Search for bottom squarks in the baryon-number violating MSSM

Debjyoti Bardhan,[1,*] Amit Chakraborty,[1,†] Debajyoti Choudhury,[2,‡] Dilip Kumar Ghosh,[3,§] and Manas Maity[4,‖]

[1]*Department of Theoretical Physics, Tata Institute of Fundamental Research, 1, Homi Bhabha Road, Mumbai 400005, India*

[2]*Department of Physics and Astrophysics, University of Delhi, Delhi 110007, India*

[3]*Department of Theoretical Physics, Indian Association for the Cultivation of Science, 2A & 2B, Raja S.C. Mullick Road, Kolkata 700032, India*

[4]*Department of Physics, Visva-Bharati, Santiniketan 731235, India*

(Received 6 March 2017; published 25 August 2017)

We consider a scenario of a minimal supersymmetric standard model with R-parity violation, where the lightest supersymmetric particle is the lighter bottom squark ($\tilde{b}_1$). We study the production of a bottom squark pair at the LHC and their subsequent decays through the baryon number violating operators leading to a top pair with two light quarks. Looking for both semileptonic and fully hadronic (no leptons) final states, we perform cut-based as well as multivariate analyses (MVA) to estimate the signal significance at the 13 TeV run of the LHC. We find that a cut-based analysis can probe bottom squark mass up to ∼750 GeV, which may be extended up to ∼850 GeV using MVA with 300 fb^{-1} integrated luminosity. The fully hadronic final state, however, is not as promising.

DOI: 10.1103/PhysRevD.96.035024

I. INTRODUCTION

The discovery of the most elusive particle of the Standard Model (SM), the Higgs boson, by the LHC during its first run completed the hunt for SM particles [1,2]. The SM has been extremely successful in explaining diverse phenomena involving elementary particles. The SM, however, fails to address many issues, two of the most prominent being the hierarchy problem and the baryon asymmetry in the Universe. Supersymmetry (SUSY), one of the most popular extensions of the SM, can not only provide plausible resolutions for these problems, but can also facilitate answers to other questions such as the unification of forces or the presence of dark matter (DM). Consequently, the ATLAS and CMS collaborations at the LHC have searched for signatures of the supersymmetric particles. So far, no signal of any SUSY particles has been seen, and this has been construed to impose severe constraints on the parameter space of the theory [3].

One should, however, note that the majority of the SUSY search strategies at the LHC assume that "R-parity," a multiplicative quantum number defined as $R = (-1)^{3B+L+2s}$ with B, L, and s in terms of the baryon number (B), the lepton number (L), and the spin (s) of the particle, is conserved. Conservation of R-parity implies that SUSY particles will always be pair produced and that a heavy SUSY particle will decay into an odd number of lighter SUSY particles, with or without other SM particles.[1] This ensures that the lightest SUSY particle (LSP) is stable. A characteristic signature of an R-parity–conserving SUSY scenario is a final state with large missing transverse energy ($\not{E}_T$) due to the presence of the LSP. Since, in the SM, neutrinos are the only real sources of missing energy apart from detector acceptance and resolution effects, $\not{E}_T$ can be used as a standard candle to search for these SUSY particles. Besides, in supersymmetric theories with conserved R-parity, the lightest SUSY particle, if colourless and electrically neutral, can always act as a good dark matter candidate.

However, the conservation of R-parity is not guaranteed, and, if one allows for its violation, an sfermion can decay to a pair of SM fermions,[2] giving rise to signatures with, at best, only a small missing transverse energy [4]. This negation of one of the standard features of SUSY searches would immediately negate much of the collider constraints on the SUSY spectrum. Welcome consequences of this are the easing of fine-tuning on the one hand [5], and, on the other, the possibility of facilitating electroweak baryogenesis [6–8] through the accommodation of a top squark lighter than what the R-conserving scenarios can allow. From its very definition, one can see that the violation of R-parity can be achieved in three ways: violation of either L or B, or both. However, if we allow both L and B to be violated, nothing prevents the proton from decaying into a

*debjyoti@theory.tifr.res.in
†amit@theory.tifr.res.in
‡debajyoti.choudhury@gmail.com
§tpdkg@iacs.res.in
‖manas.maity@cern.ch

[1]In most popular models, the decays are into a single lighter SUSY particle and one or two (and, only rarely three) SM particles.

[2]Similarly, the gauginos and higgsinos would decay into three SM fermions.

meson and a lepton, and thus the lower limit on the proton decay lifetime [9] places severe constraints on their products. It is, of course, more natural to ensure proton stability by insisting on one of the symmetries (B or L) being unbroken, and this is the route that we take. Interestingly, such R-violating scenarios can be easily motivated from supergravity models [10]. And while violating R-parity implies that we lose the DM candidate, the dark matter content of the universe can appear from other sources [11].

In this paper, we study a R-violating (RPV) SUSY scenario in the presence of baryon number (UDD-type) violating operators alone. Contrary to naive expectations, such a scenario can be well-accommodated within a grand unified theory framework [12], thereby preserving one of the successes of SUSY. A further ramification is that, unlike in the case of the L-violating couplings, the lack of any excess in the multilepton channel at the LHC does not impose any worthwhile constraint on the squark/gluino masses [13–18]. We are faced, instead, with a multijet signal [19–22], and it has been argued that the large irreducible QCD background would result in much weaker sensitivity. Performing a collider analysis of the lightest scalar superpartner of the bottom quark, namely the bottom squark ($\tilde{b}$), subsequently decaying to a top quark and a light down-type quark through nonzero λ'' couplings, we show that it is not necessarily so. Depending on the decay of the top quark, the final state can consist of only hadronic elements (jets) or may contain at least one lepton. The latter semileptonic case is easier to study at a hadronic collider environment like that of the LHC, since we can tag on the lepton. Our analysis will take into account the very different nature of these two possible final states and is thus done in two parts: first for a final state with at least one lepton, and second for a fully hadronic final state. To study the semileptonic final state, we shall use both the traditional cut-based analysis and multivariate analysis, while in the hadronic final state, we shall rely solely on the multivariate analysis.

The rest of the paper is arranged as follows: In Sec. II, we briefly introduce R-parity violating SUSY, noting down the couplings relevant for our analysis. In Sec. III, we introduce our simplified model detailing all the parameters used. The analysis of a final state with a lepton is presented in Sec. IV, and in Sec. V we perform the analysis for a completely hadronic final state. Finally, we conclude in Sec. VI.

II. THE R-PARITY VIOLATING MSSM

In terms of lepton, quark, and Higgs superfields one can write down the R-parity violating superpotential in the following form [4]: either bilinear terms or by Yukawa-like trilinear terms. The most generic RPV superpotential is given by

$$W_{\not R_p} = \mu_i \hat{H}_u \hat{L}_i + \frac{1}{2}\lambda_{kij}\hat{L}_i\hat{L}_j\hat{E}^c_k + \lambda'_{ijk}\hat{L}_i\hat{Q}_j\hat{D}^c_k + \frac{1}{2}\lambda''_{ijk}\hat{U}^c_i\hat{D}^c_j\hat{D}^c_k, \quad (1)$$

where the first three sets of operators violate L while the last set violates B. Here, i, j, and k are generation indices, whereas both $SU(2)$ and $SU(3)$ indices have been suppressed. Clearly, the couplings λ_{kij} and λ''_{ijk} are antisymmetric in the last two indices, and thus there are a total of $(3+9+27=)39$ L- and 9 B-violating interactions. Switching off the first three sets and concentrating only on the last, we have, in terms of the quark and squark fields,

$$\mathcal{L}_{UDD} = -\frac{1}{2}\lambda''_{ijk}(\tilde{u}^\star_{iR}\bar{d}_{jR}d^c_{kL} + \tilde{d}^\star_{kR}\bar{u}_{iR}d^c_{jL} + \tilde{d}^\star_{jR}\bar{u}_{iR}d^c_{kL}). \quad (2)$$

The bounds on the couplings λ''_{ijk} are varied. Some of them are strongly constrained from $n-\bar{n}$ oscillations [23] or the LEP data on Z-decays [24]. The others are only weakly restricted, for example, through the requirement of their perturbative under renormalization group flows [25]. Compendia of such constraints can be found in Refs. [4,10,26–32]. It should be noted that many of the low-energy constraints emanate from effective four-fermi interactions, and in quoting them a reference squark mass is used; these bounds need to be scaled appropriately when the squark mass differs.

As we are interested in the $\tilde{b}$, one of j, k in λ''_{ijk} must be 3. Similarly, if we demand that the bottom squark should decay into a top, we must have $i=3$. In other words, we are left with just two choices, namely λ''_{313} and λ''_{323}, leading to $\tilde{b}^* \to t+d$ and $\tilde{b}^* \to t+s$ respectively. Since the simultaneous presence of two such couplings lead to too large a size for flavor changing neutral currents (FCNC) [33,34], we assume that only one of the two is nonzero and real. For the mass range (of the squarks) that we are interested in, the strongest constraints are $\lambda''_{313} < 0.1$ [32] and $\lambda''_{323} < 1.89$ [24] respectively. Even without saturating these bounds, it is obvious that, once produced, the bottom squark may decay promptly, thereby eliminating the possibility of recognizably displaced vertices [35,36]. We shall assume that while the R-violating couplings are small enough to be both consistent with low energy phenomenology as well as having at best marginal effect on squark-production, they are large enough to prevent displaced vertices, thereby removing tell-tale signatures.

In the presence of "UDD"-type couplings, the decays (direct or cascades) of squarks and gluinos (the dominantly produced SUSY particles at the LHC) would, typically, result in multijet configurations with very little missing momentum. As these are very difficult to detect (especially in the absence of hard leptons) in the messy hadronic environment of the LHC, the strong limits on squark/gluino masses, derived in the context of R-conserving models

(or even for R-violating but B-conserving ones), do not hold. In particular, if a pair-produced squark decays directly into a pair of quarks, the resultant four-jet sample is likely to be overwhelmed by the QCD background. The situation is ameliorated somewhat if some of the quarks (rather, the corresponding jets) can be tagged, as this would allow us not only to eliminate much of the background, but also to use invariant mass combinations to increase the signal-to-noise ratio. This was used in Refs. [37,38] and, subsequently, by the ATLAS collaboration [20] to investigate the pair production of top squarks and their decays through the very same couplings that we are considering here, to a b quark and a light quark each. Here, we investigate the complementary scenario, namely where the bottom squark (rather than the top squark) is the LSP.

III. THE SIMPLIFIED MODEL SPECTRUM AND SIMULATION

As we are primarily interested in the lighter bottom squark, we simplify the spectrum by considering it to be the LSP with the other SUSY partners being much heavier. In particular, we do not include gluino production despite the fact that, for similar masses, $\sigma(\tilde{g}\,\tilde{g}) \gg \sigma(\tilde{b}\tilde{b}^*)$ and that the gluino could easily decay into a $\bar{b}\,\tilde{b}$ pair, thereby adding to the signal strength. Indeed, gluino pair production with each decaying into three quarks has been used [22] to set a limit of $m_{\tilde{g}} > 1.08$ TeV, and hence by making the $\tilde{g}$ heavy, we deliberately preclude this contribution altogether. Our assumption about the spectrum obviously means that decays through R-conserving channels are no longer possible and that the bottom squark is forced to decay to two SM quarks with 100% branching ratio. For the choice of the RPV coupling λ''_{313} (λ''_{323}), the daughters are the top and a light quark (d or s, as the case may be). The top quark can decay either leptonically or hadronically; thus, giving rise to the following final states:

(a) $2\ell\ell' + b\bar{b} + \text{jets} + \not{E}_{\rm T}$; $\ell, \ell' = e, \mu$,
(b) $1\ell + b\bar{b} + \text{jets} + \not{E}_{\rm T}$,
(c) $0\ell + b\bar{b} + \text{jets} + \not{E}_{\rm T}$.

It should be noted that all these channels will be associated with only a small missing transverse energy, if any. Final states with multiple jets are very challenging in the LHC environment and thus require dedicated studies. Several SM processes which provide similar final state signatures have been treated as background, particularly, $t\bar{t}$ + jets (up to 2), $t\bar{t}b\bar{b}$, $t\bar{t}Z$, $t\bar{t}W$, and $t\bar{t}H$, constitute the dominant SM background; QCD multijet events constitute huge background for the purely hadronic case.

Before we delve into the discussion of signals and backgrounds, let us examine the parameter space that leads to the spectrum that we czonsider. The gaugino mass parameters M_1 and M_2, as well as the higgsino mass parameter μ, are set to 1 TeV, while the value of $\tan\beta$, the ratio of the vacuum expectation values of the two Higgs doublets H_u^0 and H_d^0, is fixed at 10. The masses of the first two generations of squarks and all the three generations of sleptons lie around 3 TeV, and the mass of the right-handed top squark is set to ∼1 TeV. The left-handed third generation squark mass is set to about 1.5 TeV. While the trilinear coupling A_t is set to −2 TeV, the other trilinear couplings A_b and A_τ are set to zero. We also fix the gluino mass parameter (M_3) at 2 TeV, while varying only the right-handed bottom squark mass parameter (m_{b_R}) to obtain different bottom squark masses. In our analysis we consider six representative benchmark points with bottom squark masses 500 GeV (BP-1), 600 GeV (BP-2), 700 GeV (BP-3), 800 GeV (BP-4), 900 GeV (BP-5), and 1000 GeV (BP-6).

TABLE I. The first row presents the masses of the bottom squark for the different benchmark points. In the bottom part of the table, the values of the low-energy flavor observables are presented. These remain identical for the different benchmark points.

	BP-1	BP-2	BP-3	BP-4	BP-5	BP-6
$m_{\tilde{b}_1}$ (GeV)	500	600	700	800	900	1000
$\mathcal{B}(B_s \to \mu^+\mu^-)$			4.27×10^{-9}			
$\mathcal{B}(b \to s\gamma)$			3.19×10^{-4}			
$\Delta m_{B_s}(\text{ps}^{-1})$			18.01			
$\Delta m_{B_d}(\text{ps}^{-1})$			0.403			

The particle spectrum has been generated using SPheno (version 3.3.8) [39,40] with the trilinear R-parity violating model as implemented in SARAH (version 4.4.6) [41,42]. FlavorKit [43] is used to calculate the low-energy flavor observables $b \to s\gamma$ and $B_s \to \mu^+\mu^-$, and care has been taken to ensure that the benchmark points are consistent with the flavor physics data [44] at better than 95% C.L. In particular, the mass differences Δm_{B_d} (Δm_{B_s}) associated with B_0- and B_s-mixing (see Table I) are very close to the experimental measurements [32]. We further ensure that the spectrum we use at each benchmark point is consistent with the latest measurements of Higgs mass, Higgs couplings, and Higgs signal strength at the LHC.

It is worth noting that for our analysis, lighter squarks, consistent with the present bounds, would not be a problem. Since the contribution to various flavor processes from RPV, typically, are proportional to $(\lambda''^2/m_{\tilde{q}}^2)$, it is possible to accommodate a smaller squark mass, provided the couplings are reduced accordingly. In this scenario, we would, for example, receive additional contribution to our signal events from, say the strange squark. If the strange squark were only slightly heavier than the bottom squark, it would decay to the bottom squark along with a bottom and a strange quark via an off-shell gluino.[3] Owing to only a

[3]The only other channel available to it would be the RPV channel to the top and the bottom, which would, again, be largely indistinguishable from that we consider here.

small difference in the masses, the bottom squark would be produced almost at rest with the two other jets being very soft; this would be indistinguishable from the bottom squark pair production scenario and would thus add to our signal events. We do not consider this, and thus the analysis in this paper is quite conservative.

The signal and background events are generated using MADGRAPH (version 2.2.2) [45], properly interfaced with PYTHIA8 (version 8.210) [46,47] for parton showering and hadronization. Event sets are then passed through DELPHES (version 3.2.0) [48] in order to simulate the detector response. Jets are reconstructed using FASTJET (version 3.1.3) [49], with $R = 0.4$ using the anti-k_t algorithm [50] in the leptonic case. For the hadronic case, we intend to tag the boosted top quarks in the final state, which will necessarily be a fat jet; thus, we use $R = 1.8$ using the C/A algorithm [51], which is optimized for tagging moderately boosted tops [52].

Jets are selected with $p_T > 30$ GeV and $|\eta| < 2.5$. Leptons (electron and muon) are selected with $p_T > 20$ GeV and $|\eta| < 2.4$. To reduce the background contribution of electrons or muons from semileptonic decays of heavy flavors, a relative isolation criteria is imposed. The relative isolation parameter, $I_{\rm rel}$, defined as

$$I_{\rm rel} = \frac{\sum_{i\neq P} p_T(i)}{p_T(P)}, \tag{3}$$

with P being the particle of interest (here electron or muon), is calculated as the sum of transverse energy of all the charged and neutral particles measured in the tracker and calorimeters in an isolation cone[4] $\Delta R < 0.3$ around the lepton direction divided by the lepton transverse momentum. In our analysis, we demand $I_{\rm rel} < 0.15$.

In the semileptonic decays of the top, the final state contains multiple leptons and a significant amount of missing transverse energy, calculated using the p_T of all the visible particles. Our signal topology also includes multiple b-jet candidates, and in order to tag them as "b-jets," we require the angular distance ΔR between the parton level b-quark and the jet to be less than 0.4, as implemented in DELPHES. A p_T-dependent b-tagging efficiency (ϵ_b) for $|\eta| < 2.5$, following the CMS collaboration [53], is used to make our analysis more robust:

$$\epsilon_b = \begin{cases} 0.75 & \text{for } p_T^b \leq 30\,\text{GeV} \\ 0.85 & \text{for } 30\,\text{GeV} < p_T^b \leq 400\,\text{GeV} \\ 0.95 - 0.00025\,p_T & \text{for } 400\,\text{GeV} < p_T^b \leq 800\,\text{GeV} \\ 0.65 & \text{for } p_T > 800\,\text{GeV}. \end{cases} \tag{4}$$

[4]Here, and henceforth, $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is the usual distance measure in the rapidity(η)–azimuthal angle (ϕ) plane.

Throughout the entire p_T range, following the CMS card, a mistagging rate of 1% is assumed for the non-b-jets. Note that the b-tagging efficiency obtained by the ATLAS collaboration [54] is comparable with that of the CMS collaboration.

The cross section of the $t\bar{t}$ + jets (up to 2) process is taken from the LHC Top Quark Working Group [55], while that of the $t\bar{t}H$ is taken from the LHC Higgs Cross Section Working Group report [56]. The Next-to leading order (NLO) cross section for $t\bar{t}W$ and $t\bar{t}Z$ are taken from [57], where the results have been computed using MSTW2008 parton distribution functions (PDFs). We use PYTHIA to calculate the cross section for the $t\bar{t}b\bar{b}$ process, where the PDF used in the calculation is CTEQ6L [58] and the factorization scale has been chosen to be M_Z, the mass of the Z boson. For the signal processes, we use the bottom squark pair production cross sections at the 13 TeV LHC calculated including the resummation of soft-gluon emission at next-to-leading logarithmic accuracy matched to next-to-leading-order supersymmetric QCD corrections [59].

IV. LEPTONIC FINAL STATE

In this section we consider the final state in which there is at least one lepton; thus, this analysis includes both leptonic and semileptonic decays of the top quarks. We first perform a cut-based analysis on the data sets and then supplement it with a multivariate analysis.

In Fig. 1, we present the jet multiplicity and the p_T distribution of the two leading non-b-tagged jets. Additionally, in Fig. 2, the distributions for H_T and M_{T2}, both defined shortly, are also shown. All the distributions include three representative benchmark points: BP-1, corresponding to $m_{\tilde{b}_1} = 500$ GeV; BP-4, corresponding to $m_{\tilde{b}_1} = 800$ GeV; and BP-6, corresponding to $m_{\tilde{b}_1} = 1000$ GeV, along with the dominant SM backgrounds. Following these distributions, we can discuss the optimization of our selection cuts in order to improve the signal to background ratio.

A. Cut-based analysis

Our cut-optimization prescription resembles the one adopted by the CMS collaboration [60] in order to distinguish a $t\bar{t}b\bar{b}$ sample from a background sample of $t\bar{t}$ + jets (up to 2). In this section, we consider only leptonic events, i.e., events which have at least one lepton. In order to distinguish the signal from background, we make use of five discriminating variables: the number of jets, the p_T of the hardest and the second hardest jets in each event, the scalar H_T, and the transverse mass variable, M_{T2}. The individual variables and the cuts imposed on them are discussed below:

C1: We demand that each signal event contain at least one lepton. This particular choice of signal topology will

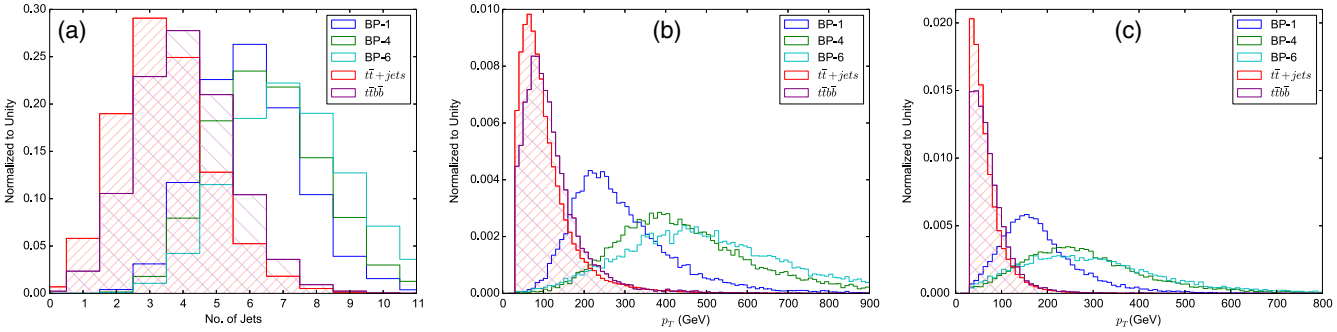

FIG. 1. In (a) we show the distribution of the number of jets, (b) the P_T distribution of the hardest non-b-tagged jet, while in (c) the same for the second hardest non-b-tagged jet. For the sake of clarity, just the two dominant SM background processes are shown, viz. $t\bar{t}+$ jets and $t\bar{t}b\bar{b}$.

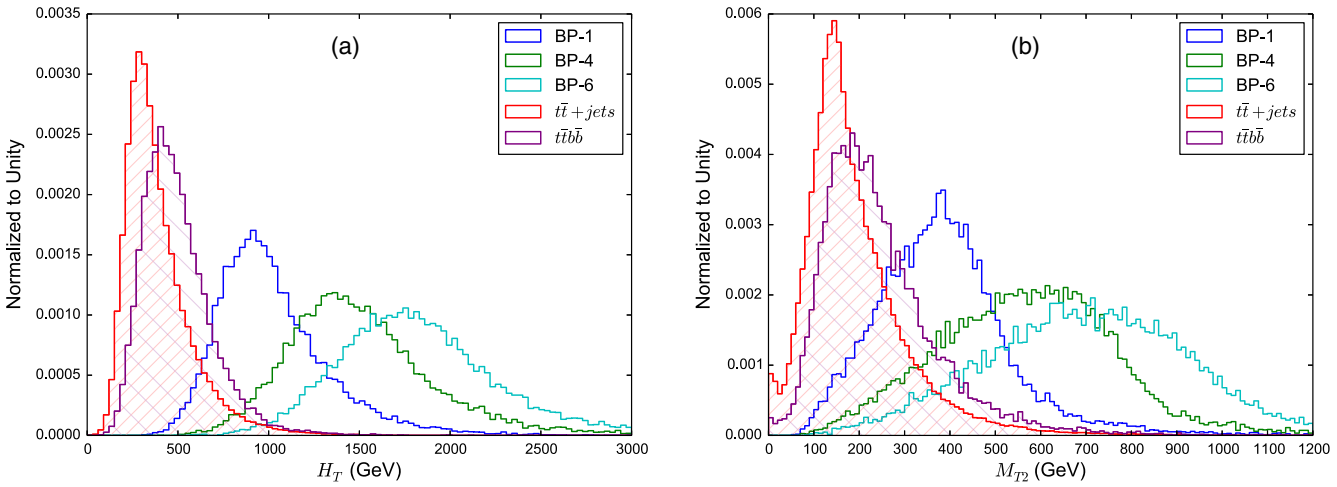

FIG. 2. Distribution of the (a) scalar H_T and (b) Stransverse mass variable, M_{T2}, are displayed.

substantially remove the most severe SM background coming from the pure QCD multijet processes.

C2: In any process with multiple jets in the final state, the number of jets (including both the b-jets and the light jets) plays a very crucial role as a discriminatory variable. Due to the large mass splitting between the bottom squark and the top, both the latter and the other daughter would, often, carry a large p_T, and hence we expect a higher multiplicity of large-p_T jets as compared to the SM backgrounds which, typically, have a significant number of softer jet progenitors. Consequently, we demand that the number of jets be greater than four.

C3: Since we expect the non-b-tagged jets coming from the bottom squark decay to have a high p_T, we can place a p_T cut on such a jet. We demand that the leading non-b-tagged jet have a $p_T > 250$ GeV.

C4: Since the bottom squark pair decay produces two light jets, we expect that the second hardest non-b-tagged jet will also be very energetic. We put a cut of $p_T > 150$ GeV on the sub-leading non-b-tagged jet. The light jets, if any, from the background processes are not expected to have such a high p_T.

C5: We calculate H_T, the scalar sum of p_T of all the visible particles, namely jets, leptons, and photons. It is defined as follows:

$$H_T = \sum_{i=e,\mu,j,\gamma} |\vec{p}_T(i)|. \tag{5}$$

The importance of this variable as a signal discriminator is very well reflected in Fig. 2. If we demand that our signal events should have substantially large value of $H_T \sim 1000$ GeV, then most of the $t\bar{t}jj$ and $t\bar{t}b\bar{b}$ events are removed. This is again taking advantage of the fact that the large mass of the bottom squark results in jets and leptons with a p_T typically much higher than those emerging from SM processes.

C6: Finally, we put a cut on the transverse mass variable $M_{T2} > 360$ GeV. The variable is defined as [61]

$$M_{T2}(\vec{p}_T^{V1}, \vec{p}_T^{V2}, \not{p}_T) = \min_{\vec{\not{p}}_T^1 + \vec{\not{p}}_T^2 = \vec{\not{p}}_T} [\max\{M_T(\vec{p}_T^{V1}, \vec{\not{p}}_T^1), M_T(\vec{p}_T^{V2}, \vec{\not{p}}_T^2)\}], \tag{6}$$

TABLE II. The surviving cross section (in fb) for the different processes after each of the cuts. For $t\bar{t}$ + jets, we consider up to 2 jets.

Cuts	$t\bar{t}$ + jets	$t\bar{t}b\bar{b}$	$t\bar{t}Z$	$t\bar{t}H$	$t\bar{t}W$	BP-1	BP-2	BP-3	BP-4	BP-5	BP-6
C0:	8.3×10^5	1.7×10^4	8.7×10^2	5.1×10^2	6.5×10^2	5.2×10^2	1.8×10^2	67.0	28.3	12.9	6.2
C1:	1.8×10^5	3.3×10^3	2.7×10^2	1.0×10^2	2.3×10^2	90.6	29.6	10.9	4.4	1.9	0.8
C2:	3.8×10^4	1.2×10^3	1.4×10^2	63.4	89.4	76.8	25.8	9.7	3.9	1.7	0.8
C3:	3.9×10^3	65.2	20.3	6.8	10.5	43.3	19.3	8.2	3.6	1.6	0.7
C4:	1.6×10^3	27.2	11.0	3.2	2.1	33.4	16.1	7.2	3.2	1.5	0.6
C5:	9.6×10^2	16.3	7.7	2.1	3.1	26.1	14.3	6.9	3.2	1.4	0.6
C6:	7.6×10^2	13.9	5.5	1.6	1.9	17.4	10.9	5.6	2.7	1.3	0.6

where $\vec{\not{p}}_T^1$ and $\vec{\not{p}}_T^2$ are two hypothetical subdivisions of the total missing transverse momentum $\not{p}_{\mathrm{T}}$. The separation of the visible particles into two sets with associated transverse momenta $\vec{p}_T^{V1}$ and $\vec{p}_T^{V2}$ is done so that the invariant masses of the two parts are as close to each other as possible.

In general, the transverse mass $M_T(\vec{p}_1, \vec{\not{p}}_2)$ of the $(\vec{p}_1, \vec{\not{p}}_2)$ system is defined as

$$M_T(\vec{p}_1, \vec{\not{p}}_2) = \sqrt{m_1^2 + 2|\vec{p}_1||\vec{\not{p}}_2|(1 - \cos\phi)}. \quad (7)$$

Here, ϕ is the azimuthal angle between the $\vec{p}_1$ and $\vec{\not{p}}_2$ vectors with $\vec{\not{p}}_2$ corresponding to a massless particle (neutrino) and $m_1^2 \equiv p_1^2$. For the process under consideration, the visible part is comprised of a b-quark, a light quark, and a lepton coming from each of the bottom squarks. Given the symmetry of the system, we group the visible entities such that the two visible parts are nearly identical in invariant mass. For calculating M_{T2}, we use the Cheng and Han bisection algorithm [62]. From the distribution shown in Fig. 2, we can easily see that this variable too has a good discriminatory power.

The event summary for the signals and backgrounds after individual selection cuts is presented in Table II. The numbers in the table denote the resulting cross sections after each selection cut is applied to both signal and background events. The first row in the table, denoted by "C0," refers to the NLO production cross section for each process.

The numbers in the subsequent rows relate to the surviving cross section for each of the cases after the relevant cut (indicated as bullet points earlier) has been imposed.

We can now estimate the signal significance corresponding to each benchmark point at the 13 TeV LHC assuming 300 fb^{-1} of integrated luminosity. We are interested in the cross section after the cut "C6" is imposed (last row of Table II). The number of signal (background) events, denoted by S (B), is given by the product of this cross section and the integrated luminosity. In Table III, we tabulate the signal significance $\mathcal{S}$ given by

$$\mathcal{S} = \frac{S}{\sqrt{S+B}}.$$

It is evident from the table that, for an integrated luminosity of 300 fb^{-1}, the LHC stands in extremely good stead to detect the bottom squark should its mass be 600 GeV or below. The LHC will graze past the exclusion limit of 95% C.L. for masses around ∼750 GeV. Given 3000 fb^{-1} integrated luminosity, we find that the discovery reach (i.e., 5σ significance) will exceed 800 GeV and the exclusion bounds might be extended to beyond the 900 GeV mass point.

B. Multivariate analysis

To achieve a better discrimination between the signal and the SM background, we perform a multivariate analysis (MVA) using the boosted decision tree (BDT) algorithm as implemented in the Toolkit for Multivariate Data Analysis (TMVA) [63] with ROOT [64]. We briefly describe the procedure, the details of which may be found in Ref. [63], along with the parameters for our analysis below.

Decision trees are used to classify events as either signal-like or backgroundlike. Each node in a decision tree uses a

TABLE III. Number of background and signal events for a integrated luminosity of 300 fb^{-1}, along with the significance for the different benchmark points. See text for details.

	Background	BP-1	BP-2	BP-3	BP-4	BP-5	BP-6
$m_{\tilde{b}_1}$ (GeV)		500	600	700	800	900	1000
$\mathcal{N}$	2.3×10^5	5.2×10^3	3.3×10^3	1.7×10^3	8.2×10^2	3.9×10^2	1.7×10^2
$\mathcal{S} = \frac{S}{\sqrt{S+B}}$		10.7	6.7	3.5	1.7	0.8	0.4

TABLE IV. Different sets of variables that can be considered for the multivariate analysis. We choose Set-2 for our analysis.

Set-1	$(p_T)_{j1}$, $(p_T)_{j2}$, $(p_T)_{j3}$, $(p_T)_{j4}$, $(p_T)_{bj1}$, $(p_T)_{bj2}$, $(p_T)_{bj3}$, $(p_T)_{bj4}$, H_T, $\not{E}_{\rm T}$, nJets, nbJets, M_{T2}, $m^h_{\tilde{b}_1}$, $m^{\ell}_{\tilde{b}_1}$
Set-2	$(p_T)_{j1}$, $(p_T)_{j2}$, $(p_T)_{j3}$, $(p_T)_{j4}$, $(p_T)_{bj1}$, $(p_T)_{bj2}$, H_T, $\not{E}_{\rm T}$, nJets, M_{T2}, $m^h_{\tilde{b}_1}$, $m^{\ell}_{\tilde{b}_1}$
Set-3	$(p_T)_{j1}$, $(p_T)_{j2}$, $(p_T)_{bj1}$, $(p_T)_{bj2}$, H_T, $\not{E}_{\rm T}$, nJets, M_{T2}
Set-4	$(p_T)_{j1}$, $(p_T)_{j2}$, H_T, $\not{E}_{\rm T}$, nJets, M_{T2}

single discriminating variable, along with a certain cut value imposed on it, to provisionally classify events as either signal-like or backgroundlike depending on the purity of the sample. The decision tree needs to be "trained" and that starts with the root node. We can think of the process as two bins originating from the root node (i.e., the zeroth node), one having events classified as signal-like and the other as backgroundlike. At the next level, each of these bins can be treated in exactly the same way as the root node, using a variable of choice and a particular value of cut on it, giving us two bins—one signal-like and the other backgroundlike—for each node. A tree is built up to a depth either determined by the remaining number of background events, or by the depth specified by the user. The final leaf nodes contain backgroundlike and signal-like events from the training sample. Generally, half of the provided sample is used for training and the other half is then used for testing.

Decision trees, however, are unstable under statistical fluctuations and cannot be used as good classifiers. Instead, the technique of boosting can be used to combine several classifiers into a single one, such that the latter is more is stable under such fluctuations and, hence, has a smaller error than the individual ones. Boosting modifies the weights of individual events and creates a new decision tree. Higher weights are preferentially assigned to the incorrectly classified events. Previously assigned weights are modified by α, given by

$$\alpha = \frac{1-\epsilon}{\epsilon}, \quad \text{where } \epsilon = \sqrt{\frac{p(1-p)}{N}}, \tag{8}$$

where N is the total number of training events in the node and $p = S/(S+B)$, called the purity of the sample. The number of decision trees in the forest we use is given by $\text{NTREES} = 400$, the maximum depth of the decision tree allowed is $\text{MAXDEPTH} = 5$ and the minimum percentage of training events in each leaf node is given by $\text{MINNODESIZE} = 2.5\%$. We choose Adaptive Boost, proven to be effective with weak classifiers and implemented as AdaBoost in TMVA, as the method for boosting the decision trees in the forest with the boost parameter $\beta \equiv \text{ADABOOSTBETA} = 0.5$. This parameter adjusts the learning rate of the algorithm simply by changing the weights $\alpha \to \alpha^{\beta}$. We have used the default values of the BDT parameters, viz. NTrees, MaxDepth, and MinNodeSize.

A challenge endemic to TMVA is finding an optimal set of observables that would lead to the best possible discrimination between signal and background events. It is important to note that a larger set of variables need not always provide better discrimination, especially if it is mostly filled with irrelevant observables. We tried four sets comprised of 15 (Set 1), 12 (Set 2), 8 (Set 3), and 6 (Set 4) variables respectively as detailed in Table IV. We then plot the receiver's operative characteristic (ROC) curve for these sets. The ROC curves signify the efficiency of the signal (ϵ_S) with respect to the efficiency of rejecting the background ($1-\epsilon_B$), with ϵ_B being the efficiency of the background. This is exemplified by the left panel of Fig. 3, wherein we plot these ROCs for the benchmark point BP-4. Whereas the use of Set 1 and Set 2 offers some improvement over Sets 3 and 4, the former are virtually indistinguishable in their efficacy. In other words, the extra variables in Set 1 are of very little relevance. Given this, we choose the largest set of variables without keeping any irrelevant variables, namely Set 2, for the rest of the analysis in this section.

The variables chosen as BDT inputs have already been introduced in the previous section (see the cut-based analysis) except for the two new variables, namely $m^h_{\tilde{b}_1}$ and $m^{\ell}_{\tilde{b}_1}$, which represent the reconstructed bottom squark mass using the hadronically (h) and leptonically (ℓ) decaying top quarks, respectively. We select events with exactly one isolated lepton (electron or muon) with two or more b-tagged jets, utilizing only the two hardest b-tagged jets in our reconstruction. Additionally, we work with the four hardest light (i.e., non-b-tagged) jets in the event.

We could have also attempted to reconstruct the bottom squark for events with two isolated leptons originating from the leptonic decay of the two top quarks. However, the presence of two neutrinos, the only source of missing energy here, renders the reconstruction nontrivial and makes it a highly involved task. With the dileptonic branching fraction being only 5% (compared to 30% for the semileptonic one), and with the pair production cross section falling rapidly with the bottom squark mass, this channel is likely to be important only in the very high luminosity run of the LHC. In this work, we thus focus on

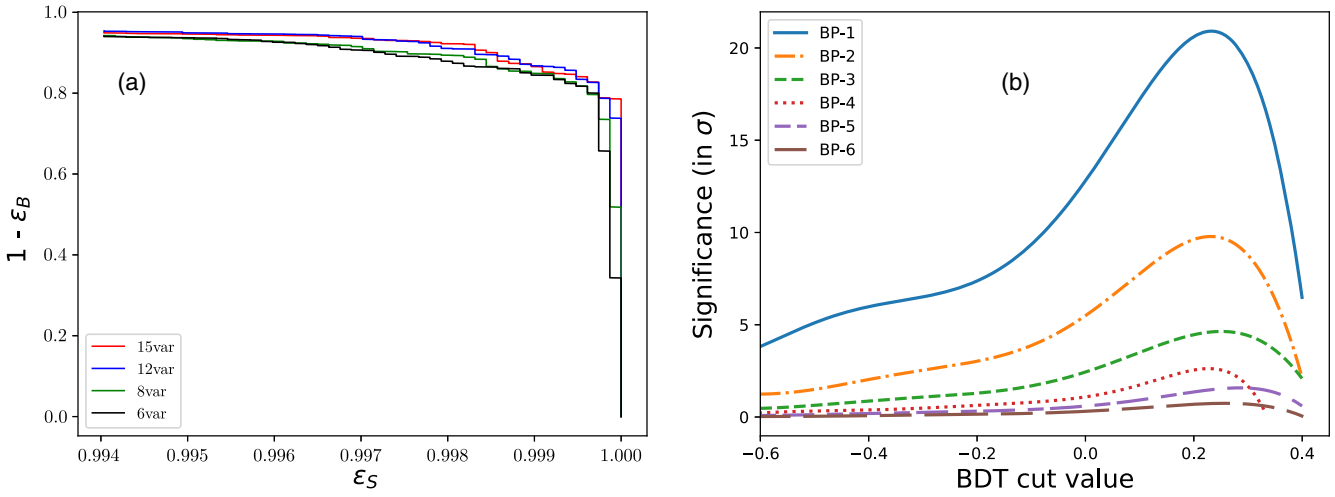

FIG. 3. On the left (a), the ROC plot is shown for BP-4 with the four sets of variables and on the right (b), the plot for the signal significance against different BDT cut values is shown.

only the semileptonic case when events contain exactly one isolated lepton with b-tagged and light jets and missing transverse energy. The interested reader can refer to [65,66] for the detailed implementation of the reconstruction of $t\bar{t}$ and heavy resonances using dileptonic modes.

Before we proceed to reconstruct the top quark, we must reconstruct the W bosons. The hadronically decaying W boson is reconstructed by choosing the pair of light non-b-tagged jets which give an invariant mass closest to the actual W boson mass with a further demand that the thus reconstructed mass lies within $M_W \pm 30$ GeV. The leptonically decaying W boson in the decay of the top quark is reconstructed, within a quadratic ambiguity, from the four momentum of the lepton p_ℓ and the missing transverse momentum $\vec{\not{p}}_T \equiv (\not{p}_x, \not{p}_y)$ by imposing the condition that the invariant mass $M_{\ell\nu} = M_W$. Note that here it is assumed that the only source of missing energy is the neutrino originating from the leptonic decay of W. Using the 4-vector of the isolated lepton $p^\mu = (E^\ell, p_x^\ell, p_y^\ell, p_z^\ell)$, arising from the decay of the W, one can construct the longitudinal component (and hence the energy) of the missing momentum as follows:

$$\not{p}_z = \frac{1}{2(E^{\ell 2} - p_z^{\ell 2})}[p_z^\ell(2p_x^\ell \not{p}_x + 2p_y^\ell \not{p}_y - m_\ell^2 + M_W^2) \pm \sqrt{\Delta}], \tag{9}$$

where the quantity Δ is given by

$$\Delta = E^{\ell 2}[(2p_x^\ell \not{p}_x + 2p_y^\ell \not{p}_y - m_\ell^2 + M_W^2)^2 - 4\not{p}_T^2(E^{\ell 2} - p_z^{\ell 2})], \tag{10}$$

with m_ℓ being the mass of the lepton and M_W being the input mass for the W boson. This provides us with two values of $\not{p}_z$ corresponding to the two signs of the square root. For certain configurations, however, one may obtain $\Delta < 0$, rendering the calculated $\not{p}_z$ complex and thus unphysical. In these cases, one can recalculate the missing energy by finding those values of $\not{p}_T$ for which $\Delta \geq 0$,

$$\begin{aligned}\not{p}_T = {} & \frac{1}{2(E^{\ell 2} - p_z^{\ell 2} - (p_x^\ell \cos\phi + p_y^\ell \sin\phi)^2)} \\ & \times [-(p_x^\ell \cos\phi + p_y^\ell \sin\phi)(m_\ell^2 - M_W^2) \\ & \pm \sqrt{(m_\ell^2 - M_W^2)^2(E^{\ell 2} - p_z^{\ell 2})}].\end{aligned} \tag{11}$$

For each sign of the square root in Eq. (11), we get a value of $\not{p}_T$, which when substituted in Eq. (9) gives two values of $\not{p}_z$ for every value of $\not{p}_T$. Thus, we end up with four values of $\not{p}_z$ in this case, instead of just two as in the earlier case.

For each value of the z component of the MET (i.e., $\not{p}_z$), we can reconstruct the leptonically decaying top quark mass by combining the 4-momenta of the lepton, b-jets, and the missing energy. Several reconstructed mass combinations can exist depending on the number of solutions of $\not{p}_z$ and the fact that there are two b-tagged jets to choose from.

To obtain the optimal values of the leptonic and hadronic top quark masses in each event, a minimum-χ^2 approach is adopted with the χ^2 defined as

$$\chi^2 = \frac{(m_{tH} - m_t)^2}{\sigma_{m_{tH}}^2} + \frac{(m_{tL} - m_t)^2}{\sigma_{m_{tL}}^2}, \tag{12}$$

where $\sigma_{m_{tL}}$ and $\sigma_{m_{tH}}$ represent the uncertainty in top quark mass measurement for leptonically and hadronically decaying tops, respectively, at the LHC. We consider $\sigma_{m_{tL}} = 2.7$ GeV and $\sigma_{m_{tH}} = 1.15$ GeV [67,68]. Using the 4-momentum information of the isolated lepton and

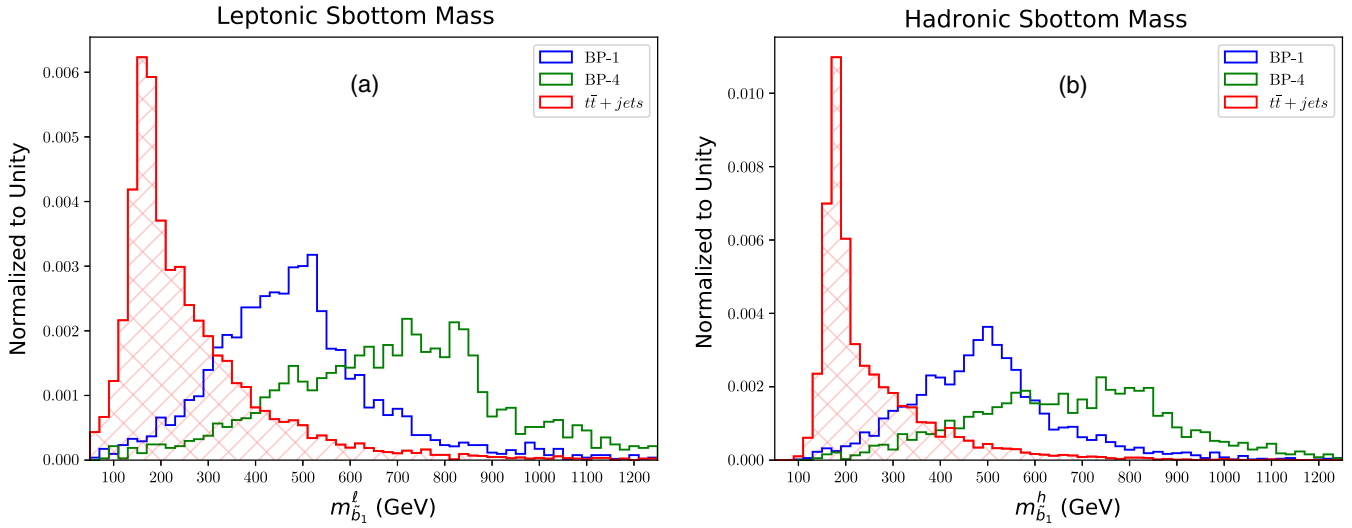

FIG. 4. Reconstructed bottom squark masses for BP-1, BP-4, and the $t\bar{t}$ events, left (a) denotes the case when the top quark decays leptonically, while the right (b) signifies the hadronically decaying top quark scenario. For details, see the text.

missing transverse energy along with the two b-tagged jets and the leading four non-b-tagged jets, we reconstruct the leptonic and hadronic top quark masses. The combination which leads to the lower χ^2 value is chosen. Nice resonance peaks around top quark mass are observed for both the leptonically and hadronically decaying tops for all the benchmark points.

After the reconstruction of two top quarks, we are now left with the final reconstruction of the bottom squark mass using these two reconstructed top quarks and the two remaining light quark jets originating from the decay of the two bottom squarks. For each reconstructed top mass, there are two possible choices to combine the light jets for the reconstruction of the bottom squark mass. We select the combination which leads to the least difference between the reconstructed mass of the leptonically decaying bottom squark and the hadronically decaying bottom squark. The plot for the reconstructed bottom squark for BP-1 (corresponding to a 500 GeV bottom squark) and for BP-4 (corresponding to a 800 GeV bottom squark) are shown in Fig. 4, where the left and right panels denote the reconstruction method involving the leptonically and hadronically decaying top quarks. The reconstructed bottom squark masses peak at the truth masses for the two benchmark points, while for $t\bar{t}$ events it peaks near the truth top quark mass. The peaks corresponding to the signal events are significantly distinct from that of the backgrounds, and this motivates us to consider the reconstructed masses as the BDT inputs.

For each benchmark point, the variation of signal significance with the BDT cut value has been shown in plot (b) of Fig. 3, and Table V shows the best signal significance with the corresponding BDT cut values assuming $\mathcal{L} = 300$ fb^{-1}. Clearly, the MVA improves the reach of the search compared to the cut-based analysis, e.g., signal significance improves from 10.7 to 20.9 for BP-1, resulting in an increase of the discovery reach of $\sim$100 GeV in the mass of the $\tilde{b}_1$. The variation of signal significance with integrated luminosity is shown in Fig. 5, with solid lines for the cut-based analysis and dashed lines for the MVA. One can observe that with 2000 fb^{-1} data, $\tilde{b}_1$ mass of up to 1 TeV can be explored at the high luminosity run of LHC.

V. THE HADRONIC FINAL STATE

We now consider the case where both the top quarks decay fully hadronically. The fully hadronic final state is difficult to investigate at the LHC because of the overwhelming QCD background. However, in our signal

TABLE V. Signal significance for the benchmark points with the choice of BDT cuts with $\mathcal{L} = 300$ fb^{-1} of integrated luminosity.

	BP-1	BP-2	BP-3	BP-4	BP-5	BP-6
$m_{\tilde{b}_1}$ (GeV)	500	600	700	800	900	1000
BDT cut	0.231	0.234	0.258	0.230	0.311	0.294
$\mathcal{S} = \frac{S}{\sqrt{S+B}}$	20.9	9.9	4.7	2.7	1.6	0.9

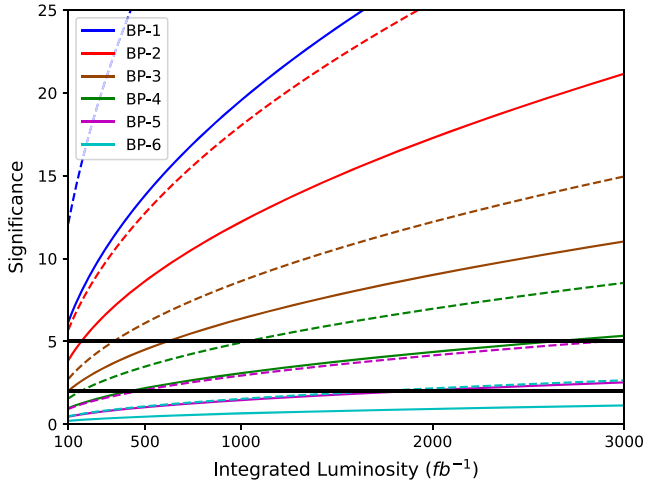

FIG. 5. Plot of significance versus the integrated luminosity. While the inset legend shows the color for the different benchmark points, the solid line corresponds to the significance corresponding to the cut-based analysis and the dashed line to that provided by the multivariate analysis. Horizontal lines at 2σ and 5σ indicate the potential for exclusion and discovery.

events, the top quarks are expected to be boosted such that the three quarks from its decay form a "fat jet" with substructure. Our plan is to exploit the substructure of such a fat jet to identify a top quark and investigate the reach for bottom squarks using 13 TeV data from the LHC.

Our final state will contain only reclustered fat jets, and we shall attempt to tag some of these jets as tops. The background should ideally have contributions from all the SM processes we considered in the leptonic counterpart—$t\bar{t}+$jets (up to two), $t\bar{t}b\bar{b}$, $t\bar{t}Z$, $t\bar{t}W$, and $t\bar{t}H$—in addition to the QCD multijet, but for all practical purposes the QCD multijet processes and $t\bar{t}+$jets (up to two) contribute so overwhelmingly to the background (even after cuts) that we really need not consider the other processes. In this section, we work with this simplifying assumption about the background. It is to be noted that while simulating the QCD multijet events, we restrict ourselves to up to four jets at the parton level (light quarks and gluons only) due to our computational limitations. However, once parton showering is switched on, the jet multiplicity can and does become larger.

Our strategy is to tag at least one top quark in each signal event. For this purpose, we use HEPTopTagger [69], which is quite efficient for tagging tops with moderate boosts ($p_T \gtrsim 200$ GeV). We avail ourselves of the energy flow of the particles, provided in the EFlow branch of the DELPHES generated ROOT file to obtain the particle information. We use FASTJET to construct fat jets of R = 1.8 using the anti-kT jet algorithm with a minimum p_T of 30 GeV. The jets with $p_T > 200$ GeV and $|\eta| < 3$ are then selected to pass through the HEPTopTagger. Before they enter the toptagger, these jets are reclustered exclusively with the Cambridge-Aachen (C/A) algorithm with the same jet radius (viz. $R = 1.8$). The default settings of HEPTopTagger were used: the mass drop required for jet splitting was set at $\min(m_{j_1}, m_{j_2})/m_j = \mu < 0.8$ with the minimum mass of a subjet $m_{\rm sub}^{\rm min} = 30$ GeV, where j_1 and j_2 are the subjets of the fat jet j. The top and W masses are reconstructed on a set of filtered subjets numbering no more than $N_{\rm filt} = 5$. Tops are tagged with masses in the range between $m_{\rm top}^{\rm min} = 140$ GeV and $m_{\rm top}^{\rm max} = 200$ GeV. We achieve an efficiency of about 30% using these conditions for a moderate (~200 GeV) to high (say 600 GeV or more) p_T regime. The choice of a large jet radius indicates that we are required to incorporate some jet grooming technique in order to get rid of soft and large angle radiations as well as underlying events. In our analysis, we use a particular technique, named jet trimming [70], which has been found to be very effective in grooming large R jets. This grooming technique involves two independent parameters, namely $R_{\rm trim}$ and $p_T^{\rm frac}$. The prescription is to essentially recluster the constituents of a given jet with a smaller jet radius $R_{\rm trim}$ and then keep those subjets with p_T greater than a fixed fraction, $p_T^{\rm frac}$ of the input jet p_T. In our analysis, we optimize these two parameters and choose $R_{\rm trim} = 0.4$ and $p_T^{\rm frac} = 1\%$. These trimmed jets, obtained after trimming the original anti-kT jets, are used for further analysis.

A. Multivariate analysis

After passing the jets to the HEPTopTagger, we select the events containing at least one top-tagged jet. The complete event information is used to construct different observables, and these, in turn, are used to perform a multivariate analysis using the TMVA framework.

Once we have successfully described the full event information in terms of jets, we classify the different types of jets as top-tagged jets, b-tagged jets, and "light" jets (nontop-tagged, non-b-tagged jets). For b-tagging, we calculate the angular distance between a jet and the b hadron, and make sure that the separation $\Delta R < 0.5$. Furthermore, we also take into account a p_T-dependent b-tagging efficiency given by [54]

$$\epsilon_b = \begin{cases} 0.5 & \text{for } p_T^b \leq 50 \text{ GeV} \\ 0.75 & \text{for } 50 \text{ GeV} < p_T^b \leq 400 \text{ GeV} \\ 0.5 & \text{for } p_T > 400 \text{ GeV}. \end{cases} \quad (13)$$

Note that the above-mentioned efficiencies are conservative estimates; with more data and improved algorithms we expect significant improvement in b-tagging efficiencies. Finally, jets which are not tagged either as "top jets" or "b jets" are called "light jets."

Not only are the light jets in the signal sample often harder than those in the background, an analogous statement also holds for the respective top jet constituent (especially for heavier bottom squarks). To utilize these

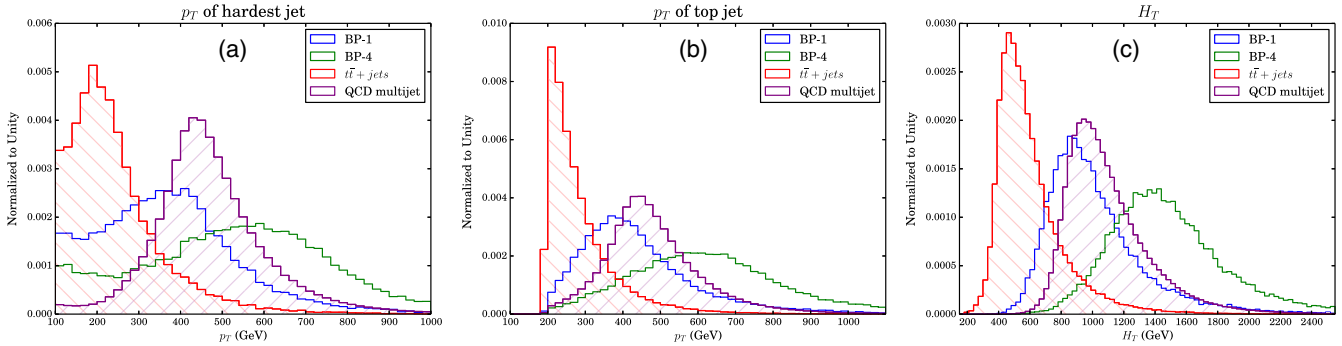

FIG. 6. The plots for a few of the variables used in the analysis. The left plot (a) shows the p_T of the hardest light jet (nontop-tagged, non-b-tagged jet), while the plot in the middle (b) shows the p_T of the hardest top-tagged jet. The rightmost figure (c) is that of the H_T, which is the scalar sum of all the jets. In all of these, only two benchmark points (BP-1 and BP-4) have been shown, and the histograms for the background processes are hatched.

characteristic differences, between the signal and background events, we consider the p_T of the hardest top jet and the p_T of the hardest and second hardest light jets as BDT input variables. Like the leptonic analysis, one of the most important variables is H_T [see Eq. (5)], with the sum, obviously, running over all the jets. Being closely associated with the center-of-mass energy of the process, it too is an important discriminator. It is important to remind our readers here that we use only trimmed jets to construct the jet observables. We use the number of b-tagged jets as a discriminator by passing it as a variable for MVA, as QCD decreases vastly if a b-tag is demanded. We could, instead, have put a cut on it before the MVA—a pre-MVA cut—but as this would decrease the background a lot, making the BDT analysis somewhat unreliable, we desist. In Fig. 6, we plot the distributions in the p_T of the hardest light jet, that of the hardest top-tagged jet, and H_T. The QCD multijet sample was generated with an imposed cut of 1 TeV on the H_T and after demanding that the two hardest jets in the sample be harder than 100 GeV. With the center-of-mass energy of the bottom squark pair production process being ∼1 TeV, this ensures ample yet relevant statistics for the QCD multijet process. The variable H_T turns out to be a good discriminator as the peak of higher mass benchmarks lies to the right of the QCD peak, while the tail of the distribution only contributes to the signal peaks.

Restricting ourselves to events with a tagged top and at least four jets overall,[5] two more useful observables are obtained by partitioning an event such that one subset contains the tagged top and a single non-b jet, while the other contains the rest of the jets. Denoting the invariant masses of the two sets by m_{tj} and m_{jets}, we retain these variables for the pairing that minimizes the difference

$$\Delta M \equiv |m(j_l, j_i) - m(j_k, j_l, \ldots)|. \tag{14}$$

Ideally, ΔM should vanish. However, owing to the vagaries of jet reconstruction algorithms as well as detector effects, this would rarely occur. Note that the requirement of the top's partner above being a non-b jet helps get rid of significant amount of the QCD background in the signal peak region. Note that, among the two invariant masses m_{tj} and m_{jets} and the mass difference ΔM, only two are independent parameters. In the MVA analysis, however, we use all three parameters simultaneously as BDT inputs. In Fig. 7, we plot the two invariant masses we talked about earlier. These seem to have moderate discriminatory powers. Furthermore, in Fig. 8, we also show the correlations in the m_{tj}-m_{jets} plane for the QCD multijets (left plot), and for two benchmark points—BP-4 (middle) and BP-6 (right). For the QCD, the points are dense in the region around the (500, 500) point, while for the signal it is dense around (800, 800) for BP4 and around (1000, 1000) for BP6. It is interesting to note that this feature, in principle, can be used for probing heavier bottom squarks.

Nsubjettiness [71] is an inclusive jet shape variable which takes into account the energy distribution within a fat jet. It is defined as

$$\tau_N = \frac{\sum_k p_{T,k} \min\left(\Delta R_{1,k}, \Delta R_{2,k} \ldots \Delta R_{N,k}\right)}{\sum_k p_{T,k} R_0}, \tag{15}$$

where $\Delta R_{j,k}$ is the angular separation between the jth candidate jet and the kth constituent particle, $p_{T,k}$ is the p_T of the kth constituent, and R_0 is the jet radius of the fat jet under consideration. Normalization ensures that $0 \leq \tau_N \leq 1$. If $\tau_N \approx 0$, it indicates that all the radiation in the jet is aligned with the subjet directions and that there is a maximum of N subjets in the considered jet. On the

[5]Our primary event selection criteria includes at least one top-tagged jet; however, for the reconstruction of invariant masses we restrict ourselves to exactly one top-tagged event. In principle, two or more top-tagged samples would give better mass peaks with negligible QCD events; however, we find a very few signal events surviving the two or more top-tagged jet selection criteria.

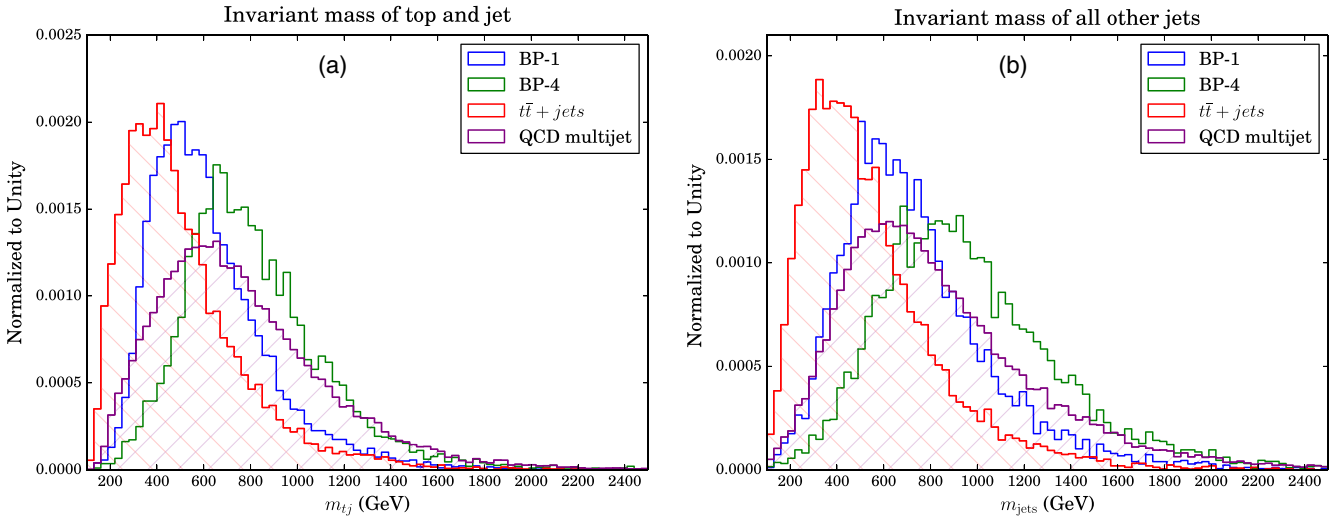

FIG. 7. The reconstructed masses that we used for the multivariate analysis. The plot on the left (a) is the invariant mass of the top with one of the light jets (in short, "tj" set), while that on the right (b) is the invariant mass of all the other jets in that event which do not correspond to the tj set. The invariant mass reconstruction technique has been discussed in the text in detail. Only two benchmarks (BP-1 and BP-4) are plotted, and the background histograms are hatched.

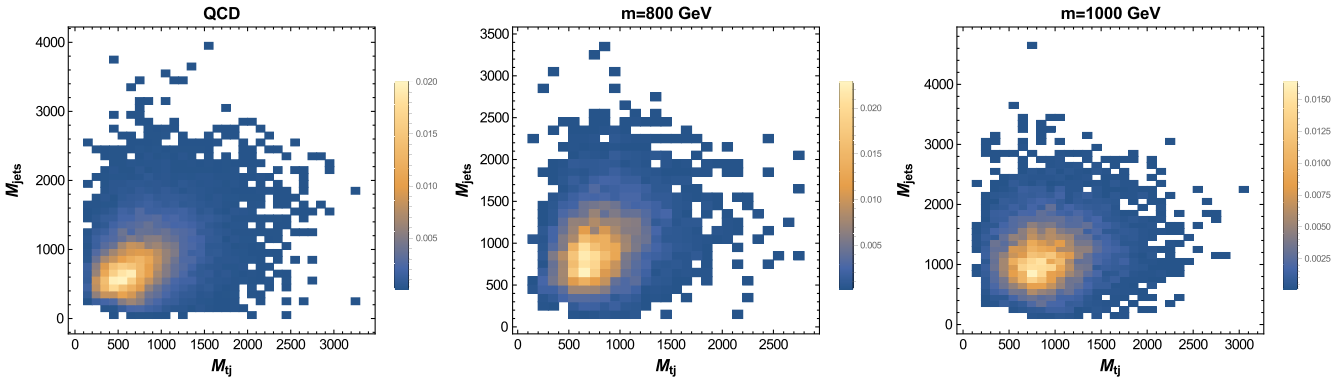

FIG. 8. Color plots showing correlation between the m_{tj} and $m_{\rm jets}$ variables in the case of QCD (left), BP-4 (middle), and BP-6 (right). Strong correlation can clearly be seen around the (500, 500) point for QCD, (800, 800) for BP-4, and (1000, 1000) for BP-6. This can be exploited to probe heavier bottom squarks as well.

other hand, if $\tau_N \gg 0$, it indicates the presence of more subjets and that the radiation is distributed far from the candidate subjets. It turns out that the ratio between two Nsubjettiness variables might have higher discriminatory power than the variables themselves. For events with at least one top-tagged jet, we calculate three such ratios, τ_{21}, τ_{31}, and τ_{32}, where $\tau_{ij} = \tau_i/\tau_j$, associated to the leading top-tagged jet. The ideal top-tagged jet should have a three-prong structure, and thus τ_{31} and τ_{32} are expected to be small, while this would not be true for the QCD background. Thus, these can be used as good discriminating variables. In Table VI, we list all the above-mentioned variables that are passed to the BDT for the MVA.

Two further observables, namely ρ and $\Phi(t, j)$, are used as BDT inputs, with the former being defined as

$$\rho = \frac{(p_T)_{j_t^{(1)}}}{(p_T)_{j_\ell^{(1)}}}, \tag{16}$$

where $(p_T)_{j_t^{(i)}}$ is the p_T of the ith top jet, while $(p_T)_{j_\ell^{(i)}}$ is the p_T of the ith light jet (i.e., a nontop-tagged, non-b-tagged jet). The quantity $\Phi(t, j)$ measures the azimuthal angular separation between the top-tagged jet and the leading light jet.

B. Results

We now proceed to discuss the details of the multivariate analysis using the BDT method implemented in the TMVA ROOT framework. The fifteen variables discussed earlier

TABLE VI. List of all the variables used in the multivariate analysis. Note that the variable ρ_2 is calculated only for events with two or more tagged tops.

	Variable	Definition
1.	nlJet	The number of light jets in the event
2.	nbJet	The number of b-tagged jets in the event
3.	ntJet	The number of top-tagged jets in the event
4.	$(p_T)_{j_1}$	p_T of the hardest light jet
5.	$(p_T)_{j_2}$	p_T of the second hardest light jet
6.	$(p_T)_{j_t^{(1)}}$	p_T of the hardest top-tagged jet
7.	H_T	scalar sum of the p_T of all the jets
8.	m_{tj}	the invariant mass of the top and jet system
9.	m_{jets}	the invariant mass of all the other jets
10.	$\Delta M = \lvert m_{tj} - m_{\text{jets}} \rvert$	the mass difference of the two reconstructed invariant masses
11.	$\tau_{21} = \tau_2/\tau_1$	Ratio of the Nsubjettiness variables
12.	$\tau_{31} = \tau_3/\tau_1$	Ratio of the Nsubjettiness variables
13.	$\tau_{32} = \tau_3/\tau_2$	Ratio of the Nsubjettiness variables
14.	$\rho = \frac{(p_T)_{j_t^{(1)}}}{(p_T)_{j_\ell^{(1)}}}$	Ratio of the hardest top jet p_T and light jet p_T
15.	$\Phi(t, j)$	Azimuthal angle separation between the top-tagged jet and the leading light jet.

TABLE VII. Showing the initial cross section (σ_0) and surviving cross section after at least one top is tagged (σ_{toptag}) for the background and all the signal benchmarks. The QCD multijet sample is generated after a cut on the H_T variable of 800 GeV and cut of 100 GeV on the p_T of the two hardest jets.

	QCD	$t\bar{t}$ + jets	BP-1	BP-2	BP-3	BP-4	BP-5	BP-6
σ_0 (fb)	1.9×10^7	8.3×10^5	5.2×10^2	1.8×10^2	6.7×10^1	2.8×10^1	1.3×10^1	6.2
σ_{toptag} (fb)	2.6×10^6	6.7×10^4	1.4×10^2	5.0×10^1	2.0×10^1	8.6	4.0	2.0

and listed in Table VI, each of which we expect to have some discrimination power, are used as the BDT inputs. The BDT parameters are same as in the leptonic case, viz. NTrees = 400, MaxDepth = 5, and MinNodeSize = 2.5% with AdaBoostBeta = 0.5. In Table VII, we show the initial cross sections (σ_0) and the cross section after at least one top is tagged (σ_{toptag}). The top-tagged events in the QCD samples are due to misidentification of fat jets as top jets; the corresponding "fake rate" is about 10% for the QCD sample. The advantage of using the multivariate analysis is that we can translate a complicated multidimensional optimization problem over all input variables into one involving a one-parameter function which is much easier to handle. We can now choose the BDT cut value such that it maximizes the signal significance. The results of the multivariate analysis are presented in Table VIII and Fig. 9.

From Table VIII, it is evident that the signal significance for the benchmark points diminishes rapidly as we proceed from BP-1 to BP-6. The primary reason is the rapid decrease of the bottom squark pair production cross section with the increase of bottom squark mass. Even though we expect to tag the top quarks originating from the heavier bottom squarks more efficiently, the impact is negligible compared to the drastic fall in the production cross section. Improved top-tagging with a smaller fake rate for the QCD multijet events is essential for better signal-background discrimination. State-of-the-art jet grooming techniques, namely Pruning [72] and SoftDrop [73] may help in reducing QCD multijet events, thus enhancing the signal significance.

It is thus clear that the hadronic channel is not the most favorable one for the discovery of the bottom squark. The exclusion limit is barely reached, for 300 fb^{-1} integrated luminosity, for the first benchmark point (500 GeV). A higher integrated luminosity of 3000 fb^{-1} from a futuristic collider like the high luminosity LHC will be able to push the exclusion limit to about 650 GeV. However, here we would like to suggest an interesting extension of this

TABLE VIII. Signal significance for the benchmark points with the choice of BDT cuts with $\mathcal{L} = 300$ fb^{-1} of integrated luminosity.

	BP-1	BP-2	BP-3	BP-4	BP-5	BP-6
$m_{\tilde{b}_1}$ (GeV)	500	600	700	800	900	1000
BDT cut	0.186	0.167	0.245	0.238	0.266	0.280
$\mathcal{S} = \frac{S}{\sqrt{S+B}}$	3.50	1.21	0.57	0.32	0.20	0.16

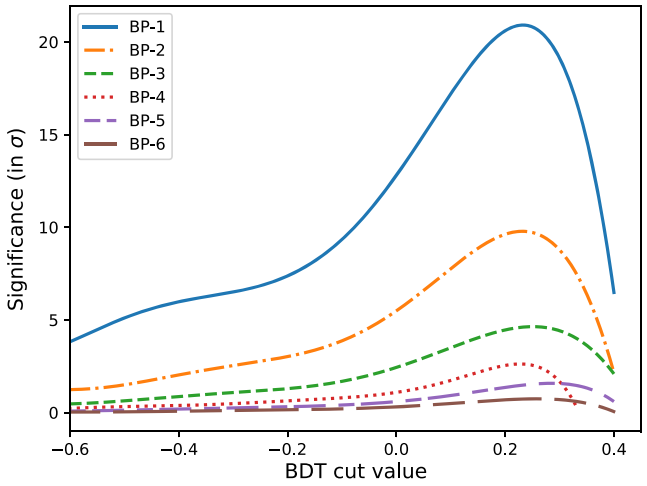

FIG. 9. Plot showing the significance of the different benchmark points with variation of the BDT cut value. All figures for 300 fb^{-1}

analysis which combines our leptonic and hadronic analyses, a successful marriage of the boosted and nonboosted analyses with semileptonic final states. We will leave this very interesting avenue for our future work.

VI. SUMMARY AND OUTLOOK

In this paper, we analyze the discovery potential of the LHC for a bottom squark (the LSP) which decays, with a 100% branching ratio, to the top and a light quark via R-parity violating UDD couplings. While relatively heavy squarks allow for large couplings, thereby opening up the possibility of significant resonance production (such as $d + b \to \tilde{t}^*$), we eschew this possibility altogether, assuming the couplings are small enough for them to be unimportant in production processes (whether resonance or pair), yet large enough to preclude recognizably displaced vertices.

Based on the final state, we devise two strategies, one for a final state which has at least one isolated lepton (electron or muon) and the other for a fully hadronic final state. For the leptonic state, two independent investigations have been performed: first, using the traditional cut-based analysis and then using a multivariate analysis. The backgrounds considered for the leptonic analysis are $t + \bar{j}$ets, $t\bar{t}b\bar{b}$, $t\bar{t}W$, $t\bar{t}H$, and $t\bar{t}Z$. After demanding an isolated lepton tag, we consider cuts on various observables, like H_T and M_{T2} among others, in order to separate signal from background. We also reconstruct the bottom squark mass with exactly one isolated lepton with two or more b-tagged jets and four hardest light (i.e., non-b-tagged) jets in the event. We use these reconstructed bottom squark masses, namely $m^h_{\tilde{b}_1}$ and $m^\ell_{\tilde{b}_1}$, representing the reconstructed masses using the hadronically and leptonically decaying top quarks, respectively, as BDT inputs. While the cut-based analysis reveals an exclusion of ∼750 GeV of the bottom squark mass, the MVA extends that range to ∼850 GeV with 300 fb^{-1} of data.

For the fully hadronic final state, we perform the MVA directly as we find through our leptonic analysis that it helps to improve the reach for heavy resonance. The dominant QCD multijet and the $(t\bar{t} + \text{jets})$ backgrounds drown out all other sources of SM backgrounds. We consider events in which we can tag at least one top jet using the HEPTopTagger framework. Furthermore, in order to reduce the effect of underlying events and soft radiation, we groom the large R anti-kT jets using the "trimming" technique. Several observables are then constructed using these trimmed jets and then passed to the MVA. The results, unfortunately, are not as good as in the leptonic channel, with the exclusion limit barely crossing 500 GeV.

The sensitivity that our analyses project can be further improved upon the inclusion of other aspects. We list a few here:

(a) Incorporating tracker information such as number of soft tracks [74], not associated with the reconstructed objects, is likely to help in improving the sensitivity, especially for the most challenging case, viz. the fully hadronic final state.

(b) As we have already mentioned, the very couplings that we have investigated here also lead to the top squark decaying to a bottom and a light quark. Although the background processes to the corresponding final state (two b + d/s pairs) have larger production rates than the one here, the simpler nature of the final state, especially the ability to reconstruct the masses [37,38], allows for a higher experimental sensitivity [20]. In this work, we have deliberately avoided this channel, assuming the top squark to be much heavier. If it is not so but is comparable to the bottom squark in mass, the sensitivities need to be compounded.

(c) The very same coupling will also lead to decays like $\tilde{d}(\tilde{s}) \to \bar{t} + \bar{b}$ (depending on the identity of the coupling). Once again, we have not included this assuming that the $\tilde{d}(\tilde{s})$ is much heavier. This assumption was partly motivated by the need to keep large FCNCs at bay. However, a second solution exists if the squark masses are relatively degenerate [75–77]. This can be motivated if the soft-supersymmetry breaking masses for the right-handed squarks are similar, and so are the small trilinear terms A_d, A_s, A_b. As can be readily appreciated, this solution is more natural than the one we have considered here.

Direct two-body decays of $\tilde{d}(\tilde{s})$ that are nearly degenerate with the $\tilde{b}$ would lead to configurations very similar to the one we have considered here, with the added advantage that the nontop jets here originate from b quarks and thus can be tagged. This would severely curtail the SM backgrounds (with the biggest effect being seen in the fully hadronic state), resulting in much improved sensitivity.

(d) Indeed, even if the $\tilde{d}(\tilde{s})$ are sufficiently heavier than the $\tilde{b}$ (on account of a possibly large A_b, the effects due to $A_{d,s}$ of similar magnitudes being smaller) so as to open up their R-conserving three-body decays (into $\tilde{b}$ accompanied by a pair of quarks), the associated quarks would lead to only soft jets. Thus, for such a cascade, one essentially comes back to the configuration that we have analyzed here.

(e) Finally, both top squarks and bottom squarks (and, similarly, the other squarks) can originate from gluinos. If the gluino is not much heavier than the quark, its production cross section is much larger. Such a gluino would decay into the squark-quark pair. The latter would lead to a soft jet, with the first suffering a R-violating decay leading to a configuration very similar to the one under consideration. This is quite analogous to the ATLAS study [22] (that set a limit of $m_{\tilde{g}} > 1.08$ TeV), except for the fact that, in the present context, some of the jets would be rather soft.

On the other hand, if the gluino is very heavy, the produced squarks will be highly boosted, providing highly boosted tops in turn. This suits top tagger algorithms favorably and has been analyzed in [78].

In view of these obvious improvements to the sensitivity that can be effected, it is quite apparent that the conclusions reached by us are only conservative.

ACKNOWLEDGMENTS

A. C. and D. B. wish to thank Gouranga Kole, Soureek Mitra, Shankha Banerjee, and Seema Sharma for helping us resolve various ROOT and TMVA-related issues. A. C. and D. B. would also like to thank Sabyasachi Chakraborty for many fruitful discussions. D. C. acknowledges partial support from the European Union's Horizon 2020 research and innovation program under Marie Sklodowska-Curie Grant No. 674896. The computations reported here were performed on the computational resources of the Department of Theoretical Physics, TIFR. D. B. thanks the Department of Theoretical Physics, IACS for hospitality during the completion of this work.

[1] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).
[2] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).
[3] A. M. Sirunyan *et al.* (CMS Collaboration), arXiv: 1705.04673; The ATLAS collaboration, Report No. ATLAS-CONF-2017-034; The ATLAS collaboration, Report No. ATLAS-CONF-2017-039; CMS Collaboration, Report No. CMS-PAS-SUS-17-001; CMS Collaboration, Report No. CMS-PAS-SUS-17-004; M. Aaboud *et al.* (ATLAS Collaboration), arXiv:1706.03731.
[4] R. Barbier *et al.*, Phys. Rep. **420**, 1 (2005).
[5] L. M. Carpenter, D. E. Kaplan, and E. J. Rhee, Phys. Rev. Lett. **99**, 211801 (2007).
[6] C. Balazs, M. Carena, and C. E. M. Wagner, Phys. Rev. D **70**, 015007 (2004).
[7] M. Carena, M. Quiros, M. Seco, and C. E. M. Wagner, Nucl. Phys. **B650**, 24 (2003).
[8] M. Carena, G. Nardini, M. Quiros, and C. E. M. Wagner, Nucl. Phys. **B812**, 243 (2009).
[9] H. Nishino *et al.* (Super-Kamiokande Collaboration), Phys. Rev. Lett. **102**, 141801 (2009); V. Takhistov *et al.* (Super-Kamiokande Collaboration), Phys. Rev. Lett. **115**, 121803 (2015).
[10] B. C. Allanach, A. Dedes, and H. K. Dreiner, Phys. Rev. D **69**, 115002 (2004); **72**, 079902(E) (2005).
[11] For example, it has been argued recently that the DM might be largely composed of ultralight axions. See, e.g., L. Hui, J. P. Ostriker, S. Tremaine, and E. Witten, Phys. Rev. D **95**, 043541 (2017).
[12] K. Tamvakis, Phys. Lett. B **382**, 251 (1996).
[13] ATLAS Collaboration, Report No. ATLAS-CONF-2016-037.
[14] ATLAS Collaboration, Report No. ATLAS-CONF-2016-075.
[15] M. Aaboud *et al.* (ATLAS Collaboration), Eur. Phys. J. C **76**, 541 (2016).
[16] CMS Collaboration, Report No. CMS-PAS-SUS-16-013.
[17] CMS Collaboration, Report No. CMS-PAS-SUS-13-010.
[18] CMS Collaboration, Report No. CMS-PAS-SUS-13-003.
[19] CMS Collaboration, Report No. CMS-PAS-SUS-14-003.
[20] The ATLAS collaboration, Report No. ATLAS-CONF-2016-022.
[21] The ATLAS collaboration, Report No. ATLAS-CONF-2016-084.
[22] The ATLAS collaboration, Report No. ATLAS-CONF-2016-057.
[23] J. L. Goity and M. Sher, Phys. Lett. B **346**, 69 (1995); **385**, 500(E) (1996).
[24] G. Bhattacharyya, D. Choudhury, and K. Sridhar, Phys. Lett. B **355**, 193 (1995); D. Bardhan, G. Bhattacharyya, D. Ghosh, M. Patra, and S. Raychaudhuri, Phys. Rev. D **94**, 015026 (2016).
[25] B. Brahmachari and P. Roy, Phys. Rev. D **50**, R39 (1994); **51**, 3974(E) (1995).
[26] J. P. Saha and A. Kundu, Phys. Rev. D **66**, 054021 (2002); **69**, 016004 (2004).
[27] A. de Gouvea, S. Lola, and K. Tobe, Phys. Rev. D **63**, 035004 (2001); M. Chaichian and K. Huitu, Phys. Lett. B **384**, 157 (1996).
[28] G. Bhattacharyya, J. R. Ellis, and K. Sridhar, Mod. Phys. Lett. A **10**, 1583 (1995).

[29] G. Bhattacharyya and P. B. Pal, Phys. Rev. D **59**, 097701 (1999); A. Abada and M. Losada, Phys. Lett. B **492**, 310 (2000).
[30] K. Agashe and M. Graesser, Phys. Rev. D **54**, 4445 (1996).
[31] D. K. Ghosh, S. Raychaudhuri, and K. Sridhar, Phys. Lett. B **396**, 177 (1997); D. K. Ghosh, X. G. He, B. H. J. McKellar, and J. Q. Shi, J. High Energy Phys. 07 (2002) 067.
[32] K. A. Olive *et al.* (Particle Data Group Collaboration), Chin. Phys. C **38**, 090001 (2014).
[33] D. Chakraverty and D. Choudhury, Phys. Rev. D **63**, 075009 (2001).
[34] D. Chakraverty and D. Choudhury, Phys. Rev. D **63**, 112002 (2001).
[35] N. Zwane, arXiv:1505.03479.
[36] A. Monteux, J. High Energy Phys. 03 (2016) 216.
[37] D. Choudhury, M. Datta, and M. Maity, Phys. Rev. D **73**, 055013 (2006).
[38] D. Choudhury, M. Datta, and M. Maity, J. High Energy Phys. 10 (2011) 004.
[39] W. Porod, Comput. Phys. Commun. **153**, 275 (2003).
[40] W. Porod and F. Staub, Comput. Phys. Commun. **183**, 2458 (2012).
[41] F. Staub, arXiv:0806.0538.
[42] F. Staub, Adv. High Energy Phys. **2015**, 840780 (2015).
[43] W. Porod, F. Staub, and A. Vicente, Eur. Phys. J. C **74**, 2992 (2014).
[44] Y. Amhis *et al.* (Heavy Flavor Averaging Group Collaboration), arXiv:1412.7515.
[45] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, J. High Energy Phys. 07 (2014) 079.
[46] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, Comput. Phys. Commun. **191**, 159 (2015).
[47] T. Sjöstrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05 (2006) 026.
[48] J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi (DELPHES 3 Collaboration), J. High Energy Phys. 02 (2014) 057.
[49] M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. C **72**, 1896 (2012).
[50] M. Cacciari, G. P. Salam, and G. Soyez, J. High Energy Phys. 04 (2008) 063.
[51] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, J. High Energy Phys. 08 (1997) 001.
[52] C. Anders, C. Bernaciak, G. Kasieczka, T. Plehn, and T. Schell, Phys. Rev. D **89**, 074047 (2014).
[53] CMS Collaboration, Report No. CMS-PAS-BTV-13-001.
[54] The ATLAS collaboration, Report No. ATLAS-CONF-2014-046.
[55] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO.
[56] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt1314TeV2014
[57] F. Maltoni, D. Pagani, and I. Tsinikos, J. High Energy Phys. 02 (2016) 113.
[58] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, J. High Energy Phys. 07 (2002) 012.
[59] C. Borschensky, M. Krmer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell, Eur. Phys. J. C **74**, 3174 (2014).
[60] V. Khachatryan *et al.* (CMS Collaboration), Phys. Lett. B **746**, 132 (2015).
[61] A. Barr, C. Lester, and P. Stephens, J. Phys. G **29**, 2343 (2003).
[62] H. C. Cheng and Z. Han, J. High Energy Phys. 12 (2008) 063.
[63] A. Hocker *et al.*, *Proc. Sci.*, ACAT2007 (2007) 040 [arXiv:physics/0703039].
[64] I. Antcheva *et al.*, Comput. Phys. Commun. **180**, 2499 (2009).
[65] V. Simk, P. Homola, J. Valenta, and R. Leitner, Report Nos. ATL-PHYS-2001-018, ATL-COM-PHYS-99-073, CERN-ATL-PHYS-2001-018.
[66] Y. Bai and Z. Han, J. High Energy Phys. 04 (2009) 056.
[67] M. Aaboud *et al.* (ATLAS Collaboration), arXiv:1702.07546.
[68] A. M. Sirunyan *et al.* (CMS Collaboration), arXiv:1701.06228.
[69] T. Plehn, G. P. Salam, and M. Spannowsky, Phys. Rev. Lett. **104**, 111801 (2010); T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, J. High Energy Phys. 10 (2010) 078; T. Plehn and M. Spannowsky, J. Phys. G **39**, 083001 (2012).
[70] M. Cacciari, J. Rojo, G. P. Salam, and G. Soyez, J. High Energy Phys. 12 (2008) 032.
[71] J. Thaler and K. Van Tilburg, J. High Energy Phys. 03 (2011) 015.
[72] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Phys. Rev. D **80**, 051501 (2009).
[73] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, J. High Energy Phys. 05 (2014) 146.
[74] A. Chakraborty, S. Chakraborty, and T. S. Roy, Phys. Rev. D **94**, 111703 (2016).
[75] D. Choudhury, F. EFberlein, A. Konig, J. Louis, and S. Pokorski, Phys. Lett. B **342**, 180 (1995).
[76] M. Misiak, S. Pokorski, and J. Rosiek, Adv. Ser. Dir. High Energy Phys. **15**, 795 (1998).
[77] A. Brignole, L. E. Ibanez, and C. Munoz, Adv. Ser. Dir. High Energy Phys. **18**, 125 (1998).
[78] B. Bhattacherjee and A. Chakraborty, Phys. Rev. D **89**, 115016 (2014).