Academic year 2025 - 2026

P. Monaco

Radioastronomy
Active Galactic Nuclei

Notes of the course for *laurea magistrale*, University of Trieste

Figure: radio imaging of M87*, the black hole at the center of the M87 galaxy.

♣ In many galaxies it is possible to notice **a very strong nuclear emission**. In other words, the central region of the galaxy, point-like within the angular resolution of the telescope, appears much brighter than what expected from the galaxy surface brightness profile. In some cases the difference in luminosity is so large that the nucleus is observed as a point source, while the rest of the galaxy remains unobserved.

♣ The optical spectra of these nuclei always show emission lines, besides continuum components. Sometimes, for the least bright nuclei, these lines are similar to those of HII regions, and can then be explained as emission from strongly star-forming regions, where massive stars are still alive together with their Stromgren spheres. Multiple supernova explosions can further complicate the spectrum, both directly and through their remnants. In this cases we talk of **starburst nuclei.**

♣ Optical spectra of the brightest nuclei cannot be explained in terms of normal or exceptional episodes of star formation. These are called **active galactic nuclei, AGNs**.

♣ AGN emission extends, as a matter of fact, to the whole electro-magnetic spectrum; the upper panel of figure 1 shows, as an example, the **Spectral Energy Distribution (SED)** of the famous 3c273 quasar, while the lower panel shows a schematic view of the various emission components. AGNs are classified in several ways, depending on the band in which they are observed. In the optical we have two parallel classifications. The first depends on the nuclear luminosity:

- We call **Seyfert** galaxies those that host an AGN with luminosity smaller than $M_V = -23$.
- More luminous AGNs are conventionally called **Quasi-Stellar Objects (QSO)**.

The difference between the two classes is historical: bright AGNs were seen as point-sources (historically identified as stars), both because they outshine their host galaxies and because these objects are absent in the local Universe. They are found at cosmological distances, where observing a galaxy is not trivial. Physically, Seyfert galaxies are just low-luminosity QSOs. QSOs are often loosely called quasars, a term that will be introduced later.

♣ A second classification is based on **emission lines** shown in the optical spectrum of AGNs (figure 2). These are of two main kinds; some are very broad: their width, interpreted in terms of Doppler effect, corresponds to velocities of the order of $\sim 10^4$ km s^{-1}. Other lines are relatively narrow, corresponding to motions of $\sim 10^3$ km s^{-1}. Broad lines are mostly permitted lines, of low and high ionization, narrow lines can be both permitted and forbidden, again of low and high ionization. AGN spectra can be divided in the following classes:

- **Type 1:** they show both broad and narrow emission lines, over a continuum emission in blue and UV bands.
- **Type 2:** they show only narrow emission lines.
- **LINER:** meaning Low-Ionization Narrow Emission Region; these AGNs only show narrow lines of low ionization.

♣ This classification also corresponds to a **luminosity sequence**, with type 1 nuclei being more luminous than type 2, and these more luminous than LINERs. In particular, almost all QSOs are of type 1, though we will later mention the existence of some type 2 QSOs. Seyfert

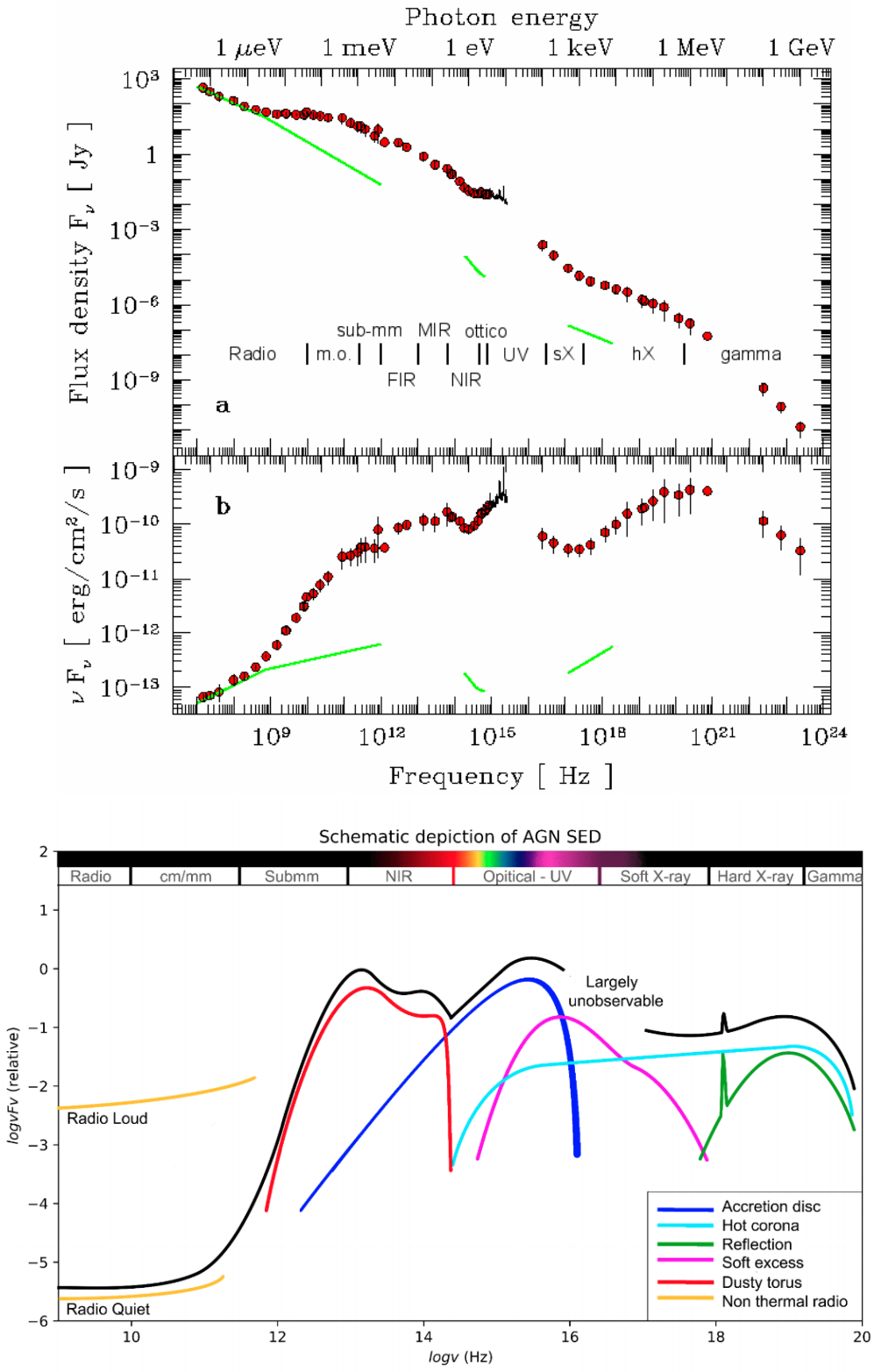

Figure 1: Upper panel: broad-band spectrum of quasar 3c273. Source: M. Turler et al. (1999). Lower panel: schematic view of AGN emission. Courtesy of F. Shankar.

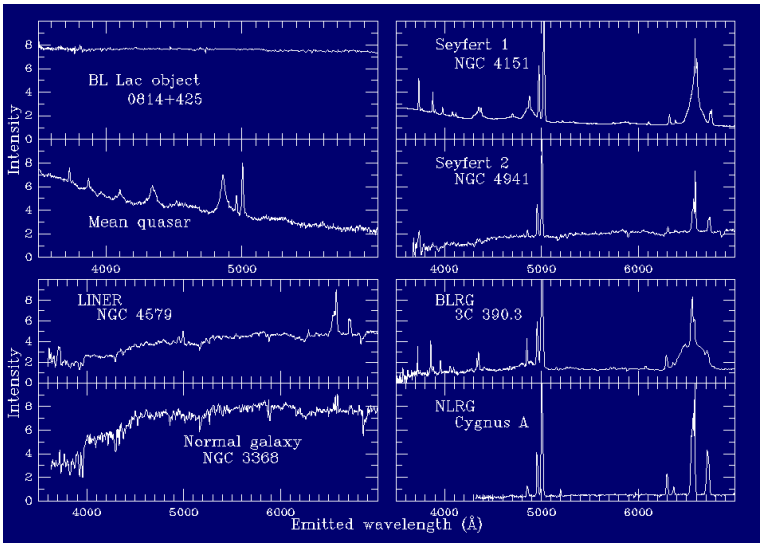

Figure 2: AGN spectra. Source: B. Keel.

can be of type 1 or 2, but intermediate types exist like, e.g., Seyfert of type 1.5. LINERs are the least luminous AGNs, and some of them could be explainable as anomalous starburst episodes.

♣ The main method to identify low-luminosity AGNs is the **BPT diagram** (after Baldwin, Phillips & Terlevich), which reports the equivalent width ratios $[OIII]/H_\beta$ and $[NII]/H_\alpha$ (figure 3). In place of $[NII]$ one can use another low ionization line like $[SII]$ o $[OI]$. HII regions and AGNs populate different regions of that diagram, as a consequence of the "hardness" of the ionizing continuum, i.e. the spectral slope in the UV, and of the ionization parameter U, i.e. the strenght of the ionizing radiation. This way, it is also possible to distinguish Seyfert from LINERs, given that the latters have lower equivalent widths of $[OIII]$, a higher ionization line.

♣ AGNs show **preference of morphological types**. At low redshift, Seyfert nuclei are found preferentially in early type spirals, Sa/Sb, but not in E/S0 galaxies, while LINERs are present in both classes. Late-type spirals, Sc/Sd, are typical hosts of starburst nuclei.

♣ AGNs are very visible at radio wavelengths, where stars do not emit significant amounts of radiation. The phenomenology of radio AGNs is very interesting and diverse. We call **radio-loud AGNs** the extragalactic sources with $L_{\rm radio} > 10^{41}$ erg s^{-1} (10^{34} W). These are conventionally divided into two classes, depending on optical identification:

- **quasar** (quasi-stellar radio sources) if the optical counterpart is point-like;
- **radiogalaxies** if the optical counterpart is not seen as a QSO but as galaxy.

♣ The radio manifestation of a radio-loud AGN (figure 4) is that of a **compact central source**, corresponding to the galaxy nucleus, from which two **radio jets** start, ending in **radio lobes**. Such structures can be as large as hundreds of kpc. It is worth noticing that almost all AGNs and starburst galaxies show radio emission, to a level lower than the 10^{41} erg s^{-1} limit; in these cases no radio lobes and jets are found.

♣ Radio-loud radio sources are classified as **Faranoff-Riley** I (FR I) sources if the jets are both visible and not well collimated, and the lobes are brighter at the center (*limb darkened*). FR

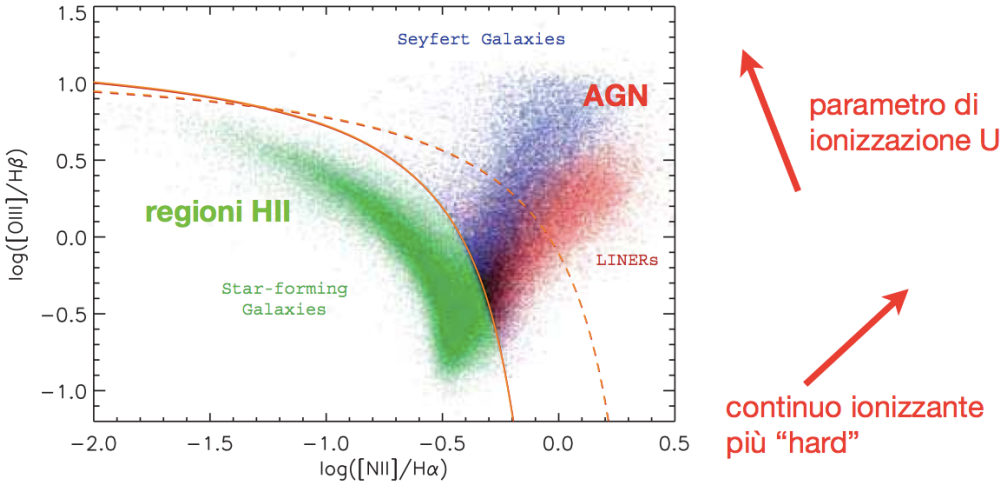

Figure 3: Example of BPT diagram from the Sloan Digital Sky Survey. Taken from the lectures of A. Marconi and G. Risaliti. "parametro di ionizzazione U": ionization parameter U; "continuo ionizzante più hard": harder ionizing continuum.

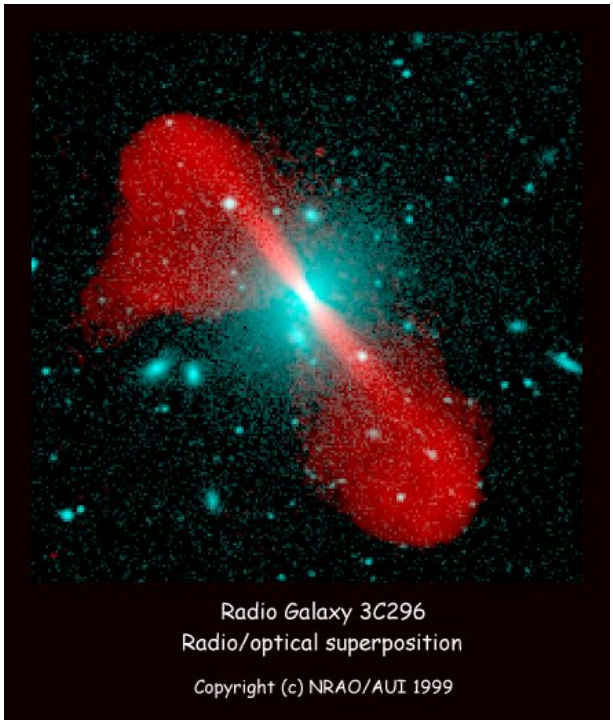

Figure 4: Radio and optical image of the radiogalaxy 3c296. Source: NRAO/AUI.

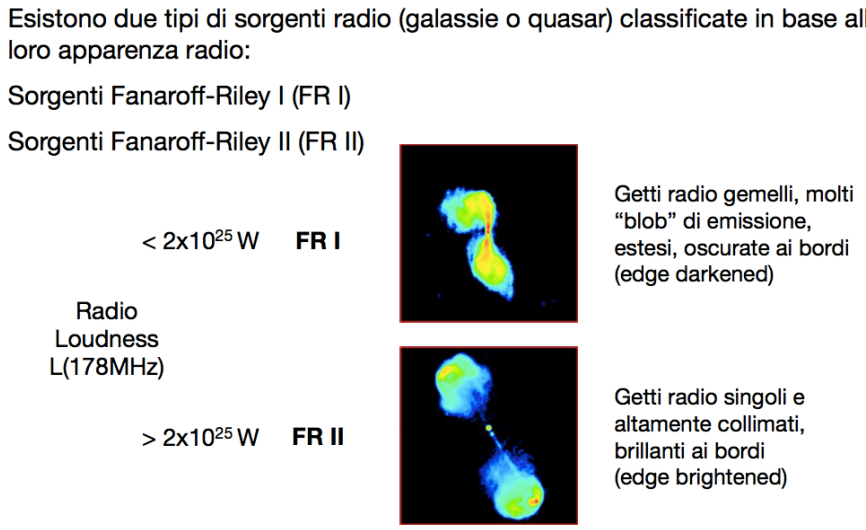

Figure 5: Faranoff-Riley classification of radio AGNs. Taken from the lectures of A. Marconi and G. Risaliti. The parts in italian correspond to what explained in the main text.

II sources, brighter than FR I, have more collimated jets, of which one is much more visible than the other as an effect of relativistic beaming. Lobes are brighter at the border (*limb brightened*) as an effect of the shock produced in the collision of the jet with the diffuse gas external to the galaxy (intra-group or intra-cluster medium).

♣ Radiogalaxies are very interesting physical systems. The magnetic energy contained in their lobes can amount to 10^{61} erg (10^{54} J)! Radio emission is due to synchrotron radiation from ultra-relativistic electrons, emitted by the central engine, that travel in a magnetic field. Detailed observation of radio lobes requires interferometric techniques, while to observe the details of the central compact source it is necessary to use the Very Long Baseline Interferometry technique. With this it is possible to observe, in correspondence of the jets, blobs that are in apparent **superluminal motion**. (figure 6). This evidence is interpreted as a relativistic effect: a jet of matter shoot along the line of sight at relativistic speed will appear to move on the sky at a larger velocity than the speed of light.

♣ **To demonstrated this**, let's consider a source at distance D that, at time t_0, emits a gas cloud at a velocity $v \sim c$ along a trajectory at an angle θ from the line of sight. The signal, emitted at time t_0 from the source, arrives at the observer at time $T_0 = t_0 + D/c$. At time t_1 the cloud will have walked a distance vt_1; a photon emitted by the cloud at that time will arrive to the observer at time $T_1 = (D - v(t_1 - t_0)\cos\theta)/c + t_1$. If the distance walked by the cloud perpendicularly to the line of sight is $\delta = v(t_1 - t_0)\sin\theta$, the apparent velocity of this motion will be $v_{\rm app} = \delta/(T_1 - T_0)$, or:

$$v_{\rm app} = \frac{v\sin\theta}{1 - \frac{v}{c}\cos\theta} \tag{1}$$

If $v \sim c$ and the angle θ (in radians) is small, we will have $v_{\rm app} \sim 2c/\theta$, that can easily be larger than c.

♣ A classification analogous to the optical one is used in **X rays**. Here AGNs are the brightest point sources, much brighter and harder than X-ray binaries of the whole host galaxy; these become visible only when the AGN is very weak or absent. The typical X-ray AGN

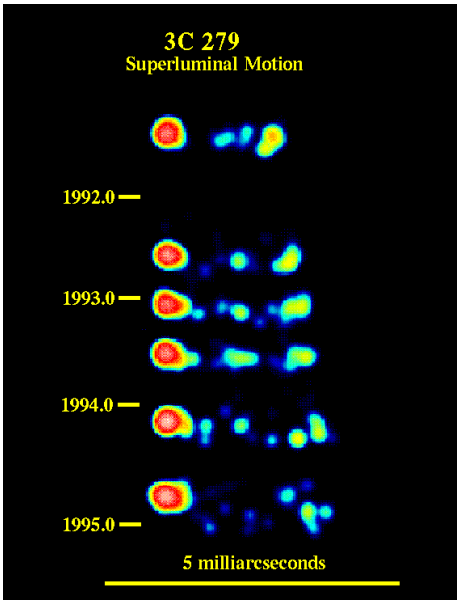

Figure 6: Superluminal motion of a radio source. Source: H.E. Smith.

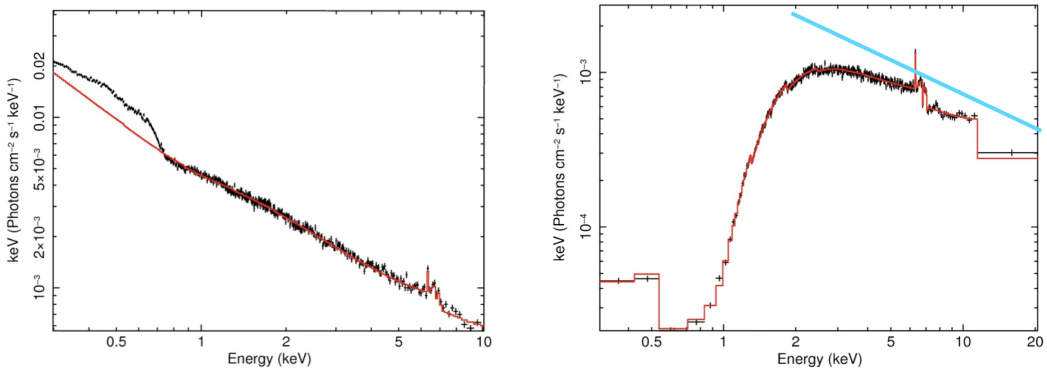

Figure 7: Typical X-ray spectra of AGNs: unobscured (left) and obscured (right). Taken from the lectures of A. Marconi and G. Risaliti.

spectrum (figure 7) can be represented by a power law, with an excess in the soft part. Figure 1 gives a schematic interpretation of the typical X-ray spectrum of an AGN; we will go in no detail here, but the presence of soft excess, hot corona and reflection components gives a hint of the complexity of the geometry of this hot gas. In other cases the spectrum can be represented by a power law attenuated in the soft part by photoelectric absorption. This attenuation allows to estimate the column density of the gas (typically in terms of N_H, the column density of hydrogen atoms) responsible for the attenuation. Small values of $N_H \lesssim 10^{21}$ cm^{-2} define type 1 AGNs, while large values define type 2 AGNs. An AGN with intrinsic X-ray luminosity of $> 10^{44}$ erg s^{-1} is called quasar, and quasars with $N_H > 10^{21}$ cm^{-2} are called type-2 quasars; few examples of this class of objects are known.

♣ Given that attenuation in the hard-X beyond 2 keV is relatively small, this band is a preferential one for detecting AGNs. But if $N_H > 10^{23}$ cm^{-2}, the object results obscured even in the hard-X. These objects, called **Compton thick**, could be absent in the catalogs selected both in the optical and in the X-rays. In this case, only mid-IR can reveal the presence of an obscured AGN, as will be illustrated below.

♣ AGN phenomenology presents a high degree of diversity, that is matched by further classifications and sub-classifications, that are beyond the purpose of these lectures. It is worth to mention another class of objects, named **BL Lac**, from the name of the prototype. The optical spectra of these objects are characterised by a continuum with no lines, by a highly variable luminosity and by a high degree of polarisation (figure 2). They are part of a larger class, the so-called **blazars**, that present similar characteristics of variability and polarisation. As we will see, these object can be interpreted as AGNs observed along the direction of a relativistic jet with a high Lorentz factor: the continuum is subject to a strong relativistic beaming that makes it so luminous to over-shine emission lines.

♣ We nowadays use call **quasar** all the AGNs with $M_V < -23$, or with $L_X > 10^{44}$ erg s^{-1}, even though only $\sim$10% of them are radio-loud. Quasars occupy an important place within astrophysics, and an even more important one in cosmology. They are to our knowledge the brightest compact and "long-lasting" objects (compared to human time scale), and they are visible to very long distances. They have long been the most distant known objects, though we now find even more distant galaxies. Their association with galaxies has been for a long time a (sometimes criticised) assumption, but know it is proven by several observations. In particular, quasars appear to be preferentially **hosted by elliptical galaxies**. Figure 8 shows the host galaxy of the famous quasar 3c273, after an accurate subtraction of the point quasar source, obtained thanks to ACS on Hubble Space Telescope; the central hole is clearly due to the subtraction.

♣ AGNs are associated with galaxy nuclei, yet they are much less numerous. This may mean that only a small fraction of galaxies host AGNs, or that AGN activity is common to all galaxies but intermittent, with short duty cycle time with respect to the Hubble time. It will be argued below that the second interpretation is correct. It is possible to estimate the duty cycle time of an AGN using the concept of **Eddington luminosity**. We assume here that AGN energy, whatever is its source, is carried in by gas infall onto a central object of mass M and luminosity L. We put this radiation source at the center of a reference frame, and assume spherical symmetry. The force per unit mass exerted by radiation pressure to matter at radius r is:

$$f_{\rm rad} = \frac{\kappa}{c}\frac{L}{4\pi r^2} \tag{2}$$

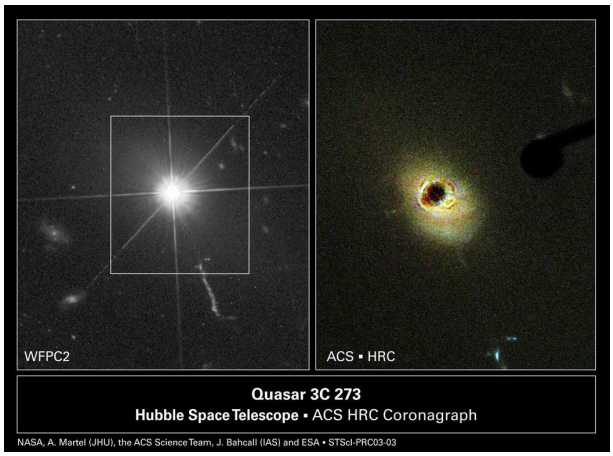

Figure 8: Host galaxy of the quasar 3c273, after accurate subtraction of the central source. Source: Hubble Space Telescope.

where κ is the gas opacity. The gravitational force per unit mass is $f_{\rm grav} = GM/r^2$. Gas can infall toward the luminous object as long as gravity is stronger than radiation pressure force, but when radiation pressure overtakes gravitational attraction the gas is pushed away from the object; without accretion the object will quickly dim. Assuming that opacity is given by the Thomson scattering, that is an unavoidable source of opacity if the gas is highly ionized (very likely with such a hard source as an AGN), and assuming that gas composition is dominated by Hydrogen, then $\kappa = \kappa_T = \sigma_T/m_H$ (where σ_T is the Thomson cross section and m_H the mass of the Hydrogen atom, very similar to the proton mass). Both force terms scale with $1/r^2$, so their balance does not depend on the distance from the luminosity source. A condition that radiation force counterbalances gravity can be translated into a condition on luminosity, the Eddington luminosity:

$$L_{\rm Edd} = \frac{4\pi GcMm_H}{\sigma_T} = 1.25 \times 10^{38} \left(\frac{M}{M_\odot}\right) \ {\rm erg\ s^{-1}} \tag{3}$$

where the second equation is obtained by assuming $M = M_\odot$ and by making the scaling with mass explicit.

♣ If the AGN source shines at $L > L_{\rm edd}$, then radiation pressure will quench further mass accretion. A source with a large reservoir of gas, ready to fall on it, will thus tend to shine at the Eddington luminosity. Under the hypothesis that the central engine has an energy efficiency η, we have that $L = \eta \dot{M} c^2$, where $\dot{M}$ is the mass accretion rate onto the central source. One can then define an **Eddington accretion rate**:

$$\dot{M}_{\rm Edd} = \frac{4\pi G m_H}{\eta c \sigma_T} M \tag{4}$$

It is easy to see that in this case the mass of the central object will grow exponentially, $M(t) = M(t_0)\exp(t/t_{\rm Edd})$, where the **Eddington time** is:

$$t_{\rm Edd} = \frac{M}{\dot{M}_{\rm Edd}} = \frac{\eta c \sigma_T}{4\pi G m_H} \tag{5}$$

The Eddington time does not depend on object mass, and can be interpreted as the typical time-scale of a shining AGN. Its value depends on knowledge of the radiation efficiency η; we will find later that $\eta \simeq 0.1$ for mass infall onto a black hole, in which case

$$t_{\rm Edd} \simeq 4.5 \times 10^7 \text{ yr} \tag{6}$$

We will assume this value as a fiducial time for the duration of AGN duty cycle.

♣ **Where does AGN energy come from?** A quasar luminosity can easily get to 10^{13} $L_\odot$, and this luminosity is sustained for $\sim 4.5 \times 10^7$ yr. In this period the quasar radiates an energy of $\sim 5 \times 10^{61}$ erg; this is of the same order than the energy contained in the radio lobes of a radiogalaxy, giving support to the validity of the assumed duty cycle time. This energy corresponds to an equivalent mass of $\sim 3 \times 10^7$ $M_\odot$. The largest achievable efficiency of nuclear processes amounts to 10^{-3} (or 10^{-2}, due to nuclear burning from hydrogen to iron, on 10% of the star mass), so if the quasar energy were due to stellar nuclear fusion we should have $\sim 3 \times 10^{10}$ $M_\odot$ of stars that shine in an extremely compact region. If the energy came from supernova explosions, assuming $\sim 10^{51}$ erg per supernova we would need $\sim 5 \times 10^{10}$ supernovae in the same compact region. In a normal stellar population we form a star of 8 $M_\odot$ (or larger) each ~ 100 $M_\odot$ of stars, so the need for stellar mass would amount to $\sim 5 \times 10^{12}$ $M_\odot$. A smaller mass (for not more than a factor of ten) can be obtained under the assumption that only massive stars form. Such mass of stars (or remnant neutron stars) packed in the limited volume of a galaxy nucleus, beyond being problematic from several points of view (it is unstable for collapse into a black hole), would easily be revealed, if not by its emission, through its dynamical effect on the nearby stars or gas. **To explain quasars we need a new mechanism to produce energy**, more efficient than nuclear fusion or supernova explosions.

♣ A very good explanation for this energetic mechanism is given by assuming that the central engine of the AGN is a **supermassive black hole**, with a mass of $\sim 10^6 - 10^{10}$ $M_\odot$. Indeed, matter falling onto the black hole heats up and emits radiation. According to the virial theorem, half of the gravitational energy acquired from gravitational contraction becomes gas thermal energy, while the other half is radiated away. Let's consider a mass m of gas that falls onto a black hole of mass M, whose **gravitational radius** is $R_{\rm G} = GM/c^2$ (in geometrical units with $G = c = 1$, it is simply M). If the black hole has no rotation, the Schwartzschild metric is valid in its neighbourhood and the event horizon is at the Schwartzschild radius $R_{\rm s} = 2R_{\rm G}$. If this mass orbits at a distance $r_{\rm min} = a \times R_{\rm G}$ (without directly falling into the black hole) for a sufficiently long time, it will radiate an energy:

$$E \sim -\frac{1}{2}\Omega = \frac{1}{2}\frac{GMm}{r_{\rm min}} = \frac{mc^2}{2a} \tag{7}$$

So the efficiency of energy production will be $\eta \sim 1/2a$. From the study of stable orbits around black holes, that will be discussed below, it results that $a \sim 3 - 6$, so $\eta \sim 0.1 - 0.2$, an order of magnitude larger that the largest efficiency of nuclear reactions. It is worth noticing that this efficiency does not depend on black hole mass.

♣ If the efficiency of energy production is $\eta = 0.1$, a quasar of 10^{13} $L_\odot$, radiating $\sim 3 \times 10^7$ $M_\odot$ of mass-energy, will leave behind a black hole of $\sim 3 \times 10^8$ $M_\odot$. The gravitational radius of such black hole is of $\sim 1.5 \times 10^{-5}$ pc (roughly 3 AU). Such an object, if no more gas falls on it, would be **practically invisible** in the nucleus of a big elliptical galaxy. The way to identify these "dormant" black holes will be discussed in the following. We lastly notice that if the activity cycle of a quasar is of 4.5×10^7 yr, the accretion rate required to sustain the above mentioned luminosity is of ~ 8 $M_\odot$ yr^{-1}.

♣ AGN activity is thus explained as accretion of matter onto supermassive black holes at galaxy centers. The first intuition of what a black hole can be is due to **Laplace and Mitchell**,

who speculated that a very compact object could have an escape velocity $V_{\rm esc} = \sqrt{2GM/r}$ larger than the speed of light c, so that light, if subject to gravity, could not get to infinite distance. Though this reasoning has some obvious limitations due to the fact that it ignores special and general relativity, it allows to recover the Schwartzschild radius by setting $V_{\rm esc} = c$: $r = R_{\rm s} = 2GM/c^2$. In general relativity, a black hole is defined as a singularity in spacetime, a point mass of infinite density, that is "dressed" by an event horizon.

♣ One of the first exact solutions of Einstein equations was obtained for a point mass of mass M in a static universe. The solution for the metric is due to **Schwartzschild**, the line element of this metric, in a spherical coordinate system, is:

$$ds^2 = -\left(1 - \frac{R_s}{r}\right) c^2\, dt + \left(1 - \frac{R_s}{r}\right)^{-1} dr^2 + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right) \tag{8}$$

where $R_{\rm s} = 2R_{\rm G}$ is the Schwartzschild radius defined above.

♣ This metric is singular not only for $r = 0$ but also for $r = R_s$. The surface at $r = R_s$ is called **event horizon**, because there are no possible trajectories for particles that can connect, in the future direction, events inside the horizon with events outside it: all particles within the event horizon can only move on trajectories that end into the central singularity. Event horizon then acts like a semi-permeable barrier: matter, energy and information can only flow from the outside toward the inside, not vice versa.

♣ Radiation coming from a source at distance $r > R_{\rm s}$ from the black hole is subject to **gravitational redshift**:

$$z = \frac{1}{\sqrt{1 - R_{\rm s}/r}} - 1 \tag{9}$$

An observer at infinity will see the radiation coming from a source approaching the event horizon more and more redshifted, and thus less and less energetic, until it becomes unmeasurable, thus making the black hole **"black"**. At the same time the distant observer will see the source take an infinite time to get to the event horizon. However, the source falling towards the central singularity will pass through the event horizon at some finite time according to its clock. This implies that it is possible to find coordinate systems where the singularity at $r = R_s$ is not present, while the central singularity cannot be removed by a coordinate transformation.

♣ At large distance from the event horizon it is possible to give this **interpretation of gravitational redshift** $z = (\lambda - \lambda_e)/\lambda_e$: if the energy of a photon is $e = hc/\lambda$ and it its "equivalent mass" is $m = e/c^2$, to first order in a Taylor expansion it is possible to find:

$$e = \frac{hc}{\lambda} = \frac{hc}{\lambda_e}\left(1 - \frac{1}{2}\frac{R_s}{r}\right) = \frac{hc}{\lambda_e} - \frac{GMm}{r} \tag{10}$$

In other words, the energy lost by the photon because of gravitational redshift can be interpreted, to a first approximation, as that necessary to climb the potential well.

♣ The **"hole"** nature of the black hole comes from the following fact. A point mass is a singularity also in Newtonian mechanics, but it is impossible that two point masses collide. Indeed, it is well known that the two-body problem of two point masses subject to mutual gravitational attraction can be recast into an equivalent problem of a test point mass subject to a central force:

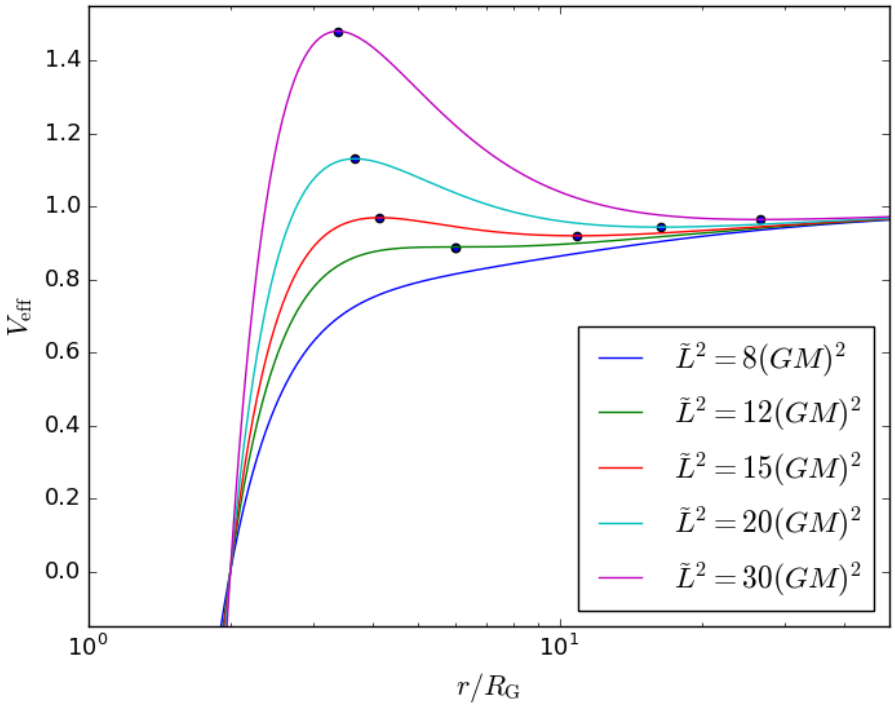

Figure 9: Effective potential for a Schwartzschild black hole, and for various values of the angular momentum $\tilde{L}$. Dots denote the positions of the potential maximum and minimum; they coincide for $\tilde{L}^2 = 12(GM)^2$, while there are no extrema for lower $\tilde{L}$.

$$\frac{1}{2}\left(\frac{dr}{dt}\right)^2 + V_{\rm eff}(r) = E \tag{11}$$

where r is the distance between the two masses, E the total energy and the effective potential $V_{\rm eff}$ is:

$$V_{\rm eff}(r) = -\frac{GM}{r} + \frac{h^2}{2r^2} \tag{12}$$

where M is the reduced mass of the system and h its specific angular momentum, that depends on the impact parameter. The second term acts as a centrifugal barrier and prevents the two point masses to collide (generating infinite energy) in all cases in which $h > 0$, that is for any non-vanishing impact parameter. The problem of the orbit of a test mass around a black hole can be expressed through a similar equation (see Cosmology I notes, lecture 10):

$$\left(\frac{dr}{dt}\right)^2 = \tilde{E}^2 - V_{\rm eff}(r) \tag{13}$$

(in Cosmology I $V_{\rm eff} = \tilde{V}^2$) but the effective potential becomes (we use $c = 1$ here):

$$V_{\rm eff}(r) = \left(1 - \frac{2GM}{r}\right)\left(1 + \frac{\tilde{L}^2}{r^2}\right) \tag{14}$$

(figure 9), where the conserved quantity $\tilde{L}$ plays the same role as the specific angular momenum h. It is easy to demonstrate that the slope of this potential gives the direction of acceleration: negative slope means repulsion, positive slope means attraction. At small distances the further

term $\propto r^{-3}$ dominates, and this in principle allows particles to fall to $r = 0$ even when $\tilde{L} > 0$. By differentiating $V_{\rm eff}(r)$ it is easy to see that the effective potential has two extrema at:

$$r = \frac{\tilde{L}^2}{2GM}\left[1 \pm \sqrt{1 - 12\left(\frac{GM}{\tilde{L}}\right)^2}\right] \tag{15}$$

This is true provided that $\tilde{L}^2 > 12(GM)^2$ (see figure 9). In this case the potential allows for closed stable orbits. Loss of angular momentum, will drive $\tilde{L}$ toward $\sqrt{12}GM$, where we have the **last stable orbit** at $r = 6R_g = 3R_s$, as anticipated above.

♣ Because of the "no-hair theorem", the only information that survives collapse into a black hole is mass, angular momentum and charge. Assuming that this last quantity is null in cases of astrophysical interest, the most general case is that of a rotating black hole. The solution of Einstein's equations leads this time to the **Kerr metric**, whose phenomenology is more complex and is beyond the interest of this course. It is however important to know that in this case the last stable orbit moves from 6 toward 3 gravitational radii (for a maximally rotating black hole), thus making the radiative efficiency of accretion discs higher.

♣ As the formation of stars is limited from the angular momentum of proto-stellar gas, that must be dissipated for the gas to concentrate sufficiently, the infall of gas on the black hole is limited by gas **angular momentum**. In this case the required loss of angular momentum is extreme: gas must lose at least 15 orders of magnitude of it to get to distances comparable to the black hole gravitational radius. Because the loss of angular momentum regulates the infall of gas, this will settle on an **accretion disc** around the black hole. The viscosity that is generated in the differential rotation of the accretion disc will allow the gas to lose the residual angular momentum until it falls onto the black hole. The infall takes place when the gas reaches the **last stable orbit**. The accretion disc heats up due to gravitational infall (and through the viscosity mechanism) and emits thermal radiation. However, because its temperature is not constant, its continuum spectrum will differ from that of a black body at a single temperature. We further notice that, being the inner parts of the accretion disc hotter, main emission at shorter wavelengths will come from the inner regions. In other words, we will observe the inner parts of the disc in the X-rays, the outer parts in the blue and near-UV.

♣ Despite the rich diversity of their phenomenology, the main AGN characteristics can be broadly understood in terms of a **unified model**. This includes the following components: (figure 10):

- A **central engine**, composed by a black hole surrounded by an accretion disc. If the black hole rotates, this engine will emit jets of highly relativistic plasma from the poles, corresponding to the rotation axis.
- A region, named **broad line region** (BLR), of dimension of $\lesssim 0.01$ pc, composed of gas clouds in very quick motion, $\sim 10^4$ km s^{-1}, and (relatively high) density and temperature such to generate low and high ionization permitted lines of type-1 AGNs.
- A toroidal region composed of **gas and dust**, of dimension of $\sim$1 pc, that surrounds the central engine and the BLR, obscuring their emission in a large solid angle.
- A region, named **narrow line region** (NLR), of dimension of $\sim$100 pc, composed of warm clouds, similar to the BLR but with the clouds in slower motion and lower density so as to generate forbidden lines.

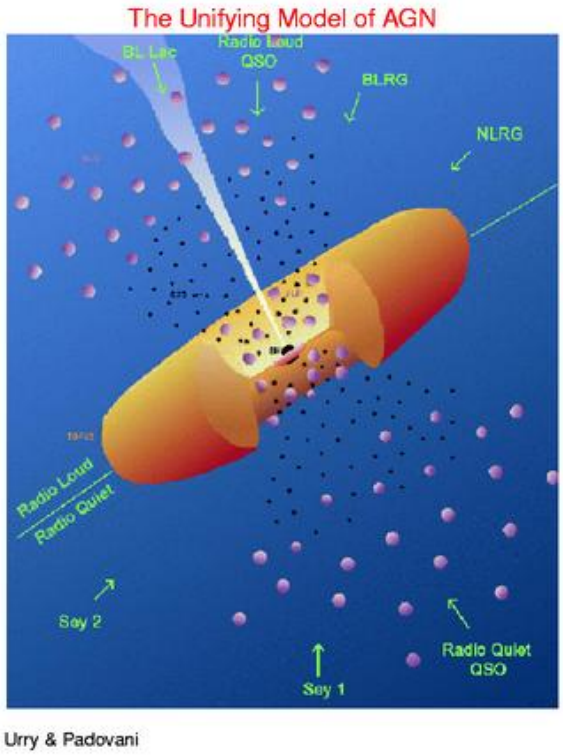

Figure 10: Scheme of the unified model of AGNs. Source: M. Urry & P. Padovani (1995).

All the distances given here scale with the black hole mass, and are referred to $M_{\rm BH} \sim 10^8\ M_\odot$. Such a system, if observed along a line of sight near to the polar one, avoiding the obscuring torus, will show a continuum from blue to X due to thermal emission of the accretion disc, the broad emission lines of the BLR and the narrow ones of the NLR; it will then appear as a type-1 AGN. Observed instead along a line of sight that crosses the obscuring torus, the same system will show only narrow lines, thus appearing as a type-2 AGN. The jets of relativistic plasma and the magnetic field associated to them would give rise to radio emission. Finally, the few AGNs observed along the polar jets will appear as blazars, because relativistic beaming will amplify the continuum emission, so as to over-shine broad and narrow emission lines.

♣ Getting back to figure 1, it is now possible to **interpret the SED** of an AGN (in terms of $\nu F(\nu)$) in the light of the unified model. It is possible to identify three main components, the first from the radio to the NIR, with a peak in the MIR; the second, named Big Blue Bump, from the optical to the soft-X, peaking in the UV; the third in the hard-X and gamma, with a peak at very high energies. A note: in the upper panel data in the extreme UV are not available, because in this band the absorption of ionizing photons from neutral hydrogen makes the Milky Way opaque; this does not allow to see the peak of the Big Blue Bump in the local universe; in the lower panel the "largely unobservable" region is shifted to higher frequencies, on the basis that most quasars are observed at cosmological distances and their SEDs are thus redshifted. Starting from radio, the IR component peaking in the MIR is associated to the thermal emission of dust in the torus, warmer (~ 100 K) than typical interstellar dust of normal or even starburst galaxies; this emission allows in principle to find even the most obscured, Compton thick AGNs. The Big Blue Bump is associated to the emission of the accretion disc, even though the details are complex. The narrow and broad lines are (barely) visible in the optical/UV spectrum. The gamma peak, that can contain an important fraction of the object total energy, is interpreted as inverse Compton emission from ultra-relativistic electrons ejected from the central engine.

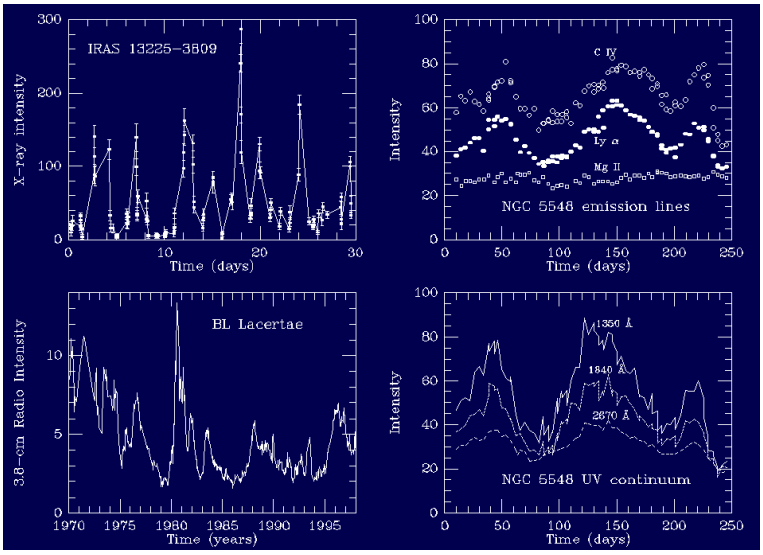

Figure 11: Variability of AGNs in various bands. Source: B. Keel.

♣ A fourth component, the radio emission from the central source discussed above, plays here a minor energetic role (radio emission from the jets and lobes is much more extended and is not represented in the nuclear SED). It is important to point out that the energy source for the emission of jets may not be accretion: in fact, thanks to the **Blandford-Znajek** mechanisms, magnetic fields entangled with the accreting gas are able to extract energy from the rotation energy of the black hole.

♣ We will see that dormant supermassive black holes in local galaxies have a preference for elliptical galaxies, it makes sense to ask **why in nearby galaxies we see AGNs mainly in (early) spiral galaxies**. A black hole alone is not a necessary condition to have nuclear activity, some gas must fall on it, and this accretion rate must be not too small with respect to the Eddington accretion rate of the black hole. This can happen in early-type spirals, rich of interstellar gas and with a sufficiently large bulge to contain a black hole of $\sim 10^6 - 10^7\ M_\odot$, but it is more difficult in elliptical galaxies, that contain much more massive black holes but are poor of gas. However, radio activity, much more powerful than that of Seyfert galaxies, is limited to some elliptical "radiogalaxies". These host much more massive black holes, so in case some gas falls on it, accretion takes place at rates much lower than the Eddington rate. In this case an accretion disc cannot form, and the infall of matter onto the black hole is characterised by a much lower energy efficiency, $\eta \sim 10^{-4}$. However, extraction of black hole rotation energy performed by magnetic fields of accreting matter can lead to emission of jets of plasma, and give rise to the phenomenology of radiogalaxies.

♣ The unified model, based on the presence of a black hole, explains well many aspects of AGN phenomenology. But what are the **proofs of existence of supermassive black holes** powering AGNs?

- The first argument is based on the **efficiency of energy generation**, and has already been presented: to have $\eta \sim 0.1$, if we exclude the process of matter-antimatter annihilation, the only known process is the infall of matter into a black hole.
- We have also seen that the phenomenology of radiogalaxies and blazars (over all superluminal apparent motion) suggests the presence of **relativistic motion** of macroscopic

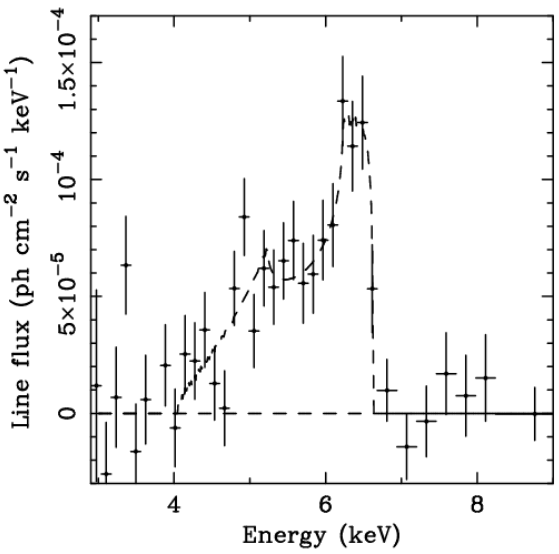

Figure 12: Iron line in X-ray of an AGN, and fit based on an accretion disc model. Source: A. Fabian et al. (2000).

amounts of matter; star formation activity, be it normal or starburst, is not able to generate such motions.

- The presence of **giant radio lobes** requires that the emission of plasma jets is stable in direction for at least $\sim 10^6$ yr. This stability is obtained naturally assuming the presence of a rotating supermassive black hole.

- The presence of **variability on short time scales** (figure 11) suggests the presence of compact objects. At relatively short wavelengths like UV or X-rays (that presumably come from the inner parts of the AGN engine), it is possible to measure significant variations of AGN luminosity on time-scales that can be, in extreme cases, as brief as 30 minutes. For the luminosity to change this rapidly, information must have time to propagate through the whole source. Because information travels at a speed no larger than c, a source of dimension D will vary on time-scales not shorter than $\Delta t \sim D/c$. If $\Delta t \sim 30$ minutes, $D \lesssim 4$ AU, of the same order of the Schwartzschild radius of a typical supermassive black hole.

- Any astrophysical object more massive than $\sim 2-3\ M_\odot$ (Oppenheimer-Volkoff limit) whose size is comparable to the Schwartzschild radius is **unstable for collapse into a black hole**. This is because in the relativistic regime pressure is a source of gravity (of curvature), so the body has no means to resist its own gravity. This makes all astrophysical alternatives to a supermassive black hole not very attractive.

- The most direct proof of strong gravity in the central engine of AGNs is the observation of the **fluorescence Kα like of Fe** (at $E \sim 6.5$ keV), that is visible in the X-rays. This is distorted in a very peculiar, asymmetric way (figure 12). Its shape is expected if this line is emitted by the hot inner parts of an accretion disc, that are subject to both Doppler and gravitational redshift. According to the argument given above, the presence of strong gravity makes a black hole the natural (and hardly avoidable) candidate for the central mass concentration.

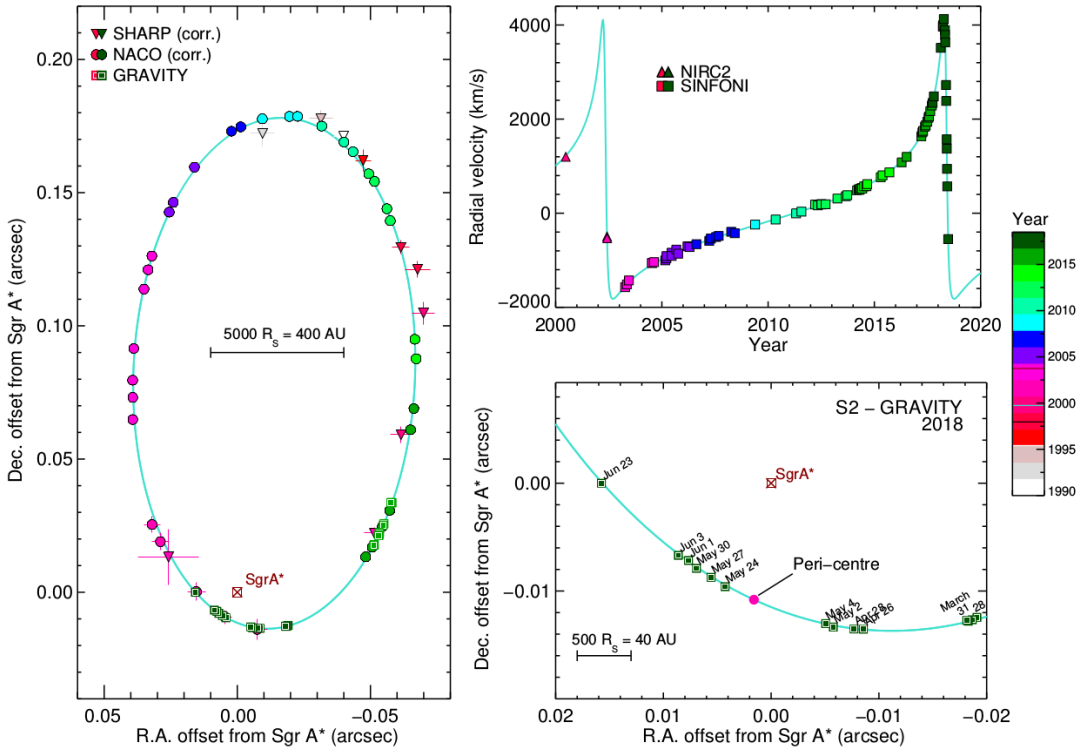

Figure 13: Orbit of star S2 around the SgrA* object at the Galaxy center. Source: GRAVITY collaboration 2018, A&A 615, L15.

♣ After decades of astronomical observations, the case for black holes powering AGNs has become very strong and **widely accepted**. The detection of gravitational waves, the observations of SgrA* and of M87 by the Event Horizon Telescope have given further support to the actual existence of black holes. Yet, a definitive proof of the existence of these object is still elusive: one needs to demonstrate that the object does not have a surface but an event horizon. However, alternatives to supermassive black holes are rather exotic, black holes offer the simplest explanation.

♣ Presently, the most convincing proof of the existence of a black hole is given by the observations of our **Galactic center**. Observing it in the radio, it is possible to reveal a point source in radio and X-rays, called SgrA*, that is associated to a potential black hole of mass $\sim 4 \times 10^6\ M_\odot$. Observing the Galactic center in the near IR, it is possible to minimise the extinction due to interstellar dust and observe a cluster of very massive and bright stars that orbits around the central object. Thanks to the adaptive optics of the instruments NAOS and GRAVITY, mounted on the VLT, it has been possible during 26 years to track the trajectory of some of these stars. It is well known that the measurement of stellar proper motion and radial velocity allows to completely determine the properties of a binary system, where in this case, due to the huge mass difference, the star can be considered as a simple test particle. In particular, it has been possible to determine with accuracy the mass of the central object, $(4.10 \pm 0.03) \times 10^6$ $M_\odot$, together with the distance from the galactic center, 8122 ± 31 pc. The reconstruction of the orbit of the S2 star, that passes at just 5.8×10^{-3} pc from the central object, allows to constrain the mass within that radius. This allows to exclude most viable alternatives to a black hole, with some exotic exceptions.

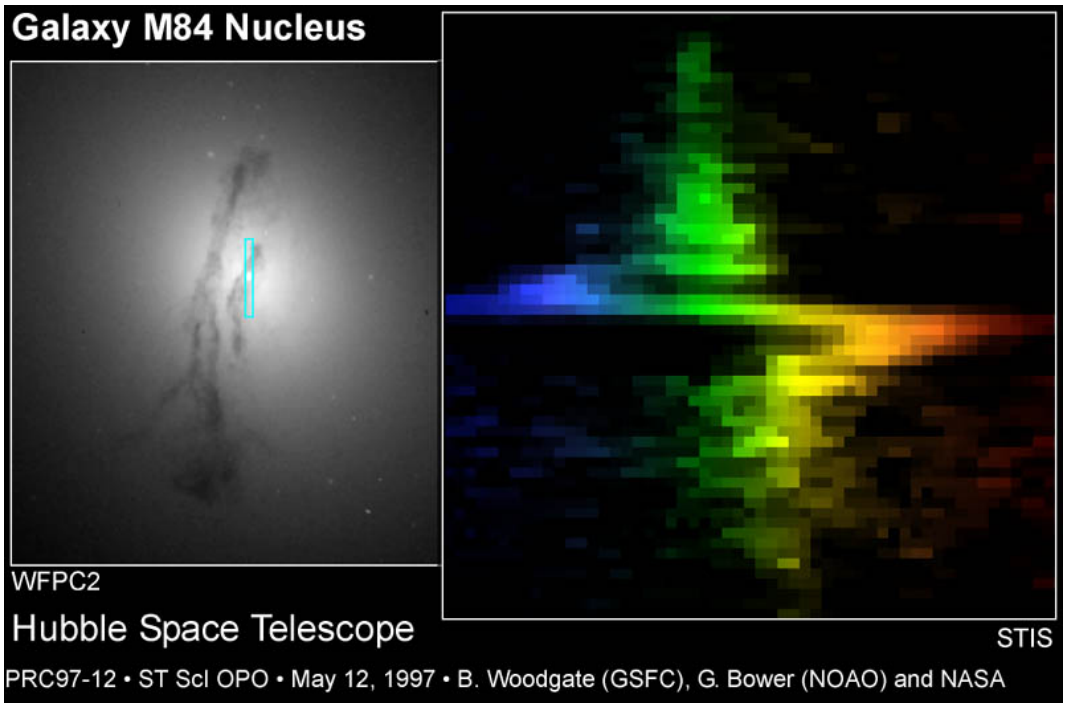

Figure 14: "Signature" of a massive dark object: Keplerian rotation curve of a nuclear disc. Source: HST.

♣ The S2 star passes through the pericenter at a velocity of 7650 km s^{-1}. Spectroscopy with the GRAVITY instruments has allowed astronomers, in 2018, to show departures from a Newtonian orbit at the pericenter, that are perfectly explained as post-Newtonian effects due to General Relativity.

♣ To get an (almost!) final killing argument on the nature of such massive and dark objects, astronomers have conceived the **Event Horizon Telescope**, a VLBI observational campaign at mm wavelengths that allows to resolve the projection of the photon capture radius of some of the most massive and nearby black holes. The collaboration has targeted SgrA* and the nucleus of M87, called M87*; this galaxy, at the center of the Virgo cluster, is well known for showing a relativistic jet, due to accretion onto a 6×10^9 $M_\odot$ putative black hole. The strong curvature of spacetime near a black hole produces a dark shadow surrounded by an observable bright ring of photons orbiting just outside the event horizon. An image of the photon ring of M87* has been published in April 2019, and is reported on the cover of these notes; see the hand-written EHT notes for more details.

♣ What happens to **remnant black holes** of AGN activity, when they stop accreting mass? The answer has been found at the end of last century. In some relatively lucky cases, like M84, NGC 4258 or M87, it is possible to observe, at the center of a passive galaxy, a ring of gas in Keplerian rotation (figure 14) that reveals the presence of a **massive dark object** (MDO). The rotation curve can be used to measure the MDO mass. More in general, detailed observation of the nuclei of nearby (passive) elliptical galaxies and bulges of spiral galaxies reveal small cusps in the surface brightness profiles, accompanied by increased stellar velocity dispersion. Such cusps might be due to the presence of an MDO that distorts stellar orbits, making them concentrate around the nucleus with correspondingly higher speed. This effect is modest, and could in principle be due to (not implausible) peculiarities in stellar orbits. Thanks to an observational campaign based on Hubble Space Telescope images and ground-based spectroscopy, it has been

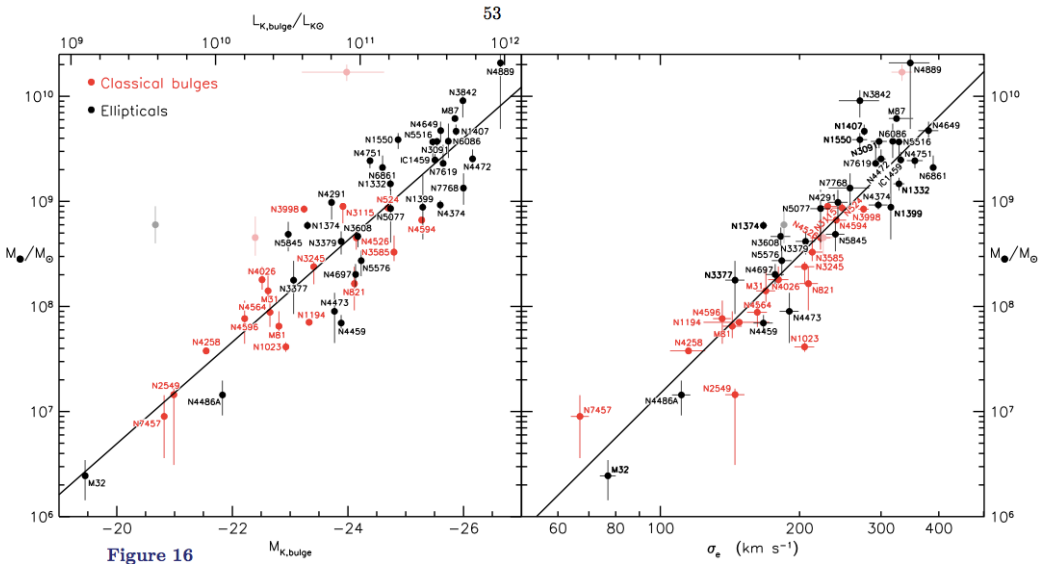

Figure 15: Relation between massive dark object mass and stellar mass (on the left, the latter being quantified through K-band luminosity) or central velocity dispersion (right panel) of the host bulge. Source: Kormendy & Ho, 2013.

possible in the late '90s to measure the presence of such cusps in a large number of nearby galaxies with significant bulge component. From these, it is possible to estimate the MDO mass. The surprising result is that MDO mass results well correlated with the stellar mass of the spheroidal component of the host galaxy, and even better with its central stellar velocity dispersion $\sigma_\star$ (figure 15).

♣ It is natural to interpret MDOs as the remnant black holes of quasar/AGN activity, observed in a quiescent phase. The correlation of bulge and black hole (MDO) mass has important implications:

- All the galaxies with a significant bulge component have passed through a phase of quasar/AGN activity.
- Quasars, that correspond to the most massive MDOs, should be associated to elliptical galaxies.
- The proportionality between the mass of the host bulge (or galaxy) and black hole mass strongly suggests that quasar activity has taken place **during the formation of the galaxy.**

♣ If $P(M_{\rm BH}|L_{\rm bul})$ is the relation that connects bulge luminosity and black hole mass, and if $\Phi(L_{\rm bul})$ is the bulge luminosity function (obtained by knowing the average fraction of light that is associated to the bulge component of galaxies), it is possible to readily obtain the **black hole mass function**, even though black hole estimates are available only for a small (but possibly unbiased) set of black holes:

$$\Phi(M_{\rm BH}) = \int_0^\infty P(M_{\rm BH}|L_{\rm bul})\Phi(L_{\rm bul})dL_{\rm bul} \tag{16}$$

The resulting black hole mass density is:

$$\rho_{\rm BH} = \int_0^\infty M_{\rm BH}\Phi(M_{\rm BH})dM_{\rm BH} \simeq 3 - 5 \times 10^5 \ {\rm M}_\odot \, {\rm Mpc}^{-3} \tag{17}$$

♣ Alternatively, one can use the central stellar velocity dispersion $\sigma_\star$ in place of the bulge luminosity. The two estimates will be consistent if the sample of measured black hole masses is representative of the full population. However, recent work has showed some inconsistencies due to a potential bias. Indeed, measuring black hole masses from MDO's requires to resolve the **black hole influence radius**, defined as:

$$r_{\rm BH} = \frac{GM_{\rm BH}}{\sigma_\star^2} = 10.7 \ {\rm pc} \left(\frac{M_{\rm BH}}{10^8 \ {\rm M}_\odot}\right) \left(\frac{\sigma_\star}{200 \ {\rm kms}^{-1}}\right)^{-2} \tag{18}$$

This translates into an angular radius of:

$$\theta_{\rm BH} = 0.11'' \left(\frac{M_{\rm BH}}{10^8 \ {\rm M}_\odot}\right) \left(\frac{\sigma_\star}{200 \ {\rm kms}^{-1}}\right)^{-2} \left(\frac{D}{20 \ {\rm Mpc}}\right)^{-1} \tag{19}$$

barely resovable by Hubble Space Telescope for the fiducial values used in the formula. This open issue keeps the uncertainty on the black hole mass density high.

♣ It is possible to connect the radiation emitted by quasars with the mass accumulated in remnant black holes through the so-called **Soltan argument**. A quasar emits light at a luminosity L as a result of matter accretion at a rate $\dot{M}_{\rm acc}$, such that, for each shining quasar:

$$L = \eta \dot{M}_{\rm acc} c^2 \tag{20}$$

$$\dot{M}_{\rm BH} = (1-\eta)\dot{M}_{\rm acc} \tag{21}$$

where η is the efficiency of radiation discussed above. Then, the total bolometric luminosity density of quasars integrated along the cosmic time can be related to the the mass density of remnant black holes:

$$\rho_{\rm BH} = \frac{1-\eta}{\eta c^2} \int_0^\infty dz \int_0^\infty dL_Q L_Q \Phi(L_Q; z) \left|\frac{dt}{dz}\right| \tag{22}$$

It is important to notice that the quasar bolometric luminosity function $\Phi(L_Q; z)$ is computed from an observed luminosity function, typically in the hard-X rays, correcting both for the fraction of bolometric light that is emitted in that specific band and for AGN obscuration. The black hole mass density so obtained is in good agreement, for reasonable values of $\eta \sim 0.1$, with the mass density obtained from local estimates of black hole mass density.

♣ If **quasars shine at Eddington luminosity**, or at a known fraction λ of it, then one can set a one-to-one correspondence between bright quasars and massive black hole; this is valid under the hypothesis that merging of black holes, that quickly follows the merging of their host galaxies, is negligible. It is then possible to work out a prediction for the expected mass function of remnant black holes, leftovers of the past quasar activity (figure 16). The match of the two determinations of the mass function allows to recover the two parameters of Eddington ratio λ and radiative efficiency η. Agreement has been found by many authors for very reasonable values of the two parameters.

♣ Supermassive black holes imply the presence of **seed black holes** at the beginning of their history. These may be formed by the collapse of massive stars at high redshift, possibly PopIII stars, a mechanism that limits their mass $< 1000 \ M_\odot$. The presence of some bright

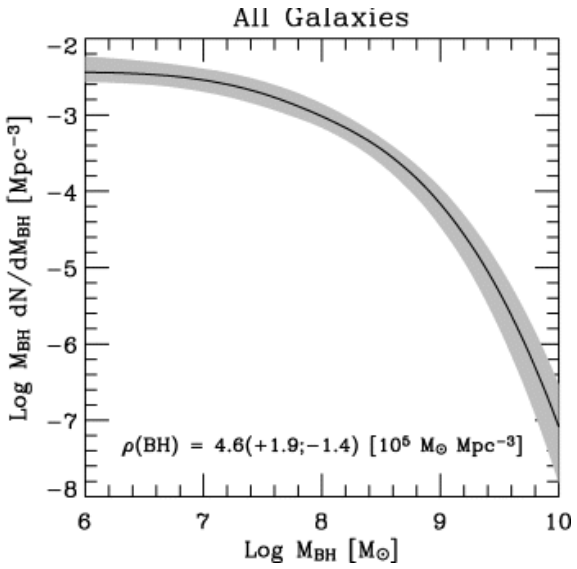

Figure 16: Mass function of remnant black holes, leftovers of past quasar activity. Source: A. Marconi et al. (2004).

quasars at very high redshift is in tension with this assumption: the age of the Universe at $z > 5$ is less than 1 Gyr, while the first halos that can host astrophysical objects are born a few 100 Myr after the Big Bang; in this limited time the seed has to grow by several orders of magnitude, xsand this is impossible for an Eddington-limited accretion. This is shown in Fig. 17, where they report a few observed bright quasars, with estimated black hole masses $\sim 10^9\ M_\odot$ at $z > 7$, together with the tracks obtained by assuming steady Eddington accretion. The horizontal axis gives the redshift, but the cosmic time scale is reported at the top. These observations rule out (at least for these objects) a scenario with stellar black hole seeds accreting at the Eddington rate.

♣ This tension has given raise to a lively discussion on the nature of seed black holes. One option is to allow for **super-Eddington accretion**. There are several examples of quasars that are accreting at a rate of up to 10× the Eddington rate, but accretion is expected to be intermittent due to the energetic feedback of the AGN on the accreting gas, so a super-Eddington episode may correspond to a lower accretion on average. Another option is to form seed black holes of $10^5 - 10^6\ M_\odot$ through **direct collapse** of a gas cloud (DCBH in the figure). For this to happen, a (zero-metallicity) cloud contained in a small dark matter halo should be prevented to form a star (whose mass is limited by the same Eddington mechanism discussed above) by an extra source of energy, that may be given by ionising radiation from a neighbour galaxy or, on more exotic physical grounds, by heating from dark matter annihilation in the first mini-halos that form at very high redshift $z \sim 30$.

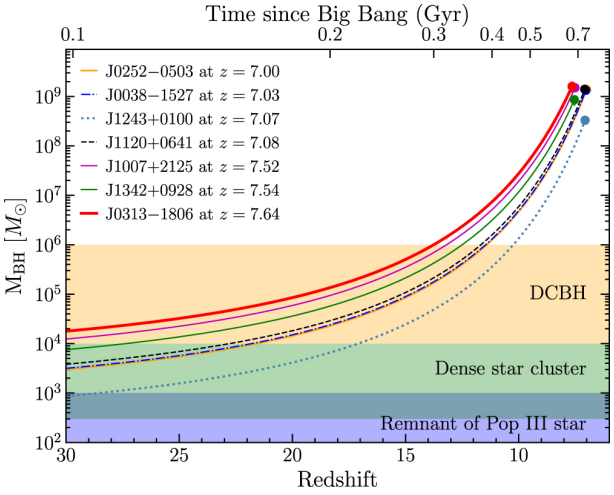

Figure 17: Bright quasar at high redshifts and their implied seed masses for Eddington accretion. Source: Wang et al. (2021), ApJ 907, L1.