Monte Carlo Generators

Rikkert Frederix
Lund University

2021 CERN-Fermilab HCP Summer School, CERN, Aug.23 - Sep.4, 2021

Success of the LHC

- Searches for New Physics are relying more and more upon high-precision comparisons between theory and data
  - Large data samples, methods to reduce systematics
  - High precision computations
- We are scrutinising the Standard Model at higher and higher precision and in smaller and smaller corners of the phase-space
  - The ultimate stress-test for our predictions
- What goes into these predictions…?

Bridging the gap

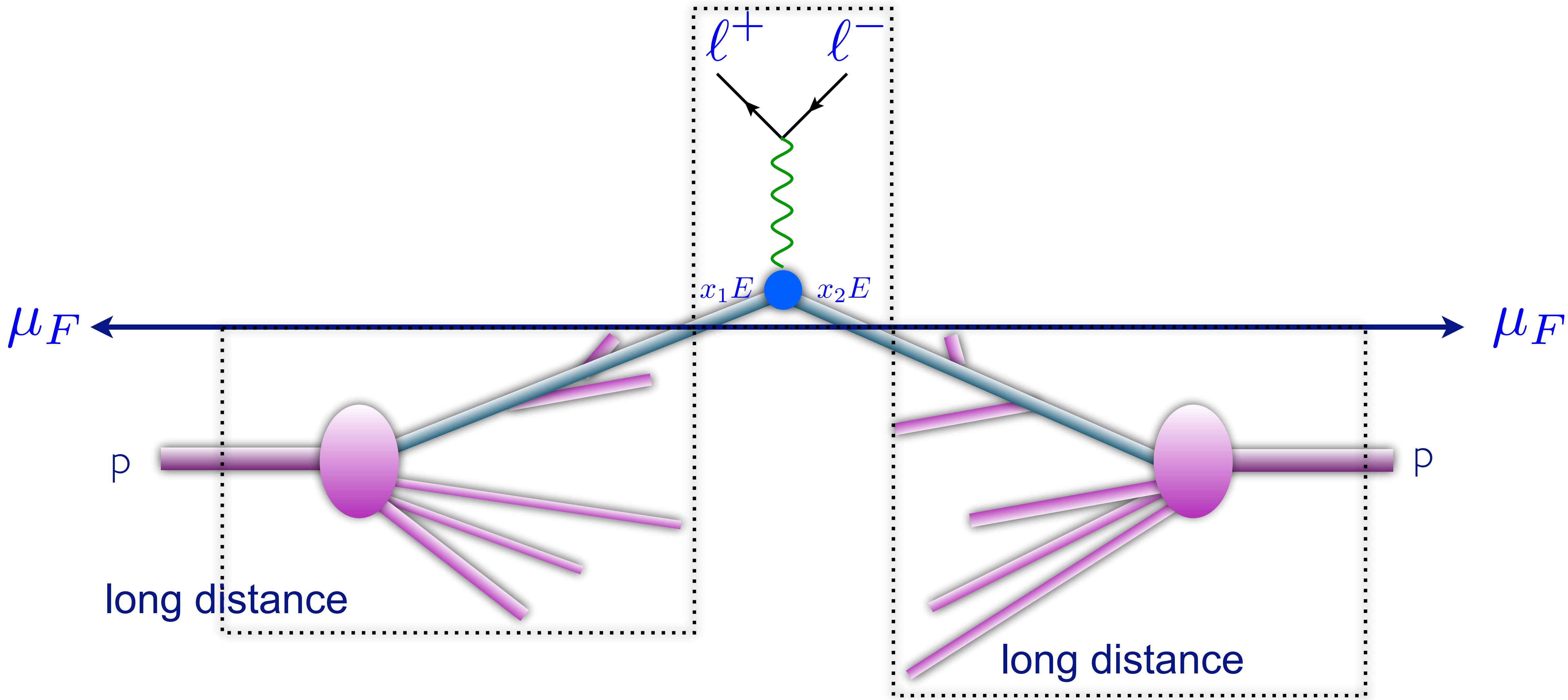
Theory

Lagrangian
Gauge Invariance
Partons
Fixed Order Corrections
Resummation
...

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

Experiment

Factorisation



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

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

Master equation for hadron colliders

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

Phase-space integral

Parton density functions

Parton-level cross section

Two ingredients necessary:

1. Parton distribution functions
(from experiment, but evolution from theory)

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

- 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

Predictions at LO

How do we calculate a LO cross section for 3 jets at the LHC?

I. 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

II. For each one, calculate the amplitude:

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

moderate

III. Square the amplitude, sum over spins & color, integrate over the phase space (D ~ 3n)

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

quite hard

Phase-space integral

- Calculations of cross section or decay widths involve integrations over phase space of very complex functions

$$Dim[\Phi(n)] \sim 3n$$

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

General and flexible method is needed:

Numerical (Monte Carlo) integration

Integrals as averages

$$I = \int_{x_1}^{x_2} f(x)dx \quad \Rightarrow \quad I_N = (x_2 - x_1)\frac{1}{N}\sum_{i=1}^{N} f(x)$$

$$V = (x_2 - x_1)\int_{x_1}^{x_2} [f(x)]^2 dx - I^2 \quad \Rightarrow \quad V_N = (x_2 - x_1)^2 \frac{1}{N}\sum_{i=1}^{N}[f(x)]^2 - I_N^2$$

- Integral as a sum: $I = I_N \pm \sqrt{V_N / N}$

☞ Convergence is slow but it can be estimated easily

☞ Error does not depend on # of dimensions!

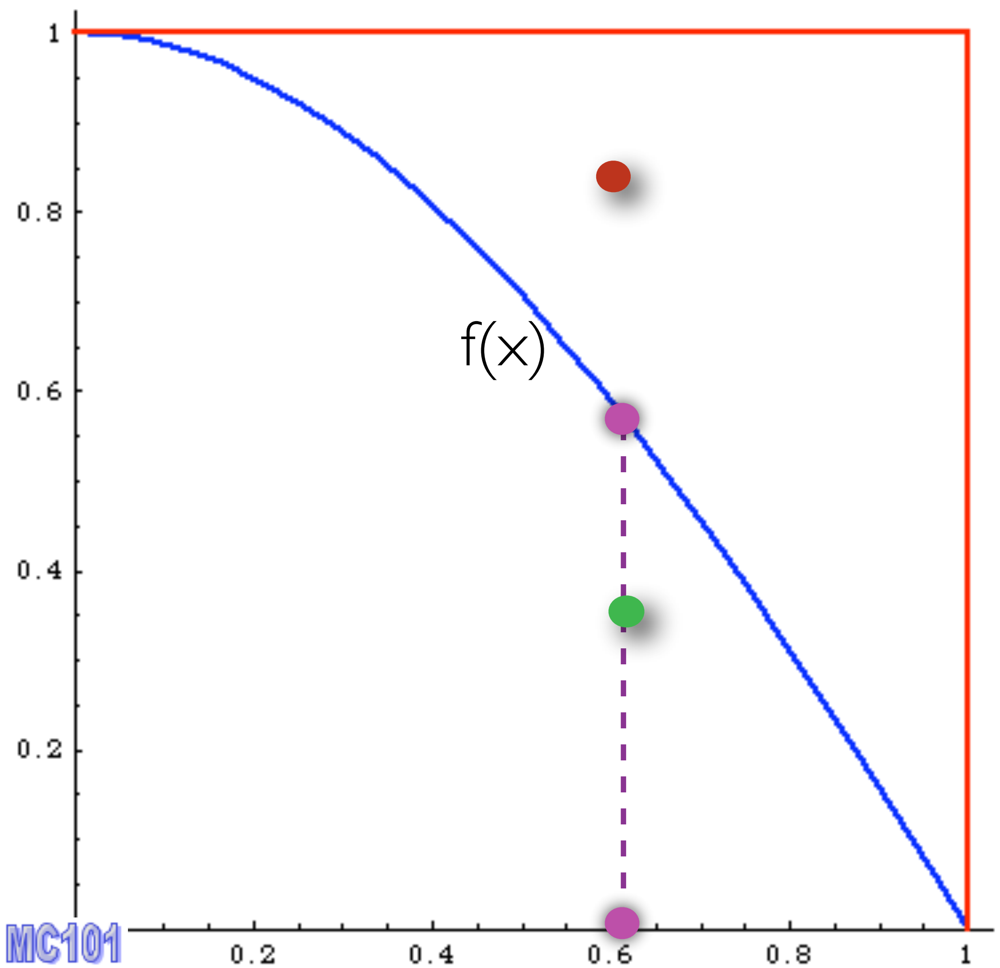
☞ Improvement by minimising V_N

☞ 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 cross section is large
  - events with small weights where the cross section is small
- In nature, the events don’t carry a weight:
  - more events where the cross section is large
  - less events where the cross section is small
- How to go from weighted events to unweighted events?

Event generation

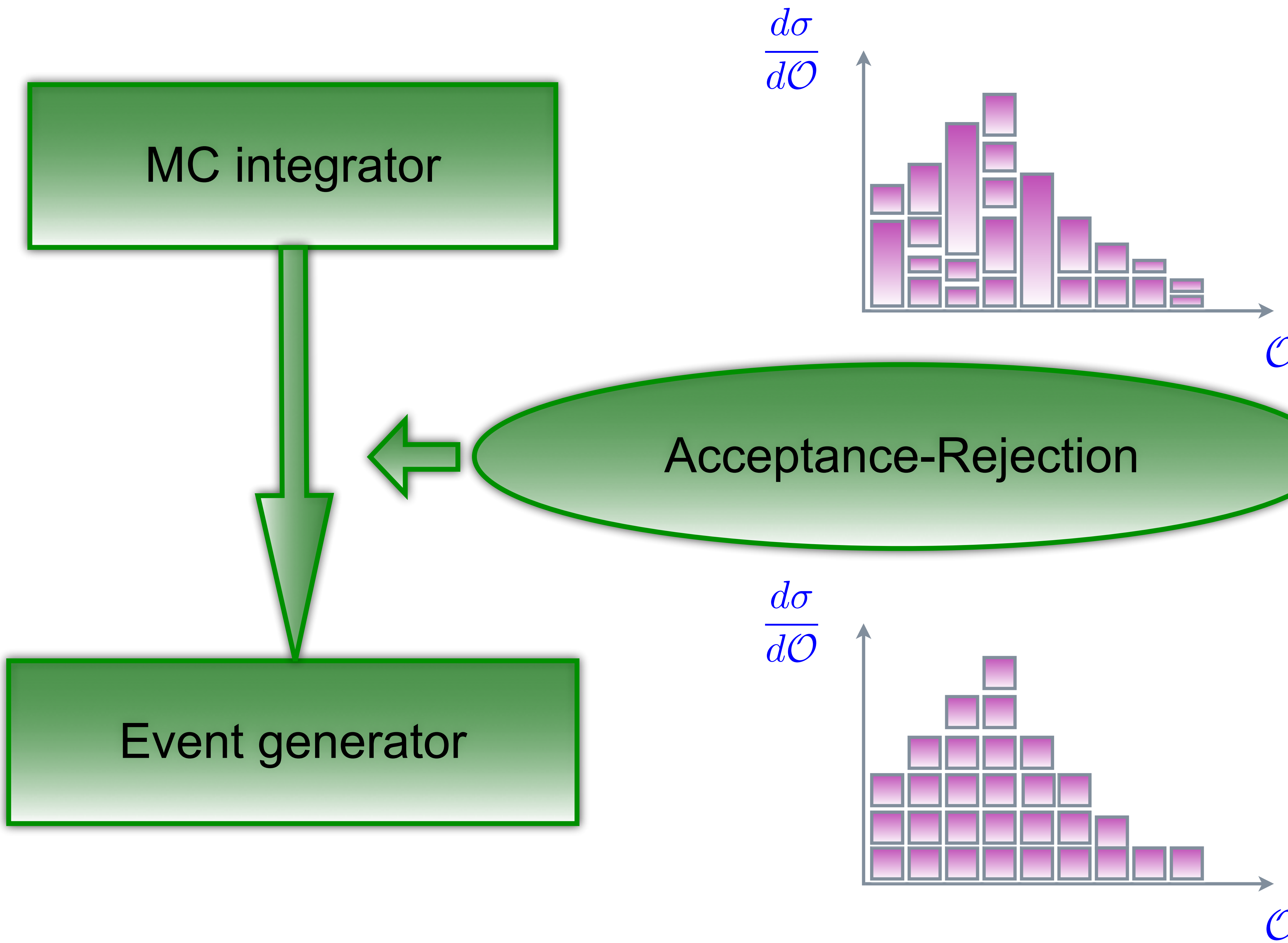


Alternative way

1. (randomly) pick x
2. calculate f(x)
3. (randomly) pick 0<y<fmax
4. Compare:
   if f(x)>y accept event,
   else reject it.

$$\frac{\text{accepted}}{\text{total tries}} = \text{efficiency}$$

Event generation



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

MC Event generator: definition

At the most basic level a Monte Carlo event generator is a program which produces particle physics events with the same probability as they occur in nature (virtual collider).

In practice it performs a large number of (sometimes very difficult) integrals and then unweights to give the four momenta of the particles that interact with the detector (simulation).

Note that, at least among theorists, the definition of a “Monte Carlo program” also includes codes which don’t provide a fully exclusive information on the final state but only cross sections or distributions at the parton level, even when no unweighting can be performed.

Higher order corrections

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:

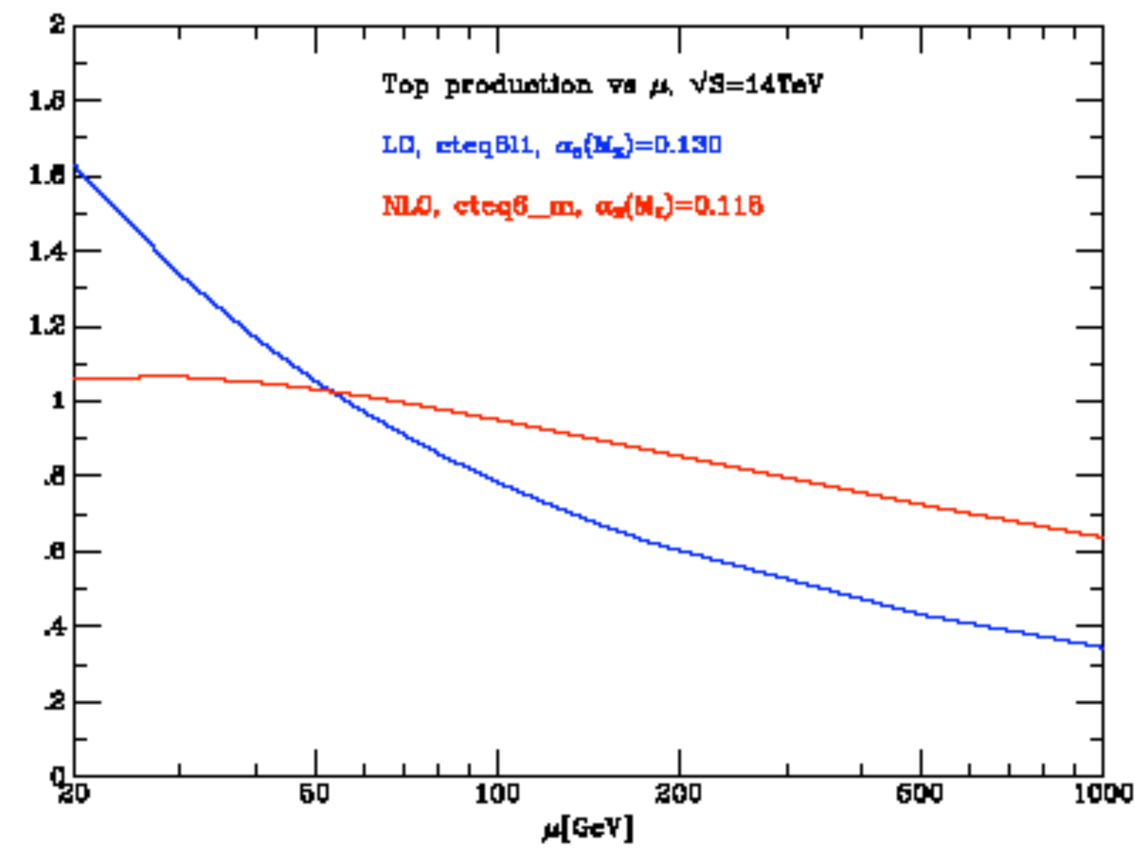
$$\hat{\sigma} = \sigma^{\mathrm{Born}}\left(1 + \frac{\alpha_s}{2\pi}\sigma^{(1)} + \left(\frac{\alpha_s}{2\pi}\right)^2\sigma^{(2)} + \left(\frac{\alpha_s}{2\pi}\right)^3\sigma^{(3)} + \ldots\right)$$

LO predictions | NLO corrections | NNLO corrections | NNNLO corrections

- Including higher corrections improves predictions and reduces theoretical uncertainties

Going NLO

- At NLO the dependence on the renormalisation and factorisation scales is reduced
  - First order where scale dependence in the running coupling and the PDFs is compensated for via the loop corrections: **first reliable estimate of the total cross section**
  - Better description of final state: impact of extra radiation included (e.g. jets can have substructure)
  - Opening of additional initial state partonic channels



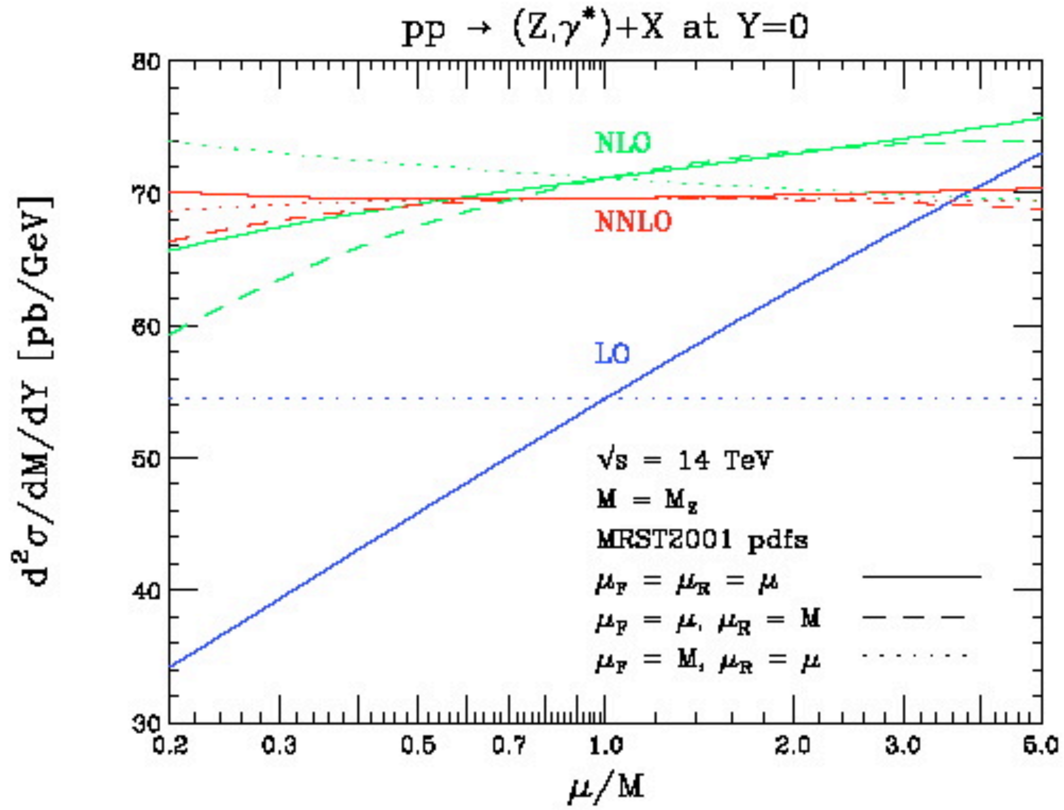
NLO corrections

- NLO corrections have three parts:
  - The Born contribution, i.e. the Leading order.
  - Virtual (or Loop) corrections: formed by an amplitude with a closed loop of particles interfered with the Born amplitudes
  - Real emission corrections: formed by amplitudes with one extra parton compared to the Born process
- Both Virtual and Real emission have one power of α_s extra compared to the Born process

$$\sigma^{\mathrm{NLO}} = \int_m d\sigma^B + \int_m d\sigma^V + \int_{m+1} d\sigma^R$$

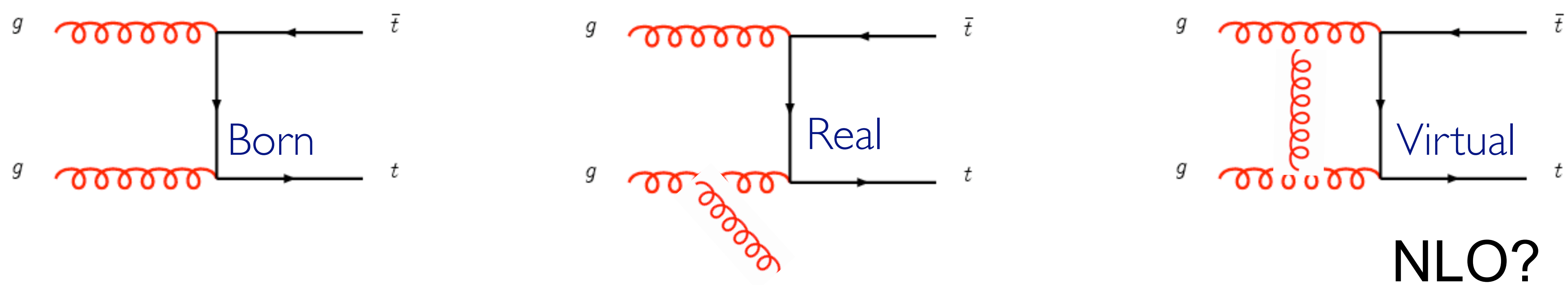
Going NNLO

- NNLO is now where NLO was a few years before automation: Nearly all relevant 2->2 have been computed, and 2->3 is underway
- Why do we need it?
  - An NNLO calculation gives control of the uncertainties in a calculation
  - It is "mandatory" if NLO corrections are exceptionally large to check the behaviour of the perturbative series
  - It is needed for Standard Candles and very precise tests of perturbation theory, exploiting all the available information, e.g. for determining NNLO PDF sets



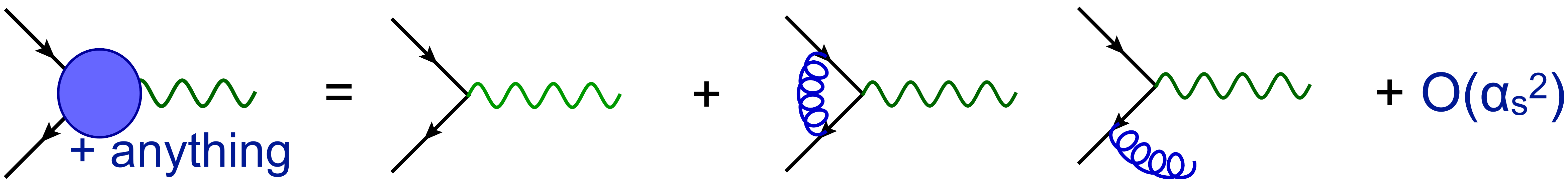
NLO...?

- Are all (IR-safe) observables that we can compute using a NLO code correctly described at NLO? Suppose we have a NLO code for pp ⟶ ttbar



	NLO?
• Total cross section	✔
• Transverse momentum of the top quark	✔
• Transverse momentum of the top-antitop pair	✘
• Transverse momentum of a jet	✘
• Top-antitop invariant mass	✔
• Azimuthal distance between the top and anti-top	✘

Obstacles

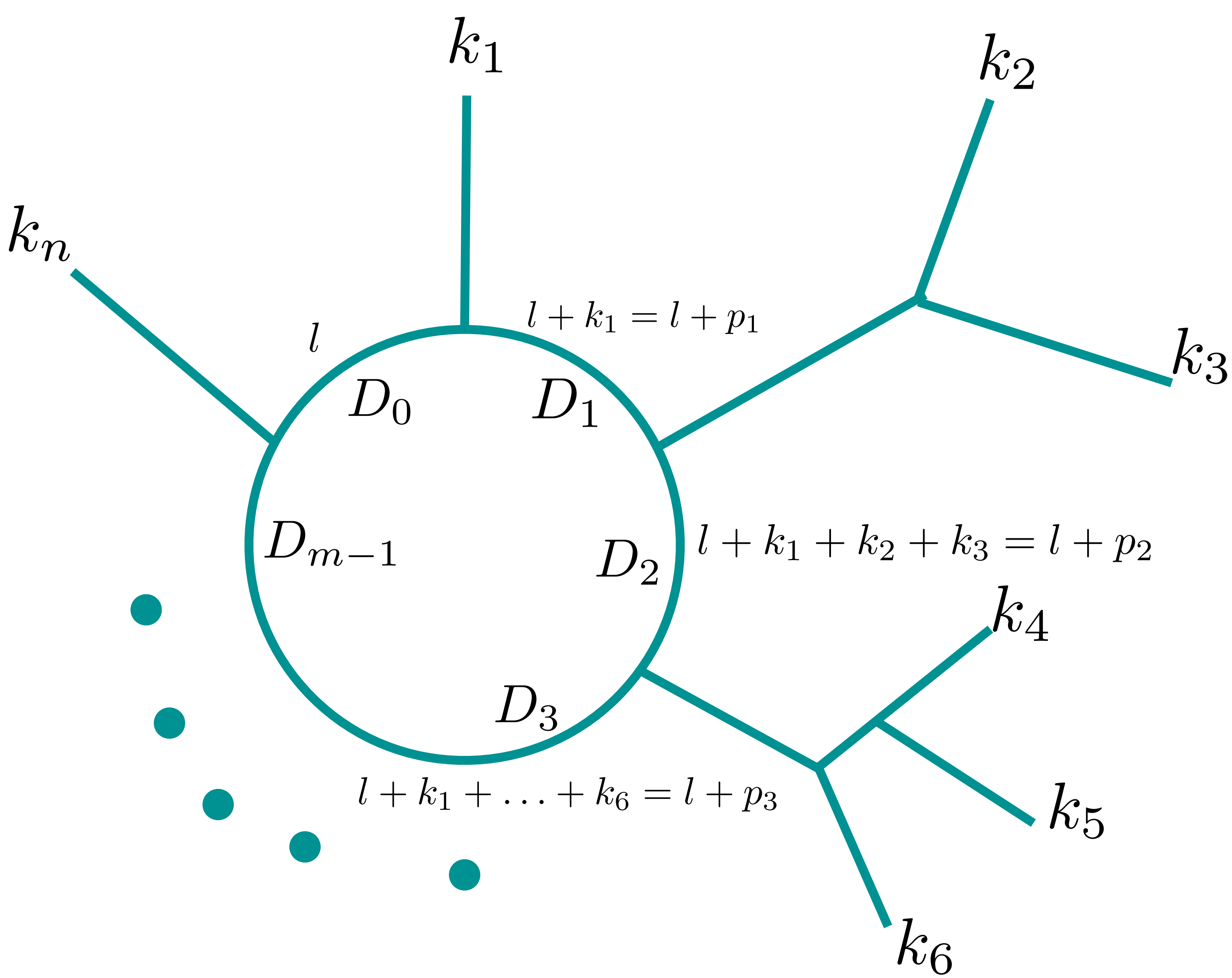


- Let us focus on NLO... there are already enough steps to be taken:
  - Virtual amplitudes: how to compute loop diagrams in a reasonable amount of time (and stable!)
  - How to deal with infra-red divergences: virtual corrections and real-emission corrections are separately divergent and only their sum is finite (for IR-safe observables) according to the KLN theorem
  - How to match these processes to a parton shower without double counting

NLO: virtual corrections

one-loop integral

- Consider this *m*-point loop diagram with *n* external momenta



- The integral to compute is

$$\int d^d l \, \frac{N(l)}{D_0 D_1 D_2 \cdots D_{m-1}}$$

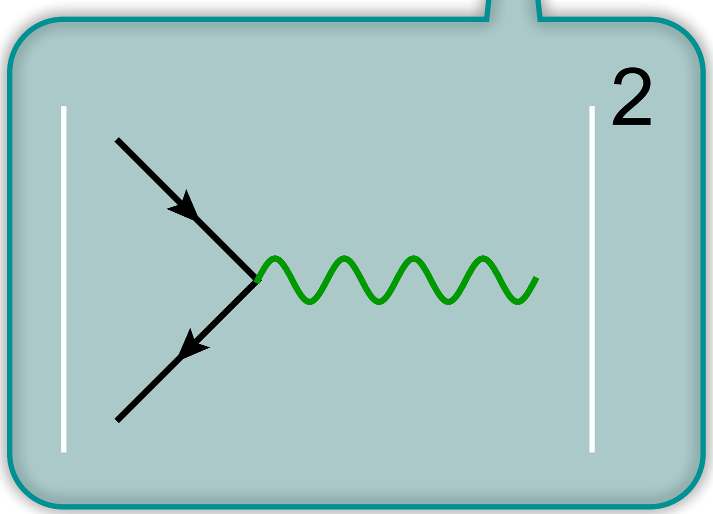
$$D_i = (l + p_i)^2 - m_i^2$$

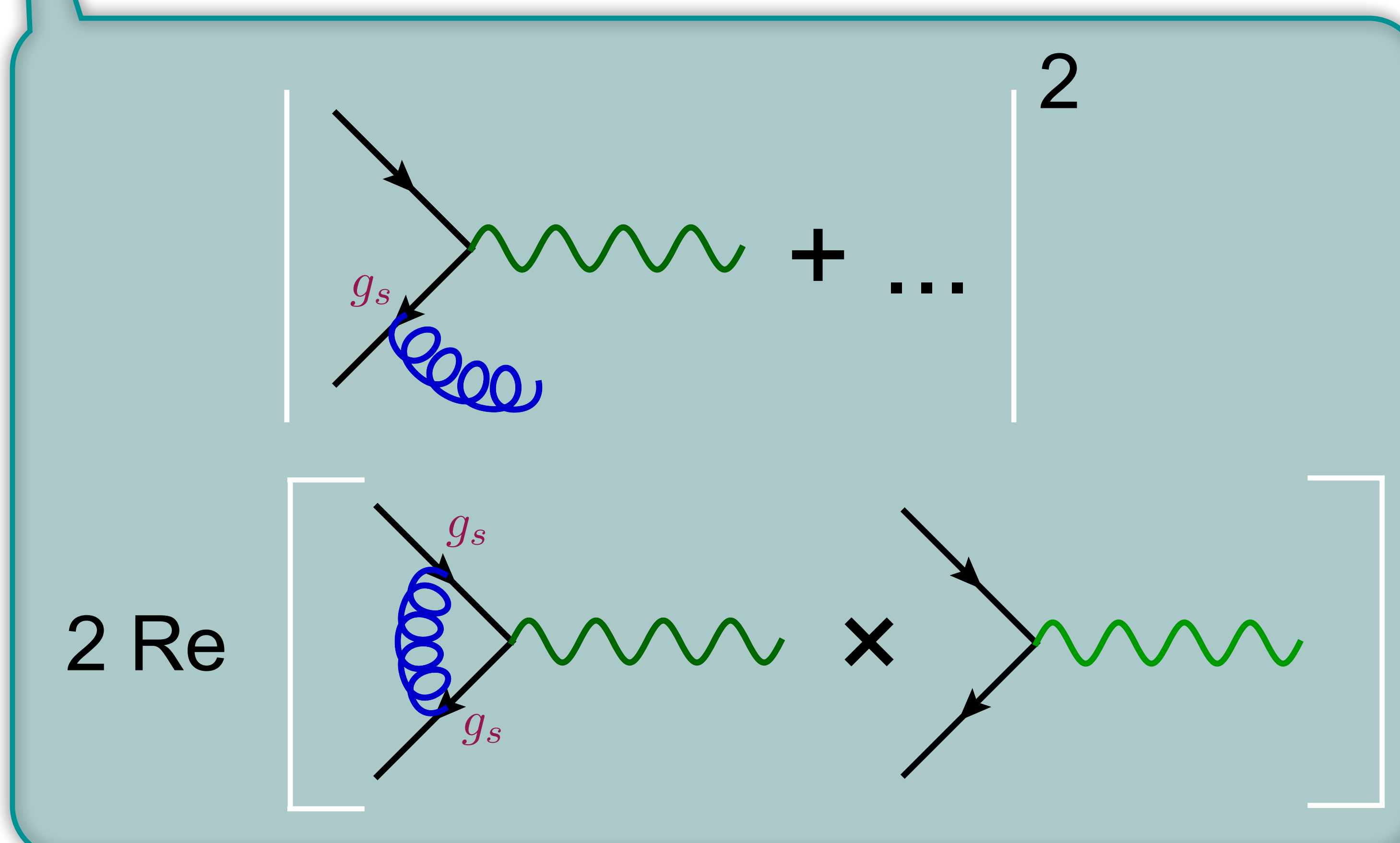
- After UV renormalisation, this gives IR poles in the dimensional regulator and a finite contribution

NLO: infrared divergences

NLO predictions

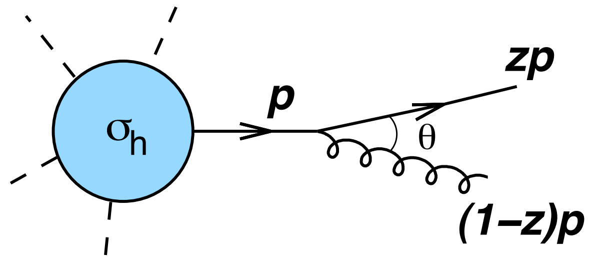
$$\hat{\sigma} = \sigma^{\text{Born}}\left(1 + \frac{\alpha_s}{2\pi}\sigma^{(1)} + \dots\right)$$





Branching

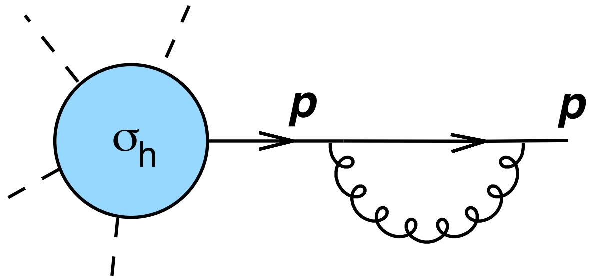
- In the collinear region, the branching of a gluon from a quark can be written as



$$\sigma_{h+g} \simeq \sigma_h \frac{\alpha_s C_F}{\pi} \frac{dz}{1-z} \frac{dk_t^2}{k_t^2}$$

where k_t is the transverse momentum of the gluon, $k_t = E \sin\theta$.

- The singularities cancel against the singularities in the virtual corrections, which result from the integral over the loop momentum of the function



$$\sigma_{h+V} \simeq -\sigma_h \frac{\alpha_s C_F}{\pi} \frac{dz}{1-z} \frac{dk_t^2}{k_t^2}$$

- The sum is finite for observables that cannot distinguish between two collinear partons ($k_t \longrightarrow 0$); a hard and a soft parton ($z \longrightarrow 1$); and a single parton (in the virtual contributions)

Infrared cancellation

$$\sigma^{\rm NLO} \sim \int d^4\Phi_m\, B(\Phi_m) + \int d^4\Phi_m \int_{\rm loop} d^d l\, V(\Phi_m) + \int d^d\Phi_{m+1}\, R(\Phi_{m+1})$$

- The KLN theorem tells us that divergences from virtual and real-emission corrections cancel in the sum for observables insensitive to soft and collinear radiation (“IR-safe observables”)
- When doing an analytic calculation in dimensional regularisation this can be explicitly seen in the cancellation of the $1/\epsilon$ and $1/\epsilon^2$ terms (with ϵ the regulator, $\epsilon \to 0$)
- In the real emission corrections, the explicit poles enter after the phase-space integration (in d dimensions)

Infrared safe observables

- For an observable to be calculable in fixed-order perturbation theory, the observable should be infrared safe, i.e., it should be insensitive to the emission of soft or collinear partons.
- In particular, if p_i is a momentum occurring in the definition of an observable, it most be invariant under the branching

  $$p_i \longrightarrow p_j + p_k,$$

  whenever p_j and p_k are collinear or one of them is soft.
- Examples
  - "The number of gluons" produced in a collision is not an infrared safe observable
  - "The number of hard jets defined using the k_T algorithm with a transverse momentum above 40 GeV," produced in a collision is an infrared safe observable

Phase-space integration

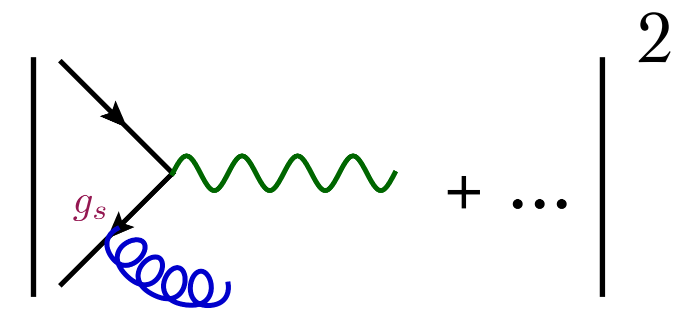
$$\sigma^{\text{NLO}} \sim \int d^4\Phi_m\, B(\Phi_m) + \int d^4\Phi_m \int_{\text{loop}} d^d l\, V(\Phi_m) + \int d^d\Phi_{m+1}\, R(\Phi_{m+1})$$

- For complicated processes we have to result to numerical phase-space integration techniques (“Monte Carlo integration”), which can only be performed in an integer number of dimensions
  - Cannot use a finite value for the dimensional regulator and take the limit to zero in a numerical code
- But we still have to cancel the divergences explicitly
- Use a subtraction method to explicitly factor out the divergences from the phase-space integrals

Example

- Suppose we want to compute the integral ("real emission radiation", where the 1-particle phase-space is referred to as the 1-dimensional *x*)

$$\int_0^1 dx\, f(x)$$



where $f(x) = \frac{g(x)}{x}$ and $g(x)$ is finite everywhere

- Let's introduce a regulator

$$\lim_{\epsilon \to 0} \int_0^1 dx\, \frac{g(x)}{x^{1+\epsilon}} = \lim_{\epsilon \to 0} \int_0^1 dx\, x^{-\epsilon} f(x)$$

for any non-integer value for ϵ this integral is finite

- We would like to factor out the explicit poles in ϵ so that they can be canceled explicitly against the virtual corrections

Phase-space slicing

$$\lim_{\epsilon \to 0} \int_0^1 dx\, x^{-\epsilon} f(x) \qquad\qquad f(x) = \frac{g(x)}{x}$$

- Introduce a small parameter δ

$$\begin{aligned}
\lim_{\epsilon \to 0} \int_0^1 dx\, x^{-\epsilon} f(x) &= \lim_{\epsilon \to 0} \left[\int_0^\delta dx\, x^{-\epsilon} f(x) + \int_\delta^1 dx\, x^{-\epsilon} f(x) \right] \\
&= \lim_{\epsilon \to 0} \left[\int_0^\delta dx\, x^{-\epsilon} \frac{g(0)}{x} + \int_\delta^1 dx\, x^{-\epsilon} \frac{g(x)}{x} \right] \\
&= \lim_{\epsilon \to 0} \frac{\delta^{-\epsilon}}{-\epsilon} g(0) + \int_\delta^1 dx\, \frac{g(x)}{x} \\
&= \lim_{\epsilon \to 0} \left[\frac{-1}{\epsilon} + \log \delta \right] g(0) + \int_\delta^1 dx\, \frac{g(x)}{x}
\end{aligned}$$

- We get the explicit pole in ϵ and a finite integral that can be computed numerically

Subtraction method

$$\lim_{\epsilon\to 0}\int_0^1 dx\, x^{-\epsilon} f(x) \qquad\qquad f(x) = \frac{g(x)}{x}$$

- Add and subtract the same term

$$\lim_{\epsilon\to 0}\int_0^1 dx\, x^{-\epsilon} f(x) = \lim_{\epsilon\to 0}\int_0^1 dx\, x^{-\epsilon}\left[\frac{g(0)}{x} + f(x) - \frac{g(0)}{x}\right]$$

$$= \lim_{\epsilon\to 0}\int_0^1 dx\left[g(0)\frac{x^{-\epsilon}}{x} + \frac{g(x)-g(0)}{x^{1+\epsilon}}\right]$$

$$= \lim_{\epsilon\to 0}\frac{-1}{\epsilon}g(0) + \int_0^1 dx\,\frac{g(x)-g(0)}{x}$$

- We have factored out the $1/\epsilon$ divergence and are left with a finite integral
- According to the KLN theorem the divergence cancels against the virtual corrections

Slicing vs Subtraction

Slicing: $\int_\delta^1 dx\, \frac{g(x)}{x} + g(0)\log\delta$

Subtraction: $\int_0^1 dx\, \frac{g(x)-g(0)}{x}$

"Plus distribution"

- Terms of order δ are neglected in the slicing method; the subtraction method is exact
  - One has to show that δ is taken small enough such that no observable depends on it; in practice well below 1 GeV
- Both methods feature cancellations between large numbers: if for an observable O, if $\lim_{x\to 0} O(x) \neq O(0)$ or we choose the bin-size too small, instabilities render the computation useless
  - We already knew that! KLN is sufficient; one must have infra-red safe observables and cannot ask for infinite resolution
- Subtraction method is more flexible -> method of choice in automation

NLO with Subtraction

$$\sigma^{\rm NLO} \sim \int d^4\Phi_m\, B(\Phi_m) + \int d^4\Phi_m \int_{\rm loop} d^d l\, V(\Phi_m) + \int d^d\Phi_{m+1}\, R(\Phi_{m+1})$$

- With the subtraction method this is replace by

$$\sigma^{\rm NLO} \sim \int d^4\Phi_m\, B(\Phi_m)$$

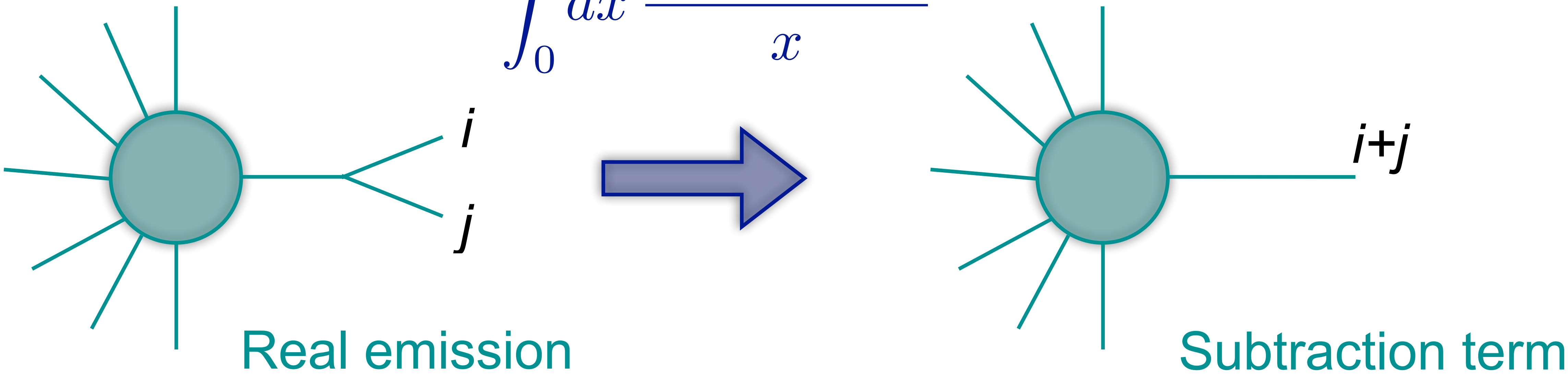
$$+ \int d^4\Phi_m \left[\int_{\rm loop} d^d l\, V(\Phi_m) + \int d^d\Phi_1 G(\overline{\Phi}_{m+1}) \right]_{\epsilon\to 0}$$

$$+ \int d^4\Phi_{m+1} \left[R(\Phi_{m+1}) - G(\overline{\Phi}_{m+1}) \right]$$

- Terms between the brackets are finite. Can integrate them numerically and independent from one another in 4 dimensions

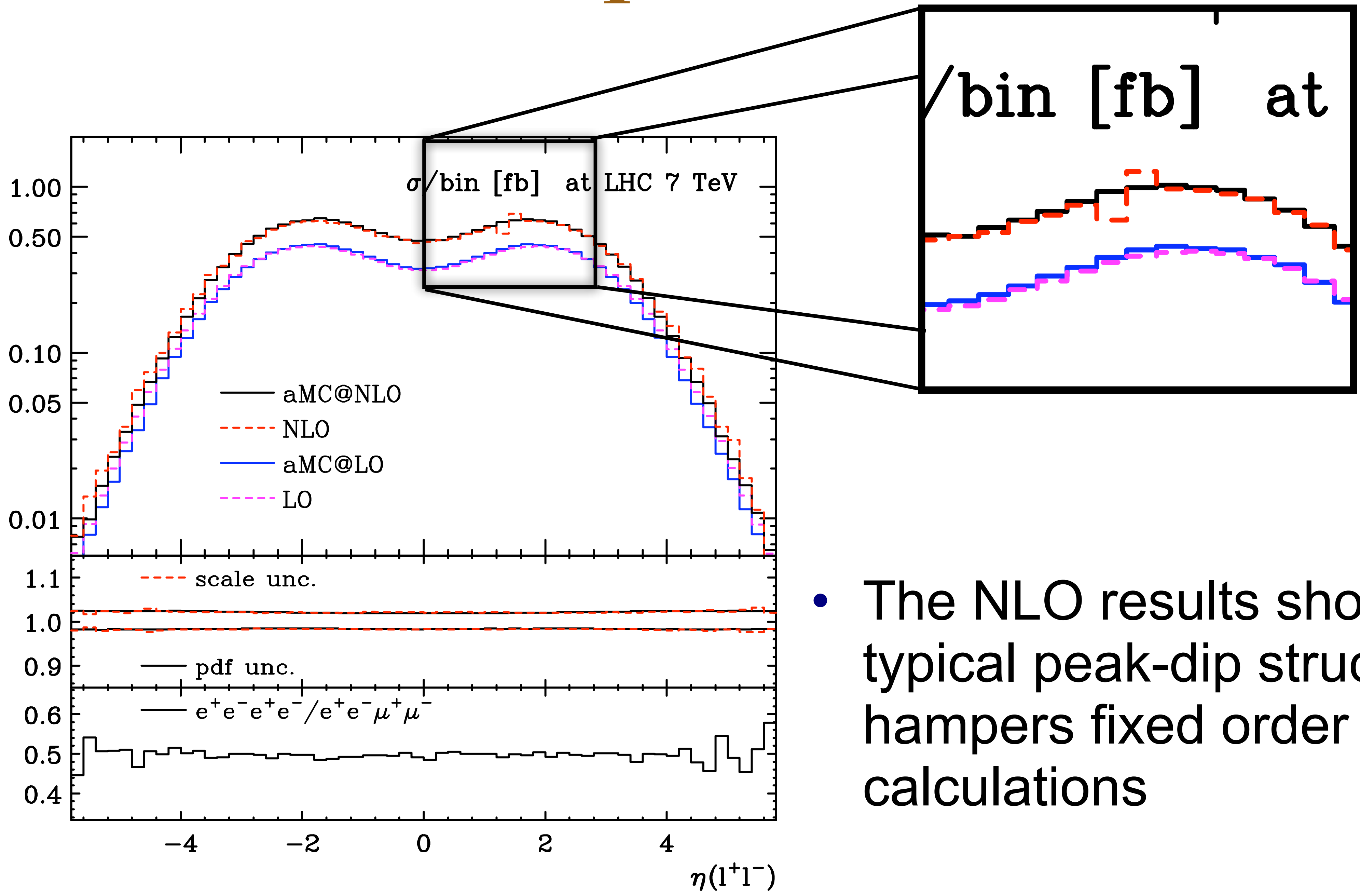
Kinematics of counter events

$$\int_0^1 dx \, \frac{g(x) - g(0)}{x}$$



- If *i* and *j* are two on-shell particles that are present in a splitting that leads to an singularity, for the counter events we need to combine their momenta to a new on-shell parton that's the sum of *i+j*
- This is not possible without changing any of the other momenta in the process
- When applying cuts or making plots, events and counter events might end-up in different bins
  - Use IR-safe observables and don't ask for infinite resolution! (KLN theorem)

Example in 4 charged lepton production



- The NLO results shows a typical peak-dip structure that hampers fixed order calculations

Two methods at NLO

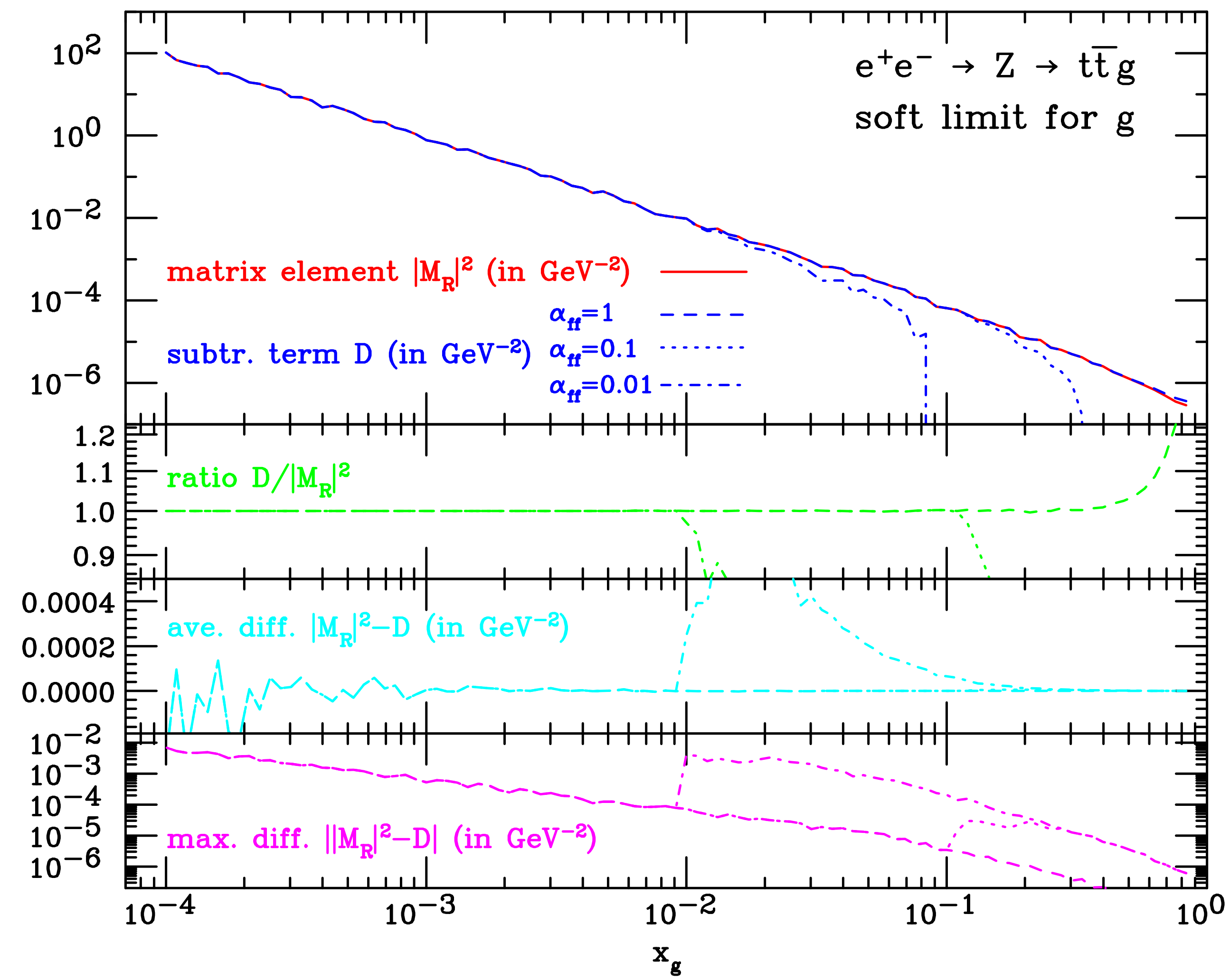
- **Dipole subtraction**
  - Recoil taken by one (colour-connected) parton: N^3 scaling
  - Method evolved from cancellation of the soft divergence
  - Proven to work for simple as well as complicated processes
  - Used in MCFM, Sherpa, …
- **FKS subtraction**
  - Recoil evenly distributed by all particles: N^2 scaling
  - Collinear divergences as a starting point
  - Proven to work for simple as well as complicated processes
  - Automated in MadGraph5_aMC@NLO & POWHEG BOX

A plethora of methods at NNLO

- Slicing:
  - q_T-subtraction
  - N-jettiness subtraction
- Subtraction
  - Antennae subtraction
  - Stripper
  - Projection to Born
  - ColoRFulNNLO
  - Nested Soft-Collinear
  - Local analytic
  - …

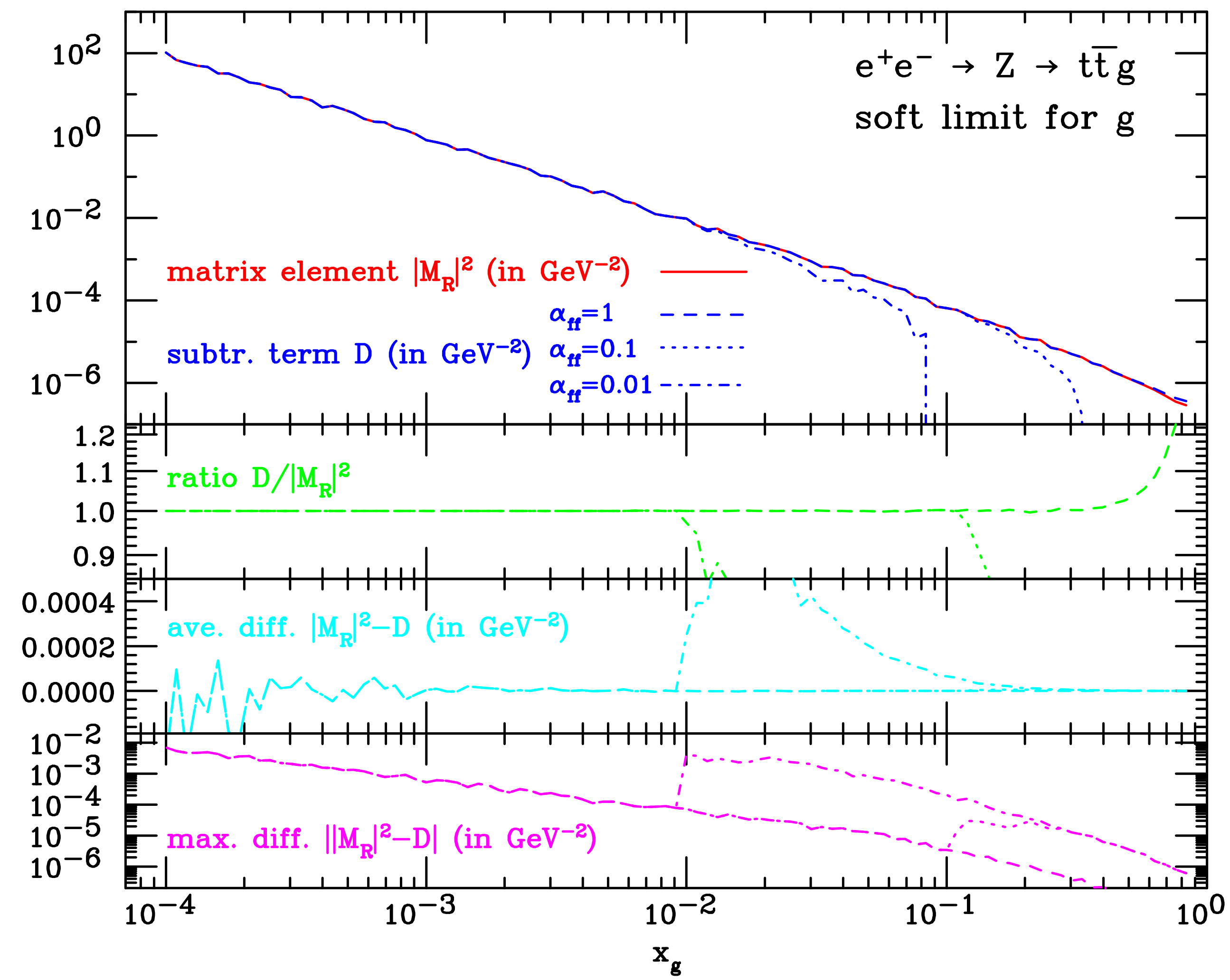
Event unweighting?

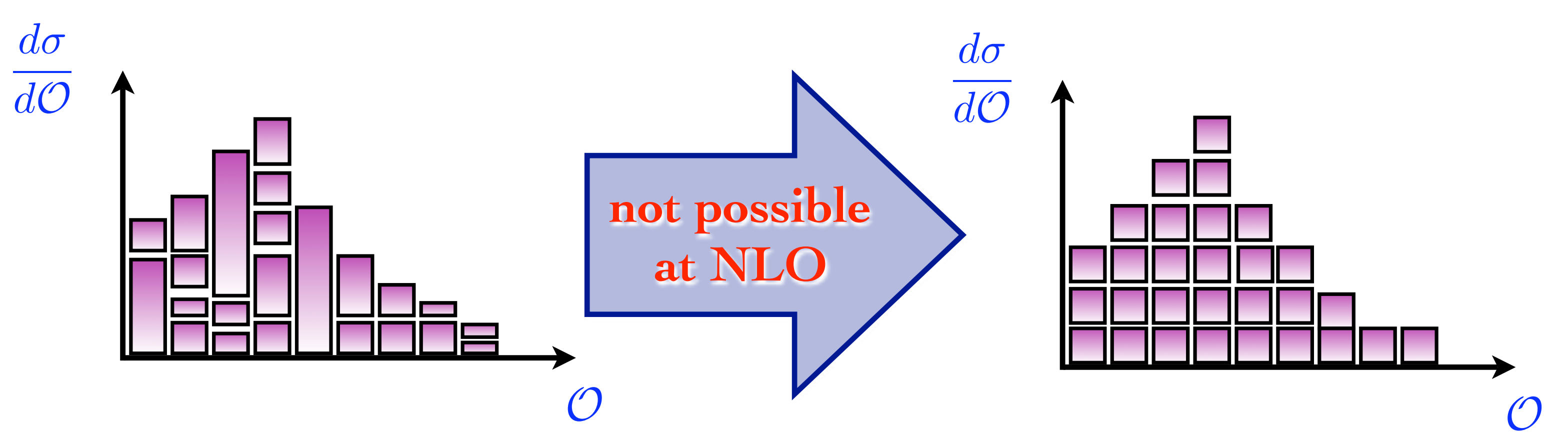
- Another consequence of this kinematic mismatch is that we cannot generate events at fixed order NLO
  - Even though the integrals are finite, they are not bounded (compare with $\int_0^1 dx \frac{1}{\sqrt{x}}$), so there is no maximum to unweight against: a single event can have an arbitrarily large weight!
  - Furthermore, event and counter event have different kinematics: which one to use for the unweighted event?



Event unweighting?

- Another consequence of this kinematic mismatch is that we cannot generate events at fixed order NLO
    - Even though the integrals are finite, they are not bounded (compare with $\int_0^1 dx\,\frac{1}{\sqrt{x}}$), so there is no maximum to unweight against: a single event can have an arbitrarily large weight!
    - Furthermore, event and counter event have different kinematics: which one to use for the unweighted event?





NLO summary

- Both the virtual and real-emission corrections are IR divergent, but their sum is finite: We can use a subtraction methods to factor the divergences in the real-emission phase-space integration and cancel them explicitly against the terms in the virtual corrections
- This generates events and counter events with slightly different kinematics. This means we cannot generate unweighed events (integrals are not bounded), but we can fill plots with weighted events: MC integrator (not an MC event generator)
- When making plots or applying cuts, use only IR safe observables with finite resolution
- NNLO is very similar, but at least a couple of orders of magnitude more intricate: double loop integrals, and a rather more involved IR structure.