Studying The Correlation In The So-called Event Mixing Method.

Xing Wang, Peking University, China

September 4, 2013

Abstract

The event mixing method is a background estimation method. For the study CMS Monte Carlo Simulation is used as an environment to easily generate Monte Carlo events and to write a toy analysis. The analysis that is used is a search for supersymmetry in the invariant mass distribution of opposite-sign lepton pairs. Leptons from the decay chains of supersymmetric particles will create a so-called mass edge in the invariant mass spectrum. The event mixing creates an estimation (a histogram) for the shape of the background but the errors of this estimate are correlated. These correlations have to be understood in order to quantify the statistical relevance of any deviation.

Contents

1 Introduction

There are many standard model process that also produce lepton pairs as in some Supersymmetry (SUSY) decays. Especially $t\bar{t}$ events or events with a Z_0 boson decay. When we look for supersymmetry we typically try to get rid of standard model events by requiring that the event has a lot of missing energy. Supersymmetry includes typically a massive, neutral stable particle. It is the lightest supersymmetric particle i.e. LSP, which could be a dark matter candidate. The LSP cannot be observed in the detector, therefore we have missing energy. Since we do not exactly know the amount of energy in a proton proton collision we only look in the transverse momentum. We use the missing transverse energy (MET) as a hint. Events with $Z_0 \to l^+ + l^-$ do not have much MET but $t\bar{t}$ events do when the top quarks decay leptonically $t \to b + W$ with $W \to l + \nu$ The events we see are typically $t\bar{t}$ decay events, not SUSY. We have to know how many $t\bar{t}$ events we see in the data to be able to say if there is an excess due to a new particle. For this we need the data driven methods i.e. we use in some way the data to predict what we see (the background) in the data. Sounds strange but is useful since we do not trust the Monte Carlo programs in all detail.

But in this report, I'll not use this method on real data. I just use it as a technique on the Monte Carlo Simulation and study the bin correlations in bins. The result may be helpful to statistics (I hope so) .

This report is organized in this way: First in the Introduction Section I'll briefly go through the Standard Model and it's shortcomings. After that it is easier to give a brief introduction to SUSY. And the introduction to the event mixing method is followed. Second, I will give the results and the corresponding analysis of the results. And finally, I'll conclude and outlook.

2 Review of the theory

2.1 Review of the Standard Model

2.1.1 Introduction to the SM

Fundamentally, the SM is nothing but an equation:
Gauge Invariance+Spontaneous Symmetry Breaking(Higgs Mechanics)→SM.
i.e.

$$SM = U(1) \otimes SU_L(2) \otimes SU(3). \tag{1}$$

Of course, it must be made of particle content, assuming leptons and quarks are the elementary particles. We believe that symmetry is the most important concept in modern physics. First we have the fundamental matter field (or the fermionic field) ψ, and the Lagrangian $\mathcal{L}$ made of ψ and $\partial_\mu \psi$. Then we assume the Lagrangian $\mathcal{L}$ is invariant under gauge transformation, then we must introduce the covariant derivative D_μ where

contains a new field, the gauge field $\mathcal{A}_\mu^a$, in place of ∂_μ, here a is just extra degree of freedom. For abelian case a is zero, for nonabelian case a is nonzero.
The $U(1) \otimes SU_L(2)$ is just the electroweak theory which unifies the electromagnetism and weak interaction. The $SU(3)$ is Quantum Qhromodynamics or QCD.
In the particle representation, fundamental matter field's quanta is fundamental fermion to make up the matter, e.g. electron, quarks and the gauge field's quanta is fundamental boson to carry force, e.g. photon , gluon. And another important part is the "God particle", the Higgs Boson which gives mass to fundamental particles. LHC has already found a Higgs-like boson.

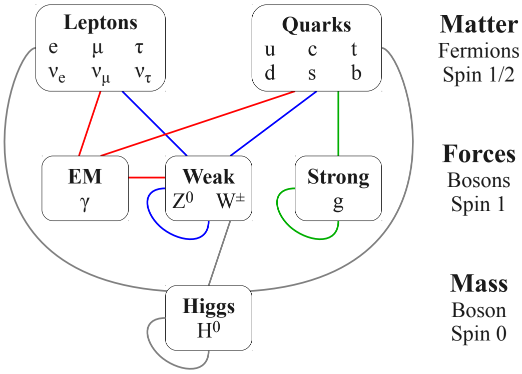

Figure 1: Review of the SM

2.1.2 Shortcomings of the SM

The SM perhaps is the most experimentally successful and theoretically clear model. But fundamentally, there are two most important things about the theory today. One is inside and the other is outside.
Inside, the SM has too many free parameters, about 26. A fundamental theory is not excepted in this way. And there are many fine-tuning in the SM. Just imagine: We see a pencil standing on its tip in the middle of a table. While this scenario is not impossible, if we were confronted with this sight we would seek an explanation, looking for some mechanism that stabilizes the pencil and prevents it from falling over. For instance, we might look to see if the pencil is secretly hanging from a string attached to the ceiling.
Outside, we are already sure that the SM is not a final theory. It's just a small visible part of the huge iceberg! Because of gravity, the SM cries at the Planck scale and the fundamental concept in the SM must be changed very much. More concretely, the SM has many problems, both experimentally and theoretically, e.g. dark matter, neutrino mass, hierarchy problems and fine-tuning problems, etc. New physics is coming but we don't know exactly the scale where it must be counted neither what the new physics looks like.
However, we are sure that a brand new physics theory must appear at so-called Planck

scale. Why? Because at Planck scale, gravity comes in and it almost disturb all. On the one hand because of the uncertainty principle, we have to use high energies to probe short distances. In a world without gravity, we could resolve arbitrarily small distances in this way, but gravity eventually and dramatically changes the picture. At minuscule distances, so much energy has to be concentrated into such a tiny region of space that the region itself collapses into a black hole, making it impossible to extract any information from the experiment. This occurs when we attempt to probe distances around 10^{-33} cm. One of the two pillars of quantum field theory that interactions take place at points don't make sense any more because of the space-time identity above. As a consequence, we don't have a meaningful measurement. On the other hand, classically, both the gravitational and electric forces vary with distance following an inverse-square law; however, at a distance of around 10^{-11} cm, this gets corrected in an important way: again because of the uncertainty principle, simply holding two electrons at shorter distances requires a huge amount of energy. The force of gravity increases with increasing mass, or with equivalently increasing energy, so the attraction between electrons begins to increase relative to the electrical repulsion. At around 10^{-31} cm, gravity surpasses the electric force, and at 10^{-33} cm, it dominates all interactions.

2.2 Introducion to SUSY

Supersymmetry is the only allowed non-trivial combination of space-time and internal symmetries.

2.2.1 Problems of the Standard Model and the Motivation of SUSY

- Dark matter and dark energy
  - We only know about 5% of our world, which the SM decribels very well, how terrible!
- Gravitation
  - An obvious problem of the Standard Model is that it provides no description of gravity.
  - Theory with gravity will be totally different from the SM now because of the most fundamental measurement even changing.
- Hierarchy problems and fine-tuning problems
  - The biggest hierarchy problem: We have found a huge contribution to the cosmological constant from quantum fluctuations, this quantum effect is 10^{120} bigger than the observers cosmological constant. So perhaps there is also a purely classical part of the cosmological constant, whose size just so happens to delicately cancel the contributions from quantum fluctuations, to an accuracy of 120 decimal places. This is a deeply unsatisfying explanation, and for

obvious reasons is referred to as unnatural fine-tuning of the parameters of the theory. The fine-tuning needed to understand why we have a big universe is also known as the cosmological constant problem.

– There is also another hierarchy problem, related to the question of why atomic scales are so much larger than the Planck length. The relatively large size of the atom is a consequence of the small mass of the electron. Or more generally, we have the gauge hierarchy, that is , there is a so big and unnatural gap between different gauge couplings.

– As briefly reviewed above, an electron acquires its mass from interacting with the Higgs field, with a typical interaction length near 10^{-17} cm. But the Higgs field itself should have enormous quantum fluctuations growing stronger toward the Planck scale, and so the typical length scale of its interactions with an electron should be closer to 10^{-33} cm. This outcome would make electrons sixteen orders of magnitude heavier than they are observed to be. To avoid this conclusion, we have to invoke another unnatural fine-tuning in the parameters of the theory, this time to an accuracy of one part in 10^{30} . And there are also some other hierarchy, for example, an extreme fine-tuning of the Higgs mass parameter.[1]

– How to explain the big hierarchy between three generations of fermions, i.e. leptons and quarks?

- Grand Unification

  – Unification is our dream. As we know that electromagnetism and weak interaction is unified as electro-weak theory. At some high scale, they are the same. Mathematically , by calculating the β functions of the two interactions, we find that the couplings will be the same at some scale. However, only in the SM, we can't unify the strong interaction.

Figure 2: Inclusion of gravity

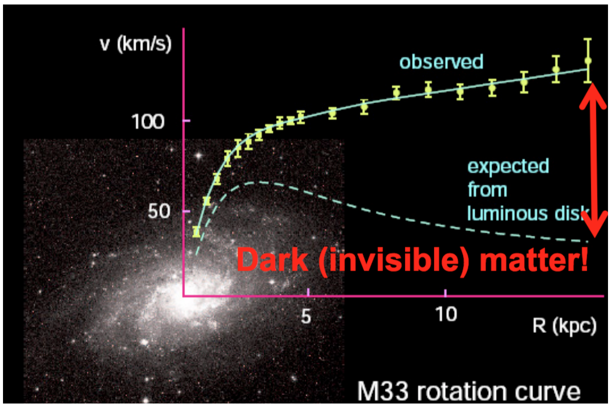

Figure 3: Dark matter

2.2.2 How about SUSY

As far as I know, SUSY is the most promising extension to the SM for both theoretical and experimental reasons. SUSY is a extension of our space-time and it's a space-time internal symmetry. Or more concretely, in the 4-dimension space-time representation, SUSY is a symmetry between bosons and fermions. Each particle of the Standard Model becomes part of a supermultiplet together with new particles called superpartners. Supersymmetry requires now these particles to have identical quantum numbers with exception of the spin that differs by 1/2. This means that also the masses of the particles in each supermultiplet should be the same. Obviously this is not the case since light superpartners with the same couplings as the known Standard-Model particles would have been discovered easily. The symmetry must be somehow broken. In many SUSY models, R-parity is conserved. The experimental lower limit on the proton lifetime is of order of $10^{3}0$ years. This is a strong argument why baryon number should also be conserved in Supersymmetry. To achieve this a new quantum number is introduced: the R-parity defined by

$$R = (-1)^{3B+L+2S} \tag{2}$$

where B is the baryon number, L the lepton number, and S the spin. All Standard Model particles have R = 1 while all SUSY particles have R = ?1. The conservation of R leads to a stable proton and to a stable LSP.[2]

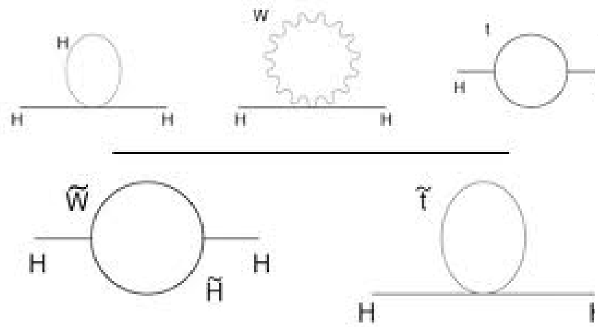

Figure 4: Higgs Mass Hierarchy solved after SUSY

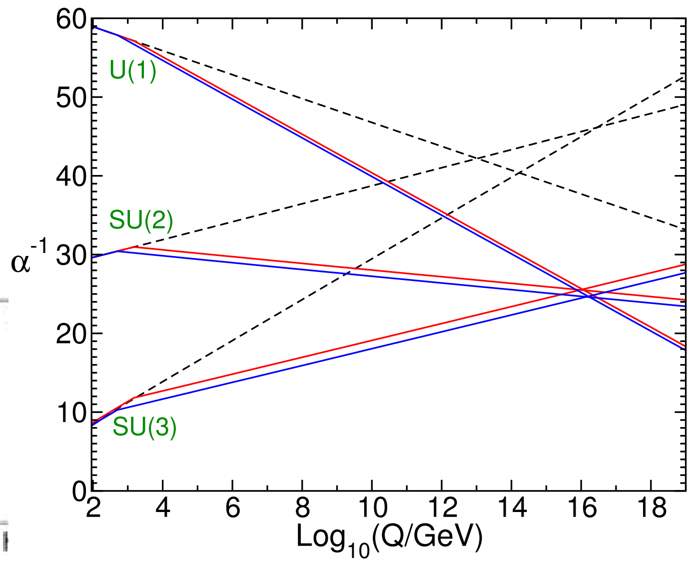

Figure 5: Unification before and after SUSY

- If R-parity is conserved, SUSY provides a perfect Dark-Matter candidate, the LSP.
- Local SUSY requires inclusion of gravity.

- SUSY relieves the hierarchy very much. For example, after including SUSY particles' cancellation, the fine-tuning of the cosmological constant can be suppressed from 10^{120} to 10^{60}. And SUSY can solve the extreme fine-tuning of the Higgs mass parameter due to stop's contribution.

2.2.3 Particle Content of SUSY

The number of new particles predicted by a supersymmetric extension depends on the actual model. The only constrain given is that the number of fermionic and the number of bosonic degrees of freedom in each supermultiplet must be the same.So-called minimal supersymmetric extension of the Standard Models (MSSM) introduce the minimal amount of the particles: for each known particle the least possible number of superpartners. For each spin-1/2 fermion of the Standard Model two scalars are introduced, each

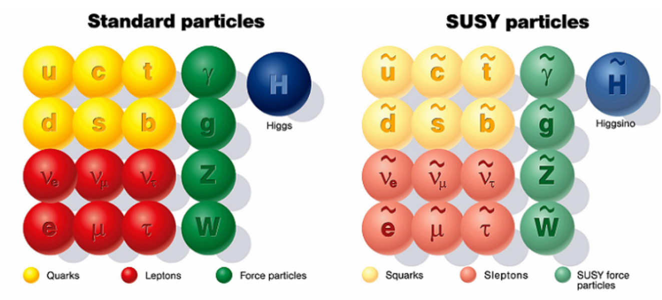

Figure 6: SUSY Particle

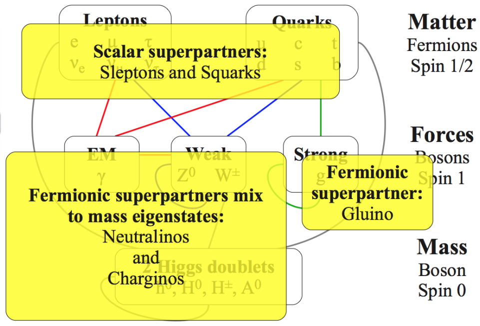

Figure 7: SUSY Particle

as partner of one helicity state of the fermion. Each gauge boson is accompanied by a supersymmetric fermion.
In supersymmetric models the Higgs sector is slightly modified. The Standard Model Higgs doublet would lead to triangular anomalies. The fermionic superpartner of the Higgs boson would also contribute to the loop, and hence cause a divergence. Therefore a second Higgs doublet has to be introduced. In order to restore the cancellation the two Higgs supermultiplets must have opposite hypercharge. Moreover, in a supersymmetric environment two Higgs doublets a necessary, one generating the masses of up-type quarks, the other of down-type quarks and charged leptons. From the two complex Higgs doublets five Higgs bosons are expected, labeled as h_0, H_0, $H_\pm$,and A_0.
SUSY mass eigenstates are kind of different from the SM sector, just as the picture shows below.

2.2.4 Breaking Mechanism of SUSY

But we have not seen any SUSY particle signal. So if SUSY is true, the particles are very heavy! How? Again the spontaneous symmetry breaking. Anyway there are many different ways to make it. So there is not only one SUSY model.

2.2.5 Searches for SUSY and Signal Scenarios

One of the motivations of LHC is to find SUSY. Generally, the supersymmetric particles are expected to be heavier than the Standard Model particles. In a collider experiment they can be produced if the center-of-mass energy is high enough and the couplings reasonable.
Here I'm working on the search for SUSY in di-lepton mass spectrum.

- New quantum number R_P distinguishes between SM and SUSY particles;
- R_P conserved in many models $\rightarrow$ SUSY particles produced in pairs, the lightest SUSY particle (LSP) is stable. And then it can be a dark matter candidate; So a typical SUSY dilepton signal cascade is like the picture. However, the $t\bar{t}$ decay events have the same decay features experimentally, and this is the main background. We have to get the shape of the $t\bar{t}$ background.

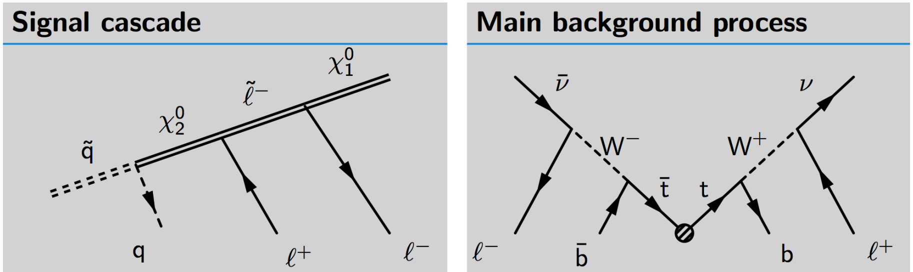

Figure 8: SUSY decay and $t\bar{t}$ background

- Di-lepton mass is constrained by involved particle masses.
- Event selections to $t\bar{t}$ events
  1. at least one $\mu^+\mu^-$ or e^+e^- pair:
     Defines the signal, rejects QCD processes, which are by far the most events at a hadron collider.
     a) p_T >20GeV,10GeV;
     b) $|\eta| < 2.4$;
     c) $R(l^+l^-) > 0.4$;
  2. $E_T^{miss} > 100$ GeV:
     Missing transverse energy expected from LSPs, rejects mainly Drell-Yan events.
  3. at least 2 jets: for generic search:
     Rejects further background events, also processes with real E_T^{miss}, such as di-boson events.
     a) p_T >40GeV;
     b) $|\eta| < 2.4$;
     c) $R(jets + l) > 0.4$;

2.3 The Event Mixing Method

Event mixing is a novel data-driven technique to decorrelate observables that can be combined to one quantity. A typical example is the kinematic correlation between particles that originate from the decay of one on-shell intermediate state. The Lorentz vectors are the observables, the two-particle mass is the combined quantity.
The correlation between the Lorentz vectors is measured by comparing the two-particle invariant mass spectrum with a second artificial mass spectrum that is obtained from taking one particle from one event and the other particle from another event. These cross-event particle combinations are called mixed events.
We can just pause to plot how to get mixing event:

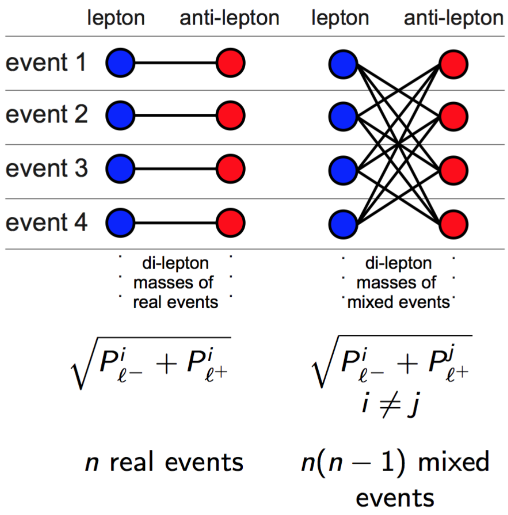

Figure 9: Mixing events

1. Combine the lepton of one event with the anti-lepton of another event to calculate the di-lepton mass;
2. Any correlation is broken;
3. The shape of the uncorrelated background is reproduced.

Assuming a correlated signal and an uncorrelated background, by comparing the real two- particle mass distribution with the mixed one, the background distribution is reproduced by the mixed events, and the signal contribution within the event sample should become visible. The difference between the spectra can be quantified with shape comparison methods like maximum likelihood ratio, χ^2, or Kolmogorov test.

2.3.1 Motivation

The event mixing method can be used to estimate the background formed from $t\bar{t}$ events in SUSY analysis. $t\bar{t}$ events are the most background, we must use MC simulation to generate it and then extract it from the data to see if there is a signal for SUSY. Although it is a data-driven technique, I will apply this method on MC simulation.

2.3.2 Is the event mixing method trusted?

As you can see in Figure 8, the two leptons in the $t\bar{t}$ events are from different decay branches. And there are many decay vertices. So it seems that these leptons don't have correlation on the Lorentz momentum, so that we can just do the mixing. Here comes a big problem! Is the physics correlation effect negligible? Not necessarily!

2.4 The physics correlation

There is physics correlations in $t\bar{t}$ events caused by common production:

- Common boost of $t\bar{t}$ system causes positive correlation in η;
  - In many events the whole $t\bar{t}$ system is boosted in z direction (i. e. along the beam axis), caused by an imbalance of energies of the colliding partons. In transverse direction no net boost of the whole event is expected, but the $t\bar{t}$ system can still be boosted if it is recoiled by initial-state radiation. However, the main direction of the boost is along the beam axis.
    Of course, such a boost induces a lepton-lepton correlation: The $t\bar{t}$ boost is inherited by all daughter particles including the lepton, causing a correlation. The two leptons tend to be in the same forward- backward hemisphere.
    The actual effect on the mixing can be described from a kinematic point of view: Assuming a scenario with only highly boosted $t\bar{t}$ event, the leptons are typically high energetic and close together. The events mixing results in a collection of mixed events with, of course, also high energetic leptons but no angular relation between them. Since the two-particle mass is larger for higher energies and smaller for smaller angles the spectrum of the mixed events are shifted to higher mass w. r. t. the spectrum of the real events.
    In another notation of the di-lepton mass the effects of angular and energetic properties are more clear:

$$M_{ll} = \sqrt{2E_{l^+}E_{l^-}(1 - \cos\theta_{ll})}, \tag{3}$$

    The larger the boost, the larger the energy, but the smaller the opening angle. They compensate exactly and lead to the well known invariance. But, of course, only in the real events and not in the mixed events.

- Di-top mass leads to a back-to-back structure of the event, visible in ϕ correlation.
  - A second reason for a correlation is found in the center-of-mass energy of the $t\bar{t}$ system. This energy can be interpreted as the di-top mass. Although there is no resonance in the mass spectrum, there is a certain structure in the center-of-mass frame of the hard interaction: At the lower end there is a threshold

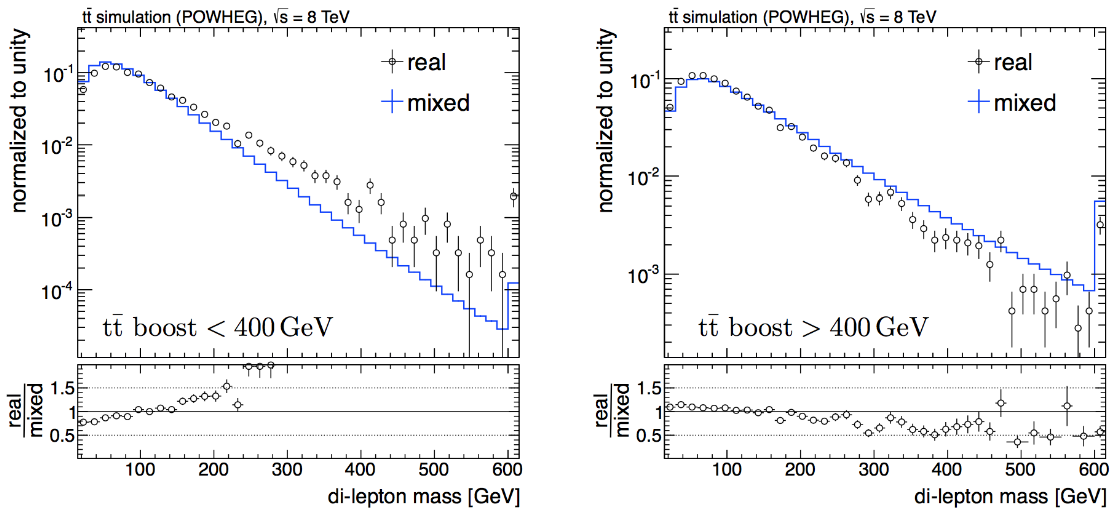

Figure 10: The boost correlation(from Hannies)

at two times the top mass, the upper tail is governed by the typically reached energies in the parton-parton collisions (depending on the proton energies and the parton density functions). The leptons as descendants of the tops carry parts of their kinematics and therefore remnants of the di-top mass. Since both leptons in an event are affected by the same di-top mass, a correlation is induced. In the mixed events the correlation is broken and the spectrum of the di-lepton mass is generally expected to differ from the spectrum in the real events.

The other point of view is again purely kinematic: The higher the energy in $t\bar{t}$rest frame the stronger the back-to-back topology of the whole event is realized. In events with two high energetic back-to-back tops, the leptons inherit the kinetic energy and have preferably also larger opening angles.

Actually, there are other correlations, e.g. the $t\bar{t}$ spin correlation, etc. It is found that these effects are small.

But luckily, some study states that these two kinds of correlations just compensate for each other, so the net effect is negligible! So we can use the event mixing method safely! [2]

2.5 Study the bins correlation in the event mixing method

Although there is almost no physics correlations net effect in the event mixing method, there may be bins correlations. That is the topic I'm studying.

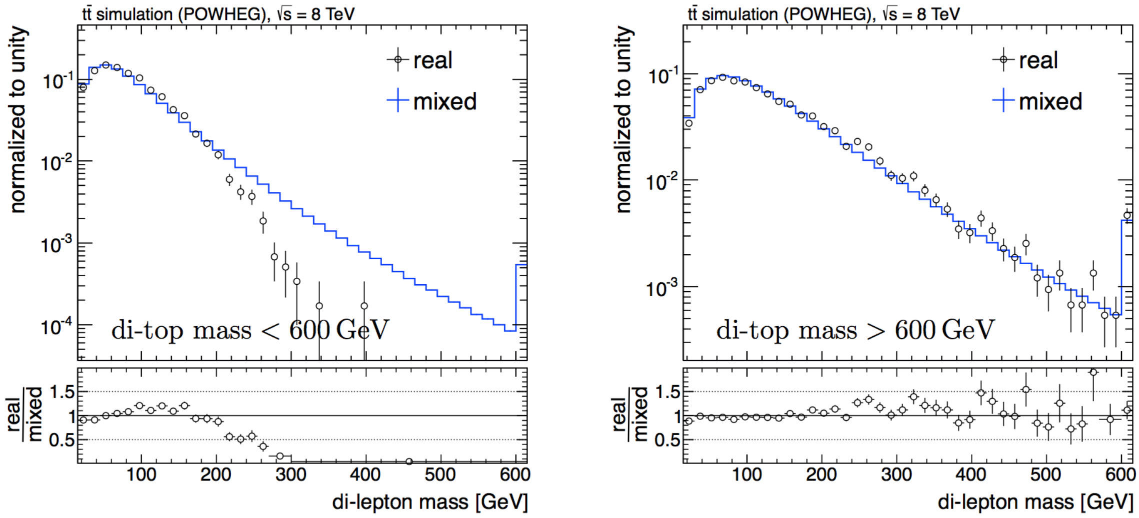

Figure 11: The do-top mass correlation (from Hannies)

3 My work on the study

3.1 Review on my study

- Generate the MC events
- Use CMS Monte Carlo Simulation to write a toy analysis
- The event mixing creates an estimation for the shape of the background but the errors of this estimate are correlated!!
- Do the χ^2 test and other tests
- Analyse the errors

3.2 $t\bar{t}$ background by pythia

At beginning, I used Pythia to generate events and then used Rivet to analyze the events to get the selected data. But the shape is not well described by this simulation, so I decided to change the MC simulation program to Madgraph.

Figure 12 is the result of pythia: χ^2=148.378762; Prob = 3.81549×10^{-6}; NDF=79; igood=2.
The result is just not so good. At least worse than that of madgaph. So next I'll just turn to madgraph case. Pythia is only first-order simulation and use Rivet only. I will only study the bins correlation in Madgraph+all detector simulation method.

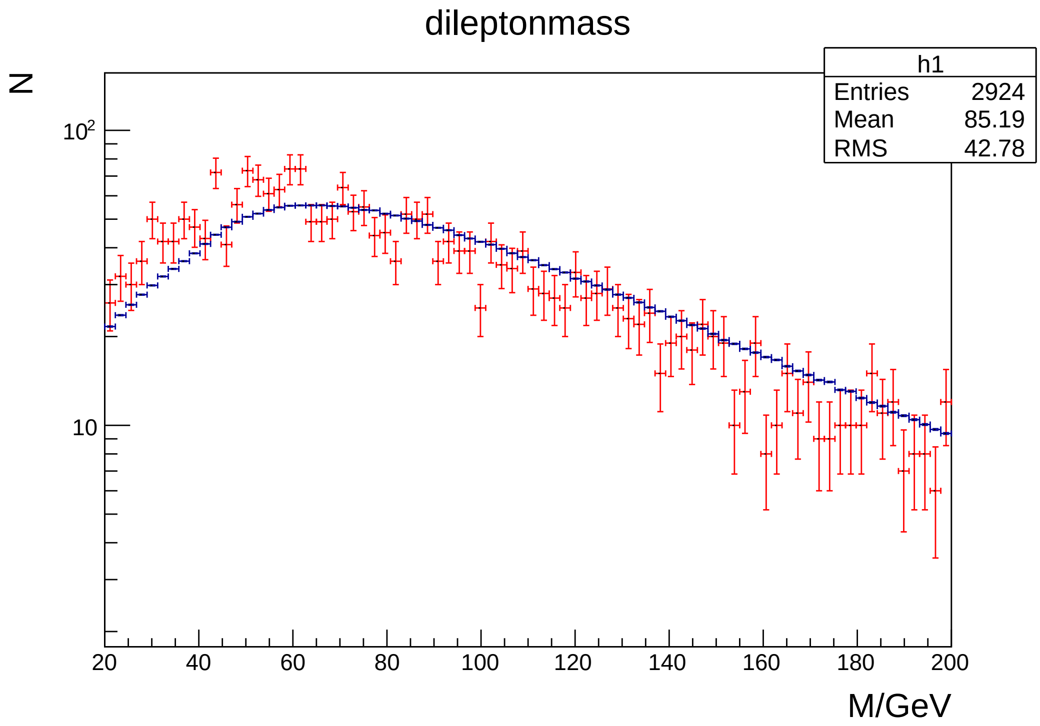

Figure 12: result of pythia

3.3 $t\bar{t}$ background by madgraph

I do the χ^2test like Figure 14 and the result is below.

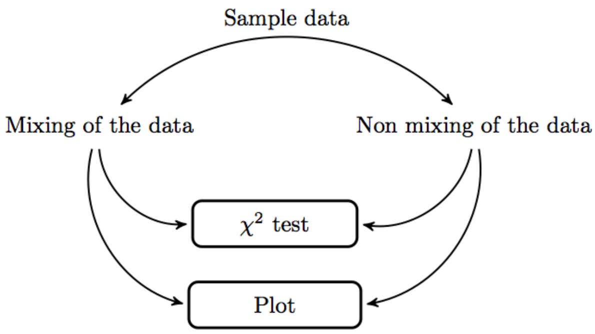

Figure 13: test work flow

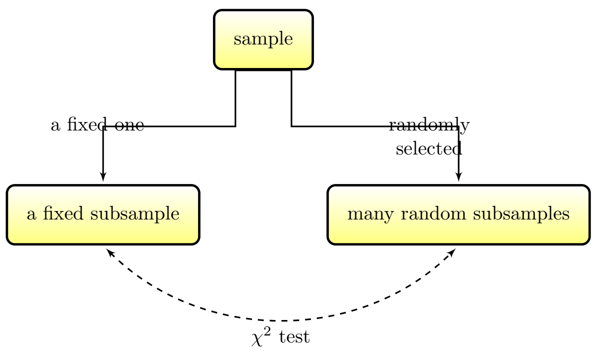

Figure 14: advanced test work flow

The sample is just the lepton pair data get from madgraph after adding the cuts. Firuge15 is the plot of the madgraph case. The χ^2 test result is $\chi^2 = 102.095969, Prob = 0.0413093, NDF = 79, igood = 2$. But the result is not really good. It should not be so different!

To see more convincingly, we have to make some advanced tests about the sample.

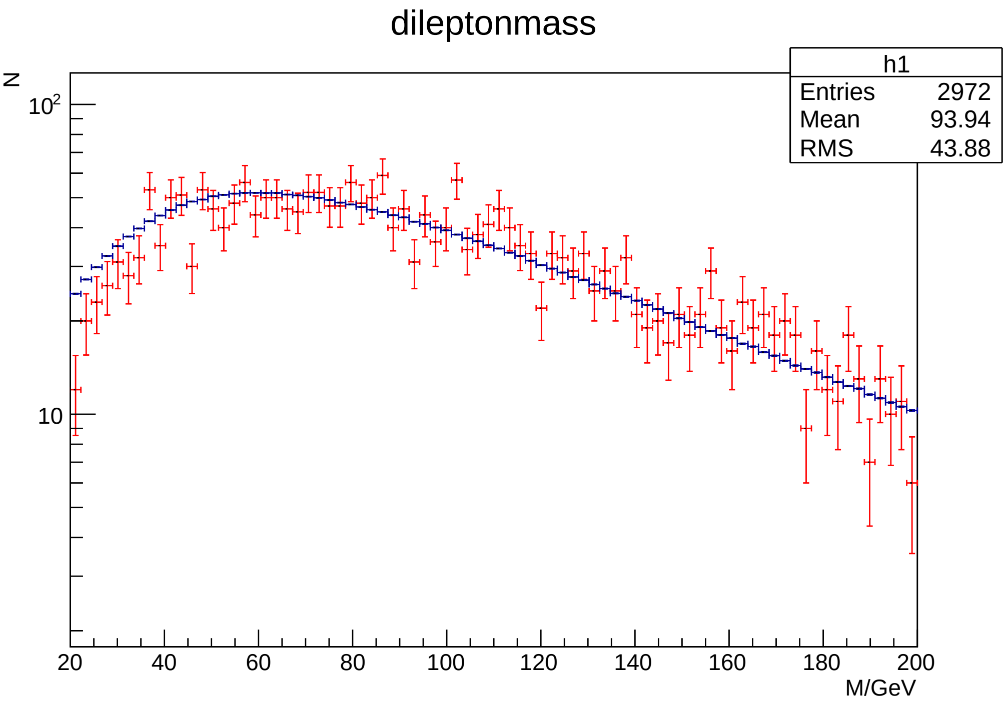

Figure 15: madgraph result

3.4 Advanced tests

3.4.1 Advanced tests results

The idea that how to do this is given in Figure 14. χ^2 tests between the first subsample(mixed or non-mixed) one with other mixed subsamples generate χ^2 distribution. And the results are just Figure 16, 17, 18 below.

3.4.2 Advanced tests analysis

The results are not good either. It is not probable that a histogram deviates randomly so much but the error is so small. As you can see from the advanced result1.
And we have a strange χ^2 with the probability. Just as the advanced result2 and result3 show. And the probability should be 1, but in the advanced result2 it fluctuates too much.

One of the problems is that we have the wrong errors after we do the mixing:

- The number of counts N in a bin is Poisson distributed so the error $\sigma \sim \frac{1}{\sqrt{N}}$
- If we Fill a histogram usually all the entries are independent then $\chi^2 \sim$ numbers of bins-1

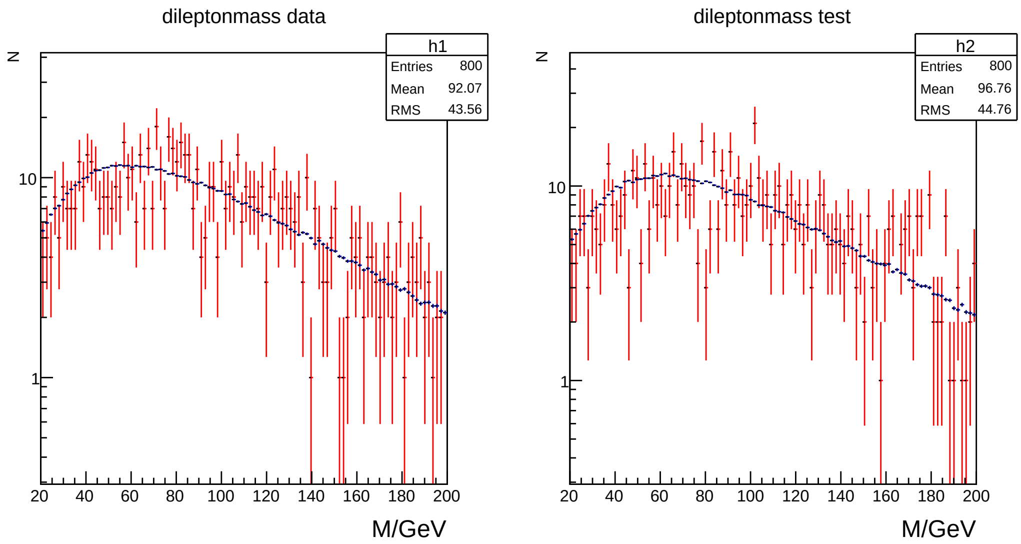

Figure 16: Left: The mixing of the fixed subsample χ^2 test with the non-mixing of itself; Right: The mixing of a random subsample χ^2 test with the non-mixing of itself.

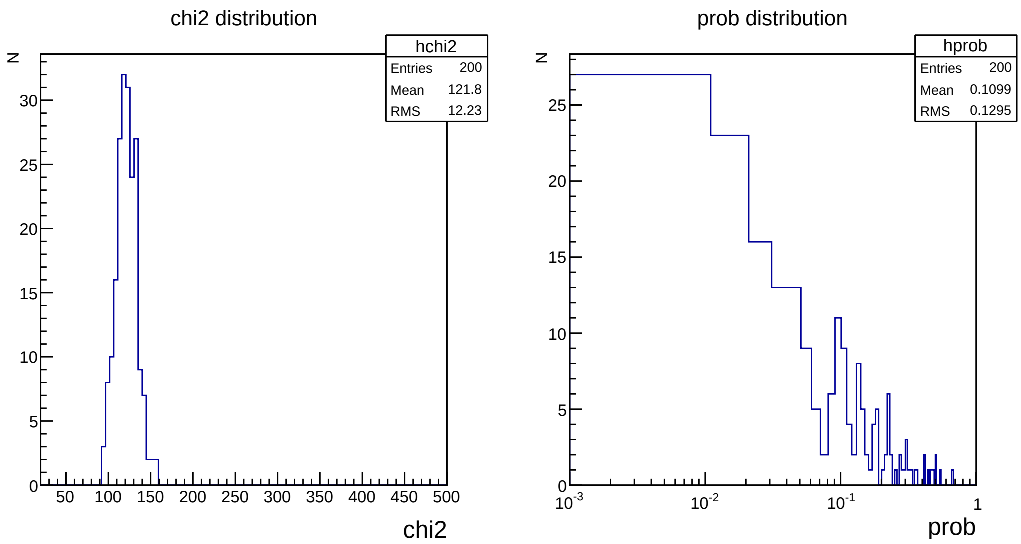

Figure 17: Left: The non-mixing of the fixed subsample χ^2 test with the mixing of the other subsamples Right: The non-mixing of the fixed subsample χ^2 test with the mixing of the other subsamples .

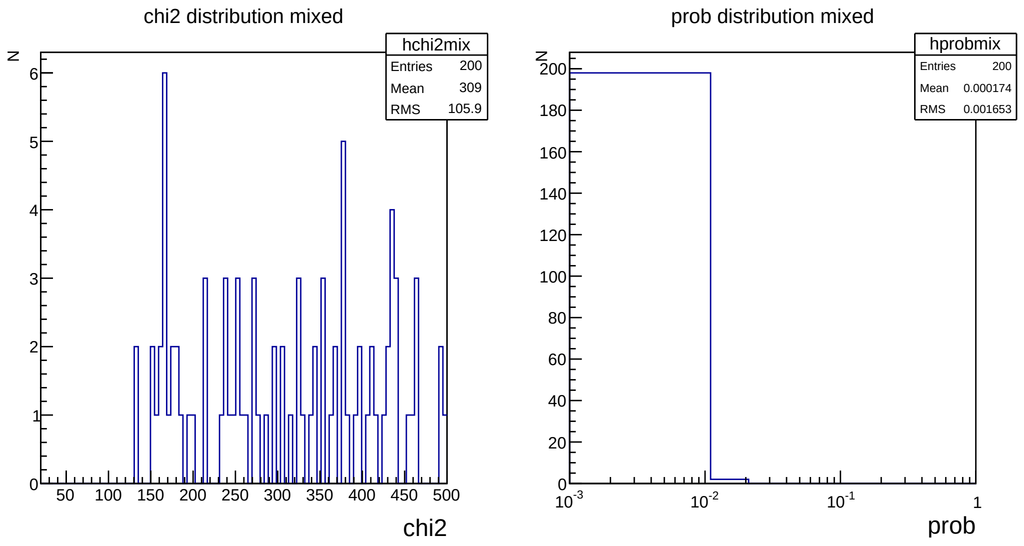

Figure 18: Left: The mixing of the fixed subsample χ^2 test with the non-mixing of itself; Right: The mixing of the fixed subsample χ^2 test with the mixing of the other subsamples.

- It become more complicated if we do the mixing. We are using the same leptons several times. For root it looks as we had $M = N(N-1)$ bins, so it thinks $\sigma \sim \frac{1}{\sqrt{M}}$ That is why the mixing error is so small.

So I have to get the correct errors!

3.5 The correlated errors

The correct error must have some statistic correlation so that it will not be so small as the above shows. Here I uses the so-called bootstrapping method to do this.
Different mixed histograms have on average the same shape but sometimes a bin will be larger and sometimes smaller. We look into the same bins in different histograms and we track these changes.(i.e. calculating the variance) Then we will get a much real error of the mixing events and the histogram i.e. the mean mixed histogram. And then we compare the mean mixed histogram with an example of the mixed subsets. We will see much more realistic error.
Figure 15 shows on the left an examplet of the mixed histogram. And the right one is the mean mixed histogram. As you can see in the first one , it fluctuates very much and somewhere the fluctuation is even bigger than the error (it is too small).
The mean mixed one looks much smoother because the error is correlated. Actually, what we see is the mean mixed distribution with an error according to the real fluctuation. The bins must move together up and down i.e. the different histogram bins are not independent. That comes from the mixing.

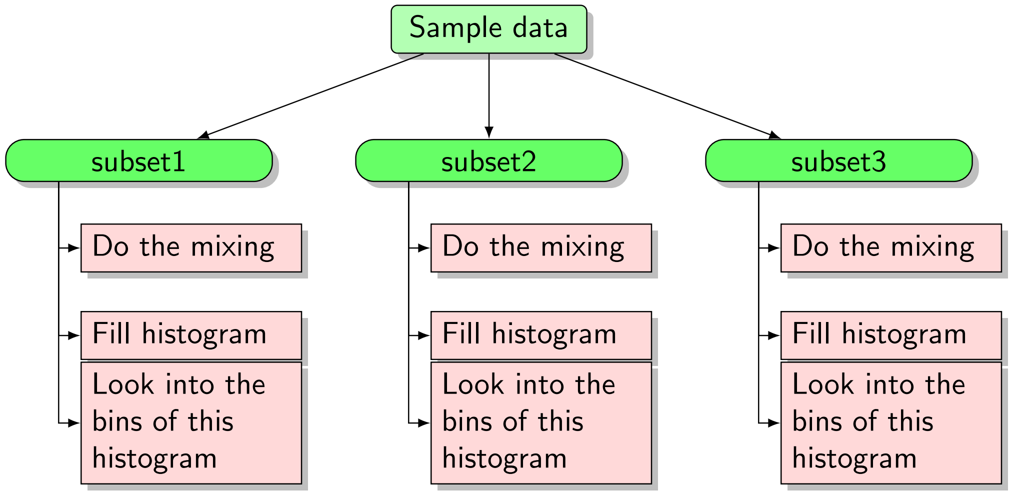

Figure 19: The example idea of this method

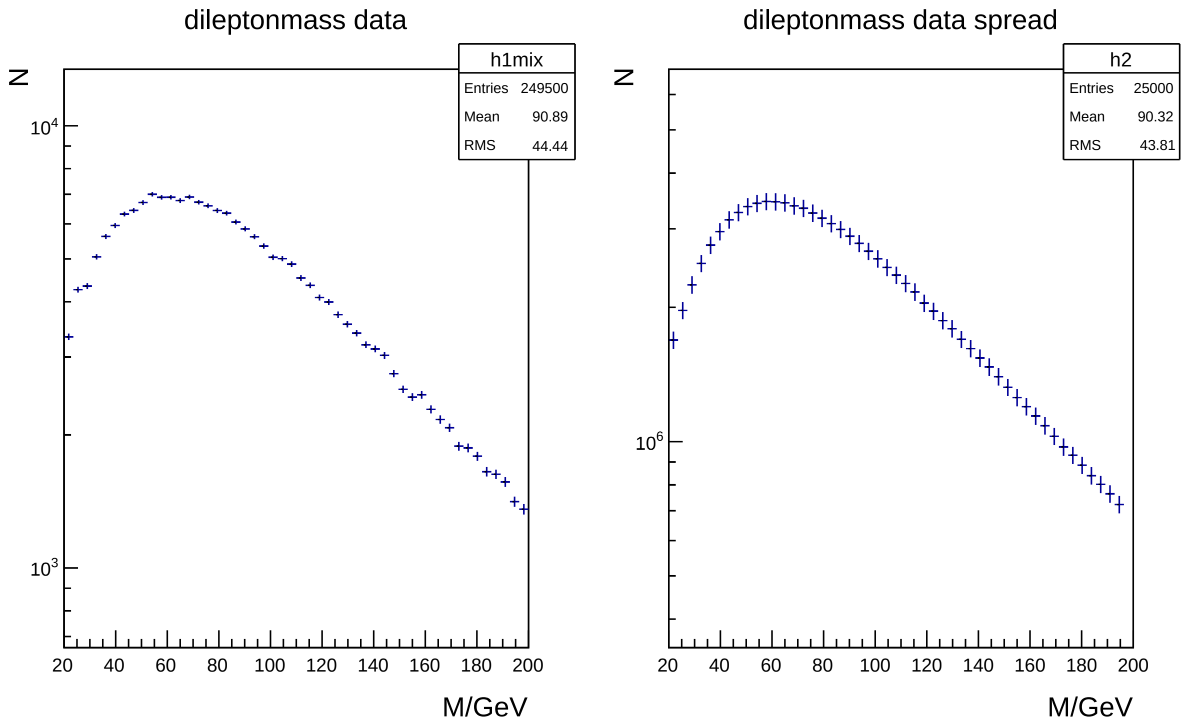

Figure 20: Bootstrapping result

4 Conclusion and outlook

- Large errors is only one part of the problem. The bins are correlated.
- And to treat such a case correctly, one has to take into account the correlation matrix not just the error, and it is still on the run.

Acknowledgements

Thanks to my supervisor Dirk Kruecker and Hannes Schettler's phD thesis. And of course, thanks to all the DESY CMS SUSY group members. You help me a lot.

References

[1] The Future of Fundamental Physics *Nima Arkani-Hamed*

[2] The Event-Mixing Technique for Modeling the $s\bar{t}$ Background in a Search for Supersymmetry in the Di-Lepton Channel *Hannes Schettler*