Introduction to Event Generators

Lecture 3: Hadronisation and Jets



MONTE CARLOS & FRAGMENTATION

PYTHIA anno 1978

(then called JETSET)

LU TP 78-18 November, 1978

A Monte Carlo Program for Quark Jet Generation

T. Sjöstrand, B. Söderberg

A Monte Carlo computer program is presented, that simulates the **fragmentation of a fast parton into a jet of mesons.** It uses an iterative scaling scheme and is compatible with the jet model of Field and Feynman.

Note:

Field-Feynman was an early fragmentation model Now superseded by the String (in PYTHIA) and Cluster (in HERWIG & SHERPA) models.

SUBROUTINE JETGEN(N) COMMON /JET/ K(100,2), P(100,5) COMMON /PAR/ PUD, PS1, SIGMA, CX2, EBEG, WFIN, IFLBEG COMMON /DATA1/ MESO(9,2), CMIX(6,2), PMAS(19) IFLSGN=(10-IFLBEG)/5 W=2.*E8EG 1=0 190=0 C 1 FLAVOUR AND PT FOR FIRST QUARK IFL1=IABS(IFLBEG) PT1=SIGMA*SQRT(-ALOG(RANF(D))) PHI1=6.2832*RANF(0) PX1=PT1*COS(PHI1) PY1=PT1*SIN(PHI1) 100 I=I+1 C 2 FLAVOUR AND PT FOR NEXT ANTIQUARK IFL2=1+INT(RANF(0)/PUD) PT2=SIGMA*SQRT(-ALOG(RANF(0))) PH12=6.2832*RANF(0) PX2=PT2*COS(PHI2) PY2=PT2*SIN(PHI2) C 3 MESON FORMED, SPIN ADDED AND FLAVOUR MIXED K(I,1)=MESO(3*(IFL1-1)+IFL2;IFLSGN) ISPIN=INT(PS1+RANF(0)) K(I,2)=1+9*ISPIN+K(I,1) IF(K(I,1).LE.6) GOTO 110 TMIX=RANF(0) KM=K(1,1)-6+3*ISPIN K(I,2)=8+9*ISPIN+INT(TMIX+CMIX(KM,1))+INT(TMIX+CMIX(KM,2)) C 4 MESON MASS FROM TABLE, PT FROM CONSTITUENTS 110 P(I,5)=PMAS(K(I,2)) P(I,1) = PX1 + PX2P(1,2) = PY1 + PY2PMTS=P(I,1)**2+P(I,2)**2+P(I,5)**2 C 5 RANDOM CHOICE OF X=(E+PZ)MESON/(E+PZ)AVAILABLE GIVES E AND PZ x = RANF(0)IF(RANF(0).LT.CX2) X=1.-X**(1./3.) P(1,3)=(X*W-PMTS/(X*W))/2. P(I,4)=(X*W+PMTS/(X*W))/2. C & IF UNSTABLE, DECAY CHAIN INTO STABLE PARTICLES 120 IPD=IPD+1 IF(K(IPD:2).GE.8) CALL DECAY(IPD:1) IF(IPD.LT.I.AND.I.LE.96) GOTO 120 C 7 FLAVOUR AND PT OF QUARK FORMED IN PAIR WITH ANTIQUARK ABOVE IFL1=IFL2 PX1 = -PX2PY1=-PY2 C 8 IF ENOUGH E+PZ LEFT, GO TO 2 W = (1, -X) * WIF(W.GT.WFIN.AND.I.LE.95) GOTO 100 N = IRETURN END



FROM PARTONS TO PIONS

Here's a fast parton





FROM PARTONS TO PIONS

Here's a fast parton



How about I just call it a hadron?

→ "Local Parton-Hadron Duality"

PARTON → HADRONS?

Early models: "Independent Fragmentation"

Local Parton Hadron Duality (LPHD) can give useful results for inclusive quantities in collinear fragmentation

Motivates a simple model:

"Independent Fragmentation" q



But ...

The point of confinement is that partons are coloured

Hadronisation = the process of **colour neutralisation**

→ Unphysical to think about independent fragmentation of a single parton into hadrons

→ Too naive to see LPHD (inclusive) as a justification for Independent Fragmentation (exclusive)

→ More physics needed

COLOUR NEUTRALISATION

A physical hadronization model

Should involve at least TWO partons, with opposite color charges (e.g., think of them as R and anti-R)*



Strong "confining" field emerges between the two charges when their separation >~ 1fm

*) Really, a colour singlet state $\frac{1}{\sqrt{3}}(|R\bar{R}\rangle + |G\bar{G}\rangle + |B\bar{B}\rangle$



RECAP: COLOUR FLOW

(leading-colour approximation) **Colour flow** in parton showers



Coherence of pQCD cascades → not much "overlap" between systems → Leading-colour approximation pretty good

(LEP measurements in $e^+e^- \rightarrow W^+W^- \rightarrow hadrons$ confirm this (at least to order 10% ~ 1/N_c²))

Note: (much) more color getting kicked around in hadron collisions. More tomorrow.



THE ULTIMATE LIMIT: WAVELENGTHS > 10^{-15} M



FROM PARTONS TO STRINGS

Motivates a model:

Let color field collapse into a narrow flux tube of uniform energy density

 $\kappa \sim 1 \text{ GeV} / \text{fm}$

Limit → Relativistic 1+1 dimensional worldsheet

In "unquenched" QCD

 $g \rightarrow qq \rightarrow$ The strings will break

→ Gaussian suppression of high $m_T^2 = m_q^2 + p_T^2$ Heavier quarks suppressed. Prob(d:u:s:c) $\approx 1 : 1 : 0.2 : 10^{-11}$

> <u>Pedagogical Review:</u> B. Andersson, *The Lund model.* Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol., 1997.







THE (LUND) STRING MODEL Main implementation: PYTHIA. (EPOS also implements a string-based hadronisation model.)

Map:

- Quarks → String Endpoints
- **Gluons** → Transverse Excitations (kinks)
- Physics then in terms of string worldsheet evolving in spacetime
- Probability of string break (by quantum tunneling) constant per unit area → AREA LAW



Simple space-time picture

Details of string breaks more complicated (e.g., baryons, spin multiplets)



FRAGMENTATION FUNCTION



LEFT-RIGHT SYMMETRY



Note: In principle, *a* can be flavour-dependent. In practice, we only distinguish between baryons and mesons



ITERATIVE STRING BREAKS

Causality → May iterate from outside-in

Note: using light-cone coordinates: $p_+ = E + p_z$



On average, expect energy of n^{th} "rank" hadron ~ E_n ~ $<\!z\!>^n\!E_0$



(NOTE ON THE LENGTH OF STRINGS)

In **Spacetime**:

String tension \approx 1 GeV/fm \rightarrow a 5-GeV quark can travel 5 fm before all its kinetic energy is transformed to potential energy in the string.

Then it must start moving the other way (\rightarrow "yo-yo" model of mesons. Note: string breaks \rightarrow several mesons)



1980: string (colour coherence) effect



1980: string (colour coherence) effect



DIFFERENCES BETWEEN QUARK AND GLUON JETS



Can be hugely important for discriminating new-physics signals (decays to quarks vs decays to gluons, vs composition of background and bremsstrahlung combinatorics)

THE CLUSTER MODEL Two main (independent) implementations: HERWIG, SHERPA



Peter Skands



JETS

Think of jets as **projections** that provide a universal view of events



 I'm not going to cover the many different types of jet clustering algorithms (k_T, anti-k_T, C/A, cones, ...) – see e.g., lectures & notes by G. Salam.
 ▶ Focus instead on the physical origin and MC modeling of jets



JETS VS PARTON SHOWERS

Jet clustering algorithms

Map event from low E-resolution scale (i.e., with many partons/hadrons, most of which are soft) to a higher E-resolution scale (with fewer, hard, IR-safe, jets)



Parton shower algorithms

Map a few hard partons to many softer ones

Probabilistic \rightarrow closer to nature.

Not uniquely invertible by any jet algorithm*

(* See "Qjets" for a probabilistic jet algorithm, <u>arXiv:1201.1914</u>) (* See "Sector Showers" for a deterministic shower, <u>arXiv:1109.3608</u>)

INFRARED SAFETY

Definition: an observable is infrared safe if it is <u>insensitive</u> to

SOFT radiation:

Adding any number of infinitely *soft* particles (zero-energy) should not change the value of the observable

COLLINEAR radiation:

Splitting an existing particle up into two *comoving* ones (conserving the total momentum and energy) should not change the value of the observable

Note: some people use the word "infrared" to refer to soft only. Hence you may also hear "infrared and collinear safety". Advice: always be explicit and clear what you mean.



EXAMPLE

Counting the number of particles/tracks is ... ?

The number of tracks, weighted by energy times angle*?

angle*: with respect to some principal axis representing the "collinear" direction (e.g., jet axis or "event-shape" axis)



WHY DO WE CARE?

(example by G. Salam)



Real life does not have infinities, but pert. infinity leaves a real-life trace

$$\alpha_{\rm s}^2 + \alpha_{\rm s}^3 + \alpha_{\rm s}^4 \times \infty \to \alpha_{\rm s}^2 + \alpha_{\rm s}^3 + \alpha_{\rm s}^4 \times \ln p_t / \Lambda \to \alpha_{\rm s}^2 + \underbrace{\alpha_{\rm s}^3 + \alpha_{\rm s}^3}_{\text{BOTH WASTED}}$$

YOU decide how to look at event

The construction of jets is inherently ambiguous

Jet	 Which particles get grouped together? JET ALGORITHM (+ size/resolution parameters)
Definition	2. How will you combine their momenta? RECOMBINATION SCHEME (e.g., 'E' scheme: add 4-momenta)

Ambiguity complicates life, but gives flexibility in one's view of events \rightarrow At what resolution / angular size are you looking for structure(s)? \rightarrow Do you prefer "circular" or "QCD-like" jet areas? (Collinear vs Soft structure) \rightarrow Sequential clustering \rightarrow substructure (veto/enhance?)



TYPES OF ALGORITHMS

1. Sequential Recombination

Take your 4-vectors. Combine the ones that have the lowest 'distance measure'

Different names for different distance measures

Durham k_T : $\Delta R_{ij}^2 \times \min(k_{Ti}^2, k_{Tj}^2)$ Cambridge/Aachen : ΔR_{ij}^2 Anti- k_T : $\Delta R_{ij}^2 / \max(k_{Ti}^2, k_{Tj}^2)$

ArClus (3-2):
$$p_{\perp}^2 = s_{ij}s_{jk}/s_{ijk}$$

$$k_{Ti}^2 = E_i^2 (1 - \cos \theta_{ij})$$
$$\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + \Delta \phi_{ij}^2$$

+ Prescription for how to combine 2 momenta into 1 (or 3 momenta into 2)

→ <u>New set of (n-1) 4-vectors</u>





WHY K_T (OR P_T OR ΔR)?

Attempt to (approximately) capture universal jet-within-jetwitin-jet... behavior

Recall: Approximate full matrix element

$$\frac{|M_{X+1}^{(0)}(s_{i1}, s_{1k}, s)|^2}{|M_X^{(0)}(s)|^2} = 4\pi\alpha_s C_F \left(\frac{2s_{ik}}{s_{i1}s_{1k}} + \dots\right)$$

"Eikonal"

(universal, always there)

by Leading-Log limit of QCD \rightarrow universal dominant terms

 $\frac{\mathrm{d}s_{i1}\mathrm{d}s_{1k}}{s_{i1}s_{1k}} \rightarrow \underbrace{\frac{\mathrm{d}p_{\perp}^{2}}{p_{\perp}^{2}}}_{\mathbf{T}} \underbrace{\frac{\mathrm{d}z}{z(1-z)}}_{\mathbf{T}} \rightarrow \underbrace{\frac{\mathrm{d}E_{1}}{\min(E_{i},E_{1})}}_{\mathbf{H}} \underbrace{\frac{\mathrm{d}\theta_{i1}}{\theta_{i1}}}_{\mathbf{H}} (E_{1} \ll E_{i}, \theta_{i1} \ll 1) , \dots$ $\mathbf{Rewritings in soft/collinear limits}$ "smallest" k_T (or p_T or θ_{ij} , or ...) \rightarrow largest Eikonal (and/or most collinear)



TYPES OF ALGORITHMS

2. "Cone" type

Take your 4-vectors. Select a procedure for which "test cones" to draw

Different names for different procedures

Seeded (obsolete): start from hardest 4-vectors (and possibly combinations thereof, e.g., CDF midpoint algorithm) = "seeds"

Unseeded : smoothly scan over entire event, trying everything

Sum momenta inside test cone \rightarrow new test cone direction

Iterate until stable (test cone direction = momentum sum direction)

Warning: to optimise speed, **seeded** algorithms were sometimes used in the past. INFRARED UNSAFE



(IR SAFE VS UNSAFE OBSERVABLES)

May look pretty similar in experimental environment ... But IR unsafe is not nice to your (perturbative) theory friends ...

- **Unsafe**: badly divergent in pQCD \rightarrow large IR corrections:
 - IR Sensitive Corrections $\propto \alpha_s^n \log^m \left(\frac{Q_{\rm UV}^2}{Q_{\rm ID}^2}\right)$, $m \le 2n$

Even if we have a hadronization model which computes these corrections, the dependence on it is larger \rightarrow uncertainty

Safe \rightarrow IR corrections power suppressed:

IR Safe Corrections $\propto \frac{Q_{\text{IR}}^2}{Q_{\text{UV}}^2}$ Can still be computed (MC) but can also be neglected (pure pQCD)

Let's look at an example ...



ICPR iteration issue

Iterative Cone Progressive Removal





Peter Skands





Collinear splitting can modify the hard jets: ICPR algorithms are collinear unsafe \implies perturbative calculations give ∞



ICPR iteration issue

 $\begin{array}{c} 500 \\ 500 \\ 400 \\ 0 \\ 100 \\ 0 \\ -1 \\ 0 \end{array} \begin{array}{c} -- \text{ cone axis} \\ \hline \text{ cone} \end{array}$

Collinear splitting can modify the hard jets: ICPR algorithms are collinear unsafe \implies perturbative calculations give ∞



ICPR iteration issue

Collinear splitting can modify the hard jets: ICPR algorithms are collinear unsafe \implies perturbative calculations give ∞





ICPR iteration issue

 $\begin{array}{c} 500 \\ 500 \\ 400 \\ 0 \\ 100 \\ 0 \\ -1 \\ 0 \end{array} \begin{array}{c} -- \text{ cone axis} \\ \hline \text{ cone} \end{array}$

Collinear splitting can modify the hard jets: ICPR algorithms are collinear unsafe \implies perturbative calculations give ∞





ICPR iteration issue

 $\begin{array}{c} 500 \\ 500 \\ 400 \\ 300 \\ 100 \\ 0 \end{array} \begin{array}{c} -1 \\ -1 \end{array} \begin{array}{c} 0 \\ 100 \\ 0 \end{array} \begin{array}{c} -1 \\ 100 \\ 100 \end{array} \begin{array}{c} -1 \\ 100 \\ 100 \end{array} \begin{array}{c} -1 \\ 100 \\ 100 \\ 100 \end{array} \begin{array}{c} -1 \\ 100 \\ 100 \\ 100 \end{array} \begin{array}{c} -1 \\ 100 \\ 100 \\ 100 \\ 100 \end{array} \begin{array}{c} -1 \\ 100 \\ 10$

Collinear splitting can modify the hard jets: ICPR algorithms are collinear unsafe \Longrightarrow perturbative calculations give ∞





ICPR iteration issue

Iterative Cone Progressive Removal





ICPR iteration issue

Iterative Cone Progressive Removal





ICPR iteration issue

Iterative Cone Progressive Removal





ICPR iteration issue

Collinear splitting can modify the hard jets: ICPR algorithms are collinear unsafe \implies perturbative calculations give ∞

Peter Skands



Slides from G. Salam

ICPR iteration issue

cone iteration cone axis — 500 cone p_T (GeV/c) Т 400 300 200 100 0 0 1 -1 rapidity

Collinear splitting can modify the hard jets: ICPR algorithms are collinear unsafe \implies perturbative calculations give ∞



ICPR iteration issue

Iterative Cone Progressive Removal





Peter Skands

ICPR iteration issue

Iterative Cone Progressive Removal





ICPR iteration issue

Iterative Cone Progressive Removal





"Seeded Cone Algorithm" **ICPR** iteration issue Start from "hardest" seeds Iterative Cone Progressive Removal cone iteration cone axis — 500 cone p_T (GeV/c) 400 300 200 100 0 0 -1 1 rapidity jet 1

Collinear splitting can modify the hard jets: ICPR algorithms are collinear unsafe \implies perturbative calculations give ∞

ICPR iteration issue

Iterative Cone Progressive Removal





ICPR iteration issue

Iterative Cone Progressive Removal



cone iteration cone axis 500 cone p_T (GeV/c) 400 300 Note: none of the jet 200 algorithms in use at 100 LHC are seeded. But worth 0 -1 0 understanding issue if/ when you consider rapidity jet 1 proposals for new observables jet 2

Collinear splitting can modify the hard jets: ICPR algorithms are collinear unsafe \implies perturbative calculations give ∞

ICPR iteration issue

Iterative Cone Progressive Removal

STEREO VISION

Use IR Safe algorithms

To study short-distance physics

Recombination-type algos → "inverse shower"

 \rightarrow can study jet substructure \rightarrow test shower properties & distinguish BSM?

"Cone-like": SiSCone (unseeded)

"Recombination-like": k_T, Cambridge/Aachen

"Hybrid": **Anti-k**_T (cone-shaped jets from recombination-type algorithm; note: clustering history not ~ shower history)

Use IR Sensitive observables



(e.g., FASTJET)

http://www.fastjet.fr/

E.g., number of tracks, identified particles, ... To explicitly study hadronisation and models of IR physics

 \rightarrow message is not to avoid IR unsafe observables at all costs. But to know when and how to use them.



SUMMARY

Jets: Discovered at SPEAR (SLAC '72) and DORIS (DESY '73): at E_{CM} ~ 5 GeV Collimated sprays of nuclear matter (hadrons). Interpreted as the "fragmentation of fast partons" -> MC generators

PYTHIA (and EPOS): Strings enforce confinement; break up into hadrons Based on linear confinement: V(r) = κr at large distances + Schwinger tunneling Powerful energy-momentum picture, with few free parameters Not very predictive for flavour/spin composition; many free parameters

HERWIG and SHERPA employ 'cluster model'

Based on **universality of cluster mass spectra** + 'preconfinement' Algorithmically simpler; flavour/spin composition largely from hadron masses

NB: many indications that confinement is more complicated in pp

- ~ well understood in "dilute" environments (ee: LEP) ~ vacuum
- LHC is providing a treasure trove of measurements on jet fragmentation, identified particles, minimum-bias, underlying event, ... tomorrow's lecture!



Extra Slides

THE EFFECTS OF HADRONISATION

Generally, expect few-hundred MeV shifts by hadronisation

Corrections to IR safe observables are "power corrections"

 $\propto \Lambda_{\rm QCD}^2/Q_{\rm OBS}^2$

Corrections for jets of radius $R = \Delta \eta \times \Delta \phi$ $\propto 1/R$

See

Korchemsky, Sterman, NPB 437 (1995) 415 Seymour, NPB 513 (1998) 269 Dasgupta, Magnea, Salam, JHEP 0802 (2008) 055

Simple analytical estimate → ~ 0.5 GeV / R correction from hadronisation (scaled by colour factor) hadronisation p_t shift (scaled by R C_F/C)



Significant differences between codes/tunes pt (parton) [GeV] → important to pin down with precise QCD hadronisation measurements at LHC



HIDDEN VALLEYS / EMERGING JETS



HIDDEN VALLEYS / EMERGING JETS



Requirements for a model to produce emerging jet phenomenology:

- Hierarchy between the mediator mass and hidden sector mass.
- Strong coupling in hidden sector \rightarrow large particle multiplicity.
- Macroscopic decay lengths of hidden sector fields back to the visible sector

Schwaller, Stolarski, Weiler JHEP 1505 (2015) 059



R-HADRONS

 \Rightarrow PYTHIA allows for hadronization of 3 generic states:

- colour octet uncharged, like g, giving gud, guud, gg, ...,
- colour triplet charge +2/3, like \tilde{t} , giving $\tilde{t}\overline{u}$, $\tilde{t}ud_0$, ...,
- colour triplet charge -1/3, like $\tilde{\mathrm{b}}$, giving $\tilde{\mathrm{b}}\overline{\mathrm{c}}$, $\tilde{\mathrm{b}}\mathrm{su}_1$, . . .



Most hadronization properties by analogy with normal string fragmentation, but glueball formation new aspect, assumed $\sim 10\%$ of time (or less).

R-hadron interactions with matter: part of detector simulation, i.e. GEANT, not PYTHIA Freight-train BSM particle surrounded by light pion/gluon cloud → little dE/dx + charge flipping ! A.C. Kraan, Eur. Phys. J. C37 (2004) 91; M. Fairbairn et al., Phys. Rep. 438 (2007) 1

