Introduction to Monte Carlos

Stefan Gieseke

Institut für Theoretische Physik
KIT

CTEQ Summer School 2013
Pittsburgh, 7–17 July 2013

Karlsruhe Institute of Technology

MCnet

PHYSICS AT THE TERA SCALE
Helmholtz Alliance

Outline

- Part I — Basics
  - Introduction
  - Monte Carlo techniques
- Part II — Perturbative physics
  - Hard scattering
  - Parton showers
- Part III — Non–perturbative physics
  - Hadronization
  - Hadronic decays
  - Comparison to data

Thanks

Thanks to my colleagues

Frank Krauss, Leif Lönnblad, Steve Mrenna, Peter Richardson, Mike Seymour, Torbjörn Sjöstrand.

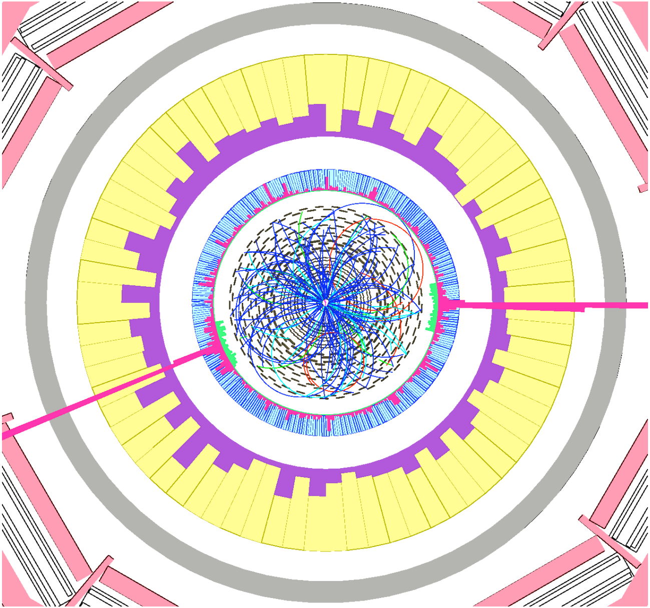
Introduction

Why Monte Carlos?

We want to understand

$$\mathscr{L}_{\text{int}} \longleftrightarrow \text{Final states}\ .$$

Can you spot the Higgs?



Why Monte Carlos?

LHC experiments require
sound understanding of signals and *backgrounds*.

$\uparrow$

Full detector simulation.

$\uparrow$

Fully exclusive hadronic final state.

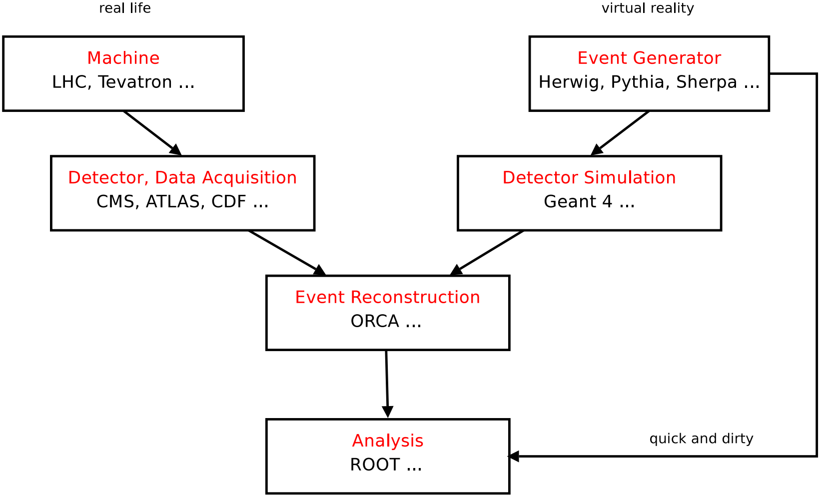
$\uparrow$

Monte Carlo event generator with
parton shower, hadronization model, decays of unstable particles.

$\uparrow$

Parton level computations.

Experiment and Simulation



Monte Carlo Event Generators

- Complex final states in full detail (jets).
- Arbitrary observables and cuts from final states.
- Studies of new physics models.

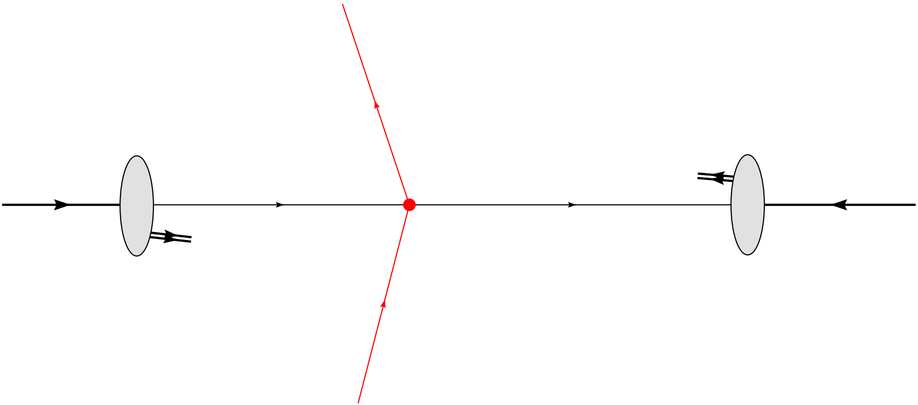
- Rates and topologies of final states.
- Background studies.
- Detector Design.
- Detector Performance Studies (Acceptance).

- *Obvious* for calculation of observables on the quantum level

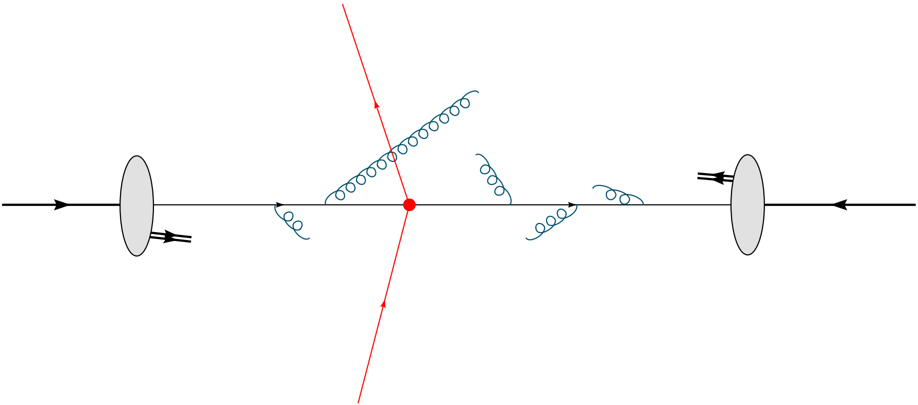
$$|A|^2 \longrightarrow \text{Probability.}$$

pp Event Generator

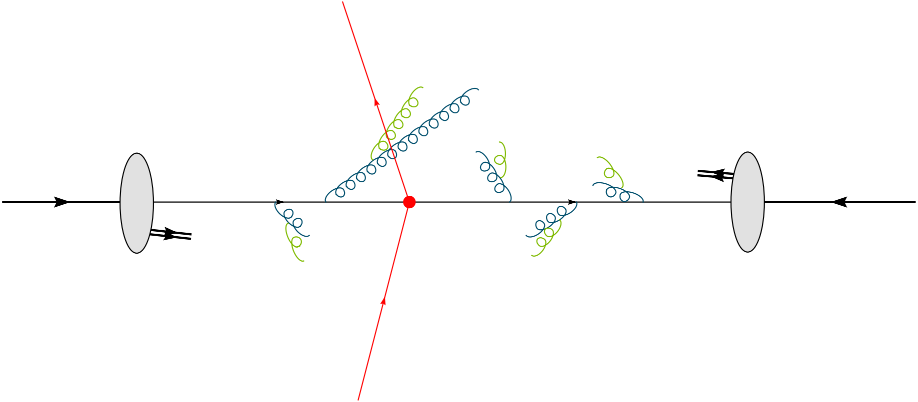
pp Event Generator



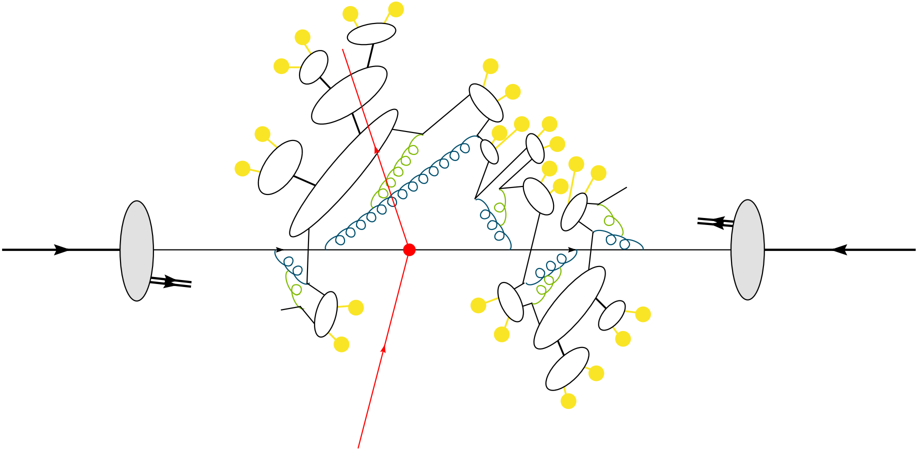
pp Event Generator



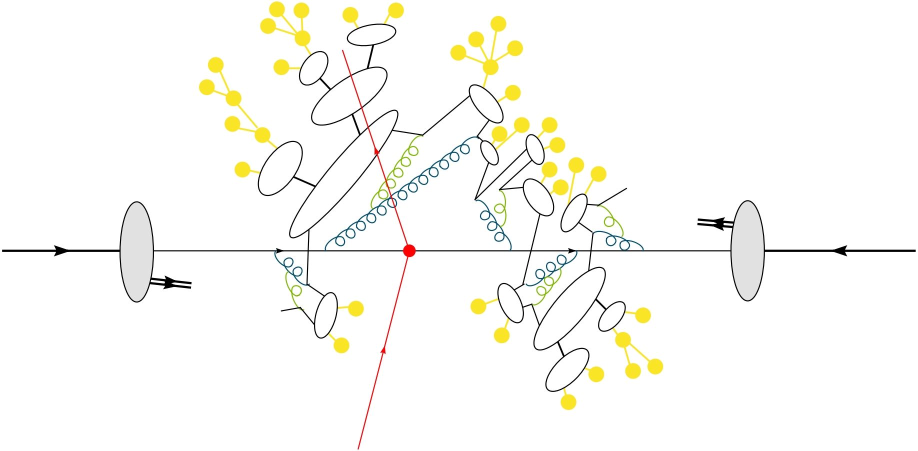
pp Event Generator



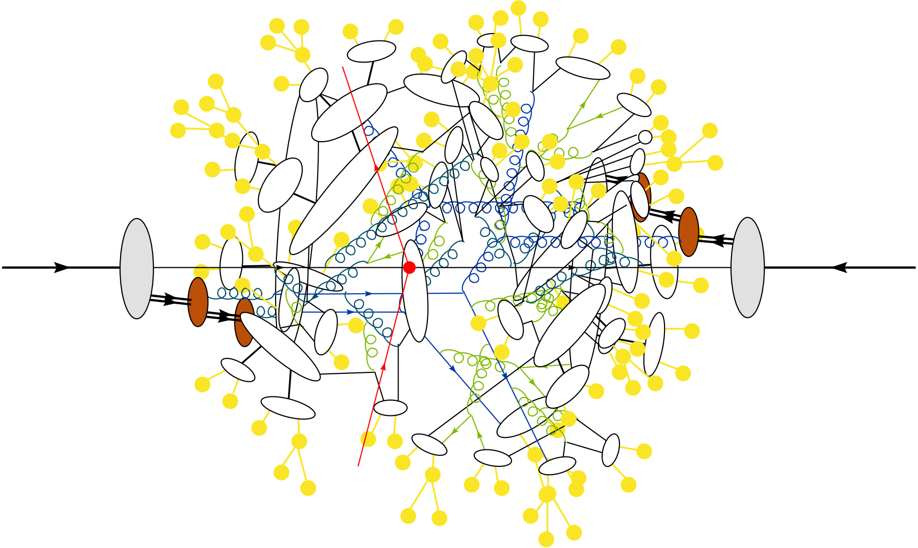
pp Event Generator



pp Event Generator



pp Event Generator



Divide and conquer

Partonic cross section from Feynman diagrams

$$\mathrm{d}\sigma = \mathrm{d}\sigma_{\mathrm{hard}}\mathrm{d}P(\mathrm{partons} \to \mathrm{hadrons})$$

Note, that

$$\int \mathrm{d}P(\mathrm{partons} \to \mathrm{hadrons}) = 1\ ,$$

- σ remains unchanged
- introduce realistic fluctuations into distributions.

Divide and conquer

Partonic cross section from Feynman diagrams

$$d\sigma = d\sigma_{\text{hard}} dP(\text{partons} \to \text{hadrons})$$

Note, that

$$\int dP(\text{partons} \to \text{hadrons}) = 1 \; ,$$

- σ remains unchanged
- introduce realistic fluctuations into distributions.

Simulation steps governed by different scales
$\longrightarrow$ separation into ($Q_0 \approx 1\,\text{GeV} > \Lambda_{\text{QCD}}$)

$$\begin{aligned} dP(\text{partons} \to \text{hadrons}) = \; & dP(\text{resonance decays}) && [\Gamma > Q_0] \\ & \times dP(\text{parton shower}) && [\text{TeV} \to Q_0] \\ & \times dP(\text{hadronisation}) && [\sim Q_0] \\ & \times dP(\text{hadronic decays}) && [O(\text{MeV})] \end{aligned}$$

Divide and conquer

$$\begin{aligned} \mathrm{d}P(\text{partons} \to \text{hadrons}) = \; & \mathrm{d}P(\text{resonance decays}) & [\Gamma > Q_0] \\ & \times \mathrm{d}P(\text{parton shower}) & [\text{TeV} \to Q_0] \\ & \times \mathrm{d}P(\text{hadronisation}) & [\sim Q_0] \\ & \times \mathrm{d}P(\text{hadronic decays}) & [O(\text{MeV})] \end{aligned}$$

Quite complicated integration.

Divide and conquer

$$
\begin{aligned}
\mathrm{d}P(\text{partons} \to \text{hadrons}) = \; & \mathrm{d}P(\text{resonance decays}) && [\Gamma > Q_0] \\
& \times \mathrm{d}P(\text{parton shower}) && [\text{TeV} \to Q_0] \\
& \times \mathrm{d}P(\text{hadronisation}) && [\sim Q_0] \\
& \times \mathrm{d}P(\text{hadronic decays}) && [O(\text{MeV})]
\end{aligned}
$$

Quite complicated integration.

Monte Carlo is the only choice.

Monte Carlo Methods

Introduction to the most important MC sampling
(= integration) techniques.

1. Hit and miss.
2. Simple MC integration.
3. (Some) methods of variance reduction.
4. Multichannel.

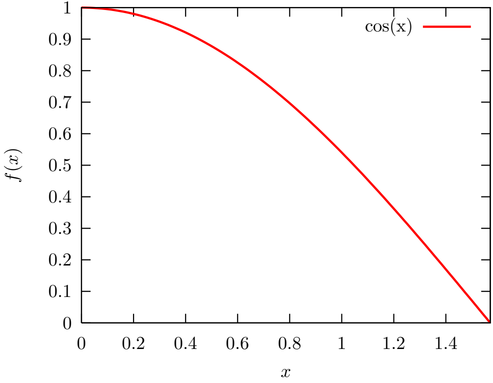
Probability

Example: $f(x) = \cos(x)$.

Probability density:

$$dP = f(x)\,dx$$

is probability to find value x.



Probability $\sim$ *Area*

Probability

Example: $f(x) = \cos(x)$.

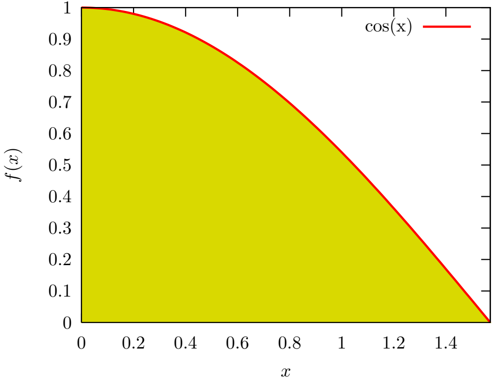
Probability density:

$$dP = f(x)\,dx$$

is probability to find value x.

$$F(x) = \int_{x_0}^{x} f(x)\,dx$$

is called *probability distribution*.



Probability $\sim$ *Area*

Hit and Miss

Hit and miss method:

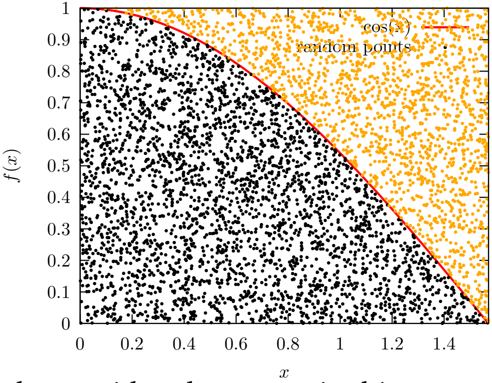
- throw N random points (x, y) into region.
- Count hits N_{hit}, i.e. whenever $y < f(x)$.

Then

$$I \approx V \frac{N_{\text{hit}}}{N}.$$

approaches 1 again in our example.

Example: $f(x) = \cos(x)$.



Every accepted value of x can be considered an event in this picture. As $f(x)$ is the 'histogram' of x, it seems obvious that the x values are distributed as $f(x)$ from this picture.

Hit and Miss

This method is used in many event generators. However, it is not sufficient as such.

- Can handle any density $f(x)$,
  however wild and unknown it is.
- $f(x)$ should be bounded from above.
- Sampling will be very *inefficient* whenever $\mathrm{Var}(f)$ is large.

Improvements go under the name variance reduction as they improve the error of the crude MC at the same time.

Simple MC integration

Mean value theorem of integration:

$$I = \int_{x_0}^{x_1} f(x)\,dx$$
$$= (x_1 - x_0)\langle f(x)\rangle$$
$$\approx (x_1 - x_0)\frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

(Riemann integral).

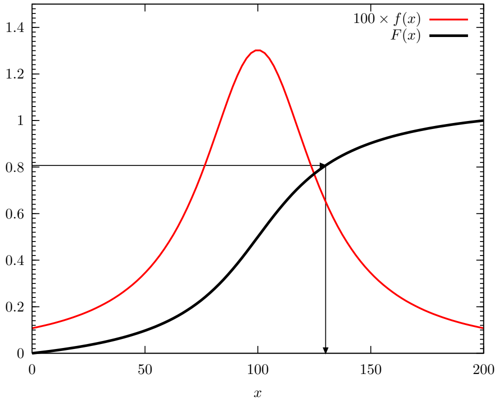
Sum doesn't depend on ordering
$\longrightarrow$ randomize x_i.

Yields a flat distribution of events x_i,
but weighted with *weight* $f(x_i)$ ($\rightarrow$ unweighting).

Inverting the Integral

- Probability density $f(x)$. Not necessarily normalized.
- Integral $F(x)$ known,
- $P(x < x_s) = F(x_s)$.
- Probability = 'area', distributed evenly,

$$\int_{x_0}^{x} dP = r \cdot \text{area}$$



Sample x according to $f(x)$ with

$$x = F^{-1}\Big[F(x_0) + r\big(F(x_1) - F(x_0)\big)\Big] .$$

Inverting the Integral

Sample x according to $f(x)$ with

$$x = F^{-1}\Big[F(x_0) + r\big(F(x_1) - F(x_0)\big)\Big] \, .$$

Optimal method, but we need to know

- The integral $F(x) = \int f(x)\,\mathrm{d}x$,
- It's inverse $F^{-1}(y)$.

That's rarely the case for real problems.

But very powerful in combination with other techniques.

Importance sampling

Error on Crude MC $\sigma_{MC} = \sigma/\sqrt{N}$.
$\Longrightarrow$ Reduce error by reducing variance of integrand.

Importance sampling

Error on Crude MC $\sigma_{MC} = \sigma/\sqrt{N}$.
$\Longrightarrow$ Reduce error by reducing variance of integrand.

Idea: *Divide out the singular structure.*

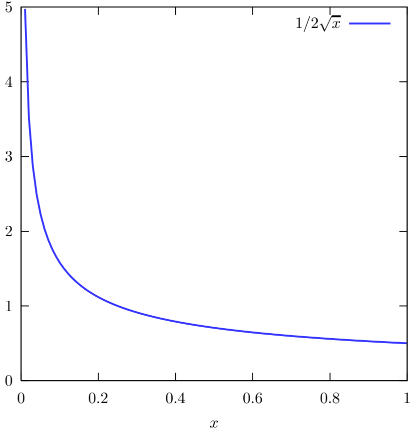
$$I = \int f \, \mathrm{d}V = \int \frac{f}{p} p \mathrm{d}V \approx \left\langle \frac{f}{p} \right\rangle \pm \sqrt{\frac{\langle f^2/p^2 \rangle - \langle f/p \rangle^2}{N}} \, .$$

where we have chosen $\int p \, \mathrm{d}V = 1$ for convenience.

Note: need to sample flat in $p \, \mathrm{d}V$, so we better know $\int p \mathrm{d}V$ and it's inverse.

More interesting for divergent integrands, eg

$$\frac{1}{2\sqrt{x}} \ ,$$



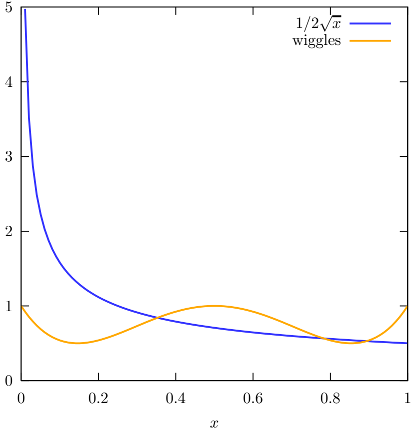
Importance sampling — better example

More interesting for divergent integrands, eg

$$\frac{1}{2\sqrt{x}} \,,$$

with some wiggles,

$$p(x) = 1 - 8x + 40x^2 - 64x^3 + 32x^4 \,.$$



Importance sampling — better example

More interesting for divergent integrands, eg

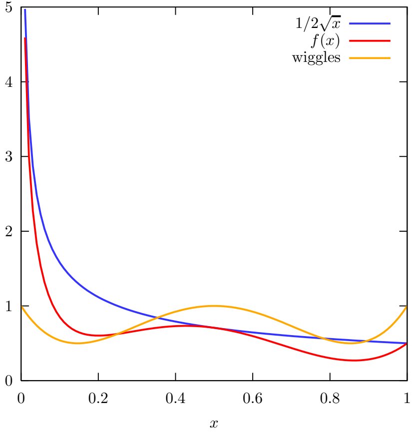
$$\frac{1}{2\sqrt{x}}\ ,$$

with some wiggles,

$$p(x) = 1 - 8x + 40x^2 - 64x^3 + 32x^4\ .$$

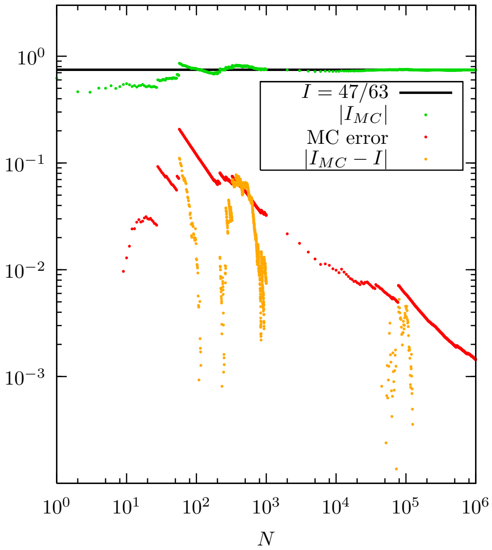
i.e. we want to integrate

$$f(x) = \frac{p(x)}{2\sqrt{x}}\ .$$



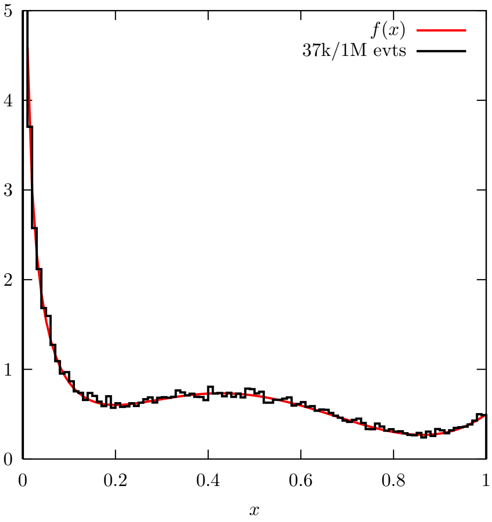
Importance sampling — better example

- Crude MC gives result in reasonable 'time'.
- Error a bit unstable.
- Event generation with maximum weight $w_{\text{max}} = 20$. (that's arbitrary.)
- hit/miss/events with $(w > w_{\text{max}}) =$ 36566/963434/617 with 1M generated events.



Want events:
use hit+mass variant here:

- Choose new random number r
- $w = f(x)$ in this case.
- if $r < w/w_{\max}$ then "hit".
- MC efficiency = hit$/N$.
- Efficiency for MC events only 3.7%.
- Note the wiggly histogram.



Importance sampling — example

Now importance sampling, i.e. divide out $1/2\sqrt{x}$.

$$\int_0^1 \frac{p(x)}{2\sqrt{x}}\,dx = \int_0^1 \left(\frac{p(x)}{2\sqrt{x}}\middle/\frac{1}{2\sqrt{x}}\right)\frac{dx}{2\sqrt{x}}$$
$$= \int_0^1 p(x)\,d\sqrt{x}$$
$$= \int_0^1 p(x(\rho))\,d\rho$$
$$= \int_0^1 1-8\rho^2+40\rho^4-64\rho^6+32\rho^8\,d\rho$$

so,

$$\rho = \sqrt{x}, \qquad d\rho = \frac{dx}{2\sqrt{x}}$$

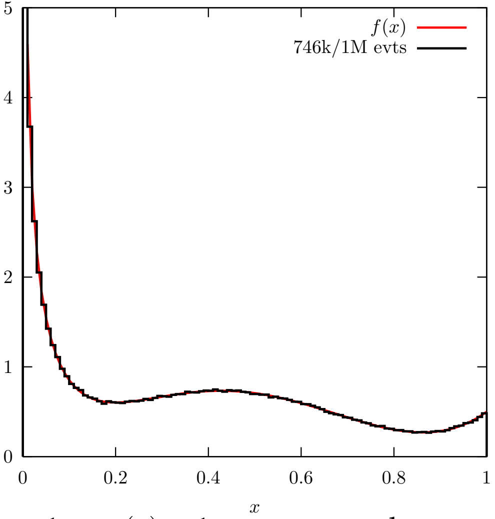
x sampled with *inverting the integral* from flat random numbers ρ, $x=\rho^2$.

Importance sampling — example

$$\int_0^1 \frac{p(x)}{2\sqrt{x}}\,dx = \int_0^1 p(x(\rho))\,d\rho$$

with

$$\rho = \sqrt{x}, \qquad d\rho = \frac{dx}{2\sqrt{x}}$$



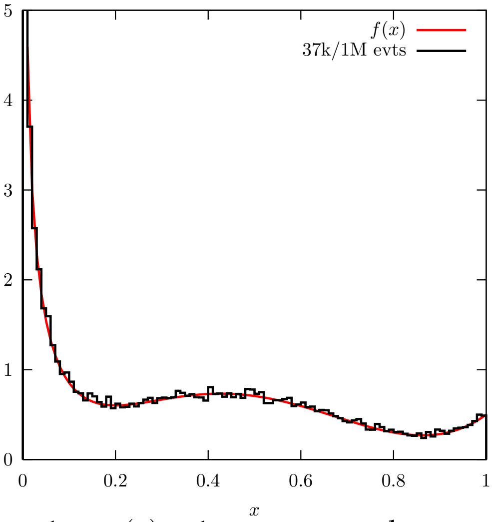
Events generated with $w_{\max} = 1$, as $p(x) \le 1$, no guesswork needed here! Now, we get 74.6% MC efficiency.

Importance sampling — example

$$\int_0^1 \frac{p(x)}{2\sqrt{x}}\, dx = \int_0^1 p(x(\rho))\, d\rho$$

with

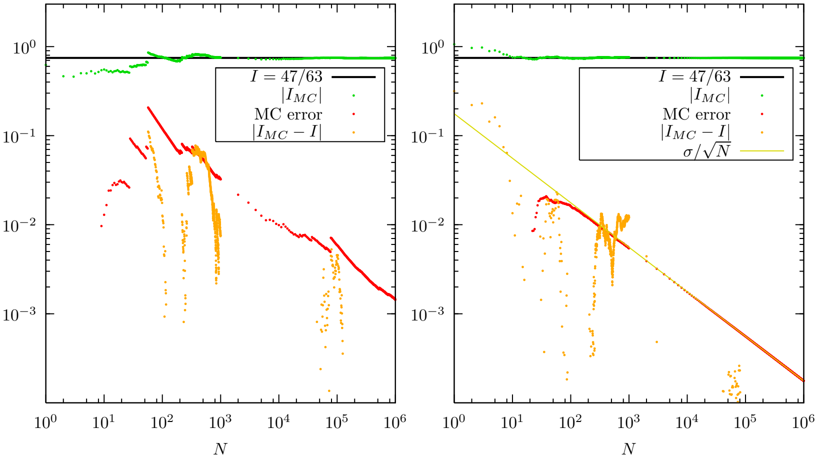
$$\rho = \sqrt{x}, \qquad d\rho = \frac{dx}{2\sqrt{x}}$$



Events generated with $w_{\max} = 1$, as $p(x) \leq 1$, no guesswork needed here! Now, we get 74.6% MC efficiency.
...as opposed to 3.7%.

Importance sampling — example

Crude MC vs Importance sampling.

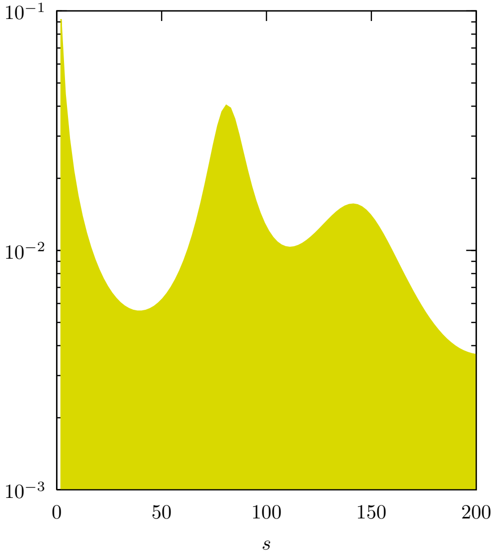


100× more events needed to reach same accuracy.

Multichannel MC

Typical problem:

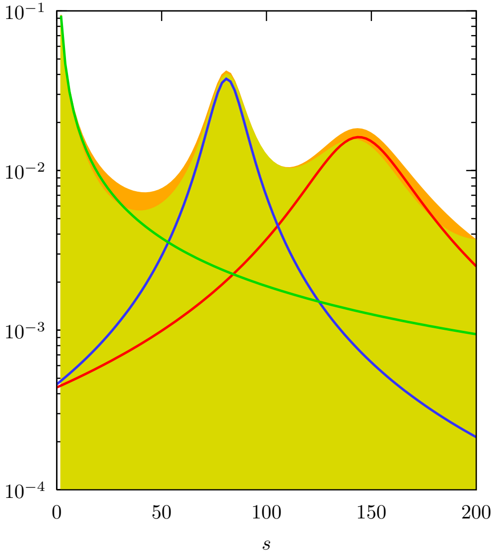
- $f(s)$ has multiple peaks ($\times$ wiggles from ME).



Multichannel MC

Typical problem:

- $f(s)$ has multiple peaks ($\times$ wiggles from ME).
- Usually have some idea of the peak structure.

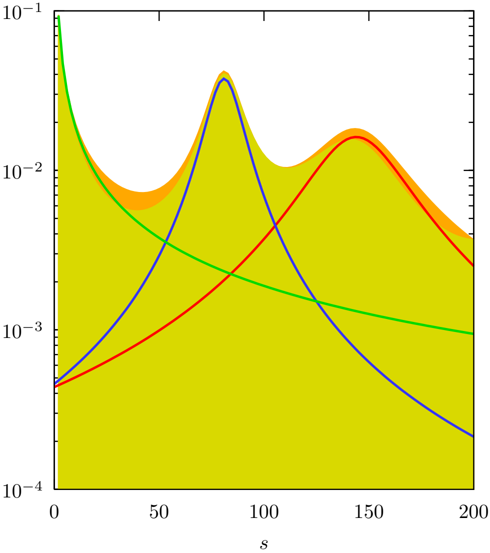


Multichannel MC

Typical problem:

- $f(s)$ has multiple peaks ($\times$ wiggles from ME).
- Usually have some idea of the peak structure.
- Encode this in sum of sample functions $g_i(s)$ with weights $\alpha_i, \sum_i \alpha_i = 1$.

$$g(s) = \sum_i \alpha_i g_i(s) \; .$$



Multichannel MC

Now rewrite

$$\begin{aligned}
\int_{s_0}^{s_1} f(s)\, ds &= \int_{s_0}^{s_1} \frac{f(s)}{g(s)} g(s)\, ds \\
&= \int_{s_0}^{s_1} \frac{f(s)}{g(s)} \sum_i \alpha_i g_i(s)\, ds \\
&= \sum_i \alpha_i \int_{s_0}^{s_1} \frac{f(s)}{g(s)} g_i(s)\, ds
\end{aligned}$$

Now $g_i(s)\, ds = d\rho_i$ (inverting the integral).

Multichannel MC

Now rewrite

$$\begin{aligned}\int_{s_0}^{s_1} f(s)\,ds &= \int_{s_0}^{s_1} \frac{f(s)}{g(s)} g(s)\,ds \\ &= \int_{s_0}^{s_1} \frac{f(s)}{g(s)} \sum_i \alpha_i g_i(s)\,ds \\ &= \sum_i \alpha_i \int_{s_0}^{s_1} \frac{f(s)}{g(s)} g_i(s)\,ds\end{aligned}$$

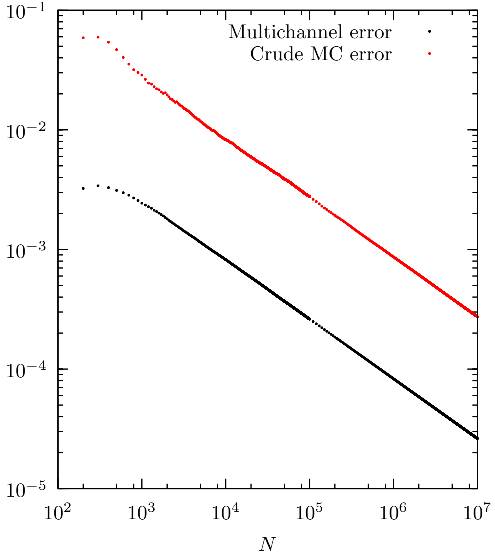
Now $g_i(s)\,ds = d\rho_i$ (inverting the integral).

Select the distribution $g_i(s)$ you'd like to sample next event from acc to weights α_i.

α_i can be optimized after a number of trials.

Multichannel MC

Works quite well:

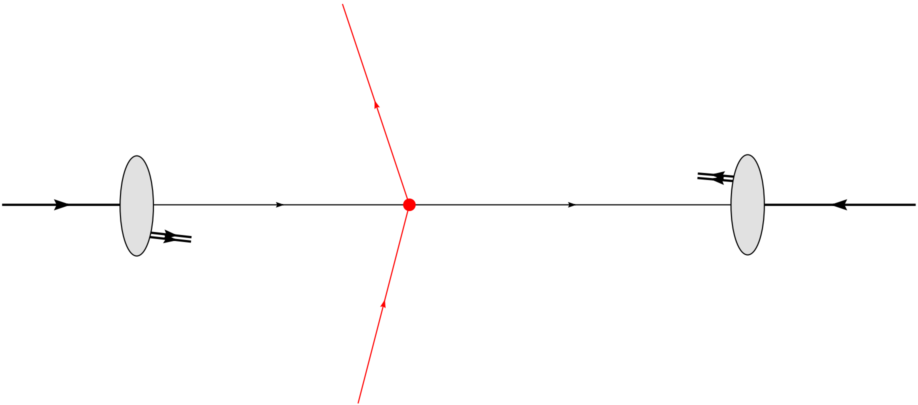


Some Remarks/Real Life MC

- Didn't discuss random number generators. Please make sure to use 'good' random numbers.
- Didn't discuss *stratified sampling* (VEGAS).
  Sample where variance is biggest.
  (not necessarily where PS is most populated).
- Only discussed one–dimensional case here. N–particle PS has $3N - 4$ dimensions...
- Didn't discuss tools geared towards this, like `RAMBO` (generates flat N particles PS).
- generalisation straightforward, particularly
  MCError $\sim \frac{1}{\sqrt{N}}$,
  compare eg Trapezium rule Error $\sim \frac{1}{N^{2/D}}$.
- Many important techniques covered here in detail! Should be good starting point.

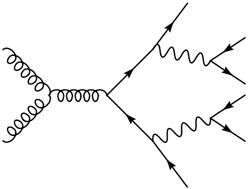
Hard Scattering

Hard scattering



Matrix elements

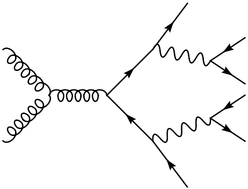
- Perturbation theory/Feynman diagrams give us (fairly accurate) final states for a few number of legs ($O(1)$).



- OK for very inclusive observables.

Matrix elements

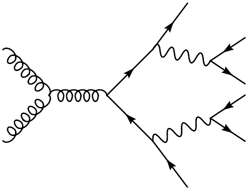
- Perturbation theory/Feynman diagrams give us (fairly accurate) final states for a few number of legs ($O(1)$).



- OK for very inclusive observables.
- Starting point for further simulation.
- Want exclusive final state at the LHC ($O(100)$).

Matrix elements

- Perturbation theory/Feynman diagrams give us (fairly accurate) final states for a few number of legs ($\mathcal{O}(1)$).



- OK for very inclusive observables.
- Starting point for further simulation.
- Want exclusive final state at the LHC ($\mathcal{O}(100)$).
- Want arbitrary cuts.
- $\rightarrow$ use Monte Carlo methods.

Matrix elements

Where do we get (LO) $|M|^2$ from?

- Most/important simple processes (SM) are 'built in'.
- Calculate yourself (≤ 3 particles in final state).
- Matrix element generators:
  - MadGraph/MadEvent.
  - Comix/AMEGIC (part of Sherpa).
  - HELAC/PHEGAS.
  - Whizard.
  - CalcHEP/CompHEP.

  generate code or event files that can be further processed.
- $\rightarrow$ FeynRules interface to ME generators.

Cross section formula

From Matrix element, we calculate

$$\sigma = \int f_i(x_1,\mu^2) f_j(x_2,\mu^2) \frac{1}{F} \overline{\sum} |M|^2 \qquad\qquad \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}\Phi_n \ ,$$

Cross section formula

From Matrix element, we calculate

$$\sigma = \int f_i(x_1, \mu^2) f_j(x_2, \mu^2) \frac{1}{F} \overline{\sum} |M|^2 \, \Theta(\text{cuts}) \, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}\Phi_n \ ,$$

Cross section formula

From Matrix element, we calculate

$$\sigma = \int f_i(x_1,\mu^2) f_j(x_2,\mu^2) \frac{1}{F} \overline{\sum} |M|^2 \, \Theta(\mathrm{cuts}) \, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}\Phi_n \ ,$$

now,

$$\frac{1}{F} \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}\Phi_n = J(\vec{x}) \prod_{i=1}^{3n-2} \mathrm{d}x_i \qquad \left(\mathrm{d}\Phi_n = (2\pi)^4 \delta^{(4)}(\ldots) \prod_{i=1}^{n} \frac{\mathrm{d}^3\vec{p}}{(2\pi)^3 2E_i} \right)$$

such that

$$\sigma = \int g(\vec{x}) \, \mathrm{d}^{3n-2}\vec{x} \ , \qquad \left(g(\vec{x}) = J(\vec{x}) f_i f_j \overline{\sum} |M|^2 \Theta(\mathrm{cuts}) \right)$$

$$= \frac{1}{N} \sum_{i=1}^{N} \frac{g(\vec{x}_i)}{p(\vec{x}_i)} = \frac{1}{N} \sum_{i=1}^{N} w_i \ .$$

Cross section formula

From Matrix element, we calculate

$$\sigma = \int f_i(x_1, \mu^2) f_j(x_2, \mu^2) \frac{1}{F} \overline{\sum} |M|^2 \, \Theta(\text{cuts}) \, \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}\Phi_n \ ,$$

now,

$$\frac{1}{F} \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}\Phi_n = J(\vec{x}) \prod_{i=1}^{3n-2} \mathrm{d}x_i \qquad \left(\mathrm{d}\Phi_n = (2\pi)^4 \delta^{(4)}(\ldots) \prod_{i=1}^{n} \frac{\mathrm{d}^3 \vec{p}}{(2\pi)^3 2E_i} \right)$$

such that

$$\begin{aligned} \sigma &= \int g(\vec{x}) \, \mathrm{d}^{3n-2}\vec{x} \ , \qquad \left(g(\vec{x}) = J(\vec{x}) f_i f_j \overline{\sum} |M|^2 \Theta(\text{cuts}) \right) \\ &= \frac{1}{N} \sum_{i=1}^{N} \frac{g(\vec{x}_i)}{p(\vec{x}_i)} = \frac{1}{N} \sum_{i=1}^{N} w_i \ . \end{aligned}$$

We generate events $\vec{x}_i$ with weights w_i.

Mini event generator

- We generate pairs $(\vec{x}_i, w_i)$.

Mini event generator

- We generate pairs $(\vec{x}_i, w_i)$.
- Use immediately to book weighted histogram of arbitrary observable (possibly with additional cuts!)

Mini event generator

- We generate pairs $(\vec{x}_i, w_i)$.
- Use immediately to book weighted histogram of arbitrary observable (possibly with additional cuts!)
- Keep event $\vec{x}_i$ with probability

$$P_i = \frac{w_i}{w_{\max}} \ .$$

Generate events with same frequency as in nature!

Mini event generator

- We generate pairs $(\vec{x}_i, w_i)$.
- Use immediately to book weighted histogram of arbitrary observable (possibly with additional cuts!)
- Keep event $\vec{x}_i$ with probability

$$P_i = \frac{w_i}{w_{\max}} \ ,$$

where $w_{\max}$ has to be chosen sensibly.
$\to$ reweighting, when $\max(w_i) = \bar{w}_{\max} > w_{\max}$, as

$$P_i = \frac{w_i}{\bar{w}_{\max}} = \frac{w_i}{w_{\max}} \cdot \frac{w_{\max}}{\bar{w}_{\max}} \ ,$$

i.e. reject events with probability $(w_{\max}/\bar{w}_{\max})$ afterwards.
(can be ignored when #(events with $w_i > \bar{w}_{\max}$) small.)

Mini event generator

- We generate pairs $(\vec{x}_i, w_i)$.
- Use immediately to book weighted histogram of arbitrary observable (possibly with additional cuts!)
- Keep event $\vec{x}_i$ with probability

$$P_i = \frac{w_i}{w_{\max}} \; .$$

Generate events with same frequency as in nature!

Matrix elements

Some comments:

- Use techniques from above to generate events efficiently. Goal: small variance in w_i distribution!

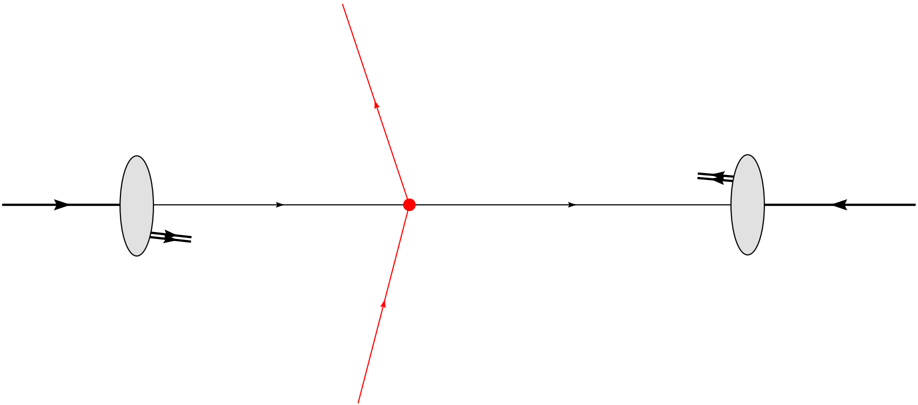
Matrix elements

Some comments:

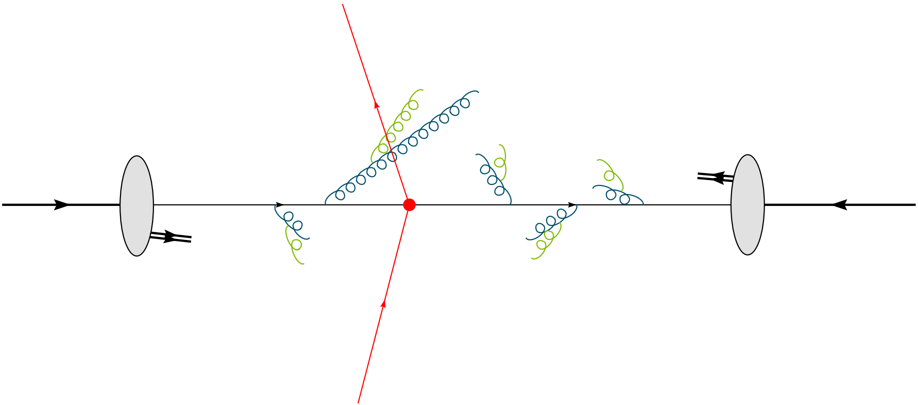
- Use techniques from above to generate events efficiently. Goal: small variance in w_i distribution!
- Clear from above: efficient generation closely tied to knowledge of $f(\vec{x}_i)$, *i.e.* the matrix element's propagator structure.
  $\to$ build phase space generator already while generating ME's automatically.

Parton Showers

Hard matrix element



Hard matrix element → parton showers



Parton showers

Quarks and gluons in final state, pointlike.

Parton showers

Quarks and gluons in final state, pointlike.

- Know short distance (short time) fluctuations from matrix element/Feynman diagrams: $Q \sim$ few GeV to O(TeV).

- Measure hadronic final states, long distance effects, $Q_0 \sim 1\,\mathrm{GeV}$.

Parton showers

Quarks and gluons in final state, pointlike.

- Know short distance (short time) fluctuations from matrix element/Feynman diagrams: $Q \sim$ few GeV to O(TeV).
- Parton shower evolution, multiple gluon emissions become resolvable at smaller scales. TeV $\to 1\,\text{GeV}$.
- Measure hadronic final states, long distance effects, $Q_0 \sim 1\,\text{GeV}$.

Parton showers

Quarks and gluons in final state, pointlike.

- Know short distance (short time) fluctuations from matrix element/Feynman diagrams: $Q \sim$ few GeV to O(TeV).
- Parton shower evolution, multiple gluon emissions become resolvable at smaller scales. TeV $\to$ 1 GeV.
- Measure hadronic final states, long distance effects, $Q_0 \sim 1\,\mathrm{GeV}$.

Dominated by large logs, terms

$$\alpha_S^n \log^{2n} \frac{Q}{Q_0} \sim 1\ .$$

Generated from emissions *ordered* in Q.

Parton showers

Quarks and gluons in final state, pointlike.

- Know short distance (short time) fluctuations from matrix element/Feynman diagrams: $Q \sim$ few GeV to O(TeV).
- Parton shower evolution, multiple gluon emissions become resolvable at smaller scales. TeV $\to 1\,\mathrm{GeV}$.
- Measure hadronic final states, long distance effects, $Q_0 \sim 1\,\mathrm{GeV}$.

Dominated by large logs, terms

$$\alpha_S^n \log^{2n} \frac{Q}{Q_0} \sim 1 \, .$$

Generated from emissions *ordered* in Q.
Soft and/or collinear emissions.

Good starting point: $e^+e^- \to q\bar{q}g$:

Final state momenta in one plane (orientation usually averaged).
Write momenta in terms of

$$x_i = \frac{2p_i \cdot q}{Q^2} \quad (i = 1,2,3) \ ,$$

$$0 \le x_i \le 1 \ , x_1 + x_2 + x_3 = 2 \ ,$$

$$q = (Q,0,0,0) \ ,$$

$$Q \equiv E_{cm} \ .$$

$(x_1, x_2) = (x_q, x_{\bar{q}})$ –plane:

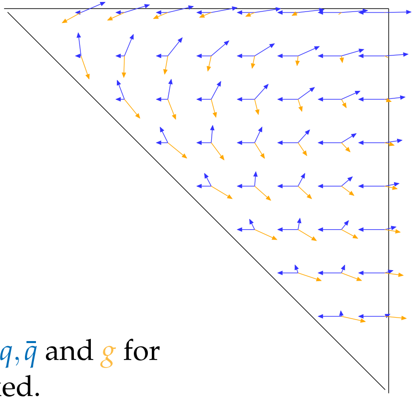
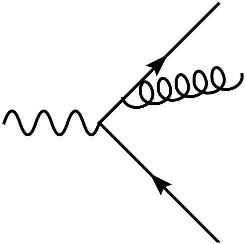


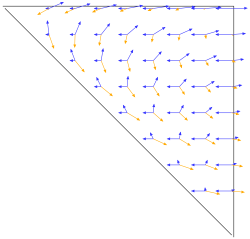
Fig: momentum configuration of $q, \bar{q}$ and g for given point (x_1, x_2), $\bar{q}$ direction fixed.

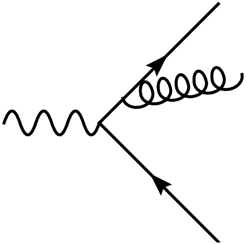


Differential cross section:

$$\frac{d\sigma}{dx_1 dx_2} = \sigma_0 \frac{C_F \alpha_S}{2\pi} \frac{x_1 + x_2}{(1 - x_1)(1 - x_2)}$$

Collinear singularities: $x_1 \to 1$ or $x_2 \to 1$.
Soft singularity: $x_1, x_2 \to 1$.

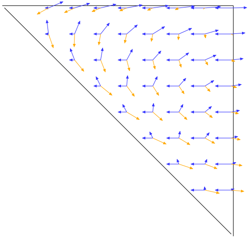




Differential cross section:

$$\frac{d\sigma}{dx_1 dx_2} = \sigma_0 \frac{C_F \alpha_S}{2\pi} \frac{x_1 + x_2}{(1-x_1)(1-x_2)}$$

Collinear singularities: $x_1 \to 1$ or $x_2 \to 1$.
Soft singularity: $x_1, x_2 \to 1$.



Rewrite in terms of x_3 and $\theta = \angle(q, g)$:

$$\frac{d\sigma}{d\cos\theta\, dx_3} = \sigma_0 \frac{C_F \alpha_S}{2\pi} \left[\frac{2}{\sin^2\theta} \frac{1+(1-x_3)^2}{x_3} - x_3 \right]$$

Singular as $\theta \to 0$ and $x_3 \to 0$.

Can separate into two jets as

$$\frac{2\mathrm{d}\cos\theta}{\sin^2\theta} = \frac{\mathrm{d}\cos\theta}{1-\cos\theta} + \frac{\mathrm{d}\cos\theta}{1+\cos\theta}$$
$$= \frac{\mathrm{d}\cos\theta}{1-\cos\theta} + \frac{\mathrm{d}\cos\bar{\theta}}{1-\cos\bar{\theta}}$$
$$\approx \frac{\mathrm{d}\theta^2}{\theta^2} + \frac{\mathrm{d}\bar{\theta}^2}{\bar{\theta}^2}$$

Can separate into two jets as

$$\frac{2\mathrm{d}\cos\theta}{\sin^2\theta} = \frac{\mathrm{d}\cos\theta}{1-\cos\theta} + \frac{\mathrm{d}\cos\theta}{1+\cos\theta}$$
$$= \frac{\mathrm{d}\cos\theta}{1-\cos\theta} + \frac{\mathrm{d}\cos\bar{\theta}}{1-\cos\bar{\theta}}$$
$$\approx \frac{\mathrm{d}\theta^2}{\theta^2} + \frac{\mathrm{d}\bar{\theta}^2}{\bar{\theta}^2}$$

So, we rewrite $\mathrm{d}\sigma$ in collinear limit as

$$\mathrm{d}\sigma = \sigma_0 \sum_{\text{jets}} \frac{\mathrm{d}\theta^2}{\theta^2} \frac{\alpha_S}{2\pi} C_F \frac{1+(1-z)^2}{z^2} \mathrm{d}z$$

e^+e^- annihilation

Can separate into two jets as

$$\frac{2\mathrm{d}\cos\theta}{\sin^2\theta} = \frac{\mathrm{d}\cos\theta}{1-\cos\theta} + \frac{\mathrm{d}\cos\theta}{1+\cos\theta}$$
$$= \frac{\mathrm{d}\cos\theta}{1-\cos\theta} + \frac{\mathrm{d}\cos\bar{\theta}}{1-\cos\bar{\theta}}$$
$$\approx \frac{\mathrm{d}\theta^2}{\theta^2} + \frac{\mathrm{d}\bar{\theta}^2}{\bar{\theta}^2}$$

So, we rewrite $\mathrm{d}\sigma$ in collinear limit as

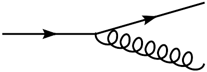
$$\mathrm{d}\sigma = \sigma_0 \sum_{\text{jets}} \frac{\mathrm{d}\theta^2}{\theta^2}\frac{\alpha_S}{2\pi} C_F \frac{1+(1-z)^2}{z^2}\mathrm{d}z$$
$$= \sigma_0 \sum_{\text{jets}} \frac{\mathrm{d}\theta^2}{\theta^2}\frac{\alpha_S}{2\pi} P(z)\mathrm{d}z$$

with DGLAP splitting function $P(z)$.

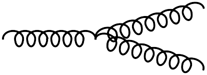
Collinear limit

Universal DGLAP splitting kernels for collinear limit:

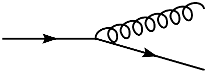
$$\mathrm{d}\sigma = \sigma_0 \sum_{\text{jets}} \frac{\mathrm{d}\theta^2}{\theta^2} \frac{\alpha_S}{2\pi} P(z) \mathrm{d}z$$



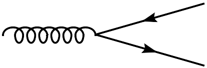
$$P_{q\to qg}(z) = C_F \frac{1+z^2}{1-z}$$



$$P_{g\to gg}(z) = C_A \frac{(1-z(1-z))^2}{z(1-z)}$$



$$P_{q\to gq}(z) = C_F \frac{1+(1-z)^2}{z}$$



$$P_{g\to qq}(z) = T_R(1-2z(1-z))$$

Collinear limit

Universal DGLAP splitting kernels for collinear limit:

$$\mathrm{d}\sigma = \sigma_0 \sum_{\mathrm{jets}} \frac{\mathrm{d}\theta^2}{\theta^2} \frac{\alpha_S}{2\pi} P(z) \mathrm{d}z$$

Note: Other variables may equally well characterize the collinear limit:

$$\frac{\mathrm{d}\theta^2}{\theta^2} \sim \frac{\mathrm{d}Q^2}{Q^2} \sim \frac{\mathrm{d}p_\perp^2}{p_\perp^2} \sim \frac{\mathrm{d}\tilde{q}^2}{\tilde{q}^2} \sim \frac{\mathrm{d}t}{t}$$

whenever $Q^2, p_\perp^2, t \to 0$ means "collinear".

Collinear limit

Universal DGLAP splitting kernels for collinear limit:

$$\mathrm{d}\sigma = \sigma_0 \sum_{\text{jets}} \frac{\mathrm{d}\theta^2}{\theta^2} \frac{\alpha_S}{2\pi} P(z) \mathrm{d}z$$

Note: Other variables may equally well characterize the collinear limit:

$$\frac{\mathrm{d}\theta^2}{\theta^2} \sim \frac{\mathrm{d}Q^2}{Q^2} \sim \frac{\mathrm{d}p_\perp^2}{p_\perp^2} \sim \frac{\mathrm{d}\tilde{q}^2}{\tilde{q}^2} \sim \frac{\mathrm{d}t}{t}$$

whenever $Q^2, p_\perp^2, t \to 0$ means "collinear".

- θ: HERWIG
- Q^2: PYTHIA ≤ 6.3, old SHERPA.
- $p_\perp$: PYTHIA ≥ 6.4, ARIADNE, Catani–Seymour showers in HERWIG++ and SHERPA.
- $\tilde{q}$: Herwig++.

Resolution

Need to introduce resolution t_0, e.g. a cutoff in $p_\perp$. Prevent us from the singularity at $\theta \to 0$.

Emissions below t_0 are unresolvable.

Finite result due to virtual corrections:

 = finite.

unresolvable + virtual emissions are included in Sudakov form factor via unitarity (see below!).

Towards multiple emissions

Starting point: factorisation in collinear limit, single emission.

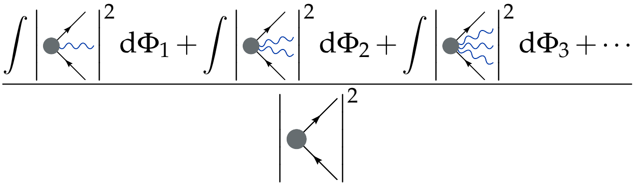
$$\sigma_{2+1}(t_0) = \sigma_2(t_0) \int_{t_0}^{t} \frac{\mathrm{d}t'}{t'} \int_{z_-}^{z_+} \mathrm{d}z \, \frac{\alpha_S}{2\pi} \hat{P}(z) \; = \sigma_2(t_0) \int_{t_0}^{t} \mathrm{d}t \, W(t) \; .$$

Starting point: factorisation in collinear limit, single emission.

$$\sigma_{2+1}(t_0) = \sigma_2(t_0) \int_{t_0}^{t} \frac{\mathrm{d}t'}{t'} \int_{z_-}^{z_+} \mathrm{d}z \, \frac{\alpha_S}{2\pi} \hat{P}(z) \; = \sigma_2(t_0) \int_{t_0}^{t} \mathrm{d}t \, W(t) \; .$$

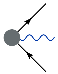
Simple example:
Multiple photon emissions, strongly ordered in t.
We want

$$W_{\rm sum} = \sum_{n=1} W_{2+n} =$$

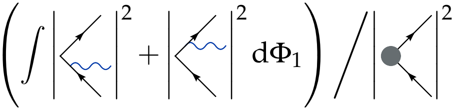


for any number of emissions.

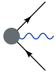
Towards multiple emissions

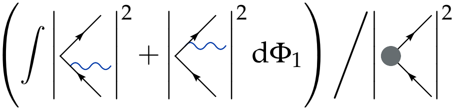
$(n = 1)$ 

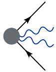
$$W_{2+1} =$$

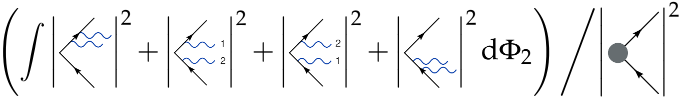


$$= \frac{2}{1!}\int_{t_0}^{t} \mathrm{d}t\, W(t)\ .$$

$(n = 1)$ 

$W_{2+1} =$  $= \dfrac{2}{1!} \displaystyle\int_{t_0}^{t} \mathrm{d}t\, W(t)$.

$(n = 2)$ 

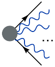
$W_{2+2} =$ 

$$= 2^2 \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t''\, W(t') W(t'') = \frac{2^2}{2!} \left(\int_{t_0}^{t} \mathrm{d}t\, W(t) \right)^2 .$$

We used

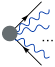
$$\int_{t_0}^{t} \mathrm{d}t_1 \ldots \int_{t_0}^{t_{n-1}} \mathrm{d}t_n \;\; W(t_1) \ldots W(t_n) = \frac{1}{n!} \left(\int_{t_0}^{t} \mathrm{d}t\, W(t) \right)^n .$$

Towards multiple emissions

Easily generalized to n emissions  by induction. *i.e.*

$$W_{2+n} = \frac{2^n}{n!} \left(\int_{t_0}^{t} \mathrm{d}t\, W(t) \right)^n$$

Towards multiple emissions

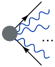
Easily generalized to n emissions  by induction. *i.e.*

$$W_{2+n} = \frac{2^n}{n!} \left(\int_{t_0}^{t} \mathrm{d}t \, W(t) \right)^n$$

So, in total we get

$$\sigma_{>2}(t_0) = \sigma_2(t_0) \sum_{k=1}^{\infty} \frac{2^k}{k!} \left(\int_{t_0}^{t} \mathrm{d}t \, W(t) \right)^k = \sigma_2(t_0) \left(\mathrm{e}^{2 \int_{t_0}^{t} \mathrm{d}t \, W(t)} - 1 \right)$$

Towards multiple emissions

Easily generalized to n emissions  by induction. *i.e.*

$$W_{2+n} = \frac{2^n}{n!} \left(\int_{t_0}^{t} \mathrm{d}t\, W(t) \right)^n$$

So, in total we get

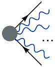
$$\sigma_{>2}(t_0) = \sigma_2(t_0) \sum_{k=1}^{\infty} \frac{2^k}{k!} \left(\int_{t_0}^{t} \mathrm{d}t\, W(t) \right)^k = \sigma_2(t_0) \left(\mathrm{e}^{2 \int_{t_0}^{t} \mathrm{d}t\, W(t)} - 1 \right)$$

$$= \sigma_2(t_0) \left(\frac{1}{\Delta^2(t_0, t)} - 1 \right)$$

Sudakov Form Factor

$$\Delta(t_0, t) = \exp\left[- \int_{t_0}^{t} \mathrm{d}t\, W(t) \right]$$

Towards multiple emissions

Easily generalized to n emissions  by induction. *i.e.*

$$W_{2+n} = \frac{2^n}{n!}\left(\int_{t_0}^{t} \mathrm{d}t\, W(t)\right)^n$$

So, in total we get

$$\sigma_{>2}(t_0) = \sigma_2(t_0)\sum_{k=1}^{\infty}\frac{2^k}{k!}\left(\int_{t_0}^{t}\mathrm{d}t\, W(t)\right)^k = \sigma_2(t_0)\left(\mathrm{e}^{2\int_{t_0}^{t}\mathrm{d}t\, W(t)} - 1\right)$$

$$= \sigma_2(t_0)\left(\frac{1}{\Delta^2(t_0,t)} - 1\right)$$

Sudakov Form Factor in QCD

$$\Delta(t_0,t) = \exp\left[-\int_{t_0}^{t}\mathrm{d}t\, W(t)\right] = \exp\left[-\int_{t_0}^{t}\frac{\mathrm{d}t}{t}\int_{z_-}^{z_+}\frac{\alpha_S(z,t)}{2\pi}\hat{P}(z,t)\mathrm{d}z\right]$$

Sudakov form factor

Note that

$$\sigma_{\text{all}} = \sigma_2 + \sigma_{>2} = \sigma_2 + \sigma_2 \left(\frac{1}{\Delta^2(t_0,t)} - 1 \right) ,$$
$$\Rightarrow \Delta^2(t_0,t) = \frac{\sigma_2}{\sigma_{\text{all}}} .$$

$$\text{Two jet rate} = \Delta^2 = P^2(\text{No emission in the range } t \to t_0) .$$

Sudakov form factor = No emission probability .

Often $\Delta(t_0,t) \equiv \Delta(t)$.

- Hard scale t, typically CM energy or $p_\perp$ of hard process.
- Resolution t_0, two partons are resolved as two entities if inv mass or relative $p_\perp$ above t_0.
- P^2 (not P), as we have two legs that evolve independently.