No tension between assembly models of super massive black hole binaries and pulsar observations

Hannah Middleton, Siyuan Chen, Walter Del Pozzo, Alberto Sesana & Alberto Vecchio

Pulsar timing arrays are presently the only means to search for the gravitational wave stochastic background from super massive black hole binary populations, considered to be within the grasp of current or near-future observations. The stringent upper limit from the Parkes Pulsar Timing Array has been interpreted as excluding (>90% confidence) the current paradigm of binary assembly through galaxy mergers and hardening via stellar interaction, suggesting evolution is accelerated or stalled. Using Bayesian hierarchical modelling we consider implications of this upper limit for a range of astrophysical scenarios, without invoking stalling, nor more exotic physical processes. All scenarios are fully consistent with the upper limit, but (weak) bounds on population parameters can be inferred. Recent upward revisions of the black hole–galaxy bulge mass relation are disfavoured at 1.6σ against lighter models. Once sensitivity improves by an order of magnitude, a non-detection will disfavour the most optimistic scenarios at 3.9σ.
Dedicated timing campaigns of ultra-stable radio pulsars lasting over a decade and carried out with the best radio telescopes around the globe have targeted the isotropic gravitational-wave (GW) background in the frequency region $10^{-10}–10^{-11}$ Hz generated by the cosmic population of merging super massive black hole binaries (SMBHBs). In the hierarchical clustering scenario of galaxy formation, galaxies form through a sequence of mergers. In this process, the SMBHs hosted at their centers will inevitably form a large number of binaries, forming an abundant population of GW sources in the Universe. Detecting and/or placing constraints on their emitted signal will thus provide an insight into the formation and evolution of SMBHBs in connection with their galaxy hosts and will help to better understand the role played by SMBHBs in galaxy evolution and the dynamical processes operating during galaxy mergers (for a review see ref. 3).

No detection at nHz frequencies has been reported so far. The most stringent constraint on an isotropic background radiation has been obtained through an 11-year-long timing of 4 radio pulsars by the Parkes Pulsar Timing Array (PPTA). It yields an upper limit on the GW characteristic amplitude of $h_{\nu y} = 1.0 \times 10^{-15}$ (at 95% confidence) at a frequency of 1 yr$^{-1}$. Consistent results, although a factor $\approx 2$ less stringent, have also been reported by the European PTA$^5$, the North American Nanohertz Observatory for Gravitational Waves$^6$ and the International PTA$^7$, an international consortium of the three regional PTA collaborations. Those values are in the range of signal amplitudes predicted by state-of-the-art SMBHB population models, and can therefore be used to constrain such a population. It has been noted, however, that these limits start to be sensitive to uncertainties in the determination of the solar system ephemeris used in the analysis. Recent unpublished work has in fact found that different ephemeris choices can result in a partial degradation of the upper limit. This is still an active area of research which may lead to a small upward revision of the upper limit, a circumstance which, if anything, will strengthen the research which may lead to a small upward revision of the upper limit, a circumstance which, if anything, will strengthen the

Results

Inference using the upper limit. For each model, we use a Bayesian hierarchical analysis to compute the model evidence (which is the probability of the model given the data and allows for the direct comparison of models) and posterior density functions on the model parameters given the observational results reported by ref. 4. We find that the upper limit is now beginning to probe the most optimistic predictions, but all models are so far consistent with the data. Figure 1, our main result, compares the predictions under different model assumptions with the observed upper limit. The dotted area shows the prior range of the GW amplitude under the model assumptions, and the orange solid line shows the 95% confidence PPTA upper limit on $h_c$. The (central) 68% and 90% posterior probability intervals on $h_c$ are shown by the shaded blue bands. The posterior density functions (PDFs) on the right hand side of each plot gives the prior (black dashed line) and posterior (blue line) for $h_c$ at a reference frequency of $f = 1/5$ yr$^{-1}$.

The difference between the dotted region and the shaded bands in the main panels in Fig. 1 indicates the constraining power of the Parkes PTA limit on astrophysical models—the greater the difference between the two regions, the smaller is the consistency of that particular model with the data. We see that although some upper portion of the allowable prior region is removed from the 90% posterior probability interval (less so for S16), none of the models can be ruled out at any significant level. The confidence bands across the frequency range are constructed by taking the relevant credibility region of the posterior distribution of $h_c$ at each frequency, and therefore the boundaries of each band do not follow any particular functional form as a function of frequency. In addition, although eccentricity is allowed by the data, the power-law spectrum of circular binaries driven by radiation reaction alone can clearly be consistently placed within these bands (see also Supplementary Fig. 1 for further details on the individual parameter including eccentricity). This can be quantified in terms of the model evidences $Z$, shown in Table 1. The normalization is chosen so that a putative model unaffected by the limit yields $Z = 1$ and therefore the values can be interpreted as Bayes factors against such a model. None of the posterior probabilities of the models with respect to this putative one show any tension. As an example,
The least favoured model in the range of those considered here is KH13; with Bayes factors in favour of the others ranging from 1 to 1.13 to \(\approx 1.76\). These are however values of order unity and no decisive inference can be made from the data\(^{21}\). Comparisons between each of the individual model parameters (see Methods) posterior and prior distribution functions are described in Supplementary Fig. 1 and Supplementary Table 1, which further support our conclusions. For KH13, the model that produces the strongest GW background, we find a probability of \(e^{-2.36} = 0.094\) with respect to a putative model that is unaffected by the limit. KH13 is therefore disfavoured at \(~1.6\sigma\). This conclusion is reflected in the value of the K-L divergence of 0.85 (this is the same K-L divergence as between two Gaussian distributions with the same variance and means \(~1.3\) standard deviation apart). We note that ref.\(^{4}\) choose in their analysis only a sub-sample of the \(^9\) models, with properties similar to KH13. Our results for KH13 are therefore consistent with the 91%-to-97% ’exclusion’ claimed by ref.\(^{4}\).

### Discussion

It is argued in ref.\(^{4}\) that the Parkes PTA upper-limit excludes at high confidence standard models of SMBH assembly—i.e., those considered in this work—and therefore these models need to be substantially revised to accommodate either accelerated mergers

---

**Fig. 1** The posterior density function on the gravitational wave characteristic amplitude. The four panels compare the prior and posterior density functions on the GW stochastic background characteristic amplitude in light of the PPTA upper limit for each of the astrophysical models considered here: a) S16; b) KH13; c) G09; d) ALL. The central 90% region of the prior is indicated by the black dotted band and the posterior is shown by the progressively lighter blue shading indicating the central 68% and 90% regions, respectively, along with the median (solid blue line). Also shown are the PPTA bin-by-bin limit (orange solid line) and the corresponding integrated limit assuming \(h_c(f) \propto f^{2/3}\) (orange star and vertical dotted line). The difference in the prior and posterior indicates how much has been learnt from the PPTA data. In each panel, the right-hand side one-dimensional distribution shows the prior (black dashed) and posterior (blue solid) at a reference frequency of \(f\sim 1/5\) yr\(^{-1}\), with the central 90% regions marked (black and blue dashed lines respectively).

---

| Model | \(h_{w} = 1 \times 10^{-15}\) (PPTA) | \(h_{w} = 3 \times 10^{-16}\) | \(h_{w} = 1 \times 10^{-16}\) |
|-------|----------------------------------|----------------------------------|----------------------------------|
|        | K-L divergence | \(\log Z\) | K-L divergence | \(\log Z\) | K-L divergence | \(\log Z\) |
| KH13  | 0.85 | -2.36 | 2.25 | -5.68 | 5.18 | -13.17 |
| G09   | 0.39 | -1.2 | 1.11 | -3.35 | 2.86 | -8.26 |
| S16   | 0.37 | -0.6 | 0.69 | -1.62 | 1.42 | -3.82 |
| ALL   | 0.62 | 1.23 | 1.33 | -2.68 | 2.50 | -5.74 |

The values in the table show the K-L divergence and natural logarithm of the evidence, \(\log Z\), for each of the four astrophysical models given the PPTA upper limit at \(h_{w} = 1 \times 10^{-15}\) and for more stringent putative limits at the levels of \(3 \times 10^{-16}\) and \(1 \times 10^{-16}\).
in finding new stable millisecond pulsars, will provide the necessary ground to improve sensitivity down to $h_{1\text{yr}} \sim 10^{-16}$, which is in line with the lower limit of the expected stochastic gravitational wave background according to our current understanding of SMBH evolution$^{26}$. Although not yet decisive, our findings highlight the potential of PTAs in informing the current debate on the SMBH-host galaxy relation. Recent discoveries of over-massive black holes in brightest cluster ellipticals$^{27,28}$ led to an upward revision of those relations$^{29,30}$. However, several authors attribute the high normalization of the recent SMBH-host galaxy relations to selection biases$^{16}$ or to the intrinsic difficulty of resolving the SMBH fingerprint in measurements based on stellar dynamics (see discussion in ref. 30). Future facilities such as the Extremely Large Telescope$^{31}$ and the Thirty Meter Telescope$^{32}$ will likely measure many more SMBH masses in elliptical galaxies$^{33}$, providing a better understanding of the SMBH-host galaxy relations. PTA limits may therefore be used to gain more information about the other underlying uncertainties in the model, in particular the massive galaxy merger rate, which is currently poorly constrained observationally (e.g. see refs. $34,35$).

An important question is: what is the sensitivity level required to really put under stress our current understanding of SMBHB assembly? If a null result persists in PTA experiments, this will in turn lead to a legitimate re-thinking of the PTA observing strategy to target possibly more promising frequencies of the GW spectrum. To address this question, we simulate future sensitivity improvements by shifting the Parkes PTA sensitivity curve down to provide 95% upper limits of $h_{1\text{yr}}$ at $3 \times 10^{-16}$ and $1 \times 10^{-16}$. The results are summarized in Table 1 and more details are provided in Supplementary Fig. 2, Supplementary Table 2 and Supplementary Note 1. At $3 \times 10^{-16}$, possibly within the sensitivity reach of PTAs in the next 5 years, S16 will be significantly favoured against KH13, with a Bayes factor of $c_{1,06}$, and only marginally favoured over G09, with Bayes factor of $c_{1,76}$. It will still be impossible to reject this model at any reasonable significant level with respect to, say, a model which predicts negligible GW background radiation at $\sim 10^{-9} - 10^{-8}$ Hz. However SMBH-host galaxy relations with high normalizations will show a $\approx 2\sigma$ tension with more conservative models. At $1 \times 10^{-16}$, within reach in the next decade with the advent of MeerKAT, FAST and SKA, models KH13, G09 and ALL are disfavoured at $3\sigma$, $2.5\sigma$ and $1.2\sigma$, respectively, in comparison with S16. K-L divergences in the range 5.18–1.42 show that the data are truly informative. S16 is also disfavoured at $2.3\sigma$ with respect to a model unaffected by the data, possibly indicating the need of additional physical processes to be included in the models.

**Methods**

**Analytical description of the GW background.** The GW background from a cosmic population of SMBHBs is determined by the binary merger rate and by the dynamical properties of the systems during their inspiral. The comoving number density of SMBHBs per unit log mass chirp ($N = (M_1 M_2)^{3/5} / (M_1 + M_2)$) and unit solid angle defines the normalization of the GW spectrum. If all binaries evolve under the influence of GW backreaction only in circular orbits, then the spectral index is fixed at $h_{\text{GW}} \propto f^{-2/3}$ and the GW background is fully determined$^{26}$. However, to get to the point at which GW emission is efficient, SMBHBs need to exchange energy and angular momentum with their stellar and/or gaseous environment$^{1}$, a process that can lead to an increase in the binary eccentricity (e.g. see refs. $37,47$). We assume SMBHBs evolve via three-body scattering against the dense stellar background up to a transition frequency $f_\text{GW}$ which GW emission takes over. According to recent studies$^{39,40}$, the hardening is dictated by the density of background stars $n_\text{bg}$ at the influence radius of the binary $r_\text{influence}$. The bulge stellar density is assumed to follow a Hernquist density profile$^{41}$ with total mass $M_\text{bg}$ and scale radius $a$ defined by the SMBHB total mass $M_\text{SMBHB}$ max mass $M_\text{SMBHB}$ via empirical relations from the literature (see full details in ref. $12$). Therefore, for each individual system, $r_\text{influence}$ is defined solely by $M_\text{SMBHB}$. In the stellar hardening phase, the binary is assumed to hold constant eccentricity $e$ up to $f_\text{GW}$, beyond which it circularizes under the effect of the now dominant GW backreaction. The GW spectrum emitted by an individual binary adiabatically inspiralling under these processes to be included in the models.
for $f > f_s$. The spectrum has a turnaround around $f_s$ and its exact location depends on the binary eccentricity $e$. The observed GW spectrum is therefore uniquely determined by the binary chirp mass $M$, redshift $z$, transition frequency $f_t$, and eccentricity at transition $e_t$.

The GW spectrum from the overall population can be computed by integrating the spectrum of each individual system over the co-moving number density of merging SMBHBs

$$h_c^2(f) = \int dM \int d\log_{10} M \int d^n \theta \left( C_1 \left( \frac{M}{C_0 M_0} \right)^{\alpha} + C_2 \left( \frac{M}{C_0 M_0} \right)^{\beta} \right) \left( \frac{f}{f_0} \right)^{5/3} \left( \frac{z}{z_0} \right)^{-1/3},$$

(1)

where $h_c$ is an analytic fit to the GW spectrum of a reference binary with chirp mass $M_0$ at redshift $z_0$ (i.e., assuming $d^n \theta / d \log_{10} M dz = 6(\delta(M - M_0)\delta(z - z_0)$, characterized by an eccentricity of $e_t$ at a reference frequency $f_0$. For these reference values, the peak frequency of the spectrum $f_0$ is computed. The contribution of a SMBHB with generic chirp mass, emission redshift, transition frequency $f_t$ and initial eccentricity $e_t$ are then simply computed by calculating the spectrum at a rescaled frequency $f_t f_0^{-5/3} e_t^{-1/3}$ and then shifting it with frequency mass and redshift as indicated in Eq. (1).

Anchoring the model before astrophysical models. Although no sub-parsec SMBHBs emitting in the PTA frequency range have been unambiguously identified to date, their cosmic merger rate can be connected to the merger rate of their host galaxies. The procedure has been extensively described in ref. 7. The galaxy merger rate can be estimated directly from observations via

$$d^3N_{\text{gal}} / dM d\text{merger} = \phi(M, z) F(z, M_0, q) dt \int M_0^{10} \left( \frac{M}{M_0} \right)^{\gamma} dM_0$$

(3)

where $M_0$ is the galaxy mass; $\phi(M, z) = (dN/d\log_{10} M_0)$ is the galaxy mass function measured at redshift $z$, $F(z, M_0, q) = (d^3N/dM_0dq)_\text{merger}$, for every $M_0$ and $z$, denotes the fraction of galaxies paired with a companion galaxy with mass ratio between $q$ and $q + \delta q$. For each pair $z$, $q$ and $M_0$, $q$ is the merger timescale of the pair as a function of the relevant parameters. We construct a library of galaxy merger rates by combining four measurements of the galaxy mass function $\phi(M_0, z)$, four estimates of the close pair fraction $F(z, M_0, q)$ and two estimates of the merger timescale $\gamma$. For each of the galaxy mass functions and pair fractions, we consider three estimates given by the best fit and the two boundaries of the $1 \sigma$ confidence interval reported by the authors. We therefore have $12 \times 12 \times 2 = 288$ galaxy merger rates. Each merging galaxy pair is assigned SMBHBs with masses drawn from 14 different SMBHB–galaxy relations found in the literature, for more details see Supplementary Table 3. SMBHBs are assumed to merge in coincidence with the host galaxies (i.e., no stalling or extra delays), but can accrete either before or after merger according to the three different prescriptions described in ref. 52. This gives a total of $14 \times 3 = 42$ distinctive SMBHB populations for a given galaxy merger model. We combine the 288 galaxy merger rates as per Eq. (3) and the 42 SMBH masses assigned via using Supplementary Table 3, plus accretion prescriptions into a grand total of 12,096 SMBHB population models. Given the uncertainties, biases, selection effects, and poor understanding on the underlying physics affecting each of the individual ingredients, we do not attempt a ranking of the models, and give each of them equal weight. The models result in an allowed SMBHB merger rate density as a function of chirp mass and redshift.

We then marginalize over mass and redshift separately to obtain the functions $dN/dM$ and $dN/dz$. We are particularly interested here in testing different SMBHB–galaxy relations. We therefore construct the function $dN/dM$ for every model $M_0$ and $z$.

We use a generic simple model for the cosmic merger rate density of SMBHBs based on an overall amplitude and two power law distributions with exponential cutoffs,

$$\frac{d^n \theta}{d \log_{10} M dz} = n_0 \left( \frac{M}{10 M_0} \right)^{-\alpha} \left( 1 + z \right)^\beta \exp \left( -\frac{z}{z_0} \right) dN dz\ \mathcal{d}$$

(2)

where $dN/dz$ is the relationship between time and redshift assuming a standard ACDM flat Universe with cosmological constant of $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The five free parameters are: $n_0$ representing the co-moving number of mergers per Mpc$^3$ per Gyr; $\sigma$ and $\alpha$ control the slope and cutoff of the chirp mass distribution respectively; $\beta$ and $z_0$ regulate the equivalent properties of the redshift distribution. Eq (2) is also used to compute the number of emitting systems per frequency resolution bin at $f > 10 \text{ Hz}$. The small number statistics of the most massive binaries determines a steepening of the GW spectrum at high frequencies, full details of the computation are found in refs. 41 and 12. The GW spectrum is therefore uniquely computed by a set of six (seven) parameters $\theta = \{n_0, \beta, z_0, \sigma, \alpha, M, c(z, \eta)\}$. The resulting distribution of characteristic amplitudes $h_0$ is consistent with that of the original models. We computed the GW background under the assumption of circular GW driven systems (i.e., $h = h_{\text{circ}}$) and compared the distributions of $h_{\text{circ}}$, i.e., the strain amplitudes at $f = 1 \text{ yr}^{-1}$. The $h_{\text{circ}}$ distributions obtained with the two techniques were found to follow each other quite closely with a difference of median values and 90% confidence regions smaller than 0.1 dex. We conclude that our analytical models provide an adequate description of the observationally inferred SMBHB merger rate and can therefore be used to constrain the properties of the cosmic

Fig. 3 Astrophysical prior on the SMBHB chirp mass and redshift distributions. Left panel: a mass density distribution $dN/dM$ of the four astrophysical priors selected in this study (see text for full description). Right panel: b redshift evolution of the SMBHB mass density for the same four models. It is noteworthy that the coloured region represent the 99% interval allowed by each model, this is why individual models can extend beyond the region associated to model ALL (which includes KH13, G09, and S16 as subsets)
SMBH population. In particular model KH13 provides an optimistic prediction of the GW background with median amplitude at $f = 1 \text{yr}^{-1}$ of $h_{150} = 1.3 \times 10^{-15}$; model G09 results in a more conservative prediction $h_{150} \approx 7 \times 10^{-16}$; model S16 result in an ultra conservative estimate with median $h_{150} \approx 4 \times 10^{-16}$; and finally the characteristic amplitude predicted by the compilation of all models (ALL) encompasses only the two orders of magnitude with median value $h_{150} \approx 8 \times 10^{-16}$.

As for the parameters defining the binary dynamics, we assume that all binaries have the same eccentricity for which we pick a flat prior in the range $10^{-5} < e < 0.999$ (see Supplementary Fig. 3). In the extended model, featuring a rescaling of the density $p_\theta$ regulating the binary hardening in the stellar phase, we assume a log flat prior for the multiplicative factor $\gamma$ in the range $0.01 < \gamma < 100$. For more detailed results of including this additional density parameter see Supplementary Table 2, Supplementary Note 1 and Supplementary Fig. 4.

**Likelihood function and hierarchical modelling.** By making use of Bayes theorem, the posterior probability distribution $p(\theta | d, H)$ of the model parameters $\theta$ inferred by the data $d$ given our model $H$ is

$$p(\theta | d, H) = \frac{p(d | \theta, H)p(\theta | H)}{Z_H}$$

(4)

where $p(\theta | H)$ is the prior knowledge of the model parameters, $p(d | \theta, H)$ is the likelihood of the data $d$ given the parameters $\theta$, and $Z_H$ is the evidence of model $H$, computed as

$$Z_H = \int p(d | \theta, H)p(\theta | H)d\theta.$$  

(5)

The evidence is the integral of the likelihood function over the multi-dimensional space of the model parameters $\theta$ weighted by the multivariate prior probability distribution of the parameters. When comparing two competitive models $A$ and $B$, the odds ratio is computed as

$$O_{AB} = \frac{Z_B P_A}{Z_A P_B} = \frac{B_A P_A}{P_B},$$

(6)

where $B_A = Z_A/Z_B$ is the Bayes factor and $P_B$ is the prior probability assigned to model $H$. When comparing the four models KH13, G09, S16 and ALL, we assign equal prior probability to each model. Therefore, in each model pair comparison, the odds ratio reduces to the Bayes factor. Above we have defined the distribution of prior parameters $p(\theta | H)$, to proceed with model comparison and parameter estimation we need to define the likelihood function $p(d | \theta, H)$. The likelihood function, $p(d | \theta, H)$, is defined following ref. 3. We take the posterior samples from the Parkes PTA analysis (courtesy of Shannon and collaborators) used to place the 95% upper limit at $\sigma / c \sim 1 \times 10^{-15}$, when a single power law background $h_\sigma = f^{−\gamma}$ is assumed. However, for our analysis we would like to convert this upper limit at $\sigma / c \sim 1 \times 10^{-15}$, where the frequency $f$ is equal to the frequency upper limit of the power spectral density as shown on the orange curve in Fig. 1. Our likelihood is constructed by multiplying all bins together, therefore the resulting overall limit from these bin-by-bin upper limits must be consistent with $h_{150} \sim 1 \times 10^{-15}$. The $f_{\text{min}}$ posterior distribution is well fitted by a Fermi function. To estimate a frequency dependent upper limit, we use Fermi function likelihoods at each frequency bin, which are then shifted, normalized and corrected to provide the correct overall upper limit. In our analysis we consider the contributions by only the first frequency bins of size $1/11$ yr$^{-1}$, as the highest frequency portion of the spectrum provides no additional constraint. We have verified that when we include additional bins the results of the analysis are unchanged. Ideally, we would take the bin-by-bin upper limits directly from the pulsar timing analysis to take account of the true shape of the posterior; however, the method we use here provides a consistent estimate for our analysis.

Having defined the population of merging binaries, the astrophysical prior and the likelihood based on the PPTA upper limit result, we use a nested sampling algorithm$^3$ to construct posterior distributions for each of the six model parameters. For the results shown here, we use $2,000$ live points and run each analysis 5 times, giving an average of around $18,000$ posterior samples.

**Data availability.** The posteriors are available from www.sr.bham.ac.uk/pta/publications/ncomms18/posteriors. The code used for the analysis in this study is available from the corresponding author on request.

Received: 17 July 2017 Accepted: 9 January 2018 Published online: 08 February 2018

**References**

1. White, S. D. M. & Rees, M. J. Core condensation in heavy haloes—a two-stage theory for galaxy formation and clustering. MNRAS 183, 341–358 (1978).

2. Begelman, M. C., Blandford, R. D. & Rees, M. J. Massive black hole binaries in active galactic nuclei. Nature 287, 307–309 (1980).

3. Sesana, A. Insights into the astrophysical and massive black hole binaries from pulsar timing observations. Class. Quantum Gravity 30, 224014 (2013).

4. Shannon, R. M. et al. Gravitational waves from binary supermassive black holes missing in pulsar observations. Science 349, 1522–1525 (2015).

5. Lentati, L. et al. European Pulsar Timing Array limits on an isotropic stochastic gravitational-wave background. MNRAS 453, 2576–2598 (2015).

6. Arzoumanian, Z. et al. The NANOGrav nine-year data set: limits on the isotropic stochastic gravitational wave background. Astrophys. J. 821, 13 (2016).

7. Verbiest, J. P. W. et al. The international pulsar timing array: first data release. MNRAS 458, 1267–1288 (2016).

8. Hobbs, G. & Dai, S. A review of pulsar timing array gravitational wave research. Preprint available at: https://arxiv.org/abs/1707.01615 (2017).

9. Sesana, A. Systematic investigation of the expected gravitational wave signal from supermassive black hole binaries in the pulsar timing band. MNRAS 433, L1–L5 (2013).

10. Sesana, A., Shankar, F., Bernardi, M. & Sheth, R. K. Selection bias in dynamically measured supermassive black hole samples: consequences for pulsar timing arrays. MNRAS 463, L6–L11 (2016).

11. Hernquist, L. An analytical model for spherical galaxies and bulges. Astrophys. J. 356, 359–364 (1990).

12. Chen, S., Sesana, A. & Del Pozzo, W. Efficient computation of the gravitational wave spectrum emitted by eccentric massive black hole binaries in stellar environments. MNRAS 470, 1734–1749 (2017).

13. Simon, J. & Burke-Spolaor, S. Constraints on black hole/host galaxy co-evolution and binary stalling using pulsar timing arrays. Astrophys. J. 826, 11 (2016).

14. Kormendy, J. & Ho, L. C. Coevolution (or not) of supermassive black holes and host galaxies. ARA&A 41, 511–653 (2013).

15. Gültekin, K. et al. The M-σ and M-L relations in galactic bulges, and determinations of their intrinsic scatter. Astrophys. J. 698, 199–221 (2009).

16. Shankar, F. et al. Selection bias in dynamically measured supermassive black hole samples: its consequences and the quest for the most fundamental relation. MNRAS 466, 3119–3142 (2016).

17. McWilliams, S. T., Ostriker, J. P. & Pretorius, F. Gravitational waves and dynamical galactic environments. MNRAS 463, 3131–3157 (2017).

18. Kang, R. E. & Rafferty, A. E. Bayes factors. J. Am. Stat. Assoc. 90, 773–795 (1995).

19. Taylor, S. R. et al. Are we there yet? time to detection of nanohertz gravitational waves based on pulsar-timing array limits. Astrophys. J. 819, L6 (2016).

20. Booth, R. S. de Blok, W. J. G., Jonas, J. L. & Fanaroff, B. MeetKAT key project science, specifications, and proposals. Preprint available at: https://arxiv.org/abs/0910.2955 (2009).

21. Nan, R. et al. The five-hundred aperture spherical radio telescope (fast) project. Int. J. Mod. Phys. D. 20, 989–1024 (2011).

22. Dewdney, P., Hall, R. J., Schlitz, R. T. & Lazio, T. J. W. The square kilometre Array. IEEE Proc. 97, 1482–1496 (2009).

23. Bonetti, M., Sesana, A., Barausse, E. & Haardt, F. Post-Newtonian evolution of massive black hole triples in galactic nuclei—III. A robust lower limit to the nHz stochastic background of gravitational waves. Preprint available at: https://arxiv.org/abs/1709.06095 (2017).

24. McConnell, N. J. et al. Two ten-billion-solar-mass black holes at the centres of giant elliptical galaxies. Nature 480, 215–218 (2011).

25. Häflecker-Larrondo, J., Fabian, A. C., Edge, A. C. & Hogan, M. T. On the hunt for ultramassive black holes in brightest cluster galaxies. MNRAS 424, 224–232 (2014).

26. McConnell, N. J. & Ma, C.-P. Revisiting the scaling relations of black hole masses and host galaxy properties. Astrophys. J. 764, 184 (2013).

27. Ruszkowski, A. & Merritt, D. Evolution of massive black hole binaries in rotating stellar nuclei: implications for gravitational wave detection. Preprint available at: https://arxiv.org/abs/1606.07484 (2016).

28. Gilmozzi, R. & Spyrisoni, J. The European Extremely Large Telescope (E-ELT). The Messenger, 127, 11–19 (2007).

29. Sanders, G. H. The thirty meter telescope (TMT): an international observatory. J. Astrophys. Astron. 34, 81–86 (2013).
33. Do, T. et al. Prospects for measuring supermassive black hole masses with future extremely large telescopes. *Astron. J.* **147**, 93 (2014).
34. Lotz, J. M. et al. The major and minor galaxy merger rates at z < 1.5. *Astrophys. J.* **742**, 103 (2011).
35. Mundy, C. J. et al. A consistent measure of the merger histories of massive galaxies using close-pair statistics—I. Major mergers at z < 3.5. *MNRAS* **470**, 3507–3531 (2017).
36. Phinney, E. A. A practical theorem on gravitational wave backgrounds. Preprint available at: https://arxiv.org/abs/astro-ph/0108028 (2001).
37. Quinlan, G. D. The dynamical evolution of massive black hole binaries I. Hardening in a fixed stellar background. *New Astron.* **1**, 35–56 (1996).
38. Cuadra, J., Armitage, P. J., Alexander, R. D. & Begelman, M. C. Massive black hole binary mergers within subparsec scale gas discs. *MNRAS* **393**, 1423–1432 (2009).
39. Sesana, A. & Khan, F. M. Scattering experiments meet N-body—1. A practical recipe for the evolution of massive black hole binaries in stellar environments. *MNRAS* **454**, L66–L70 (2015).
40. Vasiliou, E., Antonini, F. & Merritt, D. The final-parsec problem in the collisionless limit. *Astrophys. J.* **810**, 49 (2015).
41. Sesana, A., Vecchio, A. & Colacino, C. N. The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays. *MNRAS* **390**, 192–209 (2008).
42. Ilbert, O. et al. Mass assembly in quiescent and star-forming galaxies since z ~ 4 from UltraVISTA. *A&A* **556**, A55 (2013).
43. Muzzin, A. et al. The evolution of the stellar mass functions of star-forming and quiescent galaxies to z = 4 from the COSMOS/UltraVISTA survey. *Astrophys. J.* **777**, 18 (2013).
44. Tomczak, A. R. et al. Galaxy stellar mass functions from ZFOURGE/CANDELS: an excess of low-mass galaxies since z = 2 and the rapid buildup of quiescent galaxies. *Astrophys. J.* **783**, 85 (2014).
45. Bernardi, M. et al. The massive end of the luminosity and stellar mass functions and clustering from CMASS to SDSS: evidence for and against passive evolution. *MNRAS* **455**, 4122–4135 (2016).
46. Bundy, K. et al. The greater impact of mergers on the growth of massive galaxies: implications for mass assembly and Evolution since z sim 1. *Astrophys. J.* **697**, 1369–1383 (2009).
47. de Ravel, L. et al. The VIMOS VLT deep survey, evolution of the major merger rate since z from spectroscopically confirmed galaxy pairs. *Astron. Astrophys.* **498**, 379–397 (2009).
48. López-Sanjuan, C. et al. The dominant role of mergers in the size evolution of massive early-type galaxies since z. *Astron. Astrophys.* **548**, A7 (2012).
49. Xu, C. K. et al. Major-merger galaxy pairs in the COSMOS field—mass-dependent merger rate evolution since z = 1. *Astrophys. J.* **747**, 85 (2012).
50. Kitzbichler, M. G. & White, S. D. M. A calibration of the relation between the abundance of close galaxy pairs and the rate of galaxy mergers. *Mon. Not. R. Astron. Soc.* **391**, 1489–1498 (2008).
51. Lotz, J. M., Jonsson, P., Cox, T. J. & Primack, J. R. The effect of mass ratio on the morphology and time-scales of disc galaxy mergers. *Mon. Not. R. Astron. Soc.* **404**, 575–589 (May 2010).
52. Sesana, A., Vecchio, A. & Volonteri, M. Gravitational waves from resolvable massive black hole binary systems and observations with pulsar timing arrays. *MNRAS* **394**, 2255–2265 (2009).
53. Chen, S., Middleton, H., Sesana, A., Del Pozzo, W. & Vecchio, A. Probing the assembly history and dynamical evolution of massive black hole binaries with pulsar timing arrays. *MNRAS* **468**, 404–417 (2017).
54. Skillings, J. Nested sampling. In: *American Institute of Physics Conference Series*, Vol. 735 (eds Fischer, R., Preuss, R. & Toussaint, U.-V.) 395–405 (2004).
55. Del Pozzo, W. & Veitch, J. CPNest: Parallel nested sampling in python. *GitHub* https://github.com/johnveitch/cpnest (2015).

Acknowledgements

H.M. and A.V. acknowledge the support by the Science and Technology Facilities Council (STFC). S.C. acknowledges the support of the University of Birmingham via the AE Hills scholarship. A.S. is supported by a URF of the Royal Society.

Author contributions

All the authors have contributed to this work.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-02916-7.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s) 2018