Introduction to Monte Carlo Event Generation Lecture 1: Introduction to Monte Carlo Techniques

Peter Richardson

IPPP Durham

MCnet School: 5th August

Motivation

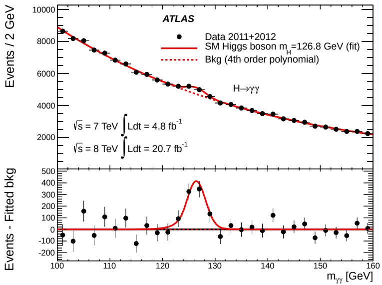
The aims of the LHC physics programme are:

- discovery, and now measurement of the properties, of the Higgs boson;
- the search for physics Beyond the Standard Model;
- the measurement of Standard Model (SM) processes at the highest energies.

All of these require accurate simulations.

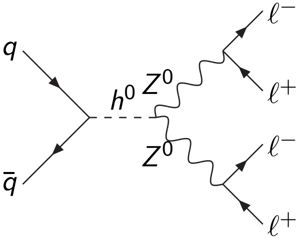
Higgs Boson

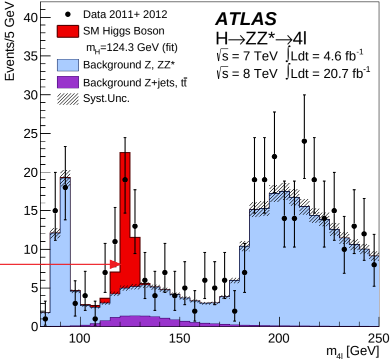
- In some searches the background can be extracted from data.
- However even for the simplest cases there is often a hidden dependence on simulation for the cuts and training of neutral nets and boosted decision trees.



Higgs Boson

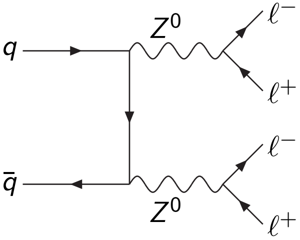
- In other cases we need very accurate simulations of complex final states to predict the background.

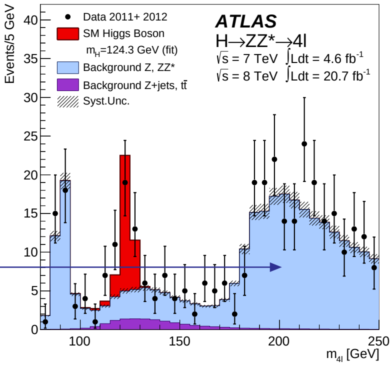




Higgs Boson

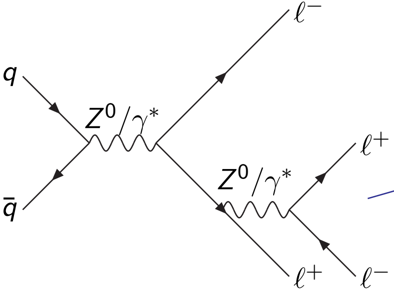
- In other cases we need very accurate simulations of complex final states to predict the background.

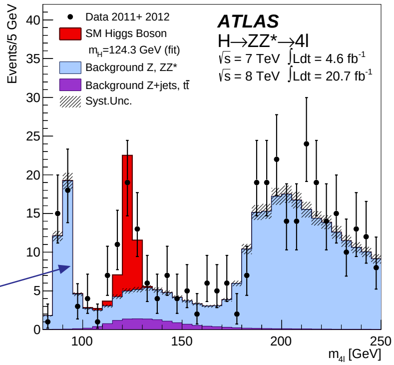




Higgs Boson

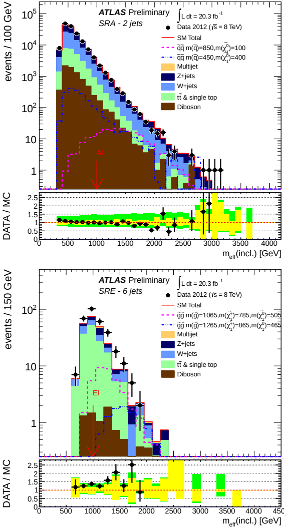
- In other cases we need very accurate simulations of complex final states to predict the background.





SUSY Searches

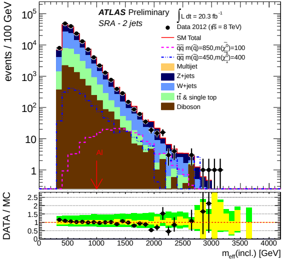
- Understanding the SM backgrounds is essential in any BSM search.
- Often try to use control regions to validate/normalize simulations.
- However MC simulations are an essential tool in these searches to predict the signal and background.

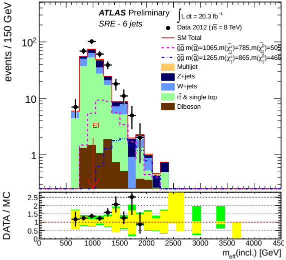


SUSY Searches

Use a wide range of simulations

- Z/γ^* and γ + jets SHERPA
- W + jets ALPGEN+HERWIG.
- $t\bar{t}$,MC@NLO+HERWIG.
- s-channel and Wt single top quark + jets MC@NLO+HERWIG
- t-channel single top quark + jets AcerMC+PYTHIA6
- $t\bar{t}$ + jets, W or Z MADGRAPH+PYTHIA6.
- WZ, ZZ and $Z\gamma$ SHERPA
- SUSY Herwig++ or MADGRAPH

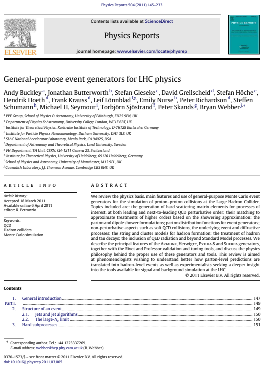




Overview

Resources

- There are a lot of lectures on Monte Carlo event generation from previous MCnet and other schools.
- Best single reference review produced by MCnet General-purpose event generators for LHC physics Buckley, et. al., Phys.Rept. 504 (2011) 145-233, arXiv:1101.2599


Physics Reports 504 (2011) 145–233

Contents lists available at ScienceDirect

Physics Reports

journal homepage: www.elsevier.com/locate/physrep

General-purpose event generators for LHC physics

Andy Buckley[a], Jonathan Butterworth[b], Stefan Gieseke[c], David Grellscheid[d], Stefan Höche[e], Hendrik Hoeth[d], Frank Krauss[d], Leif Lönnblad[f,g], Emily Nurse[b], Peter Richardson[d], Steffen Schumann[h], Michael H. Seymour[i], Torbjörn Sjöstrand[f], Peter Skands[g], Bryan Webber[j,*]

[a] PPE Group, School of Physics & Astronomy, University of Edinburgh, EH25 9PN, UK
[b] Department of Physics & Astronomy, University College London, WC1E 6BT, UK
[c] Institute for Theoretical Physics, Karlsruhe Institute of Technology, D-76128 Karlsruhe, Germany
[d] Institute for Particle Physics Phenomenology, Durham University, DH1 3LE, UK
[e] SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
[f] Department of Astronomy and Theoretical Physics, Lund University, Sweden
[g] PH Department, TH Unit, CERN, CH-1211 Geneva 23, Switzerland
[h] Institute for Theoretical Physics, University of Heidelberg, 69120 Heidelberg, Germany
[i] School of Physics and Astronomy, University of Manchester, M13 9PL, UK
[j] Cavendish Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, UK

ARTICLE INFO

Article history:
Accepted 18 March 2011
Available online 6 April 2011
editor: R. Petronzio

Keywords:
QCD
Hadron colliders
Monte Carlo simulation

ABSTRACT

We review the physics basis, main features and use of general-purpose Monte Carlo event generators for the simulation of proton–proton collisions at the Large Hadron Collider. Topics included are: the generation of hard scattering matrix elements for processes of interest, at both leading and next-to-leading QCD perturbative order; their matching to approximate treatments of higher orders based on the showering approximation; the parton and dipole shower formulations; parton distribution functions for event generators; non-perturbative aspects such as soft QCD collisions, the underlying event and diffractive processes; the string and cluster models for hadron formation; the treatment of hadron and tau decays; the inclusion of QED radiation and beyond Standard Model processes. We describe the principal features of the Ariadne, Herwig++, Pythia 8 and Sherpa generators, together with the Rivet and Professor validation and tuning tools, and discuss the physics philosophy behind the proper use of these generators and tools. This review is aimed at phenomenologists wishing to understand better how parton-level predictions are translated into hadron-level events as well as experimentalists seeking a deeper insight into the tools available for signal and background simulation at the LHC.

Contents

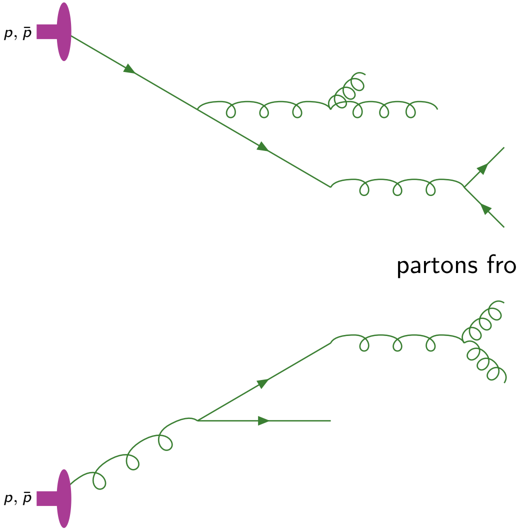
* Corresponding author. Tel.: +44 1223337269.
E-mail address: webber@hep.phy.cam.ac.uk (B. Webber).

0370-1573/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.physrep.2011.03.005

Evolution of an event

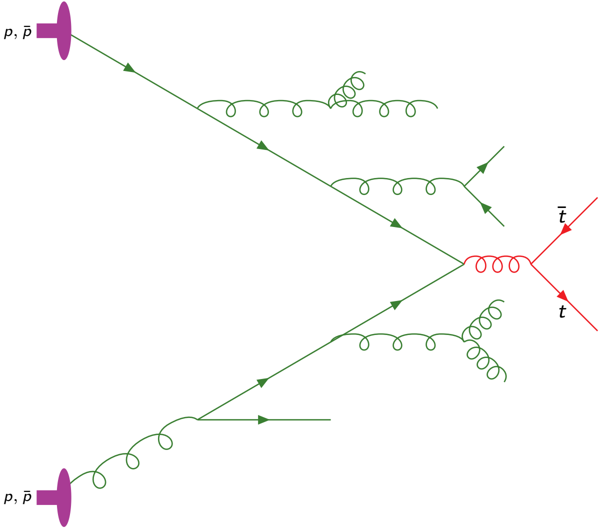
$t = -\infty$, incoming protons

Evolution of an event



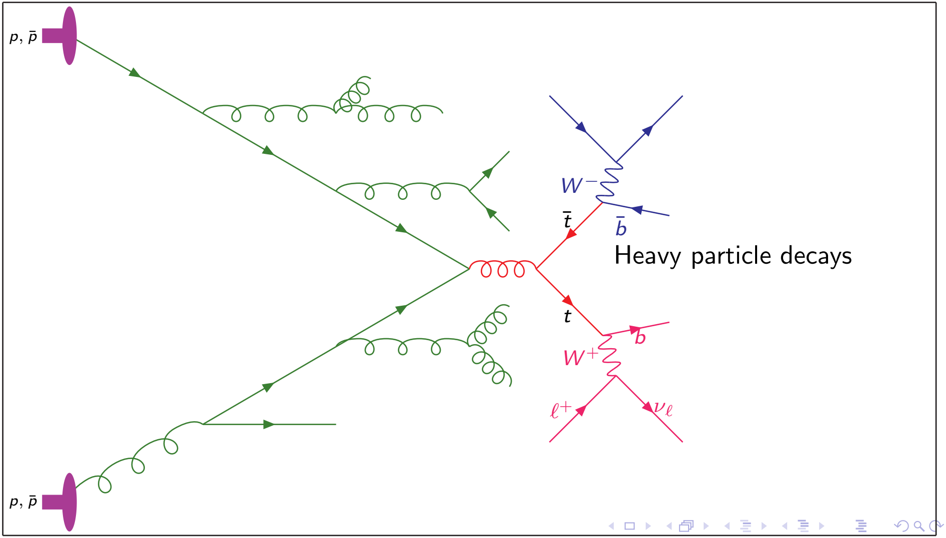
partons from the protons radiate

Evolution of an event

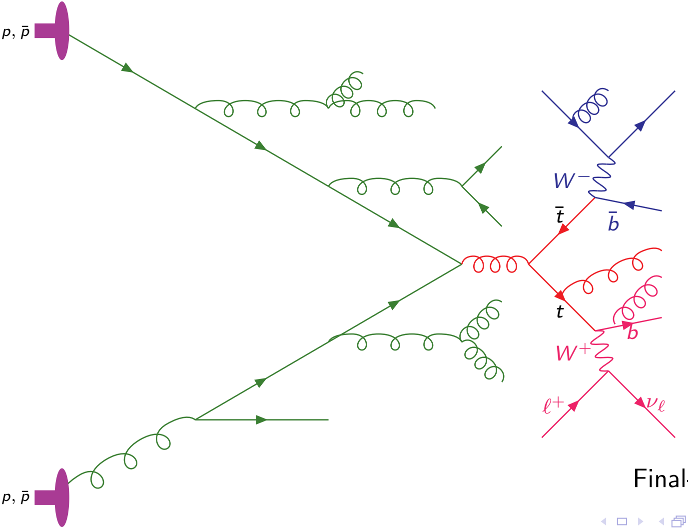


partons collide in fundamental hard process

Evolution of an event

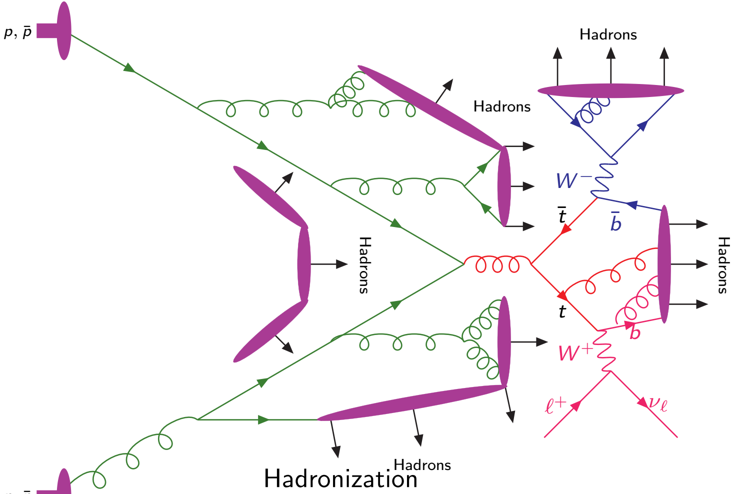


Evolution of an event



Final-state radiation

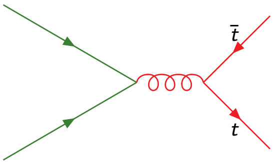
Evolution of an event



Simulation

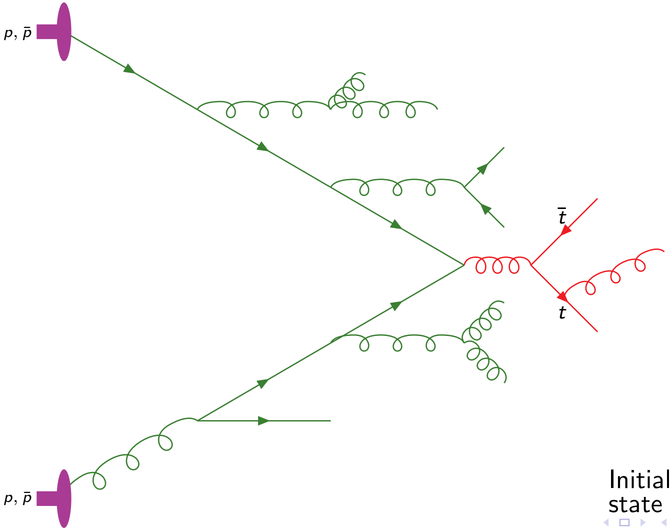
- There are a lot of different physical processes involved.
- Some we understand and can calculate from first principles.
- Some we can approximately calculate.
- For others we have to rely and phenomenological models.
- We are helped by being able to separate, at some level of approximation, different physics happening on different time/length/energy scales.
- Simulate different pieces separately, together with evolution between the different scales.

A Monte Carlo Event



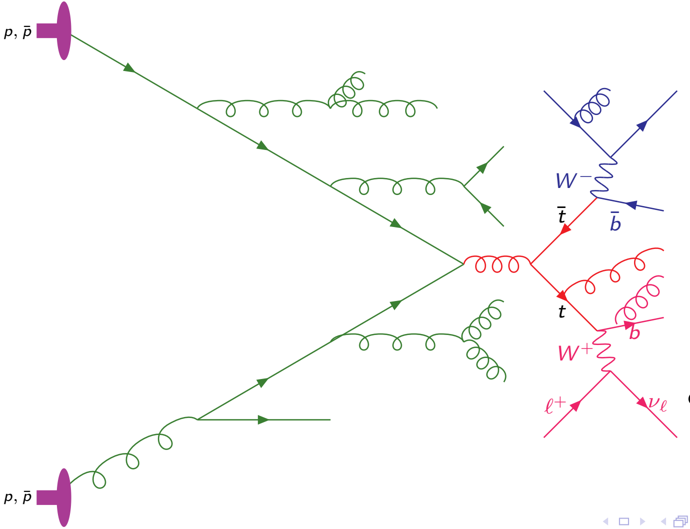
Hard Process, usually calculated at leading order

A Monte Carlo Event



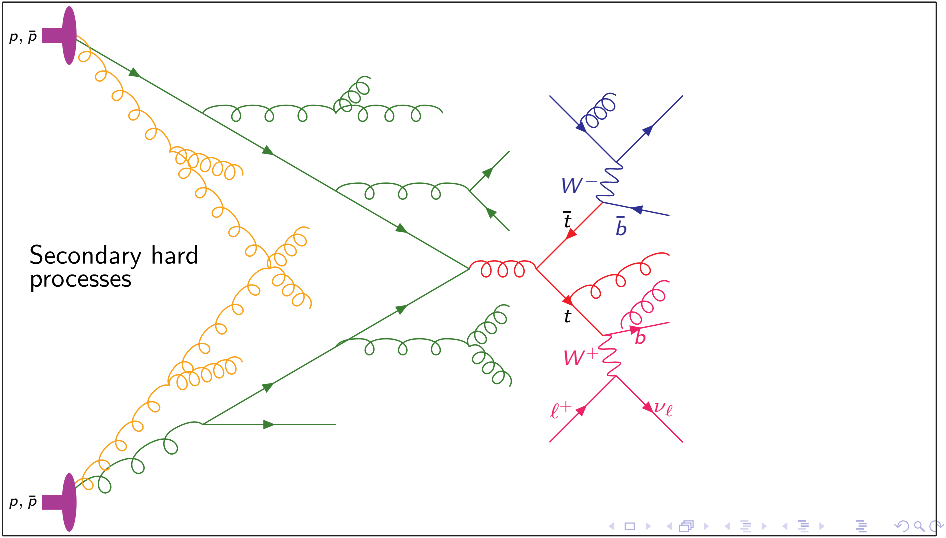
Initial- and final-state parton shower

A Monte Carlo Event

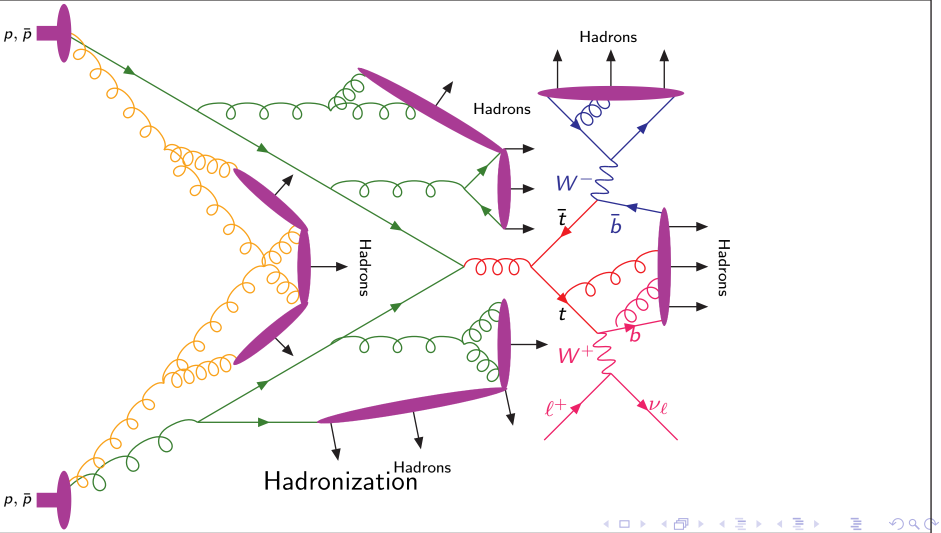


Perturbative decays of heavy particles

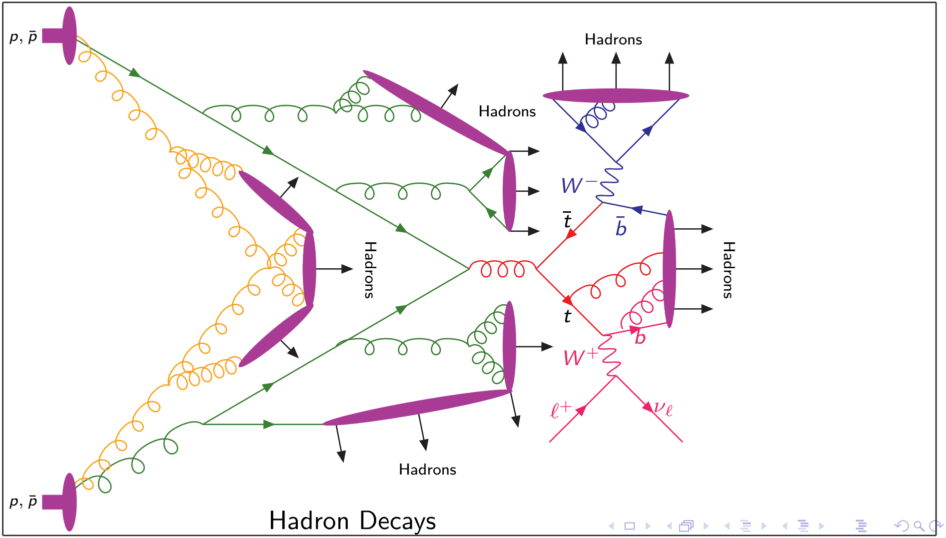
A Monte Carlo Event



A Monte Carlo Event



A Monte Carlo Event



Parton-Level event generation

- We want calculate the expectation value of an observable, $\mathcal{O}$, which is a function of the momenta of the n final-state particles.
- At the parton-level this is given by

$$\langle \mathcal{O} \rangle = \int \frac{(2\pi)^4}{2\hat{s}} \prod_{i=1}^{n} \frac{\mathrm{d}^3 p_i}{(2\pi)^3 2E_i} |\mathcal{M}(\{p_i\})|^2 \mathcal{O}(\{p_i\}).$$

- In hadronic collisions we also have to convolve with the parton distribution functions.
- There are two issues:
  1. calculating the matrix element for a given phase-space point;
  2. integrating over the phase space.

Monte Carlo integration

In general the functions aren't integrable analytically, so how do we numerical integrate an arbitrary function

$$I = \int_\Omega \prod_{i=1}^{n} \mathrm{d}x_i f(\{x_i\}),$$

where x_i are the integration variable and Ω are the limits? Problems are that discretization techniques are extremely inefficient for:

- large n, trapezium rules converges $\propto \frac{1}{n^{\frac{2}{d}}}$, Simpson's rule converges $\propto \frac{1}{n^{\frac{4}{d}}}$;
- complicated limits;
- integrands which have peaks and divergences.

All of of which occur in particle physics!

Monte Carlo Integration

- Suppose we want to evaluate

$$I = \int_{x_1}^{x_2} f(x)\mathrm{d}x.$$

  This can be written as an average

$$I = \int_{x_1}^{x_2} f(x)\mathrm{d}x = (x_2 - x_1)\langle f(x)\rangle.$$

- The average can be calculated by selecting N values randomly from a uniform distribution

$$I \approx I_N \equiv (x_2 - x_1)\frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

- Often we define a weight, $w_i = (x_2 - x_1)f(x_i)$ in which case the integral is the average of the weight.

Monte Carlo Integration

- The error on the integral can be found using the central limit theorem

$$I \approx I_N \pm \sqrt{\frac{V_N}{N}},$$

where

$$I_N = \frac{1}{N}\sum_{i=1}^{N} w_i \qquad V_N = \frac{1}{N}\sum_{i=1}^{N} w_i^2 - \left[\frac{1}{N}\sum_{i=1}^{N} w_i\right]^2$$

Generation according to a distribution

- Suppose we want to select values of x at random according to $f(x)$.
- Easy provided the function is integrable and invertible, *i.e.* we can calculate
$$F(x) = \int \mathrm{d}x f(x),$$
and its inverse $F^{-1}(x)$.
- In this case we can generate x according to $F(x)$ between $x_{\rm min}$ and $x_{\rm max}$ using
$$x = F^{-1}\left[F(x_{\rm min}) + \mathcal{R}\left(F(x_{\rm max}) - F(x_{\rm min})\right)\right]$$
- Usually impossible to find $F(x)$ or $F^{-1}(x)$.

Generation according to a distribution

- Consider the example of the Breit-Wigner

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

- Using the substitution

$$m^2 = M^2 + M\Gamma \tan\rho \qquad \Rightarrow \qquad \mathrm{d}m^2 = M\Gamma \sec^2\rho \mathrm{d}\rho$$

then

$$F(m^2) = \int \mathrm{d}m^2 f(m^2) = \int \mathrm{d}\rho \frac{M^2\Gamma^2 \sec^2\rho}{M^2\Gamma^2 \tan^2\rho + M^2\Gamma^2} = \int \mathrm{d}\rho = \rho$$

Therefore

$$F(m^2) = \tan^{-1}\left[\frac{m^2 - M^2}{M\Gamma}\right]$$

Generation according to a distribution

- The inverse

$$\rho = F(m^2) = \tan^{-1}\left[\frac{m^2 - M^2}{M\Gamma}\right] \quad \Rightarrow m^2 = M^2 + M\Gamma\tan(\rho)$$

- Hence

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

- Therefore generating according to the Breit-Wigner

$$m^2 = M^2 + M\Gamma\tan\left[\tan^{-1}\left[\frac{m^2_{\rm min} - M^2}{M\Gamma}\right] + \mathcal{R}\left(\tan^{-1}\left[\frac{m^2_{\rm max} - M^2}{M\Gamma}\right] - \tan^{-1}\left[\frac{m^2_{\rm min} - M^2}{M\Gamma}\right]\right)\right]$$

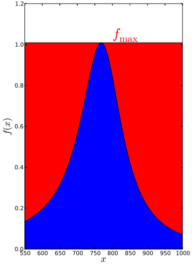
- But if we don't know $F(x)$ and $F^{-1}(x)$.

Unweighting

- Provided that we know the maximum value of the function, $f_{\rm max}$, we can also generate x according to $f(x)$.
- Randomly generate values of x in the integration region and keep them with probability

$$P = \frac{f(x)}{f_{\rm max}} \geq \mathcal{R}.$$

- Easy to implement by generating a random number between 0 and 1 and keeping the value of x if the random number is less than the probability, called unweighting.



Monte Carlo Integration

The Monte Carlo technique has a number of important advantages:

- always converges as $1/\sqrt{N}$ regardless of the number of dimensions;
- arbitrarily complex integration regions, simply use a hypercube and set the integrand to zero outside Ω;
- easy estimate of the error;
- calculation of all observables at once.

In a typical LHC event we have ~ 1000 particles so we need to do ~ 3000 phase-space integrals for the momenta. Monte Carlo integration is the only viable option.

Improving convergence

- Convergence of the integral can be improved by reducing, V_N.
- Called Importance Sampling.
- Perform a Jacobian transform so that the integral is flat in the new integration variable.
- Consider the example of a fixed width Breit-Wigner distribution

$$I = \int_{M^2_{\rm min}}^{M^2_{\rm max}} {\rm d}m^2 \frac{1}{(m^2 - M^2) + M^2\Gamma^2}$$

where M is the physical mass of the particle, m is the off-shell mass and Γ is the width.

Improving convergence

- A useful transformation is

$$m^2 = M^2 + M\Gamma \tan\rho \qquad \Rightarrow \qquad \mathrm{d}m^2 = M\Gamma \sec^2\rho \mathrm{d}\rho$$

which gives

$$I = \int_{M^2_{\rm min}}^{M^2_{\rm max}} \mathrm{d}m^2 \frac{1}{(m^2 - M^2) + M^2\Gamma^2} = \int_{\rho_{\rm min}}^{\rho_{\rm max}} \mathrm{d}\rho \frac{M\Gamma \sec^2\rho}{M^2\Gamma^2 \tan^2\rho + M^2\Gamma^2}$$

- So we have in fact reduced the error to zero.

$$I = \frac{1}{M\Gamma}(\rho_{\rm max} - \rho_{\rm min})$$

Improving convergence

- In practice few of the cases we need to deal with in real examples can be exactly integrated.
- In these cases we try and pick a function that approximates the behaviour of the function we want to integrate.
- For example suppose we have a spin-1 meson decaying to two scalar mesons which are much lighter, consider the example of the ρ decaying to massless pions.
- In this case the width

$$\Gamma(m) = \frac{\Gamma_0 M}{m}\left(\frac{p(m)}{p(M)}\right)^3 = \frac{\Gamma_0 M}{m}\left(\frac{m}{M}\right)^{\frac{3}{2}} = \Gamma_0\sqrt{\frac{m}{M}},$$

where $p(m)$ is the 3-momentum of the decay products in the ρ rest frame.

Improving convergence

- If we were just to generate flat in m^2 then the weight would be

$$w_i = \frac{M^2_{\rm max} - M^2_{\rm min}}{(m^2 - M^2)^2 + \frac{\Gamma_0^2 m^3}{M}}$$

- If we perform a Jacobian transformation the integral becomes

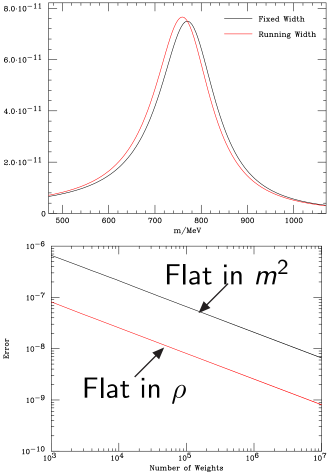
$$I = \int_{M^2_{\rm min}}^{M^2_{\rm max}} {\rm d}m^2 \frac{1}{(m^2 - M^2)^2 + \frac{\Gamma_0^2 m^3}{M}} = \frac{1}{M\Gamma_0} \int_{\rho_{\rm min}}^{\rho_{\rm max}} {\rm d}\rho \frac{(m^2 - M^2)^2 + M^2\Gamma_0^2}{(m^2 - M^2)^2 + \frac{\Gamma_0^2 m^3}{M}}$$

and the weight is

$$w_i = \frac{1}{M\Gamma_0} \left(\rho_{\rm max} - \rho_{\rm min}\right) \frac{(m^2 - M^2)^2 + M^2\Gamma_0^2}{(m^2 - M^2)^2 + \frac{\Gamma_0^2 m^3}{M}}$$

Improving Convergence

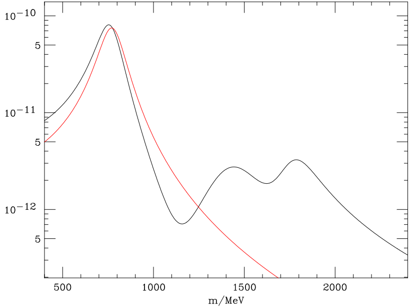
- If we perform the integral using m^2 the error is ~ 10 times larger for the same number of evaluations.
- *i.e.* Factor of 10 slower.



Improving Convergence

- Using a Jacobian transformation is always the best way of improving the convergence.
- There are automatic approaches (e.g. VEGAS) but they are never as good.

- Suppose instead of having one peak we have an integral with lots of peaks, say from the inclusion of excited ρ resonances in some process.
- Cant just use one Breit-Wigner. The error becomes large.



Multi-Channel approaches

- If we want to smooth out many peaks pick a function

$$f(m^2) = \sum_i \alpha_i g_i(m^2) = \sum_i \alpha_i \frac{1}{(m^2 - M_i^2)^2 + M_i^2 \Gamma_i^2}$$

where α_i is the weight for a given term such that $\sum_i \alpha_i = 1$.

- We can then rewrite the integral of a function

$$I = \int_{M^2_{\rm min}}^{M^2_{\rm max}} {\rm d}m^2 h(m^2) = \int_{M^2_{\rm min}}^{M^2_{\rm max}} {\rm d}m^2 h(m^2) \frac{f(m^2)}{f(m^2)}$$

$$= \int_{M^2_{\rm min}}^{M^2_{\rm max}} {\rm d}m^2 \sum_i \alpha_i g_i(m^2) \frac{h(m^2)}{f(m^2)} = \sum_i \alpha_i \int_{M^2_{\rm min}}^{M^2_{\rm max}} {\rm d}m^2 g_i(m^2) \frac{h(m^2)}{f(m^2)}$$

Multi-Channel approaches

- We can then perform a separate Jacobian transform for each of the integrals in the sum

$$I = \sum_i \alpha_i \int_{M^2_{\rm min}}^{M^2_{\rm max}} {\rm d}m^2 g_i(m^2) \frac{h(m^2)}{f(m^2)} = \sum_i \alpha_i \int_{rho_{\rm i,min}}^{\rho_{\rm i,max}} {\rm d}\rho_i \frac{h(m^2)}{f(m^2)}$$

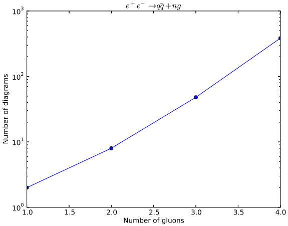
- Pick one of the integrals (channels) with probability α_i and calculate the weight as before.
- Called the Multi-Channel procedure and is used in the most sophisticated programs for integrating matrix elements in particle physics.
- There are methods to automatically optimise the choice of the channel weights, α_i.

Matrix Element Calculations

- The phase-space integration is only part of the problem of efficiently calculating observables.
- Efficient phase-space integration is usually the most important part of the problem.
- However the calculation of the matrix element is also important.

Factorial Growth

- The main issue for the evaluation of matrix elements is the factorial growth with the number of external particles.
- We need to evaluate $|\mathcal{M}|^2 = |\sum_{i=1}^{n} \mathcal{M}_i|^2$.



- Traditional squaring and and trace techniques grow like n^2.
- But, amplitudes are complex numbers, add them before squaring!

Helicity Amplitudes

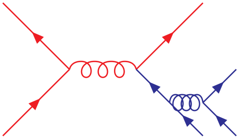
- As spinors and γ matrices have an explicit form they can be evaluated by (brute force) matrix multiplication (HELAS).
- Alternatively introduce basic helicity spinors and write everything as spinor products, *e.g.*

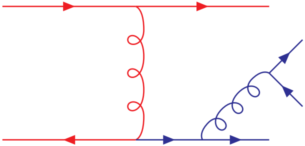
$$\bar{u}(p_1, h_1)u(p_2, h_2) = \text{complex number}$$

- Translate the Feynman diagrams into helicity amplitudes, complex-valued functions of momenta and helicities.
- Spin-correlations come essentially for free.

Recursion relations

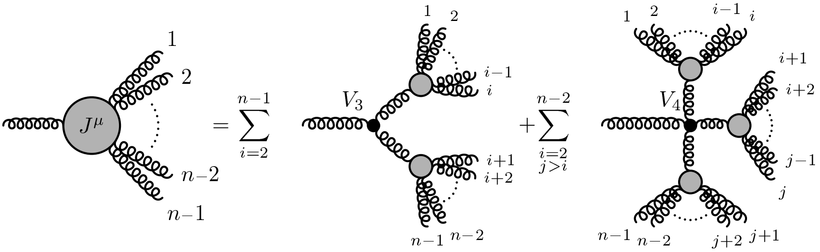
- Still have the factorial growth in the number of diagrams.
- In the helicity method
  - Reuse pieces: Only calculate them once,
  - Factoring out: reduce the number of multiplications





- Recursion relations with recycling built in are a better method
- Off-shell recursions Dyson-Schwinger, Berends-Giele, . . . best candidate so far.

Berends-Giele Recursion Relations



- In Berends-Giele relations the off-shell gluon current is recursively calculated.

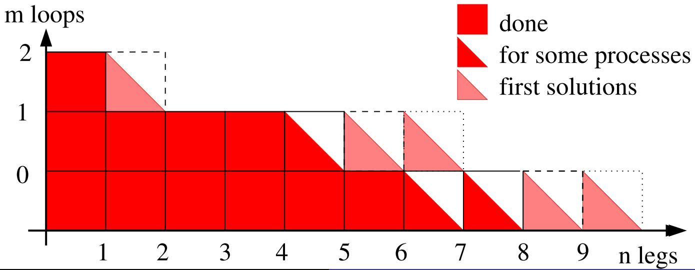
Colour Dressing

- Also a factorial growth from the colour algebra
- Sampling over colours helps
- Colour dressing F.Maltoni et. al. Rev. D67 (2003) 014026 improves things, particularly with Berends-Giele recursions C.Duhr et. al. JHEP 0608 (2006) 062

Final State	BG		BCF		CSW	
	CO	CD	CO	CD	CO	CD
$2g$	0.24	0.28	0.28	0.33	0.31	0.26
$3g$	0.45	0.48	0.42	0.51	0.57	0.55
$4g$	1.20	1.04	0.84	1.32	1.63	1.75
$5g$	3.78	2.69	2.59	7.26	5.95	5.96
$6g$	14.2	7.19	11.9	59.1	27.8	30.6
$7g$	58.5	23.7	73.6	646	146	195
$8g$	276	82.1	597	8690	919	1890
$9g$	1450	270	5900	127000	6310	29700
$10g$	7960	864	64000	-	48900	-

Current Status

- Calculation of higher order processes is more complicated.
- Tree-level is now fully automated, limits due to algorithms and computers.
- Automation of one-loop is ongoing with many new processes calculated.
- Only a few NNLO calculations

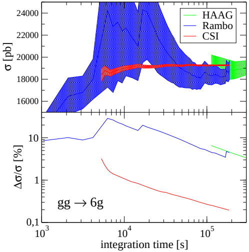


Parton-Level Tools

Program	$2 \to n$	Ampl.	Integ.	Public?	Lang.
ALPGEN	$n = 8$	rec.	Multi	yes	Fortran
AMEGIC++	$n = 6$	hel.	Multi	yes	C++
COMIX	$n = 8$	rec.	Multi	yes	C++
COMPHEP	$n = 4$	trace	1 Channel	yes	C
CALCHEP	$n = 4$	trace	1 Channel	yes	C
HELAC	$n = 8$	rec.	Multi	yes	Fortran
MADEVENT	$n = 6$	hel.	Multi	yes	Python/Fortran
WHIZARD	$n = 8$	rec.	Multi	yes	OCaml

Current Best Option

- Currently the best combination of phase-space and ME calculation, *i.e.* fastest and highest multiplicity COMIX
- Colour-dressed Berends-Giele amplitudes in the SM with fully recursive phase space generation.



σ [μb]	Number of jets						
$b\bar{b}$ + jets	0	1	2	3	4	5	6
Comix	471.2(5)	8.83(2)	1.813(8)	0.459(2)	0.150(1)	0.0531(5)	0.0205(4)
ALPGEN	470.6(6)	8.83(1)	1.822(9)	0.459(2)	0.150(2)	0.053(1)	0.0215(8)
AMEGIC	470.3(4)	8.84(2)	1.817(6)				

$gg \to ng$	Cross section [pb]				
n	8	9	10	11	12
$\sqrt{s}$ [GeV]	1500	2000	2500	3500	5000
Comix	0.755(3)	0.305(2)	0.101(7)	0.057(5)	0.026(1)
Maltoni(2002)	0.70(4)	0.30(2)	0.097(6)		
ALPGEN	0.719(19)				

Summary

- Monte Carlo sampling is a vital tool in particle physics for calculating observables.
- Modern phase-space sampling and matrix element calculation techniques allow ever higher multiplicity matrix elements to be calculated.
- However eventually we still have to use approximations and models to study LHC physics, as we will see in the rest of the lectures.