a. Extended Synopsis

Overview and Objectives

Current theories of galaxy formation and evolution strongly suggest that central, supermassive black holes (SMBHs) have a profound effect on the galaxies that they live in (e.g., Kormendy and Ho, 2013). This is not surprising since the potential energy associated with mass accretion onto a supermassive black hole is comparable to that generated via the nuclear fusion in the galaxy's stars (see e.g. Fabian, 2012). Thus when a galaxy goes through a "quasar phase" (where gas is supplied and accreted by the SMBH) there is ample energy to potentially impact the host galaxy and the surrounding intergalactic medium.

However, the critical details of the physical processes involved in how this energy escapes the inner most regions of the galaxy and then interacts with the gas, dust, stars and dark matter, is currently poorly understood. This fact, along with current observational data giving more puzzles than clues is preventing the field from moving forward. Significant further issues arise since startling new observations from my (Nicholas P. Ross; NPR) research team (MacLeod et al., 2016; Ross et al., 2018) show that *quasars vary significantly on timescales of weeks to months*, whereas the accretion disks (that supply 'fuel' for the quasar) should take thousands of years to change their optical emission; this has recently been called the "Quasar Viscosity Crisis" (e.g., Lawrence, 2018). Thus, it is unclear if we have an understanding of a physical phenomena prevalent in many astrophysical systems: the accretion disk.

The field of observational extragalactic astrophysics is poised for a fundamental and rapid change. Starting in late 2019, a fleet of new telescopes, instruments and missions will be commissioned, start data taking, and will leap-frog the quality and quantity of data we have available today. These surveys and missions include: the fifth incarnation of the Sloan Digital Sky Survey (SDSS-V[1]); the Large Synoptic Survey Telescope (LSST[2]); the Dark Energy Spectroscopic Instrument (DESI[3]) survey; the 4-metre Multi-Object Spectroscopic Telescope (4MOST[4]) survey, and the ESA *Euclid* mission[5]. Even more imminent is the launch of the *James Webb Space Telescope* (JWST[6]).

Quasars in the 4th Dimension (Q4D) has two broad and well-posed goals. First, we aim to elucidate in detail **how the energy directly associated with a supermassive black holes impacts the universal galaxy population.** We will gain a deep understanding into the physical mechanisms related to central engine black holes; their accretion disk physics, their dynamics on both human and galactic timescales and the role they might play in forming and regulating the galaxy population. Second, we anticipate **the discovery of brand new extragalactic phenomena.** By tapping into the massive raw discovery space that the new experiments will open up, there is the highly likely outcome of discovering something "brand new" (Ivezic and Tyson, 2008; LSST Science Collaborations et al., 2009), e.g. the EM counterparts to mergers of Binary SMBHs (with their associated gravitational wave chirp and ringdown), objects potentially similar to repeating fast radio bursts Spitler et al. (2016) or more objects akin to the still unexplained 'SCP 06F6' (Barbary et al., 2009). Q4D objectives are:

1. Characterize the variable extragalactic universe and quasar population.
2. Establish the energy transport mechanisms associated with the"quasar phase", and explain the relation between accretion rate, black hole mass build-up with observed light curve and spectral properties.
3. Develop and then link theoretical accretion and galaxy formation models for a fully holistic theory of active galaxies.
4. Discover new extragalactic variable objects.

We will achieve this by leveraging several of the new, large-scale surveys that will all have 'First Light' over the lifetime of the project. These critical observations are made by exploiting the large imaging and spectroscopic

[1] www.sdss.org/future/ [2] lsst.org [3] desi.lbl.gov [4] 4most.eu [5] sci.esa.int/euclid/ [6] jwst.stsci.edu

datasets that will be available from the SDSS-V, DESI, 4MOST, LSST and ESA *Euclid*. **Crucially, although these projects individually will deliver new state-of-the-art datasets, it is my project that will be the first to break down the associated data silos and combine these data in order to go beyond the state-of-the-art.**

1. Current State of the Art.

The current state-of-the-art data samples have either $\approx 10^6$ quasars with one spectral epoch, or only a few objects with repeat photometric data, i.e. light-curve information and the accompanying repeat spectra (see Figure 1). NPR has been involved in the production of both of these two types of samples (MacLeod et al., 2016; Pâris et al., 2017). We plan to collate datasets so that the 10^6 sample have *high-fidelity light-curves and ample repeat spectroscopy (necessary of emission/absorption line diagnostics*, and in doing so will kick start the new field of Variable Extragalactic Astrophysics.

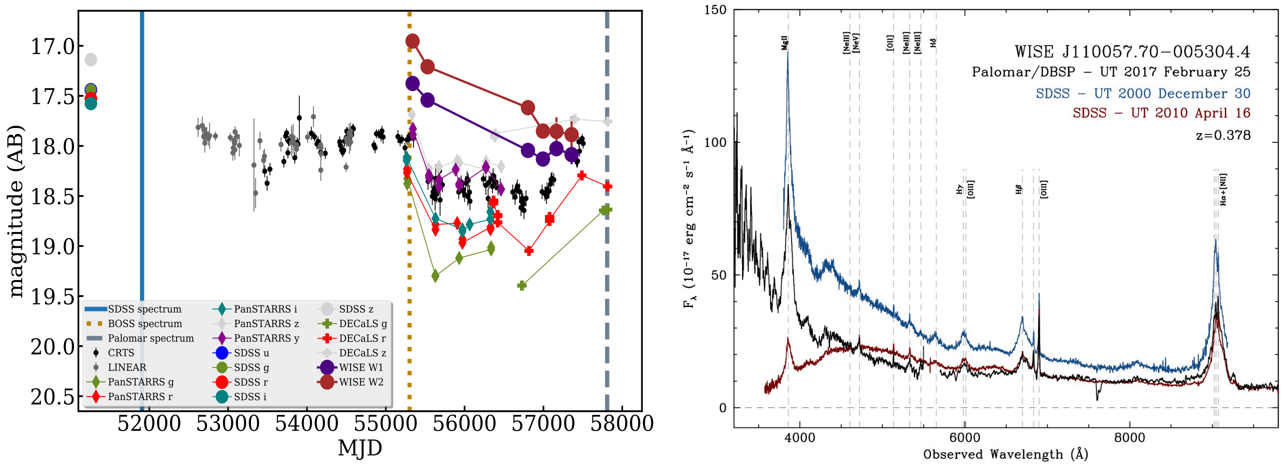

Figure 1: *(Left:)* The optical and infrared light-curve for the redshift $z = 0.378$ quasar J1100-0053 (Ross et al. 2018). Note the fall in the infrared, whereas there is a decrease, but then recovery in the optical. *(Right:)* Three epochs of spectra for J1100-0053. The spectacular downturn in the blue for the 2010 spectrum indicates a dramatic change in the accretion disk.

During its initial phases of operation the Sloan Digital Sky Survey (SDSS) obtained spectra of 1 million galaxies in the local Universe. This dataset has become the *de facto* standard for understanding the present day galaxy population, and sets the boundary conditions for all theoretical comparisons. The paradigm changing success of the SDSS was due to having 1,000,000 objects *with very high signal-to-noise photometry and spectra*, enabling multivariate analysis that is required for galaxy astrophysics investigations. *Q4D will generate the same sample size and revolutionary understanding with a new temporal dimension of the quasar population, as the SDSS had with the low-redshift $z \sim 0.1$ galaxy population.* The ground-breaking aspect of Q4D is that it takes quasar astrophysics into the 2020s, going from single objects samples, to surveys and samples of millions of objects, with massive spectroscopic monitoring giving access to the time-domain and leveraging these very large scale next generation missions, telescopes and their datasets.

The timing for this proposal could not be better. The first of the data "firehoses" turns on in late 2019, with the full datastream from our key sources fully online by mid-2022. As such, we have the time to mature our analysis techniques, and then be in the ideal position to take advantage of the initial data releases of all these new projects. Prompt ERC Consolidator level-support is also imperative since final survey design and optimization trade-off studies are being made e.g. for DESI, SDSS-V and LSST over the next ∼12-18 months. Having the ability to influence and fully optimise these decisions for our science objectives would be very powerful. The ground breaking nature of the Q4D will attract high quality PDRAs, *who would be guaranteed "First Light" data and science.*

The importance of this branch of astrophysics is already well established in Europe and is a priority for the next two decades.[7][8] This is demonstrated by noting that one of the two primary mission goals for the ATHENA

[7] ESA Cosmic Visions [8] L-Class Mission Timeline

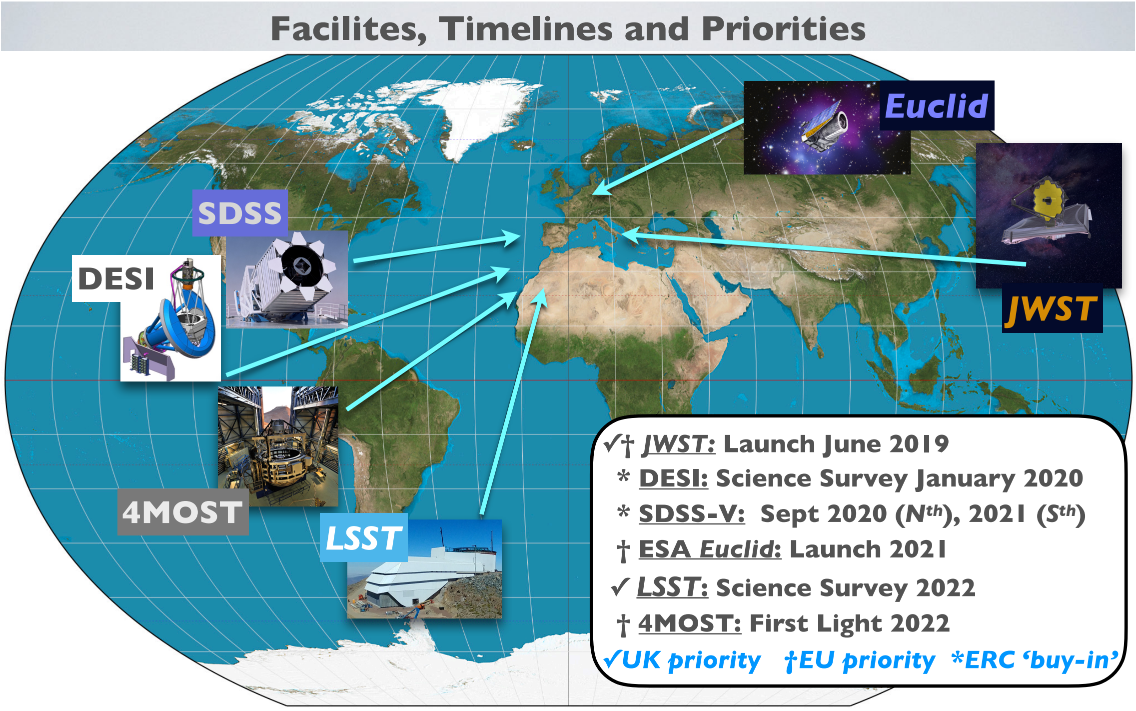

Figure 2: Facilites, Timelines and Priorities. With SDSS-V and DESI in the Northern Hemisphere and 4MOST, LSST in the South, we have full celestial sphere coverage.

mission is answering the question "How do black holes grow and shape the Universe?". ATHENA is ESA's second L-class flagship mission, due for launch in 2028.

The scope and remit of an ERC Consolidator grant will allow us to combine these data products in a manner that will not only establish the new state-of-the-art in variable extragalactic astrophysics, but it will also establish and kickstart the new field of variable extragalactic astrophysics itself.

2. Methodology

Our proposal contains six work packages that fall into three broad and complementary categories: observational studies of large numbers (millions) of objects; high-risk, very high-reward observational studies of a small number (10s) of objects; theoretical modeling investigations. Figure 3 summarises our overall WP plan. Risks and mitigation strategies are present for each WP as are Key Deliverables. NPR, three PDRAs, "PDRA1", "PDRA 2", "PDRA 3", and one PhD student, "PhD1" are the personnel required to carry out these work packages.

NPR is a world-leader in the field of extragalactic observational astrophysic. NPR's research focuses on implementing novel data science and machine learning algorithms and techniques in order to discover and study the physical processes in quasars. I have an exceptionally strong track record including being the lead of a science Working Group, with prodigious scientific output (over 400 published, peer-reviewed papers from that particular collaboration). **I was the Co-Founder and Chief Data Scientist of String Security Inc. where I built a predictive threat detection and remediation platform for cyber security teams by applying machine learning and predictive algorithms. Thus the P.I.'s research strengths, ability to quickly develop bleeding-edge software and science output are all ideally matched to this proposal.**

The skill set of PDRA1 would include development of the underlying tools and techniques necessary to extract meaning from large and/or complex data sets. PDRA1 would have a strong physical sciences background, and a PhD in astrophysics or computer science. The skill sets of PDRA2 would include expertise in time series analysis, primarily with optical data but potentially also in other wavebands. PDRA2 would have a PhD in astrophysics or a related field. The skill set of PDRA3 would include experience with fluid mechanics modelling and/or large computer simulations. PDRA3 would have a PhD in astrophysics, mathematics or computer

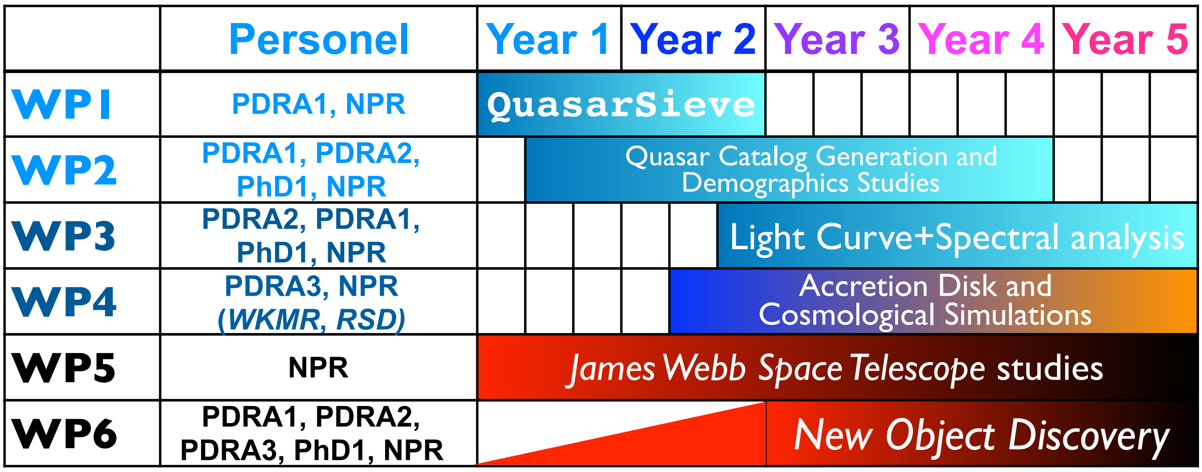

Figure 3: An overview of the WPs, with the personnel attached to each WP and a guide to their start and duration. As given by the shadings, WP1, 2 and 3 are observational studies of large numbers of objects; WP4 are theoretical modeling investigations and WP5 and 6 are high-risk, very high-reward observational studies of small numbers of objects.

science. PhD1 would have a Masters or a strong 4-year undergraduate degree in Physics or Mathematics with evidence of research-level project work.

WP1: Build QuasarSieve: In order to utilize the LSST datastream for our science goals we will build a "Stage 2 filter", which we name *QuasarSieve*. This will identify the quasars, add context, perform outburst forecasting etc. The heavy-industry computing infrastructure is being supplied by the UK LSST Data Access Center (DAC, based at the University of Edinburgh) and our task will be to build software in a timely and robust manner. This is a novel enterprise and a rate-limiting step in our overall programme, with the associated high-risk. We mitigate this risk with the data science and machine learning experience from PDRA1 and the P.I. (NPR). We thus classify **WP1 as medium-risk, high-reward. Key Deliverables:** An open-source, well-documented software package that can interact with and return data from the LSST Data Access Center.

WP2: Quasar Catalogue Generation and Demographic studies: Building the quasar corpus and cataloguing the observational data will be a vital step for our science goals. Following on from the quasar catalogue generation, a key science output will be the study of the quasar demographics. All these are vital observational tests for galaxy formation models and theory (see WP4 below). The goal of this WP is to construct a quasar catalogue and make key observational tests. Given the P.I.s experience at these specific tasks, plus the effort level of PDRA1, PDRA2 and PhD this WP is deemed medium-risk. **WP2 is medium-risk, high-reward. Key Deliverables:** A science-enabling compendium that will be the state-of-the-art quasar dataset for the 2020s. A suite of new, beyond-the-state-of-the-art quasar demographic measurements which are the boundary conditions for theoretical models.

WP3: Light-Curve and Spectral Analyses: Another major scientific output will be the full and detailed light-curve and spectral analyses of the conglomerate datasets. This WP will have a data science/machine learning aspect. The goal of this WP is to elucidate the physical processes that drive quasar variability and as such is significant high-reward science. This level of investigation is highly novel, though we envisage no major barriers outside of our control to achieving our science goals and PDRA1, PDRA2, as well as the P.I. (NPR) and PhD1 effort will be directed towards this. As such, **WP3 as medium-risk, high-reward. Key Deliverables:** Measurements, for the first time of how the light-curves and spectra of quasars depend on key physical quasar properties e.g. $M_{\rm SMBH}$, luminosity, $\lambda = \log(L/L_{\rm Edd})$, spin etc. These measurements will allow us to make direct comparisons to accretion disk models.

WP4: Accretion Disk and Quasar Feedback Simulations: New accretion models are needed to fully explain the "changing look" quasars and the "Quasar Viscosity Crisis". New radiation MHD codes begin to explain the observations here, but further development is needed to gain the desired deep understanding. Cosmological-scale simulations with stellar and quasar feedback are now also online. The exceedingly ambitious goal of WP5 is to develop new holistic accretion disk-to-cosmological scale simulations that explain our observational results and links them to "quasar feedback". WP4 is thus high-risk due to its novel nature and algorithmic complexity. We also envisage ramp-up time to get our theoretical simulations to the level that will be required by our beyond-the-state-of-the-art dataset. However, we mitigate this risk first by noting this will be the lead WP and top priority for PDRA3. We further mitigate this risk by invoking collaboration with accretion disk theorist Prof. Ken Rice (WKMR; Chair of Computational Astrophysics at the IfA, University of Edinburgh) and cosmological simulation expert Prof. Romeel Dave (RSD; Chair of Physics at the IfA, University of Edinburgh). Thus PDRA3, NPR, with guidance from WKMR and RSD would collaborate on this WP. We thus classify **WP4 as medium-to-high risk, very high-reward. Key Deliverables:** New accretion disk models and theory that explain the light curve data of our beyond-the-state-of-the-art dataset. New galaxy evolution models, describing the hydrodynamics involved on galactic scales, but related to the quasar central engine.

WP5: Observations of Quasars by the James Webb Space Telescope: What are the star-formation properties of luminous quasars at the peak of quasar activity? We aim to answer this by looking for the presence of polycyclic aromatic hydrocarbon (PAH) spectral features in infrared bright quasars with the *James Webb Space Telescope* (JWST). **WP5 is high risk, high-reward.** This is an ideal investigation for the JWST, but we classify this as high-risk since we have to apply for the telescope time and are not guaranteed the data. We note this WP does not impact in any direct way the other WPs and would lead to very-high gain science. **Key Deliverables:** State-of-the-art data products from the JWST, with the observational evidence and physical interpretation of how "quasar feedback" regulates galaxy formation in high-redshift quasars.

WP6: New Object Discovery: The LSST will scan the sky repeatedly, enabling it, and us, to both discover new, distant transient events and to study variable objects throughout our universe. The most interesting science to come may well be the discovery of new classes of objects. Suffice to say, this would be exceptionally high-reward. **WP6 is high risk, exceptionally high-reward.** We class this as high risk, since it is complex to class a WP with essentially unknown discovery potential as 'low-risk'. However, we *nota bene* that a lack of any novel discovery here would be a startling null result. **Key Deliverables:** Potential discovery of new classes of astronomical objects.

3. Resources, Survey 'buy-in' and Budget

Personnel: We request the resources and support for 100% of the time and effort for the P.I. We request the resources and support for 3 Postdoctoral Research Associates (PDRAs), for a total of 10 PDRA year equivalents (3+3+4). We request the resources and support for 1 PhD studentship.

Survey Buy-in: We request support for the "buy-in" to two of the new surveys, SDSS-V and DESI. The costs here are €184,100 for SDSS-V and €200,100 for DESI. We ask this support to come from the "additional funds that can be made available to cover access to large facilities." We request access to these funds as it gives our project access to telescopes and data in the North and South Hemispheres for complete coverage of the celestial sphere and delivers the crucial early spectroscopy that will be vital to train, test and build our data science and machine learning codes and algorithms. We emphasise that the science return is 'exponentially' dependent on the breadth of data available and heralds a brand new regime of "several-survey" or "multi-mission" astronomy. *Buy-in here would place the P.I. and the University of Edinburgh as the only group and institute in the world to be involved in SDSS-V, DESI, 4MOST, LSST and ESA Euclid and JWST.*

Computing Requirements: With the availability of the facilities at an institute (e.g. IfA Cullen), university (e.g. Edinburgh Compute and Data Facility) and at a national (The Hartree Centre) level, the rate limiting factor will be how quickly and efficiently we can deploy our codes and analysis.

Travel: We request support for travel for all 5 members of the group, including repeat medium-term (i.e., few weeks) travel to the US and ESO Chile to work with key collaborators at critical timings for the First Light of the new telescopes.

References

K. Barbary et al. Discovery of an Unusual Optical Transient with the Hubble Space Telescope. *ApJ*, 690: 1358–1362, Jan 2009. doi: 10.1088/0004-637X/690/2/1358.

A. C. Fabian. Observational Evidence of Active Galactic Nuclei Feedback. *ARA&A*, 50:455–489, Sept 2012. doi: 10.1146/annurev-astro-081811-125521.

Z. Ivezic and J. A. for the LSST Collaboration Tyson. LSST: from Science Drivers to Reference Design and Anticipated Data Products. *ArXiv e-prints*, May 2008.

J. Kormendy and L. C. Ho. Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. *ARA&A*, 51:511–653, Aug 2013. doi: 10.1146/annurev-astro-082708-101811.

A. Lawrence. Quasar viscosity crisis. *Nature Astronomy*, 2:102–103, Jan 2018.

LSST Science Collaborations, P. A. Abell, J. Allison, S. F. Anderson, J. R. Andrew, J. R. P. Angel, L. Armus, D. Arnett, S. J. Asztalos, T. S. Axelrod, and et al. LSST Science Book, Version 2.0. *ArXiv 0912.0201v1*, Dec 2009.

C. L. MacLeod, N. P. Ross, et al. A systematic search for changing-look quasars in SDSS. *MNRAS*, 457: 389–404, March 2016. doi: 10.1093/mnras/stv2997.

I. Pâris, P. Petitjean, N. P. Ross, et al. The Sloan Digital Sky Survey Quasar Catalog: Twelfth data release. *Astron. & Astrophys.*, 597:A79, Jan 2017. doi: 10.1051/0004-6361/201527999.

N. P. Ross et al. A new physical interpretation of optical and infrared variability in quasars. *Nature Astronomy*, 2018.

L. G. Spitler et al. A repeating fast radio burst. *Nat*, 531:202–205, March 2016. doi: 10.1038/nature17168.