Simulating Physics at High-p_T

Monte Carlo Experience from RunII

Stephen Mrenna

Fermilab, Computing Division, Physics Simulations Group

mailto:mrenna@fnal.gov

- Why do we need Monte Carlo?
- How will we do this?

```
Disclaimer:  based on my (limited) experience in RunII
```

- Why do we need Monte Carlo?
- How will we do this?

`Disclaimer: based on my (limited) experience in RunII`

- MC will not be used to discover

  It will play a part, but not indispensable

- It will be used to understand

  It will give confidence

  Can be used for interpretation

Most asked MC questions

1. How can I estimate the "theoretical" systematic uncertainty in a MC prediction?
   - Setting a limit, measuring a physics quantity
2. How can I add different "exclusive" MC samples to make a more inclusive one?
   - i.e. 3 jets + PS $\oplus$ 4 jets + PS
   - setting a better limit, making a more powerful discovery

These are not unrelated!

Prime goal for LHC physics

Q1: Estimating "Theoretical" Uncertainty

Dissection of a MC Prediction

$$d\sigma \sim \sigma_0 H(Q) \exp\left\{ -\int_{C_2 Q_0^2}^{C_1 Q^2} \frac{d\mu}{\mu} \left(A(\alpha_s) \ln\left(\frac{C_1 Q^2}{\mu^2}\right) + B(\alpha_s) \right) \right\} F_{\text{NP}}[C_1, C_2] \text{ SGA}$$

$$+ \left(\text{Fixed Order - Asymptotic} \right) [C_1, C_2] \text{ HGC}$$

C_1, C_2 set the infrared cutoff and hard scale

SGA $\equiv$ Soft Gluon Approximation; HGC $\equiv$ Hard Gluon Correction

- "Standard" Practice is to turn off ISR (FSR) to evaluate uncertainty
  - $C_1 \to C_2 (Q_0/Q)^2$ everywhere

1. HGC missing except for special (simple) cases
2. Refitting F_{NP} is no easy task (could be automated)

Ask the right questions

i. Given a physics description, how much can it reasonably vary?

ii. What is inherently lacking in the description? What approximations were made?

Ask the right questions

i. Given a physics description, how much can it reasonably vary?

ii. What is inherently lacking in the description? What approximations were made?

What theoretical uncertainty isn't

ISR on vs. ISR off

PYTHIA vs. HERWIG

Ask the right questions

i. Given a physics description, how much can it reasonably vary?

ii. What is inherently lacking in the description? What approximations were made?

What theoretical uncertainty isn't

ISR on vs. ISR off

PYTHIA vs. HERWIG

How are we doing better?

MC@NLO (matching a NLO calculation to HERWIG)

Tree Level-Parton Shower Matching

Q2: Adding different MC samples

W+3 partons + PS $\oplus$(?) W+4 partons + PS

W+3 hard jets + b-tags $\equiv$ B

W+4 hard jets + b-tags $\equiv$ S

How much of "top" is W+4 hard jets? Can we use W+3 hard jets?

How do I add MC samples without over/under counting?

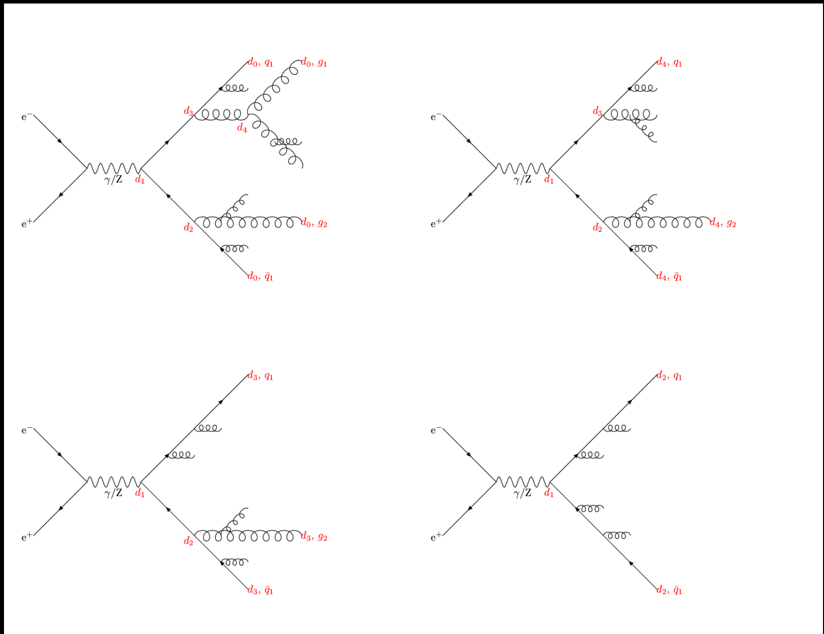
- In PS, (continuous) variation of topologies comes from Sudakov Form Factor (probability for no emission)
- Matrix Element calculation can be "mapped" into a PS history and reweighted with Sudakov FFs

  in soft/collinear limit, recover SGA

  in hard limit, apply HGC

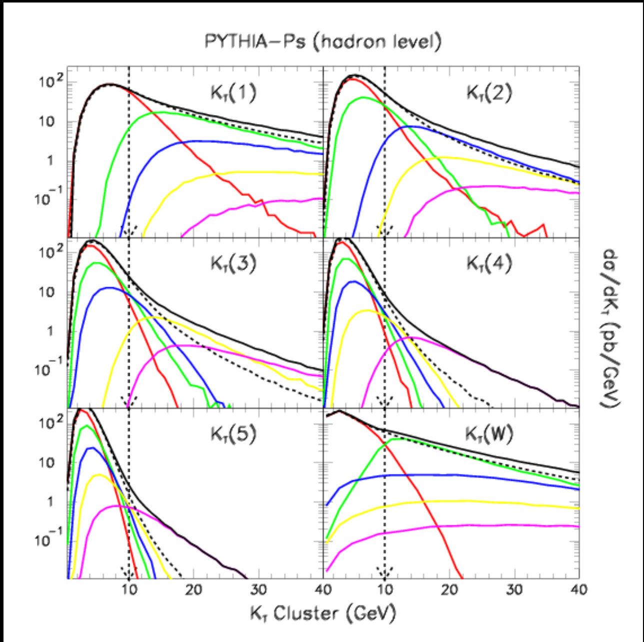
Pseudo–Showers and Sudakov Weight

Rerun the PS history and reject events with “bad” emissions



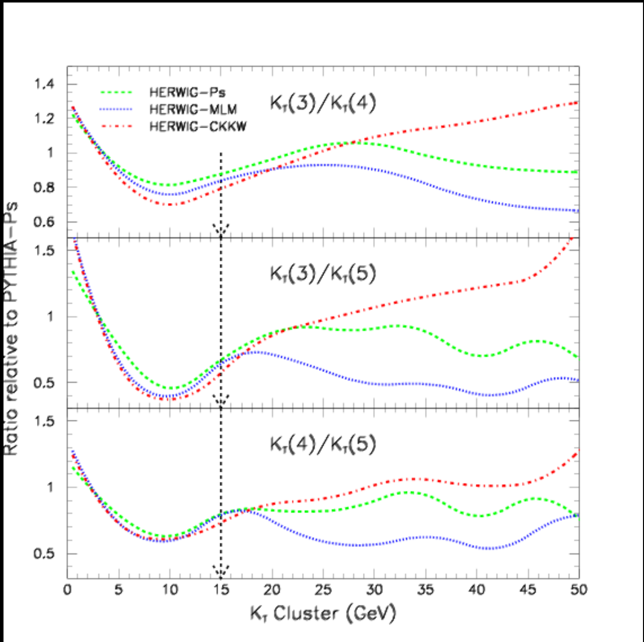
Reweighting allows smooth matching with lower topology

W+0 $\oplus \cdots \oplus$ W+4 hard partons



$$k_T^2 = 2\mathrm{min}(E_i, E_j)^2(1-\cos\theta_{ij}) \sim \mathrm{min}(E_i/E_j, E_j/E_i)m^2$$

Variation of Scheme yields a Theory Error



Variation with hard parton cutoff is also relevant

Lessons

- These calculations are not trivial

  My conviction is that experts should do expert work

  Theory/Phenos need to carry work through to where the experiment can take over

- Those who do this work are necessary and must be supported

  given resources (computing farms, mass storage, etc.)

- Nature of these calculations begs for databases and interface with experimental software

- mass storage
- standard format for files
- writeable from a computing farm
- searchable
- reasonably safe/secure
- easy to access by experiments, theorists
- files downloadable on hits

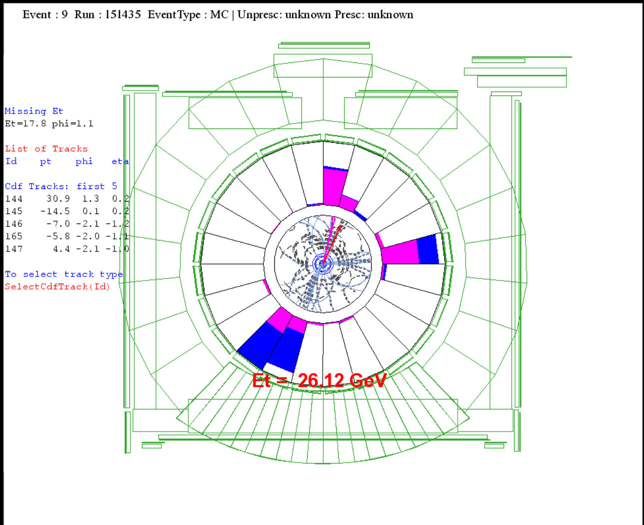
Patriot at FNAL

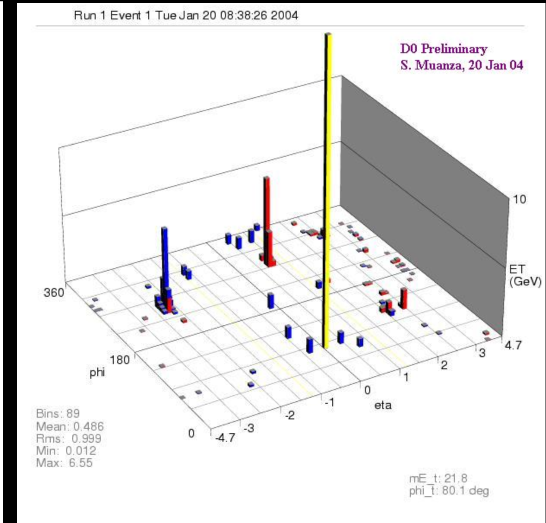
Physics Analysis Tools Required to Investigate Our Theories

- 1 TB Enstore repository
  - STDHEP + extra information + MCFIO
- Several different generators
  - Herwig, Pythia
  - Madgraph, Gr@ppa, CompHep, Alpgen
- Several different levels of generation
  - partons $\leftrightarrow$ showered partons $\leftrightarrow$ hadron level
- interface to SAM through disk cache
  - SAM $\equiv$ Sequential data Access via Meta-data
  - Oracle database (mcdb $\rightarrow$ Oracle)

Processed Events

Theorist → Patriot → Experiment





Predictions

- "Theory" databases will play an important role in LHC analyses
  - new and developing MC predictions from theorists
  - quality not quantity
- Tricks will be developed to fully exploit them
  - e.g., look tables from fully simulated events to allow a quick scan of different theory predictions
- Theory/Pheno types will organize more along the lines of experimental collaborations
  - ensure that calculations are performed, legacy is maintained
- calculations will be done differently
  - Effective field theory more suitable for parton showers will be used in HO calculations

For LHC physics analyses, MC must:

- · · · give a reasonable estimate of theoretical uncertainty
- · · · be improved beyond the present level of approximation

Important for:

Setting limits

Qualifying an anomaly

Quantifying a measurement

Progress has been made

In RunII, we are learning what we need to do this

Ideas, farms, databases