

## Chapter 7

### SEARCH FOR MASSIVE SUPERSYMMETRY AT 13 TEV

This chapter presents a search for supersymmetry involving pair-produced gluinos ( $\tilde{g}$ ) decaying via off-shell top squarks ( $\tilde{t}$ ) into the lightest neutralino ( $\tilde{\chi}_1^0$ ) and third-generation [Standard Model \(SM\)](#) particles:  $\tilde{g} \rightarrow \tilde{t}_1 \bar{t} \rightarrow t \bar{t} \tilde{\chi}_1^0$ . This analysis was performed using  $36.1 \text{ fb}^{-1}$  of data collected by the ATLAS experiment in 2015 and 2016. In the Run-I iteration of this analysis in 2014 [1], no significant excess of events over the expected [SM](#) background was observed and a lower limit on the  $\tilde{g}$  mass was set to be at 1.4 TeV at the 95% confidence level for this signal model. I discuss the work done during the Run-2 iteration of this analysis, also observing no significant excess, and setting a much stronger lower limit on the  $\tilde{g}$  mass to 2.1 TeV at the 95% confidence level.

My work described in the rest of the chapter is organized as follows: section [7.1](#) describes the signal models, the data sets, and the [monte-carlo \(MC\)](#) samples used to model the signal and backgrounds; section [7.2](#) describes the kinematic objects and event selections used in the various channels; section [7.3](#) presents the preselection, background estimation, and reweighting techniques; section [7.4](#) describes the details of the cut-and-count analysis performed, optimizing for signal discovery; section [7.5](#) describes the regions defined as a result of the previous section; section [7.6](#) describes the semi-data-driven normalization of [MC](#) to data; and section [7.7](#) describes the main sources of systematic uncertainty. The next chapter presents the results of this search.

## 7.1 Searching for New Physics: A Counting Experiment

### 7.1.1 Signal Models

I performed a search for pair-produced gluinos decaying via top squarks in events with multiple jets originating from the hadronization of  $b$ -quarks, high  $E_T^{\text{miss}}$ , and potentially additional light-quark jets and/or an isolated charged lepton. The results I present in ?? supersede the ones obtained previously using  $3.2 \text{ fb}^{-1}$  of data collected in 2015 using the same strategy [2].

Various simplified SUSY models [3, 4] are employed to optimise the event selection and/or interpret the results of the search. In terms of experimental signature, they all contain at least four  $b$ -jets originating from either gluino or top quark decays, and two  $\tilde{\chi}_1^0$ , which escape the detector unseen, resulting in high  $E_T^{\text{miss}}$ . Each top quark decays to a  $W$  boson and a  $b$ -quark as described in ?? and a final state with four tops will have at least four  $b$ -jets.

Gluinos are assumed to be pair-produced and to decay as  $\tilde{g} \rightarrow \tilde{t}_1 \bar{t} \rightarrow t \bar{t} \tilde{\chi}_1^0$ <sup>1</sup>. In all cases, the stop squarks are assumed to be off-shell in order to have simplified models with only two parameters: the gluino and  $\tilde{\chi}_1^0$  masses.<sup>2</sup> All other sparticles are decoupled.

A simplified model is used to optimize the event selection and to interpret the results. In the Gtt model, illustrated in fig. 7.1, each gluino undergoes an effective three-body decay  $\tilde{g} \rightarrow t \bar{t} \tilde{\chi}_1^0$  via off-shell top squarks, with a branching ratio of 100%. The Gtt model has four  $b$ -jets and two  $\tilde{\chi}_1^0$ , and four  $W$  bosons originating from the top quark decays,  $t \rightarrow Wb$ , in its final state.

---

<sup>1</sup>The charge conjugate process is implied.

<sup>2</sup>The analysis sensitivity is found to be mostly independent of the top squark mass, except when the top squark is very light [1].

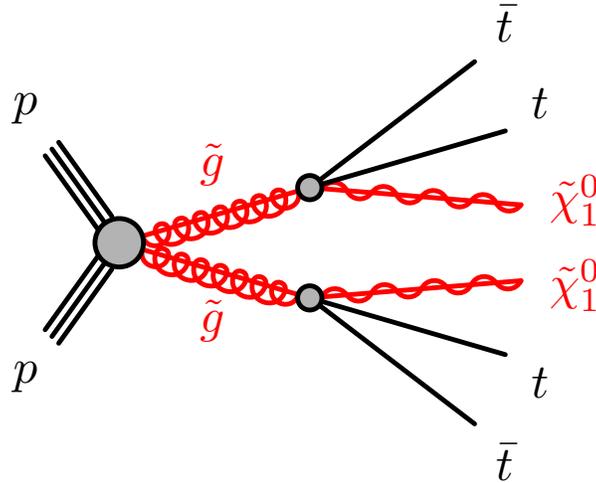


Figure 7.1: The decay topology in the off-shell Gtt simplified model.

The technical implementation of the simulated samples produced from these models is described in ??.

## 7.2 Kinematic Variables and Event Selection

In [ATLAS](#), many of the kinematics and selections are standardized through a central analysis framework. This section will detail the definitions used in the analysis from the perspective of the physics of the detector. Corrections derived from data control samples are applied to simulated events to account for differences between data and simulation in the reconstruction efficiencies, momentum scale and resolution of leptons (electrons and muons), in the efficiency and fake rate for identifying  $b$ -jets, and in the efficiency for rejecting jets originating from pile-up interactions.

There are three primary classes of event selection observables [5] that are sensitive to distinct features of [Supersymmetry \(SUSY\)](#) processes:

- **Missing energy-type:** sensitive to the properties of the invisible states, e.g. how

many neutralinos in the event, what is their mass, etc.;

- **Energy scale-type:** sensitive to the overall energy scale of the event, e.g. the mass of the gluino  $m_{\tilde{g}}$ ;
- **Energy structure-type:** sensitive to the structure of the visible energy, e.g. how many partons are generated in the decay, how that energy is partitioned across the final state visible and invisible objects.

Observables that fall into each of these classes are used in the analysis search I performed and are discussed in the following sections.

### 7.2.1 Kinematic Variables

As described in section 7.1.1, the final state of the signal is very busy. There can be 3 or 4  $b$ -jets and, depending on the mass splitting between the  $\tilde{g}$  and  $\tilde{\chi}_1^0$ , lots of missing transverse momentum with many highly energetic jets. It is useful to define some variables that discriminate the supersymmetry signal models against the standard model background. This section describes the definition of those variables that will be used scan the phase-space of the analysis in section 7.4 to form search channels in section 7.5 where the signal is expected to dominate over background.

## Object Multiplicity

The number of jets,  $N_{\text{jet}}$ , and number of  $b$ -jets,  $N_{b\text{-jets}}$ , are counting variables which count the number of selected jets and number of selected  $b$ -jets in the event.

## Effective Mass

The effective mass variable ( $m_{\text{eff}}$ ) is defined as:

$$m_{\text{eff}} = \sum_{\text{jets}}^{N_{\text{jet}}} p_{\text{T}}^{\text{jet}} + \sum_{\text{leptons}}^{N_{\text{lepton}}} p_{\text{T}}^{\ell} + E_{\text{T}}^{\text{miss}}, \quad (7.1)$$

which is a scalar sum of the  $p_{\text{T}}$  of the selected objects in the event. It typically has a much higher value in pair-produced gluino ( $\tilde{g}$ ) events than in background events.

## Transverse Mass

In regions with at least one selected lepton, the transverse mass  $m_{\text{T}}$  composed of the  $p_{\text{T}}$  of the leading selected lepton ( $\ell$ ) and  $E_{\text{T}}^{\text{miss}}$  is defined as:

$$m_{\text{T}} = \sqrt{2p_{\text{T}}^{\ell} E_{\text{T}}^{\text{miss}} \{1 - \cos[\Delta\phi(\vec{p}_{\text{T}}^{\text{miss}}, \vec{p}_{\text{T}}^{\ell})]\}}. \quad (7.2)$$

It is used to reduce the  $t\bar{t}$  and  $W$ +jets background events in which a  $W$  boson decays leptonically. Neglecting resolution effects, the  $m_{\text{T}}$  distribution for these backgrounds has an expected upper bound corresponding to the  $W$  boson mass and typically has higher values for Gtt events. Another useful transverse mass variable is  $m_{\text{T},\text{min}}^{b\text{-jets}}$ , the minimum transverse mass formed by  $E_{\text{T}}^{\text{miss}}$  and any of the three highest- $p_{\text{T}}$   $b$ -tagged jets in the event:

$$m_{\text{T},\text{min}}^{b\text{-jets}} = \min_{i \leq 3} \left( \sqrt{2p_{\text{T}}^{b\text{-jet}_i} E_{\text{T}}^{\text{miss}} \{1 - \cos[\Delta\phi(\vec{p}_{\text{T}}^{\text{miss}}, \vec{p}_{\text{T}}^{b\text{-jet}_i})]\}} \right). \quad (7.3)$$

The  $m_{\text{T},\text{min}}^{b\text{-jets}}$  distribution has an expected upper bound corresponding to the top quark mass for  $t\bar{t}$  events with a semileptonic top quark decay. Since the  $\tilde{\chi}_1^0$ , which produces  $E_{\text{T}}^{\text{miss}}$  in [SUSY](#) events, is largely independent of the  $b$ -jet kinematics, the value of  $m_{\text{T},\text{min}}^{b\text{-jets}}$  can peak

at larger values for signal processes.

## Total Jet Mass

Another powerful variable is the total jet mass variable,  $M_J^\Sigma$  [6, 7], defined as:

$$M_J^\Sigma = \sum_{i \leq 4} m_{J,i}, \quad (7.4)$$

where  $m_{J,i}$  is the mass of the large-radius re-clustered jet  $i$  in the event. The decay products of a hadronically decaying boosted top quark can be reconstructed in a single large-radius re-clustered jet, resulting in a jet with a high mass. This variable typically has larger values for Gtt events than for background events. This is because Gtt events contain as many as four<sup>3</sup> hadronically decaying top quarks while the background is dominated by  $t\bar{t}$  events with one or two semileptonic top quark decays. The presence of numerous boosted and semi-boosted top quarks in the signal events leads to the formation of high- $p_T$ , massive jets at the scale of  $R \approx 0.8$ . This variable is sensitive to the large degree of structure in signal events compared to background processes which are dominated by semi-leptonic and di-leptonic  $t\bar{t}$  events (figs. 7.5 and 7.6).

## Multi-jet Suppression

The requirement of a selected lepton, with the additional requirements on jets,  $E_T^{\text{miss}}$  and event variables described above, will make the multi-jet background negligible for the  $\geq 1$ -lepton signal regions. For the 0-lepton signal regions, the minimum azimuthal angle  $\Delta\phi_{\text{min}}^{4j}$  between  $\vec{p}_T^{\text{miss}}$  and the  $p_T$  of the four leading **small-radius (small- $R$ )** jets in the event, defined

---

<sup>3</sup>In the case of events with less than four re-clustered jets, all of them are used.

as:

$$\Delta\phi_{\min}^{4j} = \min_{i \leq 4} \left( |\phi_{\text{jet}_i} - \phi_{\vec{p}_T^{\text{miss}}}| \right), \quad (7.5)$$

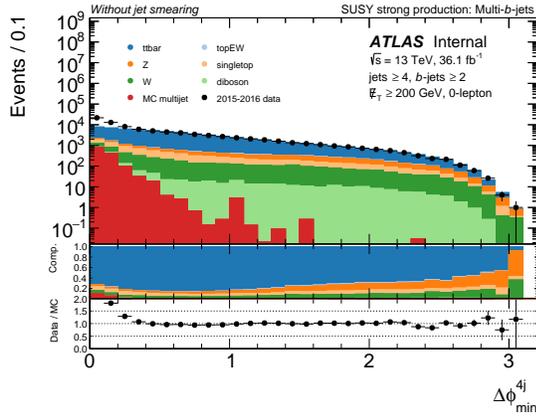
is required to be greater than 0.4. This requirement suppresses the multi-jet background, which can produce events with large  $E_T^{\text{miss}}$  if containing poorly measured jets or neutrinos emitted close to the axis of a jet. In particular, anything that is all-hadronic is typically mis-modeled<sup>4</sup> with monte-carlo simulations so contributions from a multi-jet background are estimated using a data-driven approach “jet smearing” (described in [8]). This method performs a smearing of the jet response in data events with well-measured  $E_T^{\text{miss}}$  (“seed events”) to estimate the tail of  $E_T^{\text{miss}}$  where the modeling is not as great. Figure 7.2 shows the impact of jet smearing on the modeling of multi-jet as compared to 2015+2016 data. This is even more important for  $b$ -jets because in the high tails of effective mass, most of the mis-measurement in the hadronic decays of multi-jets comes in the form of heavy-flavor decays. As fig. 7.2 also shows, a selection of  $\Delta\phi_{\min}^{4j} > 0.4$  effectively suppresses a majority of the multi-jet background in the search, so the multi-jet background is typically called “reducible”.

### 7.2.2 Event Selection

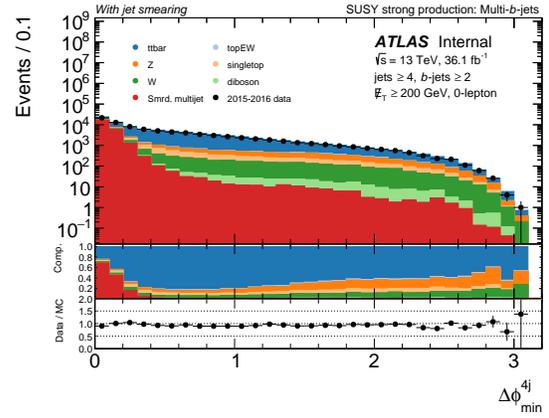
Two different types of selections are applied in succession for the analysis, event selection and then preselection. Two sets of preselection criteria targeting the 0-lepton and the 1-lepton channels are presented in section 7.3. The modeling of the data in these regions is also discussed in that section. The event selection criteria, defined based on kinematic requirements for the objects defined in ?? and discriminating event-based variables described in section 7.2.1, are used to further reject the background. There are a number of more common general cuts which are applied to all events described in the next few sections.

---

<sup>4</sup>The generators have different ways of calculating the non-perturbative QCD showers.



(a) without jet smearing



(b) with jet smearing

Figure 7.2: Distributions of the  $\Delta\phi_{\min}^{4j}$  variable before and after jet smearing is applied for 2015-2016 data using a 0-lepton preselection, without the  $\Delta\phi_{\min}^{4j}$  selection applied. The top panel shows the distribution in log-scale of the number of events in each bin of  $\Delta\phi_{\min}^{4j}$ . The second panel shows the background composition with each background contribution some fraction of the total background. The third panel shows the data/MC ratio where data is compared to the total estimated, pre-fit background. Looking at the background composition, it is clear that the multi-jet background (red) is reducible, as is evidenced by a  $\Delta\phi_{\min}^{4j} > 0.4$  selection that can be applied.

These cuts are

- [Good Runs List \(GRL\)](#): applied to data only
- [Tile calorimeter \(Tile\)](#), [LAr](#), [Semiconductor Tracker \(SCT\)](#): applied to data only
- Trigger: applied to both data and [MC](#)
- Jet Cleaning: applied to both data and [MC](#)
- Muon Cleaning: applied to both data and [MC](#)

## Good Runs

[ATLAS](#) provides an XML file containing a list of events in data that have passed data quality criteria. This removes luminosity blocks with poor detector quality. In particular, this analysis is sensitive to  $b$ -tagging and runs with the [Insertable B-Layer \(IBL\)](#) turned off are not included. Two different [GRLs](#) are used: one for 2015 data and one for 2016 data.

## Tile, LAr, and SCT

Following [ATLAS](#) recommendations, corrupted events due to errors in the [Tile](#), [LAr](#), or [SCT](#) are vetoed.

## Trigger

The unprescaled<sup>5</sup>  $E_T^{\text{miss}}$  triggers used for both the 0-lepton and 1-lepton analyses are `HLT_xe70` for 2015 data and `HLT_xe100_mht_L1XE50/HLT_xe110_mht_L1XE50` for early/late 2016 data,

---

<sup>5</sup>If a trigger is prescaled, this means the trigger rate is purposefully decreased in order to keep the output rate manageable.

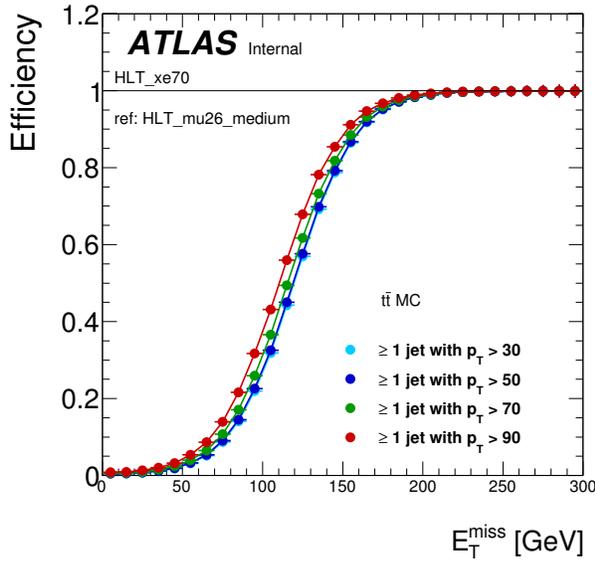
respectively. One important feature of the  $E_{\text{T}}^{\text{miss}}$  triggers is to make sure that the selections in our analysis lie above the turn-on curve. This means that the trigger does not affect the selection efficiency when searching for signal over background.

The lowest unprescaled trigger for the 2015 dataset is 70 GeV and for 2016 dataset is 100 GeV and 110 GeV. The efficiencies of these triggers are shown in fig. 7.3 for the 2015 dataset and fig. 7.4 for the 2016 dataset. These figures show the efficiency of the  $E_{\text{T}}^{\text{miss}}$  trigger in both  $t\bar{t}$  MC and 25 ns data for different requirements on the leading jet  $p_{\text{T}}$ . The trigger is plotted as a function of a modified, reconstructed  $E_{\text{T}}^{\text{miss}}$ . Typically, when  $E_{\text{T}}^{\text{miss}}$  is reconstructed offline, it incorporates information from the muon spectrometer, but this information is not available in Level-1 (L1). To properly study the L1  $E_{\text{T}}^{\text{miss}}$  trigger, muons are subtracted from the reconstructed  $E_{\text{T}}^{\text{miss}}$  to reproduce the L1  $E_{\text{T}}^{\text{miss}}$ . Since our preselection, described in section 7.3, requires  $E_{\text{T}}^{\text{miss}} > 200$  GeV, you can see that both  $E_{\text{T}}^{\text{miss}}$  triggers are fully efficient.

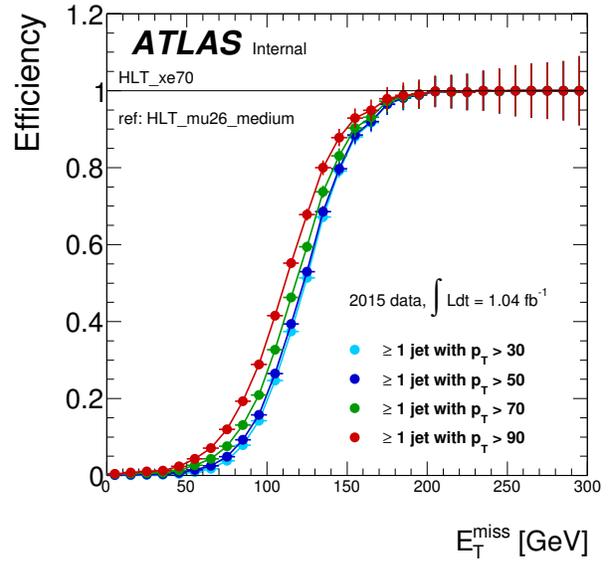
## Jet Cleaning

Fake jets can arise from non-collision background or cosmic muons with a catastrophic energy loss in the calorimeters or from fake signals in the calorimeter, arising either from noise bursts or the presence of coherent noise. A set of cuts having a high rejection against fake jets while preserving an efficiency close to 100% for selected jets has been designed. The selection criteria are based on

- the timing of the calorimeter signal with respect to that of the Bunch Crossing (BC),
- the quality of the fit on the calorimeter pulse shape,
- the fraction of jet energy belonging to specific calorimeter samples, and

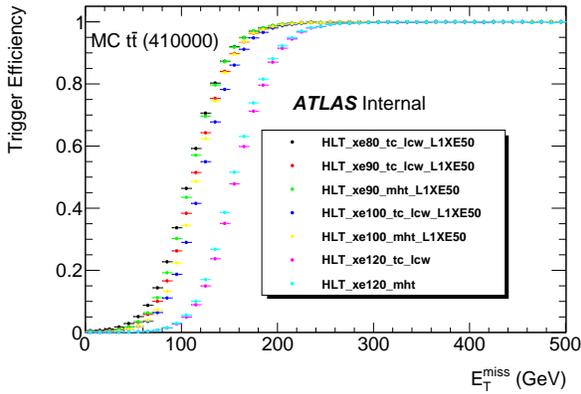


(a)  $t\bar{t}$  MC

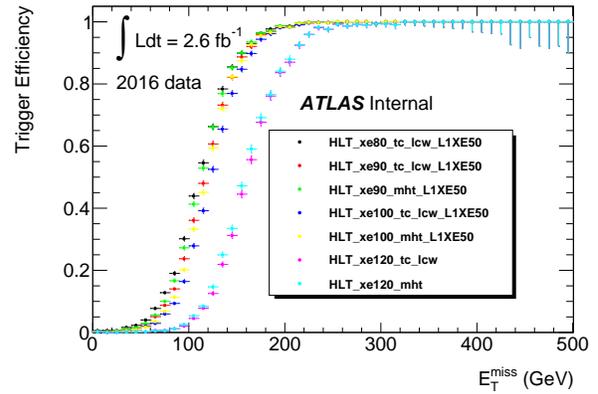


(b) 25 ns data

Figure 7.3: [9] Efficiency of the 2015  $E_T^{\text{miss}}$  trigger for different requirements on the leading jet  $p_T$ . The muons are subtracted from  $E_T^{\text{miss}}$  to reproduce the L1  $E_T^{\text{miss}}$ . The reference trigger used to select events is HLT\_mu26\_medium.



(a)  $t\bar{t}$  MC



(b) 25 ns data

Figure 7.4: [9] Efficiency of the various 2016  $E_T^{\text{miss}}$  triggers as a function of the offline  $E_T^{\text{miss}}$ . The muons are subtracted from  $E_T^{\text{miss}}$  to reproduce the L1  $E_T^{\text{miss}}$ . The reference trigger used to select events is HLT\_mu26\_ivarmedium. One of the triggers used in the analysis for early 2016 data is shown here: HLT\_xe100\_mht\_L1XE50.

- the charged momentum fraction<sup>6</sup> of the jet.

The selections are summarized at [10]. Events are vetoed if any of the baseline jets are determined to be “bad” based on the above criteria.

## Muon Cleaning

Another set of event cleaning cuts come from identifying two types of muons, “fake” muons and cosmic-ray muons as described below. If either of these muons are identified in an event, the event is excluded from the analysis search.

1. There are fake muons that can be reconstructed from high hit multiplicities in the muon spectrometer due to very energetic [punch-through](#) jets or from badly measured inner detector tracks in jets wrongly matched to muon spectrometer segments. Such fake muons contribute to fake  $E_T^{\text{miss}}$ . Events containing one or more fake muon candidates are rejected.
2. Events with muons arising from cosmic rays are also rejected to avoid the effects on the tails of the  $E_T^{\text{miss}}$ . Potential cosmic muons are identified after the muon-jet overlap removal (see ??) as muons with large longitudinal and transverse impact parameters. Events containing one or more muon candidates with

$$|d_0| > 0.2 \text{ mm} \quad \text{and} \quad (7.6a)$$

$$|z_0| > 1 \text{ mm} \quad (7.6b)$$

are rejected to suppress cosmic rays.

---

<sup>6</sup>as measured by the [Inner Detector \(ID\)](#) relative to the calorimeter

Table 7.1: Definitions of the 0-lepton and 1-lepton preselection of the cut-and-count analysis.  $m_{\text{eff}}$ ,  $m_{\text{T,min}}^{b\text{-jets}}$ , and  $m_{\text{T}}$  have no preselection cuts for both channels.

	0-lepton	1-lepton
$N_{\text{lepton}}$	= 0	$\geq 1$
Trigger (section 7.2.2)	70 GeV unprescaled	100 GeV unprescaled
$E_{\text{T}}^{\text{miss}}$	> 200 GeV	
$N_{\text{jet}}$	> 4	
$N_{b\text{-jets}}$	> 2	
$\Delta\phi_{\text{min}}^{4j}$	> 0.4	-

### 7.3 Pre-selection comparisons of Data/MC

This section contains distributions of the analysis observables described in section 7.2.1, showing the pre-fit modeling of these observables in the following loose selection regions. All are triggered with the  $E_{\text{T}}^{\text{miss}}$  triggers which are fully efficient in our offline  $E_{\text{T}}^{\text{miss}}$  cut.

Preselection criteria<sup>7</sup>, as summarized in table 7.1, in the 0-lepton and 1-lepton channels require  $E_{\text{T}}^{\text{miss}} > 200$  GeV, in addition to the  $E_{\text{T}}^{\text{miss}}$  trigger requirement, and at least four jets of which at least two must be  $b$ -tagged. The 0-lepton channel requires no selected leptons. The 1-lepton channel requires the event to contain at least one selected lepton. Because the 0-lepton channel requires no selected leptons, a  $\Delta\phi_{\text{min}}^{4j}$  selection is additionally required to reduce the multijet background as described in section 7.2.1.

In this analysis, correction factors need to be extracted to account for shape discrepancies in the  $m_{\text{eff}}$  spectrum between the data and the expected background for the 1-lepton preselection sample<sup>8</sup>. These factors are defined as the ratio of the number of observed events to

<sup>7</sup>Preselection is designed to be 100% efficient for the signal processes under study.

<sup>8</sup>This effective mass reweighting is described in more detail in [11]. This was a shape discrepancy that was not observed in 0-lepton. The underlying cause is still not understood and the current, 2017 iteration of

the predicted number of background events in a given  $m_{\text{eff}}$  bin, in a signal-depleted region. This region is defined by applying the 1-lepton preselection criteria and requiring exactly two  $b$ -tagged jets and  $m_{\text{T},\text{min}}^{b\text{-jets}} < 140$  GeV. This kinematic reweighting leads to correction factors ranging from 0.7 to 1.1. They are applied to the background prediction and the full size of the correction is taken as an uncertainty for both the background and signal events.

Figures 7.5 and 7.6 show the multiplicity of selected jets and  $b$ -tagged jets, the distributions of  $E_{\text{T}}^{\text{miss}}$ ,  $m_{\text{eff}}$ , and  $M_{\text{J}}^{\Sigma}$  for events passing the 0-lepton or the 1-lepton preselection, respectively. Figures 7.5 and 7.6 also display the distribution of transverse mass:  $m_{\text{T},\text{min}}^{b\text{-jets}}$  for 0-lepton channel and  $m_{\text{T}}$  for 1-lepton channel. The  $m_{\text{eff}}$  correction described above is applied in the 1-lepton channel.

The uncertainty bands include the statistical and experimental systematic uncertainties, as described in section 7.7, but not the theoretical uncertainties in the background modeling.

The data and the predicted background are found to agree reasonably well at the preselection level after the kinematic reweighting described above. A discrepancy between data and prediction is observed for the number of  $b$ -tagged jets, but it has a negligible impact on the background estimate after the renormalization of the simulation in dedicated control regions with the same  $b$ -tagged jets requirements as the signal regions, as described in section 7.5. Example signal models with enhanced cross-sections are overlaid for comparison.

## 7.4 Optimizations

The main analysis strategy documented in this thesis is a **cut-and-count** analysis, using partially overlapping single-bin **signal regions (SRs)**, optimized to maximize the expected discovery power for benchmark signal models, and allowing for reinterpretation of the results.

---

this analysis continues to study this more.

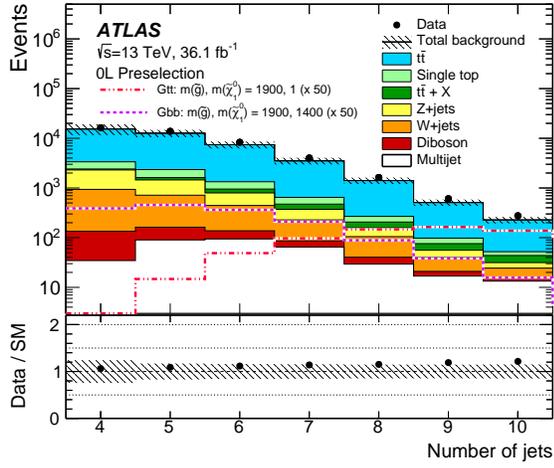
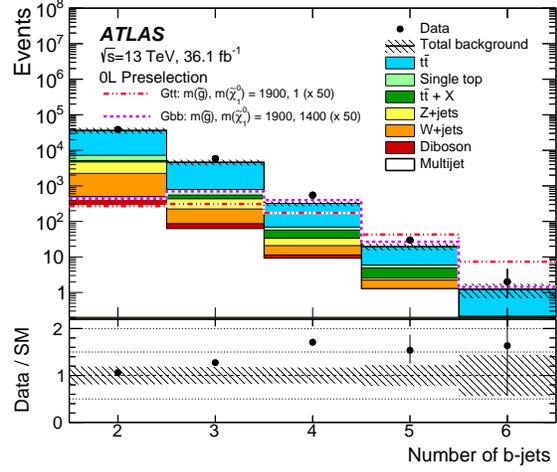
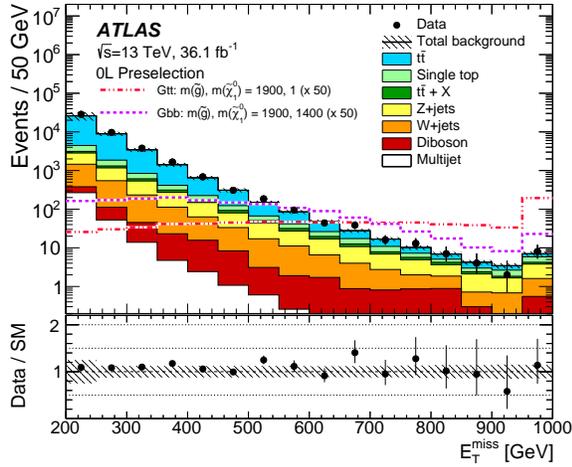
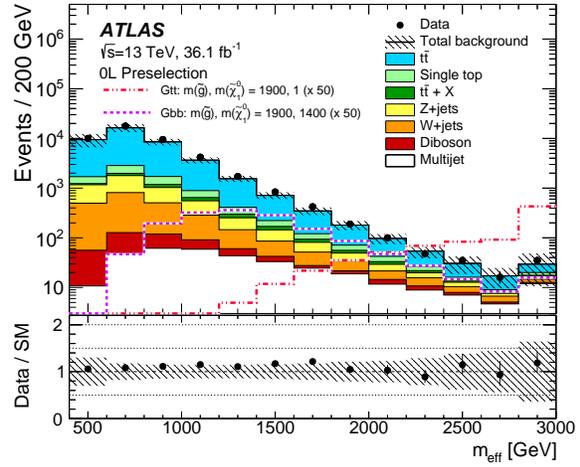
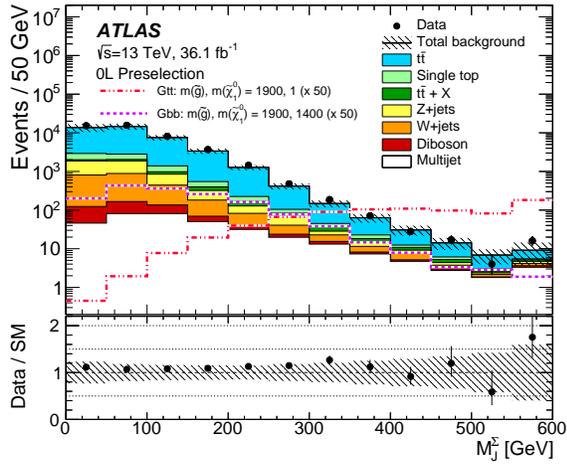
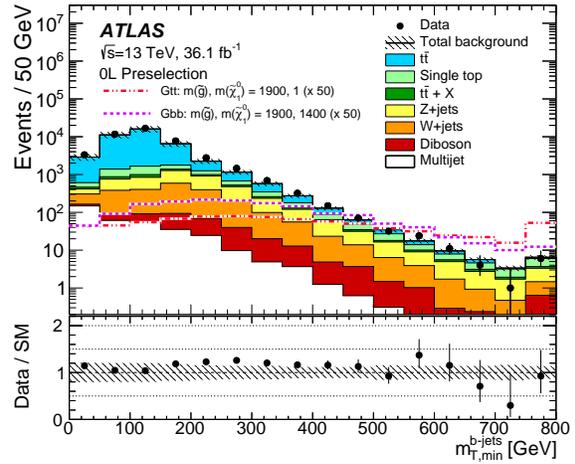
(a)  $N_{\text{jet}}$ (b)  $N_{b\text{-jets}}$ (c)  $E_T^{\text{miss}}$ (d)  $m_{\text{eff}}$ 

Figure 7.5: Distributions of discriminating variables for events passing the 0-lepton pre-selection criteria. The statistical and experimental systematic uncertainties (as defined in section 7.7) are included in the uncertainty band. The last bin includes overflow events. The lower part of each figure shows the ratio of data to the background prediction. All backgrounds (including  $t\bar{t}$ ) are normalized using the best available theoretical calculation described in ???. The background category  $t\bar{t} + X$  includes  $t\bar{t}W/Z$ ,  $t\bar{t}h$  and  $t\bar{t}t\bar{t}$  events. Example signal models with cross-sections enhanced by a factor of 50 are overlaid for comparison.



(e)  $M_J^\Sigma$



(f)  $m_{T,\min}^{b\text{-jets}}$

Figure 7.5: Distributions of discriminating variables for events passing the 0-lepton preselection criteria. The statistical and experimental systematic uncertainties (as defined in section 7.7) are included in the uncertainty band. The last bin includes overflow events. The lower part of each figure shows the ratio of data to the background prediction. All backgrounds (including  $t\bar{t}$ ) are normalized using the best available theoretical calculation described in ???. The background category  $t\bar{t} + X$  includes  $t\bar{t}W/Z$ ,  $t\bar{t}h$  and  $t\bar{t}t\bar{t}$  events. Example signal models with cross-sections enhanced by a factor of 50 are overlaid for comparison.

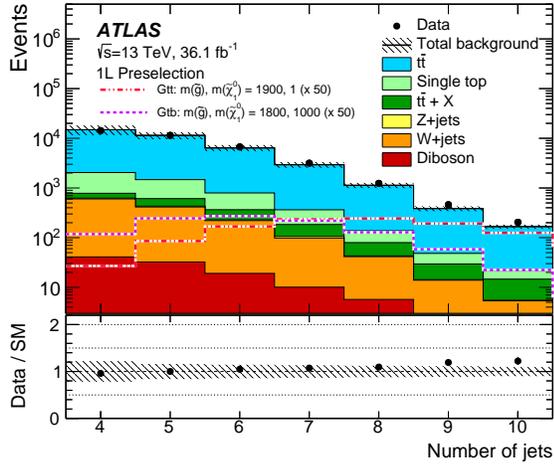
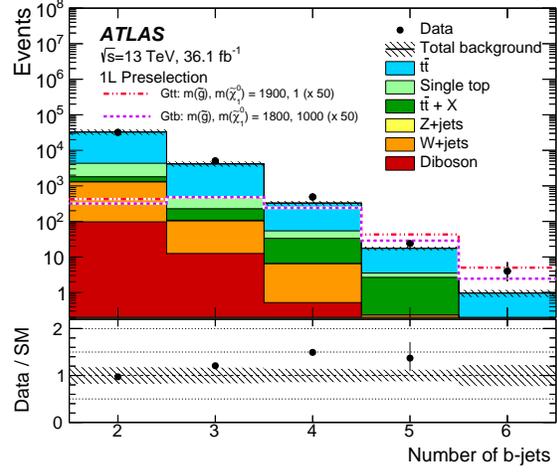
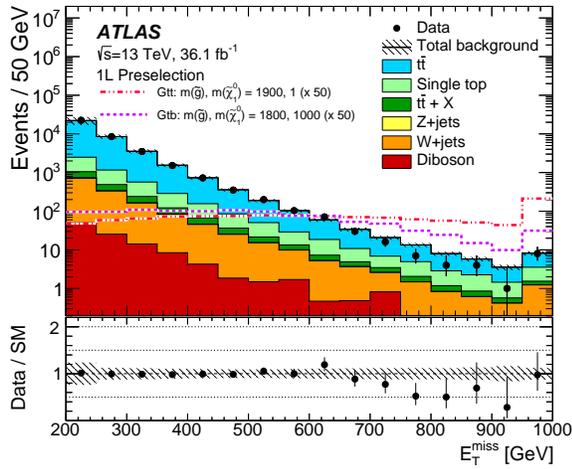
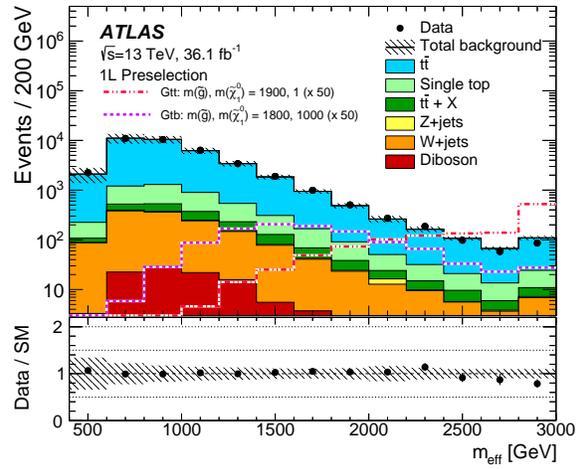
(a)  $N_{\text{jet}}$ (b)  $N_{b\text{-jets}}$ (c)  $E_T^{\text{miss}}$ (d)  $m_{\text{eff}}$ 

Figure 7.6: Distributions of discriminating variables for events passing the 1-lepton preselection criteria, after applying the kinematic reweighting to the  $m_{\text{eff}}$  distribution described in the text. The statistical and experimental systematic uncertainties (as defined in section 7.7) are included in the uncertainty band. The last bin includes overflow events. The lower part of each figure shows the ratio of data to the background prediction. All backgrounds (including  $t\bar{t}$ ) are normalized using the best available theoretical calculation described in ???. The background category  $t\bar{t} + X$  includes  $t\bar{t}W/Z$ ,  $t\bar{t}h$  and  $t\bar{t}t\bar{t}$  events. Example signal models with cross-sections enhanced by a factor of 50 are overlaid for comparison.

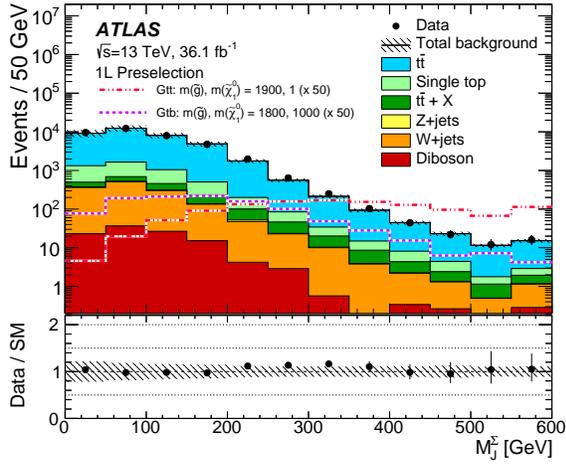
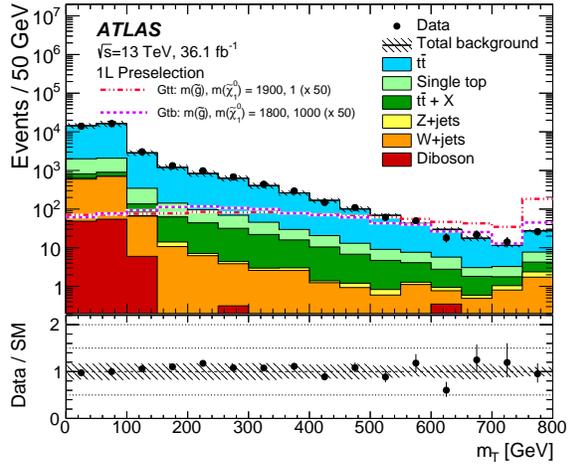
(e)  $M_J^\Sigma$ (f)  $m_T$ 

Figure 7.6: Distributions of discriminating variables for events passing the 1-lepton preselection criteria, after applying the kinematic reweighting to the  $m_{\text{eff}}$  distribution described in the text. The statistical and experimental systematic uncertainties (as defined in section 7.7) are included in the uncertainty band. The last bin includes overflow events. The lower part of each figure shows the ratio of data to the background prediction. All backgrounds (including  $t\bar{t}$ ) are normalized using the best available theoretical calculation described in ???. The background category  $t\bar{t} + X$  includes  $t\bar{t}W/Z$ ,  $t\bar{t}h$  and  $t\bar{t}t\bar{t}$  events. Example signal models with cross-sections enhanced by a factor of 50 are overlaid for comparison.

This is described in more detail in section 7.4.1. In order to define these regions, it helps to scan the phase space of the kinematic variables to determine the strongest selections to maximize signal over background and this strategy is described in more detail in section 7.4.2. After performing the optimization, the regions for the cut-and-count analysis are defined in section 7.5.

#### 7.4.1 Analysis strategy and background treatment

In order to enhance the sensitivity to the various signal benchmarks described in section 7.1.1, multiple **SRs** are defined. The **SRs** are defined to probe the existence of a signal or to assess model-independent upper limits on the number of signal events. The main background in all these regions is the production of a  $t\bar{t}$  pair in association with heavy-flavor and light-flavor jets. A normalization factor ( $\mu_{t\bar{t}}$ ) for this background is extracted for each individual **SR** from a data **control region (CR)** that has comparable background composition and kinematics. This is ensured by keeping the kinematic requirements similar in the two regions. The **CRs** and **SRs** are defined to be mutually exclusive. Signal contributions in the **CRs** are suppressed by inverting or relaxing some requirements on the kinematic variables (e.g.  $m_T$  or  $m_{T,\min}^{b\text{-jets}}$ ), leading to a signal contamination in the **CRs** of 6% at most.

The  $t\bar{t}$  normalization is cross-checked in **validation regions (VRs)** that share similar background composition, i.e. jet and lepton flavors, with the **SR**. The signal contamination in the **VRs** is found to be lower than 30% for benchmark signal mass points above the already excluded mass range. The  $t\bar{t}$  purity is superior to 73% and 53% in the **CRs** and **VRs**, respectively. All of this is summarized in fig. 7.7.

The non- $t\bar{t}$  backgrounds mainly consist of single-top,  $W$ +jets,  $Z$ +jets,  $t\bar{t} + W/Z/h$ ,  $t\bar{t}t\bar{t}$  and diboson events. Their normalization is taken from the simulation normalized using the best

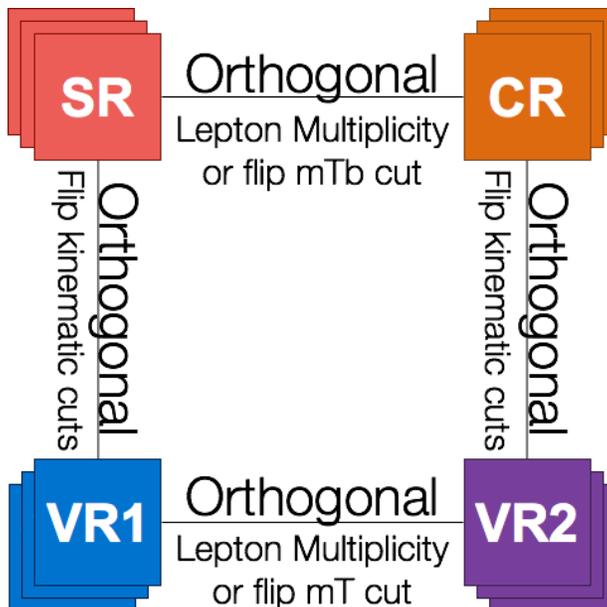


Figure 7.7: A diagrammatic overview of how the different regions **SR**, **CR**, and **VR** are defined for the cut-and-count analysis

available theory prediction. The multijet background is found to be very small or negligible in all regions. It is estimated using a procedure described in [12], in which the jet response is determined from simulated dijet events. This response function is then used to smear the jet response in low- $E_T^{\text{miss}}$  events. The jet response is cross-checked with data where the  $E_T^{\text{miss}}$  can be unambiguously attributed to the mismeasurement of one of the jets.

#### 7.4.2 Optimization Strategy

This section describes the optimization of the **SRs**, **CRs**, and **VRs**. For the **SRs**, the figure of merit is the expected significance, defined by the function `BinomialExpZ` of `RooStats` [13, 14] assuming  $35 \text{ fb}^{-1}$  of data<sup>9</sup> This function estimates a significance for a signal above **SM** background, given the Poisson likelihood, with an estimate of the impact of systematic

<sup>9</sup>2015+2016 data accounted for  $36.1 \text{ fb}^{-1}$  at the end of 2016, so this is not too far off.

uncertainties<sup>10</sup>. We require that the MC statistical uncertainty on the background is less than 30% to ensure that the optimized selection is reliable. The optimization then motivates defining a number of SRs. The goal of these SRs is to provide robust regions capable of discovery of SUSY signatures, and for ease of re-interpretation by theorists [15].

Two categories of these discovery regions are defined: Gtt-0L and Gtt-1L in sections 7.4.3 and 7.4.4. Gtt-0L targets the zero lepton channel of the  $\tilde{g} \rightarrow t\bar{t}$  process and Gtt-1L targets the one lepton channel. Determination of the signal and CRs for the cut-and-count analysis is carried out with several metrics in mind:

1. The expected significance is used as the primary metric to evaluate the performance of SRs.
2. The expected number of signal events must be  $N_S > 2$  in order to ensure possibility of an observation.
3. The total background must always be  $N_B > 0.5$ .
4. The background composition within SRs and CRs is required to be dominated by  $t\bar{t}$ . The presence of other backgrounds, whose contributions are evaluated directly from simulation, is minimized.<sup>11</sup>
5. The statistical uncertainty on  $t\bar{t}$  backgrounds must never be larger than 30%.
6. The signal contamination in validation and CRs is minimized.

I wrote a package, an Optimization Framework, that automates and streamlines this entire process in a very generic way. See ?? for more about the technical implementation.

---

<sup>10</sup>This is the standard significance optimization technique within the SUSY group in ATLAS.

<sup>11</sup>The post-fit  $t\bar{t}$  purity in all regions is  $> 50\%$ , and typically  $> 70\%$ .

### 7.4.3 *Gtt-0L Optimization*

The Gtt 0L optimization used the following variables which are summarized in table 7.2.

	inclusive		step
	start	stop	size
$N_{\text{jet}}$	6	10	1
$N_{b\text{-jets}}$	3	6	1
$E_{\text{T}}^{\text{miss}}$	200 GeV	1000 GeV	50 GeV
$m_{\text{eff}}$	800 GeV	3000 GeV	100 GeV
$M_{\text{J}}^{\Sigma}$	0 GeV	500 GeV	100 GeV
$m_{\text{T},\text{min}}^{b\text{-jets}}$	0 GeV	200 GeV	4 GeV0

Table 7.2: The 0-lepton optimization phase-space that was scanned over. This corresponded to 16,560 cuts. Each variable was scanned from “start” to “stop” in discrete step sizes. The starting and stopping values of the scan are inclusive.

Since the masses of the new particles are not known yet, the grid plotted in figs. 7.8 and 7.9 represents all the mass points for the  $\tilde{g}\text{-}\tilde{\chi}_1^0$  pair with the mass of the gluino on the  $x$ -axis in GeV and the mass of the neutralino on the  $y$ -axis in GeV. Figure 7.8 shows the optimal significance possible with the best selected variables at each grid point. Figure 7.9 shows the optimal cuts on various variables selected at each mass point. These were produced using the optimization framework that I wrote for this analysis, but is general enough to be applied to similar types of analyses doing a search and trying to understand the phase-space of their signal model. Being able to visualize the change in the optimal cut for each variable across the different gluino-neutralino masses influences our choices for signal region definitions. For example, we can see two regions that appear based on the  $b$ -jet multiplicity where most mass points prefer requiring  $N_{b\text{-jets}} \geq 3$  except for the compressed region which requires a tighter selection of  $N_{b\text{-jets}} \geq 4$ . Meanwhile, it is also nice to see a reinforcement of physics that we expect to see from this signal model as in the behavior of  $M_{\text{J}}^{\Sigma}$  where the boosted

regions towards the bottom-right (a larger mass splitting) prefer a tighter cut on the mass of the four leading reclustered jets than the compressed regions towards the top-left (a smaller mass splitting). This is consistent with reclustered jets containing more merged decays and accidental substructure in a boosted topology.

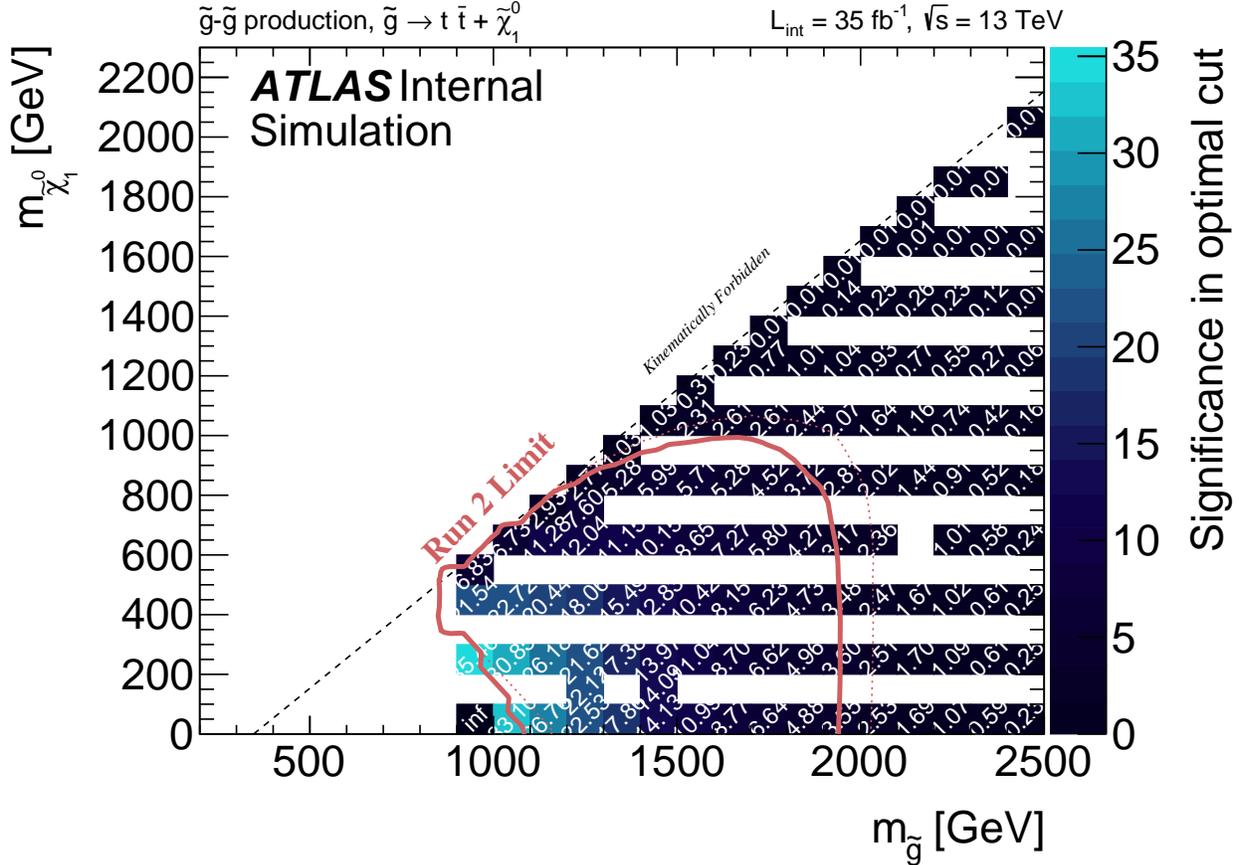


Figure 7.8: Significance of optimal cuts (fig. 7.9) for each grid point in the Gtt 0-lepton channel. The grid represents all the mass points for the  $\tilde{g}$ - $\tilde{\chi}_1^0$  pair with the mass of the gluino on the  $x$ -axis in GeV and the mass of the neutralino on the  $y$ -axis in GeV. This optimization was performed at an assumed total integrated luminosity of  $35 \text{ fb}^{-1}$ . The Run-2 limit from my first paper on this search at  $3.6 \text{ fb}^{-1}$  is overlaid in red [2]. The significance reported is an estimate of the CLs method using `BinomialExpZ` of `RooStats` [13, 14]. This significance represents the discriminating power the optimal cut has in each mass point.

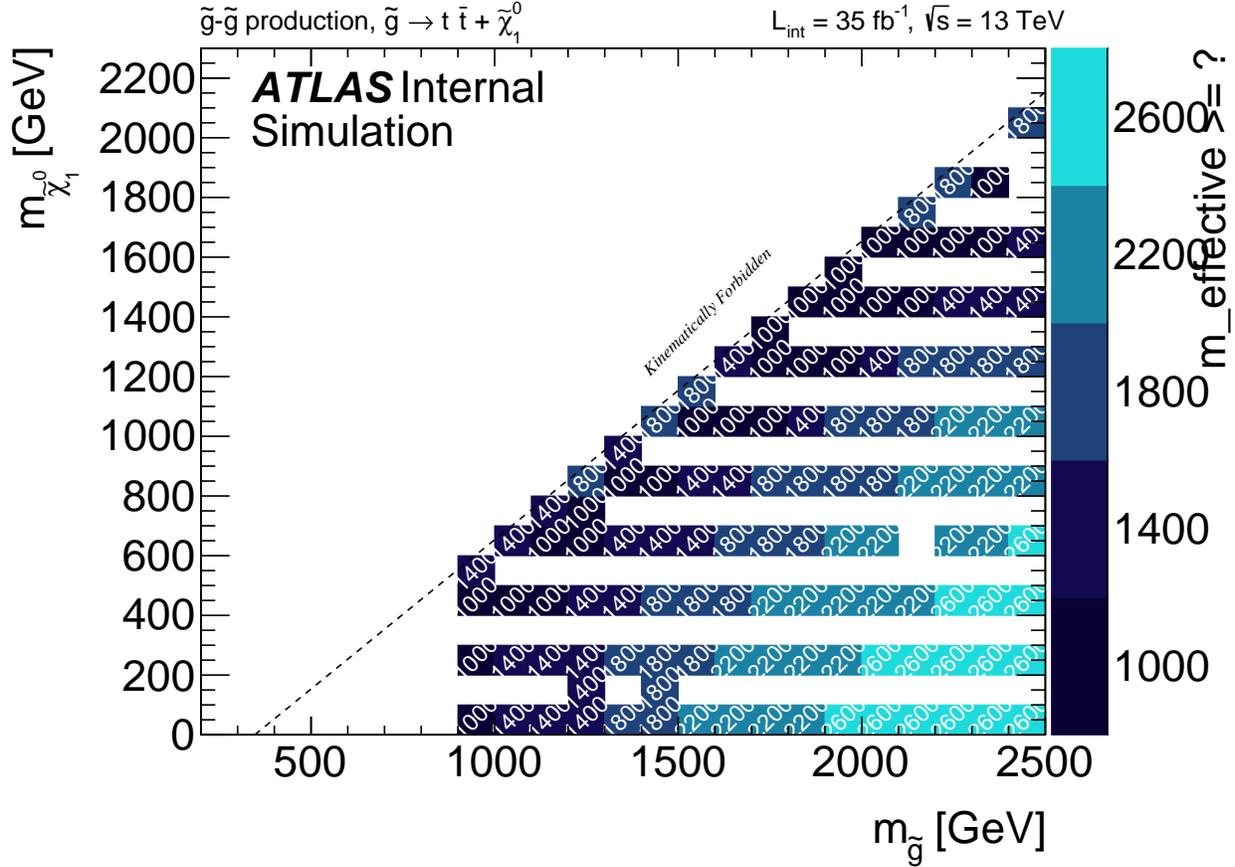


Figure 7.9: Optimal cut values for each grid point in the Gtt 0-lepton channel. The grid represents all the mass points for the  $\tilde{g}\text{-}\tilde{\chi}_1^0$  pair with the mass of the gluino on the  $x$ -axis in GeV and the mass of the neutralino on the  $y$ -axis in GeV. This optimization was performed at an assumed total integrated luminosity of  $35 \text{ fb}^{-1}$ . The kinematic variables were scanned over the values defined in table 7.2, a significance for each combination of selections was computed, and the selection that provided maximum discriminating power in each mass point is plotted. Each plot represents the cut applied to the kinematic variable (a)  $m_{\text{eff}}$ , (b)  $M_J^\Sigma$ , (c)  $E_T^{\text{miss}}$ , (d)  $N_{b\text{-jets}}$ .

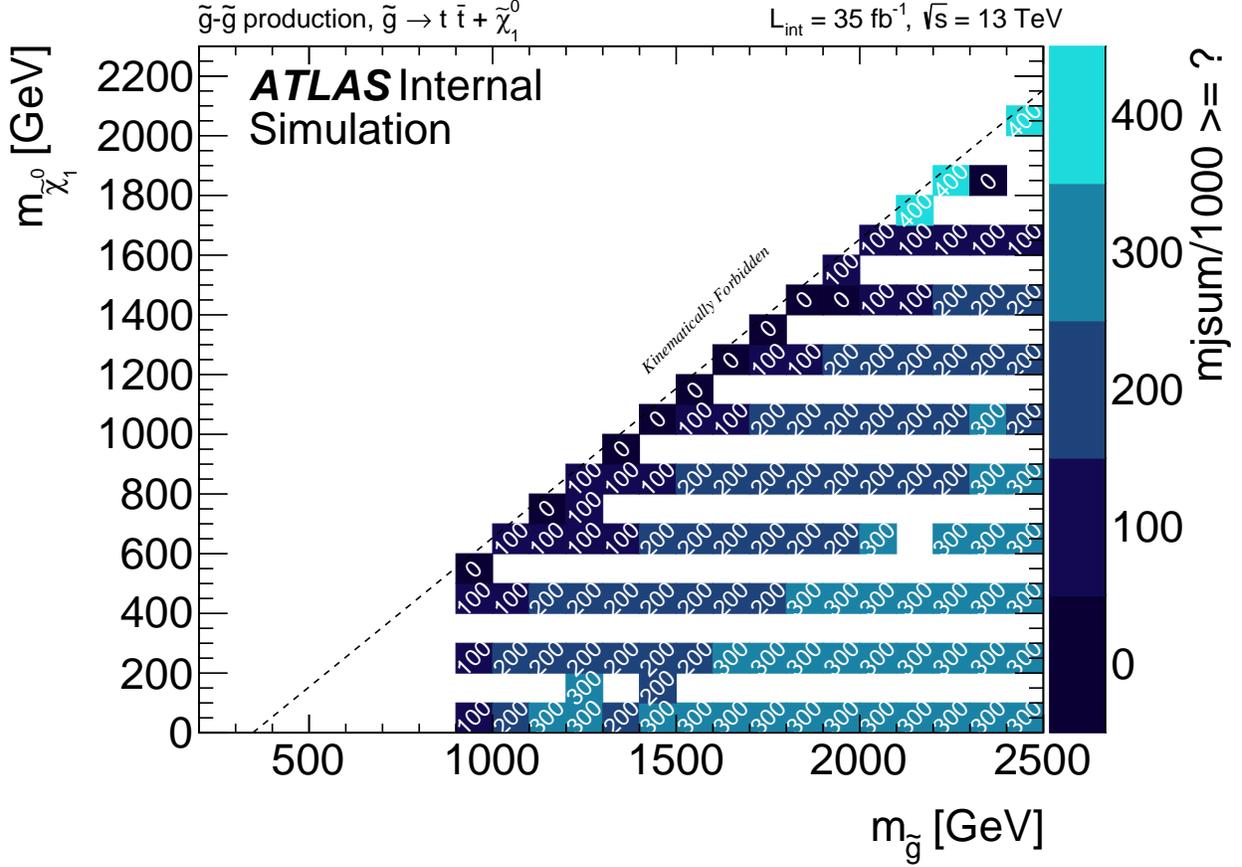


Figure 7.9: Optimal cut values for each grid point in the Gtt 0-lepton channel. The grid represents all the mass points for the  $\tilde{g}\text{-}\tilde{\chi}_1^0$  pair with the mass of the gluino on the  $x$ -axis in GeV and the mass of the neutralino on the  $y$ -axis in GeV. This optimization was performed at an assumed total integrated luminosity of  $35 \text{ fb}^{-1}$ . The kinematic variables were scanned over the values defined in table 7.2, a significance for each combination of selections was computed, and the selection that provided maximum discriminating power in each mass point is plotted. Each plot represents the cut applied to the kinematic variable (a)  $m_{\text{eff}}$ , (b)  $M_J^\Sigma$ , (c)  $E_T^{\text{miss}}$ , (d)  $N_{b\text{-jets}}$ .

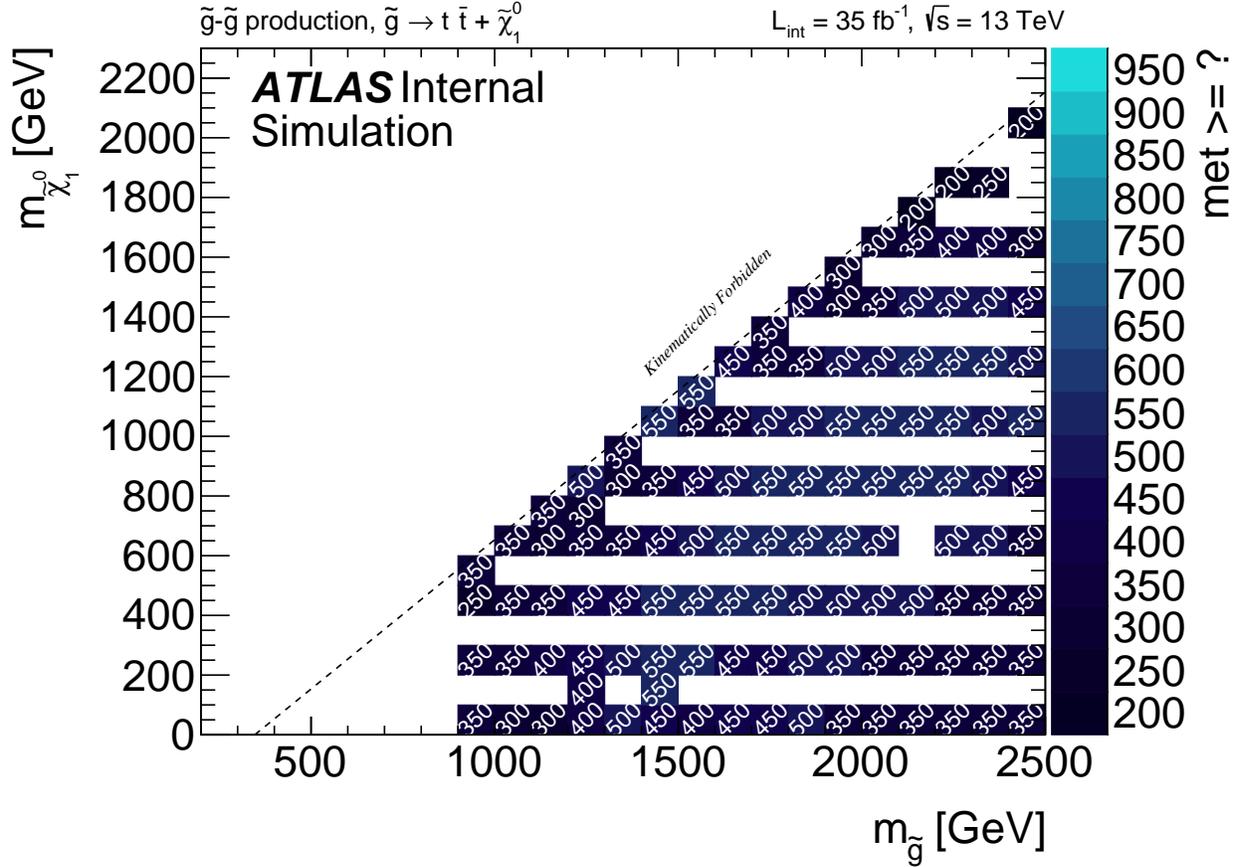


Figure 7.9: Optimal cut values for each grid point in the Gtt 0-lepton channel. The grid represents all the mass points for the  $\tilde{g}\text{-}\tilde{\chi}_1^0$  pair with the mass of the gluino on the  $x$ -axis in GeV and the mass of the neutralino on the  $y$ -axis in GeV. This optimization was performed at an assumed total integrated luminosity of  $35 \text{ fb}^{-1}$ . The kinematic variables were scanned over the values defined in table 7.2, a significance for each combination of selections was computed, and the selection that provided maximum discriminating power in each mass point is plotted. Each plot represents the cut applied to the kinematic variable (a)  $m_{\text{eff}}$ , (b)  $M_{\mathcal{J}}^{\Sigma}$ , (c)  $E_{\text{T}}^{\text{miss}}$ , (d)  $N_{b\text{-jets}}$ .

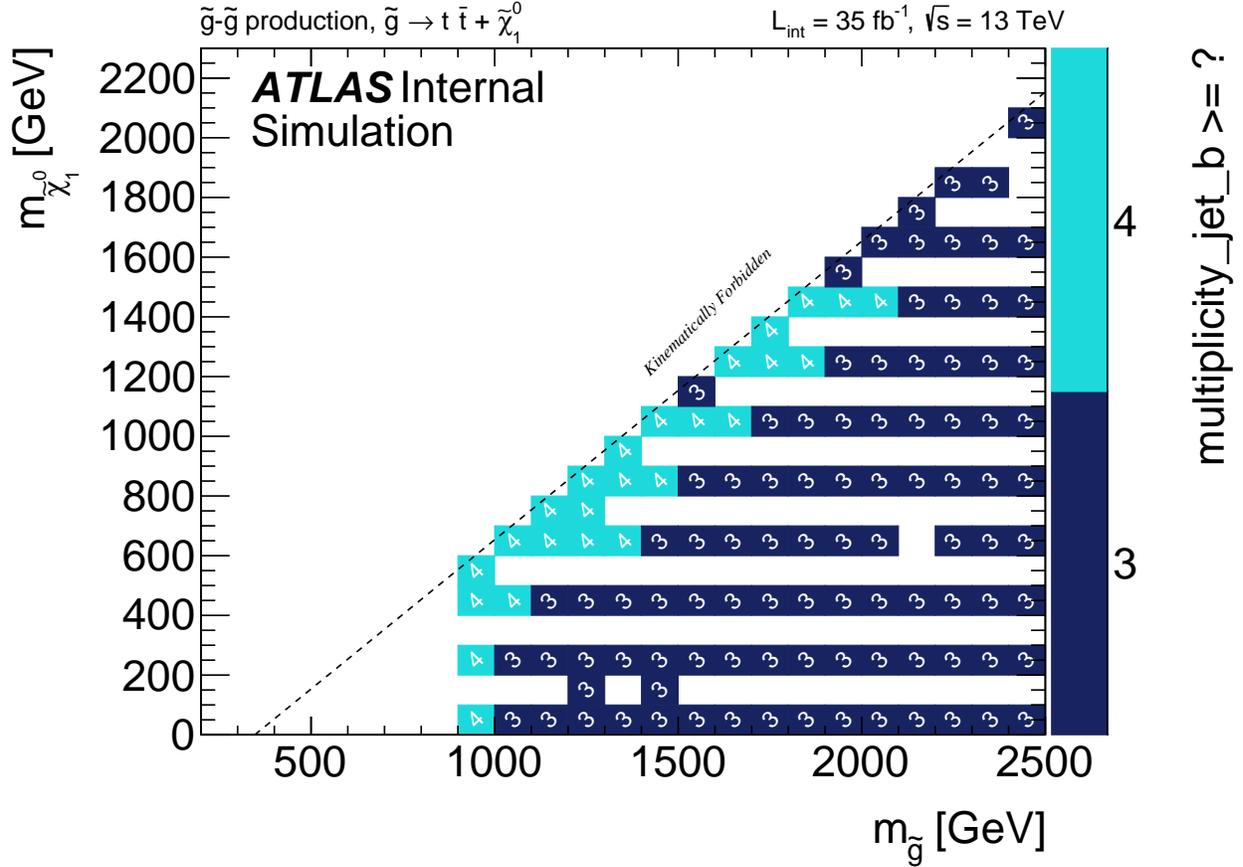


Figure 7.9: Optimal cut values for each grid point in the Gtt 0-lepton channel. The grid represents all the mass points for the  $\tilde{g}\text{-}\tilde{\chi}_1^0$  pair with the mass of the gluino on the  $x$ -axis in GeV and the mass of the neutralino on the  $y$ -axis in GeV. This optimization was performed at an assumed total integrated luminosity of  $35 \text{ fb}^{-1}$ . The kinematic variables were scanned over the values defined in table 7.2, a significance for each combination of selections was computed, and the selection that provided maximum discriminating power in each mass point is plotted. Each plot represents the cut applied to the kinematic variable (a)  $m_{\text{eff}}$ , (b)  $M_J^\Sigma$ , (c)  $E_T^{\text{miss}}$ , (d)  $N_{b\text{-jets}}$ .

## Signal Regions

From these optimization results, three **SRs** have been designed for the 0-lepton channel of the Gtt analysis, summarized in table 7.4. All 0-lepton **SRs** veto events with one or more reconstructed signal leptons, and require at least 3  $b$ -tagged **small- $R$**  jets. **SR-B**<sup>12</sup> targets the dramatic events in the region of high mass splitting between the  $\tilde{g}$  and  $\tilde{\chi}_1^0$ . This region places large requirements on the  $m_{\text{eff}}$  and the largest requirement on  $M_{\mathcal{J}}^{\Sigma}$  out of all Gtt **SRs**. **SR-M**<sup>13</sup> targets the bulk region of the signal grid, requiring large amounts of  $E_{\text{T}}^{\text{miss}}$ ,  $m_{\text{eff}}$ , and  $M_{\mathcal{J}}^{\Sigma}$ . **SR-C**<sup>14</sup> targets the “near-diagonal” region of the signal grid, where the small mass splitting between the  $\tilde{g}$  and  $\tilde{\chi}_1^0$  results in final states with jets which are typically softer than those found in the other regions of the signal plane. **SR-C** also has an additional  $b$ -tagged jet to further reject background in this portion of phase space.

The significance at each signal grid point for the optimal **SR** is plotted in fig. 7.10b. The three **SRs** each target distinct regions of phase space (fig. 7.10a), with little degradation in performance from the optimal set of cuts at each mass point (fig. 7.11).

## Control Regions

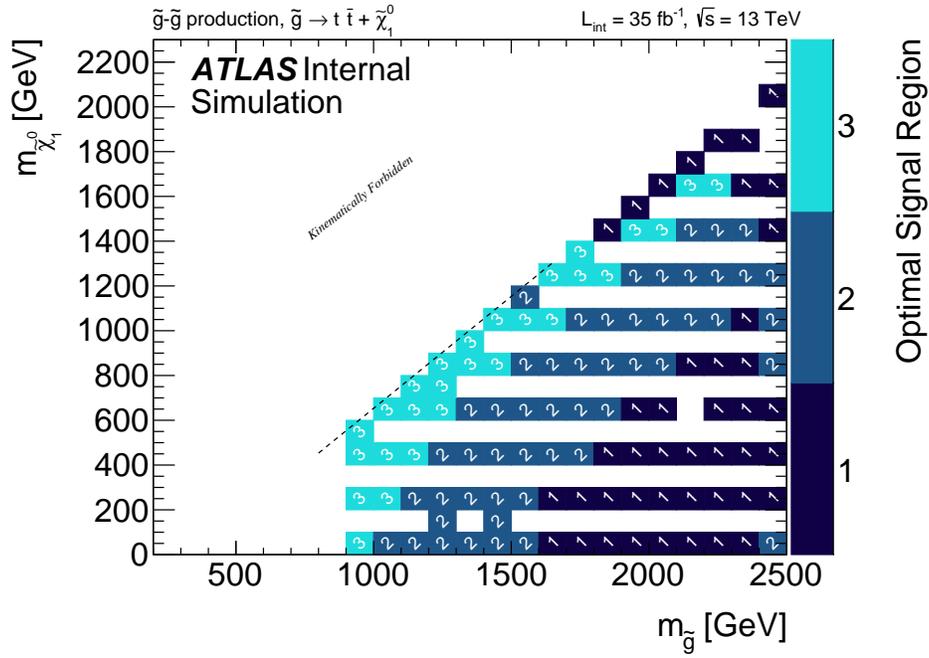
**CRs** are also defined for each **SR** in table 7.4. As the background is expected to be dominated entirely by semi-lepton  $t\bar{t}$  events, due to the  $b$ -jet requirement and the  $E_{\text{T}}^{\text{miss}}$  requirement, following the strategy of the 2015 analysis we define only a single  $t\bar{t}$  **CR** for the analysis and will take all the remaining backgrounds directly from simulation [16]. Since the background is expected to be semi-leptonic  $t\bar{t}$ , a single lepton **CR** strategy is used. An upper cut on  $m_{\text{T}}$

---

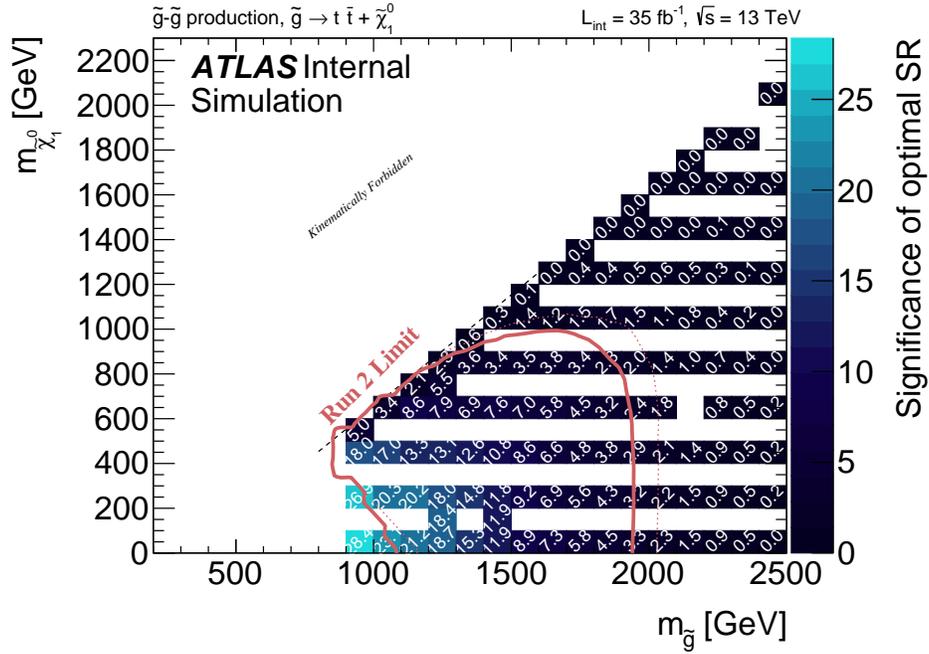
<sup>12</sup> “B” for boosted.

<sup>13</sup> “M” for moderate-boost.

<sup>14</sup> “C” for compressed.



(a) Optimal SR for each grid point in the Gtt 0-lepton channel. In this figure, regions 1-2-3 correspond to B-M-C in the text.



(b) Significance of optimal SR for each grid point in the Gtt 0-lepton channel.

Figure 7.10: (a) Optimal SR and (b) significance for the optimal region for all points of the grid in the Gtt 0-lepton analysis.

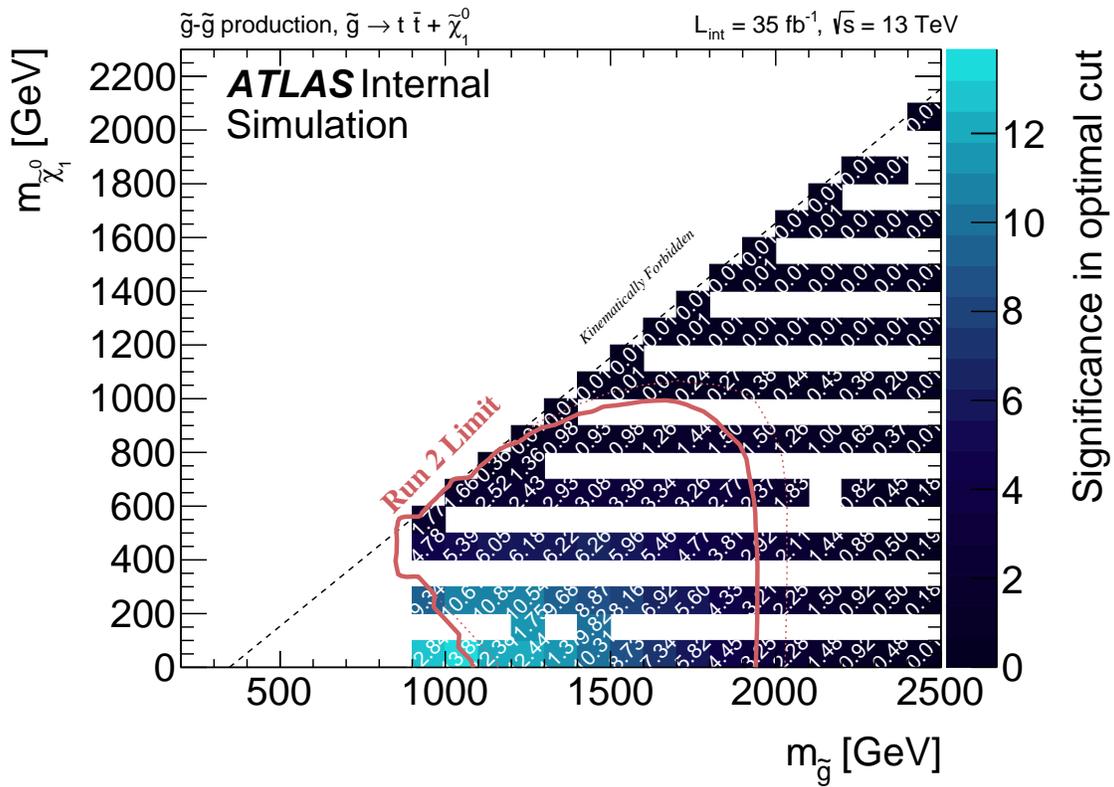


Figure 7.11: Significance of each SR at each grid point of the Gtt 0-lepton analysis.

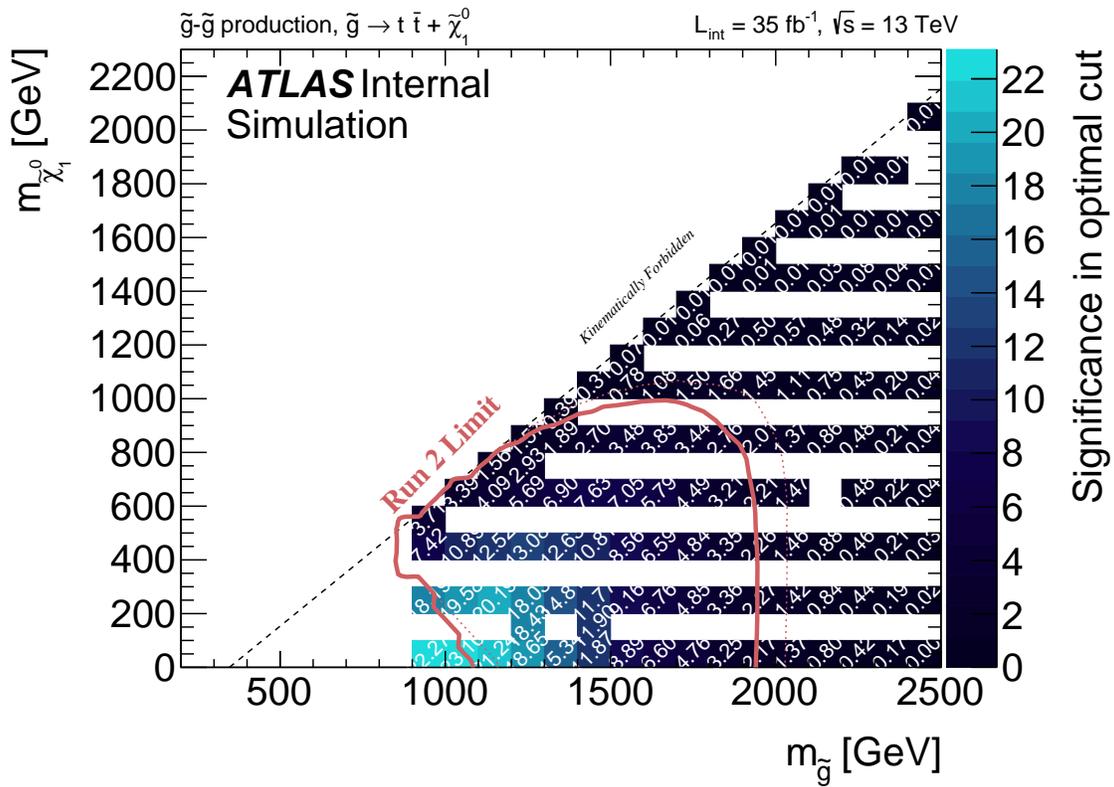
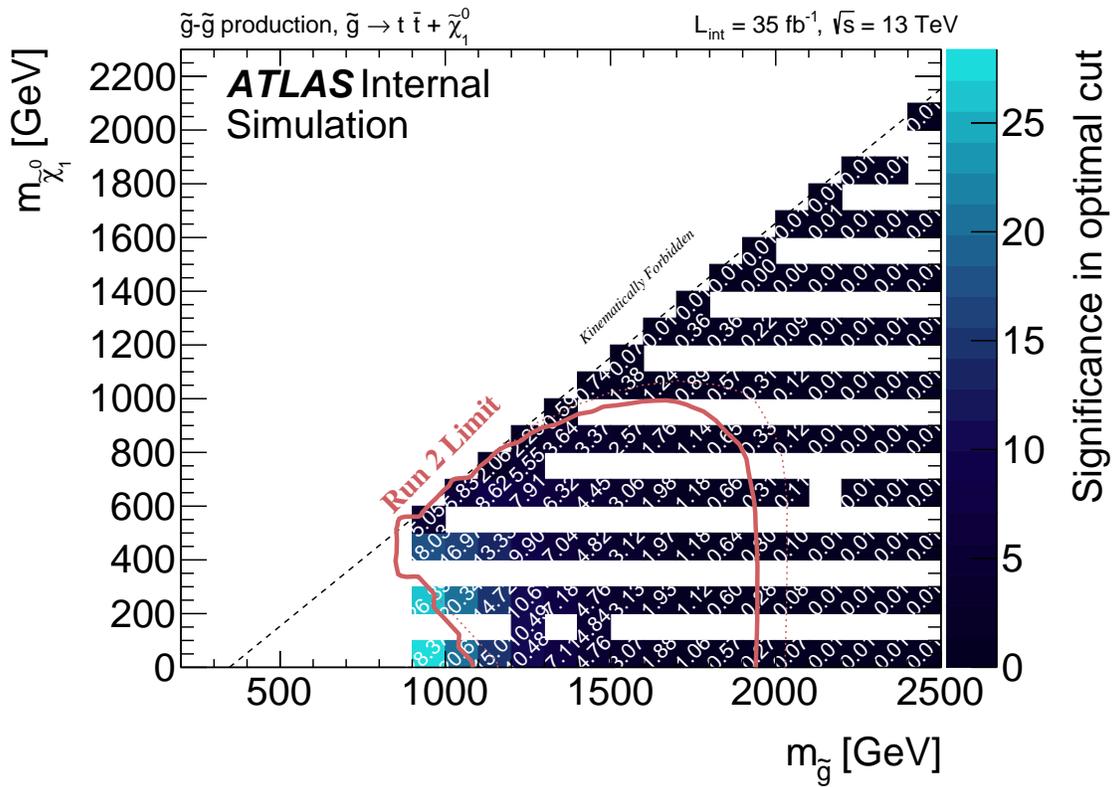


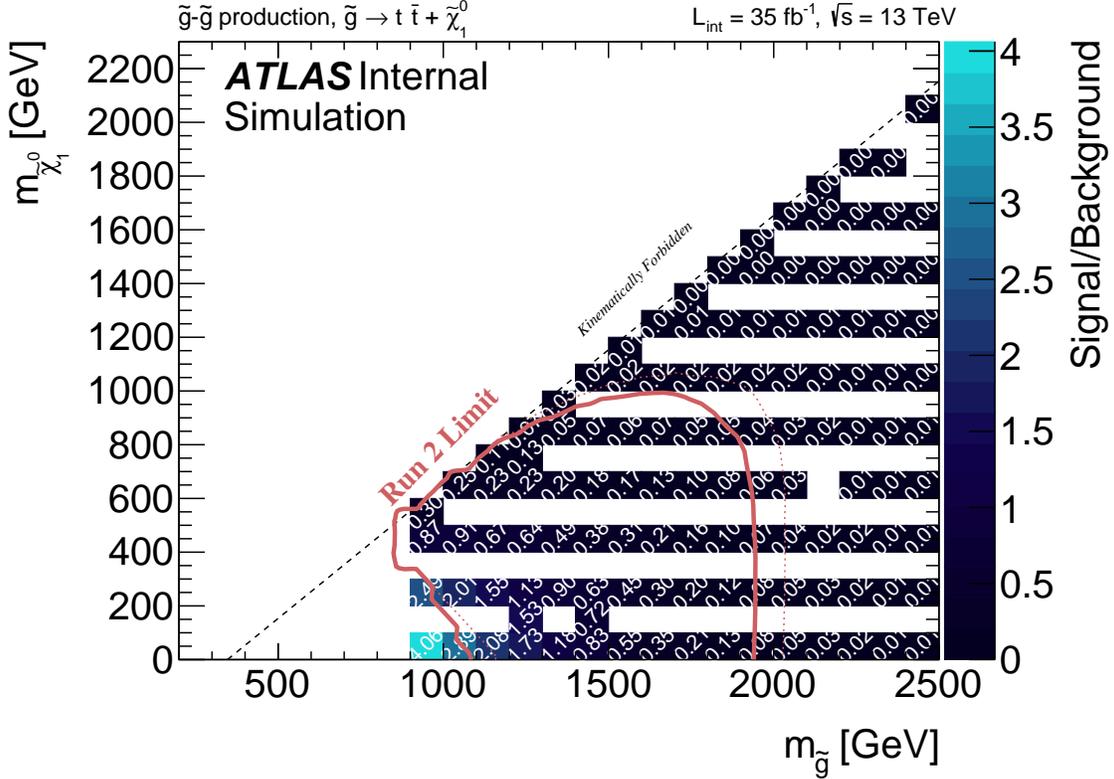
Figure 7.11: Significance of each SR at each grid point of the Gtt 0-lepton analysis.



(c) SR-C

Figure 7.11: Significance of each SR at each grid point of the Gtt 0-lepton analysis.

is enforced to prevent overlap with the Gtt SRs. All jet number related cuts are lowered by 1 between each SR and the corresponding CR (since a lepton replaces a jet), and cuts on  $E_T^{\text{miss}}$  and  $m_{\text{eff}}$  and  $m_{T,\text{min}}^{b\text{-jets}}$  are lowered to ensure sufficient statistics (a minimum of 15 expected events) in the CRs. The signal contamination in each CR is expected to be very small, and is shown in fig. 7.12.

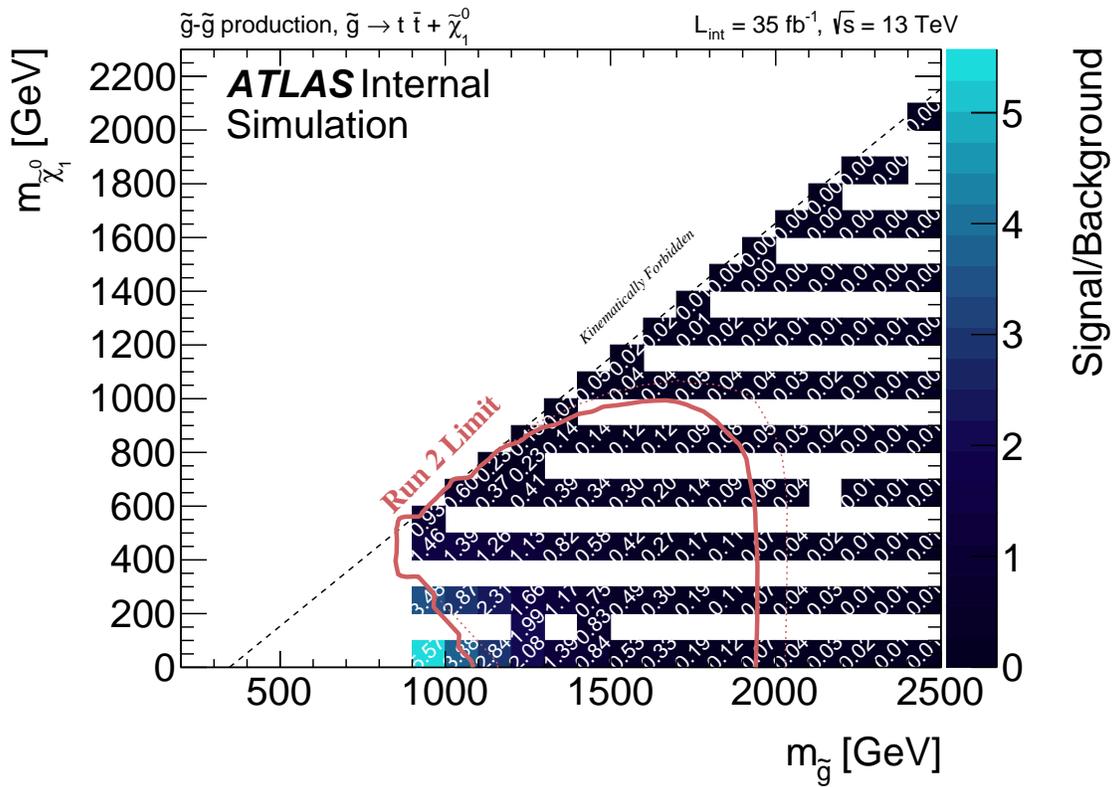


(a) CR-B

Figure 7.12: Signal contamination of each CR at each grid point of the Gtt 0-lepton analysis.

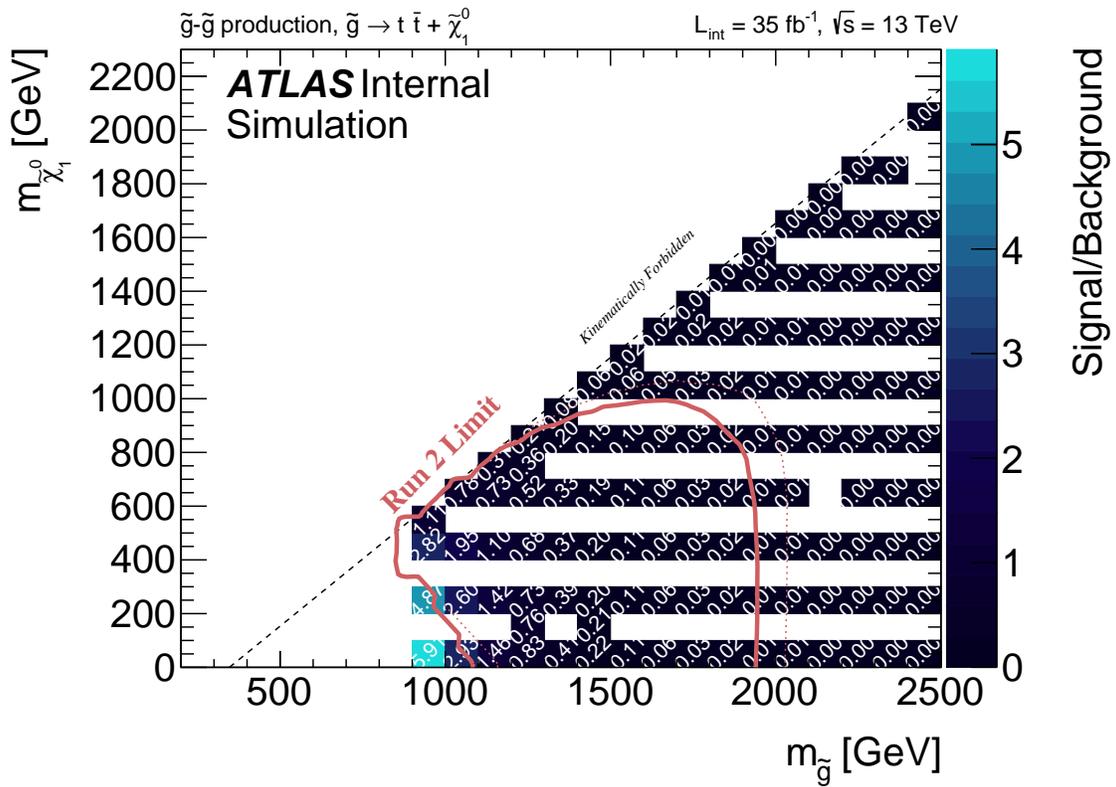
## Validation Regions

Table 7.4 also defines one VRs for each SR. The VR category validates the extrapolation from the 1L CR to the 0L SR. This is done by inverting the cut on  $M_J^\Sigma$ , which is shown to be well modeled in section 7.3. Signal contamination plots for the VRs are shown in figs. 7.13



(b) CR-M

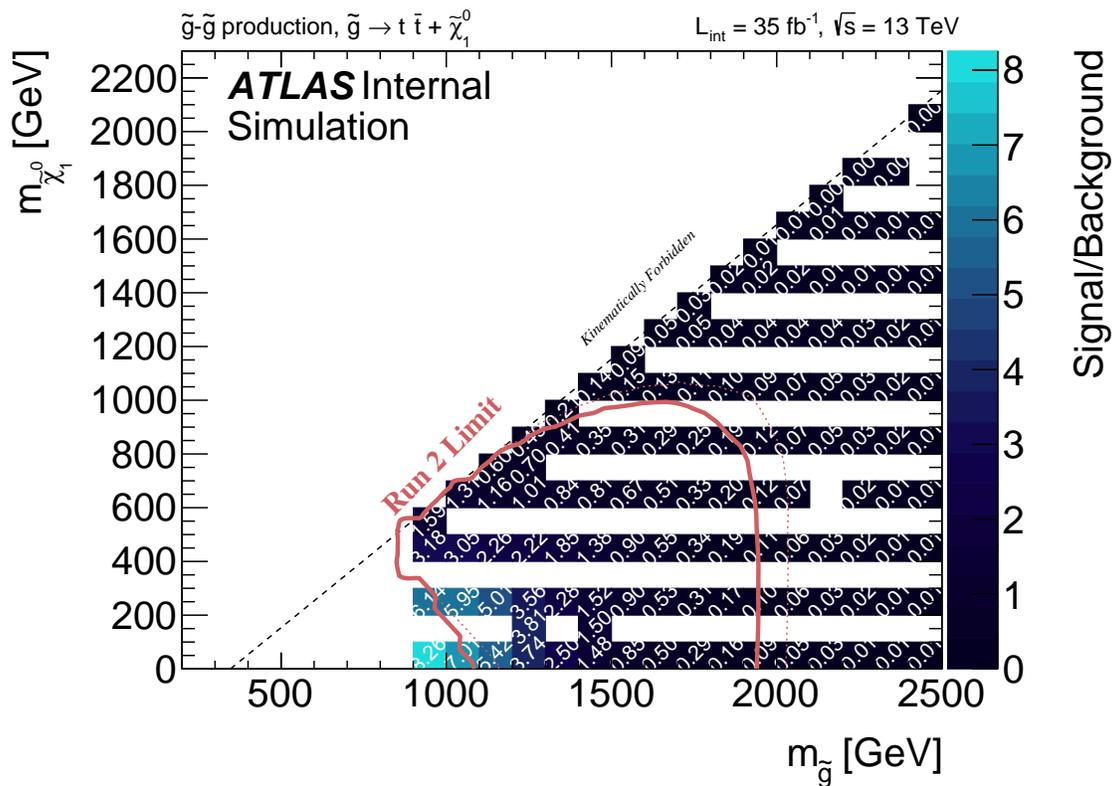
Figure 7.12: Signal contamination of each CR at each grid point of the Gtt 0-lepton analysis.



(c) CR-C

Figure 7.12: Signal contamination of each CR at each grid point of the Gtt 0-lepton analysis.

and 7.14. Both are rather low.

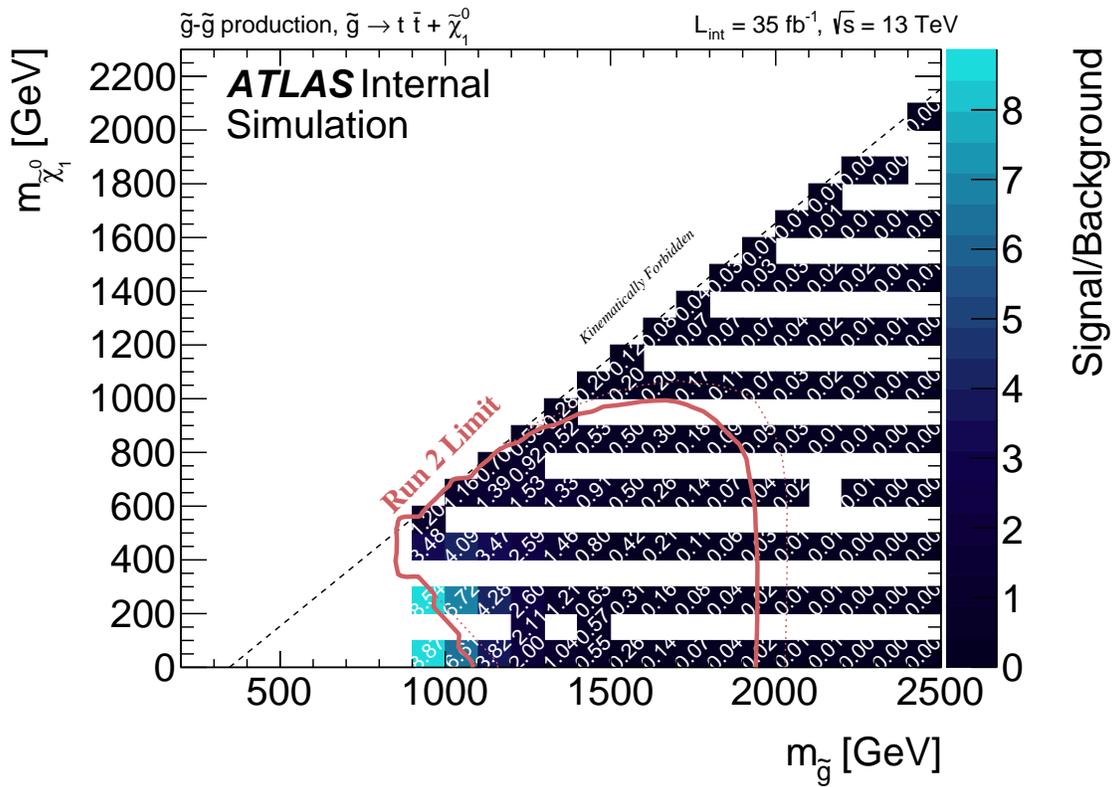


(a) VR-B

Figure 7.13: Signal contamination of each VR at each grid point of the Gtt 0-lepton analysis.

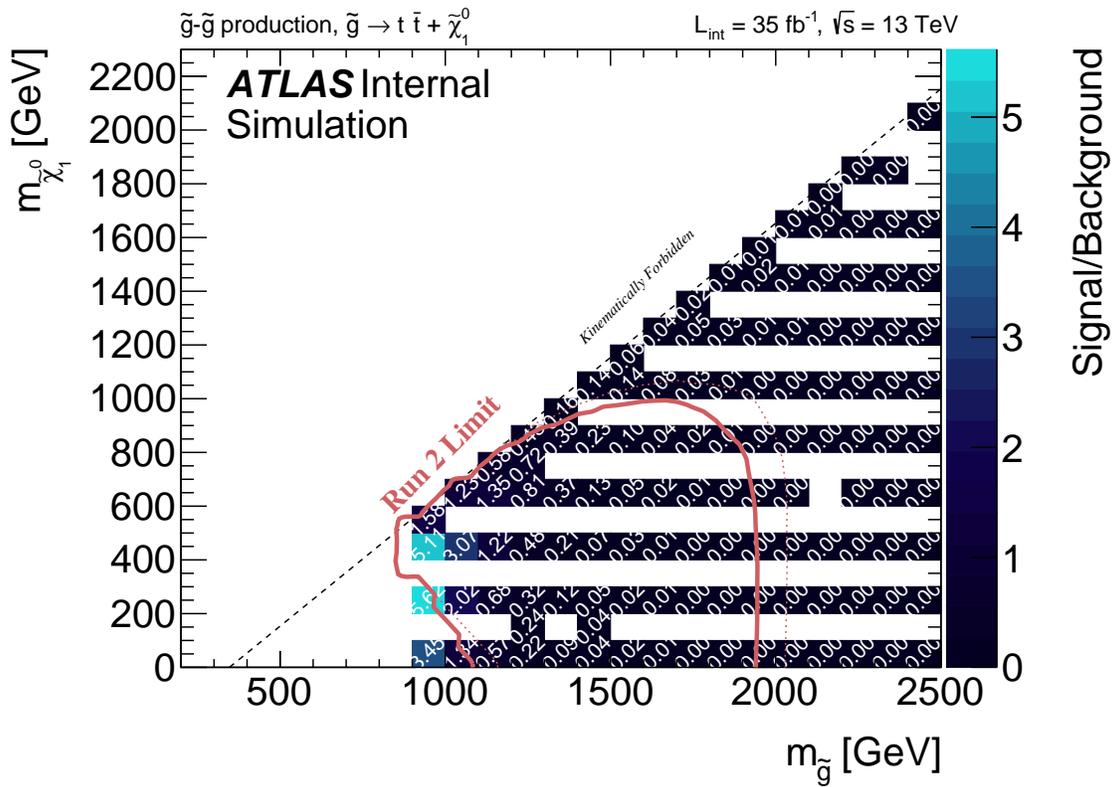
## Background Composition

The composition of the background in the optimized Gtt-0L regions are all shown in ???. Three of these composition plots are shown for the 0-lepton boosted region in fig. 7.15 showing the heavy flavor composition of the  $t\bar{t}$  background. In all cases, the SR is dominated by  $t\bar{t}$  as expected, as are the CRs and VRs. The expected yields of the CRs and VRs are also sufficient. Note all the composition plots are pre-fit.



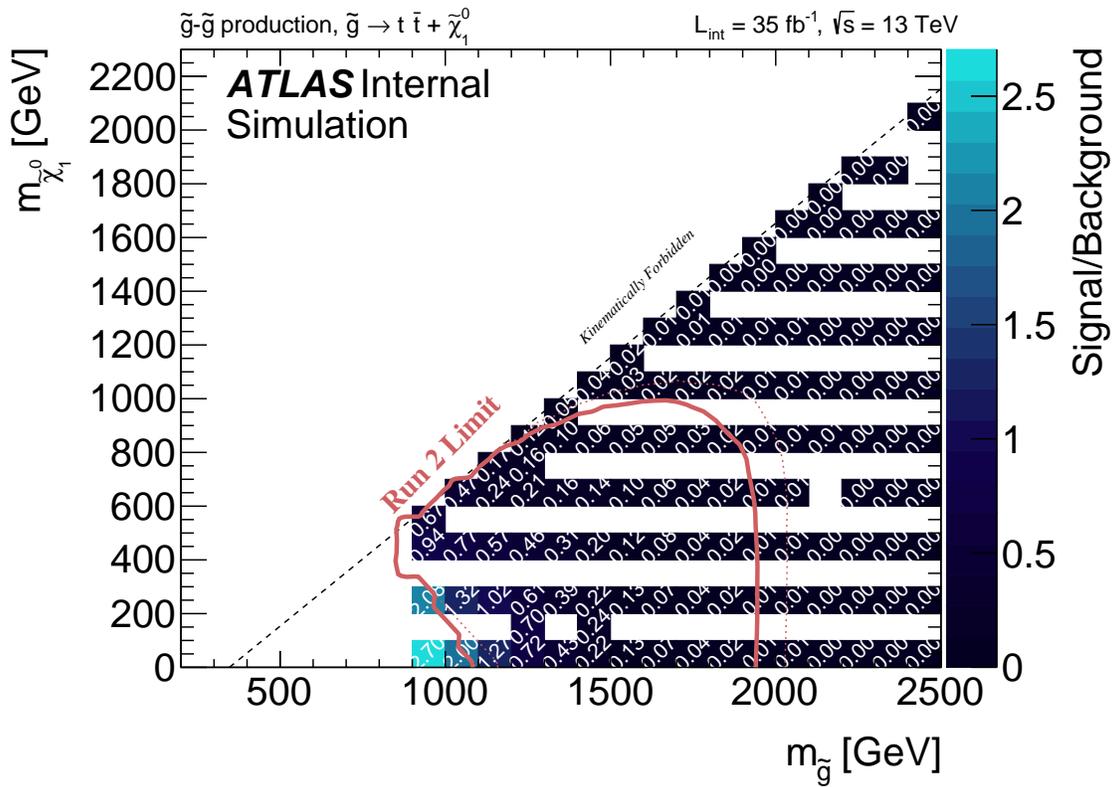
(b) VR-M

Figure 7.13: Signal contamination of each VR at each grid point of the Gtt 0-lepton analysis.



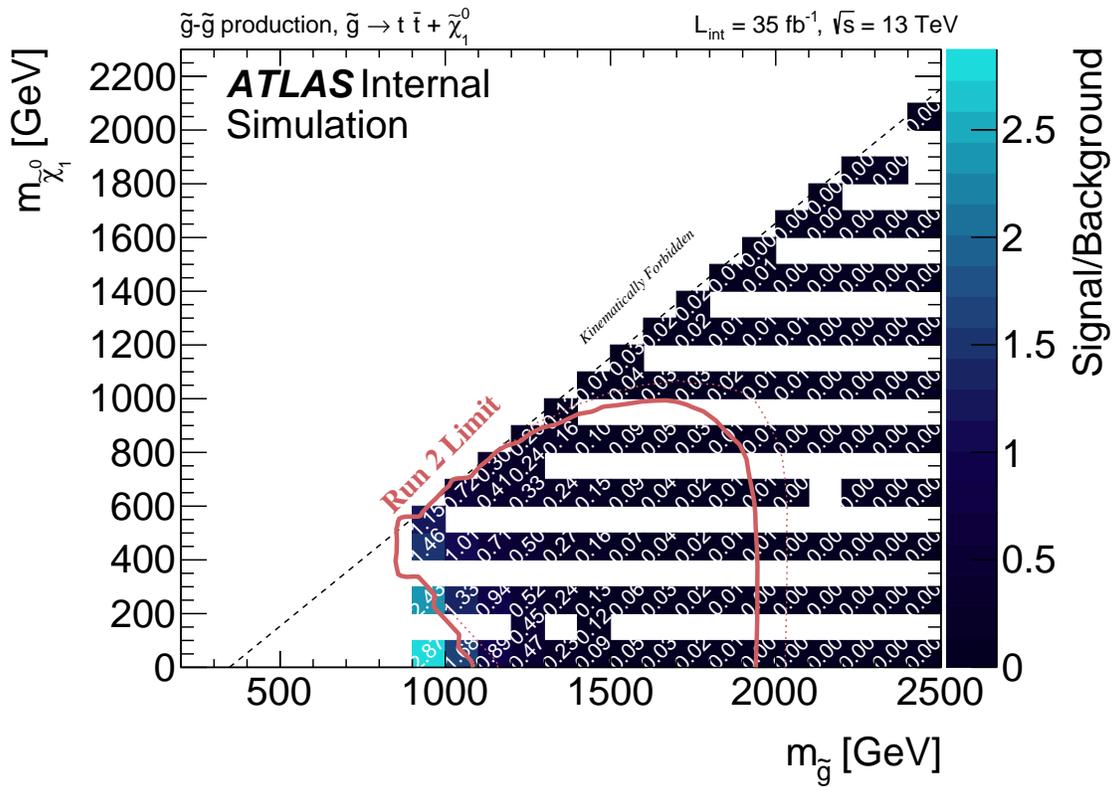
(c) VR-C

Figure 7.13: Signal contamination of each VR at each grid point of the Gtt 0-lepton analysis.



(a) VR-B

Figure 7.14: Signal contamination of each VR at each grid point of the Gtt 0-lepton analysis.



(b) VR-M

Figure 7.14: Signal contamination of each VR at each grid point of the Gtt 0-lepton analysis.

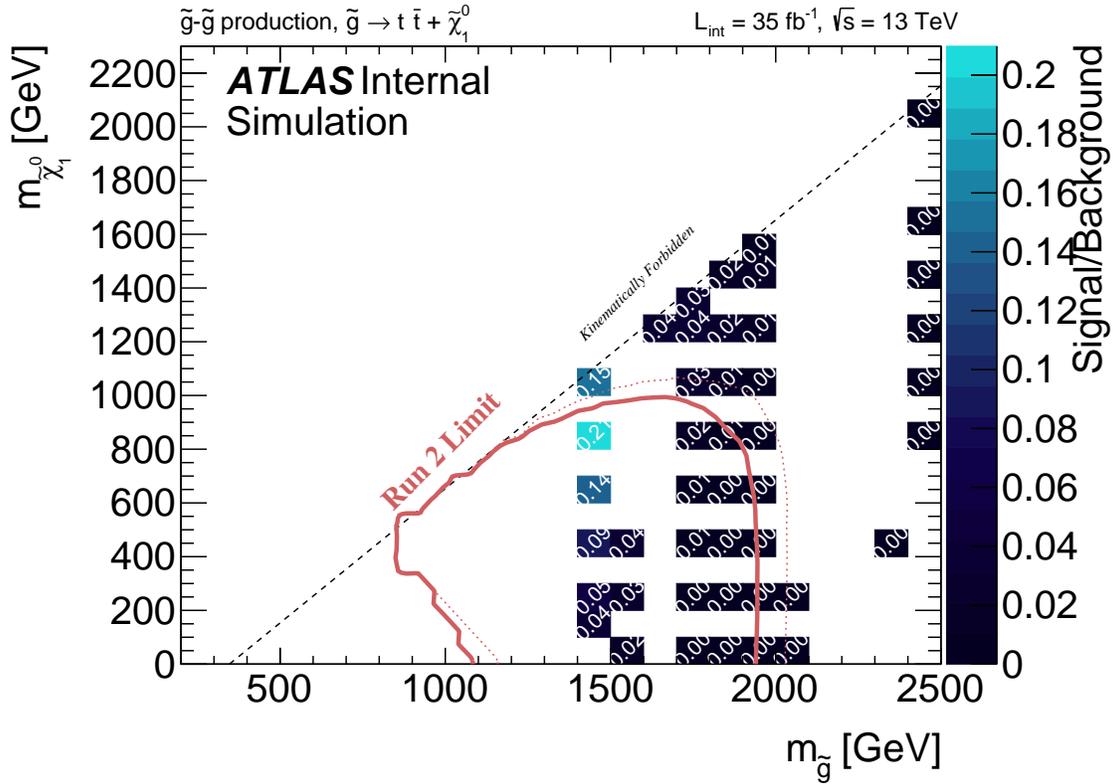


Figure 7.14: Signal contamination of each VR at each grid point of the Gtt 0-lepton analysis.

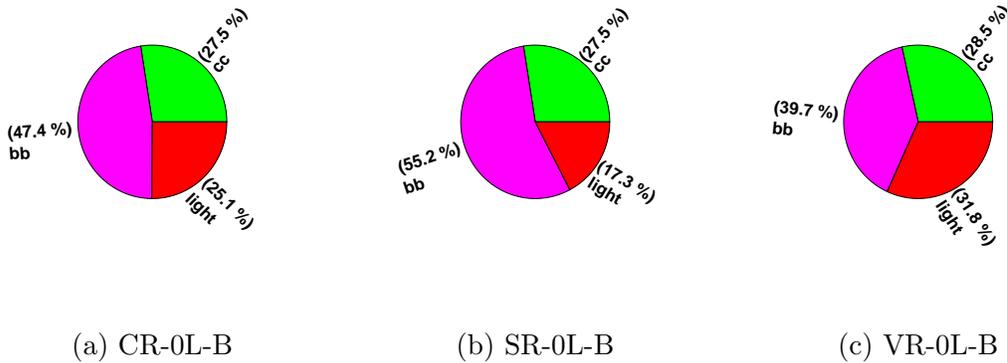


Figure 7.15: [9] Heavy flavor composition of the  $t\bar{t}$  component of the background in the optimized Gtt-0L boosted regions.

## $N-1$ Plots

Distributions of  $N-1$  plots for the kinematic variables used in the optimization of signal regions are all shown in ???. Each of the variables shows significant discrimination power for the signals shown. An example plot is shown for the 0-lepton boosted signal region in fig. 7.16 for the total jet mass variable. This plot is made by applying all of the  $N$  selections in the 0-lepton boosted signal region described in table 7.4 except for the total jet mass variable  $M_J^\Sigma$ , hence the name  $N-1$  plot.

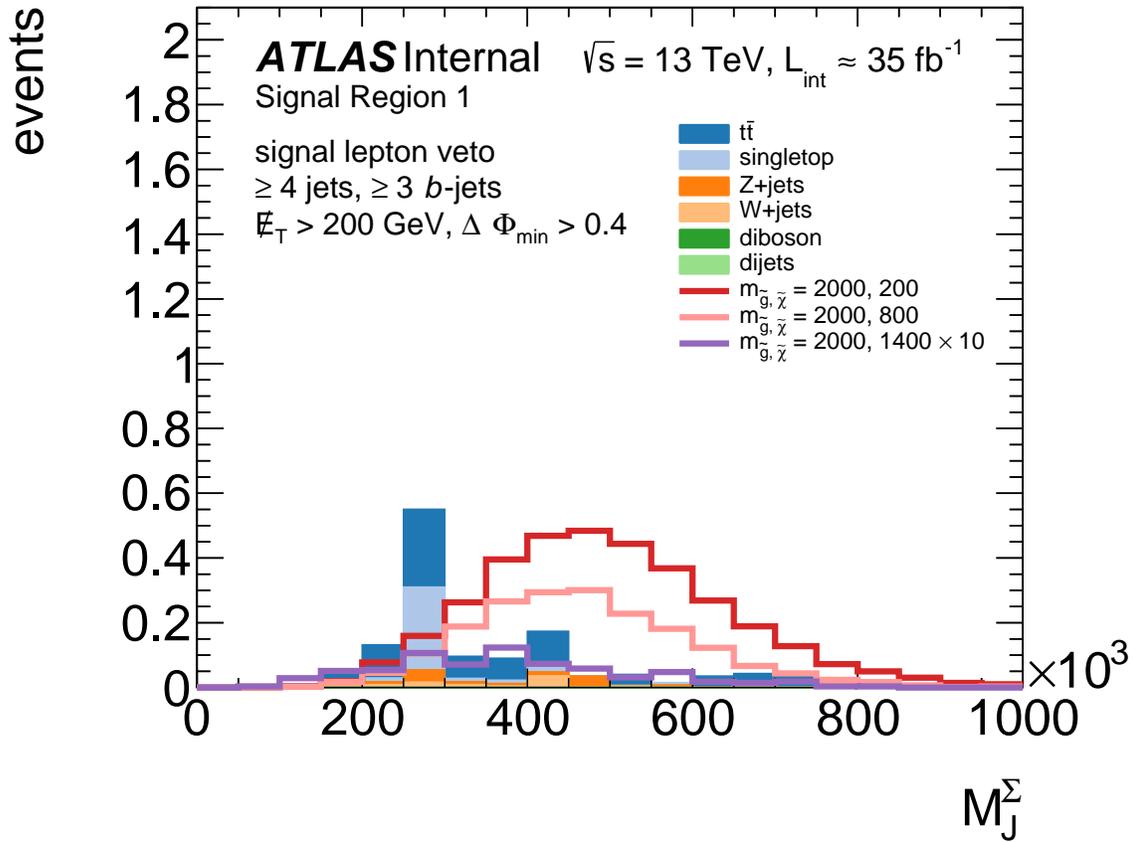


Figure 7.16:  $N-1$  plot of the total jet mass variable in the 0-lepton boosted region. The background is stacked in the histogram and three signal points are overlaid to show the shape comparisons, all normalized to  $35 \text{ fb}^{-1}$  except the highest mass signal curve which has been scaled up by a factor of 10.

#### 7.4.4 Gtt-1L Optimization

The Gtt 1-lepton optimization is very similar to 0-lepton. The main change is the lowering of the  $N_{\text{jet}}$  requirement (accounting for the extra lepton) and the addition of another transverse mass cut  $m_{\text{T}}$  which is inverted to create the CRs. To summarize, the variables used in this optimization are summarized in table 7.3.

	inclusive		step
	start	stop	size
$N_{\text{jet}}$	5	9	1
$N_{b\text{-jets}}$	3	6	1
$E_{\text{T}}^{\text{miss}}$	200 GeV	1000 GeV	50 GeV
$m_{\text{eff}}$	800 GeV	3000 GeV	100 GeV
$M_{\text{J}}^{\Sigma}$	0 GeV	500 GeV	100 GeV
$m_{\text{T},\text{min}}^{b\text{-jets}}$	0 GeV	200 GeV	4 GeV0
$m_{\text{T}}$	0 GeV	300 GeV	50 GeV

Table 7.3: The 1-lepton optimization phase-space that was scanned over. This corresponded to 115,920 cuts. Each variable was scanned from “start” to “stop” in discrete step sizes. The starting and stopping values of the scan are inclusive.

Figure 7.17 shows the optimal significance possible with the best selected variables at each grid point; Figure 7.18 shows the optimal cuts on various variables selected at each mass point.

### Signal Regions

From these optimization results, three SRs have been designed for the 1-lepton channel of the Gtt analysis, summarized in table 7.5. All 1-lepton SRs require events with one or more

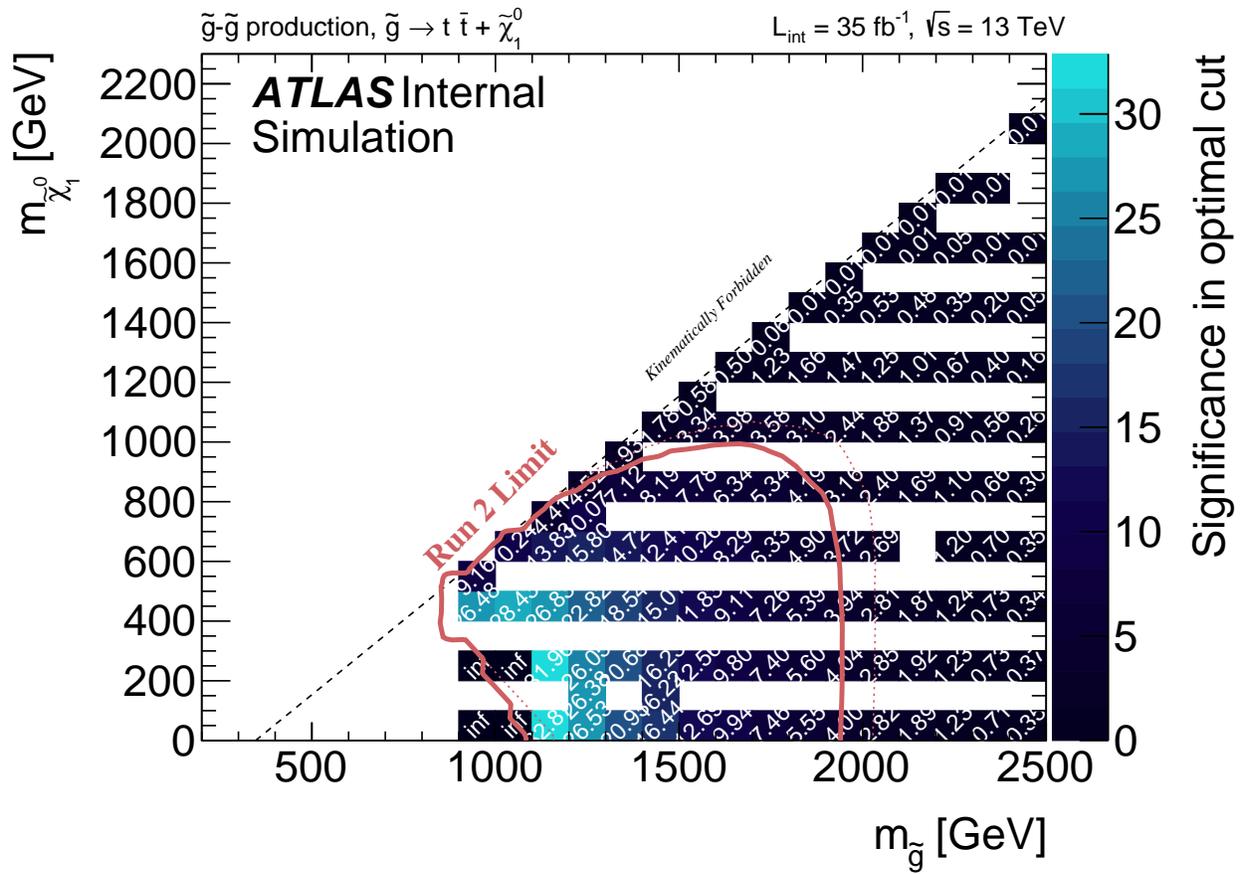
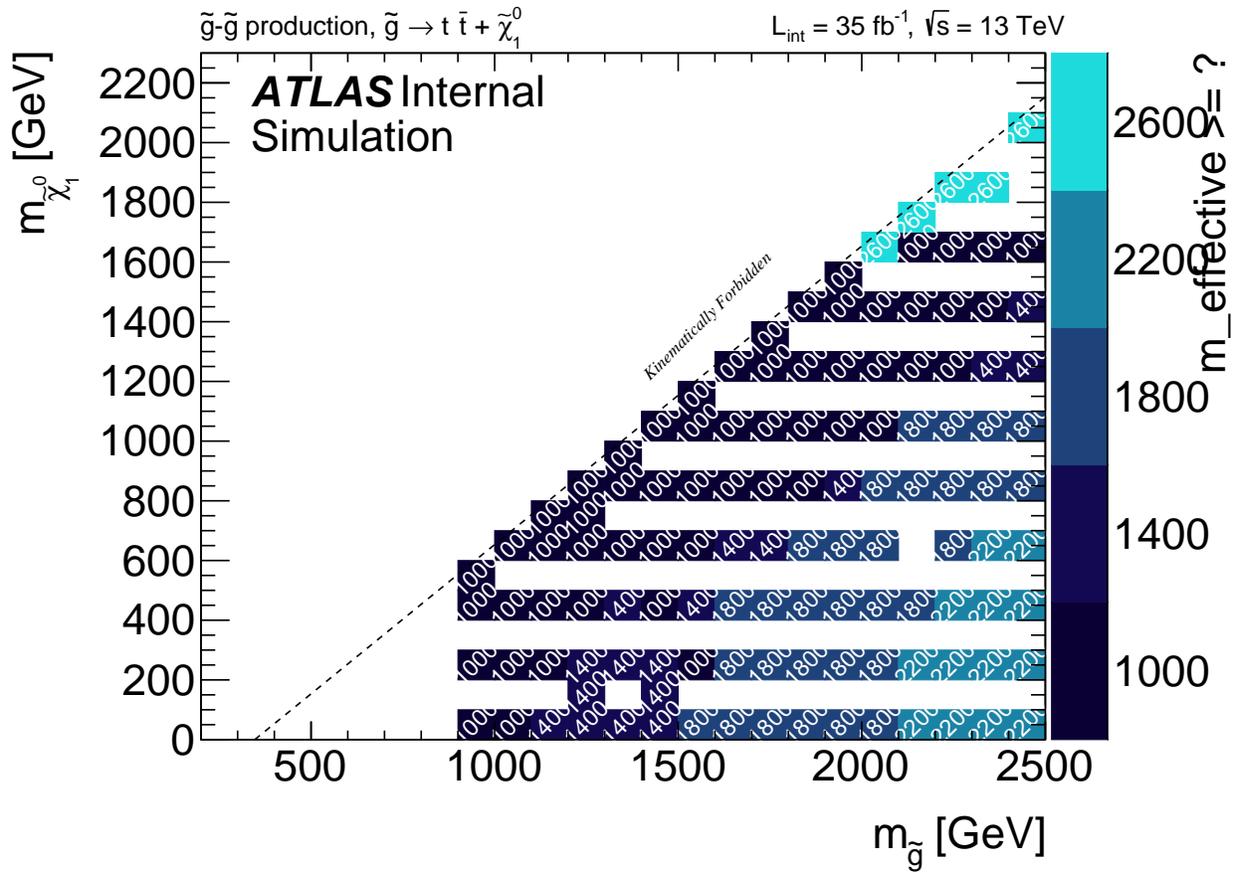
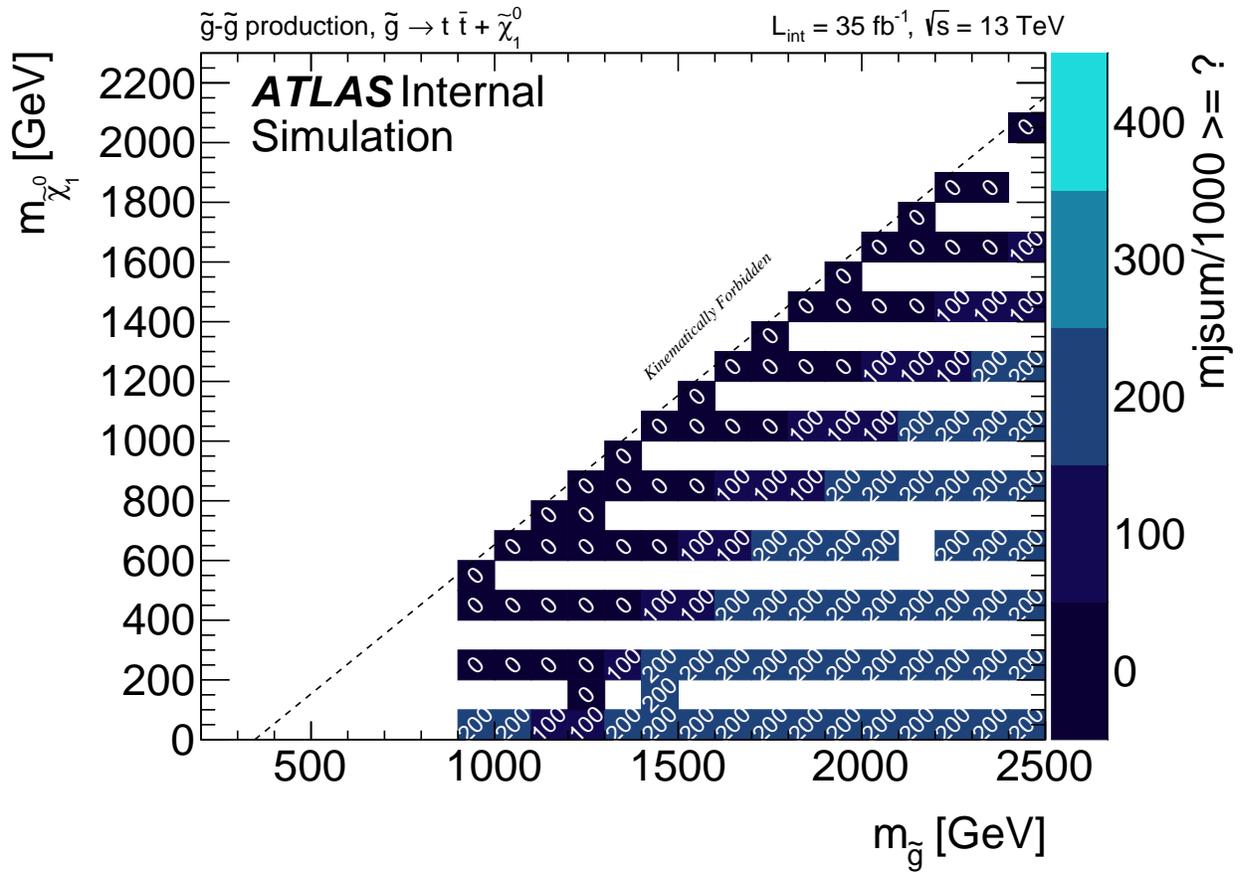


Figure 7.17: Significance of optimal cuts for each grid point in the Gtt 0-lepton channel.



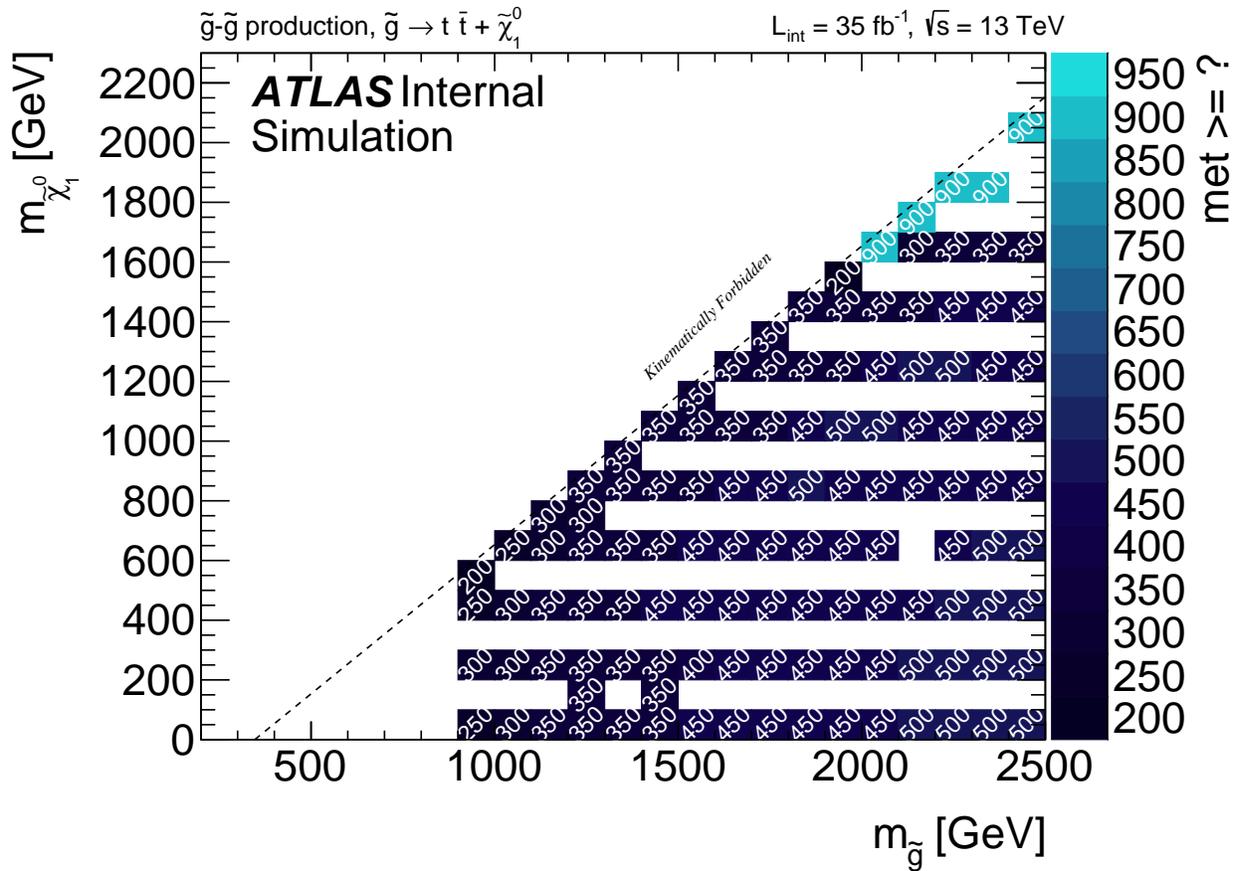
(a)  $m_{\text{eff}}$

Figure 7.18: Optimal cut values for each grid point in the Gtt 1-lepton channel.



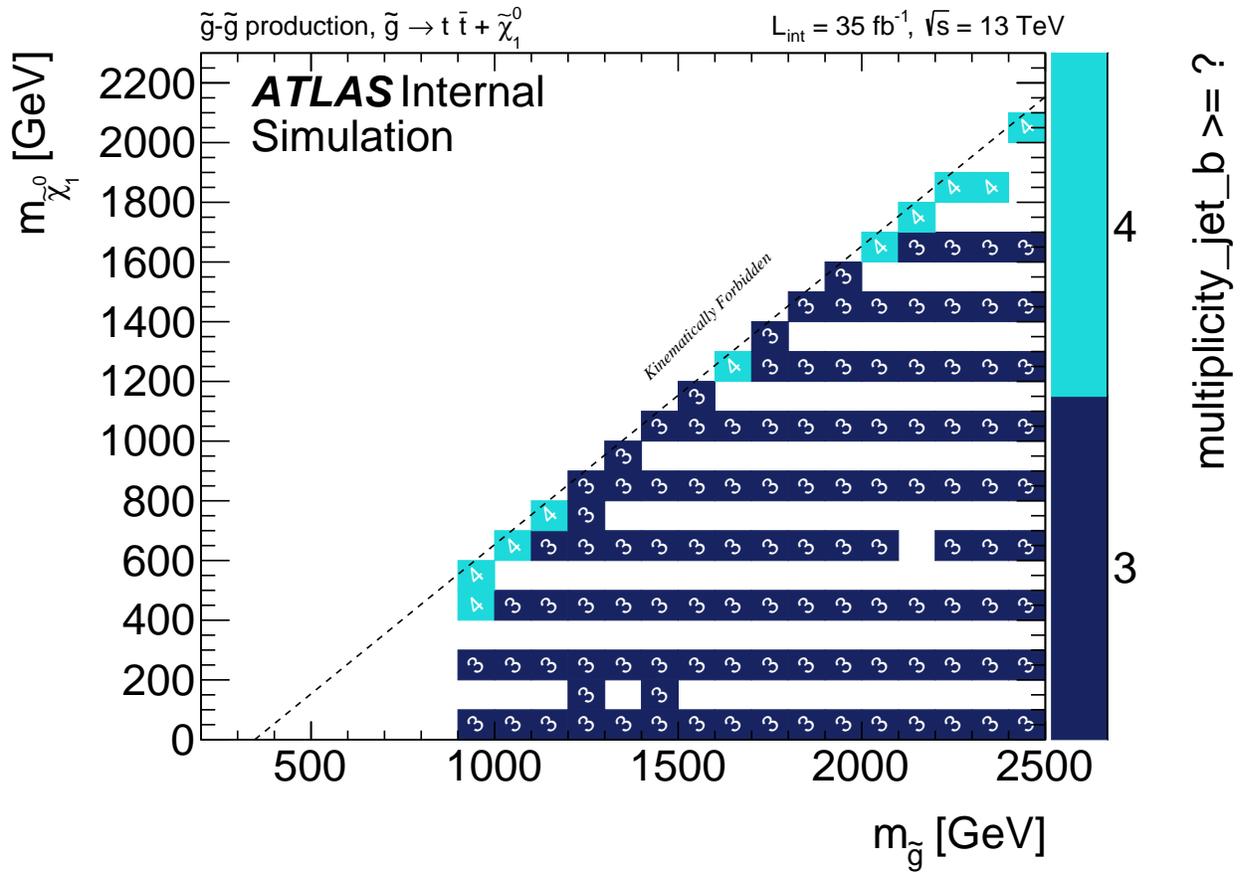
(b)  $M_J^\Sigma$

Figure 7.18: Optimal cut values for each grid point in the Gtt 1-lepton channel.



(c)  $E_T^{\text{miss}}$

Figure 7.18: Optimal cut values for each grid point in the Gtt 1-lepton channel.



(d)  $N_{b\text{-jets}}$

Figure 7.18: Optimal cut values for each grid point in the Gtt 1-lepton channel.

reconstructed signal leptons<sup>15</sup>, and require at least 3  $b$ -tagged *small- $R$*  jets.

*SR-B*<sup>16</sup> targets the dramatic events in the region of high mass splitting between the  $\tilde{g}$  and  $\tilde{\chi}_1^0$ . This region places large requirements on the  $m_{\text{eff}}$  and the largest requirement on  $M_J^\Sigma$  out of all Gtt *SRs*. The large energy splitting between the  $\tilde{g}$  and  $\tilde{\chi}_1^0$  means that most of the energy goes into the visible decay products of the event, the boosted top quarks. *SR-M*<sup>17</sup> targets the bulk region of the signal grid, requiring large amounts of  $E_T^{\text{miss}}$ ,  $m_{\text{eff}}$ , and  $M_J^\Sigma$ . *SR-C*<sup>18</sup> targets the “near-diagonal” region of the signal grid, where the small mass splitting between the  $\tilde{g}$  and  $\tilde{\chi}_1^0$  results in final states with jets which are typically softer than those found in the other regions of the signal plane as a larger fraction of the energy will be in the neutralinos, the invisible decay products, and so the visible decay products are less boosted and more resolved. *SR-C* also has an additional  $b$ -tagged jet to further reject background in this portion of phase space as the top quarks are more resolved and is more likely to have four isolated  $b$ -jets, rather than in a more boosted region where two of the  $b$ -jets might merge into a single jet.

The significance at each signal grid point for the optimal *SR* is plotted in fig. 7.19b. The three *SRs* each target distinct regions of phase space (fig. 7.19a), with little degradation in performance from the optimal set of cuts at each mass point (fig. 7.20).

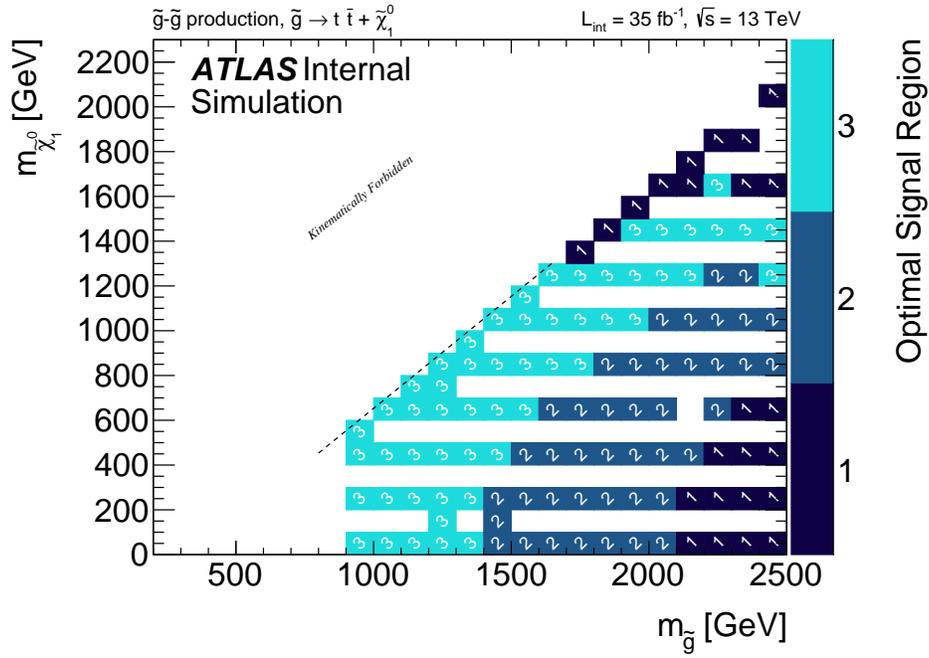
---

<sup>15</sup>In the first iteration of the analysis in 2015 [2], a study was done to determine whether baseline leptons or signal leptons would be used, and signal leptons were found to improve the sensitivity of the cut-and-count analysis.

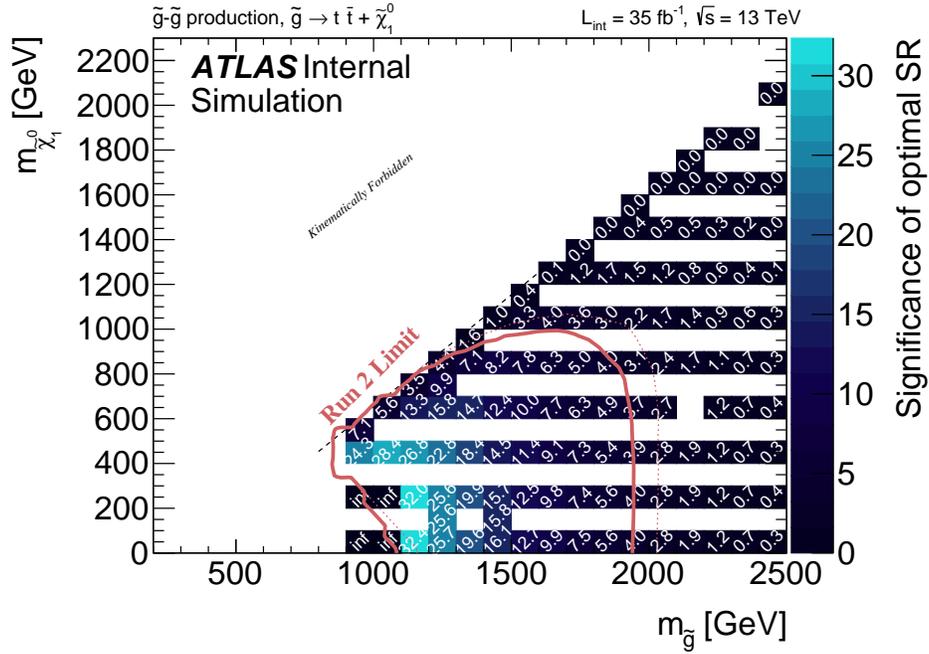
<sup>16</sup>“B” for boosted.

<sup>17</sup>“M” for moderate-boost.

<sup>18</sup>“C” for compressed.

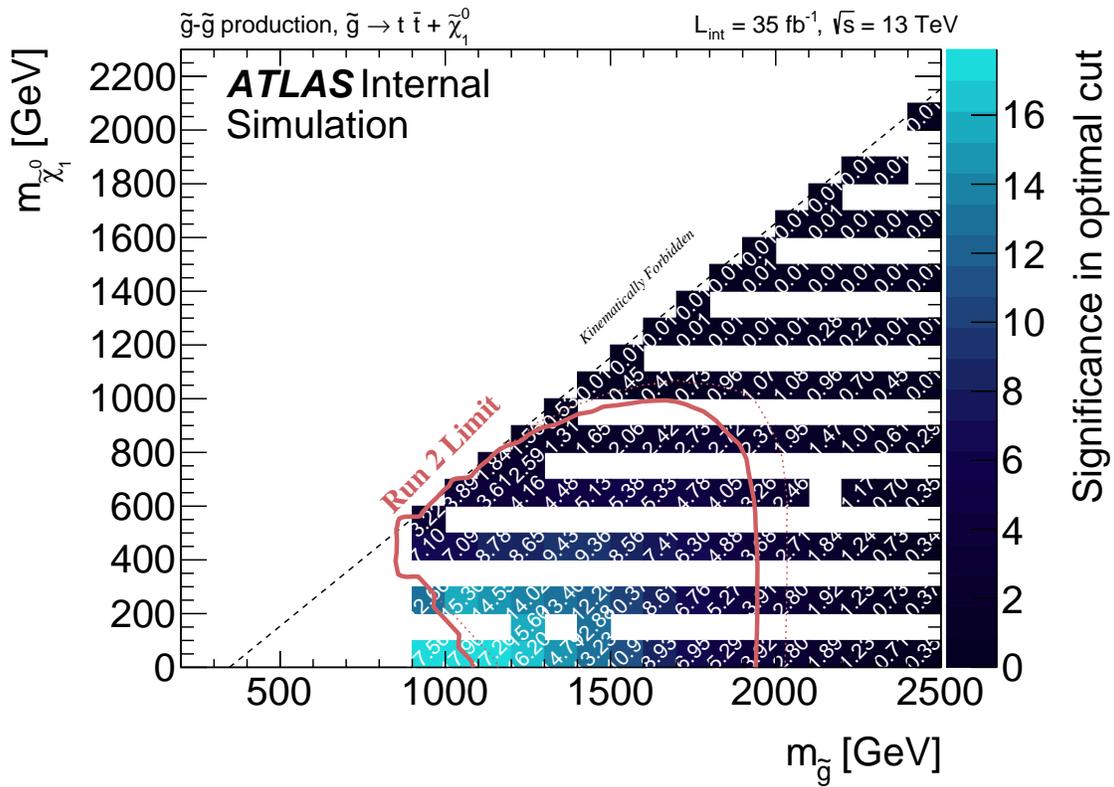


(a) Optimal SR for each grid point in the Gtt 1-lepton channel. In this figure, regions 1-2-3 correspond to B-M-C in the text.



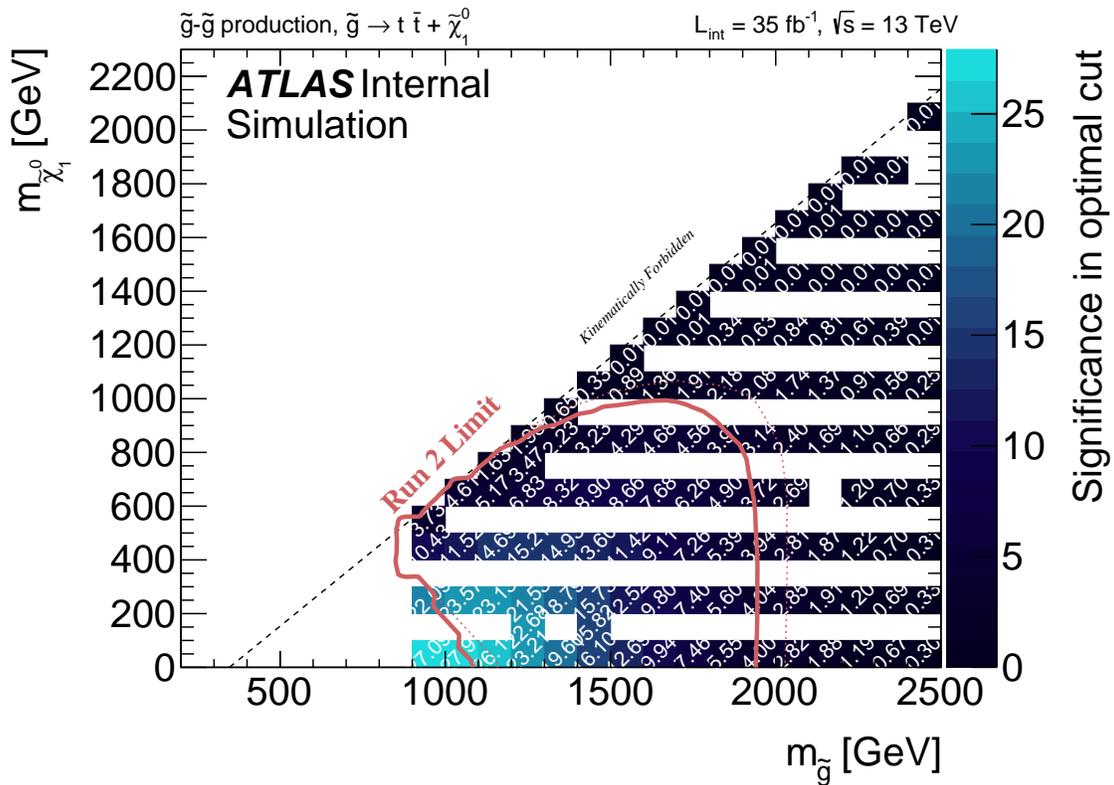
(b) Significance of optimal SR for each grid point in the Gtt 1-lepton channel.

Figure 7.19: (a) Optimal SR and (b) significance for the optimal region for all points of the grid in the Gtt 1-lepton analysis.



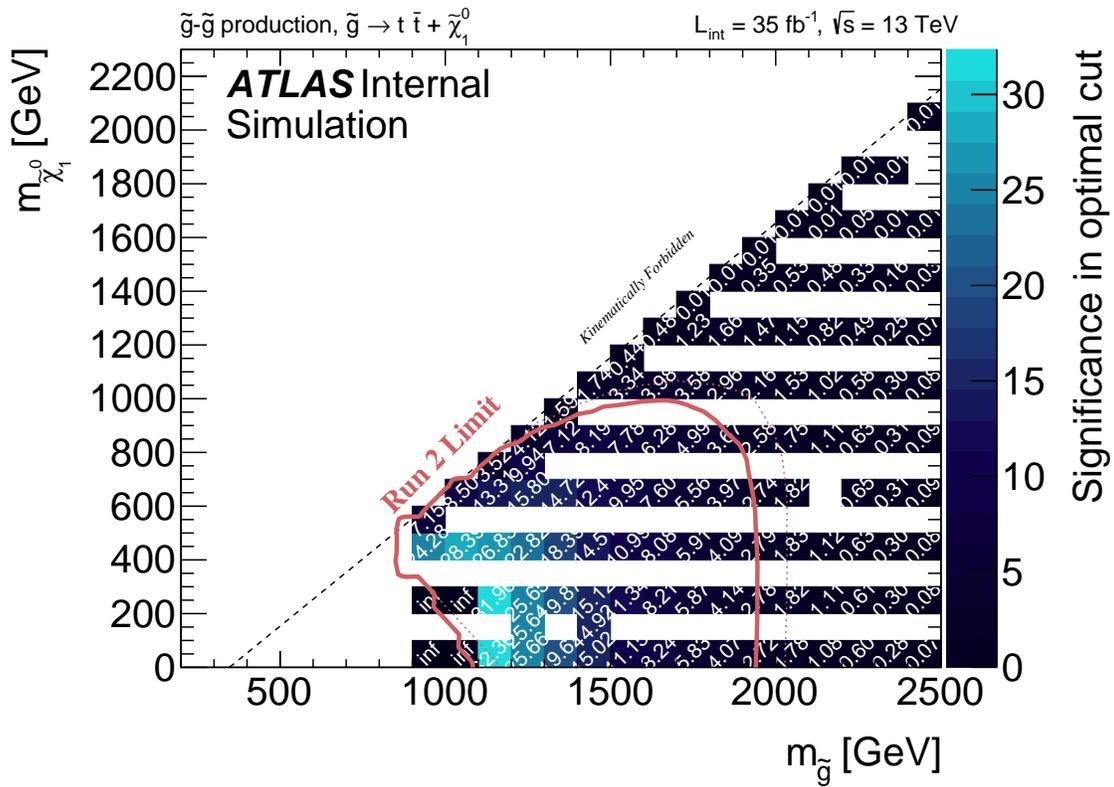
(a) SR-1L-B

Figure 7.20: Significance of each SR at each grid point of the Gtt 1-lepton analysis.



(b) SR-1L-M

Figure 7.20: Significance of each SR at each grid point of the Gtt 1-lepton analysis.



(c) SR-1L-C

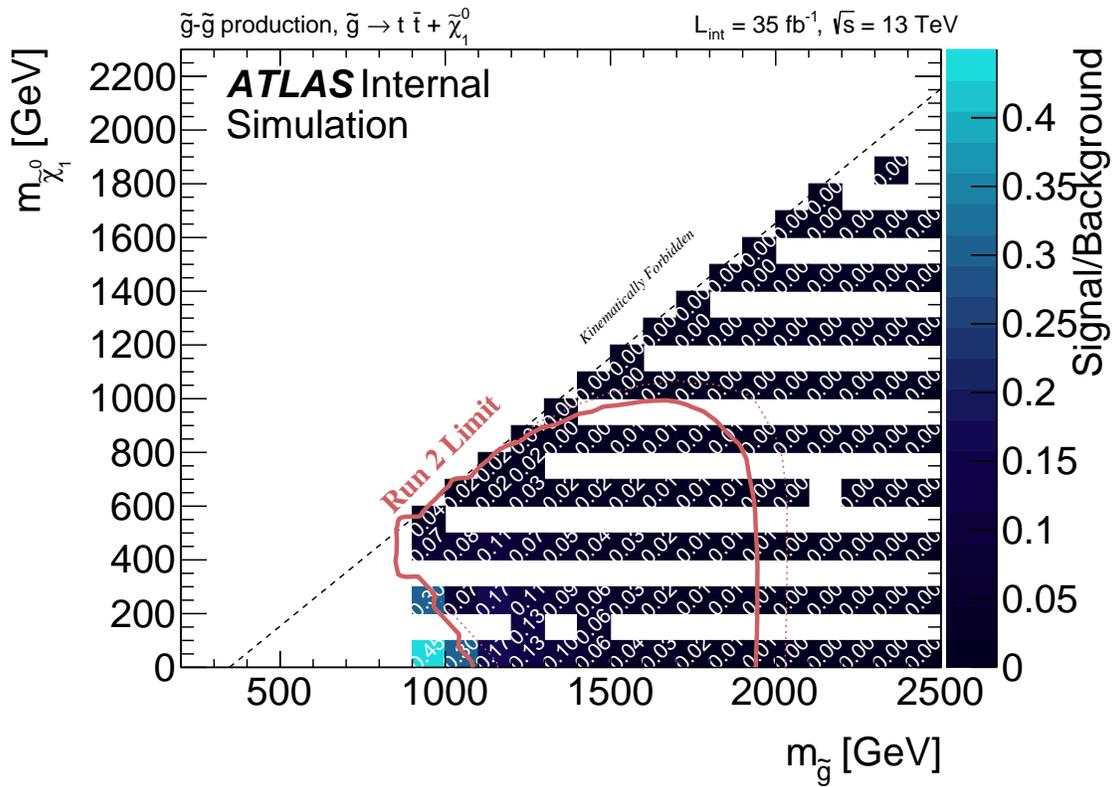
Figure 7.20: Significance of each SR at each grid point of the Gtt 1-lepton analysis.

## Control Regions

**CRs** are also defined for each **SR** in table 7.5. As the background is expected to be dominated entirely by semi-leptonic  $t\bar{t}$  events, due to the  $b$ -jet requirement and the  $E_{\text{T}}^{\text{miss}}$  requirement, following the strategy of the 2015 analysis we define only a single  $t\bar{t}$  **CR** for the analysis and will take all the remaining backgrounds directly from simulation [16]. Since the background is expected to be semi-leptonic  $t\bar{t}$ , a single lepton **CR** strategy is used. The cut on  $m_{\text{T}}$  is inverted in order to find a region with similar background composition but low signal contamination, while preventing overlap with the Gtt **SRs**. Cuts on  $E_{\text{T}}^{\text{miss}}$  and  $m_{\text{eff}}$  and  $m_{\text{T},\text{min}}^{b\text{-jets}}$  are lowered to ensure sufficient statistics (a minimum of 15 expected events) in the **CRs**. The signal contamination in each **CR** is expected to be very small, and is shown in fig. 7.21.

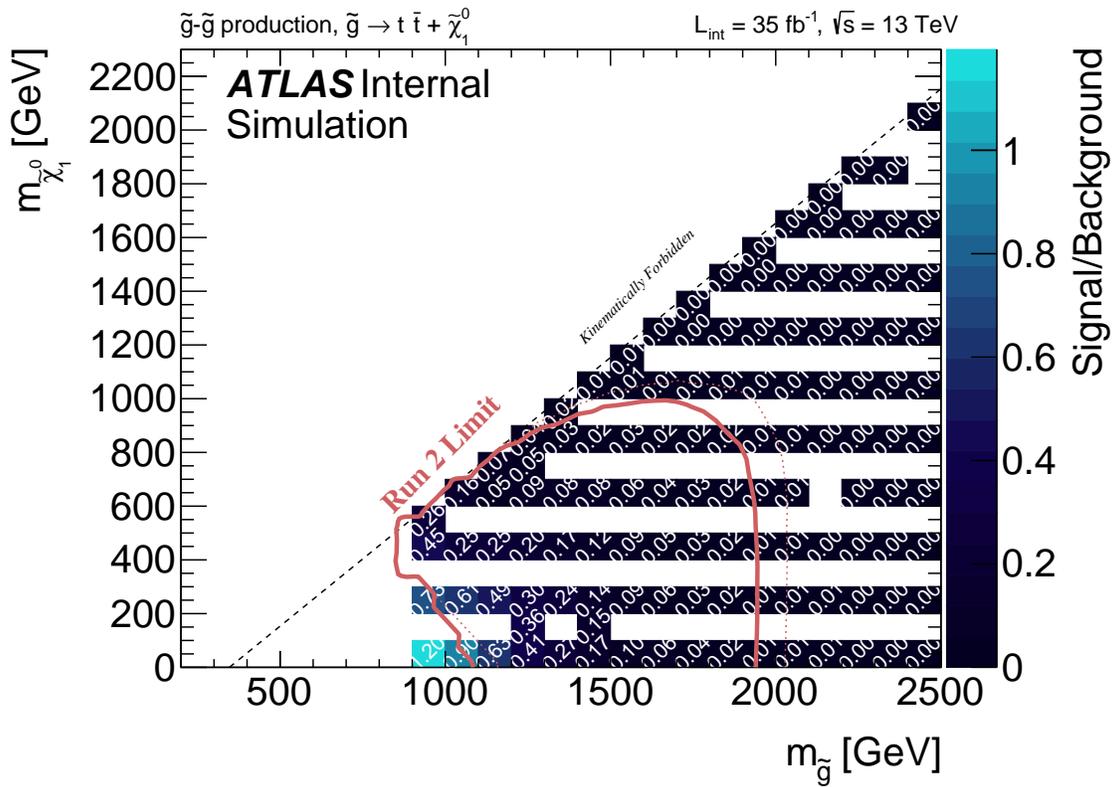
## Validation Regions

Table 7.5 also defines two **VRs** for each **SR**. The **VR- $m_{\text{T}}$**  category validates the extrapolation from low to high  $m_{\text{T}}$ ; orthogonality is enforced with the **SR** by inverting the  $M_{\text{J}}^{\Sigma}$  (in the case of regions B and M) or  $m_{\text{T},\text{min}}^{b\text{-jets}}$  cut (in the case of region C). The **VR- $m_{\text{T},\text{min}}^{b\text{-jets}}$**  category validates the extrapolation from no cut on  $m_{\text{T},\text{min}}^{b\text{-jets}}$  in the **CR** to the optimized cut on  $m_{\text{T},\text{min}}^{b\text{-jets}}$  in the **SR**. Orthogonality with the **CR** is enforced by requiring  $> N_{\text{jet}}$ , instead of exactly  $= N_{\text{jet}}$ . Signal contamination plots for the **VRs** are shown in figs. 7.22 and 7.23. Both are rather low.



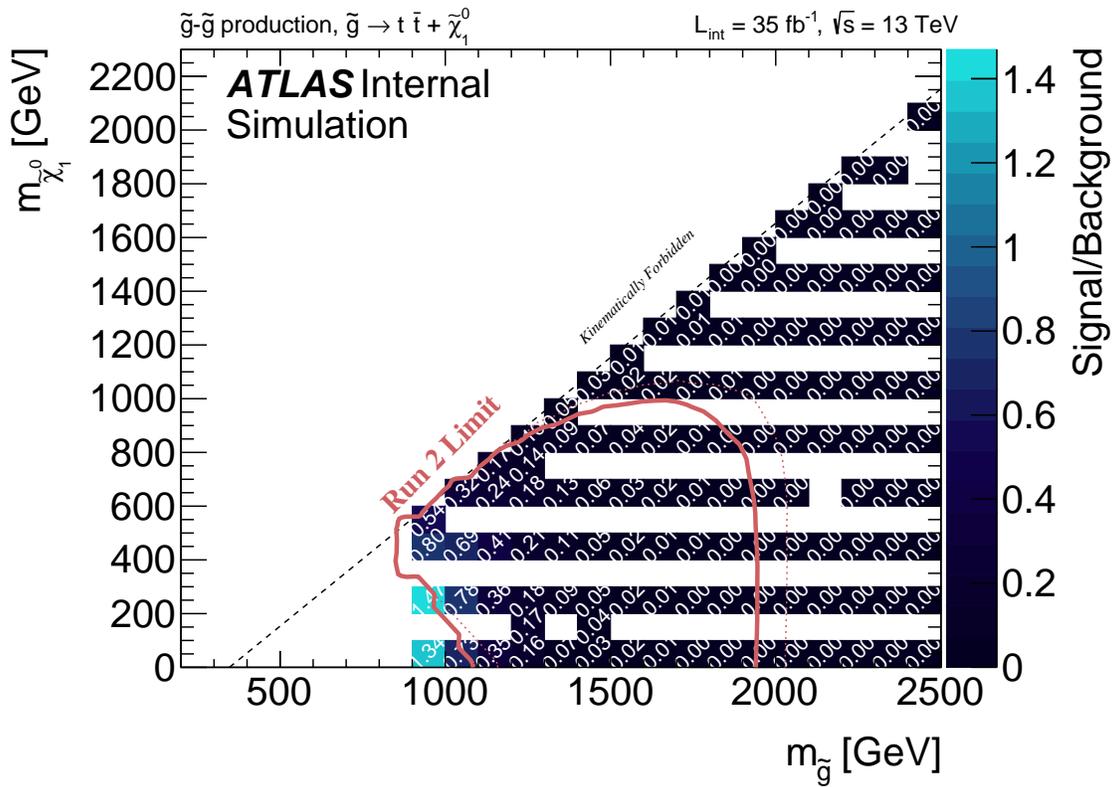
(a) CR-1L-B

Figure 7.21: Signal contamination of each CR at each grid point of the Gtt 1-lepton analysis.



(b) CR-1L-M

Figure 7.21: Signal contamination of each CR at each grid point of the Gtt 1-lepton analysis.



(c) CR-1L-C

Figure 7.21: Signal contamination of each CR at each grid point of the Gtt 1-lepton analysis.

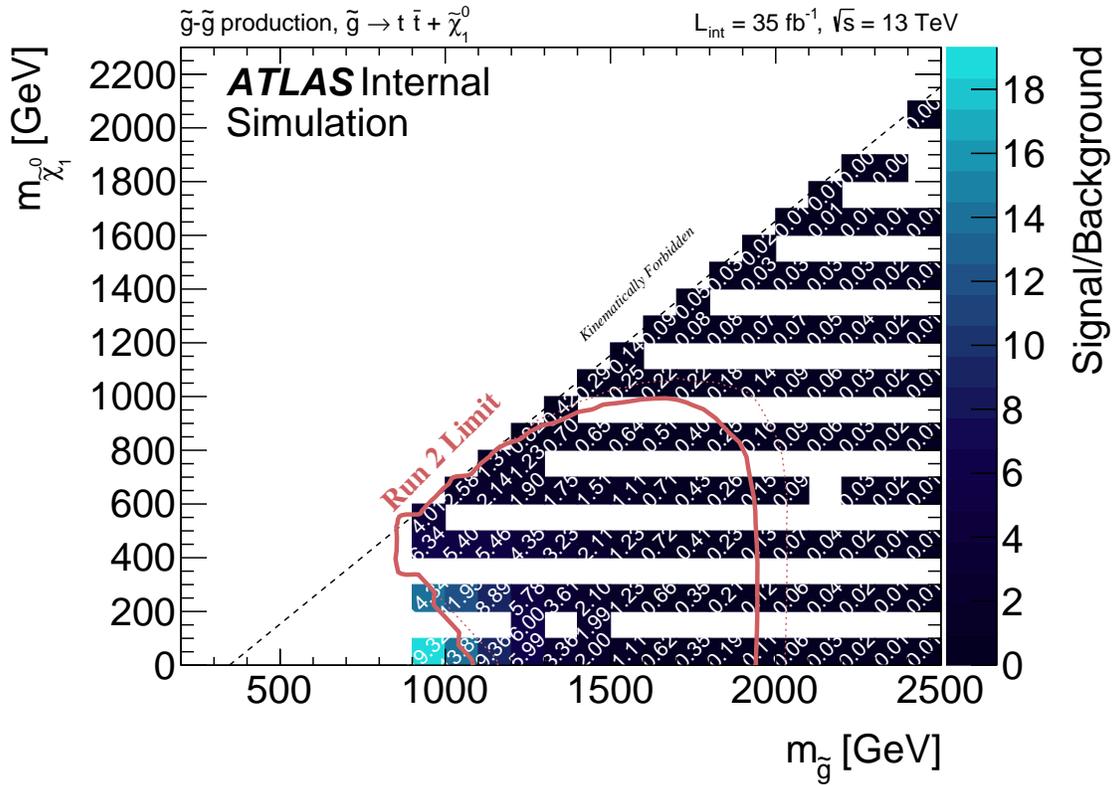


Figure 7.22: Signal contamination of each VR- $m_T$  at each grid point of the Gtt 1-lepton analysis.

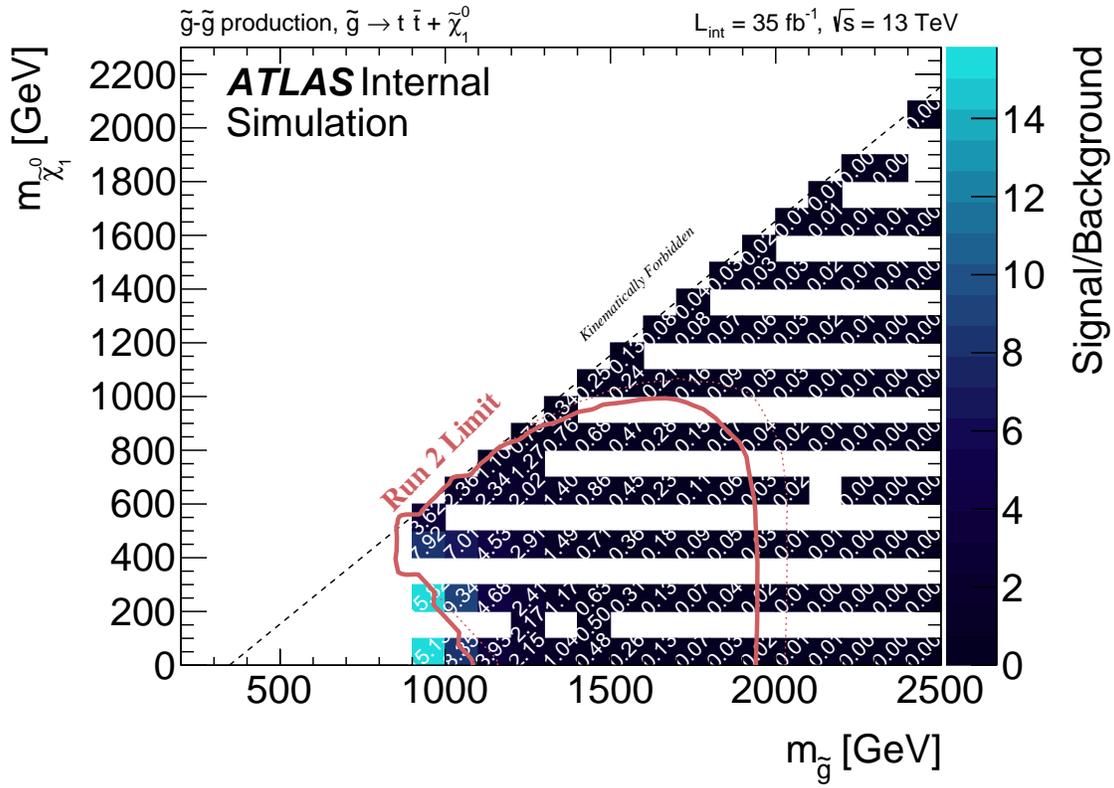
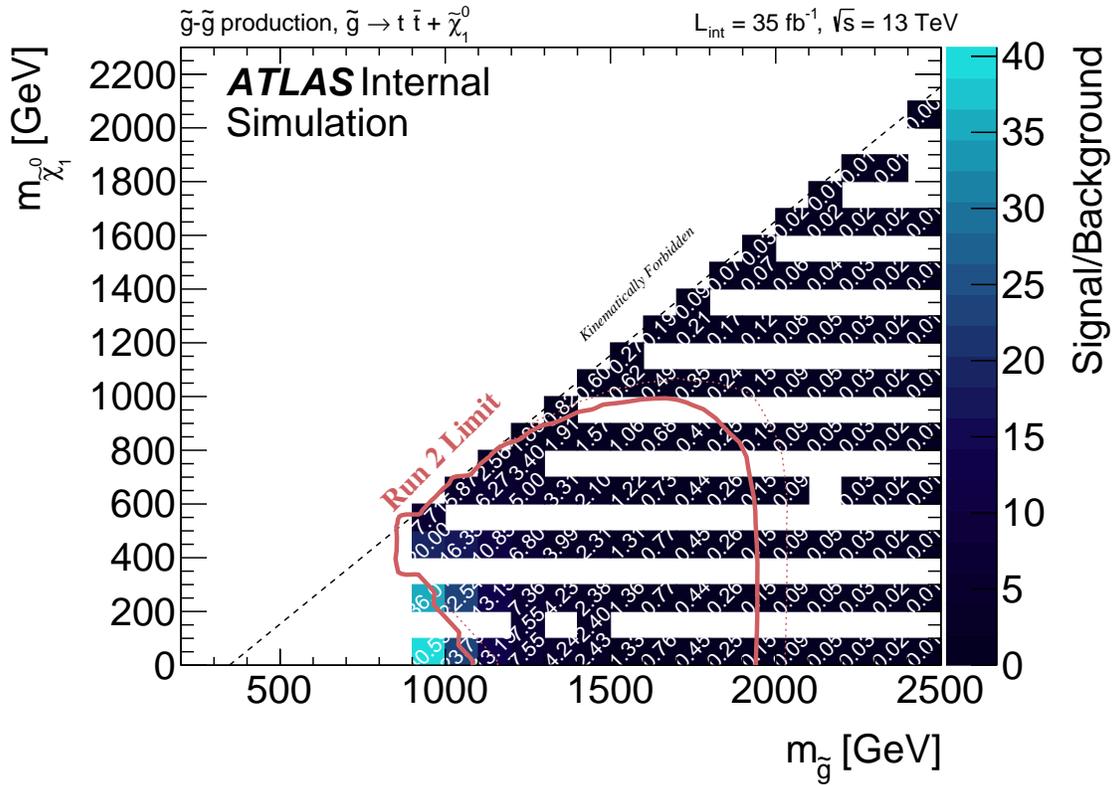
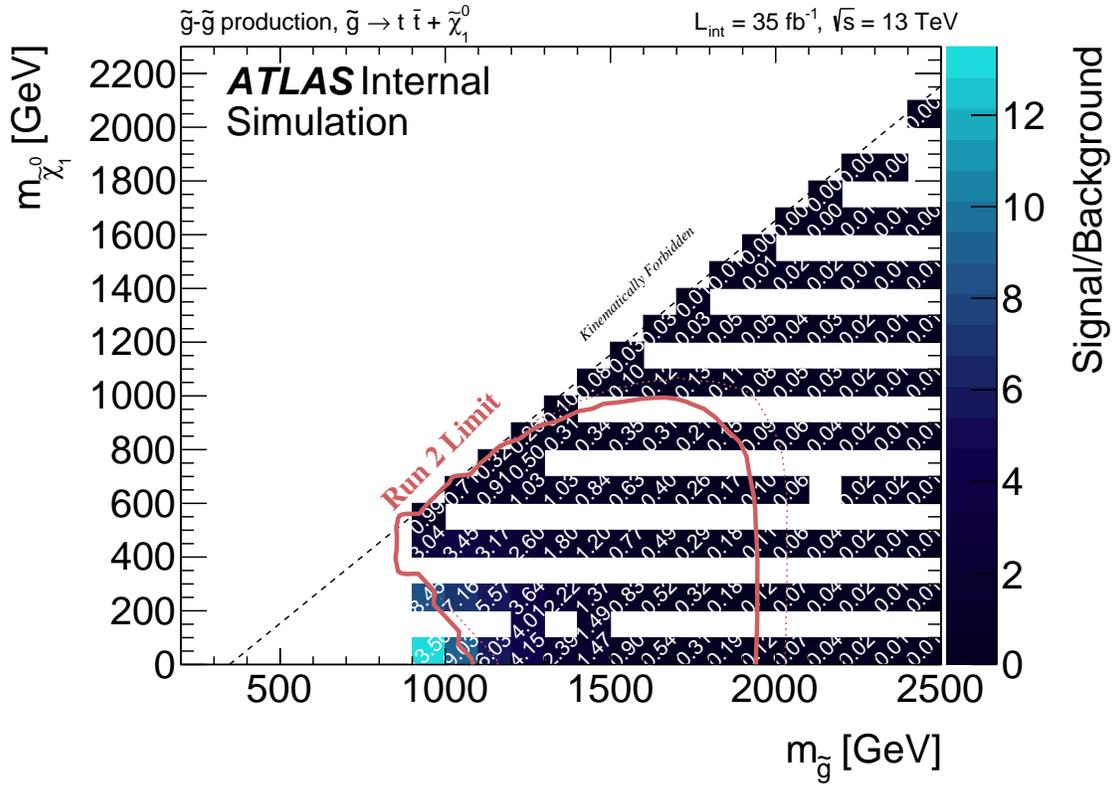


Figure 7.22: Signal contamination of each VR- $m_T$  at each grid point of the Gtt 1-lepton analysis.



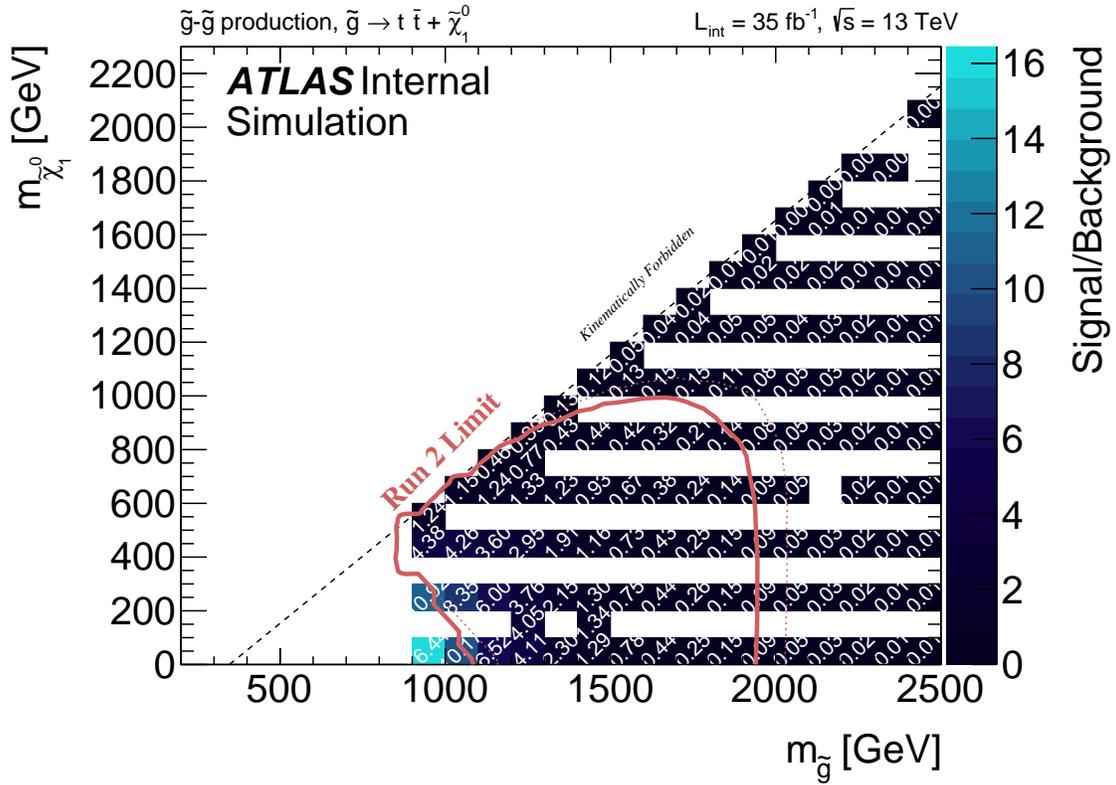
(c) VR- $m_T$ -1L-C

Figure 7.22: Signal contamination of each VR- $m_T$  at each grid point of the Gtt 1-lepton analysis.



(a) VR- $m_{T,\min}^{b\text{-jets}}$ -1L-B

Figure 7.23: Signal contamination of each VR- $m_{T,\min}^{b\text{-jets}}$  at each grid point of the Gtt 1-lepton analysis.



(b) VR- $m_{T,\min}^{b\text{-jets}}$ -1L-M

Figure 7.23: Signal contamination of each VR- $m_{T,\min}^{b\text{-jets}}$  at each grid point of the Gtt 1-lepton analysis.

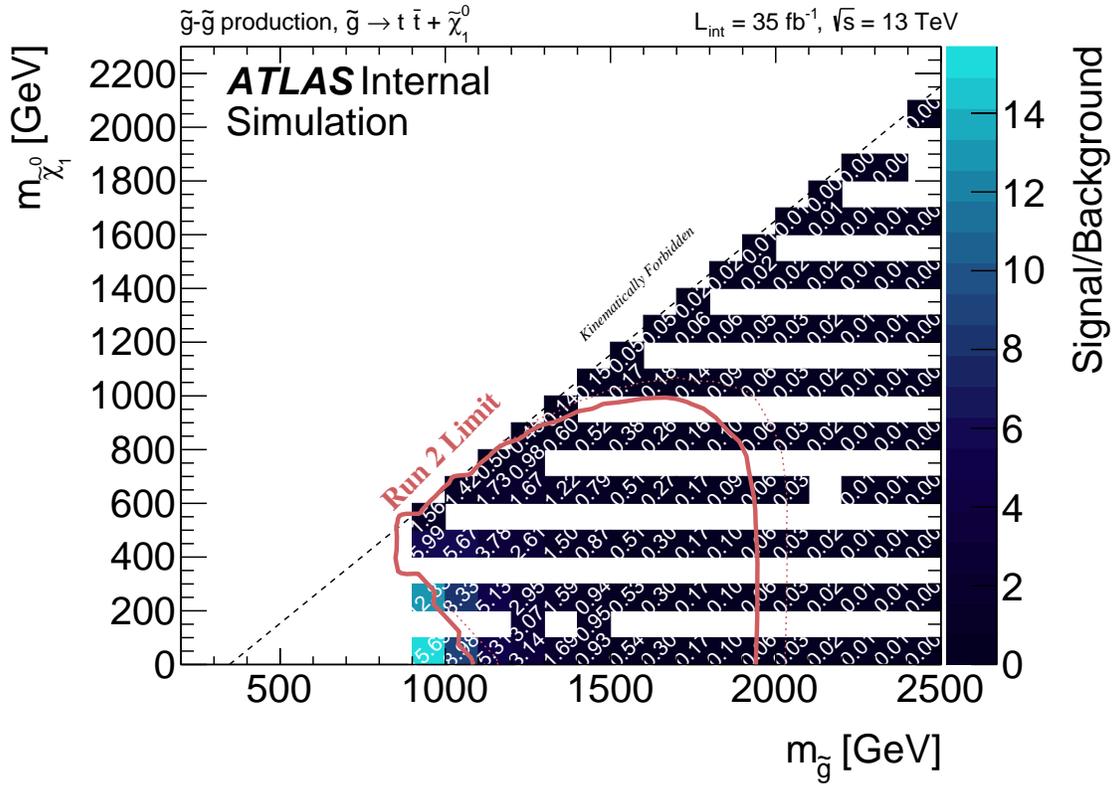


Figure 7.23: Signal contamination of each VR- $m_{T,\min}^{b\text{-jets}}$  at each grid point of the Gtt 1-lepton analysis.

## Background Composition

The composition of the background in the optimized Gtt-1L regions are all shown in ???. Three of these composition plots are shown for the 1-lepton boosted region in fig. 7.24 showing the heavy flavor composition of the  $t\bar{t}$  background. In all cases, the signal region is dominated by  $t\bar{t}$  as expected, as are the CRs and VRs. The expected yields of the CRs and VRs are also sufficient. Note all the results for comoposition are pre-fit.

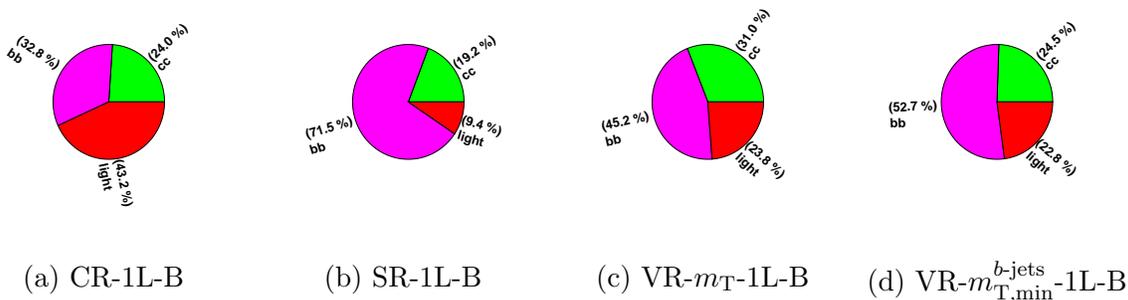


Figure 7.24: Heavy flavor composition of the  $t\bar{t}$  component of the background in the optimized Gtt-1L regions.

## $N-1$ Plots

Distributions of  $N-1$  plots for the kinematic variables used in the optimization of signal regions are all shown in ??. Each of the variables shows significant discrimination power for the signals shown. An example plot is shown for the 0-lepton boosted signal region in fig. 7.25 for the total jet mass variable. This plot is made by applying all of the  $N$  selections in the 0-lepton boosted signal region described in table 7.5 except for the total jet mass variable  $M_J^\Sigma$ , hence the name  $N-1$  plot.

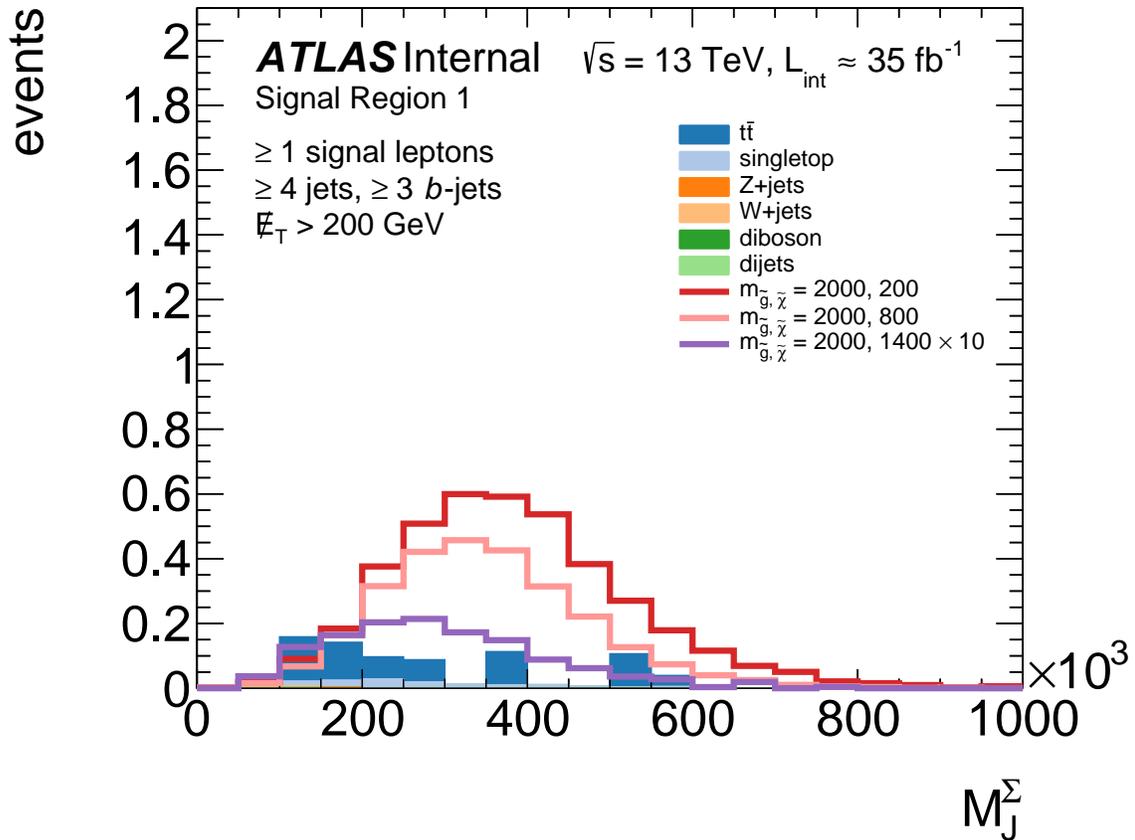


Figure 7.25:  $N-1$  plot of the total jet mass variable in the 1-lepton boosted region. The background is stacked in the histogram and three signal points are overlaid to show the shape comparisons, all normalized to  $35 \text{ fb}^{-1}$ , except for the highest mass signal curve which has been scaled up by a factor of 10.

## 7.5 Region Definitions for Cut-and-Count Analysis

The **SRs** are named in the form **SR- $X$ - $YL$ - $Z$** , where  $X$  indicates the target model,  $Y$  indicates the number of leptons and  $Z$  labels the type of region targeted. The cut-and-count regions labelled B (for “boosted”) are optimised for signals with a large mass difference between the gluino and the neutralino ( $\Delta m \gtrsim 1.5$  TeV), possibly leading to highly boosted objects in the final state. Conversely, regions C (for “compressed”) primarily focus on signals for which the gluino decay products are softer due to the small  $\Delta m$  ( $\Delta m \lesssim 300$  GeV). Regions M (for “moderate”) target intermediate values of  $\Delta m$ . **SRs** targeting the Gtt model in the 0-lepton and 1-lepton channels are presented in tables 7.4 and 7.5.

In the 1-lepton channel, these regions differ mainly in their kinematic selections thresholds:  $m_{\text{eff}}$ ,  $E_{\text{T}}^{\text{miss}}$  and  $M_{\mathcal{J}}^{\Sigma}$  selections are relaxed when going from region B to C to improve the acceptance for softer signals. The resulting background increase is compensated for by tightening the requirements on the number of ( $b$ -tagged) jets or  $m_{\text{T},\text{min}}^{b\text{-jets}}$ . **CRs** constraining the  $t\bar{t}$  background are defined in the low- $m_{\text{T}}$  region to remove overlaps with the **SRs**. The requirements on  $m_{\text{T},\text{min}}^{b\text{-jets}}$  are removed, and the selections on kinematic variables are relaxed to ensure at least about 10 events in each **CR**.

The requirement of an exclusive jet multiplicity permits the definition of **VRs** kinematically close to the **SRs** and mutually exclusive to both the **CRs** and **SRs**. **VR- $m_{\text{T}}$**  validates the background prediction in the high- $m_{\text{T}}$  region. It is kept mutually exclusive with the **SR** by an inverted selection on  $M_{\mathcal{J}}^{\Sigma}$  or  $m_{\text{T},\text{min}}^{b\text{-jets}}$ . **VR- $m_{\text{T},\text{min}}^{b\text{-jets}}$**  checks the background prediction in the high- $m_{\text{T},\text{min}}^{b\text{-jets}}$  regime, with an upper bound on  $m_{\text{T}}$  to keep the region mutually exclusive with the corresponding **SR**. The other kinematic requirements are kept as close as possible to those of the **SRs** to ensure that the event kinematics are similar, and allow sufficiently large yields.

### Gtt 0-lepton

Criteria common to all regions:  $p_T^{\text{jet}} > 30$  GeV

Targeted kinematics	Type	$N_{\text{lepton}}$	$N_{b\text{-jets}}$	$N_{\text{jet}}$	$\Delta\phi_{\text{min}}^{4j}$	$m_T$	$m_T^{b\text{-jets, min}}$	$E_T^{\text{miss}}$	$m_{\text{eff}}^{\text{incl}}$	$M_J^\Sigma$
Region B (Boosted, Large $\Delta m$ )	SR	= 0	$\geq 3$	$\geq 7$	$> 0.4$	–	$> 60$	$> 350$	$> 2600$	$> 300$
	CR	= 1	$\geq 3$	$\geq 6$	–	$< 150$	–	$> 275$	$> 1800$	$> 300$
	VR	= 0	$\geq 3$	$\geq 6$	$> 0.4$	–	–	$> 250$	$> 2000$	$< 300$
Region M (Moderate $\Delta m$ )	SR	= 0	$\geq 3$	$\geq 7$	$> 0.4$	–	$> 120$	$> 500$	$> 1800$	$> 200$
	CR	= 1	$\geq 3$	$\geq 6$	–	$< 150$	–	$> 400$	$> 1700$	$> 200$
	VR	= 0	$\geq 3$	$\geq 6$	$> 0.4$	–	–	$> 450$	$> 1400$	$< 200$
Region C (Compressed, moderate $\Delta m$ )	SR	= 0	$\geq 4$	$\geq 8$	$> 0.4$	–	$> 120$	$> 250$	$> 1000$	$> 100$
	CR	= 1	$\geq 4$	$\geq 7$	–	$< 150$	–	$> 250$	$> 1000$	$> 100$
	VR	= 0	$\geq 4$	$\geq 7$	$> 0.4$	–	–	$> 250$	$> 1000$	$< 100$

Table 7.4: Definitions of the Gtt 0-lepton SRs, CRs and VRs of the cut-and-count analysis. All kinematic variables are expressed in GeV except  $\Delta\phi_{\text{min}}^{4j}$ , which is in radians. The jet  $p_T$  requirement is also applied to  $b$ -tagged jets.

<b>Gtt 1-lepton</b>							
Criteria common to all regions: $\geq 1$ signal lepton, $p_T^{\text{jet}} > 30$ GeV, $N_{b\text{-jets}} \geq 3$							
Targeted kinematics	Type	$N_{\text{jet}}$	$m_T$	$m_T^{b\text{-jets}}_{\text{min}}$	$E_T^{\text{miss}}$	$m_{\text{eff}}^{\text{incl}}$	$M_J^\Sigma$
Region B (Boosted, Large $\Delta m$ )	SR	$\geq 5$	$> 150$	$> 120$	$> 500$	$> 2200$	$> 200$
	CR	$= 5$	$< 150$	–	$> 300$	$> 1700$	$> 150$
	VR- $m_T$	$\geq 5$	$> 150$	–	$> 300$	$> 1600$	$< 200$
	VR- $m_T^{b\text{-jets}}_{\text{min}}$	$> 5$	$< 150$	$> 120$	$> 400$	$> 1400$	$> 200$
Region M (Moderate $\Delta m$ )	SR	$\geq 6$	$> 150$	$> 160$	$> 450$	$> 1800$	$> 200$
	CR	$= 6$	$< 150$	–	$> 400$	$> 1500$	$> 100$
	VR- $m_T$	$\geq 6$	$> 200$	–	$> 250$	$> 1200$	$< 100$
	VR- $m_T^{b\text{-jets}}_{\text{min}}$	$> 6$	$< 150$	$> 140$	$> 350$	$> 1200$	$> 150$
Region C (Compressed, small $\Delta m$ )	SR	$\geq 7$	$> 150$	$> 160$	$> 350$	$> 1000$	–
	CR	$= 7$	$< 150$	–	$> 350$	$> 1000$	–
	VR- $m_T$	$\geq 7$	$> 150$	$< 160$	$> 300$	$> 1000$	–
	VR- $m_T^{b\text{-jets}}_{\text{min}}$	$> 7$	$< 150$	$> 160$	$> 300$	$> 1000$	–

Table 7.5: Definitions of the Gtt 1-lepton SRs, CRs and VRs of the cut-and-count analysis. All kinematic variables are expressed in GeV except  $\Delta\phi_{\text{min}}^{4j}$ , which is in radians. The jet  $p_T$  requirement is also applied to  $b$ -tagged jets.

The [SRs](#) of the 0-lepton channel follow a similar strategy to the 1-lepton channel. Background composition studies performed on simulated event samples show that semileptonic  $t\bar{t}$  events, for which the lepton is outside the acceptance or is a hadronically decaying  $\tau$ -lepton, dominate in the [SRs](#). Thus, [CRs](#) to normalise the  $t\bar{t}$ +jets background make use of the 1-lepton channel, requiring the presence of exactly one signal lepton. An inverted selection on  $m_T$  is applied to suppress overlaps with the 1-lepton [SRs](#). The background prediction is validated in a 0-lepton region, inverting the  $M_J^\Sigma$  selection to suppress any overlap with the [SRs](#).

## 7.6 Semi Data-Driven $t\bar{t}$ Normalization

The main source of background is the production of  $t\bar{t}$  events. The third (and fourth)  $b$ -jet required in the region definitions (section [7.5](#)) can come from

- additional  $b$ -jets produced in association with a pair of top quarks
- from a  $c$ -jet
- a  $\tau$ -lepton decaying to hadrons and a  $\nu_\tau$  is mistagged as a  $b$ -jet

The contribution from  $t\bar{t}$  events with a light or gluon jet mistagged as a  $b$ -jet is sub-dominant but not negligible. In the 0-lepton channel, most of these  $t\bar{t}$  events have a  $W$  boson decaying leptonically where the lepton is: not reconstructed, outside of acceptance, mis-identified as a jet, or a  $\tau$  which decays hadronically. In the 1-lepton channel, the high  $m_T$  requirement enhances the contribution from dileptonic  $t\bar{t}$  events with one hadronically decaying  $\tau$ . Additional sources of background are single-top production,  $t\bar{t}+W/Z/h$ ,  $W/Z$ +heavy-flavour jets, and diboson production; as already mention in ??.

The strategy used to estimate the  $t\bar{t}$  is a semi-data-driven method which relies on the renormalisation of  $t\bar{t}$  simulated events in **CRs** enriched in  $t\bar{t}$  background and with low expected yields from the targeted **SUSY** signals. The definition of these **CRs** is described in the section section 7.5. The extrapolation of the  $t\bar{t}$  yield renormalized in the **CR** to the **VRs** and **SRs** is performed by a fit based on the profile likelihood method [17, 18]. Each signal region is fitted separately. The free parameter in each fit is the  $t\bar{t}$  overall normalization scale,  $\mu_{t\bar{t}}$ , while the contributions from other background processes are set at the expected value and allowed to vary within their systematic uncertainties. The systematic uncertainties on the expected values are included in the likelihood as nuisance parameters with a Gaussian probability density function and the correlations are taken into account when appropriate. The likelihood function is built as the product of a Poisson probability density function, describing the event counts in each region, and the constraints on the nuisance parameters. The entire  $t\bar{t}$  background is normalized with one single scale factor. The fitted  $t\bar{t}$  background is normalized in one **CR** with the same  $b$ -tag requirements as the corresponding **SR**. Three different likelihood fits can be performed to extract these results:

- Background-only fit: Only the **CR** is used to constrain the fit parameters. Any potential signal contamination is neglected and the number of observed events in the signal regions is not taken into account in the fit.
- Exclusion fit: Both **CRs** and **SRs** are used to constrain the fit parameters. The signal contribution as predicted by the tested model is taken into account in both regions using an additional free parameter for the non-**SM** signal strength, constrained to be non-negative, in the likelihood fit. Since the observed event yield in the signal region is used, the background prediction can differ from the prediction on the background-only fit. The exclusion fit configuration was used to produce all the model-dependent limits.
- Discovery fit: Both **CRs** and **SRs** are used to constrain the fit parameters. A potential

signal contribution is considered in the signal regions but neglected in the CRs. This background prediction is conservative since any signal contribution in the CRs is attributed to background and thus yields a possible overestimate of the background in the signal regions. The discovery fit configuration is used to produce upper limits on the visible cross-sections.

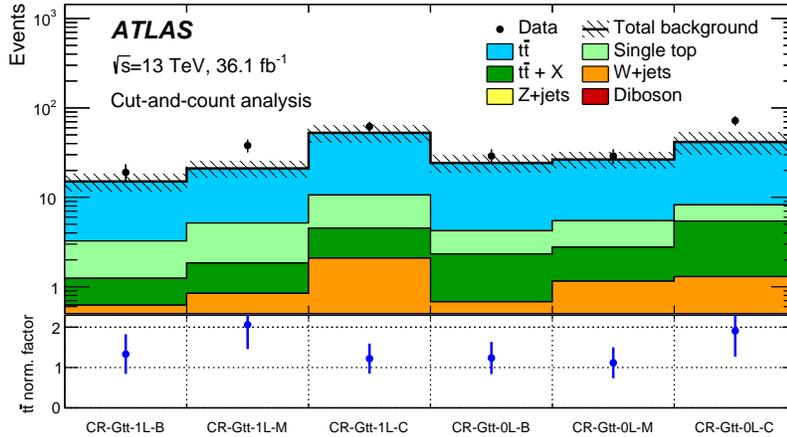


Figure 7.26: The value of  $\mu_{t\bar{t}}$ , and uncertainty, after the background-only fit, for the regions defined in section 7.5.

Figure 7.26 shows the value of  $\mu_{t\bar{t}}$  after the background-only fit in all of the SRs of the analysis described in section 7.5. ?? shows the results of the background-only fit to the VRs and the unblinded SRs. The normalization factor is higher in the 1-lepton regions because of the rising data/MC slope corrected by a kinematic reweighting described in section 7.3. Regions with 4  $b$ -tags have a higher normalization because of the data/MC disagreement in the number of  $b$ -jets, primarily due to the poor modeling in MC.

## 7.7 Systematic Uncertainties

Figure 7.27 summarizes the relative systematic uncertainties in the background estimate for the cut-and-count analysis. These uncertainties arise from the extrapolation of the  $t\bar{t}$  normalization obtained in the CRs to the SRs as well as from the yields of the minor backgrounds in the SRs, which are predicted by the simulation. The total systematic uncertainties range from approximately 20% to 80% in the various SRs.

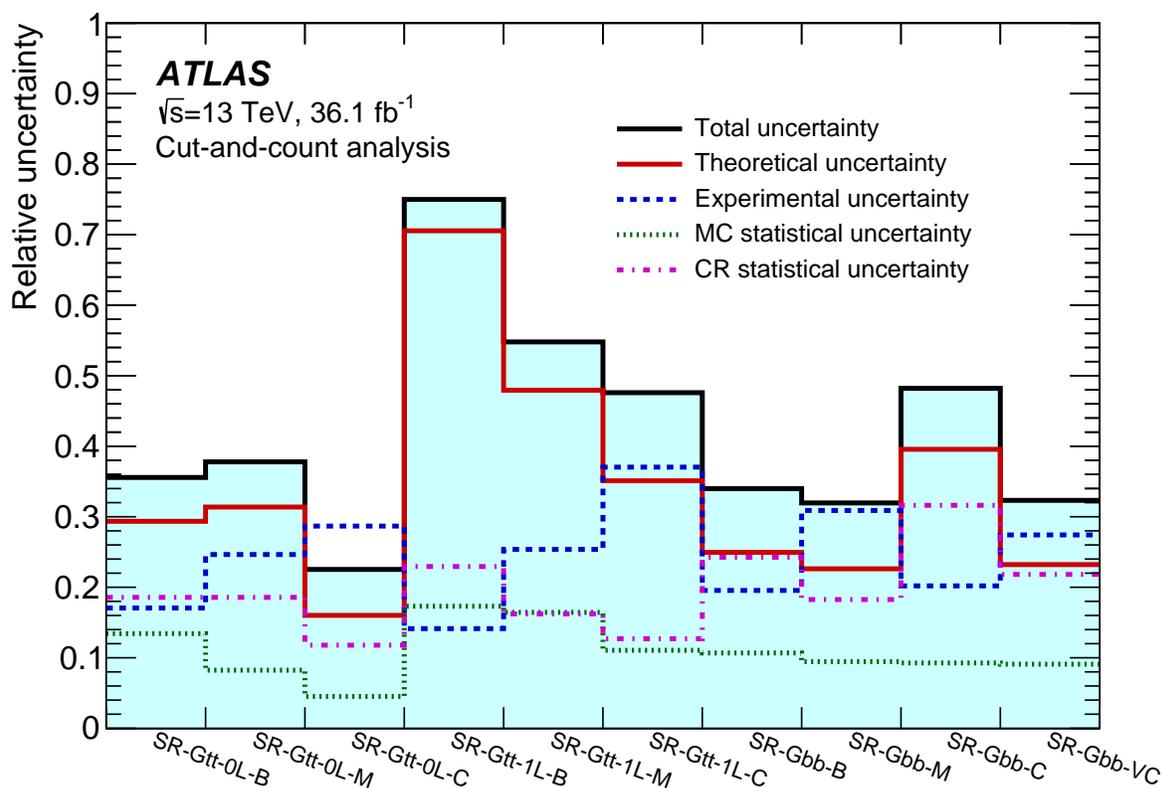


Figure 7.27: Relative systematic uncertainty in the background estimate for the cut-and-count analysis. The individual uncertainties can be correlated, such that the total background uncertainty is not necessarily their sum in quadrature.

### 7.7.1 Experimental Systematic Uncertainties

The detector-related systematic uncertainties affect both the background estimate and the signal yield. The largest sources in this analysis relate to the [jet energy scale \(JES\)](#)<sup>19</sup>, [jet energy resolution \(JER\)](#)<sup>20</sup> and the  $b$ -tagging efficiencies and mistagging rates<sup>21</sup>. The [JES](#) uncertainties for the small- $R$  jets are derived from  $\sqrt{s} = 13$  TeV data and simulations while the [JER](#) uncertainties are extrapolated from 8 TeV data using [MC](#) simulations [19]. These uncertainties are also propagated to the re-clustered [large-radius \(large- \$R\$ \)](#) jets, which use them as inputs. The jet mass scale and resolution uncertainties have a negligible impact on the re-clustered jet mass. The impact of the [JES](#) uncertainties on the expected background yields is between 4% and 35%, while [JER](#) uncertainties affect the background yields by approximately 0–26% in the various regions.

Uncertainties in the measured  $b$ -tagging efficiencies and mistagging rates are the subleading sources of experimental uncertainty. The impact of these uncertainties on the expected background yields is 3–24% depending on the considered region. The uncertainties associated with lepton reconstruction and energy measurements have a negligible impact on the final results. All lepton and jet measurement uncertainties are propagated to the calculation of  $E_{\text{T}}^{\text{miss}}$ , and additional uncertainties are included in the scale and resolution of the soft term. The overall impact of the  $E_{\text{T}}^{\text{miss}}$  soft-term uncertainties is also small. Since the normalization of the  $t\bar{t}$  background is fit to data in the [CRs](#), uncertainties in the modeling of this background only affect the extrapolation from the [CRs](#) to the [SRs](#) and [VRs](#).

---

<sup>19</sup>This is calculated by scaling up and down by  $1\sigma$  the [JES](#) provided by the Jet/Etmiss group within [ATLAS](#).

<sup>20</sup>This is calculated by smearing the  $p_{\text{T}}^{\text{jet}}$  based on the  $p_{\text{T}}^{\text{jet}}$  and  $\eta$  to account for a possible underestimate of the [JER](#) in [MC](#).

<sup>21</sup>This is calculated by varying the  $\eta$ ,  $p_{\text{T}}^{\text{jet}}$ , and flavor of each jet, separately for  $B$ -jets,  $C$ -jets, and light jets. This leads to three uncorrelated systematic uncertainties.

### 7.7.2 Theoretical Systematic Uncertainties on Background

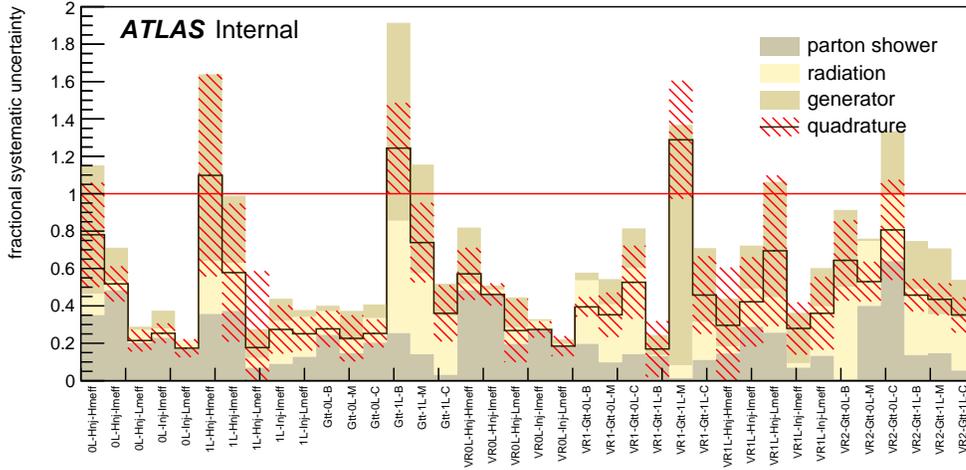
The estimation of the background systematic uncertainties is performed by variations of the MC generator parameters and the comparison of various generator predictions. These are summarized in fig. 7.28. The  $m_{T,\min}^{b\text{-jets}}$  and  $\Delta\phi_{\min}^{4j}$  cuts were relaxed in the SRs to allow for enough statistical power in calculating the theory systematics. There are three main sources of fully uncorrelated uncertainties particular to this analysis: the radiation tunes for PYTHIA v6.428, the hadronization and parton showering model, and the generator<sup>22</sup>.

Hadronization and parton showering model uncertainties are estimated using a sample generated with POWHEG and showered by HERWIG++ v2.7.1 with the UEEE5 underlying-event tune. Systematic uncertainties in the modeling of initial-state radiation (ISR) and final-state radiation (FSR) are explored with POWHEG samples showered with two alternative settings of PYTHIA v6.428. The first of these uses the PERUGIA2012radHi tune [20] and has the renormalization and factorization scales set to twice the nominal value, resulting in more radiation in the final state. In addition, it has  $h_{\text{damp}}$  set to  $2m_{\text{top}}$ . The second sample, using the PERUGIA2012radLo tune, has  $h_{\text{damp}} = m_{\text{top}}$  and the renormalization and factorization scales are set to half of their nominal values, resulting in less radiation in the event. In each case, the uncertainty is taken as the change in the expected yield of  $t\bar{t}$  background with respect to the nominal sample. The uncertainty due to the choice of event generator is estimated by comparing the expected yields obtained using a  $t\bar{t}$  sample generated with MADGRAPH5\_aMC@NLO and one that is generated with POWHEG. Both of these samples are showered with HERWIG++ v2.7.1. The total theoretical uncertainty in the inclusive  $t\bar{t}$  background is taken as the sum in quadrature of these individual components.

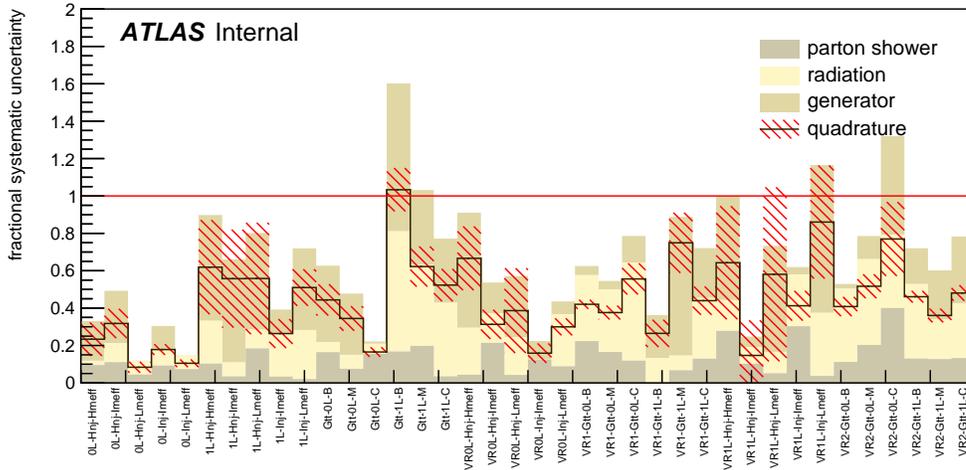
An additional uncertainty is assigned to the fraction of  $t\bar{t}$  events produced in association with additional heavy-flavour jets (i.e.  $t\bar{t}+ \geq 1b$  and  $t\bar{t}+ \geq 1c$ ), a process which suffers from

---

<sup>22</sup>e.g. how the choice of MC generator affects our predictions



(a) without truth  $b$ -tagging



(b) with truth  $b$ -tagging

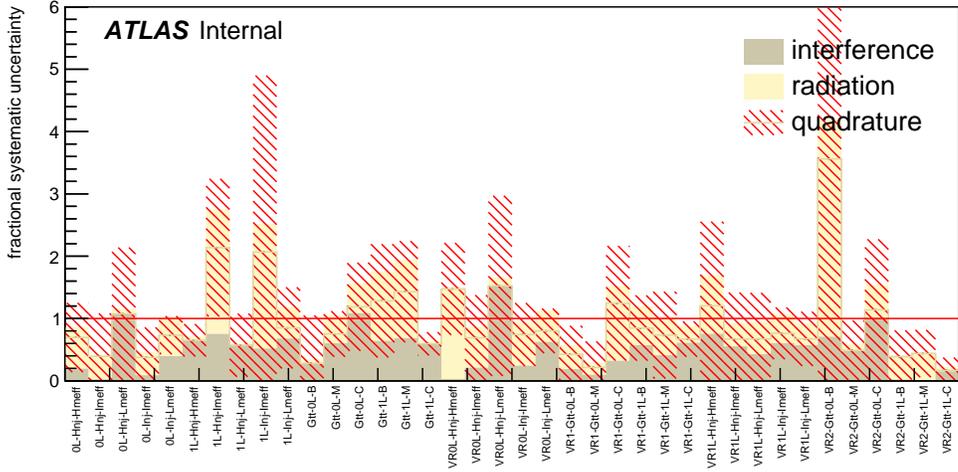
Figure 7.28: The summary of the calculated theory uncertainties by comparison of generator predictions is shown here. Two different versions are shown, without and with truth  $b$ -tagging. One of the main problems was obtaining enough statistical power in truth generator samples. Instead of vetoing truth events that do not have enough  $b$ -tags, one uses truth  $b$ -tagging which weights each event by the probability for it to have the given  $N_{b\text{-jets}}$  requirement (both inclusively and exclusively). Each component of the uncertainty is shown as stacked and the sum in quadrature is overlaid in black. The uncertainty of the systematic is shaded red. The unit is the full size of the uncertainty (i.e. 2 corresponds to a 200% uncertainty).

large theoretical uncertainties. Simulation studies show that the heavy-flavour fractions in each set of SR, CR and VR, which have almost identical  $b$ -tagged jets requirements, are similar. Therefore, the theoretical uncertainties in this fraction affect these regions in a similar way, and thus largely cancel out in the semi-data-driven  $t\bar{t}$  normalization based on the observed CR yields. The residual uncertainty in the  $t\bar{t}$  prediction is taken as the difference between the nominal  $t\bar{t}$  prediction and the one obtained after varying the cross-section of  $t\bar{t}$  events with additional heavy-flavour jets by 30%, in accordance with the results of the ATLAS measurement of this cross-section at  $\sqrt{s} = 8$  TeV [21]. This component typically makes a small contribution (0–8%) to the total impact of the  $t\bar{t}$  modeling uncertainties on the background yields, which ranges between 5% and 76% for the various regions. The statistical uncertainty of the CRs used to extract the  $t\bar{t}$  normalization factors, which is included in the systematic uncertainties, ranges from 10% to 30% depending on the SR.

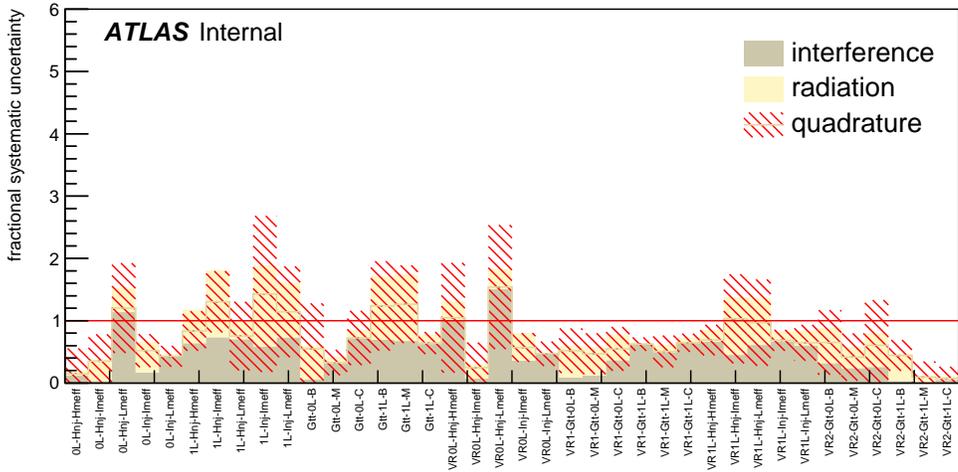
Modeling uncertainties affecting the single-top process arise especially from the interference between the  $t\bar{t}$  and  $Wt$  processes. This uncertainty in fig. 7.29 is estimated using inclusive  $WWbb$  events, generated using MADGRAPH5\_aMC@NLO, which are compared with the sum of  $t\bar{t}$  and  $Wt$  processes. Furthermore, as in the  $t\bar{t}$  modeling uncertainties, variations of PYTHIA v6.428 settings increasing or decreasing the amount of radiation are also used. An additional 5% uncertainty is included in the cross-section of single-top processes [22].

Overall, the modeling uncertainties affecting the single-top process lead to changes of approximately 0–11% in the total yields in the various regions.

Uncertainties in the  $W/Z$ +jets backgrounds are estimated by varying independently the scales for factorization, renormalization and resummation by factors of 0.5 and 2. The scale used for the matching between jets originating from the matrix element and the parton shower is also varied. The resulting uncertainties in the total yield range from approximately 0 to 50% in the various regions. A 50% normalization uncertainty is assigned to  $t\bar{t} + W/Z/h$ ,



(a) without truth  $b$ -tagging



(b) with truth  $b$ -tagging

Figure 7.29: The summary of the calculated theory uncertainties for single top is shown here. Two different versions are shown, without and with truth  $b$ -tagging. One of the main problems was obtaining enough statistical power in truth generator samples. Instead of vetoing truth events that do not have enough  $b$ -tags, one uses truth  $b$ -tagging which weights each event by the probability for it to have the given  $N_{b\text{-jets}}$  requirement (both inclusively and exclusively). Each component of the uncertainty is shown as stacked and the sum in quadrature is overlaid in black. The uncertainty of the systematic is shaded red. The unit is the full size of the uncertainty (i.e. 2 corresponds to a 200% uncertainty).

$t\bar{t}t\bar{t}$  and diboson backgrounds and are found to have no significant impact on the sensitivity of this analysis. Uncertainties arising from variations of the parton distribution functions were found to affect background yields by less than 2%, and therefore these uncertainties are neglected here. Uncertainties due to the limited number of events in the MC background samples are included if above 5%. They reach approximately 20% in regions targeting large mass-splitting.

### 7.7.3 *Systematic uncertainties on the signal*

The signal samples are normalized using the best cross-section calculations at [next-to-leading-order \(NLO\)](#) in the strong coupling constant, adding the resummation of soft gluon emission at [next-to-leading-logarithm \(NLL\)](#) accuracy [23, 24, 25, 26, 27]. The nominal cross-section and the uncertainty are taken from an envelope of cross-section predictions using different PDF sets and factorization and renormalization scales, as described in [28]. The cross-section of gluino pair-production in these simplified models is  $14 \pm 3 \text{ fb}^{-1}$  for a gluino mass of 1.5 TeV, falling to  $1.0 \pm 0.3 \text{ fb}^{-1}$  for 2 TeV mass gluino. This is also summarized in ??.

### 7.7.4 *Other systematic uncertainties*

A systematic uncertainty is also assigned to the kinematic correction described in section 7.3. The total size of the correction is used as an uncertainty, and is applied to all simulated event samples for the 1-lepton channel.

## Glossary

**ATLAS** a general-purpose detector at the [Large Hadron Collider \(LHC\)](#). [3](#), [9](#), [20](#), [72](#)

**BC** Bunch Crossing. [11](#)

**CR** control region. [19–21](#), [28](#), [33–36](#), [43](#), [54–57](#), [63](#), [65–72](#), [75](#)

**FSR** final-state radiation. [73](#)

**GRL** Good Runs List. [9](#)

**IBL** Insertable B-Layer. [9](#)

**ID** Inner Detector. [11](#)

**ISR** initial-state radiation. [73](#)

**JER** jet energy resolution. [72](#)

**JES** jet energy scale. [72](#)

**L1** Level-1. [10](#), [11](#)

**LAr** Liquid Argon Calorimeter. [9](#)

**large- $R$**  large-radius. [72](#)

**MC** [monte-carlo](#). [1](#), [9–11](#), [20](#), [70](#), [72](#), [73](#)

**monte-carlo** simulated event using random numbers. [1](#)

**NLL** next-to-leading-logarithm. [77](#)

**NLO** next-to-leading-order. [77](#)

**punch-through** For jets at very high transverse momentum it is possible that part of the energy is not deposited in the calorimeter, but leaks out to the detector components beyond the calorimeter. This leads to a systematic reduction in the measured jet energy. Jets that deposit energy beyond the hadronic Tile calorimeter and in the muon system are called punch-through jets. [29]. 12

**SCT** Semiconductor Tracker. 9

**SM** Standard Model. 1, 20, 69

**small- $R$**  small-radius. 6, 28, 49

**SR** signal region. 14, 19–21, 28–33, 36, 43, 49–54, 57, 65–73

**SUSY** Supersymmetry. 3, 5, 20, 21, 69

**Tile** Tile calorimeter. 9

**VR** validation region. 19, 20, 33, 36–41, 57–63, 65–67, 69, 70, 72

## Bibliography

- [1] ATLAS Collaboration. “Search for strong production of supersymmetric particles in final states with missing transverse momentum and at least three  $b$ -jets at  $\sqrt{s} = 8$  TeV proton–proton collisions with the ATLAS detector”. In: *JHEP* 10 (2014), p. 024. DOI: [10.1007/JHEP10\(2014\)024](https://doi.org/10.1007/JHEP10(2014)024). arXiv: [1407.0600 \[hep-ex\]](https://arxiv.org/abs/1407.0600) (cit. on pp. 1, 2).
- [2] ATLAS Collaboration. “Search for pair production of gluinos decaying via stop and sbottom in events with  $b$ -jets and large missing transverse momentum in  $pp$  collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: *Phys. Rev. D* 94 (2016), p. 032003. DOI: [10.1103/PhysRevD.94.032003](https://doi.org/10.1103/PhysRevD.94.032003). arXiv: [1605.09318 \[hep-ex\]](https://arxiv.org/abs/1605.09318) (cit. on pp. 2, 23, 49).
- [3] Johan Alwall, Philip Schuster, and Natalia Toro. “Simplified Models for a First Characterization of New Physics at the LHC”. In: *Phys. Rev. D* 79 (2009), p. 075020. DOI: [10.1103/PhysRevD.79.075020](https://doi.org/10.1103/PhysRevD.79.075020). arXiv: [0810.3921 \[hep-ph\]](https://arxiv.org/abs/0810.3921) (cit. on p. 2).
- [4] Daniele Alves. “Simplified Models for LHC New Physics Searches”. In: *J. Phys.* G39 (2012). Ed. by Nima Arkani-Hamed et al., p. 105005. DOI: [10.1088/0954-3899/39/10/105005](https://doi.org/10.1088/0954-3899/39/10/105005). arXiv: [1105.2838 \[hep-ph\]](https://arxiv.org/abs/1105.2838) (cit. on p. 2).
- [5] Timothy Cohen et al. “Dissecting Jets and Missing Energy Searches Using  $n$ -body Extended Simplified Models”. In: *JHEP* 08 (2016), p. 038. DOI: [10.1007/JHEP08\(2016\)038](https://doi.org/10.1007/JHEP08(2016)038). arXiv: [1605.01416 \[hep-ph\]](https://arxiv.org/abs/1605.01416) (cit. on p. 3).
- [6] ATLAS Collaboration. “Search for new phenomena in final states with large jet multiplicities and missing transverse momentum at  $\sqrt{s} = 8$  TeV proton–proton collisions using the ATLAS experiment”. In: *JHEP* 10 (2013), p. 130. DOI: [10.1007/JHEP10\(2013\)130](https://doi.org/10.1007/JHEP10(2013)130). arXiv: [1308.1841 \[hep-ex\]](https://arxiv.org/abs/1308.1841) (cit. on p. 6). Erratum: in: *JHEP* 01 (2014), p. 109. DOI: [10.1007/JHEP01\(2014\)109](https://doi.org/10.1007/JHEP01(2014)109).

- [7] ATLAS Collaboration. “Search for massive supersymmetric particles decaying to many jets using the ATLAS detector in  $pp$  collisions at  $\sqrt{s} = 8$  TeV”. In: *Phys. Rev. D* 91 (2015), p. 112016. DOI: [10.1103/PhysRevD.91.112016](https://doi.org/10.1103/PhysRevD.91.112016). arXiv: [1502.05686 \[hep-ex\]](https://arxiv.org/abs/1502.05686) (cit. on p. 6). Erratum: in: *Phys. Rev. D* 93 (2016), p. 039901. DOI: [10.1103/PhysRevD.93.039901](https://doi.org/10.1103/PhysRevD.93.039901).
- [8] ATLAS Collaboration. “Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using  $4.7 \text{ fb}^{-1}$  of  $\sqrt{s} = 7$  TeV proton-proton collision data”. In: *Phys. Rev. D* 87.1 (2013), p. 012008. DOI: [10.1103/PhysRevD.87.012008](https://doi.org/10.1103/PhysRevD.87.012008). arXiv: [1208.0949 \[hep-ex\]](https://arxiv.org/abs/1208.0949) (cit. on p. 7).
- [9] Jean-Francois Arguin et al. *Search for gluino-mediated stop and sbottom pair production in events with b-jets and large missing transverse momentum*. Tech. rep. ATL-COM-PHYS-2016-1592. Geneva: CERN, Nov. 2016. URL: <https://cds.cern.ch/record/2231120> (cit. on pp. 11, 41).
- [10] ATLAS Collaboration. *Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector*. ATLAS-CONF-2015-029. 2015. URL: <https://cds.cern.ch/record/2037702> (cit. on p. 12).
- [11] ATLAS Collaboration. “Search for Supersymmetry in final states with missing transverse momentum and multiple  $b$ -jets in proton-proton collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector”. In: (2017). arXiv: [1711.01901 \[hep-ex\]](https://arxiv.org/abs/1711.01901) (cit. on p. 13).
- [12] ATLAS Collaboration. “Search for squarks and gluinos with the ATLAS detector in final states with jets and missing transverse momentum using  $4.7 \text{ fb}^{-1}$  of  $\sqrt{s} = 7$  TeV proton-proton collision data”. In: *Phys. Rev. D* 87 (2013), p. 012008. DOI: [10.1103/PhysRevD.87.012008](https://doi.org/10.1103/PhysRevD.87.012008). arXiv: [1208.0949 \[hep-ex\]](https://arxiv.org/abs/1208.0949) (cit. on p. 20).
- [13] <https://www.slac.stanford.edu/grp/eg/minos/rootSYS/cvs/roofit/roostats/src/NumberCountingUtils.cxx> (cit. on pp. 20, 23).

- [14] [https://root.cern.ch/root/html526/RooStats\\_\\_NumberCountingUtils.html](https://root.cern.ch/root/html526/RooStats__NumberCountingUtils.html) (cit. on pp. 20, 23).
- [15] Kyle Cranmer and Itay Yavin. “RECAST: Extending the Impact of Existing Analyses”. In: *JHEP* 1104:038,2011 04 (2010), p. 038. DOI: [10.1007/JHEP04\(2011\)038](https://doi.org/10.1007/JHEP04(2011)038). arXiv: [1010.2506](https://arxiv.org/abs/1010.2506) [[hep-ex](#)] (cit. on p. 21).
- [16] Jean-Francois Arguin et al. *Search for gluino-mediated stop and sbottom pair production in events with b-jets and large missing transverse momentum*. Tech. rep. ATL-COM-PHYS-2015-319. Geneva: CERN, Apr. 2015. URL: <https://cds.cern.ch/record/2011623> (cit. on pp. 28, 54).
- [17] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011), p. 1554. DOI: [10.1140/epjc/s10052-011-1554-0](https://doi.org/10.1140/epjc/s10052-011-1554-0). arXiv: [1007.1727](https://arxiv.org/abs/1007.1727) [[physics.data-an](#)]. Erratum: *Eur. Phys. J. C* **73** (2013) 2501 (cit. on p. 70).
- [18] M. Baak et al. “HistFitter software framework for statistical data analysis”. In: *Eur. Phys. J. C* 75 (2015), p. 153. DOI: [10.1140/epjc/s10052-015-3327-7](https://doi.org/10.1140/epjc/s10052-015-3327-7). arXiv: [1410.1280](https://arxiv.org/abs/1410.1280) [[hep-ex](#)] (cit. on p. 70).
- [19] ATLAS Collaboration. *Jet Calibration and Systematic Uncertainties for Jets Reconstructed in the ATLAS Detector at  $\sqrt{s} = 13$  TeV*. ATL-PHYS-PUB-2015-015. 2015. URL: <https://cds.cern.ch/record/2037613> (cit. on p. 73).
- [20] Peter Zeiler Skands. “Tuning Monte Carlo Generators: The Perugia Tunes”. In: *Phys. Rev. D* 82 (2010), p. 074018. DOI: [10.1103/PhysRevD.82.074018](https://doi.org/10.1103/PhysRevD.82.074018). arXiv: [1005.3457](https://arxiv.org/abs/1005.3457) [[hep-ph](#)] (cit. on p. 74).
- [21] ATLAS Collaboration. “Measurements of fiducial cross-sections for  $t\bar{t}$  production with one or two additional  $b$ -jets in  $pp$  collisions at  $\sqrt{s} = 8$  TeV using the ATLAS detector”.

- In: *Eur. Phys. J. C* 76 (2016), p. 11. DOI: [10.1140/epjc/s10052-015-3852-4](https://doi.org/10.1140/epjc/s10052-015-3852-4). arXiv: [1508.06868](https://arxiv.org/abs/1508.06868) [[hep-ex](#)] (cit. on p. 76).
- [22] P. Kant et al. “HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions”. In: *Comput. Phys. Commun.* 191 (2015), pp. 74–89. DOI: [10.1016/j.cpc.2015.02.001](https://doi.org/10.1016/j.cpc.2015.02.001). arXiv: [1406.4403](https://arxiv.org/abs/1406.4403) [[hep-ph](#)] (cit. on p. 76).
- [23] W. Beenakker et al. “Squark and gluino production at hadron colliders”. In: *Nucl. Phys. B* 492 (1997), pp. 51–103. DOI: [10.1016/S0550-3213\(97\)00084-9](https://doi.org/10.1016/S0550-3213(97)00084-9). arXiv: [hep-ph/9610490](https://arxiv.org/abs/hep-ph/9610490) (cit. on p. 78).
- [24] A. Kulesza and L. Motyka. “Threshold resummation for squark-antisquark and gluino-pair production at the LHC”. In: *Phys. Rev. Lett.* 102 (2009), p. 111802. DOI: [10.1103/PhysRevLett.102.111802](https://doi.org/10.1103/PhysRevLett.102.111802). arXiv: [0807.2405](https://arxiv.org/abs/0807.2405) [[hep-ph](#)] (cit. on p. 78).
- [25] A. Kulesza and L. Motyka. “Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC”. In: *Phys. Rev. D* 80 (2009), p. 095004. DOI: [10.1103/PhysRevD.80.095004](https://doi.org/10.1103/PhysRevD.80.095004). arXiv: [0905.4749](https://arxiv.org/abs/0905.4749) [[hep-ph](#)] (cit. on p. 78).
- [26] Wim Beenakker et al. “Soft-gluon resummation for squark and gluino hadroproduction”. In: *JHEP* 12 (2009), p. 041. DOI: [10.1088/1126-6708/2009/12/041](https://doi.org/10.1088/1126-6708/2009/12/041). arXiv: [0909.4418](https://arxiv.org/abs/0909.4418) [[hep-ph](#)] (cit. on p. 78).
- [27] W. Beenakker et al. “Squark and gluino hadroproduction”. In: *Int. J. Mod. Phys. A* 26 (2011), pp. 2637–2664. DOI: [10.1142/S0217751X11053560](https://doi.org/10.1142/S0217751X11053560). arXiv: [1105.1110](https://arxiv.org/abs/1105.1110) [[hep-ph](#)] (cit. on p. 78).
- [28] Christoph Borschensky et al. “Squark and gluino production cross sections in pp collisions at  $\sqrt{s} = 13, 14, 33$  and  $100$  TeV”. In: *Eur. Phys. J. C* 74.12 (2014), p. 3174. DOI: [10.1140/epjc/s10052-014-3174-y](https://doi.org/10.1140/epjc/s10052-014-3174-y). arXiv: [1407.5066](https://arxiv.org/abs/1407.5066) [[hep-ph](#)] (cit. on p. 78).

- [29] ATLAS Collaboration. “Jet energy measurement with the ATLAS detector in proton–proton collisions at  $\sqrt{s} = 7$  TeV”. In: *Eur. Phys. J. C* 73 (2013), p. 2304. DOI: [10.1140/epjc/s10052-013-2304-2](https://doi.org/10.1140/epjc/s10052-013-2304-2). arXiv: [1112.6426](https://arxiv.org/abs/1112.6426) [hep-ex] (cit. on p. 80).