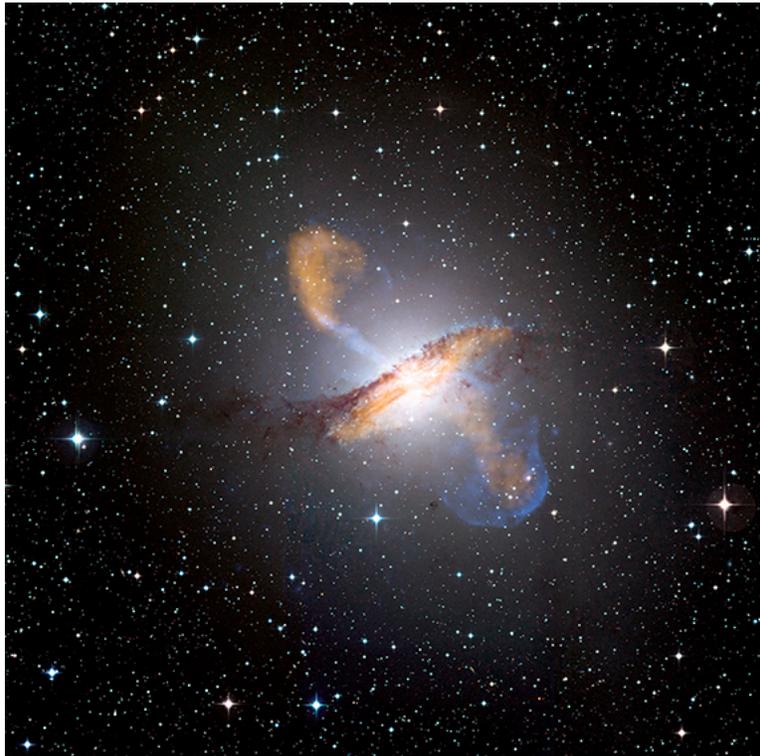


ACTIVE GALAXIES AND QUASARS, 2010-2020

Science White Paper for the 2009 NRC Decadal Review

MARTIN ELVIS¹, W. NIEL BRANDT², DIANA WORRALL³,
GIUSEPPINA FABBIANO¹, ANN HORNSCHEMEIER⁴, ROGER BRISSENDEN¹

February 16, 2009



1. Smithsonian Astrophysical Observatory.
2. Pennsylvania State University.
3. University of Bristol.
4. NASA-Goddard Space Flight Center.

Active Galaxies and Quasars in 2010-2020

Key Questions:

1. **How does the quasar central engine work?**
 - (a) How do Supermassive Black holes grow?
 - (b) How do quasars accelerate matter, up to the highest known energies?
 - (c) Why are 99% of Supermassive Black Holes quiescent?
2. **How, and when, does AGN feedback on galaxy evolution work?**

1 Scientific Opportunities and Context

The study of quasars and Active Galactic Nuclei (AGNs) is the study of the growth phase of Supermassive Black Holes (SMBH, $\sim 10^6 M_{\odot} < M_{BH} < \sim 10^9 M_{\odot}$). Quasars and AGNs have long posed fundamental questions in astrophysics but, for almost as long, AGNs have been largely detached from the rest of astrophysics. In the past decade, advances in both observation and theory have not only offered solutions to these longstanding key questions and opened up the chance to test General Relativity in the strong gravity regime, but have also implicated AGNs closely with problems in galaxy formation and evolution.

1.1 Advances in last decade

The past decade has been the era of the Great Observatories. With Chandra/XMM, Spitzer and Hubble STIS/ACS all providing qualitatively new data, and operating concurrently, a synergy has been created for rapid feedback of discoveries from one band to another. This has also been the SDSS era, which provided vast samples, by earlier standards: e.g, $\sim 100,000$ quasars rather than a few 100s. Notably, smaller facilities (e.g., FUSE, GALEX, reverberation networks) have also played crucial roles.

In parallel, theoretical developments, both numerical and conceptual, have brought the origin and growth of supermassive black holes to a level where clearly defined theories are now in circulation. Confrontation of these theories with data has become feasible.

These capabilities have translated into qualitative advances in AGN/quasar research that cohere into a few 'compelling themes':

For AGNs themselves:

1. Unobscured AGN evolution is now well defined, the soft ($E < 5\text{keV}$) X-ray background is resolved, and the Soltan[1] argument (comparing the total accreted mass implied by the quasar luminosity function with the total remnant SMBH mass locally) is known to work to first order [2, 3]. [Given by X-rays (Chandra, XMM), SDSS and HST/STIS.] However, the global question of '**How do SMBH grow?**' remains unanswered because *obscured growth of SMBH* [Spitzer, Chandra] is sketchy but potentially large [4, 5], while the *early SMBH growth phase* ($z > 7$) which is unconstrained by the Soltan argument, remains totally unknown. Early growth may need to be super-Eddington to give the observed $10^9 M_{\odot}$ SMBH by $z=6.4$ [6, 7, 8].

2. AGN research has transitioned to using *physical parameters* (e.g., mass) instead of observed ones (e.g., luminosity). The most important of these are SMBH masses [9, 14, 15], from which we measure Eddington rates, and which also showed the tight connection between SMBH growth and stellar bulge growth (the 'M- σ ' relation [11, 10, 12, 13] [Given by STIS, reverberation mapping,

X-ray power spectra and secondary methods (optical/UV)]. The broad 6.4 keV Fe-K line in AGNs is adding black hole spin and inclination angle [16] . [Given by X-rays (ASCA, XMM, Suzaku)] Other physical parameters we now possess are the radial scale for the broad (1% c) emission line region (BLR), which turns out to lie on accretion disk scales (a few 1000 Schwartzchild radii, R_g), in quasi-Keplerian motion [17], and the distance to the hottest dust [18]. [optical and near-IR reverberation] 'Eclipse mapping' by fast moving 'Compton thick' obscurers [19] has started to scan the inner accretion disk. [X-rays] As we combine these measurements we can see our way to **mapping the structure of active nuclei** in detail.

3. We now realize that *AGN winds are widespread* (and possibly universal) thanks to surveys of UV and X-ray 'warm' absorption lines near to the AGN redshift [20, 21, 22, 23]. [given by FUSE, STIS and X-ray gratings on Chandra and XMM.] Extending this result to higher velocities, spectropolarimetry showed that the 0.2 c wide broad absorption lines (BALs) have strongly polarized residual light within them, implying a highly non-symmetric structure [25][Keck], implying that BALs must be a widespread feature of AGNs (and have a small opening angle), rather than being rare and spherical. The origin (disk wind? torus?) and strengths of AGN winds are avidly discussed [26, 27, 28, 29] , but undetermined. **The role of winds in accretion and feedback** has moved center stage.

4. The superluminal expansion in relativistic jets has been traced in blazars in detail [30] [VLBA] and its occurrence and polarization connected to Gamma-ray [EGRET], optical [ground] and X-ray flares [31, 32]. Imaging of jets seen side-on (e.g., in M87) has found the collimation region at $\sim 50 R_g$ [35, 36], shown that Gamma-rays can come from knots outside the core (by time correlation with X-ray, optical brightening [33]), and shown that jets remain relativistic to Mpc distances [34] [VLBI, VLA, Chandra, HST]. Jet models have become more sophisticated (e.g. fast core - slower sheath models to explain these discoveries, and can explain many details [36, 37]. **Pinning down the jet acceleration mechanism** as a prelude to understanding the energy transport from the nucleus, now appears feasible.

5. The discovery of SMBH in almost all nearby massive galaxies implies that **the great majority of SMBH are 'quiescent'**, like our Galactic Center, as they show no sign of AGN activity above $\sim 10^{-8}$ of \dot{m}_{Edd} , or $\sim 10^{-6}$ of \dot{m}_{Bondi} [38, 39]. [Given by Chandra, Hubble imaging.] How to keep an SMBH this quiet is not well understood. Occasionally a quiescent SMBH lights up rapidly and fades over ~ 1 year [40, 41]. These are thought to be '*Tidal Disruption Events*', where a single star strays within the tidal breakup radius around the SMBH. [given by X-ray, UV imaging [ROSAT, Chandra, GALEX]

The relation of AGNs to their Large Scale surroundings has become both urgent and tractable. Contact has at last been made between observations and the theory of galaxy evolution and quasar (SMBH) evolution to the point where some form of **co-evolution of AGNs and galaxies** is now required, though not understood, from several disparate lines of evidence:

- Modelling of the development of Large Scale Structure (LSS) from initial CMB fluctuations has become immensely sophisticated, moving beyond 'Dark Matter plus large baryonic test particles', to include gas hydrodynamics and, to a limited extent, central SMBHs. This modelling has led to a fleshing out of once simple scenarios for galaxy/SMBH evolution via mergers, including feedback, and suggests a 'shrouded quasar' phase during the starburst caused by the merger, and a 'breakout'

phase where the quasar throws off the dust and gas around it to emerge as a 'naked quasar' shining strongly in the UV. [Given by the Millennium Run, GADGET and other codes [42, 43, 52, 45]

- A strong AGN-starburst link is supported by the 'downsizing' that is seen in AGN evolution [46, 47], as lower luminosity AGNs peak at lower redshifts [given by ROSAT, XMM, Chandra], a behavior similar to that seen in star formation in galaxies in the Lilly-Madau plot [48, 49, 50].
- Yet the 'red sequence' of massive passively evolving galaxies found in GEMS, GOODS and other fields challenges this picture: how can major mergers occur without star formation [51]? Where did the gas go? It has been proposed that AGNs have ejected all of the ISM in these galaxies [52]. Energetically this is easy, but a realistic mechanism has not been identified.
- The 'cooling flows' of clusters of galaxies neither cool nor flow [53, 54]. Somehow a balance of heating and cooling is maintained. This X-ray-discovered puzzle appears to be answered by Chandra/VLA observations that show the X-ray and radio images fitting together like a jigsaw - the radio lobes blow bubbles into the hot ICM, presumably injecting heat [55, 56]. Yet this apparently easy answer is complex when studied in detail [57], leaving the balance unexplained.

The beginnings of this co-evolution are also starting to become tractable. Modelling of *the origins of SMBH*, in the first metal-free Population III stars [59, 60], or by direct collapse into 'superstars' [8, 58], and their subsequent history via the 'merger tree' as dark matter halos merge [61], has become a detailed and quantitative study. Observations of $z > 6$ quasars are expanding, and the starburst link may be present, as large reservoirs of CO are seen, enabling total mass measurements [62], and [CII] measures high star formation rates [63]. [SDSS, HST, Chandra, VLA, SMA, IRAM]

1.2 Compelling Themes

The fundamental *physics* questions posed by quasars and AGNs, 'How do SMBH grow?', 'How are AGN winds and jets accelerated?', 'Why are 99% of Supermassive Black Holes quiescent?', all boil down to one theme: *How does the quasar central engine work?*. The unknowns of the *astrophysics* of quasars and AGNs all revolve around one theme: *How does AGN galaxy co-evolution work?*.

20th century astrophysics primarily studied objects in isolation: a galaxy, a star, a cluster of stars or galaxies. In the 21st century this 'island universes' approach has given way to studying the interconnections, the 'ecology', of astronomy. These interconnections are often driven by features that are far from dominating the radiative output, and so demand high sensitivity. Yet they tie together the cosmic landscape into a unified model based on physics. This is as true of AGNs as it is of galaxies or star formation.

2 Broader Scientific Context

Quasars dominate the sky away from the UV/optical/IR bands, and so pose the basic question of astronomy: 'What do we see in the sky?'. Yet only a decade ago, quasars seemed largely to be a thing apart from the rest of astrophysics. Now quasars are linked to large parts of astrophysics, and to 'fundamental' physics.

The workings of the quasar central engine works involves a broad array of physics questions.

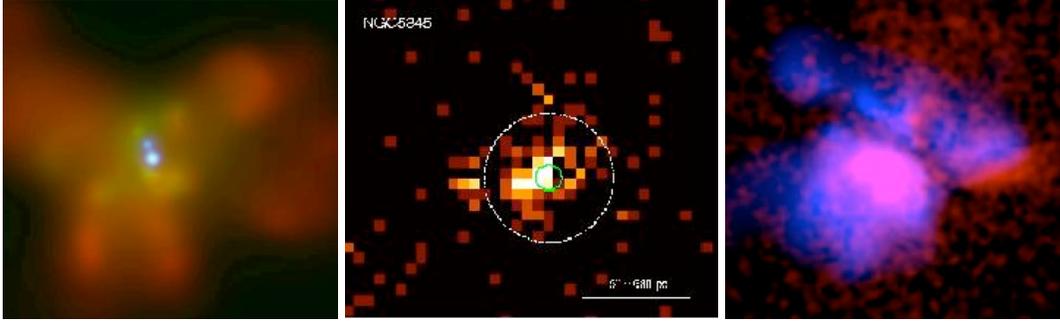


Figure 1: Chandra $\frac{1}{2}$ arcsec resolution images of: *left*, the double active nucleus in NGC 6240 [40]; *center*, the quiescent nucleus of NGC 5845 [39]; *right*, the extended X-ray nucleus of NGC 1365 (red), and the Hubble [OIII] narrow line region image (blue) [64].

These involve using known physics in new ways (e.g. to understand how accretion disks really work, to understand how winds are driven), and so find commonalities with other areas where similar questions arise: e.g. *X-ray binaries, Cataclysmic variables, Wolf-Rayet and O-stars, Red Giant Branch stars, T Tauri stars, the Solar corona*. A quantitative model of how quasar central engines work will turn 'feedback' to *galaxy evolution* into a fully physics-based field by predicting the mass loss rate from winds and jets. Quasar central engines also probe extreme conditions of gravity, that could lead us to 'new' physics: *relativistic jet acceleration, the space-time near the Schwartzchild radius*.

2.1 How do Supermassive Black holes grow?

High resolution X-ray imaging of $1 < z < 3$ AGNs in galaxies will determine the 'Black Hole Merger Tree': i.e. the history of the number of supermassive black holes in a galaxy with cosmic epoch. This history is highly sensitive to models of the growth of structure and to models of 'feedback' by AGNs on their environment. FIGURE 1A.

The *growth history of supermassive black holes* (SMBHs) is governed by two mechanisms: accretion and mergers. While AGNs trace the bulk of the accretion onto SMBHs in the Universe, as the Soltan argument showed [1, 3, 12], SMBH mergers are more elusive.

X-rays are much used to map the accretion history of SMBHs [46], as this band provides the highest surface density of AGNs, penetrates obscuring dust and gas, and can be seen back to the earliest times (pre-reionization HI absorption has no effect on them). The most obscured AGNs need surveys in the infrared [4] [Spitzer, JWST] and in high energy (>10 keV) X-rays [EXIST, IXO] up to $z \sim 6$. The $z > 7$ rapid growth phase needs more sensitive telescopes [Generation-X], and redshifting puts the hard band in the convenient 0.5-10 keV band.

High angular resolution in X-rays can explore the other SMBH growth path - mergers - by mapping out the *Black Hole Merger Tree* [61] - in the obscured environments where they are expected. Black Hole mergers occur as part of the process of the merger of Dark Matter halos, which we see as galaxies. By the present epoch, the single SMBHs seen in most, but not all (e.g. M33 [9]), massive galaxies have a constant ratio of SMBH mass to host galaxy bulge mass [10, 11, 12, 13, 14, 15]. SMBH mergers plus accretion must contrive to produce this ratio, and must produce the observed, not quite 100%, 'occupation fraction' of SMBH in massive galaxies. SMBH mergers are most directly found by gravitational wave chirps, signalling the final moments of a black hole merger

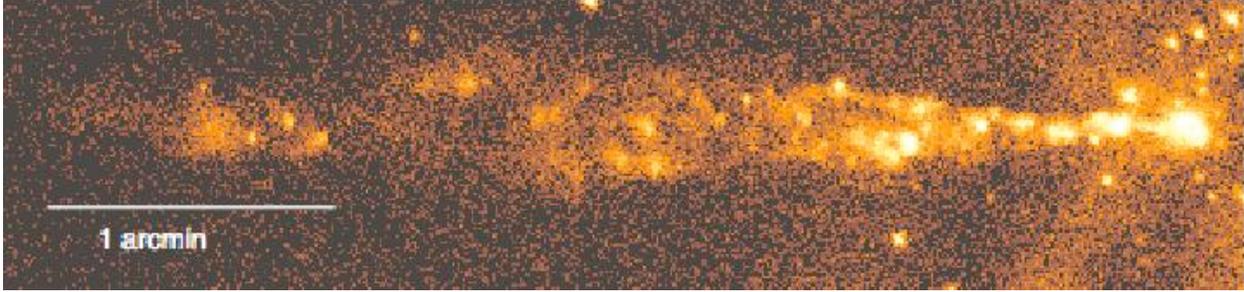


Figure 2: Chandra image of the jet in Centaurus A with 0.5 arcsec resolution showing X-ray knots that must be sites of particle acceleration [69].

[LISA]. A pair of SMBHs spends far longer at kpc separations though [61], and so will be more common in X-ray surveys. Combined, the merging time can be measured.

A few AGN pairs are already known in obscured nuclei; the most famous being NGC 6240 where the AGNs are separated by ~ 1 kpc. Their large EW Fe-K lines at 6.4 keV show the nuclei are Compton thick and seen only in scattered light. At higher redshifts AGN pairs should be much more common [61]. For concordance cosmology a 1 kpc separation is an almost constant ~ 0.17 arcsec for $z > 1$, so X-ray imaging at $6.4/(1+z) < 3$ keV with ~ 0.1 arcsec resolution and 100 times the *Chandra* area, is the best path.

2.2 How do quasars accelerate matter, up to the highest energies?

The sites of particle acceleration in jets are local and distant from the central black hole; high resolution X-ray imaging will help us learn how these likely sources of the highest energy cosmic rays are accelerated. FIGURE 2.

There is an unknown, highly efficient, process occurring in radio jets that accelerates particles up to relativistic energies, far above that of the Large Hadron Collider [34]. These particles may be the source of *the highest energy cosmic rays* now being imaged by Auger [66]. Ultimately, the source of the high particle energies and bulk velocities in jets is likely to be the spin of the central black hole (the 'Blandford-Znajek' effect [67]), but the intermediate steps remain murky.

X-ray imaging of 'radio' jets, notably of the nearest Cen-A ($D=3.7$ Mpc) [68, 69], where our physical resolution is at its best, shows that *local acceleration* must dominate: small knots of X-ray emission occur all along the jet, yet the particles radiating these X-rays have synchrotron lifetimes far shorter than the travel time from the distant cores, so much so, that many radio galaxies must have localized particle acceleration. Finer resolution will show how the particles are accelerated and then age away from these localized production sites. This is the only key we have to learn the local mechanisms at work. To reach more jets at ~ 100 pc resolution requires at least 5 times the angular resolution of *Chandra*. To measure the age of the electron population requires good spectra of the bright knots and so larger area than *Chandra*. [Generation-X]

2.3 Why are 99% of Supermassive Black Holes quiescent?

High Resolution X-ray Imaging is the best path to understanding why the great majority of Supermassive Black Holes are quiescent, radiating not only well below the Eddington rate, but well below the Bondi rate. FIGURE 1B

Almost all large galaxies contain SMBH, so why are 99% of them inactive in the current epoch? There are several dozen cases in the local universe ($D < 20$ Mpc) of known black holes in galaxies that show none of the accepted signs of activity. As $L = \epsilon \dot{m} c^2$, either there is no accretion occurring, or the accretion is radiatively inefficient, or the radiation is hidden from our view. *Chandra*, with 0.5 arcsecond resolution in X-rays, made a breakthrough, revealing weak X-ray sources and setting limits on \dot{m} at $\sim 10^{-4}$ of the Bondi limit (and $\sim 10^{-7} L_{Edd}$). Our Galactic center, SgrA* is a good example. Several of these nuclei have suggestive X-ray structures around them that could be small outflows or jets [38, 39].

However, with *Chandra* we are limited to crude spectra and uncertain images of quiescent nuclei. Larger area will provide spectra, but only in conjunction with high angular resolution, as the X-ray binary population and the hot ISM cause confusion. Spectra and variability will distinguish nuclear SMBH emission from that of X-ray binaries, and so measure \dot{m} rather than merely placing limits, and will tell us whether the SMBH emission is like that of efficient radiators (normal AGNs), or like the predictions for inefficient radiators (ADIOS, ADAFs).

JWST will measure many more quiescent SMBH masses, in a wider variety of host galaxies than HST-STIS was capable of doing. An X-ray imaging survey of all JWST discovered quiescent SMBHs will need higher angular resolution than *Chandra* and larger area [Generation-X], and will reveal the physical conditions that allow highly sub-Bondi accretion.

Candidate 'Tidal disruption events' (TDEs) should begin to be found routinely with Pan-STARRS, and in abundance with LSST. Detailed follow-up with optical, ultraviolet and X-ray spectra and light curves can settle their nature and, if TDEs, will lead to mass and spin measurements of otherwise quiescent SMBHs.

2.4 When and how does AGN feedback on galaxy evolution work?

High Resolution X-ray spectra ($R \sim 10,000$) of ionized outflows (BALs and WAs) will determine their location and establish their mass, momentum and metal loss rates, providing a firm underpinning for models of feedback. FIGURE 1C.

There are three ways in which an SMBH could cause feedback: (1) radiation; (2) relativistic jets, (3) slow (non-relativistic) outflows. All three need investigation. Jets were already discussed.

Feedback via radiation seems uncontroversial, as we see AGNs across the spectrum. However, there are complications: the geometry of the nucleus matters - does the 'obscuring torus' allow radiation to impact the host ISM? Polarimetry, near-IR interferometric imaging, and variability in the X-ray obscurer, all pin down the obscurer geometry. The Compton temperature of the Spectral Energy Distribution (SED) governs the effect on the host ISM [71]. Currently, an overly simple uniform SED is typically used [70], and the dominant EUV band is just interpolated. To map out the redshift, luminosity and other dependencies of the quasar SED needs omni-wavelength surveys from the IR to 30 keV (rest frame).

Feedback via Non-relativistic Outflows involves studying the ~ 1000 km s $^{-1}$ to $\sim 20,000$ km s $^{-1}$ winds seen in UV and X-ray absorption spectra [20, 21, 22, 23] and some 'bi-cones' of the Narrow Line Region (NLR) [24]. As these winds are seen in over 50% of AGNs, and could be universal [25, 27], they are the obvious source of feedback for most galaxies. Unfortunately the kinetic energy and mass outflow rates in these winds are uncertain by factors of $\sim 10^6$! [28, 72, 73, 74, 75].

Optical, UV and X-ray spectral imaging at \sim parsec resolution [UVOI, MRO-I, Gen-X] could measure \dot{m}_{Wind} . X-ray absorption lines contain density diagnostics [76] that will give loss rates, but only at $R \sim 5000$ and high S/N [IXO, Gen-X].

3 Key Advances Needed: 2010-2020

Quasars are the exemplars of objects in the sky that ignore our division of astrophysics into convenient wavelength 'bands'. Emitting strongly over 10 - 20 decades of the electromagnetic spectrum (10 for quasars, 20 for blazars) no single instrument can possibly answer all our questions. *Continued concurrent availability of a suite of pan-chromatic observatories through 2020*, is more essential for quasar research than for any other single area.

Quasars are also the quintessential point sources (they are 'quasi-stellar' objects, after all). So spectra, timing and polarimetry have been the main sources of our knowledge. In timing, the 'virtual imaging' of *reverberation mapping* can be developed much further [77] with ground networks [78], and a small UV space mission, while large temporal/solid angle surveys [Pan-STARRS, LSST, EXIST] will find TDEs routinely. Spectropolarimetry of faint emission lines, with 30 m class telescopes, will expose inner AGN structures. Spectroscopically, while the UV and optical have much to offer, the least explored band rich in atomic features is the X-ray. The 6.4 keV Fe-K line will probe the inner few Schwartzchild radii, while the 100s of absorption lines in the 0.2 - 2 keV bands produced by warm absorber winds have enormous diagnostic potential at high resolving power ($R \sim 10^4$). The *microcalorimeters and diffraction gratings spectrometer on IXO* will open up this field

Imaging nonetheless has great potential for quasar research. It is striking that all of the key questions need *high angular resolution X-ray imaging*, of order 0.1 arcsec, always with substantial collecting area. The technology to make 0.1 arcsec X-ray mirrors is nascent, and is dependent on the success of the IXO mirrors. (IXO mirrors will have XMM quality at 1/10 the mass/effective area ratio of XMM.) With *technological investment over the next decade may yield 0.1 arcsec resolution X-ray optics*, laying the basis for a new mission, *Generation-X*, sometime after 2020, combining 'super-Chandra' angular resolution and large collecting area.

Still higher angular resolution would directly image the inner workings of an active nucleus. Working inwards, these are: how mass reaches the inner parsec and how it sheds angular momentum to reach the accretion disk; what form the 'obscuring torus' takes; whether the gravitationally unstable disk makes stars; where the AGN winds come from (torus/disk); what the broad emission line region is and how it is moving; the shape of the 'disk' continuum source in the optical and UV and, eventually, where the 'X-ray corona' really lies. At $<100 mas$ extreme AO on 30-meter class optical telescopes [GMT, TMT, ELT] will reach the torus; at $\sim 1 milliarcsec$ resolution with interferometers, UV in space [UVOI], IR on the ground [Magdalena Ridge, Keck, Antarctica] will resolve the outer BLR. $10-100 \mu arcsec$ will image the full BLR, while imaging reverberation mapping would give full 3-D maps. To realize these capabilities soon after the 2020 horizon requires *technology investment in interferometry* in 2010-2020. Some near-IR interferometers (e.g. Keck, Magdalena Ridge) could reach AGN sensitivity within a few years, if funded.

References: <http://hea-www.harvard.edu/~elvis/Astro2010>
<http://www.cfa.harvard.edu/hea/genx/dev/astro2010/>