

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

SEARCHING FOR SUPERSYMMETRIC PARTICLES AT THE LARGE HADRON
COLLIDER USING THE ATLAS DETECTOR

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

OTHMANE RIFKI
Norman, Oklahoma
2016

SEARCHING FOR SUPERSYMMETRIC PARTICLES AT THE LARGE HADRON
COLLIDER USING THE ATLAS DETECTOR

A DISSERTATION APPROVED FOR THE
HOMER L. DODGE DEPARTMENT OF PHYSICS AND
ASTRONOMY

BY

Dr. Brad Abbott, Chair

Dr. Peter Parker, Outside Member

Dr. Mike Strauss

Dr. Chung Kao

Dr. Eric Abraham

© Copyright by OTHMANE RIFKI 2016
All Rights Reserved.

To my loving parents and sister.

Acknowledgements

I would like to acknowledge...

Table of Contents

Abstract	vii
1 Theoretical Background	4
2 The LHC and the ATLAS Experiment	5
3 The Region of Interest Builder	6
3.1 Introduction	6
3.2 Run 2 Challenges	6
3.3 ATLAS Dataflow Design	8
3.4 Region of Interest Builder	10
3.5 Performance in Run 2	11
3.6 Conclusion	13
3.7 Introduction	13
3.8 VMEbus based RoIB	16
3.8.1 Hardware implementation	16
3.8.2 System Performance and Evolution	17
3.9 PC based RoIB	18
3.9.1 The Common Readout Receiver Card	19
3.9.2 Readout System Firmware & Software	20
3.9.3 RoIB Software	21
3.10 Prototype Tests	22
3.10.1 Standalone Tests	23
3.10.2 Full System Tests	23
3.11 Outlook	24
4 Object Reconstruction	26
5 Data-driven techniques to estimate fake lepton backgrounds	27
5.1 The problem of fakes	27
5.2 Common processes for faking leptons	28
5.3 Monte Carlo Template Method	30
5.3.1 Motivation	30
5.3.2 Description of the method	31
5.3.3 Correction factors	32
5.3.4 Control regions	33
5.4 Matrix Method	35
5.4.1 Events with one object	38
5.4.2 Dynamic matrix method	40
5.4.3 Propagation of uncertainties	45
6 Search for new physics in events with two same sign leptons or three leptons and jets	46
6.1 Introduction	46
6.2 Signal models	49

6.2.1	Gluino pair production with slepton-mediated two-step decay $\tilde{g} \rightarrow q\bar{q}\ell\ell\tilde{\chi}_1^0$	51
6.2.2	Gluino pair production with gaugino-mediated two-step decay $\tilde{g} \rightarrow qq'WZ\tilde{\chi}_1^0$	53
6.2.3	Sbottom pair production with one-step decay $\tilde{b}_1 \rightarrow t\tilde{\chi}_1^\pm$	53
6.2.4	Gluino pair production with stop-mediated decay $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$	54
6.2.5	$\tilde{t}_1\tilde{t}_1^*$ with “three-same-sign leptons” signature	55
6.2.6	Non-Universal Higgs Models	56
6.2.7	Signal cross-sections and simulations	57
6.3	Data set and simulated event samples	59
6.4	Event selection	63
6.4.1	Pre-selection and event cleaning	63
6.4.2	Trigger strategy	65
6.4.3	Object definition	66
6.4.4	Jets	66
6.4.5	Electrons	68
6.4.6	Muons	69
6.4.7	Overlap removal	70
6.4.8	Missing transverse energy	70
6.5	Signal regions	71
6.6	Background estimation	81
6.6.1	Irreducible background estimation methods	81
6.6.2	Associate $t\bar{t} + W/Z$ production	81
6.6.3	Diboson $WZ, ZZ, W^\pm W^\pm$ production	83
6.6.4	Other rare processes	85
6.6.5	$W^\pm W^\pm$ production via Double Parton Scattering	86
6.6.6	Experimental uncertainties	89
6.6.7	Expected yields in the signal regions	92
6.6.8	Reducible background estimation methods	93
6.6.9	Matrix Method	93
6.6.10	Reducible background estimation methods	93
6.6.11	Matrix Method	93
6.6.12	Baseline-to-signal efficiency for fake muons	94
6.6.13	Baseline-to-signal efficiency for fake electrons	100
6.6.14	Baseline-to-signal efficiency for real leptons	104
6.6.15	MC Template Method	108
6.6.16	Background processes with charge-flipped electrons	112
6.6.17	Methodology	115
6.6.18	Measured rates	118
6.6.19	Systematic uncertainties	119
6.6.20	Expected yields in signal regions	121
6.6.21	Expected yields in signal regions	121
6.7	Systematic uncertainties	123
6.8	Results and interpretation	123
6.9	Conclusion	128
6.10	Auxiliary	129

Abstract

Well now this is my abstract.

Preface

...

In the spring of 2014, I received an ATLAS Support Center graduate fellowship to move to Argonne National Laboratory to work on hardware and physics analysis. This thesis is based on the work conducted in the period of May 2014 to May 2017. In my research, the supervision of Dr. Brad Abbott has been invaluable.

There is a very long list of people I would like to thank. I would like to thank the Argonne group for the warm welcome to the physics division at Argonne. The scientists at Argonne became colleagues and friends.

The working conditions either in Oklahoma, Argonne, or CERN have been excellent. I had the privilege of interacting with many interesting people at these institutes which I contributed significantly to building my independence as a physicist. I have been particularly fortunate of being part of Argonne National Laboratory, where I am able to work with many talented physicists. I would like to start with Alexander Paramonov and Sergei Chekanov who were the first to introduce me to physics analysis. Robert Blair, Jeremy Love, and Jinlong Zhang with whom I evolved the glorious region of interest builder of ATLAS responsible for processing every event recorded by the ATLAS detector. It was a real joy working with them and bringing the project to completion and taking data in ATLAS. It is fair to say that James Proudfoot became a mentor to me during my stay at Argonne and afterwards. Furthermore, my gratitude goes to Ximo Poveda, Julien Maurer, and Otilia Ducu who made my work within the same sign

SUSY group a real joy.

Finally, I would like to thank my family and friends for their support and acceptance of my absence from their lives. Last, but not least, I thank ... encouragement and patience during these years.

Othmane Rifki, May 2017

Introduction

Chapter 1

Theoretical Background

Chapter 2

The LHC and the ATLAS Experiment

Chapter 3

The Region of Interest Builder

3.1 Introduction

The ATLAS detector [?] is a multipurpose particle detector at the Large Hadron Collider (LHC) at CERN, Switzerland. After a 2-year shutdown for maintenance and upgrade, the LHC resumed operations starting Run 2 of the LHC in 2015. The ATLAS trigger and data acquisition (TDAQ) system was upgraded to simplify its architecture and increase its flexibility due to the increased energy and instantaneous luminosity (rate of proton collisions), and the addition of new detector systems. In Run 2 the recorded particle interactions, i.e. events, have a larger size and need to be processed at higher rates which required an upgrade of the dataflow component of the ATLAS TDAQ system. This system has been re-shaped in order to maximize the flexibility and efficiency of the data selection process leading to a different architecture of the ATLAS dataflow. In this proceeding, the Run 2 challenges motivating the upgrade will be covered along with a description of the new dataflow architecture and its performance.

3.2 Run 2 Challenges

The ATLAS TDAQ system reduces the proton interaction rate from 40 MHz to the ATLAS data storage capacity of about 1.5 kHz. A hardware First Level Trigger (L1) reduces the rate to 100 kHz and a software High Level Trigger (HLT)

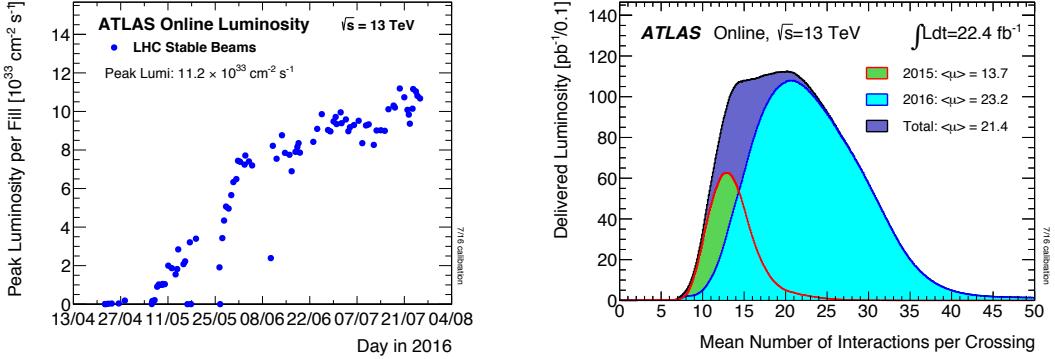


Figure 3.1: Run conditions during Run 2: ATLAS online luminosity (left), ATLAS online pileup (right) [?].

selects events for offline analysis. The function of the DAQ system is to efficiently buffer, transport, and record the events that were selected by the trigger system. Its performance is affected by the instantaneous luminosity that leads to busy events with multiple proton-proton interactions occurring in each bunch crossing, referred to as pileup. The high pileup results in a higher data volume collected by the detector that needs to be processed at the required rate to avoid exerting back-pressure on the L1 system. In Run 2, the LHC has exceeded the designed instantaneous luminosity of 10^{34} cm $^{-2}$ s $^{-1}$ leading to pileup of $\langle \mu \rangle = 30$ or more as shown in Figure 3.1. The L1 accept rate has also increased from 75 kHz in Run 1 to 100 kHz in Run 2 and the average output rate of the data logger system has increased from 400-600 Hz in Run 1 to about 3 kHz with 1.5 kHz for physics data. Moreover, there were new detectors that were added in Run 2 (Insertable B-layer (IBL), L1 topological trigger, Fast Tracker (FTK)) [?]

leading to an increase of 20% in the number of readout channels. To be able to deliver more rate to the High Level Trigger (HLT), the upgrade also targeted the Readout System (ROS)[?]. For the same reason the two level of the HLT system were collapsed into a single level which made the system more flexible allowing for incremental data retrieval and analysis. The dataflow network system was re-designed to increase its capacity and simplify its architecture[?].

3.3 ATLAS Dataflow Design

In Run 1, the farm was subdivided to several slices, with each slice managed by a dedicated supervisor. This layout has been dropped in favor of global management by a single farm master operating at 100 kHz referred to as the HLT supervisor (HLTSV). The Region of Interest Builder (RoIB) that assembles the RoIs previously implemented on a VMEbus system is now integrated with the HLTSV and the RoI building done in software. The change in the HLT architecture from two to one level required re-writing the HLT software and algorithms in such a way that each node in the farm can perform all processing steps. The handling of these processing steps is done by a single Data Collection Manager (DCM) process running on each HLT node to manage the L1 RoIs, the dataflow between the ROS and the HLT processing units (HLTPU), the event building processes, and the data logging. In the new architecture, the computing resources are managed more efficiently by balancing the utilization of all cluster nodes depending on the active HLT algorithms and by sharing the HLT code and

services to reduce memory and resource usage.

The dataflow network was simplified and upgraded to handle a larger data volume. A single network is used for RoI based access from the ROS, event building in the HLT processing nodes, and sending data for logging. A 10 GbE connectivity has been adopted throughout the dataflow system resulting in a factor of four increase in bandwidth between the data loggers and the permanent storage, and a 4×10 GbE output from each ROS PC to the core routers. The HLTSV and the HLT racks are all connected directly to each of the two core routers via 2×10 GbE connection. Each HLT rack is hosting up to 40 nodes connected by 2×1 GbE to the top-rack switches. The capacity of the routers can accommodate an increase in the number of HLT server racks and ROS PCs by a factor of two, which will be needed when the system scales as run conditions change. The core routers also provide load balancing and traffic shaping protocols [?] to distribute the data throughout the system more evenly. A duplication of core routers provide link redundancy at every level in case of link or switch failures.

To take advantage of multi-core architectures, the dataflow software is using multi-threaded software design for CPU consuming operations. The Input/Output of the dataflow is based on asynchronous communication using industry standard libraries such as the Boost::ASIO library. All the ATLAS software suite was switched to exclusively 64 bit operation in 2016.

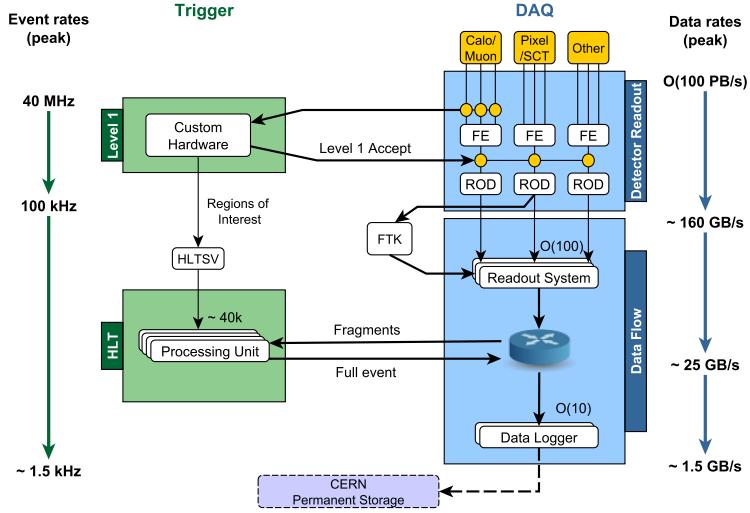


Figure 3.2: ATLAS TDAQ architecture.

3.4 Region of Interest Builder

The first step of the HLT processing is to run on the RoIs found by the L1 hardware trigger. These RoIs are collected and distributed to the HLT farm by the RoIB [?] which was the latest change to the ATLAS dataflow. The evolution of the RoIB system from a crate of custom VME-based electronics (VME-RoIB) to a commodity PC hosting a custom PCI-Express card (PC-RoIB) has been undertaken to increase the system performance, flexibility, and ease of maintenance. The functionality of the VME-RoIB previously possible only in FPGAs has now been implemented in a multi-threaded C++ software library. For each proton-proton collision that is accepted by the L1 trigger, the RoIB receives an ROI record from the custom inputs via S-Link. The RoIB assembles these records into a single record which is then forwarded to the HLTSV. The

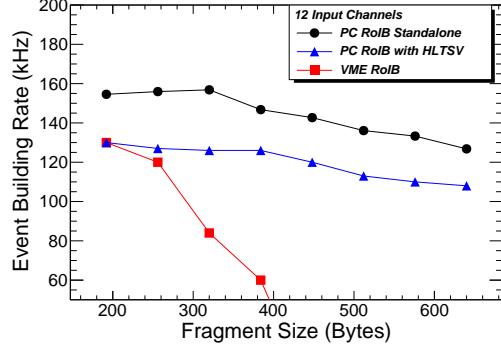


Figure 3.3: The event building rate as a function of the RoI record size in Bytes.

The rates are shown for a standalone application that implements a minimal interface for event building, the integrated RoIB software into an HLTSV process running within the full ATLAS TDAQ software suite, and for comparison the VME-RoIB performance.

HLTSV then distributes these single records to the HLT farm. The RoIB is also responsible for monitoring the data integrity of the incoming fragments and diagnostic performance of the system.

As shown in Figure 3.3, the performance of the PC-RoIB with realistic running ATLAS conditions is improved over the VME-RoIB particularly at high RoI sizes.

Figure 3.4 shows that the memory usage of the HLTSV is at the level of 5% and that the RoIB event assembly does not depend on pileup conditions.

3.5 Performance in Run 2

The reliable operation of the TDAQ system directly impacts the efficiency of the ATLAS experiment in recording the collisions delivered by the LHC. As a result, high data-taking efficiency is crucial for the ATLAS physics program. The

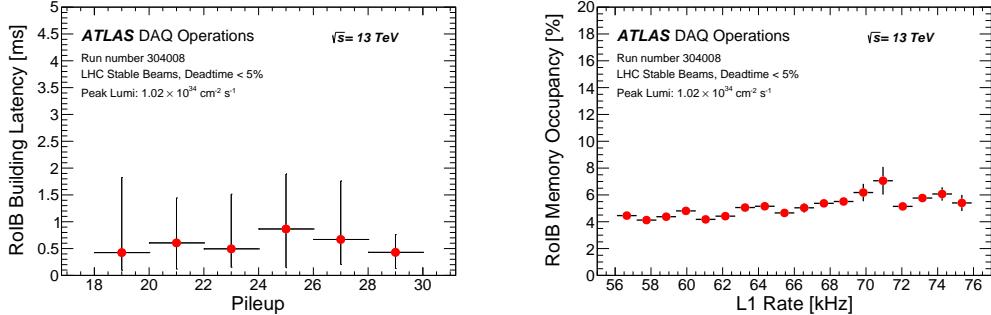


Figure 3.4: RoIB performance: RoIB building latency as a function of pileup (left), RoIB memory occupancy as a function of L1 rate (right).

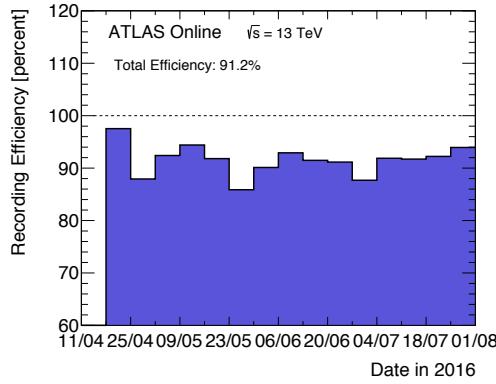


Figure 3.5: ATLAS recorded efficiency [?].

ATLAS recorded efficiency in 2016 is over 90%, as shown in Figure 3.5 with a negligible fraction of data loss due to the DAQ system. The new dataflow architecture is scaling well with the increased instantaneous luminosity during 2016 data-taking and is capable of handling larger pileup and thus larger event sizes. For illustration, Figure 3.6 shows the evolution of the average processing time per event and the event size where there is relatively mild increase as a function of pileup which will within the system capacity.

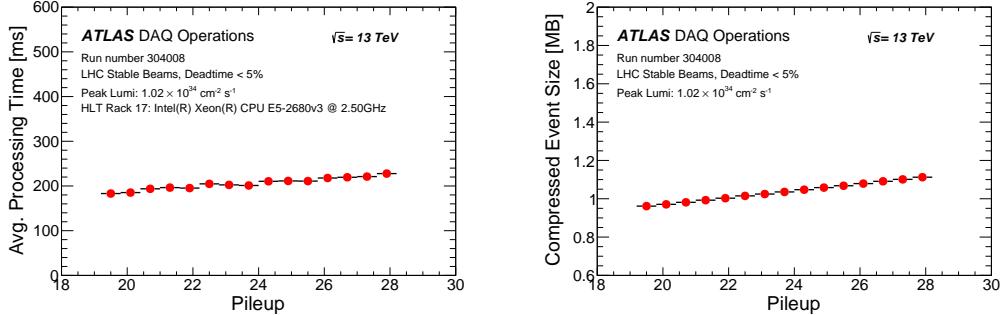


Figure 3.6: Performance in Run 2: Average processing time as a function of pileup (left), compressed event size as a function of pileup (right).

3.6 Conclusion

The dataflow system of ATLAS was re-shaped for Run 2 in order to handle the more demanding run conditions expected throughout the run. The new redesign profitted from the technological progress that took place in the last few years. As a result, the new system is considerably simplified, more performant, and scalable. Moreover, there is more headroom in performance to cope with more challenging run conditions of the LHC to ensure that ATLAS DAQ continues delivering physics data with high efficiency.

3.7 Introduction

The ATLAS [?] detector’s data acquisition system, illustrated in Figure 3.7, makes use of a multi-tiered trigger to reduce bandwidth from the LHC proton bunch crossing rate of 40 MHz to the 1 kHz written to disk [?]. The first tier (Level-1 or L1) [?], implemented in

real time with custom electronics, makes an early event selection to determine if any objects of interest are present and reduces the data flow to 100 kHz. The second tier, referred to as the High Level Trigger (HLT) [?], is implemented on a commodity computing cluster running custom triggering software. The HLT uses information from the hardware based L1 system to guide the retrieval of information from the Readout System (ROS) [?].

Jet, electromagnetic and tau clusters, missing transverse momentum (E_T^{miss}), $\sum E_T$, jet E_T , and muon candidate information from L1 determine detector Regions of Interest (RoIs) that seed HLT processing. These RoIs are provided to the HLT by a custom VMEbus based system, referred to as the Region of Interest Builder (RoIB) [?]. The RoIB collects data from L1 trigger sources and assembles the data fragments into a complete record of L1 RoIs. These RoIs are made available to the HLT to initiate event processing. In order to improve maintainability and scalability, and to minimize the amount of custom hardware needing to be supported, the RoIB will be implemented using commodity server hardware and an interface technology already deployed within the ATLAS Trigger and Data Acquisition (TDAQ) system. The approach of implementing the RoIB functionality in software has been investigated in the past and the conclusion at that time was that a software based approach is possible but requires a higher rate readout card [?]. Since data readout cards operating at high rates became available and the capabilities of computers have improved with the increase in CPU clock speed and number of cores, it became possible to implement the RoIB functionality using a PC based approach. The PC based RoIB must duplicate the

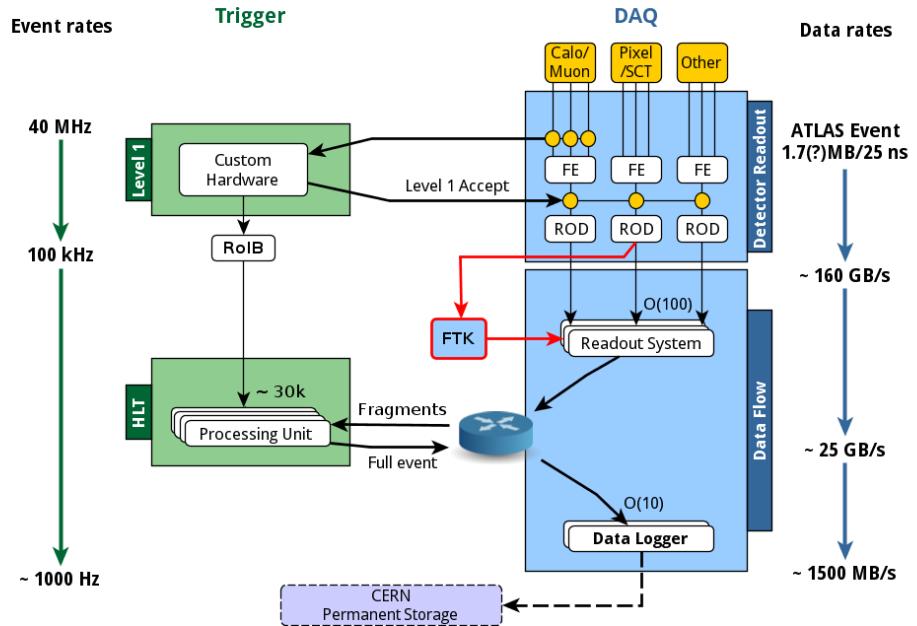


Figure 3.7: ATLAS TDAQ Architecture.

functionality of the VMEbus based RoIB which means that the PC based solution must receive and assemble the individual L1 fragments, and pass them as a single L1 result to the HLT. Modern computers have multicore CPU architectures with the possibility of running multi-threaded application, a feature which is being fully exploited in the RoIB software to achieve the desired performance of 100 kHz over 12 input links for fragment sizes of 400 bytes. This paper describes the evolution of the RoIB from the VMEbus based system to the PC based system and gives details on the hardware, firmware, and software designs used to achieve the full RoIB functionality.

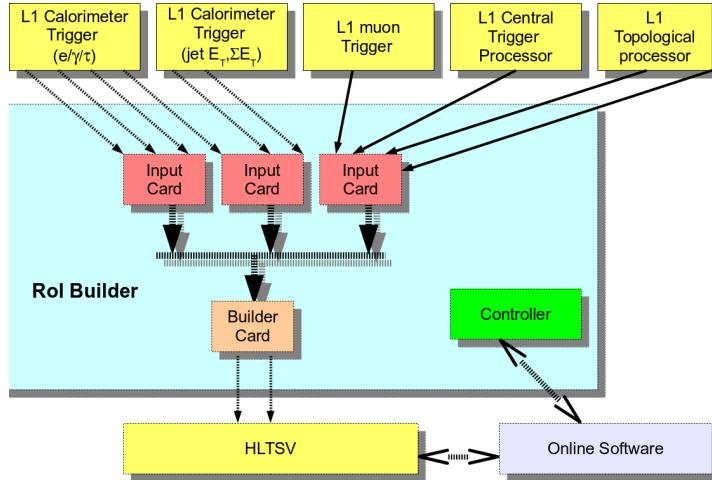


Figure 3.8: Block scheme of the RoI Builder and overview of connections to external systems. The custom input and builder cards and the controller, a commercially available single board computer, are installed in a single 9U VMEbus crate. The controller connects to the Control Network to interact with the rest of the data acquisition system.

3.8 VMEbus based RoIB

3.8.1 Hardware implementation

The RoIB is implemented as a custom 9U VMEbus system that includes a controller which configures and monitors the system along with custom cards that receive and assemble the event fragments and send them to the HLT. Figure 3.8 shows a block of the RoIB and its connection to external systems.

The RoIB contains four input cards and uses one builder card in the Run-2 configuration. Each input card accepts three inputs from L1 subsystems. The builder card assembles the input data of the events and passes the results via

two optical links to another receiver card in a PC running the HLT supervisor (HLTSV) application. The receiver card in the HLTSV is a TILAR card [?] that implements four PCIe Gen1 lanes to interface with the two optical links. The HLTSV manages the HLT processing farm by using L1 results provided by the RoIB, retrieves events from the ROS, assigns events to HLT farm nodes, and handles event bookkeeping including requesting removal of data from ROS storage when no longer required.

The fragments received by the RoIB are identified by a 32 bit identifier, the extended L1 ID (L1ID). The RoIB input cards use the L1ID and the number of outputs enabled to assign keys to the various fragments and send them to the output channel in the builder card that was assigned that key value. The input data is transferred over a custom J3 backplane. The backplane operates at 20 MHz and transfers 16 data bits per clock cycle simultaneously for up to 12 inputs. The total maximum data throughput is therefore 480 MB/s, 40 MB/s per input. The maximum size of any single fragment is limited to 512 bytes imposed by resources available in the FPGA firmware. The current RoIB input links are listed in Table 3.1.

3.8.2 System Performance and Evolution

The custom VMEbus based RoIB operated reliably during the first run of the LHC, however, it is desirable to have a more flexible RoIB. In addition, the RoIB is getting close to its design limitation, as seen in Figure 3.12. For fragments of 400 bytes and inputs from eight L1 systems, referred to as channels, the current

Table 3.1: L1 input sources to the RoIB.

Source	Links
Central Trigger Processor (CTP)	1
L1 calorimeters ($e/\gamma, \tau, \text{jet}, \sum E_T$)	6
Muon Trigger to CTP Interface (MUCTPI)	1
Topological processor (L1Topo)	2
Spare	2

RoIB rate limit is 60 kHz which is below the required 100 kHz at L1. While the current fragment size coming from L1 are around 160 bytes, the sizes are expected to grow due to the increase of instantaneous luminosity and the complexity of L1 triggers. The current VMEbus system will be replaced by a PCI-express card hosted in the HLTSV PC with the possibility to upgrade the commodity hardware (e.g. ability to upgrade CPUs). The new configuration simplifies the readout architecture of ATLAS. The targeted rate for event building is 100 kHz over 12 input channels for fragment sizes in the order of 400 bytes.

3.9 PC based RoIB

A custom PCIe card developed by the ALICE collaboration, the Common ReadOut Receiver Card (C-RORC) [?], was deployed as an upgraded detector readout interface within the ATLAS ROS with ATLAS specific firmware and software called the RobinNP [?]. The new PC based RoIB uses the RobinNP firmware and

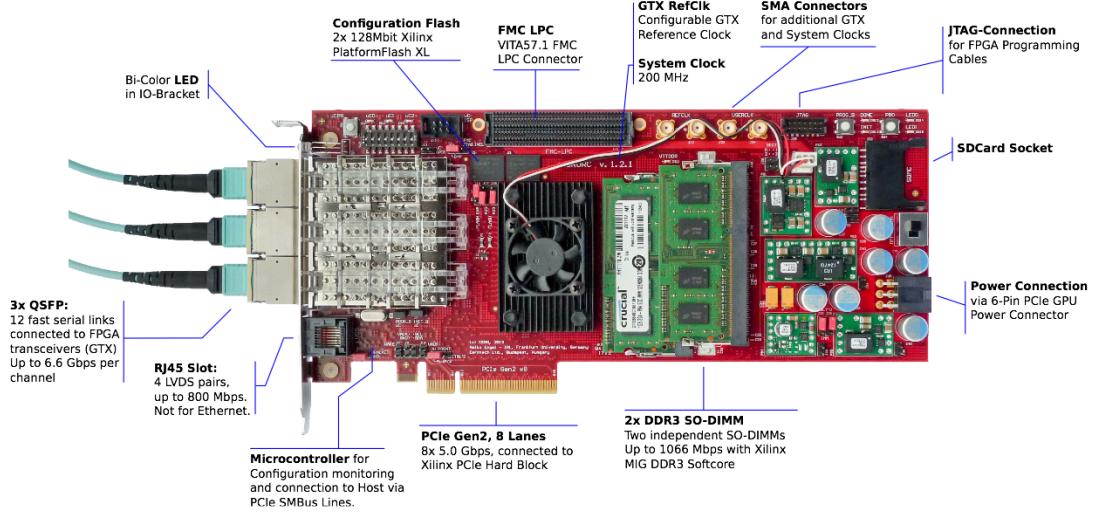


Figure 3.9: Photo of the C-RORC board with the major components and features annotated [?].

a dedicated API to facilitate the implementation of the RoIB functionality on a commodity PC. In this section, we describe the C-RORC hardware as well as the RobinNP firmware, API, and the event building software.

3.9.1 The Common Readout Receiver Card

The C-RORC implements 8 PCIe Gen1 lanes with 1.4 GB/s bandwidth to the CPU fed via 12 optical links each running 200 MB/s on 3 QSFP transceivers. It utilizes a single Xilinx Virtex-6 series FPGA that handles data input from the 12 links and buffers the data in two on-board DDR3 memories. It is also capable of processing and initiating DMA transfer of event data from the on-board memory to its host PC's memory. The major components of the C-RORC are annotated in the picture shown in Figure 3.9.

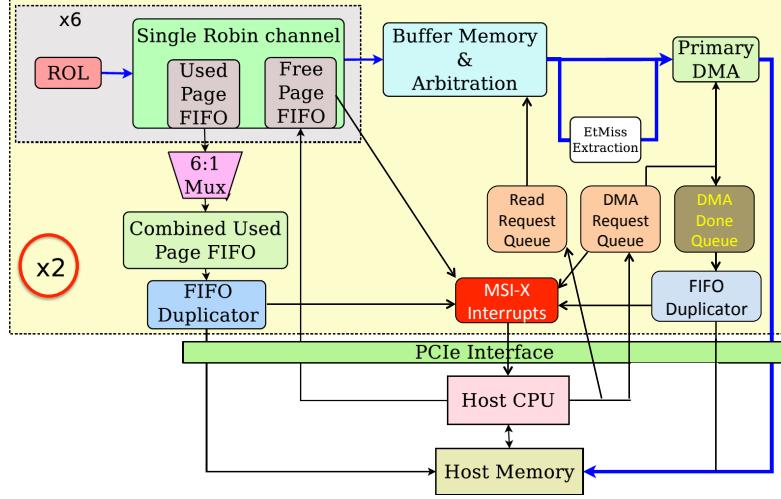


Figure 3.10: RobinNP firmware organization and flow of data from host CPU to the firmware (by means of programmed I/O) and from the firmware to the host memory (by means of DMA).

3.9.2 Readout System Firmware & Software

The RobinNP firmware used for the RoIB is identical to that used in the ATLAS ROS[?]. As shown in the schematic of Figure 3.10, the logic is divided into two functional blocks, known as sub-ROBs, each servicing six input links and one DDR3 memory module. Event data fragments arriving via a link are subjected to a range of error checks before being stored in the memory module for the relevant sub-ROB. At the same time a token representing the address of a region of the memory, referred to as a page, is passed to a listening software process via a ‘FIFO duplicator’. To avoid a costly read across the PCIe bus, data is continuously streamed from firmware to software via a chain of firmware and software FIFOs. Notification of new data arriving in the software FIFO is managed via coalesced

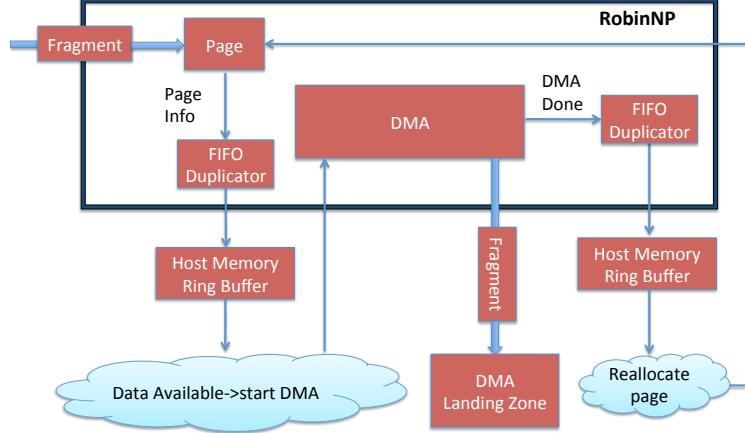


Figure 3.11: Layout of the readout system firmware and software specific to the RoIB.

interrupts to allow for efficient use of CPU resources. For the RoIB application, the receipt of page information immediately triggers a DMA of fragment data from the RobinNP memory into the host PC memory. The fragments are then passed via a queue (one per sub-ROB) to the RoIB process along with any relevant fragment error information. A schematic of this shortened dataflow path is presented in Figure 3.11. The API for the RoIB process consists of these queues, return queues for processed pages now available for re-use and a configuration interface. The software is implemented with multiple threads each handling specific tasks such as supply of free pages, receipt of used pages, DMA control and bulk receipt of fragment data.

3.9.3 RoIB Software

The HLTSV is a multi-threaded application that obtains a L1 result from a variety of possible input sources and exchanges information with the rest of the

HLT computing farm. For the RoIB, the L1 source is a RobinNP interface that performs fragment assembly and is used as a plug-in to the HLTSV application. The RobinNP plug-in has two receive threads, each thread services six channels by pulling fragments from the RobinNP on-board memories to the host PC. Fragments with the same L1ID are copied to a contiguous memory space and a queue of completed events is prepared. Upon request by the HLTSV, a pointer to the contiguous memory space is passed back to the HLTSV process for further handling. In order to optimize concurrent access to RoIB data structures, containers from the Intel threading building block (TBB) library were used. These containers allow multiple threads to concurrently access and update items in the container while maintaining high performance.

3.10 Prototype Tests

In order to understand the requirements for the underlying server PC, a validation system based on Intel(R) Xeon(R) CPU E5-1650 v2 @ 3.5 GHz with six cores is being used to perform tests of the PC based RoIB. The goal is to perform software based fragment assembly at a rate of 100 kHz over 12 channels for a typical fragment size of 400 bytes. The current system offers flexibility in terms of the fragment size allowed which was not the case in the VMEbus based RoIB. The initial tests were performed with a standalone application that implements a minimal interface for event building. Once the system was validated, the relevant code modules were integrated into an HLTSV process running within the full

ATLAS TDAQ software suite with appropriately scaled test hardware to represent the remaining elements of the system.

3.10.1 Standalone Tests

The goal was to test input/output bandwidth limitations of the RobinNP and the rate of event building. Initial performance testing used a standalone RobinNP application and an external source that emulates the L1 trigger data in the form of 32-bit word fragments with 12 channels. In this test, the host PC was running the assembly routine with a single threaded application. Figure 3.12 shows the input rate without event building as a function of fragment size. For 400 byte fragments the input rate to the RobinNP is 215 kHz. The same figure shows the event building rate which is 150 kHz. This performance shows that the event building at the required rate of 100 kHz with 12 channels is achievable in a standalone application.

3.10.2 Full System Tests

Since the HLTSV is performing tasks other than the event building, there is overhead associated with additional operations that reduces the performance. For this reason, we use the full ATLAS TDAQ software in a test environment that emulates the major components of the ATLAS data acquisition system shown in Figure 3.7. The setup includes an emulated input from L1 trigger sources, the HLTSV and other PCs to simulate the HLT computing farm, and the ROS that buffers the full event data. In this test setup, an external source sends data

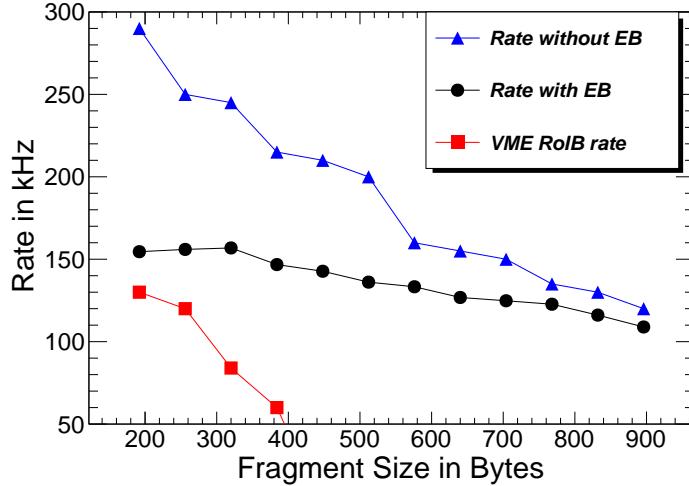


Figure 3.12: Rate as a function of the fragment size (in bytes) with external source that emulates the L1 trigger input. The rates shown are for the input rate to the RobinNP without event building (EB) (triangle), rate with EB (circle), and for comparison, the current VMEbus RoIB rate is also shown (square).

that emulates L1 RoIs via 12 links connected to the RobinNP hosted by the HLTSV. When the HLTSV requests a built RoI event, the software RoIB plug-in provides the RoI event which will be used to seed requests for the event data to be processed. Figure 3.13 shows an event building rate of 110 kHz measured with 400 byte fragments with the HLTSV application in a setup close to the ATLAS TDAQ system.

3.11 Outlook

The RoIB will evolve from the VMEbus based system to the PC based system using a PCI-Express card and firmware shared with the ATLAS ROS. The new system will add flexibility and improve maintainability of the ATLAS TDAQ

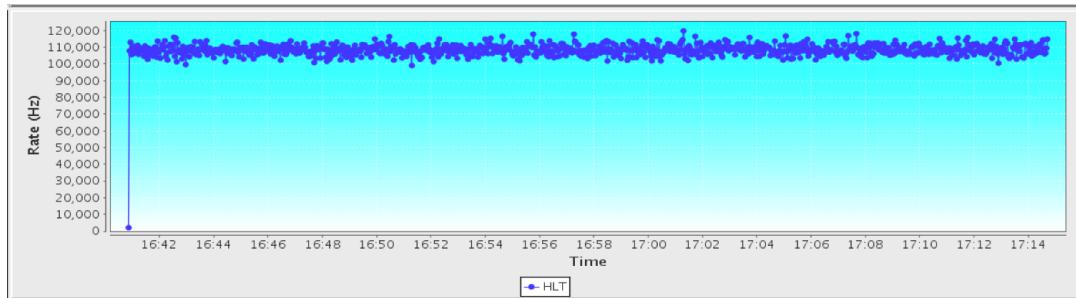


Figure 3.13: Screenshot of a monitoring tool which shows the HLTSV processing rate using the ATLAS TDAQ software .

system. As the technology evolves, the PCs and CPUs can be upgraded and more channels can be included by adding more RobinNP cards while maintaining high readout rates. A full integration test of the readout performance of the ATLAS TDAQ system with the PC based RoIB will be performed during the 2015-2016 LHC winter shutdown in preparation for a system evolution.

Chapter 4

Object Reconstruction

Chapter 5

Data-driven techniques to estimate fake lepton backgrounds

5.1 The problem of fakes

The reconstructed objects (leptons, photons, b -jets, etc.) in a collision event are used to perform a wide range of SM measurements or searches for evidence of BSM physics. The assumption is that these objects are ‘real’ representing the desired particles in the final state used in the analysis. In practice, the reconstructed objects might not be always ‘real’. In fact, they may be something completely different that was mistakenly reconstructed as the desired object, called ‘fake’. For the purpose of the analysis presented in this thesis, the focus is on ‘fake’ leptons. To illustrate the problem, a hadronic jet may deposit more energy in the electromagnetic calorimeter than the hadronic calorimeter, or that it leaves a narrow deposit of energy leading the reconstruction algorithms to mistake this jet for an electron. From the analysis point of view, the ‘fake’ electron will pass all the selection criteria and will be indistinguishable from a ‘real’ electron. It is important for the analysis that requires a reconstructed electron to model the fake electron background to get a sound result. This example was given with electrons, but can be generalized to muons as well. In short, any analysis that uses leptons in the final state must account for the ‘fake’ lepton background. This background can be more or less important depending on the detector, the analysis selection,

and the number of leptons required. To estimate this background it is important to first understand what type of processes lead to fake leptons.

5.2 Common processes for faking leptons

The reconstruction of ‘fake’ leptons can be an instrumental effect related to the inability to identify the object based on its measured properties by the detector. In this case, the reconstructed lepton is not a real lepton and the production process will be different for electrons and muons.

The reconstruction of electrons relies on the observation of well aligned particle hits in the layers of the ID that are consistent with an energy deposition in the EM calorimeter. Photons can mimick this signature since they deposit energy in the EM calorimeter that happens to be alligned with a charged track. A jet for example containing charged and neutral pions can lead to such scenarios. It is possible for the jet to have one charged pion leaving a track similar to that of an electron. The decay of π^0 mesons to photons in this jet can deposit energy in the EM calorimeter leading to the required signature. Another mechanism that can lead to fake electrons is the emission of photons via Bremsstrahlung from high energy muons. The muon track can be mistaken for that of an electron and the photons interact with the EM calorimeter leading to a signature similar to that of electrons. An additional process is that of photon conversions into a e^+e^- .

The reconstruction of muons relies on the observation of tracks from the ID matched to tracks from the muon spectrometer. It is possible for charged hadrons

with long lifetime to traverse the calorimeter layers and leave hits in the muon spectrometer. These hits may coincide with other hits from the ID due to the random activity in the event. As a result, a muon can get reconstructed. Another instance may occur when pions or kaons decay in-flight to muons in the muon spectrometer and happen to align with the primary vertex.

The leptons that are used in the physics analyses must be coming from the hard scatter, generally referred to as prompt leptons. There is another case where the reconstructed lepton is a real lepton but is not a lepton coming from the hard interaction, referred to as non-prompt leptons. Non-prompt leptons can be produced from heavy flavor meson decays with a low energy activity around the lepton which allows it to pass isolation requirements. A good example of this type of process is the semi-leptonic decay of top quark pair which contribute to final states with two leptons.

For the rest of the thesis, the fake leptons will be referred to as fake/non-prompt (FNP) leptons. There are several methods used to perform the estimation of FNP lepton backgrounds. A method that the author developed will be described next along with a standard method for estimating this type of backgrounds. The benefit of having two methods for estimating the FNP lepton background is to have enough confidence in the final estimate. The two methods use different assumptions which naturally leads to a more robust estimation of this difficult background. Moreover, the final estimate of the FNP lepton background is taken as a statistical combination of the estimates from the two methods leading to a reduction of the systematic uncertainties on the estimate.

5.3 Monte Carlo Template Method

5.3.1 Motivation

The processes leading to FNP leptons depend on the selection applied in the analysis. For instance, a selection with same-sign leptons will have contributions from top quark pair production ($t\bar{t}$) or the associated production of a vector boson and jets ($W+\text{jets}$ or $Z+\text{jets}$). These processes cannot give two leptons of the same electric charge unless there is a charge mis-measurement (mainly affecting electrons) or that a FNP lepton was produced. It is possible to generate the processes that can contribute to a FNP lepton, such as $t\bar{t}$ or $V+\text{jets}$, with Monte Carlo event generators processed through Geant4 detector simulation of the ATLAS detector. This approach will yield an estimate however it might not be reliable. For instance, the detector simulation itself might not reproduce the true behavior of the interaction of the physics objects with the detector, particularly when looking at rare processes such as the production of FNP leptons. The second limitation is in the generation of enough MC events to probe the region of the phase space targetted by the analysis which affects the statistical uncertainties in the estimates. The latter concern is addressed by ensuring that the simulations for the major backgrounds ($t\bar{t}$ and $V+\text{jets}$) have much higher event count than the corresponding number of events observed in the data sample. In fact, these backgrounds have a large number of simulated events because they are important for many analyses (including SM measurements and BSM searches). The rest of

the section will concentrate on addressing the former limitation.

5.3.2 Description of the method

The MC template method relies on the correct modelling of FNP leptons kinematics in MC simulation to extrapolate background predictions from control regions to the signal regions. The method assumes that the kinematic shapes for each source of FNP lepton is correctly modeled in the simulations, and the normalization for each source is extracted in a combined fit to data control regions. The number of normalization factors depend on the number of identified origins of the FNP lepton in the signal regions and the control regions are designed to constrain these factors in regions enriched with FNP leptons from the same origin.

To illustrate the approach, we describe the application of the method in SS/3L analysis later described in this thesis. The processes of interest that may lead to a FNP lepton or a charge flip are $t\bar{t}$ and $V+jets$. FNP leptons are classified using an algorithm that navigates the generator particle record to determine where the FNP lepton is originating from. The lepton is classified as either an electron or a muon that is prompt from decays of on-shell W and Z bosons, non-prompt from a heavy flavor b decay (HF), or fake from mis-identification of a light flavor jet or a photon (LF). In the case of an electron, we further classify the prompt electrons to prompt electrons with the correct charge or with a charge mis-measurement, commonly named charge flip. In total, five categories referred to as MC templates are constructed following the classification illustrated in figure 5.1.

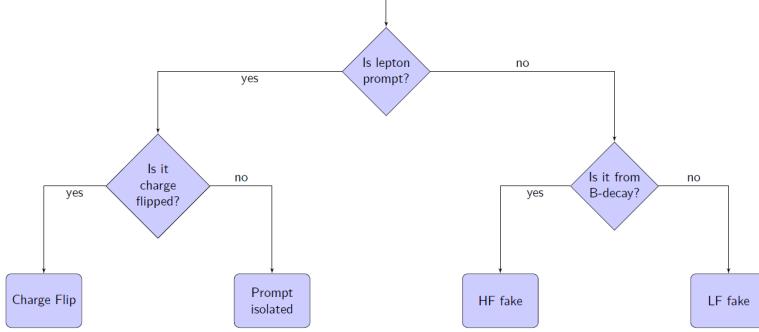


Figure 5.1: Lepton classification.

5.3.3 Correction factors

The FNP estimate relies on kinematic extrapolation using processes expected to contribute via FNP leptons from control regions with low jet multiplicity and E_T^{miss} , to the signal regions that require high jet multiplicity and E_T^{miss} . The control regions are chosen to separate FNP leptons from HF origins and FNP leptons from LF origins. For instance, a control sample characterized by the presence of a b -jet will be enriched in processes with one FNP lepton that is coming from a HF decay, while a sample characterized by the absence of a b -jet will have one FNP lepton from LF decay. The presence of one FNP lepton in the control sample allows the correction of the production rate of these FNP leptons by performing a fit to data.

For example, if a $Z \rightarrow \mu\mu + \text{LF jet}$ event is reconstructed as a $\mu^+\mu^-e^+$ event, then the electron is fake. Therefore, a correction of $\text{LF jet} \rightarrow e$ ($\text{Fr}(\text{LF} \rightarrow e)$) is applied to the rate of $\mu\mu e$ events. The correction $\text{Fr}(\text{LF} \rightarrow e)$ is constrained by a fit to data in control regions dominated by LF jet $\rightarrow e$ type fakes. Similarly, three other corrections are defined as LF jet $\rightarrow \mu$ ($\text{Fr}(\text{LF} \rightarrow \mu)$), HF jet $\rightarrow e$

(Fr(HF→e)), HF jet → μ (Fr(HF→ μ)). An additional correction is applied to correct the charge flip rate predicted by simulation. For example, a $Z \rightarrow e^+e^-$ event is reconstructed as e^+e^+ or e^-e^- . The simulation takes into account the charge flip rate but it might be off. The charge flip (Cf(e)) correction derived from a data fit is expected to recover this mis-modeling. The charge flip rate only concern electrons as the muon charge flip rate is negligable.

A likelihood fit is defined as the product of the Poisson probabilities describing the observed events in the binned distributions from the expected number of events rescaled by the five multipliers which are left free to float in the fit. These multipliers are applied to the MC predictions in the signal regions to obtain an estimation of the charge flip and FNP backgrounds.

5.3.4 Control regions

The corrections depend on the simulated sample, the reconstructed final state, and the flavor of the leptons. As a result, care must be taken when designing the control regions used to perform the fit of the FNP leptons and electron charge flip templates. For instance, each template needs to be constrained in a selection that is representative of the processes leading to FNP leptons and charge flip electrons present in the kinematic region targetted by the search for BSM physics.

In the SS/3L analysis discussed in this thesis the control regions are defined with at least two same-sign leptons, $E_T^{\text{miss}} > 40$ GeV, two or more jets. This preselection ensures that the FNP leptons are not from fakes originating from QCD like event topologies. They are further split in regions with or without

b -jets to constrain the HF and LF leptons respectively. In addition, they are also split with different flavours of the same-sign lepton pair ee, e μ , and $\mu\mu$, giving a total of six control regions. Any event entering the signal region is vetoed. The ee channel will constrain the charge flip correction factor, fake leptons from LF decays in the selection without b -jets, and non-prompt decay from HF in the selection with b -jets. The $\mu\mu$ channel will constrain the muon fake rates in the LF and HF decays for the selection without or with b -jets, respectively. The e μ channel will constrain both the electron and muon fakes for events containing both lepton flavors.

The six distributions are chosen for variables that provide the best separation between processes with prompt leptons and processes with FNP leptons and charge flip and are shown before and after the fit in Figures 6.20-6.22 and Figures 6.21-6.23, respectively.

The minimization of the negative log likelihood using the MINUIT package leads to the multipliers shown in Tables 6.17 and 6.18. The tables represent the multipliers obtained from the fit upon using two different parton showers, POWHEG-Box and SHERPA for the processes that lead to FNP leptons and charge flips. The systematic uncertainty is obtained by varying the generator from POWHEG-Box to SHERPA and evaluating the impact on the expected background from FNP and charge flip leptons. It is found to be the dominant contribution to the systematic uncertainty of the method (up to 80%). The uncertainties in the multipliers themselves correspond to how much the parameter needs to be varied for a one standard deviation change in the likelihood function. This uncertainty

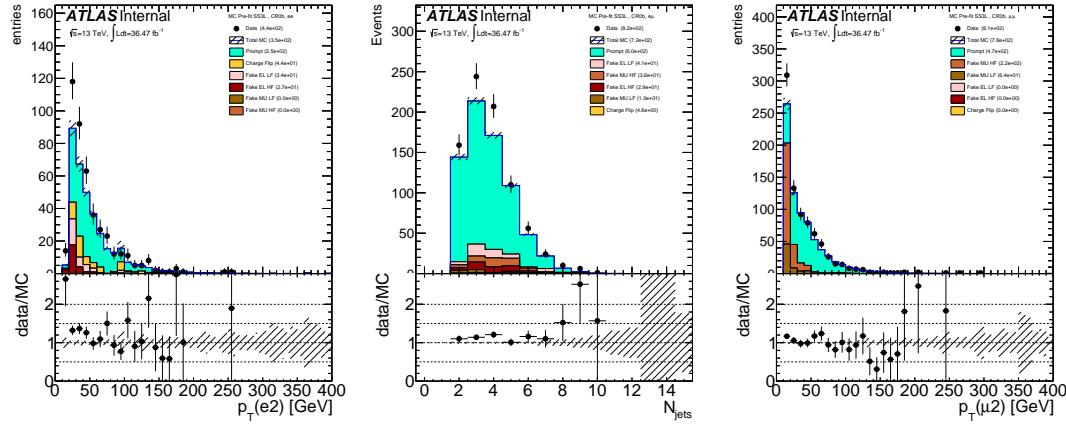


Figure 5.2: Pre-fit distributions for ee channel (left), for $e\mu$ channel (middle), and for $\mu\mu$ channel (right) from CR0b that were used in the fit to extract the FNP lepton and charge flip multipliers. The generator used in these plots is Powheg. The hashed band represents the sum of systematic uncertainties on the predictions.

takes into account the limited number of simulated events and is included as a systematic uncertainty on the expected number of background events.

5.4 Matrix Method

The FNP leptons do not often pass one of the lepton selection criteria but have non-zero impact parameter, and are often not well-isolated. These selection requirements are key ingredients to control the FNP leptons. The number of events with at least one FNP lepton is estimated using two classes of leptons: a real-enriched class of “tight” leptons corresponding to signal leptons and a fake-enriched class of “loose” leptons corresponding to candidate leptons with

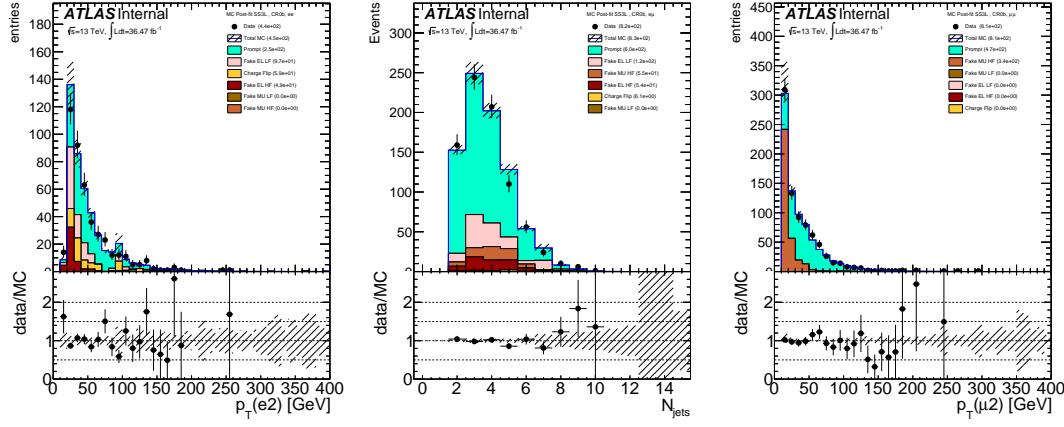


Figure 5.3: Post-fit distributions for ee channel (left), for $e\mu$ channel (middle), and for $\mu\mu$ channel (right) from CR0b that were used in the fit to extract the FNP lepton and charge flip multipliers. The generator used in these plots is Powheg. The hashed band represents the sum of systematic uncertainties on the predictions.

Table 5.1: The FNP and charge flip multipliers obtained after minimizing the likelihood function using Pythia. The uncertainty in the multipliers takes into account the limited statistics of simulated events.

Category	Multiplier	Uncertainty
chFlip	1.49	0.58
HF EL	2.80	0.98
LF EL	2.89	0.88
HF MU	1.59	0.31
LF MU	1.00	1.34

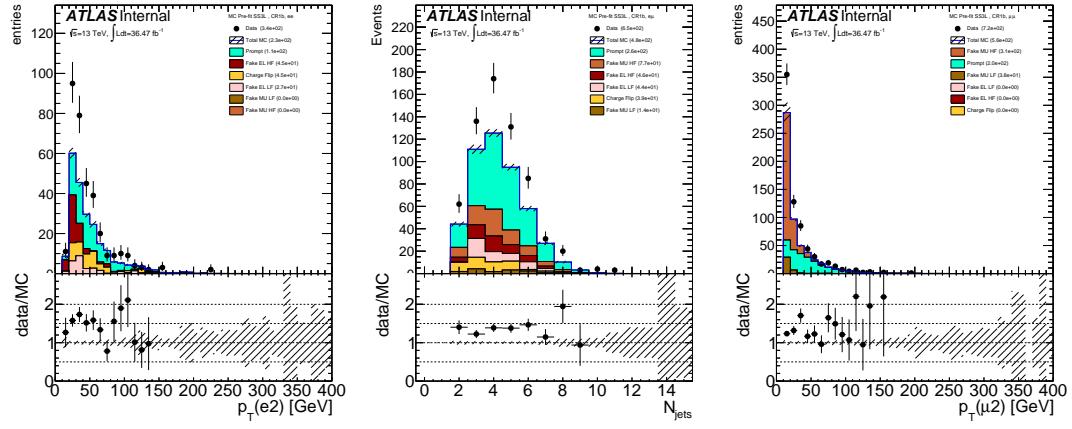


Figure 5.4: Pre-fit distributions for ee channel (left), for $e\mu$ channel (middle), and for $\mu\mu$ channel (right) from CR1b that were used in the fit to extract the FNP lepton and charge flip multipliers. The generator used in these plots is Powheg. The hashed band represents the sum of systematic uncertainties on the predictions.

Table 5.2: The FNP and charge flip multipliers obtained after minimizing the likelihood function using Sherpa. The uncertainty in the multipliers takes into account the limited statistics of simulated events.

Category	Multiplier	Uncertainty
chFlip	1.34	0.58
HF EL	2.40	0.85
LF EL	1.83	1.04
HF MU	1.17	0.16
LF MU	2.40	0.81

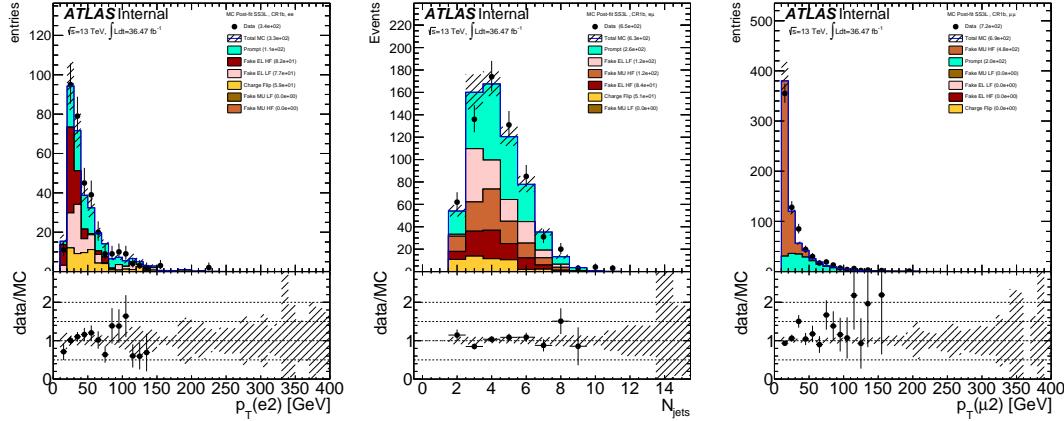


Figure 5.5: Post-fit distributions for ee channel (left), for $e\mu$ channel (middle), and for $\mu\mu$ channel (right) from CR1b that were used in the fit to extract the FNP lepton and charge flip multipliers. The generator used in these plots is Powheg. The hashed band represents the sum of systematic uncertainties on the predictions.

relaxed identification criteria¹. In the next sections, a description of the simplest form of the matrix method will be given with events containing one object. Then a generalized treatment that can handle events with an arbitrary number of leptons in the final states will be discussed.

5.4.1 Events with one object

Given the probabilities ε/ζ for a real/FNP candidate lepton to satisfy the signal lepton criteria, one can relate the number of events with one candidate lepton passing/failing signal requirements ($n_{\text{pass}}/n_{\text{fail}}$) to the number of events with one real/FNP signal leptons ($n_{\text{real}}/n_{\text{FNP}}$):

¹Signal leptons are leptons satisfying the signal lepton definition, while the candidate leptons are leptons satisfying some pre-selection cuts and usually passing the overlap removal requirements as discussed in the analysis section ??.

$$\begin{pmatrix} n_{\text{pass}} \\ n_{\text{fail}} \end{pmatrix} = \begin{pmatrix} \varepsilon & \zeta \\ 1 - \varepsilon & 1 - \zeta \end{pmatrix} \begin{pmatrix} n_{\text{real}} \\ n_{\text{FNP}} \end{pmatrix}; \quad (5.1)$$

allowing to determine the unknown number of events n_{FNP} from the observed n_{pass} and n_{fail} given measurements of the probabilities ε/ζ .

The predictive power of the matrix method comes from the fact that the real and FNP leptons have different composition in the two collections of tight and loose objects leading to $\varepsilon \neq \zeta$. In fact, the tight lepton collection will be dominated by real objects while the loose region will be dominated by fake objects. As a result, the inequality $\varepsilon >> \zeta$ will always hold true which guarantees that the matrix in Eq. 5.1 is invertible and gives positive estimates.

The next step is to invert the relation in Eq. 5.1 to obtain

$$\begin{pmatrix} n_{\text{real}} \\ n_{\text{FNP}} \end{pmatrix} = \frac{1}{\varepsilon - \zeta} \begin{pmatrix} \bar{\zeta} & -\zeta \\ -\bar{\varepsilon} & \varepsilon \end{pmatrix} \begin{pmatrix} n_{\text{pass}} \\ n_{\text{fail}} \end{pmatrix}; \quad (5.2)$$

where $\bar{\varepsilon} = 1 - \varepsilon$ and $\bar{\zeta} = 1 - \zeta$. The FNP lepton component is:

$$n_{\text{FNP}} = \frac{1}{\varepsilon - \zeta} ((\varepsilon - 1) n_{\text{pass}} + n_{\text{fail}}). \quad (5.3)$$

However, the quantity of interest is the expected FNP lepton background that passes the tight selection criteria: $n_{\text{pass} \cap \text{FNP}} = \zeta n_{\text{FNP}}$. To obtain this quantity, the identity from Eq. 5.1 is used to get:

$$n_{\text{FNP}} = \frac{\zeta}{\varepsilon - \zeta} ((\varepsilon - 1) n_{\text{pass}} + n_{\text{fail}}). \quad (5.4)$$

The linearity of Eq. 5.4 with respect to n_{pass} and n_{fail} allows the method to be applied on an event-by-event, effectively resulting into a weight being assigned to each event. By defining

$$n_{\text{pass}} = \sum_{\text{all events}} \mathbb{1}_{\text{pass}}, \quad n_{\text{fail}} = \sum_{\text{all events}} \mathbb{1}_{\text{fail}}, \quad \mathbb{1}_{\text{fail}} = 1 - \mathbb{1}_{\text{pass}},$$

where $\mathbb{1}_{\text{pass(fail)}} = 1$ if the object pass (fail) the tight selection requirement and $\mathbb{1}_{\text{pass(fail)}} = 0$ otherwise. Eq. 5.4 can be written as

$$n_{\text{FNP}} = \sum_{\text{all events}} \left\{ \frac{\zeta}{\varepsilon - \zeta} (\varepsilon - \mathbb{1}_{\text{pass}}) \right\} = \sum_{\text{all events}} \omega$$

where

$$\omega = \frac{\zeta}{\varepsilon - \zeta} (\varepsilon - \mathbb{1}_{\text{pass}}) \tag{5.5}$$

is the weight to be assigned to each event in the case of one FNP lepton in the event. The generalization of this formalism to higher dimensions with multiple objects will be covered next.

5.4.2 Dynamic matrix method

The one lepton case readily generalizes to events with more than one lepton in a formalism that can handle an arbitrary number of leptons in the event. The method should be applied event-by-event, effectively resulting into a weight being assigned to each event. The predicted yield of events with FNP leptons is simply the sum of weights. A general formula will be derived starting from the two

objects case then specific examples will be given to illustrate the application of the method.

If two objects are present in the event, the probabilities ε/ζ will depend on the kinematic properties of these objects. Typically the probabilities will vary as a function of p_T and $|\eta|$. For this reason, the probabilities will be different and will have an index to identify the object under study: ε_i/ζ_i where $i = 1, 2, \dots$. An identity similar to Eq. 5.1 can be formed for two objects with a change in notation for simplicity:

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \Lambda \times \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix}, \quad (5.6)$$

where $(N_{RR}, N_{RF}, N_{FR}, N_{FF})$ are the number of events with respectively two real, one real plus one FNP (two terms), and two FNP leptons before applying tight cuts, respectively, and $(N_{TT}, N_{TL}, N_{LT}, N_{LL})$ are the observed number of events for which respectively both lepton pass the tight cut, only one of them (two terms), or both fail the tight cut, respectively.

Λ is given by:

$$\Lambda = \begin{pmatrix} \varepsilon_1 \varepsilon_2 & \varepsilon_1 \zeta_2 & \zeta_1 \varepsilon_2 & \zeta_1 \zeta_2 \\ \varepsilon_1 (1 - \varepsilon_2) & \varepsilon_1 (1 - \zeta_2) & \zeta_1 (1 - \varepsilon_2) & \zeta_1 (1 - \zeta_2) \\ (1 - \varepsilon_1) \varepsilon_2 & (1 - \varepsilon_1) \zeta_2 & (1 - \zeta_1) \varepsilon_2 & (1 - \zeta_1) \zeta_2 \\ (1 - \varepsilon_1) (1 - \varepsilon_2) & (1 - \varepsilon_1) (1 - \zeta_2) & (1 - \zeta_1) (1 - \varepsilon_2) & (1 - \zeta_1) (1 - \zeta_2) \end{pmatrix}$$

which can also be written in terms of a Kronecker product in Eq. 5.6 to obtain:

$$\begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix} = \begin{pmatrix} \varepsilon_1 & \zeta_1 \\ \bar{\varepsilon}_1 & \bar{\zeta}_1 \end{pmatrix} \otimes \begin{pmatrix} \varepsilon_2 & \zeta_2 \\ \bar{\varepsilon}_2 & \bar{\zeta}_2 \end{pmatrix} \begin{pmatrix} N_{RR} \\ N_{RF} \\ N_{FR} \\ N_{FF} \end{pmatrix} \quad (5.7)$$

To make the notation more compact, the set of 4 numbers $(N_{TT}, N_{TL}, N_{LT}, N_{LL})$ can be represented by a rank 2 tensor $\mathcal{T}_{\alpha_1 \alpha_2}$ where α_i corresponds to one object that is either tight (T) or loose (L). Similarly the numbers $(N_{RR}, N_{RF}, N_{FR}, N_{FF})$ can be represented by $\mathcal{R}_{\alpha_1 \alpha_2}$ where α_i corresponds to one object that is either real (R) or FNP (F). With this convention, the Kronecker product of Eq. 5.7 can be obtained by contracting each index α_i of the tensors \mathcal{T} or \mathcal{R} by the 2×2 matrix $\phi_{i \beta_i}^{\alpha_i}$:

$$\mathcal{T}_{\beta_1 \beta_2} = \phi_{1 \beta_1}^{\alpha_1} \phi_{2 \beta_2}^{\alpha_2} \mathcal{R}_{\alpha_1 \alpha_2}, \quad \phi_i = \begin{pmatrix} \varepsilon_i & \zeta_i \\ \bar{\varepsilon}_i & \bar{\zeta}_i \end{pmatrix} \quad (5.8)$$

Following the same procedure as in the one object case, the matrix inversion of the $4 \times 4 \lambda$ matrix is simplified to a matrix inversion of the $2 \times 2 \phi$ matrices. The quantity of interest is the FNP lepton background that passes the tight selection criteria as in Eq. 5.4 which can be compactly written in the two objects case as:

$$\mathcal{T}_{\nu_1 \nu_2}^{\text{FNP}} = \phi_{\nu_1}^{\mu_1} \phi_{\nu_2}^{\mu_2} \xi^{\beta_1 \beta_2}_{\mu_1 \mu_2} \phi^{-1}_{\beta_1}{}^{\alpha_1} \phi^{-1}_{\beta_2}{}^{\alpha_2} \mathcal{T}_{\alpha_1 \alpha_2}. \quad (5.9)$$

The tensor ξ encodes the component of tight and FNP lepton background. In the two objects case, ξ needs to select the total background with at least one fake

lepton $N_F = N_{RF} + N_{FR} + N_{FF}$ that are also passing the tight selection criteria corresponding to the region with signal leptons. As a result, ξ takes the form:

$$\xi = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

To further illustrate, Eq. 5.9 can be written explicitly in the notation of Eq. 5.6 as:

$$N_{\text{FNP}}^{\text{signal}} = \begin{pmatrix} 0 & \varepsilon_1 \zeta_2 & \zeta_1 \varepsilon_2 & \zeta_1 \zeta_2 \end{pmatrix} \Lambda^{-1} \begin{pmatrix} N_{TT} \\ N_{TL} \\ N_{LT} \\ N_{LL} \end{pmatrix}$$

The generalization of Eq. 5.9 from the two objects case to m number of objects in the final state is straightforward:

$$\mathcal{T}_{\nu_1 \dots \nu_m}^{\text{FNP}} = \phi_{\nu_1}{}^{\mu_1} \dots \phi_{\nu_m}{}^{\mu_m} \xi_{\mu_1 \dots \mu_m}^{\beta_1 \dots \beta_m} \phi_{\beta_1}{}^{\alpha_1} \dots \phi_{\beta_m}{}^{\alpha_m} \mathcal{T}_{\alpha_1 \dots \alpha_m}. \quad (5.10)$$

The tensor ξ is of the general form

$$\xi_{\mu_1 \dots \mu_m}^{\beta_1 \dots \beta_m} = \delta_{\mu_1}^{\beta_1} \dots \delta_{\mu_m}^{\beta_m} h(\beta_1, \dots, \beta_m, \nu_1, \dots, \nu_m)$$

where the function h can take values 0 or 1 based on the tight or loose configuration being computed which is encoded in the dependence on the indices ν_i .

The application of the matrix method to multilepton final states comes with two specificities. Firstly, contributions of events with charge-flip electrons would bias a straightforward matrix method estimate (in particular for a final state formed by two leptons with same electric charge). This happens because the candidate-to-signal efficiency for such electrons is typically lower than for real electrons having a correctly-assigned charge. One therefore needs to subtract from n_{pass} and n_{fail} the estimated contributions from charge-flip. This can be performed by including as well events with pairs of opposite-sign candidate leptons in the matrix method estimate, but assigning them an extra weight corresponding to the charge-flip weight. Thanks (again) to the linearity of the matrix method with respect to n_{pass} and n_{fail} , this weight-based procedure is completely equivalent (but more practical) to the aforementioned subtraction.

Secondly, the analytic expression of the matrix method event weight depends on the lepton multiplicity of the final state. This concerns events with three or more candidate leptons: one such event takes part both in the evaluation of the FNP lepton background for a selection with two signal leptons or a selection with three signal leptons, but with different weights². Therefore, for a given event used as input to the matrix method, one should consider all possible leptons combinations, each with its own weight and its own set of kinematic variables. For example, a $e^+e^-\mu^+$ event is comptabilized in the background estimate both as an $e^+\mu^+$ event (with a weight w_1) and as an $e^+e^-\mu^+$ event (with a weight $w_2 \neq w_1$).

²This can appear for inclusive selections: for example an event with two signal leptons may or not contain additional candidate leptons, in a transparent way

5.4.3 Propagation of uncertainties

The two parameters (ε and ζ respectively) can be measured in data, and depend on the flavor and kinematics of the involved leptons. Systematic uncertainties resulting from the measurement of these two parameters, and their extrapolation to the signal regions, can be propagated to uncertainties on the event weight through standard first-order approximations. The different sources of uncertainties should be tracked separately so that correlations of uncertainties across different events can be accounted for correctly. The resulting set of uncertainties on the cumulated event weights can be then added in quadrature to form the systematic uncertainty on the predicted FNP lepton background yield. The corresponding statistical uncertainty can be taken as the RMS of the event weights.

Chapter 6

Search for new physics in events with two same sign leptons or three leptons and jets

6.1 Introduction

As discussed in Chapter ??, supersymmetry (SUSY) is a theoretically favoured extension of the Standard Model (SM), which for each degree of freedom of the SM predicts another degree of freedom with a different spin. These degrees of freedom combine into physical superpartners of the SM particles: scalar partners of quarks and leptons (squarks (\tilde{q}) and sleptons), fermionic partners of gauge and Higgs bosons (gluinos (\tilde{g}), charginos ($\tilde{\chi}_i^{\pm}$, with $i = 1,2$) and neutralinos ($\tilde{\chi}_i^0$ with $i = 1,2,3,4$)), all with identical quantum numbers to their SM partners, except spin. If R -parity is conserved the lightest supersymmetric particle (LSP) is stable and is typically the lightest neutralino $\tilde{\chi}_1^0$ which is a viable dark matter candidate.

As some of these particles are expected to be in the TeV-range, the discovery (or exclusion) of weak-scale SUSY is one of the highest physics priorities for the LHC. Both the ATLAS and CMS collaborations have carried out a vigorous search program for SUSY in various final states.

In this chapter, a detailed description will be given for the search for supersymmetric particles with an experimental signature involving multiple leptons in the final state. The search strategy will be covered along with the event selection procedure and the mechanisms by which the various backgrounds are estimated.

The results of the analysis will be interpreted in the context of simplified and popular supersymmetric models and also recast in terms of model independent limits.

In order to address the SM hierarchy problem with SUSY models , the masses of the partners of the gluons (gluinos \tilde{g}) and of the top quark chiral degrees of freedom (top squarks \tilde{t}_L and \tilde{t}_R), and the partners of the bottom quark (sbottom) are expected to be in the TeVrange. The production cross sections for gluino pairs ($\tilde{g}\tilde{g}$), top–antitop squark pairs ($\tilde{t}_1\tilde{t}_1^*$) and bottom–antibottom squark pairs ($\tilde{b}_1\tilde{b}_1^*$) are relatively large which makes them a primary target for the early search program with the LHC data citeBorschensky:2014cia. The cascade decay of these SUSY particles may proceed via intermediate neutralinos $\tilde{\chi}_{2,3,4}^0$ or charginos $\tilde{\chi}_{1,2}^\pm$ that in turn lead to W , Z or Higgs bosons, or to lepton superpartners (sleptons $\tilde{\ell}$) which will lead to isolated leptons and neutralinos $\tilde{\chi}_1^0$.

This searches focuses on final states with two leptons (electrons or muons) of the same electric charge (referred to as same-sign (SS) leptons) or three leptons (3L) in any charge combination or of same electric charge, jets and missing transverse momentum (\vec{E}_T^{miss} , whose magnitude is referred to as E_T^{miss}). Despite the penalty in signal yields due to the low branching ratio to SS/3L, the extreme background reduction achieved in this signature is a very good opportunity to discover new physics, particularly in scenarios with small mass differences between SUSY particles (“compressed scenarios”). This search is thus sensitive to a wide variety of models based on very different assumptions.

The results covered in the next sections use the full data collected in 2015 and

2016 by the ATLAS experiment in proton–proton (pp) collisions at a center-of-mass energy of $\sqrt{s} = 13\text{TeV}$, corresponding to a total integrated luminosity of 36.1 fb^{-1} .

It uses as signal benchmarks a few SUSY scenarios described in section 6.2. Observed data and predicted Standard Model yields are compared in a set of unbinned signal regions defined in section ?? (“cut and count”), and chosen to provide good sensitivity to the signal benchmarks. The background prediction relies on Monte Carlo simulation of the relevant processes with prompt same-sign leptons and theoretical computations of their cross-sections (section ??), and data-driven estimates of the contributions from processes with charge-flipped electrons (section ??) or non-prompt or fake leptons (section ??). The predictions are compared to observed data in a set of suitably chosen validation regions (section ??) enriched in the different processes. Statistical interpretation of the observations in the signal regions is performed through the framework described in section ??, which is also used by convenience to determine the total background yields in the signal regions, accounting for correlations between systematic uncertainties associated to the different background processes. These yields, compared to observed data, are reported in section ?. In the absence of significant excess in observed data over the Standard Model prediction, exclusion limits are set on the masses of SUSY partners involved in the benchmark SUSY scenarios, and are presented in section ?. Final conclusions are stated in section ??.

6.2 Signal models

Final states with two same-sign leptons or three leptons and multiple jets are sensitive to a variety of new physics scenarios. In supersymmetric models in particular, such final states can be produced in the decays of heavy superpartners involving massive gauge bosons, sleptons or top quarks. Depending on the nature of the particles accompanying the leptons in the final states, a large variety of signatures can be obtained, notably in terms of the numbers of jets and b -tagged jets in the final state. We chose to illustrate the analysis versatility by evaluating its sensitivity to six R -parity-conserving (RPC) SUSY scenarios with the lightest neutralino $\tilde{\chi}_1^0$ as lightest and stable superpartner, featuring gluino, bottom squark or top squark pair production. The first scenarios studied focus on gluino pair production with decays into on-shell (Figure ??) or off-shell (Figure ??) top quarks, as well as on-shell light quarks. The latter are accompanied by a cascade decay involving a $\tilde{\chi}_1^\pm$ and a $\tilde{\chi}_2^0$ (Figure ??) or a $\tilde{\chi}_2^0$ and light sleptons (Figure ??). The other two scenarios target the direct production of third-generation squark pairs with subsequent electroweakino-mediated decays (Figures ?? and ??). The former is characterized by final states with bottom squark pairs decaying to $t\bar{t}WW\tilde{\chi}_1^0\tilde{\chi}_1^0$. The latter, addressed here by looking at a final state with three same-sign leptons, is a model that could explain the excess seen in same-sign lepton signatures during Run 1 [?]. Finally, a full SUSY model with low fine-tuning, the non-universal Higgs model with two extra parameters (NUHM2) [?, ?], is also considered. When the soft-SUSY-breaking electroweakino mass, $m_{1/2}$, is in

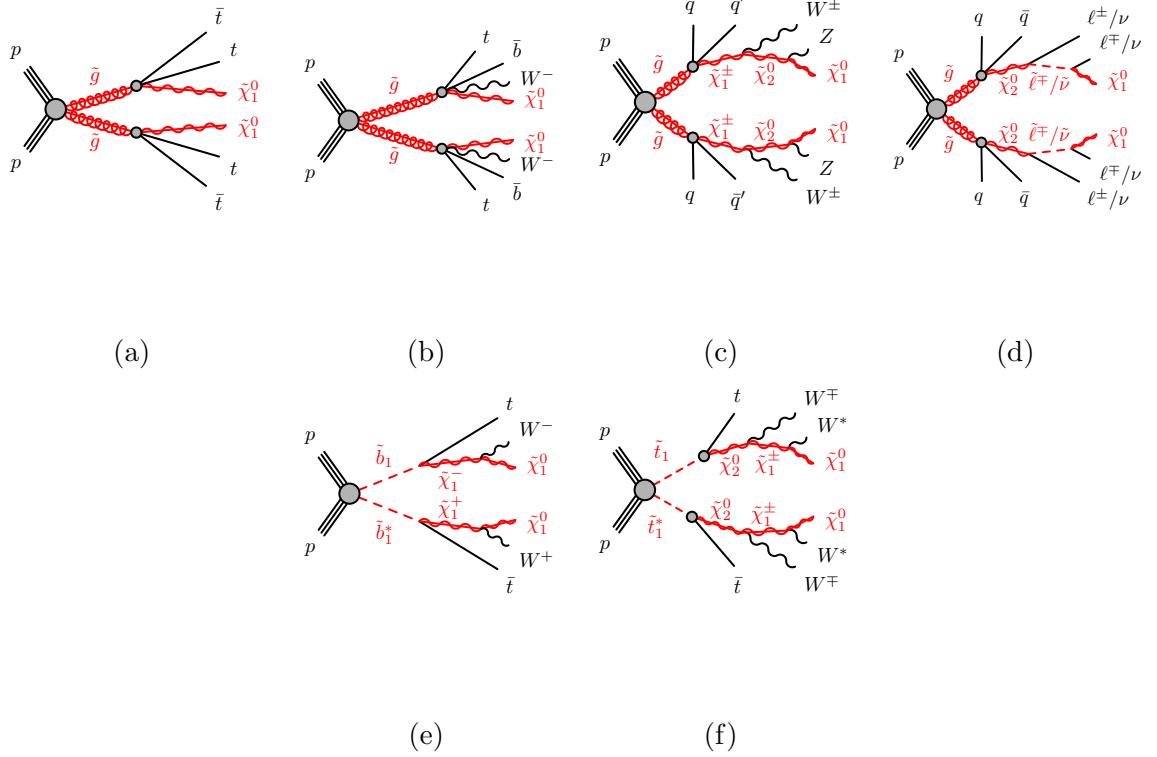


Figure 6.1: SUSY processes featuring gluino ((a), (b), (c), (d)) or third-generation squark ((e), (f)) pair production studied in this analysis. In Figure ??, $\tilde{\ell} \equiv \tilde{e}, \tilde{\mu}, \tilde{\tau}$ and $\tilde{\nu} \equiv \tilde{\nu}_e, \tilde{\nu}_\mu, \tilde{\nu}_\tau$. In Figure ??, the W^* labels indicate largely off-shell W bosons – the mass difference between $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_1^0$ is around 1 GeV.

the range 300–800 GeV, the model mainly involves gluino pair production with gluinos decaying predominantly to $t\bar{t}\tilde{\chi}_1^0$ and $tb\tilde{\chi}_1^\pm$, giving rise to final states with two same-sign leptons and E_T^{miss} .

These scenarios were used as benchmarks to identify regions of the phase space where the analysis can bring particularly useful complementarity to other SUSY searches, and subsequently define our signal regions with a particular focus on these regions. In this section, the scenarios considered are presetend with

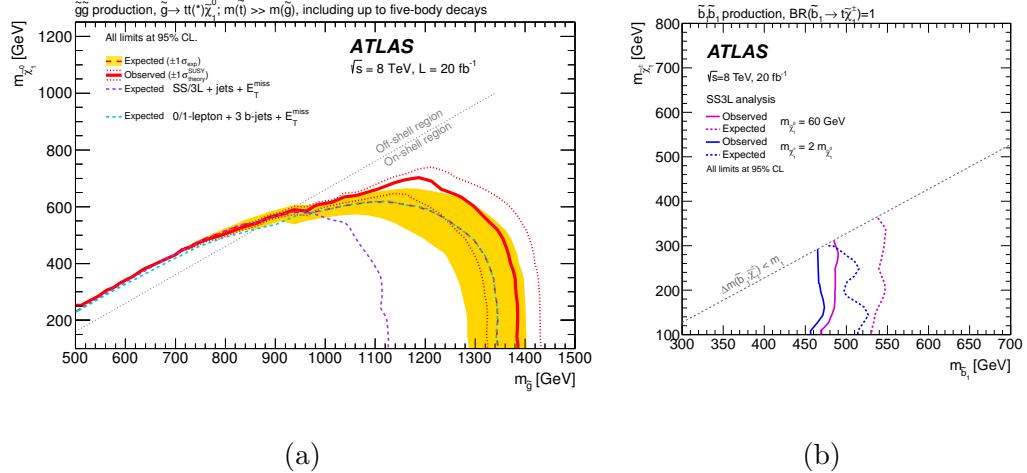


Figure 6.2: Exclusion limits on the gluino-stop offshell (left) and direct sbottom (right) scenarios set by ATLAS with the 2012 dataset [?] prior to the author’s work.

details about the assumed superpartner masses and decay modes. Exclusion limits obtained prior to the work of the author will also be shown to highlight the improvement in reach after the current analysis.

6.2.1 Gluino pair production with slepton-mediated two-step decay

$$\tilde{g} \rightarrow q\bar{q}\ell\bar{\ell}\tilde{\chi}_1^0$$

This scenario (Fig. ??) features gluino pair-production with two-step decays via neutralinos $\tilde{\chi}_2^0$ and sleptons, $\tilde{g} \rightarrow q\bar{q}'\tilde{\chi}_2^0 \rightarrow q\bar{q}'(\tilde{\ell}\ell/\tilde{\nu}\nu) \rightarrow q\bar{q}'(\ell\ell/\nu\nu)\tilde{\chi}_1^0$. The decays are mediated by generic heavy squarks, therefore the b -jet multiplicity in this scenario is low. The final state is made of charged leptons, four additional jets and invisible particles (neutrinos and neutralinos). The average jet multiplicity per event is the smallest among the four scenarios; another characteristic is the

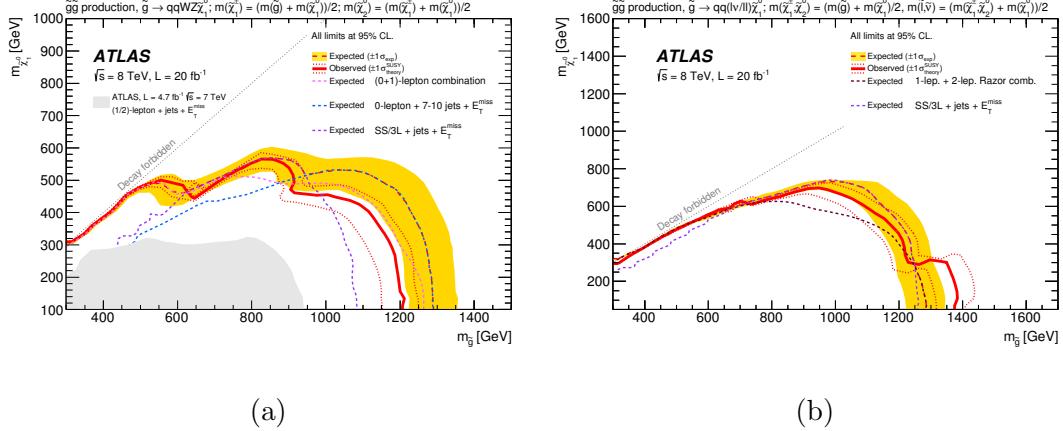


Figure 6.3: Exclusion limits on scenarios featuring gluino pair production followed by two-step decays via heavy gauge bosons or sleptons, set by ATLAS with the 2012 dataset [?] prior to the author’s work.

large fraction of events with several leptons, unlike the other scenarios that have a rather low acceptance due to the branching ratios of $W \rightarrow \ell\nu$ or $Z \rightarrow \ell\ell$. The exclusion limits obtained in run-1 (Fig. 6.3 right) show again that the SS/3L+jets final state is very competitive to probe those models. This scenario is used as a benchmark to define the signal regions with ≥ 3 leptons and no b -jet.

The signal grid is built with variable gluino and $\tilde{\chi}_1^0$ masses; the $\tilde{\chi}_2^0$ mass is chosen half-way between the gluino and LSP masses, and the sleptons masses are also set equal and half-way between the $\tilde{\chi}_2^0$ and LSP masses. The $\tilde{\chi}_2^0$ may decay to any of the six “left-handed” sleptons ($\tilde{\ell}, \tilde{\nu}$) with equal probability. “Right-handed” sleptons are assumed heavy and do not participate to the decay.

6.2.2 Gluino pair production with gaugino-mediated two-step decay

$$\tilde{g} \rightarrow q\bar{q}'WZ\tilde{\chi}_1^0$$

This scenario (Fig. ??) features gluino pair-production with two-step decays via gauginos and W and Z bosons, $\tilde{g} \rightarrow q\bar{q}'\tilde{\chi}_1^\pm \rightarrow q\bar{q}'W\tilde{\chi}_2^0 \rightarrow q\bar{q}'WZ\tilde{\chi}_1^0$, mediated by generic heavy squarks of the first and second generations. The final state is made of two W and two Z bosons (possibly offshell), four additional jets and invisible particles (neutrinos and neutralinos). This generally leads to events with large jet multiplicities and a fair branching ratio for dileptonic final states. The exclusion limits obtained in run-1 indeed illustrate the competitiveness of the SS/3L+jets search (Fig. 6.3 left) particularly the heavy- $\tilde{\chi}_1^0$ region of the phase space. This scenario is used as a benchmark to define the signal regions with many jets but none tagged as a b -jet.

The signal grid is built with variable gluino and $\tilde{\chi}_1^0$ masses, and the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_2^0$ masses are set such that the former lies half-way between the gluino and $\tilde{\chi}_1^0$ masses, and the latter half-way between $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_1^0$ masses.

6.2.3 Sbottom pair production with one-step decay $\tilde{b}_1 \rightarrow t\tilde{\chi}_1^\pm$

In this scenario (Fig. ??), bottom squarks are rather light and assumed to decay in a top quark and a chargino $\tilde{\chi}_1^\pm$, with a subsequent $\tilde{\chi}_1^\pm \rightarrow W^\pm\tilde{\chi}_1^0$ decay, providing complementarity to the mainstream search [?] which focuses on the channel $\tilde{b}_1 \rightarrow b\tilde{\chi}_1^0$. The final state resulting from the production of a $\tilde{b}_1\tilde{b}_1^*$ pair contains two top quarks, two W bosons and two neutralinos. While this final state may lead

to various experimental signatures, the model was considered in Run-1 [?] only by the same-sign leptons and jets search, leading to the exclusion limits presented in Fig. 6.2. Signal events typically contain one or two b -tagged jets, therefore this scenario is used as benchmark to define the signal regions with ≥ 1 b -jet.

The model adopts a fixed chargino-neutralino mass difference of 100 GeV, therefore always allowing on-shell W bosons in the $\tilde{\chi}_1^\pm \rightarrow W\tilde{\chi}_1^0$ decay¹. Only pair production of the lightest sbottom is considered, followed by an exclusive decay in the aforementioned channel.

6.2.4 Gluino pair production with stop-mediated decay $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$

In this scenario inspired by naturalness arguments, gluinos are coupling preferentially to stops which are lighter than the other squarks. Gluinos are however considered lighter than stops, and decay directly into a $t\bar{t}\tilde{\chi}_1^0$ triplet via a virtual stop (Fig. ??). The pair production of gluinos leads to a final state containing four top quarks and two neutralinos. This characteristic final state is accessible through various experimental signatures, which is why this model is commonly used as a benchmark to compare analyses sensitivities. The searches performed with Run-1 data [?], summarized in Fig. 6.2, showed that the same-sign leptons final state is competitive only at large neutralino mass. This region of the phase space is consequently given a particular attention in the choice of signal regions described further on. For instance, the region of phase-space with $\Delta m(\tilde{g}, \tilde{\chi}_1^0) < 2m_t$, where

¹A different chargino mass assumption is adopted in the current work compared to the Run-1 paper [?]. Fig. 6.2 is shown for illustration only. The reduced chargino-neutralino mass gap in the current analysis allows to study signal scenarios with heavy neutralinos, which were not considered previously.

gluinos decay via one or two offshell top quarks, is only accessible for this analysis.

In the signal samples referenced in this document, the mass of the lightest stop is fixed to 10 TeV and is mostly a \tilde{t}_R state. Only gluino pair production is considered, followed by an exclusive decay in the aforementioned channel. Signal events typically contain many b -tagged jets, therefore this scenario is used as benchmark to define the signal regions with ≥ 2 b -jets.

6.2.5 $\tilde{t}_1\tilde{t}_1^*$ with “three-same-sign leptons” signature

Inspired by Ref. [?], a simplified model featuring a stop pair-production with two-step decays via a neutralino $\tilde{\chi}_2^0$ and a chargino $\tilde{\chi}_1^\pm$ is added in this version of the analysis, according to the decay illustrated on Fig. ??:

$$\tilde{t}_1 \rightarrow t\tilde{\chi}_2^0 \rightarrow t\tilde{\chi}_1^\pm W^\mp \rightarrow tW^\pm W^\mp \tilde{\chi}_1^0.$$

This simplified model is a well-motivated representation of a pMSSM model. The lightest stop (\tilde{t}_1) is right-handed and $\tilde{\chi}_2^0$ is bino-like which leads to a large branching ratio in the decay $\tilde{t}_1 \rightarrow t\tilde{\chi}_2^0$. Furthermore, the decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^\pm W^\mp$ is also enhanced since $\tilde{\chi}_1^\pm$ is wino-like, as long as $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_1^0$ are nearly mass degenerate and $m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0} < m_H = 125\text{GeV}$ to suppress the decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 + H$ (the decay $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^0 + Z$ is suppressed). By respecting these conditions and evading the bottom squark limit shown in Fig. ??, c, we consider a one-dimensional grid with a \tilde{t}_1 mass varying between 550 GeV and 800 GeV with a 50 GeV gap², a two body decay to an on-shell top quark and a $\tilde{\chi}_2^0$ which has a 100 GeV mass difference from $\tilde{\chi}_1^0$. The mass difference between the $\tilde{\chi}_1^\pm$ and $\tilde{\chi}_1^0$ is taken to be 500 MeV which

²Only the points at \tilde{t}_1 mass of 550 GeV are available at the moment.

is not excluded by the disappearing track analysis. In fact, this mass gap could easily be increased by introducing a small amount of higgsino mixing [?].

While the stop pair production is similar to the sbottom pair production in terms of kinematics, the stop pair production offers a unique topology that leads to three leptons of the same electric charge. This final state benefits from an extreme reduction of the SM background while maintaining a good signal acceptance which helps loosen the kinematic cuts to access a more compressed SUSY phase space. As a result, this scenario is complementary to the search for bottom squarks.

6.2.6 Non-Universal Higgs Models

In references [?, ?, ?], theorists studied a complete two-extra-parameter non-universal Higgs model (NUHM2) that can have low fine tuning (natural) and predicts final state signatures that allow large background rejection while retaining high signal efficiency. The NUHM2 model allows the soft SUSY breaking masses of the Higgs multiplets, m_{H_u} and m_{H_d} , to be different from matter scalar masses (m_0) at the grand unification scale. The NUHM2 model is expected to form the effective theory for energies lower than m_{GUT} resulting from SU(5) or general SU(10) grand unified theories. The scalar mass m_0 , the soft SUSY breaking gaugino mass $m_{1/2}$, the pseudoscalar Higgs boson mass m_A , the trilinear SUSY breaking parameter A_0 , the weak scale ratio of Higgs field vacuum expectation values $\tan \beta$, and the superpotential Higgs mass μ are the free parameters. Both $m_{1/2}$ and μ are varied while the other parameters are fixed to $m_0 = 5$ TeV, $A_0 = -1.6m_0$, $\tan \beta = 15$,

Decay	BR	Decay	BR
$t\bar{t}\chi_1^0$	0.13	$tb\chi_1^\pm$	0.45
$t\bar{t}\chi_2^0$	0.21	$tb\chi_2^\pm$	0.04
$t\bar{t}\chi_3^0$	0.13	-	-
$t\bar{t}\chi_4^0$	0.02	-	-
$t\bar{t}\chi_i^0$	0.49	$tb\chi_i^\pm$	0.49

Table 6.1: The dominant gluino decay modes for $m_{1/2} = 400$ GeV for the NUHM2 model.

$m_A = 1$ TeV, and $\text{sign}(\mu) > 0$. These parameter choices lead directly to a Higgs mass of 125 GeV in accord with experiment. In this “radiatively-driven natural” SUSY approach, the higgsino is required to have mass below 200-300 GeV, the stop to have a mass below ~ 3 TeV, and the gluino below ~ 4 TeV. Final state topologies that include MET, $W \rightarrow \ell\nu$ and that result in same-sign dileptons can be explored. Simulated NUHM2 signal samples with mass ($m_{1/2}$) values from 300-800 GeV and $\mu = 150$ GeV were generated. The gluino mass in this model is approximately $2.5 \times m_{1/2}$. Table 6.1 shows the branching ratios of the dominant gluino decay modes for $m_{1/2} = 400$ GeV.

6.2.7 Signal cross-sections and simulations

The signal processes are generated from leading order (LO) matrix elements with up to two extra partons (only one for the grid featuring slepton-mediated gluino

Table 6.2: Signal cross-sections [pb] and related uncertainties [%] for scenarios featuring $\tilde{g}\tilde{g}$ (top table) or $\tilde{b}_1\tilde{b}_1^*$ (bottom table) production, as a function of the pair-produced superpartner mass, reproduced from Ref. [?].

Gluino mass (GeV)	500	550	600	650	700
Cross section (pb)	$27.4 \pm 14\%$	$15.6 \pm 14\%$	$9.20 \pm 14\%$	$5.60 \pm 14\%$	$3.53 \pm 14\%$
750	800	850	900	950	1000
$2.27 \pm 14\%$	$1.49 \pm 15\%$	$0.996 \pm 15\%$	$0.677 \pm 16\%$	$0.466 \pm 16\%$	$0.325 \pm 17\%$
1050	1100	1150	1200	1250	1300
$0.229 \pm 17\%$	$0.163 \pm 18\%$	$0.118 \pm 18\%$	$0.0856 \pm 18\%$	$0.0627 \pm 19\%$	$0.0461 \pm 20\%$
1350	1400	1450	1500	1550	1600
$0.0340 \pm 20\%$	$0.0253 \pm 21\%$	$0.0189 \pm 22\%$	$0.0142 \pm 23\%$	$0.0107 \pm 23\%$	$0.00810 \pm 24\%$

Sbottom mass (GeV)	400	450	500	550
Cross section (pb)	$1.84 \pm 14\%$	$0.948 \pm 13\%$	$0.518 \pm 13\%$	$0.296 \pm 13\%$
600	650	700	750	800
$0.175 \pm 13\%$	$0.107 \pm 13\%$	$0.0670 \pm 13\%$	$0.0431 \pm 14\%$	$0.0283 \pm 14\%$

decays), using the `MADGRAPH v5.2.2.3` generator [?] interfaced to `PYTHIA 8.186` [?] with the *ATLAS 14* tune [?] for the modelling of the SUSY decay chain, parton showering, hadronisation and the description of the underlying event. For the RPV models, `MADGRAPH v5.2.3.3` and `PYTHIA 8.210` [?] were used instead. Parton luminosities are provided by the `NNPDF23LO` [?] set of parton distribution functions. Jet-parton matching is realized following the CKKW-L prescription [?], with a matching scale set to one quarter of the pair-produced superpartner mass.

The signal samples are normalised to the next-to-next-to-leading order cross-

section from Ref. [?] including the resummation of soft gluon emission at next-to-next-to-leading-logarithmic accuracy (NLO+NLL), as detailed in Ref. [?]; some of these cross-sections are shown for illustration in Table 6.2. For the production of like-sign d-squark (RPV scenario), Prospino [?] is used to scale the samples to their NLO cross-section.

Cross-section uncertainties are also taken from Ref. [?] as well, and include contributions from varied normalization and factorization scales, as well as PDF uncertainties. They typically vary between 15 and 25%. We do not consider any source of uncertainties on signal acceptance, experience having shown that these are generally smaller than the uncertainties on the inclusive production cross-section. [this is currently being verified for some cases]

The dataset IDs are listed in appendix ??, Table ??, and further details on simulation and reconstruction are provided in section ??.

6.3 Data set and simulated event samples

The data used in this analysis were collected during 2015 and 2016 with a peak instantaneous luminosity of $L = 1.4 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. The mean number of pp interactions per bunch crossing (pile-up) in the data set is 24. After the application of beam, detector and data-quality requirements, the integrated luminosity considered corresponds to 36.1 fb^{-1} . The uncertainty in the combined 2015+2016 integrated luminosity is 3.2%. It is derived, following a methodology similar to that detailed in Ref. [?], from a preliminary calibration of the luminosity

scale using x - y beam-separation scans performed in August 2015 and May 2016.

Monte Carlo (MC) simulated event samples are used to model the SUSY signals and to estimate the irreducible SM background with two same-sign and/or three “prompt” leptons. Prompt leptons are produced directly in the hard-scattering process, or in the subsequent decays of W, Z, H bosons or prompt τ leptons. The reducible background, mainly arising from $t\bar{t}$ production, is estimated from the data as described in Section 6.6.10. The MC samples were processed through a detailed ATLAS detector simulation [?] based on GEANT4 [?] or a fast simulation using a parameterization of the calorimeter response and GEANT4 for the ID and MS [?]. To simulate the effects of additional pp collisions in the same and nearby bunch crossings, inelastic interactions were generated using the soft strong-interaction processes of PYTHIA 8.186 [?] with a set of tuned parameters referred to as the A2 tune [?] and the MSTW2008LO parton distribution function (PDF) [?]. These MC events were overlaid onto the simulated hard-scatter event and reweighted to match the pile-up conditions observed in the data. Table 6.3 presents, for all samples, the event generator, parton shower, cross-section normalization, PDF set and the set of tuned parameters for the modelling of the parton shower, hadronization and underlying event. In all MC samples, except those produced by the SHERPA generator, the EVTGEN v1.2.0 program [?] was used to model the properties of bottom and charm hadron decays.

The SUSY signals from Figure 6.1 are defined by an effective Lagrangian describing the interactions of a small number of new particles [?, ?, ?]. All SUSY particles not included in the decay of the pair-produced squarks and gluinos are

Physics process	Event generator	Parton shower	Cross-section normalization	PDF set	Set of tuned parameters
Signal	AMC@NLO 2.2.3 [?]	PYTHIA 8.186 [?]	NLO+NLL	NNPDF2.3LO [?]	A14 [?]
$t\bar{t} + X$					
$t\bar{t}W, t\bar{t}Z/\gamma^*$	AMC@NLO 2.2.2	PYTHIA 8.186	NLO [?]	NNPDF2.3LO	A14
$t\bar{t}H$	AMC@NLO 2.3.2	PYTHIA 8.186	NLO [?]	NNPDF2.3LO	A14
$4t$	AMC@NLO 2.2.2	PYTHIA 8.186	NLO [?]	NNPDF2.3LO	A14
Diboson					
ZZ, WZ	SHERPA 2.2.1 [?]	SHERPA 2.2.1	NLO [?]	NNPDF2.3LO	SHERPA default
Other (inc. $W^\pm W^\pm$)	SHERPA 2.1.1	SHERPA 2.1.1	NLO [?]	CT10 [?]	SHERPA default
Rare					
$t\bar{t}WW, t\bar{t}WZ$	AMC@NLO 2.2.2	PYTHIA 8.186	NLO [?]	NNPDF2.3LO	A14
$tZ, tWZ, t\bar{t}\bar{t}$	AMC@NLO 2.2.2	PYTHIA 8.186	LO	NNPDF2.3LO	A14
WH, ZH	AMC@NLO 2.2.2	PYTHIA 8.186	NLO [?]	NNPDF2.3LO	A14
Triboson	SHERPA 2.1.1	SHERPA 2.1.1	NLO [?]	CT10	SHERPA default

Table 6.3: Simulated signal and background event samples: the corresponding event generator, parton shower, cross-section normalization, PDF set and set of tuned parameters are shown for each sample. Because of their very small contribution to the signal-region background estimate, $t\bar{t}WW, t\bar{t}WZ, tZ, tWZ, t\bar{t}\bar{t}, WH, ZH$ and triboson are summed and labelled “rare” in the following.

effectively decoupled. These simplified models assume one production process and one decay channel with a 100% branching fraction. All simplified models were generated from leading-order (LO) matrix elements with up to two extra partons in the matrix element (only up to one for the $\tilde{g} \rightarrow q\bar{q}(\ell\ell/\nu\nu)\tilde{\chi}_1^0$ model) using AMC@NLO 2.2.3 [?] interfaced to PYTHIA 8 with the A14 tune [?] for the modelling of the parton shower, hadronization and underlying event. Jet–parton matching was realized following the CKKW-L prescription [?], with a matching

scale set to one quarter of the pair-produced superpartner mass. All signal models were generated with prompt decays of the SUSY particles. Signal cross-sections were calculated at next-to-leading order (NLO) in the strong coupling constant, adding the resummation of soft-gluon emission at next-to-leading-logarithmic accuracy (NLO+NLL) [?, ?, ?, ?, ?], except for the NUHM2 model where NLO cross-sections were used [?, ?]. The nominal cross-sections and the uncertainties were taken from envelopes of cross-section predictions using different PDF sets and factorization and renormalization scales, as described in Refs. [?, ?]. Typical pair-production cross-sections are: 4.7 ± 1.2 fb for gluinos with a mass of 1.7 TeV, 28 ± 4 fb for bottom squarks with a mass of 800 GeV, and 15.0 ± 2.0 fb for down squark-rights with a mass of 800 GeV and a gluino mass of 2.0 TeV.

The two dominant irreducible background processes are $t\bar{t}V$ (with V being a W or Z/γ^* boson) and diboson production with final states of four charged leptons ℓ ,³ three charged leptons and one neutrino, or two same-sign charged leptons and two neutrinos. The MC simulation samples for these are described in Refs. [?] and [?], respectively. For diboson production, the matrix elements contain the doubly resonant diboson processes and all other diagrams with four or six electroweak vertices, such as $W^\pm W^\pm jj$, with one ($W^\pm W^\pm jj$) or two (WZ , ZZ) extra partons. NLO cross-sections for $t\bar{t}W$, $t\bar{t}Z/\gamma^*(\rightarrow \ell\ell)$ ⁴ and leptonic diboson processes are respectively 600.8 fb [?], 124 fb and 6.0 pb [?]. The processes $t\bar{t}H$ and $4t$, with NLO cross-sections of 507.1 fb [?] and 9.2 fb [?] respectively, are also

³All lepton flavours are included here and τ leptons subsequently decay leptonically or hadronically.

⁴This cross-section is computed using the configuration described in Refs. [?, ?].

considered.

Other background processes, with small cross-sections and no significant contribution to any of the signal regions, are grouped into a category labelled “rare”. This category contains $t\bar{t}WW$ and $t\bar{t}WZ$ events generated with no extra parton in the matrix element, and tZ , tWZ , $t\bar{t}\bar{t}$, WH and ZH as well as triboson (WWW , WWZ , WZZ and ZZZ) production with fully leptonic decays, leading to up to six charged leptons. The processes WWW , WZZ and ZZZ were generated at NLO with additional LO matrix elements for up to two extra partons, while WWZ was generated at LO with up to two extra partons.

6.4 Event selection

6.4.1 Pre-selection and event cleaning

A sample of two same-sign or three leptons is selected applying the following criteria:

- **Jet Cleaning:** Events are required to pass a set of cleaning requirements. An event is rejected if any pre-selected jets ($|\eta| < 4.9$, after jet-electron overlap removal) fails the jet quality criteria. The cleaning requirements are intended to remove events where significant energy was deposited in the calorimeters due to instrumental effects such as cosmic rays, beam-induced (non-collision) particles, and noise. Around 0.5% of data events are lost after applying this cut.
- **Primary Vertex:** Events are required to have a reconstructed vertex [?]

with at least two associated tracks with $p_T > 400$ MeV. The vertex with the largest Σp_T^2 of the associated tracks is chosen as the primary vertex of the event. This cut is found to be 100% efficient.

- **Bad Muon Veto:** Events containing at least one pre-selected muon satisfying $\sigma(q/p)/|q/p| > 0.2$ before the overlap removal are rejected. Around 0.1% of data events are removed by this cut.
- **Cosmic Muon Veto:** Events containing a cosmic muon candidate are rejected. Cosmic muon candidates are looked for among pre-selected muons, if they fail the requirements $|z_0| < 1.0$ mm and $|d_0| < 0.2$ mm, where the longitudinal and transverse impact parameters z_0 and d_0 are calculated with respect to the primary vertex. Up to 6% of data events are lost at this cleaning cut.
- **At least two leptons:** Events are required to contain at least two signal leptons with $p_T > 20$ GeV for the two leading leptons. If the event contains a third signal lepton with $p_T > 10$ GeV the event is regarded as a three-lepton event, otherwise as a two-lepton event. The data sample obtained is then divided into three channels depending on the flavor of the two leptons forming a same-sign pair (ee , $\mu\mu$, $e\mu$). If more than one same-sign pairs can be built, the one involving the leading lepton will be considered for the channel selection.
- **Same-sign:** if the event has exactly two leptons, then these two leptons have to be of identical electric charge (“same-sign”).

The following event variables are also used in the definition of the signal and validation regions in the analysis:

- The inclusive effective mass m_{eff} defined as the scalar sum of all the signal leptons p_T (see Table ??), all signal jets p_T (see Table 6.4) and E_T^{miss} .

6.4.2 Trigger strategy

Events are selected using a combination of dilepton and E_T^{miss} triggers, the latter being used only for events with $E_T^{\text{miss}} > 250\text{GeV}$. Since the trigger thresholds have been raised between 2015 and 2016 due to the continuous increase of the instantaneous luminosity, the dilepton triggers used for:

- 2015 data: logical **or** of a trigger with two electrons of 12 GeV, with an electron of 17 GeV and a muon of 14 GeV, with two muons of 18 GeV and 8 GeV.
- 2016 data: logical **or** of a trigger with two electrons of 17 GeV, with an electron of 17 GeV and a muon of 14 GeV, with two muons of 22 GeV and 8 GeV.

The E_T^{miss} trigger was also raised from 70GeV to a 100 GeV and 110 GeV. The trigger-level requirements on E_T^{miss} and the leading and subleading lepton p_T are looser than those applied offline to ensure that trigger efficiencies are constant in the relevant phase space.

Trigger matching

For events exclusively selected via one or several of the dilepton triggers, we

require a matching between the online and offline leptons with $p_T > 20\text{GeV}$. with the exception of the dimuon trigger for which muons with $p_T > 10\text{GeV}$ are also considered. In addition, for the dimuon trigger in the 2016 configuration, the p_T requirement of the leading matched muon is raised to 23 GeVto remain on the trigger efficiency plateau.

Trigger scale factors

The simulated events are corrected for any potential differences in the trigger efficiency between data and MC simulation. Assuming uncorrelation between the E_T^{miss} and dilepton triggers, trigger scale factors are applied to MC events which were not selected by the E_T^{miss} trigger. These scale factors are computed for each event, considering the combination of fired triggers, the number and flavours of the leptons,

6.4.3 Object definition

This section presents the definitions of the objects used in the analysis: jets, electrons, muons and E_T^{miss} (the taus are not considered).

6.4.4 Jets

The jet selection is summarized in Table 6.4. Jets are reconstructed using the anti- k_t jet algorithm [?] with the distance parameter R set to 0.4 and a three dimensional input of topological energy clusters in the calorimeter [?]. The jets are kept only if they have $p_T > 20\text{ GeV}$ and lie within $|\eta| < 2.8$. To mitigate the effects of pileup, the pile-up contribution is subtracted from the expected

Table 6.4: Summary of the jet selection criteria.

Pre-selected jet	
Collection	AntiKt4EMTopo
Acceptance	$p_T > 20 \text{ GeV}$, $ \eta < 2.8$
Overlap	see section 6.4.7
Jet vertex tagger	reject jets with $p_T < 60 \text{ GeV}$, $ \eta < 2.4$ and $\text{JVT} < 0.59$ after overlap removal
b-jets	
Acceptance	$p_T > 20 \text{ GeV}$, $ \eta < 2.5$
<i>b</i> -tagging	MV2c10 algorithm 70% OP MV2c10 algorithm 85% OP for overlap removal

average energy contribution according to the jet area [?, ?]. In order to reduce the effects of pile-up, a significant fraction of the tracks in jets with $p_T < 60 \text{ GeV}$ and $|\eta| < 2.4$ must originate from the primary vertex, as defined by the jet vertex tagger (JVT) [?]. The jet calibration follows the prescription in Ref. [?].

Jets containing *b*-hadrons (commonly referred to as *b*-tagging) is performed with the MV2c10 algorithm, a multivariate discriminant making use of track impact parameters and reconstructed secondary vertices [?, ?]. The 70% efficiency operating point was chosen which corresponds to the average efficiency for tagging

b -jets in simulated $t\bar{t}$ events. This efficiency working point was favored by optimisation studies performed in simulated signal and background samples. The rejection factors for light-quark/gluon jets, c -quark jets and hadronically decaying τ leptons in simulated $t\bar{t}$ events are approximately 380, 12 and 54, respectively [?, ?]. Jets with $|\eta| < 2.5$ which satisfy the b -tagging and JVT requirements are identified as b -jets. Correction factors and uncertainties determined from data for the b -tagging efficiencies and mis-tag rates are applied to the simulated samples [?].

For the data-driven background estimations, two categories of electrons and muons are used: “candidate” and “signal” with the latter being a subset of the “candidate” leptons satisfying tighter selection criteria.

6.4.5 Electrons

Electron candidates are reconstructed from energy depositions in the electromagnetic calorimeter and required to be matched to an inner detector track, to have $p_T > 10\text{GeV}$ and $|\eta| < 2.47$, and to pass the “Loose” likelihood-based electron identification requirement [?]. Electrons in the transition region between the barrel and endcap electromagnetic calorimeters ($1.37 < |\eta| < 1.52$) are rejected to reduce the contribution from fake/non-prompt electrons. The transverse impact parameter d_0 with respect to the reconstructed primary vertex must satisfy $|d_0/\sigma(d_0)| < 5$. This last requirement helps reduce the contribution from charge mis-identification.

Signal electrons are additionally required to pass the “Medium” likelihood-based identification requirement [?]. Only signal electrons with $|\eta| < 2.0$ are

considered, to reduce the level of charge-flip background. In addition, signal electrons that are likely to be reconstructed with an incorrect charge assignment are rejected using a few electron cluster and track properties: the track impact parameter, the track curvature significance, the cluster width and the quality of the matching between the cluster and its associated track, both in terms of energy and position. These variables, as well as the electron p_T and η , are combined into a single classifier using a boosted decision tree (BDT). A selection requirement on the BDT output is chosen such as to achieve a rejection factor between 7 and 8 for electrons with a wrong charge assignment while selecting properly measured electrons with an efficiency of 97% (in $Z \rightarrow ee$ MC).

A multiplicative event weight is applied for each signal electron in MC to the overall event weight in order to correct for differences in efficiency between data and MC.

6.4.6 Muons

Muons candidates are reconstructed from muon spectrometer tracks matched to the inner detector tracks in the region $|\eta| < 2.5$. Muon candidates must pass the “Medium” identification requirements [?] and have $p_T > 10\text{GeV}$ and $|\eta| < 2.4$. Signal muons are required to pass $|d_0|/\sigma(d_0) < 3$ and $|z_0 \cdot \sin(\theta)| < 0.5\text{mm}$.

A multiplicative event weight is applied for each selected muon in MC to the overall event weight in order to correct for differences in efficiency between data and MC.

6.4.7 Overlap removal

According to the above definitions, one single final state object may fall in more than one category, being therefore effectively double-counted. For example, one isolated electron is typically reconstructed both as an electron and as a jet. A procedure to remove overlaps between final state objects was therefore put in place, and applied on pre-selected objects. Any jet within a distance $\Delta R_y \equiv \sqrt{(\Delta y)^2 + (\Delta\phi)^2} = 0.2$ of a lepton candidate is discarded, unless the jet is *b*-tagged,⁵ in which case the lepton is discarded since it probably originated from a semileptonic *b*-hadron decay. Any remaining lepton within $\Delta R_y \equiv \min\{0.4, 0.1 + 9.6\text{GeV}/p_T(\ell)\}$ of a jet is discarded. In the case of muons, the muon is retained and the jet is discarded if the jet has fewer than three associated tracks. This reduces inefficiencies for high-energy muons undergoing significant energy loss in the calorimeter.

6.4.8 Missing transverse energy

The missing transverse energy (E_T^{miss}) is computed as a negative vector sum of the transverse momenta of all identified candidate objects (electrons, photons [?], muons and jets) and an additional soft term. The soft term is constructed from all tracks associated with the primary vertex but not with any physics object. In this way, the E_T^{miss} is adjusted for the best calibration of the jets and the other identified physics objects listed above, while maintaining approximate pile-up

⁵In this case the *b*-tagging operating point corresponding to an efficiency of 85% is used.

independence in the soft term [?, ?].

6.5 Signal regions

In order to maximize the sensitivity to the signal models of Figure 6.1, 13 non-exclusive signal regions are defined in Table 6.6. The SRs are named in the form $RPCNLMBX$, where N indicates the number of leptons required, M the number of b -jets required, and X indicates the severity of the E_T^{miss} or m_{eff} requirements (Soft, Medium or Hard). All signal regions, except Rpv2L0b, allow any number of additional leptons in addition to a $e^\pm e^\pm$, $e^\pm \mu^\pm$ or $\mu^\pm \mu^\pm$ pair. Signal regions with 3 leptons can be either any charge combination or all three with the same charge (Rpc3LSS1b). For each lepton/ b -jet multiplicity, two signal regions are defined targeting either compressed spectra or large mass splittings.

The optimization of the definitions of signal regions relied on a brute-force scan of several discriminating variables in a loose classification of events in terms of number of b -jets and/or leptons in the final state, each being associated to the signal scenario favouring this final state. The other main discriminant variables (e.g number of jets above a certain p_T threshold, m_{eff} , E_T^{miss} , $E_T^{\text{miss}}/m_{\text{eff}}$ ratio) were then allowed to vary, to determine for each point of the parameter space the best configuration. The figure of merit used to rank configurations was the discovery significance for a statistical test based on a ratio of two Poisson means [?]. A realistic systematic uncertainty of 30% on the expected background yield was included in the statistical test. To preserve the discovery potential, only

configurations leading to at least two signal events were considered for a given signal point. The total number of background events should not be smaller than 1; to model in a more realistic way the effect of non-prompt and fake leptons and electron charge mis-identification backgrounds, which are determined from data in the analysis, the MC predictions for those processes in $t\bar{t}$ and $Z+jets$ MC were scaled using the factors obtained from the MC template method (see Section ??), as shown in Table 6.5. Note that different corrections are applied depending on the showering (Pythia or Sherpa) used for each sample, and for the fake and non-prompt leptons originated from heavy-flavour (HF) and light-flavour (LF).

Table 6.5: Scaling factors applied to the electron charge-flip and non-prompt/fake lepton background in the SR optimization procedure.

	Charge mis-id	HF e	HF μ	LF e	LF μ
Pythia	0.96 ± 0.08	1.80 ± 0.45	2.10 ± 0.58	1.55 ± 0.14	0.74 ± 0.81
Sherpa	1.02 ± 0.09	2.72 ± 0.57	1.81 ± 0.75	1.16 ± 0.18	1.84 ± 1.16

Since the signal regions defined out of the scanning procedure may not be mutually exclusive, the results expressed in terms of exclusion limits will be obtained by using for each signal point the SR giving the best expected sensitivity. For the latter, only the signal regions that were defined aiming for that particular signal model are considered, though.

Figures 6.4 and 6.5 show the performance of the SRs in the four RPC benchmark models with top quarks considered. The discovery significance for each

Signal region	$N_{\text{leptons}}^{\text{signal}}$	$N_{b\text{-jets}}$	N_{jets}	p_T^{jet} [GeV]	E_T^{miss} [GeV]	m_{eff} [GeV]	$E_T^{\text{miss}}/m_{\text{eff}}$	Other	Targeted Signal
Rpc2L2bS	$\geq 2\text{SS}$	≥ 2	≥ 6	> 25	> 200	> 600	> 0.25	–	Fig. ??
Rpc2L2bH	$\geq 2\text{SS}$	≥ 2	≥ 6	> 25	–	> 1800	> 0.15	–	Fig. ??, NUHM2
Rpc2Lsoft1b	$\geq 2\text{SS}$	≥ 1	≥ 6	> 25	> 100	–	> 0.3	$20, 10 < p_T^{\ell_1}, p_T^{\ell_2} < 100 \text{ GeV}$	Fig. ??
Rpc2Lsoft2b	$\geq 2\text{SS}$	≥ 2	≥ 6	> 25	> 200	> 600	> 0.25	$20, 10 < p_T^{\ell_1}, p_T^{\ell_2} < 100 \text{ GeV}$	Fig. ??
Rpc2L0bS	$\geq 2\text{SS}$	$= 0$	≥ 6	> 25	> 150	–	> 0.25	–	Fig. ??
Rpc2L0bH	$\geq 2\text{SS}$	$= 0$	≥ 6	> 40	> 250	> 900	–	–	Fig. ??
Rpc3L0bS	≥ 3	$= 0$	≥ 4	> 40	> 200	> 600	–	–	Fig. ??
Rpc3L0bH	≥ 3	$= 0$	≥ 4	> 40	> 200	> 1600	–	–	Fig. ??
Rpc3L1bS	≥ 3	≥ 1	≥ 4	> 40	> 200	> 600	–	–	Other
Rpc3L1bH	≥ 3	≥ 1	≥ 4	> 40	> 200	> 1600	–	–	Other
Rpc2L1bS	$\geq 2\text{SS}$	≥ 1	≥ 6	> 25	> 150	> 600	> 0.25	–	Fig. ??
Rpc2L1bH	$\geq 2\text{SS}$	≥ 1	≥ 6	> 25	> 250	–	> 0.2	–	Fig. ??
Rpc3LSS1b	$\geq \ell^\pm \ell^\pm \ell^\pm$	≥ 1	–	–	–	–	–	veto $81 < m_{e^\pm e^\pm} < 101 \text{ GeV}$	Fig. ??

Table 6.6: Summary of the signal region definitions. Unless explicitly stated in the table, at least two signal leptons with $p_T > 20 \text{ GeV}$ and same charge (SS) are required in each signal region. Requirements are placed on the number of signal leptons ($N_{\text{leptons}}^{\text{signal}}$), the number of b -jets with $p_T > 20 \text{ GeV}$ ($N_{b\text{-jets}}$), the number of jets (N_{jets}) above a certain p_T threshold (p_T^{jet}), E_T^{miss} , m_{eff} and/or $E_T^{\text{miss}}/m_{\text{eff}}$. The last column indicates the targeted signal model. The Rpc3L1b and Rpc3L1bH SRs are not motivated by a particular signal model and can be seen as a natural extension of the Rpc3L0b SRs with the same kinematic selections but requiring at least one b -jet.

signal point is shown, together with the contours corresponding to a 3σ discovery sensitivity, 1.64σ discovery sensitivity and 95% confidence level limits.

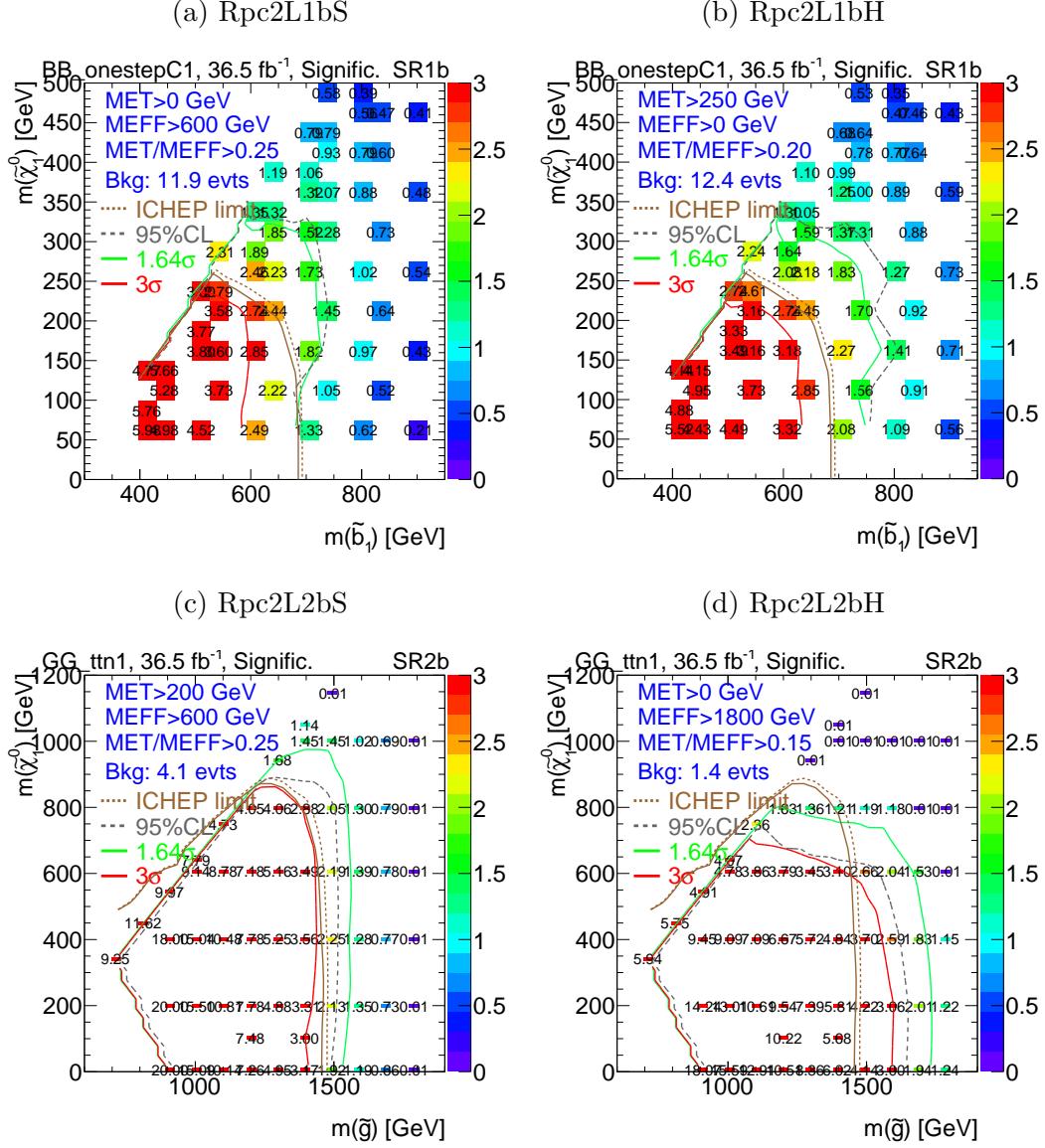


Figure 6.4: Discovery significance for the SRs with b -jets defined in Table ?? for 10 fb^{-1} : Rpc2L1bS and Rpc2L1bH in the $\tilde{b}_1\tilde{b}_1^* \rightarrow t\bar{t}\tilde{\chi}_1^+\tilde{\chi}_1^-$ grid (top), Rpc2L2bS and Rpc2L2bH in the $\tilde{g}\tilde{g} \rightarrow t\bar{t}t\bar{t}\tilde{\chi}_1^0\tilde{\chi}_1^0$ grid (bottom). The 2015 limits from the SS/3L analysis in those models are shown with a brown line, and the 95% CL, 1.64σ , and 3σ discovery contours from the proposed signal regions are shown in grey, green, and red, respectively.

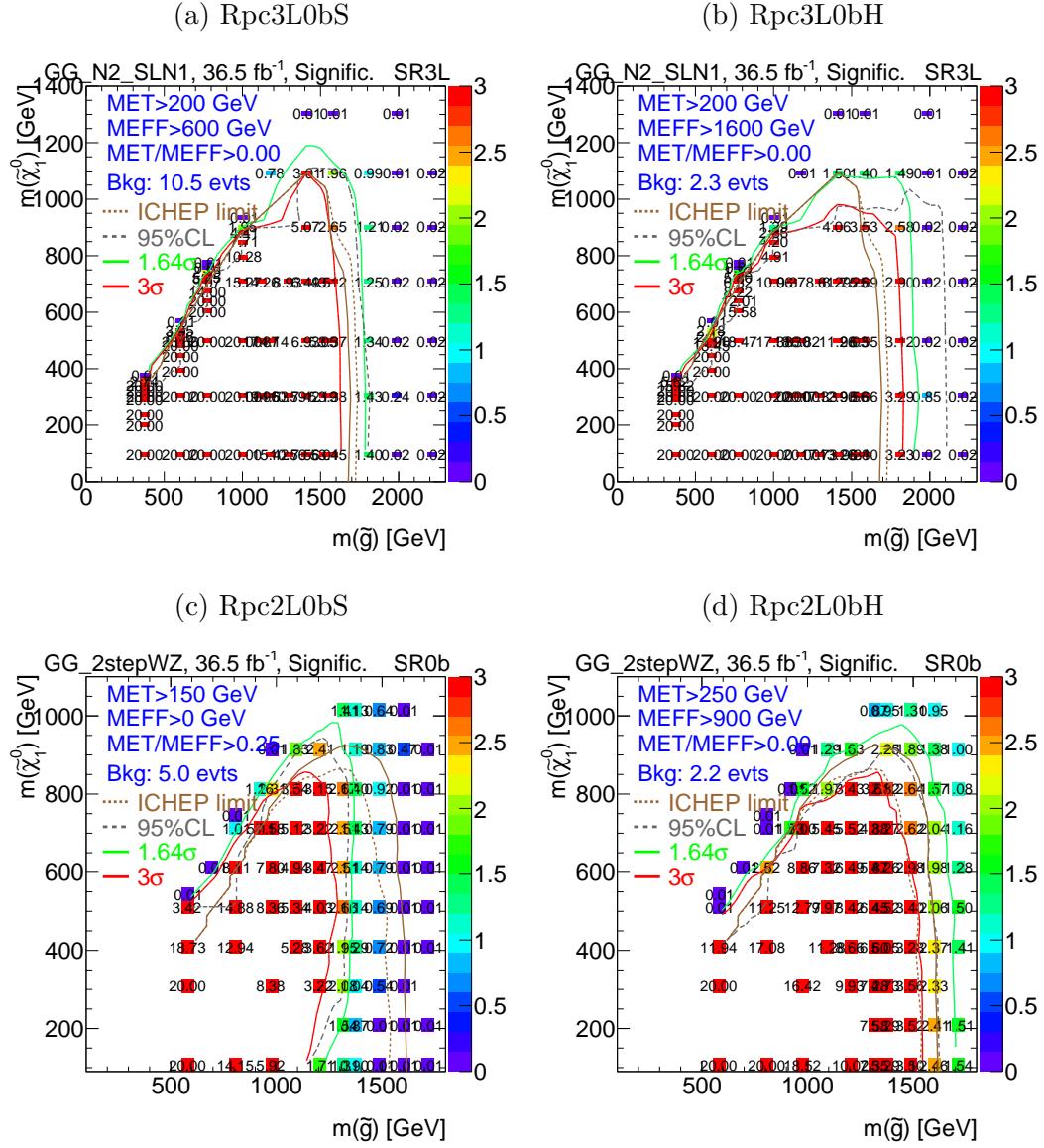


Figure 6.5: Discovery significance for the SRs without b -jets defined in Table ??

for 36.5 fb^{-1} : Rpc2L0bS and Rpc2L0bH in the $\tilde{g}\tilde{g}$ with $\tilde{g} \rightarrow q\bar{q}'WZ\tilde{\chi}_1^0$ grid (top)

and Rpc3L0bS and Rpc3L0bH in the $\tilde{g}\tilde{g}$ with $\tilde{g} \rightarrow q\bar{q}(\ell\ell/\ell\nu)\tilde{\chi}_1^0$ grid (bottom).

The ICHEP 2016 limits from the SS/3L analysis in those models are shown with a brown line, and the 95% CL, 1.64σ , and 3σ discovery contours from the proposed signal regions are shown in grey, green, and red, respectively.

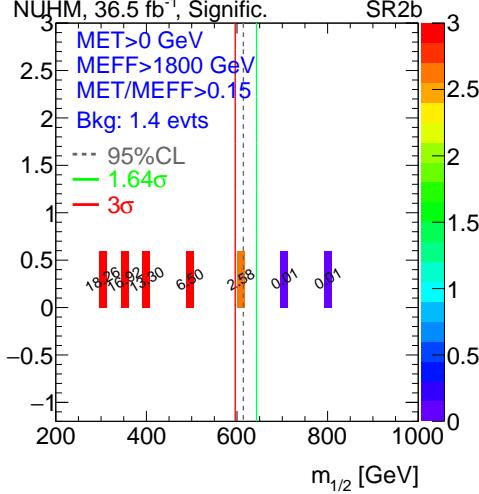
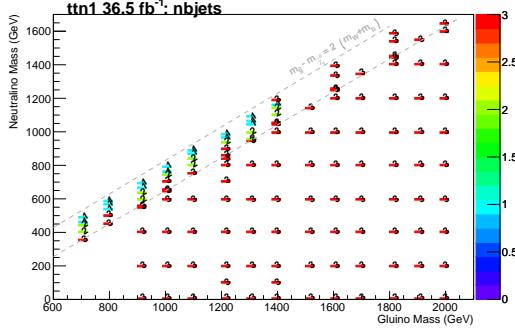


Figure 6.6: Discovery significance for Rpc2L2bH signal region for 36.5 fb^{-1} , NUHM2 model. The 95% CL, 1.64σ , and 3σ discovery contours from the proposed signal regions are shown in grey, green, and red, respectively.

Dedicated new SRs have been optimized for the gluino pair production with stop-mediated decay $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$ with off-shell tops.

The $\tilde{g}\tilde{g}$ production with $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$ scenario at low LSP masses (where the multi- b analysis has a much better sensitivity) is not the only motivation for Rpc2L2bH signal region, but also the NUHM2 model, which features large branching ratios for the $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_{1,2,3}^0$ and $\tilde{g} \rightarrow tb\tilde{\chi}_{1,2}^\pm$ decays. As shown in Figure 6.6, with the Rpc2L2bH signal region $m_{1/2}$ values of 600 GeV can be excluded at 95% CL or observed with a 3σ significance, and we plan to use this SR for the first interpretation in this model.

In addition, the SS/3L analysis has the unique potential to explore the region of phase space at high LSP masses with a more compressed spectra. This scenario leads to softer decay products, in particular softer b -jets as seen in Figure 6.7,

(a) $m_{\tilde{g}} - m_{\tilde{\chi}_1^0}$ mass plane

(b) Above the diagonal

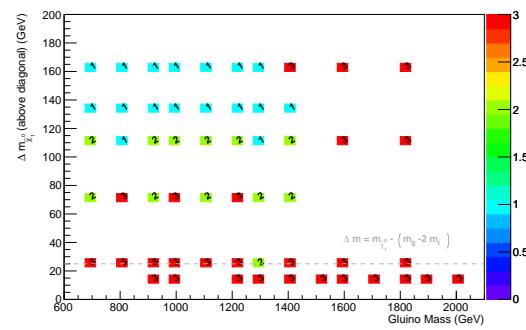


Figure 6.7: Optimal cut on the number of b -jets leading to the best discovery significance.

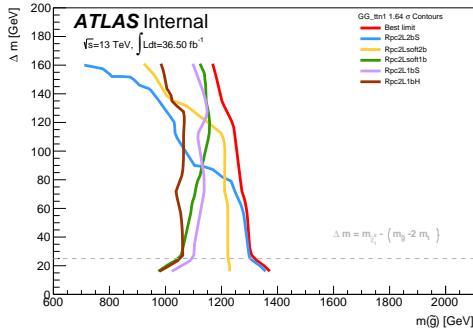
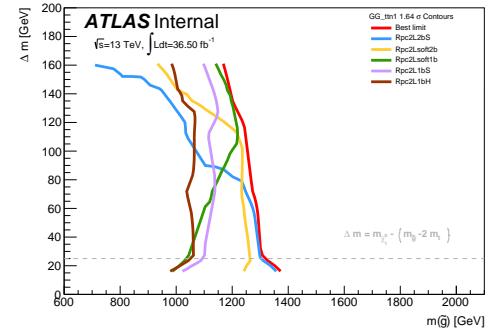
(a) Without $p_T^{\ell_1}$ uppercut(b) With $p_T^{\ell_1}$ uppercut

Figure 6.8: Comparison of significance contours at 1.64σ for 36.5 fb^{-1} between Rpc2Lsoft2b and Rpc2Lsoft1b and other signal regions in the off-diagonal region considering the option of an upper cut on the leading lepton p_T .

which makes the multi- b analysis less sensitive. For this reason, we introduce two additional signal region with at least 1 b -jet (Rpc2Lsoft1b) or 2 b -jets (Rpc2Lsoft2b) defined in Table 6.7. In addition, these signal regions are defined with an upper cut on the leading lepton p_T . The sensitivity is degraded if we remove this upper cut as shown in Figure 6.8.

Table 6.7: Definitions of the RPC signal regions for 36.5 fb^{-1}

SR	N_ℓ	$N_{b\text{-jets}}^{20}$	N_{jets}	p_T^{jets}	E_T^{miss} [GeV]	m_{eff} [GeV]	$E_T^{\text{miss}}/m_{\text{eff}}$	Other
Rpc2Lsoft2b	≥ 2	≥ 2	≥ 6	25	>200	>600	>0.25	$20 < p_T^{\ell_1} < 100 \text{ GeV}$ $p_T^{\ell_2} > 10 \text{ GeV}$
Rpc2Lsoft1b	≥ 2	≥ 1	≥ 6	25	>100	-	>0.3	
Rpc3LSS1b	$\geq 3(\text{SS})$	≥ 1	-	-	-	-	-	veto $81 < m_{e^\pm e^\pm} < 101 \text{ GeV}$
Rpc3L1bS	≥ 3	≥ 1	≥ 4	40	>200	>600	-	-
Rpc3L1bH	≥ 3	≥ 1	≥ 4	40	>200	>1600	-	-

Motivated by the \tilde{t} production with $\tilde{t}_1 \rightarrow \tilde{\chi}_2^0 W$ model in Section 6.2.5, we will explore for the first time the signature of three leptons with the same electric charge (3LSS). As shown in Figure 6.9, after an inclusive 3LSS selection, the background is dominated by dibosons and $Z+\text{jets}$ (with only one real lepton, and the two other leptons with either an electron with charge mis-identified or a fake lepton) both dominantly without b -jets. Once a b -jet requirement is applied, the background is dominated by $t\bar{t}V$, with a clear peak at $m_{\ell\ell} \approx m_Z$ showing that a large fraction of these events are originated from charge mis-identification from events containing $Z \rightarrow ee$. After applying a $81 < m_{e^\pm e^\pm} < 101 \text{ GeV}$ veto, the background is reduced to only 1.7 events for 36.5 fb^{-1} , almost removing the $Z+\text{jets}$ and diboson backgrounds completely. The final background is dominated

by $t\bar{t} + H, Z, W$, with $\approx 60\%$ originating from charge flips and $\approx 40\%$ from fakes and non-prompt leptons. With these very generic selections (Rpc3LSS1b in Table 6.7), a significance of 3.7σ can be obtained for $m_{\tilde{t}} = 550$ GeV. Figure 6.10 shows some lepton distributions, including the number of electrons, where most of the charge flip background populates the bins with 2 or 3 electrons, although cutting away those bins would also have a large impact on the signal.

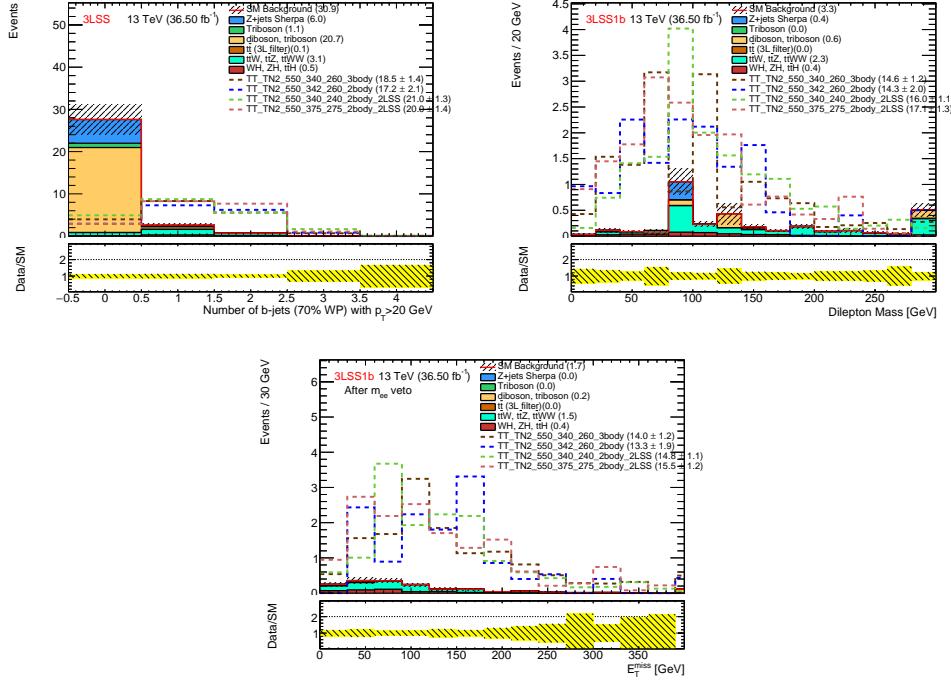


Figure 6.9: b -jet multiplicity after a 3LSS selection (top left), dilepton invariant mass distributions after a 3LSS plus ≥ 1 b -jet selection (top right), and E_T^{miss} distribution after a 3LSS, ≥ 1 b -jet and $81 < m_{e^\pm e^\pm} < 101$ GeV veto selection (bottom), all corresponding to 36.5 fb^{-1} . The background distributions are stacked, while the lines show the predictions for four signal points at \tilde{t} mass of 550 GeV.

Finally, since the SRs defined for the $\tilde{g}\tilde{g}$ production with $\tilde{g} \rightarrow q\bar{q}\ell\bar{\ell}\chi_1^0$ feature

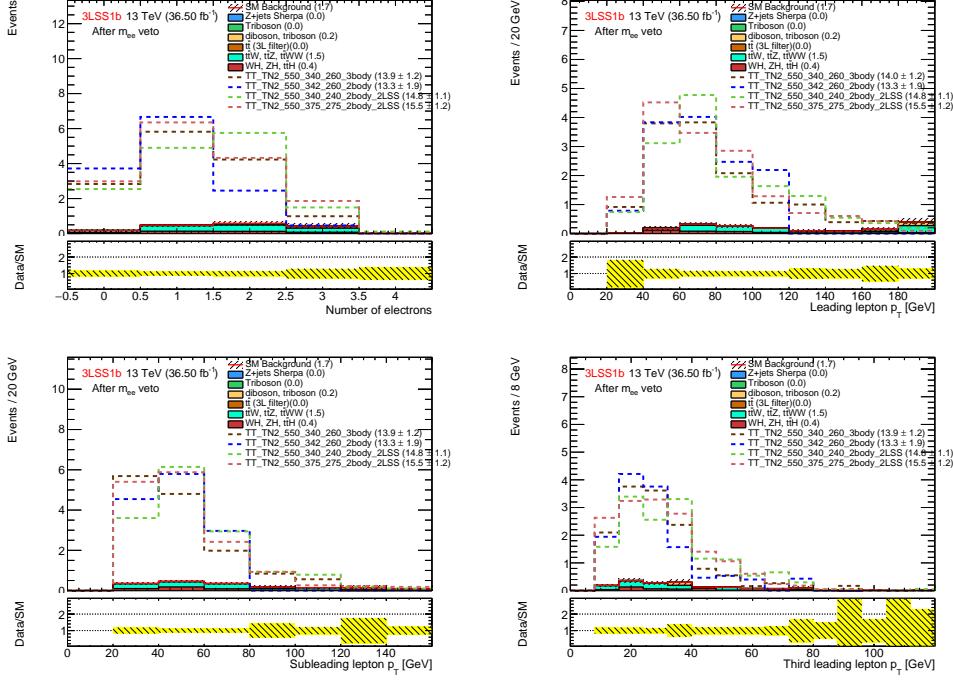


Figure 6.10: Number of electron (top left), and p_T of the leading (top right), subleading (bottom left) and third leading lepton (bottom right) after a 3LSS, ≥ 1 b -jet and $81 < m_{e^\pm e^\pm} < 101$ GeV veto selection (bottom), all corresponding to 36.5 fb^{-1} . The background distributions are stacked, while the lines show the predictions for four signal points at \tilde{t} mass of 550 GeV.

a b -jet veto (Rpc3L0bS and Rpc3L0bH), and to avoid leaving uncovered the 3 lepton plus b -jets signature, SRs with the same kinematic cuts as Rpc3L0bS and Rpc3L0bH but with a ≥ 1 b -jet requirement are also proposed in Table 6.7 as Rpc3L1bS and Rpc3L1bH.

6.6 Background estimation

In this analysis, two types of backgrounds can be distinguished. The first category is the reducible background, which includes events containing electrons with mismeasured charge, mainly from the production of top quark pairs, and events containing at least one fake or non-prompt (FNP) lepton that was discussed in detail in Chapter ???. Data-driven methods used for the estimation of this reducible background in the signal and validation regions are described in Section 6.6.10. The second category is the irreducible background from events with two same-sign prompt leptons or at least three prompt leptons and is estimated using the MC simulation samples. Since diboson and $t\bar{t}V$ events are the main backgrounds in the signal regions, dedicated validation regions (VR) with an enhanced contribution from these processes, and small signal contamination, are defined to verify the background predictions from the simulation (Section ??).

6.6.1 Irreducible background estimation methods

6.6.2 Associate $t\bar{t} + W/Z$ production

The production of a $t\bar{t}$ pair in association with a leptonically-decaying W or Z boson (including non-resonant contributions) constitutes the main source of background with prompt same-sign leptons for event selections including b -jets.

Details on event generation and normalization

Simulated events for these processes were generated at NLO with AMC@NLO

v2.2.2 [?] interfaced to PYTHIA 8, further details can be found in Ref. [?]. The samples are normalised to the inclusive process NLO cross-section using appropriate k -factors computed by the PMG group [?, ?] with a methodology similar to Ref. [?].

Theory uncertainties

The theoretical uncertainties on the ttW and $ttZ^{(*)}$ processes are evaluated following the recommendations of the Physics Modelling Group and using the LHE3 weights included within the nominal MC samples. The variation weights considered are the following:

- Normalization and factorization scales varied independently up and down by a factor of two from the central scale $\mu_0 = H_T/2$ as detailed in Ref. [?].

The largest deviation with respect to the nominal is used as the symmetric uncertainty.

- PDF variations
- Alternative SHERPA (v2.2) samples produced using LO with 1 extra parton in the matrix element for ttW and 2 extra partons for ttZ [?]. The yield comparison for all SRs is shown in Table 6.8, with negligible differences in some SRs and up to 28% in the worst case.

More details can be found in Appendix ???. Based on these studies and the cross-section uncertainties, the total theory uncertainty for these processes is at

the level of 13–33% in the signal and validation regions used in the analysis. In addition, the cross-sections used to normalise the MC samples are varied according to the uncertainty on the cross-section calculation, that is, 13% for $t\bar{t}W$ and 12% $t\bar{t}Z$ production [?].

6.6.3 Diboson $WZ, ZZ, W^\pm W^\pm$ production

The production of multiple W, Z bosons decaying leptonically constitutes the main source of background with prompt same-sign leptons for event selections vetoing b -jets.

Details on event generation and normalization

Diboson processes with four charged leptons, three charged leptons and one neutrino, or two charged leptons and two neutrinos were simulated at NLO using the SHERPA 2.2.1 generator [?], as described in detail in Ref. [?]. The main samples simulate $qq \rightarrow VV \rightarrow$ leptons production including the doubly resonant WZ and ZZ processes, non-resonant contributions as well as Higgs-mediated contributions, and their interferences; up to three extra partons were included (at LO) in the matrix elements. Note that these samples include the contributions from tri-boson production with one of the vector bosons decaying hadronically, while tri-boson production with only leptonic decays was simulated separately (see section 6.6.4).

Simulated events for the $W^\pm W^\pm jj$ process (including non-resonant contributions) were produced at LO with up to one extra parton, separately for QCD-

induced ($\mathcal{O}(\alpha_{\text{em}}^4)$) and VBS-induced ($\mathcal{O}(\alpha_{\text{em}}^6)$) production – the interferences being neglected. Additional samples for VBS-induced $qq \rightarrow 3\ell\nu jj$ and $qq \rightarrow 4\ell$ and loop-induced $gg \rightarrow WZ^{(*)}/ZZ^{(*)}$ processes were also produced with the same configuration.

The samples generated at NLO are directly normalized to the cross-sections provided by the generator. The complete list of samples and cross-sections can be found in appendix ??, Table ??.

Theory uncertainties

The theoretical uncertainties on the WZ and ZZ processes are evaluated following the recommendations of the Physics Modelling Group and using the LHE3 weights included within the nominal MC samples where available. The variation considered are the following:

- Normalization and factorization scales varied independently up and down by a factor of two from the central scale choice. The largest deviation with respect to the nominal is used as the symmetric uncertainty.
- PDF variations
- Resummation scale varied up and down by a factor of two from the nominal value.
- CKKW merging scale varied up and down to a value of 15 and 30 GeV (with a value of 20 GeV used in the nominal samples).

More details can be found in Appendix ???. Based on these studies and the cross-section uncertainties, the total theory uncertainty for these processes is at the level of 25-40% in the signal and validation regions used in the analysis. In addition, the cross-sections used to normalise the MC samples are varied according to the uncertainty on the cross-section calculation of 6%.

No theoretical uncertainties have been evaluated specifically for the $W^\pm W^\pm jj$ process, to which we assign the same uncertainties as for WZ , by lack of a better choice. But it should be noted that contributions from this process are minor in the SRs and typically smaller than those from WZ and ZZ .

6.6.4 Other rare processes

Production of a Higgs boson in association with a $t\bar{t}$ pair is simulated using **AMC@NLO** [?] (in **MADGRAPH** v2.2.2) interfaced to **HERWIG** 2.7.1 [?]. The UEEE5 underlying-event tune is used together with the CTEQ6L1 [?] (matrix element) and CT10 [?] (parton shower) PDF sets. Simulated samples of SM Higgs boson production in association with a W or Z boson are produced with **PYTHIA** 8.186, using the A14 tune and the NNPDF23LO PDF set. Events are normalised with cross-sections calculated at NLO [?].

MADGRAPH v2.2.2 [?] is used to simulate the $t\bar{t}WW$, tZ , $t\bar{t}t\bar{t}$ and $t\bar{t}t$ processes, and the generator cross-section is used for tZ and $t\bar{t}t$. **MADGRAPH** interfaced to **PYTHIA** 8 is used to generate $t\bar{t}WZ$ processes, and appropriate k -factors are taken from [?]. **AMC@NLO** interfaced to **PYTHIA** 8 is used for the generation of the tWZ process, with an alternative sample generated with **AMC@NLO** interfaced

to HERWIG used to evaluate the parton shower uncertainty.

Fully leptonic triboson processes (WWW , WWZ , WZZ and ZZZ) with up to six charged leptons are simulated using SHERPA v2.1.1 and described in Ref. [?]. The 4ℓ and $2\ell + 2\nu$ processes are calculated at next-to-leading order (NLO) for up to one additional parton; final states with two and three additional partons are calculated at leading order (LO). The $WWZ \rightarrow 4\ell + 2\nu$ or $2\ell + 4\nu$ processes are calculated at LO with up to two additional partons. The $WWW/WZZ \rightarrow 3\ell + 3\nu$, $WZZ \rightarrow 5\ell + 1\nu$, $ZZZ \rightarrow 6\ell + 0\nu$, $4\ell + 2\nu$ or $2\ell + 4\nu$ processes are calculated is calculated at NLO with up to two extra partons at LO. The CT10 [?] parton distribution function (PDF) set is used for all SHERPA samples in conjunction with a dedicated tuning of the parton shower parameters developed by the SHERPA authors. The generator cross-sections (at NLO for most of the processes) are used when normalising these backgrounds.

A conservative 50% uncertainty is assigned on the summed contributions of all these processes ($t\bar{t}H$, tZ , tWZ , $t\bar{t}t\bar{t}$, $t\bar{t}WW$, $t\bar{t}WZ$, WH , ZH , VVV), which is generally quite larger than the uncertainties on their inclusive production cross-sections, and assumes a similar level of mismodelling as for diboson or $t\bar{t}V$ processes.

6.6.5 $W^\pm W^\pm$ production via Double Parton Scattering

Double parton scattering (DPS) occurs when two partons interact simultaneously in a proton-proton collision leading to two hard scattering processes overlapping in a detector event. Accordingly, two single W production processes can lead to a

$W^\pm + W^\pm$ final state via DPS. In the 2015 version of the present analysis, the background originated from $W^\pm W^\pm$ production via DPS was roughly estimated by overlaying at analysis level events from the SHERPA W +jets samples. The estimated contribution was at the level of 0.001-0.01 events for 4 fb^{-1} in a SR similar to SR0b1-2 (b -jet veto), although slightly looser since it only required 5 jets.

DPS effects are implemented in Sherpa diboson samples used in the analysis in the case of $V + jj$, but not for $W^\pm + W^\pm$. In the current analysis, the sensitivity to this background and its contribution to the signal regions have been revisited using simulated DPS $WW \rightarrow \ell\nu\ell\nu$ events (with all possible sign combinations) generated at LO with PYTHIA8EVTGEN. In addition, a sample with single W +jets events produced via SPS has been generated with the same Monte Carlo framework. This sample was used together with a SHERPA W +jets sample to derive SHERPA/PYTHIA correction factors to address the insufficient prediction from PYTHIA 8 at high jet multiplicities. These factors were then applied to the actual DPS sample in order to provide a corrected Monte Carlo based estimation of the DPS $W^\pm W^\pm$ process in the signal regions. A detailed description of the samples used according validation plots can be found in appendix ??.

To estimate a conservative upper bound on cross-section for WW events which might arise from DPS, a standard ansatz is adopted: in this, for a collision in which a hard process (X) occurs, the probability that an additional (distinguishable)

process (Y) occurs is parametrized as:

$$\sigma_{XY}^{DPS} = \sigma_X \sigma_Y / \sigma_{eff} \quad (6.1)$$

where σ_X is the production cross section of the hard process X and σ_{eff} (effective area parameter) parameterizes the double-parton interaction part of the production cross section for the composite system (X+Y). A value of σ_{eff} is 10-20 mb is assumed in this study (as obtained from 7 TeV measurements, and with no observed dependence on \sqrt{s}), and it is independent on the processes involved. For the case of $W^\pm + W^\pm$ production:

$$\sigma_{W^\pm W^\pm}^{DPS} = \frac{\sigma_{W^+} \sigma_{W^+} + \sigma_{W^-} \sigma_{W^-} + 2\sigma_{W^+} \sigma_{W^-}}{\sigma_{eff}} \simeq 0.19 - 0.38 \text{ pb.} \quad (6.2)$$

After the application of the SR criteria, only 4 raw MC events in the DPS $WW \rightarrow \ell\nu\ell\nu$ remain. Table 6.9 shows the expected contribution in the SRs where some MC event survives all the cuts. The ranges quoted in the tables reflect the range in the predicted $\sigma_{W^\pm W^\pm}^{DPS}$ cross-section above, as well as the combinatorics for scaling the jet multiplicity⁶. Due to the large uncertainties involved in these estimates, some of them difficult to quantify (such as the modelling of DPS by PYTHIA at LO), the contribution from this background is not included in the final SR background estimates. Note that the estimated DPS contribution is typically much smaller than the uncertainty on the total background for the SRs.

⁶For instance, a DPS event with 6 jets can be due to the overlap of two events with 6+0 jets, or 5+1, 4+2 or 3+3 jets. All possible combinations are considered and the range quoted in the table shows the combinations leading to the smallest and largest correction factors.

6.6.6 Experimental uncertainties

All the experimental systematics provided by the SUSYTools `getSystInfoList()` method have been considered. The list of sources of uncertainty and the corresponding names of the variations are:

Jet energy scale (`Jet.JESNPSet:1`)

One of the strongly reduced uncertainty sets provided by the JetEtMiss group for Run-2 searches is used in this note. These sets are intended for use by analyses which are not sensitive to jet-by-jet correlations arising from changes to the jet energy scale (as expected for many early SUSY searches), and we use the scenario `JES2015_SR_Scenario1.config` (as included in the JetUncertainties package, 4 nuisance parameters). We checked that the uncertainties obtained from one of the 3 other scenarios did not lead to significant changes. The jet energy is scaled up and down (in a fully correlated way) by the $\pm 1\sigma$ uncertainty of each nuisance parameter.

Jet energy resolution (`JET_JER_SINGLE_NP_1up`)

An extra p_T smearing is added to the jets based on their p_T and η to account for a possible underestimate of the jet energy resolution in the MC simulation. This is done by the `JERSmearingTool` in the JetResolution package. We are using the `Simple` systematic mode configuration (1 nuisance parameter).

Jet vertex tagger (`JvtEfficiency{down,up}`)

The uncertainties are applied via the `JetJvtEfficiency` tool, which account for the residual contamination from pile-up jets after pile-up suppression and the MC

generator choice.

Flavor tagging (`FT_EFF_{B,C,Light}_systematics_1{up,down}`, `FT_EFF_extrapolation_1{up,down}`,
`FT_EFF_extrapolation_from_charm_1{up,down}`)

Similarly to the case of the JES, a significant reduction in the number of nuisance parameters was provided by the Flavour Tagging CP group and used in this analysis.

Egamma resolution (`EG_RESOLUTION_ALL_1{up,down}`) and **scale** (`EG_SCALE_ALL_1{up,down}`)

As the analysis is weakly sensitive to the energy scale of electrons, the special simplified correlation model (`1NPCOR_PLUS_UNCOR`) is used. It has only 2 systematic variations, one for the scale, one for the resolution. In this scheme all the effects are considered fully correlated in η and they are summed in quadrature.

Electron efficiency (`EL_EFF_{RECO, ID, Iso, CFT}_TOTAL_1NPCOR_PLUS_UNCOR_1{up,down}`)

These uncertainty sources are associated with the electron efficiency scale factors provided by the Egamma CP group. The default correlation model (`TOTAL`) is used, which provides 1 uncertainty sources, for reconstruction, identification, isolation and charge flip tagger, separately.

Muon efficiency (`MUON_EFF_{STAT, SYS}_1{up,down}`, `MUON_EFF_{STAT, SYS}_LOWPT_1{up,down}`,
`MUON_ISO_{STAT, SYS}_1{up,down}`)

This uncertainty corresponds to the statistical and systematic uncertainties in the muon efficiency scale factors provided by the Muon CP group on the muon reconstruction and isolation. The `*_LOWPT_*` sources of uncertainties are associated to muons with $p_T < 15\text{GeV}$.

Muon resolution uncertainty (`MUONS_ID_1{up,down}`, `MUONS_MS_1{up,down}`)

This is evaluated as variations in the smearing of the inner detector and muon spectrometer tracks associated to the muon objects by $\pm 1\sigma$ their uncertainty

Muon momentum scale (`MUONS_SCALE_-1{up,down}`)

This is evaluated as variations in the scale of the momentum of the muon objects.

Muon scale corrections (`MUON_SAGITTA_RHO_-1{up,down}`, `MUON_SAGITTA_RESBIAS_-1{up,down}`)

All analyses are recommended to set the “SagittaCorr” property to true. The flag corrects data for charge dependent local effects due to misalignments mainly in the ID and smaller local effects due to local misalignments in the MS. The MUON_SAGITTA_RHO systematics describe a variation in the scale of the momentum (charge dependent), based on combination of correction on combined (Z scale) and recombination of the corrections. The MUON_SAGITTA_RESBIAS systematic describe a variation in the scale of the momentum (charge dependent), based on the residual charge-dependent bias after correction.

Muon TTVA (`MUON_TTVA_{STAT,SUSY}_-1{up,down}`)

Uncertainties associated to the TTVAs (track to vertex association) SFs for 2016 data (periods A-I) to account for worse ID d_0 resolution, resulting in lower efficiency and φ -dependent SFs, 1-2% effects (new SFs are parametrized in $p_T - \eta - \varphi$, before only in $p_T - \eta$).

Bad muon veto (`MUON_BADMUON_{STAT,SYST}_-1{up,down}`)

Systematic associated to the bad-muon veto for high-pT muons (to reject tracks with poor resolution in the MS).

E_T^{miss} soft term uncertainties (`MET_SoftTrk_Reso{Pare,Perp}`, `MET_SoftTrk_Scale{up,down}`)

Note that the effect of the hard object uncertainties (most notably JES and JER)

are also propagated to the E_T^{miss} .

Pileup reweighting (PRW_DATASF_1{up,down})

This uncertainty is obtained by re-scaling the μ value in data by 1.00 and 1/1.18, covering the full difference between applying and not-applying the nominal μ correction of 1/1.09, as well as effects resulting from uncertainties on the luminosity measurements, which are expected to dominate.

Luminosity

See section ??.

Trigger

Trigger uncertainties provided by CP groups and propagated via `TrigGlobalEfficiencyCorrection` package.

The uncertainty on the beam energy is neglected.

All the experimental uncertainties are applied also on the signal samples when computing exclusion limits on SUSY scenarios.

6.6.7 Expected yields in the signal regions

The predicted event yields in the signal regions are presented in Table 6.10, while the contributions of particular rare processes to the signal regions, relative to the summed contributions of all these processes, are shown in Table 6.11.

6.6.8 Reducible background estimation methods

Two data-driven methods are used to estimate the FNP lepton background, referred to as the “matrix method” and the “MC template method”. The estimates from these methods are combined to give the final estimate. These two methods are described below. A detailed description of both methods was given in Chapter ???. In this section, more details will be given on their application for this analysis.

6.6.9 Matrix Method

6.6.10 Reducible background estimation methods

Two data-driven methods are used to estimate the FNP lepton background, referred to as the “matrix method” and the “MC template method”. The estimates from these methods are combined to give the final estimate. These two methods are described below. A detailed description of both methods was given in Chapter ???. In this section, more details will be given on their application for this analysis.

6.6.11 Matrix Method

The matrix method relates the number of events containing prompt or FNP leptons to the number of observed events with tight or loose-not-tight leptons using the probability for loose prompt or FNP leptons to satisfy the tight criteria, also referred to as the fake rate. The probability for loose prompt leptons to satisfy the tight selection criteria (ε), also referred to as the real efficiency. The procedure employed to measure the fake rate and real efficiency is described in

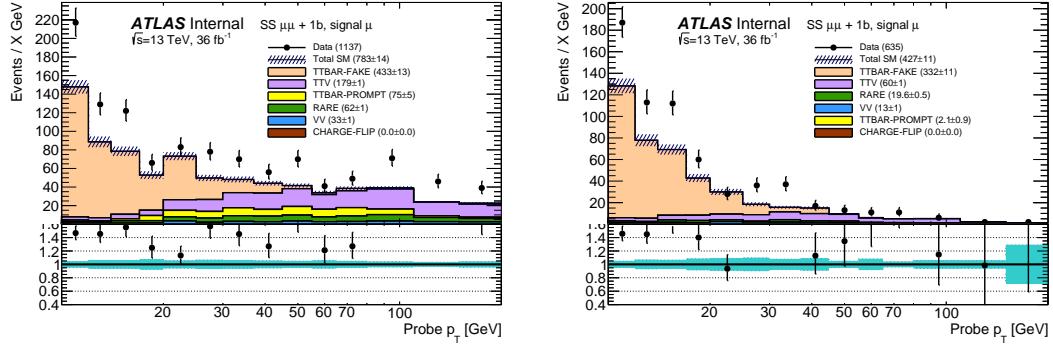


Figure 6.11: Signal probe muon p_T distribution in data and MC, after preselection (left) or further tightening of the tag muon requirements (right), as described in section 6.6.12. The yellow area indicates $t\bar{t}$ events in which the tag muon is fake and the probe real, leading to a measurement bias.

the next sections.

6.6.12 Baseline-to-signal efficiency for fake muons

Baseline-to-signal efficiency for fake leptons (further called “fake rate”) is measured in a sample enriched in fake leptons from $t\bar{t}$ processes, as simulations indicate that this is the largest contribution to fake lepton background in the signal regions, even those with b -jet vetoes, due to the requirements on jet multiplicity and E_T^{miss} . For that we select events with exactly two same-sign muons (and no extra baseline lepton), at least one b -jet, and at least 3 jets, acquired with the mu18_mu8noL1 (2015 data) or mu22_mu8noL1 (2016 data) trigger chains. One of the muons in the event (referred to as “tag”) is required to satisfy signal requirements, verify $p_T > 25\text{GeV}$, and trigger the event recording. The measurement may then be performed on the other lepton (“probe”), likely to be the fake lepton of the pair.

Figure 6.11 left shows the number of signal muon probes available after this preselection. A similar level of disagreement between data and simulation as reported elsewhere in this document is observed, with data exceeding the MC predictions; this is not a concern for the purpose of this section, of course. One can notice that at this stage, measurements above 25 GeV would be very affected by the important fraction of events in which the tag muon is fake and the probe muon is real. To overcome this issue, we considered three alternatives:

- tighten the p_T and isolation requirements of the tag muon beyond those defining our usual “signal” category, to reduce its probability of being a fake muon
- use an identical selection for tag and probe muons, and require them to be in the same (p_T, η) bin for the measurement; after subtraction of estimated contributions from processes with two prompt muons, all events have one real and one fake muon, and the symmetry in the muon selection can be taken advantage of to obtain an unbiased measurement of the fake rate:

$$\zeta = \frac{\varepsilon n_2}{\varepsilon n_1 + (2\varepsilon - 1)n_2}$$

with n_1, n_2 the number of events with 1 or 2 signal muons, and ε the efficiency for prompt

This method is limited to measurements in inclusive or wide bins. It also can't be used at too low p_T , due to contributions from processes with two fake muons (e.g. from $B\bar{B}$ meson production).

- a generalization of the former recovering events in which the two muons may

also be in different (p_T, η) bins, numerically solving a system of non-linear equations to determine simultaneously the fake rates in different bins in an unbiased way. Upgraded from asymptotic equalities to statistical estimates, this was in fact implemented as the maximization of a likelihood built from a product of multinomial PDFs.

Details on the implementation and performances of these three methods are provided in appendix ???. Comparisons made with $t\bar{t}$ MC indicated that when using a very tight isolation requirement on the tag muon ($\max(E_T^{\text{topo, cone } 40}, p_T^{\text{cone } 40}) < 0.02 \times p_T$), the level of bias is always largely inferior to the statistical uncertainty in the measurement, which itself is smaller than for the other two methods.

Figure 6.11 right shows the number of signal muon probes when applying those reinforced isolation criteria to the tag muon, as well as requiring $p_T^{\text{tag}} > \max(40, p_T^{\text{probe}} + 10)$ GeV. As expected, the number of pairs with a fake tag muons is down to a minor level, at least according to the simulation. Similar distributions for baseline probe muons can be found in appendix ???.

Muon fake rates as predicted by the simulation ($t\bar{t}$, inclusive selection of leptons via truth-matching) are shown on Fig. 6.12 as function of p_T and $|\eta|$. One can expect a moderate dependency of the fake rates to the transverse momentum, with the strongest evolution at low p_T and a slight increase toward higher p_T . The fake rates are also essentially independent of the pseudorapidity, except at the edge ($|\eta| > 2.3$) where there is a strongly pronounced increase of the rates. This motivates measurements in data as function of p_T in two $|\eta|$ bins.

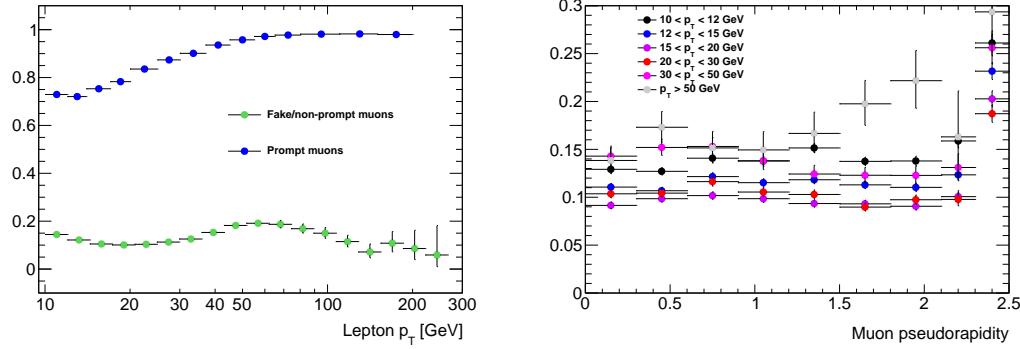


Figure 6.12: Muon fake rates in $t\bar{t}$ MC with an inclusive selection, as function of p_T (left, green markers) or $|\eta|$ in different momentum ranges (right).

Observations in data seem to indicate that the rejection of fake tag muons by the reinforced isolation criteria is quite less important than in the simulation, or that the amount of fake muons at high p_T is larger than in the simulation, or both. This leads to an unknown level of bias in measurements performed with the straightforward tag-and-probe selection at high p_T . For that reason, the final rates measured in data are provided by the tag-and-probe method below 25 GeV, and by the symmetric selection for $p_T > 25 \text{ GeV}$. The former are obtained with

$$\zeta = \frac{n_{\text{signal}}^{\text{data}} - n_{\text{signal}}^{\text{MC}}}{n_{\text{baseline}}^{\text{data}} - n_{\text{baseline}}^{\text{MC}}} \quad (6.3)$$

with $\Delta\zeta_{\text{stat}} = \frac{\sqrt{(1-2\zeta)n_{\text{signal}}^{\text{data}} + \zeta^2 n_{\text{baseline}}^{\text{data}}}}{n_{\text{baseline}}^{\text{data}} - n_{\text{baseline}}^{\text{MC}}}$

while the latter are obtained with:

$$\zeta = \frac{\varepsilon(n_{\text{both signal}}^{\text{data}} - n_{\text{both signal}}^{\text{MC}})}{\varepsilon(n_{\text{only 1 signal}}^{\text{data}} - n_{\text{only 1 signal}}^{\text{MC}}) + (2\varepsilon - 1)(n_{\text{both signal}}^{\text{data}} - n_{\text{both signal}}^{\text{MC}})} \quad (6.4)$$

with $\Delta\zeta_{\text{stat}} = \frac{\zeta}{n_{\text{both signal}}^{\text{data}} - n_{\text{both signal}}^{\text{MC}}} \sqrt{\zeta^2 n_{\text{only 1 signal}}^{\text{data}} + \left(1 - \frac{2\varepsilon - 1}{\varepsilon}\zeta\right)^2 n_{\text{both signal}}^{\text{data}}}$

the efficiency for prompt muons ε is assigned values compatible with section 6.6.14.

The measured rates are presented in Table 6.12. The central values are shown together with the associated statistical uncertainty, as well as the propagation of the uncertainty on the subtracted backgrounds normalization, which is taken as a global $\Delta B/B = 20\%$. The rates are of the order of 10% up to 30 GeV, beyond which they increase. Overall these values are not very different from those predicted by the simulation.

Complementary information (event yields for data and background processes, estimates of the level of bias and contributions from QCD double-fakes) can be found in appendix ???. The observed rates are quite in agreement with the simulation up to $p_T \sim 50\text{GeV}$.

Some of the validation and signal regions require events with 2 or more b -tagged jets, which reduces the fraction of non-prompt muons coming from B meson decays. Figure 6.13 illustrates how this impacts on the fake rates. Given the good agreement between data and simulation for the measured values, we chose to apply a correction to the measured rates for events with ≥ 2 b -jets, taken directly from simulated $t\bar{t}$ events. This correction factor varies between 1 and 2 with p_T , and the whole size of the correction is assigned as an additional systematic uncertainty (see Table 6.13).

Systematic uncertainties

To cover potential differences in the fake rates between the measurement regions and the signal regions, that could be due to different origins or kinematic properties of the fake leptons, uncertainties are set based on the extent of those differences

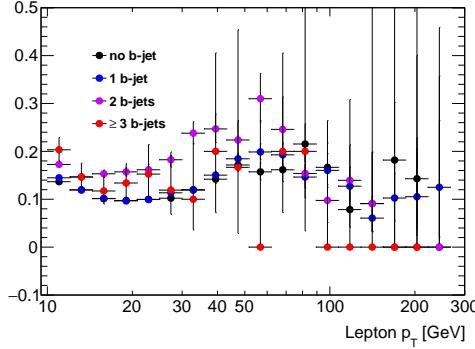


Figure 6.13: Muon fake rates in $t\bar{t}$ MC with an inclusive selection, as function of p_T and split according to the number of b -tagged jets in the event.

predicted by the simulation. The largest effect is the decrease of the fake rates with HT (especially for high- p_T muons), which likely correlates to a harder jet at the origin of the non-prompt muon, hence a reduced likelihood to satisfy isolation requirements. Details can be consulted in appendix ??, and Table 6.13 summarizes the additional systematic uncertainties applied to the muon fake rates. They vary from 30% at low p_T , to up to 85% for $p_T > 40\text{GeV}$; in that range, the uncertainties are made H_T -dependent.

As already shown, Fig. 6.13 shows the variation of the fake rate in $t\bar{t}$ MC as function of the number of b -tagged jets in the event. Unsurprisingly, the rates are very similar for $0b$ and $\geq 1b$ final states, justifying the use of the fake rates measured in this section (i.e. in a $\geq 1b$ region) to predict fake muon background in all signal regions.

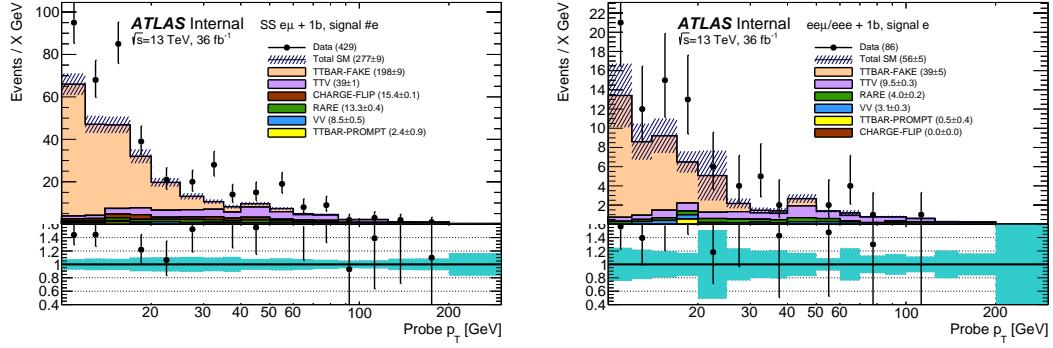


Figure 6.14: Signal probe electron p_T distribution in data and MC, for $e^\pm\mu^\pm$ pairs (left, with probe electrons satisfying CFT cut and a reinforced tag muon selection) or $\ell^\pm e^\mp e^\mp$ pairs (right, with reinforced tag electron selection), as described in section 6.6.13. The yellow area indicates $t\bar{t}$ events in which the tag lepton is fake and the probe electron real, leading to a measurement bias.

6.6.13 Baseline-to-signal efficiency for fake electrons

Electron fake rates are measured with a similar methodology, but the $e^\pm e^\pm$ channel is unusable to the presence of a large charge-flip background. This is overcome by working with $e^\pm\mu^\pm$ pairs instead (with a tag muon), but mixing leptons of different flavours brings additional complications (one can not, for example, use the unbiased measurement employed to measure muon fakes rates at higher p_T , as there is no symmetry between the leptons). To improve confidence, measurements are performed in four different ways, which complement each other:

- straightforward tag-and-probe wit $e\mu$ pairs, with the same tag muon selection as in the previous section.
- same selection, but subtracting from the numerator the number of pairs

with one fake tag muon and one prompt probe electron, itself estimated from the number of observed $e\mu$ events with a muon failing signal requirements, scaled by a “muon fake rate” taken as the ratio between $\mu\mu$ pairs with two or only one signal muon. scaled by an efficiency correction factor $e\mu/\mu\mu$ taken from $t\bar{t}$ MC (only for pairs with one fake muon). This only works if the two muons satisfy the same kinematic requirements, therefore can be used only for measurements in wide or inclusive bins.

- selecting $\ell^\mp e^\pm e^\pm + \geq 1b$ events, with a Z veto on SFOS pairs. This selection entirely suppresses contributions from charge-flip, or events with fake muons. One of the electron, with standard signal requirements, is required to satisfy the same reinforced p_T and isolation requirements as for the muon measurement (section 6.6.12), and the measurement can be performed on the other electron.
- same selection, using the symmetry between the two same-sign electrons to measure the rates in an unbiased way, similarly to the muon case.

Events are acquired with the combination of single-muon (as in previous section) and $e\mu$ triggers.

Figure 6.14 shows the number of signal probe electrons selected in the $e\mu$ and ℓee channels. There are quite less events selected in the trilepton channel. Figure 6.15 shows the electron fake rate as a function of p_T or η in $t\bar{t}$ MC. The variations of the rates as function of the pseudorapidity are not very large, therefore we only perform measurements as function of p_T . The low p_T range is

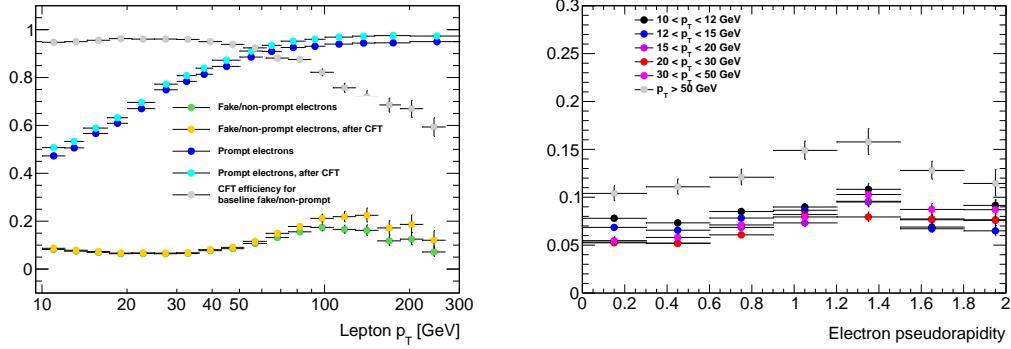


Figure 6.15: Electron fake rates in $t\bar{t}$ MC with an inclusive selection, as function of p_T (left, yellow/green markers = with/without CFT cut applied) or $|\eta|$ in different momentum ranges (right).

dominated by non-prompt electrons from heavy flavour decays, while beyond 30 GeV, electron fakes mostly come from conversions of photons produced inside jets, such as $\pi^0 \rightarrow \gamma\gamma$ decays (see appendix).

Based on the estimated levels of bias, and achievable statistical precision, of the different methods, we decided to measure electron fake rates with the tag-and-probe $e\mu$ selection up to 30 GeV, and by combining “unbiased” evaluations in both $e\mu$ and ℓee channels beyond. The measured rates are presented in Table 6.14, together with the associated statistical and background-subtraction uncertainties. The rates are here as well of the order of 10% up to 30 GeV, beyond which they increase to up to 25%.

Unlike muons, we do not so far apply MC-based correction factors for final states with $\geq 2 b$ -tagged jets. This is because there is less good agreement between the measured rates and the simulation; in particular the former take larger values in the medium- p_T range.

Systematic uncertainties

Similarly to the muon case, systematic uncertainties are assigned to cover for difference in the rates in the measurement regions in the signal regions, that would be due to different sources of fake leptons, or different kinematic properties of these sources.

Details can be consulted in appendix ???. Unlike muons, there is much less of a dependency to H_T . The dominant source of potential differences is therefore the origin of the fake electron (see Fig. 6.16); for $p_T < 20\text{GeV}$, non-prompt electrons from HF hadron decays dominate with some certainty, which is confirmed by the good agreement between MC fake rates and those measured in data. In that range, we assign a 30% uncertainty on the fake rates (inflated to 50% for final states with $\geq 2b$ -tagged jets). Beyond, the rates measured in data are larger than those predicted by the simulation, and would for example be consistent with a larger amount of electrons from photon conversions than predicted. In that range, we therefore assign a 50% uncertainty, which covers any arbitrary variation of the relative contributions of each source.

Finally, Figure 6.17 shows the variation of the fake rate in $t\bar{t}$ MC as function of the number of b -tagged jets in the event. Unsurprisingly, the rates are very similar for $0b$ and $\geq 1b$ final states, justifying the use of the fake rates measured in this section (i.e. in a $\geq 1b$ region) to predict fake electron background in all signal regions. We do not assign extra uncertainties for final states with $\geq 2b$ -jets for $p_T > 20\text{ GeV}$: if, indeed, the relative contribution of HF decays is smaller than

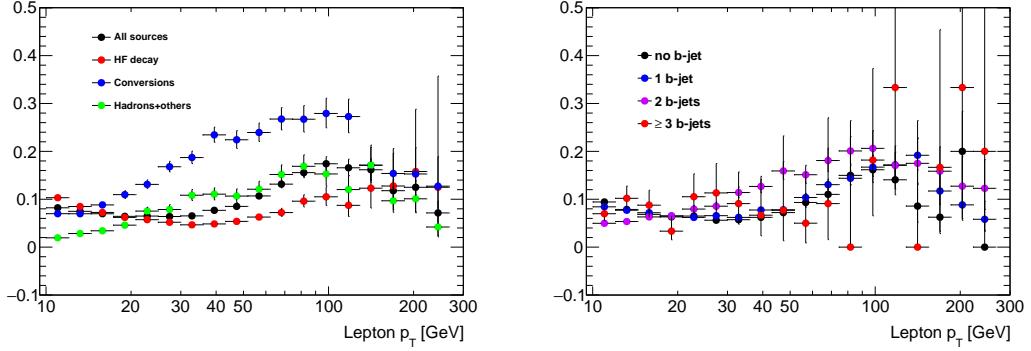


Figure 6.16: Electron fake rates in $t\bar{t}$ MC with an inclusive selection, as function of p_T and split according to the source of the fake electron (left). The relative contributions of each source (for signal electrons) are indicated on the right-hand-side.

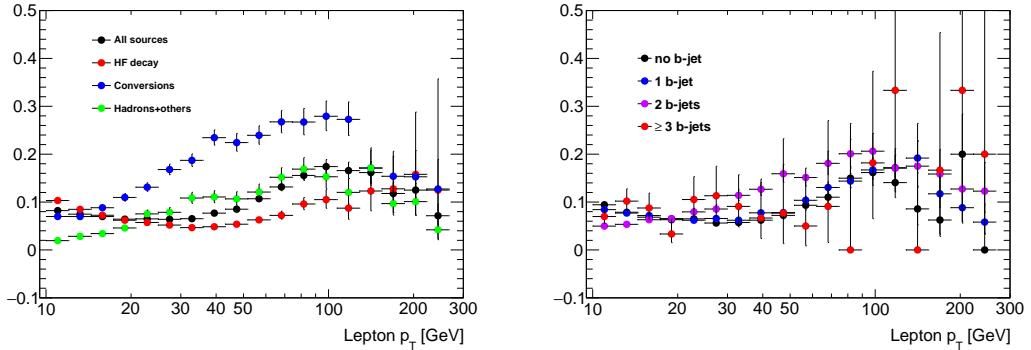


Figure 6.17: Electron fake rates in $t\bar{t}$ MC with an inclusive selection, as function of p_T and split according to the number of b -tagged jets in the event.

expected in that range, then there should be less intrinsic difference between 0/1 b and $\geq 2b$ final states, and those are already within the existing uncertainties.

6.6.14 Baseline-to-signal efficiency for real leptons

Baseline-to-signal efficiency for real leptons is measured in a high purity data sample of opposite-sign same-flavor leptons with the standard Z tag-and-probe method.

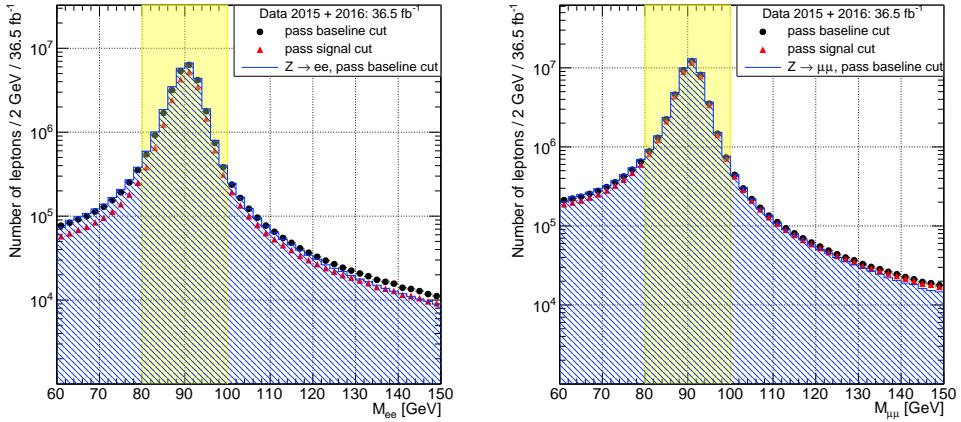


Figure 6.18: Invariant mass of opposite-sign same-flavor electrons (left) and muons (right), after the tag selection, where the probe satisfies the baseline requirements or the signal requirements.

Events are selected by a single lepton trigger, `e24_lhmedium_iloose_L1EM20VH` or `e26_lhtight_nod0_ivarloose` for electrons, `mu20_iloose_L1MU15` or `mu26_ivarmedium` for muons, respectively in 2015 or 2016 data. The tag lepton, required to have triggered the event recording, also satisfies signal requirements and verifies $p_T > 25\text{GeV}$. The probe lepton used for the efficiency measurement satisfies baseline requirements. All possible tag-and-probe combinations are considered in an event (including permutation of the tag and probe leptons), as long as the invariant mass of the pair is comprised between 80 and 100 GeV. Figure 6.18 illustrates this event selection.

A non-negligible background contamination in the electron channel affects measurements below $p_T = 20\text{GeV}$. This contamination is taken into account in the measurement using a background template method inspired by the one

used by the e/γ CP group to measure reconstruction, identification, and isolation efficiencies [?]. This template is built from the tag-and-probe invariant mass distribution for baseline-level probe electrons that fail both tighter PID cuts (`mediumLH`) and isolation requirements, smoothed by assuming an exponential shape whose parameters are determined by a fit in the interval $60 < m_{ee} < 120\text{GeV}$ excluding the $80 < m_{ee} < 100\text{GeV}$ region. The background template is then normalized to the main tag-and-probe distribution in the background-dominated tail $120 < m_{ee} < 150\text{GeV}$. More details and validation of the method will be provided in the appendix ???. The estimated level of background goes up to 4%, reached for probe electrons with $p_T < 15\text{GeV}$ and $|\eta| < 0.8$.

The efficiency is measured as a function of p_T and η , and the results are presented in Fig 6.19 for electrons and muons. The background subtraction is applied on the electron channel only. The following systematic uncertainties are assigned to the measured efficiencies:

- Background contamination: 27 variations of the tag-and-probe method are considered to assess the electron measurement systematics. Three m_{ee} windows and 9 variations of the background subtraction methods are considered. The largest contribution to the systematics arises from the m_{ee} window variation. This is expected as the proportion of electrons affected by bremsstrahlung depends on m_{ee} . The resulting relative systematics vary from 6% \sim 12% in the $10 < p_T < 15\text{GeV}$ region, 3% to 6% in the $15 < p_T < 20\text{GeV}$ region, 1% to 3% in the $20 < p_T < 40\text{GeV}$ region, and less

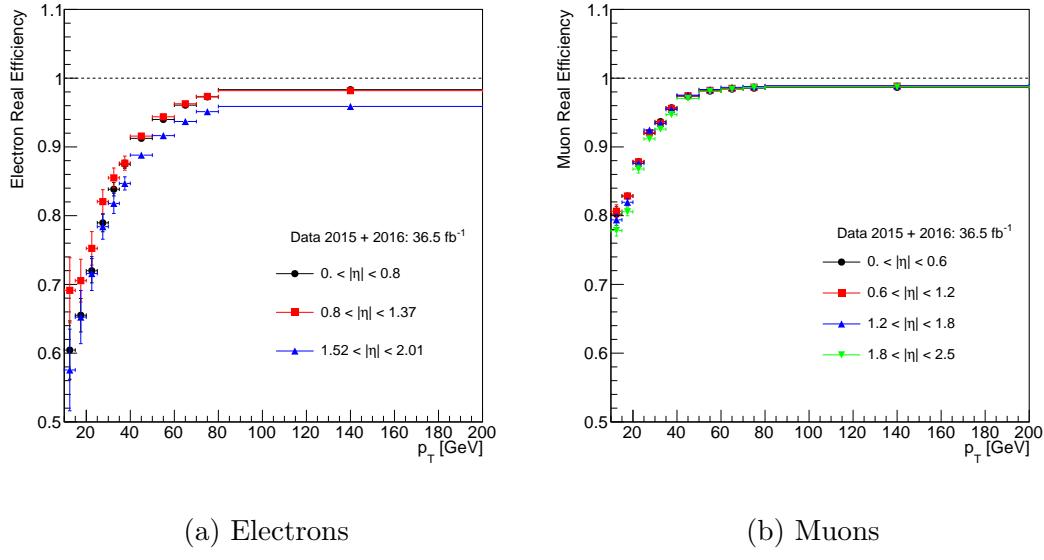


Figure 6.19: Baseline-to-signal efficiencies as a function of p_T and $|\eta|$ for real electrons (left) and muons (right), measured in 2015+2016 data. The $|\eta|$ binning used in the electron case corresponds to the geometry of the electromagnetic calorimeter. For muons a homogeneous $|\eta|$ binning is considered. The error bars corresponds to the quadratic sum of the statistical and tag-and-probe measurement systematic uncertainties.

than 1% for $p_T > 40$ GeV. The systematic uncertainties associated to the muon efficiencies measurement vary from 1% to 1.3% in the $10 < p_T < 15$ GeV region and less than 1% for $p_T > 15$ GeV.

- Trigger: a systematic uncertainty accounting for a potential bias at trigger level is considered as last year.
- Extrapolation to busy environments: efficiencies are typically lower in such environments due to the proximity of jets and leptons; an uncertainty is

assigned by comparing efficiencies in simulated $Z \rightarrow \ell\ell$ and $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$ events, for $\Delta m(\tilde{g}, \tilde{\chi}_1^0) > 1\text{TeV}$ which represents an extreme case of final states with highly boosted top quarks. The uncertainty, taken as the difference in efficiencies, is parametrized as a function of p_T and ΔR (the angular distance between the lepton and the closest jet).

The resulting systematic uncertainties are summarized in Table 6.15 and Table 6.16.

6.6.15 MC Template Method

We describe in this section a simulation-based method that provides an alternative estimate of the data-driven backgrounds described in Sections ?? and ??.

The MC-based method relies on kinematic distributions from MC simulations to extrapolate background predictions from control regions with low E_T^{miss} , m_{eff} , and jet multiplicity to the signal regions (where E_T^{miss} , m_{eff} , or jet multiplicity are required to be high). The control regions are used to reweight MC simulation of $t\bar{t}$ and $V + \text{jets}$ events to match the observed data by extracting correction factors which depend on the identified origin of the fake or non-prompt lepton. The main assumption of the method is that the MC simulations describe the kinematic distributions correctly and predict accurately the rate of fake leptons up to a global factor (for each type of fake lepton) independent of the event kinematics

and the process type. This assumption makes the method a suitable cross-check of the matrix method, that assumes that the lepton fake rates are the same in control and signal regions regardless of the selection requirements. The other assumption the MC template method makes is that the fake rates are uncorrelated in events with multiple fake leptons which is expected to be negligible.

Six non-overlapping control regions are defined by the presence of b -jets and by the flavors of the same sign lepton pair in the event:

- CR0b: events without b -jets in ee , $e\mu$, and $\mu\mu$ channels.
- CR1b: events with at least one b -jet in ee , $e\mu$, and $\mu\mu$ channels.

All the selected events contain two or more same-sign signal leptons and E_T^{miss} $> 40 \text{ GeV}$ and 2 or more jets. Events satisfying the signal regions requirements are excluded from the control regions. The purpose of the E_T^{miss} requirement is to remove multi-jet events that have two or more “fake” leptons and tend to have low E_T^{miss} .

The next step is to classify events into separate categories depending on the lepton origin. The three main categories are prompt isolated leptons, charge flipped electrons, and “fake” leptons which consist of non-prompt leptons coming from hadron decays and hadrons misidentified as leptons. The fakes are further separated by the lepton flavor (e or μ) and by the jet flavor producing the fake lepton. We separate the leptons that are coming from b -hadron decays, labelled as heavy-flavor (HF), from the rest of the fakes, labelled as light-flavor (LF), and derive a fake rate correction for each category. The purpose of this separation is

to make the b -tagging requirement orthogonal to the fake rate correction. The classification is done based on the parent particle from the generator event record using the type and origin of the lepton provided by the MCTruthClassifier.

In total we have five categories (charge flip, EL HF, MU HF, EL LF, MU LF) that we derive correction factors for using a simultaneous fit to data in six control regions (CR0b and CR1b for ee , $e\mu$, and $\mu\mu$ channels). The fit uses a likelihood function defined as the product of the Poisson probabilities describing the observed events in the binned distributions from the expected number of events rescaled by the five correction factors which are left free to float in the fit. These correction factors are applied to the MC predictions in the signal regions to obtain an estimation of the fake and charge flip backgrounds.

The six distributions are chosen for variables that provide the best separation between processes with prompt leptons and processes with fake leptons and charge flip and are shown before and after the fit in Figures 6.20-6.22 and Figures 6.21-6.23, respectively.

The minimization of the negative log likelihood using the MINUIT package leads to the correction factor shown in Tables 6.17 and 6.18. The tables represent the correction factors obtained from the fit upon using two different parton showers, Pythia and Sherpa for the processes that lead to non-prompt leptons and charge flips. The goal of varying the parton shower is to access the dependence of the fake and charge flip estimates on the choice of the parton shower. An additional systematic uncertainty is added to account for this dependency. The uncertainties in the correction factors themselves correspond to how much the

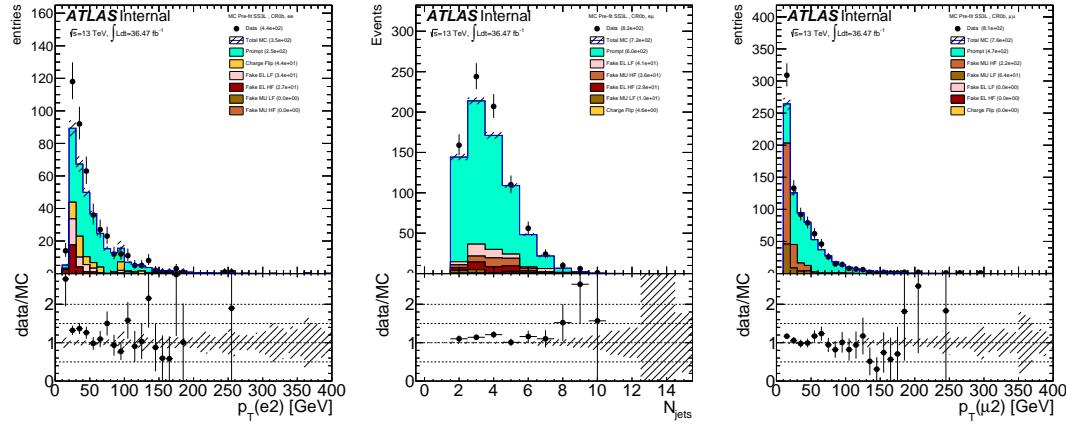


Figure 6.20: Pre-fit distributions for ee channel (left), for $e\mu$ channel (middle), and for $\mu\mu$ channel (right) from CR0b that were used in the fit to extract the fake rate and charge flip corrections. The generator used in these plots is Powheg. The hashed band represents the sum of systematic uncertainties on the predictions.

parameter needs to be varied for a one standard deviation change in the likelihood function. This uncertainty takes into account the limited number of simulated events and is included as a systematic uncertainty on the expected number of background events.

The MC template method is validated by looking at the Data/MC agreement near the signal regions where we invert one of the cuts while keeping the last bin corresponding to the SR blinded. These distributions can be found in Appendix ??.

After applying the correction factors of Table 6.17 to obtain the nominal fake estimates with the samples shown in Table ??, an additional uncertainty on the rate of fake leptons from $t\bar{t}$ and $V+jets$ in the signal region is assigned by repeating the background evaluation procedure using samples with Sherpa shown in Tables ??-?? and with the correction factors from Table 6.18. In practice, only

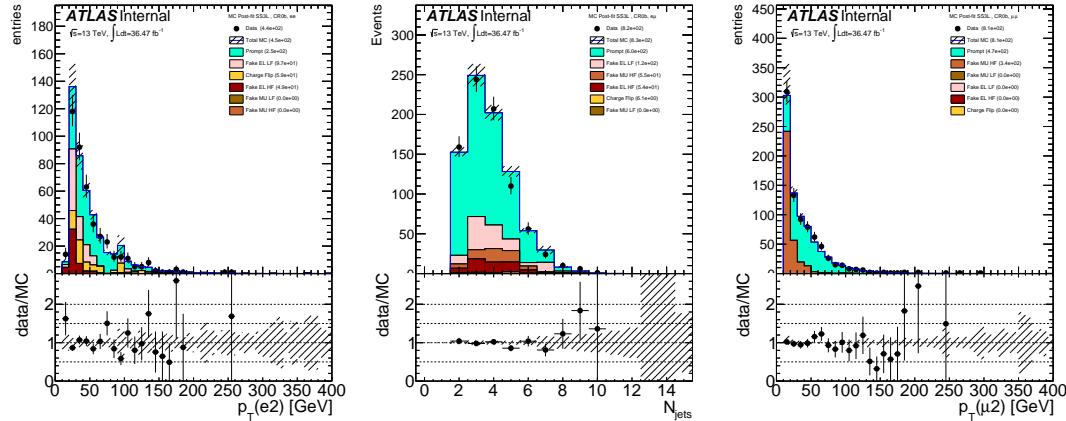


Figure 6.21: Post-fit distributions for ee channel (left), for $e\mu$ channel (middle), and for $\mu\mu$ channel (right) from CR0b that were used in the fit to extract the fake rate and charge flip corrections. The generator used in these plots is Powheg. The hashed band represents the sum of systematic uncertainties on the predictions.

$t\bar{t}$ contributes to the SRs and the final yields with systematic uncertainties from fit uncertainty, theory uncertainties on $t\bar{t}$, comparison of different showers (Pythia 6 and Sherpa) is shown in Table 6.19. This table also shows a global correction factor derived by taking the ratio of the weighted $t\bar{t}$ to raw MC $t\bar{t}$ with a global uncertainty that includes all systematic uncertainties used to obtain the final estimate.

6.6.16 Background processes with charge-flipped electrons

The lepton charge mis-measurement commonly referred to as “charge flip”, is an experimental background strongly associated to analyses relying on same-sign leptons final states. In those events, the electric charge of one of the two leptons forming an opposite-sign (OS) pair, coming from an abundant SM process ($pp \rightarrow Z$,

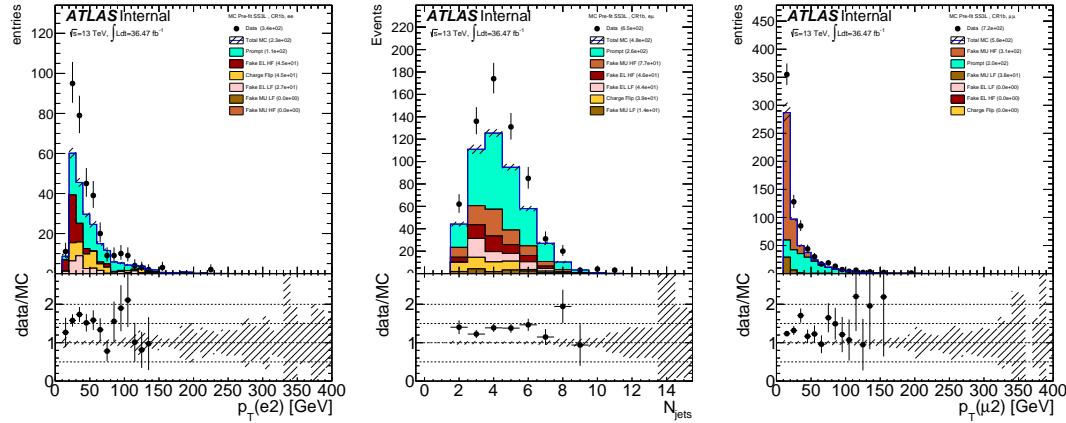


Figure 6.22: Pre-fit distributions for ee channel (left), for $e\mu$ channel (middle), and for $\mu\mu$ channel (right) from CR1b that were used in the fit to extract the fake rate and charge flip corrections. The generator used in these plots is Powheg. The hashed band represents the sum of systematic uncertainties on the predictions.

$t\bar{t}$, $W^+W^- \dots$), is mis-identified leading to a much rarer SS pair event. In most cases, the source of such a misidentification is the creation of additional close-by tracks $e^\pm \rightarrow e^\pm \gamma \rightarrow e^\pm e^\pm e^\mp$ via bremmstrahlung of the original electron when interacting with the material of the inner tracker. If one of the secondary electron tracks is subsequently preferred to the original track in the reconstruction of the electron candidate, the charge assigned to the electron might be incorrect, leading to a charge-flip event. Errors on the track charge assignment itself may occur as well, but they are much rarer. In the case of muons, charge-flip is essentially negligible due to the much smaller interaction cross-section with matter, and the requirement of identical charges to be measured for the inner tracker and muon spectrometer tracks.

We rely on a purely data-driven method (same in the previous version of

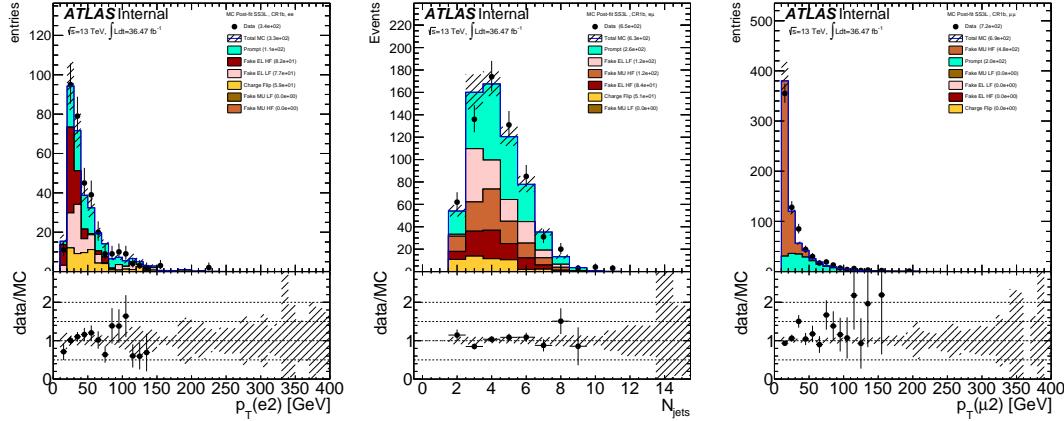


Figure 6.23: Post-fit distributions for ee channel (left), for $e\mu$ channel (middle),

and for $\mu\mu$ channel (right) from CR1b that were used in the fit to extract the fake rate and charge flip corrections. The generator used in these plots is Powheg. The hashed band represents the sum of systematic uncertainties on the predictions.

the analysis [?]) to estimate event yields for the electron charge-flip background.

Assuming one knows the electron charge flip rates $\xi(\eta, p_T)$, a simple way to predict these yields is to select events with pairs of opposite-sign leptons in data and assign them a weight:

$$w_{\text{flip}} = \xi_1(1 - \xi_2) + (1 - \xi_1)\xi_2 \quad (6.5)$$

where $\xi_{(i)} = 0$ for muons.

The advantages of this method are a good statistical precision since the charge flip rate is quite small, and the absence of dependency to the simulation and related uncertainties. Obviously, it requires a precise measurement of the rates, which is described in this section. An inconvenience of this approach is that the reconstructed momentum for charge-flipped electrons tends to be negatively biased

(too low by a few GeV), since such important bremmstrahlung topologies represent only a very small fraction of the cases used to tune electron energy calibration. Simply re-weighting electrons from opposite-sign lepton pairs therefore does not predict correctly the charge-flip background shape for variables very sensitive to the electron momentum, for example the m_{ee} lineshape. Further, we neglect this bias.

For the nominal (tight) estimate of the charge-flip background contributions, only events with exactly two OS signal electrons are considered. Corrections in the fake lepton estimate however require estimating as well charge-flip contributions for selections involving baseline electrons failing signal requirements (see section ??); for that reason, the charge-flip (loose) rate is measured for these two categories of electrons.

6.6.17 Methodology

Charge-flip rates are measured in data relying on a clean $Z \rightarrow ee$ sample ($75 < m_{ee} < 100\text{GeV}$), in which the rates can be determined from the relative proportions of OS and SS electron pairs. Figure 6.24 illustrates this event selection. The rates are measured as function of η and p_T , to follow their dependency to the distribution of material in the detector, the bremmstrahlung emission rate, and the track curvature. Because of this binned measurements, and that the two electrons in a given pair generally have different kinematic properties, it has been found that the most efficient and less biased use of the available statistics is obtained by simultaneously extracting the rates in all bins via the maximization

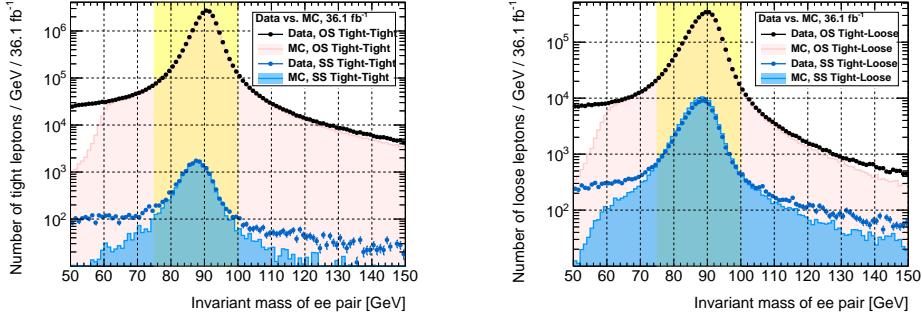


Figure 6.24: Invariant mass of opposite- and same-sign electron pairs, when both electrons satisfy signal requirements (left) or one of them fails them (right). Drell-Yan MC samples are not included, thus the drop in the MC distributions (light magenta filled area).

of the likelihood function describing the Poisson-expected yields of SS pairs:

$$L(\{N_{\varpi}^{\text{SS,obs}}\} | \{\xi(\eta, p_T)\}) = \prod_{\varpi} \mathcal{P}(N_{\varpi}^{\text{SS,obs}} | w_{\text{flip}}(\xi(\eta_1, p_{T,1}), \xi(\eta_2, p_{T,2})) \times N_{\varpi}^{\text{OS+SS,obs}}) \quad (6.6)$$

with $\varpi = (\eta_1, p_{T,1}, \eta_2, p_{T,2})$ indexing bins, where (arbitrarily) $p_{T,1} > p_{T,2}$; the expression of w_{flip} is given by (6.5). Statistical uncertainties on the extracted charge-flip rates are obtained (in a standard way) from the likelihood's numerically-computed Hessian matrix.

In the nominal charge-flip measurement, the two electrons are required to satisfy signal requirements. To measure charge-flip rates for baseline electrons failing signal (noted $\bar{\xi}$ below), we use instead pairs with only one signal electron; this provides larger statistics than applying (6.6) to electron pairs where both fail the signal cuts. However, the expression of the likelihood has to be adapted

due to the induced asymmetry between the two electrons forming the pair:

$$L(\{N_{\varpi}^{\text{SS,obs}}\} | \{\xi(\eta_1, p_{\text{T},1})\}, \{\bar{\xi}(\eta_2, p_{\text{T},2})\}) = \prod_{\varpi} \mathcal{P}(N_{\varpi}^{\text{SS,obs}} | w_{\text{flip}}(\xi(\eta_1, p_{\text{T},1}), \bar{\xi}(\eta_2, p_{\text{T},2})) \times N_{\varpi}^{\text{OS+SS,obs}}) \quad (6.7)$$

where this time $(\eta_1, p_{\text{T},1})$ corresponds to the signal electron. Using the same η and p_{T} binnings for both measurements, the number of free variables in the maximization of (6.7) – as well as the number of terms in the product forming L – is twice larger than for the nominal case (6.6). In fact, a by-product of the maximization of (6.7) is another determination of the charge-flip rates for signal electrons, although with a more limited precision than obtained in the nominal measurement (6.6); one can however as a cross-check verify that the two sets are compatible. A simultaneous maximization of the product of (6.6) and (6.7) (which rely on completely orthogonal sets of events) was studied for the 2015 analysis [?] as it could potentially improve the precision of the measurements, but it was found to be a bit less stable than the independent maximizations, and is therefore not used.

Background subtraction is performed through a simple linear extrapolation of the invariant mass distribution sidebands; it matters mostly for low p_{T} in the nominal measurement, and for the additional measurement with baseline electrons failing signal requirements, where the level of background is larger.

More details about the validation of the method are provided in Ref. [?] (section 7.2).

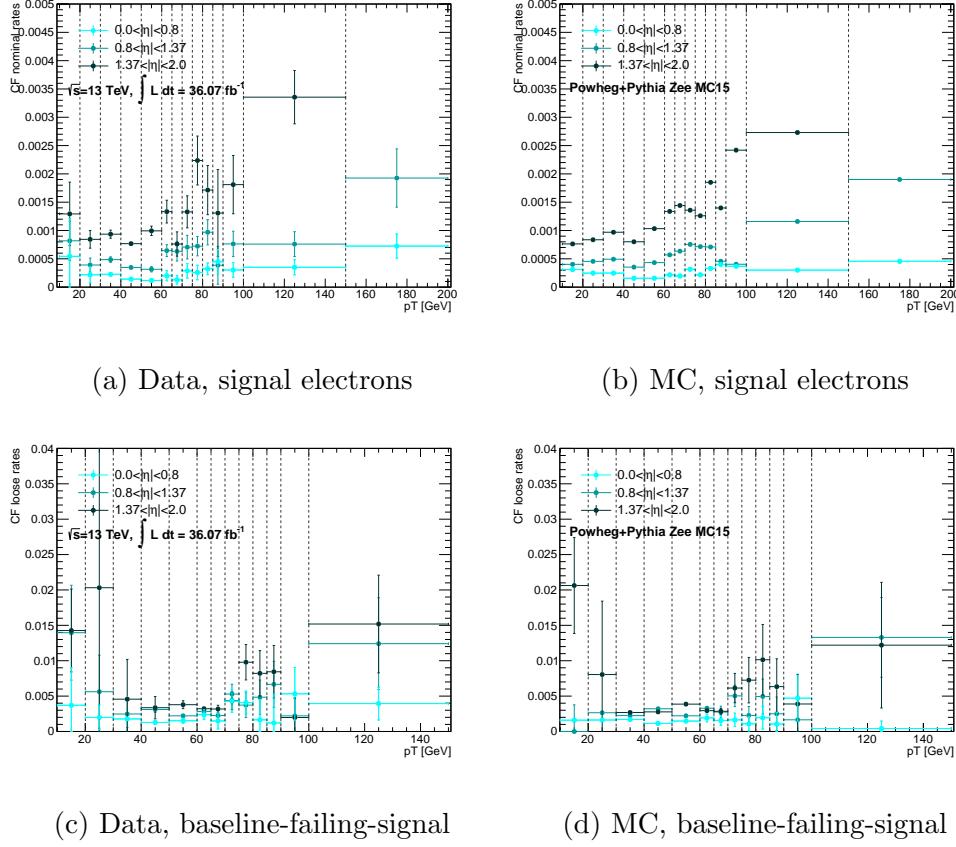


Figure 6.25: Charge-flip rate as measured in data (left) and MC (right). Only the statistical uncertainty is displayed. The last p_T bin is inclusive.

6.6.18 Measured rates

The charge-flip rates measured in data and MC are shown on Figure 6.25. In data, the nominal rates (Fig. 6.25a) go up to $\sim 0.1\%$ in the barrel region ($|\eta| < 1.37$), while it increases up to $\sim 0.2\%$ in the end-cap region ($|\eta| > 1.37$). For baseline electrons failing signal requirements (Fig. 6.25c), the rates are in general greater than the nominal ones in every bin, as expected. The charge-flip rates for these electrons go up to $\sim 0.5\%$ in the barrel region and up 1% in the end-cap region.

Compared to the rates used in the previous version of the analysis [?], the central values are much lower now. After applying the `ElectronChargeIDSelector` tool, the charge flip rates are strongly reduced for both signal and baseline-failing-signal electrons (up to a factor 20 in some bins). Below 30 GeV, the statistics is very low for the loose measurement; however, these results are used only to measure the electron fake rate and, as illustrated in Figure 6.14, in this p_T interval the charge flip background is negligible.

The charge-flip rates in MC (Figs. 6.25b, 6.25d) are obtained by applying the same methodology as in data. Generally, the rates are not very far from data, validating the use of MC to predict charge-flip background in several of the optimization studies presented in this document. In addition, a closure test is performed on $t\bar{t}$ MC, checking that weighted OS events can reproduce the distribution of SS charge-flip events (identified by truth-matching). A good overall agreement is found, largely within the assigned uncertainties; detailed plots are included in appendix ??.

6.6.19 Systematic uncertainties

The main uncertainties on the measured charge-flip rates come from the presence of background and the way it is estimated. To assess them, variations of the selection and background estimation are considered:

- 1) $75 < m_{ee} < 100 \text{GeV}$, no background subtraction;
- 2) $75 < m_{ee} < 100 \text{GeV}$, sidebands of 20 GeV;

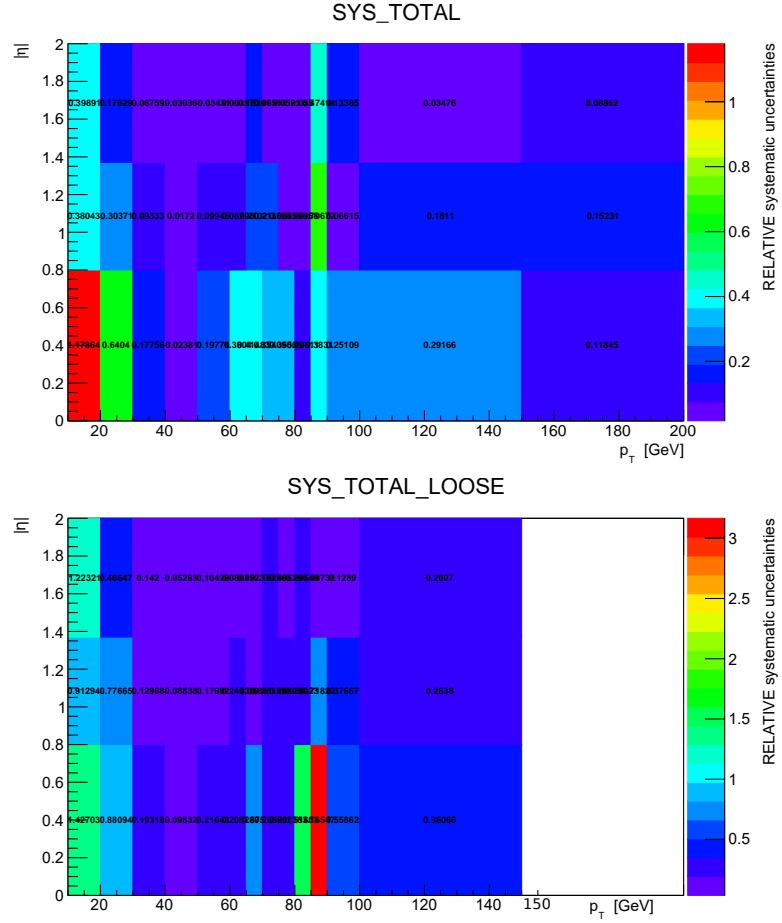


Figure 6.26: Total systematic uncertainties on the charge-flip rates for electrons satisfying the signal requirements (upper plot), and for baseline electrons failing the signal requirements (lower plot).

- 3) $75 < m_{ee} < 100 \text{ GeV}$, sidebands of 25 GeV(nominal measurement);
- 4) $75 < m_{ee} < 100 \text{ GeV}$, sidebands of 30 GeV;
- 5) $80 < m_{ee} < 100 \text{ GeV}$, sidebands of 20 GeV.

The effect of applying the background subtraction itself is evaluated by comparing configurations 1 and 3. The impact of the width of the m_{ee} chosen for the measurement is by comparing configurations 3 and 5, while the sideband width

effects are evaluated by comparing configuration 3 and 2, or 3 and 4. The largest deviation in each bin is taken as the systematic uncertainty on the charge-flip rate.

Those variations are shown in Appendix ?? (Figs ??-??) for every bin. The resulting systematic uncertainties on the charge-flip rates are presented in Fig. 6.26.

For the signal electrons charge-flip rates the systematic uncertainties vary in general between 2% and 20% (increasing up to > 50% in the region with $p_T < 10\text{GeV}$), whereas for baseline-failing-signal electrons they vary between 3% and 30% (increasing up to > 50% in the region with $p_T < 10\text{GeV}$). Part of these large values, at low p_T and in the [80,90] GeV p_T interval, can be explained by large statistical fluctuations between the different configurations.

6.6.20 Expected yields in signal regions

The expected background yields for processes with charge-flipped background, estimated with the method described in this section, are presented in Table 6.20 for the signal regions. They are compared for illustration to the MC prediction for simulated $t\bar{t}$ processes, and with the MC-template method (see section ??).

6.6.21 Expected yields in signal regions

The expected yield for processes with fake leptons, estimated with the method described in sections ?? and ??, are presented in Table 6.21 for the signal regions. They are compared for cross-check with the alternative ABCD prediction (section ??), and are all found to be consistent with each other.

The final numbers retained for the fake lepton background estimate (also shown in the tables) are taken as the weighted-average of the predictions from the matrix method and the MC template; the weights are based on the statistical component, and the systematic uncertainties are propagated assuming conservatively a full correlation between the two methods (although they are in fact largely independent!). The central value and statistical/systematic uncertainties are therefore:

$$(w\zeta_1 + (1-w)\zeta_2) \pm \sqrt{w^2 \left(\Delta\zeta_1^{(\text{stat})} \right)^2 + (1-w)^2 \left(\Delta\zeta_2^{(\text{stat})} \right)^2} \pm \left(w\Delta\zeta_1^{(\text{syst})} + (1-w)\Delta\zeta_2^{(\text{syst})} \right)$$

(6.8)

$$\text{with } w = \frac{\left(\Delta\zeta_2^{(\text{stat})} \right)^2}{\left(\Delta\zeta_1^{(\text{stat})} \right)^2 + \left(\Delta\zeta_2^{(\text{stat})} \right)^2}$$

When the estimated value is too small(below 0.15), the expected yield is set to 0.15 ± 0.15 , to cover for possibilities of an under-fluctuation of the number of baseline-not-signal leptons when applying the matrix method, as well as lack of statistics in the MC samples for the other method.

This upper bound is inflated if the original combined prediction $x \pm \Delta x$ is such that its plus-one-sigma variation exceeds the upper bound, that is $\delta = (x + \Delta x) > 0.30$; in that case the final retained number is $(\delta/2) \pm (\delta/2)$. There is no such instance in the current SR estimates, though, as can be seen in the table.

To check the validity and robustness of the FNP lepton estimate, the distributions of several discriminating variables in data are compared with the predicted

background after various requirements on the number of jets and b -jets. Examples of such distributions are shown in Figure 6.27, and illustrate that the data are described by the prediction within uncertainties. The apparent disagreement for m_{eff} above 1 TeV in Figure 6.27d is covered by the large theory uncertainty for the diboson background, which is not shown but amounts to about 30% for m_{eff} above 1 TeV.

6.7 Systematic uncertainties

6.8 Results and interpretation

Figure 6.30a shows the event yields for data and the expected background contributions in all signal regions. Detailed information about the yields can be found in Table 6.22. In all 19 SRs the number of observed data events is consistent with the expected background within the uncertainties. The contributions listed in the rare category are dominated by triboson, tWZ and $t\bar{t}WW$ production. The triboson processes generally dominate in the SRs with no b -jets, while tWZ and $t\bar{t}WW$ dominate in the SRs with one and two b -jets, respectively. Contributions from WH , ZH , tZ and $t\bar{t}t$ production never represent more than 20% of the rare background.

Figure 6.30b summarizes the contributions from the different sources of systematic uncertainty to the total SM background predictions in the signal regions. The uncertainties amount to 25–45% of the total background depending on the signal region, dominated by systematic uncertainties coming from the reducible

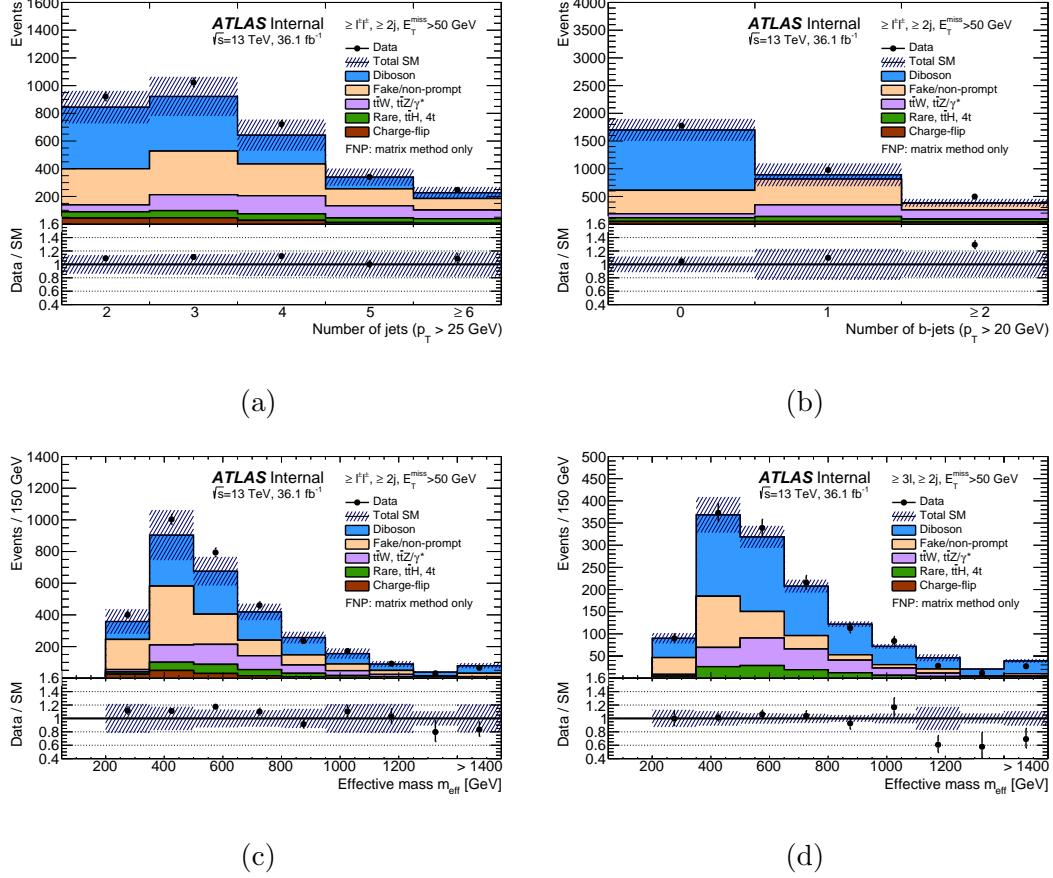


Figure 6.27: Distributions of (a) the number of jets, (b) the number of b -tagged jets and (c), (d) the effective mass. The distributions are made after requiring at least two jets ($p_T > 40\text{GeV}$) and $E_T^{\text{miss}} > 50\text{GeV}$, as well as at least two same-sign leptons ((a), (b), (c)) or three leptons (d). The uncertainty bands include the statistical uncertainties for the background prediction as well as the systematic uncertainties for fake- or non-prompt-lepton backgrounds (using the matrix method) and charge-flip electrons. Not included are theoretical uncertainties in the irreducible background contributions. The rare category is defined in the text.

background or the theory.

In the absence of any significant deviation from the SM predictions, upper

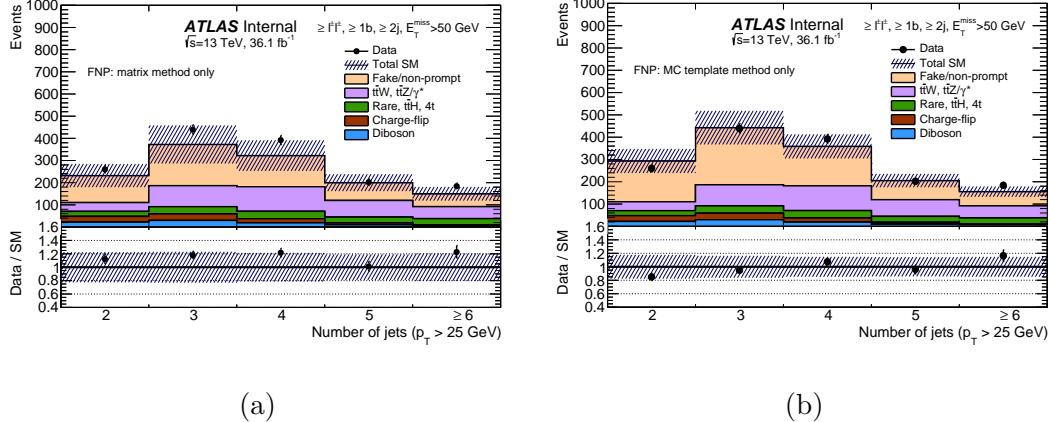


Figure 6.28: Distributions of the number of jets after requiring at least two jets

($p_T > 40\text{GeV}$) and $E_T^{\text{miss}} > 50\text{GeV}$, as well as at least two same-sign leptons.

The fake or non-prompt leptons backgrounds are estimated alternatively with the matrix method (a) or the MC template method (b). The uncertainty band includes the statistical uncertainties for the background prediction as well as the full systematic uncertainties for fake or non-prompt leptons backgrounds or charge-flip electrons. The rare category is defined in the text. In both figures, the last bin contains the overflow.

limits on possible BSM contributions to the signal regions are derived, as well as exclusion limits on the masses of SUSY particles in the benchmark scenarios of Figure 6.1. The HistFitter framework [?], which utilizes a profile-likelihood-ratio test [?], is used to establish 95% confidence intervals using the CL_s prescription [?].

The likelihood is built as the product of a Poisson probability density function describing the observed number of events in the signal region and, to constrain the nuisance parameters associated with the systematic uncertainties, Gaussian distributions whose widths correspond to the sizes of these uncertainties; Poisson

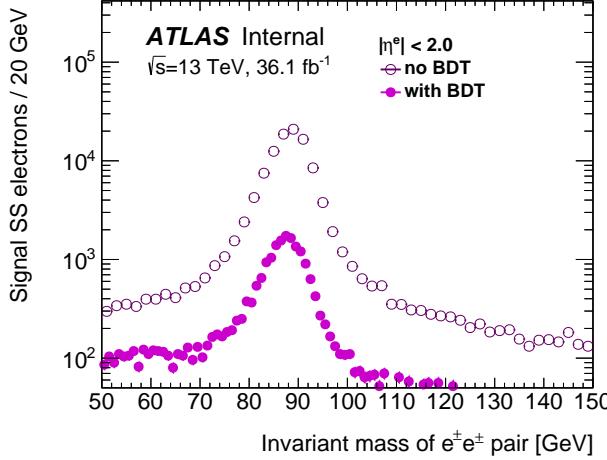


Figure 6.29: Invariant mass of the signal $e^\pm e^\pm$ pair distribution with (full markers) and without (open markers) charge-flip electron BDT selection applied.

distributions are used instead for MC simulation statistical uncertainties. Correlations of a given nuisance parameter between the backgrounds and the signal are taken into account when relevant. The hypothesis tests are performed for each of the signal regions independently.

Table 6.22 presents 95% confidence level (CL) observed (expected) model-independent upper limits on the number of BSM events, S_{obs}^{95} (S_{exp}^{95}), that may contribute to the signal regions. Normalizing these by the integrated luminosity L of the data sample, they can be interpreted as upper limits on the visible BSM cross-section (σ_{vis}), defined as $\sigma_{\text{vis}} = \sigma_{\text{prod}} \times A \times \epsilon = S_{\text{obs}}^{95}/L$, where σ_{prod} is the production cross-section, A the acceptance and ϵ the reconstruction efficiency. The largest deviation of the data from the background prediction corresponds to an excess of XX standard deviations in the YY SR.

Exclusion limits at 95% CL are also set on the masses of the superpartners

involved in the SUSY benchmark scenarios considered. Apart from the NUHM2 model, simplified models are used, corresponding to a single production mode and with 100% branching ratio to a specific decay chain, with the masses of the SUSY particles not involved in the process set to very high values. Figures 6.31 and 6.32 show the exclusion limits in all the models considered in Figure 6.1 and the NUHM2 model. The assumptions about the decay chain considered for the different SUSY particles are stated above each figure. For each region of the signal parameter space, the SR with the best expected sensitivity is chosen.

For the RPC models, the limits set are compared with the existing limits set by other ATLAS SUSY searches [?, ?]. For the models shown in Figure 6.31, the mass limits on gluinos and bottom squarks are up to 400 GeV higher than the previous limits, reflecting the improvements in the signal region definitions as well as the increase in integrated luminosity. Gluinos with masses up to 1.75 TeV are excluded in scenarios with a light $\tilde{\chi}_1^0$ in Figure 6.31a. This limit is extended to 1.87 TeV when $\tilde{\chi}_2^0$ and slepton masses are in between the gluino and the $\tilde{\chi}_1^0$ masses (Figure 6.31c). More generally, gluino masses below 1.57 TeV and bottom squarks with masses below 700 GeV are excluded in models with a massless LSP. The “compressed” regions, where SUSY particle masses are close to each other, are also better covered and LSP masses up to 1200 and 250 GeV are excluded in the gluino and bottom squark pair-production models, respectively. Of particular interest is the observed exclusion of models producing gluino pairs with an off-shell top quark in the decay (Figure ??), see Figure 6.31a. In this case, models are excluded for mass differences between the gluino and neutralino of 205 GeV (only 35 GeV

larger than the minimum mass difference for decays into two on-shell W bosons and two b -quarks) for a gluino mass below 0.9 TeV. The Rpc3LSS1b SR allows the exclusion of top squarks with masses below 700 GeV when the top squark decays to a top quark and a cascade of electroweakinos $\tilde{\chi}_2^0 \rightarrow \tilde{\chi}_1^\pm W^\mp \rightarrow W^* W^\mp \tilde{\chi}_1^0$ (see Figure 6.31e for the conditions on the sparticle masses).

Finally, in the NUHM2 model with low fine-tuning, values of the parameter $m_{1/2}$ below 615 GeV are excluded, corresponding to gluino masses below 1500 GeV (Figure 6.32).

6.9 Conclusion

A search for supersymmetry in events with two same-sign leptons or at least three leptons, multiple jets, b -jets and large E_T^{miss} and/or large m_{eff} is presented. The analysis is performed with proton–proton collision data at $\sqrt{s} = 13\text{TeV}$ collected in 2015 and 2016 with the ATLAS detector at the Large Hadron Collider corresponding to an integrated luminosity of 36.1 fb^{-1} . With no significant excess over the Standard Model prediction observed, results are interpreted in the framework of simplified models featuring gluino and squark production in R -parity-conserving and R -parity-violating scenarios. Lower limits on particle masses are derived at 95% confidence level. In the $\tilde{g}\tilde{g}$ simplified RPC models considered, gluinos with masses up to 1.85 TeV are excluded in scenarios with a light $\tilde{\chi}_1^0$. RPC models with bottom squark masses below 700 GeV are also excluded in a $\tilde{b}_1\tilde{b}_1^*$ simplified model with $\tilde{b}_1 \rightarrow tW^-\tilde{\chi}_1^0$ and a light $\tilde{\chi}_1^0$. In RPV

scenarios, masses of down squark-rights are probed up to $m_{\tilde{d}_R} \approx 500$ GeV. All models with gluino masses below 1.3 TeV are excluded, greatly extending the previous exclusion limits. Model-independent limits on the cross-section of a possible signal contribution to the signal regions are set.

6.10 Auxiliary

Table 6.8: Comparison of the event yields for the $t\bar{t}V$ background processes between aMC@NLO (default generator) and SHERPA in the SRs, as well as their relative difference.

SR	SHERPA	aMCATNLO	Relative diff.
Rpc2L0bH	0.25 ± 0.03	0.20 ± 0.05	25%
Rpc2L0bS	0.60 ± 0.06	0.82 ± 0.10	-26%
Rpc2L1bH	3.84 ± 0.14	3.86 ± 0.20	<1%
Rpc2L1bS	3.55 ± 0.13	3.94 ± 0.20	-9%
Rpc2L2bH	0.35 ± 0.04	0.41 ± 0.05	-14%
Rpc2L2bS	1.57 ± 0.08	1.57 ± 0.12	<1%
Rpc2Lsoft1b	1.01 ± 0.07	1.24 ± 0.11	-18%
Rpc2Lsoft2b	1.13 ± 0.07	1.15 ± 0.10	-1%
Rpc3L0bH	0.23 ± 0.02	0.18 ± 0.04	27%
Rpc3L0bS	0.90 ± 0.05	0.99 ± 0.09	-9%
Rpc3L1bH	1.54 ± 0.08	1.52 ± 0.11	1%
Rpc3L1bS	6.95 ± 0.16	7.02 ± 0.23	<1%
Rpc3LSS1b	0.00 ± 0.00	0.00 ± 0.00	-
Rpv2L0b	0.18 ± 0.03	0.14 ± 0.04	28%
Rpv2L1bH	0.51 ± 0.04	0.54 ± 0.07	-5%
Rpv2L1bM	1.78 ± 0.08	1.40 ± 0.12	27%
Rpv2L1bS	9.85 ± 0.19	9.91 ± 0.30	<1%
Rpv2L2bH	0.52 ± 0.04	$^{130}_{0.53 \pm 0.08}$	-1%
Rpv2L2bS	5.00 ± 0.14	4.89 ± 0.21	6%

Table 6.9: Number of raw MC events and its equivalent for 36.1 fb^{-1} with and without the correction as a function of the jet multiplicity described in Appendix ??.

Only the SRs where at least one MC event passes all the cuts are shown.

SR	Raw MC evts	Without N_{jet} correction	With N_{jet} correction
Rpc2L0bS	2	0.016-0.033	0.09-0.38
Rpc2L0bH	1	0.006-0.012	0.05-0.17
Rpv2L0b	1	0.006-0.012	0.05-0.17

Table 6.10: Expected yields for background processes with prompt leptons, in the SRs proposed in Section ??, for 36.1 fb^{-1} . Quoted uncertainties include statistical sources only. Rare category includes $t\bar{t}WW$, $t\bar{t}WZ$, $3t$, tZ , tWZ , WH , ZH and VVV , and detailed contributions of these processes can be found in Table 6.11.

	$t\bar{t}V$	VV	$t\bar{t}H$	$t\bar{t}t\bar{t}$	rare
Rpc2L0bH	0.20 ± 0.05	1.14 ± 0.23	0.08 ± 0.04	0.02 ± 0.01	0.17 ± 0.04
Rpc2L0bS	0.82 ± 0.10	3.13 ± 0.21	0.26 ± 0.05	0.01 ± 0.00	0.20 ± 0.04
Rpc2L1bH	3.86 ± 0.20	0.61 ± 0.06	1.01 ± 0.10	0.53 ± 0.03	0.97 ± 0.12
Rpc2L1bS	3.94 ± 0.20	0.48 ± 0.05	1.28 ± 0.10	0.33 ± 0.03	0.87 ± 0.12
Rpc2L2bH	0.41 ± 0.05	0.04 ± 0.01	0.10 ± 0.03	0.17 ± 0.02	0.14 ± 0.04
Rpc2L2bS	1.57 ± 0.12	0.10 ± 0.03	0.44 ± 0.06	0.25 ± 0.02	0.32 ± 0.05
Rpc2Lsoft1b	1.24 ± 0.11	0.14 ± 0.02	0.44 ± 0.06	0.09 ± 0.01	0.18 ± 0.04
Rpc2Lsoft2b	1.15 ± 0.10	0.05 ± 0.02	0.37 ± 0.06	0.20 ± 0.02	0.17 ± 0.03
Rpc3L0bH	0.18 ± 0.04	2.64 ± 0.12	0.03 ± 0.02	0.01 ± 0.00	0.29 ± 0.04
Rpc3L0bS	0.99 ± 0.09	8.95 ± 0.21	0.12 ± 0.04	0.02 ± 0.01	0.75 ± 0.07
Rpc3L1bH	1.52 ± 0.11	0.48 ± 0.05	0.25 ± 0.06	0.28 ± 0.03	0.87 ± 0.12
Rpc3L1bS	7.02 ± 0.23	1.44 ± 0.10	1.36 ± 0.10	0.69 ± 0.04	2.51 ± 0.22
Rpc3LSS1b	0.00 ± 0.00	0.00 ± 0.00	0.21 ± 0.04	0.00 ± 0.00	0.09 ± 0.01
Rpv2L0b	0.14 ± 0.04	0.52 ± 0.10	0.02 ± 0.02	0.01 ± 0.00	0.10 ± 0.04
Rpv2L1bH	0.54 ± 0.07	0.12 ± 0.02	0.07 ± 0.03	0.34 ± 0.02	0.29 ± 0.09
Rpv2L1bM	1.40 ± 0.12	0.41 ± 0.04	0.28 ± 0.06	0.53 ± 0.03	0.79 ± 0.12
Rpv2L1bS	9.91 ± 0.30	1.66 ± 0.08	1.93 ± 0.15	1.79 ± 0.06	2.40 ± 0.19
Rpv2L2bH	0.53 ± 0.08	0.04 ± 0.01	0.12 ± 0.03	0.48 ± 0.03	0.19 ± 0.05
Rpv2L2bS	4.80 ± 0.21	0.25 ± 0.03	0.87 ± 0.11	1.52 ± 0.05	1.11 ± 0.11

Table 6.11: Contributions of particular rare processes to the signal regions, relative to the summed contributions of all these processes.

	VVV	VH	$3t$	tZ	$t\bar{t}WW$	tWZ	$t\bar{t}WZ$
Rpc2L0bH	23%	0%	2%	3%	25%	43%	1%
Rpc2L0bS	50%	0%	3%	15%	14%	16%	0%
Rpc2L1bH	2%	0%	7%	4%	41%	41%	2%
Rpc2L1bS	2%	0%	6%	3%	34%	50%	2%
Rpc2L2bH	3%	0%	15%	4%	47%	27%	1%
Rpc2L2bS	2%	0%	13%	2%	42%	36%	2%
Rpc2Lsoft1b	3%	0%	9%	0%	76%	7%	2%
Rpc2Lsoft2b	2%	0%	17%	4%	54%	19%	2%
Rpc3L0bH	52%	0%	0%	3%	1%	40%	1%
Rpc3L0bS	50%	0%	0%	4%	2%	39%	1%
Rpc3L1bH	3%	0%	3%	3%	17%	70%	1%
Rpc3L1bS	2%	0%	3%	7%	18%	64%	2%
Rpc3LSS1b	25%	0%	0%	0%	0%	0%	74%
Rpv2L0b	10%	0%	3%	0%	66%	20%	0%
Rpv2L1bH	2%	0%	8%	0%	35%	52%	0%
Rpv2L1bM	1%	0%	6%	0%	19%	69%	1%
Rpv2L1bS	1%	0%	8%	4%	28%	54%	2%
Rpv2L2bH	0%	0%	17%	0%	35%	45%	0%
Rpv2L2bS	0%	0%	14%	4%	32%	45%	3%

Table 6.12: Muon fake rate measured in data and the associated statistical uncertainty. The systematic uncertainty originating from the subtraction of “backgrounds” with only prompt leptons is also displayed.

10 < p_T < 12 GeV		12 < p_T < 14	
$ \eta < 2.3$	$ \eta > 2.3$	$ \eta < 2.3$	$ \eta > 2.3$
$0.14 \pm 0.01 \pm 0.00$	$0.22 \pm 0.05 \pm 0.00$	$0.11 \pm 0.01 \pm 0.00$	$0.24 \pm 0.06 \pm 0.00$
14 < p_T < 17		17 < p_T < 20 GeV	
$ \eta < 2.3$	$ \eta > 2.3$	$ \eta < 2.3$	$ \eta > 2.3$
$0.12 \pm 0.01 \pm 0.00$	$0.09 \pm 0.05 \pm 0.00$	$0.09 \pm 0.01 \pm 0.00$	$0.21 \pm 0.07 \pm 0.00$
20 < p_T < 30	30 < p_T < 40	40 < p_T < 60	$p_T > 60$
$0.07 \pm 0.02 \pm 0.00$	$0.12 \pm 0.05 \pm 0.01$	$0.16 \pm 0.09 \pm 0.04$	$0.49 \pm 0.10 \pm 0.07$

Table 6.13: Additional systematic uncertainty on the muon fake rates, to address variations of the latter in different environments. The table also shows the correction factors and uncertainties applied to final states with ≥ 2 b -tagged jets.

p_T	< 14	14 – 20	20 – 30	30 – 40	40 – 60	> 60
$\Delta\zeta^{(\text{syst})}$	30%	30%	30%	50%	50% for $H_T < 600$ 70% for $600 < H_T < 1200$ 85% for $H_T > 1200$	
$\frac{\zeta_{\geq 2b}}{\zeta}$	1.2 ± 0.2	1.5 ± 0.5	1.7 ± 0.7	2.0 ± 1.0	1.5 ± 0.5	–

Table 6.14: Electron fake rate measured in data and the associated statistical uncertainty. The systematic uncertainty originating from the subtraction of “backgrounds” with only prompt leptons is also displayed.

$10 < p_T < 12$	$12 < p_T < 14$	$14 < p_T < 17$	$17 < p_T < 20$
$0.10 \pm 0.01 \pm 0.00$	$0.10 \pm 0.01 \pm 0.01$	$0.12 \pm 0.01 \pm 0.01$	$0.08 \pm 0.02 \pm 0.00$
$20 < p_T < 25$	$25 < p_T < 30$	$30 < p_T < 40$	$40 > p_T$
$0.07 \pm 0.02 \pm 0.01$	$0.11 \pm 0.03 \pm 0.01$	$0.20 \pm 0.07 \pm 0.03$	$0.25 \pm 0.10 \pm 0.05$

	Electrons			Muons			
	$0 < \eta < 0.8$	$0.8 < \eta < 1.37$	$1.52 < \eta < 2.01$	$0 < \eta < 0.6$	$0.6 < \eta < 1.2$	$1.2 < \eta < 1.8$	$1.8 < \eta < 2.5$
$10 \text{ GeV} < p_T < 15 \text{ GeV}$	0.047	0.063	0.089	0.014	0.010	0.008	0.011
$15 \text{ GeV} < p_T < 20 \text{ GeV}$	0.027	0.042	0.062	0.005	0.006	0.008	0.011
$20 \text{ GeV} < p_T < 25 \text{ GeV}$	0.018	0.031	0.041	0.003	0.006	0.010	0.010
$25 \text{ GeV} < p_T < 30 \text{ GeV}$	0.029	0.024	0.027	0.011	0.015	0.022	0.019
$30 \text{ GeV} < p_T < 35 \text{ GeV}$	0.023	0.021	0.023	0.007	0.009	0.014	0.011
$35 \text{ GeV} < p_T < 40 \text{ GeV}$	0.014	0.018	0.018	0.004	0.004	0.006	0.006
$40 \text{ GeV} < p_T < 50 \text{ GeV}$	0.007	0.010	0.010	0.002	0.001	0.002	0.001
$50 \text{ GeV} < p_T < 60 \text{ GeV}$	0.008	0.010	0.010	0.001	0.001	0.001	0.001
$60 \text{ GeV} < p_T < 70 \text{ GeV}$	0.007	0.010	0.010	0.001	0.001	0.001	0.002
$70 \text{ GeV} < p_T < 80 \text{ GeV}$	0.008	0.011	0.012	0.002	0.001	0.001	0.002
$80 \text{ GeV} < p_T < 120 \text{ GeV}$	0.010	0.010	0.011	0.004	0.002	0.002	0.002
$120 \text{ GeV} < p_T < 150 \text{ GeV}$	0.005	0.005	0.011	0.006	0.005	0.005	0.005
$150 \text{ GeV} < p_T < 200 \text{ GeV}$	0.005	0.002	0.019	0.005	0.005	0.005	0.006

Table 6.15: Systematic uncertainties on the measured real lepton efficiency, separating sources affecting the measurement itself (background subtraction, trigger bias, and different methods). See appendix ?? for more details about the sources of the systematic uncertainties.

electrons (busy environments)								
$\Delta R(e, jet)$	[0, 0.1]	[0.1, 0.15]	[0.15, 0.2]	[0.2, 0.3]	[0.3, 0.35]	[0.35, 0.4]	[0.4, 0.6]	[0.6, 4]
10 GeV < p_T < 20 GeV	-	-	-	-	-	-	-	25.31% 6.5%
20 GeV < p_T < 30 GeV	-	-	-	-	-	-	73.37%	10.21% 0.37%
30 GeV < p_T < 40 GeV	-	-	-	97.71%	48.22%	15.54%	7.29%	0.58%
40 GeV < p_T < 50 GeV	-	-	-	52.81%	22.80%	16.73%	7.68%	1.10%
50 GeV < p_T < 60 GeV	-	-	-	29.96%	21.49%	20.23%	6.99%	2.78%
60 GeV < p_T < 80 GeV	-	-	55.89%	24.31%	17.40%	24.77%	6.20%	2.87%
80 GeV < p_T < 150 GeV	-	57.52%	30.24%	16.45%	12.73%	20.92%	4.44%	2.73%
150 GeV < p_T < 200 GeV	88.54%	40.16%	19.34%	8.45%	14.66%	16.57%	2.57%	1.90%

muons (busy environments)								
$\Delta R(\mu, jet)$	[0, 0.1]	[0.1, 0.15]	[0.15, 0.2]	[0.2, 0.3]	[0.3, 0.35]	[0.35, 0.4]	[0.4, 0.6]	[0.6, 4]
10 GeV < p_T < 20 GeV	-	-	-	-	-	-	-	33.59% 5.18%
20 GeV < p_T < 30 GeV	-	-	-	-	-	-	82.34%	22.27% 3.39%
30 GeV < p_T < 40 GeV	-	-	-	98.54%	56.36%	31.89%	14.22%	2.24%
40 GeV < p_T < 50 GeV	-	-	-	53.10%	21.33%	13.90%	6.81%	1.45%
50 GeV < p_T < 60 GeV	-	-	-	24.98%	13.72%	9.62%	3.83%	0.79%
60 GeV < p_T < 80 GeV	-	-	44.41%	13.75%	6.14%	4.76%	2.04%	0.15%
80 GeV < p_T < 150 GeV	-	29.94%	7.14%	3.16%	1.30%	1.04%	0.07%	0.57%
150 GeV < p_T < 200 GeV	82.26%	4.14%	1.02%	0.17%	0.29%	0.62%	1.02%	1.13%

Table 6.16: Systematic uncertainties on the measured real lepton efficiency, due to the extrapolation to busy environments using $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$ events. See appendix ?? for more details.

Table 6.17: The fake-rate and charge flip corrections obtained after minimizing the likelihood function using Pythia. The uncertainty in the corrections takes into account the limited statistics of simulated events.

Category	Correction	Uncertainty
chFlip	1.49	0.58
HF EL	2.80	0.98
LF EL	2.89	0.88
HF MU	1.59	0.31
LF MU	1.00	1.34

Table 6.18: The fake-rate and charge flip corrections obtained after minimizing the likelihood function using Sherpa. The uncertainty in the corrections takes into account the limited statistics of simulated events.

Category	Correction	Uncertainty
chFlip	1.34	0.58
HF EL	2.40	0.85
LF EL	1.83	1.04
HF MU	1.17	0.16
LF MU	2.40	0.81

Table 6.19: Expected yields for background processes with fake leptons, in the signal regions with a global correction factor that represents the ratio of weighted $t\bar{t}$ to raw MC $t\bar{t}$ with a global uncertainty that includes: fit uncertainty, theory uncertainties on ttbar, comparison of different showers. The fraction of the systematic uncertainty from the comparison between two showers (Pythia and Sherpa) is also shown.

Region	MC Template method	Global correction	Shower systematic
Rpc2L0bH	$1.00 \pm 0.96 \pm 0.81$	2.80 ± 2.10	74%
Rpc2L0bS	$1.68 \pm 1.02 \pm 1.26$	2.89 ± 1.97	65%
Rpc2L1bH	$2.07 \pm 0.63 \pm 1.56$	1.22 ± 1.14	34%
Rpc2L1bS	$2.33 \pm 1.17 \pm 2.10$	1.83 ± 1.42	81%
Rpc2L2bH	< 0.5	0 ± 0	0%
Rpc2L2bS	$0.41 \pm 0.33 \pm 0.45$	1.47 ± 1.12	73%
Rpc2Lsoft1b	$2.48 \pm 1.32 \pm 1.86$	1.59 ± 1.31	68%
Rpc2Lsoft2b	$1.66 \pm 0.66 \pm 1.28$	1.72 ± 1.29	54%
Rpc3L0bH	< 0.5	0 ± 0	0%
Rpc3L0bS	$0.21 \pm 0.15 \pm 0.16$	2.90 ± 2.20	71%
Rpc3L1bH	$0.42 \pm 0.29 \pm 0.32$	1.59 ± 1.25	59%
Rpc3L1bS	$3.55 \pm 1.80 \pm 2.76$	1.76 ± 1.32	67%
Rpc3LSS1b	$0.90 \pm 0.14 \pm 0.69$	2.34 ± 1.44	56%
Rpv2L0b	$1.02 \pm 0.96 \pm 0.76$	2.80 ± 2.10	66%
Rpv2L1bH	$0.60 \pm 0.35 \pm 0.48$	1.32 ± 0.92	45%
Rpv2L1bM	$1.20 \pm 0.69 \pm 0.95$	1.59 ± 1.25	58%
Rpv2L1bS	$4.46 \pm 1.67 \pm 3.45$	138 1.33 ± 0.80	67%
Rpv2L2bH	< 0.5	0 ± 0	0%

Table 6.20: Expected yields for background processes with charge-flipped electrons, in the signal regions proposed in Section ??, shown for 36 fb^{-1} . Uncertainties include all statistical and systematic sources. Charge-flip processes do not contribute to signal regions which require ≥ 3 leptons. For illustration, charge-flip yields obtained with the MC-template method (section ??) are compared to the data-driven estimate.

Region	Weighted OS data	Template method
Rpc2L0bH	$0.01 \pm 0.00 \pm 0.00$	< 0.4
Rpc2L0bS	$0.05 \pm 0.01 \pm 0.01$	$00.02 \pm 00.02 \pm 00.00$
Rpc2L1bH	$0.25 \pm 0.03 \pm 0.04$	$00.21 \pm 00.32 \pm 00.16$
Rpc2L1bS	$0.25 \pm 0.02 \pm 0.04$	$00.35 \pm 00.37 \pm 00.26$
Rpc2L2bH	$0.02 \pm 0.01 \pm 0.00$	< 0.4
Rpc2L2bS	$0.10 \pm 0.01 \pm 0.02$	< 0.4
Rpc2Lsoft1b	$0.08 \pm 0.01 \pm 0.02$	< 0.4
Rpc2Lsoft2b	$0.08 \pm 0.01 \pm 0.02$	< 0.4
Rpc3LSS1b	$0.39 \pm 0.03 \pm 0.07$	$00.81 \pm 00.53 \pm 00.34$
Rpv2L0b	$0.03 \pm 0.02 \pm 0.00$	$00.22 \pm 00.06 \pm 00.09$
Rpv2L1bH	$0.02 \pm 0.01 \pm 0.00$	$00.02 \pm 00.01 \pm 00.01$
Rpv2L1bM	$0.10 \pm 0.01 \pm 0.02$	$00.10 \pm 00.06 \pm 00.04$
Rpv2L1bS	$0.74 \pm 0.04 \pm 0.11$	$01.30 \pm 00.66 \pm 00.54$
Rpv2L2bH	$0.03 \pm 0.01 \pm 0.01$	< 0.4
Rpv2L2bS	$0.46 \pm 0.03 \pm 0.07$	$01.55 \pm 01.07 \pm 01.25$

Table 6.21: Expected yields for background processes with fake leptons, in the signal regions proposed in Section ??, shown for 36 fb^{-1} . Uncertainties include all statistical and systematic sources for the nominal estimate (except ABCD, cf section ??).

Region	Matrix method	Template method	ABCD method	Retained estimate
Rpc2L0bH	$0.83 \pm 0.56 \pm 0.74$	$1.00 \pm 0.96 \pm 0.81$	$0.36 \pm 0.25 \pm 0.06$	$0.87 \pm 0.48 \pm 0.76$
Rpc2L0bS	$1.51 \pm 0.60 \pm 0.66$	$1.68 \pm 1.02 \pm 1.26$	$0.66 \pm 0.39 \pm 0.22$	$1.55 \pm 0.52 \pm 0.81$
Rpc2L1bH	$3.54 \pm 1.62 \pm 3.12$	$2.07 \pm 0.63 \pm 1.56$	$2.06 \pm 0.32 \pm 0.16$	$2.26 \pm 0.59 \pm 1.76$
Rpc2L1bS	$2.65 \pm 1.21 \pm 1.89$	$0.233 \pm 0.1.17 \pm 0.210$	$2.7 \pm 0.6 \pm 1.5$	$2.48 \pm 0.84 \pm 2.00$
Rpc2L2bH	$-0.11 \pm 0.11 \pm 0.18$	< 0.5	$0.10 \pm 0.05 \pm 0.03$	$0.15 \pm 0.15 \pm 0.00$
Rpc2L2bS	$1.31 \pm 1.07 \pm 1.65$	$0.41 \pm 0.33 \pm 0.45$	$1.1 \pm 0.2 \pm 0.4$	$0.49 \pm 0.32 \pm 0.55$
Rpc2Lsoft1b	$4.75 \pm 1.42 \pm 2.64$	$2.48 \pm 1.32 \pm 1.86$	$1.2 \pm 0.3 \pm 1.2$	$3.53 \pm 0.97 \pm 2.22$
Rpc2Lsoft2b	$1.91 \pm 1.18 \pm 1.63$	$1.66 \pm 0.66 \pm 1.28$	$0.72 \pm 0.19 \pm 0.26$	$1.72 \pm 0.58 \pm 1.36$
Rpc3L0bH	$-0.01 \pm 0.11 \pm 0.10$	< 0.5	< 0	$0.15 \pm 0.15 \pm 0.00$
Rpc3L0bS	$2.31 \pm 1.50 \pm 2.63$	$0.21 \pm 0.15 \pm 0.16$	< 0	$0.23 \pm 0.15 \pm 0.18$
Rpc3L1bH	$0.57 \pm 0.43 \pm 0.50$	$0.42 \pm 0.29 \pm 0.32$	< 0	$0.47 \pm 0.24 \pm 0.38$
Rpc3L1bS	$4.94 \pm 1.83 \pm 2.96$	$3.55 \pm 1.80 \pm 2.76$	$1.0 \pm 0.9 \pm 0.1$	$4.23 \pm 1.28 \pm 2.86$
Rpc3LSS1b	$-0.18 \pm 1.24 \pm 2.85$	$0.90 \pm 0.14 \pm 0.69$	$-$	$0.89 \pm 0.14 \pm 0.72$
Rpv2L0b	$0.14 \pm 0.22 \pm 0.27$	$1.02 \pm 0.96 \pm 0.76$	$0.13 \pm 0.12 \pm 0.04$	$0.18 \pm 0.21 \pm 0.29$
Rpv2L1bH	$-0.06 \pm 0.03 \pm 0.09$	$0.60 \pm 0.35 \pm 0.48$	$0.31 \pm 0.11 \pm 0.08$	$0.15 \pm 0.15 \pm 0.00$
Rpv2L1bM	$1.70 \pm 2.07 \pm 1.68$	$1.20 \pm 0.69 \pm 0.95$	$0.53 \pm 0.10 \pm 0.18$	$1.25 \pm 0.65 \pm 1.02$
Rpv2L1bS	$16.49 \pm 4.04 \pm 18.70$	$4.46 \pm 1.67 \pm 3.45$	$4.5 \pm 0.9 \pm 1.4$	$6.22 \pm 1.54 \pm 5.68$
Rpv2L2bH	$-0.04 \pm 0.02 \pm 0.04$	< 0.5	$0.26 \pm 0.08 \pm 0.07$	$0.15 \pm 0.15 \pm 0.00$
Rpv2L2bS	$9.67 \pm 3.29 \pm 9.04$	$7.24 \pm 2.36 \pm 5.43$	$2.7 \pm 0.5 \pm 1.0$	$8.07 \pm 1.92 \pm 6.66$

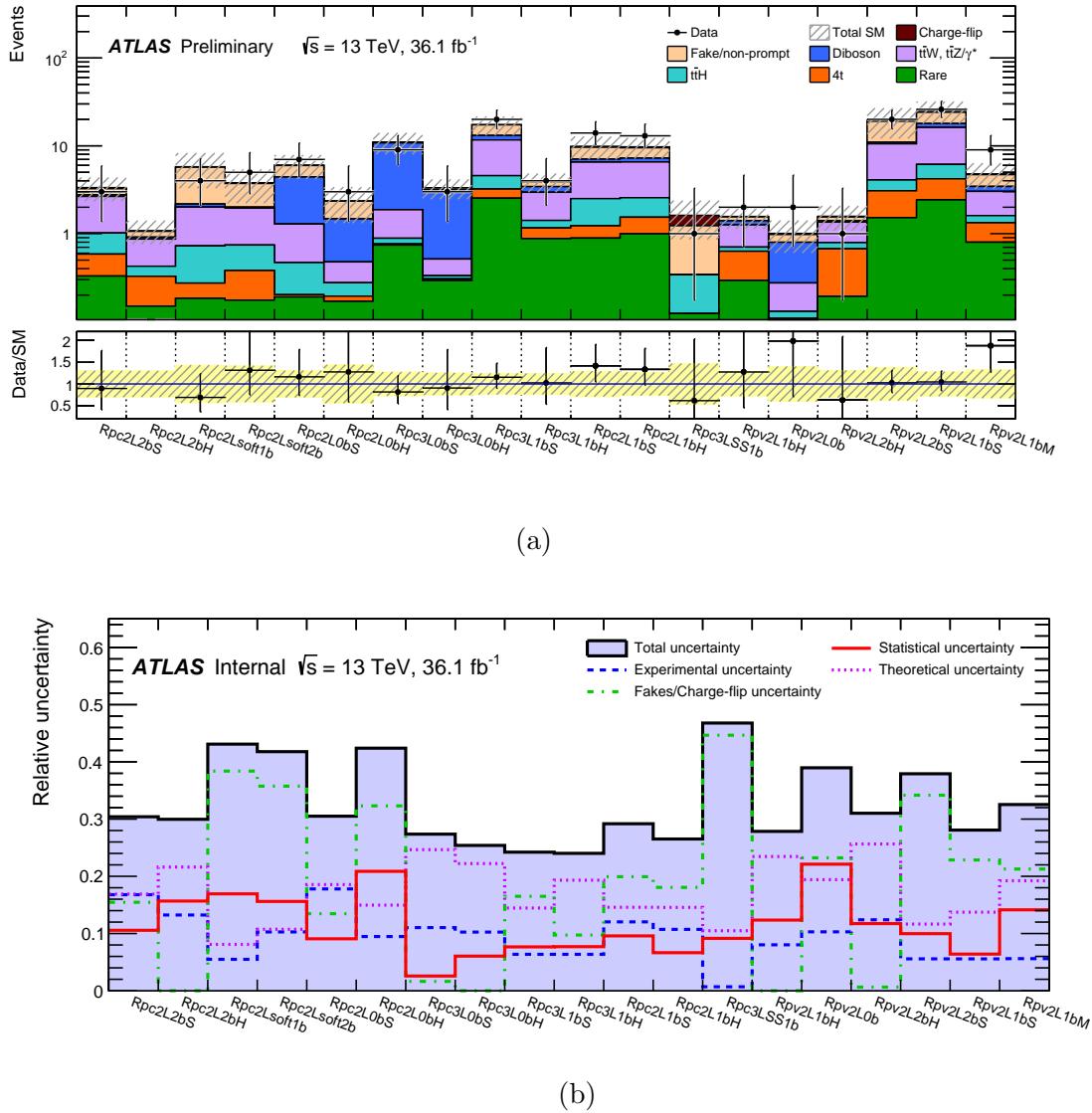
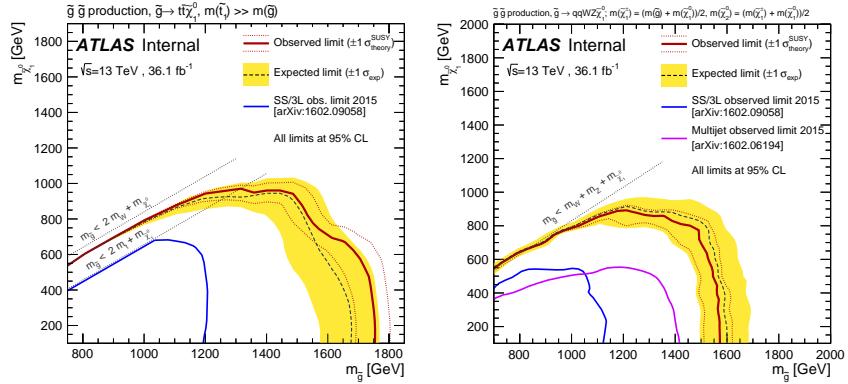


Figure 6.30: Comparison of (a) the observed and expected event yields in each signal region and (b) the relative uncertainties in the total background yield estimate. For the latter, “statistical uncertainty” corresponds to reducible and irreducible background statistical uncertainties. The background predictions correspond to those presented in Table 6.22 and the rare category is explained in the text.

Signal Region	Rpc2L2bS	Rpc2L2bH	Rpc2Lsoft1b	Rpc2Lsoft2b	Rpc2L0bS	Rpc2L0bH
$t\bar{t}W, t\bar{t}Z\gamma^*$	1.6 ± 0.4	0.44 ± 0.14	1.3 ± 0.4	1.21 ± 0.33	0.82 ± 0.31	0.20 ± 0.10
$t\bar{t}H$	0.43 ± 0.25	0.10 ± 0.06	0.45 ± 0.24	0.36 ± 0.21	0.27 ± 0.15	0.08 ± 0.07
$4t$	0.26 ± 0.13	0.18 ± 0.09	0.09 ± 0.05	0.21 ± 0.11	0.01 ± 0.01	0.02 ± 0.02
Diboson	0.10 ± 0.10	0.04 ± 0.02	0.17 ± 0.09	0.05 ± 0.03	3.1 ± 1.4	1.0 ± 0.5
Rare	0.33 ± 0.18	0.15 ± 0.09	0.18 ± 0.10	0.17 ± 0.10	0.19 ± 0.11	0.17 ± 0.10
Fake/non-prompt leptons	0.5 ± 0.6	0.15 ± 0.15	3.5 ± 2.4	1.7 ± 1.5	1.6 ± 1.0	0.9 ± 0.9
Charge-flip	0.10 ± 0.01	0.02 ± 0.01	0.08 ± 0.02	0.08 ± 0.02	0.05 ± 0.01	0.01 ± 0.01
Total Background	3.3 ± 1.0	1.08 ± 0.32	5.8 ± 2.5	3.8 ± 1.6	6.0 ± 1.8	2.4 ± 1.0
Observed	3	0	4	5	7	3
S_{obs}^{95}	5.5	3.6	6.3	7.7	8.3	6.1
S_{exp}^{95}	$5.6^{+2.2}_{-1.5}$	$3.9^{+1.4}_{-0.4}$	$7.1^{+2.5}_{-1.5}$	$6.2^{+2.6}_{-1.5}$	$7.5^{+2.6}_{-1.8}$	$5.3^{+2.1}_{-1.3}$
$\sigma_{\text{vis}} [\text{fb}]$	0.15	0.10	0.17	0.21	0.23	0.17
$p_0 (Z)$	0.71 (-)	0.91 (-)	0.69 (-)	0.30 (0.5σ)	0.36 (0.4σ)	0.35 (0.4σ)

Signal Region	Rpc3L0bS	Rpc3L0bH	Rpc3L1bS	Rpc3L1bH	Rpc2L1bS	Rpc2L1bH	Rpc3L2bS
$t\bar{t}W, t\bar{t}Z\gamma^*$	0.98 ± 0.25	0.18 ± 0.08	7.1 ± 1.1	1.54 ± 0.28	4.0 ± 1.0	4.0 ± 0.9	-
$t\bar{t}H$	0.12 ± 0.08	0.03 ± 0.02	1.4 ± 0.7	0.25 ± 0.14	1.3 ± 0.7	1.0 ± 0.6	0.22 ± 0.11
$4t$	0.02 ± 0.01	0.01 ± 0.01	0.7 ± 0.4	0.28 ± 0.15	0.34 ± 0.17	0.54 ± 0.28	-
Diboson	8.9 ± 2.9	2.6 ± 0.8	1.4 ± 0.5	0.48 ± 0.17	0.5 ± 0.3	0.7 ± 0.3	-
Rare	0.7 ± 0.4	0.29 ± 0.16	2.5 ± 1.3	0.9 ± 0.5	0.9 ± 0.5	1.0 ± 0.6	0.12 ± 0.06
Fake/non-prompt leptons	0.23 ± 0.23	0.15 ± 0.15	4.2 ± 3.1	0.5 ± 0.5	2.5 ± 2.2	2.3 ± 1.9	0.9 ± 0.8
Charge-flip	-	-	-	-	0.25 ± 0.04	0.25 ± 0.05	0.39 ± 0.30
Total Background	11.0 ± 3.0	3.3 ± 0.8	17 ± 4	3.9 ± 0.9	9.8 ± 2.9	9.8 ± 2.6	1.6 ± 1.0
Observed	9	3	20	4	14	13	10
S_{obs}^{95}	8.3	5.4	14.7	6.1	13.7	12.4	10.0
S_{exp}^{95}	$9.3^{+3.1}_{-2.3}$	$5.5^{+2.2}_{-1.5}$	$12.6^{+5.1}_{-3.4}$	$5.9^{+2.2}_{-1.8}$	$10.0^{+3.7}_{-2.6}$	$9.7^{+3.4}_{-2.6}$	$8.7^{+3.0}_{-2.5}$
$\sigma_{\text{vis}} [\text{fb}]$	0.23	0.15	0.41	0.17	0.38	0.34	0.30
$p_0 (Z)$	0.72 (-)	0.85 (-)	0.32 (0.5σ)	0.46 (0.1σ)	0.17 (1.0σ)	0.21 (0.8σ)	0.15 (0.7σ)

Table 6.22: Numbers of events observed in the signal regions compared with the expected backgrounds. The rare category is defined in the text. Background categories with yields shown as a “–” do not contribute to a given region (e.g.

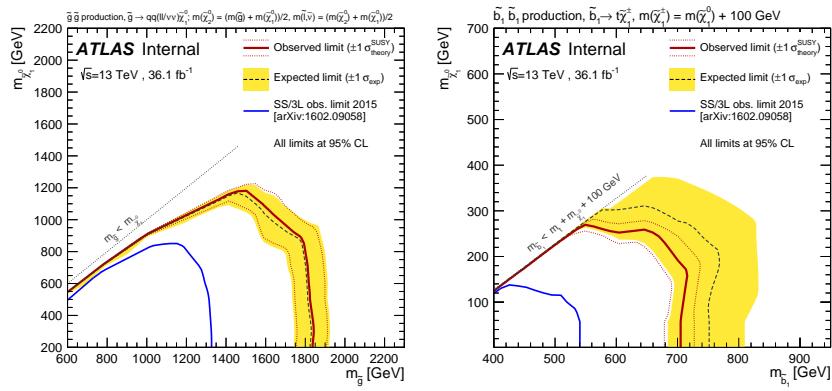


(a)

Rpc2L2bS/H,

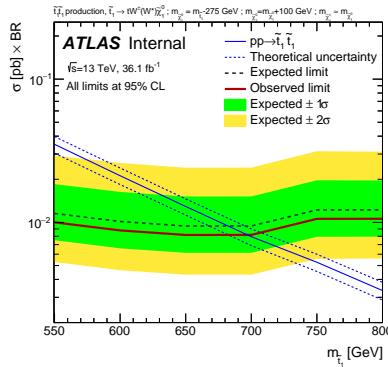
(b) Rpc2L0bS, Rpc2L0bH

Rpc2Lsoft1b/2b



(c) Rpc3L0bS, Rpc3L0bH

(d) Rpc2L1bS, Rpc2L1bH



(e) Rpc3LSS1b

Figure 6.31: Observed and expected exclusion limits on the \tilde{g} , \tilde{b}_1 , \tilde{t}_1 and $\tilde{\chi}^0$

masses in the context of RPC SUSY scenarios with simplified mass spectra.

The signal regions used to obtain the limits are specified in the subtitle of each scenario. All limits are computed at 95% CL. The dotted lines around the observed limit illustrate the change in the observed limit as the nominal signal

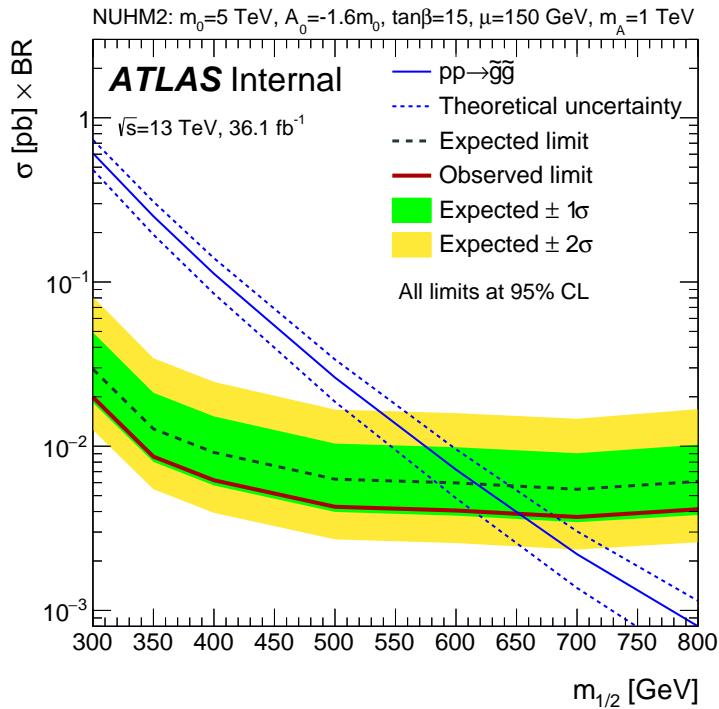


Figure 6.32: Observed and expected exclusion limits as a function of $m_{1/2}$ in the NUHM2 model [?, ?]. The signal region Rpc2L2bH is used to obtain the limits. The contours of the green (yellow) band around the expected limit are the $\pm 1\sigma$ ($\pm 2\sigma$) results, including all uncertainties. The limits are computed at 95% CL.

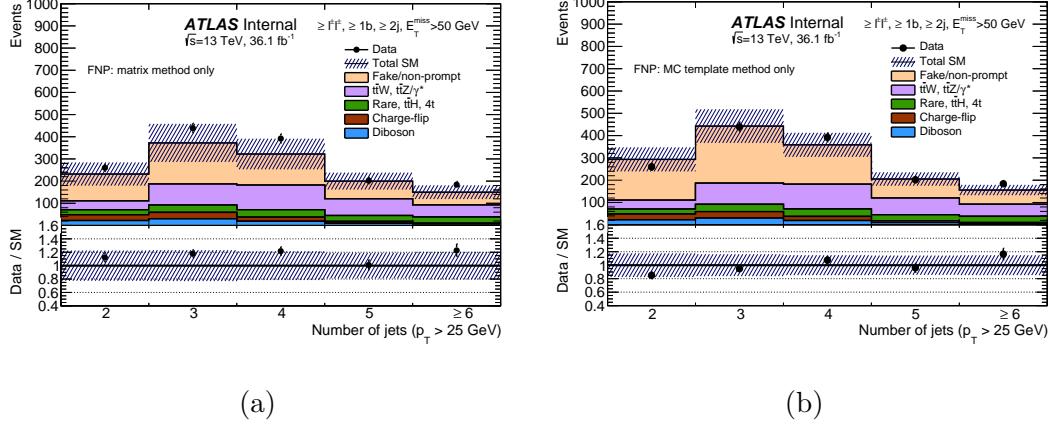


Figure 6.33: Distributions of the number of jets after requiring at least two jets ($p_T > 40\text{GeV}$) and $E_T^{\text{miss}} > 50\text{GeV}$, as well as at least two same-sign leptons.

The fake or non-prompt leptons backgrounds are estimated alternatively with the matrix method (a) or the MC template method (b). The uncertainty band includes the statistical uncertainties for the background prediction as well as the full systematic uncertainties for fake or non-prompt leptons backgrounds or charge-flip electrons. The rare category is defined in the text. In both figures, the last bin contains the overflow.

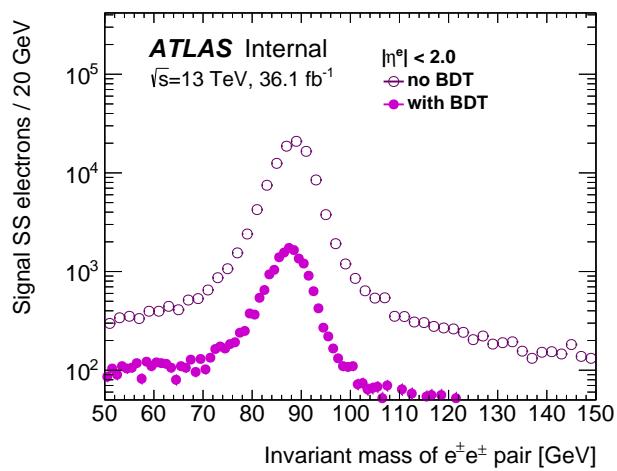


Figure 6.34: Invariant mass of the signal $e^\pm e^\pm$ pair distribution with (full markers) and without (open markers) charge-flip electron BDT selection applied.

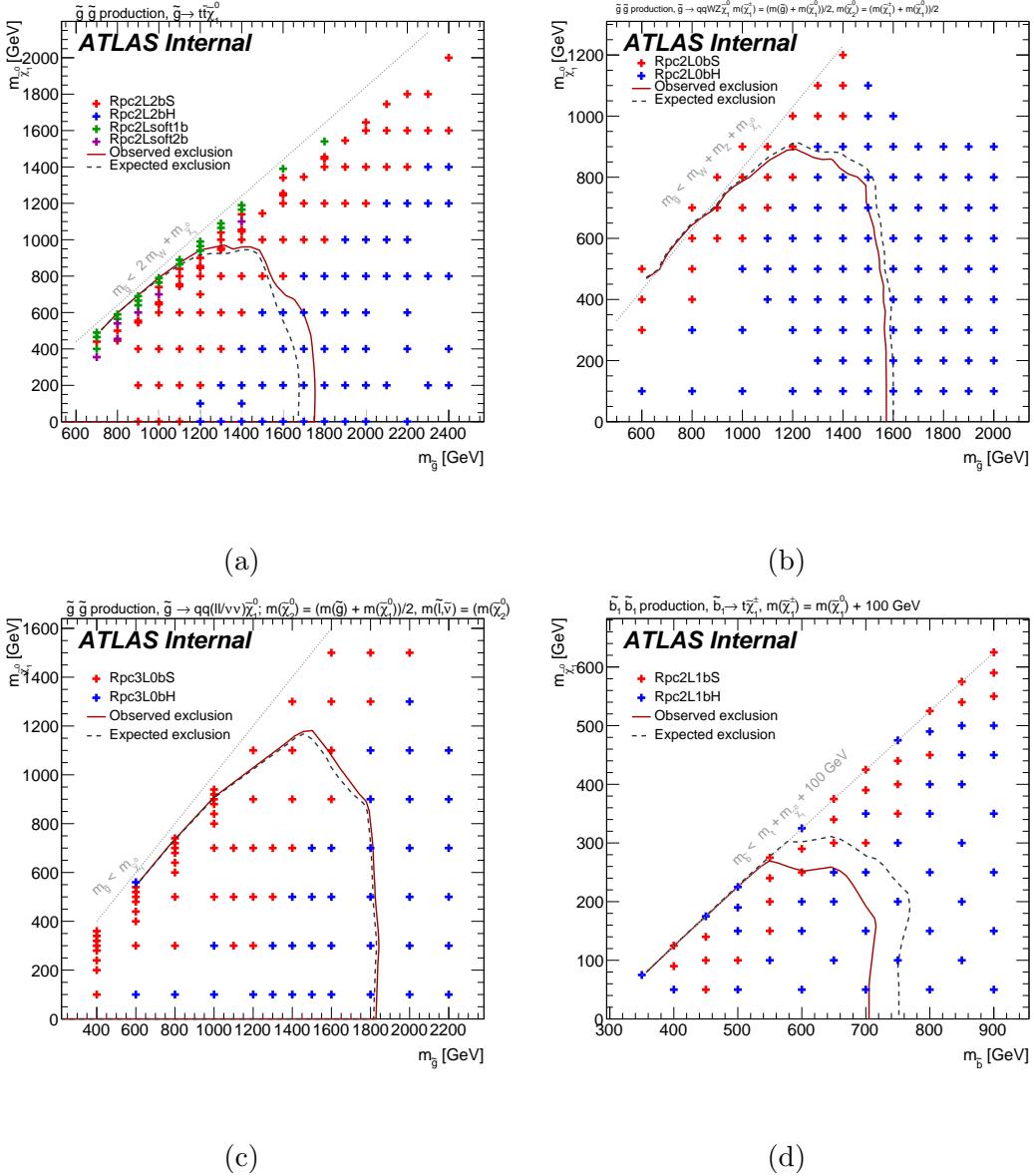


Figure 6.35: Illustration of the best expected signal region per signal grid point

for the $\tilde{g} \rightarrow q\bar{q}(\ell\ell/\nu\nu)\tilde{\chi}_1^0$ (Fig. ??), $\tilde{g} \rightarrow q\bar{q}'WZ\tilde{\chi}_1^0$ (Fig. ??), $\tilde{g} \rightarrow q\bar{q}'\ell/\nu\ell/\nu\tilde{\chi}_1^0$ (Fig. ??), $\tilde{b} \rightarrow tW\tilde{\chi}_1^0$ (Fig. ??) models. This mapping is used for the final combined exclusion limits.

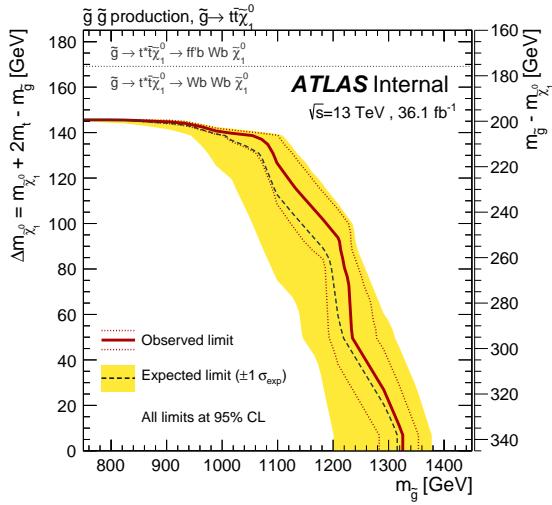


Figure 6.36: Observed and expected exclusion limits on the \tilde{g} and $\tilde{\chi}_1^0$ masses for the model in Figure ??, that is, in the mass-parameter space of Figure 6.31a where there is at least one top-quark off-shell decay. All limits are computed at 95% CL. The dotted lines around the observed limit illustrate the change in the observed limit as the nominal signal cross-section is scaled up and down by the theoretical uncertainty. The contours of the band around the expected limit are the $\pm 1\sigma$ results, including all uncertainties except the theoretical ones on the signal cross-section.

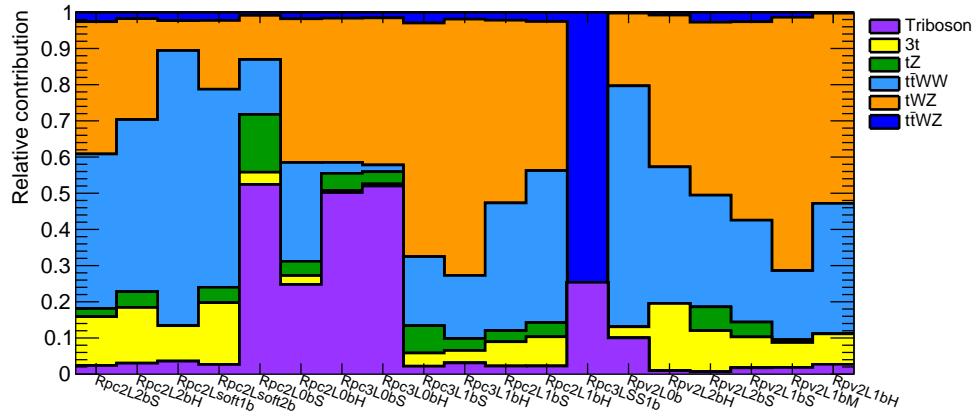


Figure 6.37: Relative contribution in each signal region from the processes in the category labelled as rare in the paper ($t\bar{t}WW$, $t\bar{t}WZ$, tZ , tWZ , $t\bar{t}\bar{t}$, WH , ZH and triboson production).

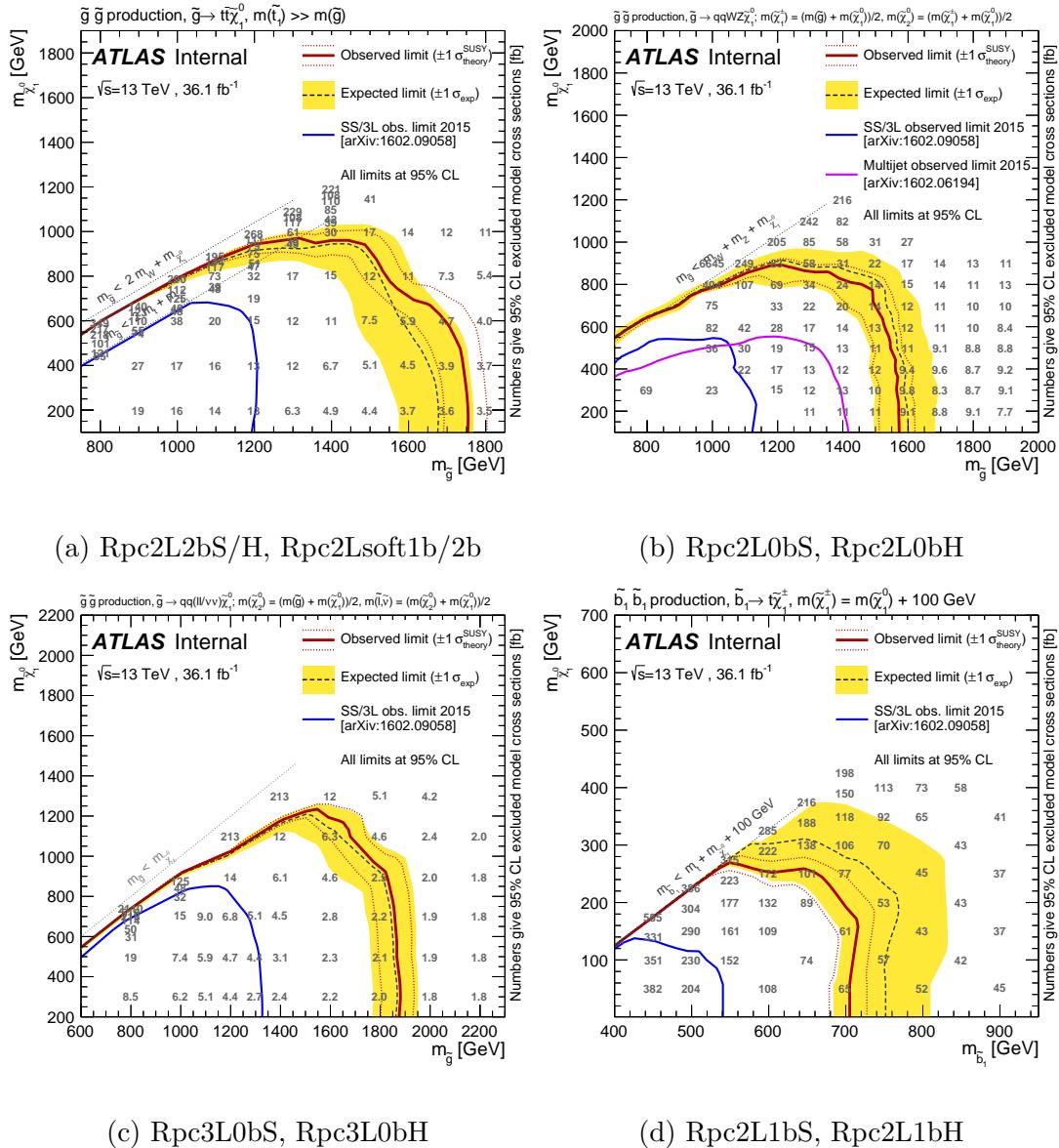


Figure 6.38: Observed and expected exclusion limits on the \tilde{g} , \tilde{b}_1 , \tilde{t}_1 and $\tilde{\chi}_1^0$ masses in the context of RPC SUSY scenarios with simplified mass spectra. The signal regions used to obtain the limits are specified in the subtitle of each scenario. All limits are computed at 95% CL. The dotted lines around the observed limit illustrate the change in the observed limit as the nominal signal cross-section is scaled up and down by the theoretical uncertainty. The contours of the band around the expected limit are the $\pm 1\sigma$ results, including all uncertainties except the theoretical ones on the signal cross-section. In Figures 6.38a–6.38d, the grey diagonal lines indicate the kinematic limit for the decays in each specified scenario.

Rpc2L2bS, $\tilde{g}\tilde{g}$ production, $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$	
$m_{\tilde{g}} = 1.5\text{TeV}$, $m_{\tilde{\chi}_1^0} = 800\text{GeV}$	
MC events generated	98000
Expected for 36.1 fb^{-1}	5.1×10^2
≥ 2 SS leptons ($p_T > 20\text{GeV}$)	19.96 ± 0.35
Trigger	19.17 ± 0.35
≥ 2 b -jets ($p_T > 20\text{GeV}$)	16.10 ± 0.32
≥ 6 jets ($p_T > 25\text{GeV}$)	13.11 ± 0.28
$E_T^{\text{miss}} > 200\text{GeV}$	10.17 ± 0.26
$m_{\text{eff}} > 0.6\text{TeV}$	10.17 ± 0.26
$E_T^{\text{miss}} > 0.25 \times m_{\text{eff}}$	5.94 ± 0.20

Table 6.23: Number of signal events at different stages of the Rpc2L2bS signal region selection. Only statistical uncertainties are shown.

Rpc2L2bH, $\tilde{g}\tilde{g}$ production, $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$	
$m_{\tilde{g}} = 1.7\text{TeV}$, $m_{\tilde{\chi}_1^0} = 200\text{GeV}$	
MC events generated	98000
Expected for 36.1 fb^{-1}	1.7×10^2
≥ 2 SS leptons ($p_T > 20\text{GeV}$)	7.32 ± 0.13
Trigger	7.19 ± 0.13
≥ 2 b -jets ($p_T > 20\text{GeV}$)	5.81 ± 0.11
≥ 6 jets ($p_T > 40\text{GeV}$)	4.92 ± 0.11
$m_{\text{eff}} > 1.8\text{TeV}$	3.93 ± 0.09
$E_T^{\text{miss}} > 0.15 \times m_{\text{eff}}$	3.12 ± 0.08

Table 6.24: Number of signal events at different stages of the Rpc2L2bH signal region selection. Only statistical uncertainties are shown.

Rpc2Lsoft1b, $\tilde{g}\tilde{g}$ production, $\tilde{g} \rightarrow tWb\tilde{\chi}_1^0$	
$m_{\tilde{g}} = 1.2\text{TeV}, m_{\tilde{\chi}_1^0} = 940\text{GeV}$	
MC events generated	50000
Expected for 36.1 fb^{-1}	3.1×10^3
≥ 2 SS leptons ($100 > p_T > 20, 10 \text{ GeV}$)	101.9 ± 2.7
Trigger	89.3 ± 2.5
≥ 1 b -jet ($p_T > 20\text{GeV}$)	75.1 ± 2.3
≥ 6 jets ($p_T > 25\text{GeV}$)	31.5 ± 1.5
$E_T^{\text{miss}} > 100\text{GeV}$	23.0 ± 1.3
$E_T^{\text{miss}} > 0.3 \times m_{\text{eff}}$	6.5 ± 0.7

Table 6.25: Number of signal events at different stages of the Rpc2Lsoft1b signal region selection. Only statistical uncertainties are shown.

Rpc2Lsoft2b, $\tilde{g}\tilde{g}$ production, $\tilde{g} \rightarrow tWb\tilde{\chi}_1^0$	
$m_{\tilde{g}} = 1.2\text{TeV}$, $m_{\tilde{\chi}_1^0} = 900\text{GeV}$	
MC events generated	50000
Expected for 36.1 fb^{-1}	3.1×10^3
≥ 2 SS leptons ($100 > p_T > 20, 10\text{ GeV}$)	91.8 ± 2.6
Trigger	79.7 ± 2.4
≥ 2 b -jets ($p_T > 20\text{GeV}$)	41.3 ± 1.7
≥ 6 jets ($p_T > 25\text{GeV}$)	21.4 ± 1.2
$E_T^{\text{miss}} > 200\text{GeV}$	8.7 ± 0.7
$m_{\text{eff}} > 0.6\text{TeV}$	8.7 ± 0.7
$E_T^{\text{miss}} > 0.25 \times m_{\text{eff}}$	6.7 ± 0.6

Table 6.26: Number of signal events at different stages of the Rpc2Lsoft2b signal region selection. Only statistical uncertainties are shown.

Rpc2L0bS, $\tilde{g}\tilde{g}$ production, $\tilde{g} \rightarrow q\bar{q}' WZ\tilde{\chi}_1^0$	
$m_{\tilde{g}} = 1.2\text{TeV}, (m_{\tilde{\chi}_1^\pm} - 150) = (m_{\tilde{\chi}_2^0} - 75) = m_{\tilde{\chi}_1^0} = 900\text{GeV}$	
MC events generated	19000
Expected for 36.1 fb^{-1}	3.1×10^3
≥ 2 SS leptons ($p_T > 20\text{GeV}$)	64 ± 4
Trigger	58.6 ± 3.3
no b -jet ($p_T > 20\text{GeV}$)	46.3 ± 3.0
≥ 6 jets ($p_T > 25\text{GeV}$)	26.6 ± 2.4
$E_T^{\text{miss}} > 150\text{GeV}$	16.3 ± 2.0
$E_T^{\text{miss}} > 0.25 \times m_{\text{eff}}$	9.0 ± 1.3

Table 6.27: Number of signal events at different stages of the Rpc2L0bS signal region selection. Only statistical uncertainties are shown.

Rpc2L0bH, $\tilde{g}\tilde{g}$ production, $\tilde{g} \rightarrow q\bar{q}' WZ\tilde{\chi}_1^0$	
$m_{\tilde{g}} = 1.6\text{TeV}$, $(m_{\tilde{\chi}_1^\pm} - 750) = (m_{\tilde{\chi}_2^0} - 375) = m_{\tilde{\chi}_1^0} = 100\text{GeV}$	
MC events generated	20000
Expected for 36.1 fb^{-1}	2.9×10^2
≥ 2 SS leptons ($p_T > 20\text{GeV}$)	12.8 ± 0.5
Trigger	12.5 ± 0.5
no b -jet ($p_T > 20\text{GeV}$)	8.5 ± 0.4
≥ 6 jets ($p_T > 40\text{GeV}$)	7.12 ± 0.35
$E_T^{\text{miss}} > 250\text{GeV}$	5.13 ± 0.29
$m_{\text{eff}} > 0.9\text{TeV}$	5.13 ± 0.29

Table 6.28: Number of signal events at different stages of the Rpc2L0bH signal region selection. Only statistical uncertainties are shown.

Rpc3L0bS, $\tilde{g}\tilde{g}$ production, $\tilde{g} \rightarrow q\bar{q}(\tilde{\ell}\ell/\tilde{\nu}\nu)$	
$m_{\tilde{g}} = 1.4\text{TeV}$, $(m_{\tilde{\chi}_2^0} - 150) = (m_{\tilde{\ell},\tilde{\nu}} - 75) = m_{\tilde{\chi}_1^0} = 1100\text{GeV}$	
MC events generated	20000
Expected for 36.1 fb^{-1}	9.1×10^2
≥ 3 leptons ($p_T > 20, 20, 10$ GeV)	76.9 ± 2.1
Trigger	76.0 ± 2.0
no b -jet ($p_T > 20\text{GeV}$)	67.5 ± 1.9
≥ 4 jets ($p_T > 40\text{GeV}$)	31.6 ± 1.3
$E_T^{\text{miss}} > 200\text{GeV}$	17.1 ± 1.0
$m_{\text{eff}} > 0.6\text{TeV}$	17.1 ± 1.0

Table 6.29: Number of signal events at different stages of the Rpc3L0bS signal region selection. Only statistical uncertainties are shown.

Rpc3L0bH, $\tilde{g}\tilde{g}$ production, $\tilde{g} \rightarrow q\bar{q}(\tilde{\ell}\ell/\tilde{\nu}\nu)$	
$m_{\tilde{g}} = 1.8\text{TeV}$, $(m_{\tilde{\chi}_2^0} - 850) = (m_{\tilde{\ell},\tilde{\nu}} - 375) = m_{\tilde{\chi}_1^0} = 100\text{GeV}$	
MC events generated	20000
Expected for 36.1 fb^{-1}	1.0×10^2
≥ 3 leptons ($p_T > 20, 20, 10 \text{ GeV}$)	9.98 ± 0.25
Trigger	9.94 ± 0.25
no b -jet ($p_T > 20\text{GeV}$)	8.44 ± 0.23
≥ 4 jets ($p_T > 40\text{GeV}$)	7.79 ± 0.22
$E_T^{\text{miss}} > 200\text{GeV}$	6.58 ± 0.21
$m_{\text{eff}} > 1.6\text{TeV}$	6.56 ± 0.21

Table 6.30: Number of signal events at different stages of the Rpc3L0bH signal region selection. Only statistical uncertainties are shown.

Rpc2L1bS, $\tilde{b}_1 \tilde{b}_1^*$ production, $\tilde{b}_1 \rightarrow t\tilde{\chi}_1^- \rightarrow tW^-\tilde{\chi}_1^0$	
$m_{\tilde{b}_1} = 600\text{GeV}$, $m_{\tilde{\chi}_1^\pm} = 350\text{GeV}$, $m_{\tilde{\chi}_1^0} = 250\text{GeV}$	
MC events generated	10000
Expected for 36.1 fb^{-1}	6.3×10^3
≥ 2 SS leptons ($p_T > 20\text{GeV}$)	221 ± 4
Trigger	201 ± 4
≥ 1 b -jet ($p_T > 20\text{GeV}$)	173 ± 4
≥ 6 jets ($p_T > 25\text{GeV}$)	66.3 ± 2.2
$E_T^{\text{miss}} > 150\text{GeV}$	36.5 ± 1.7
$m_{\text{eff}} > 0.6\text{TeV}$	36.1 ± 1.7
$E_T^{\text{miss}} > 0.25 \times m_{\text{eff}}$	15.1 ± 1.1

Table 6.31: Number of signal events at different stages of the Rpc2L1bS signal region selection. Only statistical uncertainties are shown.

Rpc2L1bH, $\tilde{b}_1 \tilde{b}_1^*$ production, $\tilde{b}_1 \rightarrow t \tilde{\chi}_1^-$	
$m_{\tilde{b}_1} = 750\text{GeV}$, $m_{\tilde{\chi}_1^\pm} = 200\text{GeV}$, $m_{\tilde{\chi}_1^0} = 100\text{GeV}$	
MC events generated	10000
Expected for 36.1 fb^{-1}	1.6×10^3
≥ 2 SS leptons ($p_T > 20\text{GeV}$)	71.1 ± 1.2
Trigger	66.4 ± 1.2
≥ 1 b -jet ($p_T > 20\text{GeV}$)	56.6 ± 1.1
≥ 6 jets ($p_T > 25\text{GeV}$)	27.7 ± 0.7
$E_T^{\text{miss}} > 250\text{GeV}$	12.5 ± 0.5
$E_T^{\text{miss}} > 0.2 \times m_{\text{eff}}$	9.5 ± 0.4

Table 6.32: Number of signal events at different stages of the Rpc2L1bH signal region selection. Only statistical uncertainties are shown.

Rpc3LSS1b, $\tilde{t}_1 \tilde{t}_1^*$ production, $\tilde{t}_1 \rightarrow t \tilde{\chi}_2^0 \rightarrow \tilde{t} W^\pm \chi_1^\mp$	
$m_{\tilde{t}_1} = 700\text{GeV}$, $m_{\tilde{\chi}_2^0} = 525\text{GeV}$, $m_{\tilde{\chi}_1^\pm} \approx m_{\tilde{\chi}_1^0} = 425\text{GeV}$	
MC events generated	5000
Expected for 36.1 fb^{-1}	2.4×10^3
≥ 3 SS leptons ($p_T > 20, 20, 10\text{ GeV}$), $Z \rightarrow e^\pm e^\pm$ veto	4.6 ± 0.5
Trigger	4.5 ± 0.5
≥ 1 b -jet ($p_T > 20\text{GeV}$)	3.6 ± 0.4

Table 6.33: Number of signal events at different stages of the Rpc3LSS1b signal region selection. Only statistical uncertainties are shown.

Rpc2L2bH, $\tilde{g}\tilde{g}$ production in the NUHM2 model							
$m_{\tilde{g}}$ [GeV]	300	350	400	500	600	700	800
Acceptance	0.8%	1.6%	2.1%	3.2%	3.5%	4.4%	4.0%
Efficiency	43%	49%	50%	49%	48%	43%	49%

Table 6.34: Rpc2L2bH signal region acceptance and reconstruction efficiency for $\tilde{g}\tilde{g}$ production in the NUHM2 model.

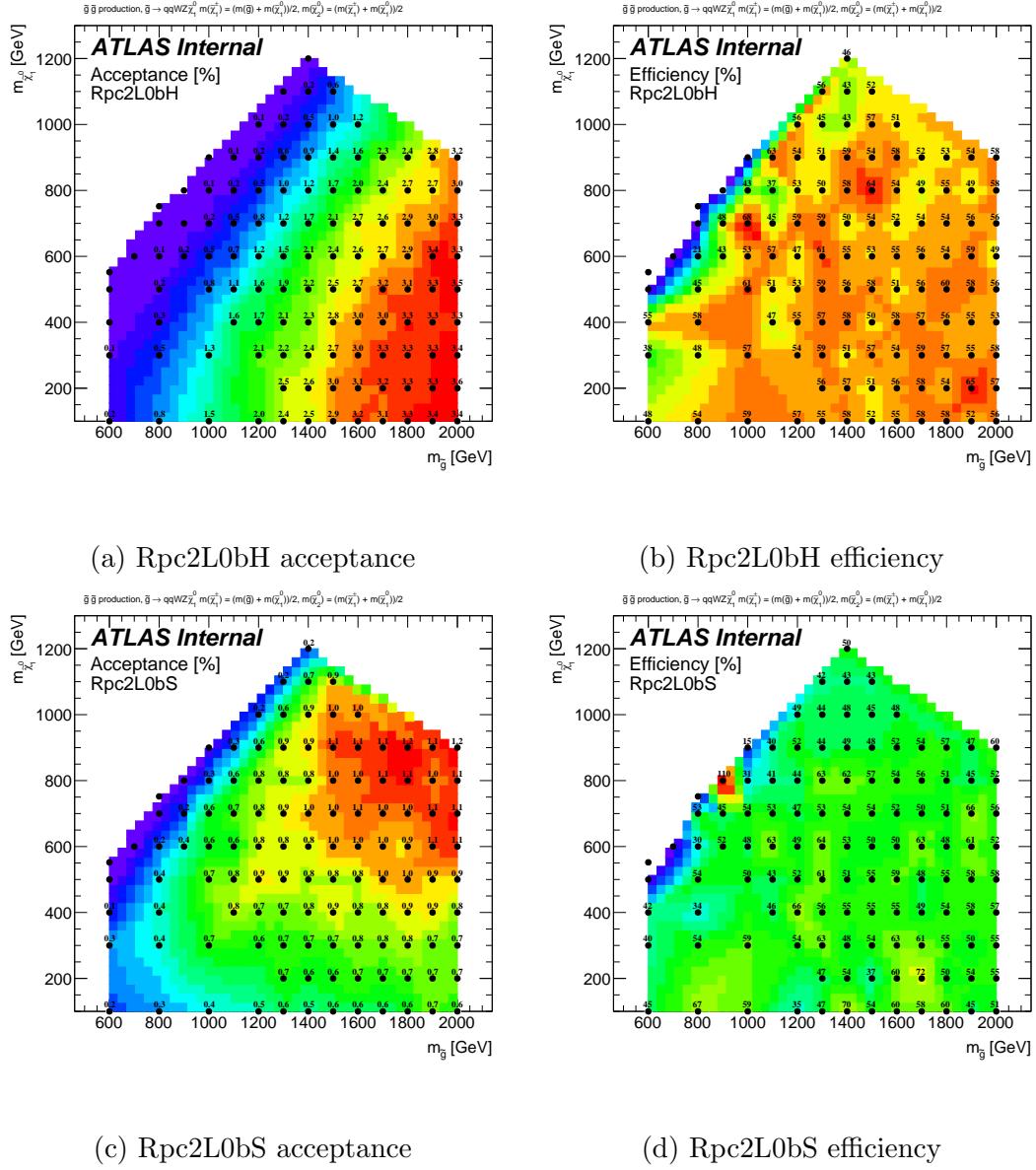


Figure 6.39: Signal acceptance (a,c) and reconstruction efficiency (b,d) for simplified models of $\tilde{g}\tilde{g}$ production with $\tilde{g} \rightarrow q\bar{q}'WZ\tilde{\chi}_1^0$ decays, in the signal regions Rpc2L0bH (a,b) and Rpc2L0bS (c,d).

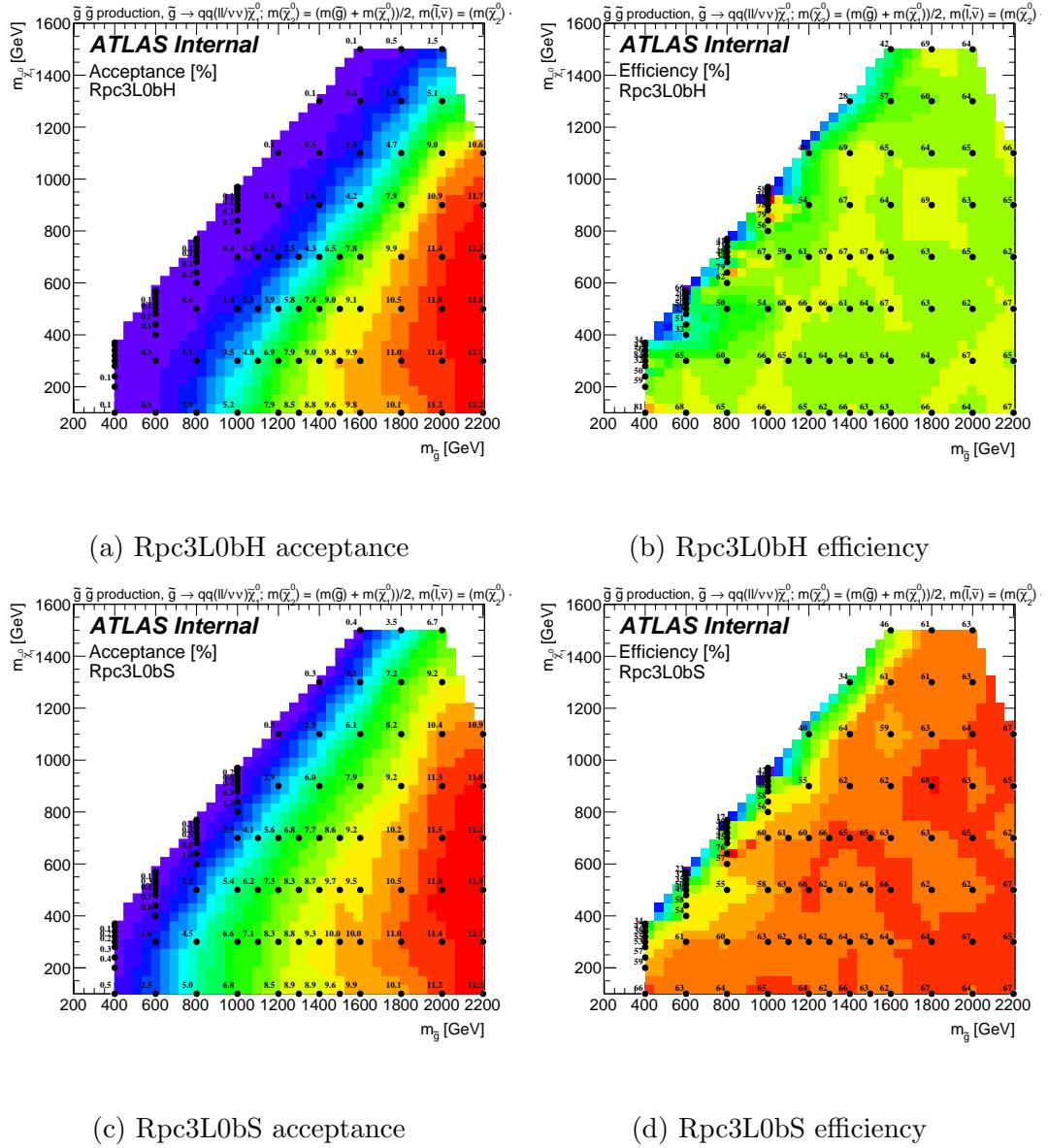


Figure 6.40: Signal acceptance (a,c) and reconstruction efficiency (b,d) for simplified models of $\tilde{g}\tilde{g}$ production with $\tilde{g} \rightarrow q\bar{q}(\ell\ell/\nu\nu)\tilde{\chi}_1^0$ decays, in the signal regions Rpc3L0bH (a,b) and Rpc3L0bS (c,d).

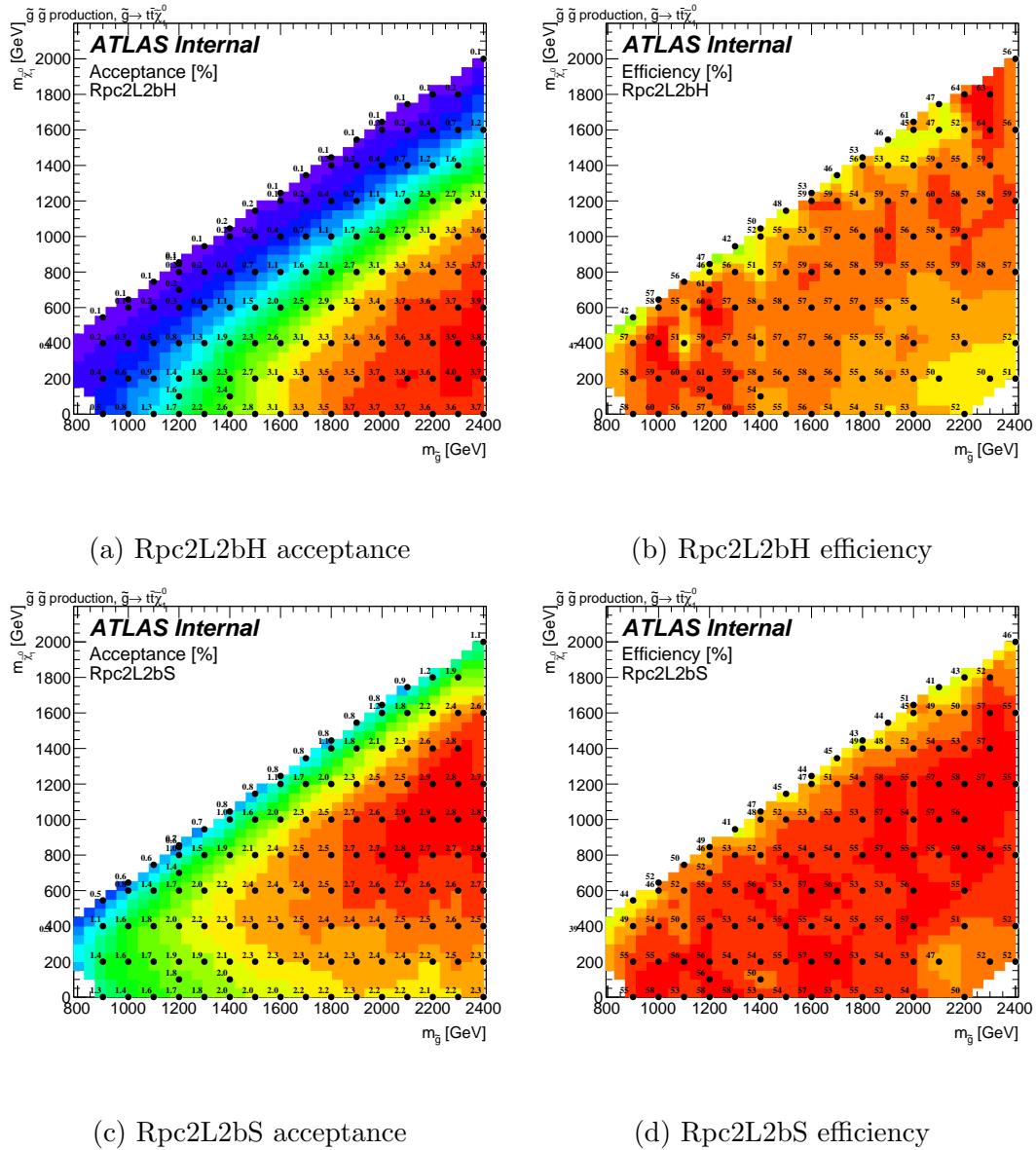


Figure 6.41: Signal acceptance (a,c) and reconstruction efficiency (b,d) for simplified models of $\tilde{g}\tilde{g}$ production with $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$ decays, in the signal regions Rpc2L2bH (a,b) and Rpc2L2bS (c,d).

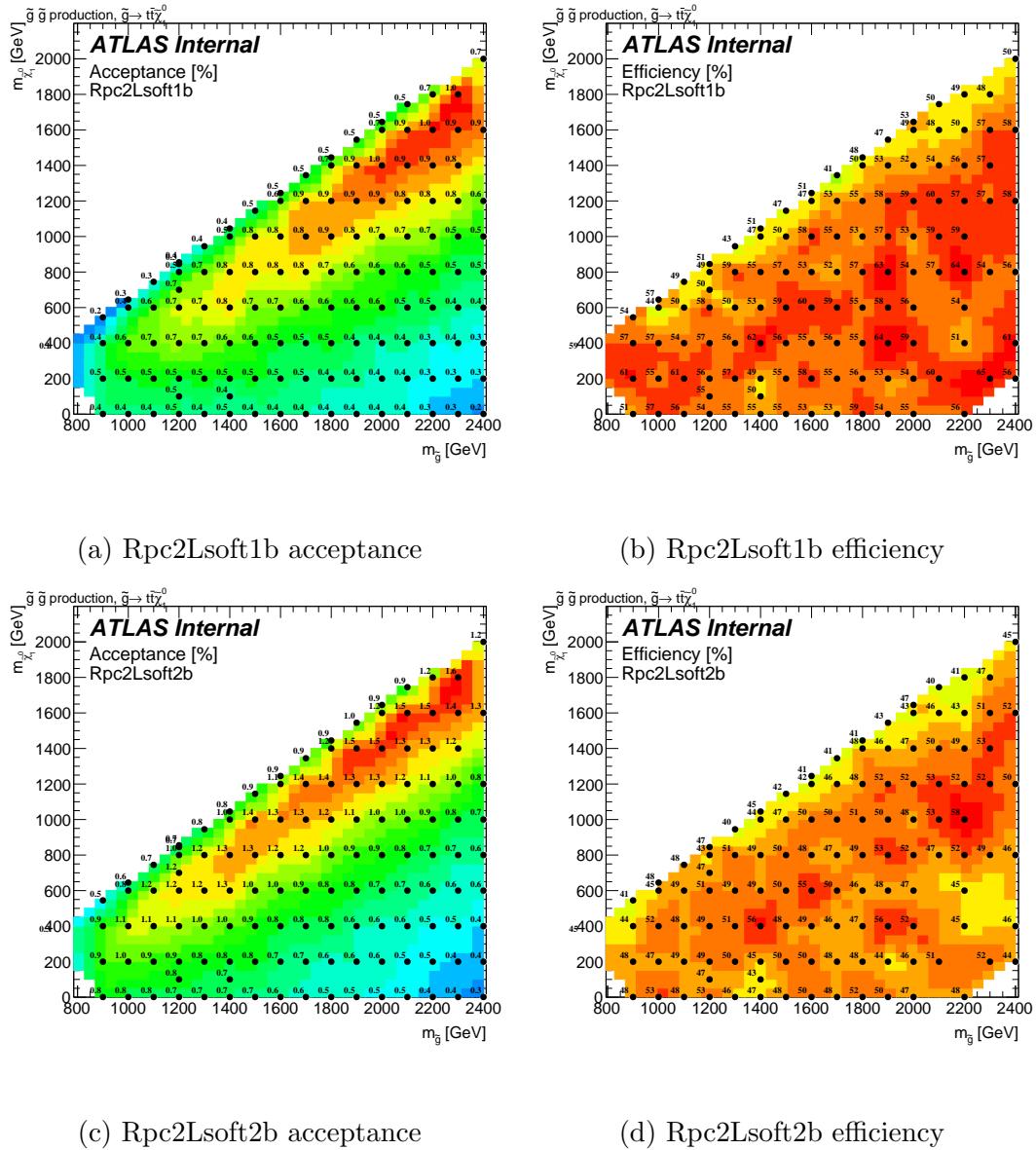


Figure 6.42: Signal acceptance (a,c) and reconstruction efficiency (b,d) for simplified models of $\tilde{g}\tilde{g}$ production with $\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$ decays, in the signal regions Rpc2Lsoft1b (a,b) and Rpc2Lsoft2b (c,d).

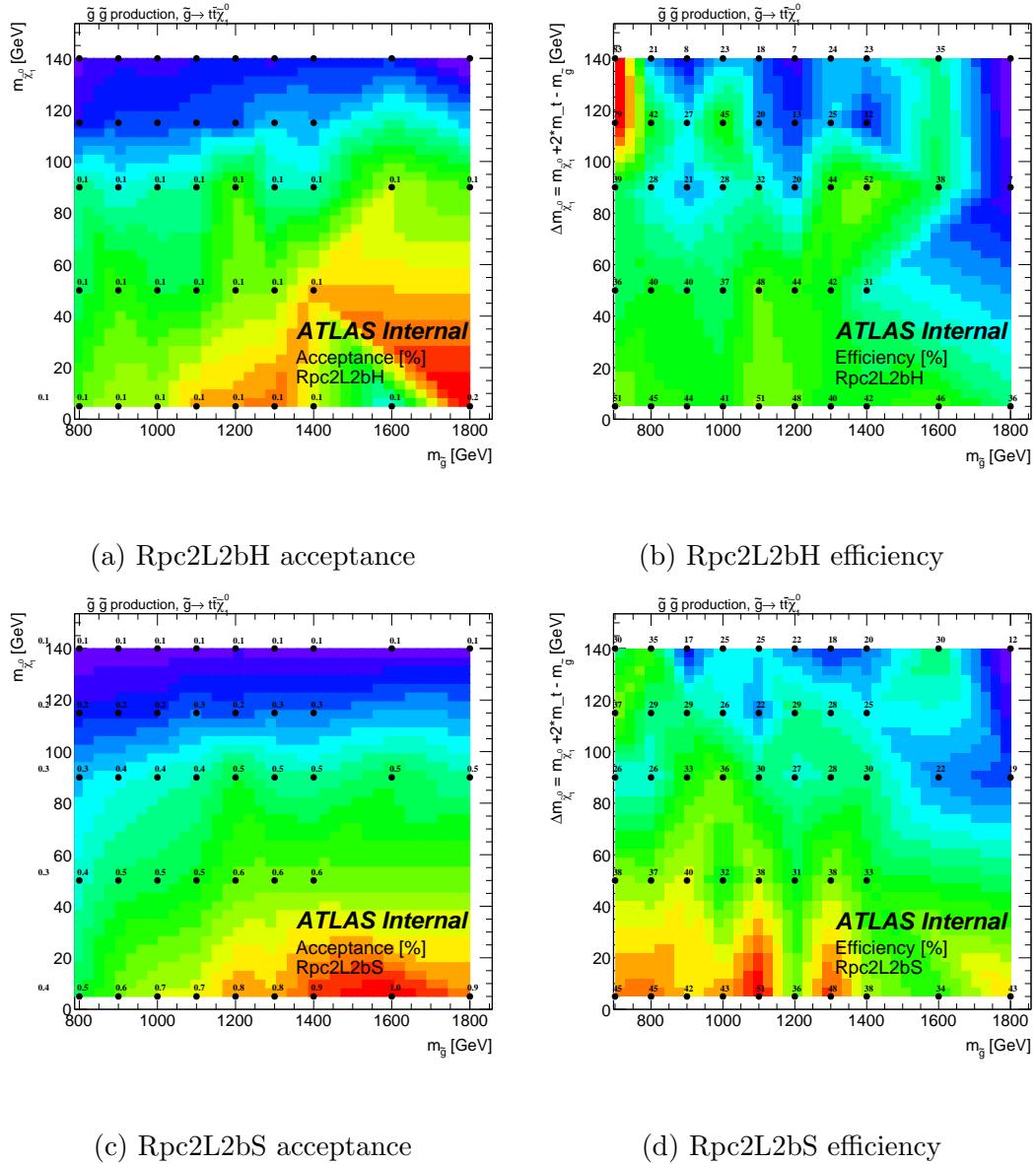


Figure 6.43: Signal acceptance (a,c) and reconstruction efficiency (b,d) for simplified models of $\tilde{g}\tilde{g}$ production with $\tilde{g} \rightarrow tWb\tilde{\chi}_1^0$ decays (region with $\Delta m(\tilde{g}, \tilde{\chi}_1^0) < 2m_t$), in the signal regions Rpc2L2bH (a,b) and Rpc2L2bS (c,d).

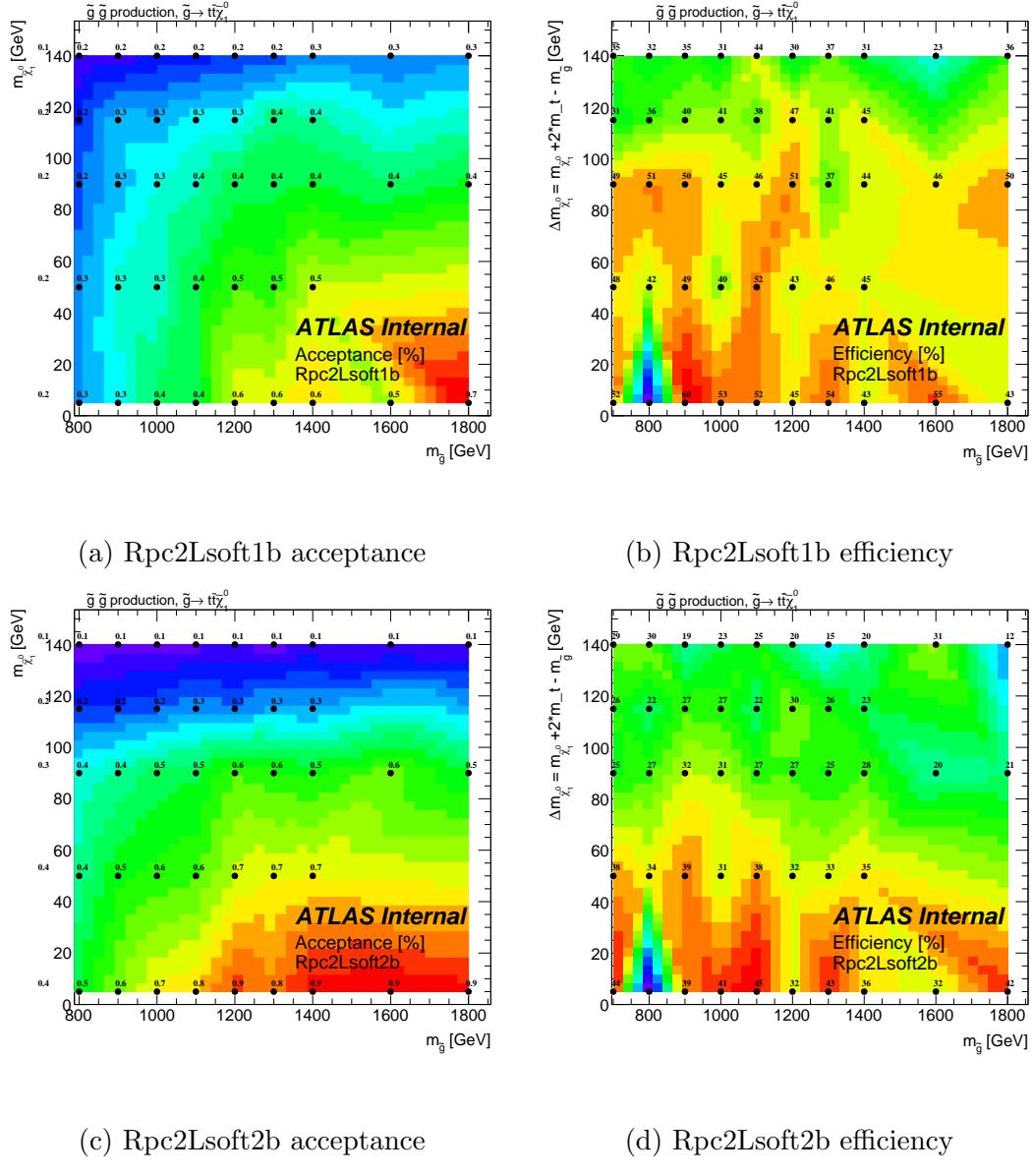


Figure 6.44: Signal acceptance (a,c) and reconstruction efficiency (b,d) for simplified models of $\tilde{g}\tilde{g}$ production with $\tilde{g} \rightarrow tWb\tilde{\chi}_1^0$ decays (region with $\Delta m(\tilde{g}, \tilde{\chi}_1^0) < 2m_t$), in the signal regions Rpc2Lsoft1b (a,b) and Rpc2Lsoft2b (c,d).

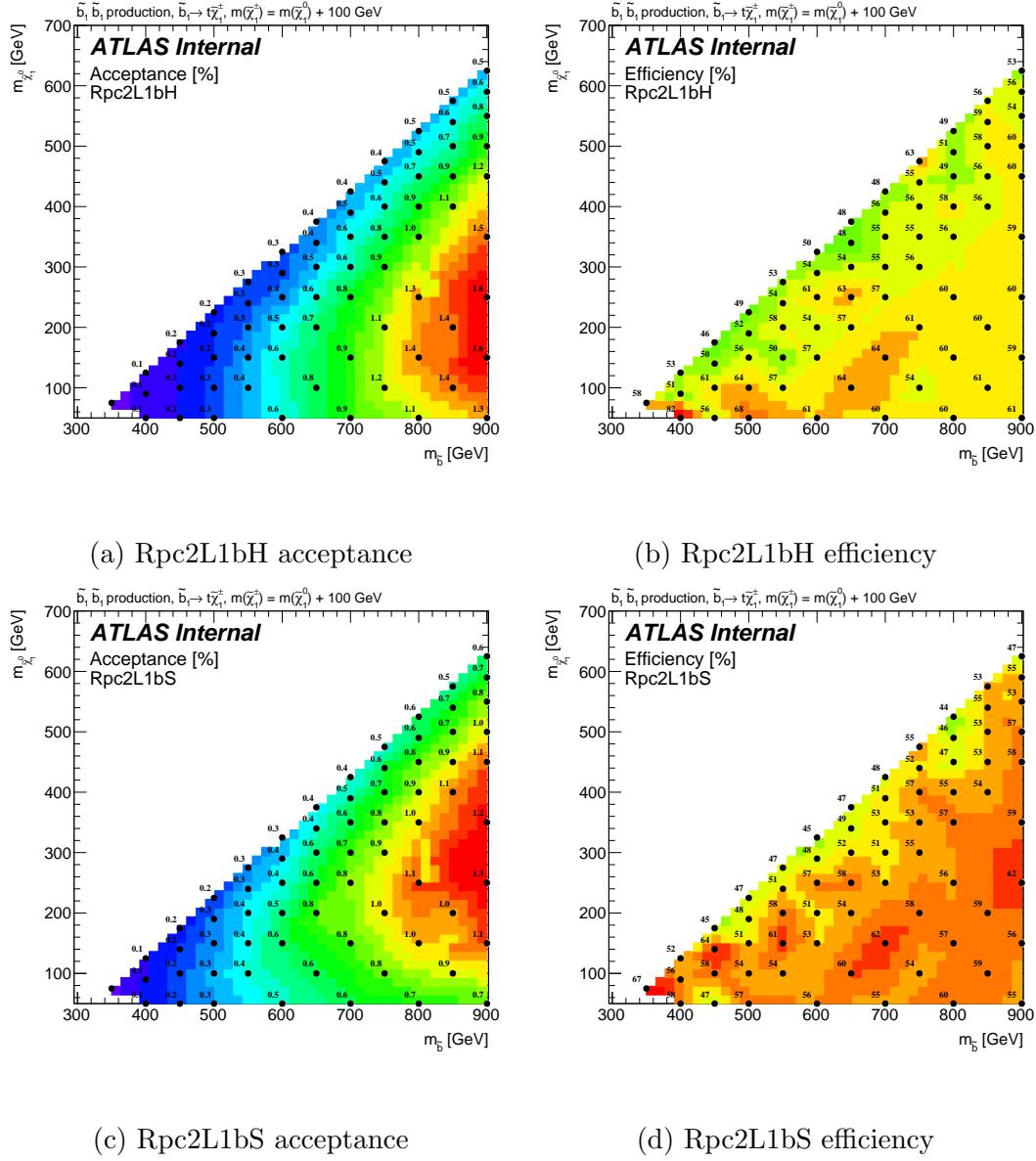


Figure 6.45: Signal acceptance (a,c) and reconstruction efficiency (b,d) for simplified models of $\tilde{b}_1 \tilde{b}_1^*$ production with $\tilde{b}_1 \rightarrow tW - \tilde{\chi}_1^0$ decays, in the signal regions Rpc2L1bH (a,b) and Rpc2L1bS (c,d).

Rpc3LSS1b, $\tilde{t}_1 \tilde{t}_1^*$ production, $\tilde{t}_1 \rightarrow t\tilde{\chi}_2^0$, $\tilde{\chi}_2^0 \rightarrow W^\mp \tilde{\chi}_1^\pm$						
$m_{\tilde{t}_1}$ [GeV]	550	600	650	700	750	800
Acceptance	0.3%	0.3%	0.3%	0.4%	0.4%	0.4%
Efficiency	36%	42%	44%	37%	33%	30%

Table 6.35: Rpc3LSS1b signal region acceptance and reconstruction efficiency for $\tilde{t}_1 \tilde{t}_1^*$ production with $\tilde{t}_1 \rightarrow t\tilde{\chi}_2^0$ ($\tilde{\chi}_2^0 \rightarrow W\tilde{\chi}_1^\pm$) decays.

Chapter 7

Statistical Analysis

Conclusion