OVERVIEW OF GENERATORS USED IN ATLAS.

James Robinson[1]

[1] DESY

Monte Carlo in ATLAS Tutorial, 28th September 2015

HELMHOLTZ ASSOCIATION

ATLAS EXPERIMENT

DESY

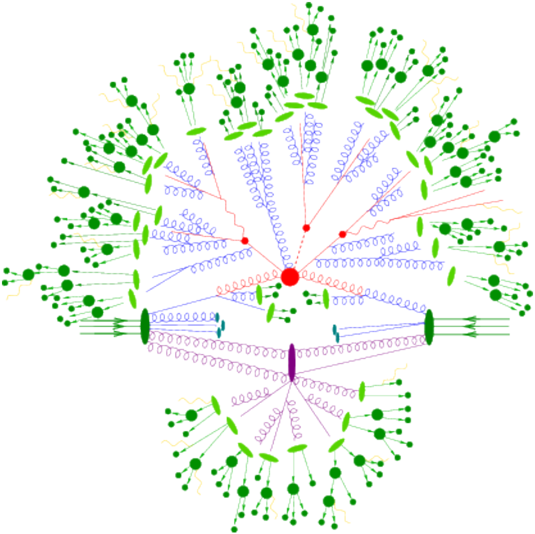
Introduction

I'll briefly discuss how we (ATLAS) use Monte Carlo generators

- designed to be informal
- please ask questions as we go along
- understanding more important than covering all material

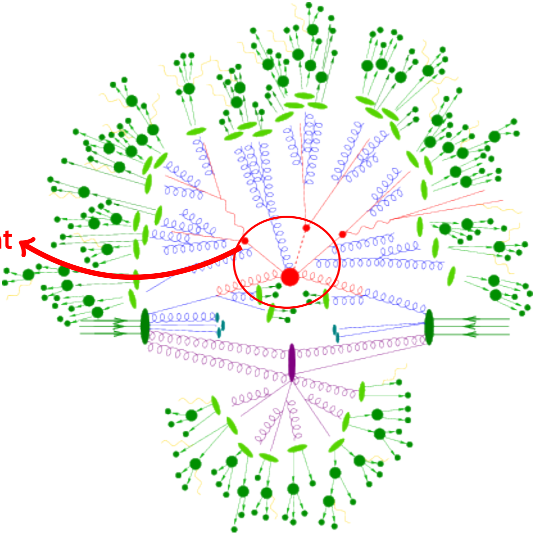
Monte Carlo basics

Simplistic Monte Carlo event structure

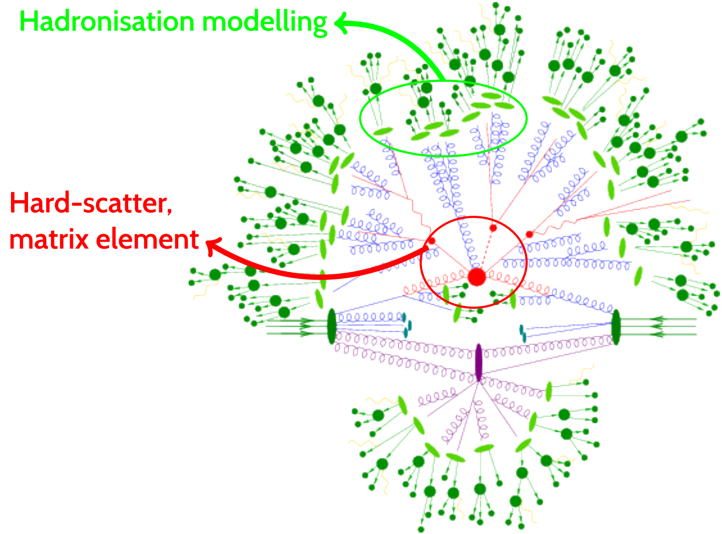


Simplistic Monte Carlo event structure

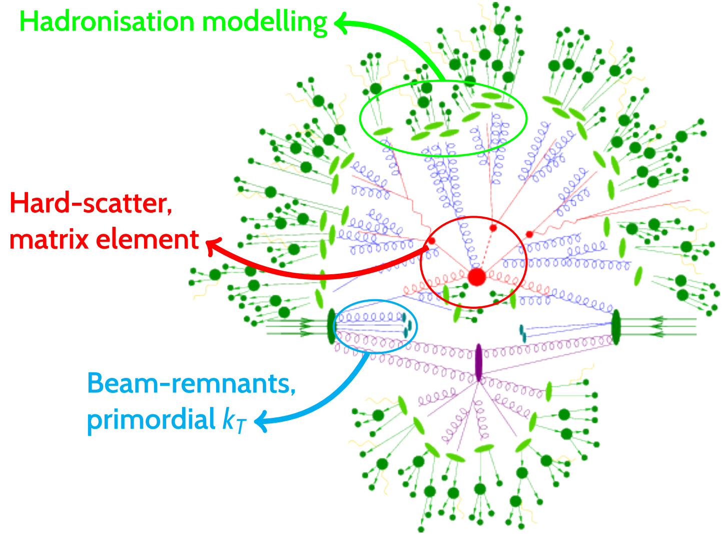
Hard-scatter,
matrix element



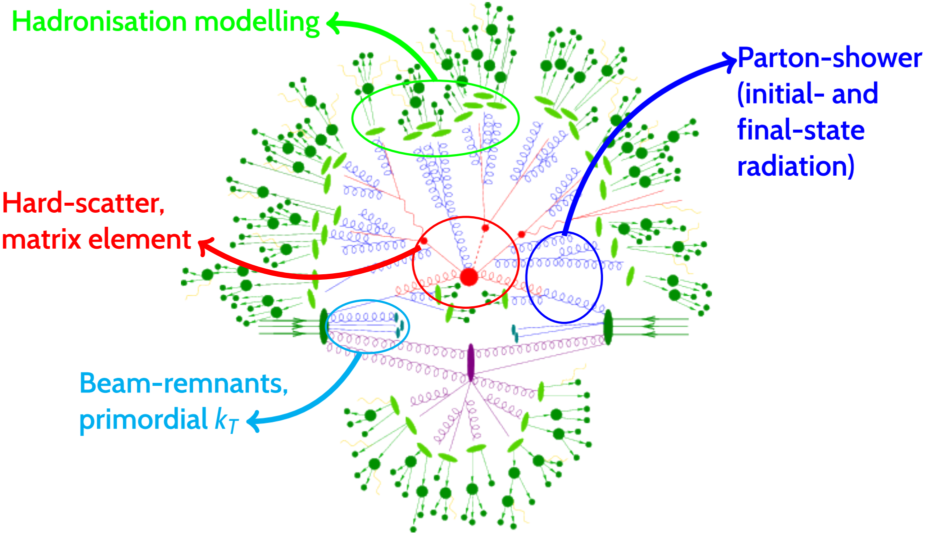
Simplistic Monte Carlo event structure



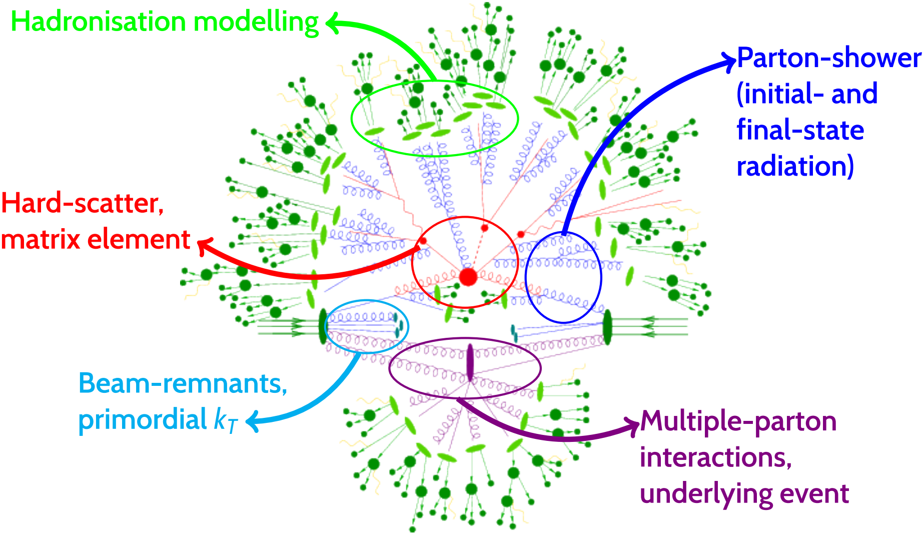
Simplistic Monte Carlo event structure



Simplistic Monte Carlo event structure

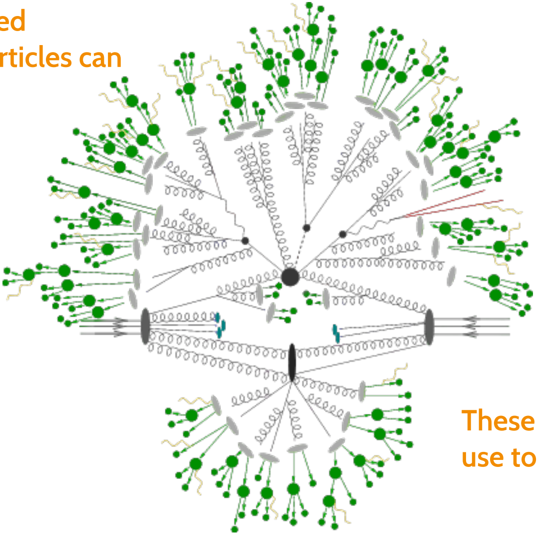


Simplistic Monte Carlo event structure



Simplistic Monte Carlo event structure

Only long-lived final-state particles can be observed



These are what we can use to compare to data

What's in a Monte Carlo event?

- **Hard-scatter** (aka matrix element):
  - exact theoretical calculation up to stated accuracy (e.g. LO or NLO)
- **Parton Shower**:
  - QCD radiation matched to the matrix element (bremsstrahlung)
- **Hadronisation**/**beam-remnants**/**MPI**:
  - phenomenological models describing non-perturbative effects

- Higher-order calculations blur these distinctions
- Complicated interplay between ME and PS
- Solutions: merging and matching (eg. CKKW, MLM)

Factorisation ansatz

To simplify event generation, we assume:

- cross section and event structure depend on hard-scatter
- parton showers/hadronisation happen at lower (softer) scales

Factorisation ansatz

Assuming the hard and soft scales are separable
$\rightarrow$ we can dress the events without changing the cross section

Monte Carlo generators and QCD

- All measurements at ATLAS need an understanding of QCD
- Even channels like $H \rightarrow 4\mu$ are affected by QCD
- Any observable prediction needs QCD corrections
- QCD is hard!

Monte Carlo generators are all about QCD

Why is QCD difficult?

- α_S is large ($\gtrsim 0.1$)
- gluon self-coupling gives us lots of gluons
- measurements rely on detecting hadrons (don't try to measure partons!)
- hadron production is non-perturbative

$\boxed{\text{We need good models for parton showering and hadronisation}}$

Matrix elements

Matrix elements

- The basis of any event generation is a $2 \to n$ matrix element
- Can (in principle) be obtained trivially from the Langrangian

Simple, right?

Matrix elements

- The basis of any event generation is a $2 \rightarrow n$ matrix element
- Can (in principle) be obtained trivially from the Langrangian

Caveats

- divergences at tree-level from soft and collinear partons
- beyond leading-order loops can give infinities
- number of possible diagrams grows exponentially with n

Tree-level calculations

Simple $2 \to 2$ matrix elements

- can be calculated from relevant Feynman diagrams
- fairly easy to generate (beware soft/collinear divergences)

Higher-order tree-level matrix elements

- large number of diagrams, but can be automated
- increasingly slow as n increases
- more divergences $\to$ harder to sample phase space

Generators such as ALPGEN and MADGRAPH give $2 \to n$ topologies

Next-to-leading order (NLO)

- Each of these tree-level calculations is inclusive
- $pp \to W + 1j$ means W plus at least one jet
- There is therefore overlap with the $pp \to W + 2j$ phase space

- Naïve combination would double-count emissions
- Correctly combining involves matching emissions
- Allows event production at next-to-leading-order (NLO)

Generators such as POWHEG and AMC@NLO generate at NLO

NLO or LO depends on what you're measuring

- Example: $\Delta\phi_{jj}$ for events generated as $W + 1j$ at NLO
  $\to$ clearly leading-order in this observable

Parton showers

Resummation

- Leading order for a given observable is:
  $\rightarrow$ lowest order in α_S which gives a non-zero cross section
- Usually the expansion in α_S does not converge quickly (or at all)
  $\rightarrow$ need to resum the additional terms

- Tree-level generators give inclusive events
- NLO generators give one extra parton

We want to approximate all terms instead of explicitly calculating

This is what parton showers can do

Parton showers

- Generate real, exclusive events down to low (but still perturbative) scale
- Order emissions in some scale ρ: $\rho_1 > \rho_2 > \rho_3 \ldots$
- Use $1 \to 2$ splitting kernels (usually DGLAP)
- In order to guarantee exclusivity:
  $\to$ multiply by probability of no emission above current scale

Without strong ordering, PS assumption breaks down

Parton shower evolution

Which variable should we order emissions in?

- PYTHIA pre-6.4, (old) SHERPA
  - virtuality ordering (Q^2): simplest conceptually
- HERWIG and HERWIG++
  - angular ordering: reduces soft gluon emissions
- ARIADNE, PYTHIA 6.4+, (new) SHERPA
  - colour dipoles: replace $1 \to 2$ splitting with $2 \to 3$
  - allows all partons to be on-shell throughout shower

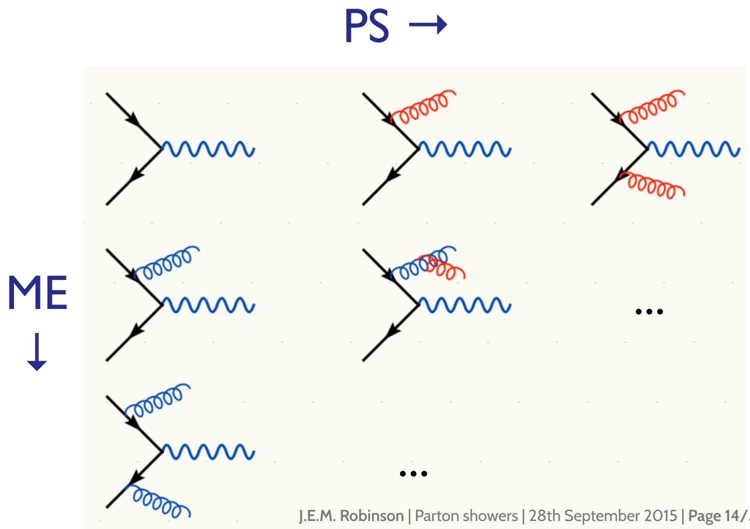
Matching and merging

- Fixed-order tree-level matrix-element generators:
  - first *n* orders in α_S exactly
  - good for: a few, hard, well-separated partons
  - bad for: many, soft/collinear partons
- Parton showers:
  - approximate (N)LL terms to all orders in α_S
  - good for: many, soft/collinear partons
  - bad for: a few, hard, well-separated partons

Can we get the best of both?

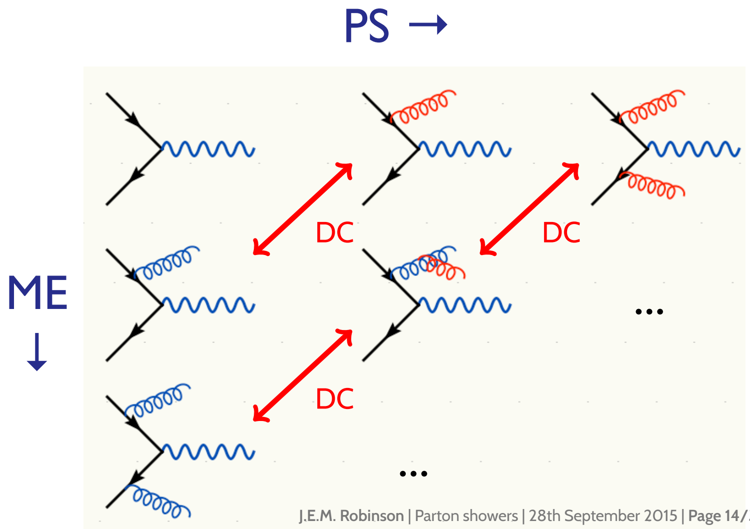
Merging matrix elements with a parton shower

Leading order extended by PS and/or higher order ME



Merging matrix elements with a parton shower

Some topologies are double counted

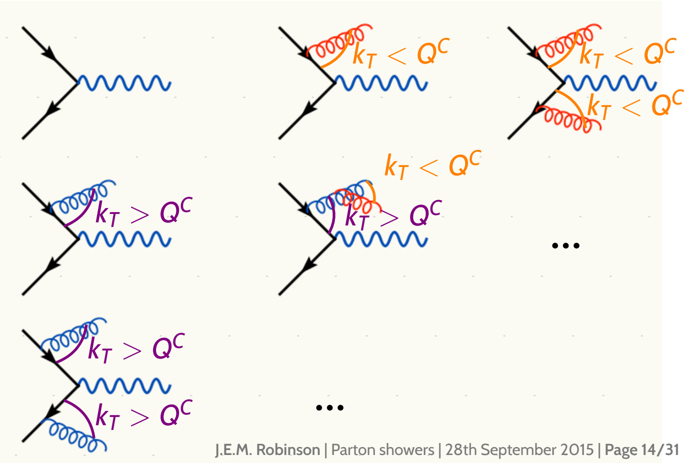


Merging matrix elements with a parton shower

Avoided using phase space cut: **ME** above cut; **PS** below

PS →

ME ↓



Matching matrix elements with a parton shower

- Merging solves the double counting problem
- Creates possible dependence on merging scale, Q^C
- Need to match the ME to the PS at Q^C

Matching matrix elements with a parton shower

- Merging solves the double counting problem
- Creates possible dependence on merging scale, Q^C
- Need to match the ME to the PS at Q^C

MLM matching (other prescriptions exist)

- generate ME events using parton-level cuts
- cluster back to a $2 \to 2$ process
- run the parton shower from this scale
- accept event if N_{jets} above Q^C is the same with and without PS

Independent of the details of the process and/or the shower

ATLAS interfaces between hard-scatter and shower

- Separate steps
  - LHE files produced by the ME generator
  - Showering step run over these files later
- On-the-fly
  - As above but with both steps in a single job
- Integrated
  - Some generators can do both steps within a single code base
    $\rightarrow$ internal HERWIG++ implementation of POWHEG

more details in Dan Hayden's talk

ATLAS afterburners

- Dedicated generators that more accurately model certain decays
  - EvtGen: *b*-hadron decays
  - Photos: photon correlations
  - Tauola: τ-lepton decays
- Some care is needed when using these
  - Only if final-state correlations are important for analysis
  - Ensure they improve on the native generator handling (not guaranteed)

more details in Dan Hayden's talk

Underlying event

What is underlying event?

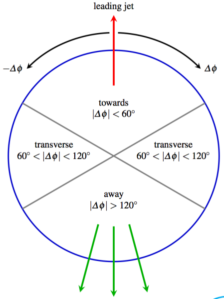
Any hadronic activity not associated with hard-scatter

- Unavoidable background in collision events
- Not well-predicted as non-perturbative effects dominate
- Need to ensure measurements not dependent on modelling

Cannot unambiguously assign particles to the hard scatter or UE

Typically modelled with

- Multiple parton interactions
- Initial/final-state radiation
- Constrained by data



Underlying event modelling

Early attempts: non-perturbative model (default in HERWIG)

- Assume that whole of two beam-remnants interact coherently

Current: perturbative models

- Assume dominated by local parton-parton interactions
- Colour Reconnection model: HERWIG++ and JIMMY
  - Partonic scatters separated into "hard" and "soft" at $\sim$5 GeV
  - Include colour correlation between scatters
- Interleaved shower: PYTHIA
  - Evolve shower in p_{T} allowing ISR or additional scatters
  - Gives colour conenction between MPI and ISR

Hadronisation

Hadronisation

- Remember that we can't ever observe partons
- To be useful, Monte Carlo generators must replicate what we see

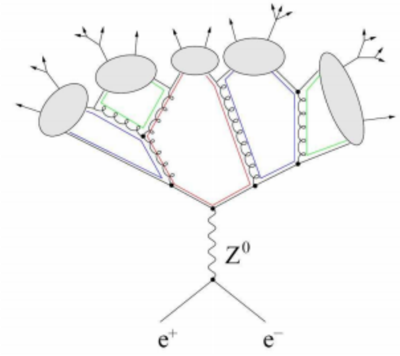
$$\boxed{\text{observable particles} \rightarrow \text{long-lived hadrons}}$$

- Non-perturbative process $\rightarrow$ empirical models
- Informed by our knowledge of non-perturbative QCD

Cluster hadronisation

Based on idea of pre-confinement

- gluons are emitted mainly between colour-connected partons
- with enough gluons, colour-dipoles will be small

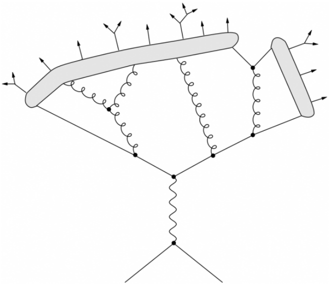


- force $g \to q\bar{q}$ splittings after the parton shower
- construct low-mass, colour-singlet clusters

String hadronisation

- QCD is Coulomb-like at small distances
- field lines are compressed at large distances

- model as massless relativistic string
- as each $q\bar{q}$ pair moves apart more energy is stored in the string
- eventually the string breaks and creates a new $q\bar{q}$ pair
- more energy between $gq/\bar{q}$ than $q\bar{q}$ $\rightarrow$ gluons form kinks on the string



Hadronisation model comparison

Model	String	Cluster
energy-momentum	powerful, predictive few parameters	simple, unpredictive many parameters
flavour composition	messy, unpredictive many parameters	simple, predictive few parameters

Both models have advantages/disadvantages

Monte Carlo Tuning

The need for tuned generators

Data constrains free parameters in non-perturbative models

- Pileup simulation
  - tuned to data with very inclusive triggers (minimum bias)
- Calibration
  - jet/τ identification and substructure
- Unfolding
  - correct for detector effects: need to reduce model dependence
- Background estimates in analysis
  - used either directly or through extrapolation from control regions

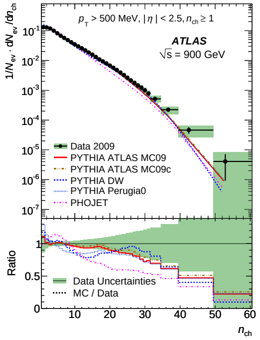
Reliable tunes essential for precision measurements/discoveries

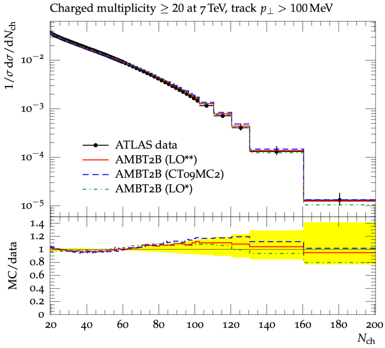
Tuning workflow

1. Choose generator parameters of interest (and ranges)
2. Choose relevant experimental data
3. Sample parameter space (PROFESSOR)
4. Generate and analyse events (RIVET)
5. Interpolate generator response (PROFESSOR)
6. Find minimum over full parameter space (PROFESSOR)

- ideally tunes should be universal
- not possible to perfectly fit to all data
- some tunes optimised for precision physics processes

Pre-LHC tunes





- Tunes to Tevatron data disagreed with 900 GeV ATLAS data
- Large deviations → new tunes needed: AMBT2
- UE no longer modelled as average subtraction
- Separated charged and neutral components

Minimum bias tunes

Especially relevant for pileup simulation

Generator	Tune	Comments
PYTHIA 6	AMBT2B	tuned to 900 GeV and 7 TeV ATLAS data
PYTHIA 8	A2(M)	tuned to 7 TeV ATLAS data
PYTHIA 8	Monash	author's tune

Process-dependent tunes

Optimised by fitting to specific ATLAS measurements

Generator	Tune	Comments
PYTHIA 8	AZ	Low Z p_{T}, precision EW measurements
POWHEG +PYTHIA 8	AZNLO	As above but matched to POWHEG
PYTHIA 8	ATTBAR	7 TeV $t\bar{t}$ measurements, ISR/FSR tune
aMC@NLO +PYTHIA 8	ATTBAR-MG5_aMC@NLO	As above but matched to aMC@NLO
POWHEG +PYTHIA 8	ATTBAR-Powheg	As above but matched to POWHEG

General purpose tunes

- There are lots! https://twiki.cern.ch/twiki/bin/view/AtlasProtected/MCTunes
- Some of the most commonly used in ATLAS are detailed below

Generator	Tune	Comments
PYTHIA 8	A14	Combined shower and MPI tune with **eigentunes**
PYTHIA 6	Perugia2012	author's tune
HERWIG++	UE-EE5	author's tune
HERWIG +JIMMY	AUET2	ATLAS MPI tune using 900 GeV data
SHERPA	default	author's tune

Choosing a Monte Carlo generator

Choice of generators

- Many generators are on the market and supported by ATLAS
- No simple prescription to help you choose (sorry!)

Things to consider

1. Which generators have the physics you need?
   - could be different for signal and background
2. When using NLO generators
   - are they NLO for the process you're interested in?
3. How important is the tune?
   - consider using eigentunes for systematics

Conclusion

Conclusion

- Many different components make up Monte Carlo generation
- Some are perturbative, some are empirical and tuned to data
- NLO and NNLO predictions increasing prevalent
  $\rightarrow$ matching/merging to parton showers
- Choice of generator not trivial
- Don't try to reconstruct/use parton-level information
  $\rightarrow$ lots of work goes into producing final-state predictions

Most importantly:

- Don't be afraid to ask questions: `atlas-phys-pmg@cern.ch`
- Think about what you're doing
- Don't blindly follow prescriptions :)