Monte Carlo navigation

Alan Barr

University of Cambridge

Well-posed questions

- What particles were in the initial state?
  - + properties
- What particles were in the final state?
  - + properties
- How are the various particles allowed to interact?

- In real QM (e.g. matrix element) this is all we can ask.
- Does not address many other interesting questions!

Typical analysis questions

Increasingly badly posed →

- "Did this lepton come from a W decay?"

- "Is this really a b-jet?"

- "Where did this jet come from?"

- Almost all are un-physical in QM!
  - "Which slit did the photon go through?"
- However physicists want (need!) to ask them.
- Various degrees of satisfactory answer…

Why ask unphysical questions?

- **Understanding, bug-checking, justifying…**
  - Analysis code
  - Detector simulation
  - The MC itself

  Various use cases to be investigated

- **Unfolding:**
  - Correct to hadron level
    - Reduce detector effects
  - Correct to parton level?
    - Reduce hadronisation effects

Can't we do this already?

- HepMC allows one to find mothers/daughters etc. from connected graph
  - Nice interface
    - Iterator / Predicate
  - Good for tracing family relationships
  - Extraction of trees from graph
  - Unambiguous answers within MC framework
- Does not attempt to address ambiguous questions
  - "Where did this jet come from?"
  - "Did this jet come from a b-quark?"

Easy to do yourself?

- **Efficiency of effort**
  - Everyone does such things differently
  - Often M.C. specific (see ATLfast example next)
  - Associations require a good deal of thought!

- **B.C. (before CLHEP) everyone wrote their own:**
  - .pt() .m()
  - 4-vector addition
- **Common methods reduce error**
  - Better there aren't 1000 different versions
 - Produce focused, well-defined alternatives

Shortcomings of status codes

"What is the state of the event
after the end of the parton shower?"

- Different MC use attach different meanings to status codes
- E.g. in ATLfast (fortran and c++) finding partons uses ISTHEP=3 i.e. documentation line!
  - Works for Pythia, not for others.
- Status codes 1 (final state) is perhaps the only thing one can really rely on!
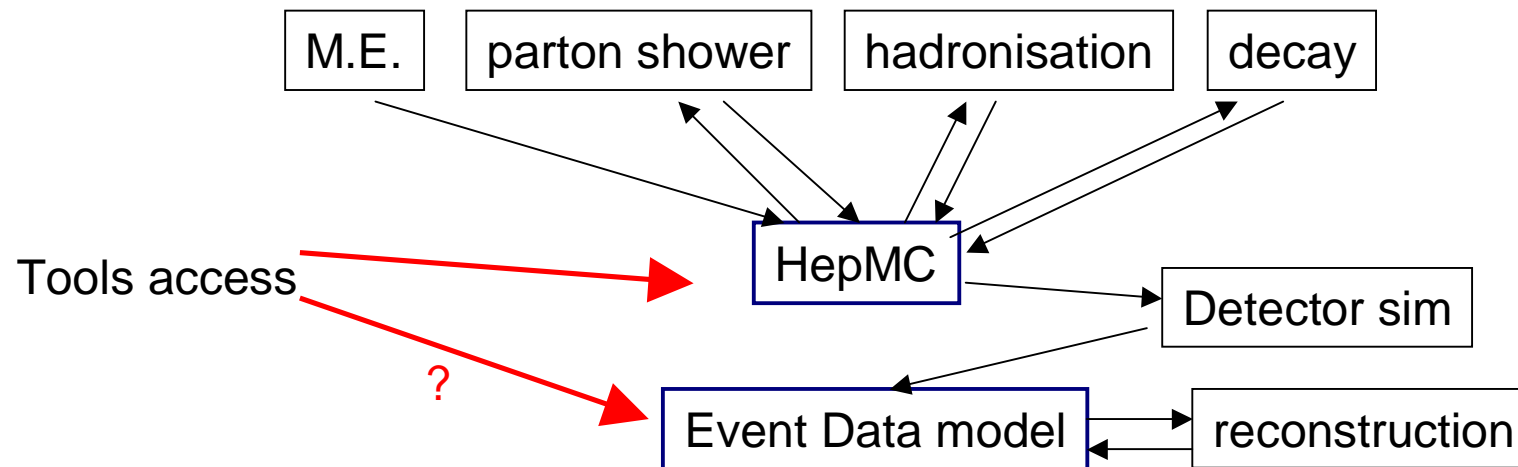
Representations

- **QM ~ real data**
  - Access to initial/final state only
- **Graph (HepMC)**
  - Appropriate for MC internals
  - Unambiguous
  - Offers "full" information
- **Association**
  - Good for answering many questions
  - Unavoidable ambiguity in answers

Parton from M.E.

Jet
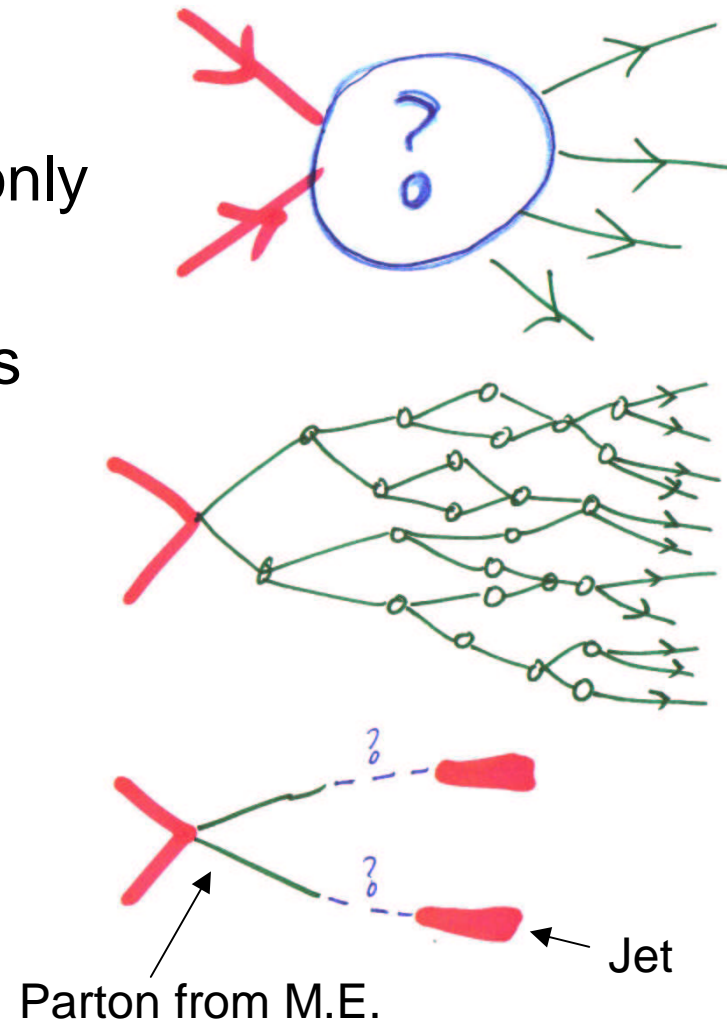
Data flow

Modularisation of
event-generator MC
following Les Houches

- Accessing objects at each stage
  - e.g. immediately before and after hadronisation
- Associating objects at different simulation stages
  - e.g. jet to parton
- NOT a replacement for HepMC or EDM

How to deal with ambiguities

- **The meat of the issue!**
  - □ No "right" answer
  - □ Alternatives with probabilities?
  - □ Best guesses?

Much work needed here!

- **Similar things done before in various contexts:**
  - □ Jet algoritm e.g. KTCLUS "assigns" particles to jets
    - Lots of possible alternatives.
    - Lots of physics
  - □ Simulation & reconstruction code associates "tracks" from inner detector hits back to MC particles.
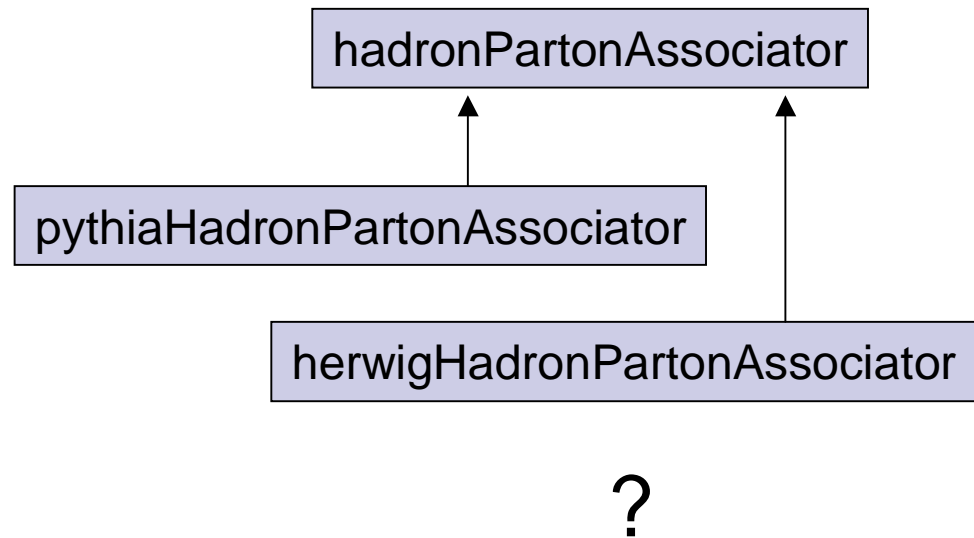
MC-independent answers?

- Could attempt to "translate" one MC to another

  - 'string' → 'cluter'

- Loss of information?

- Some questions best answered differently for different MC?

Possibility for implimentation

- **Common interface**
  - ☐ Or default method
  - ☐ Required for users
- **MC-specific implementation?**
  - ☐ E.g. cluster vs string models
  - ☐ Different MC use of status codes etc.
- **N.B. a suggestion**
  - ☐ Alternatives to be investigated!

hadronPartonAssociator

pythiaHadronPartonAssociator

herwigHadronPartonAssociator

?

Current situation

- ATLAS UK e-Science bid for 1 FTE over 3 years
- Relating matrix element, parton, hadron and detector-level information
- Ties in with larger navigation issue
- Generic LCG tool