Using JetWeb to tune Monte Carlo for hadronic backgrounds from $\gamma\gamma$ events at a linear collider

A. Buckley[1], J. M. Butterworth[2], J. W. Monk[2], E. Nurse[2], M. R. Whalley[1] and M. Wing[2]

1- University of Durham - IPPP
South Road, Durham - England

2- University College London - Dept. of Physics & Astronomy
Gower Street, London - England

We demonstrate the JetWeb system for validating and tuning Monte Carlo simulations and show how it can be applied to Monte Carlo simulations for hadronic backgrounds at a linear lepton collider.

1 Introduction

Collisions at a future linear lepton collider will exist inside a background of a very large number of photons. These photons arise from both bremsstrahlung and beamstrahlung [2]. Due to the hadronic content of the photon, these photons lead to an important QCD background. Accurately simulating this background with Monte Carlo will therefore be useful in understanding its effect.

$\gamma\gamma$ interactions in which the partonic structure of the photon is revealed are similar to resolved photoproduction at the HERA lepton-proton collider, the difference being that one of the photons is replaced by a proton. $\gamma\gamma$ events at LEP also resolve the structure of the photon, although the collision energy is lower than at either HERA or a future linear collider. Monte Carlo simulations can therefore be tuned to best fit the photon data from these two colliders and then extrapolated to the higher energies of a linear collider.

2 JetWeb

JetWeb [3] is a web application and database of Monte Carlo results that has been developed by the CEDAR [4] collaboration. JetWeb allows a user to supply a set of Monte Carlo parameters (either input by hand into the web interface or uploaded in a HepML XML file) and perform a search for models that match those parameters. If a matching model already exists within JetWeb's database then a comparison between the Monte Carlo and experimental results can immediately be made across a wide range of datasets from many experiments. The comparison allows a simple visual inspection of the plots or the χ^2 can be calculated between the Monte Carlo and experimental data for all or a subset of the plots. The latter is important because it allows one to determine if the Monte Carlo is valid for all interactions, not just those of immediate interest.

If JetWeb's database does not already contain a model that matches the search then results are not returned immediately. Instead, the user is given the opportunity to request generation of Monte Carlo data to match the new model. The data is generated on a computing grid using scripts written out by JetWeb that steer either HZTool or Rivet (section 4).

Although its primary intended use is validation of existing Monte Carlo tunings, JetWeb can be used in this case to provide a simple way of comparing tunings to photon data.

3 HepData

HepData [3] is used by JetWeb as the source of the comparison plots against which Monte Carlos are validated. The original HepData [5] has been maintained as a hierarchical database at Durham for around thirty years. However, this legacy database, which is usually accessed via Fortran routines, lacks many of the features required for a more modern Java application such as JetWeb. For this reason, the CEDAR collaboration is migrating HepData to a relational MySQL database.

The new HepData allows the data to be represented not only as a database, which can be accessed using remote SQL queries if necessary, but also as a Java object model. The Java object model is a set of Java objects that represent structures from the data, and these objects are automatically filled from the database using the Hibernate system [6]. Hibernate protects the object model from needing to know the details of the underlying database.

4 Rivet

Rivet [7, 8] has been developed by the CEDAR collaboration as a C++ replacement for HZTool [9], which was originally written in Fortran by the H1 and Zeus collaborations. The initial purpose of HZTool was to allow analyses by one experiment to be easily reproduced in Monte Carlo by the other, however it can more generally be used by anyone to reproduce any analysis that is implemented within the HZTool framework.

One of the main ways in which Rivet improves upon HZTool is by the use of what are called *projections*. A projection is a C++ object that acts on a Monte Carlo event to project out a quantity or quantities from that event. For example, a projection might do something as simple as project out only the stable final state particles from the full HepMC record, or it may do something more complicated such as run the jet finding algorithm of a particular experiment. To construct an analysis, all one needs to do is therefore chain together the appropriate set of projections. Once a projection has been applied to an event, Rivet caches the results, which are then available to all analyses. In this way, each projection need only be run once per event, even if many analyses are run as is the case with JetWeb.

Rivet provides histogram booking routines for the AIDA and ROOT histogram formats and also a generator steering package called RivetGun. RivetGun provides a simple interface that allows many different generators to be steered from the command line.

JetWeb runs HZTool or Rivet on a computing grid in order to reproduce all of the analyses that are relevant to the Monte Carlo that is requested by a user.

5 Comparison data and tuneable parameters

For this demonstration we vary the parameters `PTJIM` and the photon radius from the Jimmy [10] multiple interactions model used with Herwig [11]. In these proceedings we show only a small number of the available plots comparing values of `PTJIM` of 2 GeV and 3 GeV and photon radii of 0.47 GeV^{-1} and 0.19 GeV^{-1} (figure 1). All other parameters are set at their default values. A much larger set of plots is available from the relevant pages of the online JetWeb server for models 8, 9, 11 and 14 [12]. The data shown here are taken from two H1 photoproduction papers [13, 14] that show the dependence of the cross section on the fraction x_γ of the photon's momentum carried by the parton and the transverse energy, E_T, of the jets. Data from [13] shows the x_γ dependence above a give jet E_T, while data from

		χ^2		
PHRAD	PTJIM	data from [13]	data from [14]	total
0.47 GeV^{-1}	2 GeV	7.6	0.35	5.4
0.47 GeV^{-1}	3 GeV	0.90	0.93	0.91
0.19 GeV^{-1}	2 GeV	6.8	0.39	4.876
0.19 GeV^{-1}	3 GeV	0.86	0.83	0.85

Table 1: The χ^2 per degree of freedom from JetWeb models 8, 9, 11 and 14 with data shown in figure 1. Where PHRAD=0.19 GeV^{-1}, PTJIM, is set to 2 GeV.

[14] shows the E_T dependence in a given x_γ range. The χ^2 per degree of freedom for the plots shown is given in table 1. Note that Herwig is a leading order Monte Carlo, and as such the shape of the distributions produced is more important than the overall normalisation.

An important difference between the two papers is that in [13] the outgoing lepton is tagged, which allows the fraction of the beam lepton's momentum carried by the photon (y) to be determined. In [14], the lepton is not tagged and y must be estimated using calorimeter information. PTJIM is a cut-off scale below which there can be no radiation, so a lower value of PTJIM leads to a higher simulated calorimeter activity. This means that while the values of y determined in [13] are not sensitive to PTJIM, in [14] they are. There is a $1/y$ dependence in x_γ, so the sensitivity to changes in the underlying event model is not the same in both papers.

The analysis in [13] requires the jet E_T to be greater than 4 GeV, whereas in [14] the E_T starts at 7 GeV. This makes the top row of plots from [13] more sensitive to the underlying event, which has a greater effect on lower E_T jets. Indeed, the extent to which the underlying event is dominating the simulated jet can be seen in the second row, in which a jet pedestal for the underlying event has been removed leading to an underestimation of the data by the simulation.

6 Acknowledgments

The CEDAR collaboration thanks the UK Science & Technology Facilities Council (STFC) for their generous support.

7 Bibliography

References

[1] Slides: http://ilcagenda.linearcollider.org/contributionDisplay.py?contribId=444&sessionId=79&confId=1296

[2] D. Schulte, Thesis and TESLA Note 97-08, Study of Electromagnetic and Hadronic Background in the Interaction Region of the TESLA Collider (1996).

[3] A. Buckley, W. J. Stirling, M. R. Whalley, J. M. Butterworth, J. Monk, E. Nurse and B. Waugh, arXiv:hep-ph/0605048.

[4] A. Buckley, arXiv:0708.2655 [hep-ph].

[5] HepData reaction database: http://durpdg.dur.ac.uk/hepdata/reac.html

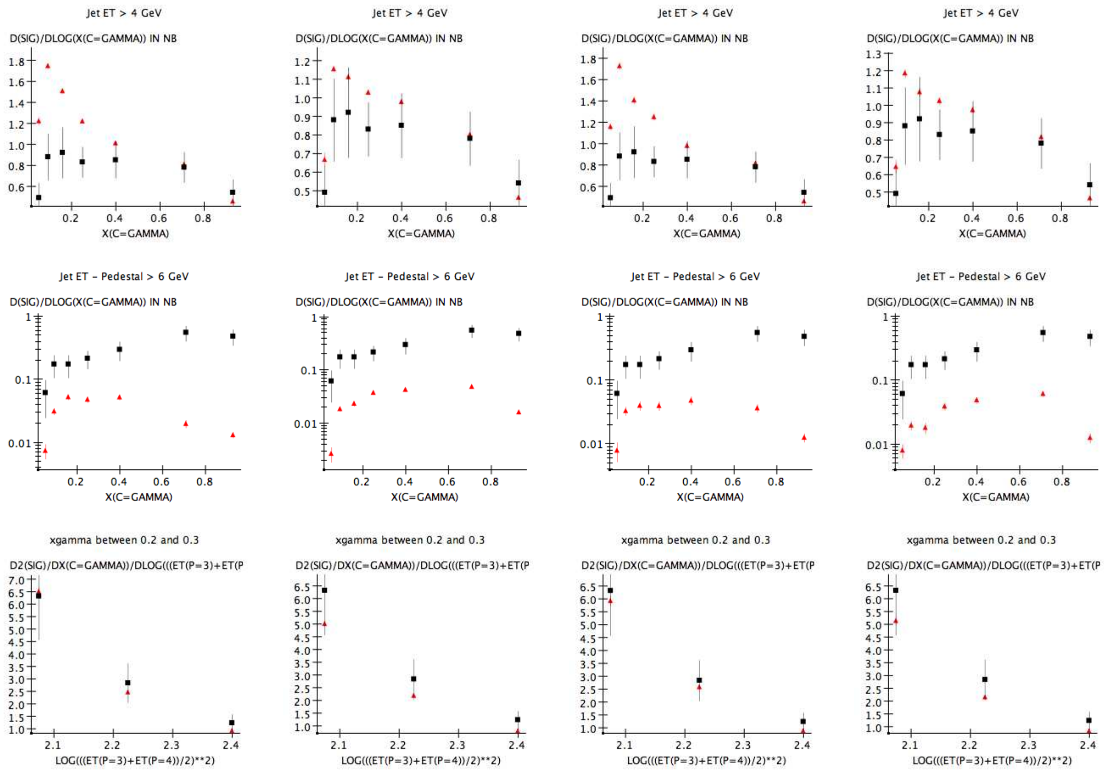

Figure 1: Plots taken from JetWeb comparing Herwig simulation to data from [13](top two rows) and [14] (bottom row) for `PTJIM` values of 2 GeV (first and third columns) or 3 GeV (second and fourth columns) and photon radii of 0.47 GeV^{-1} (left two columns) or 0.19 GeV^{-1} (right two columns).

[6] Hibernate: `http://www.hibernate.org/`

[7] B. M. Waugh, H. Jung, A. Buckley, L. Lonnblad, J. M. Butterworth and E. Nurse, arXiv:hep-ph/0605034.

[8] Rivet project: `http://projects.hepforge.org/rivet/`

[9] J. Bromley *et al.*, *Prepared for Workshop on Future Physics at HERA (Preceded by meetings 25-26 Sep 1995 and 7-9 Feb 1996 at DESY), Hamburg, Germany, 30-31 May 1996*

[10] J. M. Butterworth, J. R. Forshaw and M. H. Seymour, "Multiparton interactions in photoproduction at HERA," Z. Phys. C **72** (1996) 637 [arXiv:hep-ph/9601371].

[11] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour and B. R. Webber, JHEP **0101** (2001) 010 [arXiv:hep-ph/0011363]; hep-ph/0210213.

[12] `http://jetweb.cedar.ac.uk/jetweb-webapp/JWSearch/`

[13] C. Adloff *et al.* [H1 Collaboration], Phys. Lett. B **483** (2000) 36 [arXiv:hep-ex/0003011].

[14] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **1** (1998) 97 [arXiv:hep-ex/9709004].