An Interface to High p_t HERA Data: Quaero@H1

Sascha Caron[1] and Bruce Knuteson[2] *

1- Physikalisches Institut, Universität Freiburg, Germany;
e-mail: scaron@physik.uni-freiburg.de

2- Massachusetts Institute of Technology;
e-mail: knuteson@mit.edu

Distributions from high-p_T HERA event data analyzed in a general search for new physics at H1 have been incorporated into QUAERO, an algorithm designed to automate tests of specific hypotheses with high energy collider data. This article introduces the framework and shows examples to illustrate the algorithm's performance.

1 Introduction

On 30$^{\text{th}}$ June 2007, data taking ended at the HERA accelerator. Data analysis will still continue for a few years. In the years after HERA new theoretical scenarios will be constructed, and HERA data may yet prove to be useful in their testing. HERA data may even prove to be important in constraining theoretical interpretations of findings made at the Large Hadron Collider. Unfortunately, re-interpretation of previous data typically requires a re-analysis with expert collaboration-specific knowledge. The QUAERO framework [3] allows quick testing of any specific hypothesis against collider event data, with the analysis performed by an algorithm that encapsulates the expert knowledge of the experiment. QUAERO@H1 incorporates distributions published by the H1 Collaboration in a general search for new physics into this QUAERO framework [2]. Within this system new theoretical scenarios can be tested with HERA data and exclusion contours (or discovery regions) can be produced on demand. The system is available to users via a web interface [6].

2 Available data and the H1 General Search

The H1 General Search has been published in Ref. [4]. This search investigates events with high-p_T objects (electrons, muons, jets, photons, and the presence of missing transverse energy) produced in ep collisions at HERA. The histograms published by H1 (the invariant masses and the sums of the transverse momenta for high-p_T events) are used as input to the QUAERO algorithm in the studies described in this paper.

The H1 data available within QUAERO correspond to

- 36.4 pb^{-1} of 27.5 GeV positrons on 820 GeV protons, at a center of mass energy of 301 GeV;
- 13.8 pb^{-1} of 27.5 GeV electrons on 920 GeV protons, at a center of mass energy of 319 GeV; and
- 66.4 pb^{-1} of 27.5 GeV positrons on 920 GeV protons, at a center of mass energy of 319 GeV.

Standard object identification criteria are used to define electrons (e), muons (μ), photons (γ), and jets (j) [4]. All identified objects are required to have $p_T > 20$ GeV and

*The authors thank the H1 collaboration for supporting Quaero@H1.

$10^\circ < \theta < 140^\circ$. A neutrino object ($\not p$) is defined for missing transverse momentum above 20 GeV. The experimental sources of systematic error affecting the modeling of these data are identical to those considered in Ref. [4] and are included into QUAERO. Several Monte Carlo event generators are combined to estimate dominant Standard Model processes [4]. These generated events serve as the reference model to which hypotheses presented to QUAERO are compared.

3 Turbosim for H1 and the QUAERO algorithm

To keep QUAERO fast and standalone, a fast detector simulation algorithm (TURBOSIM@H1) is built in accordance with the H1 detector simulation. It is based on a large lookup table of one half million lines mapping particle-level objects to objects reconstructed in the detector. Validation of TURBOSIM@H1 has been performed by running an independent sample of one million events through both the H1 full simulation and TURBOSIM@H1. The event classification and the kinematic distributions of the events from the two simulation chains are compared using a Kolmogorov-Smirnov (KS) test. TURBOSIM@H1 has been found in agreement with the full simulation of H1.

QUAERO provides a convenient interface to the understanding represented by high energy collider data, backgrounds, and detector response. This interface is designed to facilitate the test of any specific hypothesis against such data. A physicist wishing to test her hypothesis against H1 data will provide her hypothesis in the form of commands to one of the built-in event generators. QUAERO uses the specified event generator to generate signal events corresponding to e^+p collisions at 301 GeV, e^-p collisions at 319 GeV, and e^+p collisions at 319 GeV. The response of the H1 detector to these events is simulated using TURBOSIM@H1. Three distinct samples of events exist at this point: the data $\mathcal{D}$; the Standard Model prediction SM; and the hypothesis $\mathcal{H}$, which is the sum of included Standard Model processes and the physicist's signal. Each sample of events is partitioned into exclusive final states, categorized by reconstructed objects with $p_T > 20$ GeV. In each exclusive final state, a pre-defined list of two variables — the summed scalar transverse momentum ($\sum p_T$) and the invariant mass of all objects (m_{all}) — are ranked according to the difference between the Standard Model prediction and the physicist's hypothesis $\mathcal{H}$. The variable showing the most difference is used and densities are estimated from the Monte Carlo events predicted by SM and $\mathcal{H}$. These densities are used to define a discriminant, which is binned to distinguish SM from $\mathcal{H}$. The likelihood ratio $\mathcal{L} = p(\mathcal{D}|\mathcal{H})/p(\mathcal{D}|\text{SM})$ is determined using this binning, and systematic errors are integrated numerically. The result returned by QUAERO is the decimal logarithm of this likelihood ratio. The measurement of model parameters using QUAERO is easily accomplished by graphing $\log_{10}\mathcal{L}$ as a function of varied parameter values, with multiple QUAERO submissions. This information, in addition to plots showing data, Standard Model prediction, and the physicist's hypothesis, are returned in an email.

A rough useful comparison of the sensitivity of QUAERO's results (which take the form of the decimal logarithm of a likelihood ratio) with previous analyses (which typically take the form of 95% confidence level exclusion limits) can be made by comparing $\log_{10}\mathcal{L} = -1$ with the 95% confidence level exclusion limit.

4 One example: Leptoquarks

QUAERO@H1 has been used to search for leptoquarks, particles possessing both lepton and baryon quantum numbers that arise naturally in Grand Unified Theories. Attention is restricted to a scalar leptoquark coupling to a positron and an up quark. The coupling λ of the LQ-e-u vertex and the leptoquark mass m_{LQ} are allowed to vary. The interaction Lagrangian is assumed to be of the form

$$\begin{aligned}\mathcal{L} &= \lambda\,\mathrm{LQ}\,\bar{u}_R\,e_L + \mathrm{h.c.} \\ &\quad + i g_s G^{*}_{\mu}(\mathrm{LQ}^{*}\overleftrightarrow{\partial}^{\mu}\mathrm{LQ}),\end{aligned} \tag{1}$$

where LQ is a scalar leptoquark field; $\bar{u}_R$ and e_L represent a right-handed anti-up quark and left-handed electron; G_μ is the gluon field; and $g_s = \sqrt{4\pi\alpha_s} \approx 1.2$. MadEvent [5] is used to generate events corresponding to these Lagrangian terms within QUAERO .

The subset of D0 Run I data made available in the first implementation of QUAERO [3] has been incorporated into the current version of QUAERO. Plots of QUAERO's result in the parameter plane of λ and m_{LQ} using the H1 and D0 data separately are shown in Fig. 1. QUAERO's result using H1 and D0 data combined is shown in Fig. 1. QUAERO is able to make use of the Tevatron's λ-independent exclusion of leptoquarks with low mass and HERA's λ-dependent exclusion at higher masses to rule out more of the parameter space than either collider is able to on its own.

The use of Quaero@H1 to search for R-parity violating supersymmetry and excited quarks can be found in Ref. [2].

5 Summary

The histograms of the invariant masses and the sum of transverse momenta from the high-p_T events selected in a general search for new physics at H1 have been incorporated into QUAERO, a framework for automating tests of hypotheses against data. The resulting interface (QUAERO@H1) will allow future users to quickly compare new theoretical scenarios to HERA data.

6 Bibliography

References

[1] slides:
`http://indico.cern.ch/contributionDisplay.py?contribId=128&sessionId=9&confId=9499`

[2] S. Caron and B. Knuteson, arXiv:hep-ph/0612201.

[3] V. M. Abazov *et al.* [D0 Collaboration], Phys. Rev. Lett. **87** (2001) 231801 [arXiv:hep-ex/0106039].

[4] A. Aktas *et al.* [H1 Collaboration], Phys. Lett. B **602** (2004) 14 [arXiv:hep-ex/0408044].

[5] F. Maltoni and T. Stelzer, JHEP **0302** (2003) 027 [arXiv:hep-ph/0208156].

[6] Quaero: `http://mit.fnal.gov/Quaero`

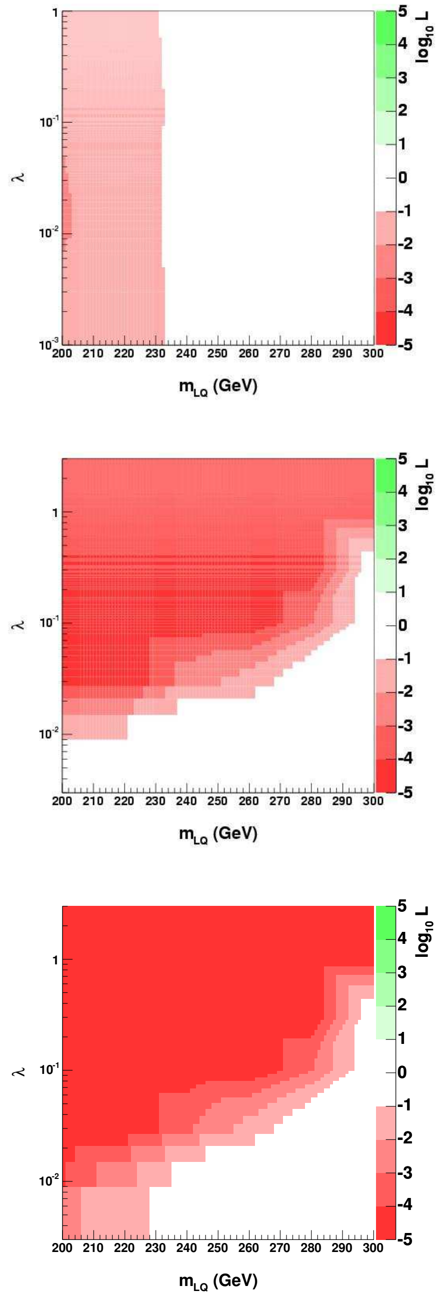

Figure 1: QUAERO's log likelihood ratio as a function of the coupling λ and leptoquark mass $m_{\rm LQ}$ for the Leptoquark example. Shown separately are results from QUAERO using data from DØ Tevatron Run I (upper plot) and using data from H1 HERA Run I (middle plot) and using data from both experiments.