Monte Carlo Generation Involving Searches for Diphoton Resonances in Association with $\tau^+\tau^-$ or b-jets at the Electroweak Scale in the ATLAS Detector at the LHC

Njokweni Mbuyiswa[1,2], **Kutlwano Makgetha**[1,2], **Vuyolwethu Kakancu**[1,2], **Paballo Ndhlovu**[1,2], **Kgothatso Ntumbe**[1,2], **Reda Mekouar**[3], **Phuti Rapheeha**[1,2,4], **Mukesh Kumar**[1], **Rachid Mazini**[1], **Bruce Mellado**[1,2]

[1]School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand, Johannesburg, Wits 2050, South Africa.
[2]iThemba LABS, National Research Foundation, PO Box 722, Somerset West 7129, South Africa.
[3]Institute of High Energy Physics, Chinese Academy of Sciences, 19B Yuquan Road, Shijingshan District, Beijing, China
[4]School of Electrical Engineering, Tshwane University of Technology, Staatsartillerie Road, Pretoria West 0001, South Africa

E-mail: 2314612@students.wits.ac.za

Abstract. A Monte Carlo generation campaign has been submitted under the HBSM subgroup to study a simplified model featuring the resonant production of scalar particles at the electroweak scale. The process under consideration is $gg \to H \to SS'$, where the scalar S (with a mass of approximately 150 GeV) decays to a diphoton final state ($\gamma\gamma$), and the accompanying scalar S' (with a mass of approximately 95 GeV) decays to either $b\bar{b}$ or $\tau^+\tau^-$. The motivation for these studies is rooted in the persistent multi-lepton anomalies observed in various channels at the LHC, as highlighted in combined searches for scalar resonances by ATLAS and CMS and further explored in the context of electroweak-scale scalar states decaying to photons, leptons, or b-jets [arXiv:2109.02650, arXiv:2306.17209, arXiv:2503.16245]. The MC production is designed for Run 3 conditions at a center-of-mass energy of $\sqrt{s} = 13.6$ TeV and will facilitate detailed kinematic studies and optimization of selection strategies in these channels.

1 Introduction

Recent anomalies observed in multi-lepton and diphoton final states at the LHC have sparked interest in extended Higgs sectors, such as the Two-Higgs-Doublet Model with a singlet scalar (2HDM+S) [1]. These anomalies, including excesses observed around 152 GeV in the diphoton channel [2, 3, 4], and in multi-lepton events, [5, 6, 7, 8, 9] suggest the possible existence of new scalar particles. In this context, a search is conducted for a heavy scalar H produced via gluon-gluon fusion ($gg \to H$) that decays into two lighter scalars: $H \to SS'$, where $S \to \gamma\gamma$ and S' decays into other Standard Model (SM) particles, as illustrated in Figure 1. The decay process that are currently being requested are for $S' \to \tau^+\tau^-$ or $b\bar{b}$. This signature provides a clean experimental handle due to the presence of two isolated photons and either a pair of taus or b-jets. This paper focuses on the Monte Carlo (MC) request and generation of simulated signal events for this Beyond Standard Model (BSM) process. Generating accurate MC samples is a critical prerequisite for any physics analysis at the Large Hadron Collider

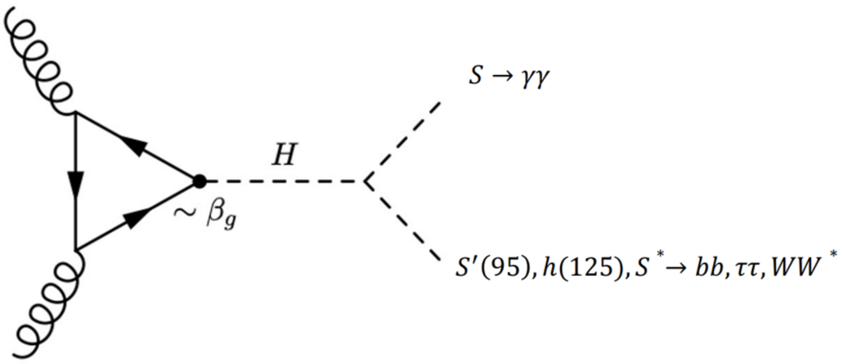

Figure 1: Feynman diagram illustrating a hypothetical heavy scalar boson H produced via gluon-gluon fusion (ggF), with subsequent decays into lighter scalar states. One branch decays into a diphoton final state ($S \to \gamma\gamma$), while the other decays into a lighter scalar or Higgs-like particle ($S'(95), h(125), S^*$) which further decays into fermions or vector bosons ($bb, \tau\tau, WW^*$), providing signatures such as b-jets, taus, leptons, and jets. This process is sensitive to new physics via the gluonic coupling $\sim \beta_g$.

(LHC). These samples allow us to model the expected signal, optimize event selection criteria, estimate detection efficiencies, and compare with observed data. The workflow follows the official ATLAS MC production pipeline, which includes event generation, detector simulation, digitization, reconstruction, and derivation of analysis-ready datasets. A key part of this process is the formal MC request submission via **JIRA**, ATLAS's project management tool. Before submission, small-scale signal samples are generated and validated to ensure the physics model is correctly implemented. A detailed request including benchmark points, cross sections, and truth-level filters is then prepared. This request was reviewed by the ATLAS MC production team, who assigned official dataset identifiers (DSIDs). Once approved, large-scale production was initiated across different pile-up conditions (mc23a/d/e) and simulation types (FullSim and AF3). The final datasets are distributed on the Worldwide LHC Computing Grid (WLCG) for use in physics analyses.

1.1 The ATLAS Detector

The ATLAS detector [10] is a multipurpose particle detector located at the LHC at CERN. It is designed to detect and measure the particles produced in high-energy proton–proton collisions at center-of-mass energies of up to 13.6 TeV. The detector has a cylindrical geometry with forward–backward symmetry and consists of several layered subdetectors that surround the interaction point. Starting from the innermost layer, the Inner Detector tracks charged particles using silicon pixel and strip detectors, as well as a transition radiation tracker. Surrounding the Inner Detector are the calorimeters, which measure the energy of particles. The electromagnetic calorimeter is made using liquid argon as a sampling medium, which is required for precise measurements of electron or photon energy. The hadronic calorimeter, on the other hand, uses scintillator tiles in some sections and liquid argon in other sections, and has the ability to measure the energy of jets produced by quarks and gluons. Outside the calorimeters is the muon spectrometer, which will identify and track all muons, since according to the electromagnetic spectrum muons are the only charged particles likely to penetrate into all of the inner detector layers. It uses large toroidal magnetic fields and multiple layers of precision tracking chambers. The entire detector is supported by a sophisticated trigger and data acquisition system that selects interesting events in real time.

1.2 Monte Carlo Simulation Procedure

The MC simulation process follows a well-defined series of steps that mimic the production, propagation, and detection of particles at the LHC. The first step is **event generation**, where the hard-scattering process is modeled using MG5 [11], which calculates the matrix element for $gg \to H \to SS'$. Next, **parton showering and hadronization** are handled by Pythia8 [12], which simulates the radiation of quarks and gluons and the transformation of partons into hadrons. For b-hadrons, decays are further refined using EvtGen. Following generation, the simulation enters the **detector simulation** stage. Two types are utilized, which are FullSim and ATLFAST3 (AF3). **GEANT4-based FullSim** [13] provides a detailed, step-by-step simulation of particle interactions with detector materials, including electromagnetic and hadronic shower modeling, multiple scattering, and energy loss. **AF3**, by contrast, uses a parameterized detector response, approximating calorimeter and tracking responses using templates and smearing functions derived from GEANT4. This drastically reduces computational time while maintaining reasonable accuracy for many analyses. The next stage is **digitization**, where the energy deposits from detector simulation are converted into electronic signals, mimicking the readout of the real detector. This includes modeling noise, pile-up effects, and readout thresholds. Subsequently, **reconstruction** algorithms process these digitized signals to identify physics objects like photons, jets, muons,electrons and missing transverse energy (MET). The same software stack used for real data is applied, ensuring consistency between simulation and experiment. Finally, the data is saved in ROOT format during the **derivation** step, producing xAOD files that can be used for physics analysis. This full pipeline ensures a comprehensive and realistic simulation of events in the ATLAS detector.

2 Monte Carlo Simulation, Validation and Request Setup

Signal samples were generated for benchmark mass points motivated by observed anomalies. The heavy scalar H is to be produced through gluon fusion with masses ranging from 250 to 400 GeV. The decay chain $H \to SS'$ involves S (decaying to $\gamma\gamma$) with mass 150–325 GeV and S' (decaying to $\tau^+\tau^-$ or $b\bar{b}$) with mass 95–125 GeV. A total of 240,000 events were requested, distributed across both channels. Three pile up conditions are considered which is low (20k), medium (20k), and high (80k) to match LHC Run 3 conditions (mc23a/d/e). We consider both simulation types, FullSim and AF3. The branching ratios used in the cross-section calculation were taken from the CERN Yellow Report benchmarks [14], ensuring compatibility with other ATLAS reinterpretation studies. The complete cross-section for the final state is calculated as:

$$\sigma_{\text{total}} = \sigma(gg \to H) \times \text{Br}(H \to SS') \times \text{Br}(S \to \gamma\gamma) \times \text{Br}(S' \to \tau\tau \text{ or } b\bar{b}) \quad (1)$$

In collider physics, kinematic variables are used to describe the motion and spatial distribution of particles produced in high-energy proton-proton collisions. The transverse momentum, denoted p_{T}, is the component of a particle's momentum perpendicular to the beam axis. Pseudo-rapidity, η, describes the angular position of a particle relative to the beam axis and is defined as $\eta = -\ln[\tan(\theta/2)]$, where θ is the polar angle. It is preferred over the polar angle because it is approximately invariant under Lorentz boosts along the beam direction. The angular separation between two particles in the ϕ–η plane is quantified by $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, which is used to associate decay products with reconstructed jets or to isolate objects from one another. Jets, which are collimated sprays of hadrons resulting from quarks or gluons, are reconstructed using the anti-k_T algorithm [15] with a distance parameter $R = 0.4$. At the truth level, jets are clustered from stable generator-level particles and stored in the `AntiKt4TruthJets` container. A jet is considered a truth b-jet if it contains a b-hadron within a cone of $\Delta R < 0.4$ and the b-hadron has $p_{\text{T}} > 5$ GeV. These definitions ensure consistent and physically meaningful object identification before detector effects are applied. The following truth-level selection criteria were applied during event generation to retain only events with the desired signal topology while reducing sample size for computational efficiency. These selections are summarized in Table 2.

Selection	**Requirement**
Photons	$\geq$ 2 photons, $p_{\text{T}} > 20$ GeV, $\lvert\eta\rvert < 2.5$
b-jets	$\geq$ 2 jets matched to b-hadrons ($p_{\text{T}} > 5$ GeV, $\Delta R < 0.4$)
Jet kinematics	$p_{\text{T}} > 20$ GeV, $\lvert\eta\rvert < 2.5$ for all jets
Taus	$\geq$ 2 visible taus from S', $p_{\text{T}}^{\tau_{had}} > 25$GeV, $\lvert\eta\rvert < 2.5$
Jet algorithm	anti-k_T, R=0.4
Jet container	`AntiKt4TruthJets`

Table 1: Truth-level selection criteria.

3 Results

The generated signal samples were validated using kinematic distributions at the truth level. These distributions confirm that the event generation and decay chains are correctly implemented and that the samples reflect the expected physical behavior. The p_{T} distributions for the leading and subleading photon shown in Figure 2 further validate the simulation by showing the expected transverse momentum spectra for the two photons. The sum of these will show a clear peak at 150 GeV, which corresponds to the input mass of the scalar resonance S. This confirms that the $S \to \gamma\gamma$ decay is correctly modeled in the simulation. The distributions shown in Figure 3 show the correct implementation of the truth level selections. These plots collectively demonstrate that the MC generation accurately reproduces the physics processes of interest, making the samples suitable for further analysis. The derived information, as shown in Table 3, was a key part of our MC request and has been reviewed and approved by the ATLAS production team.

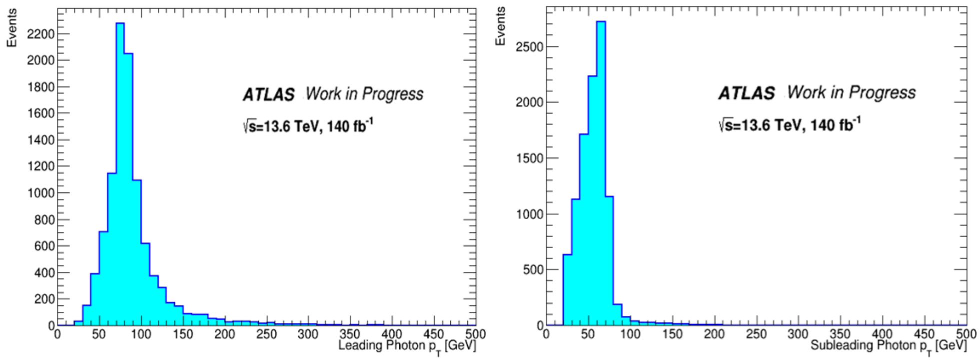

Figure 2: p_{T} distribution of the leading photon (left) and the subleading photon (right) for the same benchmark point. The distributions begin after 20 GeV, reflecting the correct implementation of the truth level selections.

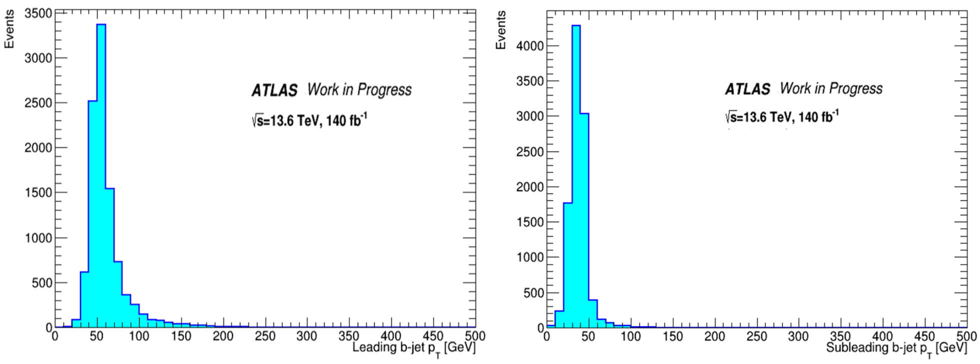

Figure 3: p_{T} distribution of the leading b-jet (left) and the subleading b-jet (right) for the same benchmark mass. The distributions begin after 20 GeV, reflecting the correct implementation of the truth level selections.

4 Conclusion

We have successfully completed the Monte Carlo request and validation for a new physics signal involving a heavy scalar H decaying to SS', with $S \to \gamma\gamma$ and $S' \to \tau^+\tau^-$ or $b\bar{b}$. Small-scale samples were generated, validated using truth-level kinematic distributions, and used to submit a formal request via JIRA for large-scale production under LHC Run 3 conditions. The request included detailed specifications for mass benchmarks, pile-up scenarios, and simulation types (FullSim and AF3). The datasets are now being produced and will soon be available on the WLCG. These MC samples will enable the search for diphoton resonances associated with τ's or b-jets, probing anomalies at the electroweak scale. Future work includes the integration of these signals into the full analysis framework, estimation of background processes, and sensitivity studies for exclusion or discovery.

Channel	m_H [GeV]	m_S [GeV]	$m_{S'}$ [GeV]	Efficiency (%)	σ [fb]
$\gamma\gamma + \tau^+\tau^-$	250	150	95	18.6	1.154
	260	160	100	22.6	0.412
	300	175	125	33.6	0.053
	350	225	125	34.2	0.010
	400	275	125	35.1	0.004
$\gamma\gamma + b\bar{b}$	250	150	95	44.7	11.1
	260	160	100	47.7	3.904
	300	175	125	56.2	0.483
	350	225	125	57.9	0.092
	400	275	125	60.1	0.003

Table 2: Summary of simulation details for benchmark mass points, including total selection efficiency and cross section after truth-level filtering. These values are used in the ATLAS MC request and will inform sensitivity studies.

References

[1] A. Drozd, B. Grzadkowski, J. F. Gunion, and Y. Jiang, "Extending two-higgs-doublet models by a singlet scalar field—the case for dark matter," *Journal of High Energy Physics*, vol. 2014, no. 11, pp. 1–40, 2014.

[2] S. Bhattacharya, B. Lieberman, M. Kumar, A. Crivellin, Y. Fang, R. Mazini, and B. Mellado, "Emerging Excess Consistent with a Narrow Resonance at 152 GeV in High-Energy Proton-Proton Collisions," 3 2025.

[3] G. Coloretti, A. Crivellin, S. Bhattacharya, and B. Mellado, "Searching for low-mass resonances decaying into W bosons," *Phys. Rev. D*, vol. 108, no. 3, p. 035026, 2023.

[4] A. Crivellin, Y. Fang, O. Fischer, S. Bhattacharya, M. Kumar, E. Malwa, B. Mellado, N. Rapheeha, X. Ruan, and Q. Sha, "Accumulating evidence for the associated production of a new Higgs boson at the LHC," *Phys. Rev. D*, vol. 108, no. 11, p. 115031, 2023.

[5] S. Buddenbrock, A. S. Cornell, Y. Fang, A. Fadol Mohammed, M. Kumar, B. Mellado, and K. G. Tomiwa, "The emergence of multi-lepton anomalies at the LHC and their compatibility with new physics at the EW scale," *JHEP*, vol. 10, p. 157, 2019.

[6] S. von Buddenbrock, R. Ruiz, and B. Mellado, "Anatomy of inclusive tt w production at hadron colliders," *Physics Letters B*, vol. 811, p. 135964, 2020.

[7] Y. Hernandez, M. Kumar, A. S. Cornell, S.-E. Dahbi, Y. Fang, B. Lieberman, B. Mellado, K. Monnakgotla, X. Ruan, and S. Xin, "The anomalous production of multi-lepton and its impact on the measurement of Wh production at the LHC," *Eur. Phys. J. C*, vol. 81, no. 4, p. 365, 2021.

[8] S. von Buddenbrock, A. S. Cornell, A. Fadol, M. Kumar, B. Mellado, and X. Ruan, "Multi-lepton signatures of additional scalar bosons beyond the Standard Model at the LHC," *J. Phys. G*, vol. 45, no. 11, p. 115003, 2018.

[9] A. Crivellin, Y. Fang, O. Fischer, S. Bhattacharya, M. Kumar, E. Malwa, B. Mellado, N. Rapheeha, X. Ruan, and Q. Sha, "Accumulating evidence for the associated production of a new higgs boson at the large hadron collider," 2023.

[10] G. Aad *et al.*, "The ATLAS experiment at the CERN Large Hadron Collider: a description of the detector configuration for Run 3," *JINST*, vol. 19, no. 05, p. P05063, 2024.

[11] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations," *Journal of High Energy Physics*, vol. 2014, no. 7, pp. 1–157, 2014.

[12] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, "An introduction to pythia 8.2," *Computer physics communications*, vol. 191, pp. 159–177, 2015.

[13] S. Bein, P. Connor, K. Pedro, P. Schleper, and M. Wolf, "Refining fast simulation using machine learning," *EPJ Web Conf.*, vol. 295, p. 09032, 2024.

[14] LHC Physics (CERN Yellow Report), "Cern yellow report working group on bsm at 13tev — gluon-gluon fusion process (twiki)," https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBSMAt13TeV#gluon_gluon_Fusion_Process, 2025, accessed July 30, 2025.

[15] A. Behring *et al.*, "Flavoured jet algorithms: a comparative study," 6 2025.