UNIVERSITÀ DEGLI STUDI DI UDINE
hic sunt futura

ATLAS EXPERIMENT

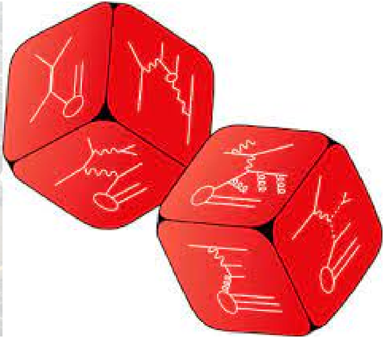
INFN

Third HEP Graduate Workshop

Physics Modeling

May 25, 2022

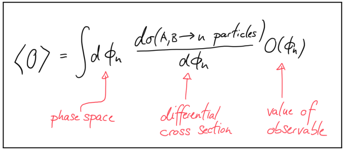
Giancarlo Panizzo Università degli Studi di Udine & INFN - Gruppo collegato di Udine



Observables and averages

An observation in particle physics is

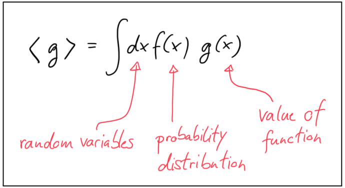
$$\langle O \rangle = \int d\phi_n \frac{d\sigma(A,B \to n \text{ particles})}{d\phi_n} O(\phi_n)$$



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

Compare to expectation value in statistics:

$$\langle g \rangle = \int dx\, f(x)\, g(x)$$



$\Rightarrow$ Calculate "theory predictions" for O with statistical methods.

Monte-Carlo vs analytic computations

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

Integration and averages

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

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

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

Ideally suited for our types of integrals

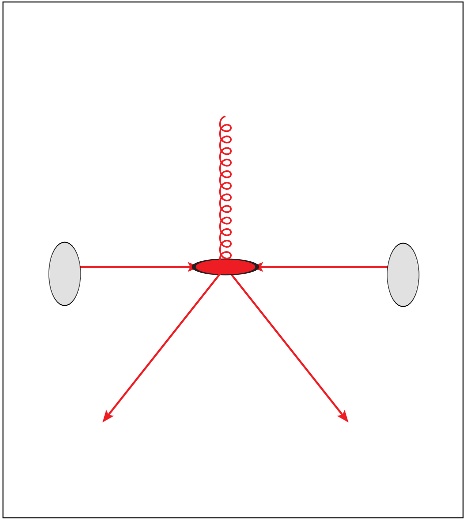
$$\langle O\rangle = \int d\Phi_n \frac{d\sigma_n}{d\Phi_n} O(\Phi_n) \propto \frac{1}{N}\sum_{i=1}^{N} \frac{d\sigma_n}{d\Phi_n}(\Phi_n^{(i)}) O(\Phi_n^{(i)})$$

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

NB: Les Houches Event Files are effectively that.

Collisions: a pictorial view

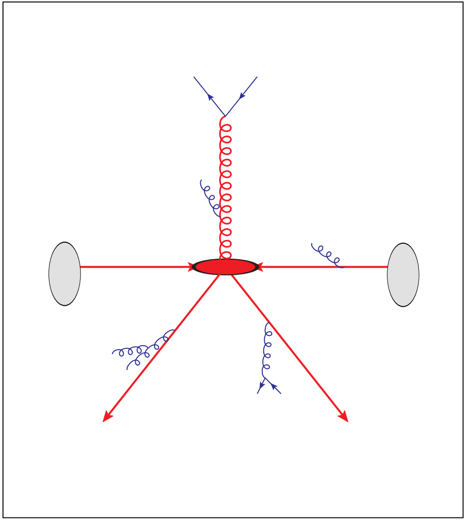
A high-energy scattering breaks the beams apart



Collisions: a pictorial view

A high-energy scattering breaks the beams apart

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

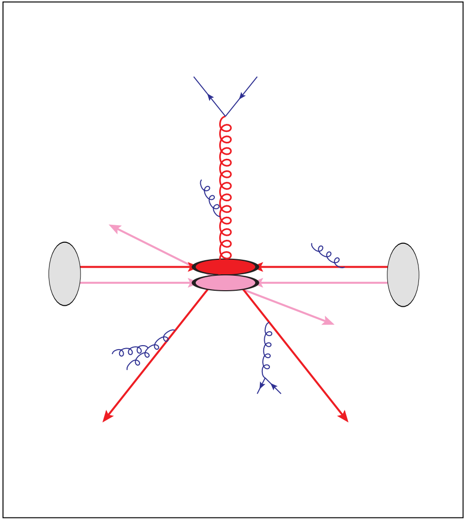


Collisions: a pictorial view

A high-energy scattering breaks the beams apart

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

Secondary interactions might occur at the same time



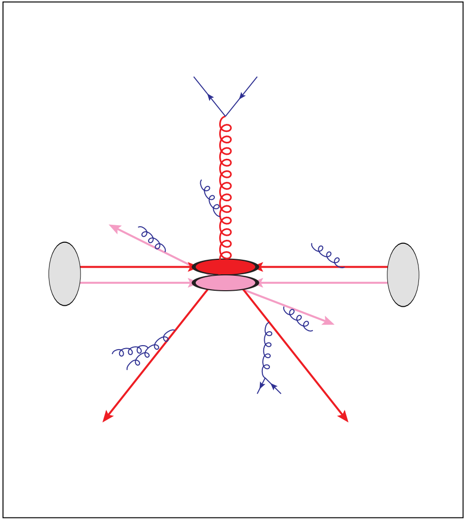
Collisions: a pictorial view

A high-energy scattering breaks the beams apart

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

Secondary interactions might occur at the same time

...and initiate further radiation "showers".



courtesy of S.Prestel

Collisions: a pictorial view

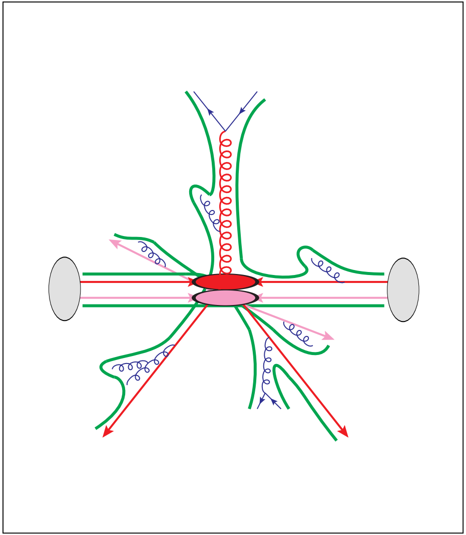
A high-energy scattering breaks the beams apart

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

Secondary interactions might occur at the same time

...and initiate further radiation "showers".

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



Collisions: a pictorial view

A high-energy scattering breaks the beams apart

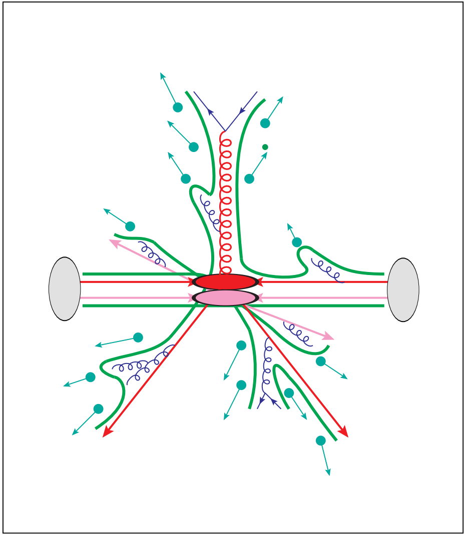
...which initiates a cascade of radiation in the vacuum.

Secondary interactions might occur at the same time

...and initiate further radiation "showers".

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

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



courtesy of S.Prestel

Collisions: a pictorial view

A high-energy scattering breaks the beams apart

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

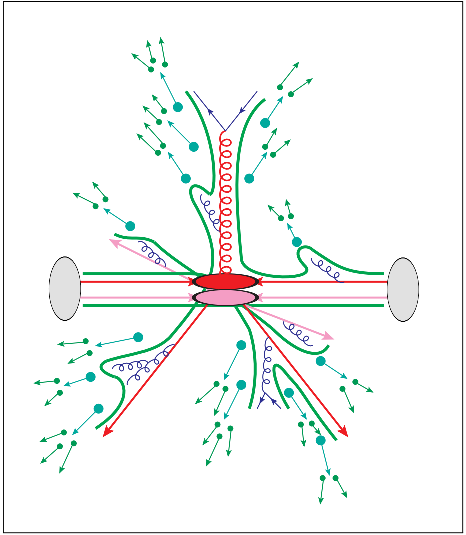
Secondary interactions might occur at the same time

...and initiate further radiation "showers".

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

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

...which decay into stable states



courtesy of S.Prestel

Collisions: a pictorial view

A high-energy scattering breaks the beams apart

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

Secondary interactions might occur at the same time

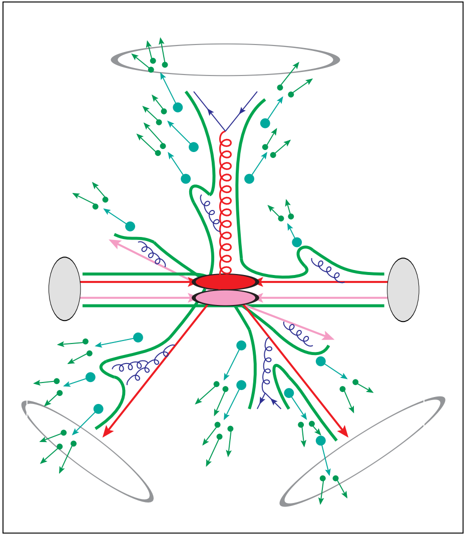
...and initiate further radiation "showers".

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

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

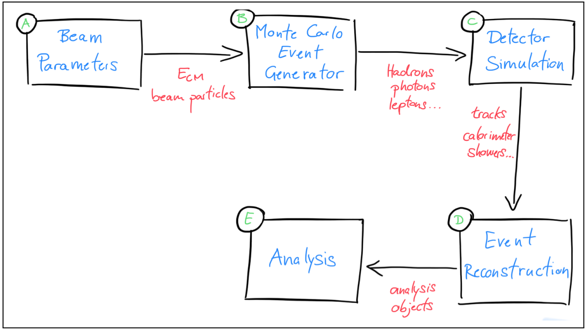
...which decay into stable states.

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



Divide et impera

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



A factorized at LHC, but not for neutrino experiments
C often factorized – but not for decays of long-lived particles

courtesy of S.Prestel

Monte-Carlo generators

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

Neutrino physics:	Cosmic rays:
Genie, GiBUU, NuWro, NEUT...	EPOS, QGSJET and SIBYLL

Heavy ions:	LHC physics:
HIJING, AMPT, JEWEL...	Herwig, Pythia, Sherpa Madgraph, Whizard, Alpgen...

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

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

courtesy of S.Prestel

Split into smaller problems

From a technical viewpoint, this chain of phenomena looks like

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

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

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

Basic numerical techniques

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

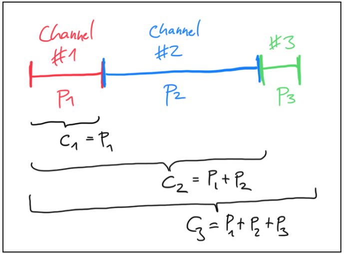
In the following, we'll now look at

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

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

Discrete transformation

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



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

channel #1 if $0 < RC_3 < C_1$

channel #2 if $C_1 < RC_3 < C_2$

channel #3 if $C_2 < RC_3 < C_3$

Repeat as often as you like.

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

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

Inversion sampling

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

The cumulative distribution becomes

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

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

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

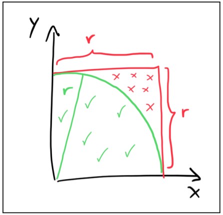
This is called inversion sampling.

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

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

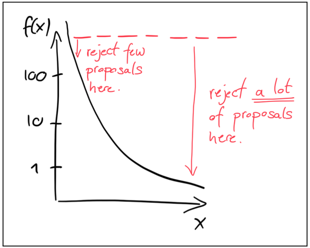
Rejection sampling

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



Example: Calculate π by random sampling:

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



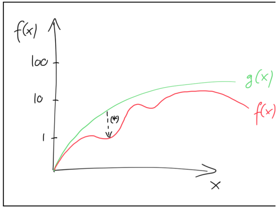
In practise, "uniform sampling" often not sufficient – efficiency very bad!

Combining the two

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

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

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



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

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

Comparison: Uniform sampling

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

error worse in regions of large variance...

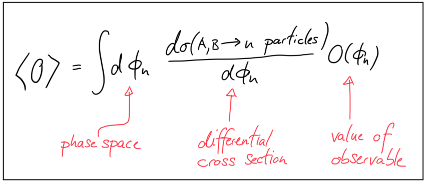
Importance sampling

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

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

Measurements

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



...so we have to worry about

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

Sampling factorization

When sampling phase space,

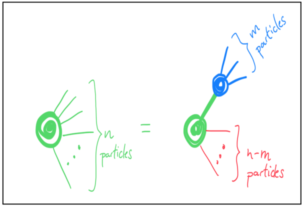
avoid large event weight fluctuations

avoid excessive rejection rate

⇒ Phase space generation separates enthusiasts from experts.

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

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



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

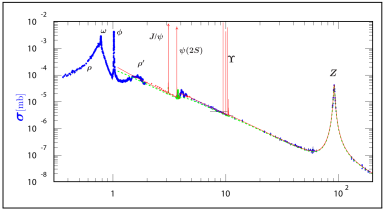
...we can continue until only simple integrations $(d\Phi_2, d\Phi_3)$ remain, and then find a clever parameterization for those.

Example

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

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

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



The cumulative function is

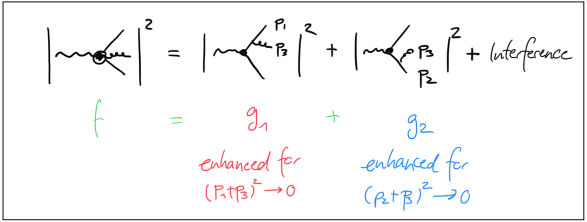
$$\begin{aligned} C(\hat{s}_{min},\hat{s}_{max}) &\propto I(\hat{s}_{max}) - I(\hat{s}_{min}) \\ &= \frac{1}{M_V\Gamma_V}\left[\mathrm{atan}\left(\frac{\hat{s}_{max}-M_V^2}{M_V\Gamma_V}\right) - \mathrm{atan}\left(\frac{\hat{s}_{min}-M_V^2}{M_V\Gamma_V}\right)\right] \end{aligned}$$

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

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

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

Multichannel sampling



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

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

NB: also heavily used in parton showers.

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

Vegas

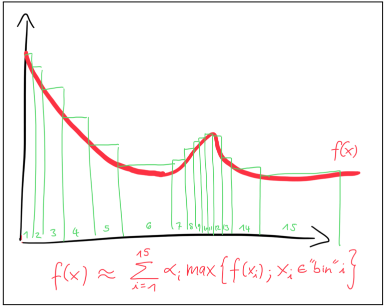
All of these methods require (analytical) knowledge of the differential cross section – which is often hard to come by.

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

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

This is the construction principle of VEGAS.

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



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

courtesy of S.Prestel

Hands on: session I

Worksheet link