A LIKELIHOOD METHOD FOR SEPARATING SUSY EVENTS FROM BACKGROUND

T.Kamon and K.Kondo
Institute of Physics, University of Tsukuba, Japan

SUMMARY

A likelihood method is examined for separating light gluino (200 GeV/c^2) pair production from QCD background at SSC energy.

1 INTRODUCTION

Searches for supersymmetric particles (SUSY) at the SSC have been studied by Aronson et.al. in 1982 Snowmass Workshop [1]. At the SPS collider the experimental searches and analyses in the low mass range (< 100 GeV/c^2) have been continued [2]. At Tevatron collider the search for SUSY's in the intermediate mass range (~ a few hundred GeV/c^2) will start in the near future. SUSY searches have been made also at e^+e^- colliders. Relevant SUSY species to be searched for with each type of colliders are different; At e^+e^- colliders sleptons and photino are suitable to search, while the most favorable SUSY's to be studied in hadron colliders are squarks and gluino.

A general problem with a new particle search in hadron colliders is how to identify the wanted particle in a huge background . The lack of information about the initial momenta in the parton-parton collision results in a loss of a useful kinematical constraint and complicates the event reconstruction. Nevertheless, there are signatures characteristic to individual processes. A typical signature of gluinos or squarks is one or more jets with large transverse momenta accompanied by a large missing transverse momentum. The large missing momentum comes from the fact that the quantum number R should be conserved in productions and decays involving SUSY's, so that the lightest SUSY, generally assumed to be photino, should carry away the momentum without interacting with ordinary matter of detectors. In case of gluino pair production at the SSC, the process is

$$pp \rightarrow \tilde{g}\tilde{g} + \text{anything},$$

followed by decay of each gluino a la $\tilde{g} \rightarrow q\tilde{q} + \tilde{\gamma}$.

In Ref.1, the authors pointed out that the kinematical parameters representing missing transverse momentum ($E_{T,miss}$, x_E, and pout) are efficient to reject the QCD backgrounds. In addition, to reject background coming from heavy quark pair production, the identification of leptons which are expected to be produced by the semileptonic decay of heavy quarks has been suggested to be efficient to suppress the background.

It is true that these parameters and the condition are efficient in separating the signal from the background events. But in principle there are much more informations available in the process, simply because the number of particles involved is huge. The present work was motivated by questions: Are there any other variables which are useful in identifying SUSY production in hadron colliders ? Is there any general method to find out kinematical variables or parameters relevant for telling signals from background for events with (a) jet(s), (a) lepton(s) and missing momentum ? How can we integrate all useful informations to identify the SUSY events ? Works with similar motivation were reported in past Snowmass Workshops [3]. A likelihood method to be discussed in the present report is along this line and could be useful to various reactions. Here we choose an examle of gluino search assuming the high transverse momentum QCD process associated with finite calorimeter resolution as a main background.

2 ANALYSIS ALGORITHM

The present analysis is based on the event generation program (ISAJET Version 5.2) of CDF off-line package [4] added by a simple detector simulation. Actually we assume a calorimeter with tower size of $\Delta\eta$ = 0.1 units of pseudorapidity and $\Delta\varphi$ = 15 degrees and no magnetic field. We also assume that the lepton identification can be made perfectly. We identify the energy flow into a calorimeter cell, if we know from ISAJET that it is not of a lepton, as a hadron hitting the center of a calorimeter tower with the same energy.

Although the number of hadrons involved in the process is large, it would be no doubt unwise to handle kinematical variables of individual hadrons. We follow the conventional wisdom and group hadrons into clusters. In the present analysis we use a clustering algorithm CLUST-2 [4] of the CDF off-line software. The clusters thus identified do not necessarily correspond to partons in the hard collision, but approximately to 'final' partons coming from gluon bremsstrahlungs in the final stage. In addition, because of empirical nature of clustering algorithm, it is possible to group particles coming from two partons into one cluster as well as to count particles from a final parton into two clusters.

The method proposed here is to proceed with gluino search in the following steps.

1. We start with kinematical variables of a set of clusters, leptons and a missing E_T, and define kinematical parameters xi (i= 1, 2, ..., n) to describe the events. They can be any variables; for example, total transverse energy ΣEt, total missing transverse energy $E_{T,miss}$ or parameters characterizing momentum unbalanc of clusters such as P_E, Pout, x_E, and xout.
2. We generate the signal events and calculate the distribution of each variable and find out the mean value <xi> of the variables for the signal events.
3. We calculate the covariance matrix Cij of variables around its mean value;

$$Cij = \frac{\Sigma (xi^{\lambda} - \langle xi \rangle)(xj^{\lambda} - \langle xj \rangle)}{N \sigma i \ \sigma j}, \tag{1}$$

   where N is the total number of events, and σ i and σ j are variances of xi and xj.
4. We diagonalize the covariance matrix and find out transformation matrix Uij, Eigenvalues Eii, and the Eigenvectors or a new set of variable zi (i=1,2,...,n) which are given by

$$zi = \Sigma U^{-1}_{ij} (yj - \langle yj \rangle) / \sqrt{Eii}, \tag{2}$$

   where yj = (xj−<xj>) / σ j.
5. We get the probability function Psi(zi) of each variable zi for signal events. By generating the background events we also get the probability function Pbi(zi) of each variable zi for background events.
6. To see the effectiveness of a variable zi for the purpose of separating signal from background, we introduce a separation parameter Si by

$$Si = 0.5 \int |Psi(z) - Pbi(z)| \, dz. \tag{3}$$

   The separation parameter Si defined by Eq(3) takes a value between 0 and 1. If Si = 1, it means that the signal and backgroud separation can be done perfectly with zi, while Si = 0 means that the variable zi is totally useless for the separation. Thus we pick up an appropriate number (m) of parameters which give large values of Si.
7. We generate another set of samples of signal and background events to test the algorithm.
8. Now, we calculate for each event in the new samples, no matter whether it is a signal or a background event, a ratio of the probability that the event is a signal to the probability that it is a background by

$$r = \frac{Ps1(z1) x Ps2(z2) x \ldots\ldots\ldots\ldots x Psm(zm)}{Pb1(z1) x Pb2(z2) x \ldots\ldots\ldots\ldots x Pbm(zm)}, \tag{4}$$

   where m is the maximum number of parameters to be considered.
9. We define the likelihood that the event is a signal and not a background by

$$L = \ln r. \tag{5}$$

If the likelihood distribution of events is well separated between signal and background, it means that our method of the event identification works.

Since the initial kinematical variables are chosen somewhat arbitrarily, they are in general not mutually independent. To estimate the overall probability by multiplying the individual probability Psi or Pbi for a given parameter, all parameters should be mutually independent. This is the reason why we do the diagonalization of the covariance matrix Cij and get a set of transformed variables zi's. If an original parameter xi is strongly dependent on another, the variance of one of the new parameters zi should be very small.

It should be noted that the effectiveness of the method strongly depends on the choice of initial parameters. Since the transformed variables are only linear combinations of initial variables, the variables to be generated by multiplication or division of the initial parameters can never be produced in the proposed procedure.

3 SIMULATION AND RESULTS

3.1 Event Generation

The gluino pairs and gluon pairs were generated by using ISAJET V5.2. The gluino mass was assumed to be 200 GeV/c^2. The transverse energy range accepted in the genration was 10 < Pt < 200 GeV/c for gluinos and 100 < Pt < 500 GeV /c for gluons. The energy resolution was assumed to be 15 % / $\sqrt{E}$ for EM and 80 % / $\sqrt{E}$ for hadron calorimeter. The event was supressed if the reconstructed missing transverse energy was below 20 GeV. The fraction of events which passed this cut was 96 % for gluino pairs and 35 % for gluon pairs. The cross sections under these conditions are 27.1 nb for gluino and 8.40 μb for gluon pair production respectively. Total number of events to get probability functions was 500 and 2000 for signal and background respectively. The same number of events for each process were generated to study the effectiveness of the method.

3.2 Initial Parameters

Parameters initially chosen are as follows.

$$x1 = [1/(N-1)] \sum_{i=2} E_{T,i} \cos\varphi_{T1i} / E_{T,1} \qquad (E_{T,i} > 10\,\text{GeV})$$

$$x2 = [1/(N-1)] \sum_{i=2} E_{T,i} \cos^3\varphi_{T1i} / E_{T,1} \qquad (E_{T,i} > 10\,\text{GeV})$$

$$x3 = [1/(N-1)] \sum_{i=2} E_{T,i} \cos^5\varphi_{T1i} / E_{T,1} \qquad (E_{T,i} > 10\,\text{GeV})$$

$$x4 = (1/(N-1)] \sum_{i=2} E_{T,i} \sin\varphi_{T1i} / E_{T,1} \qquad (E_{T,i} > 10\,\text{GeV})$$

$$x5 = \sqrt{(\sum |Pi|)^2 - |\sum Pi|^2} \qquad (i = \text{clusters within } R_{1i} < 1)$$

$$x6 = \sqrt{(1/(N-1)] \sum_{i=2} P_i^2 \sin^2\theta_{1i}} \qquad (E_{T,i} > 10\,\text{GeV})$$

$x7 = P_E$
$x8 = \text{Pout}$
$x9 = E_{T,miss}$
$x10 = \sum Et$
$x11 = N$ (Number of energy clusters with $E_T > 10$ GeV)

where $E_{T,i}$ is transeverse energy of i-th cluster. φ_{T1i} and θ_{1i} are opening angle between the leading energy cluster and the i-th energy cluster . The subscript T in φ_{T1i} means the angle calculated in the transverse plane to the beam axis. R_{1i} is distance between the leading cluster and the i-th cluster and defined as $\sqrt{\Delta\theta^2+\Delta\varphi^2}$.

3.3 Diagonarization And Distribution Of Transformed Parameters

The covariance matrix elements were calculated by Eq(1) and was diagonalized. By a standard procedure Eigenvectors or a new set of parameters zi were obtained. Figures 1(a) and 1(b) show correlations of typical two parameters before and after diagonarization respectively. We observe the correlation is removed between the parameters zi's. Table 1 shows new parameters zi's as linear combinations of original parameters xi's. The event distribution and probability function for typical new parameters are shown in Figs.2(a) and 2(b). To get probability functions, we did smoothing for event distribution using Spline method [5].

To evaluate the efficiency of separating signal from background, we calculated the separation parameter Si for each transformed parameter zi. The values of separation parameters are also listed in Table 1.

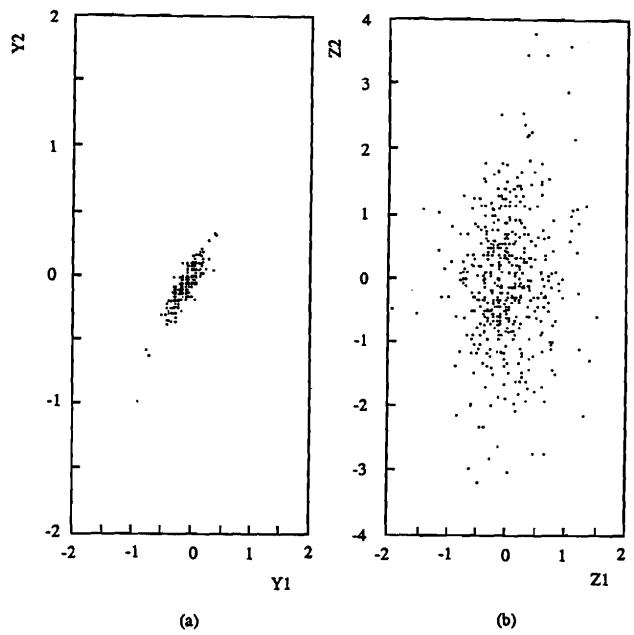

Fig.1 Correlation plots between typical two parameters; (a) x1 and x2, (b) y1 and y2.

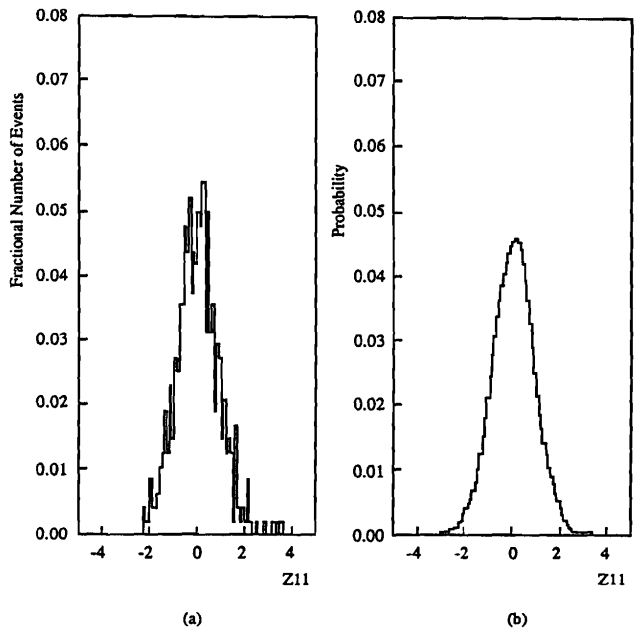

Fig.2 (a) Events distribution of z11 and (b) Probability function Ps11(z11) obtained by using Spline method [5].

Table 1 Transformation Matrix U^{-1}_{ij}

		----- xj -----											
		1	2	3	4	5	6	7	8	9	10	11	Si
	1:	-0.2638	0.7921	-0.5495	0.0088	-0.0037	-0.0004	0.0261	-0.0062	0.0088	-0.0142	-0.0061	0.194
	2:	-0.0350	0.0126	0.0887	0.0890	0.0315	0.0202	0.6614	-0.0904	0.2356	-0.6826	0.1198	0.236
	3:	-0.7947	0.1014	0.5199	-0.0332	0.0327	0.0174	-0.1751	-0.1449	0.1141	0.0218	0.1425	0.184
	4:	0.0474	-0.1546	-0.2666	0.0043	0.0546	0.1570	-0.4567	-0.0337	0.0311	-0.3287	0.7459	0.257
	5:	0.0143	0.1724	0.2273	-0.3948	0.0825	0.2855	-0.0338	0.4664	-0.5925	-0.3173	-0.0626	0.303
zi	6:	0.0805	0.0638	0.0564	-0.3892	-0.8385	0.2611	0.0549	-0.0670	0.2066	0.0719	0.0906	0.229
	7:	0.0110	-0.0881	-0.1329	-0.3459	0.3692	0.5538	-0.1809	-0.2546	0.3827	-0.1098	-0.3946	0.268
	8:	-0.0046	-0.0480	-0.0757	-0.6839	0.1319	-0.6756	-0.0500	0.0399	0.1901	-0.0982	0.0197	0.182
	9:	-0.1550	-0.1113	-0.0868	0.1657	-0.0560	0.0347	-0.0873	0.8128	0.4963	-0.0231	-0.0959	0.299
	10:	-0.1723	-0.1759	-0.1660	0.2072	-0.3565	-0.2141	-0.3899	-0.1404	-0.1720	-0.5145	-0.4795	0.343
	11:	0.4847	0.5011	0.4854	0.1606	0.0336	-0.1213	-0.3515	-0.0305	0.2796	-0.1806	-0.0392	0.393

3.4 Likelihood Values

For the final step of event identification, we calculated a likelihood value L for each event according to Eq(4). For this purpose we used parameters whose separation parameter S is larger than 0.25. This threshold is rather arbitrary, but we confirmed that addition of other parameters with lower values of S does not improve the separation. The likelihood distributions for signal and background are shown in Fig.3. Since the background cross section is about 300 times larger than the signal cross section and we don't have enogh statistics of background events at high values of L, we evaluated the background in the higher L region by using a Gaussian fit. The rejection power against background and the acceptance for signal are shown in Table 2 with various cut values on likelihood.

As we pointed out in Section 2, the effectiveness of the method strongly depends on the initial coice of kinematical parameters. The result may be improved by adding better parameters, i.e. parameters with large S values. In the present study, we looked the gluino Pt region of 100 to 200 GeV/c. If we take the higher Pt region, the S/N ratio will be also improved.

4 CONCLUSION

We tried a likelihood method to separate the gluino pair production from QCD background associated with detector resolution. The gluino mass was assumed to be 200 GeV and the gluino Pt region of 100 to 200 GeV/c was studied. The rejection power against the background was 1.9×10^3 with signal acceptance of 6.6 %, which correpond the S/N ratio of 6.2.

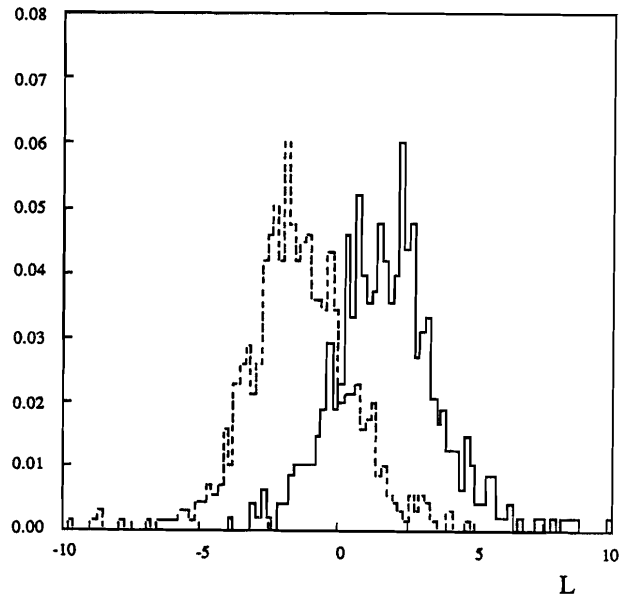

Fig.3 Likelihood distributions for signal (solid line) and background (dashed line). Both distributions are normalized by the number of events with $E_{T,miss} > 20$ GeV.

REFERENCES

1) S.H.Aronson, et al., Proceedings Of The 1982 DPF Summer Study On Elementary Particle Physics And Future Facilities (1982), p505.
I.Hinchliffe and L.Littenberg, ibid, p242.

2) G.Arnison, et al., CERN-EP/85-108 (1985).
L.DiLella, Proceedings of the 1985 International Symposium on Lepton and Photon Interactions at High Energies, Kyoto (1985), p280.
R.M.Barnett, H.E.Haber and G.L.Kane, Nucl. Phys. B267 (1986) 625.

3) R.Odorico, Proceedings Of The 1982 DPF Summer Study On Elementary Particle Physics And Future Facilities (1982), p478.
J.M.Hauptman, Proceedings of the 1984 Summer Study on the Design and Utilization of the Superconducting Super Collider (1984), p121.
E.Fernandez, et al., ibid, p107.

4) J.E.Freeman, R.Perchonock, Y.Takaiwa, and J.K.Yoh, private communication.

5) B.Schoor, CERN-DD/EE/78-3.

Table 2 Acceptance and Rejection Power

	$\tilde{g}\tilde{g}$		$g\,g$			
Pt range	10 - 200 GeV/c		100 - 500 GeV/c			
σ(nb)	27.1 nb		8.40×10^3 nb			
Ngen	500		2000			
N(ETmiss > 20 GeV)	481		698			
Lcut	Accep.	σ(nb)	Accep.	σ(nb)	Rejection Power	S/N
4.8	6.6×10^{-2}	1.8	3.5×10^{-5}	0.29	1.9×10^3	6.2
4.2	9.8×10^{-2}	2.7	1.4×10^{-4}	1.13	7.0×10^2	2.4
3.6	14.4×10^{-2}	3.9	4.7×10^{-4}	3.90	3.1×10^2	1.0