(1) Submission Title or Survey Name:
Exploring the high redshift obscured accretion with Roman-WFIRST: A legacy survey leveraging on existing space and ground based facilities and tools.

(2) Contact
Nico Cappelluti, Department of Physics, University of Miami, Coral Gables, FL 33124, USA, ncappelluti@miami.edu

(3) **Co-authors**
Meg Urry (Yale University), David Sanders (IfA-University of Hawaii), Alessandro Peca (University of Miami), Priyamvada Natarajan (Yale University)

(4) **Do you support the selection of a Roman Early-Definition Astrophysics Survey**
Yes. Our support comes from the fact that an early definition survey would leverage on existing widely explored areas of sky by multiple facilities that a) might not be longer available at the time of Telescope operation and b) would require a significant amount of time and resources to assemble datasets that might be promptly available.
In particular, survey science cannot rely on a single instrument data but they are a rather panchromatic effort. Specifically at the time of writing this RfI there is no planned new X-ray facility capable of obtaining the imaging and throughput of Chandra and XMM-Newton. The availability of large X-ray dataset and broadband band multiwavelength coverage together with Roman revolutionary features will significantly improve the output of the Roman program both in the AGN/SMBH physics and in lensing studies.

(5) Describe the science investigations enabled by the survey

Our proposed study addresses how the youngest, biggest, most obscured black holes grew. These black holes grew at galaxy centers over billions of years, often hidden from view by thick columns of gas and dust along the line of sight. Theorists believe the enormous amounts of energy released during this growth provide feedback that drives galaxy evolution, including in the Milky Way. Much of that feedback energy may come from the most luminous Active Galactic Nuclei (AGN), i.e., quasars, which are sufficiently rare that they are found only in surveys spanning most of the extragalactic sky. We propose a multi-wavelength photometric and spectroscopic optical + IR + X-ray survey at the 50 deg^2 scale complemented with 4 filters NIR Roman-WFIRST photometry, 45 minutes deep. Our survey will make use of a large amount of existing and planned data sets, constituting the largest "layer" of the "wedding cake" probes of the full AGN and galaxy population (see Table 1). At present, only large optical surveys like the 10,000 deg^2 Sloan Digital Sky Survey (SDSS) probe volumes large enough to detect high redshift SMBH/AGN (e.g., Fan 2006, Mortlock et al. 2011, Venemanset al. 2013). But rest-frame UV light is easily extinguished by small amounts of obscuration, and popular theories of quasar activation and massive galaxy evolution, based on gas-rich mergers, imply an extended stage of obscured black hole growth (e.g., Hopkins et al. 2006). During that phase, luminous X-ray and IR emission would be prominent but bright UV/optical emission from the active nucleus would not be visible. We know obscuration makes optical surveys at lower luminosities miss most AGN (e.g., Treister et al. 2004,Vito et al. 2018, Ananna et al. 2019) and the incidence of obscuration appears to increase with redshift (e.g.,Treister et al. 2006, Ananna et al. 2019). Thus it is very likely that SDSS quasars represent only the "tip of the iceberg" of luminous and/or early black hole growth (Güver and Özel 2009). We propose to assemble the best optical-IR-X-ray 50 deg^2 survey to push the census of blackhole growth to moderate luminosity, high redshift, and high obscuration. We will then use three complementary approaches to assess the co-evolution of black holes and galaxies: decomposition of quasar spectral energy distributions (SEDs) using X-CIGALE (Yang et al. 2020), thanks to the superb angular resolution of Roman and ancillary data we will be measuring quasar host galaxy properties (redshift, morphology, star formation history, size, mass) using machine learning techniques (e.g., Salvato et al. 2019; Ghosh et al. 2020); and abundance matching analysis of quasar large-scale environments (e.g., Powell et al. 2018, 2020. Finally, we will use the quasar properties to constrain the growth of massive black holes at high luminosity and redshift, extending the work of Ananna et al. (2019). The availability of X-ray data, together with deep Subaru-HSC imaging and Subaru-PFS spectroscopy, SDSS, Spitzer and Roman will allow us to select and characterize obscured AGN at $z>4$ hence opening the window on the study of a large number of early obscured AGN and move below "the tip of the iceberg" of SMBH growth unveiled by SDSS Type-I Quasars (see Fig. 1). While the high latitude survey (HLS) of Roman will be magnificent in detecting these objects, only the proposed combination of multiwavelength data will be able to achieve this goal. In particular we want to stress that X-ray data are fundamental and no wide areas of the HLS can rely on a wide and deep coverage from Chandra and XMM but only from eROSITA with shallow coverage and a poor angular resolution of the order of 0.5 arcmin. While our main interest for this RfI is in AGN science the proposed survey with its ancillary dataset will be beneficial for the hot topics of Roman. We will be able to X-ray preselect hundreds of galaxy clusters and groups on a contiguous area. In this way we will be able to obtain a large sample of regions for weak lensing studies and to perform very deep, Frontier-Field-like studies of the faintest lensed background galaxies/AGN in the early Universe for pushing the faint end coverage of the survey thanks to lensing magnification.

(6) **Provide a possible observational outline of the survey**

We propose to cover the Stripe-82 XL area (Peca et al., in prep see cartoon) with four photometric filters (F129, F158, F184, F213) for a total of 179 pointings/filter, with 3600s exposure time each (including overheads) and a tiling strategy identical to the HLS to avoid gaps between the WFI chips. For high-z AGN (z~4-6) this choice corresponds to the rest frame wavelength [5-15] μm to enhance our selecting capabilities. Using the provided exposure time calculator we estimated the 27.8, 27.7, 27.3 and 26.1 AB magnitude limits in the point source case (which is what we expect for Quasar) and in the worst case Zodiacal light scenario for filters F129, F158, F184, F213, respectively. We also provide a second scenario with a exposure of 1200 s/pointing.

Why Roman in the Stripe 82-XL?

Stripe 82-XL represents a rich investment from both ground-and space-based observatories, including far more near- and far-IR coverage (including no longer available Herschel and Spitzer) than any other survey of comparable area.

One of the main limitations of the current design of the Stripe 82-XL is the lack of NIR high resolution imaging that is needed to perform a detailed analysis of the AGN host galaxy with our convolutional neural networks. This approach was basically limited to small area surveys with HST and now thanks to Roman this can be extended. Moreover the photometric bands of Roman are strategically located in a region of the spectrum where, at high-z, the SED of obscured and unobscured quasars show most of the discrepancy. Therefore by completing with Roman the survey of this unique region of the sky we can both open a new window in accretion science and create a unique dataset that fits most of the Roman scientific goals.

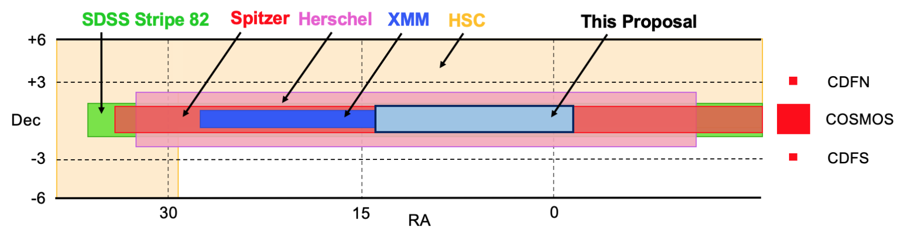
(7) Describe specific preparatory activities enabled by early definition

At the time of writing the proposal the Stripe 82-XL area has been covered substantially by most of the major facilities at any wavelength. Remarkably the project has been awarded XMM-Newton time and is leveraging on a Chandra archival project to reach 50 deg^2 in the X-ray band with imaging in the arcsec angular resolution regime. Our team has access to Subaru PFS and we are currently planning a vast spectroscopic campaign to target all the known AGN in the field and galaxies likely achieving an unprecedented spectroscopic completeness. Once paired with Roman we are confident in extending the analysis to Roman selected AGN. With these observing campaigns in place and with the potential of a Roman observation we will be in a position to competitively obtain a full Chandra coverage of the field as a Legacy program.

We are currently developing the infrastructure to store and analyze large datasets with machine learning. Among them, worth of notice, we want to mention GAMORNET-2 and PSF, convolutionary neural networks tools that allow the removal of the central point source in a quasar image and classifying the host. Within the same project we are developing AGN-finder, a deep learning tool aimed to discern AGN from regular galaxies and stars using multidimensional photometric and imaging that uses as a training set all the known AGN collected by our team in AGNDB. AGNDB (http://agndb.physics.miami.edu) is a large database that stores all the properties of all the know AGN in the sky by creating through a sophisticated matching a single entry per source that contains all the known photometric, morphological, spectroscopic and time domain information. The database, currently in beta version is fully interfaced with all the NASA archives and provides as data products machine learning friendly catalogs and images. These tools and observations would be of high value for the Roman program and having this survey planned in advance would allow us to leverage on the time before launch to complete the ancillary observational programs and the software and database development.

APPENDIX Figures, tables and references

Figure 1 Place holder



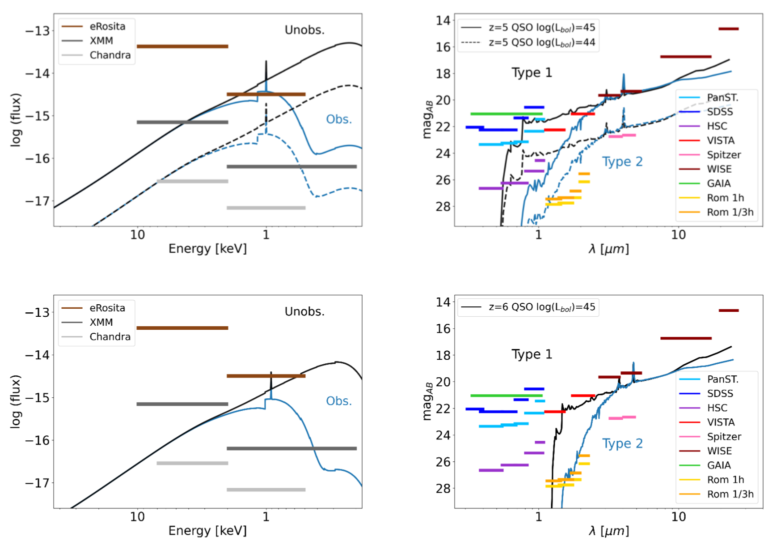


Fig. 2 Survey sensitivity (bars) as a function of energy (X-ray, left) or wavelength (UV-IR, right). Model spectra of Type 1/unobscured (black dashed lines) and Type 2/obscured quasars (brown dashed lines) with $L_{bol}=10^{46}$ erg/s ($L_X=10^{45}$) are also plotted. For the UV-IR, we used Type 1 and Type 2 quasar SEDs from Ananna et al. (2017). For the X-ray, we assumed a power-law model with photon index 1.8 and a 30% Compton reflection component; the obscured spectrum has an absorbing column density of $N_H=3\times10^{23}$ cm−2. Quasars at this low end of the targeted luminosity range are easily detected at redshift z=4 with the available surveys.

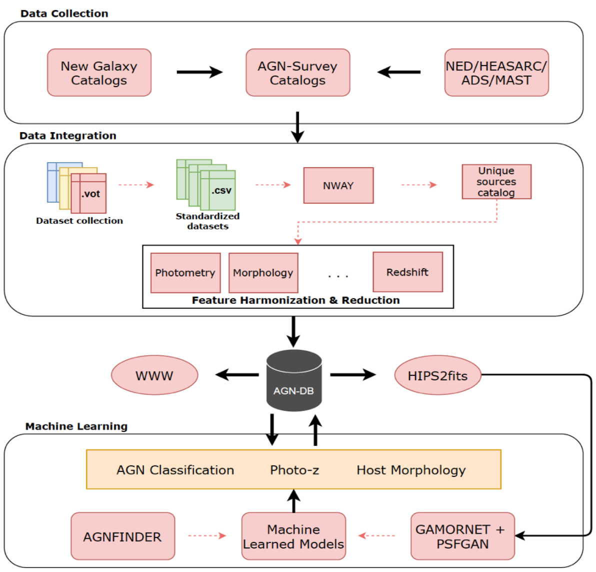


Table 1: Multiwavelength Coverage in Stripe 82

Band	Survey	Reference
X-ray	Chandra/XMM) - 45 deg^2; $F_X \sim 10^-15.5$ (cgs)	LaMassa+16
UV	GALEX(2 bands) - 300 deg^2; $m_{fuv} \sim 23.5$(AB)	Morrissey+07
Optical	SDSS coadd (5 bands): 300 deg^2; $r \sim 24.7$ (AB)	Annis+14, Jiang+14, Fliri+16
	CFHT (i band): 140 deg^2; $i \sim 23.5$ (AB)	PI: A. Leauthaud
	Subaru HSC (5 bands): 270 deg^2; $r \sim 26$ (AB)	Miyazaki+12
	DES: (4 bands) 300 deg^2 deg^2; $r \sim 25.1$ (AB)	astro-ph/0510346
	PanSTARRS (5 bands): 300 deg^2; $r \sim 21.6$ (AB)	Kaiser+10
	Spectroscopy: SDSS, BOSS, eBOSS, WiggleZ, PRIMUS, DEEP2, 2SLAQ, 6dF, VIMOS VLT Deep Survey (VVDS); 52% coverage of S82XL sources	Ahn+12, 14, Drinkwater+10, Coil+11, Newman+13, Croom+09, Jones+09, Garilli+08
	S-PLUS survey 12 optical bands	PI: C. Mendes de Oliveira
NIR	UKIDSS - LAS (4 bands): 300 deg^2; K$\sim$20.3 (AB)	Hewett+06, Lawrence+07
	VHS (3 bands): 300 deg^2; K$\sim$20 (AB)	McMahon+13
	VISTA-CFHT S82 (140 deg^2 to $J = 22.2$ mag and $Ks = 21$ mag (AB)	VICS82; Geach+submitted
MIR	*Spitzer*-SpIES (2 bands): 115 deg^2; [3.6]μm$\sim$21.9 (AB)	Timlin+16
	Spitzer-SHELA (2 bands) 28 deg^2; [3.6]μm$\sim$22.8 (AB)	Papovich+16
FIR	*Herschel* SPIRE (3 bands) 112 deg^2; 250μm$\sim$13 mJy/beam	Oliver+12, Viero+14; Hurley+16
Millimeter	ACT/ACTPol: 300 deg^2; 148, 218, 277 GHz; 2,3,7 mJy	Fowler+07; Swetz+11; Niemack+10
Radio	FIRST (1.4 GHz): 300 deg^2; 0.75-1 mJy	Becker+95; Helfand+15
	VLA-L (1.4 GHz): 92 deg^2; 52 μJy	Hodge+11
	VLA-B (3 GHz): 270 deg^2; 40 μJy	Mooley+16
	VLA-C (1-2 GHz): 70 deg^2; 40-50 μJy	Heywood+16
Planned		
Chandra	50 deg^2, $F_X \sim 10^{-16}$	
Subaru-PSC	50 deg^2, Spectroscopy	

Deleted: *<object><object>*