Supermassive black holes from runaway mergers and accretion in nuclear star clusters

Konstantinos Kritos,[1][*] Emanuele Berti,[1] and Joseph Silk[1,2,3]

[1]*William H. Miller III Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA*
[2]*Insitut d'Astrophysique de Paris, UMR 7095 CNRS and UPMC, Sorbonne Université, F-75014 Paris, France*
[3]*University of Oxford, Oxford OX1 3RH, United Kingdom*

ABSTRACT

Rapid formation of supermassive black holes occurs in dense nuclear star clusters that are initially gas-dominated. Stellar-mass black hole remnants of the most massive cluster stars sink into the core, where a massive runaway black hole forms as a consequence of combined effects of repeated mergers and Eddington-limited gas accretion. The associated gravitational-wave signals of high-redshift extreme mass-ratio inspirals are a unique signature of the nuclear star cluster scenario.

Key words: (transients:) black hole mergers, galaxies: nuclei, (galaxies:) quasars: supermassive black holes

Introduction. The emergence of supermassive black holes (SMBHs) in the early Universe is an outstanding problem in modern cosmology and astrophysics. Many active galactic nuclei (AGN) powered by SMBHs have been identified at redshift $z \gtrsim 6$, with spectroscopically inferred black hole (BH) masses in the range 10^6–$10^8 \, M_\odot$ (Greene *et al.* 2023; Juodžbalis et al. 2024; Maiolino et al. 2024). As suggested by recent spectroscopic surveys with the James Webb Space Telescope (JWST), the abundance of red-selected AGN is a hundred times that expected in this redshift range (Matthee *et al.* 2023; Maiolino et al. 2023). The consensus is that SMBHs have grown out of massive seed black holes (Volonteri 2012), and the problem is then reduced to explaining the formation of those intermediate-mass BH seeds (Greene et al. 2020). One possible formation channel is the direct collapse of pristine gas clouds at $z > 10$ (Latif et al. 2022; Begelman et al. 2006). Alternative seeding mechanisms involve repeated BH mergers in dense star clusters (Antonini et al. 2019; Fragione & Silk 2020; Atallah et al. 2023; Kritos et al. 2023) and feeding by debris of successive tidal disruption events (TDEs; Stone et al. 2017; Rizzuto et al. 2023; Inayoshi et al. 2020). Moreover, several works have studied the rapid formation of supermassive stellar objects with masses in the range 10^3–$10^5 M_\odot$ within metal-poor massive star clusters through runaway collisions of stars and gas accretion (Tagawa et al. 2020; Vergara et al. 2023; Reinoso et al. 2023; Escala 2021; Schleicher et al. 2022; Das et al. 2021; Boekholt et al. 2018). These supermassive stellar objects can then collapse into intermediate-mass BH seeds.

Nuclear star clusters (NSCs) are extremely dense and massive stellar systems that have been observed to reside in the centers of most galaxies in the local Universe (Neumayer et al. 2020). In the nearby Universe, NSCs span the mass range from $10^5 M_\odot$ to $10^9 M_\odot$, with evidence for in situ star formation in the more massive systems (Fahrion et al. 2022). At high redshift, a number of lensed star clusters with effective radii as small as a parsec (pc) and with masses of order $10^6 M_\odot$ have been identified (Vanzella *et al.* 2023; Adamo et al. 2024). There is some theoretical support for the fragmentation of the central region of protogalaxies into millions of metal-poor stars, forming clusters with half-mass radii $\lesssim 0.5$ pc (Devecchi et al. 2010).

In this paper, we use a semianalytic NSC model and demonstrate the possibility of rapidly forming SMBHs through the combined processes of repeated BH mergers and Eddington-limited gas accretion, assuming that a substantial amount of gas can be retained within the cluster for a few tens of millions of years (Myr).

Nuclear star cluster model. Our numerical model for the NSC implements a two-mass system consisting of $N_\star$ stars, each of mass $m_\star$, and $N_{\rm BH}$ stellar BHs, each of mass $m_{\rm BH}$ and nonspinning, with an initial mass $M_{\rm g}$ of residual gas not turned into stars. We find that typically one of the centrally located and initially stellar mass BHs grows dramatically via gas, star, and BH accretion, and we denote its mass and spin by $M_{\rm BH}$ and $\chi_{\rm BH}$, respectively. The total cluster mass is therefore $M_{\rm cl} = m_\star N_\star + m_{\rm BH} N_{\rm BH} + M_{\rm g} + M_{\rm BH}$. Since the growing BH is one of the stellar BHs, initially $M_{\rm BH,0} = m_{\rm BH}$ and $\chi_{\rm BH,0} = 0$. This "oligarchic" growth scenario can occur in dynamical environments (see e.g. Matsubayashi et al. 2004; Kovetz et al. 2018), where the formation of multiple massive BHs within the same system is significantly suppressed by their preferential accretion by the most massive object (Kritos et al. 2023).

The gas, composed primarily of ionized hydrogen with sound speed $c_{\rm s} \simeq 10 \, {\rm km \, s^{-1}}$ (Krumholz & Matzner 2009), is removed exponentially over a characteristic gas expulsion timescale $\tau_{\rm ge}$ (Baumgardt & Kroupa 2007). The growing BH may accrete some of this gas. We assume an Eddington-limited Bondi accretion rate onto the BH seed $\dot{M}_{\rm acc}$ with Salpeter timescale of 50 Myr, and a time-evolving Plummer gas density $\rho_{\rm g} = 3M_{\rm g}/[4\pi (r_{\rm h}/1.3)^3]$, where $r_{\rm h}$ is the half-mass radius. As a consequence of gradual gas removal, $r_{\rm h}$ expands adiabatically (Hills 1980).

The half-mass relaxation time $\tau_{\rm rh}$, modified for two-mass systems following Eq. (10) from Antonini & Gieles (2020), is the characteristic time over which a fraction $\zeta \simeq 0.1$ of the total cluster energy

[*] E-mail: kkritos1@jhu.edu

$E_{cl} = -\kappa G M_{cl}^2/(2r_h)$ (with $\kappa \simeq 0.4$, see Spitzer & Hart 1971) can be transferred throughout the cluster via two-body relaxation (Breen & Heggie 2013). The core is dominated by BHs, which partially decouple from stars due to the impossibility of energy equipartition ("Spitzer instability"). The core collapses on a time-scale τ_{cc} given by Eq. (5) from Portegies Zwart & McMillan (2002).

At this point, three-body hard binaries (3bb) can form and heat the system. The rate of energy production $\dot{E}_{cl} = -\zeta E_{cl}/\tau_{rh}$ depends on r_h, and is controlled by the efficiency of two-body relaxation ("Hénon's principle"). We compute the energy generated by a single 3bb Q_{3bb}, and use this condition to calculate the 3bb rate $R_{3bb} = \dot{E}_{cl}/Q_{3bb}$ (as in Heggie & Hut 2003, page 244). If all BHs but one have been ejected from the system, we assume that energy generation proceeds via tidal disruption of stars accreting onto the runaway BH. The TDE rate is $R_{TDE} = \dot{E}_{cl}/Q_{TDE}$, where $Q_{TDE} = G M_{BH} \overline{m}_\star/(2r_T)$ is the heat released from a single TDE and r_T is the tidal radius (Rees 1988). The main advantage of our method, as in Antonini & Gieles (2020), is that it is not necessary to calculate the properties of the core, while we can incorporate the effect of a cusp around the SMBH.

The differential equations that govern the evolution of the system are given by

$$\frac{d\overline{m}_\star}{dt} = -\nu \frac{\overline{m}_\star}{t} \Theta(t - t_{se}), \tag{1a}$$

$$\frac{dN_\star}{dt} = -\xi_e \frac{N_\star}{\tau_{rh}} - R_{TDE}, \tag{1b}$$

$$\frac{dN_{BH}}{dt} = -k_{ejs} R_{3bb}, \tag{1c}$$

$$\frac{dM_g}{dt} = -\frac{M_g}{\tau_{ge}} - \dot{M}_{acc}, \tag{1d}$$

$$\frac{dM_{BH}}{dt} = \dot{M}_{acc} + f_{TDE} \overline{m}_\star R_{TDE}, \tag{1e}$$

$$\frac{dr_h}{dt} = \left(\frac{\zeta r_h}{\tau_{rh}} + \frac{2r_h}{M_{cl}} \frac{dM_{cl}}{dt} \right) \Theta(t - \tau_{cc})$$
$$- \frac{r_h}{M_{cl}} \left(N_\star \frac{d\overline{m}_\star}{dt} - \frac{M_g}{\tau_{ge}} \right). \tag{1f}$$

Equations (1a) and (1b) can be compared with Alexander et al. (2014). The main difference is a loss term due to TDEs, R_{TDE}, on the right-hand side of Eq. (1b). The first three terms on the right-hand side of Eq. (1f) follow Antonini & Gieles (2020), but we also include an adiabatic expansion term due to gas expulsion. The quantity k_{ejs} in Eq. (1c) is the number of single BHs ejected per unit binary. We set $\nu = 0.07$, $t_{se} = 2$ Myr (Antonini & Gieles 2020), and $\xi_e = 0.0074$ (Alexander et al. 2014).

At each iteration, we sample k_{3bb}, the number of 3bbs formed, from a Poisson distribution with parameter $\lambda = R_{3bb} dt$, where dt is the time-step of the simulation. Every 3bb is formed with components M_{BH} and m_{BH}, with its initial semimajor axis a_0 determined by sampling the initial hardness parameter $\eta_0 \geq 1$ (Morscher et al. 2015). Binaries harden until they reach their ejection (a_{ejb}) or gravitational-wave (GW) radius (a_{GW}), assuming thermal eccentricity. To compute the merger remnant mass m_r, spin χ_r, and relativistic recoil—or GW kick v_k—we implement the equations in Sec. V.A-C of Gerosa & Kesden (2016) and references therein, which are based on numerical relativity fits and assume isotropic spin directions. If $a_{GW} > a_{ejb}$ the binary merges within the cluster through the emission of GWs, and if $v_k < v_{esc}$, the merger remnant is retained in the cluster. In this case, we discretely update the runaway BH mass, spin, and number by the substitutions $M_{BH} \rightarrow m_r$, $\chi_{BH} \rightarrow \chi_r$, and $N_{BH} \rightarrow N_{BH} - 1$, while in any case of ejection, $M_{BH} \rightarrow m_{BH}$, $\chi_{BH} \rightarrow 0$, and $N_{BH} \rightarrow N_{BH} - 2$.
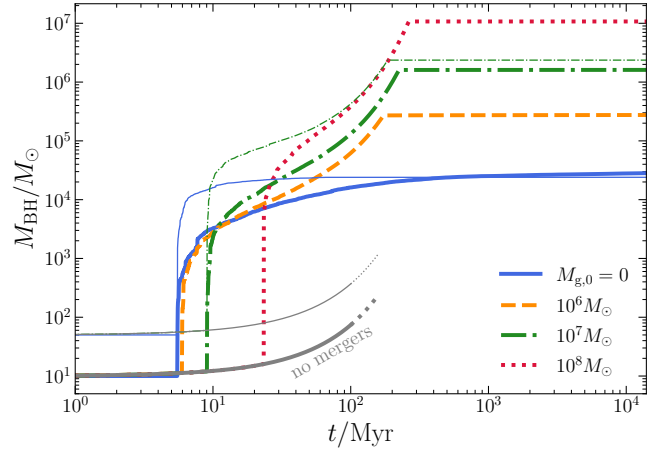
Figure 1. Time evolution of the growing BH mass M_{BH} for different initial amounts of gas in the system. The gas expulsion time-scale is set to $\tau_{ge} = 100$ Myr, the initial stellar number to $N_{\star,0} = 10^7$, and the initial half-mass radius to $r_{h,0} = 0.1$ pc. Thick (thin) lines correspond to simulations with $m_{BH} = 10 M_\odot$ ($m_{BH} = 50 M_\odot$). For comparison, we also show how a $10 M_\odot$ seed (thick gray) and a $50 M_\odot$ seed (thin gray) would grow under the sole effect of Eddington accretion.

Black hole growth. We implement a fourth-order Runge-Kutta scheme to numerically solve the system of equations (1) using a time step of $dt = 0.01$ Myr (chosen to be much smaller than τ_{rh}). We set $N_{\star,0} = 10^7$, $r_{h,0} = 0.1$ pc (where the "0" refers to initial values) and $\tau_{ge} = 100$ Myr. Using a Kroupa initial mass function in the range $[0.08, 150] M_\odot$, the average initial stellar mass is $\overline{m}_{\star,0} = 0.59 M_\odot$, while the fraction of BH progenitors with mass $> 20 M_\odot$ is $\simeq 0.18\%$. We assume $m_{BH} = 10 M_\odot$, but to quantify uncertainties, we also repeat the analysis with $m_{BH} = 50 M_\odot$. In Fig. 1, we show the growth of M_{BH} for four selected values of $M_{g,0} = 0, 10^6, 10^7, 10^8 M_\odot$.

All simulated systems have $v_{esc,0} > 300$ km s^{-1}, allowing repeated mergers to form a runaway BH (Antonini et al. 2019). The BH rapidly grows by hundreds of repeated BH mergers (Fig. 2, left panel) with a merger rate that peaks at ~ 100 Myr^{-1} right after core collapse (see inset in the left panel of Fig. 2) and steadily drops as the system expands.

Significant amounts (20%–30%) of gas can be accreted within 200 Myr, rapidly bringing the BH mass into the SMBH regime. The residual gas does not fragment because of heating by UV radiation from cluster stars. Our accretion channel is similar to the direct collapse scenario but occurs on the Salpeter timescale, rather than on the dynamical timescale. The combination of mergers and accretion helps accelerate the growth because Eddington-limited accretion cannot itself grow a stellar-mass BH seed above $1000 M_\odot$ on a similar timescale (compare the gray lines in Fig. 1, which ignore BH mergers).

Starting from the initial radius $r_{h,0} = 0.1$ pc, the cluster inflates by a factor of 10–100 (depending on $M_{g,0}$), so that $r_{h,0}$ is on the parsec scale after a Hubble time (see Fig. 2, right panel). When $m_{BH} = 50 M_\odot$, a larger amount of energy is produced, causing r_h to expand to above 100 pc (a feature that resembles the bulges of present-day dwarf galaxies). The rapid mass loss in the first few hundred Myr is due to gas expulsion (see the inset of the right panel of Fig. 2). After core collapse, the gas-richest system (dotted red line) undergoes a phase of contraction due to balanced evolution that starts at $\simeq 22$ Myr. Once the gas has been expelled, the energy production
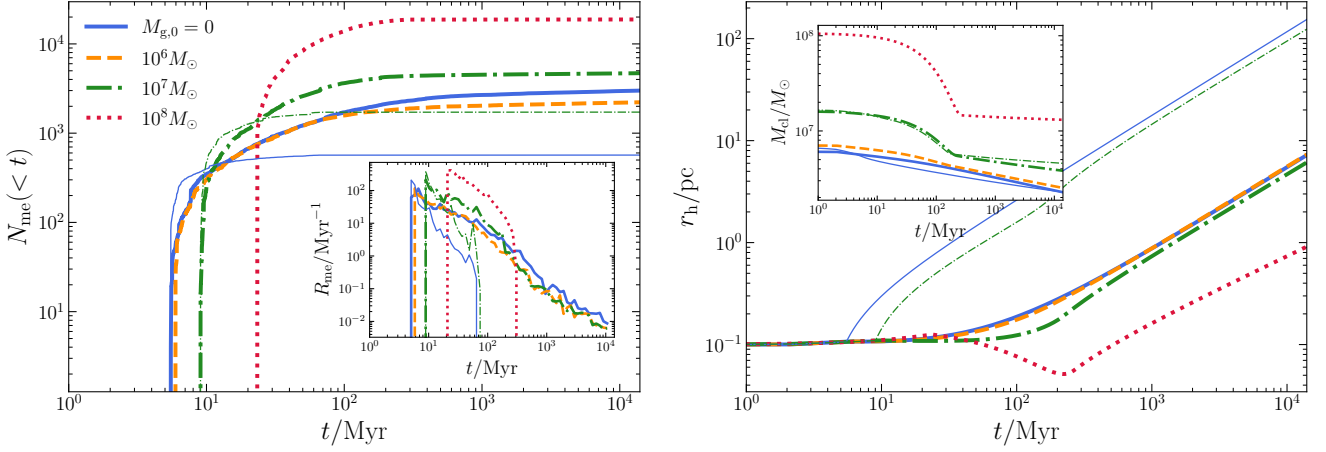
Figure 2. Time evolution of the cumulative number of BH mergers N_{me} (left panel), merger rate R_{me} (left panel, inset), half-mass radius r_h (right panel), and cluster mass M_{cl} (right panel, inset) under the same initial conditions used in Fig. 1.
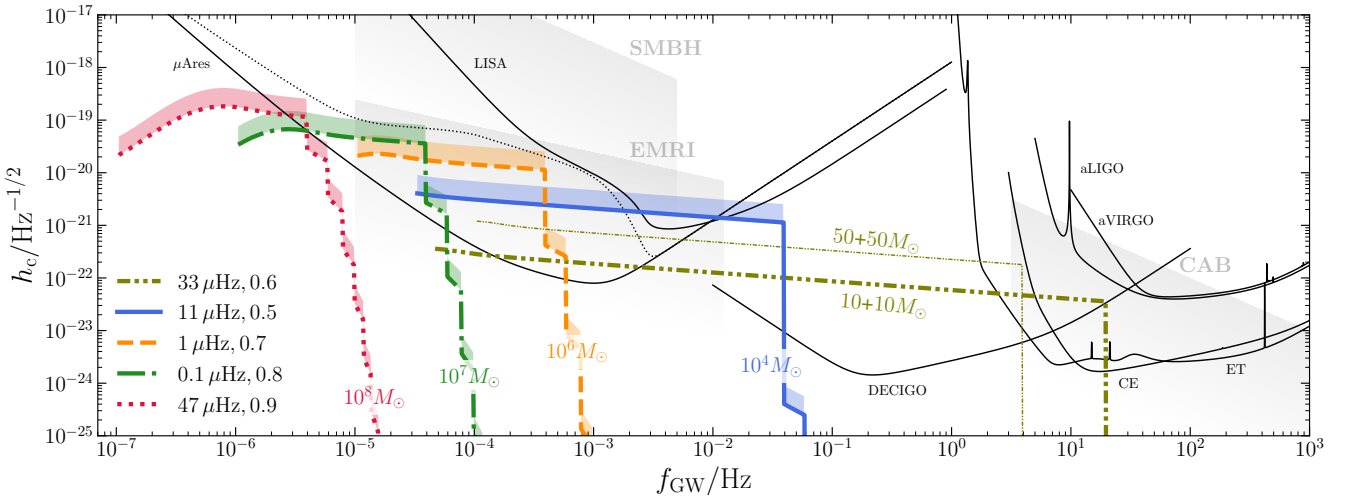

Figure 3. Characteristic GW strain amplitude h_c (computed as in Bonetti & Sesana 2020) as a function of detector-frame GW frequency for five extreme mass-ratio inspirals (EMRIs) at redshift $z = 10$ (colored lines). We use cosmological parameters from Planck 18 (Aghanim et al. 2020). The initial orbital frequency and eccentricity are indicated in the legend; labels on the colored curves refer to the mass of the primary. The bands indicate the h_c range as the secondary mass is varied between $10 M_\odot$ (lower bound) and $50 M_\odot$ (upper bound). The black solid curves correspond to the noise amplitudes of present and planned GW observatories: aLIGO (https://dcc.ligo.org/LIGO-T1800044/public), aVIRGO (https://git.ligo.org/sensitivity-curves/observing-scenario-paper-2019-2020-update), ET (https://www.et-gw.eu/index.php/etsensitivities), CE (https://cosmicexplorer.org/sensitivity.html), DECIGO (Yagi & Seto 2011), LISA (Robson et al. 2019), and μAres (Sesana et al. 2021). Finally, the shaded regions correspond to lower redshift ($z = 1$–2) binary sources from stellar-mass compact astrophysical binaries (CAB), EMRIs, and SMBH binaries.

term in Eq. (1f) dominates, and the system starts to expand again at $t \simeq 200$ Myr.

We have demonstrated with an example that a combination of gas accretion and BH mergers is capable of growing stellar-mass BH seeds into the SMBH regime within only 200 Myr. Most notably, if such a cluster forms at $z = 15$, then (depending on the initial gas content) an SMBH can assemble by $z \simeq 10$. These SMBH seeds can then further grow by subsequent gas accretion episodes during gas inflows into the NSC following a galaxy merger. This hierarchical scenario can be studied by combining our formalism with galaxy merger trees.

We recall that very massive NSCs are observed locally in many galaxies. Many of them are inferred to have gas because we observe young stars. We find that NSC evolution inevitably results in expansion by an order of magnitude or more (see the right panel of Fig. 2). In contrast, at formation, many NSCs were massive and compact. This supports our contention that NSCs are a promising environment for generating SMBHs via an enhanced merger rate density in the high-redshift Universe.

The ultimate test of our model would be the observation of individual GW signals from SMBH–BH mergers in NSC cores, which may be possible with future GW observatories (Fig. 3). The GW background produced by these sources is a distinctive signature of the NSC scenario that deserves further study. Our predictions motivate the development of the science case for next-generation GW observatories spanning the μHz frequency range between LISA and Pulsar

Timing Arrays. Our proposed GW background should be taken into account when estimating astrophysical confusion noise sources that can affect the sensitivity of proposed μHz detectors, such as μAres. For comparison, the black-dotted line in Fig. 3 shows an estimate of the confusion noise from SMBH and white dwarf binaries in the μAres band (Sesana et al. 2021).

Data availability. The output of the simulations is available upon a reasonable request to the corresponding author. A Python implementation of our model through the code Nuce (for "NUclear Cluster Evolution") is publicly available on GitHub at the URL `https://github.com/Kkritos/Nuce`.

REFERENCES

Adamo A., et al., 2024, arXiv e-prints, p. arXiv:2401.03224

Aghanim N., et al., 2020, Astron. Astrophys., 641, A6

Alexander P. E. R., Gieles M., Lamers H. J. G. L. M., Baumgardt H., 2014, MNRAS, 442, 1265

Antonini F., Gieles M., 2020, Mon. Not. Roy. Astron. Soc., 492, 2936

Antonini F., Gieles M., Gualandris A., 2019, Mon. Not. Roy. Astron. Soc., 486, 5008

Atallah D., Trani A. A., Kremer K., Weatherford N. C., Fragione G., Spera M., Rasio F. A., 2023, Mon. Not. Roy. Astron. Soc., 523, 4227

Baumgardt H., Kroupa P., 2007, Mon. Not. Roy. Astron. Soc., 380, 1589

Begelman M. C., Volonteri M., Rees M. J., 2006, Mon. Not. Roy. Astron. Soc., 370, 289

Boekholt T. C. N., Schleicher D. R. G., Fellhauer M., Klessen R. S., Reinoso B., Stutz A. M., Haemmerlé L., 2018, MNRAS, 476, 366

Bonetti M., Sesana A., 2020, Phys. Rev. D, 102, 103023

Breen P. G., Heggie D. C., 2013, Mon. Not. Roy. Astron. Soc., 432, 2779

Das A., Schleicher D. R. G., Leigh N. W. C., Boekholt T. C. N., 2021, MNRAS, 503, 1051

Devecchi B., Volonteri M., Colpi M., Haardt F., 2010, MNRAS, 409, 1057

Escala A., 2021, ApJ, 908, 57

Fahrion K., Leaman R., Lyubenova M., van de Ven G., 2022, A&A, 658, A172

Fragione G., Silk J., 2020, Mon. Not. Roy. Astron. Soc., 498, 4591

Gerosa D., Kesden M., 2016, Phys. Rev. D, 93, 124066

Greene J. E., Strader J., Ho L. C., 2020, ARA&A, 58, 257

Greene *et al.* J. E., 2023, arXiv e-prints, p. arXiv:2309.05714

Heggie D., Hut P., 2003, The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics. Cambridge Univ. Press

Hills J. G., 1980, ApJ, 235, 986

Inayoshi K., Visbal E., Haiman Z., 2020, Ann. Rev. Astron. Astrophys., 58, 27

Juodžbalis I., et al., 2024, arXiv e-prints, p. arXiv:2403.03872

Kovetz E. D., Cholis I., Kamionkowski M., Silk J., 2018, Phys. Rev. D, 97, 123003

Kritos K., Berti E., Silk J., 2023, Phys. Rev. D, 108, 083012

Krumholz M. R., Matzner C. D., 2009, ApJ, 703, 1352

Latif M. A., Whalen D. J., Khochfar S., Herrington N. P., Woods T. E., 2022,] 10.1038/s41586-022-04813-y

Maiolino R., et al., 2023

Maiolino R., et al., 2024, Nature, 627, 59

Matsubayashi T., Shinkai H., Ebisuzaki T., 2004, Astrophys. J., 614, 864

Matthee *et al.* J., 2023, arXiv e-prints, p. arXiv:2306.05448

Morscher M., Pattabiraman B., Rodriguez C., Rasio F. A., Umbreit S., 2015, Astrophys. J., 800, 9

Neumayer N., Seth A., Böker T., 2020, A&ARv, 28, 4

Portegies Zwart S. F., McMillan S. L. W., 2002, Astrophys. J., 576, 899

Rees M. J., 1988, Nature, 333, 523

Reinoso B., Klessen R. S., Schleicher D., Glover S. C. O., Solar P., 2023, MNRAS, 521, 3553

Rizzuto F. P., Naab T., Rantala A., Johansson P. H., Ostriker J. P., Stone N. C., Liao S., Irodotou D., 2023, Mon. Not. Roy. Astron. Soc., 521, 2930

Robson T., Cornish N. J., Liu C., 2019, Class. Quant. Grav., 36, 105011

Schleicher D. R. G., et al., 2022, MNRAS, 512, 6192

Sesana A., et al., 2021, Exper. Astron., 51, 1333

Spitzer Lyman J., Hart M. H., 1971, ApJ, 164, 399

Stone N. C., Küpper A. H. W., Ostriker J. P., 2017, Mon. Not. Roy. Astron. Soc., 467, 4180

Tagawa H., Haiman Z., Kocsis B., 2020, ApJ, 892, 36

Vanzella *et al.* E., 2023, ApJ, 945, 53

Vergara M. C., Escala A., Schleicher D. R. G., Reinoso B., 2023, MNRAS, 522, 4224

Volonteri M., 2012, Science, 337, 544

Yagi K., Seto N., 2011, Phys. Rev. D, 83, 044011

This paper has been typeset from a TEX/LATEX file prepared by the author.