



# Past, Present and Future of the PYTHIA Event Generator

**Torbjörn Sjöstrand**

Department of Astronomy and Theoretical Physics  
Lund University  
Sölvegatan 14A, SE-223 62 Lund, Sweden

Jyväskylä, 16 April 2021

# 1994: First ATLAS/CMS Technical Proposals

## CMS Technical Proposal:

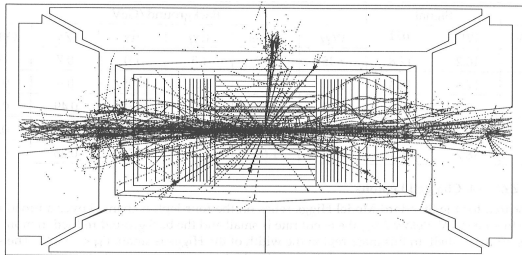


Fig. 12.9: Full GEANT simulation of  $H(150 \text{ GeV}) \rightarrow ZZ^* \rightarrow 2 e^+ 2 e^-$ .

## ATLAS T. P.:

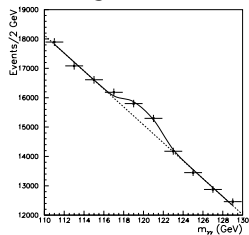


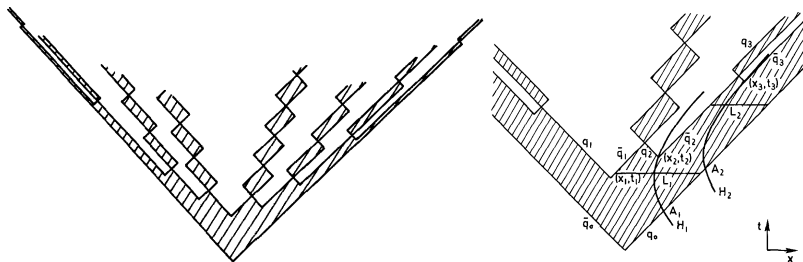
Figure 11.6: Expected  $H \rightarrow \gamma\gamma$  signal for  $m_H = 120 \text{ GeV}$ , combined with the prompt  $\gamma\gamma$  background, assuming an integrated luminosity of  $10^5 \text{ pb}^{-1}$ .

For ATLAS/CMS/LHCb detector design studies in the 1990'ies, PYTHIA was providing input for most GEANT 3 simulations!

How did that come about? What has happened since?

Many of the basic ideas came early, and are “easy” to present. Later additions are very important, but less transparent.

# 1977: Lund studies of hadronization begin



B.Andersson, G. Gustafson, C. Peterson, Z. Physik C1 (1979) 105  
(begun 1977, preprint 1978, published 1979):

- constant string tension  $\kappa \approx 1 \text{ GeV/fm}$
- particle production (approximately) along hyperbola
- lightcone kinematics ( $p^\pm = E \pm p_z$ )
- analytic, recursive procedure from one end
- **no complete systems**
- **$f(z) = 1$  not left-right symmetric**

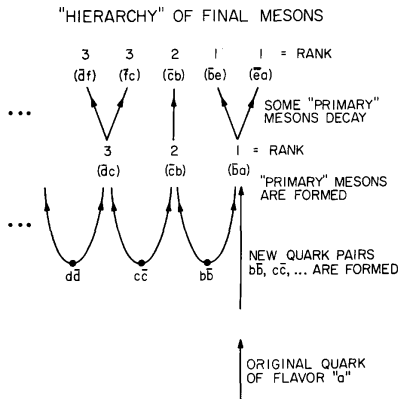
# 1978: The beginning of jet Monte Carlo

R.D. Field and R.P. Feynman,  
A Parametrization of the  
Properties of Quark Jets,  
Nucl. Phys. B136 (1978) 1

- recursive procedure, with
- **Monte Carlo implementation**
- **only one jet**
- **no space-time picture**

starting point for  $e^+e^-$  generators:

- Hoyer et al.
- Ali et al.



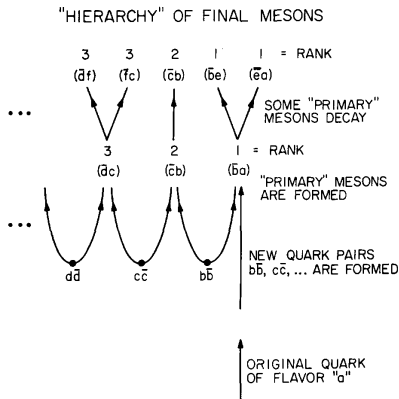
# 1978: The beginning of jet Monte Carlo

R.D. Field and R.P. Feynman,  
A Parametrization of the  
Properties of Quark Jets,  
Nucl. Phys. B136 (1978) 1

- recursive procedure, with
- **Monte Carlo implementation**
- **only one jet**
- **no space-time picture**

starting point for  $e^+e^-$  generators:

- Hoyer et al.
- Ali et al.



Lund: Bengt E.Y. Svensson suggests Monte Carlo implementation of current Lund analytic equations in Field–Feynman spirit, carried out by TS and B. Söderberg

# 1978: JETSET version 1

```

SUBROUTINE JETGEN(N)
COMMON JET/ K(100:2), P(100:5)
COMMON /PAR/ PUD, PSI, SIGMA, CX2, EBEG, WFJN, IFLBEG
COMMON /DATA2/ IVD(1:2), CBN(2), KDP(29:3)
DIMENSION U(1:3), BE(3)
IFLSEN=(1D-IFLBEJ)/5
I=2, EBEG
1=0
LPD=0
C 1 FLAVOUR AND PT FOR FIRST QUARK
IFL1=IABS(IFLBE)
PT1=SIGMA*SQRT(-ALOG(RANF(D)))
PHI1=0.2832*PI*P(1)
PY1=PT1*COS(PHI1)
PX1=PT1*SIN(PHI1)
100 I=1+I
C 2 FLAVOUR AND PT FOR NEXT ANTIQUARK
IFL2=1+INT(RANF(D)/PUD)
PT2=SIGMA*SQRT(-ALOG(RANF(D)))
PHI2=0.2832*PI*P(2)
PX2=PT2*COS(PHI2)
PY2=PT2*SIN(PHI2)
C 3 MESON FORMED, SPIN ADDED AND FLAVOUR MIXED
N(1,1)=MEO(3*(IFL1-1)+IFL2,IFLSEN)
ISPIN=INT(PSI+RANF(D))
K(1,1)=N(1,ISPIN+1(1,1))
IF(K(1,1),L(6) GOTO 150
    TMI=RAMF(D)
    NTK(K(1,1)-6+3*ISPIN
    K(1,2)=8**ISPIN*INT(TMI+CMIX(K(1,1)))+INT(TMI+CMIX(KH(2)
C 4 MESON MASS FROM TABLE, PT FROM CONSTITUENTS
110 P(1)=PMAS(K(1,2))
P(1,2)=P(1)+P2
P(1,3)=P(1)+P2
P(1,4)=P(1)+P2
P(1,5)=P(1)+P2
C 5 RANDOM CHOICE OF I=(E=PI)MESON/(E=PI) AVAILABLE IS E AND PI
IFRANF(D).LT.C12) I=1,XXX(1,3)
P(1,3)=X(N-PRTS(X(N))/2.
P(1,4)=X(N+PRTS(X(N))/2.
C 6 IF UNSTABLE, DECAY CHAIN INTO STABLE PARTICLES
120 I=IPD+1
IF(K(IPD-2),GE,6) CALL DECAY(IPD,I)
IF(3P,D,LT,1.AND,I,LE,96) GOTO 120
C 7 FLAVOUR AND PT OF QUARK FORMED IN PAIR WITH ANTIQUARK ABOVE
IFL1=IFL2
PX1=-PX2
PY1=-PY2
C 8 IF ENOUGH E=PI LEFT, GO TO 2
W(1,1)=X(N)
IF(W(1,6),WFJN,AND,I,LE,95) GOTO 100
1=1
RETURN
END

SUBROUTINE LIST(N)
COMMON JET/ K(100:2), P(100:5)
COMMON /DATA3/ CHA(1:9), CHA2(19), CHA3(2)
WRITE(6,1)
DO 100 I=1,N
IF(K(I,1).GT,0) C1=C-CHA(K(I,1))
IF(K(I,1),LE,0) IC1=K(I,1)
C2=CHA2(K(I,2))
C3=CHA3(K(I,3)-2)/2D)
IF(K(I,1),1.GT,0) WRITE(6,120) I, C1, C2, C3, (P(I,J), J=1:5)
100 IF(K(I,1),LE,0) WRITE(6,130) I, IC1, C2, C3, (P(I,J), J=1:5)
RETURN
110 FORMAT('////',I,' ',I3,' ',O1,' ',T2,' ',PART,' ',T32,' ',STAB',
    &T4,' ',PV,' ',T56,' ',PV,' ',T65,' ',P2,' ',T80,'E',',',T92,'M')
120 FORMAT('10:12:AX:10:13:2:AX:11:12:AX:FS:1)
130 FORMAT('10:12:AX:11:12:2:AX:AA:15:AX:FS:1)
END

SUBROUTINE DECAY(IPD,I)
COMMON JET/ K(100:2), P(100:5)
COMMON /DATA1/ MESO(9:2), CM1(6:2), PMAS(19)
COMMON /DATA2/ IVD(1:2), CBN(2), KDP(29:3)
DIMENSION U(1:3), BE(3)
C 1 DECAY CHANNEL CHOICE, GIVEN DECAY PRODUCTS
TS=RAMF(D)
IDC=IDCD(K(IPD-2))
100 IDC=IDC+1
IF(TB,GT,CBN(IDC)) GOTO 100
NB=(9+M*P(IDC))/2D
DO 150 I=1+I,NM
K(I,1)=I-1P
K(I,2)=ND*P(IDC,1-I)
150 P(I,1)=M*P(CBN(I,2))
C 2 IN THREE-PARTICLE DECAY CHOICE OF INVARIANT MASS OF PRODUCTS 2+3
IF(IND,EQ,2) GOTO 130
SA=(P(IPD,5)+P(1+1,5))**2
SB=(P(IPD,5)-P(1+1,5))**2
SC=(P(1+2,5)+P(1+3,5))**2
SD=(P(1+2,5)-P(1+3,5))**2
TDO=(SA-SB)*(SB-SD)/4,+B*RT(SB*SD)
IF(K(IPD-2),GE,11) TDO=SB*(SB+SC)+TDO**3
120 SX=(SB-SD)*RAMF(D)
TDF=SB*RT((SE+SA)*SE-SB*(SE-SC)+(SX-SB))/SX
IF(K(IPD-2),GE,11) TDF=SB*TDF**3
IF(RANF(D)+TDF,GT,TDO) GOTO 120
P(100,5)=SB*RT(SX)
C 3 TWO-PARTICLE DECAY IN CH: TWICE TO SIMULATE THREE-PARTICLE DECAY
130 DO 160 I=1,3-N(1)
ID=(IL-I)*50D-(IL-2)*ID
I=I+1L
IS=ND-IL-I+1D0-IND-IL-2*(C1+L+5)
PA=SB*RT((P(10,5)**2-(P(1,5)-P(12,5))**2)/C2)+P(10,5)
140 U(1)=C,=RANF(D)
PHI=0.2832*PI*P(1)
P(1,2)=P(1)+RANF(D)
P(1,3)=P(1)+RANF(D)
P(1,4)=P(1)+RANF(D)
TDA=1-(U(1)+P(1,2)+U(2)+P(1,2)+U(3)+P(1,2))**2
IF(1D,I,1)**2*(P(1,2)+P(1,3)+P(1,4))**2
IF(K(IPD-2),GE,11.AND,I,LE,2.AND,RANF(D),ST,TD) GOTO 140
DO 150 I=1,3
P(1,1)=M*P(I)
150 P(1,2)=P(1)+P(1,1)
P(1,3)=P(1)+P(1,1)
P(1,4)=P(1)+P(1,1)
P(1,5)=P(1)+P(1,1)
P(1,2,4)=B*RT(P(1,1)+P(1,5)**2)
C 4 DECAY PRODUCTS LORENTZ BOOST TRANSFERRED TO LAB SYSTEM
DO 180 I=1,N(1-I)
ID=(IL-I)*50D-(IL-2)*ID
DO 190 J=1,3
170 BE(J)=P(ID)*/P(ID,4)
BEP(10,4)=P(10,5)
DO 180 I=1,3+I,N
BEP=BE(1)*P(1,1)+BE(2)*P(1,2)+BE(3)*P(1,3)
DO 190 J=1,3
180 P(1,1)=I(1,1)+J*(1,1)*GA*(1,1)+GA*(1,1)+BEP*(1,1,1)+BEP(J)
190 P(1,4)=GA*(P(1,1,4)+BEP)
I=I+ND
RETURN
END

SUBROUTINE EDIT(N)
COMMON /JET/ K(100:2), P(100:5)
COMMON /EDPAR/ ITHROW, PMIN, PHIN, THETA, PHI, BETA(3)
REAL ROT(3:3), PH
C 5 THREW ABOVE NEUTRAL OR UNSTABLE OR WITH TOO LOW PZ OR P
DO 100 I=1,N
IF(ITHROW,GE,1.AND,K(I,2),GE,8) GOTO 150
IF(ITHROW,GE,2.AND,K(I,2),GE,6) GOTO 150
IF(ITHROW,GE,3.AND,K(I,2),GE,3) GOTO 150
IF(P(1,3),LT,PMIN,OR,P(1,4)**2+P(1,5)**2,LT,PHIN**2) GOTO 150
I(1,1)=IDIM(K(I,1),D)
K(I,2)=M(K(I,1),2)
DO 100 J=1+J,3
100 P(1,1)=J*(1,1)
110 CONTINUE
N(1,1)
C 2 ROTATE TO GIVE JET PRODUCED IN DIRECTION THETA, PHI
IF(THETA,LT,1E-4) GOTO 140
ROT(1,1)=COS(THETA)*COS(PHI)
ROT(1,2)=-SIN(PHI)
ROT(1,3)=SIN(THETA)*COS(PHI)
ROT(2,1)=COS(THETA)*SIN(PHI)
ROT(2,2)=COS(PHI)
ROT(2,3)=SIN(THETA)*SIN(PHI)
ROT(3,1)=-SIN(THETA)
ROT(3,2)=0
ROT(3,3)=COS(THETA)
DO 130 I=1,3
DO 120 J=1+J,3
120 PR(J)=P(I,J)
DO 130 J=1+J,3
130 P(1,1)=ROT(1,1)*PR(J,2)+PR(1,2)+PR(2,1)*PR(J,3)+PR(1,3)
C 3 OVERALL LORENTZ BOOST GIVEN BY BETA AND ROT
140 IF(BETA(1)**2+BETA(2)**2+BETA(3)**2,LT,1E-8) RETURN
SA=1/5B*RT(1-BETA(1)**2+BETA(2)**2+BETA(3)**2)
DO 140 I=1,N
BEP=BE(1)*P(1,1)+BETA(2)*P(1,2)+BETA(3)*P(1,3)
DO 150 J=1+J,3
150 P(1,1)=P(I,1)+GA*(P(1,1,1)+GA)*BEP*(P(1,1,4)+BETA(J,1)
160 P(1,4)=GA*(P(1,1,4)+BEP)
RETURN
END

BLOCK DATA
COMMON /PAR/ PUD, PSI, SIGMA, CX2, EBEG, WFJN, IFLBEG
COMMON /EDPAR/ ITHROW, PMIN, PHIN, THETA, PHI, BETA(3)
COMMON /DATA1/ MESO(9:2), CM1(6:2), PMAS(19)
COMMON /DATA2/ IVD(1:2), CBN(2), KDP(29:3)
COMMON /DATA3/ CHA(1:9), CHA2(19), CHA3(2)
DATA PUD/4., PSI/D.5/, SIGMA/280./, CX2/D.777/,
EBEG/1000./, WFJN/100./, IFLBEG/1/
DATA ITHROW/4., PMIN/D.5/, PHIN/D.5/, THETA,PHI, BETA/50./,
DATA MESO/7.1:3:2:8:5:4:6:9:7:2:4:1:1:1:8:3:5:7/
DATA CM1/2:0:5:1., 2:0:5:1., 2:0:5:1., 2:0:5:1., 2:0:5:1., 2:0:5:1.,
DATA IVD/2., 2:13:5., 2:24:5., 7:24:9:7:12:5., 9:4:8:5:9:7:8.,
8:2:7:5:2:4:8:2:2:2:4:8:9:3:7:7:2., 2:7:8:2., 5:10:16./
DATA IVD(10)/1:6:11:12:13:15:17:19:21:22:25/
DATA CBN/1., 0.381, 0.681, 0.763, 0.769, 1., 0.426, 0.642, 0.755,
0.780, 1., 1., 0.687, 1., 0.687, 1., 0.687, 1., 0.687, 1., 0.687, 1., 1.,
0.899, 0.987, 1., 0.986, 0.937, 0.986, 1./
DATA KDP/1.1:6:2:1:1:2:1:1:2:3:6:1:6:4:6:5:7:6:4:6:5:7:2:2:,
8:5:1:6:2:1:5:1:8:3:1:2:1:3:8:1:7:1:8:1:6:8:1:2:1:5:1:3:8:2:2:,
8:2:3:3:3:5:7:3:5:0:1:0:8:3:8:7:9:1:4:0:8:4:0:8:0/
DATA CHA1/'UB', 'DU', 'US', 'SU', 'DB', 'BD', 'UB', 'BU', 'SU', 'SU',
DATA CHA2/'GAMB', 'PI+', 'PI-', 'KA+', 'KA-', 'K0+', 'K0-', 'KSD+', 'KSD-', 'ETA',
'ETAP', 'RH0+', 'RH0-', 'K*+', 'K*-', 'K*0+', 'K*0-', 'RH0+', 'RH0-', 'ONES' /PHI /
DATA CHA3/' ', 'STAB' /
END

```

≈ 200 punched cards  
Fortran code

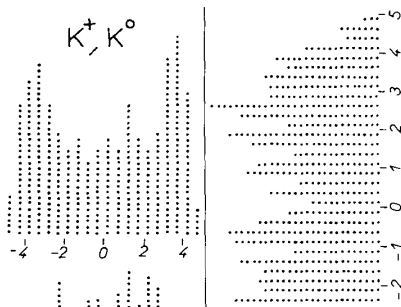
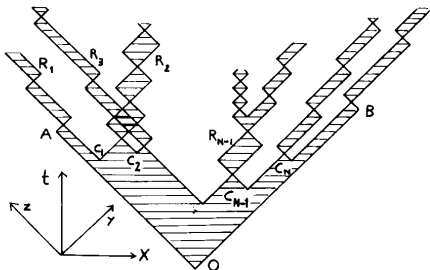
# 1973: The forgotten Artru–Mennessier model

X. Artru and G. Mennessier,  
String Model  
and Multiproduction  
Nucl. Phys.B70 (1974) 93

- exponential decay in area
- complete two-jet system
- Monte Carlo code
- **off-shell hadrons**
- **no transverse d.o.f.**
- **not salesmen**

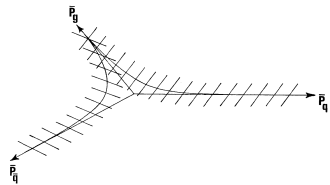
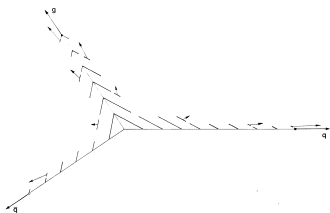
1982: Lund symmetric  
fragmentation function

$$f(z) = \frac{(1-z)^a}{z} \exp\left(-\frac{bm_{\perp}^2}{z}\right)$$



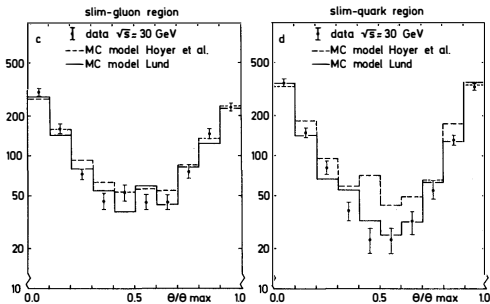
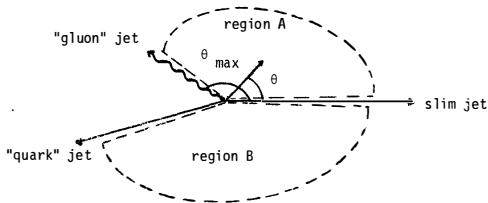
# 1980: The string effect

Lund December 1979



⇒ JADE,  
Moriond, March 1980

... but not by TASSO





# 1982: The strong coupling

## CELLO: The influence of Fragmentation Models in the Determination of the Strong Coupling Constant in $e^+e^-$ Annihilation into Hadrons

Value of  $\alpha_s$  obtained at  $\sqrt{s} = 34$  GeV with the Lund model (LM) and the Hoyer model (HM).  
(first order in QCD)

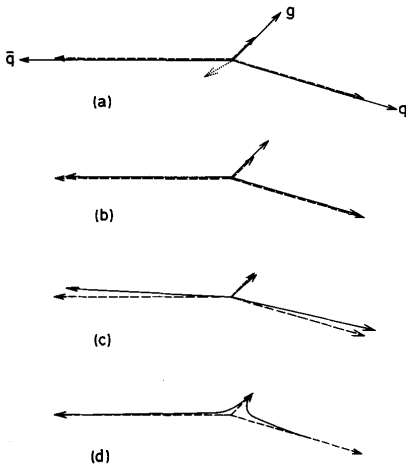
Method	Lund model	Hoyer model	$\frac{\alpha_s(\text{LM})}{\alpha_s(\text{HM})}$
$S \geq 0.25 \ A \leq 0.1$	$0.280 \pm 0.045$	$0.190 \pm 0.030$	1.47
$O \geq 0.20$	$0.260 \pm 0.040$	$0.190 \pm 0.020$	1.37
$O \geq 0.30$	$0.255 \pm 0.050$	$0.200 \pm 0.035$	1.28
# of 3-clusters	$0.235 \pm 0.025$	$0.145 \pm 0.020$	1.62
Cluster Thrust	$0.235 \pm 0.025$	$0.155 \pm 0.015$	1.52
EWAC*	$0.250 \pm 0.040$	$0.150 \pm 0.020$	1.67

The error in the determination of  $\alpha_s$  using the 3-jet fraction (see text) is statistical only (including statistical Monte Carlo error).

\*Energy-weighted angular correlation.

String fragmentation  
increases  $\alpha_s$  by  $\sim 50\%$ !

JETSET 3:  $\sim 1000$  lines



**Fig. 2a-d.** A slightly exaggerated picture of momentum conservation effects. In **a** the momenta of initial partons are full arrows and of jets after fragmentation dashed, with dotted indicating final momentum imbalance: In **b-d** the momenta before conservation are dashed (as in **a**), after full Hoyer rescaling in **b**, Ali boost in **c**, Lund strings (along which particles are sitting) in **d**

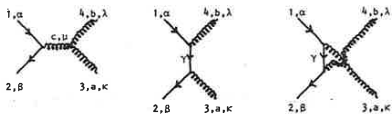
# 1982: The beginning of PYTHIA

LEPTO: colour flow in ep DIS (G. Ingelman & TS)

Compton + High- $p_{\perp}$ : colour flow in pp  
(Hans-Udo Bengtsson)

Process:  $q_i \bar{q}_i \rightarrow gg$

Diagrams:



Colour flows:

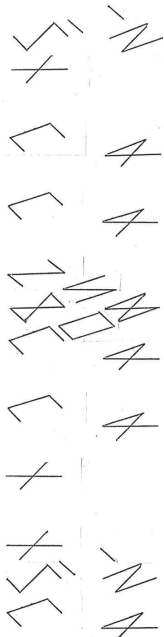


String configurations:



$N_C \rightarrow \infty$  classifies colour topologies

$\Rightarrow$  ~~Cassandra~~  $\Rightarrow$  PYTHIA



# Delphi and Pythia



Delphi: 120 km west of Athens, on the slopes of Mount Parnassus.

Python: giant snake killed by Apollon.

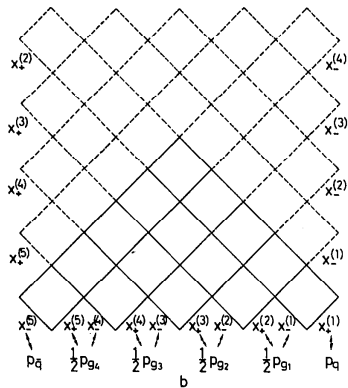
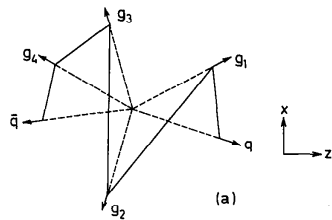
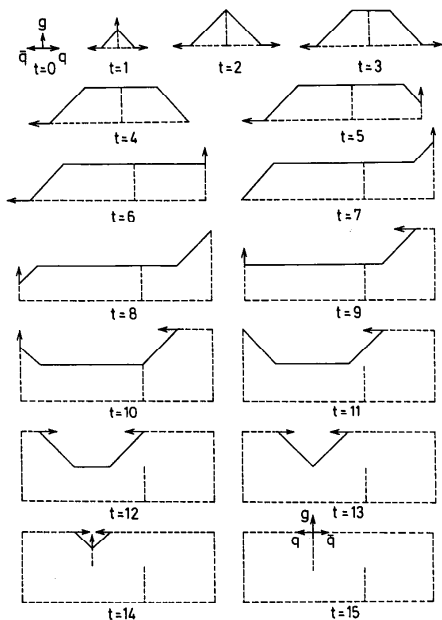
**The Oracle of Delphi:** ca. 1000 B.C. – 390 A.D.

**Pythia:** local prophetess/priestess.

Key role in myths and history, notably in

“The Histories” by Herodotus of Halicarnassus (~482 – 420 B.C.)

# 1983: Complicated string topologies



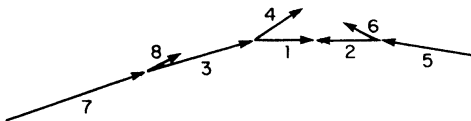
# 1984: Backwards evolution of ISR

Final-state radiation (FSR) intensely studied, two coded up:

- Kajantie–Pieterinen (incoherent) and
- Marchesini–Webber (coherent)

Initial-state radiation (ISR) big hurdle

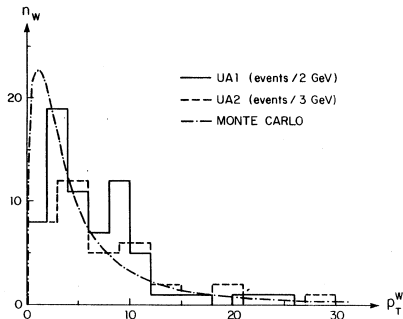
- forward evolution in time and  $Q^2$  may not “hit right”
- backwards evolution reverses order



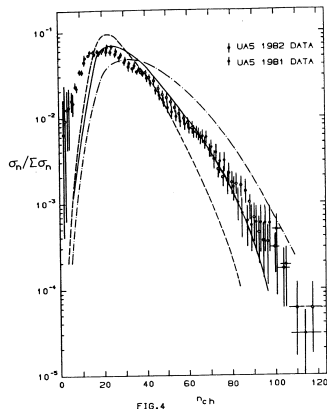
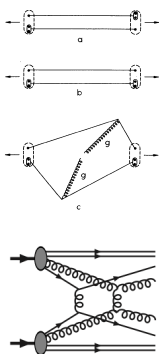
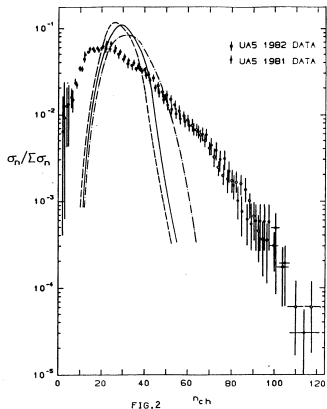
$$dP_b = \frac{df_b(x,t)}{f_b(x,t)} = |dt| \frac{\alpha_s(t)}{2\pi} \sum_a \int \frac{dx'}{x'} \frac{f_a(x',t)}{f_b(x,t)} P_{a \rightarrow bc}(\frac{x}{x'}) \quad (2)$$

This probability exponentiates, so that one may define a form factor

$$S_b(x, t_1; t) = \exp \left\{ - \int_t^{t_1} dt' \frac{\alpha_s(t')}{2\pi} \sum_a \int \frac{dx'}{x'} \frac{f_a(x', t')}{f_b(x, t')} P_{a \rightarrow bc}(\frac{x}{x'}) \right\} \quad (3)$$



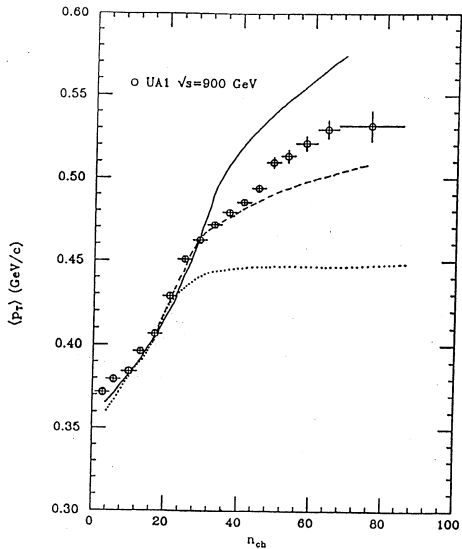
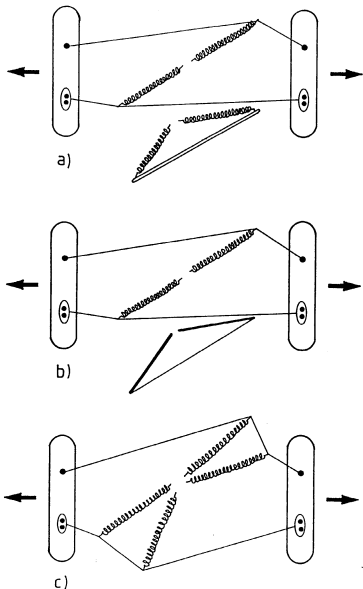
# 1985: Multiparton interactions



without MPI: low- $p_{\perp}$   
 + QCD  $p_{\perp \min} = 1.6$  GeV  
 + ISR+FSR

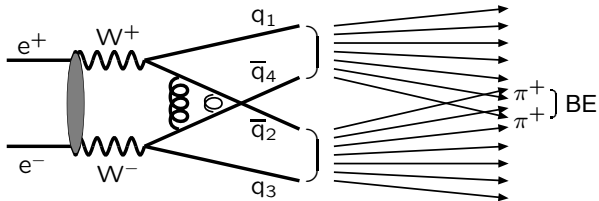
with MPI,  
 $p_{\perp \min} = 2.0, 1.6, 1.2$  GeV

# 1986: Colour reconnection



extremes all or no colour reconnection

# 1996: Colour reconnection in $e^+e^-$ annihilation



At LEP 2 search for effects in  $e^+e^- \rightarrow W^+W^- \rightarrow q_1\bar{q}_2 q_3\bar{q}_4$ :

- **perturbative**  $\langle \delta M_W \rangle \lesssim 5$  MeV : negligible!
- **nonperturbative**  $\langle \delta M_W \rangle \sim 40$  MeV :

**favoured**; no-effect option ruled out at 99.5% CL.

Best description for reconnection in  $\approx 50\%$  of the events.

- **Bose-Einstein**  $\langle \delta M_W \rangle \lesssim 100$  MeV : full effect ruled out (while models with  $\sim 20$  MeV barely acceptable).



# 1986: Dipole showers

Gösta Gustafson: dual description of partonic state:  
partons connected by dipoles  $\Leftrightarrow$  dipoles stretched between partons  
**parton branching**  $\Leftrightarrow$  **dipole splitting**

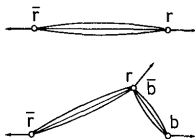
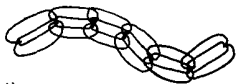


Fig. 3

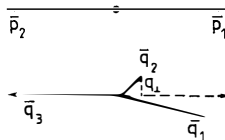


Fig. 4

$p_{\perp}$ -ordered dipole emissions  $\Rightarrow$  **coherence** (cf. angular ordering)

- Originally implemented in **ARIADNE**
- Now basis for three different implementations in **PYTHIA**:  
old simple, **VINCIA** and **DIRE**
- plus showers in **HERWIG**, **SHERPA**, ...

Huge enterprise with many people over many years,  
aiming for increased precision, NLO+NLL and beyond

# 1986: Matrix element corrections

Consider  $e^+e^- \rightarrow \gamma^*/Z^0 \rightarrow q\bar{q} \rightarrow q\bar{q}g$   
with  $d\mathcal{P}_{\text{ME}} = d\sigma_{q\bar{q}g}^{\text{LO}}/\sigma_{q\bar{q}}^{\text{LO}}$

$$\begin{aligned}d\sigma_{q\bar{q}g} &= \sigma_{q\bar{q}}^{\text{NLO}} d\mathcal{P}_{\text{PS}} \exp\left(-\int_{Q^2}^{Q_{\text{max}}^2} d\mathcal{P}_{\text{PS}}\right) \\ &\times \frac{d\mathcal{P}_{\text{ME}}}{d\mathcal{P}_{\text{PS}}} \exp\left(-\int_{Q^2}^{Q_{\text{max}}^2} (d\mathcal{P}_{\text{ME}} - d\mathcal{P}_{\text{PS}})\right) \\ &= \sigma_{q\bar{q}}^{\text{NLO}} d\mathcal{P}_{\text{ME}} \exp\left(-\int_{Q^2}^{Q_{\text{max}}^2} d\mathcal{P}_{\text{ME}}\right)\end{aligned}$$

using the veto algorithm, assuming  $d\mathcal{P}_{\text{PS}} > d\mathcal{P}_{\text{ME}}$  everywhere.

Later extended to (almost) all resonance decays  $a \rightarrow bc \rightarrow b c g$   
and some ISR like  $q\bar{q} \rightarrow \gamma^*/Z^0/W^\pm/\dots$

Rediscovered as the POWHEG method,  
now commonly used for NLO processes.

# 1992: Unified full-length manual

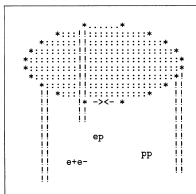
CERN-TH.6488/92

W5035/W5044

## PYTHIA 5.6 and JETSET 7.3 Physics and Manual

Torbjörn Sjöstrand

Theory Division, CERN  
CH-1211 Geneva 23  
Switzerland



- good documentation  
key to early success
- 12 published manuals
- from 1992: steadily updated  
big manual (280 pp)
- PYTHIA 6.4 in JHEP 2006,  
(480 pp →) 580 pp  
> 11,500 citations
- in total > 35,000 citations
- now bulk of documentation  
as xml/html manual
- but big new publication  
in preparation

does not stop a **HUGE**  
amount of mail/questions

CERN-TH.6488/92  
May 1992

# 1996: SPYTHIA

No.	Subprocess	No.	Subprocess	No.	Subprocess	No.	Subprocess	No.	Subprocess	No.	Subprocess
<b>Hard QCD processes:</b>											
11	$f_i f_j \rightarrow f_i f_j$	36	$f_i \gamma \rightarrow f_i W^\pm$	<b>New gauge bosons:</b>				<b>Higgs pairs:</b>			
12	$f_i \bar{f}_i \rightarrow f_i \bar{f}_i$	69	$\gamma \gamma \rightarrow W^+ W^-$	141	$f_i \bar{f}_i \rightarrow \gamma/Z^0/\mathbb{Z}'^0$	297	$f_i \bar{f}_i \rightarrow H^\pm H^0$	146	$e \gamma \rightarrow e^+ e^-$	210	$f_i \bar{f}_j \rightarrow f_i \ell \rho_j^+ +$
13	$f_i \bar{f}_i \rightarrow f_i g$	70	$\gamma W^\pm \rightarrow Z^0 W^\pm$	142	$f_i \bar{f}_i \rightarrow W^\pm Z^0$	298	$f_i \bar{f}_i \rightarrow H^\pm H^0$	147	$dq \rightarrow d^+ Z^0$	211	$f_i \bar{f}_j \rightarrow \tau_i \rho_j^+ +$
28	$f_i g \rightarrow f_i g$	<b>Prompt photons:</b>				144	$f_i \bar{f}_i \rightarrow R$	299	$f_i \bar{f}_i \rightarrow A^0 H^0$	212	$f_i \bar{f}_j \rightarrow \tau_i \nu_j^+ +$
53	$gg \rightarrow f_i \bar{f}_i$	14	$f_i \gamma \rightarrow g \gamma$	<b>Heavy SM Higgs:</b>				300	$f_i \bar{f}_i \rightarrow A^0 H^0$	213	$f_i \bar{f}_j \rightarrow \nu_i \nu_j^+ +$
68	$gg \rightarrow g g$	18	$f_i \bar{f}_i \rightarrow \gamma \gamma$	5	$Z^0 Z^0 \rightarrow h^0$	8	$W^+ W^- \rightarrow h^0$	301	$f_i \bar{f}_i \rightarrow H^+ H^-$	214	$f_i \bar{f}_j \rightarrow \rho_i^+ \rho_j^+ +$
<b>Soft QCD processes:</b>											
91	elastic scattering	29	$f_i g \rightarrow f_i \gamma$	71	$Z_L^0 Z_L^0 \rightarrow Z_L^0 Z_L^0$	145	$q_i \bar{q}_j \rightarrow L Q$	165	$f_i \bar{f}_j (\rightarrow \gamma^*/Z^0) \rightarrow f_i \bar{f}_k$	217	$f_i \bar{f}_i \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0$
92	single diffraction ( $XB$ )	114	$gg \rightarrow \gamma \gamma$	72	$Z_L^0 Z_L^0 \rightarrow W_L^\pm W_L^\mp$	162	$gg \rightarrow \ell L Q$	166	$f_i \bar{f}_j (\rightarrow W^\pm/Z^0) \rightarrow f_i \bar{f}_k$	218	$f_i \bar{f}_i \rightarrow \tilde{\chi}_2^0 \tilde{\chi}_2^0$
93	single diffraction ( $AX$ )	115	$gg \rightarrow g \gamma$	73	$Z_L^0 W_L^\pm \rightarrow Z_L^0 W_L^\pm$	163	$gg \rightarrow L Q L Q$	219	$f_i \bar{f}_i \rightarrow \tilde{\chi}_1^+ \tilde{\chi}_1^-$	219	$f_i \bar{f}_i \rightarrow \tilde{\chi}_3^0 \tilde{\chi}_3^0$
94	double diffraction	99	$\gamma^* \gamma^* \rightarrow q \bar{q}$	76	$W_L^\pm W_L^\pm \rightarrow Z_L^0 Z_L^0$	164	$q_i \bar{q}_j \rightarrow L Q L Q$	<b>Extra Dimensions:</b>			
95	low- $p_\perp$ production	<b>Photon-induced:</b>				77	$W_L^\pm W_L^\pm \rightarrow W_L^\pm W_L^\pm$	391	$f \bar{f} \rightarrow G^*$	221	$f_i \bar{f}_i \rightarrow \tilde{\chi}_1^+ \tilde{\chi}_1^-$
<b>Open heavy flavour: (also fourth generation)</b>											
81	$f_i \bar{f}_i \rightarrow Q_k \bar{Q}_k$	33	$f_i \gamma \rightarrow f_i g$	<b>BSM Neutral Higgs:</b>				392	$gg \rightarrow G^*$	222	$f_i \bar{f}_i \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0$
82	$gg \rightarrow Q_k \bar{Q}_k$	34	$f_i \gamma \rightarrow f_i \gamma$	151	$f_i \bar{f}_i \rightarrow H^0$	191	$f_i \bar{f}_i \rightarrow \rho_{bc}^0$	393	$q \bar{q} \rightarrow g G^*$	223	$f_i \bar{f}_i \rightarrow \tilde{\chi}_2^0 \tilde{\chi}_2^0$
83	$q_i \bar{q}_j \rightarrow Q_k \bar{Q}_k$	54	$g \gamma \rightarrow f_i \bar{f}_i$	152	$gg \rightarrow H^0$	192	$f_i \bar{f}_i \rightarrow \rho_{bc}^+$	394	$qg \rightarrow q G^*$	224	$f_i \bar{f}_i \rightarrow \tilde{\chi}_2^\pm \tilde{\chi}_2^\pm$
84	$g \gamma \rightarrow Q_k \bar{Q}_k$	58	$\gamma \gamma \rightarrow f_i \bar{f}_i$	153	$\gamma \gamma \rightarrow H^0$	193	$f_i \bar{f}_i \rightarrow \omega_{bc}^0$	395	$gg \rightarrow g G^*$	225	$f_i \bar{f}_i \rightarrow \tilde{\chi}_3^\pm \tilde{\chi}_3^\pm$
85	$\gamma \gamma \rightarrow F_k \bar{F}_k$	58	$\gamma \gamma \rightarrow f_i \bar{f}_i$	171	$f_i \bar{f}_i \rightarrow Z^0 H^0$	194	$f_i \bar{f}_i \rightarrow f_k \bar{f}_k$	<b>Left-right symmetry:</b>			
<b>Closed heavy flavour:</b>											
86	$gg \rightarrow J/\psi g$	131	$f_i \gamma \gamma \rightarrow f_i g$	172	$f_i \bar{f}_i \rightarrow W^\pm H^0$	195	$f_i \bar{f}_i \rightarrow f_i \bar{f}_i$	341	$\ell_i \bar{\ell}_j \rightarrow H_{LR}^{\pm\pm}$	226	$f_i \bar{f}_i \rightarrow \tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$
87	$gg \rightarrow \chi_{c0} g$	132	$f_i \gamma \gamma \rightarrow f_i g$	173	$f_i \bar{f}_i \rightarrow f_i f_j H^0$	361	$f_i \bar{f}_i \rightarrow W_L^\pm W_L^\mp$	342	$\ell_i \bar{\ell}_j \rightarrow H_{LR}^{\pm\pm} e^\mp$	227	$f_i \bar{f}_i \rightarrow \tilde{\chi}_2^\pm \tilde{\chi}_2^\mp$
88	$gg \rightarrow \chi_{c1} g$	133	$f_i \gamma \gamma \rightarrow f_i \gamma$	174	$f_i \bar{f}_i \rightarrow f_k f_l H^0$	362	$f_i \bar{f}_i \rightarrow W_L^\pm \pi_{bc}^\pm$	343	$\ell_i^+ \gamma \rightarrow H_{LR}^{\pm\pm} e^+$	228	$f_i \bar{f}_i \rightarrow \tilde{\chi}_1^+ \tilde{\chi}_1^-$
89	$gg \rightarrow \chi_{c2} g$	134	$f_i \gamma \gamma \rightarrow f_i \gamma$	181	$gg \rightarrow Q_k \bar{Q}_k H^0$	363	$f_i \bar{f}_i \rightarrow \pi_{bc}^+ \pi_{bc}^-$	344	$\ell_i^+ \gamma \rightarrow H_{LR}^{\pm\pm} e^+$	229	$f_i \bar{f}_i \rightarrow \tilde{\chi}_2^+ \tilde{\chi}_2^-$
104	$gg \rightarrow \chi_{c0}$	135	$g \gamma \gamma \rightarrow f_i \bar{f}_i$	182	$q_i \bar{q}_j \rightarrow Q_k \bar{Q}_k H^0$	364	$f_i \bar{f}_i \rightarrow \gamma \pi_{bc}^0$	345	$\ell_i^+ \gamma \rightarrow H_{LR}^{\pm\pm} \mu^+$	230	$f_i \bar{f}_i \rightarrow \tilde{\chi}_3^+ \tilde{\chi}_3^-$
105	$gg \rightarrow \chi_{c2}$	136	$g \gamma \gamma \rightarrow f_i \bar{f}_i$	183	$f_i \bar{f}_i \rightarrow g H^0$	365	$f_i \bar{f}_i \rightarrow \gamma \pi_{bc}^0$	346	$\ell_i^+ \gamma \rightarrow H_{LR}^{\pm\pm} \tau^+$	231	$f_i \bar{f}_i \rightarrow \tilde{\chi}_3^0 \tilde{\chi}_3^0$
106	$gg \rightarrow J/\psi \gamma$	137	$\gamma^* \gamma^* \rightarrow f_i \bar{f}_i$	184	$f_i \bar{f}_i \rightarrow f_i H^0$	366	$f_i \bar{f}_i \rightarrow Z^0 \pi_{bc}^0$	347	$\ell_i^+ \gamma \rightarrow H_{LR}^{\pm\pm} \tau^+$	232	$f_i \bar{f}_i \rightarrow \tilde{\chi}_4^+ \tilde{\chi}_4^-$
107	$g \gamma \rightarrow J/\psi g$	138	$\gamma^* \gamma^* \rightarrow f_i \bar{f}_i$	185	$gg \rightarrow g H^0$	367	$f_i \bar{f}_i \rightarrow Z^0 \pi_{bc}^+$	348	$\ell_i^+ \gamma \rightarrow H_{LR}^{\pm\pm} \tau^+$	233	$f_i \bar{f}_i \rightarrow \tilde{\chi}_4^0 \tilde{\chi}_4^0$
108	$\gamma \gamma \rightarrow J/\psi \gamma$	139	$\gamma^* \gamma^* \rightarrow f_i \bar{f}_i$	186	$f_i \bar{f}_i \rightarrow A^0$	368	$f_i \bar{f}_i \rightarrow W^\pm \pi_{bc}^\pm$	349	$f_i \bar{f}_i \rightarrow H_L^\pm + H_R^\mp$	234	$f_i \bar{f}_i \rightarrow \tilde{\chi}_5^+ \tilde{\chi}_5^-$
		140	$q_i \gamma \rightarrow q_i \pi^\pm$	187	$gg \rightarrow A^0$	370	$f_i \bar{f}_i \rightarrow W^\pm Z_L^0$	350	$f_i \bar{f}_i \rightarrow W_L^\pm + H_R^\mp$	235	$f_i \bar{f}_i \rightarrow \tilde{\chi}_6^+ \tilde{\chi}_6^-$
		180	$q_i \gamma \rightarrow q_i \pi^\pm$	158	$\gamma \gamma \rightarrow A^0$	371	$f_i \bar{f}_i \rightarrow W_L^\pm \pi_{bc}^0$	351	$f_i \bar{f}_i \rightarrow f_k f_l H_{LR}^{\pm\pm}$	236	$f_i \bar{f}_i \rightarrow \tilde{\chi}_7^+ \tilde{\chi}_7^-$
				176	$f_i \bar{f}_i \rightarrow Z^0 A^0$	372	$f_i \bar{f}_i \rightarrow \pi_{bc}^\pm Z_L^0$	352	$f_i \bar{f}_i \rightarrow f_k f_l H_{LR}^{\pm\pm}$	237	$f_i \bar{f}_i \rightarrow \tilde{\chi}_8^+ \tilde{\chi}_8^-$
				177	$f_i \bar{f}_i \rightarrow W^\pm A^0$	373	$f_i \bar{f}_i \rightarrow \pi_{bc}^\pm Z_L^0$	353	$f_i \bar{f}_i \rightarrow Z_L^0$	238	$f_i \bar{f}_i \rightarrow \tilde{\chi}_9^+ \tilde{\chi}_9^-$
				178	$f_i \bar{f}_i \rightarrow f_i f_j A^0$	374	$f_i \bar{f}_i \rightarrow \pi_{bc}^\pm Z_L^0$	354	$f_i \bar{f}_i \rightarrow W_H^\pm$	239	$f_i \bar{f}_i \rightarrow \tilde{\chi}_{10}^+ \tilde{\chi}_{10}^-$
				179	$f_i \bar{f}_i \rightarrow f_k f_l A^0$	375	$f_i \bar{f}_i \rightarrow Z^0 \pi_{bc}^\pm$	<b>SUSY:</b>			
				186	$gg \rightarrow Q_k \bar{Q}_k A^0$	376	$f_i \bar{f}_i \rightarrow W^\pm \pi_{bc}^0$	201	$f_i \bar{f}_i \rightarrow e \bar{L} e_L^c$	240	$f_i \bar{f}_i \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0$
				187	$q_i \bar{q}_j \rightarrow Q_k \bar{Q}_k A^0$	377	$f_i \bar{f}_i \rightarrow W^\pm \pi_{bc}^+$	202	$f_i \bar{f}_i \rightarrow e \bar{R} e_R^c$	241	$f_i \bar{f}_i \rightarrow \tilde{\chi}_1^\pm \tilde{\chi}_1^\mp$
				188	$f_i \bar{f}_i \rightarrow g A^0$	378	$f_i \bar{f}_i \rightarrow W^\pm \pi_{bc}^-$	203	$f_i \bar{f}_i \rightarrow e \bar{L} e_L^c$	242	$f_i \bar{f}_i \rightarrow \tilde{\chi}_2^0 \tilde{\chi}_2^0$
				189	$f_i \bar{f}_i \rightarrow f_i A^0$	381	$q_i \bar{q}_j \rightarrow q_i q_j$	204	$f_i \bar{f}_i \rightarrow \tilde{\mu}_L \tilde{\mu}_L^c$	243	$f_i \bar{f}_i \rightarrow \tilde{g} \tilde{g}$
				190	$gg \rightarrow g A^0$	382	$q_i \bar{q}_j \rightarrow q_i \bar{q}_k$	205	$f_i \bar{f}_i \rightarrow \tilde{\mu}_R \tilde{\mu}_R^c$	244	$gg \rightarrow \tilde{g} \tilde{g}$
						383	$q_i \bar{q}_j \rightarrow g q_k$	206	$f_i \bar{f}_i \rightarrow \tilde{\mu}_L \tilde{\mu}_R^c$	245	$gg \rightarrow b \tilde{b}_1$
						384	$f_i \bar{f}_i \rightarrow g g$	207	$f_i \bar{f}_i \rightarrow \tilde{\nu}_i \tilde{\nu}_i^c$	246	$gg \rightarrow b_1 \tilde{b}_2$
						385	$f_i \bar{f}_i \rightarrow g g$	208	$f_i \bar{f}_i \rightarrow \tau_i^+ \tau_i^c$	247	$f_i \bar{f}_i \rightarrow \tilde{q}_i L \tilde{\chi}_1^+$
						386	$gg \rightarrow g g$	209	$f_i \bar{f}_i \rightarrow \tau_i^+ \tau_i^c$	248	$f_i \bar{f}_i \rightarrow \tilde{q}_i L \tilde{\chi}_2^+$
						387	$f_i \bar{f}_i \rightarrow Q_k \bar{Q}_k$			249	$f_i \bar{f}_i \rightarrow \tilde{q}_i R \tilde{\chi}_2^+$
						388	$gg \rightarrow Q_k \bar{Q}_k$				

(Snowmass 1984, 1986; Aachen 1990; ...)

# 1996: Parton-level interfaces

- originally: each generator is an island,  
with hard-coding only feasible for  $2 \rightarrow 2$  and a few  $2 \rightarrow 3$
- 1988: PDG particle codes (1 = d, 2 = u, 11 =  $e^-$ , 21 = g, ...)
- 1989: HEPEVT commonblock for final (LEP) events
- 1996: LEP2 4-fermion generator parton input to JETSET
- (1989  $\rightarrow$ )  $\sim$ 1998: CompHEP
- (1994  $\rightarrow$ )  $\sim$ 2000: MadGraph
- 2001: Les Houches Accord, transfer of event information  
using Fortran commonblocks

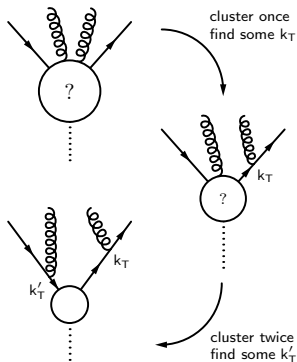
```
INTEGER MAXPUP
PARAMETER (MAXPUP=100)
INTEGER IDBMUP, PDFGUP, PDFSUP, IDWTUP, NPRUP, LPRUP
DOUBLE PRECISION EBMUP, XSECUP, XERRUP, XMAXUP
COMMON/HEPRUP/IDBMUP(2), EBMUP(2), PDFGUP(2), PDFSUP(2),
&IDWTUP, NPRUP, XSECUP(MAXPUP), XERRUP(MAXPUP),
&XMAXUP(MAXPUP), LPRUP(MAXPUP)
```

```
INTEGER MAXNUP
PARAMETER (MAXNUP=500)
INTEGER NUP, IDPRUP, IDUP, ISTUP, MOTHUP, ICOLUP
DOUBLE PRECISION XWGTUP, SCALUP, AQEDUP, AQCDUP, PUP, VTIMUP,
&SPINUP
COMMON/HEPEUP/NUP, IDPRUP, XWGTUP, SCALUP, AQEDUP, AQCDUP,
&IDUP(MAXNUP), ISTUP(MAXNUP), MOTHUP(2, MAXNUP),
&ICOLUP(2, MAXNUP), PUP(5, MAXNUP), VTIMUP(MAXNUP),
&SPINUP(MAXNUP)
```

- 2006: Les Houches Event Files 1.0, ditto, using file format
- several other standards: SLHA, LHAPDF, HepMC, ...

# 2000: Match and Merge

- Match: transition from (one) ME at high  $Q$  to PS at low
- Merge: combine several ME topologies:  $X, X + 1, X + 2, \dots$
- Use shower Sudakovs to provide missing virtual corrections
- Increasingly technical sophistication over 20 years!
- Main research topic of larger event generator community



many methods, several from Lund  
(Leif Lönnblad, Stefan Prestel)

- Match: MC@NLO, POWHEG
- Merge: CKKW, CKKW-L, MLM, FxFx
- M&M: UMEPS, NL<sup>3</sup>, UNLOPS

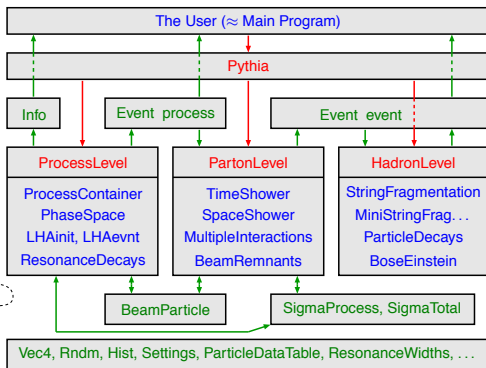
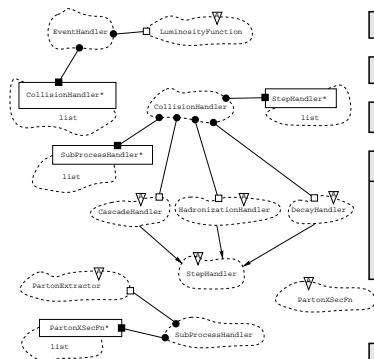
~ 10 alternatives in PYTHIA,  
all rely on LHEF input

# 2004: PYTHIA 8

All early codes written in Fortran 77

1998: PYTHIA 7 in C++, sophisticated platform → ThePEG

2004: PYTHIA 8 in C++, simpler approach but physics focus

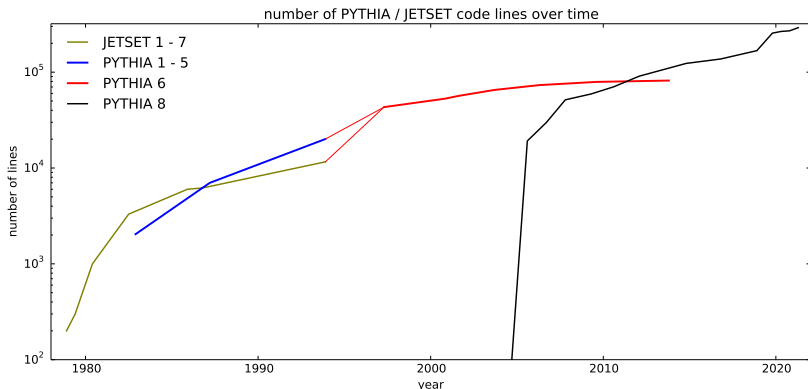


2007: 8.1 first public release

2014: 8.2 some systematization ⇒ minor incompatibility

2019: 8.3 C++98 → C++11, significant internal changes

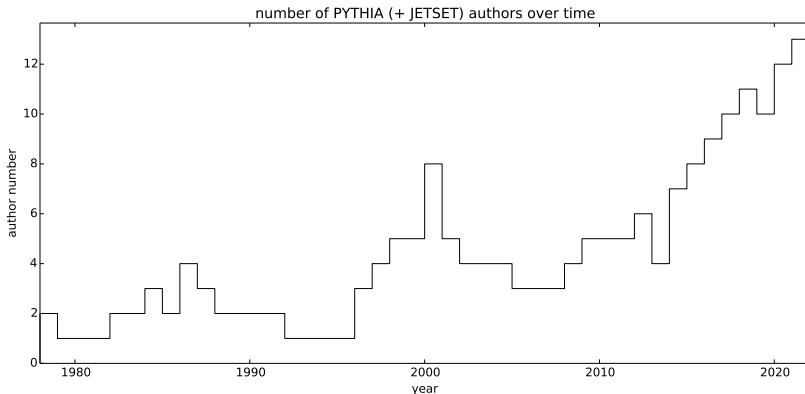
# Code size expansion



- 1997: JETSET fused into PYTHIA
- size includes comment lines and blank lines
- for C++: source, headers, example main programs, but not data (PDF, LHEF), xml/html manual, ME libraries, ...
- **currently ~ 300,000 lines**



# Group size expansion



Does not include:

- non-coding collaborators, like Bo Andersson and Gösta Gustafson
- authors of other “Lund” programs built on top, like LEPTO, ARIADNE, FRITIOF, LDC, DIPSY, POMPYT, ...
- many authors of other non-Lund programs built on top

# Administrative structure

## Current authors:

Christian Bierlich

Nishita Desai

Leif Gellersen

Ilkka Helenius

Philip Ilten

Leif Lönnblad

Stephen Mrenna

Stefan Prestel

Christian Preuss

Torbjörn Sjöstrand

Peter Skands

Marius Utheim

Rob Verheyen

Exploding collaboration size new problem;  
still finding our way.

Main tasks crystallized in recent years,  
notably Philip Ilten as codemaster.

Future organization discussed this week,  
resulting in triumvirate:

- spokesperson: Peter Skands  
(deputy: Ilkka Helenius)
- code master: Philip Ilten  
(deputy: Stephen Mrenna)
- web master: Christian Bierlich

Physics studies based on personal interest,  
so far little to no central planning,  
but now begun discussion of common projects.

# Showers and matching&merging

Strive towards NLO + NLL by improved showers,  
combined with higher-order matrix elements



VINCIA – VIRTual Numerical Collider  
with Interleaved Antennae

Skands, Preuss, Verheyen

- antenna-dipole:  $2 \rightarrow 3$  splittings with both recoiling
- sector shower: unique path to given final state
- full electroweak cascade module



DIRE – DIpole REsummation

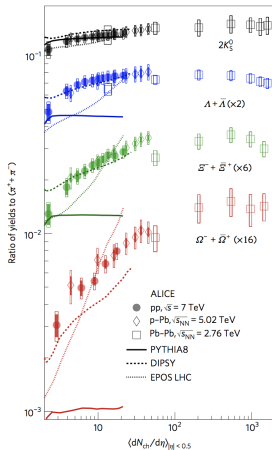
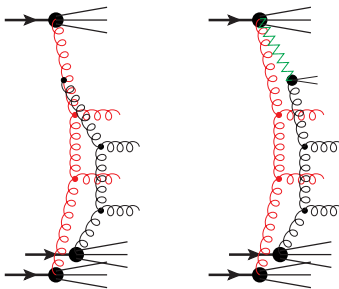
Prestel, Gellersen

- developed jointly with SHERPA
- NLO splitting kernels (negative weights!)
- scale and scheme variations in merging
- Dark Matter emission in shower

# Heavy-ion collisions

Bierlich, Lönnblad (+ Gösta Gustafson, students, postdocs)

- 1984: FRITIOF, successful at low energies, but not for higher
- 2016: ANGANTYR for complete pA and AA collisions
  - full nuclear geometry
  - subdivide collisions into binary ones
  - ropes with higher string tension
  - shove between strings gives flow



# Hadronic rescattering and applications

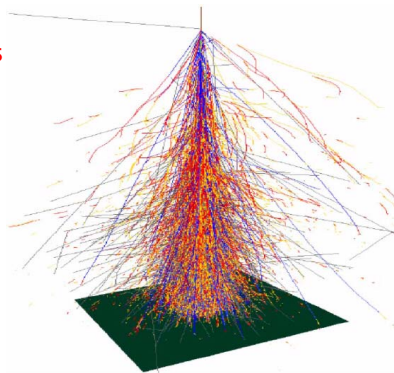
Utthem, TS (Bierlich, Ilten)

- space-time picture of hadronization
- low-energy hadron-hadron collisions
- hadronic rescattering in pp, pA, AA

Future (?):

- formation of pentaquarks etc.
- Bose-Einstein
- extend to arbitrary energies
- component of cosmic ray cascades

(PYTHIA already heavily used  
for cosmic ray production,  
e.g. by Dark Matter annihilation)



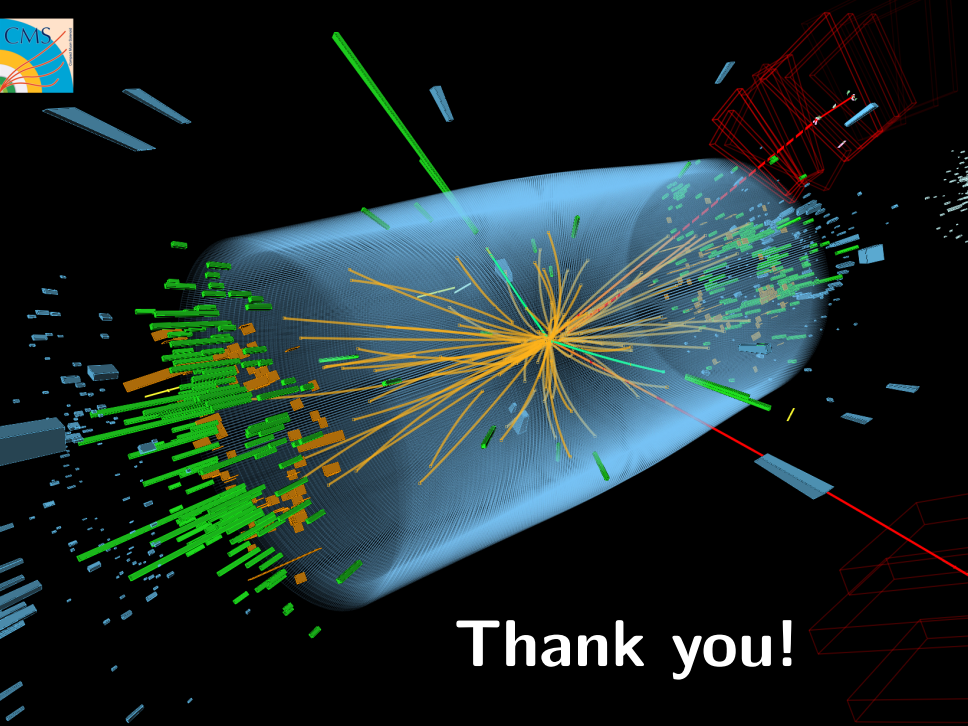
high-energy cosmic ray  
in atmosphere,  
*not* with PYTHIA

- $\gamma p, \gamma\gamma$ , notably in AA collisions;  
UPC = UltraPeripheral Collisions (Helenius)
- DIS and photoproduction transition e.g. at EIC  
(Helenius, Prestel, Bierlich)
- BSM physics, e.g. Dark Matter (Desai, Prestel, Skands, ...)
- bottom/charm/ $\tau$  physics (Ilten)
- FCC and other future accelerators (all)
- Rivet and other common tools (Bierlich, ...)
- code development (all), e.g. parallelization (Utheim, ...)

New topics tend to come along when least you expect it.  
No lack of work to be done!

# Five-year plan?

- Dark showers
- DM annihilation spectra
- New BSM models
- BSM in hadron decay
- EW evolution
- Precision physics
- Become NNLO generator
- Heavy ions
- Photon-ion collisions
- Smooth DIS transition
- Nonperturbative models
- B physics
- QED at hadronic scales
- Cosmic rays
- Improved code structure
- Better interfaces
- New tunes
- Machine learned ME generation
- Native code rather than external links, like PHOTOS
- Parallel processing
- GPU's and other new computing
- Interact with numerous experimental collaborations, old and new



**Thank you!**