Kinematic reconstruction performance in $t\bar{t}$ and $t\bar{t}$ + DM production in dileptonic channel

Aliya Nigamova, MEPhI, Moscow, Russia

September 7, 2016

Abstract

To increase our chances of success in such non-trivial task as a search for dark matter(DM) the performance of the kinematic reconstruction of the $t\bar{t}$ system has been studied. Performance of kinematic reconstruction has been benchmarked on SM $t\bar{t}$ production, and it was shown that SM $t\bar{t}$ events are well reconstructed. The performance of the default kinematic reconstruction for $t\bar{t}$ + DM sample. A possible update of the kinematic reconstruction algorithm, designed to improve the performance on $t\bar{t}$ + DM samples, is presented.

Contents

1 Introduction **3**

2 Simplified model **3**

3 Data sample and event selection **4**

3.1 Simulation 4

3.2 Event selection 4

4 Kinematic reconstruction **5**

4.1 Description 5

4.2 Kinematic reconstruction performance check on SM $t\bar{t}$ 5

4.3 Attempt to apply default kinematic reconstruction on DM sample 7

5 Modified algorithm of the kinematic reconstruction **10**

5.1 Description 10

5.2 Comparison with the results obtained by default algorithm 10

6 Conclusion **18**

7 A. Simulation samples **19**

1 Introduction

A lot of astronomical observations propose the existence of dark matter, which is the substance that appears to interact with the ordinary matter via gravity and seems to make up more than a quarter of the total matter-energy in the universe. There are many DM models, but in this research we will consider a simplified model, which proposes a spin-0 interaction between DM and SM particles. The coupling to the SM particles is considered to be Yukawa which is proportional to the mass of the quark. Due to the fact that the top quark is the heaviest one, the search of new physics in association with top pair production seems to be reasonable. This report describes the study of kinematic reconstruction in $t\bar{t}$ and $t\bar{t}$ + DM production in the dilepton channel. The main disadvantage of the dilepton channel is the low statistics available. But leptons are accurately reconstructed, that gives us more accurate spin correlation information, which can be used in the search for sensitive variables.

2 Simplified model

In this section the brief description of simplified model is given, while more detailed description can be found in [3].
The simplified model of DM proposes two mediator particles: scalar (S) and pseudo-scalar (PS), and two DM fermion particles. The model can be characterized by four parameters: the mass of the DM fermion (Mχ), the mass of the mediator (Mϕ), the coupling strength of the mediator to the SM fermion and the coupling strength of the mediator to the DM fermion.
In Figure 1 a Feynman diagram describing this process is shown.

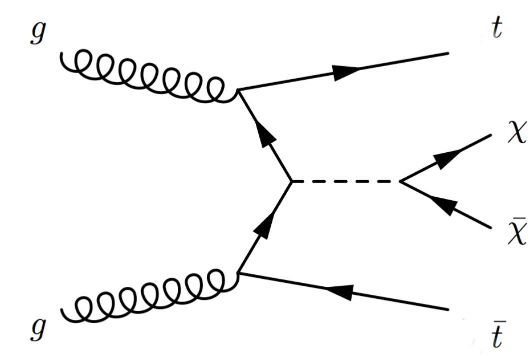

Figure 1: Feynman diagram illustrating the simplified $t\bar{t}$ + DM model

3 Data sample and event selection

This analysis is performed in the CMS framework. A full description of the CMS detector can be found here [1].

3.1 Simulation

This analysis is based on the MC simulation which was produced by the CMS MC group. A list of used simulation samples can be found in appendix A.

3.2 Event selection

In this analysis we are interested in $t\bar{t}$ pair production in the dilepton channel which is illustrated in Figure 2. In the final state we have 2 oppositely-charged leptons, 2 b-quarks and 2 undetected neutrinos, events with τ-> $l\nu_l\nu_\tau$ decays are not considered as a signal.
In the search for DM this process will be considered as a background and $t\bar{t}$ + DM will be considered as a signal process.

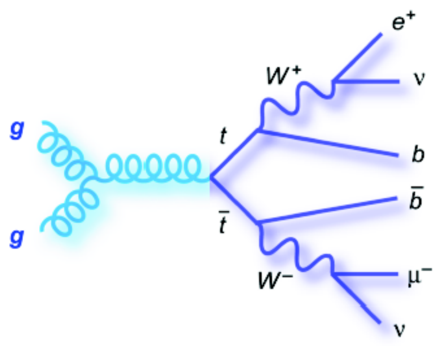

Figure 2: Dilepton channel t$\bar{t}$ decay

The main constraints:

- 2 leptons: events with two oppositely-charged leptons, m_{ll} >20 GeV, and events with 76 GeV < m_{ll} < 106 GeV are rejected for e^+e^- and $\mu^+\mu^-$ channels (to remove background from Z+jets)
- at least 2 jets and at least one of them should be b-jet
- E_t^{miss} > 40 GeV(for e^+e^- and $\mu^+\mu^-$ channels)

The missing transverse energy (MET) E_t^{miss} is defined as the magnitude of the projection on the plane perpendicular to the beam, of the negative sum of momenta of all reconstructed particles.

4 Kinematic reconstruction

4.1 Description

One of the goals of this project is to test the $t\bar{t}$ kinematic reconstruction on $t\bar{t}$ + DM production sample. The method needs to be tested on $t\bar{t}$ SM sample, to be sure that these events are well reconstructed. This method is based on a algebraic approach and fully described in [2]. In this section only a brief description will be given. The event reconstruction is based on solving the following set of equations:

$$
\begin{aligned}
&E_x^m = p_{x,\nu} + p_{x,\bar{\nu}} \\
&E_y^m = p_{y,\nu} + p_{y,\bar{\nu}} \\
&m_{W^+}^2 = (E_{l^+} + E_\nu)^2 - (p_{x,\nu} + p_{x,l})^2 - (p_{y,\nu} + p_{y,l})^2 - (p_{z,\nu} + p_{z,l})^2 \\
&m_{W^-}^2 = (E_{l^-} + E_{\bar{\nu}})^2 - (p_{x,\bar{\nu}} + p_{x,l})^2 - (p_{y,\nu} + p_{y,l})^2 - (p_{z,\nu} + p_{z,l})^2 \\
&m_t^2 = (E_{l^+} + E_\nu + E_b)^2 - (p_{x,l^+} + p_{x,b} + p_{x,\nu})^2 - (p_{y,l^+} + p_{y,b} + p_{y,\nu})^2 - (p_{z,l^+} + p_{z,b} + p_{z,\nu})^2 \\
&m_{\bar{t}}^2 = (E_{l^-} + E_{\bar{\nu}} + E_{\bar{b}})^2 - (p_{x,l^-} + p_{x,\bar{b}} + p_{x,\bar{\nu}})^2 - (p_{y,l^-} + p_{y,\bar{b}} + p_{y,\bar{\nu}})^2 - (p_{z,l^-} + p_{z,\bar{b}} + p_{z,\bar{\nu}})^2
\end{aligned}
$$

This system can be solved considering the following assumptions:

- the b-jet mass is equal to the b-quark mass
- leptons and neutrinos are considered to be mass-less
- $\mathrm{m}_t = \mathrm{m}_{\bar{t}}$ = 172.5 GeV
- the invariant lepton-neutrino mass is equal to the m_W=80.4 GeV

In a significant fraction of events there is no solution. In order to solve this problem each event is reconstructed 100 times, each time smearing the kinematic variables within detector resolution. The smearing factors were obtained from the MC simulation.
During the kinematic reconstruction an event weight is calculated. The event weight value represents the quality of the reconstruction, a larger weight value corresponds to a better solution. In this analysis the weight value is used as a criterium of reconstruction quality.

4.2 Kinematic reconstruction performance check on SM $t\bar{t}$

To check kinematic reconstruction performance the energy resolution and ΔR have been studied, which represents the difference between generator level and reconstructed particles.
Where $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.
In Figure 3 you can see the ΔR distribution for the top quark decay products: leptons, b-quarks, neutrinos, and the top-quark itself. We can see that lepton kinematics are accurately reconstructed. The ΔR distribution for b-quarks is broader than the same distribution for a leptons because we reconstruct b-jets in a finite cone. Because we

obtain neutrino kinematic information only as a solution of the kinematic reconstruction algorithm the neutrinos are the most poorly reconstructed particles, as one can see. Top-quark resolution graphs fully reflects the behaviour of the decays products.
In general one can say that we reconstruct the top-quark kinematics well, despite the fact that neutrinos are invisible in the detector.

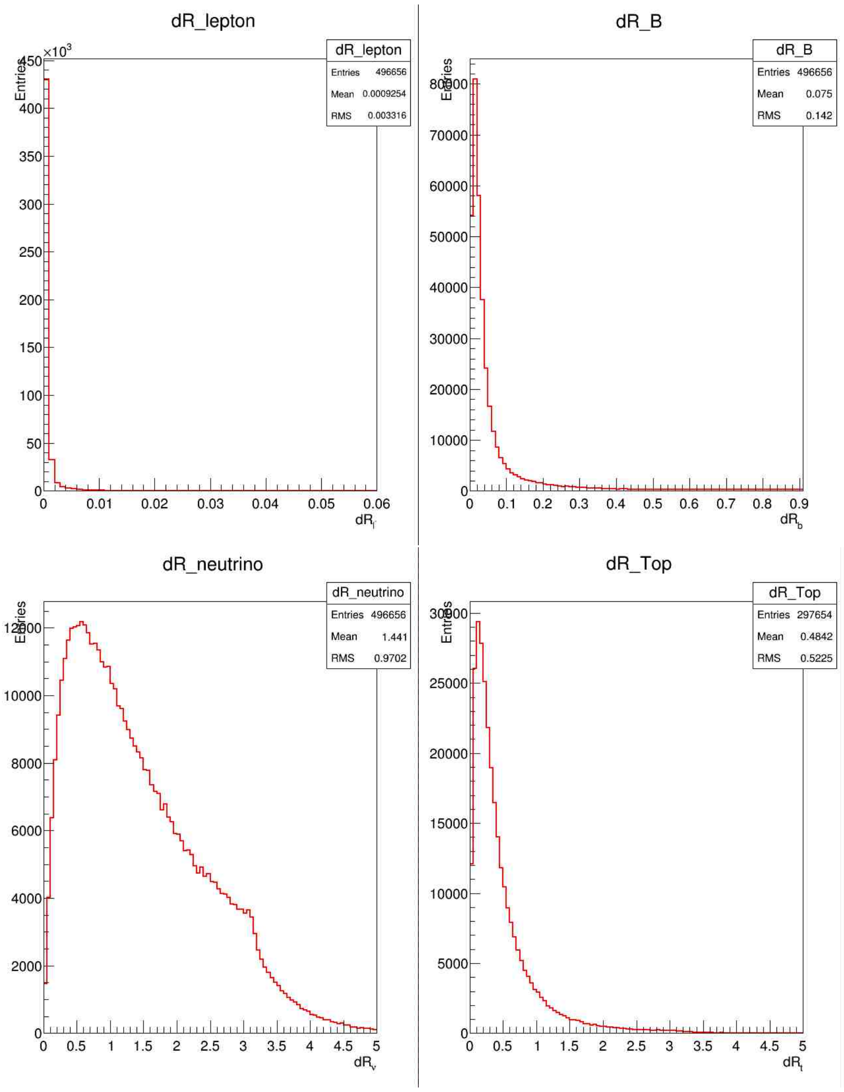

Figure 3: ΔR distributions for l(top left), b-quarks(top right), ν(bottom left), t-quarks(bottom right)

4.3 Attempt to apply default kinematic reconstruction on DM sample

Once we know that the reconstruction method works on a $t\bar{t}$ SM sample, we can try to apply the same method on a $t\bar{t}$ + DM sample. In this case we should expect worse performance, due to the fact that we have larger MET values compatible with $t\bar{t}$ production. In Figure 3 MET distributions are shown. For the DM sample the mass point $(M_\phi = 10GeV, M_\chi = 1GeV)$ has been chosen due to the larger statistics. Even though only the lowest mass point is presented, one already can see difference between SM and DM MET distributions.

In Figure 5 a comparison of the weight distributions is presented. The weight distribution for scalar case looks very similar as the weight distribution for SM, but still it is possible to distinguish them. The weight distribution for PS has a narrow peak near zero, this fact tells us that there is a significant amount of poorly reconstructed events. Hence we can check if the weight value can be used to distinguish the DM and SM processes. That means that we can try to apply a cut on the weight value to check if we can increase the signal over background ratio (S/B).

The initial weight distributions were normalised to the same luminosity of 20 fb^{-1} for each of the mass points, and then the number of events that we obtain after we apply cut on the weight value was calculated, i.e. we throw away events with a weight above the cut value.

In Figure 6 S/B as a function of the weight cut value for all mass points is shown.

From these distributions one should conclude that it is impossible to significantly increase S/B using a cut on the weight of the default reconstruction method.

In the next section of my report a modification of the reconstruction method is described.

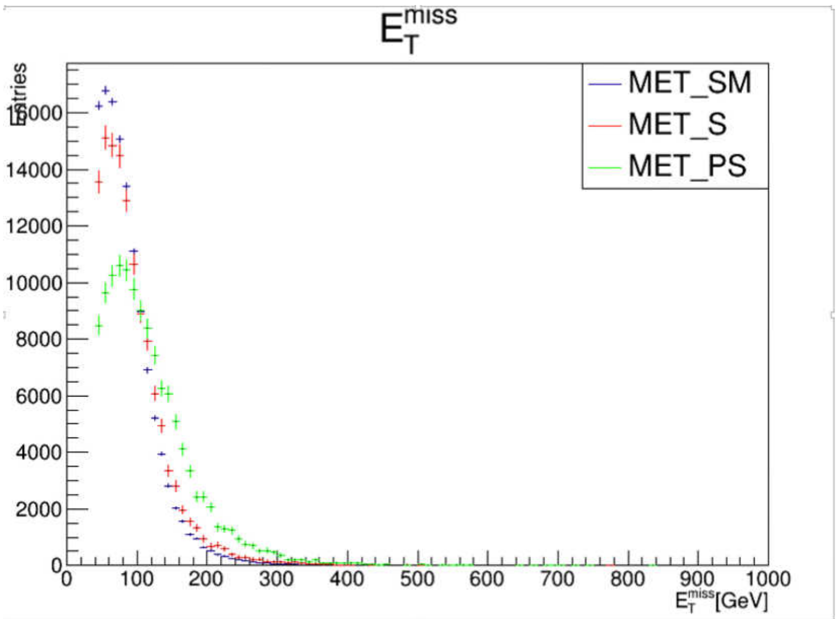

Figure 4: E_T^{miss} distributions comparison

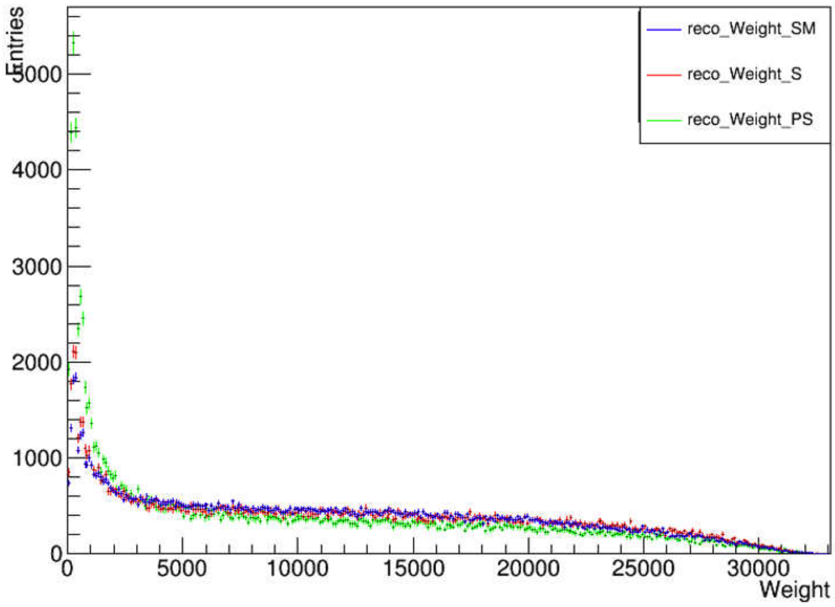

Figure 5: Weight distributions comparison

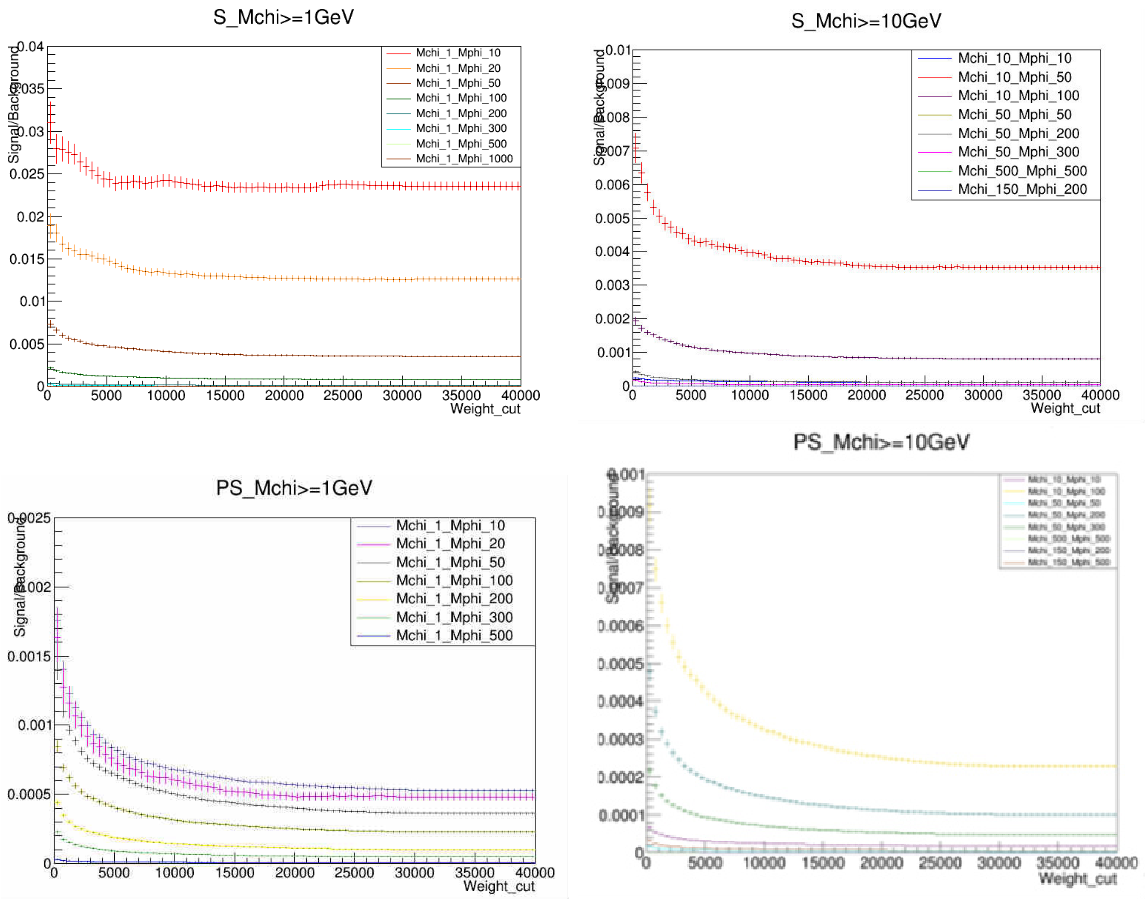

Figure 6: Weight cut distributions for scalar mediator

5 Modified algorithm of the kinematic reconstruction

5.1 Description

The main problem of the DM search in association with $t\bar{t}$ production in dilepton channel is that neutrinos and DM particles are both invisible in the detector. That give us a problem with separating the MET between neutrinos and DM particles. The one of the possible ways to solve this problem is to let the kinematic reconstruction algorithm decide which fraction of the observed MET is most compatible with $t\bar{t}$ production. This could be done by scanning over possible MET vectors as input for the kinematic reconstruction algorithm and consider the weight with the best reconstruction as one that coming from $t\bar{t}$.

We vary the components of the vector that we give to the algorithm with 1 GeV steps in the following ranges :

- $-1.2|E_x^{miss}| < p_x < 1.2|E_x^{miss}|$
- $-1.2|E_y^{miss}| < p_y < 1.2|E_y^{miss}|$

Where $E_{x,y}^{miss}$ are x and y components of the observed MET in the event. For each of these vectors the kinematic reconstruction algorithm has been applied, and the best solution has been chosen(based on the maximum weight value). One could choose wider ranges for x and y components of the momenta, but these ranges has been chosen to speed up the computing process.

One should expect that resolution for $t\bar{t}$+DM events will improve.

5.2 Comparison with the results obtained by default algorithm

In this section the results obtained with modified kinematic reconstruction algorithm are presented.

In Figure 7 the weight distributions are presented. The weight values are expected to become larger even in SM $t\bar{t}$, because a choice of the best solution is based on the maximum weight value. In Figure 9 the ΔR distributions for the neutrinos are shown. There is no significant improvement in comparison with the results obtained using default reconstruction method.

In this research E_x^{ν}/E_x^{miss} and E_y^{ν}/E_y^{miss} correlation has not been taken into account, but one can clearly see some correlation in Figure 8. ($E_{x,y}^{\nu}$ - x and y components of neutrinos momenta; $E_{x,y}^{miss}$ - x and y components of the MET that we give as an input information to the algorithm) The fact that we don't put any constraints on angle variables in our method could give poor $\Delta\phi$ resolution, which could be the reason of this broad structure that we can see in ΔR distributions.

In Figure 10 the neutrino energy resolution plots are presented. One can see that energy resolution has slightly changed after the new algorithm has been applied.

In Figure 11 ΔR distributions for the top-quark are shown. One can see that for S and PS cases distributions has slightly changed, but it should be checked on higher statistics. ΔR distribution for SM became broader, but in SM model we should get the same resolution. One should conclude that probably the weight criteria is not the best. In Figure 12 the energy resolution plots for top-quark are presented, here we can see that we achieved a much improved resolution with the new reconstruction method.

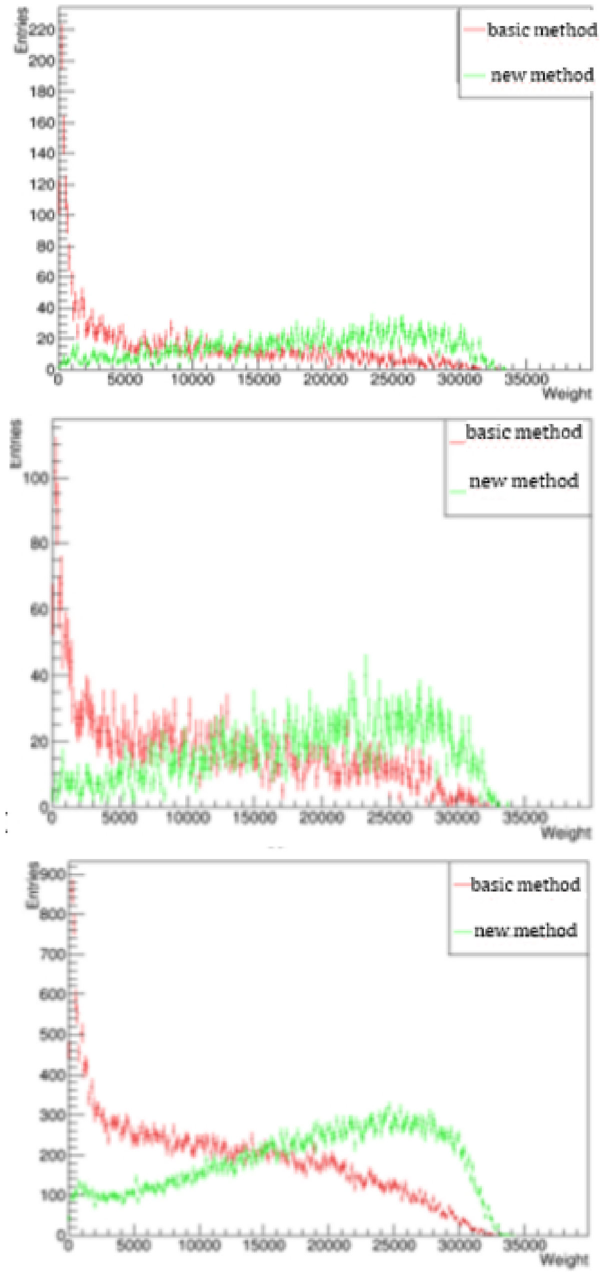

Figure 7: Weight distributions for a new algorithm; upper-PS, middle-S, bottom-SM

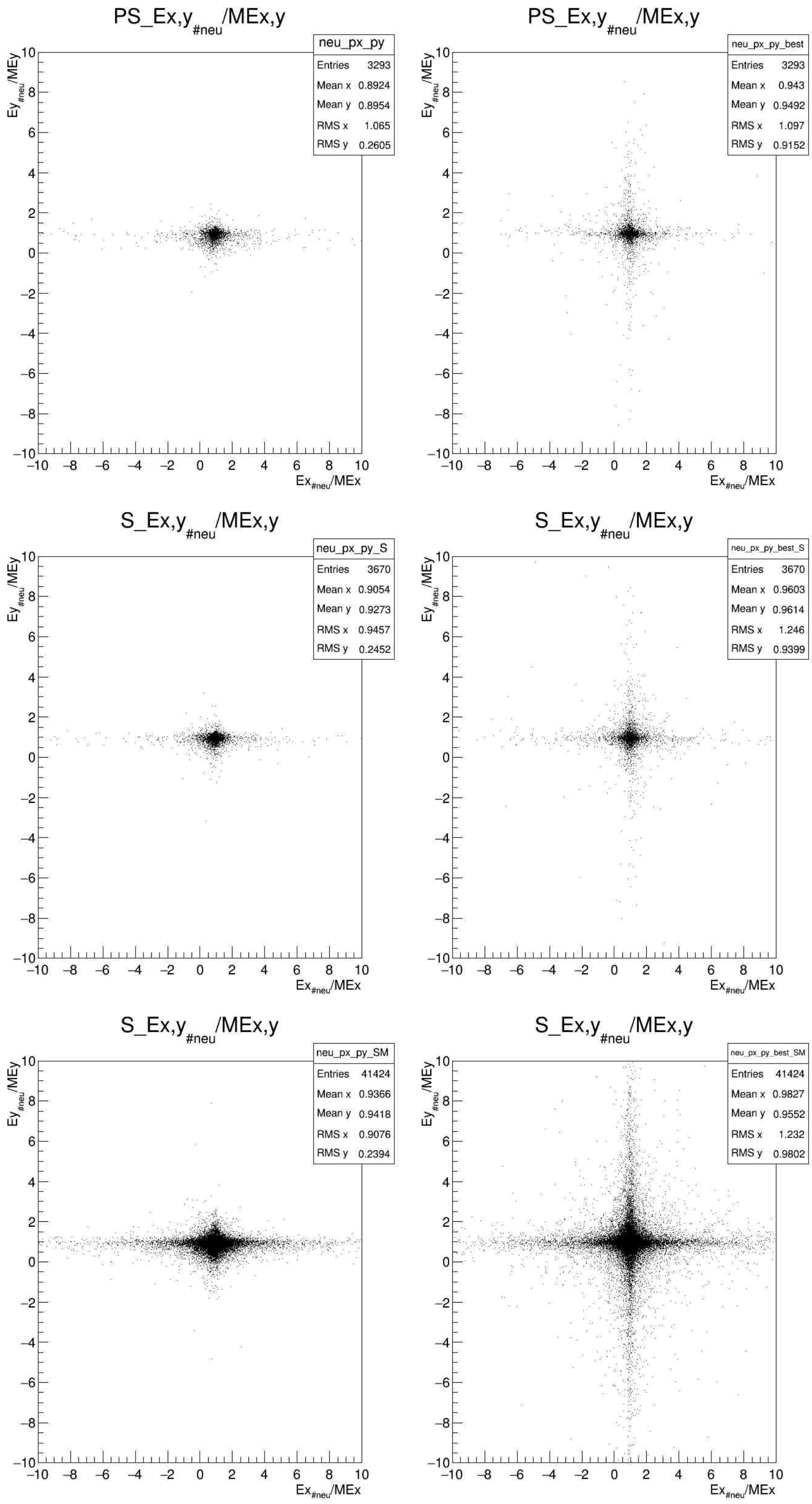

Figure 8: Distributions illustrating correlation between E_x^{ν}/E_x^{miss} and E_y^{ν}/E_y^{miss}
x axis:E_x^{ν}/E_x^{miss}, y axis:E_y^{ν}/E_y^{miss}
on the left:default method, on the right: new method;
upper-PS, middle-S, bottom-SM

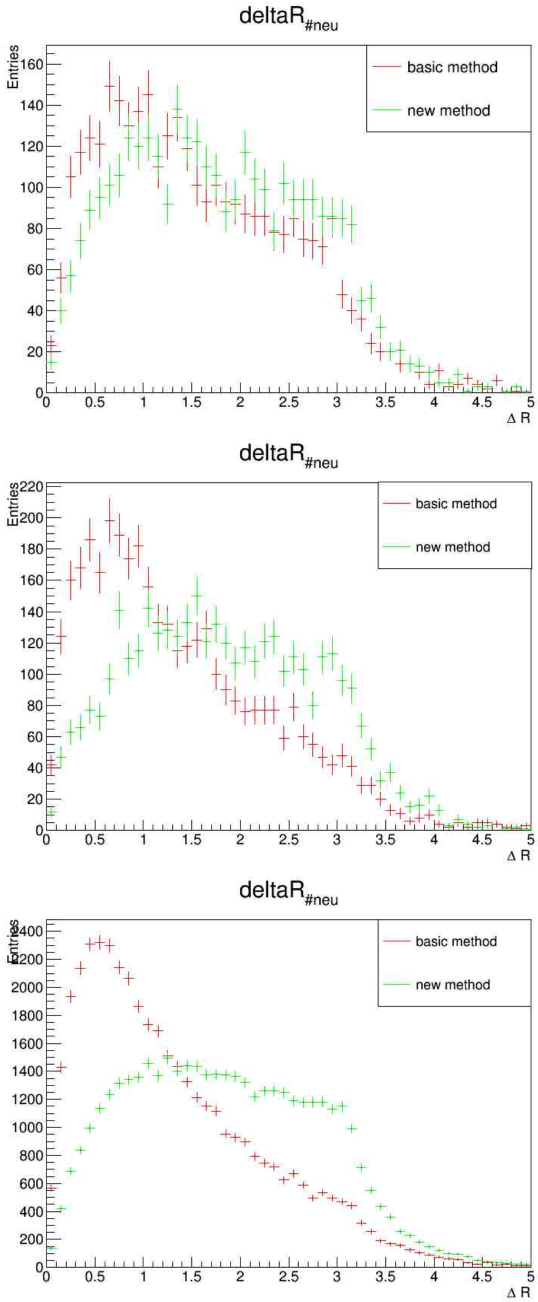

Figure 9: ΔR_ν for a new algorithm; upper-PS, middle-S, bottom-SM

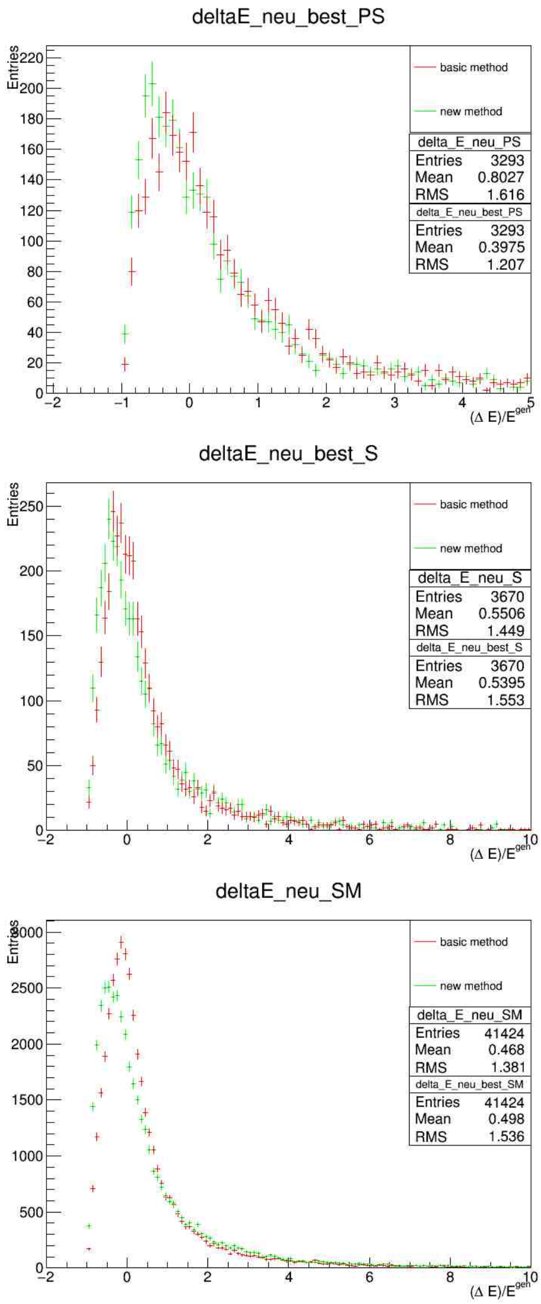

Figure 10: Neutrino energy resolution distributions for a new algorithm; upper-PS, middle-S, bottom-SM

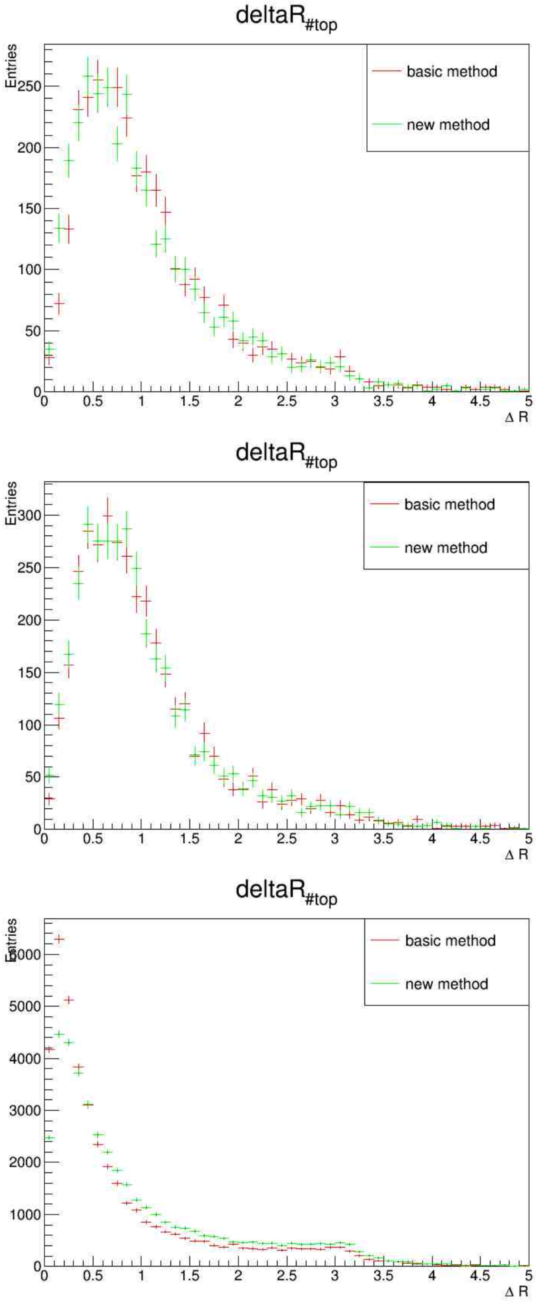

Figure 11: ΔR_t for a new algorithm; upper-PS, middle-S, bottom-SM

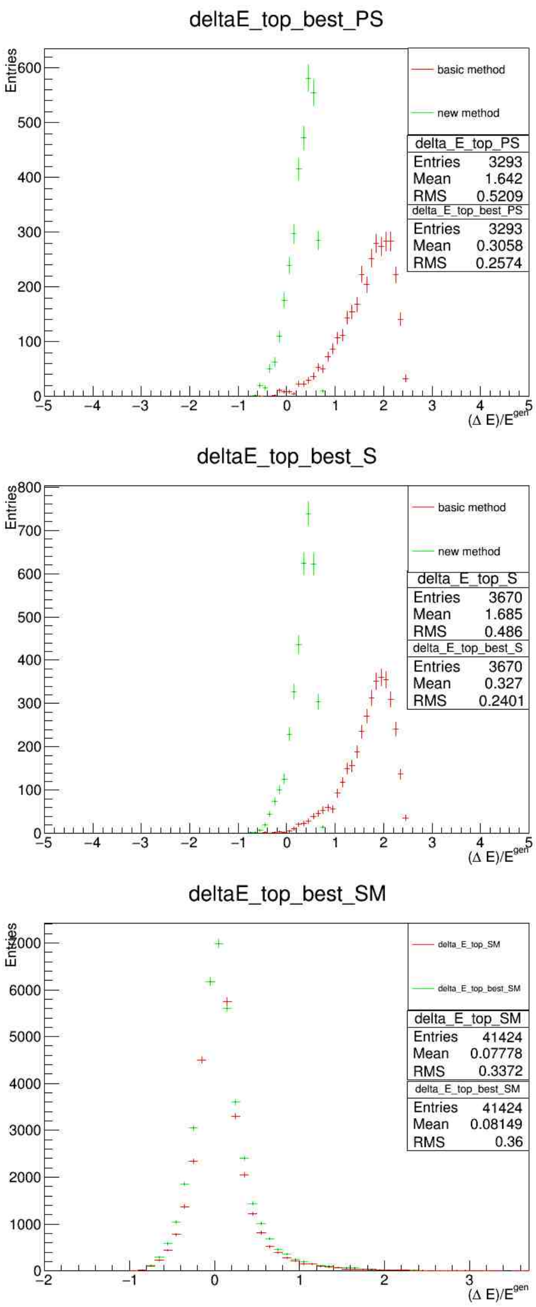

Figure 12: Energy resolution distributions for a new algorithm; upper-PS, middle-S, bottom-SM

6 Conclusion

It has been checked that the default kinematic reconstruction performs well in $t\bar{t}$ SM production. As a result of probing the kinematic reconstruction on the $t\bar{t}$ + DM production we concluded that it's impossible to significantly increase s/b using only weight cut. Then an updated kinematic reconstruction has been tested. We can see most significant improvement in the resolution plots for PS sample, and for S sample there is also noticeable improvement in the resolution plots.
These results give us a hope for possible future of this project. For example, we could try to use different criteria of the reconstruction quality even though weight criteria already gives us good results, for example m_{lb} - invariant lepton-b-quark mass, or $\mathrm{m}t\bar{t}$ - $t\bar{t}$ invariant mass. Also one could try to use E_x^{miss} and E_y^{miss} correlation information to specify the area where the best solution could be found.

7 A. Simulation samples

Sample name	$\sigma[pb]$
/TTbarDMJets pseudoscalar Mchi-10 Mphi-100 TuneCUETO8M1 13TeV-madgraphMLM-pythia8	0.1901
/TTbarDMJets pseudoscalar Mchi-10 Mphi-10 TuneCUETO8M1 13TeV-madgraphMLM-pythia8	0.01499
/TTbarDMJets pseudoscalar Mchi-150 Mphi-200 TuneCUETO8M1 13TeV-madgraphMLM-pythia8	0.0004124
/TTbarDMJets pseudoscalar Mchi-150 Mphi-500 TuneCUETO8M1 13TeV-madgraphMLM-pythia8	0.004611
/TTbarDMJets pseudoscalar Mchi-1 Mphi-100 TuneCUETO8M1 13TeV-madgraphMLM-pythia8	0.1909
/TTbarDMJets pseudoscalar Mchi-1 Mphi-10 TuneCUETO8M1 13TeV-madgraphMLM-pythia8	0.4409
/TTbarDMJets$_pseudoscalarMchi-1Mphi-200TuneCUETO8M113TeV-madgraphMLM-pythia8$	0.0836
/TTbarDMJets pseudoscalar Mchi-1 Mphi-20 TuneCUETO8M1 13TeV-madgraphMLM-pythia8	0.3992
/TTbarDMJets$_pseudoscalarMchi-1Mphi-300TuneCUETO8M113TeV-madgraphMLM-pythia8$	0.03999
/TTbarDMJets pseudoscalar Mchi-1 Mphi-500 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.005408
/TTbarDMJets pseudoscalar Mchi-1 Mphi-50 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.3032
/TTbarDMJets pseudoscalar Mchi-50 Mphi-200 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.08382
/TTbarDMJets pseudoscalar Mchi-50 Mphi-300 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.03989
/TTbarDMJets pseudoscalar Mchi-50 Mphi-50 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.002979
/TTbarDMJets pseudoscalar Mchi-500 Mphi-500 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	3.275 x 10^{-6}
/TTbarDMJets scalar Mchi-10 Mphi-100 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.6732
/TTbarDMJets scalar Mchi-10 Mphi-10 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.09487
/TTbarDMJets scalar Mchi-10 Mphi-50 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	2.942
/TTbarDMJets scalar Mchi-150 Mphi-200 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.00013
/TTbarDMJets scalar Mchi-1 Mphi-1000 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.0003687
/TTbarDMJets scalar Mchi-1 Mphi-10 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	19.59
/TTbarDMJets scalar Mchi-1 Mphi-20 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	10.48
/TTbarDMJets scalar Mchi-1 Mphi-500 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.00518
/TTbarDMJets scalar Mchi-1 Mphi-50 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	2.941
/TTbarDMJets scalar Mchi-50 Mphi-300 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.02901
/TTbarDMJets scalar Mchi-50 Mphi-50 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.002329
/TTbarDMJets scalar Mchi-1 Mphi-100 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.6723
/TTbarDMJets scalar Mchi-1 Mphi-200 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.09327
/TTbarDMJets scalar Mchi-1 Mphi-300 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.0295
/TTbarDMJets scalar Mchi-50 Mphi-200 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	0.09224
/TTbarDMJets scalar Mchi-500 Mphi-500 TuneCUETP8M1 13TeV-madgraphMLM-pythia8	9.894 x 10^{-7}
$t\bar{t}$ SM-/TT TuneCUETP8M1 13TeV-powheg-pythia8	831.76

References

[1] The CMS experiment at the CERN LHC J. Instrum., 3 (2008), Article S08004 http://dx.doi.org/10.1088/1748-0221/3/08/S08004 *CMS Collaboration*

[2] Measurement of the differential cross section for top qaurk pair production in the dilepton final state at $\sqrt{s} = 13TeV$ CMS Analyis Note, CMS AN -2015/309, 10 December 2015(v7, 17 March 2016) *CMS Collaboration*

[3] Recommendation on presenting LHC searches for missing transverse energy signals using simplified s-channel models of dark matter, arXiv:1603.04156 [hep-ex] *LHC Dark Matter Working Group*