FIG. 17. Event displays of $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ candidates in the $n_j = 0$ (top) and $n_j \geq 2$ VBF-enriched (bottom) categories.

The neutrinos are represented by missing transverse momentum ($\text{met}$, dotted line) that points away from the $e\mu$ system.

The properties of the first event are $p_{e\mu} = 33$ GeV, $p_{\text{met}} = 37$ GeV, and $m_{\nu\nu} = 48$ GeV.

The properties of the second event are $p_{e\mu} = 51$ GeV, $p_{\text{met}} = 59$ GeV, and $m_{\nu\nu} = 127$ GeV. Both events have a small value of $\tau$, which is characteristic of the signal. The second event shows two well-separated jets that are characteristic of VBF production.
Acknowledgements

These slides have been prepared using heavily the material prepared by Christian Schultz Coulon and clarifying few points plus updating to recent developments in proton-proton physics.
Why MC simulation?

1) to extract an interesting signal we need to subtract the expectation from known processes;

2) signal needs also to be modelled in order to compute detection efficiency and estimate production cross sections and couplings.
The simulation chain

\[ \sigma = \sum_{ij} \int dx_1 dx_2 \ f_i(x_1, Q^2) f_j(x, Q^2) \hat{\sigma}(Q^2) \]

\(\otimes\) Hadronization

Hard Process [calculable]

PDFs

Proton

\(p_1\)

\(x_j p_1\)

\(f_j(x_j)\)

\(f_k(x_k)\)

Proton

\(p_2\)

\(x_k p_2\)

Parton Shower

Hadron-Jets Leptons ...

B. Di Micco

Università degli Studi di Roma Tre
MC simulations in particle physics

How Monte Carlo simulation works

• Numerical process generation based on random numbers

• Method very powerful in particle physics

Event generation programs:

Pythia6, Pythia8, Herwig, Herwig++, Sherpa ...

Hard partonic subprocess + fragmentation and hadronisation ...

Detector simulation:

Geant4
Fluka low energy hadron interactions...

interaction & response of all produced particles ...

Event Generator
simulate physics process (quantum mechanics: probabilities!)

Detector Simulation
simulate interaction with detector material

Digitisation
translate interactions with detector into realistic signals

Reconstruction/Analysis
as for real data
Baseline of the simulation process

Typically, we need to generate a continuous variable following some distribution i.e. energy loss of a particle in a given material segment; angle of a photon in the h reference frame for the h → γγ decay

\[ dP = f(x, \ldots)dx \]

\( dP = f(\theta, \phi)d\theta d\phi = \text{sen} \theta d\theta d\phi \)

flat distribution in \( \phi \)

non flat in \( \theta \)
Distribution function transformation properties

1) software libraries provide basic functions to produce flat distributed random numbers in the interval [0,1] (ex. root TRandom3 class), they are typically fast and accurate uniform random numbers generators;

2) starting from uniform distributed random numbers, it is possible to generate numbers following any distribution using different techniques

\[ dP_x = f(x)dx \quad y = g(x) \]
\[ x \in [x_a, x_b] \]
\[ dP_y = h(y)dy = h(y)g'(x)dx \]

\[ h(y)g'(x) = f(x) \Rightarrow h(y) = \frac{f(x)}{g'(x)} = \frac{f(g^{-1}(y))}{g'(g^{-1}(y))} \]

Ex.: range map

\[ [0, 1] \rightarrow [a, b] \quad y = (b - a)x + a \]
\[ f(x) = 1 \quad g'(x) = b - a \quad h(y) = \frac{1}{b - a} \]

How “y” distributes in \([g(x_a), g(x_b)]\)?

Because \(y\) is a monotonic function of \(x\) the probability to have \(y\) between \(g(x)\) and \(g(x+dx)\) is equal to the probability to have \(x\) between \(x\) and \(x+dx\)
Distribution function transformation properties

Ex. 2: integration method:

\[ g(x) = \frac{1}{\int_a^b f(x') \, dx'} \int_a^x f(x') \, dx' \quad g'(x) = \frac{f(x)}{\int_a^b f(x') \, dx'} \]

\[ h(y) = \frac{f(x)}{g'(x)} = \frac{f[g^{-1}(y)]}{f[g^{-1}(y)]} \cdot \int_a^b f(x') \, dx' = \int_a^b f(x') \, dx' \]

y is uniformly distributed:

1) generate y flat in \([f_{\text{min}}, f_{\text{max}}] \);  
2) compute \( x = g^{-1}(y) \), \( x \) will be distributed in \( g^{-1}(f_{\text{min}}), g^{-1}(f_{\text{max}}) \)

Finding \( g^{-1}(y) \) is equivalent to solve the equation:

\[ \frac{1}{\int_a^b f(x') \, dx'} \int_a^x f(x') \, dx' = y \]
Hit or miss method.

1) generate \( x \) flat in \( x_{\text{min}}, x_{\text{max}} \)
2) generate \( y \) flat in \( 0, f_{\text{max}} \)
3) if \( y < f(x) \) accept the event, otherwise ignore it

for a given \( x \) in \( x_{\text{min}}, x_{\text{max}} + dx \) the fraction of accepted events is proportional to \( f(x)dx \rightarrow dP_x = f(x)dx \)

1) advantages:
   - can be used for all functions, even non continuous …
   - can be extended to N-dimension (generate \( x_1, x_2, \ldots, x_n \), y accept if \( y < f(x_1, x_2, \ldots, x_n) \))

2) disadvantages
   - can be extremely slow

points generated uniformly in the square
points accepted only below the curve

MC generators implement “smart” generation techniques to increase efficiencies
Comparison between real and simulated events

"Real"

LHC collisions
Events

Detector, DAQ
ATLAS, CMS, LHCb, ALICE

"Virtual"

Event Generator
Pythia8, Pythia, Herwig++

Detector Simulation
Geant4, ...

Events

Produce events

Observe/store events

Event Reconstruction
Athena (ATLAS), …

Physics Analysis
Root based analysis packages

Compare data & simulation

"Quick & Inaccurate"

[after T. Sjöstrand]
Use specialised programs

Now fully automatised in programs like Madgraph5_aMC@NLO (from the lagrangian to the full simulation)

- **ME Generator**
- **ME Expression**
- **SUSY/... spectrum calculation**

**Process Selection**
- **Resonance Decays**
- **Parton Showers**
- **Multiple Interactions**
- **Beam Remnants**

**Phase Space Generation**

**PDF Library**
- **$\tau$ Decays**
- **$B$ Decays**

**Hadronization**

**Ordinary Decays**

**Detector Simulation**

The tricky part of the calculations is the virtual corrections. NLO is now state-of-the-art, with NNLO still in its infancy. If one is content with Born-level diagrams only, i.e. without any loops, it is possible to go to quite high orders, with up to something like eight partons in the final state. These
Simulation elements

- **ME Generator**
- **ME Expression**
- **SUSY/... spectrum calculation**
- **Process Selection**
  - **Resonance Decays**
- **Parton Showers**
- **Multiple Interactions**
- **Beam Remnants**
- **Hadronization**
- **Ordinary Decays**

**Phase Space Generation**

- **PDF Library**
- **τ Decays**
- **B Decays**

**Detector Simulation**

Use specialised programs

---

*Fig. 1: Example how different programs can be combined in the event-generation chain.*

Information and form factors require special encoding. Even a sophisticated event-simulation program some parts of the generator may use during simulation of secondary interactions and decays. Several standards have been developed to further this interoperability. The Les Houches Accord (LHA) for user processes [10] specifies how parton-level information about the hard process and sequential decays can be encoded and passed on to a general-purpose generator. Originally it was defined in terms of two Fortran commonblocks, but more recently a standard Les Houches Event File format [11] offers a language-independent alternative approach. The Les Houches Accord Parton Density Functions (LHAPDF) library [12] makes different PDF sets available in a uniform framework. The SUSY Les Houches Accord (SLHA) [13] allows a standardized transfer of masses, mixing, couplings and branching ratios from spectrum calculators to other programs. Finally, the HepMC C++ event record [14] succeeds the HEPEVT Fortran one [15] as a standard way to transfer information from a generator on to the detector-simulation stage. One of the key building blocks for several of these standards is the PDG codes for all the most common particles [16], also in some scenarios for physics beyond the Standard Model.

The $2 \rightarrow 2$ processes we started out with above are about the simplest on a hadron collider. In reality one needs to go on to higher orders. In $O(\alpha_s^3)$ two new kind of graphs enter. One kind is where one additional parton is present in the final state, i.e. $2 \rightarrow 3$ processes. The cross section for such processes is almost always divergent when one of the parton energies vanish (soft singularities) or two partons become collinear (collinear singularities). The other kind is loop graphs, with an additional intermediate parton not present in the final state, i.e. a correction to the $2 \rightarrow 2$ processes. Strictly speaking, at $O(\alpha_s^3)$ one picks up the interference between the lowest-order graph and the loop graph, and this interference has negative divergences that exactly cancel the positive ones above, with only finite terms surviving. For inclusive event properties such next-to-leading order (NLO) calculations lead to an improved accuracy of predictions, but for more exclusive studies the mathematical cancellation of singularities has to be supplemented by more physical techniques, which is far from trivial.

The tricky part of the calculations is the virtual corrections. NLO is now state-of-the-art, with NNLO still in its infancy. If one is content with Born-level diagrams only, i.e. without any loops, it is possible to go to quite high orders, with up to something like eight partons in the final state. These

---

**GEANT** Geometry And Tracking

Detailed description of detector **geometry**
[sensitive & insensitive volumes]

**Tracking** of all particles through detector material ...

[geant4.kek.jp/~tanaka/GEANT4/ATLAS_G4_GIFFIG/]

→ **Detector response**

Developed at CERN since 1974 (FORTRAN)
[Today: Geant4; programmed in C++]
Several standards have been developed to further this interoperability. The Les Houches Accord (LHA) for user processes [10] specifies how parton-level information about the hard process and sequential decays can be encoded and passed on to a general-purpose generator. Originally it was defined in terms of two Fortran commonblocks, but more recently a standard Les Houches Event File format [11] offers a language-independent alternative approach. The Les Houches Accord Parton Density Functions (LHAPDF) library [12] makes different PDF sets available in an uniform framework. The SUSY Les Houches Accord (SLHA) [13] allows a standardized transfer of masses, mixings, couplings and branching ratios from spectrum calculators to other programs. Finally, the HepMC C++ event record [14] succeeds the HEPEVT Fortran one [15] as a standard way to transfer information from a generator on to the detector-simulation stage. One of the key building blocks for several of these standards is the PDG codes for all the most common particles [16], also in some scenarios for physics beyond the Standard Model.

The processes we started out with above are about the simplest on a hadron collider. In reality one needs to go on to higher orders. In $O(\alpha_s^3)$ two new kind of graphs enter. One kind is where one additional parton is present in the final state, i.e. $2 \rightarrow 3$ processes. The cross section for such processes is almost always divergent when one of the parton energies vanish (soft singularities) or two partons become collinear (collinear singularities). The other kind is loop graphs, with an additional intermediate parton not present in the final state, i.e. a correction to the $2 \rightarrow 2$ processes. Strictly speaking, at $O(\alpha_s^3)$ one picks up the interference between the lowest-order graph and the loop graph, and this interference has negative divergences that exactly cancel the positive ones above, with only finite terms surviving. For inclusive event properties such next-to-leading order (NLO) calculations lead to an improved accuracy of predictions, but for more exclusive studies the mathematical cancellation of singularities has to be supplemented by more physical techniques, which is far from trivial.

The tricky part of the calculations is the virtual corrections. NLO is now state-of-the-art, with NNLO still in its infancy. If one is content with Born-level diagrams only, i.e. without any loops, it is possible to go to quite high orders, with up to something like eight partons in the final state. These...
Strong interactions:

No free Quarks

Expect jets
i.e. bundles of particles at high energies
[hadron $p_T$ range limited w.r.t. initial parton]

First observation of jets
in $e^+e^-$ collisions @ $E_{\text{CMS}} > 6$ GeV
[SPEAR, SLAC, 1975]

Later also observed in hadron-hadron collisions
[e.g. @ CERN ISR]

Goal: Infer parton properties from jet properties
[need to calculate and/or model fragmentation & hadronisation process]
Pure matrix element (ME) simulation:

MC integration of cross section & PDFs, no hadronisation
(recall: cross section = \(|\text{matrix element}|^2 \otimes \text{phase space}\))

Useful for theoretical studies, no exclusive events generated

[Example: MCFM (http://mcfm.fnal.gov); many LHC processes up to NLO,
HNNLO (http://theory.fi.infn.it/grazzini/codes.html) Higgs production at NNLO]

Event generators:

Combination of ME and parton showers ...

Typical: generator for leading order ME
combined with leading log (LL) parton shower MC (see later)

Exclusive events ➔ useful for experimentalists ...
Parton showers

A realistic simulation needs many particles in the final state, it is quite difficult (sometimes impossible) to compute a pp \((2 \rightarrow n)\) → many particles process

\[(2 \rightarrow n) = \ldots = (2 \rightarrow 2) \oplus \text{ISR} \oplus \text{FSR}\]

FSR: Final state radiation
\(Q^2 \sim m^2 > 0\) decreasing
[time-like shower]

ISR: Initial state radiation
\(Q^2 \sim -m^2 > 0\) increasing
[space-like shower]

Hard process \([2 \rightarrow 2]\):

\[
\sigma = \int \int \int dx_1 \, dx_2 \, d\hat{t} \, f_i(x_1, Q^2) \, f_j(x_2, Q^2) \frac{d\hat{\sigma}_{ij}}{d\hat{t}}
\]

Shower evolution:

Viewed as probabilistic process, which occurs with unit total probability; cross section not directly affected; only indirectly via changed event shape.
Parton showers

e^+e^− → qqg

\[
x_i = \frac{2E_i}{E_{cm}} \quad x_1 + x_2 + x_3 = 2
\]

Cross section:
\[
\frac{d\sigma_{qqg}}{dx_1 dx_2} = \frac{4 \alpha_s}{3 \cdot 2\pi} \cdot \sigma_0 \cdot \frac{x_1^2 + x_2^2}{(1 - x_1)(1 - x_2)}
\]

[m_q = 0; see e.g. Halzen/Martin]

Cross section has large contributions for \(x_1, x_2 \rightarrow 1\)

from \(p_T\) balance
\[
1 - x_2 = \frac{m_{13}^2}{E_{cm}^2} = \frac{Q^2}{E_{cm}^2}
\]

\(m_{13}^2 \sim 2E_1 E_2 (1 - \cos \theta) x_2 \rightarrow 1 \Rightarrow m_{13}^2 \rightarrow 0 \Rightarrow \theta \rightarrow 0\) collinear limit

Rewrite for \(x_2 \rightarrow 1\):

\[
dx_2 = -\frac{dQ^2}{E_{cm}^2}
\]

[qg collinear limit]

\[
E_q = E_1 = zE_b \quad E_g = E_3 = (1 - z)E_b
\]

\(x_3 \approx 1 - z\)

\[
d\mathcal{P} = \frac{d\sigma_{qqg}}{\sigma_0} = \frac{4 \alpha_s}{3 \cdot 2\pi} \cdot \frac{dx_2}{(1 - x_2)} \cdot \frac{x_1^2 + x_2^2}{(1 - x_1)} dx_1 \approx \frac{\alpha_s}{2\pi} \cdot \frac{dQ^2}{Q^2} \cdot \frac{4}{3} \left[ \frac{1 + z^2}{1 - z} \right] dz
\]

\(z \rightarrow 1 \Rightarrow E_g \rightarrow 0\) soft divergence
Splittung probability determined by splitting functions $P_{q\rightarrow qg}$

Analogous splitting functions used in PDF evolution

\[
dP_{a\rightarrow bc} = \frac{\alpha_s}{2\pi} \frac{dQ^2}{Q^2} P_{a\rightarrow bc}(z)dz
\]

\[
P_{q\rightarrow qg} = \frac{4}{3} \frac{1 + z^2}{1 - z}
\]

\[
P_{g\rightarrow gg} = 3 \frac{(1 - z(1 - z))^2}{z(1 - z)}
\]

\[
P_{g\rightarrow q\bar{q}} = \frac{n_f}{2} (z^2 + (1 - z)^2)
\]

$z$ : fractional momentum of radiated parton

$n_f$ : number of quark flavours

In NLO calculations soft and collinear divergencies cancelled by virtual contributions: they persist in LO calculations.

Iteration yields

parton shower ...

Need soft/collinear cut-offs to avoid non-perturbative regions ...

[divergencies!]

Details model-dependent

e.g. $Q > m_0 = \min(m_{ij}) \approx 1 \text{ GeV}$,

$z_{\text{min}}(E,Q) < z < z_{\text{max}}(E,Q)$ or

$p_{\perp} > p_{\perp \text{min}} \approx 0.5 \text{ GeV}$
Parton shower evolution 1

Conservation of total probability:

\[ \mathcal{P}(\text{nothing happens}) = 1 - \mathcal{P}(\text{something happens}) \]

Time evolution:

\[ \mathcal{P}_{\text{nothing}}(0 < t \leq T) = \mathcal{P}_{\text{nothing}}(0 < t \leq T_1) \mathcal{P}_{\text{nothing}}(T_1 < t \leq T) \]

\[ \mathcal{P}_{\text{nothing}}(0 < t \leq T) = \lim_{n \to \infty} \prod_{i=0}^{n-1} \mathcal{P}_{\text{nothing}}(T_i < t \leq T_{i+1}) \]

\[ = \lim_{n \to \infty} \prod_{i=0}^{n-1} (1 - \mathcal{P}_{\text{something}}(T_i < t \leq T_{i+1})) \]

\[ = \exp \left( - \lim_{n \to \infty} \sum_{i=0}^{n-1} \mathcal{P}_{\text{something}}(T_i < t \leq T_{i+1}) \right) \]

\[ = \exp \left( - \int_0^T \frac{d\mathcal{P}_{\text{something}}(t)}{dt} dt \right) \]

\[ \Rightarrow d\mathcal{P}_{\text{first}}(T) = d\mathcal{P}_{\text{something}}(T) \exp \left( - \int_0^T \frac{d\mathcal{P}_{\text{something}}(t)}{dt} dt \right) \]
Parton shower evolution 2

Instead of evolving to later and later times
need to evolve to smaller and smaller $Q^2$ ...
[Heisenberg: $Q \sim 1/t$]

\[
dP_{a\rightarrow bc} = \frac{\alpha_s}{2\pi} \frac{dQ^2}{Q^2} P_{a\rightarrow bc}(z) \, dz \, \exp \left( - \sum_{b,c} \int_{Q^2}^{Q_{\text{max}}^2} \frac{dQ'^2}{Q'^2} \int \frac{\alpha_s}{2\pi} P_{a\rightarrow bc}(z') \, dz' \right)
\]

Sudakov Form Factor

Probability to radiated with virtuality $Q^2$

Note that $\sum_{b,c} \iint dP_{a\rightarrow bc} = 1$...
[Convenient for Monte Carlo]

Sudakov form factor ...

... provides “time” ordering of shower ...
[lower $Q^2 \leftrightarrow$ longer times]

... regulates singularity for first emission ...

But in the limit of repeated soft emissions $q \rightarrow qg$ (but no $g \rightarrow gg$)
one obtains the same inclusive $Q$ emission spectrum as for ME,
i.e. divergent ME spectrum $\leftrightarrow$ infinite number of PS emissions
Sudakov picture of parton showers

Basic algorithm: Markov chain
[each step requires only knowledge only of previous step]
(i) Start with virtuality $Q_1$ and momentum fraction $x_1$
(ii) Generate target virtuality $Q_2$ with random number $R_t$ uniform distributed in [0,1]

Probability to not have $Q_x > Q_2$
using:
$$\Delta(Q_i^2) = \exp \left( - \sum_{b,c} \int_{Q_i^2}^{Q_{max}} \frac{dQ'^2}{Q'^2} \int \frac{\alpha_s}{2\pi} P_{a \rightarrow bc}(z') \, dz' \right)$$
solve the equation for $Q_2$
$$R_t = \frac{\Delta(Q_2^2)}{\Delta(Q_1^2)}$$

[probability to evolve from $t_1$ to $t_2$ without radiation]

(iii) $Q_2$ known ($x_2$ known), need to compute $x_1 \sim z$
$$P_{q \rightarrow qg} = \frac{4 \, 1 + z^2}{3 \, 1 - z} \quad R_z = \frac{\int_z^1 P(z') \, dz'}{\int_0^1 P(z') \, dz'} \quad \text{flat distributed} \quad R_z \in [0,1]$$

(iv) Generate random azimuthal angle $\Phi$ flat distributed

Process ends when partons are below threshold ($p_T, Q$)
Parton shower and logarithmic resummation

If $\alpha_s$ is small higher contributions are power suppressed, but... $\alpha_s$ increases at small $Q^2$

$$\alpha_s(Q_n) \sim \alpha_s(Q_1) \ln(Q_1/Q_n)$$

$$\alpha_s(Q_1) + \alpha_s(Q_1)\alpha_s(Q_2) + \ldots + \alpha_s(Q_1) \cdot \ldots \cdot \alpha_s(Q_n)$$

$$\sim [\alpha_s(Q_1) \ln(Q_1)]^2 \sim [\alpha_s(Q_1) \ln(Q_1)]^n$$

if $\alpha_s(Q_1) \ln(Q_1)$ is large, the expansion is broken, PS allow to sum up all the large contribution [Leading Log resummation]
Parton shower ordering

\[
dP_{a\rightarrow bc} = \frac{\alpha_s}{2\pi} \frac{dQ^2}{Q^2} P_{a\rightarrow bc}(z) \, dz \, \exp \left( -\sum_{b,c} \int_{Q^2}^{Q_{\text{max}}^2} \frac{dQ'^2}{Q'^2} \int \frac{\alpha_s}{2\pi} P_{a\rightarrow bc}(z') \, dz' \right)
\]

In the splitting function appears only \(dQ^2/Q^2\), therefore if \(P = f(z)Q^2 \, dP/P = dQ^2/Q^2\)

Three main approaches to showering in use:

\[p_\perp^2 \approx z(1 - z)m^2 \quad \text{p}_T \text{ ordered showers} \quad E^2\theta^2 \approx m^2/(z(1 - z)) \quad \text{angular ordered showers}\]

Two are based on the standard shower language of a \(\rightarrow\) bc successive branchings:

- HERWIG, HERWIG++ : \(Q^2 \approx E^2(1 - \cos\theta) \approx E^2\theta^2/2\)
- PYTHIA, 8 (basic) : \(Q^2 = m^2\) (timelike) or \(-m^2\) (spacelike)
- PYTHIA6, 8 (\(p_T\)ordered) : mixture: collinear splitting but di-pole kinematic

One is based on a picture of dipole emission:

- Ariadne : \(Q^2 = p_\perp^2\); FSR mainly, ISR is primitive ...

consider the full recoil and not only the branching

B. Di Micco

Università degli Studi di Roma Tre
**PYTHIA**: $Q^2 = m^2$

**HERWIG/++**: $Q^2 \sim E^2 \theta^2$

**ARIADNE/Pythia8**: $Q^2 = p^2_{\perp}$

- Large mass first
  - "hardness" ordered
  - Covers phase space
  - ME merging simple
  - $g \rightarrow qq$ simple
  - not Lorentz invariant
  - no stop/restart
  - ISR: $m^2 \rightarrow -m^2$

- Large angle first
  - not "hardness" ordered
  - Gaps in coverage
  - ME merging messy
  - $g \rightarrow qq$ simple
  - not Lorentz invariant
  - no stop/restart
  - ISR: $\theta \rightarrow \theta$

- Large $p_{\perp}$ first
  - "hardness" ordered
  - Covers phase space
  - ME merging simple
  - $g \rightarrow qq$ messy
  - Lorentz invariant
  - can stop/restart
  - ISR: complicated
Color coherence

**QED: Chudakov effect (mid-fifties)**

1. soft gluons see the pair of split gluons as a whole, color screening reduce their emission

2. angular ordered and $p_T$ ordered PS reproduce the correct color coherence

3. Pythia Q$^2$ needs aposteriori corrections

**QCD: colour coherence for soft gluon emission**

solved by:

- requiring emission angles to be decreasing
- requiring transverse momenta to be decreasing
Comparisons to LHC data

4 jets cross section: $p_T^{(1)} > p_T^{(2)} > p_T^{(3)} > p_T^{(4)}$

- ATLAS
- $\sqrt{s}=8$ TeV, 95 pb$^{-1}$ - 20.3 fb$^{-1}$
- Theory/Data
- $d\sigma / dp_T^{(1)}$ [fb/GeV]
- $d\sigma / dp_T^{(2)}$ [fb/GeV]
- $d\sigma / dp_T^{(4)}$ [fb/GeV]
- $d\sigma / dm_{4j}$ [fb/GeV]

- Data
- Pythia 8 ($\times 0.6$)
- Herwig++ ($\times 1.4$)
- MadGraph+Pythia ($\times 1.1$)
- $p_T^{(1)}>100$ GeV
- Total experimental systematic uncertainty

- $p_T^{(1)}$ [GeV]
- $p_T^{(2)}$ [GeV]
- $p_T^{(4)}$ [GeV]
- $m_{4j}$ [GeV]
Example of processes implemented in Pythia6
Process simulation

Many specialized processes already available in Pythia8/Herwig++

but, processes usually only implemented in lowest non-trivial order ...

Need external programs that ...

1. include higher order loop corrections or, alternatively, do kinematic dependent rescaling

3. allow matching of higher order ME generators [otherwise need to trust parton shower description …]

5. provide correct spin correlations often absent in PS …[e.g. top produced unpolarized, while $t \rightarrow bW \rightarrow blv$ decay correct]

7. simulate newly available physics scenarios …[appear quickly; need for many specialised generators]

Les Houches Accord ...

Specifies how parton-level information about the hard process and sequential decays can be encoded and passed on to a general-purpose generator.

Les Houches: regular annual meeting between theoreticians and experimentalists on MC generator developments.
**Specialised Generators [some examples]**

<table>
<thead>
<tr>
<th>Specialized Generator</th>
<th>AcerMC</th>
<th>ALPGEN</th>
<th>AMEGIC++</th>
<th>CompHEP</th>
<th>GRACE</th>
<th>GR@PPA</th>
<th>MadCUP</th>
<th>HELAS &amp; MadGraph</th>
<th>MCFM</th>
<th>O’Mega &amp; WHIZARD</th>
<th>VECBOS</th>
<th>HRES</th>
<th>DYNNNLO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ttbb, .sinlr top</td>
<td>W/Z + ≤ 6j, nW + mZ + kH + ≤ 3j, ...</td>
<td>generic LO</td>
<td>generic LO</td>
<td>generic LO</td>
<td>bbbb</td>
<td>W/Z+ ≤ 3j, ttbb</td>
<td>generic LO</td>
<td>NLO W/Z+ ≤ 2j, WZ, WH, H+ ≤ 1j</td>
<td>generic LO</td>
<td>Higgs boson production</td>
<td>W/Z production @NNLO</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Type I: Leading order matrix element & leading log parton shower

LO ME for hard processes
[2→1 or 2→2]

1) 2)

1. Leading order matrix element
2. Leading log parton shower
**Type I: Leading order matrix element & leading log parton shower**

**LO ME for hard processes**

[2→1 or 2→2]

1) \hspace{2cm} 2)

- Parton Shower: attaches gluons at each leg.
- only in soft-collinear approximation
- typically underestimate large angle/hard emission
- 1) or 2) at ME (different generations, different accuracy: cannot be combined
**Type 2**: Leading order matrix element & leading log parton shower + merging

**LO ME for hard processes**

\[2 \rightarrow 1 \text{ or } 2 \rightarrow 2\]

1) from PS  
2) from ME

- Type 1 can be improved using 1) + 2)
- Use ME calculation for hard/large angle jets
- But needs to remove double-counting: merging (CKKW, MLM)
- Very good description of high jet multiplicity kinematics

---

**Software Tools**

- ME+PS
  - Herwig++/Pythia6/8
    - Single Top, TopRex
    - Phantom
    - AcerMC
    - GRAPPA
    - CompHEP
- ME+PS+merging
  - Alpgen
  - MadGraph
  - Sherpa
- NLO+PS
  - MC@NLO
  - POWHEG

---

B. Di Micco  
Università degli Studi di Roma Tre  
[F. Maltoni]
Merging @LO

MLM matching (simplified)

1) define matching cuts:
   for example $p_T^J > 20$ GeV, $\Delta R=0.4$
Merging @LO

MLM matching (simplified)

1) define matching cuts: for example $p_T^J > 20$ GeV, $\Delta R=0.4$

2) generate ME with 1, 2, …n jets
Merging @LO

MLM matching (simplified)

1) define matching cuts:
   for example $p_T^J > 20$ GeV, $\Delta R = 0.4$

2) generate ME with 1, 2, …n jets

3) shower all events
Merging @LO

MLM matching (simplified)

1) define matching cuts:
   for example $p_T^J > 20$ GeV, $\Delta R=0.4$

2) generate ME with 1, 2, …n jets

3) shower all events

4) select only events where jets above the $p_T$ threshold match with final partons
Merging @LO

MLM matching (simplified)

1) define matching cuts:
   for example $p_T^J > 20$ GeV, $\Delta R=0.4$

2) generate ME with 1, 2, …n jets

3) shower all events

4) select only events where jets above the $p_T$ threshold match with final partons

Consequences:
   all jets with $p_T > 20$ GeV and $\Delta R>0.4$ to other jets come from ME
collinear and soft jets come from PS
Use each of them where they are best.
W+jets distributions

Type III: Next-to-leading order ME & leading-log parton shower

hard processes simulated at NLO accuracy including real & virtual corrections ...

improved description of cross sections & kinematic distributions

Herwig++/Pythia6/8

Herwig++/Pythia6/8
Single Top, TopRex
Phantom
AcerMC
GRAPPA
CompHEP

ME+PS+merging

Alpgen
MadGraph
Sherpa

NLO+PS

MC@NLO
POWHEG
Type III: Next-to-leading order ME & leading-log parton shower

hard processes simulated at NLO accuracy including real & virtual corrections ...

improved description of cross sections & kinematic distributions

2 Matching methods:

1. Powheg
   - first emission produced by the ME;
   - don’t allow the PS to produce patrons harder than the first emission;
   - not exact at NLO (contains unbalanced higher order terms)

2. MC@NLO:
   \[ |\text{ME}|^2 = |\text{ME} + \text{PS} - \text{PS}(\text{up to } \alpha_s^2)|^2 \]
   - Result is exact at NLO...
   - produce some negative weights, need retuning for each PS
Merging @NLO (quite new, going to be used at 13 TeV)

FxFx (Frederix-Frixione) merging

1) define a matching scale $\mu_Q$;

2) don’t allow $S$ events with $p_T > \mu_Q$ (those will be provided by $H$ events of n-1 partons NLO real emission); the restriction is imposed both at ME and on the shower starting scale $\mu < \mu_Q$

3) treat the obtained events as LO ones and apply an LO-style merging (this allow to produce smoother distributions)
Let’s recap
From partons to color neutral hadrons

**Fragmentation:**
Parton splitting into other partons
[QCD: re-summation of leading-logs]
[“Parton shower”]

**Hadronization:**
Parton shower forms hadrons
[non-perturbative, only models]

**Decay** of unstable hadrons
[perturbative QCD, electroweak theory]
Modelling relies on phenomenological models available

Models based on MC simulations very successful:

- Generation of complete final states ...
  [Needed by experimentalists in detector simulation]

Caveat: tunable ad-hoc parameters

Most popular MC models:

- Pythia/8 : Lund string model
- Herwig/++ : Cluster model
Simplest approach:  

Start with original quark
Generate quark-antiquark pairs from vacuum

- form “primary meson” with energy fraction z
Continue with leftover quark with energy fraction 1−z
Stop at low energies (cut-off)
Include flavour non-perturbative fragmentation functions D(z)

D(z): probability to find a meson/hadron with energy fraction z in jet ...
Lund String Model


QCD potential:

\[ V(r) = -\frac{4}{3} \alpha_s \left( \frac{1}{r^2} \right) + kr \]

neglected

String formation between initial quark-antiquark pair

- String breaks up if potential energy large enough to produce a new quark-antiquark pair
- Gluons = 'kinks' in string
- At low energy: hadron formation
- Very widely used ...
  [default in Pythia 6/8]

After: Ellis et al.,
QCD and Collider Physics
Lund String Model

Repeated string breaks for large system with pure $V(r) = \kappa \cdot r$, i.e. neglect Coulomb part

$$\left| \frac{dE}{dz} \right| = \left| \frac{dp_z}{dz} \right| = \left| \frac{dE}{dt} \right| = \left| \frac{dp_z}{dt} \right| = \kappa$$

Energy-momentum quantities can be read from space-time quantities ...

Simple but powerful picture of hadron production

[with extensions to massive quarks, baryons, ...]

Yields: Common Gaussian $p_{\perp}$ spectrum

Heavy quark suppression
Repeated string breaks for large system with pure $V(r) = \kappa \cdot r$, i.e. neglect Coulomb part

\[
\frac{dE}{dz} = \left| \frac{dp_z}{dz} \right| = \left| \frac{dE}{dt} \right| = \left| \frac{dp_z}{dt} \right| = \kappa
\]

Energy-momentum quantities can be read from space-time quantities ...

Simple but powerful picture of hadron production

[with extensions to massive quarks, baryons, ...]

\[
P \propto \exp \left( -\frac{\pi m_{\perp q}^2}{\kappa} \right) \exp \left( -\frac{\pi p_{\perp q}^2}{\kappa} \right)
\]

Yields: Common Gaussian $p_{\perp}$ spectrum

Heavy quark suppression
Cluster Model


Color flow confined during hadronisation process

- Formation of color-neutral parton clusters

Gluons (color-anticolor) split to quark-antiquark pairs

Clusters decay into 2 hadrons according to phase-space, i.e. isotropically

- no free tuning parameters parton clusters

Very widely used ...
 [default in Herwig/Herwig++]

After: Ellis et al., QCD and Collider Physics
## Hadronisation models summary

<table>
<thead>
<tr>
<th>Model</th>
<th>Pythia6/8 (string)</th>
<th>Herwig/Herwig++ / Sherpa(cluster)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy-mom. picture</td>
<td>powerful</td>
<td>simple</td>
</tr>
<tr>
<td></td>
<td>predictive</td>
<td>unpredictable</td>
</tr>
<tr>
<td>Parameters</td>
<td>few</td>
<td>many</td>
</tr>
<tr>
<td>Flavour composition</td>
<td>messy</td>
<td>simple</td>
</tr>
<tr>
<td></td>
<td>unpredictable</td>
<td>in-between</td>
</tr>
<tr>
<td>Parameters</td>
<td>many</td>
<td>few</td>
</tr>
</tbody>
</table>

[from G.Herten, T.Sjöstrand]
Structure of basic generator process [by order of consideration]

From the 'simple' to the 'complex' or from 'calculable' at large scales to 'modelled' at small scales.

Matrix elements (ME)
1. Hard subprocess: $|M|^2$, Breit Wigners, PDFs
   \[ \bar{q} \rightarrow Z^0 \rightarrow h^0 \]
2. Resonance decays: Includes particle correlations
   \[ Z^0 \rightarrow \mu^+ \mu^- \quad W^+ \rightarrow c \bar{s} \quad W^- \rightarrow \tau^- \bar{\nu}_\tau \]

Parton Shower (PS)
3. Final-state parton showers:
   \[ q \rightarrow qg \quad g \rightarrow gg \quad q \rightarrow q\gamma \]
4. Final-state parton showers:
   \[ q \rightarrow qg \quad g \rightarrow gg \quad q \rightarrow q\gamma \]

B. Di Micco
Università degli Studi di Roma Tre

[from G. Herten]
Conclusions: Structure of basic generator process

From the 'simple' to the 'complex' or from 'calculable' at large scales to 'modelled' at small scales.

Underlying Event (UE)

5. Multi-parton interaction:

6. Beam remnants:

Stable Particle State

7. Hadronisation:

8. Decays:
The DGLAP evolution equation is said to resum large collinear logarithms. So where are these logarithm, and where is the resummation?

Let's perform the integration of the DGLAP equation and expand the result:

\[ f(x, t) = f_0(x) + \int_{t_0}^{t} \frac{dt'}{t'} \frac{\alpha_s(t')}{2\pi} \int_{x}^{1} \frac{dz}{z} P(z) q\left(\frac{x}{z}, t'\right) \]

\[ = f_0(x) + \int_{t_0}^{t} \frac{dt'}{t'} \frac{\alpha_s}{2\pi} \int_{x}^{1} \frac{dz}{z} P(z) \left\{ f_0\left(\frac{x}{z}\right) + \right. \]

\[ + \int_{t_0}^{t'} \frac{dt''}{t''} \frac{\alpha_s}{2\pi} \int_{x/z}^{1} \frac{dz'}{z'} P(z') \left[ f_0\left(\frac{x}{zz'}\right) + \ldots \right] \}

\[ = f_0(x) + \frac{\alpha_s}{2\pi} \ln \left(\frac{t}{t_0}\right) \int_{x}^{1} \frac{dz}{z} P(z) f_0\left(\frac{x}{z}\right) + \]

\[ + \frac{1}{2!} \left[ \frac{\alpha_s}{2\pi} \ln \left(\frac{t}{t_0}\right) \right]^2 \int_{x}^{1} \frac{dz}{z} P(z) \int_{x/z}^{1} \frac{dz'}{z'} P(z') f_0\left(\frac{x}{zz'}\right) + \ldots \]

As suggested by the last step, it is indeed a resummation of all terms proportional to \( \left[ \frac{\alpha_s}{2\pi} \ln \left(\frac{t}{t_0}\right) \right]^n \).