Ephemeris Errors and the Gravitational-wave Signal: Harmonic Mode Coupling in Pulsar Timing Array Searches

Elinore Roebber
School of Physics and Astronomy and Institute for Gravitational Wave Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
e.roebber@bham.ac.uk
Received 2019 January 16; revised 2019 March 5; accepted 2019 March 13; published 2019 May 3

Abstract

Any unambiguous detection of a stochastic gravitational wave background (GWB) by a pulsar timing array will rest on the measurement of a characteristic angular correlation between pulsars. The ability to measure this correlation will depend on the geometry of the array. However, spatially correlated sources of noise, such as errors in the planetary ephemeris or clock errors, can produce false-positive correlations. The severity of this contamination will also depend on the geometry of the array. This paper quantifies these geometric effects with a spherical harmonic analysis of the pulsar timing residuals. At least nine well-spaced pulsars are needed to simultaneously measure a GWB and separate it from ephemeris and clock errors. Uniform distributions of pulsars can eliminate the contamination for arrays with large numbers of pulsars, but pulsars following the galactic distribution of known millisecond pulsars will always be affected. We quantitatively demonstrate the need for arrays to include many pulsars and for the pulsars to be distributed as uniformly as possible. Finally, we suggest a technique to cleanly separate the effect of ephemeris and clock errors from the gravitational wave signal.

Key words: ephemerides – gravitational waves – methods: data analysis – pulsars: general – time

1. Introduction

Pulsar timing arrays (PTAs) are gravitational wave (GW) detectors composed of radio telescopes on Earth and a collection of millisecond pulsars across the sky. Each pulsar regularly emits radio pulses with periods on the order of milliseconds. These periods are extremely stable—comparable to atomic clocks over timescales of years to decades (Lorimer 2008).

These millisecond pulsars are used as clocks to probe nanohertz gravitational perturbations. GWs can be found by comparing the difference between the expected and actual times of arrival of pulses (“timing residuals”) for multiple pulsars.

The primary signal in the PTA band is expected to be produced by a large population of slowly inspiraling binary systems of supermassive black holes (Rajagopal & Romani 1995). Such systems likely form in the aftermath of galaxy mergers (Begelman et al. 1980), and are thought to be numerous enough to lead to a nearly stochastic GW signal with a large degree of source confusion (Sesana et al. 2008).

This astrophysical signal is commonly referred to as the stochastic gravitational wave background (hereafter GWB), and is often assumed to take the form of a statistically unpolarized and isotropically distributed Gaussian random field. Other possible signals, such as a background produced by inflation, may have similar statistics, but with a different spectrum (e.g., Lasky et al. 2016). For simplicity, this paper will assume signals with these statistical properties.

There are currently three main PTA consortia: the European Pulsar Timing Array (EPTA) (Lentati et al. 2015), the North American Nanohertz Observatory for Gravitational Waves (Arzoumanian et al. 2018), and the Parkes Pulsar Timing Array (Shannon et al. 2015), all of which cooperate to form the International Pulsar Timing Array (IPTA) (Verbiest et al. 2016). Upper limits from each PTA are within a factor of a few of each other, and are now cutting into an astrophysically interesting range of parameter space (Taylor et al. 2017b; Arzoumanian et al. 2018; Chen et al. 2018).

Looking into the future, pulsar surveys—especially those performed by the forthcoming Square Kilometer Array—are expected to find many more millisecond pulsars (Keane et al. 2015). Furthermore, new regional PTAs in India (Joshi et al. 2018) and China (Lee 2016) are being formed, and new telescopes will contribute to pulsar discovery and timing (Hobbs & Dai 2017).

PTAs are affected by a large number of noise sources, including intrinsic white and red noise, as well as extrinsic and instrumental effects such as dispersion measure variation, ephemeris errors, and clock errors (Edwards et al. 2006; Cordes 2013). Although many sources of noise can be largely removed by cross-correlating timing residuals from different pulsars, there are also sources of noise that are correlated between pulsars. Important sources of correlated noise include errors in the clock used to calibrate timing residuals and errors in the solar system ephemeris used to correct for Earth’s orbit around the solar system barycenter (Foster & Backer 1990; Tiburzi et al. 2016).

Because these correlated sources of noise induce large-scale correlations between pulsars, they can mimic the effects of gravitational waves. The ephemeris errors are of particular concern, because current measurements of the location of the solar system barycenter are not sufficiently accurate (Champion et al. 2010; Lazio et al. 2018). Arzoumanian et al. (2018) have recently shown that using recent ephemerides in a PTA GW search can produce a systematic bias on the upper limit of the GWB amplitude. It is well-known that clock errors produce instantaneously monopolar signals, ephemeris errors produce instantaneously dipolar signals, and GWs produce signals that contain quadrupolar and higher-order terms (Foster & Backer 1990; Tiburzi et al. 2016). However, the underlying geometry of this contamination has not yet been completely explored.

In this paper, we will build on these established facts to show how the contamination of the gravitational signal by the clock and ephemeris errors can be expressed in a spherical harmonic...
analysis of pulsar timing residuals. This framework can be used to make decisions about which pulsars should be included in PTAs and to guide efforts to separate the signal due to GWs from that produced by clock or ephemeris errors.

Previous work to model and remove ephemeris errors has included simple geometric models (Foster & Backer 1990), detailed models of the solar system (e.g., Taylor et al. 2017a), and algebraic removal techniques based on the principles of time-delay interferometry (Tinto 2018). The framework proposed in this paper is related to that in Tinto (2018), but differs in our geometric focus and description of the origin of the contamination.

For simplicity, we will limit the discussion to a single bin of frequency space. The focus of this paper is on the spatial geometry, which is orthogonal to time–frequency formulations, so analysis can be straightforwardly extended to multiple frequency bins or a time-domain analysis. However, this analysis relies on simultaneous combinations of data from many pulsars, so it will not be trivial to extend it to the case of unevenly sampled data.

The organization of this paper is as follows. Section 2 presents a basic description of PTA measurements of GWs. Section 3 reviews previous work by the author on using harmonic analysis to describe ideal PTA measurements. Section 4 shows how incomplete coverage of the sky by pulsars in the PTA leads to coupling between different spatial modes. Section 5 shows how different configurations of PTAs affect the coupling. Section 6 discusses how knowledge of the coupling can be used to remove contamination during a search for GWs. Section 7 presents our conclusions.

2. Pulsar Timing Arrays

A gravitational wave signal can be written in the form

$$h_{ab}(t, x) = \int_{-\infty}^{\infty} df \int d^2\Omega_\hat{n} \, h_{ab}(f, \hat{n}) \, e^{2\pi i (\hat{n} \cdot \hat{x} / c)},$$

where $h_{ab}$ represents a perturbation in the metric describing the curvature of spacetime, $a$ and $b$ are indices representing spatial coordinates, $\hat{n}$ gives the direction from an observer toward the GW source, and $d^2\Omega_\hat{n}$ represents an infinitesimal region of sky in the direction $\hat{n}$. The GWs will delay or advance pulses from a given pulsar by an amount given by (Romano & Cornish 2017):

$$\Delta T(t, \hat{p}) = \int_{-\infty}^{\infty} df \int d^2\Omega_\hat{n} \, \frac{1}{2} \frac{p^a p^b}{1 + \hat{n} \cdot \hat{p}} \, h_{ab}(f, \hat{n}) \, e^{2\pi i \hat{p} \cdot \hat{n}} \times \frac{1}{2\pi i f} \left[ 1 - e^{-2\pi i (1 + \hat{n} \cdot \hat{p}) / c} \right].$$

Here, $\hat{p}$ gives the direction to the pulsar, $L$ is the distance to the pulsar, and the observer is assumed to be at the location of the solar system barycenter.

This equation can be broken into two terms, with one subtracted from the other. The first term, with no dependence on $L$, is known as the Earth term; it represents the effect of the GWs on pulses arriving at the Earth. The second term, which is dependent on $L$, is known as the pulsar term; it represents the effect of the GWs on pulses leaving the pulsar. If $L$ is known to a precision better than the GW wavelength, the pulsar term can be used to improve sky localization of individual GW sources (Zhu et al. 2016). However, this is not the case for most millisecond pulsars. Moreover, for a stochastic GWB, the pulsar term averages to a random phase and can be treated as a noise term. For simplicity, we will neglect the pulsar term and work with the Earth-term timing residuals:

$$r(t, \hat{p}) = \int_{-\infty}^{\infty} df \int d^2\Omega_\hat{n} \, \frac{1}{2\pi i f} \frac{p^a p^b}{1 + \hat{n} \cdot \hat{p}} \, h_{ab}(f, \hat{n}) \, e^{2\pi i \hat{p} \cdot \hat{n}}.$$  

(3)

In a single pulsar, it will be impossible to distinguish between this GW signal and the timing noise of the pulsar. Instead, a GW search looks for a correlated signal between different pulsars

$$\langle r(\hat{p}_i) r^*(\hat{p}_j) \rangle = P_{\text{GWB}}(f) \, C(\vartheta_{ij}),$$

(4)

where $P_{\text{GWB}}(f)$ is the frequency power spectrum of the GWB, and $C(\vartheta_{ij})$ is the characteristic angular correlation pattern of GWs (Hellings & Downs 1983)

$$C(\vartheta_{ij}) = \frac{1}{2} \left[ 1 - \frac{3}{4} (1 - \cos \vartheta_{ij}) \left( \ln \left( \frac{1 - \cos \vartheta_{ij}}{2} \right) - \frac{1}{6} \right) + \frac{1}{2} \delta_{ij} \right],$$

(5)

where $\vartheta_{ij}$ represents the angle between pulsars $i$ and $j$. This is known as the Hellings and Downs curve or the overlap reduction function. It has been normalized such that $C(0) = 1$, and subsequent angular correlation functions shall be as well.

The timing residuals will also be affected by any error in the terrestrial time standard, which is used to measure the time between pulses (Tiburzi et al. 2016). Any such error will affect all pulsars measured at the same time:

$$r_{\text{clk}}(t, \hat{p}) = e^{r_{\text{clk}}(t)},$$

(6)

where $e^{r_{\text{clk}}(t)}$ is the clock error as a function of time. Because all pulsars are affected equally, a clock error will have a monopolar characteristic correlation pattern

$$C(\vartheta_{ij}) = 1.$$  

(7)

Similarly, because Equation (3) is calculated assuming that observations are taken at the solar system barycenter, it is necessary to translate pulse times of arrival to the solar system barycenter frame when calculating the timing residuals. If the planetary ephemerides are not known correctly, this conversion will induce an extra term in the timing residuals (Tiburzi et al. 2016):

$$r_{\text{eph}}(t, \hat{p}) = \frac{1}{c} e^{r_{\text{eph}}(t)} \cdot \hat{p},$$

(8)

where $c$ is the speed of light and $e^{r_{\text{eph}}(t)}$ is the error in the position of the solar system barycenter with respect to time. The characteristic correlation function for this is dipolar:

$$C(\vartheta_{ij}) = \cos \vartheta_{ij}. $$

(9)

Under ideal conditions, the characteristic correlation functions for the GW signal and the clock and ephemeris errors are different enough to separate out these effects. However, this does not seem to be true in practice; we will explore why in Sections 4 and 5.
3. Harmonic Space Analysis

Measuring the Hellings and Downs curve in the cross-correlation of pulsar timing residuals will be crucial to the detection of GWs by PTAs. A previous paper by the author (Roebber & Holder 2017) developed an equivalent analysis using harmonic-space correlations of the pulsar timing residuals. The results of this paper will be based on that technique, so it is summarized here.

Consider an idealized PTA with an infinite number of pulsars smoothly covering the entire sky. In this case, we can make a map of the Earth-term timing residuals measured at each pulsar. If the time-domain residuals are used, the map will be real and smoothly varying with time. The frequency-domain residuals can be shown as a set of independent complex maps, with one for each frequency bin. Figure 1 shows an example of the latter case for a single frequency bin.

A smooth, full-sky timing residual map \( r(\Omega) \) can be decomposed into spherical harmonics:

\[
r(\Omega) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\Omega),
\]

where \( Y_{\ell m}(\Omega) \) are the spherical harmonics.

Figure 1. A map of the noiseless Earth-term timing residuals across the sky for a single random realization of a statistically isotropic Gaussian GWB in a single frequency bin. Upper and lower figures show the real and imaginary components of the complex Fourier-space timing residuals. The pulsars used in the IPTA initial data release (Verbiest et al. 2016) are marked with black circles. Each acquires the complex timing residual shown inside the circle. The mostly quadrupolar nature of the GW signal can be seen by eye: there are roughly large blobs of positive and negative residuals, and pulsars are correlated.

Related Image
GWBS. Any single measurement of the two-point function (in a single frequency bin) will differ from the underlying form due to sample variance (see discussion in Roebber & Holder 2017), but this can be reduced by measuring more frequency bins.

As expected, because $C_2$ is a steeply decreasing function of $\ell$, most of the power is in the quadrupole moment ($\ell = 2$), with smaller contributions from higher modes and none in the monopole or the dipole. This can also be seen in the example timing residual map in Figure 1: the dominant fluctuations are those covering about 1/4 of the sky, smaller-scale fluctuations are much less important, and fluctuations on scales smaller than $\sim 10^\circ$ are practically absent. The higher-than-quadrupole terms are due to the factor of $(1 + \hat{n} \cdot \hat{\rho})^{-1}$ in Equation (3) and are unrelated to any non-quadrupolar components of the strain.

The quadrupole term of $C_2$ contains nearly all of the signal-to-noise of the total GWB, so a successful detection of the GWB will require the detection of $C_2$. Most contributions to the noise are uncorrelated between pulsars and should affect all terms of $C_2$ equally.

The spatially correlated noise terms discussed in Section 2, however, have their own characteristic angular power spectra. In both cases, these are very simple: clock errors produce a monopole ($\ell = 0$) signal, and ephemeris errors produce a dipole ($\ell = 1$) signal. Harmonic analysis of the timing residuals naturally separates these spatially correlated noise signals from the GW signal ($\ell \geq 2$).

Given this natural separation, why are we concerned that measurements of the GW signal might be contaminated by ephemeris or clock errors? So far, we have assumed that our array of pulsars covers the sky sufficiently completely that we may accurately measure spherical harmonics up to $\ell = 2$. However, a real PTA will have a limited number of pulsars distributed non-uniformly across the sky; as a result, measurements of spherical harmonics will suffer from convolution with the geometry of the pulsar array. This effect will be the subject of the next section.

4. The Mode Coupling Matrix

We previously assumed an ideal PTA that can measure timing residuals at every point on the sky. However, a real PTA is limited to only tens of points on the sky: namely, the locations of each pulsar in the array. This section generalizes the results from the previous section to include the effects of a PTA that instead sparsely samples the underlying timing residual map.

In an ideal PTA, the measured timing residuals form a smooth field, which can be straightforwardly transformed into spherical harmonics as in Equation (10). This is possible because the spherical harmonics form an orthonormal basis for fields on a sphere. However, for a finite number of pulsars, the spherical harmonics sampled at the locations of the pulsars will no longer form an orthogonal basis function for this space.

Attempting to measure the amplitudes of the standard spherical harmonics will have mixed results: some harmonics will be poorly sampled and power will bleed between different modes. This is particularly a concern if the pulsars are not evenly distributed across the sky, as is the case for the population of known millisecond pulsars (see Figure 2).

The degree to which power from one mode can bleed into another can be represented by a coupling matrix (Peebles 1973; Gorski 1994; Wandelt et al. 2001; Hivon et al. 2002; Mortlock et al. 2002; Efstathiou 2004), which shows how much each spherical harmonic overlaps with all other harmonics:

$$K_{\ell m i j} = \int d\Omega Y_{\ell m}^* (\Omega) W (\Omega) Y_{\ell m i}^* (\Omega) \Omega$$

$$= \frac{4\pi}{\sum w_i \sum w_j Y_{\ell m} (\hat{\rho}_j) Y_{\ell m i} (\hat{\rho}_j)}.$$ (18)

The window function $W (\Omega)$ is used to define the part of the sky measured by pulsars:

$$W (\Omega) = \sum w_i \delta (\Omega - \hat{\rho}_i),$$

where the pulsar sky locations are given by $\hat{\rho}_i$ and the weights for each pulsar are given by:

$$w_i \propto \frac{\sigma_i^2 \sum}{\sigma_i^2} \propto \frac{\sigma_i^2 \sum}{\sigma_i^2}$$

We will discuss the first case (pulsars have different levels of intrinsic noise) in Section 5.3, and otherwise assume the second case (all pulsars are equal).

If complex spherical harmonics are used to calculate $K_{\ell m i j}$, the coupling matrix will be Hermitian. For real spherical harmonics, it will be real and symmetric.

When the integral in Equation (18) is taken over the entire sphere with $W (\Omega) = 1$ everywhere, the coupling matrix becomes $K_{\ell m i j} = \delta_{\ell \ell'} \delta_{m m'}$ and the spherical harmonics are orthonormal once more.

The properties of the coupling matrix determine whether or not we will be able to algebraically disentangle the different modes. When the different harmonics are orthogonal, the coupling matrix is diagonal and the problem is trivial. In the opposite limit, where all harmonics are maximally entangled, the coupling matrix is singular and the problem is impossible. In general, if the coupling matrix is well-conditioned (invertible), it should be possible to disentangle the GW signal from dipolar ephemeris and monopolar clock errors.

Given the PTA response function $\propto (1 + \hat{n} \cdot \hat{\rho})^{-1}$, the spherical harmonic transform of the timing residuals will contain GW signal at all $\ell \geq 2$, even for a strain signal that is strictly quadrupolar. However, because the majority of the
signal is contained in the quadrupole, we will neglect these higher-order terms.

Using spherical harmonic product identities, Equation (18) can also be written (e.g., Hivon et al. 2002):

\[
K_{(m_1)(m')}(\ell_1 \ell_2 \ell_3 \rightarrow m_1 m_2 m_3) = \frac{4\pi}{\sum \sum_i \sqrt{(2l+1)(2l'+1)(2L+1)}} \times \left( \begin{array}{ccc} \ell & \ell' & L \\ m & m' & M \end{array} \right) \left( \begin{array}{ccc} \ell & \ell' & L \\ 0 & 0 & 0 \end{array} \right) w_i Y_{LM}(\hat{\eta}_i),
\]

where terms of the form

\[
\left( \begin{array}{ccc} \ell_1 & \ell_2 & \ell_3 \\ m_1 & m_2 & m_3 \end{array} \right)
\]

are known as Wigner 3j symbols and are closely related to the Clebsch–Gordan coefficients. They generally cannot be written in closed form, but obey the following symmetry rules (e.g., Hivon et al. 2002):

\[
m_1 + m_2 + m_3 = 0
\]

\[
|\ell_1 - \ell_2| \leq \ell_3 \leq \ell_1 + \ell_2 \quad \text{(triangle inequality)}
\]

\[
\ell_1 + \ell_2 + \ell_3 \text{ is even if } m_1 = m_2 = m_3 = 0
\]

This form of the coupling matrix is useful because it is explicitly written in terms of the pulsar sky locations \(\sum_i w_i Y_{LM}(\hat{\eta}_i)\). This will allow us to gain insight into how the latter affects the former (see Section 5.2).

5. PTA Configuration and Mode Coupling

The coupling matrix \(K_{(m_1)(m')}(\ell_1 \ell_2 \ell_3 \rightarrow m_1 m_2 m_3)\) is entirely determined by the properties of the pulsar array. Important properties include the number of pulsars in the array, the spatial distribution of the pulsars, and the quality of each pulsar.

The effects of varying these properties will be illustrated by a set of six toy model PTAs. The models include arrays with 10 pulsars, with 100 pulsars, and with 10 good pulsars (with weight 1) and 90 poor pulsars (with weight 0.05). The pulsars in each case are generated by random sampling from two spatial distributions: “uniform” and “galaxy.”

The “uniform” distribution represents the case where every direction is equally likely to contain a pulsar. The “galaxy” distribution is based on the known population of millisecond pulsars (as shown in Figure 2). The distribution was created by selecting millisecond pulsars from the ATNF pulsar catalog (Manchester et al. 2005), removing multiples associated with globular clusters (in order to avoid biasing the distribution toward those sky locations), and smoothing with a Gaussian beam with full width at half maximum of 45°.

The distribution of pulsars for the “uniform” and “galaxy” cases are shown in Figure 3. The subsets used for the “10 equal pulsars” and the “10 good pulsars” are marked with dark stars. Examples of harmonics that are not well-measured or are coupled to other harmonics are shown in Figure 4. The coupling matrix for each model PTA is shown in Figure 5, and their eigenvalues in Figure 6.

5.1. Pulsar Number

The first important consideration is the number of pulsars in the array, or the number of points at which the signal is sampled. To successfully disentangle GW signal and spatially
correlated clock and ephemeris errors, it will be necessary to measure the $\ell = 0$ (monopole), $\ell = 1$ (dipole), and $\ell = 2$ (quadrupole) moments of the PTA response. With too few pulsars, some modes of interest will be poorly sampled and $K_{\ell m}^{(\ell m)}$ will not be invertible. This problem is closely related to the Nyquist sampling theorem.

A lower limit on the number of pulsars required can be calculated using a mode-counting argument. There are $2\ell + 1$ harmonic modes $m$ for every multipole $\ell$. For a frequency-domain signal, the timing residuals are complex, and all modes are independent. If we assume that the GW signal is approximately band-limited with $\ell \leq \ell_{\text{max}}$, we can count the number of modes needed:

$$\sum_{\ell=0}^{\ell_{\text{max}}} 2\ell + 1 = (\ell_{\text{max}} + 1)^2. \quad (26)$$

Measuring the leading-order term of the GW signal (the quadrupole) and disentangling it from the potentially nonzero monopole and dipole terms would therefore require a minimum of nine pulsars. Measuring up to the octopole term of the PTA response would require a minimum of 16 pulsars.

However, larger arrays may perform significantly better. As can be seen in Figure 5, the coupling matrices for the 10-pulse arrays have large off-diagonal elements, indicating strong coupling, and occasional weak diagonal elements where modes are not measured well (see also Figure 4). These aspects improve for arrays with more pulsars.

To determine whether the coupling can be removed, we want to know if the coupling matrix is well-conditioned. This can be evaluated by looking at its eigenvalue decomposition, as in Figure 6.

A matrix with at least one zero-valued eigenvalue is singular. If the eigenvalues are nonzero but very small, it may be possible to numerically invert the matrix; however, using the inverse will magnify errors in the data. In this scenario, the matrix is ill-conditioned. A useful figure of merit is the condition number, i.e., the ratio between the largest and smallest eigenvalue. When the condition number is close to unity, the matrix is well-conditioned; when it is large, the matrix is ill-conditioned; and when it is infinite, the matrix is singular. In this paper, we will not address the issue of choosing where to draw the line between well- and ill-conditioned matrices. The range of acceptable values will likely depend on the details of the analysis chosen, and should be assessed using simulations with a more thorough treatment of the various sources of noise than are considered here.

Although the mode-counting argument above suggested that arrays with $\geq 9$ pulsars should have well-conditioned coupling matrices, the 10-pulse examples seen in Figure 6 are marginal, particularly in the case of the galaxy-distributed PTA. A coupling matrix for a PTA with a small number of pulsars can

\[ \text{Figure 5. Coupling matrices } K_{\ell m}^{(\ell m)} \text{ for six different toy PTAs with the sky distributions given in Figure 3. The left column shows uniformly sampled PTAs, and the right shows galaxy-sampled PTAs. The top row shows arrays with 10 pulsars. The middle row shows arrays where the first 10 pulsars are “good” (weight = 1) and the other 90 are “bad” (weight = 0.05). The bottom row shows the coupling matrix for the full equally weighted array of 100 pulsars. The gray boxes highlight terms that represent coupling between the monopole or dipole modes and the quadrupole modes. Any nonzero terms within a box can lead to clock or ephemeris errors being mistaken for a GW signal.} \]

\[ \text{Figure 6. Sorted eigenvalues of the coupling matrices from Figure 5. Large arrays (dotted lines) perform well, whereas small arrays (dashed lines) perform poorly. Coupling matrices for arrays with 10 “good” pulsars and 90 “bad” pulsars (solid lines) are nearly as good as the equally weighted case, despite the appearance of Figure 5. This is because their largest eigenvalues follow those of the 10-pulsar arrays, but the smallest are raised. Uniformly drawn arrays (blue, thick) give better results than the arrays drawn from a galactic distribution (red, thin).} \]
Figure 7. Distributions (using a kernel density estimator) of the ratio between largest and smallest eigenvalue of the coupling matrix for 1000 realizations of different random PTA configurations. When this ratio is close to unity, the coupling matrix is well-conditioned and can be safely inverted. When this ratio is large (but not infinite), the inverse coupling matrix can be computed, but using it will amplify errors. An array of nine pulsars is the smallest array that can have a well-conditioned coupling matrix but using it will amplify errors. An array of nine pulsars is the smallest array more likely to do so.

have a wide range of condition numbers, depending on the locations of the different pulsars.

Figure 7 shows the condition number for 1000 different random realizations of arrays with a variable number of pulsars. From the top plot, it is clear that the coupling matrix of any array with \( \leq 8 \) pulsars is numerically singular, whereas for larger arrays, the coupling matrix is not singular. However, 9- or 10-pulsar arrays show a spread of many orders of magnitude in condition number. Some distributions will be well-conditioned, and some will certainly not. As the number of pulsars increases, the mean and spread decrease, and an array with \( \geq 20 \) pulsars will be more likely to have a well-conditioned coupling matrix.

This result agrees with previous work (Siemens et al. 2013; Taylor et al. 2016) arguing that large arrays are needed to detect a stochastic GWB, albeit for different reasons. However, other researchers (Foster & Backer 1990; Cornish & Sampson 2016; Tinto 2018) have suggested that a five-pulsar array could be sufficient, which contradicts the nine-pulsar minimum found in this paper.

In the cases of Foster & Backer (1990) and Tinto (2018), this is due to a difference in the counting arguments used. Their five pulsars can be understood to include one pulsar to measure and remove the monopole clock error, three pulsars to measure and remove the dipole ephemeris error, and a final pulsar to measure the GWB. However, if the GWB is assumed to have the form of a Gaussian-random quadrupole, it will have five independent components, bringing the minimum to nine pulsars.

On the other hand, Cornish & Sampson (2016) show that small five-pulsar arrays are adequate for detecting a sufficiently isotropic GWB in the absence of correlated noise terms. This does not contradict our results, however, because they also find that small arrays perform less well for GW signals dominated by small numbers of sources or in the presence of correlated noise.

The number of pulsars needed for the coupling matrix to be well-conditioned is a useful lower limit on the necessary size of a PTA. However, as can be seen in Figure 5, large arrays have less intrinsic coupling between the quadrupole and monopole/ dipole modes than small arrays.

This leads to another important threshold: the number of pulsars needed for the coupling between the GW signal and clock/ephemeris error to become unimportant. It happens that the distribution of pulsars across the sky has a significant impact on this threshold, so it will be discussed in the next section.

5.2. Pulsar Sky Distribution

The second important aspect is the sky distribution of the pulsars. Millisecond pulsars are not located uniformly across the sky (see Figure 2). As a result, the distribution of pulsar separations will not be uniform in $\cos \alpha$, but will be biased toward smaller separations (see figures in Verbiest et al. 2016 and Arzoumanian et al. 2018). Furthermore, they are predominantly located in an area that surrounds the center of the galaxy and is also nearly centered on the ecliptic.

The effect of the pulsar spatial distribution on the form of the coupling matrix is perhaps most clear in Equation (22). In this equation, terms of the form $w_l Y_{LM}(\hat{\mathbf{r}})$ contain information about the pulsars in the array. The Wigner symmetry rules (Equations (24) and (25)) determine which $Y_{LM}$ terms are important for coupling. In particular, we are interested in coupling between $\ell = 2$ and $\ell = 1$ modes and between $\ell = 2$ and $\ell = 0$ modes. In the first case, harmonics with $L = 1, 3$ will contribute, but only harmonics with $L = 2$ will matter in the second case.

This means that any PTA distribution that has significant dipole, quadrupole, or octopole components will lead to coupling between the quadrupole GW signal and dipole (monopole) ephemeris (clock) error. In particular, the distribution of known millisecond pulsars correlates with the galaxy; from Figure 2, it is clear that distributions of pulsars will naturally tend to cluster around the center of the galaxy and will therefore have non-negligible dipole, quadrupole, and octopole components. As a result, PTAs naively chosen from a collection of the best pulsars will be predisposed toward strong coupling between the quadrupolar signal and the undesired dipole and monopole sources.

Figure 8, which shows the underlying structure of the amplitudes of the coupling matrices, confirms this. The uniformly distributed case has no underlying structure, while the galaxy-distributed case has an underlying tendency for certain modes to be coupled.
While Figure 8 shows the median coupling matrix for many different random realizations, Figure 9 shows how the amplitudes of the problematic terms of the coupling matrix are distributed across many random realizations. This figure shows that, while larger arrays have better coupling matrices, the worst coupling term for galaxy-distributed arrays remains large no matter how many pulsars are included. This worst coupling term has a median value \( \sim 0.7 \), but can be much higher. From Figure 8, we can see that the strongest coupling term is typically between the \((\ell, m) = (2, 2)\) mode and the \((1, -1)\) mode.

For galaxy-distributed PTAs with many pulsars, most potentially problematic terms in the coupling matrix are small, but none of the quadrupole modes are free from coupling. The \((2, 1)\) mode shows the least coupling overall (\(\sim 0.2\) to the monopole), but it is also the most poorly sampled of the quadrupole modes. These effects can also be seen in the bottom-right coupling matrix in Figure 5. This coupling matrix represents an array of 100 pulsars and is well-conditioned, as expected. Nonetheless, it shows the same structure as Figure 8.

PTAs with uniformly distributed pulsars show similarly large couplings when the number of pulsars is small. However, as the number of pulsars in the array increases, the overall coupling between quadrupole and monopole/dipole modes decreases and the matrix becomes increasingly diagonal. With 50 pulsars in the array, the mode with the worst coupling is always \(\sim 0.5\), and typical modes are \(\sim 0.1\) (see Figure 9).

Futuristic PTAs with \(\geq 50\) high-quality pulsars may be able to eliminate contamination by solar system ephemeris or clock errors, but only if the pulsars are distributed nearly uniformly across the sky. Current PTAs have been constructed to be more isotropic than our naive galaxy-distributed model, but they are not isotropic enough to have low levels of contamination.

In addition to the large-scale distribution of pulsars (dipole–octopole moments), the small-scale distribution of pulsars can also have an effect. Tightly clustered pulsars (separations \(< 180^\circ / \ell_{\max}\)) will make highly correlated measurements of the large-scale modes, so they cannot really be considered independent. This means that adding an additional pulsar close to one already present will not improve the coupling matrix as well as adding a widely spaced pulsar would.

As an example, the galaxy-distributed 10-pulsar model (shown in the bottom plot of Figure 3) has two sets of pulsars that are very close to one another. It might be expected to act somewhat like a seven-pulsar array. This is reflected in the eigenvalues of its associated coupling matrix (Figure 6); there is a drop of two orders of magnitude between the seventh- and the eighth-largest eigenvalues.

However, if there are enough widely separated pulsars in the array, tightly clustered pulsars may not be a problem. In particular, poor-quality pulsars with large amounts of white or red noise might not be useful on their own. However, if there are several of them all sampling the same underlying GW (and ephemeris error) signal, they could behave similarly to a single higher-quality pulsar. This effect could be employed in order to build a more uniform array of pulsars, with multiple poor-quality pulsars used to fill large gaps between the better ones.

In summary, a PTA must have \(\geq 9\) pulsars in order to be able to separate quadrupolar GW signal from dipolar ephemeris error and monopolar clock error, but there are also requirements on the distribution of those pulsars. Ideally, the pulsars should be arranged uniformly across the sky, with separations \(\sim 180^\circ / \ell_{\max}\). Non-uniformly distributed pulsars will increase
the coupling. An array with more closely spaced pulsars than optimal will need to have more pulsars in order to properly sample the relevant modes.

It is worth noting that this spacing recommendation is different from the spacing that maximizes the detectability of the Hellings and Downs curve. The most important difference is that pulsars with small separations are not useful for improving the coupling matrix. This is because correlations between tightly coupled pulsars may be useful for detecting a correlated signal, but they will be useless for determining whether this correlation is due to a GW signal, an error in the planetary ephemeris, or a clock error.

5.3. Pulsar Quality

The final important effect is the quality of different pulsars. This can be treated using the weights defined in Equation (20), so that “good” pulsars have a large weight and “bad” pulsars a small one.

The example for this scenario has the full set of 100 pulsars in Figure 3, but the weights are arranged such that the first 10 have \( w_i = 1 \) and the other 90 have \( w_i = 0.05 \). This simulates having a small array of “good” pulsars with many “bad” pulsars added to it. The weights were chosen to make the effects in Figures 5 and 6 visible.

Comparing the middle row of Figure 5 to the top and bottom shows that the coupling matrix is more similar to the case with 10 equal pulsars than to the case with 100 equal pulsars. We can show this from Equation (19), assuming a simple distribution of weights, as in our example.

Consider two populations of pulsars: \( N_g \), “good” pulsars with associated weights \( w_g \) and coupling matrix \( K_{(lm)(lm)'}^{g} \), and \( N_b \) “bad” pulsars with associated weights \( w_b \) and coupling matrix \( K_{(lm)(lm)'}^{b} \). Because there are more “bad” pulsars than good ones, \( K_{(lm)(lm)'}^{b} \) will typically have stronger diagonal terms and weaker off-diagonal terms than \( K_{(lm)(lm)'}^{g} \), although this will depend on the sky distribution of both sets of pulsars.

The contribution of each of these populations to the coupling matrix of the total population (from Equation (19)) can be separated:

\[
K_{(lm)(lm)'}^{g+b} = \frac{w_g N_g K_{(lm)(lm)'}^{g} + w_b N_b K_{(lm)(lm)'}^{b}}{w_g N_g + w_b N_b}.
\] (27)

The overall coupling matrix is a weighted average between the coupling matrices calculated respectively for each population. Unless \( N_b/N_g > w_b/w_g \), “good” pulsars will be more important for the overall form of the coupling matrix. Real PTAs tend to have sensitivities that are dominated by a small number of pulsars (e.g., Babak et al. 2016, where six out of 41 pulsars contribute more than 90% of the total array’s signal-to-noise squared), so similar results are likely to hold in practice.

Although the form of the coupling matrix is mostly set by the subarray of “good” pulsars, adding a spatially extended distribution of “bad” pulsars can have a significant effect on the conditioning of the coupling matrix (see Figure 6). The best eigenvalues of \( K_{(lm)(lm)'}^{g+b} \) are very similar to the best eigenvalues of \( K_{(lm)(lm)'}^{g} \), but the worst are considerably improved. This suggests that a large number of “bad” pulsars could be used to compensate for not having enough “good” pulsars to measure the desired \( \ell_{\text{max}} \).

Equation (27) and Weyl’s theorem can be used to place bounds on the amplitude of the smallest eigenvalue \( \lambda_{\text{min}} \) for the combined arrays of “good” and “bad” pulsars:

\[
\lambda_{\text{min}}^{g+b} \geq \frac{w_g N_g \lambda_{\text{min}}^{g} + w_b N_b \lambda_{\text{min}}^{b}}{w_g N_g + w_b N_b},
\] (28)

\[
\lambda_{\text{min}}^{g+b} \leq \frac{w_g N_g \lambda_{\text{min}}^{g} + w_b N_b \lambda_{\text{max}}^{b}}{w_g N_g + w_b N_b}.
\] (29)

This can be rewritten in terms of the improvement in \( \lambda_{\text{min}} \) between the “good” array and the combined array:

\[
F_b(\lambda_{\text{min}}^{b} - \lambda_{\text{min}}^{g} \leq \Delta \lambda_{\text{min}} \leq F_b(\lambda_{\text{max}}^{b} - \lambda_{\text{min}}^{g})
\] (30)

where \( F_b = w_b N_b/(w_g N_g + w_b N_b) \).

The difference between the two bounds is related to the spatial distributions of the “good” and the “bad” pulsars. The worst improvement \( \Delta \lambda_{\text{min}} \) will occur when the two arrays have very similar distributions and so make degenerate measurements. In this case, \( \lambda_{\text{min}}^{g} \) may be similar in amplitude to \( \lambda_{\text{max}}^{b} \), so the improvement can be very small.

By contrast, the best improvement \( \Delta \lambda_{\text{max}}^{b} - \lambda_{\text{min}}^{g} \) will occur when the two arrays have complementary distributions, such that adding the “bad” pulsars add new information. In this case, because \( \lambda_{\text{max}}^{b} \approx 1 \), even a relatively small array of low-weight pulsars can yield a significant improvement.

Although adding “bad” pulsars to the “good” ones typically will not significantly change the amount of coupling, they are most likely to be useful when they have a sky distribution that complements the “good” ones. In this case, the off-diagonal terms of their coupling matrices will partially cancel.

In summary, second-rate pulsars are generally less useful for making the sky distribution more uniform, because they have a weaker effect on the coupling matrix, but they are useful for improving the conditioning of the coupling matrix. The best effect will be produced by a collection of “bad” pulsars with nearly the opposite distribution on the sky from the “good” pulsars, in sufficient numbers that \( \sum w_b \sim \sum w_g \).

6. Correcting for Mode Coupling

Large numbers of uniformly distributed pulsars have very little coupling between the GW quadrupole (and higher) terms and spatially correlated clock and ephemeris errors (monopole and dipole terms). However, because millisecond pulsars do not have a uniform sky distribution, it may not be possible in practice to construct a sensitive PTA without this coupling.

Therefore, it is useful to have techniques for removing the coupling. The standard analysis uses the Hellings and Downs correlation between different pulsars to search for a GWB. An equivalent harmonic space calculation could be made using the angular correlation in Equation (13).

This is both optimal and unbiased if the monopole and dipole terms are both zero. However, any single attempted measurement of the GWB where either the monopole or dipole is nonzero will produce biased results (see Tiburzi et al. 2016; Vigeland et al. 2018) because power can leak from the monopole or dipole into the GW modes.

Knowledge of the coupling matrix can be used to explicitly remove the contamination, but care must be taken to avoid methods that are unbiased only for an ensemble of observations. This section will explore a single example of how this can be done, but other techniques are possible (see Tinto 2018).
6.1. Orthogonalized Spherical Harmonics

One way to deal with the coupling is to build a new set of orthogonal basis functions based on the spherical harmonics. As shown in Gorski (1994) and Mortlock et al. (2002) in the context of analysis of the cosmic microwave background anisotropies, this approach allows us to identify and excise components that are contaminated by the noisy monopole and dipole. This section and the next will apply their strategy in the context of PTAs. For mathematical ease, the notation of the previous sections will be used interchangeably with linear algebra notation: \( y = Y_{\text{tot}} \).

The orthogonalization will be performed using a Cholesky decomposition of the coupling matrix:

\[
K_{(\ell m)(\ell' m')} = K = LL^T,
\]  

where \( L \) is a lower-triangular matrix, and \( L^T \) represents the Hermitian transpose of \( L \).

The new orthogonalized harmonics are given by:

\[
y' = L^{-1}y.
\]  

By construction, they are orthonormal on the space of the pulsar sky positions:

\[
\int_{\text{pars}} W(\Omega) y'(\Omega)'d\Omega = \int_{\text{pars}} W(\Omega) yy'd\Omega \left(L^{-1}\right)' = L^{-1}K(L^{-1})' = L^{-1}LL^TL^{-1} = 1.
\]  

The new coefficients \( a' \) are similarly related to the true harmonic coefficients:

\[
a' = L' a
\]  

and can be estimated from the measured timing residuals (see Equation (11)):

\[
a' = \int_{\text{pars}} r(\hat{\Omega})[y'(\hat{\Omega})]'d\Omega.
\]

The spherical harmonics \( y \) are best described with two indices \((\ell, m)\), but the transformation in Equation (32) does not preserve these symmetries. As a result, the orthogonalized spherical harmonics \( y' \) have only a single index (denoted hereafter by \( b \) or \( c \)), which is related to the \( \ell \) components of \( y \).

Because the transformation in Equation (34) is based on a Cholesky decomposition, \( L' \) is upper-triangular. This means that the components of \( a' \) contain contributions from a decreasing number of modes of \( a \).

The first term of \( a' \), \( a'_0 \), is made up of contributions from all \( a_{\ell m} \). If the components of \( K_{(\ell m)(\ell' m')} \) are arranged in ascending order of the \( \ell \)'s, then \( a'_0 \) will contain contributions from all the \( a_{\ell m} \) except for the monopole. Similarly, the next three terms of \( a' \) progressively drop the dipole terms. Therefore, only the first four components of \( a' \) contain any contribution from the monopole and dipole modes.

By dropping those first four terms, we can cleanly separate the GW signal from the effect of the ephemeris and clock errors (see Figure 10). Because the ambiguous components are also dropped, some amount of signal will be lost. This can be calculated (on average) from the form of \( L' \). In general, coupling matrices with large coupling terms and large condition numbers will lose more signal.

This process can be visualized as a switch to a new basis of functions that are built from spherical harmonics. This new basis consists of (1) functions that contain the well-measured components of the quadrupole and (2) functions that contain the poorly-measured components of the quadrupole plus the monopole and dipole terms. For an array with a significant amount of coupling, functions of type (1) will not cover much of the sky, and so a significant amount of signal will be lost. By contrast, for an array with a nearly diagonal coupling matrix, functions of type (1) will approximately recover the quadrupole spherical harmonics and very little signal will be lost.

Note that, in ecliptic coordinates, we can expect that the ephemeris error will only contaminate the \( m = \pm 1 \) modes, so we could limit ourselves to removing the contributions from only three modes: \( Y_{\ell 0}, Y_{\ell 1}, \) and \( Y_{\ell -1} \). This would reduce the number of ambiguous modes, so more signal could be recovered.

This technique is related to the one described by Tinto (2018), but leads to a very different set of uncontaminated modes, with very different covariance properties. There are also

Figure 10. The measured and true amplitudes of the standard and new harmonics, calculated for a stochastic GWB (with the standard normalization), a monopole (amplitude 3), a dipole (amplitude 10) pointed in a random direction along the ecliptic, and the sum of all three fields. For visual clarity, no pulsar noise is included. Solid lines indicate true values and dashed lines indicate the amplitudes measured by an array with 100 equal galaxy-distributed pulsars. The top panel shows the leakage of the monopole and dipole into the quadrupole terms. The new harmonics do not allow any leakage, but some of the power in the true quadrupole is lost.
other techniques that can be used to construct orthogonal modes from spherical harmonics. The Cholesky decomposition was used for simplicity in constructing the clean modes $a'$, but there are many other options. Mortlock et al. (2002) discusses this problem more generally, including the case of a singular value decomposition, as well as techniques for extending the analysis to higher $\ell$, where the coupling matrix becomes singular.

6.2. Likelihood Analysis

The vector $a'$ can be thought of as a set of five linear combinations of the pulsar timing residuals, each of which is insensitive to the noisy monopole and dipole terms (note that we have now dropped the four contaminated modes). In this section, we will discuss the statistical properties of $a'$ that are needed for GWB searches.

For a Gaussian GWB, the spherical harmonic amplitudes $a_{lm}$ are, by definition, normally distributed. The amplitudes $a'$ of the orthogonalized spherical harmonics can each be written as a sum of $a_{lm}$ components. Therefore, the new amplitudes are also Gaussian random variables:

$$P(a') = \frac{\exp\left[-\frac{1}{2}(a')^T C^{-1} a'\right]}{(2\pi)^{N/2} \det C},$$

where $N = (\ell_{\text{max}} + 1)^2 - N_{\text{removed}}$ is the number of $a'$ modes. If $\ell_{\text{max}} = 2$ and all monopole and dipole mode terms are removed, then $N = 5$. The covariance matrix $C$ can be separated into a signal component and pulsar noise component: $C = C_{\text{GWB}} + C_N$.

The signal covariance matrix can be defined in terms of the coefficients of the orthogonalized harmonics:

$$C_{\text{GWB}} = \langle a (a')^* \rangle = L' (aa') L = \sum_{(lm)} L_{b(lm)}^* C_l L_{c(lm)},$$

where $b$ and $c$ now represent the index of the orthogonal modes where all monopole and dipole terms have been dropped. The second line is found using Equation (12) and noting that $\langle a_{lm} \rangle$ is the same for all $m$ at fixed $\ell$.

In a full-sky analysis, the orthogonality of the spherical harmonics means that $C_{\text{GWB}}$ is diagonal (with amplitudes given by $C_l$). However, this is not generally the case for the orthogonalized spherical harmonics. As the array becomes larger and more uniform, the coupling matrix $K$ will approach the identity matrix, and $C_{\text{GWB}}$ will become approximately uniform.

Assuming that the noise in each pulsar is given by $\sigma_i$ and is uncorrelated with other pulsars, the noise covariance matrix can be written:

$$C_N = \langle n n^* \rangle = \frac{4\pi}{\sum W_i} \langle \sigma_i w_i y_i^* (\sigma_i w_i y_i')^* \delta_{ij} \rangle$$

$$= \frac{4\pi}{\sum W_i} \langle \sigma_i^2 w_i^2 y_i^* (y_i')^* \rangle.$$

This is diagonal and exactly equal to the identity matrix if all pulsars are equal, or if each pulsar weight is equal to the inverse variance of its noise, as in Equation (21).

7. Discussion

The form of the antenna response function ensures that PTA searches for GWs have an inherently geometric component. This is fortunate, because it gives us a way to decouple the GW signal from correlated noise sources such as planetary ephemeris errors and clock errors.

A natural space in which to do this is provided by a spherical harmonic analysis of the pulsar timing residuals. In this space, GWs appear as quadrupole and higher-order terms, while ephemeris errors appear as dipole terms, and clock errors appear as a monopole.

However, real PTAs are composed of a moderate number of pulsars with varying intrinsic noise levels, distributed non-uniformly across the sky. Attempts to measure the different modes are therefore subject to aliasing-type effects, and the measured quadrupole GW signal may be contaminated by the clock error monopole and ephemeris error dipole.

This contamination can be quantified in terms of a coupling matrix that measures the extent to which different spherical harmonics are nonorthogonal for a given array of pulsars. The coupling matrix depends only on the spatial arrangement and noise properties of the pulsars in the array, and so provides a means to assess the degree to which contamination by ephemeris errors or clock errors is likely to be an issue for a particular configuration of pulsars.

Measuring the monopole, dipole, and quadrupole modes requires an array with a minimum of nine pulsars, but such arrays are typically subject to considerable coupling between the GW signal and ephemeris or clock errors. The arrays with the fewest coupling problems are those with many isotropic, high-quality pulsars. However, because the underlying distribution of millisecond pulsars is anisotropic, future PTAs will probably still experience significant amounts of coupling.

A biased array can be made less biased by keeping only the best pulsars from the region around galactic center and by adding clusters of second-rate pulsars wherever possible in the undersampled regions. In particular, covering the ecliptic plane is especially useful for minimizing coupling with the dipole modes most likely to be induced by ephemeris errors.

The distribution of pulsars that allows the GW signal to be most cleanly separated from the clock and ephemeris errors is not the same as the one that is most sensitive to detecting GWs, because the latter case favors closely spaced pulsars, which tend to increase coupling. Moreover, choosing pulsars such that the sky coverage is as uniform as possible will lead to a loss of sensitivity to GWs, if good pulsars are dropped.

An array that can make a clean detection of GWs will need to balance these two conflicting guidelines by ensuring that all pulsars are evaluated both in terms of their own quality and also in terms of their effect on the coupling matrix. As long as a detection has not yet been made, increasing the sensitivity of the array will remain the priority. However, improving the sky coverage can significantly improve the conditioning of the coupling matrix, even if only second-rate pulsars are used. Such arrays will still exhibit significant amounts of coupling, but should allow the GW signal to be disentangled from the effects of ephemeris and clock errors.

A PTA with significant coupling will need to correct for that coupling. If the array contains a large enough number of pulsars, the coupling matrix will be invertible and the coupling can be removed algebraically. Making a clean cut between the coupling and the signal will necessitate some loss of signal.
More signal will be retained when the coupling matrix is well-conditioned and has small off-diagonal components. This paper and Tinto (2018) present different examples of how the analysis can be done in a space that nulls contributions from the monopole and dipole terms. These algebraic techniques are complementary to physical models such as the planetary ephemeris model BAYESEPHEM (Arzoumanian et al. 2018), but it is likely that any analysis will benefit if pulsar arrays are chosen to minimize coupling while maintaining sensitivity to GWs.

Throughout this paper, we have assumed that ephemeris and clock errors produce pure monopolar and dipolar correlations between all pulsars in the array, but the situation will be more complicated in reality. In particular, timing residuals from different pulsars are irregularly sampled, there are gaps in the timing of certain pulsars, and new pulsars are occasionally added to the array. As a result, a real PTA cannot be treated as a static collection of pulsars, and a frequency-domain analysis becomes difficult.

For this reason, PTA analyses are typically performed in the time domain. It is possible to extend the analysis of this paper to the time domain. The time-domain coupling matrix can be calculated with Equation (19), using real spherical harmonics and time-varying pulsar positions and weights. The single coupling matrix that we have been considering becomes instead a series of different coupling matrices as pulsars are added and subtracted. As a result, the degree of contamination by ephemeris and clock errors will also vary as a function of time. If the analysis above is followed, this will also lead to a choice of \( g(t) \) that is not smoothly varying. However, because this analysis requires simultaneously combining data from all the pulsars in the array, rather than two at a time as in the standard analysis, irregular sampling may make this difficult to implement.

The full time-domain Bayesian search for GWs is extremely computationally expensive, because the combination of timing residuals from all pulsars results in a very large covariance matrix. This covariance matrix cannot be precomputed, because it depends on the model parameters and must be inverted for each likelihood evaluation. In search of a more computationally efficient analysis, recent work often treats the model parameters for the correlated components with a reduced-rank analysis that, through the application of known transformations, approximates the full time-domain covariance matrix (Lentati et al. 2013, 2015; van Haasteren & Vallisneri 2014; Arzoumanian et al. 2018). This is typically an approximate Fourier decomposition in which the low-frequency components (which contain most of the information) are retained—however, cf. van Haasteren & Vallisneri (2015).

The pulsar correlations due to GWs are treated in this stage of the analysis, through our knowledge of their expected frequency and spatial correlations. The analysis presented in this work could potentially be used in the place of the traditional Hellings and Downs correlations to treat the spatial part. This would lead to a new transformation between the orthogonalized spherical harmonics and the individual pulsar correlations, similar to the transformation between the frequency-domain and the time-domain correlations. These transformations would probably need to be done with careful treatment of pulsar weights at the different frequencies to avoid reintroducing monopole or dipole contamination via spectral leakage.

Implementing the techniques described in this paper within the reduced-rank variant of the full time-domain Gaussian analysis overcomes some of the difficulties presented by irregular sampling, because all pulsars can be transformed to a consistent set of frequencies. It could also have the benefit of further reducing the rank of the modeled covariance matrix from \( N_{\text{freq}} \times N_{\text{psr}} \) to \( N_{\text{freq}} \times 5 \), where \( N_{\text{freq}} \) is the number of frequency components considered. This reduction comes from the fact that the spatial correlations can be described using the \( 5 \times 5 \) orthogonalized spherical harmonic covariance matrix from Equations (37) and (39), rather than the standard \( N_{\text{psr}} \times N_{\text{psr}} \) covariance matrix.

The analysis presented in this paper has also treated only the case where the GWB is an isotropic Gaussian random field, but this may be an oversimplification. It is plausible that strong individual GW sources will dominate the signal in some frequency bins (Sesana et al. 2008; Ravi et al. 2012; Roeber et al. 2016). Although there are better means to detect single sources (e.g., Lee et al. 2011; Babak & Sesana 2012; Boyle & Pen 2012; Zhu et al. 2016; Goldstein et al. 2018), their presence should not significantly bias our analysis (see Cornish & Sampson 2016).

Single sources of GWs also produce Hellings and Downs correlations between pulsars, and equivalently, they produce the same angular power spectrum as in Equation (13). For multiple sources, as with individual realizations of the stochastic GWB, this correlation represents an ensemble average and does not hold exactly (Roeber & Holder 2017). The primary difference between these cases is that, for a stochastic GWB, the power in each multipole is random and uncorrelated, while for a GW signal produced by a few loud sources, the power in different multipoles becomes correlated due to interference between the pulsar antenna response to each source. These correlations appear as small ripples in the angular power spectrum (Roeber & Holder 2017). Because the underlying form of the angular power spectrum is similar in all cases, it will almost always be appropriate to approximate it as a quadrupole. As a result, the technique described in this paper should work even when our assumption of a statistically isotropic and Gaussian random GWB is violated by the presence of a few loud sources.

Large-scale anisotropies in the distribution of GW sources (Mingarelli et al. 2013; Taylor & Gair 2013), however, could produce a GWB with a very different angular power spectrum. In this case, the assumption that the GW signal is largely contained in the quadrupole modes will no longer hold. However, this scenario is unlikely in practice, as it represents a significant departure from cosmological isotropy.

The problems discussed in this paper will also affect astrometric GW searches using Gaia and similar experiments (Darling et al. 2018; Klioner 2018; Mihaylov et al. 2018; O’Beirne & Cornish 2018). Gaia observes many stars, but star locations are biased toward the direction of the galaxy, so similar coupling problems will arise.

Similarly, measurements of just the clock (Hobbs et al. 2012) or the planetary ephemeris (Champion et al. 2010) errors could experience significant coupling with the other effect or with GWs. A variant of the method described here may be useful for attempts to measure these individual quantities.

Although the problem of ephemeris errors is not trivial, it is solvable, especially for future arrays with many high-quality pulsars.
I thank Alberto Vecchio and Alberto Sesana for useful comments.

Software: Python with the following packages: numpy (van der Walt et al. 2011), scipy (Jones et al. 2001), pandas (McKinney 2010), healpy (Górski et al. 2005), matplotlib (Hunter 2007), seaborn (Waskom et al. 2017).

ORCID iDs
Elinore Roebber @ https://orcid.org/0000-0002-5709-4840

References
Arzoumanian, Z., Baker, P. T., Brazier, A., et al. 2018, ApJ, 859, 47
Babak, S., Peiteau, A., Sesana, A., et al. 2016, MNRAS, 455, 1665
Babak, S., & Sesana, A. 2012, PhRvD, 85, 044034
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Natur, 287, 307
Boyle, L., & Pen, U.-L. 2012, PhRvD, 86, 124028
Babak, S., Petiteau, A., Sesana, A., et al. 2016, MNRAS, 455, 1665
Beirne, L., & Cornish, N. J. 2018, PhRvD, 98, 024020
Cordes, J. M. 2013, CQGra, 30, 224002
Cornish, N. J., & Sampson, L. 2016, PhRvD, 93, 104047
Darling, J., Truenenbach, A. E., & Paine, J. 2018, ApJ, 861, 113
Dodelson, S. 2003, Modern Cosmology (New York: Academic)
Edwards, R. T., Hobbs, G. B., & Manchester, R. N. 2006, MNRAS, 372, 1549
Edwards, R. T., Hobbs, G. B., & Manchester, R. N. 2006, MNRAS, 372, 1549
Elston, A., 2004, Nature, 430, L85
Foster, R. S., & Backer, D. C. 1990, ApJ, 361, 300
Gair, J., Romano, J. D., Taylor, S., & Migliarelli, C. M. F. 2014, PhRvD, 90, 082001
Goldstein, J. M., Veitch, J., Sesana, A., & Vecchio, A. 2018, MNRAS, 477, 5447
Górski, K. M. 1994, ApJL, 430, L85
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Hellings, R. W., & Downs, G. S. 1983, ApJL, 265, L39
Hivon, E., Görski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2
Hobbs, G., Coles, W., Manchester, R. N., et al. 2012, MNRAS, 427, 2780
Hobbs, G., & Dai, S. 2017, arXiv:1707.01615
Hunter, J. D. 2007, CSE, 9, 90
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open Source Scientific Tools for Python, http://www.scipy.org/
Josh, B. C., Arumugasamy, P., Bagchi, M., et al. 2018, JApA, 39, 51
Keane, E., Bhattacharyya, B., Kramer, M., et al. 2015, in Proc of Advancing Astrophysics with Square Kilometre Array (AASKA14) (Trieste: SISSA), 40
Klioner, S. A. 2018, CQGra, 35, 045005
Lasky, P. D., Migliarelli, C. M. F., Smith, T. L., et al. 2016, PhRvX, 6, 011035
Lazio, T. J. W., Bhaskaran, S., Cutler, C., et al. 2018, in IAU Symp. 337, Pulsar Astrophysics the Next Fifty Years, ed. P. Weltevrede et al. (Cambridge: Cambridge Univ. Press), 150
Lee, K. J. 2016, in ASP Conf. Ser. 502, Frontiers in Radio Astronomy and FAST Early Sciences Symp. 2015, ed. L. Qai & D. Li (San Francisco, CA: ASP), 19
Lee, K. J., Wex, N., Kramer, M., et al. 2011, MNRAS, 414, 3251
Lentati, L., Alexander, P., Hobson, M. P., et al. 2013, PhRvD, 87, 104021
Lentati, L., Taylor, S. R., Migliarelli, C. M. F., et al. 2015, MNRAS, 453, 2576
Lorimer, D. R. 2008, LRR, 11, 8
Manchester, R. N., Hobbs, G. B., Teoh, A., & Hobbs, M. 2005, AJ, 129, 1993
McKinney, W. 2010, in Proc. 9th Python in Science Conf., ed. S. van der Walt & J. Millman, 51, http://conference.scipy.org/proceedings/scipy2010/mckinney.html
Mihaylov, D. P., Moore, C. J., Gair, J. R., Lasenby, A., & Gilmore, G. 2018, PhRvD, 97, 124058
Mingarelli, C. M. F., Sidery, T., Mandel, I., & Vecchio, A. 2013, PhRvD, 88, 062005
Mortlock, D. J., Challinor, A. D., & Hobson, M. P. 2002, MNRAS, 330, 405
O’Beirne, L., & Cornish, N. J. 2018, PhRvD, 98, 024020
Peebles, P. J. E. 1973, ApJ, 185, 413
Rajagopal, M., & Romani, R. W. 1995, ApJ, 446, 543
Ravi, V., Wylie, J. S. B., Hobbs, G., et al. 2012, ApJ, 761, 84
Roebber, E., & Holder, G. 2017, ApJ, 835, 21
Roebber, E., Holder, G., Holz, D. E., & Warren, M. 2016, ApJ, 819, 163
Romano, J. D., & Cornish, N. J. 2017, LRR, 20, 2
Sesana, A., Vecchio, A., & Colacino, C. N. 2008, MNRAS, 390, 192
Shannon, R. M., Ravi, V., Lentati, L. T., et al. 2015, Sci, 349, 1522
Siemens, X., Ellis, J., Jenet, F., & Romano, J. D. 2013, CQGra, 30, 224015
Taylor, S. R., & Gair, J. R. 2013, PhRvD, 88, 084001
Taylor, S. R., Lentati, L., Babak, S., et al. 2017a, PhRvD, 95, 042002
Taylor, S. R., Simon, J., & Sampson, L. 2017b, PhRvL, 118, 181102
Taylor, S. R., Vallisneri, M., Ellis, J. A., et al. 2016, ApJL, 819, L6
Tiburzi, C., Hobbs, G., Kerr, M., et al. 2016, MNRAS, 455, 4339
Tinto, M. 2018, PhRvD, 97, 084047
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
van Haasteren, R., & Vallisneri, M. 2014, PhRvD, 90, 104012
van Haasteren, R., & Vallisneri, M. 2015, MNRAS, 446, 1170
Verbiest, J. P. W., Lentati, L., Hobbs, G., et al. 2016, MNRAS, 458, 1267
Vigeland, S. J., Islo, K., Taylor, S. R., & Ellis, J. A. 2018, PhRvD, 98, 044003
Wandelt, B. D., Hivon, E., & Görski, K. M. 2001, PhRvD, 64, 083003
Waskom, M., Botvinnik, O., O’Kane, D., et al. 2017, mwaskom/seaborn: v0.8.1, Zenodo, doi:10.5281/zenodo.1313201
Zhu, X.-J., Wen, L., Xiong, J., et al. 2016, MNRAS, 461, 1317