Available on CMS information server **CMS IN 2002/XXX**

CMS Internal Note

The content of this note is intended for CMS internal use and distribution only

May 17, 2002

On Optimization of the Trigger Selection for the Low-Mass Scale SUSY

S. Abdullin[a)]

Experimental High Energy Physics Group, Department of Physics, University of Maryland, Regents Dr., College Park, MD 20742, USA

Abstract

We have used a hybrid genetic algorithm to optimize Level-1 and Level-2 cuts at low luminosity for the generic SUSY signatures : multijet + E_T^{miss}. 3 probing points in the framework of mSUGRA model at the upper mass reach of the Tevatron II are studied. The case of R-parity violation scenario is considered, as the most challenging one, along with the "standard" case of R-parity conservation. The conservative L1 and L2 scenarios are considered : 2 kHz and 3 Hz for jets + E_T^{miss} channels at L1 and L2 respectively. The main background channels involved in the rate evaluation are QCD, $t\bar{t}$ and Wj (W $\rightarrow l\,\nu$).

D R A F T

[a)] On leave of absence from ITEP, Moscow, Russia. Email: abdullin@mail.cern.ch

1 Introduction

One of the main aims of the LHC collider is the search for physics beyond SM. One of the direction of this search is to look for superpartners of ordinary particles expected in various SUSY models developed over last 20 years. SUSY implies the existence of a symmetry between bosons and fermions and seems to be a natural high-energy scale extension of the Standard Model (SM). The "classical" SUSY signature with a few jets plus fairly high missing E_T in the final state is usually considered as a most promising SUSY discovery channel via excess of the SUSY signal over SM expectations [1]. One of the popular SUSY models is so-called mSUGRA with the limited number of parameters due to constraints imposed Supergravity unification [2] :

- a common gaugino mass ($m_{1/2}$) ;
- a common scalar mass (m_0);
- a common trilinear interaction amongst the scalars (A_0);
- the ratio of the vacuum expectation values of the Higgs fields that couple to $T_3 = 1/2$ and $T_3 = -1/2$ fermions (tanβ);
- a Higgsino mixing parameter μ which enters only through its sign (sign(μ)).

For a given choice of model parameters all the masses and couplings, and thus production cross sections (up to structure function) and decay branching ratios are fixed. The important feature of the standard mSUGRA scenario with conserved R-parity is that the lightest supersymmetric particle (LSB) is stable and neutral ($\tilde{\chi}_1^0$). So there are at least 2 such detector-escaping particles in the final state resulting in a significant missing E_T. A typical multijet mSUGRA event in the CMS calorimetry is shown in Fig.1. Two $\tilde{\chi}_1^0$ (not necessarily back-to-back and balanced!) from cascade decays of primary sparticles plus some neutrinos (visible with accompanying muons from b-jet approximately "at 8 o'clock") create significant energy imbalance in the transverse plane.

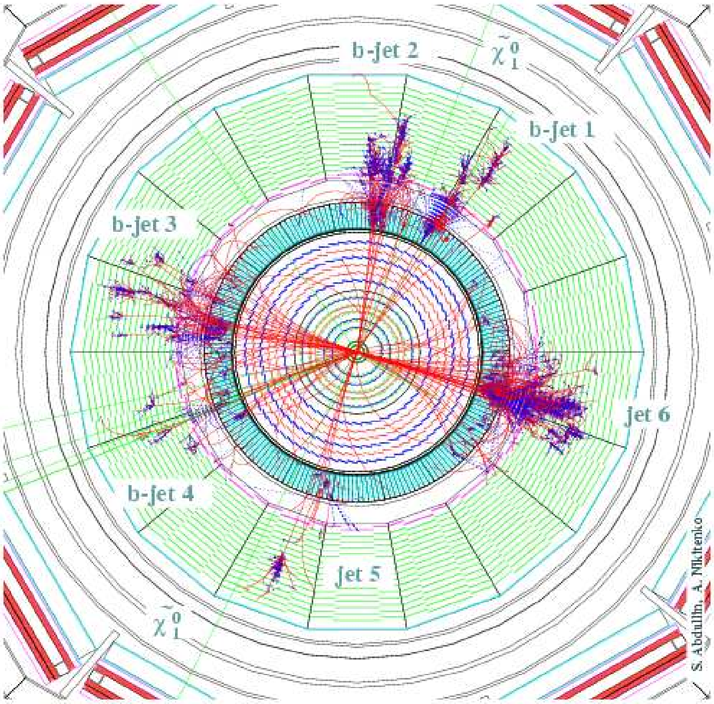

Figure 1: GEANT transverse view of a typical multijet + E_T^{miss} SUSY event in the CMS calorimetry.

The case of R-parity violation is commonly considered as more challenging as $\tilde{\chi}_1^0$ is not stable anymore and this reduces missing E_T, unless $\tilde{\chi}_1^0$ decays to leptons with the significant branching ratio. The latter circumstance may even facilitate the trigger in particular and SUSY observability in general. The most challenging scenario involves purely hadronic decay of $\tilde{\chi}_1^0$, when final states contain mainly jets without significant missing E_T. Some modest missing E_T still may result from (semi)leptonic decays of W/Z bosons and b-jets. Here we have high QCD background rate to straggle with.

The LHC starts in 2007 when a significant integrated luminosity (20 fb^{-1}) is supposed to be collected at Tevatron II. The possible sparticle mass reach in the mSUGRA framework with conserved R-parity at Tevatron is considered in [3]. This maximal reach is assumed to be a starting point for LHC. Basically it is defined by masses of the strongly-interacting sparticles (squarks and gluino), the choice of other parameters is not crucial for generic SUSY signature : jets + E_T^{miss} in most of the cases. There are some arguments in favour of positive value of μ and "intermediate" value of tanβ = 5–20 [4]. So for the sake of simplicity we consider three parameters of mSUGRA model are fixed in our study : A_0=0, tanβ=10, μ >0.

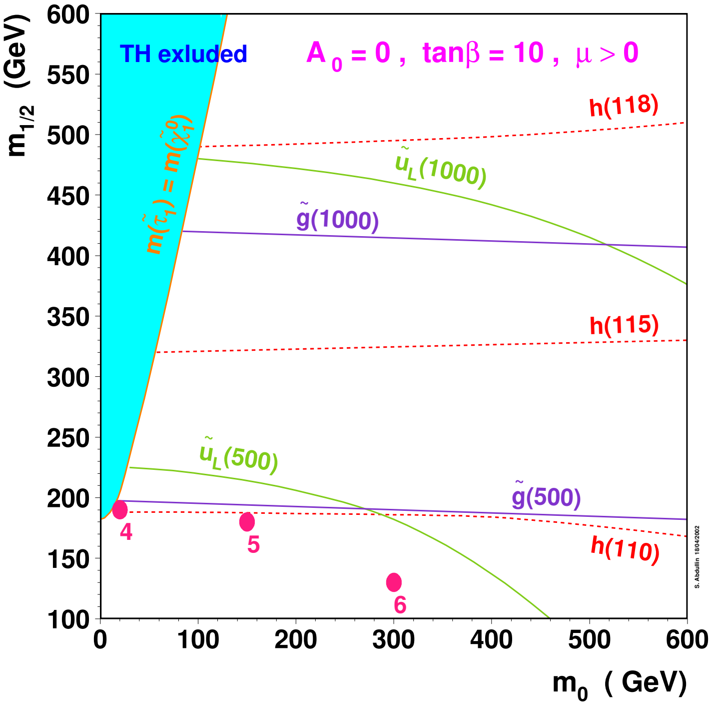

Figure 2: mSUGRA parameter space in terms of m_0, $m_{1/2}$ with the rest of mSUGRA parameters fixed : A_0=0, tanβ=10, μ >0. Parameters in parentheses show the mass of the corresponding sparticles in GeV. Three enumerated filled circles denote the probing points used in the study. See also comments in the text.

Fig.2 shows a typical representation of the mSUGRA parameter space in terms of m_0, $m_{1/2}$ with other 3 parameter fixed. One can see isomass curves for lightest scalar Higgs (one of 5 Higgses in Minimal Supersymmetry model from which mSUGRA is derived), gluino and left u-squark (which is often the heaviest squark). The inscription "TH excluded" stands for theoretically excluded (in R-parity conserving model) region of the parameter space where $\tilde{\chi}_1^0$ is not the lightest sparticle anymore. One may see the boundary line of this region where m($\tilde{\chi}_1^0$)=m($\tilde{\tau}_1$). The 3 filled enumerated circles lie approximately on the maximal Tevatron reach boundary, which is close to the 400 GeV squark isomass curve. The numeration of the points is quite arbitrary and just follows the order in which they were generated and processed (among many other mSUGRA points) in the CMS production. More details about these 3 points are shown in Table 1. In case of R-parity is broken in such a way that $\tilde{\chi}_1^0 \rightarrow$ jjj one might expect up to 6 additional jets in the final state with a substantial reduction of E_T^{miss}.

The LSB decay of course does not affect neither production cross section nor mass spectrum of sparticles and Higgses. Keeping in mind that the mass of LSB is fairly small (see Table 1) one might suppose that these additional jets are quite weak, so that the loss of E_T^{miss} is not fully compensated by additional jets and this might hamper effective triggering on such events. One of the goal of this study is to quantitatively check this guess.

It is quite clear that from the Trigger point of view higher masses of squarks and gluino should generally result in the efficiency increase, as jets from cascade decays of primary sparticles become harder and E_T^{miss} rises also. Of course, the production cross section decreases with the mass scale increase and the SUSY signal observability as a signal excess over SM background for a given luminosity might degrade, but this question is out of the frames of our trigger-oriented study.

In this paper we make an attempt to optimize Level-1 and Level-2 trigger cuts to get the maximal mSUGRA efficiency at the starting line : low luminosity regime and mass scale about maximal reachable one for Tevatron. We consider both R-parity conserved and R-parity broken scenarios.

Table 1: Production cross section and masses of main sparticles for probing mSUGRA points

mSUGRA point	m_0 (GeV)	$m_1/2$ (GeV)	σ_{tot} (pb)	$m_{\tilde{g}}$ (GeV)	$m_{\tilde{u}_L}$ (GeV)	$m_{\tilde{\chi}_1^0}$ (GeV)	m_h (GeV)
4	20	190	181	466	410	70	110
5	150	180	213	447	415	66	110
6	300	150	500	349	406	45	106

2 Available Data Samples

We use, along with the mSUGRA data, 3 main jets + E_T^{miss} rate contributors : W+jets (with W $\rightarrow l\nu$), $t\bar{t}$ and QCD di-jet events. The latter data sample is produced in 22 $\hat{p}_T$ bins to get sufficient statistics of events with high transverse momentum. All these 3 SM data samples were generated with PYTHIA 6.152 [5]. The data is described in Tables 2 and 3. All the data samples are produced at FNAL (Fall 2001 production) and finally selected information is stored to Ntuples using Ntuple maker of the HCAL Jets/MET group [6]. These Ntuples contain L2 jets found with Iterative Cone Algorithm with cone size of R=0.5 and seed cut E_T^{seed}=1 GeV on EcalPlusHcalTower objects. We use a minimal cut of E_T=30 GeV on jets in our study. It's worth also mentioning that both L1 and L2 jets are corrected using low-luminosity corrections [8].

No preselection (off-line cuts) was applied to the data samples. So all the efficiencies are evaluated with respect to the full generated signal samples, unless explicitly stated otherwise.

Table 2: Wj and $t\bar{t}$ data samples used in the analysis

Data Sample	σ (pb)	Number of Events	Comment
W + jets	4680 ($\sigma \cdot$ Br for $\hat{p}_T > 30$ GeV)	146397	W $\rightarrow l\nu$
$t\bar{t}$	800	46000	no forced decays or cuts

3 Calculational Details

The rate calculation is based on the LHC cycle parameters [9] : the number of non-empty bunches per second is 40.079 MHz (LHC frequency) x 2808 /3565 (ratio of filled/total bunches in the LHC cycle) = $3.158 \cdot 10^7$ Hz. So we have about 32 MHz collision rate.

To calculate the rate contribution from QCD bins we apply bunch crossing weighting technique from [10] which uses in-time $\hat{p}_T >$ values of all events, their relative cross section fraction f_i from Table 3 and number of events in each $\hat{p}_T <$ bin to calculate the weight (expressed in terms of rate) of each bunch crossing.

We also apply a filter described in [11] to eliminate some events (less than 1 % of the total statistics) with abnormal energy deposits due to improper simulation of δ-rays in CMSIM 121 [12].

To facilitate the processing of the information, all the data samples (Ntuples) were further suppressed to the binary files containing only the very essential information, like : four leading (descending E_T order) L1 and L2 jets and missing E_T and a few other quantities which are thought to be useful in our analysis. This compresses the information down to the size of a couple of hundred Megabytes which makes possible the keeping of all this

Table 3: QCD binned data samples.

Bin No.	$\hat{p}_T$ (GeV)	f$_i$ = $\sigma_i/\sigma^*_{total}$	Number of Events
1	0-10	$7.880 \cdot 10^{-1}$	87115**
2	10-15	$1.606 \cdot 10^{-1}$	95546
3	15-20	$3.357 \cdot 10^{-2}$	129450
4	20-30	$1.416 \cdot 10^{-2}$	106270
5	30-50	$3.348 \cdot 10^{-3}$	185299
6	50-80	$4.410 \cdot 10^{-4}$	143005
7	80-120	$6.083 \cdot 10^{-5}$	137992
8	120-170	$1.024 \cdot 10^{-5}$	48210
9	170-230	$2.106 \cdot 10^{-6}$	9500
10	230-300	$5.092 \cdot 10^{-7}$	20080
11	300-370	$1.421 \cdot 10^{-7}$	8000
12	370-480	$4.340 \cdot 10^{-8}$	7999
13	480-600	$1.675 \cdot 10^{-8}$	3000
14	600-800	$5.257 \cdot 10^{-9}$	4300
15	800-1000	$1.003 \cdot 10^{-9}$	2000
16	1000-1400	$3.307 \cdot 10^{-10}$	6000
17	1400-1800	$3.635 \cdot 10^{-11}$	2000
18	1800-2200	$5.561 \cdot 10^{-12}$	1000
19	2200-2600	$1.028 \cdot 10^{-12}$	1000
20	2600-3000	$2.037 \cdot 10^{-13}$	1000
21	3000-3500	$4.690 \cdot 10^{-14}$	500
22	3500-4000	$6.230 \cdot 10^{-15}$	500

* The total QCD cross section is assumed to be 55.2 mb.

** The cut $\hat{p}_T < 10$ GeV is applied
as original sample does not have the upper $\hat{p}_T$ limit.

data permanently in the memory thus drastically speeding up the manipulation of the data in the optimization algorithms. We will see later (Section 4) that this is really an issue in the optimization procedure.

3.1 Simulation of R-parity Violation

As we mentioned above (Section 1) we choose the most challenging (for the trigger) case of R-parity violation with $\tilde{\chi}^0_1 \to jjj$. The mass spectrum and branching ratio were calculated with the help of ISAWIG 1.104 [13], which in turn is using ISAJET 7.51 [14] routine SUGRUN. ISAJET version 7.51 is generally treated as a generator which provides benchmark masses and branching ratios [4]. The resulting masses and branchings (including $\tilde{\chi}^0_1 \to jjj$ decay) were imported by HERWIG 6.301 generator [15] and the full set of sparticle pair production was generated and stored in HEPEVT Ntuples as an input for the full chain of the CMS detector simulation with CMSIM 121 [12] and event reconstruction with ORCA 5 [16].

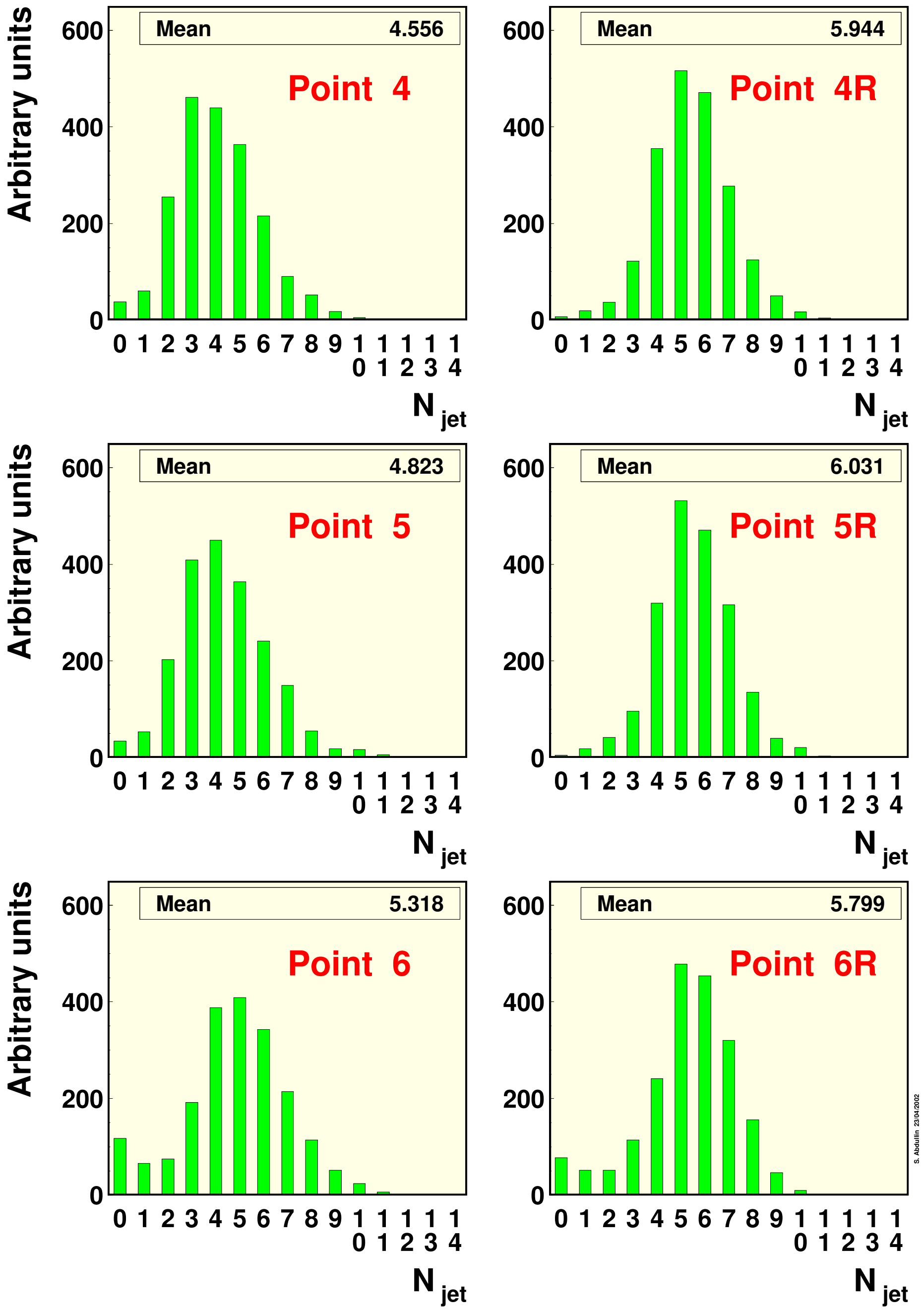

Figure 3: Multiplicity of jets with E_T^{jet} = 30 GeV in mSUGRA probing points with/without R-parity conservation. The latter case (R-parity broken) is denoted by "R" accompanying mSUGRA point number.

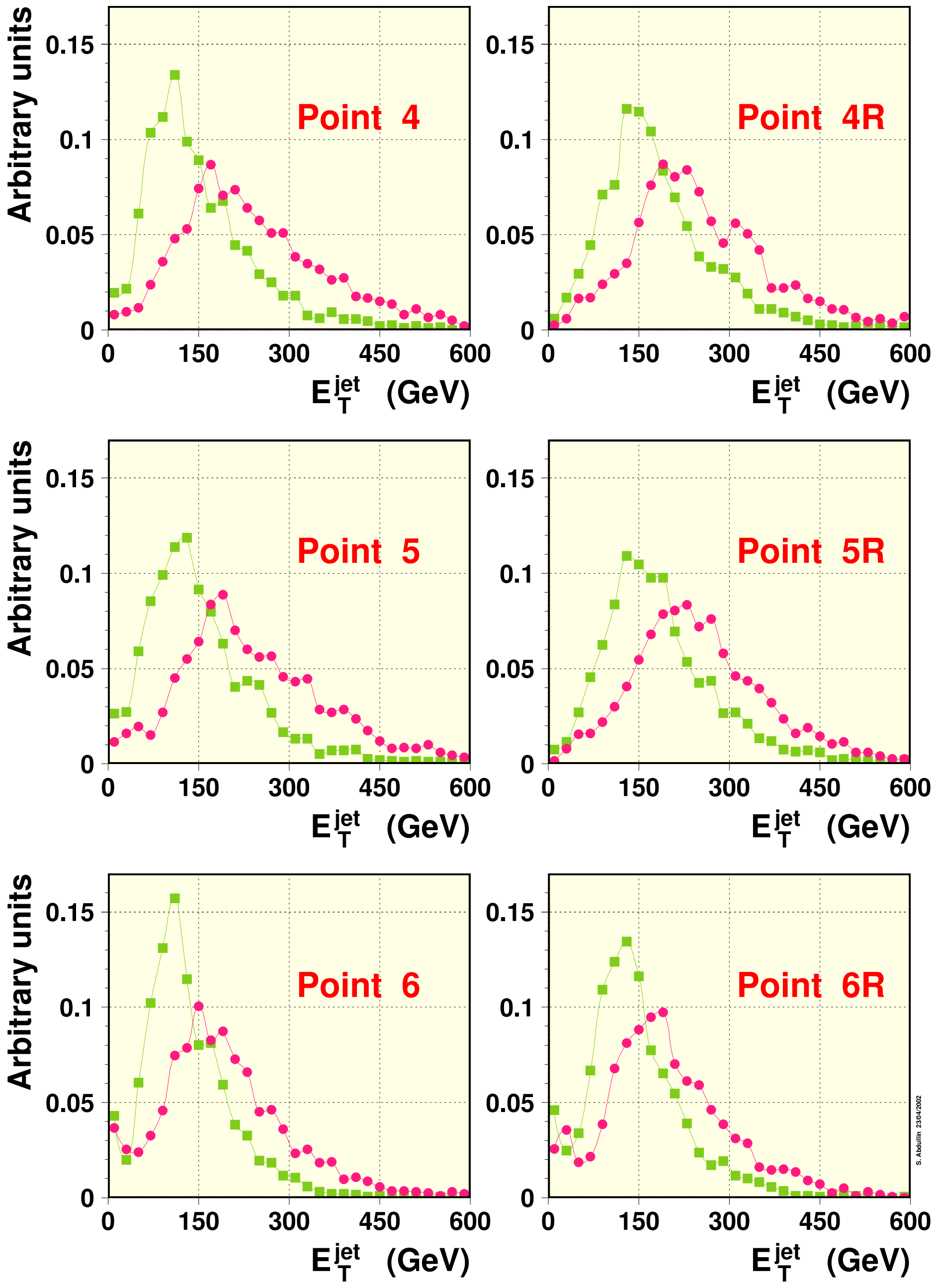

Figure 4: E_T distributions of the leading two jets in probing mSUGRA data samples. Pink filled circles denote the leading jet, while green filled squares - next-to-leading one.

3.2 Characteristic Features of the Probing mSUGRA Points

As mentioned in Section 1 and shown in Table 1, the mass of LSB is fairly small, below 70 GeV, in the mSUGRA region of interest. So 6 additional jets resulting from LSB decay are expected to be quite soft and not all of them can be detected, even at low luminosity. Fig. 3 shows the multiplicity of jets above the cut of E_T^{jet} = 30 GeV for R-parity conserving scenario (left column of 3 plots corresponding to various probing points) and for R-parity breaking scenario (right column with "R" notations on the plots). It is clearly seen that maximal increase in a number of jets between two R-parity cases is observed in case of the highest (of 3 probing points) mass of LSB in point 4. LSB decay into jets adds only ≈ 0.5 jet to point 6 original jet multiplicity, as the mass of LSB in this point is about 40 GeV only. Besides, point 6 "suffer" from having a significant fraction of the total production cross section taken by chargino-neutralino production.

Once the jets from LSB decay are fairly soft, one might assume there should not be a big contribution from those jets to the leading ones, which normally originate from decays of primary sparticles. The Fig. 4 just illustrates this fact. Still some small right shift, especially at relatively low E_T values, is observed in "R" data with respect to the corresponding standard mSUGRA due to occasional hardness of additional LSB-originated jets.

We see that we do not gain much in terms of additional jets in mSUGRA with R-parity broken with respect to the standard mSUGRA. But as there are no LSB anymore in R-parity broken scenario, missing E_T is expected to be small. Among the remaining sources of missing E_T in mSUGRA with broken R-parity are :

- copiously produced b-jets
- semi(leptonic) decays of W(Z)
- significant fraction of taus in decays of chargino and neutralino (especially at high tanβ)

One can see in Fig. 5 how E_T^{miss} "shrinks" if R-parity is broken the way we assume it. Still, even in point 6R the average E_T^{miss} is about that in $t\bar{t}$ events. Nevertheless, the loss of E_T^{miss} in R-parity broken scenario does not seem to be compensated by additional jets from LSB. So one might expect a priori more problems with trigger on mSUGRA events with broken R-parity.

3.3 Azimuthal Angle between Two Leading Jets

In paper [11] is shown that in general the distribution on the azimuthal angle between two leading L2 jets might be useful in the QCD rate suppression. In the same time, the authors concluded that this cut is not really helpful in case of relatively high mSUGRA mass scale (above 500 GeV) where simple cut on missing E_T works better. Fig. 6 shows the distribution of $\Delta\phi$(jet1,jet2) for a few set of conditions imposed on the QCD data as well as for Wj, $t\bar{t}$ and mSUGRA point 5 events. It is seen that even mSUGRA events tend to have back-to-back leading jets, so it is not clear a priori if the cut on this quantity can be more helpful at L2 than just the cut on missing E_T. We will try to consider our trigger optimization both with and without $\Delta\phi$(jet1,jet2) cut.

3.4 Baseline Trigger Rates

Currently a few L1 Trigger scenarios [17] and the final tape bandwidth are under consideration [18]. For our study we take rather conservatively 2 kHz for jets + E_T^{miss} channels at L1 and 3 Hz at L2.

One can easily calculate Wj and $t\bar{t}$ production rate at low luminosity ($2 \cdot 10^{33}$ cm^{-2} s^{-1} = 0.2 pb^{-1} Hz) using the numbers in the Table 2 : 9.4 and 1.6 Hz respectively. So at L1 with typical rates about few kHz QCD dominates. While at L2 these 2 channels may have a sizeable contribution to the total rate.

4 Trigger Selection Criteria

As it is not known a priori the concrete SUSY scenario, we need to cope with a variety of them. Though we do not consider here some "pathological" mass hierarchy possible in more general (than mSUGRA) MSSM model, when mass of LSB is quite close to mass of strongly interacting sparticles, which leaves no phase space for jets, thus hampering (or excluding at all) the generic SUSY observability. Here we use our 6 probing points in an attempt to find and to optimize L1 and L2 cuts in a way which results in a maximal trigger efficiency for all the points. Figs. 7 – 9 show some simple L2 correlations between jets and E_T^{miss} as efficiency contours for mSUGRA signal or as

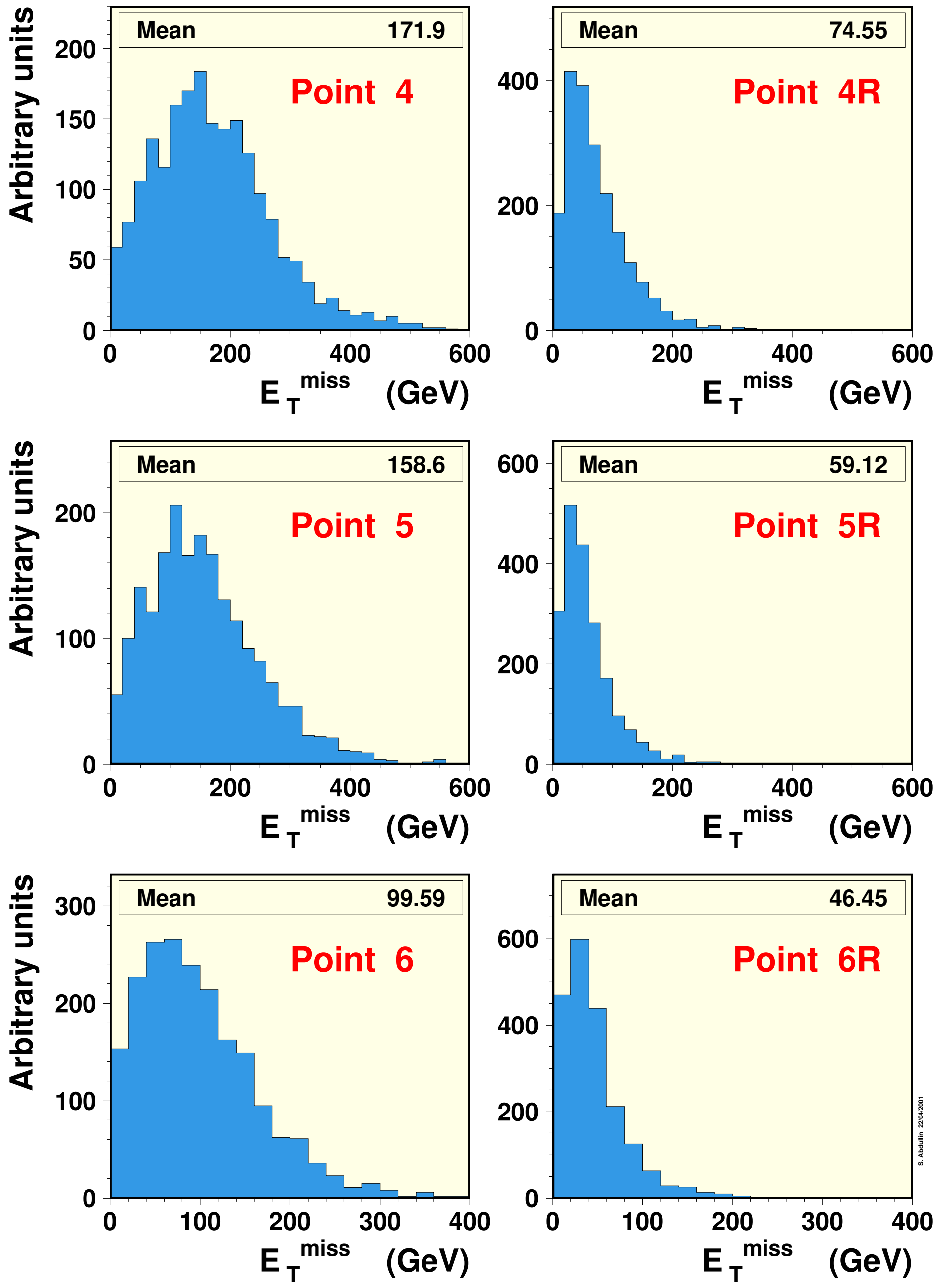

Figure 5: E_T^{miss} distributions in probing mSUGRA data samples.

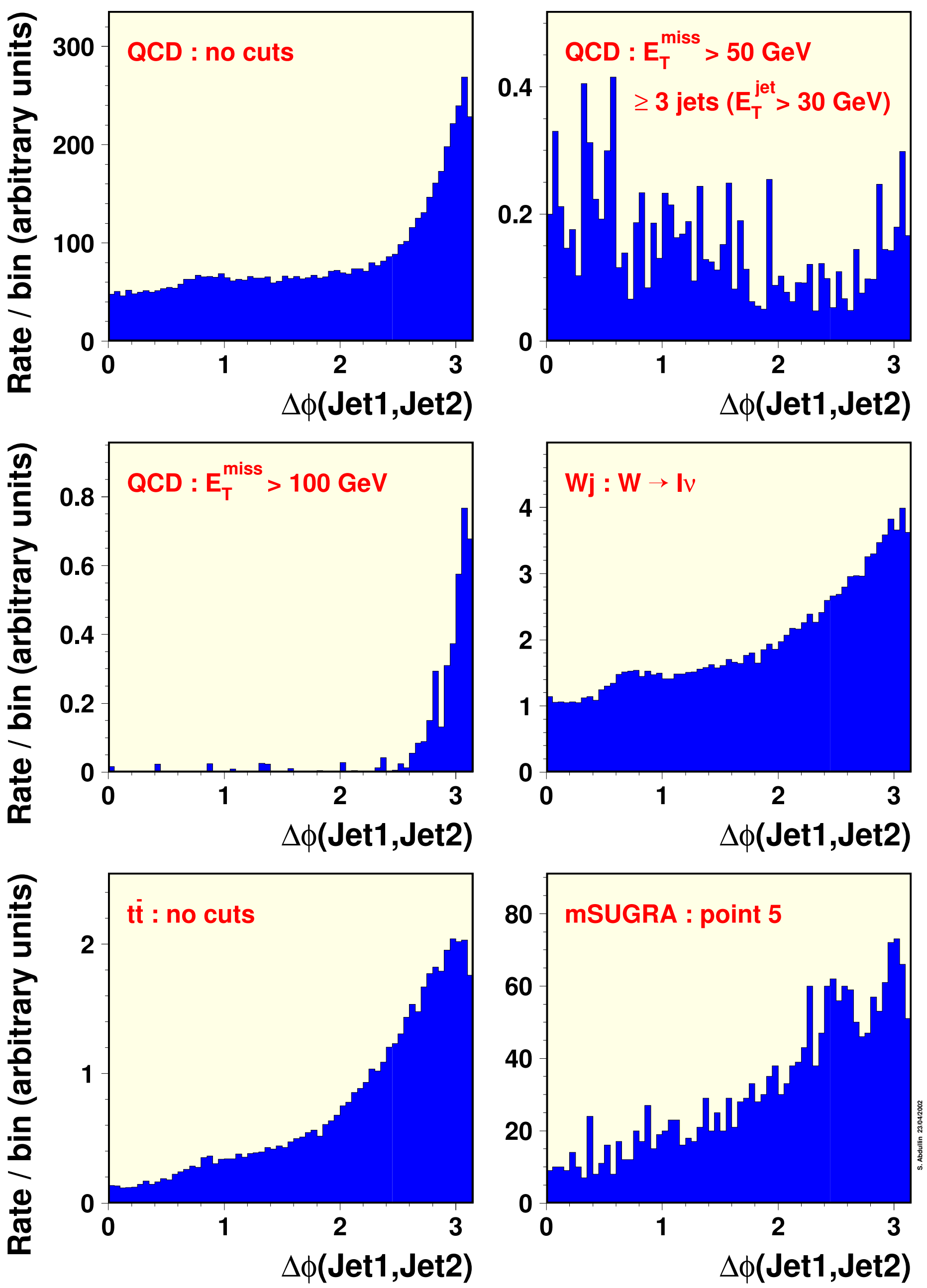

Figure 6: The distributions on the azimuthal angle between two leading jets for various processes. The distribution for mSUGRA point 5 is quite similar to those for other mSUGRA points (w/wo R-parity violation).

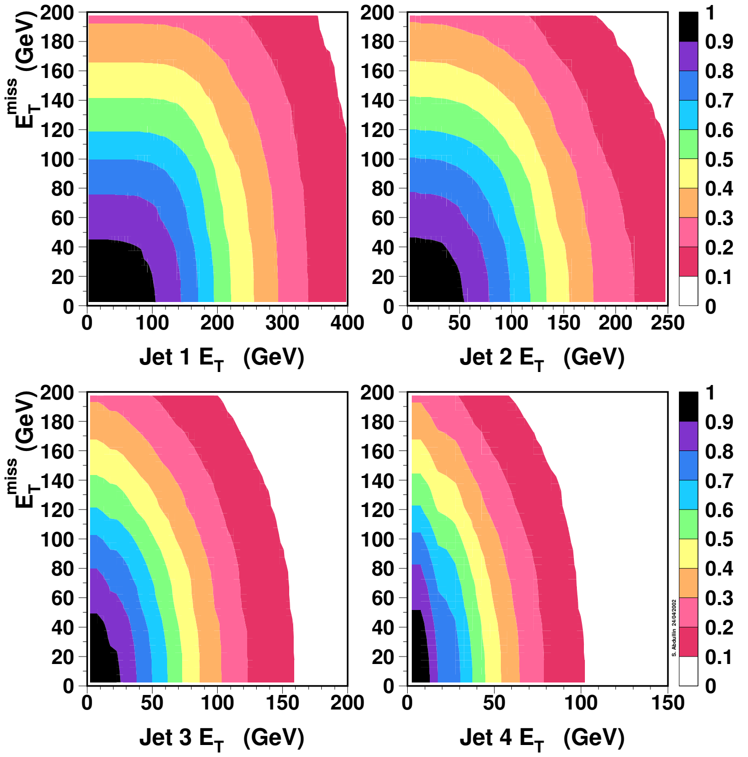

Figure 7: mSUGRA point 5 L2 efficiency as a function of E_T^{miss} and 4 leading jets E_T.

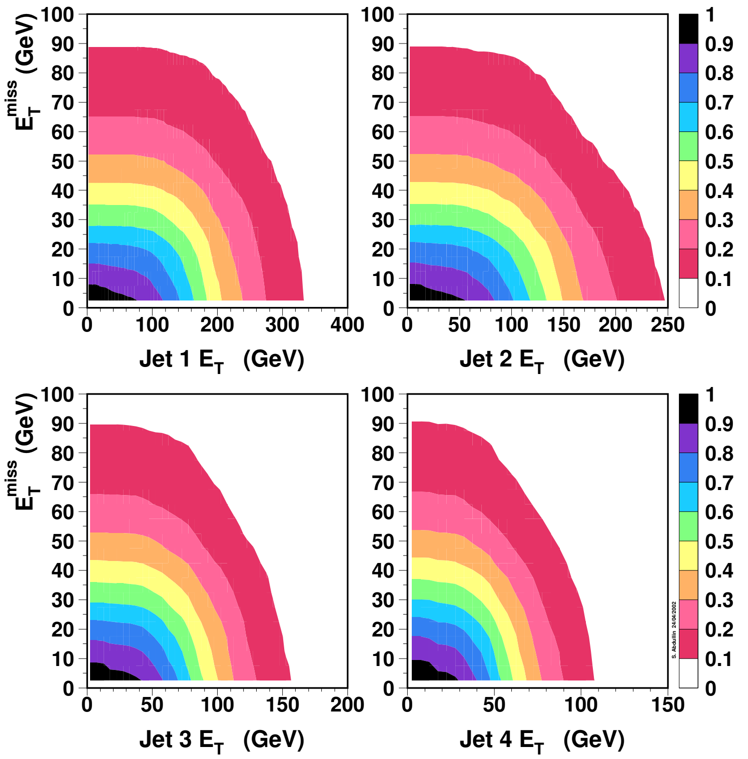

Figure 8: MSUGRA point 6R L2 efficiency as a function of E_T^{miss} and 4 leading jets E_T.

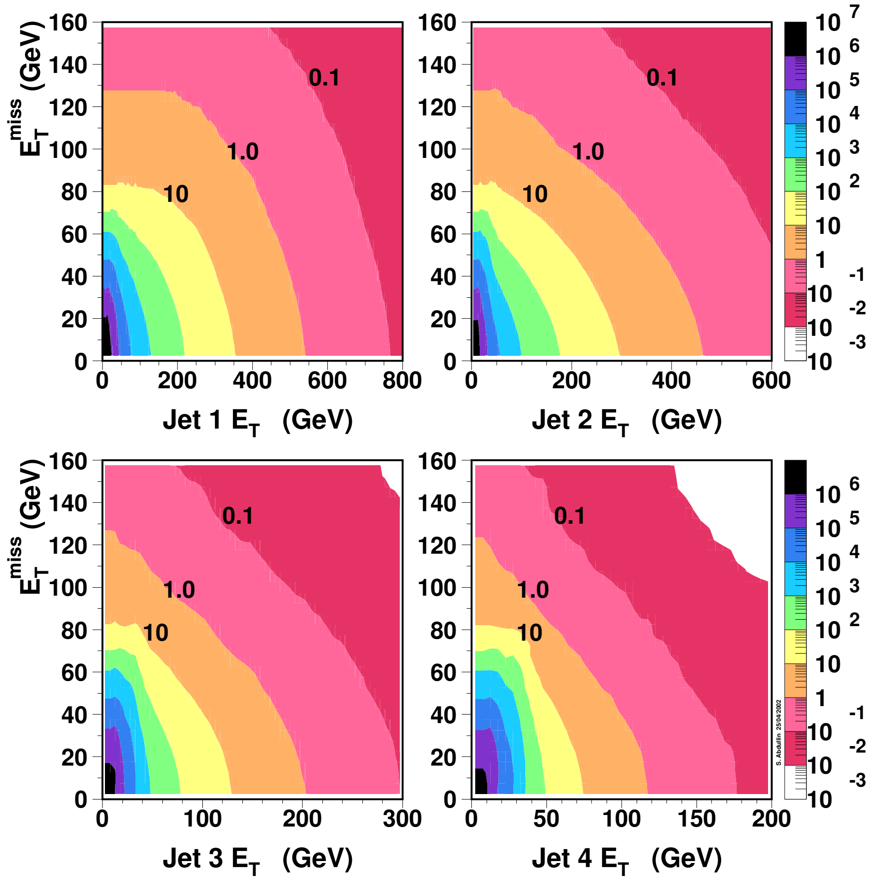

Figure 9: QCD L2 rate (Hz) as a function of E_T^{miss} and 4 leading jets E_T.

rate contours for QCD events. No L1 cuts are applied for the simplicity. Our goal is to choose multidimensional criteria space and to get the maximal signal efficiency at the given (within some narrow corridor) background rate.

Despite the fact that already at L1 we have an access to some topological variables, like $\Delta\phi$(jet1,jet2) or H (scalar sum of all jet E_T above some threshold), we use only cuts on jets and E_T^{miss}, composing them in 3 combinations :

- jet1.OR.jet2.OR.jet3.OR.jet4 (L1A)
- L1A.OR.(jet1.AND.E_T^{miss}) (L1B)
- L1B.OR.E_T^{miss} (L1C)

where logical expression jetN means L1 E_T^{jetN} $>$ appropriate cut etc. So event passes L1 trigger if satisfies the whole string of conditions. The criterion L1B is specially constructed to accommodate invisible Higgs trigger [19]. The absence of mentioned above topological cuts in our L1 consideration can be explained in particular by the inefficiency of H cut and, more generally, by sufficiently high efficiency of simple combination of jets and E_T^{miss} alone.

At L2 we have more freedom and can use more quantities, including topological cuts in order to keep the total rate at the appropriate (almost 3 orders of magnitude lower than L1 rate) level. Running ahead, some tried cuts turned out to be inefficient, such as scalar sum of jets E_T^{jet} + E_T^{miss}. Among the L2 criteria we applied in our study are :

- E_T^{miss}.OR.
(jet1.AND.E_T^{miss}).OR.(jet2.AND.E_T^{miss}).OR.(jet3.AND.E_T^{miss}).OR.(jet4.AND.E_T^{miss}) (L2A)
- $\Delta\phi$(jet1,jet2).AND.(L2A) (L2B)
- jet1.OR.jet2.OR.jet3.OR.jet4.OR.L2A (L2C)

- L2C.OR.($\Delta\phi$(jet1,jet2).AND.L2C) (L2D)

It's worth mentioning that 5 E_T^{miss} parameters entering expression L2A are independent, similar comment is valid for L2D (here 10 E_T^{miss}).

5 Optimization Functional (Evaluation Function)

In the Table 4 we put the number of parameters (often correlated) for each of L1 and L2 selection criteria listed above, as well as the number of possible combination of the parameters taking into account .OR./.AND. logic in each criterion. Given the time required to evaluate the background rate vs signal efficiency for a particular choice of parameters of certain criterion, we estimate the total CPU time (Pentium III 650 MHz) required for a simple scanning of all the combinations. Only L1A criteria look more or less doable within a reasonable time scale.

Table 4: Number of parameters and CPU time required to scan all possible parameter combinations

Selection criterion	Number of parameters (cuts)	Total number of combinations*	CPU time (sec) per combination	Total CPU time
L1A	4	$1.6 \cdot 10^5$	0.5 s	1 day
L1B	6	$6 \cdot 10^5$	-"-	1 year
L1C	7	$1.3 \cdot 10^5$	-"-	20 years
L2A	9	$\approx 10^{12}$	0.1 s	1000 years
L2B	10	$\approx 10^{13}$	-"-	10000 years
L2C	13	$\approx 10^{17}$	-"-	10^7 years
L2D	27	$\approx 10^{35}$	-"-	10^{26} years **

* Range of 20 (integer) bins is assumed
** The age of Universe $\approx 1.5 \cdot 10^{10}$ years.

So, generally speaking, the task can be formulated as an optimization of some functional which includes both total rate and signal efficiency in such a way that the rate is kept at some level and the efficiency is maximized. To perform "global" optimization of all 6 (both with/without R-parity violation) mSUGRA points we consider as a signal the sum of all 6 efficiencies with equal weights. To keep the rate at the given level we introduce the optimizable functional (which serves as evaluation function for the subsequent optimization, see Section 6) as follows :

$$F(\overrightarrow{P}) = S_{eff}(\overrightarrow{P}) - \frac{W}{R_{max} - R(\overrightarrow{P})} \qquad for R < R_{max} \tag{1}$$

$$F(\overrightarrow{P}) = S_{eff}(\overrightarrow{P}) - R(\overrightarrow{P}) \qquad for R \geq R_{max}, \tag{2}$$

where $S_{eff}(\overrightarrow{P})$ is weighted sum of signal efficiencies, *W* is a tunable weight which controls the closeness of the current rate $R(\overrightarrow{P})$ to some maximal rate limit R_{max}. The $\overrightarrow{P}$ denotes the vector of parameters (cuts). R_{max} and *W* are fixed in the optimization process.

The Fig. 10 show a sketch of the functional (1-2) behaviour as a function of $R(\overrightarrow{P})$. S stands for $S_{eff}(\overrightarrow{P})$, B_1 and B_2 for the second term (subtracted) in (1) and (2) respectively, while F is a shortcut for $F(\overrightarrow{P})$. The filled circle shows the maximal value of $F(\overrightarrow{P})$ and the bar on R-axis beneath F_{max} shows the supposed working range of optimization procedure, at least at the very end of optimization. The second half of the definition (2) helps to establish the hierarchy at the beginning of the optimization process, where most (or all) of the cuts combinations gives $R(\overrightarrow{P}) > R_{max}$, as this part of the parameter space is typically much broader than the region of interest denoted by the bar in Fig. Besides, (2) pushes optimization towards R_{max} and further (due to the gap narrowness). Once optimization reaches R < R_{max}, the functional increases significantly and it is highly "profitable" to stay on the left side of the narrow gap at R_{max}. Parameters W and R_{max} serve to keep a proper balance between S slope and B_1 slope in the vicinity of the F_{max} and to keep, thanks to the steepness of B_1, this point at the desirable R position.

It should be emphasized that this type of functional can be used only with some non-negligible step size of parameters, otherwise problems can arise when approaching the singularity point (R_{max}) from the right side.

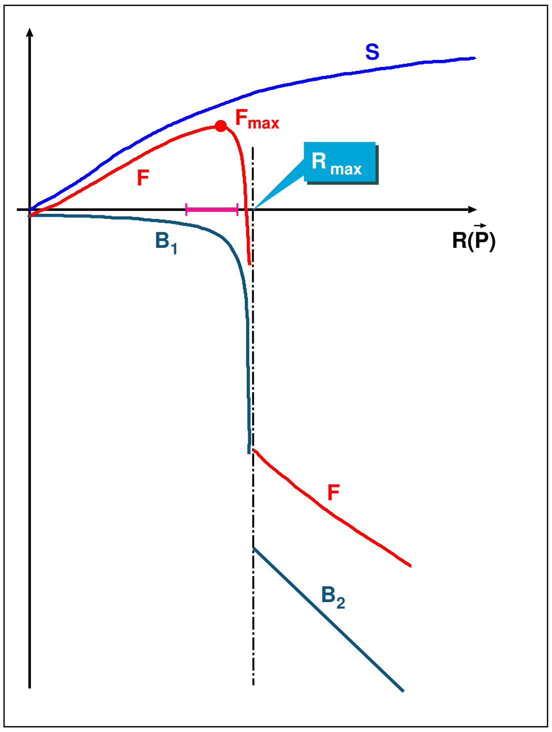

Figure 10: Sketch of the functional behaviour, see comments in the text.

6 Optimization Algorithm

As the simple scan over parameter space does not look feasible, we need some optimization algorithm which enables us to perform our optimization within reasonable time schedule operating with up to almost 30 parameters. We failed to achieve our goal using traditional MINUIT [20] either with continuous (using functionals different from one given by formulae (1,2)) or discrete parameter space. So we tried to look for some other methods of optimization, like genetic algorithms [21]-[22].

The code we wrote from scratch contains basic features of the algorithm of the sort. We consider a "society" of individuals of fixed-size population (100-1000) with some established hierarchy. Individuals are characterized by their own unique combinations of parameters-genes $\overrightarrow{P}$, which in fact are cuts of the selection criteria listed in Section 4. The hierarchy is maintained according to the value of F($\overrightarrow{P}$) from equations 1,2. A few basic "behaviour" functions are defined for the whole population of individuals :

- Ceation (initialization)
- breeding
- mutation
- selection (external with respect to population) with merciless regulation of the population size...
- cataclismic update

The optimization is a process involving generation-by-generation evolution of the population. Let us consider in some detail these basic behaviour operations.

Creation is intended to populate the list of individuals with randomly generated members before cyclic genetic algorithm starts.

Breeding involves a random crossover of genes (usually denoted as "uniform corssover") between two individuals with the creation of one offspring, while "parents" continue to exist as well. Probability to participate in the breeding depends linearly (our choice) on the place in the current hierarchy. Offsprings are added to the population.

Mutation means that some randomly chosen, independently from hierarchy, individuals in each generation (some significant fraction of the entire population) undergo a random change "up" or "down" by one quantum of the one randomly chosen gene. Each mutated individual produces an offspring added (not replacing a parent) to the population.

Selection takes place after previous steps are performed in order to finalize the generation. First, a removal of clones is performed. Clone (exact replica) of existing individual which might appear as a result breeding or mutation is replaced with a completely new individual Then population is ordered according to evaluation function and worst (redundant) individuals are eliminated to keep the nominal size of population. The best individuals may enter "Top 10" list if they are better than former members of this separate list. The last (redundant) individuals are removed from the list to maintain the fixed size of the entire population.

Cataclismic update is applied to the entire population when no progress occurs during a number of generations (stagnation). This means all individulas except the first (currently the best) one are replaced with completely new randomly created individuals. This idea comes from CHC algorithm [23], though we use not a heavy mutations of all the individuals, as in the original work, but rather a complete renewal of the population (but not the "Top 10" list !)

The algorithm described above does not follow one of the existing recipes, it is rather a combination of them. Probably this kind of genetic algorithm is close to GENITOR [24], but with some specific deviations and extensions. It's worth also mentioning we use non-binary coding of the problem, unlike most of the genetic algorithms do, but instead some final grid of states (final "alphabet") for each parameter (gene).

This cyclic process continues some given number of evolutions, then the process terminates. The best from "Top 10" is considered as the "summit of evolution" and its genes (parameters) are assumed to be the best ...

In fact, there is no guarantee that the best choice is not just a "local best", especially when the optimization is performed in quite a complicated parameter space, when comparable optimums can be found at a significant distance from each other. Moreover, in this algorithm there are no means to judge if this is a local or global best. So, a typical strategy consists in running a number of independent optimization processes in parallel, starting them from different random numbers.

A typical (single) optimization job with L2D criterion (Table 4) takes about 2-3 CPU hours and requires a few hundred generations. The algorithm demonstrates its abilities moving insignificant parameters to the values which do not bring a contribution to the rates/efficiencies, thus excluding them out of the process of optimization. In case of insufficient range of variation assigned to the parameter, the algorithm brings the parameter to the limit, so that the contribution is significant, but the limiting value of the parameter is reached. So the ranges of parameters should be defined wide enough to avoid this "borderline" effect.

7 L1 Optimization

Fig. 11 shows detailed break of the L1 trigger optimization results for the selection criterion L1C (see Section 5). Numbers denote either signal efficiency (%) or QCD trigger rate (kHz). Numbers in parentheses stand for partial values for given cut in each column, while numbers out of parentheses show cumulative values, increasing from left to right with each additional cut. Numbers in the rightmost column on the dark background show the total cumulative efficiency/rate. One can see that at the given rate of ≈ 2 kHz the signal efficiency is fairly (and sufficiently !) high for all types of the signals. One may also notice that MET, J1 and J3 cuts do not play a big role in the selection, the significant cuts are J1, J3 and combination of J1 and MET, see e.g. QCD rates line.

Table 5 shows only the resulting signal efficiencies at L1 for all 3 types of L1 criteria from Table 4. The rate is kept at the level of 2.03-2.05 kHz in all the cases. One may conclude that missing E_T cut alone is not really needed once J1 & MET combination is already included, cf. L1B and L1C results. Still even pure jet L1A criterion (without

MET cuts) gives sufficiently high efficiencies, which are just a few percent lower (and only for R-parity conserved scenario) than those provided by MET-included L1B and L1C.

		MET .OR.	J1 .OR.	J2 .OR.	J3 .OR.	J4 .OR.	(J1 .AND. MET)
	Cuts (GeV) →	90	160	140	60	50	90 + 50
signal efficiency (%)	Point 4	71 (71)	84 (76)	85 (47)	89 (58)	89 (31)	92 (86)
	Point 5	67 (67)	83 (77)	84 (51)	89 (63)	89 (35)	92 (86)
	Point 6	41 (41)	67 (61)	68 (39)	80 (66)	80 (43)	84 (67)
	Point 4R	33 (33)	86 (85)	87 (71)	93 (89)	93 (70)	94 (63)
	Point 5R	22 (22)	85 (85)	86 (72)	94 (89)	94 (71)	95 (55)
	Point 6R	14 (14)	68 (67)	69 (53)	87 (82)	87 (65)	88 (43)
Background rate (kHz)	QCD	0.02(0.02)	0.58 (0.58)	0.63 (0.36)	1.31 (0.85)	1.32 (0.14)	2.04 (0.91)
	$t\bar{t}$, W j ($l\nu$)	irrelevant !					

Figure 11: L1 SUSY trigger optimization results.

Table 5: Comparison of the signal efficiency for various L1 criteria

Criterion		E_T^{miss}	Jet1	Jet2	Jet3	Jet4	Jet1 & E_T^{miss}
L1A	cuts		130	-	60	-	
	Signal eff.(%) = 88.6 90.1 81.5 94.1 95.0 88.2						Bkg.rate (kHz) = 2.039
L1B	cuts		150	-	55	-	70 & 60
	Signal eff.(%) = 92.5 92.1 84.1 94.2 94.9 87.9						Bkg.rate (kHz) = 2.050
L1C	cuts	90	160	140	60	50	90 & 50
	Signal eff.(%) = 92.3 92.3 84.3 94.2 94.7 87.9						Bkg.rate (kHz) = 2.039

So, the conclusion is that simple E_T cuts on the first and the third jet (L1A) with the thresholds of 130 and 60 GeV respectively are quite suitable for SUSY selection at L1.

8 L2 Optimization

Fig. 12 shows detailed break of the L2 trigger optimization results for the selection criteria L2D (with L1C prior to). (see Section 5). Numbers denote either signal efficiency (%) with respect to the data samples passed L1 or the background trigger rate (Hz).

First of all, it should be mentioned that single jet cuts turned out to be inefficient at L2, as well as some other cuts included in the most complete criterion L2D. Only relevant (efficient) cuts are shown in Fig. 12, so that only 12 parameters left from initial 27 of L2D. The $\Delta\phi$(jet1,jet2) cut, as shown in the definition of criterion L2D in Section 4, is applied (logical .AND.) to 3 columns below the surrounding line with the inscription "$\Delta\phi$(jet1,jet2) $< 150^0$".

The numbers shown have the same meaning as those in Fig. 11, just some additional info is given. First, the rightmost column shows the final (absolute) rate of various signals. Then one may find, in addition to the QCD rate (2.52 Hz), also rates for $t\bar{t}$ and Wj channels. So after L2 (optimized) cuts QCD takes $\approx$ 80-90 % of the bandwidth, $t\bar{t}$ and Wj take about 3 % each and mSUGRA signals occupy 4 – 15 % depending on the signal nature. The total rate at L2 is a sum of background (2.74 Hz) rate and one of the signal rates (0.13 – 0.45 Hz) thus about 3 $\pm$ 0.2 Hz.

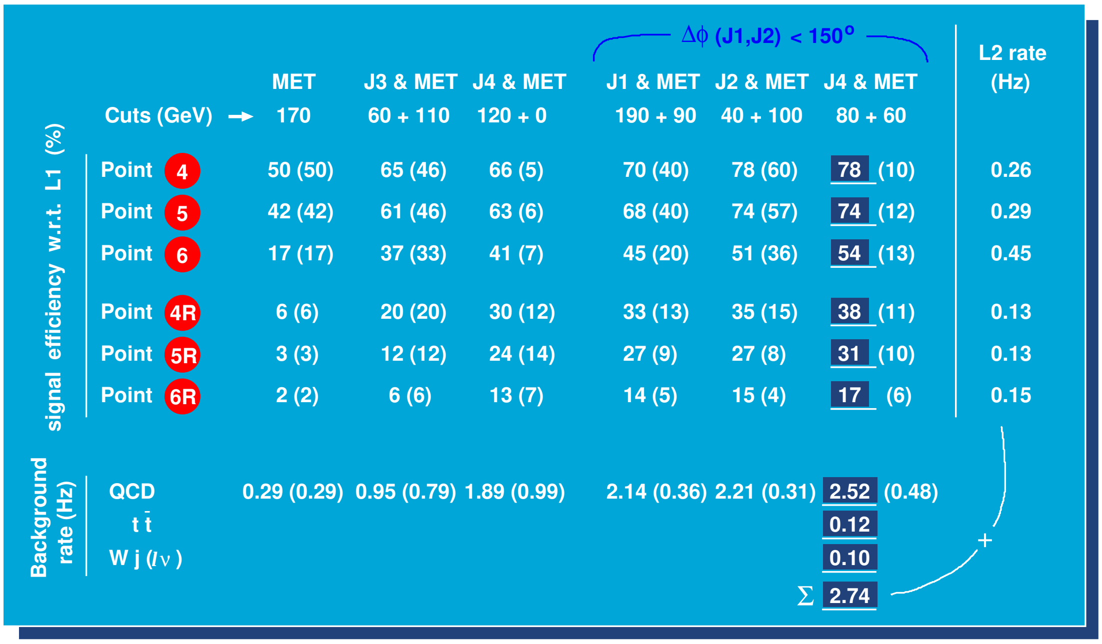

Figure 12: L2 SUSY trigger optimization results (L2D + L1C criteria)

Table 6: Comparison of the signal efficiency for various L2 criteria (with respect to L1 selection)

Criterion	Signal efficiency (%)						Background rate (Hz)
	4	5	6	4R	5R	6R	
L2A (+ L1C)	76.7	73.4	53.4	32.4	21.4	13.4	2.77
L2D (+ L1B)	79.1	76.2	54.5	36.3	28.3	17.1	2.67
L2D (+ L1C)	78.3	74.4	54.1	38.2	30.7	16.8	2.74

Fig. 13 is similar to Fig 12, except L2D criterion is combined with L1B. The criterion L2C effectively coincide with L2A due to mentioned ineffectiveness of the single jets cuts. Then, criterion L2B tends to coincide with L2A as well, as $\Delta\phi$(jet1,jet2) cut applied in such a way (before all other cuts) seems to be ineffective and optimization algorithm pulls it to 180^o. So the final L2 signal efficiencies (with respect to L1) are given in Table 6 only for criteria L2A and L2D. Due to the big total number of parameters involved in L2 study (max. 27) we omit cut values in the table, presenting just signal efficiencies. One can see that $\Delta\phi$(jet1,jet2) cut might help to improve signal efficiency for mSUGRA with R-parity violation (points 4R-6R) being combined with selected cuts as it is shown in Fig. 12.

To better understand events from which concrete L1 channel stream to various L2 channels we consider the sequence of quite simple criterion L1B (for traceability) followed by L2D. The Fig. 14 show the break of the partial QCD rates in each L2 channel according to the L1 origin. Once L1B have 3 conditions (cuts), then the criterion can be satisfied in 7 combinations : 3 exclusive, 3 pair combinations (2 cuts passed simultaneously) and 1 triple combination. One can see that none of the 3 L1 channels dominates in L2 cumulative rate, see bottom row of values on the dark background. L2 events mostly pass all the L1 cuts in combinations or just simultaneously all of them (last column), L1 cuts values are reminded in the upper left corner of the partial rate parallelogram as 150,

	MET	J3 & MET	J4 & MET	$\Delta\phi$ (J1,J2) $< 160^o$ MET	$\Delta\phi$ (J1,J2) $< 160^o$ J2 & MET	$\Delta\phi$ (J1,J2) $< 160^o$ J4 & MET	L2 rate (Hz)
Cuts (GeV) →	170	80 + 110	100 + 40	110	90 + 100	60 + 80	
signal efficiency w.r.t. L1 (%)							
Point 4	50 (50)	61 (33)	62 (9)	76 (65)	78 (51)	**79** (19)	0.26
Point 5	43 (43)	56 (34)	59 (10)	72 (60)	74 (52)	**76** (23)	0.30
Point 6	17 (17)	32 (24)	38 (12)	48 (35)	51 (31)	**54** (21)	0.46
Point 4R	6 (6)	18 (17)	28 (16)	31 (15)	33 (16)	**36** (16)	0.12
Point 5R	3 (3)	10 (10)	23 (16)	25 (9)	25 (9)	**28** (12)	0.11
Point 6R	2 (2)	6 (5)	13 (10)	14 (4)	15 (5)	**17** (7)	0.15
Background rate (Hz)							
QCD	0.29 (0.29)	0.73 (0.51)	1.79 (1.12)	2.0 (0.44)	2.14 (0.41)	**2.37** (0.50)	
$t\bar{t}$						**0.12**	
W j ($l\nu$)						**0.17**	
Σ						**2.67**	

Figure 13: L2 SUSY trigger optimization results (L2D + L1B criteria)

L2 Cuts (GeV)		150 J1 excl.	55 J3 excl.	70 60 (J1 & MET) excl.	J1 & J3	J1 & (J1&MET)	J3 & (J1&MET)	J1 & J3 & (J1&MET)
		Partial QCD rate (L1 origin), Hz						
MET	170	-	-	-	-	0.15	-	0.14
J3 & MET	80 & 110	-	-	-	-	0.01	-	0.50
J4 & MET	100 & 40	-	-	-	0.42	-	0.02	0.67
$\Delta\phi$ (J1,J2) $< 160^o$								
MET	110	-	-	0.02	-	0.15	0.02	0.24
J2 & MET	90 & 100	-	0.02	-	-	0.03	-	0.36
J4 & MET	70 & 60	-	-	-	0.04	-	0.04	0.42
Cumulative QCD rate (L1 origin), Hz		-	0.03	0.02	0.46	0.31	0.07	1.47

Figure 14: L2 QCD rate break according to L1 origin (L2D + L1B criteria)

55, 70 and 60 GeV over L1 cuts notations.

Now let us consider all possible L1+L2 trigger combinations for L1B + L2D (see Tab. 5 and Fig. 14) from the point of view of the signal efficiency. More precisely, we want to find the most signal-efficient L1+L2 triggers among all 18 combinations. The Table 7 shows the L1+L2 combinations we consider here. The two-digit notations are used (as matrix elements indices) for L1+L2 trigger combinations. The first index denote one of 3 L1 cuts and the second index - one of 6 L2 cuts. We consider all partial contributions for all 6 signal points, order them and evaluated the cumulative efficiency. It was found that only 7 combinations (out of 18) provide most significant contribution to the signal efficiency for all 6 signal samples. These 7 combinations are given in bold in Table 7.

Tables 8-9 show the resulting partial and cumulative signal efficiency (with respect the entire data sample) for chosen 7 combinations from Table 7 and comparison with the efficiency evaluated with all 18 combinations. Tables 10 is similar to Tables 8-9, except the background rates are shown. Due to the exclusion of the least significant L1 cut (first jet $E_T > 150$ GeV) we saved 0.5 kHz of bandwidth at L1 (now 1.55 kHz instead of former 2.05 kHz). Besides, the use of only 7 L1+L2 triggers instead of full set of 18 facilitates the trigger efficiency evaluation for the real detector, while the signal efficiency is kept almost untouched.

Table 7: L1+L2 trigger combinations for L1B + L2D case

	L2 Trigger cuts (GeV)					
				$\Delta\phi$(jet1,jet2) $<$ 160 deg.		
	MET 170	J3 && MET 80 && 110	J4 && MET 100 && 40	MET 110	J2 && MET 90 && 100	J4 && MET 60 && 80
L1 J1 (150)	1 1	1 2	1 3	1 4	1 5	1 6
L1 J3 (55)	2 1	**2 2**	**2 3**	2 4	2 5	**2 6**
L1 J1&&MET (70&&60)	**3 1**	**3 2**	3 3	**3 4**	**3 5**	3 6

Table 8: Partial and cumulative efficiencies (%) for signal points 4,5,6

4			5			6		
L1 L2	partial	cumulative	L1 L2	partial	cumulative	L1 L2	partial	cumulative
3 4	60.0	60.0	3 4	54.9	54.9	3 4	28.8	28.8
3 5	46.7	61.7	3 5	47.0	57.7	3 5	25.6	31.7
3 1	45.8	67.7	3 1	39.2	62.7	3 2	19.8	36.7
3 2	30.4	70.7	3 2	31.6	66.1	2 2	19.8	36.9
2 2	30.3	70.8	2 2	31.3	66.2	2 6	17.8	40.9
2 6	17.6	71.9	2 6	20.6	68.3	3 1	14.1	41.7
2 3	8.2	72.8	2 3	9.7	69.9	2 3	10.3	45.6
All 18 combinations :		73.0	All 18 combinations :		69.9	All 18 combinations :		45.7

9 Conclusions

We used a simple genetic algorithm to optimize the cuts for low mass-scale SUSY signal at low luminosity in CMS. Two simple jet cuts already provide high efficiency for all types of probing signals at L1. For soft E_T^{miss} + jets signatures, like invisible Higgs, an additional loose cut on the leading jet combined with E_T^{miss} is considered.

Unfortunately, at L2 R-parity violation scenario yields fairly low efficiency, though it is difficult to say a priori

Table 9: Partial and cumulative efficiencies (%) for signal points 4R,5R,6R

4R			5R			6R		
L1 L2	partial	cumulative	L1 L2	partial	cumulative	L1 L2	partial	cumulative
3 2	16.0	16.0	2 3	15.4	15.4	2 3	8.4	8.4
2 2	16.0	16.1	2 6	11.1	21.9	2 6	5.9	12.0
2 6	14.7	23.0	2 2	9.6	25.0	3 2	4.5	13.7
2 3	14.6	30.4	3 2	9.6	25.1	2 2	4.4	13.8
3 5	14.6	32.6	3 5	8.8	26.1	3 5	4.2	14.4
3 4	13.8	33.6	3 4	8.3	26.5	3 4	3.6	14.8
3 1	5.6	34.0	3 1	3.2	26.7	3 1	1.4	15.0
All 18 combinations :		34.0	All 18 combinations :		26.7	All 18 combinations :		15.0

Table 10: Partial and cumulative rates (Hz) for background processes

QCD			$t\bar{t}$			Wj		
L1 L2	partial	cumulative	L1 L2	partial	cumulative	L1 L2	partial	cumulative
2 3	1.12	1.12	3 4	0.08	0.08	3 4	0.16	0.16
3 2	0.51	1.56	3 5	0.04	0.09	3 1	0.04	0.17
2 2	0.50	1.56	3 1	0.02	0.09	3 5	0.01	0.17
2 6	0.46	1.81	2 6	0.02	0.10		-	
3 4	0.42	2.03	3 2	0.02	0.11		-	
3 5	0.36	2.09	2 2	0.02	0.11		-	
3 1	0.29	2.30	2 3	0.01	0.11		-	
All 18 combinations :		2.37	All 18 combinations :		0.12	All 18 combinations :		0.17

if what we get is enough for such signal observability in off-line analysis. It requires additional study. R-parity conservation scenario provides very high signal efficiency at L2 even with relatively tight requirements on the total rate of 3 Hz dedicated to jets + E_T^{miss} channels.

10 Acknowledgements

The author would like to thank Nikita Stepanov for the idea and some hints how to write and use genetic algorithm and Sarah Eno for useful discussions.

References

[1] S. Abdullin, F. Charles, Nucl.Phys.**B547** (1999) 60-81.

[2] S. Abel *et al.,* Report of the SUGRA Working Group for Run II of the Tevatron, hep-ph/0003154.

[3] H. Baer *et al.,* Phys.Rew.**D58**:075008, 1998.

[4] M. Battaglia *et al.,* hep-ph/0106204.

[5] T. Sjöstrand, *Computer Physics Commun.* 39 (1986) 347; T. Sjöstrand and M. Bengtsson, *Computer Physics Commun.* 43 (1987) 367; H.U. Bengtsson and T. Sjöstrand, *Computer Physics Commun.* 46 (1987) 43; T. Sjöstrand, CERN-TH.7112/93.
See also PYTHIA home page at http://www.thep.lu.se/ torbjorn/Pythia.html .

[6] P. Hidas, A. Nikitenko, S. Kunori, H.-P. Wellisch, CMS IN-2001/044.

[7] H.-P. Wellisch, CMS Note 1999/034.

[8] http://computing.fnal.gov/cms/jpg/fall20012e33/

[9] http://sl.web.cern.ch/SL/sli/Cycles.htm

[10] J. G. Branson, E. Trepagnier, CMS IN 2000/038.

[11] P. Hidas, S. Kunori, CMS IN 2002/015.

[12] http://cmsdoc.cern.ch/cmsim/cmsim.html

[13] http://www-thphys.physics.ox.ac.uk/users/PeterRichardson/HERWIG/isawig.html

[14] F. Paige and S. Protopopescu, in *Supercollider Physics*, p. 41, ed. D. Soper (World Scientific, 1986); H. Baer, F. Paige, S. Protopopescu and X. Tata, in *Proceedings of the Workshop on Physics at Current Accelerators and Supercolliders*, ed. J. Hewett, A. White and D. Zeppenfeld (Argonne National Laboratory, 1993).
See also F. Paige home page at http://paige.home.cern.ch/paige/ .

[15] G. Corcella *et al.*, hep-ph/0201201.

[16] http://cmsdoc.cern.ch/orca/

[17] P. Chumney, S. Dasu, F. di Lodovico, W. H. Smith, CMS IN-2002/019.

[18] P. Sphicas, talk at PRS joint/plenary session, CPT week, April 19, 2002.

[19] D.Cavalli *et al.*, hep-ph/0203056.

[20] CERN Program Library Long Writeup D506, http://consult.cern.ch/writeup/minuit/.

[21] John H. Holland. *Adaptation in natural and artificial systems*. The University of Michigan Press, Ann Arbor, 1975.

[22] David E. Goldberg. *Genetic algorithms in search, optimization and machine learning*. Addison Wesley, 1989.

[23] L. Eshelman. The CHC Adaptive search Algorithm. *Foundations of Genetic Algorithms*, G. Rawlings, ed. Morgan-Kaufmann (1991) 256-283.

[24] D. Whitley. The GENITOR Algorithm and Selective Pressure. *Proc. 3rd Int. Conf. on Genetic Algorithms*, Morgan-Kaufmann (1989) 116-121.