Introduction to Monte Carlo Event Generators

Stefan Gieseke

Institut für Theoretische Physik
KIT

Lectures at MCnet Beijing School 2021
University of Chinese Acedemy of Sciences, Beijing, China
28 June-2 July 2021

Karlsruhe Institute of Technology

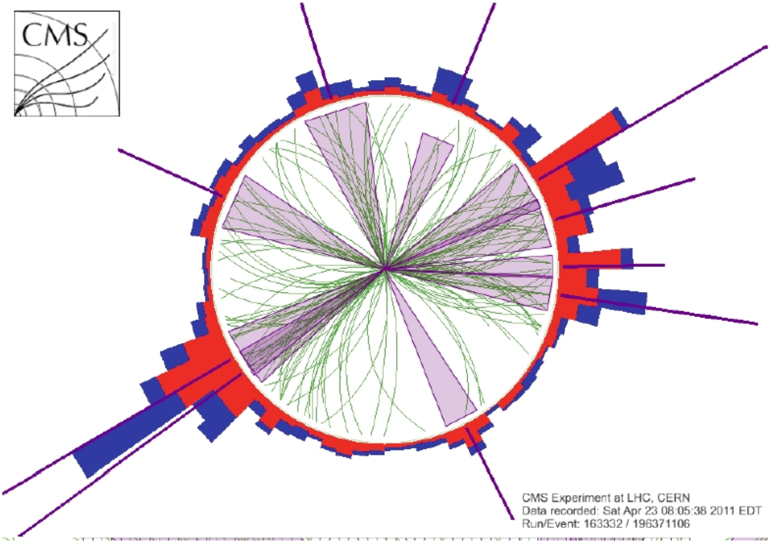
Bundesministerium für Bildung und Forschung

TRR 257 - Particle Physics Phenomenology after the Higgs Discovery

Motivation: jets

[Google Images]

Motivation: jets (at LHC of course)



[CMS 2011]

Why Monte Carlos?

We want to understand

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

Why Monte Carlos?

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

$\uparrow$

Full detector simulation.

$\uparrow$

Fully exclusive hadronic final state.

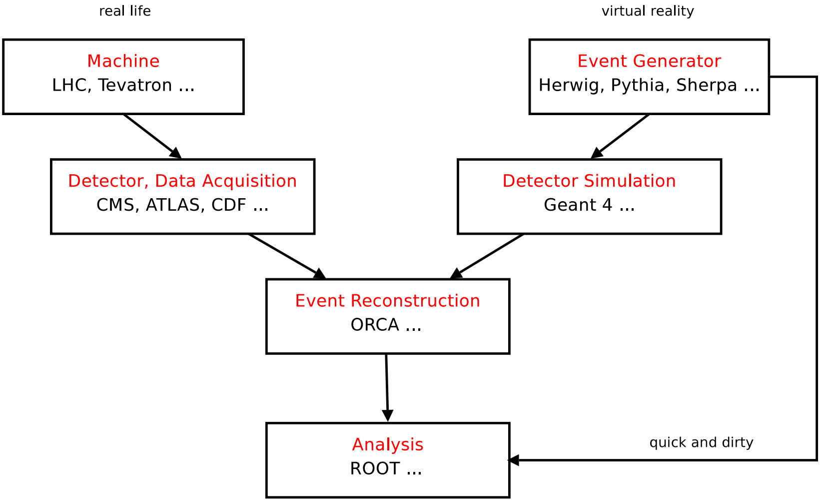
$\uparrow$

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.

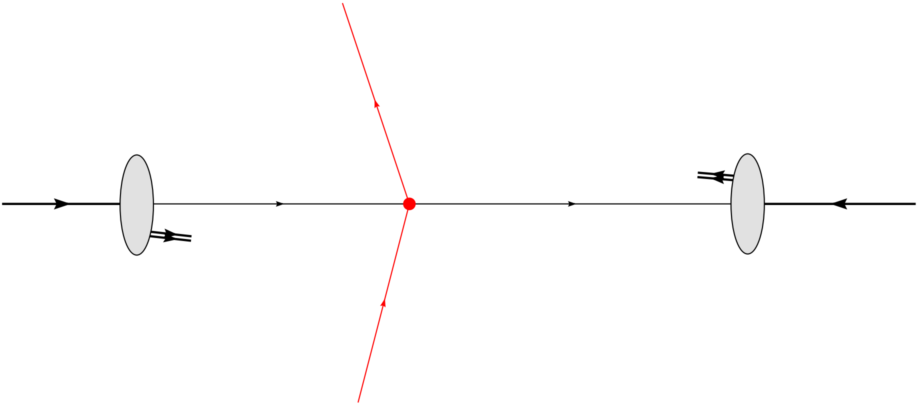
- Rates and topologies of final states.
- Background studies.
- Detector Design.
- Detector Performance Studies (Acceptance).

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

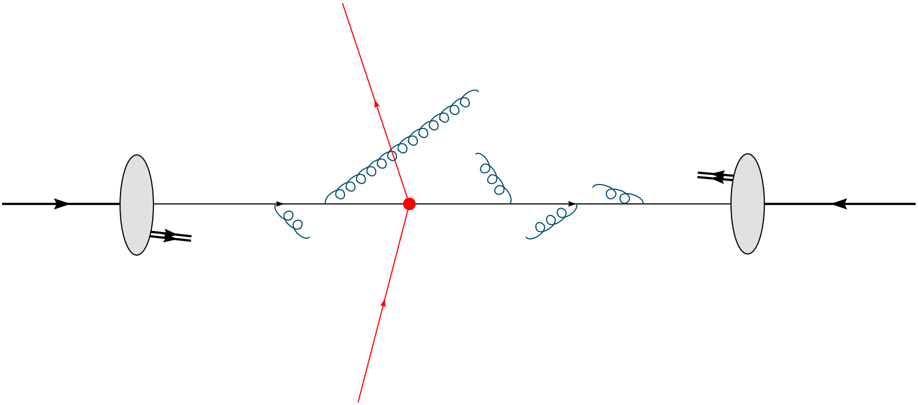
$$|A|^2 \longrightarrow \text{Probability.}$$

pp Event Generator

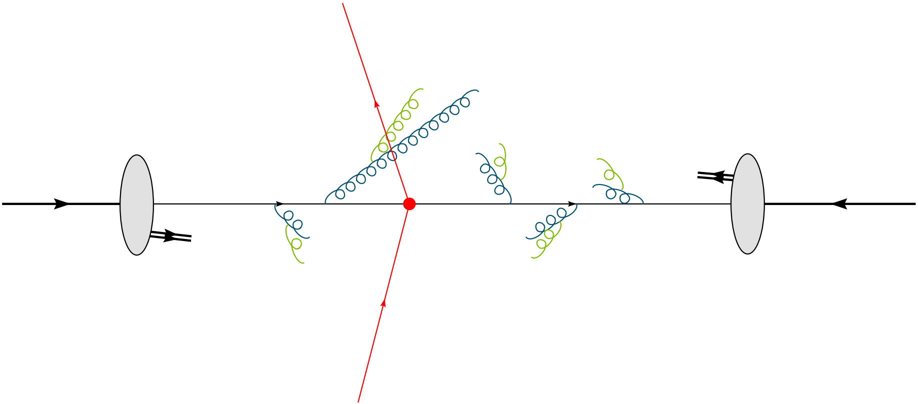
pp Event Generator



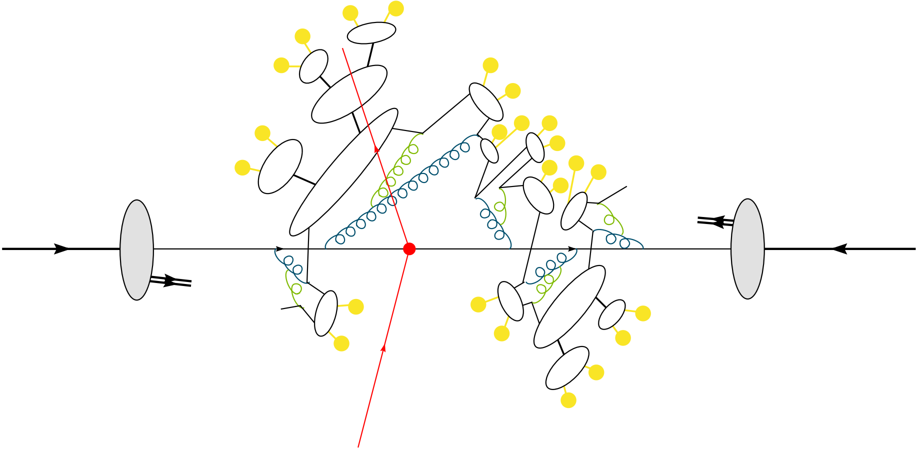
pp Event Generator



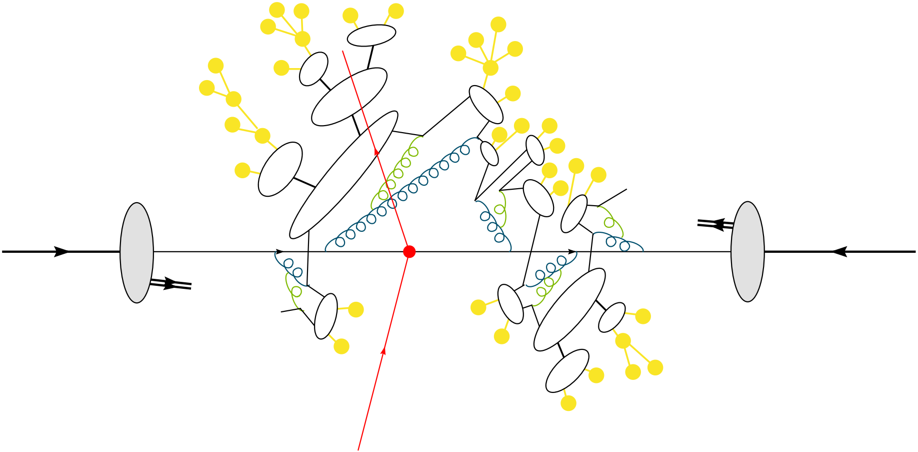
pp Event Generator



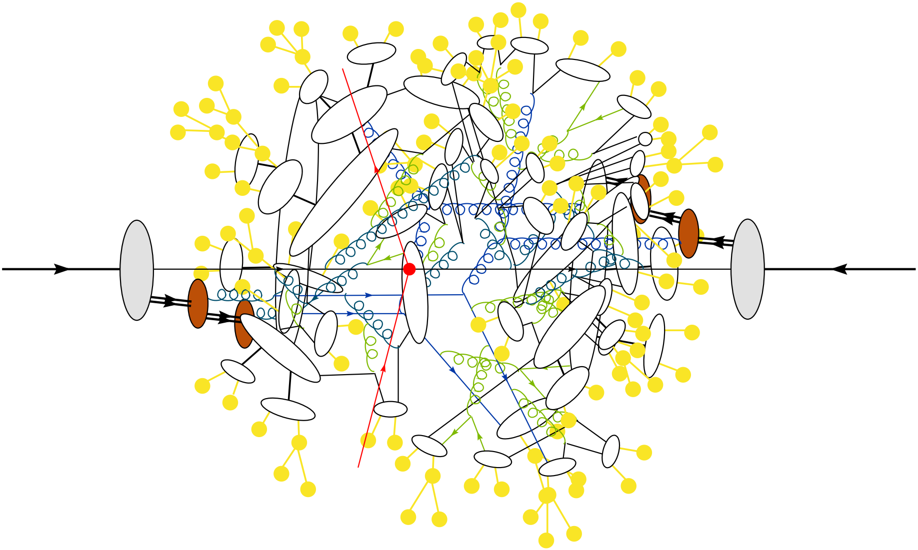
pp Event Generator



pp Event Generator



pp Event Generator



Divide and conquer

Partonic cross section from Feynman diagrams

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

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

Underlying event from multiple partonic interactions

$$d\sigma \longleftarrow d\sigma(\text{QCD } 2 \to 2)$$

Plan for these lectures

- Monte Carlo Methods
- Hard Scattering
- Parton Showers

Monte Carlo Methods

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. Adaptive MC, VEGAS.
5. Multichannel.
6. Mini event generator in particle physics.

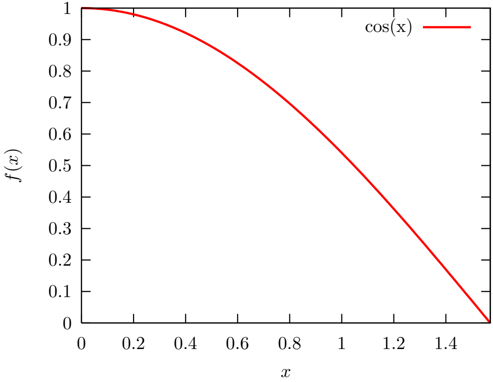
Probability

Probability density:

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

is probability to find value x.

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



Probability

Probability density:

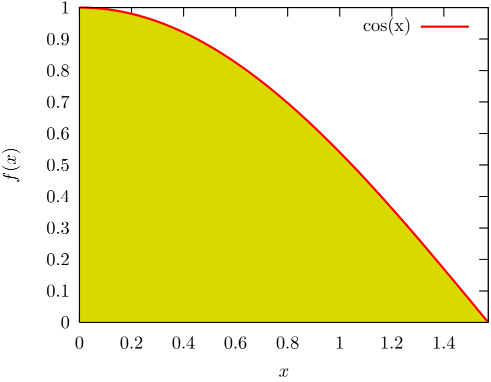
$$dP = f(x)\,dx$$

is probability to find value x.

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

is called *probability distribution*.

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



Probability

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

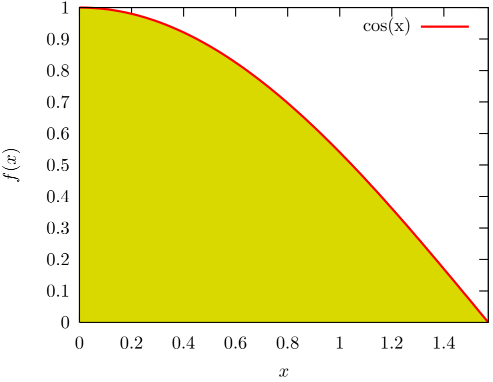
Probability density:

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

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.

Hit and Miss

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

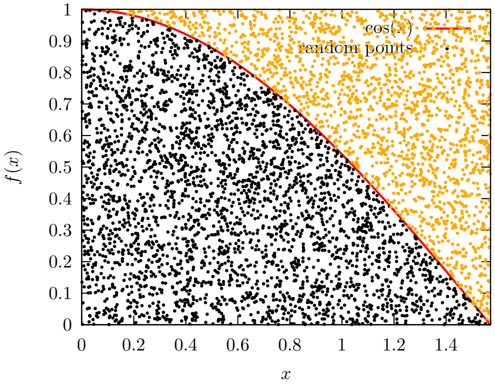
Hit and miss method:

- 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.



Hit and Miss

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

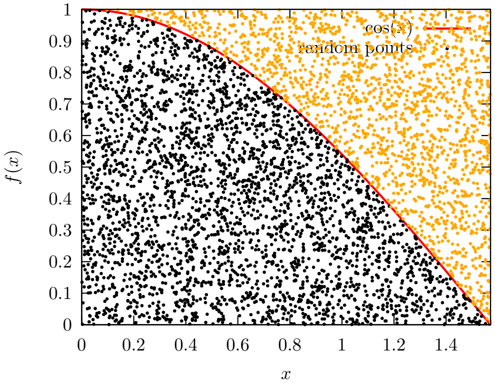
Hit and miss method:

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

Then

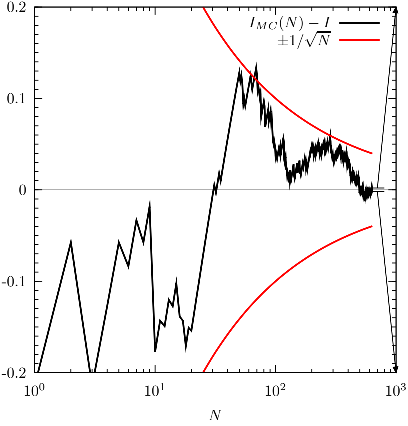
$$I \approx V \frac{N_{\text{hit}}}{N}.$$

approaches 1 again in our example.



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.

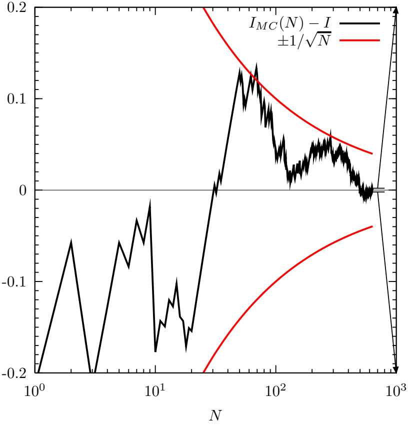
Hit and Miss



How well does it converge?

Error $1/\sqrt{N}$.

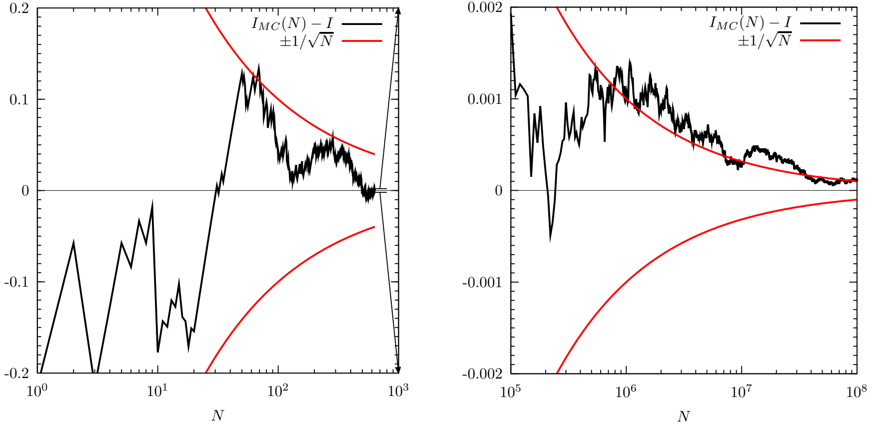
Hit and Miss



More points, zoom in…

Error $1/\sqrt{N}$.

Hit and Miss



Error $1/\sqrt{N}$.

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:

$$
\begin{aligned}
I &= \int_{x_0}^{x_1} f(x)\,dx \\
&= (x_1 - x_0)\langle f(x)\rangle
\end{aligned}
$$

(Riemann integral).

Simple MC integration

Mean value theorem of integration:

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

(Riemann integral).

Simple MC integration

Mean value theorem of integration:

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

(Riemann integral).

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

Simple MC integration

Mean value theorem of integration:

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

(Riemann integral).

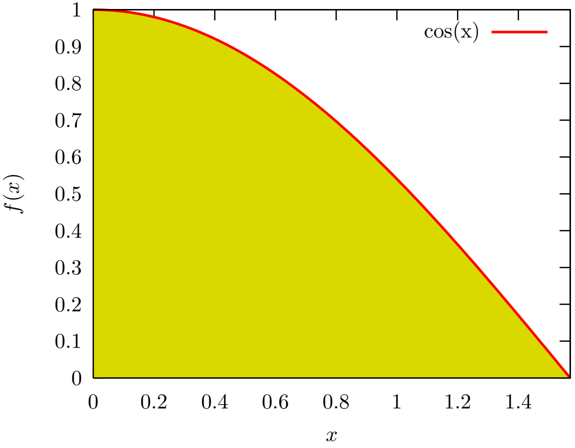
Sum doesn't depend on ordering
$\longrightarrow$ randomize x_i.

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

Simple MC integration

Pictorially:

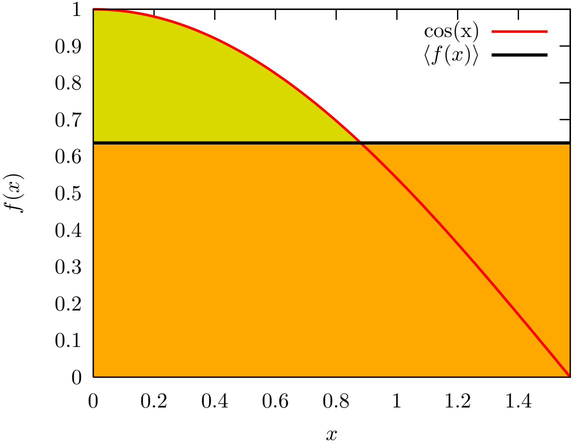
$$I = \int_{x_0}^{x_1} f(x)\, dx$$
$$= (x_1 - x_0)\langle f(x)\rangle$$



Simple MC integration

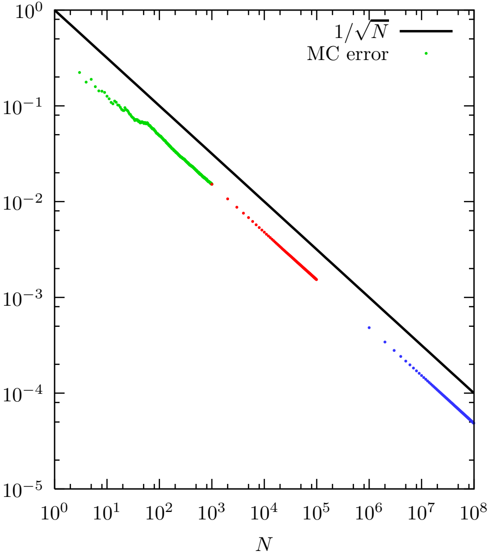
Pictorially:

$$I = \int_{x_0}^{x_1} f(x)\, dx$$
$$= (x_1 - x_0)\langle f(x)\rangle$$



Simple MC integration

What's the error?



Again, looks like

$$\sigma \sim \frac{1}{\sqrt{N}}$$

Simple MC integration

What's the error?

We can calculate it (central limit theorem for the average):

In general: *Crude MC*

$$
\begin{aligned}
I &= \int f dV \\
&\approx V\langle f\rangle \pm V\sqrt{\frac{\langle f\rangle^2 - \langle f^2\rangle}{N}} \\
&\approx V\langle f\rangle \pm V\frac{\sigma}{\sqrt{N}}
\end{aligned}
$$

Simple MC integration

What's the error?

We can calculate it (central limit theorem for the average):

Our example: $\cos(x), 0 \leq x \leq \pi/2$,
compute σ_{MC} from

$$\langle f \rangle = \frac{1}{N} \sum_{i=1}^{N} f(x_i)$$
$$\langle f^2 \rangle = \frac{1}{N} \sum_{i=1}^{N} f^2(x_i).$$

Simple MC integration

What's the error?

We can calculate it (central limit theorem for the average):

Compute σ directly ($V = \pi/2$):

$$\langle f \rangle = \int_0^{\pi/2} \cos(x)\, dx = 1$$
$$\langle f^2 \rangle = \int_0^{\pi/2} \cos^2(x)\, dx = \frac{\pi}{4}$$

then

$$\sigma = \sqrt{1^2 - \frac{\pi}{4}} \approx 0.4633.$$

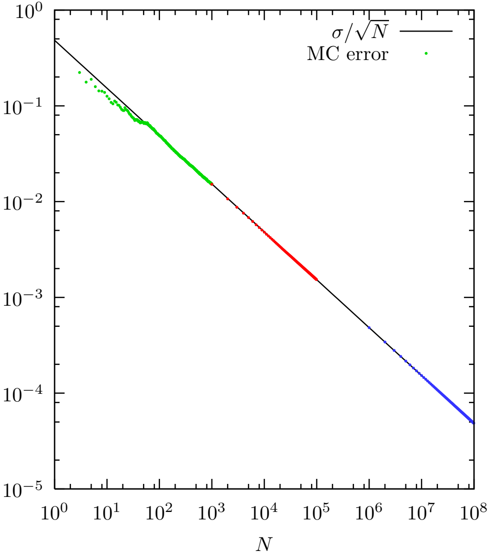
Simple MC integration

What's the error?

Now, compare

$$\sigma_{MC} = \frac{0.4633}{\sqrt{N}}$$

with error estimate from MC.



Simple MC integration

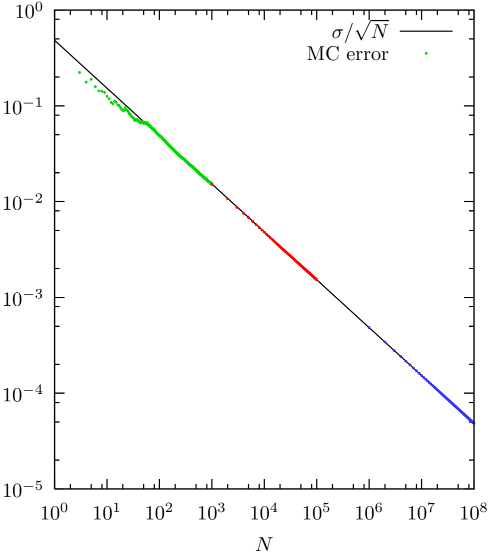
What's the error?

Now, compare

$$\sigma_{MC} = \frac{0.4633}{\sqrt{N}}$$

with error estimate from MC.

Spot on.



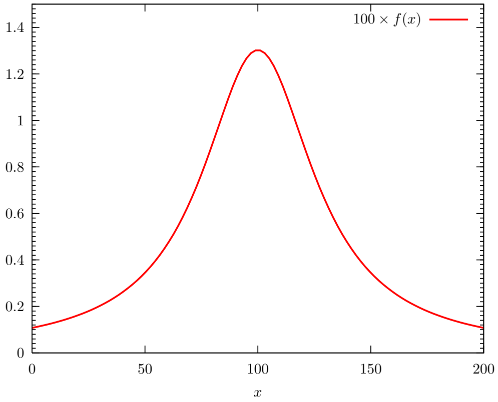
Inverting the Integral

Another basic MC method, based on the observation that

Probability $\sim$ *Area*

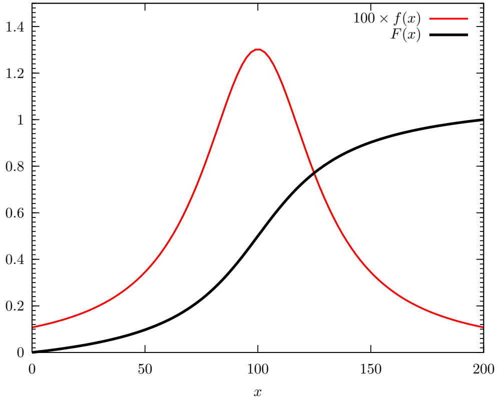
Inverting the Integral

- Probability density $f(x)$. Not necessarily normalized.



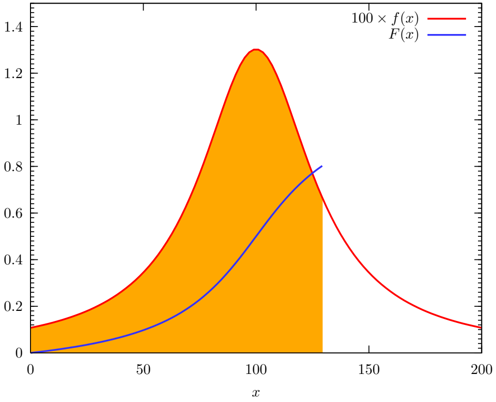
Inverting the Integral

- Probability density $f(x)$. Not necessarily normalized.
- Integral $F(x)$ known,



Inverting the Integral

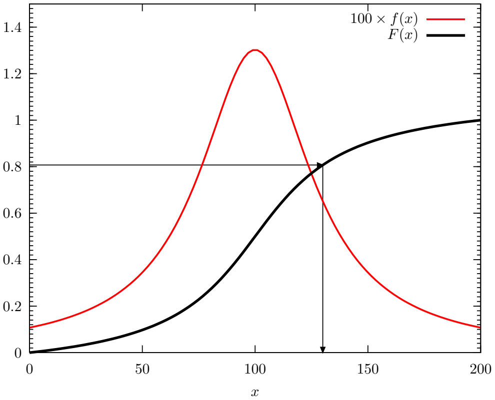
- Probability density $f(x)$. Not necessarily normalized.
- Integral $F(x)$ known,
- $P(x < x_s) = F(x_s)$.



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,

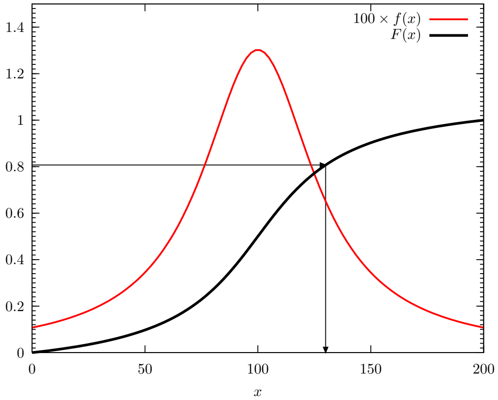
$$\int_{x_0}^{x} dP = r$$



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,

$$\int_{x_0}^{x} dP = r$$



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

Another basic MC method, based on the observation that

Probability ∼ Area

Sample x according to $f(x)$ with

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

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.

Importance sampling

Consider error term:

$$E = \left\langle \frac{f^2}{p^2} \right\rangle - \left\langle \frac{f}{p} \right\rangle^2 = \int \frac{f^2}{p^2} p \mathrm{d}V - \left[\int \frac{f}{p} p \mathrm{d}V \right]^2$$

$$= \int \frac{f^2}{p} \mathrm{d}V - \left[\int f \, \mathrm{d}V \right]^2 .$$

Importance sampling

Consider error term:

$$E = \int \frac{f^2}{p}\,\mathrm{d}V - \left[\int f\,\mathrm{d}V\right]^2 .$$

Best choice of p? Minimises $E \to$ functional variation of error term with (normalized) p:

$$0 = \delta E = \delta \left(\int \frac{f^2}{p}\,\mathrm{d}V - \left[\int f\,\mathrm{d}V\right]^2 + \lambda \int p\,\mathrm{d}V \right)$$

$$= \int \left(-\frac{f^2}{p^2} + \lambda \right) \mathrm{d}V\,\delta p\ ,$$

Importance sampling

Consider error term:

$$E = \int \frac{f^2}{p} \, \mathrm{d}V - \left[\int f \, \mathrm{d}V \right]^2 .$$

Best choice of p? Minimises $E \to$ functional variation of error term with (normalized) p:

$$0 = \delta E = \int \left(-\frac{f^2}{p^2} + \lambda \right) \mathrm{d}V \delta p ,$$

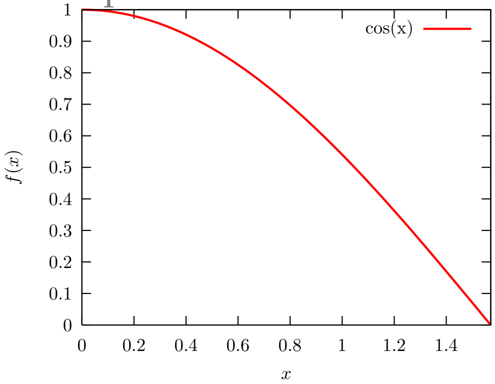
hence

$$p = \frac{|f|}{\sqrt{\lambda}} = \frac{|f|}{\int |f| \, \mathrm{d}V} .$$

Choose p as close to f as possible.

Importance sampling — example

Improving $\cos(x)$ sampling,



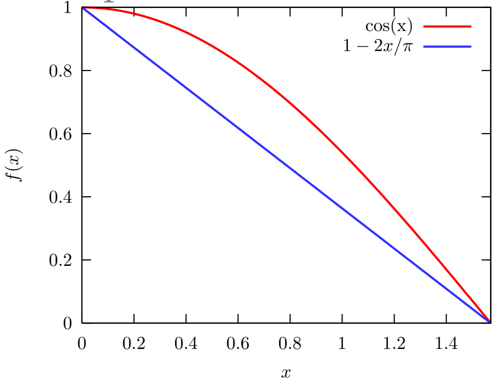
Importance sampling — example

Improving $\cos(x)$ sampling,

$$I = \int_0^{\pi/2} \cos(x)\, dx$$

$$= \int_0^{\pi/2} \frac{\cos(x)}{1-\frac{2}{\pi}x} \left(1-\frac{2}{\pi}x\right) dx$$

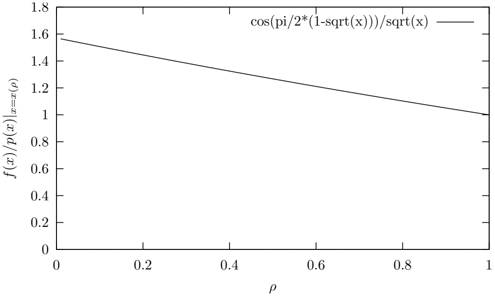
$$= \int_0^{1} \left.\frac{\cos(x)}{1-\frac{2}{\pi}x}\right|_{x=x(\rho)} d\rho \ .$$



Importance sampling — example

Improving $\cos(x)$ sampling,

$$I = \int_0^{\pi/2} \cos(x)\, dx$$
$$= \int_0^{\pi/2} \frac{\cos(x)}{1-\frac{2}{\pi}x} \left(1-\frac{2}{\pi}x\right) dx$$
$$= \int_0^1 \left. \frac{\cos(x)}{1-\frac{2}{\pi}x} \right|_{x=x(\rho)} d\rho\ .$$



Sample x with *inverting the integral* technique (flat random number ρ),

$$x = \frac{\pi}{2}\left(1-\sqrt{1-\rho}\right) \hat{=} \frac{\pi}{2}(1-\sqrt{\rho}) \quad \left(I = \int_0^1 \frac{\cos\left(\frac{\pi}{2}\left(1-\sqrt{\rho}\right)\right)}{\sqrt{\rho}}\, d\rho.\right)$$

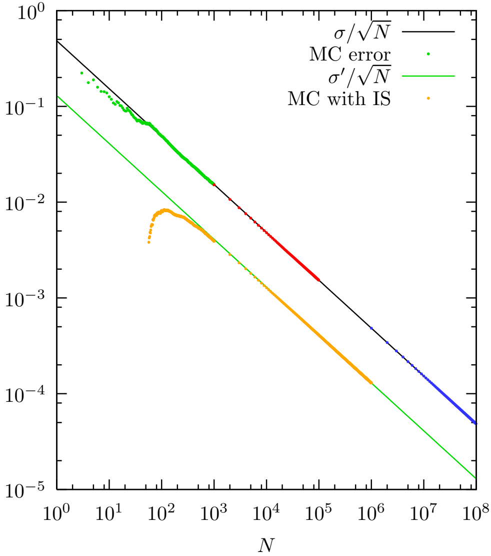
Importance sampling — example

Improving $\cos(x)$ sampling,

much better convergence,

about 80% "accepted events".

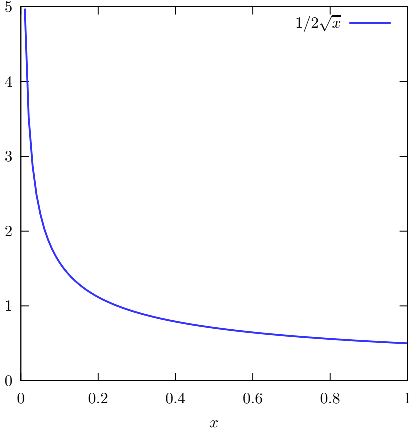
Reduced variance ($\sigma' = 0.027$)
$\Rightarrow$ better efficiency.



Importance sampling — better example

More interesting for divergent integrands, eg

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



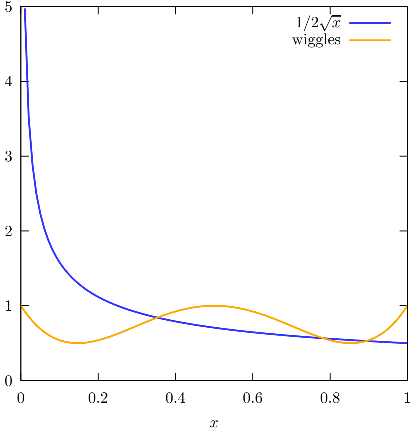
Importance sampling — better example

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

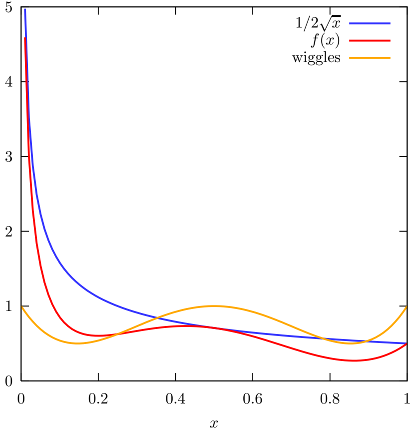
$$\frac{1}{2\sqrt{x}} \; ,$$

with some wiggles,

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

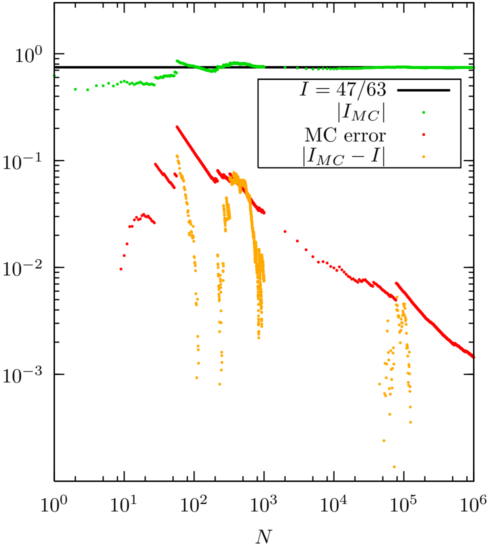
i.e. we want to integrate

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



Importance sampling — better example

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

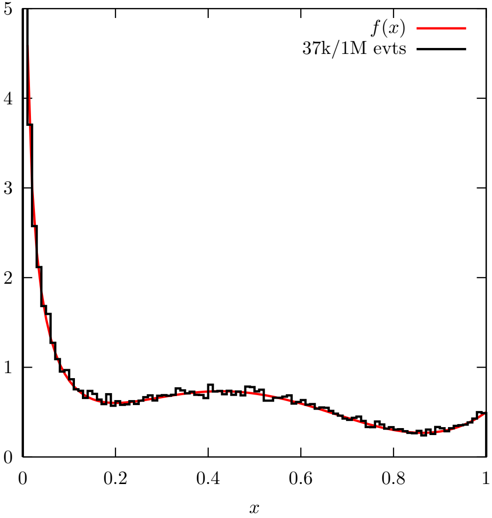


Importance sampling — better example

Want events:

use hit+mass variant here:

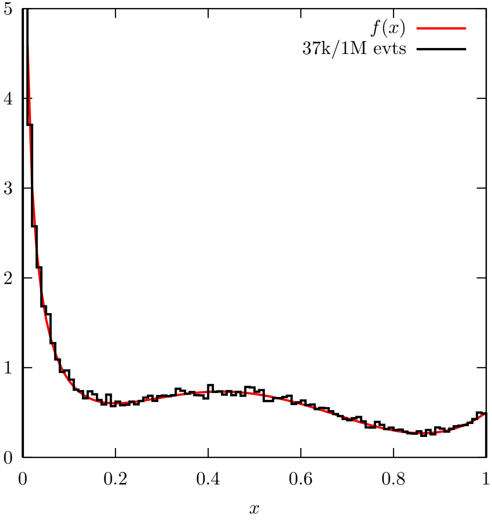
- Choose new random number r
- $w = f(x)$ in this case.
- if $r < w/w_{\max}$ then "hit".
- MC efficiency = hit$/N$.



Importance sampling — better example

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 — better example

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

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

so,

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

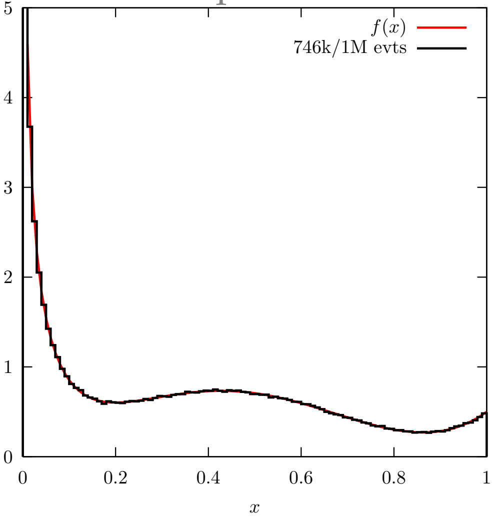
x sampled with *inverting the integral* from flat random numbers ρ, $x = \rho^2$.

Importance sampling — better example

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

with

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



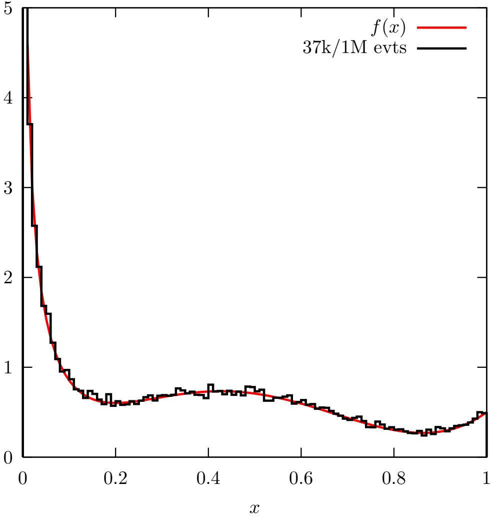
Events generated with $w_{\max} = 1$, as $p(x) \le 1$, no guesswork needed here! Now, we get 74.6% MC efficiency.

Importance sampling — better example

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

with

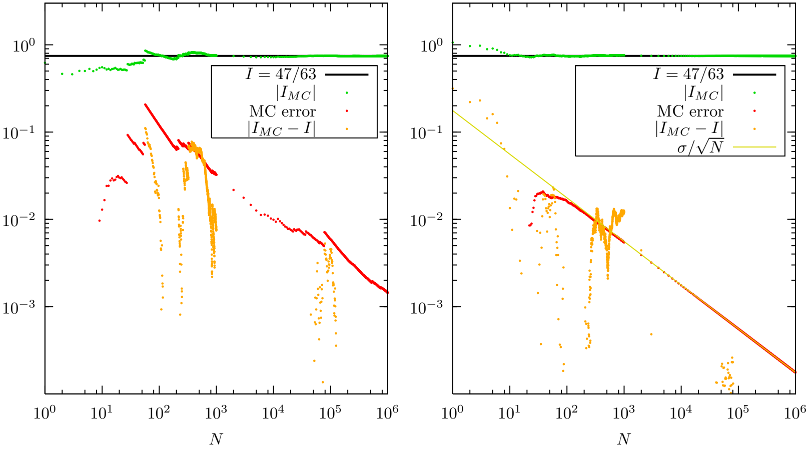
$$\rho = \sqrt{x}, \qquad d\rho = \frac{dx}{2\sqrt{x}}$$



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 — better example

Crude MC vs Importance sampling.



100× more events needed to reach same accuracy.

Importance sampling — another useful example

Breit–Wigner peaks appear in many realistic MEs for cross sections and decays.

$$I = \int_{s_0}^{s_1} \frac{ds}{(s-m^2)^2 + m^2\Gamma^2}$$

Importance sampling — another useful example

Breit–Wigner peaks appear in many realistic MEs for cross sections and decays.

$$I = \int_{s_0}^{s_1} \frac{ds}{(s-m^2)^2 + m^2\Gamma^2} = \frac{1}{m\Gamma}\int_{y_0}^{y_1} \frac{dy}{y^2+1} \qquad (y = \frac{s-m^2}{m\Gamma})$$

$$= \frac{1}{m\Gamma} \arctan \frac{s-m^2}{m\Gamma}\Bigg|_{s_0}^{s_1}$$

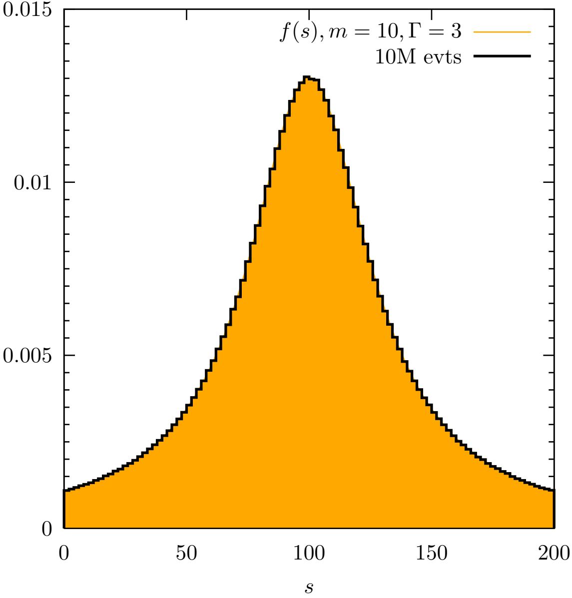
Inverting the integral gives ("tan mapping").

$$f(s) = \frac{m\Gamma}{(s-m^2)^2 + m^2\Gamma^2} \,,$$

$$F(s) = \arctan \frac{s-m^2}{m\Gamma} = \rho \,,$$

$$F^{-1}(\rho) = m^2 + m\Gamma \tan\rho \,.$$

Importance sampling — another useful example



VEGAS

- Classic algorithm.
- Automatic impotance sampling.
- Adopt grid size.
- Often used for multidimensional integration.
- Very robust.

VEGAS

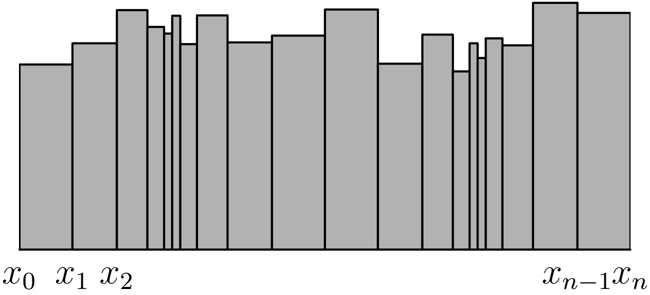
- start with equidistant grid $x_0, x_1, \ldots, x_N$.
- Sample a number of points $\big(x_{s,i}, f(x_{s,i})\big)$, compute first estimate of integral as $\langle f \rangle$.
- Resize grid:
  choose x'_i such that contribution from partial areas inside $x_i < x < x_{i+1}$ to integral is $\langle f \rangle / N$.
- Remember, optimal $p(x) \sim |f(x)|$.
- Sample again with same number of points into every bin $x_i < x < x_{i+1}$. Results in step weight function with steps

$$p_i = \frac{1}{N(x_i - x_{i-1})}\,, \qquad x_i < x < x_{i+1}\,.$$

- $\Rightarrow$ Sample often where density is high.

VEGAS

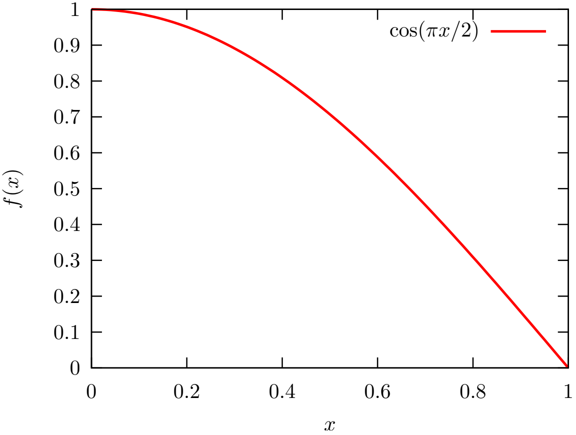
Rebinning:



[from T. Ohl, VAMP]

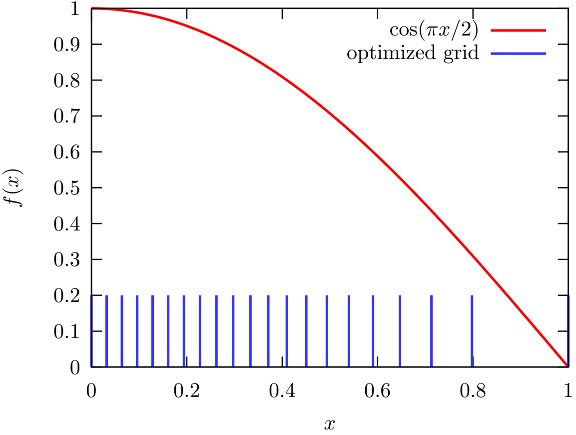
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20, 100$
Convergence improved.



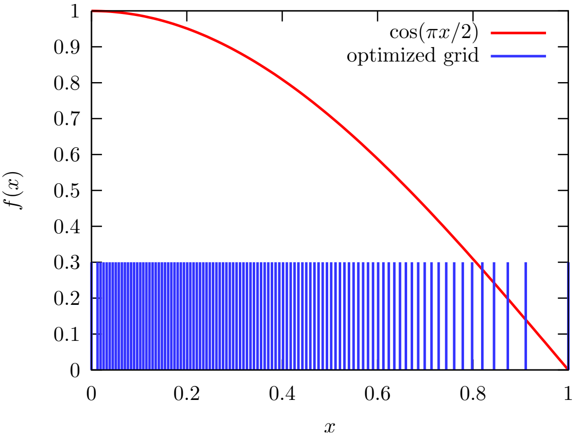
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20, 100$
Convergence improved.



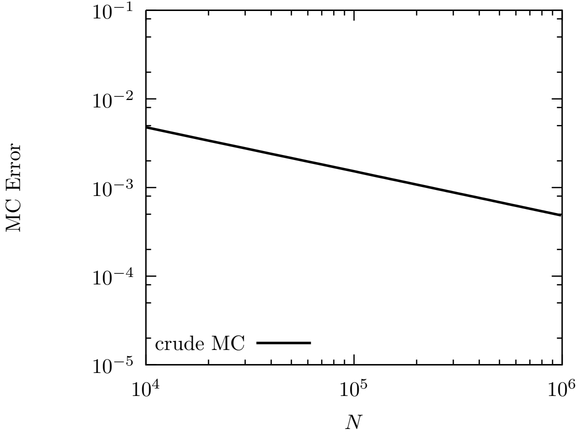
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20, 100$
Convergence improved.



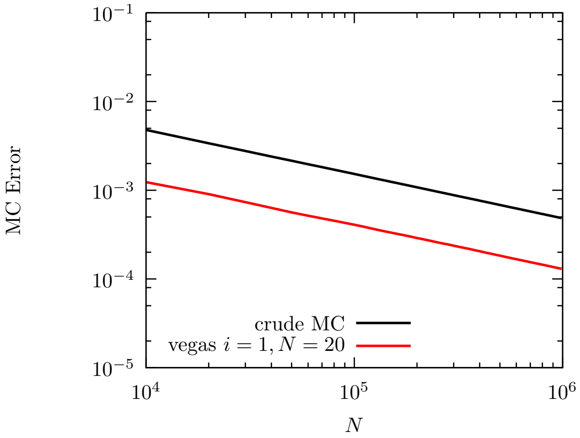
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20,100$
Convergence improved.



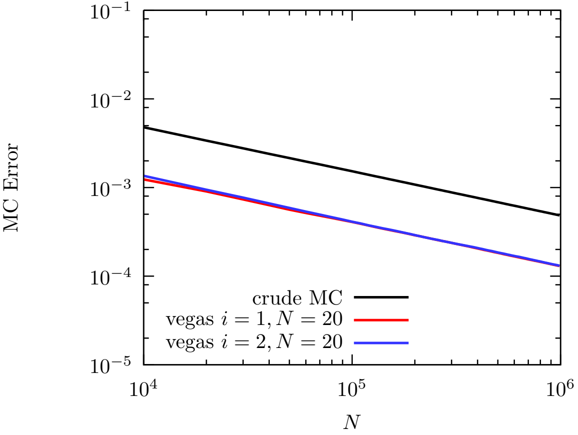
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20, 100$
Convergence improved.



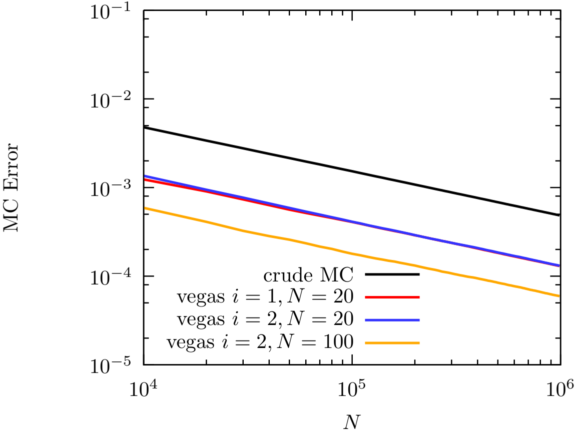
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20, 100$
Convergence improved.



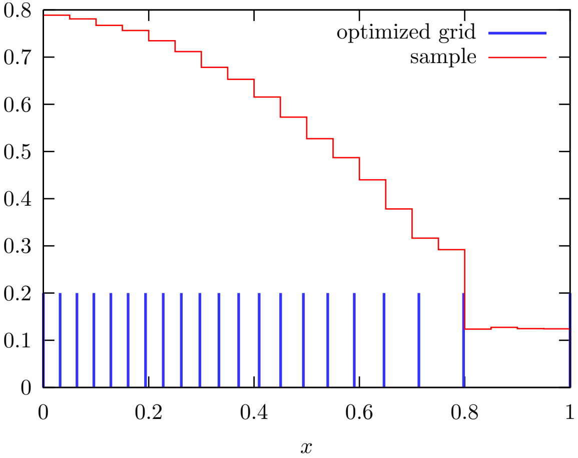
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20, 100$
Convergence improved.



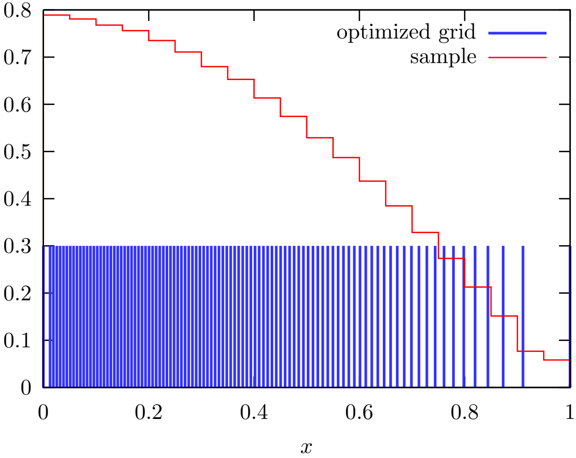
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20, 100$
Convergence improved.



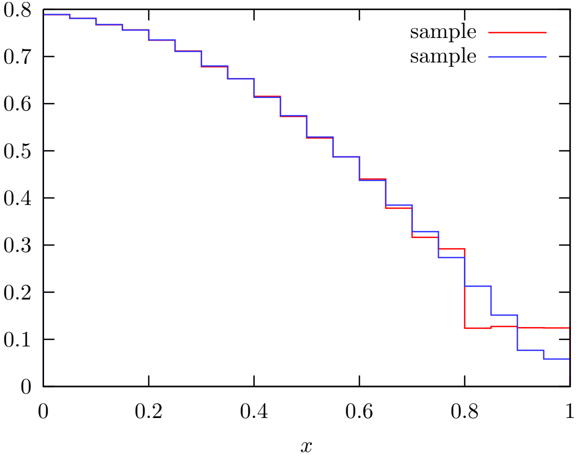
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20, 100$
Convergence improved.



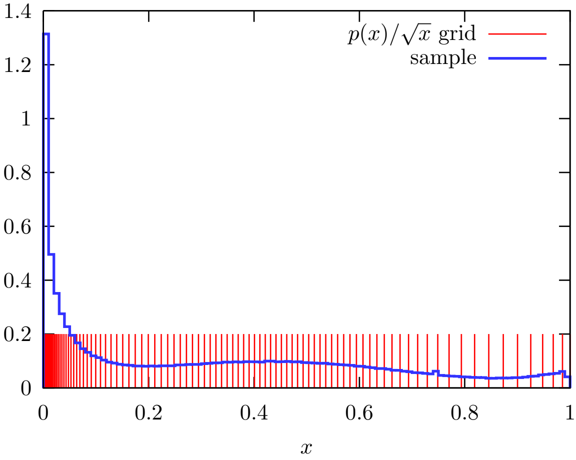
VEGAS

Example: $\cos(\frac{\pi x}{2})$
$N_{\text{grid}} = 20, 100$
Convergence improved.



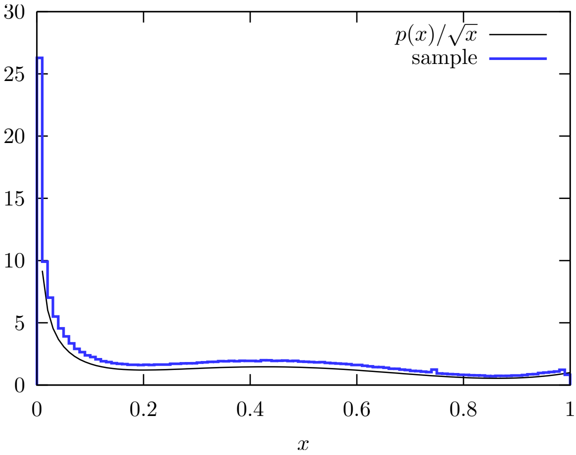
VEGAS

Second example: $p(x)/\sqrt{x}$ (divergence with wiggles)



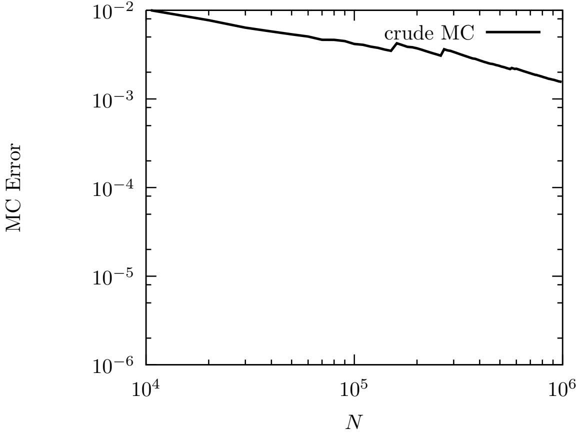
VEGAS

Second example:
$p(x)/\sqrt{x}$
(divergence with wiggles)



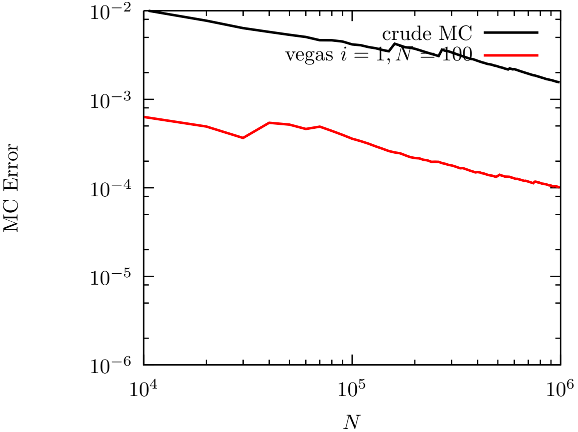
VEGAS

Second example: $p(x)/\sqrt{x}$ (divergence with wiggles)



VEGAS

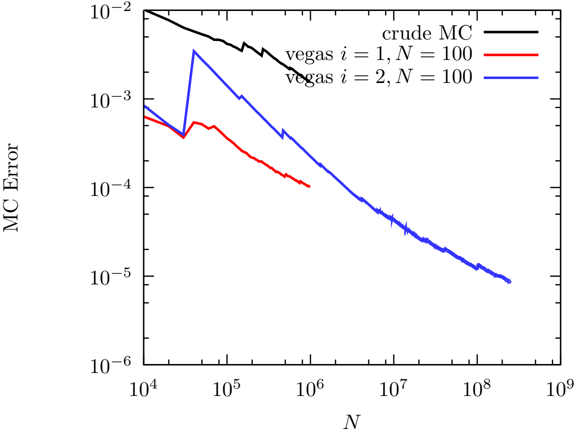
Second example:
$p(x)/\sqrt{x}$
(divergence with wiggles)



Acc 10^{-4} after $N = 10^6$ comparable with 'inverting the integral'.

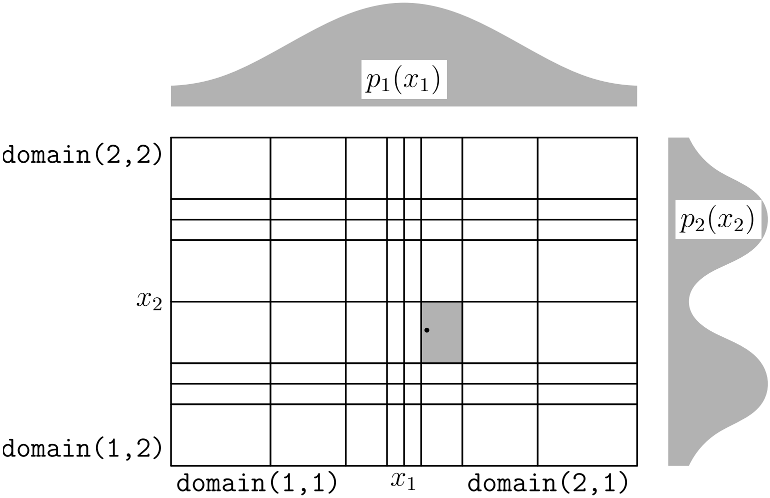
VEGAS

Second example:
$p(x)/\sqrt{x}$
(divergence with wiggles)



VEGAS

Problem to adapt in multiple dimensions:

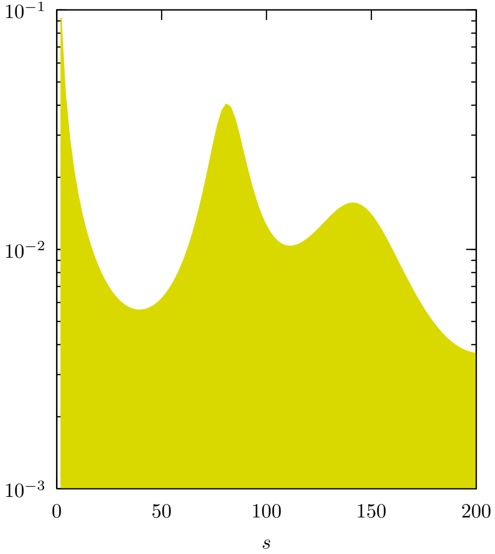


[from T. Ohl, VAMP]

Multichannel MC

Typical problem:

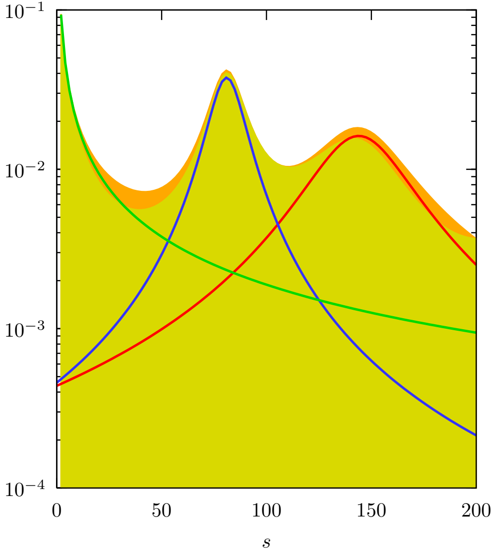
- $f(s)$ has multiple peaks ($\times$ wiggles from ME).



Multichannel MC

Typical problem:

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

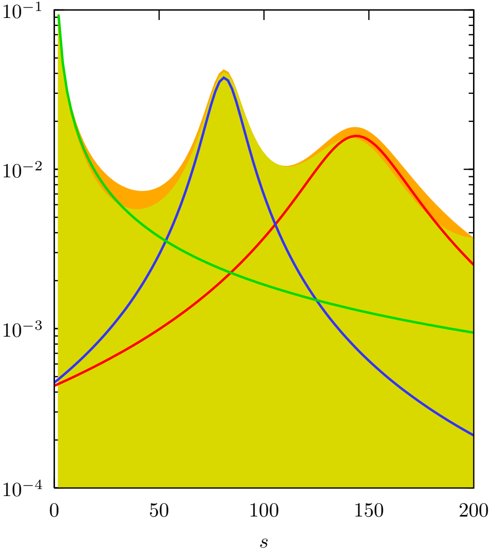


Multichannel MC

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

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

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

Multichannel MC

Now rewrite

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

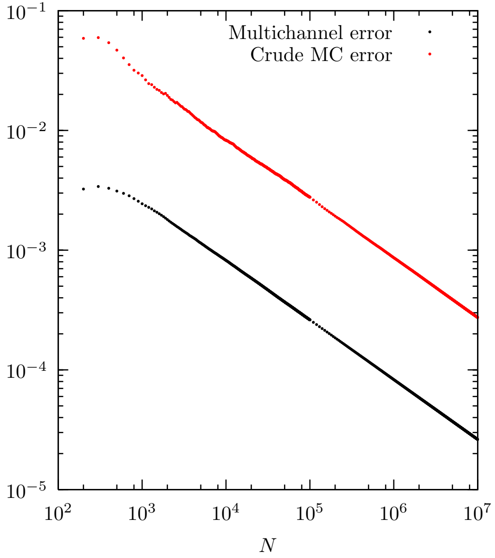
Now $g_i(s)\,ds = d\rho_i$ (inverting the integral).

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

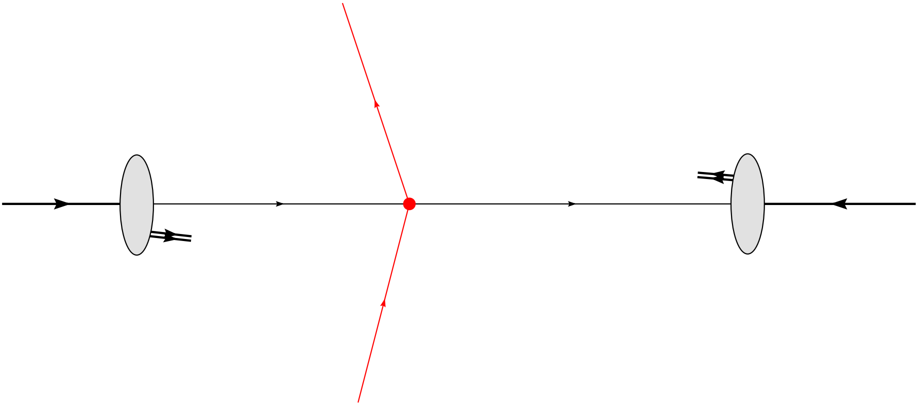
Works quite well:



Hard Scattering

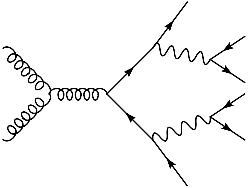
Hard scattering

Hard scattering



Matrix elements

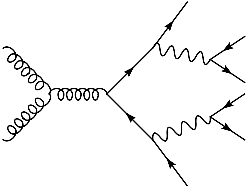
- Perturbation theory/Feynman diagrams give us (fairly accurate) final states for a few number of legs ($O(1)$).



- OK for very inclusive observables.

Matrix elements

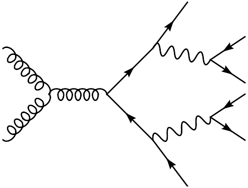
- Perturbation theory/Feynman diagrams give us (fairly accurate) final states for a few number of legs ($O(1)$).



- 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 ($O(1)$).



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

Matrix elements

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

- Most/important simple processes (SM and BSM) 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.

Also NLO mostly automatically available.
See "Matching and Merging".

Cross section formula

From Matrix element, we calculate

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

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

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

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

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

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

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

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

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

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.

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 common Monte Carlo techniques to generate events efficiently. Goal: small variance in w_i distribution!

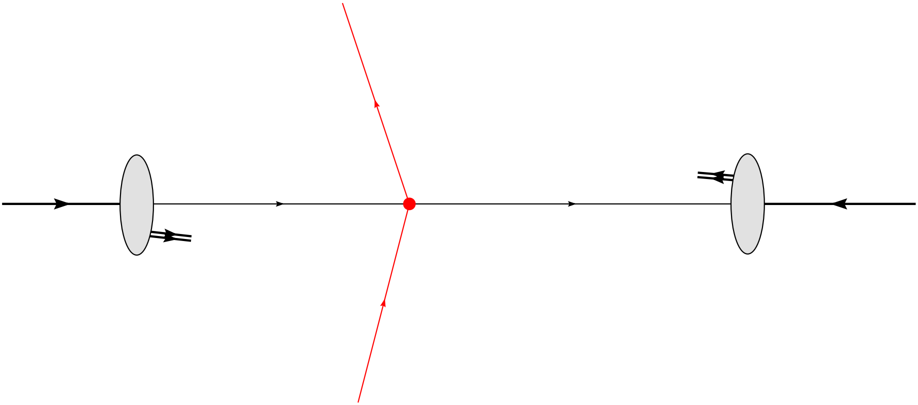
Matrix elements

Some comments:

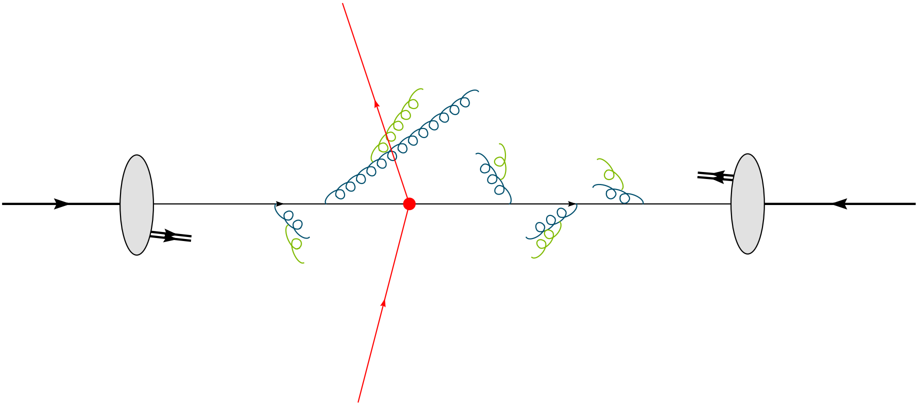
- Use common Monte Carlo techniques to generate events efficiently. Goal: small variance in w_i distribution!
- Efficient generation closely tied to knowledge of $f(\vec{x}_i)$, *i.e.* the matrix element's propagator structure.
  $\rightarrow$ build phase space generator already while generating ME's 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, $Q_0 \sim 1\,\text{GeV}$.

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, $Q_0 \sim 1\,\mathrm{GeV}$.

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.

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, $Q_0 \sim 1\,\mathrm{GeV}$.

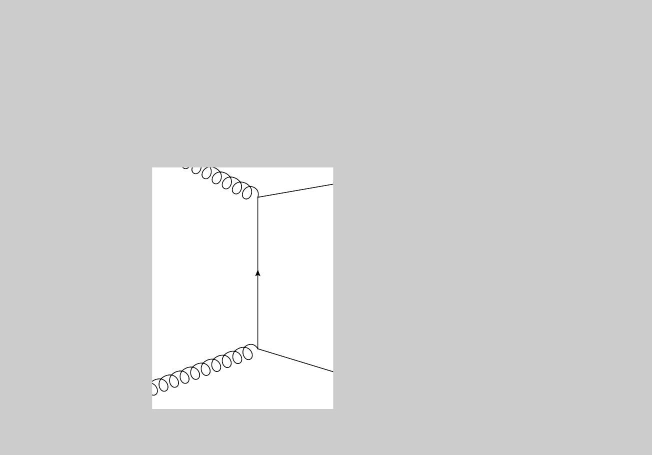
Dominated by large logs, terms

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

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

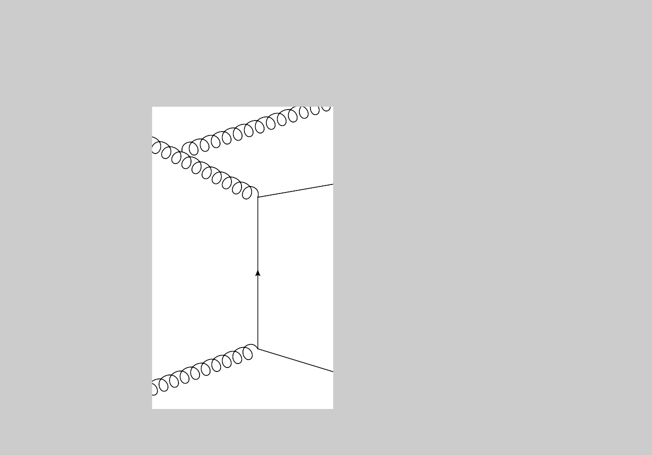
ME approximated by parton cascade

Evolution in scale, typically $Q \sim 1\,\text{TeV}$ down to $Q \sim 1\,\text{GeV}$.



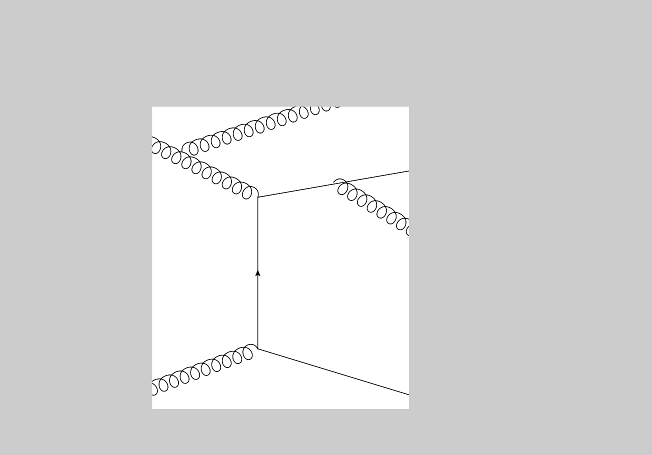
ME approximated by parton cascade

Evolution in scale, typically $Q \sim 1\,\text{TeV}$ down to $Q \sim 1\,\text{GeV}$.



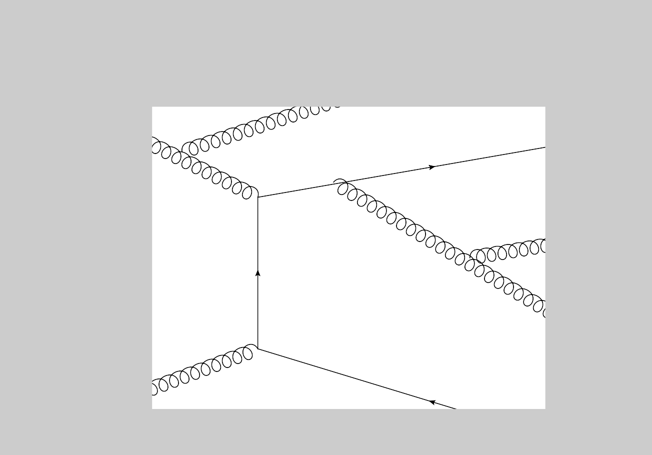
ME approximated by parton cascade

Evolution in scale, typically $Q \sim 1\,\text{TeV}$ down to $Q \sim 1\,\text{GeV}$.



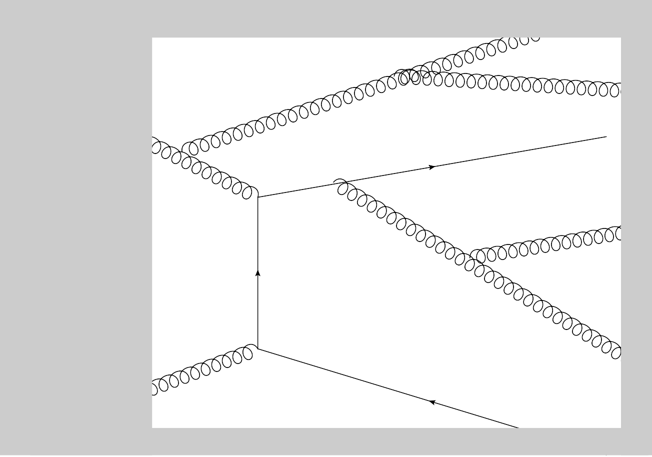
ME approximated by parton cascade

Evolution in scale, typically $Q \sim 1\,\text{TeV}$ down to $Q \sim 1\,\text{GeV}$.



ME approximated by parton cascade

Evolution in scale, typically $Q \sim 1\,\text{TeV}$ down to $Q \sim 1\,\text{GeV}$.



e^+e^- annihilation

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:

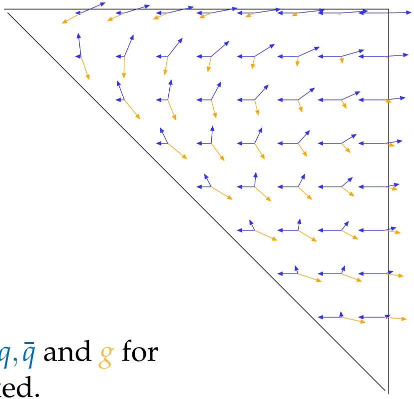


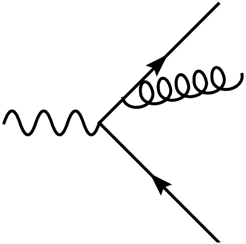
Fig: momentum configuration of $q, \bar{q}$ and g for given point (x_1, x_2), $\bar{q}$ direction fixed.

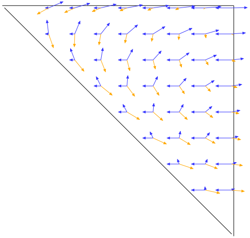
e^+e^- annihilation

Differential cross section:

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

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

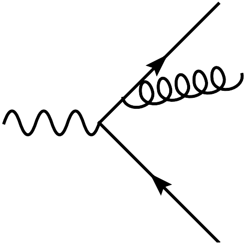




e^+e^- annihilation

Differential cross section:

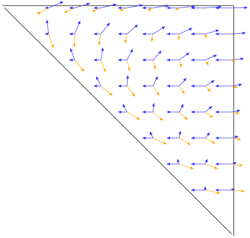
$$\frac{\mathrm{d}\sigma}{\mathrm{d}x_1\mathrm{d}x_2} = \sigma_0 \frac{C_F \alpha_S}{2\pi} \frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)}$$



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 $\theta = \angle(q, g)$:

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



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

Can separate into two jets as

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

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

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

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

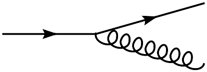
$$\mathrm{d}\sigma = \sigma_0 \sum_{\text{jets}} \frac{\mathrm{d}\theta^2}{\theta^2} \frac{\alpha_S}{2\pi} C_F \frac{1+(1-z)^2}{z} \mathrm{d}z$$
$$= \sigma_0 \sum_{\text{jets}} \frac{\mathrm{d}\theta^2}{\theta^2} \frac{\alpha_S}{2\pi} P(z) \mathrm{d}z$$

with DGLAP splitting function $P(z)$.

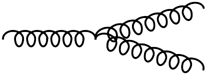
Collinear limit

Universal DGLAP splitting kernels for collinear limit:

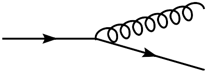
$$\mathrm{d}\sigma = \sigma_0 \sum_{\text{jets}} \frac{\mathrm{d}\theta^2}{\theta^2} \frac{\alpha_S}{2\pi} P(z) \mathrm{d}z$$



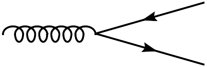
$$P_{q\to qg}(z) = C_F \frac{1+z^2}{1-z}$$



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



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



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

Collinear limit

Universal DGLAP splitting kernels for collinear limit:

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

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".

Collinear limit

Universal DGLAP splitting kernels for collinear limit:

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

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

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

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

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

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

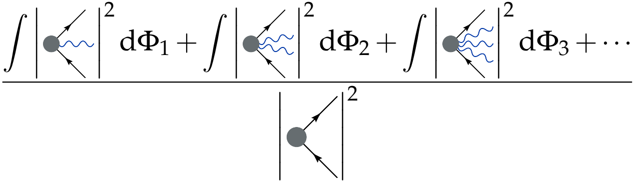
Towards multiple emissions

Starting point: factorisation in collinear limit, single emission.

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

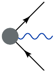
Simple example:
Multiple photon emissions, strongly ordered in t.
We want

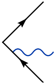
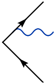
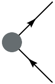
$$W_{\mathrm{sum}} = \sum_{n=1} W_{2+n} =$$



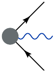
for any number of emissions.

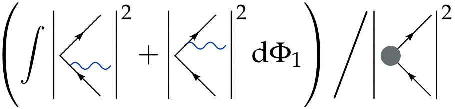
Towards multiple emissions

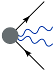
$(n=1)$ 

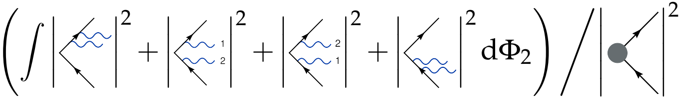
$W_{2+1} = \Bigg(\int \Big|$  $\Big|^2 + \Big|$  $\Big|^2 \mathrm{d}\Phi_1 \Bigg) \Big/ \Big|$  $\Big|^2 = \frac{2}{1!} \int_{t_0}^{t} \mathrm{d}t\, W(t)\ .$

Towards multiple emissions

$(n=1)$ 

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

$(n=2)$ 

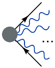
$W_{2+2} =$ 

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

We used

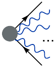
$$\int_{t_0}^{t} \mathrm{d}t_1 \ldots \int_{t_0}^{t_{n-1}} \mathrm{d}t_n\;\; W(t_1)\ldots W(t_n) = \frac{1}{n!}\left(\int_{t_0}^{t} \mathrm{d}t\, W(t)\right)^n \;.$$

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

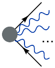
Easily generalized to n emissions  by induction. *i.e.*

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

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.*

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

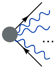
So, in total we get

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

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.

Sudakov form factor from Markov property

Unitarity

$$P(\text{“some emission”}) + P(\text{“no emission”}) \\ = P(0 < t \leq T) + \bar{P}(0 < t \leq T) = 1\ .$$

Multiplication law (no memory)

$$\bar{P}(0 < t \leq T) = \bar{P}(0 < t \leq t_1)\bar{P}(t_1 < t \leq T)$$

Sudakov form factor from Markov property

Unitarity

$$P(\text{“some emission”}) + P(\text{“no emission”}) = P(0 < t \leq T) + \bar{P}(0 < t \leq T) = 1 \, .$$

Multiplication law (no memory)

$$\bar{P}(0 < t \leq T) = \bar{P}(0 < t \leq t_1)\bar{P}(t_1 < t \leq T)$$

Then subdivide into n pieces: $t_i = \frac{i}{n}T, 0 \leq i \leq n$.

$$\bar{P}(0 < t \leq T) = \lim_{n\to\infty} \prod_{i=0}^{n-1} \bar{P}(t_i < t \leq t_{i+1}) = \lim_{n\to\infty} \prod_{i=0}^{n-1} \big(1 - P(t_i < t \leq t_{i+1})\big)$$

$$= \exp\left(- \lim_{n\to\infty} \sum_{i=0}^{n-1} P(t_i < t \leq t_{i+1}) \right) = \exp\left(- \int_0^T \frac{\mathrm{d}P(t)}{\mathrm{d}t} \mathrm{d}t \right) \, .$$

Sudakov form factor

Again, no–emission probability!

$$\bar{P}(0 < t \leq T) = \exp\left(-\int_0^T \frac{\mathrm{d}P(t)}{\mathrm{d}t}\mathrm{d}t\right)$$

So,

$$\begin{aligned}\mathrm{d}P(\text{first emission at } T) &= \mathrm{d}P(T)\bar{P}(0 < t \leq T) \\ &= \mathrm{d}P(T)\exp\left(-\int_0^T \frac{\mathrm{d}P(t)}{\mathrm{d}t}\mathrm{d}t\right)\end{aligned}$$

That's what we need for our parton shower! Probability density for next emission at t:

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

Parton shower Monte Carlo

Probability density:

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

Conveniently, the probability distribution is $\Delta(t)$ itself.

Parton shower Monte Carlo

Probability density:

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

Conveniently, the probability distribution is $\Delta(t)$ itself.
Hence, parton shower very roughly from (HERWIG):

1. Choose flat random number $0 \le \rho \le 1$.
2. If $\rho < \Delta(t_{\max})$: no resolbable emission, stop this branch.
3. Else solve $\rho = \Delta(t_{\max})/\Delta(t)$
   (= no emission between $t_{\max}$ and t) for t.
   Reset $t_{\max} = t$ and goto 1.

Determine z essentially according to integrand in front of exp.

Parton shower Monte Carlo

Probability density:

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

Conveniently, the probability distribution is $\Delta(t)$ itself.

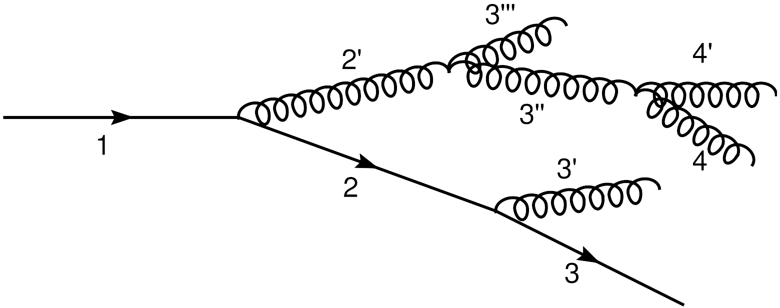
- That was old HERWIG variant. Relies on (numerical) integration/tabulation for $\Delta(t)$.
- Pythia, now also Herwig++, use the Veto Algorithm.
- Method to sample x from distribution of the type

$$\mathrm{d}P = F(x)\exp\left[-\int^{x}\mathrm{d}x'\,F(x')\right]\mathrm{d}x\;.$$

Simpler, more flexible, but slightly slower.

Parton cascade

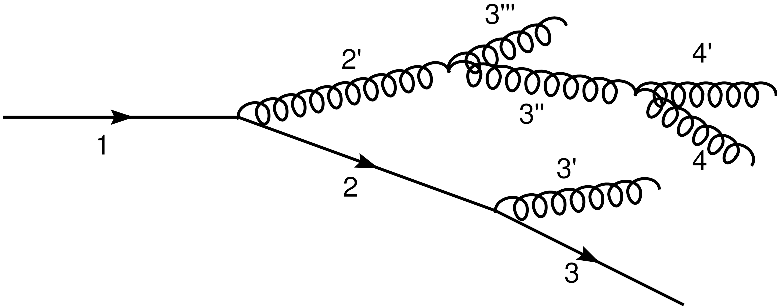
Get tree structure, ordered in evolution variable t:



Here: $t_1 > t_2 > t_3$; $t_2 > t_{3'}$ etc.
Construct four momenta from (t_i, z_i) and (random) azimuth ϕ.

Parton cascade

Get tree structure, ordered in evolution variable t:

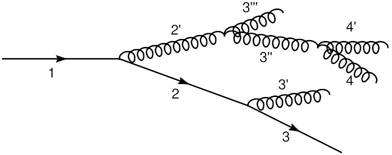


Here: $t_1 > t_2 > t_3$; $t_2 > t_{3'}$ etc.
Construct four momenta from (t_i, z_i) and (random) azimuth ϕ.

Not at all unique!
Many (more or less clever) choices still to be made.

Parton cascade

Get tree structure, ordered in evolution variable t:



- t can be θ, Q^2, $p_\perp$, ...
- Choice of hard scale t_{max} not fixed. "Some hard scale".
- z can be light cone momentum fraction, energy fraction, ...
- Available parton shower phase space.
- Integration limits.
- Regularisation of soft singularities.
- ...

Good choices needed here to describe wealth of data!

Soft emissions

- Only *collinear* emissions so far.
- Including *collinear+soft*.
- *Large angle+soft* also important.

Soft emissions

- Only *collinear* emissions so far.
- Including *collinear+soft*.
- *Large angle+soft* also important.

Soft emission: consider *eikonal factors*,
here for $q(p+q) \to q(p)g(q)$, soft g:

$$u(p)\ \not{\varepsilon} \frac{\not{p} + \not{q} + m}{(p+q)^2 - m^2} \longrightarrow u(p) \frac{p \cdot \varepsilon}{p \cdot q}$$

soft factorisation. Universal, *i.e.* independent of emitter.
In general:

$$d\sigma_{n+1} = d\sigma_n \frac{d\omega}{\omega} \frac{d\Omega}{2\pi} \frac{\alpha_S}{2\pi} \sum_{ij} C_{ij} W_{ij} \quad \text{(“QCD–Antenna”)}$$

with

$$W_{ij} = \frac{1 - \cos\theta_{ij}}{(1 - \cos\theta_{iq})(1 - \cos\theta_{qj})} \ .$$

Soft emissions

We define

$$W_{ij} = \frac{1 - \cos\theta_{ij}}{(1 - \cos\theta_{iq})(1 - \cos\theta_{qj})} \equiv W_{ij}^{(i)} + W_{ij}^{(j)}$$

with

$$W_{ij}^{(i)} = \frac{1}{2}\left(W_{ij} + \frac{1}{1 - \cos\theta_{iq}} - \frac{1}{1 - \cos\theta_{qj}} \right) .$$

$W_{ij}^{(i)}$ is only collinear divergent if $q \| i$ etc .

Soft emissions

We define

$$W_{ij} = \frac{1-\cos\theta_{ij}}{(1-\cos\theta_{iq})(1-\cos\theta_{qj})} \equiv W_{ij}^{(i)} + W_{ij}^{(j)}$$

with

$$W_{ij}^{(i)} = \frac{1}{2}\left(W_{ij} + \frac{1}{1-\cos\theta_{iq}} - \frac{1}{1-\cos\theta_{qj}}\right) .$$

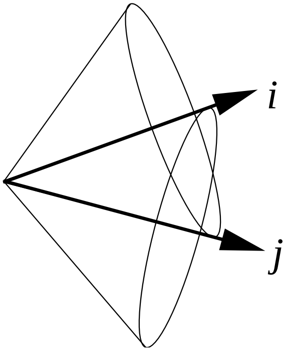
$W_{ij}^{(i)}$ is only collinear divergent if $q \| i$ etc .
After integrating out the azimuthal angles, we find

$$\int \frac{d\phi_{iq}}{2\pi} W_{ij}^{(i)} = \begin{cases} \dfrac{1}{1-\cos\theta_{iq}} & (\theta_{iq} < \theta_{ij}) \\ 0 & \text{otherwise} \end{cases}$$

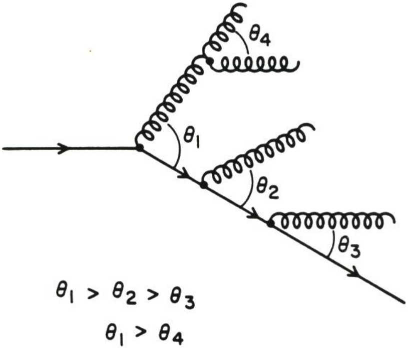
That's angular ordering.

Angular ordering

Radiation from parton *i* is bound to a cone, given by the colour partner parton *j*.



Results in angular ordered parton shower and suppresses soft gluons viz. hadrons in a jet.



Colour coherence from CDF

Events with 2 hard (> 100 GeV) jets and a soft 3rd jet (∼ 10 GeV)

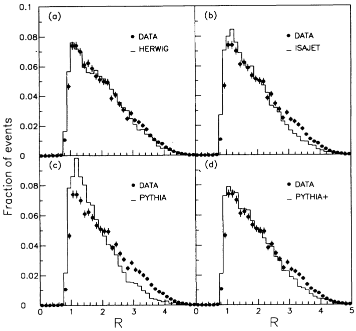


FIG. 14. Observed R distribution compared to the predictions of (a) HERWIG; (b) ISAJET; (c) PYTHIA; (d) PYTHIA+.

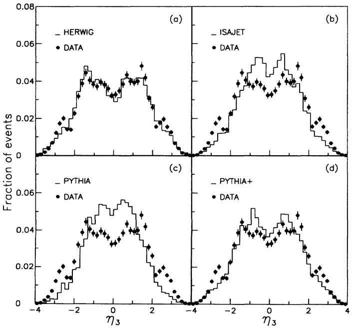


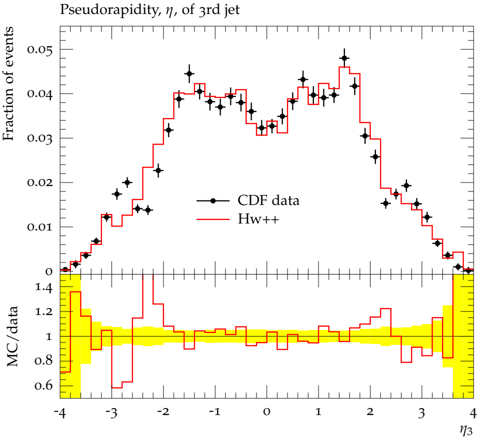
FIG. 13. Observed η_3 distribution compared to the predictions of (a) HERWIG; (b) ISAJET; (c) PYTHIA; (d) PYTHIA+.

F. Abe *et al.* [CDF Collaboration], Phys. Rev. D **50** (1994) 5562.

Best description with angular ordering.

Colour coherence from CDF

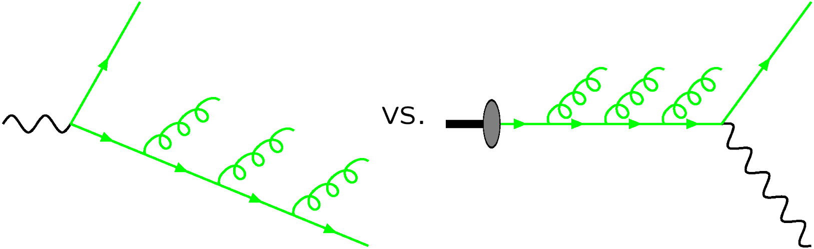
Events with 2 hard (> 100 GeV) jets and a soft 3rd jet (~ 10 GeV)



F. Abe *et al*. [CDF Collaboration], Phys. Rev. D **50** (1994) 5562.

Best description with angular ordering.

Initial state radiation



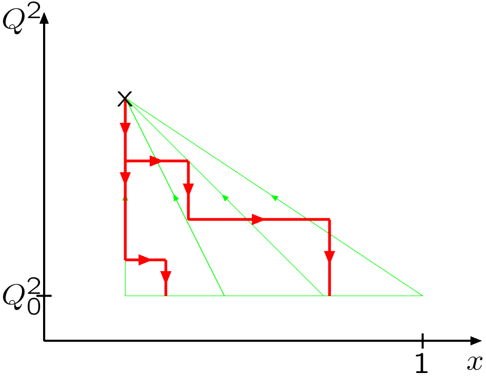
Similar to final state radiation. Sudakov form factor ($x' = x/z$)

$$\Delta(t, t_{\max}) = \exp\left[-\sum_b \int_t^{t_{\max}} \frac{\mathrm{d}t}{t} \int_{z_-}^{z_+} \mathrm{d}z \frac{\alpha_S(z,t)}{2\pi} \frac{x' f_b(x',t)}{x f_a(x,t)} \hat{P}_{ba}(z,t) \right]$$

Have to divide out the pdfs.

Initial state radiation

Evolve backwards from hard scale Q^2 *down* towards cutoff scale Q_0^2. Thereby increase x.

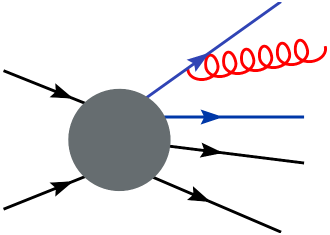


With parton shower we *undo* the DGLAP evolution of the pdfs.

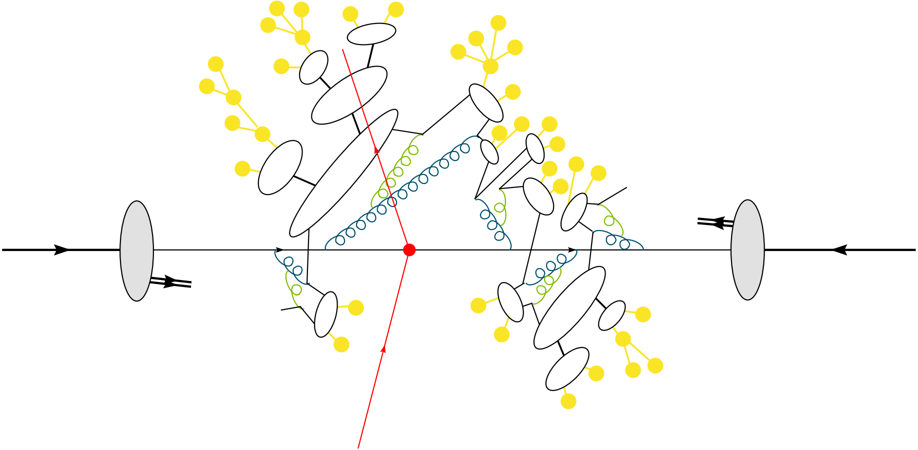
Dipoles

Exact kinematics when recoil is taken by spectator(s).

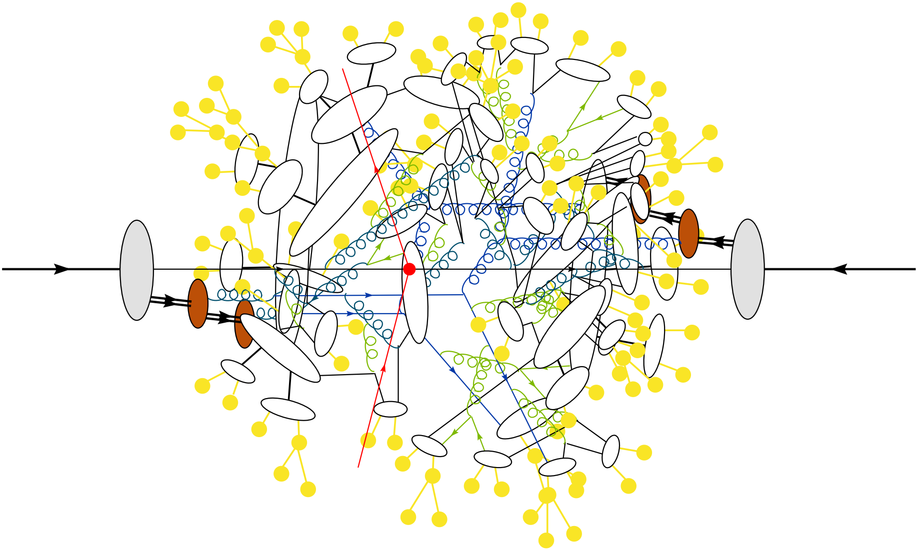
- Dipole showers.
- Ariadne.
- Recoils in Pythia.
- New dipole showers, based on
    - Catani Seymour dipoles.
    - QCD Antennae.
    - Herwig, Sherpa, Vincia, Dire, …
    - Goal: matching with NLO.
- Generalized to IS–IS, IS–FS.



Brief graphical summary



Brief graphical summary



A few plots

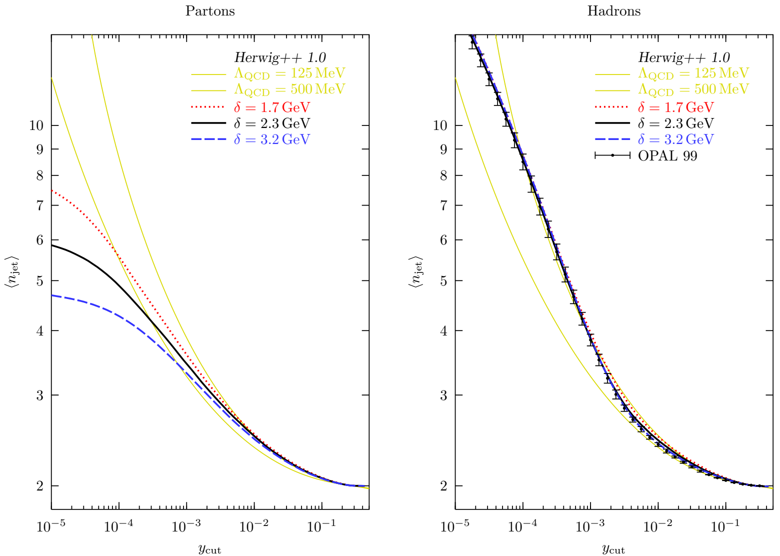
How well does it work?

- $e^+e^- \to$ hadrons, mostly at LEP.
- Jet shapes, jet rates, event shapes, identified particles…
- 'Tuning' of parameters.
- Use *all* analyses available in Rivet.
- Want to get *everything* right with *one* parameter set.
- Compare to literally $\approx$ 20000 plots.

- Check out `http://herwig.hepforge.org` ($\to$ Plots) for many more and comparisons with the latest release.

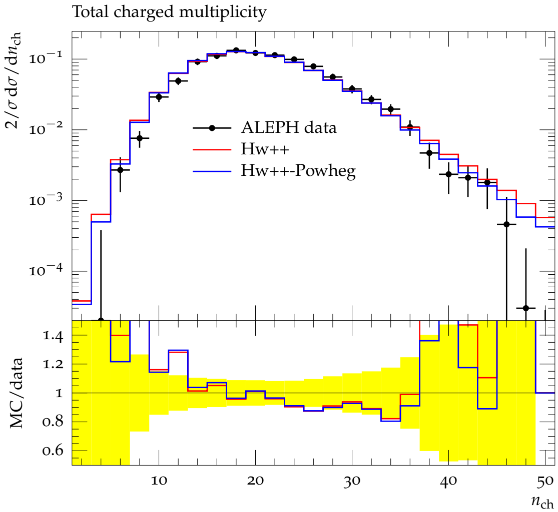
How well does it work?

Smooth interplay between shower and hadronization.



How well does it work?

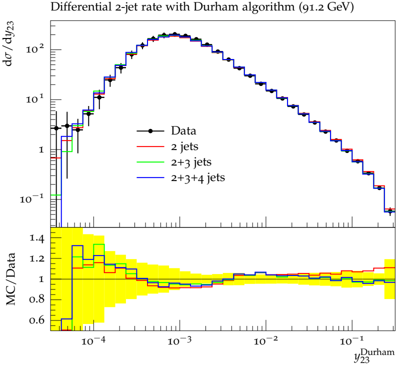
N_{ch} at LEP. Crucial for t_0 (Herwig++ 2.5.2)

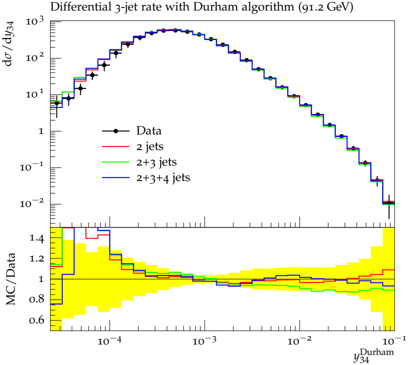


How well does it work?

How well does it work?

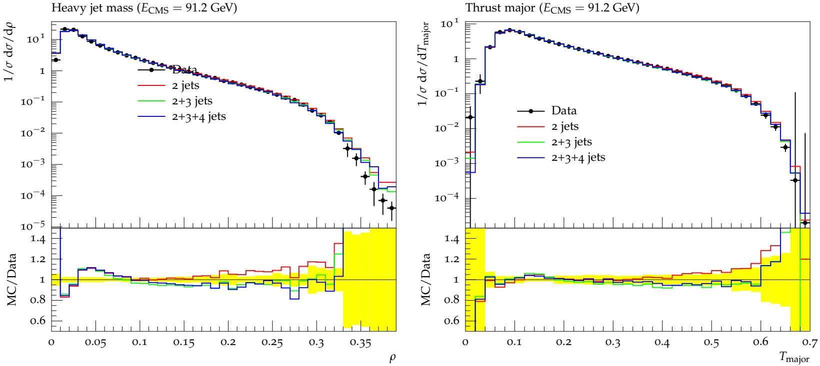
Differential Jet Rates at LEP (Herwig++ pre-3.0). Dipole shower + some merging





How well does it work?

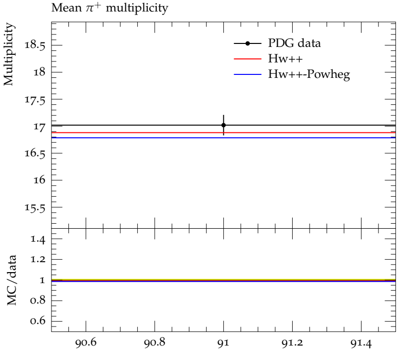
Event Shapes at LEP (Herwig++ pre-3.0).
Dipole shower + some merging

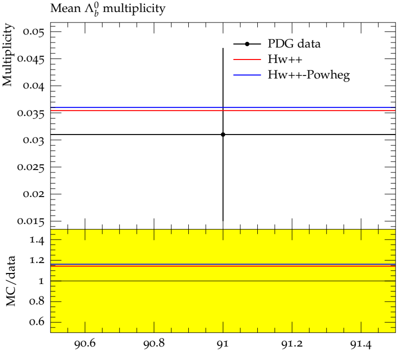


Parton showers do very well, today!

How well does it work?

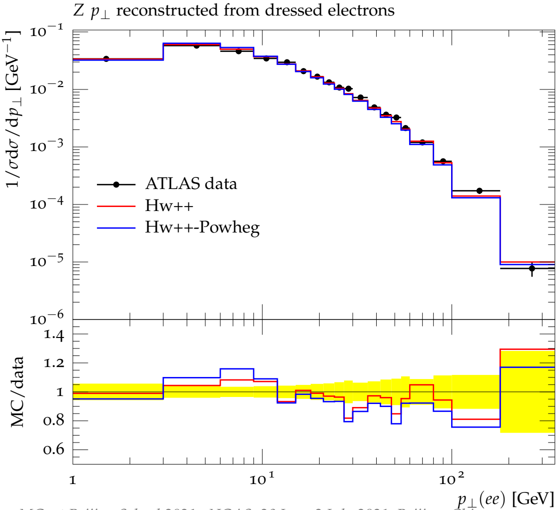
Hadron Multiplicities at LEP (e.g. π^+, Λ_b^0).



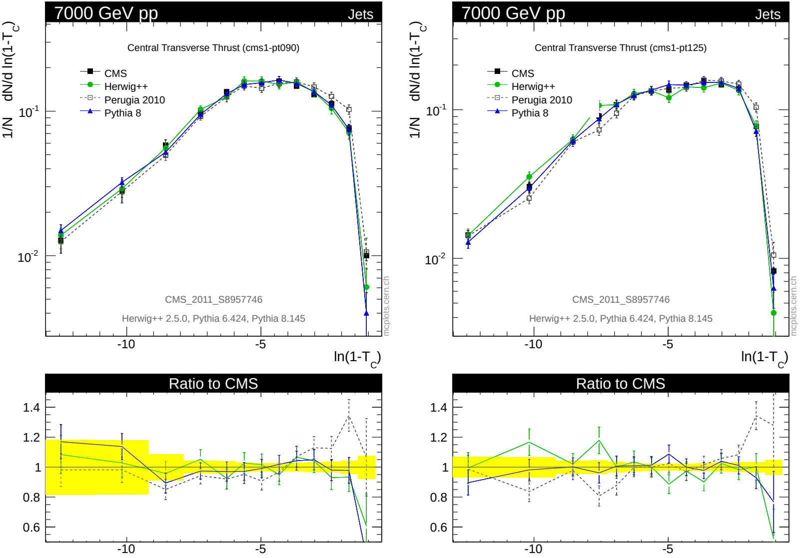


How well does it work?

$p_\perp(Z^0) \to$ intrinsic $k_\perp$ (LHC 7 TeV).
See also in context of matching/marging.

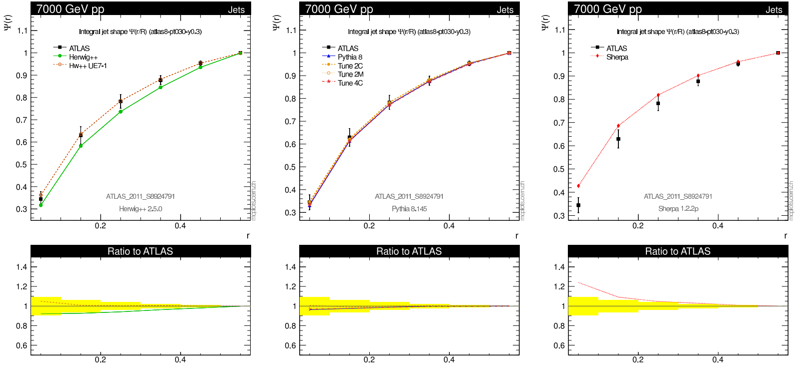


Transverse thrust



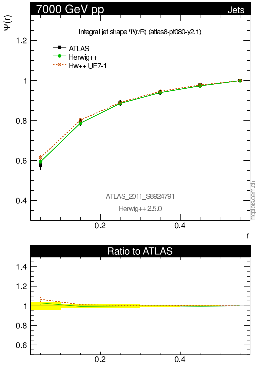
Integral jet shapes

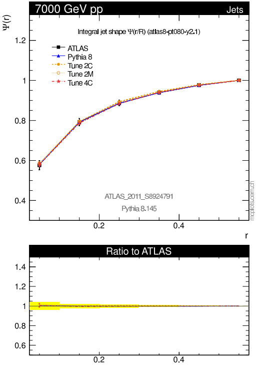
not too hard, central ($30 < p_T/\mathrm{GeV} < 40; 0 < |y| < 0.3$)

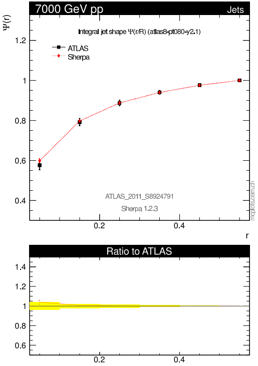


Integral jet shapes

harder, more forward ($80 < p_T/\text{GeV} < 110; 1.2 < |y| < 2.1$)

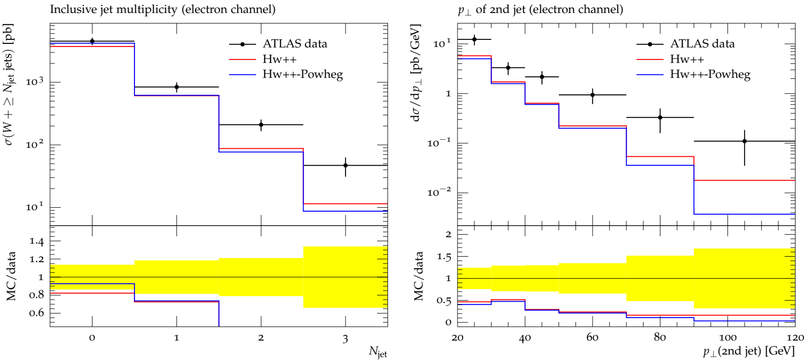






Limits of parton showers

$W + \text{jets}$, LHC 7 TeV.



Higher jets not covered by parton shower only → merging.