



## Model-Independent (VISTA) and Quasi-Model-Independent (SLEUTH) Search for New High- $p_T$ Physics at CDF

The CDF Collaboration  
URL <http://www-cdf.fnal.gov>  
(Dated: May 4, 2007)

Data collected in Run II of the Fermilab Tevatron are searched for indications of new electroweak scale physics. Rather than focusing on particular new physics scenarios, CDF data are analyzed for discrepancies with respect to the Standard Model prediction. A model-independent approach (VISTA) considers the gross features of the data, and is sensitive to new large cross section physics. A quasi-model-independent approach (SLEUTH) emphasizes the high- $p_T$  tails, and is particularly sensitive to new electroweak scale physics. This global search for new physics in  $927 \text{ pb}^{-1}$  of  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96 \text{ TeV}$  reveals no indication of physics beyond the Standard Model.

*Preliminary Results for Summer 2007 Conferences*

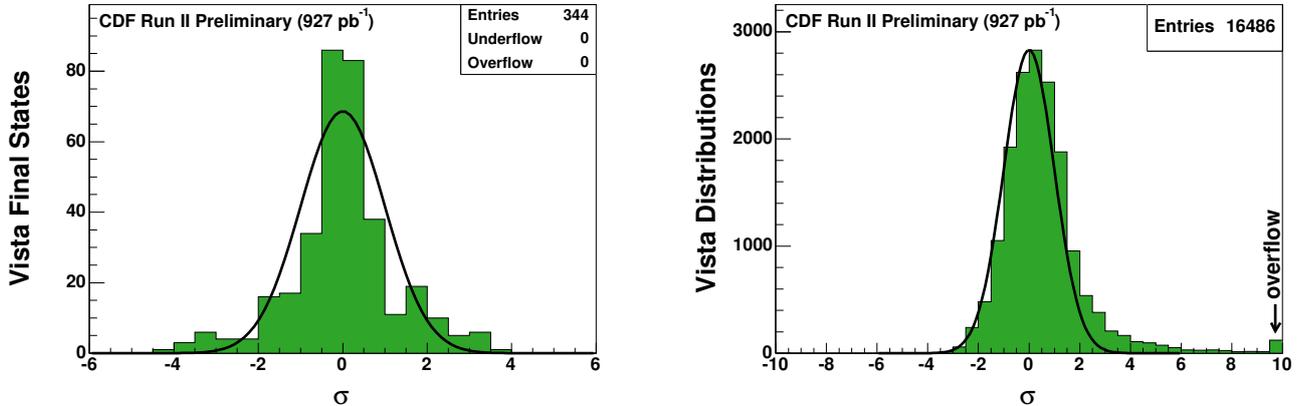


FIG. 1: Distribution of VISTA discrepancy between data and the Standard Model prediction, measured in units of standard deviation ( $\sigma$ ). The left pane shows the distribution of discrepancies between the total number of events observed and predicted in the final states considered. The right pane shows the distribution of discrepancies between the observed and predicted shapes of roughly  $10^4$  kinematic distributions. In the right pane, distributions in which data and the Standard Model prediction are in relative agreement correspond to negative  $\sigma$ , and distributions in which the data and the Standard Model prediction are in relative disagreement correspond to large positive  $\sigma$ . The two distributions have a roughly Gaussian core, centered at zero with unit width. Interest is focused on the entries in the tails of the distribution at left and the high tail of the distribution at right.

## I. INTRODUCTION

The particle physics Standard Model (SM) is remarkably successful, but is believed to require expansion above the electroweak scale. A variety of possible extensions have been proposed. Many analyses optimized for specific signatures have been performed to search for evidence of these possibilities. Limits have been set on cross sections for postulated processes and on masses of hypothetical particles, but no clear evidence of physics beyond the Standard Model has yet been seen.

This Letter presents a broad search for new physics at the electroweak scale. Rather than focusing on any specific proposed scenario, high- $p_T$  data collected by the CDF experiment in Run II of the Fermilab Tevatron are analyzed for discrepancies relative to the Standard Model prediction. A model-independent approach (VISTA) considers gross features of the data, and is sensitive to new large cross section physics. A quasi-model-independent approach (SLEUTH) emphasizes high- $p_T$  tails, and is particularly sensitive to new electroweak scale physics. Searches in a similar spirit have previously been performed by the  $D\bar{O}$  Collaboration [1, 2, 3] in Tevatron Run I and by the H1 Collaboration [4] at HERA-I.

This analysis uses data corresponding to an integrated luminosity of  $927 \text{ pb}^{-1}$  of  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96 \text{ TeV}$ , recorded by the CDF detector [5]. CDF consists of a charged particle tracking system composed of silicon strip detectors and a gas drift chamber inside a 1.4 T magnetic field, surrounded by electromagnetic and hadronic calorimeters and enclosed by muon detectors. The central section of the CDF detector is defined to be  $|\eta_{\text{det}}| < 1.0$  [17]; the plug region corresponds to  $1.0 < |\eta_{\text{det}}| < 2.5$ .

## II. VISTA

VISTA is a framework for obtaining a panoramic view of the bulk of an experiment's high- $p_T$  data. The algorithm involves the identification of isolated and energetic objects; the incorporation of events passing an offline trigger; the generation of a complete Standard Model background estimate; the simulation of the detector response; the partitioning of events into exclusive final states; the development of a correction model; and a global comparison of data to the Standard Model prediction.

A standard set of object identification criteria are used to identify isolated and energetic objects produced in the hard collision. Electrons ( $e^\pm$ ) are characterized by an isolated track in the central gas tracking chamber and a narrow shower in the electromagnetic calorimeter. Muons ( $\mu^\pm$ ) are identified by a track in the central tracking chamber matched to a muon stub in the central muon detectors, with energy consistent with minimum ionizing deposition in the electromagnetic and hadron calorimeters along the muon trajectory. Narrow central jets with a single charged

CDF Run II Preliminary									
Final State	Data	Background	Final State	Data	Background	Final State	Data	Background	Final State
3j $\tau$ +	71	113.7 $\pm$ 3.6	2e+j	13	9.8 $\pm$ 2.2	e+ $\gamma$ $\not{p}$	141	144.2 $\pm$ 6	
5j	1661	1902.9 $\pm$ 50.8	2e+e-	12	4.8 $\pm$ 1.2	e+ $\mu$ - $\not{p}$	54	42.6 $\pm$ 2.7	
2j $\tau$ +	233	296.5 $\pm$ 5.6	2e+	23	36.1 $\pm$ 3.8	e+ $\mu$ + $\not{p}$	13	10.9 $\pm$ 1.3	
be+j	2207	2015.4 $\pm$ 28.7	2b $\Sigma p_T > 400$ GeV	327	335.8 $\pm$ 7	e+ $\mu$ -	153	127.6 $\pm$ 4.2	
3j $\Sigma p_T < 400$ GeV	35436	37294.6 $\pm$ 524.3	2b $\Sigma p_T < 400$ GeV	187	173.1 $\pm$ 7.1	e+j	386880	392614 $\pm$ 5031.8	
e+3j $\not{p}$	1954	1751.6 $\pm$ 42	2b3j $\Sigma p_T < 400$ GeV	28	33.5 $\pm$ 5.5	e+j2 $\gamma$	14	15.9 $\pm$ 2.9	
be+2j	798	695.3 $\pm$ 13.3	2b2j $\Sigma p_T > 400$ GeV	355	326.3 $\pm$ 8.4	e+j $\tau$ +	79	79.3 $\pm$ 2.9	
3j $\not{p}$ $\Sigma p_T > 400$ GeV	811	967.5 $\pm$ 38.4	2b2j $\Sigma p_T < 400$ GeV	56	80.2 $\pm$ 5	e+j $\tau$ -	162	148.8 $\pm$ 7.6	
e+ $\mu$ +	26	11.6 $\pm$ 1.5	2b2j $\gamma$	16	15.4 $\pm$ 3.6	e+j $\not{p}$	58648	57391.7 $\pm$ 661.6	
e+ $\gamma$	636	551.2 $\pm$ 11.2	2b $\gamma$	37	31.7 $\pm$ 4.8	e+j $\gamma$ $\not{p}$	52	76.2 $\pm$ 9	
e+3j	28656	27281.5 $\pm$ 405.2	2bj $\Sigma p_T > 400$ GeV	415	393.8 $\pm$ 9.1	e+j $\mu$ - $\not{p}$	22	13.1 $\pm$ 1.7	
b5j	131	95 $\pm$ 4.7	2bj $\Sigma p_T < 400$ GeV	161	195.8 $\pm$ 8.3	e+j $\mu$ -	28	26.8 $\pm$ 2.3	
j2 $\tau$ +	50	85.6 $\pm$ 8.2	2bj $\not{p}$ $\Sigma p_T > 400$ GeV	28	23.2 $\pm$ 2.6	e+e-4j	103	113.5 $\pm$ 5.9	
j $\tau$ + $\tau$ -	74	125 $\pm$ 13.6	2bj $\gamma$	25	24.7 $\pm$ 4.3	e+e-3j	456	473 $\pm$ 14.6	
b $\not{p}$ $\Sigma p_T > 400$ GeV	10	29.5 $\pm$ 4.6	2be+2j $\not{p}$	15	12.3 $\pm$ 1.6	e+e-2j $\not{p}$	30	39 $\pm$ 4.6	
e+j $\gamma$	286	369.4 $\pm$ 21.1	2be+2j	30	30.5 $\pm$ 2.5	e+e-2j	2149	2152 $\pm$ 40.1	
e+j $\not{p}$ $\tau$ -	29	14.2 $\pm$ 1.8	2be+j	28	29.1 $\pm$ 2.8	e+e- $\tau$ +	14	11.1 $\pm$ 2	
2j $\Sigma p_T < 400$ GeV	96502	92437.3 $\pm$ 1354.5	2be+	48	45.2 $\pm$ 3.7	e+e- $\not{p}$	491	487.9 $\pm$ 12	
be+3j	356	298.6 $\pm$ 7.7	$\tau$ + $\tau$ -	498	428.5 $\pm$ 22.7	e+e- $\gamma$	127	132.3 $\pm$ 4.2	
8j	11	6.1 $\pm$ 2.5	$\gamma$ $\tau$ +	177	204.4 $\pm$ 5.4	e+e-j	10726	10669.3 $\pm$ 123.5	
7j	57	35.6 $\pm$ 4.9	$\gamma$ $\not{p}$	1952	1945.8 $\pm$ 77.1	e+e- $\not{p}$	157	144 $\pm$ 11.2	
6j	335	298.4 $\pm$ 14.7	$\mu$ + $\tau$ +	18	19.8 $\pm$ 2.3	e+e- $\gamma$	26	45.6 $\pm$ 4.7	
4j $\Sigma p_T > 400$ GeV	39665	40898.8 $\pm$ 649.2	$\mu$ + $\tau$ -	151	179.1 $\pm$ 4.7	e+e-	58344	58575.6 $\pm$ 603.9	
4j $\Sigma p_T < 400$ GeV	8241	8403.7 $\pm$ 144.7	$\mu$ + $\not{p}$	321351	320500 $\pm$ 3475.5	b6j	24	15.5 $\pm$ 2.3	
4j2 $\gamma$	38	57.5 $\pm$ 11	$\mu$ + $\not{p}$ $\tau$ -	22	25.8 $\pm$ 2.7	b4j $\Sigma p_T > 400$ GeV	13	9.2 $\pm$ 1.8	
4j $\tau$ +	20	36.9 $\pm$ 2.4	$\mu$ + $\gamma$	269	285.5 $\pm$ 5.9	b4j $\Sigma p_T < 400$ GeV	464	499.2 $\pm$ 12.4	
4j $\not{p}$ $\Sigma p_T > 400$ GeV	516	525.2 $\pm$ 34.5	$\mu$ + $\gamma$ $\not{p}$	269	282.2 $\pm$ 6.6	b3j $\Sigma p_T > 400$ GeV	5354	5285 $\pm$ 72.4	
4j $\gamma$ $\not{p}$	28	53.8 $\pm$ 11	$\mu$ + $\mu$ - $\not{p}$	49	61.4 $\pm$ 3.5	b3j $\Sigma p_T < 400$ GeV	1639	1558.9 $\pm$ 24.1	
4j $\gamma$	3693	3827.2 $\pm$ 112.1	$\mu$ + $\mu$ - $\gamma$	32	29.9 $\pm$ 2.6	b3j $\not{p}$ $\Sigma p_T > 400$ GeV	111	116.8 $\pm$ 11.2	
4j $\mu$ +	576	568.2 $\pm$ 26.1	$\mu$ + $\mu$ -	10648	10845.6 $\pm$ 96	b3j $\gamma$	182	194.1 $\pm$ 8.8	
4j $\mu$ + $\not{p}$	232	224.7 $\pm$ 8.5	j2 $\gamma$	2196	2200.3 $\pm$ 35.2	b3j $\mu$ + $\not{p}$	37	34.1 $\pm$ 2	
4j $\mu$ + $\mu$ -	17	20.1 $\pm$ 2.5	j2 $\gamma$ $\not{p}$	38	27.3 $\pm$ 3.2	b3j $\mu$ +	47	52.2 $\pm$ 3	
3 $\gamma$	13	24.2 $\pm$ 3	j $\tau$ +	563	585.7 $\pm$ 10.2	b2 $\gamma$	15	14.6 $\pm$ 2.1	
3j $\Sigma p_T > 400$ GeV	75894	75939.2 $\pm$ 1043.9	j $\not{p}$ $\Sigma p_T > 400$ GeV	4183	4209.1 $\pm$ 56.1	b2j $\Sigma p_T > 400$ GeV	8812	8576.2 $\pm$ 97.9	
3j2 $\gamma$	145	178.1 $\pm$ 7.4	j $\gamma$	49052	48743 $\pm$ 546.3	b2j $\Sigma p_T < 400$ GeV	4691	4646.2 $\pm$ 57.7	
3j $\not{p}$ $\Sigma p_T < 400$ GeV	20	30.9 $\pm$ 14.4	j $\gamma$ $\tau$ +	106	104 $\pm$ 4.1	b2j $\not{p}$ $\Sigma p_T > 400$ GeV	198	209.2 $\pm$ 8.3	
3j $\gamma$ $\tau$ +	13	11 $\pm$ 2	j $\gamma$ $\tau$ -	913	965.2 $\pm$ 41.5	b2j $\gamma$	429	425.1 $\pm$ 13.1	
3j $\gamma$ $\not{p}$	83	102.9 $\pm$ 11.1	j $\mu$ +	33462	34026.7 $\pm$ 510.1	b2j $\mu$ + $\not{p}$	46	40.1 $\pm$ 2.7	
3j $\gamma$	11424	11506.4 $\pm$ 190.6	j $\mu$ + $\tau$ -	29	37.5 $\pm$ 4.5	b2j $\mu$ +	56	60.6 $\pm$ 3.4	
3j $\mu$ + $\not{p}$	1114	1118.7 $\pm$ 27.1	j $\mu$ + $\not{p}$ $\tau$ -	10	9.6 $\pm$ 2.1	b $\tau$ +	19	19.9 $\pm$ 2.2	
3j $\mu$ + $\mu$ -	61	84.5 $\pm$ 9.2	j $\mu$ + $\not{p}$	45728	46316.4 $\pm$ 568.2	b $\gamma$	976	1034.8 $\pm$ 15.6	
3j $\mu$ +	2132	2168.7 $\pm$ 64.2	j $\mu$ + $\gamma$ $\not{p}$	78	69.8 $\pm$ 9.9	b $\gamma$ $\not{p}$	18	16.7 $\pm$ 3.1	
3bj $\Sigma p_T > 400$ GeV	14	9.3 $\pm$ 1.9	j $\mu$ + $\gamma$	70	98.4 $\pm$ 12.1	b $\mu$ +	303	263.5 $\pm$ 7.9	
2 $\tau$ +	316	290.8 $\pm$ 24.2	j $\mu$ + $\mu$ -	1977	2093.3 $\pm$ 74.7	b $\mu$ + $\not{p}$	204	218.1 $\pm$ 6.4	
2 $\gamma$ $\not{p}$	161	176 $\pm$ 9.1	e+4j	7144	6661.9 $\pm$ 147.2	b $j$ $\Sigma p_T > 400$ GeV	9060	9275.7 $\pm$ 87.8	
2 $\gamma$	8482	8349.1 $\pm$ 84.1	e+4j $\not{p}$	403	363 $\pm$ 9.9	b $j$ $\Sigma p_T < 400$ GeV	7236	7030.8 $\pm$ 74	
2j $\Sigma p_T > 400$ GeV	93408	92789.5 $\pm$ 1138.2	e+3j $\tau$ -	11	7.6 $\pm$ 1.6	bj2 $\gamma$	13	17.6 $\pm$ 3.3	
2j2 $\gamma$	645	612.6 $\pm$ 18.8	e+3j $\gamma$	27	21.7 $\pm$ 3.4	bj $\tau$ +	13	12.9 $\pm$ 1.8	
2j $\tau$ + $\tau$ -	15	25 $\pm$ 3.5	e+2 $\gamma$	47	74.5 $\pm$ 5	bj $\not{p}$ $\Sigma p_T > 400$ GeV	53	60.4 $\pm$ 19.9	
2j $\not{p}$ $\Sigma p_T > 400$ GeV	74	106 $\pm$ 7.8	e+2j	126665	122457 $\pm$ 1672.6	bj $\gamma$	937	989.4 $\pm$ 20.6	
2j $\not{p}$ $\Sigma p_T < 400$ GeV	43	37.7 $\pm$ 100.2	e+2j $\tau$ -	53	37.3 $\pm$ 3.9	bj $\gamma$ $\not{p}$	34	30.5 $\pm$ 4	
2j $\gamma$	33684	33259.9 $\pm$ 397.6	e+2j $\tau$ +	20	24.7 $\pm$ 2.3	bj $\mu$ + $\not{p}$	104	112.6 $\pm$ 4.4	
2j $\gamma$ $\tau$ +	48	41.4 $\pm$ 3.4	e+2j $\not{p}$	12451	12130.1 $\pm$ 159.4	bj $\mu$ +	173	141.4 $\pm$ 4.8	
2j $\gamma$ $\not{p}$	403	425.2 $\pm$ 29.7	e+2j $\gamma$	101	88.9 $\pm$ 6.1	be+3j $\not{p}$	68	52.2 $\pm$ 2.2	
2j $\mu$ + $\not{p}$	7287	7320.5 $\pm$ 118.9	e+ $\tau$ -	609	555.9 $\pm$ 10.2	be+2j $\not{p}$	87	65 $\pm$ 3.3	
2j $\mu$ + $\gamma$ $\not{p}$	13	12.6 $\pm$ 2.7	e+ $\tau$ +	225	211.2 $\pm$ 4.7	be+ $\not{p}$	330	347.2 $\pm$ 6.9	
2j $\mu$ + $\gamma$	41	35.7 $\pm$ 6.1	e+ $\not{p}$	476424	479572 $\pm$ 5361.2	be+ $\not{p}$ $\not{p}$	211	176.6 $\pm$ 5	
2j $\mu$ + $\mu$ -	374	394.2 $\pm$ 24.8	e+ $\not{p}$ $\tau$ -	48	35 $\pm$ 2.7	be+e-j	22	34.6 $\pm$ 2.6	
2j $\mu$ +	9513	9362.3 $\pm$ 166.8	e+ $\not{p}$ $\tau$ +	20	18.7 $\pm$ 1.9	be+e-	62	55 $\pm$ 3.1	

TABLE I: A subset of the VISTA@CDF comparison between Tevatron Run II data and the Standard Model prediction, showing those final states populated with ten or more data events. Events are partitioned into exclusive final states based on standard CDF particle identification criteria. Final states are labeled in this table according to the number and types of objects present, and are ordered according to decreasing discrepancy between the total number of events expected and the total number observed in the data. (Final states that do not exhibit notable discrepancies have been listed in inverted alphabetical order, to allow the reader to quickly find a particular final state of interest.) Only statistical uncertainties on the background prediction have been included.

track are identified as tau leptons ( $\tau^\pm$ ), with additional cuts to discriminate against electrons or muons. Photons ( $\gamma$ ) exhibit showers in the electromagnetic calorimeter, with a veto on associated charged particle tracks. A tagging algorithm based on finding displaced vertices is used to identify jets likely resulting from the fragmentation of a bottom quark ( $b$ ). Jets ( $j$ ) are reconstructed using a clustering algorithm with a cone of  $\Delta R = 0.4$ , unless the event contains one or more jets with  $p_T > 200$  GeV and no leptons or photons, in which case a cone of 0.7 is used. Missing momentum ( $\not{p}$ ) is calculated as the negative vector sum of the 4-vectors of all identified objects and any energy visible in the detector but not clustered into an identified object (`unc1`). All objects are required to have  $p_T > 17$  GeV. Collision events are incorporated into the VISTA analysis if they contain at least one high- $p_T$  electron, muon, photon, jet, or  $b$ -tagged jet.

The Standard Model prediction is obtained using a variety of event generators. PYTHIA [6] is used for the generation of large samples of inclusive  $W$ ,  $Z$ ,  $\gamma\gamma$ ,  $\gamma j$ , and  $jj$  production, and smaller samples corresponding to  $WW$ ,  $WZ$ , and  $ZZ$  production. MADEVENT [7] provides events modeling  $W+n$  jets, for  $0 \leq n \leq 4$ , and  $Z+n$  jets, for  $0 \leq n \leq 3$ , which are matched [8, 9, 10] in order to avoid double-counting of events produced by the matrix element and by the showering algorithm, implemented with PYTHIA. MADEVENT is also used to generate the processes  $W\gamma$  and  $Z\gamma$ , and to generate additional statistics at large  $\sum p_T$ . HERWIG [11] is used to model top quark pair production. All generated events are passed through CDF's GEANT-based simulation of the detector response (CDFSIM [12]).

VISTA partitions data and Monte Carlo events into exclusive final states labelled according to the objects ( $e^\pm$ ,  $\mu^\pm$ ,

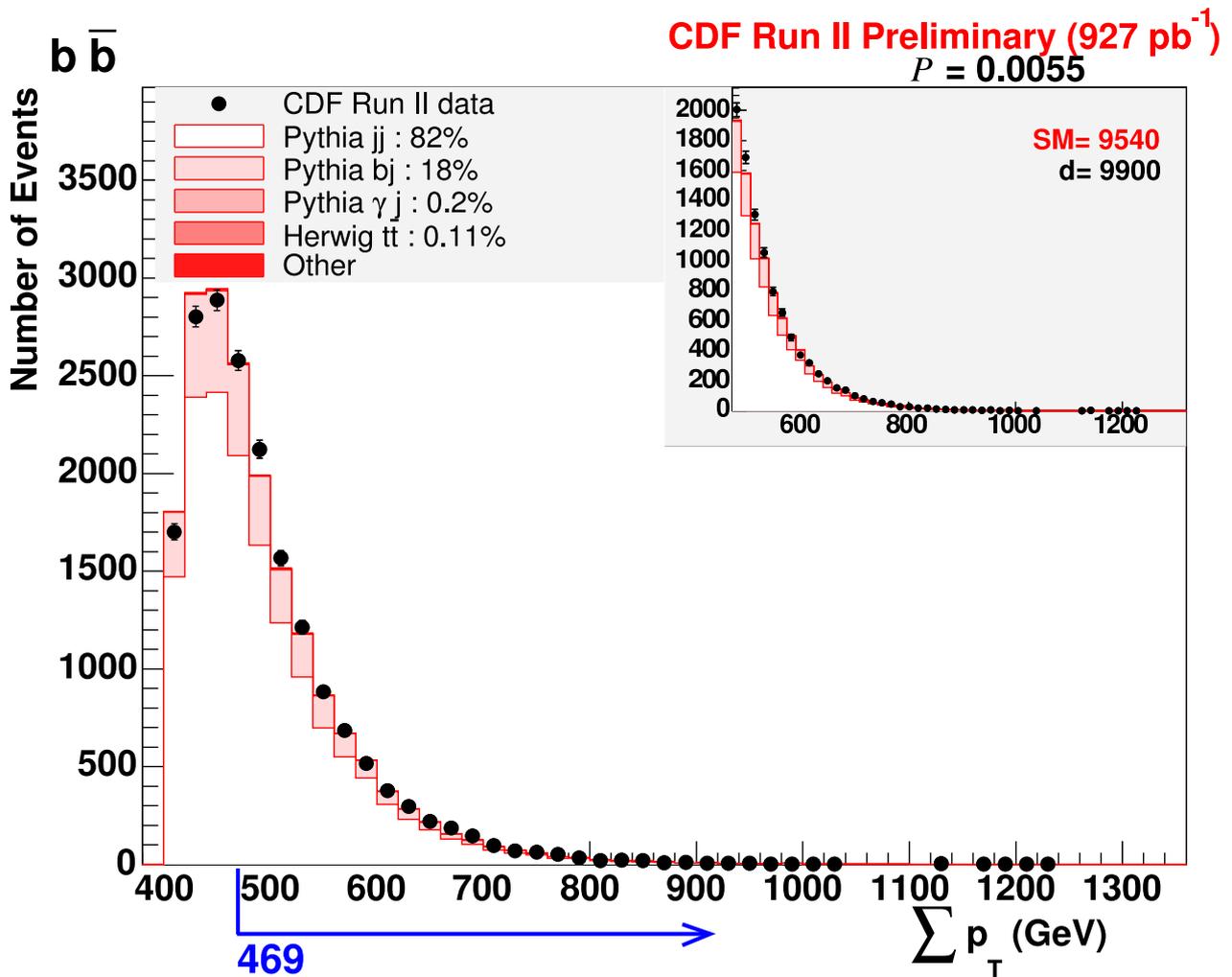


FIG. 2: The most interesting SLEUTH final state, corresponding to  $\tilde{\mathcal{P}} = 0.46$ . Filled circles (black) show CDF Run II data; the (red) histogram shows the Standard Model prediction. The horizontal axis shows the summed scalar transverse momentum of all objects in each event in the histogram, with the number of events per 20 GeV bin shown on the vertical axis. SLEUTH considers semi-infinite regions on the horizontal axis stretching from each individual data point up to infinity. For each region, SLEUTH computes the Poisson probability of the Standard Model prediction to fluctuate up to or beyond the number of events observed in the data. The region SLEUTH determines to be most interesting is shown with the (blue) arrow to be the region corresponding to  $\sum p_T > 469$  GeV. A magnified view of this region is shown in the inset, together with the number of events predicted (SM) and the number of events observed ( $d$ ). Taking into account the trials factor associated with looking at many different regions, SLEUTH determines the fraction of hypothetical similar experiments that would produce something in this final state as interesting as the region shown to be  $\mathcal{P} = 0.0055$  (number at upper right). Taking into account the further trials factor associated with looking at many SLEUTH final states, the number of hypothetical similar CDF experiments that would produce a region in any final state as interesting as the region shown is  $\tilde{\mathcal{P}} = 0.46$ .

$\tau^\pm$ ,  $\gamma$ ,  $j$ ,  $b$ ,  $\not{p}$ ) identified in each event. This partitioning is orthogonal, with every event placed into exactly one exclusive final state.

A correction model is developed to improve known or suspected systematic deficiencies in either the Standard Model prediction or the simulation of the CDF detector response. The details of this correction model are motivated by individual discrepancies noted in a global comparison of CDF data to the Standard Model prediction. The development of this correction model is the heart of VISTA.

The correction model includes specific correction factors for the integrated luminosity of the sample, the ratio ( $k$ -factor) of the all-orders calculation performed by Nature and the usually leading order calculation performed by event generators, object identification efficiencies, object fake rates, and trigger efficiencies. A total of 44 correction factors are used, of which 26 are constrained from external information. A global  $\chi^2$  is formed by comparison of CDF data

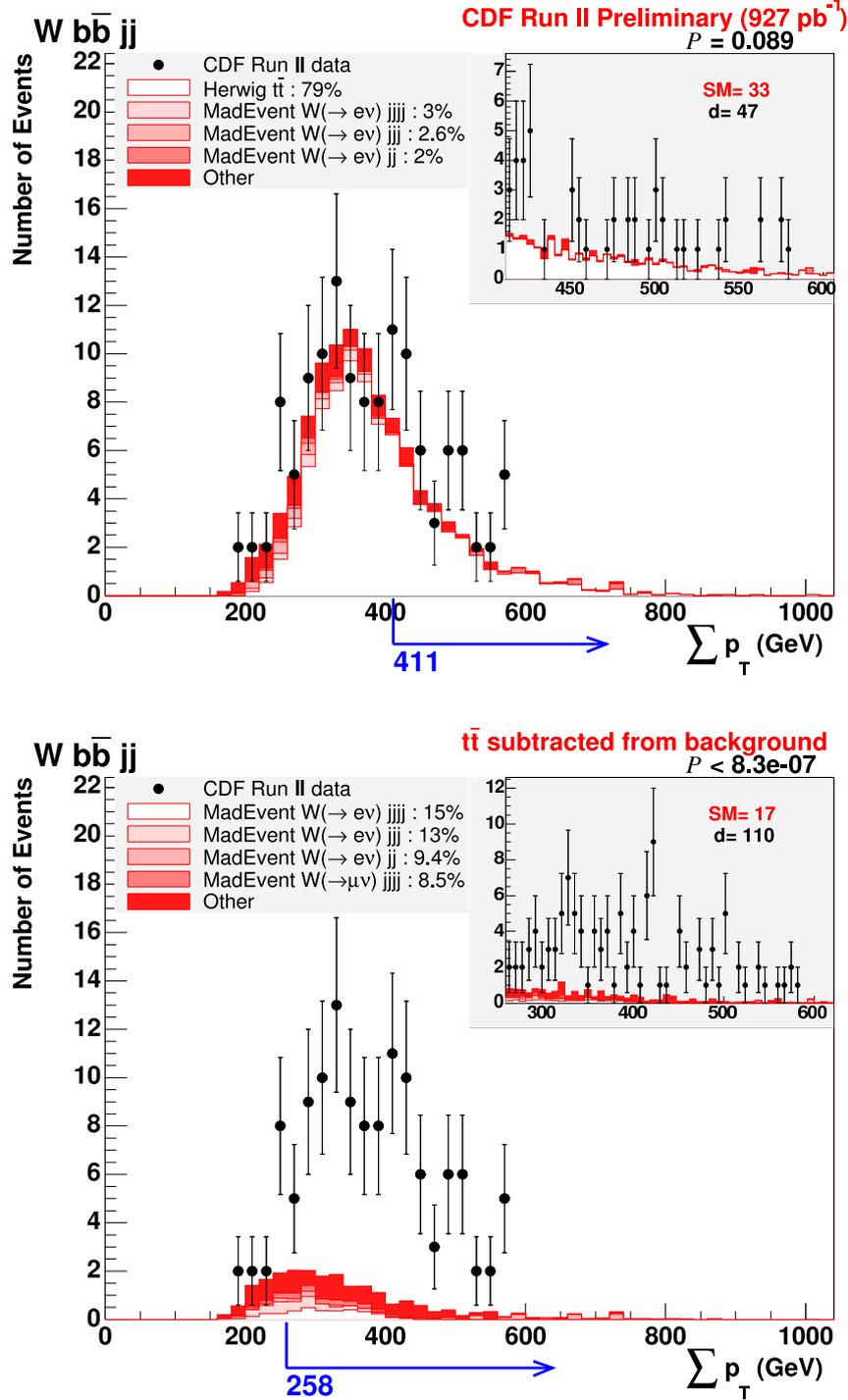


FIG. 3: (Top) The SLEUTH  $Wb\bar{b}jj$  final state, consisting of events with one electron or muon, missing transverse energy, and  $\geq 4$  jets, at least one of which is  $b$ -tagged, and corresponding to  $927 \text{ pb}^{-1}$  of integrated luminosity. The CDF Run IIa are shown as (black) filled circles, with the Standard Model prediction shown as the (red) shaded stacked histogram. SLEUTH chooses the region with  $\sum p_T > 411 \text{ GeV}$ , shown by the (blue) arrow, and displayed in the inset. In this region, the number of predicted Standard Model events is  $SM = 33$ , and the number of observed data events is  $d = 47$ . The interestingness of this most interesting region in this final state is measured by SLEUTH to be  $\mathcal{P}_{Wb\bar{b}jj} = 0.089$ . (Bottom) The same SLEUTH final state with the  $t\bar{t}$  contribution subtracted from the Standard Model prediction. SLEUTH chooses the region with  $\sum p_T > 258 \text{ GeV}$ , shown by the arrow, and displayed in the inset. In this region, the number of predicted Standard Model events is  $SM = 17$ , and the number of observed data events is  $d = 110$ . SLEUTH quantifies the fraction of hypothetical similar experiments that would have produced a region more interesting than the region chosen in this final state, and finds  $\mathcal{P}_{Wb\bar{b}jj} < 8.3 \times 10^{-7}$ , corresponding to a value of  $\tilde{\mathcal{P}}$  that surpasses SLEUTH's discovery threshold of  $\tilde{\mathcal{P}} < 0.001$ .

to the Standard Model prediction, and minimized as a function of these correction factors. This correction model is intentionally crude in the interest of transparency, recognizing the desire for simplicity in making an eventual argument for the possible presence of new physics.

A global comparison of data to the Standard Model prediction is made in a total of 16486 kinematic distributions in 344 populated exclusive final states. In each final state, the number of events observed is compared with the Standard Model prediction. In each kinematic distribution, the shape of the data is compared to the shape of the Standard Model prediction using the Kolmogorov-Smirnov statistic. A discovery claim can be made by VISTA if the statistical significance of the discrepancy corresponds to a probability  $< 0.001$  after accounting for the appropriate number of final states or distributions considered, if the discrepancy appears inexplicable as a deficiency in the modeling of the CDF detector response or in the calculation of the Standard Model prediction, and if the discrepancy has a new physics interpretation.

The number of events observed is in agreement with the Standard Model prediction; assuming the data have been drawn from our prediction, the probability of observing the most discrepant excess is 8%. The number of distributions showing a significant ( $> 5\sigma$ ) difference in shape between data and the Standard Model prediction is 409. Most of these differences are attributed to modeling the parton radiation and underlying event in the data, and none are believed to have a new physics interpretation.

### III. SLEUTH

With VISTA having seen nothing on which to base a discovery claim, SLEUTH [1, 2, 3, 4, 13, 14] is used to search for evidence of new physics on the high- $p_T$  tails. SLEUTH is a quasi-model-independent search technique, based on the assumption that new electroweak-scale physics will manifest itself as an excess of data over the Standard Model expectation in a particular final state at large summed scalar transverse momentum ( $\sum p_T$ ). The strengths and limitations of SLEUTH follow directly from these assumptions.

Object identification criteria in SLEUTH are the same as used in VISTA. SLEUTH final states are formed by merging some VISTA final states, with the goal of increasing statistical significance without compromising sensitivity to new physics: final states with electrons and muons are merged under the assumption that new physics treats the first two generations similarly; and jets and  $b$ -jets are identified as pairs, since Feynman diagrams drawn for most new physics signals have an even number of quarks in the final state.

SLEUTH considers a single variable, the summed scalar transverse momentum ( $\sum p_T$ ) of all objects in the event, with the Standard Model prediction for the distribution of  $\sum p_T$  determined using the correction factors found by VISTA. For each final state, SLEUTH determines the most interesting region on the tail of this distribution. A final state populated by  $d$  data points contains  $d$  regions, where the  $i^{\text{th}}$  region is defined as the semi-infinite interval with lower bound equal to the  $i^{\text{th}}$  largest data  $\sum p_T$ . The  $i^{\text{th}}$  region contains  $i$  events, with the Standard Model prediction obtained by integrating the predicted  $\sum p_T$  distribution over this semi-infinite region.

The interestingness  $p_N$  of a region containing  $N$  data points is defined as the Poisson probability that the integrated Standard Model prediction would fluctuate up to or above  $N$ . The most interesting region  $\mathcal{R}$  is defined as the  $N$  data points for which  $p_N$  is minimal. Pseudo experiments are performed by drawing pseudo data from the Standard Model  $\sum p_T$  distribution, and the most interesting region is found for each pseudo experiment. The fraction  $\mathcal{P}$  of these pseudo experiments producing a region more interesting than the region  $\mathcal{R}$  found in the data quantifies the interestingness of this final state. In the interest of speed, possible variations in the correction factor values determined by VISTA are not explicitly propagated into SLEUTH. Propagation is instead implicit: an interesting excess identified by SLEUTH that can be explained away in terms of a reasonable adjustment in the VISTA correction model (consistent with all available external information) leads to an adjustment in the VISTA correction model. The operation of SLEUTH is illustrated in Fig. 2.

Considering all final states, SLEUTH determines the most interesting region in the CDF high- $p_T$  data, and calculates  $\tilde{\mathcal{P}}$ , the fraction of hypothetical similar CDF experiments that would have produced a region in any final state more interesting than the most interesting  $\mathcal{R}$ . In calculating  $\tilde{\mathcal{P}}$ , SLEUTH rigorously accounts for the number of final states that have been considered. With an accurate correction model and in the absence of new physics, the distribution of  $\tilde{\mathcal{P}}$  is uniform between zero and unity; in the presence of new physics, small  $\tilde{\mathcal{P}}$  is expected. The threshold for pursuit of a possible discovery case is  $\tilde{\mathcal{P}} < 0.001$ .

In 927 pb $^{-1}$  of CDF Run IIa data, SLEUTH finds  $\tilde{\mathcal{P}} = 0.46$ . The most interesting region observed is shown in Fig. 2. The fraction of hypothetical similar CDF experiments (assuming an accurate Standard Model prediction, detector simulation, and correction model) that would observe something as interesting as the most interesting region observed in the CDF Run IIa data is 46%.

#### IV. SENSITIVITY

Figure 3 shows a sensitivity test in which the Standard Model process  $p\bar{p} \rightarrow t\bar{t}$  is subtracted from the Standard Model background and observed as an excess in the CDF data. Based on this test, we would expect SLEUTH to discover the top quark with an integrated Run II luminosity comparable to that accumulated by CDF and DØ in Tevatron Run I when the top quark discovery was announced [15, 16]. Several other sensitivity tests have been conducted with pseudo signals injected into pseudo data drawn from the Standard Model prediction. On these sensitivity tests, SLEUTH performs comparably to targeted searches for phenomena satisfying SLEUTH's basic assumptions that new physics will appear as an excess of data over the Standard Model prediction at larged summed scalar transverse momentum in one primary final state.

Targeted searches typically out-perform SLEUTH on signals with particularly distinct kinematic features. Production of a 115 GeV Standard Model Higgs boson decaying to two  $b$ -tagged jets in association with a heavy electroweak gauge boson at the Tevatron is best studied with a targeted search: SLEUTH treats the final states  $Wb\bar{b}$ ,  $\ell^+\ell^-b\bar{b}$ , and  $p\bar{b}\bar{b}$  independently, while a targeted search can join the distinct signals arising from  $p\bar{p} \rightarrow W(\rightarrow \ell\nu)h(\rightarrow b\bar{b})$ ,  $p\bar{p} \rightarrow Z(\rightarrow \ell^+\ell^-)h(\rightarrow b\bar{b})$ , and  $p\bar{p} \rightarrow Z(\rightarrow \nu\bar{\nu})h(\rightarrow b\bar{b})$ ; the 115 GeV resonance in the  $b\bar{b}$  invariant mass is a better indicator than SLEUTH's  $\sum p_T$ ; and the integral role played by the Higgs boson in the Standard Model makes this particle a uniquely well-motivated search target.

#### V. SUMMARY

Data collected in Run II of the Fermilab Tevatron have been searched for indications of new electroweak scale physics. Rather than focusing on particular new physics scenarios, CDF data are analyzed for discrepancies with respect to the Standard Model prediction. A model-independent approach (VISTA) sensitive to new large cross section physics has considered the gross features of the data. A quasi-model-independent approach (SLEUTH) has emphasized the high- $p_T$  tails, and is designed to be particularly sensitive to new electroweak scale physics. This global search for new physics in  $927 \text{ pb}^{-1}$  of  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96 \text{ TeV}$  reveals no indication of physics beyond the Standard Model. This analysis of course does not prove the absence of new physics in these data; continued Tevatron running may yet reveal a remarkable new discovery.

- 
- [1] DØ Collaboration (2001), *Phys. Rev. Lett.* **86** 3712.
  - [2] DØ Collaboration (2000), *Phys. Rev.* **D 62** 092004.
  - [3] DØ Collaboration (2001), *Phys. Rev.* **D 64** 012004.
  - [4] H1 Collaboration (2004), *Phys. Lett.* **B 602** 14.
  - [5] CDF Collaboration (2005), [hep-ex/0508029](#).
  - [6] T. Sjöstrand et al. (2001), *Comput. Phys. Commun.* **135** 238.
  - [7] F. Maltoni and T. Stelzer (2003), *J. High Energy Phys.* **0302** 027, URL <http://madgraph.physics.uiuc.edu/>.
  - [8] S. Mrenna and P. Richardson (2003), [hep-ph/0312274](#).
  - [9] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber (2001), *J. High Energy Phys.* **11** 063.
  - [10] F. Krauss (2002), *J. High Energy Phys.* **08** 015.
  - [11] G. Corcella et al. (2002), [hep-ph/0210213](#).
  - [12] E. Gerchtein and M. Paulini (2003), [physics/0306031](#).
  - [13] B. Knuteson (2004), *Nucl. Instrum. Meth.* **A534** 7, [hep-ex/0402029](#).
  - [14] B. Knuteson (2004), [hep-ex/0504041](#).
  - [15] CDF Collaboration (1995), *Phys. Rev. Lett.* **74** 2626.
  - [16] DØ Collaboration (1995), *Phys. Rev. Lett.* **74** 2632.
  - [17] The symbol  $\eta_{\text{det}}$  denotes pseudorapidity in a coordinate system in which the origin lies at the center of the CDF detector.