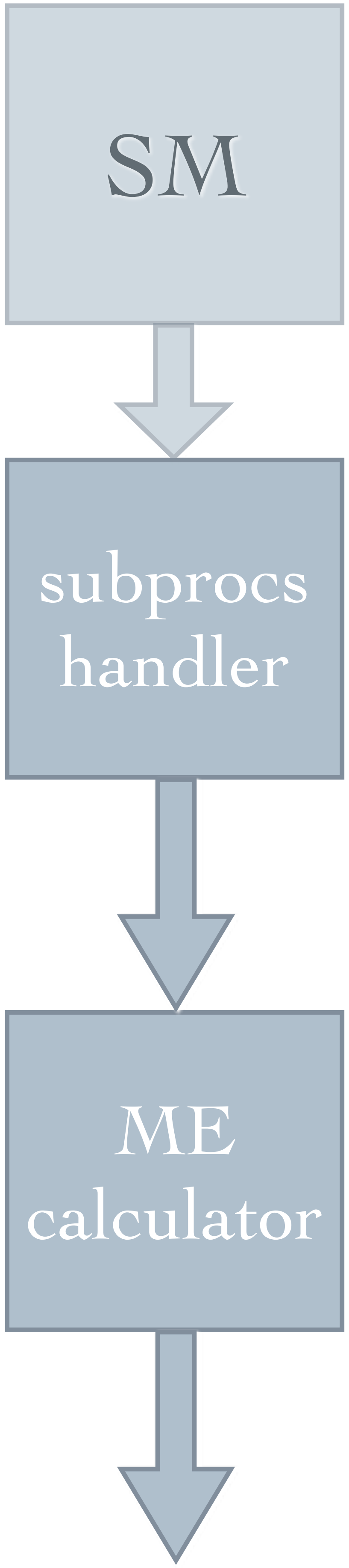
Introduction to Monte Carlo generators: Parton Showers

Rikkert Frederix
Lund University

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

LUNDS UNIVERSITET

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.

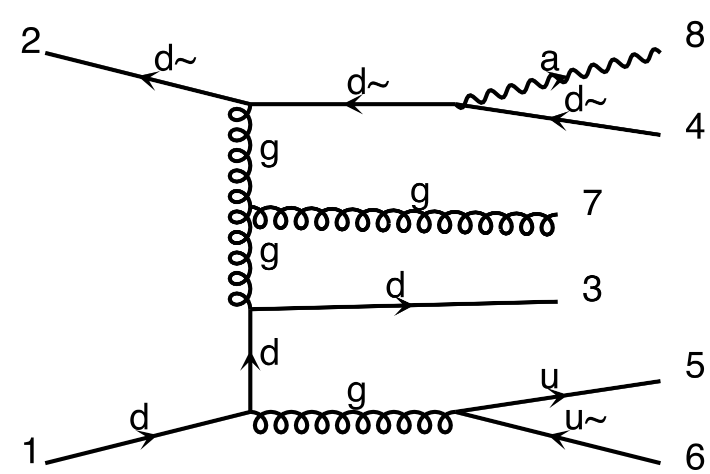
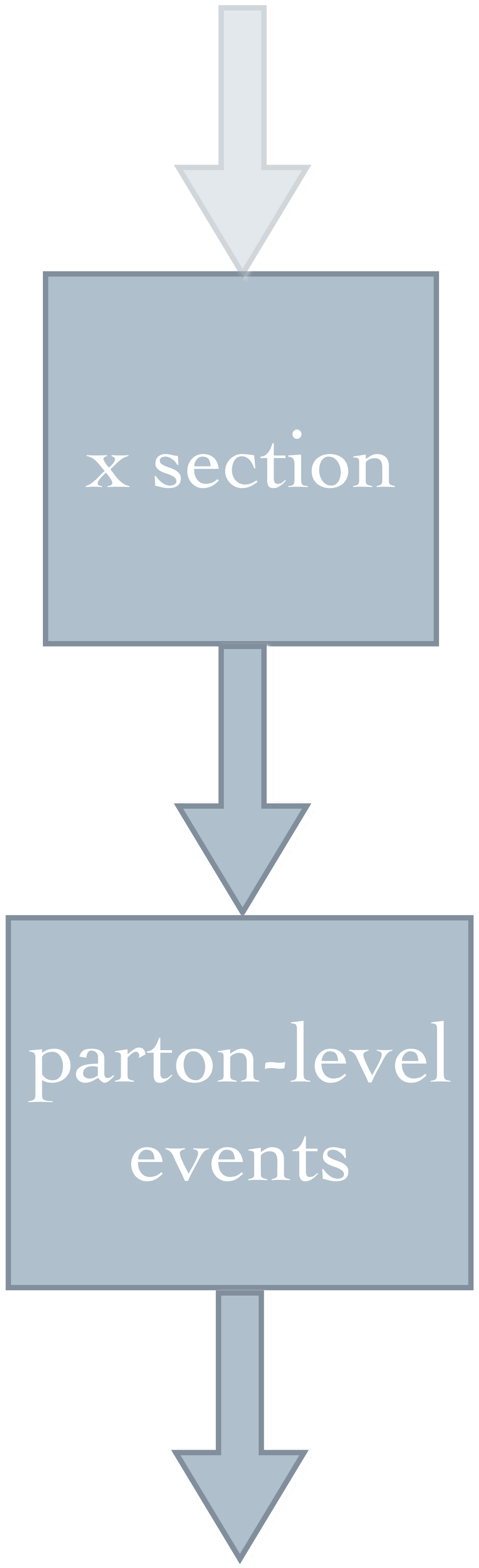
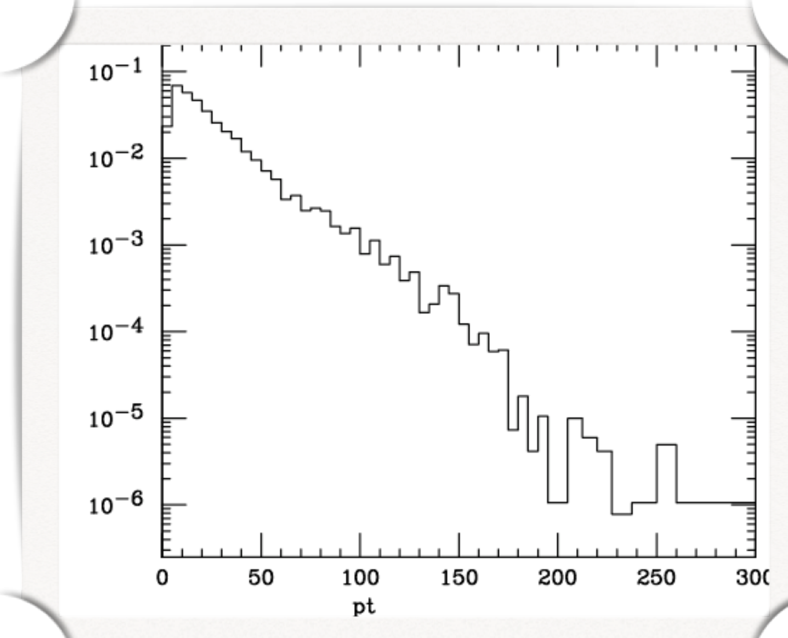


diagram 949 QCD=5, QED=1

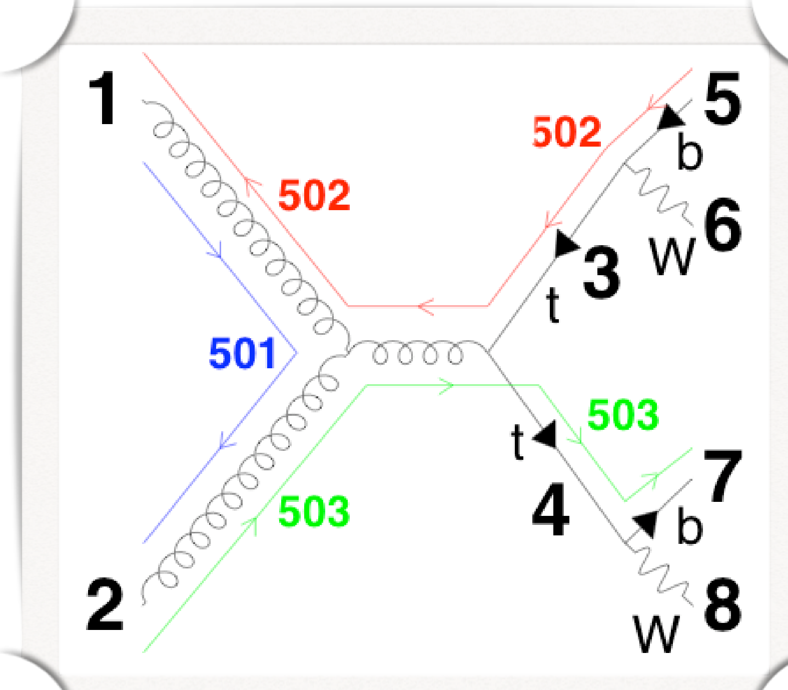
ME generators: general structure



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

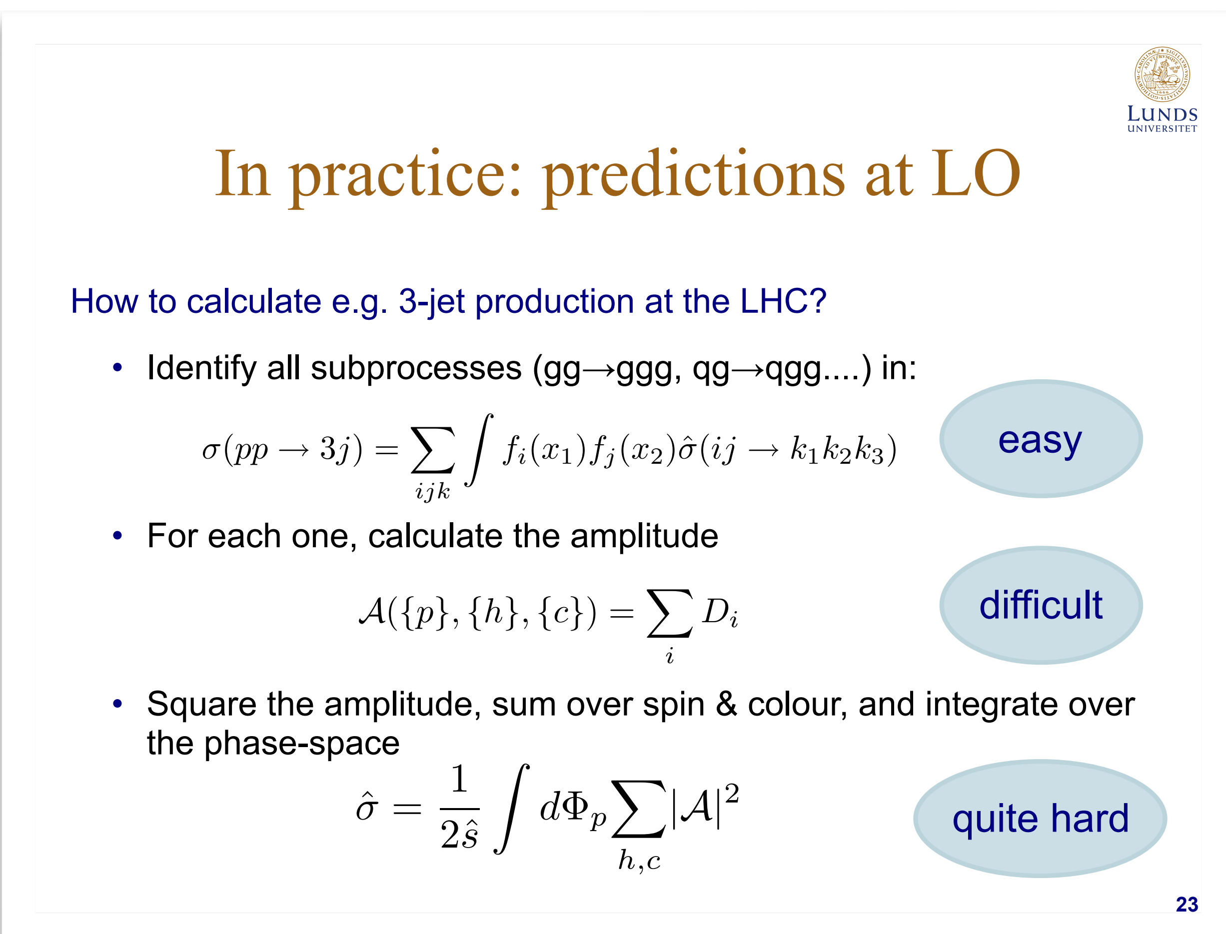


Events are obtained by unweighting.
These are at the parton-level.
Information on particle id, momenta, spin, color is given in the Les Houches Event (LHE) File format.



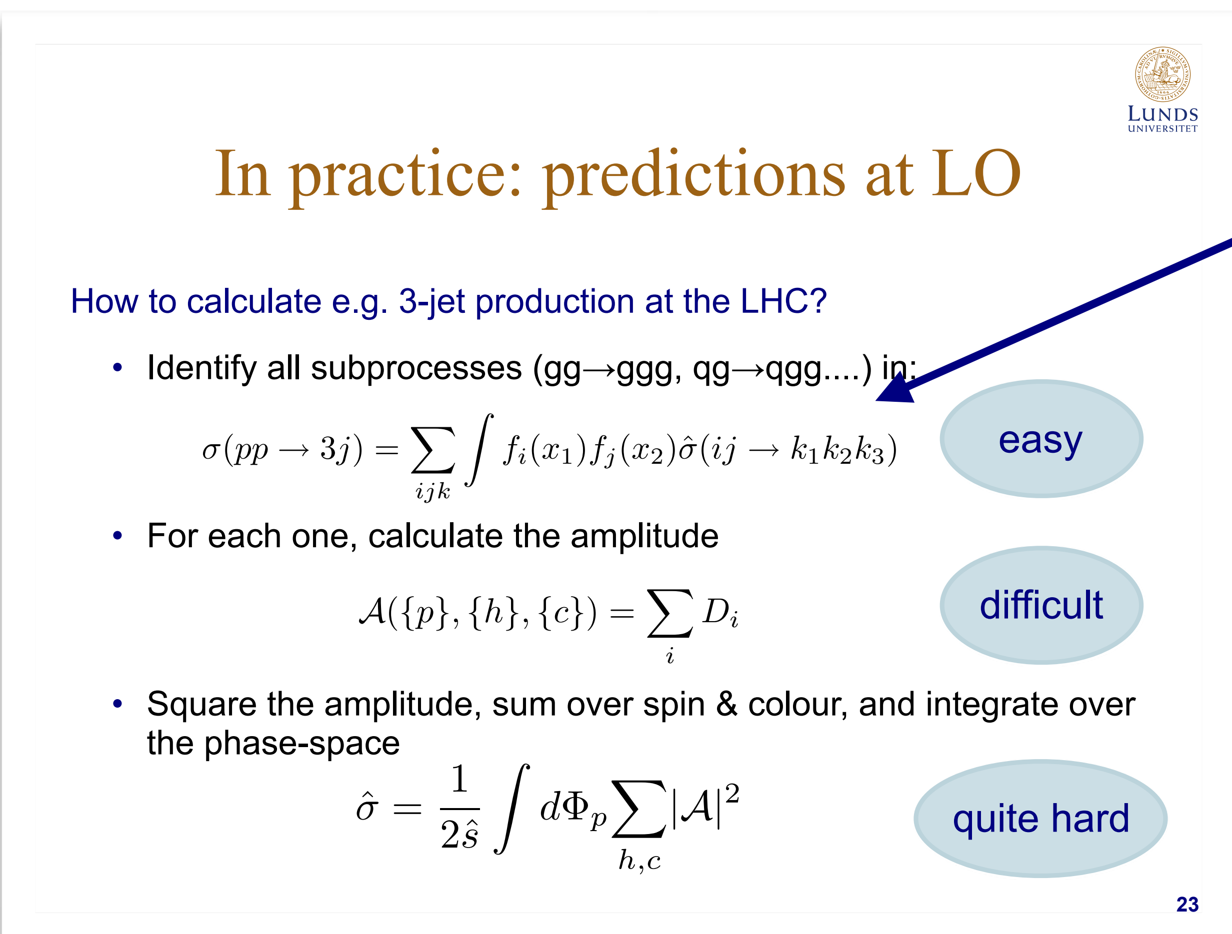
What about higher orders?

- 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)



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

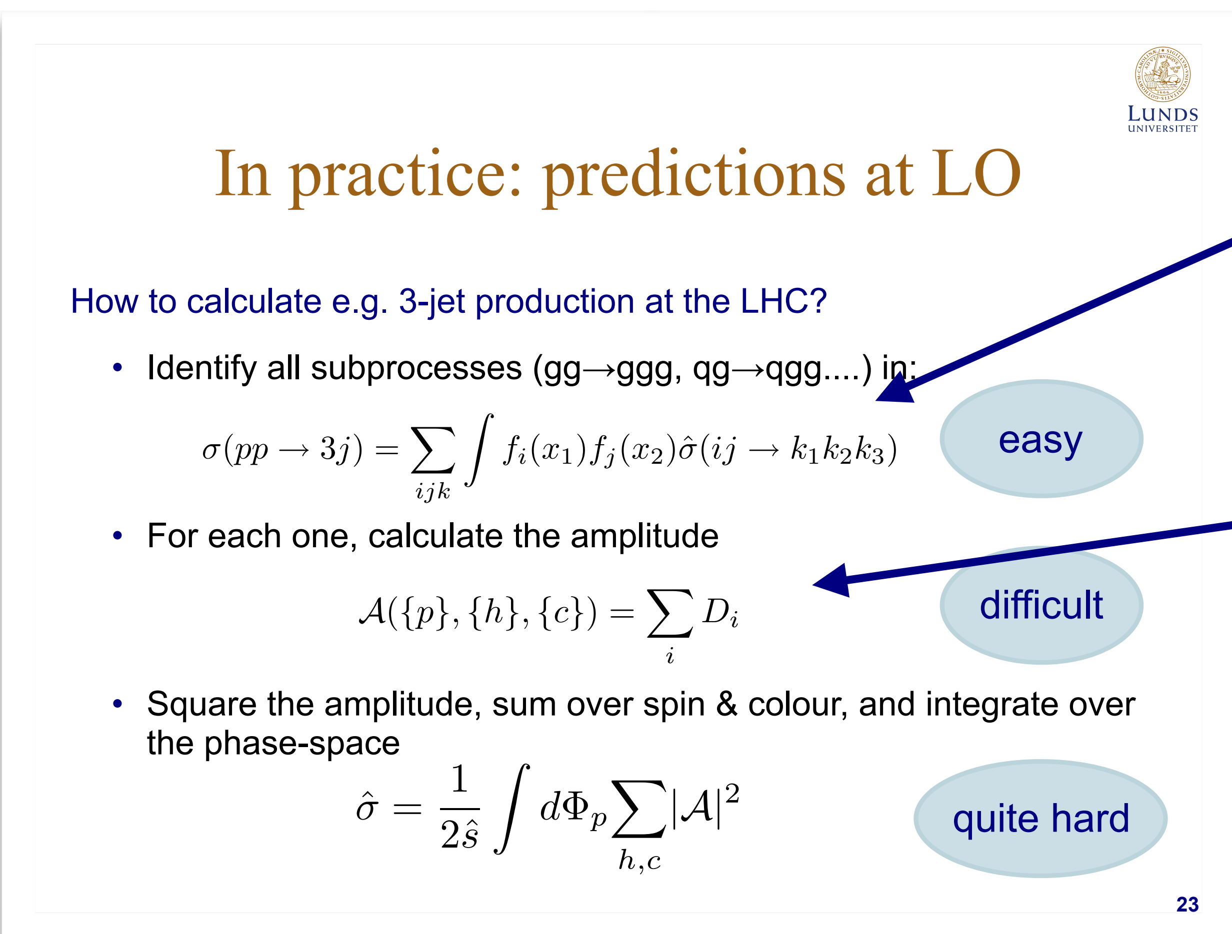
23

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

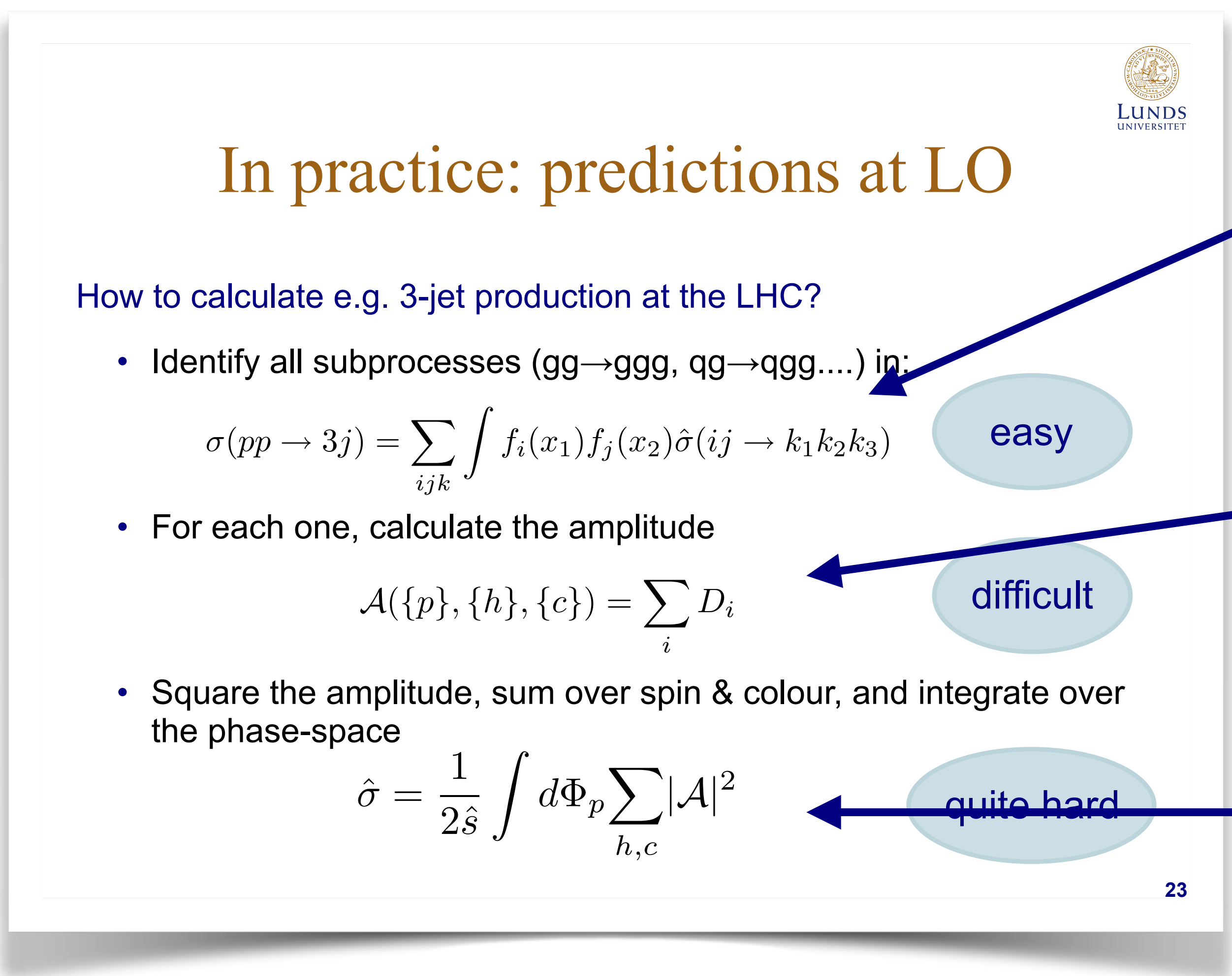
- need also subprocesses with one more parton

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,

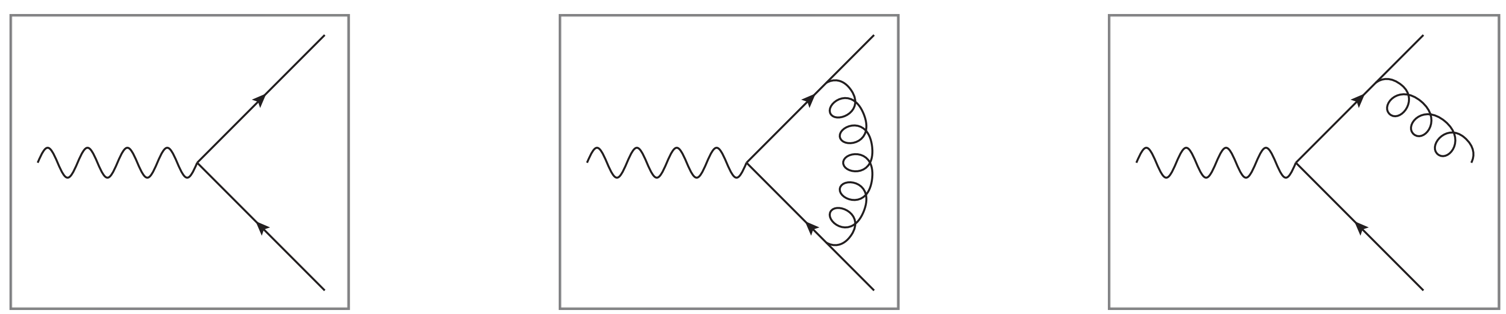
- need also to cancel divergencies

NLO: how to?

- Three ingredients need to be computed at NLO

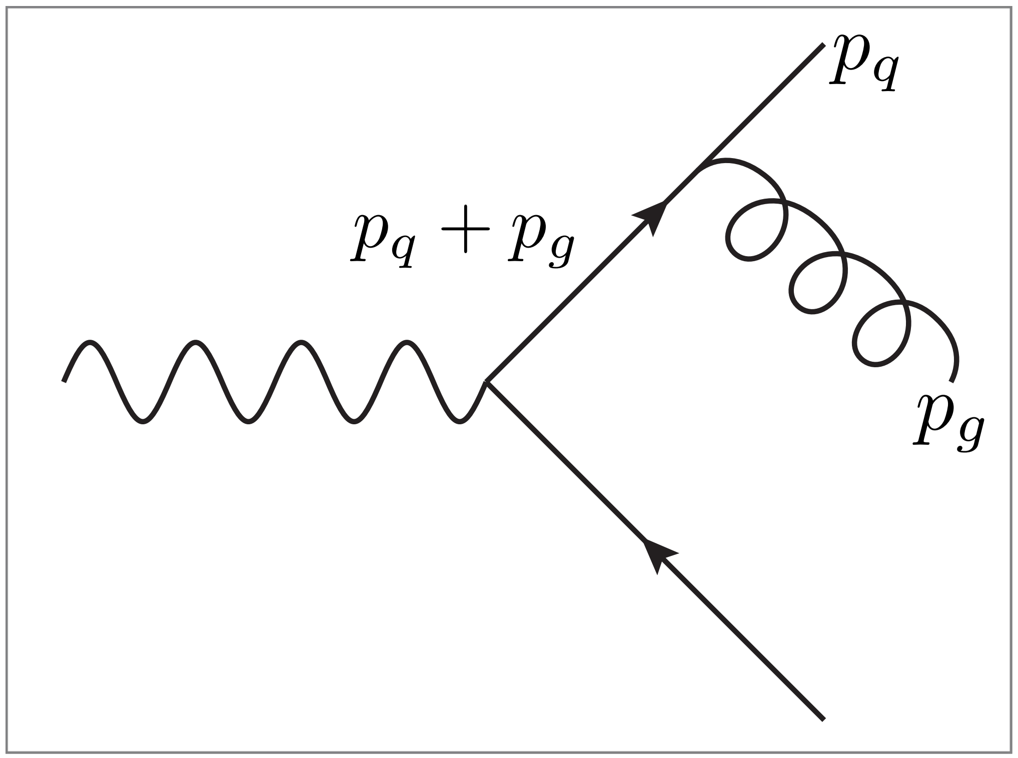
$$\sigma_{NLO} = \int_n \alpha_s^b d\sigma_0 + \int_n \alpha_s^{b+1} d\sigma_V + \int_{n+1} \alpha_s^{b+1} d\sigma_R$$

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
$$\int_0^1 f(x)\,\mathrm{d}x, \text{ with } f(x) = \frac{g(x)}{x} \text{ and } g(x) \text{ a regular function}$$

- 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

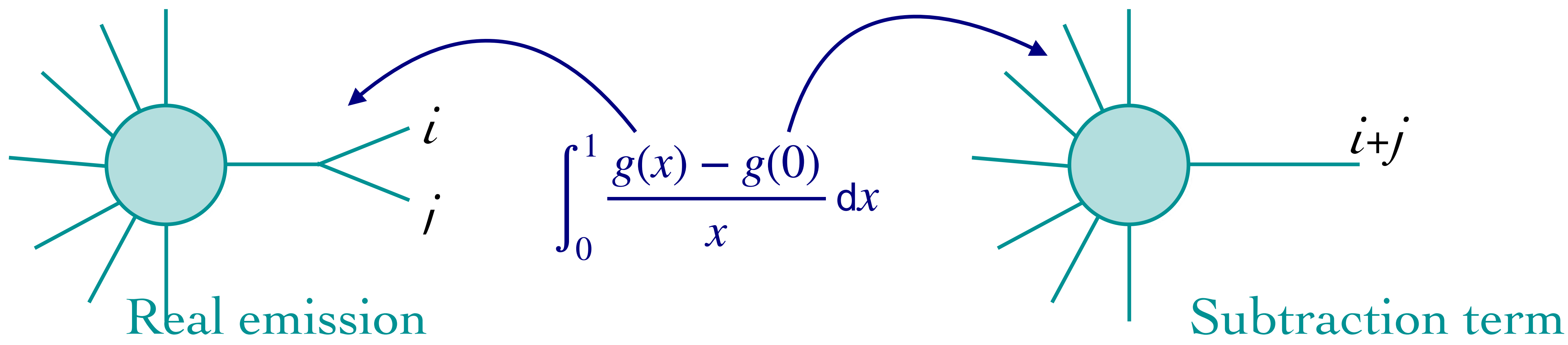
Subtraction method

- 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

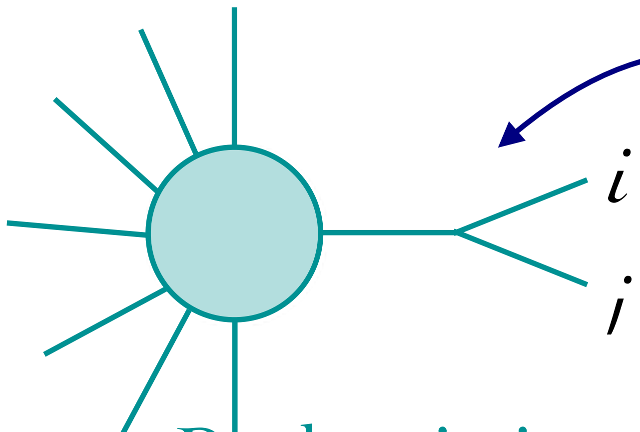
NLO: kinematics of subtraction terms



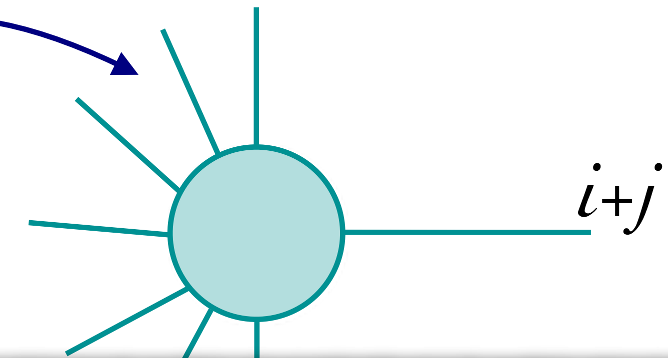
- Real emission and subtraction term cannot be separated (individually, they are divergent!)
- i and j are on-shell in the real emission, but $i + j$ is not: $x \sim m_{i+j}^2$

  $i + j$ 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

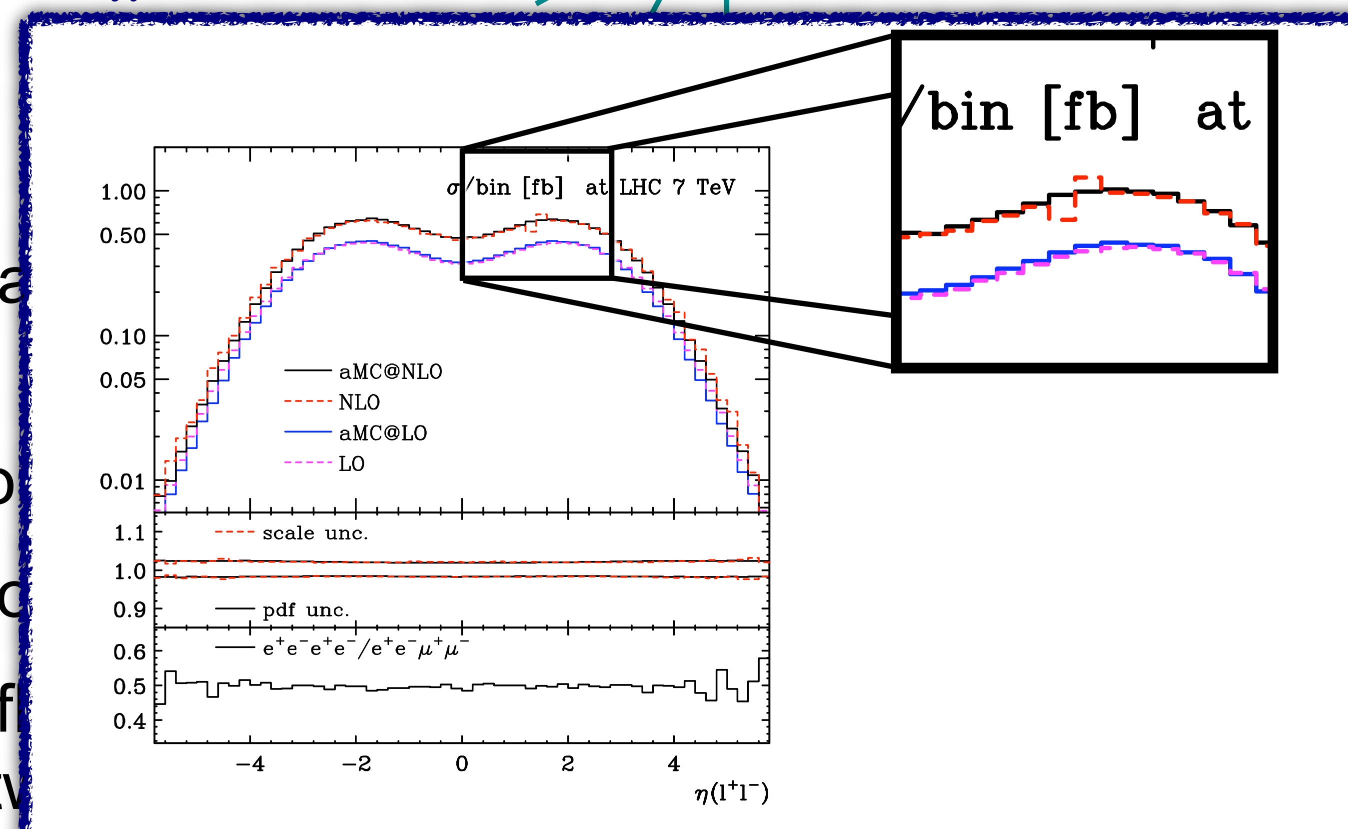


$$\int_0^1 \frac{g(x) - g(0)}{x}\,\mathrm{d}x$$



- 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 subtractio
  - This is not possible without reshuff 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?

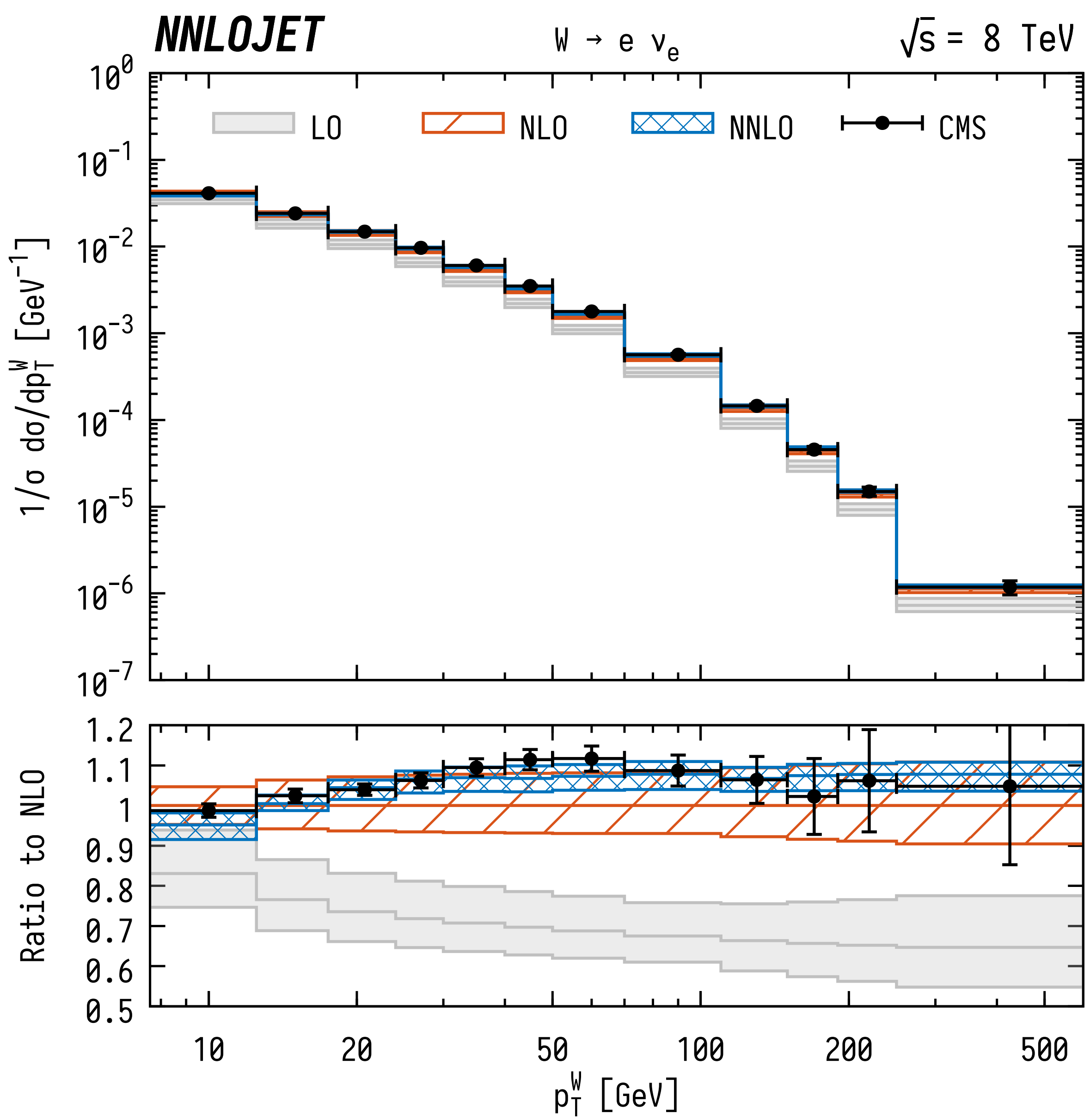
- Another consequence of the kinematic mismatch is that we cannot generate unweighted events at NLO
  - $n+1$-body contribution and n-body contribution are not bounded from above → unweighting not possible
  - Further ambiguity on which kinematics to use for the unweighted events

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)

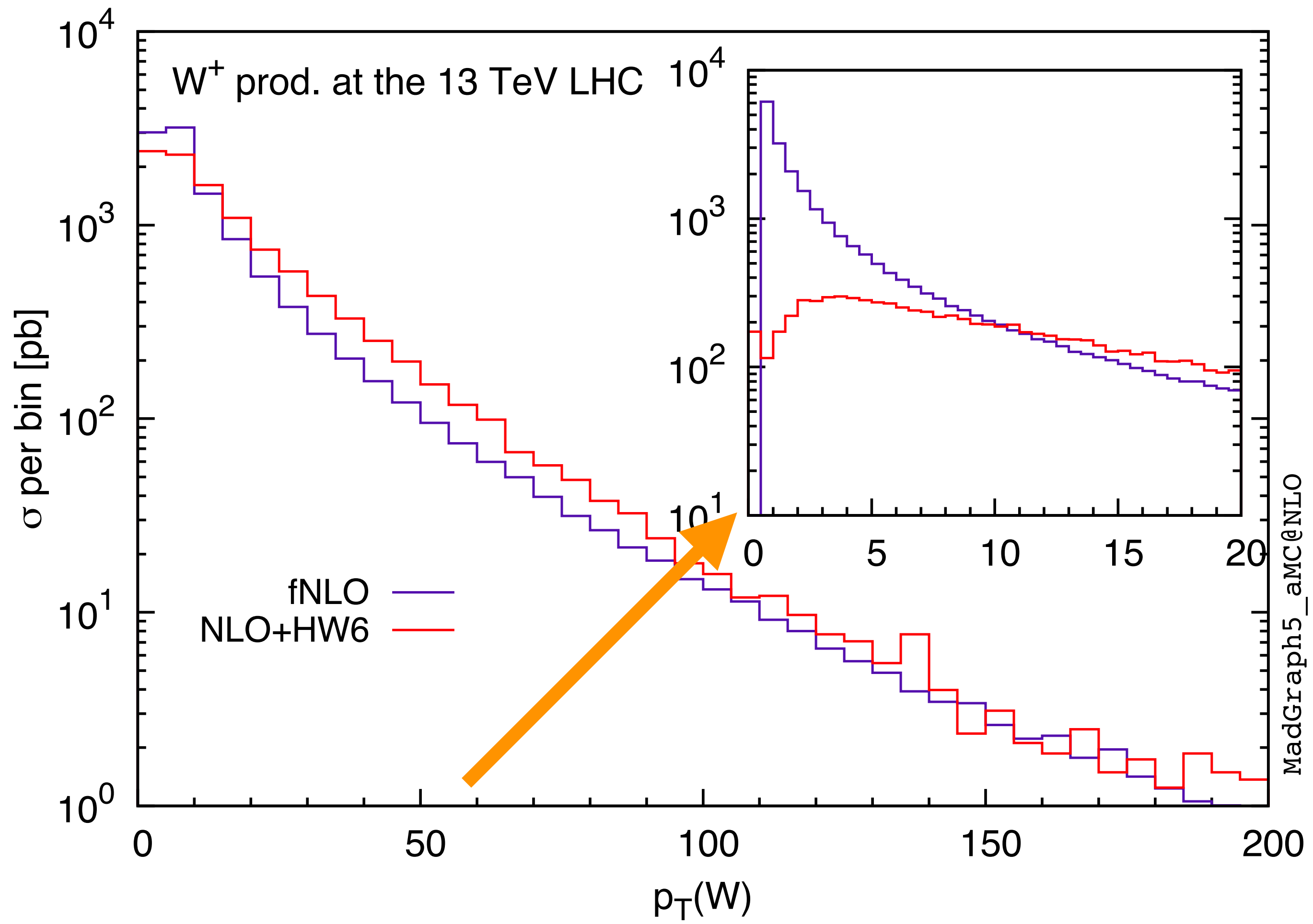
Example: W+j production

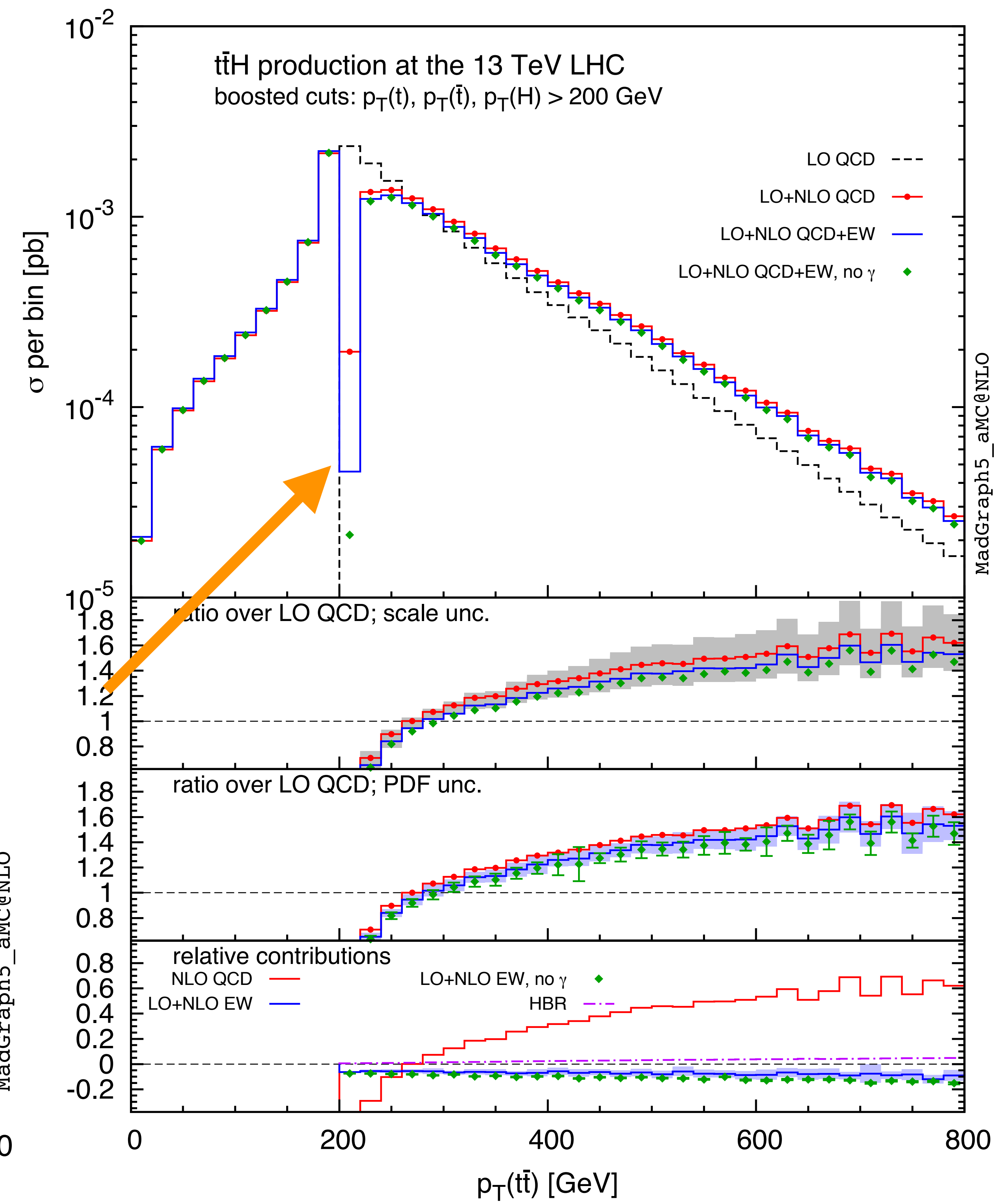


- Both NLO and NNLO agree with the CMS data (8 TeV collisions),
  - NNLO has significantly smaller uncertainties
- LO uncertainties underestimated
  - In general: NLO accuracy required to describe LHC data

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

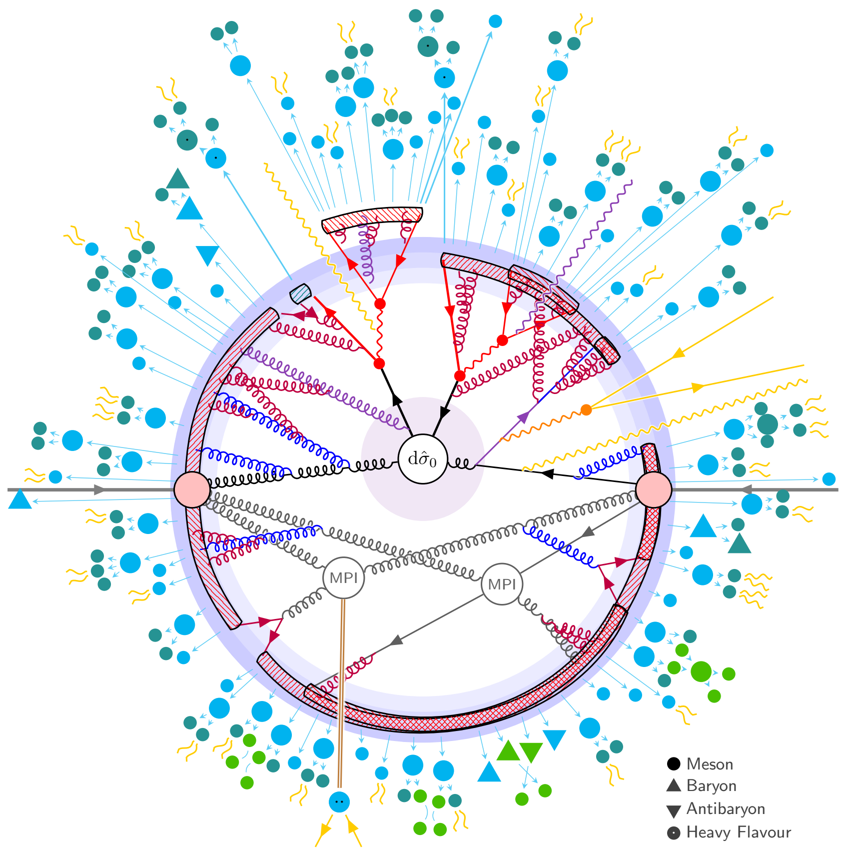




Summary: the hard interaction

- 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

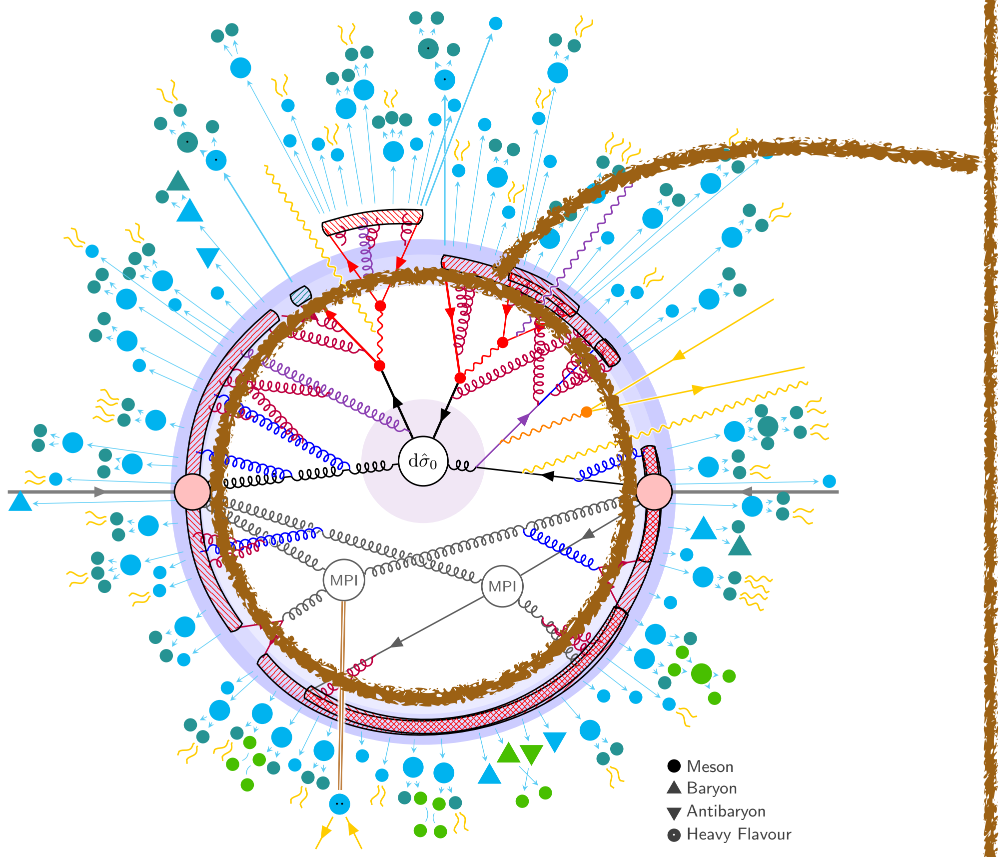


- Only discussed the central part of the collision.
- Sometimes this is enough!
  - No matching to parton shower
  - Easy to go beyond LO
  - Analytic resummation (instead of resummation with PS also a way forward, and possibly higher accuracy)

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

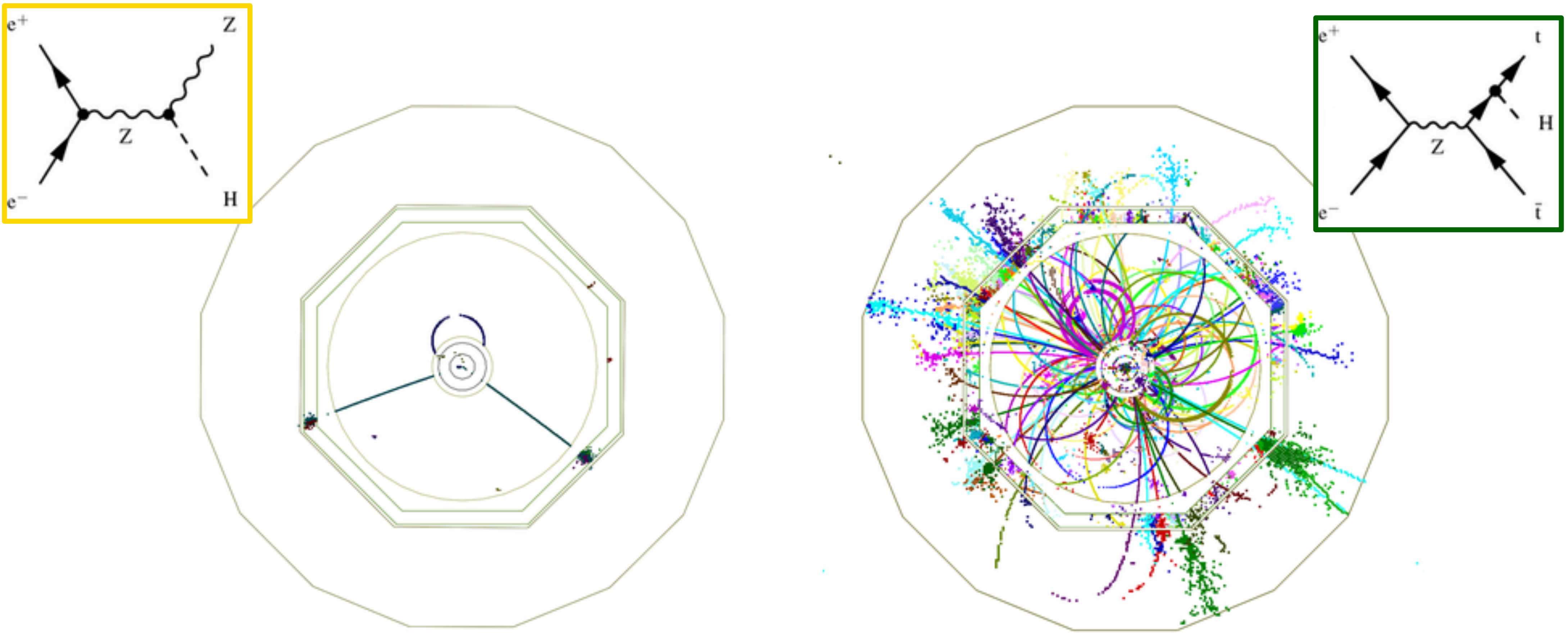
Phase-space integral — Parton density functions — Parton-level cross section

The Parton Shower



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

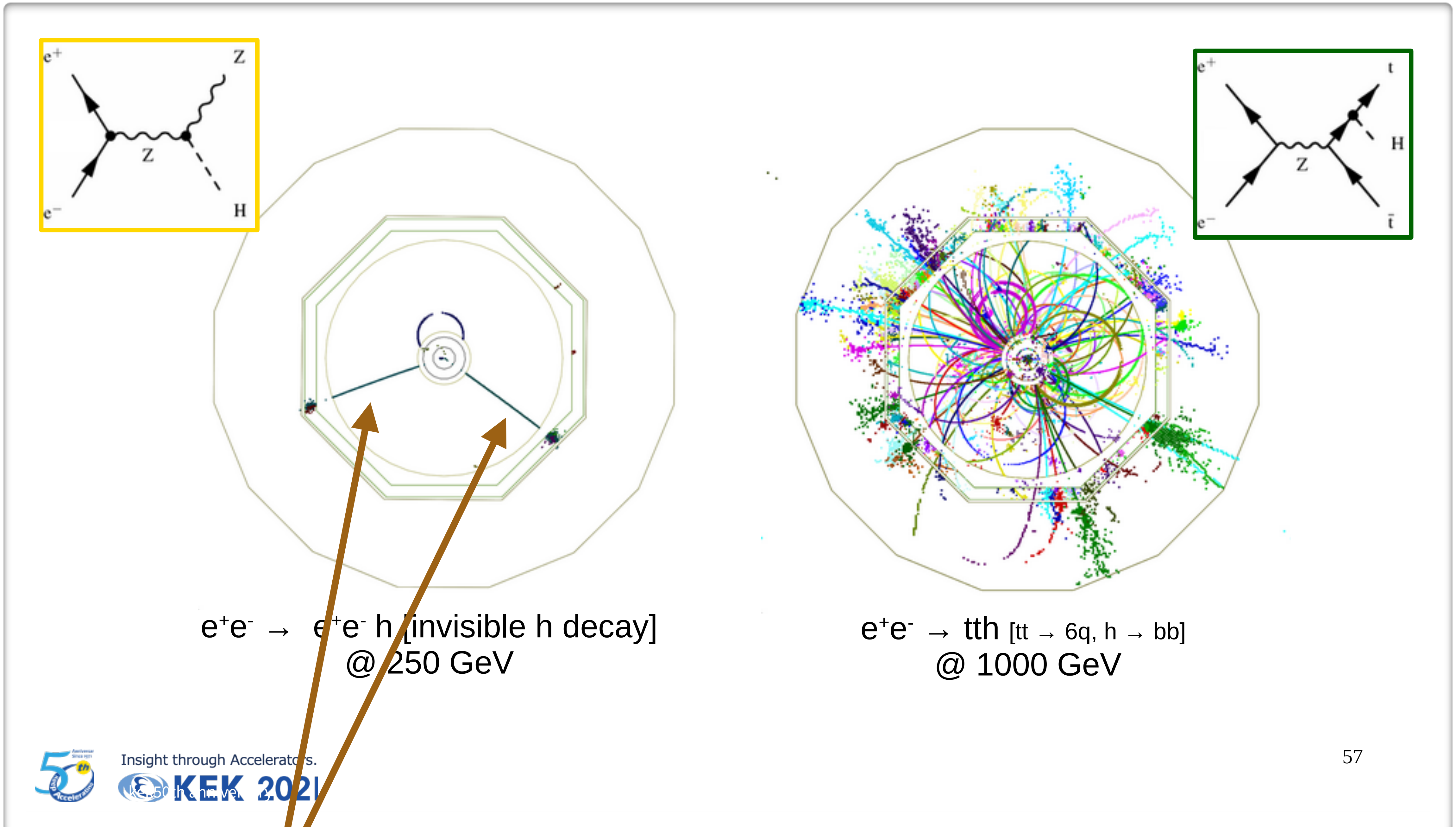
Slide adapted from Daniel Jeans (ICS, 2024)



$e^+e^- \rightarrow e^+e^- h$ [invisible h decay] @ 250 GeV

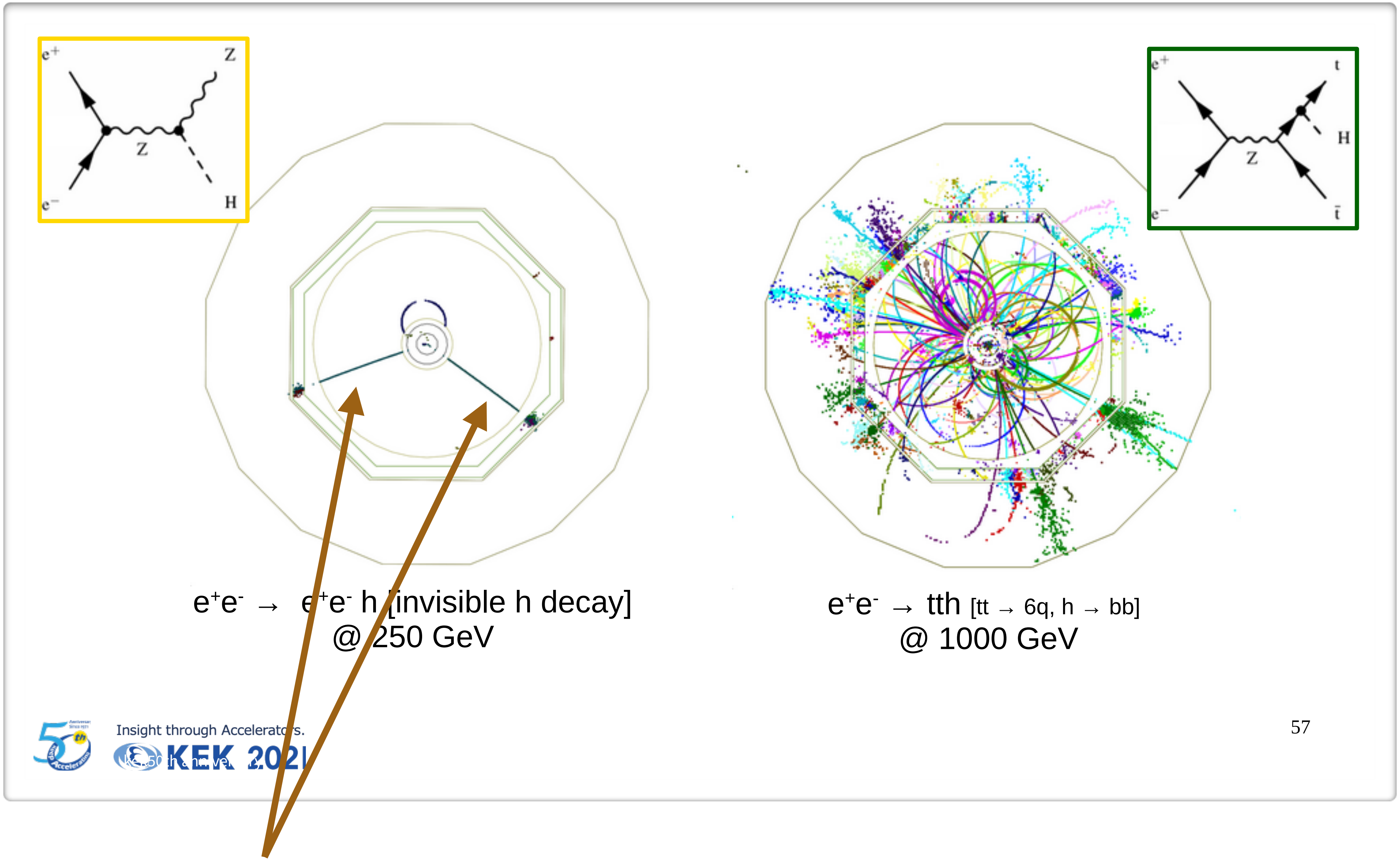
$e^+e^- \rightarrow tth$ [$tt \rightarrow 6q$, $h \rightarrow bb$] @ 1000 GeV

Slide adapted from Daniel Jeans (ICS, 2024)



Fixed-order calculation yields the two final state leptons (and the invisible Higgs boson decay products). **They corresponds directly two the two observed particles**

Slide adapted from Daniel Jeans (ICS, 2024)

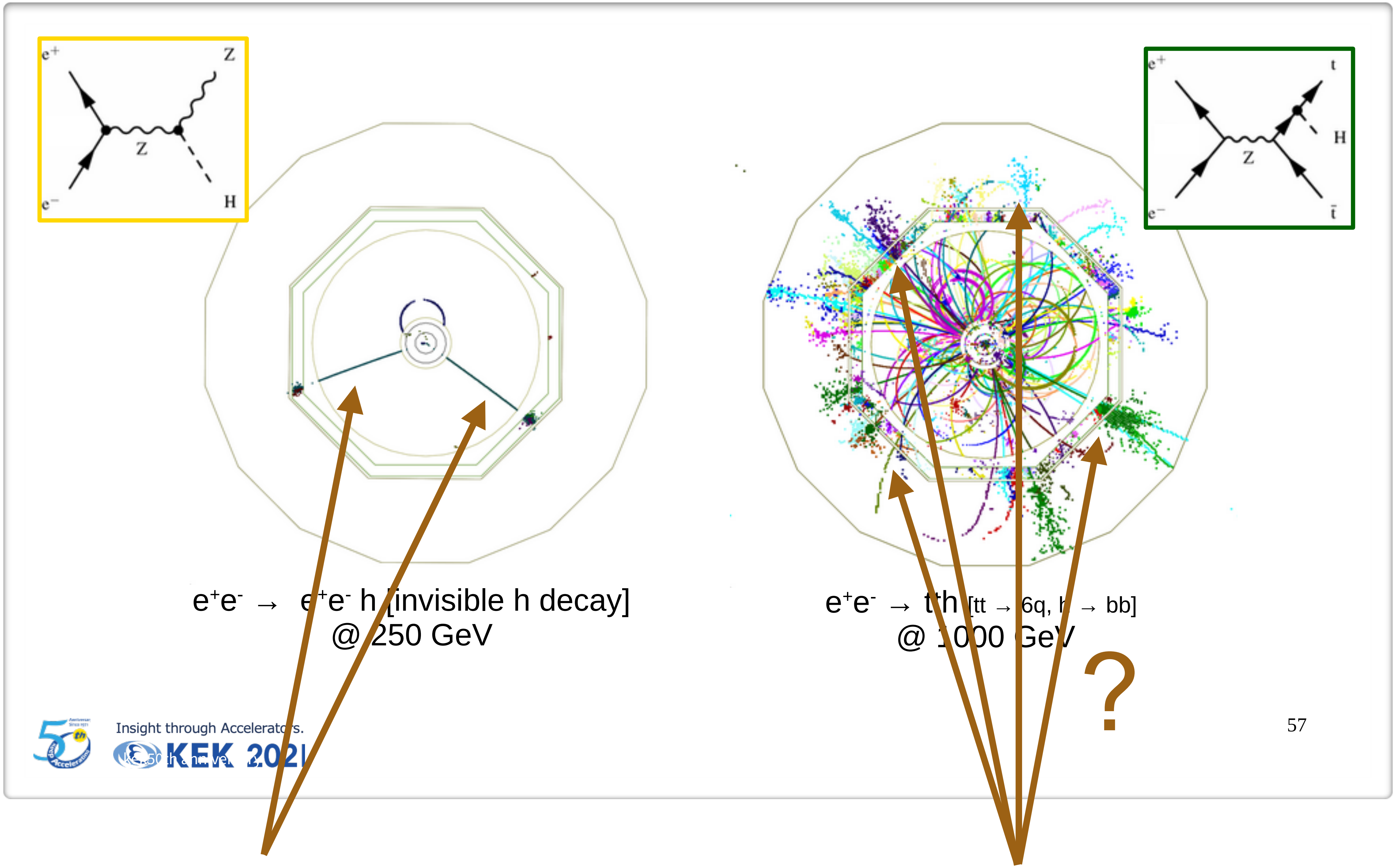


Fixed-order calculation yields the two final state leptons (and the invisible Higgs boson decay products). **They corresponds directly two the two observed particles**

Fixed-order calculation generates the 8 final state quarks (of which 4 are (anti-)bottom quarks).
How does this corresponds to the observed particles?

Slide adapted from Daniel Jeans (ICS, 2024)

LUNDS UNIVERSITET



Fixed-order calculation yields the two final state leptons (and the invisible Higgs boson decay products). **They corresponds directly two the two observed particles**

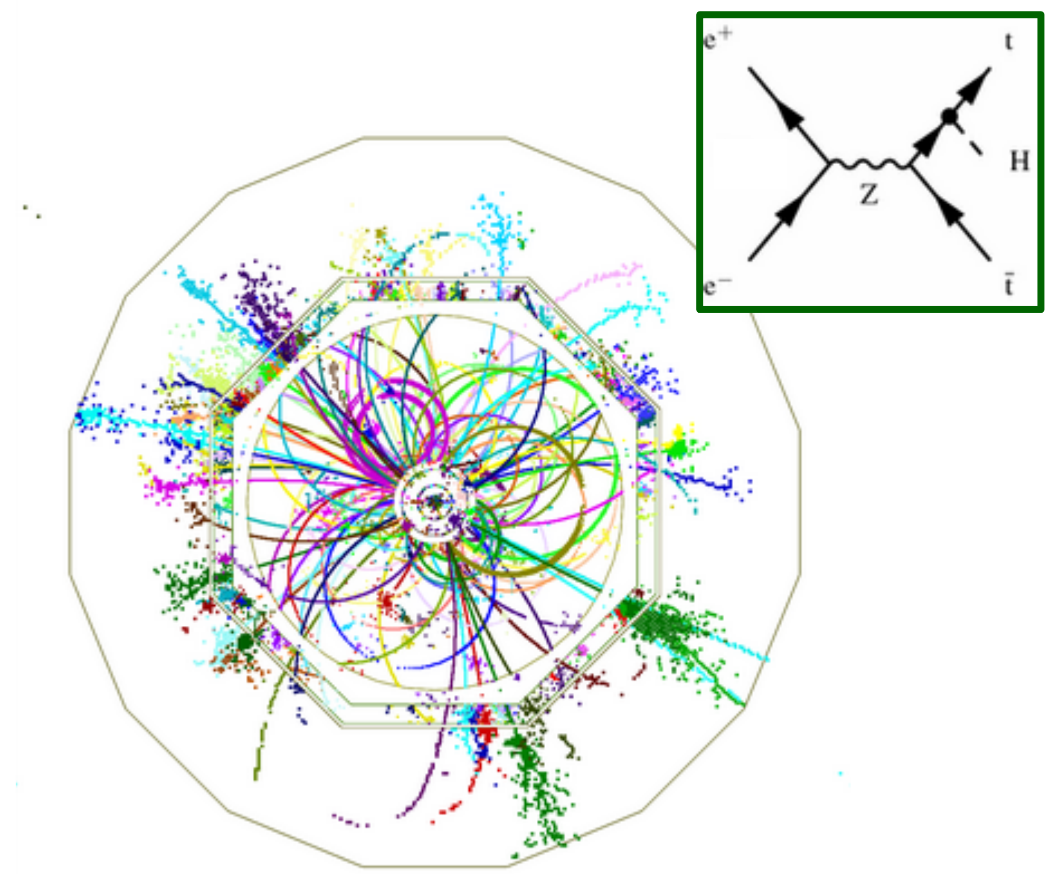
Fixed-order calculation generates the 8 final state quarks (of which 4 are (anti-)bottom quarks).
How does this corresponds to the observed particles?

Two possibilities

- Parton showering!

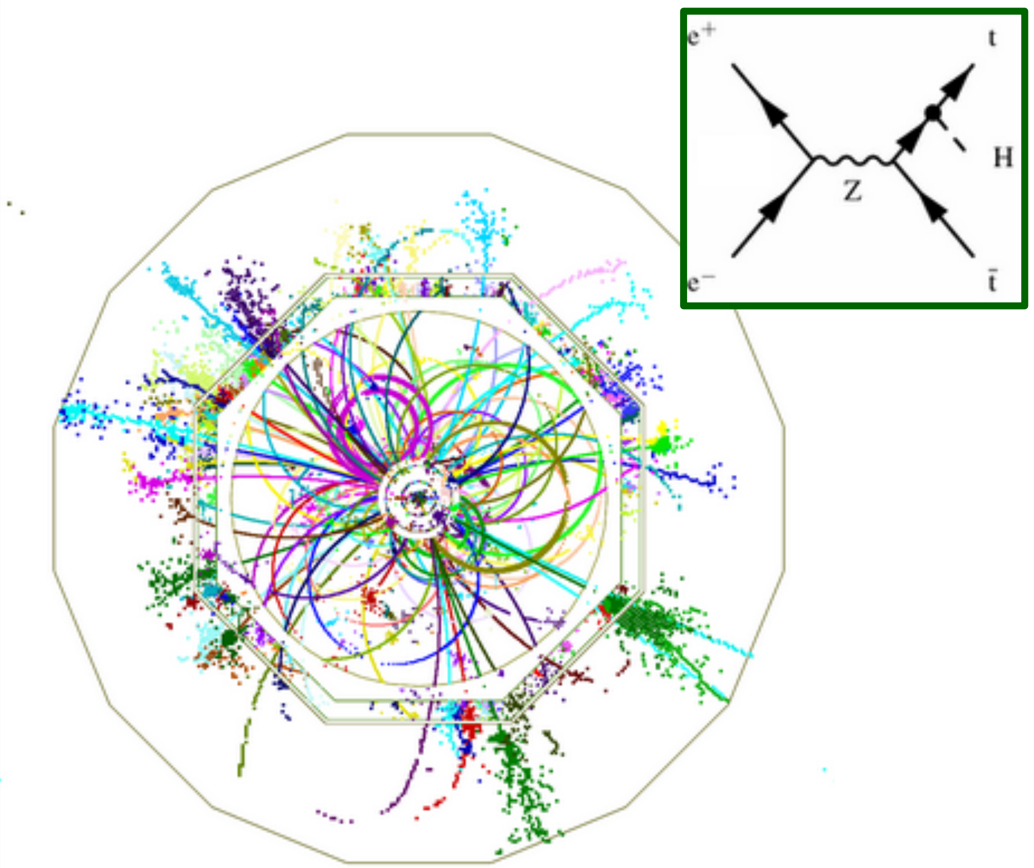
or

- Jet clustering!



$e^+e^- \rightarrow t\bar{t}h$ [$t\bar{t} \rightarrow 6q$, $h \rightarrow b\bar{b}$]
@ 1000 GeV

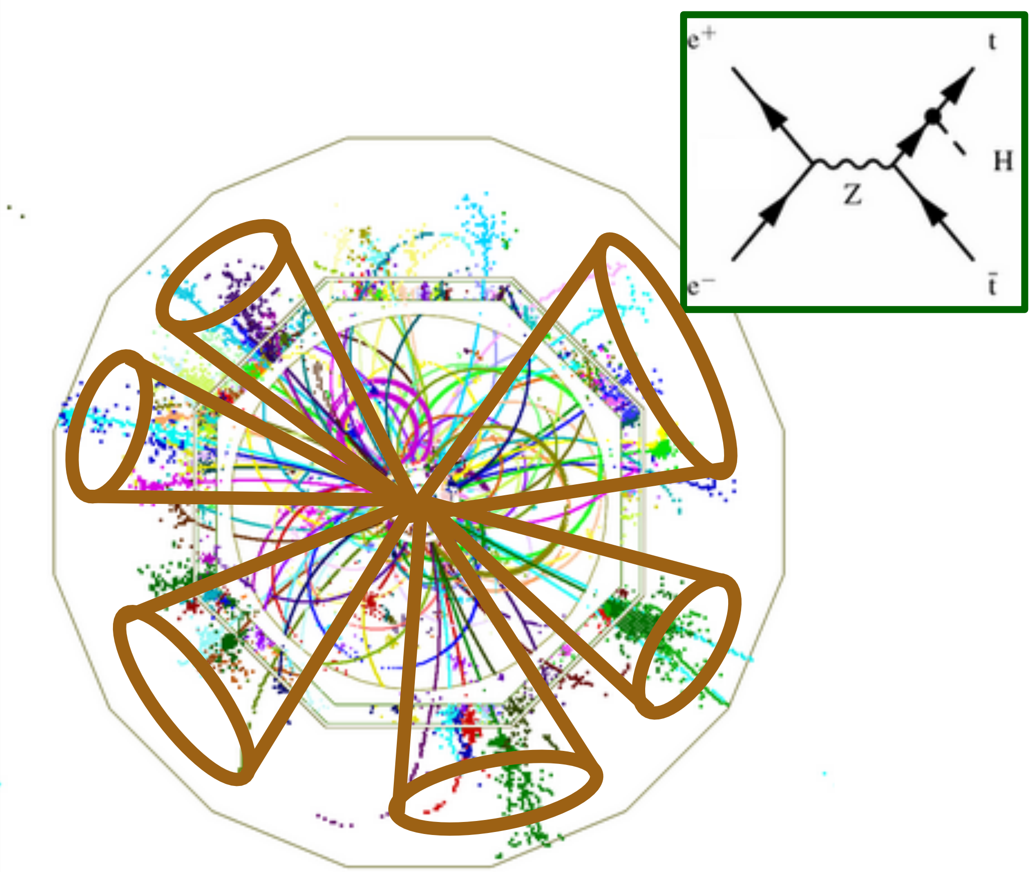
Jet clustering



$e^+e^- \rightarrow tth$ [tt → 6q, h → bb]
@ 1000 GeV

- Goal:
  - Cluster particles "that are close in phase-space" into single objects: jets
  - These jets correspond to the quarks (or gluons) generated at fixed order

Jet clustering



$e^+e^- \rightarrow$ tth [tt → 6q, h → bb]
@ 1000 GeV

- Goal:
  - Cluster particles "that are close in phase-space" into single objects: jets
  - These jets correspond to the quarks (or gluons) generated at fixed order

Jet clustering

- Sequential algorithm
  - Define all the distances
    - d_{ij} (between particles i and j) and
    - d_{iB} (between particle i and the beam)
  - If d_{ij} is the smallest, replace particles i and j by a new (pseudo) particle
  - If d_{iB} is the smallest, call particle i a jet and remove it from the list
  - Keep going until no particles are left

Jet clustering

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2} \,,$$
$$d_{iB} = k_{ti}^{2p} \,,$$

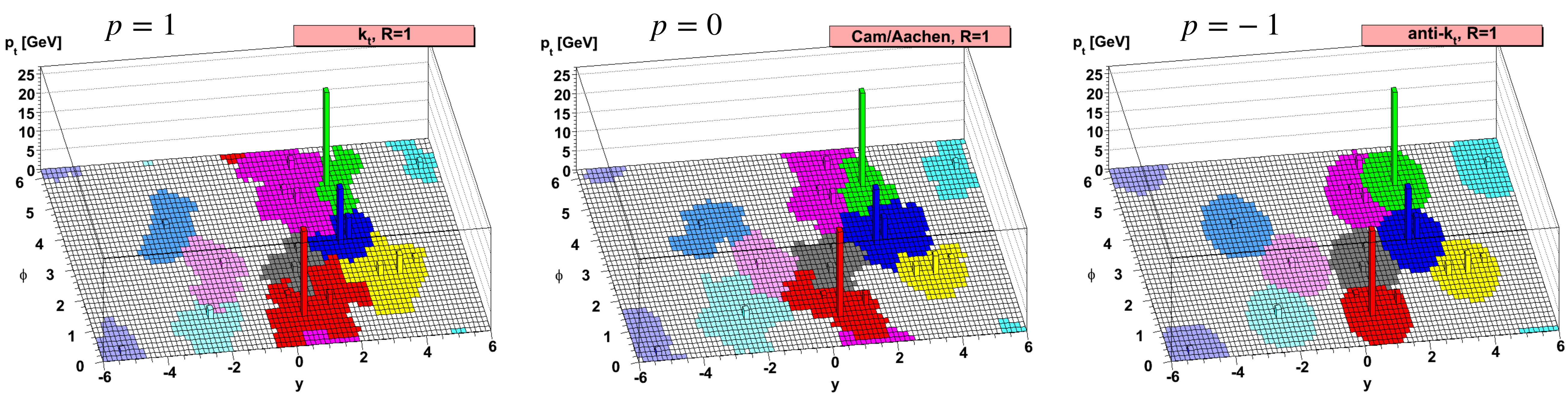
- Sequential algorithm
  - Define all the distances
    - d_{ij} (between particles i and j) and
    - d_{iB} (between particle i and the beam)
  - If d_{ij} is the smallest, replace particles i and j by a new (pseudo) particle
  - If d_{iB} is the smallest, call particle i a jet and remove it from the list
  - Keep going until no particles are left

LUNDS UNIVERSITET

Jet algorithms

- Different clustering algorithms exist
- Same event clustered with different algorithms gives slightly different jet(s) and shapes

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2},$$
$$d_{iB} = k_{ti}^{2p},$$



- Anti-k$_T$ ($p = -1$) algorithm most popular at the LHC
- All implemented in the fastjet package (https://fastjet.fr)

Be careful!

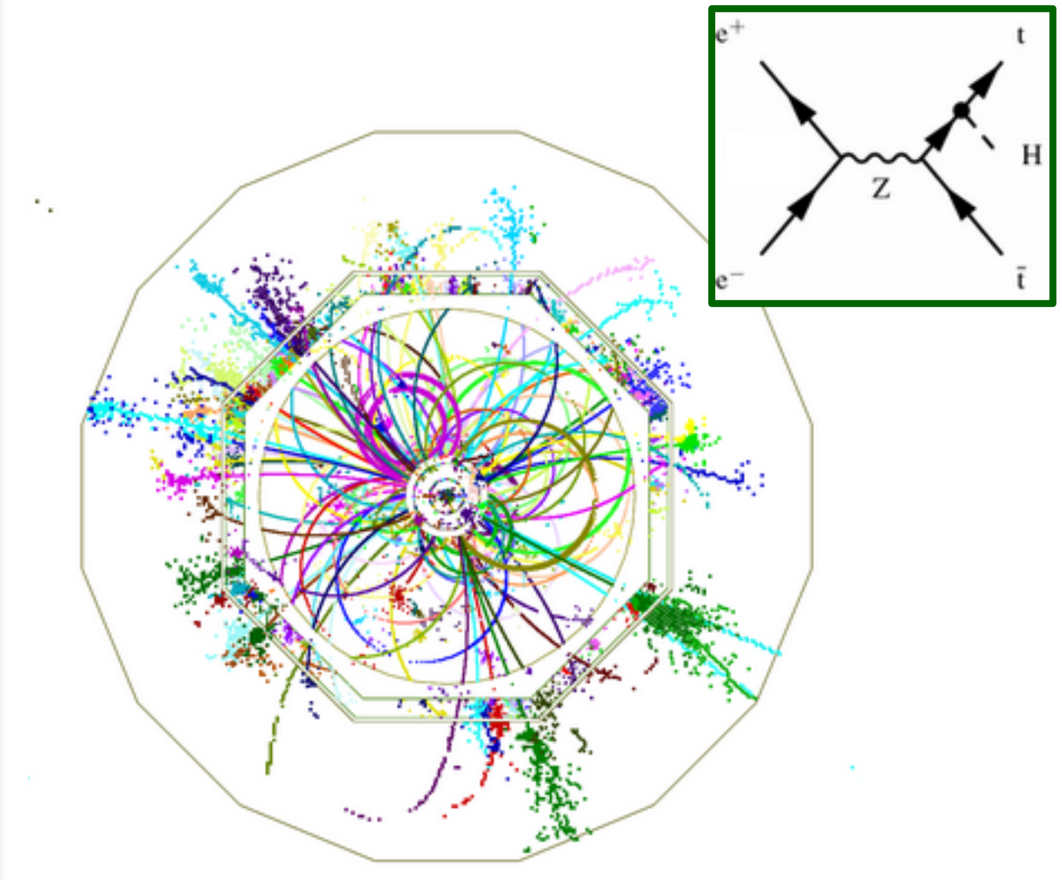
- The correspondence between quarks/gluons and jets works well
- However:
  - The jets that come out of the jet algorithm can be arbitrarily soft (i.e., with a very small energy or transverse momentum)
  - For them to correspond to quarks/gluons computed by by a matrix element event generator, they need to be "hard" and "well-separated"
    - Only consider jets above a threshold
    - But this is somewhat arbitrary…
      - How hard does "hard" need to be to be fine?
        - No general rule here… depends on the rest of the event!
        - In practice, in your calculation you get a large logarithms that hamper the convergence of perturbation theory (in the expansion of the strong coupling, each order is larger than the previous)

Two possibilities

- Parton showering!

  or

- Jet clustering!



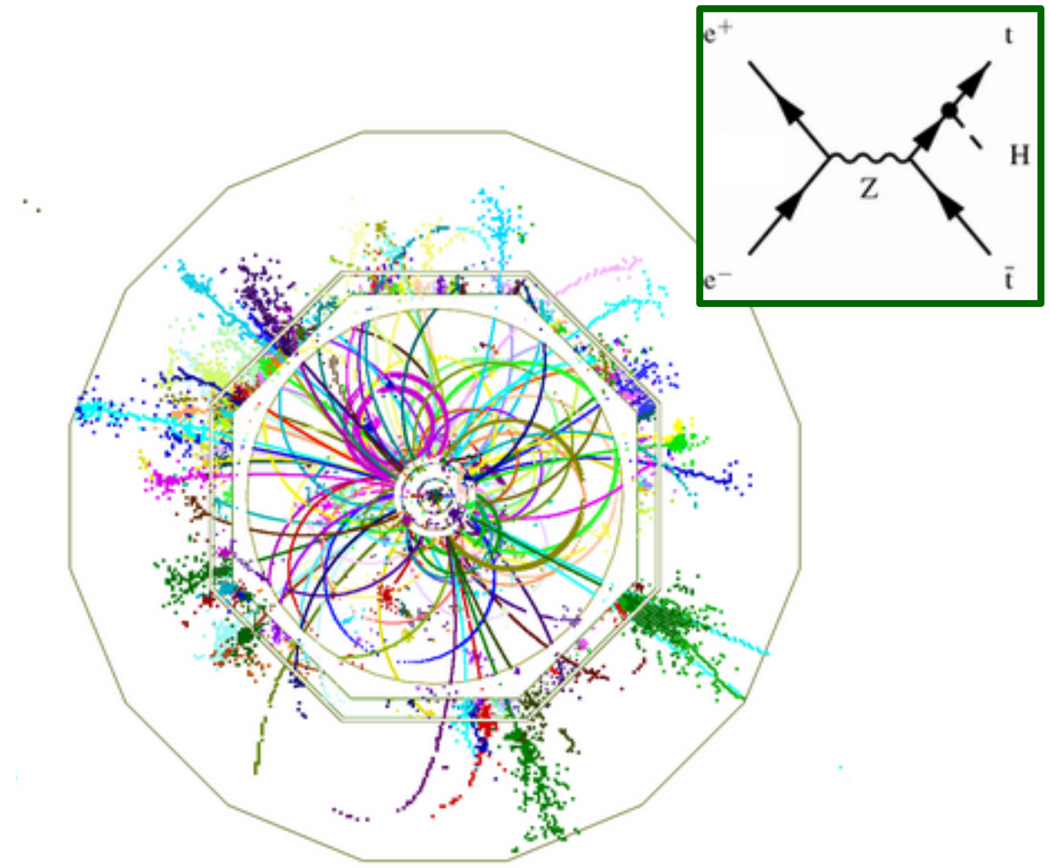
$e^+e^- \rightarrow t\bar{t}h$ [$t\bar{t} \rightarrow 6q$, $h \rightarrow b\bar{b}$]
@ 1000 GeV

Two possibilities

- Parton showering!

and

- Jet clustering!



$e^+e^- \rightarrow t\bar{t}h$ [$t\bar{t} \rightarrow 6q$, $h \rightarrow b\bar{b}$]
@ 1000 GeV

An LHC collision, factorised

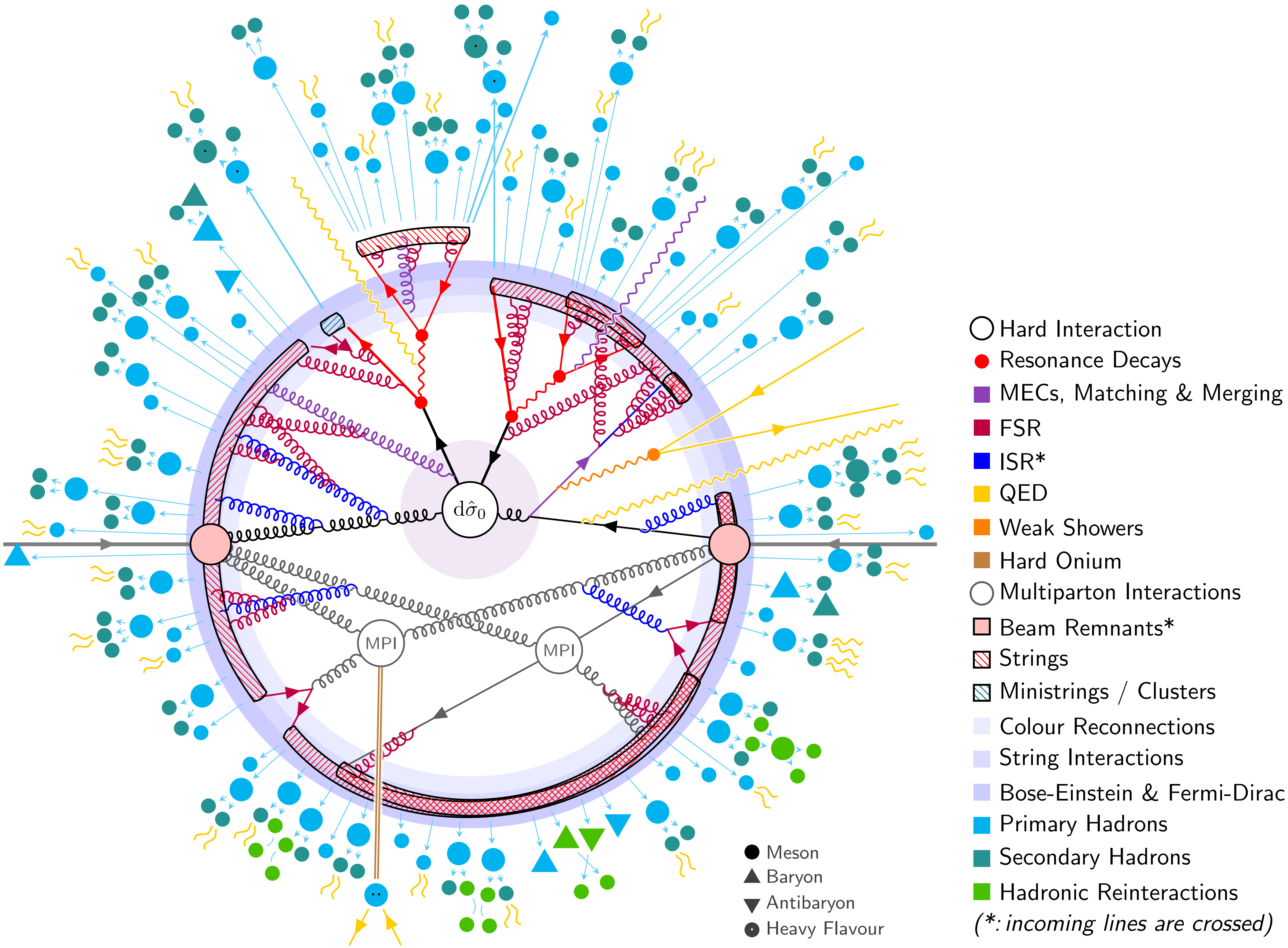
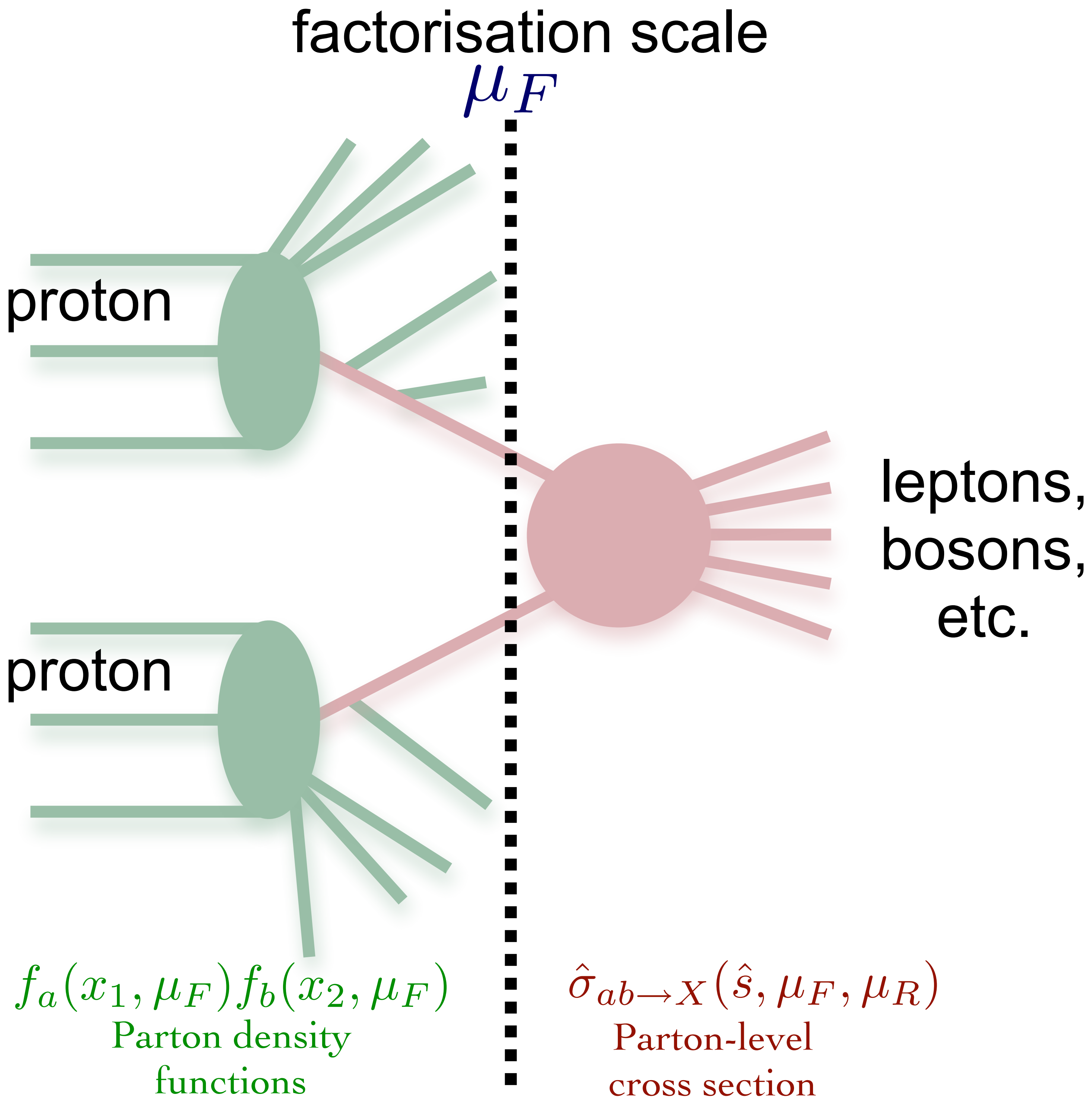


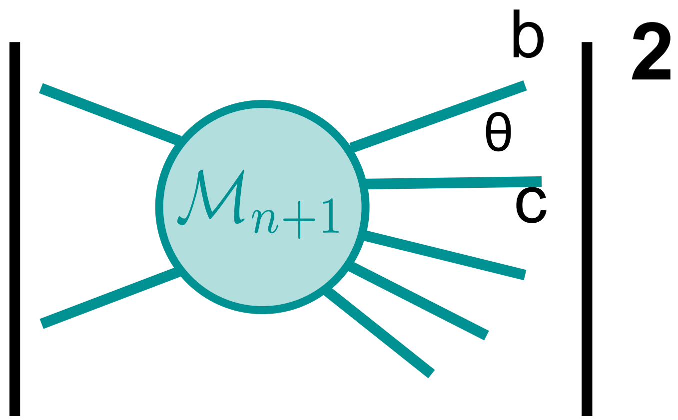
Figure taken from Bierlich et al., 2022 *(Pythia8.3 manual)*

Inclusiveness



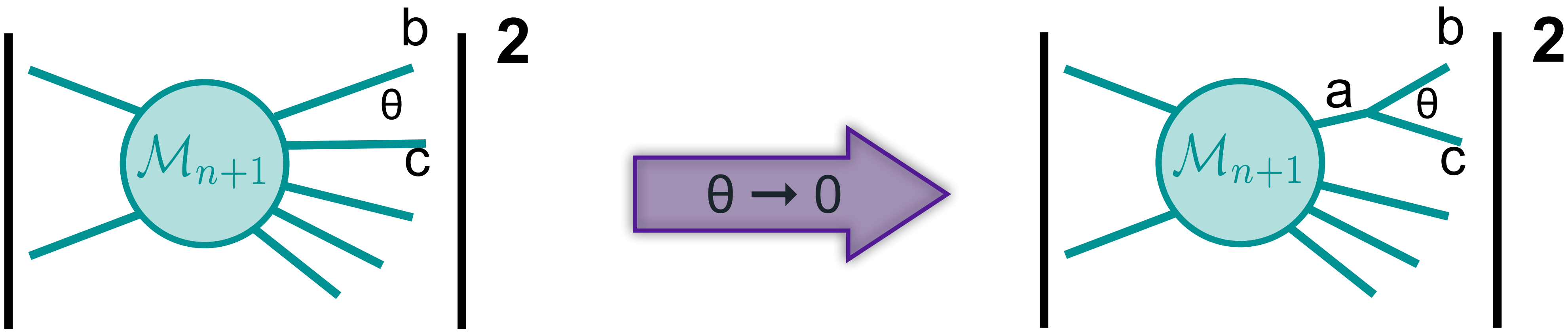
- In matrix element calculations in perturbation theory
  - "initial state QCD radiation" is included inclusively ("resummed") in the PDFs (and through strong coupling definition) and
  - "final state QCD radiation" is included through the parton-jet duality (and through strong coupling definition)
- Hence… all is already there! What to do…?
  - "Undo" this resummation and make it explicit

Collinear factorisation



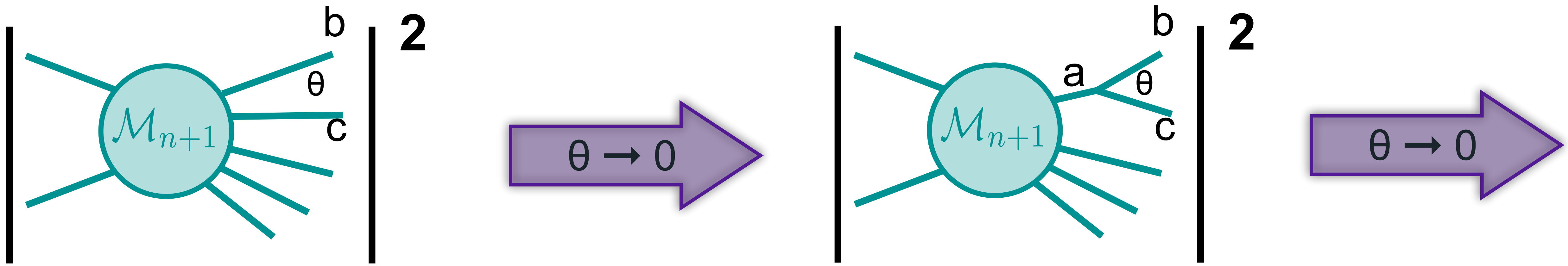
- Consider a process for which two particles are separated by a small angle θ
- In the limit of θ → 0, the contribution is coming from a single parent particle going on shell: therefore its branching is related to time scales which are very long with respect to the hard subprocess
- The inclusion of such a branching cannot change the picture set up by the hard process: the whole emission process must be writable in this limit as the simpler one times a branching probability

Collinear factorisation



- Consider a process for which two particles are separated by a small angle θ
- In the limit of θ → 0, the contribution is coming from a single parent particle going on shell: therefore its branching is related to time scales which are very long with respect to the hard subprocess
- The inclusion of such a branching cannot change the picture set up by the hard process: the whole emission process must be writable in this limit as the simpler one times a branching probability

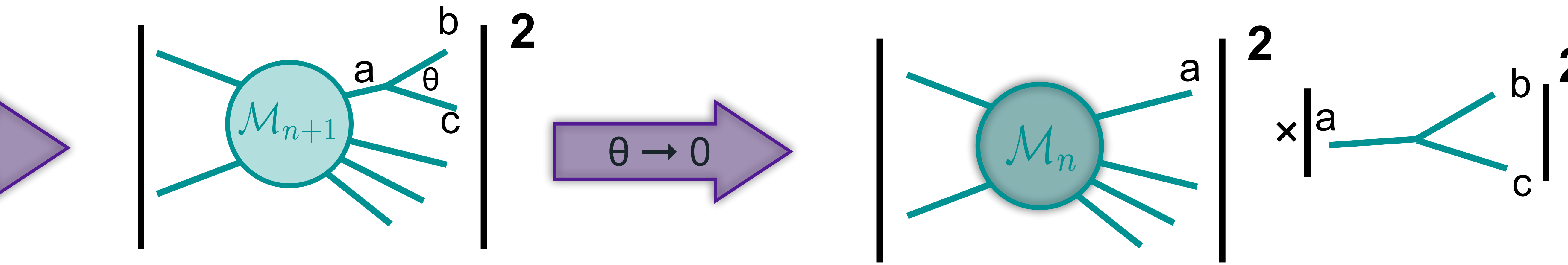
Collinear factorisation



- Consider a process for which two particles are separated by a small angle θ
- In the limit of θ → 0, the contribution is coming from a single parent particle going on shell: therefore its branching is related to time scales which are very long with respect to the hard subprocess
- The inclusion of such a branching cannot change the picture set up by the hard process: the whole emission process must be writable in this limit as the simpler one times a branching probability

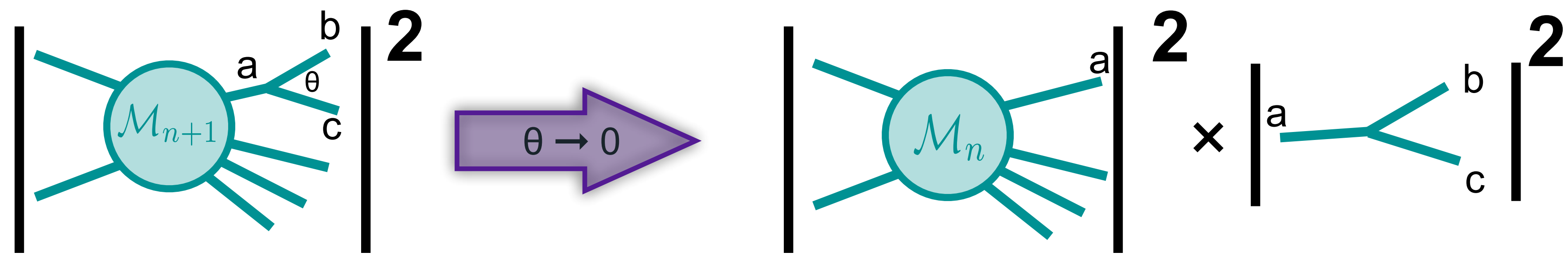
LUNDS UNIVERSITET

Collinear factorisation



- Consider a process for which two particles are separated by a small angle θ
- In the limit of θ → 0, the contribution is coming from a single parent particle going on shell: therefore its branching is related to time scales which are very long with respect to the hard subprocess
- The inclusion of such a branching cannot change the picture set up by the hard process: the whole emission process must be writable in this limit as the simpler one times a branching probability

Collinear factorisation



- The process factorises in the collinear limit. This procedure is universal

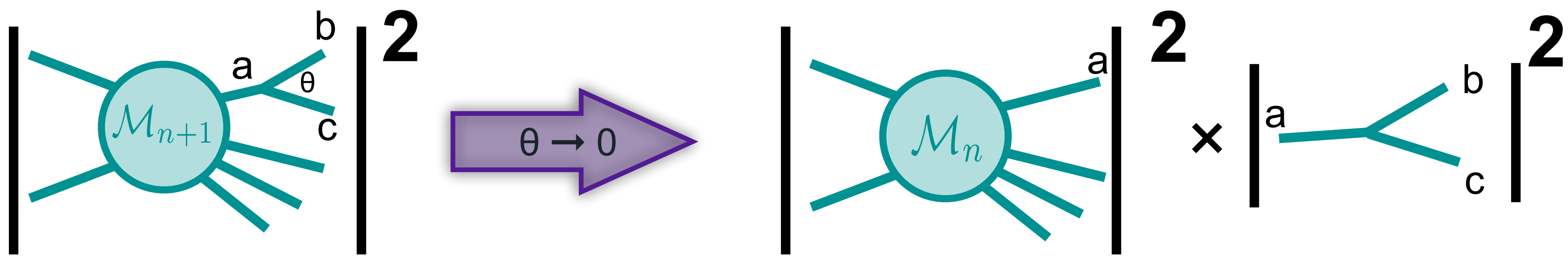
$$|\mathcal{M}_{n+1}|^2 d\Phi_{n+1} \simeq |\mathcal{M}_n|^2 d\Phi_n \frac{dt}{t} dz \frac{d\phi}{2\pi} \frac{\alpha_s}{2\pi} P_{a\to bc}(z)$$

- Notice that what has been roughly called 'branching fraction' is actually a singular factor, so one will need to make sense of this definition.
- At the leading contribution to the (n+1)-body cross section the DGLAP splitting kernels are defined as:

$$P_{g\to qq}(z) = T_R\left[z^2 + (1-z)^2\right], \qquad P_{g\to gg}(z) = C_A\left[z(1-z) + \frac{z}{1-z} + \frac{1-z}{z}\right],$$

$$P_{q\to qg}(z) = C_F\left[\frac{1+z^2}{1-z}\right], \qquad P_{q\to gq}(z) = C_F\left[\frac{1+(1-z)^2}{z}\right].$$

Collinear factorisation



- The process factorises in the collinear limit. This procedure is universal

$$|\mathcal{M}_{n+1}|^2 d\Phi_{n+1} \;\simeq\; |\mathcal{M}_n|^2 d\Phi_n \frac{dt}{t} dz \frac{d\phi}{2\pi}\frac{\alpha_s}{2\pi} P_{a\to bc}(z)$$

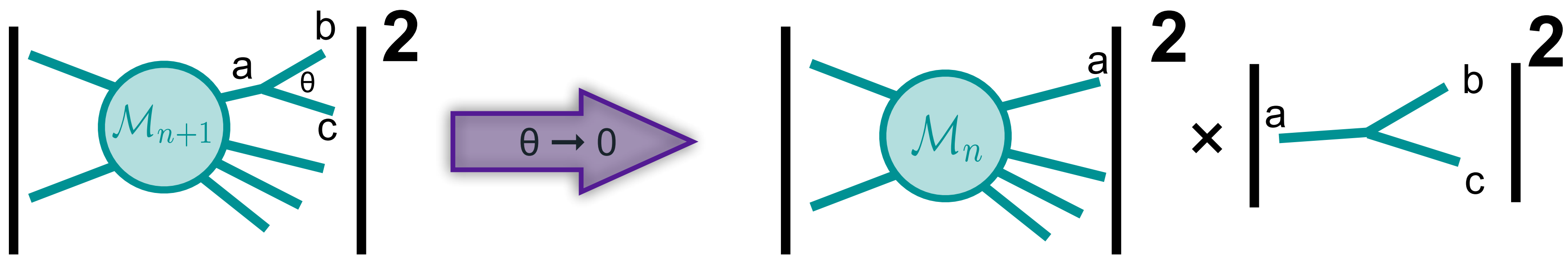
- t can be called the 'evolution variable': it can be the virtuality m^2 of particle a, or its p_T^2, or $E^2\theta^2$ …
  - It represents the hardness of the branching and tends to 0 in the collinear limit.

$$m^2 \simeq z(1-z)\theta^2 E_a^2$$
$$p_T^2 \simeq z m^2$$

- Indeed in the collinear limit one has: so that the factorisation takes place for all these definitions:

$$d\theta^2/\theta^2 \;=\; dm^2/m^2 \;=\; dp_T^2/p_T^2$$

Collinear factorisation

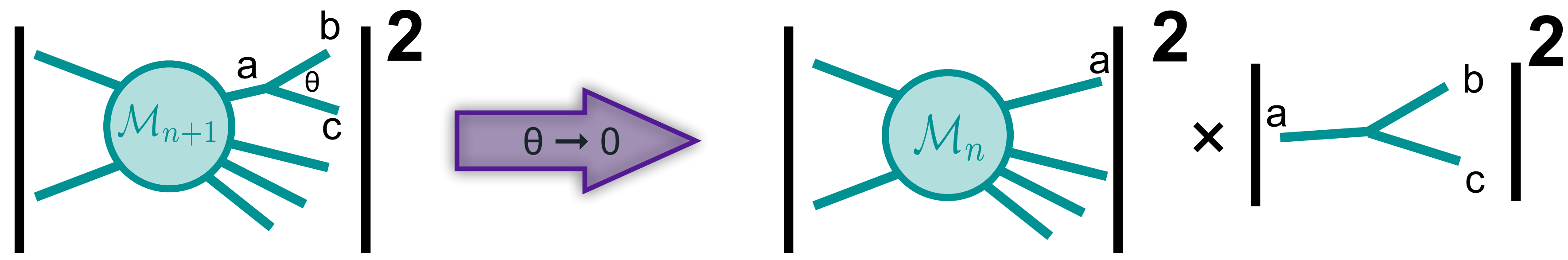


- The process factorises in the collinear limit. This procedure is universal

$$|\mathcal{M}_{n+1}|^2 d\Phi_{n+1} \simeq |\mathcal{M}_n|^2 d\Phi_n \frac{dt}{t} dz \frac{d\phi}{2\pi} \frac{\alpha_s}{2\pi} P_{a\to bc}(z)$$

- z is the "energy variable": it is defined to be the energy fraction taken by parton b from parton a
  - It represents the energy sharing between b and c and tends to 1 in the soft limit (parton c going soft)
- ϕ is the azimuthal angle. It can be chosen to be the angle between the polarisation of a and the plane of the branching

Collinear factorisation

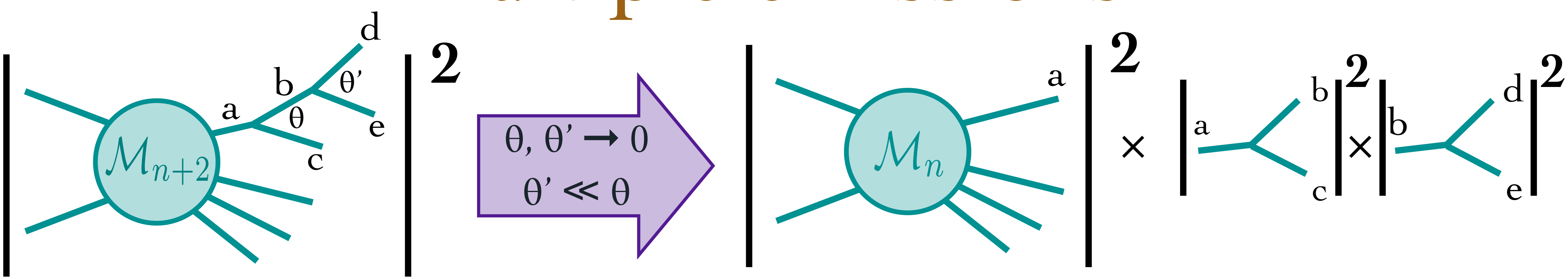


- The process factorises in the collinear limit. This procedure is universal

$$|\mathcal{M}_{n+1}|^2 d\Phi_{n+1} \simeq |\mathcal{M}_n|^2 d\Phi_n \frac{dt}{t} dz \frac{d\phi}{2\pi} \frac{\alpha_\mathrm{S}}{2\pi} P_{a\to bc}(z)$$

- This is an amplitude squared: naively one would maybe expect $1/t^2$ dependence. Why is the square not there?
    - It's due to angular-momentum conservation.
    E.g., take the splitting $q \longrightarrow qg$: helicity is conserved for the quarks, so the final state spin differs by one unity with respect to the initial one. The scattering happens in a p-wave (orbital angular momentum equal to one), so there is a suppression factor as $t \longrightarrow 0$.
        - Indeed, a factor $1/t$ is always cancelled in an explicit computation

Multiple emissions

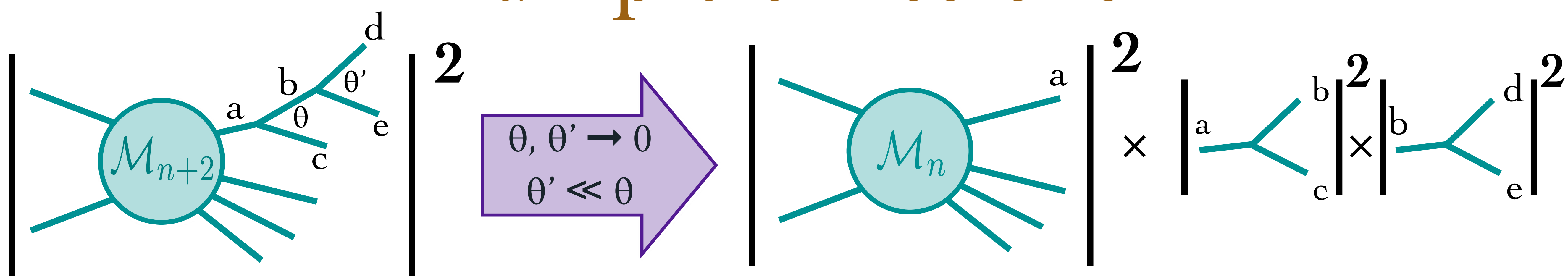


- Now consider M_{n+2} as the new core process and use the recipe we used for the first emission in order to get the dominant contribution to the (n+2)-body cross section: add a new branching at angle much smaller than the previous one:

$$|\mathcal{M}_{n+2}|^2 d\Phi_{n+2} \simeq |\mathcal{M}_n|^2 d\Phi_n \frac{dt}{t} dz \frac{d\phi}{2\pi} \frac{\alpha_s}{2\pi} P_{a\to bc}(z) \times \frac{dt'}{t'} dz' \frac{d\phi'}{2\pi} \frac{\alpha_s}{2\pi} P_{b\to de}(z')$$

- This can be done for an arbitrary number of emissions. The recipe to get the leading collinear singularity is thus cast in the form of an iterative sequence of emissions whose probability does not depend on the past history of the system: a ‘Markov chain’.

Multiple emissions



- The dominant contribution comes from the region where the subsequently emitted partons satisfy the strong ordering requirement: θ » θ' » θ''...

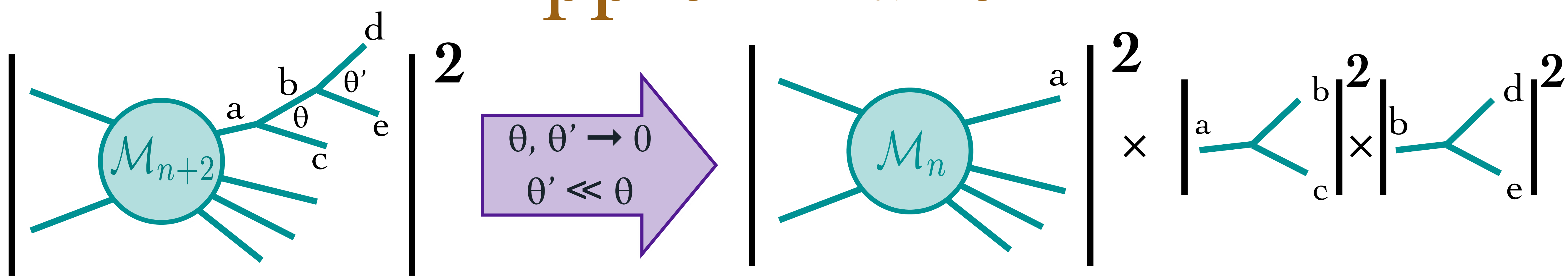
  For the rate for multiple emission we get

$$\sigma_{n+k} \propto \alpha_S^k \int_{Q_0^2}^{Q^2} \frac{dt}{t} \int_{Q_0^2}^{t} \frac{dt'}{t'} \cdots \int_{Q_0^2}^{t^{(k-2)}} \frac{dt^{(k-1)}}{t^{(k-1)}} \propto \sigma_n \left(\frac{\alpha_S}{2\pi}\right)^k \log^k(Q^2/Q_0^2)$$

  where Q is a typical hard scale and Q_0 is a small infrared cutoff that separates perturbative from non perturbative regimes.

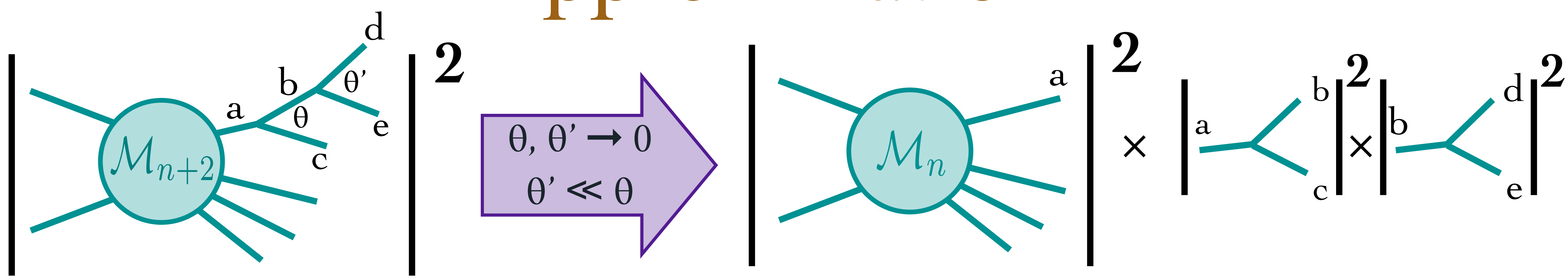
- Each power of α_s comes with a logarithm. The logarithm can easily be large, and therefore we see a breakdown of perturbation theory

Approximation



- We have an approximation of the matrix elements for multiple emissions
- We know that the $|M_n|^2$ is *inclusive* over all radiation (due to parton/jet duality and PDF evolution)
  - However, in our approximation we multiply $|M_n|^2$ by k "branching fractions" to get $|M_{n+k}|^2$. These branching fractions are actually singular factors
  - How to make sense of this? How to enforce that summing over all branching fractions adds up to one?

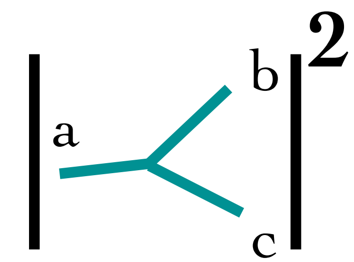
Approximation



- We have an approximation of the matrix elements for multiple emissions
- We know that the $|M_n|^2$ is *inclusive* over all radiation (due to parton/jet duality and PDF evolution)
  - However, in our approximation we multiply $|M_n|^2$ by k "branching fractions" to get $|M_{n+k}|^2$. These branching fractions are actually singular factors
  - How to make sense of this? How to enforce that summing over all branching fractions adds up to one?

We are missing the contributions with no emission

(No-)emission probability



- The probability for the branching $a \longrightarrow bc$ between scales t and $t+dt$ is equal to

$$dp(t) = \sum_{bc} \frac{dt}{t} \int dz \frac{d\phi}{2\pi} \frac{\alpha_S}{2\pi} P_{a\to bc}(z)$$

- The probability that a parton does NOT split between the scales t and $t+dt$ is given by $1-dp(t)$

- Probability that particle a does not emit between scales Q^2 and t

$$\Delta(Q^2,t) = \prod_k \left[1 - \sum_{bc} \frac{dt_k}{t_k} \int dz \frac{d\phi}{2\pi} \frac{\alpha_S}{2\pi} P_{a\to bc}(z)\right] =$$

$$\exp\left[-\sum_{bc} \int_t^{Q^2} \frac{dt'}{t'} dz \frac{d\phi}{2\pi} \frac{\alpha_S}{2\pi} P_{a\to bc}(z)\right] = \exp\left[-\int_t^{Q^2} dp(t')\right]$$

$\Delta(Q^2,t)$ is the Sudakov form factor

Sudakov form factor

- The Sudakov form factor is the heart of the parton shower. It gives the probability that a parton does not branch between two scales
  *Initial state shower also requires PDF contributions
- This no-emission probability needs to be included to interpret the branchings as probabilities that add up to 1
- Define dP_k as the probability for k ordered splittings from leg a at given scales

$$
\begin{aligned}
dP_1(t_1) &= \Delta(Q^2, t_1)\ dp(t_1)\Delta(t_1, Q_0^2), \\
dP_2(t_1, t_2) &= \Delta(Q^2, t_1)\ dp(t_1)\ \Delta(t_1, t_2)\ dp(t_2)\ \Delta(t_2, Q_0^2)\Theta(t_1 - t_2), \\
\ldots &= \ldots \\
dP_k(t_1, \ldots, t_k) &= \Delta(Q^2, Q_0^2) \prod_{l=1}^{k} dp(t_l)\Theta(t_{l-1} - t_l)
\end{aligned}
$$

- Q_0^2 is the hadronisation scale (~1 GeV^2). Below this scale we do not trust the perturbative description for parton splitting anymore
- This is what is implemented in a parton shower, taking the scales for the splitting t_i randomly (but weighted according to the no-emission probability)

Unitarity

$$dP_k(t_1, \ldots, t_k) \quad = \quad \Delta(Q^2, Q_0^2) \prod_{l=1}^{k} dp(t_l)\Theta(t_{l-1} - t_l)$$

- The parton shower has to be unitary (the sum over all branching trees should be 1). We can explicitly check this by integrating the probability for k splittings

$$P_k \equiv \int dP_k(t_1, \ldots, t_k) = \Delta(Q^2, Q_0^2) \frac{1}{k!} \left[\int_{Q_0^2}^{Q^2} dp(t) \right]^k , \quad \forall k = 0, 1, \ldots$$

- Summing over all number of emissions

$$\sum_{k=0}^{\infty} P_k = \Delta(Q^2, Q_0^2) \sum_{k=0}^{\infty} \frac{1}{k!} \left[\int_{Q_0^2}^{Q^2} dp(t) \right]^k = \Delta(Q^2, Q_0^2) \exp \left[\int_{Q_0^2}^{Q^2} dp(t) \right] = 1$$

- Hence, the total probability is conserved

Physical interpretation

- Because we are including both the emission and no-emission contributions…
- …we should **not** interpret the parton shower to be generating an approximation of the $|M_{n+k}|^2$ matrix elements
  - Rather it is an approximation of the N^kLO computation of $|M_n|^2$
  - That is, including the real-emission contributions, but also virtual no-emission corrections

Initial-state parton showers

- To simulate parton radiation from the initial state, we start with the hard scattering, and then “devolve” the DGLAP evolution to get back to the original hadron: backwards evolution!
  - i.e. we undo the analytic resummation and replace it with explicit partons (e.g. in Drell-Yan this gives non-zero p_T to the vector boson)
- In backwards evolution, the Sudakovs include also the PDFs -- this follows from the DGLAP equation and ensures conservation of probability:

$$\Delta_{Ii}(x, t_1, t_2) = \exp\left\{-\int_{t_1}^{t_2} dt' \sum_j \int_x^1 \frac{dx'}{x'} \frac{\alpha_s(t')}{2\pi} P_{ij}\left(\frac{x}{x'}\right) \frac{f_i(x', t')}{f_j(x, t')}\right\}$$

  This represents the probability that parton i will stay at the same x (no splittings) when evolving from t_1 to t_2.
- The shower simulation is now done as in a final state shower

An LHC collision, factorised

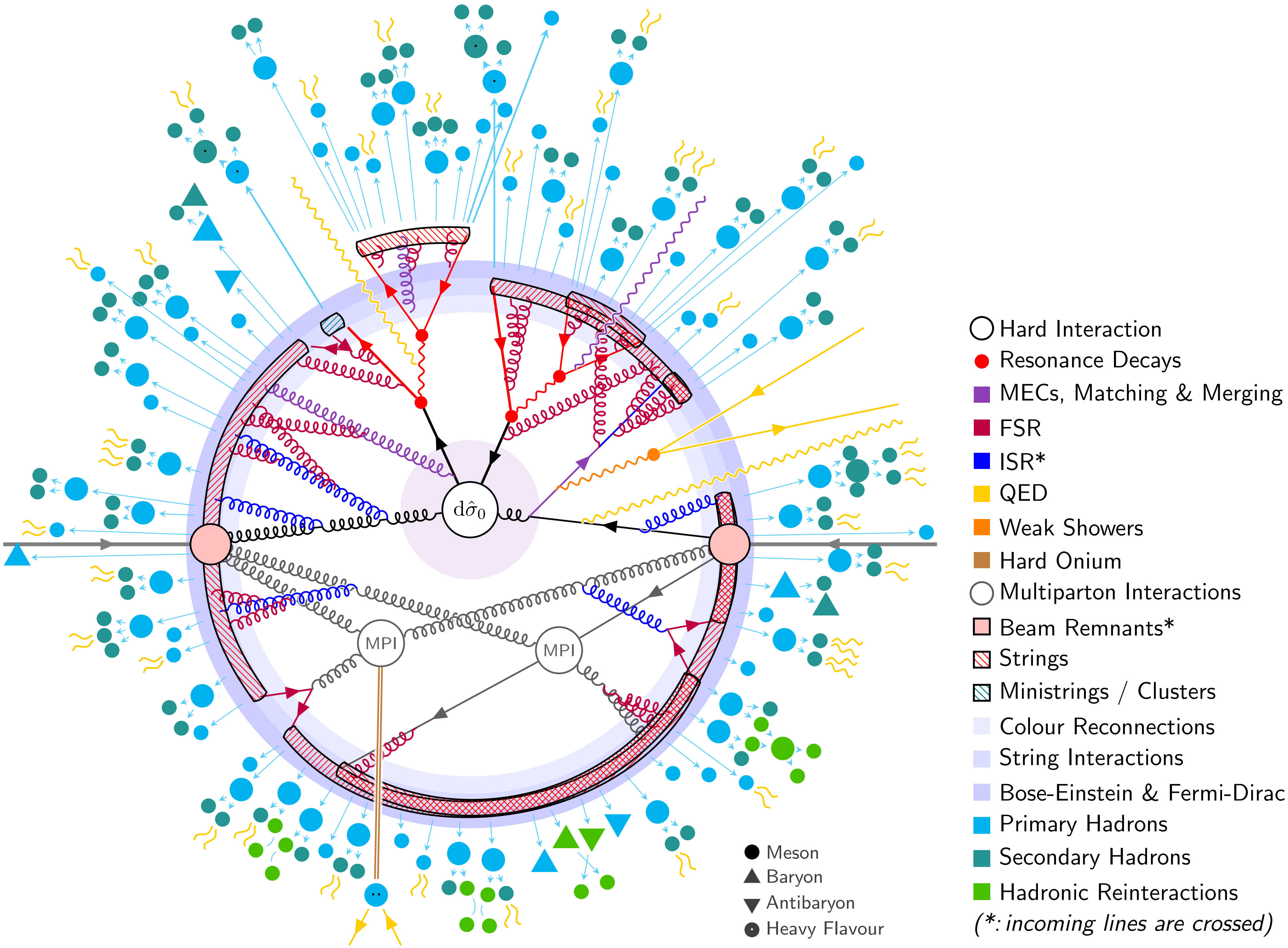


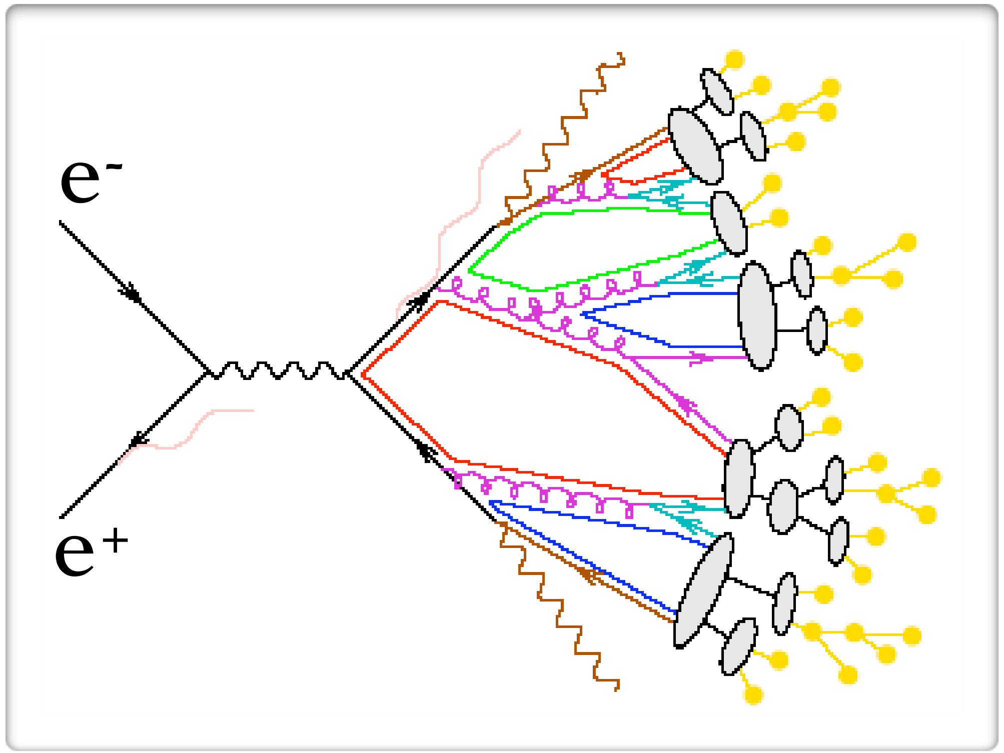
Figure taken from Bierlich et al., 2022 *(Pythia8.3 manual)*

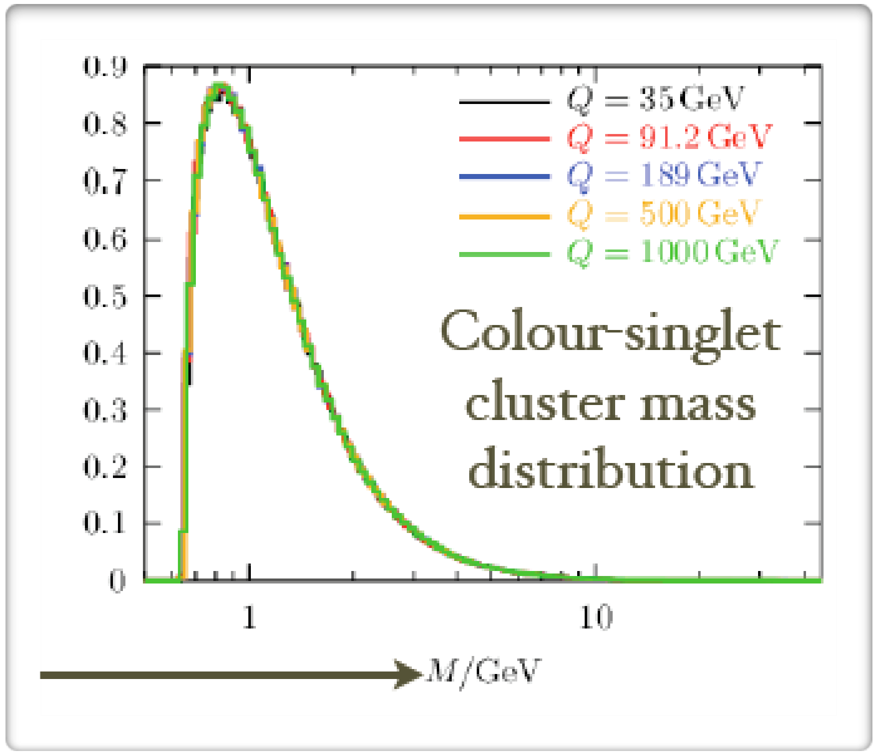
Hadronisation

- The shower stops if all partons are characterised by a scale at the IR cut-off: $Q_0 \sim 1$ GeV
- Physically, we observe hadrons, not (coloured) partons
- We need a non-perturbative model in passing from partons to colourless hadrons
- There are two models, based on physical and phenomenological considerations

Cluster model

The structure of the perturbative evolution including angular ordering, leads naturally to the clustering in phase-space of colour-singlet parton pairs (pre-confinement). Long-range correlations are strongly suppressed. Hadronisation will only act locally, on low-mass colour singlet clusters.





Lund string model

From lattice QCD one sees that the colour confinement potential of a quark-antiquark grows linearly with their distance: $V(r) \sim kr$, with $k \sim$ 0.2 GeV, This is modelled with a string with uniform tension (energy per unit length) k that gets stretched between the $q\bar{q}$ pair.

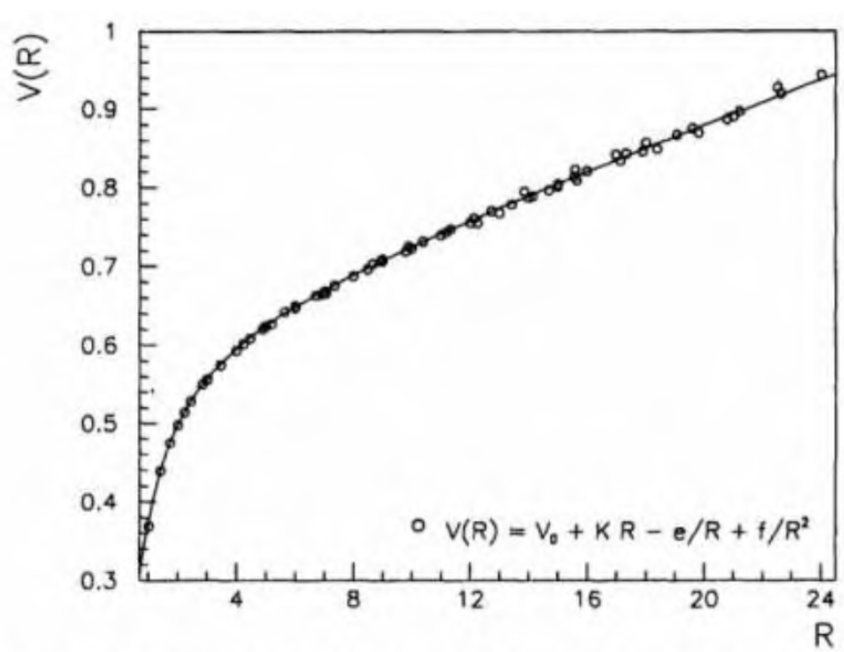
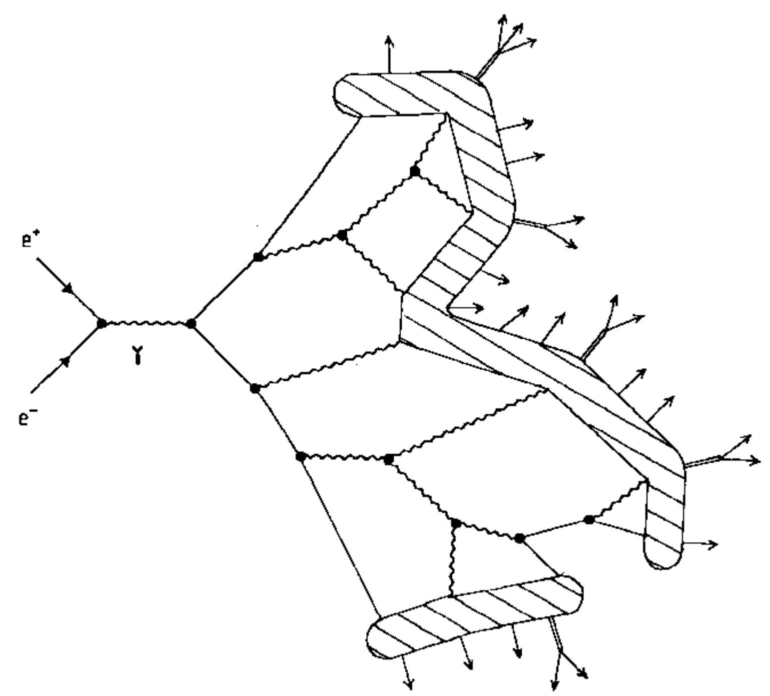


Fig. 2.9. QCD potential *vs.* *R* (in lattice units) from lattice QCD. Figure from ref. [23].



When quark-antiquarks are too far apart, it becomes energetically more favourable to break the string by creating a new $q\bar{q}$ pair in the middle.

An LHC collision, factorised

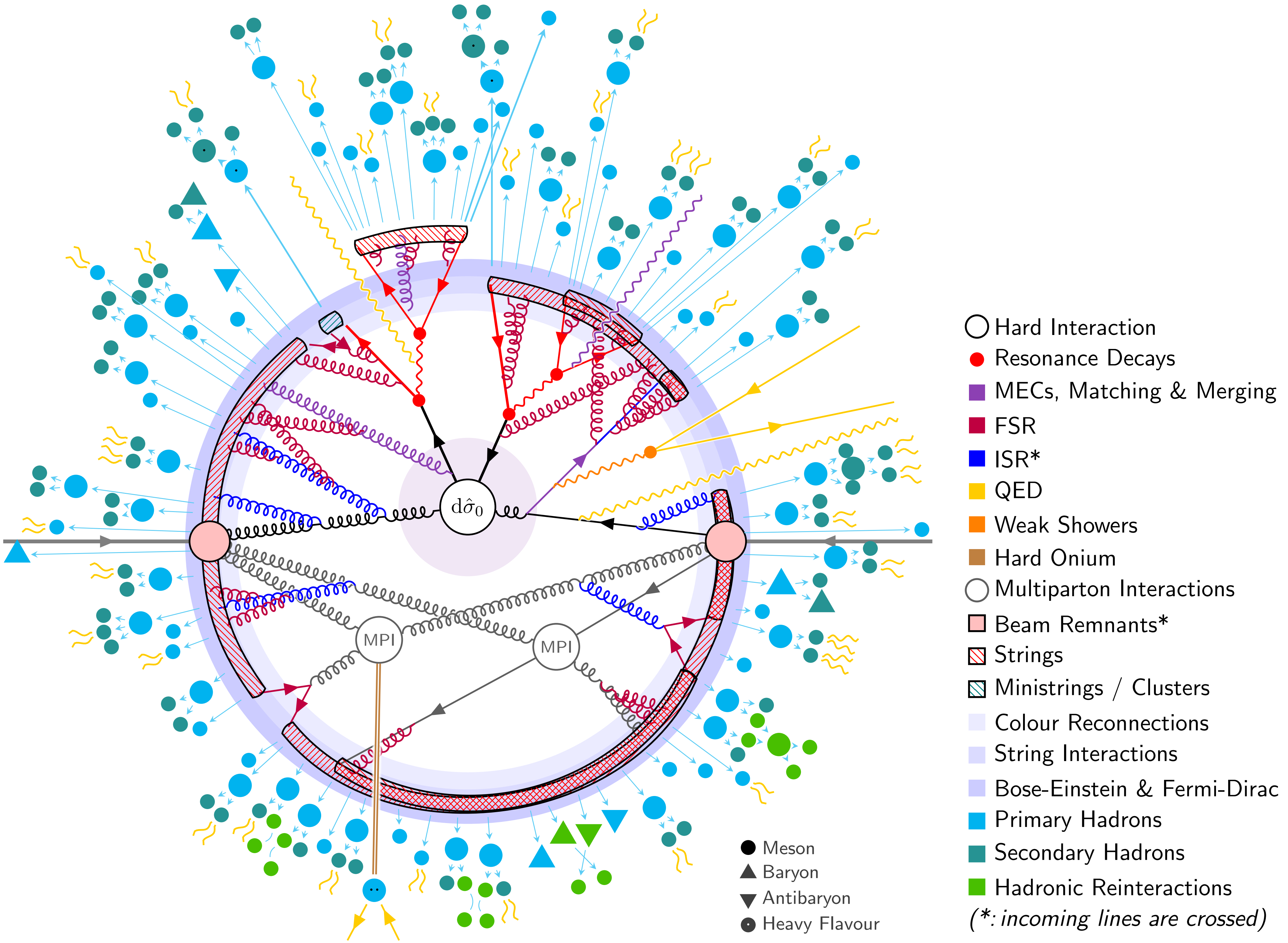
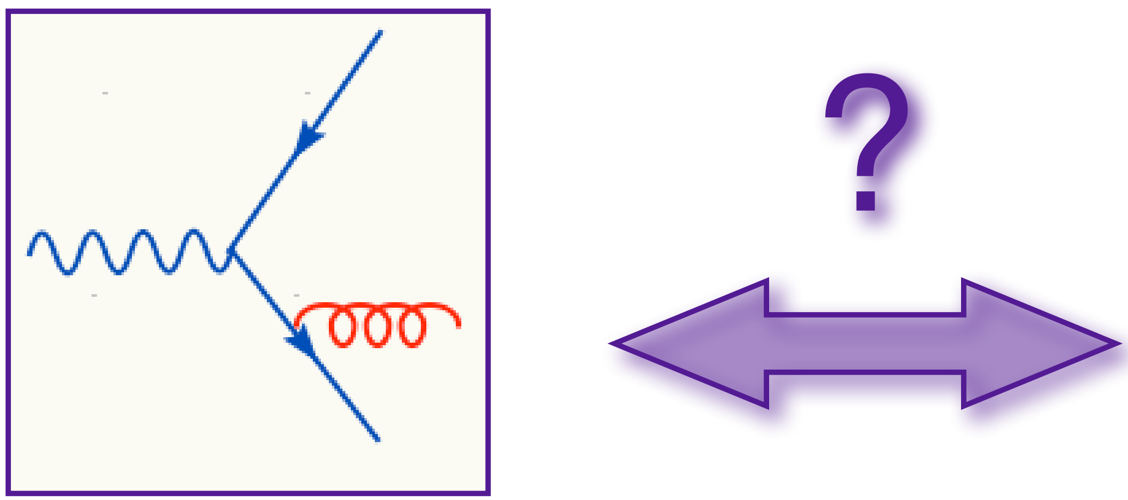
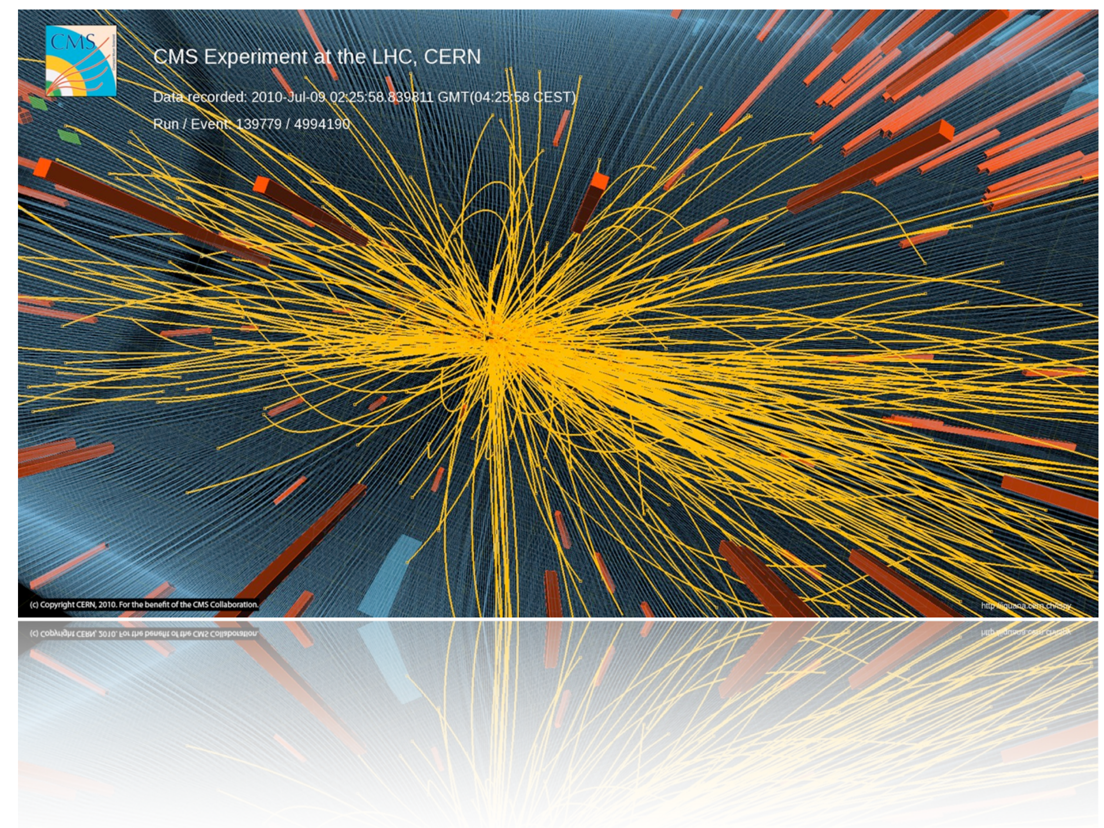


Figure taken from Bierlich et al., 2022 *(Pythia8.3 manual)*

Exclusive observable





A parton shower program associates one of the possible histories (and pre-histories in case of pp collisions) of an hard event in an explicit and fully detailed way, such that the sum of the probabilities of all possible histories is unity.

Parton Shower Monte-Carlo event generators

A parton shower program associates one of the possible histories (and pre-histories in case of pp) of an hard event in an explicit and fully detailed way, such that the sum of the probabilities of all possible histories is unity.

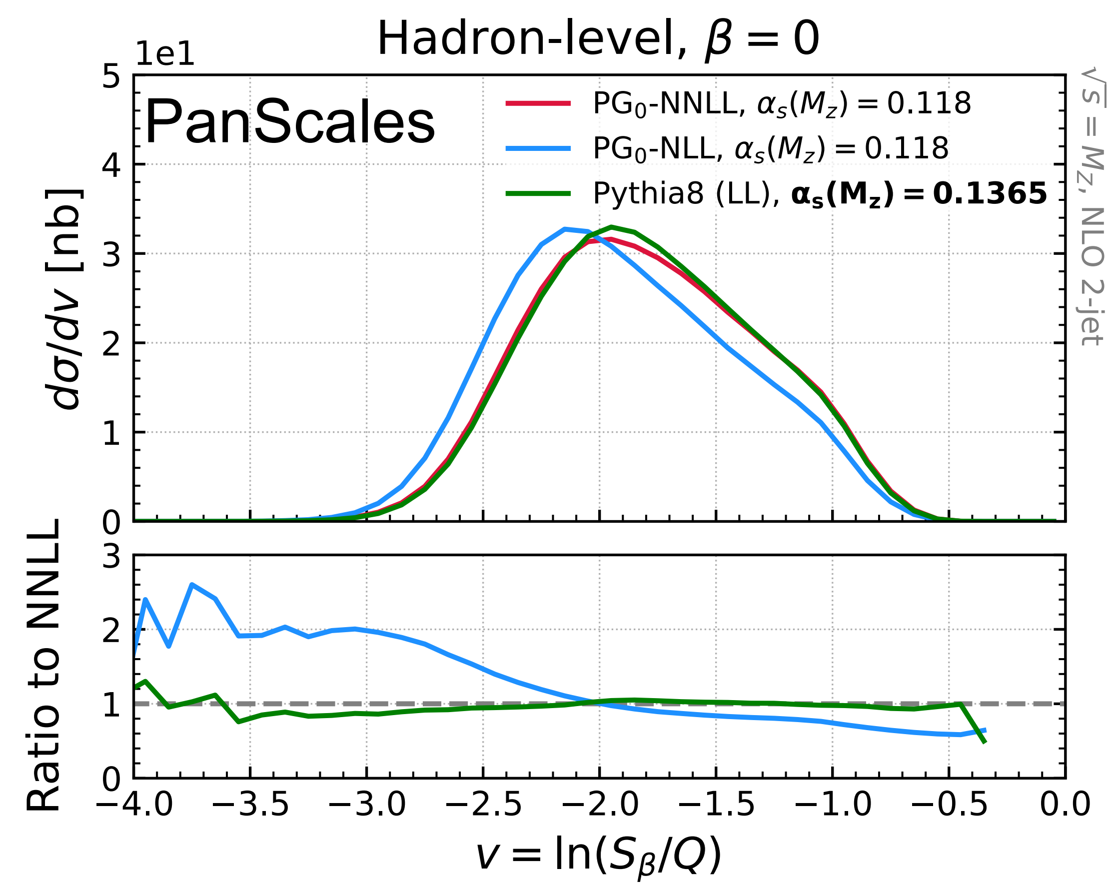
- General-purpose tools
- Always the first experimental choice
- Complete exclusive description of the events: hard scattering, showering & hadronisation (and underlying event)
- Reliable and well-tuned tools
- *Significant and intense progress in the development of new showering algorithms with the final aim to go beyond (N)LL in QCD*

Pythia, Herwig & Sherpa

- Significant differences between shower implementations (choice of evolution variable and kernel, momentum mappings, phase-space boundaries, massive quarks, photon emissions, etc.)
- All are tuned to data, and describe it reasonably well (typically better than expected from their formal accuracy)
- Some are (formally) more correct than others
  - However, not easy to assess accuracy for a general observable
  - Assessment (and improvement!) of formal accuracy is an active field of research

New approaches

- Hot topic: the accuracy of current parton showers is one of the limiting factors in our understanding of LHC data
  - They seem to be working much better than one should expect!
- New approaches such as *Deductor*, *PanScales*, etc., aim at understanding and improving the current accuracy of parton shower implementations
  - Turns out, that, heuristically, for e^+e^- collisions, almost all beyond-LL effects can be taken care of by rescaling the value of α_s
    - Not obvious if this also holds true for hadron-hadron collisions (probably not!)



energy-correlation moment @ LEP

Extra material

Exclusive observables

- Very exclusive observables are poorly described in perturbation theory.
  - One could take the conservative attitude of considering only perturbatively well-behaved observables. But one would miss an extremely rich variety of observables which may play important roles in experimental analyses.
- If fixed-order perturbation theory breaks down for an observable, this does NOT mean that observable is useless/unimportant: it is just that one is not using the right tools to describe it.
- It is better to try and find a way to reorganise the computation in order to take into account emissions close to the singular regions of the phase space, to all orders in perturbation theory.
  - This can be done in a systematic way: "resummation"!

Absence of interference

- The collinear factorisation picture gives a branching sequence for a given leg starting from the hard subprocess all the way down to the non-perturbative region.
- Suppose you want to describe two such histories from two different legs:
  - these two legs are treated in a completely uncorrelated way. And even within the same history, subsequent emissions are uncorrelated.
- The collinear picture completely misses the possible interference effects between the various legs
  - the extreme simplicity comes with the price of quantum inaccuracy.
- **Smart choices improve upon this: soft enhancement (which is purely an interference contribution) can be included. For this, the evolution variable must be related to the angle of the emission**
- Nevertheless, the collinear picture captures the leading contributions: it gives an excellent description of an arbitrary number of (collinear) emissions:
  - it is a “resummed computation” and
  - it bridges the gap between fixed-order perturbation theory and the non-perturbative hadronisation.

Cancellation of singularities

- We have shown that the shower is unitary. However, how are the IR divergences cancelled explicitly? Let's show this for the first emission: Consider the contributions from (exactly) 0 and 1 emissions from leg a:

$$\frac{d\sigma}{\sigma_n} = \Delta(Q^2, Q_0^2) + \Delta(Q^2, Q_0^2) \sum_{bc} dz \frac{dt}{t} \frac{d\phi}{2\pi} \frac{\alpha_s}{2\pi} P_{a\to bc}(z)$$

- Expanding to first order in α_s gives

$$\frac{d\sigma}{\sigma_n} \simeq 1 - \sum_{bc} \int_{Q_0^2}^{Q^2} \frac{dt'}{t'} dz \frac{d\phi}{2\pi} \frac{\alpha_s}{2\pi} P_{a\to bc}(z) + \sum_{bc} dz \frac{dt}{t} \frac{d\phi}{2\pi} \frac{\alpha_s}{2\pi} P_{a\to bc}(z)$$

- Same structure of the two latter terms, with opposite signs: cancellation of divergences between the approximate virtual and approximate real emission cross sections.
- The probabilistic interpretation of the shower ensures that infrared divergences will cancel for each emission.

Argument of α_S

- Each choice of argument for α_S is equally acceptable at the leading-logarithmic accuracy. However, there is a choice that allows one to resum certain classes of subleading logarithms.
- The higher order corrections to the partons splittings imply that the DGLAP splitting kernels should be modified: $P_{a\to bc}(z) \longrightarrow P_{a\to bc}(z) + \alpha_S P'_{a\to bc}(z)$
- For $g \to gg$ branchings $P'_{a\to bc}(z)$ diverges as $-b_0 \log[z(1-z)]\ P_{a\to bc}(z)$ (just z or $1-z$ if quark is present)
- Recall the one-loop running of the strong coupling

$$\alpha_S(Q^2) = \frac{\alpha_S(\mu^2)}{1+\alpha_S(\mu^2)b_0\log\frac{Q^2}{\mu^2}} \sim \alpha_S(\mu^2)\left(1-\alpha_S(\mu^2)b_0\log\frac{Q^2}{\mu^2}\right)$$

- We can therefore include the P'(z) terms by choosing $p_T^2 \sim z(1-z)Q^2$ as argument of α_S:

$$\begin{aligned}\alpha_S(Q^2)\left(P_{a\to bc}(z)+\alpha_S(Q^2)P'_{a\to bc}\right) &= \alpha_S(Q^2)\left(1-\alpha_S(Q^2)b\log z(1-z)\right)P_{a\to bc}(z)\\ &\sim \alpha_S(z(1-z)Q^2)P_{a\to bc}(z)\end{aligned}$$

Choice of evolution parameter

$$\Delta(Q^2,t) = \exp\left[-\sum_{bc}\int_t^{Q^2}\frac{dt'}{t'}dz\frac{d\phi}{2\pi}\frac{\alpha_s}{2\pi}P_{a\to bc}(z)\right]$$

- There is a lot of freedom in the choice of evolution parameter t. It can be the virtuality m^2 of particle a or its $p_T{}^2$ or $E^2\theta^2$ ... For the collinear limit they are all equivalent

$$d\theta^2/\theta^2 \;=\; dm^2/m^2 \;=\; dp_T^2/p_T^2$$

- However, in the soft limit ($z \longrightarrow 1$) they behave differently
- Can we chose it such that we get the correct soft limit?

Choice of evolution parameter

$$\Delta(Q^2, t) = \exp\left[-\sum_{bc}\int_t^{Q^2}\frac{dt'}{t'}dz\frac{d\phi}{2\pi}\frac{\alpha_S}{2\pi}P_{a\to bc}(z)\right]$$

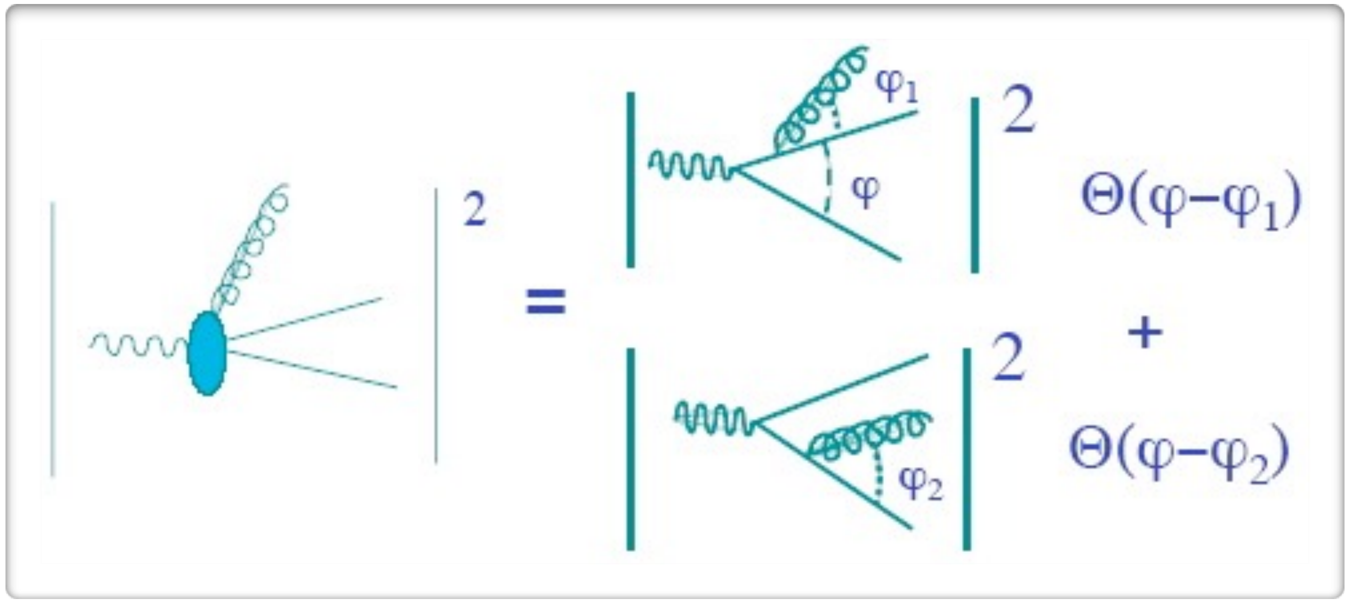
- There is a lot of freedom in the choice of evolution parameter t. It can be the virtuality m^2 of particle a or its p_T^2 or $E^2\theta^2$ ... For the collinear limit they are all equivalent

$$d\theta^2/\theta^2 \;=\; dm^2/m^2 \;=\; dp_T^2/p_T^2$$

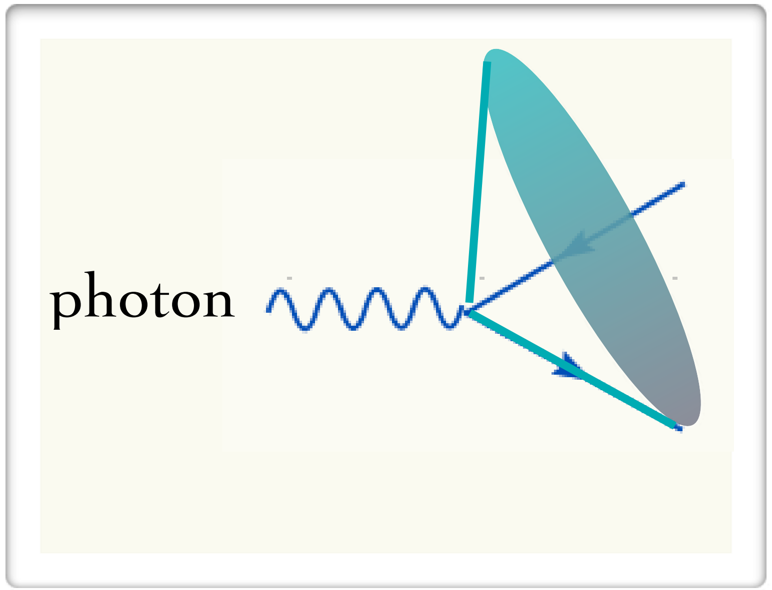
- However, in the soft limit ($z \longrightarrow 1$) they behave differently
- Can we chose it such that we get the correct soft limit?

YES! It should be (proportional to) the angle θ

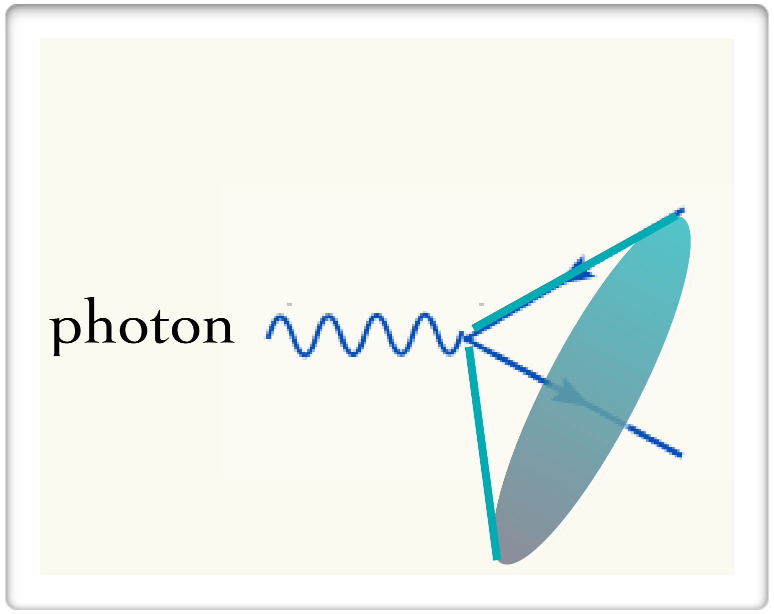
Angular ordering



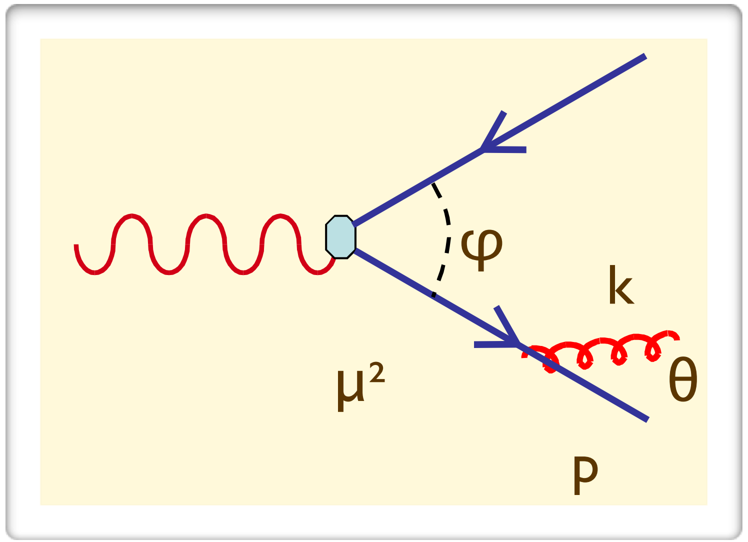
- Radiation inside cones around the original partons is allowed (and described by the eikonal approximation), outside the cones it is zero (after averaging over the azimuthal angle)



+



Intuitive explanation



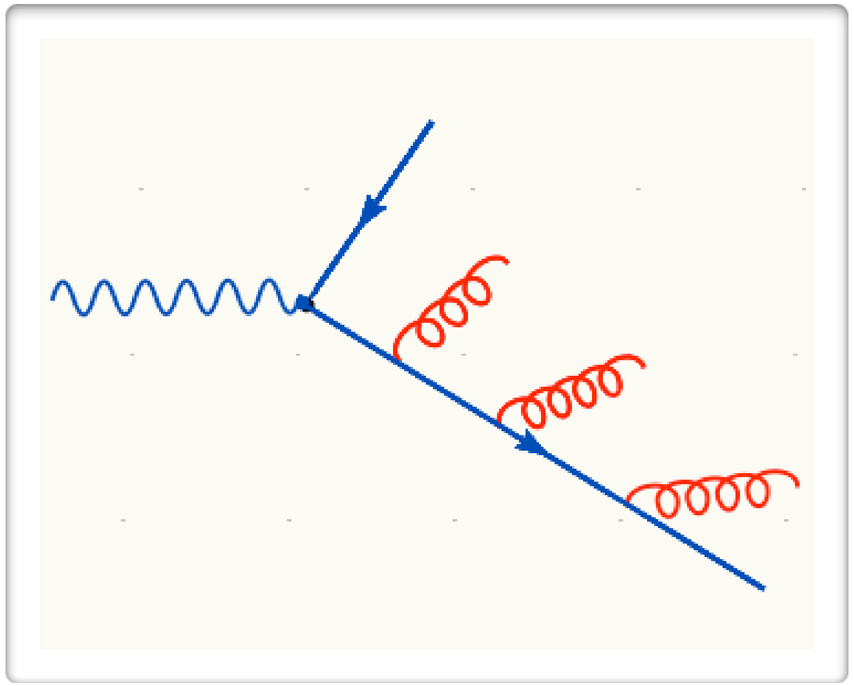
$\mu^2 = (p+k)^2 = 2E\, k_0\, (1-\cos\theta)$
$\sim E\, k_0\, \theta^2 \sim E\, k_\perp\, \theta$

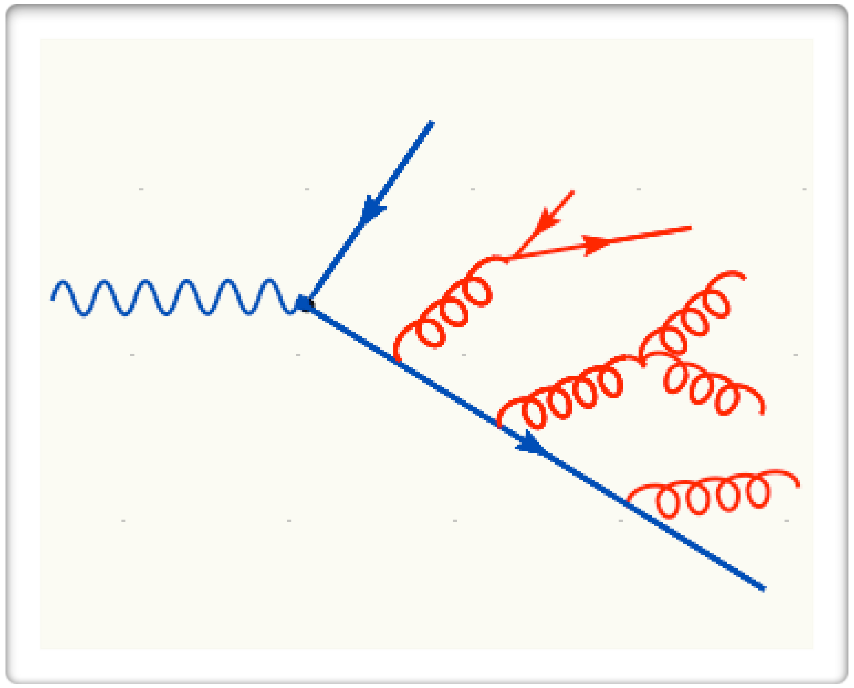
- Lifetime of the virtual intermediate state:
  $\tau < \gamma/\mu = E/\mu^2 = 1/(k_0\theta^2) = 1/(k_\perp\theta)$
- Distance between q and qbar after τ:
  $d = \varphi\tau = (\varphi/\theta)\ 1/k_\perp$

If the transverse wavelength of the emitted gluon is longer than the separation between q and qbar, the gluon emission is suppressed, because the q qbar system will appear as colour neutral (i.e. dipole-like emission, suppressed)

Therefore $d>1/k_\perp$, which implies $\theta < \varphi$

Angular ordering





- The construction can be iterated to the next emission, with the result that the emission angles keep getting smaller and smaller.
- One can generalise it to a generic parton of colour charge Q_k splitting into two partons i and j, $Q_k=Q_i+Q_j$. The result is that inside the cones i and j emit as independent charges, and outside their angular-ordered cones the emission is coherent and can be treated as if it was directly from colour charge Q_k.
- Angular ordering is automatically satisfied in θ ordered showers! (and straight-forward to account for in p_T ordered showers)

Angular ordering

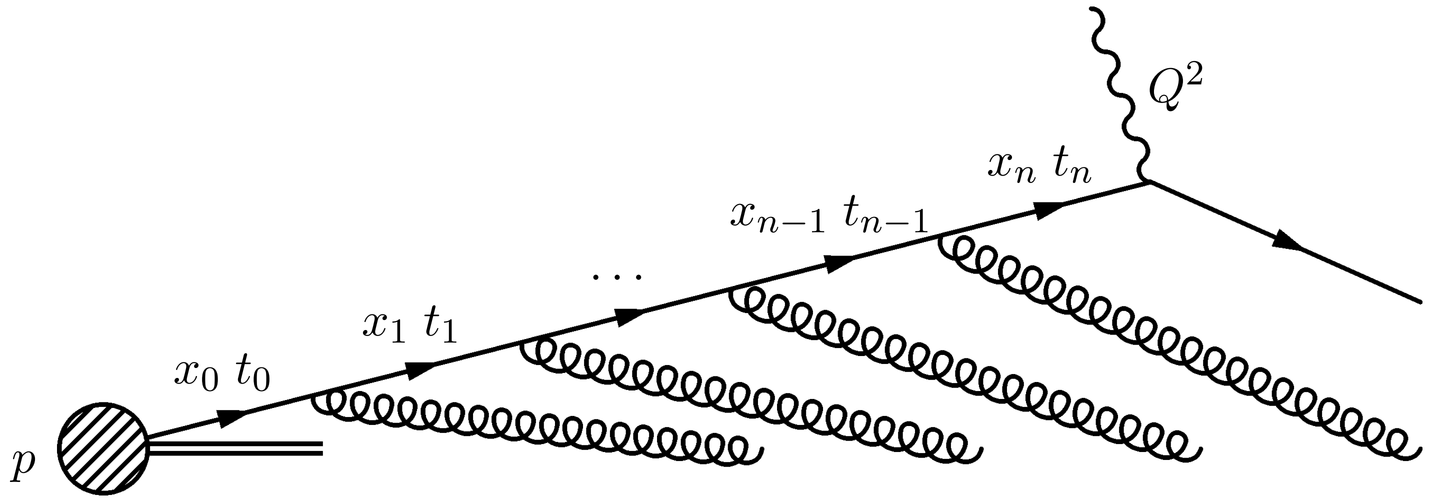
Angular ordering is:

1. A quantum effect coming from the interference of different Feynman diagrams.
2. Nevertheless it can be expressed in “a classical fashion” (square of an amplitude is equal to the sum of the squares of two special “amplitudes”). The classical limit is the dipole-radiation.
3. It is not an exclusive property of QCD (i.e., it is also present in QED) but in QCD produces very non-trivial effects, depending on how particles are colour connected.

Initial-state parton splittings

- So far, we have looked at final-state (time-like) splittings
- For initial state, the splitting functions are the same
- However, there is another ingredient:
  the parton density (or distribution) functions (PDFs)
  - ➡ Naively: Probability to find a given parton in a hadron at a given momentum fraction $x = p_z/P_z$ and scale t
- How do the PDFs evolve with increasing t?

Initial-state parton splittings



- Start with a quark PDF $f_0(x)$ at scale t_0. After considering a single parton emission, the probability to find the quark at virtuality $t > t_0$ is

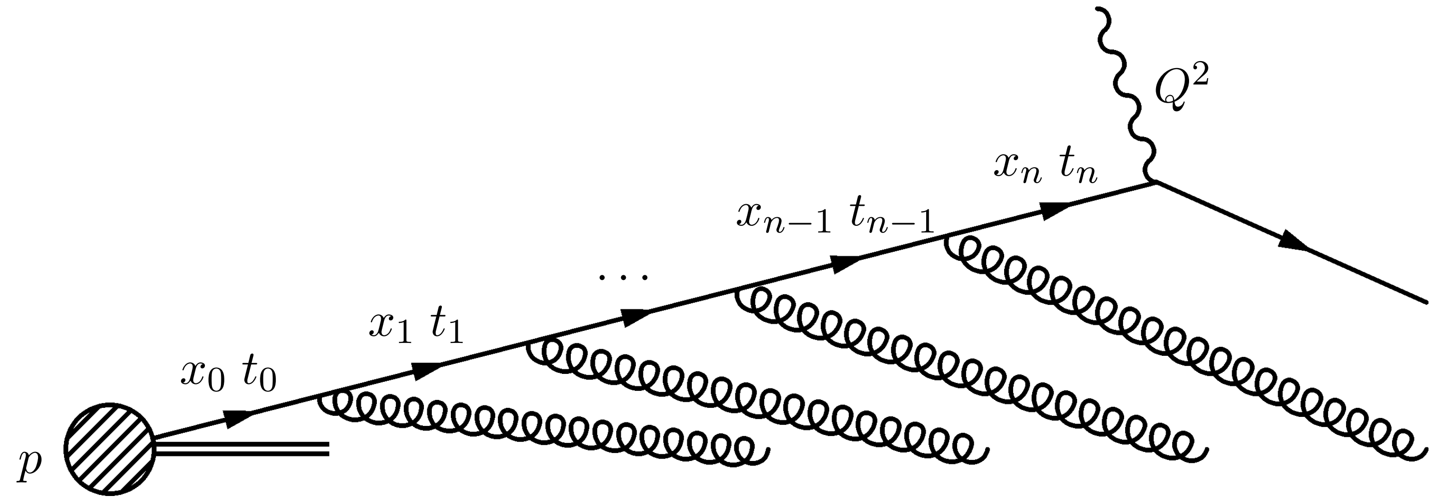
$$f(x,t) = f_0(x) + \int_{t_0}^{t} \frac{dt'}{t'} \frac{\alpha_s}{2\pi} \int_x^1 \frac{dz}{z} P(z) f_0\left(\frac{x}{z}\right)$$

- After a second emission, we have

$$f(x,t) = f_0(x) + \int_{t_0}^{t} \frac{dt'}{t'} \frac{\alpha_s}{2\pi} \int_x^1 \frac{dz}{z} P(z) \left\{ f_0\left(\frac{x}{z}\right) + \int_{t_0}^{t'} \frac{dt''}{t''} \frac{\alpha_s}{2\pi} \int_{x/z}^1 \frac{dz'}{z'} P(z') f_0\left(\frac{x}{zz'}\right) \right\}$$

$f(x/z, t')$

The DGLAP equation



- So for multiple parton splittings, we arrive at an integral-differential equation:

$$t\frac{\partial}{\partial t} f_i(x,t) = \int_x^1 \frac{dz}{z}\frac{\alpha_s}{2\pi} P_{ij}(z) f_j\left(\frac{x}{z}, t\right)$$

- This is the famous DGLAP equation (where we have taken into account the multiple parton species i, j). The boundary condition for the equation is the initial PDFs $f_{i0}(x)$ at a starting scale t_0 (around 2 GeV).
- These starting PDFs are fitted to experimental data.