Search for Displaced Photons from Exotic Decays of the Higgs Boson with the ATLAS Detector

Devin Mahon

Submitted in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2021

Abstract

Search for Displaced Photons from Exotic Decays of the Higgs Boson
with the ATLAS Detector

Devin Mahon

A search for displaced photons from exotic decays of the Higgs boson is presented, specifically targeting the relatively weak constraints on the branching ratio of Higgs boson decays to invisible particles imposed by other searches. 139 fb^{-1} of $p-p$ collision data at center-of-mass energy $\sqrt{s}$ = 13 TeV, collected between 2015 and 2018, is analyzed. Exploiting the excellent timing performance and longitudinal segmentation of the ATLAS electromagnetic calorimeter, the photon arrival time and extrapolation to the beam axis, or pointing, can be used to extract the displaced photon signature from Standard Model background processes. The background photon timing and pointing distributions are predicted using carefully constructed, data-derived templates. Signal contributions are analyzed in the context of a Gauge-Mediated Supersymmetry Breaking model in which the Higgs boson decays into two, long-lived next-to-lightest supersymmetric particle neutralinos, which each subsequently decay into a photon and a gravitino. Agreement with background is observed with no significant excesses for any signal models considered. Limits at the 95% confidence level are placed on the branching ratio of the Higgs boson to two neutralinos in the context of the various signal models.

Table of Contents

List of Figures

List of Tables

Acknowledgments

For the invaluable guidance, I owe an enormous thank you to my advisor, John Parsons. From my time at Nevis, to all of my years at CERN, and through the entire thesis writing process, I've had nothing but the sincerest of support from him—not only in my academic pursuits, but in my personal life as well. His knowledge, expertise, and kindness have undoubtedly been instrumental in the successes of my graduate career and beyond. I couldn't have asked for a better mentor.

I also have to thank all of the members of the Columbia ATLAS group over the years and my fellow Columbia graduate students for all of the helpful discussions and support and for being part of a very welcoming and friendly environment. Thanks to all of the staff at the Columbia Physics Department and especially to those at Nevis Labs over the years, who have helped to make my graduate career and all the traveling back and forth between New York, Nevis, and Geneva as rewarding as possible.

Thanks to those Columbia colleagues who worked with me at Nevis, shared an office at CERN, and/or contributed to the work comprising this thesis, including Alex Emerman, Jue Chen, Kiley Kennedy, Elena Busch, Andrew Smith, Gustaaf Brooijmans, Mike Tuts, Kalliopi Iordanidou, Tim Andeen, Inês Ochoa, Julia Gonski, Alan Kahn, Emily Simon, and honorary Columbia summer student, Elena Villhauer. And thanks to my fellow graduate students at Columbia

A special thanks to Nikiforos Nikiforou for his help and patience, in whose esteemed footsteps I ended up following quite closely. His expertise in the LAr system and displaced photons served as an admirable model for my own work. And his uncanny ability to produce results sometimes even before I could finish asking never did get old.

Thanks to the entire LAr operations team, and especially to my qualification task supervisor, Emmanuel Monnier. I'm proud to have been a part of this incredible, dedicated team of people passionate about their work. Their welcome and support during my operations work is deeply appreciated.

The analysis presented in this thesis was, by no means, a one-person job. In addition to the Columbia ATLAS members who contributed and to Nikiforos, I'd like to acknowledge

the many contributions from all of the members of the ATLAS Run 2 displaced photon analysis from UC Berkeley and Indiana University as well.

Besides John Parsons, I'd also like thank the other members of my thesis examining committee, Frederik Denef, Georgia Karagiorgi, Frederik Paerels, and Hong Ma for their review of this work.

Finally, this thesis would never have been possible without the loving and unwavering support of my family and friends. And so I have to express the utmost gratitude to my parents, Dan and Mary, my sister, Marcy, my grandmother, Sally, and my aunt and uncle, Kay and Glen. A big thank you as well to all of my friends from back in Texas, especially to Michael, and to those in France and Switzerland throughout my time at CERN. You all certainly have a part in the words that follow, and I cannot express strongly enough how much your support has meant to me.

Foreword

With the discovery of the Higgs boson in 2012, the long saga of cataloguing the particles in the Standard Model (SM) was finally complete. Numerous, rigorous tests of the full predictive power of the SM were carried out with astonishingly precise agreement, lending authoritative credence to its validity. But along the way, a collection of anomalies and mathematical puzzles was being amassed that sowed doubts about the completeness of its description of nature. Despite its tremendous successes, it became clear that the SM could not, in fact, be a full description of nature.

When the LHC began operation in 2008, there was great anticipation not only for the discovery of the final, most elusive of the SM particles, but also for what other mysterious it may uncover. With the Higgs boson under its belt, the physics world looked forward to a new, wide-open frontier. Over the past decades of particle hunting, accessing new energy scales had nearly always revealed surprises, leading the field in exciting new directions, requiring novel formulations and insights to explain the newly observed phenomena. But as of late, physics has had a new problem: it's too predictable. To an unnerving degree, we keep seeing what we expect to see. Many scientific disciplines would be thrilled to, as it were, have such a complete understanding that their experiments keep coming up with all the right answers. But with the nagging questions casting doubt on the SM's very foundations, physicists took their repeated, decisive successes with a sense of bittersweet unease.

Despite the efficient operation of the LHC and its various experiments over the past decade or so, no new physics has yet definitively emerged. A glut of pet theories have come and gone over the years, attempting to resolve the SM's shortcomings. Thus far, some have been ruled out experimentally; others have been modified with every new, faint fluctuation; and still new ones are being churned out, hoping to uncover some new, unexplored phase space of which the LHC could yet provide a glimpse. Given the immense logistical and scientific hurdles that await the construction of the next accelerator complex capable of accessing a new energy regime, as well as the lack of any obvious sign of new physics from the current machine, it has become the experimental community's mission to both test, as precisely as possible, the SM's predictions and to leave no phase space untouched in the

search for new phenomena. My own research efforts over the past six years in the pursuit of my PhD have—I'd like to think, at least in some small way—contributed to this pursuit.

My ATLAS journey began in the summer of 2015 at Nevis Labs, before I began my first graduate courses. There, I contributed to the study of the analog-to-digital converters (ADCs) used to digitize the LAr calorimeter ionization pulses on the front-end boards (FEBs) for the Phase I upgrade. These ADCs are custom, radiation-hard 12-bit pipeline ASIC ADCs, developed with 130 nm CMOS technology. Specifically, I carried out a performance study measuring the contribution of the ADC to the fractional energy resolution. By generating simulated ionization pulses, varying their amplitudes, and reconstructing their peaks, I was able to determine that, for typical pulse heights, the total resolution was $\lesssim 0.8$ ADC counts, comfortably meeting the design specifications.

The following summer, I investigated a technical issue observed on a small fraction of ADCs in which output code errors occurred at a rate of a few Hz. These errors manifested as spikes during sine wave scans of the full dynamic range. Studying the spike rate as a function of the supplied voltages, we determined that the issue was attributable to the common-mode conversion between the digital and analog modules of the ADC. As a result, the operating analog voltage was slightly raised, eliminating the spikes while increasing power dissipation by only a few percent. With this adjustment and a new, associated quality-control procedure, the ADCs met all specifications. These ADCs are now being installed and commissioned on-detector for use in the upcoming LHC run. The paper on the performance and quality control of the ADC, of which I am one of the authors, can be found in Ref. [1].

After completing all of my graduate courses in New York, I moved to Geneva in the summer of 2017 to join the LAr online operations team. I began as an online software expert, solving time-sensitive issues with the LAr software infrastructure and serving 24/7 on-call shifts during data-taking. Concurrently, I developed tools for online monitoring to improve the efficiency of identifying data-taking issues. After some time on the job, I became a so-called LAr Super Shifter, serving as a LAr representative in the training of ATLAS Control Room calorimeter shifters. In this role, I helped to organize and lead training sessions and improved shifter documentation to ensure clear communication with experts. After more time on the job, I also served several weeks as the LAr Run Coordinator, being on-call as

the first point of contact for the LAr system for daily operations. The full LAr operations team, of which I was a small part, had an extraordinarily successful Run 2, achieving a 99.6% overall data-taking efficiency. I also had the opportunity to present this impressive performance at the INSTR2020 conference in Novosibirsk, Russia [2].

In parallel with my online operations commitments, I began work on the LAr offline timing calibration. This task involved improving the roughly 1 ns timing resolution achieved online using prompt data from W and Z boson decays. Starting with an established procedure used in Run 1, I obtained corrections for the full Run 2 dataset while making a number of improvements, including extending the range of the corrections to lower energies, implementing smoothed fits for better fits and automation, and adding statistical checks to guard against fluctuations with minimal sacrifices in efficiency. In the end, in the high-energy limit, timing resolutions of $\sim 190 - 280$ ps are achieved, reaching nearly down to the calculated, irreducible beam spread timing component of $\sim$190 ps. This impressive performance, in fact, provides the most precise timing measurement in ATLAS for Run 2. Since this timing behavior is not modeled in simulation, a smearing tool was constructed to match the calibration resolutions observed in data. The full details of the timing calibration procedure are documented in Ref. [3].

Little did I know as I worked through the timing calibration that I would be its first customer. As I began the search for a thesis topic, I soon found myself drawn to unconventional signatures. In the face of the disappointing lack of evidence for BSM physics, I was intrigued by the argument that we might be looking in the wrong places. I was especially intrigued by the tantalizing, unexplored regions of phase space occupied by displaced photons from the decays of long-lived particles. It soon became clear that my extensive calorimeter expertise through my experiences with the LAr system could be neatly synthesized with my interest in long-lived particle to perform a search for displaced photons. This thesis represents the product of that realization. Though I certainly did not know it at the time, the work I began when I arrived at Nevis Labs that first summer morning in 2015 would lead in a remarkably and quite satisfyingly continuous thread through all of the pages that follow.

Introduction

This thesis searches for evidence of displaced photons produced from exotic decays of the Higgs boson by examining the data collected by the ATLAS detector. The search is designed to be largely model-independent, maintaining sensitivity to any soft, displaced photons; however, in order to investigate sensitivity to proposed models, results are also evaluated in the context of a supersymmetry (SUSY) model in which a long-lived neutralino decays radiatively. Background estimation is conducted in an entirely data-driven manner, using templates derived from prompt background in order to model the variable photon purity among different analysis regions.

Chapter 1 introduces the Standard Model (SM)—the theoretical framework undergirding the field of high-energy physics—including its foundations, successes, and limitations. Chapter 2 offers SUSY as a potential solution to several shortcomings of the SM, discussing the general formalism and phenomenology. In particular, SUSY theories that may give rise to the displaced photons targeted by the analysis presented in this thesis are given special emphasis.

Background into the experimental apparatuses used to collect the data are then presented. The accelerator complex, the Large Hadron Collider (LHC), and its basic principles of operation are outlined in Chapter 3. A detailed description of the many systems that comprise the ATLAS detector, which measures the product of the collisions engineered by the LHC, follows in Chapter 4. A particular focus is placed on the Liquid Argon (LAr) electromagnetic calorimeter, which is responsible for the measurements of the proposed displaced photons. The various measurements and algorithms used by the ATLAS detector to identify and characterize different particles are discussed in Chapter 5.

The selections targeting the signature of interest are given in Chapter 6. The full analysis strategy, including region definitions, background estimation, and optimization can be found in Chapter 7. Systematic uncertainties and their incorporation into the analysis likelihood model are discussed in Chapter 8. The expected sensitivities to the target SUSY model based on a validation region are presented in Chapter 9.

The results from the signal region are given in Chapter 10, including analyses of the level

of evidence for the proposed signal and limits on the SUSY model based on this evidence. Chapter 11 briefly summarizes the entire analyses, discusses the results in context, and assesses the future outlook.

Chapter 1: The Standard Model

The SM is a local quantum field theoretical framework describing the strong, weak, and electromagnetic interactions. Here, the general principles of the model are discussed, including a brief illustration of its theoretical derivation, an overview of the resulting interactions, its successes in empirical validation, and its limitations.

1.1 Foundations & Assumptions

A general, underlying mathematical principle of the quantum mechanical formulation of the SM is the principle of least action. This principle states that the evolution of a physical system follows the path in configuration space for which the action is stationary, i.e. unchanged to first order in small perturbations of the system configuration. In mathematical terms, the action S obeys $\delta S = 0$. For generic fields $\phi_i(x)$, the action can be defined as the spacetime integral of the Lagrangian density $\mathcal{L}$ (often simply known as the Lagrangian):

$$S = \int d^4x \mathcal{L}(\phi_i(x), \partial_\mu \phi_i(x)) \tag{1.1}$$

For descriptive convenience, the Lagrangian is considered rather than the action. Several well-motivated assumptions about the form of the Lagrangian are generally made to ensure that it has properties consistent mathematically, as well as physically, with what is observed in nature. These assumptions include that $\mathcal{L}$ is invariant under the Poincaré group (spacetime translations, rotations, and boosts), that it is a function of the fields and their derivatives at a single spacetime point x (making it a local field theory), and that it is real (to conserve probability). In order to avoid unphysical infinite quantities in the integral to compute the action, $\mathcal{L}$ must contain only terms of dimension four or fewer in the fields

and their derivatives. This dimensionality requirement is equivalent to the requirement of so-called renormalizability [4]. Finally, in the vein of perturbation theory, $\mathcal{L}$ is considered to be analytic in the fields, which can be written as some polynomial expansion.

Given these assumptions, the entirety of structure of the SM can be constructed by applying the following rule: add all possible terms to $\mathcal{L}$ not forbidden by a spontaneously broken $SU(3)\times SU(2)\times U(1)$ gauge symmetry. These three factors correspond to the strong, weak, and electromagnetic interactions, respectively. As will be illustrated in Section 1.2, these symmetries determine the general properties of these interactions, leaving to experiment only the determination of the values of fundamental constants. It should be noted that this symmetry group was developed over many years through careful matching of experimental observations. However, once this mathematical symmetry is chosen, the behavior of the theory falls out, meaning all that is needed to construct the properties is the simple decision of the proper symmetry group. Thus, though it is in some sense a post-hoc model constructed to fit data, the simplicity and strictly mathematical nature of the inputs lends great power to its soundness and credibility as a theory.

1.2 Symmetries

1.2.1 Consequences of Symmetries

To illustrate the way in which symmetries essentially determine the SM Lagrangian, let us consider the electromagnetic interaction, whose quantum formulation is known as quantum electrodynamics (QED) and is defined by a local $U(1)$ gauge symmetry.

As a starting point, with foreknowledge of the end result, the Lagrangian based on the Dirac equation is chosen:

$$\mathcal{L}_{\text{Dirac}} = \bar{\psi}\Big(i\gamma^{\mu}\partial_{\mu} - m\Big)\psi \tag{1.2}$$

where ψ is a relativistic spin-1/2 field with a particle of rest mass m and γ^{μ} denotes the

so-called gamma matrices[1] [5]. Natural units[2] have also been employed.

It should also be noted that the Dirac equation itself can be derived from representations of the Lorentz group, and thus using it as a starting point does not reduce the generality of the results. Imposing the local $U(1)$ gauge symmetry is equivalent to requiring invariance under a local phase transformation. This transformation only changes the arbitrary phase of the particle wave functions and thus has no observable consequence. That is to say that the local $U(1)$ gauge symmetry is physically justified. Denoting this phase transformation as $\psi(x) \rightarrow e^{iq\Lambda(x)}\psi(x)$ transforms the Dirac Lagrangian in Equation 1.2 in the following manner:

$$\mathcal{L}_{\text{Dirac}} \rightarrow \bar{\psi}\Big(i\gamma^{\mu}(\partial_{\mu} + iq\partial_{\mu}\Lambda(x)) - m\Big)\psi \tag{1.3}$$

Comparing to Equation 1.2, it is clear that the derivative does not transform covariantly. Therefore, let us recast the Lagrangian, introducing a gauge field $A_{\mu}(x)$ that transforms under the $U(1)$ symmetry as $A_{\mu}(x) \rightarrow A_{\mu}(x) + \partial_{\mu}\Lambda(x)$ with a so-called covariant derivative defined as: $D_{\mu} \equiv \partial_{\mu} - iqA_{\mu}$. D_{μ} is constructed explicitly such that $\bar{\psi}D_{\mu}\psi$ is invariant under the $U(1)$ gauge transformation, which then allows us to write a new invariant Lagrangian $\mathcal{L}'_{\text{Dirac}}$ as follows:

$$\mathcal{L}'_{\text{Dirac}} = \bar{\psi}\Big(i\gamma^{\mu}D_{\mu} - m\Big)\psi = \mathcal{L} + q\bar{\psi}\gamma^{\mu}\psi A_{\mu} \tag{1.4}$$

Thus, the gauge invariance requirement necessitates an interaction term of the spinor field with a vector field A_{μ}. Having introduced this new vector field, and keeping with the principle that all terms must be present that are not forbidden by a symmetry, it becomes necessary to consider kinetic and mass terms associated with A_{μ}. Let us denote the electromagnetic field tensor $F_{\mu\nu} \equiv \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ such that the kinetic term may be written as $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$. The mass term must be of the form $\frac{1}{2}m^2A_{\mu}A^{\mu}$. However, this term is only gauge invariant if $m = 0$, so this mass term cannot be present.

---

[1]The gamma matrices are defined by the anticommutation relation $\{\gamma^{\mu}, \gamma^{\nu}\} = -2\eta^{\mu\nu}$ (using the Minkowski metric with signature +2).

[2]In natural units, all kinematic units are expressed in terms of energy by setting the speed of light c and Planck's constant $\hbar$ to 1.

Now that the Lagrangian has been formulated in a gauge-invariant representation and all possible terms not disallowed by the gauge symmetry have been accounted for, the new Lagrangian may be written as:

$$\begin{aligned}\mathcal{L}_{\text{QED}} &= \bar{\psi}\Big(i\gamma^\mu D_\mu - m\Big)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \bar{\psi}\Big(i\gamma^\mu \partial_\mu - m\Big)\psi + q\bar{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\end{aligned} \tag{1.5}$$

This Lagrangian is that of QED, and many of its properties may be read off of the equation itself. Equation 1.5 describes the interactions of relativistic particles described by the spin-1/2 field ψ with charge q and mass m with a vector field A_μ mediated by a massless spin-1 vector boson (i.e. the photon). Thus, purely by arguments of symmetry and naturalness, the quantum nature of electromagnetism can be determined, leaving only the measurements of the mass m and coupling constant q to experiment.

In this same manner, the entire structure of the SM may be constructed, extending the local $U(1)$ gauge symmetry of QED to the full SM $SU(3) \times SU(2) \times U(1)$ gauge symmetry with spontaneous symmetry breaking of $SU(2) \times U(1)$ via the Higgs mechanism, which is discussed in Section 1.2.2.

1.2.2 Spontaneous Symmetry Breaking

A key feature of the symmetry defining the SM is the spontaneous breaking of the $SU(2) \times U(1)$ symmetry describing the unified electroweak (EW) interaction into the separate symmetries of the weak force and electromagnetism.

As discussed for the photon in Section 1.2.1, the explicit photon mass term in the Lagrangian violated the gauge symmetry, thus requiring the mass to be zero. The same argument applies to other gauge bosons. And in a similar manner, for fermions, explicit fermionic mass terms of the form $m\bar{\psi}\psi$ are forbidden by imposed symmetries. Namely, since the chirality of fermionic fields causes the left- and right-handed components to transform differently under the $SU(2)$ symmetry, such an explicit mass term is disallowed for violating chiral symmetry. However, the phenomenon of EW symmetry breaking via the so-called Higgs mechanism provides a theoretical reparametrization of the Lagrangian that gives mass

to some of these particles, as is observed in nature [6–11].

Under the constraints the $SU(2) \times U(1)$ gauge symmetry, a complex scalar potential may be added to the SM Lagrangian. The most general, renormalizable form of such a potential is:

$$V(\phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \tag{1.6}$$

where Φ is a complex scalar $SU(2)$ doublet.

In the case that $\mu^2 > 0$, the minimum occurs at $\Phi = 0$, and thus the so-called vacuum expectation value (VEV) of the field is 0. However, as is observed experimentally, nature prefers the case where $\mu^2 < 0$. In this case, the VEV is nonzero and the degeneracy of the minimum in the complex plane allows us to arbitrarily choose the following:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.7}$$

This nonzero VEV breaks the $SU(2) \times U(1)$ symmetry into the $U(1)$ symmetry of electromagnetism, implying that three massless Goldstone bosons are produced. To make the structure of the physical particles most clear, one can choose the so-called unitarity gauge in which the Goldstone boson components of the scalar fields are zeroed or "rotated away". In other words, all fields can be made to be orthogonal to the Goldstone fields, leaving three of the four degrees of freedom of the potential Φ to be absorbed into the longitudinal components of three gauge bosons which thus now acquire mass. In this way, the four generators of $SU(2) \times U(1)$ mix with the neutral and two charged degrees of freedom of the Goldstone bosons to yield four bosons in the unitarity gauge: one neutral massive boson (labelled Z), two charged massive bosons ($W^\pm$), and one remaining neutral boson (the photon) associated with the $U(1)$ symmetry of electromagnetism that remains massless since there are no Goldstone degrees of freedom left to absorb. This acquisition of mass by these three gauge bosons via the broken $SU(2) \times U(1)$ symmetry is known as the Higgs mechanism. The fourth degree of freedom of Φ not assumed by the three now massive gauge bosons is left standing as a neutral scalar boson, known as the Higgs boson.

This spontaneous breaking of the $SU(2) \times U(1)$ symmetry allows for the structure of

the weak and electromagnetic interactions observed in nature to become apparent. The consequences of the Higgs mechanism also provide a clear interpretation for the fermion masses since fermions can couple to the scalar Higgs field via Yukawa interactions, endowing them with mass.

1.3 Survey of Particles & Interactions

Following the symmetry principles outlined and demonstrated in Section 1.2, the full $SU(3)\times SU(2)\times U(1)$ symmetry of the SM establishes the structure of particles and interactions present in the model. The SM symmetries, however, allow for multiple so-called generations of particles, which share the same properties as other particles across generations except for their mass and flavor quantum numbers (conserved by all but the weak interaction). The number of generations is thus not constrained by symmetries. However, strong experimental evidence suggests that there are only three such generations. The SM symmetries also allow for each particle to have a so-called antiparticle, which has the same mass but inverted under charge conjugation (all charges, i.e. internal quantum numbers). To date, all particles of the SM have been discovered experimentally (though not all of their expected properties have been fully verified). Figure 1.1 shows all of these particles and many of their expected or measured properties. In this section, the general structure of the particles and interactions will be discussed.

The $SU(2)\times U(1)$ symmetry of the SM is identified with the EW interaction. Generally speaking, as discussed in Section 1.2, spontaneous symmetry breaking divides this symmetry into the $SU(2)$ symmetry of the weak interaction and the $U(1)$ symmetry of the electromagnetic interaction. The symmetries are often written more explicitly as $SU(2)_L$ and $U(1)_Y$, where L indicates that the weak interaction couples only to fermions of left-handed chirality (a property determined by experimental observation) and Y denotes weak hypercharge. Weak hypercharge is defined as $Y \equiv 2(Q - I_3)$, where Q denotes electric charge and I_3 is the third component of weak isospin[3].

The spin-1 vector bosons mediating the EW interaction are the $W^{\pm}$ and Z^0 bosons

[3]Weak isospin corresponds to the weak interaction gauge symmetry and groups left-handed doublets with $I_3 = \pm\frac{1}{2}$ in each generation of both quarks and leptons.

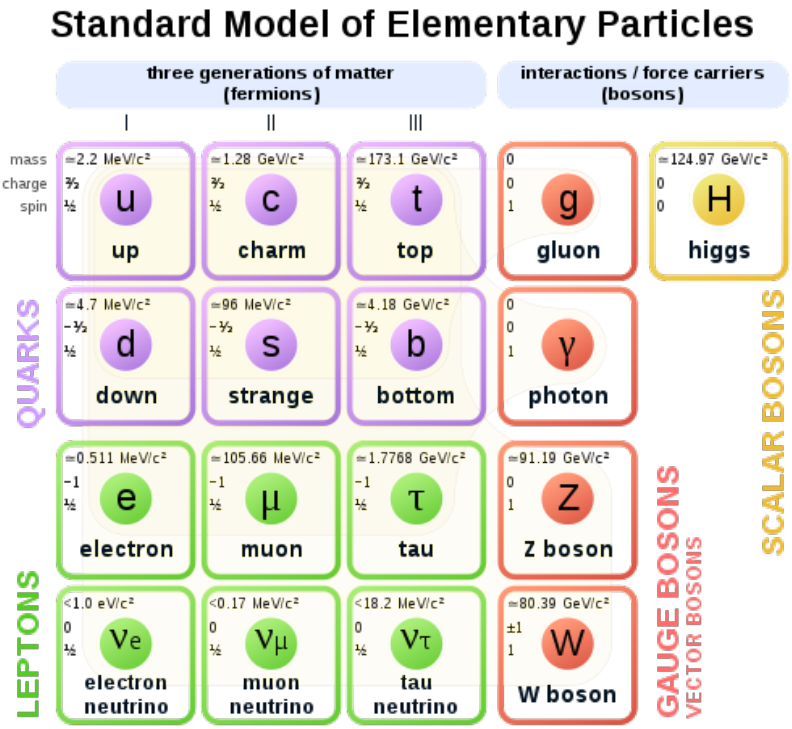


Figure 1.1: A diagram of the particles of the SM and their properties.

and the photon. The phenomenon of EW symmetry breaking via the Higgs mechanism gives mass to the two $W^{\pm}$ bosons and one Z boson (80 GeV and 91 GeV, respectively) and allows them to be identified with the weak force. The fact that these bosons are massive supresses their strength (hence the "weak" label) and makes their interactions short range. As previously mentioned, the weak force is coupled only to left-handed fermions, making it a maximally parity violating interaction. The massless photon mediates the electromagnetic force between all particles with electric charge. This masslessness facilitates the infinite range of this force. The scalar Higgs boson, whose field breaks the EW symmetry, was discovered with a mass of 125 GeV in 2012 by the CMS and ATLAS collaborations independently, thus completing the experimental observation of all SM particles [12].

There are six spin-1/2 leptons (and their corresponding antiparticles) that interact via the weak interaction, three of which are also electrically charged and thus couple to the photon. The electron, muon, and tau are the charged leptons with associated electron, muon, and tau neutrinos. Each respective pair of charged lepton and neutrino is known as a generation, and the number of particles minus the number of antiparticles in each generation is generally conserved in the SM[4]. The observation of neutrino oscillations necessitates small ($\lesssim$ 20 MeV) masses for the neutrinos, and the precise values and ordering of the masses among the generations is currently unknown.

The remaining force carriers, the eight massless, spin-1 gluons, are associated with the eight generators of the $SU(3)$ SM symmetry and the strong interaction, also referred to as quantum chromodynamics (QCD) [13, 14]. This interaction mediates the forces between gluons themselves and spin-1/2 quarks and thus explains the dynamics of hadrons, composite particles of two or more quarks. The charges of the strong interaction are known as color charges, of which there are three, denoted red, green, and blue (and antired, antigreen, and antiblue for the respective antiparticles). The eight gluons can be represented by eight linearly independent color states, forming an octet in the adjoint representation. The fact that gluons themselves carry color charge leads to a qualitative difference in the nature of the strong interaction in comparison to, for example, the electromagnetic force which also has a

[4]There are known processes such as neutrino oscillations that violate the conservation of lepton number for each generation, but the total lepton number across all generations is exactly conserved in the SM.

massless gauge boson. The $SU(3)$ symmetry allows for gluon self-interaction terms, which, roughly speaking, prevent gluons from spreading out infinitely in space, instead leading them to stick together and cluster between the quarks they bind. The structure imposed by the $SU(3)$ symmetry leads to two important features of the strong interaction: color confinement[5] and asymptotic freedom. Color confinement refers to the fact that particles with color charge (i.e. quarks and gluons) are never observed alone but rather in color-neutral bound states. This phenomenon can be understood as a consequence of the roughly constant force between two bound quarks generated by gluon exchange, which then requires increasing energy to separate them. The energy increases very rapidly over distances of order one femtometer until the point that there is sufficient energy to produce a quark-antiquark pair from the vacuum. Thus, two color-neutral bound states are produced rather than two isolated color charges. Asymptotic freedom refers to the decrease in the coupling constant with increasing energy [15, 16]. Such a property is a technical feature related to the requirement of renormalizability within the context of an $SU(3)$ symmetry. In the high energy limit, asymptotic freedom implies that quarks are loosely bound and can even form a phase known as a quark-gluon plasma in which quarks and gluons are free in an equilibrium state. But in the low energy limit—such as energies present in atomic nuclei—the coupling strength increases, and quantum computations based on perturbation theory fail, making calculations challenging to perform.

There are six spin-1/2 quarks (and their corresponding antiparticles) that interact via the strong, weak, and electromagnetic interactions. The so-called up-type quarks of electric charge +2/3 (in units of the elementary charge, i.e. that of a proton) are called the up, charm, and top quarks. The down-type quarks of electric charge -1/3 are known as the down, strange, and bottom quarks. Like the leptons, the quarks are grouped into three generations (paired up- and down-types in the order listed). The quark masses among each type are larger with each increasing generation—a pattern matched with the charged leptons and possibly with the neutrinos.

[5]Strictly speaking, color confinement is not a proven property, but the mechanism that allows it is well-motivated theoretically, and no confirmed experimental evidence for isolated color charge exists.

1.4 Empirical Validation

The incredibly accurate and precise predictions of the SM lend credence to its descriptive power. The SM successfully predicted the existence of the gluon, the top and charm quarks, the tau neutrino, and the W, Z, and Higgs bosons—as well as many of their properties—prior to their discoveries. Its prediction for the magnitude of the anomalous magnetic moment of the electron matches experimental observation beyond 10 significant figures, making it one of the most accurately-verified physical predictions of all time [17]. Any theory with such beautiful structure and remarkable predictive power deserves to be taken seriously. However, the SM as a description of the natural world cannot be considered complete and comes with its own set of mathematical peculiarities that suggest deeper structural issues.

1.5 Limitations

Several fundamental natural phenomena are completely absent from the SM. Some prominent omissions include any consideration of gravity as well as dark matter and dark energy—each of which has strong experimental evidence for its existence. Additionally, the allowed mechanisms by which the SM can produce an asymmetry between matter and antimatter are not sufficient to explain the apparent dominance of matter over antimatter observed in the universe.

Beyond these deficiencies, the SM also has several peculiar features that point to possible issues with its theoretical integrity. Experiments trying to measure the electric dipole moment of the neutron place strong limits that indicate that CP violation is extremely small, if not nonexistent, in the strong sector [18]. However, there is no theoretical constraint on QCD imposed by the SM that requires that CP symmetry be conserved. Such apparent fine tuning suggests some possible, as-of-yet unknown structure that would more naturally impose such a constraint. In the quantum corrections that enter the computation of the Higgs boson mass, contributions scale up to the order of the Planck scale ($\sim 10^{19}$ GeV), beyond which the effects of gravity begin to compete with those of the SM, and the assumptions of the effective field theory break down. The size of these contributions suggest that the mass of the Higgs boson should be of a similar magnitude, but since its mass is observed to

be immensely smaller, similar hints of unnatural fine tuning are suggested by the intricate cancellations seemingly required for these corrections to settle on the observed mass [19–23]. This issue is known as the hierarchy problem due to the dramatic difference in scales, which establishes a hierarchy that presents issues when considering the interplay of these scales. In light of the compelling search for a so-called grand unified theory in which, at some high energy, all gauge interactions become a single force, it may also be considered unappealing that the SM's gauge couplings seem not to unify at some higher energy scale.

In addition to these unexplained phenomena and theoretical puzzles, there are experimental hints that may present direct challenges to the predictions of the SM. These include measurements of the anomalous magnetic moment of the muon and deviations from lepton universality in B meson decays, both of which disagree with the SM predictions—though not yet at the consensus five standard deviation level required for a claim of observation [24, 25].

Many theories introducing new physical structures beyond the SM (BSM) have been proposed to account for the apparent deficiencies and discrepancies of the SM. Many of these theories do so in ways motivated to some extent by the principle of naturalness in conjunction with limitations imposed by experimental searches for new phenomena. None have thus far made predictions that have been experimentally verified. However, there remains a great deal of unexplored phase space for BSM physics that is still built upon physically and mathematically well-motivated foundations. In Chapter 2, one of these theories is presented that resolves many SM issues and presents as-of-yet untested predictions for signatures that may be observed in modern experiments. A particular signature motivated in part by this type of theory will be explored in the course of this thesis.

Chapter 2: Supersymmetry

Theories of supersymmetry (SUSY) offer solutions to many SM problems laid out in Chapter 1. In this chapter, the foundations of SUSY theories are discussed with particular emphasis on gauge-mediated SUSY breaking (GMSB) models. Some GMSB models allow for the production of long-lived heavy SUSY particles, or sparticles, which can then decay to produce the signatures of the type investigated in this analysis—namely, delayed and non-pointing photons. Particular emphasis is placed on a specific, well-motivated model focused on the production of these SUSY particles from decays of the Higgs boson.

2.1 Foundations

SUSY posits a new symmetry relating fermions and bosons that is generated by an operator that transforms fermionic states into bosonic ones and vice versa [26–32]. Such an operator constructs so-called supermultiplets that form irreducible representations, with each SM fermion pairing with a supersymmetric bosonic partner and each SM boson with a supersymmetric fermionic partner. It can be shown that, in order to produce a consistent algebra that allows for the parity violation observed in the SM, particles in the same supermultiplet pair must have both the same mass and the same quantum numbers under all SM symmetries [33, 34].

It can further be demonstrated that the number of degrees of freedom of the fermions and bosons comprising a supermultiplet must be equal. In standard, realistic SUSY theories[1], such a constraint defines the nature of particles in these supermultiplets. Below, a

[1]Namely, so-called non-extended SUSY theories that have only one pair of generators. Adding more generators is mathematically valid, but such models in four dimensions cannot produce the maximal parity violation observed for SM fermions [35]. Theories proposing extra dimensions can circumvent this obstacle, but such theories are not considered in this thesis.

particular, convenient choice of properties is laid out, but any other choices that preserve the relation of the number of degrees of freedom between fermions and bosons and that allow for renormalizability may be reduced to combinations of the structures presented here.

Given the constraints and caveats discussed above, there are three possible supermultiplets. The first configuration is known as a chiral supermultiplet and contains a spin-1/2 fermion and a complex scalar field. The second is known as a gauge supermultiplet and contains a spin-1 gauge boson and a spin-1/2 fermion. For gauge supermultiplets, it is worth noting that, since gauge bosons transform under the adjoint representation of the gauge group, the spin-1/2 fermion in this supermultiplet must also do so [35]. And since the adjoint representation is its own conjugate, right- and left-handed components of the fermionic fields must transform identically under gauge transformations. The third configuration allows for the addition of spin-2 graviton with a spin-3/2 superpartner known as a gravitino.

Given that all SM fermions are maximally parity violating and that fermions in gauge supermultiplets forbid such a property, if SUSY were to be realized, SM fermions must belong to chiral supermultiplets. Thus, the superpartner of each SM fermion is a supersymmetric scalar boson. Conversely, the scalar SM Higgs boson also belongs to a chiral supermultiplet. But here, in order to avoid a gauge anomaly that would render the theory inconsistent, there must in fact be two Higgs supermultiplets with $Y = \pm 1/2$ [35]. Additionally, only the couplings to up-type quarks provided by the $Y = 1/2$ Higgs supermultiplet can endow them with mass. Likewise, only the $Y = -1/2$ Higgs supermultiplet couplings can give mass to the down-type quarks and the charged leptons. These $Y = 1/2$ and $Y = -1/2$ supermultiplets have associated bosons denoted as H_u^+, H_u^0 and H_d^0, H_d^-, respectively. Finally, the only room for a gauge boson is in a gauge supermultiplet, so the SM gauge bosons each have a supersymmetric spin-1/2 fermion superpartner. The chiral and gauge supermultiplets just discussed corresponding to all of the SM particles comprise the particles of the Minimal Supersymmetric Standard Model (MSSM) [36, 37]. The notation and basic properties of these particles are summarized in Table 2.1. For generality, the SM gauge bosons are represented before EW symmetry breaking as the three bosons W^+, W^0, and W^- associated with the $SU(2)_L$ fields; and B^0, the vector boson associated with $U(1)_Y$.

Chiral Supermultiplets						
				Quantum Numbers		
Name	**Spin 0**	**Name**	**Spin 1/2**	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
Squarks (x3 generations)	$(\widetilde{u}_L\ \widetilde{d}_L)$	Quarks	$(u_L\ d_L)$	3	2	1/6
	$\widetilde{u}_R^*$		$u_R^\dagger$	3	1	-2/3
	$\widetilde{d}_R^*$		$d_R^\dagger$	3	1	1/3
Sleptons (x3 generations)	$(\widetilde{\nu}\ \widetilde{e}_L)$	Leptons	$(\nu\ e_L)$	1	2	-1/2
	$\widetilde{e}_R^*$		e_R^*	1	1	1
Higgs Bosons	$(H_u^+\ H_u^0)$	Higgsinos	$(\widetilde{H}_u^+\ \widetilde{H}_u^0)$	1	2	1/2
	$(H_d^0\ H_d^-)$		$(\widetilde{H}_d^0\ \widetilde{H}_d^-)$	1	2	-1/2
Gauge Supermultiplets						
				Quantum Numbers		
Name	**Spin 1/2**	**Name**	**Spin 1**	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
Gluino	$\widetilde{g}$	Gluon	g	8	1	0
Winos	$\widetilde{W}^\pm\ \widetilde{W}^0$	W Bosons	$W^\pm\ W^0$	1	3	0
Bino	$\widetilde{B}^0$	B Boson	B^0	1	1	0
Gravity Supermultiplet						
				Quantum Numbers		
Name	**Spin 2**	**Name**	**Spin 3/2**	$SU(3)_C$	$SU(2)_L$	$U(1)_Y$
Graviton	G	Gravitino	$\widetilde{G}$	1	1	0

Table 2.1: Summary of the naming conventions and quantum numbers of the particles in the MSSM chiral and gauge supermultiplets as well as the graviton supermultiplet.

There remains an additional, as of yet unmentioned symmetry of the traditional MSSM. Recall that in the SM, baryon number B and lepton number L are conserved. Such a conservation prevents such processes as proton decay, for which there are high experimental limits on the lifetime [38]. This conservation is due to the fact that no renormalizable terms that violate B or L can be added to the SM Lagrangian. However, by introducing SUSY, new terms that do violate the conservation of these quantities can be introduced. Thus, it is conventional to impose a new symmetry called R-parity—a quantity defined as:

$$P_R = (-1)^{3(B-L)+2s} \tag{2.1}$$

where s is the particle's spin [39]. All SM particles have even R-parity ($P_R = +1$), and all SUSY particles have odd R-parity ($P_R = -1$). R-parity conservation has a number of key effects on the types of interactions permitted in the MSSM. First, mixing of SM and SUSY particles is prohibited. Second, the lightest supersymmetric particle (LSP) is stable, i.e. SUSY particle decays must always result in an odd number of LSPs. And thirdly, in considering the production of SUSY particles from SM ones (such as what would be done at the LHC), an even number of SUSY particles must be produced.

2.2 SM Limitations and SUSY

Now that the general mathematical structure of SUSY has been defined, some of its attractive features can be appreciated. In Section 1.5, several problems with the SM were outlined. Some of these issues will now be reexamined in the context of SUSY.

As previously discussed, the SM seems to require incredible fine tuning in the quantum corrections to the Higgs boson mass from radiative loop diagrams. However, the introduction of SUSY implies that each of these diagrams now comes with a new contribution containing the supersymmetric particle differing in spin by 1/2. Since fermions and bosons of equal mass contribute equally in magnitude but with opposite signs to such diagrams, each of these corrections conveniently cancels out, eliminating the need for fine tuning to deal with enormous, Planck-scale terms to sum up to the EW scale. Thus, SUSY can, in a very structurally satisfying way, solve the hierarchy problem [40–43].

The introduction of R-parity into SUSY results in a stable LSP that cannot decay to SM particles. If the LSP is not electrically charged, it would only interact weakly with SM particles. Such properties are considered likely features of dark matter— a subject about which the SM is silent. Thus, SUSY can fairly naturally provide a viable dark matter candidate.

Finally, the running of the gauge couplings is modified by the introduction of SUSY. It can be shown that in the MSSM, the three gauge couplings corresponding to $SU(3)$, $SU(2)$, and $U(1)$ may now converge at some high energy, renewing hope that the forces may indeed be unified [44–48].

2.3 Breaking of SUSY

As presented thus far, SUSY is an unbroken symmetry, a consequence of which is that superpartners of SM particles have the same mass and internal quantum numbers as SM particles. However, it is clear experimentally that no such particles exist, as they would be readily detectable. Therefore, if SUSY is in fact a true symmetry of nature, it must be a broken symmetry.

It can be shown that unbroken SUSY automatically ensures that the dimensionless couplings in the quantum corrections to the Higgs boson mass discussed in Section 2.2 are related such that no quadratic divergences occur [35]. However, if SUSY is broken, in order to maintain this desirable property, the concept of so-called soft SUSY breaking will now be introduced [49]. In soft SUSY breaking, the SUSY Lagrangian may be separated into two terms: one containing all the gauge and Yukawa interactions of SUSY and another with SUSY-violating terms but all of which have positive mass dimension. Such a requirement on the SUSY-violating term is not excessively contrived, as it is in fact satisfied by many viable theoretical models, including GMSB, as will be discussed in Section 2.4. It can in fact be shown that, with such a constraint, the corrections to the Higgs boson mass have no quadratic divergences and can be written in the following form:

$$\Delta m_H^2 = m_{\text{soft}}^2 \left(\frac{\lambda}{16\pi^2} \ln \left(\frac{\Lambda_{\text{UV}}}{m_{\text{soft}}} \right) + \dots \right) \tag{2.2}$$

where m_{soft} is the largest mass scale in the SUSY-violating part of the Lagrangian, λ represents the different possible couplings (combined into one variable for illustrative purposes), and Λ_{UV} is the ultraviolet (UV) cutoff. The UV cutoff occurs at the energy scale at which new physics emerges. For SUSY models, it is assumed that this would be near the Planck scale.

In order for SUSY to still provide the desired cancellations, Equation 2.2 requires that m_{soft} not be excessively large compared to the EW breaking scale ($\sim$174 GeV), lest they dominate the SM corrections and introduce the need for fine tuning. In fact, for many of the most mathematically sound and natural SUSY models with seemingly reasonable assumptions, m_{soft} is estimated to be around the TeV scale. Thus, should these models prove correct, the lightest SUSY particles should be of this scale and therefore potentially observable at the energies attained at the LHC.

One other critical phenomenological feature of SUSY breaking is that the states detailed in Table 2.1 are not necessarily the mass eigenstates that would be observed. In a manner analogous to the mixing that occurs after EW symmetry breaking, after the breaking of both EW symmetry and SUSY, mixing may occur between the gauginos (winos and bino) and the higgsinos and among the different squarks, sleptons, and the neutral Higgs bosons[2]. This full resulting set of mass eigenstates in the MSSM will now be briefly discussed.

EW symmetry breaking as discussed in Section 1.2.2 involved a single complex scalar Higgs doublet. With SUSY breaking, a second Higgs doublet is introduced, doubling the number of degrees of freedom from four to eight. The resulting mixing proceeds in a manner similar to that of EW symmetry breaking in the SM with three massless Goldstone bosons produced from the non-zero VEV being absorbed into the longitudinal components of the now massive Z and $W^{\pm}$ bosons and leaving a massless photon. But now there are five remaining degrees of freedom from the Higgs potentials instead of one. After mixing among the Higgs gauge eigenstates, this results in two CP-even neutral scalars h^0 and H^0 (with

[2]Is is interesting to note that while mass terms for SM fermions and the photon and gluon are prohibited by symmetries, this is not true for squarks, sleptons, or Higgs bosons since mass terms for scalars are not prohibited by any symmetry. For higgsinos and gauginos, which are in a real representation of the gauge group (unlike SM fermions), there is likewise no prohibition. All MSSM particles then already have mass before EW symmetry breaking and interaction with the Higgs field. Therefore, it is reasonable to expect that all MSSM particles would be heavier than SM particles, and—should the MSSM reflect reality—that all SM particles would be discovered first.

$m_{h^0} < m_{H^0}$ by convention), one CP-odd neutral scalar A^0, and two oppositely-charged scalars $H^{\pm}$.

Similarly for SUSY particles, EW symmetry breaking causes mixing between the higgsinos and gauginos. The neutral higgsinos ($\widetilde{H}_u^0$ and $\widetilde{H}_d^0$) and gauginos ($\widetilde{B}$ and $\widetilde{W}^0$) mix into four neutral mass eigenstates known as neutralinos labelled $\widetilde{\chi}_i^0$ ($i = 1, 2, 3, 4$ in order of increasing mass). The charged higgsinos ($\widetilde{H}_u^+$ and $\widetilde{H}_d^-$) and winos ($\widetilde{W}^+$ and $\widetilde{W}^-$) mix into two, oppositely-charged mass eigenstates known as charginos labelled $\widetilde{\chi}_i^{\pm}$ ($i = 1, 2$ in order of increasing mass).

As for the gluons, gluinos form a color octet of linearly independent color states. The carrying of a color charge prevents them from mixing with other MSSM particles. Thus, their gauge and mass eigenstates are identical.

Analogously to the SM, mixing can occur within the up-type squarks, down-type squarks, charged sleptons, and sneutrinos. The degree of mixing is characterized by three 6×6 matrices for the first three groupings and one 3×3 matrix for the sneutrinos (due to the lack of right-handed (s)neutrinos). In order to limit the contributions of virtual squarks and sleptons to the observed rates of flavor-changing processes in the SM, the mixing parameters among the first- and second-generations are generally taken to be negligible [35]. Mixing among the left- and right-handed third-generation squarks and sleptons, however, is not similarly constrained.

In an attempt to incorporate gravity into the theory, SUSY—whose breaking thus far has been considered to be a global symmetry—it becomes necessary to promote SUSY to a local symmetry. This local SUSY theory known as supergravity has a gravity supermultiplet (see Table 2.1) with a spin-2 graviton and spin-3/2 gravitino, which are both massless before SUSY breaking [50, 51]. Thus far, the breaking of SUSY has been considered explicitly as a necessity to relieve tension with empirical evidence. The fundamental nature of soft SUSY breaking as discussed may be considered to generally emerge from the structure of the spontaneous breaking of SUSY. In such a case, the spontaneous breaking of SUSY implies the production of a massless Goldstone fermion known as a goldstino. After SUSY breaking, the gravitino absorbs the goldstino into its longitudinal components and acquires mass.

A summary of all the new particles introduced by the MSSM in their mass eigenstates

and their corresponding gauge eigenstates are shown in Table 2.2 following the discussion and assumptions presented above.

Name	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs Bosons	0	+1	$H_u^0\ H_d^0\ H_u^+\ H_d^-$	$h^0\ H^0\ A^0\ H^\pm$
Squarks	0	-1	$\widetilde{u}_L\ \widetilde{u}_R\ \widetilde{d}_L\ \widetilde{d}_R$	same
			$\widetilde{s}_L\ \widetilde{s}_R\ \widetilde{c}_L\ \widetilde{c}_R$	same
			$\widetilde{t}_L\ \widetilde{t}_R\ \widetilde{b}_L\ \widetilde{b}_R$	$\widetilde{t}_1\ \widetilde{t}_2\ \widetilde{b}_1\ \widetilde{b}_2$
Sleptons	0	-1	$\widetilde{e}_L\ \widetilde{e}_R\ \widetilde{\nu}_e$	same
			$\widetilde{\mu}_L\ \widetilde{\mu}_R\ \widetilde{\nu}_\mu$	same
			$\widetilde{\tau}_L\ \widetilde{\tau}_R\ \widetilde{\nu}_\tau$	$\widetilde{\tau}_1\ \widetilde{\tau}_2\ \widetilde{\nu}_\tau$
Neutralinos	1/2	-1	$\widetilde{B}^0\ \widetilde{W}^0\ \widetilde{H}_u^0\ \widetilde{H}_d^0$	$\chi_1^0\ \chi_2^0\ \chi_3^0\ \chi_4^0$
Charginos	1/2	-1	$\widetilde{W}^\pm\ \widetilde{H}_u^+\ \widetilde{H}_d^-$	$\chi_1^\pm\ \chi_2^\pm$
Gluino	1/2	-1	$\widetilde{g}$	same
Gravitino	3/2	-1	$\widetilde{G}$	same

Table 2.2: Summary of the gauge and mass eigenstates of the non-SM particles in the MSSM including the gravitino.

2.4 Gauge-Mediated SUSY Breaking

In GMSB models, soft SUSY breaking is realized via new gauge interactions resulting from the introduction of new chiral supermultiplets which couple to both the source of SUSY breaking and the MSSM particles [52–54]. The members of the new chiral supermultiplets are known as messengers, as they communicate the presence of SUSY breaking to the so-called visible sector of the MSSM via the familiar $SU(2) \times SU(2))L \times U(1)_Y$ gauge and gaugino interactions[3]. While the messengers themselves reside at the higher mass scale somewhere below the SUSY breaking scale, they provide mass to the gauginos and scalar sparticles via one- and two-loop radiative corrections, respectively.

One theoretically appealing aspect of GMSB models lies in the built-in degeneracy of squarks and sleptons—a property that limits the contribution of flavor-changing processes

[3]GMSB models modify the running of the gauge couplings by the addition of a new scale for the messengers. However, the unification of gauge couplings achieved by the general MSSM discussed in Section 2.2 is maintained as long as the messengers both form complete multiplets of the $SU(5)$ global symmetry of the unified theory that contains the gauge group of the SM and do not vary excessively in mass.

that are not observed. Another interesting phenomenological aspect of GMSB models is that generally the gravitino is the LSP since it only interacts gravitationally, as opposed to other models in which supergravity facilitates the breaking of SUSY. Assuming R-parity conservation, all SUSY decays are then expected to eventually results in stable gravitinos. GMSB models may be generalized in a number of different ways including arbitrary messenger masses and messenger supermultiplets that do not form complete multiplets in an $SU(5)$ unified theory [55]. However, the gravitino LSP and degeneracy of squarks and sleptons remain general features of such models.

In many of these models the lightest neutralino is the next-to-lightest supersymmetric particle (NLSP). Since this neutralino contains some (possibly dominant) bino component, it could then decay to a photon and a gravitino. Another unique feature of this particular configuration is that the lifetime of the neutralino NLSP is essentially a free parameter in the model, as it depends directly on the SUSY breaking scale [56]. Therefore, with sufficiently high SUSY breaking scale, the neutralino becomes macroscopically long-lived, and the photons produced in its decay would be significantly displaced from the neutralino production vertex. Conversely, observation of such unusual photons would provide a key window into the nature of SUSY breaking. In a collider experiment such at ATLAS, heavy neutralinos produced at the IP travelling at velocities significantly less than the speed of light and producing photons that recoil off the the gravitino would result in photons that arrive at the calorimeter delayed and with their showers not pointing back to the IP. In this case, it is possible that nearly all SUSY decays would result in photons that are displaced in both space and time in an empirically measurable manner.

2.5 Long-Lived Particles

Before introducing the concept of long-lived particles in the context of SUSY, it is worth noting that many SM particles are relatively long lived. This property arises due to various reasons inherent to the nature of the SM, including hierarchies of scale, small mass splittings, and weak couplings.

A hierarchy of energy scales—manifested as large mass differences between particles

and the mediating gauge bosons—leads to decays that are highly off shell and thus highly suppressed in their decay rates, leading to significant longevity. For example, the only decay mode of the charged pion ($\pi^{\pm}$) allowed in the SM is via the W boson of the weak interaction, while the neutral pion (π^0) decays via the electromagnetic interaction. The difference in scales between QCD and the weak interaction ($\Lambda_{\text{QCD}} \ll M_W$) establishes a hierarchy that suppresses the decay rate of the $\pi^{\pm}$. But the π^0, interacting via the massless photon, has no such dramatic suppression. Accordingly, the lifetime of the $\pi^{\pm}$ is measured to be about 26 ns—more than 9 orders of magnitude longer than that of the π^0 [38].

Small mass splittings can likewise suppress decay rates, which can be thought of as due to the pseudo-symmetries they imply. For example, consider the decay of the free neutron. This decay can only occur via a weak interaction, as for the $\pi^{\pm}$. But in addition, the low momentum transfer due to the isospin symmetry (tied to the near-degeneracy of the neutron and proton) suppresses this decay rate even further such that its lifetime is about 879 s [38].

The coupling strength also directly influences the decay rate. The bottom quark may only decay via the weak interaction, and it couples overwhelmingly to the top quark (as measured empirically), and its couplings to the up and charm quarks are accordingly very small. However since $m_b \ll m_t$, only decays to up and down quarks are kinematically allowed. Thus, the couplings that govern the bottom quark's decay rate are very small, and as a result, b-hadrons have relatively long lifetimes (of order picoseconds). In fact, this phenomena allows for so-called b-tagging in which b-hadrons can be distinguished from other quark decay products partly by their displaced decay relative to the interaction point at collider experiments.

The above reasons—hierarchies of scale, small mass splittings, and weak couplings—were discussed in the context of SM particles, but for the same phenomenological reasons, BSM particles may also be long-lived. Specifically in the context of SUSY, potentially heavy sparticles are posited as mediators of lighter BSM interactions, establishing a hierarchy of scales. Similarly, the coupling strengths and potential mass splittings between particles in many models are often not strictly constrained and may thus be quite small. It is therefore well-motivated theoretically to consider the possibility of non-prompt signatures of BSM physics.

2.6 Exotic Higgs Boson Decays

As discussed in Section 1.2.2, the Higgs boson couples to mass and consequently, from an experimental perspective, offers a compelling window to hypothetical new heavy particles. It is possible that, because of this direct coupling, the first hints of BSM physics will come in the form of Higgs-mediated signatures. Dedicated searches for Higgs boson decays to invisible particles have been carried out in order to determine the limits of how completely the Higgs boson branching ratio can be detected given the uncertainties in measuring $E_{\mathrm{T}}^{\mathrm{miss}}$. Such a search recently completed by ATLAS set the 95% CL upper limit on the branching ratio of the Higgs boson to invisible particles at 11% [57]. Such a relatively weakly constrained branching ratio leaves a significant amount of available phase space and thus remains a promising potential source of BSM physics still reachable by the LHC.

This thesis specifically targets this open phase space of Higgs decays with displaced photons in the final state from some intermediate long-lived particle. While the goal is to remain as model-independent as possible, a specific signal model is employed in order to investigate the sensitivity to SUSY models with long-lived particles of the type described in this section and Section 2.5. Namely, the model involves the exotic decay of the SM Higgs boson to two long-lived neutralino NLSPs, which each subsequently decay to a photon and an LSP. This LSP may either be the nearly massless gravitino LSP of a GSMB-like model[4] or another, lighter neutralino. A range of LSP masses are considered in this analysis in order to more fully explore different potential kinematic scenarios and thus to remain more model-independent[5]. Since there are four final decay products from the 125 GeV Higgs boson, the photons produced are expected to be relatively soft. Given that the abundance of SM background processes that produce photons increases dramatically with decreasing energy, it is untenable to trigger on events with extremely soft photons due to the limits of the readout capabilities of the detector. Therefore, leptons produced from associated Higgs

[4]Traditional GMSB models have $m_{\tilde{G}} \lesssim 1$ keV [35].

[5]Different models exist in which the gravitino is more massive and the lightest neutralino is the LSP, including so-called next-to-minimal supersymmetric models (NMSSMs) [58–60]. However, we do not incorporate any additional, specific model constraints in our evaluation of the sensitivity. We simply explore the full space of Higgs boson decays to long-lived NLSPs, which then decay to photons and LSPs of any kinematically available masses.

boson production are used for triggering—namely, Higgs bosons produced in association with W bosons, Z bosons, or $t\bar{t}$ that contain an electron or muon in their decays. An example of a Feynman diagram for these processes is illustrated in Figure 2.1.

The parameter space of m_{NLSP}, m_{LSP}, and τ_{NLSP} is explored with m_{NLSP} and $m_{\text{LSP}} < m_{\text{NLSP}}$ varying from 0.5 to 60 GeV (i.e. $\sim m_H/2$ since two NLSPs are produced from the 125 GeV Higgs boson). As discussed in Section 2.4, the lifetime τ_{NLSP} is essentially a free parameter as it depends on the SUSY breaking scale. Since τ_{NLSP} can thus, in principle, be arbitrarily large, it is possible that neutralinos produced at the IP would decay outside the LAr calorimeter, leaving the photons undetected. Therefore, despite the requirement (by R-parity conservation) that two neutralinos and thus two photons be produced, the final state measured in the detector may include either one or two photons.

It may seem that only a very fine-tuned model would allow for τ_{NLSP} in the narrow window of $0.1 \lesssim c\tau \lesssim 1.5$ m such that the photons would both be sufficiently displaced to be distinguished from those resulting from background processes and be produced before the LAr calorimeter. However, due to the exponential nature of particle decay, despite large lifetime values, a substantial fraction of decays still may occur in this range, as illustrated in Figure 2.2. Thus, as will be demonstrated, the present search has significant sensitivity to a relatively wide range of NLSP lifetimes.

Searches in the prompt regime for such exotic Higgs boson decays using one and/or two photons relying on $E_{\text{T}}^{\text{miss}}$ cuts have found no excesses [62–64]. Long-lived searches have also been carried out by both ATLAS and CMS with SM agreement observed in all [65–67]. However, these searches cover a different corner of the phase space, targeting a specific benchmark GMSB model known as the Snowmass Points and Slopes 8 (SPS8). SPS8 is generally more constrained than more modern models of general gauge mediation, as discussed in Section 2.4. For example, in SPS8, the mass scales for gluinos, squarks, and the bino are all the same. More generalized GMSB models grant greater flexibility by allowing for the strong production of heavy sparticles.

As a consequence of specifically targeting exotic Higgs boson decays and of achieving more model flexibility by moving away from the SPS8 model, the resulting photons are significantly softer than those examined in previous displaced photon searches. The present

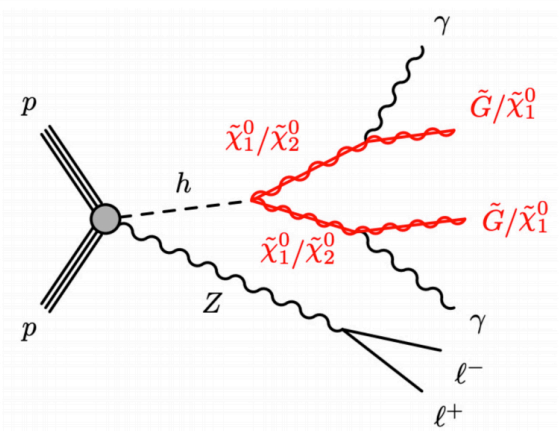


Figure 2.1: Feynman diagram of the process examined in this thesis. Namely, the exotic decay of the Higgs boson to two long-lived neutralinos that each then decay to a gravitino and a displaced photon. For triggering purposes, the Higgs boson must be produced in association with a W boson, Z boson, or $t\bar{t}$ that contains an electron or muon in its decay.

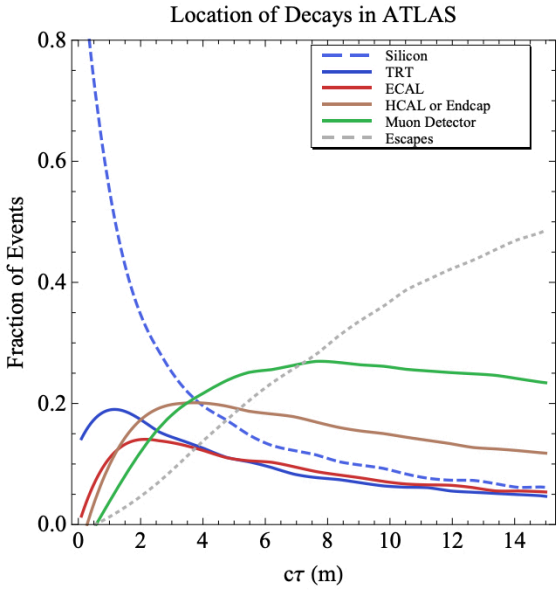


Figure 2.2: Fraction of decays that occur in different subdetectors of ATLAS as a function of χ_1^0 decay length. A model with $m_{\chi_1^0} = 250$ GeV produced from 600 GeV gluinos was used, but the impact of these specific model parameters is small compared to the qualitative dependencies due to the detector geometry and decay length itself [61].

search thus occupies an unexplored corner of the phase space in the context of a more widely interpretable signal model.

Chapter 3: The Large Hadron Collider

The Large Hadron Collider (LHC) is a high-energy particle collider with both proton ($p-p$) heavy ion ($A-A$) modes that delivers collisions with a center of mass energy up to $\sqrt{s} = 13$ TeV at four interaction points (IPs). Since this thesis concerns data from $p-p$ collisions, the description of the LHC complex will be from that perspective, and ion collisions will not be discussed further. The basics of accelerator physics and the design details that allow for the operation of such a machine are outlined below.

3.1 Accelerator Physics

The full LHC complex contains a wide variety of specialized devices that take gaseous hydrogen and produce two, focused, colliding 6.5 TeV proton beams. The main components to be discussed are the radiofrequency (RF) cavities, dipole magnets, and multipole magnets. These three elements make up the primary functional elements of the linear accelerators and synchrotrons of the LHC complex.

RF cavities are specialized devices that produce oscillating electric fields at specific frequencies and locations such that charged particles can be optimally accelerated. The frequency of these RF cavities creates so-called RF buckets in the wave troughs in which protons accumulate and form what are called bunches [68]. A proton arriving slightly out of sync with the characteristic frequency will experience a relative force towards the center of the nearest bucket and thus a stable equilibrium is maintained. In the LHC, during nominal running, each proton bunch contains about 10^{11} protons.

Dipole magnets are employed in the various synchrotrons of the LHC complex to bend protons in order to maintain them in orbit around the machine. These magnets generally

aim to produce a constant, vertical magnetic field such that a centripetal force is exerted on the circulating protons. As RF cavities ramp up the energies of the protons, the strength of the dipole field are simultaneously increased in order to maintain the same orbit.

Multipole magnets are used to focus and fine-tune the optical properties of the beam. The main multipole magnet type used at the LHC is the quadrupole magnet, which can be arranged to squeeze the beam tightly together to optimize collision conditions. Many more complex, higher-order multipole magnets are also employed for higher-order effects on beam optics. As a side effect of their focusing action, multipole magnets induce so-called betatron oscillations in the transverse direction. The magnitude of these oscillations are described by the beta function β. The basic principles of the optical properties of colliding beams are discussed below, which further illustrates the importance of the focusing characteristics afforded by the various multipole magnets.

The goal of a particle collider like the LHC is to deliver the maximum number of collision events, which can be represented by the following simple equation:

$$N_\alpha = \sigma_\alpha \int \mathcal{L}(t)dt \tag{3.1}$$

where N_α is the expected number of events for a given process α, σ_α is the cross-section for α, and $\mathcal{L}$ is the instantaneous luminosity. Cross sections represent probabilities of specific processes during collisions and thus depend on the nature of the physical processes involved. Luminosity is proportional to the total number of collisions that occur and depends on many beam parameters under the control of the operator. Thus, along with maximizing the energy to investigate new, unexplored physical regimes, maximizing luminosity is considered the primary goal of an effective collider.

The luminosity may be expressed in the following manner:

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi\varepsilon_n \beta^*} F \tag{3.2}$$

where N_b is the number of particles per bunch, n_b is the total number of bunches per beam, f_{rev} is the revolution frequency, γ_r is the relativistic gamma factor, ε_n is the normalized transverse emittance, β^* is the beta function at the IP, and F is the geometric luminosity

reduction factor [69]. The transverse emittance is a measure of the spread of the beam in the transverse direction taking into account both the RMS of the transverse beam width σ and the betatron oscillations in that direction:

$$\varepsilon \equiv \pi \frac{\sigma^2}{\beta} \tag{3.3}$$

with normalized emittance $\varepsilon_n = \beta_r \gamma_r \varepsilon$. F effectively reduces the luminosity due to the fact that the beams do not collide head-on but rather at a crossing angle θ_c at the IP:

$$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)\right)^{-1/2} \tag{3.4}$$

where σ_z is RMS of the bunch length and σ^* is the RMS of the transverse beam width at the IP.

Equations 3.1–3.4 make clear that, in order to maximize the number of observed events for collisions at a given energy, the beams must contain many, densely-populated bunches that are focused as tightly as possible in the transverse direction and, to a lesser extent, in the longitudinal direction. Such are the goals of the various components of the LHC complex and the multipole magnets that finely tune the beam parameters to optimize the luminosity delivered at the IPs.

A precise measurement of the luminosity is critical to nearly all physics activities associated with the LHC since this quantity informs the expected frequency with which each physics process—SM or not—is observed. Computing the luminosity values from the quantities in Equation 3.2 turns out to not be particularly precise, mainly due to relatively large uncertainties on β^* and other complex beam-beam interactions. Instead, special van der Meer scans are performed in which the beams are separated from each other in the transverse direction, and the resulting event rates are measured as a function of this separation [70]. The absolute luminosity can then be computed based on the rate of inelastic scatters R_{inel} and the inelastic cross section σ_{inel} as follows[1]:

[1]Under the assumption that the beam densities in the x- and y-directions do not contain any complex, non-separable interdependencies.

$$\mathcal{L} = \frac{R_{\text{inel}}}{\sigma_{\text{inel}}} = \frac{n_b f_{\text{rev}} n_1 n_2}{2\pi\sigma_x\sigma_y} \tag{3.5}$$

where n_1 and n_2 are the populations of each beam, and σ_x and σ_y are the standard deviations of the spread of event rate vs. beam separation in the x- and y-directions, respectively.

3.2 LHC Design

The LHC accelerator ring lies within the previously existing tunnel used for the electron-positron-colliding LEP experiment near Geneva, Switzerland on the border with France [69]. The approximately circular underground tunnel is about 26.7 km in circumference and consists of eight short straight sections (∼150 m long) and eight arcs connecting them. Two beams circulate in opposite directions around the ring and are then steered into each other at the IPs. The eight straight sections contain various beam services such as beam cleaning, injection, dump, and RF acceleration, as well as four experimental caverns at which beam collisions occur and the resulting products are studied. The arcs contain about 1,200 superconducting dipole magnets that bend the protons along the arcs, about 400 superconducting quadrupole magnets that focus the beam into concentrated bunches, and many other secondary magnets that perform various tasks for ensuring optimal beam optics.

A novel, compact, twin-bore superconducting dipole magnet system was designed specifically for the LHC, which provides two oppositely directed magnetic fields for the two beams while sharing a single cold vessel. The magnets are coupled for the two beams, reducing flexibility for the optimization of beam conditions, but their design is relatively compact (a requirement for the restrictive 3.7 m tunnel size) and cost-effective. Niobium-titanium superconducting cables are cooled to 1.9 K using superfluid helium to produce the magnetic field of about ∼8 T to maintain 6.5 TeV protons in orbit.

The chain of the major machines used by the LHC for proton injection is illustrated schematically in Figure 3.1 and described briefly below. An electric field is applied to gaseous hydrogen, stripping the electrons and leaving the bare protons. These protons are then guided by focusing and bending magnets into a linear accelerator known as Linac 2, which uses RF cavities to accelerate groups of protons up to an energy of 50 MeV. The

beams are then sent through a series of four synchrotrons: the Proton Synchrotron Booster, the Proton Synchrotron, the Super Proton Synchrotron, and finally, the LHC itself. These machines accelerate the protons up to energies of 1.4 GeV, 25 GeV, 450 GeV, and 6.5 TeV, respectively.

In the LHC, the RF cavities operate at a frequency of 400 MHz, corresponding to 2.5 ns spacings between RF buckets. Buckets are organized around the full LHC ring to allow for 25 ns between bunches with empty buckets in between[2]. This then corresponds to a bunch crossing frequency of 40 MHz—the data rate at which all detectors at the LHC were designed to collect data.

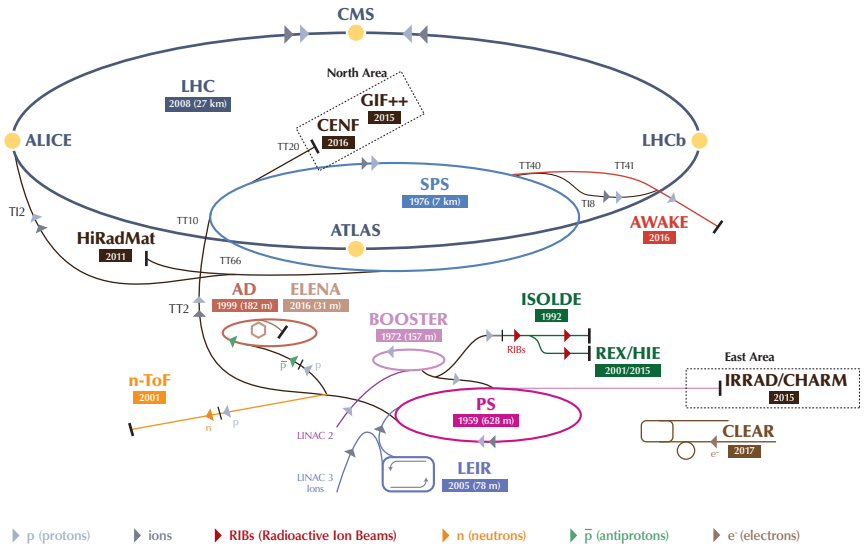


Figure 3.1: Diagram of the LHC complex illustrating the full beam injection chain [71].

The full procedure from Linac 2 to injection of oppositely-circulating bunches into the LHC is repeated multiple times until the desired beam intensity and bunch configuration is achieved. Then, after the beams are ramped up the nominal 6.5 TeV each, focusing magnets near the four IPs optically adjust and squeeze the beams and prepare them for collision. The beams are then carefully bent until they cross at the desired location. After final adjustments are made, the declaration of stable beams is made, and while conditions

[2]To be precise, the RF cavity frequency, bunch crossing frequency, and bunch spacings are in fact 400.8 MHz, 40.08 MHz, and 24.95 ns, respectively. However, for simplicity, in the rest of this thesis, 400 MHz, 40 Hz, and 25 ns will be used.

remain nominal, collisions between proton bunches proceed for several hours, and the various experiments collect data. Depending on beam conditions and configuration, $\sim 10-70$ proton-proton interactions occur during each bunch crossing at the ATLAS IP (see more details in Section 3.3). This number of interactions per bunch crossing is known as pileup. As collisions occur, the bunches slowly de-populate over time until it is more efficient to dump the beams and re-fill than to continue circulating beams of such intensity. The cycle from LHC injection to beam dump is depicted schematically in Figure 3.2. This entire process is then repeated constantly, day and night, during scheduled operation of the machine.

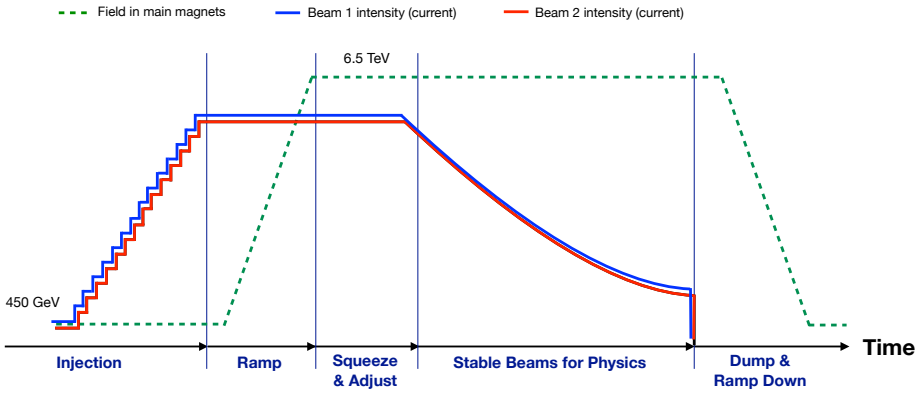


Figure 3.2: Diagram illustrating the typical operation of the LHC beam [72].

3.3 Run 2 Conditions & Operations

The LHC was designed to operate up to a peak instantaneous luminosity of 2.1×10^{34} cm^{-2}s^{-1}. Due to improvements in beam operations, the need to adjust beam configurations to handle certain beam issues, and various other operational reasons, a range of different rates of integrated luminosities as well as pileup values and were observed during collisions throughout Run 2 (2015-2018, the period during which the data used for this thesis was collected). Figure 3.3 shows the cumulative integrated luminosity over time. It also includes the amount recorded by ATLAS (one of the detectors at the LHC, to be discussed in detail in Chapter 4) and the amount considered good for physics. The former is reduced compared to the delivered amount due to inefficiencies in data acquisition and ramping up of

inner detector systems after stable beams are declared (see Chapter 4). The latter requires the proper functioning of all subsystems, including many additional data quality checks. In total, about 139 fb^{-1} of good physics data was collected by ATLAS throughout Run 2 and will be considered in this thesis. Figure 3.4 illustrates the frequency of different pileup values observed by ATLAS. Generally, over time with constantly improving operations of the LHC and the detectors at the IPs, pileup was increased to allow for more data to be collected[3]. This presented significant data acquisition challenges as it significantly stressed the processing and filtering power of the detectors (see Section 4.7).

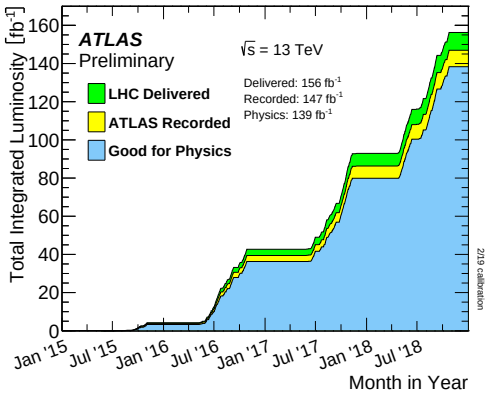


Figure 3.3: Cumulative integrated luminosity over time for all of Run 2 [73].

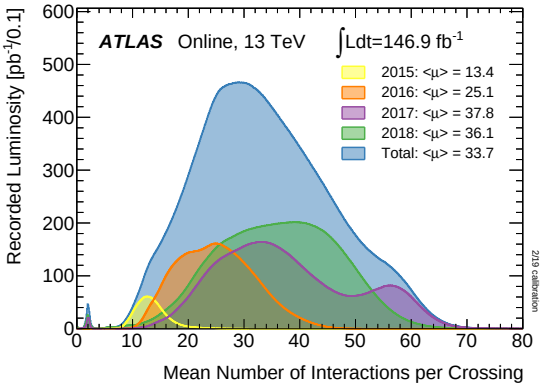


Figure 3.4: Mean number of $p-p$ interactions per bunch crossing weighted by luminosity throughout Run 2 and separated by year [73].

The next chapter will focus specifically on one detector at the LHC, namely the ATLAS experiment, which collected all of the data used in this thesis.

[3]The shoulder on the right side of the distribution for 2017 corresponds to a specific non-nominal bunch configuration that was used for much of the year to mitigate problems with the magnet system. Pileup was increased to compensate for the resulting loss in luminosity.

Chapter 4: The ATLAS Detector

The ATLAS detector is one of two general-purpose particle detectors at the LHC, the other being CMS. A detailed description of the detector, its design, installation, and expected performance is presented in Ref. [74]. A general summary of the detector and its subsystems follows, with a special emphasis on the liquid argon (LAr) calorimeter, which plays a central role in the analysis presented in this thesis.

4.1 Detector Overview

The ATLAS detector was designed to study the products of $p-p$ and $A-A$ collisions at energies previously unachieved at any particle collider. The overall goal of the detector is to measure, as precisely as possible, all products of such collisions while optimizing for the anticipated signatures of the Higgs boson and possible signatures of physics beyond the SM. Doing all of this in a high-radiation environment with many essentially simultaneous collisions occurring at a rate of 40 MHz presents a significant challenge that drove many decisions of the detector design.

The detector generally consists of different layers surrounding the beam in a cylindrical manner that make up subsystems that perform various detector functions. Before beginning to describe each subsystem, let us define a set of commonly-used terms and conventions. The intended IP for the collision of beams from the LHC is considered the origin of the coordinate system used for the ATLAS detector. The z-axis is defined to be aligned along the beam direction at the IP, the x-axis points to the center of the LHC ring, and the y-axis points vertically upward. The transverse direction, used for quantities such as the transverse momentum p_{T} and transverse energy E_{T}, is defined as the direction transverse

to the beam, i.e. the $x - y$ plane. This transverse direction carries particular importance since the momentum of the colliding beams is zero in this direction. Thus, the negative net transverse momentum in the detected collision products corresponds to so-called missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), which is a crucial variable in the study of BSM physics that may allow for the production of new particles that escape undetected. Using cylindrical coordinates, the azimuthal angle ϕ is measured from the x-axis around the z-axis, and the polar angle θ is measured from the z-axis.

A more convenient parameter to describe the polar displacement in the context of high energy physics is known as pseudorapidity (η), defined in Equation 4.1.

$$\eta \equiv \ln\left(\tan\frac{\theta}{2}\right) \tag{4.1}$$

The convenience of pseudorapidity over the polar angle lies in the Lorentz invariance of differences in pseudorapidity, as well as the fact that particles in hadron collisions are distributed roughly uniformly in pseudorapidity in the high-energy (massless) limit. Central, low-η regions of the detector are often described as the barrel, while outer, high-η regions are known as the end-caps.

A diagram illustrating the coordinate system presented is shown in Figure 4.1.

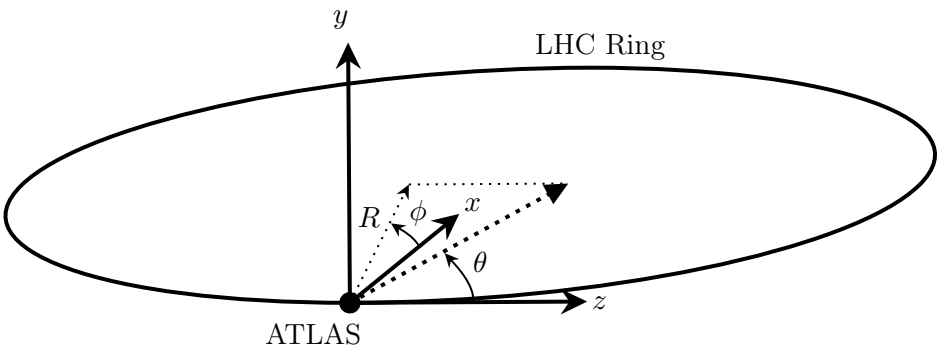


Figure 4.1: ATLAS coordinate system.

To describe angular distance in the $\eta - \phi$ plane, a quantity ΔR is defined in Equation 4.2.

$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{4.2}$$

An overall view of the ATLAS detector and its main subsystems can be found in Figure 4.2, and some key, general performance goals of the ATLAS detector are shown in Table 4.1.

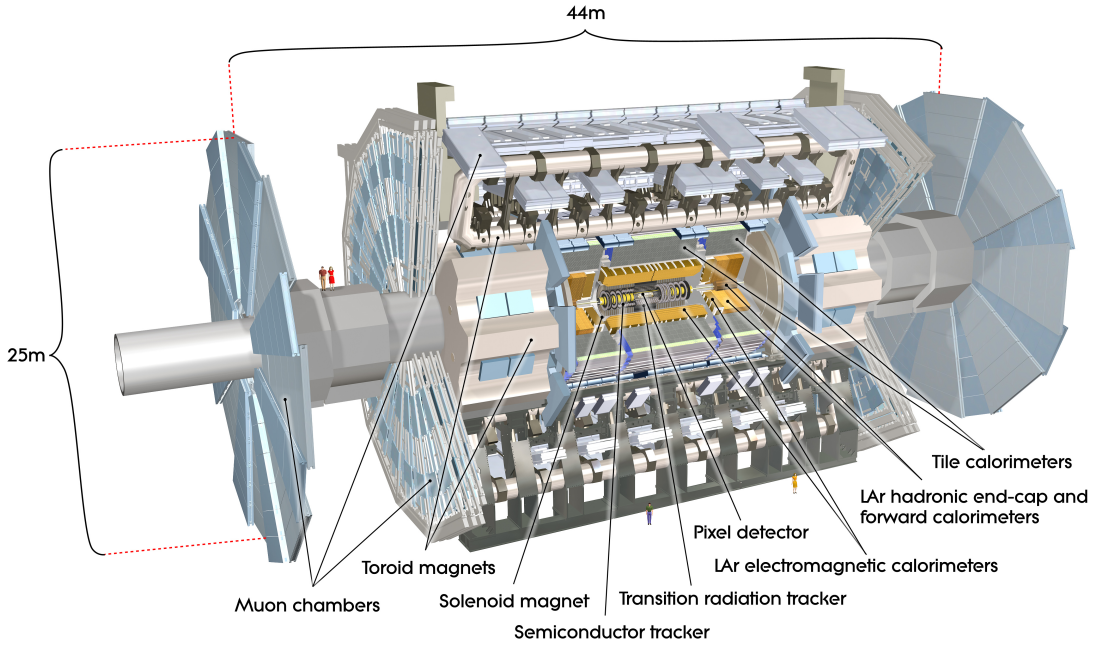


Figure 4.2: An overall view of the ATLAS dector with major subsystems labeled [74].

Detector Component	**Required Resolution**	**η Coverage**	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\% \, p_T \oplus 1\%$	± 2.5	–
EM Calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic Calorimetry (jets)			
Barrel and End-Cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
Forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon Spectrometry	$\sigma_{p_T}/p_T = 10\%$ at $p_T = 1$ TeV	± 2.7	± 2.4

Table 4.1: Expected design resolutions and $|\eta|$ coverage of the main detector components of the ATLAS detector. E and p_T are in units of GeV [74].

Now each subsystem of the ATLAS detector will be presented, from the innermost region near the IP outward.

4.2 The Inner Detector

The inner detector (ID), located directly outside the beam pipe where collisions take place, is responsible for precise momentum and vertex position measurements for charged particles and also contributes to the identification of electrons. The ID achieves these goals by employing high-resolution semiconductor pixel detectors, strip detectors, and straw-tube tracking detectors, which are all immersed in a 2 T magnetic field produced by the central superconducting solenoid magnet to induce the bending of charged particles for momentum measurements (see Section 4.5 for details on the magnet system).

An overview of the layout of the ID can be found in Fig. 4.3. Some key parameters of the ID can be found in Table 4.2.

Item		**Radial Extension (mm)**	**Length (mm)**	**No. of Channels**
Overall ID envelope		$0 < R < 1150$	$0 < \|z\| < 3512$	–
Beam-pipe		$29 < R < 36$	–	
IBL	Overall envelope	$30.0 < R < 40.0$	–	6.02×10^6
	Sensitive	$\langle R \rangle = 25.7$	$0 < \|z\| < 332$	
Pixel	Overall envelope	$45.5 < R < 242$	$0 < \|z\| < 3092$	80.4×10^6
3 cylindrical layers	Sensitive barrel	$50.5 < R < 122.5$	$0 < \|z\| < 400.5$	
2×3 disks	Sensitive end-cap	$88.8 < R < 149.6$	$495 < \|z\| < 650$	
SCT	Overall envelope	$255 < R < 549$ (barrel)	$0 < \|z\| < 805$	6.3×10^6
		$251 < R < 610$ (end-cap)	$810 < \|z\| < 2797$	
4 cylindrical layers	Sensitive barrel	$299 < R < 514$	$0 < \|z\| < 749$	
2×9 disks	Sensitive end-cap	$275 < R < 560$	$839 < \|z\| < 2735$	
TRT	Overall envelope	$554 < R < 1082$ (barrel)	$0 < \|z\| < 780$	3.51×10^5
		$617 < R < 1106$ (end-cap)	$827 < \|z\| < 2744$	
73 straw planes	Sensitive barrel	$563 < R < 1066$	$0 < \|z\| < 712$	
160 straw planes	Sensitive end-cap	$644 < R < 1004$	$848 < \|z\| < 2710$	

Table 4.2: Design parameters of the various ID subsystems [74].

The high-precision measurements are carried out by the Insertable B-Layer (IBL), pixel, and semiconductor (SCT) trackers in the region $|\eta| < 2.5$. The trackers are arranged in concentric cylinders around the beam line in the barrel region and parallel disks around the beam line in the end-cap region. In the barrel, the granularity is highest closest to the IP with the silicon pixel detectors, consisting of about 80.4 million readout channels. Each sensor has a pixel size of approximately 50×400 μm^2 in $R - \phi \times z$ and are arranged such

that a typical particle track will pass through three pixel layers. Such an optimized layout allows for intrinsic accuracies of 10 μm ($R-\phi$) and 115 μm (z) in the barrel and 10 μm ($R-\phi$) and 115 μm (R) in the end-caps.

In 2014, the IBL was installed for preparation for expected higher luminosities and to compensate for aging effects on the pixel detector due to radiation damage [75, 76]. Located between the beam pipe and the first pixel layer, the IBL adds additional redundancy with one extra hit in this critical small-radius region. It contains about 6 million pixel readout channels as is comprised of 14 carbon-fiber support staves with 200 μm doped planar silicon sensors in the central region and 230 nm, more radiation-resistant 3D silicon sensors in the forwards regions. The sensor resoluation are about 10 μm in $R-\phi$ and 75 μm in z.

The SCT consists of about 6.3 million readout channels arranged into eight strip layers, which are paired into four measurements in space with a 40 mrad stereo angle between pairs in order to resolve both $R-\phi$ and z. In the barrel region, 6.4 cm square sensors are aligned along the z direction with a strip pitch of 80 μm. In the end-caps, wedge-shaped sensors are oriented radially with the same 80 μm strip pitch as in the barrel. These arrangements afford intrinsic accuracies of 17 μm ($R-\phi$) and 580 μm (z) in the barrel and 17 μm ($R-\phi$) and 580 μm (R) in the end-caps.

Outside of the pixel and SCT detectors lies the transition radiation tracker (TRT), consisting of 351,000 straw-tube readout channels. Each straw has a diameter of 4 mm and is filled with a $Xe/CO_2/O_2$ gas mixture with a wire anode through the center that collects ionized particles from interactions between incident charged particles and the gas. The drift times from the particle location to the anode along a particular straw provide measurements in $R-\phi$ with an intrinsic accuracy of 130 μm. The straws are aligned in the z direction in the barrel with lengths of 144 cm and radially in the end-caps with lengths of 37 cm. The TRT collects a much larger number of hits (typically 36 per track) that provides valuable information for track-following.

The combination of the extremely precise measurements from the pixel and SCT detectors near the beam line and the large number of measurements over a larger length from the TRT further out allows for very precise momentum measurements, $R-\phi$ and z localization, vertexing and impact parameter measurements, and pattern recognition for particle

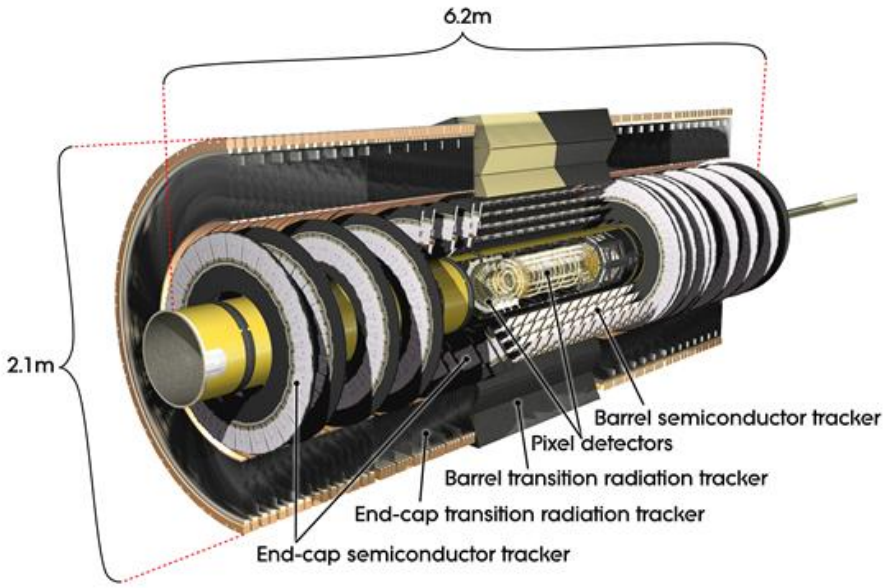


Figure 4.3: A schematic diagram of the ATLAS ID, showing the layouts of the pixel, strip, and transition radiation detectors [74].

identification and triggering.

4.3 Calorimeters

The ATLAS calorimeter systems measure the energy of electromagnetic (EM) and hadronic particles and contribute to particle identification. The layout of the various calorimeter systems, which cover the region of $|\eta| < 4.9$, are shown in Fig. 4.4. Based on the different physics and detector requirements and environments in different regions, the calorimeters feature several different subsystems with different designs and layouts. The central regions must provide relatively high-precision energy and spatial measurements to coordinate with information from the ID (especially concerning electron and photon measurements), and the outer end-cap regions must have sufficient performance for proper jet reconstruction and $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement while remaining radiation tolerant in the radiation-intense forwards regions. Additionally, the calorimeters must provide sufficient containment such that all SM particles except muons and neutrinos are absorbed in the calorimeters. This is both to limit punch-through of other particles to the muon system and to achieve precise $E_{\mathrm{T}}^{\mathrm{miss}}$

measurements critical to identifying potential BSM physics.

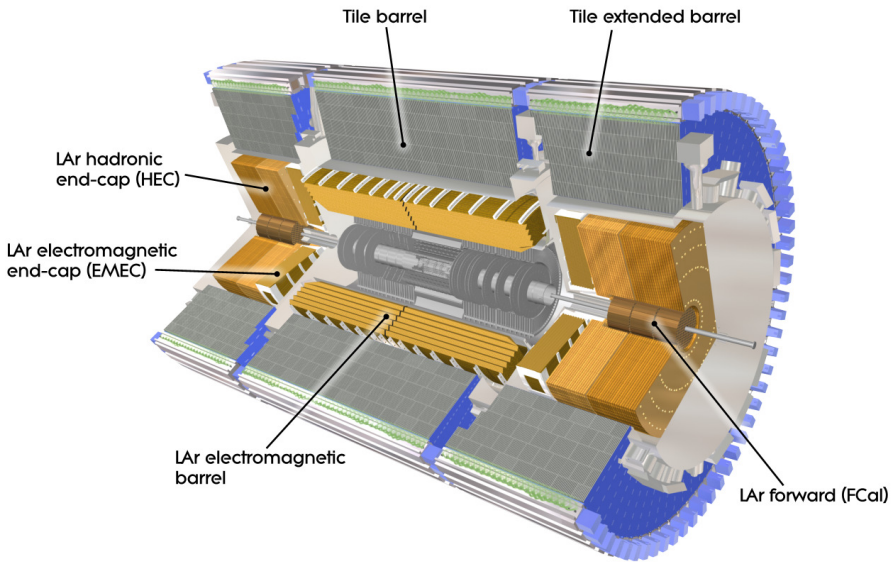


Figure 4.4: Diagram of the liquid argon and tile calorimeters, which sit outside the ID and solenoid magnet [74].

4.3.1 Liquid Argon Calorimeter

An overview of the LAr calorimeter itself and the collection and processing of the signal are discussed below. A more complete description is listed in Ref. [77].

4.3.1.1 Design & Structure

The LAr calorimeter consists of four sections: the EM barrel (EMB), EM end-cap (EMEC), hadronic end-cap (HEC), and forward (FCal) calorimeters, covering the regions $|\eta| < 1.475$, $1.375 < |\eta| < 3.2$, $1.5 < |\eta| < 3.2$, and $3.1 < |\eta| < 4.9$, respectively.

The LAr calorimeter is a sampling calorimeter with cryogenically-cooled liquid argon as the active medium and various materials interspaced as absorbers. In the EM sections (EMB and EMEC), lead plates are used as the passive material; in the FCal, there are copper and tungsten absorber matrices; and in the HEC, parallel copper plates form the absorbers. Cryogenic cooling is provided by three cryostats: one serving the EMB and one

each for the two sides that serve the EMEC, HEC, and FCal.

LAr is advantageous for calorimetry for many reasons. First, its reliably linear response allows for precise, well-understood measurements. Secondly, since it is a liquid, it can fill a large volume uniformly without cracks. Thirdly, when properly purified, it is chemically stable over time. And finally, it is very radiation-tolerant, capable of operating under LHC collision conditions without significant degradation in performance.

In the EMB and EMEC, the lead plate thickness is optimized as a function of $|\eta|$ to maximize the energy resolution while ensuring that the calorimeter is capable of sufficiently absorbing high-energy EM objects. The electrodes within the LAr gaps are arranged in accordion-like structures as illustrated in Fig. 4.5. This unique structure is self-supporting with full, uniform azimuthal coverage and no additional "dead" material for supports. The design also significantly reduces the inductance of the readout electrodes, allowing for a faster signal rise time and better timing performance.

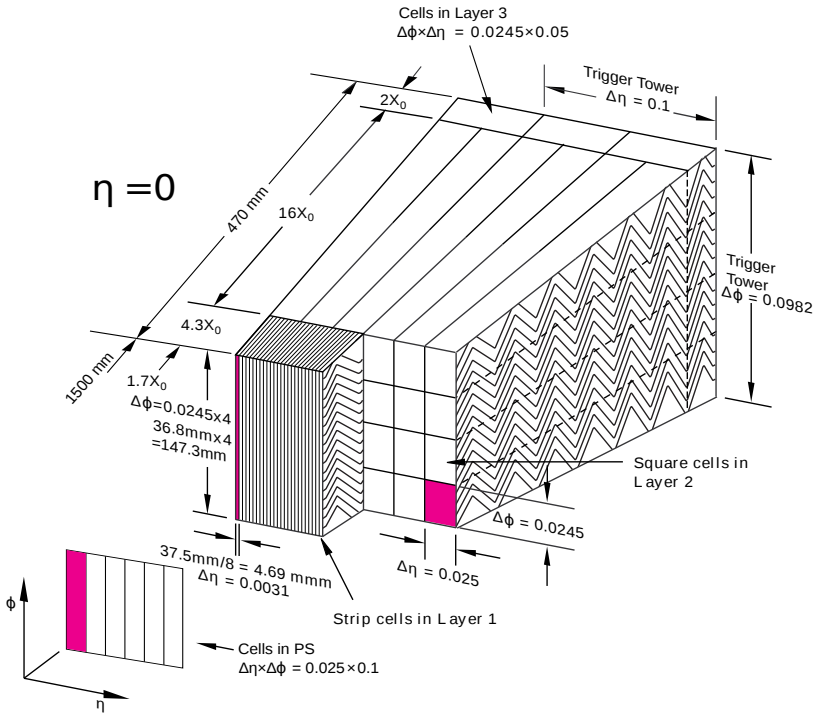


Figure 4.5: Diagram of the accordion electrode structure of the EMB and EMEC calorimeters. [77].

The LAr calorimeter cells are segmented with varying granularity both laterally in ϕ and

η as well as longitudinally, i.e. in depth. The granularity in $\Delta\eta$ and $\Delta\phi$ as a function of η for the EM sections are detailed in Table 4.3. In the precision-physics region with $|\eta| < 2.5$, the lateral granularity is finest. Additionally, a so-called presampler layer exists in the region $|\eta| < 1.8$ in order to account for radiative losses of EM objects before the calorimeter. In the first layer, there is especially fine segmentation in η in order to resolve the pairs of photons from π^0 decays and thus achieve efficient rejection of such objects. Such fine η segmentation also allows for precise measurements of the pointing of photons (see Section 5.2.1), which reduces the mass resolution of diphoton Higgs boson decays by improving the identification of the vertex from which the photons originated. Photon pointing also provides a valuable discriminating variable for the identification of potential BSM long-lived particles decaying to EM objects, which would tend to produce electrons/photons that do not point back to the IP. Such reasoning is a crucial discriminating feature in the analysis presented in this thesis.

η range	0 to 1.4	1.4 to 1.8	1.8 to 2.0	2.0 to 2.5	2.5 to 3.2
Presampler	0.025×0.1	0.025×0.1			
Sampling 1	0.003×0.1	0.003×0.1	0.004×0.1	0.006×0.1	0.1×0.1
Sampling 2	0.025×0.025	0.025×0.025	0.025×0.025	0.025×0.025	0.1×0.1
Sampling 3	0.050×0.025	0.050×0.025	0.050×0.025	0.050×0.025	
Trigger	0.1×0.1	0.1×0.1	0.1×0.1	0.1×0.1	0.2×0.2
Readout channels	110 208	25 600	12 288	24 064	1792

Table 4.3: Granularity in $\Delta\eta \times \Delta\phi$ of cells in the EM sections of the LAr calorimeter [77].

The HEC consists of wedge-shaped conventional parallel copper plate electrodes arranged in two wheels per end-cap. A schematic of one of the 32 modules in each of these wheels is shown in Figure 4.6. The HEC shares cryostats with the EMEC and FCal and overlaps slightly in η with the FCal and tile calorimeters in order to achieve a higher density to compensate for the loss that occurs in the transition region between different detectors.

In order to improve uniformity and reduce the radiation in the muon system, the FCal is built into the end-cap cryostat system itself. The FCal consists of copper and tungsten absorber matrices with rod electrode structures and LAr filling the gaps between the two. This module layout is illustrated in Figure 4.7.

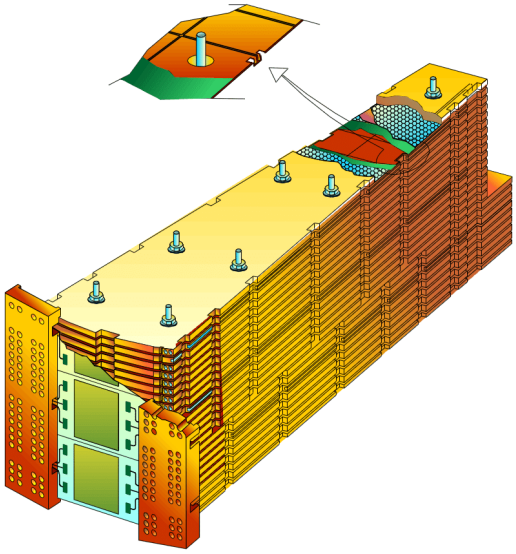


Figure 4.6: LAr HEC module diagram showing the wedge shape and readout structure [74].

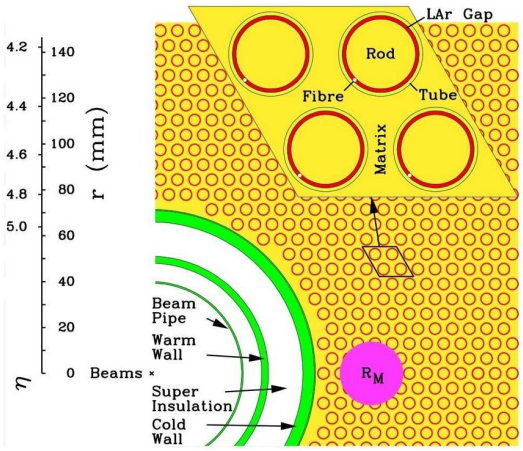


Figure 4.7: Diagram of the LAr FCal rod electrode and absorber matrix layout. The Molière radius is labelled R_M [74].

4.3.1.2 LAr Signal Measurement & Readout

An incoming particle hits the absorber material and produces an ionizing shower into the LAr. Electrodes placed in the LAr gaps collect the resulting current, as illustrated in Fig. 4.8 for the EMB and EMEC.

The electrodes send this signal to on-detector electronics, where it is amplified, split in to three overlapping linear gain scales, and shaped as shown in Fig. 4.9. The shaped pulse is sampled at the LHC bunch crossing frequency of 40 MHz and stored in an analog pipeline, awaiting a trigger decision. If triggered, the appropriate gain is selected, and four samples are digitized and read out via optical fiber.

In order to extract the peak amplitude (which is directly proportional to the energy) and arrival time of the incident particle, calibrations are carried out using an optimal filtering technique to maximize the precision of these values of interest [79]. This technique exploits the knowledge of the expected shape and auto-correlation matrix of the sampled ionization pulse. Then, using a Lagrange multiplier optimization approach, the so-called optimal filtering coefficients (OFCs) are computed. These coefficients, denoted as a_i and b_i

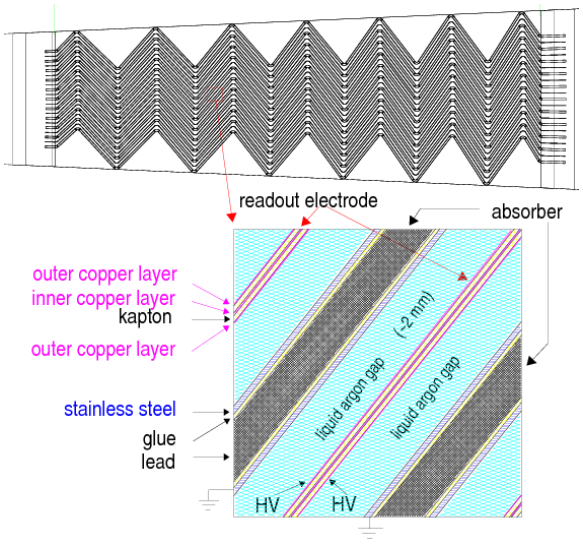


Figure 4.8: Diagram of the structure of the lead absorbers and liquid argon gaps in the EMB and EMEC calorimeters [78].

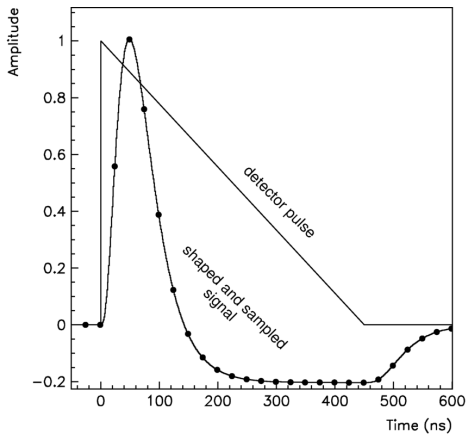


Figure 4.9: Illustration of the shaped and unshaped ionization pulses from LAr ionization [74].

in Equation 4.3, serve as weights by which to multiply the samples in order to obtain the optimal pulse amplitudes and times.

$$A = \sum_{i=0}^{n-1} a_i S_i$$
$$t = \frac{1}{A} \sum_{i=0}^{n-1} b_i S_i \tag{4.3}$$

where A is the peak amplitude, t is the peak time, and S_i are the measured amplitude of sample i. For the Run 2 data used in this thesis, the value of n, the total number of samples taken from the pulse, was four.

4.3.2 Tile Calorimeter

The tile calorimeter is a sampling calorimeter that surrounds the EM calorimeters and that is optimized to absorb and measure the energy of hadronic particles. The absorbing material is steel, and scintillating tiles are used for the active medium. Scintillation light is

collected by wavelength-shifting fibers on each side of the tiles and sent to photomultiplier tubes (PMTs). The tile calorimeter system consists of one barrel and two end-cap sections covering the regions $|\eta| < 1.0$ and $0.8 < |\eta| < 1.7$, respectively. The small crack region between the barrel and end-cap sections is filled with cables, power supplies, and other services for the ID and LAr calorimeter as well as a set of special, small scintillators similar to those in the rest of the tile calorimeter that help compensate for the energy lost in this crack region.

The tile calorimeter sections are segmented longitudinally into three layers and stretch from a radius of 2.28 m to 4.25 m, corresponding to a total depth of ∼7.4 interaction lengths. The cells comprising the first two layers and the last layer have granularities of 0.1×0.1 and 0.2×0.1 in $\Delta\eta \times \Delta\phi$, respectively. The approximate uniformity of cell granularity in pseudorapidity is achieved by an optimized grouping of cells into PMTs. Each section contains 64 radial wedge modules with a supporting steel girder on the outer edge that both houses the readout electronics and serves as a return yoke for the magnetic flux of the solenoid (as discussed further in Section 4.5). A schematic of one of these modules, showing the arrangement of the scintillators, absorbers, and readout system is shown in Figure 4.10.

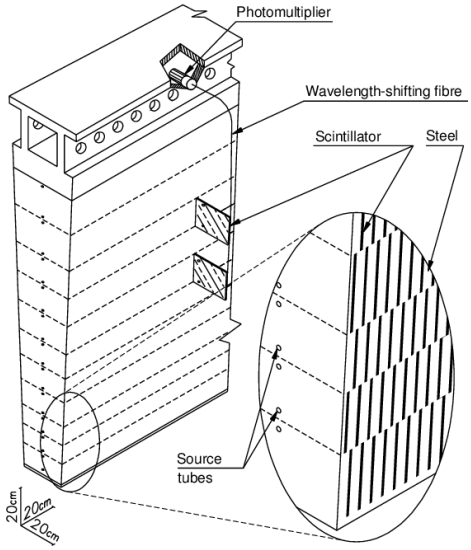


Figure 4.10: A schematic of one of the 64 tile calorimeter modules (each with $\Delta\phi \sim 0.1$) that make up a full annular section centered on the beam axis [74].

4.4 Muon Spectrometer

Located in the outermost extent of the ATLAS detector is the muon spectrometer (MS) system, which detects muons, the primary SM particles (besides the effectively undetectable neutrinos) that escape all other layers of the detector. The MS aims to detect these particles and their charges and to perform precise momentum measurements, in addition to providing information to the trigger. Muon measurements are obtained from bending in the magnetic field provided the toroid magnet system (see Section 4.5). These magnets are designed such that the magnetic field is mostly perpendicular to the muon path, but also such that multiple scattering is suppressed in order to prevent a subsequent reduction in resolution. To accommodate varying muon fluxes, there are four separate MS subsystems, with varying design parameters tailored to the different expected environments in different $|\eta|$ regions. These subsystems are referred to as the Monitored Drift Tubes (MDTs), the Cathode Strip Chambers (CSCs), the Resistive Plate Chambers (RPCs), and the Thin Gap Chambers (TGCs). Each of these subsystems is described below, and their main parameters are summarized in Table 4.4. Their spatial arrangement is illustrated in a profile diagram in the y-z plane in Figure 4.11.

	MDT	**CSC**	**RPC**	**TGC**
Function	Tracking	Tracking	Triggering & second coordinate	Triggering & second coordinate
Coverage	$\|\eta\| < 2.7$	$2.0 < \|\eta\| < 2.7$	$\|\eta\| < 1.05$	$1.05 < \|\eta\| < 2.7$ (2.4 for triggering)
No. of Chambers	1150	32	606	373k
No. of Channels	354k	30.7k	373k	318k
Chamber Resolution				
- z or R	35 μm (z)	40 μm (R)	10 mm (z)	2-6 mm (R)
- ϕ	–	5 mm	10 mm	3-7 mm
- time	–	7 ns	1.5 ns	4 ns
Measurements/Track				
- barrel	20	–	6	–
- end-cap	20	4	–	9

Table 4.4: Parameters describing the four subsystems comprising the muon spectrometer system [74].

Precision tracking is provided by two subsystems: the MDTs in the region $|\eta| < 2.7$ (ex-

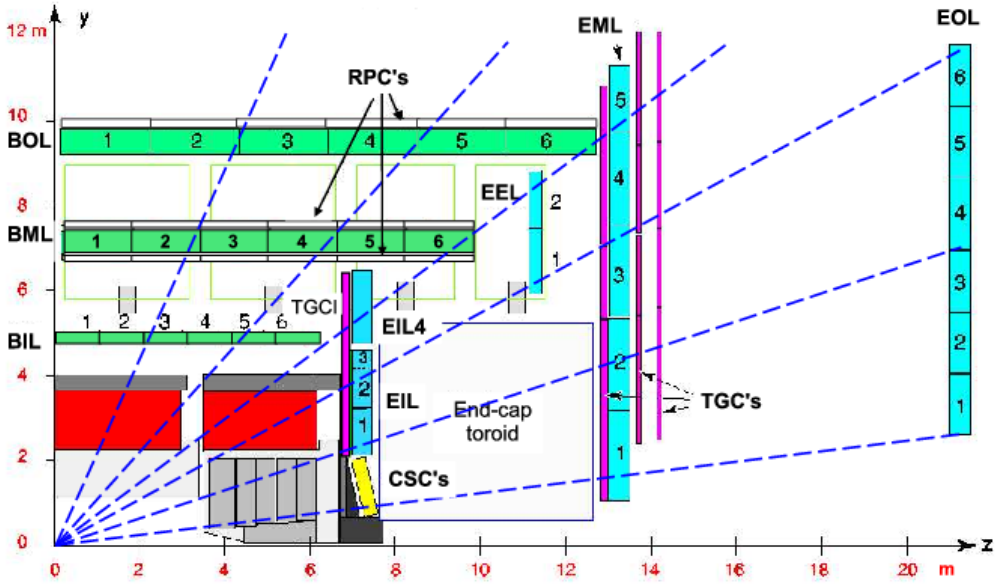


Figure 4.11: A profile view in the $y - z$ plane of the muon spectrometer system. The MDT chambers are shown in green in the barrel and in teal in the end-caps [74].

cept the innermost end-cap layer which only extends to $|\eta| < 2.0$) and the CSCs in the region $2.0 < |\eta| < 2.7$. The MDTs are composed of individually isolated wires within drift tubes, which provides reliable, precise tracking information. The CSCs are multiwire proportional chambers arranged in perpendicular strips, allowing for both η and ϕ measurements. The CSC detectors have higher granularity and more precise timing performance relative to the MDTs, which make them more suitable to the more forward region with more intense rates.

Efficient triggering on muons is provided in the region $|\eta| < 2.4$. To achieve this goal as well as to identify the bunch crossing and to obtain a second coordinate for tracking in the direction perpendicular to the bending plane (i.e. the ϕ direction), the MS employs RPCs in the region $|\eta| < 1.05$ and TGCs for $1.05 < |\eta| < 2.7$. After precise alignment of all the subsystems of the MS, matching of the hits from the precision tracking and triggering parts of the muon system can be achieved with negligible probability for overlapping tracks from uncorrelated particles. In the correlated case, matching with inner detector tracks resolves the ambiguity.

4.5 Magnet System

The ATLAS magnet system consists of four superconducting magnet systems: a central solenoid, barrel toroids, and two sets of end-cap toroids. The geometric layout of the system is illustrated in Fig. 4.12.

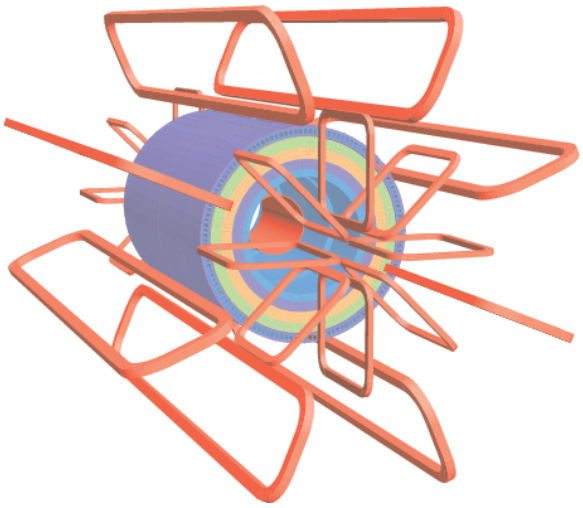


Figure 4.12: The geometric layout of the magnet system with central solenoid, tile calorimeter layers and return yoke, eight barrel toroid coils, and two sets of eight end-cap toroid coils. The magnets themselves are shown in red while the tile calorimeter structure is shown in different colors representing the different magnetic properties of the different layers and outer yoke structure [74].

The solenoid provides an axial magnetic field along the z-axis for the ID with a strength of 2 T at the center and was designed to limit radiative losses for traversing particles so as not to distort subsequent calorimetric measurements. Accordingly, the solenoid is only $\sim$0.66 radiation lengths thick for normally-incident particle. Such a small loss is partly achieved by the sharing the vacuum vessel with the LAr calorimeter. The 5.4-metric-ton solenoid consists of a single Al-stabilized NbTi conductor coil supported within a 12 mm-thick aluminum alloy cylinder with inner and outer diameters of 2.46 and 2.56, respectively. The steel of the tile calorimeter and its supporting girders serve as return yokes for the solenoid's magnetic flux.

The barrel and end-cap toroids provide $\sim$0.5 and $\sim$1 T magnetic fields, respectively, for the MS. The coils are each housed within in racetrack-shaped vacuum vessels arranged radi-

ally around the beam axis. The eight barrel toroid coils lie just outside the tile calorimeter, stretching 25.3 m along the beam axis and out from a diameter of 9.4 m to 20.1 m. The eight end-cap toroid coils on each sides are placed 5 m along the beam pipe out from a diameter of 1.65 m to 10.7 m. All toroid coils use an Al-stabilized Nb/Ti/Cu conductor wound into flat coils. The toroid magnet systems provide bending powers of $\sim 1.5 - 5.5$ Tm in the region $0 < |\eta| < 1.4$ and $\sim 1 - 7.5$ Tm in the region $1.6 < |\eta| < 2.7$, with lower power in the intermediate region where there is a transition between the barrel and end-cap systems.

4.6 Forward Detectors

The primary function of the forward detectors is to perform precise measurements of the delivered luminosity, which is critical to understanding and measuring the cross sections of various interactions. The three primary forward detectors used for this purpose are LUminosity measurement using Cerenkov Integrating Detector (LUCID), Absolute Luminosity For ATLAS (ALFA), and Zero-Degree Calorimeter (ZDC). The coverage of these detectors is summarized in Table 4.5. By measuring inelastic $p-p$ scatters, LUCID provides the main online relative-luminosity measurement. The number of particles detected from such scatters is proportional to the pileup in a given bunch crossing. LUCID achieves a time resolution capable of measuring individual bunch crossings each 25 ns. The LUCID detectors consist of twenty aluminum tubes on each side of the detector, lying just outside the beam pipe, that are filled with C_4F_{10} at 1.2-1.4 bar. PMTs collect Cerenkov light with a half-angle of 3° that is produced in these tubes to count the number of incident particles, which is then used to predict the luminosity. ALFA measures the absolute luminosity from elastic $p - p$ collisions using so-called Roman pots containing scintillating fibers that lie within a few millimeters of the beam. After calibration with ALFA, LUCID is designed to measure the luminosity to a precision of $< 5\%$. The ZDC is used during heavy-ion collisions to measure the centrality of such collisions by measuring neutral particles passing through alternating quartz and tungsten layers.

Detector	Coverage	Dist. from IP ($\|z\|$ [m])
LUCID	$5.6 < \|\eta\| < 5.9$	17
ZDC	$8.3 < \|\eta\|$	140
ALFA	$10.6 < \|\eta\| < 13.5$	240

Table 4.5: Location and coverage of the forward detectors.

4.7 Trigger & Data Acquisition

The Trigger and Data Acquisition (TDAQ) system is responsible for managing the selection and flow of data to be considered for further processing and possible long-term storage. It performs this task via a series of compartmentalized filtering stages divided into three levels: Level 1 (L1), Level 2 (L2), and the event filter. The L2 trigger and the event filter are collectively known as the High-Level Trigger (HLT). A diagram illustrating the general structure of the TDAQ systems is shown in Figure 4.13.

At the 40 MHz LHC bunch crossing frequency, select data from the calorimeter and muon spectrometer systems is sent to the L1 trigger, whose readout and processing is performed via dedicated on-detector hardware. The L1 trigger system performs some limited processing to make a trigger decision in less than 2.5 μs, reducing the data rate down to $\sim$100 kHz. The L1 readout granularity in the calorimeter is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, summing the activity through all layers to form so-called trigger towers. The readout and buffering of data from the front end to the HLT, as directed by the L1 trigger, is facilitated by Readout Drivers (RODs). RODs provide a common technical infrastructure with which many subsystems can coherently collect, organize, buffer, and transmit data for further processing.

The RPC and TGC systems provide triggering for high-p_{T} muons to the L1 trigger. Likewise, using a lower-granularity readout from groupings of cells in the calorimeters allows for quick, efficient identification of high-p_{T} electrons, photons, jets, and hadronically-decaying τ-leptons. Additionally at L1, simplified computations of the total E_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ for the event (reconstructed from the muon spectrometer and calorimeter systems together) are used for triggering. A central trigger processor takes all of these inputs and carries out a so-called trigger menu by performing a set of logical operations on individual triggers in

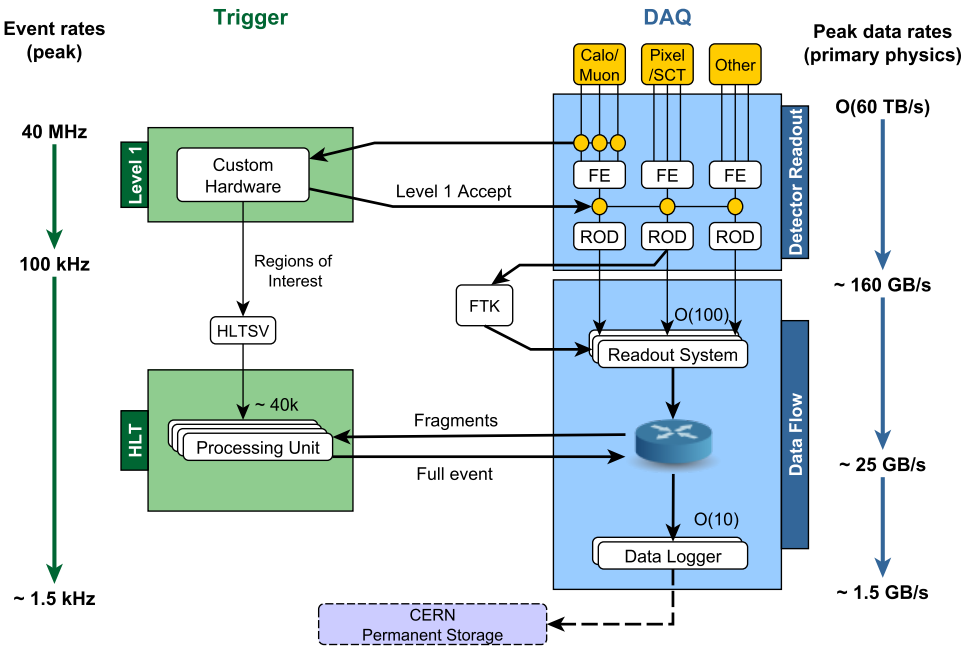


Figure 4.13: Overview of the TDAQ systems for Run 2 with approximate event and data rates at each stage of the filtering process [80].

combination with pre-scaling of the rates themselves to best optimize the available bandwidth for different run conditions. The resulting configurable menu represents the L1 trigger selection that then triggers data from the rest of the detector (ID, forward detectors, etc.) to be read out from the front-end electronics and send on for further processing by the HLT. The L1 trigger also passes information on to the HLT about the Regions-of-Interest (RoIs) in which the triggering decision was made.

At L2, using the full readout granularity and precision in the RoIs identified at L1, more sophisticated selections are applied to make another triggering decision. This reduced data stream is then fed into the event filter, which fully reconstructs the event and applies another level of selections based on offline analysis methods. Unlike the hardware-based L1 system, both the L2 and event-filter stages are entirely software-based. These stages, which comprise the HLT system, reduce the data flow down to $\sim$1.5 MHz, with a final event size of $\sim$1 MB, all of which is then sent off for permanent storage.

Chapter 5: Object Reconstruction & Identification

This chapter describes the algorithms used to reconstruct physics objects from measurements collected by the various subdetectors. After reconstruction, objects can be identified with a known physical particle using algorithms with associated efficiencies and background rejection rates. Object identification (ID) requirements utilize various likelihood (LH) models that prescribe set criteria useful in different analysis contexts. These criteria are based on an optimization of cuts to distinguish the target particle from backgrounds that mimic, or "fake", that particle. Additionally, isolation requirements further improve background rejection by examining the surrounding detector activity. The impact of such ID and isolation requirements, as well as the nature of other special measurements—namely, pointing and timing—in the context of the present analysis, are discussed below.

5.1 Physics Object Reconstruction

In this section, the general algorithms for reconstructing the various objects relevant to this analysis are discussed.

5.1.1 Primary Vertices

As will be discussed in more detail in the following sections, hits in the various subdetectors are grouped into tracks that reconstruct the path different particles follow. The precise extrapolation of these tracks back to the IP leads to convergences that are labelled as vertices corresponding to the approximate location at which the corresponding collision products emerged. Due to pileup, multiple interaction are known to take place and thus multiple vertices are often reconstructed. The hardest of these vertices, that is, the one with

the largest $\sum p_T^2$, is denoted the primary vertex (PV). During the reconstruction procedures described below, objects are associated to the same interaction if it can be determined that they are consistent with production at the PV.

It is worth noting that the standard tracking-based choice of PV as described is used in the analysis, without considering photon pointing information. Since the analysis triggers on a hard, prompt lepton, the PV identification is already quite robust due to the presence of the relatively more precise tracking information.

5.1.2 Topological Clusters

EM and hadronic showers produce regions of activity over multiple cells in the calorimeters. Topological clusters, or topo-clusters, are the groupings of this activity into coherent clusters, which are constructed using an optimized, noise-resistant algorithm [81]. Topo-clusters serve as the basic building blocks to construct more complex objects in a variety of contexts, as described in the following sections.

Topo-clusters are seeded by cells with energies four times higher than their expected noise thresholds, considering both electronic noise and pileup. From these seeds, the topo-clusters are expanded in three dimensions across both the EM and hadronic calorimeters and their respective layers, adding adjacent cells with at least twice the expected noise energy threshold. These so-called proto-clusters may merge with others if cells border both. Once expansion stops, an additional shell of cells surrounding the cluster is added. An energy threshold removes proto-clusters below a set threshold.

Since showers may overlap, the construction of proto-clusters can often result in clusters covering large areas of the calorimeters. Therefore, the next step considers the splitting of proto-clusters if it can be determined that there are multiple, resolvable objects contained therein. In general, if local maxima can be reliably picked out, the cluster may be split and the cells assigned to separate objects. This reassignment is achieved in an iterative manner similar to the initial proto-cluster construction, but this time only the original cells are used, no energy thresholds are applied, and cluster merging does not occur. When a conflict emerges over cells joining multiple clusters, a sharing algorithm weights them according to

their energies and distances and splits them between the clusters accordingly[1]. Any clusters not assigned to a local maximum are assigned to separate clusters.

5.1.3 Isolation

Isolation requirements examine the neighborhood around the object to determine the degree to which it can be considered to be isolated from surrounding detector activity [82]. The concept of isolation is valuable in distinguishing, for example, electrons from converted photons and jets—both of which tend to have more surrounding activity than cleaner, focused electron tracks and showers. Isolation measurements consist of both track and calorimeter isolation, which are then combined into fixed working points with given target applications, efficiencies, and background rejection rates that can be chosen based on the relevant analysis context.

Track isolation examines the momentum of other tracks within a cone around the target object. The p_{T} of tracks within the cone (excluding those matched to the target object) are summed. For higher energies, it is often the case that background processes that fake the target signature can lie very close to the target itself, due to the reduced bending in the magnetic field. Therefore, a variable-sized cone is often employed. The radius (in ΔR) of this cone scales inversely with p_{T} and has a maximum size generally set at 0.2[2]. In order to mitigate noise effects, tracks must have $p_{\mathrm{T}} > 1$ GeV to be included in the sum, and they must satisfy specific quality requirements related to the nature of their hits in the ID subdetectors and their associated vertices. Further details can be found in Ref. [83].

Calorimeter isolation considers the E_{T} of topo-clusters in surrounding regions and applies cuts to increase the target purity. Calorimeter isolation is generally quantified by examining the energy outside of a small core of cells around the target particle but still within a larger ΔR cone. The E_{T} of topo-clusters whose barycenters lie within this fixed ΔR cone is summed. Then, the E_{T} of middle-layer cells within the core of dimensions $N_\eta \times N_\phi = 5 \times 7$ is subtracted. Corrections to account for leakage of energy outside the immediate core, pileup, and energy contributions from the underlying event are applied to obtain the final calorimeter

[1]The weights are required to sum to unity, so there is no double counting.

[2]Specifically, $\Delta R = \min\left(\frac{10}{p_{\mathrm{T}}\,[\mathrm{GeV}]}, \Delta R_{\mathrm{max}}\right)$, where R_{max} is usually 0.2.

isolation measurement. An illustration of the isolation regions is shown in Figure 5.1.

Different standard isolation working points can choose different sizes for the ΔR cone considered, but generally, a ΔR value of 0.2 is used for electrons and either 0.2 or 0.4 is used for photons.

Isolation working points combine track and calorimeter isolation measurements and achieve different compromises between efficient ID of the target object and efficient rejection of backgrounds. These working points can set fixed ΔR cone sizes, or they can use what is called a gradient working point that aims to achieve a constant efficiency, uniform in η, at a given energy by adjusting the cuts accordingly. The working point is chosen based on the desired level of performance in a given context, taking into account the cleanliness of the relevant dominant background. The details of the available cuts for electrons and photons can be found in Ref. [83]; for muons, Ref. [84]; and for jets, Ref. [85]. Another method of combining track and calorimeter isolation information is known as particle flow, which attempts to take advantage of the improved energy and angular resolutions and reduced pileup dependence of the tracker compared to the calorimeter at low p_{T}, while using the calorimeter's ability to measure neutral particles. The full details of the particle flow algorithm are presented in Ref. [85] and are summarized in the context of electron, photon, and muon isolation in Ref. [86]. The working points chosen for the present analysis will be discussed in Chapter 6.

5.1.4 Electrons & Photons

For the purposes of reconstruction, electrons and photons are generally treated quite similarly, as they both result in similar EM showers in the LAr calorimeter. Their reconstruction involves the clustering of calorimeter activity and matching with tracks in the ID, if present. The reconstruction of tracks in the ID is performed similarly for all charged particles, including both electrons and converted photons—that is, photons that undergo pair production into electron-positron pairs within the tracking volume. An optimized tracking algorithm that efficiently handles reconstruction from the many hits in the various ID subdetectors is used to obtain ID tracks. The full details of this algorithm can be found in Refs. [87, 88]. Superclusters, along with any associated tracks from the ID, are the primary

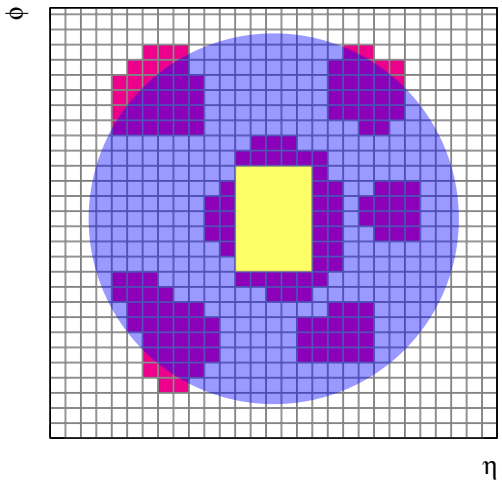


Figure 5.1: Illustration of the isolation calculation in the calorimeter, centered on the cluster in question. Topo-clusters are shown in red, the blue ΔR region shows the surrounding region over which topo-cluster E_{T} is summed the yellow rectangle depicts the $\Delta\eta \times \Delta\phi = 0.125 \times 0.175$ window of the core, and the blue ΔR region shows the surrounding isolation region [82].

building blocks of electron and photon reconstruction [83]. They are variable-size clusters that maintain a linear energy response, resistance to pileup effects, and flexibility to account for bremsstrahlung effects. Both electrons and converted photons can have hits in the ID matched to superclusters. However, significant separation can be achieved with detailed track analyses that identify supercluster-matched ID activity as either a track or a photon conversion vertex. Photons that do not undergo pair production—known as unconverted photons—can be distinguished from electrons by their lack of matching ID hits. However, in all cases, the reconstruction of electrons and photons begins with the construction of these superclusters, which is described below.

The overall procedure to form superclusters consists of seeding by topo-clusters, adding secondary clusters, and applying calibrations. Supercluster construction begins with the testing of EM topo-clusters for use as seeds in order of descending energies. Energy thresholds are set, and track matching is required when considering electrons. Then, a search for potential satellite clusters around the seeds begins with a fixed window of $N_\eta \times N_\phi = 3 \times 5$ cells around the seed cluster barycenter. For electrons, an expanded $N_\eta \times N_\phi = 5 \times 12$ window is also considered if the two clusters have the same matched track. For converted

photons with conversion vertices with hits only in the pixel detector, a satellite cluster is added if its matched track contains the conversion vertex of the seed cluster. An overview of supercluster construction as described is illustrated in Figure 5.2. As a final step, the size of all topo-clusters within each supercluster is limited to a maximum of 0.075 and 0.125 in η in the barrel and endcaps, respectively, in order to mitigate pileup effects.

Once the superclusters are constructed, preliminary energy and position calibrations are applied and the matching of ID tracks and conversion vertices is performed, as was done with topo-clusters. Based on the nature and quality of the matching between ID hits and superclusters, electron and photon objects are created. If there is no ambiguity, only one is created. Otherwise, both are, and they are flagged as ambiguous for further consideration. Finally, another energy calibration that uses $Z \to ee$ and $Z \to ll\gamma$ data is performed based on the new information from the matching, as described in Ref. [89].

Next, LH identification models optimized for different candidate particles are applied in order to distinguish the target particle from similar backgrounds that may fake its signature. Different ID working points exist that balance signal efficiency and background rejection, which can be selected depending on the analysis context. Variables that enter the LH identification for both electrons and photons include shower shape measurements involving hadronic leakage and first- and second-layer shape information. For electrons, additional inputs come from the third layer, track conditions, and the matching of tracks to superclusters.

There are generally three photon ID requirements that are made available by standard object processing: loose, medium, and tight, in order of increasing strictness. These requirements are designed to increase the purity of true photons to different degrees, according to the appropriate context. Loose ID requirements set cuts on the level of hadronic leakage and the shapes of the showers in the middle layer of the LAr calorimeter. The medium requirements add one additional cut having to do with the sizes of the two largest energy deposits in the finely segmented strip layer. This cut is used primarily as a way to filter out photons originating from jets with a leading π^0 meson that decays to two photons. The tight requirements impose yet more constraints on middle and strip layer shower shapes for even higher purity. The full details of these different ID categories can be found in Ref. [90]. A

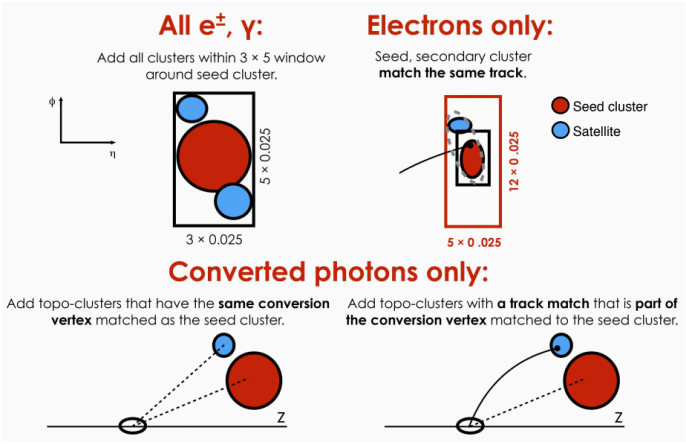


Figure 5.2: Illustration of the superclustering algorithm for electrons and both converted and unconverted photons [83].

summary of the difference between the types of requirements for each is shown in Table 5.1. The efficiency of the different photon IDs varies with E_{T}, converted/unconverted status, and pileup. See Ref. [83] for more details.

Electron reconstruction proceeds in much the same manner as that for photons. However, special emphasis is placed on the matching of tracks in the ID to superclusters in the calorimeter [91]. Tracks are matched via extrapolation from the last measured point to the second calorimeter layer, which, after accounting for possible bremsstrahlung losses, must differ from the second-layer seed cluster by less than a set threshold. In the case of multiple matches, the one with the lowest ΔR between impact point and seed cluster is selected, after giving preference to tracks with silicon hits[3]. The electron energy is computed via a weighted average of the track momentum and the total cluster energy. The η and ϕ coordinates are generally taken from track measurements, but if certain ID measurements are missing, they may be taken from the cluster.

Jets, converted photons, and semi-leptonic heavy-quark hadrons can fake the signature of an electron. Similar to what is done for photons, a multivariate LH optimization procedure is used that takes into account many of the same kinds of shower shape and hadronic leakage

[3]Those without silicon hits being more likely to be associated with a converted photon rather than an electron.

Type of Requirement	Name	Description	Loose	Medium	Tight
Hadronic Leakage	R_{had_1}	Ratio of E_{T} in the first sampling of the hadronic calorimeter to the E_{T} of the EM cluster (used over the range $\lvert\eta\rvert < 0.8$ and $\lvert\eta\rvert > 1.37$	✓	✓	✓
	R_{had}	Ratio of E_{T} in the full hadronic calorimeter to the E_{T} of the EM cluster (used over the range $0.8 < \lvert\eta\rvert < 1.37$)	✓	✓	✓
EM Middle (Second) Layer	R_{η}	Ratio in η of cell energies in 3×7 versus 7×7	✓	✓	✓
	w_2	Lateral width of the shower	✓	✓	✓
	R_{ϕ}	Ratio in ϕ of cell energies in 3×3 and 3×7 cells			✓
EM Strip (First) Layer	$w_{s,3}$	Shower width for 3 strips around the maximum energy strip			✓
	$w_{s,\text{tot}}$	Total lateral shower width			✓
	F_{side}	Fraction of energy outside the core of 3 central strips but within 7 strips			✓
	ΔE	Difference between the energy in the second maximum in the strip layer and the energy in the strip with the minimal values found between the first and second maxima			✓
	E_{ratio}	Ratio of the energy in the largest and second-largest energy deposits to the sum of these energies		✓	✓

Table 5.1: Description of the loose, medium, and tight photon ID requirements [90].

considerations as is done for photons, in addition to more detailed track and track-cluster matching information, as previously mentioned [82]. Also like photons, there are three standard working points for electrons denoted, in order of increasing rejection power, as loose, medium, and tight IDs. The loose ID requirements deal mostly with shower shape variables and achieve a jet-rejection factor of roughly 500. Medium ID folds in more stringent shower shape cuts as well as some track-related cuts to reach a jet-rejection factor of about 5,000. Tight ID adds even more shower shape and track criteria to push the jet-rejection factor up to 50,000. For typical electroweak processes, the electron LH ID achieves, on average, efficiencies of 93%, 88%, and 80% for loose, medium, and tight IDs, respectively[4].

5.1.5 Jets

The hadronization of quarks and gluons produces fragmented but collimated hadronic showers of particles in the detector that can contain a large number of decay products spread out

[4]A gradual increase in efficiency as a function of E_{T} is also observed. See Ref. [83] for more details.

among multiple clusters that all together comprise what is known as a jet. Jets are not used directly in the analysis, but they do enter into the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$ and overlap removal considerations. Jet are reconstructed combining information from the calorimeters and the ID tracker using the previously mentioned particle flow algorithm, described now in more detail below.

The particle flow algorithm iterates through a series of optimized steps to improve the performance of jet reconstruction [85]. Energy deposited in the calorimeter by charged particles is subtracted and replaced by the momenta of tracks that are matched to those topological clusters. The jet four-momentum is corrected for the non-compensating calorimeter response[5], signal losses due to noise threshold effects, energy lost in non-instrumented regions, and contributions from pileup [92, 93].

The combination of noise-suppressed topo-clusters to form a jet object is done using the so-called anti-k_T algorithm[6] with a radius parameter $R = 0.4$ [94]. A description of this algorithm is presented briefly below.

The anti-k_T algorithm defines two distance parameters:

$$d_{ij} = \min\left(\frac{1}{p_{\mathrm{T}i}^2}, \frac{1}{p_{\mathrm{T}j}^2}\right)\frac{\Delta R_{ij}^2}{R^2} \tag{5.1}$$

$$d_{iB} = \frac{1}{p_{\mathrm{T}i}^2} \tag{5.2}$$

where d_{ij} is a distance between topo-clusters i and j, and d_{iB} is a distance between topo-cluster i and the beam line. Additionally, in this case, ΔR_{ij} is defined as $\sqrt{\Delta(y_{ij})^2 + (\phi_{ij})^2}$, that is, using rapidity rather than pseudorapidity, as is used elsewhere in this thesis.

These distances d_{ij} and d_{iB} are computed for all topo-clusters being considered. The algorithm then iterates over all distances from smallest to largest. For each distance, if it is of the first type (d_{ij}), clusters i and j are combined. If it is of the second type (d_{iB}), cluster i is considered a jet and is removed from further iterations.

---

[5] This refers to the difference in response of the calorimeter to EM vs. hadronic showers.

[6] Here, k_t is the transverse momentum, which, in the notation used throughout this thesis, is denoted p_{T}. However, anti-k_T is the standard name, so it will be used when referring to the name of the algorithm.

The anti-k_T algorthm has a few key properties and features that justify its use for jet reconstruction. Since harder clusters are given preference via smaller distances, softer clusters will be combined with harder ones before they combine with themselves. This property ensures that if there is collinear splitting of the jet or soft emmissions—both of which occur frequently during hadronization—the algorithm will be able to recognize that this activity belongs to the same jet. This insensitivity to such soft and collinear changes makes the algorithm infrared and collinear safe. In the language of perturbative QCD, it ensures that this method of constructing jets leads to finite cross sections at all orders in perturbation theory. The anti-k_T algorthm also ensures that hard jets will be conical in shape, which matches the general expected shape from well-collimated hadronization.

5.1.6 Muons

In addition to being used in the triggering of events in the analysis, muons also enter into $E_{\mathrm{T}}^{\mathrm{miss}}$ calculations and are used for vetoing cosmic rays. Muons are first reconstructed independently in the ID and the MS and then combined in a systematic manner [84]. The ID reconstruction is performed in the same manner as other charged particles like electrons, as discussed in Section 5.1.4 and Refs. [87, 88]. In the MS, tracks are constructed based on matching between hit segments in the different MS subdetectors (MDTs, RPCs, TPCs, and CSCs). Each subdetector has a dedicated algorithm to construct the segments therein. A segment-seeded combinatorial search algorithm then assembles the different segments into a candidate muon track. Several refinements and consistency checks are performed, accounting for potential track overlapping, false hits, and missed hits.

Four muon types are then defined based on the subdetectors used in both the ID and MS track reconstructions. Combined (CB) muons have tracks produced from a global refit of hits from both ID and MS tracks, after these tracks are matched via either extrapolation inward from the MS or outward from the ID. Muons are labelled as segment-tagged (ST) if an ID track extrapolated to the MS matches with at least one MDT or CSC segment. ST muons generally consist of muons passing through just a single layer of MS detectors due to either low p_{T} or having passed through regions of low MS acceptance. Calorimeter-tagged (CT) muons have ID tracks that are matched to calorimeter activity consistent with

minimum-ionizing radiation, which is the expectation for most muons of interest. Compared to other muon types, CT muons are relatively impure; however, they are primarily used to recover muons in regions where the MS lacks full instrumentation due to the presence of services for other detectors ($|\eta| \lesssim 0.1$). Finally, extrapolated (ME) muons use only MS tracks that, once extrapolated inwards, are consistent with an origin near the IP. ME muons are generally used to recover muons in the region $2.5 < |\eta| < 2.7$, for which there is no ID coverage. Overlap between the four types is resolved by preferring first muons of type CB, then ST, then CT, while ME overlaps are resolved by assessing fit quality.

Muon ID defines several quality cuts based on the consistency of measurements in the ID and MS and the goodness-of-fit of the full track in order to achieve high efficiency and high rejection of backgrounds—mostly from pion and kaon decays. In this analysis the so-called medium muon ID is used, which is designed to minimize the systematic uncertainties related to reconstruction and calibration. Medium ID considers only CB and ME tracks and applies a series of quality cuts to achieve an efficiency of about 95-96%, while rejecting over 99.6% of hadronic backgrounds.

Muon isolation requirements aim to suppress muons produced in jets from semileptonic decays, which are less isolated than those from the target W, Z, and Higgs boson decays. In the typical manner, track and calorimeter isolation measurements are combined into fixed working points for use in different analysis contexts, as appropriate.

5.1.7 Missing Transverse Energy

As previously discussed in Section 4.1, the total transverse momentum in collisions is essentially zero, since the colliding beams have little momentum in this direction, thanks to their highly focused nature. Any non-zero $E_{\mathrm{T}}^{\mathrm{miss}}$ can indicate the presence of undetected particles. These particles include SM neutrinos, which pass through the entire detector without any detectable interactions. Practically, $E_{\mathrm{T}}^{\mathrm{miss}}$ can also arise from instrumental effects, imperfect calibration, and incomplete detector coverage. However, it may also contain contributions from unknown particles, including those predicted in many BSM physics models.

$E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the negative vector sum of the transverse momenta of fully reconstructed hard objects and the remaining soft, hadronic activity associated to the PV [95].

Hard objects include the selected electron, photon, muon, and jet objects in the event. The soft term accounts for soft hadronic activity from tracks in the ID that are consistent with an origin at the PV but that are not matched to any hard objects. Equation 5.3 schematically summarizes the different $E_{\mathrm{T}}^{\mathrm{miss}}$ contributions:

$$\begin{aligned} \overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}} &= -\sum_{e,\gamma,\mu,\mathrm{jets}} \overrightarrow{p_{\mathrm{T}}} - \sum_{\mathrm{unused\ soft\ tracks}} \overrightarrow{p_{\mathrm{T}}} \\ E_{\mathrm{T}}^{\mathrm{miss}} &= |\overrightarrow{E_{\mathrm{T}}^{\mathrm{miss}}}| = \sqrt{(E_x^{\mathrm{miss}})^2 + (E_y^{\mathrm{miss}})^2} \end{aligned} \tag{5.3}$$

Double counting of momentum from reconstructed objects and soft tracks is accounted for via a signal ambiguity resolution procedure, which establishes rules for adding up different contributions from different objects and ensuring that they consist of separate signals. Details can be found in Ref. [95].

5.1.8 Overlap Removal

As already discussed in the context of clustering and $E_{\mathrm{T}}^{\mathrm{miss}}$ calculations, care must be taken to ensure that detector activity is attributed to only one object to avoid double counting. Thus, it is necessary to implement a series of rules known as overlap removal rules that depend on the algorithm efficiencies and the analysis context in order to determine which ID will be chosen.

For the analysis considered in this thesis, the following rules are implemented, prioritizing photons and otherwise following standard recommendations and maintaining consistency. First, electrons overlapping with photons within $\Delta R < 0.4$ are removed. Then, jets overlapping with photons ($\Delta R < 0.4$) and electrons ($\Delta R < 0.2$) are removed. Electrons overlapping with the remaining jets ($\Delta R < 0.4$) are removed to match the requirements imposed when measuring isolated electron efficiencies. Finally, muons overlapping with photons or jets ($\Delta R < 0.4$) are removed. The resulting, non-duplicated set of physics objects then enter the full event selection described in Chapter 6.

5.2 Calorimeter Pointing & Timing Measurements

Photon pointing and timing, as measured by the LAr calorimeter, provide the main discriminating variables for the present analysis. Pointing and, to an even greater extent, timing are measured quite precisely, with heavily suppressed tails for background processes. Additionally, background processes have very little correlation between these two variables. On the other hand, photons produced from LLPs decaying significantly away from the PV are likely to have both high pointing and timing measurements. Thus, there is very high sensitivity in the extreme tails of both distributions, where background is heavily suppressed but signal can be abundant. The nature of pointing and timing measurements is detailed below.

5.2.1 Calorimeter Pointing

Calorimeter pointing is computed from measurements in the LAr calorimeter and is entirely tracker-independent—a key feature allowing for increased sensitivity to more highly displaced photons that do not rely on tracker-based hit requirements. Pointing relies on the longitudinal segmentation of the LAr calorimeter and is computed from the position of the cluster barycenters in the front and middle layers. The calculation is done entirely in the z-R plane in order to exploit the excellent resolution of the calorimeter cells in the η direction due to fine segmentation in this direction, especially in the front layer[7]. The calculation consists of an extrapolation in the z-R plane of the line connecting the two cluster barycenters to the z-axis. This yields a formula for the pointing value z_{poiut} of the form shown in Equation 5.4.

$$z_{\text{point}} = \frac{z_1 R_2 - z_2 R_1}{R_2 - R_1} \tag{5.4}$$

[7]On the other hand, the ϕ measurements are much poorer, especially in the front layer. The emphasis on the η direction is by design. The distribution of collisions in space is extremely narrow in x and y due to beam focusing, but it is much larger in z. Thus, since many collisions occur in each bunch crossing, identifying the z coordinate is sufficient to identify the correct vertex. This capability provided by calorimeter pointing allows for improved PV identification, particularly for $H \to \gamma\gamma$ decays with unconverted photons—for which pointing with the tracker is not possible—where proper PV identification is associated with improved mass resolution.

where the 1 and 2 subscripts denote coordinates in the first and second layers, respectively. The R coordinates used in this calculation have a very small correction applied to them to account for the position of the beam in the x-y plane, which is known very precisely[8]. A diagram illustrating the pointing computation in the context of the signal considered in this thesis is shown in Figure 5.3.

In the present analysis, in which photons are produced in decays that occur some significant distance from the PV, the true variable of interest is not the raw z_{point} value but rather the difference between z_{point} and z_{PV}. This subtraction operation then establishes pointing as a proxy for the displacement of photon production. For background processes, in which the photons are produced at the PV, pointing is expected to be consistent with zero. Detector and instrumental effects contribute noise to such measurements and, in fact, give rise to significant non-Gaussian tails in the pointing distribution. However, the core of the pointing distribution has a resolution of $\sim$20 mm for prompt photons.

Since the sign of the pointing value is physically insignificant (due to the detector's symmetry about $z = 0$), the pointing value is often taken in absolute value. For the remainder of this thesis, the term pointing will refer to the absolute value after shifting by z_{PV}, and it will often simply be denoted as $|z|$ or just z.

5.2.2 Calorimeter Timing

The LAr Calorimeter provides a timing measurement which, after calibration, serves as the most powerful discriminating variable for distinguishing displaced from prompt photons. Here, the nature of the measurement, the offline timing calibration procedure, and the smearing of simulated data is discussed in detail.

5.2.2.1 Timing Measurement

As discussed in Section 4.3.1.2, the arrival time of the LAr ionization pulse in a given cell is computed using OFCs, which apply the optimal weights to each of the four measured pulse

[8]The beams are highly focused in the x- and y-directions, and during collisions, the position of the beam in these directions is regularly measured to ensure efficient collisions. The correction is of the form $R' = R - R_{\text{beam}} \cos \Delta\phi_{\text{beam}}$. The size of this correction is generally negligible, as $R_{\text{beam}}/R \lesssim 10^{-3}$.

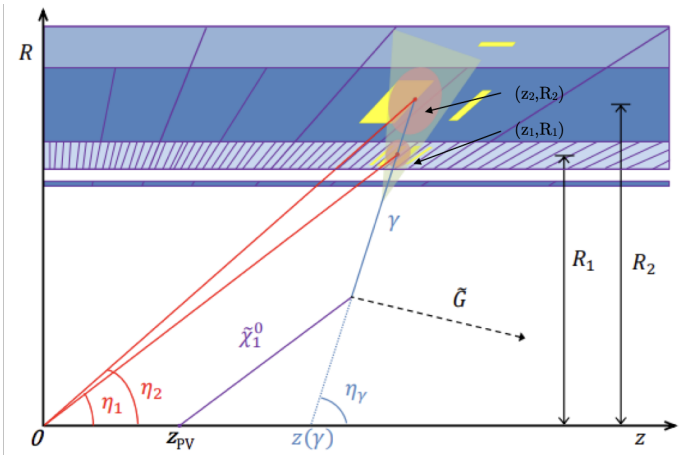


Figure 5.3: Schematic of the calorimeter pointing computation. The position of the PV is indicated on the z-axis. The decay of a long-lived neutralino to a displaced gravitino and photon is depicted with the calorimeter energy deposits depicted in yellow and the cluster barycenters indicated with arrows.

samples [9]. There are three separate readout gains known as (in order of increasing energy scale) low, medium, and high gains. Since these gains have different readout chains in the electronics, they are treated independently for timing considerations.

For the purposes of determining the arrival time of a given EM object, only the timing from the cell in the middle layer with the maximum energy deposit (E_{cell}) is used. There are three reasons for this simple choice. First, most of the energy in an EM cluster is deposited in the middle layer, and in fact, generally ∼15–40% of the energy in the entire EM cluster is deposited in the $\Delta\eta \times \Delta\phi = 0.025 \times 0.0245$ middle layer cell with the maximum energy. Thus, this cell alone contains a high proportion of the information about the object. Second, not only do cells in the middle layer have higher energy, but they also have less cross-talk than the front layer cells, which are very finely segmented in η. It is also worth noting that cells in the presampler and back layers generally do not have sufficient energy depositions for precision timing, so the middle layer cells are the most viable option for high-quality timing information. And thirdly, although in principle, other cells contain additional information that may be combined in some optimal way, studies have shown no significant improvement

[9] OFCs are optimized for prompt signals, and there is a small timing bias for non-prompt signals. For delays relevant to this analysis, the effect is ∼10% and can be accounted for as a systematic uncertainty. See Section 8.2.1 for further discussion.

in the timing resolution with this method, likely due to cross-talk effects.

5.2.2.2 Timing Calibration

The LAr calorimeter achieves an online timing resolution of ∼1 ns. This level of precision is sufficient for determining the bunch crossing (spaced by 25 ns) for triggering. However, the timing performance can be greatly improved offline via a calibration procedure to account for known kinematic and detector effects [3].

Timing is calibrated offline via a series of passes to synchronize cells and to improve resolutions, using electrons from $W \to e\nu$ and $Z \to ee$ events for calibration and validation, respectively. Studies have shown that (as expected from MC simulations) electrons and photons behave similarly for timing purposes. There are only small deviations at very low energies, which can be corrected, as will be discussed later in this section.

Calibrations are obtained separately for high and medium gains, but there is insufficient low gain data to obtain accurate low gain corrections. In the search presented in this thesis, however, the photons from the SM Higgs decay are expected to be relatively soft and thus very few low gain signal events are expected.

In this calibration procedure, anomalous behavior in individual channels such as large mean timing shifts without corresponding changes in OFCs is flagged. Such behavior is thought to be due to hardware failures. The small handful of about 10 such channels are excluded from the calibration, and corrections are not provided.

There are a total of seven independent calibration passes that are applied sequentially to empirically correct averaged or fitted variations in timing. The first pass applies a correction for the time-of-flight from the PV to the cell (calculated geometrically) as well as a zeroing out of the average time per run. Then, there are two hardware-level corrections: one for the average time per FEB and one for the average time per channel. The channel correction in fact has the most significant effect on the timing resolution, due to subtle, individual variations in the electronic components in the electronics readout chain. Next are passes that remove timing variations with a smoothing function based on energy (in the maximum energy cell) and variables associated with cross-talk. As a proxy for middle-layer cross-talk, $\delta\eta$ and $\delta\phi$ are used, where δ indicates a difference between the respective coordinate

of the EM cluster and the center of the maximum energy cell. As a proxy for inter-layer cross-talk, the fractions of energy deposited in the front and back layers are used. Then, another pass correcting for individual channel differences is performed since subtle patterns re-emerge after applying the other passes. The reason for this re-emergence stems from the correlations between the variables used in the different passes, which are not accounted for due to the independence of the passes. Finally, as previously mentioned, small differences are seen between the timing behavior of electrons and photons at very low energies. This phenomenon is understood to be a consequence of subtle differences in shower development between electrons and photons, including bending effects on electrons in the magnetic field. These differences consist of mean timing shifts of photon timing to $\sim$ 100–200 ps lower than that for electrons when $E_{\text{cell}} \lesssim 10$ GeV. To correct for this deviation, a smoothed fit to the cluster energy profile[10] is performed and subtracted out from the timing only for photon objects, centering the times around zero and bringing them into reasonable agreement with electrons. More details can be found in Appedix A Section A.3. After all of these passes, the variations in timing for all of these variables is relatively flat, as desired.

A series of statistical checks was implemented into the calibration procedure to ensure the reliability of the corrections. For the passes involving the zeroing out of simple average (by run, FEB, and channel), the uncertainty on the mean from the calibration sample was required to be less than 100 ps. For the passes implementing a smoothing function of the timing dependence (energy and cross-talk passes), a rebinning procedure was carried out, weighting the bins by their uncertainty for the smoothing fit. Then, for any inputs for which the uncertainty is too large or the input for a given pass lies outside the fit range, the correction is considered invalid. These checks result in a small efficiency drop with 99% (93%) of high (medium) gain events in the validation sample passing all checks. They also provide valuable protection against statistical fluctuations and possible unexpected behavior in unstudied regions that could artificially "correct" times out of the prompt regime.

The timing resolution σ_t is expressed as a function of E_{cell}. Analogous to the form of the conventional energy resolution curve as a function of energy for a sampling calorimeter,

[10]Cluster energy is used here instead of E_{cell} since the behavior of different samples of varying photon purity was more consistent as a function of cluster energy rather than E_{cell}.

there are three terms to consider: the noise, stochastic, and constant terms. The noise term proportional to $1/E_{\text{cell}}$ is attributable to noise in the readout electronics, which depends on the specific design of the full chain and all components. The stochastic term proportional to $1/\sqrt{E_{\text{cell}}}$ is important for the energy measurements of sampling calorimeters since the presence of passive material makes the fraction of the full energy deposited in the active medium a stochastic quantity. This property arises from the fact that the amount of collected signal depends on the number of ionized particles collected in the active medium, which in turn depends on fluctuations in the shower development and the nature of the sampling. However, the timing measurement does not have such high sensitivity to stochastic fluctuations as energy does. This conclusion is essentially due to the fact that the arrival time of the shower does not depend on the total amount of collected signal (as energy does) but rather on the determination of the ionization pulse peak—a relative measure independent of sampling. Thus, the stochastic term is considered negligible for timing performance. Finally, the constant term encompasses any other detector effects that are not energy-dependent. For timing at a collider like the LHC, this constant term includes variation in collision times of protons from the finite length of the colliding bunches. Other contributions in the constant term are understood to come from the intrinsic performance of the LAr calorimeter and from the calibration procedure itself. The dependence of the timing resolution σ_t as a function of E_{cell} is then given by:

$$\sigma_t(E_{\text{cell}}) = \frac{p_0}{E_{\text{cell}}} \oplus p_1 \tag{5.5}$$

where $\oplus$ denotes addition in quadrature and p_0 and p_1 are the noise and constant term coefficients, respectively.

The typical timing resolution as a function of E_{cell} is shown is Figure 5.4. From this figure, it is clear that the energy dependence is well modeled by the expected form of Equation 5.5 for both high and medium gains. It is also worth noting that the timing resolutions across measured energies are markedly less than the delays of several nanoseconds that dominate the sensitivity to many proposed signal models—like those examined in this thesis.

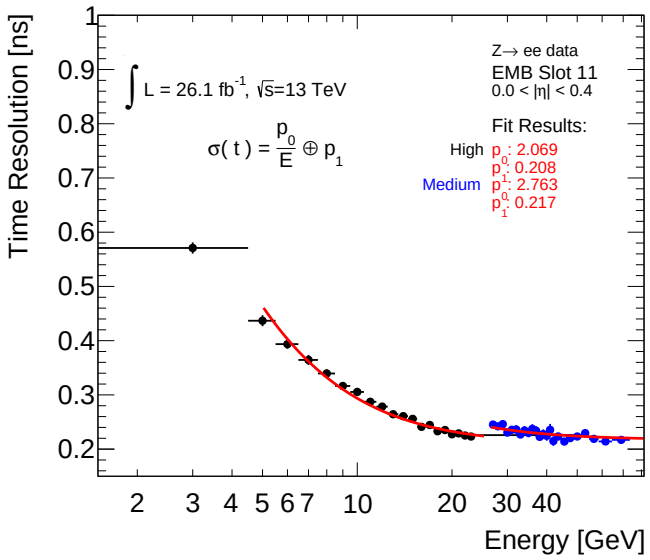


Figure 5.4: Timing resolution versus E_{cell} for high and medium gain in the range $0 < |\eta| < 0.4$ for a subset of Run 2 data. Superimposed in red are the results of fits to Equation 5.5. Qualitatively similar results are obtained for all Run 2 data and detector regions [3].

5.2.2.3 Timing Smearing

MC simulations used to generate data modeling potential signal models do not properly model the full timing behavior. Some modeling of the electronic noise is included, but it does not fully account for all sources of noise, including global effects that are seen in the calibration. It also omits the timing spread from the beam. Therefore, additional MC smearing must be added to simulated MC data to match the resolutions observed in data after the calibration described in Section 5.2.2.2. As will be discussed in more detail below, the timing measurement can be thought of as containing both a correlated and an uncorrelated component. To properly smear times in MC events with multiple EM objects from the same PV, both components must be determined[11].

After calibration with single-electron $W \to e\nu$ events, the two-electron $Z \to ee$ events can be used to extract the correlated component of the timing spread due to the variation

[11]For events in the analysis presented in this thesis, only the timing information from a single photon in each event is considered, so the distinction between correlated and uncorrelated components is not relevant. However, to properly model the smearing in general, this distinction must be made.

in collision times. In these Z decays, both electrons share the same collision time since they originate from the same PV. Thus, by studying the sum and difference between the measured time of the two electrons in these events, the correlated and uncorrelated components may be determined. These studies indicate that the correlated component attributable to the LHC beam spread is ∼190 ps, in good agreement with expectations based on the performance of the accelerator. Comparing this value to the resolutions seen in Figure 5.4, it can be seen that in the high-energy limit, the resolution is dominated by the beam spread, as the curves flatten out to a value of ∼ 210–220 ps.

In order to tune the additional smearing to match calibrated data, a correlated timing contribution in the form of a single Gaussian random sample per event is added to account for the beam spread. Then, the energy-dependence of the timing resolution is compared with $Z \to ll\gamma$ photon data in each barrel slot per gain to produce a binned profile of the additional uncorrelated smearing that must be added. The quadrature difference between the two curves is used as the required additional uncorrelated smearing component to be added to MC in order to achieve closure with data. Details of the tuning of the smearing can be found in Appendix A.4.

Chapter 6: Event Selection

The signal event selection aims to select events with displaced photons from Higgs boson decays, such as those that would result from the signal model described in Section 2.6, while filtering as much background as possible. An electron or muon from the associated Higgs boson production with W bosons, Z bosons, or $t\bar{t}$ is used for triggering due to the expected softness of the photons. Thus, the final state contains at least one electron or muon, one photon, and $E_{\mathrm{T}}^{\mathrm{miss}}$. Specific additional requirements for each of the objects involved in the analysis are described in detail in the following sections.

6.1 Event Cleaning

Before analysis-specific selections are applied, a set of standard event cleaning cuts filter out events in which issues with data collection or quality are observed [96]. These cleaning cuts ensure that data was collected when the LHC beams were operating as expected and when the detector was fully operational. Proper detector operation depends on the status of individual subdetectors, which may occasionally experience temporary down-times, excessive noise, or data corruption. Standard data quality checks also ensure that there are no major data collection anomalies.

6.2 Object Selections

In order to select physics objects of sufficient quality for this analysis to achieve high signal efficiency and background rejection, the following photon and lepton selections are made. A summary of these selection can be found at the end of this section in Table 6.1

6.2.1 Photons

The loose photon ID was chosen for this analysis, based on studies carried out in the previous non-pointing photon analyses from Run 1 that found a significant drop in efficiency for tight compared to loose ID for photon objects that were non-pointing [65]. This drop can be understood to be a consequence of the additional cuts on the strip layer variables for tight ID not required for loose ID. The strip layer, which is finely segmented in η, is especially sensitive to changes in the "non-pointingness" of EM objects, and distortions in the shower shape are thus more apparent, resulting in a large efficiency drop.

Due to poorer pointing resolutions for photons in the endcaps, the tendency for the relatively low energy of photons considered in this analysis to be in the barrel, and other issues associated with the end-cap regions, only photons in the barrel are considered. It is, however, not required that all photons in the event be in the barrel, only that at least one of them that passes all other cuts be there.

As discussed in Section 5.1, isolation requirements are imposed on the loose photons to increase purity by ensuring that the cluster does not contain excess activity in the surrounding area—a potential sign that the cluster may be a jet rather than a photon. The selected calorimeter isolation requires that topo-clusters within a fixed $\Delta R = 0.2$ cone, excluding the core, be less then 6.5% of the total cluster p_{T}. Track isolation requires that the fixed-radius $R = 0.2$ cone—excluding the photon candidate itself—must contain less than 5% of the total object p_{T}. Several other standard cleaning and quality cuts are applied to photon candidates to mitigate effects from noise and pileup [90]. After passing these ID and isolation requirements, photons are then selected if they have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.37$ (excluding the crack region between barrel and end-cap[1]).

Photons likely originating from out-of-time pileup (i.e. from neighboring bunch crossings) become evident in the timing distribution of calibrated cluster times beyond $|t| > 17.5$ ns and may even extend down to $|t| > 12.5$ ns in certain kinematic regimes. Additionally, the modeling of delayed particles in simulation is only valid for $|t| < 12.5$ ns, after which point simulated signals are not properly handled, and the timing information cannot be

[1]Objects are defined to be in the crack region if the pseudorapidity of its second-layer cluster barycenter is between 1.37 and 1.52.

trusted. Out of caution for other issues that may be caused by this mismodeling, the limit for EM object times allowed in this analysis is moved slightly lower such that the final requirement on calibrated photon times is chosen to be $|t| < 12$ ns. Despite placing the cut so far from the hard limit of 25 ns at the next bunch crossing and limiting the extent to which more significantly delayed photons can be considered, studies have provided evidence that in fact the sensitivity gained by extending beyond 12 ns is marginal and may actually decrease due to the contamination from the timing tail of this other bunch crossing. More details on the simulation mismodeling and the associated significance studies can be found in Appendices A.1 and A.2.

From the timing calibration, it is known that the photon timing resolution has a strong dependence on E_{cell} such that low-energy photons have very high resolutions and thus degrade the analysis sensitivity. E_{cell} is also closely linked to the final state photon kinematics. Therefore, the E_{cell} cut required for the analysis was examined in an optimization procedure described in detail in Section 7.3.1.3. In the end, two different E_{cell} cuts at 7 and 10 GeV are used depending on the NLSP and LSP masses of the signal point being considered.

Finally, photon candidates are required to have had their most energetic middle-layer cell (the one used for the timing measurement) read out with either medium or high gain. The exclusion of low gain reflects the lack of calibrated timing corrections for these events due to low statistics in the calibration sample. However, given the fact that the photons in this analysis are of relatively low energies, virtually no photons of interest are read out with low gain, and thus this requirement has a negligible impact on signal acceptance.

In order to exploit the substantial background reduction observed in events with at least two photons compared to just one, events are categorized into either the one-photon or two-photon channel[2]. The one-photon channel comprises events with only one photon meeting all of the aforementioned selection requirements. The two-photon channel comprises events with at least two photons meeting all selection criteria; however, only the leading barrel photon in this event is used in the fit itself. That is, the photon multiplicity is only used as a useful background-reduction tool—the timing and pointing information from other photons

[2]More details about the optimization performed and the benefit of separating events in this way is discussed in Section 7.3.1.3 and Appendix B

does not enter the analysis.

6.2.2 Leptons

Electron candidates are required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.47$ (again excluding the crack region). They are required to pass the medium electron ID requirements discussed in Section 5.1.4. The selected calorimeter isolation requires that the topo-clusters within a fixed $\Delta R = 0.2$ cone, excluding the core, contain less than 20% of the total cluster p_{T}. For track isolation, the variable-radius cone with a maximum R of 0.2, excluding the target object, must contain less than 15% of the total object p_{T}. The reconstructed track associated to the electron candidate must also be consistent with the PV by requiring that its longitudinal impact parameter z_0 relative to the vertex satisfy $|z_0 \sin\theta| < 0.5$ mm. Additionally, the relative uncertainty of the electron transverse impact parameter d_0 with respect to the beam axis must be less than 5, i.e. $|d_0|/\sigma_{d_0} < 5$.

As discussed in Section 5.1.6, medium muon ID is used in this analysis. Muon candidates are also required to have $p_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.7$. The specific isolation working point used for muons utilizes a particle flow isolation working point[3]. Muons are required to satisfy calorimeter- and track-based isolation requirements that are 95-97% efficient for muons with p_{T} between 10 and 60 GeV and 99% efficient for those with $p_{\mathrm{T}} > 60$ GeV. Finally, muon tracks must satisfy $|z_0 \sin\theta| < 0.5$ mm and $|d_0|/\sigma_{d_0} < 3$.

The leading lepton, i.e. that with the highest p_{T}, regardless of flavor, is additionally required to have $p_{\mathrm{T}} > 27$ GeV in order to achieve high efficiency with the lowest-p_{T} standard lepton triggers available. Subleading leptons must otherwise simply satisfy the aforementioned $p_{\mathrm{T}} > 10$ GeV requirement.

[3]Specifically, the PFlowLoose_FixedRad (FixedCutPFlowLoose) working point. Details and performance can be found in Ref. [86].

	Lepton (e/μ)	Photon
Multiplicity	>=1	>=1
ID	Medium	Loose
Calo. Isolation	e: $p_{\mathrm{T},\Delta R=0.2}/p_{\mathrm{T}} < 20\%$ μ: PFlowLoose_FixedRad (see Ref. [86])	$p_{\mathrm{T},\Delta R=0.2}/p_{\mathrm{T}} < 6.5\%$
Track Isolation	e: $p_{\mathrm{T},\Delta R_{\mathrm{max}}=0.2}/p_{\mathrm{T}} < 15\%$ $\lvert z_0 \sin\theta\rvert < 0.5$ mm $\lvert d_0\rvert/\sigma_{d_0} < 5$ μ: PFlowLoose_FixedRad (see Ref. [86]) $\lvert z_0 \sin\theta\rvert < 0.5$ mm $\lvert d_0\rvert/\sigma_{d_0} < 3$	$p_{\mathrm{T},\Delta R=0.2}/p_{\mathrm{T}} < 5\%$
Leading (Subleading) p_{T} [GeV]	> 27 (> 10)	> 10 (> 10)
$\lvert\eta\rvert$	e: < 1.37 or [1.52, 2.47] μ: < 2.7	At least one: < 1.37 Otherwise: < 1.37 or [1.52, 2.37]

Table 6.1: Summary of the selection criteria for leptons and photons.

Chapter 7: Analysis Strategy

This chapter outlines the analysis strategy, detailing the definition of the different analysis regions, the design of the data-driven background estimation method and its validation, the optimization of the analysis regions and binning, and the results of fit validation studies.

7.1 Analysis Region Definitions

Control (CR), validation (VR), and signal (SR) regions are established to systematically study the background and background prediction strategy and to blind the analysis to the region of interest before finalizing the full strategy to minimize bias. $E_{\mathrm{T}}^{\mathrm{miss}}$ is the primary variable used to distinguish the various regions, with photon timing providing an additional, useful distinction for an even fuller understanding of the fitting strategy.

7.1.1 Missing Transverse Energy

The final state of interest in this analysis contains particles that escape the detector—either neutralinos that have not yet decayed, gravitinos, or neutrinos from W or $t\bar{t}$ decays—and thus carry away energy that can be measured as $E_{\mathrm{T}}^{\mathrm{miss}}$. Therefore, since more $E_{\mathrm{T}}^{\mathrm{miss}}$ is to be expected in signal-like events, a simple set of $E_{\mathrm{T}}^{\mathrm{miss}}$ cuts is used to distinguish between the CR, VR, and SR. The low-$E_{\mathrm{T}}^{\mathrm{miss}}$ CR contains a negligible amount of signal contamination and is thus used to study and understand in detail the SM background processes that pass the analysis event selection. The high-$E_{\mathrm{T}}^{\mathrm{miss}}$ SR is the final region of interest, whose events will be examined to determine whether there is evidence for the signature of interest. The intermediate VR is used to verify an understanding of the background on a set of events that is statistically independent from the CR and that also more closely resembles the SR

kinematically. A successful description of the VR using the CR serves as a key indicator of the statistical power of the background estimation procedure and engenders confidence in its use to describe the SR. The exact ranges of $E_{\mathrm{T}}^{\mathrm{miss}}$ that distinguish the CR, VR, and SR are determined via an optimization procedure that is described in detail later in this chapter in Sections 7.3 and 7.3.1.2.

7.1.2 Photon Timing

One additional observation about the timing of photons in the final state is useful to consider in the context of establishing the various analysis regions. Since signal photons always arrive delayed—never early—and the background, prompt photon distribution is centered at 0 ns, no sensitivity to the signal of interest is expected at negative times, i.e. times measured before the average LHC bunch crossing. The timing resolution for a photon with an energy typical in the context of this analysis is $\gtrsim 200$ ps, so in fact, the sensitivity to photons with measured positive times roughly of this order are also expected to yield negligible sensitivity.

Examining the signal contamination in the negative-timing region confirms this reasoning, where the statistical uncertainty of the background data overwhelmingly dominates any signal contribution. The studies demonstrating this fact can be found in Appendix D.2. Thus, it is safe to consider only positive times for the final signal region and to unblind events with negative times. This observation presents a unique opportunity to unblind all events with negative photon timing—even in the SR. Thus, the so-called negative-timing SR, or $\mathrm{SR}_{t<0}$, can be used as, in effect, a second VR, but with the added benefit that the events in this region are even more kinematically similar to the true SR with positive timing. The timing distribution in $\mathrm{SR}_{t<0}$ is often reflected across $t < 0$ to model the expected background for the positive-timing SR. This operation is justified by the symmetry of the templates, which is demonstrated in Appendix A.6.

7.2 Background Estimation

This section describes the overall fitting strategy using various background templates derived from data. The processing, selection, and handling of uncertainties for the templates is

described in detail, followed by a validation of the strategy on several different samples.

7.2.1 Overview

Background estimation is done in an entirely data-driven manner using templates. Data-driven methods are employed due to the fact that the signal of interest lies in the tails of the timing and pointing distributions, and the modeling in such extreme regimes presents such great challenges that simulated data is not considered trustworthy. It is thus necessary to construct templates that can then be used to fit the background component in the SR.

Background events are expected to arise from electroweak and QCD multijet processes. Each of the three reconstructed objects used in this analysis—photons, leptons, and $E_{\mathrm{T}}^{\mathrm{miss}}$—can arise from either a genuine source or from instrumental effects. Leptons and $E_{\mathrm{T}}^{\mathrm{miss}}$ can originate from leptonic decays of W or Z bosons or of $t\bar{t}$ in electroweak processes. However, jets may also be misidentified as leptons, which also results in artificially high $E_{\mathrm{T}}^{\mathrm{miss}}$ due to a misapplied jet energy scale correction. Since we loose photon identification requirement is applied and a relatively low p_{T} threshold, a significant fraction of selected photons are expected to in fact be fake photons arising from neutral hadron decays or other jet backgrounds. Therefore, background in the SR can be generally described as originating from one of two sources: genuine photons and fake photons.

In order to best exploit the pointing discriminant in the background modelling, selected data events are divided into multiple categories with varying relative quantities of signal, genuine photon background, and fake photon background. In each pointing category, the timing distribution is modeled by an extended probability density function (PDF) with a signal component derived from MC simulation and a background component derived from a data-driven method. When performing a model-dependent statistical test, the normalization of signal between categories is correlated by introducing the signal strength parameter, while the total normalization of background in each pointing category is left as a free parameter, determined by data in the fit.

7.2.2 Template Fit Strategy

While the ultimate source of the background is some unknown mixture of electroweak and QCD processes, the only parameter that impacts the background modeling itself is the composition in terms on photon ID, i.e. genuine and fake photons. As known from Run 1 studies, the timing distribution of genuine EM objects is somewhat narrower than that of fake photons [65, 66]. Since the background in the SR is some unknown mixture of genuine and fake photons, its timing shape can be modeled as a linear combination of two timing distributions with differing genuine and fake photon fractions. Thus, two timing templates are constructed to span the space of genuine- and fake-photon purities. One template is derived from samples with selections designed to increase the fraction of genuine photons; and the other, to increase the fraction of fake photons. The samples used to construct these two templates are entirely orthogonal to the SR.

1. The genuine-photon-enriched region is defined as the sum of two independent data selections. The first selection is the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ CR with a specific photon ID requirement (e.g. medium or tight, see Table 5.1). The level of ID required is an optimization question addressed in Section 7.5. The second selection targets radiative $Z \to \ell\ell + \gamma$ events. This selection requires a single-lepton trigger, two electrons or two muons and at least one loose photon, a dilepton invariant mass $m_{ll} < 83$ GeV, and a three-body invariant mass $81 < m_{ll\gamma} < 101$ GeV. These criteria are adapted from standard radiative Z boson decay selections, as detailed in Ref. [97]. The two templates are shape averaged, that is, they are each normalized, scaled by 0.5, and added together to yield the final template.

2. The fake-photon-enriched region is defined by a different subset of photons in the low-$E_{\mathrm{T}}^{\mathrm{miss}}$ CR, namely those that fail the ID requirement from genuine-photon-enriched region but still pass loose ID requirements.

The overall fit equation that predicts N_{ij}^{data}, the number of observed events in pointing bin i and timing bin j is as follows:

$$N_{ij}^{\mathrm{pred}} = N_i^{\mathrm{bkg}} \left(\alpha_i f_{ij}^{\gamma} + (1 - \alpha_i) f_{ij}^{\mathrm{fake}} \right) + \mathrm{BR}(H \to \chi_1^0 \chi_1^0) N_i^{\mathrm{sig}} f_{ij}^{\mathrm{sig}} \tag{7.1}$$

where f_{ij}^{γ} and f_{ij}^{fake} are the genuine-photon-enriched and fake-photon-enriched templates, respectively, α_i denotes the genuine-photon-enriched template-mixing fraction[1], and f_{ij}^{sig} is the signal shape template. N_i^{bkg} and N_i^{sig} (given a particular BR) scale the background and signal predictions, respectively.

The overall likelihood function $\mathcal{L}$ governing the full fit comprises the product of Poisson probabilities for every timing bin j in each pointing category i, as written in Equation 7.2.

$$\mathcal{L} = \prod_{ij} \text{Pois}(N_{ij}^{\text{data}} | N_{ij}^{\text{pred}}) = \prod_{ij} \frac{\left(N_{ij}^{\text{pred}}\right)^{N_{ij}^{\text{data}}} e^{-N_{ij}^{\text{pred}}}}{N_{ij}^{\text{data}}!} \tag{7.2}$$

For the purposes of optimization, it is conventional to deal with a quantity known as the negative log likelihood (NLL) instead of $\mathcal{L}$ directly.

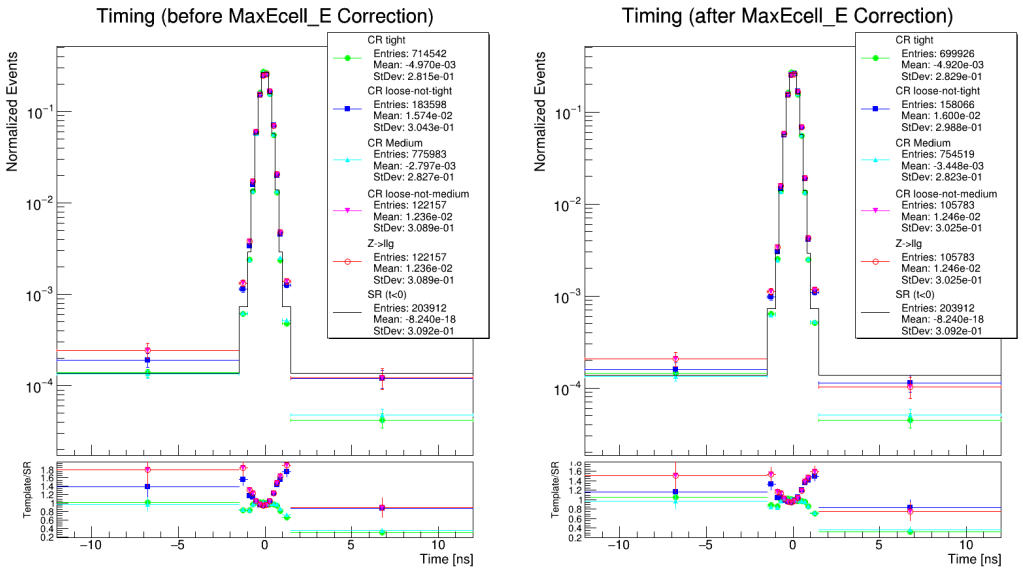
7.2.2.1 Energy Reweighting

From the timing calibration procedure, it is known that the timing resolution of EM objects varies as a function of E_{cell}. Thus, the difference in E_{cell} distributions between the samples comprising the template samples and the SR requires a correction. To achieve this, events in the control samples are reweighted such that their E_{cell} distributions match that of the SR. A reweighting factor is computed as a ratio between a given template and the target region (e.g. SR) in a fixed set of E_{cell} bins, chosen to cover the full E_{cell} range while minimizing statistical uncertainties. Figure 7.1 shows the effect of the reweighting procedure on various template components. The standard deviations are generally slightly reduced, reflecting the tightening of the distributions to account for the slightly harder photons in the high-$E_{\text{T}}^{\text{miss}}$ SR.

7.2.2.2 Mean Shifting

The timing calibration procedure is performed with W and Z boson data in order to correct for the major dependencies of various variables with timing such that these corrections can

[1]That is, the fraction of the total background prediction taken from f_{ij}^{γ}. Note, α_i is simply a nuisance parameter associated with—but not equivalent to—the photon purity of the background prediction, since the two templates are themselves only enriched in their target components. The actual photon purity is, in fact, not relevant to the analysis. It is only important that the two templates can sufficiently span the purity space to describe the SR.



(a) (b)

Figure 7.1: (a) Timing PDFs without E_{cell} reweighting. (b) Timing PDFs after E_{cell} reweighting.

be applied to most electrons and photons independent of their production mode. However, imperfections in this procedure as well as other instrumental and detector effects can have subtle effects on the timing performance. One such instance can be observed in the location of the means of the timing distributions of the various samples used in this analysis. The means of these distributions are relatively small across all samples, with values generally less than a few tens of ps. However, since the fit strategy relies on fine binning with spacing of 200 ps near the cores of the distributions, where the vast majority of the data lies, subtle shifts in the means of these distributions can significantly impact the fit performance. Therefore, each distribution is shifted such that its mean across the full timing range of ± 12 ns is precisely zero. It is worth emphasizing that this simple shift of the distributions is entirely justified in the vein of a second-order timing calibration. The general calibration procedure performs operations of exactly this type, i.e. trying to center the distributions around zero for prompt, background events. Thus, due to the subtle effects described above, the centering of each of the distributions independently before fitting is simply correcting for the subtleties that emerge from the kinematic differences and other complex effects that are unique to each selection.

7.3 Analysis Optimization

As discussed in Section 7.2.1, the final likelihood fit of the analysis is performed over bins of photon timing in each pointing bin. The choice of binning used for both variables affects the distribution of signal and background events in the fit and therefore has a large impact on the final sensitivity. Dedicated studies are performed to determine the optimal binning for both variables in conjunction with the choices for the region definitions based on $E_{\mathrm{T}}^{\mathrm{miss}}$ and and cuts on E_{cell}. Such studies improve sensitivity to the signal of interest, while maintaining sufficient statistics for all bins in the background templates in order to ensure a robust shape fit. A description of this optimization procedure and its results follows.

7.3.1 Overview

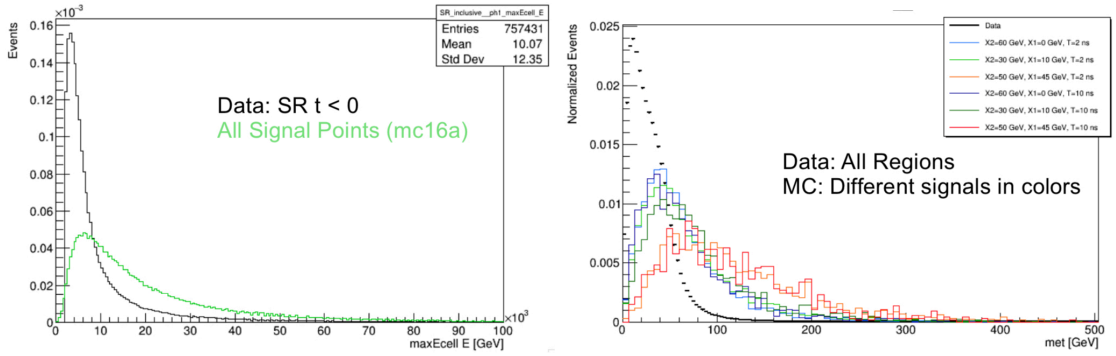
The optimization procedure considers many aspects of the analysis strategy simultaneously and determines the configuration that maximizes sensitivity to the signal model of interest. In all, it optimizes the $E_{\mathrm{T}}^{\mathrm{miss}}$ cuts that define the analysis regions, the E_{cell} event selection cut, the number of timing and pointing bins, and the placement of the edges for each of these bins. An overview of each of these items in the context of the full optimization framework is discussed below.

7.3.1.1 $E_{\mathrm{T}}^{\mathrm{miss}}$ and E_{cell}

$E_{\mathrm{T}}^{\mathrm{miss}}$ and E_{cell} both have particularly strong influences on the pointing and timing shapes and yields. As discussed in Section 7.1.1, $E_{\mathrm{T}}^{\mathrm{miss}}$ arises from escaping gravitinos and/or neutralinos, in addition to any neutrinos produced from W boson or $t\bar{t}$ decays. If the neutralino does not decay before the EM calorimeter, the neutralino itself contributes directly to the $E_{\mathrm{T}}^{\mathrm{miss}}$ of the event. Thus, it is clear that $E_{\mathrm{T}}^{\mathrm{miss}}$ is correlated with both the sparticle masses and lifetimes, and therefore with the timing and pointing shapes of the signal. Additionally, the mass difference between the NLSP and LSP constrains the energy of the associated photon and thus E_{cell}, which has a significant influence on the timing resolution of the photon, particularly at lower energies (see Section 5.2.2). The masses also influence the speed at which the intermediate SUSY particle travels before decaying and producing a photon,

which contributes to a delay in timing and an increase in non-pointingness of the resulting photon.

In addition to these effects on the shapes of the timing and pointing distributions, the E_{cell} and $E_{\text{T}}^{\text{miss}}$ cuts also greatly affect the statistics in the analysis regions and in the background templates. As demonstrated in Figure 7.2, the vast majority of background data events are concentrated at relatively low $E_{\text{T}}^{\text{miss}}$ and E_{cell} values. Since the background prediction is entirely data driven, raising these cuts improves signal yield relative to background yield but also depletes template statistics, which increases the uncertainty in the background prediction, counteracting the gain in sensitivity. Any gain in template statistics from a lower E_{cell} cut must also be balanced against the rapidly deteriorating time resolution, which widens the background templates and reduces sensitivity in the tails.



(a) (b)

Figure 7.2: Comparison of the E_{cell} and $E_{\text{T}}^{\text{miss}}$ distributions for data and signal MC. (a) shows compares the E_{cell} distributions for the $\text{SR}_{t<0}$ region (black) and an average of many signal points (green). (b) compares all analysis regions (black) and different individual signal models (different colors).

To account for the interplay of all of these factors, the optimization of $E_{\text{T}}^{\text{miss}}$ and E_{cell} is performed simultaneously with the optimization of the timing and pointing binning (to be discussed in Section 7.3.1.3).

7.3.1.2 Signal Mass Splittings

By considering the space of NLSP and LSP masses, an important observation can be made regarding kinematic similarities between different regions in the mass plane. The so-called

mass splitting between NLSP and LSP masses ($\Delta m \equiv m_{\text{NLSP}} - m_{\text{LSP}}$) has important connections to $E_{\text{T}}^{\text{miss}}$ and E_{cell}. For signal points with lower mass splittings, there tends to be relatively higher $E_{\text{T}}^{\text{miss}}$, leaving little energy for the associated photon, which is then relatively soft. These points would thus be expected to benefit from a higher SR $E_{\text{T}}^{\text{miss}}$ cut and a lower E_{cell} cut. On the other hand, for signal points with higher mass splittings, the opposite is true, and a lower $E_{\text{T}}^{\text{miss}}$ cut and higher E_{cell} cut would be preferred. Due to these opposing trends, an optimization was performed, which found significantly improved results by establishing two separate $E_{\text{T}}^{\text{miss}}$ and E_{cell} cuts for two particular signal groupings with respect to sparticle mass splittings, namely $\Delta m \leq 10$ GeV and $\Delta m > 10$ GeV. The study to determine the optimal choice of how to group signal points is described in Appendix B Section B.3. For signal points with $\Delta m \leq 10$ GeV, a requirement of $E_{\text{cell}} > 7$ GeV is used, and the SR is defined by $E_{\text{T}}^{\text{miss}} > 80$ GeV. For signal points with $\Delta m > 10$ GeV, the E_{cell} requirement is raised to > 10 GeV, and the SR is defined by $E_{\text{T}}^{\text{miss}} > 50$ GeV. In both cases, the CR and VR are defined identically, with $E_{\text{T}}^{\text{miss}}$ ranges of <30 GeV and 30–50 GeV, respectively. Table 7.1 summarizes these optimized cuts for the two mass splitting analyses.

Requirement	**Δm $\leq$ 10 GeV**	**Δm $>$ 10 GeV**
E_{cell} [GeV]	7	10
CR MET [GeV]	<30	<30
VR MET [GeV]	30–50	30–50
SR MET [GeV]	>80	>50

Table 7.1: Optimized $E_{\text{T}}^{\text{miss}}$ and E_{cell} cuts describing the different analysis region definitions and distinctions based on signal point mass splittings $\Delta m \equiv m_{\text{NLSP}} - m_{\text{LSP}}$.

It is important to note that since these two groupings by mass splitting entail different event selections, they comprise entirely independent analyses that require their own, separate strategies and validation procedures. More details about the optimization will be discussed in Section 7.3, where the full framework is discussed. However, the results for the division of the analysis into these two, separate mass splittings is important for the understanding of the background estimation and template choice studies in Section 7.2.

7.3.1.3 Pointing & Timing Binning

The optimization procedure employs a simplified version of the background estimation procedure described in Section 7.2 in order to analyze the expected significance as a function of different cut and binning configurations. For each of the two, independent mass splitting analyses, there are six types of choices to be made: the $E_{\mathrm{T}}^{\mathrm{miss}}$ cuts defining the CR and VR, the E_{cell} cut, the number of timing bins, the number of pointing bins, and the placement of the individual bins in both timing and pointing.

In principle, increasing the number of bins and combining their individual significances should always increase total significance or leave it unchanged. However, one must also consider that using too many bins increases the statistical uncertainty of the background templates and can even leave some bins empty, leading to a background prediction of zero events for which the uncertainty is ill-defined. The same argument applies to $E_{\mathrm{T}}^{\mathrm{miss}}$ and E_{cell} cuts, as discussed in Section 7.3.1.1. Accordingly, the optimization procedure simultaneously considers all of the aforementioned factors involving both timing and pointing binning and event selection cuts.

For simplicity, a full template fit was not performed to make a background prediction in the signal region. Instead, a template mixing fraction of $\alpha = 0.5$ (i.e. equally mixing the two normalized templates) was used, after studies showed that there was generally little dependence of the results on α (see Appendix B.6).

7.3.2 Optimization Procedure

The input to the optimizer is the two-dimensional pointing verus timing histograms[2] for the background templates and the signal—with events having positive timing and $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV blinded. For each possible choice of E_{cell} and $E_{\mathrm{T}}^{\mathrm{miss}}$ cuts, the templates are reweighted by E_{cell} in order to account for the known energy-dependence of the timing resolution (see Section 5.2.2.2). The normalized sum of the two templates is then scaled to the total number of events in the $\mathrm{SR}_{t<0}$ to get the total background prediction. The signal templates from MC were scaled to a chosen value of $\mathrm{BR}(H \to \chi_1^0\chi_1^0)$ to obtain the signal prediction. For

[2]The histograms have 0.1 ns granularity in timing and 10 mm in pointing.

the purposes of this optimization study, a BR of 5% was used as a reasonable target[3].

The input templates were separated by the two analysis channels, namely one- and two-photon channels. Initially, studies were performed by making an additional distinction by the number of leptons in the event. However, it was found that removing this distinction left the total significance roughly the same and eliminated some issues with a lack of statistics in the two-lepton channels, which were suppressed by a factor of ∼20 relative to the respective one-lepton channels. More details on dropping the lepton number distinction are presented in Appendix B.1.

In order to simplify the optimization procedure and to avoid overfitting individual signal points, all available signal points with lifetimes 2 and 10 ns were averaged together. The averaging was performed by simply adding together all events and dividing by the total number of signal points, thus creating an "average signal" with the distributions from higher-yield points contributing relatively more to this average[4].

One additional nuance to the optimization procedure lies in the need for fine timing binning near 0 ns. This need arises from studies that indicate that most of the information about the shape of the timing distribution lies in the core. Therefore, fine sampling in low-timing bins allows for a more robust fit. To satisfy this requirement, timing edges at 0, 0.2, 0.4, 0.6, 0.8, and 1 ns are fixed before optimization.

The optimization runs separately in a brute-force manner over every combination of E_{cell}, upper CR $E_{\text{T}}^{\text{miss}}$ cut, and lower SR $E_{\text{T}}^{\text{miss}}$ cut. The full search space is outlined in Table 7.2. Within each choice of these cuts, a number of allowed bins in timing and pointing is chosen, and an optimization of the placement of the bin edges is performed. The full available search space of all possible timing bin configurations and all possible pointing bin configurations is impractically large to explore completely. Thus, a number of reasonable simplifications are made.

First, the placement of the bin edges is restricted to multiples of 0.5 ns in timing (beyond

[3]The effect of the chosen BR is generally fairly subtle for the purposes of determining the optimal analysis configuration. Differences typically do not arise unless signal statistical uncertainties dominate. However, this situation is not observed in the samples tested.

[4]See Appendix B.2 for a discussion of the justification for using such an average signal in lieu of the individual signal points.

Parameter	Allowed Values
E_{cell} Cut [GeV]	2, 3, ..., 10
Upper CR $E_{\text{T}}^{\text{miss}}$ Cut [GeV]	30, 40
Lower SR $E_{\text{T}}^{\text{miss}}$ Cut [GeV]	50, 60, 70, 80
$n_{\text{bins,t}}$	6, 7, ..., 11
$n_{\text{bins,z}}$	1, 2, ..., 6

Table 7.2: Optimization search space, detailing the allowed values of the floating parameters. Within each choice of $n_{\text{bins,t}}$ and $n_{\text{bins,z}}$, an independent optimization of the placement of bin edges is performed.

1 ns) and 50 mm in pointing. In addition to shrinking the parameter space, this also prevents the creation of bins that may be smaller than the observed resolutions (∼0.3 ns in timing and ∼20 mm in pointing), which may in turn help to prevent overfitting.

Secondly, the placement of the bins is done sequentially rather than simultaneously. That is, the optimizer iterates over all possible combinations of a single timing and a single pointing bin edge placement and places this pair of edges wherever sensitivity is maximized. A simplified illustration of this process in the pointing dimension is shown in Figure 7.3. The optimizer iterates in this way, accumulating bin edges sequentially until it reaches the maximum number of allowed timing or pointing edges being considered, whichever is smaller. Then, once there are no more simultaneous optimizations possible, it performs one-dimensional optimizations in timing or pointing until there are no more edges to be placed.

The significance formula used by the optimizer is a standard asymptotic formula for exclusion significance, modified to fold in non-negligible shape uncertainties in the background and signal shapes [98]. The full significance formula is given in Equation 7.3.

$$Z' = \sqrt{\sum_{i,j} \left(\frac{-1}{2\left(B_{ij}\log\left(1+\frac{S_{ij}}{B_{ij}}\right) - S_{ij}\right)} + \frac{B_{ij}}{S_{ij}^2} + \frac{1}{S_{ij}} \right)^{-1}} \tag{7.3}$$

where $S_{i,j}$ and $B_{i,j}$ are the number of signal and background events, respectively, in pointing bin i and timing bin j. This formula is discussed in more detail in Appendix B.4.

The asymptotic form of the first term of Equation 7.3 accounts for the wide variation in the relative abundances of signal and background that occur in the different pointing

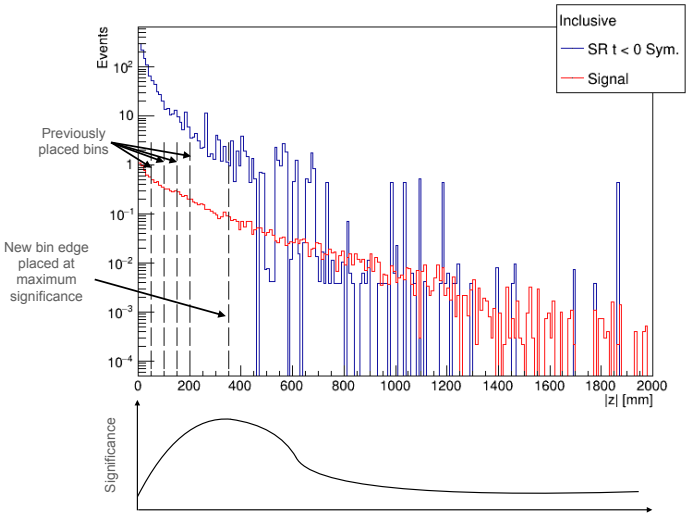


Figure 7.3: A simplified illustration of the bin edge placement strategy for sensitivity optimization, showing the data and signal events as well as the significance corresponding to placing a new edge at the given $|z|$ pointing value. The desired new edge is then placed at the location at which the maximum significance occurs. In the actual implementation, this iteration is done simultaneously in timing and pointing.

and timing bins. The added signal and background shape uncertainty terms correct for assumptions made about low uncertainties in these shapes, which are violated in very low-yield high pointing and high timing bins.

After obtaining the results of the optimization procedure, the binnings are slightly adjusted by hand to harmonize bin edges and to achieve a more intuitive and simple binning strategy. After making these minor adjustments, it was verified that the significance was within ∼5% of the optimal significance for each of the individual signal points.

7.3.3 Optimization Results

The final analysis binning following from the full procedure described above is detailed in Table 7.3. The results generally highlight the importance of raising the E_{cell} and $E_{\text{T}}^{\text{miss}}$ cuts to take advantage of their significant background-rejection power. Then relatively coarse bins in pointing and timing are placed with the available remaining statistics in each photon channel. The different mass splittings call for different E_{cell} and SR $E_{\text{T}}^{\text{miss}}$ thresholds, with higher mass splittings having higher E_{cell} and $E_{\text{T}}^{\text{miss}}$ cuts, as expected from the final state kinematics. The different photon channels differ only in one additional bin edge in pointing

at 1.5 ns for the one-photon channel. Such a distinction reflects the relatively higher yield in the one-photon channel, which allows for the placement of this extra edge. Other than these distinction, the rest of the cuts and bins are fully harmonized.

A comparison of the sensitivity achieved with these optimized cuts and bins and those from another proposed based on initial, informed guesses is shown in Appendix B Section B.5, where improvements in sensitivity by factors of 2–4 are achieved for most signal points.

7.4 Yields

Tables 7.4 and 7.5 show the yields for 139 fb^{-1}, sorted into the different analysis regions, after applying the trigger and analysis event selections for data and some select MC points, respectively. For MC, BR($H \to \chi_1^0\chi_1^0$) is assumed to be 20%, which is roughly the highest BR not already constrained by Higgs boson to invisible particle searches. The final yields in each timing and pointing bin for data in the SR$_{t<0}$ are presented in Appendix D.1.

7.5 Template Selection

In this section, the procedure to choose the optimal photon ID requirements that distinguish the two background templates is described.

	Optimized Cuts/Bins	
Parameter	$\Delta m \leq 10$ GeV	$\Delta m > 10$ GeV
E_{cell} Cut [GeV]	7	10
CR $E_{\text{T}}^{\text{miss}}$ [GeV]	<30	<30
VR $E_{\text{T}}^{\text{miss}}$ [GeV]	30–50	30–50
SR $E_{\text{T}}^{\text{miss}}$ [GeV]	>80	>50
$\lvert z\rvert$ bins [mm]	[0,50,100,200,300,2000]	
γ channel : t bins [ns]	[0,0.2,0.4,0.6,0.8,1.0,1.5,12.0]	
$\gamma\gamma$ channel: t bins [ns]	[0,0.2,0.4,0.6,0.8,1.0,12.0]	

Table 7.3: Final, optimized cuts and bins used in the analysis, as determined by the optimization procedure.

Region	Selection	Number of Events
Total	–	119017973
Trigger	e/μ Trigger + ≥ 1 lepton	111664787
	e trigger	52718791
	μ trigger	58997831
Pre-selection	Trigger + ≥ 1 γ	13852973
	γ	13731614
	$\gamma\gamma$	121359
CR	γ	940707
	$\gamma\gamma$	14884
VR	γ	466277
	$\gamma\gamma$	6044
Low Δm $\mathrm{SR}_{t<0}$	γ	86319
	$\gamma\gamma$	1247
High Δm $\mathrm{SR}_{t<0}$	γ	205756
	$\gamma\gamma$	2857

Table 7.4: Yields for different selections in data.

	Signal Model					
Selections	(60,0.5,2)	(30,0.5,2)	(50,20,2)	(60,0.5,10)	(30,0.5,10)	(50,20,10)
Total	15105.45	15105.45	15105.45	15105.45	15105.45	15105.45
Trigger	6994.984	6957.499	7037.954	6955.48	7023.664	7016.068
e trigger	3285.758	3258.465	3303.859	3271.679	3285.445	3282.965
μ trigger	3776.713	3761.407	3805.26	3752.863	3802.824	3802.043
Trigger + ≥ 1 γ	3269.706	2935.154	3163.726	1609.433	1128.328	1457.546
γ	1948.963	2159.161	2133.148	1339.953	1026.825	1271.615
$\gamma\gamma$	1320.743	775.993	1030.578	269.480	101.503	185.931
Trigger + ≥ 1 γ + SR	1303.949	1172.796	1192.774	594.6473	432.0876	499.2941
γ	745.080	869.816	799.151	492.165	397.112	433.623
$\gamma\gamma$	558.870	302.980	393.624	102.483	34.975	65.671

Table 7.5: Yields for six sample signal points. Signals are labelled by (m_{NLSP} [GeV], m_{LSP} [GeV], τ [ns]). The yield shown here is normalized to 139 fb^{-1} assuming $\mathrm{BR}(H \to \chi_1^0\chi_1^0) = 20\%$.

7.5.1 Candidate Templates

After selection, two versions of analyses proceed in parallel, based on the signal mass splittings (see Section 7.3.1.2): one that targets signal points with small mass splittings ($\Delta m \leq 10$ GeV), and the other targeting signal points with high mass splittings ($\Delta m > 10$ GeV). The background modeling is studied separately for these two analyses, but the procedure and evaluation are the same.

Three choices are examined: so-called medium, tight, and mixed templates. Tight templates refers to the tight photon ID requirement as the dividing line between the two templates. That is that the genuine-enriched template includes photons passing the radiative Z requirement and photons in the CR passing the tight ID requirements, while the fake-enriched template includes CR photons that fail the tight identification requirements (also referred to as anti-tight). Medium templates still place all radiative Z photons in the genuine-enriched template, but the division among the templates for the rest of the CR photons is, correspondingly, the medium photon ID requirements. A third choice known as mixed templates arose as the result of dedicated studies indicating a difference in performance of the medium and tight templates among the different photon channels. Mixed templates use medium templates for the one-photon channel and tight templates for the two-photon channel[5]. Table 7.6 summarizes the definition of the three sets of templates.

The two templates in each pointing bin for each of the photon channels are shown in Appendix C.1. Since these two templates in each pointing bin are used as a basis, the shape differences are essential to the fitting of the SR background expectation. In order for this

Templates	Photon Channel(s)	Genuine-Enriched	Fake-Enriched
Medium	γ & $\gamma\gamma$	$Z \to ll\gamma$ + CR Medium	CR Anti-Medium
Tight	γ & $\gamma\gamma$	$Z \to ll\gamma$ + CR Tight	CR Anti-Tight
Mixed	γ	$Z \to ll\gamma$ + CR Medium	CR Anti-Medium
	$\gamma\gamma$	$Z \to ll\gamma$ + CR Tight	CR Anti-Tight

Table 7.6: Definition of medium, tight, and mixed photon timing templates.

[5]In general, it was found that tight templates are a bit too narrow in the one photon channel and that medium templates were a bit too wide for the two photon channel. Intuitively, the template preference among photon channels can be explained by differences in purity. For example, in the two photon channel, which is relatively more pure in genuine photons, a tigher ID requirement allows for a better spanning of the purity space.

fit to be stable, there must be a statistically significant difference in shape between the two templates. Note, the lack of difference is not in itself an issue, as long as the template mixture can describe the data. However, if there is ambiguity between their shapes, this can lead to fit instabilities that can produce discontinuities in the optimization space that hinder the optimizer's ability to find the true minimum. Ambiguity between the templates can also produce strange effects in which the two templates that could have radically different statistics seem like equally valid choices, leading to unstable fluctuations due to large differences in the statistical uncertainties. In order to combat this issue, a standard measure of the statistical difference between two distributions, known as the Kolmogorov-Smirnov (KS) test, was applied in each analysis category. The values of this KS test represent probabilities that the given distributions are drawn from the same distribution. A summary of the KS test values is shown in Table 7.7 for the tight and medium templates in the low mass splitting and high mass splitting analyses[6]. Almost all of the KS values among the different pointing bins are very near 0, indicating that there is no ambiguity. However, a few have much higher values nearer to 1, indicating that there is little statistical difference between the two templates. It has been observed that fits experience instabilities when trying to fit templates in the categories with these very high KS values. Thus, a threshold was chosen at 0.9, above which the two templates are considered to have non-distinguishable shapes, and the purity nuisance parameter α_i is fixed in the fit to 0.5, i.e. mixing the two templates together equally. Otherwise, α_i is left as a floating parameter in the fit. This criteria requires purity fixing in only two pointing bins for the high mass splitting analysis. The low mass splitting analysis requires no purity fixing at all. This limited and targeted purity fixing treatment ensures the fit to be stable in the fit validation regions.

7.5.2 Template Performance

The three template choices are evaluated on the $\mathrm{SR}_{t<0}$ region first and then later with the $\mathrm{VR}_{t>0}$ data as a cross check (see Section 7.5.4). A signal-plus-background fit is performed to this region with the available signal points. Given that this region is essentially purely background, the optimal background templates should produce no large deficits or excesses.

[6]These values are also displayed on the template comparison plots themselves in Appendix C.1.

		$\Delta m \leq 10$ GeV		$\Delta m > 10$ GeV	
Channel	$\|z\|$ Bin [mm]	Medium	Tight	Medium	Tight
γ	$\|z\| < 50$ mm	0.000	0.000	0.000	0.000
	$50 < \|z\| < 100$	**0.945**	**1.000**	0.001	0.000
	$100 < \|z\| < 200$	0.000	0.000	0.000	0.000
	$200 < \|z\| < 300$	0.000	0.000	0.000	0.000
	$300 < \|z\| < 2000$	0.744	0.602	0.093	0.004
$\gamma\gamma$	$\|z\|<50$	0.000	0.000	0.000	0.000
	$50 < \|z\| < 100$	0.703	**0.999**	0.003	0.000
	$100 < \|z\| < 200$	0.000	0.000	0.000	0.000
	$200 < \|z\| < 300$	0.000	0.000	0.000	0.000
	$300 < \|z\| < 2000$	0.853	0.511	0.075	0.002

Table 7.7: $SR_{t<0}$ KS test values for genuine-enriched and fake-enriched template comparisons for medium and tight choices for high and low mass splitting selections. $\Delta m \leq 10$ GeV selections include SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV. $\Delta m > 10$ GeV selections include SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV. KS values reaching the threshold of 0.9 for purity fixing are in bold.

The observed significance of this fit is taken as the primary figure of merit for evaluating the template performance. If the deficit or excess is small enough to be generally consistent with zero, the NLL value itself can then also be considered as an additional measure of the fit quality.

The signal-plus-background fit has the normalization of the signal component as a free parameter. The observed deficit or excess corresponds to the signed two-sided significance of this fitted signal normalization parameter. Among the available 2 and 10 ns signal points, timing and pointing distribution shapes are expected to differ enough such that deviations may be picked up in different categories. For example, higher lifetimes have lower yields in the two-photon channel. However, the optimal background templates should generally not detect deviations from background with the testing data sets, regardless of the signal model used.

Figure 7.4 summarizes the results. For the high mass splitting analyses, the mixed and medium templates both perform well with relatively small significances with no major outliers. For the low mass splitting analyses, all of the templates do a reasonable job. The general consistency of the template performance across lifetimes attests to the fit stability and engenders confidence in applying the same method to other lifetimes as well.

The final template choice among those having acceptable small significances is then determined by the NLL value. In both the low and high mass splitting analyses, for every individual signal point, it is the case that the mixed templates always have the lowest NLL value. Since the structure and parameter space of the fit does not change with a different template choice, a lower NLL value is an indicator of the fit quality and suggests that the corresponding template describes the data the best. Therefore, since they both do not produce large significances on fits to $SR_{t<0}$ and they produce the best NLL values, mixed templates are the optimal choice for this analysis.

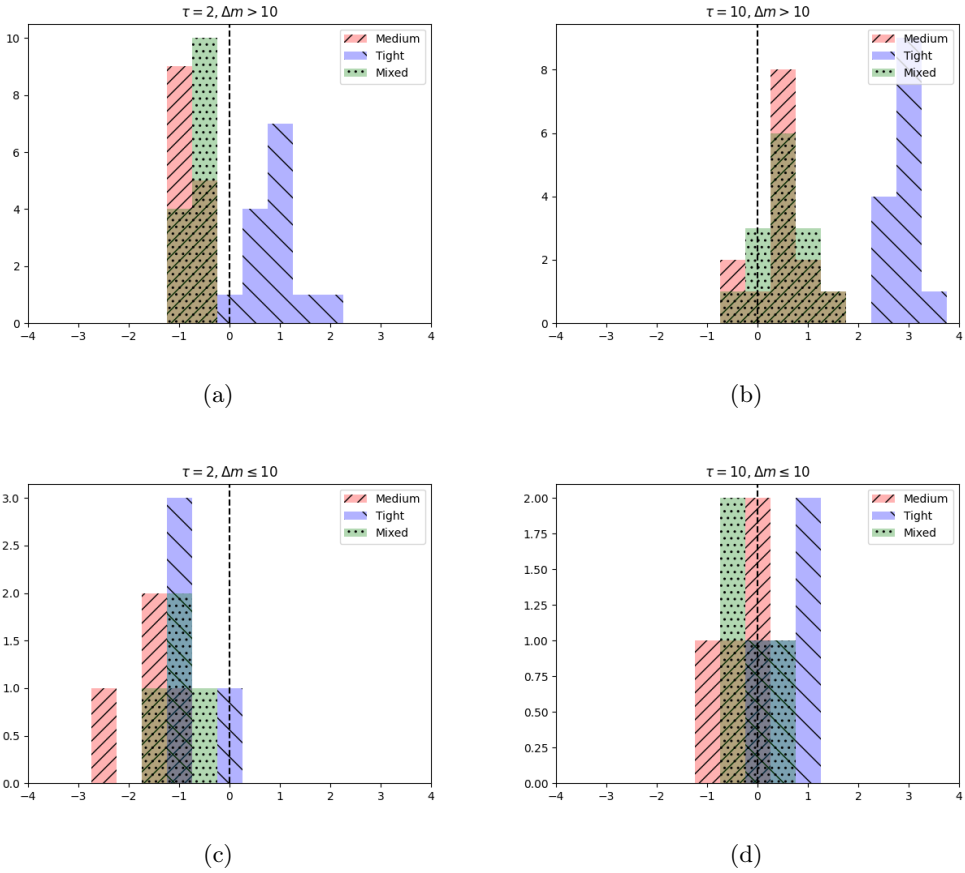
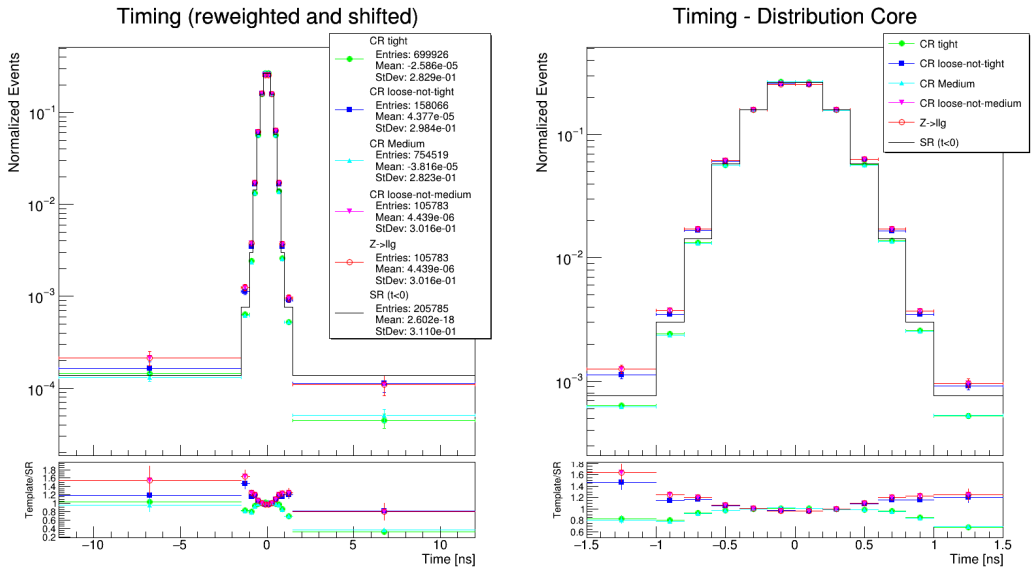


Figure 7.4: Distributions of the significance values of deficits or excesses in signal-plus-background fits to the $SR_{t<0}$ region for the high mass splitting selection for (a) 2 ns lifetimes and (b) 10 ns lifetimes. The same for the low mass splitting selection is also shown for (c) 2 ns lifetimes and (d) 10 ns lifetimes.

7.5.3 Final Template Shape Comparison

This section presents comparisons of the timing distributions in various regions and categories used in the analysis after applying the energy reweighting and mean time shifting described in Section 7.2.2. The templates displayed are the optimal mixed templates chosen via the evaluation in Section 7.5.2.

Figure 7.5 shows the mixed timing templates in data comprising the genuine- and fake-photon regions, as well as the timing distributions in the VR and the $SR_{t<0}$ region (blinded, reflected across $t = 0$). Generally, the fake-photon-enriched template is observed to be slightly wider than the $SR_{t<0}$—as expected due to its reduced purity that produces a broader timing tail—while the photon-enriched template is narrower. This sandwiching of the SR by the two templates illustrates a key property exploited by the primary fitting strategy of the analysis, as it suggests that the SR can indeed be described by some linear combination of the two templates. In this way, the different photon purities of the two templates produce a basis that can, in principle, describe any other region expected to be composed of some mixture of genuine and fake photons.



(a) (b)

Figure 7.5: Timing distributions in data for the real photon template, fake photon template, validation region, and negative timing signal region (reflected over t=0), in the full timing range (left) and zoomed on the central bins (right).

Figure 7.6 compares the timing shapes of the two templates and the $SR_{t<0}$ for the one-photon channel in the each of the pointing bins. Figure 7.7 does the same for the two-photon channel. From these plots, it can be observed that the two templates generally model the shape of the $SR_{t<0}$ and that, in many cases across pointing and timing bins, the two templates surround the $SR_{t<0}$ yields, clearly illustrating that the target distribution is indeed an intermediate mixture of the two templates.

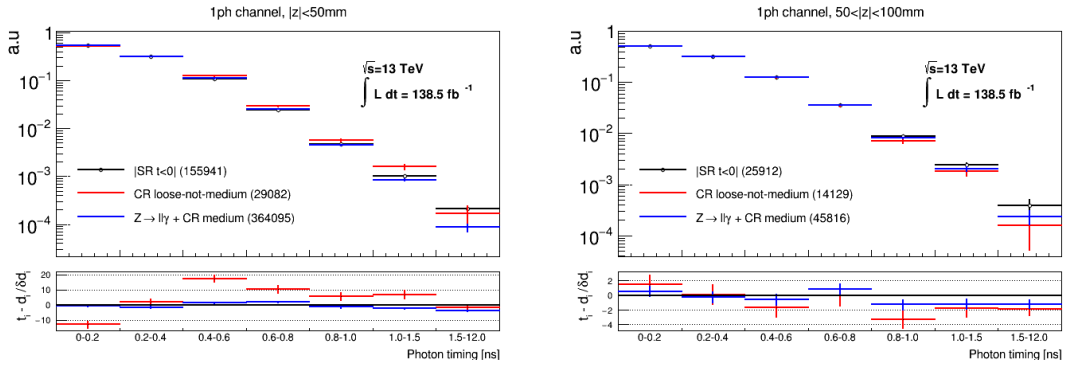
7.5.4 Template Selection Validation

The $VR_{t>0}$ region—now looking at the postive side of the timing distribution—is used as another independent sample to test the background modeling strategy. The CRs are energy-reweighted to the VR, and the studies of the observed deficits and excesses of the medium, tight, and mixed templates are re-evaluated.

The significance along with NLL values for signal-plus-background fits to the $VR_{t>0}$ region are listed in Table 7.8. Consistent with the $SR_{t<0}$ results, none of the signal points produce large deviations from the background expectation of zero. Evaluating the various templates, it is found that mixed templates also produce the lowest NLL values, as was the case for the $SR_{t<0}$ region. Therefore, the evaluation of the templates in the $VR_{t>0}$ region confirms the choice of the mixed templates that was made based on the $SR_{t<0}$ region and attests to the robustness of the fitting procedure across analysis regions.

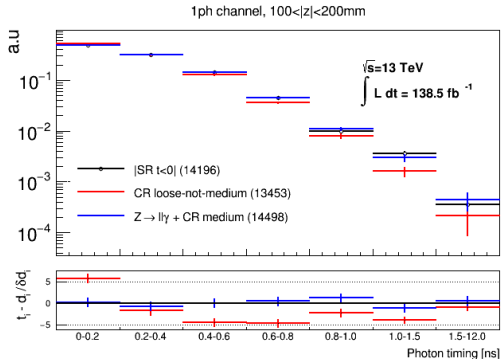
Signal Point			Fit Results			
m_{NLSP} [GeV]	m_{LSP} [GeV]	τ [ns]	p-value	Significance	NLL_{min} - $\text{NLL}_{\text{BR=0}}$	BR_{min}
30	0	2	0.418	0.208	-0.021	0.010
30	10	2	0.907	-1.32	-0.87	-0.076
40	10	2	0.276	0.594	-0.176	0.007
40	20	2	0.615	-0.292	-0.042	-0.005
50	0	2	0	0	[-0.001, 0.000]	0.0
50	10	2	0	0	[-0.001, 0.000]	0.0
50	20	2	0.459	0.103	-0.005	0.0009
50	30	2	0.339	0.414	-0.085	0.008
60	0	2	0.442	0.145	-0.010	0.0006
60	10	2	0.447	0.134	-0.009	0.0005
60	20	2	0.471	0.0738	-0.002	0.0003
60	30	2	0.525	-0.0635	-0.002	-0.0005
30	0	10	0.842	-1	-0.50	-0.056
30	10	10	0.583	-0.209	-0.021	-0.043
40	0	10	0.823	-0.927	-0.429	-0.017
40	10	10	0.768	-0.731	-0.267	-0.013
40	20	10	0.8	-0.841	-0.35	-0.03
50	0	10	0.719	-0.581	-0.169	-0.006
50	10	10	0.702	-0.531	-0.141	-0.004
50	20	10	0.885	-1.2	-0.718	-0.014
60	0	10	0.762	-0.713	-0.254	-0.004
60	30	10	0.872	-1.14	-0.645	-0.008

Table 7.8: Fit results for the high mass splitting points in the $\text{VR}_{t>0}$ for mixed templates.

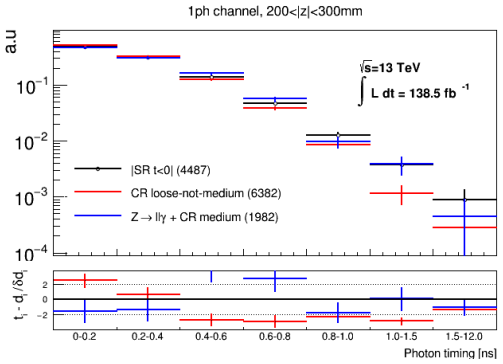


(a)

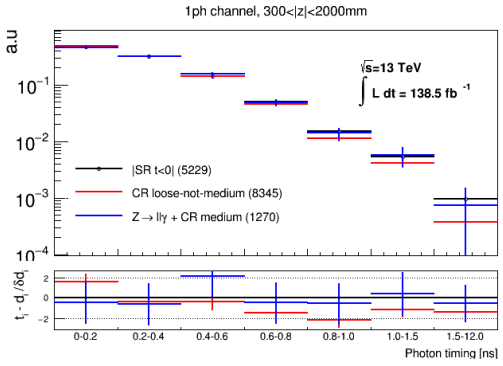
(b)



(c)

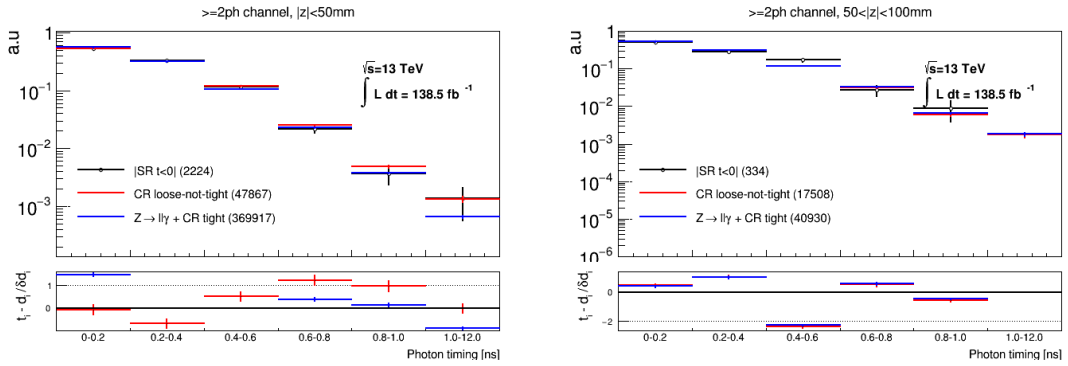


(d)



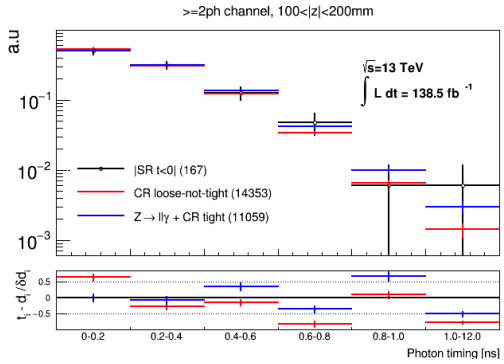
(e)

Figure 7.6: Comparison between the real-enriched and fake-enriched templates to the $\mathrm{SR}_{t<0}$ region for the one-photon channel with the high Δm selections (SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV) separated into pointing bins (a) $|z| < 50$ mm (b) $50 < |z| < 100$ mm (c) $100 < |z| < 200$ mm (d) $200 < |z| < 300$ mm (e) $300 < |z| < 2000$ mm.

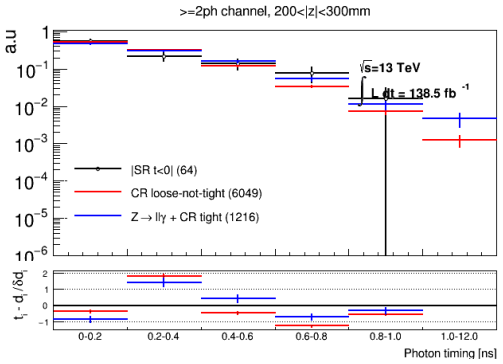


(a)

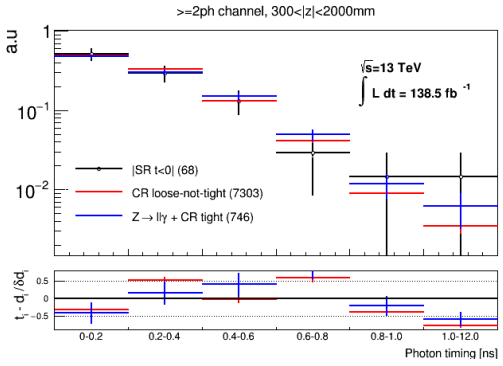
(b)



(c)



(d)



(e)

Figure 7.7: Comparison between the real-enriched and fake-enriched templates to the $\mathrm{SR}_{t<0}$ region for the two-photon channel with the high Δm selections (SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV) separated into pointing bins (a) $|z| < 50$ mm (b) $50 < |z| < 100$ mm (c) $100 < |z| < 200$ mm (d) $200 < |z| < 300$ mm (e) $300 < |z| < 2000$ mm.

Chapter 8: Systematic Uncertainties

The systematic uncertainties to be considered in this analysis can be divided into two categories: those affecting the total yield of signal events (see Section 8.1) and those affecting the shapes of the pointing and timing distributions of signal and/or background templates (see Section 8.2). Many of these uncertainties are generally considered standard, as they are applied, to some extent, in nearly all ATLAS analyses. However, due to the displaced nature of the EM objects under study, this analysis must also consider additional uncertainties that may arise due to this unique event topology. As will be discussed, the impact of systematic uncertainties to this analysis is generally quite modest. It is therefore expected that, even with further implementation of other minor uncertainties before publication, the analysis will remain limited by statistical rather than systematic uncertainties.

8.1 Signal Yield Uncertainties

Several standard uncertainties dealing with the measured luminosity and instrumental uncertainties are included in the analysis results. These uncertainties impact only the overall signal normalization. Since variations in the signal yields between pointing categories arise from similar sources, they are treated as correlated uncertainties and are incorporated into the likelihood model via a single nuisance parameter.

8.1.1 Integrated Luminosity Uncertainty

A standard integrated luminosity uncertainty on the combined 2015–2018 dataset of 1.7% is taken into account. The magnitude of this uncertainty is derived from measurements using the LUCID detector [99], using the methodology outlined in Section 3.1 and described in

detail in Ref. [100]. This uncertainty is correlated across all pointing categories.

8.1.2 Instrumental Uncertainties

Instrumental uncertainties arise from many different experimental systematic variations including trigger efficiencies, energy scales, energy resolutions, ID efficiencies, and $E_{\mathrm{T}}^{\mathrm{miss}}$ efficiencies. These variations are correlated across all categories and are taken from standard ATLAS recommendations.

Systematic uncertainties on the lepton trigger efficiencies are evaluated via so-called tag-and-probe methods in which events with two lepton resonances (e.g. $Z \to ll$) are triggered with one of the leptons, and the other is used as an unbiased probe for efficiency studies [82, 84]. To get the trigger efficiency, the number of reconstructed probe leptons satisfying a set of isolation and requirements is compared to the number of those leptons that result in a corresponding trigger.

For electrons and photons, uncertainties related to energy scales and resolutions arise from many different sources including pileup, calibration of the different EM calorimeter layers, modeling of the detector material in front of the EM calorimeter, non-linearities in cell energy measurements, modeling of the lateral shower shape, and misclassification of converted and unconverted photons. Each of these are quantified with dedicated studies, as described in Ref. [89].

Since reconstruction algorithms and the ATLAS detector itself are optimized for prompt, pointing objects, the loose photon ID efficiency drops for increasingly delayed and non-pointing photons, impacting signal yields. The magnitude of this decreased efficiency is determined by first confirming that photons from SM and SUSY sources achieve equal efficiencies in the pointing regime and then measuring the SUSY efficiencies out to much larger pointing values. Figure 8.1 shows a sample result, illustrating the effect of the falling efficiency. The maximum residual from 100% efficiency can be taken as a systematic uncertainty on the loose photon ID scale factors. This uncertainty is not implemented in the results presented in this thesis since it is not expected to have a significant impact on the results, but implementation is in progress and will be completed before the results are published.

The preceding discussion of sources of systematic uncertainties for EM objects also ap-

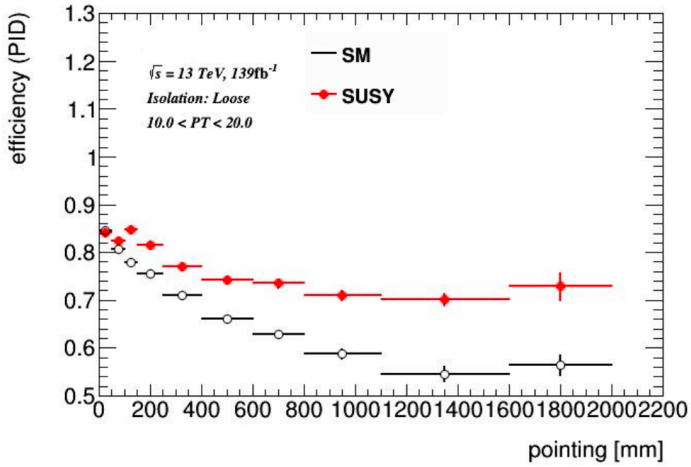


Figure 8.1: Photon ID efficiency as a function of pointing for photons with $10 < p_{\mathrm{T}} < 20$ GeV. The black SM points show simulated photons from radiative $Z \to ll\gamma$ decays. The red SUSY points show photons produced from signal MC using the analysis selection outlined in Chapter 6.

plies to other objects entering this analysis, including jets and $E_{\mathrm{T}}^{\mathrm{miss}}$. Similar studies to quantify the magnitude of analogous uncertainties (barring the ones specifically addressing displacement) are also carried out for these objects, as described in Refs. [85] and [95].

Table 8.1 provides an example table, breaking down the magnitude of the effect of the systematic variations on the signal yield for a sample signal point. As for all signal points, the uncertainties are generally quite small, being on the order of a few percent.

8.2 Shape Uncertainties

Uncertainties associated with statistical fluctuations, timing measurements, and the background modeling strategy can all affect the shapes of the timing distributions. They must therefore be modelled with appropriate shape systematic uncertainties.

8.2.1 Signal Shape Uncertainties

Signal timing and pointing distributions are taken from signal MC, and thus the shape uncertainties for signal include the MC statistical fluctuations in each pointing category and

Systematic Source		γ		$\gamma\gamma$	
Description	Name	Up [%]	Down [%]	Up [%]	Down [%]
e/γ Energy Resolution	EG_RESOLUTION_ALL	-0.0757	0.0163	-0.1531	0.3326
e/γ Energy Scale	EG_SCALE_ALL	-0.2416	-0.1073	0.8533	-0.6288
μ Isolation	MUON_EFF_ISO_STAT	0.1262	-0.5830	0.1242	-0.6159
	MUON_EFF_ISO_SYS	0.2034	-0.1827	0.2096	-0.1819
μ Reco.	MUON_EFF_RECO_STAT	0.0307	-0.0307	0.0354	-0.0354
	MUON_EFF_RECO_STAT_LOWPT	0.0014	-0.0018	0.0020	-0.0027
	MUON_EFF_RECO_SYS	0.1351	-0.1345	0.1387	-0.1377
	MUON_EFF_RECO_SYS_LOWPT	0.0019	-0.0019	0.0031	-0.0031
μ Track to Vertex Assoc.	MUON_EFF_TTVA_STAT	0.0213	-0.0214	0.0222	-0.0222
	MUON_EFF_TTVA_SYS	0.0267	-0.0246	0.0258	-0.0242
μ ID Track Smearing	MUON_ID	-0.0584	-0.0067	-0.0028	0.0673
μ MS Track Smearing	MUON_MS	0.0823	-0.0047	0.0145	-0.0338
μ Sagitta (p_{T})	MUON_SAGITTA_RESBIAS	0.0003	0.0003	0.0000	0.0000
	MUON_SAGITTA_RHO	0.0003	0.0003	0.0000	0.0272
μ Energy Scale	MUON_SCALE	-0.0237	0.0372	-0.0093	0.0269
e Efficiencies	EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	0.3410	-0.3404	0.3040	-0.3037
	EL_EFF_Iso_TOTAL_1NPCOR_PLUS_UNCOR	0.0235	-0.0235	0.0174	-0.0174
	EL_EFF_Reco_TOTAL_1NPCOR_PLUS_UNCOR	0.0699	-0.0699	0.0634	-0.0634
	EL_EFF_TriggerEff_TOTAL_1NPCOR_PLUS_UNCOR	0.0001	-0.0001	0.0000	-0.0000
	EL_EFF_Trigger_TOTAL_1NPCOR_PLUS_UNCOR	0.0823	-0.0823	0.0770	-0.0770
$E_{\mathrm{T}}^{\mathrm{miss}}$ Soft Track	MET_SoftTrk_ResoPara	-0.7479	-0.7479	-0.0899	-0.0899
	MET_SoftTrk_ResoPerp	-0.3325	-0.3325	0.3150	0.3150
	MET_SoftTrk_Scale	-0.3572	0.0760	0.0451	0.1317

Table 8.1: Signal yield systematic uncertainty breakdown for the $m_{\mathrm{NLSP}} = 60$ GeV, $m_{\mathrm{LSP}} = 0.5$ GeV, $\tau = 2$ ns signal point. Columns show up and down variations by photon channel. Rows corresponds to a variation from the indicated source alone, expressed as percentages of the signal yield in the SR.

timing bin. Since each individual timing bin has its own statistical uncertainty, they each have independent nuisance parameters to account for these MC fluctuations.

Uncertainties in the characterization of the timing resolution affect how the signal MC is smeared to match the resolutions observed in data after calibration, as described in Section 5.2.2.3. This uncertainty is estimated by implementing alternative degrees of timing smearing in order to cover any discrepancies between smeared MC and data. Details can be found in Appendix A.4. This resolution uncertainty is fully correlated between pointing categories and timing bins. In the likelihood function, it is parameterized by a single nuisance parameter with correlated variations in timing and pointing bins.

The timing measurement, as described in Section 5.2.2.1, is optimized for prompt signals using OFCs. As was studied in previous displaced photon analyses from Run 1, a small bias exists in the measurement of delayed objects [65, 66]. In those analyses, a 10% systematic uncertainty was applied to all MC timing measurements. However, as the timing bins used in this analysis are quite coarse, with the most sensitive final timing bins being over 10 ns wide, the effect of this uncertainty on the migration of events between timing bins is effectively negligible. Thus, since it is not expected to appreciably impact the final result, this uncertainty is not implemented in the results presented in this thesis. It will, however, be implemented prior to publication [1].

8.2.2 Background Shape Uncertainties

As described in Section 7.2, the predicted background timing and pointing distributions in the SR are constructed from templates derived from data. Similar to the case for signal shapes, a shape uncertainty for the two background templates arises due to statistical fluctuations. Likewise, this is an uncorrelated uncertainty.

Additionally, a correlated shape systematic uncertainty is implemented in order to account for uncertainties associated with the overall shape of the constructed templates. This uncertainty is correlated within each pointing category across all timing bins. The mag-

[1] In fact, with new data collected in 2018, an improved method of characterizing this bias has become available that will allow for a precise measurement of the bias as a function of true arrival time. See Appendix A.7 for more details on this study. It is intended that, with this improved understanding, MC times will be adjusted to mimic the bias seen in data, and a new systematic uncertainty (smaller than 10%) will be implemented to cover any residual variation.

nitude of this uncertainty is calculated based on the differences between the nominal and alternate templates based on photon ID, as discussed in Section 7.5.

The template reweighting procedure, detailed in Section 7.2.2.1, has uncertainties associated with the weights, as calculated from binned E_{cell} distributions. These uncertainties can be derived from the statistical errors in each E_{cell} bin and applied event-by-event. Since the reweighting procedure was optimized to maximize the available statistics among the different bins, the weight uncertainties are quite small. Thus, this uncertainty is expected to cause very little event migration between the most sensitive timing bins due to their relatively large widths. This uncertainty is, therefore, not implemented in the results presented in this thesis, but it will be implemented as described before publication.

Finally, to account for any remaining residual intrinsic shape differences between the templates and validation data, an uncorrelated, non-closure uncertainty is introduced. To estimate these intrinsic differences, a background-only fit is first performed to the $\text{SR}_{t<0}$ data. If the data and background prediction agree within their statistical uncertainties for a given timing bin, no non-closure uncertainty is assigned. Any residual difference, after accounting for the statistical uncertainties (one standard deviation), is taken as the non-closure uncertainty in that bin, which is then incorporated into the final fit. Visualizations and tables of the non-closure uncertainties in each bin can be found in Appendix C.2. These non-closure uncertainties are generally quite small or zero, as the original fits tend to already achieve relatively good closure.

Chapter 9: Expected Sensitivity

This chapter presents post-fit plots and sensitivity projections using the $SR_{t<0}$ data, which serves as a proxy for the final $SR_{t>0}$. The results are derived using the full analysis strategy outlined in Chapter 7 along with the systematic uncertainties discussed in Chapter 8.

9.1 $SR_{t<0}$ Post-Fit Plots

The post-fit plots comparing data against the fitted estimated background for the high and low mass splitting analyses are available in Figures 9.1 and 9.2, respectively. Good agreement within one standard deviation is observed between the data in the $SR_{t<0}$ and the estimated background across nearly all pointing categories and timing bins. The fitted purity values, reflecting the fraction taken from the genuine-photon timing template, can be found in Appendix C.3.

9.2 $SR_{t<0}$ Sensitivity

Expected sensitivity results are presented as upper limits at the 95% CL on BR($H \rightarrow \chi_1^0\chi_1^0$). These results for the full signal grid are shown in Figure 9.3. For signal points with high mass splittings, sensitivity at the sub-5% level is expected for most points, with even sub-1% expectations for the highest mass splittings. For the low mass splitting signal points, with much lower photon acceptance due to the softer nature of the photons, sensitivities of $\sim$10–20% are expected for most points.

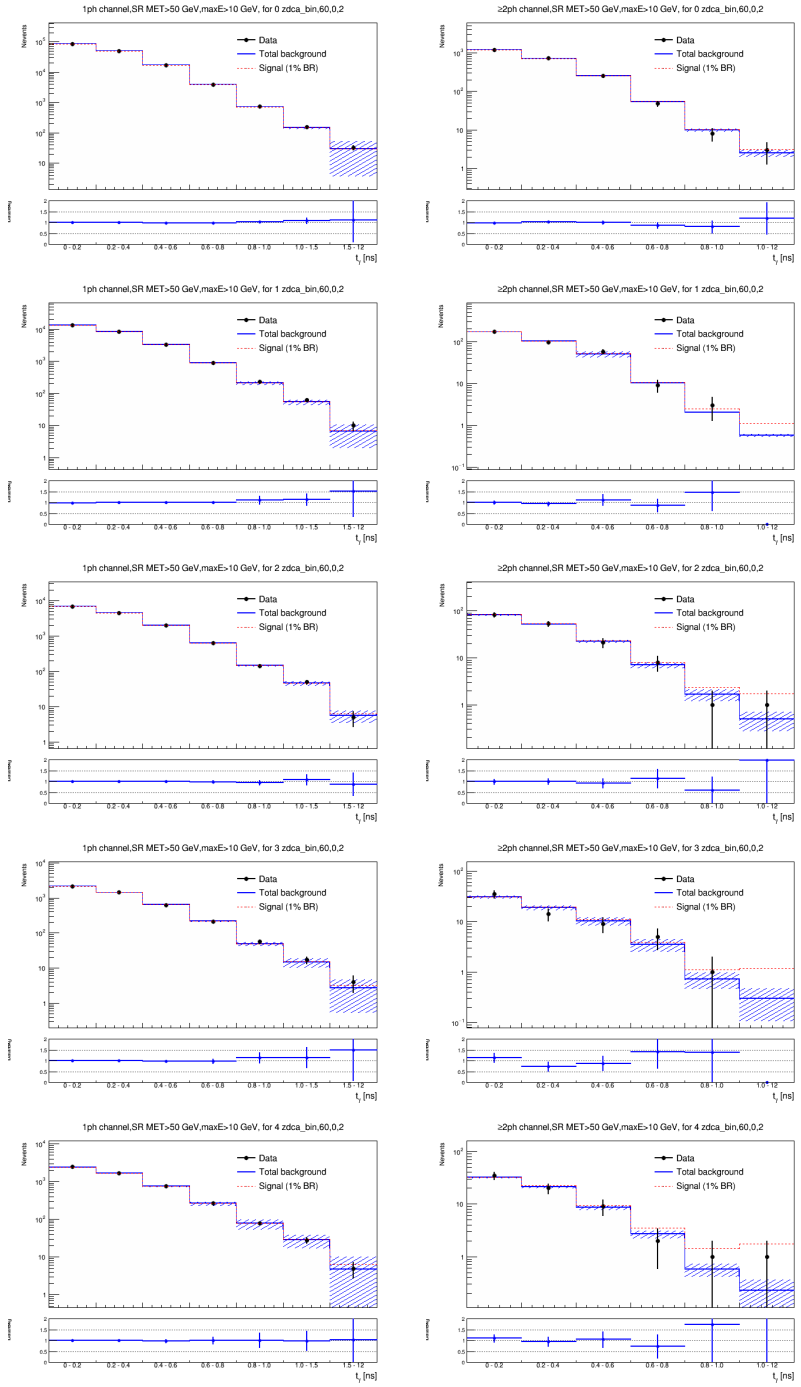


Figure 9.1: Post-fit plots comparing $\mathrm{SR}_{t<0}$ data (black) against a background-only fit with BR = 0 (blue) for the high mass splitting analysis in the one-photon (left) and two-photon (right) channels. The signal template for the $m_{\mathrm{NLSP}} = 60$ GeV, $m_{\mathrm{LSP}} = 0.5$ GeV, $\tau = 2$ ns signal model with BR = 1% is added on top of the background fit (red).

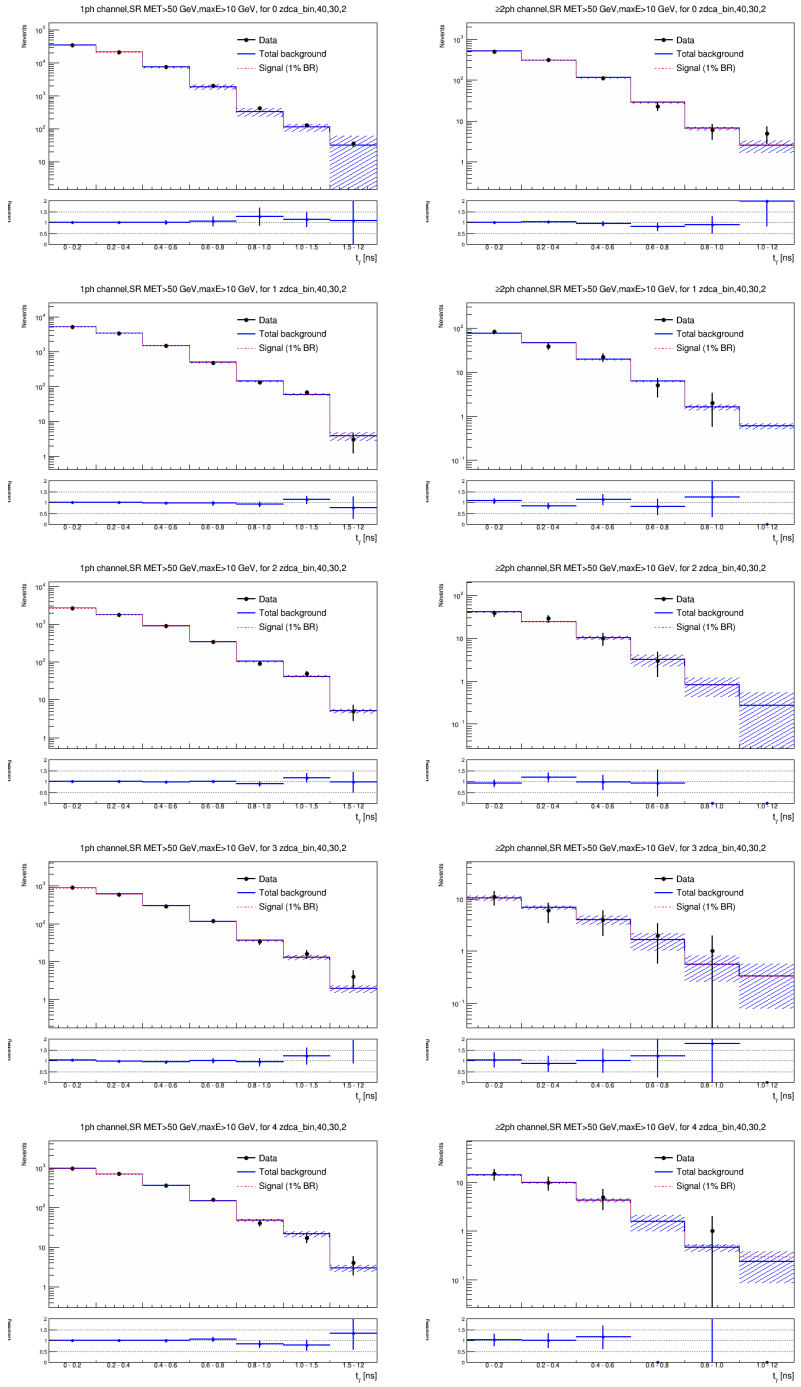


Figure 9.2: Post-fit plots comparing $\mathrm{SR}_{t<0}$ data (black) against a background-only fit with BR = 0 (blue) for the low mass splitting analysis in the one-photon (left) and two-photon (right) channels. The signal template for the $m_{\mathrm{NLSP}} = 60$ GeV, $m_{\mathrm{LSP}} = 0.5$ GeV, $\tau = 2$ ns signal model with BR = 1% is added on top of the background fit (red).

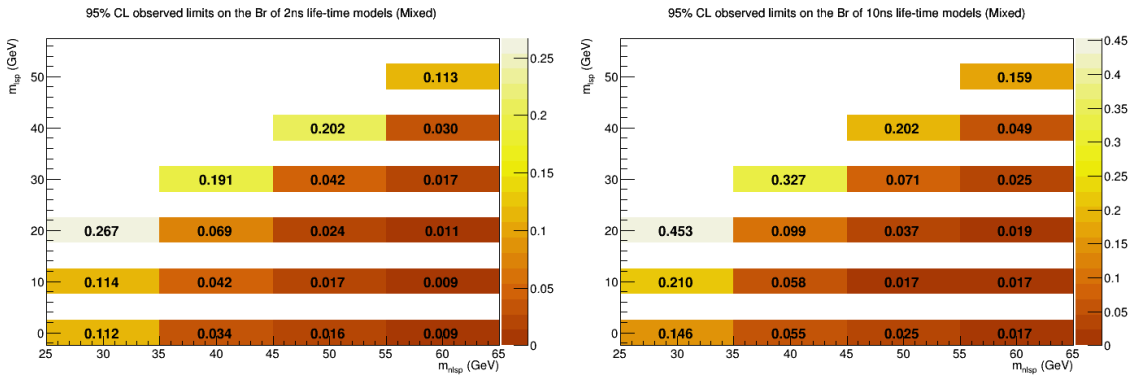


Figure 9.3: Expected upper limits at 95% CL of signal branching ratio under different assumptions of signal models based on the $SR_{t<0}$ timing shape. Each signal point uses the appropriate low or high mass splitting SR selections, as given in Table 7.3.

Chapter 10: Results

Having thoroughly validated the full analysis strategy presented in Chapter 7 with $SR_{t<0}$ data as a proxy for the blinded $SR_{t>0}$ data, the final, unblinded results are now presented here. Post-fit and pull plots of the timing distributions and the observed best-fit BRs are shown first. Since no significant excesses are observed, an evaluation of the BR exclusion limits follows.

10.1 $SR_{t>0}$ Post-Fit Plots

The post-fit plots comparing data against the background-only fits for the high and low mass splitting analyses are given in Figures 10.1 and 10.2, respectively [1]. Two representative signal points are stacked on top of the fit to illustrate the impact of any potential signal contribution. In contrast to the increasing signal contributions in higher pointing categories and higher timing bins, the $SR_{t>0}$ data remains consistent with background predictions, maintaining good agreement within small statistical fluctuations across all bins for both the high and low mass splitting analyses. Such consistency strongly suggests that background modeling strategy describes the $SR_{t>0}$ data very well without any observed signal component.

Figure 10.3 displays the observed significances for signal-plus-background fits to the unblinded $SR_{t>0}$ data for all the full signal grid. These results confirm that there are indeed no large excesses observed for any of the available signal points, with all points having signal contribution significances of $\lesssim 1$ standard deviation.

[1] The results of the KS tests to determine purity fixing for the unblinded $SR_{t>0}$ data can be found in Appendix E.1. The purity fixing required is identical to that in the $SR_{t<0}$. The fitted (or fixed) purity values in each pointing category can be found in Appendix E.2

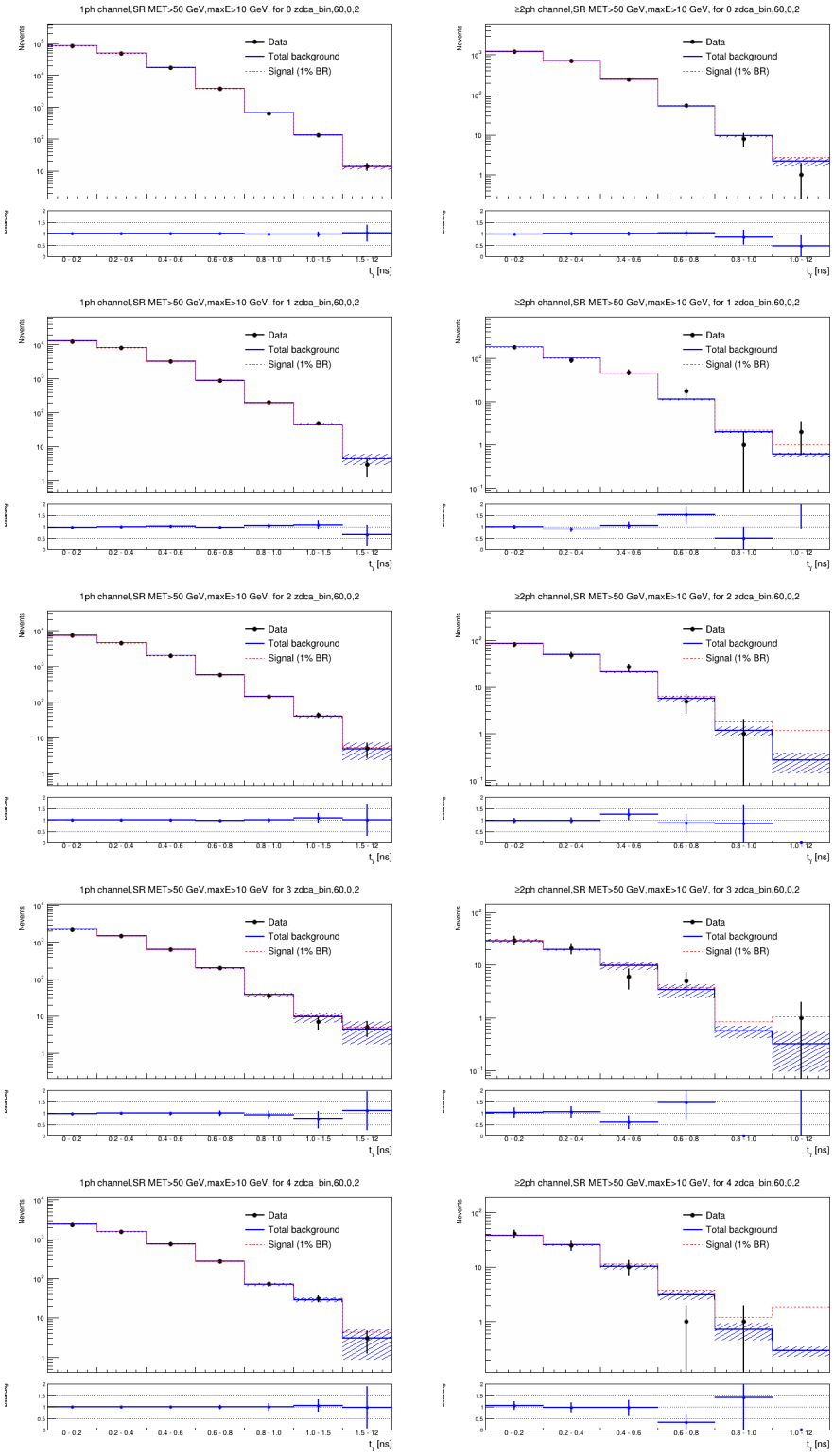


Figure 10.1: Post-fit plots comparing $SR_{t>0}$ data (black) against a background-only fit with BR = 0 (blue) for the high mass splitting analysis in the one-photon (left) and two-photon (right) channels. The signal template for the $m_{\text{NLSP}} = 60$ GeV, $m_{\text{LSP}} = 0.5$ GeV, $\tau = 2$ ns signal model with BR = 1% is added on top of the background fit (red).

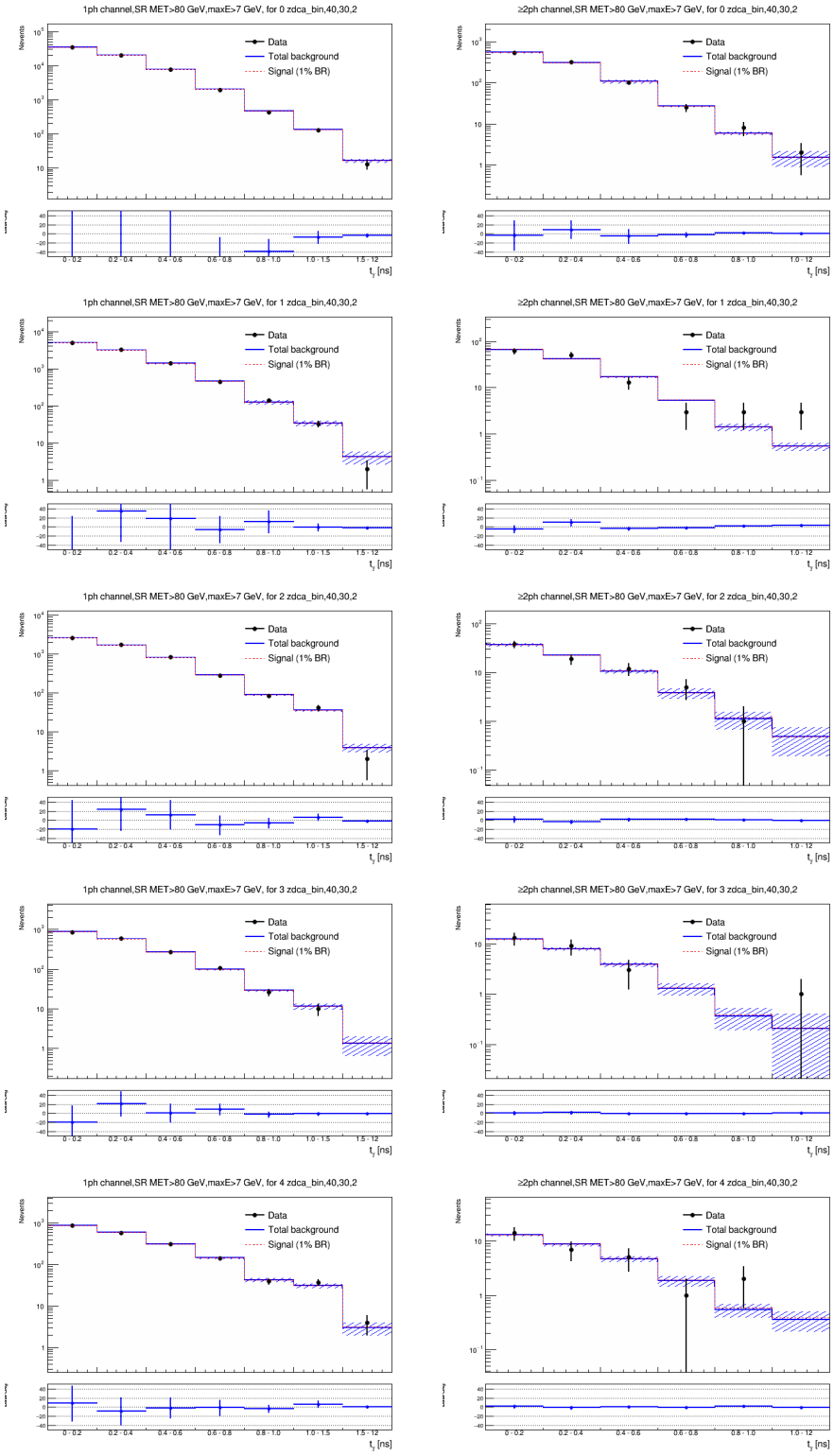


Figure 10.2: Post-fit plots comparing $SR_{t>0}$ data (black) against a background-only fit with BR = 0 (blue) for the low mass splitting analysis in the one-photon (left) and two-photon (right) channels. The signal template for the $m_{\mathrm{NLSP}} = 40$ GeV, $m_{\mathrm{LSP}} = 30$ GeV, $\tau = 2$ ns signal model with BR = 1% is added on top of the background fit (red).

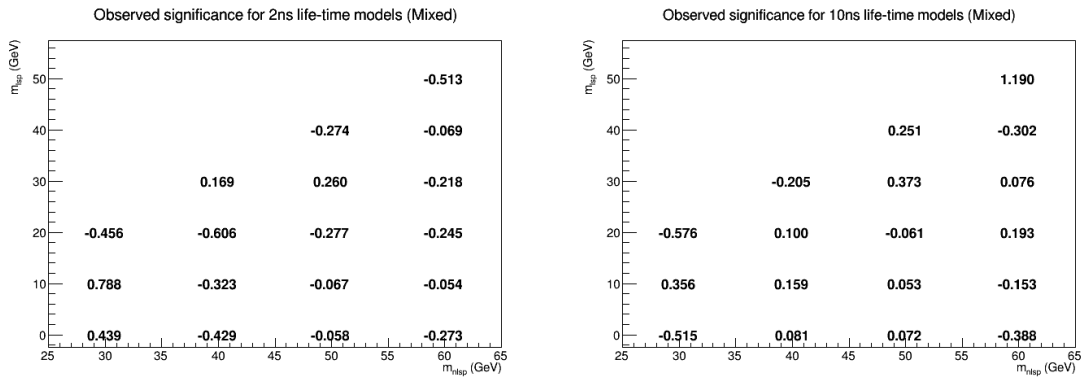


Figure 10.3: Observed significance for signal BR under different assumptions of signal models in fits to $SR_{t>0}$. Each signal point uses the appropriate low or high mass splitting SR selections, as given in Table 7.3.

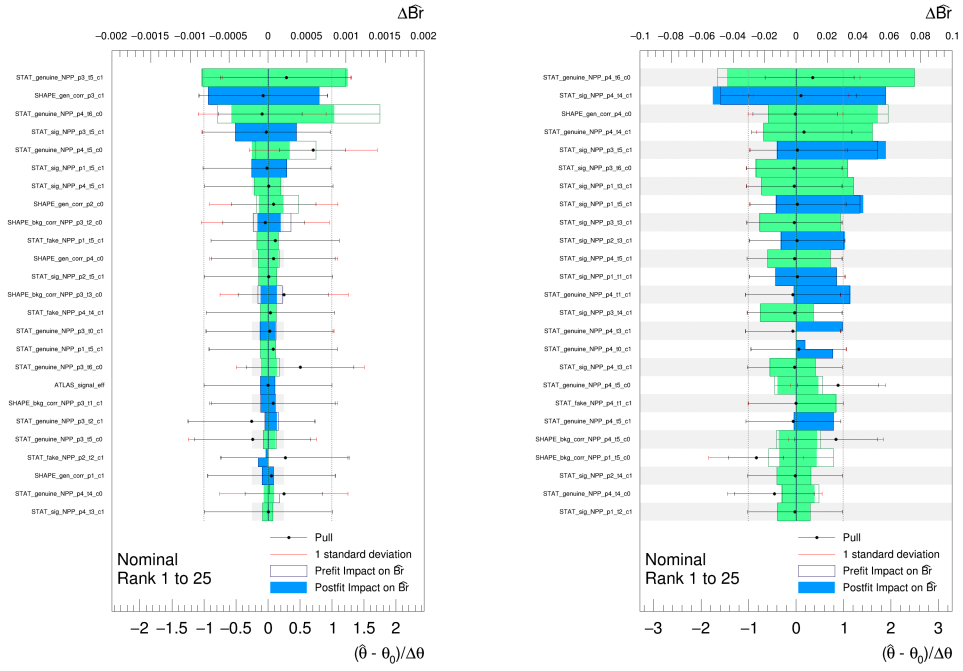
10.2 Pulls

Post-fit statistical pulls quantify discrepancies between fitted nuisance parameters and data. A signal-plus-background fit is performed, and then the nuisance parameters are independently varied by one standard deviation in each direction to quantify their impact on the parameter of interest. Significant deviations beyond uncertainties or identifiable directional trends in the pulls may indicate template mismodeling and/or of systematic uncertainties. Pulls thus help to evaluate the overall goodness-of-fit. The largest post-fit statistical pulls for two representative signal points are summarized in Figure 10.4. As shown, the pulls are generally well-behaved, with relatively small deviations and no clear patterns among nuisance parameters. Thus, the pulls provide evidence that the fits and the uncertainties incorporated therein behave as expected for proper modeling of the data.

10.3 BR Limits

In light of the consistency of the $SR_{t>0}$ results with background, exclusion limits on $BR(H \to \chi_1^0\chi_1^0)$ are evaluated at the 95% CL and shown in Figure 10.5. These observed limits can be compared to the expected ones based on $SR_{t<0}$ pseudodata in Figure 9.3. As can be seen, no large deviations from sensitivity expectations is observed.

These BR limits represent the lowest available limits in the soft, displaced photon sce-



(a) (b)

Figure 10.4: Summary of the post-fit statistical pulls for the nuisance parameters with the largest impact on the fitted BR in signal-plus-background fits to the $\mathrm{SR}_{t>0}$ for signal points (a) $m_{\mathrm{NLSP}} = 60$ GeV, $m_{\mathrm{LSP}} = 0.5$ GeV, $\tau = 2$ ns (high mass splitting) and (b) $m_{\mathrm{NLSP}} = 40$ GeV, $m_{\mathrm{LSP}} = 30$ GeV, $\tau = 2$ ns (low mass splitting).

nario considered. No other search has explored this parameter space of Higgs boson decays. In doing so, for most of the available signal grid, this search has set limits well below the best available ones provided by Higgs boson to invisible particle searches in ATLAS, which set BR limits of ∼11% [57]. For the signal points with the highest mass splittings, to which there is the highest sensitivity, limits extend all the way down to ∼1%.

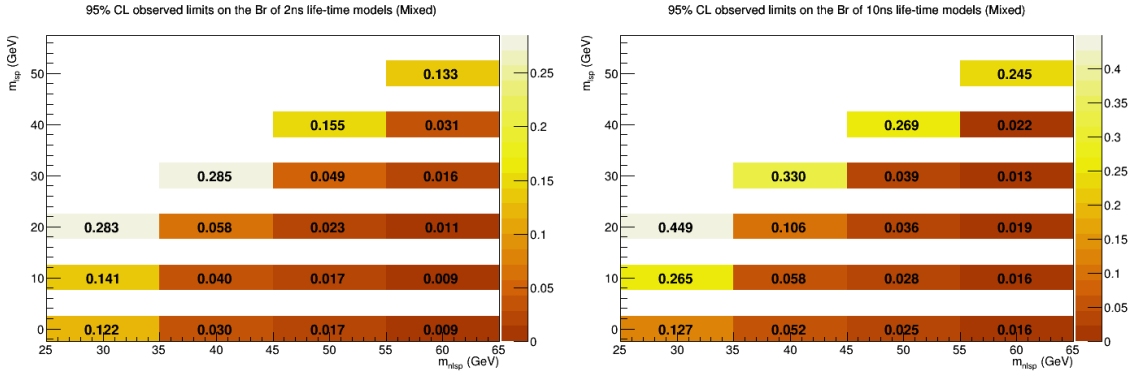


Figure 10.5: Observed upper limits at 95% CL of signal branching ratio under different assumptions of signal models based on the $SR_{t>0}$ timing shape. Each signal point uses the appropriate low or high mass splitting SR selections, as given in Table 7.3.

Chapter 11: Conclusions

This thesis has presented a novel search for displaced photons from exotic decays of the Higgs boson, specifically targeting the relatively weakly constrained branching ratio of the decay of the Higgs boson to invisible particles. The analysis exploits both the excellent timing performance and longitudinal segmentation of the LAr calorimeter to achieve unprecedented sensitivity to signals of this kind.

The analysis strategy revolves around an entirely data-driven background estimation method, using carefully constructed templates to model the pointing and timing distributions of genuine and fake photons in the high-$E_{\mathrm{T}}^{\mathrm{miss}}$ signal region. The analysis aims to remain generally model independent so as to maintain sensitivity to any soft, displaced photon signatures that can be triggered with an associated lepton. However, in order to more fully investigate sensitivity, a SUSY model in the context of GMSB was considered in which the Higgs boson decays into two, long-lived NLSP neutralinos, which each subsequently decay into a photon and a gravitino. The energy cuts, analysis region definitions, timing and pointing binning, were all optimized with respect to a model of this type, while preserving sensitivity to the more general signature of interest.

A full statistical framework, implementing a Poisson likelihood model, was developed to perform the simultaneous fitting of the timing distributions in each pointing category. A variety of systematic uncertainties were implemented into this framework, including both those impacting the signal yield, such as luminosity and instrumental uncertainties, and those impacting the shapes of the signal and background templates, including timing and modeling uncertainties. Full validations of the fitting procedure were performed to ensure that the fits were well-understood in both the VR and the $\mathrm{SR}_{t<0}$ prior to the fixing of the analysis strategy and unblinding of the $\mathrm{SR}_{t>0}$.

The final unblinded results reveal agreement with background predictions, with no significant signal component observed for any model configuration considered. Exclusion limits at the 95% CL on BR($H \to \chi_1^0\chi_1^0$) were computed for all available signal points. Some future work remains to more fully quantify the systematic uncertainties, though these will not impact the consistency of the result with background and are expected to have a relatively small effect on the final sensitivity.

This analyses represents the first displaced photon analysis performed using the full Run 2 dataset, utilizing seven times more integrated luminosity compared to the previous ATLAS result. It also obtains substantial sensitivity in the unexplored phase space of Higgs boson decays to soft, displaced photons. Accordingly, it achieves the best current limits on the Higgs boson branching ratio in the model scenario considered—as low as ∼1% for the points with the highest sensitivity. These limits lie well beyond indirect ones from Higgs boson to invisible particles searches—set at ∼11%—for much of the signal grid.

It is the hope that this analysis will serve as the pioneer Run 2 displaced photon analysis, laying a foundation for the understanding of the displaced photon object, upon which subsequent novel analyses may be built. Beyond Run 2, work is already underway for future upgrades of the LHC and the ATLAS detector to deliver and record ∼3000 fb^{-1} data from $p-p$ collisions at $\sqrt{s} = 14$ TeV. Such a large increase in available data will further expand the reach of many existing searches and open up new possibilities for probing uncharted regions of BSM phase space.

This search has illuminated a previously untouched part of Nature. And while Nature did not divulge its most alluring secrets, hope yet remains that new physics lies within the reach of the LHC. It may be the case that Nature has placed the keys to its mysteries out of our current reach, but it must be the hope of any true, warm-blooded physicist that they lie just around the corner. Perhaps we will find Nature has left a light on for us when we get there.

Bibliography

[1] T. Andeen, J. Ban, G. Brooijmans, A. Emerman, P. Kinget, J. Kuppambatti, D. Mahon, I. Ochoa, W. Sippach, and Q. Wang, Journal of Instrumentation **15** no. 04, (2020) P04012–P04012. https://doi.org/10.1088/1748-0221/15/04/p04012.

[2] D. Mahon, Journal of Instrumentation **15** no. 06, (2020) C06045–C06045. https://doi.org/10.1088/1748-0221/15/06/c06045.

[3] J. Chen, D. J. Mahon, and J. Parsons, tech. rep., CERN, Geneva, Jan, 2020. https://cds.cern.ch/record/2707701.

[4] G. 't Hooft and M. J. G. Veltman, Nucl. Phys. **B44** (1972) 189–213.

[5] P. A. M. Dirac, Proc. Roy. Soc. Lond. A **117** (1928) 610–624.

[6] F. Englert and R. Brout, Phys. Rev. Lett. **13** (1964) 321–323.

[7] P. W. Higgs, Phys. Lett. **12** (1964) 132–133.

[8] P. W. Higgs, Phys. Rev. Lett. **13** (1964) 508–509.

[9] G. Guralnik, C. Hagen, and T. Kibble, Phys. Rev. Lett. **13** (1964) 585–587.

[10] P. W. Higgs, Phys. Rev. **145** (1966) 1156–1163.

[11] T. Kibble, Phys. Rev. **155** (1967) 1554–1561.

[12] G. Aad et al., Physics Letters B **716** no. 1, (2012) 1 – 29. http://www.sciencedirect.com/science/article/pii/S037026931200857X.

[13] M. Gell-Mann, Phys. Lett. **8** (1964) 214–215.

[14] G. Zweig„ *An SU(3) model for strong interaction symmetry and its breaking. Version 2.* 2, 1964.

[15] D. J. Gross and F. Wilczek, Phys. Rev. Lett. **30** (1973) 1343–1346. https://link.aps.org/doi/10.1103/PhysRevLett.30.1343.

[16] H. D. Politzer, Phys. Rev. Lett. **30** (1973) 1346–1349. https://link.aps.org/doi/10.1103/PhysRevLett.30.1346.

[17] D. Hanneke, S. Fogwell Hoogerheide, and G. Gabrielse, Phys. Rev. A **83** (2011) 052122. https://link.aps.org/doi/10.1103/PhysRevA.83.052122.

[18] C. Abel et al., Phys. Rev. Lett. **124** (2020) 081803. https://link.aps.org/doi/10.1103/PhysRevLett.124.081803.

[19] G. 't Hooft, C. Itzykson, A. Jaffe, H. Lehmann, P. K. Mitter, I. M. Singer, and R. Stora, eds., *Recent Developments in Gauge Theories. Proceedings, Nato Advanced Study Institute, Cargese, France, August 26 - September 8, 1979*, vol. 59. 1980.

[20] S. Weinberg, Phys. Rev. D **13** (1976) 974–996. https://link.aps.org/doi/10.1103/PhysRevD.13.974.

[21] S. Weinberg, Phys. Rev. D **19** (1979) 1277–1280. https://link.aps.org/doi/10.1103/PhysRevD.19.1277.

[22] E. Gildener, Phys. Rev. D **14** (1976) 1667–1672. https://link.aps.org/doi/10.1103/PhysRevD.14.1667.

[23] L. Susskind, Phys. Rev. D **20** (1979) 2619–2625. https://link.aps.org/doi/10.1103/PhysRevD.20.2619.

[24] Muon g-2 Collaboration, G. W. Bennett et al., Phys. Rev. D **73** (2006) 072003. https://link.aps.org/doi/10.1103/PhysRevD.73.072003.

[25] G. Ciezarek et al., Nature **546** no. 7657, (2017) 227–233. https://doi.org/10.1038/nature22346.

[26] Y. Golfand and E. Likhtman, JETP Lett. **13** (1971) 323. [Pisma Zh. Eksp. Teor. Fiz. **13** (1971) 452].

[27] D. Volkov and V. Akulov, Phys. Lett. B **46** (1973) 109.

[28] J. Wess and B. Zumino, Nucl. Phys. B **70** (1974) 39.

[29] J. Wess and B. Zumino, Nucl. Phys. B **78** (1974) 1.

[30] S. Ferrara and B. Zumino, Nucl. Phys. B **79** (1974) 413.

[31] A. Salam and J. Strathdee, Phys. Lett. B **51** (1974) 353.

[32] S. P. Martin, Adv. Ser. Direct. High Energy Phys. **18** (1998) 1, arXiv:hep-ph/9709356.

[33] S. Coleman and J. Mandula, Phys. Rev. **159** (1967) 1251–1256. https://link.aps.org/doi/10.1103/PhysRev.159.1251.

[34] R. Haag, J. T. Lopuszanski, and M. Sohnius, Nucl. Phys. B **88** (1975) 257.

[35] S. P. Martin, Adv. Ser. Direct. High Energy Phys. **21** (2010) 1–153, arXiv:hep-ph/9709356.

[36] P. Fayet, Phys. Lett. B **64** (1976) 159.

[37] P. Fayet, Phys. Lett. B **69** (1977) 489.

[38] Particle Data Group, Progress of Theoretical and Experimental Physics **2020** no. 8, (2020), https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf. https://doi.org/10.1093/ptep/ptaa104. 083C01.

[39] G. R. Farrar and P. Fayet, Phys. Lett. B **76** (1978) 575.

[40] N. Sakai, Z. Phys. C **11** (1981) 153.

[41] S. Dimopoulos, S. Raby, and F. Wilczek, Phys. Rev. D **24** (1981) 1681.

[42] L. E. Ibáñez and G. G. Ross, Phys. Lett. B **105** (1981) 439.

[43] S. Dimopoulos and H. Georgi, Nucl. Phys. B **193** (1981) 150.

[44] P. Langacker, arXiv: High Energy Physics - Phenomenology (1993).

[45] J. R. Ellis, S. Kelley, and D. V. Nanopoulos, Phys. Lett. B **260** (1991) 131–137.

[46] U. Amaldi, W. de Boer, and H. Furstenau, Phys. Lett. B **260** (1991) 447–455.

[47] P. Langacker and M. Luo, Phys. Rev. D **44** (1991) 817–822. https://link.aps.org/doi/10.1103/PhysRevD.44.817.

[48] C. GIUNTI, C. W. KIM, and U. W. LEE, Modern Physics Letters A **06** no. 19, (1991) 1745–1755, https://doi.org/10.1142/S0217732391001883. https://doi.org/10.1142/S0217732391001883.

[49] L. Girardello and M. T. Grisaru, Nucl. Phys. B **194** (1982) 65.

[50] S. Deser and B. Zumino, Phys. Lett. B **62** (1976) 335.

[51] D. Z. Freedman, P. van Nieuwenhuizen, and S. Ferrara, Phys. Rev. D **13** (1976) 3214.

[52] M. Dine and W. Fischler, Phys. Lett. B **110** (1982) 227.

[53] L. Alvarez-Gaumé, M. Claudson, and M. B. Wise, Nucl. Phys. B **207** (1982) 96.

[54] C. R. Nappi and B. A. Ovrut, Phys. Lett. B **113** (1982) 175.

[55] S. P. Martin, Phys. Rev. D **55** (1997) 3177–3187. https://link.aps.org/doi/10.1103/PhysRevD.55.3177.

[56] J. D. Mason, D. E. Morrissey, and D. Poland, Phys. Rev. D **80** (2009) 115015. https://link.aps.org/doi/10.1103/PhysRevD.80.115015.

[57] ATLAS Collaboration, Tech. Rep. ATLAS-CONF-2020-052, CERN, Geneva, Oct, 2020. http://cds.cern.ch/record/2743055.

[58] J. Ellis, J. F. Gunion, H. E. Haber, L. Roszkowski, and F. Zwirner, Phys. Rev. D **39** (1989) 844–869. https://link.aps.org/doi/10.1103/PhysRevD.39.844.

[59] U. Ellwanger, C. Hugonie, and A. M. Teixeira, Phys. Rept. **496** (2010) 1–77, arXiv:0910.1785 [hep-ph].

[60] M. Maniatis, Int. J. Mod. Phys. A **25** (2010) 3505–3602, arXiv:0906.0777 [hep-ph].

[61] P. Meade, M. Reece, and D. Shih, JHEP **10** (2010) 067, arXiv:1006.4575 [hep-ph].

[62] ATLAS Collaboration, C. Bernius, PoS **EPS-HEP2015** (2015) 158.

[63] K. P. Mistry, C. Meyer, and H. Williams, Tech. Rep. ATL-COM-PHYS-2018-196, CERN, Geneva, Mar, 2018. https://cds.cern.ch/record/2306702.

[64] V. Khachatryan et al., Physics Letters B **753** (2016) 363 – 388. http://www.sciencedirect.com/science/article/pii/S037026931500965X.

[65] ATLAS Collaboration, G. Aad et al., Phys. Rev. D **88** (2013) 012001. https://link.aps.org/doi/10.1103/PhysRevD.88.012001.

[66] ATLAS Collaboration, G. Aad et al., Phys. Rev. D **90** (2014) 112005. https://link.aps.org/doi/10.1103/PhysRevD.90.112005.

[67] CMS Collaboration, A. Sirunyan et al., Phys. Rev. D **100** (2019) 112003. https://link.aps.org/doi/10.1103/PhysRevD.100.112003.

[68] S. Baird, tech. rep., CERN, Geneva, Feb, 2007. https://cds.cern.ch/record/1017689.

[69] L. Evans and P. Bryant, JINST **3** (2008) S08001.

[70] S. van der Meer, Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968. https://cds.cern.ch/record/296752.

[71] E. Mobs,. https://cds.cern.ch/record/2636343. General Photo.

[72] ATLAS Collaboration, JINST **15** no. 10, (2020) P10004–P10004. https://doi.org/10.1088%2F1748-0221%2F15%2F10%2Fp10004.

[73] ATLAS Collaboration, "Luminosity public results run 2." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2.

[74] ATLAS Collaboration, JINST **3** (2008) S08003. http://stacks.iop.org/1748-0221/3/i=08/a=S08003.

[75] ATLAS Collaboration, M. Capeans, G. Darbo, K. Einsweiller, M. Elsing, T. Flick, M. Garcia-Sciveres, C. Gemme, H. Pernegger, O. Rohne, and R. Vuillermet, tech. rep., Sep, 2010. https://cds.cern.ch/record/1291633.

[76] B. Abbott et al., JINST **13** (2018) T05008, arXiv:1803.00844 [physics.ins-det].

[77] ATLAS Collaboration, Tech. Rep. CERN-LHCC-96-041 ATLAS-TDR-2, CERN, Geneva, 1996.

[78] G. Aad et al., The European Physical Journal C **70** no. 3, (2010) 755–785. https://doi.org/10.1140/epjc/s10052-010-1403-6.

[79] W. Cleland and E. Stern, NIM A **338** no. 2, (1994) 467 – 497. http://www.sciencedirect.com/science/article/pii/0168900294913323.

[80] ATLAS Collaboration, "Approved plots DAQ." https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ.

[81] ATLAS Collaboration, G. Aad et al., Eur. Phys. J. C **77** (2016) 490. 87 p, arXiv:1603.02934. https://cds.cern.ch/record/2138166.

[82] ATLAS Collaboration, M. Aaboud et al., Eur. Phys. J. C **79** no. arXiv:1902.04655. 8, (2019) 639. 40 p. https://cds.cern.ch/record/2657964.

[83] ATLAS Collaboration, JINST **14** (2019) P12006, arXiv:1908.00005 [hep-ex].

[84] ATLAS Collaboration, Eur. Phys. J. C **76** (2016) 292, arXiv:1603.05598 [hep-ex].

[85] ATLAS Collaboration, Eur. Phys. J. C **77** (2017) 466, arXiv:1703.10485 [hep-ex].

[86] R. Hyneman et al., Tech. Rep. ATL-COM-PHYS-2019-432, CERN, Geneva, Apr, 2019. https://cds.cern.ch/record/2672803.

[87] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, E. Moyse, and A. Salzburger, Tech. Rep. ATL-SOFT-PUB-2007-007. ATL-COM-SOFT-2007-002, CERN, Geneva, Mar, 2007. https://cds.cern.ch/record/1020106.

[88] ATLAS Collaboration, Tech. Rep. ATLAS-CONF-2010-072, CERN, Geneva, Jul, 2010. https://cds.cern.ch/record/1281363.

[89] ATLAS Collaboration, M. Aaboud et al., JINST **14** no. 03, (2019) P03017, arXiv:1812.03848 [hep-ex].

[90] Tech. Rep. ATL-PHYS-PUB-2011-007, CERN, Geneva, Apr, 2011. https://cds.cern.ch/record/1345329.

[91] Tech. Rep. ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr, 2011. https://cds.cern.ch/record/1345327.

[92] ATLAS Collaboration, ATLAS Collaboration, arXiv:2007.02645 [hep-ex].

[93] ATLAS Collaboration, Eur. Phys. J. C **76** (2016) 581, arXiv:1510.03823 [hep-ex].

[94] M. Cacciari, G. Salam, and G. Soyez, JHEP **04** (2008) 063, arXiv:0802.1189 [hep-ph].

[95] ATLAS Collaboration, M. Aaboud et al., Eur. Phys. J. C **78** no. arXiv:1802.08168. 11, (2018) 903. 66 p. https://cds.cern.ch/record/2305380.

[96] M. Baak et al., tech. rep., CERN, Geneva, Mar, 2009. https://cds.cern.ch/record/1168026.

[97] G. Aad et al., The European Physical Journal C **80** no. 1, (2020) 47. https://doi.org/10.1140/epjc/s10052-019-7500-2.

[98] E. Gross, CERN Yellow Rep. School Proc. **3** (2018) 199–221.

[99] G. Avoni et al., Journal of Instrumentation **13** no. 07, (2018) P07017–P07017. https://doi.org/10.1088%2F1748-0221%2F13%2F07%2Fp07017.

[100] ATLAS Collaboration, The European Physical Journal C **76** no. 12, (2016) 653. https://doi.org/10.1140/epjc/s10052-016-4466-1.

Appendices

Appendix A: Timing Studies

A.1 Neighboring Bunch Crossings in Timing Distributions

Photon timing distributions in the CR show evidence of objects from neighboring bunch crossings beyond ∼12.5 ns, as shown in Figure A.1.

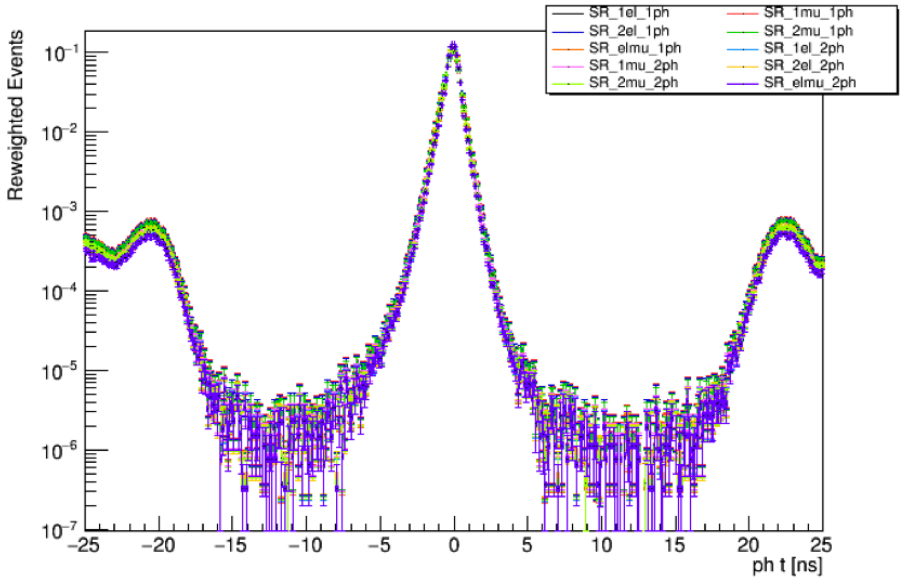


Figure A.1: Photon timing distributions in the CR reweighted to different SR E_{cell} distributions (in different colors).

A.2 Mismodeling of Timing in Simulation

A mismodeling of the MC timing appears at 12.5 ns, halfway to the neighboring bunch crossing. A clear discontinuity in the relationship between the reconstructed time and truth

time exists, as shown in Figure A.2. Before $|t| < 12.5$ ns, the reconstructed and truth times correspond very well in a linear fashion. However, at 12.5 ns, there is a sudden discontinuity followed by an unusual nonlinear trend thereafter. Looking at the difference between the times vs. the truth time, it is also clear that there is no modeling of the known bias in timing that occurs for pulses that arrive significantly delayed since the relationship is flat below 12.5 ns. The derivative from the digitized samples is used to extrapolate to delayed samples. However, at 12.5 ns, a switch is made to the next sample, but the same OFCs are used, introducing a discontinuity and bias.

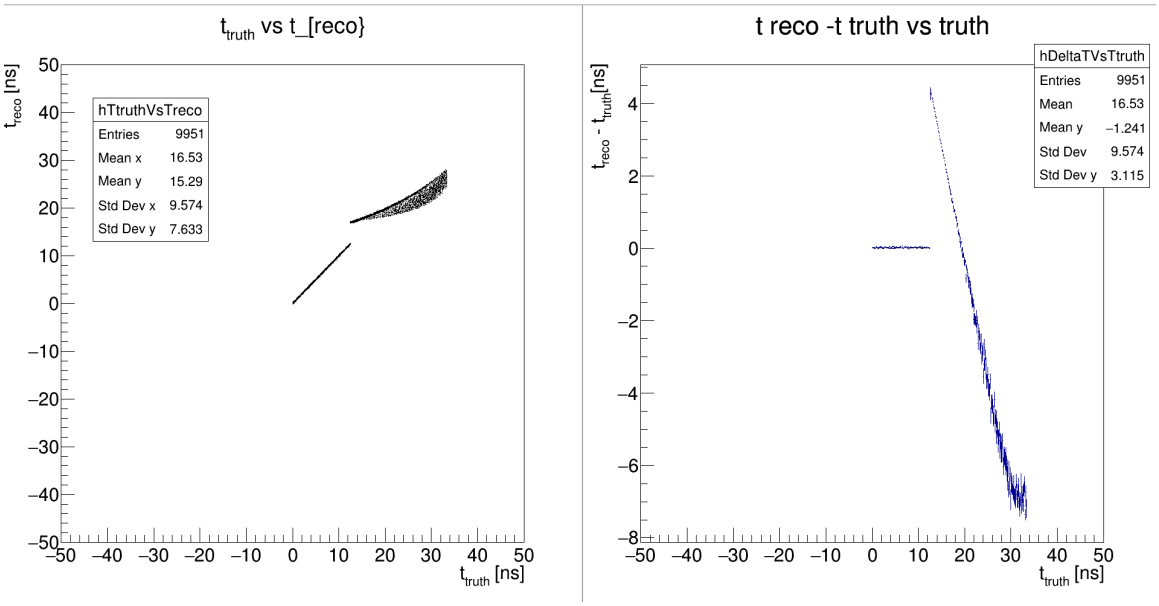


Figure A.2: Reconstructed time vs. truth time for a single photon gun simulation sample (left). The difference between these times vs. truth time for the same sample (right).

Further evidence of this strange behavior after 12.5 ns can be seen in the overall timing distribution from officially-produced delayed signal samples can be seen in Figure A.3, where the distribution has a distinct, qualitative change at 12.5 ns despite the fact that the truth time comes from an exponential distribution.

A.3 Low-Energy Photon Correction

The timing calibration procedure, described in Section 5.2.2.2, uses electrons to derive the timing corrections, but they are applied to both electrons and photons. Past studies have

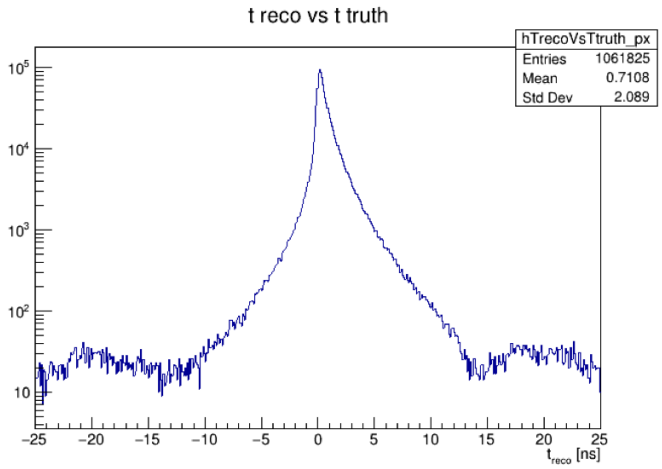
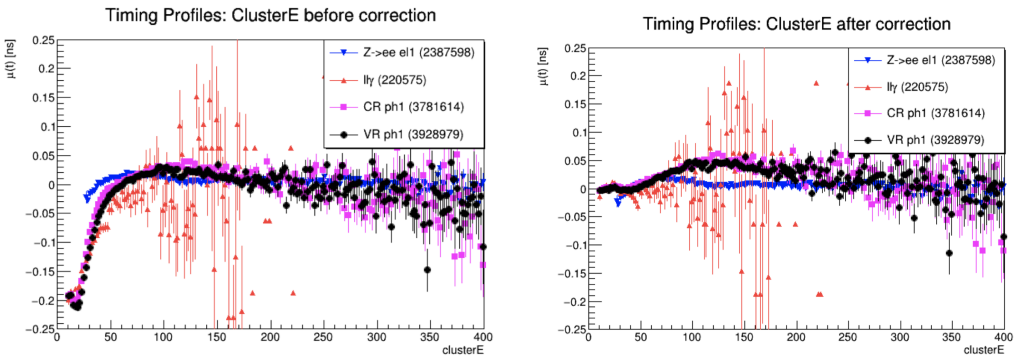


Figure A.3: Reconstructed time from the sum of all available signal samples.

confirmed that, for timing purposes, electrons and photons behave similarly. However, at extremely low E_{cell}, noticeable non-zero photon timing means emerge while electron timing remains well-behaved, presumably reflecting differences in the EM shower development of electrons and photons at low energies.

For this divergence, a new correction was devised. It was observed that the mean timing behavior of photons various analysis samples—including photons from the CR, VR, and the radiative Z selection—was qualitatively similar as a function of E_{cell}, but significant differences in the magnitude of the non-zero deviation existed. These differences are likely related to the purities of the various samples, some of which contain a much more jet contamination. However, for this analysis, a more or less universal set of corrections for all EM objects passing at least the loose photon ID requirements is desired since they all enter into the final fit. Thus, E_{cell} was not considered the best variable as a basis for the correction since a single correction would not account for the variation among different photon selections. Dedicated studies determined that this effect is more uniformly modeled be the energy of the full EM cluster rather than a single cell. A plot of the mean times for several selection before the new correction was applied is shown in Figure A.4(a).

Since the mean timing behavior as a function of cluster energy of photons and jets appears to be similar, it was decided to extract an additional set of corrections based on this variable. This correction is applied to objects identified as photons on top of the electron-



(a) (b)

Figure A.4: Timing profiles in cluster energy for several samples (a) before and (b) after applying the timing correction. Electrons from a $Z \to ee$ selection are shown in blue; loose photons from a radiative Z decay (enriched in genuine photons) in red; loose photons with $E_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV (enriched in jets) in magenta; and loose photons with $20 < E_{\mathrm{T}}^{\mathrm{miss}} < 40$ GeV in black.

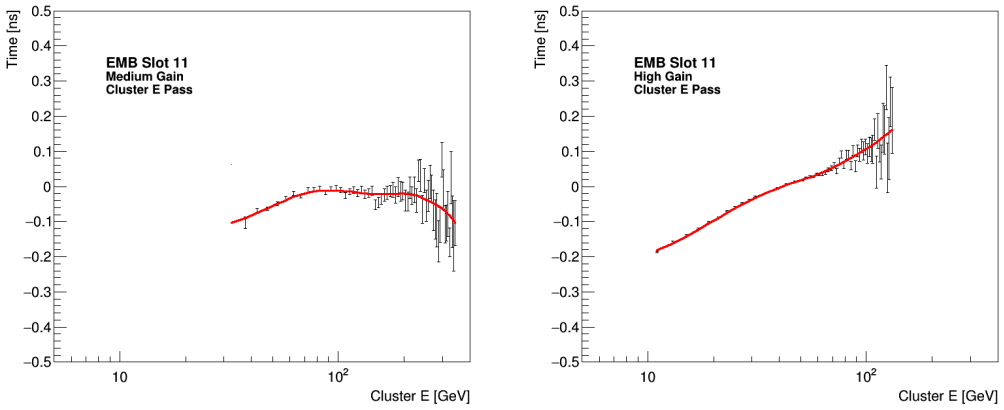
based timing calibrations. The calibration sample used (due to its abundant statistics) includes all photons in the analysis with $E_{\mathrm{T}}^{\mathrm{miss}} < 40$ GeV but excluding events compatible with radiative $Z\gamma$ decays. The corrections are provided using a smoothing algorithm[1] on the binned mean time vs. EM cluster energy relationship. Deviations from zero are subtracted out such that the final corrected curve is flattened. Corrections are derived independently in bins of $|\eta|$ (by FEB slot) and by gain (medium and high).

Figure A.5 shows an example of the cluster energy profiles for a slice of low-η photons (one slot) from which the corrections are derived. As shown in Figure A.4(b), the new corrections flatten the timing profiles. Table A.1 summarizes the effect on the mean of the background timing distributions.

Sample	$\bar{t}$ before correction [ps]	$\bar{t}$ after correction [ps]
VR photons	-79.9	11.6
CR photons	-59.9	14.8
$Z \to ll\gamma$ photons	-109.5	-7.6

Table A.1: Effect of the additional cluster-energy-dependent timing correction for low-energy photons on the distribution means.

[1]Specifically, a variable span smoother based on local linear fits



(a) (b)

Figure A.5: Timing profiles in cluster energy for photons with $E_{\mathrm{T}}^{\mathrm{miss}} < 40$ GeV reconstructed in (a) medium and (b) high gain.

A.4 MC Smearing

The analysis uses a fully data-driven prediction for the prompt SM background. MC samples are therefore only used in the determination of the timing distributions expected in signal. Studies in $Z \to ee$ data have measured the timing resolution for electrons after the timing calibration is applied [3]. The timing resolution depends on the energy deposited in the calorimeter cell, on the cell position in η, and on the front-end board (FEB) electronics gain in which the calorimeter cell is read out. The timing performance is expected to have a dependence on E_{cell} of the form of Equation 5.5. The measured resolutions are provided in η ranges which are serviced by the same type of FEB which, on the detector, corresponds to a particular slot number in which the FEB is installed within the front-end crate. The measured values for the noise (p_0) and constant (p_1) terms for each slot in the EMB, as well as for the entire EMB, are shown in Table A.2 for the high and medium gains. No results are shown for the low gain, for which calibration was not performed due to lack of statistics.

The measured resolutions include a correlated component of order 190 ps, which is understood to originate from the uncertainty in the p-p collision time among the many protons in a bunch. MC samples contain a very small correlated component of about 55 ps of unknown origin. Subtracting this already baked-in component from the desired total

Calorimeter Detector	Slot Number	Appox. $\|\eta\|$ Range	High Gain		Medium Gain	
			Value of p_0 [ps/GeV]	Value of p_1 [ps]	Value of p_0 [ps/GeV]	Value of p_1 [ps]
EMB	11	[0.0, 0.4]	2069	208	2763	217
	12	[0.4, 0.8]	2145	223	3302	227
	13	[0.8, 1.2]	2796	242	4982	222
	14	[1.2, 1.37]	3170	237	4527	224
	All	[0.0, 1.37]	1962	262	3650	223

Table A.2: Values of the fit parameters p_0 (coefficient of noise term) and p_1 (constant term) when fitting the expression of Equation 5.5 of the time resolution versus cell energy for the various slots of the EMB and EMEC calorimeters.

correlated component of 190 ps, the value of 180 ps is used to be added in quadrature to the native calorimeter timing performance. The spread in these collision times is observed to be quite constant over time, with deviations of no more than ∼5% from the nominal value over all IOVs in Run 2, as shown in Figure A.6.

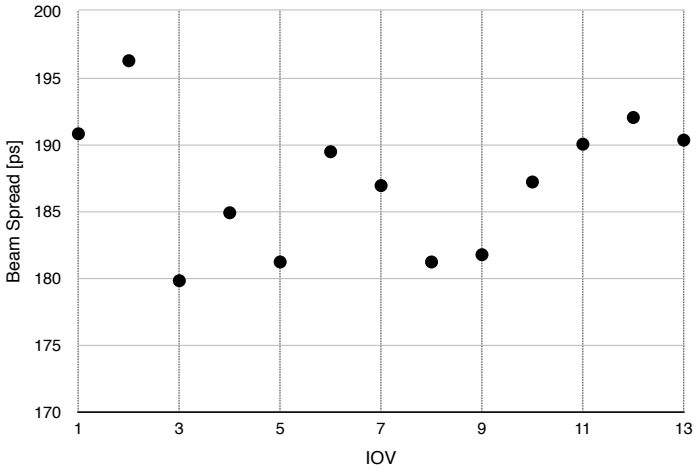


Figure A.6: Correlated timing resolution component attributable to the LHC beam spread for each IOV in Run 2 computed from $Z \to ee$ electrons, both with energy > 20 GeV.

In ATLAS MC simulation samples, no ad-hoc smearing is applied to simulate the beam spread in time, as is done for the beam spread in space along the z-axis. However, smearing is applied during the digitization of the simulated pulse ADC samples to account for the electronic and expected pileup noise, which affects both the energy and timing resolution performance. These facts were first verified in simulated electrons in the signal MC samples

used in the analysis by employing the fact that those electrons are produced promptly in the associated SM-Higgs production. Subsequently, a full study was performed with simulated prompt $Z \to ll\gamma$ samples in order to make sure no bias was introduced in the photon timing given the small differences between the electron and timing performance at low cell energies observed previously (see A.3).

Figure A.11 shows the photon timing resolution as a function of cell energy as measured in data from a $Z \to ll\gamma$ sample selected from the pre-selected analysis data. The timing resolution was measured after the corrections from the electron-based timing calibration procedure and the additional low-energy photon correction were applied. The resolution is estimated by a Gaussian fit to the core of the distribution of each energy slice. Also shown overlaid are similar resolution curves measured in the prompt $Z \to ll\gamma$ MC samples mentioned above, after the expected 180 ps beam spread component was added to the reconstructed photon time. A random "collision time" was selected once per event from a Gaussian distribution with mean $\mu = 0$ and $\sigma = 180$ ps and subsequently added to each photon time. This procedure results in a global shift of the MC resolution curve upwards. No calibration is needed to be applied to MC, however a minor correction to the time-of-flight of the photon due to position of the primary vertex was applied. As can be seen, after the addition of the random collision time, the MC follows the shape of the data resolution curve relatively well. However, some non-constant residuals remain across the energy range.

An additional residual smearing component was therefore calculated as a function of energy and slot by subtracting, bin-by-bin in quadrature, the resolution curves shown above in MC from the corresponding resolutions in data. For each resulting bin, an uncertainty on the central value of the residual was assigned by propagating the uncertainties on the resolutions in data and MC. Smoothed curves of the residuals as a function of energy were obtained using a smoothing function. Uncertainty bands for each residual curve were defined by repeating the smoothing procedure on the nominal residual points shifted by two times their statistical uncertainties. The residual smearing remaining after this subtraction is shown in Figure A.8.

These smoothed residual curves as a function of E_{cell} are then used to provide the additional uncorrelated component to be added to MC. For the up (down) systematic un-

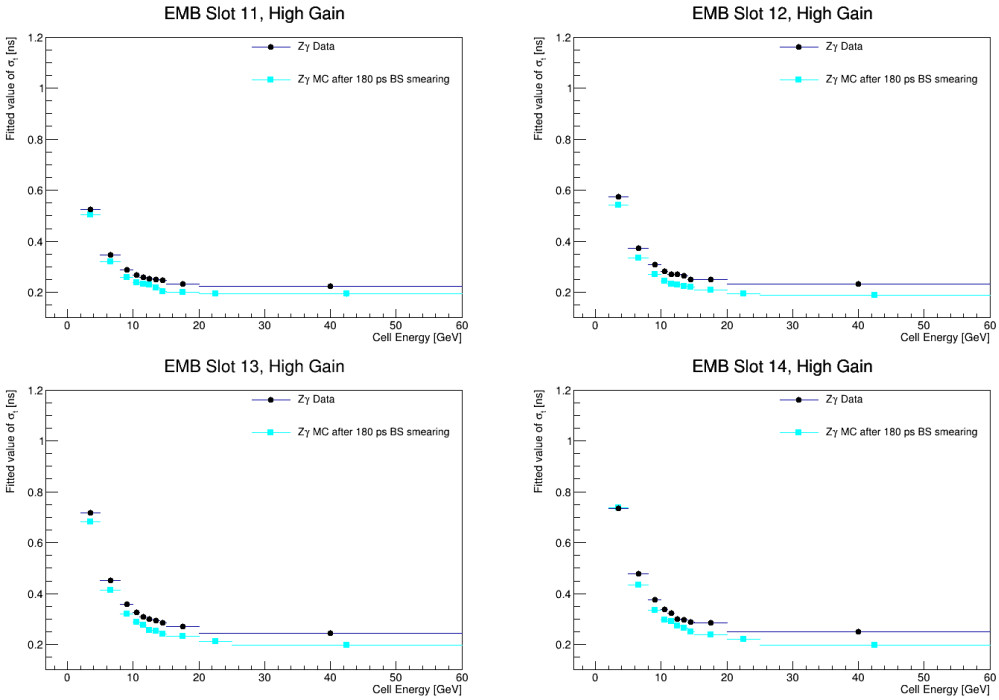


Figure A.7: Timing resolution as a function of cell energy in high gain for the four EMB slots, for photons in $Z \to ll\gamma$ data and MC. The MC is shown after smearing with the 180 ps beam-spread.

certainties associated with this smearing procedure, the upper (lower) extent of the residual envelope (indicated with dotted lines) is used.

The smearing procedure was applied to the same prompt MC sample and the resulting measured timing resolution curves are shown in Figure A.9 in comparison with the photon timing resolution measured in data in an inclusive analysis CR with $E_{\mathrm{T}}^{\mathrm{miss}} < 40$ GeV. As can be seen, with this smearing procedure, the prompt MC timing resolution matches the resolution measured in data reasonably well, and the systematic variations provide adequate coverage of any fluctuations.

A similar procedure was performed to obtain smearing parameters for the medium gain. However, binning in energy as not performed, and instead just a single constant parameter was used per slot. This was done for two reasons—one practical and one physically motivated. First, since the statistics are relatively limited at these higher energies, multiple bins result in large statistical fluctuations. And second, the energy-dependence for medium gain objects is relatively flat due to the dominance of the constant term over the noise term at

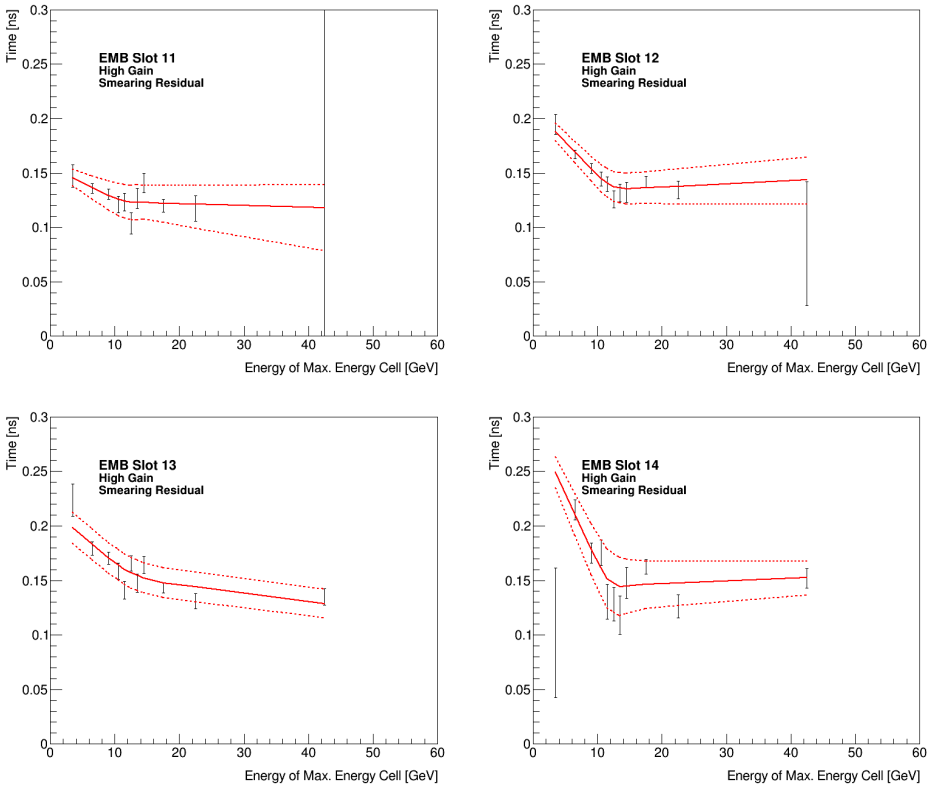


Figure A.8: Additional smearing as a function of cell energy required to match the timing resolution for photons observed in data for each of the four barrel calorimeter slots in high gain. The band defined with the dotted lines shows the up/down systematic variations of the smearing procedure.

these energies.

A procedure similar to that for high gain objects described above was carried out for medium gain, including the application of the beam spread and calculation of residuals. In order to ensure adequate coverage, residual values were inflated by three times the propagated fit uncertainty ($\pm 3\sigma$) to define the uncertainty bands. The calculated additional smearing parameters for the medium gain as well as their systematic variations are listed for each slot in the EMB on Table A.3. For completeness, the table includes the approximate effective residuals for high gain and their systematic variations with the additional requirement that $E_{\text{cell}} > 20$ GeV (beyond the plateau of the timing resolution curve).

Figure A.10 shows the results of applying the smearing procedure and its systematic variations to the prompt $Z \to ll\gamma$ sample for simulated photons reconstructed with medium

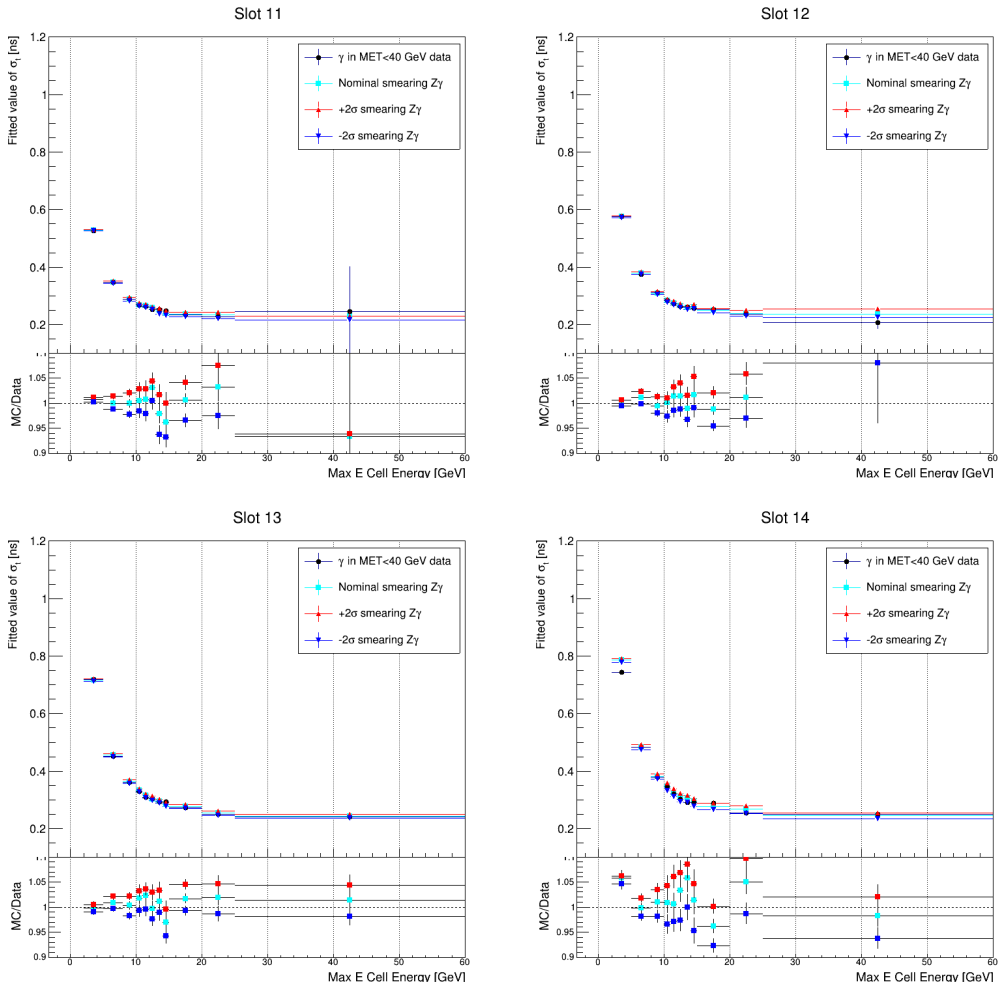


Figure A.9: Timing resolution as a function of cell energy in high gain for the four EMB slots, for photons in $Z \to ll\gamma$ data (black points) and MC. The MC is shown after applying the smearing procedure which includes the 180 ps beam spread and the nominal values of the smoothed slot- and energy-dependent residuals (cyan) as well as its up/down systematic variations (red/blue). The bottom panels show the ratios of the MC resolution curves (nominal and up/down variations) with respect to the data.

gain. The fitted timing resolution σ is shown as a function of E_{cell} and is compared to photons in data from an inclusive analysis CR with $E_{\text{T}}^{\text{miss}} < 40$ GeV, also reconstructed in medium gain. A relatively good agreement with the data resolutions and an adequate systematic variation coverage is observed.

Calorimeter Detector	Slot Number	Approx. $\lvert\eta\rvert$ Range	Additional smearing [ps]			
			High Gain (approx. for $E_{\text{cell}} > 20$ GeV)		Medium Gain (constant)	
			Nominal	Down/Up Variation	Nominal	Down/Up Variation
EMB	11	[0,0.4]	120	100/138	145	132/158
	12	[0.4,0.8]	137	122/152	138	124/152
	13	[0.8,1.2]	146	132/159	108	76/140
	14	[1.2,1.37]	147	126/168	105	68/142

Table A.3: Additional MC photon time smearing on top of the 180 ps beam spread needed to match the photon timing resolution measured in data. For high gain, the approximate values for high energy depositions are shown, while for medium gain the single constant values are shown for each slot. The Up/Down systematic variations are also listed.

A.5 Investigation of Prompt Electron Timing Correction to Photon Timing

The beam spread adds a collision time uncertainty of about 190 ps [3]. Such a variation could be significant for this analysis, where the shape of the photon timing distribution is a crucial element of the analysis construction. The associated production mechanism is expected to have prompt electrons or muons. Electron timing is well understood and calibrated, and since the electron and photon share the same collision time (originating from the PV of the event), it is possible that the electron time can be used to, in effect, anchor the photon time. Thus, rather than considering the photon time alone, which contains both the collision time and the time to reach the calorimeter, the prompt electron time can be subtracted from the photon time to yield a more precise value. In other words, it is possible, in principle, that using the difference between electron and photons times has a lower timing resolution than the photon alone. Let us first examine this possibility theoretically to get a conceptual idea of the effects at play.

We assume that the measured time t for prompt events can be written as follows:

$$\begin{aligned} t &= t_{\text{true}} + t_{\text{corr}} + t_{\text{uncorr}} \\ &= 0 + t_{\text{corr}} + t_{\text{uncorr}} \\ &= \mathcal{N}(\sigma_{\text{beam}}) + \mathcal{N}(\sigma_{\text{uncorr}}) \end{aligned}$$

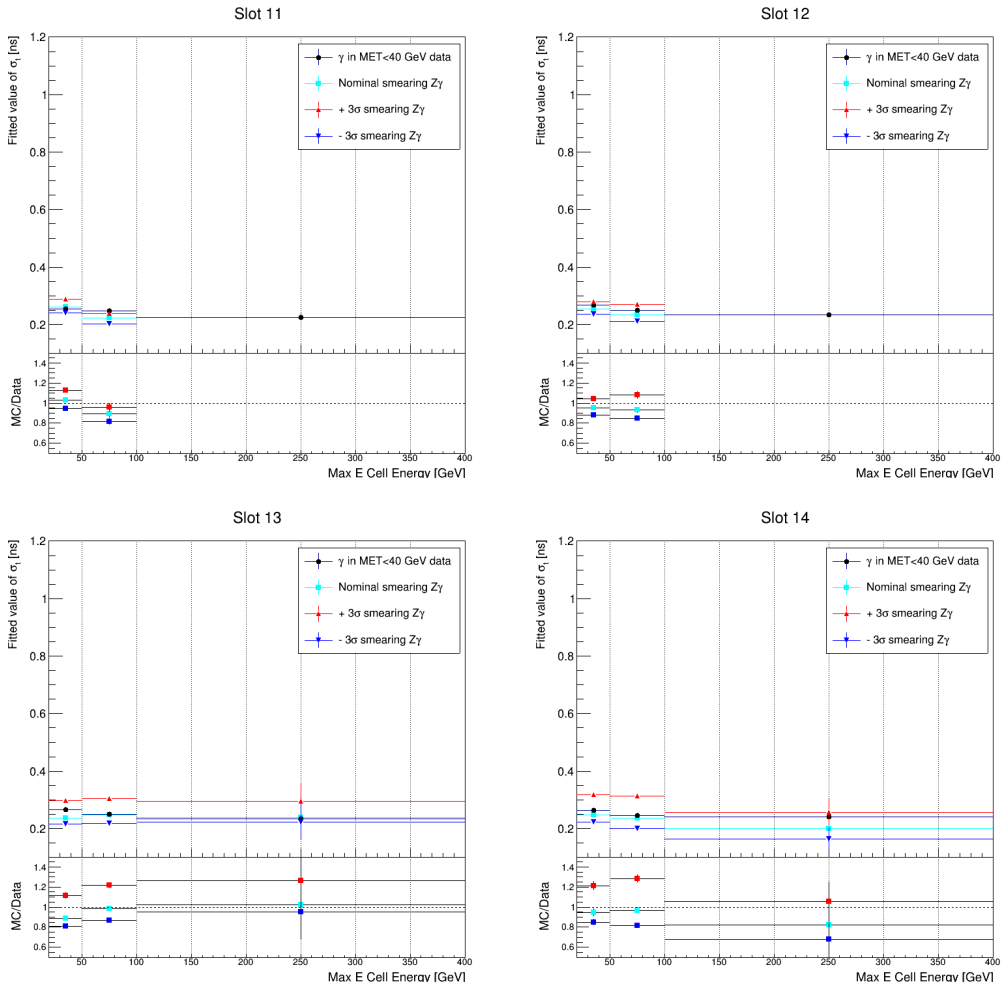


Figure A.10: Timing resolution as a function of cell energy in medium gain for the four EMB slots, for photons in $Z \to ll\gamma$ data (black points) and MC. The MC is shown after applying the smearing procedure which includes the 180 ps beam spread and the nominal values of the constant residuals (cyan) as well as its up/down systematic variations (red/blue) for each slot. The bottom panels show the ratios of the MC resolution curves (nominal and up/down variations) with respect to the data.

where $\mathcal{N}(\sigma)$ denotes a Gaussian distribution with mean zero and standard deviation σ.

Consider the time measurements of a prompt electron and a prompt photon: t_e and t_γ, respectively. The covariance of these two time measurements is given by:

$$\begin{aligned}
\sigma_{e\gamma} &= < (t_e - < t_e >)(t_\gamma - < t_\gamma >) > \\
&= < t_e t_\gamma > \\
&= < (t_{\mathrm{corr,e}} + t_{\mathrm{uncorr,e}})(t_{\mathrm{corr},\gamma} + t_{\mathrm{uncorr},\gamma}) > \\
&= < t^2_{\mathrm{corr}} > \\
&= \sigma^2_{\mathrm{corr}} + < t_{\mathrm{corr}} >^2 \\
&= \sigma^2_{\mathrm{corr}} = \sigma^2_{\mathrm{beam}}
\end{aligned} \tag{A.1}$$

If we subtract the electron time from the photon time, the time resolution of the resulting difference is:

$$\sigma_{t_\gamma - t_e} = \sqrt{\sigma^2_\gamma + \sigma^2_e - 2\sigma_{e\gamma}}$$

If we require that the resolution improves by using the electron time, i.e. $\sigma_{t_\gamma - t_e} < \sigma_\gamma$, then we must have:

$$\sigma_{e\gamma} > \frac{1}{2}\sigma^2_e \tag{A.2}$$

Combining equations A.5 and A.5, we get:

$$\sigma_e < \sqrt{2}\sigma_{\mathrm{beam}}$$

As determined in the timing calibration procedure, the resolution σ_e can be understood as the quadrature sum of an E_{cell}-dependent noise term and a constant term (p_0 and p_1, respectively). We can then write the following requirement:

$$\sqrt{\left(\frac{p_0}{E_{\text{cell}}}\right)^2 + p_1^2} < \sqrt{2}\sigma_{\text{beam}}$$

$$E_{\text{cell}} > \frac{p_0}{\sqrt{2\sigma_{\text{beam}}^2 - p_1^2}}$$

From the timing calibration with $Z \to ee$ data, $\sigma_{\text{beam}} = 190$ ps, and for electrons in the barrel, $p_0 \sim 2700$ GeV $\cdot$ ps and $p_1 \sim 230$ ps. Therefore, in order for an electron to improve the timing resolution of photons in the same event, we must have $E_{\text{cell}} \gtrsim 19$ GeV. Using similar reasoning now considering the use of the average electron time for events with two electrons, we arrive at an approximate requirement of $E_{\text{cell}} \gtrsim 11$ GeV in order for two approximately equal-energy electrons in the barrel to improve the timing resolution of photons in the same event.

E_{cell} comprises roughly 15–40% of the total energy of the photon cluster. Therefore, in the single photon case, a total electron energy of ∼50–125 GeV would be required. However, in this analysis, electrons originate from W, Z, and $t\bar{t}$ decays, which do not often have such high energy electrons. Therefore, it is expected that using the lepton to anchor the photon time will in fact degrade the photon resolution simply due to the fact that the electrons in our final states are too soft to provide precise enough timing to serve as a reliable anchor.

Studies with data confirm this expectation. Figures A.11 and A.12 compare the uncorrected and corrected timing distributions of the leading photon, in a control region with exactly one electron and one or two photons respectively.

Figures A.13 and A.14 show the correlations in timing between the electron and photon, or leading and subleading electron, respectively.

These studies indicate that the electron and photon in a typical analysis event are weakly correlated in timing. In addition to the energy-dependent considerations discussed, there is also the fact that at such relatively low energies, the electron and photon may not reliably come from the same bunch crossing. Therefore, the electron-corrected photon timing is not used in the analysis.

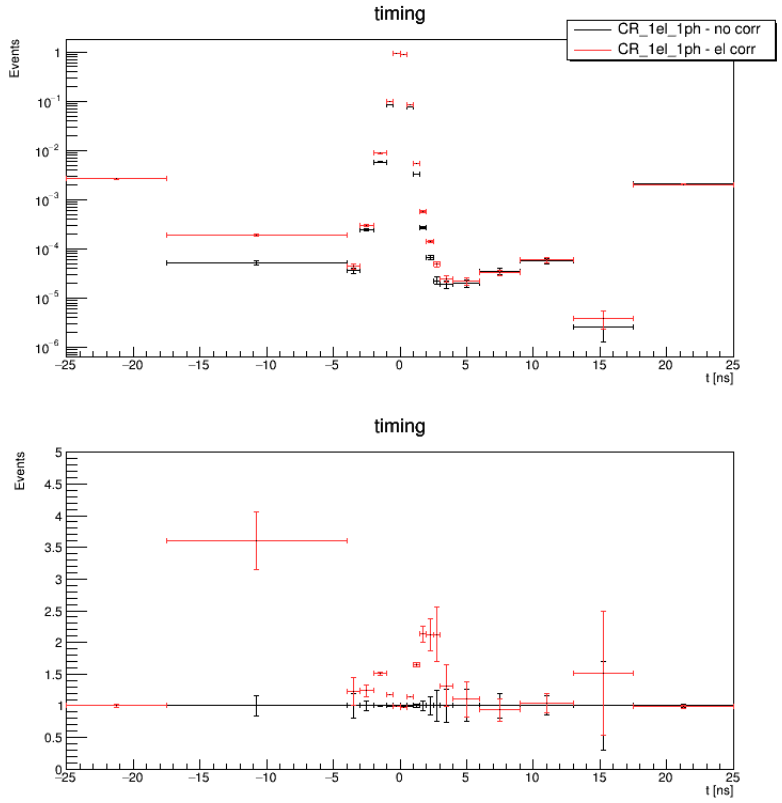


Figure A.11: Comparison of uncorrected vs. electron-corrected timing distributions for the leading photon in the 1 electron 1 photon control region.

A.6 Symmetry of Timing Distributions

The $SR_{t<0}$ region is used throughout the analysis for various validation purposes. In doing so, the timing distribution is reflected across $t = 0$, which is taken as a prediction of the background timing distribution for positive times. However, this operation assumes that there are no significant shape difference between the negative and positive sides of the timing distribution. After the additional photon cluster energy correction, E_{cell} reweighting, and the shifting of the means of the overall timing distributions to 0, it is expected that this symmetry is in fact present. This assumption is validated in Figure A.15 for the various regions considered in the analysis. The lower ratio plot confirms that the positive and negative sides are generally consistent with each other. Some fluctuations are present in the tails; however, the statistical uncertainties are large, and the fit behavior is dominated by

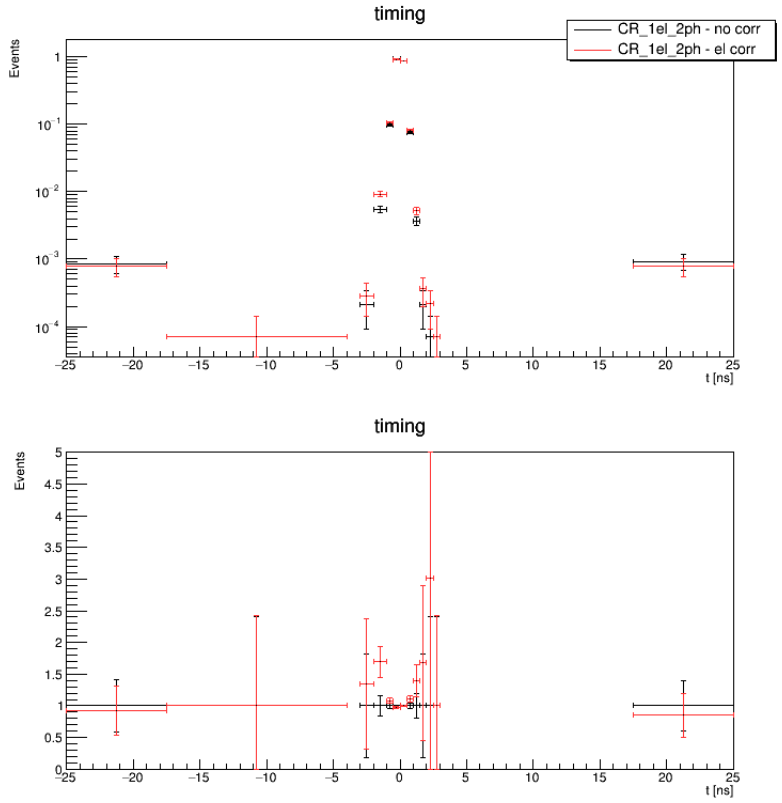


Figure A.12: Comparison of uncorrected vs. electron-corrected timing distributions for the leading photon in the 1 electron 2 photon control region.

the statistics-dominant core of the distributions.

A.7 Timing Bias for Non-Prompt Signals

Since the timing measurements are optimized for prompt signals via the computation of OFCs during calibration, a small bias exists for non-prompt signals. Data from special runs in 2018 when the solenoid magnet was off allow for the reconstruction of LAr pulses with a fine sampling of different trigger delays. This data allows for an unprecedented understanding of the timing bias as a function of signal delay. This study can be cross-checked in physics data recorded in nominal conditions by examining the satellite-satellite collision peaks known to occur at ±5 ns. The actual time of such collisions is known to be quite precisely at ±5 ns due to the required operating tolerances of the RF cavities that

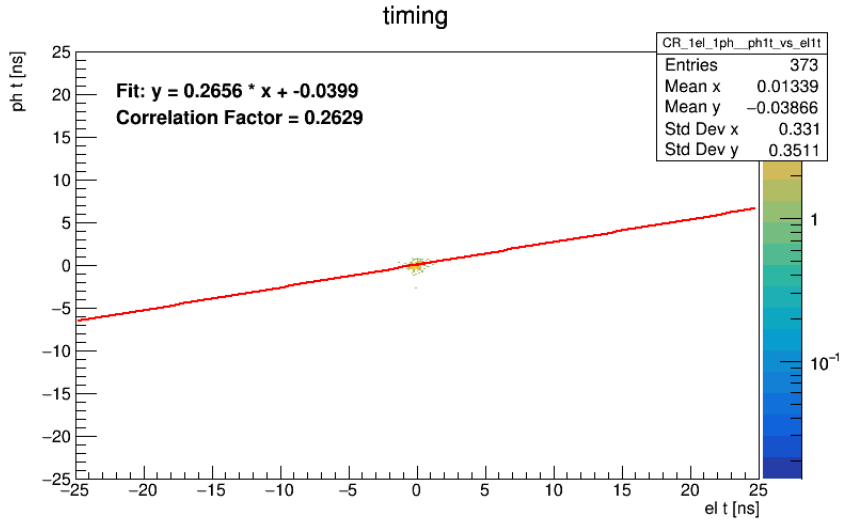


Figure A.13: Correlation between electron and photon timing in the 1 electron 1 photon CR.

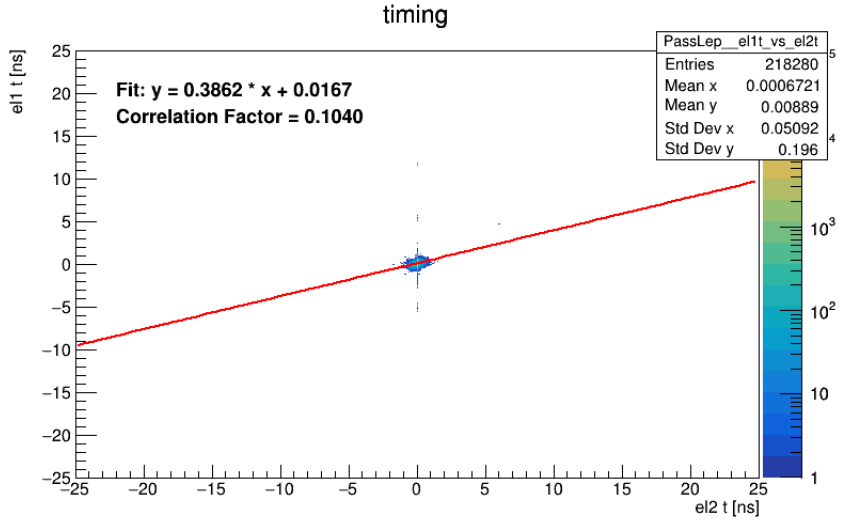
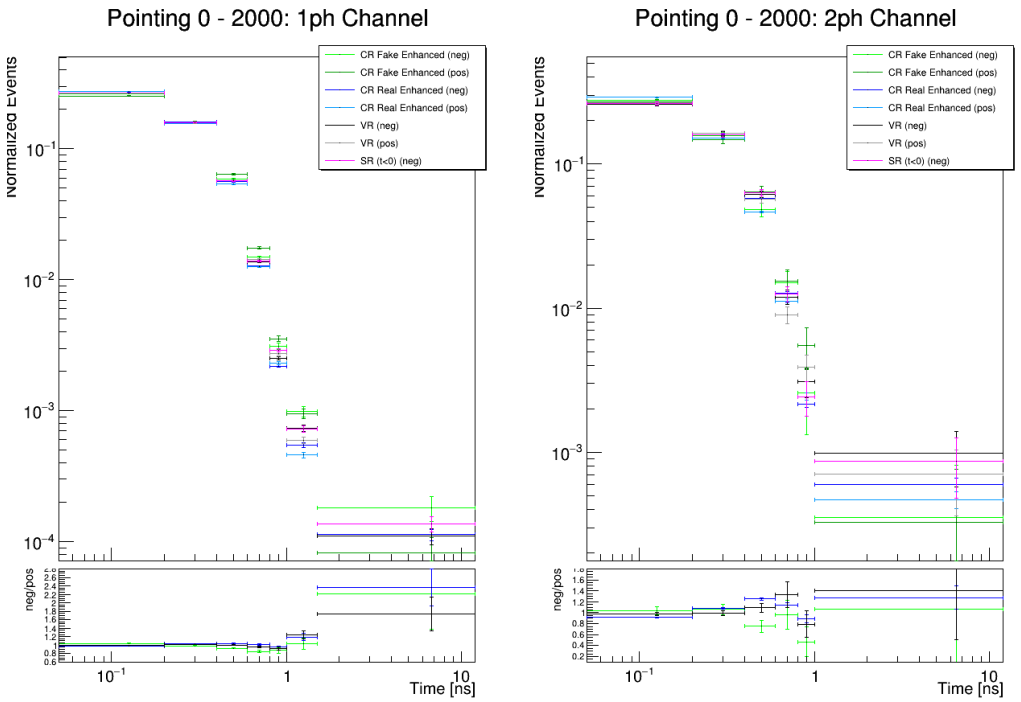


Figure A.14: Correlation between leading and subleading electron timing in the 2 electron CR.

determine the collision times of these bunches. Data from the calibration of the timing using W boson decays is used for this cross-checking since it is plentiful and well-studied. The timing bias for one calorimeter slot as a function of delay from the special run studies is shown in Figure A.16. Note that, for a delay of +5 ns, the measured timing bias is ∼+0.5 ns; and for a delay of -5 ns, it is ∼-0.2 ns. The ±5 ns satellite-satellite collision peaks, measured



(a) (b)

Figure A.15: Positive timing distributions overlaid with reflected negative timing distributions for the A.15(a) one-photon and A.15(b) two-photon channels. Bottom ratio plot shows that shapes are generally consistent between negative and positive sides.

from W boson decays, are shown in Figure A.17. Here, note that the positive and negative peaks occur at 5.55 and -5.17 ns, respectively, achieving excellent agreement with the special run bias measurements.

Plots such as Figure A.16 for each slot can thus be used to understand the nature of the timing bias quite precisely, and MC times can be adjusted accordingly. Using the uncertainties from these profile plots can help to define a systematic to be applied to account for any residual effects.

Additionally, it should be emphasized that this adjustment of MC times adds the effect of this known bias to match data. It is thus not a correction, and any investigations into the actual arrival times of signal-like events would require the unfolding of this bias from reconstructed times using the same parametrization with the opposite sign.

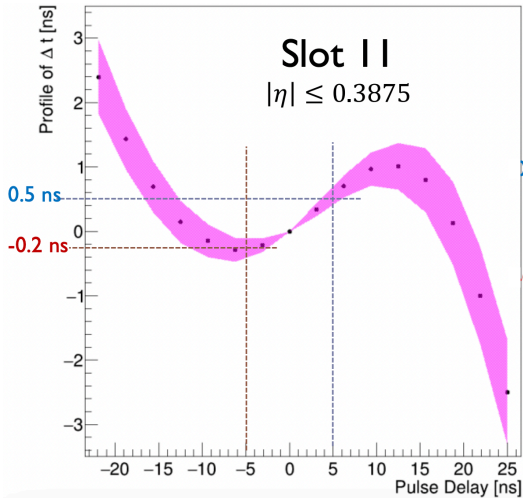
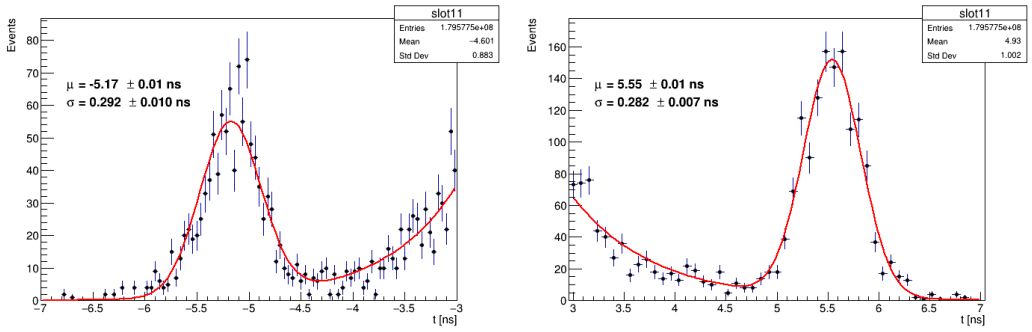


Figure A.16: Timing bias vs. pulse delay measured in special pulse studies carried out in 2018 for slot 11 ($|\eta| \leq 0.3875$). Positive values indicate that the measured time is later than the actual arrival time.



(a) (b)

Figure A.17: (a) −5 ns and (b) +5 ns satellite-satellite collision peaks, as measured in slot 11 ($|\eta| \leq 0.3875$) with electron times from the Run 2 $W \to e\nu$ timing calibration sample. The data is fit to an exponential distribution plus a Gaussian distribution with fitted mean μ and standard deviation σ.

Appendix B: Optimization Studies

B.1 Lepton Multiplicity Distinction

In the events passing our selections and those in the final state of the signal model considered in this analysis, the leptons present originate from the decays of $Z/W/t\bar{t}$ produced in association with the Higgs. The possible decay modes of all of these associated particles may produce differing numbers of leptons. However, the photons we consider in this analysis originate from Higgs boson decays and not from the associated $Z/W/t\bar{t}$ decays. The purpose of choosing of these associated production modes was to trigger on lepton(s) from their decays in order to reach manageable trigger rates due to the softness of the photons. But which particular channel is considered that results in differing numbers of leptons should not intuitively have any impact on the photons that results from a completely independent decay chain.

Therefore, it is not expected that making any distinction on the number of leptons in our events will have any significant impact on the sensitivity. This intuition was confirmed by a simplified significance study shown in Figure B.1. By comparing the total signifances for making the lepton number distinction and for combining them and only considering photon number, one can see that combining lepton channels, less than 10% of significance is lost across many different E_{cell} cuts.

In addition to the relatively small change in significance from dropping this lepton number distinction, combining lepton channels also helps to alleviate statistics issues in the two lepton channels, which are heavily suppressed (roughly by a factor of 20) compared to the corresponding one lepton channels. It also reduce the number of fits that must be handled properly in the statistical framework, greatly simplifying this procedure.

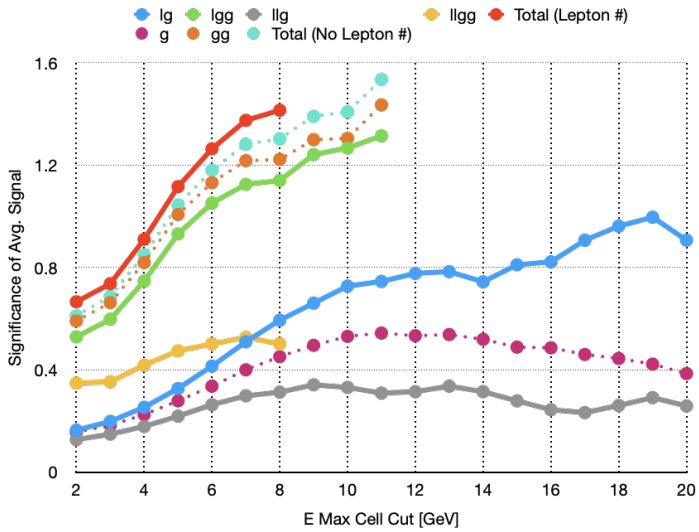


Figure B.1: Significances by channel and total significances for distinguishing by the number of leptons and for combining them vs. E_{cell} cut. The timing binning was [0,0.2,0.4,0.8,1,12] ns, and there was only one z bin: [0, 2000] mm. The background was estimated by reflecting the SR negative timing events to the positive timing side. The signal was an average of all available signal points with lifetime 2 and 10 ns.

Therefore, the decision to drop the lepton number distinction was made to simplify the analysis in these ways while maintaining high sensitivity.

B.2 Performance of Average Signal

The optimization procedure uses an average signal composed of all available signal points with lifetimes 2 and 10 ns added together and divided by the total number of signal points. In studies of separately optimizing 2 and 10 ns signal points, no significant differences were seen between the two lifetimes. Thus, in the signal averaging, they were combined indiscriminately with respect to lifetime.

By synthesizing such a potentially unphysical average signal that may, in principle, not model any of the individual signal points well, it must be verified that an optimization performed on such a signal does indeed provide good results for the individual points of which it is composed. This verification is performed by simply taking the results of the optimization with the average signal and applying them to each of the individual signal points to verify that they perform reasonably well.

In addition to verifying that the average signal optimization performs well, one would also want to check that the general behavior of the average signal mirrors that of the individual points as the optimizer explores the parameter space. Such a check was performed by holding the $E_{\mathrm{T}}^{\mathrm{miss}}$ cuts constant (CR: $E_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV, SR: $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV) and allowing the binning and E_{cell} cuts to float. This optimization procedure (see details in Section 7.3) was carried out on both the average signal and on each of the individual signal points independently. These optimizations were carried out with the fixed binning of $[0, 0.2, 0.4, 0.6, 0.8, 1, 12]$ ns in timing and $[0, 2000]$ mm in pointing as well as with one additional timing bin between 1 and 12 ns and with one additional pointing bin between 0 and 2000 mm. The results are shown in Figure B.2, which demonstrate that the E_{cell} peak in significance of the average signal generally matches the desired peak of the individual points when optimized separately for two different binning configurations. Thus, it was taken that the average signal was a reasonable proxy to study the general behavior of the individual signal points, vastly simplifying the optimization search space.

B.3 Separate Signal Grid Region Optimizations

Since the average signal across all points is dominated by higher-yield points, i.e. points with large mass splittings, it was found that the results especially favored signal points with already very high significances and did not sufficiently improve points with lower Δm's for which exclusion limits may not fall below the current limit on $BR_{H\to\mathrm{inv}}$. In order to focus more optimization resources on such points near the potential edge of exclusion, separate optimizations with averages across different groups of signal points were considered. For $f_\gamma = 0.5$, the full optimization procedure was performed separately with the following groupings of signal points (both 2 and 10 ns lifetimes together): $\Delta m \leq 5$ GeV, $\Delta m > 5$ GeV, $\Delta m \leq 10$ GeV, $\Delta m > 10$ GeV, $\Delta m \leq 20$ GeV, $\Delta m > 20$ GeV, $\Delta m \leq 30$ GeV, and $\Delta m > 30$ GeV. The validation plots showing the significances achieved at a branching ratio of 30% for each individual signal point for each grouping choice are shown in Figures B.4 and B.3.

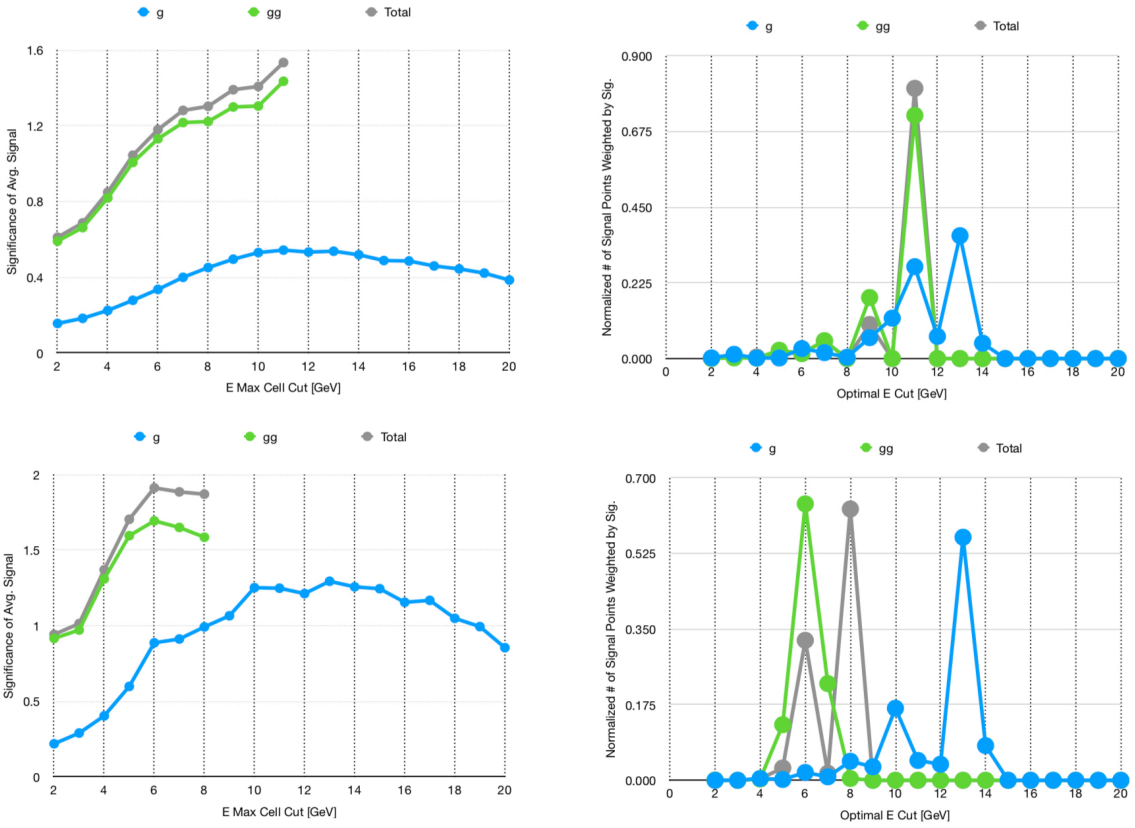
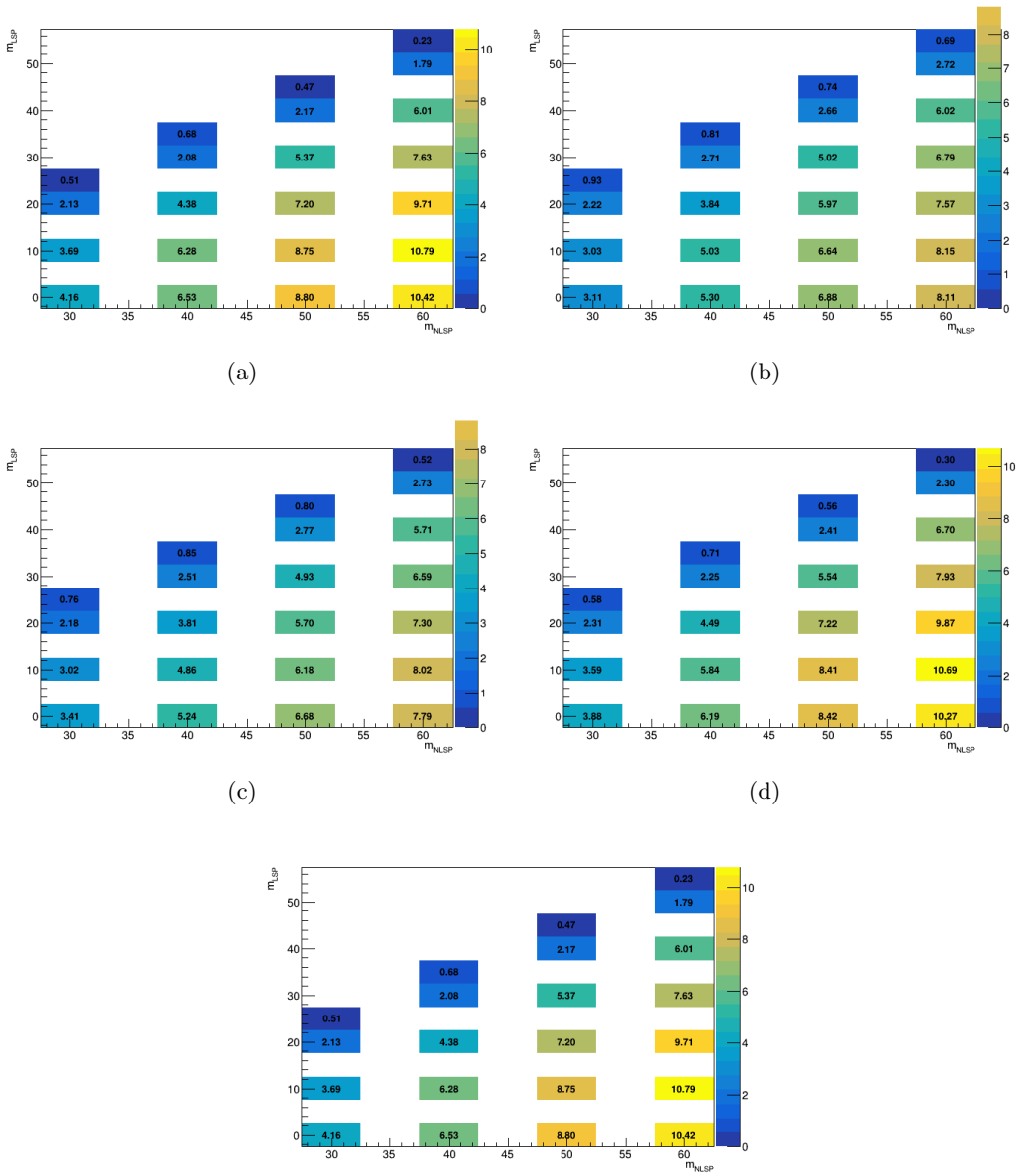


Figure B.2: At left, optimized significance vs. E_{cell} for the average signal with fixed binning $[0, 0.2, 0.4, 0.6, 0.8, 1, 12]$ ns in timing and $[0, 2000]$ mm in pointing (top left) and with one additional bin allowed in timing after 1 ns and in pointing (bottom left). At right, histograms of the optimal E_{cell} cut weighted by the significance of each of the 44 available signal points (2 and 10 ns) for the previously mentioned fixed binning (top right) and for the previously mentioned finer binning (bottom right).

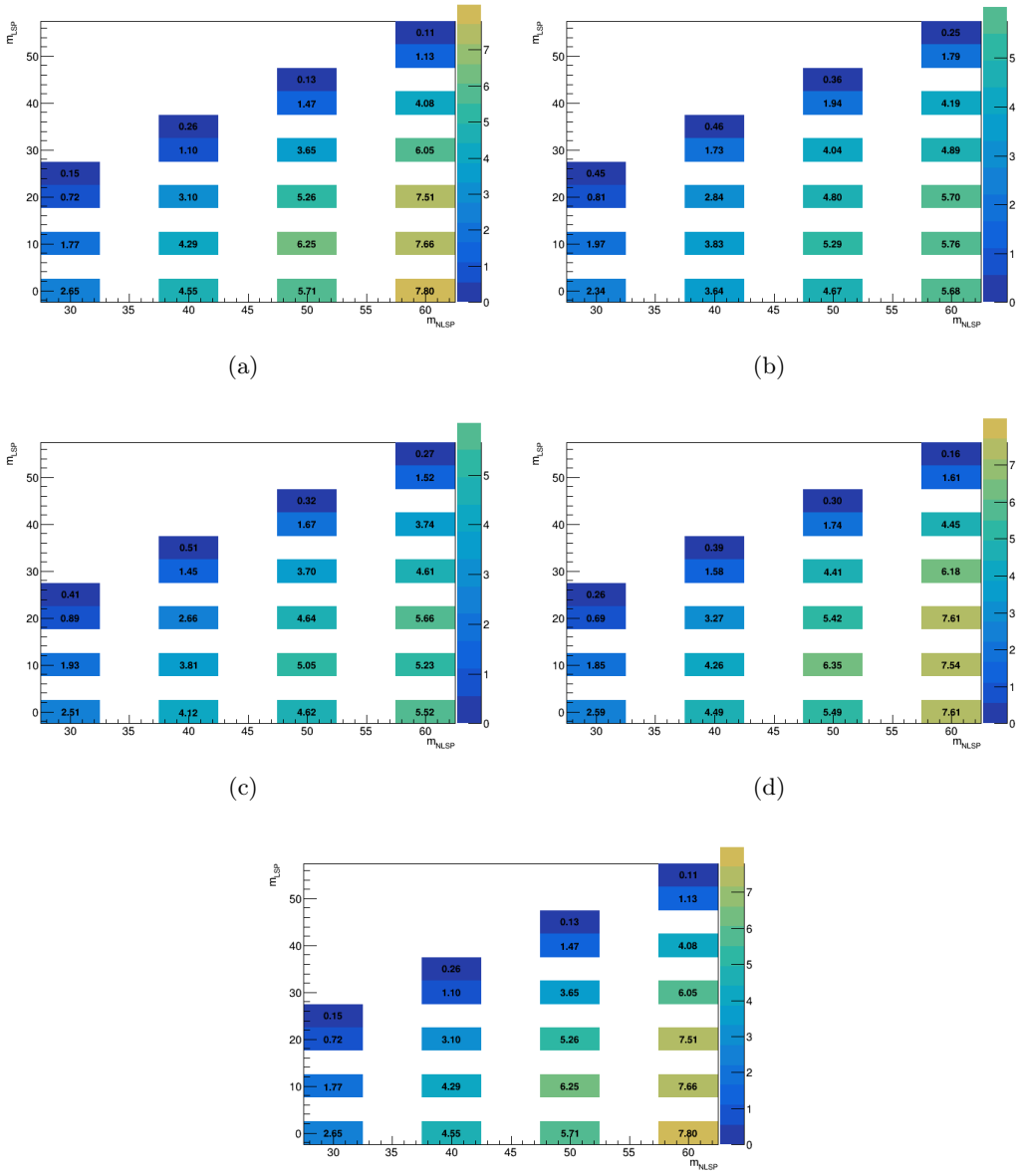


(a) (b)

(c) (d)

(e)

Figure B.3: Significances inflated for a branching ratio of 30% for 2 ns signal points using the optimized cuts and bins of the optimization procedure using a signal average of points with (a) all points, (b) $\Delta m < 5$ GeV, (c) $\Delta m < 10$ GeV, (d) $\Delta m < 20$ GeV, and (e) $\Delta m < 30$ GeV. The results with $\Delta m > 5$, 10, 20, and 30 GeV are identical to those with all points, i.e. (a).



(a) (b)

(c) (d)

(e)

Figure B.4: Significances inflated for a branching ratio of 30% for 10 ns signal points using the optimized cuts and bins of the optimization procedure using a signal average of points with (a) all points, (b) $\Delta m < 5$ GeV, (c) $\Delta m < 10$ GeV, (d) $\Delta m < 20$ GeV, and (e) $\Delta m < 30$ GeV. The results with $\Delta m > 5$, 10, 20, and 30 GeV are identical to those with all points, i.e. (a).

In general, it was found that the resulting cuts and binnings followed some general patterns. Firstly, all "greater than" optimizations were identical, indicating that the large

Δm points dominate anything below $\Delta m = 30$ GeV. These "greater than" optimizations were also identical to the optimization with the full signal grid. Secondly, lower Δm's were associated with lower E_{cell} cuts and higher upper VR $E_{\text{T}}^{\text{miss}}$ cuts. This is expected since lower Δm's have softer photons and higher $E_{\text{T}}^{\text{miss}}$.

By grouping each signal point into the optimization containing that point, we considered choosing two separate optimizations drawing a line in the mass plane at either $\Delta m = 5$, 10, 20, or 30 GeV. It was found that choosing $\Delta m = 10$ GeV achieved the best compromise, balancing losing significance at lower Δm with higher boundaries and losing significance at intermediate Δm with lower boundaries.

B.4 Calculation of Significance

The timing and pointing bins in which we calculate the significance vary widely in sensitivity, i.e. signal and background vary by orders of magnitude between low pointing, low timing bins and high pointing, high timing bins. Estimating significance with the heuristic $S/\sqrt{B}$ will not provide accurate results since this approximation only holds for $S \ll B$, and this is not true in high pointing and timing bins.

We begin by considering the general asymptotic formula for the exclusion significance Z_{ij} in pointing bins i, timing bins j, B_{ij} predicted background events, and S_{ij} signal events [98]:

$$Z_{ij} = \sqrt{-2\left(B_{ij}\log\left(1+\frac{S_{ij}}{B_{ij}}\right) - S_{ij}\right)} \tag{B.1}$$

We choose exclusion rather than discovery significance since exclusion significance folds in shape uncertainties of the signal model and thus allows us to account for uncertainty from our limited MC statistics. This formula has an additional assumption, namely that the uncertainties in the background and signal shapes are small. However, in the higher timing and pointing bins where statistics in the tails of both the background templates and our limited MC samples drop, this is not the case. To account for this uncertainty, we consider additional so-called shape terms corresponding to the statistical uncertainties of the background ($\sigma_{ij}^{B,\text{shape}} = \sqrt{B_{ij}}$) and signal ($\sigma_{ij}^{S,\text{shape}} = \sqrt{S_{ij}}$) in each bin. These shape terms are then added in quadrature to S_{ij}/Z_{ij} (modeling Z_{ij} as S_{ij}/σ_{ij}). We denote this

new, modified significance Z'_{ij}:

$$Z'_{ij} = \frac{S_{ij}}{\sqrt{\left(\frac{S_{ij}}{Z_{ij}}\right)^2 + \left(\sigma_{ij}^{B,\text{shape}}\right)^2 + \left(\sigma_{ij}^{S,\text{shape}}\right)^2}} \tag{B.2}$$

$$= \left(\frac{-1}{2\left(B_{ij}\log\left(1+\frac{S_{ij}}{B_{ij}}\right) - S_{ij}\right)} + \frac{B_{ij}}{S_{ij}^2} + \frac{1}{S_{ij}}\right)^{-1/2} \tag{B.3}$$

The final, total significance Z' across all timing and pointing bins is then simply:

$$Z' = \sqrt{\sum_{i,j} {Z'_{ij}}^2} \tag{B.4}$$

B.5 Evaluation of Optimizer Performance

To evaluate the performance of the optimized cuts and bins, an alternate set of cuts and bins is presented here. This set represents initial guesses made before the optimization procedure was carried out based on an understanding from previous non-pointing photon analyses from Run 1 [65, 66], the unique conditions of Run 2, and the target final state in the present analysis. These choices are summarized in Table B.1. This can be compared to the final, optimized cuts and bins in Table 7.3.

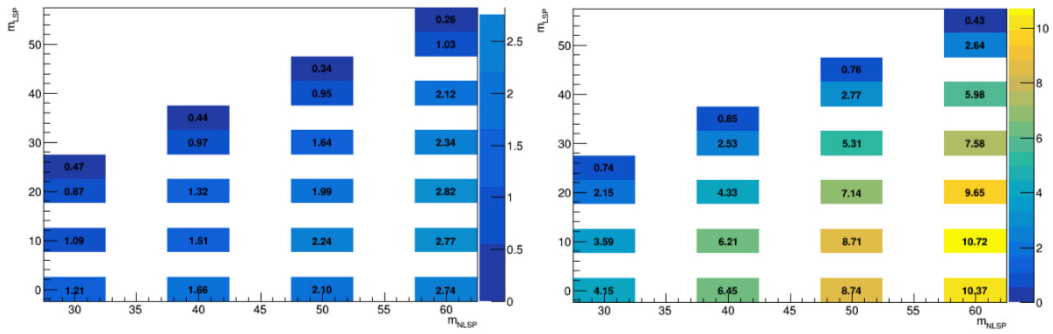
The low E_{cell} cut was designed to target the relatively soft final-state photons compared to the Run 1 analyses, which used a 5 GeV cut and had much harder photons. The pointing and timing binning was also based on the Run 1 strategy but adding additional bins based

Parameter	Cuts/Bins
E_{cell} Cut [GeV]	2
CR $E_{\text{T}}^{\text{miss}}$ [GeV]	<30
VR $E_{\text{T}}^{\text{miss}}$ [GeV]	30-50
SR $E_{\text{T}}^{\text{miss}}$ [GeV]	>50
$\lvert z\rvert$ bins [mm]	[0,50,100,150,200,250,500,2000]
t bins [ns]	[0,0.2,0.4,0.6,0.8,1.0,1.5,2.5,4.5,6.5,12.0]

Table B.1: Informed initial set of cuts and bins established before the optimization procedure used as a baseline to evaluate its performance, as shown in Figure B.5.

on the increased statistics available in Run 2 and extending the timing out to 12 ns from 4 ns[1].

A comparison of the sensitivities between the two choices are shown in Figure B.5, where gains in sensitivity post-optimization by factors of 2-4 are observed for many signal points.



(a) (b)

Figure B.5: Sensitivities shown for BR($H \to \chi_1^0\chi_1^0$) = 30% for 2 ns signal points using an equal mixture ($\alpha = 0.5$) of photon- and fake-photon-enhanced templates. (a) shows the results for the initial, informed guesses for cuts and bins summarized in Table B.1. (b) shows the results for the full optimization procedure discussed in Section 7.3.1.3, whose cuts and bins are summarized in Table 7.3.

B.6 Dependence on Template Mixing Fraction

Studies of the dependence of the optimization procedure on α, the fraction of the real-photon-enhanced template used in the background prediction, are summarized here. The full optimization, as described in Section 7.3 was carried out independently with $\alpha = 0$, 0.25, 0.5, 0.75, and 1. The results are summarized in Table B.2. As can be seen, the results are relatively insensitive to α, with many having identical results (bolded) for many of the optimization parameters including not only the cuts and the number of bins, but the placements of the bin edges as well. Even in the extreme cases of $\alpha = 0$ and 1, where only the fake- or real-photon-enhanced templates are considered, the results are generally quite similar, with no large, qualitative differences observed. It should be noted that the differences that are observed–one or two extra bins and small shifts of their edges—are

[1]the 4.5-6.5 ns bin was chosen specifically to target the satellite-satellite collision events.

not considered to have significant effects on the significance. This fact was verified in the arrival at the final analysis binning in which the optimization results were harmonized across categories and bins, and no significant impact was seen from these small adjustments.

		α (Fraction of Real-Photon-Enriched Template)				
		0	0.25	0.5	0.75	1
	Max. Significance	1.04868936044	1.36420610387	1.38026779764	1.46238929765	1.93569925559
	Energy Cut	8	**10**	**10**	**10**	9
	VR Low $E_{\mathrm{T}}^{\mathrm{miss}}$	40	**30**	**30**	**30**	**30**
	VR High $E_{\mathrm{T}}^{\mathrm{miss}}$	**50**	**50**	**50**	**50**	**50**
γ	Significance	0.637439179791	0.873354904171	0.910057533562	0.999603772476	1.32931468176
	# t bins	9	**7**	**7**	**7**	6
	t bins	[1.5, 2.0, 2.5, 12.0]	**[1.5, 12.0]**	**[1.5, 12.0]**	**[1.5, 12.0]**	[12.0]
	# z bins	3	**5**	**5**	**5**	6
	z bins	[0, 50, 150, 2000]	**[0, 50, 100, 200, 250, 2000]**	**[0, 50, 100, 200, 250, 2000]**	**[0, 50, 100, 200, 250, 2000]**	[0, 100, 150, 250, 350, 800, 2000]
$\gamma\gamma$	Significance	0.832718838394	1.04800262938	1.03775454283	1.06741499426	1.40707288194
	# t bins	**6**	**6**	**6**	**6**	**6**
	t bins	**[12.0]**	**[12.0]**	**[12.0]**	**[12.0]**	**[12.0]**
	# z bins	4	**6**	**6**	**6**	**6**
	z bins	[0, 50, 100, 150, 2000]	**[0, 50, 100, 150, 200, 350, 2000]**	**[0, 50, 100, 150, 200, 350, 2000]**	**[0, 50, 100, 150, 200, 350, 2000]**	**[0, 50, 100, 150, 200, 350, 2000]**

Table B.2: Optimal parameter values for the optimization procedure for a set of different values of α, the template mixing fraction. Timing bin edges listed are those beyond the fixed low-timing bins [0.0, 0.2, 0.4, 0.6, 0.8, 1.0] ns. Results are also separated by the optimal cuts and the results by the one (γ) and two ($\gamma\gamma$) photon channels. Results that are identical among different α values are grouped and shown in bold.

Appendix C: Template Studies

C.1 Template Shape Comparison

Figures C.1 to C.10 show the template timing shape comparisons for different pointing and one-photon channel and two-photon channel for signal region with SR $E_{\mathrm{T}}^{\mathrm{miss}}$ >50 GeV, E_{cell} >10 GeV, and Figures C.11 to C.20 show similar comparison for signal region with SR $E_{\mathrm{T}}^{\mathrm{miss}}$ >80 GeV, E_{cell} >7 GeV.

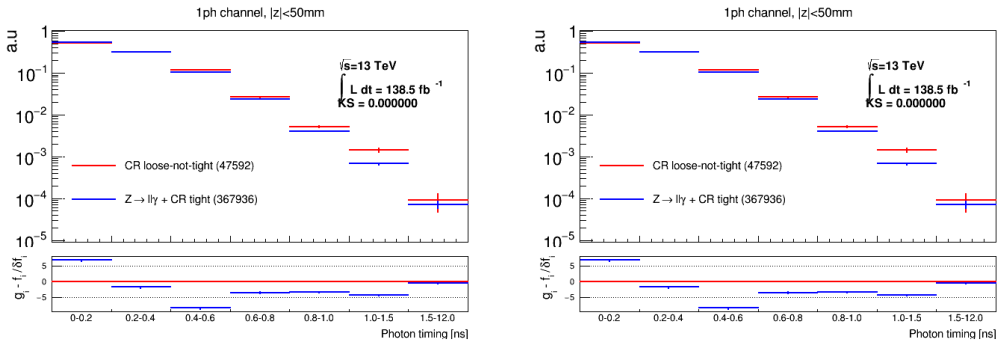


Figure C.1: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $|z| < 50$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

C.2 Non-Closure Uncertainties

Figure C.21 shows the fitted background PDF and data in the $\mathrm{SR}_{t<0}$ region. In each bin, the residual difference between the data point and the fitted PDF reflects the level of the intrinsic discrepancy between data distributions and timing templates.

Tables C.1 and C.2 summarize the non-closure uncertainties for the high mass splitting

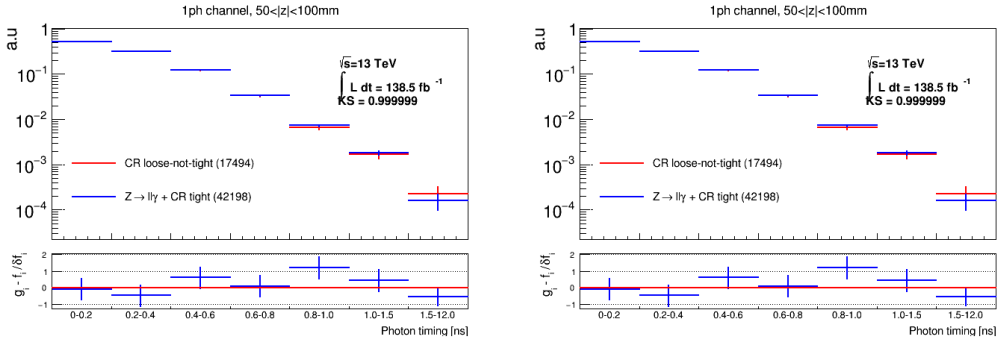


Figure C.2: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $50 < |z| < 100$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

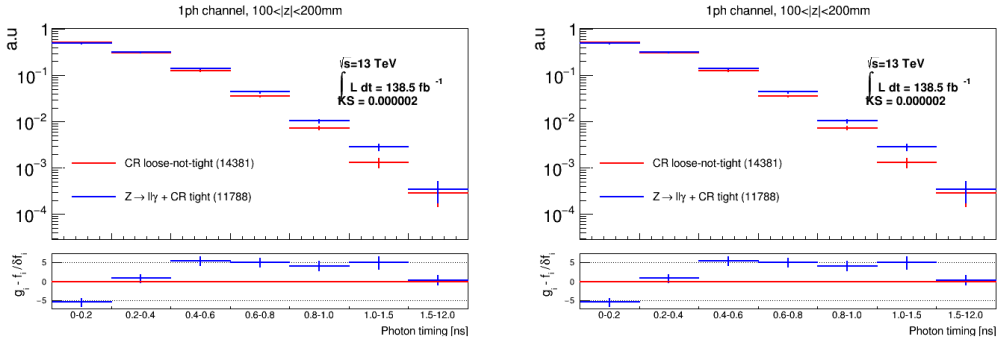


Figure C.3: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $100 < |z| < 200$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

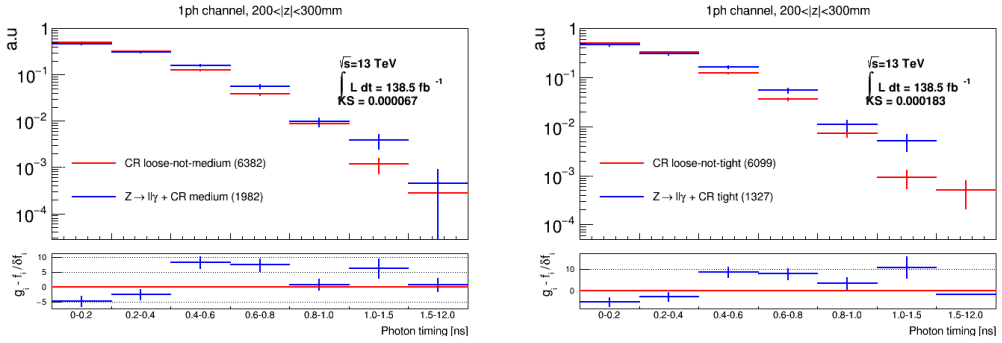


Figure C.4: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $200 < |z| < 300$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

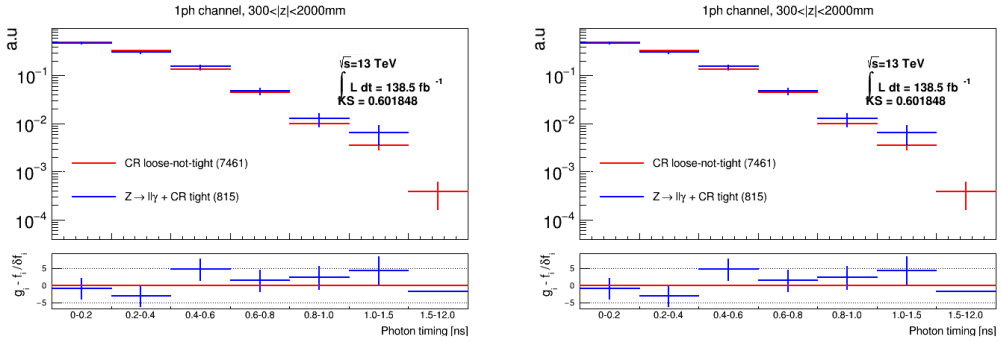


Figure C.5: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $300 < |z| < 2000$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

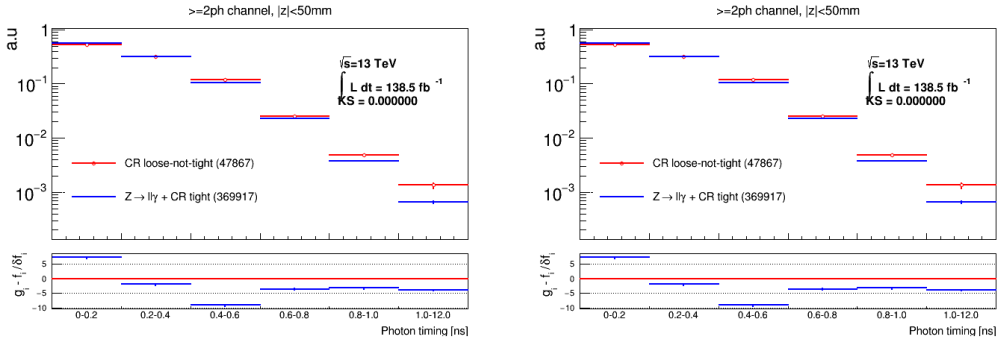


Figure C.6: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $|z| < 50$ mm in the two photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

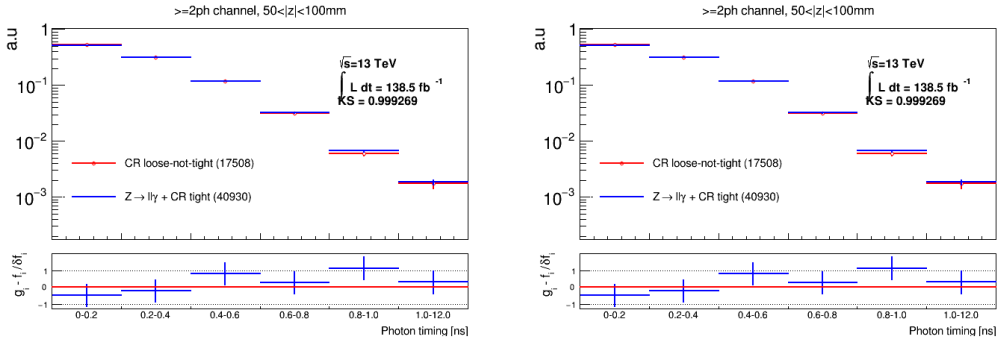


Figure C.7: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $50 < |z| < 100$ mm in the two photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

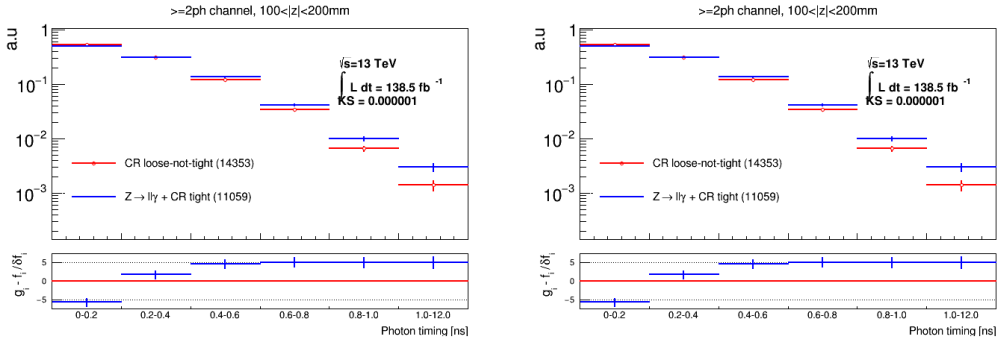


Figure C.8: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $100 < |z| < 200$ mm in the two photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

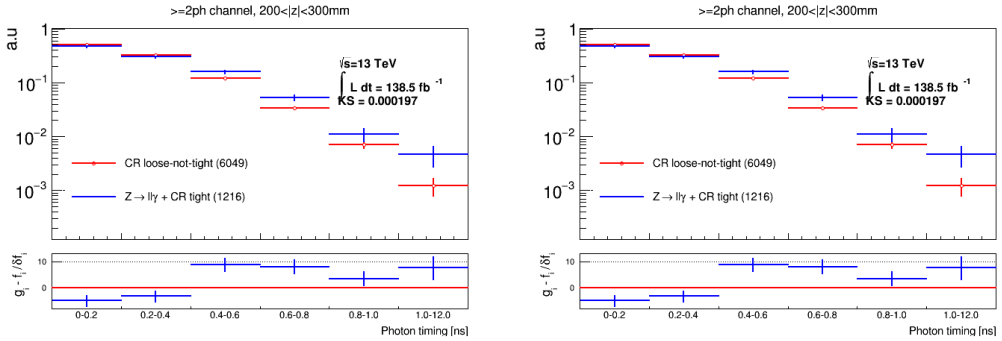


Figure C.9: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $200 < |z| < 300$ mm in the two photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

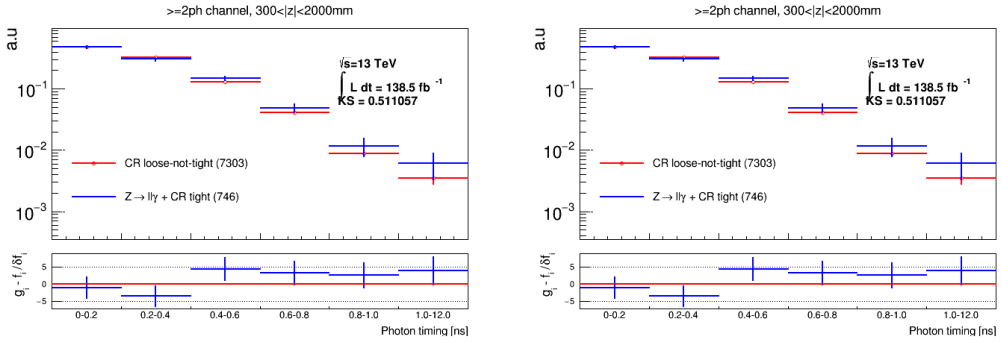


Figure C.10: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $300 < |z| < 2000$ mm in $\geq$ 2ph channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV selection.

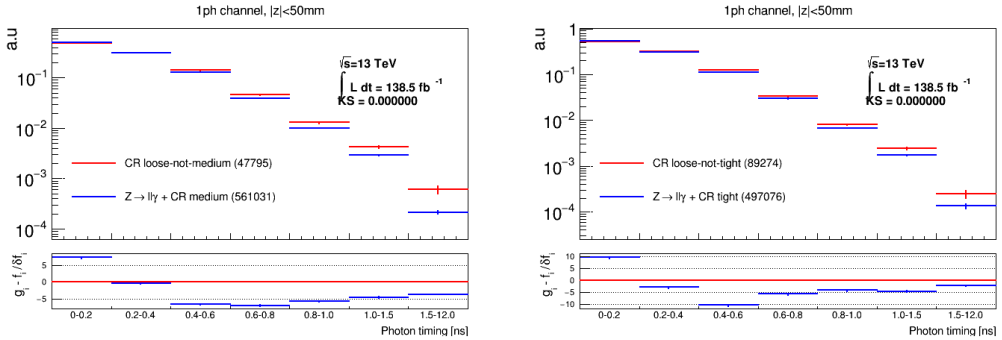


Figure C.11: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $|z| < 50$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

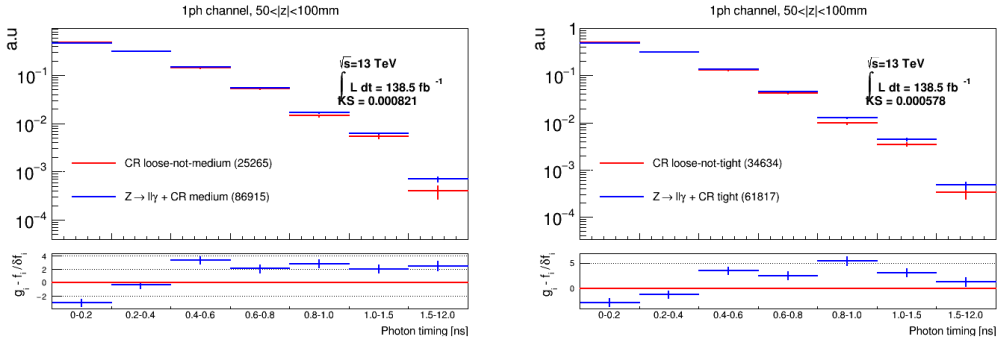


Figure C.12: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $50 < |z| < 100$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

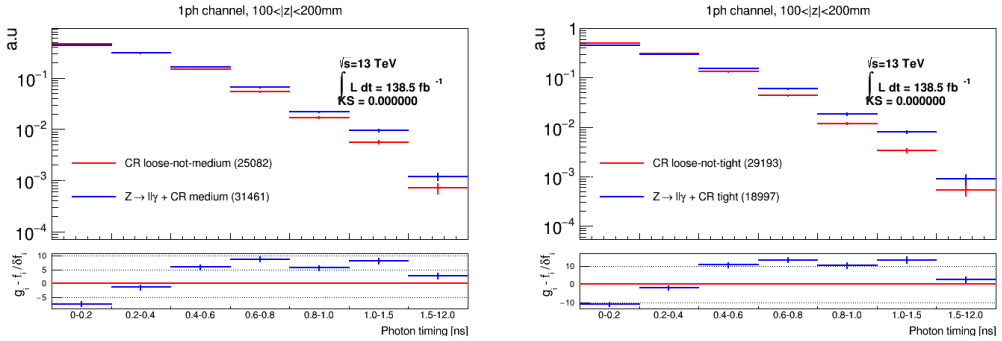


Figure C.13: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $100 < |z| < 200$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

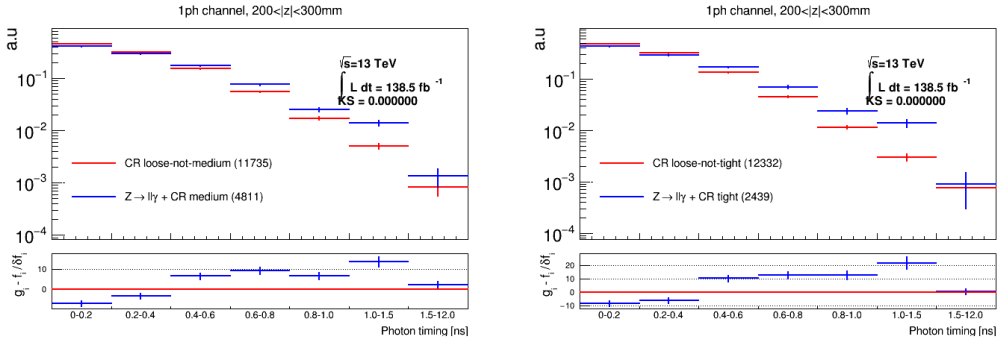


Figure C.14: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $200 < |z| < 300$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}}$ > 80 GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

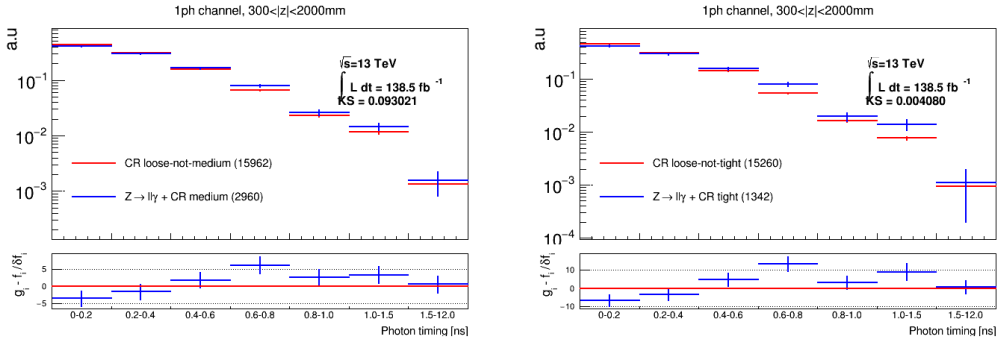


Figure C.15: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $300 < |z| < 2000$ mm in the one photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}}$ > 80 GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

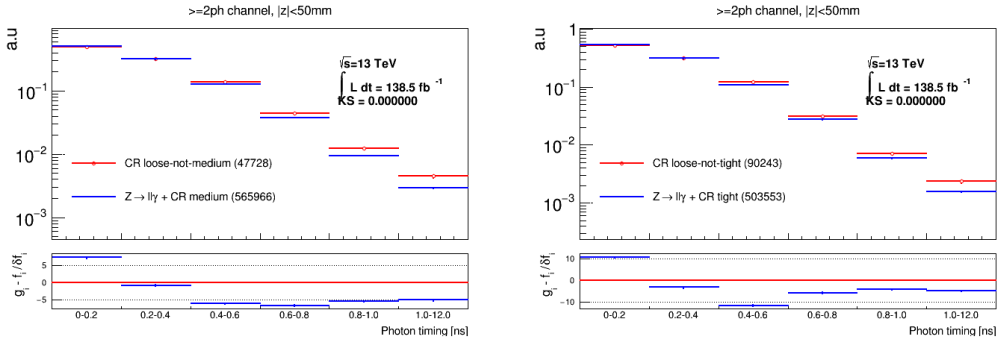


Figure C.16: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $|z| < 50$ mm in the two photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

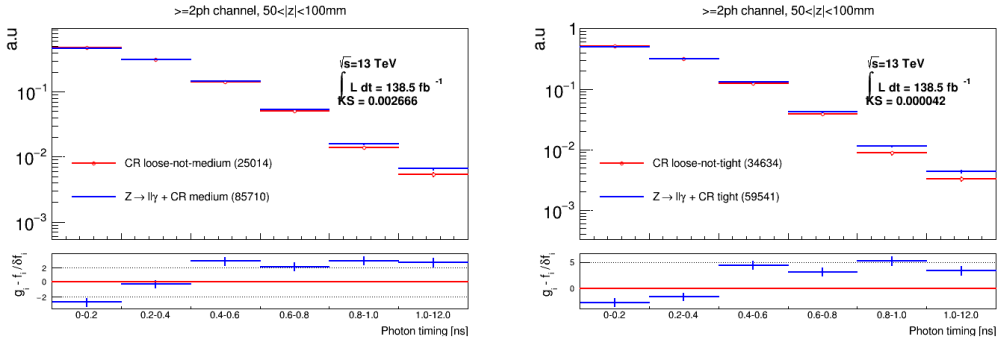


Figure C.17: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $50 < |z| < 100$ mm in the two photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

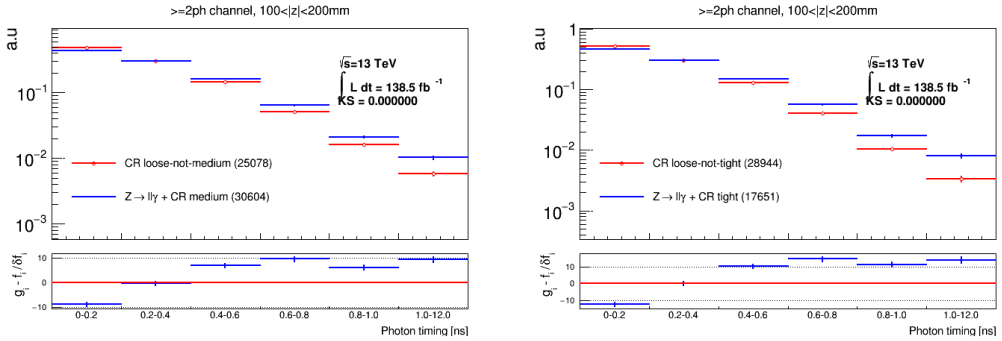


Figure C.18: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $100 < |z| < 200$ mm in the two photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

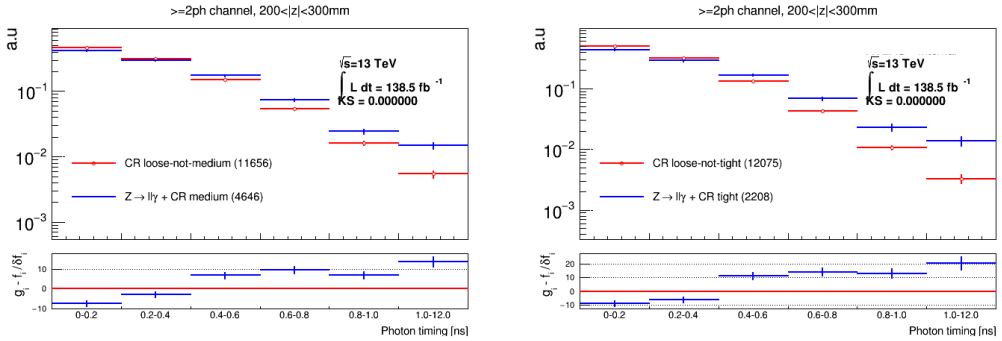


Figure C.19: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $200 < |z| < 300$ mm in the two photon channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

and low mass splitting analyses, respectively.

C.3 $SR_{t<0}$ Fitted Purity Values

The fitted purity values, reflecting the fraction taken from the genuine-photon-enriched timing template, for the $SR_{t<0}$ background-only fit are shown in Table C.3.

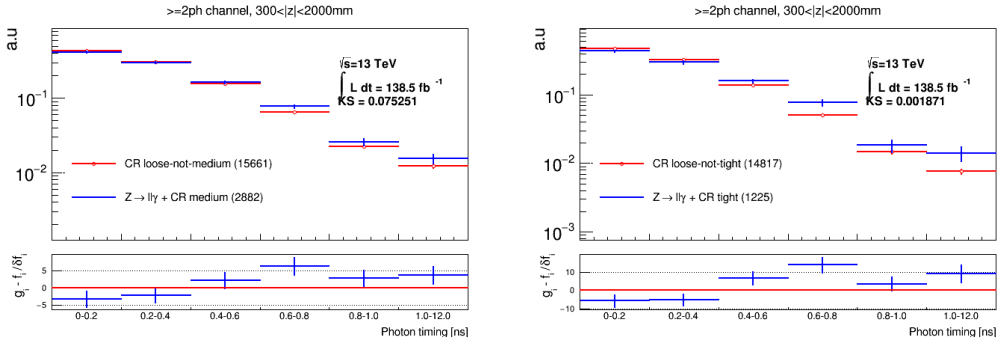


Figure C.20: Comparison between medium and loose-not-medium templates are available on the left side and comparisons between tight and loose-not-tight templates are present on the right side for $300 < |z| < 2000$ mm in $\geq$ 2ph channel for signals with SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV selection.

One-Photon Channel

	Timing Bin [ns]						
Pointing Bin [mm]	0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1	1-1.5	1.5-12
0-50	0	0.00011	0.0047	0.017	0	0.074	0.86
50-100	0	0	0	0	0.12	0	0
100-200	0	0	0	0	0	0	0
250-300	0	0	0	0	0.087	0	0
300-2000	0	0	0	0.0020	0	0	0

Two-Photon Channel

	Timing Bin [ns]						
Pointing Bin [mm]	0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1	1-1.5	1.5-12
0-50	0	0	0	0	0	0	—
50-100	0	0.0040	0.25	0	0	0	—
100-200	0	0	0	0	0	0	—
200-300	0	0.080	0	0	0	0	—
300-2000	0	0	0	0	0	0	—

Table C.1: The non-closure relative uncertainties are summarized here for mixed templates and the high mass splitting selection ($\mathrm{SR}_{t<0}$ $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV).

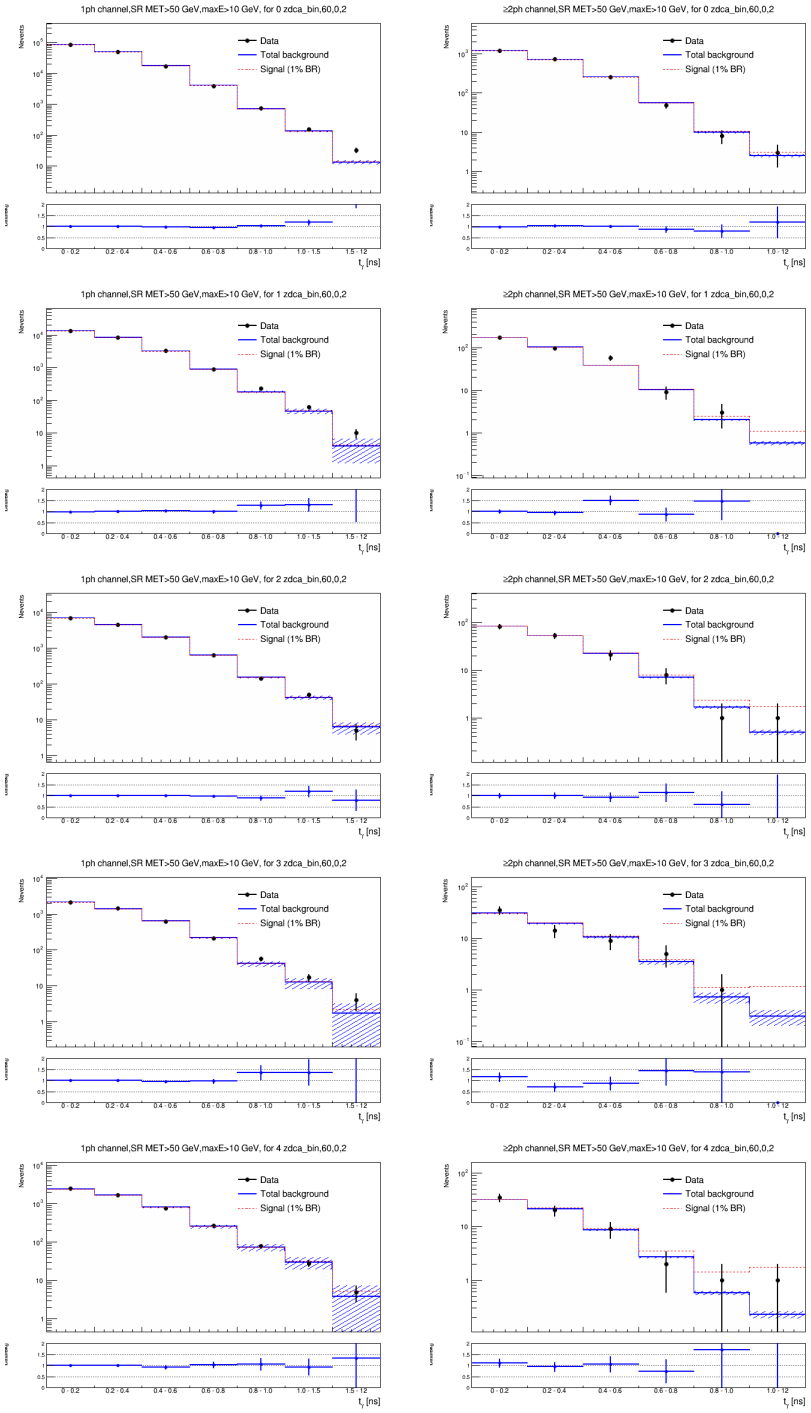


Figure C.21: Plots comparing $\mathrm{SR}_{t<0}$ data (black) against background fit with BR=0 (blue) without non-closure uncertainties using mixed templates for the high mass splitting selection (SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV) in the one-photon channel on the left and the two-photon channel on the right. Ratio of data to each of the background estimation is depicted in the bottom panes.

One-Photon Channel

	Timing Bin [ns]						
Pointing Bin [mm]	0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1	1-1.5	1.5-12
0-50	0.042	0	0.096	0.20	0.32	0.27	0.99
50-100	0.0040	0	0	0.075	0.059	0.022	0
100-200	0	0	0	0	0.058	0.035	0
200-300	0.0043	0	0.0034	0	0	0	0
300-2000	0	0	0	0	0.090	0.17	0

Two-Photon Channel

	Timing Bin [ns]						
Pointing Bin [mm]	0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1	1-1.5	1.5-12
0-50	0	0	0	0.044	0	0.31	—
50-100	0	0.044	0	0	0	0.19	—
100-200	0	0	0	0	0	0.19	—
200-300	0	0	0	0	0	0.19	—
300-2000	0	0	0	0	0	0.19	—

Table C.2: The non-closure relative uncertainties are summarized here for mixed templates and the low mass splitting selection (SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV).

Analysis	Channel	Pointing Category [mm]	Purity
High Δm	γ	0–50	1.000 ± 0.048
		50–100	0.5
		100–200	0.998 ± 0.036
		200–300	0.593 ± 0.091
		300–2000	0.988 ± 0.139
	$\gamma\gamma$	0–50	0.310 ± 0.318
		50–100	0.5
		100–200	1.000 ± 0.817
		200–300	1.000 ± 0.734
		300–2000	0.000 ± 0.824
Low Δm	γ	0–50	1.000 ± 0.002
		50–100	0.000 ± 0.048
		100–200	0.349 ± 0.127
		200–300	0.180 ± 0.130
		300–2000	0.000 ± 0.205
	$\gamma\gamma$	0–50	0.257 ± 0.621
		50–100	0.705 ± 0.602
		100–200	0.000 ± 0.738
		200–300	1.000 ± 0.986
		300–2000	0.000 ± 0.752

Table C.3: The measured purity values in $\mathrm{SR}_{t<0}$ pseudo-data. In the 50–100 mm pointing categories for the high mass splitting analysis, the purity parameter is fixed to 0.5.

Appendix D: $SR_{t<0}$ Statistics

D.1 Data Yields

The yields for the $SR_{t<0}$ region are shown in Figure D.1, and the associated statistical errors in Figure D.2.

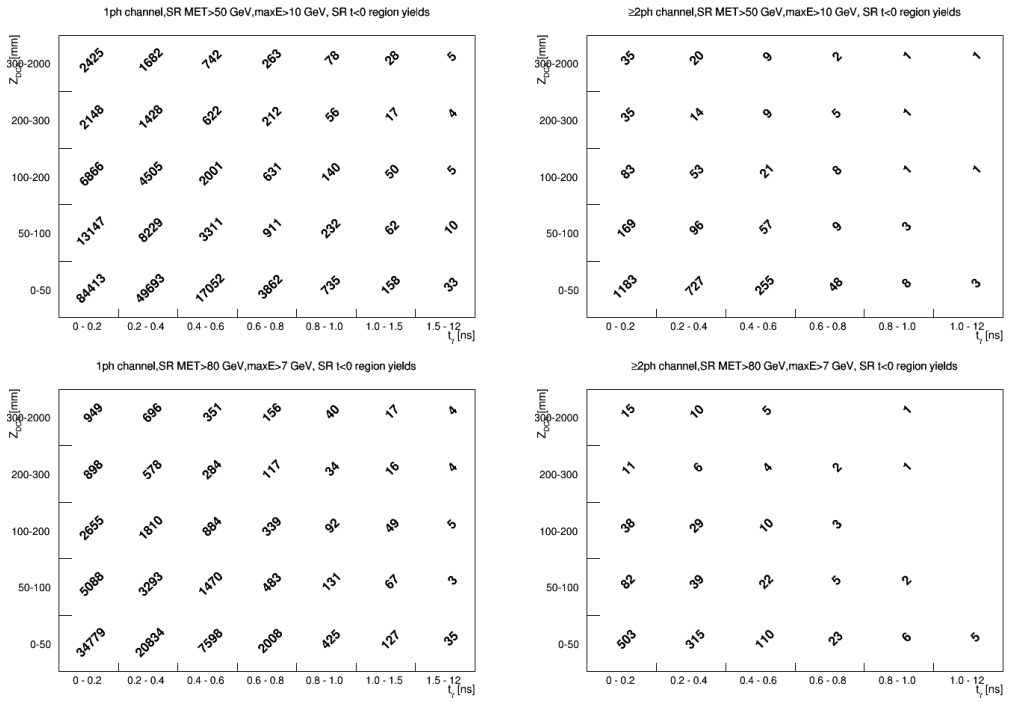


Figure D.1: Yields in data $SR_{t<0}$ region are presented here. Plots in first and second rows correspond to the high and low mass splitting selections, respectively. The left and right columns correspond to the one- and two-photon channels, respectively. These plots list the yield in the $SR_{t<0}$ reflected to the positive timing side to allow easy comparison to signal.

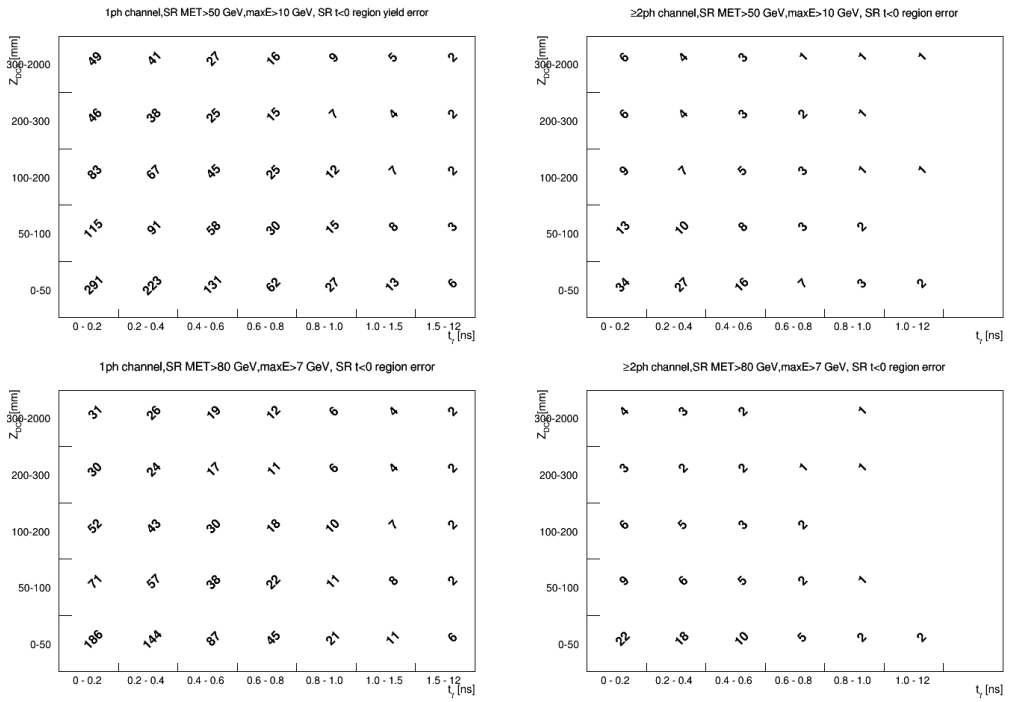


Figure D.2: Statistical uncertainties for data $SR_{t<0}$ region are presented here. Plots in first and second rows correspond to the high and low mass splitting selections, respectively. The left and right columns correspond to the one- and two-photon channels, respectively. These plots list the uncertainty in the $SR_{t<0}$ reflected to the positive timing side to allow easy comparison to signal.

D.2 Signal Contamination

The signal point with $m_{NLSP} = 60$ GeV, $m_{LSP} = 0.5$ GeV and $\tau = 2$ ns has the highest yields of all signal points. Figure D.3 presents the signal yield in $\mathrm{SR}_{t<0}$region for this signal point using a BR of 10% (roughly the expected sensitivity for many signal points). Note that the actual expected sensitivity for this particular signal point is much smaller at around 0.8%, but it is inflated here to demonstrate that signal contamination is negligible even in this extreme case.

Comparing the signal yield in $\mathrm{SR}_{t<0}$ to the data statistical uncertainty in Figure D.2, it is clear that the signal yield is much smaller than the statistical uncertainty of the corresponding data yield. This justifies the use the $\mathrm{SR}_{t<0}$ region as a validation region to develop our background estimation technique.

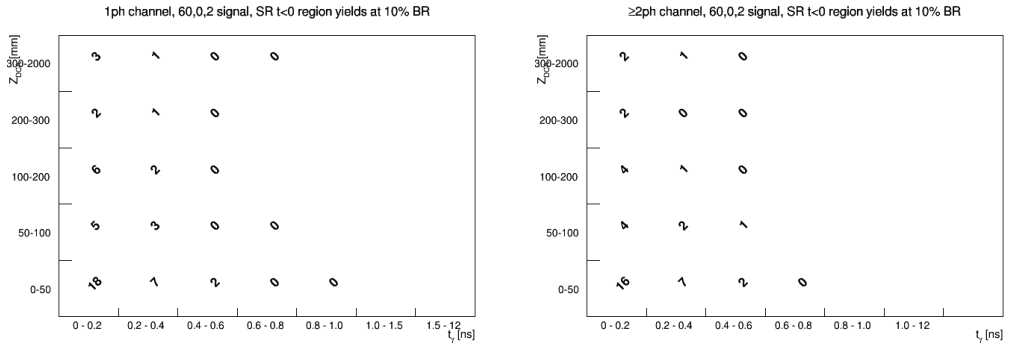


Figure D.3: Yields for the signal point with $m_{NLSP} = 60$ GeV, $m_{LSP} = 0.5$ GeV and $\tau = 2$ ns in the $\mathrm{SR}_{t<0}$ are presented here in each timing and pointing bin using $\mathrm{BR}(H \to \chi_1^0\chi_1^0) = 10\%$.

Appendix E: $SR_{t>0}$ Results

E.1 KS Test Results for Template Distinguishability

A summary of the KS test values for the unblinded $SR_{t>0}$ data is shown in Table E.1. As described in Section 7.5.1, the analysis strategy dictates that, for pointing categories in which the KS values are above 0.9, the purity nuisance parameter is to be fixed to 0.5. Therefore, purities are to be fixed in two pointing categories for the high mass splitting analysis, while the low mass splitting analysis requires no purity fixing. Identical purity fixing was prescribed during validation on $SR_{t<0}$ data based on the same requirements.

E.2 $SR_{t>0}$ Fitted Purity Values

The fitted purity values, reflecting the fraction taken from the genuine-photon-enriched template, for the $SR_{t>0}$ background-only fit are shown in Table E.2.

		$\Delta m \leq 10$ GeV		$\Delta m > 10$ GeV	
Channel	$\|z\|$ Bin [mm]	Medium	Tight	Medium	Tight
γ	$\|z\| < 50$ mm	0.000	0.000	0.000	0.000
	$50 < \|z\| < 100$	**1.000**	**1.000**	0.000	0.001
	$100 < \|z\| < 200$	0.000	0.000	0.000	0.000
	$200 < \|z\| < 300$	0.000	0.000	0.000	0.000
	$300 < \|z\| < 2000$	0.499	0.649	0.000	0.001
$\gamma\gamma$	$\|z\|<50$	0.000	0.000	0.000	0.000
	$50 < \|z\| < 100$	0.212	**0.998**	0.000	0.000
	$100 < \|z\| < 200$	0.000	0.000	0.000	0.000
	$200 < \|z\| < 300$	0.000	0.001	0.000	0.000
	$300 < \|z\| < 2000$	0.576	0.612	0.000	0.004

Table E.1: $SR_{t>0}$ KS test values for genuine-enriched and fake-enriched template comparisons for medium and tight choices for high and low mass splitting selections. $\Delta m \leq 10$ GeV selections include SR $E_{\mathrm{T}}^{\mathrm{miss}} > 50$ GeV and $E_{\mathrm{cell}} > 10$ GeV. $\Delta m > 10$ GeV selections include SR $E_{\mathrm{T}}^{\mathrm{miss}} > 80$ GeV and $E_{\mathrm{cell}} > 7$ GeV. KS values reaching the threshold of 0.9 for purity fixing are in bold.

Analysis	Channel	Pointing Category [mm]	Purity
High Δm	γ	0–50	0.80 ± 0.07
		50–100	0.5
		100–200	0.85 ± 0.09
		200–300	0.49 ± 0.12
		300–2000	1.00 ± 0.05
	$\gamma\gamma$	0–50	0.62 ± 0.34
		50–100	0.5
		100–200	0.00 ± 0.51
		200–300	1.00 ± 0.78
		300–2000	0.00 ± 0.88
Low Δm	γ	0–50	0.72 ± 0.09
		50–100	1.00 ± 0.63
		100–200	0.75 ± 0.10
		200–300	0.44 ± 0.11
		300–2000	1.00 ± 0.20
	$\gamma\gamma$	0–50	0.97 ± 0.98
		50–100	1.00 ± 0.70
		100–200	0.76 ± 0.87
		200–300	0.47 ± 0.73
		300–2000	1.00 ± 0.81

Table E.2: The measured purity values in $SR_{t>0}$ data. In the 50–100 mm pointing categories for the high mass splitting analysis, the purity parameter is fixed to 0.5.