**ERC Consolidator Grant 2020**
**Research Proposal [Part B1]**

# REALDARK

Principal Investigator : Dr. Caterina Doglioni
Host Institution : Lund University
Proposal duration in months : 60

**Summary**

The Standard Model of Particle Physics (SM) describes the fundamental particles and interactions in ordinary matter. Despite the SM's success in predicting experimental results, it fails to account for the large excess of unobservable (Dark) matter in the Universe. A compelling hypothesis is that Dark Matter (DM) processes can be created from SM particle collisions such as those produced by the Large Hadron Collider and recorded by the ATLAS detector at the CERN laboratory.

In the REALDARK proposal, I will consolidate my leadership in dark matter searches with innovative data taking techniques, with a team of postdoctoral researchers and students at Lund University.

This research program makes use of real-time data analysis techniques to make the most of the vast amount of LHC data, expanding the ATLAS experiment's physics program in a resource-constrained, data-rich environment. The technical outcomes of REALDARK will be shared with the wider experimental community, providing valuable input in terms of technological advancements.

We will deploy and significantly extend techniques I have pioneered at the ATLAS detector to to search for evidence of processes related to dark matter at the LHC. We will pursue broad yet sensitive searches for dark matter models that can be discovered at the LHC (weakly interacting massive particles and dark sector particles) at a crucial time for the global quest for dark matter. The results of this proposal will provide outputs that define the direction of future experiments and theoretical efforts. The planned searches will yield either a discovery of a dark matter particle candidate ready to study in connection with astrophysical observations, or constraints on dark matter's particle nature.

## Section A: Extended synopsis of the proposal

### 1 Aims and impact of this research project

The first years of data taking at the Large Hadron Collider (LHC) [1] at CERN yielded the discovery of a new fundamental particle, the Higgs boson [2]. With this and other notable results, the LHC data has confirmed the predictive power of the Standard Model (SM) of particle physics, the theory of fundamental particles and non-gravitational interactions. However, the amount of ordinary matter described by the SM is exceeded by a factor of five by a kind of unknown matter as determined by cosmological observations, called Dark Matter (DM) [3].

Theoretical models that explain the abundance of DM in the universe include massive DM particles that interact weakly with ordinary particles, called WIMPs. These can be produced at the LHC in collisions of ordinary matter (see e.g. [4] and references therein). Creating DM in controlled conditions enables the study of its interactions with ordinary matter, complementing searches for cosmological DM in direct and indirect detection experiments and astrophysical observations. WIMP Dark matter searches have been a flagship of the physics programmes of LHC experiments. In my Starting Grant (StG) I have led novel and comprehensive searches for particles that could reveal SM-DM interactions. I have ensured that the world-leading LHC constraints that resulted had a widespread impact in the global quest for DM, within a consistent theoretical framework that has been adopted by most LHC WIMP searches so far [5].

The lack of evidence for WIMPs to date motivates a two-prong approach for DM searches at the LHC. In this Consolidator Grant proposal I will:

- advance the state of the art for WIMP searches by enhancing their sensitivity to probe even rarer interactions.
- deliver a new set of searches for models beyond the WIMP paradigm. These models postulate a new force akin to the strong interaction in the SM that could have so far escaped detection.

The discovery of new, rare processes, at a time when traditional data-taking methods consistenty agree with the SM, mandates technical innovation. The LHC collides bunches of protons up to 30 million times per second. Recording and processing all this data is unfeasible: only a small fraction of interesting data can be selected by the experiment's *trigger systems* due to constraints on both processing and storage; the rest is discarded forever. This leads to a loss of sensitivity to large areas of parameter space for DM models.

The searches in this proposal will be enabled by innovative data-taking techniques at the earliest stage of data selection and processing, increasing the utility of the data recorded by the ATLAS experiment as a whole. These techniques are called Trigger Level Analysis (TLA[1], whose ATLAS proof-of-principle was delivered by my StG [8]) and Partial Event Building in ATLAS. Combined, they reduce the size of the data used for physics analysis by a factor of X, overcoming the storage limitations that would otherwise force ATLAS and other LHC experiments to discard the majority of data of interest for many DM searches.

The physics results from this project, derived analysis products, and tools for their interpretation will be disseminated to the broader DM community in order to generate interdisciplinary impact. The outcomes of this proposal will have a transformative effect in terms of both data taking innovations and in the global quest for DM. Data storage requirements are a widespread concern for LHC upgrades, as well as for many experiments where the increase in data collection is not matched by a proportional increase in resources [9, 10]. REALDARK will directly address this challenge, and develops solutions that can be ported beyond the LHC. Future collider projects will be prioritized in this decade [11], and a number of astroparticle and non-collider experiments will start taking data [12, 13]. This proposal delivers as outputs discoveries or constraints that are valuable input to these scientific communities by using Open Science tools, defining the future direction of DM research.

I am uniquely suited to lead a team to deliver this ambitious and timely research program. As evidenced by my CV and track record, my profile combines both technical and scientific proficiency with leadership of large groups of scientists in both physics outputs and data analysis tool development. With my international collaborators and within my StG, I have led a paradigm shift in data taking techniques in ATLAS from software concept to publication. I coordinate synergistic activities towards more efficient data selection and analysis that span the entire high energy physics and beyond [9]. I have authored a number of LHC searches for DM and new phenomena, and I have coordinated ATLAS- and LHC-wide working groups instrumental to the design of DM search strategies, such as the Dark Matter

[1] Analogue to the techniques of *Data Scouting* in CMS [6] and *Turbo stream* in LHCb [7]

Working Group [14], and contributed to the prioritization of future experiments strategies in light of their sensitivity to uncovering the particle nature of DM.

## 2 Innovation to the state of the art from this project

The presence of DM does not have an explanation in the SM. The absence of a particle explaining DM at the LHC and other experiments indicates that, if DM interacts with SM particles, these interactions must be very feeble and/or the experimental signals of DM must be subtle. At the same time, the enormous data rates of modern particle experiments present a challenge: with traditional data taking methods, it is not possible to record, process and store the large datasets required to reveal DM signals, and the majority of the data has to be discarded forever milliseconds after being taken.

Two notable examples of discoveries that are impossible with traditional data taking methods are:

- Rare processes where a new particle decays into ordinary matter, mimicked by more frequent SM-only processes. These new processes are discarded together with the vast amount of irreducible SM background;
- Processes where new particles leave non-standard signals in the detector. In these cases, the exact content of the collision is too time-consuming to reconstruct within the timing budget with which a decision to keep an event must be made. Therefore, the features that distinguish signal events from the more frequent background events are missing, and the event is discarded.

These discoveries map to two classes of models that explain the particle nature of dark matter, both able to reproduce the amount of DM measured in the universe (relic density).

The first class of models corresponds to the current state of the art for LHC searches, also as a result of my StG [4, 5]. In these models, DM is a massive particle that interacts only weakly with SM particles – a WIMP. Due to their feeble interactions with the detector material, WIMPs are traditionally sought at collider experiments by selecting events with an energy imbalance where an invisible particle has escaped detection. WIMPs can be produced from LHC proton-proton collisions through a new massive particle mediating the SM-DM interaction, analogously to the W and Z mediating the weak force. In addition to decaying into DM particles, this mediator will also decay into ordinary matter, through the same interactions responsible for its production. In my StG, I delivered a new set of searches where the DM mediator decays to two jets, using LHC Run-2 data. These searches have set the most stringent constraints to date for DM mediators with masses between 250 GeV and a TeV (see shaded color areas in Fig. 1). The search in the mediator mass range 450 GeV – 1 TeV would not have been possible without the application of the TLA technique to jets, which overcomes the experiment's storage limitations by performing most of the initial data analysis and calibration already at the trigger level, retaining only a small amount of information for further analysis and discarding raw detector data. In this proposal my CoG team and I will significantly extend this technique by enabling its use for photons, electrons and muons for the first time, in order to gain sensitivity to a larger range of mediator masses as well as to new dark matter models. We [2] will validate these data taking techniques with early LHC data by constraining mediators with masses above 400 GeV, via a world-first measurement of two-jet events using the TLA technique. We will subsequently exploit the full dataset containing photons and jets at the trigger level for more sensitive and lower-mass WIMP mediator searches (see shaded color areas in Fig. 1).

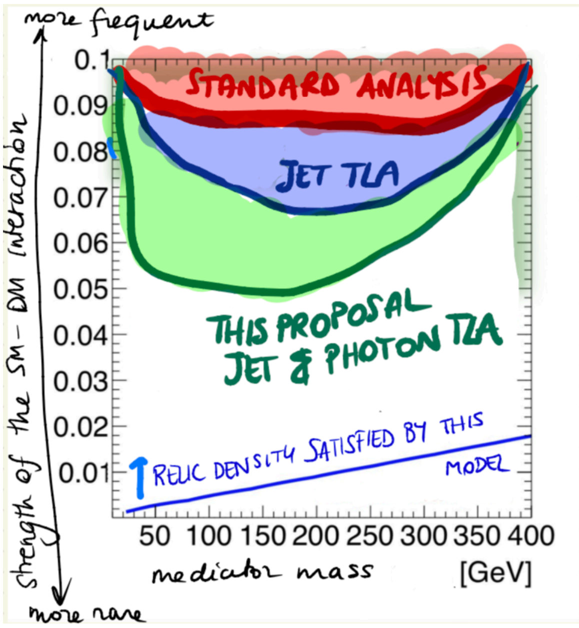


Figure 1: Sketch of sensitivity of WIMP searches in this proposal. Make this sketch into a plot with the expected sensitivity of the searches.

Motivated by constraints set on WIMPs, the second class of models postulates interactions that are much feebler, with much lighter mediators. These models [15, 16] predict a multitude of new particles in addition to the DM candidate,

[2] When the term "we" is used, it refers to myself as PI and to the personnel hired within REALDARK.

mirroring the complexity of the SM in theories similar to the strong force. An unambiguous signal of these models are "dark QCD" jets which, in addition to visible particles, are comprised of invisible particles and light dark matter mediators which decay into low-energy electrons and muons [17]. Standard searches in two-jet final states are not optimally sensitive to these scenarios. This is because of the very large SM QCD backgrounds (a problem shared by WIMP mediator searches), and because the trigger system is unable to build and identify the characteristic features of dark QCD jets in time and they are considered noise. In this proposal, we will combine the TLA and the Partial Event Building technique for the first time in ATLAS to solve this problem, after extending TLA to electrons and muons. By augmenting the limited trigger level information with the full set of raw data but only in the region of interest behind the dark jets, we will overcome storage and processing limitations, and enable the reconstruction of features that distinguish signal from background at a later processing stage where more resources are available. I will use the dataset recorded with this technique to map the parameter space of dark QCD theories and discover or set stringent constraints on models never tested at the LHC. The unique dataset recorded with this new technique will be made available for other searches and measurements, extending the potential for discovery of the entire ATLAS physics search programme.

It is crucial that the results obtained in those searches are put into the broader context of the global search for dark matter. For this reason, I will continue leading the way to define community standards for LHC searches that highlight the complementarity of the results in the searches in this project to astroparticle, non-collider and cosmological measurements. Building on the success of the Dark Matter Working Group that under my leadership set the standards for LHC WIMP search targets and for the presentation of results, I will lead initiatives that bring together the flourishing experimental and theoretical research communities studying non-WIMP DM scenarios [18].

## 3 Project organization and description

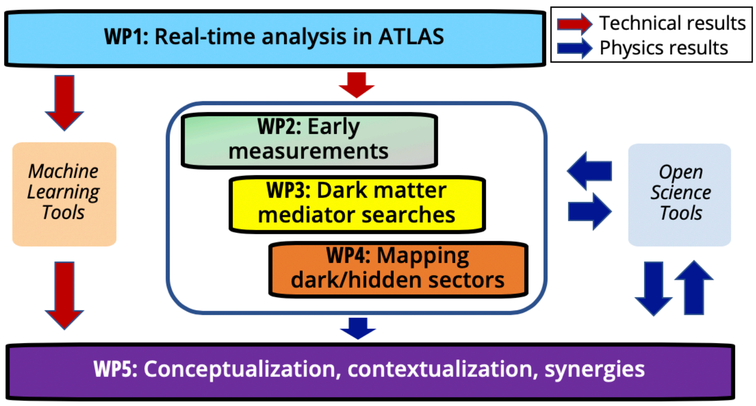


Figure 2: Schema of work packages and expected results.

The proposal consists of five logically interconnected work packages. The work packages and their interconnections are indicated in Fig. 2. In **WP1** we will overcome the technological constraints that limit the sensitivity to a variety of physics phenomena including DM, with the use of non-standard data taking and recording techniques, and employ machine learning techniques for forward-looking improvements. In **WP2**, we will lead a programme of searches and measurements using the first LHC Run 3 data and in so doing commission the techniques developed in WP1. In **WP3 and WP4**, we will use the dataset recorded with the fully commissioned techniques developed in WP1 and WP2 to perform world-best and world-first searches for current and state-of-the-art dark matter models. In **WP5**, we will interpret the results of those searches in coherent frameworks that go beyond model boundaries with the inclusion of input from LHC measurements and non-collider DM searches using Open Science tools, and work in synergy with the broader community for optimal contextualization and dissemination of results and tools.

### WP1: Real-time analysis in ATLAS

As the main technological advancement delivered by this proposal **my team will break the traditional paradigm of first recording the entirety of raw detector data and then analyzing it, by enabling ATLAS data to be reduced and processed in real-time** using the two techniques of TLA and Partial Event Building. As a pioneer of real-time analysis, I have delivered proof-of-concept studies for these techniques, but they must now be developed and commissioned on a much larger scale in WP1 in order for me to exploit them for groundbreaking DM searches with the Run-3 dataset.

**TLA** events only contain high level information reconstructed in the trigger, and they are considerably smaller than events with full detector data. This technique therefore enables the recording of a much larger number of events within the same data rate, as shown in Fig. 3. With my StG team, I have shown that this technique is effective in a proof-of-concept use case of dark matter mediators [8]. Within this proposal I intend to extend this paradigm to photons, electrons and muons so that it can be used to extend the ATLAS experiment reach to a much larger set of searches and measurements.

In the **Partial Event Building** technique, the information available to real-time analyses is augmented by a subset of the raw detector information in reduced regions of the detector. This paradigm combines the data rate enhancements of real-time analysis with the added precision of full offline reconstruction in a user-configurable manner. Non-ordinary features signaling the presence of new phenomena can then be detected *a posteriori*. This proposal will commission and use this technique for physics searches for the first time in ATLAS.

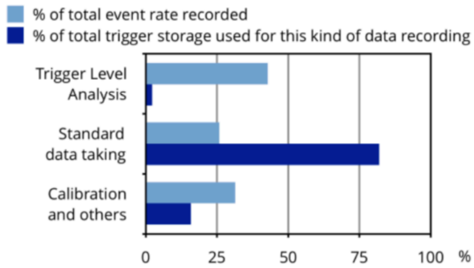


Figure 3: Fraction of event rates recorded and trigger storage used for standard and TLA technique. The event rate recorded with TLA is much higher than the rate of events for standard data taking, for a much smaller amount of storage resources, as measured in 2017 ATLAS data-taking [19].

Both of these techniques reduce data storage requirements, but they can be complemented by **further gains in data compression**. I will therefore pursue a method for data compression using machine learning (ML) algorithms. My students and I have obtained preliminary results in collaboration with other ATLAS ML and computing experts using autoencoder deep neural networks, achieving a compression of a factor of two with minimal performance loss. This is a forward-looking activity targeting the high-luminosity LHC run starting in 2026 as well as experiments beyond the LHC, as demonstrated by my ongoing collaborations with gravitational wave experiments.

### WP2: Early measurements and searches

During the shutdown between the Run-2 and Run-3 LHC data taking periods, the ATLAS experiment trigger system is undergoing significant upgrades [20], and the trigger and reconstruction software is being rewritten [21]. Thorough commissioning and testing is mandatory, first using simulation and cosmic ray data, and subsequently with LHC commissioning data. In WP2, **together with my team, I will test a completely new software trigger framework and its performance via early measurements and searches**, using the already-established TLA techniques from Run-2 [8]. This work will be documented by physics and technical publications.

We will select events where at least two jets are detected and reconstructed in real-time within the trigger system, and compare the performance of their reconstruction and calibration at the trigger level with traditional data taking. Beyond the technical publications documenting the new software framework and its commissioning with LHC data, this WP will lead to two physics publications: We will use this dataset to perform a world-first measurement of the jet cross section with trigger level data, and for the first Run-3 search for low-mass DM mediators. We will make extensive use of the Open Science tool RECAST [22] to preserve the end-to-end analysis workflow, from detector data to final plots so that we can take advantage of this work for the searches in WP3.

### WP3: WIMP dark matter mediator searches

In WP3, **my team and I will use the TLA technique to search for decays of WIMP mediators in a scenario that is not fully constrained by existing searches**. This search targets mediators with masses between 150 and 350 GeV, a mass range close to where the SM massive force mediators reside that is still poorly constrained, as neither traditional searches nor jet-only TLA searches have optimal sensitivity (see Fig. 1).

We will employ the detector signature that I introduced to ATLAS during Run-2 using traditional data-taking methods, where the mediator is produced in association with an additional jet or photon [23] and yields the world-leading constraint in the upper mass range of this region. **Performing this search with the TLA method, after the implementation of photons planned for WP1, will significantly increase its discovery potential**, allowing for an order of magnitude more signal events to be recorded.

During the 2022-2025 grant period, we will also analyse data in the two-jet signature to search for mediators with masses above 350 GeV with the entire Run-3 dataset, together with a postdoc. This analysis will capitalize on the end-to-end analysis workflows set up in WP2 that make the analysis more efficient and ensure that the Run-3 results can be combined with the full Run-2 results and used to constrain a wide set of physics models [24].

### WP4: Mapping dark Quantum Chromodynamics

While the searches in WP2 are powerful probes of WIMPs, they are not sensitive to DM mediators whose interaction with the SM is even feebler, or to GeV-scale DM mediators. For this reason, in WP4 **my team and I will deploy the combination of TLA and Partial Event Building developed in WP1 to discover signs of DM in a different class of models yielding new detector signatures**. This proposal focuses on two new searches.

Firstly, we will search for evidence of models where the DM candidate particles are produced in association with a large number of other dark sector particles. Since these dark sector particles also interact through the SM's strong force, while the DM particles escape detection, this leads to *semi-visible dark jets* [16]. Prototypes of these searches are being developed with Run-2 data, but they are limited to very high masses and to a narrow range of parameters of the model. In this proposal, we will scan a much larger part of the parameter space with the unique sensitivity enabled by this dataset, focusing on the region where the model can explain the relic density of DM [25].

The second search builds on the semi-visible jet results and exploits the availability of electrons and muons at the trigger level enabled by WP1, targeting models where low-energy leptons from the decays of a light dark mediator or a dark Higgs boson are also found within dark jets [17, 26]. Current searches are restricted to higher-energy leptons, and this dataset will extend the sensitivity to low-mass, low-energy and exotic objects overlooked in searches so far.

### WP5: Contextualization, conceptualization, synergies

WP5 defines frameworks and optimal search regions in the context of the global DM search community, resulting in enhanced conceptualization and dissemination of tools and results. This dedicated WP ensures that the results of the project have a broad research impact that exploits both local synergies and synergies with other communities.

Within WP5, my team and I will pinpoint the optimal parameter space to be targeted for the new searches in WP4, and adapt the selection of events to be stored in the dataset accordingly. This will be delivered in collaboration with the local Lund theory group, who are authors of the most widely used implementation of the theory in event generation software [27], and using software that enables the use of precision results released by LHC collaborations [28] that has never been used in the context of these models before. In WP5 we will ensure that all physics results from WP2-4 are disseminated and implemented in Open Science tools (e.g. RECAST, HEPData) that meet the needs of other communities in terms of reproducibility, usability and complementarity. The successful experience of the Dark Matter Working Group for the LHC DM community is the stepping stone for a new, ambitious initiative that is being brought to the attention of the community [18] that will enable results from this project to reach a much broader context, including non-collider experiments, astrophysics, cosmology and multimessenger astronomy. An innovative aspect of this project is that its dissemination strategy is not limited to physics results but includes algorithms and tools. As I and many others advocated during the process leading to the update of the European Strategy of Particle Physics [29], challenges related to data acquisition, selection and analysis can be tackled more effectively and efficiently by going beyond a single experiment's boundaries. For this reason, WP5 also includes collaborative work on data compression with other experiments where a reduced storage footprint is necessary to increase the physics potential within the same resources.

## 4 Timeliness and timeline of the research program

I will lead a research team of two postdoctoral researchers and two students, working in pairs on the two lines of physics analysis in this proposal, and frequently collaborating to share technical tasks and software. An additional software engineer will augment the team during the demanding time of the LHC Run-3 startup. For WP1, the team will be joined by talented Lund University undergraduates that I have a track record of recruiting and training who will work on the forward-looking ML compression activities. The research program spans the entire upcoming LHC data taking period (Run-3). The 2021-2026 period is the ideal time to advance the state-of-the-art in processing of large datasets and DM searches, as it ensures continued impact through a significant extension of my current successful StG research program that pioneered proof-of-principle real-time searches in ATLAS. The current LHC schedule includes an initial commissioning period where innovations can be deployed and tested via early measurements with start-up data, which forms the first part of this proposal (WP1 and WP2), readying them for the second phase of the proposal where the LHC will be in production mode. The second (WP3 and WP4) phase of this proposal will exploit the wealth of LHC data delivered by WP1 and WP2 for novel DM searches, with a dataset that will be more than twice as sensitive to the physics observables of this proposal than the data collected so far. Throughout the timeline of entire proposal, in WP5 we will share tools and results within the global DM community, to answer one of the most pressing questions of our universe with a discovery or by defining future search directions in which such a discovery will be made.

### References


[1] L. Evans and P. Bryant, JINST **3** (2008) S08001.
[2] ATLAS and CMS Collaborations, JHEP **08** (2016) 045, arXiv:1606.02266.
[3] G. Bertone and D. Hooper, Rev. Mod. Phys. **90** (2018), no. 4 045002, arXiv:1605.04909.
[4] A. Boveia and C. Doglioni, Ann. Rev. Nucl. Part. Sci. **68** (2018) 429, arXiv:1810.12238.
[5] LHC Dark Matter Forum, D. Abercrombie *et al.*, arXiv:1507.00966.
[6] CMS Collaboration, Phys. Rev. Lett. **117** (2016), no. 3 031802, arXiv:1604.08907.
[7] LHCb Collaboration, Comput. Phys. Commun. **208** (2016) 35, arXiv:1604.05596.
[8] ATLAS Collaboration, Phys. Rev. Lett. **121** (2018), no. 8 081801, arXiv:1804.03496.
[9] HEP Software Foundation, J. Albrecht *et al.*, Comput. Softw. Big Sci. **3** (2019), no. 1 7, arXiv:1712.06982.
[10] G. Allen *et al.*, 2018. arXiv:1807.04780.
[11] R. K. Ellis *et al.*, arXiv:1910.11775.
[12] https://www.appec.org/wp-content/uploads/Documents/Current-docs/APPEC-Strategy-Book-Proof-19-Feb-2018.pdf, 2018. [Online; accessed 7-January-2019].
[13] J. Beacham *et al.*, J. Phys. **G47** (2020), no. 1 010501, arXiv:1901.09966.
[14] LHC Dark Matter Working Group, https://lpcc.web.cern.ch/content/lhc-dm-wg-wg-dark-matter-searches-lhc, 2016. [Online; accessed 7-January-2019].
[15] M. J. Strassler and K. M. Zurek, Phys. Lett. **B651** (2007) 374, arXiv:hep-ph/0604261.
[16] T. Cohen, M. Lisanti, H. K. Lou, and S. Mishra-Sharma, JHEP **11** (2017) 196, arXiv:1707.05326.
[17] D. Curtin, R. Essig, S. Gori, and J. Shelton, JHEP **02** (2015) 157, arXiv:1412.0018.
[18] https://indico.cern.ch/e/iDMEu, 2019. [Online; accessed 7-January-2019].
[19] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ComputingandSoftwarePublicResults, 2019. [Online; accessed 7-January-2019].
[20] ATLAS, ATLAS Collaboration, .
[21] ATLAS, G. A. Stewart *et al.*, J. Phys. Conf. Ser. **762** (2016), no. 1 012024.
[22] A. Schuy, L. Heinrich, K. Cranmer, and S.-C. Hsu, in *Proceedings of the Meeting of the Division of Particles and Fields of the American Physical Society (DPF2019)*, 2019. arXiv:1910.10289.
[23] ATLAS Collaboration, Phys. Lett. **B795** (2019) 56, arXiv:1901.10917.
[24] J. H. Kim, K. Kong, B. Nachman, and D. Whiteson, arXiv:1907.06659.
[25] E. Bernreuther, F. Kahlhoefer, M. Krämer, and P. Tunney, Submitted to: J. High Energy Phys. (2019) arXiv:1907.04346.
[26] A. Falkowski, J. T. Ruderman, T. Volansky, and J. Zupan, Phys. Rev. Lett. **105** (2010) 241801, arXiv:1007.3496.
[27] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput. Phys. Commun. **178** (2008) 852, arXiv:0710.3820.
[28] J. M. Butterworth *et al.*, JHEP **03** (2017) 078, arXiv:1606.05296.
[29] C. Doglioni, in *Proceedings of the European Physical Society Conference on High Energy Physics (EPS-HEP) 2019, special ECFA-EPS session*, 2019. arXiv:1912.12745.


**Note:** The PI is the editor of Refs. [4, 5, 8, 23, 29], the initiator of the activities in Refs. [14, 18] and an author of Refs. [9, 11] as well as of all publications by the ATLAS collaboration.