Post-Newtonian-accurate pulsar timing array signals induced by inspiralling eccentric binaries: accuracy and computational cost

Abhimanyu Susobhanan*

Center for Gravitation, Cosmology, and Astrophysics,
University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA

(Dated: October 21, 2022)

Pulsar Timing Array (PTA) experiments are expected to be sensitive to gravitational waves (GWs) emitted by individual supermassive black hole binaries (SMBHBs) inspiralling along eccentric orbits. We compare the computational cost of different methods of computing the PTA signals induced by relativistic eccentric SMBHBs, namely approximate analytic expressions, Fourier series expansion, post-circular expansion, and numerical integration. We show that the fastest method for evaluating PTA signals is by using the approximate analytic expressions, providing up to a ~50 times improvement in computational performance over the alternative methods. We investigate the accuracy of the approximate analytic expressions by employing a mismatch metric valid for PTA signals. We show that this method is accurate within the region of the binary parameter space that is of interest to PTA experiments. We introduce a spline-based method for further accelerating the PTA signal evaluations for narrow-band PTA datasets. These results are crucial for searching for eccentric SMBHBs in large PTA datasets. We have implemented these results in the G\textit{Wecc} package and can be readily accessed from the popular \textit{ENTERPRISE} package to search for such signals in PTA datasets.

I. INTRODUCTION

Routine detections of gravitational waves (GWs) from stellar mass compact object merger events by ground-based GW detectors have ushered in the era of GW astronomy [1–3]. Pulsar Timing Arrays [PTAs: 4] are experiments that aim to detect GWs in the nanohertz frequency range by routinely monitoring ensembles of millisecond pulsars using some of the world’s most sensitive radio telescopes. Ongoing PTA campaigns include the North American Nanohertz Observatory for Gravitational waves [NANOGrav: 5], the European Pulsar Timing Array [EPTA: 6], the Parkes Pulsar Timing Array [PPTA: 7], the Indian Pulsar Timing Array [InPTA: 8], MeerTime [9], and the Chinese Pulsar Timing Array [CPTA: 10]. The International Pulsar Timing Array consortium

* abhimanyu.susobhanan@nanograv.org
[IPTA: 11] aims to combine data and resources from different PTA campaigns to accelerate the discovery of nanohertz GWs and improve the prospects of post-discovery science. PTA experiments have grown in sensitivity over the years, and are expected to achieve their first detection in the near future [12–15].

The most prominent sources of nanohertz GWs are expected to be supermassive black hole binaries (SMBHBs), usually hosted by active galactic nuclei (AGNs) [16]. The first PTA detection is expected to be that of a stochastic GW background (GWB) formed via the incoherent addition of GWs emitted by an ensemble of unresolved SMBHBs, followed by the detection of individual SMBHBs that stand out above the GWB [17–19]. Several promising SMBHB candidates have been identified through electromagnetic observations of AGNs [e.g. 20–23], and PTA experiments have already put increasingly stringent constraints on the presence of SMBHB signals in their datasets [24–27].

SMBHBs are believed to form through galaxy mergers, where the central black holes of the merging galaxies sink to the center of the merger remnant, eventually forming a bound system [28]. Such binary systems shrink due to energy and angular momentum exchange with the surrounding stars and gas until the orbital evolution is dominated by GW emission [29], and can retain significant eccentricities as they enter the PTA frequency band [e.g. 30, 31]. The SMBHB candidate OJ 287 is believed to host a binary system with eccentricity $\sim 0.6$ [21]. It is therefore desirable to search for signals induced by eccentric SMBHB systems in PTA datasets.

The PTA responses to GW signals (known as PTA signals) due to inspiralling eccentric SMBHB systems were modeled by Refs. [24, 32, 33] (see Section II for details). This usually involves modeling the relativistic motion of the binary system using the post-Newtonian (PN) formalism, where general relativistic corrections to Newtonian dynamics are expressed in powers of $(v/c)^2 \sim GM/(c^2r)$, and where $M$ is the total mass of the binary, $r$ is the relative separation, and $v$ is the relative speed [34]. The GW strain amplitudes in the two orthogonal polarizations $h_{+,x}$ can then be expressed in terms of the polar coordinates $r$ and $\phi$ in the orbital plane and their time derivatives. Finally, the PTA signal $R(t)$ involves the time integrals of $h_{+,x}$.

A significant challenge in searching for such signals in PTA datasets is posed by the cost of computing the PTA signal given a set of pulse times of arrival (TOAs) of an ensemble of pulsars. (See Refs [32, 35] for estimates of the sensitivity degradation experienced when using the computationally inexpensive circular PTA signals [24] to search for eccentric sources.) This computational cost is mainly incurred in two stages: (a) solving the orbital evolution, and (b) computing the PTA signal as a function of the orbital variables. Refs [33, 36] provided an analytic solution to the
quadrupolar-order orbital evolution equations governing the motion of inspiralling non-spinning eccentric binaries, thereby mitigating the computational cost of solving the orbital evolution of such systems. A few different approaches have been presented in the literature for computing the PTA signals given the orbital variables as a function of time, including an approximate analytic integral [24], a Fourier series expansion [32], a post-circular expansion, and numerical integration [33]. In this work, we compare the computational cost of these methods and investigate the accuracy of the most efficient method. We then introduce a new spline-based method for further improving the computational performance of PTA signal evaluations by leveraging the properties of real PTA datasets. These results are implemented in the GWecc package [33, 37], and can readily be used with the widely-used ENTERPRISE package [38] for searching PTA datasets.

This paper is arranged as follows. In Section II, we introduce the PTA signals induced by inspiralling eccentric binaries, and the GW phasing approach used for their computation. We compare the computational cost of different approaches for evaluating the PTA signals given the orbital evolution as a function of time in III. We investigate the accuracy of the most computationally efficient PTA signal computation method in IV. In Section V, we introduce a new spline-based method for further reducing the computational cost of evaluating PTA signals by exploiting the structure of typical PTA datasets. Finally, we summarize our results in Section VI.

II. THE PTA SIGNAL MODEL FOR ECCENTRIC BINARIES

A. The gravitational waveform and the PTA signal

We begin by briefly describing the PN-accurate PTA signal model for inspiralling eccentric binaries. GWs traveling across the line of sight to a pulsar induces modulations on the observed TOAs of its pulses. These modulations are given by [39]

\[ R(t_E) = \int_{t_0}^{t_E} \left( h(t'_E) - h(t'_E - \Delta_p) \right) dt'_E , \]

where the time variables \( t_E \) and \( t'_E \) are measured in the solar system barycenter (SSB) frame, \( \Delta_p \) is a geometric time delay given by

\[ \Delta_p = \frac{D_p}{c} \left( 1 - \cos \mu \right) , \]

where \( D_p \) is the pulsar distance, \( \mu \) is the angle between the lines of sight to the pulsar and the GW source, and \( t_0 \) is an arbitrary fiducial time. The coordinate time \( t_E \) measured in the SSB frame
relates to the coordinate time \( t \) measured in the GW source frame via the cosmological redshift as
\[
t_E - t_0 = (1 + z)(t - t_0),
\]
(3)
and the PTA signal can be re-written as a function of \( t \) as
\[
R(t) = (1 + z) \int_{t_0}^{t} (h(t') - h(t' - \Delta'_p)) \, dt',
\]
(4)
where \( \Delta'_p = \Delta_p/(1 + z) \).

The dimensionless GW strain \( h(t) \) is given by
\[
h(t) = \begin{bmatrix} F_+ & F_x \end{bmatrix} \begin{bmatrix} \cos 2\psi - \sin 2\psi \\ \sin 2\psi & \cos 2\psi \end{bmatrix} \begin{bmatrix} h_+(t) \\ h_\times(t) \end{bmatrix},
\]
(5)
where \( h_+, h_\times(t) \) are the two GW polarization amplitudes, \( F_+, F_\times \) are the antenna pattern functions that depend on the sky locations of the pulsar and the GW source (see, e.g., Ref. [40] for explicit expressions), and \( \psi \) is the GW polarization angle. Hence, \( R(t) \) can be expressed in terms of functions \( s_+, s_\times(t) \) defined as
\[
s_+, s_\times(t) = (1 + z) \int_{t_0}^{t} h_+, h_\times(t') \, dt',
\]
(6)
such that
\[
R(t) = \begin{bmatrix} F_+ & F_x \end{bmatrix} \begin{bmatrix} \cos 2\psi - \sin 2\psi \\ \sin 2\psi & \cos 2\psi \end{bmatrix} \begin{bmatrix} s_+(t) - s_+(t - \Delta'_p) \\ s_\times(t) - s_\times(t - \Delta'_p) \end{bmatrix}.
\]
(7)
The \( s_+, s_\times(t - \Delta'_p) \) contributions are known as the Earth term and the pulsar term respectively.

The quadrupolar-order \( h_+, h_\times \) expressions valid for binary systems inspiralling along eccentric orbits are given by [33]
\[
h_+ = H \left( c_t^2 + 1 \right) \left( \frac{2e_t^2 - \chi^2 + \chi - 2}{(1 - \chi)^2} \cos(2\phi) - \frac{2\sqrt{1 - e_t^2\xi}}{(1 - \chi)^2} \sin(2\phi) \right) + s_t^2 \frac{\chi}{(1 - \chi)},
\]
(8a)
\[
h_\times = H 2c_t \left( \frac{2\sqrt{1 - e_t^2\xi}}{(1 - \chi)^2} \cos(2\phi) + \frac{(2e_t^2 - \chi^2 + \chi - 2)}{(1 - \chi)^2} \sin(2\phi) \right),
\]
(8b)
where \( H = \frac{GM\eta}{D_Lc} \), \( \eta \) is the symmetric mass ratio, \( D_L \) is the luminosity distance to the binary, \( x = (GM(1 + k)n/c^3)^{2/3} \) is a dimensionless PN parameter, \( n \) is the mean motion of the orbit, \( k \) is the relativistic advance of periapsis per orbit, \( e_t \) is the time eccentricity, \( c_t = \cos \iota \), \( s_t = \sin \iota \), \( \iota \) is the orbital inclination, \( \chi = e_t \cos u \), \( \xi = e_t \sin u \), \( u \) is the true anomaly, and \( \phi \) is the angular coordinate in the orbital plane also known as the orbital phase. The variables \( \phi \) and \( u \) can be obtained as functions of time using the gravitational wave phasing approach [41], and this is what we discuss in the next subsection.
B. The GW phasing formalism for computing the orbital motion

The conservative dynamics of the binary system can be expressed using the PN-accurate quasi-Keplerian parametrization [42, 43]. We begin by defining the mean anomaly $l$ as

$$l(t) = \int_{t_0}^{t} n(t')dt'.$$

(9)

The eccentric anomaly $u$ can be written implicitly as a function of $l$ using the PN-accurate Kepler equation [43, 44]

$$l = u - e_t \sin u + \Phi_l(u),$$

(10)

where $e_t$ is known as the time eccentricity and $\Phi_l(u)$ is a periodic function of $u$. Although this transcendental equation admits an analytic Fourier series solution [44], it is usually solved numerically in practice due to the computationally expensive nature of the analytic solution. An efficient numerical method for solving the Newtonian Kepler equation was introduced by Ref. [45], and can be adapted to solve the PN-accurate Kepler equation.

The orbital phase $\phi$ can be written as

$$\phi = \gamma + (1 + k)(f - l) + \Phi_\phi(u),$$

(11)

where $f$ is the true anomaly given by

$$f = 2 \arctan \left( \frac{\sqrt{1 + e_\phi}}{1 - e_\phi} \tan \frac{u}{2} \right),$$

(12)

$k$ is the advance of periapsis per orbit, $\gamma$ is the angle of periapsis defined as

$$\gamma(t) = \int_{t_0}^{t} k(t')n(t')dt',$$

(13)

e_\phi is known as the angular eccentricity, and $\Phi_\phi(u)$ is a periodic function of $u$. The periastron angle $\gamma$ is not to be confused with the argument of periastron $\omega$, which is defined as $\omega = \phi - f$; instead, $\gamma(t)$ represents the orbital-averaged secular evolution of $\omega$. Explicit expressions for $k$, $e_\phi$, $\Phi_l(u)$ and $\Phi_\phi(u)$ in terms of $e_t$, $n$, $M$, and $\eta$ can be found in, e.g., Ref. [43]. In the Newtonian limit $x \to 0$, we have $e_\phi, e_t \to e$, $k \to 0$, $\Phi_l \to 0$, $\Phi_\phi \to 0$, and $\omega \to \gamma$.

The reactive evolution of the orbit due to gravitational radiation reaction can be incorporated into this formalism by allowing $n$ and $e_t$ to vary slowly with time. Up to the leading quadrupolar order, this can be expressed as a system of four coupled ordinary differential equations (ODEs) [33, 41]

$$\frac{dn}{dt} = \frac{1}{5} \left( \frac{GM_{ch}n}{c^3} \right)^{\frac{3}{2}} n^2 \frac{\left(96 + 292e_t^2 + 37e_t^4\right)}{(1 - e_t^2)^{7/2}},$$

(14a)
\[
\frac{de_t}{dt} = \frac{-1}{15} \left( \frac{GM_{ch} n}{c^3} \right)^{\frac{2}{3}} ne_t \left( \frac{304 + 121e_t^2}{(1 - e_t^2)^{5/2}} \right),
\]
\[
\frac{d\gamma}{dt} = kn,
\]
\[
\frac{dl}{dt} = n.
\]

An analytic solution for this set of equations was derived by Refs. [33, 36], and involves hypergeometric functions. (We do not provide this solution here explicitly due to its length.) This solution allows us to evolve the binary orbit efficiently without the need for numerical ODE solvers.

C. Computing the PTA signal

Given the complicated PN-accurate functional forms of \( h_{+,*}(t) \), it is in general not possible to obtain exact expressions for their integrals \( s_{+,*}(t) \), which are necessary for computing the PTA signal \( R(t) \). Four different approaches for computing \( s_{+,*}(t) \) given the dynamic variables \( n(t), e_t(t), \gamma(t), l(t) \) have been presented in literature, and these methods are listed below.

**Method 1 (Analytic):** Ref. [24] derived analytic expressions for \( s_{+,*}(t) \) assuming slow advance of periapsis and orbital decay compared to the orbital motion, as well as the Newtonian limit for the eccentricities \( e_\phi \to e_t \). These expressions, re-written in terms of \( u \) and \( \omega \) in our notation, are given by

\[
s_{A+}(t) = \frac{(1 + z)H}{n} \left( \left( c_t^2 + 1 \right) \left( -P \sin(2\omega) + Q \cos(2\omega) \right) + s_t^2 R \right),
\]
\[
s_{A*}(t) = \frac{(1 + z)H}{n} 2c_t \left( P \cos(2\omega) + Q \sin(2\omega) \right),
\]

where

\[
P = \frac{\sqrt{1 - e_t^2} (\cos(2u) - e_t \cos(u))}{1 - e_t \cos(u)},
\]
\[
Q = \frac{\left( e_t^2 - 2 \right) \cos(u) + e_t \sin(u)}{1 - e_t \cos(u)},
\]
\[
R = e_t \sin(u).
\]

Specifically, \( s_{A+,*}(t) \) obey the relation

\[
h_{+,*} = \frac{n}{(1 + z)} \left. \frac{\partial s_{+,*}}{\partial t} \right|_{e_\phi \to e_t; n, e_t, \omega}.
\]
Note that these expressions are exact when only the Newtonian motion is considered, where $\omega$ is a constant. In the case of relativistic binaries with a non-zero advance of periapsis, the PN-accurate values of $u$ and $\omega$ can be substituted.

**Method 2 (Fourier):** Ref. [32] derived Fourier series expansions for $s_{+\times}(t)$ using the Fourier series expansions of $h_{+\times}(t)$ [46], assuming slow advance of periapsis and orbital decay, as well as the Newtonian limit for the eccentricities $e_\phi \to e_t$, and neglecting the orbital-timescale variations of $\omega$. In our notation, these expressions can be written as

$$s_{+}^F(t) = \frac{(1+z)H}{n} \sum_{p=1}^{\infty} \left[ (c_t^2 + 1) \left( -\mathcal{P} \cos(pl) \sin(2\gamma) + \mathcal{Q} \sin(pl) \cos(2\gamma) \right) + s_t^2 \mathcal{R} \sin(pl) \right], \quad (17a)$$

$$s_{\times}^F(t) = \frac{(1+z)H}{2c_t} \sum_{p=1}^{\infty} \left( \mathcal{P} \cos(pl) \cos(2\gamma) + \mathcal{Q} \sin(pl) \sin(2\gamma) \right), \quad (17b)$$

where

$$\mathcal{P} = \sqrt{1 - e_t^2} \left[ J_{p-2}(pe_t) + J_{p+2}(pe_t) - 2J_p(pe_t) \right], \quad (18a)$$

$$\mathcal{Q} = -J_{p-2}(pe_t) + 2e_t J_{p-1}(pe_t) - \frac{2}{p} J_p(pe_t) - 2e_t J_{p+1}(pe_t) + J_{p+2}(pe_t), \quad (18b)$$

$$\mathcal{R} = \frac{2}{p} J_p(pe_t), \quad (18c)$$

and $J_n(x)$ represent Bessel functions of the first kind. It is straightforward to show that the above Fourier series expansions $s_{+\times}^F(t)$ are equivalent to Eqs. (15) in the limit $\omega \to \gamma$.

**Method 3 (Post-circular):** Ref. [33] derived a post-circular expansion (expansion in powers of $e_t$) for $s_{+\times}(t)$ valid for low eccentricities ($e_t < 0.3$) incorporating fully 3PN-accurate advance of periapsis using the post-circular expansion of $h_{+\times}(t)$ derived in Ref.[44]. This expansion has the form

$$s_{+\times}^{PC}(t) = \frac{(1+z)H}{n} \sum_{p,q} \left\{ A_{p,q}^{+\times} \cos(pl) \cos(q\lambda) + B_{p,q}^{+\times} \sin(pl) \cos(q\lambda) \right.$$

$$+ C_{p,q}^{+\times} \cos(pl) \sin(q\lambda) + D_{p,q}^{+\times} \sin(pl) \sin(q\lambda) \left\} , \quad (19)$$

where $\lambda = l + \gamma$ and $A_{p,q}^{+\times}, B_{p,q}^{+\times}, C_{p,q}^{+\times}$, and $D_{p,q}^{+\times}$ are polynomials of $e_t$. The primed sum excludes the $p = q = 0$ term and truncates after some finite number of harmonics $P$ depending on the order of the post-circular expansion.
Method 4 (Numerical): Ref. [33] computed $s_{+,\times}(t)$ by numerically integrating $h_{+,\times}(t)$ incorporating fully 3PN-accurate orbital dynamics. We denote the $s_{+,\times}(t)$ computed via this method as $s_{+,\times}^N(t)$.

Each of these methods has its own advantages and disadvantages. While evaluating $s_{+,\times}(t)$ given $u(t)$ and $\phi(t)$ using Method 1 is relatively inexpensive, it requires the numerical solution of the Kepler equation which can be computationally expensive. Method 2 does not require a solution of the Kepler equation but requires expensive Bessel function evaluations while computing the Fourier coefficients. This Fourier series converges slowly for high eccentricities, and its truncation at some finite number of harmonics can affect the accuracy of the computed PTA signal (see Appendix A for an estimate of the required number of harmonics as a function of eccentricity). Furthermore, both Method 1 and Method 2 assume a slow advance of periapsis and can give inaccurate results for more relativistic systems. This is especially true of Method 2, where only the secular variations of $\omega$, given by $\gamma$, are incorporated (see Section IV for a theoretical estimate of the error incurred in Method 1 due to this assumption). Method 3 requires neither the solution of the Kepler equation nor the evaluation of expensive special functions but is valid only for low eccentricities. Method 4, while being the most accurate, is also the most computationally expensive due to the application of numerical integration.

These considerations motivate us to investigate the computational cost and accuracy of the PTA signals computed using these methods, and this is what we pursue in the forthcoming sections.

III. COMPUTATIONAL COST OF EVALUATING $s_{+,\times}(t)$

We now proceed to compare the cost of computing the PTA signal $R(t)$ using the approximate analytic expressions, Fourier series expansion, post-circular approximation, and numerical integration. Figure 1 shows the execution time per TOA for the four methods for different eccentricities, normalized by the execution time per TOA for computing circular PTA signals. Figure 1 shows that for Analytic, Post-circular, and Numerical computations of $R(t)$, the run time does not strongly depend on the eccentricity, whereas the Fourier series expansion shows an increasing trend. The number of harmonics required by the Fourier expansion to achieve a given degree of precision is an increasing function of eccentricity, and the trend seen in the run time for the Fourier series expansion reflects this relationship (see Appendix A).

The most important observation from Figure 1 is that the computation using the analytic
FIG. 1. Computational performance comparison between different methods of computing $R(t)$ as a function of the initial eccentricity $e_{i0}$. Only the Earth terms are used for the comparison. The binary parameters used are $M = 5 \times 10^9 M_\odot$, $\eta = 0.25$, $P_{b0} = 2\pi/n_0 = 2$ yrs, $l_0 = 0$, $\gamma_0 = 0$, $i = 0$, $\psi = 0$. The execution times are normalized using the execution time for the circular case. The figure shows that the analytic expressions provide $\sim 50$ times improvement in performance over numerical integration, but are still $\sim 200$ times slower than computing the circular PTA signals. The increasing trend in the computational cost for the Fourier expansion reflects the fact that the number of harmonics of the Fourier expansion required to achieve a given accuracy is an increasing function of eccentricity. (See Eq. (A6).) The plot for the post-circular method is truncated because it is not valid for $e_{i0} \geq 0.3$.

expressions (Eqs. (15)) is faster than all other methods. Specifically, it is $\sim 50$ times faster than the computation using numerical integration, which is the most accurate method. However, the fastest method for computing the eccentric PTA signals is still $\sim 200$ times slower than computing the circular PTA signals.

It should be evident from Eqs. (15) and 17 that Method 2 is an approximation of Method 1, which in turn is an approximation of Method 4. Hence, it will be advantageous to use the analytic expressions to compute $s_{+,x}(t)$ provided they are accurate enough within the region of the binary parameter space that is of interest to PTAs. We investigate the accuracy of the analytic expressions, as compared to numerical integration, in the next section.
IV. ACCURACY OF THE APPROXIMATE ANALYTIC $s_{+,\times}(t)$ EXPRESSIONS

While the analytic $s_{+,\times}(t)$ expressions given in Eqs. (15) are exact in the Newtonian limit, it is not so when PN effects, especially the advance of periastris, are considered. In the PN case, we evaluate $s_{+,\times}^{A}(t)$ using PN-accurate values of $u$ and $\omega$, where $\omega = \phi - f$. The error incurred in this prescription can be estimated as follows:

$$|s_{+,\times} - s_{+,\times}^{A}| = |s_{+,\times} - \int dt \frac{ds_{+,\times}^{A}}{dt}|$$

$$= |s_{+,\times} - \int dt \frac{\partial s_{+,\times}^{A}}{\partial l} - \int dt \frac{d\omega}{dt} \frac{\partial s_{+,\times}^{A}}{\partial \omega}|$$

$$\approx |s_{+,\times} - \int dt h_{+,\times}(t) - kn \int dt \frac{\partial s_{+,\times}^{A}}{\partial \omega}|$$

$$= kn \left| \int dt \frac{\partial s_{+,\times}^{A}}{\partial \omega} \right|. \quad (20)$$

It is straightforward to show that the functions $P$ and $Q$ (Eqs. (16)) have zero orbital average. This ensures that the orbital average of $\frac{\partial s_{+,\times}^{A}}{\partial \omega}$ vanishes, and the error $|s_{+,\times} - s_{+,\times}^{A}|$ as a function of time for a given set of orbital parameters does not grow secularly with time. While we incur an amplitude error at the 1PN order in $s_{+,\times}^{A}(t)$ (specifically, the error is $O(k)$), the phase is accurate to whichever PN order $u$ and $\phi$ are computed. We may expect, then, that this approximation will be valid within some low-to-moderately relativistic region of the parameter space. To establish this, we proceed by comparing the PTA signals computed using the analytic $s_{+,\times}$ expressions (Method 1) and via numerical integration (Method 4) using a particular mismatch metric in the following subsection.

A. Mismatch metric for comparing PTA signals

In this section, we investigate the accuracy of the analytic $s_{+,\times}$ expressions against numerically integrated $s_{+,\times}$ using a mismatch metric applicable to PTA signals. Recall that the mismatch between two signal vectors $a$ and $b$ is usually defined as

$$\text{mismatch}[a,b] = 1 - \frac{(a,b)}{\sqrt{(a,a)(b,b)}}, \quad (21)$$

where $(a,b)$ represents an appropriately defined inner product that takes into account the detector response. In the case of a PTA dataset, $a$ and $b$ will be some signals evaluated at each TOA, and the appropriate notion of the inner product is given by Ref. [47] as

$$(a,b) = a^{T}Kb, \quad (22)$$
where

\[ K = C^{-1} - C^{-1}F(F^T C^{-1} F)^{-1} F^T C^{-1}, \]  

(23)

\( F \) is the \( N \times p \) dimensional pulsar timing design matrix, \( C \) is the \( N \times N \) dimensional TOA covariance matrix, \( p \) is the number of timing model parameters, and \( N \) is the number of TOAs. For simplicity, we assume an ideal pulsar dataset with no correlated noise and equal TOA measurement uncertainties \( \sigma \). In this case, \( C = \sigma^2 I \), where \( I \) is the \( N \times N \) dimensional identity matrix, and \( K \) simplifies to

\[ K = \sigma^{-2} (I - F(F^T F)^{-1} F^T), \]

and we note that mismatch\([a, b]\) is independent of \( \sigma \). Therefore, we set \( \sigma = 1 \) and write

\[ K = I - M(M^T M)^{-1} M^T. \]  

(24)

The most important contributions to \( F \) arise from fitting TOAs for the pulse phase offset, pulsar spin frequency, spin frequency derivative, and the sky location of the pulsar. The effects of fitting these parameters are, approximately, to remove a constant offset, a straight line, a quadratic function, and a sinusoid of period 1 year respectively. Hence, given some TOAs \( t_1 \ldots t_N \), the design matrix \( M \) can be approximated as

\[
\begin{bmatrix}
1 & 1 & \ldots & 1 \\
t_1/t_m - 1 & t_2/t_m - 1 & \ldots & t_N/t_m - 1 \\
(t_1/t_m - 1)^2 & (t_2/t_m - 1)^2 & \ldots & (t_N/t_m - 1)^2 \\
\sin\left(\frac{2\pi t_1}{1 \text{ yr}}\right) & \sin\left(\frac{2\pi t_2}{1 \text{ yr}}\right) & \ldots & \sin\left(\frac{2\pi t_N}{1 \text{ yr}}\right) \\
\cos\left(\frac{2\pi t_1}{1 \text{ yr}}\right) & \cos\left(\frac{2\pi t_2}{1 \text{ yr}}\right) & \ldots & \cos\left(\frac{2\pi t_N}{1 \text{ yr}}\right)
\end{bmatrix},
\]

where \( t_m \) is the median TOA. Note that this idealized scenario corresponds to a highly sensitive PTA dataset and will therefore produce worse mismatch values for a given pair of signals as compared to a realistic dataset. Hence, this is a more stringent comparison between signals than one done using a realistic dataset.

We now plot in Figure 2 the mismatch between approximate analytic PTA signals given by Eqs. (15) and numerically integrated PTA signals for different values of \( M, P_{50} \) and \( e_i \). We fix the other binary parameters and the pulsar and GW source sky locations since the largest contribution to the mismatch arises from the non-exact treatment of the advance of periapsis while doing the integrals in Eqs. (16), and the advance of periapsis per orbit \( k \) only depends on \( M, n \) and \( e_t \) at the leading order [42].
FIG. 2. Mismatch between approximate analytic PTA signals and numerically integrated PTA signals for different values of $M$, $P_{b0}$, and $e_{t0}$. The plots correspond to RA$_{psr} = 4^h37^m15.81476^s$, DEC$_{psr} = -47^\circ15'8.6242''$, RA$_{gw} = 4^h0^m0^s$, DEC$_{gw} = -45^\circ0'0''$, $\eta = 0.25$, $\psi = 0$, $\iota = 0$, $l_0 = 0$, $\gamma_0 = 0$. Only the Earth terms are compared. The mismatch is only $\sim 0.01$ even for $M = 10^9M_\odot$, $e_{t0} = 0.85$, and $P_{b0} = 0.5$ year.

In Figure 2, the mismatch grows worse for increasing $M$ and $e_{t0}$ and decreasing $P_{b} = 2\pi/n$ as expected. Interestingly, even for high eccentricities ($e \sim 0.85$), the mismatch is only $\sim 0.01$, indicating the good quality of the approximation. Hence, we may conclude that this approximation can be used for masses up to $10^9M_\odot$, eccentricities up to 0.85, and orbital periods as small as 0.5 years. This can also be seen in Figure 3, which shows a visual comparison of the approximate analytic and numerically integrated waveforms.
FIG. 3. Comparison between the analytic PTA signal \( s_{+}^A(t_E) \) (red dashed line) and the numerically integrated PTA signal \( s_{+}^N(t_E) \) (blue solid line) for different eccentricities, \( M = 10^9 M_\odot, \eta = 0.25, P_b = 1.5 \) years, \( \psi = 0, \iota = 0, l_0 = 0, \) and \( \gamma_0 = 0 \). The difference \( s_{+}^A(t_E) - s_{+}^N(t_E) \) is plotted using black dotted lines with a separate Y axis on the right hand side of each plot. Only the Earth-term contributions are shown. The left panels show \( s_+ \) and the right panels show \( s_\times \). The difference between the curves is visually discernible from the red and blue curves only for \( e = 0.6 \) where the curves show sharp features.

V. EFFICIENT COMPUTATION OF \( R(t) \) FOR REALISTIC PTA DATASETS

The relatively efficient method discussed above for computing \( R(t) \) still requires costly hypergeometric function evaluations to solve the orbital evolution analytically (See Eqs. 30 of Ref. [33]) as well as the inversion of the Kepler equation (Eq. (10)). Furthermore, it is in general not possible to obtain a fully analytic prescription for computing \( R(t) \) when higher order reactive PN effects [e.g. 41] are taken into account. In this section, we further reduce the computational cost of evaluating \( R(t) \) given a set of TOAs by leveraging the properties of realistic PTA datasets.

PTA experiments measure TOAs from their radio observations via two different techniques, namely the traditional narrow-band technique [e.g. 5] and the modern wide-band technique [e.g. 48], resulting in two types of datasets. These techniques differ in how they model the observing frequency-dependent effects in the pulsar signal, such as the interstellar dispersion and the frequency-dependent evolution of the integrated pulse profile. The narrow-band technique mea-
sures multiple TOAs per observation epoch by splitting broadband observations into multiple frequency sub-bands, such that the frequency-dependent effects are reduced in each sub-band [49]. The wide-band technique employs principal component decomposition of the frequency-resolved pulse profiles to model the interstellar dispersion and the frequency-dependent profile evolution, resulting in one TOA and one DM measurement per observation [50, 51]. Further, the PTA datasets usually have observational cadences of the order of weeks, with time spans up to decades.

In narrow-band datasets, the sub-banded TOAs derived from the same observation are typically only a fraction of a second apart, and therefore the $R(t)$ values evaluated at those TOAs will not show any appreciable variation. There can be multiple TOAs per epoch even in the case of wide-band datasets, e.g., in the case of simultaneous multi-band observations [52]. In such cases, it is not necessary to evaluate $R(t)$ independently at each TOA. Rather, one can evaluate $R(t)$ only at unique epochs $t_i$ and interpolate the $R(t)$ values for the TOAs $t_{ij}$ at that epoch as

$$R(t_{ij}) = R(t_i) + h(t_i)(t_i - t_{ij}).$$

Note that this computation requires $h(t_i)$ at each epoch, which is cheap to compute together with $R(t_i)$. The correction $h(t_i)(t_i - t_{ij})$ has a magnitude that is $n|t_i - t_{ij}|$ times less than $R(t)$, which is only $\sim 10^{-2}$ for $n \sim 100$ nHz and $|t_i - t_{ij}| \sim 1$ day. The relative error incurred in this approximation is of the order of $(n|t_i - t_{ij}|)^2 \sim 10^{-4}$ for $n \sim 100$ nHz and $|t_i - t_{ij}| \sim 1$ day. Therefore, $R(t)$ has to be computed only once every day, even if the TOAs on that day were not obtained from simultaneous observations. If required, the approximation can be improved by including the $O((n|t_i - t_{ij}|)^2)$ term, which requires the computation of $h'(t) = dh/dt$. $h'(t)$ can be derived easily by differentiating Eq. (8) and is cheap to compute together with $R(t)$ like $h(t)$. In practice, we implement this by interpolating the $R(t)$ values computed for every observation epoch using a cubic Hermite spline [53] to compute $R(t)$ at every TOA.

This method can, in an ideal scenario, reduce the computational cost of evaluating $R(t)$ by a factor equal to the average number of TOAs per day, without any appreciable loss of accuracy. For example, the NANOGrav 12.5-year data release has $\sim 58$ TOAs per epoch for the narrow-band dataset and $\sim 1.4$ TOAs per epoch for the wide-band dataset on average for PSR J1909$-$3744 [5, 48]. This method can be applied to TOAs of multiple pulsars if only the Earth term is considered, providing a further performance improvement of a factor that is proportional to the number of pulsars in the best-case scenario. The performance gained by employing this method for different numbers of TOAs per epoch is plotted in Figure 4. This figure clearly shows the expected decreasing trend in the ratio between the execution times of computing $R(t)$ using the interpolation method
FIG. 4. Ratio between execution times of computing $R(t)$ at every TOA using the interpolation method vs independently at each TOA.

and computing $R(t)$ independently at each TOA.

VI. SUMMARY AND DISCUSSION

In this work, we have investigated the computational cost of four different methods of computing the PTA signal $R(t)$ given the orbital variables (the mean motion $n$, eccentricity $e$, mean anomaly $l$, and the periastron angle $\gamma$) as functions of time. These four methods are (1) using approximate analytic expressions, (2) using a truncated Fourier series expansion, (3) using a post-circular expansion, and (4) using numerical integration. Our performance comparison revealed that the approximate analytic expressions provide the best execution times out of the four methods, and is $\sim 50$ times faster than numerical integration, which is the most accurate method.

We showed that the approximate analytic expressions incur an amplitude error at the 1PN level. We characterized this error by introducing a mismatch metric valid for comparing PTA signals. This comparison revealed that the mismatch between the PTA signals computed using the analytic expressions and numerical integration is only $\sim 0.01$ even for $M = 10^9 M_\odot$, $e_{i0} = 0.85$ and $P_{b0} = 0.5$ year; i.e., the computationally efficient analytic expressions can be safely used within the
region of the binary parameter space that is of interest to PTA experiments.

Although the analytic PTA signal expressions significantly improve the computational performance, they still require the evaluation of expensive hypergeometric functions for solving the orbital evolution. To address this, we introduced a spline interpolation method that exploits the typical structure of the narrow-band PTA datasets, where each epoch has multiple TOAs, to achieve further gains in computational performance. This method computes $R(t)$ and $h(t)$ only once every epoch and interpolates the $R(t)$ values at every TOA using a cubic Hermite spline. This method can in principle provide an improvement in performance by a factor equal to the average number of TOAs per epoch.

The prohibitively high cost of computing the PTA signals has been a major bottleneck in searching for and constraining the presence of inspiralling eccentric SMBHBs in PTA datasets. The results presented in this paper will provide a significant boost to such endeavors. These methods have been incorporated into the G\textit{wecc} package, and we plan to apply them to search for such eccentric SMBHBs in the upcoming NANOGrav 15-year dataset and the IPTA Data Release 3.

ACKNOWLEDGMENTS

AS is supported by the NANOGrav NSF Physics Frontiers Center (awards #1430284 and 2020265). We acknowledge the use of the following software packages: \texttt{numpy} \cite{numpy}, \texttt{scipy} \cite{scipy}, \texttt{astropy} \cite{astropy}, \texttt{matplotlib} \cite{matplotlib}, GSL \cite{gsl}, and \texttt{ENTERPRISE} \cite{enterprise}. AS thanks David Kaplan and Sarah Vigeland for providing valuable comments and suggestions on this manuscript, Stephen Taylor for fruitful discussions, and Amit Jit Singh, Nidhi Pant, Lankeswar Dey, and Belinda Cheeseboro for providing valuable feedback on the G\textit{wecc} package.

Appendix A: Number of harmonics required for accurate computation of quadrupolar-order gravitational waveforms using its Fourier series expansion

The total GW power emitted by an eccentric binary, accurate up to the leading quadrupolar order, is given by \cite{bib1}

$$P(e) = \frac{32}{5} \frac{e^5}{G} \left( \frac{GMn}{c^3} \right)^{10/3} \eta^2 f(e),$$  \hspace{1cm} (A1)
where
\[
\begin{align*}
    f(e) &= \frac{(1 + \frac{73}{24} e^2 + \frac{37}{36} e^4)}{(1 - e^2)^{7/2}}, \quad (A2)
\end{align*}
\]
and the corresponding spectral power associated with the \( p \)th harmonic is given by \[46\]
\[
    P_p(e) = \frac{32}{5} e^5 \left( \frac{GMn}{c^3} \right)^{10/3} \eta^2 g_p(e), \quad (A3)
\]
where
\[
    g_p(e) = \frac{n^4}{32} \left\{ \left[ J_{p-2}(pe) - 2eJ_{p-1}(pe) + \frac{2}{p} J_p(pe) + 2eJ_{p+1}(pe) - J_{p+2}(pe) \right]^2 \\
    + (1 - e^2) \left[ J_{p-2}(pe) - 2J_p(pe) + J_{p+2}(pe) \right]^2 + \frac{4}{3p^2} \left[ J_p(pe) \right]^2 \right\}. \quad (A4)
\]
We define \( \mathcal{N}(e) \) as the number of harmonics that captures 99.9% of the total emitted power for a given eccentricity \( e \). i.e., \( \mathcal{N}(e) \) is the smallest integer such that
\[
    \sum_{p=1}^{\mathcal{N}(e)} g_p(e) f(e) \geq 0.999. \quad (A5)
\]
We plot in Figure 5 numerically estimated values of \( \mathcal{N}(e) \) for different values of \( e \). We have found that the following formula fits the numerically estimated \( \mathcal{N}(e) \) well:
\[
    \mathcal{N}(e) \approx \alpha (1 - e^2)^{-3/2} + \beta, \quad (A6)
\]
with \( \alpha = 18.64801851 \) and \( \beta = -14.04695398 \). The \( \mathcal{N}(e) \) values predicted by this fitting formula are plotted in Figure 5 using a solid grey line. We have found that the fitting function provides accurate estimates for \( \mathcal{N}(e) \) for \( e \gtrsim 0.063 \).
FIG. 5. Number of harmonics $N$ required to capture 99.9% of the total emitted power in a Peters-Mathews waveform. Blue circles represent numerically estimated values and the grey solid curve represents the fitting function.

[6] G. Desvignes, R. N. Caballero, L. Lentati, J. P. Verbiest, D. J. Champion, et al., Monthly Notices of the Royal Astronomical Society 458, 3341 (2016).

[7] M. Kerr, D. J. Reardon, G. Hobbs, R. M. Shannon, R. N. Manchester, et al., Publications of the Astronomical Society of Australia 37, e020 (2020).

[8] P. Tarafdar, K. Nobleson, P. Rana, J. Singha, M. A. Krishnakumar, et al., arXiv e-prints, 2206.09289 (2022).

[9] R. Spiewak, M. Bailes, M. T. Miles, A. Parthasarathy, D. J. Reardon, et al., Publications of the Astronomical Society of Australia 39, e027 (2022), arXiv:2204.04115 [astro-ph.HE].

[10] K. J. Lee (Astronomical Society of the Pacific, 2016) p. 19.

[11] B. B. Perera, M. E. DeCesar, P. B. Demorest, M. Kerr, L. Lentati, et al., Monthly Notices of the Royal Astronomical Society 490, 4666 (2019).

[12] B. Goncharov, R. M. Shannon, D. J. Reardon, G. Hobbs, A. Zic, et al., The Astrophysical Journal Letters 917, L19 (2021).

[13] S. Chen, R. N. Caballero, Y. J. Guo, A. Chalumeau, K. Liu, et al., Monthly Notices of the Royal Astronomical Society 508, 4970 (2021).

[14] Z. Arzoumanian, P. T. Baker, H. Blumer, B. Bécsy, A. Brazier, et al., Phys. Rev. Lett. 127, 251302 (2021).
[15] J. Antoniadis, Z. Arzoumanian, S. Babak, M. Bailes, A.-S. Bak Nielsen, et al., Monthly Notices of the Royal Astronomical Society 510, 4873 (2022).
[16] S. Burke-Spolaor, S. R. Taylor, M. Charisi, T. Dolch, J. S. Hazboun, et al., Astronomy and Astrophysics Review 27, 5 (2019).
[17] L. Z. Kelley, L. Blecha, L. Hernquist, A. Sesana, and S. R. Taylor, Monthly Notices of the Royal Astronomical Society 471, 4508 (2017).
[18] L. Z. Kelley, L. Blecha, L. Hernquist, A. Sesana, and S. R. Taylor, Monthly Notices of the Royal Astronomical Society 477, 964 (2018).
[19] N. S. Pol, S. R. Taylor, L. Z. Kelley, S. J. Vigeland, J. Simon, et al., The Astrophysical Journal Letters 911, L34 (2021).
[20] S. Iguchi, T. Okuda, and H. Sudou, The Astrophysical Journal 724, L166 (2010).
[21] L. Dey, A. Gopakumar, M. Valtonen, S. Zola, A. Susobhanan, et al., Universe 5, 108 (2019).
[22] B. X. Hu, D. J. D’Orazio, Z. Haiman, K. L. Smith, B. Sniros, et al., Monthly Notices of the Royal Astronomical Society 495, 4061 (2020).
[23] C. Xin, C. M. F. Mingarelli, and J. S. Hazboun, The Astrophysical Journal 915, 97 (2021).
[24] F. A. Jenet, A. Lommen, S. L. Larson, and L. Wen, The Astrophysical Journal 606, 799 (2004).
[25] S. Babak, A. Petiteau, A. Sesana, P. Brem, P. A. Rosado, et al., Monthly Notices of the Royal Astronomical Society 455, 1665 (2016).
[26] K. Aggarwal, Z. Arzoumanian, P. T. Baker, A. Brazier, M. R. Brinson, et al., The Astrophysical Journal 880, 116 (2019).
[27] Z. Arzoumanian, P. T. Baker, A. Brazier, P. R. Brook, S. Burke-Spolaor, et al., The Astrophysical Journal 900, 102 (2020).
[28] M. C. Begelman, R. D. Blandford, and M. J. Rees, Nature 287, 307 (1980).
[29] M. Dotti, A. Sesana, and R. Decarli, Advances in Astronomy 2012, 940568 (2012), arXiv:1111.0664 [astro-ph.CO].
[30] P. J. Armitage and P. Natarajan, The Astrophysical Journal 634, 921 (2005).
[31] C. Roedig and A. Sesana, Journal of Physics: Conference Series 363, 012035 (2012).
[32] S. R. Taylor, E. A. Huerta, J. R. Gair, and S. T. McWilliams, The Astrophysical Journal 817, 70 (2016).
[33] A. Susobhanan, A. Gopakumar, G. Hobbs, and S. R. Taylor, Physical Review D 101, 043022 (2020).
[34] L. Blanchet, Living Reviews in Relativity 17, 2 (2014).
[35] X.-J. Zhu, L. Wen, G. Hobbs, Y. Zhang, Y. Wang, D. R. Madison, R. N. Manchester, M. Kerr, P. A. Rosado, and J.-B. Wang, Monthly Notices of the Royal Astronomical Society 449, 1650 (2015).
[36] B. Moore, T. Robson, N. Loutrel, and N. Yunes, Classical and Quantum Gravity 35, 235006 (2018).
[37] A. Susobhanan, Gwec: Calculator for pulsar timing array signals due to eccentric supermassive binaries (2020).
[38] J. A. Ellis, M. Vallisneri, S. R. Taylor, and P. T. Baker, Enterprise: Enhanced numerical toolbox
enabling a robust pulsar inference suite, Zenodo (2020).

[39] M. Anholm, S. Ballmer, J. D. E. Creighton, L. R. Price, and X. Siemens, Physical Review D 79, 084030 (2009).

[40] K. J. Lee, N. Wex, M. Kramer, B. W. Stappers, C. G. Bassa, et al., Monthly Notices of the Royal Astronomical Society 414, 3251 (2011).

[41] T. Damour, A. Gopakumar, and B. R. Iyer, Physical Review D 70, 064028 (2004).

[42] T. Damour and N. Deruelle, Annales de l'I.H.P. Physique théorique 43, 107 (1985).

[43] R. M. Memmesheimer, A. Gopakumar, and G. Schäfer, Physical Review D 70, 17 (2004).

[44] Y. Boetzel, A. Susobhanan, A. Gopakumar, A. Klein, and P. Jetzer, Physical Review D 96, 044011 (2017).

[45] S. Mikkola, Celestial Mechanics 40, 329 (1987).

[46] P. C. Peters and J. Mathews, Physical Review 131, 435 (1963).

[47] R. van Haasteren and Y. Levin, Monthly Notices of the Royal Astronomical Society 401, 2372 (2010).

[48] M. F. Alam, Z. Arzoumanian, P. T. Baker, H. Blumer, K. E. Bohler, et al., The Astrophysical Journal Supplement Series 252, 5 (2020).

[49] J. H. Taylor, Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences 341, 117 (1992).

[50] T. T. Pennucci, P. B. Demorest, and S. M. Ransom, The Astrophysical Journal 790, 93 (2014).

[51] T. T. Pennucci, The Astrophysical Journal 871, 34 (2019).

[52] B. C. Joshi, A. Gopakumar, A. Pandian, T. Prabu, L. Dey, et al., arXiv e-prints , arXiv:2207.06461 (2022), arXiv:2207.06461 [astro-ph.HE].

[53] E. Kreyszig, Numerics in general (Wiley, 2005) p. 816, 9th ed.

[54] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, et al., Nature 585, 357 (2020).

[55] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, et al., Nature Methods 17, 261 (2020).

[56] A. M. Price-Whelan, B. M. Sipőcz, H. M. Günther, P. L. Lim, S. M. Crawford, et al., The Astronomical Journal 156, 123 (2018).

[57] J. D. Hunter, Computing in Science & Engineering 9, 90 (2007).

[58] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, et al., GNU Scientific Library Reference Manual, 3rd ed. (Network Theory Ltd., 2009).