ERC Starting Grant 2020
Research Proposal [Part B2]

Part B2: *The scientific proposal*

1 State of the art and objectives

Experiments at the CERN laboratory in Geneva, studying collisions from the Large Hadron Collider (LHC) [1], have verified the predictive power of the Standard Model (SM) of particle physics. Nevertheless, the SM must be an incomplete theory. It describes all known fundamental particles and their non-gravitational interactions, but it completely lacks any particle consistent with the astrophysical evidence for Dark Matter (DM) [2]. The **failure of the SM to account for DM**, which in the universe is about five times as abundant as the matter described by the SM, remains one of the most important puzzles in high energy physics and astrophysics research.

Solving this puzzle requires new experimental data. Absent clues from the LHC and other particle physics experiments, astrophysics observations provide evidence of DM interacting gravitationally, and hints of very feeble interactions between DM and SM matter [3, 4]. **Probing for these interactions requires larger and larger datasets to reveal.** LHC experiments can play a key role in discovering DM-SM interactions, complementary to other particle physics and astrophysics experiments [5], since DM particles can emerge from high-rate collisions of SM particles. At the same time, LHC experiments and many modern particle physics experiments face a data acquisition challenge. With traditional data acquisition methods, it is not possible to record and store the extremely large datasets needed to discover DM or other rare processes buried in substantially larger backgrounds.

As a senior lecturer at Lund University, I will lead a research team to search for signs of DM and other new phenomena in LHC data to be collected in 2021–2026 by the ATLAS experiment. We will deploy a **new data-taking paradigm** that significantly increases the discovery potential of LHC data for the entire ATLAS experiment, even though the LHC center-of-mass energy and dataset size are expected to just be comparable to previous data-taking runs. We will use data collected with new techniques to **search for signals of broad classes of compelling DM models** and new phenomena that are rare or have so far been neglected, leading to discoveries or constraints with an impact on the global DM community. We will disseminate the results of this research and its innovations through working groups and cross-experimental collaborations of theorists and experimentalists from collider and astrophysics experiments, and to others outside academia. This five-year program will advance the state-of-the-art in data-taking and data-processing of enormous datasets at scientific experiments and shed further light on one of the greatest mysteries of our universe.

1.1 The state of the art in theoretical frameworks for dark matter

Gravitational and cosmological observations [2] point to the existence of dark matter, for which the SM does not provide any explanation. Many mechanisms explaining the present amount of dark matter in our universe (called *relic density*) imply that, if DM is a new particle, it must have a connection with SM particles [3, 4, 6], allowing for its observation in experiments made of SM matter [1].

A popular DM candidate satisfying freeze-out relic density is the WIMP, a TeV-scale stable particle with interaction strength comparable to the weak force. Owing to these characteristics, WIMPs are detectable by a variety of experiments. Indirect detection experiments (ID) could observe excesses of SM particles over astrophysical backgrounds due to WIMP annihilation in DM-rich regions. Direct detection experiments (DD) could detect the interaction between incoming WIMPs and recoiling target nuclei within the detector. The LHC participates in the quest for DM: while it needs DD and ID to ascertain the cosmological origin of a DM-like particle discovery [2], it would allow the production of DM in controlled conditions for a deeper exploration of DM-SM interactions.

[1] Alternative mechanisms and explanation for DM exist, e.g. [7–10]. Given our ignorance on the genesis of DM, pursuing a broad theoretical and experimental approach that considers different possibilities is motivated and advocated in WP5.

[2] Particles that look like DM in LHC experiments may decay after leaving the detector, with a lifetime incompatible with DM's cosmological timescales.

Exploiting **synergies between different experiments** in terms of common discovery potential requires a **common theoretical ground**.

WIMP DM models range from complex yet fully-specified theories such as supersymmetry (SUSY) [11] to effective field theories (EFT) at energy scales where the exact details of the interaction can be ignored [12]. During the course of my StG, I co-led the Dark Matter Working Group and produced a series of recommendations that would define the state-of-the-art for benchmark models for generic LHC DM searches [13], focusing the community around **simplified models** that reproduce relevant experimental features using a limited number of parameters. Simplified models introduce a new particle acting as the mediator of the interaction between ordinary matter and dark sector particles, as a step forward from EFTs. This mediator can be a spin-1 particle analogous to the Z boson (a Z' boson) [14–16], a new spin-0 scalar much like a new Higgs boson [17–19], or a spin-2 graviton [20]. While additional interactions and modifications are required to make simplified models self-consistent (see e.g. [21]), they do not change the main experimental feature: given that the mediator has been produced from the interaction of SM particles, then it can also decay into SM particles. Sensitivity to these visible mediator decays gives LHC experiments unique insight into the dark sector. An astrophysics discovery of DM could be further characterized in terms of the interactions between the mediator and its SM decay products, and a LHC discovery of a mediator particle aided by relic density measurements could aim attention at particular regions in DM parameter space. Models with a vector or scalar mediator have been among the DM models chosen to benchmark the sensitivity of future facilities and compare it to the next-generation DD and ID experiments, as input to the update of the European Strategy of Particle Physics [22]. This complementarity of DD, ID and LHC experiments [23] is fully rooted in the LHC DM search program and has been established during the course of my StG [24].

Simplified models are well-motivated benchmarks also at a time when no experimental hints are observed, as they allow for a systematic exploration of experimental search targets by scanning the parameter space (DM mass m_{DM}, mediator spin and mass m_{Med}, couplings to DM g_{DM} and SM g_{SM}). The systematic exploration of the mediator decays in the left-hand side of Fig. 1 motivates the searches in WP2 and WP3 in this proposal.

In parallel to considering WIMP DM models, it is also imperative to **consider alternative DM models that could have escaped detection** so far. An example is a class of models including mediator-like particles that do not directly connect DM and SM, but rather provide a **hidden portal** with much weaker interactions between a more complex dark sector and SM particles [25]. The theoretical and experimental ground for hidden sector models is much less explored than for WIMPs, due to the additional complexity in the parameter space and to their non-standard detector signatures requiring new, dedicated experimental techniques.

The symmetries of the SM can be used as guiding principle to construct the particle spectrum and the interactions of the new dark sector. A new $U(1)$ symmetry introduces a portal particle called *dark photon* [26, 27]. The dark photon mixes with the ordinary photon, leading to rare but observable signatures from its decays to SM particles. This model is currently used as a benchmark for many non-LHC experiments and for future facilities [22]. Models with a new SU(2), also termed *dark QCD* models, predict that their fundamental components (dark quarks, q_{dark}) are confined at an energy scale comparable to that of QCD. This often leads to signatures of highly collimated particles, resembling hadronic jets that include dark sector particles and their decay products (*dark jets*).

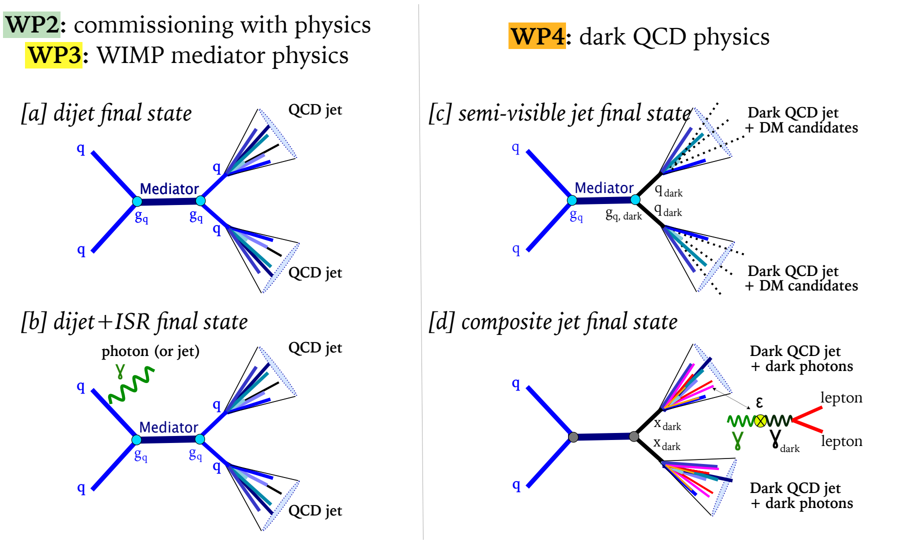

Figure 1: Sketches of search signatures in this proposal.

A number of concrete realizations of theories including one or more of those new symmetries exist (e.g. asymmetric

DM [28], electroweak SUSY models [29], self-interacting DM [30], strongly-interacting DM [31]). However, the large number of parameters and particles, and their subsequent variability in terms of experimental manifestations, mandates **signature-driven searches** rather than searches targeting specific theory benchmarks. For this reason, the data-taking techniques in REALDARK are designed to **record data that could contain a variety of these signatures, and that would otherwise be discarded**. The searches in WP4, using the dark jets signatures shown in the right-hand side of Fig. 1 **target two uncovered dark sector signatures** as concrete use cases for this dataset. As shown in the right-hand side of Fig. 1, we will first search for dark jets containing a sizable fraction of thermal relic DM particles (semi-visible jets), using the Z' boson simplified model as building block [31, 32] and dark jets where the showering process interleaves dark photons and dark hadrons, leading to hadronic jets with an anomalous leptonic content (composite jets) [29, 33].

The emergence of new models for DM and new ways to look for them indicates that the DM community is thriving. Depending on the SM-DM interaction strength and on the DM characteristics, a wealth of new experiments can contribute to the quest for WIMP and non-WIMP DM. These range from accelerator-based to table-top experiments [34], to experiments sensitive to gravitational wave signals [35].It is clear that, given the breadth of possible explanations for DM, an equally broad experimental and theoretical approach is needed, and that new synergies can be exploited beyond the already-mentioned complementarity between DD/ID and colliders searches. This consideration inspires the work planned for WP5, taking place within a newly-established common platform for dissemination of results and tools using open science tools [36].

1.2 The state of the art in experimental tools for discovering dark matter and new phenomena

1.2.1 LHC and ATLAS: overview and schedule

The LHC will restart operations in Summer 2021 and continue delivering proton-proton collision data until the end of 2024. In this data taking period (called *Run-3*), up to 250/fb will delivered to the ATLAS experiment [35].It is expected that the bulk of the data will be collected in 2022 onwards, after a first period of commissioning for the machine and the experiments.

In Run-3, the LHC collision energy will not be significantly increased, and the dataset size will be only slightly larger than the Run-2 dataset (189/fb). It is clear that the discovery potential warranted by the upcoming dataset will not improve by statistics alone. **New methods and technical innovations are required to discover rare processes**, at a time when traditional data-taking confirms the SM.Innovating the ATLAS data acquisition and selection system is an integral part of this proposal, as the key to extract a much larger amount of useful information from the upcoming LHC dataset. This proposal is especially timely given that the improvements within REALDARK can be tested with physics in early data and are required to be made in advance of the LHC production phase for full exploitation of the LHC dataset.

1.2.2 The ATLAS trigger and its upgrades

Recording the entirety of the LHC data (up to 30 million collision events/second) is unfeasible due to storage and computing limitations. **Only a small fraction of interesting events can be selected** by the ATLAS trigger system.

The ATLAS trigger is composed of two levels [35]. The first hardware level (called Level-1) performs an initial coarse selection within 2.5 μs. Selected events are passed to the software-based High Level Trigger (HLT), implemented on a computing farm. The HLT performs a more refined yet still resource-constrained event reconstruction and data analysis *online*. This informs the final decision on whether to discard the event or keep the full detector information for further *offline* reconstruction and analysis. The total amount of events that can be retained after the HLT decision is directly limited by the available amount of storage space, since traditional offline analysis currently requires the full set of detector information for reconstruction of physics objects (e.g. electrons, muons, photons, jets...). This motivates the implementation of alternative data-taking techniques to overcome these limitations, as proposed in WP1 and discussed in the next section.

The rates of useful events recorded for offline analysis are also indirectly limited by the available HLT computing: a number of high-level features that would help in distinguishing signal from background are **too expensive to reconstruct given the available HLT resources**. An example of such a feature is the track-collision vertex association for physics objects, as this enables removing spurious contributions from simultaneous proton-proton collisions other the one of interest (pile-up). In Run-3, the computing power of the ATLAS HLT farm will increase significantly and work is undergoing to optimize the trigger tracking algorithms [3]. This will enable pile-up rejection techniques, and more extended tracking at the trigger level, to be available. The ATLAS L1 trigger is also undergoing significant improvements before the start of Run-3, allowing a more efficient first-pass selection. As part of my StG, my team has developed control software for the board which will permit using information from the entire hadronic calorimeter at once (gFEX), and contributed to the software for the new jet board (jFEX). The board that reads information from the electromagnetic calorimeters (eFEX) will allow for more refined distinction between electrons and photons already at L1. The combination of these planned **Run-3 ATLAS trigger improvements enable and motivate the work** in WP1 and the searches in WP2-4.

1.2.3 Review of analysis strategies for DM and related particles at the LHC

Traditional searches for **WIMP Dark Matter** at colliders seek an excesses of imbalanced events where weakly or non-interacting particles escape the detector unseen [5]. If the processes sought in traditional searches via a mediator particle as described in Ref. 1.1, fully-visible signals should also be present from the decays of the mediator into SM particles. Searches for visible signatures is an established complementary alternative to searches for invisible particles, with different challenges and sensitivity [35].

Visible decays of the Z' also connect DM mediator searches with the well-known class of searches for new resonances at particle colliders. New resonant particles have a large array of theoretical motivations beyond DM, and are not yet fully constrained by existing searches [35]The searches in WP2 and WP3 in this proposal focus on a well-motivated yet experimentally challenging category of decays (see e.g. [16]). Having been produced via a quark-quark vertex (Fig. 1), the mediator will also decay back into quarks, leading to a two-jet (dijet) detector signature (Fig. 1 [a]).

A mediator decaying into two quarks leads to a resonant signal in the dijet invariant mass spectrum, peaking atop the smooth QCD background. The relatively low background rates of high-mass hadronically-decaying resonances permit collecting all events in full above a threshold of approximately 1 TeV [35]. Below 1 TeV, large amounts of QCD backgrounds overwhelm the experiment's data storage resources, forcing to discard background together with signal. For this reason, dijet resonances with masses around the electroweak scale (≈100-200 GeV) have been very difficult to probe with standard data-taking techniques. To date, searches in this region are only sensitive to coupling strengths well above those of the weak force. This is a region where the weak force mediators reside, and that is favored in WIMP and non-WIMP Z' models fitting ID excesses [35]. Pushing the collider sensitivity to be comparable to DD in this region will help validate and characterize a possible DM interpretation of these ID results [20, 21].

Discovering **resonances with sub-TeV masses require novel data taking techniques and/or targeted signatures**. During my StG, my team and I have pursued both directions within ATLAS. Following the example of the Data Scouting technique in CMS [35] and the Turbo Stream in LHCb [35], we prototyped the **Trigger Level Analysis (TLA) technique**. In TLA, high-level jet information is retained for all dijet events reconstructed by the HLT, discarding raw data. TLA allows for for a much smaller event size and much higher event rates [35].However, TLA dijet searches are limited by the L1 event rate to resonances with masses above 450 GeV. Other detector signatures with reduced background rates are required to reach lower mediator masses. One such signature that my collaborators and I introduced to the LHC is the dijet+ISR signature. In this search, the resonance recoils against a high-p_T photon or gluon radiated from the initial-state quarks (Fig. 1, [b]). The QCD background is reduced owing to the requirement of an energetic object, but so is the signal acceptance. The combination of ATLAS and CMS dijet+ISR searches [35][4] with searches exclusively looking for resonances decaying in heavy flavor quarks [35] and searches where the resonance is boosted [35] improves the sensitivity of LHC experiments but still leaves significant room for discovery around the EW scale.

While the landscape of DM mediators is well mapped even though not yet fully constrained, the state-of-the-art of **hidden sector and dark QCD searches** leaves much room for joint experimental and experimental improvement.

Hidden sector models generally imply more difficult detector signatures than more established benchmarks. An example are particles with a long lifetime, whose decays away from the interaction point are not suited to standard re-

[3] See Ref. [35] for an example of the improvements from optimizing tracking algorithms on simulated data containing 200 simultaneous collisions for the ATLAS high luminosity upgrade

[4] The CMS search in Ref. [35] combines the dijet+ISR signature with the Scouting technique, but has a similar sensitivity as the offline ATLAS search due to being restricted to a dataset with special triggers.

construction techniques [35]. Unusual features, such as displaced decays or unexpected particles within the busy environment of a hadronic jet, are too expensive to reconstruct at the HLT and require specific data formats. This forces searches to rely on triggers that let large amounts of background through (e.g. missing transverse energy or purely hadronic triggers) and subsequently induce limitations on the signal rates, or to trigger on specific signatures and partially lose model-independence. It is more advantageous to only rely on the broad features of this class of models to apply a coarse initial L1 selection, store high-level objects reconstructed at the HLT, and record only the raw detector information needed to defer reconstruction of distinctive features at a later date. This can be done with a combination of TLA and a technique called **Partial Event Building (PEB)**. PEB allows recording selected raw detector information in a limited region of the detector, and in Run-2 it was implemented in ATLAS for for B-physics and calibrations [35].This allows to systematically target a variety of signatures, especially in view of a multi-year shutdown at the end of Run-3. This is part of the work in WP1 described in the next section.

The approach for choosing and designing Run-3 searches to be performed within the timeframe of this proposal is signature-driven rather than model-driven. As specified in Sec. 1.1, we adopt the grounding assumption that the confinement scale is sufficient to produce dark jets. As an initial guide to the experimental state-of-the-art for the searches in WP4, dark jets can be classified according to their main characteristics. Extending Ref. [32,33], Fig. 2 characterizes dark jets according to:

- how promptly the full content of the dark jets appears in the detector, indicating that they contain long-lived constituents (emerging/trackless jets);
- what is the fraction of invisible particles in the dark jet (R_{inv}), indicating the presence of stable dark sector particles and DM candidates (semi-visible jets);
- whether jets contain an unusual number of leptons, from leptonically-decaying dark photons (composite jets).

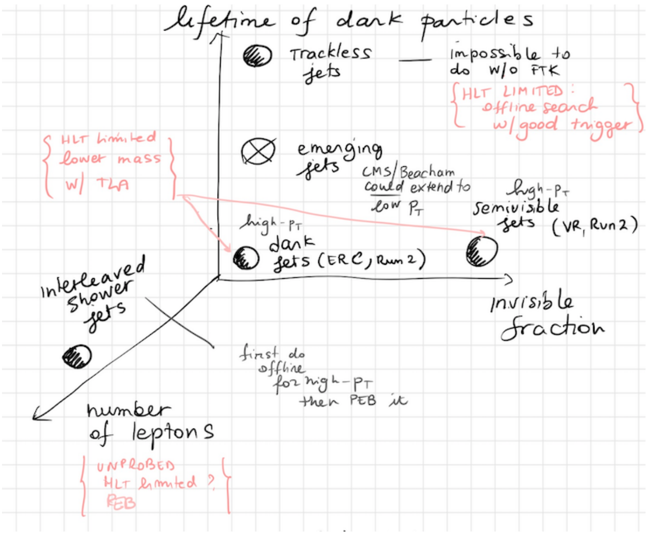

Figure 2: Sketches of search signatures in this proposal. **This will be a sketch in powerpoint indicating what is covered and what is uncovered**

Searches for dark shower models in ATLAS and CMS currently cover the cases of:

- jets where the constituents are prompt and fully visible, but the fragmentation is different [33]. I am pursuing the first-ever LHC search for this signature and expect a publication by Fall 2020.
- jets containing a large number of displaced tracks (emerging jets, [35], trackless jets [35]),
- jets composed exclusively of leptons (lepton-jets, [35]).

From this classification, **Semi-visible jets** (Fig. 1 [c]) are experimentally uncovered at the time of writing, and so are **composite jets** where dark QCD and dark photon showers are interleaved (Fig. 1 [d]). This motivates the choice of the two WP4 searches.

Experimental attention generally stimulates theoretical development and extensions of unprobed models, as well as efforts towards their systematic classification. As part of this proposal, we will also open a number of discussion channels with the theory, astroparticle, non-collider and nuclear physics communities, within WP5. This cross-talk will be useful for input on search targets, on the simulation of benchmark models and for the contextualization of search results.

1.3 Objectives

As outlined in part B1, this proposal is composed of 5 interconnected work packages. The proposal's overarching aim is to discover or constrain the particle nature of dark matter via its production at the LHC and the contextualization of these results in the global DM theoretical and experimental landscape. The basis for reaching this aim is the implementation of data-taking techniques that enable the ATLAS experiment to obtain a much bigger physics output from the upcoming LHC dataset without requiring a significant increase in resources. Physics results and software tools resulting this proposal will be shared with the broader DM community.

The main objectives of the REALDARK WPs are:

(WP1) To extend the capabilities of the ATLAS trigger system with a comprehensive set of real-time analysis techniques. In WP1, we will:
1. implement photons, electrons and muons in TLA, starting from the jet prototype developed in my StG;
2. contribute to a suite of reconstruction and calibration techniques for HLT objects that can also be used offline;
3. implement a combination of PEB and TLA techniques targeting complex detector signatures;
4. study ML data compression techniques to be used for TLA data and beyond in future LHC runs.

(WP2) To commission the upgraded ATLAS trigger with early Run-3 data, using searches that already have a solid methodological basis from Run-2. In WP2 we will:
1. study the performance of physics objects reconstructed and calibrated at the HLT and offline;
2. deploy and validate new calibrations and real-time analysis techniques with well-established generic searches for new phenomena in the dijet spectrum;
3. prepare end-to-end analysis code to be used for faster analysis turnaround and as part of the dissemination strategy.

(WP3-WP4) To use the LHC dataset recorded with novel data taking techniques to perform searches sensitive to two broad classes of DM models:
- in WP3, we will search for hadronic decays of WIMP DM mediators in the dijet and dijet+ISR final states using a TLA that including both photons and jets, to probe electroweak-scale mediator masses not fully explored by traditional searches.
- in WP4, we will search for dark matter candidates and dark photons within the jets as predicted from dark QCD, using a combination of TLA and PEB to recover sensitivity to signals that escape traditional detector reconstruction techniques.

(WP5) To disseminate and communicate physics results and tools to make them possible to the broader DM and experimental community. In WP5 we will:
1. collaborate with theory experts to identify the most promising parameter space for searches in WP4.
2. contribute to organizing and leading the new cross-community initiative for DM that I co-founded in 2019 (iDMEu), including nuclear physics, astroparticle physics and particle physics.
3. disseminate technical outcomes of WP1 to other experiments at the LHC and beyond.
4. disseminate the outcome of the searches in WP3 and WP4, whether they are a discovery to be characterized or constraints that will guide future DM experiments. The latter two objectives will be supported by my role in the HEP Software Foundation and by my involvement in the ESCAPE project, as detailed below.

The objectives of each WP are developed in more detail in the following sections.

1.3.1 Objectives of WP1

The first objective of WP1 is to **deploy a comprensive TLA stream in ATLAS within the new multithreaded ATLAS HLT software**. In Run-3, this project will allow offline-quality photons, muons and electrons to be used for physics analysis, in addition to jets. The success of the jet prototype has been demonstrated in Run-2 within my StG. The use of a jet TLA allowed to lower the HLT jet threshold from 420 GeV in traditional analysis to 220 GeV in TLA, bringing orders-of-magnitude improvements in the number of recorded events. A TLA implementation of HLT photons will bring significant improvements to the sensitivity of the dijet+ISR DM mediator search. The HLT analyzes all photon and electron candidates that have a p_{T} above 30 GeV, while only single-photon events with a p_{T} above 150 GeV are retained in traditional analysis. The work done and lessons learned in the case of trigger-level jets will be the stepping stone for adding photons to the TLA stream, and doubling the signal acceptance for the dijet+ISR search. The implementation of electrons and muons in TLA will follow that of photons. While electron and muon TLAs can still bring significant improvement to e.g. dark photon searches below the Z peak [35], the aim of this work is to gain confidence with the challenges of reconstruction and calibration of different physics objects with constrained computing resources.

The second objective of WP1 is to **implement and reconstruct a data stream combining physics objects reconstructed at the trigger level and selected raw information in restricted regions of the detector**, through the combination of the TLA and Partial Event Building (PEB) techniques. Deployed for the first time in ATLAS in this proposal, this combination will maintain a sufficiently small data format with an amount of information equivalent to traditional techniques. It will be a necessary ingredient to enable searches in WP4, and can be used for the characterization of any excesses observed in TLA searches. Preliminary estimates based on the Run-2 $B-$physics PEB place the size of TLA+PEB events to less than half of a standard event. To remain advantageous, TLA and PEB must have the minimal possible footprint in terms of both storage and computing power. Joint work betwen WP1 and WP2-4 will ensure that these constraints are met with dedicated trigger selections and selections. The budget of this proposal also includes storage servers that will be used to store this data in 2023-2024.

As a third objective of WP1, we will **develop the calibration techniques that are necessary to enable physics analysis** with a reduced HLT-level data format, achieving near-parity with offline performance. With my StG team, I have been responsible for reaching a 1-permille agreement between the jet energy scale of HLT jets with respect to offline jets [35].Within REALDARK we will ensure that this is still the case for jets (as our main observables for WP3), as well as for other physics objects. Differences in the energy scale of HLT and offline objects also lead to inefficiencies in the trigger selection, as events that should pass the trigger are rejected due to HLT miscalibrations. As an example, the minimum p_T threshold applied to jets used in physics analysis is set to be higher than the HLT threshold due to these differences, leading to a substantial waste in terms of events that are recorded but not used for analysis, as shown in Fig. 3. Preliminary studies performed within my StG and with collaborators show that, due to the steeply falling jet energy spectrum, the amount of events that are recorded but never used by offline analysis can be as high as 60% for certain trigger chains. The application of TLA-motivated improved calibration constants at the HLT during e.g. LHC technical stops (allowing sufficient time for the reoptimization of the trigger menu) will improve the fraction of useful data recorded and impact the overall ATLAS data taking strategies.

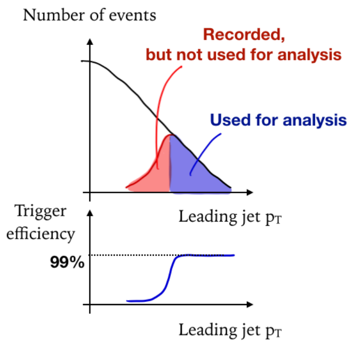

Figure 3: Sketch of how trigger inefficiencies caused by a mismatch between the HLT and offline jet energy scales causes events to be recorded but never used for analysis.

The fourth objective of WP1 is to **further reduce the storage load of TLA events (and more generally ATLAS data and simulation) by compressing this data**. TLA data is ideally suited for studies of more aggressive (lossy) compression, since it has already been shown that it can be made robust against loss of information using dedicated calibrations. If compression and decompression algorithms are sufficiently fast (order of milliseconds) they do not significantly increase the amount of resources needed for data processing. The use of machine learning techniques for fast and performant compression (e.g. of images) is widespread: inspired by this, my ATLAS collaborators and I have been supervising Lund University Master's students in preliminary tests to compress TLA data using deep autoencoders [35]. When compressing 2017 TLA jet data, these studies demonstrated that a compression of factor better than 2 can be achieved with a negligible performance loss. In this proposal we plan to continue this work and deploy this compression algorithm in a proof-of-principle yet realistic emulation of the trigger system using raw detector data. This is a future-looking study targeting HL-LHC, but if the results are ready to be deployed during the course of Run-3 we will use them for compression of the data recorded with techniques in WP1.

As the trigger system is crucial for this research program and for the overall ATLAS physics output, the Lund group will maintain a leading role in its operations. Throughout the course of this project, the team members will be involved in the development and monitoring of trigger software. This is a particularly crucial responsibility during the run up to first data taking, after the complete overhaul of the ATLAS software. The Lund group has already taken responsibility roles in the trigger, with the StG postdoc William Kalderon (now a postdoctoral fellow at BNL), who has served for two years as the convenor of the jet trigger and has now moved on to be overall ATLAS trigger menu coordinator.

1.3.2 Objectives of WP2

The overall objective of WP2 is the commissioning of the new trigger techniques using early LHC data, and their validation using well-established physics observables. These observables will be used to evaluate the performance of the calibration techniques and take corrective measures if needed, still to be implemented in time for the production phase of the LHC.

The first focus of WP2 is the **determination of the performance of HLT jets and photons, and of electrons and muons** at a later date. Together with our collaborators in ATLAS, we will evaluate the energy scale and the uncertainty of HLT objects using early data and simulation. I am an expert in those topics, having derived the very first iteration of the energy scale uncertainty [35] and having supervised a number of students on this topic since. Within my StG, we measured the performance of both offline and trigger jets in early Run-2 data using the same metrics and techniques, and we intend to do the same in Run-3 for an impact on all ATLAS early measurements and searches.

We will use **early data to commission the new Run-3 trigger software** in the HLT dijet mass spectrum. The experience at the beginning of Run-2, where my collaborators and I observed a L1 trigger misconfiguration in early jet data that limited measurements and searches [37], showed that this is a mandatory validation step to take after upgrades. Subsequently, we will repeat this process for PEB muons and electrons, using the dimuon and dielectron mass peaks from the decay of standard candles (Z and J/ψ) **[10]**.

If the sensitivity of mediator searches in the dijet mass spectrum improves with respect to Run-2 results, we will publish dijet and dijet+ISR searches using early data. As an example, if the LHC delivers more than ≈ 18/fb of low-luminosity runs, dijet searches can profit from much lower jet trigger thresholds as shown in Fig. 4 and surpass sensitivity with respect to current searches[5].

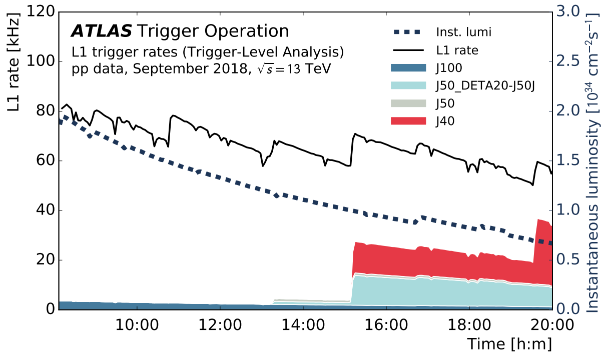

Figure 4: In Run-2, the underutilization of HLT resources at low LHC instantaneous luminosity allowed my team and collaborators to design TLA L1 triggers with reduced p_{T} thresholds [35].

Another important component of this proposal developed in WP2 is **reproducible end-to-end analysis software for dijet and dijet+ISR DM mediator searches [12]**. We will use the RECAST framework to package the entire software stack used for early data analyses so that many of the steps can be executed automatically rather than manually. Using RECAST at this early stage has three advantages. Firstly, it makes intermediate testing and monitoring much faster, as this code can be executed on new data and compared to already-tested data. Secondly, it shortens the time required to go from calibrated data to final plots and analysis results in the analysis iteration with the full LHC dataset in WP3. Thirdly, it can be used as back-end to the REANA framework [35]. so that theorists and members of the DM community can test their own signals, as discussed in WP5.

1.3.3 Objectives of WP3

The overall objective of WP3 is to use Run-3 LHC data recorded with the TLA technique to search for new resonances in the dijet mass spectrum, motivated by DM mediators. In this proposal we will focus on the decays of the mediators to light quarks, but the work in WP1 also enables dedicated searches for mediators decaying preferentially to heavy quarks.

The main innovation in this WP is the **dijet+ISR photon search using the TLA technique**, alongside the gluon ISR signature. The two channels are complementary: the jet ISR channel is more sensitive due to higher signal rates, but selecting events using photon ISR can reach lower mediator masses. Both channels will be much more sensitive than current Run-2 traditional searches, since the threshold on the associated object is lowered from ≈ 400 GeV to ≈ 220 GeV and from ≈ 150 GeV to ≈ 40 GeV for jet and photon cases respectively. In turn, this increases the signal acceptance of more than one order of magnitude for a mediator mass of 250 GeV [6], and it lowers the minimum mediator mass to which these searches are sensitive, as shown e.g. in Ref. **??**.We will also maintain our involvement in the **full Run-2 + Run-3 dataset dijet TLA** towards a legacy TLA publication covering both LHC runs.

1.3.4 Objectives of WP4

The objective of WP4 is to use Run-3 LHC data recorded with the TLA+PEB technique for dark QCD searches. The most promising parameter space to target will be determined by discussing the theory and broader DM community (WP5), and by using jet measurements to constrain dark QCD phenomenology. The rejection of QCD background at the HLT will be crucial for these searches, as it will allow recording higher signal rates in TLA+PEB events. The **search for semi-visible jets** will be performed first, adapting calibration and performance studies in WP1 and WP2 and analysis techniques in WP3 to this signature. Subsequently, we will augment the standard reconstruction techniques in WP2 to correctly identify anomalous content (e.g. leptons in jets) needed for the **composite jet search**. An additional advantage of the PEB+TLA data stream selected for these searches is that it enables searches for a wide variety of non-standard jet topologies (e.g. photon-jets [35], jets containing long-lived particles [35]) at a later date. This will have an

[5] The analysis of NN/fb of low-threshold Run-2 data is undergoing as part of my VR project grant, see Funding ID in Part B1

[6] The background also increases, but since it is estimated using data-driven techniques as explained in **??** its increase can be managed without a significant loss in sensitivity.

impact especially after the end of this proposal when the long LHC shutdown is foreseen before HL-LHC and data already taken will be analyzed in further detail.

1.3.5 Objectives of WP5

The overall objective of WP5 is to connect the results in this proposal to the broader experimental community, with a focus on the search for DM, to maximize their impact. This WP naturally includes the contextualization, communication and dissemination of results, in terms of tools that can enhance the physics potential of experiments and of DM discoveries or constraints. WP5 also covers my work within synergistic initiatives involving LHC experiments and the broader DM search communities that directly benefit the objectives of WP3 and WP4, started with the Dark Matter Forum and in the context of the update of the European Strategy of Particle Physics. Collectively furthering the DM community's understanding of the theoretical and experimental landscape for WIMP and dark QCD models will sharpen the LHC search targets. WP5 spans the entire course of this project, and has four goals.

Firstly, we will collaborate with theory experts in Lund, Heidelberg and worldwide to **identify the most promising parameter space for the searches in WP4**, in order to develop trigger chains that can best target it **[18]**. Concretely, this will take place through shared supervision of PhD students and dedicated workshops in Lund, similar to the 2019 Dark Jets workshop [35]. These workshops will allow cross-talk with the community interested in comparing and contrasting regular and dark QCD, leading to concrete improvements for MC generation.

Secondly, I will **continue co-organizing the Initiative for DM in Europe and beyond (iDMEu)**, a cross-community effort focused on DM that includes the nuclear physics, astroparticle physics and particle physics. This initiative has more than 200 endorsers at the time of writing, and will have its first kick-off meeting in early Summer 2020. It has the ambition of becoming a permanent forum so that all different communities can identify opportunities to work together and exploit synergies and complementarities. Practical outcomes of work within this initiative are summary plots with commonly used benchmarks that enhance the complementarity of the LHC results in WP3 and WP4 with other experiments and astrophysical observations **[19]**. iDMEu also includes a component of communication to the general public, also part of WP5 **[20]**.

Thirdly, we will **disseminate the technical outcomes of WP1 to other experiments at the LHC and beyond**. This will be achieved via technical peer-reviewed papers that include links to prototype open source software implementations as auxiliary materials and presentation in conferences. I have recently become part of the coordination team of the cross-experiment HEP Software Foundation (HSF). With its goal of helping experiments meeting the challenges posed by new experimental programmes for HL-LHC, the HSF is an ideal platform to connect the solutions in this proposal to experimental needs **[21]**.

The fourth goal of WP5 is to **make results and data from WP3 and WP4 as accessible as possible**, both inside and outside the ATLAS collaboration. We will publish the final analysis likelihoods from WP2-4 and implement the end-to-end RECAST analyses in the REANA hub [35], so that they can be scrutinized and reproduced without a need to directly access ATLAS data (which will only be available in open format at a later date) **[22]**. These analyses will become part of a broader effort the cross-experiment *Virtual DM environment* for end-to-end data analysis within the European Science Cluster of Astronomy and Particle Physics ESFRI research infrastructure (ESCAPE), which I will start leading at the beginning of 2020. All software implementations within this project will become part of the ESCAPE Software Catalogue.

2 Methodology

The work within REALDARK proposal builds novel data acquisition and selection tools and techniques in WP1 and applies them to obtain the physics results in WP2-4. This modus operandi follows my proven track record of bringing challenging improvements to the data taking system and to the combined performance of the entire experiment, and successfully exploiting them for world-best physics measurements and searches. Sec. 2.1 and Sec. 2.3 describe the common methodologies required for reaching the objectives of the WPs described in Sec. 1.3. Their description follows logically the procedures designed to record, treat and analyze large amounts of raw data towards the dissemination of results that can shed light on the particle nature of DM. Specific methodologies are described in Sec. 2.2. How all these methods fit together in the project planning is described in Sec. **??**.

2.1 Methodologies: technical tools

2.1.1 Common methodologies: From LHC collisions to physics objects for analysis

In this section outlines the methods used to record data with the techniques proposed, and transform these data into physics quantities to be used for the searches in WP2-4.

From LHC collisions to storage The ATLAS core software infrastructure receives raw trigger and detector information from collisions and assembles it into different data streams. Full events that passed the selections determined by the ATLAS trigger menu are placed in the `physics_Main` stream for subsequent reconstruction, while partially-built detector data (including TLA and TLA+PEB objects) need to be selected by dedicated software algorithms and placed in specific data streams. In WP1, the team will rely on my expertise as one of the main authors of the TLA framework in Run-2, and deploy new multithreaded software algorithms that can write out jets, photons, muons and electrons, as well as user-defined partial detector information. These algorithms are executed if an event is selected by a given L1+HLT trigger "*seed*" chain. Given that a large part of the searches in this proposal occurs at the HLT, a resource constrained environment, an important part of the planning prior to data taking is to optimize the CPU and storage resources needed for each of the planned data streams [35], and determine the optimal stream content and seed chains for the searches in WP2-4.

Reconstruction Full events are passed through the ATLAS reconstruction software, so that raw detector data can be reconstructed into final physics objects (muons, electrons, jets...). While TLA events are already reconstructed within the HLT software using inputs and algorithms that are as close as possible to offline, TLA+PEB events require the offline software reconstruction algorithms to be adapted, to cope with receiving partial detector data as input. We can take advantage of trigger-level reconstruction algorithms, where regional reconstruction has already been used to optimize HLT resources [35], and extend existing reconstruction and identification techniques to non-isolated muons and electrons in hadronic environments [35].

Object identification and calibration Correct object identification and calibration is a crucial point to be able to use HLT and custom objects for physics analysis. I am an expert of jet calibration In the case of TLA, these can be improved at a later date using dedicated algorithms and more refined corrections that are not applied at the HLT. Relying on the success of HLT jet calibration in my StG, we will keep the procedures to identify and calibrate HLT objects in this proposal as close as possible to those of offline objects, with additional data-derived corrections and scale-factors applied to HLT objects covering the remaining differences. A major improvement in the Run-3 calibration procedures is the addition of tracking information computed by the HLT software, as described in Sec. 1.2. This allows for improvements in both jet reconstruction and calibration, as *particle flow* [35] jets combining calorimeter and tracking information can be used at the trigger level, and the same track-based correction steps can be applied to HLT and offline jets [35]. The use of tracks is not yet planned for HLT photons, electrons and muons, and will be investigated in WP1. For example, the use of tracks could improve the rejection of fake photons and facilitate identification and correct calibration for converted photons in TLA. Preliminary studies that I supervised [35] show that the HLT photon calibration in Run-2 was already well-understood without further optimizations in the energy range used for the searches in WP3 [7]. We expect a similar situation in Run-3 for photons and electrons alike, given that the HLT ATLAS electron and photon group plans to align reconstruction, identification and calibration as much as possible, and that residual differences (e.g. from different online/offline calorimeter corrections) can be covered with data-derived factors e.g. from analysis of $Z \to ee(\gamma)$ events. Identification and calibration of HLT muons will also rely on the offline procedures (reaching muon momenta as low as 3 GeV [35]) and additional corrections from $Z \to /J/\psi \to \mu\mu$ events.

Performance evaluation Once the reconstruction and calibration algorithms are in place, their performance needs to be evaluated in data using figures of merit for probe objects such as response and resolution, tested against well-measured references. In WP2, we will prepare a comprehensive software toolkit for quantitative comparisons between physics objects reconstructed in TLA, TLA+PEB and offline, using automated problem detection techniques. This toolkit will be deployed online, so that HLT problems can be corrected as early as possible, and used offline to determine the performance of HLT objects. The extent to which a mismatch between tested object and reference objects requires intervention (e.g. modification of the reconstruction software, re-derivation of calibration constants...) varies depending on how this affects the sensitivity of WP3 and WP4 searches, estimated case by case as discussed in

[7] We expect the situation to improve in Run-3, especially since the preliminary studies did not subtract any of the QCD background that can be removed using a combination of calorimeter variables (e.g. as in **??**)

Sec. ??.

2.2 Specific methodologies for WP1: data compression

Compressing data can lead to further gains in terms of reducing the storage requirements of TLA and TLA+PEB. For this reason, this proposal investigates using deep autoencoders for lossy data compression, in a forward-looking study targeting HL-LHC. Preliminary studies by a Lund engineering (LTH) student that I co-supervised have shown that compressing data using deep autoencoders could lead to large space savings for TLA data and ATLAS data and simulation as a whole, leading to compression factors of 2 with negligible performance loss in TLA data [35]. The advantage of autoencoders over methods such as Principal Component Analysis is that they are able to learn non-linear correlations among the nodes, and that, once the network is trained, the compression is fast (order 2-5 μs), making them a good candidate to be used in the HLT (latency $\approx$ 300 μ s).

2.3 Common methodologies: analysis and interpretation tools

This section outlines the data analysis and reinterpretation methods that are common to the searches in WP2-4, mentioning specific details for each of the searches.

Sensitivity studies Prior to starting any of the physics analyses in WP2-4, studies are needed to determine the sensitivity of the analysis and refine the parameter space to be targeted (e.g. for producing fully simulated signal grids). These studies are implemented as a simplified version of the analysis, where trigger, detector and calibration effects are implemented using parameterizations and only employ full simulation wherever needed. For the implementation of the analyses in this proposal, we will use the RIVET software [35]as a starting point to analyze generator-level signal and background simulation. The advantage of using RIVET is that it easily interfaces with the CONTUR software [35]which allows to use a large number of LHC measurements to constrain new phenomena. This is particularly crucial for the searches in WP4, since measurements of jet fragmentation can already constrain dark QCD showering parameters. Sensitivity studies also provide a first testing ground for analysis techniques prior to data taking, and can be used during the data processing stage to understand the level of precision needed for HLT objects in order to be sensitive to certain kinds of signals.

Background reduction and estimate Standard Model backgrounds to all searches in WP2-WP4 need to be reduced and estimated.

The background reduction techniques employed for the searches in WP2-4 select events based on one or more variables that discriminate signal from background. These selections will be initially determined in simulation samples, and verified in "validation regions" with limited amounts of signal contamination. Simplicity is a driving factor in the choice of background reduction techniques in this proposal. Simpler selections allow searches to remain relatively agnostic to the specific model details, and ease the reinterpretation of our results using models others than the ones we employed, enabling the dissemination objectives in WP5. In the searches in WP2 and WP3, we will use the angular distribution of the jets as a discriminant between s-channel processes over the QCD backgrounds. In the semi-visible and composite jets searches in WP4, we will apply loose selection to reduce the QCD background already at the HLT. We will investigate the use and CPU cost of particle flow jet properties [35], event shape variables [35], jet substructure techniques and variables [35], as well as new dedicated theoretical and phenomenological metrics for measuring QCD jet showering and identifying anomalies [35]. To complement this *cut-based* approach and to further increase the generality of the searches in WP4, exploratory studies will be performed in the direction of detecting anomalies over the known QCD background using ML techniques. This will benefit from the work done within my VR Project Grant, which focuses on unsupervised search strategies specifically targeting resonance searches.

The estimation of the backgrounds, first tested in simulated data, will be derived directly from data wherever possible. The background of the dijet and dijet+ISR TLA searches in WP2 and WP3 will be derived directly from data, since the enormous number of events collected excludes the possibility of generating and simulating an equivalent amount, and because QCD simulations are subject to large theoretical and modeling uncertainties. For the semi-visible and lepton-in-jets searches in WP4, we will use a combination of data-derived and simulation-driven backgrounds, depending on whether the backgrounds are mainly from QCD (e.g. fake leptons, jets containing heavy flavor quarks decaying leptonically) or from QED and weak interactions (e.g. Z/W+jets where the vector boson decays into leptons and neutrinos, boosted top quarks containing a lepton from the W decay). Three background estimation methods are foreseen for the searches in this project:

- For searches in WP2 and WP3 will estimate the smoothly falling QCD background with a functional form or with a continuous parameterization that does not accommodate narrow local excesses such as those from a new resonant particle. The experience gained with the high-statistics Run-2 dijet TLA search within my StG will be crucial to choose the appropriate method among a number of available techniques [35].
- For searches in WP4, we will use the ABCD method (e.g. Ref. [35]) to estimate the background events in the signal region from normalization and control regions that have a minimal background contamination, defined in terms of independent variables.
- For the smaller W/Z backgrounds in the WP4 searches, we will make use of MC simulation using normalization factors derived from control regions in data.

The robustness of the background estimation techniques will be evaluated in partial data scaled to the full expected dataset, using signal injection studies.

Cross-checks of performance within the analysis selection The performance of the physics objects in the search will be monitored using the WP2 toolkit, after applying the background reduction selection.

2.3.1 Statistical analysis

Analyses in this project will be blinded, meaning that the region of phase space or the distribution where a signal is expected is not analysed until the full analysis strategy has been verified from data. We will use control regions and limited amounts of data where a signal would not have a significant impact to disentangle performance issues. Before unblinding the signal region, we will finalize background estimation, systematic uncertainties and expected limits satisfactorily, according to an internal peer-review committee. Once this is done, a statistical test will be employed to compare the observed data and estimated background. In this phase, we will also incorporate the systematic uncertainties from the reconstruction and calibration and on the background estimation techniques. Given the data-driven nature of the background estimation employed, the systematic uncertainties will mostly impact the signal yields and therefore have a limited impact on the search sensitivity. The statistical compatibility of background and data can then be extracted from either the CLs technique [35], or from tests that are well suited for looking for a resonance or for an excess in the tails of the distribution, such as the BumpHunter and TailHunter [35]. We will use a statistical toolkit that enables sharing the final likelihood function, e.g. pyhf citeToBeCited.

2.3.2 Interpretation and dissemination of results

The outcome of the statistical analysis will be either compatibility between the data and the background estimation, or an excess above the background. In both cases, the implementation of the analysis code will be available via RIVET/RECAST/REANA and within the ESCAPE software catalogue, and the digitized results will be made available on HEPData [35].

Null result: The compatibility of data and background indicates that no new phenomena are found within the dataset used for the search. Therefore, constraints can be set on benchmark models in the limit setting phase. Frequentist or Bayesian techniques [35] can be used to set limits on the maximum allowed cross-section of the benchmark process versus the model parameters, e.g. the mass of new resonances. To extend the results of these searches to any resonant particle, we will also set limits on generic Gaussian-shaped resonances of different widths as done in all the dijet searches in my StG, and a responsibility of one of the students for the Run-2 dijet+ISR search [35]. The outcomes of these searches will be plotted alongside results of other LHC searches and non-LHC experiments within a common theoretical framework. For WIMP searches, this will initially follow the methodology of Ref. [35], and will evolve following input from the broader community (e.g. from dedicated discussion within workshops and iDMEU).

Discovery: If an excess above the background prediction is observed, it will need to be characterized. Following the procedure for blind analyses, the observation will be made public as soon as the excess has been reviewed by the entire ATLAS Collaboration. It will be cross-correlated with other ATLAS searches which might see hints of the same process in different final states. Identifying the source of the excess and the nature of the resonance (e.g. distinguishing its parton content [35]) will become an outmost priority for the team [8] and for the Collaboration. In this case, the results of astrophysics and non-LHC experiments will be crucial to ascertain whether these results can be attributed to processes involving DM.

[8] This may lead to changes in the and work sharing of this project, as the entire team will be working first-hand with the tools and datasets employed up to that point for a result that has a world-wide impact on the field.

2.4 Project planning

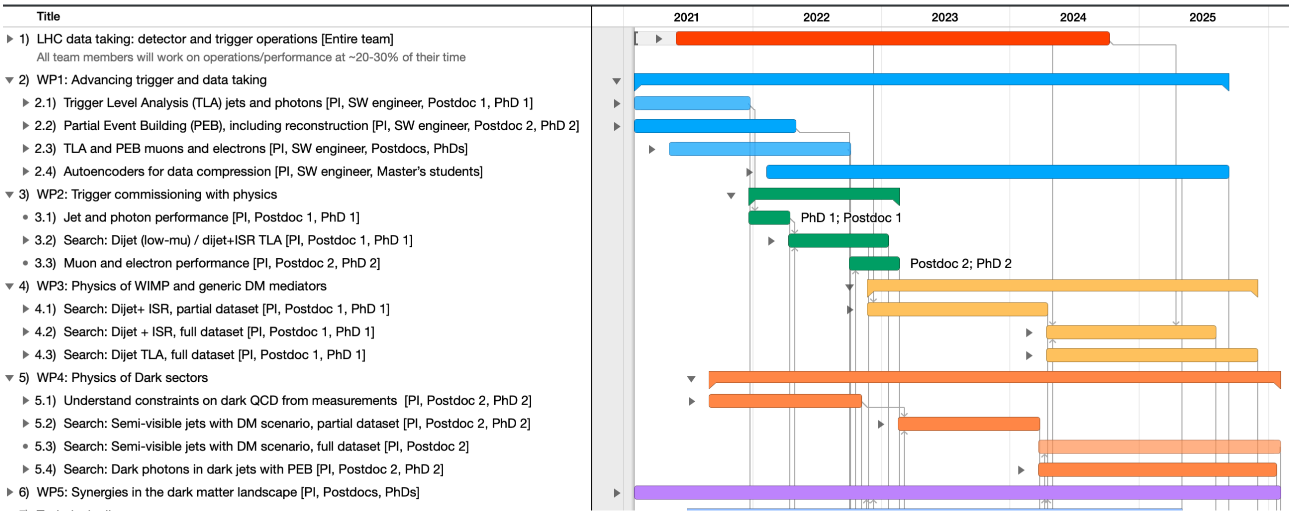

Figure 5: Sketch of project planning and division of work across team members.

The research team of REALDARK that I will lead as a PI is composed by two postdoctoral researchers, two PhD students and one software engineer. The postdoctoral researchers (PD1 and PD2) will be employed for the entirety of the project. The PhD students (PhD1 and PhD2) will work within REALDARK for their 4-year thesis period. Each of the postdocs will work with one of the PhD students on the two physics objectives of WP3 and WP4 respectively, and related technical and commissioning tasks in WP1 and WP2. As PI, I will oversee all tasks in the project with day-to-day supervision and regular meetings, and have a hands-on involvement in more complex tasks as specified below. The software engineer will have a two-year contract at the beginning of the project to ensure the feasibility of the more software-intensive tasks in WP1. Fig. 5 provides a project planning schedule, including the breakdown of the work between the team members and intermediate milestones and deliverables marked with **[N]**. It has been prepared using the OmniPlan software **??**,taking into account the interdependencies between the WPs and scheduling the work allocation in a way that ensures the feasibility of this ambitious project without overcommitting the team members.

2.5 WP1: Real-time analysis and data compression in ATLAS

TLA software implementation and calibration of trigger objects As a preliminary input to this proposal, I have started working on a new Run-3 multithreaded HLT software algorithm that selects and writes out partially-built physics objects in the TLA stream. This algorithm will be more flexible than its Run-2 counterpart, and can handle any physics objects. By the start of this proposal, we will have preliminary results of the use of this algorithm on TLA jets. In 2021, PI, PD1, PhD1 and the software engineer will work on the readiness of the core software for photons and jets, on commissioning with cosmic data, and on their subsequent calibration **[1]** The experience gained with photons will be used by P1 and PhD1 to implement and commission TLA electrons as well **[2]**, while PD2 and PhD2 will focus on TLA muons **[3]**. Initially, the TLA stream will contain both jets and photons to enable analyses in WP3 in a simple manner. As part of this work, we will contribute to the design and implementation of seed trigger chains and data streams that selectively record one or more TLA objects simultaneously, optimizing according to the use cases in WP3 and WP4. The finalized TLA software streams will be ready in Q3 2022, allowing sufficient time for validation with data prior to the LHC production period.
PEB software implementation and reconstruction Prior to the implementation of TLA muons (Q1-Q2 2021), PD2 and PhD2 will deploy the core software to write out user-defined regions of the detector together with TLA objects, taking into account the optimization between CPU and storage costs mentioned in Sec. 2.1.1. Subsequently, they will test the TLA+PEB implementation on early data and work together with the software engineer on the implementation of dedicated reconstruction algorithms for partial detector input, taking advantage of the experience gained with TLA electrons and muons **[4]**. At the end of this work, expected for Q3 2022, we will publish a technical paper describing the combined TLA and PEB implementation within the ATLAS trigger (Q4 2022) **[5]**.
Identification and calibration of TLA jets, photons, electrons and muons Throughout 2021 and 2022, PhDs and postdocs will contribute to the algorithms for the identification and calibration of the HLT physics objects that they have implemented in TLA and PEB. In this proposal, we will focus on new HLT pile-up suppression techniques using calorimeter information against techniques using tracking information. The performance of this calibration will

be studied in WP2 and documented in the technical papers in **[5]** and **[10]**.
Together with the performance studies, this work will be part of the students' ATLAS authorship qualification task.
Data compression with autoencoders Throughout the course of this proposal, the software engineer and I will supervise Master's students that study the performance, implementation and characteristics of autoencoder networks used to compress ATLAS data. The outcome of this work will be a test implementation in system that emulates a HLT computing node, preparing the ground for an implementation in the HL-LHC trigger and computing system **[7]**. This work will be documented in technical publications, detailing preliminary and final work **[6,8]**.

2.6 WP2: Commissioning the upgraded ATLAS trigger with physics

HLT object performance Over the course of the early LHC data taking (2021-2022), PD1 and PhD1 will determine the performance of HLT jets and photons with early data, while PD2 and PhD2 will focus on the performance of HLT and PEB electrons and muons.
Searches and measurements of the dijet mass spectrum After the trigger jet and photon performance is well understood, PD1 and PhD1 and I will measure the dijet mass spectrum using a limited amount of data. We will select inclusive dijet events, and events where dijets produced in association with a photon, first using offline jets and TLA jets, and then using TLA+PEB jets, for comparison with simulation and Run-2 data. This will lead to a prototype of the dijet and dijet+ISR TLA searches within the RECAST framework **[9]**. PD2 and PhD2 will repeat this exercise with electrons and muons in a measurement of the Z and J/ψ peaks, to be included with the jet and photon performance work in a dedicated technical publication **[10]**. If the LHC provides a dataset that surpasses the Run-2 sensitivity for either dijet and dijet+ISR searches, PD1, PhD1 and I will work with ATLAS collaborators towards publication of the result of these searches in Q2 2023 **[11]**.

2.7 WP3: Dark matter mediator searches

WIMP mediator searches and the interpretation of their results will be the main physics focus of PD1 and PhD1. **Dijet+ISR TLA search** This search relies on calibrations and data analysis code developed and tested in WP1 and WP2. PD1 and PhD1 will refine the WP1 jet calibration for a high-statistics search (e.g. for pile-up suppression suppression techniques for the low-p_{T} resonance jets) and repeat the performance studies from WP2 with a larger dataset. PD1 and PhD1 will take advantage of WP2 RECAST implementation of the dijet+ISR mass spectrum to estimate backgrounds and produce inputs to run the statistical analysis. PD1 and PhD1 will collaborate with colleagues from OSU, Oregon and other ATLAS institutes for two publications, one using the first part of the dataset where the Run-2 sensitivity is surpassed **[12]** in 2024, and one using the full Run-2 dataset **[13]** in Q2 2025.
Dijet TLA search PD1 and I will remain involved in the full ATLAS Run-3 TLA with an advisory role. We will provide the WP2 RECAST implementation, and combine Run-2 and Run-3 results for a legacy TLA publication covering both LHC runs in Q3 2025 **[14]**. Dijet searches are sensitive to a variety of new physics signals (see Sec. 1.1): code and results from WP3 searches will be input to WP5 for broad dissemination.

2.8 WP4: Dark QCD searches

Dark QCD searches and the interpretation of their results will be the main physics focus of PD2 and PhD2. Both searches in WP4 will require preliminary work and collaborations to decide on the optimal parameter space to target, described in WP5 below. This will be done within the first and second year of this proposal and culminate in a public document (ATLAS PUB note) on the reinterpretation of measurements and existing searches for a variety of dark QCD benchmark models in Q3 2022. This will inform the configuration for the TLA+PEB stream to be implemented upon the LHC restart in Q1-Q2 2023, together with the preliminary studies on the suppression of QCD background to implement at the HLT.
Semi-visible jet search This search will be the first to be tackled. PD2 and PhD2 will bring their expertise to search for low-mass particles decaying into semi-visible jets in TLA+PEB, and work with colleagues from the University of Witswatersrand who will focus on the high-mass region with traditional techniques. This continues an existing collaboration started within my StG, which we expect to be supported by the ERC "Implementing Agreements" program during 2020. We will publish this search with an intermediate dataset in Q2 2024 **15**, and continue our involvement in the full-dataset search with an advisory role.
Composite jet search This shares many of the common PEB+TLA performance tools with the semi-visible jet search and proceeds in parallel during 2023. The bulk of the work will be done during 2024, first adapting WP1 and WP2 code and techniques for leptons within jets and then running the analysis chain, leading to a publication in Q4 2025 **16**. Throughout this period we will also work to make the TLA+PEB stream and the analysis tools available within ATLAS

for other dark sector searches.

2.9 WP5: Interpretation, dissemination and synergies

The work in WP5 is spread over all the timeline of this project and its hands-on components will involve the PI, the postdocs and the PhDs. A **two-weeks workshop** with other non-LHC communities, involving experts and authors of the theory benchmarks for searches in this proposal who agreed to participate [35] will be organized in Q3 2021 within the iDMEu initiative. The workshop will be hosted at Lund for the first week and at CERN for the second week (with videoconference available in both), to discuss the state-of-the-art of dark QCD theory and experiment and understand current constraints including dark meson searches, non-collider and astrophysical constraints.
Together with each technical and physics publication we will share our results on HEPData and on plots that summarize searches from different techniques and experiments constraining the same model. Within iDMEu, we will also extend the benchmarks for these plots to dark QCD models, with a first iteration expected after the Lund/CERN workshop.
We will also make common-interest software in WP1 (e.g. compression, non-standard object reconstruction) and WP2-4 (RECAST analyses in REANA) available in the ESCAPE Software Catalogue as part of the DM Virtual Environment, and on the HSF webpages. This will include documentation and working examples for enhancing the usability and impact of the shared software.
Finally, I will write a review of the state of the art of LHC searches at colliders at the end of this project **[17]**.

References

[1] L. Evans and P. Bryant, JINST **3** (2008) S08001.
[2] G. Bertone and D. Hooper, Rev. Mod. Phys. **90** (2018), no. 4 045002, arXiv:1605.04909.
[3] N. Bernal *et al.*, Int. J. Mod. Phys. **A32** (2017), no. 27 1730023, arXiv:1706.07442.
[4] G. Steigman, B. Dasgupta, and J. F. Beacom, Phys. Rev. **D86** (2012) 023506, arXiv:1204.3622.
[5] A. Boveia and C. Doglioni, Ann. Rev. Nucl. Part. Sci. **68** (2018) 429, arXiv:1810.12238.
[6] L. J. Hall, K. Jedamzik, J. March-Russell, and S. M. West, JHEP **03** (2010) 080, arXiv:0911.1120.
[7] S. S. McGaugh, F. Lelli, and J. M. Schombert, Physical Review Letters **117** (2016) .
[8] O. Lennon, J. March-Russell, R. Petrossian-Byrne, and H. Tillim, JCAP **1804** (2018), no. 04 009, arXiv:1712.07664.
[9] S. Bird *et al.*, Phys. Rev. Lett. **116** (2016), no. 20 201301, arXiv:1603.00464.
[10] D. J. E. Marsh, Phys. Rept. **643** (2016) 1, arXiv:1510.07633.
[11] S. P. Martin, Adv. Ser. Direct. High Energy Phys **18** (1997) 1, arXiv:hep-ph/9709356.
[12] J. Goodman *et al.*, Phys. Rev. **D82** (2010) 116010, arXiv:1008.1783.
[13] LHC Dark Matter Forum, D. Abercrombie *et al.*, arXiv:1507.00966.
[14] I. M. Shoemaker and L. Vecchi, Phys. Rev. **D86** (2012) 015023, arXiv:1112.5457.
[15] O. Buchmueller, M. J. Dolan, and C. McCabe, JHEP **01** (2014) 025, arXiv:1308.6799.
[16] M. Chala *et al.*, JHEP **07** (2015) 089, arXiv:1503.05916.
[17] M. R. Buckley, D. Feld, and D. Goncalves, Phys. Rev. **D91** (2015) 015017, arXiv:1410.6497.
[18] D. Egana-Ugrinovic, S. Homiller, and P. R. Meade, Phys. Rev. **D100** (2019), no. 11 115041, arXiv:1908.11376.
[19] LHC Dark Matter Working Group, T. Abe *et al.*, Phys. Dark Univ. (2019) 100351, arXiv:1810.09420.
[20] Y.-J. Kang and H. M. Lee, arXiv:2001.04868.
[21] J. Ellis, M. Fairbairn, and P. Tunney, arXiv:1807.02503.
[22] R. K. Ellis *et al.*, arXiv:1910.11775.
[23] Snowmass 2013 Cosmic Frontier Working Groups 1–4, D. Bauer *et al.*, Phys. Dark Univ. **7-8** (2015) 16, arXiv:1305.1605.
[24] G. Busoni *et al.*, Phys. Dark Univ. (2019) 100365, arXiv:1603.04156.
[25] M. J. Strassler and K. M. Zurek, Phys. Lett. **B651** (2007) 374, arXiv:hep-ph/0604261.
[26] B. Holdom, Phys. Lett. **166B** (1986) 196.
[27] D. Curtin, R. Essig, S. Gori, and J. Shelton, JHEP **02** (2015) 157, arXiv:1412.0018.
[28] K. M. Zurek, Phys. Rept. **537** (2014) 91, arXiv:1308.0338.
[29] C. Cheung, J. T. Ruderman, L.-T. Wang, and I. Yavin, JHEP **04** (2010) 116, arXiv:0909.0290.
[30] S. Tulin and H.-B. Yu, Phys. Rept. **730** (2018) 1, arXiv:1705.02358.
[31] E. Bernreuther, F. Kahlhoefer, M. Krämer, and P. Tunney, Submitted to: J. High Energy Phys. (2019) arXiv:1907.04346.
[32] T. Cohen, M. Lisanti, H. K. Lou, and S. Mishra-Sharma, JHEP **11** (2017) 196, arXiv:1707.05326.
[33] M. Park and M. Zhang, Phys. Rev. **D100** (2019), no. 11 115009, arXiv:1712.09279.
[34] J. Beacham *et al.*, J. Phys. **G47** (2020), no. 1 010501, arXiv:1901.09966.
[35] CERN, https://cern.ch, 2016. [Online; accessed 21-January-2019].
[36] https://indico.cern.ch/e/iDMEu, 2019. [Online; accessed 7-January-2019].
[37] A. Collaboration, ATLAS-COM-CONF-2015-042, CERN, Geneva, Jul, 2015.