SOFTWARE
SUSTAINABILITY
INSTITUTE

MCnet

Event Generators for Collider Physics

Christian Gütschow

RAL Advanced Graduate Lectures

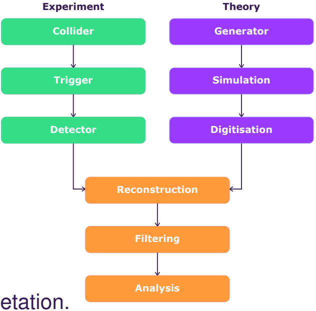
17 June 2026

UCL

Introduction

chris.g@cern.ch

Overview



- Event generation is the simulation of the fundamental particle-collision process.
  - Theory predicts amplitudes (not observables!), detectors measure tracks and calorimeter deposits.
  - Monte Carlo **event generators bridge this gap**
- Indispensable in HEP – from experiment design to interpretation.
- This lecture focuses on the parton-shower generators that produce fully exclusive hadron-level events, ready to be fed into the detector simulation.
- These tools are often treated as black boxes – rarely a good idea!
- A little extra understanding goes a long way, so let's see what we can do . . .

The dream (and why it fails)

- In an ideal world, we'd compute the full event from first principles:
  - Write down the squared amplitude from initial beams to every final-state particle hitting the detector
  - Integrate over phase space, sample, done.
- Quantum field theory gives us the framework – so why not **just do it**?

chris.g@cern.ch

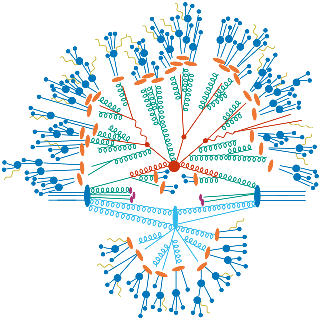
The dream (and why it fails)

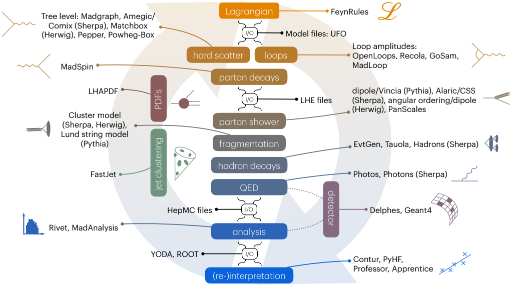
- In an ideal world, we'd compute the full event from first principles:
  - Write down the squared amplitude from initial beams to every final-state particle hitting the detector
  - Integrate over phase space, sample, done.
- Quantum field theory gives us the framework – so why not **just do it**?

- Three obstacles force us back to reality:
  - ✗ **Non-perturbative effects**: the transition to and from hadrons isn't calculable from first principles (beyond even lattice QCD).
  - ✗ **Perturbative tractability**: the number of terms grows factorially – we're limited to < 10 final-state partons, but need hundreds.
  - ✗ **Phase-space dimensionality**: $\sim 3N_{\text{fs}} = \mathcal{O}(1000)$ dimensions, with probability concentrated in unknown islands.

Principled approximation

[arXiv:2605.16036]





Lagrangian — FeynRules

I/O — Model files: UFO

Tree level: Madgraph, Amegic/ Comix (Sherpa), Matchbox (Herwig), Pepper, Powheg-Box — hard scatter

loops — Loop amplitudes: OpenLoops, Recola, GoSam, MadLoop

MadSpin — parton decays

LHAPDF — PDFs

I/O — LHE files

parton shower — dipole/Vincia (Pythia), Alaric/CSS (Sherpa), angular ordering/dipole (Herwig), PanScales

Cluster model (Sherpa, Herwig), Lund string model (Pythia) — fragmentation

hadron decays — EvtGen, Tauola, Hadrons (Sherpa)

FastJet — jet clustering

QED — Photos, Photons (Sherpa)

HepMC files — I/O

detector — Delphes, Geant4

Rivet, MadAnalysis — analysis

YODA, ROOT — I/O

(re-)interpretation — Contur, PyHF, Professor, Apprentice

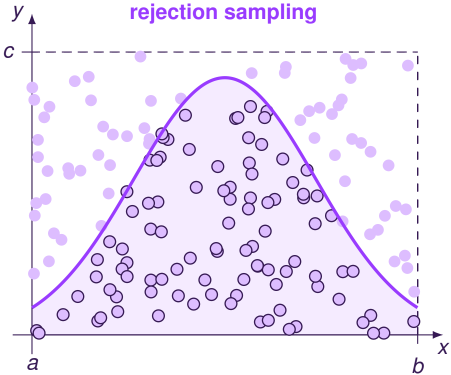
- But good approximations are not far away:
  - Most partonic multiplicity can be emulated by a Markov-chain **parton shower** – iterated splittings in place of full correlations.
  - NP effects yield to modelling too: factorised beam structure, beam remnants, multiple interactions, hadronisation.

✔ Realistic events become possible – via a **modular, multi-stage chain**, each stage addressing one piece of the physics.

From theory to partons

chris.g@cern.ch

Monte Carlo integration



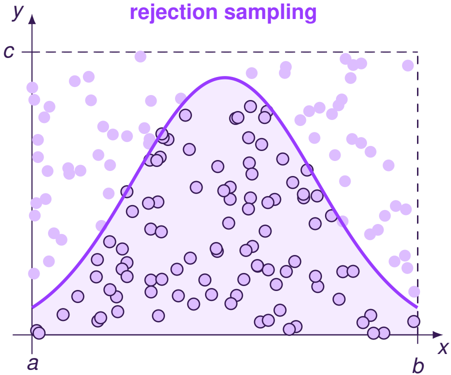
rejection sampling

● rejected (miss): $y > f(x)$

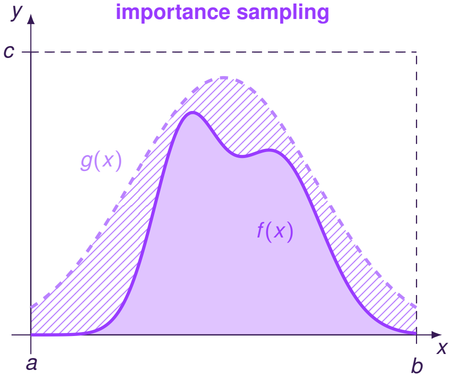
○ accepted (hit): $y < f(x)$

$$\int_a^b f(x)\,\mathrm{d}x \;\approx\; \frac{N_{\text{hits}}}{N_{\text{total}}} \times c\,(b-a)$$

Monte Carlo integration



rejection sampling



importance sampling

- ● rejected (miss): $y > f(x)$
- ○ accepted (hit): $y < f(x)$

$$\int_a^b f(x)\,\mathrm{d}x \;\approx\; \frac{N_{\text{hits}}}{N_{\text{total}}} \times c\,(b-a)$$

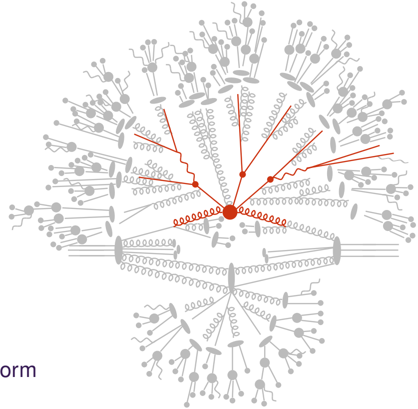
construct simple $g(x) > f(x)$ with $\frac{f(x)}{g(x)} < 1$

draw x from $g(x)$, reject if $\frac{f(x)}{g(x)} < R \in [0,1]$

VEGAS decomposes into multiple channels:

$g(x) = \sum_i a_i\, g_i(x)$

Hard scattering and phase space



- Core of all event generation is an integral of PDFs and partonic matrix elements over phase space:

$$\mathrm{d}\sigma_n^{\mathrm{LO}} = f_a^{h_1}(x_a,\mu_{\mathrm{f}}) f_b^{h_2}(x_b,\mu_{\mathrm{f}}) \times \frac{1}{2\hat{s}} \left|\mathcal{M}_{ab\to n}(\Phi_n)\right|^2 \times \mathrm{d}\Phi_n(\hat{s};\, p_1,\ldots,p_n)$$

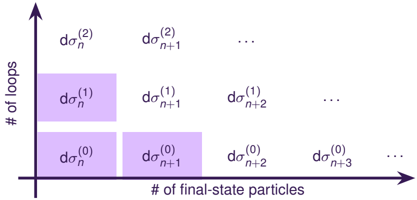
- Integral is done by Monte Carlo sampling
  - integration uncertainty scales as $1/\sqrt{N}$
- Easy to spend all CPU on atypical points!
- Better strategy to perform a Jacobian transform and use importance/adaptive sampling or even ML to ‘learn the space’.
- MC integration can take months, unless massively parallelised.
  - Save integration as ‘gridpack’ for event generation.

chris.g@cern.ch

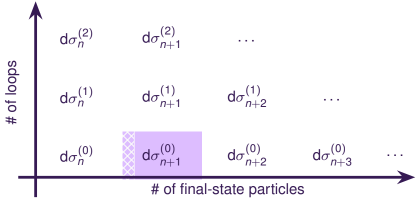
Beyond leading order

- Extra partonic emissions at tree-level increase the final state multiplicity and change the event kinematics directly.
- An NLO cross-section has three parts: $\mathrm{d}\sigma_n^{\mathrm{NLO}} = \mathrm{d}\sigma_n^{(0)} + \mathrm{d}\sigma_n^{(1)} + \mathrm{d}\sigma_{n+1}^{(0)}$
- The LO prediction for the process $\mathcal{P}_{n+1}$ is exactly the real-emission contribution entering the NLO correction to $\mathcal{P}_n$.
- Kinematic cuts (hatched region) are required to ensure that this extra radiation remains resolved.
- Many NLO frameworks, but only few automated
  - aMC@NLO and Sherpa fully automated
  - POWHEG-BOX is a framework to assist manual implementation
  - Virtual terms from dedicated libraries (OpenLoops, BlackHat, NJETS, GoSAM, ...)

$\mathcal{P}_n$ @ NLO



$\mathcal{P}_{n+1}$ @ LO



Splitting functions and resummation

- Partons are not jets: fixed-order partonic events are only comparable to data for IR-safe observables.

- Soft and collinear phase-space (1 GeV $< p_\mathrm{T} \lesssim$ 20 GeV) receives large resummation corrections from multiple QCD emissions.

- Analytic resummation is process-specific but factorises in the collinear limit: cross-section for process σ_0 with parton i to be accompanied by a collinear parton j with momentum fraction z:

$$\mathrm{d}\sigma \approx \sigma_0 \sum_{\text{partons } i} \frac{\alpha_\mathrm{s}}{2\pi} \frac{\mathrm{d}\theta^2}{\theta^2} \,\mathrm{d}z\, P_{ji}(z,\phi)\mathrm{d}\phi$$

where θ is the angle between i and j and

$$P_{qq}(z) = C_F \frac{1+z^2}{1-z} \qquad P_{gq}(z) = C_F \frac{1+(1-z)^2}{z}$$

$$P_{gg}(z) = C_A \frac{z^4+1+(1-z)^4}{z(1-z)} \qquad P_{qg}(z) = T_R(z^2+(1-z)^2)$$

- These are the spin-averaged QCD collinear splitting functions, or DGLAP kernels.

Splitting functions and resummation

- If the emission probability is proportional to

$$C\frac{\mathrm{d}\theta}{\theta}\mathrm{d}z$$

then in a sequence of **strongly ordered** soft/collinear radiations, the probability is given by

$$\prod_{i=0}^{n} C\frac{\mathrm{d}\theta_i}{\theta_i}\mathrm{d}z_i$$

- Quiz time: If integration is the natural continuous limit of a sum:

$$\sum \rightarrow \int$$

What's the equivalent continuum limit of a product?

Splitting functions and resummation

- If the emission probability is proportional to

$$C \frac{\mathrm{d}\theta}{\theta} \mathrm{d}z$$

then in a sequence of **strongly ordered** soft/collinear radiations, the probability is given by

$$\prod_{i=0}^{n} C \frac{\mathrm{d}\theta_i}{\theta_i} \mathrm{d}z_i$$

- Quiz time: If integration is the natural continuous limit of a sum:

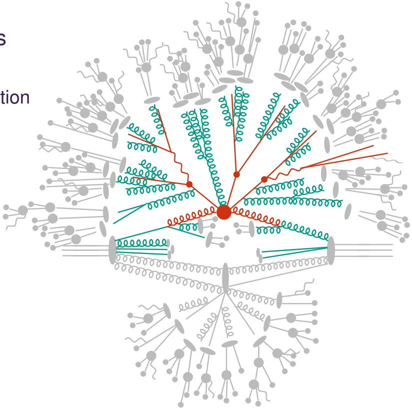
$$\sum \to \int$$

What's the equivalent continuum limit of a product?

$$\prod = \exp\left\{ \sum \ln \right\} \to \exp\left\{ \int \ln \right\}$$

- Every scale-step has a differential probability of no-emission; the **Sudakov exponentiation** multiplies them together.

Parton shower



- Parton showers are Markov chain algorithms based on the QCD splitting functions
  - process-independent, approximate resummation
- From splitting functions, one can calculate the probability of no emission above q_0:
  $\Delta(q_0, Q) = \exp\left\{-C' \int^{Q} \frac{\mathrm{d}\theta}{\theta} \;\int \mathrm{d}z\, P_{ji}(z)\right\}$
  The famous **Sudakov form factor**.
- Main showers used:
  - Pythia8 (most widely used shower model in ATLAS/CMS)
  - Herwig7 (angularly ordered shower with coherent soft emissions)
  - Sherpa's CSS (dipole-style shower used internally)
- Controls event topology and jet multiplicities

Matching

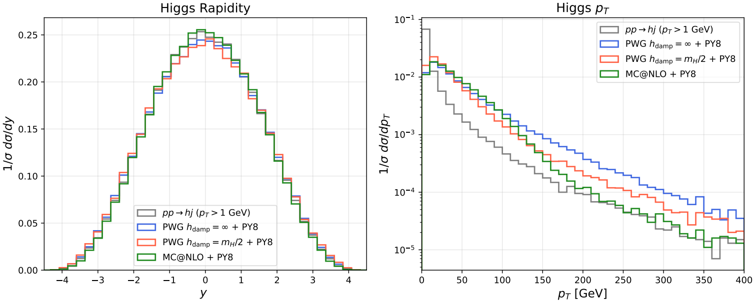
- NLO includes the Born, one-loop, and one-real emission amplitudes (and all their interferences), and smoothly matches the real emissions between shower and ME:

$$\sigma^{\mathrm{NLO}} = \int \mathrm{d}\Phi_n \, [B(\Phi_n) + V(\Phi_n)] + \int \mathrm{d}\Phi_{n+1} \, R(\Phi_{n+1}) \, ,$$

- Main benefit: NLO scale stability – ensures $n+1$ emission shapes now also stable without disrupting formal accuracy of parton shower.
  - rule of thumb: shapes from real, normalisation from virtual
- Matching methods split the $(n+1)$-particle real-emission correction R into an infrared (S) and a hard (H) part: $R = S + H$.
- MC@NLO defines a **subtractive** procedure for $H = R - S$, identifying S with the splitting kernels
  - process specific, prone to negative weights
- POWHEG uses a **multiplicative** procedure to define $S = FR$ such that $H = (1 - F)R$ where F is defined to approach 1 in the infrared-singular limits.
  - shower independent, closer to all-positive weights (but not actually positive only)

chris.g@cern.ch

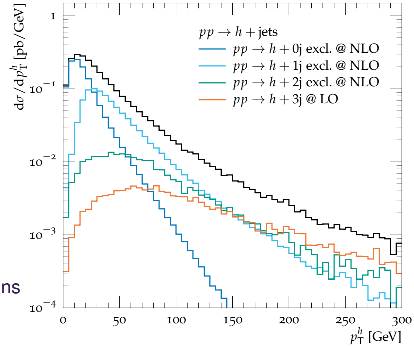
Comparing matching schemes in Higgs production



- Higgs rapidity already well described by LO $pp \to hj$, no sensitivity to extra emissions, but Higgs needs to recoil against a jet for non-zero p_T.
- At low energies, matched distributions feature the so-called “Sudakov peak”, corresponding to the recoil against shower emissions.
- Beyond this, MC@NLO begins to follow the $pp \to hj$@LO calculation, POWHEG is very sensitive to the choice of h_damp in $F = h^2_\mathrm{damp}/(p^2_\mathrm{T} + h^2_\mathrm{damp})$
  - can lead to significant effects away from the shower-dominated region

chris.g@cern.ch

Multi-jet merging



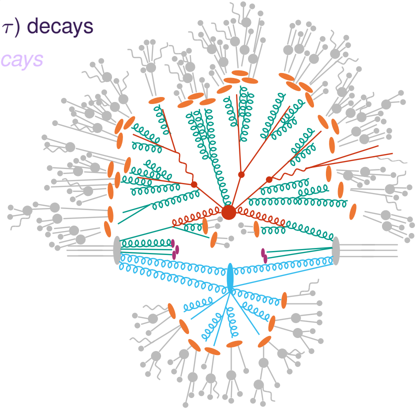
- Consistently combine matrix-elements for a range of final-state multiplicities.
  - Let the shower freely populate the soft region below a **merging cut** Q_{cut}, leaving the hard domain above this cut to the matrix elements.
  - Cluster the partons according to the inverse splitting probabilities of the shower.
  - Evaluate no-emission probability between successive splittings, making these contributions exclusive in the resolved emissions.
- CKKW(-L) and MLM schemes designed to replace the shower's collinear splitting functions with proper matrix elements in the relevant (hard) phase space
- Extended to automatic merging of many NLO and/or LO multi-leg matrix elements in MEPS@NLO (Sherpa) and FxFx (MG5_aMC@NLO).
- **Huge book-keeping exercise!**

Non-perturbative modelling

chris.g@cern.ch

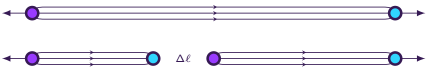
From partons to hadrons

Fragmentation = (Final-state) parton showers
Hadronisation
Prompt hadron (and τ) decays
Non-prompt decays

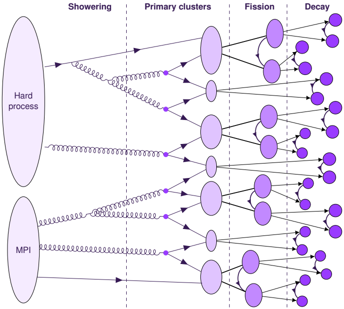


- Hadronisation = non-perturbative transition from partons to hadrons around Λ_{QCD}
- Underlying event = beam remnants + multiple parton interactions (MPI)
- Main models used:
  - Pythia8: Lund string model + MPI model (interleaved with ISR/FSR)
  - Herwig7: Cluster model + MPI
  - Sherpa: Cluster hadronisation + UE model; plugin to Pythia8 for Lund-style hadronisation
- Requires **model tuning** to MB/UE observables (ATLAS A14/AZ, CMS CP5/CH3 etc.)

Hadronisation models



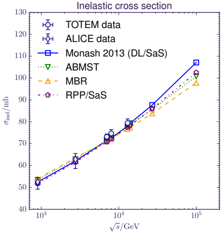
- Lund string (Pythia):
  - Inspired by linear scaling of QCD potential at large distances
  - Break colour strings to produce new quark pairs; gluons form kinks in strings
  - Lorentz invariance and left-right symmetry give Lund Symmetric Fragmentation Function
  - Kinematics well-described, but flavour (especially baryons) not natural
- Cluster hadronisation (Herwig, Sherpa):
  - Colour preconfinement, seen in colour-connected neighbour parton mass spectrum
  - Non-perturbative $g \to q\bar{q}$, then cluster colour singlets: requires finite gluon constituent mass
  - Clusters treated as meson resonances
- Both models also contain colour reconnection heuristics
- Requires **model tuning**!

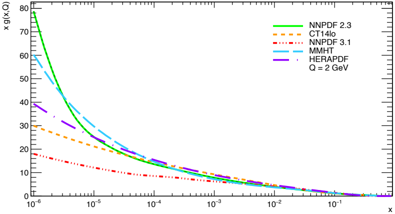


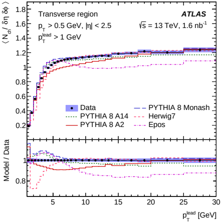
Underlying event and MPI

[arXiv:1808.07224]

[arXiv:1701.05390]

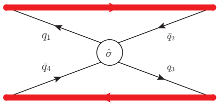


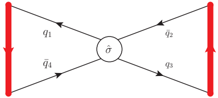




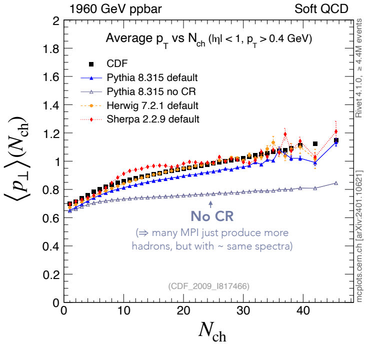
- Inelastic cross-section rises strongly with centre-of-mass energy due to low-x PDFs
- We can think of two colliding hadrons as two beams of partons colliding
  - the probability for a high-p_T interaction is low (→ calculate using perturbation theory)
  - the probability for low-p_T scatterings is very large (dominated by t-channel gluon exchanges)
- Sample $\langle n_{\text{parton-parton}} \rangle \sim \sigma_{\text{parton-parton}} / \sigma_{\text{hadron-hadron}}$ interactions from a Poisson
- Effective matter overlap between the two hadrons characterised by “overlap factors”
  - small $\langle n_{\text{parton-parton}} \rangle$ in peripheral collisions, large $\langle n_{\text{parton-parton}} \rangle$ in central collisions
- **Pedestal effect** in UE as “minimum bias” sample transitions to a “maximum bias” one
  - Not possible to be more central than fully central!
- Requires **model tuning**!

Colour reconnection





- Correlations in colour space could cause non-trivial rearrangements of the hadronising systems.
- Most modern CR models seek to minimise a measure of the total potential energy of the hadronising system.
  - e.g. total string length in string CR models or (summed) cluster masses in cluster CR models
- Such minimisations typically result in fewer but more energetic hadrons.
- Shower radiation patterns could also be induced by colour connections to partons outside the resonance-decay systems (e.g. top quarks).



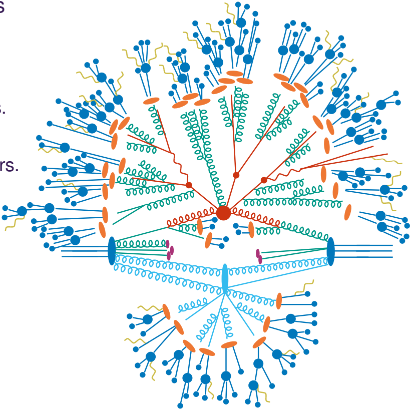
No CR

(⇒ many MPI just produce more hadrons, but with ~ same spectra)

chris.g@cern.ch

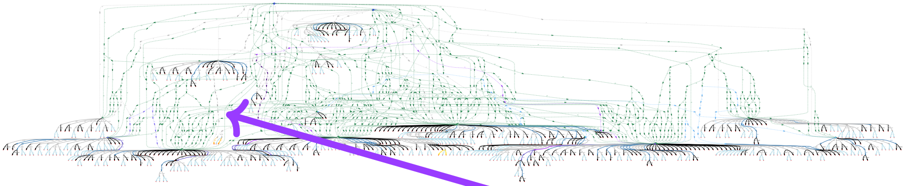
Hadron/tau decay modelling

- Accurate modelling of hadron decays can be complex, but in many cases sufficiently accurate results can be obtained with tabulated decay modes and branching ratios.
  - Decay kinematics sampled (in the decaying particle's rest frame) from the appropriate dimensionality of phase-space configurations.
  - Models with more specific matrix-element influences can be implemented as form factors.
- Main tools used:
  - Pythia/Herwig/Sherpa's internal decayers
  - EvtGen (used for *B*-hadron decays with custom decay tables on top of Pythia/Herwig-based samples)
  - Tauola (optional for τ decays with polarisation modelling used on top of Pythia/Herwig-based samples)
- Spin correlations sometimes retained depending on interface

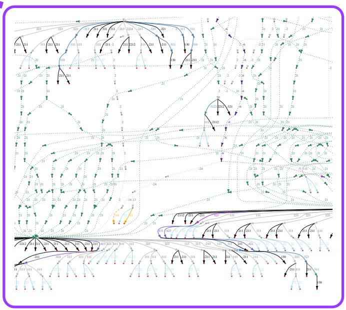


Using Monte Carlo events

The HepMC event graph



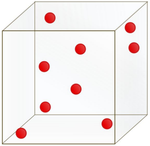
- A MC event graph showing all the particles (lines) and interactions (vertices) in the Pythia event record.
- Guess the event type?
  (Hint: 2212 = proton, 1–6 = quarks,
  11–16 = leptons, 21–25 = bosons)
- Ouff . . . how to safely analyse *that*?
  - There is no unambiguous truth in quantum mechanics!
  - Event records are half physics, half debug info
    – and zero indication of amplitude interference!
  - Kinematic frames aren't defined until the final state, momentum not necessarily conserved at all vertices.
  - **Beware!**



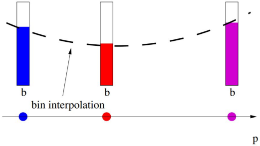
chris.g@cern.ch

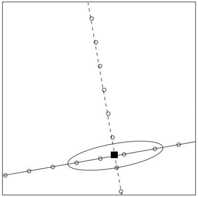
Model tuning

- Tuning was historically a mix of brute force and expert intuition.
- Professor/Apprentice accelerates convergence by building surrogate models:
  - Sample parameter points p_n from a hypercube/sphere.
  - Generate MC run-sets (for beams, processes, etc.) at each p_n.
  - Run jobs in parallel on batch/grid systems to produce histograms.
  - Fit surrogate models $\hat{b}(p)$ for each histogram bin using the MC outputs – traditionally 3rd/4th-order polynomials, though other form are possible (e.g. rational approximants via Apprentice).
  - Construct a surrogate goodness-of-fit from these models and optimise the parameters efficiently.
- Expert knowledge is still essential – but surrogates dramatically reduce the cost of scanning high-dimensional parameter spaces.
- Machine learning? Sure – but if polynomials (possibly after a variable transformation) capture the behaviour well, they remain simple, transparent, and robust.



[Andy Buckley]

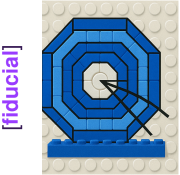




Tactics for tuning

- Factorise the parameter space
  - Traditionally split hadron flavours/spectra, jet structure, event topologies, underlying event – typically $\mathcal{O}(10)$ parameters
  - Approximate but practical; reduces dimensionality and stabilises fits.
  - Possible to automate grouping via mutual sensitivities / parameter-bin influence matrices.
- Weighting, observable balance, and uncertainties
  - Some data types dominate the fit unless reweighted — balance is essential.
  - Models cannot fully describe all bins: examine envelopes, sensitivity maps, parameter ranges, and consider per-bin weighting.
  - Custom goodness-of-fit is common, but regularisation weakens strict statistical interpretation.
  - Even "χ^2" is non-classical in practice: eigentunes, empirical tolerances... still room for more principled approaches.
- Future directions
  - Improved treatment of heavy flavour; incorporating matching/merging systematics.
  - Systematic-uncertainty handling via event weights.
  - More flexible surrogate models and robust optimisation strategies.

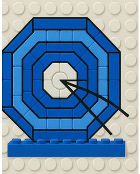
Fiducial cross-sections: simple idea, big impact



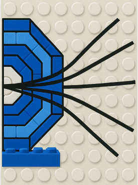
- A surprisingly powerful idea
  - 👍 Define cross-sections using observable final-state particles, within a clearly specified kinematic region.
  - 👍 Makes it easy to reproduce key plots, enabling real understanding, catching issues early, and improving MCs.
  - 👍 Establishes a shared language between theory and experiment – essential for tuning, fits, and reinterpretation.

Fiducial cross-sections: simple idea, big impact

- A surprisingly powerful idea
  - 👍 Define cross-sections using observable final-state particles, within a clearly specified kinematic region.
  - 👍 Makes it easy to reproduce key plots, enabling real understanding, catching issues early, and improving MCs.
  - 👍 Establishes a shared language between theory and experiment – essential for tuning, fits, and reinterpretation.
- But it's tempting to cheat . . .
  - 👎 Partons, bosons, and other "truth" objects in the event record look easy to use.
  - 👎 In practice, they're often ambiguous, model-dependent, or even non-existent (e.g. in higher-order simulations).
  - ⚠ Focus on physical final states (hadrons, leptons, photons) as the reliable basis for comparisons.



[fiducial]



[extrapolated]

Fiducial cross-sections: simple idea, big impact

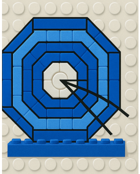
➜ A surprisingly powerful idea

- 👍 Define cross-sections using observable final-state particles, within a clearly specified kinematic region.
- 👍 Makes it easy to reproduce key plots, enabling real understanding, catching issues early, and improving MCs.
- 👍 Establishes a shared language between theory and experiment – essential for tuning, fits, and reinterpretation.

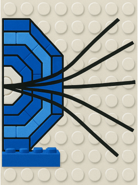
➜ But it’s tempting to cheat . . .

- 👎 Partons, bosons, and other “truth" objects in the event record look easy to use.
- 👎 In practice, they’re often ambiguous, model-dependent, or even non-existent (e.g. in higher-order simulations).
- ⚠ Focus on physical final states (hadrons, leptons, photons) as the reliable basis for comparisons.

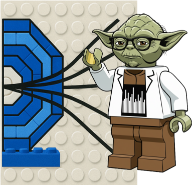
😊 Who knows – **nature might still surprise us**!



[fiducial]



[extrapolated]



[nature]

Why it matters

- Standardisation sounds boring – but agreeing on things like status codes PDG IDs, and weight schemes saves huge headaches later.
- Avoiding unphysical shortcuts has led to genuine physics insight, including:
  - Recognition of hadronisation as a “decoherence barrier”, guiding analysis definitions
  - Better truth tagging strategies: e.g. using fragmenting heavy-flavour hadrons rather than ancestry of soft partons
  - The shift to dressed leptons by default, i.e. truth leptons with their photon halos attached
- Common pitfalls to avoid:
  - Coloured objects aren’t final-state particles: “final-state tops” aren’t physical!
  - Electroweak bosons aren’t final states either: prefer leptonic or hadronic decay products.
  - Missing energy $\neq$ sum of neutrino momenta: use particle-level missing transverse momentum instead.
  - Hidden cuts hide physics: all vetoes and selections should be encoded in the fiducial definition, not just the code.

🔗 [arXiv:1901.05407]

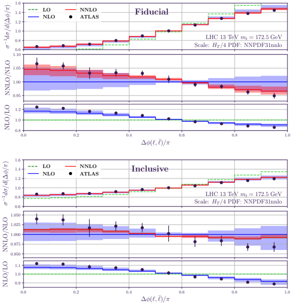
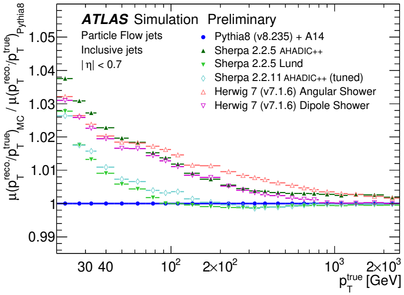


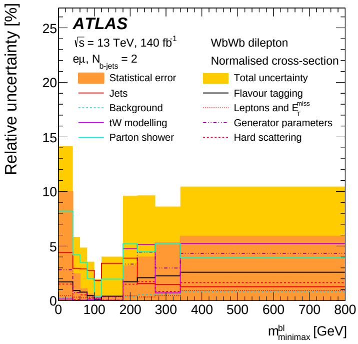
FIG. 1: NNLO QCD predictions for the fiducial (top) and inclusive selections (bottom) of the normalized $\Delta\phi_{\ell\ell}$ distribution versus ATLAS data [20]. Uncertainty bands are from 7-point scale variation.

Generator uncertainties

- Main systematics: scale variation, PDF choice, matching/merging systematics
- Alternative generators often used to estimate modelling systematics (2-point systematics)
- Tuning systematics: UE/hadronisation model dependence (often neglected)

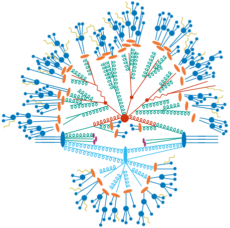


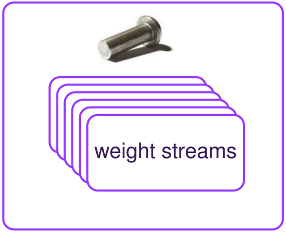
[ATL-PHYS-PUB-2022-021]

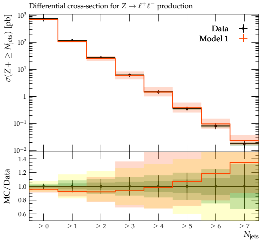


[arXiv:2506.14700]

Reweighting techniques







- Compute alternative event weights during generation for:
  - Renormalisation/factorisation scale variations (e.g. $\mu_\text{r}, \mu_\text{f} = 0.5, 1, 2\times$ central)
  - PDF error sets (e.g. eigenvectors or replicas from CT18, NNPDF, MSHT etc.)
  - NLO electroweak virtual corrections
- Advantages:
  - Exact same phase-space sampling for all variations → reduced statistical noise
  - No need to rerun generation → huge cost saving

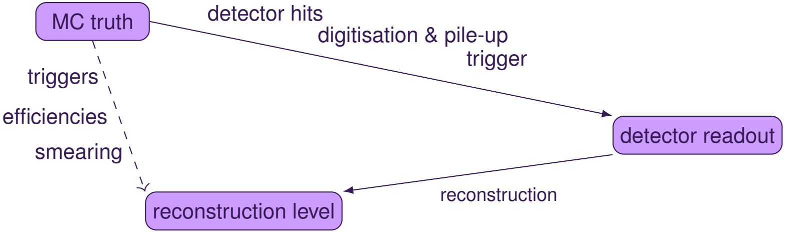
chris.g@cern.ch

Monte Carlo production at scale

chris.g@cern.ch

Event generators in experiment data processing

- Typical experimental use of generators is to feed their output into a detector simulation (e.g. based on Geant4)
- Then apply the same reconstruction & analysis as for data:

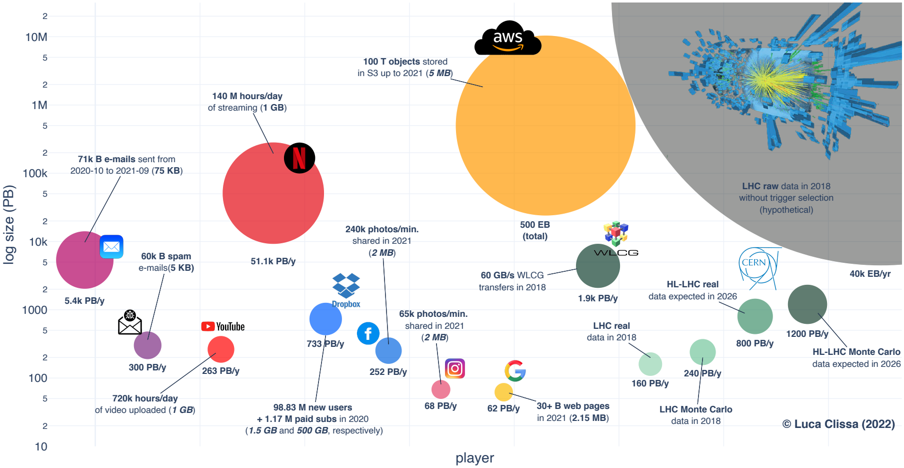


- The generator bit of this chain was historically cheap.
  - CPU/memory requirements much less than detector geometry, *B*-field stepping, material interaction and secondaries.

⚠ **Not true these days!**

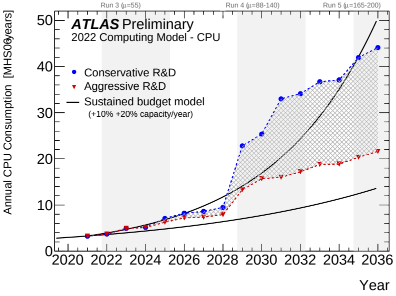
LHC = Big Data

[arXiv:2202.07659]

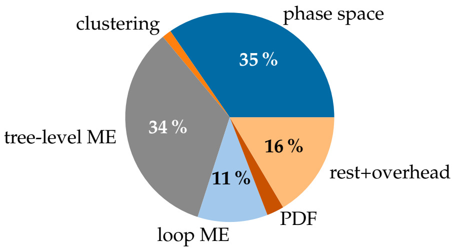


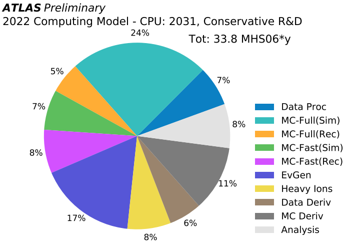
Expected computing requirements

- ➜ Projected evolution of computing resources saw cost of event generation on par with detector simulation
- ➜ LHC measurements in danger of being limited by Monte Carlo statistics



$$pp \to e^+e^- + 0{,}1{,}2j@\text{NLO} + 3{,}4{,}5j@\text{LO}$$



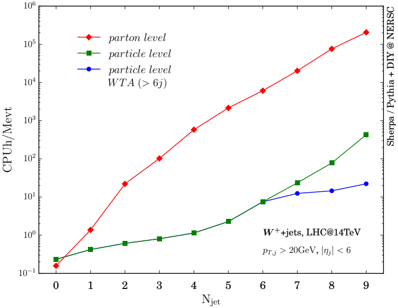


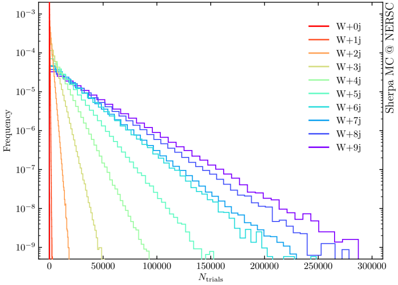
[CERN-LHCC-2022-005]

$$\mathrm{d}\sigma_{pp\to n} = f_a(x_a,\mu_\mathrm{f}^2)f_b(x_b,\mu_\mathrm{f}^2) \times |\mathcal{M}_{ab\to n}|^2 \times \mathrm{d}\Phi_n(p_1,\ldots,p_n)$$

Understanding the cost of precision Monte Carlo

- Cost of multi-leg production **dominated by parton level** matrix elements.
  - Number of diagrams grows factorially with every additional emission (at best exponentially when exploiting recursions a la Comix).
- High-multiplicity matrix elements lend themselves well to GPU-based bulk production.
- Low-multiplicity matrix elements cheaper to regenerate entirely than to store on disk.



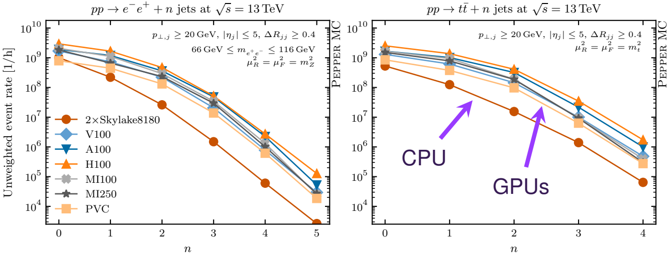


[PRD 100 (2019) 1]

chris.g@cern.ch

Architecture-based acceleration

- Computing architectures are shifting rapidly, driven in part by AI-driven hardware trends
  - HPC is entering the exascale era, requiring both scalability and portability across diverse platforms
- Pepper and MadGraph4GPU are designed with these goals in mind, using:
  - MPI for parallelisation, HDF5 for I/O, and Kokkos for performance portability
- These tools parallelise the entire parton-level event generation chain
  - Supports both multi-core CPUs and a wide range of GPU architectures



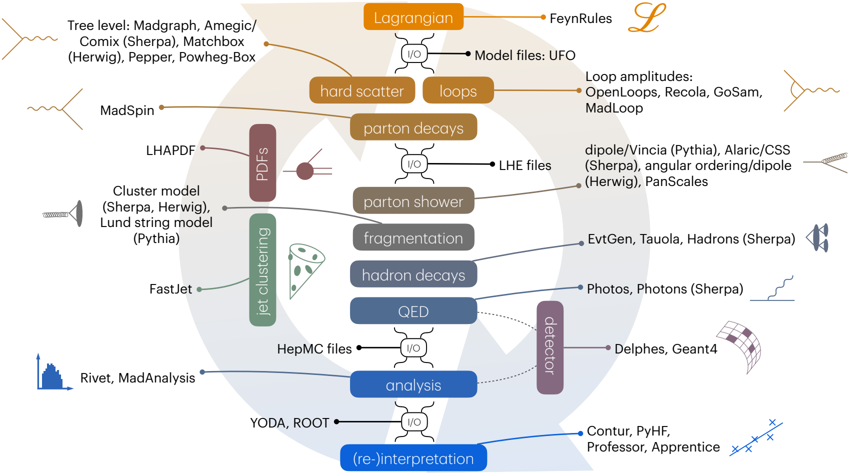
[SciPost Phys. 17 (2024) 3, 081]

Final thoughts

Open-source tools & community infrastructure

- Monte Carlo (support) tools are community-driven and open-source
  - Transparent, inspectable, and modifiable code
- Widely used and maintained by both theorists and experimentalists:
  - FastJet: Standard library for jet clustering in collider physics
  - LHAPDF: Standardised interface for accessing and interpolating PDF sets
  - HepMC: Common event record format for storing full generator output
  - Rivet: Portable, experiment-independent analysis toolkit for generator validation
  - YODA: Histogramming and statistical accumulator library used by Rivet
  - Contur: Constraint setting from LHC measurements and beyond using Rivet
  - Professor/Apprentice: Generator tuning framework using fast parameterisation
  - HEPData: Archival and retrieval of experimental data for reuse in MC validation
- Shared infrastructure = less duplication, better validation, reproducibility
- Open tools = easier training and broader contributions
- Cross-talk between experiments, theorists, and developers is essential.

The Monte Carlo Ecosystem in High-Energy Physics



Lagrangian

I/O

hard scatter

loops

parton decays

I/O

parton shower

fragmentation

hadron decays

QED

I/O

analysis

I/O

(re-)interpretation

PDFs

jet clustering

detector

FeynRules

Model files: UFO

Tree level: Madgraph, Amegic/
Comix (Sherpa), Matchbox
(Herwig), Pepper, Powheg-Box

Loop amplitudes:
OpenLoops, Recola, GoSam,
MadLoop

MadSpin

LHAPDF

LHE files

dipole/Vincia (Pythia), Alaric/CSS
(Sherpa), angular ordering/dipole
(Herwig), PanScales

Cluster model
(Sherpa, Herwig),
Lund string model
(Pythia)

EvtGen, Tauola, Hadrons (Sherpa)

FastJet

Photos, Photons (Sherpa)

HepMC files

Delphes, Geant4

Rivet, MadAnalysis

YODA, ROOT

Contur, PyHF,
Professor, Apprentice

🔗 [arXiv:2605.16036]

Summary

- Event generation is **principled approximation**: a modular chain turning theory into detector-ready events
    - **Hard process**: matrix elements at LO/NLO, matched and merged with the shower
    - **Parton shower**: resums the dominant soft/collinear radiation
    - **Hadronisation + UE**: non-perturbative models (string, cluster), tuned to data
    - **Decays, detector, analysis**: the path to something comparable with measurement
- Every stage carries **choices and uncertainties** – not a black box!
    - Know which knobs you're turning, and which sources of uncertainty they probe.
- These tools are long-lived, shared infrastructure – you'll use them for years, and you can contribute back!
- Further reading:
    - **MCnet review** (2011) 🔗 [**arxiv:1101.2599**]
    - **Höche review** (2014) 🔗 [**arxiv:1411.4085**]
    - **Practical Collider Physics** (2021) 🔗 [**Buckley, White ($\times$2)**]
    - **MC primer** (2026) 🔗 [**arxiv:2605.16036**]

MCnet short-term studentships

- 3-month PhD placements with MC developer teams, in collaboration with the LHC Physics Centre at CERN
- What is it?
  - A chance to work closely with MC generator developers on a focused project.
  - Similar to past MCnet-funded short-term placements — but now open more broadly.
  - Not limited to MCnet member nodes — external projects and students welcome.
- Who can apply?
  - PhD students interested in event generators, theory/phenomenology, or experimental MC use.
  - Especially suited for students looking to deepen their expertise in LHC simulation tools.
- More info & available projects [**click me**]
  - New rounds of placements will be announced 2–3 times per year.
  - Feel free to reach out with ideas for new projects – external proposals encouraged!

Backup

QED corrections

- Collinear QED radiation affecting lepton isolation and mass resolution
- Main tools used:
  - Pythia8 / Herwig7 include basic QED FSR in their shower, good enough for many purposes
  - Photos used as afterburner to Herwig/Pythia-based precision samples
  - Sherpa's YFS-style soft photon resummation
- Important for precision Drell-Yan and 4ℓ production etc.
- Can shift kinematic distributions subtly – affects fiducial phase-space matching!