Observing the dynamics of super-massive black hole binaries with Pulsar Timing Arrays

C. M. F. Mingarelli,¹ K. Grover,¹ T. Sidery,¹ R. J. E. Smith,¹,² and A. Vecchio¹
¹School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
²Perimeter Institute for Theoretical Physics, Waterloo, Ontario N2L 2Y5, Canada
(Dated: May 2, 2014)

Pulsar Timing Arrays are a prime tool to study unexplored astrophysical regimes with gravitational waves. Here we show that the detection of gravitational radiation from individually resolvable supermassive black hole binary systems can yield direct information about the masses and spins of the black holes, provided that the gravitational-wave induced timing fluctuations both at the pulsar and at the Earth are detected. This in turn provides a map of the non-linear dynamics of the gravitational field and a new avenue to tackle open problems in astrophysics connected to the formation and evolution of super-massive black holes. We discuss the potential, the challenges and the limitations of these observations.

PACS numbers: 04.80.Nn, 04.25.dg, 95.85.Sz, 97.80.-d 97.60.Gb 04.30.-w

Introduction—

Gravitational waves (GWs) provide a new means for studying black holes and addressing open questions in astrophysics and fundamental physics: from their formation, evolution and demographics, to the assembly history of galactic structures and the dynamical behaviour of gravitational fields in the strong non-linear regime. Specifically, GW observations through a network of radio pulsars used as ultra-stable clocks – Pulsar Timing Arrays (PTAs) [1–3] – represent the only direct observational avenue for the study of supermassive black hole (SMBHB) systems in the \( \sim 10^{8} \) – \( 10^{9} \) \( M_{\odot} \) mass range, with orbital periods between \( \sim 1 \) month and a few years, see e.g. [6,7] and references therein. Ongoing observations [4, 5, 8, 9] and future instruments, e.g. the Square Kilometre Array [10], are expected to yield the necessary timing precision [11,12] to observe the diffuse GW background. This background is likely dominated by the incoherent superposition of radiation from the cosmic population of massive black holes [13–21] and within it, we expect a handful of sources that are sufficiently close, massive and high-frequency to be individually resolvable [22–30].

Massive black hole formation and evolution scenarios [31–34] predict the existence of a large number of SMBHBs. Furthermore, SMBHBs are expected to be (possibly rapidly) spinning [35,36]. In fact the dynamics of such systems – which according to general relativity are entirely determined by the masses and spins of the black holes [37] – leave a direct imprint on the emitted gravitational waveforms. From these, one could measure SMBHB masses and their distribution, yielding new insights into the assembly of galaxies and the dynamical processes in galactic nuclei [26]. Moreover, measuring the magnitude and/or orientation of spins in SMBHBs would provide new information on the role of accretion processes [23, 32]. Finally, detections of SMBHBs could allow us to probe general relativistic effects in the non-linear regime in an astrophysical context not directly accessible by other means, see [43] and references therein.

The observation of GWs with PTAs relies on the detection of the small deviation induced by gravitational radiation in the times of arrival (TOAs) of radio pulses from millisecond pulsars that function as ultra-stable reference clocks. This deviation, called the residual, is the difference between the expected (without GW contribution) and actual TOAs once all the other physical effects are taken into account. The imprint of GWs on the timing residuals is the result of how the propagation of radio waves is affected by the GW-induced space-time perturbations along the travel path. It is a linear combination of the GW perturbation at the time when the radiation transits at a pulsar, the so-called “pulsar term”, and then when it passes at the radio receiver, the “Earth term” [1–3]. The two terms reflect the state of a GW source at two different times of its evolution separated by \( \tau \equiv (1 + \hat{\Omega} \cdot \hat{p}) L_{\nu} \sim 3.3 \times 10^{3} (1 + \hat{\Omega} \cdot \hat{p}) (L_{\nu}/1 \text{ kpc}) \text{ yr} \) where \( \hat{\Omega} \) and \( \hat{p} \) are the unit vectors that identify the GW propagation direction and the pulsar sky location at a distance \( L_{\nu} \) from the Earth, respectively, see e.g. [24]. [We use geometrical units in which \( G = c = 1 \).] In a network (array) of pulsars all the perturbations at the Earth add coherently and therefore boost the signal-to-noise ratio (S/N) of the signal. Each pulsar term is at a slightly different frequency since the orbital period of the binary evolves over the time \( \tau \).

Measuring the key physics of SMBHBs is hampered by the short (typically \( T = 10 \) yr) observation time compared to the typical orbital evolution timescale \( f/\dot{f} = 1.6 \times 10^{3} (M/10^{9} M_{\odot})^{-5/3} (f/50 \text{ nHz})^{-8/3} \) yr of binaries that are still in the weak field adiabatic inspiral regime, with an orbital velocity \( v = 0.12 (M/10^{9} M_{\odot})^{1/3} (f/50 \text{ nHz})^{1/3} \) [44]. Here \( M = m_{1} + m_{2}, \mu = m_{1} m_{2}/M \) and \( M = M_{\odot}^{2/3} \mu^{3/5} \) are the total, reduced and chirp mass, respectively, of a binary with component masses \( m_{1,2}, \) and \( f \) is the GW emission frequency at the leading quadrupole order. The chirp mass determines the frequency evolution at the leading Newtonian order. In the post-Newtonian (pN) expansion of the binary evolution [45] in terms of \( v \ll 1 \), the second mass parameter enters at pN order (\( \mathcal{O}(v^{2}) \) correction); spins
contribute at \( p^{1.5} \)N order and above (\( O(\nu^3) \)) causing the orbital plane to precess through spin-orbit coupling, at leading order. These contributions are therefore seemingly out of observational reach.

The GW effect at the pulsar – the pulsar term – may be detectable in future surveys, and for selected pulsars their distance could be determined to sub-parsec precision \([27, 46, 47]\). If this is indeed the case, it opens the opportunity to coherently connect the signal observed at the Earth and at pulsars, therefore providing snapshots of the binary evolution over \( \sim 10^3 \) yr. These observations would drastically change the ability to infer SMBHB dynamics, and study the relevant astrophysical process and fundamental physics.

In this Letter we show that for SMBHBs at the high end of the mass and frequency spectrum observable by PTAs, say \( m_{1,2} = 10^9 M_\odot \) and \( f = 10^{-7} \) Hz, the observations of a source still in the weak-field regime become sensitive to post-Newtonian contributions up to \( p^{1.5}N \), including spin-orbit effects, if both the pulsar and Earth term can be detected. This in principle enables the measurement of the two mass parameters and a combination of the spin’s magnitude and relative orientation. We also show that the Earth-term only can still be sensitive to spin-orbit coupling due to geometrical effects produced by precession. We discuss the key factors that enable these measurements, and future observational prospects and limitations.

**Signals from SMBHBs**— Consider a radio pulsar emitting radio pulses at frequency \( \nu_0 \) in the source rest-frame. GWs modify the rate at which the radio signals are received at the Earth \([1–3]\), inducing a relative frequency shift \( \delta \nu(t)/\nu_0 = h(t - \tau) - h(t) \), where \( h(t) \) is the GW strain. The quantities that are actually produced at the end of the data reduction process of a PTA are the timing residuals, \( \int dt \; \delta t(t)/\nu_0 \), although without loss of generality, we will base the discussion on \( h(t) \). The perturbation induced by GWs is repeated twice, and carries information about the source at time \( t \), the “Earth term”, and at past time \( t - \tau \), the “pulsar term”.

We model the radiation from a SMBHB using the so-called restricted pN approximation, in which pN corrections are included only in the phase and the amplitude is retained at the leading Newtonian order, but we include the leading order modulation effects produced by spin-orbit coupling. The strain is given by

\[
\begin{align*}
    h(t) &= -A_{gw}(t)A_p(t) \cos(\Phi(t) + \varphi_p(t) + \varphi_T(t)) \tag{1}
\end{align*}
\]

where \( A_{gw}(t) = 2[\pi f(t)]^{3/2} M^{5/3} \) is the Newtonian order GW amplitude, \( \Phi(t) \) is the GW phase, see e.g. Eq. (232, 234) in \([45]\) and Eq. (8.4) in \([48]\), and \( D \) is the distance to the GW source. \( A_p(t) \) and \( \varphi_p(t) \) are the time-dependent polarisation amplitude and phase and \( \varphi_T(t) \) is an additional phase term, analogous to Thomas precession, see Eq. (29) in \([49]\).

The physical parameters leave different observational signatures in the GW strain \( h(t) \) and are therefore found in the TOA residuals. At the leading Newtonian order, \( \mathcal{M} \) drives the frequency and therefore the phase \( \Phi(t) \) evolution, with the second independent mass parameter entering from the \( p^2 \)N onwards. SMBHBs are believed to be rapidly spinning, and the spins are responsible for three distinctive imprints in the waveform: (i) they alter the phase evolution through spin-orbit coupling and spin-spin coupling at \( p^{1.5}N \) and \( p^2 \)N order, respectively \([50]\), (ii) they cause the orbital plane to precess due to (at lowest order) spin-orbit coupling and therefore induce amplitude and phase modulations in the waveform through \( A_p(t) \) and \( \varphi_p(t) \); and (iii) through orbital precession they introduce an additional secular contribution \( \varphi_T(t) \) to the waveform phase. Astrophysically we expect PTAs to detect SMBHBs of comparable component masses \([24]\). We therefore model the spin-orbit precession using the simple precession approximation \([49]\), which formally applies when \( m_1 = m_2 \), or when one of the two spins is negligible with respect to the other. Let \( S_{1,2} \) and \( L \) be the black holes’ spins and the orbital angular momentum, respectively. Then both \( S = S_1 + S_2 \) and \( L \) precess around the (essentially) constant direction of the total angular momentum, \( J = S + L \), at the same rate \( d\alpha/dt = \pi^2 (2 + 3m_2/(2m_1)) (|L + S|) f^2(t)/M [49] \), where \( \alpha \) is the precession angle, while preserving the angle of the precession cone, \( \lambda_L \), see Fig. 4 of Ref. \([49]\). This approximation is adequate to conceptually explore these effects, however in the case of real observations, one will need to consider the exact expressions \([51]\).

The detection and particularly the measurement of the aforementioned parameters relies on coherently matching the signal with a template that faithfully reproduces its amplitude and, importantly, its phase evolution. We therefore consider the contribution to the total number of wave cycles a proxy for the significance of a specific parameter. Individual terms that contribute \( \sim 1 \) GW cycle or more mean that the effect is in principle detectable, hence one can infer information about the associated parameter(s). We show that information about the parameters can only be inferred for SMBHBs at the high end of the mass spectrum and PTA observational frequency range. Having a sufficiently high-mass and high-frequency GW source is also essential to ensure sufficient frequency evolution over the time \( \tau \), so that the Earth and pulsar term are clearly separated in frequency space cf. Table \([1]\). We therefore take fiducial source parameters of \( m_1 = m_2 = 10^9 M_\odot \), frequency at the Earth at the beginning of the observation \( f_E = 10^{-7} \) Hz and an observational time \( T = 10 \) years to illustrate the main results. We provide scaling relations as a function of the relevant quantities, allowing the reader to rescale the results for different astrophysical and/or observational values.

**Observations using the Earth-term only** — We start by considering analyses that rely only on the Earth-term contribution to the residuals, as done in Ref. \([25, 52]\). The case of a coherent analysis based both on the Earth- and pulsar-term, introduced in Ref. \([22]\), is discussed later in this Letter. Table \([1]\) shows that, in general, the frequency...
change over 10 yr is small compared to the frequency bin width, 3.2/(10 yr/T) nHz [24, 27]. The observed signal is effectively monochromatic, making the dynamics of the system impossible to infer. However, the presence of spins affects the waveform not only through the GW phase evolution, but also via the modulations of $A_p(t)$ and $\varphi_p(t)$ that are periodic over the precession period, and also introduces the secular contribution $\varphi_T(t)$. For $m_{1,2} = 10^9 M_\odot$ and $f_E = 10^{-7}$ Hz the orbital angular momentum precesses by $\Delta \alpha = 2$ rad (for dimensionless spin parameter $a \equiv S/M^2 = 0.1$) and $\Delta \alpha = 3$ rad (for $a = 0.98$). Furthermore, and therefore the additional modulation effect on $A_p(t)$ and $\varphi_p(t)$ is small, and likely undetectable. However, the overall change of $\varphi_T(t)$ over 10 yrs could be appreciable: the average contribution for each precession cycle of this additional phase term is $\langle \Delta \varphi_T \rangle = 4\pi$ or $4\pi (1 - \cos \lambda)$, depending on whether $\Omega$ lies inside or outside the precession cone, respectively [49]. If $\Omega$ lies inside the precession cone, and given that the observation will cover between a third and a half of a full precession cycle, then $\langle \Delta \varphi_T \rangle \sim \pi$, which could surely indicate the presence of spins. On the other hand, the precession cone will be small in general since $|S/L| \sim a v (M/\mu) \lesssim 0.1 a (M/\mu) (M/10^9 M_\odot)^{1/3} (f/100 \text{ nHz})^{1/3}$, therefore the likelihood of $\Omega$ lying inside the precession cone is small, assuming an isotropic distribution and orientation of sources. In this case the Thomas precession contribution (per precession cycle) is suppressed by a factor $\langle 1 - \cos \lambda \rangle \sim \lambda^2/2 \sim 5 \times 10^{-3} a^2 (M/\mu)^2 (M/10^9 M_\odot)^{2/3} (f/100 \text{ nHz})^{2/3}$, which sets a negligible contribution $\Delta \varphi_T(t) \ll 1$. However, unlikely, spins may still introduce observable effects that need to be taken into account.

Measuring SMBHB evolution using the Earth and pulsar term—With more sensitive observations and the increasing possibility of precisely determining $L_p$, see e.g. [47], the prospect of also observing the contribution from the pulsar term from one or more pulsars becomes more realistic. We show below that if at least one of the pulsar terms can be observed together with the Earth-term, this opens opportunities to study the dynamical evolution of SMBHBs and, in principle, to measure their masses and spins. This is a straightforward consequence of the fact that PTAs become sensitive to $\sim 10^3$ yr of SMBHB evolutionary history, in “snippets” of length $T \ll L_p$ that can be coherently concatenated.

The signal from each pulsar term will be at a S/N which is significantly smaller than the Earth-term by a factor $\sqrt{N_p}$, where $N_p$ is the number of pulsars that effectively contribute to the S/N of the array. For example, if the Earth-term yields an S/N of $\sim 30 \sqrt{N_p}/20$, then each individual pulsar term would give an S/N $\sim 8$. The possibility of coherently connecting the Earth-term signal with each pulsar term becomes therefore a question of S/N, prior information about the pulsar-Earth baseline and how accurately the SMBHB location in the sky can be reconstructed, as part of a “global fit”, e.g. [27]. Assuming for simplicity that the uncertainties on $L_p$ and $\Omega$ are uncorrelated, this requires that the distance to the pulsar and the location of the GW source are known with errors $\lesssim 0.01(100 \text{ nHz}/f)$ pc and $\lesssim 3(100 \text{ nHz}/f)(1 \text{ kpc}/L_p)$ arcsec, respectively. These are very stringent constraints [24, 28, 47], and a detailed analysis is needed in order to assess the feasibility of reaching this precision. Clearly if an electromagnetic counterpart to the GW source were to be found [53, 54], it would enable the identification of the source location in the sky, making the latter constraint unnecessary.

We can now consider the contribution from the different terms in the pN expansion to the total number of cycles in observations that cover the GW source evolution over the time $\tau$ that are encoded in the simultaneous analysis of the Earth and pulsar terms. The results are summarised in Table II for selected values of $m_{1,2}$ and $f_E$ and for a fiducial value $\tau = 1$ kpc. The wavecycle contributions from the spin-orbit parameter are normalised to $\beta = (1/12) \sum_{i=1}^{113} (m_i/M)^2 + 75\eta \hat{L} \cdot \hat{S}_i$, which has a maximum value of 7.8. Contributions from the pN order spin-spin terms are negligible. The results clearly show that despite the fact that the source is in the weak field regime the extended Earth-pulsar baseline requires the p$^{1.5}$N, and in some rare cases the p$^2$N contribution, to accurately (i.e. within $\sim 1$ GW cycle) reproduce the full phase evolution.

For $m_{1,2} = 10^9 M_\odot$ and $f_E = 10^{-7}$ Hz there is a total of 4305 GW cycles over a 1 kpc light travel time evolution, with the majority (4267) accounted for by the leading order Newtonian term, providing information about the chirp mass, and tens of cycles due to the p$^1$N and p$^{1.5}$N terms (77 and 45, respectively), that provide information about a second independent mass parameter. Spins contribute to phasing at p$^{1.5}$N with $\sim 3/3$ cycles. Therefore their total contribution is smaller than the p$^{1.5}$N mass contribution by a factor between a few and $\sim 10$. The additional Thomas precession phase contribution may become comparable to the p$^1$N mass contribution in some cases. In fact, for $a = 0.1(0.98)$ the binary undergoes 24 (34) precession cycles. This corresponds to a total Thomas precession phase contribution of 306 (426) rad if $\Omega$ lies outside the precession cone.

The modulations of $A_p(t)$ and $\varphi_p(t)$ are characterised by a small $\lambda_1$, because for most of the inspiral $S \ll L$, and are likely to leave a smaller imprint on the waveform than those discussed so far. We can indeed estimate the importance of this effect for the most favourable parameter combinations. The value of $\varphi_p(t)$ oscillates over time with an amplitude which depends on the time to coalescence, $S$, $L$, $\Omega$, and $\hat{p}$. We choose the orientation of $\hat{S}$ such that $\lambda_1$ is maximised, and we vary $\Omega$ and $\hat{p}$, each of which is drawn from a uniform distribution on the two-sphere.

In Figure 1 we show that for rapidly spinning ($a = 0.98$)
TABLE I: Frequency change $\Delta f$, total number of GW cycles and individual contributions from the leading order terms in the pN expansion over the two relevant time scales – a 10 yr period starting at the Earth and the time period $L_p(1 + \Omega \cdot \hat{p})$ between the Earth and pulsar term (hence the negative sign) – for selected values of $m_{1,2}$ and $f_E$.

| $m_1 (M_\odot)$ | $m_2 (M_\odot)$ | $f_E$ (nHz) | $\frac{v/c \times 10^{-2}}{\text{timespan}}$ | $\Delta f$ (nHz) | Total pN | $p^{1,N}$ | $p^{1.5,N}$ spin-orbit/$\beta$ | $p^{2,N}$ |
|-----------------|-----------------|-------------|-----------------------------------------|-----------------|----------|----------|-------------------------------|----------|
| $10^9$          | $10^9$          | 14.6        | 10 yr                                   | 3.22            | 32.1     | 31.7     | 0.9 -0.7 0.06 0.04             | 0.04     |
|                 |                 | 9.6         | -1 kpc                                  | 71.2            | 4305.1   | 4267.8   | 77.3 -45.8 3.6 2.2            | 2.2      |
|                 |                 | 11.6        | 10 yr                                   | 0.24            | 15.8     | 15.7     | 0.3 -0.2 0.01 < 0.01          | < 0.01   |
|                 |                 | 9.4         | -1 kpc                                  | 23.1            | 3533.1   | 3504.8   | 53.5 -28.7 2.3 1.2            | 1.2      |
| $10^8$          | $10^8$          | 6.8         | 10 yr                                   | 0.07            | 31.6     | 31.4     | 0.2 -0.07 < 0.01 < 0.01       | < 0.01   |
|                 |                 | 6.4         | -1 kpc                                  | 15.8            | 9396.3   | 9355.7   | 58.3 -19.9 1.6 0.5            | 0.5      |
|                 |                 | 5.4         | 10 yr                                   | 0.005           | 15.8     | 15.7     | 0.06 -0.02 < 0.01 < 0.01      | < 0.01   |
|                 |                 | 5.3         | -1 kpc                                  | 1.62            | 5061.4   | 5045.8   | 20.8 -5.8 0.5 0.1             | 0.1      |

SMBHBs this effect could introduce modulations larger than $\pi/2$ in $\varphi_p(t)$ over 30% of the parameter space of possible $\hat{\Omega}$ and $\hat{p}$ geometries. The amplitude would correspondingly change over the same portion of the parameter space by at most 60% with respect to its unmodulated value. Since these effects are modulated, they will not be easily identifiable.

Conclusions—We have established that the coherent observation of both the Earth and pulsar term provides information about the dynamical evolution of a GW source. The question now is whether they can be unambiguously identified. A rigorous analysis would require extensive simulations based on the actual analysis of synthetic data sets. We can however gain the key information with a much simpler order of magnitude calculation. The phase (or number of cycles) error scales as $\sim 1/(S/N)$. Assuming $S/N \sim 40$ means that the total number of wave cycles over the Earth-pulsar baseline can be determined with an error $\sim 4300/40 \sim 100$ wave cycles. This is comparable to the $p^1N$ contribution to the GW phase and, in very favourable circumstances, to the Thomas precession phase contribution, and larger by a factor of a few or more than all the other contributions. It may therefore be possible to measure the chirp mass and, say, the symmetric mass ratio of a SMBHB, and possibly a combination of the spin parameters. Effects due to the $p^{1.5}N$ and higher phase terms are likely to remain unobservable, as well as amplitude and phase modulations. Correlations between the parameters, in particular masses and spins, will further degrade the measurements. The details will depend on the actual $S/N$ of the observations, the GW source parameters, and the accuracy with which the source location and the pulsar distance can be determined. We plan to explore these issues in detail in a future study.

Acknowledgements—We would like to thank the referees for their suggestions and comments. Furthermore, we would like to thank I. Mandel, K. J. Lee, A. Sesana, J. Verbiest and our colleagues of the European Pulsar Timing Array Collaboration for many useful discussions and comments. This work has been supported by the UK Science and Technology Facilities Council. RJES acknowledges the support of a Perimeter Institute Visiting Graduate Fellowship. CMFM acknowledges the support of the Royal Astronomical Society and the Institute of Physics.

[1] F. B. Estabrook and H. D. Wahlquist, General Relativity and Gravitation, 6, 439 (1975).
[2] M. V. Sazhin, Sov. Astron., 22, 36 (1978).
[3] S. Detweiler, Astrophys. J., 234, 1100 (1979).
[4] J. P. W. Verbiest et al., Class. Quant. Grav., 27, 084015 (2010).
[5] R. D. Ferdman et al., Class. Quant. Grav., 27, 084014 (2010).
[6] A. Sesana, Adv. Astron. 2012, 805402 (2012) [arXiv:1110.6445[astro-ph.CO]].
