

YETI'06–SM IPPP, Durham, UK 27–29 March 2006

LUND UNIVERSITY

Monte Carlo Event Generators

Torbjörn Sjöstrand Lund University

1. (Monday) Introduction and Overview; Monte Carlo Techniques

2. (Monday) Matrix Elements; Parton Showers I

3. (yesterday) Parton Showers II; Matching Issues

4. (yesterday) Multiple Interactions and Beam Remnants

5. (today) Hadronization and Decays; Summary and Outlook

Event Physics Overview

Repetition: from the "simple" to the "complex", or from "calculable" at large virtualities to "modelled" at small

Matrix elements (ME):

Parton Showers (PS):

1) Hard subprocess: $|\mathcal{M}|^2$, Breit-Wigners, parton densities.



3) Final-state parton showers.



2) Resonance decays: includes correlations.



4) Initial-state parton showers.



5) Multiple parton–parton interactions.



6) Beam remnants, with colour connections.



5) + 6) = Underlying Event

7) Hadronization



8) Ordinary decays: hadronic, τ , charm, ...



Hadronization/Fragmentation models

Perturbative \rightarrow nonperturbative \implies not calculable from first principles!

Model building = ideology + "cookbook"

Common approaches:

- 1) **String** Fragmentation (most ideological)
- 2) **Cluster** Fragmentation (simplest?)
- 3) **Independent** Fragmentation (most cookbook)
- 4) Local Parton–Hadron Duality (limited applicability)
- Best studied in $e^+e^- \rightarrow \gamma^*/Z^0 \rightarrow q\overline{q}$



The Lund String Model

In QED, field lines go all the way to infinity



since photons cannot interact with each other.

Potential is simply additive:

$$V(\mathbf{x}) \propto \sum_i rac{1}{|\mathbf{x} - \mathbf{x}_i|}$$

In QCD, for large charge separation, field lines seem to be compressed to tubelike region(s) \Rightarrow **string(s)**



by self-interactions among soft gluons in the "vacuum". (Non-trivial ground state with quark and gluon "condensates". Analogy: vortex lines in type II superconductor)

Gives linear confinement with string tension:

 $F(r) \approx \text{const} = \kappa \approx 1 \text{ GeV/fm} \iff V(r) \approx \kappa r$

Separation of transverse and longitudinal degrees of freedom ⇒ simple description as 1+1-dimensional object – string – with Lorentz invariant formalism

V(r)V(R) T(6S)bb-0.9 T(5S) Verter de este de ter de de linear part 0.8<u>T(4S</u>) $2M_{B}$ total 0.7 <u>T(35</u>) _{(°}) r(1D) r(2S) 0.6 $h_b(1P) \chi_b(1P)$ 0.5 **Coulomb part** 0.4 T(15) $^{O}~V(R)=V_{\mathfrak{s}}+K~R-e/R+f/R^{2}$ 0.3 <u>- L -</u> 16 20 12 24 8 R 1+- $(0,1,2)^+$ 1 $\kappa \approx 1 \text{ GeV/fm.}$ $V(r) \approx -\frac{4\alpha_s}{3r} + \kappa r \approx -\frac{0.13}{r} + r$ (for $\alpha_s \approx 0.5$, r in fm and V in GeV) $V(0.4 \text{ fm}) \approx 0$: Coulomb important for internal structure of hadrons, not for particle production (?)

Linear confimenent confirmed e.g. by quenched lattice QCD

Real world (??, or at least unquenched lattice QCD) \implies nonperturbative string breakings $gg \ldots \rightarrow q\overline{q}$



Repeat for large system \Rightarrow Lund model which neglects Coulomb part:

$$\left|\frac{\mathrm{d}E}{\mathrm{d}z}\right| = \left|\frac{\mathrm{d}p_z}{\mathrm{d}z}\right| = \left|\frac{\mathrm{d}E}{\mathrm{d}t}\right| = \left|\frac{\mathrm{d}p_z}{\mathrm{d}t}\right| = \kappa$$

Motion of quarks and antiquarks in a $q\overline{q}$ system:



gives simple but powerful picture of hadron production (with extensions to massive quarks, baryons, ...)

How does the string break?



String breaking modelled by tunneling:

$$\mathcal{P} \propto \exp\left(-\frac{\pi m_{\perp q}^2}{\kappa}\right) = \exp\left(-\frac{\pi p_{\perp q}^2}{\kappa}\right) \exp\left(-\frac{\pi m_q^2}{\kappa}\right)$$

1) common Gaussian p_{\perp} spectrum

2) suppression of heavy quarks $u\overline{u} : d\overline{d} : s\overline{s} : c\overline{c} \approx 1 : 1 : 0.3 : 10^{-11}$

3) diquark \sim antiquark \Rightarrow simple model for baryon production

Hadron composition also depends on spin probabilities, hadronic wave functions, phase space, more complicated baryon production, . . . ⇒ "moderate" predictivity (many parameters!) Fragmentation starts in the middle and spreads outwards:





f(z), a = 0.5, b= 0.7

but breakup vertices causally disconnected \Rightarrow can proceed in arbitrary order

 \Rightarrow *left–right symmetry*

$$\mathcal{P}(1,2) = \mathcal{P}(1) \times \mathcal{P}(1 \rightarrow 2)$$

= $\mathcal{P}(2) \times \mathcal{P}(2 \rightarrow 1)$

 \Rightarrow Lund symmetric fragmentation function $f(z) \propto (1-z)^a \exp(-bm_{\perp}^2/z)/z$



The iterative ansatz



Scaling in lightcone $p_{\pm} = E \pm p_z$ (for $q\overline{q}$ system along z axis) implies flat central rapidity plateau + some endpoint effects:



 $\langle n_{\rm Ch} \rangle \approx c_0 + c_1 \ln E_{\rm Cm}$, ~ Poissonian multiplicity distribution

The Lund gluon picture



Gluon = kink on string, carrying energy and momentum Force ratio gluon/ quark = 2, cf. QCD $N_C/C_F = 9/4$, $\rightarrow 2$ for $N_C \rightarrow \infty$ No new parameters introduced for gluon jets!, so:

- Few parameters to describe energy-momentum structure!
 - Many parameters to describe flavour composition!

Independent fragmentation

Based on a similar iterative ansatz as string, but



Further numerous and detailed tests at LEP favour string picture but much is still uncertain when moving to hadron colliders.

Lund news: fragmentation of junction topology





Q=168.3 GeV

Q=349.0 GeV

Q=4845.4 GeV

 10^{2}

103

 10^{1}



1) Introduce forced $g \rightarrow q\overline{q}$ branchings 2) Form colour singlet clusters 3) Clusters decay isotropically to 2 hadrons according to phase space weight $\sim (2s_1 + 1)(2s_2 + 1)(2p^*/m)$ simple and clean, but ...

1) Tail to very large-mass clusters (e.g. if no emission in shower); if large-mass cluster \rightarrow 2 hadrons then

incorrect hadron momentum spectrum, crazy four-jet events

 \implies split big cluster into 2 smaller along "string" direction;

daughter-mass spectrum \Rightarrow iterate if required;

 \sim 15% of primary clusters are split, but give \sim 50% of final hadrons

2) Isotropic baryon decay inside cluster

 \implies splittings g \rightarrow qq + \overline{qq}

3) Too soft charm/bottom spectra \implies anisotropic leading-cluster decay

4) Charge correlations still problematic \implies all clusters anisotropic (?)

5) Sensitivity to particle content \implies only include complete multiplets



String vs. Cluster



"There ain't no such thing as a parameter-free good description"

Local Parton–Hadron Duality

Analytic approach: Run shower down to to $Q \approx \Lambda_{QCD}$ (or m_{hadron} , if larger) "Hard Line": each parton \equiv one hadron "Soft Line": local hadron density \propto parton density describes momentum spectra dn/dx_p and semi-inclusive particle flow, but fails for identified particles + "renormalons" (power corrections) $\langle 1 - T \rangle = a \alpha_{s}(E_{cm}) + b \alpha_{s}^{2}(E_{cm})$ $+c/E_{\rm CM}$



Not Monte Carlo, not for arbitrary quantities

Decays

Unspectacular/ungrateful but necessary: this is where most of the final-state particles are produced! Involves hundreds of particle kinds and thousands of decay modes.



- $B^{*0} \rightarrow B^0 \gamma$: electromagnetic decay
- $B^0 \rightarrow \overline{B}^0$ mixing (weak)

•
$$\overline{B}^0 \to D^{*+} \overline{\nu}_e e^-$$
: weak decay, displaced vertex, $|\mathcal{M}|^2 \propto (p_{\overline{B}} p_{\overline{\nu}})(p_e p_{D^*})$

- $D^{*+} \rightarrow D^0 \pi^+$: strong decay
- $D^0 \rightarrow \rho^+ K^-$: weak decay, displaced vertex, ρ mass smeared
- $\rho^+ \rightarrow \pi^+ \pi^0$: ρ polarized, $|\mathcal{M}|^2 \propto \cos^2 \theta$ in ρ rest frame
- $\pi^0 \rightarrow e^+e^-\gamma$: Dalitz decay, $m(e^+e^-)$ peaked

Dedicated programs, with special attention to polarization effects:

- EVTGEN: B decays
- TAUOLA: au decays

Jet Universality

Question: are jets the same in all processes?

Answer 1: no, at LEP mainly quarks jets, often b/c,

at LHC mainly gluons, if quarks then mainly u/d.

Answer 2: no, perturbative evolution gives calculable differences.





- Less perturbative evolution \Rightarrow strings less "wrinkled"?
- Many overlapping strings ⇒ collective phenomena?

Other program tasks/elements



- Colour reconnection: how well can we trust "perturbatively" calculable colour flow in soft region?
- Bose-Einstein: must we use amplitudes to describe production of identical particles? (\sim 50 π^+ , \sim 50 π^- , \sim 70 π^0 per event)
- Event measures; jet clustering routines; other utilities

...and more

Event Generator Practicalities



Event generation structure

1) Initialization step

- select process(es) to study
- modify physics parameters: m_t, m_h, \ldots
 - set kinematics constraints
 - modify generator performance
 - initialize generator
 - book histograms
 - 2) Generation loop
 - generate one event at a time
 - analyze it (or store for later use)
 - add results to histograms
 - print a few events
 - 3) Finishing step
 - print deduced cross-sections
 - print/save histograms etc.

How to run event generators

Often forced to use what is allowed by constricted collaboration framework, but for maximal power and minimal bugs run raw generator:

• HERWIG, ISAJET: supplied but modifiable main program, calling user-written routines



• PYTHIA: generator is subroutine package, user writes main program

C...Arithmetic in double precision; integer functions; PYDATA. IMPLICIT DOUBLE PRECISION(A-H, O-Z) INTEGER PYK,PYCHGE,PYCOMP EXTERNAL PYDATA

C...The event record and other common blocks.

COMMON/PYJETS/N,NPAD,K(4000,5),P(4000,5),V(4000,5) COMMON/PYDAT2/KCHG(500,4),PMAS(500,4),PARF(2000),VCKM(4,4) COMMON/PYSUBS/MSEL,MSELPD,MSUB(500),KFIN(2,-40:40),CKIN(200) COMMON/PYPARS/MSTP(200),PARP(200),MSTI(200),PARI(200)

C...Physics scenario.

MSEL=0 ! Mix subprocesses freely
MSUB(102)=1 ! g + g -> h0
MSUB(123)=1 ! f + f' -> f + f' + h0
MSUB(124)=1 ! f + f' -> f" + f"' + h0
PMAS(25,1)=300D0 ! Nominal Higgs mass.

C...Run parameters.

NEV=1000 ! Number of events ECM=14000D0 ! CM energy of run CKIN(1)=200D0 ! Minimum Higgs mass. CKIN(2)=400D0 ! Maximum Higgs mass.

C...Initialize and book histogram(s).

CALL PYINIT('CMS','p','p',ECM)

CALL PYBOOK(1, 'Higgs mass distribution',80,200D0,400D0)

C...Generate events and look at first few.

DO 200 IEV=1,NEV

CALL PYEVNT

IF(IEV.LE.1) CALL PYLIST(1)

C...Find Higgs and fill its mass. End event loop.

DO 150 I=7,9

IF(K(I,2).EQ.25) CALL PYFILL(1,P(I,5),1D0)

150 CONTINUE

200 CONTINUE

C...Final output.

CALL PYSTAT(1) ! Print cross section table CALL PYHIST ! Print histogram(s) END



On To C++

Currently HERWIG and PYTHIA are successfully being used, also in new LHC environments, using C++ wrappers

> Q: Why rewrite? A1: Need to clean up! A2: Fortran 77 is limiting

Q: Why C++? A1: All the reasons for ROOT, Geant4, ... ("a better language", industrial standard, ...) A2: Young experimentalists will expect C++ (educational and professional continuity) A3: Only game in town! Fortran 90

So far mixed experience:

- Conversion effort: everything takes longer and costs more (as for LHC machine, detectors and software)
- The physics hurdle is as steep as the C++ learning curve

C++ Players

PYTHIA7 project ⇒ **ThePEG** Toolkit for High Energy Physics Event Generation (L. Lönnblad; S. Gieseke, A. Ribon, P. Richardson)

HERWIG++: complete reimplementation (S. Gieseke, D. Grellscheid, A. Ribon, P. Richardson, M.H. Seymour, P. Stephens, B.R. Webber; M. Bähr, M. Gigg, K. Hamilton, S. Latunde-Dada, S. Plätzer, A. Sherstnev)

ARIADNE/LDC: to do ISR/FSR showers, multiple interactions (L. Lönnblad; N. Lavesson)

SHERPA: partly wrappers to PYTHIA Fortran; has CKKW (F. Krauss; T. Fischer, T. Gleisberg, S. Hoeche, T. Laubrich, R. Matyszkiewicz, A. Schaelicke, C. Semmling, F. Siegert, S. Schumann, J. Winter)

PYTHIA8: restart to write complete event generator (T. Sjöstrand, (S. Mrenna, P. Skands))

What is ThePEG?

Toolkit for High Energy Physics Event Generation



not SHERPA

Running ThePEG

• ThePEG defines a set of abstract Handler classes for hard partonic sub-processes, parton densities, QCD cascades, hadronization, ...

• These handler classes interacts with the underlying structure using a special Event Record and a pre-defined set of virtual functions.

• The procedure to implement e.g. a new hadronization model, is to write a new (C++) class *inheriting* from the abstract HadronizationHandler base class, implementing the relevant virtual functions.

• The end-user will use a setup program to be able to pick objects corresponding to different physics models to build up an EventGenerator which then can be run interactively or off-line, or as a special slave program e.g. for Geant4.

• The setup program is used to choose between a multitude of predefined generators, to modify parameters and options of the selected models and, optionally, to specify the analysis to be done on the generated events.

• The Repository is the central part of the setup phase. It handles a structured list of all available objects and allows the user to manipulate them.

The new generator Herwig++

A completely new event generator in $C{++}$

- Aiming at full multi-purpose generator for LHC and future colliders.
- Preserving main features of HERWIG such as
 - angular ordered parton shower
 - cluster hadronization
- New features and improvements
 - covariant shower formulation
 - improved parton shower evolution for heavy quarks
 - consistent radiation from unstable particles (multiscale evolution)

Growth of Fortran HERWIG

Hard interactions

• Basic ME's included in ThePEG, such as:

$$e^+e^-
ightarrow qar q$$
, partonic $2
ightarrow 2$,

we use them.

- Soft and hard matrix element corrections imlemented for $e^+e^- \rightarrow q\bar{q}g$.
- AMEGIC++ will provide arbitrary ME's for multiparton final states via AMEGICInterface.
- LesHouchesFileReader enables to read in and process *any* hard event generated by parton level event generators (MadGraph/MadEvent, AlpGen, CompHEP,...).
- CKKW ME+PS foreseen.
- Other authors can easily include their own matrix elements (\rightarrow safety of OO code)

New/Future: HELAS like structures are already implemented for decays and spin correlations \longrightarrow allows us to code simple processes efficiently.

... and New Decays!

- Better decayers are being developed for almost all decay modes.
- $\rightarrow B$ decays.
- Spin correlations will be included.
- Major effort ongoing
 - a universal database is being set up.
 - contains 448 particles and 2607 decay modes at present.
 - possibility to generate configuration files for different generators (they need to write their own code however. . .).
- Particle data book as guideline.

 \longrightarrow look at examples. . .

Herwig++ Particle Properties DataBase	Page 1 of 2	Herwig++ Particle Properties DataBase	Page 2 of 2
Herwig++ Particle Properties This is the development version of the Herwig++ particle properties data	DataBase	Check baryon number conservation in decays Check lepton number conservation in decays	
replace the storage of particle properties as a text file to improve maintai	inance and accessiblity.	Check the spin is consistent with the id code	
This version is for the Herwig authors only and much of the information	is preliminary.		
The database currently contains 448 particles and 2607 decay modes.		<u>Peter Richardson</u> Last modified: Mon Jan 31 17:56:08 GMT 2005	
The information is available in a number of forms			
 The particles <u>numerically listed</u> according to the <u>PDC code</u> The particles <u>listed</u> according to the multiplets taken from the <u>PDC</u> The <u>decayers</u> The <u>width Generators</u> The <u>Mass Generators</u> The <u>references</u> Generate the <u>input files</u> for event generation 	3		
The contents of the database can be altered by following the links in the descriptions or by selecting an option below	particle table or particle		
• Add or modify a particle: 0			
• Add a decay mode for particle with id: 0			
Add a meson multiplet			
Add a decayer			
Add a width generator			
Add a mass generator			
Add a reference			
Set the multiplets			
• Set the decay modes for a particle to the charge conjugates of the a	antiparticle:0		
Add a Baryon 10plet			
Add a Baryon 35plet			
Add a Baryon 40plet			
Simple checks on the contents of the database			
Check charge conservation in decays			
file://C:\cygwin\home\seymout\Tex\Slides\Stefan\steffan\database.html	15/03/2005	file://C:\cygwin\home\seymour\Tex\Slides\Stefan\steffan\database.html	15/03/2005

What's next?

Near Future. . .

- ★ Initial state shower:
 - Complete implementation and tests.
- **★** Refine e^+e^- :
 - Full CKKW ME+PS matching.
 - Precision tune to LEP data should be possible.
- ★ with IS and FS showers running:
 - we can start to test Drell-Yan and jets in pp collisions.
 - cross check with Tevatron data and finally make predictions for the LHC.
- ★ Underlying Event.
- **★** Hadronic Decays: *NEW!* many new decayers, τ -decays, Spin correlations (P Richardson).
- ★ New Ideas: soft gluons, improved shower algorithm, NLO, . . .

Schedule?

• Ready for LHC!

Status of SHERPA

Scope:

• Full simulation of high energetic particle reactions at existing and future collider experiments, incl. e^+e^- , $\gamma\gamma$, $e\gamma$, ep, $p\bar{p}$ and pp collisions

Method:

- Account for multi-jet production through tree level matrix elements
- Combine them with the parton showers and hadronization according to the CKKW prescription

First α -version was released during MC4LHC workshop 2003, current version is SHERPA α -1.0.7

Sources:

- T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. S. and J. Winter, JHEP 0402:056,2004
- downloads, manual, bug reports etc. under

http://www.sherpa-mc.de

SHERPA including the ME's of AMEGIC++ and the CKKW prescription to combine them with the PS is a powerful tool to attempt the description of present-day Tevatron data and to study the extrapolation to LHC energies.

New features of version 1.0.7:

- Revised SUSY sector, including the SLHA interface
- τ -lepton and first hadron decays
- Improved decay chain treatment

Things to do:

- Alternative underlying event model
- Include and tune the cluster fragmentation model
- Supplement the hadron decays, special emphasis on b-hadrons
- **9** ...

PYTHIA8: A fresh start

Problem: PYTHIA7 stalled, no other manpower
Solution?: take a sabbatical and work "full-time"!
(⇒ baseline model, S. Mrenna & P. Skands join later ?)

٦	[ent	tati	ve	sc	hec	lul	le:
		lan	V C	30		T	С.

time	date	processes	final states
0 =	1 Sept. 2004	—	
1 =	1 Sept. 2005	LHA-style input	incomplete draft
2 =	1 Sept. 2006	a few processes	complete, buggy(?)
3 =	1 Sept. 2007	more processes	stable, debugged

... but don't forget Murphy's law

Objectives:

- clean up, keep the most recent models
- Les Houches Accord style input central
- independent of ThePEG (or anything else), but
- interface to ThePEG later written by L. Lönnblad (?)

Current PYTHIA8 structure

Vec4, Random, Settings, ParticleData, StandardModel, ...

Current PYTHIA8 status

	Existing classes		Missing classes/topics
Process Level	LHAinit LHAevnt (PYTHIA 6.3)	** ** * * *	Cross section administration Phase space selection Process matrix elements
Parton Level	TimeShower SpaceShower MultipleInteractions BeamRemnants	** ** **	Parton density libraries Resonance decays ThePEG input (?)
Hadron Level	StringFragmentation MiniStringFrag ParticleDecays	** ** **	colour flow models ME/PS matching
	Event BeamParticle	** **	updated decay tables Bose-Einstein
Vec4, Random Settings ParticleData	* * * ** **	event analysis routines	

Outlook

Generators in state of continuous development: * better & more user-friendly general-purpose matrix element calculators+integrators * * new libraries of physics processes, also to NLO * * more precise parton showers * \star better matching matrix elements \Leftrightarrow showers \star * improved models for underlying events / minimum bias * * upgrades of hadronization and decays * \star moving to C++ \star

 \Rightarrow always better, but never enough

But what are the alternatives, when event structures are complicated and analytical methods inadequate?

Final Words of Warning

[...] The Monte Carlo simulation has become the major means of visualization of not only detector performance but also of physics phenomena. So far so good. But it often happens that the physics simulations provided by the Monte Carlo generators carry the authority of data itself. They look like data and feel like data, and if one is not careful they are accepted as if they were data.

[...] I am prepared to believe that the computer-literate generation (of which I am a little too old to be a member) is in principle no less competent and in fact benefits relative to us in the older generation by having these marvelous tools. They do allow one to look at, indeed visualize, the problems in new ways. But I also fear a kind of "terminal illness", perhaps traceable to the influence of television at an early age. There the way one learns is simply to passively stare into a screen and wait for the truth to be delivered. A number of physicists nowadays seem to do just this.

J.D. Bjorken

from a talk given at the 75th anniversary celebration of the Max-Planck Institute of Physics, Munich, Germany, December 10th, 1992. As quoted in: Beam Line, Winter 1992, Vol. 22, No. 4