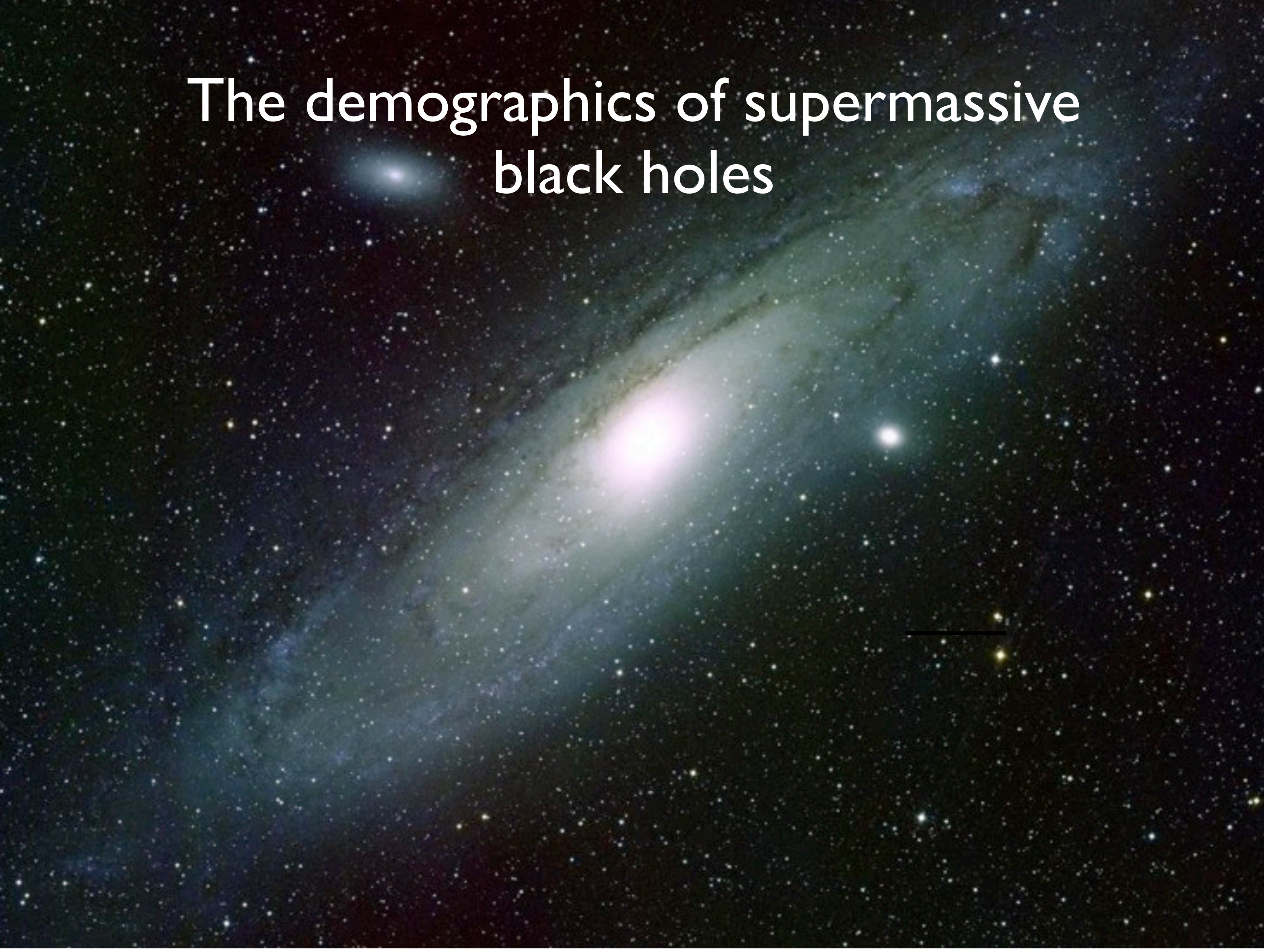
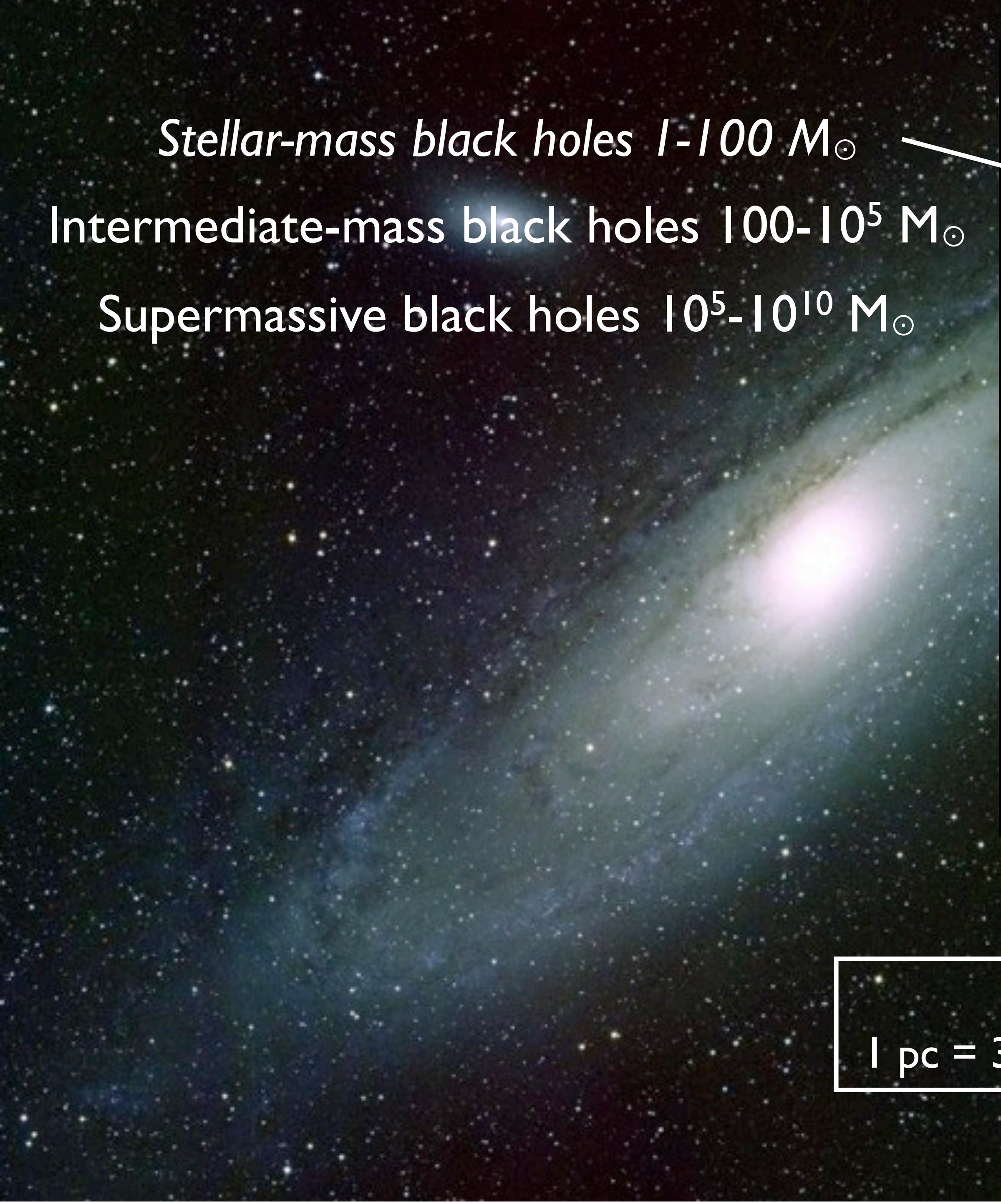
The demographics of supermassive black holes

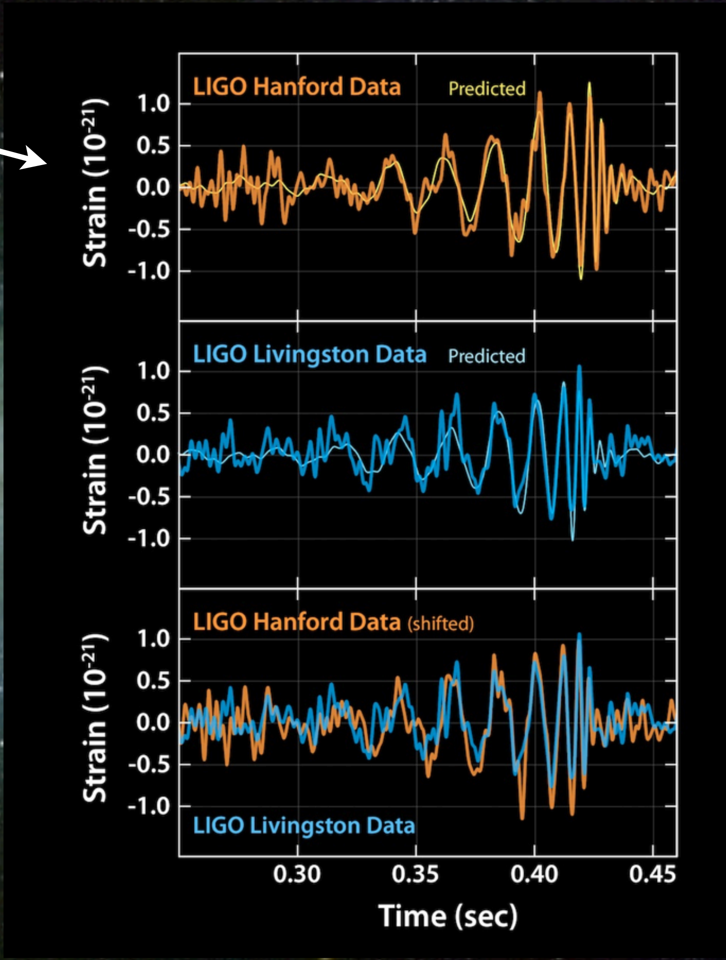




Stellar-mass black holes 1-100 $M_\odot$

Intermediate-mass black holes 100-10^5 $M_\odot$

Supermassive black holes 10^5-10^{10} $M_\odot$



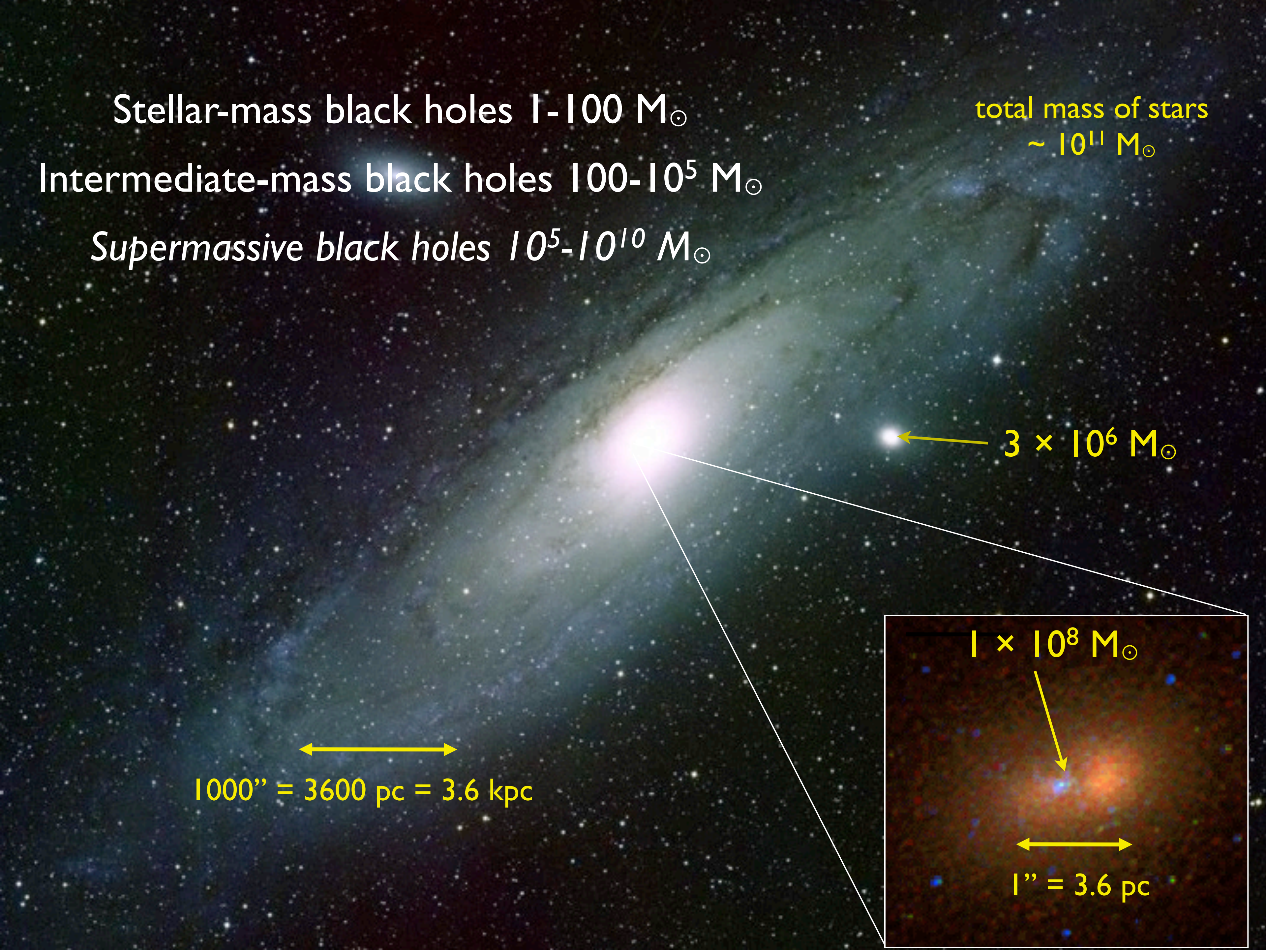
$M_\odot = 2 \times 10^{30}$ kg

1 pc = 3.09×10^{16} m = 3.26 light years

Stellar-mass black holes 1-100 $M_\odot$

Intermediate-mass black holes 100-10^5 $M_\odot$

Supermassive black holes 10^5-10^{10} $M_\odot$



total mass of stars
~ 10^{11} $M_\odot$

3×10^6 $M_\odot$

1×10^8 $M_\odot$

1000" = 3600 pc = 3.6 kpc

1" = 3.6 pc

Supermassive black holes and quasars

- quasars are the most luminous active galactic nuclei (AGN)
- emit up to ~10^{13} $L_\odot$, or 100–1000 times a typical galaxy luminosity
- Eddington luminosity (radiation pressure = gravitational force on a proton)

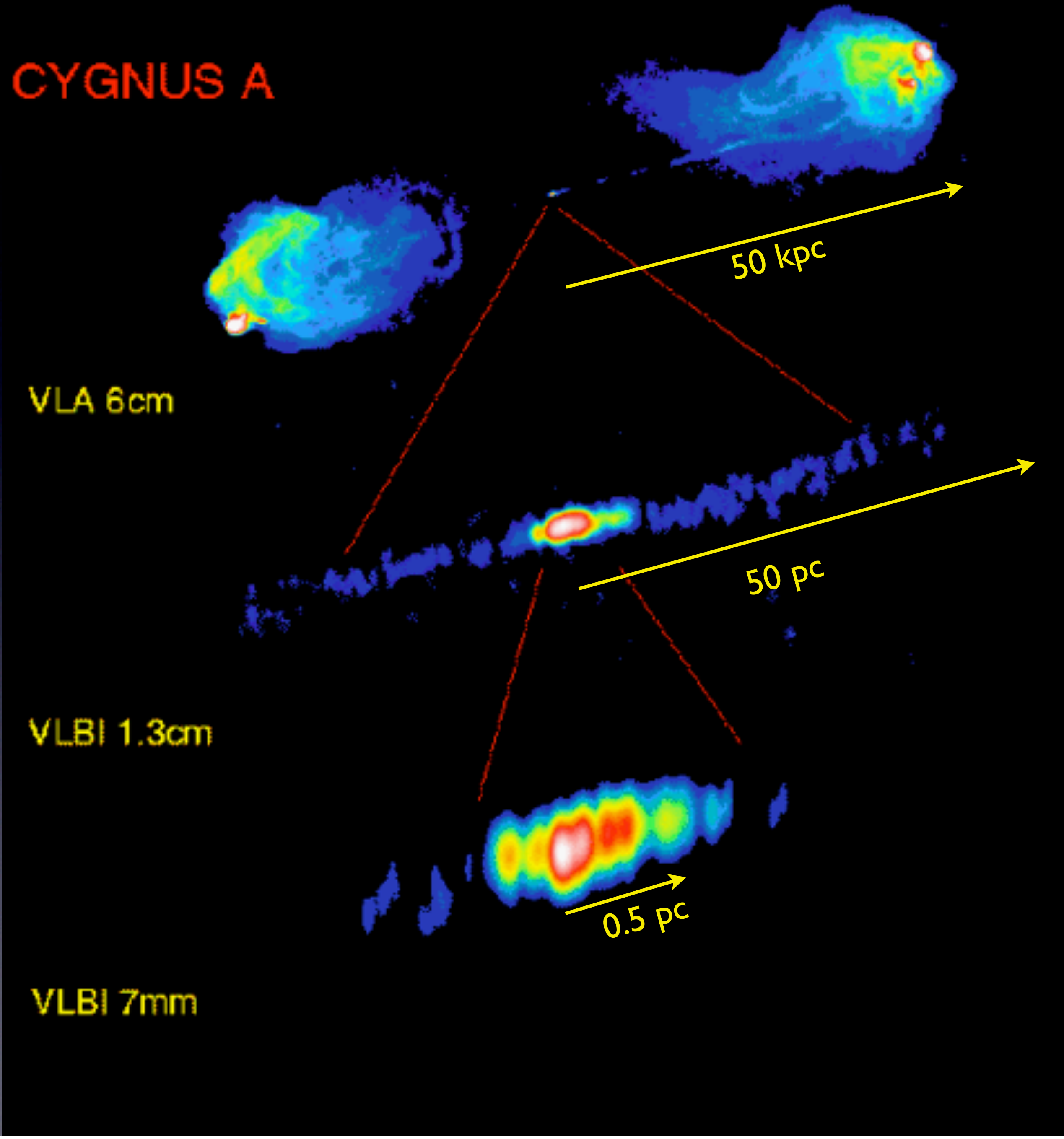
$$L_{\rm Eddington} = \frac{4\pi G M m_p c}{\sigma_{\rm Thomson}} = 1.26\times10^{31}\,{\rm W}\frac{M}{M_\odot} = 3.2\times10^4\,L_\odot\frac{M}{M_\odot}$$

- Eddington luminosity is usually an upper limit to the steady-state luminosity of an object of mass M; therefore quasar engine must be at least M = 3×10^8 $M_\odot$
- energy source for all AGN is believed to be accretion of material onto a black hole
- corresponding Schwarzschild radius

$r_S = 2GM/c^2 = 3\times10^{13}$ cm $(M/10^8\,M_\odot)$ ~ 2 × Earth-Sun distance

why quasars require black holes

1. Directional stability of radio jets over timescales of ~10^5 yr requires a gyroscope that could be provided by a spinning black hole

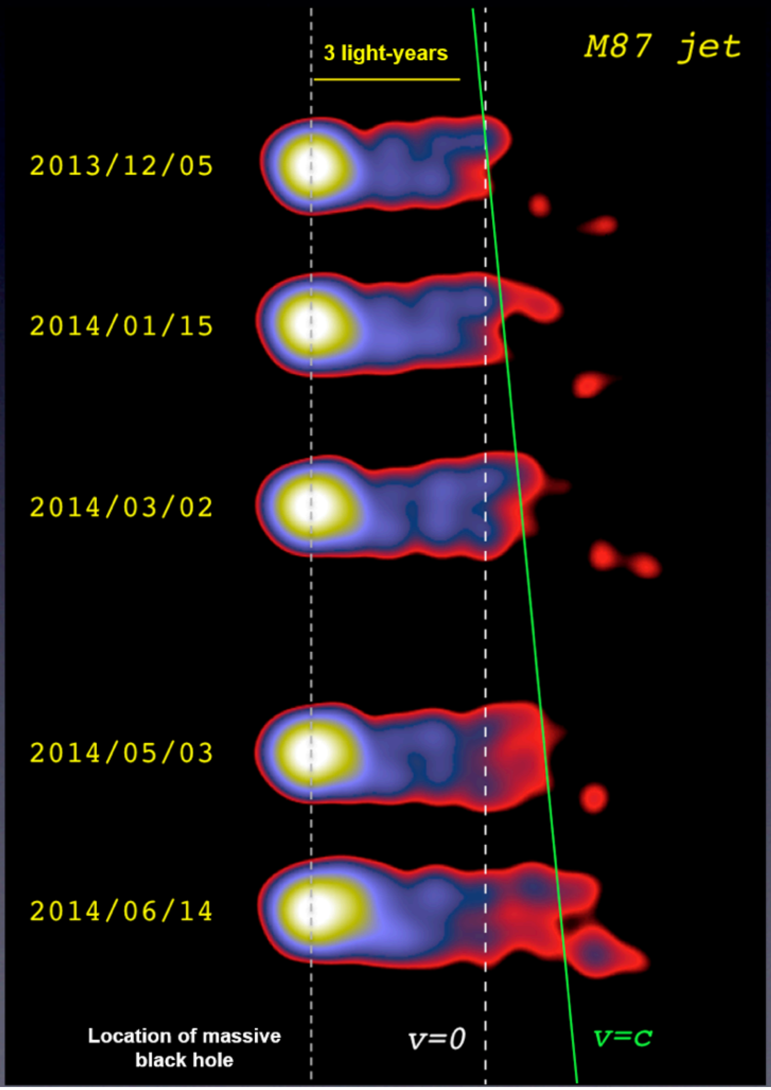


Krichbaum + (1998)

why quasars require black holes

Hubble Space Telescope image

2. Apparent superluminal motion of radio jets



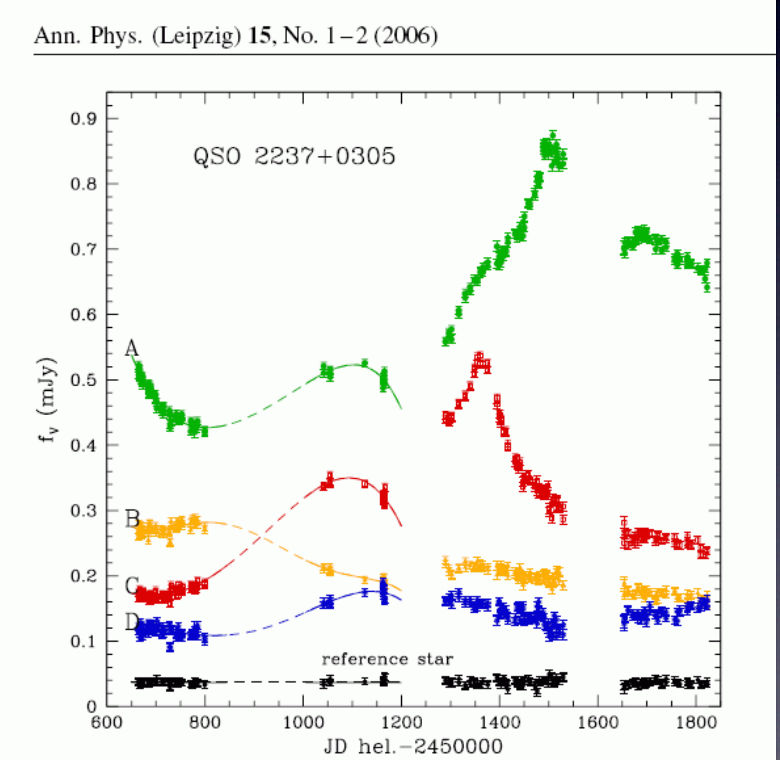
Niinuma + (2016)

why quasars require black holes

3. time variability on timescales of weeks

$$\text{size} < ct \sim 2\times10^{16}\ \text{cm} \times (t/1\ \text{week})$$

4. gravitational lensing by individual stars implies emitting region smaller than Einstein radius of the star, < 10^{14} cm

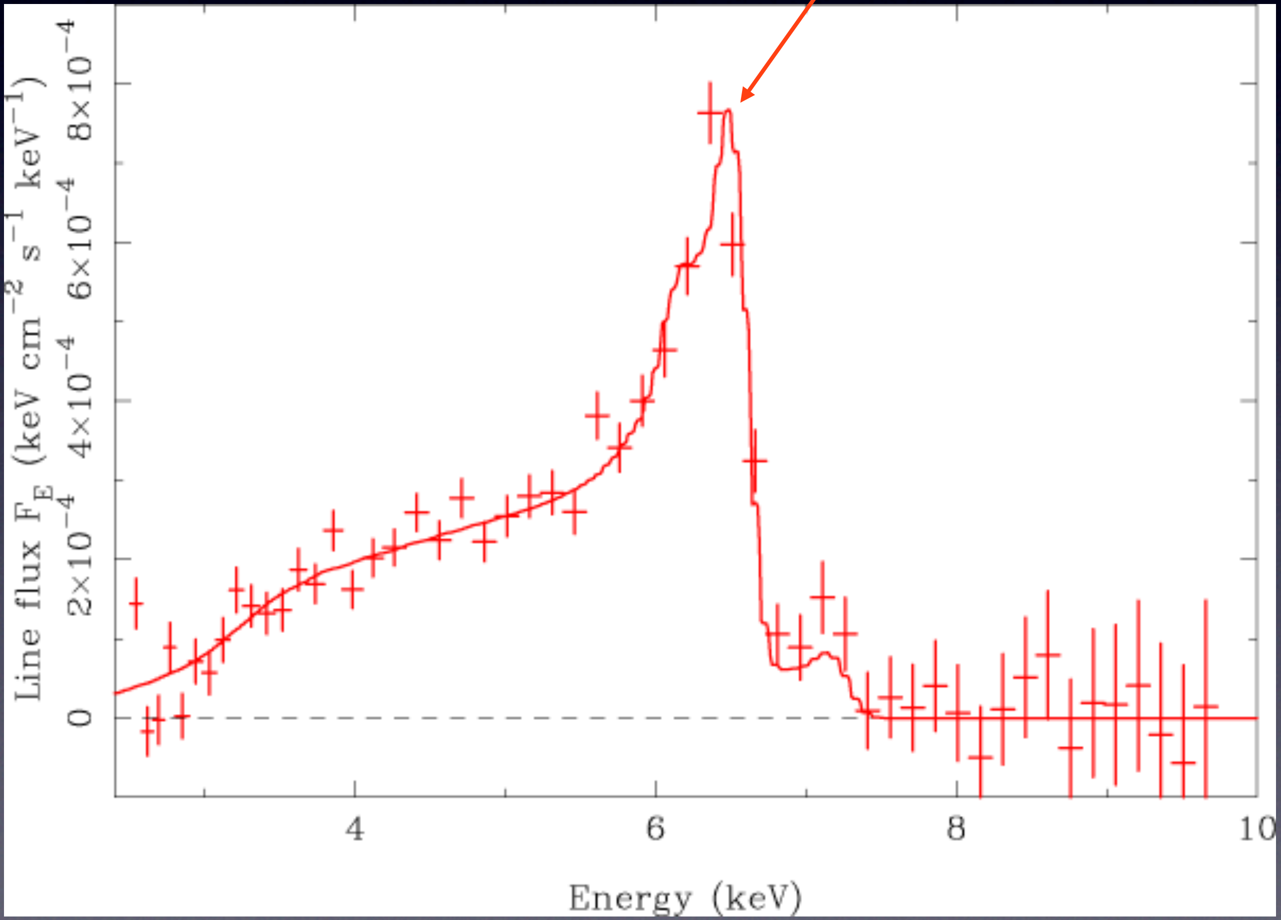


Wambsganss (2006)

why quasars require black holes

5. relativistically broadened and redshifted
X-ray emission lines

iron Kα rest wavelength



why quasars require black holes

Burning a mass ΔM produces energy ΔE with efficiency

$$\epsilon = \frac{\Delta E}{\Delta M c^2}$$

$\epsilon < 0.008$ for nuclear reactions

$\epsilon = 0.057$ for accretion onto a non-rotating black hole

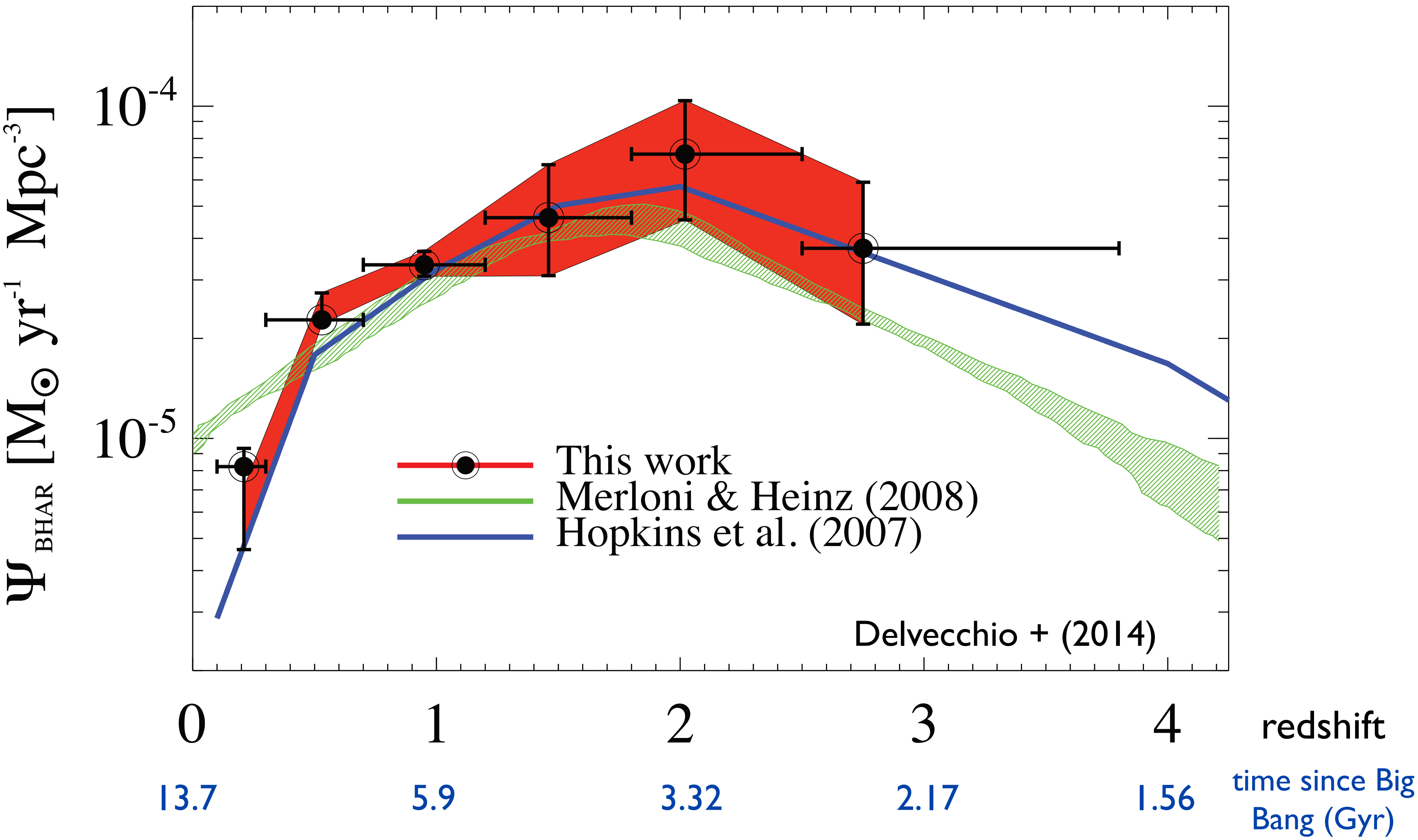
$\epsilon = 0.3$ for accretion onto a black hole in equilibrium spin state

$\epsilon = 0.423$ for accretion onto a maximally rotating black hole

Emission of energy ΔE produces "ash" of mass

$$\frac{1-\epsilon}{\epsilon}\frac{\Delta E}{c^2}$$

6. Sources other than black holes are not efficient enough to produce required energy with a reasonable supply of fuel



Delvecchio + (2014)

- comoving luminosity density of quasars is strongly peaked at redshift ~2
- current luminosity density is ~ 30 × smaller than at the peak

if

- supermassive black holes are the power source for quasars
- the present comoving number density of quasars is much less than the density at earlier epochs
- quasars are found in galaxies

then

- many nearby galaxies must contain supermassive black holes or "dead quasars" (Lynden-Bell 1969)
- in a homogeneous universe the local density of quasar photons is directly related to the expected density of quasar ashes

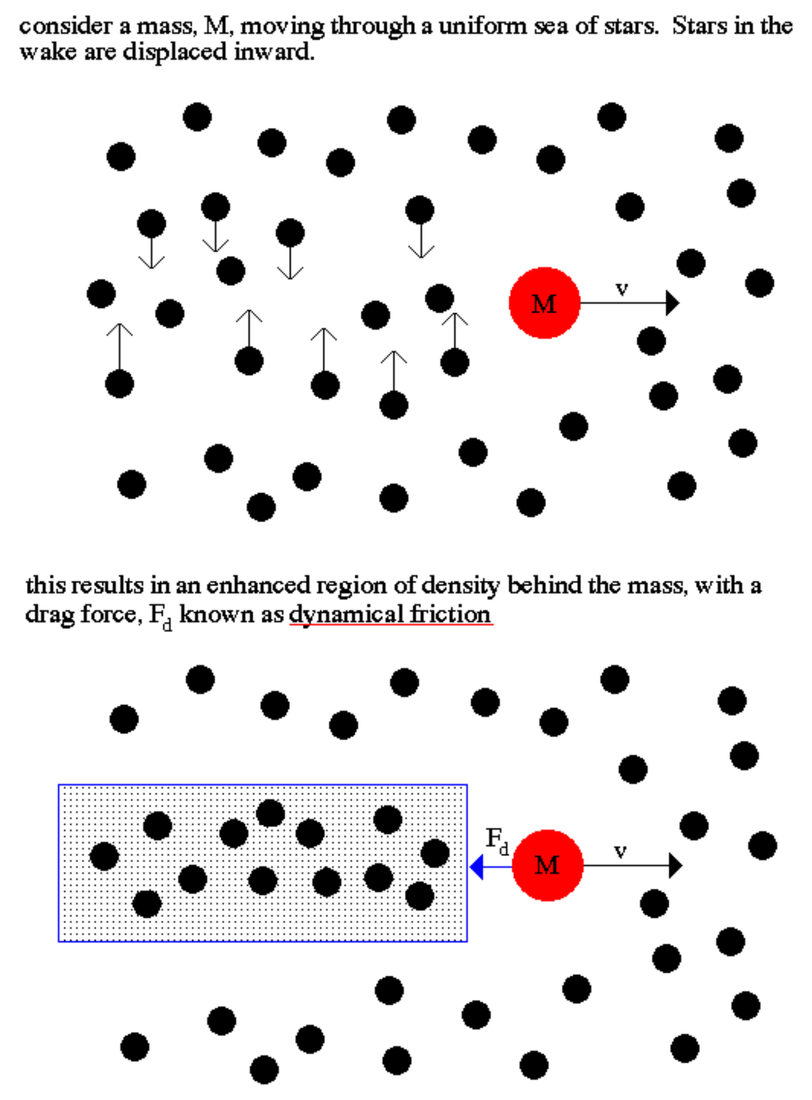
$\sim 3 \times 10^5 \, (\varepsilon/0.1) \, M_\odot/\mathrm{Mpc}^3$ (Sołtan 1982)

Dead quasars should be at the centers of galaxies

why look at the center?

- that’s the bottom of the galaxy’s gravitational potential well
- that’s the only place we can find them
- dynamical friction causes orbits of massive bodies to spiral to the center

(J. Schombert, University of Oregon)



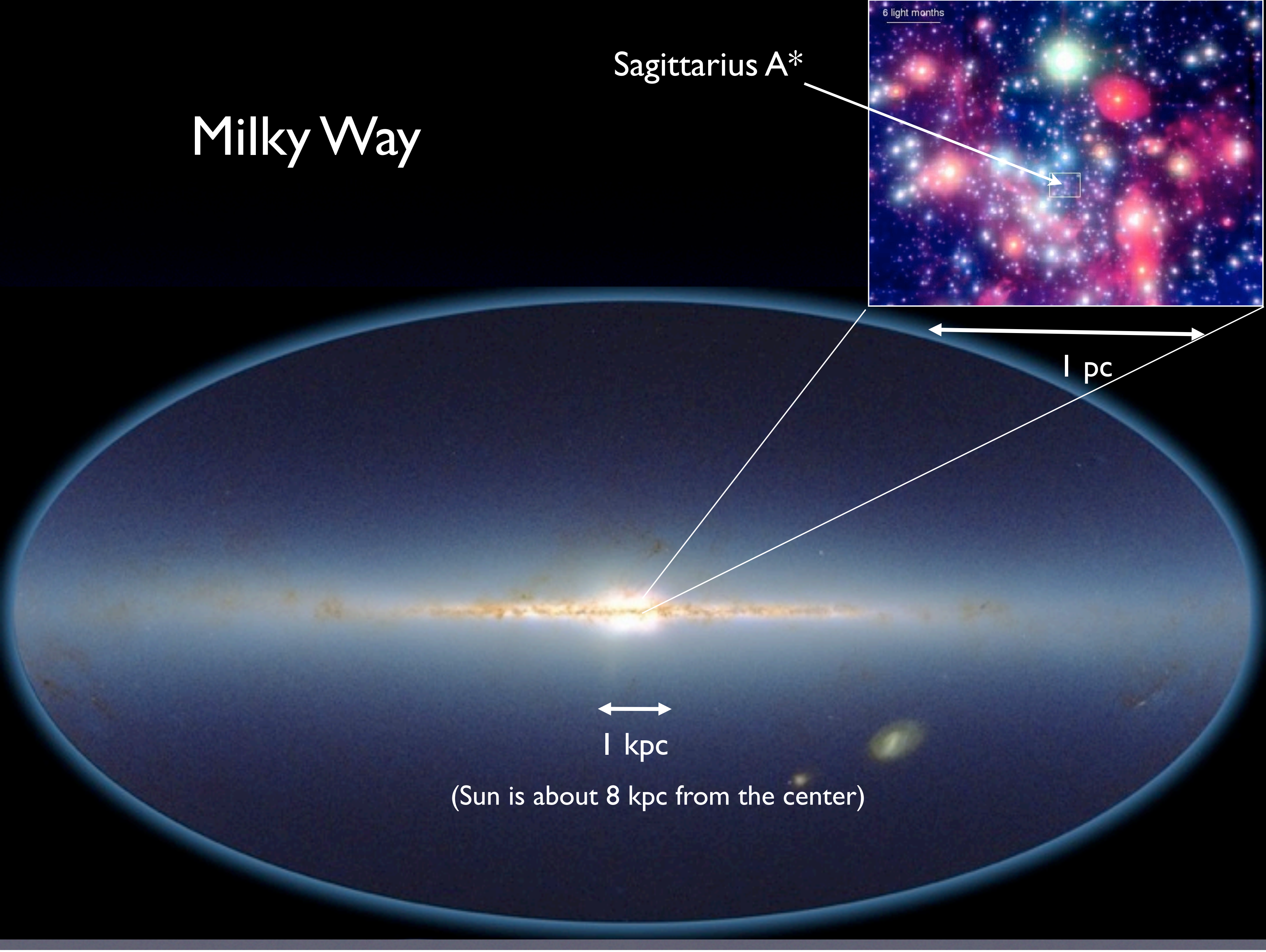
consider a mass, M, moving through a uniform sea of stars. Stars in the wake are displaced inward.

this results in an enhanced region of density behind the mass, with a drag force, F_d known as dynamical friction


10^8 $M_\odot$
10^7 $M_\odot$
10^6 $M_\odot$

Supermassive black holes in the local universe

Milky Way



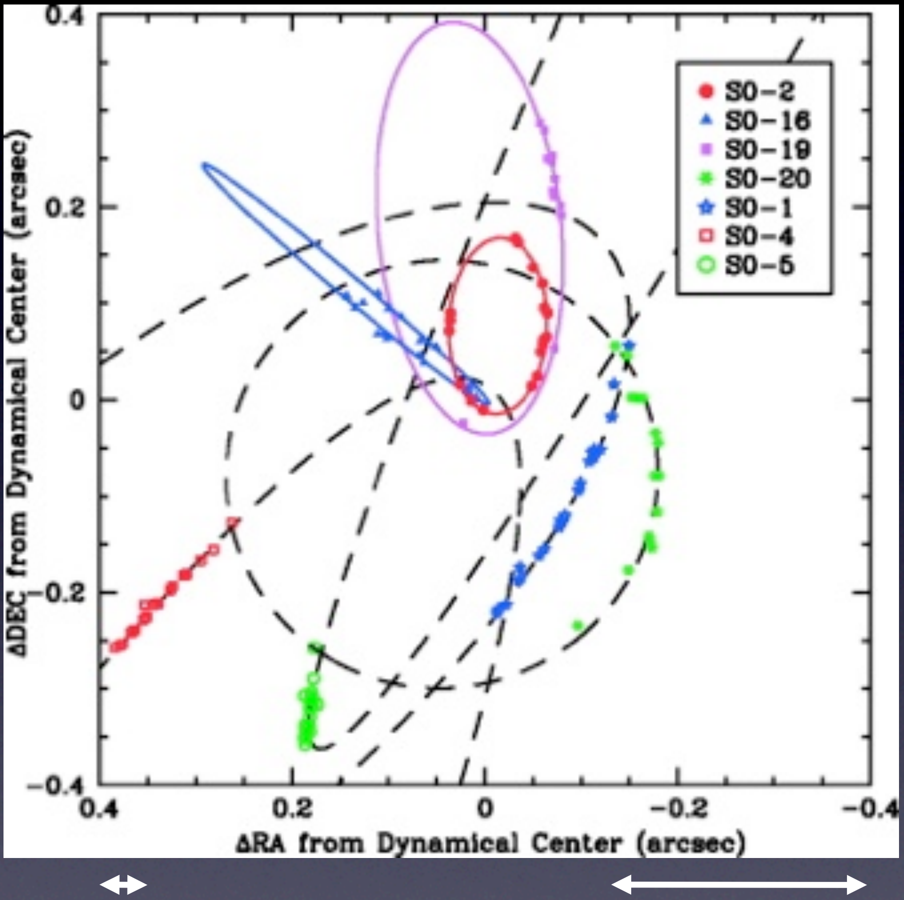
Sagittarius A*

1 pc

1 kpc

(Sun is about 8 kpc from the center)

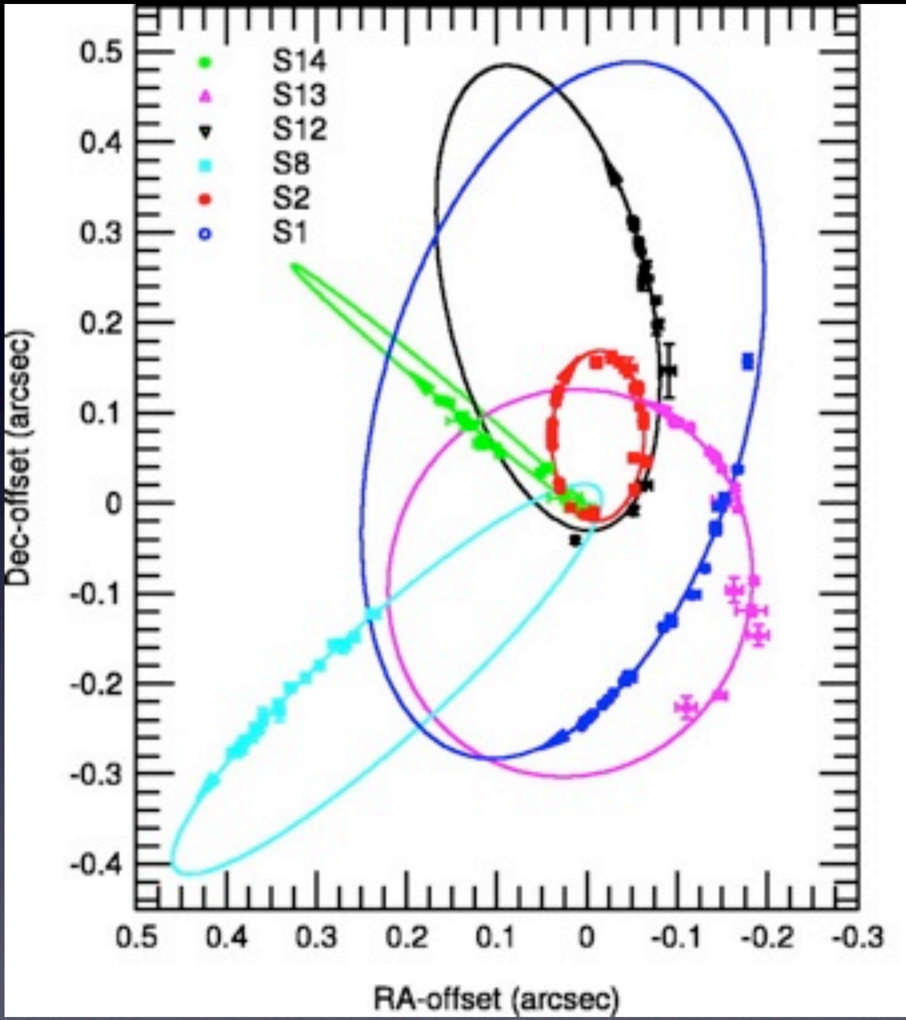
The black hole in the Galactic center



10 X Sun-Pluto distance

0.01 pc

Ghez + (2005)

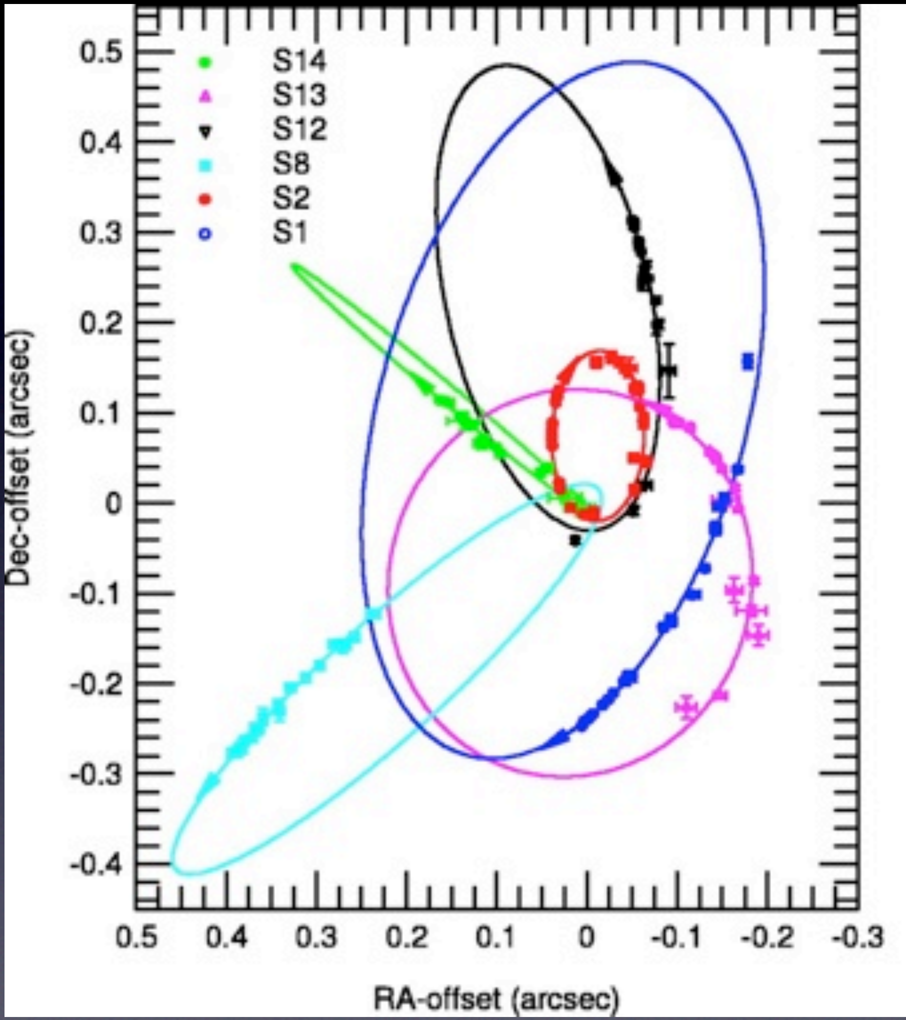


Eisenhauer + (2005)

The black hole in the Galactic center

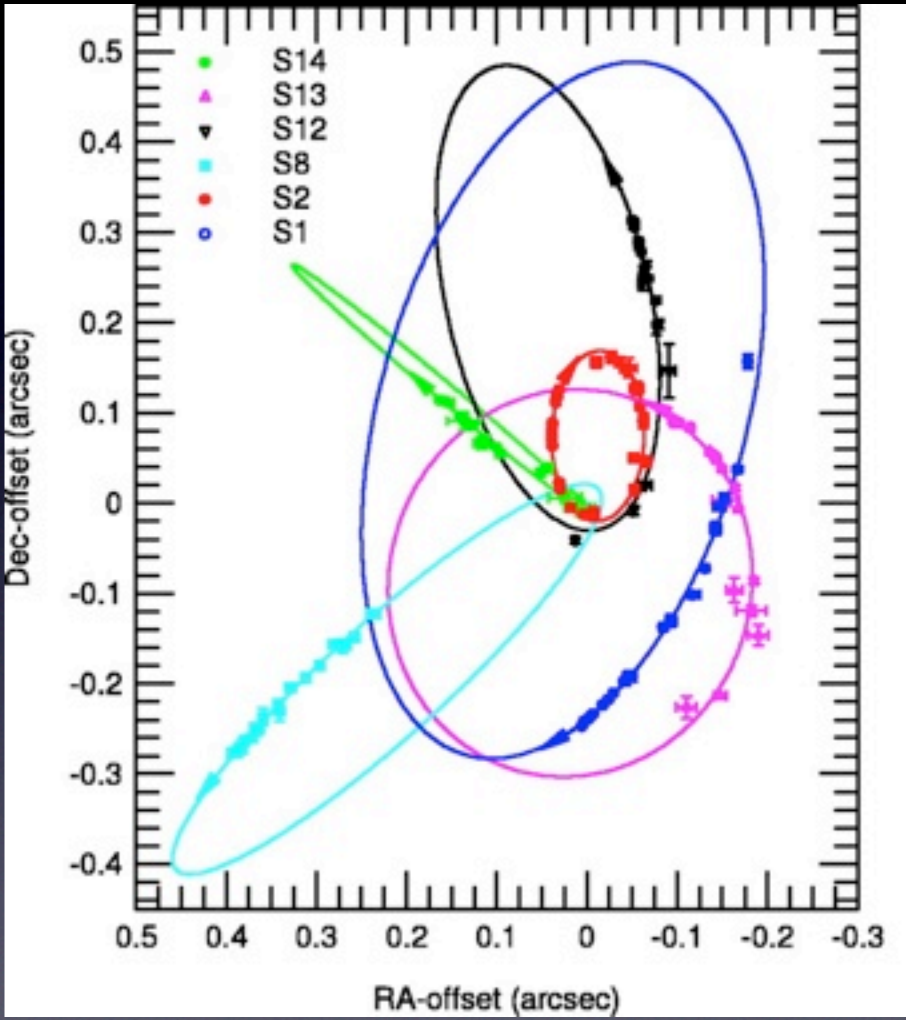
- center of attraction is located at the radio source Sagittarius A* which is presumably the black hole
- smallest pericenter is only ~ 0.0005 pc ~ 3× distance to Neptune; and smallest orbital period is only 16 yr
- orbits are closed ellipses so central mass must be smaller than pericenter
- $M = (3.95 \pm 0.06) \times 10^6\ M_\odot$ if distance R_0 = 8 kpc
- $R_0 = 8.33 \pm 0.35$ kpc

(Gillessen + 2009)



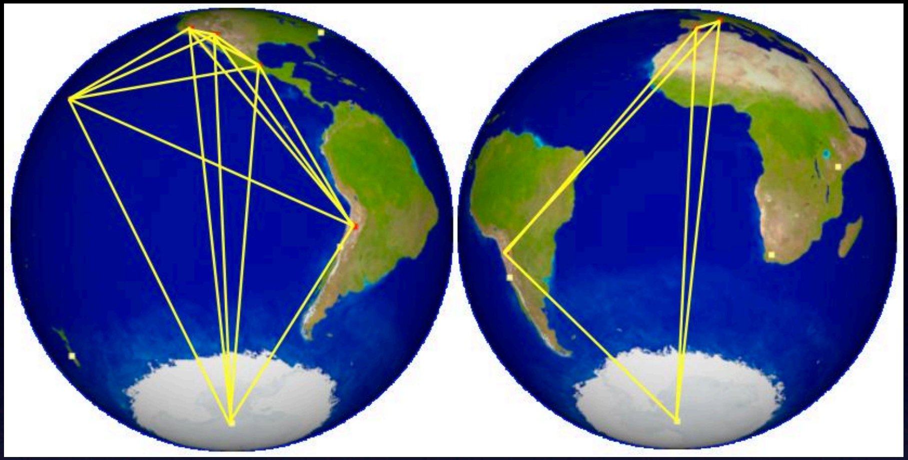
The black hole in the Galactic center

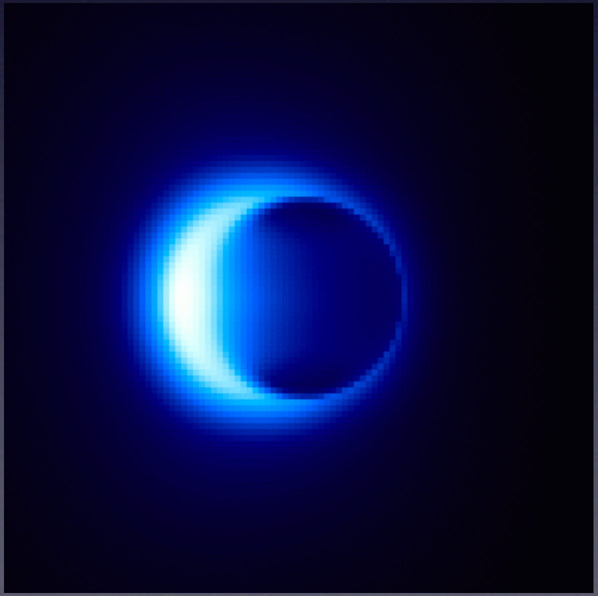
- Galactic center contains $4 \times 10^6 M_\odot$ in < 0.0005 pc ~ 3× distance to Neptune ~ 1000 r_S
- all plausible alternatives to a black hole have very short lifetimes (e.g., cluster of neutron stars)
- some implausible alternatives can survive:
  - cluster of 10^{10} Saturn-mass black holes
  - Bose-Einstein condensate of some unknown elementary particle



The black hole in the Galactic center

- Sagittarius A* emits synchrotron radiation from the accretion flow onto the black hole
- intrinsic size should be dominated by material near Schwarzschild radius r_S = 10 μas
- there should be a central "shadow" of about $2.5r_S$ = 25 μas
- requires mm-wavelength interferometry over global baselines (Event Horizon Telescope)
- preliminary experiments have already revealed highly polarized emission varying on timescales of 1 hour or ct ~ 100 r_S (Johnson + 2015)
- first images ~ 2017



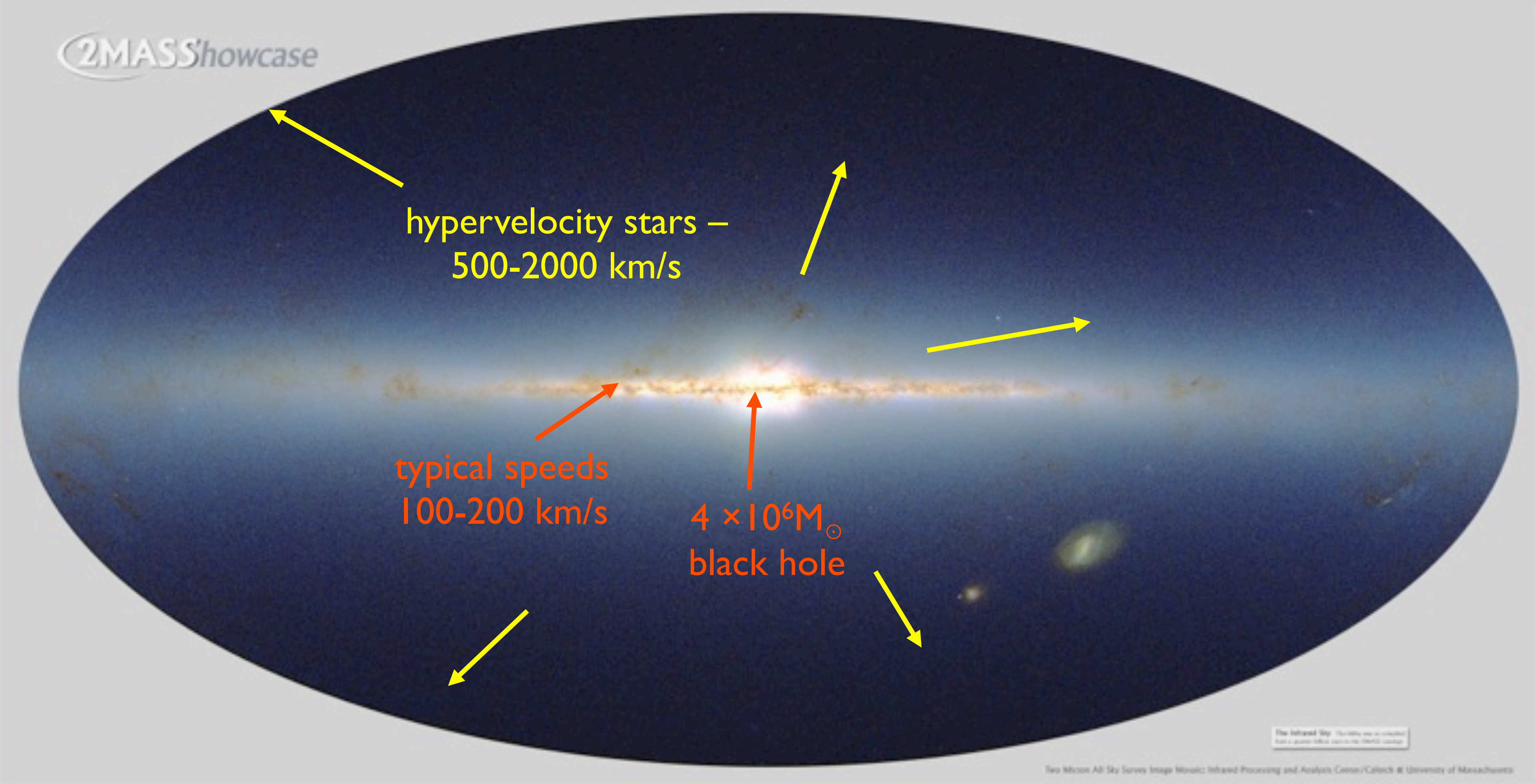


Hypervelocity stars

Hills (1988):

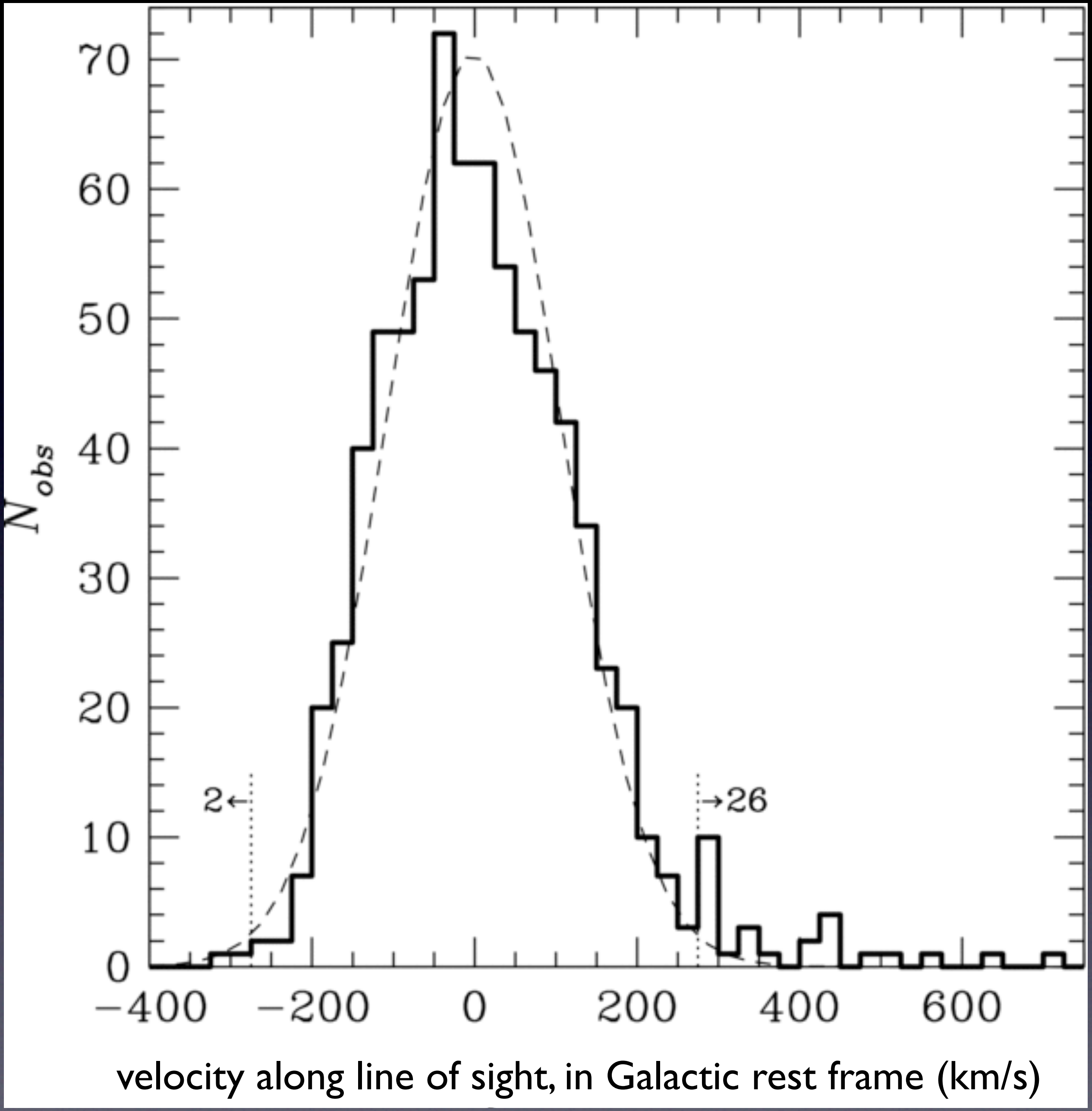
“A close…encounter between a tightly bound binary star and a $10^6 M_\odot$ black hole causes one binary component to become bound to the black hole and the other to be ejected at up to 4,000 km/s. The discovery of even one such hyper-velocity star coming from the Galactic center would be nearly definitive evidence for a massive black hole”

- ejection velocity scales as $v \sim v_{binary}(M_{bh}/M_*)^{1/6}$ where v_{binary} is the binary orbital speed, M_{bh} is the black-hole mass, and M_* is the star mass



look for stars that are:

- high above Galactic plane (less confusion)
- young (bright, so easy to find; also normal halo stars are all old)
- moving at high speed
- if at large distances and on escape orbit, must be moving away from us



line-of-sight velocities of a sample of young stars at 25-100 kpc

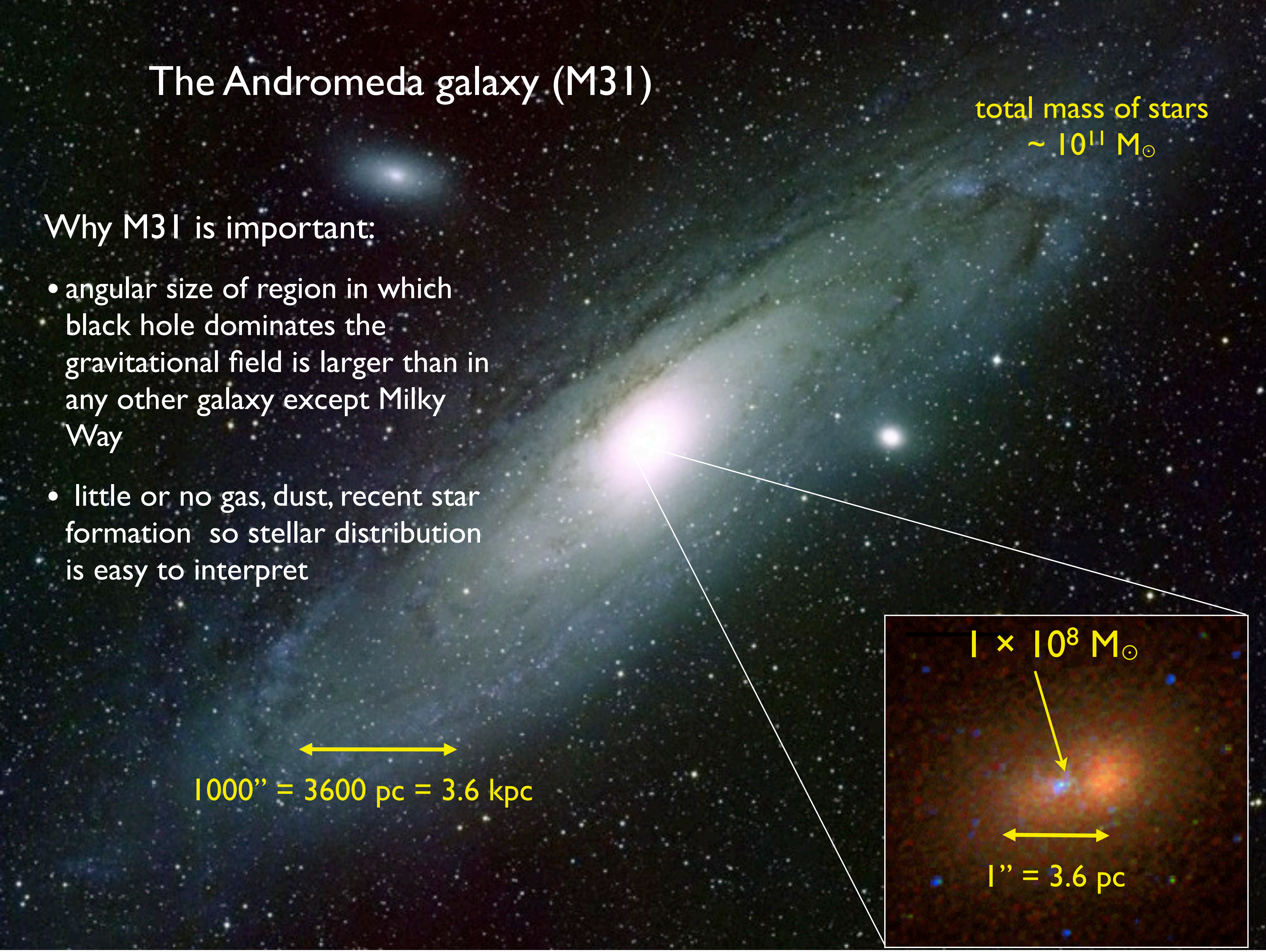
Brown + (2009)

Hypervelocity stars

Were these stars really ejected by the central black hole?

- NOT runaway stars (produced if a supernova goes off in a close binary) --- these have kick velocity < 200 km/s ✓
- travel times from the Galactic center are < stellar lifetime (200 Myr) so formation in Galactic center is possible ✓
- rate (~ 1/Myr) is roughly consistent with theoretical predictions ✓
- when orbit is unbound velocities are outward ✓
- $N(< r) \sim r$, as expected for uniform ejection rate ✓
- should be isotropically distributed ?
- when proper motions are determined by Gaia, the velocity vectors should point away from the Galactic center ?

The Andromeda galaxy (M31)

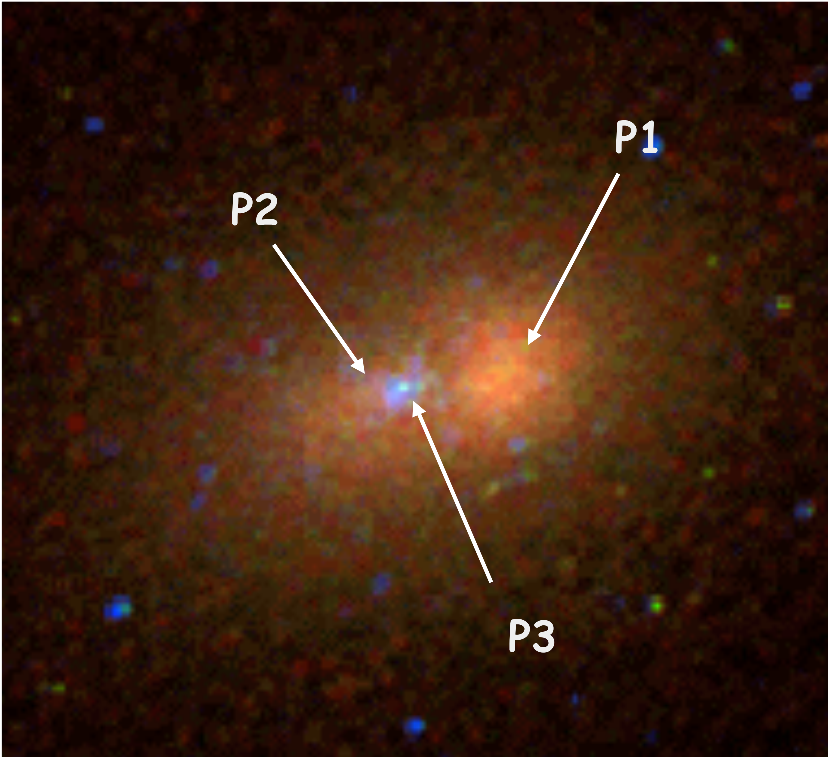


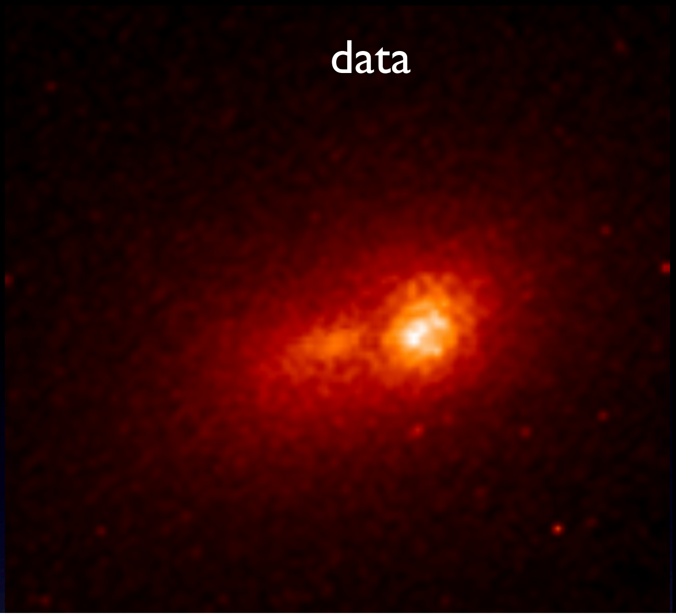
Why M31 is important:

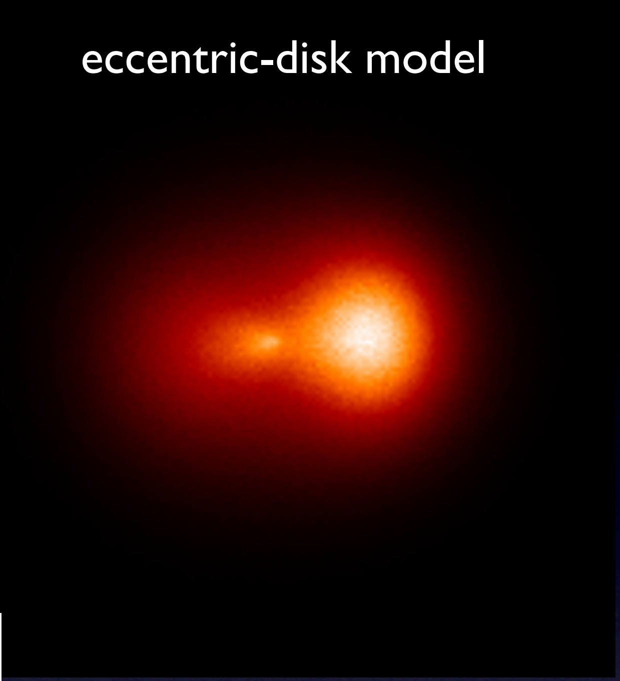
- angular size of region in which black hole dominates the gravitational field is larger than in any other galaxy except Milky Way
- little or no gas, dust, recent star formation so stellar distribution is easy to interpret

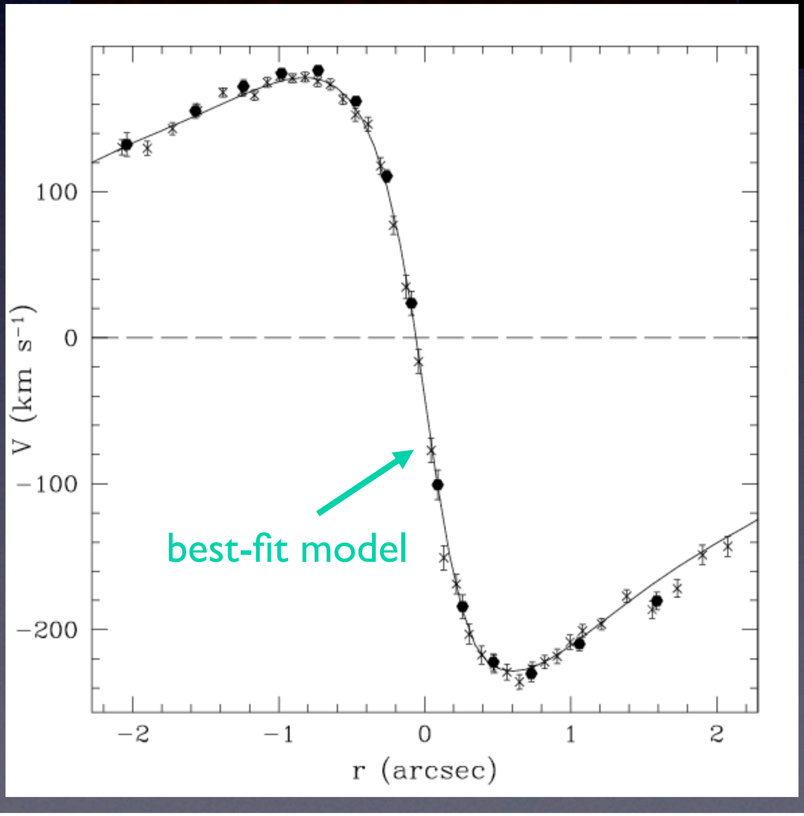
The nucleus of M31

- there is a supermassive black hole at P3, surrounded by a disk of young, blue stars
- the black hole is also orbited by a larger disk of old, reddish stars
- the stars in the disk are on eccentric, nearly Keplerian orbits which are aligned so that apocenters point in the same direction
- P1 is the portion of the disc close to apocenter; stars move slowly near apocenter so most of them are found in this region at any given time
- P2 is the portion of the disk close to pericenter
- black hole dominates gravitational potential so orbits are approximately closed, and the self-gravity of the stars ensures that they precess in lockstep




data


eccentric-disk model

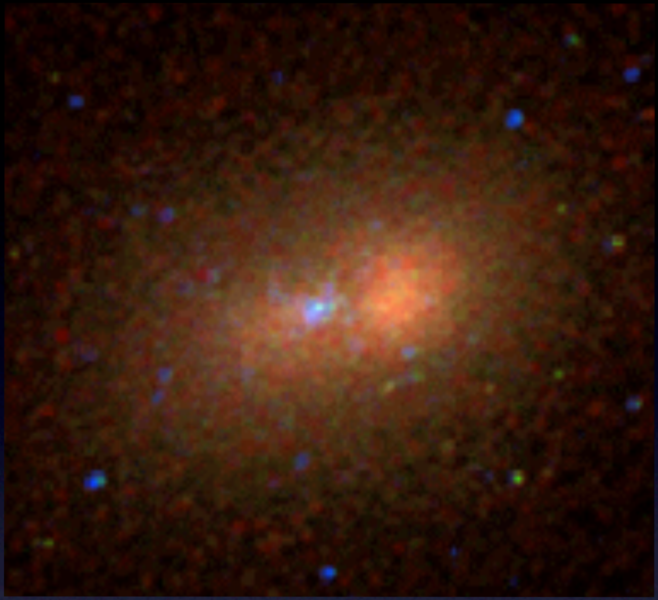


best-fit model

- eccentric disk model that matches photometry correctly and automatically reproduces rotation and dispersion curves
- correctly predicts that P3 should be at the center of the galaxy
- required black-hole mass is $M_{BH} = 1 \times 10^8\, M_\odot$

data from Kormendy & Bender (1999)
models from Peiris & Tremaine (2003)

Peiris & ST
(2003)



The blue nucleus is a cluster of stars with age ~ 200 Myr. Its mass is about 5000 $M_\odot$

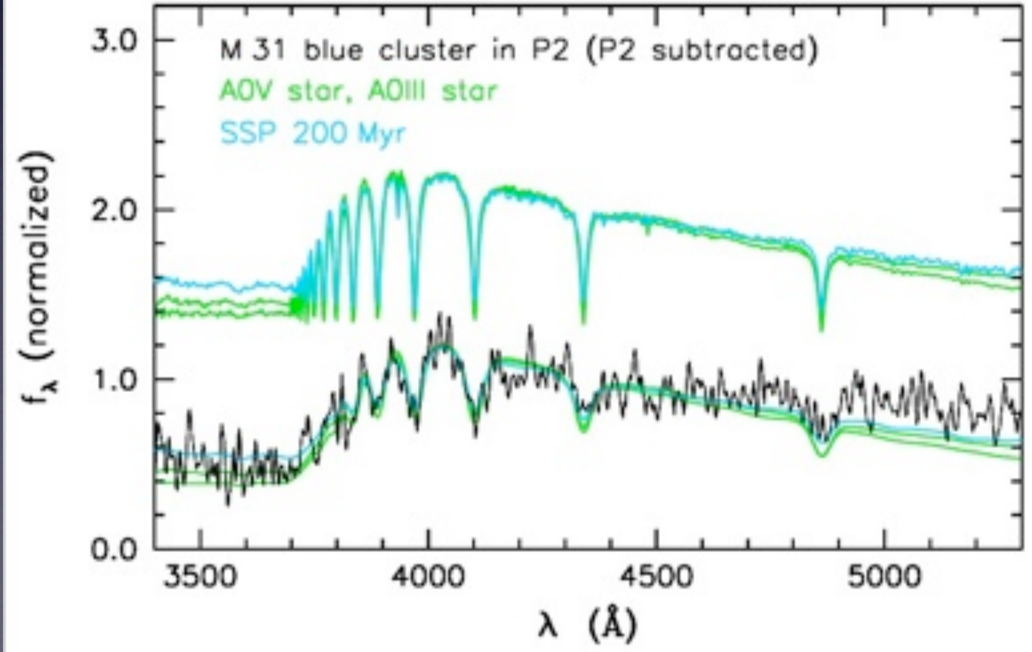
Its velocity dispersion within 0.1" = 0.3 pc is:

σ = 960 ± 106 km/s

compared to average M31 dispersion of 150 km/s

The best kinematic fit is obtained for a point mass, i.e. a black hole of mass

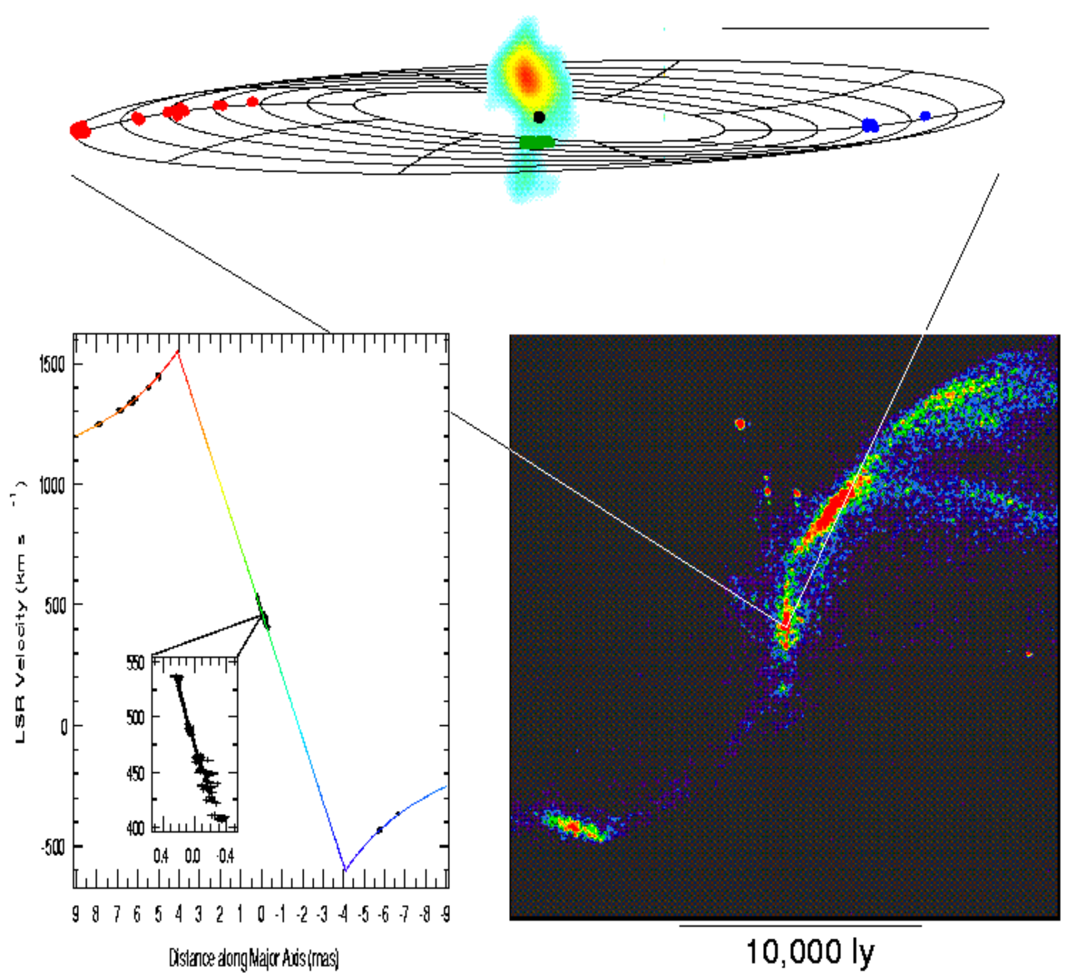
$M \sim 1.4 \times 10^8 M_\odot$



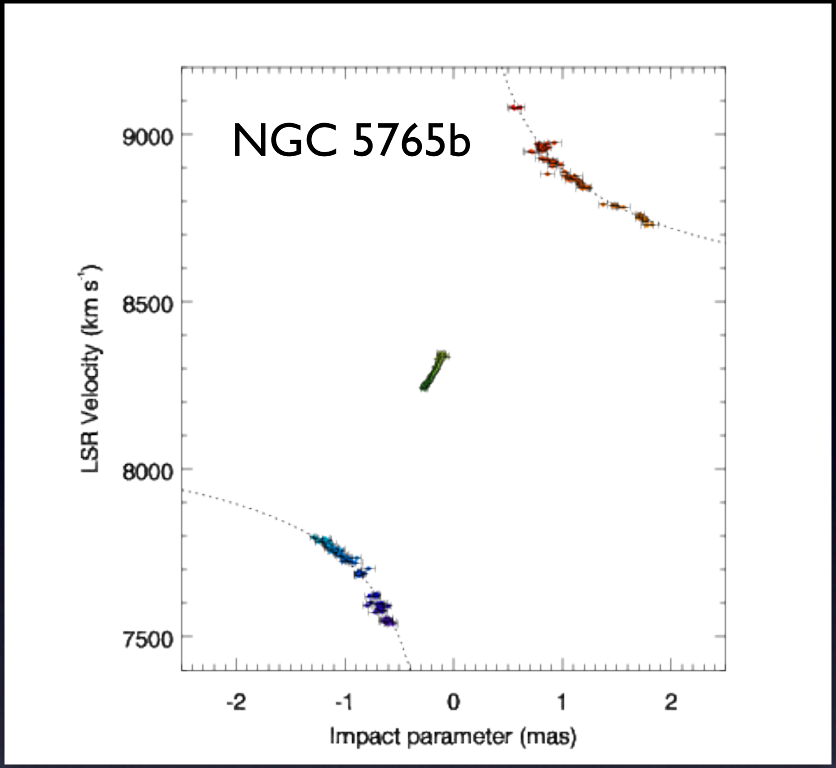
Bender + (2005)

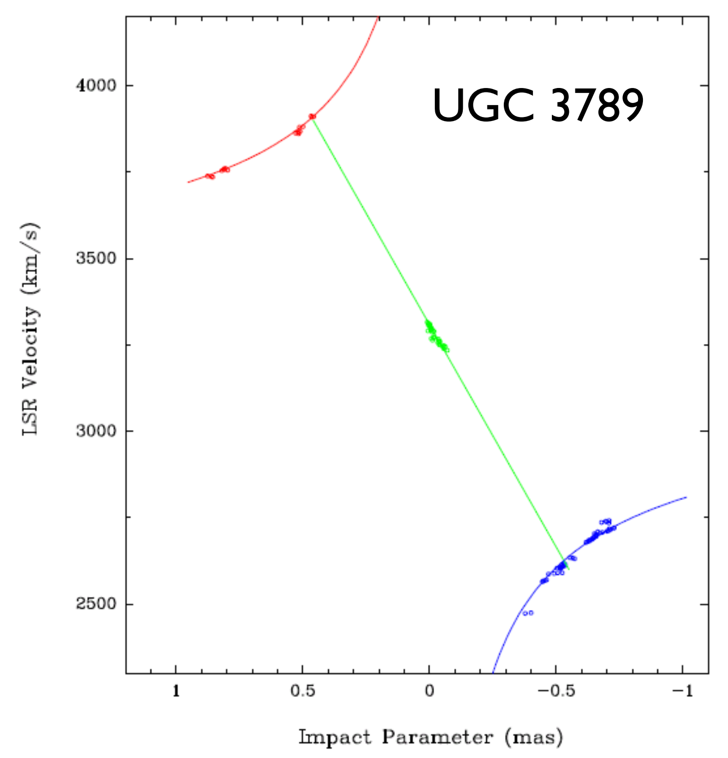
NGC 4258

NGC 4258



- four observational data:
  - amplitude of Keplerian rotation curve
  - proper motion of central masers, 31.5±1 milliarcseconds/yr
  - acceleration of central masers, 9.3±0.3 km/s/yr
  - velocity versus distance for central masers
- three unknown parameters:
  - radius of systemic masers
  - distance of galaxy d
  - black-hole mass M
- $M = (3.9 \pm 0.1) \times 10^7 M_\odot$
- $d = 7.1 \pm 0.2$ Mpc





Gao + (2016)

M = 4.55 ± 0.40 × 10^7 $M_\odot$

Reid + (2012)

M = (1.11 ± 0.08) × 10^7 $M_\odot$

~ 19 accurate black-hole masses from maser disks (Kuo + 2016)

finding black holes in "normal" nearby galaxies

- measure optical spectrum of light from the galaxy at a given position
- if typical star as a spectrum $F^*(\lambda)$ and the number of stars as a function of line-of-sight velocity is $n(v)dv$, the actual spectrum will be

$$F(\lambda) = \int F^*(\lambda - v/c)n(v)dv$$

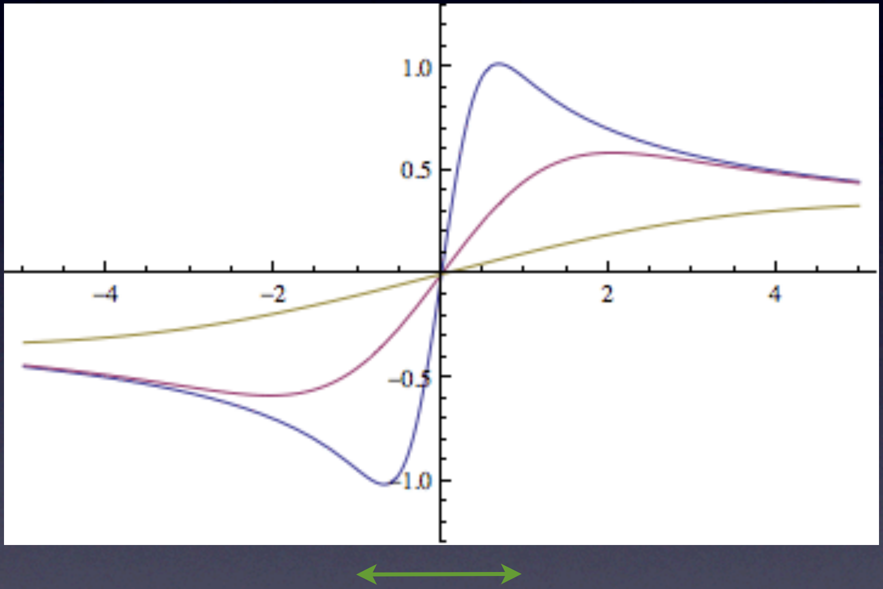
- knowing F and F^* gives $n(v)$ -- parametrize by mean velocity $\langle v \rangle$ and dispersion σ
- central black hole of mass M influences kinematics inside a radius r such that

$$GM/r > \max\,[\sigma^2, \langle v \rangle^2] \quad \text{"sphere of influence"}$$

- crucial problem is to resolve the sphere of influence -- number of galaxies in which a black hole can be detected varies as *cube* of full-width half-maximum (FWHM) of telescope point-spread function
  - typical ground-based telescope at excellent site FWHM = 0.5-1"
  - Hubble Space Telescope FWHM = 0.08"
  - 8-meter ground-based telescope with adaptive optics FWHM = 0.1"

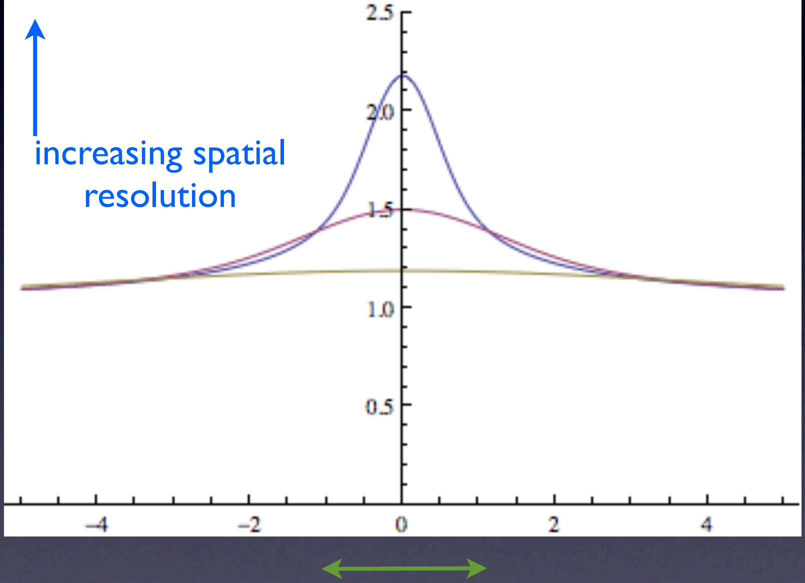
finding black holes in “normal” nearby galaxies

rotation profile <v>



sphere of influence

velocity dispersion profile σ

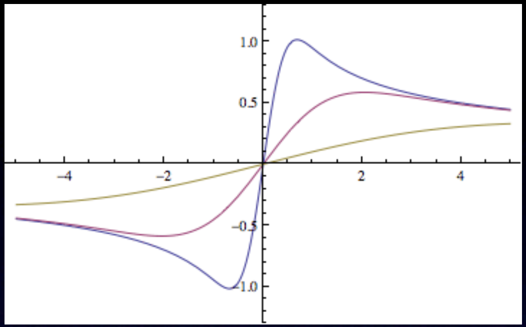


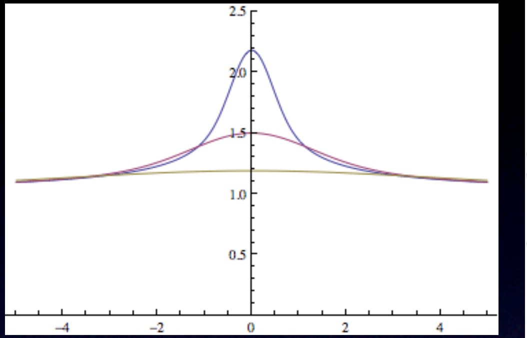
increasing spatial resolution

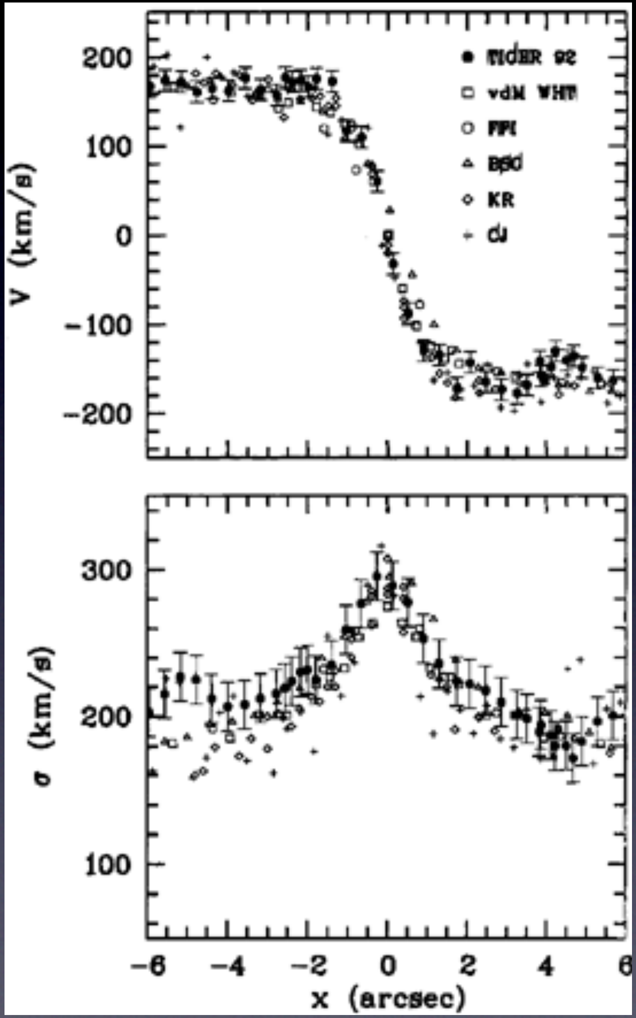
sphere of influence

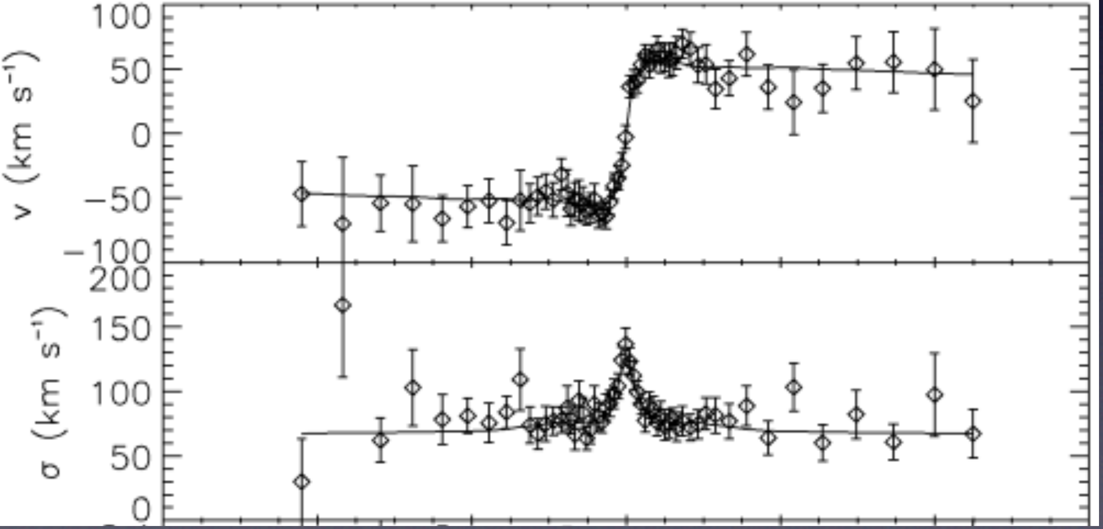
radius of “sphere of influence” = $GM/\max[\sigma^2, \langle v\rangle^2]$

finding black holes in “normal” nearby galaxies









M32 (Verolme + 2002)

NGC 3115 (Emsellem + 1999)

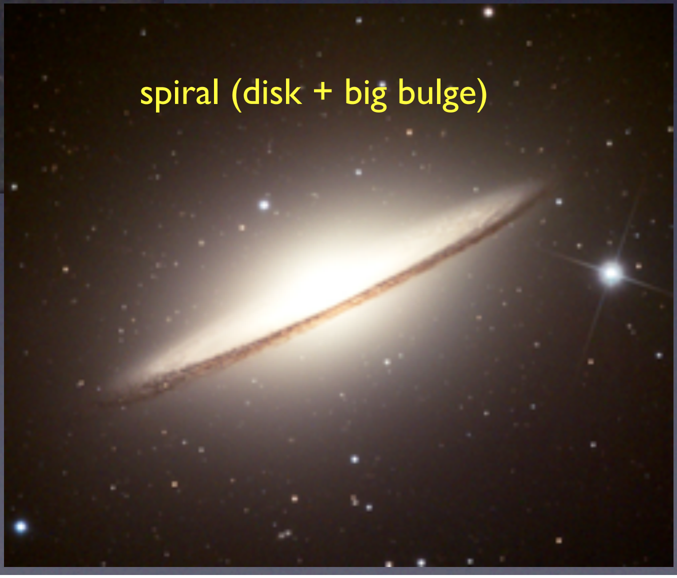
The relation between supermassive black holes and galaxies

elliptical

spiral (disk)

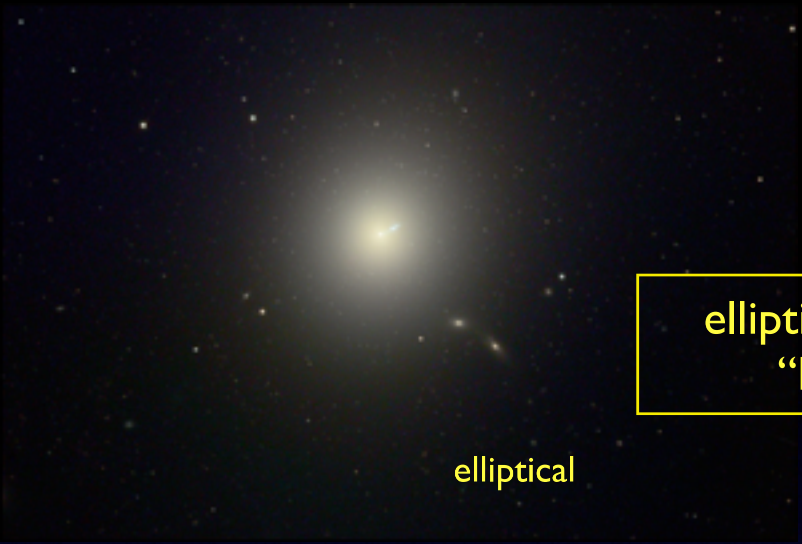
elliptical

spiral (disk + small bulge)



spiral (disk + big bulge)

spiral (disk + medium bulge)

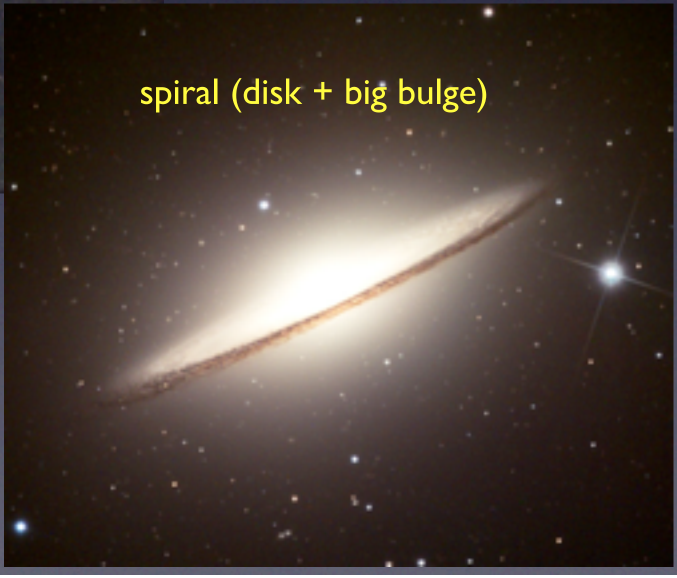

elliptical

ellipticals + spiral bulges =
“hot” components

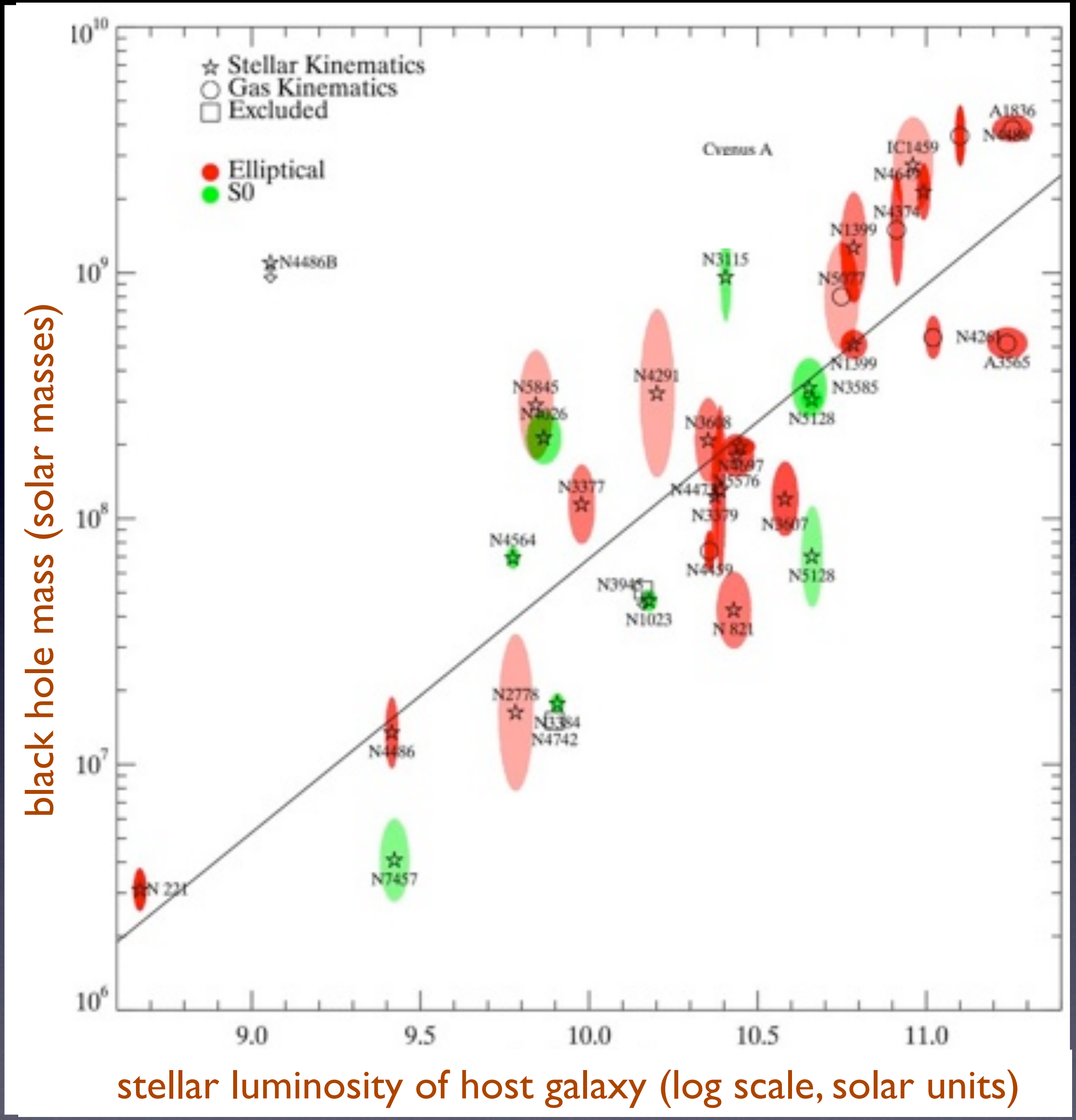

spiral (disk)


elliptical


spiral (disk + small bulge)


spiral (disk + big bulge)


spiral (disk + medium bulge)



- mass correlates with luminosity of hot component; roughly

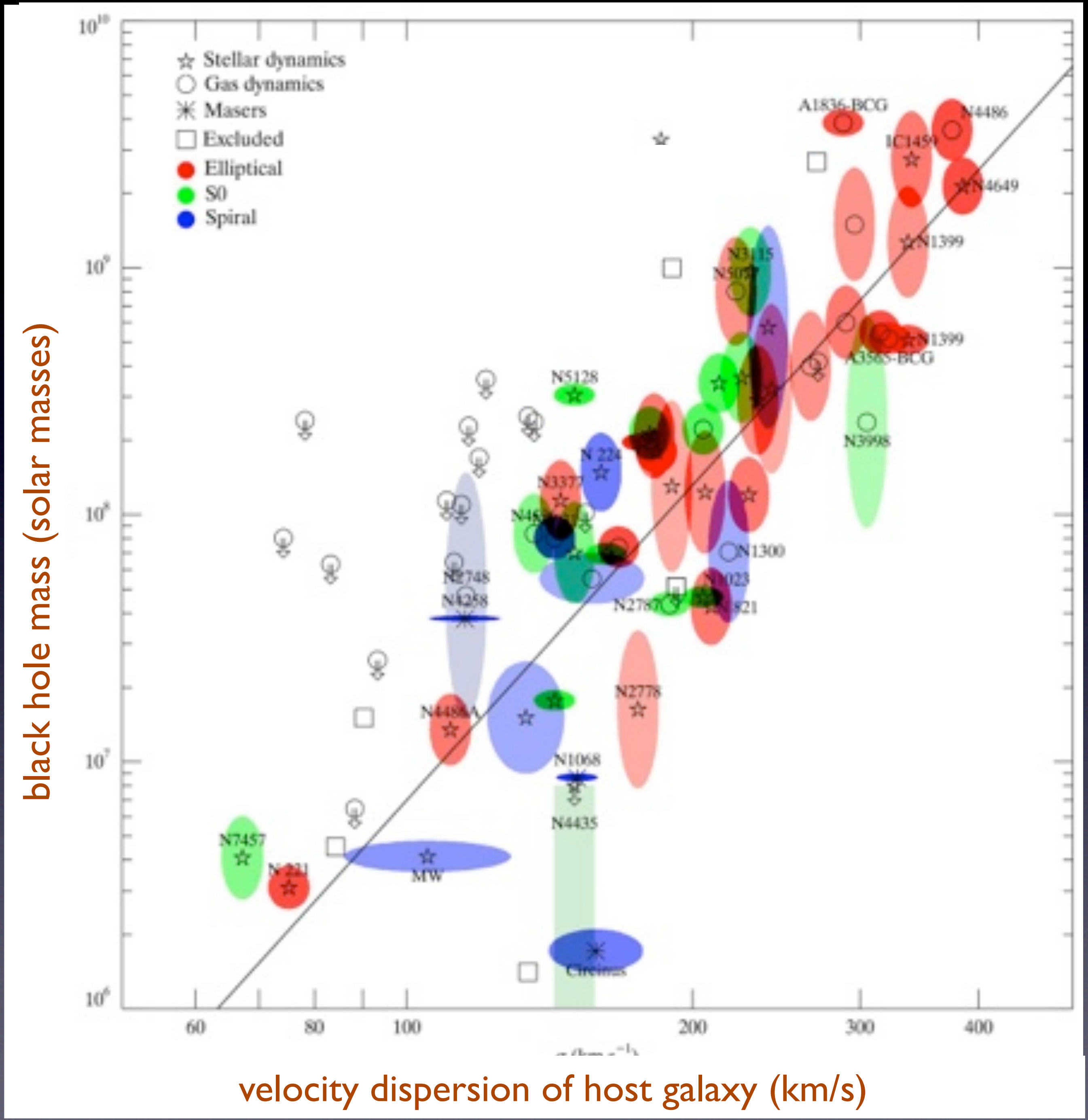
$$M \propto L$$

- in terms of stellar mass

$$M/M_{stars} \sim 0.2\text{-}0.5\ \%$$

- almost all hot components contain black holes

Gültekin + (2009)

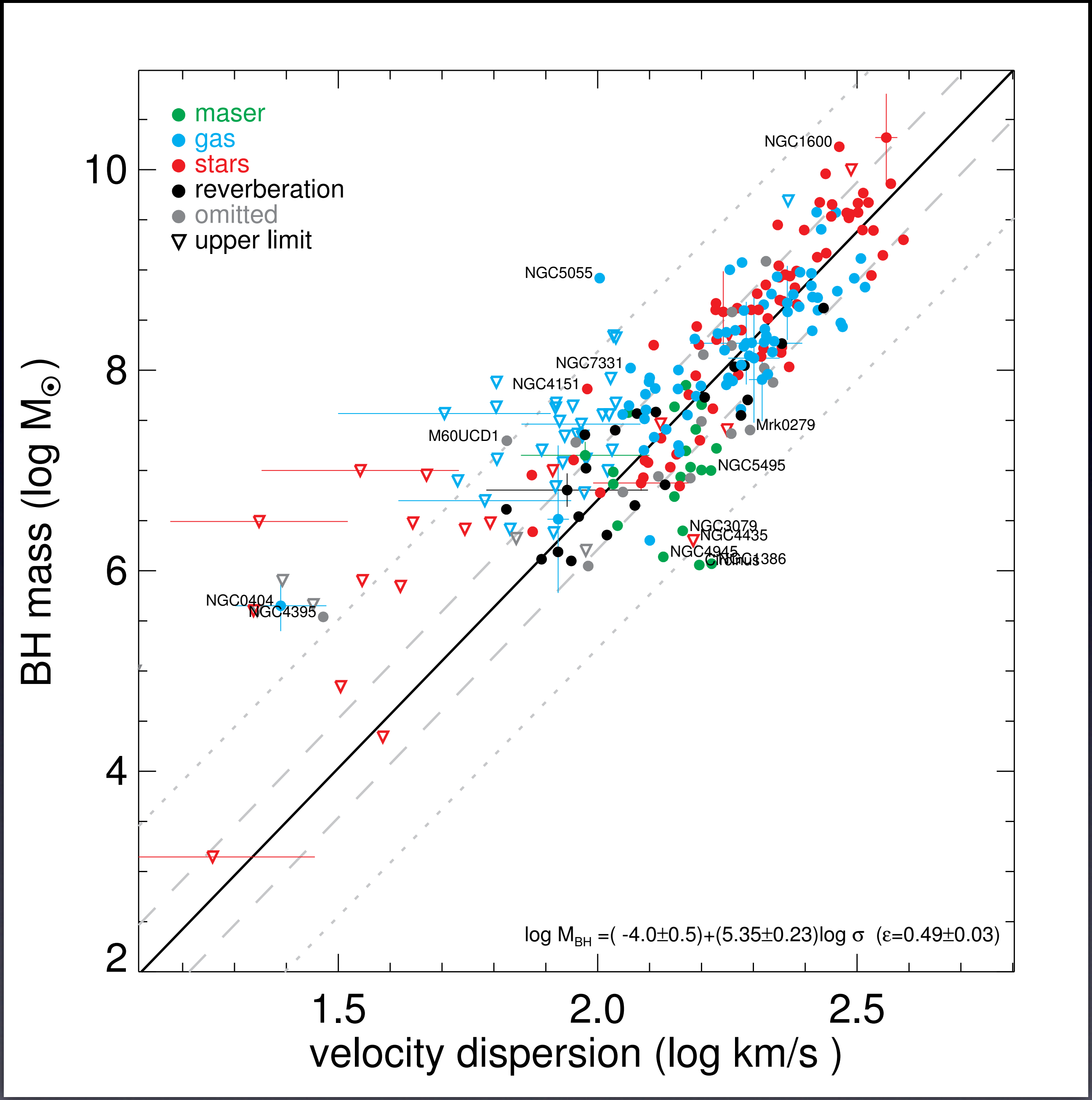


- tighter correlation is with velocity dispersion σ of hot component of host galaxy; roughly

$$M \propto \sigma^4$$

with scatter of 0.3 in $\log_{10}M$ for elliptical galaxies

Gültekin + (2009)



- including a wide range of galaxy types increases the scatter to 0.5 in $\log_{10}M$ but the correlation with dispersion is still present

van den Bosch (2016)

- masses of central black holes in nearby galaxies are about 0.2 - 0.5% of the stellar mass in hot component
- knowing the average mass density in stellar component of galaxies we can estimate the average density in black holes,

$$\rho_{BH} = 4 \pm 1 \times 10^5\ M_\odot\ \mathrm{Mpc}^{-3}$$

if

- supermassive black holes are the power source for quasars
- the present comoving number density of quasars is much less than the density at earlier epochs
- quasars are found in galaxies

then

- many nearby galaxies must contain supermassive black holes or “dead quasars” (Lynden-Bell 1969)
- in a homogeneous universe the local density of quasar photons is directly related to the expected density of quasar ashes

$$\rho_{AGN} \sim 3\times10^{5}\,(\varepsilon/0.1)\,M_\odot\,\mathrm{Mpc}^{-3}$$ (Sołtan 1982)

- masses of central black holes in nearby galaxies are about 0.2 - 0.5% of the stellar mass in hot component
- knowing the average mass density in stellar component of galaxies we can estimate the average density in black holes,

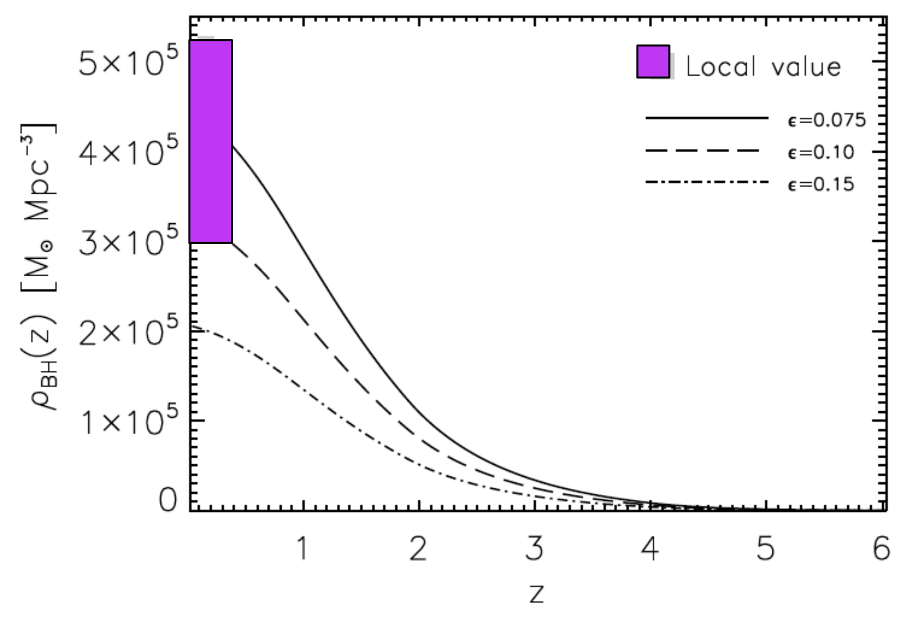
$$\rho_{BH} = 4 \pm 1 \times 10^5 \, M_\odot \, \mathrm{Mpc}^{-3}$$

- in a homogeneous universe the local density of quasar photons is directly related to the expected density of quasar ashes

$$\rho_{AGN} \sim 3 \times 10^5 \, (\varepsilon/0.1) \, M_\odot \, \mathrm{Mpc}^{-3}$$ (Sołtan 1982)

where ε = 0.06 - 0.3 is the expected efficiency for accretion from a thin disk onto a black hole

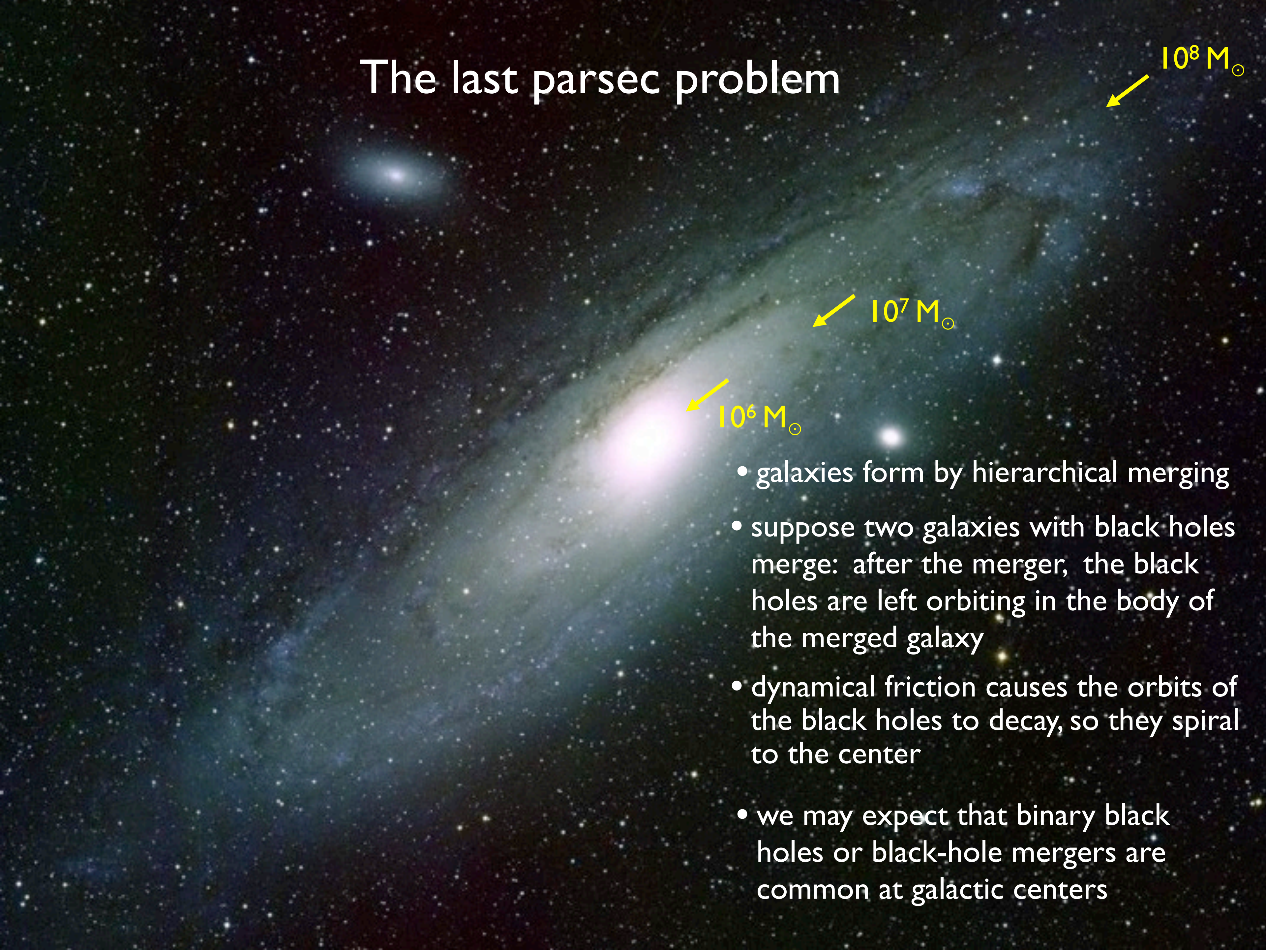
Shankar (2009)



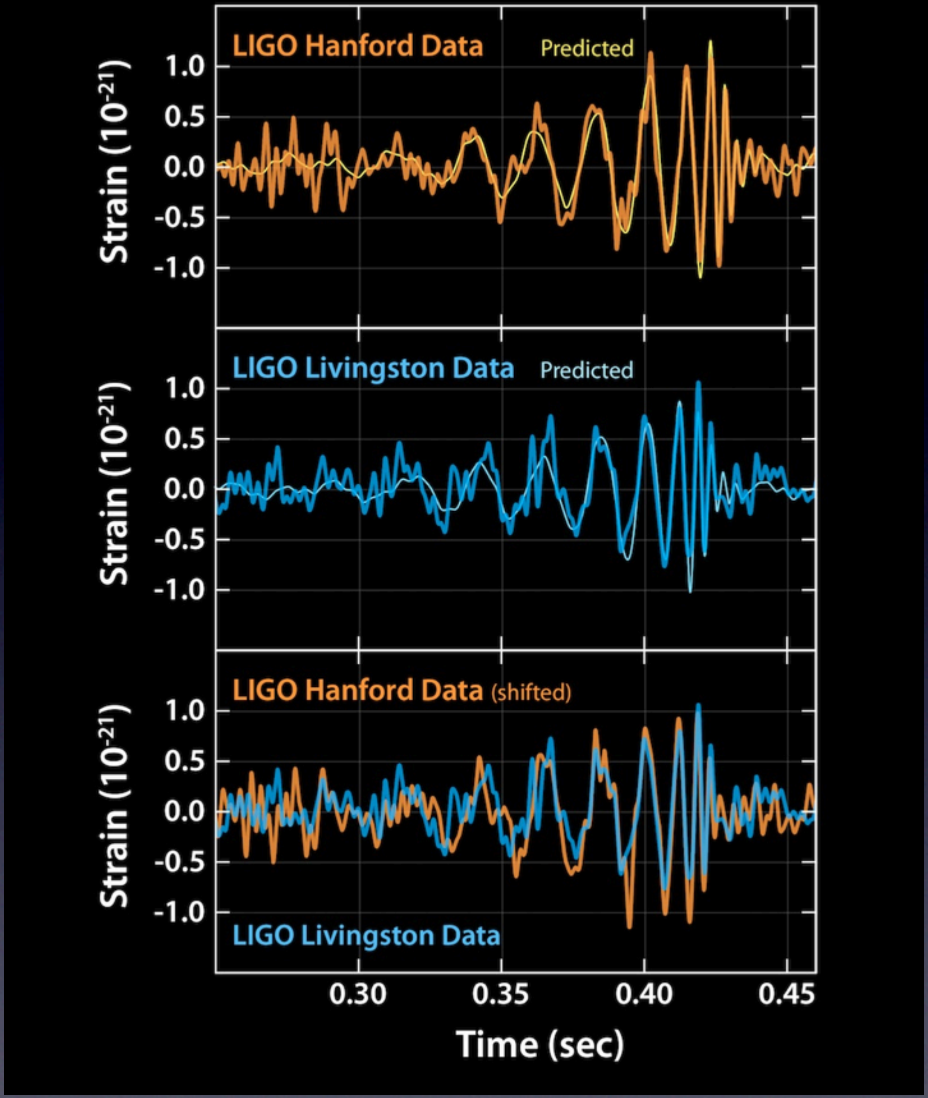
Implications for black hole and galaxy formation

- dark, compact objects of 10^6-10^{10} solar masses are present at the centers of most galaxies
- their properties are inconsistent with any plausible, long-lived astrophysical system except black holes
- arguments based on energy budgets strongly suggest that these are dead quasars
- black holes grow mostly by accretion of gas during luminous quasar phase and not by mergers, dark matter accretion, radiatively inefficient accretion, etc.
- correlation between galaxy velocity dispersion and black-hole mass implies a deep connection between the formation of galaxies and the formation of black holes
- black-hole mass is only ~0.5% of stellar mass in galaxy but energy released in forming the black hole is 1000‘s of times the energy released in forming the rest of the galaxy --- maybe the black hole quenches the formation of the galaxy (“feedback”)

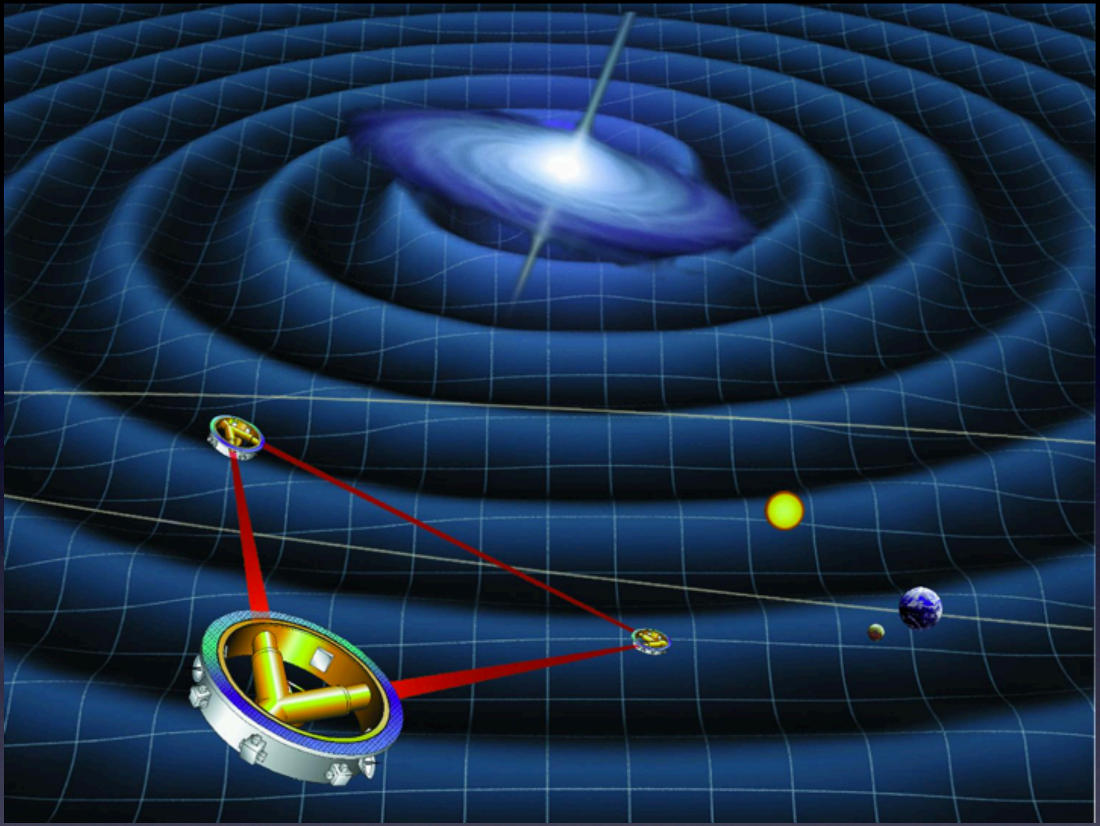
The last parsec problem



- galaxies form by hierarchical merging
- suppose two galaxies with black holes merge: after the merger, the black holes are left orbiting in the body of the merged galaxy
- dynamical friction causes the orbits of the black holes to decay, so they spiral to the center
- we may expect that binary black holes or black-hole mergers are common at galactic centers



- mergers of two supermassive black holes are the most energetic events in the universe
- release millions of solar masses of rest-mass energy within a few hours
- produce gravitational-wave signals similar to the LIGO events but on timescales longer by the ratio of masses (10^6-10^7)

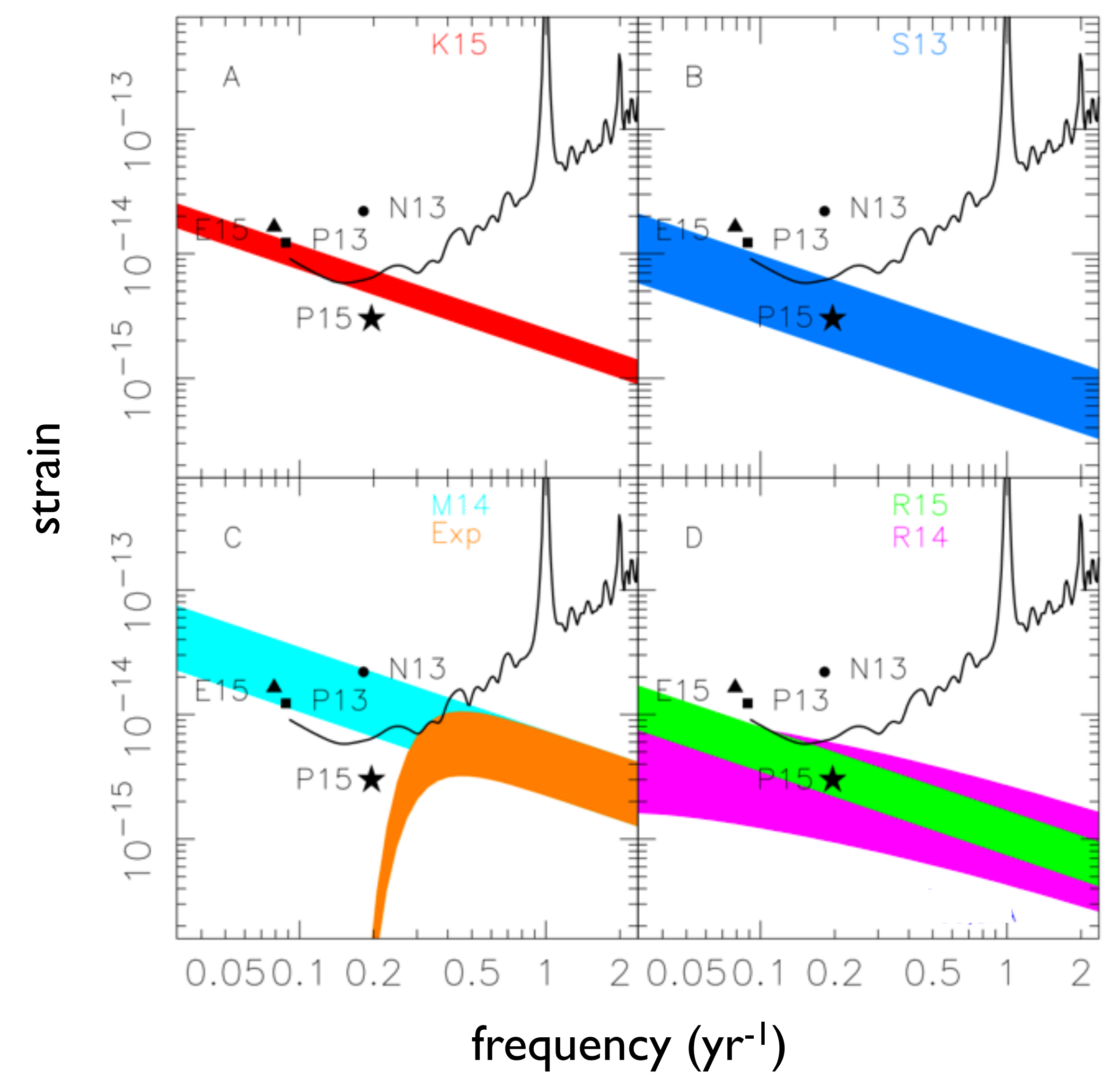


eLISA:

- proposed ESA mission to launch in 2030's
- 3 spacecraft form an interferometer with arm length of 10^6 km
- each spacecraft floats around a proof mass. LISA Pathfinder (launched December 2015) has successfully tested this technology

pulsar timing arrays:

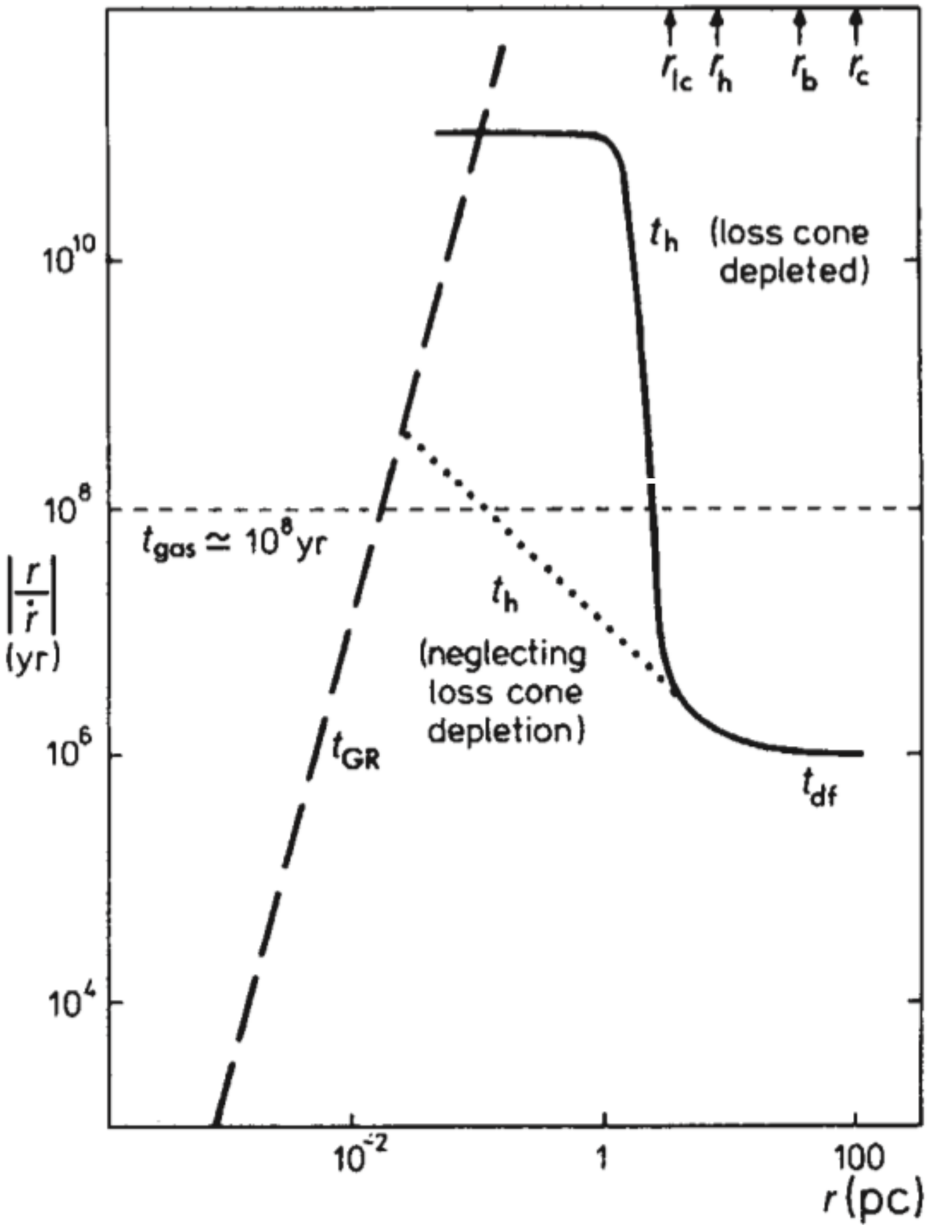
- measure stochastic background of gravitational waves from inspiral phase of binary supermassive black holes
- based on correlated timing residuals from dozens of pulsars
- arrays of radio telescopes in Europe (EPTA), U.S. (NANOGrav), and Australia (PPTA)



black points and curves: limits from various pulsar timing arrays

colored bands: theoretical models

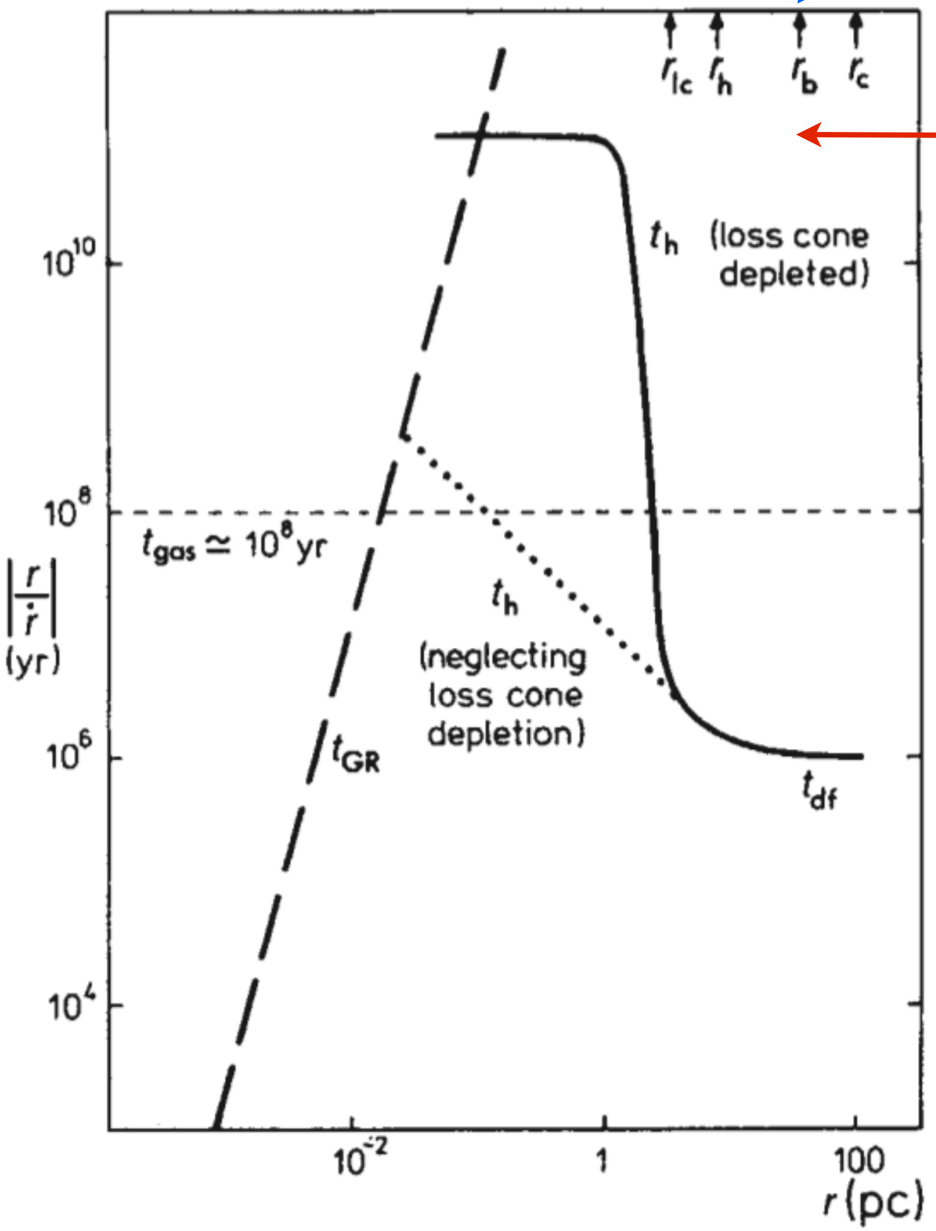
Shannon + (2015)



Begelman, Blandford & Rees (1980)

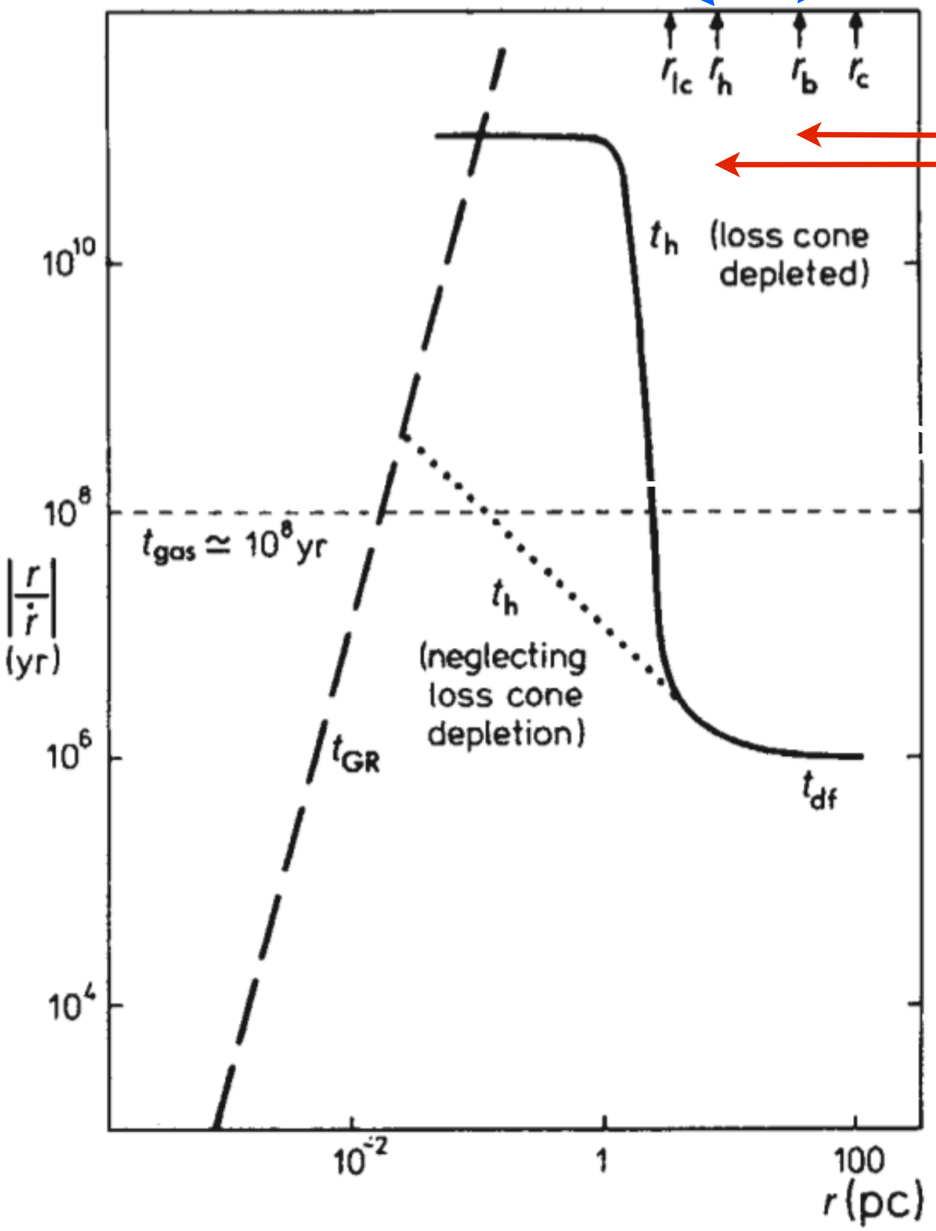
2. binary becomes bound

1. dynamical friction



Begelman, Blandford & Rees (1980)

$r_h \sim \frac{GM_\bullet}{\sigma^2}$



4. binary becomes "hard"

2. binary becomes bound

1. dynamical friction

3. more dynamical friction

Begelman, Blandford & Rees (1980)

$r_h \sim \frac{GM_\bullet}{\sigma^2}$

1. dynamical friction

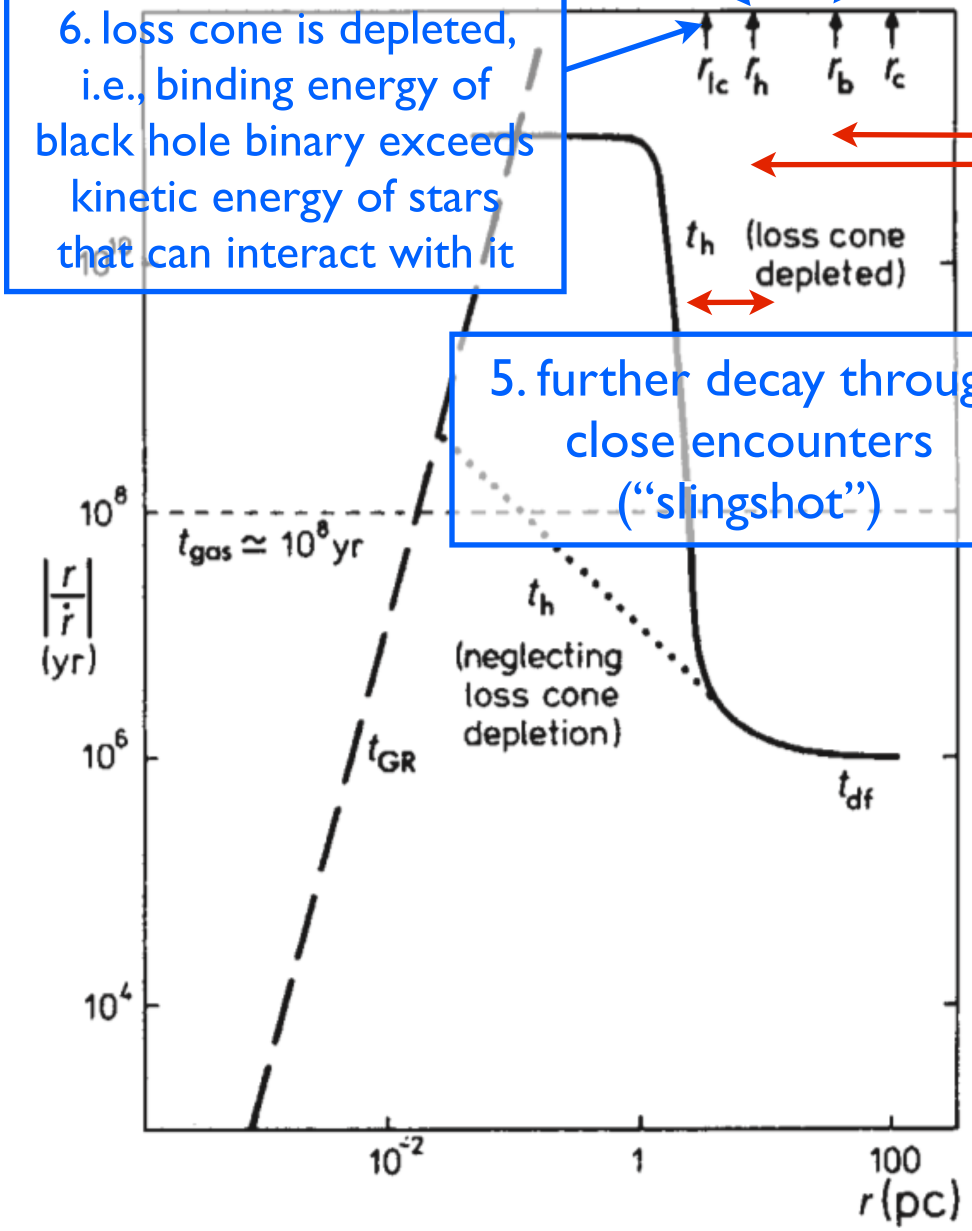
2. binary becomes bound

3. more dynamical friction

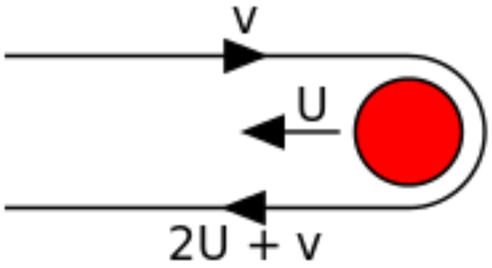
4. binary becomes "hard"

5. further decay through close encounters ("slingshot")

6. loss cone is depleted, i.e., binding energy of black hole binary exceeds kinetic energy of stars that can interact with it

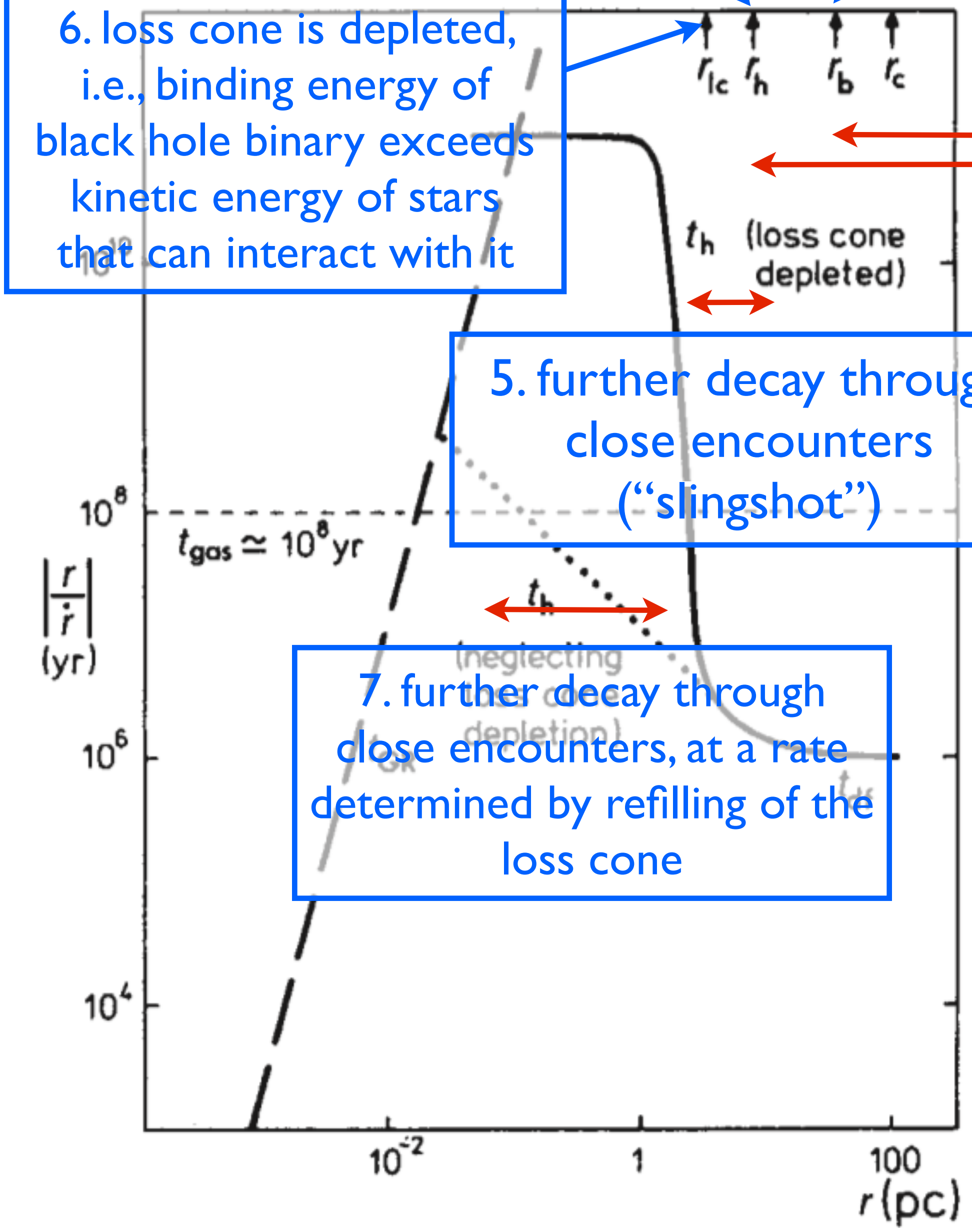


slingshot: encounters in which ejection velocity is comparable to the binary orbital velocity



loss cone: region of phase space in which stars have close encounters with the binary black hole

$r_h \sim \frac{GM_\bullet}{\sigma^2}$



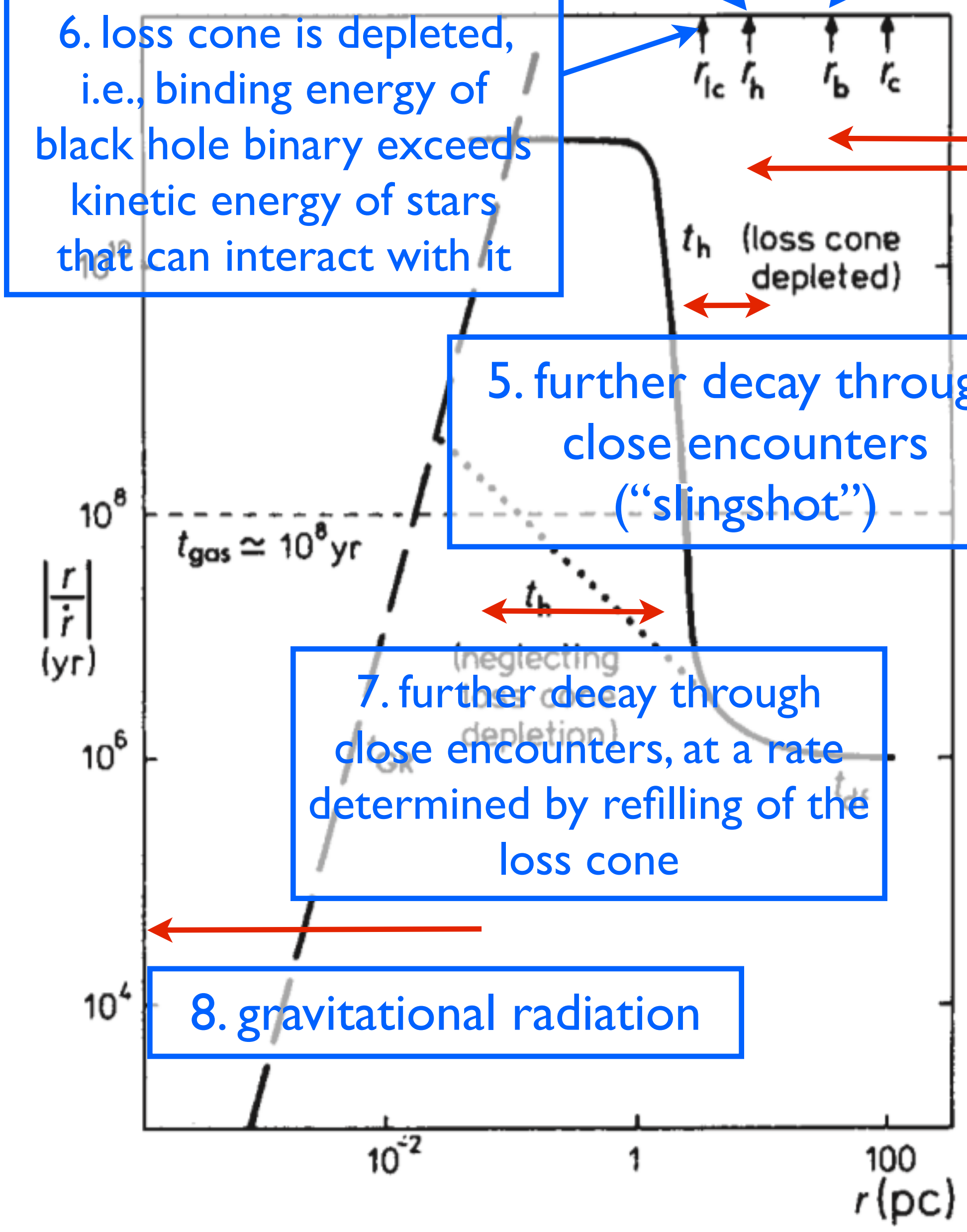
1. dynamical friction
2. binary becomes bound
3. more dynamical friction
4. binary becomes “hard”
5. further decay through close encounters (“slingshot”)
6. loss cone is depleted, i.e., binding energy of black hole binary exceeds kinetic energy of stars that can interact with it
7. further decay through close encounters, at a rate determined by refilling of the loss cone

Begelman, Blandford & Rees (1980)

$$r_h \sim \frac{GM_\bullet}{\sigma^2}$$

4. binary becomes "hard"

2. binary becomes bound



6. loss cone is depleted, i.e., binding energy of black hole binary exceeds kinetic energy of stars that can interact with it

1. dynamical friction

3. more dynamical friction

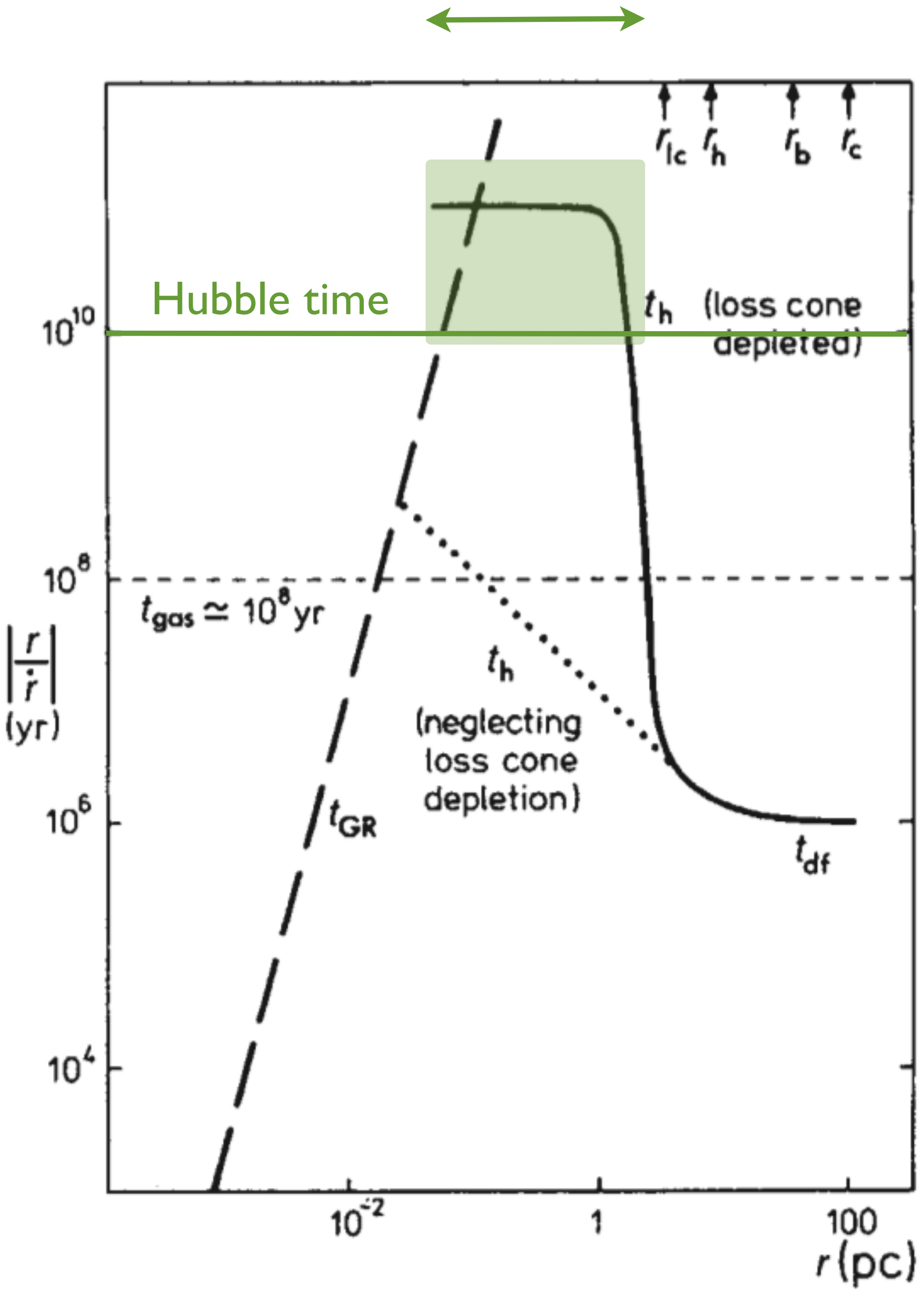
5. further decay through close encounters ("slingshot")

7. further decay through close encounters, at a rate determined by refilling of the loss cone

8. gravitational radiation

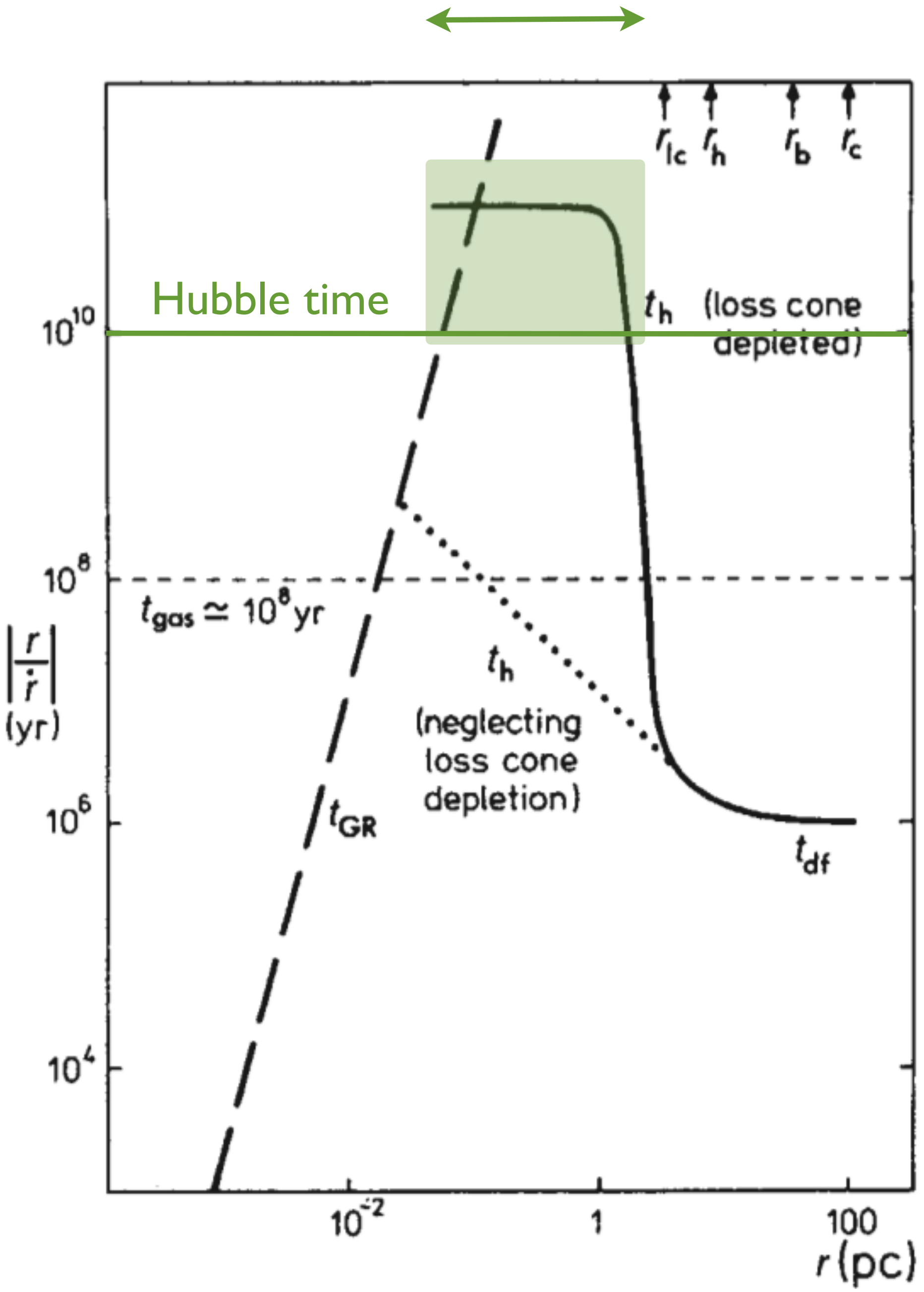
Begelman, Blandford & Rees (1980)

the "bottleneck"



Begelman, Blandford & Rees (1980)

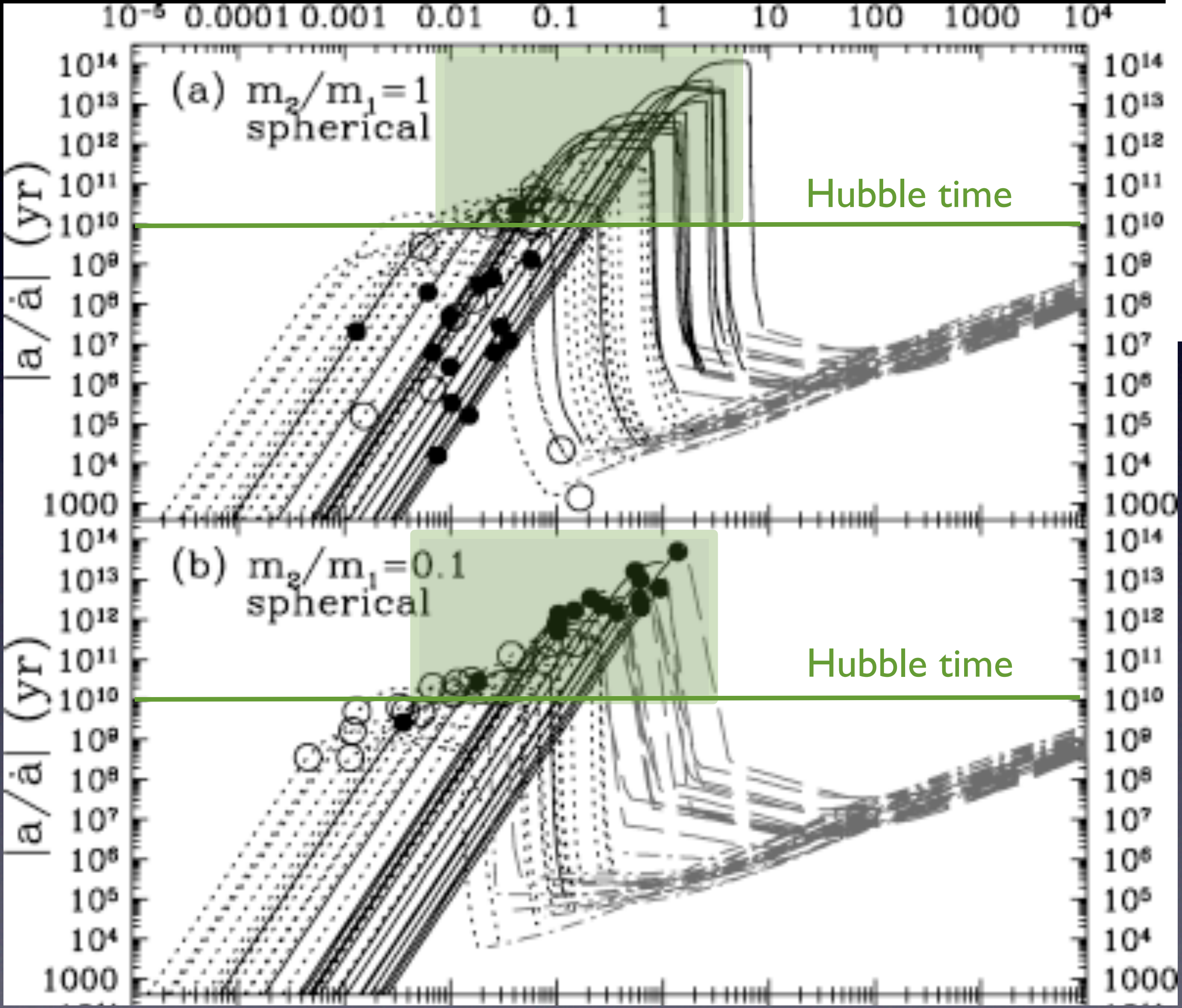
the “bottleneck”

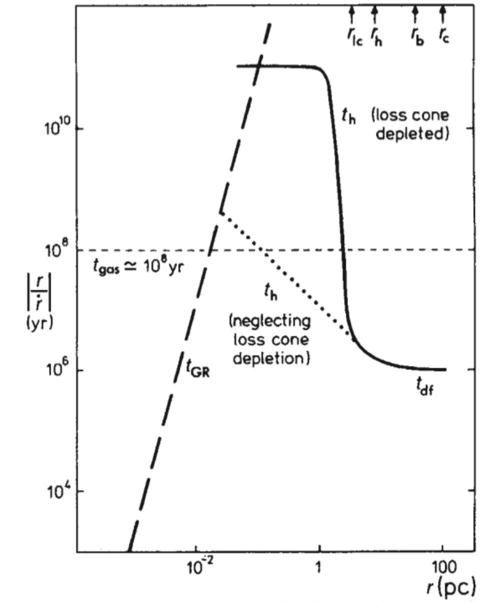


does the binary inspiral stall?

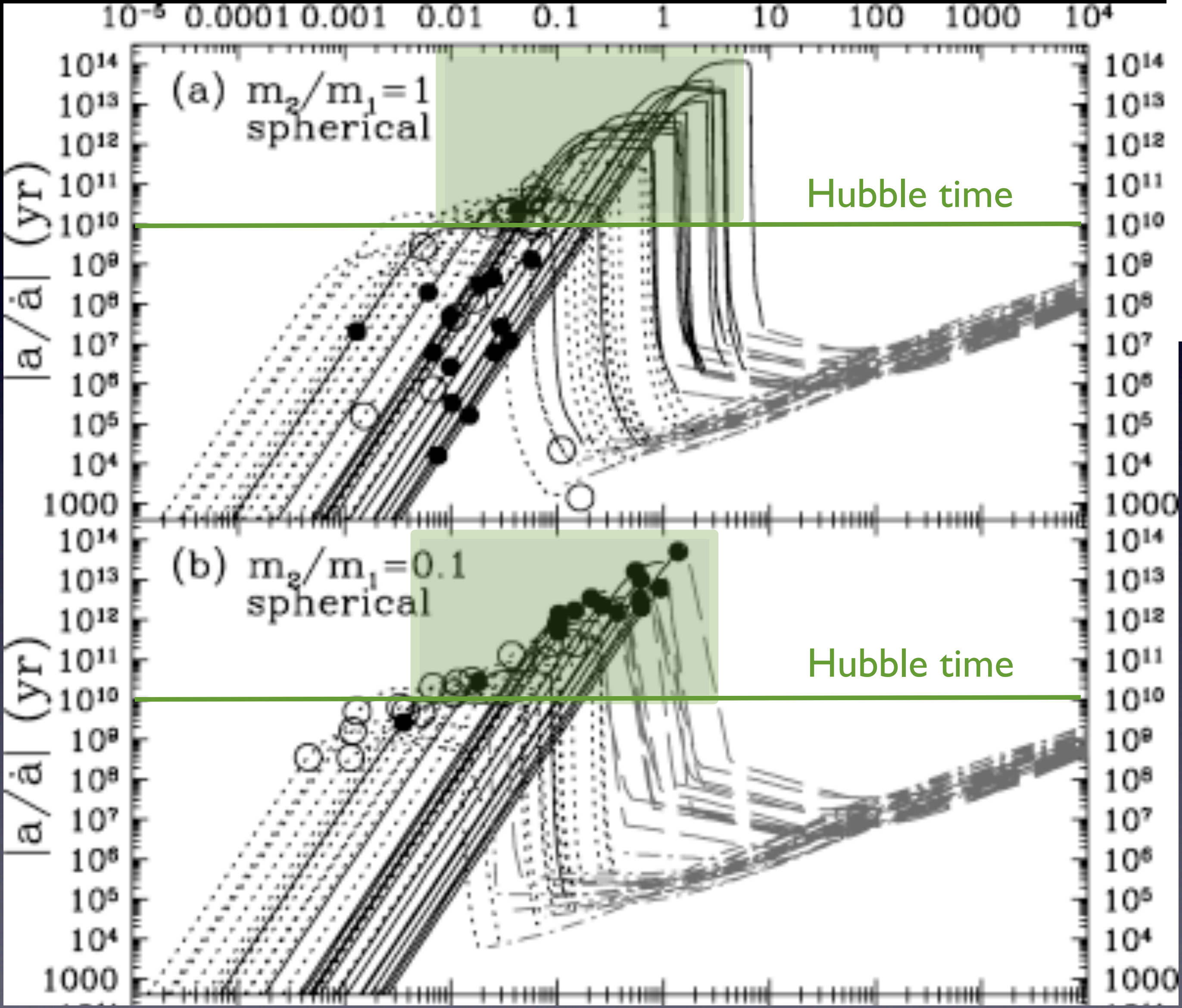
Begelman, Blandford & Rees (1980)

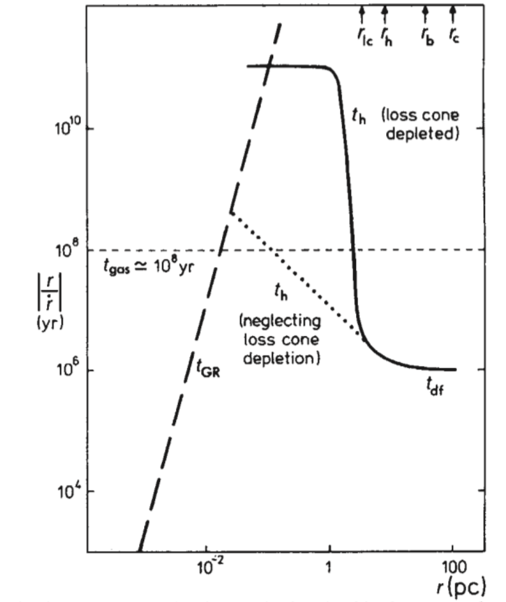
parsecs (pc)


10^-5
0.0001
0.001
0.01
0.1
1
10
100
1000
10^4
(a) m_2/m_1=1 spherical
|a/ȧ| (yr)
10^14
10^13
10^12
10^11
10^10
10^9
10^8
10^7
10^6
10^5
10^4
1000
Hubble time
(b) m_2/m_1=0.1 spherical
Hubble time


r_lc
r_h
r_b
r_c
10^10
10^8
10^6
10^4
|r/ṙ| (yr)
t_h (loss cone depleted)
t_gas ≃ 10^8 yr
t_h (neglecting loss cone depletion)
t_GR
t_df
10^-2
1
100
r (pc)

parsecs (pc)





calculations using models of real galaxies confirm there is a bottleneck of ~ 1-2 orders of magnitude in radius between 10 pc and 0.01 pc in which orbital decay time exceeds the Hubble time (the last parsec problem)

Possible solutions to the last parsec problem

- the problem arises because energy loss required to shrink the black hole binary exceeds the energy needed to eject all the stars that can interact with it (“the loss cone is emptied”)
    - in standard model loss cone is replenished only by star-star scattering so this is the rate-limiting step
    - other processes may refill the loss cone more rapidly
        - tidal forces, if the galaxy is non-spherical
        - massive perturbers such as star clusters, stellar-mass black holes, molecular clouds, etc. (Perets + 2007)
    - BUT even if the loss cone is full, the decay time exceeds the age of the universe in many galaxies
- gas drag
    - needs gas mass in the central parsec at least as large as the black hole mass
    - BUT gas drag can also accelerate inspiral in last stages and so reduces the number of binaries contributing to the stochastic background
- multiple black holes
    - arrival of a third black hole can remove energy from the binary through three-body interactions (Hoffman & Loeb 2007)

Possible solutions to the last parsec problem

- the problem arises because energy loss required to shrink the black hole binary exceeds the energy needed to eject all the stars that can interact with it (“the loss cone is emptied”)
  - in standard model loss cone is replenished only by star-star scattering so this is the rate-limiting step
  - other processes may refill the loss cone more rapidly
  - e
    universe in many galaxies

whether or not these processes permit most supermassive black hole binaries to merge, there should be a large population of sub-parsec binary black holes at the centers of galaxies

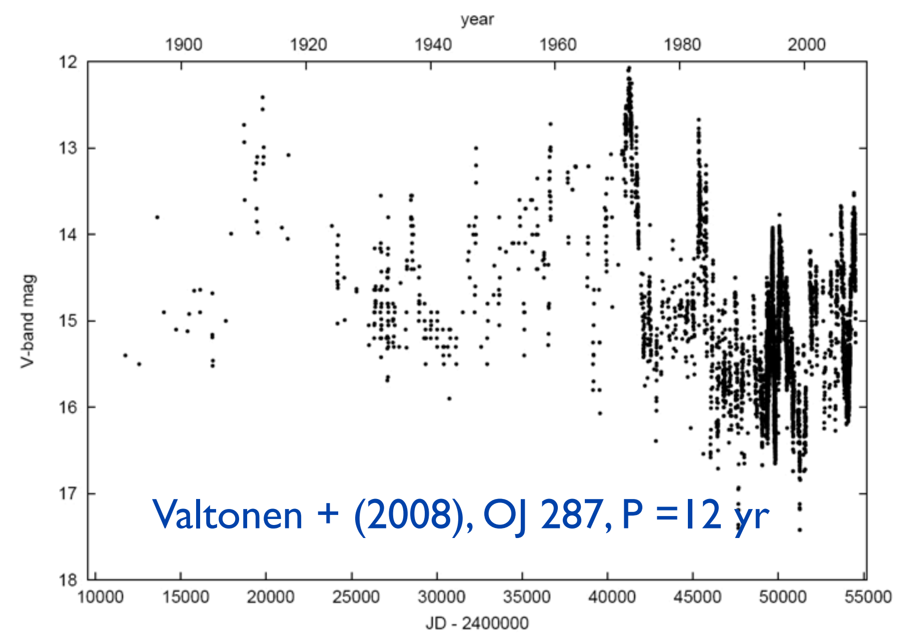
- gas drag
  - needs gas mass in the central parsec at least as large as the black hole mass
  - BUT gas drag can also accelerate inspiral in last stages and so reduces the number of binaries contributing to the stochastic background
- multiple black holes
  - arrival of a third black hole can remove energy from the binary through three-body interactions (Hoffman & Loeb 2007)

Looking for sub-parsec binary black holes

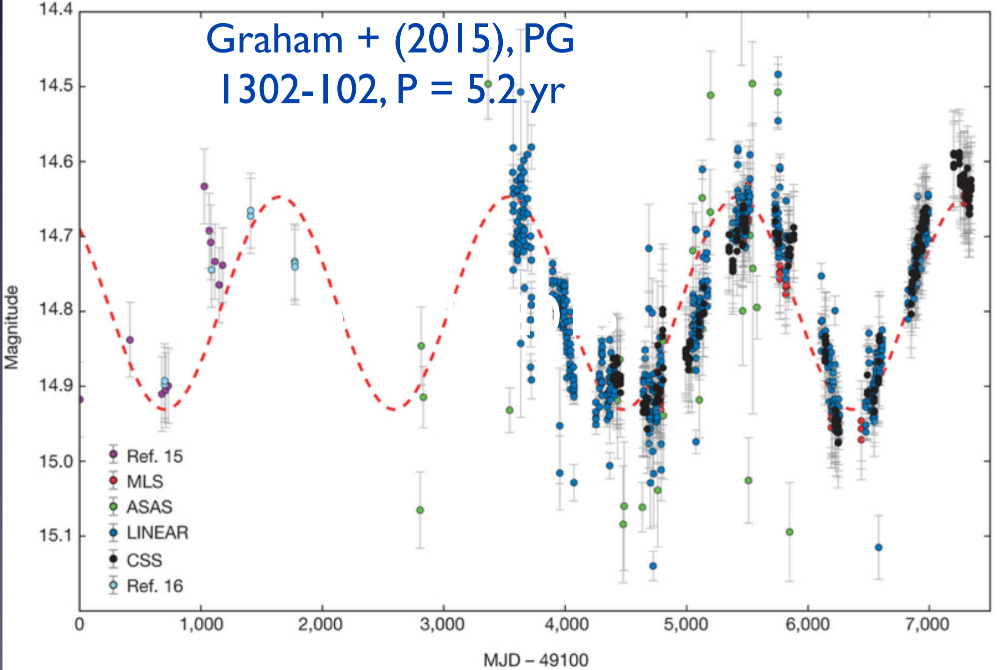
- even in the nearest galaxies, the expected separation of binary black holes in the bottleneck is too small to resolve
- search for double AGN inlong-baseline interferometry data yields only one example, at separation of 7 pc (Rodriguez + 2006, Burke-Spolaor 2011)
- look for double-peaked broad emission lines in AGN (Boroson & Lauer 2009, Tsai + 2013)
- look for offsets between broad emission line from AGN and velocity of its host galaxy (Bogdanović + 2009, Tsalmantza + 2011, Eracleous + 2012, Runnoe + 2015)
- look for drift in velocity of broad-line region relative to narrow lines (Eracleous + 2012, Decarli + 2012, Shen + 2013, Ju + 2013).
- look for periodic photometric variability in AGN

Looking for sub-parsec binary black holes

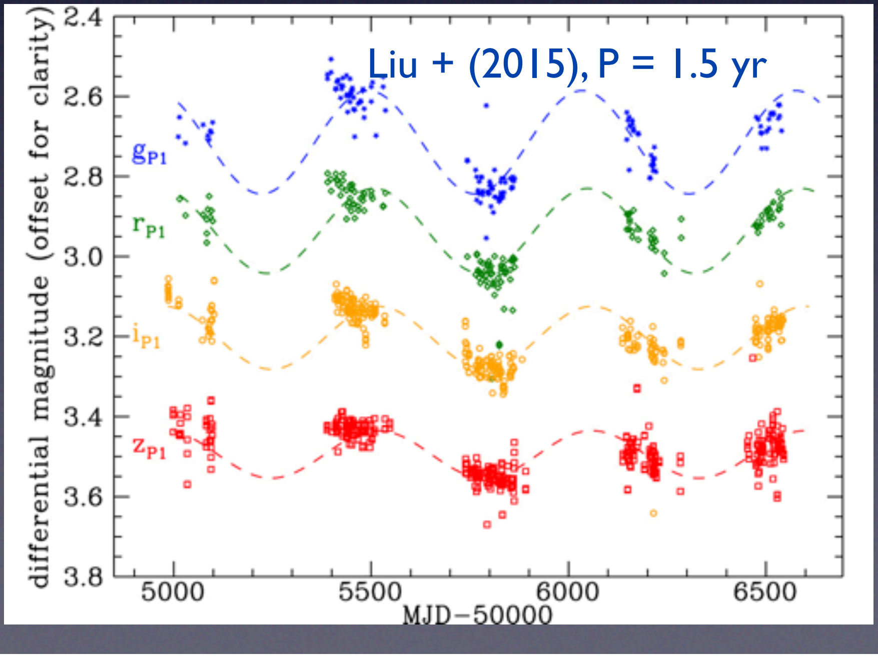
- look for periodic photometric variability in AGN



Valtonen + (2008), OJ 287, P =12 yr



Graham + (2015), PG 1302-102, P = 5.2 yr

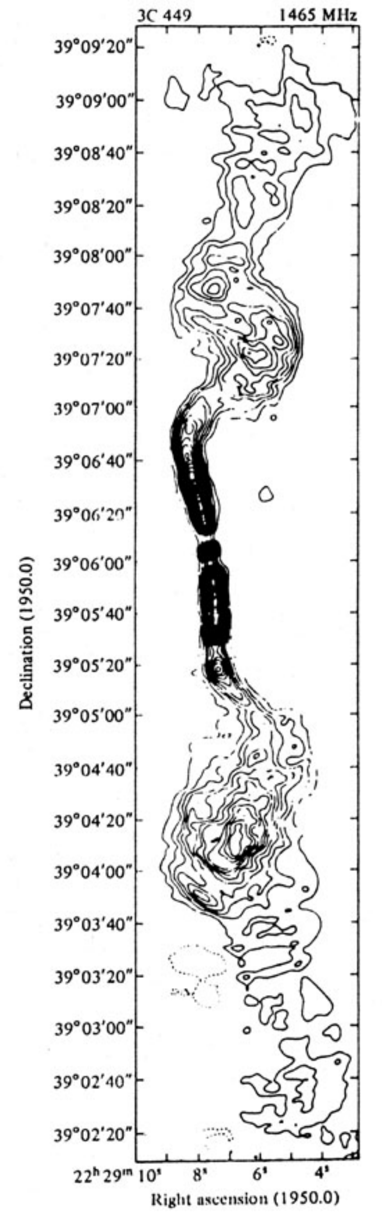


Liu + (2015), P = 1.5 yr

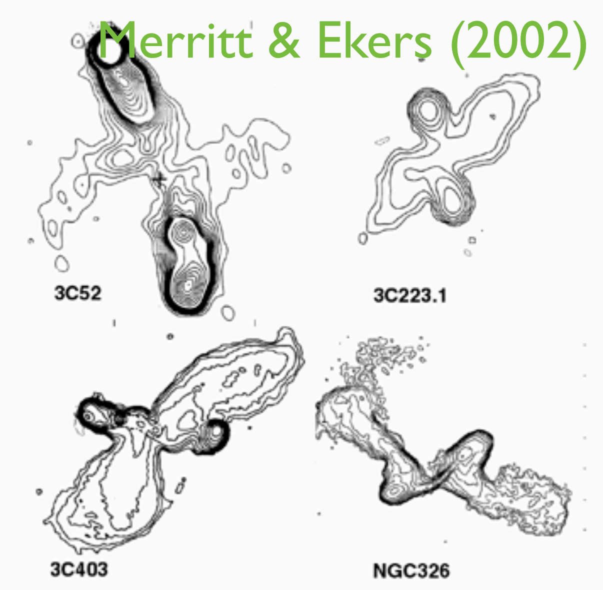
Looking for sub-parsec binary black holes

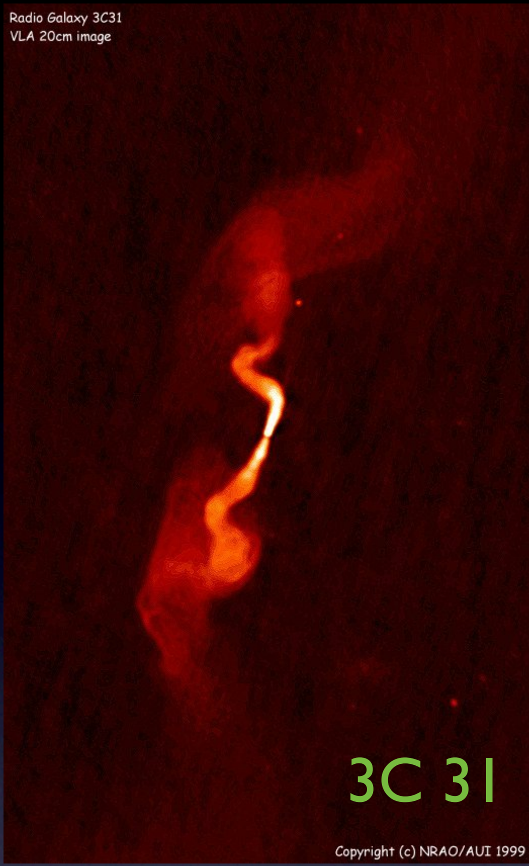
- indirect evidence from various anomalies in radio jets from AGN (Komossa 2006):
  - helical radio jets
  - double-double radio galaxies

3C 449



Merritt & Ekers (2002)





3C 31

Copyright (c) NRAO/AUI 1999

Summary:

- there is no direct evidence that supermassive black holes generally merge
- some black hole mergers probably do occur, but the rate and the mechanism for getting through the bottleneck are unknown
- binary black holes are more likely to merge in low-mass, gas-rich galaxies (bad for pulsar timing arrays, good for eLISA)
- a large fraction of galaxies are likely to contain sub-pc binary black holes, there is not yet any single convincing case
- the current data are consistent with a wide range of scenarios, including those in which all binaries stall in the sub-parsec bottleneck