Introduction to Event Generators

Frank Krauss

Institute for Particle Physics Phenomenology
Durham University

February 2, 2010

Topics of the lectures

1. Lecture 1: *The Monte Carlo Principle*
2. Lecture 2: *Parton level event generation*
3. Lecture 3: *Dressing the Partons*
4. Lecture 4: *Modelling beyond Perturbation Theory*

Thanks to

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

Menu of lecture 2

- Prelude: Orientation
- Stating the problem: Factorial growth
- Efficient matrix element calculation and phase space evaluation at leading order (tree-level)
- Survey of leading order tools
- Next-to leading order

Prelude: Orientation

Event generator paradigm

Divide event into stages, separated by different scales.

- Signal/background:

  Exact matrix elements.

- QCD-Bremsstrahlung:

  Parton showers (also in initial state).

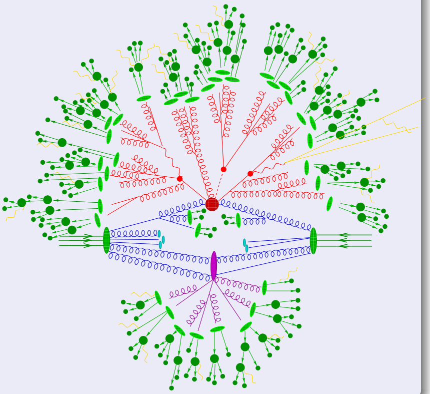
- Multiple interactions:

  Beyond factorisation: Modelling.

- Hadronisation:

  Non-perturbative QCD: Modelling.

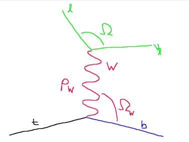
Sketch of an event



Simulation of the hard bits (signals & backgrounds)

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

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



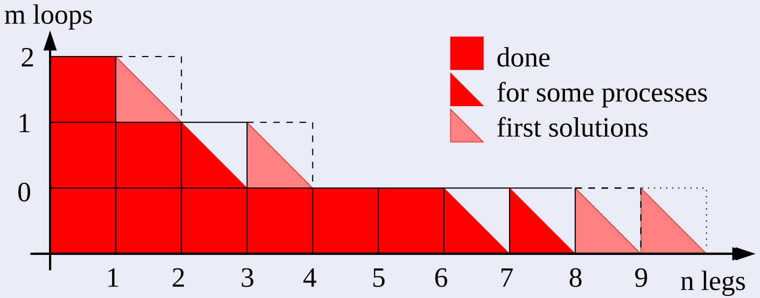
- Phase space integration (5-dim):

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

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

Availability of exact calculations (hadron colliders)

- Fixed order matrix elements ("parton level") are exact to a given perturbative order. (and often quite a pain!)
- Important to understand limitations:
  Only tree-level fully automated, 1-loop-level ongoing.



Parton level simulations

Stating the problem(s)

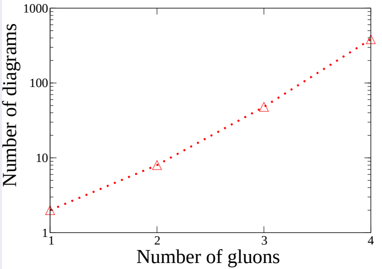
- Multi-particle final states for signals & backgrounds.
- Need to evaluate $\mathrm{d}\sigma_N$:

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

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

Example for factorial growth: $e^+e^- \longrightarrow q\bar{q} + ng$

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



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

Helicity method

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

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

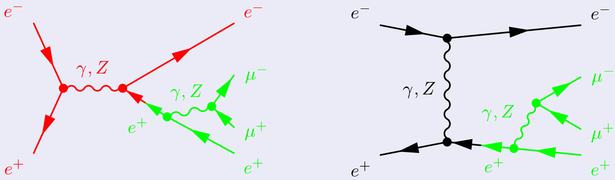
Completeness rel'n: $(\not{p}+m) \implies \frac{1}{2}\sum_h\left[\left(1+\frac{m^2}{p^2}\right)\bar{u}(p,\, h)u(p,\, h)+\left(1-\frac{m^2}{p^2}\right)\bar{v}(p,\, h)v(p,\, h)\right]$

- There are other genuine expressions ...
- Translate Feynman diagrams into "helicity amplitudes": complex-valued functions of momenta & helicities.
- Spin-correlations etc. nearly come for free.

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.



- Better method: Recursion relations (recycling built in).
  Best candidate so far: Off-shell recursions

(Dyson-Schwinger, Berends-Giele etc.)

Colour-dressing: Fighting factorial growth in colour

- In principle: sampling over colours improves situation.

(But still, e.g. naively $\simeq (n-1)!$ permutations/colour-ordering for n external gluons).

- Improved scheme: colour dressing.

F.Maltoni, K.Paul, T.Stelzer & S.Willenbrock Phys. Rev. **D67** (2003) 014026

- Works very well with Berends-Giele recursions:

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

Final	BG		BCF		CSW	
State	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 & colour.

Efficient phase space integration

("Amateurs study strategy, professionals study logistics")

- Democratic, process-blind integration methods:
  - Rambo/Mambo: Flat & isotropic

    R.Kleiss, W.J.Stirling & S.D.Ellis, Comput. Phys. Commun. **40** (1986) 359;
  - HAAG/Sarge: Follows QCD antenna pattern

    A.van Hameren & C.G.Papadopoulos, Eur. Phys. J. C **25** (2002) 563.
- Multi-channeling: Each Feynman diagram related to a phase space mapping (= "channel"), optimise their relative weights. R.Kleiss & R.Pittau, Comput. Phys. Commun. **83** (1994) 141.
- Main problem: practical only up to $\mathcal{O}(10\mathrm{k})$ channels.
- Some improvement by building phase space mappings recursively: More channels feasible, efficiency drops a bit.

Monte Carlo integration: Unweighting efficiency

- Want to generate events "as in nature".
- 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 $f_{\rm max,eff} = Kf_{\rm max}$ 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 $f(\vec{x}_j)/f_{\rm max,eff} - k$.

Best answer at the moment: COMIX (personal bias)

T.Gleisberg & S.Hoeche, JHEP **0812** (2008) 039

- Colour-dressed Berends-Giele amplitudes in the SM.
- Fully recursive phase space generation.
- Example results (cross sections):

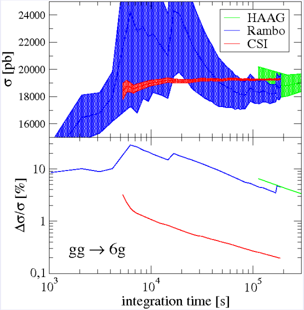
$gg \longrightarrow ng$	Cross section [pb]				
n	8	9	10	11	12
$\sqrt{s}$ [GeV]	1500	2000	2500	3500	5000
COMIX	0.755(3)	0.305(2)	0.101(7)	0.057(5)	0.019(2)
Maltoni (2002)	0.70(4)	0.30(2)	0.097(6)		
ALPGEN	0.719(19)				

σ [μb]	Number of jets						
$b\bar{b}$ + QCD jets	0	1	2	3	4	5	6
COMIX	470.8(5)	8.83(2)	1.826(8)	0.459(2)	0.1500(8)	0.0544(6)	0.023(2)
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)				

Best answer at the moment: COMIX (personal bias)

T.Gleisberg & S.Hoeche, JHEP **0812** (2008) 039

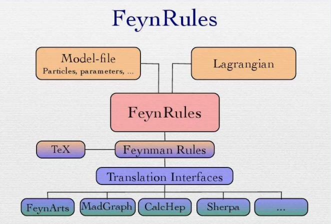
- Colour-dressed Berends-Giele amplitudes in the SM.
- Fully recursive phase space generation.
- Example results (phase space performance):



FeynRules: Implementing new models made easy

Aim

- Portable, transparent & reproducible implementation of (nearly arbitrary) new physics models.
- In most codes: New models given by new particles, their properties & interactions.
- Output to standard ME generators enabled (MadGraph, Sherpa, ...)



- Various models already implemented & validated for a list: http://feynrules.phys.ucl.ac.be

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++
COMIX	SM	$n = 8$	rec.	Multi	yes	C++
COMPHEP	SM,MSSM	$n = 4$	trace	1Channel	yes	C
HELAC	SM	$n = 8$	rec.	Multi	yes	Fortran
MADEVENT	SM,MSSM,UED	$n = 6$	hel.	Multi	yes	Fortran
WHIZARD	SM,MSSM,LH	$n = 8$	rec.	Multi	yes	O'Caml

Limitations of parton level simulation

Factorial growth

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

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

- Solution: Sampling over colours,

  but correlations with phase space

  $\Longrightarrow$ Best recipe not (yet) found.

- New scheme for colour: colour dressing

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

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)

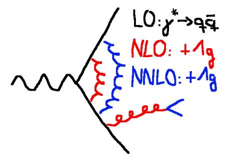
General

- 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 is parton level ...
  experimental definitions rely on observable hadrons.

Therefore: Need hadron level event generators!

A short detour to NLO calculations

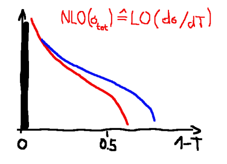
Nomenclature (example: $\gamma^* \to$ hadrons)



- In general: N^{n}LO $\leftrightarrow$ $\mathcal{O}(\alpha_s^n)$
- But: only for inclusive quantities

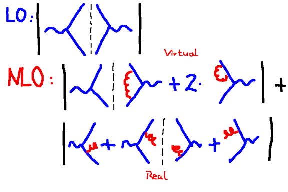
(e.g.: total xsecs like γ^* $\to$hadrons).

Counter-example: thrust distribution



- In general, distributions are HO.
- Distinguish real & virtual emissions:
  Real emissions $\to$ mainly distributions,
  virtual emissions $\to$ mainly normalisation.

Anatomy: Virtual and real corrections



NLO corrections: $\mathcal{O}(\alpha_s)$

Virtual corrections	=	extra loops
Real corrections	=	extra legs

- UV-divergences in virtual graphs → renormalisation
- But also: IR-divergences in real & virtual contributions
  Must cancel each other, non-trivial to see:
  N vs. $N+1$ particle FS, divergence in PS vs. loop

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.

Illustrative 1-dim example

- $|\mathcal{M}^R_{m+1}|^2 = \frac{1}{x}R(x)$, where x=gluon energy or similar.
- $|\mathcal{M}^V_m|^2 = \frac{1}{\epsilon}V$, regularised in $d = 4 - 2\epsilon$ dimensions.
- Cross section in d dimensions with jet measure F^J:
  $$\sigma = \int\limits_0^1 \frac{\mathrm{d}x}{x^{1+\epsilon}} R(x) F^J_1(x) + \frac{1}{\epsilon} V F^J_0$$
- Infrared safety of jet measure: $F^J_1(0) = F^J_0$
  $\Longrightarrow$ "A soft/collinear parton has no effect."
  (Tricky issue - without it, no reliable NLO calculation!)
- KLN theorem: $R(0) = V$.

Phase space slicing in 1-dim example

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

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

$$\begin{aligned}
\sigma &= \int\limits_0^\delta \frac{\mathrm{d}x}{x^{1+\epsilon}} R(x) F_1^J(x) + \frac{1}{\epsilon} V F_0^J + \int\limits_\delta^1 \frac{\mathrm{d}x}{x^{1+\epsilon}} R(x) F_1^J(x) \\
&\approx \int\limits_0^\delta \frac{\mathrm{d}x}{x^{1+\epsilon}} V F_0^J + \frac{1}{\epsilon} V F_0^J + \int\limits_\delta^1 \frac{\mathrm{d}x}{x} R(x) F_1^J(x) \\
&= \log(\delta) V F_0^J + \int\limits_\delta^1 \frac{\mathrm{d}x}{x} R(x) F_1^J(x)
\end{aligned}$$

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

Subtraction method in 1-dim example

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

- Rewrite

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

- 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).

Parton level tools: Loop level

Specific solutions

- Currently only process-specific codes/calculations, e.g.:
  - NLOJET++ (jets only),
  - VBFNLO (VBF-type processes),
  - and MCFM (the interesting rest)

  Common to both: Reach so far to $2 \to 3$ processes.
- Bottleneck so far: virtual contributions
  $\Longrightarrow$ a general solution is in sight
  keywords: OPP-method, generalised unitarity
  (both calculated $W+3$ jets in hadronic collisions)

R.K.Ellis, K.Melnikov and G.Zanderighi, JHEP **0904** (2009) 077; C.F.Berger *et al.*, arXiv:0902.2760

Automated real subtraction algorithms

- Remaining major nuisance in NLO calculations: real contributions & subtraction $\Longrightarrow$ has been "solved", i.e. automated.
- In principle: simple ("only" tree-level) & general (process-independent subtraction schemes).
- A problem that begs for automation.
- Status by now:
  - Various implementations documented in different stages, all building on Catani-Seymour subtraction.
  - Only one fully public version (MADDIPOLE), but lacks integrated dipoles.
  - A private version in SHERPA has been used to calculate the $W + 3$ jet cross section in BLACKHAT.

C. F. Berger *et al.*, arXiv:0902.2760

Summary of lecture 2

- A first level of simulation: parton level.
- Brief review of state-of-the-art there.
- Discussed automated generation of matrix elements and their phase space integration.
- Many tools available for tree-level multi-leg.
- Going to loop-level in an automated way just started now.
- Discussed some intricacies of NLO calculations.