Study of production of Z' boson in the process $pp \to Z't\bar{t} \to t\bar{t}t\bar{t}$ with top jet reconstruction techniques.

Submitted by: Shubham P. Raghuvanshi
Guided by : Prof. Monoranjan Guchait
Department of High Energy Physics, TIFR, Mumbai.

June 14, 2019

Contents

1 Introduction

The hypothetical Z' boson is required by many extensions of the Standard Model [1, 2], simplest of these models are U(1) extension of SM which require an additional massive vector boson which couples to standard model fermions. It is electrically-neutral, color singlet spin 1 particle. The mass of Z' in principle depends upon it's coupling to various SM and BSM particles, which are model dependant. Some of these model predict the existance of top-philic Z' which only couples to top quarks [32], while in some other models Z' is leptophobic which interacts with SM quarks. Top quark having mass close to electroweak symmetry breaking scale has strongest coupling to heavy particles beyond standard model and as the luminosity increases at LHC we expect top quarks to be produced in huge numbers giving opportunity to probe new physics.

In this work we study jet reconstruction techiques to reconstruct top quarks and use them to analyse events in which Z' is produced in pp collision at $\sqrt{s} = 14$ TeV in assosiation with top anti-top pair 1. We chose B-L model [12] for this study since it is one of the simplest U(1) extensions of the SM. However the methods used for analysis are fairly general. Using MSTW2008lo68cl_nf3 [15] parton distribution functions(PDF), with $M_{Z'} = 1$ TeV and $M_t = 172$ GeV, MADGRAPH$_5$aMC@NLO [19, 20] provides the cross section for this process to be $\sigma_{B-L}(pp \to Z't\bar{t}) = 0.00612 \pm 3.1 \times 10^{-6}$(stat) pb. In this analysis we will be interested mainly in the $Z' \to t\bar{t}$ decay channel, the final state therefore contains 4 top quarks. Table 1 lists cross sections for some of the relavant processes in pp collisions at $\sqrt{s} = 14$ TeV. For the Z' processes cross sections for B-L model is shown. The whole source code for this work as well as some relavant figures can be found at the Github repository of the project here.

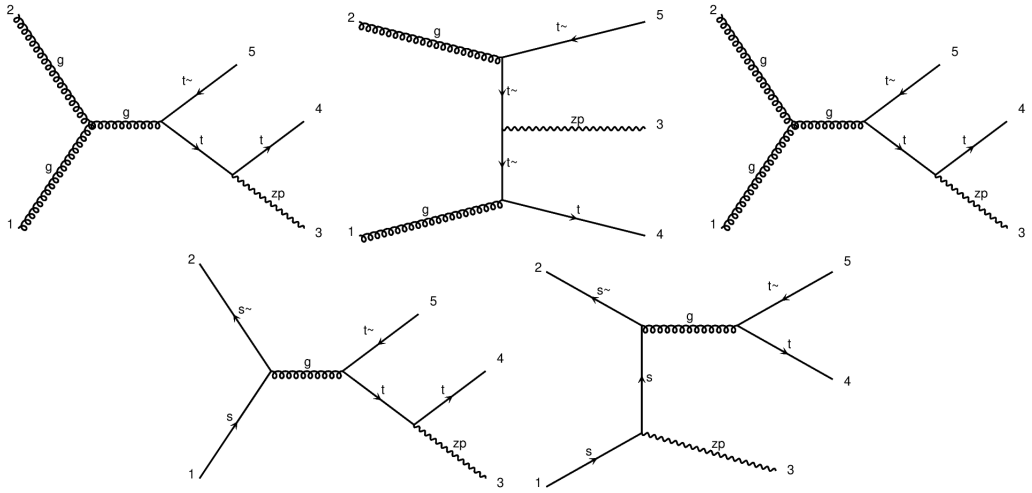

Figure 1: LO Fynnman diagrams for the parton level processes $gg \to Z't\bar{t}$, $qq \to Z't\bar{t}$.

Table 1: cross sections for some of the relavant processes in pp collisions at $\sqrt{s} = 14$ TeV. Here $m'_Z = 1000$ GeV.

Process	cross sections in pb
$pp \to t\bar{t}$	597.9
$pp \to t\bar{t}t\bar{t}$	0.001182
$pp \to t\bar{t} \to 2$ top jets	265.74
$pp \to t\bar{t}t\bar{t} \to 4$ top jets	2.3×10^{-4}
$pp \to Z'$	1.401
$pp \to Z' \to t\bar{t}$	0.00612
$pp \to Z't\bar{t}$	0.0091
$pp \to Z't\bar{t} \to t\bar{t}t\bar{t} \to 4$ top jets	7.19×10^{-5}

2 Z' phenomenonology and current mass constraints

Z' boson is predicted in many extensions of the Standard Model (SM). Each with different coupling to the SM fermions. In general generation-dependant couplings of Z' to the SM fermions can be written as

$$Z'_\mu \left[\sum_{i=1,j=1}^{i=3,j=3} \left(g_f^L\right)_{ij} \bar{f}_L^i \gamma^\mu f_L^j + \left(g_f^R\right)_{ij} \bar{f}_R^i \gamma^\mu f_R^j \right] \tag{1}$$

where the indices i and j run over all three generations and f = (u, d, e, ν). The coefficients $\left(g_f^L\right)_{ij}$ and $\left(g_f^R\right)_{ij}$ are real dimentionless parameters, and analogous to the weak hypercharge of elecro-weak interations the gauge charges z_{fi}^L, z_{fi}^R determine all the couplings and mass and decay width of Z' [8, 9]. Some of these models we discuss here in brief.

- Simple U(1) extensions of the standard model, also known as Sequential Standard Model(SSM) [7], which assumes Z' couplings to SM fermions to be the same as Z^0 but with a polemass of the order of few TeV. These theories have the conserved gauge charges (as in column 1) proportional to the baryon number minus x times the lepton number $(B - xL)$, and are in general known as $U(1)_{B-xL}$ extensions of the standard model [2].

fermion	$U(1)_{B-xL}$	$U(1)_{10+x\bar{5}}$	$U(1)_{d-xu}$	$U(1)_{q+xu}$
(u_L, d_L)	$1/3$	$1/3$	0	$1/3$
u_R	$1/3$	$-1/3$	$-x/3$	$x/3$
d_R	$1/3$	$-x/3$	$1/3$	$(2-x)/3$
(ν_L, e_L)	$-x$	$x/3$	$(-1+x)/3$	-1
e_R	$-x$	$-1/3$	$x/3$	$-(2+x)/3$

Figure 2: Examples of generation-independant U(1) gauge charges for quarks and leptons, here x is any rational number.

  Z' is leptophobic for $x = 0$, i.e. it only decays to quarks and the interactions preserve baryon number, for $x >> 1$ it is quark-phobic. The latter possibilitiy is being ruled out by various collider physics experiments. The case where $x = 1$ is particularly important, it gives $U(1)_{B-L}$, the only anomaly-free global symmetry of the Standard Model, easily promoted to a local symmetry by introducing three right-handed neutrinos, which automatically make neutrinos massive. $B - L$ symmetry remains the only symmetry of the SM which is not observed to be broken [12].

- The second column shows charges for $U(1)_{10+x5}$, these symmetry groups are part of E_6 Grand Unified Theories(GUT) [11]. This set leads to $Z - Z'$ mass mixing close to electroweak scale. Under the third set,$U(1)_{dxu}$, the weak-doublet quarks are neutral, and the ratio of u_R and d_R charges is x. For $x = 1$ this is the right-handed group $U(1)_R$.

The current analysis at LHC are motivated by pure SM contribution in which Z', couples to all the SM fermions, namely the minimal U(1) extensions of the SM. Along these lines, the ATLAS and CMS collaborations have searched for Z' bosons by investigating dilepton and dijet final states with 13 TeV data. The ATLAS collaboration [3] set the mass exclusion limits $M_{Z'} > 4.5$ TeV(SSM) and $M_{Z'} > 3.8 - 4.1$ TeV (GUT-inspired models), whereas CMS obtained $M_{Z'} > 4.0$ TeV(SSM) and $M_{Z'} > 3.5$ TeV (GUT-inspired models) [4]. For dijets, the limits are much milder and read $M_{Z'} > 2.1 - 2.9$ TeV (ATLAS) [5] and $M_{Z'} > 2.7$ TeV (CMS) [6]. The U(1) supersymmetric models however by including supersymmetric decay modes lowers the branching ratios into SM final states and impact the Z' mass limits by $200 - 300$ GeV [14].

3 Top quark reconstruction

The phenomenology of top quark is driven by it's large mass, which also gives it very short lifetime $\sim 10^{-24}$s and makes it the only quark that decays semi-weakly into W boson and b quark. The brancing ratio for the top decay $t \to Wb$ is close to unity, and for this analysis we assume it to be unity. The signature of top quarks which decay hadrnoically i.e. for which the decay $t \to Wb$ is followed by $W \to q\bar{q}$ where $q = (u, d, s, c)$, is in general three quark-like jets in the final state from the hadronization of the quarks. If the parent top is at rest then these three jets are well separated from each other in (y, ϕ) plane and can in general be identified as thee distinct quark jets in an event. On the other hand is the parent top is boosted i.e. has high p_T then the three quark jets will tend to be collinear along the direction of the parent top, and consenquently instead of having three well separated jets in final state we may get only one fatjet. Substructure of this fatjet gives information of the corresponding decay. 3.

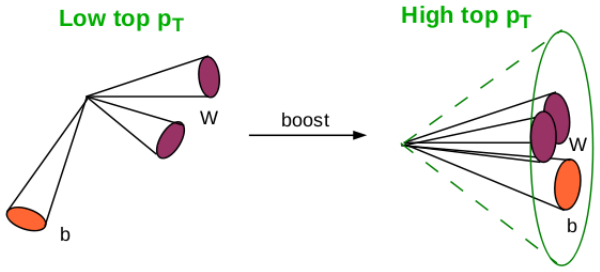

Figure 3

To quantify things lets consider the angular separations between partons in the process $t \to wb \to q\bar{q}b$ as a function of the p_T^{top} fig:4. The angular separation in rapidity-azimuth (y, ϕ) plane $\left(\Delta R = \sqrt{\Delta y^2 + \Delta\phi^2}\right)$ between the decay products in a two-body depends on the mass and boost of the decaying particle and in the high p_T approximation it is given by .

$$\Delta R \approx \frac{2m}{p_T} \tag{2}$$

Where m and p_T are the mass and transverse momenta of the decaying particle. For a top $(m = 173)$ GeV [21] with $p_T = 200$ GeV the separation is $\Delta R = 1.73$, similarly for a W boson $(m = 80)$ [22] with same p_T the separation is $\Delta R = 0.8$. Fig 4 shows the relation between ΔR and p_T for the two body decays of top and w respectively. A falling trend in the angular separation as we go to higher p_T^{top} is apparent, from the plot one can also see that roughly $\Delta R_{wb}^{avg} < 1.7$ for $p_T^{top} > 200$, and $\Delta R_{qq}^{avg} < 0.8$ for $p_T^{w} > 200$ GeV.

The top tagging algorithms that are aimed at reconstructing low p_T top work on the assumption that it would show up in an event as one b-tag and two distinct w-tag jets, this assumption grows weaker as p_T of the top increases and the top instead of corresponding to three quark jets which are well separated from each other, corresponds to one fatjet in which all three jets are merged. The substructure of this fatjet contains the information about the undelying decay. The top reconstruction effeciency of such algorithms is also expected to be going downhill for boosteed tops. We will be looking at the tops in the decay $Z' \to t\bar{t}$. These tops in general may be quite boosted in laboratory frame depending upon the mass of the Z' fig:7 (bottom right). Jet substructure techniques are required to reconstruct top in such cases.

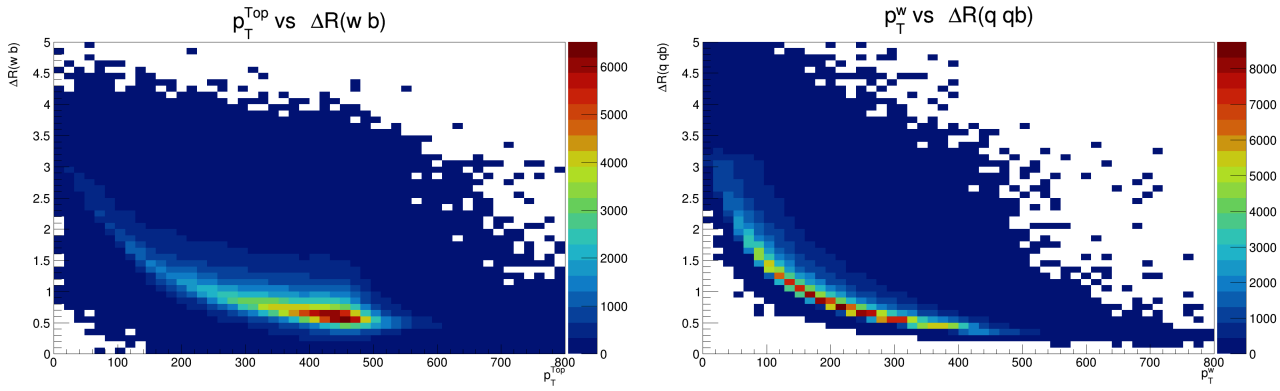

Figure 4: A plot between angular separation between decaying particles of the top and w decay, in (y, ϕ) plane, vs p_T of top and w respectively.

3.1 Top tagging algorithms

A Jet [1]is an artifact of all the high energy collider physics experiments, it's a collimated spray of particles in a given direction which can all be traced back to a common vetrex. These particles result from hadronization of a quark therefore they also share the energy and momenta of the original quark. In the experiments these particles show up as a set of hadrons and photons which are concentrated within some annular region of the detector. There is no universally accepted definition of a jet however the most natural definition is based on a cone radius in pseudorapidity-azimuth plane within which most of the jet particles are contained.

A jet algorithm operates on the four momenta of the final state particles or some detector quantities like the energy deposits in calorimeters, these quantities are referred to as 'constituents'of the jet. A jet algorithm clusters two consituents i and j into one new constituent if the quantity Δd_{ij} which is related to the angular separation ΔR_{ij} between the between the constituents and their transverse momenta , satisfies the condition $d_{ij} < d_{iB}$, i.e. if they are closer to each other than they are to the beam in some distance measure. Clustering is done recursively untill there are no clusters left within R distance from one another. For our analysis we use sequantial anti-k_T and Cambridge/Aachen(C/A) algorithms which differ only in the choice of the distance parameter d_{ij} [2]. More details on jet finding algorithms can be found here [23].

3.1.1 Tagging of low p_T top

The top when produced with low momentum can be reconstructed via three angularly resolved jets. These jets have to satisfy certain properties in order for them to be considered as 'top candidates'. Top tagging algorithms for moderately boosted top are based on this property. We describe two such algorithms which expolit this property of low p_T tops and identify the 3-pronged structure by putting mass constraints on the reconstructed jets. Jets which do not satisfy $p_T^{jet} > 20$ GeV are discarded.

1. Kinematic mass cut
2. χ^2 minimization.

First algorithms takes reconstructed jets and first identifies a pair of w-tagged jets based on the criterion $|m_{ij} - m_w| < \Delta_w$, where m_{ij} is the invariant mass of jet i and jet j, Δ_w is some positive number. If two different pairs of jets satisfy this criteria then pair with invariant mass closest to w

[1]By Jet we mean QCD Jet.

[2] $d_{ij} = \frac{\Delta R_{ij}}{R}$ and $d_{iB} = 1$ for C/A and $d_{ij} = min\left(\frac{1}{p^2_{\mathrm{T,i}}}, \frac{1}{p^2_{\mathrm{T,j}}}\right) \times \frac{\Delta R_{ij}}{R}$ and $d_{iB} = \frac{1}{p^2_{T,i}}$ for anti-k_T. The two consituents i,j are clustered together if $d_{ij} < d_{iB}$.

mass is chosen, and if two different pairs have same invariant mass till two significant digits then the one with highest p_T is chosen to be w candidate. After identifying two jets(i and j) as w tagged the next step is the identification of the triplet (ijk) as top tagged if $|m_{ijk} - m_{top}| < \Delta_{top}$. Again the triplet which has m_{ijk} closest to top mass is chosen in case two pairs satisfying this criteria. The process is repeated untill all the jets have been analysed. For this study we take $\Delta_w = 5.0\Gamma_w$ and $\Delta_{top} = 5.0\Gamma_{top}$, where $\Gamma_w = 2.0$ [22] and $\Gamma_{top} = 1.5$[21] are the decay widths of w and top respectively. These tagged top jets are then matched with the parton level top quark. Reconstruction efficency is defined to be the fraction of number of parton level top quarks for which a match with tagged top jets is found within a cone radius of 0.3.

$$\epsilon_{\text{top reco}} = \frac{N(R_{\text{top-quark, tagged-topjet}} \leq 0.3)}{N(\text{parton level top})} \tag{3}$$

The χ^2 minimization procceds by identifying the triplet (ijk) from the list of reconstructed jets which gives the minimum value of the quantity

$$\chi^2_{ijk} = \frac{(m_{ij} - m_w)^2}{\sigma_w^2} + \frac{(m_{ijk} - m_{top})^2}{\sigma_{top}^2} \tag{4}$$

where σ_w and σ_{top} are related to the decay widths of w and top respectively. Note that the same combination of jets (ijk) will give the minimum value for χ^2_{ijk} if σ_w and σ_{top} have fixed ratio. For the sake of simplicity we take them to be $\sigma_w = 5.0\Gamma_w$, $\sigma_{top} = 5.0\Gamma_{top}$, and the top tagging criteria becomes $\chi^2_{ijk} \leq 2$. These tagged top jets are again matched with a parton level top and the reocnstruciton effciency is given by equation 3.

3.1.2 Tagging of boosted tops

Top tagging algorithms for boosted tops make use of the fact that the decay products of top have large fraction on their momenta along the direction of the original top quark and analyse the top decay from geometrically large objects called fatjets. The siz e of this fatjet should be large enough to contain all three decay products of top. Fig 5 shows the size of the fatjet which can contain the top mass for a given top p_T. It's clear that for tagging high p_T tops we should chooose smaller fatjet size. Many algorithms are available for tagging of boosted tops such as mass drop tagger, Jon Hopkins tagger etc. , details about these algorithms can be found here [25]. For our analysis we use HepTopTagger(Heidelberg-Eugene-Paris) with C/A fatjets of radius $R_{\text{fat}} = 1$.

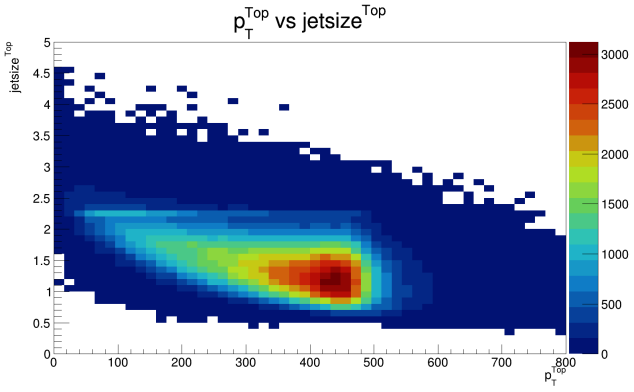

Figure 5: Size of fatjet around top which can contain the top mass vs p_T^{top}.

HepTopTagger algorithms procceds in the following way [26].

1. Reconstruct C/A fatjets with R=1 and $p_T^{jet} > 200$ GeV.
2. Undo the last stage of clustering into jets j_1 and j_2 with $m_{j_1} > m_{j_2}$. Keep the two jets if they satisfy the mass drop criteria $m_{j_1} < 0.8 m_j$. Otherwise keep only j_1. This declustering is done iteratively untill we have subjets with $m_{j_i} > 30$ GeV.

3. Construct exactly three subjets (123) from five hardest subjets. Accept them as top candidates if their invariant masses (m_{12}, m_{13}, m_{23}) satisfy one the following criteria.

$$0.2 < \tan^{-1}\left(\frac{m_{13}}{m_{12}}\right) \text{ and } R_{min} < \frac{m_{23}}{m_{123}} < R_{max}$$

$$R^2_{min}\left(1+\left(\frac{m_{13}}{m_{12}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R^2_{max}\left(1+\left(\frac{m_{13}}{m_{12}}\right)^2\right) \text{ and } \frac{m_{23}}{m_{123}} > 0.35$$

$$R^2_{min}\left(1+\left(\frac{m_{12}}{m_{13}}\right)^2\right) < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R^2_{max}\left(1+\left(\frac{m_{12}}{m_{13}}\right)^2\right) \text{ and } \frac{m_{23}}{m_{123}} > 0.35$$

Where $R_{min} = 0.85\frac{m_w}{m_{top}}$ and $R_{max} = 1.15\frac{m_w}{m_{top}}$

4. Require the resultant p_T of the three subjets to be greater than 200 GeV.

In order to have a comparision between the three algorithms, we require the fatjets to be top tagged in the mass range $m_{top} \pm \Delta_{top}$ and w-tagged in the mass range $m_w \pm \Delta_w$. Then as a consistancy check these tagged jets are matched with the corresponding parton level top, and the efficiency is defined same as equation 3.

3.2 Single top reconstruction

In order to compare the top reconstruction efficiencies of the algorithms we generated 5×10^5 events of the type $Z' \to tt \to$ wb wb $\to q_1\bar{q}_2$b $l\nu_l$b, to Leading Order(LO) using MADGRAPH$_5$aMC@NLO [19, 20] for $M_{Z'} = 500, 1000, 2000$ GeV. PYTHIA8.2 [27] is used for further decays and hadronization of these particles and for evolution of QCD showers. Final state of the signal contains 4 jets, one lepton and missing energy. We select events in which there are atleast three AK5(anti-kT with R=0.5) jets and one lepton each with $p_T > 20$ GeV and $|\eta| < 2.5$. The jets are groomed using SoftDrop[28] which removes soft componenent of the radiation. Top jets are tagged using kinematic mass cut, and HepTopTagger in the Top and W mass range $m_{top} \pm \Delta_{top}$, $m_w \pm \Delta_w$ respectively, and with χ^2 minimization where $\chi^2_{min} < 2$ is the tagging criteria. Fig 6 shows the invariant mass distributions of the tagged top jets. We can see that the tagging efficiency of HepTopTagger for higher $M_{Z'}$ is higher while for the other two methods the opposite is true, due to the larger boost of the top quark.

The quantity $\Delta R(top^{tagged}, top^{parton}) = \sqrt{\left(y_{top^{tagged}} - y_{top^{parton}}\right)^2 + \left(\phi_{top^{tagged}} - \phi_{top^{parton}}\right)^2}$ which is the angular separation between the tagged top jets and the corresponding parton level top gives a measure of precision of the reconstructed momentum of top tagged jets [3]. For a top tagged jet to have reconstructed the parton level top we require it to have $\Delta R < 0.3$ with the corresponding parton level top. Fig 8 shows the distribution of ΔR. The tagged top jets which also match with the corresponding parton level top are called reconstructed top jet. Fig 7 shows the p_T distribution for reconstructed top jets. For moderately boosted top quarks ($pT \approx 500GeV$) , conventional top quark reconstruction methods, which exploit the decay chain topology, remain adequately efficient whereas the reconstruction of high p_T tops is better for HepTopTagger .The top reconstruction efficiency as a function of p_T^{top} is shown in fig.9.

[3]Here by momemtum we mean 3 dimentional vector momentum $\vec{P}_{top}$.

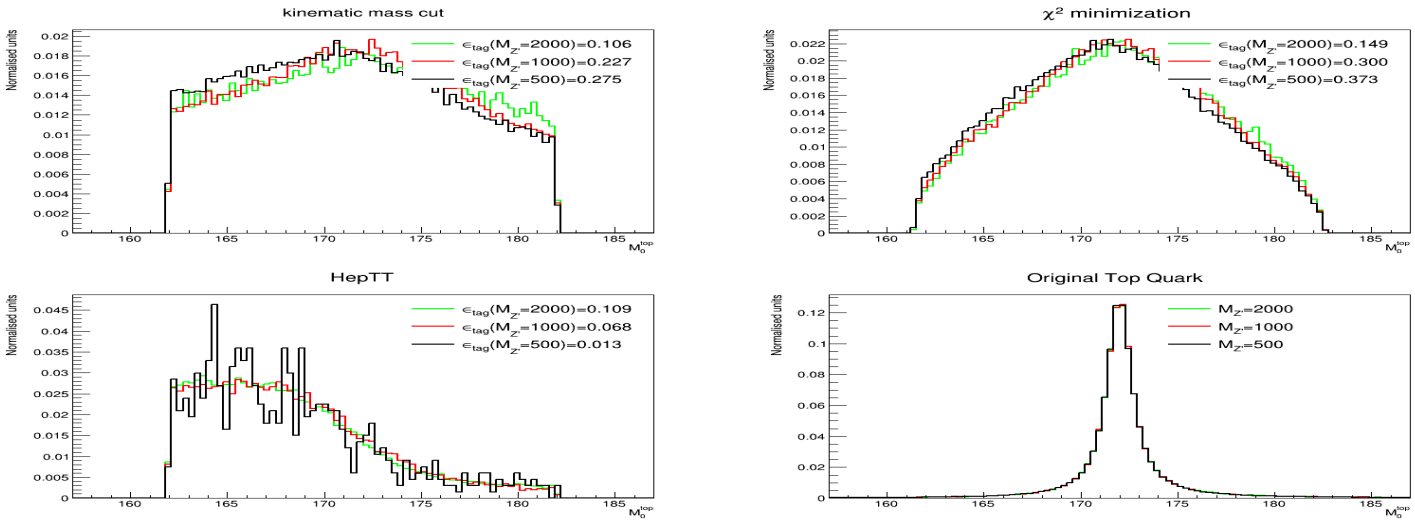

Figure 6: Invariant mass distribution of the tagged top jets in the process $Z' \to t\bar{t} \to qqbl\nu_l b$ using kinematic mass cut(top left), χ^2 minimization (top right), HepTopTagger (bottom left) and of the parton level top quark (bottom right). For M'_Z = 0.5 TeV(black), 1.0 TeV(red), 2.0 TeV(green) respectively.

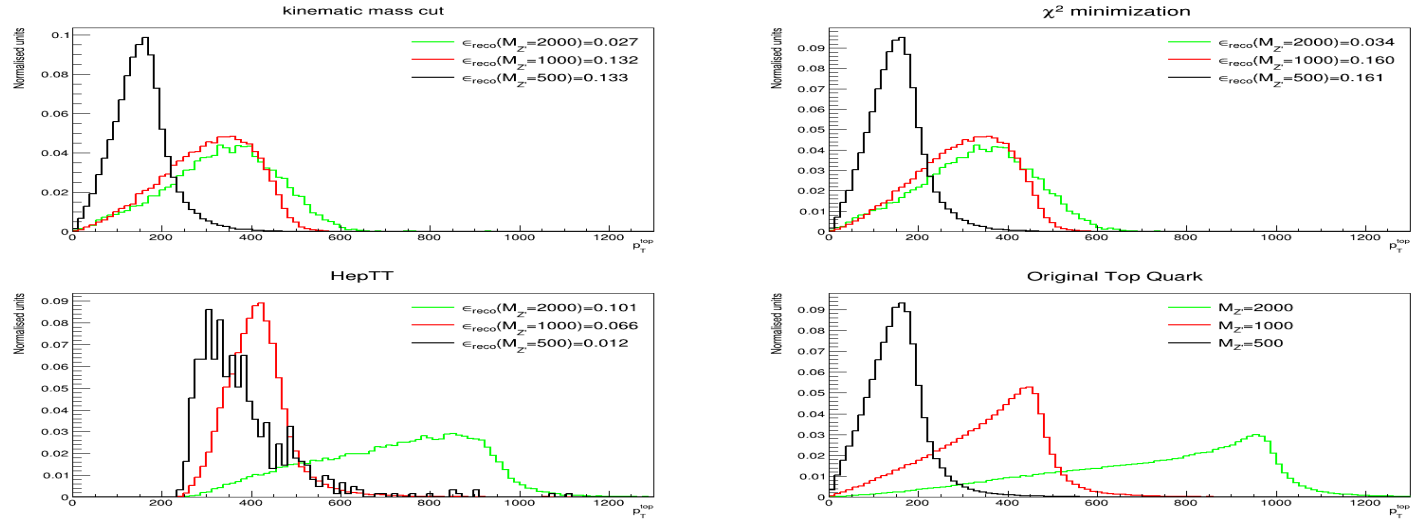

Figure 7: p_T distribution of the reconstructed tops viz tagged top jets which have also matched with it's parton level top, in the process $Z' \to t\bar{t} \to qqbl\nu_l b$ using kinematic mass cut(top left), χ^2 minimization (top right), HepTopTagger (bottom left) and of the parton level top quark (bottom right). For M'_Z = 0.5 TeV(black), 1.0 TeV(red), 2.0 TeV(green) respectively. Reconstruction efficiency is defined as fraction of input top quarks that are reconstructed.

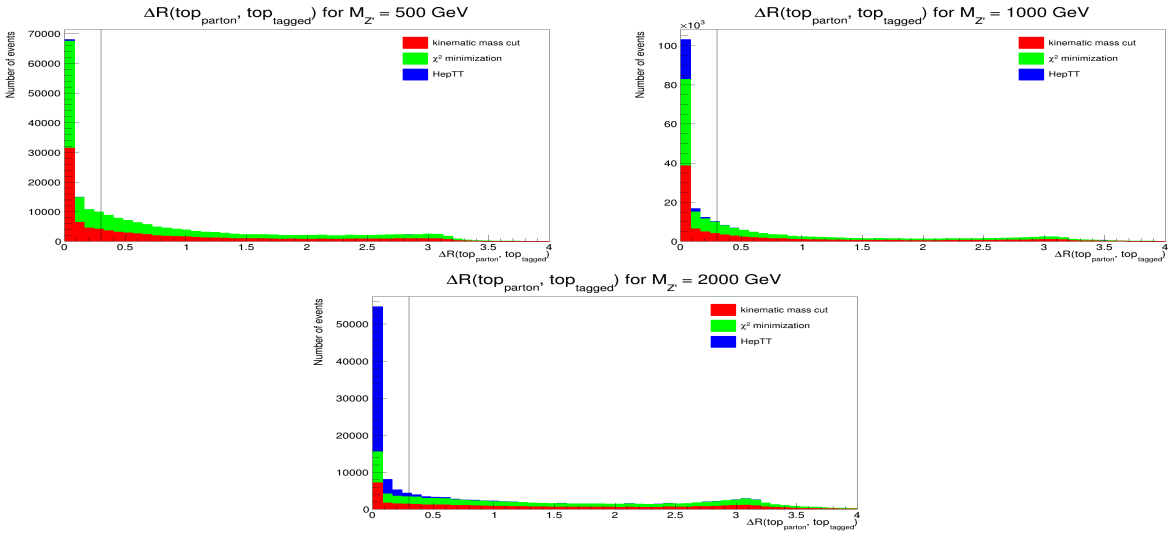

Figure 8: ΔR between the tagged top jets and the parton level top quark. For $M'_Z = 500$GeV (top left), $M'_Z = 1000$ GeV (top right), $M'_Z = 2000$ GeV (bottom). Vertical black line in all the histograms denote the value $\Delta R = 0.3$. A top is termed 'reconstructed' if the tagged top jet has $\Delta R < 0.3$ with the corresponding parton level top i.e. left of the line.

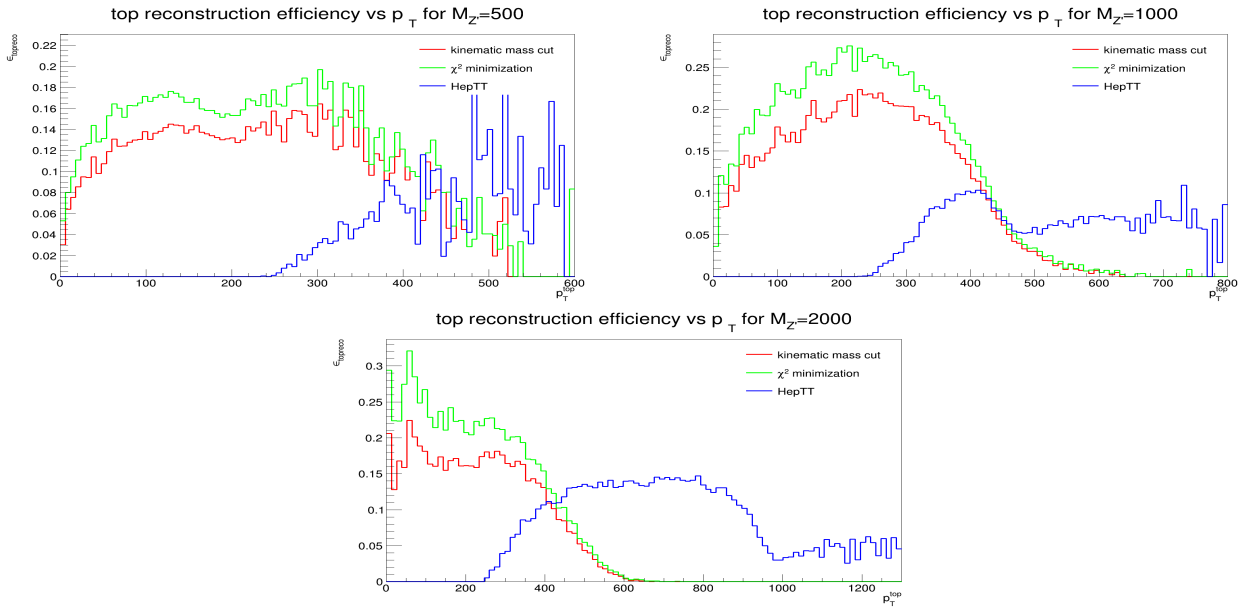

Figure 9: Top reconstruction efficency vs p_T^{top}. For $M'_Z = 500$GeV (top left), $M'_Z = 1000$ GeV (top right), $M'_Z = 2000$ GeV (bottom).

3.3 QCD multijet background

Having analysed the top recontruction efficiency for all three methods we now calculate their mistagging rate with respect to QCD multijet background [4]. QCD multijet processes have inclusive large cross section of about 300 mb in pp collisions at $\sqrt{s} = 14$ TeV [27] and it clouds all the physics searches in hadronic channels. These multijet events contain light and energetic jets in the final state due to hadronization of quark and gluon from the QCD quark-quark, quark-gluon and gluon-gluon interaction processes like $gg \to gg$, $gg \to q\bar{q}$, $qg \to qg$, $qq \to qq$, $q\bar{q} \to gg$. The background from light quark jets to boosted top decays can be removed completely by putting a mass contraint on the reconstructed jets. These light quark jets can however fake the 3 pronged decay topology of low p_T top decay. The gluon jets on the other hand can infiltrate the top mass window and be mistagged as boosted top fatjet, and the cross section for acquiring mass in top mass window increases with increase in R and p_T of the jet. Figure 10 shows the QCD jet multiplicity and the invariant mass distribution of C/A jets with R=1. We define the fractional fake rate (or mistag rate if you will) as the fraction of input jets which are top tagged in the mass range $m_{top} \pm \Delta_{top}$, where $\Delta_{top} = 5 \times \Gamma_{top} = 7.5$ GeV .

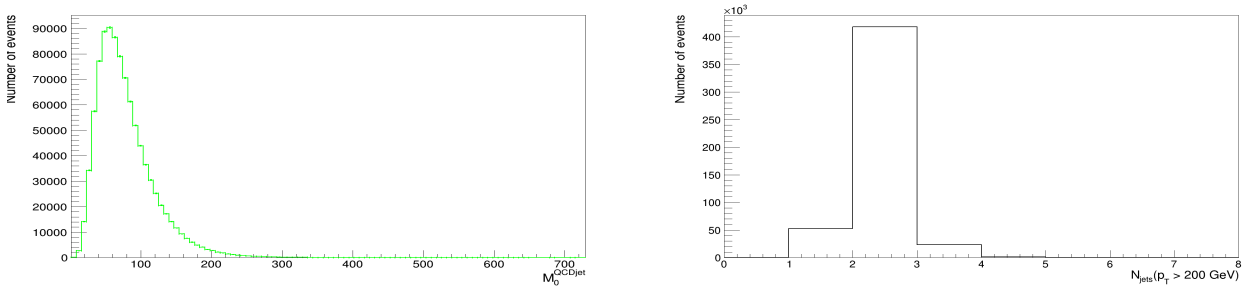

Figure 10: Invariant mass and multiplicity distributions for C/A jets with R=1 in 500000 QCD events. The peak at lower masses is due to light quark jets to LO whereas the long tail of the distribution are gluon jets which penetrate to about 700 GeV and hence contaminate top mass window. Note that the histogram is drawn for the whole range of jet mass that we got in the simulation.

A relative comparision among the top tagging algorithms and their perfomance to QCD multijet background is summarized in Table.2, where we have shown the number of events surviving each selection criteria for top signal as well as for QCD background. We didn't put any p_T^{lepton} cut in QCD background in order to get maximum statistics. For low p_T top quark reconstruction χ^2 minimization works better among the two and we will use this method in the following sections to reconstruct low p_T top quarks.

[4] In this context multijet stands for > 2 inclusive jets

Table 2: Cut flow for $Z' \to t\bar{t} \to qqbl\nu_l b$ events for $M_{Z'}$ = (500, 1000, 2000) GeV and for QCD multijet background.

Selection cut	$\mathbf{Z' \to t\bar{t}}$			**QCD**
	$\mathbf{M'_Z}$**=500 GeV**	$\mathbf{M'_Z}$**=1 TeV**	$\mathbf{M'_Z}$**=2 TeV**	
Number of events generated	500000	500000	500000	500000
Atleast one lepton with $p_T^{lepton} > 20$ GeV	388148	431552	463515	-
Atleast three AK5 jets with $p_T^{jet} > 20$ GeV	341632	400695	419343	401554
Atleast one CA10 fatjet with $p_T^{jet} > 200$ GeV	51411	343875	448661	495149
Tagged top jets (kinematic mass cut)	94030, $\epsilon_{tag}^{kin} = 0.275$	91106, $\epsilon_{tag}^{kin} = 0.227$	44387, $\epsilon_{tag}^{kin} = 0.105$	20542, $\epsilon_{mistag}^{kin} = 0.034$
Tagged top jets (χ^2 minimization)	127441, $\epsilon_{tag}^{\chi^2} = 0.373$	120344, $\epsilon_{tag}^{\chi^2} = 0.3$	62288, $\epsilon_{tag}^{\chi^2} = 0.148$	39120, $\epsilon_{mistag}^{\chi^2} = 0.065$
Tagged top jets (HEPTopTagger)	668, $\epsilon_{tag}^{hepTT} = 0.012$	23532, $\epsilon_{tag}^{hepTT} = 0.068$	48974, $\epsilon_{tag}^{hepTT} = 0.109$	707, $\epsilon_{mistag}^{hepTT} = 0.0007$
Matched top jets (kinematic mass cut)	45549, $\epsilon_{tag}^{kin} = 0.133$	53084, $\epsilon_{tag}^{kin} = 0.132$	14497, $\epsilon_{tag}^{kin} = 0.027$	-
Matched top jets (χ^2 minimization)	55034, $\epsilon_{tag}^{\chi^2} = 0.0161$	64099, $\epsilon_{tag}^{\chi^2} = 0.159$	14338, $\epsilon_{tag}^{\chi^2} = 0.034$	-
Matched top jets (HEPTopTagger)	616, $\epsilon_{tag}^{hepTT} = 0.012$	22815, $\epsilon_{tag}^{hepTT} = 0.066$	45489, $\epsilon_{tag}^{hepTT} = 0.101$	-

4 Event Selection

So far we have discussed varius top reconstruction methods and tools, compared their tagging and mistagging rate for different phase space regimes. In this section we will use these methods to reconstruct the line shape of Z' boson in it's all hadronic top decays. For this purpose we generate events in which Z' boson produced in assosiation with a top pair which itself susequently decays to top anti top pair, i.e. events of the type $pp \to Z't\bar{t} \to t\bar{t}t\bar{t}$ [5] . Assuming top decay to W and b to have 100% branching ratio, the hadronic decay of top has a branching ratio of about 2/3 while for leptonic decay the branching fraction is about 1/3. Therefore in order to get maximum statistics we do the analysis where all 4 final state top decay hadronically. For this study we choose $M'_Z = 1000$ GeV, therefore out of the 4 top quarks the two which come from Z' decay will be in general more boosted ($\sim \frac{M'_Z}{2}$)as compared to other two quarks which mainly come from quark or gluon interaction vertex. Fig. 11 shows the p_T distribution of all 4 top quarks in the final state, the tops from Z' are indeed boosted more and the distribution pf p_T tends to peak around 400 GeV. Note in the fig 9 the reconstruction effciencies of χ^2 and HEPTopTagger cross also cross around $p_T^{top} \approx 400$ GeV and beyond this value HEPTopTagger performes better. Therefore we use HEPTopTagger in it's default settings to reconstruct two hardest top fatjets in the final state of the event, and for the other two top quarks we do a χ^2 minimization (with tagging criteria being $\chi^2_{min} < 2$) to reconstruct them from the remainder of the jets. Note that the decay width of Z' is about 80 GeV in the B-L model that we are working with therefore in order to tag Z' the invaiant mass of the two hardest top tagged jets should lie the mass window $(600, 1400)$ GeV.

[5] We could in principle do the reconstruction in the processes of the type $pp \to zp \to t\bar{t}$, but the 2 top QCD background has a very large cross section of about 598 pb.

The analysis strategy for reconstructing the events is as follows:

- Cluster C/A fatjets with $p_T > 300$ GeV and R=1. Perform top tagging on these fatjets using HEPTopTagger if there are more than one fatjet in the event otherwise reject the event.
- If in an event there are more than two tagged top jets then calculate their invariant mass $M_{j1,j2}$. Keep the event if this value lies in the Z' mass range $(600, 1400)$ GeV and remove the two tagged top jets from the jet list.
- Recluser the remaining jets with anti-kT clustering with R= 0.5. Perform top tagging on these AK5 jets using chi^2 minimization.
- Events in which all 4 tops are tagged along with Z' are the events which are reconstructed, reconstruction efficiency is denied to be the fraction of total number input events which reconstructed.

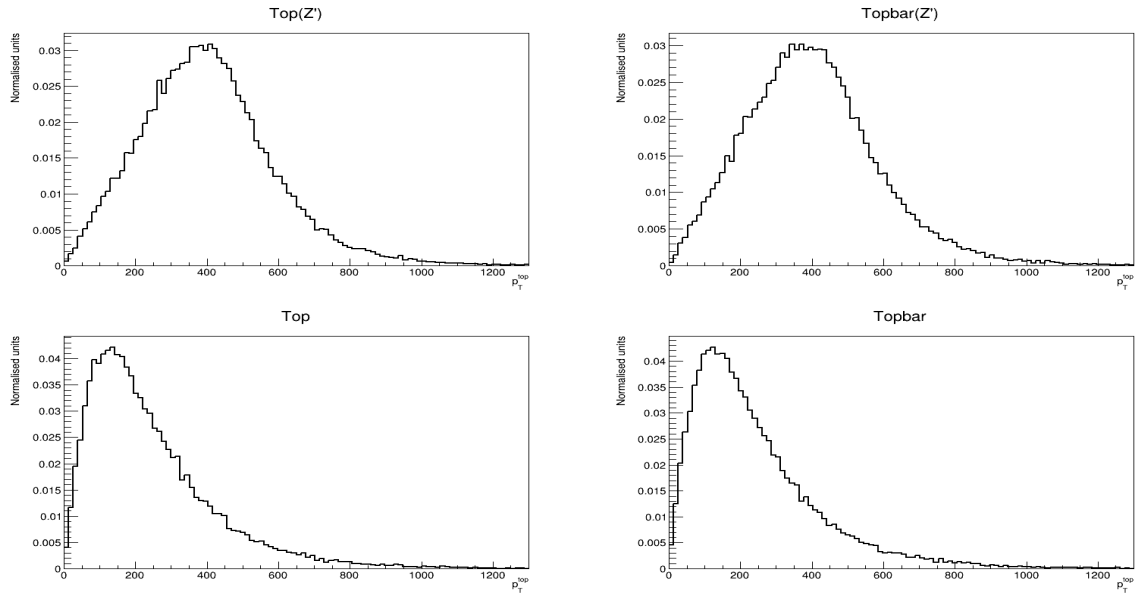

Figure 11: p_T distribution of the top final state top quarks in the events $pp \to Z't\bar{t} \to t\bar{t}t\bar{t}$. The upper two histograms labelled Top(Z') and Topbar(Z') are distributions of the tops from the Z' decay.

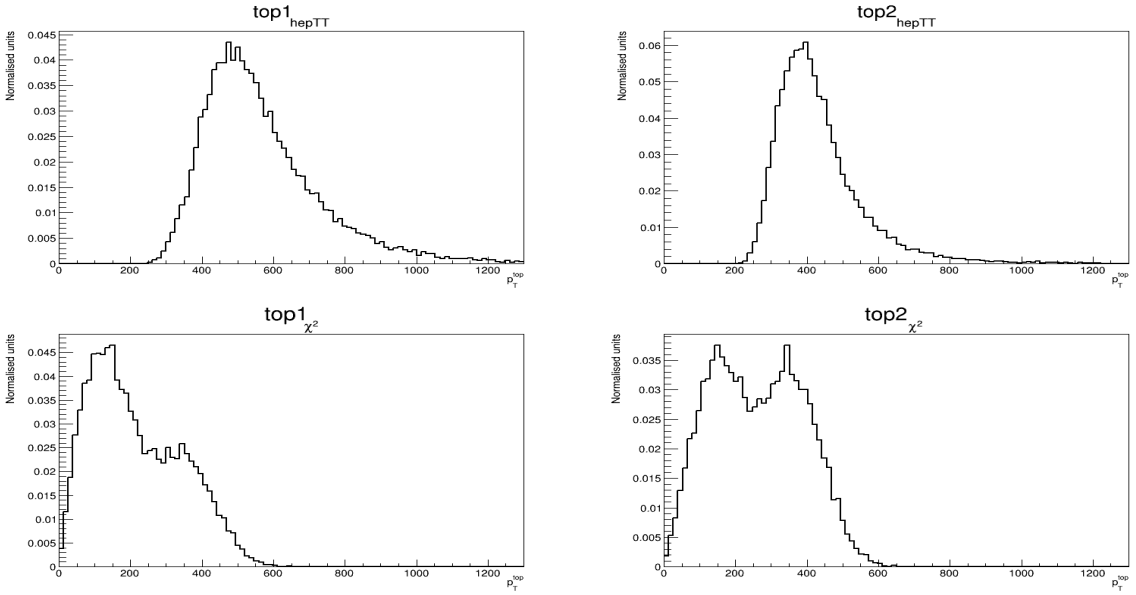

Figure 12: p_T distribution of the tagged top jets. Upper two plots are of the tops which are tagged by HepTopTagger, these also happen to be the p_T distrubtion of leading and subleading jet just with a scale factor.

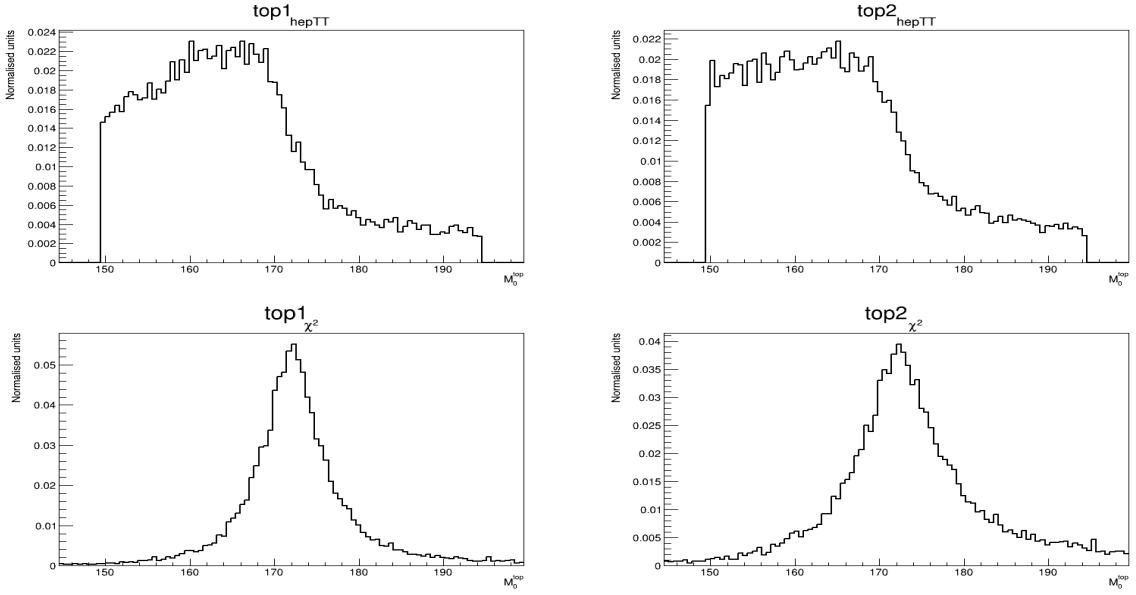

Figure 13: Invariant mass distribution of the tagged top jets. Upper two plots are of the tops which are tagged by HepTopTagger. Note that HEPTopTagger actually reconstructs top mass to be somewhat less than the actual value.

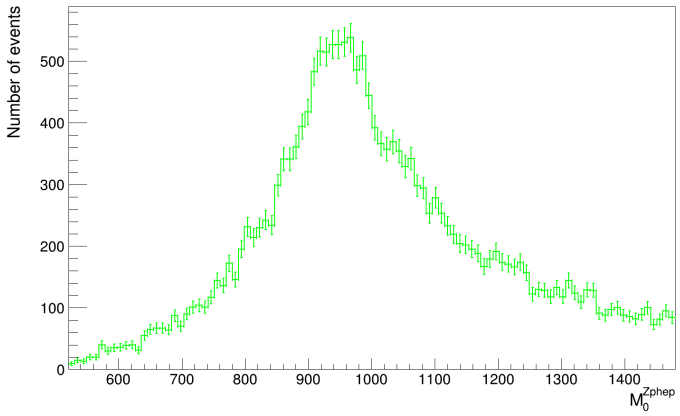

Figure 14: Distribution of the reconstructed mass of Z' in signal events.

5 QCD 4 Top background

The 4 top background from QCD at LHC has a total cross section of about 2 fb to leading order [6] and is the most dominant source of SM background $O(\alpha_s^4)$ when it comes to study BSM physics in 4 top channels, with a production threshold of $E > 4m_t \sim 700$ GeV, these processes contribute about 90% to the total background. The second most dominant background is from production of Higgs in assosiation with top quark pair $O(\alpha_s^2 y_t^2)$ contributing the rest 10% . This also goes to show why one would want to study assosiated production of Z' with top pairs instead of simply analysing events of the type $pp \to Z' \to t\bar{t}$. This process has two top QCD background with a cross section of about 300pb [30] which is five orders of magnitude more than 4 top QCD cross section. The cross section of $Z't\bar{t}$ events in pp collisions at $\sqrt{s} = 14$ TeV is 0.18 fb [7]. Fig 16 shows the p_T distribution of the 4 hardest jets in 4 top events. [29]. Background levels are suppressed by a factor of about 500 using all the criteria for signals.

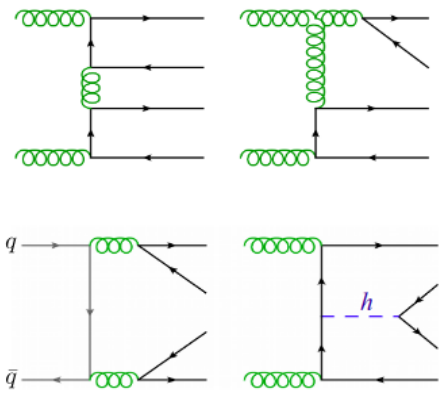

Figure 15: Underlying event topologies to LO for 4 top production in standard model.

[6] And about 12fb to NLO , but we only only generated events to leading order accuracy.
[7] All the cross sections are calculated in MADGRAPH@NLO, unless otherwise cited.

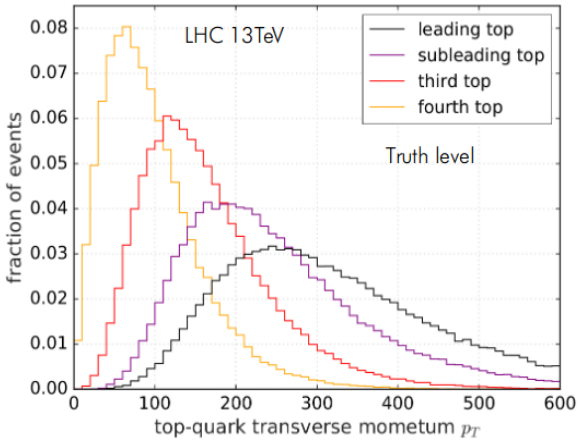

Figure 16: p_T distribution for 4 leading jets in $pp \to t\bar{t}t\bar{t} \to 4$ top jets + X [31].

6 Results and conclusion

In this work we analysed in some details top quark reconstruction techniques for hadronically decaying top quakrs, both for low boosted tops as well as tops with high p_T. We used jet substructure techniques for reconstructing highly boosted top quarks . We compared their reconstruction efficiencies as well as their mistagging rate with respect to QCD background and found jet substructure method to perform better for top with high boost. We then apply these techniques to recontruct line shape of hypothetical Z' boson in the events $pp \to Z't\bar{t} \to t\bar{t}$ and study the corresponding 4 top QCD background. The following table shows the results of the selection cuts applied on signal and background. Total number of events generated for signal and background are 5×10^5, and 2×10^5 respectively. In order to have good statistics we generate only events where all the final state top quarks decay hadronically. Signal to background ratio for is found to have a value of 0.051 .

Table 3: Selection cut flow for signal and QCD background events.

Selection cut	**Z′tt̄ → 4 top jets**	**QCD 4 top events**
Number of events generated	5×10^5	2×10^5
Atleast two CA10 fatjets with $p_T^{jet} > 300$ GeV	459328	128471
Atleast two top Tagged fatjets with HEPTopTagger	22362	1655
$\lvert M'_Z - m_{j1,j2}\rvert < 400$ GeV	19533	687
Atleast two top tagged jets from the list of remaining jets with χ^2 minimization	18179	640
Number of selected events	18179	640
Acceptence	3.6%	0.3%

Finally we scale the singal and background levels for different luminosities and calculate the signal purity and the significance for integrated luminosity of 100 fb^{-1} and 3000 fb^{-1}. Table 4 shows the results. The discovery potential for luminosities 100 fb^{-1} and 3000 fb^{-1} has a significance of 3.08 and 16.86 respectively.

Table 4: Singal purity and significance for different luminosities. For $M'_Z = 1000$ GeV.

	$\mathcal{L}_I = 100\ fb^{-1}$	$\mathcal{L}_I = 3000\ fb^{-1}$
Expected signal events	7.19	215.7
Expected background events	23	690
Singal purity $\left(\frac{s}{s+b}\right)$	0.97	0.97
Significance $\left(\frac{s}{\sqrt{b}}\right)$	3.08	16.86

7 Aknowledgement

I would like to thank everyone who helped me succesfully complete the work, and helpled me with various technical and non-technical issues. Specially my seniors Arvind, Soham bhattacharya, Nairit sur, Suman chattargee. My batch mates Mintu kumar, Arnab bhattacharya. Finally I would like to express my utmost gratitude for my guide for this project Prof. Monoranjan Guchait for giving me the opportunity to work and for the motivation.

References

[1] Langacker P., The Physics of Heavy Z' Gauge Bosons, arXiv:0801.1345, 7-16

[2] B.A. Dobrescu (Fermilab) and S. Willocq(Univ. of Massachusetts), Z'-Boson Searches, September 2017.

[3] The ATLAS Collaboration, Search for new high-mass phenomena in the dilepton final state using 36 fb^1 of protonprotoncollision data at $\sqrt{s}$ = 13 TeV with the ATLAS detector., arXiv:1707.02424v2 [hep-ex] 15 Nov 2017.

[4] CMS Physics Analysis Summary, Search for a high-mass resonance decaying into a dilepton final state in 13 fb^1 of pp collisions at $\sqrt{s} = 13$ TeV., 2016/08/05

[5] ATLAS Collaboration, Search for new phenomena in dijet events using $37fb^{-1}$ of pp collision data collected at $\sqrt{s} = 13$ TeV with the ATLAS detector., 28 March 2017.

[6] The CMS Collaboration, Search for dijet resonances in protonproton collisions at and constraints on dark matter and other models,14 February 2017

[7] R. Foot(a),X.-G. He1(b),H. Lew2(c)and R. R. Volkas3(d), Model for a LightZBoson, arXiv:hep-ph/9401250v1, 13 Jan 1994

[8] F. DEL AGUILA, The Physics of Z' bosons,Departamento de F sica Te orica y del Cosmos, Universidadde GranadaGranada, 18071, Spain, arXiv:hep-ph/9404323v1 22 Apr 1994

[9] Thomas G. Rizzo, Z' Phenomenology and the LHC, Stanford Linear Accelerator Center,2575 Sand Hill Rd., Menlo Park, CA, 94025, arXiv:hep-ph/0610104v1, 9 Oct 2006

[10] Joseph D. Lykken, Z' Bosons and Supersymmetry , Theoretical Physics Dept., MS106Fermi National Accelerator Laboratory, arXiv:hep-ph/9610218v3 25 Oct 1996

[11] David London, Jonathan L.Rosne Extra gauge bosons in E_6, 01/09/1986

[12] Julian Heeck, Unbroken B L Symmetry, arXiv:1408.6845v2 10 Nov 2014, Phys. Lett. B739, 256262 (2014)

[13] Jack Y. Araza, Gennaro Corcellab, Mariana Frankaand Benjamin Fuksc,d,e , Loopholes in Z' searches at the LHC:exploring supersymmetric and leptophobic scenarios, arXiv:1711.06302v2, 16 Feb 2018

[14] Gennaro Corcella, Searching for supersymmetry in Z' decays, arXiv:1307.1040v2, 5 Jul 2013

[15] PDF sets LHAPDF 6.2.3, https://lhapdf.hepforge.org/pdfsets

[16] LHC Constraints on a BL Gauge Model using Contur, S. Amrith, J. M. Butterworth, F. F. Deppisch, W. Liu, A. Varma, D. Yallup arXiv:1811.11452

[17] Long-lived Heavy Neutrinos from Higgs Decays, Frank F. Deppisch, Wei Liu, Manimala Mitra, arXiv:1804.04075

[18] Lorenzo Basso, Alexander Belyaev, Stefano Moretti, Claire H. Shepherd-Themistocleous, Phenomenology of the minimal B-L extension of the Standard model: Z' and neutrinos, arXiv:0812.4313

[19] The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations, J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, M. Zaro, arXiv:1405.0301

[20] Matching NLO QCD computations and parton shower simulations, S. Frixione, B.R. Webber, arXiv:hep-ph/0204244

[21] M. Tanabashiet al.(Particle Data Group), Phys. Rev. D98, 030001 (2018) and 2019 update Quarks data in PDG 2019

[22] M. Tanabashiet al.(Particle Data Group), Phys. Rev. D98, 030001 (2018) and 2019 update, W boson data in PDG 2019

[23] G. P. Salam, Towards Jetography arXiv:0906.1833v2

[24] Matteo Cacciari, Gavin P. Salam, Gregory Soyez, The anti-k˙t jet clustering algorithm arXiv:0802.1189

[25] Tilman Plehn1and, Michael Spannowsky Top Tagging , arXiv:1112.4441v1 [hep-ph] 19 Dec 2011

[26] Tilman Plehn,1Michael Spannowsky,2Michihisa Takeuchi,1and Dirk Zerwas Stop Reconstruction with Tagged Tops, arXiv:1006.2833v2 [hep-ph] 31 Aug 2010

[27] Present program version: PYTHIA8.2, T. Sjstrand, S. Mrenna and P. Skands, JHEP05 (2006) 026, Comput. Phys. Comm. 178 (2008) 852.

[28] Andrew J. Larkoski, Simone Marzani, Gregory Soyez, Jesse Thaler, SoftDrop 2.0.0 , arXiv:1402.2657

[29] PHYSICAL REVIEW D79,074012 (2009), Leandro G. Almeida,1Seung J. Lee,1,2Gilad Perez,1,2Ilmo Sung,1and Joseph Virzi, Top quark jets at the LHC

[30] ATLAS-CMS recommended predictions for single-top cross sections using the Hathor v2.1 program , SingleTopRefXsec (2017-09-19, CarlosEscobar)

[31] Ezequiel Alvarez, Darius A. Faroughy, Jernej F. Kamenik, Roberto Morales, Alejandro Szynkman Four tops for LHC. Nuclear Physics B Volume 915, February 2017, Pages 19-43

[32] Patrick J. Fox, Ian Low, Yue Zhang, Top-philic Z'forces at the LHCarXiv:1801.03505

List of Figures

List of Tables