QCD & Monte Carlo Tools

Frank Krauss

Institute for Particle Physics Phenomenology
Durham University

CERN, 6.-15.6.2007

Topics of the lectures

1. Lecture 1: *The Monte Carlo Principle*
   - Monte Carlo as integration method
   - Hard physics simulation: Parton Level event generation
2. Lecture 2: *Dressing the Partons*
   - Hard physics simulation, cont'd: Parton Showers
3. Lecture 3: *Modelling beyond Perturbation Theory*
   - Soft physics simulation: Hadronization
   - Beyond factorization: Underlying Event
4. Lecture 4: *Higher Orders in Monte Carlos*
   - Some nomenclature: Anatomy of HO calculations
   - Merging vs. Matching

Thanks to

- the other Sherpas: T.Gleisberg, S.Höche, S.Schumann, F.Siegert, M.Schönherr, J.Winter;
- other MC authors: S.Gieseke, K.Hamilton, L.Lonnblad, F.Maltoni, M.Mangano, P.Richardson, M.Seymour, T.Sjostrand, B.Webber, ...

Simulation's paradigm

Basic strategy

Divide event into stages, separated by different scales.

- Signal/background:

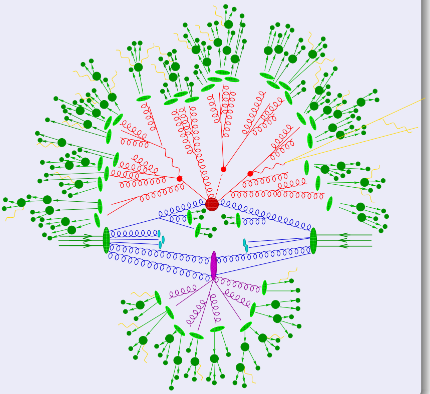
  Exact matrix elements.
- QCD-Bremsstrahlung:

  Parton showers (also in initial state).
- Multiple interactions:

  Beyond factorization: Modeling.
- Hadronization:

  Non-perturbative QCD: Modeling.

Sketch of an event



Outline of today's lecture

- Basics of MC:
  Integrals as averages
- MC integration in particle physics:
  Parton level event generation
- Limitations on the parton level
  (and attempts to overcome them)
- Survey of existing parton level tools
- Short introduction to NLO tools

Prelude: Selecting from a distribution

"Exact" case

- Usually random numbers # "flat" in [0, 1].
- But: Want random numbers x according to density $f(x)$.
- Solution:
  - Use integral F of $f(x)$ and its inverse F^{-1}
  - x given by $$\int\limits_{x_{\min}}^{x} \mathrm{d}x' f(x') = \# \int\limits_{x_{\min}}^{x_{\max}} \mathrm{d}x' f(x')$$

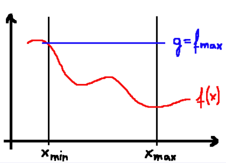
  leading to $$x = F^{-1}\left[F(x_{\min}) + \#\left(F(x_{\max}) - F(x_{\min})\right)\right].$$

Prelude: Selecting from a distribution

Hit-or-miss

What, if no exact case?

- Take an "over-estimator" $g(x)$: $g(x) > f(x) \;\; \forall x \in [x_{\rm min},\, x_{\rm max}]$, with G and G^{-1} known;
- select an x according to g;
- accept or reject with $f(x)/g(x)$;
- Obvious "guaranteed" $g(x)$: $g(x) = \mathrm{Max}\{f(x)\}$.



Monte Carlo integration

Convergence of numerical integration

- Consider $I = \int\limits_0^1 \mathrm{d}x^D f(\vec{x})$.
- Convergence behavior crucial for numerical evaluations. For integration (N = number of evaluations of f):
  - Trapezium rule $\simeq 1/N^{2/D}$
  - Simpson's rule $\simeq 1/N^{4/D}$
  - Central limit theorem $\simeq 1/\sqrt{N}$.
- Therefore: Use central limit theorem.

Monte Carlo integration

Monte Carlo integration

- Use random vectors $\vec{x}_i \longrightarrow$:
  Evaluate estimate of the integral $\langle I \rangle$ rather than I.
  $$\langle I(f) \rangle = \frac{1}{N} \sum_{i=1}^{N} f(\vec{x}_i).$$
  (This is the original meaning of Monte Carlo: Use random numbers for integration.)
- Quality of estimate given by error estimator (variance)
  $$\langle E(f) \rangle^2 = \frac{1}{N-1} \left[\langle I^2(f) \rangle - \langle I(f) \rangle^2 \right].$$
- Name of the game: Minimize $\langle E(f) \rangle$.
- Problem: Large fluctuations in integrand f
- Solution: Smart sampling methods

Monte Carlo integration

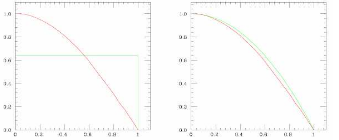
Importance sampling

Basic idea: Put more samples in regions, where f largest

$\Longrightarrow$ improves convergence behavior

(corresponds to a Jacobian transformation).

- Assume a function $g(\vec{x})$ similar to $f(\vec{x})$;
- obviously then, $f(\vec{x})/g(\vec{x})$ is comparably smooth, hence $\langle E(f/g)\rangle$ is small.



$$I = \int_0^1 dx\, \cos\tfrac{\pi}{2}x = 0.637 \pm 0.308/\sqrt{N}$$

$$I = \int_0^1 dx(1-x^2)\,\frac{\cos\frac{\pi}{2}x}{1-x^2} = \int d\rho\, \frac{\cos\frac{\pi}{2}x}{1-x^2}[x(\rho)] = 0.637 \pm 0.032/\sqrt{N}$$

Monte Carlo integration

Stratified sampling

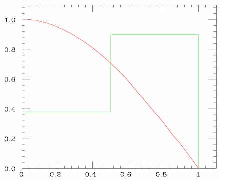
Basic idea: Decompose integral in M sub-integrals

$$\langle I(f)\rangle = \sum_{j=1}^{M}\langle I_j(f)\rangle, \quad \langle E(f)\rangle^2 = \sum_{j=1}^{M}\langle E_j(f)\rangle^2$$

Then: Overall variance smallest, if "equally distributed".

$\Longrightarrow$ Sample, where the fluctuations are.

- Divide interval in bins;
- adjust bin-size or weight per bin such that variance identical in all bins.



$\langle I\rangle = 0.637 \pm 0.147/\sqrt{N}$

Monte Carlo integration

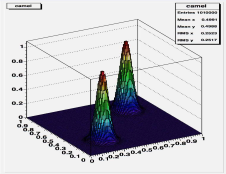
Example for stratified sampling: VEGAS

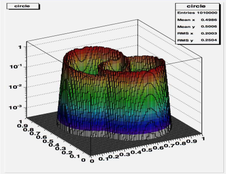
- Assume m bins in each dimension of $\vec{x}$.
- For each bin k in each dimension $\eta \in [1, n]$ assume a **weight (probability)** $\alpha_k^{(\eta)}$ for x_k to be in that bin.

  Condition(s) on the weights:
  $\alpha_k^{(\eta)} \in [0, 1]$, $\sum_{k=1}^{m} \alpha_k^{(\eta)} = 1$.
- For each bin in each dimension calculate $\langle I_k^{(\eta)} \rangle$ and $\langle E_k^{(\eta)} \rangle$.

  Obviously, for all η, $\langle I \rangle = \sum_{k=1}^{m} \langle I_k^{(\eta)} \rangle$, but error estimates different.
- In each dimensions, iterate and update the $\alpha_k^{(\eta)}$; example for updating:

  $$\alpha_k^{(\eta)}(\text{rm new}) \propto \alpha_k^{(\eta)}(\text{rm old}) \left(\frac{E_k^{(\eta)}}{E_{\text{tot.}}(\eta)} \right)^{\kappa} .$$
- Problem with this simple algorithm:
  Gets a hold only on fluctuations $\parallel$ to binning axes.





Monte Carlo integration

Multichannel sampling

Basic idea: Use a sum of functions $g_i(\vec{x})$ as Jacobian $g(\vec{x})$.

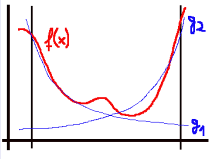
$\Longrightarrow$ $g(\vec{x}) = \sum_{i=1}^{N} \alpha_i g_i(\vec{x})$;

$\Longrightarrow$ condition on weights like stratified sampling;

("Combination" of importance & stratified sampling).

Algorithm for one iteration:

- Select g_i with probability $\alpha_i \longrightarrow \vec{x}_j$.
- Calculate total weight $g(\vec{x}_j)$ and partial weights $g_i(\vec{x}_j)$
- Add $f(\vec{x}_j)/g(\vec{x}_j)$ to total result and $f(\vec{x}_j)/g_i(\vec{x}_j)$ to partial (channel-) results.
- After N sampling steps, update a-priori weights.



This is the method of choice for parton level event generation!

Monte Carlo integration

Selecting after sampling: Unweighting efficiency

Basic idea: Use hit-or-miss method;

Generate $\vec{x}$ with integration method,

compare actual $f(\vec{x})$ with maximal value during sampling;

$\Longrightarrow$ "Unweighted events".

Comments:

- unweighting efficiency, $w_{\rm eff} = \langle f(\vec{x}_j)/f_{\rm max}\rangle$ = number of trials for each event.
- Good measure for integration performance.
- Expect $\log_{10} w_{\rm eff} \approx 3-5$ for good integration of multi-particle final states at tree-level.
- Maybe acceptable to use $f_{\rm max,eff} = K f_{\rm max}$ with $K > 1$.
  Problem: what to do with events where $f(\vec{x}_j)/f_{\rm max,eff} > 1$?
  Answer: Add $\mathrm{int}[f(\vec{x}_j)/f_{\rm max,eff}] = k$ events and perform hit-or-miss on $f(\vec{x}_j)/f_{\rm max,eff} - k$.

Monte Carlo integration

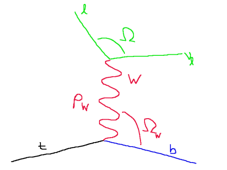
Particle physics example: Evaluation of cross sections

- Simple example: $t \longrightarrow bW^+ \longrightarrow b\bar{l}\nu_l$:

$$|\mathcal{M}|^2 = \frac{1}{2}\left(\frac{8\pi\alpha}{\sin^2\theta_W}\right)^2 \frac{p_t \cdot p_\nu \; p_b \cdot p_l}{(p_W^2 - M_W^2)^2 + \Gamma_W^2 M_W^2}$$

- Phase space integration (5-dim)

$$\Gamma = \frac{1}{2m_t}\frac{1}{128\pi^3}\int \mathrm{d}p_W^2 \frac{\mathrm{d}^2\Omega_W}{4\pi}\frac{\mathrm{d}^2\Omega}{4\pi}\left(1 - \frac{p_W^2}{m_t^2}\right)|\mathcal{M}|^2$$



Advantages

- Throw 5 random numbers, construct four-momenta ($\Longrightarrow$ full kinematics, “events”)
- Apply smearing and/or arbitrary cuts.
- Simply histogram any quantity of interest - no new calculation for each observable

Parton level simulations

Stating the problem(s)

- Multi-particle final states for signals & backgrounds.
- Need to evaluate $\mathrm{d}\sigma_N$:

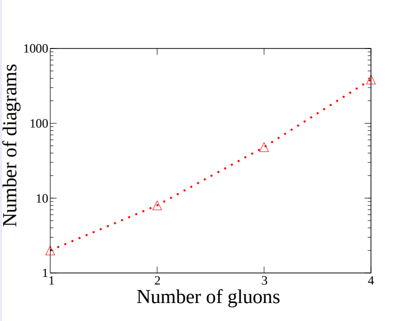
$$\int\limits_{\text{cuts}} \left[\prod_{i=1}^{N} \frac{\mathrm{d}^3 q_i}{(2\pi)^3 2E_i}\right] \delta^4\left(p_1 + p_2 - \sum_i q_i\right) \left|\mathcal{M}_{p_1 p_2 \to N}\right|^2 .$$

- Problem 1: Factorial growth of number of amplitudes.
- Problem 2: Complicated phase-space structure.
- Solutions: Numerical methods.

Parton level simulations

Factorial growth: $e^+e^- \to q\bar{q} + ng$

n	$\#_{\rm diags}$
0	1
1	2
2	8
3	48
4	384



Parton level simulations

Basic ideas of efficient ME calculation

Need to evaluate $|\mathcal{M}|^2 = \left|\sum_i \mathcal{M}_i\right|^2$

- Obvious: Traditional textbook methods (squaring, completeness relations, traces) fail
  $\Longrightarrow$ result in proliferation of terms $(\mathcal{M}_i\mathcal{M}_j^*)$
  $\Longrightarrow$ Better: Amplitudes are complex numbers,
  $\Longrightarrow$ add them before squaring!
- Remember: spinors, gamma matrices have explicit form could be evaluated numerically (brute force)
  But: Rough method, lack of elegance, CPU-expensive

Parton level simulations

Helicity method

- Introduce basic helicity spinors (needs to "gauge"-vectors)
- Write everything as spinor products, e.g.

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

- Also: $(\not{p}+m) \implies \frac{1}{2}\sum_h \left[\left(1+\frac{m^2}{p^2}\right)\bar{u}(p,\,h)u(p,\,h) + \left(1-\frac{m^2}{p^2}\right)\bar{v}(p,\,h)v(p,\,h)\right]$

(completeness relation)

- Find other genuine expressions:

$$\begin{aligned} Y(p_1, h_1,\; p_2, h_2) &:= \bar{u}(p_1, h_1)u(p_2, h_2) \\ X(p_1, h_1,\; p_2, h_2,\; p_3) &:= \bar{u}(p_1, h_1)\not{p}_3 u(p_2, h_2) \\ Z(p_1, h_1,\; p_2, h_2;\; p_3, h_3,\; p_4, h_4) &:= \bar{u}(p_1, h_1)\gamma^\mu u(p_2, h_2)\bar{u}(p_3, h_3)\gamma^\mu u(p_4, h_4)\,, \end{aligned}$$

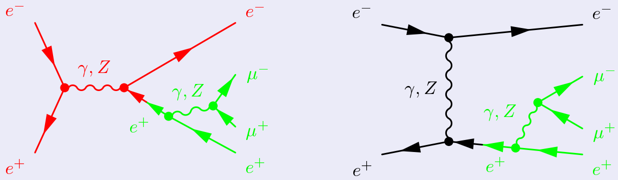
all complex-valued functions of momenta & helicities.

Parton level simulations

Taming the factorial growth in the helicity method

- Reusing pieces: Calculate only once!
- Factoring out: Reduce number of multiplications!

Implemented as a-posteriori manipulations of amplitudes.



Parton level simulations

Recursion methods (off-shell)

Basic idea: Recursively build one-particle off-shell currents (various versions of this: Berends-Giele, Alpha etc.).

"Classical" example: n-gluon amplitudes:

- Start with two on-shell gluons, represented by their polarization vectors, hence the currents associated with them are $J^\nu(k) = \varepsilon^\nu(k)$.
- Then the two-gluon current reads (no colors) $J^\mu(k = k_1 + k_2) = \frac{ig_3}{(k_1+k_2)^2} V^{\mu\nu\rho} J_\nu(k_1) J_\rho(k_2)$.
- From this, larger and larger currents can be built recursively.
- For quarks, the currents are given by spinors, and similar reasoning applies for the construction of the one-particle off-shell currents.
- Treatment of color: Color-ordering the amplitudes $\Longrightarrow \mathcal{C}^{(1,\,\dots,\,n)} = \mathrm{Tr}\,[T^{a_1}\,\dots\,T^{a_n}]$, where T^a are color matrices in fundamental representation.
- Problem: Need to sum over all allowed permutations.

Parton level simulations

Integration methods: Multi-channeling

Basic idea: Translate Feynman diagrams into channels

$\Longrightarrow$ decays, s- and t-channel props as building blocks.

R.Kleiss and R.Pittau, Comput. Phys. Commun. **83** (1994) 141

Integration methods: "Democratic" methods

- Rambo/Mambo: Flat & isotropic

  R.Kleiss, W.J.Stirling and S.D.Ellis, Comput. Phys. Commun. **40** (1986) 359;

- HAAG: Follows QCD antenna pattern

  A.van Hameren and C.G.Papadopoulos, Eur. Phys. J. C **25** (2002) 563.

Limitations of parton level simulation

Factorial growth

- . . . persists due to the number of color configurations

  (e.g. $(n-1)!$ permutations for n external gluons).

- Solution: Sampling over colors,
  but correlations with phase space
  $\Longrightarrow$ Best recipe not (yet) found.

- New scheme for color: color dressing

  (C.Duhr, S.Hoche and F.Maltoni,JHEP **0608** (2006) 062)

Limitations of parton level simulation

Factorial growth

- Off-shell vs. on-shell recursion relations:

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	-

Time [s] for the evaluation of 10^4 phase space points, sampled over helicities & color.

Limitations of parton level simulation

Efficient phase space integration

- Main problem: Adaptive multi-channel sampling translates "Feynman diagrams" into integration channels $\Longrightarrow$ hence subject to growth.
- But it is practical only for 1000-10000 channels.
- Therefore: Need better sampling procedures $\Longrightarrow$ open question with little activity.

(Private suspicion: Lack of glamour)

Limitations of parton level simulation

General

- Fixed order parton level (LO, NLO, . . .) implies fixed multiplicity
- No control over potentially large logs
  (appear when two partons come close to each other).
- Parton level is parton level
  experimental definition of observables relies on hadrons.

Therefore: Need hadron level event generators!

Survey of existing parton-level tools

Standard Model (and beyond) tools @ tree-level

- AlpGen M.L.Mangano *et al.*, JHEP **0307** (2003) 001;
- AMEGIC++ F.K., R.Kuhn, G.Soff, JHEP **0202** (2002) 044;
- CompHEP/CalcHEP E.Boos *et al.* Nucl. Instrum. Meth. A **534** (2004) 250;
- HELAC A.Kanaki and C.G.Papadopoulos, Comput. Phys. Commun. **132** (2000) 306;
- MadGraph/MadEvent F.Maltoni and T.Stelzer, JHEP **0302**, 027 (2003);
- O'Mega+WHIZARD M.Moretti, T.Ohl and J.Reuter, arXiv:hep-ph/0102195.

All tools here are completely self-contained and automated and provide amplitudes and integrators of their own.

Survey of existing parton-level tools

Comparison of tree-level tools

	Models	$2 \longrightarrow n$	Ampl.	Integ.	public?	lang.
Alpgen	SM	$n = 8$	rec.	Multi	yes	Fortran
Amegic	SM,MSSM,ADD	$n = 6$	hel.	Multi	yes	C++
CompHep	SM,MSSM	$n = 4$	trace	1Channel	yes	C
HELAC	SM	$n = 8$	rec.	Multi	yes	Fortran
MadEvent	SM,MSSM	$n = 6$	hel.	Multi	yes	Fortran
O'Mega	SM,MSSM,LH	$n = 8$	rec.	Multi	yes	O'Caml

A short detour to NLO calculations

MC calculations at NLO QCD

- Calculate two separate, divergent integrals
$$\sigma_{NLO} = \int\limits_{m+1} d\sigma_R + \int\limits_{m} d\sigma_V$$
- Real emission in $d\sigma_R$, virtual loop in $d\sigma_V$.
- Divergent structures due to soft/collinear particles.
- Combine before numerical integration to cancel divergences (KLN theorem guarantees cancellation).
- Two solutions: Phase space slicing and subtraction.

A short detour to NLO calculations

Illustrative 1-dim example

- $|\mathcal{M}^R_{m+1}|^2 = \frac{1}{x}R(x)$, where x=gluon energy or similar.
- $|\mathcal{M}^V_m|^2 = \frac{1}{\epsilon}V$, regularized in $d = 4 - 2\epsilon$ dimensions.
- Cross section in d dimensions with jet measure F^J:

  $$\sigma = \int\limits_0^1 \frac{\mathrm{d}x}{x^{1+\epsilon}} R(x) F_1^J(x) + \frac{1}{\epsilon} F_0^J$$

- Infrared safety of jet measure: $F_1^J(0) = F_0^J$

  $\Longrightarrow$ "A soft/collinear parton has no effect." (Tricky!)
- KLN theorem: $R(0) = V$.

A short detour to NLO calculations

Phase space slicing

W.T.Giele and E.W.N.Glover, Phys. Rev. D **46** (1992) 1980.

- Introduce arbitrary cutoff $\delta \ll 1$:

$$\begin{aligned}
\sigma &= \int_0^\delta \frac{\mathrm{d}x}{x^{1+\epsilon}} R(x) F_1^J(x) + \frac{1}{\epsilon} F_0^J + \int_\delta^1 \frac{\mathrm{d}x}{x^{1+\epsilon}} R(x) F_1^J(x) \\
&\approx \int_0^\delta \frac{\mathrm{d}x}{x^{1+\epsilon}} V F_0^J + \frac{1}{\epsilon} F_0^J + \int_\delta^1 \frac{\mathrm{d}x}{x} R(x) F_1^J(x) \\
&= \log(\delta) V F_0^J + \int_\delta^1 \frac{\mathrm{d}x}{x} R(x) F_1^J(x)
\end{aligned}$$

- Two separate finite integrals - both numerically large
  $\Longrightarrow$ error blows up (trial and error for stability)

A short detour to NLO calculations

Subtraction method

S.Catani and M.H.Seymour, Nucl. Phys. B **485** (1997) 291

- Rewrite

$$\begin{aligned}\sigma &= \int_0^1 \frac{\mathrm{d}x}{x^{1+\epsilon}} R(x) F_1^J(x) - \int_0^1 \frac{\mathrm{d}x}{x^{1+\epsilon}} V F_0^J + \int_0^1 \frac{\mathrm{d}x}{x^{1+\epsilon}} V F_0^J + \frac{1}{\epsilon} F_0^J \\ &= \int_0^1 \frac{\mathrm{d}x}{x^{1+\epsilon}} \left(R(x) F_1^J(x) - V F_0^J \right) + \mathcal{O}(1) F_0^J \,.\end{aligned}$$

- Two separate finite integrals, with no large numbers to be added/subtracted.
- Subtraction terms are universal (analytical bit can be calculated once and for all).

Summary of lecture 1

- MC as convenient numerical integration method.
- Well-suited to particle physics (integrand with large fluctuations in many dimensions)
- At LO: integrand positive definite
  $\Longrightarrow$ straight path to simulation/event generation
- Can be fully exclusive - direct comparison with data
- But: Parton level not always adequate, need hadron level!
- Subject of next lectures.