Introduction to Event Generators

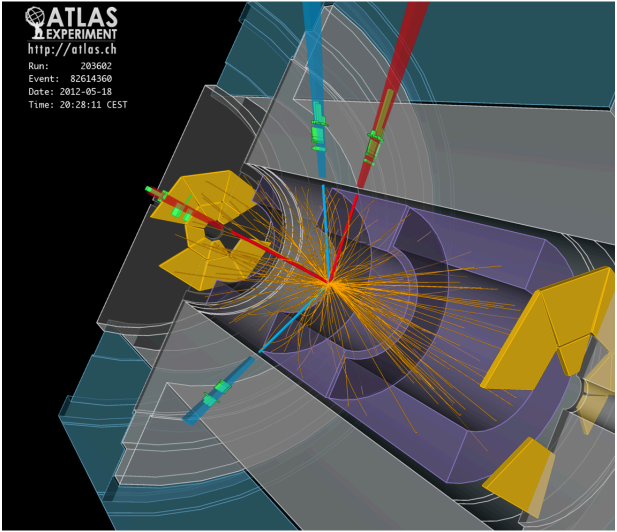
Part 1: Introduction and Monte Carlo Techniques

Torbjörn Sjöstrand

Department of Physics
Lund University

Terascale Monte Carlo School 2024, DESY

Motivation



LHC collision event:

Four leptons clearly visible.

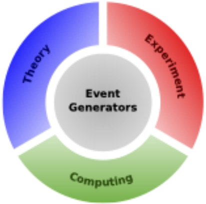
Maybe $\mathrm{H} \to \mathrm{Z}^0\mathrm{Z}^0 \to \mathrm{e}^+\mathrm{e}^-\mu^+\mu^-$.

But what about rest of tracks?

Why and how are they produced?

Course Plan

Event generators: model and understand particle collisions
Complementary to the "textbook" picture of particle physics, since event generators are close to how things work "in real life".



Lecture 1	Introduction to QCD (and the Standard Model)
	Introduction to generators and Monte Carlo techniques
Lecture 2	Parton showers and jet physics
Lecture 3	Multiparton interactions and hadronization

Apologies: PYTHIA-centric,
but most of it generic, or else options will be mentioned

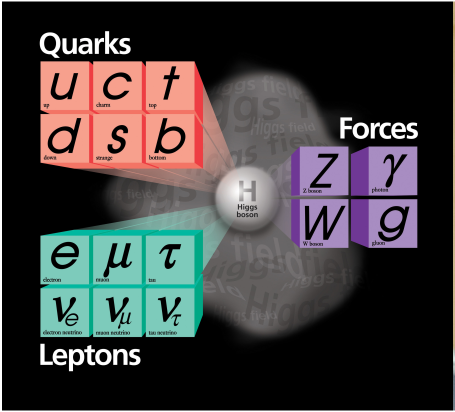
Textbook literature examples

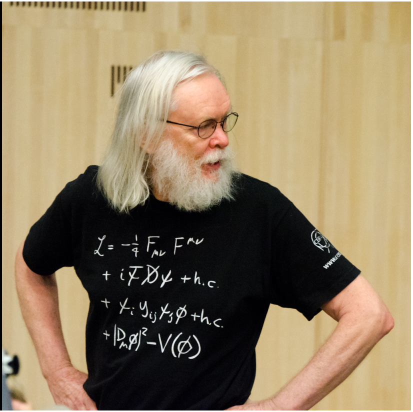
- B.R. Martin and G. Shaw, “Particle Physics”, Wiley (2017, 4th edition)
- G. Kane, “Modern Elementary Particle Physics”, Cambridge University Press (2017, 2nd edition)
- D. Griffiths, “Introduction to Elementary Particles”, Wiley (2008, 2nd edition)
- M. Thomson, “Modern Particle Physics”, Cambridge University Press (2013)
- A. Rubbia, “Phenomenology of Particle Physics”, Cambridge University Press (2022) (1100 pp!)
- P. Skands, “Introduction to QCD”, arXiv:1207.2389 [hep-ph] (v5 2017)
- G. Salam, “Toward Jetography”, arXiv:0906.1833 [hep-ph]

Event generator literature

- A. Buckley et al.,
  "General-purpose event generators for LHC physics",
  Phys. Rep. 504 (2011) 145, arXiv:1101.2599 [hep-ph], 89 pp
- J.M. Campbell et al.,
  "Event Generators for High-Energy Physics Experiments",
  for Snowmass 2021, arXiv:2203.11110 [hep-ph], 153 pp
- C. Bierlich et al., "A comprehensive guide
  to the physics and usage of PYTHIA 8.3",
  accepted by SciPost, arXiv:2203.11601 [hep-ph], 315 pp
- MCnet annual summer schools
  Monte Carlo network from ~ 10 European universities,
  see further `https://www.montecarlonet.org/`,
  with 2024 school at CERN, 10 - 14 June
- Other schools arranged by CTEQ, DESY, CERN, ...

The Standard Model in a nutshell





The Standard Model = “particles” + “interactions”
with well-defined properties and behaviour.

Particles

Particles are spin 1/2 fermions, and

- obey Fermi–Dirac statistics and Pauli exclusion principle,
- can have two spin states, "left" and "right",
- carry unique quantum numbers that are more-or-less well conserved in interactions,
- can be separated into quarks ($\Rightarrow$ hadrons) and leptons,
- come in three generations, distinguished by mass:

	first	second	third
quarks	$\begin{pmatrix} u \\ d \end{pmatrix}$	$\begin{pmatrix} c \\ s \end{pmatrix}$	$\begin{pmatrix} t \\ b \end{pmatrix}$
leptons	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$

- have each an antiparticle with opposite quantum numbers but same mass, and
- can only be created or destroyed in fermion–antifermion pairs.

Interactions

Interactions (= forces) come in different kinds.
In the Standard Model these are

- electromagnetism, QED, mediated by the photon γ,
- weak interactions, mediated by the Z^0, W^+ and W^-,
- strong interactions, QCD, mediated by eight gluons g, and
- mass generation, mediated by Higgs condensate (+ particle).

Among these, only the $W^\pm$ does **not** conserve the number of fermions minus antifermions of each type.
E.g. $u + \overline{d} \to W^+ \to e^+\nu_e$ but **not** $u + \overline{c} \to Z^0 \to e^+\mu^-$.

Gravitation, mediated by gravitons, is not included since
(a) it is too weak for any influence on particle physics processes,
(b) attempts to formulate it as a quantum field theory have failed.

Units and scales

$1\,\mathrm{fm} = 10^{-15}\,\mathrm{m} \approx r_{\mathrm{proton}}$ basic distance scale

$1\,\mathrm{GeV} \approx 1.6 \cdot 10^{-10}\,\mathrm{J} \approx m_{\mathrm{proton}} c^2$ basic energy scale

$c = 1 \approx 3 \cdot 10^{23}\,\mathrm{fm/s}$, so that t in fm, and p and m in GeV

$\hbar = 1 = \hbar c \approx 0.2\,\mathrm{GeV} \cdot \mathrm{fm}$, e.g. to use in $e^{-ipx/\hbar} \to e^{-ipx}$

$1\,\mathrm{mb} = 10^{-31}\,\mathrm{m}^2 \Rightarrow 1\,\mathrm{fm}^2 = 10\,\mathrm{mb}$

$\hbar^2 = (\hbar c)^2 \approx 0.4\,\mathrm{GeV}^2 \cdot \mathrm{mb}$

$N = \sigma \int \mathcal{L}\,\mathrm{d}t$ (“experiment = theory × machine”)

e.g. if $\sigma = 1\,\mathrm{fb} = 10^{-12}\,\mathrm{mb}$,

$\mathcal{L} = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1} = 10^{38}\,\mathrm{m}^{-2}\mathrm{s}^{-1} = 10^{7}\,\mathrm{mb}^{-1}\mathrm{s}^{-1}$,

$T = \int \mathrm{d}t = 24\,\mathrm{hours} \approx 10^5\,\mathrm{s}$,

then $N \approx 10^{-12} \cdot 10^7 \cdot 10^5 = 1$

Lagrangians

Classical Lagrangian $L = T - V = E_{\mathrm{kinetic}} - E_{\mathrm{potential}}$.

Action $S = \int L \, \mathrm{d}t$ should be at minimum, $\delta S = 0$:

$$\frac{\partial L}{\partial q} = \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}} \right) \qquad \text{(Euler – Lagrange)}$$

with q a generalized coordinate and $\dot{q}$ a generalized velocity.

In quantum field theory instead Lagrangian density $\mathcal{L}$:

$$L = \int \mathcal{L} \, \mathrm{d}^3 x \quad \Rightarrow \quad S = \int \mathcal{L} \, \mathrm{d}^4 x \quad \Rightarrow \quad \frac{\partial \mathcal{L}}{\partial \varphi} = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi)} \right)$$

E.g. for a scalar field φ

$$\mathcal{L} = \frac{1}{2} \left(\partial_\mu \varphi \partial^\mu \varphi - m^2 \varphi^2 \right) \quad \Leftrightarrow \quad (\partial^\mu \partial_\mu + m^2)\varphi = 0$$

i.e. the Klein-Gordon equation.

For $\varphi = e^{-ipx}$ this gives $(-p^2 + m^2)\varphi = (-E^2 + \mathbf{p}^2 + m^2)\varphi = 0$.

Electromagnetism

The electromagnetic potential $A^\mu = (V; \mathbf{A})$ gives

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu = \begin{pmatrix} 0 & -E_x & -E_Y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix}$$

The pure QED Lagrangian is

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} = \frac{1}{2}(\mathbf{E}^2 - \mathbf{B}^2)$$

Adding (Dirac four-component) fermion fields ψ_f with charges Q_f

$$\mathcal{L} = \sum_f \overline{\psi}_f \left[\gamma^\mu i\partial_\mu - m_f\right] \psi_f - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - e \sum_f Q_f \overline{\psi}_f \gamma^\mu \psi_f A_\mu$$

where the last term gives the interactions between the fermions and the electromagnetic field.

The Standard Model groups (1)

Examples:

- U(1): group elements $g = e^{i\theta}$ are complex numbers on the unit circle. Abelian.
- SU(n): the set of all complex $n \times n$ matrices M that are unitary ($M^\dagger M = 1$) and have determinant $+1$. Non-Abelian.
- SU(2): has three generators T_j – the Pauli matrices:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

- SU(3) has eight generators T_j – the Gell-Mann matrices:

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{etc}$$

The Standard Model groups (2)

Group elements M can operate on column vectors.
In the fundamental representation these are of dimension n.
For infinitesimal "rotations", where all θ_j are small,

$$M = \exp\left(i\sum_j \theta_j T_j\right) \approx 1 + i\sum_j \theta_j T_j$$

so the interesting transformations are given by the T_j operations, e.g. in SU(2)

$$\sigma_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In the Standard Model the column vectors represent the fermion particles and the T_j generators the interaction mediators.

The Standard Model groups (3)

Standard Model "=" $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ at high energies, which is reduced to $\mathrm{SU}(3)_C \times \mathrm{U}(1)_{em}$ at low energies.

Colour group $\mathrm{SU}(3)_C$: each quark q comes in three "colours", "red", "green" and "blue"

$$\mathrm{q}_r = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathrm{q}_g = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathrm{q}_b = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Eight gluon states can be defined from the Gell-Mann matrices, e.g.

$$g_{r\overline{g}} = \frac{\lambda_1 + i\lambda_2}{2} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

And then matrix multiplication gives that

$$g_{r\overline{g}}\mathrm{q}_g = \mathrm{q}_r$$

The Standard Model Unbroken Lagrangian

At high energies the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ is exact.
Applying our knowledge, its Lagrangian can be written as

$$\mathcal{L} = \sum_f \overline{\psi}_f \gamma^\mu i\mathcal{D}_\mu \psi_f - \frac{1}{4} G^a_{\mu\nu} G^{a\mu\nu} - \frac{1}{4} W^i_{\mu\nu} W^{i\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}$$

$$\mathcal{D}_\mu = \partial_\mu + ig_3 \frac{\lambda^a}{2} G^a_\mu + ig_2 \frac{\sigma^i}{2} W^i_\mu + ig_1 \frac{Y}{2} B_\mu$$

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + g f^{abc} A^b_\mu A^c_\nu$$

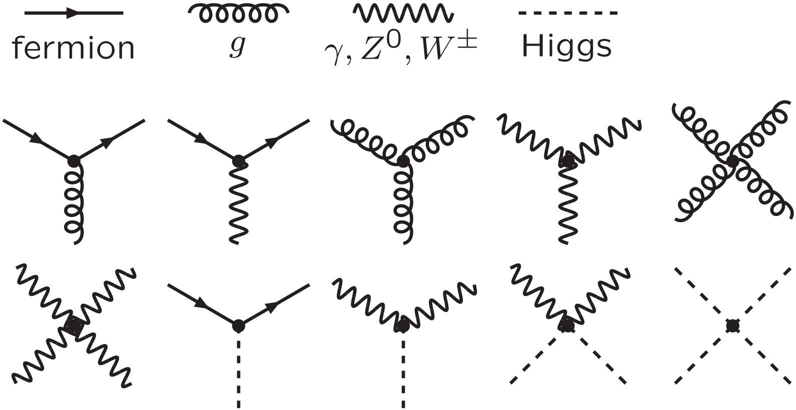
where the G^a only act on quarks,
and the W^i only on the lefthanded fermions.
A represents the potential, F the field tensor and g the coupling of the respective interaction.
The F require an additional third term for non-Abelian groups, where f^{abc} are group constants.
The Higgs mechanism breaks the electroweak part,
but QCD is unaffected, except that quarks gain mass.

Using the Standard Model Lagrangian

- Fermion wave function: $\psi_f(x) = u_f(p)\, e^{-ipx}$.
  $u_f(p)$ destroys a fermion f or creates an antifermion $\overline{f}$,
  $\overline{u}_f(p)$ creates a fermion f or destroys an antifermion $\overline{f}$,
  where $u_f(p)$ and $\overline{u}_f(p)$ are represented by Dirac spinors.
- Vector boson wave function: $A^\mu(x) = \epsilon^\mu(p)\, e^{-ipx}$,
  where ϵ^μ is a polarization vector;
  can create or destroy depending on context.
- Scalar boson wave function: $\phi(x) = 1\, e^{-ipx}$;
  can create or destroy.
- Bilinear field combinations describe propagation of "free" particles, e.g. $\overline{\psi}_f \gamma^\mu i\partial_\mu \psi_f$.
- Trilinear field combinations describe triple vertices, e.g. $\overline{\psi}_f \gamma^\mu e Q_f A_\mu \psi_f$.
- Tetralinear field combinations describe quartic vertices.

Spin handling major complicating factor!

Particle lines and vertices

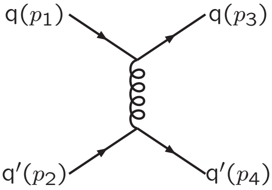


Some $\gamma/Z^0/W^\pm$ combinations not allowed, e.g. $\gamma\gamma\gamma$ or $\gamma\gamma H$.
Quantum number preservation, notably colour and charge.
Arbitrary time order, with fermion in $\equiv$ antifermion out.

Feynman diagrams

A Feynman graph is a useful pictorial representation of a process. It can be converted into a matrix element $\mathcal{M}$, ≈ an amplitude, by combining

- incoming and outgoing wave function normalizations,
- internally exchanged particle "propagators", and
- vertex coupling strengths.



Neglecting spin:

$$\mathcal{M} \sim (\overline{u}_q(p_3)u_q(p_1))(\overline{u}_{q'}(p_4)u_{q'}(p_2))\,\frac{1}{p_g^2}\,g_3^2$$

$$\sim (2E_q)(2E_{q'})\,\frac{1}{p_g^2}\,g_3^2 = g_3^2\,\frac{\hat{s}}{\hat{t}}$$

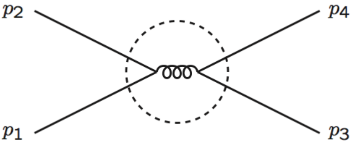
$$\hat{s} = (p_1 + p_2)^2 = (p_3 + p_4)^2$$

$$\hat{t} = (p_1 - p_3)^2 = (p_2 - p_4)^2$$

The basic QCD processes

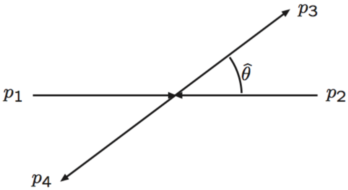
Mandelstam variables

$$\hat{s} = (p_1 + p_2)^2 = (p_3 + p_4)^2$$
$$\hat{t} = (p_1 - p_3)^2 = (p_2 - p_4)^2$$
$$\hat{u} = (p_1 - p_4)^2 = (p_2 - p_3)^2$$



In rest frame, massless limit: $m_1 = m_2 = m_3 = m_4 = 0$

$$\hat{s} = E_{\mathrm{CM}}^2$$
$$\hat{t} = -\frac{\hat{s}}{2}(1 - \cos\hat{\theta}) \approx -p_{\perp}^2$$
$$\hat{u} = -\frac{\hat{s}}{2}(1 + \cos\hat{\theta})$$
$$\hat{s} + \hat{t} + \hat{u} = 0$$



Six basic $2 \to 2$ QCD processes:

$\mathrm{qq}' \to \mathrm{qq}'$ $\quad \mathrm{q\overline{q}} \to \mathrm{q'\overline{q}'}$ $\quad \mathrm{q\overline{q}} \to \mathrm{gg}$

$\mathrm{qg} \to \mathrm{qg}$ $\quad \mathrm{gg} \to \mathrm{q\overline{q}}$ $\quad \mathrm{gg} \to \mathrm{gg}$

Cross sections

Consider subprocess $a + b \to 1 + 2 + \ldots + n$.
If $m_a^2, m_b^2 \ll \hat{s} = (p_a + p_b)^2$ then

$$\mathrm{d}\hat{\sigma} = \frac{|\mathcal{M}|^2}{2\hat{s}} \, \mathrm{d}\Phi_n$$

$$\mathrm{d}\Phi_n = (2\pi)^4 \, \delta^{(4)} \left(p_a + p_b - \sum_{i=1}^{n} p_i \right) \prod_{i=1}^{n} \frac{\mathrm{d}^3 p_i}{(2\pi)^3 2E_i}$$

$$\mathrm{d}\Phi_2 = \frac{\mathrm{d}\hat{t}}{8\pi\hat{s}}$$

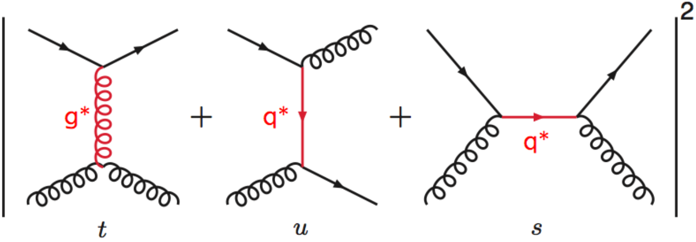
so for process $qq' \to qq'$ on preceding page

$$\mathrm{d}\hat{\sigma} \approx \left(g_3^2 \frac{\hat{s}}{\hat{t}} \right)^2 \frac{1}{2\hat{s}} \frac{\mathrm{d}\hat{t}}{8\pi\hat{s}} = \pi \left(\frac{g_3^2}{4\pi} \right)^2 \frac{\mathrm{d}\hat{t}}{\hat{t}^2} = \pi\alpha_s^2 \frac{\mathrm{d}\hat{t}}{\hat{t}^2}$$

$$\propto \frac{\mathrm{d}\cos(\hat{\theta})}{\sin^4(\hat{\theta}/2)} \ \ \text{(Rutherford scattering)} \propto \frac{\mathrm{d}p_\perp^2}{p_\perp^4}$$

Closeup: $\mathrm{qg} \to \mathrm{qg}$

Consider $\mathrm{q}(1)\,\mathrm{g}(2) \to \mathrm{q}(3)\,\mathrm{g}(4)$:

$|\mathcal{M}|^2 =$ 

$t: p_{\mathrm{g}^*} = p_1 - p_3 \Rightarrow m^2_{\mathrm{g}^*} = (p_1 - p_3)^2 = \hat{t} \Rightarrow \mathrm{d}\hat{\sigma}/\mathrm{d}\hat{t} \sim 1/\hat{t}^2$

$u: p_{\mathrm{q}^*} = p_1 - p_4 \Rightarrow m^2_{\mathrm{q}^*} = (p_1 - p_4)^2 = \hat{u} \Rightarrow \mathrm{d}\hat{\sigma}/\mathrm{d}\hat{t} \sim -1/\hat{s}\hat{u}$

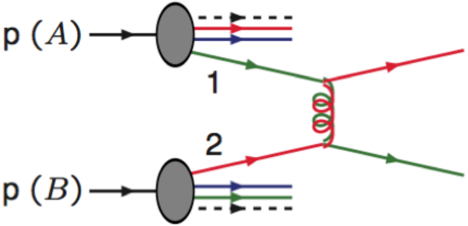
$s: p_{\mathrm{q}^*} = p_1 + p_2 \Rightarrow m^2_{\mathrm{q}^*} = (p_1 + p_2)^2 = \hat{s} \Rightarrow \mathrm{d}\hat{\sigma}/\mathrm{d}\hat{t} \sim 1/\hat{s}^2$

Contribution of each sub-graph is gauge-dependent,
only sum is well-defined:

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}\hat{t}} = \frac{\pi\alpha_{\mathrm{s}}^2}{\hat{s}^2}\left[\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} + \frac{4}{9}\frac{\hat{s}}{(-\hat{u})} + \frac{4}{9}\frac{(-\hat{u})}{\hat{s}}\right]$$

Composite beams

In reality all beams are composite:
$p : q, g, \overline{q}, \ldots$
$e^- : e^-, \gamma, e^+, \ldots$
$\gamma : e^\pm, q, \overline{q}, g$



Factorization

$$\sigma^{AB} = \sum_{i,j} \iint \mathrm{d}x_1 \, \mathrm{d}x_2 \, f_i^{(A)}(x_1, Q^2) \, f_j^{(B)}(x_2, Q^2) \int \mathrm{d}\hat{\sigma}_{ij}$$

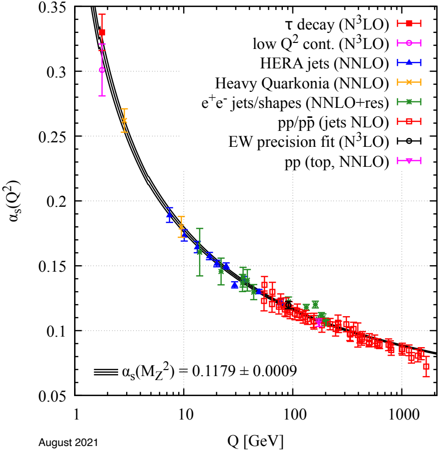
x: momentum fraction, e.g. $p_i = x_1 p_A; p_j = x_2 p_B$
Q^2: factorization scale, "typical momentum transfer scale"

Factorization only proven for a few cases, like γ^*/Z^0 prodution, and strictly speaking not correct e.g. for jet production,

but good first approximation and unsurpassed physics insight.

Couplings

Divergences in higher-order calculations ⇒ renormalization
⇒ couplings run, i.e. depend on energy scale of process.
Small effect for α_{em} (and α_1, α_2, $\sin^2\theta_W$), but big for $\alpha_s = \alpha_3$.



Small Q:
large α_s,
"infrared slavery"
= "confinement",
perturbation theory fails

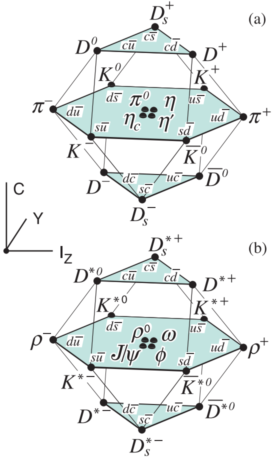
Large Q:
small α_s,
"asymptotic freedom",
perturbation theory applicable

Also quark masses run!

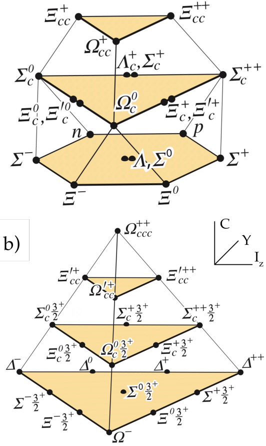
Hadrons

Confinement: no free quarks or gluons,
but bound in colour singlets — hadrons.

mesons: $q\bar{q}$



baryons: qqq



Examples mesons:

$\pi^+ = u\bar{d}$

$\pi^0 = (u\bar{u} - d\bar{d})/\sqrt{2}$

$\pi^- = d\bar{u}$

$K^+ = u\bar{s}$

Exampels baryons:

$p = uud$

$n = udd$

$\Lambda^0 = sud$

$\Omega^- = sss$

+ spin, orbital and radial excitations.

QCD scales

Renormalization group equations $\Rightarrow$

$$\alpha_{\mathrm{S}}(Q^2) = \frac{12\pi}{(33 - 2n_f)\ln(Q^2/\Lambda^2_{\mathrm{QCD}})} + \cdots$$

where n_f is the number of quarks with $m_{\mathrm{q}} < Q$, usually 5.
α_{S} continuous at flavour thresholds $\Rightarrow \Lambda_{\mathrm{QCD}} \to \Lambda^{(n_f)}_{\mathrm{QCD}}$.

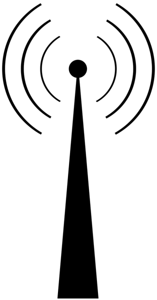
Confinement scale $\Lambda_{\mathrm{QCD}} \approx 0.2\,\mathrm{GeV}$; $\alpha_{\mathrm{S}}(\Lambda_{\mathrm{QCD}}) = \infty$

$1/\Lambda_{\mathrm{QCD}} \approx 0.2\,\mathrm{GeV} \cdot \mathrm{fm}/0.2\,\mathrm{GeV} = 1\,\mathrm{fm}$

hard QCD: $Q \gg \Lambda_{\mathrm{QCD}}$ such that $\alpha_{\mathrm{S}}(Q) \ll 1$; say $Q \geq 10\,\mathrm{GeV}$

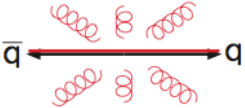
soft QCD: $Q \leq \Lambda_{\mathrm{QCD}}$; in reality $Q \leq 2\,\mathrm{GeV}$

Higher orders and parton showers



In QED, accelerated charges give rise to radiation;
this is the principle of a radio transmitter!
Also for deceleration: **bremsstrahlung**.

Dipole in QCD:



The more violent the acceleration/deceleration,
the higher frequencies/energies can be emitted.

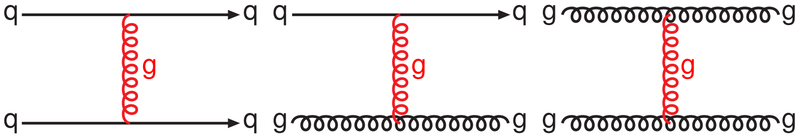
Track emission process as repeated branchings,
where each can take a non-negligible energy fraction.

QED: $\mathrm{f} \to \mathrm{f}\gamma, \gamma \to \mathrm{f}\overline{\mathrm{f}}$ (f any charged fermion)
QCD: $\mathrm{q} \to \mathrm{qg}, \mathrm{g} \to \mathrm{q}\overline{\mathrm{q}}$, $\mathrm{g} \to \mathrm{gg}$ (q any quark)

Matrix element: exact as method, but limited by complexity.
Parton showers: approximation to construct "complete" events.
Match & merge: combine the best of the two.

Multiparton interactions (MPIs)

In pp collisions t-channel exchange of gluons dominate:

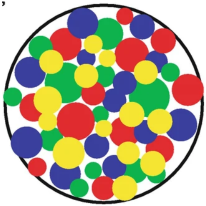


Diverges like $\mathrm{d}p_\perp^2/p_\perp^4$, also with PDF included.

At LHC, with $p_\perp > 5$ GeV, $\sigma_{2\to2} \approx 100\,\mathrm{mb} \approx \sigma_{\mathrm{total}}$

(cf. $\sigma_{\mathrm{total}} \sim \pi(2r_{\mathrm{p}})^2 \approx \pi(2 \cdot 0.85\,\mathrm{fm})^2 \approx 9\,\mathrm{fm}^2 = 90\,\mathrm{mb}$).

Implies multiple $2 \to 2$ processes: **multiparton interactions**.



Naively $p_{\perp min} \sim 1/r_{\mathrm{p}} \sim \Lambda_{\mathrm{QCD}}$,
but more relevant is typical separation
between colour and anticolour,
which if $r_{\mathrm{sep}} \sim r_{\mathrm{p}}/10$ implies
$p_{\perp min} \sim 2\,\mathrm{GeV}$, a better data fit.

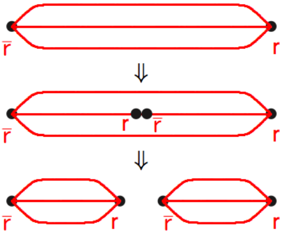
Hadronization

QCD does not allow free colour charges!

In the decay of a colour singlet, say $(\mathrm{e}^+\mathrm{e}^-) \to \mathrm{Z}^0 \to \mathrm{q}\overline{\mathrm{q}}$, the q and $\overline{\mathrm{q}}$ move apart but remain connected by a "string".

Can be viewed as an elongated hadron with radius $r_{\mathrm{string}} \approx r_{\mathrm{p}}$ ($\times\sqrt{2/3}$ since 3 $\to$ 2 dimensions).

Pulling out string costs energy: string tension $\kappa \approx 1$ GeV/fm.



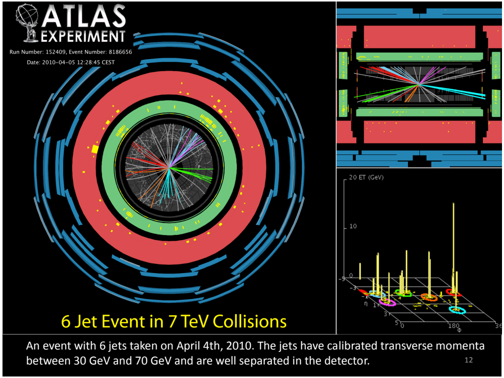
String fragmentation: a new $\mathrm{q}'\overline{\mathrm{q}}'$ pair is created inside the field between the original $\mathrm{q}\overline{\mathrm{q}}$ one, with colours screening these endpoints. Thus the big string breaks into two smaller ones.

This can be repeated to give a sequence of "small" strings $\approx$ hadrons.

In sum: each quark remains confined during string fragmentation, but the partner will change.

Jets

A jet: a spray of hadrons moving out in $\sim$ the same direction.



No unique definition, but "in the eye of the beholder".

At the LHC most commonly found in the $(\eta, \varphi, E_{\perp})$ space with the anti-$k_{\perp}$ algorithm.

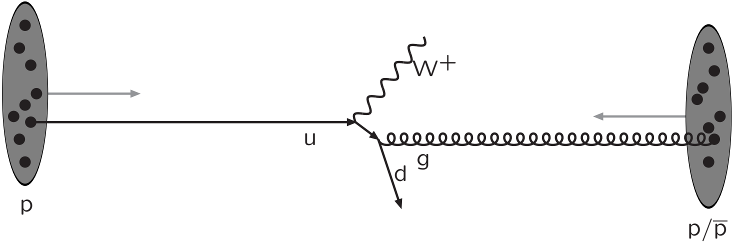
Naively a jet is associated with an outgoing quark or gluon of the hard process, but modified by ISR, FSR, MPI, hadronization.

The structure of an event – 1

Warning: schematic only, everything simplified, nothing to scale, . . .

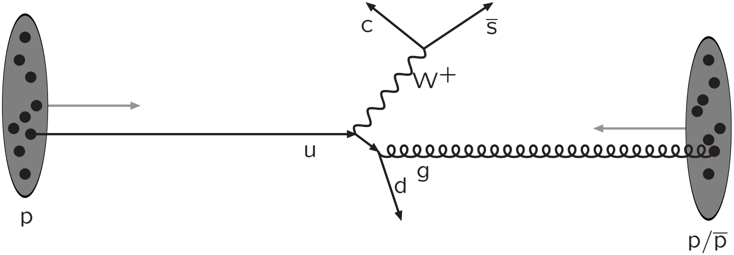
Incoming beams: parton densities

The structure of an event – 2



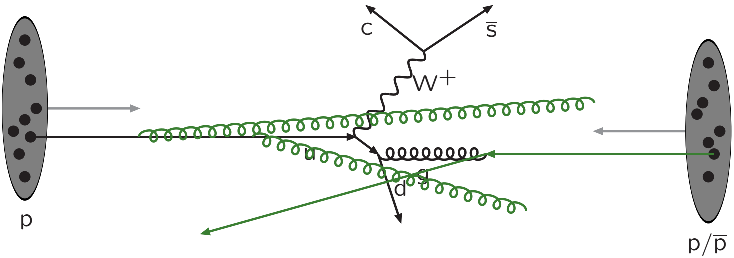
Hard subprocess: described by matrix elements

The structure of an event – 3



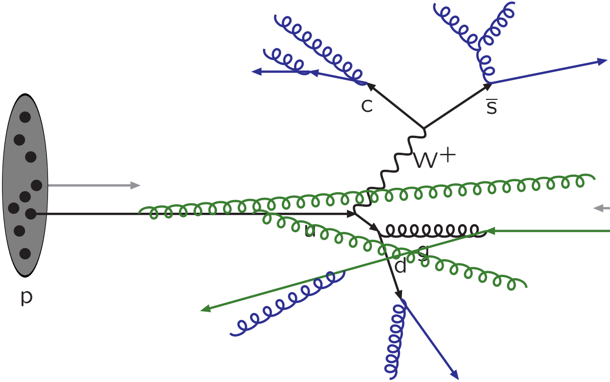
Resonance decays: correlated with hard subprocess

The structure of an event – 4



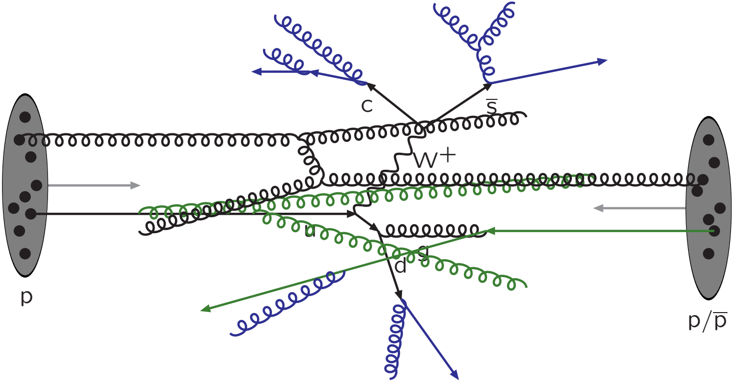
Initial-state radiation: spacelike parton showers

The structure of an event – 5



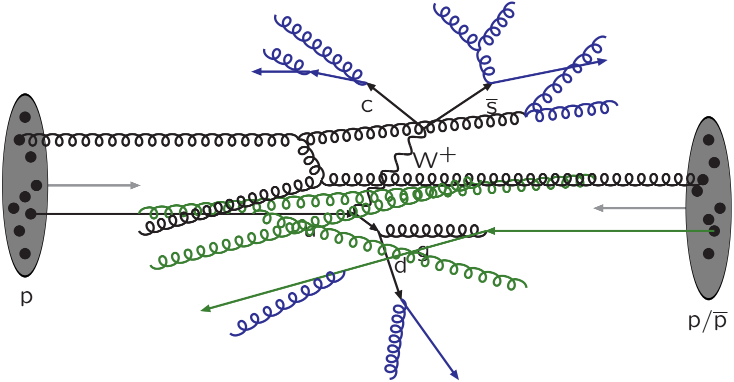
Final-state radiation: timelike parton showers

The structure of an event – 6



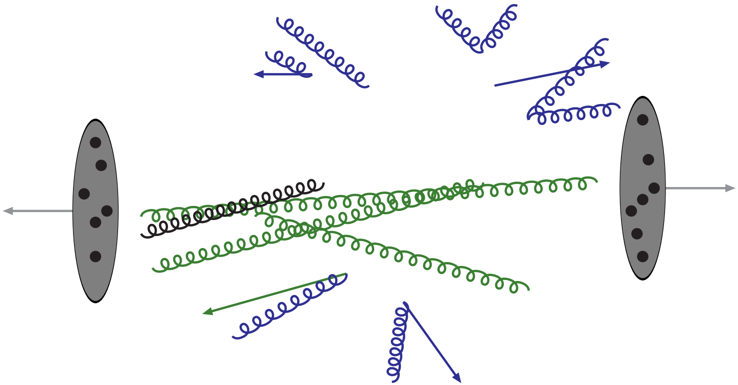
Multiple parton–parton interactions . . .

The structure of an event – 7



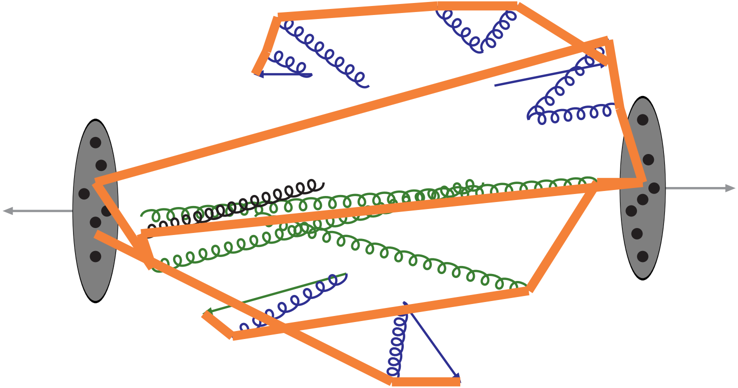
. . . with its initial- and final-state radiation

The structure of an event – 8



Beam remnants and other outgoing partons

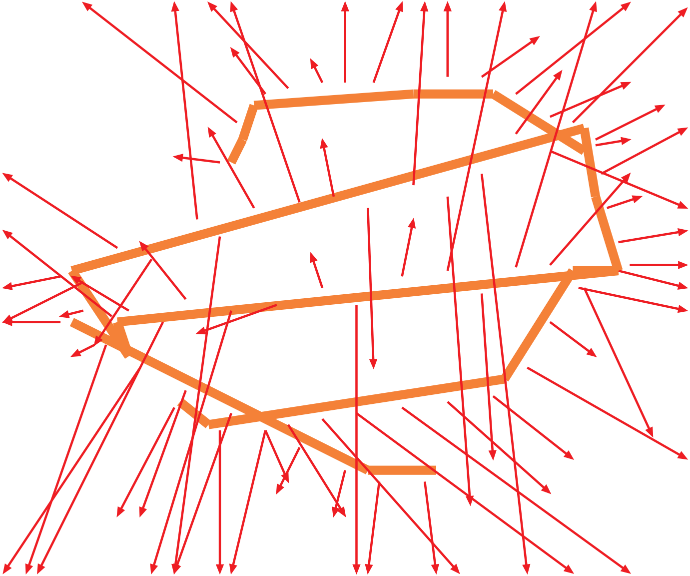
The structure of an event – 9



Everything is connected by colour confinement strings

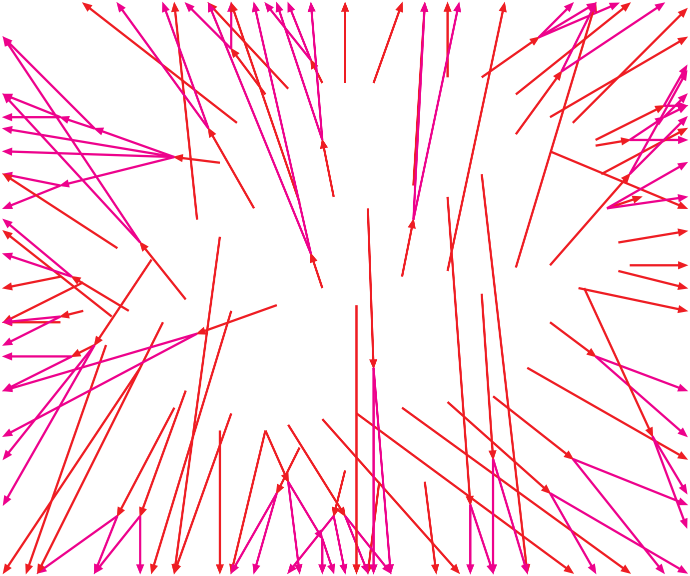
Recall! Not to scale: strings are of hadronic widths

The structure of an event – 10



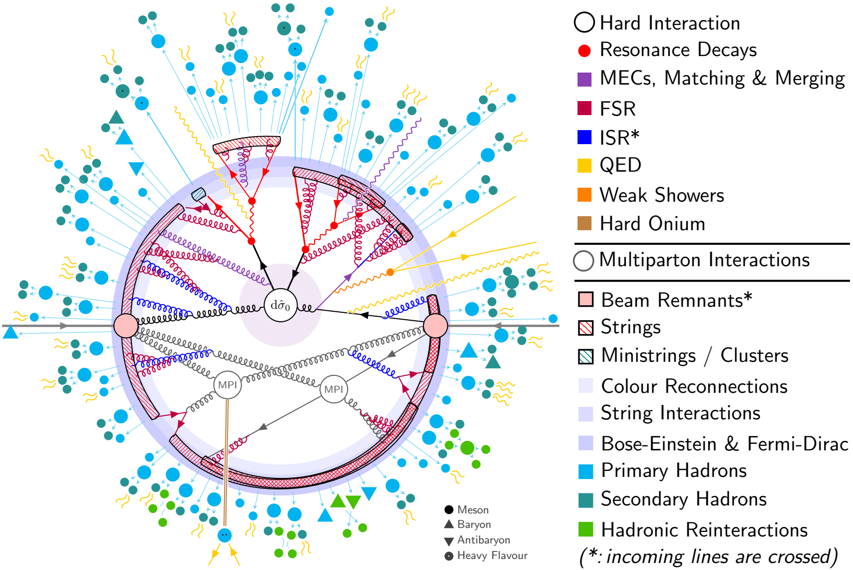
The strings fragment to produce primary hadrons

The structure of an event – 11

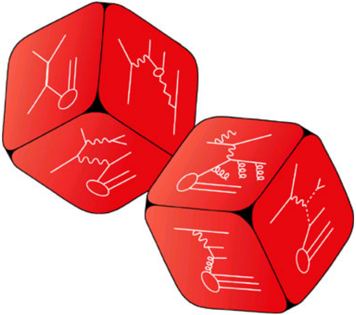


Many hadrons are unstable and decay further

A collected event view



A tour to Monte Carlo



. . . because Einstein was wrong: God does throw dice!
Quantum mechanics: amplitudes $\Longrightarrow$ probabilities
Anything that possibly can happen, will! (but more or less often)

Event generators: trace evolution of event structure.
Random numbers $\approx$ quantum mechanical choices.

The Monte Carlo method

Want to generate events in as much detail as Mother Nature
$\Longrightarrow$ get average *and* fluctutations right
$\Longrightarrow$ make random choices, $\sim$ as in nature

$$\sigma_{\mathrm{final\ state}} = \sigma_{\mathrm{hard\ process}} \mathcal{P}_{\mathrm{tot,hard\ process}\to\mathrm{final\ state}}$$

(appropriately summed & integrated over non-distinguished final states)

where $\mathcal{P}_{\mathrm{tot}} = \mathcal{P}_{\mathrm{res}} \mathcal{P}_{\mathrm{ISR}} \mathcal{P}_{\mathrm{FSR}} \mathcal{P}_{\mathrm{MPI}} \mathcal{P}_{\mathrm{remnants}} \mathcal{P}_{\mathrm{hadronization}} \mathcal{P}_{\mathrm{decays}}$

with $\mathcal{P}_i = \prod_j \mathcal{P}_{ij} = \prod_j \prod_k \mathcal{P}_{ijk} = \ldots$ in its turn

$\Longrightarrow$ **divide and conquer**

an event with n particles involves $\mathcal{O}(10n)$ random choices,
(flavour, mass, momentum, spin, production vertex, lifetime, ...)
LHC: $\sim$ 100 charged and $\sim$ 200 neutral (+ intermediate stages)
$\Longrightarrow$ several thousand choices
(of $\mathcal{O}(100)$ different kinds)

Why generators?

- Allow theoretical and experimental studies of *complex* multiparticle physics
- Large flexibility in physical quantities that can be addressed
- Vehicle of ideology to disseminate ideas from theorists to experimentalists

Can be used to

- predict event rates and topologies
  $\Rightarrow$ can estimate feasibility
- simulate possible backgrounds
  $\Rightarrow$ can devise analysis strategies
- study detector requirements
  $\Rightarrow$ can optimize detector/trigger design
- study detector imperfections
  $\Rightarrow$ can evaluate acceptance corrections

The workhorses: what are the differences?

Herwig, PYTHIA and Sherpa offer convenient frameworks for LHC pp physics studies, covering all aspects above, but with slightly different history/emphasis:

PYTHIA (successor to JETSET, begun in 1978):
originated in hadronization studies,
still special interest in soft physics.

Herwig (successor to EARWIG, begun in 1984):
originated in coherent showers (angular ordering),
cluster hadronization as simple complement.

Sherpa (APACIC++/AMEGIC++, begun in 2000):
had own matrix-element calculator/generator
originated with matching & merging issues.

Delphi and Pythia

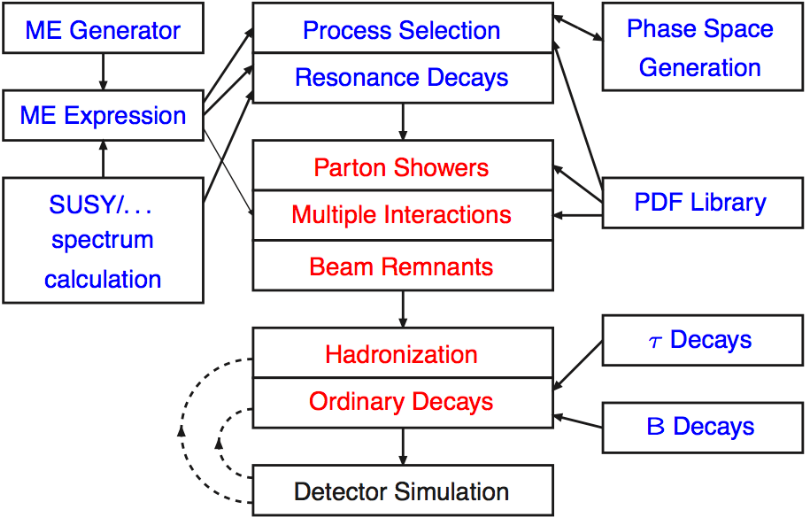

Delphi: 120 km west of Athens, on the slopes of Mount Parnassus.
Python: giant snake killed by Apollon.
The Oracle of Delphi: ca. 1000 B.C. – 390 A.D.
Pythia: local prophetess/priestess.
Key role in myths and history, notably in
"The Histories" by Herodotus of Halicarnassus ($\sim$482 – 420 B.C.)

Other Relevant Software

Some examples (with apologies for many omissions), usually combined for maximum effect:

- Event generators: EPOS, HIjing, Sibyll, DPMjet, Genie
- Matrix-element generators: MadGraph_aMC@NLO, Sherpa, Helac, Whizard, CompHep, CalcHep, GoSam
- Matrix element libraries: AlpGen, POWHEG BOX, MCFM, NLOjet++, VBFNLO, BlackHat, Rocket
- Special BSM scenarios: Prospino, Charybdis, TrueNoir
- Mass spectra and decays: SOFTSUSY, SPHENO, HDecay, SDecay
- Feynman rule generators: FeynRules
- PDF libraries: LHAPDF
- Resummed ($p_{\perp}$) spectra: ResBos
- Approximate loops: LoopSim
- Parton showers: Ariadne, Vincia, Dire, Deductor, PanScales
- Jet finders: anti-$k_{\perp}$ and FastJet
- Analysis packages: Rivet, Professor, MCPLOTS
- Detector simulation: GEANT, Delphes
- Constraints (from cosmology etc): DarkSUSY, MicrOmegas
- Standards: PDG identity codes, LHA, LHEF, SLHA, Binoth LHA, HepMC

Putting it together



Standardized interfaces essential!

PDG particle codes

A. Fundamental objects

1	d	11	e^-	21	g	32	Z'^0	39	G
2	u	12	ν_e	22	γ	33	Z''^0	41	R^0
3	s	13	μ^-	23	Z^0	34	W'^+	42	LQ
4	c	14	ν_μ	24	W^+	35	H^0	51	DM_0
5	b	15	τ^-	25	h^0	36	A^0		
6	t	16	ν_τ			37	H^+	...	...

add − sign for antiparticle, where appropriate

+ diquarks, SUSY, technicolor, ...

B. Mesons

$100\,|q_1| + 10\,|q_2| + (2s+1)$ with $|q_1| \geq |q_2|$

particle if heaviest quark u, $\bar{s}$, c, $\bar{b}$; else antiparticle

111	π^0	311	K^0	130	K^0_L	221	η^0	411	D^+	431	D_s^+
211	π^+	321	K^+	310	K^0_S	331	η'^0	421	D^0	443	J/ψ

C. Baryons

$1000\,q_1 + 100\,q_2 + 10\,q_3 + (2s+1)$

with $q_1 \geq q_2 \geq q_3$, or Λ-like $q_1 \geq q_3 \geq q_2$

2112	n	3122	Λ^0	2224	Δ^{++}	3214	Σ^{*0}
2212	p	3212	Σ^0	1114	Δ^-	3334	Ω^-

Les Houches LHA/LHEF event record

At initialization:

- beam kinds and E's
- PDF sets selected
- weighting strategy
- number of processes

Per process in initialization:

- integrated σ
- error on σ
- maximum $\mathrm{d}\sigma/\mathrm{d}(\mathrm{PS})$
- process label

Per event:

- number of particles
- process type
- event weight
- process scale
- α_{em}
- α_{s}
- (PDF information)

Per particle in event:

- PDG particle code
- status (decayed?)
- 2 mother indices
- colour & anticolour indices
- $(p_x, p_y, p_z, E), m$
- lifetime τ
- spin/polarization

Monte Carlo techniques

"Spatial" problems: no memory/ordering

1. Integrate a function
2. Pick a point at random according to a probability distribution

"Temporal" problems: has memory

1. Radioactive decay: probability for a radioactive nucleus to decay at time t, given that it was created at time 0

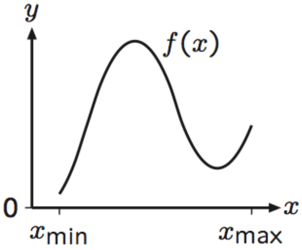
In reality combined into multidimensional problems:

1. Random walk (variable step length and direction)
2. Charged particle propagation through matter (stepwise loss of energy by a set of processes)
3. **Parton showers** (cascade of successive branchings)
4. Multiparticle interactions (ordered multiple subcollisions)

Assume algorithm that returns "random numbers" R, uniformly distributed in range $0 < R < 1$ and uncorrelated.

Integration and selection

Assume function $f(x)$,
studied range $x_{\mathrm{min}} < x < x_{\mathrm{max}}$,
where $f(x) \geq 0$ everywhere



Two connected standard tasks:

1. Calculate (approximatively)

$$\int_{x_{\mathrm{min}}}^{x_{\mathrm{max}}} f(x')\,\mathrm{d}x'$$

2. Select x at random according to $f(x)$

In step 2 $f(x)$ is viewed as "probability distribution"
with implicit normalization to unit area,
and then step 1 provides overall correct normalization.

Integral as an area/volume

Theorem

An n-dimensional integration $\equiv$ an $n+1$-dimensional volume

$$\int f(x_1,\ldots,x_n)\,\mathrm{d}x_1\ldots\mathrm{d}x_n \equiv \int\int_0^{f(x_1,\ldots,x_n)} 1\,\mathrm{d}x_1\ldots\mathrm{d}x_n\,\mathrm{d}x_{n+1}$$

since $\int_0^{f(x)} 1\,\mathrm{d}y = f(x)$.

Integral as an area/volume

Theorem

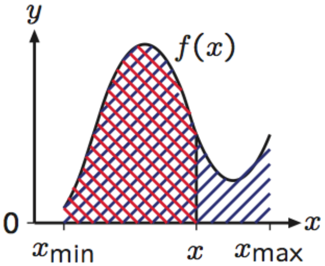
An n-dimensional integration $\equiv$ an $n+1$-dimensional volume

$$\int f(x_1, \ldots, x_n)\,\mathrm{d}x_1 \ldots \mathrm{d}x_n \equiv \int \int_0^{f(x_1,\ldots,x_n)} 1\,\mathrm{d}x_1 \ldots \mathrm{d}x_n\,\mathrm{d}x_{n+1}$$

since $\int_0^{f(x)} 1\,\mathrm{d}y = f(x)$.
So, for $1+1$ dimension, selection of x according to $f(x)$ is equivalent to uniform selection of (x, y) in the area $x_{\mathrm{min}} < x < x_{\mathrm{max}}$, $0 < y < f(x)$.
Therefore

$$\int_{x_{\mathrm{min}}}^{x} f(x')\,\mathrm{d}x' = R \int_{x_{\mathrm{min}}}^{x_{\mathrm{max}}} f(x')\,\mathrm{d}x'$$

(area to left of selected x is uniformly distributed fraction of whole area)



Analytical solution

If **know primitive function** $F(x)$ and **know inverse** $F^{-1}(y)$ then

$$\begin{aligned} F(x) - F(x_{\min}) &= R\,(F(x_{\max}) - F(x_{\min})) = R\,A_{\mathrm{tot}} \\ \Longrightarrow x &= F^{-1}(F(x_{\min}) + R\,A_{\mathrm{tot}}) \end{aligned}$$

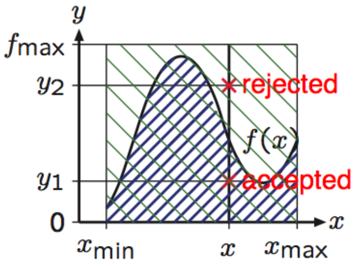
Proof: introduce $z = F(x_{\min}) + R\,A_{\mathrm{tot}}$. Then

$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}x} = \frac{\mathrm{d}\mathcal{P}}{\mathrm{d}R}\,\frac{\mathrm{d}R}{\mathrm{d}x} = 1\,\frac{1}{\frac{\mathrm{d}x}{\mathrm{d}R}} = \frac{1}{\frac{\mathrm{d}x}{\mathrm{d}z}\,\frac{\mathrm{d}z}{\mathrm{d}R}} = \frac{1}{\frac{\mathrm{d}F^{-1}(z)}{\mathrm{d}z}\,\frac{\mathrm{d}z}{\mathrm{d}R}} = \frac{\frac{\mathrm{d}F(x)}{\mathrm{d}x}}{\frac{\mathrm{d}z}{\mathrm{d}R}} = \frac{f(x)}{A_{\mathrm{tot}}}$$

Hit-and-miss solution

If $f(x) \leq f_{\max}$ in $x_{\min} < x < x_{\max}$
use interpretation as an area

1. select
   $x = x_{\min} + R\,(x_{\max} - x_{\min})$
2. select $y = R\,f_{\max}$ (new R!)
3. while $y > f(x)$ cycle to 1



Integral as by-product:

$$I = \int_{x_{\min}}^{x_{\max}} f(x)\,\mathrm{d}x = f_{\max}\,(x_{\max} - x_{\min})\,\frac{N_{\mathrm{acc}}}{N_{\mathrm{try}}} = A_{\mathrm{tot}}\,\frac{N_{\mathrm{acc}}}{N_{\mathrm{try}}}$$

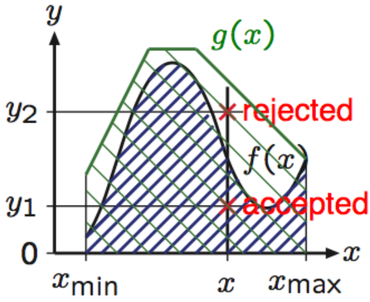
Binomial distribution with $p = N_{\mathrm{acc}}/N_{\mathrm{try}}$ and $q = N_{\mathrm{fail}}/N_{\mathrm{try}}$,
so error

$$\frac{\delta I}{I} = \frac{A_{\mathrm{tot}}\,\sqrt{p\,q/N_{\mathrm{try}}}}{A_{\mathrm{tot}}\,p} = \sqrt{\frac{q}{p\,N_{\mathrm{try}}}} = \sqrt{\frac{q}{N_{\mathrm{acc}}}} < \frac{1}{\sqrt{N_{\mathrm{acc}}}}$$

Importance sampling

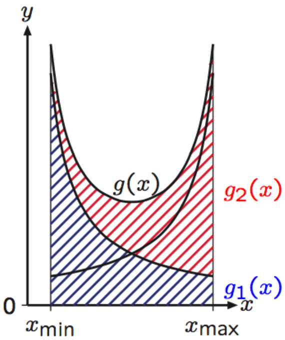
Improved version of hit-and-miss:
If $f(x) \leq g(x)$ in
$x_{\mathrm{min}} < x < x_{\mathrm{max}}$
and $G(x) = \int g(x') \, \mathrm{d}x'$ is simple
and $G^{-1}(y)$ is simple

1. select x according to $g(x)$ distribution
2. select $y = R \, g(x)$ (new R!)
3. while $y > f(x)$ cycle to 1



Multichannel

If $f(x) \leq g(x) = \sum_i g_i(x)$,
where all g_i "nice" ($G_i(x)$ invertible)
but $g(x)$ not



1 select i with relative probability

$$A_i = \int_{x_{\min}}^{x_{\max}} g_i(x') \, \mathrm{d}x'$$

2 select x according to $g_i(x)$

3 select $y = R \, g(x) = R \sum_i g_i(x)$

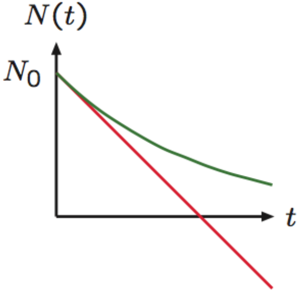
4 while $y > f(x)$ cycle to 1

Works since

$$\int f(x) \, \mathrm{d}x = \int \frac{f(x)}{g(x)} \sum_i g_i(x) \, \mathrm{d}x = \sum_i A_i \int \frac{g_i(x) \, \mathrm{d}x}{A_i} \, \frac{f(x)}{g(x)}$$

Temporal methods: radioactive decays – 1

Consider "radioactive decay":
$N(t)$ = number of remaining nuclei at time t
but normalized to $N(0) = N_0 = 1$ instead, so equivalently
$N(t)$ = probability that (single) nucleus has not decayed by time t
$P(t) = -\mathrm{d}N(t)/\mathrm{d}t$ = probability for it to decay at time t



Naively $P(t) = c \Longrightarrow N(t) = 1 - ct$.
Wrong! Conservation of probability driven by depletion:
a given nucleus can only decay once

Correctly
$P(t) = cN(t) \Longrightarrow N(t) = \exp(-ct)$
i.e. exponential dampening
$P(t) = c\exp(-ct)$

There is memory in time!

Temporal methods: radioactive decays – 2

For radioactive decays $P(t) = cN(t)$, with c constant, but now generalize to time-dependence:

$$P(t) = -\frac{\mathrm{d}N(t)}{\mathrm{d}t} = f(t)\,N(t)\;; \quad f(t) \geq 0$$

Standard solution:

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = -f(t)N(t) \iff \frac{\mathrm{d}N}{N} = \mathrm{d}(\ln N) = -f(t)\,\mathrm{d}t$$

$$\ln N(t) - \ln N(0) = -\int_0^t f(t')\,\mathrm{d}t' \implies N(t) = \exp\left(-\int_0^t f(t')\,\mathrm{d}t'\right)$$

$$F(t) = \int^t f(t')\,\mathrm{d}t' \implies N(t) = \exp\left(-(F(t) - F(0))\right)$$

Assuming $F(\infty) = \infty$, i.e. always decay, sooner or later:

$$N(t) = R \implies t = F^{-1}(F(0) - \ln R)$$

The veto algorithm: problem

What now if $f(t)$ has no simple $F(t)$ or F^{-1}?
Hit-and-miss not good enough, since for $f(t) \leq g(t)$, g "nice",

$$t = G^{-1}(G(0) - \ln R) \implies N(t) = \exp\left(-\int_0^t g(t')\,\mathrm{d}t'\right)$$

$$P(t) = -\frac{\mathrm{d}N(t)}{\mathrm{d}t} = g(t)\,\exp\left(-\int_0^t g(t')\,\mathrm{d}t'\right)$$

and hit-or-miss provides rejection factor $f(t)/g(t)$, so that

$$P(t) = f(t)\,\exp\left(-\int_0^t g(t')\,\mathrm{d}t'\right)$$

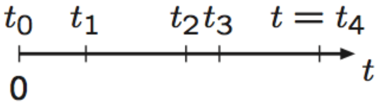
(modulo overall normalization), where it ought to have been

$$P(t) = f(t)\,\exp\left(-\int_0^t f(t')\,\mathrm{d}t'\right)$$

The veto algorithm: solution

The veto algorithm

1. start with $i = 0$ and $t_0 = 0$
2. $i = i + 1$
3. $t = t_i = G^{-1}(G(t_{i-1}) - \ln R)$, i.e $t_i > t_{i-1}$
4. $y = R\, g(t)$
5. while $y > f(t)$ cycle to 2



That is, when you fail, you keep on going from the time when you failed, and *do not* restart at time $t = 0$. (Memory!)

The veto algorithm: proof – 1

Study probability to have i intermediate failures before success:

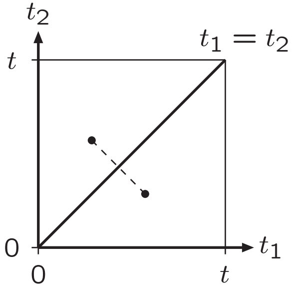
Define $S_g(t_a, t_b) = \exp\left(-\int_{t_a}^{t_b} g(t')\,\mathrm{d}t'\right)$ ("Sudakov factor")

$$P_0(t) = P(t = t_1) = g(t)\,S_g(0,t)\,\frac{f(t)}{g(t)} = f(t)\,S_g(0,t)$$

$$\begin{aligned}
P_1(t) &= P(t = t_2) \\
&= \int_0^t \mathrm{d}t_1\, g(t_1) S_g(0,t_1)\left(1 - \frac{f(t_1)}{g(t_1)}\right) g(t)\,S_g(t_1,t)\,\frac{f(t)}{g(t)} \\
&= f(t)\,S_g(0,t)\int_0^t \mathrm{d}t_1\,(g(t_1) - f(t_1)) = P_0(t)\,I_{g-f}
\end{aligned}$$

$$\begin{aligned}
P_2(t) &= \cdots = P_0(t)\int_0^t \mathrm{d}t_1\,(g(t_1) - f(t_1))\int_{t_1}^t \mathrm{d}t_2\,(g(t_2) - f(t_2)) \\
&= P_0(t)\int_0^t \mathrm{d}t_1\,(g(t_1) - f(t_1))\int_0^t \mathrm{d}t_2\,(g(t_2) - f(t_2))\,\theta(t_2 - t_1) \\
&= P_0(t)\,\frac{1}{2}\left(\int_0^t \mathrm{d}t_1\,(g(t_1) - f(t_1))\right)^2 = P_0(t)\,\frac{1}{2}\,I_{g-f}^2
\end{aligned}$$

The veto algorithm: proof – 2



Generally, i intermediate times corresponds to $i!$ equivalent ordering regions.

$$P_i(t) = P_0(t)\,\frac{1}{i!}\,I_{g-f}^i$$

$$\begin{aligned}
P(t) &= \sum_{i=0}^{\infty} P_i(t) = P_0(t)\sum_{i=0}^{\infty}\frac{I_{g-f}^i}{i!} = P_0(t)\exp(I_{g-f}) \\
&= f(t)\,\exp\left(-\int_0^t g(t')\,\mathrm{d}t'\right)\,\exp\left(\int_0^t (g(t') - f(t'))\,\mathrm{d}t'\right) \\
&= f(t)\,\exp\left(-\int_0^t f(t')\,\mathrm{d}t'\right)
\end{aligned}$$

The winner takes it all

Assume "radioactive decay" with two possible decay channels 1&2

$$P(t) = -\frac{\mathrm{d}N(t)}{\mathrm{d}t} = f_1(t)N(t) + f_2(t)N(t)$$

Alternative 1:
use normal veto algorithm with $f(t) = f_1(t) + f_2(t)$.
Once t selected, pick decays 1 or 2 in proportions $f_1(t) : f_2(t)$.

Alternative 2:

The winner takes it all

select t_1 according to $P_1(t_1) = f_1(t_1)N_1(t_1)$
and t_2 according to $P_2(t_2) = f_2(t_2)N_2(t_2)$,
i.e. as if the other channel did not exist.
If $t_1 < t_2$ then pick decay 1, while if $t_2 < t_1$ pick decay 2.

Equivalent by simple proof.

Radioactive decay as perturbation theory

Assume we don't know about exponential function.
Recall wrong solution, starting from $N(t) = N_0(t) = 1$:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -cN = -cN_0(t) = -c \Rightarrow N(t) = N_1(t) = 1 - ct$$

Now plug in $N_1(t)$, hoping to find better approximation:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -cN_1(t) \Rightarrow N(t) = N_2(t) = 1 - c\int_0^t (1 - ct')\mathrm{d}t' = 1 - ct + \frac{(ct)^2}{2}$$

and generalize to

$$N_{i+1}(t) = 1 - c\int_0^t N_i(t')\,\mathrm{d}t' \Rightarrow N_{i+1}(t) = \sum_{k=0}^{i+1} \frac{(-ct)^k}{k!}$$

which recovers exponential e^{-ct} for $i \to \infty$.

Reminiscent of (QED, QCD) perturbation theory with $c \to \alpha f$.

Summary

Main event components:

- parton distributions
- hard subprocesses
- initial-state radiation
- final-state interactions
- multiparton interactions
- beam remnants
- hadronization
- decays
- total cross sections

Main Monte Carlo methods:

- integration as an area
- analytical solution
- hit-and-miss
- importance sampling
- multichannel
- **the veto algorithm**
- the winner takes it all