Submitted to

International Europhysics Conference on High Energy Physics, EPS03, July 17-23, 2003, Aachen
(Abstract **118** Parallel Session **13**)

XXI International Symposium on Lepton and Photon Interactions, LP03, August 11-16, 2003, Fermilab

www-h1.desy.de/h1/www/publications/conf/conf_list.html

General Search for New Phenomena in ep scattering at HERA

H1 Collaboration

Abstract

A model-independent search for deviations from the Standard Model prediction has been performed in e^+p and e^-p collisions at HERA using H1 data corresponding to an integrated luminosity of 115 pb^{-1}. All experimentally measurable event topologies involving electrons, photons, muons, neutrinos and jets with high transverse momenta have been investigated. Events are classified into exclusive event classes according to their final state. A new algorithm has been developed to look for regions with large deviations from the Standard Model in the invariant mass and sum of transverse momenta distributions and to quantify the significance of the fluctuations observed. A good agreement with the Standard Model prediction is found in most of the event classes. The largest deviation was found in μ-j-ν topologies where deviations have been previously observed. About 2% of hypothetical Monte Carlo experiments would produce deviations more significant than the one observed in the corresponding sum of transverse momenta distribution.

1 Introduction

At HERA electrons[1] and protons collide with a centre-of-mass energy of 319 GeV. The high centre-of-mass energy and the unique types of colliding particles provide an ideal testing ground for the Standard Model (SM). The H1 experiment at HERA has accumulated more than 100 pb^{-1} of integrated luminosity in the first period of HERA running (HERA I) at a centre-of-mass energy of 319 and 301 GeV. These data can in particular be used for the search for signals of new physics. Various tests of possible extensions of the SM have already been performed and upper limits on cross-sections of new processes have been derived. All these exotic models lead to different final state event topologies. These analyses have been optimised to detect the anticipated experimental signatures of these extensions.

A complementary approach is described in this paper consisting in a broad-range search for deviations from the SM prediction at large transverse momentum. The analysis covers phase space regions where the SM prediction is sufficiently precise to detect anomalies and does not rely on special models for physics beyond the SM. A prior related strategy for such a model-independent search was presented by the D0 collaboration [1–3].

In this work all final states are analysed having at least two objects with a transverse momentum (P_T) above 20 GeV and in the polar angle range $10° < \theta < 140°$. The considered objects are electron (e), muon (μ), photon (γ), jet (j) and neutrino (ν) (or non-interacting particles). Moreover the objects are required to be isolated versus each other by a minimum distance $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ of 1 unit in the pseudorapidity-azimuth ($\eta - \phi$) plane[2]. In order to avoid bias, the object phase space requirements have been defined a priori and no additional cuts (except topological background finding) on the event classes are applied. The object quality criteria are defined according to our knowledge of SM processes and detector performances. The complete HERA I data sample (1994-2000) has been used corresponding to an integrated luminosity of 115.3 pb^{-1}. All selected events are then classified into exclusive event classes (e.g. e-j , j-j , j-ν) according to the number and types of objects detected in the final state. Exclusive event classes ensure a good separation of final states and an unambiguous statistical interpretation.

In a first step the global yields of all measurable event classes are compared to the SM expectation. Efficiencies and purities are derived for all classes with a sizeable SM prediction. These efficiencies can further be used to derive limits on signals from processes beyond the SM.

In a second step the invariant mass M_{all} and the scalar sum of transverse momenta $\sum P_T$ of high P_T final state objects are considered. A new algorithm is presented to search for the largest deviation from the SM prediction in these distributions. Finally the likelihood to find the deviation with the algorithm in a given event class and in all studied event classes is derived.

[1]In this paper "electron" refers to both electrons and positrons, if not otherwise stated.

[2]The origin of the H1 coordinate system is the nominal ep interaction point, while the direction of the proton beam defines the positive z-axis (forward region). The transverse momenta are measured in the xy plane. The polar θ and azimuthal ϕ angles are measured with respect to the reference system. The pseudorapidity is a function of the polar angle: $\eta = -\log(\tan(\theta/2))$.

2 SM processes and their Monte Carlo generation

A search for deviations from the SM requires a precise and reliable estimate of all relevant processes. Hence several Monte Carlo generators are used to generate a large number of events in all event classes carefully avoiding double-counting of processes. All events are passed through a full detector simulation [13]. All processes are generated with a luminosity at least 20 times higher than the one of the data.

The main contributions to the SM prediction at high transverse momentum are due the photoproduction of jets and neutral current (NC) deep-inelastic scattering (DIS).

To simulate the direct and resolved photoproduction of jets ($ep \to jjX$) and prompt photon production ($ep \to \gamma jX$) the PYTHIA 6.1 event generator [14] is used. The abbreviation X stands for a not specified low P_T system of the reaction products. Light and heavy flavours are generated. The simulation contains the Born level QCD hard scattering matrix elements and radiative QED corrections.

The order α_s matrix elements are used in the RAPGAP event generator to model NC DIS events. The QED radiative effects arising from real photon emission from both incoming and outcoming lepton are simulated using the HERACLES [16] generator. Hence the NC DIS prediction contains the processes $ep \to ejX$, $ep \to ejjX$ and $ep \to e\gamma jX$. In the case the electron is undetected RAPGAP contributes to $ep \to jjX$ and $ep \to \gamma jX$.

Charged current (CC) deep-inelastic scattering events are calculated using the DJANGO [15] program, which includes first order QED corrections based on HERACLES. Parton cascades are generated using the colour-dipole model in ARIADNE [17]. This prediction contains the processes $ep \to \nu jX$, $ep \to \nu jjX$ and processes with an additional photon radiation.

Multi-lepton events are generated with the GRAPE [18] generator, which includes all exact electroweak matrix elements at tree level. The multi-lepton production via $\gamma\gamma$, γZ, ZZ collisions, internal photon conversion and via the decay of virtual or real Z bosons is considered. Initial and final state QED radiation is included. The complete hadronic final state is obtained via interfaces to PYTHIA (SOPHIA) [19] for the inelastic (quasi-elastic) regime. Consequently GRAPE predicts $ep \to \mu\mu X$ and $ep \to eeX$ as well as additional jet production.

Elastic and quasi-elastic Compton processes ($ep \to e\gamma X$) are simulated with the WABGEN [20] generator. The inelastic contribution is already included in the used NC DIS RAPGAP prediction.

The production of W bosons ($ep \to WX$ and $ep \to WjX$) is modelled using the EPVEC [21] Monte Carlo generator. QCD corrections [22] to these processes are taken into account by reweighting the events as a function of transverse momentum and rapidity of the W boson.

Processes with additional jets (e.g. $ep \to jjjX$ or $ep \to jjjjX$) are accounted for in all models using leading logarithmic parton showers as representation of higher order QCD radiation. The prediction of processes with two or more high transverse momentum jets (e.g. $ep \to jj$) was scaled by a factor of 1.2 to re-weight the normalisation of the leading order Monte Carlos to that of next-to-leading order QCD calculations [6].

3 Experimental Technique

3.1 The H1 detector

The H1 detector components relevant to the present analysis are briefly described here. The Liquid Argon (LAr) [11] and SpaCal [12] calorimeters are used to trigger events and to measure the energy of the final state. The LAr calorimeter is used to identify jets, photons and electrons and covers the polar angle range $4^\circ < \theta < 154^\circ$ with full azimuthal acceptance. The energy calibration uncertainty for jets with high transverse momentum is determined to be 2% and varies between 0.7% and 3% for electrons and photons. The central and forward tracking detectors are used to measure charged particles' trajectories, to reconstruct the interaction vertex and to supplement the measurement of the hadronic energy flow. Two concentric drift chambers (CIP and COP) complement the vertex measurement. Since the CIP is the innermost proportional chamber it is used to veto charged particles for the photon identification. The CIP angular coverage is $7^\circ < \theta < 175^\circ$. The core of the detector is enclosed by a superconducting magnetic coil with a strength of 1.15 T. The iron return yoke builds the outermost part of the detector and is equipped with streamer tubes forming the central muon detector ($4^\circ < \theta < 171^\circ$).

The luminosity determination is based on the measurement of the Bethe-Heitler process $ep \rightarrow ep\gamma$, where the electron and photon are detected in calorimeters located downstream of the interaction point.

The main trigger for events with high transverse momentum is provided by the LAr calorimeter. The trigger efficiency is close to 100% for events having an electromagnetic deposit in the LAr (electron or photon) with transverse momentum greater than 20 GeV. Events triggered only by jets have a trigger efficiency above 90% for $P_T > 20$ GeV and nearly 100% for $P_T > 25$ GeV. For events with high missing transverse momentum, determined from an imbalance in transverse momentum measured in the calorimeter P_T^{calo}, the trigger efficiency is $\sim 90\%$ for $P_T^{\mathrm{calo}} > 20$ GeV. Events only triggered by muons have a trigger efficiency above 70% [7].

3.2 Event Selection

The event sample studied consists of the full 1994-2000 HERA I data set. The data selection requires at least an isolated electromagnetic cluster, muon or jet to be found anywhere in the used detector components. To reduce several kinds of background events it is demanded that the event vertex is reconstructed within 35 cm of the nominal z position of the vertex[3] and that $\sum_i E_i - P_{z,i} < 75$ GeV, where E_i is the energy and $P_{z,i}$ is the z component of the particle momentum. The index i runs over all hadronic objects, electromagnetic clusters and muons found in the event. Due to energy-momentum conservation a typical HERA event is expected to have a value of $\sum_i E_i - P_{z,i} = 55$ GeV if the complete final state has been detected, or if only longitudinal momentum along the proton direction has been undetected. The non-ep background is rejected as described in [8] by searching for event topologies typical for cosmic

[3]This is not required for the event classes containing only photons.

ray and beam-induced background. Moreover the event timing is required to be consistent with the HERA clock.

Identification criteria for each type of particle are widely based on previous analyses performed on specific final states [4, 5, 7, 8]. Additional requirements have been chosen to ensure an unambiguous identification of particles, still keeping high efficiencies. The following paragraphs describe the object identification criteria.

Electron identification The electron identification is based on the measurement of a compact and isolated electromagnetic shower in the LAr calorimeter. The hadronic energy within 0.75 units in the $\eta - \phi$ plane around the electron is required to be below 2.5% of the electron energy. This calorimetric electron identification is complemented by tracking conditions. It is required that a high quality track geometrically matches the electromagnetic cluster within a distance of closest approach (DCA_{cl}^{tk}) to the cluster center-of-gravity of 12 cm. No other good track is allowed within 0.5 units in the $\eta - \phi$ plane around the electron direction. In the region $20^\circ < \theta < 140^\circ$ the starting radius of the measured track, defined as the distance between the first measured point in the central drift chambers and the beam axis, is required to be below 30 cm in order to reject photons that convert late in the central tracker material. In addition, the transverse momentum measured from the associated track $P_T^{e_{tk}}$ is required to match the calorimetric measurement P_T^e with $1/P_T^{e_{tk}} - 1/P_T^e < 0.02$ GeV^{-1}. Due to higher material density in the forward region, the electrons are more likely to shower and therefore some conditions depend on the polar angle of the electron candidate. In the region not fully covered by the central drift chambers ($10^\circ < \theta < 37^\circ$) a tighter calorimetric isolation cone of 1 unit in the $\eta - \phi$ plane is required to reduce the contribution of fake electrons from hadrons. The identification is then complemented by the requirement of hits in the central inner proportional chamber (CIP) within a distance $\Delta z_{\text{CIP}} < 10$ cm to the extrapolated z-impact of the electromagnetic cluster to the CIP surface. The resulting electron finding efficiency is $\approx 85\%$ in the central region and 70% in the forward region.

Photon identification The photon identification relies on the same measurement of an electromagnetic shower and the same calorimetric isolation criteria against hadrons as the electron identification. In addition no jet with a $P_T > 5$ GeV in the vicinity of the photon candidate, i.e. within a distance below 1 unit in the $\eta - \phi$ plane, should be present. Vetoes on any charged track pointing to the electromagnetic cluster are applied. No track with a DCA_{cl}^{tk} below 24 cm or within 0.5 units in the $\eta - \phi$ plane should be present. To account for possible inefficiencies of the inner tracking system an additional veto on any hits in the CIP is applied, i.e. $\Delta z_{\text{CIP}} > 10$ cm.

Muon identification The muon identification is based on a track in the forward muon system or a charged track in the inner tracking systems associated with a track segment or an energy deposit in the instrumented iron [4]. The muon momentum is measured from the track curvature in the toroidal or solenoidal magnetic field. A muon candidate should have no more than 8 GeV deposited in the LAr calorimeter in a cylinder of radius 0.5 in the $\eta - \phi$ plane, centered on the muon direction associated with its track. Cosmic background is rejected by applying several cuts. A cut on the track opening angle and polar angle sum of muon pairs is intended to discard events with muons coming from cosmic rays. The event timing condition for muon events is more restricted and the cut on the track timing difference rejects the cosmic muons residing at high values of this quantity. Beam halo events are rejected by requiring that the muons originate from the vertex. Finally misidentified hadrons are almost completely suppressed by requiring that the muon candidate is seperated from the closest jet by at least one unit in the $\eta - \phi$ plane. The efficiency to identify muons is established to be greater than 90% [4].

Jet identification Jets are defined using the inclusive $k_\perp$ algorithm as proposed in [23, 24]. It is applied in the laboratory frame with the separation parameter set to 1 and using a P_T weighted recombination scheme in which the jets are treated as massless. Energy deposits in the calorimeters and tracks in the inner tracking system are combined to reconstruct the hadronic energy of events. The jet algorithm is run on all energy deposits, not previously identified as a electron or a photon candidate. Due to inefficiencies of the electron finder, especially in detector regions where the amount of dead material is important, the scattered electron may fake or be part of a jet. This effect is important for multi-jet events, especially at high transverse momenta of the jets. To reject these fake jets, jets are required to have a radial energy moment greater than 0.02 and the quantity $M^{\text{Jet}}/P_T^{\text{Jet}}$ above 0.1. Here P_T^{Jet} denotes the transverse momentum of the jet and its invariant mass M^{Jet} is derived using the four-momenta of the objects belonging to the jet. If the fraction of the jet energy contained in the electromagnetic part of the LAr calorimeter is greater than 0.9, the above criteria are tightened to 0.04 and 0.15, respectively. About 80% of the non-genuine jets and 1% of the genuine jets are rejected by these requirements.

Neutrino identification A neutrino candidate is defined in events with a missing transverse momentum above 20 GeV. The missing momentum is derived from all identified particles and energy deposits in the event. A neutrino candidate is only assigned to the event if $\sum_i E_i - P_{z,i} < 48$ GeV. Missing transverse momentum may arise due to mismeasurement of an identified object. This effect is reduced by isolating the four-vector of the neutrino against all identified objects with a transverse momentum above 20 GeV. However an additional criteria is applied to reduce NC and lepton pair background events where one particle's energy is mismeasured. These events typically have values of $\Delta\phi(l - X_{tot})$ of $180°$. $\Delta\phi(l - X_{tot})$ is the azimuthal angle difference between the charged lepton and the direction of the system X_{tot} build of all energies measured in the calorimeters. If one electron or muon is found in the event a neutrino candidate is only assigned to the event if $\Delta\phi(l - X_{tot}) < 170°$.

The common phase space of electron, photons, muons and jets is defined by $10° < \theta < 140°$ and $P_T > 20$ GeV. The neutrino phase space is defined by the missing transverse momentum above 20 GeV and $\sum_i E_i - P_{z,i} < 48$ GeV. All particles with $P_T > 20$ GeV, including the neutrino defined by its reconstructed four-vector, are required to be isolated versus each other by a minimum distance of 1 unit in the $\eta - \phi$ plane. Based on these object definitions the events are subdivided, according to the number and types of objects, into exclusive event classes. Events with a compact isolated object in the considered phase space which is not identified as photon, electron or jet are discarded from the analysis in order to minimise wrong classifications of events.

Purities and efficiencies resulting from these identification criteria are derived as a function of the sum of transverse momenta $\sum P_T$ and as a function of the invariant mass M_{all} of the objects. Mean values are given in Table 5 for all event classes with a sizeable SM expectation. Most purities and efficiencies are found to be above 60%. The highest efficiencies are found to be above 90% for the j-j and j-j-j event classes.

3.3 Systematic Uncertainties

This section describes the sources of experimental and theoretical systematic uncertainties that are considered.

Experimental systematic uncertainties arising from the measurement of the objects are presented in Table 1. The energy scale, the polar angle θ and the azimuthal angle ϕ are varied by the specified numbers for all measured objects.

- The electromagnetic energy scale uncertainty is determined to be 1% if the z position of the electromagnetic particle's impact on the LAr calorimeter is in the backward part ($z < -145$ cm), 0.7% in the central part ($-145 < z < 20$ cm), 1.5% for $20 < z < 100$ cm and 3% in the forward part ($z > 100$ cm) [8]. The angular uncertainty of collimated electromagnetic clusters varies θ dependent between 1 and 3 mrad [8]. The identification efficiency arises from uncertainties in the Monte Carlo description of tracks and of hits in the CIP. The measured tracking efficiency is described by the simulation to a precision ranging from 2% for polar angles above 37° to 7% in the forward region. The number of electron or photon candidates with associated CIP hits or tracks is varied by the specified percent numbers.
- The hadronic energy scale of the LAr calorimeter is varied by 2%. The uncertainty on the jet polar angle determination is 5 mrad for $\theta < 30°$ and 10 mrad for $\theta > 30°$.
- The muon energy scale uncertainty amounts to 5%. The uncertainty on the polar angle determination is 3 mrad.
- The trigger uncertainties are taken into account according to the object with the highest trigger efficiency.
- The uncertainty in the integrated luminosity results in an overall normalisation error of 1.5%.

Object	Energy Scale unc.	θ unc. (mrad)	ϕ unc. (mrad)	Identification efficiency unc.	Trigger efficiency unc.
Jet	2%	5-10	–	–	3%
Electron	0.7-3%	1-3	1	2-7%(Tracking)+0.5%(CIP)	–
Photon	0.7-3%	1-3	1	2-7%(Tracking)+0.5%(CIP)	–
Muon	5%	3	1	5%	5%

Table 1: The uncertainties attributed to the particle measurements.

Depending on the dominant production process different theoretical uncertainties are used. They are listed in Table 2. An additional theoretical uncertainty of 20% is applied for each jet dominantly produced by parton shower processes (e.g. j-j-j event class). A model uncertainty of 50% is added to NC DIS events with missing transverse momentum above 20 GeV and a high P_T electron. This uncertainty is estimated by a comparison of low P_T NC DIS events with the SM prediction.

All systematic errors are added in quadrature. The resulting total uncertainty ranged, e.g. for the e-j event class between 10% and 35% and for the j-j event class between 20% and 60% increasing with P_T.

Process(es)	Uncertainty
$ep \rightarrow jjX$ and $ep \rightarrow j\gamma X$	15%
$ep \rightarrow j\nu X$ and $ep \rightarrow jeX$	10%
$ep \rightarrow jj\nu X$ and $ep \rightarrow jjeX$	15%
$ep \rightarrow \mu\mu$ and $ep \rightarrow ee$	3%
$ep \rightarrow WX$ and $ep \rightarrow WjX$	15%
$ep \rightarrow e\gamma X$ and $ep \rightarrow e\gamma j$	10%
$ep \rightarrow e\gamma p$	5%

Table 2: The uncertainties attributed to the different processes of the SM expectation.

4 Results

All possible event classes have been investigated. Among them only those experimentally measurable and having observed data events or a SM expectation greater than 0.1 events will be considered in the following discussion. The $\mu - \nu$ event class was found to be overwhelmed by background from low P_T photoproduction and was unmeasurable. It was discarded from the analysis procedure.

The event yields subdivided in event classes are presented for the data and SM expectation in Figure 1. A good overall agreement between data and SM expectation is observed for most of the event classes. The predominant processes at HERA, i.e. photoproduction, NC and CC DIS processes, can be found in the j-j , e-j and j-ν event class, respectively.

The data event yields in these event classes are in good agreement with the SM expectation. Likewise the e-j-j , e-j-j-j , j-j-j and j-j-ν event classes correspond to the same dominant processes with the inclusion of additional jet production due to higher order QCD processes. The event yields of these event classes are also well described by the SM prediction. Event classes containing a radiative photon are e-j-γ , j-j-γ and ν-γ , and correspond, respectively, to NC DIS, photoproduction and CC DIS processes with the radiation of a photon. These event classes are found to agree with the expectation. The e-γ event class is dominated by QED Compton scattering processes (94%) and again a good agreement is observed. No radiative CC DIS event is observed in the j-ν-γ event class for 1.0 ± 0.2 expected. The j-γ event class is well described by the prediction, but, due to the still high NC DIS background in this event class, the purity is low ($20 - 40\%$).

A discrepancy between data and SM expectation is observed in the μ-j-ν event class where 4 events are observed for an expectation of 0.7 ± 0.2. This event class corresponds to typical event topologies arising from W production. The deviation was already investigated in Ref. [4] and will be further discussed in Section 5. Similarly, the event classes e-ν , e-j-ν are also populated by events arising from W production. In the e-ν event class a slight deficit of 8 data events compared to an expectation of 19.9 ± 8.0 is observed. This event class is dominated by background events from NC DIS, where possible fluctuations in the hadronic energy measurement or limited detector acceptance can produce missing transverse momentum. In the e-j-ν event class 2 data events are observed for an expectation of 0.9 ± 0.2. Most of the interesting e-j-ν events mentioned in [4] have an electron with a transverse momentum below 20 GeV and are therefore not selected in the present analysis.

Another discrepancy to the SM expectation has been reported by H1 in multi-electron events from lepton pair production [5]. In the present analysis the e-e event class is populated at about 85% by electron pair production with 8 measured data events for an expectation of 10.7 ± 1.1. All di-electron events mentioned in the dedicated multi-electron search and available in the phase space of this analysis are selected; no tri-electron event is identified due to the requirement of high transverse momentum. In the region $M_{\text{all}} > 100$ GeV 3 events are observed and 1.15 ± 0.25 are expected. The higher SM prediction compared to Ref. [5] is due to background coming from fake electrons with $\theta < 20°$ and a higher $ep \rightarrow eeX$ selection efficiency due to the increased phase space.

The event classes e-μ and μ-μ are dominated by muon pair production ($\approx 95\%$ and 100%, respectively). The e-μ event class is populated when the scattered electron and only one of the muons is selected. In the e-μ and μ-μ event classes 4 and 5 events are observed compared to an expectation of 4.9 ± 0.6 and 2.6 ± 0.6, respectively. Muon pair production processes also contribute to $\approx 85\%$ in the μ-j event class, where again a good agreement is found.

Some discrepancies on the total event yields can be observed in the j-j-j-j and e-j-j-j event classes between data and SM expectation. For the e-j-j-j event class a low SM expectation of ≈ 0.05 is estimated. Events with four high P_T jets are investigated for the first time at HERA. Since these spectacular events can — in the current Monte Carlo programs — only be produced via parton shower, it can not be ensured that the prediction is reliable. No events are found in all other event classes in good agreement with the SM expectation. The expectation of ≈ 1 event in the γ-γ event class is dominated by the $ep \rightarrow e\gamma X$ process, where the electron is misidentified because of tracking inefficiency. Contributions of higher order QED processes, which could lead to two high transverse momentum (P_T) photons, are negligible.

5 Search for deviations from the Standard Model

5.1 Search algorithm

In order to quantitatively determine the level of agreement between the measured data and the SM expectation and to identify regions of possible deviations, a new search algorithm has been developed. The calculation of the global significance per event class was inspired by Ref. [1].

Quantities sensitive to new physics signals and easy to measure are the sum of transverse momenta and the invariant mass of all objects. Detailed studies have shown that both quantities have a large finding potential by mixing various signals of new physics in data and Monte Carlo distributions. The algorithm described in the following was run on these pseudo data samples and was successful in finding the signals. Hence the invariant mass M_{all} and the scalar sum of transverse momenta $\sum P_T$ of all particles have been investigated, considering that signals of new physics are likely to manifest themselves at certain transverse momentum or invariant mass. It turned out that both quantities complement each other in the search for these signals.

The basic design of the search algorithm is very simple:

Definition of regions All possible connected regions in a given distribution (a histogram) which have at least the size of twice the resolution are considered. The data N_{obs} and SM expectation $N_b + \delta N_b$ of a region are defined as the sum of all entries found in the bins of the histogram which is to be tested.

Estimation of the probability for each region A statistical estimator p is defined to judge which region is of most interest. This estimator is derived from the convolution of the Poisson probability density function (pdf) to account for statistical errors with a Gaussian pdf, $G(b; N_b, \delta N_b)$, to include the effect of non negligible systematics uncertainties.

$$p = \begin{cases} A \int\limits_0^\infty db\, G(b; N_b, \delta N_b) \sum\limits_{i=N_{\text{obs}}}^{\infty} \frac{e^{-b} b^i}{i!} & \text{if } N_{\text{obs}} \geq N_b \\ A \int\limits_0^\infty db\, G(b; N_b, \delta N_b) \sum\limits_{i=0}^{N_{\text{obs}}} \frac{e^{-b} b^i}{i!} & \text{if } N_{\text{obs}} < N_b \end{cases}$$

Here N_b denotes the mean of the Gaussian pdf and δN_b the corresponding width. A is a normalisation factor to ensure that the pdf is normalised to 1. Hence p gives an estimate of the probability that the SM expectation fluctuates upwards or downwards to the data.

Determination of the most interesting region A possible sign of new physics is found (in our ansatz) if the expectation significantly disagrees with the data. This disagreement is quantified with the estimator p. The region of greatest interest (of greatest deviation) is the region having the smallest p-value, p_{min}. This method finds narrow resonances, single outstanding events as well as signals spread over large regions of phase space.

Global significance per event class The fact that somewhere in the studied distribution a fluctuation with a value p_{min} occurs is taken into account. $\hat{P}$ is defined as the probability to observe a deviation with a p-value p_{min} with this algorithm in the investigated distribution (similar to: at any position in the distribution). $\hat{P}$ is a measure of the significance of the found deviation. To determine $\hat{P}$ hypothetical data histograms are produced by dicing in each bin a random event number according to the pdfs of the expectation (again a convolution of Poisson and Gaussian pdfs). For each hypothetical data histogram the algorithm is run to find the region of greatest deviation and to calculate p_{min}. The probability $\hat{P}$ can then be defined as the fraction of hypothetical data histograms with a p_{min}-value smaller than $p_{\mathrm{min,\,data}}$. This fraction, $\hat{P}$, can be used to combine results of different event classes, if the event classes are independent. An event class with small $\hat{P}$ is of more interest than an event class with large $\hat{P}$. Consequently the event class of most interest for a search is the one with the smallest $\hat{P}$ value.

To compare the obtained $\hat{P}$ values with an expectation, the data distributions are replaced by distributions from Monte Carlo experiments. These Monte Carlo (MC) distributions are again hypothetical data distributions. The complete algorithm is applied on these independent sets of MC experiments. In the case that deviations from the SM arise only from statistical or systematical fluctuations the distribution of $\hat{P}$ values obtained in data events are compatible with the distribution of $\hat{P}$ arising from these MC experiments.

5.2 Search Results

The final $\hat{P}$ values obtained for all event classes are summarised in Table 3 (M_{all}) and Table 4 ($\sum P_T$). The $\hat{P}$ values of the event classes with no data event and a SM expectation $\lesssim 1$ are 1. Distributions of the invariant mass M_{all} and the sum of transverse momenta $\sum P_T$ together with the regions selected by the algorithm are presented in Figure 4-7. The j-j-j-j and e-j-j-j-j event classes are not considered in this statistical analysis as their SM prediction is less reliable. The values are compared to the distribution of $\hat{P}$ values obtained in MC experiments in Figure 2 for the invariant mass distributions and in Figure 3 for the $\sum P_T$ distributions. The negative decade logarithm of the $\hat{P}$ values, $-\log \hat{P}$ is presented. Most $\hat{P}$ values range from 0.01 to 0.99, corresponding to event classes where no significant discrepancy between data and the SM expectation is observed. These results are in agreement with the expectation from MC experiments.

The largest deviation of the analysis is located in the μ-j-ν event class, where $\hat{P}$ values of 0.010 and 0.0008 are found corresponding to the high invariant mass and high $\sum P_T$ region, respectively. The mass region contains two data events for an expectation of 0.05 ± 0.02. In the chosen $\sum P_T$ region three data events are found while only 0.07 ± 0.03 are expected. This discrepancy was studied in Ref. [4].

This analysis studies a large number of event classes. Thus there is some chance that small $\hat{P}$ values may occur. The likeliness can be calculated that the smallest probability found in the investigated M_{all} or $\sum P_T$ distributions may occur. This is the fraction of MC experiments with a smaller $\hat{P}$ value than the smallest found in the data $\hat{P}_{\mathrm{data}}$. This value is found to be ≈ 0.25 for the set of M_{all} distributions and about 0.02 for the $\sum P_T$ distributions.

6 Conclusions

The data collected with the H1 experiment during the years 1994-2000 (HERA I) has been searched for deviations from the SM prediction at high transverse momentum. All possible event topologies have been investigated in a coherent and model independent way. Many event classes are analysed herein for the first time at HERA. A good agreement between data and SM expectation has been found in most event classes. The invariant mass and sum of transverse momenta distributions of the event classes have been systematically searched for deviations with a novel algorithm. The most significant deviation is found in the μ-j-ν event class, a topology where deviations have also been previously observed. About 2% of hypothetical Monte Carlo experiments would produce deviations more significant than the one observed in the corresponding sum of transverse momenta distribution.

Acknowledgements

We are grateful to the HERA machine group whose outstanding efforts have made this experiment possible. We thank the engineers and technicians for their work in constructing and now maintaining the H1 detector, our funding agencies for financial support, the DESY technical staff for continual assistance and the DESY directorate for support and for the hospitality which they extend to the non DESY members of the collaboration. We wish to thank B. Knuteson for useful discussions.

References

[1] B. Abbott *et al.* [D0 Collaboration], Phys. Rev. D **62** (2000) 092004 [arXiv:hep-ex/0006011].

[2] B. Abbott *et al.* [D0 Collaboration], Phys. Rev. Lett. **86** (2001) 3712 [arXiv:hep-ex/0011071].

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

[4] V. Andreev *et al.* [H1 Collaboration], Phys. Lett. B **561** (2003) 241 [arXiv:hep-ex/0301030].

[5] C. Adloff *et al.* [H1 Collaboration], Multi-Electron Production at High Transverse Momenta in ep Collisions at HERA, hep-ex/0307015.

[6] C. Adloff *et al.* [H1 Collaboration], Eur. Phys. J. C **25** (2002) 13 [arXiv:hep-ex/0201006].

[7] H1 Collaboration, contributed paper 1021 to XXXIst ICHEP, July 2002, Amsterdam.

[8] C. Adloff *et al.* [H1 Collaboration], arXiv:hep-ex/0304003.

[9] I. Abt *et al.* [H1 Collaboration], Nucl. Instrum. Meth. A **386** (1997) 310.

[10] I. Abt *et al.* [H1 Collaboration], Nucl. Instrum. Meth. A **386** (1997) 348.

[11] B. Andrieu *et al.* [H1 Calorimeter Group Collaboration], Nucl. Instrum. Meth. A **336** (1993) 460.

[12] R. D. Appuhn *et al.* [H1 SPACAL Group Collaboration], Nucl. Instrum. Meth. A **386** (1997) 397.

[13] R. Brun, F. Bruyant, M. Maire, A. C. McPherson and P. Zanarini, CERN-DD/EE/84-1.

[14] T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, Comput. Phys. Commun. **135** (2001) 238 [arXiv:hep-ph/0010017].

[15] G. A. Schuler and H. Spiesberger, "Django: The Interface For The Event Generators Heracles And Lepto,"

[16] A. Kwiatkowski, H. Spiesberger and H. J. Mohring, Comput. Phys. Commun. **69** (1992) 155.

[17] L. Lonnblad, Comput. Phys. Commun. **71** (1992) 15.

[18] T. Abe, Comput. Phys. Commun. **136** (2001) 126 [arXiv:hep-ph/0012029].

[19] A. Mucke, R. Engel, J. P. Rachen, R. J. Protheroe and T. Stanev, Comput. Phys. Commun. **124** (2000) 290 [arXiv:astro-ph/9903478].

[20] C. Berger and P. Kandel, *Prepared for Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting), Hamburg, Germany, 27-30 Apr 1998*

[21] U. Baur, J. A. Vermaseren and D. Zeppenfeld, Nucl. Phys. B **375** (1992) 3.

[22] K. P. Diener, C. Schwanenberger and M. Spira, Eur. Phys. J. C **25** (2002) 405 [arXiv:hep-ph/0203269].

[23] S. D. Ellis and D. E. Soper, Phys. Rev. D **48** (1993) 3160 [arXiv:hep-ph/9305266].

[24] S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, Nucl. Phys. B **406** (1993) 187.

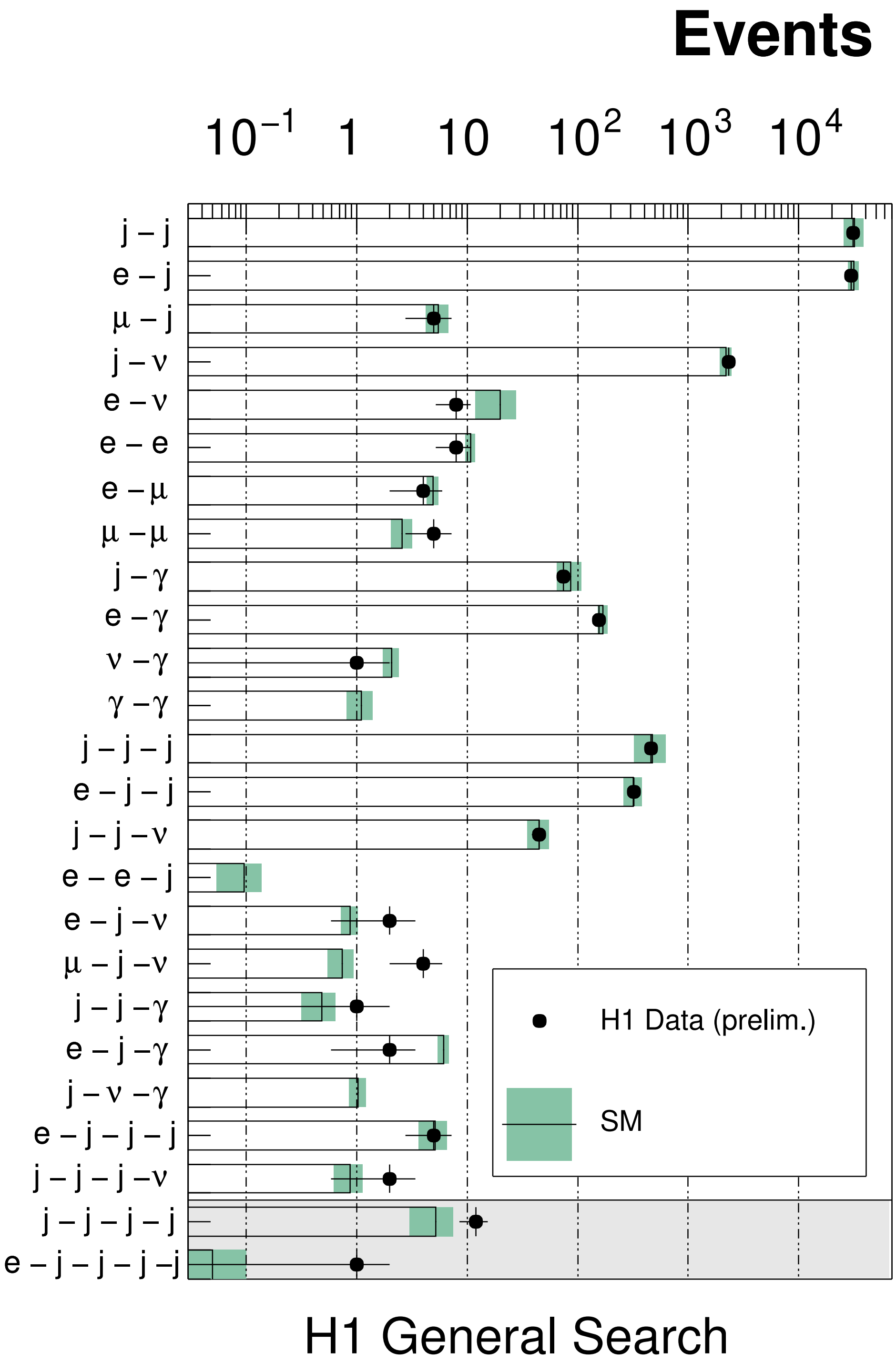

Figure 1: The data and SM expectation for all event classes with data events or with a SM expectation greater than 0.1 events. The j-j-j-j and e-j-j-j-j event classes (grey area) is less reliable and not passed through the statistical analysis.

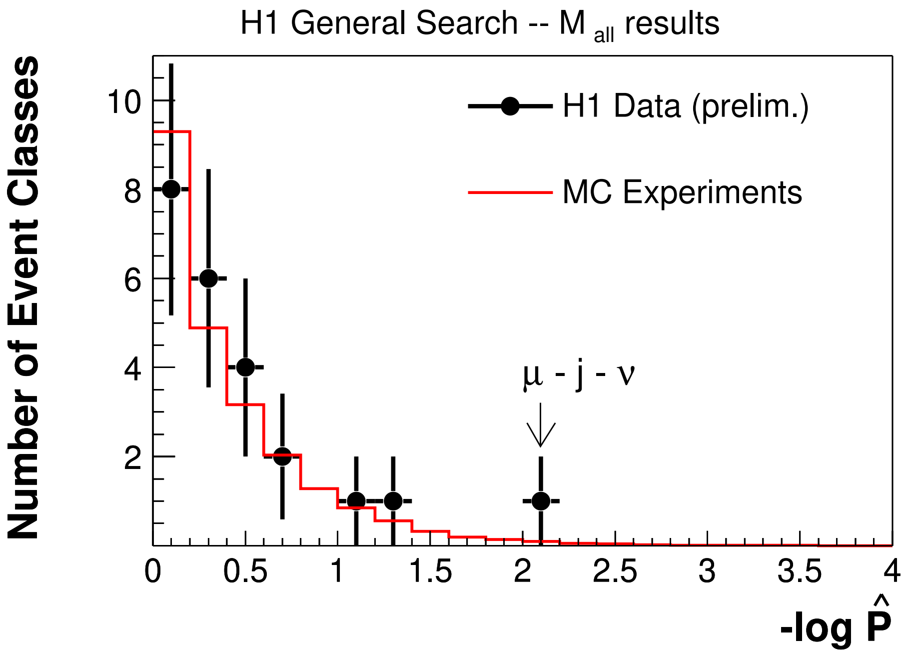

Figure 2: The $-\log \hat{P}$ values for the data event classes and the expected distribution from MC experiments. The M_{all} distributions are tested with the search algorithm. All event classes with a SM expectation greater than 0.1 events, except the j-j-j-j and the e-j-j-j-j event class, are presented.

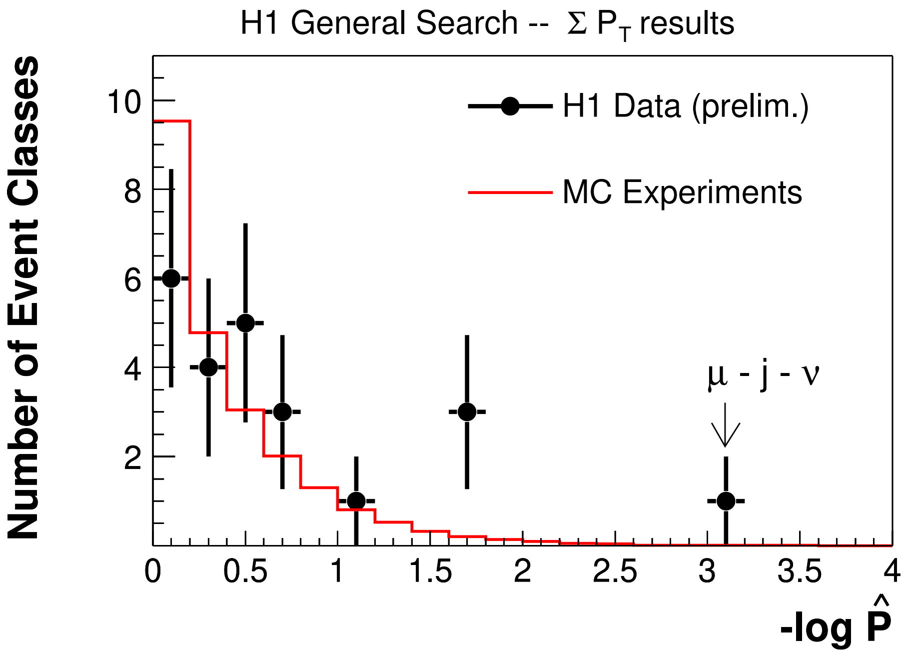

Figure 3: The $-\log \hat{P}$ values for the data event classes and the expected distribution from MC experiments. The $\sum P_T$ distributions are tested with the search algorithm. All event classes with a SM expectation greater than 0.1 events, except the j-j-j-j and the e-j-j-j-j event class, are presented.

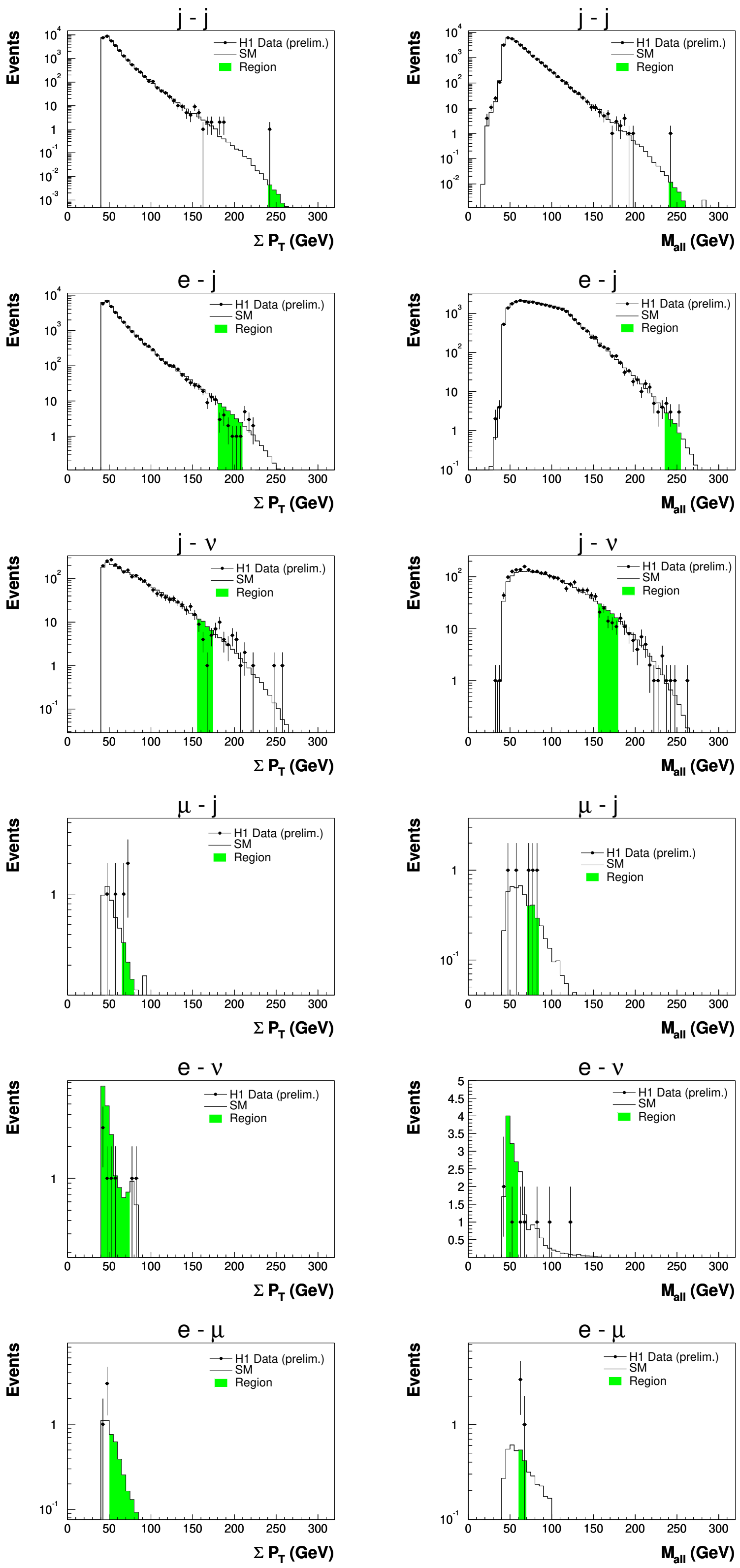

Figure 4: The number of data events and the SM expectation for various event classes as a function of $\sum P_T$ and of $M_{\rm all}$. The shaded region shows the regions of greatest deviation chosen by the search algorithm.

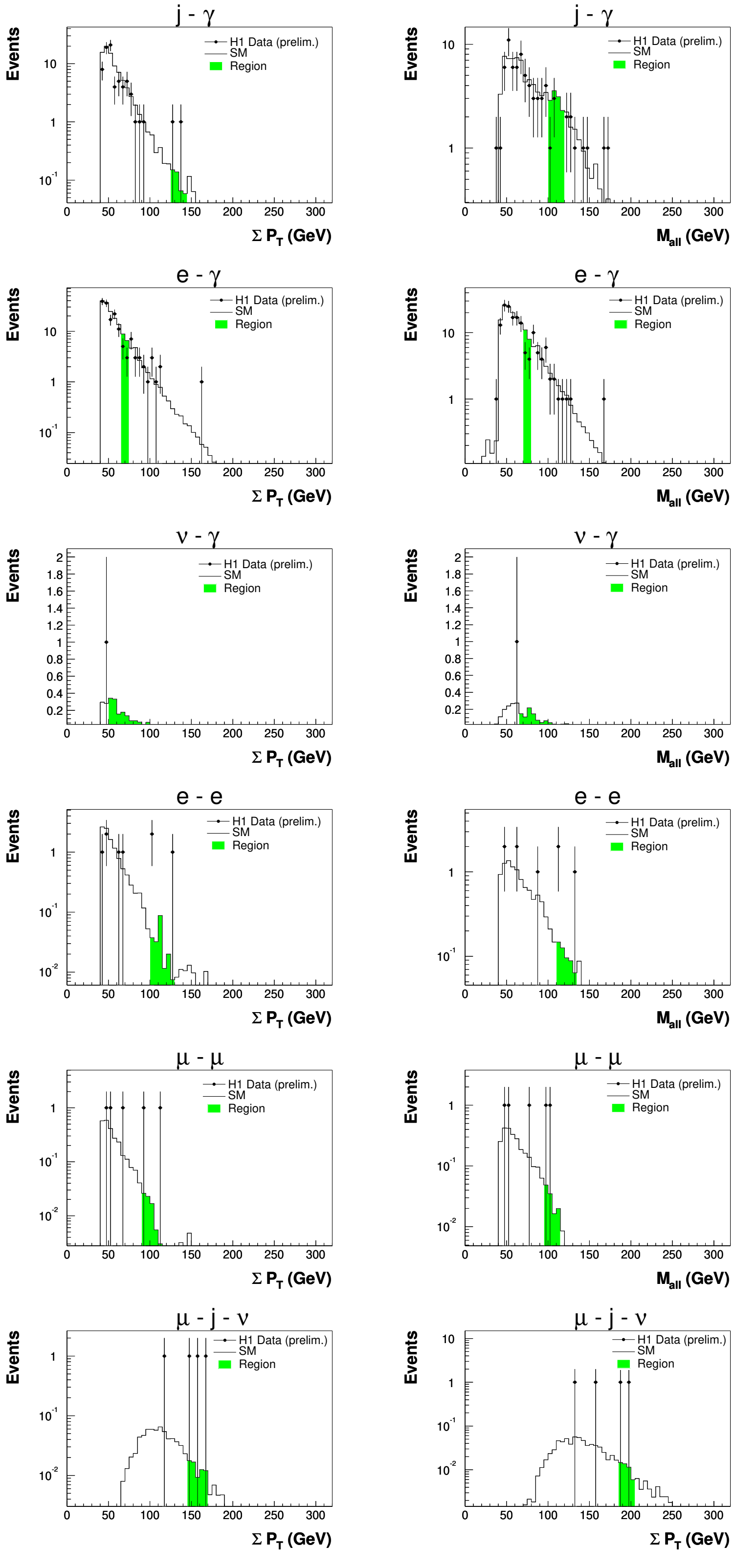

Figure 5: The number of data events and the SM expectation for various event classes as a function of $\sum P_T$ and of $M_{\rm all}$. The shaded region shows the regions of greatest deviation chosen by the search algorithm.

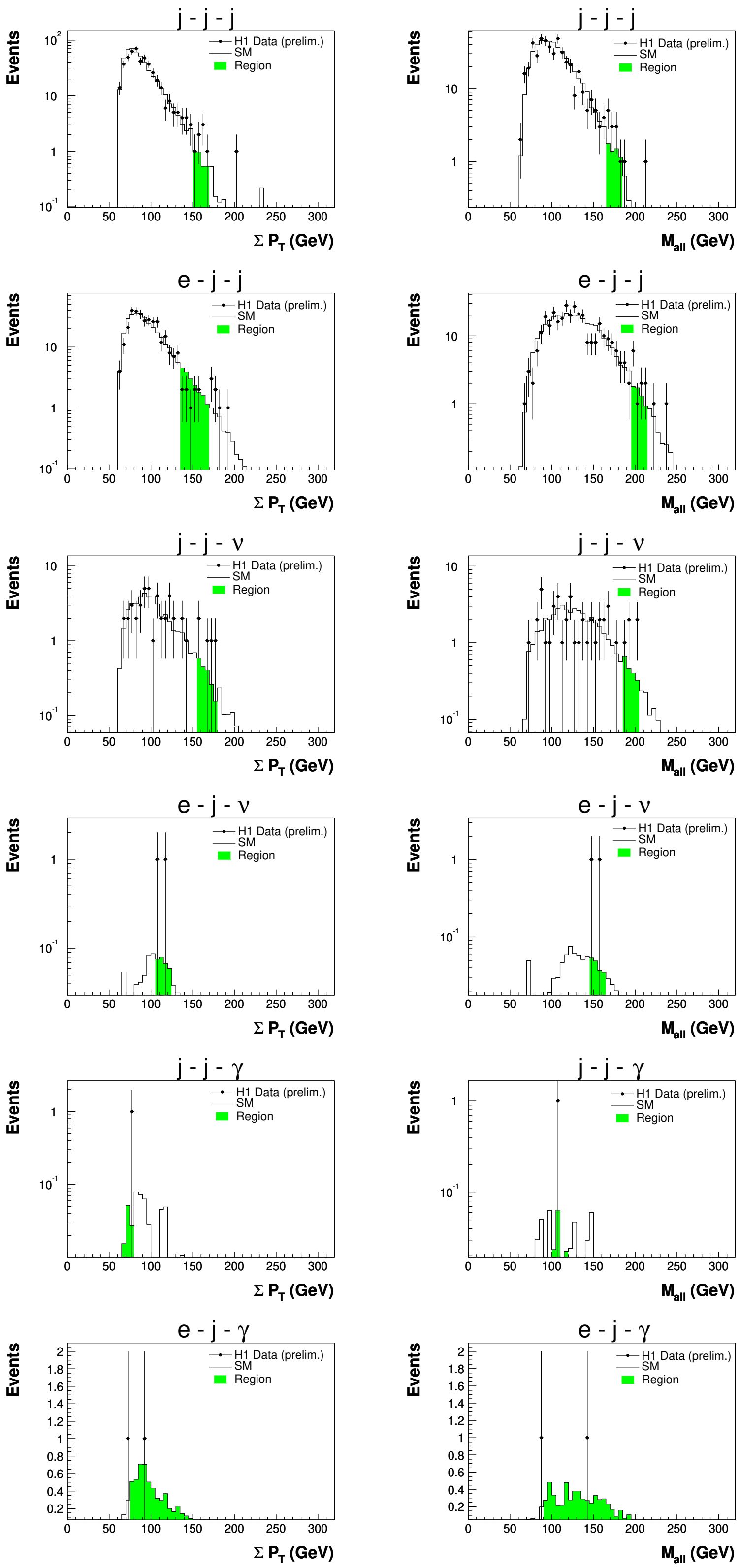

Figure 6: The number of data events and the SM expectation for various event classes as a function of $\sum P_T$ and of M_{all}. The shaded region shows the regions of greatest deviation chosen by the search algorithm.

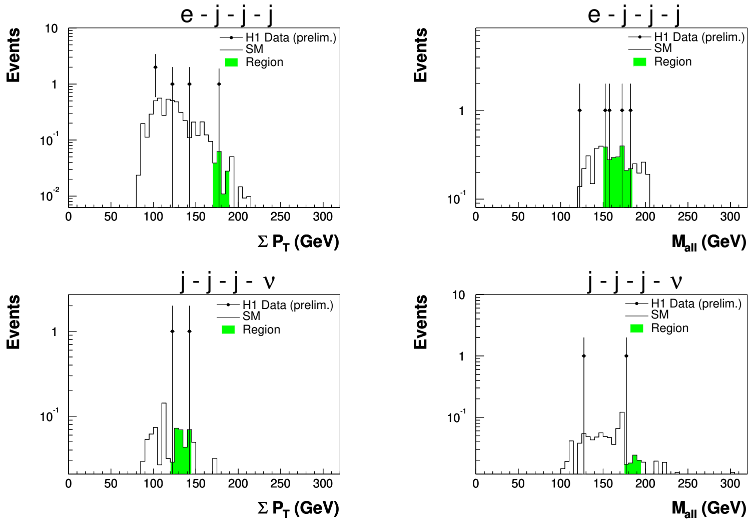

Figure 7: The number of data events and the SM expectation for various event classes as a function of $\sum P_T$ and of M_{all}. The shaded region shows the regions of greatest deviation chosen by the search algorithm.

Event class	$\hat{P}$	N_{obs}	N_{exp}	$\pm$	δN_{exp}	p
j-j	0.25	1	0.025	$\pm$	0.014	0.026
e-j	0.97	11	7.2	$\pm$	1.7	0.15
μ-j	0.68	3	1.10	$\pm$	0.27	0.11
j-ν	0.51	84	114.7	$\pm$	14.3	0.041
e-ν	0.39	1	9.9	$\pm$	5.2	0.044
e-e	0.28	3	0.52	$\pm$	0.11	0.017
e-μ	0.20	4	0.95	$\pm$	0.12	0.017
μ-μ	0.05	2	0.12	$\pm$	0.04	0.007
j-γ	0.62	4	11.9	$\pm$	3.8	0.062
e-γ	0.41	9	19.0	$\pm$	2.0	0.015
ν-γ	1.	0	0.96	$\pm$	0.37	0.406
j-j-j	0.38	12	5.8	$\pm$	2.0	0.047
e-j-j	0.57	11	5.7	$\pm$	1.3	0.053
j-j-ν	0.66	5	1.86	$\pm$	0.45	0.050
e-j-ν	0.10	2	0.18	$\pm$	0.04	0.014
μ-j-ν	0.01	2	0.046	$\pm$	0.02	0.0012
j-j-γ	0.39	1	0.12	$\pm$	0.07	0.112
e-j-γ	0.43	1	5.76	$\pm$	1.6	0.049
e-j-j-j	0.91	4	2.10	$\pm$	0.9	0.19
j-j-j-ν	0.40	1	0.08	$\pm$	0.07	0.091

Table 3: The $\hat{P}$ values, data events N_{obs} and the SM expectation N_{exp} of the region derived by the search algorithm using the M_{all} distributions for various event classes.

Event class	$\hat{P}$	N_{obs}	N_{exp}	$\pm$	δN_{exp}	p
j-j	0.065	1	0.010	$\pm$	0.005	0.010
e-j	0.02	12	30.4	$\pm$	5.0	0.0032
μ-j	0.29	3	0.70	$\pm$	0.20	0.038
j-ν	0.17	19	36.6	$\pm$	6.4	0.020
e-ν	0.38	6	18.0	$\pm$	7.4	0.073
e-e	0.02	3	0.20	$\pm$	0.08	0.0015
e-μ	0.51	0	2.70	$\pm$	0.4	0.07
μ-μ	0.02	2	0.074	$\pm$	0.03	0.0031
j-γ	0.52	2	0.41	$\pm$	0.2	0.071
e-γ	0.76	8	15.3	$\pm$	2.3	0.056
ν-γ	0.76	0	1.51	$\pm$	0.53	0.252
j-j-j	0.35	7	3.01	$\pm$	0.96	0.055
e-j-j	0.49	9	18.4	$\pm$	3.4	0.040
j-j-ν	0.56	5	1.86	$\pm$	0.55	0.054
e-j-ν	0.16	2	0.28	$\pm$	0.07	0.035
μ-j-ν	0.0008	3	0.07	$\pm$	0.03	0.00007
j-j-γ	0.31	1	0.10	$\pm$	0.07	0.104
e-j-γ	0.37	1	5.64	$\pm$	1.50	0.050
e-j-j-j	0.75	1	0.14	$\pm$	0.08	0.135
j-j-j-ν	0.22	2	0.28	$\pm$	0.20	0.048

Table 4: The $\hat{P}$ values, data events N_{obs} and the SM expectation N_{exp} of the region derived by the search algorithm using the $\sum P_T$ distributions of various event classes.

Event class	Purity $\mathcal{P}$ $(\%)$	Efficiency $\mathcal{E}$ $(\%)$
j-j	80-100	80-100
e-j	90-100	70
μ-j	80-90	50-55
j-ν	85-95	75-90
e-ν	50-80	40-50
e-e	30-70	40-50
e-μ	90	40-50
μ-μ	95-100	25-30
j-γ	20-40	30-50
e-γ	70	50-60
ν-γ	70-80	20-40
j-j-j	70-80	80-95
e-j-j	70-90	60
j-j-ν	60-80	60-85
e-e-j	40-70	20
e-e-e	30-70	20-70
j-μ-μ	50-90	20-50
e-μ-μ	60-70	20-40
e-μ-j	30-50	20-40
e-j-ν	50-70	40-50
μ-j-ν	65	40-50
j-j-γ	5-25	10-20
e-j-γ	40-60	30-45
j-ν-γ	70-100	30-50
j-j-j-j	70-90	60-80
e-j-j-j	50-80	30-70
j-j-j-ν	30-90	30-80
e-j-j-j-j	100	20

Table 5: The mean values of the efficiencies $\mathcal{E}$ and purities $\mathcal{P}$ in the M_{all} and $\sum P_T$ distributions for some event classes.