# MC event generator introduction

CTEQ-MCnet summer school 2021
September 06/07, 2021
Stefan Prestel (Lund University)

The goal of these lectures is to

give you basic background on different aspects and algorithms in high-energy physics event generators

the lectures are split into four parts:

- General overview, basic sampling algorithms
- Phase space and hard scattering
- Sudakov algorithm and its application in showers and MPI
- Hadronization

I aim for a broad, but not too detailed overview. Overlap with the other lectures is expected :)

An observation in particle physics is

$$\langle O\rangle = \int \underbrace{d\phi_n}_{\text{phase space}} \underbrace{\frac{d\sigma(A,B\to n\ \text{particles})}{d\phi_n}}_{\text{differential cross section}} \underbrace{O(\phi_n)}_{\text{value of observable}}$$

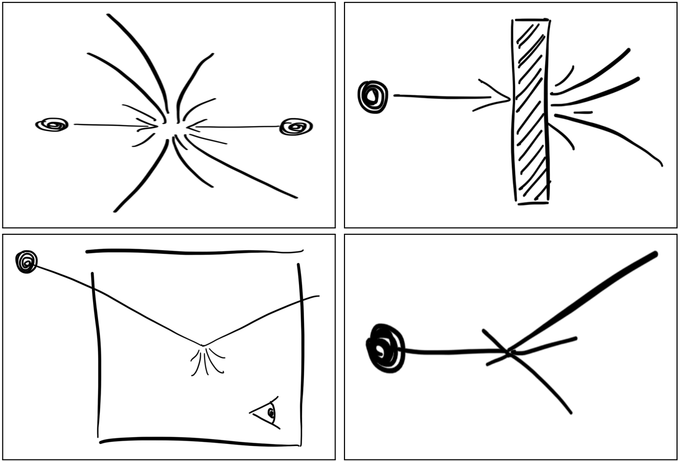
phase space: sample of all quantum numbers (momentum, flavor...) of particles in scattering final state
differential cross section≈ transition probability to scattering final state

Compare to expectation value in statistics:

$$\langle g\rangle = \int \underbrace{dx}_{\text{random variables}} \underbrace{f(x)}_{\text{probability distribution}} \underbrace{g(x)}_{\text{value of function}}$$

⇒ Calculate "theory predictions" for $O$ with statistical methods.

In an experiment, we create $\Phi_n$ states in various ways:



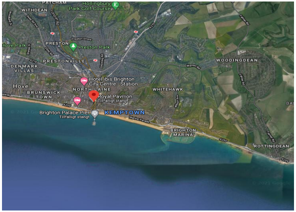
First three use *event generators* in experiment design and data analysis.
Event Generators are statistical tools to create “theory predictions”.

| | | |
|---|---|---|
| Dedicated calculations | : | Evaluate analytic expressions on paper…or very likely a computer. Safe & fast, but only viable for "simple" problems |
| Monte Carlo generators | : | Approximate analytic expressions numerically, by statistical sampling on a computer. Use Monte-Carlo methods to handle complex scattering final states and/or observations. |

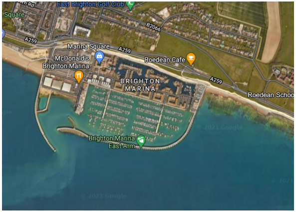
**Question:** What's the coastline of Britain?
**Reply:** How close do you look, i.e. at what *resolution*?

Coarse: Straight line

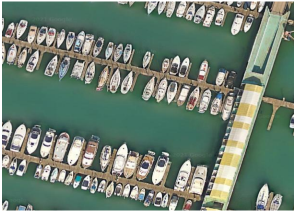

Finer: Marina appears




Finer: Quays appear




Finer: Ships add structure



...as do barnacles

...and the sand on them

Monte-Carlo algorithms are simple enough to have wide applicability, e.g. in integration

$$\int_{x_-}^{x_+} dx f(x) = (x_+ - x_-)\langle f\rangle \approx \frac{(x_+ - x_-)}{N}\sum_{i=1}^{N} f(x_i)$$

The approximation errors is $\propto \frac{1}{\sqrt{N}}$, independent of number of integrations $(dx \to dx_1 \cdots dx_n)$

Ideally suited for our types of integrals

$$\langle O\rangle = \int d\Phi_n \frac{d\sigma_n}{d\Phi_n} O(\Phi_n) \propto \frac{1}{N}\sum_{i=1}^{N} \frac{d\sigma_n}{d\Phi_n}(\Phi_n^{(i)}) O(\Phi_n^{(i)})$$

May even store the events $\Phi_n^{(i)}$ with event weight $\frac{d\sigma_n}{d\Phi_n}(\Phi_n^{(i)})$ and evaluate $O(\Phi_n^{(i)})$ later!

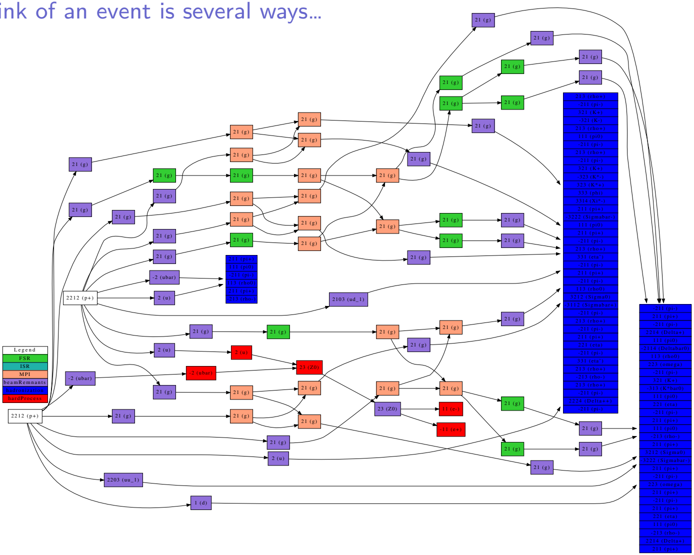
NB: Les Houches Event Files are effectively that.

You can think of an event is several ways...

```
<event>
 7  1  0.91E+00  0.80E+02  0.79E-01  0.12E+00
    2  -1  0  0  501    0  0.000E+00  0.000E+00  0.405E+03  0.405E+03  0.00
   -2  -1  0  0    0  501  0.000E+00  0.000E+00 -0.583E+02  0.583E+02  0.00
   24   2  1  2    0    0  0.979E+02 -0.643E+02  0.218E+03  0.259E+03  0.79
  -11   1  3  3    0    0  0.783E+02 -0.437E+01  0.116E+03  0.140E+03  0.00
   12   1  3  3    0    0  0.195E+02 -0.599E+02  0.101E+03  0.119E+03  0.00
    1   1  1  2  502    0 -0.558E+02  0.465E+02  0.139E+03  0.157E+03  0.00
   -2   1  1  2    0  502 -0.420E+02  0.177E+02 -0.100E+02  0.467E+02  0.00
</event>
```

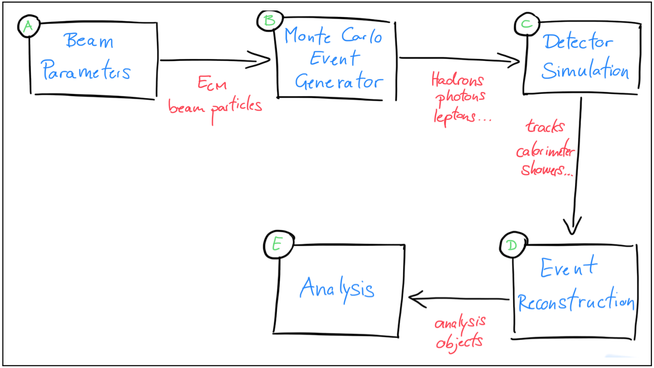
...e.g. as a list of quantum numbers in a Les Houches Event file.

## You can think of an event is several ways...




...e.g. as a list of particles linked by the evolution of the system's state.

The sampling (=event generation) of complicated phase space points $\Phi_n^{(i)}$, and the calculation of $\frac{d\sigma_n}{d\Phi_n}(\Phi_n^{(i)})$ can (with some theory, and some hand-waving) be factorized into smaller problems:




A factorized at LHC, but not for neutrino experiments

C often factorized – but not for decays of long-lived particles

The Monte-Carlo generator landscape is rich! Just to name a few:

**Neutrino physics:**
Genie, GiBUU, NuWro, NEUT...

**Cosmic rays:**
EPOS, QGSJET and SIBYLL

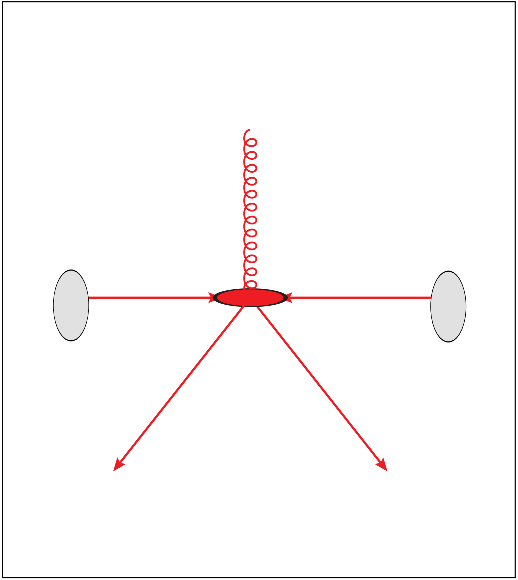
**Heavy ions:**
HIJING, AMPT, JEWEL...

**LHC physics:**
Herwig, Pythia, Sherpa
Madgraph, Whizard, Alpgen...

All of them amazing tools to learn about phenomenology. Focus here $\approx$ LHC-type physics

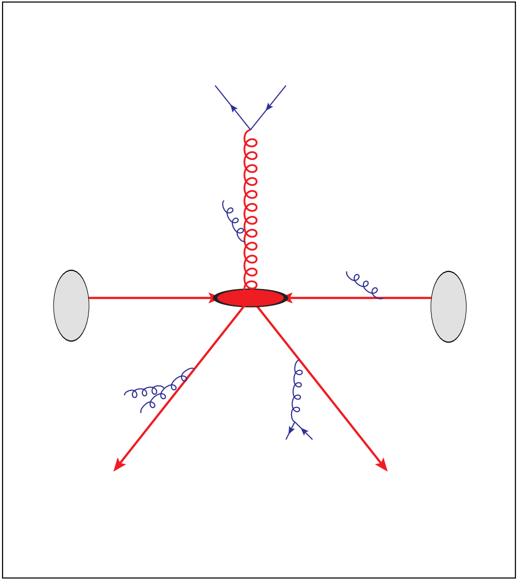
**Exercise:** Get together with friends and chat about an event generator in an unfamiliar field.

A high-energy scattering breaks the beams apart



A high-energy scattering breaks the beams apart

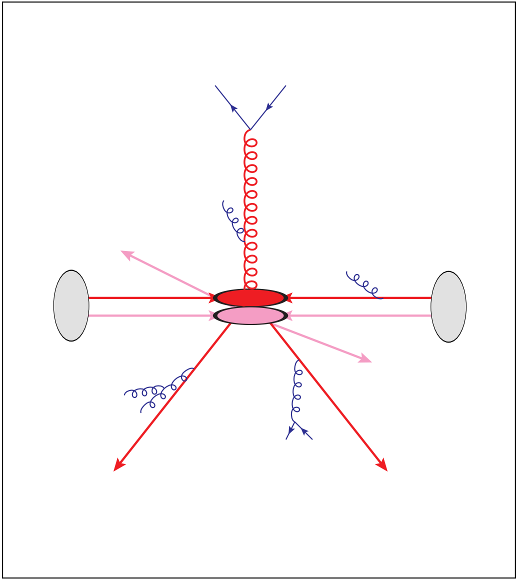
...which initiates a cascade of radiation in the vacuum.



A high-energy scattering breaks the beams apart

...which initiates a cascade of radiation in the vacuum.

Secondary interactions might occur at the same time

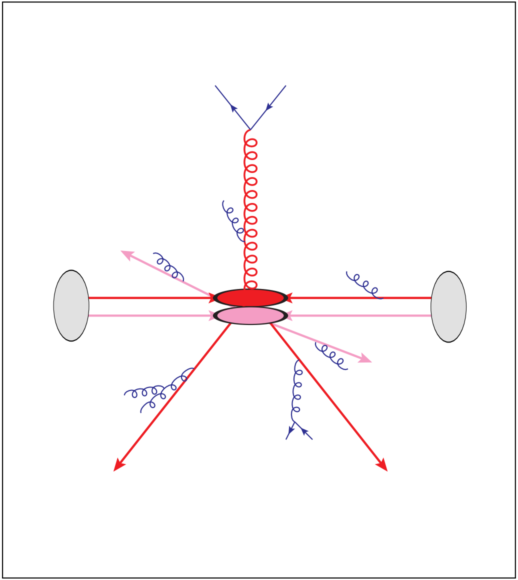


A high-energy scattering breaks the beams apart

...which initiates a cascade of radiation in the vacuum.

Secondary interactions might occur at the same time

...and initiate further radiation "showers".



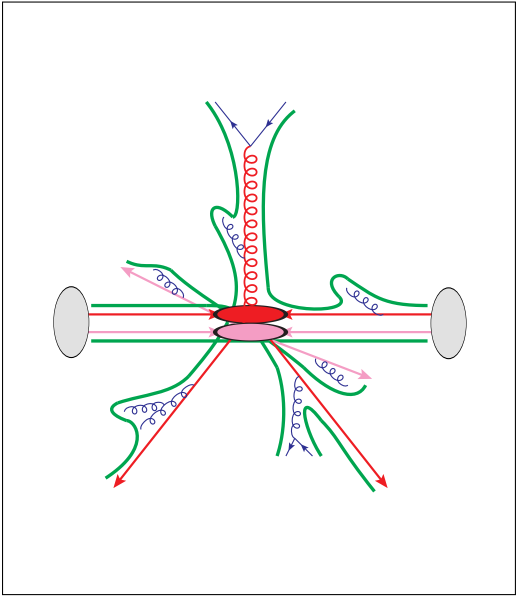
A high-energy scattering breaks the beams apart

...which initiates a cascade of radiation in the vacuum.

Secondary interactions might occur at the same time

...and initiate further radiation "showers".

Confining potentials form, once the $\langle E \rangle$ per particle is small



A high-energy scattering breaks the beams apart

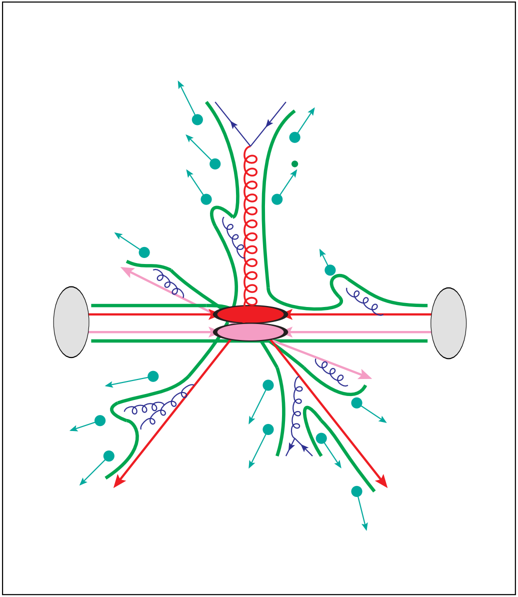
...which initiates a cascade of radiation in the vacuum.

Secondary interactions might occur at the same time

...and initiate further radiation "showers".

Confining potentials form, once the $\langle E \rangle$ per particle is small

...leading to the nucleation of excited or unstable hadrons



A high-energy scattering breaks the beams apart

...which initiates a cascade of radiation in the vacuum.

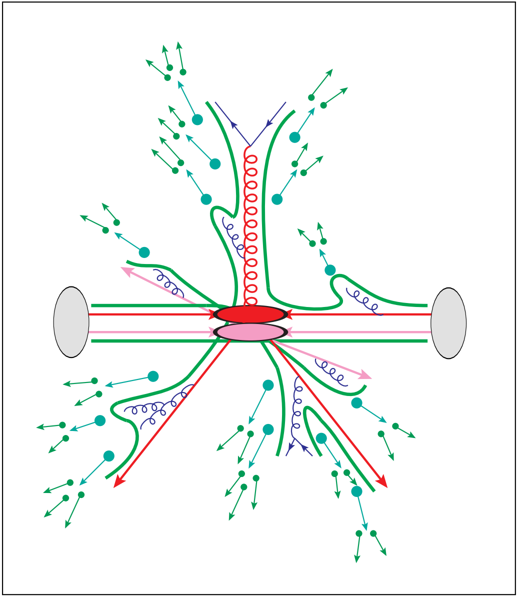
Secondary interactions might occur at the same time

...and initiate further radiation "showers".

Confining potentials form, once the $\langle E \rangle$ per particle is small

...leading to the nucleation of excited or unstable hadrons

...which decay into stable states



A high-energy scattering breaks the beams apart

...which initiates a cascade of radiation in the vacuum.

Secondary interactions might occur at the same time

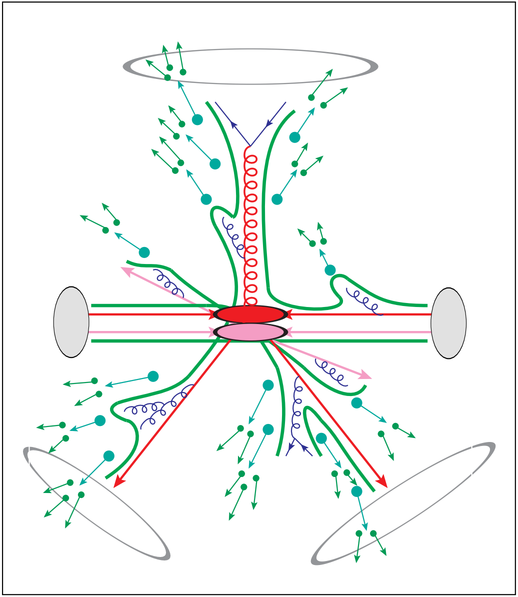
...and initiate further radiation "showers".

Confining potentials form, once the $\langle E \rangle$ per particle is small

...leading to the nucleation of excited or unstable hadrons

...which decay into stable states.

[outside MCEG: interactions with the detector material occur, analysis objects are reconstructed]



From a technical viewpoint, this chain of phenomena looks like

$$
\begin{aligned}
& dP(\text{beams} \to \text{final state}) \\
& \quad = \; dP(\text{beams} \to A, B) \\
& \quad \otimes \; dP(A, B \to \text{few partons}) \\
& \quad \otimes \; dP(\text{few parton} \to \text{many partons}) \\
& \quad \otimes \; dP(\text{many partons} \to \text{hadrons}) \\
& \quad \otimes \; dP(\text{hadrons} \to \text{stable particles})
\end{aligned}
$$

Very high integration dimension. Traditionally, only Monte-Carlo viable
$\to$ Need to learn about numerical methods

Nowadays, deep nets can be used to simulate special cases.

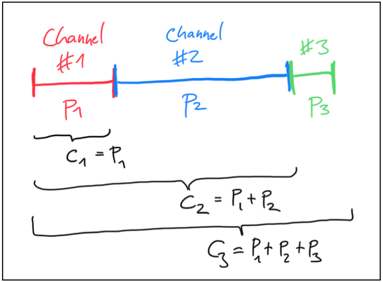
An overview of some basic numerical techniques gives a feeling about how to tackle event generation.

In the following, we'll now look at

- Picking from a probability distribution, a.k.a. inversion sampling
- Hit-or-miss sampling, a.k.a. rejection sampling

...and we'll learn more tricks in the next lectures

Imagine several changes to a state could occur, e.g. different particle decays. How do you pick one?




Draw a random number $R \in [0, 1]$. Pick

channel #1 if $0 < RC_3 < C_1$

channel #2 if $C_1 < RC_3 < C_2$

channel #3 if $C_2 < RC_3 < C_3$

Repeat as often as you like.

Q: Why go through the hassle?
A: Now, the rate of channel #i is given by its population in the sample, and no longer by an "event weight". Every "event" has identical weight ($C_3$).

This is the discrete transformation method. It may be used to pick between different hard scattering processes, decay channels, or for unweighting.

The same algorithm applies when picking a continuous "index" $y$, i.e. picking a random variable according to a distribution (e.g. a phase-space point)

The cumulative distribution becomes

$$C(y) = \int_{-\infty}^{y} dx p(x) \qquad \text{with} \int_{-\infty}^{\infty} dx p(x) = 1$$

which allows using $R \in [0, 1]$ and

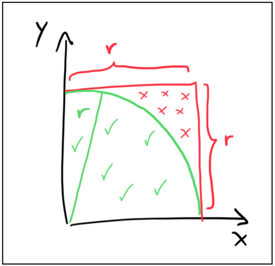
$$C(y) = R \quad \Rightarrow \quad y = C^{-1}(R)$$

This is called inversion sampling.

Often, we're not so lucky that a uniquely invertible primitive function $C^{-1}$ exists ...but we can often still use this method as part of a more flexible algorithm.

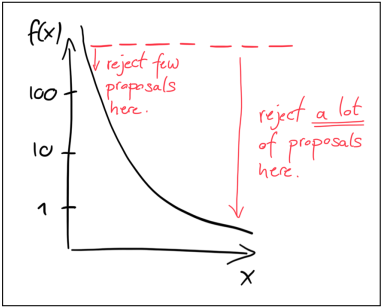
**Exercise:** Generate random variables $x > 0$ with distribution $f(x) = e^{-x}$

We can circumvent the issue with rejection sampling (a.k.a. hit-or-miss).
Basic idea: Use a simple distribution to pick $x$ from, adjust rate once $x$ is generated.




Example: Calculate $\pi$ by random sampling:

- Draw $x, y \in [0, r]$
- Accept pair if $x^2 + y^2 < r^2$
- (fraction of accepted pairs) will be $\propto \pi/4$



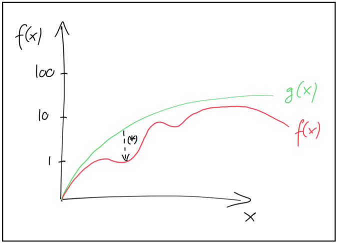

In practise, "uniform sampling" often not sufficient – efficiency very bad!

Rejection sampling will be much more efficient if combined with inversion sampling:

- Assume a simple distribution $g(x) > f(x)$, i.e.

$$f(x) = g(x)\,\underbrace{\frac{f(x)}{g(x)}}_{<1}$$

- Use inversion sampling to draw $x$ from $g(x)$.
- Draw $R \in [0,1]$. Reject $x$ if $\frac{f(x)}{g(x)} < R$




$\Rightarrow$ Accepted $x$ now distributed according to $f(x)$. This algorithm is excessively used in Monte Carlo generators.

Comparison: Uniform sampling

$$\mathrm{var}(f)_{MC} \approx \frac{\mathrm{var}(f)}{\sqrt{N}}$$

error worse in regions of large variance...

Importance sampling

$$\int dx g(x)\frac{f(x)}{g(x)} \approx \langle \frac{f}{g} \rangle \pm \sqrt{\frac{\langle f^2/g^2\rangle - \langle f/g\rangle^2}{N}}$$

**Exercise:** Generate random variables $0 < z < 1-\epsilon$ with distribution $P(z) = \frac{1+z^2}{1-z}$. Hint: Use a simpler numerator to get a simple $g(z)$...

End of lecture 1

Start of lecture 2:

- Phase space and phase space sampling
- Hard scattering cross section
- Factorization of matrix elements

Let's get back to physics for a bit :)
The measurement of an observable is

$$\langle O \rangle = \int d\phi_n \, \frac{d\sigma(A,B \to n \text{ particles})}{d\phi_n} \, O(\phi_n)$$

phase space

differential cross section

value of observable

...so we have to worry about

- sampling phase space points $\Phi_n$
- calculating the differential cross section $\frac{d\sigma_n}{d\Phi_n}$
- evaluating the observable

When sampling phase space,

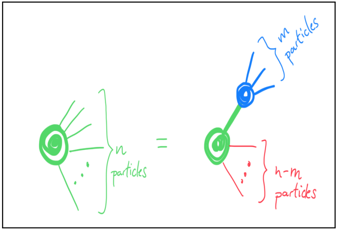
avoid large event weight fluctuations

avoid excessive rejection rate

$\Rightarrow$ Phase space generation separates enthusiasts from experts.

$$d\Phi_n = \left[\prod_{i=1}^{n} \frac{d\vec{p}_i}{(2\pi)^3 2E_i}\right] \delta(p_A + p_B - \sum_1^n p_i)$$

This $(3n-4)$ dimensional integration can be sampled in factorized steps:



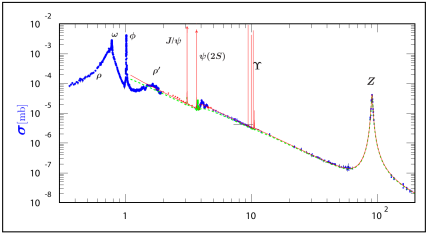

$$d\Phi_n = d\Phi_{n-m+1} \frac{ds_{1m}}{2\pi} d\Phi_m$$

...we can continue until only simple integrations $(d\Phi_2, d\Phi_3)$ remain, and then find a clever parameterization for those.

"Clever" parameterizations need knowledge about $d\sigma$.

Example: Sampling of $d\Phi_2$ stemming from decay of resonance $V$:

$$\frac{d\sigma_n}{d\Phi_n} \supset \frac{M_V\Gamma_V}{\left(\underbrace{(p_1+p_2)^2}_{=\hat{s}} - M_V^2\right)^2 + M_V^2\Gamma_V^2}$$



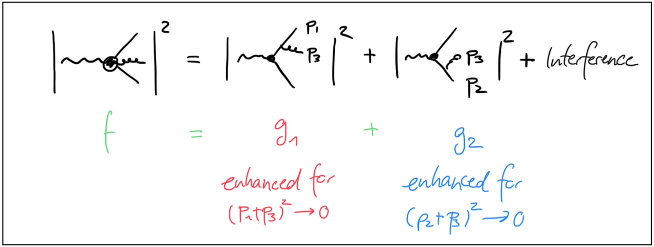

The cumulative function is

$$\begin{aligned} C(\hat{s}_{min}, \hat{s}_{max}) &\propto I(\hat{s}_{max}) - I(\hat{s}_{min}) \\ &= \frac{1}{M_V\Gamma_V}\left[\mathrm{atan}\left(\frac{\hat{s}_{max} - M_V^2}{M_V\Gamma_V}\right) - \mathrm{atan}\left(\frac{\hat{s}_{min} - M_V^2}{M_V\Gamma_V}\right)\right] \end{aligned}$$

Finding the inverse, and using $R \in [0,1]$, we may draw $\hat{s}$ according to

$$\hat{s} = M_V^2 + M_V\Gamma_V \tan\left(M_V\Gamma_V\left[I(\hat{s}_{max}) - RC(\hat{s}_{min}, \hat{s}_{max})\right]\right)$$

Basic thought: know your integrand & generate variables more often close to peaks.




Differential cross sections have a rich structure. In that case, importance sampling can be combined with the discrete transformation method into multichannel sampling:

- Use $f(x) \leq g_1(x) + g_2(x)$
- Choose index $i \in \{1, 2\}$ [using $P_i = \int dx g_i(x)$]
- Draw $x$ from $g_i(x)$. Overall, $x$ is now distributed according to $g_1 + g_2$
- Draw $R \in [0, 1]$, and accept if $(i, x)$ pair if $\frac{f(x)}{g_1(x)+g_2(x)} > R$. Else reject & restart.

NB: also heavily used in parton showers.

**Exercise:** Draw $x$ from the distribution $f(x) = \frac{1}{\sqrt{x(1-x)}}$ using two integration channels.

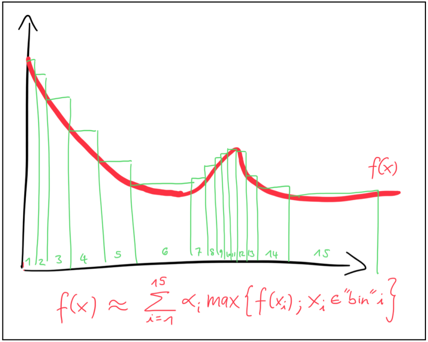
All of these methods require (analytical) knowledge of the differential cross section – which is often hard to come by.

Another way of "generating variables in integration regions where they matter most" is stratified sampling:

◦ Multichannel with $g_i \propto \max\{f\}$ in small integration region (=bin).
◦ Put more bins where variance of $f(x)$ is large.

This is the construction principle of VEGAS.

NB: Need to evaluate the function very often to learn good "integration grids".




Phase-space integrators in MCs are a mix of all of these methods, and recently also more modern machine learning techniques.

Once we have a phase-space point, it's time to evaluate the differential cross section

$$\frac{d\sigma}{d\phi_n} = \int \underbrace{dx_A\, dx_B\, f_A(x_A, \mu_F^2) f_B(x_B, \mu_F^2)}_{\text{distribution of interacting particles in beams}} \frac{1}{F} |\mathcal{M}|^2$$

$\Longrightarrow$ Parametrized measurement

Transition probability

The calculation of the transition probability $|\mathcal{M}|^2$ relies on perturbative methods:

**Pen & paper:** Calculate Feynman diagrams,
use completeness relations to square,
sum over external quantum numbers (helicity, color...)

**Real life:** Assemble helicity amplitudes for fixed color
add & square $\Rightarrow$ less complicated intermediate expressions, better scaling

Color is not a dynamic quantum number, i.e. the color algebra does not depend on parton momenta.

$\Rightarrow$ QCD amplitudes can be stripped of color. For an $n$-gluon amplitude

$$\begin{aligned}
\mathcal{M}(p_1,\ldots,p_n) &= \sum_{\vec{\sigma}\in P(2,n-1)} \mathrm{Tr}(f^{a_{\sigma_2}}\ldots\lambda^{a_{\sigma_{n-1}}}\cdots\lambda^{a_{\sigma_n}})M(p_1,p_{\sigma_2}\ldots,p_{\sigma_{n-1}},p_n) \\
&= \sum_{\vec{\sigma}\in P(2,n)} \mathrm{Tr}(\lambda^{a_1}\lambda^{a_{\sigma_2}}\cdots\lambda^{a_{\sigma_n}})M(p_1,p_{\sigma_2}\ldots,p_{\sigma_n}) \\
&= \ldots \text{ and many more ways of } \underline{\text{color ordering}} \ldots
\end{aligned}$$

So precalculate the $\mathrm{Tr}(\cdots)$ color factors, and recycle $M$ as much as possible.
Alternatively, can fix color at each vertex by random sampling (a.k.a. color dressing)

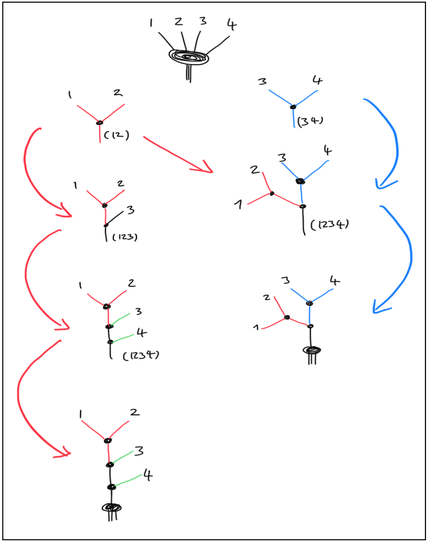
In matrix-element generators (MADGRAPH, COMIX...), the matrix elements $M$ are calculated from the outside $\rightarrow$ inwards.

You can assign the helicities first and contract spinors (polarization vectors) with fixed helicity (polarization).

◦ very efficient due to recycling parts of the amplitude.

◦ basis of helicity amplitudes methods

◦ whole research field of finding efficient method to construct amplitudes. By now, basically solved (?)



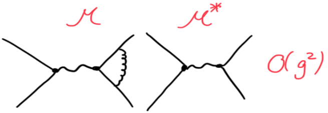
But tree-level calculations on their own are questionable: **Beware** of how to count coupling powers (and particle number).

Infrared (IR) singularities abound in tree-level diagrams ...because the "particle number" operator is **ill-defined** in perturbative QFT!

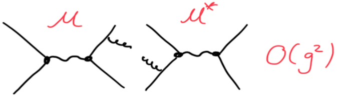
Singularities cancel between *different multiplicities* when introducing virtual corrections.

Notes:
- The result are *inclusive* cross sections.
- Measurements that ensure singularity cancellation are called *IR safe*.

Example:




... divergent due to loop integral

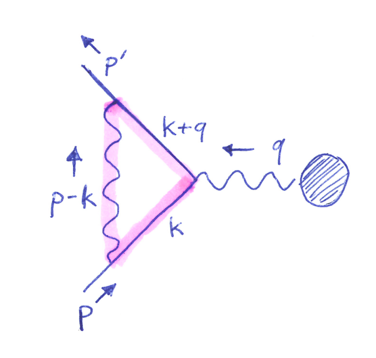



... divergent due to extra phase space integral

KLN theorem: Divergences cancel!

Virtual corrections include loop integration.

This integration is typically not performed numerically. Instead, map integrals onto master integrals after a lot of algebra. Tough problem – be clever!




Integrands $\sim \frac{\text{polynomials in the momenta}}{\prod_i \text{simple polynomials/monomials}}$

Can be reduced to easier integrals, e.g.

- find ways to cancel numerators, e.g. subtract & add sum of numerators
- many new coupled equations
- e.g. use Gauss-elimination inspired methods to solve

The devil's in the details, but $\sim$ solved at 1-loop. General algorithms implemented in ME generators or loop providers.

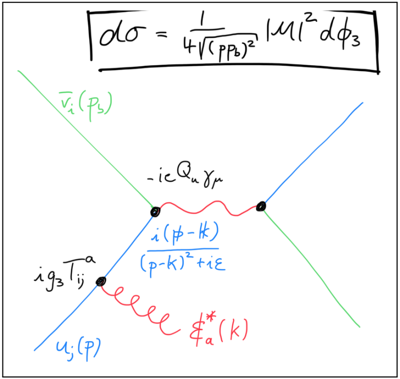
Infrared singularities in multi-parton amplitudes have a profound consequence: Nature will dress partons with many more partons to take advantage of the enhancement!

For small $p_{\perp\text{gluon}}$ and $E_{p-k} \approx zE_p$, the internal quark is almost on-shell, and

$$\frac{i(\not{p}-\not{k})}{(p-k)^2+i\varepsilon} \approx \frac{u(p_a)\bar{u}(p_a)}{p_a^2}$$

$$d\Phi_3 \approx d\Phi_2 \frac{d\phi dz dp_\perp}{4(2\pi)^2(1-z)}$$

$$\frac{1}{4\sqrt{(pp_b)^2}} \approx z\frac{1}{4\sqrt{(p_a p_b)^2}}$$




All components of the x-section factorize, and we're left with

$$d\sigma_3 \approx d\sigma_2 \int d\phi dz dp_\perp P(\phi, z, p_\perp)$$

where the universal splitting function $P$ contains the singularities due to gluon emission.

Once the divergences have been factorized, we may attempt to calculate an observable to next-to-leading order accuracy

$$\begin{aligned}\langle O\rangle_{\mathrm{NLO}} &= \int d\Phi_n \left\{\frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} + \frac{d\sigma_n^{\mathrm{Virt}}}{d\Phi_n} + \frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} \otimes \int d\Phi_1 S\right\} O(\Phi_n) \\ &+ \int d\Phi_{n+1} \left\{\frac{d\sigma_{n+1}^{\mathrm{Tree}}}{d\Phi_{n+1}} O(\Phi_{n+1}) - \frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} \otimes S O(\Phi_n)\right\}\end{aligned}$$

where $\frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} \otimes S$ captures the singularities of real-emission and – by the KLN theorem – virtual corrections alike.

This allows numerical predictions for IR-safe observables, i.e. when $O_{n+1} \to O_n$ when the additional particle becomes unresolvable.

However, it does not allow the generation of "NLO events".

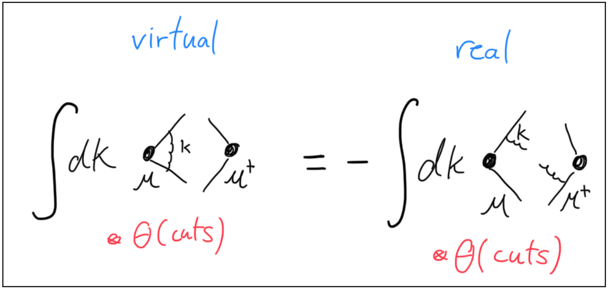
End of lecture 2

Start of lecture 3:

- The (Sudakov) veto algorithm
- Parton showers and very basic matching
- Multiparton interactions

Remember the KLN theorem: Infrared singularities arising in real-emission diagrams cancel against alike divergences in virtual corrections.[1]

For the (most) enhanced parts, we can devise a radical interpretation of KLN:




"The rate for # particles remaining the same is (negative) the rate for the # particles increasing at any scale $t$ – even in the presence of cuts/regularization".

This is the first building block of a parton shower.

[1] This is a popularized account; there are subtleties. Kinoshita's paper highly recommended.

The behavior of partons is similar to that of radioactive elements.

The $\#$ particles $n$ can only change $n \to n+1$ (due to decay or splitting) at scale $t$ if it has not already changed at $t' > t$.

The probability to not change in a finite interval $\Delta t$ is

$$1 - \Delta t P(t)$$

where $P$ is the splitting kernel containing the enhanced parts of the real correction. This is simply statement about unitarity: The rate of no change and the rate of all possible changes add to unity.

The probability not to change in any very small sub-interval $\Delta t/n$ is

$$\left(1 - \frac{\Delta t}{n} P(t)\right)^n \quad \xrightarrow{n\to\infty} \quad \exp\left(-\int_0^{\Delta t} dt P(t)\right)$$

This exponential suppression of not splitting is called the <u>Sudakov factor</u>.

[no splitting] $\leftrightarrow$ [fixed $\#$ particles]. Thus, the Sudakov introduces virtual corrections.

Combined, the decay/splitting probability at scale $t$ is

$$\mathcal{P}(t) = P(t) \exp\left( -\int_0^t d\bar{t} P(\bar{t}) \right) = P(t)\Delta(t)$$

Retains a memory: the "next" decay may only happen at scale $t$ if it had not happened before. Conservation of total probability means that the process develops a "memory".

Note that this means that the no-decay probability follows the differential equation

$$\underbrace{-\frac{d\Delta(t)}{dt}}_{\text{change of \#particles by decay}} = P(t)\Delta(t) \qquad \leftrightarrow \qquad -\frac{d\ln\Delta(t)}{dt} = P(t)$$

It is possible to rewrite the DGLAP equation in this form:

$$\frac{df(x,t)}{dt} = \int_x^1 \frac{dz}{z} \frac{\alpha_s}{2\pi} \frac{1}{t} \left[P(z)\right]_+ f\left(\frac{x}{z}, t\right) \quad \Longleftrightarrow \quad \frac{d\ln\left(\Pi(x, t_{\mathrm{Max}}, t)\right)}{dt} = \int_x^{1-\varepsilon} \frac{dz}{z} \frac{\alpha_s}{2\pi} \frac{1}{t} P(z) \frac{f\left(\frac{x}{z}, t\right)}{f(x,t)}$$

We can use differential equation to define an inversion sampling algorithm that correctly includes the "memory":

$$-\frac{d\ln\Delta(t)}{dt} = P(t) \qquad , \quad \Delta(t) = \exp\left(-\int_0^t P(t)\right) = \exp\left(-F(t) + F(0)\right)$$

Note that $\Delta(t)$ is the cumulative function of $\frac{d\Delta}{dt}$, i.e. of the probability density that defines the distribution of $t$ values. Thus, draw $R \in [0,1]$ and

$$R = \Delta(t) = \exp\left(-F(t) + F(0)\right) \qquad \Rightarrow \qquad t = F^{-1}(F(0) - \ln R)$$

...and we've produced a sample of decay scales (with memory). This is the basic algorithm used in parton showers.

In this way, parton showers can solve evolution equations. The result incorporates exponential Sudakov factors, i.e. is an all-order "resummed" prediction.

However, for most cases of interest, $F^{-1}$ does not exist – rejection sampling to the rescue. However, it's important to retain the memory.

This is achieved by the Sudakov veto algorithm:

- Assume a simple distribution $g(t) > f(t)$, i.e. $f(t) = g(t)\frac{f(t)}{g(t)}$

1 Set $t_0 = 0$

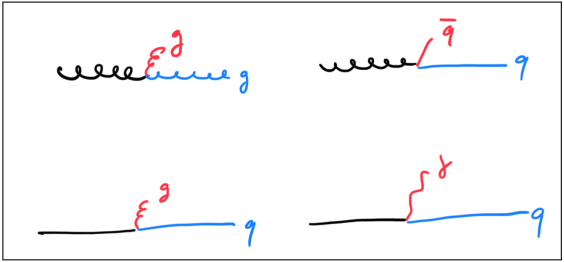
2 Use inversion sampling to draw $t$ from $g(t)$ (using $t_0$ as lower bound).

3 Draw $R \in [0, 1]$. Reject $t$ if $\frac{f(t)}{g(t)} < R$
Wrong: Restart at 1 $\leftarrow$ this would erase the memory!
Correct: Set $t_0 \to t$, restart at 2.

In this way, parton showers can solve complicated evolution equations.

NB: Typically, the algorithm is rearranged to move from large $t$-values ($\mathcal{O}(\mu_f)$) to small $t$-values ($\mathcal{O}(1\mathrm{GeV})$).

In nature, many different "decay channels" may compete




- could use Sudakov veto algorithm with $f(t) = f_1(t) + f_2(t)$
  then pick channel with proportions $f_1(t) \; : \; f_2(t)$ [discrete transformation method]

- another algorithm is winner-takes-all: generate $t_1$ as if $f = f_1$, and $t_2$ as if $f = f_2$
  then pick the channel $i$ with the smallest $t_i$ to happen[1]

The "right" competition algorithm can be very important for efficiency/speed.

[1] If algorithm is rearranged to move from large $\to$ small $t$, then pick the $i$ with the largest $t_i$.

With this, we're finally able to construct a parton shower, since

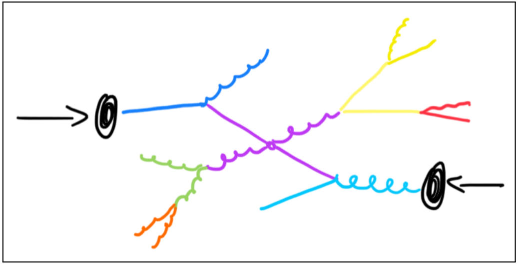
- Within the simplest approximation, the splitting functions are universal, and fully factorized from the "hard" cross section
- Within the simplest approximation, decays are independent (apart from being ordered in a decreasing sequence of scales)

⇒ The splitting process can be iterated, with the result after $n$ splittings forming the "hard" scattering for the $(n+1)$th emission.

The effect of the shower $\mathcal{F}$ on an observable $O$ is, symbolically,

$$
\begin{aligned}
\mathcal{F}_n(O, \Phi_n, t_{\max}, t_{\min}) &= \Delta_n(t_{\max}, t_{\min}) O(\Phi_n) \\
&+ \int_{t_{\min}}^{t_{\max}} d\Phi_1 \Delta_n(t_{\max}, t) P(\phi, z, t) \mathcal{F}_{n+1}(O, \Phi_{n+1}, t, t_{\min})
\end{aligned}
$$

Through $\Delta$, the shower is an "all-order" calculation, and each term in the formula is individually finite.



The parton shower will develop from high propagator virtuality and large angles to small virtuality and angle.

Several choices will influence the sequence: how are the emissions ordered? how is the phase space for emissions mapped? how are quantum interferences approximated?

The most prominent features of the event will be determined by the hardest emissions.

The different choices can give large uncertainties in the rate & distribution of hard jets. Best to improve the event generator for hard jets $\Rightarrow$ goal of matching & merging

Compare a next-to-leading order calculation and an expanded version of the shower:

$$
\begin{aligned}
\langle O\rangle_{\mathrm{NLO}} &= \int d\Phi_n \left\{ \frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} + \frac{d\sigma_n^{\mathrm{Virt}}}{d\Phi_n} + \frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} \otimes \int d\Phi_1 S \right\} O(\Phi_n) \\
&+ \int d\Phi_{n+1} \left\{ \frac{d\sigma_{n+1}^{\mathrm{Tree}}}{d\Phi_{n+1}} O(\Phi_{n+1}) - \frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} \otimes S O(\Phi_n) \right\} \\
\langle O\rangle_{\mathrm{PS}} &= \int d\Phi_n \left\{ \frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} - \frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} \int_{t_{\min}}^{t_{\max}} d\Phi_1 P(\phi, z, t) + \mathcal{O}(\alpha^2) \right\} O(\Phi_n) \\
&+ \iint_{t_{\min}}^{t_{\max}} d\Phi_n d\Phi_1 \left\{ \frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n} P(\phi, z, t) + \mathcal{O}(\alpha^2) \right\} O(\Phi_{n+1})
\end{aligned}
$$

As expected, the calculations overlap (the shower gives an approximation of NLO).

Suggestion: Subtract the PS result from the NLO, and use the result as starting point of the shower, instead of $\frac{d\sigma_n^{\mathrm{Tree}}}{d\Phi_n}$

The big advantage of this suggestion is that we can (finally!) generate NLO events – just add a couple for zeros:

$$
\begin{aligned}
\langle O\rangle_{\rm NLO} \;=\; & \int d\Phi_n \Big\{ \frac{d\sigma_n^{\rm Tree}}{d\Phi_n} + \frac{d\sigma_n^{\rm Virt}}{d\Phi_n} + \frac{d\sigma_n^{\rm Tree}}{d\Phi_n} \otimes \int d\Phi_1 S \\
& \qquad - \frac{d\sigma_n^{\rm Tree}}{d\Phi_n} \otimes \int d\Phi_1 S + \frac{d\sigma_n^{\rm Tree}}{d\Phi_n} \int_{t_{\rm min}}^{t_{\rm max}} d\Phi_1 P(\phi, z, t) \Big\} O(\Phi_n) \\
+ & \int d\Phi_{n+1} \Big\{ \frac{d\sigma_{n+1}^{\rm Tree}}{d\Phi_{n+1}} - \frac{d\sigma_n^{\rm Tree}}{d\Phi_n} P(\phi, z, t) \Theta(t_{\rm min}, t_{\rm max}) \Big\} O(\Phi_{n+1}) \\
+ & \int d\Phi_n \frac{d\sigma_n^{\rm Tree}}{d\Phi_n} \int_{t_{\rm min}}^{t_{\rm max}} d\Phi_1 P(\phi, z, t) \Big\{ O(\Phi_{n+1}) - O(\Phi_n) \Big\}
\end{aligned}
$$

Both $\{\cdots\}$ are separately finite. $\{\cdots\}$ is just the 1st-order expansion of the shower – which we would produce by showering the three first lines.

Removing $\{\cdots\}$ allows to generate events. Showering the result produces a consistent NLO matched calculation, in the MC@NLO approach.

Implementing the matching formula naively can have disadvantages:

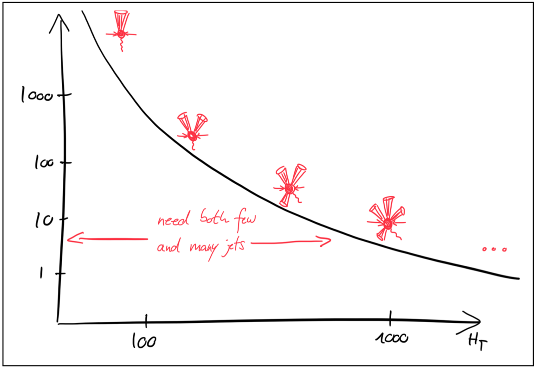
- The contributions are not necessarily positive definite
- The shower might act over an uncomfortably large phase space region
- Looking carefully, some of the differences might not be completely free of singularities
- ...

There will be a whole lecture devoted to matching & merging.

NLO matched calculations will describe one additional jet with tree-level accuracy.

Analyses of experimental data often depend on multi-jet final states, e.g. to expose Beyond-the-SM signals.

In this case, NLO (or NNLO or N3LO) matching is often not sufficient.




Instead, consistently "stack" simpler (tree-level or NLO) calculations on top of each other, with the help of the shower. This defines a merging scheme.

The task for a tree-level merging scheme is to describe events for

[simple final state $X$] + $\{0, 1, \ldots, N\}$ well-separated jets

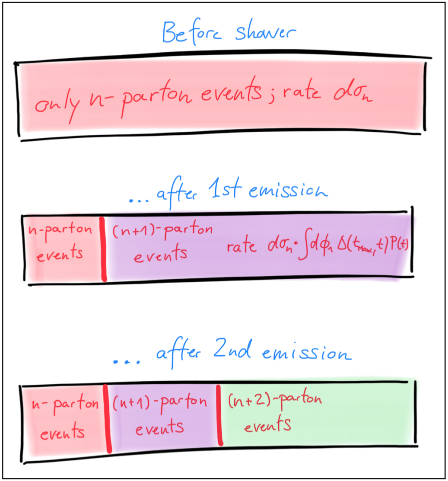
through a combined calculation, with tree-level accurate $X + \{0, 1, \ldots, N\}$ parton rates, and the jets' structure determined by the parton shower.

Simply adding several showered tree-level calculations is inconsistent, since the results overlap.

Take inspiration from PS to avoid overlap:

- Showers produce (all-order) real emission corrections
- The lower-multiplicity (inclusive) cross section is preserved by removing the emission rate from the rate of lower-multiplicity events.

An idealized merging method could handle overlap in exactly the same way.



The chain of reasoning is

$$
\begin{aligned}
& \int d\Phi_n O(\Phi_n) \frac{d\sigma_n}{d\Phi_n} + \int d\Phi_{n+1} O(\Phi_{n+1}) \frac{d\sigma_{n+1}}{d\Phi_{n+1}} + \ldots \\
\xrightarrow{\text{make } (n+1) \text{ PS-like}} \quad & \int d\Phi_n O(\Phi_n) \frac{d\sigma_n}{d\Phi_n} \\
& + \int d\Phi_{n+1} O(\Phi_{n+1}) \frac{d\sigma_{n+1}}{d\Phi_{n+1}} \Delta_n(t_n, t_{n+1}) + \ldots \\
\xrightarrow{\text{remove real from Born}} \quad & \int d\Phi_n O(\Phi_n) \frac{d\sigma_n}{d\Phi_n} - \int d\Phi_{n+1} O(\Phi_n) \frac{d\sigma_{n+1}}{d\Phi_{n+1}} \Delta_n(t_n, t_{n+1}) \\
& + \int d\Phi_{n+1} O(\Phi_{n+1}) \frac{d\sigma_{n+1}}{d\Phi_{n+1}} \Delta_n(t_n, t_{n+1}) + \ldots \\
\overset{\text{make more PS-like}}{\approx} \quad & \int d\Phi_n O(\Phi_n) \frac{d\sigma_n}{d\Phi_n} \Delta_n(t_n, t_{\min}) + \int d\Phi_{n+1} O(\Phi_{n+1}) \frac{d\sigma_{n+1}}{d\Phi_{n+1}} \Delta_n(t_n, t_{n+1}) \\
\overset{\text{effective description}}{\approx} \quad & \int d\Phi_n O(\Phi_n) \frac{d\sigma_n}{d\Phi_n} [\text{veto events with more than } n \text{ hard jets}] \\
& + \int d\Phi_{n+1} O(\Phi_{n+1}) \frac{d\sigma_{n+1}}{d\Phi_{n+1}} [\text{veto events jets harder than in ME}] + \ldots
\end{aligned}
$$

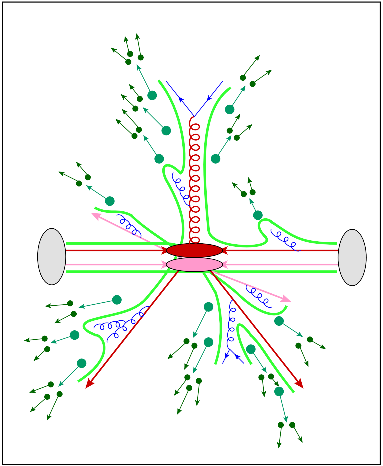
Several tree-level and NLO merging prescriptions have been implemented, with various approximations of "preserving the inclusive cross section".

Let's take a step back, and look at the bigger picture.

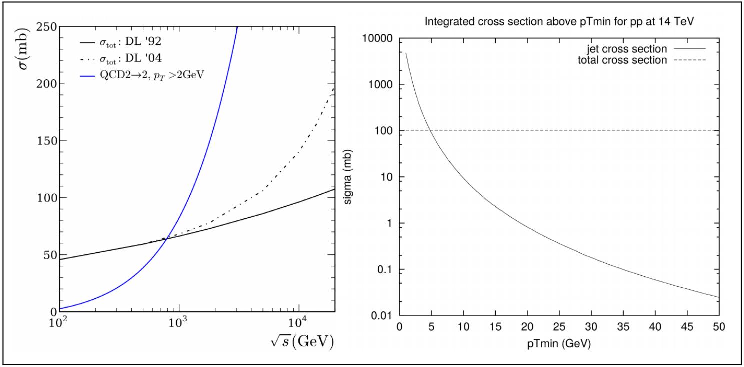
At hadron colliders, the initial state is complex.

There is no reason to expect only one parton-parton interaction to occur.

Does the inclusion of multiple interactions change the inclusive single-interaction cross section?

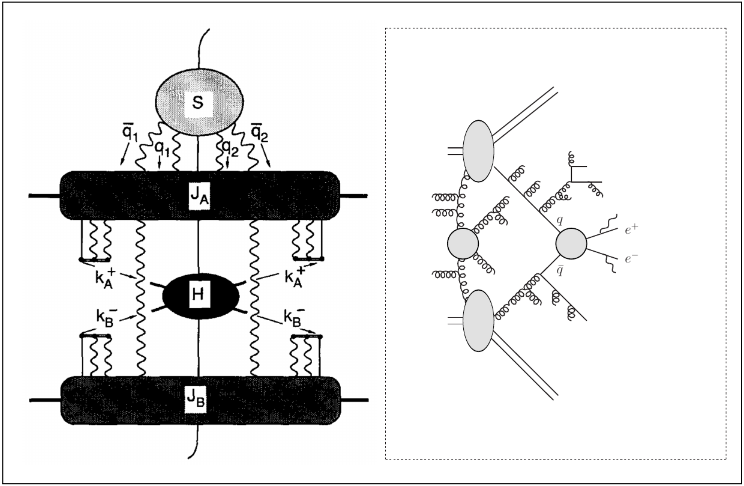


The naive inclusive cross section for parton-parton scattering is often divergent already at leading order.




This simply hints at a too literal interpretation of the concept of "inclusive cross section".

The crux lies in the definition of the parton distribution functions: These give the inclusive probability to find a parton at $x$m with all other interactions above $x \approx \frac{p_{\perp \min}}{E_{\mathrm{CM}}}$ integrated out.

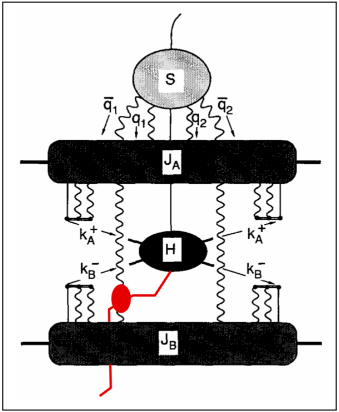


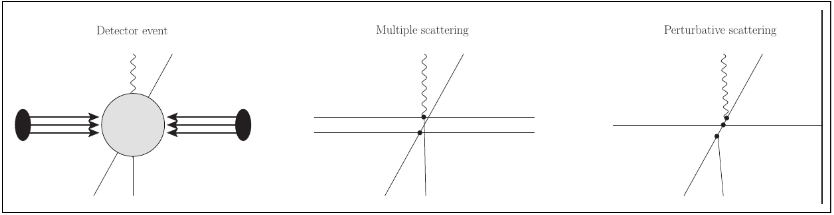
Detailed enough measurements will probe the integrand, i.e. be sensitive to multiple interactions.

In this case, we should interpret the cross section as

$$\sigma_{\text{inclusive}}(p_{\perp\min}, E_{\text{CM}}) = \langle n(p_{\perp\min})\rangle \cdot \sigma_{\text{inelastic}}(p_{\perp\min}, E_{\text{CM}})$$

$$\sigma_{\text{inelastic}}(p_{\perp\min}, E_{\text{CM}}) < \sigma_{\text{total}}(E_{\text{CM}})$$






Take a four-jet event as an example:

- jets might not be separated and emerge from showering
- jets might be well-separated and emerge from one scattering
- jets might be well-separated and emerge from two scatterings

It is important to understand the measurement in order to understand the cocktail of phenomena.

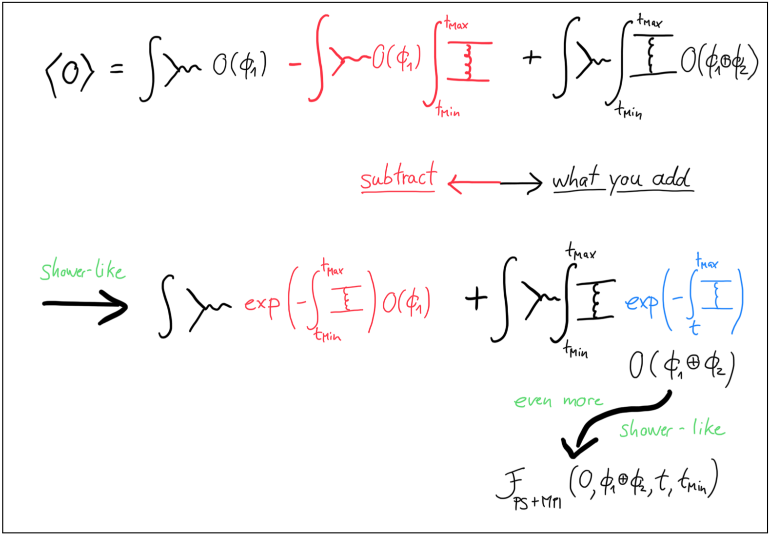
**Argument:** Want inclusive x-section to be calculable in perturbation theory + PDFs. Multiple interactions should not change this. Simply overlaying scatterings will not work.

**Realization:** Multiple interactions are not additive – just as tree-level calculations are not!

**Solution:** The rate for not having a second interaction is correlated with the rate for having a second interaction.

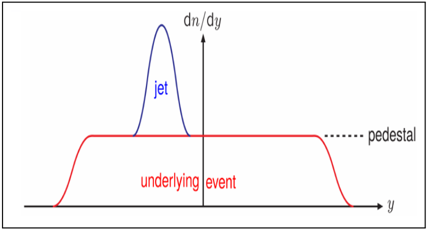
Note the similarity to loops $\leftrightarrow$ reals and shower emission rate $\leftrightarrow$ Sudakov factor

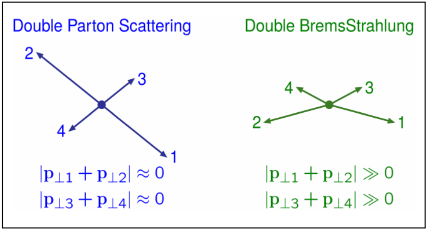
Unitarity (= conservation of probability) suggests a phenomenological model:




In fact, this is basically the same algorithm as for parton showering.

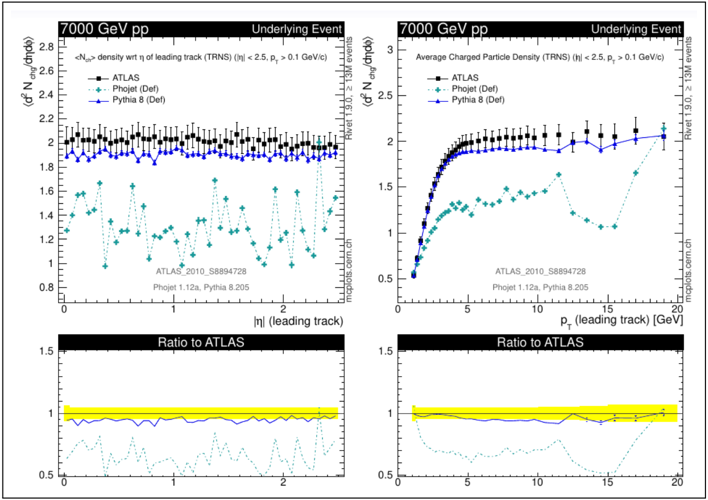
We may expose multiple interaction topologies using jet (or particle) correlations:




Multiple interactions $\approx$ fill the regions between the hardest jets.

Data indeed shows a mostly uniform rapidity coverage (blue: w/ MPI; cyan: no MPI)




...but also that harder primary particles (i.e. interactions) lead to more secondary interactions.

So as always, the proof is in the pudding

- no reason to expect primary and secondary partons to be in the "same place" in the proton
  Multiple interactions introduce impact parameter dependence
- some inelastic scattering cross sections (evaluated at fixed order) still require regularization for small momentum transfer
- the correlation and competition between multiple interactions and showers is non-trivial

Excellent field to apply your wit. Dedicated lecture later in the school.

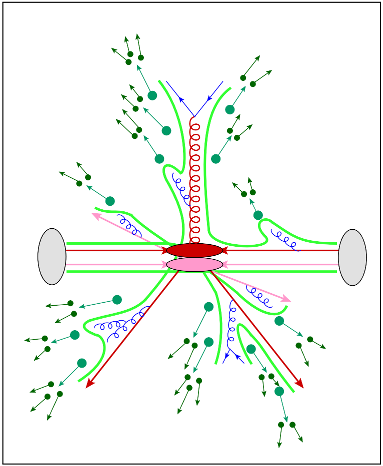
End of lecture 3

Start of lecture 4:

- Color reconnection
- Converting partons to hadrons (=hadronization)
  (hadron and particle decays)

We started from an overview of event generation at microscopic detail.

$$
\begin{aligned}
& dP(\text{beams} \to \text{final state}) \\
= \;& dP\text{beams} \to A, B) \\
\otimes \;& dP(A, B \to \text{few partons}) \\
\otimes \;& dP(\text{few parton} \to \text{many partons}) \\
\otimes \;& dP(\text{many partons} \to \text{hadrons}) \\
\otimes \;& dP(\text{hadrons} \to \text{stable particles})
\end{aligned}
$$



The last steps are typically responsible for a vast increase in particle multiplicity.

Phenomenological models & data parameterization are employed here.

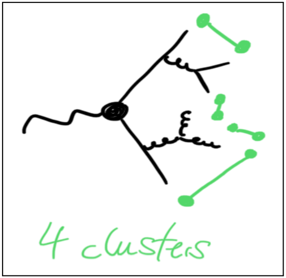
Nobody has solved strong-coupling QFTs yet. Until then, we require a model to translate set of partons to sets of hadrons.

So how do partons coalesce?

| Individual partons $\rightarrow$ hadrons | All partons $\rightarrow$ all hadrons | Subset of partons $\rightarrow$ subset of hadrons |
|---|---|---|
| ...as e.g. introduced by Feynman & Field<br><br>What about flavor and momentum conservation?<br><br>Not ideal, but still partially used ($\sim$ fragmentation functions) | In conflict with perturbative QCD (& non-universal)<br><br>Difficult to imagine "jetty" behavior.<br><br>Still useful for extremely high-multiplicity $\oplus$ low $\langle E \rangle$ events | Middle ground between the extremes<br><br>Basis of the most successful high-energy physics models – the string and cluster model.<br><br>Main approach in Event Generators. |

Partons "close to" each other hadronize coherently.

There are two main schools of thought of what "close to" means:





**Cluster hadronization**

- create clusters from color-connected partons (gluons branch to two quarks)
- invoking color preconfinement

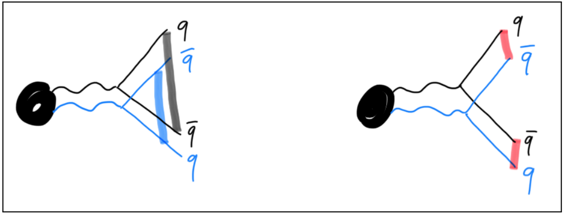
**String hadronization**

- create strings from color string, with gluons "stretching the string" locally
- invoking non-perturbative insights

Note already here: real-life models borrow traits and phenomena from both – depending e.g. on available phasespace for hadrons.

The notion of closeness determines which partons hadronize collectively.

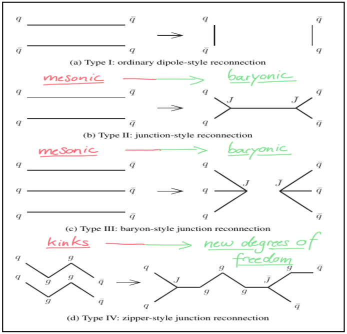
In busy systems – like LHC collisions – definitions of closeness are typically less obvious



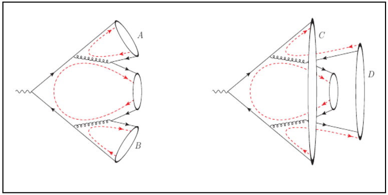

Previously independent systems might undergo color reconnection, e.g. to neutralize flavor more locally.

Color reconnection is not a completely random process: Minimizing some measure of energy ($\sim \sum_{i,j \in \text{partons}} \ln(p_i p_j)$) is likely to occur.

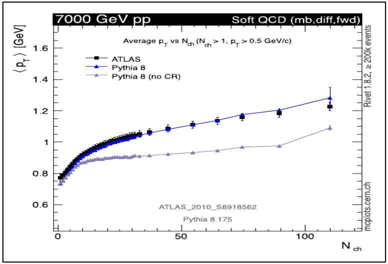
The perturbative picture of color reconnection imagines ultra-soft gluons rearranging color. CR occurs before forming the initial state (cluster/strings) for hadronization.




arXiv:1505.01681: CR can introduce new baryon production mechanisms



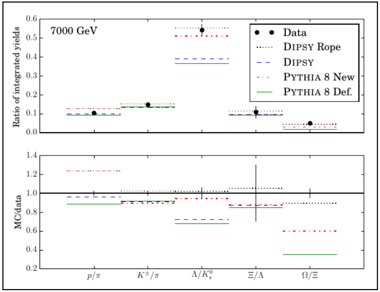

arXiv:1206.0041: CR can significantly alter the initial cluster mass distribution.

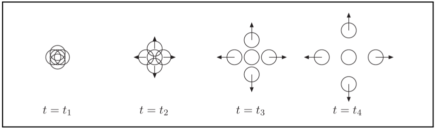



Color reconnection is needed to describe data. Color reconnection models introduce a lot of unknowns.

The non-perturbative picture of color reconnection imagines strings undergoing non-perturbative dynamics:

- Strings interact by fusing, repelling, swapping string ends before settling into a steady state for hadronization
- Implement models for individual non-perturbative effects




arXiv:1612.05132: Repelling strings (a.k.a. shoving) produces a "pressure gradient", thus producing collective effects.

arXiv:1710.04464: Combined strings (a.k.a. ropes) have a higher tension, i.e. smaller suppression for heavy hadrons.

Perturbative and non-perturbative pictures may lead to similar results. Reality will be a mixture of both.

It is an unspoken assumption of CR models that the total cross section is unaffected by any rearrangement.

Similarly, the transition partons $\rightarrow$ hadrons does not change the total cross section, i.e. colored partons coalesce into hadrons with unit probability

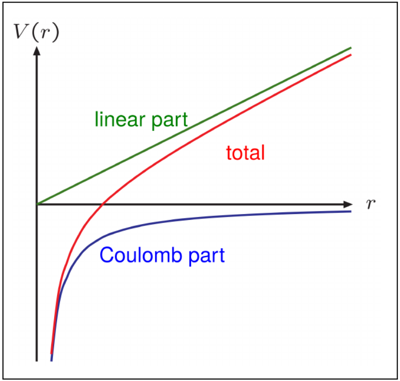
Having discussed the sets of partons that collectively hadronize, we may now discuss the string (PYTHIA) and cluster (HERWIG, SHERPA) models.

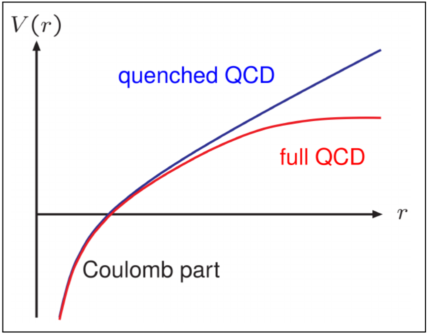
Although non-perturbative QCD is hard, some results are known e.g. from lattice QCD.

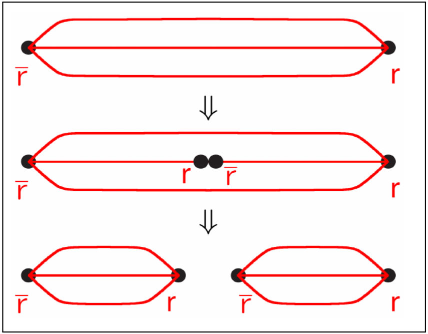
The potential between two quarks is linear, since the force per unit length is constant.

The force is confining, and similar to the force on a stretched string.

This is the basis of the string model.






In reality, the force between quarks will drop eventually: It is energetically favorable for the string to break.

Mesons are ≈ oscillating strings – so-called yo-yo modes.

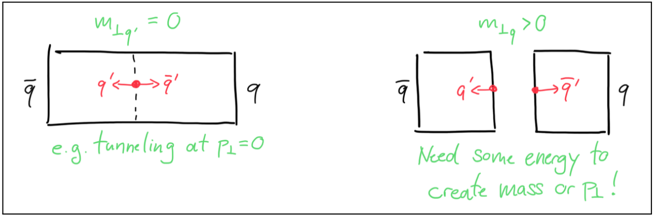
High-energy strings break through pair creation.

Strings break through $f\bar{f}$ creation through a tunneling mechanism (Heisenberg & Euler, Schwinger – yes, that old).

QCD strings break through $q\bar{q}$ creation with tunneling probability

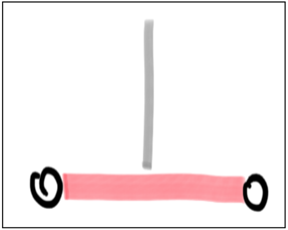
$$\mathcal{P} \propto \exp\left(-\frac{\pi m_{\perp q}^2}{\kappa}\right) \qquad \kappa = \text{string tension}$$

Tunneling of heavy quarks suppressed by $m_{\perp q}^2$ dependence. $c\bar{c}$ almost negligible.



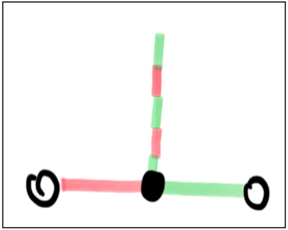

High transverse momentum suppressed. Breaking yields ≈ back-to-back particle production in string CM frame.

QCD contains both quarks and gluons, i.e. realistic model should consider gluons as well.



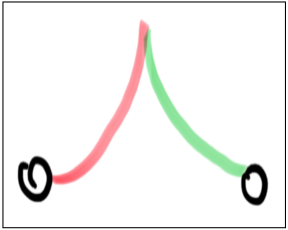
Gluon does not change color field.

Very unlikely



Gluon induces new type of string, attached by junction.

Adds new, unknown parameters.

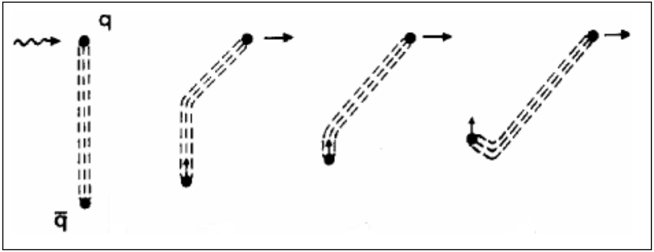


Gluon is a “kink” on the string.

Kinks are present on massless relativistic strings.

No additional parameters needed.

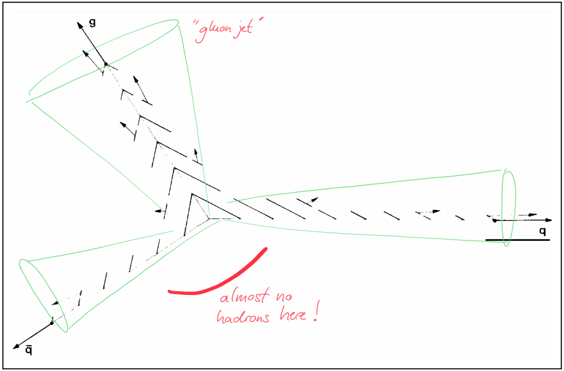
A “kink” is a large, instantaneous momentum transfer at the initial time. It stretches the string in some direction.




The kink is connected to two string segments. Thus, it looses energy twice as fast as the endpoints, in accordance with QCD, where $C_A/C_F \xrightarrow{N_C \to \infty} 2$

Causality dictates that the string + kink system fragment like any other string.

The interpretation of gluons as kinks has an important consequence: the string effect




There are almost no hadrons in the region opposite the jet formed by the gluon kink.

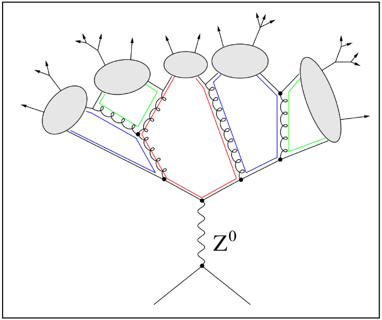
The gluon kink and the quark endpoints act coherently to deplete that region.

Coherence effects are already found in perturbative QCD: Gluon production at comparable angles is suppressed by destructive interference.

Thus, color-singlet parton pairs end up "close" in phase space. This is called preconfinement. Preconfinement mimics the string effect at perturbative level.

This is the basis of the cluster model:

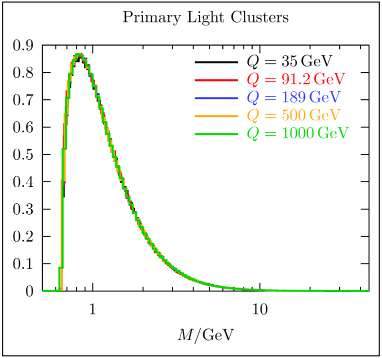
- use perturbative calculation that enforces coherence
- convert gluons to $q\bar{q}$ pairs with heuristic model
- collect $q\bar{q}$ pairs into color-singlet clusters
- clusters decay isotropically into two hadrons
- heavy clusters need to be treated separately



Indeed, the mass of color-singlet clusters is very small, and independent of the CM energy $Q$. Thus, the cluster model is relatively universal.

Light clusters decay into resonances & stable hadrons with $\approx$ flat phase-space distribution. Heavy hadron production is thus suppressed.

However, long tail to high cluster mass values.

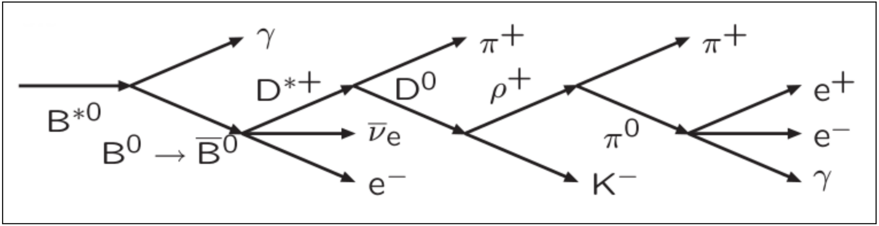



Heavy clusters undergo fission to lighter clusters ($\rightarrow$ similar to string breaking)
$\approx$ 15% of primary clusters split
$\approx$ 50% of hadrons emerge from split clusters

It would now be customary to compare string and cluster models. I'll pawns this off to subsequent lecturers.

We are now approaching the final steps in the event generation chain.

Hadronization models often produce excited hadrons, which will decay within typical detectors. For example:




Note that some of these decays will leave displaced vertices, which may be important to "tag" heavy jets.

Majority of particles will be produced here; comprehensive machinery very important:
◦ Implement as many hadronic matrix elements as possible, especially for $\tau$
◦ Include as many QED effects as possible
◦ Use PDG decay tables for rest. If incomplete, be creative.

# Let us end on “If incomplete, be creative”.

## Summary of the lectures: Event generators are not magic.

Monte Carlo Event Generators use inversion and rejection sampling algorithms to produce events.

Events are pseudo-data that looks and feels very similar to real data.

Sophisticated pert. calculations used to predict inclusive x-sections, parton showers + multiple interactions to distribute these over many-parton states, using best insights into all-order QFT.

The parton → hadron conversion is based both on perturbative and non-perturbative insights.

This level of detail does, however, come with a large number of parameters.