Sleuth: A Quasi-Model-Independent Search Strategy for New High p_T Physics

B. Knuteson

Department of Physics, University of California at Berkeley, Berkeley, California 94720

Abstract. We present a quasi-model-independent search for the physics responsible for electroweak symmetry breaking. We define final states to be studied, and construct a rule that identifies a set of relevant variables for any particular final state. A new algorithm ("Sleuth") searches for regions of excess in those variables and quantifies the significance of any detected excess. After demonstrating the sensitivity of the method, we apply it to the semi-inclusive channel $e\mu X$ collected in 108 pb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV at the DØ experiment during 1992–1996 at the Fermilab Tevatron. We find no evidence of new high p_T physics in this sample.

INTRODUCTION

It is generally recognized that the standard model, an extremely successful description of the fundamental particles and their interactions, must be incomplete; yet the possibilities beyond the current paradigm are sufficiently broad that the first hint could appear in any of many different guises. This suggests the importance of performing searches that are as model-independent as possible. In these proceedings I describe a novel search strategy ("Sleuth") for physics beyond the standard model that assumes nothing about the expected characteristics of new processes other than that they will produce an excess of events at high transverse momentum (p_T). The search strategy has been applied to the exclusive final states within $e\mu X$ in $\approx$ 108 pb^{-1} of $p\bar{p}$ collisions collected by DØ at the Fermilab Tevatron. Space limitations allow only a brief overview of the algorithm and a terse statement of the results on the final states within $e\mu X$; a more complete description can be found in the original Sleuth article. (1)

SLEUTH

We assume that the production and subsequent decay of massive new particles will yield events containing final state objects with large transverse momenta. The data are partitioned into exclusive final states using standard criteria that identify isolated high p_T electrons (e), muons (μ), and photons (γ), as well as jets (j), missing transverse energy ($\not\!\!E_T$), and the presence of W and Z bosons. For each exclusive final state, we consider a small set of variables that are easily related to the transverse momenta of the final state objects.

Although the details of the Sleuth algorithm are complicated, the concept is straightforward. What is needed is a data sample, a set of events modeling each background process i, and the number of background events $\hat{b}_i \pm \delta\hat{b}_i$ expected from each background process. From these we determine the region of greatest excess and quantify the degree to which that excess is interesting.

The Sleuth algorithm, applied to each final state individually, comprises the following steps:

(i) We begin by constructing a mapping from the original d-dimensional variable space into the d-dimensional unit box (i.e., $[0,1]^d$) that flattens the background distribution. We use this to map the data into the unit box.

(ii) We rigorously define the notion of a "region" about a set of N data points.

(iii) Each region R contains an expected number of background events $\hat{b}_R \pm \delta\hat{b}_R$. The probability that the background in the region fluctuates up to or above the observed number of events in the region is our first measure of the degree of interest of a particular region.

(iv) Our assumption that new physics is most likely to appear at high p_T translates to a preference for regions in a particular corner of the unit box. Such regions are searched to determine the region $\mathcal{R}$ of greatest excess.

(v) In any reasonably-sized data set, there will always be regions in which the probability for $\hat{b}_R \pm \delta\hat{b}_R$ to fluctuate up to or above the observed number of events is small. The relevant issue is what fraction $\mathcal{P}$ of *hypothetical similar experiments* contains a region as interesting as $\mathcal{R}$. In computing $\mathcal{P}$ we rigorously "take

Table 1. The number of background events expected for the populated final states within $e\mu X$. $\not{e}\mu$ denotes any process with a real muon and a jet that is mistakenly identified as an electron.

Data set	$\not{e}\mu$	$Z/\gamma^* \to \tau\tau$	WW	$t\bar{t}$	Total	Data
$e\mu\not{E}_T$	18.4±1.4	26.1 ± 6.5	3.9±1.0	0.011±0.003	48.5±7.6	39
$e\mu\not{E}_T j$	8.7±1.0	3.2 ± 0.8	1.1±0.3	0.4±0.1	13.2±1.5	13
$e\mu\not{E}_T jj$	2.7±0.6	0.5 ± 0.2	0.18±0.05	1.8±0.5	5.2±0.8	5
$e\mu\not{E}_T jjj$	0.4±0.2	0.07 ± 0.05	0.032±0.009	0.7±0.2	1.3±0.3	1

Table 2. Summary of results on the $e\mu\not{E}_T$, $e\mu\not{E}_T j$, $e\mu\not{E}_T jj$, and $e\mu\not{E}_T jjj$ channels as WW and then $t\bar{t}$ are added to the background estimate.

	$\mathcal{P}$ when the background includes		
Data set	$\not{e}\mu + Z/\gamma^*$	$\not{e}\mu + Z/\gamma^* + WW$	$\not{e}\mu + Z/\gamma^* + WW + t\bar{t}$
$e\mu\not{E}_T$	0.008 ($+2.4\sigma$)	0.16 ($+1.0\sigma$)	0.14 ($+1.1\sigma$)
$e\mu\not{E}_T j$	0.34 ($+0.4\sigma$)	0.45 ($+0.1\sigma$)	0.45 ($+0.1\sigma$)
$e\mu\not{E}_T jj$	0.01 ($+2.3\sigma$)	0.03 ($+1.9\sigma$)	0.31 ($+0.5\sigma$)
$e\mu\not{E}_T jjj$	0.38 ($+0.3\sigma$)	0.41 ($+0.2\sigma$)	0.71 (-0.5σ)
$\tilde{\mathcal{P}}$	0.03 ($+1.9\sigma$)	0.11 ($+1.2\sigma$)	0.72 (-0.6σ)

into account" the many regions that have been considered within this final state.

(vi) After obtaining $\mathcal{P}$ for each final state, the values are combined into a single quantity $\tilde{\mathcal{P}}$, which takes into account the many final states that have been considered. $\tilde{\mathcal{P}}$ takes on values in the half-open interval $(0,1]$; the potential presence of new high p_T physics would be indicated by finding $\tilde{\mathcal{P}}$ to be small.

APPLICATION TO $e\mu X$

The $e\mu X$ data set and basic selection criteria are identical to those used in the published $t\bar{t}$ cross section analysis for the dilepton channels (2), requiring one or more isolated electrons and one or more isolated muons, each with $p_T > 15$ GeV. The number of events expected for the dominant backgrounds in the populated final states are given in Table 1. Other final states that are analyzed, but which happen to contain zero events, include $e\mu\not{E}_T jjjj$, $ee\mu\not{E}_T$, $e\mu\mu$, and $e\mu\not{E}_T\gamma$.

Top quark and W boson pair production are excellent examples of the types of processes that we would expect Sleuth to find, and we use these to test the sensitivity of the method. We find $\tilde{\mathcal{P}} > 2.0$ standard deviations (σ) in over 50% of an ensemble of mock samples when the background includes $\not{e}\mu + Z/\gamma^*$ only (i.e., all but WW and $t\bar{t}$), and we find $\tilde{\mathcal{P}} > 2.0\sigma$ in nearly 30% of an ensemble of mock samples when the background includes $\not{e}\mu + Z/\gamma^* + WW$ (i.e., all but $t\bar{t}$). The results for DØ data are shown in Table 2. It is instructive to compare Table 2 to the results obtained from a dedicated top quark search in the same $e\mu X$ data (2), in which an excess corresponding to 2.75σ is observed, recalling that Sleuth does not "know" anything about the signals (WW and $t\bar{t}$) that we have asked it to find. (1)

CONCLUSIONS

We have developed a quasi-model-independent technique for searching for new high p_T physics. After defining final states and constructing a rule that identifies a set of relevant variables for any particular final state, we systematically search for regions of excess within the space of those variables, and rigorously quantify the significance of any observed excess. Sleuth has been applied to DØ $e\mu X$ data. Removing WW and $t\bar{t}$ from the estimated background, we find indications of these signals in an ensemble of mock experiments and in the data. Including all background processes, we find no evidence of new physics at high p_T in these data.

REFERENCES

1. DØ Collaboration, B. Abbott *et al.*, *Phys. Rev. D* **62**, 92004 (2000).
2. DØ Collaboration, S. Abachi *et al.*, *Phys. Rev. Lett.* **79**, 1203 (1997).