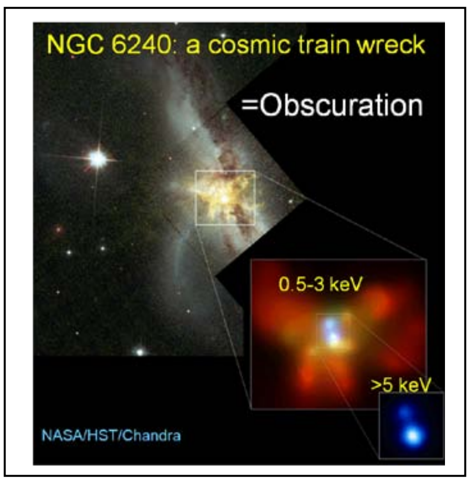
Lifting the Veil on the Black Hole-Galaxy Connection: Opportunities for 2010-2020

Science Whitepaper for the 2009 Decadal Review



Paolo Coppi (Yale), Loredana Bassani (ISAF, Bologna), Roberto Della Ceca (INAF, Brera), Alessandra De Rosa (IASF, Roma), Martin Elvis (SAO-CFA), Andy Fabian (Cambridge), Fabrizio Fiore (INAF, Roma), Suvi Gezari (Johns Hopkins), Josh Grindlay (CfA), Lars Hernquist (CfA), Phil Hopkins (CfA), Greg Madejski (Stanford), Lexi Moustakas (JPL), Francesca Panessa (IASF, Roma), Iossif Papadakis (Univ. Crete), Brad Peterson (Ohio State) , Gianpiero Tagliaferri (INAF-Brera), Fabrizio Tavecchio (INAF, Brera) , Jack Tueller (GSFC)

Associated mission: EXIST (http://exist.gsfc.nasa.gov)

1. Science problem overview and context: Since their discovery ~50 years ago, supermassive black holes, particularly the accreting (AGN) variety, have appealed to the more physics oriented subset of the astronomical community because of the extreme physical conditions and energetics they represent. Supermassive black holes (SMBH) can accelerate particles to at least 20 TeV, produce relativistic outflows with Lorentz factors ~40, and are routinely cited as possible sources for the highest energy cosmic rays. Detection of gravitational waves from their mergers and their captures of smaller objects could ultimately enable high precision tests of strong gravity. The very existence of SMBH, especially at early times, continues to challenge structure formation theory. Historically, though, others have not been as impressed by them. Yes, as some of the brightest objects in the known universe, detectable out to cosmological distances, AGN could be used as interesting probes. They made prime targets for gravitational lensing studies and by laboriously studying their AGN hosts one could get a glimpse of galaxies and their immediate environments at early times. The brighter ones could even be used as powerful searchlights to light up and characterize intervening matter. Still, many eyes glazed over when confronted by discussions of the convoluted AGN taxonomy or of what seemed to be black hole "weather." Simply put, supermassive black holes were a curiosity of nature, exotic and seemingly rare objects with little apparent connection to the "real" universe of optically bright stars and galaxies we see around us today. Interest in AGN for some declined further once astronomers gained the ability to directly find galaxies at high redshift, and AGN eventually lost the redshift race.

The mainstream perception of AGN is now arguably very different. Chandra and XMM show conclusively that the bulk of the 2-10 keV cosmic X-ray Background (CXB) comes from AGN and that AGN are much more common than previously thought, with densities on the sky of thousands per square degree. Correcting as best we can for the obscuration that has caused us to miss much of the AGN and star formation activity in the universe, we also find that the inferred star formation and AGN luminosities as a function of redshift appear to behave in very similar ways. The apparent coincidence between galaxy and SMBH evolution strengthens if one asks which objects are responsible for this luminosity as a function of redshift. Contrary to the naïve expectations of hierarchical CDM theory, both AGN and star formation activity show "downsizing," where progressively less massive and luminous objects dominate as time goes on. The nail in the coffin that black hole and star formation activity are closely interrelated, and perhaps inevitable outcomes of the same process (gas collapse), comes from looking closely at the centers of nearby galactic bulges: they all contain a black hole whose mass correlates very well with the velocity dispersion of stars and gas outside the black hole's gravitational sphere of influence (the "M-σ" relation) as well as with the overall mass and luminosity of the bulge. The rare monsters of days past are now the 600 pound gorillas sitting today, more or less quietly, in every room -- and they seem to know something about galaxy and star formation, topics that do interest the majority of astronomers.

The plot thickens as we begin to look back in time, using HST, Spitzer, and ground-based NIR telescopes, to probe the actual details of galaxy assembly. While today's dominant star formation mode, the relatively quiescent consumption of gas in disks like that of the Milky Way, does not seem particularly conducive to SMBH feeding and growth, this apparently is not true of star formation in the past. The massive "red and dead" galaxies that we now know exist at z~2 [1], for example, appear much more compact than their present-day giant elliptical descendants, and first surveys indicate they show relatively more signs of recent or on-going AGN/LINER activity [2]. Currently, the only plausible scenarios to create such unexpectedly compact objects involve rapid mergers of very gas rich progenitors, where strong dissipation drives gas to very high densities – presumably also creating an ideal environment for black hole feeding and growth. Indeed, the recent detection of hard X-rays, a unique signature of black hole activity, from stacked sub-millimeter galaxies (the most intense sites of star formation in the universe and possible "red and dead" galaxy progenitors) would seem to confirm this [3]. Conversely, as our multi-wavelength follow-up of AGN improves, we are finding that many of them also live in sites of recent or ongoing massive star formation. At some basic level, then, the study of galaxy formation in 2010 *is* the study of SMBH growth and vice versa.

The connection between intense star formation and AGN activity may be much more profound, however, than the fact the two tend to happen at roughly the same time and place. *Not only may SMBH be tracers of massive galaxy assembly, they may actually be key actors in the process.* The discovery of the M-σ relation spawned a vast literature on how "feedback" from an accreting central black hole can simultaneously shut down accretion and truncate star formation in a bulge, e.g., by blasting gas out of the nascent galaxy [e.g., see 4]. Indeed, note that if we magically donate just a few percent of the rest mass energy of our Galaxy's central black hole to the Galaxy's baryons, an amount of energy comparable to what we think was released as radiation during the black hole's accretion phase, this would unbind the baryons.

While SMBH clearly have the potential to profoundly influence their surroundings, e.g., as seen in the interactions of AGN radio jets with cluster gas, it is not actually obvious they influence galaxy formation in the manner envisioned by most M-σ speculators. Star formation simulations, for example, can produce Salpeter-like IMFs, where a certain fraction of gas mass ends up in very massive stars, purely from gas dynamical considerations [5]. Moreover, a one-time removal of gas is *not* sufficient in a hierarchical structure formation scenario. Gas which is initially unbound from the halo of a galaxy can find itself bound in the potential of a larger halo that eventually collapses, bringing that gas plus the additional gas contained in the larger halo back to the galaxy, reactivating accretion and star formation[6]. A key challenge for understanding galaxy formation in the next decade is thus to understand not just why massive galaxies suddenly become "red and dead" but why they, and perhaps their surroundings, *stay* dead during subsequent structure formation. What happens at late times is then just as important as what happens at early times, and in an extreme irony, the greatest impact of AGN on their surroundings may turn out to be posthumous.

The strong shock heating and turbulence that arises during the collapse of massive halos may contribute significantly to the late-time reduction of the cold gas supply (required for eventual star formation) that is available to galaxies, but a SMBH that is "mostly" but not quite dead or that periodically reactivates may also be part of the answer. There is much discussion now, for example, of a low power but steady "radio mode" of SMBH activity where perhaps, as in black hole X-ray binaries, accretion at low rates proceeds in a radiatively inefficient manner, with most of the power released during the accretion going into a kinetic outflow [7]. Something like this certainly appears to be happening in M87, an optically boring elliptical galaxy that in radio reveals a relativistic jet that may be surprisingly powerful, with a kinetic luminosity $\sim 10^{43-44}$ erg/sec by some estimates [e.g., 8]. Such a jet luminosity is well below the Eddington limit of $\sim 10^{47}$ erg/sec for M87's black hole, but if this low-level output persists for gigayears, the total energy accumulated is significant from a structure formation point of view. While compelling, this scenario is currently just that, a scenario to be tested. M87, for example, is a radio galaxy, i.e., it has an abnormally high radio luminosity compared to other ellipticals and thus perhaps a much higher kinetic luminosity that is not typical of other objects. *An accurate census and characterization of SMBH activity at low luminosities and late times, not just high luminosities and early times, will be required to fully unravel the story of galaxy formation.*

The pressing need for such a census and the means to achieve it, and then effectively exploit it, are the main points of this paper.

2. Realizable science opportunities for 2010-2020: Strategies, Lessons, and Prospects

Despite Chandra and a million quasar candidates from Sloan, we have much to do given the SMBH-galaxy story just told. Theoretical ignorance combined with prior limitations on our view of the universe, until recently dominated by observations at optical wavelengths, did not allow us to even ask the right questions. The black hole-galaxy formation problem is a complex one, involving a variety of physics and scales. The full solution will require a coordinated multi-wavelength, multi-technique approach that also focuses on getting at the relevant physics (that can then be incorporated into simulations) rather than the more traditional approach of just having separate communities count sources and compare notes. An important implication of this strategy is that the approach must necessarily be balanced or one stumbling block will significantly slow down progress and waste resources. For example, the science impact of an ultra-deep survey at one wavelength will be delayed if it produces sources that are not well-localized or that are too faint for other techniques, i.e., that cannot be followed up. Another important implication is that the need to target the physics means that one needs to do more than just chase after ever fainter and distant sources, the more typical mode of AGN surveys. Structure evolution in the universe does not run exactly in lock-step but rather it seems to be a function of local cosmic density: lower density regions evolve later. This means that even today there may be a few close analogs of systems at higher redshift – we just have to find them because they are rare. These are the systems, though, we can study in great detail to provide physics clues and rules of thumb to be applied to more distant system, e.g., as is being done today in the SINGS and THINGS galaxy surveys which choose objects such that high quality observations are available at comparable spatial resolution from radio to X-rays. A final consideration when dealing with galaxies is that the relevant evolutionary timescales are long compared to our timescales, even for their compact black holes. (The inner regions of the black hole system can vary on a multiple of the light crossing time, i.e., several hours, but global accretion rate or state changes could take much longer, i.e., 10^{5+} years if the analogy with stellar mass black hole binaries holds.) Any inferences about evolutionary sequences or states must necessarily be statistical and we must sample enough objects to get a fair representation of all the possible behaviors.

What do we need to really make progress on the black hole-galaxy problem and is now the right time to push on it? Some of the important ingredients for progress are: 1) spatial kinematics and temperature structure of hot and cold gas components in host galaxies, both near black hole and in the halo of galaxy, to understand feeding, star formation, and to search for outflows, 2) spatial kinematics and properties (abundances, ages, morphology) of stellar populations in host galaxies, 3) spatial measurement of star formation rates and properties in galaxies, including supernovae rates and bolometric luminosity, 4) measurement of magnetic field strengths and topologies, and of relativistic particle (cosmic ray) densities in and near host galaxies, 5) measurement of the host dark matter halo properties and the immediate galaxy environment to search for large scale structure effects (e.g., rich cluster vs. group, hot vs. cold gas fractions), 6) determination of the nature and geometry of the material that often blocks or colors our view of black holes and galaxies, and 7) measurement and correlation of SMBH properties (spectra, mass, spin, mass accretion rate) with all of the above in order to understand the physics of how a black hole responds to changes in its environment, e.g., what determines the fraction of accretion luminosity that emerges as radiation vs. a kinetic outflow.

This incomplete list is daunting, but the next decade in fact will see major progress in many of the areas, even without new project starts. Herschel, ALMA, JWST, SCUBA-2, and WISE will open up the mid-IR and FIR wavelength range, overcoming the dust obscuration problem and providing new line diagnostics. ALMA, E-VLA, and E-Merlin will provide spatially resolved gas temperature and kinematics maps. Particularly exciting, with a resolution of ~0.01", ALMA will finally reveal the kinematics of obscured nuclei on scales smaller than 100 pc (the size of the putative molecular torus in unified AGN model) for sources at $z<1$. E-Merlin and the E-VLA plus LOFAR will probe star formation in the radio with unprecedented sensitivity and independently from the FIR instruments as well as search for water masers in obscured AGN with large gas columns along the line of sight. These radio instruments together with Planck and VLBI and Fermi (in gamma-rays) will place improved constraints on cosmic ray populations and relativistic outflows as well as on magnetic field strengths (e.g., via Faraday rotation and Zeeman splitting). LOFAR, in particular, because it samples emission from long-lived electrons, will provide much better constraints on the time-integrated acceleration of relativistic particles due to AGN. Large area NIR surveys, e.g., by the VST, will photometrically identify many new high redshift galaxy candidates out z~2-3. Imaging of these candidates with ground-based 8m telescopes equipped with laser guided adaptive optics and JWST will provide significantly improved morphological information. For the brighter objects, multi-slit NIR spectrographs and IFUs will provide spatially resolved stellar population and emission line maps for significantly more galaxies at z~2-3. Dynamically determined SMBH masses will become possible for many more systems… This being said, we do see some significant bottlenecks, mainly on the black hole end, that could be ameliorated significantly without major technology development, i.e., within the decade. These are the goals of the EXIST mission, but the considerations are more general.

First, if we want to understand the physics of the central engine and how it responds to its environment, we must be able to work at X-ray energies, and in particular at *hard X-ray* (>10 keV) energies. Hard X-rays provide the cleanest and most direct probe of the environment immediately next to the black hole and penetrate through all but the most extreme (very Compton thick) obscuration. [Note that gas obscuration can completely wipe out the "blue bump" and soft X-ray emission from the SMBH accretion disk, and that intense star formation can mimic many of the signatures of AGN including emission up to ~3 keV, see Fig. 1b and image of the gas rich merger NGC 6240 on cover. FIR/mid-IR data can be useful for identifying obscured objects and constraining the reprocessed AGN luminosity.] As impressive as they have been, however, Chandra and XMM are unfortunately not hard X-ray instruments unless they look at objects at z>2. Very obscured objects clearly exist, including interesting ones which are possible progenitors for massive galaxies (e.g., NGC 6240), and it is very likely that large numbers have been missed.

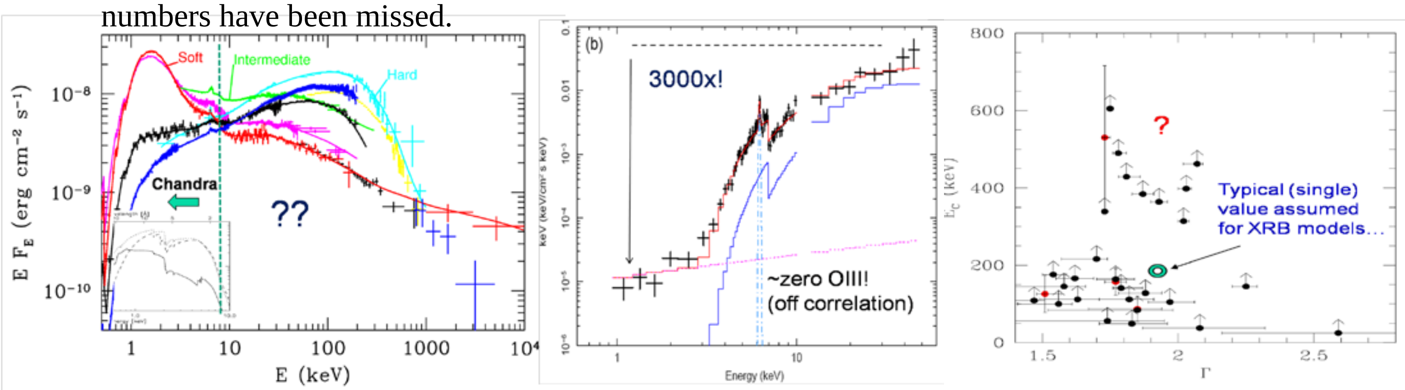


Fig. 1: *(left-a)* Compilation of broad band black hole binary spectra for various spectral states of Cyg X-1 from [9]. Inset shows Chandra effective area as function of energy. Very few AGN spectra of this quality and broad energy coverage exist (see rightmost panel for collection of fit values for many of them). *(middle-b)* unfolded Suzaku spectrum [10] of an extremely obscured source detected by SWIFT/BAT hard X-ray survey that shows very little evidence of scattered light or ambient gas emission at soft X-rays and none of the OIII narrow line emission expected in the unified model. *(right-c)* Compilation of spectral fit values (cutoff energy, EC vs. power law spectral index, Γ) for the *intrinsic* (before absorption and reflection) spectra of several Seyfert 2 galaxies observed by BeppoSAX [10], an instrument that did have good broadband coverage. Note the large spread vs. the standard value (*blue ring*) assumed in many AGN fits and X-ray background models.

Sensitive hard X-ray detectors are thus clearly required but have been difficult to achieve. The fact that hard X-rays are so penetrating also makes them also impossible to focus using conventional X-ray optics. Without good imaging capability, background rejection and thus sensitivity plummets. No hard X-ray mission flown to date has achieved a >10 keV sensitivity much beyond ~1×10^{-11} erg/cm^2/sec. At these flux levels, only the nearest non-blazar AGN are detectable. (NGC 1068 is barely seen). Luckily, technology has finally come to the rescue in the form of multi-layer foil optics, which allow imaging up to ~60 keV (to be deployed in NuSTAR and Astro-H and proposed for Simbol-X), collimated detectors with very low background (deployed in Suzaku and eventually Astro-H), and coded aperture mask CZT arrays with very large collection areas and a fine pixelation which allows good position resolution (prototype deployed in SWIFT, proposed for EXIST). The imaging and collimated detectors have narrow fields of view (<1 square degree), while the CZT arrays can have very wide fields of view and are suitable for all-sky applications. Widespread deployment and further development of hard X-ray technology will help the field significantly.

If one wants to do SMBH physics as opposed to merely SMBH detection, however, the ideal X-ray detector must satisfy one further, not always appreciated requirement: it must provide simultaneous broadband energy coverage, from well below the 6 keV fluorescent iron line to ideally to well beyond 100 keV. (Chandra completely fails this test, as does the Suzaku spectrum in Fig. 1b.) If it does not, experience with stellar mass black hole binary spectra shows that significant spectral model degeneracies exist, for example, between the amount of reflection (reprocessing) and the maximum cutoff energy of the spectrum. If one knows *a priori* the value of a particular model parameter (e.g., a cutoff energy was assumed in the fit shown in Fig. 1b), then other parameters are constrained well, but Fig. 1c demonstrates that nearby Seyferts, like black hole binaries, show a (different) *range* of intrinsic spectral parameters when one has enough statistics and energy coverage to constrain them. Dramatically increasing the number of objects with high quality, well-constrained spectra must be a goal of the next decade so that we can find out what is typical, say, for a given accretion rate or state. (EXIST would do this for ~1000 objects.) Future X-ray missions that rely on accurate continuum modeling but cannot provide or acquire the required energy coverage (e.g., via coordinated observations) will be impacted. Note that because AGN vary significantly, the soft and hard X-ray band coverage must be *simultaneous* (and ideally provided by the same mission if one wants to survey a large number of objects).

Looking at the upcoming suite of instrumentation, one sees that it consists mostly of narrow field instruments. This is problematic if one wants to find the most extreme (rare) objects or find the more typical objects (sparsely distributed on the sky) that are close enough to study their physical conditions in detail, e.g., with ALMA. Galaxy-black hole science would thus benefit significantly from a greatly improved *hard X-ray, all-sky survey* that can survey a volume large enough to sample a cosmologically representative range of environmental conditions. (The EXIST mission would probe a volume ~100 times larger than that probed by the current INTEGRAL and SWIFT/BAT surveys.) Note that for a hard X-ray survey to be effective, it must provide localizations, especially for the most obscured objects, that are accurate enough for unique counterpart identifications. Next generation hard X-ray technology provides ~15”-1’ localizations, not quite good enough. This can be improved by adding an onboard soft X-ray imaging telescope. (Weight/size requirements for such a telescope have dropped considerably.)

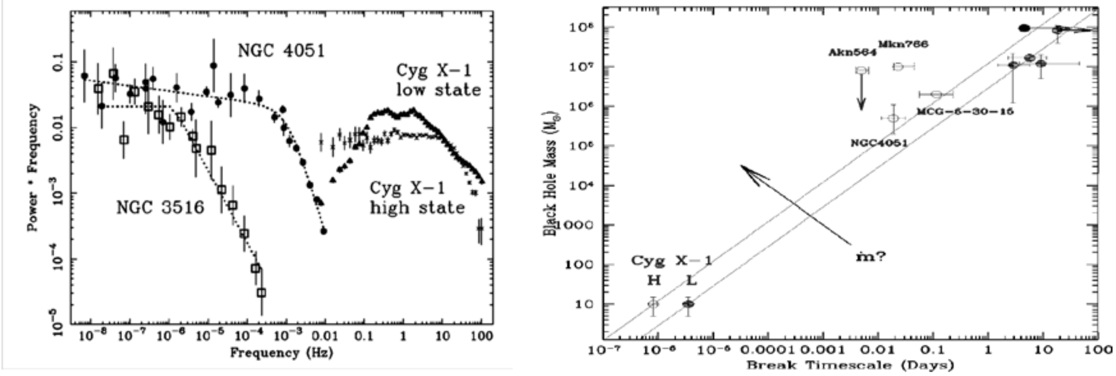


Fig. 2: *(left –a)* comparison of SMBH and binary black hole power spectra; *(right-b)* comparison of binary black hole and SMBH masses determined from power spectrum breaks with those determined by other means (dynamical measurements or reverberation mapping), from [12].

Variability is a hallmark of black hole accretion and relativistic outflows and can provide very useful information about the underlying black properties, break model degeneracies, or prove causal relationships between emission components. In black hole binary systems, for example, the shape of the X-ray variability power spectrum correlates with spectral/accretion state (high/radiatively efficient or low/radiatively inefficient), and the QPO and power spectrum break frequencies tell us about the black hole mass. AGN monitoring, mainly with RXTE - a flexibly scheduled satellite that has had a large impact on the field - indicates the same may be true for SMBH (see Fig. 2). If confirmed, this would have powerful implications, allowing hard X-ray mass determinations for highly obscured objects and providing cross check on whether low accretion, radiatively inefficient flows work as they do in X-ray binaries, e.g., by showing the same characteristic spectra or variability behavior. A flexible *timing/monitoring mission* in the next decade would thus be very useful (and straightforward to implement with a wide field survey instrument). Other applications of such a mission would be to find or at least monitor tidal disruption events (TDEs) where, as for black hole binary outbursts, the accretion rate systematically sweeps through a large range of values, allowing one to probe how the black hole system responds. It could also systematically monitor gravitationally lensed quasars to place constraints on the accretion disk structure via microlensing fluctuations [14].

Finally, the technology for NIR/optical space telescopes has improved to the point that off-the-shelf hardware (partly developed for JWST) allows cost effective deployment of small, ~1m class telescopes (e.g., on the same platform as a hard X-ray detector). NIR follow-up for AGN hosts is a major bottleneck, particularly spectroscopy which is extremely expensive for ground-based telescopes. Deployment of such a small telescope with a R~3000 spectrograph or IFU would allow studies of stellar populations and kinematics for a large sample of AGN hosts and field galaxies out to z~2.

References and supplementary materials may be found at [*http://www.astro.yale.edu/coppi/existwp*](http://www.astro.yale.edu/coppi/existwp) .