The Physics of Event Generators

A. Sznajder

UERJ
Instituto de Fisica

January - 2019

Outline

Course Contents

1) Introduction to Statistics for HEP Data Analysis

- Basic concepts of statistics
- Parameter and interval estimation
- Hypothesis testing and goodness of the fit

2) The Physics of Event Generators

- Physics processes , Feynman diagrams and cross sections
- Monte Carlo event generator
- Madgraph

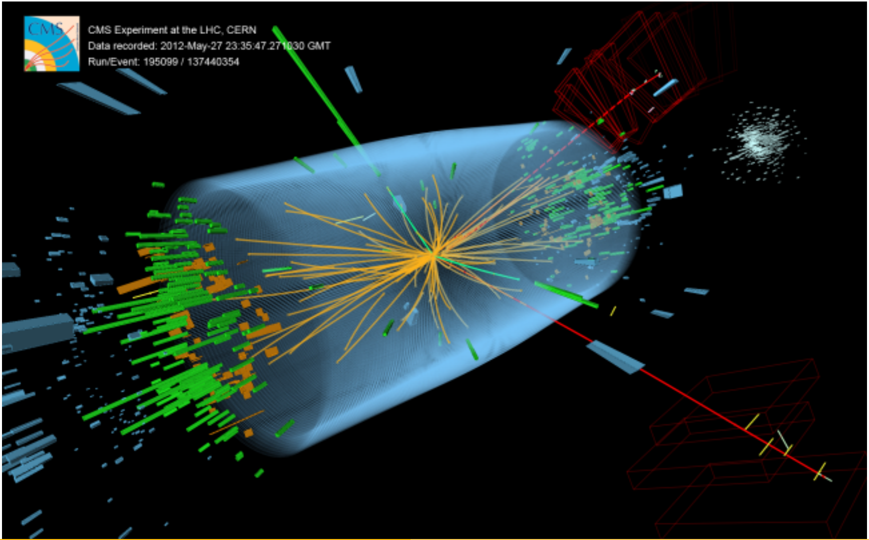
3) Machine Learning

- Introduction to machine learning: classification and regression
- The multilayer perceptron (MLP)
- Universal approximation , vanishing gradient and deep learning
- Convolutional network, autoencoder and adversarial network

Bibliography

- M.Seymour , Introduction to Monte Carlo for the LHC , CERN 2003 , https://indico.cern.ch/event/412017
- F.Maltoni , Predective Monte Carlo Tools for the LHC , CERN 2012 , https://indico.cern.ch/event/181765
- F.Maltoni , Monte Carlo , CERN 2013 , https://indico.cern.ch/event/243711
- B.Webber , QCD and Monte Carlo techniques , CERN 2017, https://indico.cern.ch/event/598530/contributions/2547052

CMS Collison Event

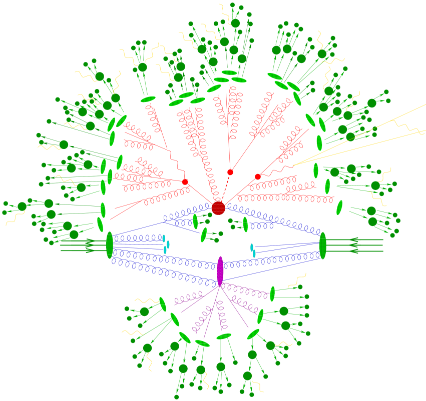


MC Event Generators

The need for MC event generators is related to how we extract predictions (stochastic processes) [1] from fundamental theories QCD and EW and BSM models

Event Generation

1. **Hard scattering** (where new physics lies)
   - high Q^2 scale (short time scale)
   - first principles decription
   - can be improved (NLO, NNLO ...)
   - process dependent
2. **Parton Shower**
   - Well known QCD physics
   - first principles description
   - process independent
3. **Hadronization**
   - low Q^2 physics (long time scale)
   - model based
   - process independent
4. **Underlying Event**
   - low Q^2 physics
   - model based
   - energy and process dependent



[1] Tossing a coin is random but given large # trials a pattern emerges: 50% heads or tails (underlying law or symmetry)

Cross Section

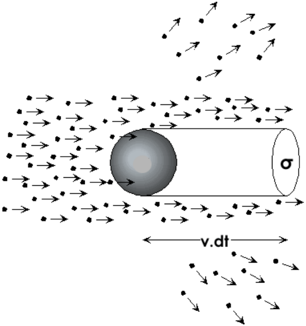
Most experimental observations in particle physics corresponds to cross section and decay rates measurements. From an experimental point of view the cross section can be defined considering a beam of particles colliding with a target of cross sectional area σ

Cross section

The number of collisions dN observed in an time interval dt is given by

$$dN = \rho\sigma(vdt) \quad \Rightarrow \quad \sigma = \frac{W}{F} \qquad (1)$$

, where σ is the cross section , $F = \rho v$ is the incident flux and $W = \frac{dN}{dt}$ the collision rate



For a quantum scattering, the cross section σ represents an effective area of interaction between the particles and it measures the probability of a scattering to occur

HARD SCATTERING

Hard Scattering

From the theory side we can relate the cross section (observable) with the process matrix element (quantum scattering amplitude)

Scattering Golden Rule

The relation between the cross section σ and matrix element M_{fi} can be obtained from Fermi golden rule. In the case of a scattering process $1 + 2 \rightarrow 3, 4, \ldots, n$ we have

$$d\sigma = \frac{||M_{fi}||^2}{F} d\Phi$$

, where $d\Phi$ is the phase space and F is the incoming flux factor

The phase space $d\Phi$ can be written in the following Lorentz invariant form

$$d\Phi = \left[\left(\frac{d^3 p_3}{(2\pi)^3 2E_3} \right) \left(\frac{d^3 p_4}{(2\pi)^3 2E_4} \right) \cdots \left(\frac{d^3 p_n}{(2\pi)^3 2E_n} \right) \right] (2\pi)^4 \delta^4 (p_1 + p_2 - p_3 - p_4 \ldots - p_n)$$

, while the flux factor can also be written in the invariant form

$$F = 4\sqrt{(p_1.p_2)^2 - (m_1 m_2)^2}$$

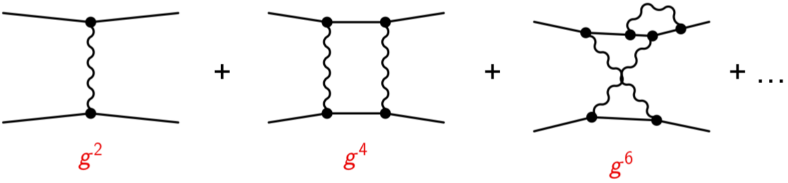
Matrix Elements and Feynman Diagrams

Theory gives the solutions of scattering problems as a perturbative series in the coupling constant. The terms of the series can be obtained directly from the Feynman diagrams through the application of the Feynman rules, without the need to calculate them explicitly (QFT).

Matrix Element

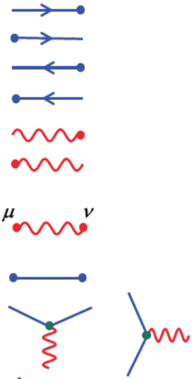
Matrix elements can be written as a power series, where each series term can be associated with a diagram

$$M_{fi} = M_{fi}^{(1)} + M_{fi}^{(2)} + M_{fi}^{(3)} + \dots$$

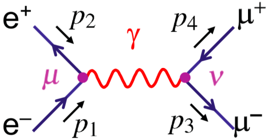


Example: QED Feynman Rules

- **External Lines**
  - spin 1/2
    - incoming particle $u(p)$
    - outgoing particle $\bar{u}(p)$
    - incoming antiparticle $\bar{v}(p)$
    - outgoing antiparticle $v(p)$
  - spin 1
    - incoming photon $\varepsilon^{\mu}(p)$
    - outgoing photon $\varepsilon^{\mu}(p)^*$
- **Internal Lines (propagators)**
  - spin 1 photon $-\dfrac{ig_{\mu\nu}}{q^2}$
  - spin 1/2 fermion $\dfrac{i(\gamma^{\mu}q_{\mu}+m)}{q^2-m^2}$
- **Vertex Factors**
  - spin 1/2 fermion (charge -$|e|$) $ie\gamma^{\mu}$
- **Matrix Element** $-iM$ = product of all factors



Matrix Element (ex: $e^+ \; e^- \;\rightarrow\; \mu^+ \; \mu^-$)



Matrix element calculation ($e^+ \; e^- \;\rightarrow\; \mu^+ \; \mu^-$)

1. Draw the Feynman diagrams
2. Use Feynman rules to get matrix element:

$$\mathcal{M} = e^2 \left[\bar{u}(p_2)\gamma^\mu v(p_1)\right] \left(\frac{g_{\mu\nu}}{q^2}\right) \left[\bar{u}(p_4)\gamma^\mu v(p_1)\right]$$

3. Average and sum over polarizations (trace theorems):

$$\frac{1}{4}\sum_{pol} \|\mathcal{M}\|^2 = \frac{8e^4}{q^4}\left[(p_1.p_3)(p_2.p_4) + (p_1.p_4)(p_2.p_3)\right]$$

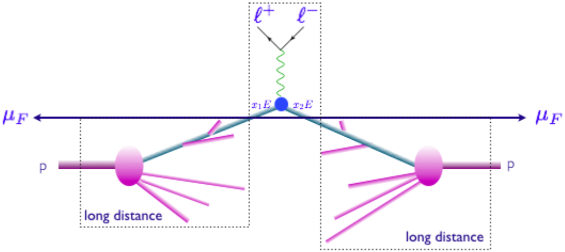
4. Number of terms rise as N^2 (efficient only for small number of final states)
5. Automatically performed by Madgraph, Sherpa ...

Hard Scattering

LHC beams are made of protons and not partons ! QCD factorization theorem relates the short-distance partonic cross section $\hat{\sigma}_{ab\to X}$ and long-distance parton distribution function (PDF) , obtained from experiment

QCD Factorization Theorem

$$\sigma_X = \sum_{a,b} \int_0^1 dx_1 dx_2 \; f_a(x_1, \mu_F) \; f_b(x_2, \mu_F) \; \hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$$



MC Event Generation

Calculations of cross section or decay widths involve integrations over high-dimensional phase space of very complex (peaked) functions

Cross Section Integral

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

, where for N final states the integration over phase space has dimension $3N$

Definition of MC Event Generator

Monte Carlo event generator is a program which produces particle physics events with the same probability as they occur in nature (virtual collider). It performs a cross section integrals calculations and unweights events to give the four momenta of the particles.

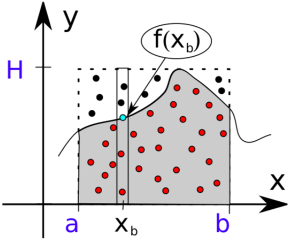
Among theorists “Monte Carlo” also includes codes which don’t provide a fully exclusive information on the final state, but only cross sections or distributions at the parton level (typically at NLO or higher). These codes are also referred as “MC integrators” or ”Cross Section integrators”.

MC Integration: from Acceptance-Rejection to Averages

Acceptance-Rejection Method

The MC integration by the Acceptance-Rejection method has a simple geometrical interpretation as

$$I = \left(\frac{N_{accept}}{N_{total}} \right) A_{box}$$



From Acceptance-Rejection to Averages

The Aceptance-Rejection method is not optimal ! Consider a narrow strip (bin) of height H , around a point x_b

1. For this bin $N_{accept}/N_{total}H$ is just an estimate for $f(x_b)$, which we know how to calculate directly !
2. We can estimate integral $f(x)$ in $[a, b]$ by sampling N points $\{x_i\}$ uniformly in the interval and summing over each bin contibution

$$I \simeq \sum_{i=1}^{N} \underbrace{\left(\frac{b-a}{N} \right)}_{\Delta x} f(x_i) = (b-a) \left[\frac{1}{N} \sum_{i=1}^{N} f(x_i) \right] = (b-a) \langle f \rangle$$

MC Integration as Averages

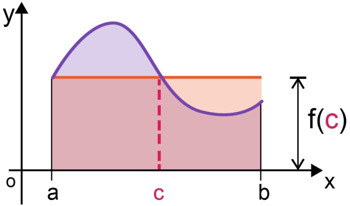
Another way of understanding the MC integration as an average is through the mean value theorem

Mean Value Theorem for Integrals

Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then there exists c in (a, b) such that

$$\int_a^b f(x)dx = (b-a)f(c)$$

The value of $f(c)$ is the mean value in $[a, b]$



We can estimate the mean value of $f(x)$ in $[a, b]$ directly by sampling the function uniformly in the interval and taking the average $f(c) \simeq \frac{1}{N}\sum_{i=1}^{N} f(x_i)$

Integrals as an Average

The integral of $f(x)$ and its variance can be estimated by sampling N points $\{x_i\}$ in the interval $[a, b]$

$$\begin{cases} I_N = \int_a^b f(x)dx \simeq (b-a)\left[\frac{1}{N}\sum_{i=1}^{N} f(x_i)\right] \\ V_N \simeq \frac{(b-a)^2}{N}\sum_{i=1}^{N}[f(x_i)]^2 - I^2 \end{cases} \Rightarrow I = I_N \pm \sqrt{\frac{V_N}{N}} \quad \text{(CLT Theorem)}$$

MC Integration in d-dimensions

The generalization of the MC integration to d-dimension is given by

MC Integration d-dimensions

Consider the d-dimentional integration of a function $f(\vec{x})$ over a volume V

$$I = \int dV\, f(\vec{x}) \simeq \frac{1}{N} \sum_{i=1}^{N} \underbrace{Vf(\vec{x}_i)}_{w_i} \quad \text{, where } w_i \text{ is the event weight}$$

Comparing the convergence of different integration methods, we can see the power of MC integration for high dimensional cases

Numerical Integration Methods in d-dimensions

Integration error as a function of the number of points N and the dimension d

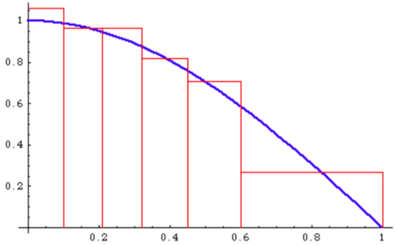
- Monte Carlo: $\frac{1}{\sqrt{N}}$ (dimension independent)
- Trapezium : $\frac{1}{N^{\frac{2}{d}}}$
- Simpson : $\frac{1}{N^{\frac{4}{d}}}$

MC Integration : VEGAS

To increase the integration efficiency, instead of sampling $[a, b]$ uniformly, we sample more points where $f(x)$ is large. For that we use an importance sampling distribution $g(x)$ that aproximates the integrand

VEGAS Algorithm [2]

VEGAS is an adaptative importance sampling algorithm. It determines the sampling distribution iteractively while sampling the integrand $f(x)$, by building a step function approximation(histogram)



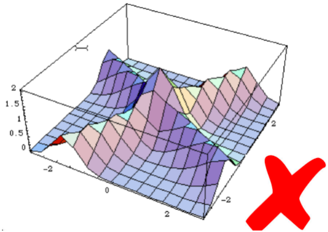
[2] Lepage, G.P. (May 1978). "A New Algorithm for Adaptive Multidimensional Integration". Journal of Computational Physics. 27: 192–203

MC Integration : VEGAS

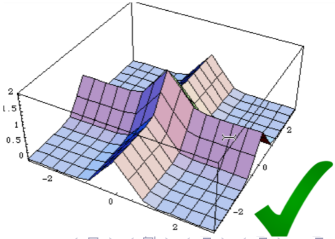
VEGAS

- The number of histogram bins in d-dimensions grows as K^d (dimensional curse)
- Vegas avoids this by approximating the distribution by a separable(factorizable) function like: $g(x_1, x_2, \ldots x_n) = g_1(x_1).g_2(x_2) \ldots g_n(x_n)$
- For factorizable functions the number of bins grows only as $K.d$
- The efficiency of VEGAS depends on the function factorization !

Separability(factorization) is equivalent to aligning the integrand peaks with the coordinate axes (change of variables)



$\Longrightarrow$



MC Integration : Multichanel

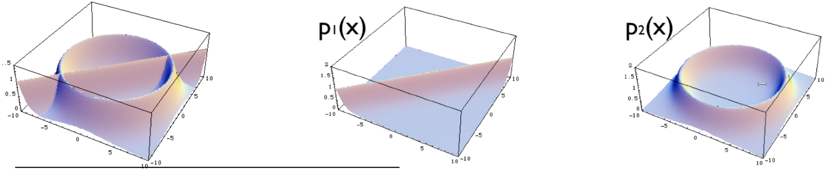
If there is no transformation that aligns the integrand peaks to the coordinate axes, VEGAS is bound to fail !

Multichannel Integration [3]

Multichannel integration use different transformations (channels) for separating non-factorizable singularities. Each channel takes care of one "peak" at the time, such that $p(x) = \sum_{i=1}^{n} \alpha_i p_i(x)$,

where $\sum_{i=1}^{n} \alpha_i = 1$

$$I = \int f(x)dx = \sum_{i=1}^{n} \alpha_i \int p_i(x) \frac{f(x)}{p(x)} dx$$

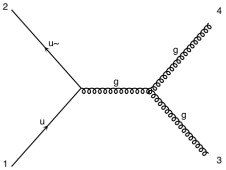


[3] R.Kleiss and R.Pittau., Weight optimization in multichannel Monte Carlo, Computer Physics Communications 83.2-3 (1994), pg.141

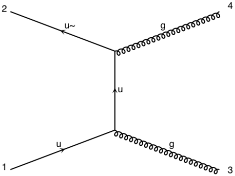
Example : QCD 2 $\to$ 2 process

Multichannel Integration

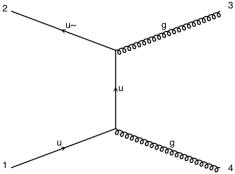
QCD 2 $\to$ 2 process shows three very different pole structures contributing to the same matrix element. One can not find a transformation that factorizes the $\|\mathcal{M}_{tot}\|^2$. Must use multichannel integration to integrate them separately and then combine them.



$$\propto \frac{1}{\hat{s}} = \frac{1}{(p_1 + p_2)^2}$$



$$\propto \frac{1}{\hat{t}} = \frac{1}{(p_1 - p_3)^2}$$



$$\propto \frac{1}{\hat{u}} = \frac{1}{(p_1 - p_4)^2}$$

MC Integration: Madgraph Method

Madgraph Method

Madgraph uses a multichannel integration based on single diagrams [4] , where the integrand $|\mathcal{M}_{tot}|^2$ is decomposed as $|\mathcal{M}_{tot}|^2 = \sum_i f_i$. The basis f_i is defined in terms of individual diagrams amplitude $\mathcal{M}_i$ by

$$f_i = \frac{|\mathcal{M}_i|^2}{\sum_j |\mathcal{M}_j|^2} |\mathcal{M}_{tot}|^2 = |\mathcal{M}_i|^2 \underbrace{\left(\frac{|\sum_l \mathcal{M}_l|^2}{\sum_j |\mathcal{M}_j|^2} \right)}_{=1 \pm \epsilon_{interf}}$$

- The peak structure of each f_i is the same as of that of $|\mathcal{M}_i|^2$
- The mapping g_i can be derived from diagram propagator structure
- Easy to reweight the channels, so large contributions are evaluated with greater number of MC points
- Decomposes the amplitude into n independent integrations (parallelization)

[4] F.Maltoni and T.Stelzer , MadEvent:Automatic event generation with MadGraph , JHEP0302:027,2003

MC Event Generation

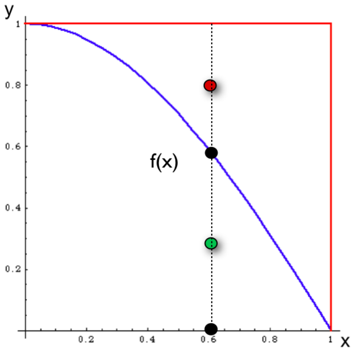
Event Generation

Event generation is done by sampling phase space points (event kinematics). The cross section associates different weights to distinct phase events

For detector simulation and analysis we want unweighted events and they are obtained by applying Acceptance-Rejection method to perform unweighting

Unweighting (Acceptance-Rejectection)

- Random sort x
- Calculate $f(x)$
- Random sort $0 < y < f_{max}$
- If $y \leq f(x)$ accept the event , otherwise reject it

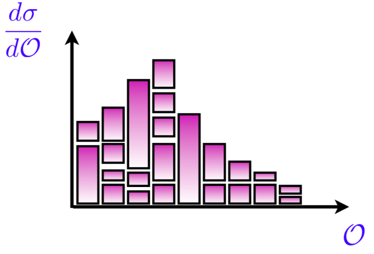


MC Event Generation

Comparison between distribution made with weighted versus unweighted events

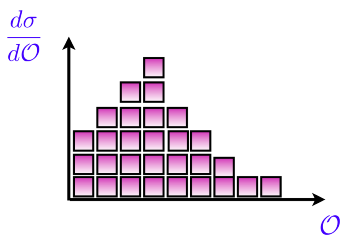
Weighted Events

Aproximately the same # of events in areas of phase space with very different probabilities, so events must have different weights



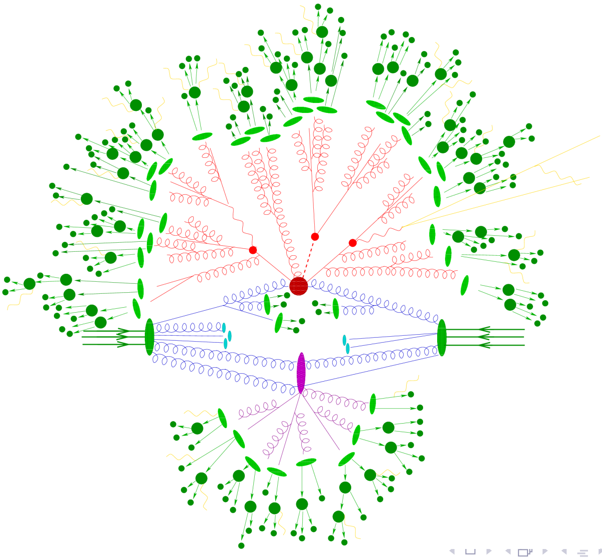
Unweighted Events

The # of events is proportional to the probability of phase space region, so all events have the same weight. Events distributed as we observe in nature !



PARTON SHOWER

Parton Shower



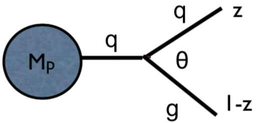
Parton Shower

Acelerated color charged radiates gluons, and as gluons themselves carry colour charges, they can emit further radiation, leading to parton showers [5]

Soft and Collinear Emissions

The QCD matrix elements enhances radiation for soft and collinear emissions (most singular).

$$\frac{1}{(p_q + p_g)^2} \simeq \frac{1}{2E_q E_g(1 - cos\theta)}$$



If the n-parton differential cross section before splitting is $d\sigma_n$ the expressions for the collinear and soft aproximations factorizes. For the splittings of a parton (ex: $q \rightarrow q + g$) we have

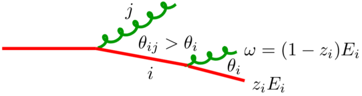
Parton Branching (Collinear)

$$d\sigma_{n+1} \simeq d\sigma_n \frac{\alpha_s}{2\pi} \sum_i dz P_{ii}(z_i, \phi_i) \frac{d\xi_i}{\xi_i} \frac{d\phi_i}{2\pi}$$

Parton Branching (Soft)

$$d\sigma_{n+1} \simeq d\sigma_n \frac{\alpha_s}{2\pi} \frac{d\omega}{\omega} \sum_i \left(-\vec{T}_i . \vec{T}_j \right) \Theta(\xi_{ij} - \xi_i) \frac{d\xi_i}{\xi_i}$$

, where θ_i is the emission angle , $\xi_i = 1 - cos\theta_i$ and z_i is the parton energy fraction after emission



Parton Shower

The $P_{ij}(z)$ splitting function are known as Altarelli-Parisi splitting kernels, where $T_R = \frac{1}{2}$, $C_A = 3$ and $C_F = \frac{4}{3}$

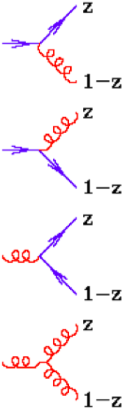
Altarelli-Parisi Splitting Kernels

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

$$P_{qg}(z) = T_R \left[z^2 + (1-z)^2 \right]$$

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

$$P_{gg}(z) = 2C_A \left[z(1-z) + \frac{z}{1-z} + \frac{1-z}{z} \right]$$



The splitting functions diverges for $z = 1$ and $z = 0$, except for P_{gq} !

Parton Shower

Sudakov Form Factor

The probability that parton does not emit [6] when evolving from scale Q_{min} to Q is given by the Sudakov form factor [7]

$$\Delta(Q, Q_{min}) \equiv 1 - [Prob(1\ emission) + Prob(2\ emissions) + \ldots]$$

$$\Rightarrow \quad \Delta(Q, Q_{min}) \simeq exp\left[-C_i \frac{\alpha_s}{\pi} ln^2\left(\frac{Q_0}{Q_{min}}\right)\right]$$

, where $C_q = C_F = 4/3$ and $C_g = C_A = 3$

The probability of a parton to evolve from scale Q_i to Q_f without resolvable emissions is given by the following ratio of Sudakov factors

Probability of no Emission between Q_i and Q_f

$$Prob(Q_i \to Q_f\ without\ emissions) = \frac{\Delta(Q_i, Q_{min})}{\Delta(Q_f, Q_{min})}$$

[6] Resolvable gluons

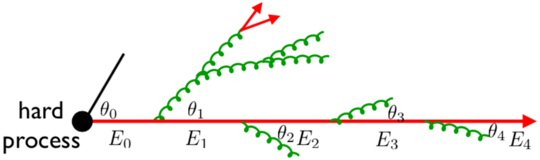
[7] Analogous to radioactive decay , where the probability of no decay before T is $exp[-\int^T dt\lambda]$

Parton Shower

By sequential application of Sudakov factors ratio, and Monte Carlo method to generate z values for each splitting, a parton shower is developed for each hard subprocess parton. The shower is terminated when the virtualities have fallen to the hadronization scale Q_{min}

Parton Shower Algorithm

1. Generate a uniform random number $R \in [0, 1]$
2. If $R < \Delta(Q_i, Q_{min})$ there's no resolvable emission (stop branching)
3. Given initial Q_i , equate $R = \dfrac{\Delta(Q_i, Q_{min})}{\Delta(Q_{i+1}, Q_{min})}$ and solve for Q_{i+1} (next emission scale)
4. Use MC method to sample z fraction from the apropriate $P_{ij}(z)$ distribution
5. Continue iterating (goto 1)

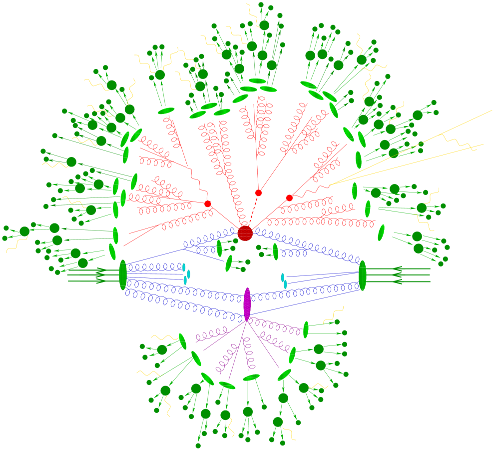


$$z = \frac{E_{i+1}}{E_i}$$

HADRONIZATION

Hadronization

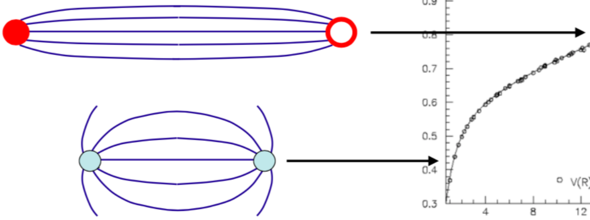
- Degrees of freedom that reaches detector are hadrons
- After parton shower is terminated quarks and gluons must turn into hadrons
- Need a phenomenological model to describe QCD confinement

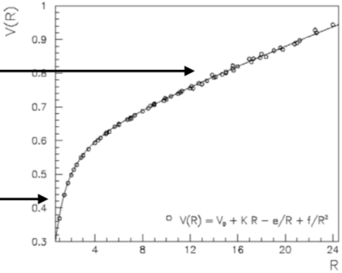


Hadronization - String Model

String Hadronization Model

- At short distances (large Q), QCD is like QED: colour field lines spread out ($1/r$ potential)
- At long distances, gluon self-attraction gives rise to colour string (linear potential ⇒ quark confinement)
- Intense colour field induces quark-antiquark pair creation, which combines into color neutral bound states (hadronization)

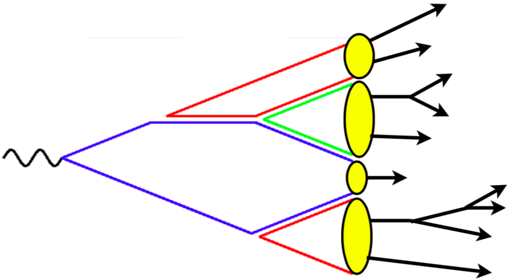




Hadronization - Cluster Model

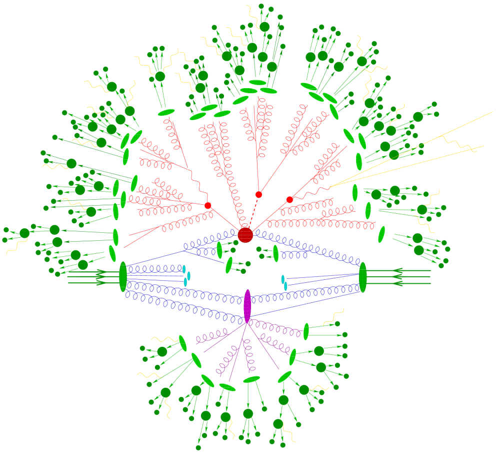
Cluster Hadronization Model

- In parton shower, relative transverse momenta evolve from a high scale Q towards lower values
- At a scale near $\Lambda_{qcd} \simeq 200 MeV$, perturbation theory breaks down and hadrons are forme
- Before that, at scales of aproximatley few $x\Lambda_{qcd}$, there is universal preconfinement of colour
- Decay of preconfined clusters provides a direct basis for hadronization



UNDERLYING EVENT

Underlying Event



Underlying Event

Underlying Event

- Multiple parton interactions in same collision
- Assume QCD 2-to-2 secondary collisions (need low pT cutoff)
- Need to model colour flow (color reconnections)