# Search for Heavy Stops with Merged Top Jets

Yang Bai, Joshua Berger, James Osborne, and Ben A. Stefanek

# Search for Heavy Stops with Merged Top-Jets

Yang Bai, Joshua Berger, James Osborne and Ben A. Stefanek

*Department of Physics, University of Wisconsin-Madison, Madison, WI 53706, USA*

### Abstract

We study an interesting region of phase space at the LHC for pair-produced stops decaying into hadronic top quarks and light neutralinos. After imposing a sizeable cut on the missing transverse energy, which is the key variable for reducing backgrounds, we have found that the two hadronic tops are likely to merge into a single fat jet. We develop a jet-substructure-based strategy to tag the two merged top-jets and utilize the $M_{T2}$ variable to further reduce the backgrounds. We obtain about a 40% increase to the ratio of the signal over background and a mild increase on the signal discovery significance, based on a signal with a 1.2 TeV stop and a 100 GeV neutralino, for the 13 TeV LHC with 100 fb$^{-1}$. The general event kinematics could also occur and be explored for other new physics signatures with large missing transverse energy.

# 1 Introduction

While the discovery of the Higgs boson [1, 2] was a major success for the Standard Model (SM), open theoretical questions remain. If the Higgs boson is an elementary scalar field, the stability of the electroweak scale against radiative corrections is not understood, since fields of this type receive quadratically divergent corrections to their mass-squared. Of the proposed solutions to this well known hierarchy problem, one of the best motivated is supersymmetry (SUSY), where every SM fermion(boson) is complemented with a superpartner boson(fermion). Elementary scalar masses are protected because a cancellation occurs between the quadratically divergent correction coming from loops with SM particles and their corresponding superpartners enforced by the enlarged symmetry of the model. Nevertheless, SUSY must be broken in order to be consistent with non-observation of superpartners to date. This breaking lifts the superpartner masses to a relatively high scale set by the SUSY breaking sector and messenger mass [3, 4].

As the superpartners become more massive, the expected mass of elementary scalars goes up and fine-tuning is likely required to explain light scalars. Since the Higgs boson receives the largest correction to its mass from a loop involving top quarks, the most important SUSY particle for protecting the Higgs boson mass is the superpartner of the top quark–the top squark or stop [5, 6]. In the Minimal Supersymmetric Standard Model (MSSM), the electroweak scale is dominantly set by stop masses and mixing, in addition to the tree level mass scales in the superpotential and soft SUSY breaking potential. Thus, the natural expectation is that the lightest third generation squarks should be observable at the Large Hadron Collider (LHC), since the mass scales that set the electroweak scale in the MSSM must not be too far away from the observed electroweak scale in order to avoid large amounts of fine tuning. While a considerable amount of natural MSSM parameter space has been excluded by the most recent results from the LHC [7, 8], viable regions remain, such as the compressed region [9–17]. These regions are likely difficult to achieve in a top down SUSY breaking model.

In this paper, we accept some fine tuning and focus on a new method to search for pair-produced stops with a mass near 1 TeV which decay to top quarks and light neutralinos, a scenario which has not yet been excluded by the LHC Run 2. We assume the neutralino to be the lightest supersymmetric particle (LSP) with a mass around 100 GeV, which also makes it a possible thermal relic weakly interacting massive particle (WIMP) dark matter candidate. The signal that our new method will be applicable to is $\bar{t}t + \not{E}_{\mathrm{T}}$ (see Ref. [18–23] for earlier studies of two hadronic tops), where the tops decay hadronically and the missing transverse energy $\not{E}_{\mathrm{T}}$ comes from the two neutralinos which leave the detector. Currently, for searches with two hadronic tops, the ATLAS collaboration has imposed a

constraint which requires the stop mass to be above 820 GeV with 13.3 fb$^{-1}$ of integrated luminosity, assuming 100% decay branching for the channel $\widetilde{t} \to t \, \widetilde{\chi}_0$ [24] and a light neutralino mass below around 200 GeV. Similarly, the CMS collaboration has obtained a limit of 860 GeV with 12.9 fb$^{-1}$ data [25]. As the $\not{E}_T$ cut is increased to optimize for heavier stops, it becomes increasingly likely that the two tops in the decay have a small angular separation and that the signals of their decay products overlap in the detectors.

We develop a new boosted top tagging procedure that recovers merged top jets, which we call "Merged Top Tagger." Our starting point is the well-known HEPTopTagger algorithm [18,26]. Rather than simply trying to find a single combination of subjets in a fat jet that looks like a top, the algorithm searches through various combinatoric possibilities in an attempt to find a total of two top-like groups of subjets. Further sensitivity is gained by using the new information of both tops' kinematics to construct and cut on the $M_{T2}$ variable [27]. Based on simulations we have performed, our strategy shows $\sim 40\%$ improvement in $S/B$ and slight improvement in discovery sensitivity relative to a strategy based on the HEPTopTagger algorithm for 1.2 TeV stops decaying to tops and 100 GeV detector-stable neutralinos.

The remainder of this paper is structured as follows. In Section 2, we motivate the need for a new analysis strategy to deal with merged hadronic tops. We present our Merged Top Tagger algorithm in Section 3. We then study the increased sensitivity to search for heavy stops in Section 4. Finally, we discuss and conclude in Section 5. We confirm that standard jet analysis strategies break down in Appendix A. The simulation details are presented in Appendix B.

## 2  Motivation for Merged Top-Jets

For the region of parameter space with $m_{\widetilde{t}} - m_{\widetilde{\chi}_0} \gg m_t$, the top quark is boosted and the three partons from the top quark decay are collimated. In similar scenarios with boosted top quarks, both collaborations at the LHC have opted to reconstruct events using jet-substructure techniques such as top-tagging. Assuming conservation of R-parity, the two neutralinos in the final state will be stable and contribute significantly to the total missing transverse energy, $\not{E}_T$. Because of this fact, imposing a large $\not{E}_T$ cut is a very efficient way to increase the signal over background ratio. However, we find that imposing a large $\not{E}_T$ cut isolates the sub-region of phase space of the signal where the neutralinos are approximately aligned to provide large $\not{E}_T$, resulting in the tops (which recoil off the neutralino momenta) also being approximately aligned. Because the pair-produced stops are not very relativistic, the top and neutralino pairs end up approximately back to back. As a result, one should anticipate that a significant fraction of signal events will have geometric overlapping between the six partons

from two hadronic tops, a scenario we call *merged top-jets*. The schematic picture of the region of phase space for signal events under discussion is shown in Fig. 1. The subject of this paper will be a new search strategy for the merged top-jets case, with the goal of improving the discovery potential of stops at the LHC.
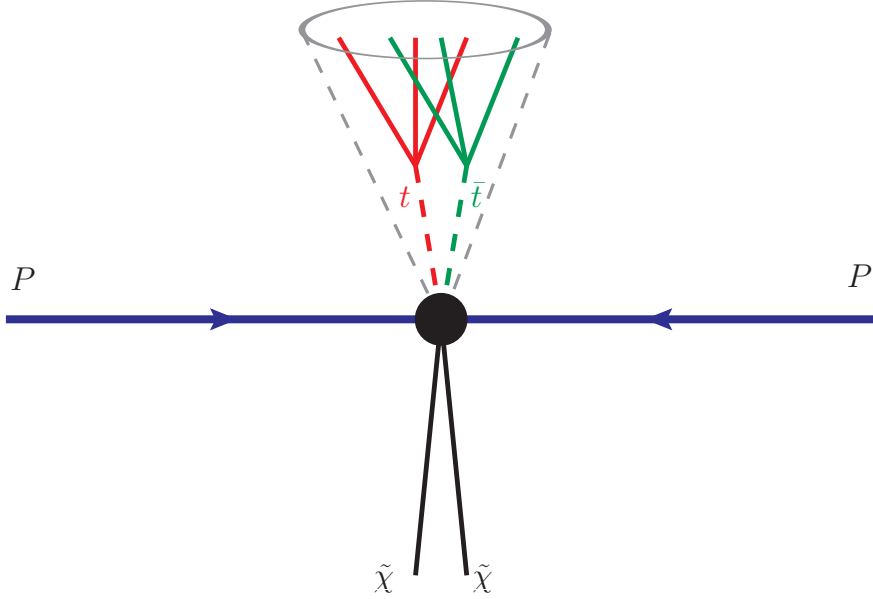


Figure 1: The schematic plot of the region of phase space for signal events with a large cut on $\not{E}_{\mathrm{T}}$. To provide large $\not{E}_{\mathrm{T}}$, the neutralino momenta must approximately align. This also results in alignment of the tops, which recoil off the sum of the neutralino momenta. When performing jet-substructure analysis with a large value of the jet clustering parameter $R$, the six partons from the two hadronic tops are likely to be included in a single fat jet.

To quantitatively understand the fraction of signal events with two merged tops, we first study the signal events at the parton level. Throughout this paper, we will assume benchmark masses of $m_{\tilde{t}} = 1.2$ TeV and $m_{\tilde{\chi}} = 100$ GeV, and $\mathrm{Br}(\tilde{t} \to t + \tilde{\chi}) = 100\%$. The lightest stop mass is chosen to be close to the reach of the LHC Run 2 with 100 fb$^{-1}$. In our simulation, we have the stop to be mainly right-handed, although our later kinematic analysis is insensitive to this choice (for the detailed analysis for left-handed and right-handed stops, see Ref. [28]). In the left panel of Fig. 2, we define a measure to demonstrate the alignment of the two tops as the cut on $\not{E}_{\mathrm{T}}$ is increased. To show this, we define $\Delta R_{\mathrm{min},t\bar{t}}$ as the minimum geometrical separation among all pairs with one parton from the top and one parton from the anti-top. Then we define $\max(\Delta R_{\mathrm{max},tt}, \Delta R_{\mathrm{max},\bar{t}\bar{t}})$ as the maximum separation among all pairs with two partons belonging only to the top or anti-top. The ratio of these two quantities, $\Delta R_{\mathrm{min},t\bar{t}} / \max(\Delta R_{\mathrm{max},tt}, \Delta R_{\mathrm{max},\bar{t}\bar{t}})$, will be small when the smallest distance between
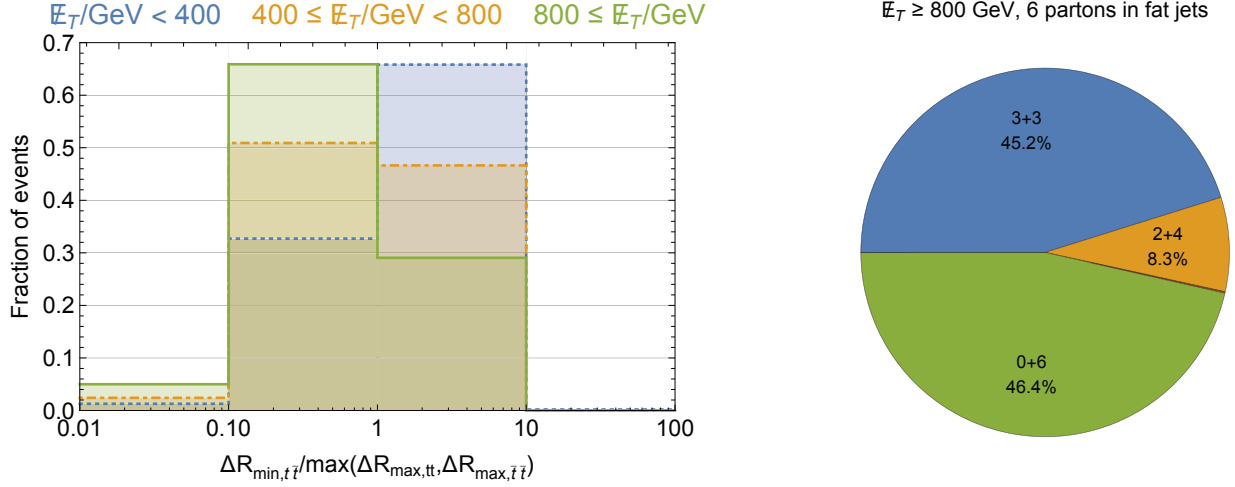
3

Figure 2: Left panel: The fraction of signal events as a function of a variable (see text for definition) which measures the overlapping of partons from the two tops: a smaller value means a larger overlap. Right panel: The distribution of partons in a $R = 1.5$ Cambridge/Aachen fat jet and $p_T > 200$ GeV. Not shown are the negligible fractions of events with 1+5 and 1+2+3.

one parton from the top and one parton from the anti-top is smaller than the largest distance between partons in the same top or anti-top, which is the aligned or merged scenario. As a result, we expect the event distribution of this measure to shift to smaller values as the cut on the missing transverse energy is increased, and this expectation is confirmed in the left panel of Fig. 2.

It is immediately plausible that the high $p_{\mathrm{T}}$ tops produced in these events cannot have their decay products resolved into a total of 6 standard jets, motivating the use of a fat jet analysis in this case. We confirm that standard jet analysis strategies break down in Appendix A. When two partons from the top and anti-top are geometrically close to each other, the standard jet substructure analysis to tag boosted top jets becomes problematic. This is because the large fat jet analysis employed frequently fails to isolate the two tops in separate fat jets. To demonstrate this issue at parton level, we apply the Cambridge/Aachen clustering algorithm [29] with $R = 1.5$ to the signal events and count the number of partons belonging to the top or anti-top contained in each fat jet. In the right panel of Fig. 2, one can see that the fully merged 0+6 case (all partons in one fat jet) and the partially merged 2+4 case (2 partons in one fat jet, 4 partons in another) account for 54.7% of the events, with the separated 3+3 case accounting for 45.2% of the events. The 1+5 and 1+2+3 cases are extremely sub-dominant due to the fat jet definition including a transverse momentum cut of $p_T > 200$ GeV.

It is now clear that the standard jet-substructure analysis is not optimal for reconstructing our signal events after a large $\not{E}_{\mathrm{T}}$ cut, since standard top taggers (which assume well isolated tops by only

4

searching for one top tag per fat jet) frequently fail to tag two tops in the fully or partially merged cases which constitute a majority of our signal events after a large $\displaystyle{\not}E_{\mathrm{T}}$ cut. To fully recover the signal kinematics and define additional transverse-mass variables like $M_{T2}$, it is important to resolve and tag the two top-jets with high efficiency. In next section of this paper, we develop a search strategy based on the HEPTopTagger [18, 26] algorithm and check the improvement of signal over background after the particle-level simulations.

## 3   The Merged Top Tagger Algorithm

Having established our motivation for tagging two merged top-jets, we now develop a jet-substructure based algorithm to reconstruct both boosted hadronic tops for the signal and background. Our algorithm for tagging two merged top-jets is based on the existing top-tagging algorithm of HEP-TopTagger in Ref. [18], although it can also be implemented using other top-tagging methods such as N-subjettiness [30]. Before discussing our algorithm, we briefly summarize the features of HEP-TopTagger. By default, HEPTopTagger takes a single fat jet as an input, on which it will perform a mass drop [31] operation to obtain a list of the relevant hard substructure. It then iterates through all pairings of three hard subjets and filters [31] each triplet in order to define the triplet mass. It then keeps *only* the triplet which has a filtered mass closest to the top mass $m_t$. This single remaining triplet is then required to pass additional mass constraints which ensures the triplet does not behave like a QCD event and that it satisfies some top and $W$ gauge boson mass constraints.

For our signal events, after a large $\displaystyle{\not}E_{\mathrm{T}}$ cut, the two top-jets have a significant probability of being merged, and the original HEPTopTagger algorithm becomes less efficient. This is because the original HEPTopTagger algorithm can find at most one top per fat jet, even in the case where the input fat jet contains two tops. Furthermore, even in the one top per fat jet case, the original algorithm will fail when the triplet with the best filtered mass fails the mass criteria, even though there may exist other triplets with filtered masses still not that far away from the top quark mass, but a better chance to pass all mass criteria. To address those issues, we introduce the following "Merged Top Tagger" algorithm with an aim to tag two merged top-jets:

1. Using the Cambridge/Aachen algorithm with $R = 1.5$, identify initial fat jets with $p_T(J) > 200$ GeV.

2. For the leading $p_T$ fat jet, find all hard subjets using a mass drop criterion: when undoing the last clustering of the jet $j$, into two subjets $j_1$, $j_2$ with $m_{j_1} > m_{j_2}$, we require $m_{j_1} < 0.8\, m_j$ to

keep $j_1$ and $j_2$. Otherwise, we keep only $j_1$. For each subjet $j_i$, we further de-cluster it until its jet mass is 30 GeV or below.

3. Iterate through all possible triplets of hard subjets (HSJ triplets) found in Step 2: first, filter them with resolution $R_{\text{filter}} = \min(0.3, \Delta R_{jk}/2)$. Next, use the five hardest filtered constituents and calculate the triplet mass $m_{\text{filtered}}$ (for less than five filtered constituents use all of them). Keep all HSJ triplets which satisfy $m_{t_{\min}} < m_{\text{filtered}} < m_{t_{\max}}$ (Default: $m_{t_{\min}} = 140$ GeV and $m_{t_{\max}} = 250$ GeV).

4. For each HSJ triplet saved from Step 3, re-cluster the five filtered constituents into exactly three subjets $j_1$, $j_2$, $j_3$, ordered by $p_T$. If the masses ($m_{12}$, $m_{13}$, $m_{23}$) satisfy one of the following three criteria, save the filtered HSJ triplet as a top candidate:

$$0.2 < \arctan\left(\frac{m_{13}}{m_{12}}\right) < 1.3 \qquad \text{and} \qquad R_{\min} < \frac{m_{23}}{m_{123}} < R_{\max},$$

$$R_{\min}^2 \left[1 + \left(\frac{m_{13}}{m_{12}}\right)^2\right] < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_{\max}^2 \left[1 + \left(\frac{m_{13}}{m_{12}}\right)^2\right] \qquad \text{and} \qquad \frac{m_{23}}{m_{123}} > 0.35,$$

$$R_{\min}^2 \left[1 + \left(\frac{m_{12}}{m_{13}}\right)^2\right] < 1 - \left(\frac{m_{23}}{m_{123}}\right)^2 < R_{\max}^2 \left[1 + \left(\frac{m_{12}}{m_{13}}\right)^2\right] \qquad \text{and} \qquad \frac{m_{23}}{m_{123}} > 0.35.$$

with $R_{\min} = (1 - f_W) \times M_W/m_t$ and $R_{\max} = (1 + f_W) \times M_W/m_t$ (we will choose the value $f_W = 0.5$ [32] to increase signal efficiencies). The number 0.35 is chosen to help remove QCD events. The above selection criteria are identical to HEPTopTagger [18].

5. If there is more than one top candidate from Step 4, check all two pairings of top candidates and keep any for which their HSJ triplets share no subjets (unique pairs). If only one unique pair is found, return it as two tagged tops. If more than one unique pair is found, return the pair of top candidates which minimizes the quantity $|m_{J_{1,\text{filtered}}} - m_t| + |m_{J_{2,\text{filtered}}} - m_t|$ as tagged tops. Otherwise, continue to Step 6.

Up to this point, the modified HEPTopTagger algorithm has only been extended to deal with the fully merged case when two tops are contained within a single fat jet. To deal with the partially merged case when some partons from the first top are clustered into the fat jet of the second top, we also introduce the following steps to capture particles in the vicinity of the leading fat jet, which we accomplish by removing the particles belong to the leading top candidate(s) and reclustering the event. Specifically, we extend the algorithm in the following way:

6. If at least one top candidate exists from Step 4, remove the particles associated with one top candidate from the final state particles. The remaining particles are then reclustered using the same Cambridge/Aachen algorithm with $R = 1.5$. For the leading $p_T$ fat jet found after this reclustering (if any), repeat the above Steps 2-4 to identify more top candidates. Of all top candidates found in the reclustered fat jet, keep the one which has a filtered mass closest to $m_t$ (if any), and pair it with the top candidate before the new reclustering procedure. For the cases with more than one top candidate from Step 4, perform this procedure on all of them, and return the pair which minimizes the quantity $|m_{J_{1,\text{filtered}}} - m_t| + |m_{J_{2,\text{filtered}}} - m_t|$ as tagged tops. If no top candidate pairs are found, continue to Step 7.

7. If the two leading $p_T$ fat jets from the initial event clustering have not both already been analyzed, repeat Steps 2-6 on the next leading initial fat jet by $p_T$, if it exists. If no initial fat jets remain or if the two leading $p_T$ initial fat jets have already been analyzed, then the algorithm has failed to tag two tops. Of all the top candidates from the leading $p_T$ initial fat jet, return the one that minimizes $|m_{J_{\text{filtered}}} - m_t|$ as a single tagged top. If there are no top candidates in the leading $p_T$ initial fat jet, use the next leading $p_T$ initial fat jet. If there are no top candidates in any initial fat jet, the algorithm has failed to tag a single top.

Not only does our modified algorithm allow two tops to be tagged in the fully and partially merged cases, it also finds the combination of hard subjets such that both tops are as close to the true top mass as possible. To quantify the improvement our algorithm offers for tagging merged top-jets, we define the following efficiency parameter

$$\mathcal{E}_{N_t \geq 2} \left( \not{E}_T \text{ range} \right) \equiv \frac{\text{Events with} \geq 2 \text{ top tags in } \not{E}_T \text{ range}}{\text{All events in } \not{E}_T \text{ range}}, \tag{1}$$

which we can use to gauge the relative performance of our modified algorithm versus the original HEP-TopTagger algorithm over different ranges of missing transverse energy. According to our hypothesis, a larger cut on $\not{E}_T$ should increase the number of merged top signal events, pushing more signal events into the region where the original algorithm cannot tag two tops. As a result, we expect the ability of the modified algorithm to tag two tops to increase relative to the original algorithm for higher missing energy windows, an expectation which is confirmed in Table 1. Also shown in Table 1 are the two top-tagging efficiencies for the leading backgrounds, $\bar{t}t$ s.l. + jets (one hadronic top and one leptonic top) and $\bar{t}t + Z$ (two hadronic tops, $Z$ decays to two neutrinos). For the $\bar{t}t$ s.l. + jets background, s.l. indicates semi-leptonic decay of the tops. Futhermore, we analyze the cases where the lepton is an electron or muon vs. a tau separately and denote these backgrounds as $\bar{t}t$ s.l. + jets (e,$\mu$) and $\bar{t}t$ s.l. +

| | Signal | | $\bar{t}t$ s.l. + jets (e, $\mu$) | | $\bar{t}t$ s.l. + jets ($\tau$) | | $\bar{t}t + Z$ | |
|---|---|---|---|---|---|---|---|---|
| Algorithm | $\not{E}_\mathrm{T} > 400$ GeV | $\not{E}_\mathrm{T} > 800$ GeV | $\not{E}_\mathrm{T} > 400$ GeV | $\not{E}_\mathrm{T} > 800$ GeV | $\not{E}_\mathrm{T} > 400$ GeV | $\not{E}_\mathrm{T} > 800$ GeV | $\not{E}_\mathrm{T} > 400$ GeV | $\not{E}_\mathrm{T} > 800$ GeV |
| Original | 18% | 14% | 6.0% | 7.1% | 21% | 24% | 8.4% | 11% |
| Modified | 48% | 55% | 18% | 26% | 12% | 18% | 23% | 41% |
| Ratio | 2.7 | 3.9 | 3.0 | 3.7 | 0.57 | 0.75 | 2.7 | 3.7 |

Table 1: The efficiencies, $\mathcal{E}_{N_t \geq 2}$ (after vetoing events with isolated leptons or tagged taus), to tag two top-jets using the original HEPTopTagger and our Merged Top Tagger. The signal events have $m_{\tilde{t}} = 1.2$ TeV and $m_{\tilde{\chi}} = 100$ GeV. Only the dominant backgrounds are shown here for comparison.

jets ($\tau$), respectively. Having discussed the tagging efficiency for our new algorithm, we turn now to estimate the improvement for the signal over background ratio and the signal discovery significance.

## 4   Estimation of Discovery Significance

In this section, we estimate the discovery sensitivity of stops in the fully-hadronic channel at the LHC Run 2 with 100 fb$^{-1}$. The details of our numerical simulations are presented in Appendix B. After the application of a large cut on missing transverse energy, we find the leading backgrounds are $\bar{t}t$ s.l. + jets (one leptonic top and one hadronic top) and $\bar{t}t + Z$ (two hadronic tops, $Z$ decays to two neutrinos). The basic cuts to select events impose the following requirements:

(a) Missing energy $\not{E}_\mathrm{T} > 400$ GeV.

(b) Using our modified algorithm, require at least two top-tagged jets with $p_T(J) > 200$ GeV.

(c) At least two anti-$k_\mathrm{T}$, $R = 0.4$ $b$-tagged jets with $p_T$ above 30 GeV.

(d) Veto events with isolated leptons with $p_{T,e,\mu} > 15$ GeV and $|\eta_{e,\mu}| < 2.4$.

(e) Veto events with a tagged tau with $p_{T,\tau} > 15$ GeV and $|\eta_\tau| < 2.4$.

For simplicity, we have implemented a $b$-tagging efficiency of 0.8 (mistag efficiency of 0.2 for charm quarks and 0.05 for light quarks and gluon) [24, 33] in our simulation. One could also perform direct $b$-tagging on the subjets of the fat jets as in Ref. [32]. Here, the main requirement of our simplistic and conservative $b$-tagging approach is to sufficiently reduce the $W/Z$ + jets backgrounds by requiring two $b$-tags in the event. Information on our tau tagging and lepton isolation criteria can be found in Appendix B.

After the basic cuts, we note three powerful kinematical variables which are useful for reducing backgrounds. The most obvious one is $\not{E}_\mathrm{T}$, which is shown in the left panel of Fig. 3. Another variable

is the transverse mass associated with the $b$-jets and $\not{E}_T$, which has already been adopted by both CMS [25] and ATLAS [24]. It is given by

$$M_T^b = \min\left[M_T(\vec{p}_{b_1}, \vec{\not{E}}_T), M_T(\vec{p}_{b_2}, \vec{\not{E}}_T)\right].\tag{2}$$

The transverse mass is defined as $M_T(\vec{p}_{b_i}, \vec{\not{E}}_T) = 2p_T^{b_i}\not{E}_T[1-\cos(\phi_{b_i} - \phi_{\not{E}_T})]$, where $\phi_{b_i}$ and $\phi_{\not{E}_T}$ are the azimuthal angles of the $b$-jet and $\vec{\not{E}}_T$. In the right-panel of Fig. 3, we show the signal and background distributions for this variable. As one can see, the $\bar{t}t$ s.l. + jets background has an endpoint around
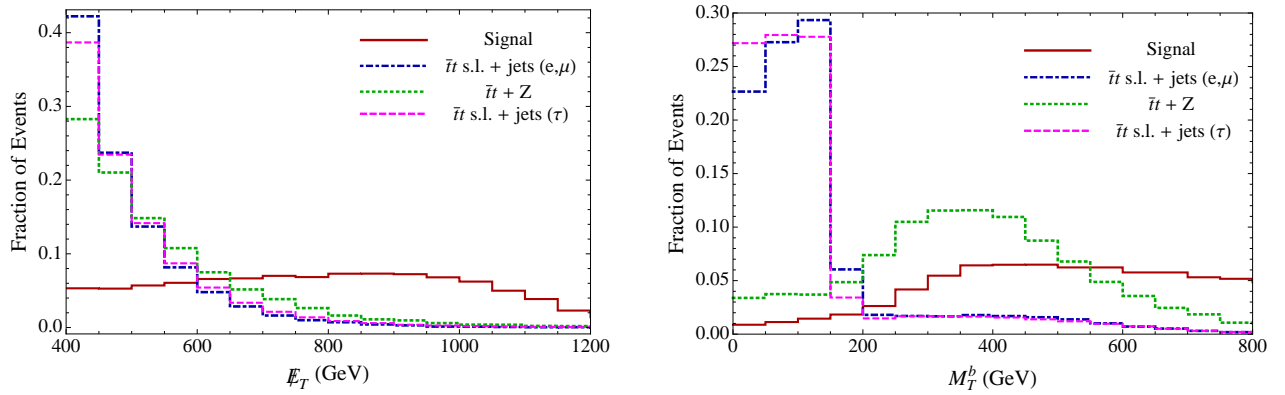


Figure 3: Left panel: The normalized event distributions in $\not{E}_T$ for the signal and leading backgrounds after basic cuts. Right panel: Same as the left but in terms of $M_T^b$. The signal has $m_{\tilde{t}} = 1.2$ TeV and $m_{\tilde{\chi}} = 100$ GeV.

the top quark mass, so requiring a sufficiently large $M_T^b$ can dramatically reduce this background. We also note that the tail of the $M_T^b$ distribution for $\bar{t}t$ s.l. + jets comes from mis-tagged $b$-jets, which do not come from top quark decays and therefore are not subject to this kinematic cutoff. The final variable is the minimum transverse mass that can be computed between the event jets and the missing energy $\not{E}_T$. We define this variable as

$$M_T^j = \min\{M_T(\vec{p}_{j_1}, \not{E}_T), ..., M_T(\vec{p}_{j_N}, \not{E}_T)\},\tag{3}$$

where $N$ is the number of untagged anti-$k_T$ jets in the event with $p_T > 15$ GeV. This variable is useful for reducing the $\bar{t}t$ s.l. + jets ($\tau$) background where the transverse mass between a jet corresponding to an untagged hadronic tau and the missing energy from the tau neutrino is limited by the $W$ boson mass. It is also useful for reducing $\bar{t}t$ s.l. + jets ($e, \mu$) events which pass the basic cuts because in order to fake a hadronic top, there should be two jets close to the leptonic top and therefore the missing energy. Since the signal topology prefers to have all jets opposite the missing energy, this is a useful kinematic variable. The distribution for $M_T^j$ is shown in the left panel of Figure 4.

9

With our new tagging algorithm, one can define additional kinematic variables based on the two top quarks in the final state. With the large missing transverse energy at hand, one natural choice is an $M_{T2}$-like variable. Specifically for events with two tops and missing transverse energy, one can define the following $M_{T2}^{\bar{t}t}$ variable as [27, 34, 35]

$$M_{T2}^{\bar{t}t} = \min \left\{ \bigcup_{\vec{p}_1^T + \vec{p}_2^T = \vec{\not{E}}_T} \max\left[ M_T(\vec{p}_{t_1}, \vec{p}_1^T), M_T(\vec{p}_{t_2}, \vec{p}_2^T) \right] \right\} . \qquad (4)$$

In Fig. 4, we show the distribution for $M_{T2}^{\bar{t}t}$ for both the signal and leading backgrounds. The signal event distribution has a peak feature with an end-point of around 1.2 TeV, which can be easily understood from the stop mass and the large missing energy cut of $\not{E}_T > 800$ GeV. The $\bar{t}t + Z$ background has a similar distribution to the signal events and behaves as an irreducible background. For the $\bar{t}t$ s.l. + jets background, two peak structures appear in the distribution. The peak at higher values is similar to the signal region of phase space with two top-jets moving in the same direction, while the peak at lower values is due to the region of phase space where one top-jet approximately aligns with the missing transverse momentum, leading to a smaller transverse mass. Imposing a lower
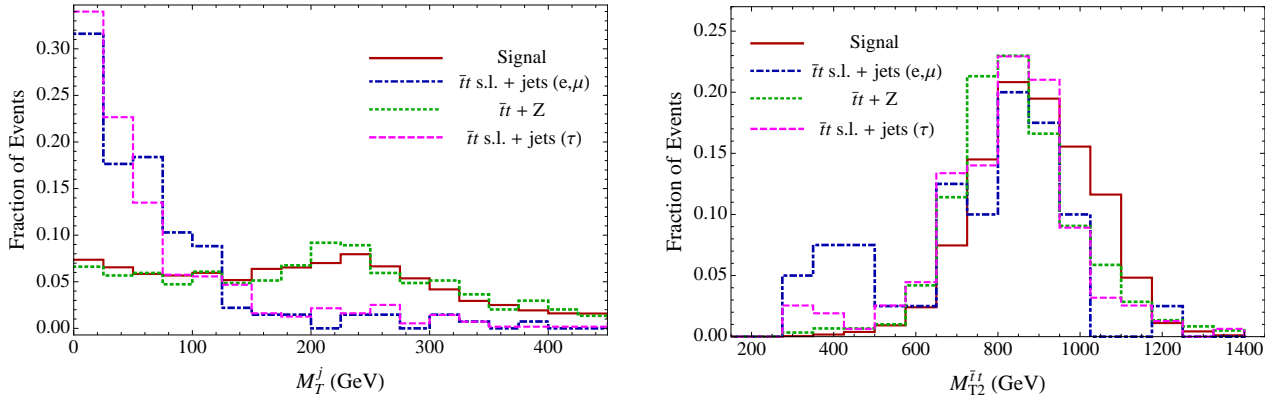


Figure 4: Left: Distribution for the $M_T^j$ variable after the optimized cuts of $\not{E}_T > 800$ GeV and $M_T^b > 220$ GeV. Right: Distribution for the $M_{T2}^{\bar{t}t}$ variable after the optimized cuts of $\not{E}_T > 800$ GeV, $M_T^b > 220$ GeV, and $M_T^j > 80$ GeV. Both plots are for a 1.2 TeV stop mass.

limit cut on $M_{T2}^{\bar{t}t}$ can therefore reduce the $\bar{t}t$ s.l. + jets background.

Using the kinematic variables outlined above, we perform an optimization of cuts to increase the expected signal discovery significance. Since we expect small numbers of signal and background events,

we use the following expected log likelihood discovery significance

$$\sigma_{\mathrm{LL}} \equiv \sqrt{-2\left[S + (B+S)\log\frac{B}{B+S}\right]}, \tag{5}$$

which approaches $S/\sqrt{B}$ for large signal and background. After optimizing the cuts on the four variables $\not{E}_{\mathrm{T}}$, $M_T^b$, $M_{T2}^{\bar{t}t}$, and $M_T^j$, we show the cut-flow of events in Table 2, which shows that requiring two top-jets from our modified tagging algorithm along with the rest of the basic cuts already reduces the dominant $\bar{t}t$ s.l. + jets background substantially. The cuts on $\not{E}_{\mathrm{T}}$, $M_T^b$, and $M_T^j$ are the most

| Modified top-tagging | after basic cuts | $\not{E}_{\mathrm{T}} > 800$ GeV | $M_T^b > 220$ GeV | $M_{T2}^{\bar{t}t} > 500$ GeV | $M_T^j > 80$ GeV | events (100 fb$^{-1}$) |
|---|---|---|---|---|---|---|
| signal | 0.17 fb | 0.086 fb | 0.084 fb | 0.083 fb | 0.066 fb | 6.6 |
| $\bar{t}t$ s.l. + jets (e, $\mu$) | 4.77 fb | 0.092 fb | 0.011 fb | $9.3 \times 10^{-3}$ fb | $2.6 \times 10^{-3}$ fb | 0.26 |
| $\bar{t}t$ s.l. + jets ($\tau$) | 9.13 fb | 0.25 fb | 0.023 fb | 0.020 fb | $6.2 \times 10^{-3}$ fb | 0.62 |
| $\bar{t}t + Z$ | 0.17 fb | 0.010 fb | $9.4 \times 10^{-3}$ fb | $9.2 \times 10^{-3}$ fb | $7.5 \times 10^{-3}$ fb | 0.75 |
| $Z$ + jets | 0.032 fb | $3.5 \times 10^{-3}$ fb | $3.2 \times 10^{-3}$ fb | $2.5 \times 10^{-3}$ fb | $1.6 \times 10^{-3}$ fb | 0.16 |
| $W$ + jets | $7.0 \times 10^{-3}$ fb | $6.0 \times 10^{-4}$ fb | $6.0 \times 10^{-4}$ fb | $6.0 \times 10^{-4}$ fb | $2.0 \times 10^{-4}$ fb | 0.02 |
| | | $S/B = 0.24$ | $S/B = 1.8$ | $S/B = 2.0$ | $S/B = 3.6$ | $S/B = 3.6$ |
| | | $\sigma_{\mathrm{LL}} = 1.4$ | $\sigma_{\mathrm{LL}} = 3.2$ | $\sigma_{\mathrm{LL}} = 3.3$ | $\sigma_{\mathrm{LL}} = 3.6$ | $\sigma_{\mathrm{LL}} = 3.6$ |

Table 2: Signal and background cross sections after cuts using our modified algorithm at the $\sqrt{s} = 13$ TeV LHC. The choice of cuts is optimized to increase $\sigma_{\mathrm{LL}}$ in Eq. (5). Here, we have $m_{\tilde{t}} = 1.2$ TeV and $m_{\tilde{\chi}} = 100$ GeV.

efficient for increasing $S/B$. As a result, large values of $S/B$ and $\sigma_{\mathrm{LL}}$ are obtained. The chosen model parameter with $m_{\tilde{t}} = 1.2$ TeV and $m_{\tilde{\chi}} = 100$ GeV will be tested at the LHC Run 2 with 100 fb$^{-1}$.

To compare to a search using the original HEPTopTagger algorithm, we replace part (b) of the basic cuts by

(b$'$) Require at least one top-tagged jet with $p_T(J) > 200$ GeV based on the original HEPTopTagger algorithm,

and show the re-optimized cut-flow of signal and backgrounds in Table 3. Comparing the results from Table 2 and Table 3, one can see that our modified top-tagging algorithm yields an obvious improvement for $S/B$ and a mild increase for $\sigma_{\mathrm{LL}}$.

# 5 Discussion and Conclusions

We note that the signal acceptance obtained when requiring two top tagged jets is not that high compared to the original HEPTopTagger algorithm with the requirement of only one top tagged jet.

11

| Original top-tagging | after basic cuts | $\not{E}_{\mathrm{T}} > 800$ GeV | $M_T^b > 280$ GeV | $M_T^j > 80$ GeV | events (100 fb$^{-1}$) |
|---|---|---|---|---|---|
| signal | 0.25 fb | 0.11 fb | 0.10 fb | 0.081 fb | 8.1 |
| $\bar{t}t$ s.l. + jets (e, $\mu$) | 14.1 fb | 0.20 fb | 0.02 fb | $6.7 \times 10^{-3}$ fb | 0.67 |
| $\bar{t}t$ s.l. + jets ($\tau$) | 26.7 fb | 0.51 fb | 0.040 fb | 0.011 fb | 1.1 |
| $\bar{t}t + Z$ | 0.45 fb | 0.016 fb | 0.014 fb | 0.011 fb | 1.1 |
| $Z$ + jets | 0.18 fb | $5.1 \times 10^{-3}$ fb | $4.8 \times 10^{-3}$ fb | $2.5 \times 10^{-3}$ fb | 0.25 |
| $W$ + jets | 0.042 fb | $8.0 \times 10^{-4}$ fb | $4.0 \times 10^{-4}$ fb | $2.0 \times 10^{-4}$ fb | 0.02 |
| | | $S/B = 0.15$ $\sigma_{\mathrm{LL}} = 1.2$ | $S/B = 1.3$ $\sigma_{\mathrm{LL}} = 3.1$ | $S/B = 2.6$ $\sigma_{\mathrm{LL}} = 3.5$ | $S/B = 2.6$ $\sigma_{\mathrm{LL}} = 3.5$ |

Table 3: The same as Table 2 but based on the original HEPTopTagger algorithm.

This is simply due to the fact that a non-negligible fraction of top quarks from 1.2 TeV stop decays are not boosted. These events can have the partons from top decays well separated from each other. For instance, the $b$-hadron could be well separated from the hadronic $W$ and may not have enough transverse momentum to satisfy the fat jet $p_T > 200$ GeV requirement. In principle, one could include those events with an ordinary $b$-jet that does not belong to any fat jet. If one wants to further increase the discovery sensitivity, one should combine events with boosted and non-boosted top quarks [25].

In addition, our algorithm should be validated with a full detector simulation before use. The simplified approach to detector effects, as well as to $b$-tagging, adopted in this work is justified by the fact that the dominant smearing of jet kinematics is coming from QCD physics, but there is some effect of calorimetric smearing as well. Our work demonstrates the gains to be had by adopting our algorithm even neglecting these effects and knowing that the final projections for discovery sensitivity are an estimation.

In this paper, we have concentrated on simplified models with stops and neutralinos. However, the key observation illustrated in Fig. 1 is general and could be applied to other models, under the condition that a sizable cut on the missing transverse energy prefers to have the visible particles collimated. Light intermediate particles in the decay chain are not necessary for our strategy to be applicable, so that could be extended to pair-produced gluinos with off-shell squark mediated decay, $\widetilde{g} \rightarrow jj\widetilde{\chi}_0$, for example. Nevertheless, our approach is ideally suited for situations in which there are a large number of partons, increasing the likelihood of particles from different sides of the event overlapping.

In summary, we have identified a new and interesting region of phase space for the heavy stop plus light neutralino model. The subset of signal events with large missing transverse momentum has a region of phase space where the partons coming from the decay of the two hadronic top quarks have a large overlapping. We have developed a jet-substructure based algorithm to identify these two

merged or semi-merged hadronic tops. Based on our estimation of signal discovery significance with only statistical errors, we have found that our algorithm can dramatically increase $S/B$ by $\sim 40\%$ and yield a mild increase in the discovery significance relative to the strategy based on the original HEPTopTagger algorithm. It is also important to point out that our new algorithm is ideal for use with additional kinematic variables which require the complete reconstruction of the two hadronic top momenta to further increase $S/B$ and the discovery significance. A stop mass of 1.2 TeV and a light neutralino will be concretely tested at the LHC Run 2 with 100 fb$^{-1}$.

## Acknowledgments

## A    Standard Jet Analysis Breakdown

In this Appendix, we provide a demonstration of the breakdown of standard jet analysis techniques. In earlier analyses, when only a smaller $\slashed{E}_{\mathrm{T}}$ cut could be sustained and lighter stops probed, it was possible to do an all hadronic stop analysis with jets of $R = 0.4$ or 0.5 as are used in many LHC analyses. At the current edge of sensitivity, this fails to adequately capture the kinematics of tops arising in stop production. To illustrate this is a simple context, we generated stop events in 13 TeV $pp$ collisions. We studied hadronic stop decays at parton level, following a similar strategy to the parton-level fat jet analyses discussed in the main text. After clustering the partons into $R = 0.4$ anti-$k_{\mathrm{T}}$ jets with a $p_{\mathrm{T}}$ cut of 50 GeV and imposing a $\slashed{E}_{\mathrm{T}}$ cut of 800 GeV, we counted the number of jets in the remaining events. The results are shown in Fig. 5. Since only a small fraction of the events contain 6 or even 5 standard jets, we conclude that standard jet analysis techniques are insufficient for studying high mass stops at the large $\slashed{E}_{\mathrm{T}}$ cuts required to reduce the SM background to competitive levels.

## B    Simulation Details

We use `MadGraph5_aMC@NLO` [36] package to simulate all parton level events. Showering of the parton level events was done in `Pythia8` [37]. After showering, a $0.1 \times 0.1$ detector granularization was applied for particles in the final state, which were then analyzed and clustered using `FastJet` [38]. The lepton

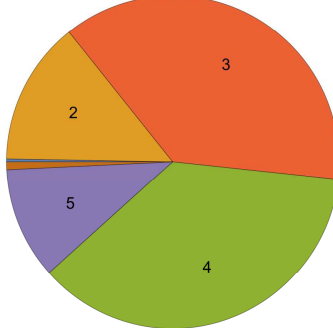Jet count, R = 0.4, $p_T$ > 50 GeV, $\not{E}_T \geq$ 800 GeV

Figure 5: Fraction of events with $n = 1, \ldots, 6$ anti-$k_\mathrm{T}$ jets after imposing standard cuts. The jet counts start with 1 in blue on the left and increase clockwise to 6 in brown.

isolation criteria used for electrons was: $p_T > 15$ GeV, $|\eta| < 2.4$, Sum of hadronic energy within an $R = \min(0.2, 10\,\mathrm{GeV}/E_T)$ cone about the lepton is $< 15\%$ of the lepton transverse energy $E_T$ [25]. Muons are treated similarly, except $R = \min(0.3, 10\,\mathrm{GeV}/E_T)$ is used. Tau tagging is implemented by looking for a jet within $R = 0.5$ of the truth parton level tau. If such a jet is found, we tag it as a tau with 60% probability as a simplistic model of the tau tagging efficiency in [40]. Events with tagged taus are then vetoed if the tagged jet satisfies $p_T > 15$ GeV, $|\eta| < 2.4$.

For the signal, we simulated 50000 events at a center-of-mass energy of 13 TeV. We assume a $K$-factor of 1.5 for the signal production cross section [39]. The desired signal topology where the stops decay with a branching ratio of 100% to tops and neutralinos, with the tops decaying hadronically, was forced in the simulation. All other parameters in the signal simulation were left at their default values.

For all background events, we simulate events at a center-of-mass energy of 13 TeV, using the default SM model file, with default model values. For the $\bar{t}t$ s.l. $+$ jets (e,$\mu$) background, we simulated $4.975 \times 10^7$ events. For the $\bar{t}t$ s.l. $+$ jets ($\tau$) background, we simulated $4.905 \times 10^7$ events. We used an unmatched sample involving 2 jets, which we checked against the variable distributions for a matched sample after the final cuts. We assume a $K$-factor of 1.3 [36]. The desired semi-leptonic background topology where one top decays leptonically and the other hadronically was forced in the simulation. All other parameters in the $\bar{t}t$ s.l. $+$ jets simulation were left at their default values, except for imposing a generator level cut requiring the missing transverse energy (sum of the neutrino momenta) to be above 200 GeV. For the $\bar{t}t + Z$ background, we simulated $7.5 \times 10^5$ events. The desired background topology where both tops decay hadronically and the $Z$ boson decays to neutrinos was forced in the simulation. All other parameters in the $\bar{t}t + Z$ simulation were left at their default values, except

14

requiring a baseline cut of $\not{E}_{\mathrm{T}} > 200$ GeV. For the $Z + $ jets background, we simulated $3.75 \times 10^7$ events. The desired background topology where the $Z$ boson decays to neutrinos was forced in the simulation. All other parameters in the $Z + 3j$ simulation were left at their default values, except requiring $\not{E}_{\mathrm{T}} > 200$ GeV. For the $W + $ jets background, we simulated $4.975 \times 10^7$ events. The desired background topology where the $W$ boson decays to $\ell\bar{\nu}$ or $\bar{\ell}\nu$, where $\ell$ does not include taus, was forced in the simulation. All other parameters in the $W + 3j$ simulation were left at their default values, except a cut of $\not{E}_{\mathrm{T}} > 200$ GeV.

# References

[1] **CMS** Collaboration, S. Chatrchyan et al., *A New Boson with a Mass of 125 GeV Observed with the CMS Experiment at the Large Hadron Collider*, Science **338** (2012) 1569–1575.

[2] **ATLAS** Collaboration, G. Aad et al., *A particle consistent with the Higgs Boson observed with the ATLAS Detector at the Large Hadron Collider*, Science **338** (2012) 1576–1582.

[3] S. P. Martin, *A Supersymmetry primer*, `hep-ph/9709356`. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[4] D. J. H. Chung, L. L. Everett, G. L. Kane, S. F. King, J. D. Lykken, and L.-T. Wang, *The Soft supersymmetry breaking Lagrangian: Theory and applications*, Phys. Rept. **407** (2005) 1–203, [`hep-ph/0312378`].

[5] S. Dimopoulos and G. F. Giudice, *Naturalness constraints in supersymmetric theories with nonuniversal soft terms*, Phys. Lett. **B357** (1995) 573–578, [`hep-ph/9507282`].

[6] A. G. Cohen, D. B. Kaplan, and A. E. Nelson, *The More minimal supersymmetric standard model*, Phys. Lett. **B388** (1996) 588–598, [`hep-ph/9607394`].

[7] M. Cahill-Rowley, J. L. Hewett, A. Ismail, and T. G. Rizzo, *Lessons and prospects from the pMSSM after LHC Run I*, Phys. Rev. **D91** (2015), no. 5 055002, [`arXiv:1407.4130`].

[8] H. Baer, V. Barger, D. Mickelson, and M. Padeffke-Kirkland, *SUSY models under siege: LHC constraints and electroweak fine-tuning*, Phys. Rev. **D89** (2014), no. 11 115019, [`arXiv:1404.2277`].

[9] Z. Han, A. Katz, D. Krohn, and M. Reece, *(Light) Stop Signs*, JHEP **08** (2012) 083, [`arXiv:1205.5808`].

[10] A. Delgado, G. F. Giudice, G. Isidori, M. Pierini, and A. Strumia, *The light stop window*, *Eur. Phys. J.* **C73** (2013), no. 3 2370, [arXiv:1212.6847].

[11] B. Dutta, W. Flanagan, A. Gurrola, W. Johns, T. Kamon, P. Sheldon, K. Sinha, K. Wang, and S. Wu, *Probing compressed top squark scenarios at the LHC at 14 TeV*, *Phys. Rev.* **D90** (2014), no. 9 095022, [arXiv:1312.1348].

[12] W. S. Cho, J. S. Gainer, D. Kim, K. T. Matchev, F. Moortgat, L. Pape, and M. Park, *Improving the sensitivity of stop searches with on-shell constrained invariant mass variables*, *JHEP* **05** (2015) 040, [arXiv:1411.0664].

[13] G. Ferretti, R. Franceschini, C. Petersson, and R. Torre, *Spot the stop with a b-tag*, *Phys. Rev. Lett.* **114** (2015) 201801, [arXiv:1502.01721].

[14] H. An and L.-T. Wang, *Opening up the compressed region of top squark searches at 13 TeV LHC*, *Phys. Rev. Lett.* **115** (2015) 181602, [arXiv:1506.00653].

[15] G. Belanger, D. Ghosh, R. Godbole, and S. Kulkarni, *Light stop in the MSSM after LHC Run 1*, *JHEP* **09** (2015) 214, [arXiv:1506.00665].

[16] A. Kobakhidze, N. Liu, L. Wu, J. M. Yang, and M. Zhang, *Closing up a light stop window in natural SUSY at LHC*, *Phys. Lett.* **B755** (2016) 76–81, [arXiv:1511.02371].

[17] H.-C. Cheng, L. Li, and Q. Qin, *Second Stop and Sbottom Searches with a Stealth Stop*, arXiv:1607.06547.

[18] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, *Stop Reconstruction with Tagged Tops*, *JHEP* **10** (2010) 078, [arXiv:1006.2833].

[19] X.-J. Bi, Q.-S. Yan, and P.-F. Yin, *Probing Light Stop Pairs at the LHC*, *Phys. Rev.* **D85** (2012) 035005, [arXiv:1111.2250].

[20] T. Plehn, M. Spannowsky, and M. Takeuchi, *Stop searches in 2012*, *JHEP* **08** (2012) 091, [arXiv:1205.2696].

[21] D. S. M. Alves, M. R. Buckley, P. J. Fox, J. D. Lykken, and C.-T. Yu, *Stops and $\not{E}_T$: The shape of things to come*, *Phys. Rev.* **D87** (2013), no. 3 035016, [arXiv:1205.5805].

[22] D. E. Kaplan, K. Rehermann, and D. Stolarski, *Searching for Direct Stop Production in Hadronic Top Data at the LHC*, *JHEP* **07** (2012) 119, [arXiv:1205.5816].

[23] M. R. Buckley, T. Plehn, and M. Takeuchi, *Buckets of Tops*, *JHEP* **08** (2013) 086, [arXiv:1302.6238].

[24] **ATLAS** Collaboration, *Search for the Supersymmetric Partner of the Top Quark in the Jets+Emiss Final State at sqrt(s) = 13 TeV*, Tech. Rep. ATLAS-CONF-2016-077, CERN, Geneva, Aug, 2016.

[25] **CMS** Collaboration, *Search for direct top squark pair production in the fully hadronic final state in proton-proton collisions at sqrt(s) = 13 TeV corresponding to an integrated luminosity of 12.9/fb*, Tech. Rep. CMS-PAS-SUS-16-029, CERN, Geneva, 2016.

[26] G. Kasieczka, T. Plehn, T. Schell, T. Strebler, and G. P. Salam, *Resonance Searches with an Updated Top Tagger*, *JHEP* **06** (2015) 203, [arXiv:1503.05921].

[27] C. G. Lester and D. J. Summers, *Measuring masses of semiinvisibly decaying particles pair produced at hadron colliders*, *Phys. Lett.* **B463** (1999) 99–103, [hep-ph/9906349].

[28] I. Low, *Polarized charginos (and top quarks) in scalar top quark decays*, *Phys. Rev.* **D88** (2013), no. 9 095018, [arXiv:1304.0491].

[29] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better jet clustering algorithms*, *JHEP* **08** (1997) 001, [hep-ph/9707323].

[30] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing N-subjettiness*, *JHEP* **02** (2012) 093, [arXiv:1108.2701].

[31] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet substructure as a new Higgs search channel at the LHC*, *Phys. Rev. Lett.* **100** (2008) 242001, [arXiv:0802.2470].

[32] **CMS** Collaboration, C. Collaboration, *Boosted Top Jet Tagging at CMS*, .

[33] **CMS** Collaboration, *Identification of b quark jets at the CMS Experiment in the LHC Run 2*, Tech. Rep. CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[34] A. Barr, C. Lester, and P. Stephens, *m(T2): The Truth behind the glamour*, *J. Phys.* **G29** (2003) 2343–2363, [hep-ph/0304226].

[35] H.-C. Cheng and Z. Han, *Minimal Kinematic Constraints and m(T2)*, *JHEP* **12** (2008) 063, [arXiv:0810.5178].

[36] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [arXiv:1405.0301].

[37] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1, JHEP05 (2006) 026, Comput. Phys. Commun.* **178** (2008) 852–867, [arXiv:0710.3820].

[38] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual, Eur. Phys. J.* **C72** (2012) 1896, [arXiv:1111.6097].

[39] C. Borschensky, M. Krmer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell, *Squark and gluino production cross sections in pp collisions at $\sqrt{s}$ = 13, 14, 33 and 100 TeV, Eur. Phys. J.* **C74** (2014), no. 12 3174, [arXiv:1407.5066].

[40] **CMS** Collaboration, V. Khachatryan et al., *Reconstruction and identification of tau lepton decays to hadrons and tau neutrino at CMS, JINST* **11** (2016), no. 01 P01019, [arXiv:1510.07488].