Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **435,** 3521–3540 (2013) doi:10.1093/mnras/stt1543
Advance Access publication 2013 September 12

Expectations for extreme-mass-ratio bursts from the Galactic Centre

C. P. L. Berry★ and J. R. Gair

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

Accepted 2013 August 14. Received 2013 July 19; in original form 2013 June 10

ABSTRACT

When a compact object on a highly eccentric orbit about a much more massive body passes through periapsis, it emits a short gravitational wave signal known as an extreme-mass-ratio burst (EMRB). We consider stellar mass objects orbiting the massive black hole (MBH) found in the Galactic Centre (GC). EMRBs provide a novel means of extracting information about the MBH; an EMRB from the Galactic MBH could be highly informative regarding the MBH's mass and spin if the orbital periapsis is small enough. However, to be a useful astronomical tool, EMRBs must be both informative and sufficiently common to be detectable with a space-based interferometer. We construct a simple model to predict the event rate for Galactic EMRBs. We estimate that there could be on average ∼2 bursts in a two-year mission lifetime for the Laser Interferometer Space Antenna. Stellar mass black holes dominate the event rate. Creating a sample of 100 mission realizations, we calculate what we could learn about the MBH. On average, we expect to be able to determine the MBH mass to ∼1 per cent and the spin to ∼0.1 using EMRBs.

Key words: black hole physics – gravitational waves – celestial mechanics – Galaxy: centre.

1 INTRODUCTION

The most compelling evidence for the existence of astrophysical black holes (BHs) comes from the measurement of stellar orbits at the centre of the Galaxy. The stars are found to orbit an object of mass $M_\bullet \simeq 4 \times 10^6\,\mathrm{M}_\odot$ coincident with the compact radio source Sagittarius A* (Reid & Brunthaler 2004; Ghez et al. 2008; Gillessen et al. 2009; Meyer et al. 2012). This is the nearest member of a population of massive black holes (MBHs; Volonteri 2010) that are believed to occupy the centres of galaxies (Lynden-Bell 1969; Rees 1984; Ferrarese & Ford 2005). The Galactic Centre (GC) is an ideal laboratory for investigating the properties of an MBH and its surrounding nuclear star cluster (Genzel, Eisenhauer & Gillessen 2010).

One means of investigating the properties of MBHs is through gravitational waves (GWs). A stellar mass compact object (CO), such as a main-sequence (MS) star, white dwarf (WD), neutron star (NS) or stellar mass BH, emits gravitational radiation as it orbits the MBH. A space-borne detector, such as the *Laser Interferometer Space Antenna* (*LISA*) or the *evolved Laser Interferometer Space Antenna*, is designed to be able to detect GWs in the frequency range of interest for these encounters assuming an MBH mass of $\sim 10^4$–$10^7\,\mathrm{M}_\odot$ (Bender et al. 1998; Danzmann & Rüdiger 2003; Jennrich et al. 2011; Amaro-Seoane et al. 2012). There are currently no funded space-borne detector missions. However, the European Space Agency's *LISA Pathfinder* will be launched in 2015 and demonstrate the key technologies required for a successful space-borne mission (Anza et al. 2005; Antonucci et al. 2012). Following on from previous work, we use the classic *LISA* design; this should provide a sensible benchmark for any future detectors.

The gravitational waveforms emitted from extreme-mass-ratio systems have been much studied (Glampedakis 2005; Barack 2009). On account of the extreme mass ratio between the two bodies, we can approximate the CO as moving in the background space–time of the MBH. The GWs carry away energy and angular momentum, causing the orbit to shrink until eventually the object plunges into the MBH. The primary focus has been upon the later stages of the orbital evolution, the last 1–2 yr immediately preceding the plunge. By this point, the orbit has nearly circularized and emits continuously within the detector's frequency band. These signals are extreme-mass-ratio inspirals (EMRIs; Amaro-Seoane et al. 2007). EMRIs can be observed over many orbits, allowing exquisitely high signal-to-noise ratios (SNRs) to accumulate. This makes them excellent probes of the background geometry permitting precise measurements of the system parameters and tests of general relativity.

EMRIs evolve from more eccentric orbits. These initial orbits may be the results of scattering from two-body encounters. Rather than emitting a continuously detectable signal, highly eccentric orbits only emit significant radiation in a burst around the point of closest approach to the MBH. These are extreme-mass-ratio bursts (EMRBs; Rubbo, Holley-Bockelmann & Finn 2006).

EMRBs are much shorter in duration than EMRIs. This means they do not accumulate as high SNRs or produce as detailed maps of the space–time. They are therefore less valued prizes. However, they may still be an interesting signal. As an object inspirals, it

★ E-mail: cplb2@cam.ac.uk

emits many bursts before eventually settling into a low-eccentricity EMRI. Some objects shall be scattered by two-body encounters and never reach the EMRI phase (Alexander & Hopman 2003). Thus, there are many potential EMRBs per EMRI, although this does not necessarily translate to there being more detectable EMRBs than EMRIs.

For EMRBs to be a useful astronomical signal, we require three things: the bursts contain sufficient information to improve our knowledge of their source systems, their event rate is sufficiently high that we expect to observe them over a mission lifetime and the signals can be successfully extracted from the data stream.

We have previously addressed the first requirement: EMRBs can give good constraints on the key parameters describing the Galaxy's MBH if the periapse distance is $r_\mathrm{p} \lesssim 10 r_\mathrm{g}$, where $r_\mathrm{g} = GM_\bullet/c^2$ is a gravitational radius (Berry & Gair 2013a). This would allow us to improve upon the current uncertainty in the mass measurement of 8 per cent (Gillessen et al. 2009). In addition, we could also measure the spin magnitude to a precision of better than 0.1.

The second requirement shall be the subject of this work. Previously, the best estimate for the event rate was given by Hopman, Freitag & Larson (2007); they predicted that the event rate for *LISA* is $\sim 1\,\mathrm{yr}^{-1}$. We follow a similar approach, but, significantly, we improve the calculation of SNR by using numerical kludge (NK) waveforms (Babak et al. 2007). Our models differ in a number of small ways; most notably, we allow for bursts to come from objects in the earliest stages of GW inspiral. In addition to this, we extend the analysis by not only considering the number of events that would be detectable, but also how many would be informative.

We begin in Section 2 by recapping how to generate and analyse burst waveforms. In Section 3, we present our model for calculating event rates and discuss its ingredients. Whilst we have been careful in trying to account for all the relevant effects, the model remains approximate and can only be trusted to give order-of-magnitude estimates. Our results, including the predicted event rate along with a quantification of the estimated information content, are presented in Section 4 and discussed further in our conclusion, Section 5.

2 WAVEFORMS AND PARAMETER UNCERTAINTIES

To establish the detectability and usefulness of EMRBs, it is necessary to calculate model waveforms. This is done using the NK approximation. We use exactly the same construction as is described in Berry & Gair (2013a); the key details are outlined below.

The detectability of a burst is dependent upon its SNR. To save calculating SNRs directly, it is possible to estimate the value from the periapse radius using a simple scaling relation. This is introduced in Section 2.2. We make use of this when selecting orbits of potential interest, but calculate the SNR from waveforms for more accurate results.

Once we have determined which bursts are of interest, we can evaluate the accuracy to which parameters can be determined, should that burst be observed. To do so, we perform Markov chain Monte Carlo (MCMC) simulations (MacKay 2003, chapter 29).

2.1 NK waveforms

Extreme-mass-ratio signals can be simulated in a computationally efficient manner using a semirelativistic approximation (Ruffini & Sasaki 1981): we assume that the CO moves along a geodesic in the Kerr geometry, but radiates as if it were in flat space–time. This technique is known as an NK. The justification of this technique is through comparison with more accurate, and computationally intensive, methods (Gair, Kennefick & Larson 2005; Babak et al. 2007). Using a geodesic for the trajectory should ensure that the signal has the correct frequency components but neglecting the effects of background curvature means that these need not have the correct amplitudes. The typical errors in the amplitude profile of a waveform can be a few per cent (Tanaka et al. 1993; Gair et al. 2005): the total amplitude error in our waveforms, integrating over all frequencies, has been estimated as typically $\sim$5 per cent, increasing to $\sim$10 per cent for the most relativistic orbits ($r_\mathrm{p} \lesssim 4r_\mathrm{g}$; Berry & Gair 2013a). We only consider parabolic (marginally bound) orbits, where the orbiting body would be at rest at infinity. When calculating event rates, we consider a range of highly eccentric orbits, not just those that are exactly parabolic. This introduces an additional error into the waveforms, but simplifies the analysis. The modification to the burst waveform should be small, as the trajectories are close to being parabolic; the largest differences are at large radii, which are unimportant for the waveforms.

To construct our NK waveforms, we first integrate the Kerr geodesic equations of motion. To avoid difficulties with turning points in the trajectory, we employ angular variables in place of the radial and polar coordinates (Drasco & Hughes 2004)

$$r = \frac{2r_\mathrm{p}}{1+\cos\psi}; \tag{1}$$

$$\cos^2\theta = \frac{Q}{Q+L_z^2}\cos^2\chi, \tag{2}$$

where Q is the Carter constant and L_z is the angular momentum about the z-axis.

Once the Kerr geodesic is constructed, we identify the Boyer–Lindquist coordinates with flat-space spherical polars (Gair et al. 2005; Babak et al. 2007). Using the relativistic trajectory ensures that the waveform incorporates the correct frequency components; translating to flat space means we can make use of the flat-space wave-generation formula. The downside of this is that the waveform amplitude shall not be precisely correct.

We use the quadrupole–octupole formula for the gravitational strain (Bekenstein 1973; Press 1977; Yunes et al. 2008). This is the familiar quadrupole formula (Misner, Thorne & Wheeler 1973, section 36.10; Hobson, Efstathiou & Lasenby 2006, section 17.9), plus the next order terms. The higher order terms modify the amplitudes of some frequency components by a few tens of per cent for the more relativistic orbits, although the integrated effect is smaller.

2.2 SNR scaling

The SNR ρ of a particular burst depends upon the precise shape of its trajectory (as specified by initial conditions) and the position of the detector. However, the most important parameter is the periapse distance.

The ρ–r_p relation is largely determined by the shape of the noise curve. For our simulations, we employ the noise model of Barack & Cutler (2004). For bursts from the GC, over much of the range of interest, the curve can be approximated as a simple power law (Berry & Gair 2013a)

$$\log\rho \simeq -2.7\log\left(\frac{r_\mathrm{p}}{r_\mathrm{g}}\right) + \log\left(\frac{\mu}{\mathrm{M}_\odot}\right) + 4.9, \tag{3}$$

where μ is the mass of the CO.

We assume a detection threshold of $\rho = 10$. This gives expected detection limits on the periapse radius. With a $1\,\mathrm{M}_\odot$ CO, bursts

should be detectable for $r_p \lesssim 27r_g$ and with a $10\,M_\odot$ CO for $r_p < 65r_g$.

2.3 Parameter estimation

Once we have a detected signal $s(t)$, we can consider the inference of the source parameters $\boldsymbol{\lambda}$. The waveform depends on the properties of the MBH, the properties of the CO, the orbit of the CO and the detector.

We assume that the position of the detector is known, and the MBH is coincident with the radio source of Sgr A* which is thought to be within $20r_g$ of the MBH (Reid et al. 2003; Doeleman et al. 2008). We use the J2000.0 coordinates which are determined to high accuracy (Reid et al. 1999; Yusef-Zadeh, Choate & Cotton 1999).

The parameters left to infer are as follows (Berry & Gair 2013a).

(1) The MBH's mass $M_\bullet$. This is well constrained by the observation of stellar orbits about Sgr A* (Ghez et al. 2008; Gillessen et al. 2009); we employ the estimate $M_\bullet = (4.31 \pm 0.36) \times 10^6\,M_\odot$ whenever a mass is required.

(2) The spin parameter a_*. This is constrained to the range $|a_*| < 1$.

(3,4) The orientation angles for the BH spin Θ_K and Φ_K. These are measured relative to the polar axis of the ecliptic coordinate system commonly used for describing the positions of celestial objects.

(5) The ratio of the GC distance and the CO mass $\zeta = R_0/\mu$. This scales the amplitude of the waveform. The distance has been measured using stellar orbits to be $R_0 = 8.33 \pm 0.35$ kpc (Gillessen et al. 2009).

(6,7) The angular momentum of the CO, parametrized by the magnitude at infinity $L_\infty = \sqrt{L_z^2 + Q}$ and the orbital inclination $\iota = \tan^{-1}(\sqrt{Q}/L_z)$.

(8–10) Coordinates specifying the trajectory. We use the angular phases at periapse, ϕ_p and χ_p, as well as the time of periapse t_p.

The first four are specific to the MBH, and we shall attempt to quantify the constraints we can expect to place on these from EMRBs.

The probability that the source is described by parameters $\boldsymbol{\lambda}$ is given by the posterior distribution

$$p(\boldsymbol{\lambda}|\boldsymbol{s}(t)) = \frac{p(\boldsymbol{s}(t)|\boldsymbol{\lambda})p(\boldsymbol{\lambda})}{p(\boldsymbol{s}(t))}. \tag{4}$$

Here $p(\boldsymbol{s}(t)|\boldsymbol{\lambda})$ is the likelihood of the parameters, $p(\boldsymbol{\lambda})$ is the prior probability distribution for the parameters and the evidence $p(\boldsymbol{s}(t))$ is a normalization factor.

If the parameter set $\boldsymbol{\lambda}_0$ defines a waveform $\boldsymbol{h}_0(t) = \boldsymbol{h}(t;\boldsymbol{\lambda}_0)$, the likelihood of the parameters is

$$p(\boldsymbol{s}(t)|\boldsymbol{\lambda}_0) \propto \exp\left[-\frac{1}{2}(\boldsymbol{s}-\boldsymbol{h}_0|\boldsymbol{s}-\boldsymbol{h}_0)\right]. \tag{5}$$

Here $(\boldsymbol{s}-\boldsymbol{h}_0|\boldsymbol{s}-\boldsymbol{h}_0)$ is the overlap between waveforms defined by the standard signal inner product (Cutler & Flanagan 1994). This is the probability of the realization of a noise signal $\boldsymbol{n}(t) = \boldsymbol{s}(t) - \boldsymbol{h}_0(t)$, assuming stationary, Gaussian noise.

We use uninformative priors on all the parameters: no existing constraints are used. We adopt uniform priors for all the parameters except $M_\bullet$, ζ and L_∞, which are positive definite, where we use a prior uniform in the logarithm of the parameter, and Θ_K and ι, where we use a prior uniform in the cosine of the angle.

To assess the accuracy to which parameters can be determined, we must find the width of the posterior distribution. MCMC methods are widely used for inference problems; they are a class of algorithms used for integrating over complicated probability distributions. The parameter space is explored by constructing randomly a chain of samples, with an acceptance rate dependent upon their relative probabilities (Metropolis et al. 1953; Hastings 1970). The distribution of points visited by the chain maps out the underlying distribution.

We employ the same semi-adaptive algorithm as was previously used in Berry & Gair (2013a). In this scheme, there is an initial phase where the proposal distribution (used in the selection of new points) is adjusted to match the distribution of points previously accepted; this tailors the width of the proposal to the particular posterior under consideration (Haario, Saksman & Tamminen 1999). The proposal is then fixed for the main phase to ensure ergodicity (Roberts & Rosenthal 2007; Andrieu & Thoms 2008) and only accepted points from this phase are used to characterize the posterior. This allows us to efficiently recover posteriors for a wide range of bursts.

3 CALCULATING EVENT RATES

Having determined how to generate a waveform and extract the information from it, we must now consider how likely it is that such a waveform would be observed. We wish to calculate the event rate for EMRBs, the probability that there is an encounter between a CO, on an orbit described by eccentricity e and periapse radius r_p, and the MBH. To do so we must build a model to describe the distribution of COs about the MBH. The number density of stars in the six-dimensional phase space of position and velocity is described by the distribution function (DF) f (Binney & Tremaine 2008, section 4.1). We introduce approximate forms for the DF appropriate for describing the Galactic core in Section 3.1. These are calibrated using the simulations of Alexander & Hopman (2009), the parameters of which, together with the others used to describe the Galactic core, are given in Section 3.2. Having set the distribution of COs, we explain how to convert this to an event rate in Section 3.3. In (28) we give an expression that relates the event rate for an orbit $\Gamma(e, r_p)$ and the DF. There is one final consideration before we can calculate the total event rate: that there is an inner periapsis below which orbits become depopulated. This is carefully explained in Section 3.4. We consider tidal disruption and collisions, which we assume truncate the DF at a finite periapsis so that the event rate inside these cut-offs is zero. We also consider GW inspiral, which we assume alters the event rate by modifying the distribution of COs away from its relaxed state. With these inner cut-offs established, we have completely defined the event rate distribution. This can then give the probability of an EMRB and the total event rates, which are presented in Section 4.1.

3.1 The distribution function

Following the work of Bahcall & Wolf (1976, 1977), we assume that the DF within the Galactic core is only a function of the orbital energy (Shapiro & Marchant 1978). The energy per unit mass of the orbit is

$$\mathcal{E} = \frac{v^2}{2} - \frac{GM_\bullet}{r}, \tag{6}$$

where v is the orbital velocity. The number of stars is

$$N = \int \mathrm{d}^3 r \int \mathrm{d}^3 v f(\mathcal{E}). \tag{7}$$

Close to the centre of the Galactic core, the dynamics are dominated by the influence of the MBH as it is significantly more massive than the surrounding stars. Its radius of influence is

$$r_{\mathrm{c}} = \frac{GM_\bullet}{\sigma^2}, \tag{8}$$

where σ^2 is the line-of-sight velocity dispersion (Frank & Rees 1976). We assume that the mass of stars enclosed within r_{c} is greater than $M_\bullet$, which, in turn, is much greater than the mass of a typical star $M_\star$ (Bahcall & Wolf 1976). We define a reference number density $n_\star$ from the enclosed mass $m_\star(r)$ such that

$$m_\star(r_{\mathrm{c}}) = \frac{4\pi r_{\mathrm{c}}^3}{3} n_\star M_\star. \tag{9}$$

Within the core, the DF can be calculated using the approximation of the Fokker–Planck formalism (Binney & Tremaine 2008, section 7.4). The population of bound stars is evolved numerically until a steady state is reached, whilst the unbound stars form a reservoir with an assumed Maxwellian distribution. Denoting a species of star by its mass M, the unbound DF is

$$f_M(\mathcal{E}) = \frac{C_M n_\star}{(2\pi\sigma_M^2)^{3/2}} \exp\left(-\frac{\mathcal{E}}{\sigma_M^2}\right), \quad \mathcal{E} > 0, \tag{10}$$

where C_M is a normalization constant.[1] If different stellar species are in equipartition, as assumed by Bahcall & Wolf (1976, 1977), we expect

$$M\sigma_M^2 = M_\star \sigma_\star^2. \tag{11}$$

However, if the unbound stellar population has reached equilibrium by violent relaxation, all mass groups are expected to have similar dispersions:

$$\sigma_M = \sigma_\star = \sigma, \tag{12}$$

and we have equipartition of energy per unit mass (Lynden-Bell 1967). This is assumed here following Alexander & Hopman (2009) and O'Leary, Kocsis & Loeb (2009). The steady-state DF is largely insensitive to this choice (Bahcall & Wolf 1977; Alexander & Hopman 2009).

For bound orbits, the DF can be approximated as a power law (Peebles 1972)

$$f_M(\mathcal{E}) = \frac{k_M n_\star}{(2\pi\sigma^2)^{3/2}} \left(-\frac{\mathcal{E}}{\sigma^2}\right)^{p_M}, \quad \mathcal{E} < 0. \tag{13}$$

The exponent p_M varies depending upon the mass of the object, determining mass segregation. For a system with a single mass component, $p = 1/4$ (Bahcall & Wolf 1976; Young 1977). The normalization constant k_M reflects the relative abundances of the different species.[2]

These cusp profiles should exist if the system has had sufficient time to become gravitationally relaxed. There is current debate about whether this may be the case, both for the GC and galaxies in general. This is discussed further in Appendix A6. For concreteness, we assume that a cusp has formed. If a cusp has not formed, we expect there to be a shallower core profile, with fewer objects passing close to the MBH. Our results are therefore an upper bound on possible event rates (Merritt 2010; Antonini & Merritt 2012; Gualandris & Merritt 2012).

[1] C_M determines the population ratios of species M far from the MBH (Alexander & Hopman 2009).

[2] For a single mass population ($p = 1/4$), $k = 2C$ gives a fit correct to within a factor of two (Bahcall & Wolf 1976; Keshet, Hopman & Alexander 2009); we assume that this holds for the dominant species of stars as, although it changes slightly with p, variation is small compared to errors introduced by fitting a simple power law (Hopman & Alexander 2006b; Alexander & Hopman 2009).

Table 1. Stellar model parameters for the Galactic core using the results of Alexander & Hopman (2009). The MS star is used as a reference for the normalization constants. The number fractions for unbound stars are estimates corresponding to a model of continuous star formation (Alexander 2005); O'Leary et al. (2009) arrive at the same proportions.

Star	$M/\mathrm{M}_\odot$	$C_M/C_\star$	p_M	$k_M/k_\star$ [a]
MS	1.0	1	−0.1	1
WD	0.6	0.1	−0.1	0.09
NS	1.4	0.01	0.0	0.01
BH	10	0.001	0.5	0.008

[a] Toonen, Hopman & Freitag (2009).

3.2 Model parameters

We use the Fokker–Planck model of Hopman & Alexander (2006a,b) and Alexander & Hopman (2009). This includes four stellar species: MS stars, WDs, NSs and stellar mass BHs. Their properties are summarized in Table 1. The behaviour of the Fokker–Planck model has been verified by N-body simulations (Baumgardt, Makino & Ebisuzaki 2004; Preto & Amaro-Seoane 2010). The steeper power law for BHs means they segregate about the MBH.[3]

Binaries may form in the Galactic core, encouraged by its high stellar density (O'Leary et al. 2009). However, the binary fraction is still expected to be small (Hopman 2009). Binaries are also disrupted by the MBH for periapses smaller than

$$r_{\mathrm{B}} \simeq \left(\frac{M_\bullet}{M_1 + M_2}\right)^{1/3} a_{\mathrm{B}}, \tag{14}$$

where M_1 and M_2 are the masses of the binary's components, and a_{B} is the binary's semimajor axis, cf. (29) below. Thus, we ignore the possible presence of binaries.

We assume that $M_\bullet = (4.31 \pm 0.36) \times 10^6\,\mathrm{M}_\odot$ (Gillessen et al. 2009) and $\sigma = (103 \pm 20)\,\mathrm{km\,s^{-1}}$ (Tremaine et al. 2002). This gives a core radius of $r_{\mathrm{c}} = (1.7 \pm 0.7)\,\mathrm{pc}$. Using the results of Ghez et al. (2008), we would expect the total mass of stars in the core to be $m_\star(r_{\mathrm{c}}) = 6.4 \times 10^6\,\mathrm{M}_\odot$, which is within 5 per cent of the value obtained similarly from Genzel et al. (2003). This gives a reference stellar density of $n_\star = 2.8 \times 10^5\,\mathrm{pc}^{-3}$.

3.3 The event rate in terms of eccentricity and periapsis

We characterize orbits by their eccentricity e and periapse radius r_{p}. The latter, unlike the semimajor axis, is always well defined regardless of eccentricity. For Keplerian orbits, the energy $\mathcal{E}$ and angular momentum $\mathcal{J}$ per unit mass are entirely characterized by these parameters

$$\mathcal{E} = -\frac{GM_\bullet(1-e)}{2r_{\mathrm{p}}}; \tag{15a}$$

$$\mathcal{J}^2 = GM_\bullet(1+e)r_{\mathrm{p}}. \tag{15b}$$

The DF is defined per element of phase space: it is necessary to change variables from position and velocity to eccentricity and

[3] Extrapolating, they would dominate in place of MS stars for radii $r < 10^{-4} r_{\mathrm{c}}$.

periapsis. We decompose the velocity into three orthogonal components: radial v_r, azimuthal v_ϕ and polar v_θ. We assume that the core is spherically symmetric (Genzel et al. 2003; Schödel et al. 2007); therefore, we are interested in the combination

$$v_\perp^2 = v_\phi^2 + v_\theta^2 = v^2 - v_r^2. \tag{16}$$

Under this change of variables

$$\mathrm{d}^3 v = \mathrm{d}v_r\, \mathrm{d}v_\phi\, \mathrm{d}v_\theta \rightarrow 2\pi v_\perp\, \mathrm{d}v_r\, \mathrm{d}v_\perp. \tag{17}$$

The specific energy and angular momentum are given by

$$\mathcal{E} = \frac{v_r^2 + v_\perp^2}{2} - \frac{GM_\bullet}{r}; \tag{18a}$$

$$\mathcal{J}^2 = r^2 v_\perp^2. \tag{18b}$$

Combining these with our earlier expressions in terms of e and r_p,

$$v_\perp^2 = \frac{GM_\bullet(1+e)r_\mathrm{p}}{r^2}, \tag{19}$$

$$v_r^2 = GM_\bullet\left[\frac{2}{r} - \frac{(1-e)}{r_\mathrm{p}} - \frac{(1+e)r_\mathrm{p}}{r^2}\right]. \tag{20}$$

From the latter, we can verify that the turning points of an orbit occur at

$$r = r_\mathrm{p},\ \frac{1+e}{1-e}r_\mathrm{p}; \tag{21}$$

the periapse is the only turning point for orbits with $e > 1$. Since we now have expressions for $\{v_r, v_\perp\}$ in terms of $\{e, r_\mathrm{p}\}$, we can rewrite our velocity element as

$$\mathrm{d}^3 v \rightarrow \frac{\pi e}{v_r r_\mathrm{p}}\left(\frac{GM_\bullet}{r}\right)^2 \mathrm{d}e\, \mathrm{d}r_\mathrm{p}. \tag{22}$$

As a consequence of our assumed spherical symmetry, the phase space volume element can be expressed as

$$\mathrm{d}^3 r\, \mathrm{d}^3 v \rightarrow \frac{4\pi^2 (GM_\bullet)^2 e}{v_r r_\mathrm{p}}\, \mathrm{d}r\, \mathrm{d}e\, \mathrm{d}r_\mathrm{p}. \tag{23}$$

The number of stars in an element $\mathrm{d}r\, \mathrm{d}e\, \mathrm{d}r_\mathrm{p}$ is

$$n(r, e, r_\mathrm{p}) = \frac{4\pi^2 (GM_\bullet)^2 e}{v_r r_\mathrm{p}} f(\mathcal{E}). \tag{24}$$

From this, we can construct the expected number of stars on orbits defined by $\{e, r_\mathrm{p}\}$. The number of stars found in a small radius range δr with given orbital properties is

$$n(r, e, r_\mathrm{p})\delta r = N(e, r_\mathrm{p}; r)\frac{\delta t}{P(e, r_\mathrm{p})}, \tag{25}$$

where $N(e, r_\mathrm{p}; r)$ is the total number of stars with orbits given by $\{e, r_\mathrm{p}\}$ defined at r, δt is the time spent in δr and $P(e, r_\mathrm{p})$ is the period of the orbit. We defer the definition of this time for unbound orbits for now. The time spent in the radius range is

$$\delta t = 2\frac{\delta r}{v_r}, \tag{26}$$

where the factor of 2 accounts for inward and outward motion. Hence,

$$\begin{aligned} N(e, r_\mathrm{p}; r) &= \frac{1}{2} v_r P(e, r_\mathrm{p}) n(r, e, r_\mathrm{p}) \\ &= \frac{2\pi^2 (GM_\bullet)^2 e P(e, r_\mathrm{p})}{r_\mathrm{p}} f(\mathcal{E}). \end{aligned} \tag{27}$$

The right-hand side is independent of position, subject to the constraint that the radius is in the allowed range for the orbit $r_\mathrm{p} \le r \le (1+e)r_\mathrm{p}/(1-e)$, and so $N(e, r_\mathrm{p}) \equiv N(e, r_\mathrm{p}; r)$. This is a consequence of the DF being dependent only upon a constant of the motion.[4]

If a burst of radiation is emitted each time a star passes through periapse, the event rate for burst emission from orbits with parameters $\{e, r_\mathrm{p}\}$ is given by

$$\Gamma(e, r_\mathrm{p}) = \frac{N(e, r_\mathrm{p})}{P(e, r_\mathrm{p})} = \frac{2\pi^2 (GM_\bullet)^2 e}{r_\mathrm{p}} f(\mathcal{E}). \tag{28}$$

The orbital period drops out from the calculation, so we do not have to worry about an appropriate definition for unbound orbits.

To generate a representative sample for the orbital parameters e and r_p, we use $\Gamma(e, r_\mathrm{p})\, \mathrm{d}e\, \mathrm{d}r_\mathrm{p}$ as the rate for a Poisson distribution.

The total event rate can be found by integrating $\Gamma(e, r_\mathrm{p})$, but before we can do this we must set the limits of the integral. The maximum periapse is set by the limit of detectability. It is particular to the detector. The inner periapse is set by physical processes. These are explained in the following section.

3.4 The inner cut-off

From (28) we see that the event rate is highly sensitive to the smallest value of the periapsis. Ultimately, the orbits cannot encroach closer to the MBH than its last stable orbit. This depends upon the spin of the MBH, but is of the order of its Schwarzschild radius. Before we reach this point, there are other processes that may intervene to deplete the orbiting stars. Our treatment of these is approximate, but should produce reasonable estimates. We consider three processes: tidal disruption by the MBH (Section 3.4.1), GW inspiral (Section 3.4.3) and collisional disruption (Section 3.4.4). Tidal disruption imposes a definite (albeit approximate) cut-off, while the others use statistical arguments. For these methods, we need to define a reference time-scale for relaxation. This is done in Section 3.4.2, with further details found in Appendix A.

The calculated inner cut-offs for the four stellar species across the range of bound orbits are shown in Fig. 1. The tidal and collisional disruption cut-offs are hard boundaries, inside of which we assume that there are no bursting sources. The transition to the GW inspiral-dominated regime marks the end of the relaxed distribution of stars; inside of this there are only inspiralling stars.

3.4.1 Tidal disruption

Tidal forces from the MBH can disrupt stars. This occurs at the tidal radius

$$r_\mathrm{T} \simeq \left(\frac{M_\bullet}{M}\right)^{1/3} R_M, \tag{29}$$

where R_M is the radius of the star (Hills 1975; Rees 1988; Kobayashi et al. 2004).[5] Any star on an orbit with $r_\mathrm{p} < r_\mathrm{T}$ is disrupted in the course of its orbit. Parametrizing orbits by their periapsis allows us to easily determine which stars should be disrupted. We do not include the full effects of the loss cone (Frank & Rees 1976; Lightman & Shapiro 1977; Cohn & Kulsrud 1978) as these were not incorporated into the Fokker–Planck calculations (Hopman 2009).[6]

[4] See Bahcall & Wolf (1976, equation 9) for a similar result.

[5] See Kesden (2012) for a general relativistic treatment.

[6] The loss cone is a region in velocity space where orbits are depleted because stars are disrupted more rapidly than they can be replenished by two-body scattering.

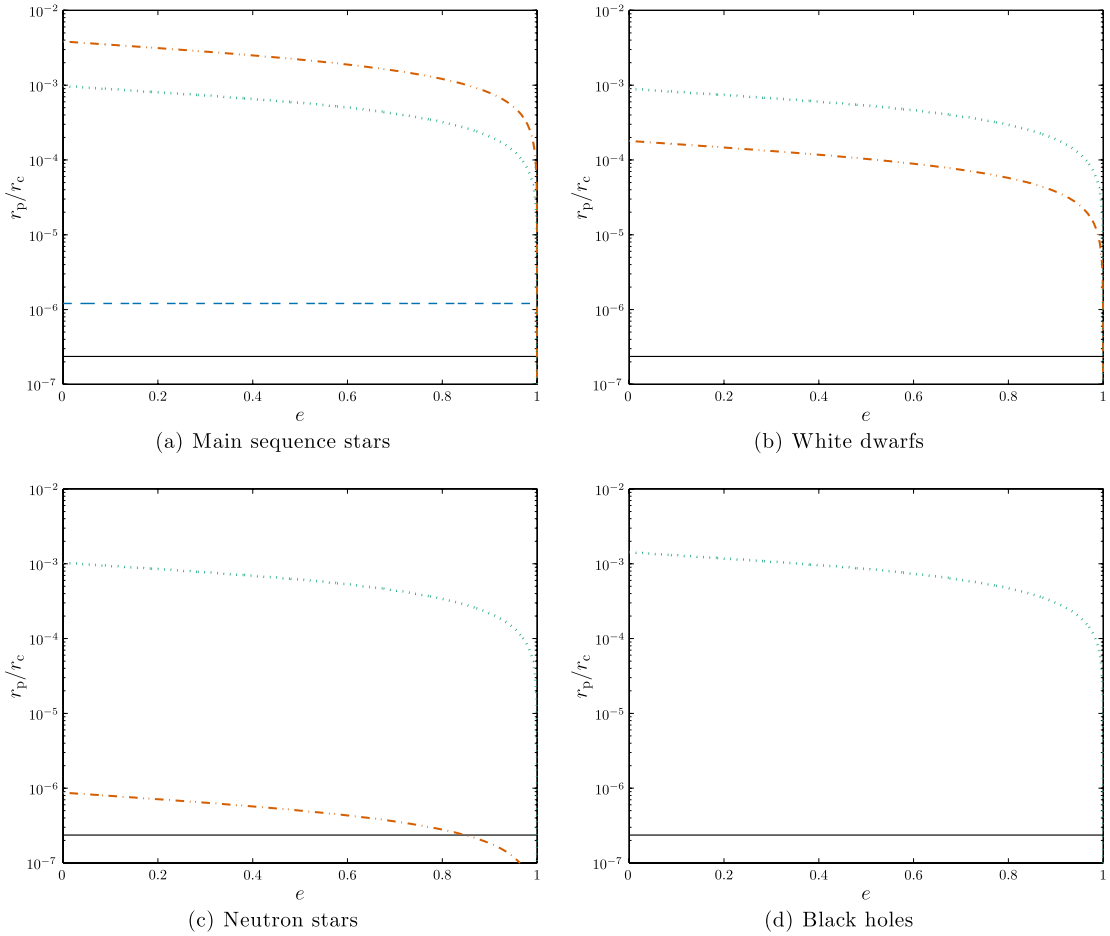


Figure 1. Inner cut-off radii for the GC as a function of eccentricity. The solid line shows the Schwarzschild radius of the MBH; this gives an indication of the innermost possible orbit which actually varies with MBH spin as well as orbital eccentricity and inclination. The dashed line shows the tidal radius which is a hard cut-off inside of which there should be no undisrupted stars. The dot–dashed line shows the collisional cut-off which is a statistical cut-off inside of which we do not expect any stars. The dotted line shows the transition to the GW-dominated inspiral regime; inside of this we expect inspiralling stars in place of the relaxed distribution.

The effect of the loss cone should be small, only modifying the DF by a logarithmic term (Bahcall & Wolf 1977; Lightman & Shapiro 1977; Cohn & Kulsrud 1978). Its effects are diluted by resonant relaxation (Hopman et al. 2007; Toonen et al. 2009; Merritt et al. 2011). Furthermore, the loss cone could be refilled by the wandering within the core of the MBH because of perturbations from the inhomogeneities in the stellar potential (Sigurdsson & Rees 1997; Chatterjee, Hernquist & Loeb 2002; Merritt, Berczik & Laun 2007).

Tidal disruption is significant for MS stars since they are least dense: calculated in this way, only MS stars are tidally disrupted outside of the MBH's event horizon (Sigurdsson & Rees 1997). The tidal radius defines the cut-off for periapsis of high-eccentricity ($e \gtrsim 1$) orbits (Lightman & Shapiro 1977).

3.4.2 Relaxation time-scale

The motion of a star is determined not only by the dominant influence of the central MBH, but also by the other stars. The gravitational potential of the stars may be split into two components: a smooth background representing the average distribution of stars and statistical fluctuations from random deviations in the stellar distribution because of individual stellar motions. The former only contributes to the stars' orbits: we neglect this since we are more interested in the influence of the MBH. The latter may be approximated as a series of two-body encounters. These lead to scattering, in a manner much like Brownian motion (Bekenstein & Maoz 1992; Maoz 1993; Nelson & Tremaine 1999).

The two-body interactions mostly lead to small deflections. Over time, these may accumulate into a significant change in the dynamics. The relaxation time-scale characterizes the time taken for this to happen (Binney & Tremaine 2008, section 1.2.1). It therefore quantifies the time over which an orbit may be repopulated by scattering. There are a variety of definitions for the relaxation time-scale. For a system with a purely Maxwellian distribution, the time-scale has the form

$$\tau_\mathrm{R}^\mathrm{Max} \simeq \kappa \frac{\sigma^3}{G^2 M_\star^2 n_\star \ln \Lambda}, \tag{30}$$

where the Coulomb logarithm is $\ln \Lambda = \ln (M_\bullet/M_\star)$ (Bahcall & Wolf 1976) and κ is a dimensionless number. In his pioneering work, Chandrasekhar (1941a, 1960) defined the time-scale as the period over which the squared change in energy was equal to the kinetic energy squared; this gives $\kappa = 9/16\sqrt{\pi} \simeq 0.32$. Subsequently, Chandrasekhar (1941b) described relaxation statistically, treating fluctuations in the gravitational field probabilistically; this gives $\kappa = 9/2(2\pi)^{3/2} \simeq 0.29$. Bahcall & Wolf (1977)

define a reference time-scale from their Boltzmann equation with $\kappa = 3/4\sqrt{8\pi} \simeq 0.15$; this is equal to the reference time-scale defined as the reciprocal of the coefficient of dynamical friction by Chandrasekhar (1943a,b). Spitzer & Harm (1958) define a reference time-scale from the gravitational Boltzmann equation of Spitzer & Schwarzschild (1951), where $\kappa = \sqrt{2}/\pi \simeq 0.45$. Following Spitzer & Hart (1971), Binney & Tremaine (2008, section 7.4.5) estimate the time-scale from the velocity diffusion coefficient of the Fokker–Planck equation yielding $\kappa \simeq 0.34$.

All these approaches yield consistent values, suggesting, as a first approximation, that any is valid. We follow the classic treatment of Chandrasekhar (1960, chapter 2) which is transparent in its assumptions, adapting from a Maxwellian distribution of velocities to one derived from the DFs (10) and (13); this makes the model self-consistent. The derivation of the relaxation time-scale along with a discussion of its shortcomings is included in Appendix A. An average time-scale for the entire system $\overline{\tau_{\rm R}}$ is defined in (A32), and an average for an orbit $\langle \tau_{\rm R} \rangle$ is defined in (A38).

Since there is uncertainty in the astrophysical parameters, we will not be concerned by small discrepancies in the numerical prefactor that result from the simplifying approximations of this approach. We defer any investigation of the consequences of using an alternative formulation for future work, as differences may well be negligible whilst the computation is complicated.

Two-body interactions lead to diffusion in both energy and angular momentum. When considering a single (bound) orbit, over a relaxation time-scale the energy changes by order of itself while the angular momentum changes by the angular momentum of a circular orbit with that energy $\mathcal{J}_{\rm circ}(\mathcal{E})$ (Lightman & Shapiro 1977; Rauch & Tremaine 1996; Hopman & Alexander 2005; Madigan, Hopman & Levin 2011):[7]

$$\left(\frac{\Delta\mathcal{E}}{\mathcal{E}}\right)^2 \approx \left[\frac{\Delta\mathcal{J}}{\mathcal{J}_{\rm circ}(\mathcal{E})}\right]^2 \approx \frac{t}{\tau_{\rm R}}. \tag{31}$$

We may define another angular momentum relaxation time-scale as the time taken for the angular momentum to change by order of itself (Merritt et al. 2011)

$$\tau_{\mathcal{J}} = \left[\frac{\mathcal{J}}{\mathcal{J}_{\rm circ}(\mathcal{E})}\right]^2 \tau_{\rm R} = \left(1 - e^2\right) \tau_{\rm R}. \tag{32}$$

This can be much shorter than the energy relaxation time-scale: diffusion in angular momentum can proceed more rapidly than diffusion in energy.

3.4.3 GW inspiral

Stars orbiting the MBH continually emit gravitational radiation; this carries away energy and angular momentum, causing the stars to inspiral. Using the analysis of Peters & Mathews (1963) and Peters (1964) for Keplerian binaries, it is possible to define a characteristic inspiral time-scale from the rate of change of energy. For consistency with the relaxation time-scale, we define this as (Miralda-Escude & Gould 2000; Merritt et al. 2011)

$$\tau_{\rm GW} \simeq \mathcal{E} \left\langle \frac{{\rm d}\mathcal{E}}{{\rm d}t} \right\rangle^{-1}, \tag{33}$$

where the term in angular brackets is the orbit-averaged rate of energy radiation. Using (15) and equation 16 of Peters & Mathews (1963),

$$\tau_{\rm GW} \simeq \frac{5}{64} \frac{c^5 r_{\rm p}^4}{G^3 M M_\bullet (M + M_\bullet)} \frac{(1+e)^{7/2}}{(1-e)^{1/2}} \times \left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right)^{-1} \tag{34}$$

$$\approx \frac{5}{64} \frac{c^5 r_{\rm p}^4}{G^3 M M_\bullet^2} \frac{(1+e)^{7/2}}{(1-e)^{1/2}} \left(1 + \frac{73}{24}e^2 + \frac{37}{96}e^4\right)^{-1}. \tag{35}$$

For comparison, the total time taken for the inspiral, if undisturbed, is given in (B7).

The time-scale associated with changes in angular momentum is (Peters 1964)

$$\tau_{{\rm GW},\mathcal{J}} \simeq \mathcal{J} \left\langle \frac{{\rm d}\mathcal{J}}{{\rm d}t} \right\rangle^{-1} \tag{36}$$

$$\simeq \frac{5}{32} \frac{c^5 r_{\rm p}^4}{G^3 M M_\bullet (M + M_\bullet)} \frac{(1+e)^{5/2}}{(1-e)^{3/2}} \left(1 + \frac{7}{8}e^2\right)^{-1} \tag{37}$$

$$\approx \frac{5}{32} \frac{c^5 r_{\rm p}^4}{G^3 M M_\bullet^2} \frac{(1+e)^{5/2}}{(1-e)^{3/2}} \left(1 + \frac{7}{8}e^2\right)^{-1}. \tag{38}$$

This is always greater than the energy time-scale; hence, we only consider changes in energy from GW emission as important for evolution of the system (Hopman & Alexander 2005).

Unbound stars only undergo a single periapse passage and only radiate one burst of radiation; we therefore neglect any evolution in their orbital parameters.[8]

The $(1 - e)^{-1/2}$ dependence of $\tau_{\rm GW}$ for bound orbits connects the two regimes. The rate of change of energy goes to zero as a consequence of assuming that the orbital parameters do not change over the course of an orbit. It is a valid approximation since the large mass ratio ensures a slow evolution of the system (Appendix B2).

When comparing with the relaxation time-scale we are comparing rates of change, with the shorter time-scale highlighting the more rapid process that dominates the evolution (Amaro-Seoane et al. 2007). We therefore compare $\tau_{\rm GW}$ with the orbital relaxation time-scale $\tau_{\mathcal{J}}$ (Merritt et al. 2011). Orbits with $\tau_{\rm GW} < \tau_{\mathcal{J}}$ become depleted by GW emission faster than they are replenished by scattering. The cusp does not extend to these orbits. Yet, these orbits are not totally depopulated as an object may pass through during its inspiral from greater periapse and eccentricity. In their calculations, Hopman et al. (2007) did not include these inspiralling stars as potential burst sources. We calculate the density of stars in this region by following the evolution of inspirals beginning at the inner edge of the cusp (where the two time-scales are equal), weighting by the rate of change of the periapse and eccentricity in each element of e–$r_{\rm p}$ space (Peters 1964). The net effects are the high-eccentricity distributions of MS stars, WDs and NSs are relatively unchanged from their cusp states, but the BH distribution is significantly depleted.

3.4.4 Collisions

As a consequence of the high densities in the Galactic core, stars may undergo a large number of close encounters with other stars (Cohn

[7] $\mathcal{J}_{\rm circ}(\mathcal{E})$ is the maximum value for orbits of that energy.

[8] Changes are only important for very high eccentricity orbits (Appendix B2). These are of high energy and are exponentially suppressed because of the Boltzmann factor in (10).

& Kulsrud 1978). These may lead to their destruction. MS stars, WDs and NSs may be pulled apart by tidal forces if they stray too close to a more massive object. As MS stars are diffuse, they would not tidally disrupt another star (Murphy, Cohn & Durisen 1991; Freitag & Benz 2005). Close encounters would result in some mass transfer; the cumulative effect of 20–30 grazing collisions could destroy an MS star (Freitag, Amaro-Seoane & Kalogera 2006). The number of collisions a star undergoes in a time interval δt is

$$\delta K = n(r) A v(r, e, r_\mathrm{p}) \delta t, \tag{39}$$

where A is the collisional cross-sectional area. For tidal disruption, where the encounter is with a collapsed object (WD, NS or BH), we set $A = \pi r_{\mathrm{T},\,M'}^2$, where $r_{\mathrm{T},\,M'}$ is the appropriate tidal radius: like (29) but with $M_\bullet$ replaced with the mass of the collapsed object M'. For collisions between MS stars, the cross-sectional area is simply the geometric $A = \pi R_\star^2$.[9]

For circular orbits, we can find the radius at which collisions lead to disruptions by setting $\delta K = 1$ for tidal disruption or $\delta K = 20$ for grazing collisions, and $\delta t = \overline{\tau_{\mathrm{R},\,M}}$. We use the system average relaxation time-scale for the species of mass M as this is the time over which stars are replenished from the reservoir. For non-circular orbits, we must consider variation with position. Using $\delta r = v_r \delta t$, and then converting to an integral, for bound orbits

$$K = 2A \frac{\overline{\tau_{\mathrm{R},\,M}}}{P(r_\mathrm{p}, e)} \int_{r_\mathrm{p}}^{(1+e)r_\mathrm{p}/(1-e)} n(r) \frac{v(r, e, r_\mathrm{p})}{v_r(r, e, r_\mathrm{p})} \,\mathrm{d}r, \tag{40}$$

where P is the period of the orbit. Again we set $K = 1$ or $K = 20$, and then numerically solve (40) to find the orbits for which stars are disrupted within $\overline{\tau_{\mathrm{R},\,M}}$. For unbound orbits, we are only interested in stars that would become disrupted before their periapse passage, so

$$K = A \int_{r_\mathrm{p}}^{r_\mathrm{c}} n(r) \frac{v(r, e, r_\mathrm{p})}{v_r(r, e, r_\mathrm{p})} \,\mathrm{d}r, \tag{41}$$

assuming that the stars in the reservoir external to the core are unlikely to undergo close collisions.

Orbits within the collisional cut-off are assumed to be depopulated and do not contribute to the event rate. Our treatment is similar to that of Hopman et al. (2007), but they only considered collisions between MS stars. Collisions provide the cut-off for bound MS stars and are significant for bound WDs.

4 RESULTS

4.1 Number of events

As a first approximation for the number of events expected in a 2 yr mission lifetime, we numerically integrated the event rate. This estimate is denoted by $\mathcal{N}_2^{\mathrm{int}}$. The lower limit on r_p was set to be the largest of the tidal cut-off, the collisional cut-off or the MBH's Schwarzschild radius.[10] The Schwarzschild radius $r_\mathrm{S} = 2r_\mathrm{g}$ is used as a proxy for an averaged innermost orbit's periapse; the innermost parabolic orbit for non-spinning MBHs has $r_\mathrm{p} = 4r_\mathrm{g}$ and the innermost parabolic orbit for a maximally spinning MBH has $r_\mathrm{p} = r_\mathrm{g}$ for a prograde equatorial orbit and $r_\mathrm{p} = (3 + 2\sqrt{2})r_\mathrm{g}$ for a retrograde equatorial orbit. The upper limit was the detection threshold as determined from (3). The lower bound on eccentricity was set to 0.9, below which we do not trust the parabolic approximation for burst waveforms; since the DF decays exponentially with eccentricity for unbound orbits, the upper limit does not influence our results.

To obtain a more accurate estimate, we performed 2×10^4 mission realizations. For each mission, we randomly selected a set of parameters to describe the MBH, and then picked orbits with probabilities defined by their event rates. The orbital position of the *LISA* detector was also chosen randomly.[11] The SNRs of the resulting bursts were calculated and a detection was recorded if $\rho > 10$. By averaging the number of events per mission, we can estimate the expected number of bursts we would detect. This is denoted by $\mathcal{N}_2^{\mathrm{run}}$.

The calculated numbers of events are shown in Table 2. The two approaches are in good agreement indicating that the average relation (3) is sufficiently accurate for this type of calculation, and that the Schwarzschild radius is a reasonable absolute inner cut-off averaged over all MBH spins. The total number of events per mission is plotted in Fig. 2. This is consistent with being Poisson distributed as expected.

Only BHs and NSs contribute to the event rate significantly. Only MS stars have a non-negligible (relative) contribution from unbound

Table 2. Expected number of events per two-year mission. $\mathcal{N}_2^{\mathrm{int}}$ is an estimate using the average SNR–periapsis scaling and $\mathcal{N}_2^{\mathrm{run}}$ is calculated by averaging results from 2×10^4 mission realizations.

Star	$\mathcal{N}_2^{\mathrm{int}}$	$\mathcal{N}_2^{\mathrm{run}}$
MS	9.5×10^{-4}	1.3×10^{-3}
WD	1.0×10^{-2}	1.0×10^{-2}
NS	5.0×10^{-1}	5.0×10^{-1}
BH	1.2×10^{0}	1.2×10^{0}
Total	1.7×10^{0}	1.7×10^{0}

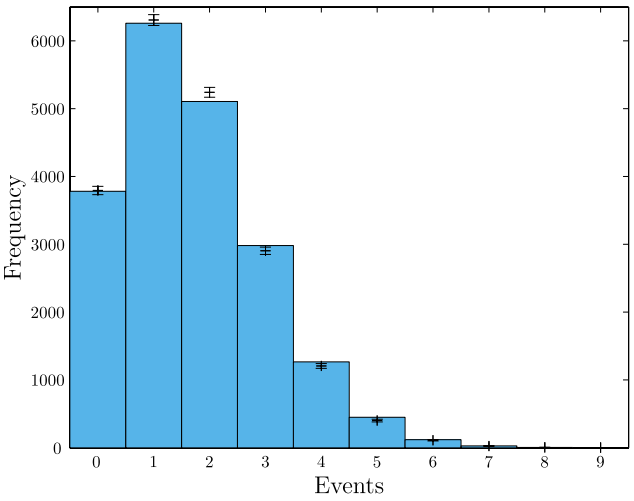


Figure 2. Calculated number of detectable EMRBs over a two-year mission. The histogram shows the number of events for 2×10^4 realizations. The points show a Poisson distribution with a mean set by $\mathcal{N}_2^{\mathrm{int}}$.

[9] Here we assume that the relative velocity of the colliding stars is much greater than the escape velocity of the star, so we may neglect the effects of gravitational focusing.

[10] The transition to the GW inspiral regime is not a cut-off, but reflects a change in the form of the stellar distribution; hence, it is not included amongst the other inner periapses as a lower limit.

[11] We use the same condition on the initial orientation as in Berry & Gair (2013a), following Cutler (1998). This does not qualitatively influence our results.

orbits. The event rates are not high, but there is an ∼4/5 (81 per cent) chance of observing at least one burst in a mission.

The overall rates are similar to those presented in Hopman et al. (2007). The MS rate is lower because of a larger collisional cut-off. This also influences the WD rate, but the overall rate is little changed. The NS rate is enhanced because of the inclusion of bursts from inspiralling objects. The physics for BHs is least changed; the (small) difference in the event rate is partly a consequence of our more realistic SNRs.

4.2 Information content

We wish to quantify what we could learn over a mission about the MBH's mass and spin. We use the parameter set $\boldsymbol{\lambda}_\bullet = \{\ln(M_\bullet/\mathrm{M}_\odot), a_*, \cos\Theta_\mathrm{K}, \Phi_\mathrm{K}\}$ as each of these has a uniform prior.

The information carried by a burst is encoded in its posterior probability distribution. This can be recovered using an MCMC as explained in Section 2.3. We ran MCMCs for bursts from the first 100 of our mission realizations that had periapses $r_\mathrm{p} < 16r_\mathrm{g}$. There were a total of 96 interesting bursts (57 from BHs and 39 from NSs) across 63 missions. For examples of recovered posterior distributions, see Berry & Gair (2013a,b).[12] Ideally, we would use information from all detectable bursts, but this would be computationally expensive and we do not expect to glean much useful information from orbits with larger periapses.

During an individual mission there may be either zero, one or multiple bursts of interest. In the first case, we learn nothing. In the second, we have to only consider the posterior from our MCMC. In the third, we must combine the posteriors of all the bursts. This is easy in theory: as the priors are uniform, we have to only multiply the individual posteriors,

$$p(\boldsymbol{\lambda}_\bullet|\{\boldsymbol{s}_i(t)\}_\mathrm{mission}) = \prod_i p(\boldsymbol{\lambda}_\bullet|\boldsymbol{s}_i(t)), \tag{42}$$

where $\{\boldsymbol{s}_i(t)\}_\mathrm{mission}$ is the set of bursts for the mission. However, since we have a sampled posterior rather than an analytic function, this is difficult in practice.

The simplest thing to do is bin the points and then multiply the numbers in each bin together (dividing by the area of the bin to convert back to a probability density). The question is then what is an appropriate bin size? Bins that are too big give insufficient resolution, whilst those that are too small may not encompass any sampled points.

One means of creating bins with sizes that reflect the structure of the distribution is using a k-d tree. This is a type of binary space partitioning tree (de Berg et al. 2008, sections 5.2, 12.1 and 12.3); it is constructed finding the median in one dimension and then splitting the parameter space into two at this location. This process is then repeated in another dimension, then another, and so on recursively until the desired number of partitions, known as leaves, have been created. Constructing a k-d tree from a sampled probability distribution creates smaller leaves in the regions of high probability, which are of most interest, and larger leaves in the areas of low probability (Weinberg 2012).

[12] The former shows posteriors for Galactic bursts where the CO is $10\,\mathrm{M}_\odot$. The latter shows posteriors for extragalactic bursts and a $10\,\mathrm{M}_\odot$ CO. The CO mass is degenerate with the distance to the source; hence, the extragalactic bursts give an indication of what happens when the CO mass is reduced (although changing from a $10\,\mathrm{M}_\odot$ BH to a $1.4\,\mathrm{M}_\odot$ NS is not as extreme as moving to another galaxy).

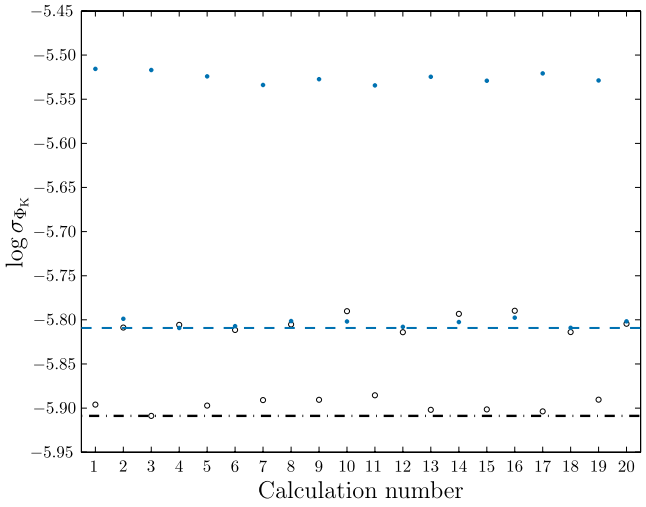


Figure 3. Example of calculated values for the standard deviation σ_SD (indicated by the solid points) and the 68-percentile half-width $\sigma_{0.68}$ (indicated by the open points) that show clear biasing. The odd numbered calculations use the k-d tree constructed from the first burst for binning and the even numbered calculations use the tree from the second. Both of these processes were repeated 10 times to check there were no numerical problems. The values used as final results for σ_SD and $\sigma_{0.68}$ are indicated by the dashed and dot–dashed lines, respectively. The posterior used here is for the orientation angle Φ_K but the effect may be seen in posteriors for the other parameters.

Taking each burst posterior in turn, we construct a k-d tree using the two-step method (Berry & Gair 2013b; Sidery, Gair & Mandel, in preparation): the data are (randomly) divided in two parts – the first half is used to create the leaves and the second half is used to populate the leaves. This reduces biasing due to fluctuations in the sampled data. We use this tree to bin the other posteriors and multiply the totals together. This gives us one estimate for the combined posterior for each of the input bursts. We resample the final distributions (sampling each leaf uniformly) to create sets of points that can be treated in the same way as the output from an MCMC.

The precision to which a parameter can be constrained may be quantified by the width of the distribution. We shall use the standard deviation σ_SD and the half-width of the 68-percentile range constructed from one-dimensional k-d trees $\sigma_{0.68}$. This is constructed by adding together the smallest leaves of the tree until they encompass a total probability of 0.68. The widths coincide for Gaussian posteriors.

The widths calculated from multiple burst posteriors may be biased to too large values. This can happen when combining distributions of significantly different widths. When using the k-d tree of a distribution that is much broader than the others, a small number of leaves can contain the majority of the final posterior probability and we cannot resolve the final width. When using the k-d tree from a narrow distribution, there can be large leaves at the edges of the parameter ranges; because the resampled points from these leaves are uniformly distributed, they can skew the overall distribution. Since the bias increases the width, we use the smallest of the calculated values.[13] Fig. 3 shows an example where there is clear biasing, it shows multiple calculations of the widths alternating between using the k-d from the first burst posterior and the second.

[13] The distributions were checked to ensure that they did not have anomalously small widths due to a computational error.

In many cases, the variation is comparable to the intrinsic scatter from random sampling. The 68-percentile half-width appears more robust against biasing.

Combining the results from the set of realizations, Fig. 4 shows the fraction of missions $\mathcal{F}(\sigma > \varsigma)$ that have posterior widths larger than ς. It appears that there is a 33 per cent chance of determining $\ln(M_\bullet/\mathrm{M}_\odot)$ to a precision of 10^{-2} or a 10 per cent chance of determining it to 10^{-4}. The current uncertainty of $\sim$0.08 is bettered in over half (52 per cent) of the missions. The spin a_* could be determined to a precision of 10^{-2} in 30 per cent of missions and there is a 10 per cent chance of determining it to better than 3×10^{-4}.

The distribution widths work for describing parameter estimation accuracies of an individual mission; however, they are less useful for calculating an average since they are undefined when no bursts are detected. There is an alternative means of characterizing what we could learn: the information entropy of the posterior distribution.

Shannon (1948a,b) introduced the idea of information entropy, which quantifies the expectation for information gained from an outcome or, equivalently, the amount of uncertainty regarding a system (MacKay 2003, chapters 2 and 4). For a discrete ensemble of probabilities $\{p_i\}$,

$$H(\{p_i\}) = -\sum_i p_i \ln p_i \tag{43}$$

is the entropy measured in nats.[14] This is identical to its counterpart in statistical physics up to a factor of the Boltzmann constant. Generalizing from discrete to continuous probabilities is not quite as simple as exchanging the sum for an integral; it is also necessary to introduce a measure function in the logarithm, otherwise the entropy would not be invariant under a simple parameter rescaling. For a continuous probability distribution $p(\lambda)$, we work in terms of the relative entropy (Ihara 1993, section 1.4)

$$H(p|q) = \int p(\lambda) \ln\left(\frac{p(\lambda)}{q(\lambda)}\right) \mathrm{d}\lambda, \tag{44}$$

where $q(\lambda)$ is another probability distribution, and we have changed the sign compared to the discrete case so that the entropy is non-negative. The relative entropy, or Kullback-Leibler divergence, measures the difference between distributions and is zero only if $p(\lambda) = q(\lambda)$ everywhere; with $p(\lambda)$ as the posterior and $q(\lambda)$ as the prior, it quantifies the information gained (Kullback & Leibler 1951).

The relative entropy is perfect for our purpose. It is zero when we do not observe a burst or the burst is uninformative such that we do not learn anything. Otherwise it scales approximately with the (logarithm of the) posterior width, giving an indication of how much could be learnt. For example, if $p(\lambda)$ and $q(\lambda)$ were both uniform distributions, with $p(\lambda)$ having $1/z$ the width of $q(\lambda)$, $H(p|q) = \ln z$; if they were both Gaussian with equal means, and $p(\lambda)$ had $1/z$ the width of $q(\lambda)$, $H(p|q) = \ln z - (1/2)(1 - z^{-2})$.

There is one complication in using the relative entropy. We have used an improper prior for $\ln(M_\bullet/\mathrm{M}_\odot)$; it is uniform over the entire real line and so cannot be normalized. As an alternative, we can use a Gaussian with parameters set by the current observations (Gillessen et al. 2009). The relative entropy then compares constraints from bursts with those from observing stellar motions.[15]

[14] The unit is set by the base of the logarithm; the more familiar bit is calculated using base 2, 1 bit $\equiv \ln(2)$ nats.

[15] Whilst this is a useful comparison, this does mean that the results are specific to the current state of knowledge and cannot be simply translated should we obtain updated measurements.

In practice, if we were trying to infer the mass of the MBH, we would combine all our data together to form a best estimate. Then a positive entropy would indicate that the final posterior is narrower than the current observational distribution. We do not incorporate our current knowledge of the MBH mass into our prior here because we are interested in what information is contained in EMRBs alone. Therefore, our posterior from EMRBs can be broader than this observational prior. In this case, the relative entropy can still be positive (since the distributions are different) even though we are not gaining information. In these cases, we set the entropy to zero by hand as we have not improved our relative state of knowledge.

The entropies calculated from multiple burst posteriors may show a similar bias to the distribution widths. This would reduce the value; hence, we use the largest calculated entropy. However, the entropy appears much less sensitive to the choice of the k-d tree used for the multiplication than the distribution widths. The entropies corresponding to the distributions in Fig. 3 differ only by 0.04 nats, less than 1 per cent of the value.

The entropies are well correlated with the logarithm of the distribution widths as expected. The distributions of the fraction of missions with entropies smaller than a given value are shown in Fig. 5. They closely mirror those in Fig. 4 (but the scale on the ordinate axis is now linear). There is a clustering at small entropies; the largest entropies are $\sim$13 nats for $\ln(M_\bullet/\mathrm{M}_\odot)$ and $\sim$15 nats for the other parameters.

Taking the average across all 100 mission realizations, we can calculate the expected information gain for each parameter:

$$\langle H_{\ln(M_\bullet/\mathrm{M}_\odot)}\rangle_{\mathrm{mission}} = 2.2 \pm 0.3 \,\mathrm{nats};$$

$$\langle H_{a_*}\rangle_{\mathrm{mission}} = 3.0 \pm 0.4 \,\mathrm{nats};$$

$$\langle H_{\cos\Theta_\mathrm{K}}\rangle_{\mathrm{mission}} = 2.8 \pm 0.4 \,\mathrm{nats};$$

$$\langle H_{\Phi_\mathrm{K}}\rangle_{\mathrm{mission}} = 3.7 \pm 0.4 \,\mathrm{nats}, \tag{45}$$

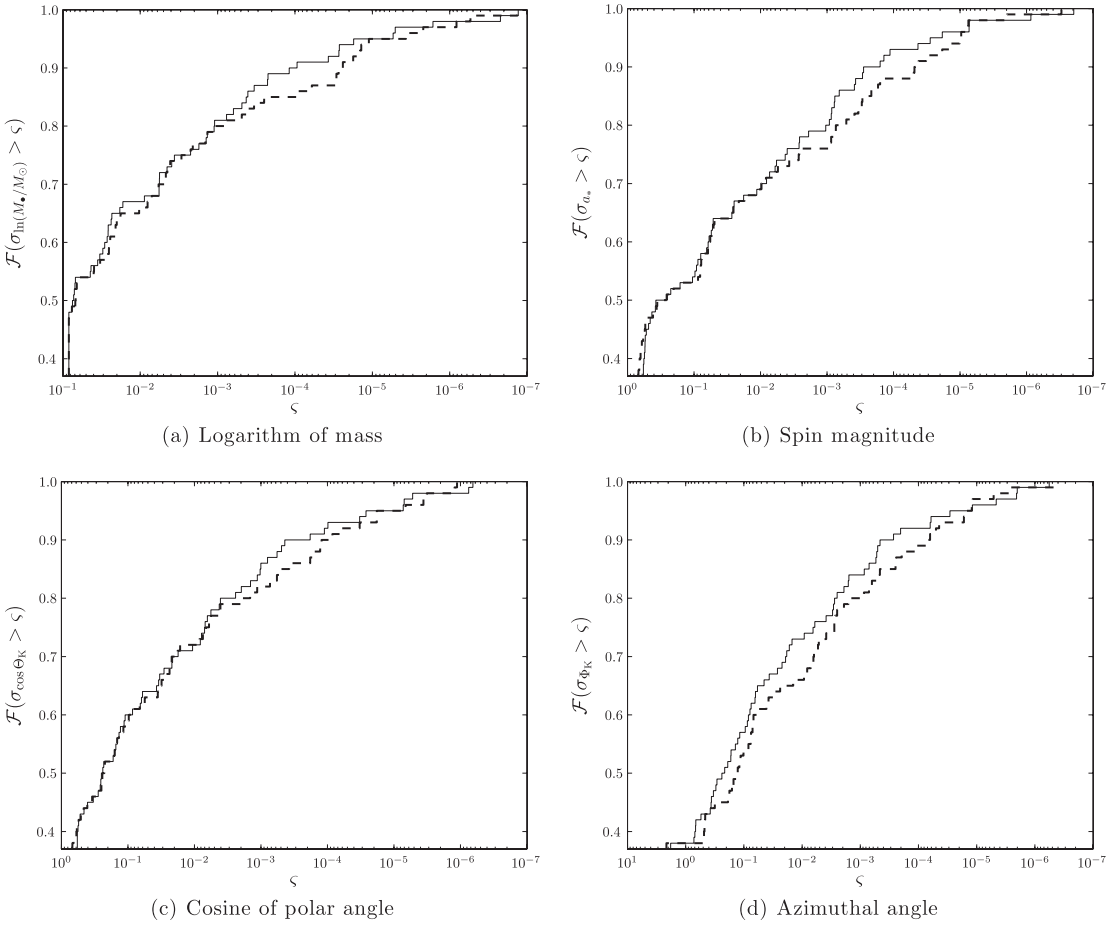
using $H_\lambda \equiv H(p(\lambda)|q(\lambda))$ for brevity. The quoted uncertainties are just the standard errors calculated from the scatter of entropies and do not include any of the other uncertainties. The typical entropy is about 3 nats; this corresponds to an improvement in the precision to which we know parameters by a factor of approximately 20.

5 CONCLUSION

EMRBs are a potentially interesting signal for a future space-based GW detector. The most promising source for EMRBs is the GC. We built a simple theoretical model to predict the Galactic EMRB event rate. The event rate is dominated by stellar mass BHs which form a cusp about the central MBH as a consequence of mass segregation. We calculate that there could be on average $\sim$1.7 detectable bursts over a two-year mission lifetime assuming a *LISA*-like interferometer, of which $\sim$1.2 are from BHs. The number of events scales linearly with the mission lifetime.

The event rate is not high: EMRBs shall not be a prolific GW source; however, the rate is not negligible. We are not guaranteed to have a burst in a mission lifetime, but it seems more likely than not that we shall have at least one.

The detectability of EMRBs is of little interest astrophysically unless we can extract information about their source systems. We investigated what could be learnt about the Galaxy's MBH. We created bursts for 100 mission realizations and characterized the posterior probability distributions for the MBH's parameters using MCMC sampling. In a large minority ($\sim$40 per cent) of



(a) Logarithm of mass (b) Spin magnitude

(c) Cosine of polar angle (d) Azimuthal angle

Figure 4. Cumulative proportion of mission realizations that produced posteriors with standard deviation (solid line) or 68-percentile half-width (dashed line) larger than the abscissa value.

realizations, we cannot improve upon our existing knowledge. However, in most cases we can, and it may be possible to gain a highly precise measurement of the MBH's mass and spin.

To quantify the information gained during a mission, we used the relative entropy with respect to our current knowledge. Averaging across all the missions, we found that we can expect to gain 2.2 nats of information about the logarithm of the mass, 3.0 nats about the spin magnitude, 2.8 nats about the cosine of the polar angle for the spin axis and 4.2 nats for the azimuthal angle. The entropy scales with the logarithm of the width of the posterior distribution; hence, these entropies represent improvements in the precision of our knowledge of the parameters by factors between $\sim$9 and $\sim$40. For the mass, this would mean that the uncertainty would become $\sim$1 per cent; we could expect to know the spin to a precision of the order of $\sim$0.1.

It must be stressed that, whilst these results are computed accurately based on the assumptions of the model, they are only to be trusted to an order of magnitude because of the significant uncertainties in the underlying assumptions. There are a number of sources of uncertainty found throughout our analysis. First, to calculate the EMRB waveforms, we employed the NK approximation assuming parabolic trajectories. The waveforms are easy to compute, but may have inaccuracies in their amplitude profiles of a few per cent at small periapses (Berry & Gair 2013a). This should not significantly influence detectability but may lead to differences in the shape of the posterior distributions. Since the errors in the waveforms are small, this should not qualitatively affect our results. Second, in calculating the event rate we made both mathematical and physical approximations. The former are correct to a few per cent and so are negligible compared to the latter. Our model, however, does include all the relevant physical processes, and further advances the previous work of Rubbo et al. (2006) and Hopman et al. (2007). Third, the astrophysical parameters used as inputs for our event rate calculation are themselves uncertain. Fourth, when calculating the constraints for our mission realizations, we only considered EMRBs from orbits with periapses smaller than $16r_{\rm g}$, yet whilst little information is expected from other bursts, the amount is not zero. This may lead us to slightly underestimate the total information that could be extracted from EMRBs. Finally, when combining posteriors from multiple EMRBs from the same mission, we binned our posterior distributions. This could lead to a small bias, making us underestimate the usefulness of bursts. Overall, the uncertainty in astrophysical parameters is likely to be the greatest source of error. As there are many unknowns regarding the physical assumptions, it is difficult to quantify the uncertainty in our results. As we learn more about the GC, we can be more confident in our prediction.

The centre of the Galaxy is a wonderful laboratory for testing our understanding of astrophysics, in particular for learning about MBHs and their influence on their environments. EMRBs could be a new means of probing this system.

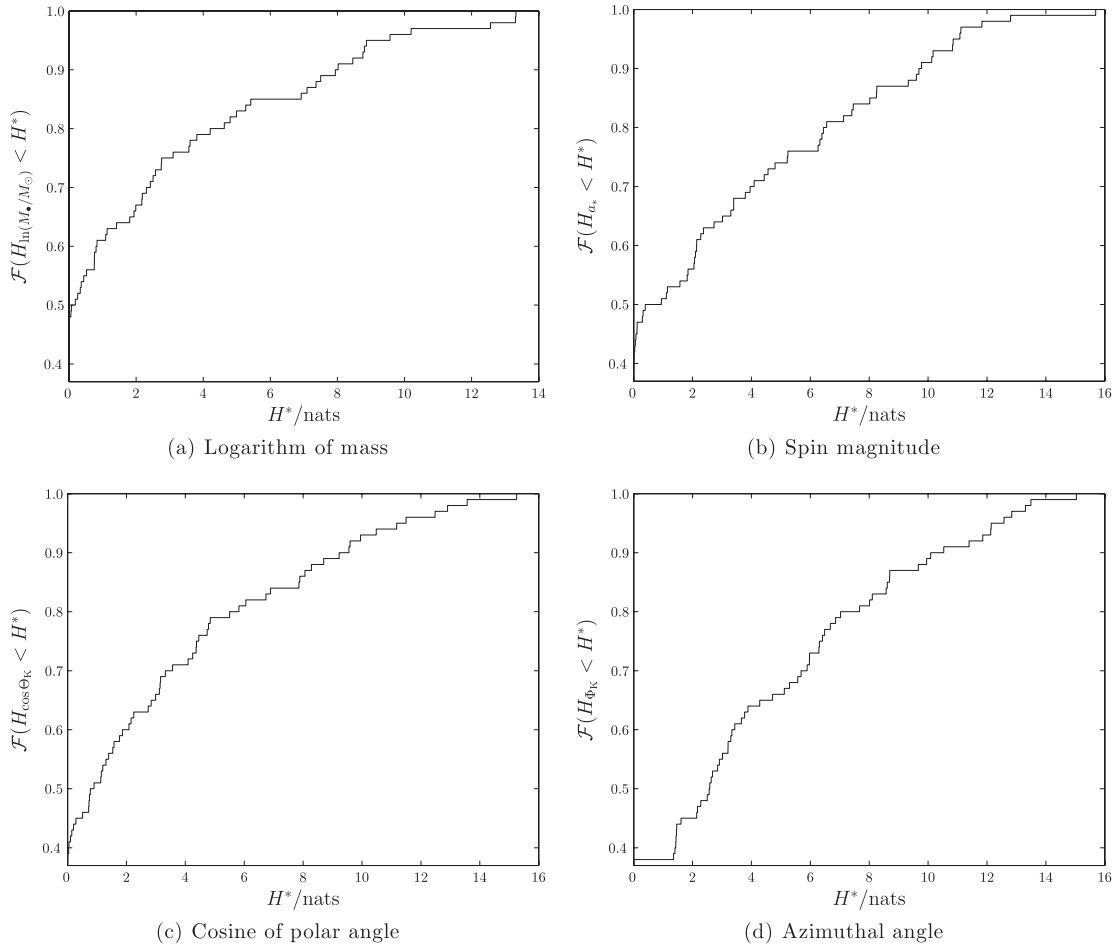


Figure 5. Cumulative proportion of mission realizations that produced posteriors with relative entropies larger than the abscissa value.

ACKNOWLEDGEMENTS

The authors are grateful for insightful discussions with Sverre Aarseth. They thank Tal Alexander and Clovis Hopman for useful correspondence. They also thank Rob Cole for productive suggestions and Fabio Antonini for useful comments. CPLB is supported by STFC. JRG is supported by the Royal Society. The MCMC simulations were performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (http://www.hpc.cam.ac.uk/), provided by Dell Inc. using Strategic Research Infrastructure Funding from the Higher Education Funding Council for England.

REFERENCES

Alexander T., 2005, Phys. Rep., 419, 65
Alexander T., Hopman C., 2003, ApJ, 590, L29
Alexander T., Hopman C., 2009, ApJ, 697, 1861
Amaro-Seoane P., Preto M., 2011, Class. Quantum Grav., 28, 094017
Amaro-Seoane P., Gair J. R., Freitag M., Miller M. C., Mandel I., Cutler C. J., Babak S., 2007, Class. Quantum Grav., 24, R113
Amaro-Seoane P. et al., 2012, Class. Quantum Grav., 29, 124016
Andrieu C., Thoms J., 2008, Stat. Comput., 18, 343
Antonini F., 2013, ApJ, 763, 62
Antonini F., Merritt D., 2012, ApJ, 745, 83
Antonini F., Capuzzo-Dolcetta R., Mastrobuono-Battisti A., Merritt D., 2012, ApJ, 750, 111
Antonucci F. et al., 2012, Class. Quantum Grav., 29, 124014
Anza S. et al., 2005, Class. Quantum Grav., 22, S125
Babak S., Fang H., Gair J., Glampedakis K., Hughes S., 2007, Phys. Rev. D, 75, 024005
Bahcall J. N., Wolf R. A., 1976, ApJ, 209, 214
Bahcall J. N., Wolf R. A., 1977, ApJ, 216, 883
Bar-Or B., Kupi G., Alexander T., 2013, ApJ, 764, 52
Barack L., 2009, Class. Quantum Grav., 26, 213001
Barack L., Cutler C., 2004, Phys. Rev. D, 69, 082005
Bartos I., Haiman Z., Kocsis B., Márka S., 2013, Phys. Rev. Lett., 110, 221102
Baumgardt H., Makino J., Ebisuzaki T., 2004, ApJ, 613, 1133
Bekenstein J. D., 1973, ApJ, 183, 657
Bekenstein J. D., Maoz E., 1992, ApJ, 390, 79
Bender P. et al., 1998, Technical report, LISA Pre-Phase A Report. Max-Planck-Institut für Quantenoptik, Garching
Berry C. P. L., Gair J. R., 2013a, MNRAS, 429, 589
Berry C. P. L., Gair J. R., 2013b, MNRAS, 433, 3572
Binney J., Tremaine S., 2008, Princeton Series in Astrophysics: Galactic Dynamics, 2nd edn. Princeton Univ. Press, Princeton, NJ
Chandrasekhar S., 1941a, ApJ, 93, 285
Chandrasekhar S., 1941b, ApJ, 94, 511
Chandrasekhar S., 1943a, ApJ, 97, 255
Chandrasekhar S., 1943b, ApJ, 98, 54
Chandrasekhar S., 1960, Principles of Stellar Dynamics, enlarged edn. Dover Publications, New York
Chatterjee P., Hernquist L., Loeb A., 2002, ApJ, 572, 371
Cohn H., Kulsrud R. M., 1978, ApJ, 226, 1087

Cutler C., 1998, Phys. Rev. D, 57, 7089
Cutler C., Flanagan E. E., 1994, Phys. Rev. D, 49, 2658
Danzmann K., Rüdiger A., 2003, Class. Quantum Grav., 20, S1
de Berg M., Cheong O., van Kreveld M., Overmars M., 2008, Computational Geometry: Algorithms and Applications, 3rd edn. Springer, Berlin
Doeleman S. S. et al., 2008, Nat, 455, 78
Drasco S., Hughes S., 2004, Phys. Rev. D, 69, 044015
Eilon E., Kupi G., Alexander T., 2009, ApJ, 698, 641
Ferrarese L., Ford H., 2005, Space Sci. Rev., 116, 523
Frank J., Rees M. J., 1976, MNRAS, 176, 633
Freitag M., Benz W., 2005, MNRAS, 358, 1133
Freitag M., Amaro-Seoane P., Kalogera V., 2006, ApJ, 649, 91
Gair J. R., Kennefick D. J., Larson S. L., 2005, Phys. Rev. D, 72, 084009
Genzel R. et al., 2003, ApJ, 594, 812
Genzel R., Eisenhauer F., Gillessen S., 2010, Rev. Mod. Phys., 82, 3121
Ghez A. M. et al., 2008, ApJ, 689, 1044
Gillessen S., Eisenhauer F., Trippe S., Alexander T., Genzel R., Martins F., Ott T., 2009, ApJ, 692, 1075
Glampedakis K., 2005, Class. Quantum Grav., 22, S605
Gualandris A., Merritt D., 2012, ApJ, 744, 74
Gürkan M. A., Hopman C., 2007, MNRAS, 379, 1083
Haario H., Saksman E., Tamminen J., 1999, Comput. Stat., 14, 375
Hastings W. K., 1970, Biometrika, 57, 97
Hills J. G., 1975, Nat, 254, 295
Hobson M. P., Efstathiou G., Lasenby A., 2006, General Relativity: An Introduction for Physicists. Cambridge Univ. Press, Cambridge
Hopman C., 2009, ApJ, 700, 1933
Hopman C., Alexander T., 2005, ApJ, 629, 362
Hopman C., Alexander T., 2006a, ApJ, 645, L133
Hopman C., Alexander T., 2006b, ApJ, 645, 1152
Hopman C., Freitag M., Larson S. L., 2007, MNRAS, 378, 129
Ihara S., 1993, Information Theory for Continuous Systems. World Scientific Publishing Company, London
Jennrich O. et al., 2011, Technical report, NGO Revealing a Hidden Universe: Opening a New Chapter of Discovery. ESA, Noordwijk
Just A., Khan F. M., Berczik P., Ernst A., Spurzem R., 2011, MNRAS, 411, 653
Kesden M., 2012, Phys. Rev. D, 85, 024037
Keshet U., Hopman C., Alexander T., 2009, ApJ, 698, L64
Kobayashi S., Laguna P., Phinney E. S., Meszaros P., 2004, ApJ, 615, 855
Kullback S., Leibler R. A., 1951, Ann. Math. Stat., 22, 79
Landau L. D., Lifshitz E. M., 1958, Course of Theoretical Physics: Statistical Physics. Pergamon Press, London
Li S., Liu F. K., Berczik P., Chen X., Spurzem R., 2012, ApJ, 748, 65
Lightman A. P., Shapiro S. L., 1977, ApJ, 211, 244
Lynden-Bell D., 1967, MNRAS, 136, 101
Lynden-Bell D., 1969, Nat, 223, 690
Lynden-Bell D., Kalnajs A. J., 1972, MNRAS, 157, 1
MacKay D. J. C., 2003, Information Theory, Inference and Learning Algorithms. Cambridge Univ. Press, Cambridge
Madigan A.-M., Hopman C., Levin Y., 2011, ApJ, 738, 99
Maoz E., 1993, MNRAS, 263, 75
Merritt D., 2010, ApJ, 718, 739
Merritt D., Berczik P., Laun F., 2007, AJ, 133, 553
Merritt D., Alexander T., Mikkola S., Will C. M., 2011, Phys. Rev. D, 84, 044024
Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, J. Chem. Phys., 21, 1087
Meyer L. et al., 2012, Sci, 338, 84
Miralda-Escude J., Gould A., 2000, ApJ, 545, 847
Misner C. W., Thorne K. S., Wheeler J. A., 1973, Gravitation. Freeman, New York
Mulder W. A., 1983, A&A, 117, 9
Murphy B. W., Cohn H. N., Durisen R. H., 1991, ApJ, 370, 60
Nelson R. W., Tremaine S., 1999, MNRAS, 306, 1
O'Leary R. M., Kocsis B., Loeb A., 2009, MNRAS, 395, 2127
Olver F. W. J., Lozier Daniel W., Boisvert R. F., Clark C. W., eds, 2010, NIST Handbook of Mathematical Functions. Cambridge Univ. Press, Cambridge
Peebles P. J. E., 1972, ApJ, 178, 371
Peters P. C., 1964, Phys. Rev., 136, B1224
Peters P. C., Mathews J., 1963, Phys. Rev., 131, 435
Press W., 1977, Phys. Rev. D, 15, 965
Preto M., Amaro-Seoane P., 2010, ApJ, 708, L42
Rauch K. P., Ingalls B., 1998, MNRAS, 299, 1231
Rauch K. P., Tremaine S., 1996, New Astron., 1, 149
Rees M. J., 1984, ARA&A, 22, 471
Rees M. J., 1988, Nat, 333, 523
Reid M. J., Brunthaler A., 2004, ApJ, 616, 872
Reid M. J., Readhead A. C. S., Vermeulen R. C., Treuhaft R. N., 1999, ApJ, 524, 816
Reid M. J., Menten K. M., Genzel R., Ott T., Schödel R., Brunthaler A., 2003, Astron. Nachr., 324, 505
Roberts G. O., Rosenthal J. S., 2007, J. Appl. Probab., 44, 458
Rubbo L. J., Holley-Bockelmann K., Finn L. S., 2006, ApJ, 649, L25
Ruffini R., Sasaki M., 1981, Prog. Theor. Phys., 66, 1627
Schödel R. et al., 2007, A&A, 469, 125
Shannon C. E., 1948a, Bell Syst. Tech. J., 27, 379
Shannon C. E., 1948b, Bell Syst. Tech. J., 27, 623
Shapiro S. L., Marchant A. B., 1978, ApJ, 225, 603
Sigurdsson S., Rees M. J., 1997, MNRAS, 284, 318
Spitzer L., Jr, 1987, Princeton Series in Astrophysics: Dynamical Evolution of Globular Clusters. Princeton Univ. Press, Princeton, NJ
Spitzer L., Jr, Harm R., 1958, ApJ, 127, 544
Spitzer L., Hart M. H., 1971, ApJ, 164, 399
Spitzer L., Jr, Schwarzschild M., 1951, ApJ, 114, 385
Spitzer L., Shapiro S. L., 1972, ApJ, 173, 529
Tanaka T., Shibata M., Sasaki M., Tagoshi H., Nakamura T., 1993, Prog. Theor. Phys., 90, 65
Toonen S., Hopman C., Freitag M., 2009, MNRAS, 398, 1228
Tremaine S., Weinberg M. D., 1984, MNRAS, 209, 729
Tremaine S. et al., 2002, ApJ, 574, 740
Turner M., 1977, ApJ, 216, 610
Volonteri M., 2010, A&AR., 18, 279
Weinberg M. D., 1986, ApJ, 300, 93
Weinberg M. D., 2012, Bayesian Anal., 7, 737
Wyse R. F. G., 2008, in Kodama T., Yamada T., Aoki K., eds, ASP Conf. Ser. Vol. 399, Panoramic Views of Galaxy Formation and Evolution. Astron. Soc. Pac., San Francisco, p. 445
Young P. J., 1977, ApJ, 217, 287
Yunes N. N., Sopuerta C. F., Rubbo L. J., Holley-Bockelmann K., 2008, ApJ, 675, 604
Yusef-Zadeh F., Choate D., Cotton W., 1999, ApJ, 518, L33

APPENDIX A: CHANDRASEKHAR'S RELAXATION TIME-SCALE

Chandrasekhar (1960, chapter 2) defined a relaxation time-scale for a stellar system by approximating the fluctuations in the stellar gravitational potential as a series of two-body encounters. The time over which the squared change in energy is equal to the squared (initial) kinetic energy of the star is the time taken for relaxation. Relaxation is mediated by dynamical friction (Chandrasekhar 1943a; Binney & Tremaine 2008, section 1.2). This can be understood as the drag induced on a star by the overdensity of field stars deflected by its passage (Mulder 1983). In the interaction between the star and its gravitational wake, energy and momentum are exchanged, accelerating some stars, decelerating others.

Chandrasekhar's approach has proved exceedingly successful despite the number of simplifying assumptions inherent in the model which are not strictly applicable to systems such as the Galactic core. We will not attempt to fix these deficiencies; the only modification is to substitute the velocity distribution.

Other authors have built upon the work of Chandrasekhar by considering inhomogeneous stellar distributions, via perturbation theory (Lynden-Bell & Kalnajs 1972; Tremaine & Weinberg 1984; Weinberg 1986), modelling energy transfer as anomalous dispersion, which adds higher order moments to the transfer probability (Bar-Or, Kupi & Alexander 2013), or using the tools of linear response theory and the fluctuation-dissipation theory (Landau & Lifshitz 1958, chapter 7), which allows relaxation of certain assumptions, such as homogeneity (Bekenstein & Maoz 1992; Maoz 1993; Nelson & Tremaine 1999). We will not attempt to employ such sophisticated techniques at this stage.

A1 Chandrasekhar's change in energy

We consider the interaction of a field star, denoted by 1, with a test star, 2; the change in energy squared from interaction over time δt is approximately (Chandrasekhar 1960, chapter 2)[16]

$$\Delta E^2(v_1) \simeq \frac{8\pi}{3} n(v_1) G^2 m_1^2 m_2^2 \ln\left(q v_2^2\right) \left\{ \begin{array}{ll} \dfrac{v_1^2}{v_2} & v_1 \leq v_2 \\ \dfrac{v_2^2}{v_1} & v_1 \geq v_2 \end{array} \right\} \mathrm{d}v_1 \delta t. \tag{A1}$$

Here v_1 and v_2 are the initial velocities, and m_1 and m_2 are the masses; $n(v_1)$ is the number of stars per velocity element $\mathrm{d}v_1$ which is calculated assuming that the density of stars is uniform.[17] The logarithmic term includes

$$q = \frac{D_0}{G\left(m_1 + m_2\right)}, \tag{A2}$$

where D_0 is the maximum impact parameter (Weinberg 1986). To eliminate the dependence upon v_1 requires a specific form for the velocity distribution.

A2 Velocity distributions

The velocity space DF can be obtained by integrating out the spatial dependence in the full DF. As we are restricting our attention to the core and assuming spherical symmetry,

$$f(v) = 4\pi \int_0^{r_\mathrm{c}} r^2 f(\mathcal{E})\, \mathrm{d}r, \tag{A3}$$

where r_c is defined by (8).

The DF for unbound stars is assumed to be Maxwellian as in (10). We assume violent relaxation such that $\sigma_M = \sigma$. Performing the integral

$$f_{\mathrm{u},\,M}(v) = \frac{n_\star}{\left(2\pi\sigma^2\right)^{3/2}} C_M \epsilon\left(\frac{v^2}{2\sigma^2}\right), \tag{A4}$$

introducing

$$\epsilon(w) = \frac{1}{2}\left\{\exp(-w)\left[4\exp(1) + \mathrm{Ei}(w) - \mathrm{Ei}(1)\right] - \frac{2 + w + w^2}{w^3}\right\}, \tag{A5}$$

where $\mathrm{Ei}(x)$ is the exponential integral.

The DF for bound stars is approximated as a simple power law as in (13). The integral gives

$$f_{\mathrm{b},\,M}(v) = \frac{n_\star}{\left(2\pi\sigma^2\right)^{3/2}} k_M \left(\frac{v^2}{2\sigma^2}\right)^{p_M - 3} \left\{ \begin{array}{ll} 3\mathrm{B}\left(\dfrac{v^2}{2\sigma^2}; 3 - p_M, 1 + p_M\right) & \dfrac{v^2}{2\sigma^2} \leq 1 \\ 3\mathrm{B}\left(3 - p_M, 1 + p_M\right) & \dfrac{v^2}{2\sigma^2} \geq 1 \end{array} \right., \tag{A6}$$

where $\mathrm{B}(x; a, b)$ is the incomplete beta function (Olver et al. 2010, section 8.17) and $\mathrm{B}(a, b) \equiv \mathrm{B}(1, a, b)$ is the complete beta function.

The velocity space density is related to the DF by

$$\frac{4\pi r_\mathrm{c}^3}{3} n_M(v_1) = 4\pi v_1^2 \left[f_{\mathrm{u},\,M}(v_1) + f_{\mathrm{b},\,M}(v_1)\right]. \tag{A7}$$

[16] As stressed by Antonini & Merritt (2012), it is important to include both the piece for $v_1 \leq v_2$ and $v_1 \geq v_2$ when calculating the effects of dynamical friction.

[17] The error introduced by this assumption can be partially absorbed by the appropriate choice of the Coulomb logarithm, which shall be introduced later (Just et al. 2011).

A3 Defining the relaxation time-scale

Using the specific forms for the velocity space density, we can calculate ΔE^2. The functional form depends upon the velocity of the test star. If $v_2^2/2\sigma^2 < 1$, then

$$\Delta E^2 \simeq \frac{16}{3}\sqrt{2\pi}\frac{G^2 m_1^2 m_2^2 n_\star}{\sigma^3}\ln\left(q v_2^2\right)\left(\frac{v_2^2}{2\sigma^2}\right)\left[k\frac{3}{(2-p)(1+p)}{}_3F_2\left(-1-p, 2-p, \frac{3}{2}; 3-p, \frac{5}{2}; \frac{v_2^2}{2\sigma^2}\right) + C\right]\delta t, \tag{A8}$$

where ${}_3F_2(a_1, a_2, a_3; b_1, b_2; x)$ is a generalized hypergeometric function (Olver et al. 2010, section 16).[18] The contribution from bound and unbound stars can be identified by the coefficients k and C, respectively. It is necessary to sum over all the species to get the total value.

If $v_2^2/2\sigma^2 > 1$,

$$\Delta E^2 \simeq \frac{16}{3}\sqrt{2\pi}G^2 m_1^2 m_2^2 n_\star \sigma \ln\left(q v_2^2\right)\left(\frac{v_2^2}{2\sigma^2}\right)^{-1/2}\left[k\beta\left(\frac{v_2^2}{2\sigma^2}; p\right) + C\alpha\left(\frac{v_2^2}{2\sigma^2}\right)\right]\delta t, \tag{A9}$$

where

$$\alpha(w) = \frac{1}{2}\left\{3w^{-1/2} + 5 + \left[4\exp(1) - \mathrm{Ei}(1) + \mathrm{Ei}(w)\right]\left[\frac{3\sqrt{\pi}}{4}\mathrm{erf}\left(w^{1/2}\right) - \frac{3}{2}w^{1/2}\exp(-w)\right] - 3\sqrt{\pi}\exp(1)\mathrm{erf}(1)\right.$$
$$\left. + 3\left[{}_2F_2\left(\frac{1}{2}, 1; \frac{3}{2}, \frac{3}{2}; 1\right) - w^{1/2}{}_2F_2\left(\frac{1}{2}, 1; \frac{3}{2}, \frac{3}{2}; w\right)\right]\right\}; \tag{A10}$$

$$\beta(w; p) = \begin{cases} \dfrac{3}{1/2-p}\left[\mathrm{B}\left(\dfrac{5}{2}, 1+p\right) - \dfrac{3w^{p-1/2}}{2(2-p)}\mathrm{B}\left(3-p, 1+p\right)\right] & p < \dfrac{1}{2} \\ \dfrac{\pi}{32}\left[12\ln(2) - 1 + 6\ln(w)\right] & p = \dfrac{1}{2} \end{cases}. \tag{A11}$$

Here ${}_2F_2(a_1, a_2; b_1, b_2; x)$ is another generalized hypergeometric function which originates from the integral

$$\int^w \frac{\exp(w')\mathrm{erf}\left(w'^{1/2}\right)}{w'}\mathrm{d}w' = \frac{4w^{1/2}}{\sqrt{\pi}}{}_2F_2\left(\frac{1}{2}, 1; \frac{3}{2}, \frac{3}{2}; w\right). \tag{A12}$$

Combining the two regimes for $v^2/2\sigma^2$, we can simplify using approximate forms. For the bound contribution

$$\Delta E_\mathrm{b}^2 \approx 16\sqrt{2\pi}G^2 m_1^2 m_2^2 n_\star \sigma \ln\left(q v_2^2\right) k\gamma\left(\frac{v_2^2}{2\sigma^2}; p\right)\delta t, \tag{A13}$$

where

$$\gamma(w; p) = \left(1 + w^4\right)^{-1}\left\{\left[\frac{3}{(1+p)(2-p)}w - \frac{9}{5(3-p)}w^2 + \frac{9p}{14(7-p)}w^3\right] + w^{7/2}\beta(w; p)\right\}. \tag{A14}$$

The resulting error, ignoring variation from $\ln\left(q v_2^2\right)$, is less than 3 per cent.

The unbound contribution is

$$\Delta E_\mathrm{u}^2 \approx \frac{16}{3}\sqrt{2\pi}G^2 m_1^2 m_2^2 n_\star \sigma \ln\left(q v_2^2\right) C\Xi\frac{v_2^2}{2\sigma^2}\left[\Xi^2 + \left(\frac{v_2^2}{2\sigma^2}\right)^3\right]^{-1/2}\delta t, \tag{A15}$$

where

$$\Xi = \lim_{w\to\infty}\{\alpha(w)\} \simeq 4.31. \tag{A16}$$

This reproduces the full function to better than 5 per cent, ignoring variation from $\ln\left(q v_2^2\right)$.

The relaxation time-scale is the time interval δt over which the squared change in energy becomes equal to the kinetic energy of the test star squared (Bar-Or et al. 2013)

$$\tau_\mathrm{R} = \left(\frac{m_2 v_2^2}{2}\right)^2\frac{\delta t}{\Delta E^2} \tag{A17}$$

$$\approx \frac{3v_2^4}{16\sqrt{2\pi}G^2 n_\star \sigma \ln\left(q v_2^2\right)}\left(\sum_M M^2\left\{k_M\gamma\left(\frac{v_2^2}{2\sigma^2}; p_M\right) + C_M\Xi\left(\frac{v_2^2}{2\sigma^2}\right)\left[\Xi^2 + \left(\frac{v_2^2}{2\sigma^2}\right)^3\right]^{-1/2}\right\}\right)^{-1}. \tag{A18}$$

[18] We have suppressed subscript M for brevity.

A4 Averaged time-scale

The relaxation time-scale (A18) is for a particular velocity v_2. This is not of much use to describe the core or even a (non-circular) orbit where there is a velocity range. It is necessary to calculate an average. Both the change in energy squared and the kinetic energy are averaged. We use two averages: over the distribution of bound velocities to give the relaxation time-scale for the system and over a single orbit. The former is of use when considering the inner cut-off of stars due to collisions, and the latter when considering the transition to GW inspiral.

A4.1 System relaxation time-scale

The total number of bound stars in the core is

$$N_{\mathrm{b},M} = \frac{3}{3/2 - p_M} \frac{\Gamma(p_M + 1)}{\Gamma(p_M + 7/2)} N_\star k_M, \tag{A19}$$

where $\Gamma(x)$ is the gamma function. Using this as a normalization constant, the probability of a bound star having a velocity in the range $v \to v + \mathrm{d}v$ is

$$4\pi v^2 p_{\mathrm{b},M}(v)\,\mathrm{d}v = \sqrt{\frac{2}{\pi}} \frac{v^2}{\sigma^3} \frac{(3/2 - p_M)\,\Gamma(p_M + 7/2)}{\Gamma(p_M + 1)} \left(\frac{v^2}{2\sigma^2}\right)^{p_M - 3} \begin{cases} \mathrm{B}\left(\dfrac{v^2}{2\sigma^2}; 3 - p_M, 1 + p_M\right) & \dfrac{v^2}{2\sigma^2} \leq 1 \\ \mathrm{B}(3 - p_M, 1 + p_M) & \dfrac{v^2}{2\sigma^2} \geq 1 \end{cases} \mathrm{d}v. \tag{A20}$$

The mean square velocity for bound stars in the core is then

$$\overline{v_M^2} = 3\sigma^2 \frac{3/2 - p_M}{1/2 - p_M}, \tag{A21}$$

assuming $p_M < 1/2$.

In the case $p_M = 1/2$, we encounter a logarithmic divergence. This reflects there being a physical cut-off.[19] We use $v_{\max} = c/2$, which is the maximum speed reached on a bound orbit about a Schwarzschild BH. Marginally higher speeds can be reached for prograde orbits about a Kerr BH, but the maximal velocity for retrograde orbits is marginally lower. In reality, we expect the maximum velocity to be lower due to a depletion of orbits. We also suspect that a simple Newtonian description of these orbits is imprecise, but a full relativistic description is beyond the scope of this analysis. For $p_M = 1/2$,

$$\overline{v_M^2} = \frac{\sigma^2}{2}\left[12\ln(2) - 5 + 6\ln\left(\frac{v_{\max}^2}{2\sigma^2}\right)\right]. \tag{A22}$$

Using a typical value of $\sigma = 10^5\,\mathrm{m\,s^{-1}}$,

$$\overline{v_M^2} \simeq 43\sigma^2. \tag{A23}$$

The mean square velocity is an order of magnitude greater than that for a Maxwellian distribution.

For the average of ΔE^2, we replace $\ln\left(q v_2^2\right)$ by a suitable average, so it may be moved outside the integral (Chandrasekhar 1960, chapter 2). We replace it by the Coulomb logarithm (Bahcall & Wolf 1976)

$$\ln\left(q\overline{v_2^2}\right) = \ln\Lambda_M \simeq \ln\left(\frac{M_\bullet}{M}\right). \tag{A24}$$

Just et al. (2011) find an extremely similar result fitting a Bahcall–Wolf cusp self-consistently. We calculate the averages for the bound and unbound populations individually and then combine these to obtain the total change for each species. We must distinguish between the bound population of field stars and the distribution of test stars over which we are averaging. We use subscripts M and M', respectively.[20] The bound average may be approximated to about 10 per cent accuracy as

$$\overline{\Delta E^2_{\mathrm{b},M'}} \approx \sum_M \frac{2^{11/2}}{3} G^2 M^2 M'^2 n_\star \sigma \ln(\Lambda_{M'}) k_M \frac{(3/2 - p_{M'})\Gamma(p_{M'} + 7/2)}{\Gamma(p_{M'} + 1)} \left[\varpi(p_M, p_{M'}) + \iota(p_M, p_{M'})\right]\delta t, \tag{A25}$$

introducing

$$\varpi(p_M, p_{M'}) = \frac{30 + 36p_M + 25p_M^2 - p_{M'}\left(13 + 15p_M + 7p_M^2\right) + p_{M'}^2\left(6 + 9p_M + 8p_M^2\right)}{210}; \tag{A26}$$

$$\iota(p_M, p_{M'}) = \mathrm{B}(3 - p_{M'}, 1 + p_{M'}) \begin{cases} \dfrac{3}{1/2 - p_M}\left[\dfrac{\mathrm{B}(5/2, 1 + p_M)}{2 - p_{M'}} - \dfrac{3\mathrm{B}(3 - p_M, 1 + p_M)}{2(2 - p_M)(5/2 - p_M - p_{M'})}\right] & p_M < \dfrac{1}{2} \\ \dfrac{\pi}{32}\dfrac{4 + p_{M'} + 12(2 - p_{M'})\ln(2)}{(2 - p_{M'})^2} & p_M = \dfrac{1}{2} \end{cases}. \tag{A27}$$

[19] A similar diverge necessitates the introduction of D_0 in Appendix A1.

[20] In a slight abuse of notation, we use for masses $m_M \equiv M$ and $m_{M'} \equiv M'$, and hope that it is clear that the summation is over the species.

The unbound component is approximately

$$\overline{\Delta E^2_{\mathrm{u},\,M'}} \approx \sum_M \frac{2^{11/2}}{3} G^2 M^2 M'^2 n_\star \sigma \ln(\Lambda_{M'})\, C_M \frac{(3/2 - p_{M'})\Gamma(p_{M'} + 7/2)}{\Gamma(p_{M'} + 1)} \times \left[\nu(p_{M'}) + \Xi \frac{\mathrm{B}(3 - p_{M'}, 1 + p_{M'})}{2 - p_{M'}} {}_2F_1\left(\frac{1}{2}, \frac{2 - p_{M'}}{3}; \frac{5 - p_{M'}}{3}; -\Xi^2\right)\right] \delta t, \tag{A28}$$

where

$$\nu(p) = \begin{cases} \dfrac{1}{1/2 - p}\left[\mathrm{B}\left(\dfrac{5}{2}, 1 + p\right) - \mathrm{B}(3 - p, 1 + p)\right] & p < \dfrac{1}{2} \\ \dfrac{\pi}{96}\left[12\ln(2) - 5\right] & p = \dfrac{1}{2} \end{cases}, \tag{A29}$$

and we have used ${}_2F_1(a_1, a_2, ; b_1; x)$, another hypergeometric function (Olver et al. 2010, equation 15.6.1). For consistency with the bound case, we have continued to use subscript M'.

The total relaxation time for a species is

$$\overline{\tau_{\mathrm{R},\,M'}} = \left(\frac{M' \overline{v^2_{M'}}}{2}\right)^2 \frac{\delta t}{\overline{\overline{\Delta E^2_{\mathrm{b},\,M'}}} + \overline{\overline{\Delta E^2_{\mathrm{u},\,M'}}}} \tag{A30}$$

$$\approx \frac{3}{2^{15/2}} \frac{\Gamma(p_{M'} + 1)}{(3/2 - p_{M'})\Gamma(p_{M'} + 7/2)} \frac{\overline{v^2_{M'}}^2}{G^2 n_* \sigma \ln(\Lambda_{M'})} \left\{\sum_M k_M M^2 \left[\varpi(p_M, p_{M'}) + \iota(p_M, p_{M'})\right] + C_M M^2 \left[\nu(p_{M'}) + \Xi \frac{\mathrm{B}(3 - p_{M'}, 1 + p_{M'})}{2 - p_{M'}} {}_2F_1\left(\frac{1}{2}, \frac{2 - p_{M'}}{3}; \frac{5 - p_{M'}}{3}; -\Xi^2\right)\right]\right\}^{-1}. \tag{A31}$$

Combining these to form an average for the entire system gives

$$\overline{\tau_{\mathrm{R}}} = \frac{\sum_{M'} N_{\mathrm{b},\,M'} \overline{\tau_{\mathrm{R},\,M'}}}{\sum_M N_{\mathrm{b},\,M}}. \tag{A32}$$

The relaxation time-scale for individual components is used in determining the collisional cut-off as described in Section 3.4.4.

A4.2 Orbital average

We calculate the time-scale for an orbit, parametrized by e and r_p, by averaging over one period.[21] The mean square velocity is

$$\left\langle v^2\left(e, r_\mathrm{p}\right)\right\rangle = \frac{G M_\bullet (1 - e)}{r_\mathrm{p}}. \tag{A33}$$

The orbital average is calculated according to (Spitzer 1987, section 2.2b)

$$\langle X \rangle = \frac{1}{T}\int_0^T X(t)\,\mathrm{d}t, \tag{A34}$$

where T is the orbital period. Despite our best efforts, we have been unsuccessful in obtaining analytic forms for the averaged changes in energy squared. Therefore, we compute them numerically. Switching to the orbital phase angle ϑ, we define

$$I_\mathrm{b}(e, \varrho, p) = \int_0^\pi \frac{1}{(1 + e\cos\vartheta)^2} \gamma\left(\frac{1}{2(1 + e)\varrho}\left(1 + e^2 + 2e\cos\vartheta\right); p\right) \mathrm{d}\vartheta \tag{A35}$$

$$I_\mathrm{u}(e, \varrho, \Xi) = \int_0^\pi \frac{\Xi}{(1 + e\cos\vartheta)^2} \left[\frac{1}{2(1 + e)\varrho}\left(1 + e^2 + 2e\cos\vartheta\right)\right] \left\{\Xi^2 + \left[\frac{1}{2(1 + e)\varrho}\left(1 + e^2 + 2e\cos\vartheta\right)\right]^3\right\}^{-1/2} \mathrm{d}\vartheta. \tag{A36}$$

The orbital relaxation time-scale is then

$$\left\langle \tau_{\mathrm{R},\,M'}\left(e, r_\mathrm{p}\right)\right\rangle = \left(\frac{G M_\bullet (1 - e) M'}{2 r_\mathrm{p}}\right)^2 \frac{\delta t}{\left\langle \Delta E^2_{\mathrm{b},\,M'}\right\rangle + \left\langle \Delta E^2_{\mathrm{u},\,M'}\right\rangle} \tag{A37}$$

$$\approx \frac{3}{64}\sqrt{\frac{\pi}{2}} \frac{M^2_\bullet (1 - e)^{1/2}}{n_\star \sigma r^2_\mathrm{p} (1 + e)^{3/2} \ln(\Lambda_{M'})} \left\{\sum_M \left[k_M M^2 I_\mathrm{b}\left(e, \frac{r_\mathrm{p}}{r_\mathrm{c}}, p_M\right) + C_M M^2 I_\mathrm{u}\left(e, \frac{r_\mathrm{p}}{r_\mathrm{c}}, \Xi\right)\right]\right\}^{-1}. \tag{A38}$$

This time-scale is defined similarly to the inspiral time-scale (33).

[21] We only consider bound orbits. The orbital relaxation time-scale is compared against the GW time-scale; the evolution of unbound orbits due to GW emission is negligible.

Diffusion in angular momentum proceeds over a shorter time, as defined by (32). Combining this with (A38) gives the orbital angular momentum relaxation time-scale.

A5 Discussion of applicability

In deriving the relaxation time-scales it has been necessary to make a number of approximations, both mathematical and physical. We have been careful to ensure that the mathematical inaccuracies introduced are of the order of a few per cent, and subdominant to the errors inherent from the physical assumptions and uncertainties in astronomical quantities. There are two key physical approximations that may limit the validity of the results.

First, it was assumed that the density of stars is uniform. This is a pragmatic assumption necessary to perform integrals over the impact parameter and angular orientation. This is not the case; however, as a star travels on its orbit, it moves through regions of different densities, sampling a range of different density–impact parameter distributions. Since we are only concerned with averaged time-scales, this partially smears out changes in density (cf. Just et al. 2011). To incorporate the complexity of the proper density distribution would greatly obfuscate the analysis.

Second, we have only considered transfer of angular momentum based upon the diffusion of energy, and not through resonant relaxation which enhances (both scalar and vector) angular momentum diffusion (Rauch & Tremaine 1996; Rauch & Ingalls 1998; Gürkan & Hopman 2007; Eilon, Kupi & Alexander 2009; Madigan, Hopman & Levin 2011). This occurs in systems where the radial and azimuthal frequencies are commensurate. Orbits precess slowly leading to large torques between the orbits. These torques cause the angular momentum to change linearly with time over a coherence time-scale set by the drift in orbits. Over longer time periods, the change in angular momentum again proceeds as a random walk, increasing with the square root of time, as for non-resonant relaxation, but is still enhanced because of the change in the basic step size. Diffusion of energy remains unchanged; there could be several orders of magnitude difference in the two relaxation time-scales.

Resonant relaxation is important in systems with (nearly) Keplerian potentials, but is quenched when relativistic precession becomes significant: inside the Schwarzschild barrier (Merritt et al. 2011). It is less likely to be of concern for the orbits influenced by GW emission (Sigurdsson & Rees 1997) and should not be significant for our purposes.

The optimal approach would be to perform a full N-body simulation of the Galactic core. This would dispense with all the complications of considering relaxation time-scales and estimates for cut-off radii. Unfortunately, such a task still remains computationally challenging at the present time (e.g., Li et al. 2012).

A6 Time-scales for the Galactic core

Evaluating $\overline{\tau_{\mathrm{R}}}$ for the Galactic core (Section 3.2) and comparing with $\tau_{\mathrm{R}}^{\mathrm{Max}}$, (30) using $\kappa = 0.34$, shows a broad consistency:

$$\overline{\tau_{\mathrm{R}}} \simeq 2.0\tau_{\mathrm{R}}^{\mathrm{Max}}. \tag{A39}$$

This is reassuring since the standard Maxwellian approximation has been successful in characterizing the properties of the Galactic core. We calculated $\tau_{\mathrm{R}}^{\mathrm{Max}}$ for the dominant stellar component alone, which gives $\tau_{\mathrm{R}}^{\mathrm{Max}} \simeq 4.5 \times 10^9$ yr.

Looking at the time-scales for each species in turn:

$$\overline{\tau_{\mathrm{R,\,MS}}} \simeq 1.7\tau_{\mathrm{R}}^{\mathrm{Max}}; \quad \overline{\tau_{\mathrm{R,\,WD}}} \simeq 1.6\tau_{\mathrm{R}}^{\mathrm{Max}}; \quad \overline{\tau_{\mathrm{R,\,NS}}} \simeq 2.1\tau_{\mathrm{R}}^{\mathrm{Max}}. \tag{A40}$$

Again there is good agreement.[22] For BHs,

$$\overline{\tau_{\mathrm{R,\,BH}}} \simeq 48\tau_{\mathrm{R}}^{\mathrm{Max}}. \tag{A41}$$

The time-scales for the lighter components are of the order of the Hubble time; the BH time-scale is much longer on account of the higher mean square velocity. This may indicate that the BH population is not fully relaxed (cf. Antonini & Merritt 2012): there has not been sufficient time for objects to diffuse on to the most tightly bound orbits (in which case, the mean square velocity would be lower). We expect that many of the most tightly bound BHs are not in a relaxed state, since GW inspiral is the dominant effect in determining the profile. This would deplete some of the innermost orbits and lower the mean square velocity for the population. Since we do not consider the collisional disruption of BHs, we do not use $\overline{\tau_{\mathrm{R,\,BH}}}$ in our model; it therefore has no influence on our results.

The long BH time-scale also inevitably includes an artefact of our approximation that the system is homogeneous: in reality the BHs, being more tightly clustered towards the centre, pass through regions with greater density (both because of a higher number density and a greater average object mass). Therefore, we expect the true relaxation time-scale to be reduced.

Formation of the cusp can occur over shorter time than the relaxation time-scale (Bar-Or et al. 2013). It should proceed on a dynamical friction time-scale $\tau_{\mathrm{DF}} \approx (M_\star/M')\overline{\tau_{\mathrm{R},\,M'}}$ (Spitzer 1987, section 3.4). This reduces the difference between the different species, but does not make it obvious that the cusp has had sufficient time to form, especially if there has been a merger in the Galaxy's history which disrupted the central distribution of stars (Gualandris & Merritt 2012). Fortunately, observations of the thick disc indicate that there has not been a major merger in the last 10^{10} yr (Wyse 2008). Minor mergers, where (globular) clusters spiral in towards the MBH, have been suggested as a means of building the stellar population that is consistent with current observations (Antonini et al. 2012; Antonini 2013). These could prevent the

[22] Freitag, Amaro-Seoane & Kalogera (2006) found that using a consistent velocity distribution for the population of stars (from an η-model), instead of relying on the Maxwellian approximation, made negligible change to the dynamical friction time-scale. They did not consider a cusp as severe as $p = 0.5$.

cusp from forming if there has not been sufficient time for the stars to relax post-merger. In any case, the time taken to form a cusp depends upon the initial configuration of stars, and so depends upon the Galaxy's history.

The existence of a cusp is a subject of debate. Preto & Amaro-Seoane (2010) conducted *N*-body simulations to investigate the effects of strong mass segregation (Alexander & Hopman 2009; Keshet et al. 2009) and found that cusps formed in a fraction of a (Maxwellian) relaxation time (Amaro-Seoane & Preto 2011). Gualandris & Merritt (2012) conducted similar computations and found that cores are likely to persist for the dominant stellar popular; intriguingly, cusp formation amongst BHs is quicker, but still takes at least a (Maxwellian) relaxation time. We cannot add further evidence to settle the matter. Our state of understanding may be improved following the passage through periapse of the gas cloud G2 this year (Bartos et al. 2013). For definiteness, we have assumed that a cusp has formed in our calculations.

Time-scales for individual orbits range over many orders of magnitude. The longest are for the most tightly bound: the cusp forms from the outside-in, and these orbits may not yet be populated. The shortest time-scales are for the most weakly bound orbits, those with large periapses and eccentricities. The orbital period can be much shorter than these time-scales, highlighting the fringe where the Fokker–Planck approximation is not appropriate (Spitzer & Shapiro 1972). The variation in the time-scale is exaggerated by neglecting the spatial variation in the stellar population.

When comparing GW inspiral time-scales and orbital angular momentum time-scales, equality can occur for times far exceeding the Hubble time. This only occurs for lower eccentricities, which are not of interest for bursts. However, it may be interesting to consider the stellar distribution in this region, which is not relaxed but dominated by GW inspiral. Since inspiral takes such a huge time to complete, it is possible that there is a pocket of objects currently mid-inspiral that reflect the unrelaxed distribution.

APPENDIX B: EVOLUTION OF ORBITAL PARAMETERS FROM GW EMISSION

B1 Bound orbits

For bound orbits, we can define a GW inspiral time from the orbit-averaged change in the orbital parameters. Using the analysis of Peters (1964) for Keplerian binaries, the averaged rates of change of the periapsis and eccentricity are

$$\left\langle \frac{\mathrm{d}r_\mathrm{p}}{\mathrm{d}t} \right\rangle = -\frac{64}{5}\frac{G^3 M_\bullet M(M_\bullet + M)}{c^5 r_\mathrm{p}^3}\frac{(1-e)^{3/2}}{(1+e)^{7/2}}\left(1 - \frac{7}{12}e + \frac{7}{8}e^2 + \frac{47}{192}e^3\right) \tag{B1}$$

$$\left\langle \frac{\mathrm{d}e}{\mathrm{d}t} \right\rangle = -\frac{304}{15}\frac{G^3 M_\bullet M(M_\bullet + M)}{c^5 r_\mathrm{p}^4}\frac{e(1-e)^{3/2}}{(1+e)^{5/2}}\left(1 + \frac{121}{304}e^2\right). \tag{B2}$$

For a circular orbit, the inspiral time from initial periapsis $r_{\mathrm{p}0}$ is

$$\tau_\mathrm{c}(r_{\mathrm{p}0}) = \frac{5}{256}\frac{c^5 r_{\mathrm{p}0}^4}{G^3 M_\bullet M(M_\bullet + M)}. \tag{B3}$$

For an orbit of non-zero eccentricity ($0 < e < 1$), we can solve for the periapsis as a function of eccentricity

$$r_\mathrm{p}(e) = \mathcal{R}(1+e)^{-1}\left(1 + \frac{121}{304}e^2\right)^{870/2299} e^{12/19}, \tag{B4}$$

where $\mathcal{R}$ is fixed by the initial conditions: for an orbit with initial eccentricity e_0,

$$\mathcal{R}(e_0) = (1+e_0)\left(1 + \frac{121}{304}e_0^2\right)^{-870/2299} e_0^{-12/19} r_{\mathrm{p}0}. \tag{B5}$$

The inspiral is complete when the eccentricity has decayed to zero; the inspiral time is (Peters 1964)

$$\tau_\mathrm{insp}(r_{\mathrm{p}0}, e_0) = \int_0^{e_0} \frac{15}{304}\frac{c^5\mathcal{R}^4}{G^3 M_\bullet M(M_\bullet + M)}\frac{e^{29/19}}{(1-e^2)^{3/2}}\left(1 + \frac{121}{304}e^2\right)^{1181/2299}\mathrm{d}e. \tag{B6}$$

This is best evaluated numerically, but it may be written in closed form as

$$\tau_\mathrm{insp}(r_{\mathrm{p}0}, e_0) = \tau_\mathrm{c}(r_{\mathrm{p}0})(1+e_0)^4\left(1 + \frac{121}{304}e_0^2\right)^{-3480/2299} F_1\left(\frac{24}{19}; \frac{3}{2}, -\frac{1181}{2299}; \frac{43}{19}; e_0^2, -\frac{121}{304}e_0^2\right), \tag{B7}$$

using the Appell hypergeometric function of the first kind $F_1(\alpha; \beta, \beta'; \gamma; x, y)$ (Olver et al. 2010, equation 16.15.1).[23]

B2 Unbound orbits

Unbound objects only pass through periapsis once. We therefore expect the orbital change from gravitational radiation to be small. Following the approach of Turner (1977), we can calculate the evolution in the eccentricity and periapse of an unbound Keplerian binary. The change in fractional eccentricity over an orbit, approximating the orbital parameters as constant, is

$$\frac{\Delta e}{e} = -\frac{608}{15}\Sigma\left[\frac{1}{(1+e)^{5/2}}\left(1 + \frac{121}{304}e^2\right)\cos^{-1}\left(-\frac{1}{e}\right) + \frac{(e-1)^{1/2}}{e^2(1+e)^2}\left(\frac{67}{456} + \frac{1069}{912}e^2 + \frac{3}{38}e^4\right)\right], \tag{B8}$$

[23] For small eccentricities, $\tau_\mathrm{insp}(r_{\mathrm{p}0}, e_0) \simeq \tau_\mathrm{c}(r_{\mathrm{p}0})[1 + 4e_0 + (273/43)e_0^2 + \mathcal{O}(e_0^3)]$.

introducing the dimensionless parameter

$$\Sigma = \frac{G^{5/2} M_\bullet M (M_\bullet + M)}{c^5 r_{\rm p}^{5/2}}. \tag{B9}$$

Similarly, the fractional change in periapsis is

$$\frac{\Delta r_{\rm p}}{r_{\rm p}} = -\frac{128}{5}\Sigma \left[\frac{1}{(1+e)^{7/2}} \left(1 - \frac{7}{12}e + \frac{7}{8}e^2 + \frac{47}{192}e^3 \right) \cos^{-1}\left(-\frac{1}{e}\right) - \frac{(e-1)^{1/2}}{e(1+e)^3} \left(\frac{67}{288} - \frac{13}{8}e + \frac{133}{576}e^2 - \frac{1}{4}e^3 - \frac{1}{8}e^4 \right) \right]. \tag{B10}$$

Both of these changes obtain their greatest magnitudes for large eccentricities, then

$$\frac{\Delta e}{e} \simeq \frac{\Delta r_{\rm p}}{r_{\rm p}} \simeq -\frac{16}{5}\Sigma e^{1/2}. \tag{B11}$$

For extreme-mass-ratio binaries, as is the case here, the mass ratio is a small quantity

$$\eta = \frac{M}{M_\bullet} \ll 1. \tag{B12}$$

The smallest possible periapsis is of the order of the Schwarzschild radius of the MBH, such that

$$r_{\rm p} = \alpha \frac{G M_\bullet}{c^2}; \quad \alpha > 1. \tag{B13}$$

These give

$$\Sigma = \frac{\eta}{\alpha^{5/2}} < \eta \ll 1. \tag{B14}$$

Hence, the changes in the orbital parameters become significant for

$$e \sim \frac{25}{256}\frac{\alpha^5}{\eta^2} > \frac{25}{256}\frac{1}{\eta^2}. \tag{B15}$$

Such orbits should be exceedingly rare, and so it is safe to neglect inspiral for unbound orbits.

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