18 January 2024

Search for displaced top quark from a new massive particle decay in the tracker of CMS

CMS collaboration

Summary

A search for massive long-lived particle (LLP) decaying to a top quark in proton-proton collisions at $\sqrt{s} = 13$ TeV is presented in this paper. New long-lived particles are predicted in several extensions of the Standard Model (SM). In the R-parity violated Minimal SuperSymmetric Model (RPV-MSSM) considered, the lightest SuperSymmetric particle (LSP) is long-lived and decays into a top and a virtual stop quark which couples to a down and strange quark pair. Machine learning is used to distinguish signal displaced tracks from Standard Model prompt tracks for the reconstruction of displaced vertices.

ATLAS Doc:	EXO-24-ZZZZ	
EDMS Id:	123456	
Version:	0.1	
Last modified:	06:12 on 18 January 2024	
Prepared by:	**Checked by:**	**Approved by:**
A.Paul Vaucelle	D. Person E. Person	F. Person G. Person

Contents

List of Figures

List of Tables

1 Introduction

Many extensions of the Standard Model (SM) predict the existence of long-lived particles through weak couplings, high masses making the production for these new particle to be highly suppressed. However, long-lived particles are highly motivated by R-Parity Violated Minimal SuperSymmetry Model (RPV-MSSM) [13, 25, 40, 23], Split-SUSY [28, 29, 32, 33, 14, 22], weakly interacting massive particles (WIMPs) [41, 20, 19], Gauge Mediated supersymmetry breaking (GMSB) [27, 34, 30], hidden sector [37, 38, 39]

In this paper, we search for long-lived neutral SUSY particles decaying in the tracker into SM particles, as allowed by the RPV-MSSM. These long lived SUSY particles are pair-produced from p-p collisions at a center-of-mass energy of 13 TeV in 2016, 2017 and 2018 corresponding to an integrated luminosity of 137 fb^{-1}. This analysis looks for displaced vertices coming from the decay of the pair-produced long-lived particles where the latter decay into SM particles producing jets and tracks in the tracker volume. The event topology, as well as the tracks and the secondary displaced vertices can be used to discriminate the displaced signature from SM backgrounds. The analysis is focused on the pair-production of neutral long-lived SUSY particles in the RPV-MSSM and does not address the direct pair-production of neutralinos.

Searches about displaced jets/vertices at $\sqrt{13}$ TeV have been reported by ATLAS [17] and [18] and CMS [24] in an inclusive approach, where the displaced vertices are mainly search in the pixel detector due to the tracking efficiency and masses mainly at the TeV scale. Limits ... and cross section ...

2 MC simulation

2.1 Signal

The development of the search is performed using privately simulated signal samples. These samples are generated using the MADGRAPH5_aMC@NLO generator [21] at leading order (LO) plus 1 jet from an initial state radiation. The generation is produced using a private model described in Appendix.13.1. This model allows the decay of the neutralino into SUSY and SM particles following the RPV-MSSM. In this analysis, we focus on the pair-production of neutral long-lived SUSY particles as shown in Fig.1.

The signal process is characterized by the production of a Z/γ^* boson from p-p collisions at the LHC. Then, the boson produced decays into a pair of sleptons, being the Next-to-Lightest SuperSymmetric-Particle (NLSP), where the considered sleptons are the smuons. Since the slepton is the NLSP, it decays into a muon and the Lightest-SuperSymmetric-Particle (LSP) that is the neutralino. The branching ratio for this decay is set to 100% in order to reduce the complexity of the model. The neutralino can only be a bino-like neutralino to allow the decay through the $\lambda^{''}$ RPV-coupling to lead to violation of the baryonic number and the production of a top quark in the final state. The sleptons are short-lived particle leaving prompt leptons as final state particles of the signal process that can be use to trigger the signal events. Since the muons can be easily triggered on and easier to identify, the muon channel is chosen. The selectron channel can be added in the analysis in the future and the stau channel could also be considered from theory but is not studied in this analysis due to the complexity of the tau decay and the consequences on the event reconstruction.

The long-lived neutralino is assumed to decay into the tracker volume into a virtual stop and a top quark. The latter follows its SM decay channel where both hadronic and leptonic decays are considered for the W boson. The stop couples to SM quarks with the $\lambda^{''}$ RPV-coupling where the specific decay is given by $\lambda^{''}_{312}$ with 3,1 and 2 being the generations of quark considered. The stop couples to a down and a strange quark giving a heavy hadronic activity to the final state of the signal process. No mixing between particles of different generations is considered in the mixing matrices of sleptons and squarks, and mass states are in an equal proportion of left-handed and right-handed states also to reduce the complexity of the model.

For our signal, we consider a wide range of combination of values between the following parameters : the mass of the smuon $M_{\tilde{\mu}}$, the mass of the neutralino $M_{\tilde{\chi}^1_0}$, the proper lifetime of the neutralino $\beta\gamma c\tau$, the RPV coupling $\lambda^{''}_{312}$ and the mass of the virtual stop $M_{\tilde{t}}$. All envisaged combinations are described in Table.1. The upper limit on $M_{\tilde{\mu}}$ is due to a low cross section ($\sim 0.1fb$) but could be extended with more data-taking and the lower limit is due to the production of a top quark in the decay channel of the neutralino and also from previous experimental limits [12, 6, 8, 9, 7, 11, 10]. The cross-section σ of the pair-production of $\tilde{\mu}$ is entirely determined by $M_{\tilde{\mu}}$ as shown in Fig.2. For simplicity, all other RPV-couplings are null.

The phase space is described in more details in Fig.2 and Fig.3.

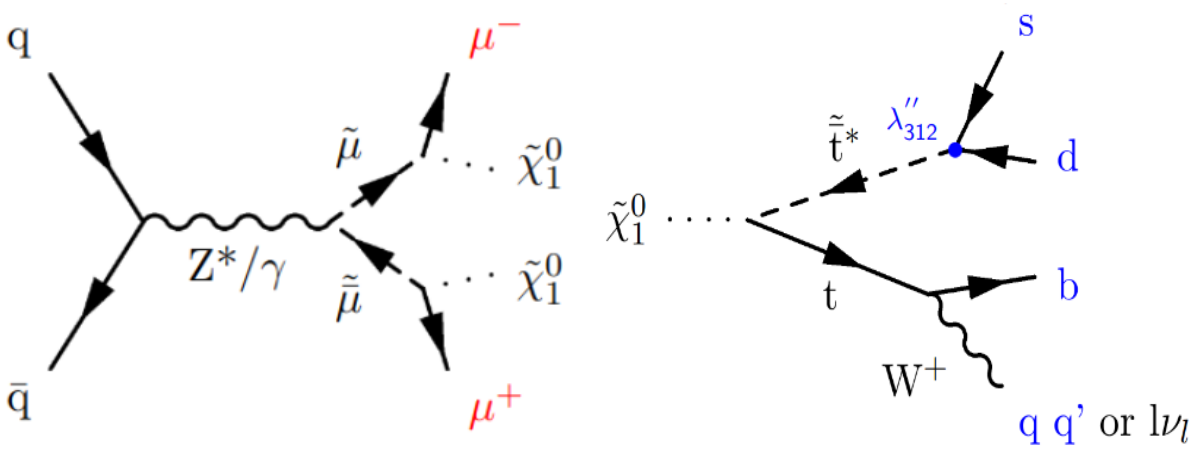

Figure 1: Neutralino production and decay channel

$\beta\gamma c\tau$(cm)	$M_{\tilde{\mu}}$ (GeV)	$M_{\tilde{\chi}_0^1}$ (GeV)	$M_{\tilde{t}}$(GeV)	λ''_{312}
0.1 to 100	**200 to 500**	**180 to 480**	1000	10^{-3} to 10^{-1}

Table 1: Masses of the SUSY particles and the associated couplings for the different benchmarks in order to probe the tracker volume

[THEN SHOW GENRETOR LEVEL VARIABLES??]

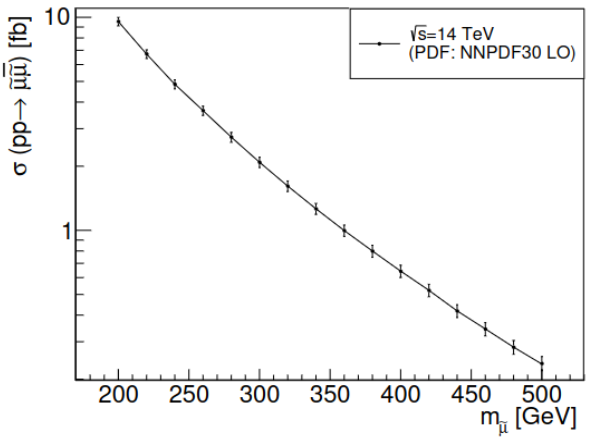

Figure 2: Cross-section of the smuons pair production as a function of the smuon mass

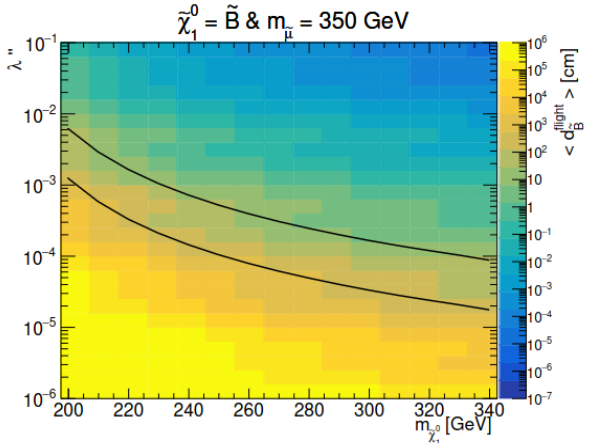

Figure 3: Mean decay length of the neutralino as a function of the neutralino mass and the coupling. The black lines are the limits of the volume of the tracker of CMS. The tracker volume puts constraints on the phase space to be probed.

2.2 Background

For the development of the analysis, MC samples generated in the Summer20 campaign for 2018 stored in the MiniAODv2 format, MiniAOD format that is used for the whole analysis. These samples are listed for each year of Run 2 in the Tables [2], [3] and [4]. [to be changed]

Sample 2016	No. events	σ [pb]
/TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8	546996	53.07
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	960000	88.3
/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	963768	10.8908
/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	920153	10.8707
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8	1836747	15910.0
/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8	800000	5379
/WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	780000	11.09
/WZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	994094	6.535
/ZZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	780170	3.676
/ttWJetsToLNu_5f_EWK_TuneCP5_13TeV_amcatnlo-pythia8	825000	0.290
/TTZToLL_5f_TuneCP5_13TeV-madgraphMLM-pythia8	660694	0.05188
/TTToHadronic_TuneCP5CR1_13TeV-powheg-pythia8	1067000	687.1
/TTWW_TuneCP5_13TeV-madgraph-pythia8	570000	0.006992

Table 2: Background MC samples fro 2016 with their associated number of events and cross-sections

3 Datasets

Since the experimental signature has two prompt leptons, this analysis can make use of several datasets depending on the flavor of the prompt leptons. For the muon channel :

- SingleMuon
- DoubleMuon

For the electron channel :

- SingleElectron, called EGamma in 2018
- DoubleElectron

For each dataset, the set of trigger that are used is given in Table5 for 2016, in Table6 for 2017 and in Table7 for 2018.

[ADD TABLE GLOBALT TAGS FOR EACH PERIODS data and MC + golden json files]

Sample 2017	No. events	σ [pb]
/TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8	546996	53.07
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	960000	88.3
/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	963768	10.8908
/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	920153	10.8707
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8	1836747	15910.0
/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8	800000	5379
/WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	780000	11.09
/WZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	994094	6.535
/ZZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	780170	3.676
/ttWJetsToLNu_5f_EWK_TuneCP5_13TeV_amcatnlo-pythia8	825000	0.290
/TTZToLL_5f_TuneCP5_13TeV-madgraphMLM-pythia8	660694	0.05188
/TTToHadronic_TuneCP5CR1_13TeV-powheg-pythia8	1067000	687.1
/TTWW_TuneCP5_13TeV-madgraph-pythia8	570000	0.006992

Table 3: Background MC samples for 2017 with their associated number of events and cross-sections

Sample 2018	No. events	σ [pb]
/TTJets_DiLept_TuneCP5_13TeV-madgraphMLM-pythia8	546996	53.07
/TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	960000	88.3
/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	963768	10.8908
/ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	920153	10.8707
/DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8	1836747	15910.0
/DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8	800000	5379
/WWTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	780000	11.09
/WZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	994094	6.535
/ZZTo2Q2L_mllmin4p0_TuneCP5_13TeV-amcatnloFXFX-pythia8	780170	3.676
/ttWJetsToLNu_5f_EWK_TuneCP5_13TeV_amcatnlo-pythia8	825000	0.290
/TTZToLL_5f_TuneCP5_13TeV-madgraphMLM-pythia8	660694	0.05188
/TTToHadronic_TuneCP5CR1_13TeV-powheg-pythia8	1067000	687.1
/TTWW_TuneCP5_13TeV-madgraph-pythia8	570000	0.006992

Table 4: Background MC samples fro 2018 with their associated number of events and cross-sections

Dataset	Trigger
SingleMuon	HLT_IsoMu27_v
DoubleMuon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL(_DZ)_v **OR** HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL(_DZ)_v **OR** HLT_TkMu17_TrkIsoVVL_TkMu8_TrkIsoVVL_(DZ)_v **OR** HLT_IsoMu24_v **OR** HLT_IsoTkMu24_v
SingleElectron	HLT_Ele27_WPTight_Gsf_v
DoubleElectron	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v **OR** HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v **OR** HLT_Ele27_WPTight_Gsf_v

Table 5: Trigger for the datasets of 2016

Dataset	Trigger
SingleMuon	HLT_IsoMu27_v
DoubleMuon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v **OR** HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8_v **OR** HLT_IsoMu24_v
SingleElectron	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v **OR** HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v
DoubleElectron	HLT_Ele32_WPTight_Gsf_v **OR** HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v **OR** HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v

Table 6: Trigger for the datasets of 2017

Dataset	Trigger
SingleMuon	HLT_IsoMu24_v
DoubleMuon	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v **OR** HLT_IsoMu24_v
SingleElectron	HLT_Ele32_WPTight_Gsf_v
DoubleElectron	HLT_Ele32_WPTight_Gsf_v **OR** HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v **OR** HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v

Table 7: Trigger for the datasets of 2018

4 Object Selection

Particle-Flow [36] is the algorithm used by the CMS experiment to reconstruct objects by using the information of all sub-detectors. Selections are applied on the following reconstructed objects : on muons for the muon channel, electrons for the electron channel, on both leptons when looking at the electron-muon channel (used for background estimation) and jets for all channels. The analysis is looking for prompt leptons with potential displaced leptons coming from heavy-flavour hadrons decay from the decay channel of the neutralino. Plus, the prompt leptons can overlap with the displaced jets increasing the complexity of the experimental signature of the signal. All these constraints have been taken into account for the selections of the different objects and of the events.

4.1 Muons

Muons are reconstructed using the information from the tracker and the muon chambers. In this analysis, muons are required to be in the acceptance of the muon detector,$|\eta| < 2.4$. All muons are required to have a $p_T > 3$ GeV. In the muon channel, there are two prompt muons coming from the decay of the smuons. These prompt muons are required to have constraints on their impact parameters, to be of opposite charge and also to follow the muonID and Isolation given in the Table.8.

Selection	Prompt	Other
GlobalMuon	True	True
$\|\eta\|$	2.4	2.4
p_T	3	3
$\|d_{dxy}\|$	0.1 cm	none
$\|d_{dz}\|$	0.2 cm	none
MediumID	True	True
TightID	True	False
TkIsoTight	True	False

Table 8: Selections for the Muons in the muon channel

[Add distribution of Muons at gen level]

Concerning muon corrections, the Rochester Correction [3] have been applied on the Monte-Carlo samples according to the Muon POG recommendations in order to correct the momentum scale and resolution of muons. [complete this part]

In addition to the Rochester Correction, the scale factors for the muns are computed according to the Muon POG [add ref : https://indico.cern.ch/event/1247210/sessions/478562/attachme ref : https://gitlab.cern.ch/cms-muonPOG/spark_tnp/] using the Spark T&P tool fro all three years of Run 2. This tool is used to correct the discrepancy between data and MC for the muons but also for the trigger used (only for the muon channel). Scale factors are given in bins of p_T and η.

[May be a more detailed part for the muon POG for the review :] In order to match the online and offline selection of the analysis, the following selection is applied for the

computation of the scale factors with Spark on the Drell-Yan resonance:

1. for 2016 post-APV :
2. for 2016 pre-APV :
3. for 2017 :
4. for 2018 : "$|probe\_dxy| < 0.1 and abs(probe\_dz) < 0.2 and abs(tag\_dxy) < 0.1 and abs(tag\_dz) < 0.2 and TightID == 1 and TkIsoTight == 1 and tag\_pt > 25 and probe\_pt > 10 and tag\_abseta < 2.4 and probe\_abseta < 2.4 and (tag\_HLT\_IsoMu24\_v == 1 or tag\_HLT\_Mu17\_TrkIsoVVL\_Mu8\_$ 1) $and pair\_mass > 10$"

The final scale factor is defined as the combination of three efficiencies :

$$SF^{\mu} = \epsilon^{\mu}_{TRK} * \epsilon^{\mu}_{ID|TRK} * \epsilon^{\mu}_{ISO|ID} * \epsilon^{\mu}_{Trigger|ISO} \quad (1)$$

where :

- ϵ^{μ}_{TRK} is the tracking efficiency really close to unity set to 1 by the MUON POG
- $\epsilon^{\mu}_{ID|TRK}$ is the ratio of muons passing a given Identification flag with respect to the number of tracker muons (to follow the computation of the Muon POG)
- $\epsilon^{\mu}_{ISO|ID}$ is the ratio of muons passing a given Isolation flag with respect to the number of muon passing the Identification flag
- $\epsilon^{\mu}_{Trigger|ISO}$ is the ratio of muons passing the trigger requirement with respect to the number of muons passing the Isolation criteria (and Identification flag).

These efficiencies are computed for each combination of ID/ISO/trigger of muons indicated in the Table.8. [add ref : https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2 https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs2018]

4.2 Electrons

Electrons are objects that have a dedicated tracking algorithm since their track can be modify by possible bremsstrahlung. The Gaussian Sum Filter tracking [add ref] algorithm takes into account the information of the tracker and the ECAL to reconstruct the track of electrons taking into account the bremsstrahlung.

These prompt electrons in the electron channel are required to have constraints on their impact parameters, to be of opposite charge and also to follow the electronID and Isolation given in the Table.9.

[Add distribution of Electrons at gen level]
talk about the electron corrections

Selection	$Prompt_{\|\eta\|<1.479}$	$Prompt_{\|\eta\|>1.556}$	Others
$\|\eta\|$	1.479	2.4	2.4
p_T	10	10	10
$\|d_{dxy}\|$	0.05 cm	0.10 cm	none
$\|d_{dz}\|$	0.10 cm	0.20 cm	none
TightID	True	True	???

Table 9: Selections for the prompt Electrons in the electron channel. The TightID refers to "cutBasedElectronID-Fall17-94X-V2-tight" working point from the EGamma POG.

4.3 Jets

This analysis aims at looking at Run 2 and Run 3 data. This involves different types of jets between the two data-taking periods. For Run 2, AK4PF jets are used where AK4PF indicates the clustering of Particle-Flow candidates through the anti-k_T algorithm [31] with a distance parameter of 0.4.The selection applied on jets are summarized in Table.10. Further selections will be applied on jets for the event reconstruction.

Selection	Value
p_T	20
TightLepVetoId	True

Table 10: Selections for the jets

The TightID refers to the recommendation of the Jets POG[1]. As a note, one could wonder the effect of the TightJetID on displaced vertices and on the final vertex reconstruction efficiency since jets are the main constituents of this physics process. It has been checked that the TightJetID and TightJetIDLepVeto select more than 90% and 80% respectively of the jets and that the final reconstruction efficiency of the vertices is not affected by this selection on the jets. The event reconstruction and vertex reconstruction algorithm are then robust with respect to the number of jets in the final state.

[add information about th jet energy correction,everything that is put in the python file for jets corrections]

[Add distribution of Jets at gen level]

5 Event Selection

This analysis can take advantage of the two prompt leptons to select the events compared to previous analyses that are mainly looking at an only-hadronic final state [add ref]. One lepton is always selected by the online trigger selection and the second lepton can come from the dilepton trigger mentioned in Table.7 or it can also be selected offline following a specific selection on the invariant mass of the two leptons selected, detailed below. The τ channel is not studied since its decay can significantly change the experimental signature and the use of leptonic triggers would be compromised as well as jets from the τ decay could overlap with the displaced jets and potentially in due a bias in the reconstruction of the event detailed in section6. In this analysis, the two prompt leptons must be of opposite charge since the neutralino is considered to be of Dirac nature. Adding to the selections introduced on prompt leptons in the previous section and the trigger mentioned in section3, Concerning the selection, the first selected muon is required to have a $p_T > 25$ GeV and the second muon selected is required to have a $p_T > 10$ GeV. The invariant mass of the two selected muons is required to be above 10 GeV in order to reduce the amount of low-mass resonances in the search.

This first Online+Offline selection allows to reduce backgrounds up to 90% but it is not enough in order to observe the signal in any part of the phase space. Therefore, a step further is implemented in order to reduce the background at event level using the topology of the event.

A multivariate analysis selection based on a Boosted Decision Tree (BDT) from the Toolkit for Multivariate Analysis (TMVA) [15] is implemented. The list of variables given as an input is given in Table 11. The distribution of the input variables between signal and backgrounds are given in the Appendix.13.4.

Variable	Definition
MET	Total Missing Transverse Energy
H_T	Sum of the p_T of the jets that passed through the selections in Table.10
S_T	Sum of the p_T of the muons that passed through the selections for **Prompt** muons in Table.8
nJets	Total number of jets that have passed the selections in Table.10
nPromptMuons	Total number of prompt muons that have passed the selections for **Prompt** muons in Table.8
nAllMuons	Total number of muons that have passed the selections for **other** muons in Table.8
NTracks	Total number of tracks in the event
$p_{T_{j1}}$	Jet leading p_T
$p_{T_{j2}}$	Jet sub-leading p_T
$p_{T_{j1}}$	Jet leading η
$p_{T_{j2}}$	Jet sub-leading η
$\Delta R_{j_1-j_2}$	ΔR between the leading and sub-leading jets
$\Delta\phi_{j_1-j_2}$	$\Delta\phi$ between the leading and sub-leading jets
$\Delta\eta_{j_1-j_2}$	$\Delta\eta$ between the leading and sub-leading jets
Hemi1 njet nomu	Number of jets from the first hemisphere having a ΔR above 0.4 with respect to the two prompt leptons
Hemi1 p_T	p_T of the first reconstructed hemisphere
Hemi1 eta	η of the first reconstructed hemisphere
Hemi1 phi	ϕ of the first reconstructed hemisphere
Hemi1 nTrks	(to be removed) number of tracks of the first reconstructed hemisphere
Hemi2 njet nomu	Number of jets from the second hemisphere having a ΔR above 0.4 with respect to the two prompt leptons
Hemi2 pt	p_T of the second reconstructed hemisphere
Hemi2 eta	η of the second reconstructed hemisphere
Hemi2 phi	ϕ of the second reconstructed hemisphere
Hemi2 nTrks	(to be removed) number of tracks of the second reconstructed hemisphere
Mass of Hemisphere 1	Mass of the first reconstructed hemisphere using the jets
Mass of Hemisphere 2	Mass of the second reconstructed hemisphere using the jets

Table 11: Input variables for the Event selection BDT

It is clear that the topology of the event depends on the point in the phase space of the signal that is chosen where the signal can be very displaced (dozens of centimeters down to a millimeter) and also the mass spectra can change the kinematic of the event considering the signal samples mentioned in Table.1.

...

One important feature of the signal is the $t\bar{t}$-like experimental signature where two b-jets are produced in the event. B-tagging is applied on slightly displaced vertices depending on the decay length of the B-Hadron. But b-tagging is not built to work on highly displaced jets (dozens of centimeters) and we show in Fig. the discrepancy in the behavior of the B-tagging DeepJet algorithm [add ref] with respect to the decay length and the mass spectra in Fig. and Fig. respectively. Therefore, B-Tagging is not a variable that is used in this analysis since its efficiency depends on the point in the phase space we want to look for.

[ADD all infos about the EVT BDT SEL OR CUT]

5.0.1 Dependence of the Event selection BDT with respect to the phase space

6 Event Reconstruction

The final state of the neutralino signal being heavily hadronic, the event reconstruction is based on the jets from the decay of the neutralino. These jets can come from the coupling between the stop and the down and strange quarks as well as the SM decay of the top. No particular restrictions are applied on the decay of the W boson from the top, both leptonic and hadronic decays are considered in this analysis and leptons originating from the W or heavy-hadron flavour decays are also considered in the event reconstruction. The event topology and kinematics depends on the phase space of the signal studied: the smuon and neutralino masses, as well as the lifetime of the neutralino.

In the signal, neutrinos are pair-produced from the decay of the smuons, and decay further away in the tracker. Both neutralinos tend to be back-to-back in azimuth (ϕ) but close in η, see Fig. 5, meaning that there are two clusters of tracks, mainly coming from jets, back-to-back in ϕ. The 3D-space is then divided into two hemispheres, one for each neutralino. A precaution is applied as the prompt muons can overlap with the displaced jets, so the momentum-vector of the prompt leptons are discarded in this building procedure to avoid any bias, then jets are re-ordered by decreasing value of p_T. The hemispheres are defined as follows :

- From the collections of jets ordered by decreasing value of p_T, the first jet is taken as the first axis of a first hemisphere
- Then, by looking at the other jets, if the ΔR between the jet and the first axis is below 1.5, we add the momentum-vectors of the two jets to correct on the first axis, else we build a second axis.
- Finally, we iterate over all the jets and assign them to one among both hemispheres and recompute the hemisphere momentum vector for each new jet
- If any jets is not involved in this procedure, =¿ nothing is currently done in the analysis

The goal of the analysis is to reconstruct one vertex per hemisphere.

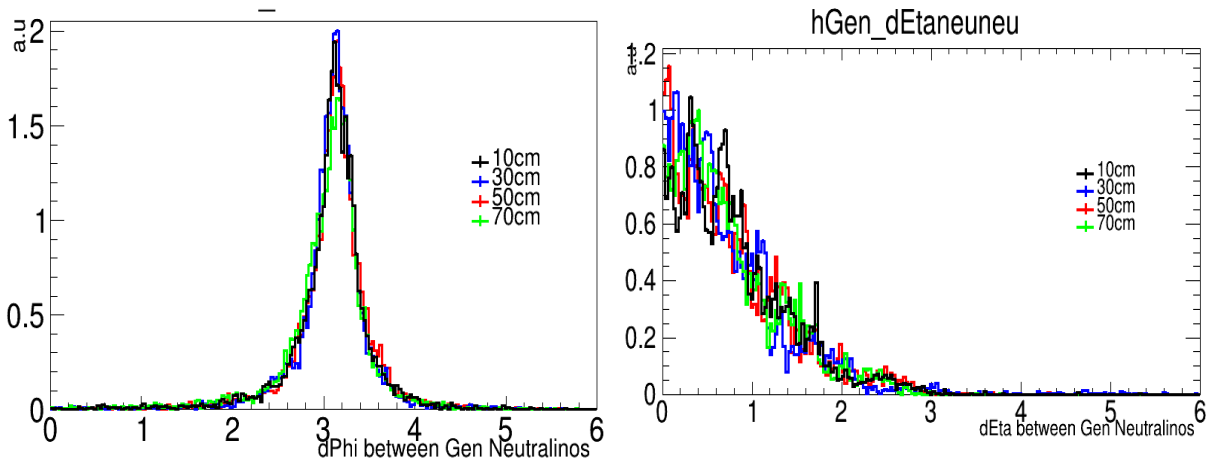

Figure 4: $\Delta\phi$ between the two generated neutralinos on the left and $|\Delta\eta|$ of the two generated neutralinos on the right

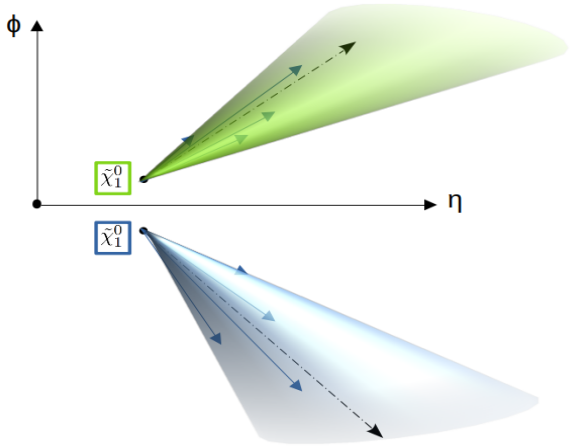

Figure 5: Mass plots of the V0 candidates from a $t\bar{t}$ sample

6.1 Quality of the reconstructed axes

The quality of the reconstructed axes can be estimated by comparing the generated axis of the neutralino and the axis of the reconstructed hemisphere associated. The axis reconstruction depends on the number of jets available, their ordering (by decreasing value of p_t) and also the ΔR criteria that is imposed for the jets to axis association (ΔR <1.5). The ΔR distribution between the reconstructed axis of the hemisphere and the generated neutralino is given in Fig. 6 as well as a function of the phase space. The plot on the left shows that about 10% of the hemispheres are not well reconstructed. This can happen for many reasons: the decay length of the neutralino is very high reaching the end of the tracker volume, hence leaving not many jets in the tracker. It can also happened that there is a mis-association of the jets between the two hemispheres if the two neutralinos are closer than $\Delta R = 3$, therefore inducing a bias in the hemisphere building procedure. The plot on the right indicates that for higher $\Delta M_{\tilde{\mu}-\tilde{\chi}_0^1} = M_{\tilde{\mu}} - M_{\tilde{\chi}_0^1}$, the hemisphere reconstruction efficiency is slightly lower, still reaching 85%, while for lower $\Delta M_{\tilde{\mu}-\tilde{\chi}_0^1}$, this

efficiency reaches 95%. In the former case where the $\Delta M_{\tilde{\mu}-\tilde{\chi}_0^1}$ is lower, jets from the decay of the neutralino are less energetic, hence not passing the selections inducing a lack of information to complete the building procedure. Finally, the top from the neutralino can be produced at large angles with respect to the neutralino, therefore, the jets from the top are still used for the building procedure but are pointing towards the top direction and not the neutralino one. All these effects can explain why the hemisphere building procedure is not reaching 100%even though the efficiency remains high.

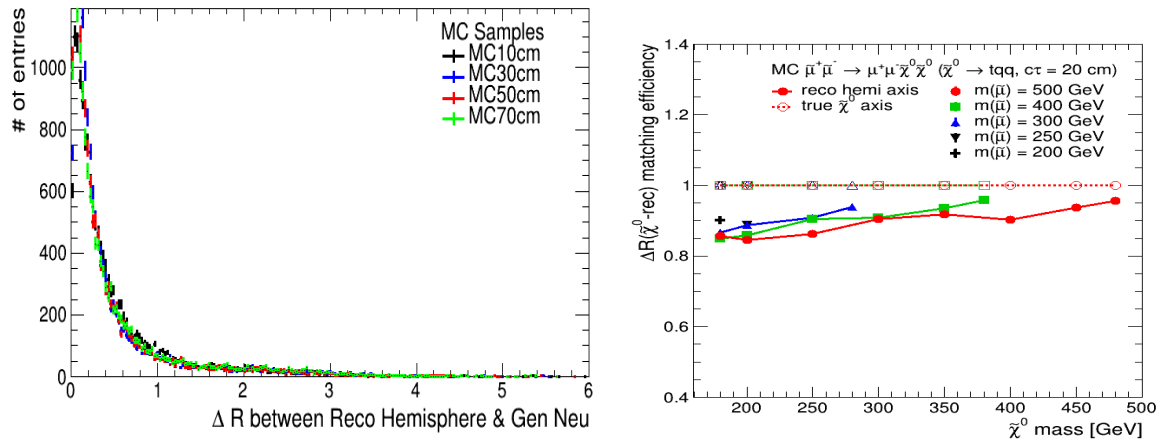

Figure 6: ΔR between the generated neutralino and the reconstructed axis

7 Selection of displaced Tracks

This analysis aims at reconstructing displaced vertices from the decay of a neutral long-lived particle. An important part of this analysis is devoted to the selection of displaced tracks coming from the decay of the neutralino. The displacement is dependant on the signal sample that we want to look for, therefore the selections are applied as independently as possible from the phase space.

The background coming from the SM is mostly prompt backgrounds with prompt tracks associated to them but secondary displaced vertices can be observed due to different sources such as V^0 candidates, photon conversions and other nuclear interactions within the material budget of the tracker of CMS.

In this analysis, we are aiming to first reduce the amount of tracks coming from these backgrounds (after the event selection mentioned in section.5) then select the displaced tracks. In the CMS Software (CMSSW), there are already collections containing the V^0 candidates and the photon conversions. These collections will not be directly used for selections but to control the selections that are applied below.

7.1 Tracks from V^0 candidates

V^0 candidates tands for two hadrons : the K^0_S meson and the Λ^0 baryon. These hadrons are produced from the hadronisation of the strange quark. They have a long lifetime (about 10^{-10} s) and produce displaced vertices with a lower track multiplicity (2-tracks vertices) compared to signal vertices (between 6 and 10 jets). Therefore, these vertices can be identified and the tracks associated to these vertices are removed from the workflow and will not be used later.

However, this is done in a specific way in this analysis. As stated above, CMSSW already gives a collection [add ref] of the V^0 candidates but there are not directly used. Using a copy of the CMSSW code to built the V^0 candidates and changing it to MiniAOD, see Appendix.13.2 and 13.3 , we reconstruct the V^0 candidates using the Kalman Filter [26] by looking at any pair of tracks of opposite signs, potentially signal tracks.
If all the conditions from CMSSW are passed and a V^0 candidate is built, the tracks associated to this candidate are removed from the analysis considering a tighter selection in mass compared to the CMSSW collection, see Fig.7.

The monitoring of this method is done by comparing the CMSSW collection of V^0 candidates and the collection that we build in the workflow, see Fig[figure ratio plots] but also by comparing these two collections between data and MC.

From the collection of reconstructed V^0 candidates (not the one of CMSSW), we get 1.1 V^0 candidates per signal and $t\bar{t}$ event.

[add the necessary figures] ...

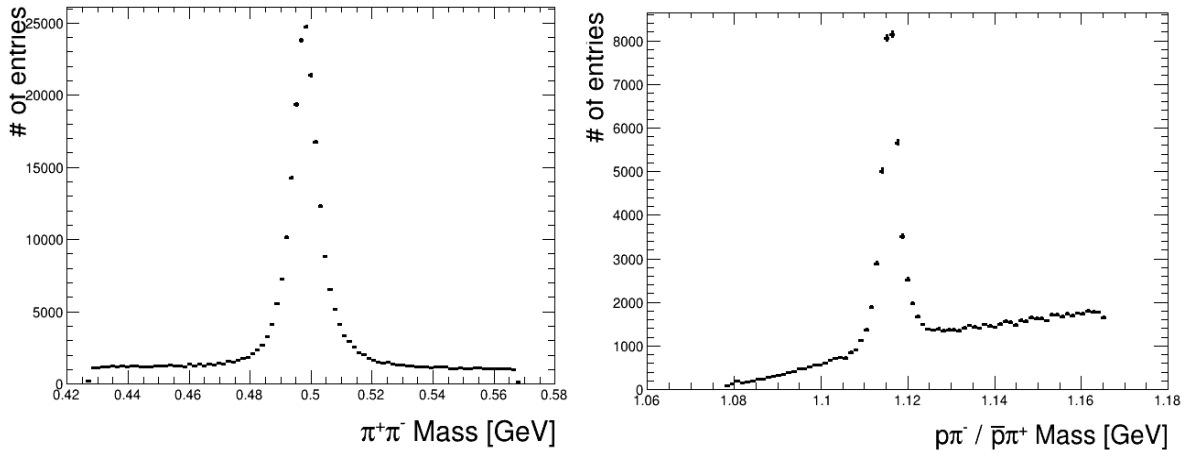

Figure 7: Mass plots of the V0 candidates from a $t\bar{t}$ sample

7.2 Tracks from Secondary Interactions

In order to remove the maximum of background tracks, the procedure of removing tracks associated to V^0 candidates vertices is extended to all kind of secondary interactions happening in the tracker volume, photon conversions and nuclear interactions with the material budget of CMS. The same algorithm is therefore used but now considering all kind of pair of tracks with selections detailed in Table[secInt sel], that are specific for the reconstruction of such interactions.

However, the veto applied on the pair of tracks belonging to secondary interactions is implemented differently. In order to remove those tracks, we match the secondary interaction vertex position with the position of the layers of the tracker. These layers can be passive (Pixel Barrel inner and outer support, Pixel Barrel rails, beam pipe) or active (Tracker Layers and Rings). The position of the layers of the tracker is defined in the RECO data format where the first hit of the tracks is used to define each layer of the tracker by their position in the tracker, with the first hit of the tracks going up to the Tracker Outer Barrel Layer 1 (TOBL1), about 70 centimeters in the transverse plane from the center of CMS. The example for the Pixel Barrel is given in Fig.8. Once the matching is done between the two positions, the tracks associated to the secondary interactions are removed from the workflow. The applied veto is shown in Fig.9 and Fig.10.

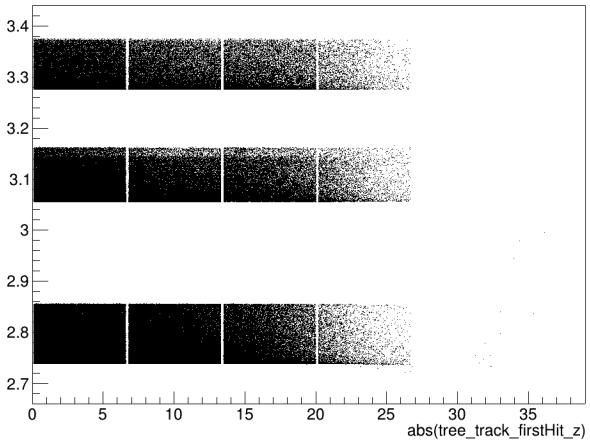

Figure 8: Example of the tracker TIBL1 geometry built using the RECO dataformat with the first hits of the tracks

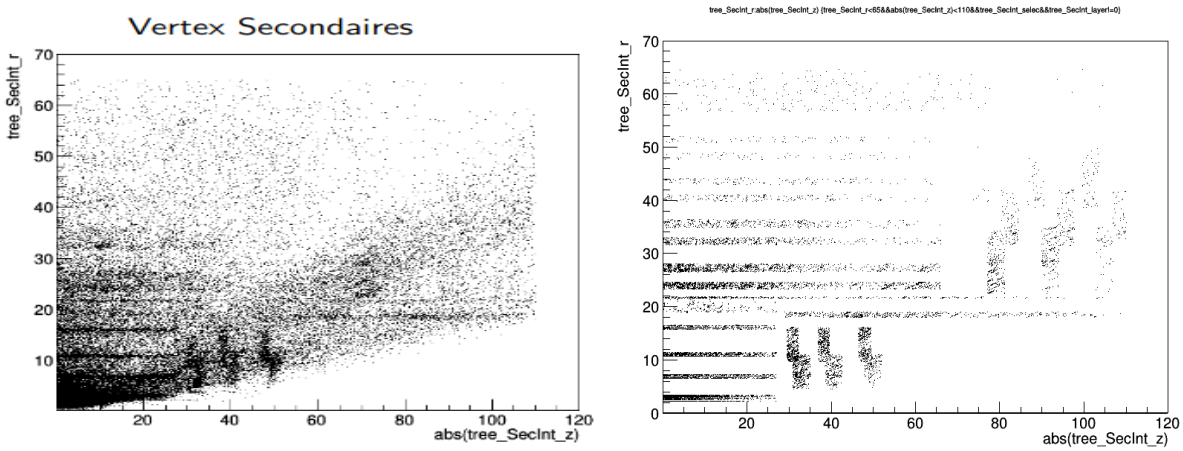

Figure 9: Mass plots of the V0 candidates from a $t\bar{t}$ sample

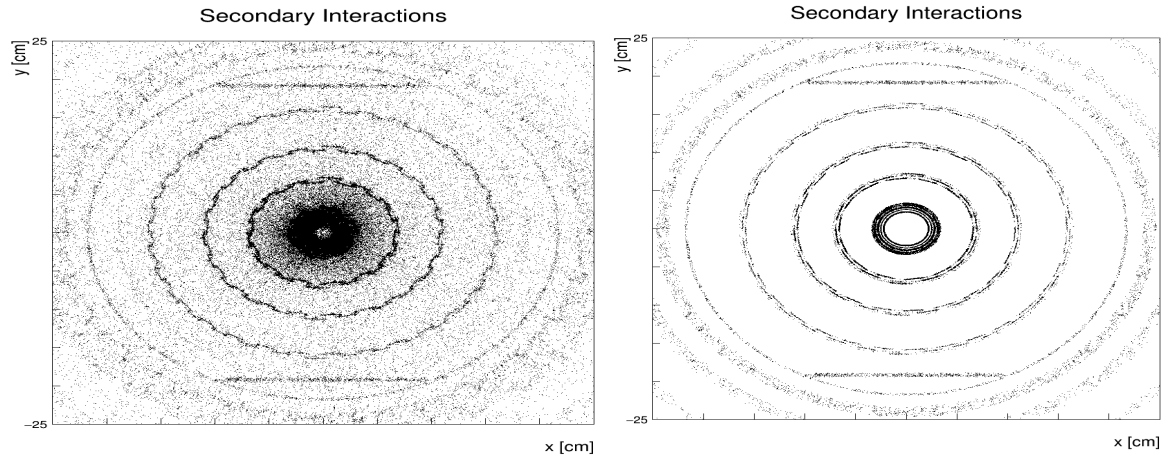

Figure 10: Mass plots of the V0 candidates from a $t\bar{t}$ sample

7.3 Tracks from signal

Since the analysis is performed in the MiniAOD data format, there are several flags that have to be taken care of. The first one is that in this analysis, both packedPFCandidate and lostTracks collections are considered, more details in [4]. The use of MiniAOD data-format also induces the use of High Purity tracks (to be described), however, it has been observed that all tracks are not High Purity tracks in the two collections. Therefore, several tests have been performed to look for the best combination of selections, see Table.12. The final requirement of the tracks are that lostTracks collection is added without requiring the HighPurity flag on. As a note, most tracks that will be selected by the track selection BDT are highPurity tracks.

Selection	Signal Track selection efficiency	Bkg track selection efficiency (in Signal Events)	$t\bar{t}$ track selection efficiency
LostTracks and without HighPurity flag on	94.5	10	6.5
With LostTracks and HighPurity flag on	93.3	8.7	5.4
Without LostTracks and without HighPurity flag on	81.5	8	4.9
Without LostTracks and with HighPurity flag on	81.4	8	4.9

Table 12: Track selection BDT input variables

Tracks from the neutralino have a displacement that depends on the decay length of the latter and it can be possibly hard to distinguish background tracks from signal tracks depending on the phase space. A first selection is optimised to reduce the amount of background tracks by requiring a $p_T > 1$ GeV, $\frac{\chi^2}{dof} < 5$ and a transverse impact parameter significance $|\frac{d_{xy}}{\sigma_{xy}}| > 5$ where σ_{xy} is the error on d_{xy}. This pre-selection allows to reduce a significant amount of background (about 90% for $t\bar{t}$) while keeping 95% of the signal displaced tracks.

[add generator level tracks plot]

However, these first cuts do not allow to reduce the background at a high enough amount to observe the signal, therefore a multivariate analysis based on a BDT is again performed to discriminate between signal tracks and background tracks. All the details related to this BDT are shown in the Appendix13.5.
The input variables are given in Table.13

Out of this BDT, one working point (WP) is defined for the rest of the workflow : Tight. The corresponding signal and background efficiencies associated to this WP are

Variables	Definition
d_z	track impact parameter in the longitudinal axis
d_{xy}	track impact parameter in the transverse plane
Sig_z	track longitudinal impact parameter significance defined as $\frac{d_z}{\sigma_z}$ where σ_z is the error on d_z
Sig_{xy}	track longitudinal impact parameter significance defined as $\frac{d_{xy}}{\sigma_{xy}}$ where σ_{xy} is the error on d_{xy}
p_T	p_T of the tracks
η	η of the tracks
$\frac{\chi^2}{dof}$	$\frac{\chi^2}{dof}$ of the tracks
n_{hits}	n_{hits} of the tracks
ntrk10-20-30-40	Number of tracks having their first hit within 10-20-30-40 centimeters from the first hit of the track considered
IsInJet	If the belong to a jet or not
isLost	If the track belongs to the lost track collection
ΔR_{min}	ΔR between the track considered and the axis of the closest hemisphere
ΔR_{max}	ΔR between the track considered and the axis of the most distant hemisphere

Table 13: Track selection BDT input variables

given in the Table.14. The tight WP is defined to reduced the background tracks by a factor 10^3.

WP	BDT value	Signal Efficiency (%)	Background Rejection (%)
Tight	0.85	73	99.5

Table 14: Working Point for the track selection BDT

7.3.1 Dependence of the track selection BDT with respect to the phase space

[tracks assignment ot hemisphere]

8 Displaced Vertices Reconstruction

The main goal of this analysis is the reconstruction of displaced vertices coming from the decay of a long-lived neutralino decaying in the tracker volume. For the vertexing, only one vertex per hemisphere is reconstructed using the Adaptive Vertex Fitter (AVF) [35]. The AVF is a more robust iteration of the Kalman Fitter, more efficient with high-track multiplicity vertices allowing to reach good reconstruction efficiencies for our signal. A comparison between the two algorithms is shown in Fig.[add figure]

The parameters given as an input of the AVF are given in the Table.15. After testing different values for the input parameters, no major changes in the final efficiencies were observed therefore the default values given in the Table.15 are kept.

Parameter	Value	Definition
maxshift	0.0001	Convergence criterion (maximum transverse distance between vertex computed in the previous and the current iterations)
maxstep	30	Maximum number of iterations to perform
maxlpshift	0.1	Criterion for the relinearization of the tracks
weightthreshold	0.001	Minimum track weight for a track to be considered "significant". If fewer than two tracks are significant, an invalid vertex is returned.
sigmacut	5.	Tracks this number of sigma from the vertex are given a weight of 0.5
Tini	256.	Parameter used to compute the weight of a given track associated to a vertex
ratio	0.25	Parameter used to compute the weight of a given track associated to a vertex

Table 15: Input parameters for the Adaptive Vertex Fitter

A first collection of tracks is built for each hemisphere out of the tight WP. The vertexing will then be performed in 2 steps for each hemisphere in order to reconstruct one vertex per hemisphere. A detail of the vertexing workflow is given below.

8.1 Multi-step vertexing

In this analysis, one vertex is reconstructed per hemisphere knowing the complexity of the sub-structure of the displaced jets. Further improvement could be potentially obtained by looking for tertiary vertices coming from the b-jets but these jets can be hard to identify since b-tagging does not apply to largely displaced jets (dozens of centimeters) and also due to the decrease in resolution of the vertices.

For each hemisphere, we first use the collection of the tight WP tracks to build the vertices where we expect to have a collection enriched in signal displaced tracks. Then,

we apply two different steps to try to build a vertex in a hemisphere. The criteria for the reconstruction of a vertex are :

- a normalized $\frac{\chi^2}{dof}$ between 0 and 10
- for signal samples, the relative distance between the reconstructed and generated vertices must be lower than 10% of the generated decay length.

1. The first iteration is the direct production of a vertex out of the tracks from the tight WP
2. if no vertex is produced in the previous step, an iterative implementation of the AVF (IAVF, see details in Appendix.13.7) is applied by requesting a $\frac{\chi^2}{dof}$ while building the vertex

In this multi-step vertexing, some tracks are removed if they do not respect the criteria that the first hit of any track must come after the position of the vertex that is built. Then, the step 2 corresponds to the IAVF, each iteration corresponding to adding a new track in the procedure, if the next iteration provides a vertex with a $\frac{\chi^2}{dof}$ below 0 or above 10, the track is removed. Out of these removed tracks, there is a possibility that these tracks belong to tertiary displaced vertices coming from b-jets. However, applying the AVF on the removed tracks has not shown any improvement in the vertex reconstruction efficiency.

8.2 Vertex Selection

The two steps mentioned above allow to maximise the reconstruction of signal vertices and minimising the reconstruction of background vertices, see Fig.XX [add figure of efficiency vs step]. However, the amount of backgrounds events at this stage is still orders of magnitude above the number of signal events, so a new BDT-based multivariate analysis is performed to discriminate signal vertices from background vertices. This can be done thanks to the high-track multiplicity of the final state and the choice of the multi-step vertexing. All the details about the BDT are described in the Appendix.13.8

The input variables are listed in Table 16.

Variables	Definition
Mass of the Hemispere at construction level	Invariant mass of all the jets and letpons (except the prompt ones) contained in the hemisphere
Mass of the Hemisphere at the vertex level	Invariant mass of all the tracks associated with the vertex
ntrk10	Number of tracks having their first hit closer than 10 cm from the vertex
ntrk20	Number of tracks having their first hit closer than 20 cm from the vertex
Mean of Distance of Closest Approach	Computation of the mean of the distance of closest approach of the tracks associated to the vertex and the vertex
N. of tracks	Number of tracks associated with the vertex
$\frac{\chi^2}{dof}$	χ^2 per degree of freedom of the vertex
Step	The step of reconstruction of the vertex (1 or 2)
Mean Weight of the tracks	Each track is associated a weight corresponding to the probability that this track belongs to the vertex. The mean value of the weights of all the tracks associated to the vertex is computed. A discussion about this variable is given in Appendix.13.8.1.

Table 16: Vertex selection BDT input variables

8.2.1 Dependence of the Vertex selection BDT with respect to the phase space

Since this BDT is based on track-multiplicity quantities, for decay length reaching the outer part of the tracker where the tracking efficiency is at its lowest, the track-multiplicity of the signal vertex can decrease and reach the track-multiplicity of a SM background vertex. Therefore, a drop of efficiency can be expected for high decay length of the neutralino for this BDT. A comparison between two BDTs is shown in Fig.[Add fig].

8.2.2 Vertex reconstruction efficiency

The reconstruction of the signal is expressed as the vertex reconstruction efficiency as a function of the decay length of the neutralino. This efficiency can also be expressed as a function of the transverse distance or $|\eta|$. In this analysis, a vertex is considered reconstructed when the $\frac{\chi^2}{dof}$ is between 0 and 10. For signal MC samples, a criteria is added on the matching between the reconstructed vertex and the generated one: the relative distance between the two vertices must be lower than 10% of the generated decay length (called matched vertices).

- The efficiency is computed as the ratio between the number of matched vertices by the number of vertices that should be reconstructed in theory, i.e twice the amount of events passing the Online+Offline selection.

- The second quantity that can be defined in signal MC samples is the purity, that is the ratio between the number of matched vertices and the number of vertices having a $\frac{\chi^2}{dof}$ between 0 and 10.

Both quantities are shown in Fig.11

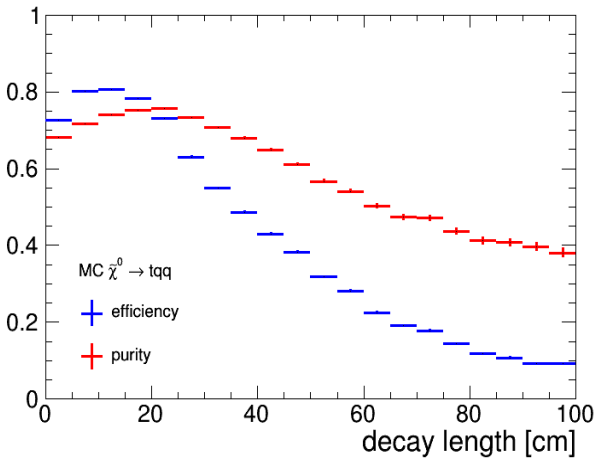

Figure 11: Vertex reconstruction efficiency and purity as a function of the decay length of the neutralino.

The vertex reconstruction efficiency is highly dependant on the tracking efficiency causing the negative slopes after a decay length of 10 cm. The purity is also dependant on the resolution on the vertices for high decay length.

Finally, the vertex reconstruction efficiency is shown as a function of the smuon and neutralino masses in Fig.12 for a given $c\tau$.

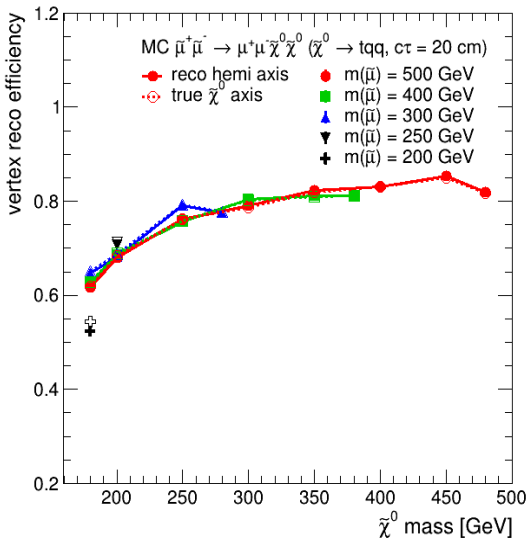

Figure 12: Vertex reconstruction efficiency as a function of the mass of the neutralino and the smuon for a $c\tau$ of 20 cm. The comparison is made between the reconstructed hemisphere and the true neutralinos axis used as the reconstructed axis. No major deviation is observed showing the good behavior of the multi-step vertexing with respect to the workflow.

8.2.3 Resolution on the reconstructed vertices

The resolution on the position of the vertices is an indicator of the quality of the reconstructed vertices. Since the reconstructed vertices are displaced, the resolution is shown as a function of the decay length in the transverse plane in the barrel region and on the longitudinal axis in the endcaps. The resolution is computed as the distance between a vertex of signal and a vertex of secondary interaction and shown in Fig.13

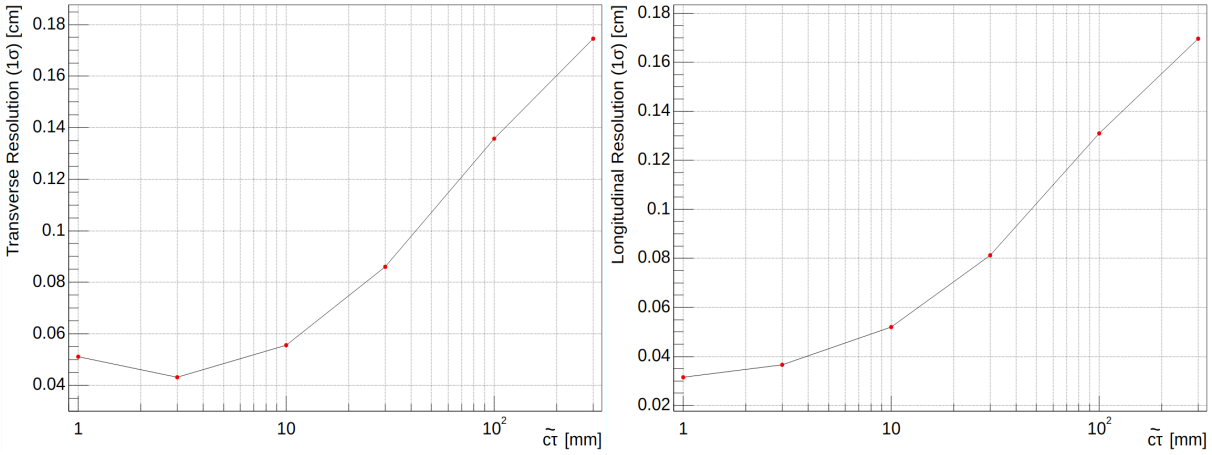

Figure 13: Transverse (longitudinal) resolution in the transverse plane (longitudinal axis) as a function of the decay length

9 Background Estimation

9.1 Validations using MC samples

9.2 Validations using Data samples

9.3 ABCD Method

The background estimation is implemented using two discriminating variables to define background enriched regions (control regions) and signal-enriched regions (signal regions). The current implementation of the analysis uses the event selection and vertex selection BDTs in order to define these regions. The final categorisation is shown for exactly one vertex and two-vertices reconstructed in Fig.14

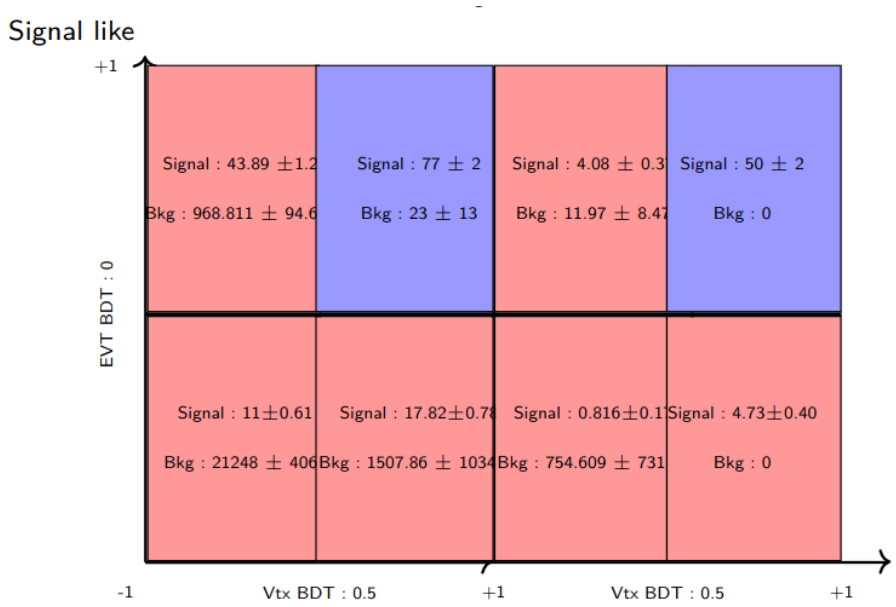

Figure 14: Event yields for Run 2 for the one vertex and two vertices categories. The two categories are defined as : "Exactly one vertex reconstructed" and "Exactly two vertices reconstructed" where a vertex is considered reconstructed when its follows $0 < \frac{\chi^2}{DoF} < 10$. For signal samples, the reconstructed vertex is required to be within 10% matching with the generated vertex such that : the difference in position between the reconstructed and generated vertices is below 10% of the generated vertices decay length. The signal region is defined by a BDT value above 0 for the event selection BDT and above 0.5 for the vertex selection BDT.

10 Systematic Uncertainties

This analysis relies on the study of simulated Monte-Carlo samples. Given that all physic quantities are not perfectly modelled and simulated, systematic effects arise from the study of the Monte-Carlo samples. These effects can come from the trigger efficiencies, pile-up, detector effects, particle identification and potentially changing the shape of the different distributions of this analysis.

The systematic parameters mentioned in the following sections are varied both positively and negatively of one sigma from the nominal value. These variations will be referenced as up and down, respectively. The final analysis of these systematic effect is shown in Table.[add table from combine].

10.1 PDF uncertainties

10.2 Luminosity Measurement

10.3 Trigger systematics

10.4 Pileup uncertainties

10.5 Jet energy scale

10.6 Track parameters : p_T, χ^2, IP_{xy}

10.7 Event selection BDT mis-identification

10.8 Track selection BDT mis-identification

10.9 Vertex selection BDT mis-identification

10.10 Monte-Carlo statistics

11 Results

col1	col2	col3	col4
Multiple row	cell2	cell3	cell4
	cell5	cell6 and cell7	
	cell8	cell9	cell10
cell11	cell12	cell13	cell14
Multicolumn			

Table 17: A table is as happy about a caption as a fiugre.

12 Summary

In this paper, a search for displaced top quark through the decay of a massive long-lived particle in the tracker of CMS for the Run 2 of the LHC is introduced. This search is mainly based on machine learning to select displaced tracks in order to reconstruct signal displaced vertices reaching a signal vertex reconstruction efficiency of about 50% at a decay length of 50 cm for the neutralino.

13 Appendix

13.1 Monte-Carlo generation of signal samples

The generation of the RPV-Process is made using the MADGRAPH5_aMC@NLO (MG5) generator. Since there is no explicit Universal FeynRules Output (UFO, see [16]) model dedicated to the RPV-process studied, a private model has been developed taking into account the constraints from MG5 such as the non-possibility of the lightest neutralino (LSP) to decay. In the following, the lightest neutralino at the MG5 level will be χ_0^2 since MG5 allows for its decay. To proceed to the generation, two steps are required. First, the merging between Pythia8 [add ref to pythia 8] and MG5 for the hadronisation part then the generation of events with the needed input parameters. For the first part, the value of merging is given the MC& I contact. The generation is detailed below.

13.1.1 Generation

The event generation of the pair-production of neutralino is done at LO+1 jet as an initial state radiation (ISR). The commands given are used to generate the signal : import model DisplacedTopUDD –modelname
generate p p > sl2+ sl2- / h01 h02 a0 n1 n2 n3 n4, (sl2+ > n2 mu+, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6), (sl 2- > n2 mu-, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6) @1
add process p p > sl2+ sl2- / h01 h02 a0 n1 n2 n3 n4, (sl2+ > n2 mu+, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6), (sl2- > n2 mu-, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd 5 sd5 sd6 sd6) @2
add process p p > sl2+ sl2- / h01 h02 a0 n1 n2 n3 n4, (sl2+ > n2 mu+, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6), (sl2- > n2 mu-, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd 5 sd5 sd6 sd6) @3
add process p p > sl2+ sl2- / h01 h02 a0 n1 n2 n3 n4, (sl2+ > n2 mu+, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6), (sl2- > n2 mu-, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6) @4
add process p p > sl2+ sl2- j / h01 h02 a0 n1 n2 n3 n4, (sl2+ > n2 mu+, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6), (sl 2- > n2 mu-, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6) @1
add process p p > sl2+ sl2- j / h01 h02 a0 n1 n2 n3 n4, (sl2+ > n2 mu+, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6), (sl2- > n2 mu-, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd 5 sd5 sd6 sd6) @2
add process p p > sl2+ sl2- j / h01 h02 a0 n1 n2 n3 n4, (sl2+ > n2 mu+, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6), (sl2- > n2 mu-, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd 5 sd5 sd6 sd6) @3
add process p p > sl2+ sl2- j / h01 h02 a0 n1 n2 n3 n4, (sl2+ > n2 mu+, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6), (sl2- > n2 mu-, n2 > t d s / sd1 sd1 sd2 sd2 sd3 sd3 sd4 sd4 sd5 sd5 sd6 sd6) @4

Then, the parameters used for generating the LHE file are the following:

Parameter	Value	Definition
shower program	Pythia8	Hadronisation
pdflabel	lhapdf	PDF set
lhaid	306000	lhapdf number (NNPDF3.1)
ickkw	1	MLM
asrwgtflavor	5	highest quark flavor for α_S reweight
xqcut	20	minimum kt jet measure between partons

Table 18: List of tuning parameters for Pythia8.

The merging value (xqcut) between MG5 and Pythia8 (better xplain this) is the one provided by the Monte-Carlo contact : between 20 and 30. This value was also retrieved by generating signal samples for different values of xqcut, from 1 to 100. The systematic uncertainties at the generator level are extracted from the **NNPDF3.1** pdf ... .

13.2 Covariance Matrix

This analysis used to be performed on the RECO data-format but it was moved the MiniAOD data-format following the request of the CMS group and to reduce computing time/file size. However, this change induces several changes such as the approximation of the covariance matrix of the tracks. This approximation may not have direct consequences for the track selection but a drop of 10% of the total vertex reconstruction efficiency was observed due to the MiniAOD data-format. The covariance matrix of the track is gathering :

1. $\frac{q}{abs(p)}$: signed inverse of momentum [1/GeV]
2. $\lambda = \frac{\pi}{2}$: polar angle at the given point
3. ϕ : azimuth angle at the given point
4. d_{xy} : -v_x*sin(ϕ) + v_y*cos(ϕ) [cm]
5. d_{sz} : v_z*cos(λ) - (v_x*cos(ϕ)+v_y*sin(ϕ))*sin(λ) [cm]

In the MiniAOD format, due to the approximation of this matrix, the matrix is not positive-definite with matrix element being equal to 0 or infinite. A simple correction is applied to correct for these non-defined values and make the matrix positive-definite. The correction was developed by the B Physics group [2] in the release CMSSW_13_0_0, but it can also be added by hand in the analysis framework. The correction allows to recover for the 10% vertex reconstruction efficiency drop observed.

13.3 First hit retrieval

This analysis used to be performed on the MiniAOD data-format following the request of the CMS group to reduce computing time/ file size. However, this change induces several changes such as the loss of the position of the first-hit of the tracks. This loss can be characterized in two ways : one for the packedPFCandidate and one for the lostTracks collection. In MiniAOD data-format, the only information about the position of the first-hit of the tracks is the layer of the first-hit that is registered as an 11-bit number and that can be interpreted as a 4-digits number also called hitpattern (see [add hitpattern ref] or table). Concerning the lostTracks collection, the 4-digits number is not registered correctly leading to no first-hit information at all. A correction is applied on Run 3 for 2023 and beyond. Concerning the packedPFCandidate collection, this 4-digit number is stored correctly and can be used to retrieve the first-hit position of the tracks using track parameters and propagators. [add plots for layers and resolution plots]. The needed track parameters are the following : momentum , η, ϕ, the distance of closest approach of the center of CMS (reference point) along the longitudinal axis, the state of the track at the reference point defined as a 6-dimensional state matrix with the associated error matrix, and the hitpattern mentioned above. The 6-dimensional state matrix parameters are the position (defined in the global frame), the momentum, the charge, the signed inverse momentum (charge divided by the magnitude of the momentum) and the transverse curvature (Transverse curvature kappa (which is the inverse radius of curvature in the transverse plane) in cm^{-1}. Sign convention is such that positive kappa means counter-clockwise rotation of the track with respect to the global z-axis.).

With this set of parameters with their associated errors, propagators are then used to retrieve the position of the first-hit. There are two different propagators used for this procedure, depending on the region of propagation : barrel or endcap. For the endcap region, the StraightLinePlaneCrossing propagator is used and the analytical propagator is used for the barrel region. If the propagators fail, a geometric estimation of the position of the first-hit of the track is done using η, ϕ and the distance of closest approach of the center of CMS (reference point) along the longitudinal axis. The output of the propagators is shown in Fig. for TIB L1. The resolution on the first-hit of the tracks is given in Table.[add table]

13.4 Event Selection BDT

After the selection on the events using the triggers and the invariant mass of the two prompt muons in Section.5, there are still many variables (Table.11) that display discriminating power between signal and background at event level. These variables are mainly based on the kinematics of the objects of the events, hence their distributions vary with respect to the phase space [add plots]. Therefore, a simple cut-based selection based on these variables would not allow for a global selection of the signal but it would select a certain part of the signal phase space. In order to select the maximum of signal events in the whole phase space allowed, a BDT is implemented thanks to the Toolkit For Multivariate Data Analysis (TMVA) available in ROOT.

For the selection of events, the training has been performed using all signal samples (re-weighted by their cross-section) as a single signal input to the BDT. A total of 635 000 signal events (-6000 for each sample of signal, the same for the test) are used for the training part. Concerning the backgrounds, about 400 000 background events are used for the training (same amount for the test) with the background being listed in Table.4.

The distributions for these variables are given in Fig.[XX] and [XY] with the associated correlations matrices in Fig.[YY]. [don't forget to talk about the dxyerror data/mc disagreement that is expected and could modify the output of this BDT]. Since no difference of efficiency is observed between training and test, no over-training is obtained but it is expected that this BDT is trained for specific regions of the phase space and therefore "over-trained" for these regions. The plot on the right of Fig.18 shows a good discrimination power that can be used to select signal events with one working point.

[I, Paul, son of Brittany, will update the plots when the BDT's will be more "finalized", thank you!]

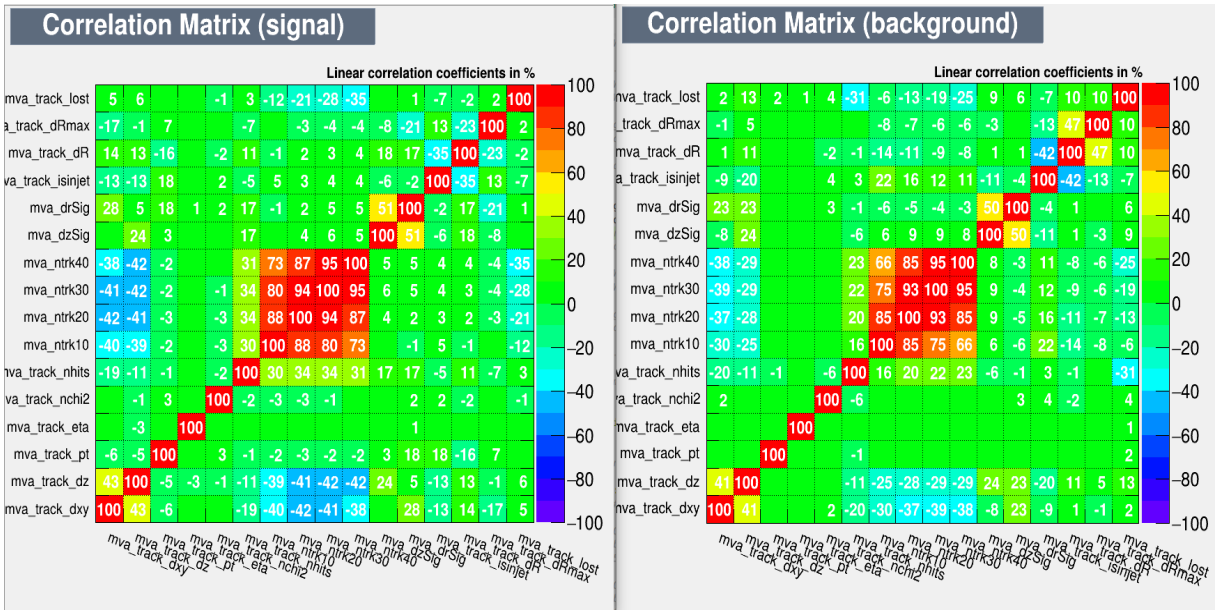

Figure 15: Correlation matrices between input variables for the track selection BDT for signal on the left and background on the right

.

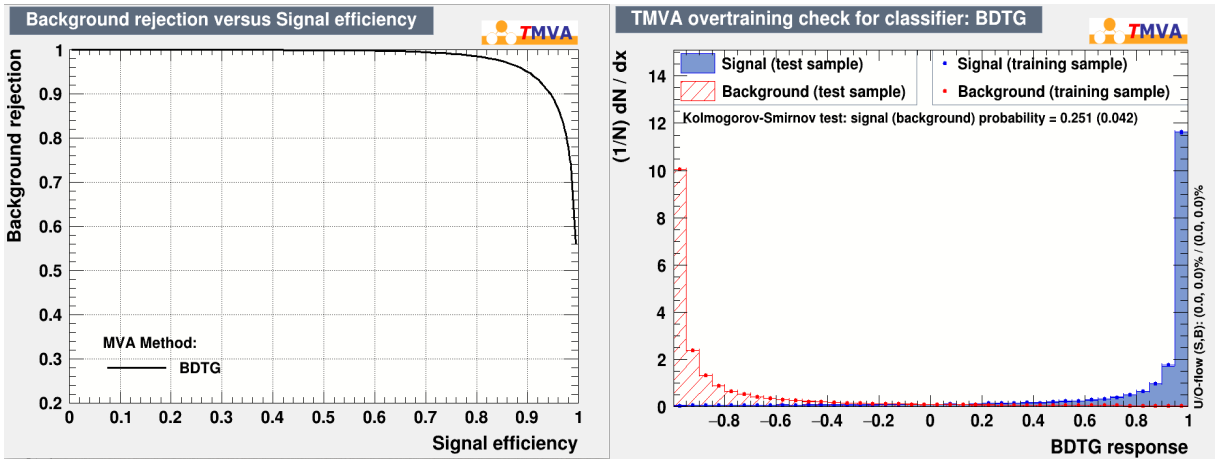

Figure 16: On the left, ROC Curve showing the background rejection with respect to signal efficiency. On the right, the test (histograms) and training (dots) output distributions for signal and background

.

13.5 Track Selection BDT

The final selection of displaced tracks for the reconstruction of vertices is performed using the Toolkit For Multivariate Data Analysis (TMVA) available in ROOT. This toolkit allows for many implementations of multivariate data analysis techniques, from the BDT to Convolution Neural Networks. In this analysis, we make use of the BDT to distinguish signal displaced tracks with respect to prompt tracks from SM backgrounds. The BDT implementation needs three steps : training, testing and "verification". The first two are performed by root while the third one has to be performed by the analyzers to monitor the theoretical and observed outputs.
For the selection of tracks, the training has been performed using one signal sample (to be extended to more signal samples in the future) of parameters $M_{\tilde{\mu}} = 275$ GeV, $M_{\tilde{\chi}_0^1} = 225$ GeV and $\beta\tilde{\gamma}c\tau = 50$ cm for the neutralino using 40 000 tracks for the training (40 000 for testing). For the backgrounds, the training has been performed with the samples mentioned in Table.4. A total of 200 000 tracks have been used for the background samples for the training (200 000 for testing). For the training, each sample has been re-weighted by their cross-section using the XSDB "database" [5] for the backgrounds and the Fig.2 for the signal sample.
The input variables for this BDT are the one given in Table.13 with the distribution for the signal and background given in Fig. XX [Add figure input variables for TRKBDT]. The correlations between these variables are given in Fig.17. The results of the training are provided in Fig.18. Since no difference of efficiency is observed between training and testing, no over-training is obtained but it is expected that this BDT is trained for specific regions of the phase space and therefore "over-trained" for these regions.
The plot on the right of Fig.18 shows a good discrimination power that can be used to select displaced tracks. Therefore, one working point is defined for this BDT, a tight one (TWP). The TWP is defined such that the rejection power for the background is about 10^3 (this factor depends on the background sample). This factor is obtained for a value of 0.85 for the BDT value giving a signal efficiency of 73% and background rejection of 99.5%. This working point will be used to build one collection of tight tracks (TT) that will be then given as an input for the vertexing workflow.

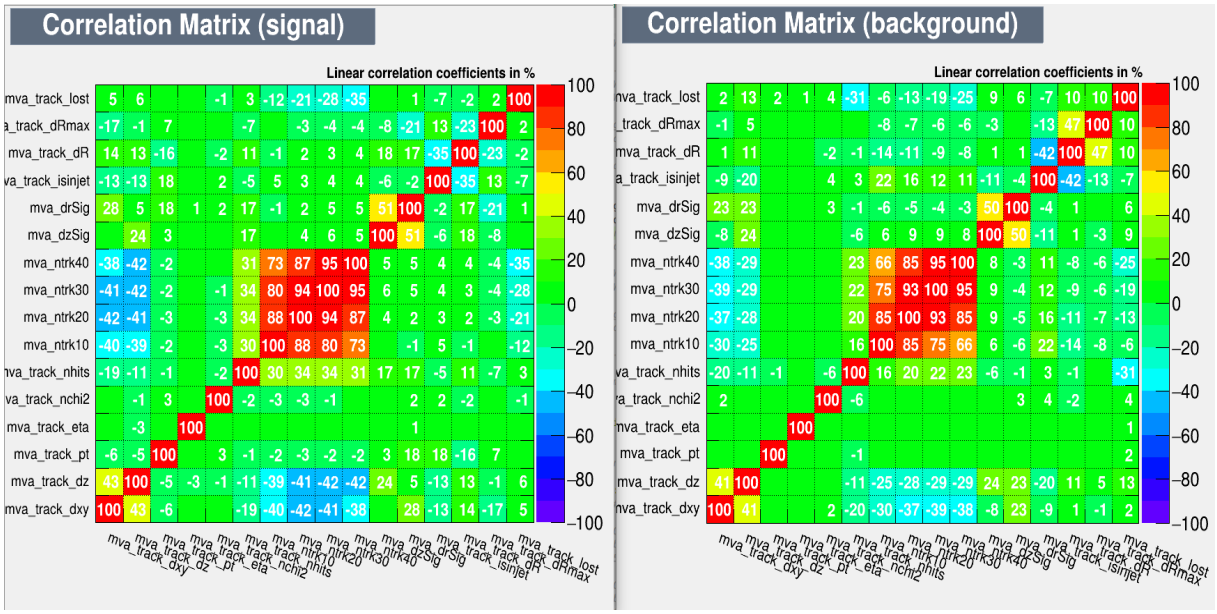

Figure 17: Correlation matrices between input variables for the track selection BDT for signal on the left and background on the right

.

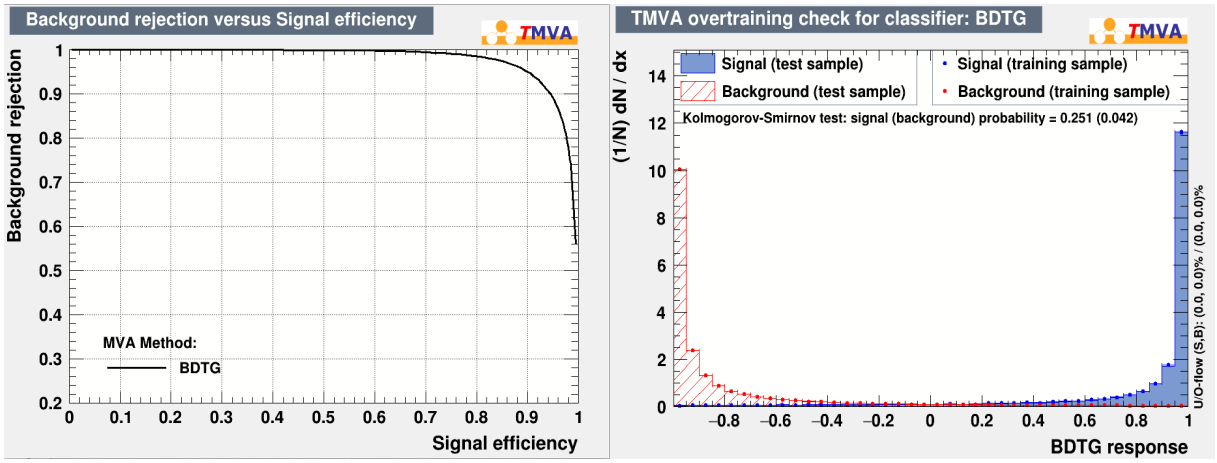

Figure 18: On the left, ROC Curve showing the background rejection with respect to signal efficiency. On the right, the test (histograms) and training (dots) output distributions for signal and background

.

13.6 Adaptive Vertex Fitter

The goal of the analysis is to reconstruct displaced vertices coming from the decay of a long-lived neutralino in each of the two hemispheres. The AVF is used to reconstruct the vertices in order to increase the robustness with respect to track-multiplicity and displacement compared to the basic Kalman Fitter. The first step of the vertexing makes use of the AVF to reconstruct vertices out of the TT collection.
The AVF uses all the tracks from the TT collection but each track is weighted according to their chi2 distance to the vertex. The weight is computed as followed :

$$w_i(\chi_i^2) = \frac{\exp^{-\chi_i^2/2T_{ini}}}{\exp^{-\chi_i^2/2T_{ini}} + \exp^{-\chi_{cut}^2/2T_{ini}}} \tag{2}$$

where χ_i^2 is a χ^2-type criterion for statistical compatibility between the vertex position and track i. The temperature parameter T_{ini} introduced before defines the shape of the function. Then, the AVF uses a deterministic annealing schema to avoid local minimum where T_{ini} is decreased by a factor defined by the parameter *ratio*, so that the temperature converges to 1. The weight is therefore defined between 0 and 1 and can be interpreted as the probability for a track to belong to the associated vertex.

A vertex is built if :

1. the AVF converges before the maximum of iteration is reached
2. the fitted position of the vertex is within the tracker bounds
3. there are at-least two tracks with a "high-enough" weight (above 0.8).

Therefore, one has to check the weight of the tracks when looking at the number of tracks associated to a vertex. A deeper analysis of the weight of the tracks is provided in Appendix.13.8.1

13.7 Iterative Adaptive Vertex Fitter

The goal of the analysis is to reconstruct displaced vertices coming from the decay of a long-lived neutralino in each of the two hemispheres. The AVF is used to reconstruct the vertices in order to increase the robustness with respect to track-multiplicity and displacement compared to the basic Kalman Fitter. The first step of the vertexing makes use of the AVF to reconstruct vertices out of the TT collection. However, when comparing the number of vertices reconstructed in this first step with the number of vertices that can be reconstructed, the first step reaches about 80% of the total vertices to be reconstructed. In the 20% of the vertices remaining,

1. either there are some missing tracks as those tracks may not be part of the TT
2. or the first step of the vertexing procedure failed with the TT collection.

The first issue is not addressedsince it would require to lower the working point of the BDT of selections of tracks and largely increase the amount of background tracks. The second issue is addressed by implementing an Iterative Adaptive Vertex Fitter (IAVF). The Adaptive Vertex Fitter is already an iterative procedure by itself but it can be changed in order to facilitate the building procedure of a vertex. In the first step of the vertexing, the AVF is simply used by giving the TT collection as an input and the output is either a vertex with a good $\frac{\chi^2}{dof}$, a vertex with a bad $\frac{\chi^2}{dof}$ or no vertex at all. A vertex is considered good when the $\frac{\chi^2}{dof}$ is between 0 and 10, see Fig.[chi2 distribution of vertices], where the dof is defined as the sum of the sum of the tracks weights minus 3. For the step 2 using the IAVF, the TT are ordered by decreasing value of BDT score of the track selection BDT, making the most signal-like tracks to be selected first. The IAVF can be defined as follow :

1. a first good vertex ($0 < \frac{\chi^2}{dof} < 10$) is built out of any pair of tracks from the TT collection. This vertex is called the seed. This step makes use of the ordering of the tracks by decreasing value of BDT score. If the first two tracks are not used to build the seed, it iterates over all the remaining tracks but no track is discarded from the vertexing during this procedure. If no seed is found, no vertex is built.
2. When the seed is built, we add one track after the other to the seed by making sure that after each new track, the newly fitted vertex has a good $\frac{\chi^2}{dof}$. If this criteria is satisfied, it iterates over the remaining tracks. Else, the track that makes the vertex not satisfying the $\frac{\chi^2}{dof}$ criteria is removed from the vertexing.
3. The final possible outputs are the same as step 1, where a good vertex can be obtained, a bad vertex or no vertex at all if no seed is built
4. Note : if a track has a before the vertex position, the track is discarded. There is an exception for tracks from the lostTracks collection where the information of the first-hit is not stored correctly for Run 2 Monte-Carlo, see Appendix.13.3

The step 2 is adding between 5 and 10% to the total efficiency depending on the decay length of the neutralino, see Table [add table total efficiency per step]

13.8 Vertex Selection BDT

The goal of the analysis is the reconstruction of displaced vertices. The use of the multi-step workflow detailed in Appendix.13.6 and 13.7 aims at reconstructing the majority of the vertices. However, even after the reconstruction of vertices using the TT collection of tracks, the level of rejection of background is still not high enough to consider observing the signal, further selections need to be applied. To do so, a BDT is implemented to select signal vertices from background vertices. There are three main aspects that discriminates signal from background vertices : the track-multiplicity associated to the vertex, the invariant mass of the vertex and the step of reconstruction of the vertex.

1. Signal vertices have more tracks associated to them compared to background vertices thanks to the coupling of the virtual $\tilde{t}$ to a down and a strange quark.

2. One can reconstruct the mass of the displaced vertex using the jets used to build the hemisphere to have the neutral component or use only the charged components using the tracks used to build the vertex

From these observations, several variables are built to discriminate signal from background vertices as shown in Table.16. The input variables distributions are given in Fig. [add fig input variables distribution]. The correlations between these variables are given in Fig.19. The training has been performed using all the signal samples generated (used as a single signal sample) as an input to the BDT signal sample using $\sim$6 000 signal vertices per sample for the training (and 6000 for testing). For the backgrounds, the training has been performed with the samples mentioned in Table.4. For the training, each sample has been re-weighted by their cross-section using the XSDB "database" [5] for the backgrounds and using the Fig.2 for the signal sample.
The results for the training are shown in Fig.20. The plot on the right of Fig.20 shows a good discrimination power that can be used to select displaced vertices. Since no difference of efficiency is observed between training and testing, no over-training is obtained but it is expected that this BDT is trained for a specific region of the phase space and therefore "over-trained" for this region of the phase space. By looking at the efficiency of this BDT with respect to the phase space, ... add [plot BDT output vs phase space]...
There is one working point that is defined such that [to be defined after all the tests] in Fig. 20 (left plot). This working point is used to define the final category of the event selection.

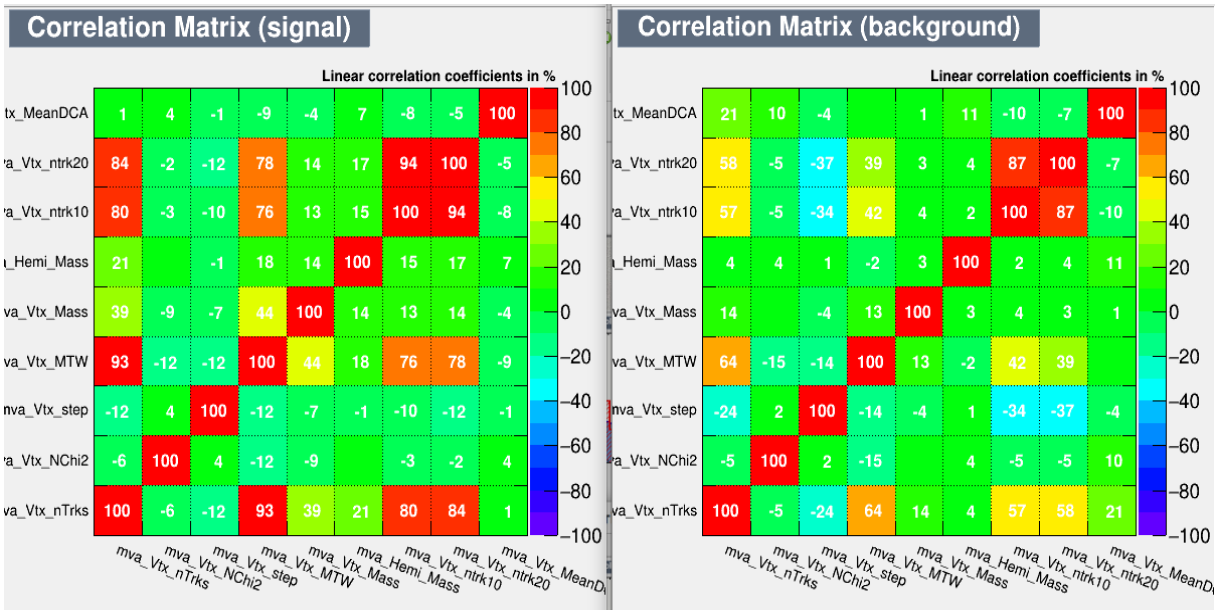

Figure 19: Correlation matrices between input variables for the track selection BDT for signal on the left and background on the right

.

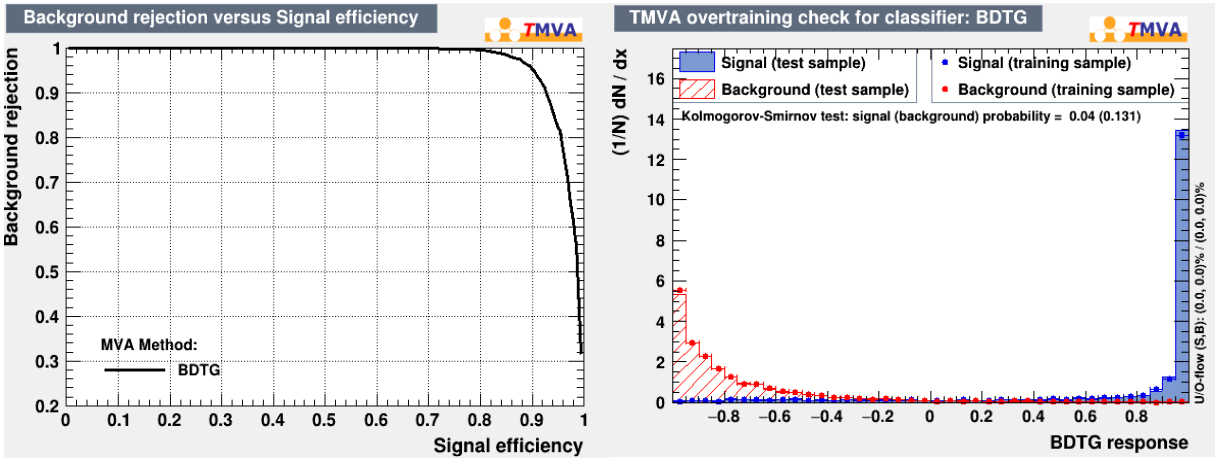

Figure 20: On the left, ROC Curve showing the background rejection with respect to signal efficiency. On the right, the test (histograms) and training (dots) output distributions for signal and background

.

13.8.1 Mean Weight of the Tracks associated to a vertex

During the multi-step vertexing procedure, tracks are re-fitted in order to build a vertex. If a vertex is effectively built, the associated tracks are assigned a weight representing the probability of the track to effectively belong to the vertex, details in [35]. Therefore, the weight is defined between 0 and 1 for each track where the weights are mostly close to 0 or 1.
A this state of the analysis, the level of rejection of background is not high enough to observe the signal [see Table or plot, selection efficiency vs step]. A vertex selection step is then needed to reach higher level of rejection of background.
Vertex level variables are then needed to select signal displaced vertices from SM background vertices. One vertex level variable can be built using the mean value of the weight of the tracks associated to a vertex. This variable can be interpreted as the effective number of tracks that belong to a vertex since the weight of a track is either close to 0 or to 1. Hence, for a vertex with N tracks associated, there are $N * \sum w$ effective tracks with w being the weight of a track. The difference between the two quantities is important : the first is a integer that only depicts the track used to build a vertex (directly linked to the track-multiplicity of the final-state), the second is a float that as mentioned above, is the effective number of tracks that is also linked to the track-multiplicity of the final-state but also describes the quality of the association of the tracks. To illustrate the last point, see Fig. 22. The best case scenario is obtained when $N = \lceil N * \sum w \rceil$ meaning all the BDT selected tracks are well associated to a vertex, plus the higher the value of N, the better since the track-multiplicity is expected to be higher in the signal process.
This newly-formed vertex level variable that is the mean weight of the tracks associated to a vertex can then be used as a discriminating variable between signal and background vertices as it carries multiple information : track multiplicity and the quality of the reconstructed vertex.

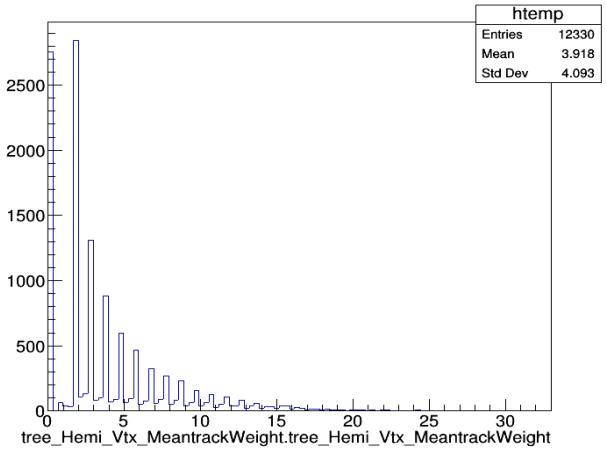

Figure 21: Example of the distribution of MWT. The peaks are observed around integers where a large fraction of these peaks is due to $N = \lceil N_{\sum w} \rceil$. But with N tracks, N-M tracks can have a high weight and M tracks have a low weight. Therefore, for these N tracks, there will be a peak at N-M effective tracks

.

This contamination from N-M tracks can be seen when looking at the efficiency of a cut on MWT, see Fig.22.

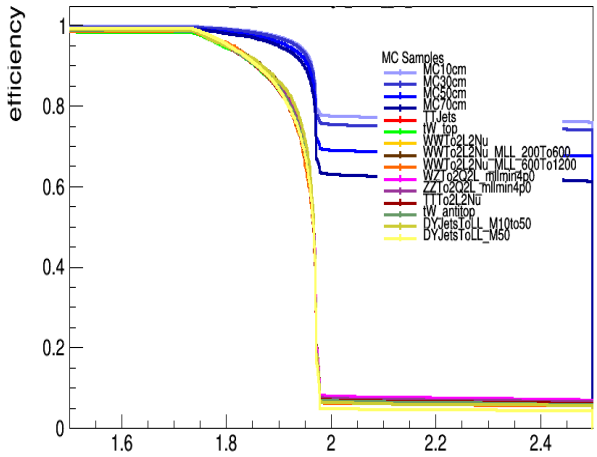

Figure 22: Efficiency of a cut applied on MWT for signal and background samples. The difference in shape between signal and background is due to the contamination of the vertices with N-M effective tracks. The difference in efficiency with respect to the MWT is due to the track-multiplicity of the background vertices being lower, mainly two to 3 tracks compared to signal vertices having 10+ tracks.

.

13.9 Parameters of the Boosted Decision Tree

This analysis is performing a multivariate analysis base on the TMVA framework of Root. TMVA allows for the implementation of a Boosted Decision Tree. A Boosted decision tree is a binary classifier based on a set of nodes making Signal-like/Background-like decisions using one input variable as shown in Fig.. The depth of the tree is given as a free parameter. The Root implementation of the BDT is a forest, the average of a set of trees to enhance the robustness with respect to the input sample. A boost is applied to also increase the stability of the response of the final classifier with respect to fluctuations in the input samples. The different parameters that are used to tune the BDT are given in Table19

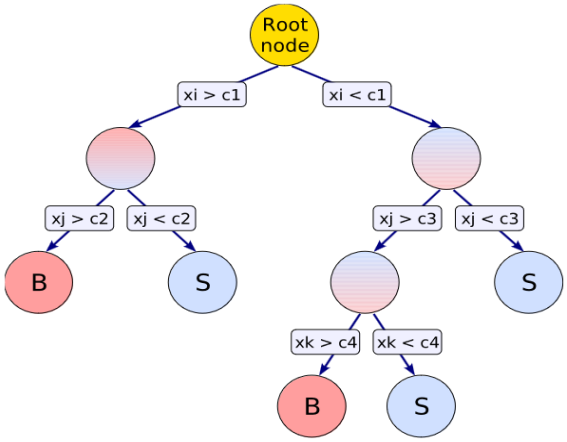

Figure 23: Example of a Boosted Decision Tree with the final leaves labelled as "S" for signal and "B" for background, extracted from [15]
. The x_i and x_j are variables used as an input to the BDT

The SeparationType GiniIndex is a cell splitting algorithm that maximises the gain of a node based on the following function :

$$gain(node) = p * (1 - p) \tag{3}$$

where $p = \frac{N_S}{N_S+N_B}$ with N_S and N_B being the number of signal and background events respectively.
The Gradient boosting (Grad) is a boosting procedure only available for decision trees that is defined as followed :

1. As mentioned above, the implementation of the BDT is a forest made of N trees called weak-learners.:

2. The function $F(x) = \sum_{m=0}^{M} = \beta * f(x)$ where β is the weight given to a single tree. To adjust the weight such that the difference between the final response of the BDT F(x) and the true value obtained from the training sample y is minimised, a loss function is implemented.

3. The Gradient Boosting uses a loss function robust to noisy samples/ samples with outliers :

$$L(F, y) = \ln\left(1 + \exp^{-2F(x)y}\right) \tag{4}$$

4. The use of Gradient Boosting is efficient for small-depth trees, between 2 and 4, reducing the possibility of overtraining.

5. The robustness can be increased by having a low learning rate (Shrinkage) between 0.1 and 0.3 but this requires a high number of trees.

Parameter	Value	Definition
NTrees	500	Number of trees in the forest
MinNodeSize	2.5%	Minimum percentage of training events required in a leaf node (default value is 5% for classification)
MaxDepth	4	Max depth of the decision tree allowed
BoostType	Grad	Boosting type for the trees in the forest
UseBaggedBoost	True	Use only a random subsample of all events for growing the trees in each iteration
GradBaggingFraction	0.6	Defines the fraction of events to be used in each iteration
Shrinkage	0.1	Learning rate for GradBoost algorithm
SeparationType	GiniIndex	Separation criterion for node splitting
nCuts	20	Number of grid points in variable range used in finding optimal cut in node splitting (default value). Increasing this value can be time consuming
UseYesNoLeaf	False	Use Sig or Bkg categories, or the purity=S/(S+B) as classification of the leaf node
UseRandomisedTrees	False	Determine at each node splitting the cut variable only as the best out of a random subset of variable
DoBoostMonitor	True	Create control plot with ROC integral vs tree number

Table 19: List of tuning parameters for the Boosted Decision Trees.

6. The efficiency of the BDT can be increased by using the BaggingSampleFraction where the best values are between 0.5 and 0.8.

Finally, the Boost Monitor helps verifying that the BDT converges well and "quickly" to a final value as shown in the Fig.[add figure as an example]. More details about these parameters are given in [15].

13.9.1 Output

References

[1] Jet identification for the 13 tev ul data.

[2] Option to make track covariance pos-def in packedcandidate.

[3] Rochester corrections for muon momentum scale and resolution.

[4] Workbook miniaod 2017.

[5] Xsdb.

[6] Search for direct production of charginos, neutralinos and sleptons in final states with two leptons and missing transverse momentum in pp collisions at $\sqrt{s} = 8$ tev with the atlas detector. *JHEP*, 05:071, 2014.

[7] Search for direct slepton and gaugino production in final states with two leptons and missing transverse momentum with the atlas detector in pp collisions at $\sqrt{s} = 7$tev. *Phys. Lett. B*, 718:879, 2014.

[8] Search for electroweak production of supersymmetric particles in final states with two or three leptons at $\sqrt{s} = 13$ tev with the atlas detector. *Eur. Phys. J. C*, 78:995, 2014.

[9] Searches for electroweak production of charginos, neutralinos, and sleptons decaying to leptons and w, z, and higgs bosons in pp collisions at 8 tev. *Eur. Phys. J. C*, 74:3036, 2014.

[10] Search for electroweak production of supersymmetric states in scenarios with compressed mass spectra at $\sqrt{s} = 13$ tev with the atlas detector. *Phys. Rev. D*, 97:052010, 2018.

[11] Search for supersymmetric partners of electrons and muons in proton–proton collisions at $\sqrt{s} = 13$ tev. *Phys*. Lett. B*, 790:140, 2019.

[12] Search for electroweak production of charginos and sleptons decaying into final states with two leptons and missing transverse momentum in $\sqrt{s} = 13$ tev pp collisions using the atlas detector. *Eur. Phys. J. C*, 80:123, 2020.

[13] S. Dimopoulos A. Arvanitaki, N. Craig and G. Villadoro. Supergauge invariant extension of the higgs mechanism and a model for the electron and its neutrino. *Nucl. Phys. B*, 90:104, 1975.

[14] S. Dimopoulos A. Arvanitaki, N. Craig and G. Villadoro. Mini-split. *JHEP*, 02:970, 2013.

[15] K. Albertssonet al. Tmva - toolkit for multivariate data analysis. 2020.

[16] B.Fuks D.Grellscheid O.Mattelaer T.Reiter C.Degrande, C.Duhr. Ufo - the universal feynrules output. 2012.

[17] ATLAS Collaboration. Reinterpretation of searches for supersymmetry in models with variable r-parity-violating coupling strength and long-lived r-hadrons. 2018.

[18] ATLAS Collaboration. Search for long-lived, massive particles in events with displaced vertices and multiple jets in pp collisions at s = 13 tev with the atlas detector. *JHEP*, 2023.

[19] Y. Cui and B. Shuve. Probing baryogenesis with displaced vertices at the lhc. *JHEP*, 02:049, 2015.

[20] Y. Cui and R. Sundrum. Baryogenesis for weakly interacting massive particles. *Phys. Rev. D*, 2013.

[21] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07(079), 2014.

[22] N. Arkani-Hamed et al. Simply unnatural supersymmetry. 2012.

[23] R. Barbier et al. R-parity violating supersymmetry. *Phys. Rept.*, 420:1, 1982.

[24] Sirunyan et. al. Search for long-lived particles decaying to jets with displaced vertices in proton-proton collisions at s = 13 tev. *Physical Review*, 81:800, 2021.

[25] G. R. Farrar and P. Fayet. Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry. *Phys. Lett. B*, 76:575, 1978.

[26] R. Frühwirth. Application of kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research*, 1987.

[27] G. F. Giudice and R. Rattazzi. Theories with gauge mediated supersymmetry breaking. *Phys. Rept.*, 322:419, 1999.

[28] G. F. Giudice and A. Romanino. Split supersymmetry. *Nucl. Phys. B*, 2004.

[29] M. Masip J. L. Hewett, B. Lillie and T. G. Rizzo. Signatures of long-lived gluinos in split supersymmetry. *JHEP*, 09:070, 2004.

[30] N. Seiberg M. Buican, P. Meade and D. Shih. Exploring general gauge mediation. *JHEP*, 03:016, 2009.

[31] G. P. Salam M. Cacciari and G. Soyez. The anti-kt jet clustering algorithm. *JHEP*, 2008.

[32] G. F. Giudice N. Arkani-Hamed, S. Dimopoulos and A. Romanino. Aspects of split supersymmetry. *Nucl. Phys. B*, 709:3, 2004.

[33] G. F. Giudice P. Gambino and P. Slavich. Gluino decays in split supersymmetry. *Nucl. Phys. B*, 726:35, 2004.

[34] N. Seiberg P. Meade and D. Shih. General gauge mediation. *Prog. Theor. Phys.*, 177:143, 2009.

[35] P. Vanlaer R. Frühwirth, W. Waltenberger. Adaptive vertex fitting. 2006.

[36] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12:P10003, 2017.

[37] M. J. Strassler and K. M. Zurek. Echoes of a hidden valley at hadron colliders. *Phys. Lett. B*, 651:374, 2007.

[38] M. J. Strassler and K. M. Zurek. Discovering the higgs through highly-displaced vertices. *Phys. Lett. B*, 661:263, 2008.

[39] K. M. Zurek T. Han, Z. Si and M. J. Strassler. Phenomenology of hidden valleys at hadron colliders. *JHEP*, 07:008, 2008.

[40] S. Weinberg. Supersymmetry at ordinary energies. 1. masses and conservation laws. *Phys. Rev. D*, 26:287, 1982.

[41] L. Randall Y. Cui and B. Shuve. A wimpy baryogenesis miracle. *JHEP*, 04:075, 2012.