Importance of Standard Model Simulation for BSM Physics

Peter Richardson
IPPP, Durham

- Introduction
- Basics of Monte Carlo Simulations
- Matrix Element Corrections
- Conclusions

Introduction

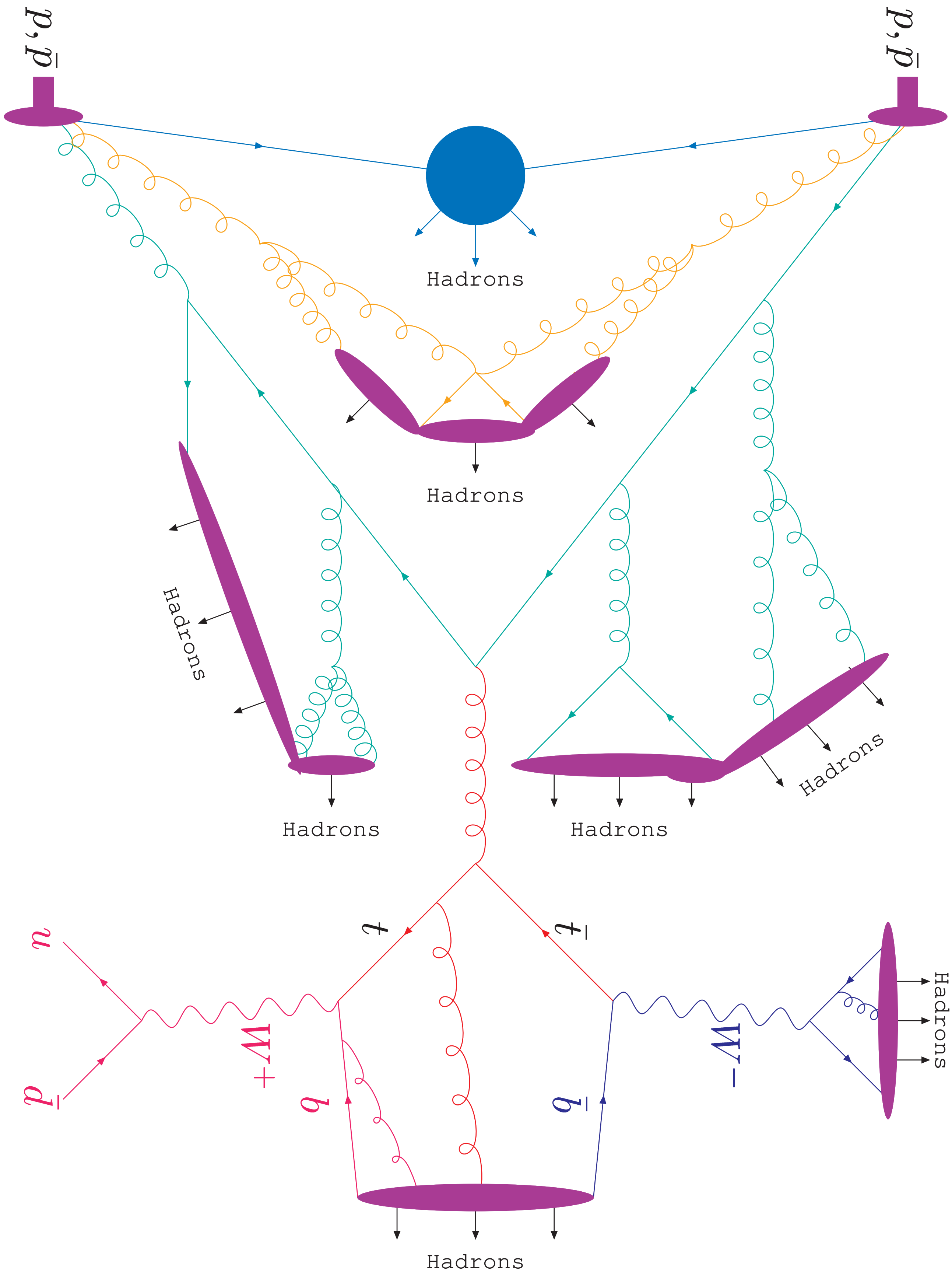
- Monte Carlo simulations are an essential part of most experimental analysis, either of the Standard Model or for searches for new physics.
- They are often treated as black boxes by experimentalists and not understood by theorists.
- It's important to understand the assumptions and approximations involved in these simulations.
- Experimental physicists need to be able to answer the following questions:
  - Is the effect I'm seeing due to different models, or approximations, or is it a bug?
  - I'm I measuring a fundamental quantity or merely a parameter in the simulation code?
  - Should the simulation describe what I'm looking at?
- Theorists need to understand enough to be able to understand problems due to the simulations used in getting experimental results.

Introduction

- There is a lot of interest in the simulation of processes involving large numbers of outgoing jets/leptons:
  - Top quark production including decays;
  - Signal Processes in BSM physics;
- In general the simulation of the background to these process is much more important than the simulation of the signal.
- Often the general purpose generators are run to give the backgrounds to these processes with little thought.
- I will explain why this is dangerous and often wrong and then present some recent ideas to improve things.
- This is important for physics at both the Tevatron and LHC.

Basics of Monte Carlo Generators

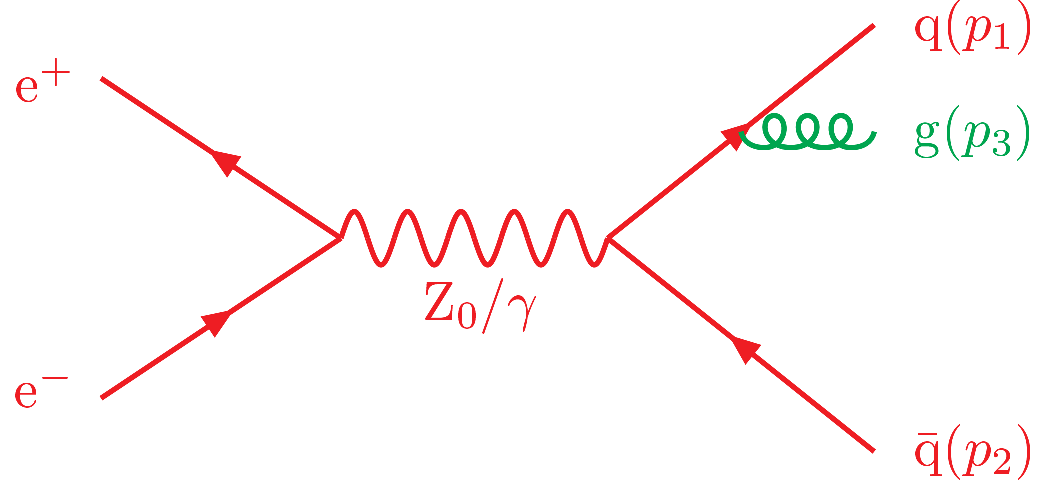
- All event generators split the simulation up into a number of parts:
  - Hard Process;
  - Parton Shower;
  - Secondary Decays;
  - Multiple Scattering/Soft Underlying Event;
  - Hadronization;
  - Hadron Decays.
- The general approach is the same in different programs but the models and approximations used are different.
- I will try and discuss the various models and approximations used in the different programs and compare them.
- I will try and give a fair objective comparison, but bear in mind I'm one of the authors of HERWIG.



$p, \bar{p}$
$p, \bar{p}$
Hadrons
Hadrons
Hadrons
Hadrons
Hadrons
Hadrons
Hadrons
t
$\bar{t}$
W^+
W^-
b
$\bar{b}$
u
$\bar{d}$
Hadrons
Hadrons

QCD Radiation

- It is impossible to calculate and integrate the matrix elements for processes involving large numbers of partons.
- Instead we treat the regions where the emission of QCD radiation is enhanced.
- Consider $\mathrm{e}^+\mathrm{e}^- \to \mathrm{q}\bar{\mathrm{q}}\mathrm{g}$



- The cross section is given by

$$\frac{d\sigma_{q\bar{q}g}}{dx_1 dx_2} = N_c \sigma_0 Q_q^2 C_F \frac{\alpha_s}{2\pi} \frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)}$$

- This diverges as $x_{1,2} \to 1$ where $x_i = 2E_i/\sqrt{s}$.

QCD Radiation

- Physically this happens when

1. The gluon and one of the quarks are collinear

$$(1-x_1) = \frac{E_2 E_3}{Q^2}(1-\cos\theta_{2g})$$

Collinear singularity

2. The energy of the gluon goes to zero

$$x_1 + x_2 + x_3 = 2, x_1, x_2 \to 1 \Rightarrow x_3 \to 0$$

Soft Singularity

Treatment of Singularities

- The general purpose event generators only treat the leading either the leading collinear emission or the leading soft and collinear emission.
  - ISAJET only treats the leading collinear region.
  - HERWIG treats both soft and collinear emission using angular ordering.
  - PYTHIA starts from the collinear approach with an angular veto to include the soft.
  - ARIADNE includes both soft and collinear emission using the dipole formalism.

- None of these approaches is reliable for the simulation of hard large angle emission.

- These programs dramatically under estimate the rate high p_T emission.

Hadronization

- ISAJET uses the original independent fragmentation model of Feynman and Field.
- PYTHIA uses the Lund string model.
- HERWIG uses the cluster hadronization model.
- ARIADNE uses the Lund string model.
- The independent fragmentation model cannot fit the LEP data.
- The cluster model gives good agreement with the LEP data on event shapes but doesn't fit the identified particle spectrum as well.
- The Lund string model give the best agreement with the LEP data.

Underlying Event

- In hadron-hadron collisions in addition to modelling the perturbative physics in the hard collision and parton showers we need to describe the underlying event.

- We can think of this as the interaction of the hadron remnants after the hard collision.

- The different event generators use different models.
  - ISAJET uses an approach based on pomeron exchange.
  - PYTHIA uses a model based on multiple parton collisions.
  - HERWIG uses a soft model based on the UA5 event generator, (there is also a multiple scattering model available.)

- Both the HERWIG and PYTHIA multiple scattering models can be tuned to give good agreement with CDF data.

Matching matrix elements and Parton Showers

- Parton Shower (PS) simulations use the soft/collinear approximation:
  - Good for simulating the internal structure of a jet;
  - Can't produce high p_T jets.
- Matrix Elements (ME) compute the exact result at fixed order:
  - Good for simulating many high p_T jets;
  - Can't give the structure of a jet.
- We want to use both in a consistent way, *i.e.*
  - ME gives hard emission,
  - PS gives soft/collinear emission,
  - Smooth matching between the two,
  - No double counting of radiation.

Matching matrix elements and Parton Showers

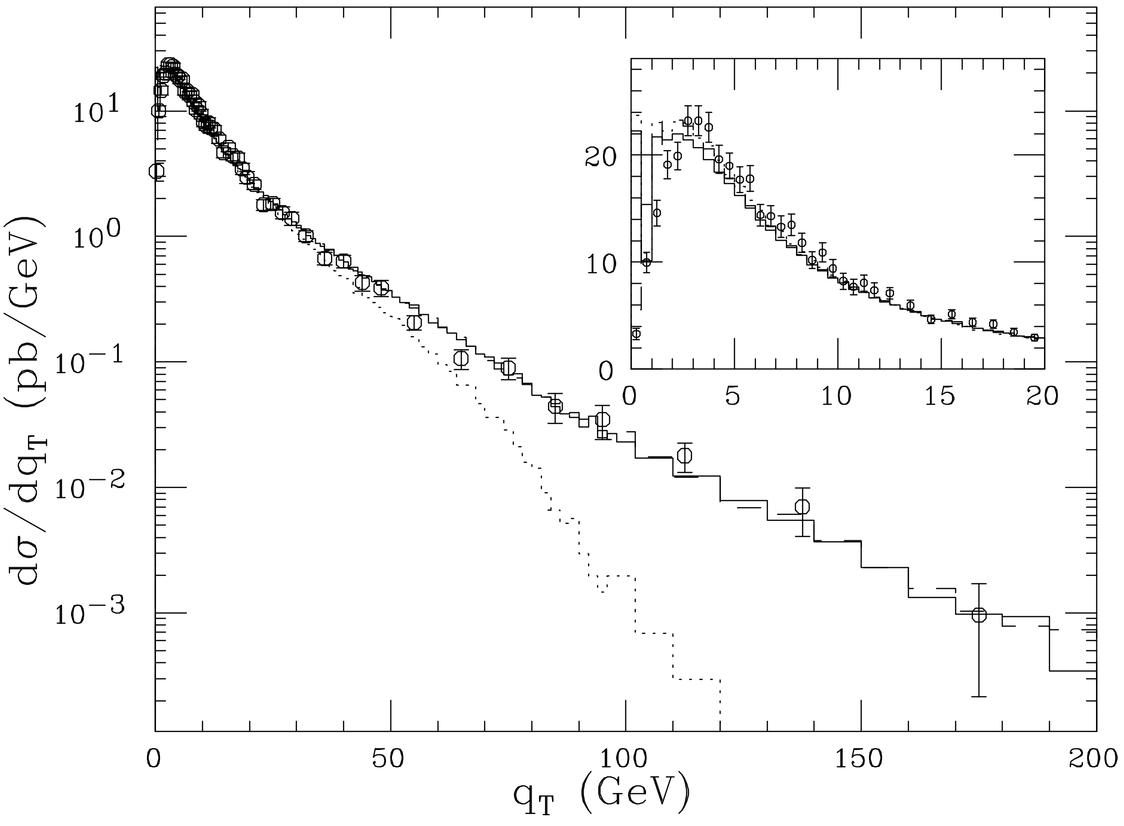
- There have been a number of attempts to do this.

- The first attempts got the hardest emission correct for simple processes, for example:
  - $e^+e^- \to q\bar{q}$;
  - DIS;
  - Drell-Yan;
  - Top Decay.

- There were problems with this:
  - Only the hardest emission was correctly described;
  - The leading-order normalization was retained.

Matching matrix elements and Parton Showers

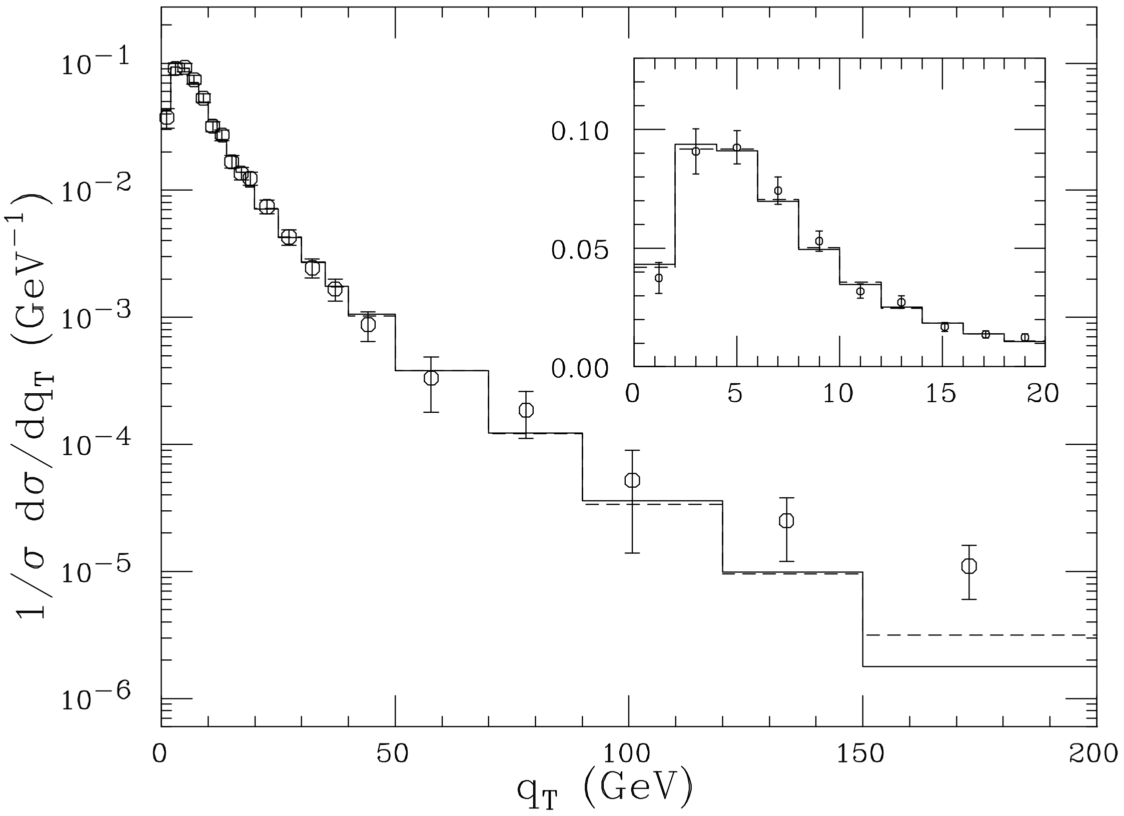
- Transverse momentum distribution of Z at the Tevatron (CDF) compared with HERWIG.



- Taken from Nucl. Phys. B565:227-244,2000, G. Corcella and M. H. Seymour.

Matching matrix elements and Parton Showers

- Transverse momentum distribution of W at the Tevatron (D0) compared with HERWIG.



- Taken from Nucl. Phys. B565:227-244,2000, G. Corcella and M. H. Seymour.

Matching matrix elements and Parton Showers

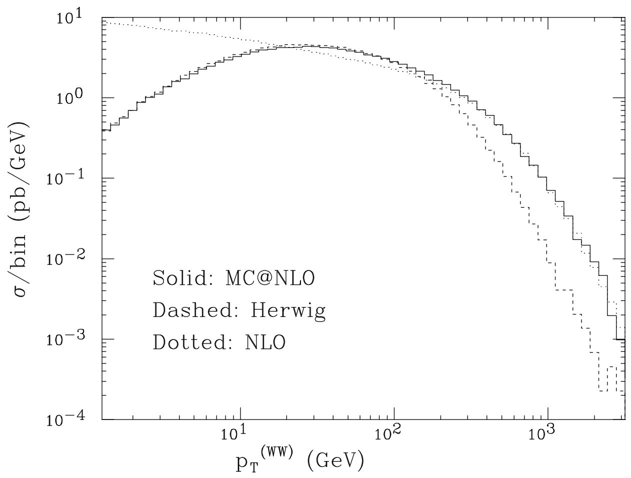
- Recently there has been renewed interest.
- Two types of approach have emerged.

1. NLO Simulation
   - NLO normalization of the cross section.
   - Gets the hardest emission correct.
   - Only programs by Frixione/Webber and Dobbs.
2. Multi-jet Leading Order
   - Still leading order.
   - Gets many hard emissions correct.
   - Programs by Krauss et. al. and Lonnblad for e^+e^-.

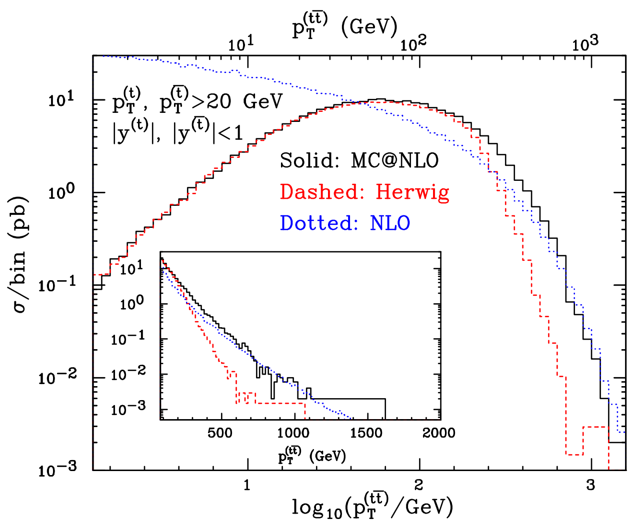
NLO Simulation

- These simulations are now available for a range of processes using the method of Frixione and Webber.

- Transverse momentum of W^+W^- pair at the LHC.



- Transverse momentum of $t\bar{t}$ pairs at the LHC.



Multi-Jet Leading Order Simulation

- While the NLO approach is good for one hard additional jet and the overall normalization can not be used to give many extra hard jets, at least at the moment.

- Therefore to simulate these processes use matching at leading order to get many hard emissions correct.

- I will briefly review the general idea behind this approach and then show some results from work I have been doing implementing it using MADGRAPH and HERWIG.

Procedure

- In order to match the ME and PS we need to separate the phase space:
  - One region should contain the soft/collinear region and be filled by the PS;
  - The other is filled by the matrix element.

- In Catani-Krauss-Kuhn-Webber (CKKW) prescription this is done via the Durham jet algorithm.
  - For all final-state particles compute the resolution variables
    * $d_{kB} \simeq E_k^2\theta_{kB}^2 \simeq k_{\perp kB}^2, \quad \theta_{kB} \to 0.$
    * $d_{kl} \simeq \min(E_k^2, E_l^2)\theta_{kl}^2, \quad \theta_{kl} \to 0$
  - The smallest of these d_{kB}, d_{kl} is then selected. If d_{ij} is the smallest the two particles are merged, if d_{iB} is the smallest the particle is merged with the beam.
  - This procedure is repeated until the minimum value is above some stopping parameter $d_{\rm cut}$.
  - The remaining particles and pseudo-particles are then the hard jets.

Procedure

1. Select the jet multiplicity with probability

$$P_n = \frac{\sigma_n}{\sum_{k=0}^{k=N} \sigma_k}$$

   where σ_n is the n-jet matrix element evaluated at resolution $d_{\rm ini}$ using $d_{\rm ini}$ as the scale for the P.D.F's and α_S, n is the number of additional jets.

2. Distribute the jet momenta according to the ME.

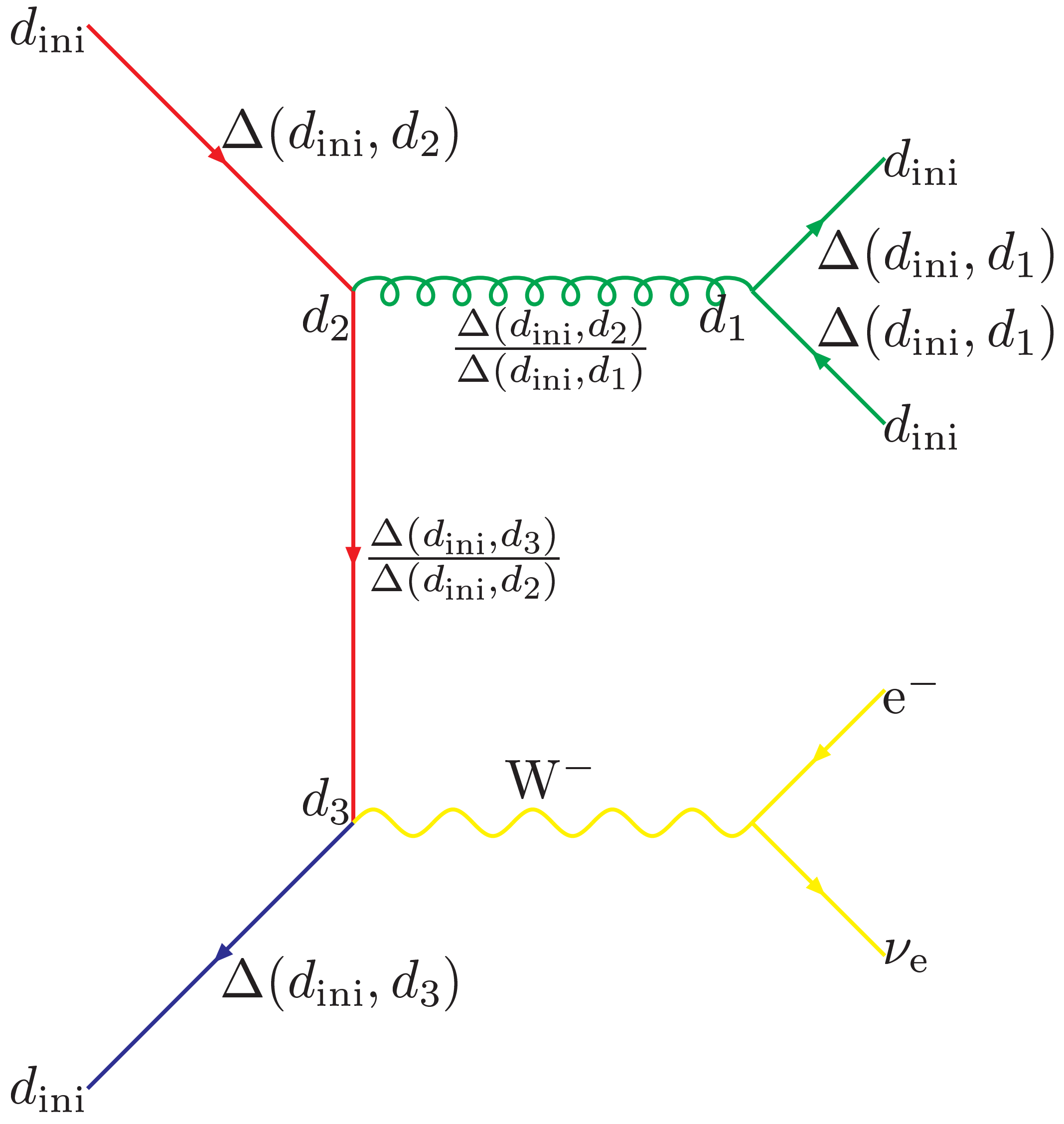
3. Cluster the partons to determine the values at which 1,2,..n-jets are resolved. These give the nodal scales for a tree diagram.

4. Apply a coupling-constant reweighting $\alpha_S(d_1)\alpha_S(d_2)..\alpha_S(d_n)/(\alpha_S(d_{\rm ini})^n \leq 1$.

5. Reweight the lines by a Sudakov factor $\Delta(d_{\rm ini}, d_j)/\Delta(d_{\rm ini}, d_k)$.

6. Accept the configuration if the product of the α_S and Sudakov weight is less the $\mathcal{R} \in [0,1]$., otherwise return to step 1.

Procedure



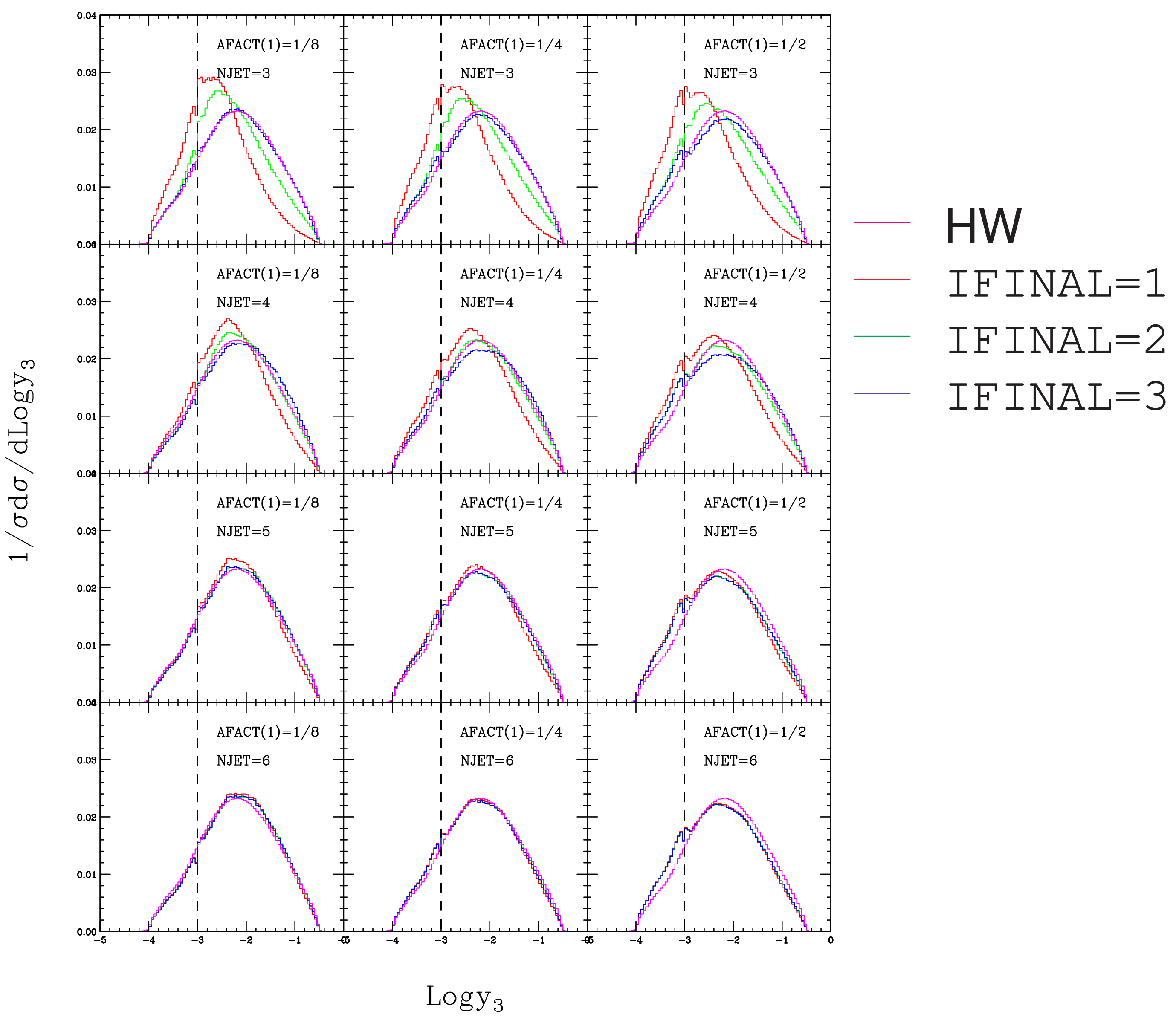
7. Generate the parton shower from the event starting the evolution of each parton at the scale it was created and vetoing emission above the scale $d_{\rm ini}$

Free Parameters

- Although the algorithm above ensures smooth matching at NLL log level there are a number of free choices
  - Sudakov Form Factors
    * HERWIG or NLL Sudakovs
    * Definition of the scale
    * Prefactor for the scale
  - α_S
    * LO or NLO definition
    * Definition of the scale
    * Prefactor for the scale
  - Treatment of Highest Multiplicity Matrix Element
    * Same as other multiplicities?
    * Different treatment to describe more emissions?
  - Definition of k_T-algorithm away from soft/collinear limit
    * Definition of the k_T-measure.
    * Recombination scheme.

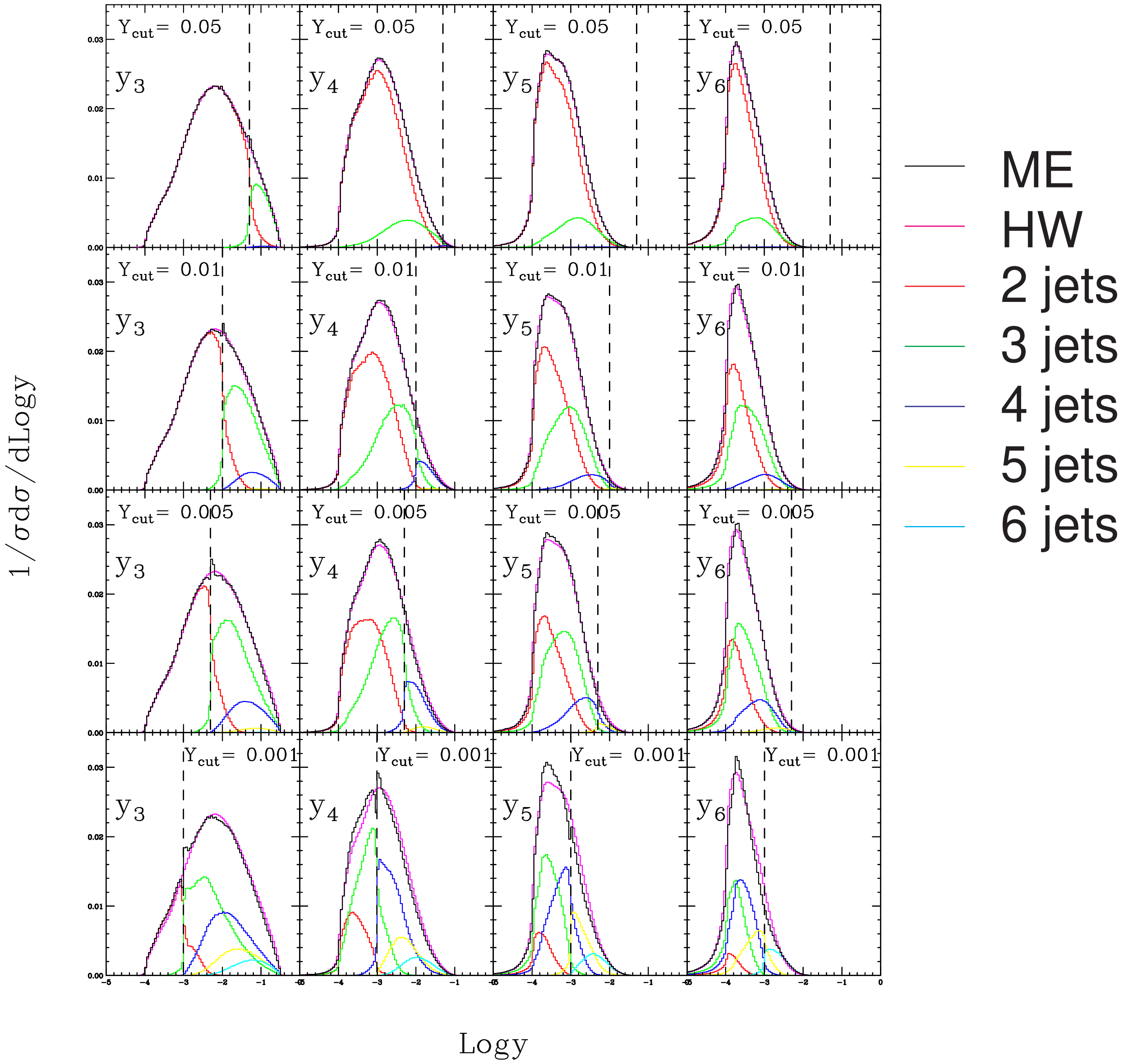
Treatment of Highest Multiplicity

- y_3 distribution.



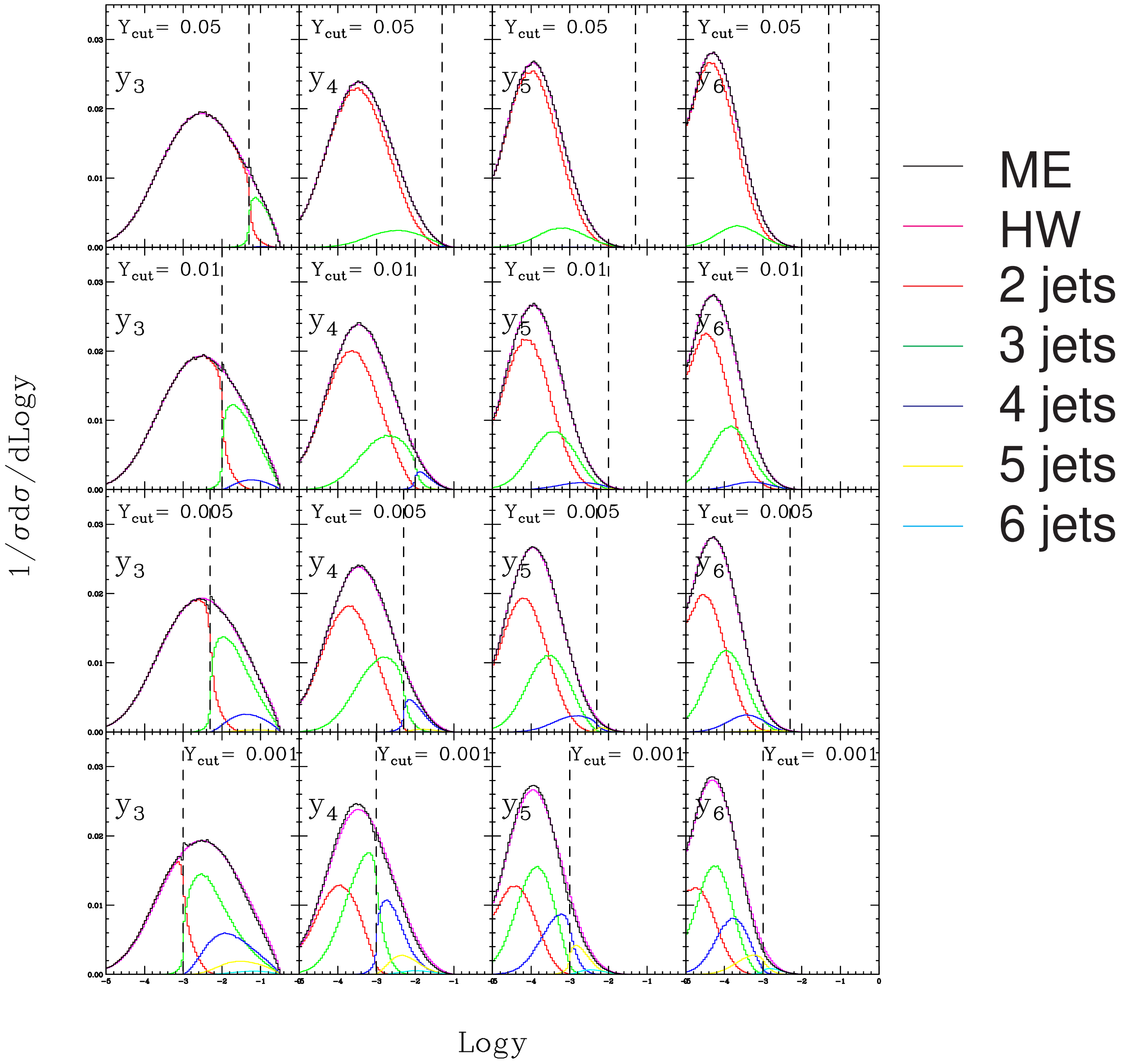
Effect of Varying the cut-off at M_Z

- Parton-level distributions.



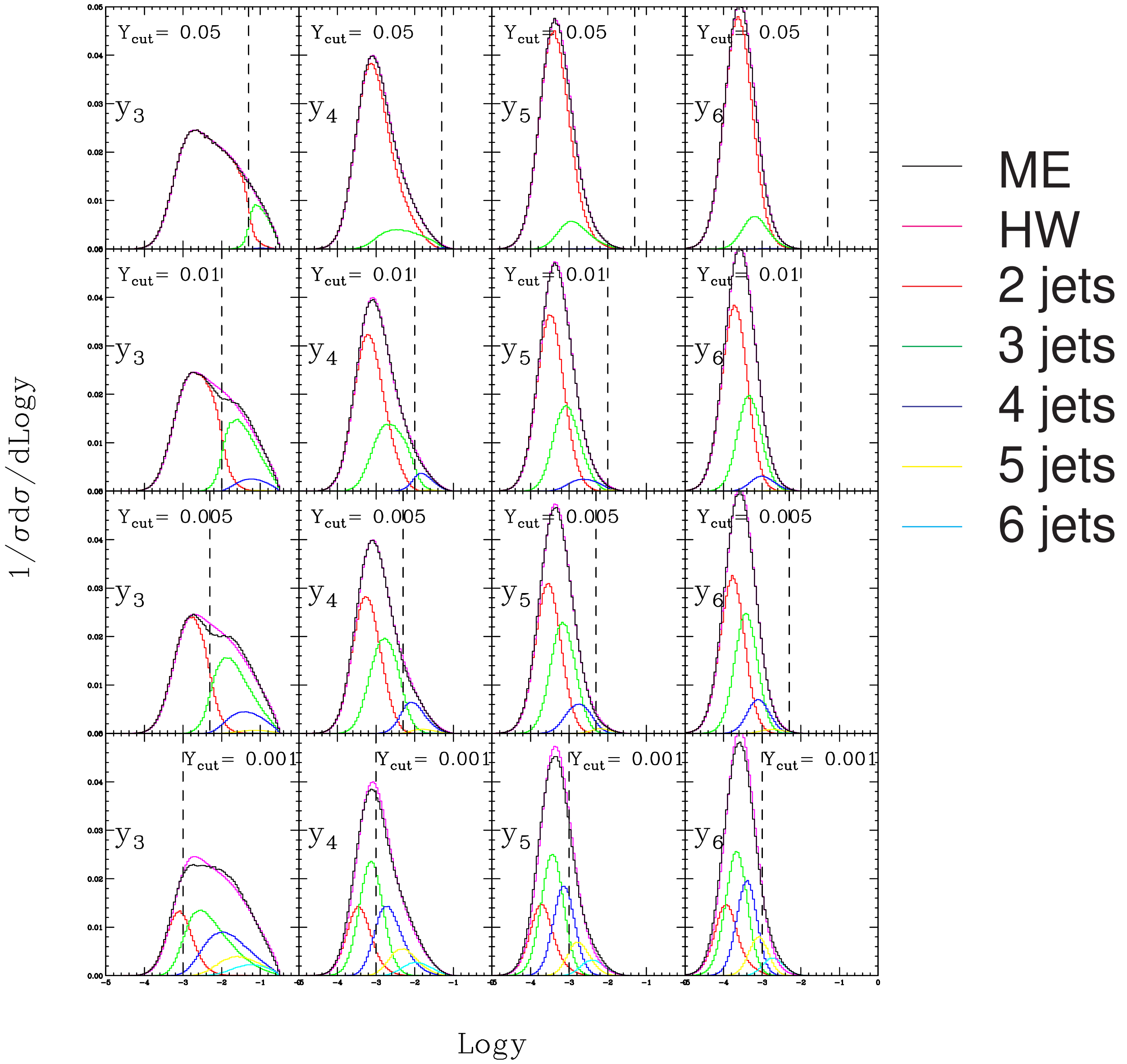
Effect of Varying the cut-off at $\sqrt{s} = 500$ GeV

- Parton-level distributions.



Effect of Varying the cut-off at M_Z

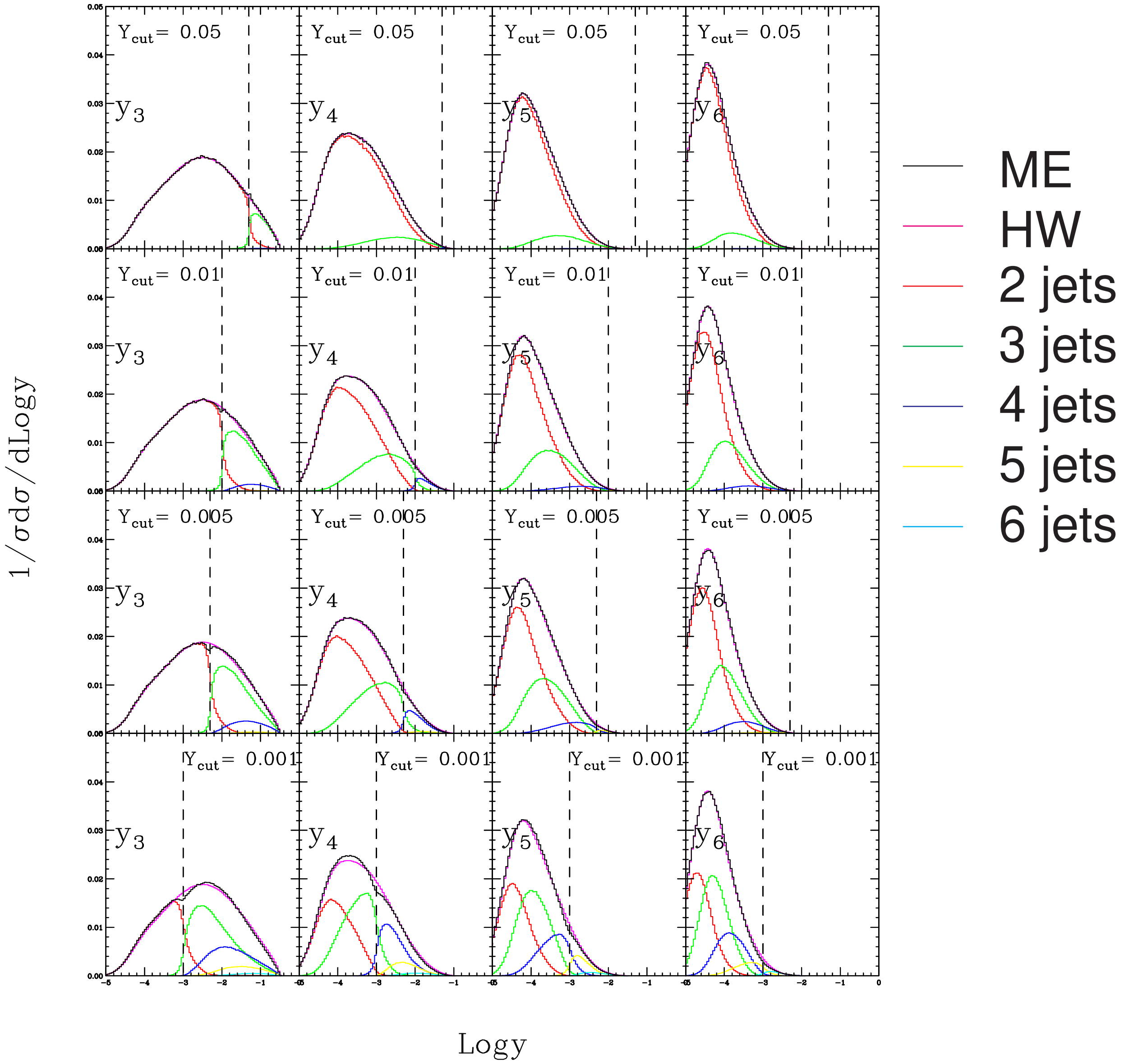
- Hadron-level distributions.
- Problems with hadronization model which remain to be solved.



Effect of Varying the cut-off at $\sqrt{s} = 500$ GeV

- Hadron-level y_3 distribution.

- Problems with hadronization model reduced.



Hadron Collisions

- In hadron collisions in addition to final-state radiation we have the problems of initial-state radiation.
- I will consider W production in hadron collisions
- All the events were generated with MADGRAPH including the leptonic decays of the gauge bosons.
- For the W samples no cuts were applied on the leptons.

Hadron Collisions

- We also have a greater choice of the clustering algorithm.

- I will only show one choice here.
  - Covariant monotonic p_T^2 scheme

$$
\begin{aligned}
d_{iB} &= p_{ti}^2, \\
d_{ij} &= \min(p_{ti}^2, p_{tj}^2) R_{ij}^2,
\end{aligned}
$$

Define p_t and R for pseudoparticles as

$$
\begin{aligned}
p_{t(ij)} &= p_{ti} + p_{tj}, \\
R_{(ij)k}^2 &= \frac{p_{ti}^2 R_{ik}^2 + p_{tj}^2 R_{jk}^2}{p_{ti}^2 + p_{tj}^2}.
\end{aligned}
$$

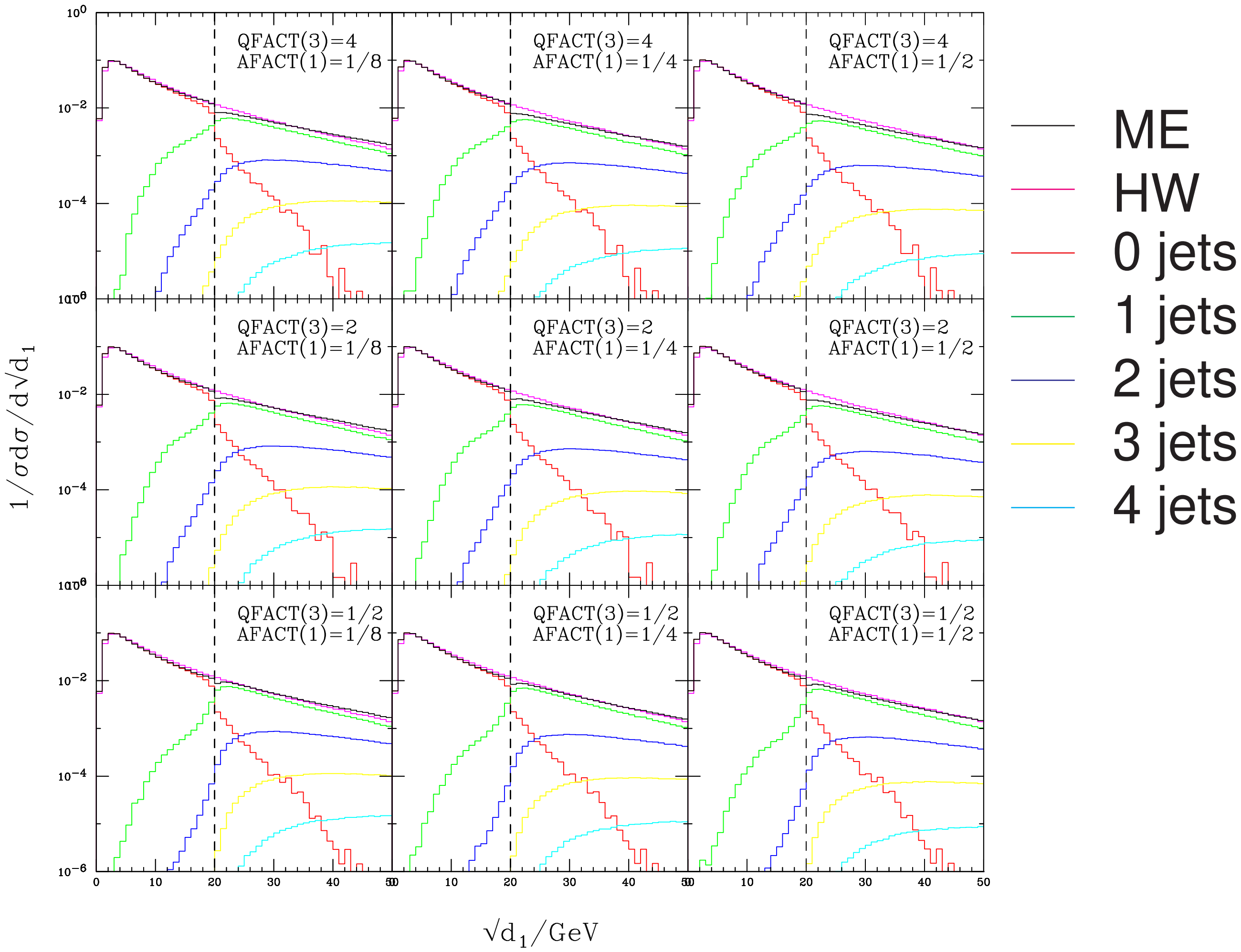
- Other schemes give similar results, (provided they are longitudinally boost invariant.)

Hadron Collisions

- Start by considering W production at the Tevatron.
- Use the Covariant monotonic p_T^2 scheme.
- and a matching scale of $400\ \mathrm{GeV}^2$.
- Look at the differential cross section with respect to $\sqrt{d_n}$ where d_n is the scale at which n jets are resolved.
- This is like the E_T in a cone algorithm.

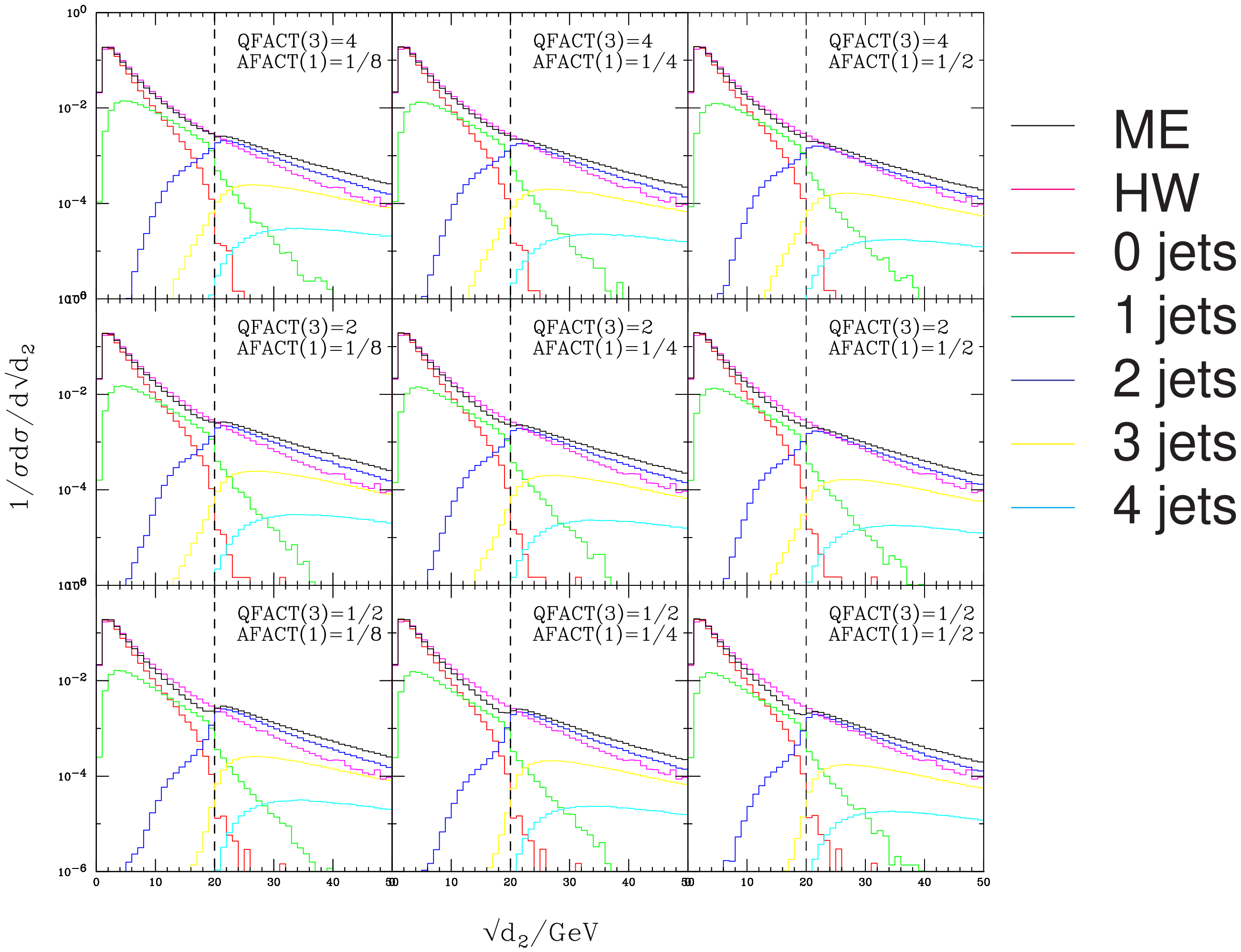
W Production

- Variation of results with scale of α_S and starting scale of shower.
- d_1 distribution at the Tevatron



W Production

- Variation of results with scale of α_S and starting scale of shower.
- d_2 distribution at the Tevatron

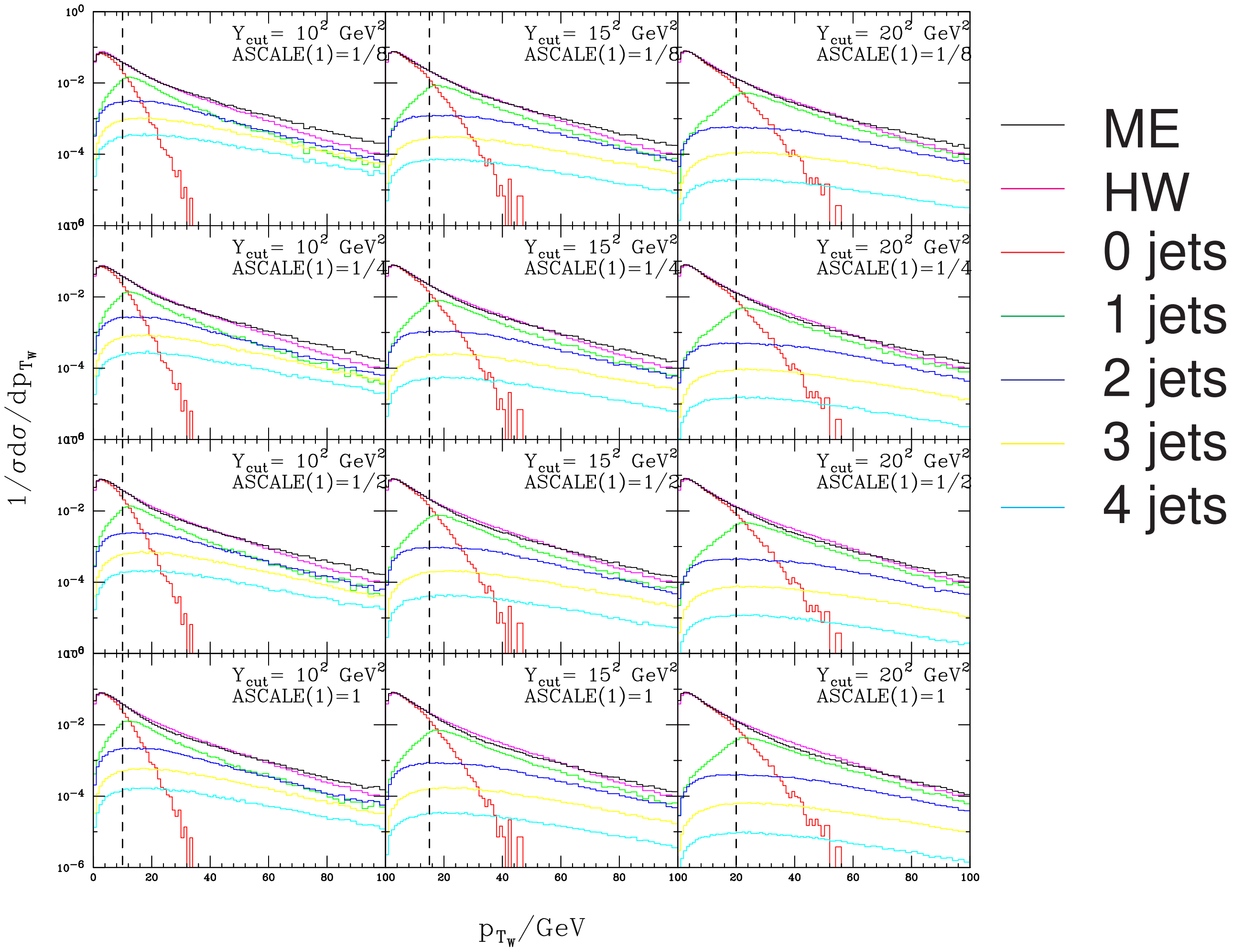


Hadron Collisions

- There is a radiation dip at the matching scale.
- There is also a problem with events migrating below the matching scale.
- Hard to get the HERWIG shower to produce more radiation than before below the matching scale.
- Solution is to decouple the scale used in the Sudakovs and that used to start the shower. Use a choice closer to the HERWIG variables to start the shower.

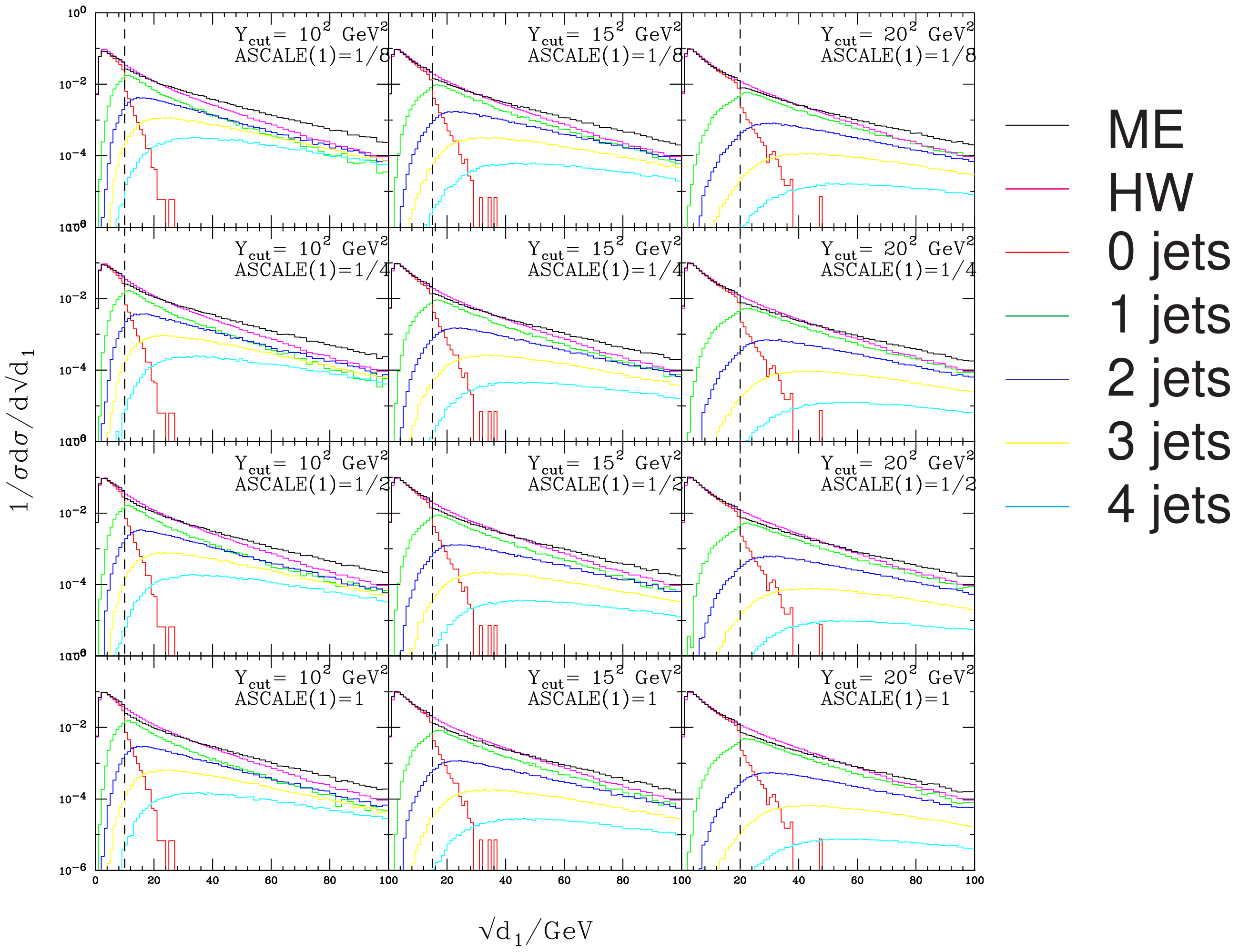
W Production

- Variation of results with scale of α_S and matching scale.
- p_T of the W at the Tevatron



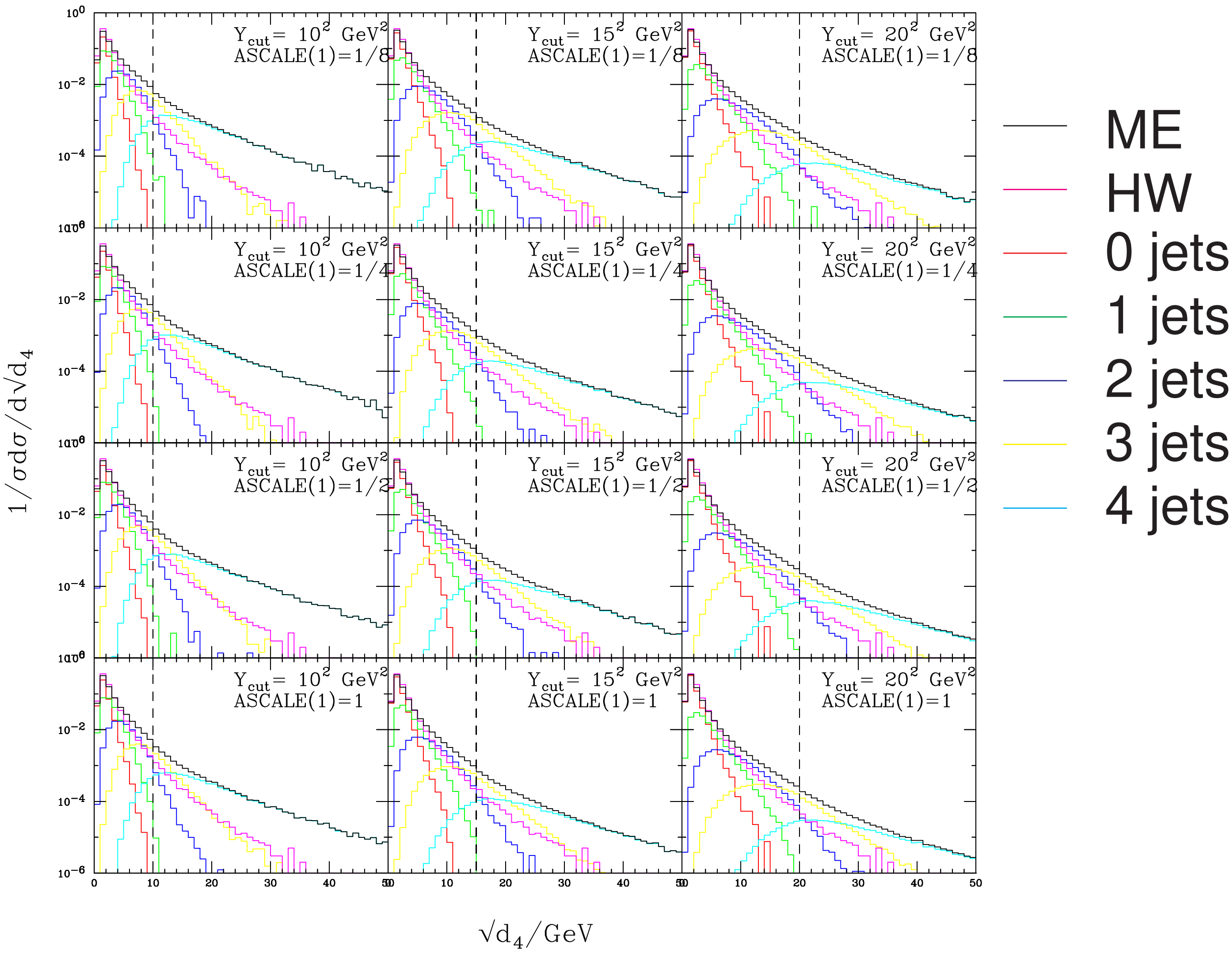
W Production

- Variation of results with scale of α_S and matching scale.
- d_1 distribution at the Tevatron



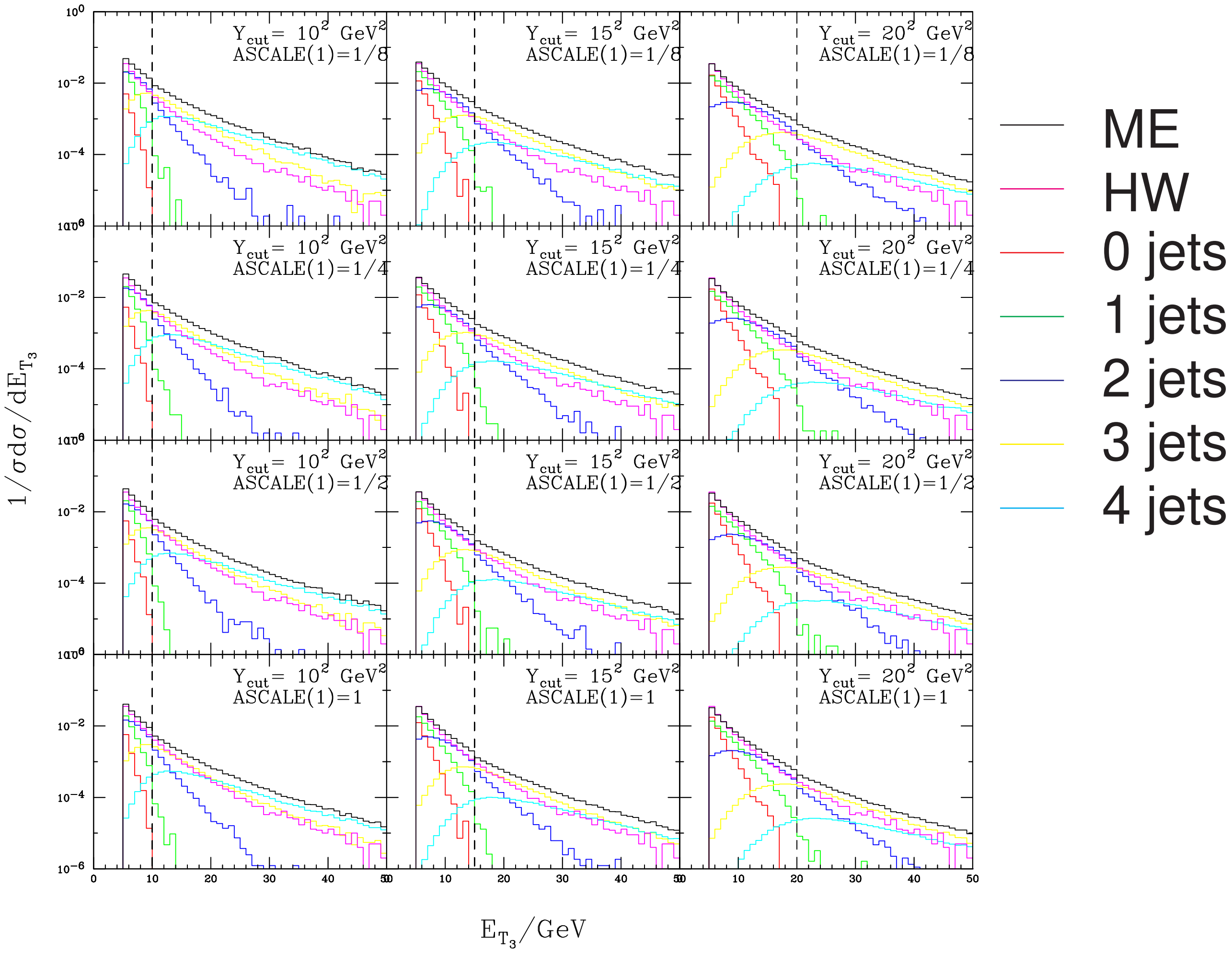
W Production

- Variation of results with scale of α_S and matching scale.
- d_4 distribution at the Tevatron



W Production

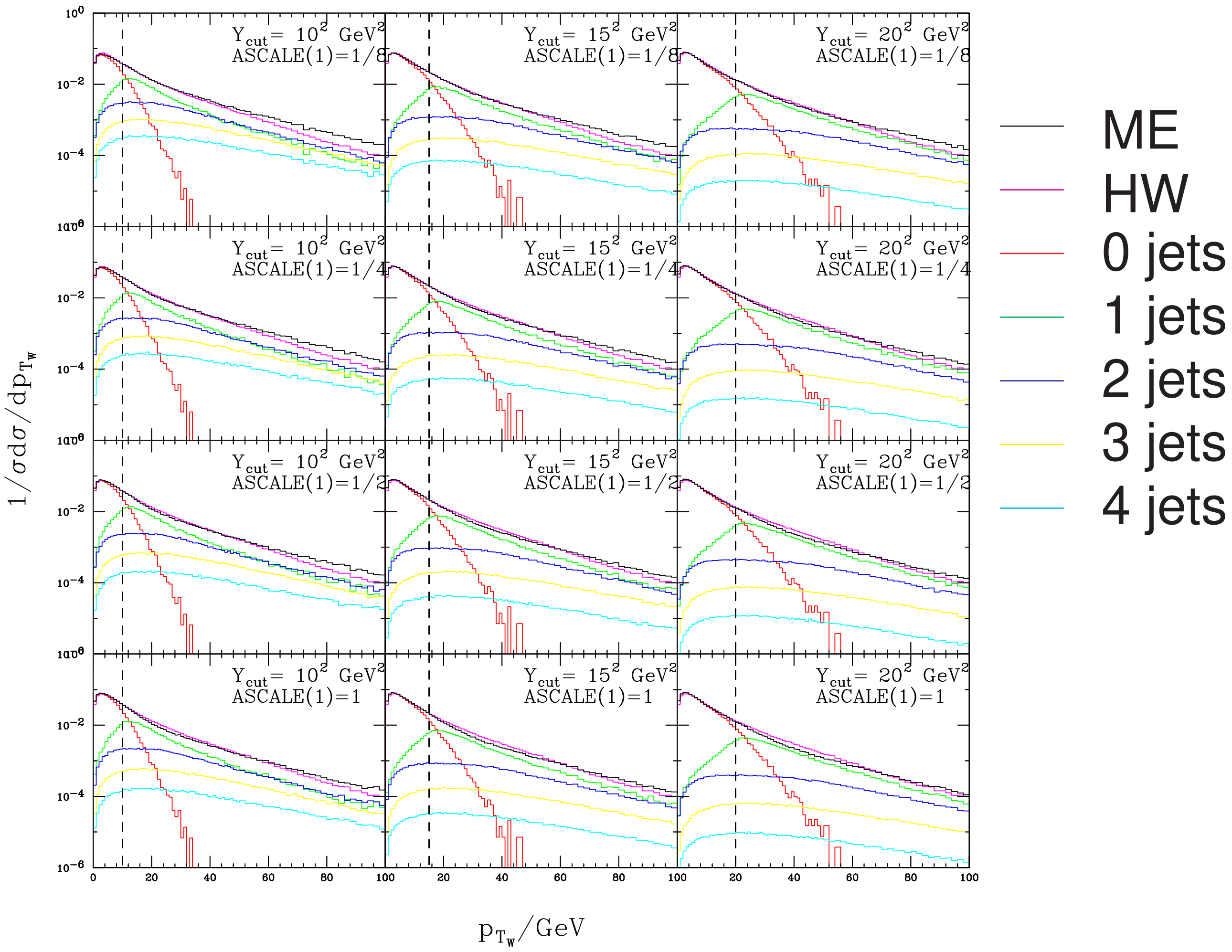
- Variation of results with scale of α_S and matching scale.
- E_T of the third jet at the Tevatron with cone $R = 0.7$.



W Production

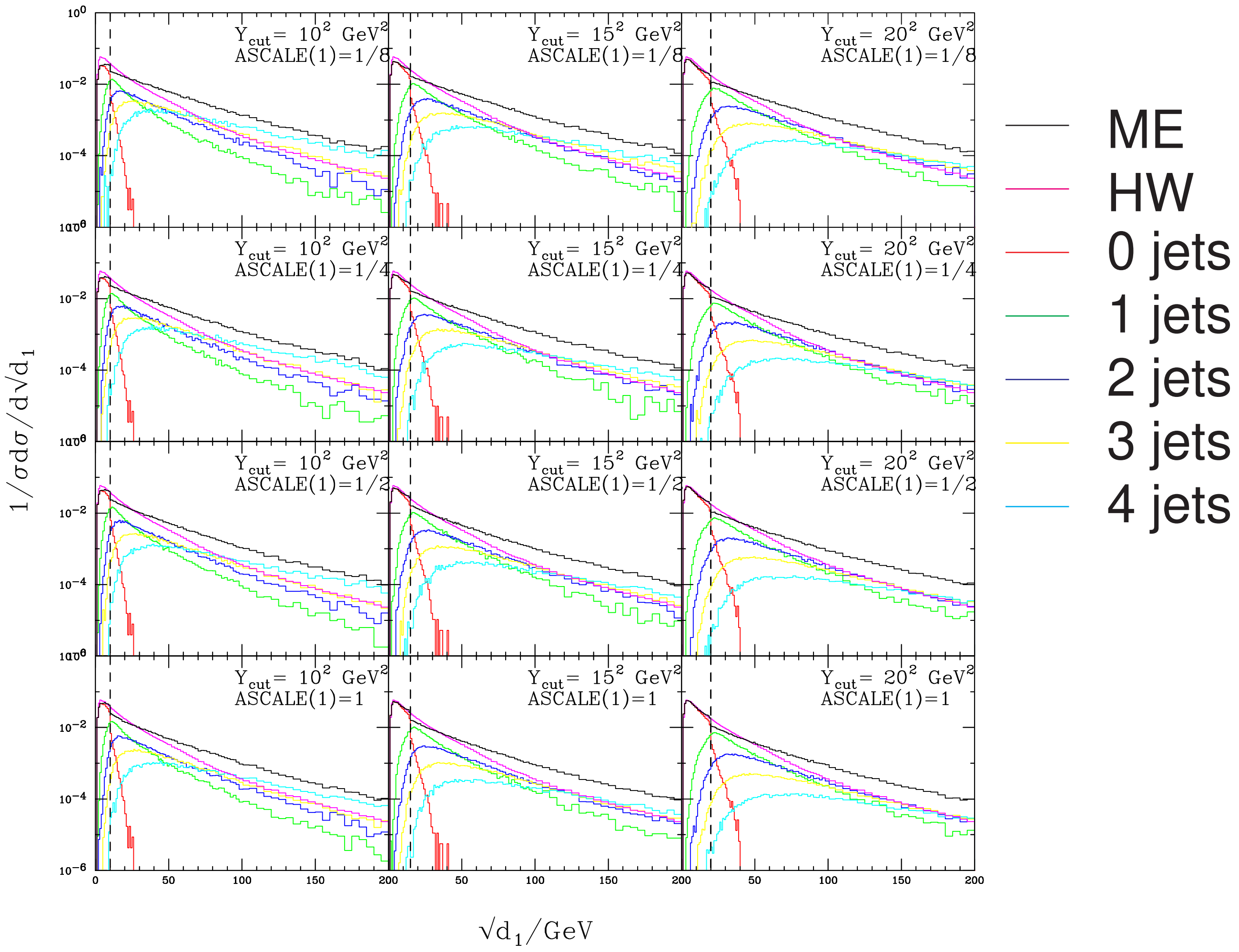
- Variation of results with scale of α_S and matching scale.

- p_T of the W at the LHC.



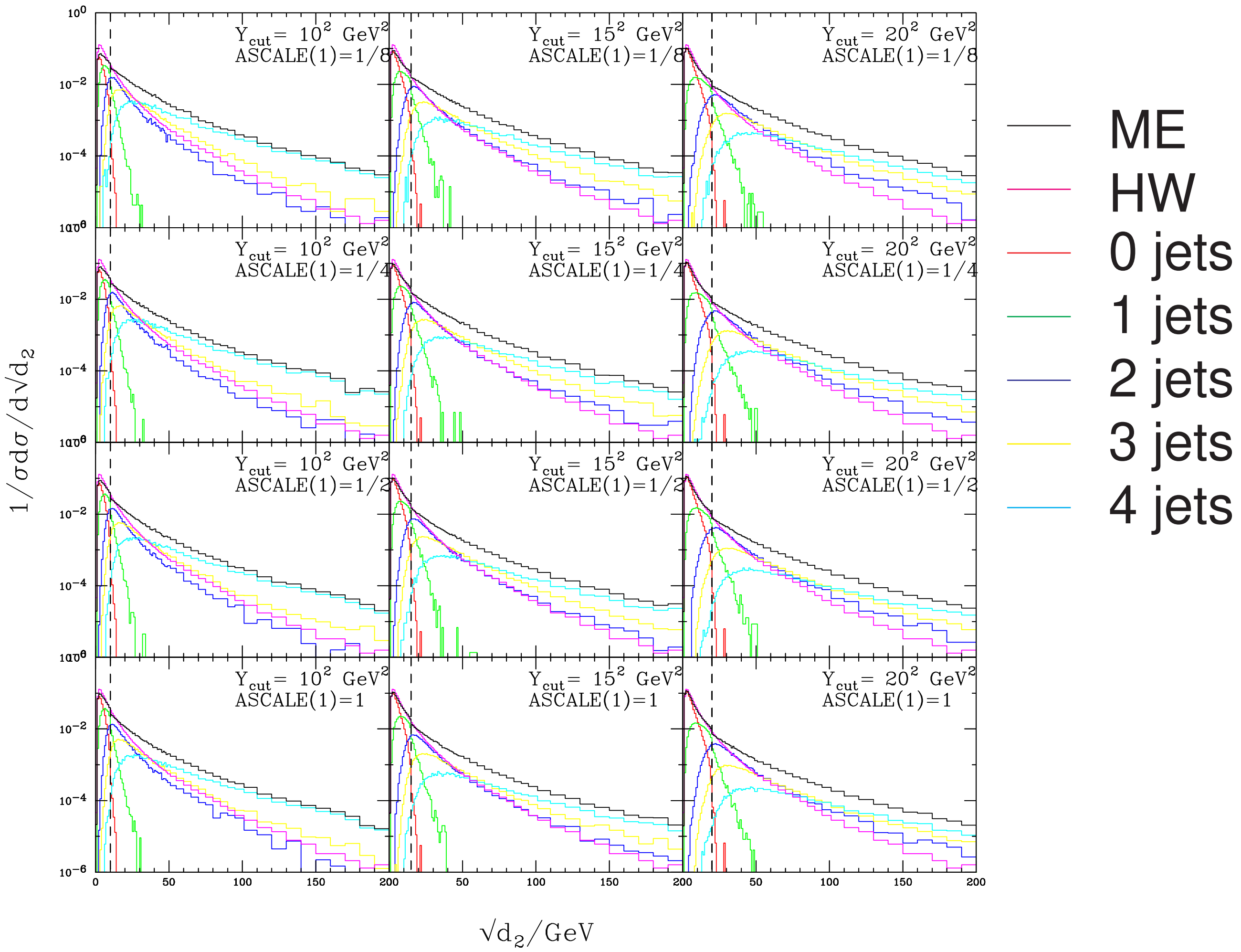
W Production

- Variation of results with scale of α_S and matching scale.
- d_1 distribution at the LHC.



W Production

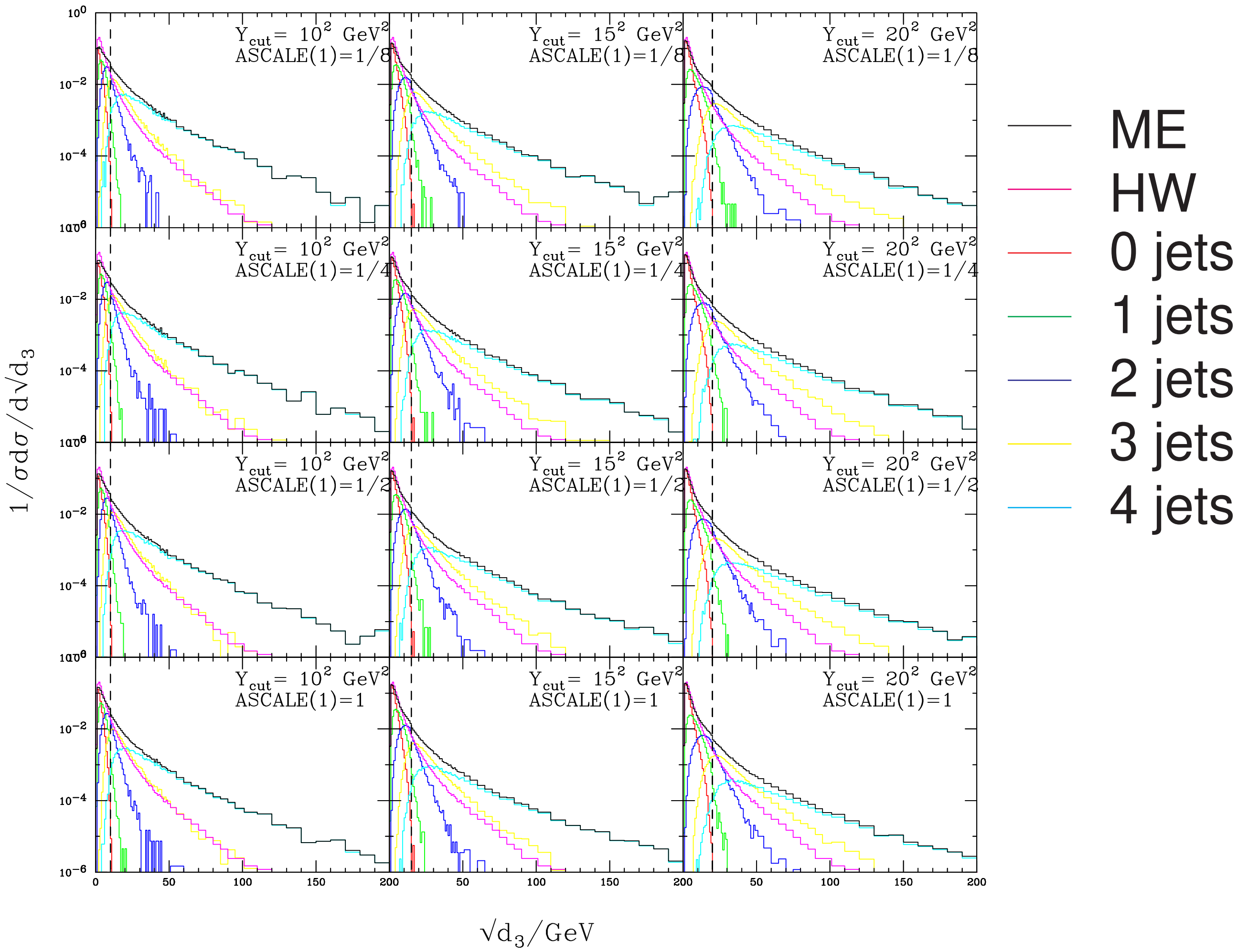
- Variation of results with scale of α_S and matching scale.
- d_2 distribution at the LHC.



W Production

- Variation of results with scale of α_S and matching scale.

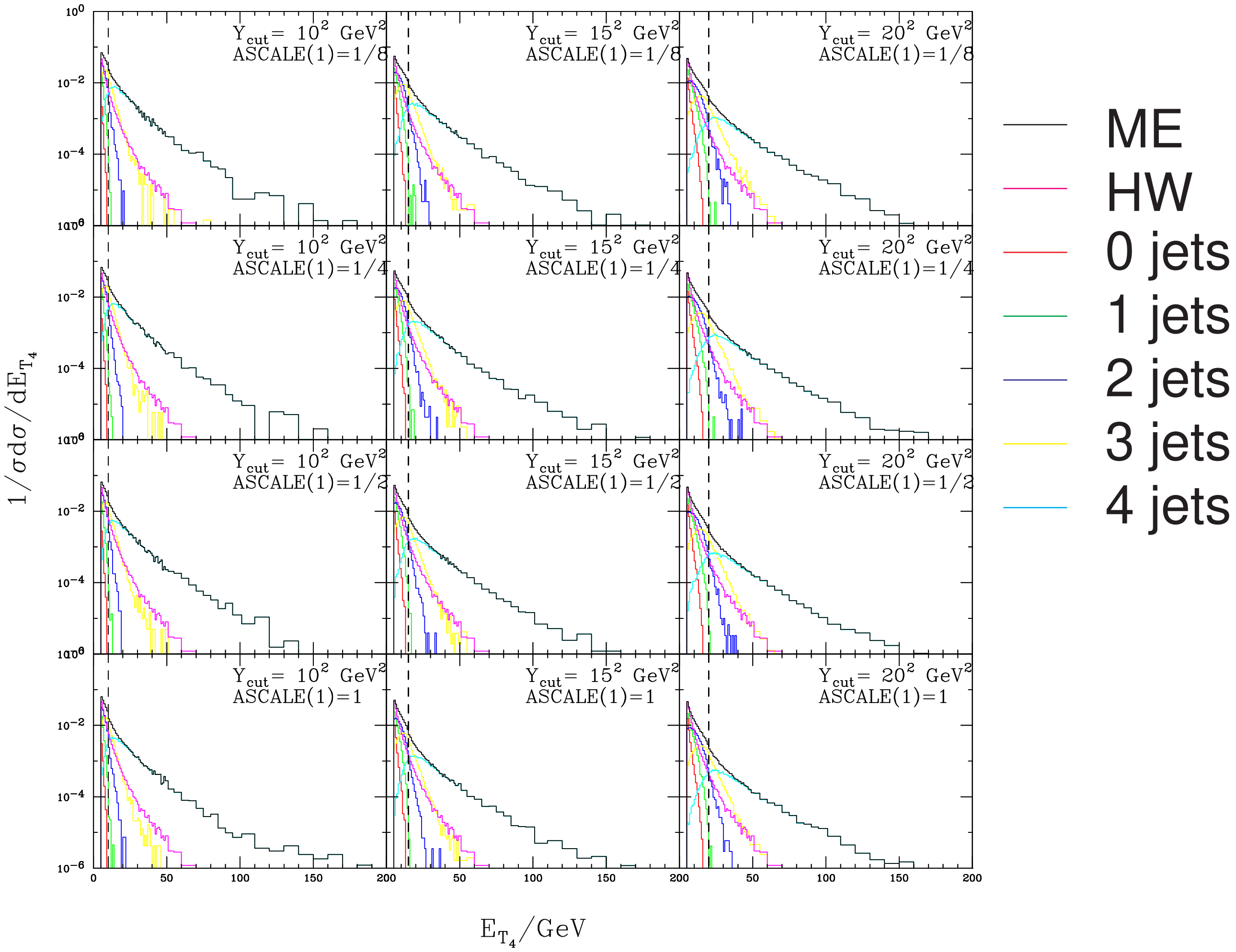
- d_3 distribution at the LHC.



W Production

- Variation of results with scale of α_S and matching scale.

- E_T of the fourth jet at the LHC with cone $R = 0.4$.



Conclusions

- Existing Monte Carlos significantly underestimate the rate for high p_T jet production.
- This means that a number of backgrounds for the LHC have probably been significantly underestimated.
- New tools are becoming available which should give a much better estimate of these rates.
- These shoulf be used rather than just running HERWIG or PYTHIA out of the box,