General Searches for New Physics

A. Soha (on behalf of the CDF and H1 Collaborations)
University of California, Department of Physics, One Shields Avenue, Davis, California 95616, USA

A model-independent global search for new physics has been performed at the CDF experiment. This search examines nearly 400 final states, looking for discrepancies between the observed data and the standard model expectation in populations, kinematic shapes, and the tails of the summed transverse momentum distributions. A new approach also searches in approximately 5000 mass variables looking for 'bumps' that may indicate resonant production of new particles. The results of this global search for new physics in $2\,\mathrm{fb}^{-1}$ are presented. In addition, a model-independent search for deviations from the Standard Model prediction is performed in e^+p and e^-p collisions at HERA II using all H1 data recorded during the second running phase. This corresponds to integrated luminosities of $178\,\mathrm{pb}^{-1}$ and $159\,\mathrm{pb}^{-1}$ for e^+p and e^-p collisions, respectively. A statistical algorithm is used to search for deviations in the distributions of the scalar sum of transverse momenta or invariant mass of final state particles, and to quantify their significance.

1. INTRODUCTION

In stark contrast to most searches for new physics, which optimize for a particular model or signature, the general searches presented here are model-independent and include many final state particle combinations in an effort to be highly inclusive. The CDF global search results will be presented first, followed by the H1 general search results.

2. GLOBAL SEARCH AT CDF

The model-independent global search for new high-p_T physics in $p\overline{p}$ collisions at the Tevatron using CDF has three components [1]: VISTA examines populations and kinematic features of the high-p_T data; the BUMP HUNTER [2] searches for resonances in invariant mass combinations; and SLEUTH looks for excesses at high sum-p_T (Σp_T).

The CDF results use data corresponding to a luminosity of $2\,\mathrm{fb}^{-1}$, acquired through inclusive high-p_T electron, muon, photon, and jet triggers. Standard criteria are imposed to identify electrons, muons, taus, photons, jets, b-jets, and missing transverse energy ($\not\!\!E_T$), all with thresholds equivalent to $p_T > 17\,\mathrm{GeV}/c$. Events are further selected to meet offline requirements such as $E_T(e) > 25\,\mathrm{GeV}$, $p_T(\mu) > 25\,\mathrm{GeV}/c$, or $E_T(\gamma) > 60\,\mathrm{GeV}$ [3]. Approximately 4.3 million events are partitioned into 399 exclusive final states, and new categories are created as needed.

The strategy is to use Monte Carlo event generators such as PYTHIA and MADEVENT to represent the Standard Model (SM), and to pass the resulting events through a GEANT-based simulation of the CDF detector response. The simulation is then used in a global fit to the CDF data, to extract 43 corrections factors. The fit is performed simultaneously to all final states and is subjected to external constraints. The correction factors, which include corrections to leading order theory cross sections, object reconstruction efficiencies, and mis-identification rates, are then used to improve the SM prediction. The three components (VISTA, BUMP HUNTER, and SLEUTH) of the global comparison between the data and SM prediction are performed, and the procedure is iterated by feeding information back into the simulation and correction factors until there is either a clear case for new physics or all discrepancies have known sources.

2.1. Population and Kinematic Distribution Results

The VISTA comparison of final state populations between data and SM predictions is shown in Fig. 1(a). The histogram shows the Poisson probability that the SM population in a final state would fluctuate above or below the observed population in data, expressed in units of standard deviations. The plotted probabilities do not include a trials factor, which accounts for the large number of final states that are examined and reduces the significance of

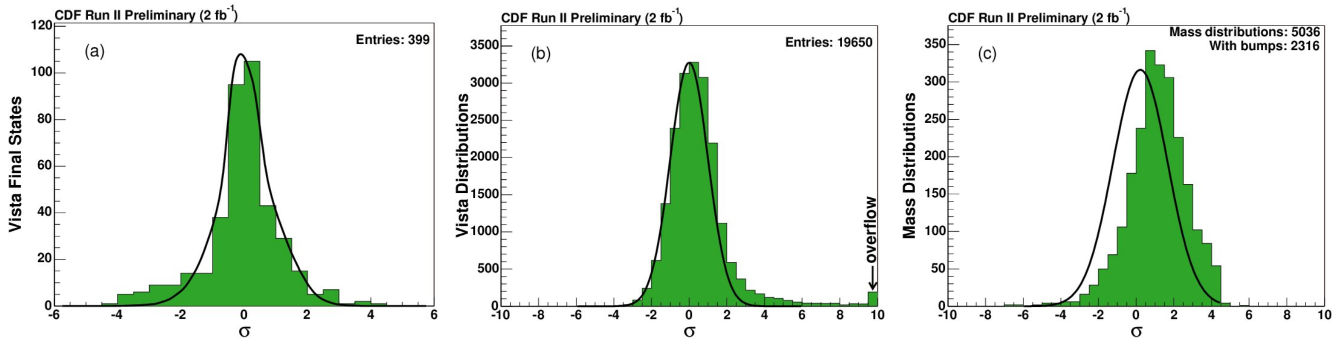

Figure 1: Results for the CDF (a) population and (b) kinematic shape comparisons, as well as the (c) bump hunter, as described in the text.

each observed discrepancy. The greatest observed discrepancy is in the final state $be^{\pm}\not\!\!E_T$ where 817.7 ± 9.2 events are expected and 690 events are observed, for a discrepancy of -4.3σ before the trials factor and -2.7σ after including the trials factor. Therefore, no population shows a significant discrepancy.

VISTA also automatically produces and examines 19650 kinematic distributions. The results are summarized in Fig. 1(b), where the histogram shows the Kilmogorov-Smirnov probability that the distributions in the data and SM prediction are consistent, expressed in units of standard deviations. The trials factor due to examining thousands of distributions has not yet been accounted for in the plot. Distributions are considered discrepant if they disagree by more than 5σ (approximately $> 3\sigma$ after including the trials factor). The 555 distributions that meet this criteria are examined more closely. It turns out that 81% of the discrepancies can be explained by a deficiency in modeling soft jet emission in QCD parton showering. An additional 16% are due to inadequate modeling of the transverse boost of the colliding system and 3% are due to residual crudeness in the correction factor model, mostly from using simplified p_T-dependencies in fake rate correction factors. Therefore, there are no claims for new physics based on the kinematic distribution comparisons.

2.2. Bump Hunter Results

A new resonance might appear as a bump in an invariant mass distribution. The CDF BUMP HUNTER uses the final states from VISTA to form all invariant mass combinations and perform a comparison between data and SM backgrounds. A search window of $2\Delta M$, where ΔM is the expected detector mass resolution, is scanned across each invariant mass distribution. A candidate bump must have at least five data events and side-bands that are in better agreement than the central search window. Pseudo-experiments are then used to estimate the significance of any qualifying bumps. The results are shown in Fig. 1(c), which shows the probability for a corresponding bump from pseudo-data to have a larger significance than the one found in data, cast in terms of standard deviations. Of the 5036 scanned distributions, 2316 have qualifying bumps. The visible shift in the histogram is caused by local deficiencies in the SM prediction, but does not invalidate the method since the shift makes it more likely that a bump surpasses the threshold for further study. The threshold for further investigation is 5σ, which corresponds to 3σ after including the trials factor for 5036 mass distributions. There is one bump beyond this threshold, in a final state with four jets and low Σp_T, but it is found to be due to the same soft jet modeling problem mentioned in section 2.1. Hence, no new physics is found by the BUMP HUNTER in $2\,\mathrm{fb}^{-1}$.

2.3. Search at High Sum-p_T

SLEUTH assumes that new physics will appear as an excess, and that the excess will be at high Σp_T and in one final state. The Σp_T is the scalar sum of the p_T of the individual objects, unclustered energy, and $\not\!\!E_T$. For each final

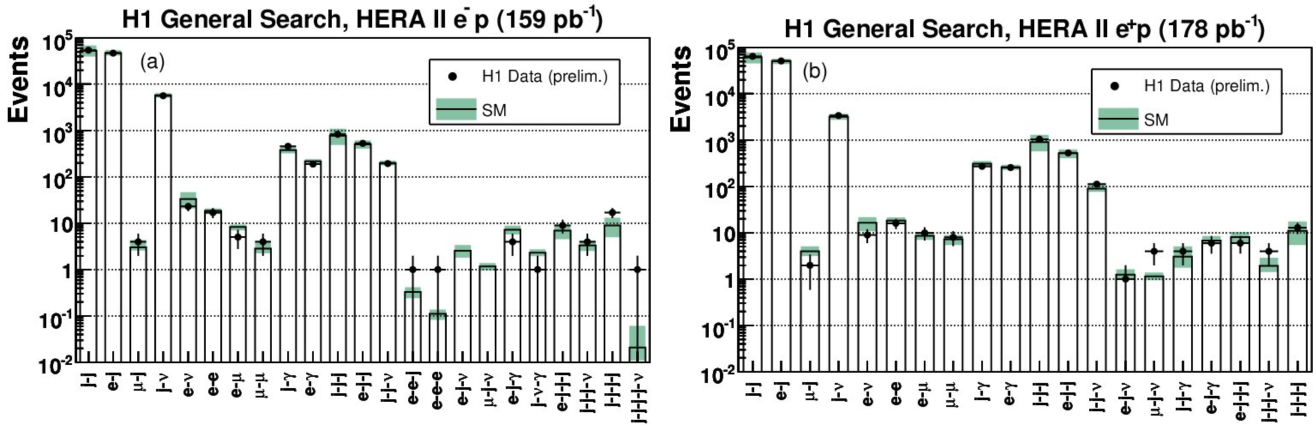

Figure 2: Event yield comparisons between data and SM predictions using (a) e^-p and (b) e^+p collisions. The plots include categories with ≥ 1 data event or > 1 expected SM events. Error bars include theory uncertainty and experimental systematics.

state, the Σp_T distribution is scanned, and the one-sided region with the most significant excess of data is selected. The significance, $\mathcal{P}$, is determined as the fraction of pseudo-experiments that find a region at least as discrepant as the one observed in data. The final state with the largest discrepancy is $e^{\pm}\mu^{\pm}$, with $\mathcal{P} = 0.00055$, corresponding to 3.26σ before including a trials factor. It is found that 8% of experiments like CDF would find an excess at least as large as this most discrepant Sleuth final state. There are no claims of new physics using Sleuth with $2\,\mathrm{fb}^{-1}$.

3. GENERAL SEARCH AT H1

The global search for new physics at H1 uses data corresponding to luminosities of $178\,\mathrm{pb}^{-1}$ and $159\,\mathrm{pb}^{-1}$ from e^+p and e^-p collisions, respectively [4, 5]. Isolated electrons, muons, photons, jets, and neutrinos are included if they have $p_T > 20\,\mathrm{GeV}/c$ and a polar angle satisfying $10^\circ < \theta < 140^\circ$. The event selection requires exclusive final states with two or more objects, and events are classified by the number and types of objects. This procedure examines all combinations and finds that 23 final states are populated. Simulations are used for all contributing SM processes, including photoproduction, deep-inelastic scattering, QED Compton scattering, electroweak production, and QCD. The resulting predictions are used in comparisons to the event yields, Σp_T, and invariant mass distributions found in data. A statistical algorithm is then employed to identify the largest deviations and evaluate the associated probabilities.

3.1. Event Yields, Sum-p_T, and Invariant Mass

The event yield comparisons show good agreement in all final states, in both the e^-p and e^+p data, as shown in Fig. 2. The Σp_T and invariant mass distribution comparisons are conducted by finding regions of greatest deviation between data and the SM expectation for each final state. All groups of neighboring 5 GeV bins are tested. A measure, p, is determined as the probability for a positive or negative fluctuation of the SM expectation to be at least as large as that observed in data. The procedure accounts for Poisson statistical errors and Gaussian systematic uncertainties. For each final state, the region with the smallest p-value is selected. The results for the invariant mass comparisons using e^-p data are shown in Fig. 3(a) and the results for the Σp_T comparison using e^+p data are shown in Fig. 3(b).

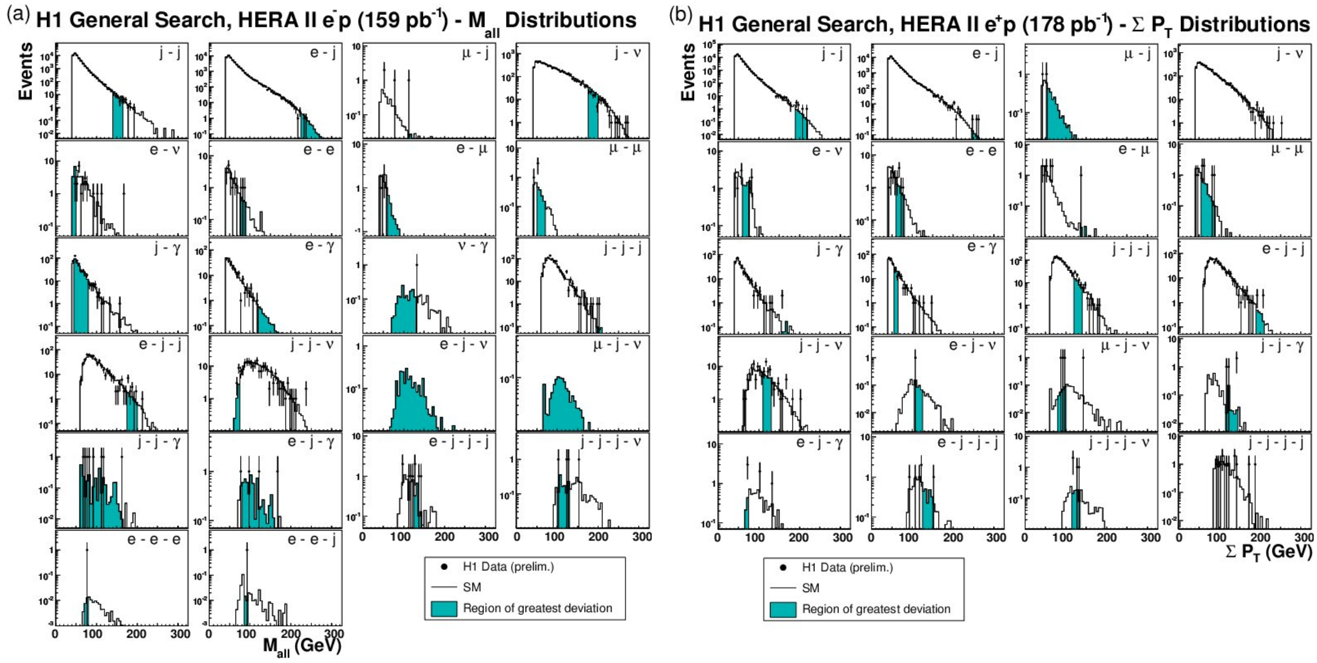

Figure 3: Distributions of data events and SM expectations for the (a) invariant mass in e^-p collisions and (b) Σp_T in e^+p collisions. The regions of greatest deviation are shaded.

3.2. Significance

The significance of each deviation in the H1 Σp_T and invariant mass comparison is evaluated using pseudo-data. The method determines the probability, $\hat{\mathrm{P}}$, to observe a region with a p-value less than the smallest p-value seen in data. Calculating this $\hat{\mathrm{P}}$ allows for the comparison of deviations across different final state categories. Fig. 4(a) shows the $-\log_{10}\hat{\mathrm{P}}$ distribution for the invariant mass comparison using e^-p data, while Fig. 4(b) shows the $-\log_{10}\hat{\mathrm{P}}$ distribution for the Σp_T comparison using e^+p data. Note that a 5σ discrepancy would correspond to a value of $-\log_{10}\hat{\mathrm{P}}$ between 5 and 6. No such significant discrepancies between data and SM expectations are observed. The largest deviation is in the $\mu j \nu$ final state category.

4. CONCLUSIONS

The CDF and H1 general searches for new physics have probed large datasets for indications of new physics in population and kinematic distributions, using a large number of final states. These searches provide broad views of the high-p_T data samples and demonstrate understanding of the detectors and SM simulation. They do not rule out all sources of new physics, thus leaving open the possibility for future discoveries.

Acknowledgments

The author wishes to thank the 2008 ICHEP organizers and hosts, as well as the CDF and H1 collaborations and funding sources.

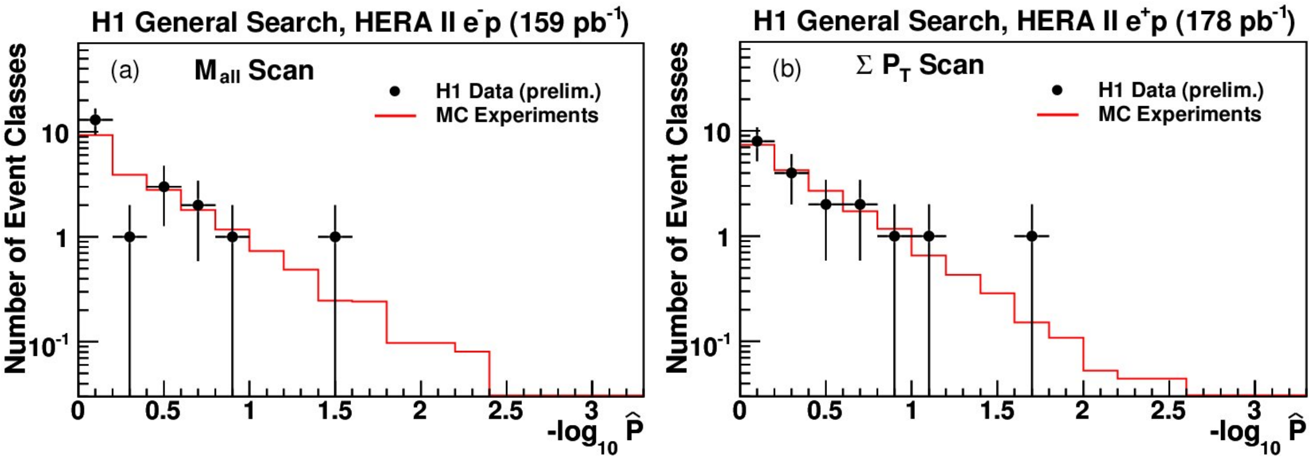

Figure 4: Distributions of the quantity $-\log_{10}\hat{\mathrm{P}}$ for all of the final states used in the comparisons of the (a) invariant mass in e^-p data and (b) Σp_T in e^+p data. The plotted lines show the expectation from Monte Carlo experiments.

References

[1] T. Aaltonen *et al.* (CDF Collaboration), Submitted to Phys. Rev. D (2008), arXiv:0809.3781.

[2] G. Choudalakis, Ph.D. thesis, Massachusetts Institute of Technology (2008), arXiv:0805.3954.

[3] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **78**, 012002 (2008).

[4] E. Sauvan *et al.* (H1 Collaboration), “A General Search for New Phenomena at HERA”, Proc. of 15th Int. Workshop on Deep-Inelastic Scattering and Related Subjects, Munich, April 2007, arXiv:0706.4110.

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