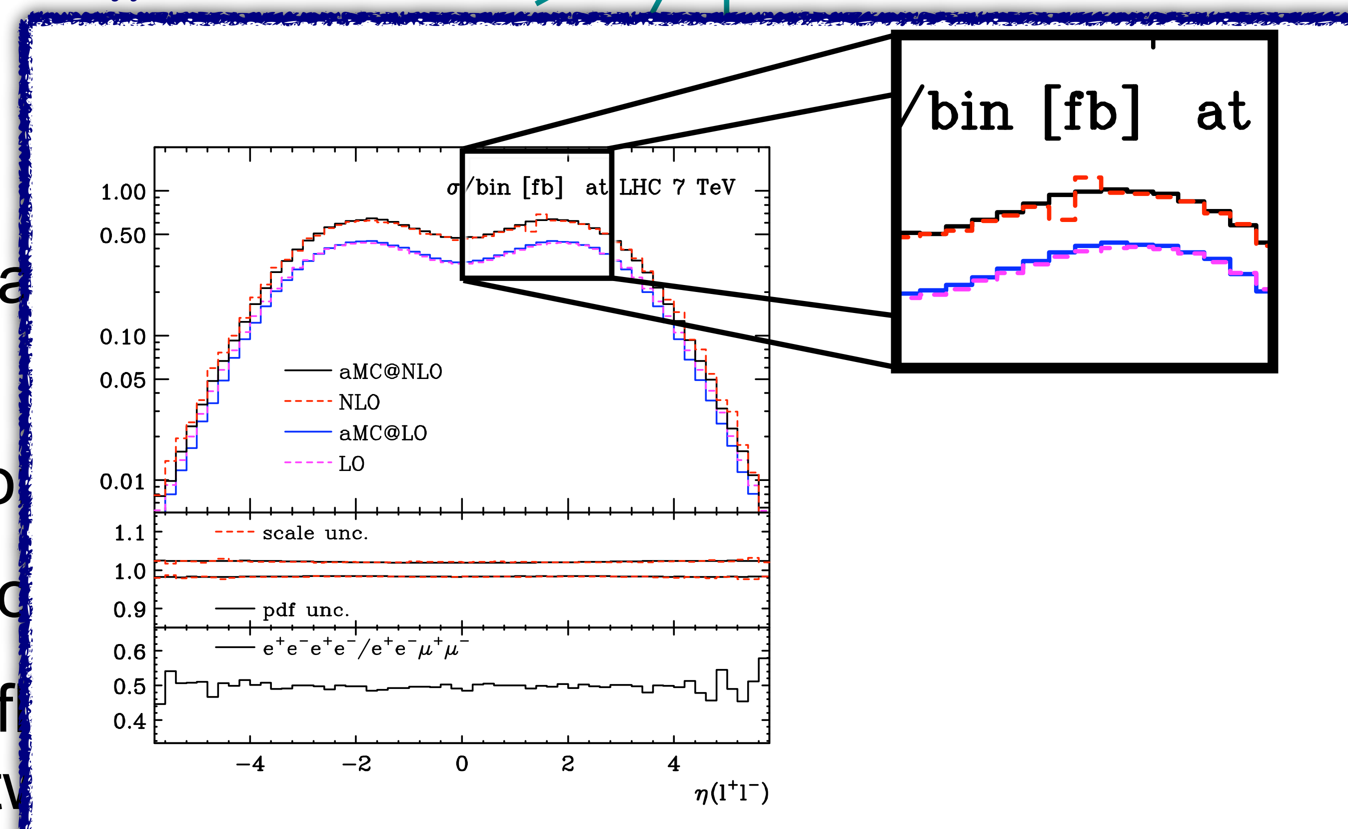
  $i + j$ must be on-shell in the subtraction term
  - This is not possible without reshuffling the momenta of other particles in the process: hence each "event" has two sets of kinematics
  - If can happen, real-emission and the subtraction terms end-up in different histogram bins
    - Use IR-safe observables and don't ask for infinite resolution! (KLN theorem)

NLO: kinematics of subtraction terms



$$\int_0^1 \frac{g(x) - g(0)}{x} \mathrm{d}x$$





- Real emission and subtraction term ca divergent!)
- i and j are on-shell in the real emissio

  $i + j$ must be on-shell in the subtractic
  - This is not possible without reshuffl process: hence each "event" has tv
  - If can happen, real-emission and the subtraction terms end-up in different histogram bins
    - Use IR-safe observables and don't ask for infinite resolution! (KLN theorem)

NLO event unweighting?

- Another consequence of the kinematic mismatch is that we cannot generate unweighted events at NLO
  - $n + 1$-body contribution and n-body contribution are not bounded from above $\rightarrow$ unweighting not possible
  - Further ambiguity on which kinematics to use for the unweighted events

NLO event unweighting?

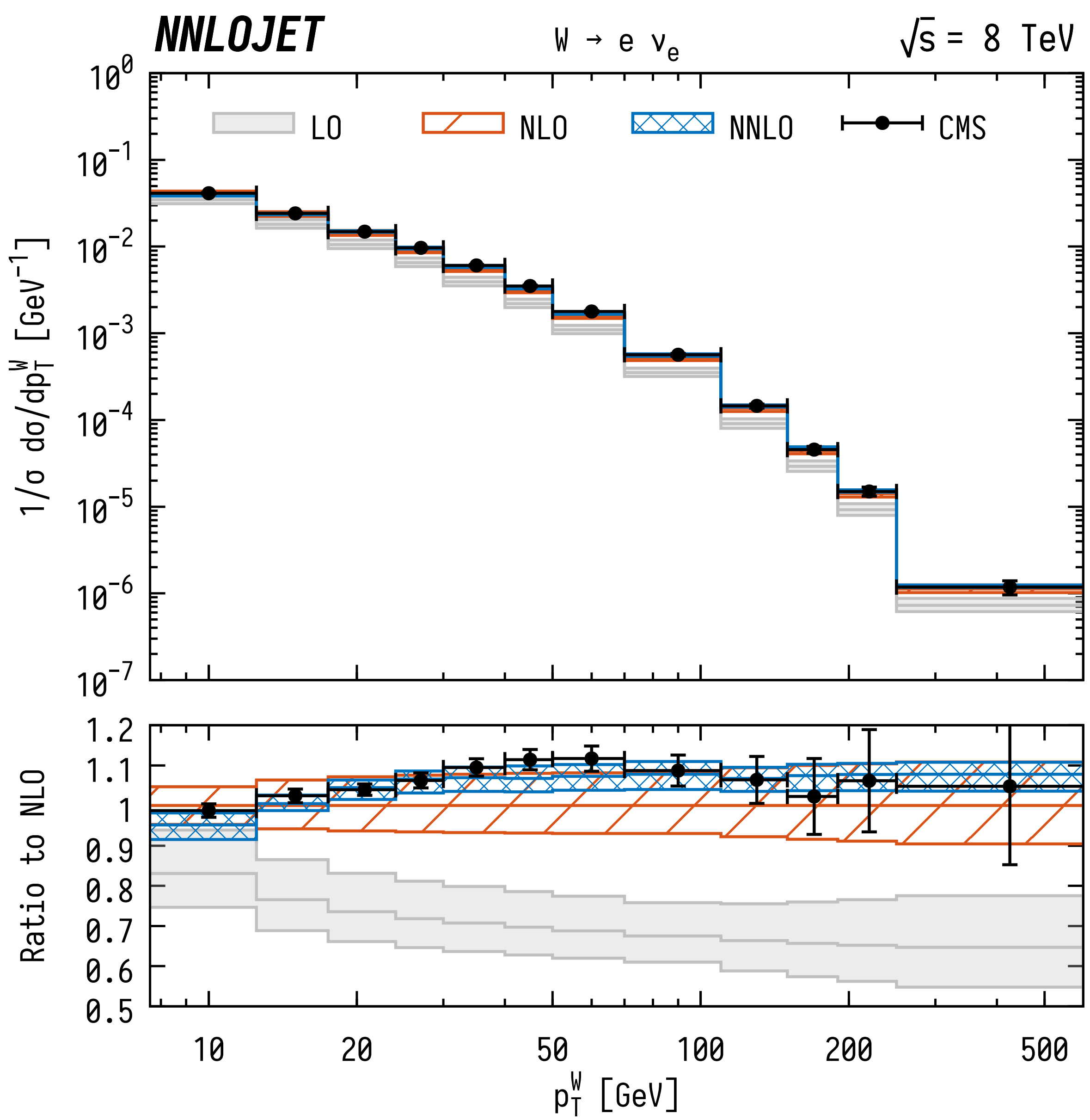
- Another consequence of the kinematic mismatch is that we cannot generate unweighted events at NLO
  - $n+1$-body contribution and n-body contribution are not bounded from above → unweighting not possible
  - Further ambiguity on which kinematics to use for the unweighted events

NLO event unweighting?

- Another consequence of the kinematic mismatch is that we cannot generate unweighted events at NLO
  - $n+1$-body contribution and n-body contribution are not bounded from above $\rightarrow$ unweighting not possible
  - Further ambiguity on which kinematics to use for the unweighted events

For NLO event generation (and parton-shower matching) we need additional work
more on this in the next lecture(s)

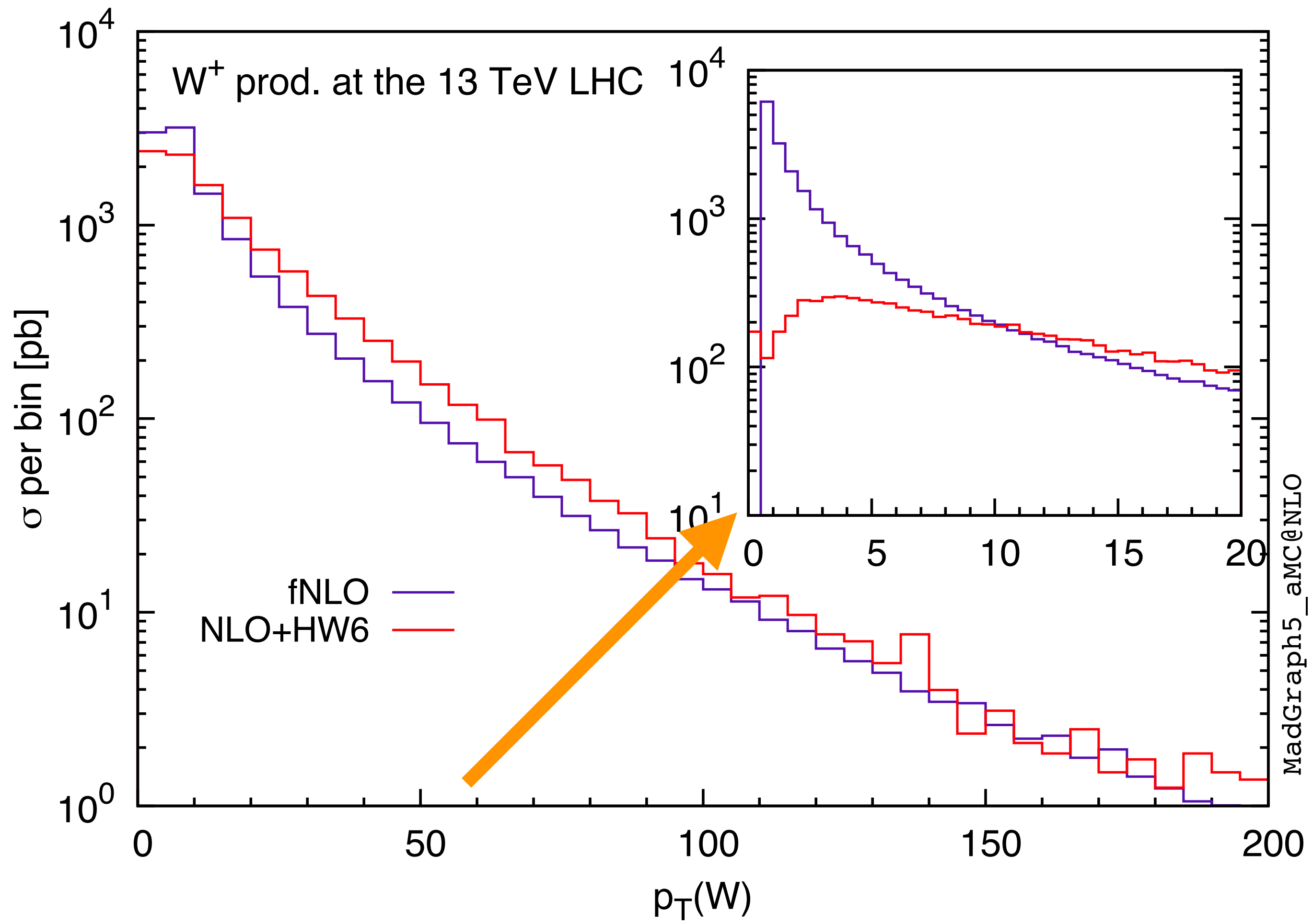
Example: W+j production

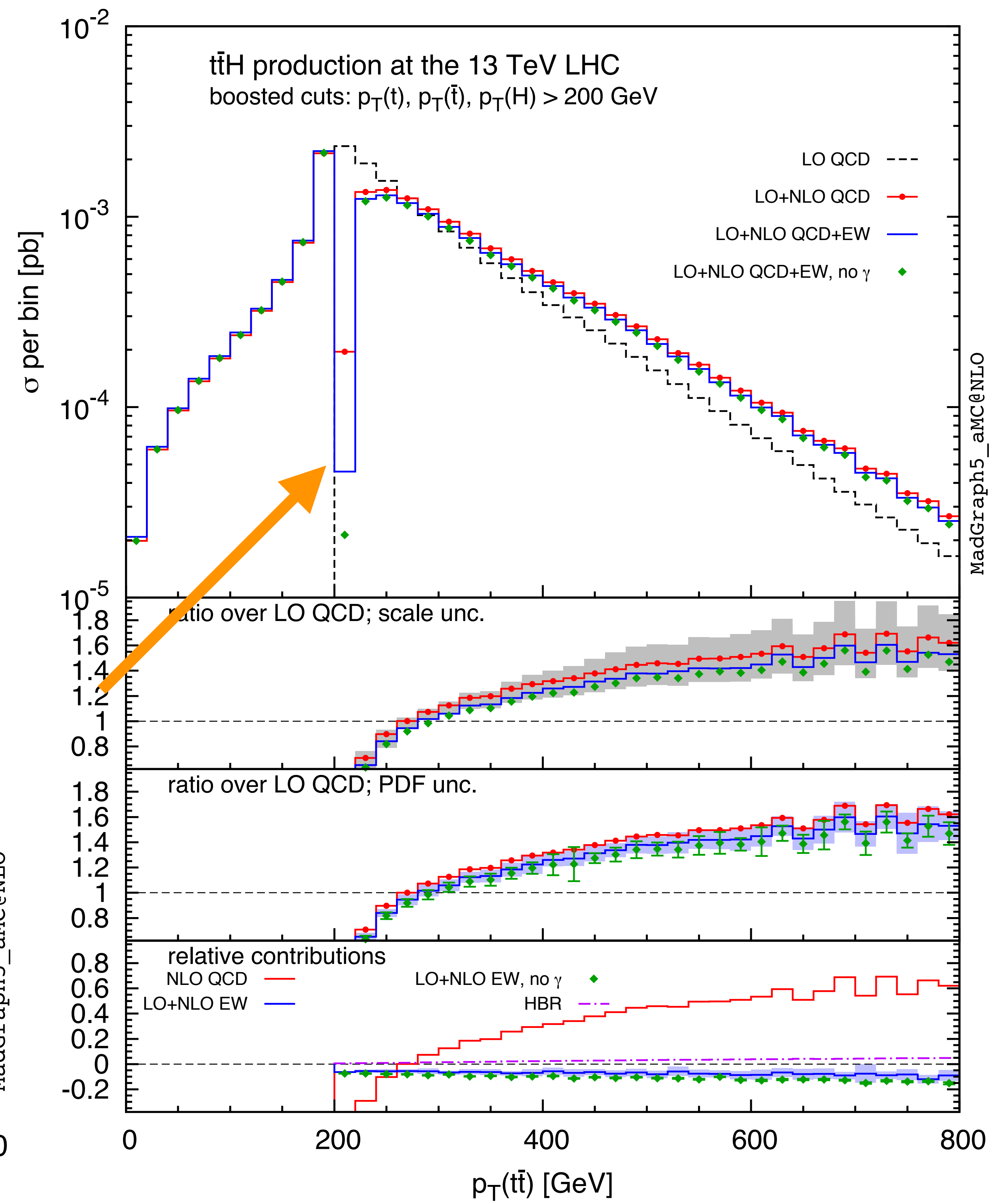


- Both NLO and NNLO agree with the CMS data (8 TeV collisions),
  - NNLO has significantly smaller uncertainties
- LO uncertainties underestimated
  - In general: NLO accuracy required to describe LHC data

Instabilities at fixed order

- Besides the mis-binning problem, the kinematics mismatch can lead to odd behaviours of certain observables, in particular when some constraint coming from the n-body kinematics is relaxed in the $n+1$-body one

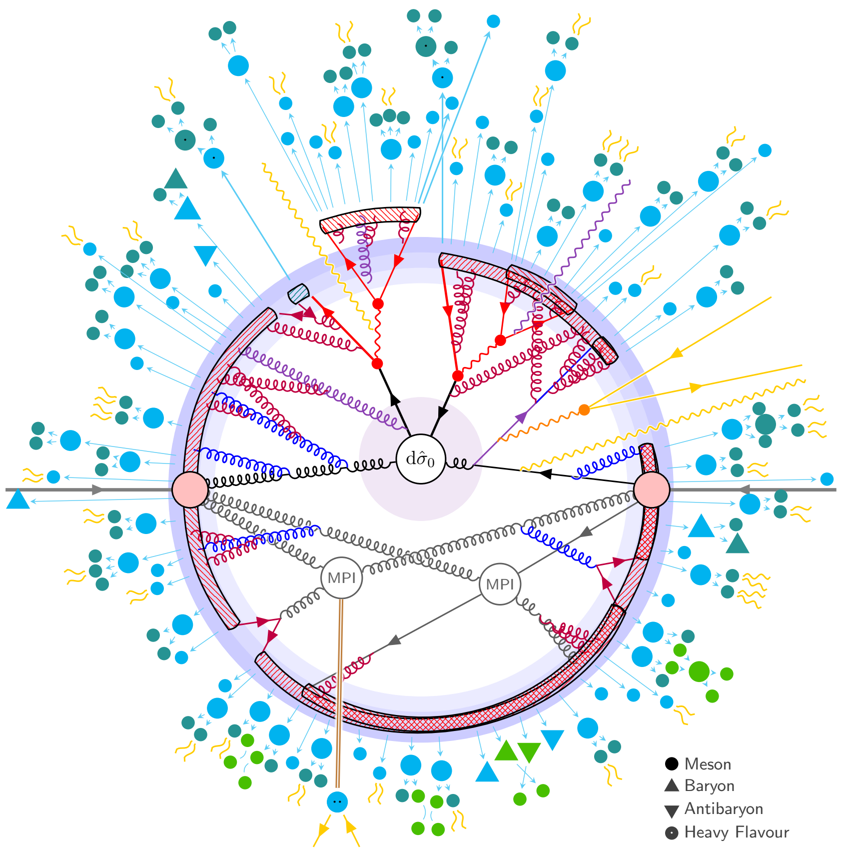




Summary: the hard interaction

- Event generators are there to bridge the gap between theory concepts and experimental concepts
- At the heart, we have a *matrix-element generator*
- Most-difficult part: Phase-space integration by using Monte-Carlo techniques
  - scales very good with number of dimensions
  - also works with involved integration boundaries (cuts!)
  - allows for event simulation
- For the generation of "unweighted" events, an acceptance/ rejection step needs to be performed

Summary: the hard interaction



- Only discussed the central part of the collision.
- Sometimes this is enough!
  - No matching to parton shower
  - Easy to go beyond LO
  - Analytic resummation (instead of resummation with PS also a way forward, and possibly higher accuracy)

$$\sum_{a,b} \int dx_1 dx_2 d\Phi_{\mathrm{FS}}\, f_a(x_1, \mu_F) f_b(x_2, \mu_F)\, \hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$$

Phase-space integral

Parton density functions

Parton-level cross section